

Note di statistica descrittiva e primi elementi di calcolo delle probabilità

Temi d'esame (soluzioni estese)

G. Boari, G. Cantaluppi

Dipartimento di Scienze statistiche
Università Cattolica del Sacro Cuore
Milano

Giuseppe Boari
Gabriele Cantaluppi



UNIVERSITÀ
CATTOLICA
del Sacro Cuore



T271	Prova del 10 gennaio 2013
T272	Prova del 24 gennaio 2013
T273	Prova del 7 febbraio 2013
T274-1	Prova del 6 giugno 2013
T274	Prova del 27 giugno 2013
T274-2	Prova del 11 luglio 2013
T275	Prova del 5 settembre 2013
T275-1	Prova del 16 settembre 2013
T276	Prova del 16 gennaio 2014
T277	Prova del 30 gennaio 2014
T278	Prova del 13 febbraio 2014
T279-1	Prova del 5 giugno 2014
T279	Prova del 26 giugno 2014
T279-2	Prova del 10 luglio 2014
T280-1	Prova del 4 settembre 2014
T280	Prova del 18 settembre 2014
T281	Prova del 15 gennaio 2015
T282	Prova del 29 gennaio 2015
T283	Prova del 12 febbraio 2015
T284-1	Prova del 4 giugno 2015
T284	Prova del 25 giugno 2015
T284-2	Prova del 9 luglio 2015
T285-1	Prova del 1 settembre 2015
T285	Prova del 15 settembre 2015

T286 Prova del 14 gennaio 2016
T287 Prova del 28 gennaio 2016
T288 Prova del 11 febbraio 2016
T289-1 Prova del 9 giugno 2016
T289 Prova del 23 giugno 2016
T289-2 Prova del 14 luglio 2016
T290-1 Prova del 23 agosto 2016
T290 Prova del 8 settembre 2016
T291 Prova del 20 gennaio 2017
T292 Prova del 6 febbraio 2017
T293 Prova del 21 febbraio 2017
T294-1 Prova del 8 giugno 2017
T294 Prova del 22 giugno 2017
T294-2 Prova del 13 luglio 2017
T295-1 Prova del 21 agosto 2017
T295 Prova del 7 settembre 2017
T296 Prova del 17 gennaio 2018
T297 Prova del 31 gennaio 2018
T298 Prova del 20 febbraio 2018
T299-1 Prova del 4 giugno 2018
T299 Prova del 18 giugno 2018
T299-2 Prova del 10 luglio 2018
T300-1 Prova del 23 agosto 2018
T300 Prova del 6 settembre 2018





T301 Prova del 16 gennaio 2019
T302 Prova del 30 gennaio 2019
T303 Prova del 19 febbraio 2019
T304-1 Prova del 3 giugno 2019
T304 Prova del 17 giugno 2019
T304-2 Prova del 10 luglio 2019
T305-1 Prova del 29 agosto 2019
T305 Prova del 5 settembre 2019
T305-2 Prova del 12 settembre 2019
T306 Prova del 15 gennaio 2020
T307 Prova del 29 gennaio 2020
T308 Prova del 19 febbraio 2020

Tema T271

Prova scritta di STATISTICA del 10.01.2013

Prova scritta di
STATISTICA del
10.01.2013

Giuseppe Boari
Gabriele Cantaluppi



UNIVERSITÀ
CATTOLICA
del Sacro Cuore

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.

Esercizio 5.

Giuseppe Boari
Gabriele Cantaluppi
Università Cattolica del Sacro Cuore, Milano



Con riferimento a 210 soggetti sono stati rilevati il genere, X , (con modalità m e f) e la spesa mensile, Y ripartita in classi. I seguenti prospetti riportano la distribuzione delle frequenze assolute di X e delle densità di frequenza di Y :

x_i	n_i
m	30
f	180

$h_{i-1} \rightarrow h_i$	d_i
10 $\rightarrow$ 20	6
20 $\rightarrow$ 40	4.5
40 $\rightarrow$ 50	6

1. Si rappresentino graficamente le distribuzioni delle frequenze di X e Y e si riassumano le due variabili con opportuni indici di posizione e di variabilità.

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.

Esercizio 5.



Si considera la distribuzione della variabile X ,

x_i	n_i
m	30
f	180

Si tratta di un carattere qualitativo sconnesso che può essere rappresentato graficamente mediante un grafico a bastoncini o a canne d'organo.

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.

Esercizio 5.



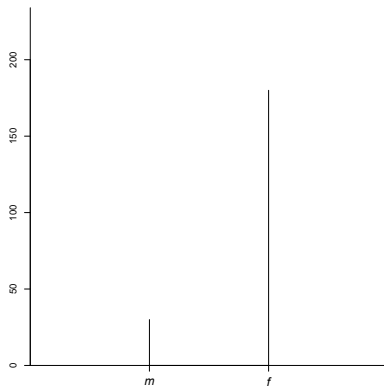
Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.

Esercizio 5.





Si tratta di un carattere qualitativo sconnesso.

L'unico indice di posizione che può essere definito è la moda (categoria alla quale è associata la massima frequenza) che risulta f .

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.

Esercizio 5.



In presenza di un carattere qualitativo la variabilità può essere misurata mediante gli indici di Gini e di Frosini normalizzati

$$G_N = \frac{1 - \sum_{i=1}^k f_i^2}{1 - \frac{1}{k}}$$

e

$$F_N = 1 - \sqrt{1 - G_N}.$$

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.

Esercizio 5.



In presenza di un carattere qualitativo la variabilità può essere misurata mediante gli indici di Gini e di Frosini normalizzati

$$G_N = \frac{1 - \sum_{i=1}^k f_i^2}{1 - \frac{1}{k}}$$

e

$$F_N = 1 - \sqrt{1 - G_N}.$$

Si ottiene:

x_i	n_i	f_i	f_i^2
m	30	0.1429	0.0204
f	180	0.8571	0.7347
	210	1	0.7551

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.

Esercizio 5.



In presenza di un carattere qualitativo la variabilità può essere misurata mediante gli indici di Gini e di Frosini normalizzati

$$G_N = \frac{1 - \sum_{i=1}^k f_i^2}{1 - \frac{1}{k}}$$

e

$$F_N = 1 - \sqrt{1 - G_N}.$$

Si ottiene:

x_i	n_i	f_i	f_i^2
m	30	0.1429	0.0204
f	180	0.8571	0.7347
	210	1	0.7551

Da cui segue:

$$G_N = \frac{1 - 0.7551}{1 - \frac{1}{2}} = 0.4898$$

e

$$F_N = 1 - \sqrt{1 - 0.4898} = 0.2857.$$

Sussiste, quindi, un livello molto basso di eterogeneità.



Moltiplicando le densità di frequenza, d_i , per le ampiezze di classe, a_i , si ottengono le frequenze assolute n_i della variabile Y ,

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.

Esercizio 5.



Moltiplicando le densità di frequenza, d_i , per le ampiezze di classe, a_i , si ottengono le frequenze assolute n_i della variabile Y ,

y_i	d_i	a_i	n_i
10 ÷ 20	6	$20 - 10 = 10$	60
20 ÷ 40	4.5	$40 - 20 = 20$	90
40 ÷ 50	6	$50 - 40 = 10$	60

Esercizio 1.

Esercizio 2.

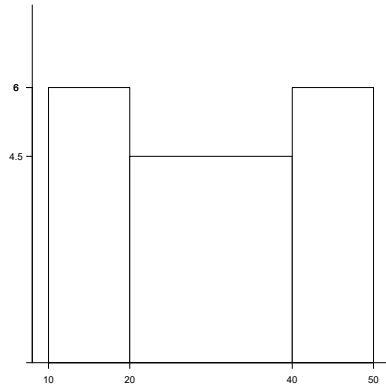
Esercizio 3.

Esercizio 4.

Esercizio 5.



Si tratta di un carattere quantitativo con modalità raggruppate in classi che può essere rappresentato graficamente mediante un istogramma.



Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.

Esercizio 5.



Si tratta di un carattere quantitativo con modalità raggruppate in classi.

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.

Esercizio 5.



La media aritmetica

$$M(X) = \frac{1}{n} \sum_{i=1}^k x_i n_i$$

risulta 30.

Si osserva, infatti, che la distribuzione è simmetrica.

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.

Esercizio 5.



La varianza

$$\text{Var}(X) = M[(X - \mu_X)^2] = M(X^2) - \mu_X^2 = \frac{1}{n} \sum_{i=1}^k x_i^2 n_i - \mu_X^2$$

risulta 128.5714.

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.

Esercizio 5.



La varianza

$$\text{Var}(X) = M[(X - \mu_X)^2] = M(X^2) - \mu_X^2 = \frac{1}{n} \sum_{i=1}^k x_i^2 n_i - \mu_X^2$$

risulta 128.5714.

Si ottiene (considerando come modalità i valori centrali delle classi)

x_i	n_i	$x_i n_i$	$x_i^2 n_i$
15	60	900	13500
30	90	2700	81000
45	60	2700	121500
	210	6300	216000

da cui segue

$$\begin{aligned} \text{Var}(X) = \sigma_X^2 &= M(X^2) - \mu_X^2 = \\ &= \frac{216000}{210} - 30^2 = 1028.5714 - 900 = 128.5714. \end{aligned}$$

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.

Esercizio 5.



Per effettuare confronti in termini di variabilità tra caratteri quantitativi occorre utilizzare una misura relativa, ad esempio il coefficiente di variazione

$$CV = \frac{\sigma}{\mu} = \frac{\sqrt{\text{Var}(X)}}{M(X)} = \frac{\sqrt{128.5714}}{30} = \frac{11.3389}{30} = 0.378.$$

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.

Esercizio 5.



2. Si calcolino la media armonica e una misura di asimmetria per Y .

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.

Esercizio 5.



Si tratta di un carattere quantitativo con modalità raggruppate in classi.

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.

Esercizio 5.



La media armonica

$$\left[M(X^{-1}) \right]^{-1} = \left(\frac{1}{n} \sum_{i=1}^k x_i^{-1} n_i \right)^{-1}$$

risulta 25.2.

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.

Esercizio 5.



La media armonica

$$\left[M(X^{-1}) \right]^{-1} = \left(\frac{1}{n} \sum_{i=1}^k x_i^{-1} n_i \right)^{-1}$$

risulta 25.2. Si ottiene (considerando come modalità i valori centrali delle classi)

x_i	n_i	$x_i^{-1} n_i$
15	60	4.0000
30	90	3.0000
45	60	1.3333
210		8.3333

da cui segue

$$\left[M(X^{-1}) \right]^{-1} = \frac{1}{\frac{1}{n} \sum_{i=1}^k x_i^{-1} n_i} = \frac{1}{\frac{8.3333}{210}} = 25.2.$$

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.

Esercizio 5.



Indice di asimmetria: 0 (la distribuzione è simmetrica)

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.

Esercizio 5.



Si consideri la seguente serie storica delle vendite di un'azienda tra il 2006 e il 2010:

t	x_t
2006	205
2007	290
2008	315
2009	340
2010	225

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.

Esercizio 5.



1. Si costruisca la corrispondente serie dei numeri indici a base mobile.

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.

Esercizio 5.



1. Si costruisca la corrispondente serie dei numeri indici a base mobile.

Si ricorda che ${}_{t-1}I_t = \frac{x_t}{x_{t-1}}$.

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.

Esercizio 5.



1. Si costruisca la corrispondente serie dei numeri indici a base mobile.
Si ricorda che ${}_t I_t = \frac{x_t}{x_{t-1}}$. Si ottiene quindi

Anno	x_t	NIBM
2006	205	
2007	290	1.4146
2008	315	1.0862
2009	340	1.0794
2010	225	0.6618

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.

Esercizio 5.



2. Si calcoli il tasso medio di variazione tra il 2008 e il 2010, commentando il risultato ottenuto.

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.

Esercizio 5.



2. Si calcoli il tasso medio di variazione tra il 2008 e il 2010, commentando il risultato ottenuto.

Si ricorda che

$${}_{2008}I_{2010} = {}_{2008}I_{2009} \cdot {}_{2009}I_{2010}$$

Si deve trovare il numero indice a base mobile medio.

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.

Esercizio 5.



Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.

Esercizio 5.

2. Si calcoli il tasso medio di variazione tra il 2008 e il 2010, commentando il risultato ottenuto.

Si ricorda che

$${}_{2008}I_{2010} = {}_{2008}I_{2009} \cdot {}_{2009}I_{2010}$$

Si deve trovare il numero indice a base mobile medio. Applicando il criterio di scelta della media secondo Chisini si ottiene

$${}_{2008}I_{2010} = {}_{2008}I_{2009} \cdot {}_{2009}I_{2010} = \alpha \cdot \alpha = \alpha^2.$$



Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.

Esercizio 5.

2. Si calcoli il tasso medio di variazione tra il 2008 e il 2010, commentando il risultato ottenuto.

Si ricorda che

$${}_{2008}I_{2010} = {}_{2008}I_{2009} \cdot {}_{2009}I_{2010}$$

Si deve trovare il numero indice a base mobile medio. Applicando il criterio di scelta della media secondo Chisini si ottiene

$${}_{2008}I_{2010} = {}_{2008}I_{2009} \cdot {}_{2009}I_{2010} = \alpha \cdot \alpha = \alpha^2.$$

Quindi

$$\alpha = \sqrt[2]{{}_{2008}I_{2010}} = \sqrt[2]{{}_{2008}I_{2009} \cdot {}_{2009}I_{2010}} = 0.8452.$$

Da cui segue una variazione relativa media pari a -0.1548 , corrispondente a un decremento medio annuo tra il 2008 e il 2010 del 15.48%.



Il seguente prospetto riporta i valori rilevati in corrispondenza a 5 unità statistiche con riferimento a due caratteri quantitativi X e Y

x_i	6	6	7	8	8
y_i	62	58	55	45	55

1. Si rappresentino graficamente le coppie di dati e la funzione di regressione di Y in funzione di X .

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.

Esercizio 5.



La funzione di regressione associa a ogni valore x appartenente all'insieme di definizione della variabile X la media di Y condizionata a x , $M(Y|x)$

x	$M(Y x)$
6	$\frac{1}{2}(58 + 62) = 60$
7	55
8	$\frac{1}{2}(45 + 55) = 50$

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.

Esercizio 5.



Esercizio 1.

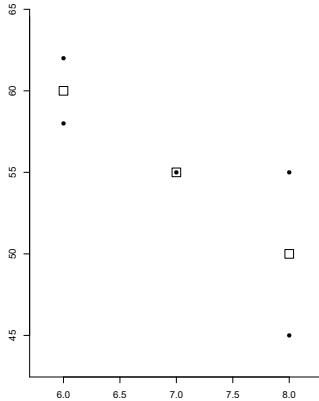
Esercizio 2.

Esercizio 3.

Esercizio 4.

Esercizio 5.

Le coppie di dati possono essere rappresentate mediante un grafico a dispersione.





2. Si calcolino, secondo il metodo dei minimi quadrati, i parametri e l'indice di adattamento dei seguenti modelli di regressione:

$$I) Y^* = a + bX \quad II) Y^* = cX$$

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.

Esercizio 5.

Si riportano i calcoli necessari¹ per ottenere le stime dei parametri e gli indici di adattamento del modello

$$Y = a + bX + E$$

x_i	y_i	$x_i y_i$	x_i^2	y_i^2
6	62	372	36	3844
6	58	348	36	3364
7	55	385	49	3025
8	45	360	64	2025
8	55	440	64	3025
35	275	1905	249	15283
7	55	381	49.8	3056.6
$M(X)$	$M(Y)$	$M(XY)$	$M(X^2)$	$M(Y^2)$

$$\hat{b} =$$



Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.

Esercizio 5.

¹Si osserva che per ragioni di stampa alcuni calcoli potrebbero essere duplicati con riferimento ai diversi modelli oggetto di studio

Si riportano i calcoli necessari¹ per ottenere le stime dei parametri e gli indici di adattamento del modello

$$Y = a + bX + E$$

x_i	y_i	$x_i y_i$	x_i^2	y_i^2
6	62	372	36	3844
6	58	348	36	3364
7	55	385	49	3025
8	45	360	64	2025
8	55	440	64	3025
35	275	1905	249	15283
7	55	381	49.8	3056.6
$M(X)$	$M(Y)$	$M(XY)$	$M(X^2)$	$M(Y^2)$

$$\begin{aligned}\hat{b} &= \frac{\text{Cov}(X, Y)}{\text{Var}(X)} = \frac{M(XY) - M(X)M(Y)}{M(X^2) - [M(X)]^2} = \\ &= \frac{381 - 7 \cdot 55}{49.8 - 7^2} = \frac{-4}{0.8} = -5 \\ \hat{a} &= \end{aligned}$$

¹Si osserva che per ragioni di stampa alcuni calcoli potrebbero essere duplicati con riferimento ai diversi modelli oggetto di studio



Si riportano i calcoli necessari¹ per ottenere le stime dei parametri e gli indici di adattamento del modello

$$Y = a + bX + E$$

x_i	y_i	$x_i y_i$	x_i^2	y_i^2
6	62	372	36	3844
6	58	348	36	3364
7	55	385	49	3025
8	45	360	64	2025
8	55	440	64	3025
35	275	1905	249	15283
7	55	381	49.8	3056.6
$M(X)$	$M(Y)$	$M(XY)$	$M(X^2)$	$M(Y^2)$

$$\hat{b} = \frac{\text{Cov}(X, Y)}{\text{Var}(X)} = \frac{M(XY) - M(X)M(Y)}{M(X^2) - [M(X)]^2} =$$

$$= \frac{381 - 7 \cdot 55}{49.8 - 7^2} = \frac{-4}{0.8} = -5$$

$$\hat{a} = M(Y) - \hat{b}M(X) = 90$$

$$\rho^2 =$$



Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.

Esercizio 5.

¹Si osserva che per ragioni di stampa alcuni calcoli potrebbero essere duplicati con riferimento ai diversi modelli oggetto di studio

Si riportano i calcoli necessari¹ per ottenere le stime dei parametri e gli indici di adattamento del modello

$$Y = a + bX + E$$

x_i	y_i	$x_i y_i$	x_i^2	y_i^2
6	62	372	36	3844
6	58	348	36	3364
7	55	385	49	3025
8	45	360	64	2025
8	55	440	64	3025
35	275	1905	249	15283
7	55	381	49.8	3056.6
$M(X)$	$M(Y)$	$M(XY)$	$M(X^2)$	$M(Y^2)$

$$\hat{b} = \frac{\text{Cov}(X, Y)}{\text{Var}(X)} = \frac{M(XY) - M(X)M(Y)}{M(X^2) - [M(X)]^2} =$$

$$= \frac{381 - 7 \cdot 55}{49.8 - 7^2} = \frac{-4}{0.8} = -5$$

$$\hat{a} = M(Y) - \hat{b}M(X) = 90$$

$$\rho^2 = \frac{[\text{Cov}(X, Y)]^2}{\text{Var}(X) \text{Var}(Y)} = \frac{(-4)^2}{0.8 \cdot 31.6} = 0.6329$$

$$M(\hat{E}^2) = \text{Var}(\hat{E}) = \text{varianza residua} =$$

¹Si osserva che per ragioni di stampa alcuni calcoli potrebbero essere duplicati con riferimento ai diversi modelli oggetto di studio



Si riportano i calcoli necessari¹ per ottenere le stime dei parametri e gli indici di adattamento del modello

$$Y = a + bX + E$$

x_i	y_i	$x_i y_i$	x_i^2	y_i^2
6	62	372	36	3844
6	58	348	36	3364
7	55	385	49	3025
8	45	360	64	2025
8	55	440	64	3025
35	275	1905	249	15283
7	55	381	49.8	3056.6
$M(X)$	$M(Y)$	$M(XY)$	$M(X^2)$	$M(Y^2)$

$$\hat{b} = \frac{\text{Cov}(X, Y)}{\text{Var}(X)} = \frac{M(XY) - M(X)M(Y)}{M(X^2) - [M(X)]^2} =$$

$$= \frac{381 - 7 \cdot 55}{49.8 - 7^2} = \frac{-4}{0.8} = -5$$

$$\hat{a} = M(Y) - \hat{b}M(X) = 90$$

$$\rho^2 = \frac{[\text{Cov}(X, Y)]^2}{\text{Var}(X) \text{Var}(Y)} = \frac{(-4)^2}{0.8 \cdot 31.6} = 0.6329$$

$$M(\hat{E}^2) = \text{Var}(\hat{E}) = \text{varianza residua} = \text{Var}(Y) \cdot (1 - \rho^2) = 11.6$$

¹Si osserva che per ragioni di stampa alcuni calcoli potrebbero essere duplicati con riferimento ai diversi modelli oggetto di studio





Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.

Esercizio 5.

Si riportano i calcoli necessari² per ottenere le stime dei parametri e la misura di adattamento del modello

$$Y = cX + E$$

x_i	y_i	$x_i y_i$	x_i^2	y_i^2
6	62	372	36	3844
6	58	348	36	3364
7	55	385	49	3025
8	45	360	64	2025
8	55	440	64	3025
		1905	249	15283
		381	49.8	3056.6
		$M(XY)$	$M(X^2)$	$M(Y^2)$

$$\hat{c} =$$

²Si osserva che per ragioni di stampa alcuni calcoli potrebbero essere duplicati con riferimento ai diversi modelli oggetto di studio



Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.

Esercizio 5.

Si riportano i calcoli necessari² per ottenere le stime dei parametri e la misura di adattamento del modello

$$Y = cX + E$$

x_i	y_i	$x_i y_i$	x_i^2	y_i^2
6	62	372	36	3844
6	58	348	36	3364
7	55	385	49	3025
8	45	360	64	2025
8	55	440	64	3025
		1905	249	15283
		381	49.8	3056.6
		$M(XY)$	$M(X^2)$	$M(Y^2)$

$$\hat{c} = \frac{M(XY)}{M(X^2)} = \frac{381}{49.8} = 7.6506$$

$$M(\hat{E}^2) =$$

²Si osserva che per ragioni di stampa alcuni calcoli potrebbero essere duplicati con riferimento ai diversi modelli oggetto di studio



Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.

Esercizio 5.

Si riportano i calcoli necessari² per ottenere le stime dei parametri e la misura di adattamento del modello

$$Y = cX + E$$

x_i	y_i	$x_i y_i$	x_i^2	y_i^2
6	62	372	36	3844
6	58	348	36	3364
7	55	385	49	3025
8	45	360	64	2025
8	55	440	64	3025
		1905	249	15283
		381	49.8	3056.6
		$M(XY)$	$M(X^2)$	$M(Y^2)$

$$\hat{c} = \frac{M(XY)}{M(X^2)} = \frac{381}{49.8} = 7.6506$$

$$M(\hat{E}^2) = M[Y \cdot (Y - \hat{c}X)] = M(Y^2) - \hat{c}M(XY) = 3056.6 - 7.6506 \cdot 381 = 141.7205$$

²Si osserva che per ragioni di stampa alcuni calcoli potrebbero essere duplicati con riferimento ai diversi modelli oggetto di studio



Il modello migliore è $Y^* = 90 - 5X$.

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.

Esercizio 5.



3. Si determini il valore del rapporto di correlazione $\eta^2_{Y|X}$.

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.

Esercizio 5.



3. Si determini il valore del rapporto di correlazione $\eta^2_{Y|X}$.

$$\eta^2_{Y|X} = 0.6329.$$

Le medie condizionate sono allineate; la retta coincide con la funzione di regressione,

$$\rho^2 = \eta^2_{Y|X}.$$

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.

Esercizio 5.

Esercizio 4

Prova scritta di
STATISTICA del
10.01.2013

Giuseppe Boari
Gabriele Cantaluppi



UNIVERSITÀ
CATTOLICA
del Sacro Cuore

Si completino, giustificando opportunamente la risposta, le seguenti tabelle in modo tale che:

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.

Esercizio 5.



a) $\rho^2 = 1$ e $\eta_{X|Y}^2 = 0$

	$y_1 = 2$	$y_2 = \text{—}$	$y_3 = 12$
$x_1 = 2$		3	
$x_2 = 4$	5		5

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.

Esercizio 5.



Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.

Esercizio 5.

a) $\rho^2 = 1$ e $\eta_{X|Y}^2 = 0$

	$y_1 = 2$	$y_2 = \text{—}$	$y_3 = 12$
$x_1 = 2$		3	
$x_2 = 4$	5		5

Impossibile in quanto $\rho^2 \leq \eta_{X|Y}^2$

b) $\rho^2 = \eta_{Y|X}^2 = 0$ e $\eta_{X|Y}^2 = 1$

	$y_1 = 2$	$y_2 = \text{—}$	$y_3 = 12$
$x_1 = 2$	0	3	0
$x_2 = 4$	5	0	5



Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.

Esercizio 5.



Esercizio 1.

Esercizio 2.

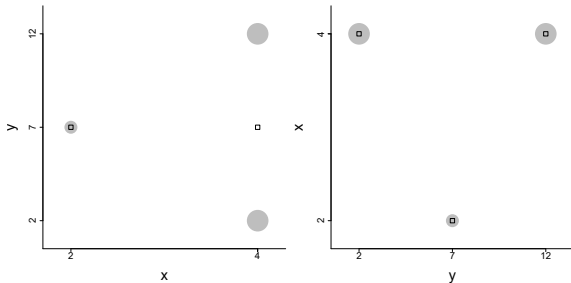
Esercizio 3.

Esercizio 4.

Esercizio 5.

$$b) \rho^2 = \eta_{Y|X}^2 = 0 \quad \text{e} \quad \eta_{X|Y}^2 = 1$$

	$y_1 = 2$	$y_2 = \text{—}$	$y_3 = 12$
$x_1 = 2$	0	3	0
$x_2 = 4$	5	0	5



Dipendenza funzionale di X da $Y \rightarrow X|y_j$ degeneri

Indipendenza in media di Y da $X \rightarrow M(Y|x_1) = M(Y|x_2) = M(Y) = y_2$

Esercizio 5

Si consideri l'esperimento di estrazione di 4 palline da un'urna contenente 21 palline rosse e 29 palline blu.

1. Si calcoli la probabilità di ottenere almeno 3 palline blu nell'ipotesi che le 4 estrazioni siano effettuate senza reimmissione.



Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.

Esercizio 5.

Esercizio 5



Si consideri l'esperimento di estrazione di 4 palline da un'urna contenente 21 palline rosse e 29 palline blu.

1. Si calcoli la probabilità di ottenere almeno 3 palline blu nell'ipotesi che le 4 estrazioni siano effettuate senza reimmissione.

$$\begin{aligned}P(\text{almeno } 3 B) &= P(3B) + P(4B) = \\&= \binom{4}{3} \frac{29}{50} \frac{28}{49} \frac{27}{48} \frac{21}{47} + \binom{4}{4} \frac{29}{50} \frac{28}{49} \frac{27}{48} \frac{26}{47} = \\&= 0.3332 + 0.1031 = 0.4363\end{aligned}$$

o mediante la v.c. ipergeometrica

$$P(X = 3) + P(X = 4) = \frac{\binom{29}{3} \binom{21}{1}}{\binom{50}{4}} + \frac{\binom{29}{4} \binom{21}{0}}{\binom{50}{4}} = 0.4363$$

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.

Esercizio 5.



2. Si calcoli la probabilità di ottenere almeno 36 palline blu nell'ipotesi che siano effettuate 60 estrazioni con reimmissione.

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.

Esercizio 5.



2. Si calcoli la probabilità di ottenere almeno 36 palline blu nell'ipotesi che siano effettuate 60 estrazioni con reimmissione.

Indicata con X la variabile casuale 'numero di successi' nelle 60 prove indipendenti risulta

$$X \sim$$

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.

Esercizio 5.



Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.

Esercizio 5.

2. Si calcoli la probabilità di ottenere almeno 36 palline blu nell'ipotesi che siano effettuate 60 estrazioni con reimmissione.

Indicata con X la variabile casuale 'numero di successi' nelle 60 prove indipendenti risulta

$$X \sim \text{Bin}(n = 60, p = 0.58).$$

$$M(X) = 34.8; \text{Var}(X) = 14.6160;$$

utilizzando l'approssimazione della variabile casuale Binomiale con la variabile casuale Normale

$$X \sim$$



2. Si calcoli la probabilità di ottenere almeno 36 palline blu nell'ipotesi che siano effettuate 60 estrazioni con reimmissione.

Indicata con X la variabile casuale 'numero di successi' nelle 60 prove indipendenti risulta

$$X \sim \text{Bin}(n = 60, p = 0.58).$$

$$M(X) = 34.8; \text{Var}(X) = 14.6160;$$

utilizzando l'approssimazione della variabile casuale Binomiale con la variabile casuale Normale

$$X \sim N(\mu = 34.8, \sigma^2 = 14.6160)$$

si ottiene

$$\begin{aligned} P(X > 36) &= P\left(\frac{X - \mu}{\sigma} > \frac{36 - \mu}{\sigma}\right) = P\left(Z > \frac{36 - 34.8}{3.8231}\right) = \\ &= P(Z > 0.31) = 1 - P(Z \leq 0.31) = 1 - 0.6217 = 0.3783. \end{aligned}$$

Tema T272

Prova scritta di STATISTICA del 24.01.2013

Prova scritta di
STATISTICA del
24.01.2013

Giuseppe Boari
Gabriele Cantaluppi



UNIVERSITÀ
CATTOLICA
del Sacro Cuore

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.

Esercizio 5.

Giuseppe Boari
Gabriele Cantaluppi
Università Cattolica del Sacro Cuore, Milano



Su 225 soggetti è stata rilevata la variabile Y , ripartita in classi, le cui densità di frequenza sono riportate nel prospetto.

$h_{i-1} \vdash h_i$	d_i
10 $\vdash$ 20	6
20 $\vdash$ 40	4.5
40 $\vdash$ 50	6
50 $\vdash$ 55	3

1. Si ricostruisca la tabella delle frequenze relative e si rappresenti graficamente la distribuzione di Y .

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.

Esercizio 5.



Moltiplicando le densità di frequenza, d_i , per le ampiezze di classe, a_i , si ottengono le frequenze assolute n_i della variabile X ,

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.

Esercizio 5.



Moltiplicando le densità di frequenza, d_i , per le ampiezze di classe, a_i , si ottengono le frequenze assolute n_i della variabile X ,

x_i	d_i	a_i	n_i
10 - 20	6	$20 - 10 = 10$	60
20 - 40	4.5	$40 - 20 = 20$	90
40 - 50	6	$50 - 40 = 10$	60
50 - 55	3	$55 - 50 = 5$	15

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.

Esercizio 5.



Le frequenze relative possono essere ottenute dividendo le frequenze assolute per la numerosità totale $n = 225$ delle unità statistiche,

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.

Esercizio 5.



Le frequenze relative possono essere ottenute dividendo le frequenze assolute per la numerosità totale $n = 225$ delle unità statistiche,

x_i	d_i	a_i	n_i	f_i
10 - 20	6	$20 - 10 = 10$	60	0.2667
20 - 40	4.5	$40 - 20 = 20$	90	0.4
40 - 50	6	$50 - 40 = 10$	60	0.2667
50 - 55	3	$55 - 50 = 5$	15	0.0667

Esercizio 1.

Esercizio 2.

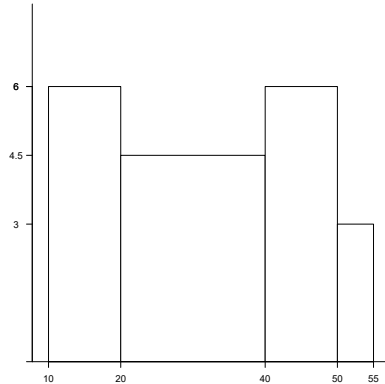
Esercizio 3.

Esercizio 4.

Esercizio 5.



Si tratta di un carattere quantitativo con modalità raggruppate in classi che può essere rappresentato graficamente mediante un istogramma.



Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.

Esercizio 5.



2. Si calcolino moda, media quadratica e coefficiente di variazione. Si individui su un opportuno grafico la mediana.

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.

Esercizio 5.



Si tratta di un carattere quantitativo con modalità raggruppate in classi.

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.

Esercizio 5.



La moda corrisponde alla classe alla quale è associata la massima densità di frequenza

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.

Esercizio 5.



La moda corrisponde alla classe alla quale è associata la massima densità di frequenza

y_i	n_i	a_i	d_i
10 – 20	60	$20 - 10 = 10$	6
20 – 40	90	$40 - 20 = 20$	4.5
40 – 50	60	$50 - 40 = 10$	6
50 – 55	15	$55 - 50 = 5$	3

e risulta 10 – 20, 40 – 50 (situazione bimodale).

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.

Esercizio 5.



La mediana (percentile di ordine 0.50) può essere ottenuta considerando l'inversa della Funzione di Ripartizione in corrispondenza di 0.50.

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.

Esercizio 5.



La mediana (percentile di ordine 0.50) può essere ottenuta considerando l'inversa della Funzione di Ripartizione in corrispondenza di 0.50.

y_i	n_i	N_i	F_i
10 - 20	60	60	0.2667
20 - 40	90	150	0.6667
40 - 50	60	210	0.9333
50 - 55	15	225	1

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.

Esercizio 5.



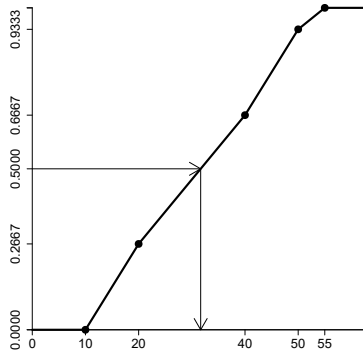
Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.

Esercizio 5.





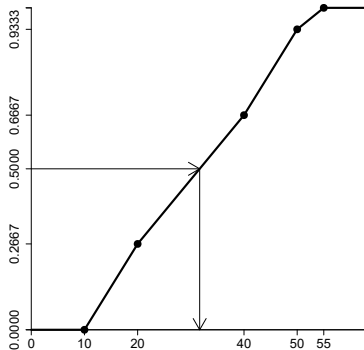
Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.

Esercizio 5.



Si ottiene

$$y_{0.5} = 20 + \frac{0.5 - 0.2667}{0.6667 - 0.2667} \cdot (40 - 20) = 31.6667.$$



La media aritmetica

$$M(Y) = \frac{1}{n} \sum_{i=1}^k y_i n_i$$

risulta 31.5.

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.

Esercizio 5.



La media aritmetica

$$M(Y) = \frac{1}{n} \sum_{i=1}^k y_i n_i$$

risulta 31.5.

Si ottiene (considerando come modalità i valori centrali delle classi)

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.

Esercizio 5.



La media aritmetica

$$M(Y) = \frac{1}{n} \sum_{i=1}^k y_i n_i$$

risulta 31.5.

Si ottiene (considerando come modalità i valori centrali delle classi)

y_i	n_i	$y_i n_i$
15.0	60	900.0
30.0	90	2700.0
45.0	60	2700.0
52.5	15	787.5
225		7087.5

da cui segue

$$M(Y) = \frac{1}{n} \sum_{i=1}^k y_i n_i = \frac{7087.5}{225} = 31.5.$$

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.

Esercizio 5.



La varianza

$$\text{Var}(Y) = M[(Y - \mu_Y)^2] = M(Y^2) - \mu_Y^2 = \frac{1}{n} \sum_{i=1}^k y_i^2 n_i - \mu_Y^2$$

risulta 151.5.

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.

Esercizio 5.



La varianza

$$\text{Var}(Y) = M[(Y - \mu_Y)^2] = M(Y^2) - \mu_Y^2 = \frac{1}{n} \sum_{i=1}^k y_i^2 n_i - \mu_Y^2$$

risulta 151.5.

Si ottiene (considerando come modalità i valori centrali delle classi)

y_i	n_i	$y_i n_i$	$y_i^2 n_i$
15.0	60	900.0	13500.00
30.0	90	2700.0	81000.00
45.0	60	2700.0	121500.00
52.5	15	787.5	41343.75
	225	7087.5	257343.75

da cui segue

$$\begin{aligned} \text{Var}(Y) = \sigma_Y^2 &= M(Y^2) - \mu_Y^2 = \\ &= \frac{257343.75}{225} - 31.5^2 = 1143.75 - 992.25 = 151.5. \end{aligned}$$

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.

Esercizio 5.



Per effettuare confronti in termini di variabilità tra caratteri quantitativi occorre utilizzare una misura relativa, ad esempio il coefficiente di variazione

$$CV = \frac{\sigma}{\mu} = \frac{\sqrt{\text{Var}(Y)}}{M(Y)} = \frac{\sqrt{151.5}}{31.5} = \frac{12.3085}{31.5} = 0.3907.$$

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.

Esercizio 5.



La media quadratica

$$\left[M(Y^2) \right]^{\frac{1}{2}} = \left(\frac{1}{n} \sum_{i=1}^k y_i^2 n_i \right)^{\frac{1}{2}}$$

risulta 33.8194.

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.

Esercizio 5.



La media quadratica

$$\left[M(Y^2) \right]^{\frac{1}{2}} = \left(\frac{1}{n} \sum_{i=1}^k y_i^2 n_i \right)^{\frac{1}{2}}$$

risulta 33.8194. Si ottiene (considerando come modalità i valori centrali delle classi)

y_i	n_i	$y_i^2 n_i$
15.0	60	13500.00
30.0	90	81000.00
45.0	60	121500.00
52.5	15	41343.75
	225	257343.75

da cui segue

$$\left[M(Y^2) \right]^{\frac{1}{2}} = \sqrt{\frac{1}{n} \sum_{i=1}^k y_i^2 n_i} = \sqrt{\frac{257343.75}{225}} = 33.8194.$$

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.

Esercizio 5.



Della serie storica X_t delle vendite di un'azienda tra il 2006 e il 2010 abbiamo le seguenti informazioni:

t	$NIBM$	$NIBF$
2006		100
2007	113	
2008		116
2009	106	
2010		131

1. Si riempiano le caselle vuote della tabella e sapendo che $x_{2010}=650$ si calcoli x_{2006} .

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.

Esercizio 5.

Occorre lavorare con le grandezze relative e, quindi, dividere i valori presenti nella tabella per 100.

$$2007/2008 =$$



Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.

Esercizio 5.

Occorre lavorare con le grandezze relative e, quindi, dividere i valori presenti nella tabella per 100.

$$\begin{aligned} \frac{2007}{2006} I_{2008} &= \frac{\frac{2006}{2006} I_{2008}}{\frac{2006}{2006} I_{2008}} = \frac{1.16}{1.13} = 1.0265 \\ \frac{2006}{2006} I_{2009} &= \end{aligned}$$



Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.

Esercizio 5.

Occorre lavorare con le grandezze relative e, quindi, dividere i valori presenti nella tabella per 100.

$${}_{2007}I_{2008} = \frac{{}_{2006}I_{2008}}{{}_{2006}I_{2008}} = \frac{1.16}{1.13} = 1.0265$$

$${}_{2006}I_{2009} = {}_{2006}I_{2008} \cdot {}_{2008}I_{2009} = 1.16 \cdot 1.06 = 1.2296$$

$${}_{2009}I_{2010} =$$



Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.

Esercizio 5.

Occorre lavorare con le grandezze relative e, quindi, dividere i valori presenti nella tabella per 100.

$$\begin{aligned}
 {}_{2007}I_{2008} &= \frac{{}_{2006}I_{2008}}{{}_{2006}I_{2008}} = \frac{1.16}{1.13} = 1.0265 \\
 {}_{2006}I_{2009} &= {}_{2006}I_{2008} \cdot {}_{2008}I_{2009} = 1.16 \cdot 1.06 = 1.2296 \\
 {}_{2009}I_{2010} &= \frac{{}_{2006}I_{2010}}{{}_{2006}I_{2009}} = \frac{1.31}{1.2296} = 1.0654
 \end{aligned}$$

t	$NIBM$	$NIBF$
2006	–	100
2007	113	113
2008	102.65	116
2009	106	122.96
2010	106.54	131

Ricordando che

$${}_{2010}I_{2006} = 1.31 = \frac{X_{2010}}{X_{2006}}$$

si ricava



Occorre lavorare con le grandezze relative e, quindi, dividere i valori presenti nella tabella per 100.

$$\begin{aligned}
 {}_{2007}l_{2008} &= \frac{{}_{2006}l_{2008}}{{}_{2006}l_{2008}} = \frac{1.16}{1.13} = 1.0265 \\
 {}_{2006}l_{2009} &= {}_{2006}l_{2008} \cdot {}_{2008}l_{2009} = 1.16 \cdot 1.06 = 1.2296 \\
 {}_{2009}l_{2010} &= \frac{{}_{2006}l_{2010}}{{}_{2006}l_{2009}} = \frac{1.31}{1.2296} = 1.0654
 \end{aligned}$$

t	$NIBM$	$NIBF$
2006	–	100
2007	113	113
2008	102.65	116
2009	106	122.96
2010	106.54	131

Ricordando che

$${}_{2010}l_{2006} = 1.31 = \frac{X_{2010}}{X_{2006}}$$

si ricava

$$X_{2006} = \frac{X_{2010}}{{}_{2010}l_{2006}} = \frac{650}{1.31} = 496.1832.$$





2. Si calcoli il tasso medio di variazione tra il 2007 e il 2010, commentando il risultato ottenuto.

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.

Esercizio 5.



2. Si calcoli il tasso medio di variazione tra il 2007 e il 2010, commentando il risultato ottenuto.

Si ricorda che

$${}_{2007}I_{2010} = {}_{2007}I_{2008} \cdot {}_{2008}I_{2009} \cdot {}_{2009}I_{2010}.$$

Si deve trovare il numero indice a base mobile medio.

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.

Esercizio 5.



2. Si calcoli il tasso medio di variazione tra il 2007 e il 2010, commentando il risultato ottenuto.

Si ricorda che

$${}_{2007}I_{2010} = {}_{2007}I_{2008} \cdot {}_{2008}I_{2009} \cdot {}_{2009}I_{2010}.$$

Si deve trovare il numero indice a base mobile medio. Applicando il criterio di scelta della media secondo Chisini si ottiene

$${}_{2007}I_{2010} = {}_{2007}I_{2008} \cdot {}_{2008}I_{2009} \cdot {}_{2009}I_{2010} = \alpha \cdot \alpha \cdot \alpha = \alpha^3.$$

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.

Esercizio 5.



Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.

Esercizio 5.

2. Si calcoli il tasso medio di variazione tra il 2007 e il 2010, commentando il risultato ottenuto.

Si ricorda che

$${}_{2007}l_{2010} = {}_{2007}l_{2008} \cdot {}_{2008}l_{2009} \cdot {}_{2009}l_{2010}.$$

Si deve trovare il numero indice a base mobile medio. Applicando il criterio di scelta della media secondo Chisini si ottiene

$${}_{2007}l_{2010} = {}_{2007}l_{2008} \cdot {}_{2008}l_{2009} \cdot {}_{2009}l_{2010} = \alpha \cdot \alpha \cdot \alpha = \alpha^3.$$

Quindi

$$\alpha = \sqrt[3]{{}_{2007}l_{2010}} = \sqrt[3]{{}_{2007}l_{2008} \cdot {}_{2008}l_{2009} \cdot {}_{2009}l_{2010}} = 1.0505.$$

Da cui segue una variazione relativa media pari a 0.0505 , corrispondente a un incremento medio annuo tra il 2007 e il 2010 del 5.05%.



Il seguente prospetto riporta i valori rilevati in corrispondenza a 6 unità statistiche con riferimento a due caratteri quantitativi X e Y

x_i	1	2	3	4	5	6
y_i	100	100	130	133	180	200

1. Si rappresentino graficamente le coppie di dati.

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.

Esercizio 5.



Esercizio 1.

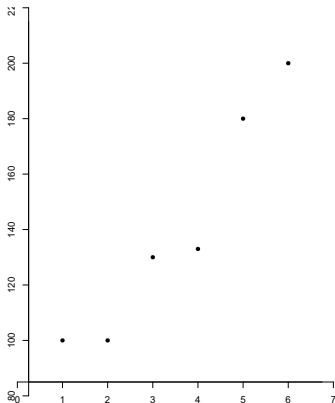
Esercizio 2.

Esercizio 3.

Esercizio 4.

Esercizio 5.

Le coppie di dati possono essere rappresentate mediante un grafico a dispersione.





2. Si calcolino, secondo il metodo dei minimi quadrati, i parametri e l'indice di adattamento dei seguenti modelli di regressione:

$$I) Y^* = a + bX \quad II) Y^* = c + dX^2$$

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.

Esercizio 5.



Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.

Esercizio 5.

Si riportano i calcoli necessari¹ per ottenere le stime dei parametri e gli indici di adattamento del modello

$$Y = a + bX + E$$

x_i	y_i	$x_i y_i$	x_i^2	y_i^2
1	100	100	1	10000
2	100	200	4	10000
3	130	390	9	16900
4	133	532	16	17689
5	180	900	25	32400
6	200	1200	36	40000
21	843	3322	91	126989
3.5	140.5	553.6667	15.1667	21164.8333
$M(X)$	$M(Y)$	$M(XY)$	$M(X^2)$	$M(Y^2)$

$$\hat{b} =$$

¹Si osserva che per ragioni di stampa alcuni calcoli potrebbero essere duplicati con riferimento ai diversi modelli oggetto di studio



Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.

Esercizio 5.

Si riportano i calcoli necessari¹ per ottenere le stime dei parametri e gli indici di adattamento del modello

$$Y = a + bX + E$$

x_i	y_i	$x_i y_i$	x_i^2	y_i^2
1	100	100	1	10000
2	100	200	4	10000
3	130	390	9	16900
4	133	532	16	17689
5	180	900	25	32400
6	200	1200	36	40000
21	843	3322	91	126989
3.5	140.5	553.6667	15.1667	21164.8333
$M(X)$	$M(Y)$	$M(XY)$	$M(X^2)$	$M(Y^2)$

$$\begin{aligned}\hat{b} &= \frac{\text{Cov}(X, Y)}{\text{Var}(X)} = \frac{M(XY) - M(X)M(Y)}{M(X^2) - [M(X)^2]} = \\ &= \frac{553.6667 - 3.5 \cdot 140.5}{15.1667 - 3.5^2} = \frac{61.9167}{2.9167} = 21.2283 \\ \hat{a} &= \end{aligned}$$

¹Si osserva che per ragioni di stampa alcuni calcoli potrebbero essere duplicati con riferimento ai diversi modelli oggetto di studio



Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.

Esercizio 5.

Si riportano i calcoli necessari¹ per ottenere le stime dei parametri e gli indici di adattamento del modello

$$Y = a + bX + E$$

x_i	y_i	$x_i y_i$	x_i^2	y_i^2
1	100	100	1	10000
2	100	200	4	10000
3	130	390	9	16900
4	133	532	16	17689
5	180	900	25	32400
6	200	1200	36	40000
21	843	3322	91	126989
3.5	140.5	553.6667	15.1667	21164.8333
$M(X)$	$M(Y)$	$M(XY)$	$M(X^2)$	$M(Y^2)$

$$\begin{aligned}\hat{b} &= \frac{\text{Cov}(X, Y)}{\text{Var}(X)} = \frac{M(XY) - M(X)M(Y)}{M(X^2) - [M(X)^2]} = \\ &= \frac{553.6667 - 3.5 \cdot 140.5}{15.1667 - 3.5^2} = \frac{61.9167}{2.9167} = 21.2283\end{aligned}$$

$$\hat{a} = M(Y) - \hat{b}M(X) = 66.2008$$

$$\rho^2 =$$

¹Si osserva che per ragioni di stampa alcuni calcoli potrebbero essere duplicati con riferimento ai diversi modelli oggetto di studio



Si riportano i calcoli necessari¹ per ottenere le stime dei parametri e gli indici di adattamento del modello

$$Y = a + bX + E$$

x_i	y_i	$x_i y_i$	x_i^2	y_i^2
1	100	100	1	10000
2	100	200	4	10000
3	130	390	9	16900
4	133	532	16	17689
5	180	900	25	32400
6	200	1200	36	40000
21	843	3322	91	126989
3.5	140.5	553.6667	15.1667	21164.8333
$M(X)$	$M(Y)$	$M(XY)$	$M(X^2)$	$M(Y^2)$

$$\begin{aligned}\hat{b} &= \frac{\text{Cov}(X, Y)}{\text{Var}(X)} = \frac{M(XY) - M(X)M(Y)}{M(X^2) - [M(X)^2]} = \\ &= \frac{553.6667 - 3.5 \cdot 140.5}{15.1667 - 3.5^2} = \frac{61.9167}{2.9167} = 21.2283\end{aligned}$$

$$\hat{a} = M(Y) - \hat{b}M(X) = 66.2008$$

$$\rho^2 = \frac{[\text{Cov}(X, Y)]^2}{\text{Var}(X) \text{Var}(Y)} = \frac{61.9167^2}{2.9167 \cdot 1424.5833} = 0.9226$$

¹Si osserva che per ragioni di stampa alcuni calcoli potrebbero essere duplicati con riferimento ai diversi modelli oggetto di studio

Si riportano i calcoli necessari² per ottenere le stime dei parametri e gli indici di adattamento del modello

$$Y = c + dX^2 + E$$

x_i^2	y_i	$x_i^2 y_i$	$x_i^{2^2}$	y_i^2
1	100	100	1	10000
4	100	400	16	10000
9	130	1170	81	16900
16	133	2128	256	17689
25	180	4500	625	32400
36	200	7200	1296	40000
91	843	15498	2275	126989
15.1667	140.5	2583	379.1667	21164.8333
$M(X^2)$	$M(Y)$	$M(X^2 Y)$	$M(X^{2^2})$	$M(Y^2)$

$$\hat{d} =$$



Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.

Esercizio 5.

²Si osserva che per ragioni di stampa alcuni calcoli potrebbero essere duplicati con riferimento ai diversi modelli oggetto di studio

Si riportano i calcoli necessari² per ottenere le stime dei parametri e gli indici di adattamento del modello

$$Y = c + dX^2 + E$$

x_i^2	y_i	$x_i^2 y_i$	$x_i^{2^2}$	y_i^2
1	100	100	1	10000
4	100	400	16	10000
9	130	1170	81	16900
16	133	2128	256	17689
25	180	4500	625	32400
36	200	7200	1296	40000
91	843	15498	2275	126989
15.1667	140.5	2583	379.1667	21164.8333
$M(X^2)$	$M(Y)$	$M(X^2 Y)$	$M(X^{2^2})$	$M(Y^2)$

$$\begin{aligned}\hat{d} &= \frac{\text{Cov}(X^2, Y)}{\text{Var}(X^2)} = \frac{M(X^2 Y) - M(X^2) M(Y)}{M(X^{2^2}) - [M(X^2)]^2} = \\ &= \frac{2583 - 15.1667 \cdot 140.5}{379.1667 - 15.1667^2} = \frac{452.0833}{149.1389} = 3.0313 \\ \hat{c} &= \end{aligned}$$

²Si osserva che per ragioni di stampa alcuni calcoli potrebbero essere duplicati con riferimento ai diversi modelli oggetto di studio



Si riportano i calcoli necessari² per ottenere le stime dei parametri e gli indici di adattamento del modello

$$Y = c + dX^2 + E$$

x_i^2	y_i	$x_i^2 y_i$	x_i^{22}	y_i^2
1	100	100	1	10000
4	100	400	16	10000
9	130	1170	81	16900
16	133	2128	256	17689
25	180	4500	625	32400
36	200	7200	1296	40000
91	843	15498	2275	126989
15.1667	140.5	2583	379.1667	21164.8333
$M(X^2)$	$M(Y)$	$M(X^2 Y)$	$M(X^{22})$	$M(Y^2)$

$$\hat{d} = \frac{\text{Cov}(X^2, Y)}{\text{Var}(X^2)} = \frac{M(X^2 Y) - M(X^2) M(Y)}{M(X^{22}) - [M(X^2)]^2} =$$

$$= \frac{2583 - 15.1667 \cdot 140.5}{379.1667 - 15.1667^2} = \frac{452.0833}{149.1389} = 3.0313$$

$$\hat{c} = M(Y) - \hat{d} M(X^2) = 94.5254$$

$$\rho^2 =$$

²Si osserva che per ragioni di stampa alcuni calcoli potrebbero essere duplicati con riferimento ai diversi modelli oggetto di studio



Si riportano i calcoli necessari² per ottenere le stime dei parametri e gli indici di adattamento del modello

$$Y = c + dX^2 + E$$

x_i^2	y_i	$x_i^2 y_i$	$x_i^{2^2}$	y_i^2
1	100	100	1	10000
4	100	400	16	10000
9	130	1170	81	16900
16	133	2128	256	17689
25	180	4500	625	32400
36	200	7200	1296	40000
91	843	15498	2275	126989
15.1667	140.5	2583	379.1667	21164.8333
$M(X^2)$	$M(Y)$	$M(X^2 Y)$	$M(X^{2^2})$	$M(Y^2)$

$$\begin{aligned}\hat{d} &= \frac{\text{Cov}(X^2, Y)}{\text{Var}(X^2)} = \frac{M(X^2 Y) - M(X^2) M(Y)}{M(X^{2^2}) - [M(X^2)]^2} = \\ &= \frac{2583 - 15.1667 \cdot 140.5}{379.1667 - 15.1667^2} = \frac{452.0833}{149.1389} = 3.0313 \\ \hat{c} &= M(Y) - \hat{d} M(X^2) = 94.5254 \\ \rho^2 &= \frac{[\text{Cov}(X^2, Y)]^2}{\text{Var}(X^2) \text{Var}(Y)} = \frac{452.0833^2}{149.1389 \cdot 1424.5833} = 0.962\end{aligned}$$

²Si osserva che per ragioni di stampa alcuni calcoli potrebbero essere duplicati con riferimento ai diversi modelli oggetto di studio





Il modello migliore è $Y^* = 94.5254 + 3.0313X^2$.

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.

Esercizio 5.



3. Si determinino i valori dei rapporti di correlazione $\eta_{Y|X}^2$ e $\eta_{X|Y}^2$.

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.

Esercizio 5.



3. Si determinino i valori dei rapporti di correlazione $\eta_{Y|X}^2$ e $\eta_{X|Y}^2$.
 $\eta_{Y|X}^2 = 1$ in quanto sussiste dipendenza funzionale di Y da X ;

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.

Esercizio 5.



3. Si determinino i valori dei rapporti di correlazione $\eta_{Y|X}^2$ e $\eta_{X|Y}^2$.
 $\eta_{Y|X}^2 = 1$ in quanto sussiste dipendenza funzionale di Y da X ;
Si procede al calcolo del rapporto di correlazione $\eta_{X|Y}^2$.

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.

Esercizio 5.



Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.

Esercizio 5.

Il rapporto di correlazione

$$\eta_{X|Y}^2 = \frac{\text{Var}_Y[M(X|Y)]}{\text{Var}(X)}$$

può essere ottenuto considerando il rapporto tra la varianza spiegata $\text{Var}_Y[M(X|Y)]$,
varianza della variabile statistica medie condizionate

y	$M(X y)$
100	1.5
130	3.0
133	4.0
180	5.0
200	6.0

e la varianza della variabile dipendente $\text{Var}(X) = 2.9167$.

La varianza spiegata risulta $\text{Var}_Y[M(X|Y)] = 2.8333$.



Si ottiene

y	$M(X y)$	$n(y)$	$M(X y) \cdot n(y)$	$M(X^2 y) \cdot n(y)$
100	1.5	2	3	4.5
130	3.0	1	3	9.0
133	4.0	1	4	16.0
180	5.0	1	5	25.0
200	6.0	1	6	36.0
		6	21	90.5

da cui segue

$$\text{Var}_Y[M(X|Y)] = \frac{90.5}{6} - 3.5^2 = 15.0833 - 12.25 = 2.8333.$$

Il rapporto di correlazione $\eta_{X|Y}^2$ risulta:

$$\eta_{X|Y}^2 = \frac{\text{Var}_Y[M(X|Y)]}{\text{Var}(X)} = \frac{2.8333}{2.9167} = 0.9714.$$

Sussiste, quindi, un livello di dipendenza molto alto.



Il rapporto di correlazione $\eta^2_{X|Y}$ può anche essere ottenuto nel modo seguente.

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.

Esercizio 5.



Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.

Esercizio 5.

Il rapporto di correlazione

$$\eta_{X|Y}^2 = 1 - \frac{M_Y[Var(X|Y)]}{Var(X)}$$

può essere ottenuto considerando il complemento a 1 del rapporto tra la varianza residua $M_Y[Var(X|Y)]$, media della variabile statistica varianze condizionate

y	$Var(X y)$
100	0.25
130	0.00
133	0.00
180	0.00
200	0.00

e la varianza della variabile dipendente $Var(X) = 2.9167$.

La varianza residua risulta $M_Y[Var(X|Y)] = 0.0833$.



Si ottiene

y	$Var(X y)$	$n(y)$	$Var(X y) \cdot n(y)$
100	0.25	2	0.5
130	0.00	1	0.0
133	0.00	1	0.0
180	0.00	1	0.0
200	0.00	1	0.0
		6	0.5

da cui segue

$$M_Y[Var(X|Y)] = \frac{0.5}{6} = 0.0833.$$

Il rapporto di correlazione $\eta_{X|Y}^2$ risulta:

$$\eta_{X|Y}^2 = 1 - \frac{M_Y[Var(X|Y)]}{Var(X)} = 1 - \frac{0.0833}{2.9167} = 0.9714.$$

Sussiste, quindi, un livello di dipendenza molto alto.

Esercizio 4

Prova scritta di
STATISTICA del
24.01.2013

Giuseppe Boari
Gabriele Cantaluppi



UNIVERSITÀ
CATTOLICA
del Sacro Cuore

Si completi la tabella, giustificando opportunamente la risposta, in modo che $\rho^2 < \eta_{Y|X}^2$

	$y_1 = 2$	$y_2 = \text{—}$	$y_3 = 12$
$x_1 = 2$		3	
$x_2 = 4$	5		5

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.

Esercizio 5.



Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.

Esercizio 5.

Si completi la tabella, giustificando opportunamente la risposta, in modo che $\rho^2 < \eta_{Y|X}^2$

	$y_1 = 2$	$y_2 = \text{—}$	$y_3 = 12$
$x_1 = 2$		3	
$x_2 = 4$	5		5

Impossibile, si hanno due distribuzioni condizionate $Y|x = 2$ e $Y|x = 4$, le 2 medie condizionate sono allineate e quindi $\rho^2 = \eta_{Y|X}^2$

Esercizio 5

Si consideri l'esperimento di lancio di 10 monete equilibrate e 6 dadi equilibrati.

1. Si calcoli la probabilità di ottenere 5 teste e 5 facce con numero pari.



Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.

Esercizio 5.

Esercizio 5



Si consideri l'esperimento di lancio di 10 monete equilibrate e 6 dadi equilibrati.

1. Si calcoli la probabilità di ottenere 5 teste e 5 facce con numero pari.

Siano:

X la variabile casuale 'n. di teste nel lancio delle 10 monete' e

Y la variabile casuale 'n. di facce con numero pari nel lancio dei 6 dadi'.

$$X \sim$$

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.

Esercizio 5.

Esercizio 5



Si consideri l'esperimento di lancio di 10 monete equilibrate e 6 dadi equilibrati.

1. Si calcoli la probabilità di ottenere 5 teste e 5 facce con numero pari.

Siano:

X la variabile casuale 'n. di teste nel lancio delle 10 monete' e

Y la variabile casuale 'n. di facce con numero pari nel lancio dei 6 dadi'.

$$X \sim \text{Bin}(n = 10, p = 0.5) \text{ e } Y \sim$$

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.

Esercizio 5.



[Esercizio 1.](#)

[Esercizio 2.](#)

[Esercizio 3.](#)

[Esercizio 4.](#)

[Esercizio 5.](#)

Si consideri l'esperimento di lancio di 10 monete equilibrate e 6 dadi equilibrati.

1. Si calcoli la probabilità di ottenere 5 teste e 5 facce con numero pari.

Siano:

X la variabile casuale 'n. di teste nel lancio delle 10 monete' e

Y la variabile casuale 'n. di facce con numero pari nel lancio dei 6 dadi'.

$$X \sim \text{Bin}(n = 10, p = 0.5) \text{ e } Y \sim \text{Bin}(n = 6, p = 0.5).$$

Il risultato del lancio dei dadi non dipende da quello delle monete, X e Y sono stocasticamente indipendenti, quindi

$$P\{(X = 5) \cap (Y = 5)\} =$$



Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.

Esercizio 5.

Si consideri l'esperimento di lancio di 10 monete equilibrate e 6 dadi equilibrati.

1. Si calcoli la probabilità di ottenere 5 teste e 5 facce con numero pari.

Siano:

X la variabile casuale 'n. di teste nel lancio delle 10 monete' e

Y la variabile casuale 'n. di facce con numero pari nel lancio dei 6 dadi'.

$$X \sim \text{Bin}(n = 10, p = 0.5) \text{ e } Y \sim \text{Bin}(n = 6, p = 0.5).$$

Il risultato del lancio dei dadi non dipende da quello delle monete, X e Y sono stocasticamente indipendenti, quindi

$$\begin{aligned} P\{(X = 5) \cap (Y = 5)\} &= P(X = 5) \cdot P(Y = 5) = \\ &= \end{aligned}$$



Si consideri l'esperimento di lancio di 10 monete equilibrate e 6 dadi equilibrati.

1. Si calcoli la probabilità di ottenere 5 teste e 5 facce con numero pari.

Siano:

X la variabile casuale 'n. di teste nel lancio delle 10 monete' e

Y la variabile casuale 'n. di facce con numero pari nel lancio dei 6 dadi'.

$$X \sim \text{Bin}(n = 10, p = 0.5) \text{ e } Y \sim \text{Bin}(n = 6, p = 0.5).$$

Il risultato del lancio dei dadi non dipende da quello delle monete, X e Y sono stocasticamente indipendenti, quindi

$$\begin{aligned} P\{(X = 5) \cap (Y = 5)\} &= P(X = 5) \cdot P(Y = 5) = \\ &= \binom{10}{5} 0.5^5 0.5^5 \cdot \binom{6}{5} 0.5^5 0.5 = \\ &= 0.2461 \cdot 0.0938 = 0.0231. \end{aligned}$$



Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.

Esercizio 5.

2. Ipotizzando di lanciare 100 monete e 100 dadi, si calcoli la probabilità di ottenere almeno 50 teste e almeno 39 numeri pari.



Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.

Esercizio 5.

2. Ipotizzando di lanciare 100 monete e 100 dadi, si calcoli la probabilità di ottenere almeno 50 teste e almeno 39 numeri pari.

Siano X la variabile casuale 'n. di teste nel lancio delle 100 monete' e

Y la variabile casuale 'n. di facce con numero pari nel lancio dei 100 dadi'.

$$X \sim$$



2. Ipotizzando di lanciare 100 monete e 100 dadi, si calcoli la probabilità di ottenere almeno 50 teste e almeno 39 numeri pari.

Siano X la variabile casuale 'n. di teste nel lancio delle 100 monete' e

Y la variabile casuale 'n. di facce con numero pari nel lancio dei 100 dadi'.

$$X \sim \text{Bin}(n = 100, p = 0.5) \text{ e } Y \sim$$

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.

Esercizio 5.



Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.

Esercizio 5.

2. Ipotizzando di lanciare 100 monete e 100 dadi, si calcoli la probabilità di ottenere almeno 50 teste e almeno 39 numeri pari.

Siano X la variabile casuale 'n. di teste nel lancio delle 100 monete' e

Y la variabile casuale 'n. di facce con numero pari nel lancio dei 100 dadi'.

$$X \sim \text{Bin}(n = 100, p = 0.5) \text{ e } Y \sim \text{Bin}(n = 100, p = 0.5).$$

$$M(X) =$$



2. Ipotizzando di lanciare 100 monete e 100 dadi, si calcoli la probabilità di ottenere almeno 50 teste e almeno 39 numeri pari.

Siano X la variabile casuale 'n. di teste nel lancio delle 100 monete' e

Y la variabile casuale 'n. di facce con numero pari nel lancio dei 100 dadi'.

$$X \sim \text{Bin}(n = 100, p = 0.5) \text{ e } Y \sim \text{Bin}(n = 100, p = 0.5).$$

$$M(X) = M(Y) = 50 \quad \text{Var}(X) =$$

[Esercizio 1.](#)

[Esercizio 2.](#)

[Esercizio 3.](#)

[Esercizio 4.](#)

[Esercizio 5.](#)



Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.

Esercizio 5.

2. Ipotizzando di lanciare 100 monete e 100 dadi, si calcoli la probabilità di ottenere almeno 50 teste e almeno 39 numeri pari.

Siano X la variabile casuale 'n. di teste nel lancio delle 100 monete' e

Y la variabile casuale 'n. di facce con numero pari nel lancio dei 100 dadi'.

$$X \sim \text{Bin}(n = 100, p = 0.5) \text{ e } Y \sim \text{Bin}(n = 100, p = 0.5).$$

$$M(X) = M(Y) = 50 \quad \text{Var}(X) = \text{Var}(Y) = 25.$$

Utilizzando l'approssimazione della variabile casuale Binomiale con la variabile casuale Normale

$X \sim$



Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.

Esercizio 5.

2. Ipotizzando di lanciare 100 monete e 100 dadi, si calcoli la probabilità di ottenere almeno 50 teste e almeno 39 numeri pari.

Siano X la variabile casuale 'n. di teste nel lancio delle 100 monete' e

Y la variabile casuale 'n. di facce con numero pari nel lancio dei 100 dadi'.

$$X \sim \text{Bin}(n = 100, p = 0.5) \text{ e } Y \sim \text{Bin}(n = 100, p = 0.5).$$

$$M(X) = M(Y) = 50 \quad \text{Var}(X) = \text{Var}(Y) = 25.$$

Utilizzando l'approssimazione della variabile casuale Binomiale con la variabile casuale Normale

$$X \sim N(\mu = 50, \sigma^2 = 25) \text{ e } Y \sim$$



Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.

Esercizio 5.

2. Ipotizzando di lanciare 100 monete e 100 dadi, si calcoli la probabilità di ottenere almeno 50 teste e almeno 39 numeri pari.

Siano X la variabile casuale 'n. di teste nel lancio delle 100 monete' e

Y la variabile casuale 'n. di facce con numero pari nel lancio dei 100 dadi'.

$$X \sim \text{Bin}(n = 100, p = 0.5) \text{ e } Y \sim \text{Bin}(n = 100, p = 0.5).$$

$$M(X) = M(Y) = 50 \quad \text{Var}(X) = \text{Var}(Y) = 25.$$

Utilizzando l'approssimazione della variabile casuale Binomiale con la variabile casuale Normale

$$X \sim N(\mu = 50, \sigma^2 = 25) \text{ e } Y \sim N(\mu = 50, \sigma^2 = 25)$$

Il risultato del lancio dei dadi non dipende da quello delle monete, X e Y sono stocasticamente indipendenti, quindi

$$P\{(X \geq 50) \cap (Y \geq 39)\} =$$



Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.

Esercizio 5.

2. Ipotizzando di lanciare 100 monete e 100 dadi, si calcoli la probabilità di ottenere almeno 50 teste e almeno 39 numeri pari.

Siano X la variabile casuale 'n. di teste nel lancio delle 100 monete' e

Y la variabile casuale 'n. di facce con numero pari nel lancio dei 100 dadi'.

$$X \sim \text{Bin}(n = 100, p = 0.5) \text{ e } Y \sim \text{Bin}(n = 100, p = 0.5).$$

$$M(X) = M(Y) = 50 \quad \text{Var}(X) = \text{Var}(Y) = 25.$$

Utilizzando l'approssimazione della variabile casuale Binomiale con la variabile casuale Normale

$$X \sim N(\mu = 50, \sigma^2 = 25) \text{ e } Y \sim N(\mu = 50, \sigma^2 = 25)$$

Il risultato del lancio dei dadi non dipende da quello delle monete, X e Y sono stocasticamente indipendenti, quindi

$$P\{(X \geq 50) \cap (Y \geq 39)\} = P(X \geq 50) \cdot P(Y \geq 39) = 0.5 \cdot 0.9861 = 0.4930.$$

essendo:

$$P(X > 50) =$$



Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.

Esercizio 5.

2. Ipotizzando di lanciare 100 monete e 100 dadi, si calcoli la probabilità di ottenere almeno 50 teste e almeno 39 numeri pari.

Siano X la variabile casuale 'n. di teste nel lancio delle 100 monete' e

Y la variabile casuale 'n. di facce con numero pari nel lancio dei 100 dadi'.

$$X \sim \text{Bin}(n = 100, p = 0.5) \text{ e } Y \sim \text{Bin}(n = 100, p = 0.5).$$

$$M(X) = M(Y) = 50 \quad \text{Var}(X) = \text{Var}(Y) = 25.$$

Utilizzando l'approssimazione della variabile casuale Binomiale con la variabile casuale Normale

$$X \sim N(\mu = 50, \sigma^2 = 25) \text{ e } Y \sim N(\mu = 50, \sigma^2 = 25)$$

Il risultato del lancio dei dadi non dipende da quello delle monete, X e Y sono stocasticamente indipendenti, quindi

$$P\{(X \geq 50) \cap (Y \geq 39)\} = P(X \geq 50) \cdot P(Y \geq 39) = 0.5 \cdot 0.9861 = 0.4930.$$

essendo:

$$P(X > 50) = P\left(\frac{X - \mu}{\sigma} > \frac{50 - \mu}{\sigma}\right) = P\left(Z > \frac{50 - 50}{5}\right) = P(Z > 0) = 0.5$$

$$P(Y > 39) =$$



Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.

Esercizio 5.

2. Ipotizzando di lanciare 100 monete e 100 dadi, si calcoli la probabilità di ottenere almeno 50 teste e almeno 39 numeri pari.

Siano X la variabile casuale 'n. di teste nel lancio delle 100 monete' e

Y la variabile casuale 'n. di facce con numero pari nel lancio dei 100 dadi'.

$$X \sim \text{Bin}(n = 100, p = 0.5) \text{ e } Y \sim \text{Bin}(n = 100, p = 0.5).$$

$$M(X) = M(Y) = 50 \quad \text{Var}(X) = \text{Var}(Y) = 25.$$

Utilizzando l'approssimazione della variabile casuale Binomiale con la variabile casuale Normale

$$X \sim N(\mu = 50, \sigma^2 = 25) \text{ e } Y \sim N(\mu = 50, \sigma^2 = 25)$$

Il risultato del lancio dei dadi non dipende da quello delle monete, X e Y sono stocasticamente indipendenti, quindi

$$P\{(X \geq 50) \cap (Y \geq 39)\} = P(X \geq 50) \cdot P(Y \geq 39) = 0.5 \cdot 0.9861 = 0.4930.$$

essendo:

$$P(X > 50) = P\left(\frac{X - \mu}{\sigma} > \frac{50 - \mu}{\sigma}\right) = P\left(Z > \frac{50 - 50}{5}\right) = P(Z > 0) = 0.5$$

$$\begin{aligned} P(Y > 39) &= P\left(\frac{Y - \mu}{\sigma} > \frac{39 - \mu}{\sigma}\right) = P\left(Z > \frac{39 - 50}{5}\right) = \\ &= P(Z > -2.2) = P(Z \leq 2.2) = 0.9861. \end{aligned}$$

Tema T273

Prova scritta di STATISTICA del 07.02.2013

Prova scritta di
STATISTICA del
07.02.2013

Giuseppe Boari
Gabriele Cantaluppi



UNIVERSITÀ
CATTOLICA
del Sacro Cuore

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.

Esercizio 5.

Giuseppe Boari
Gabriele Cantaluppi
Università Cattolica del Sacro Cuore, Milano



Su 50 famiglie si sono rilevate le variabili X 'nazionalità capofamiglia' (I =italiana, S =straniera) e Y 'n. figli'.

$X \setminus Y$	1	2	3
I	11	19	5
S	3	7	5

1. Si rappresenti graficamente il carattere Y e se ne calcolino moda e mediana.

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.

Esercizio 5.



Si considera la distribuzione marginale della variabile Y , ottenuta sommando le frequenze di colonna

y_j	$n_{\bullet j}$
1	14
2	26
3	10

Esercizio 1.

Esercizio 2.

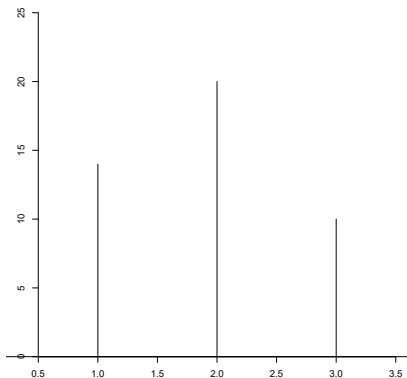
Esercizio 3.

Esercizio 4.

Esercizio 5.



Si tratta di un carattere quantitativo con modalità non raggruppate in classi che può essere rappresentato graficamente mediante un grafico a bastoncini.



Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.

Esercizio 5.



Si tratta di un carattere quantitativo con modalità non raggruppate in classi.

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.

Esercizio 5.



La moda corrisponde alla modalità alla quale è associata la massima frequenza) e risulta 2.

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.

Esercizio 5.



La mediana (percentile di ordine 0.50) in presenza di un numero pari di osservazioni corrisponde a uno degli elementi che occupano le posizioni $\frac{n}{2} = 25$ ed $\frac{n}{2} + 1 = 26$ nella serie dei dati ordinati. Dalla distribuzione delle frequenze cumulate

y_j	$n_{\bullet j}$	$N_{\bullet j}$
1	14	14
2	26	40
3	10	50

i due elementi coincidono e risultano $y_{0.5} = 2$

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.

Esercizio 5.



Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.

Esercizio 5.

2. Si confrontino con opportuni indici di posizione e di variabilità le variabili condizionate $Y|X = I$ e $Y|X = S$.



Si considera la distribuzione condizionata $Y|x_1$

y_j	n_{1j}
1	11
2	19
3	5

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.

Esercizio 5.



Si tratta di un carattere quantitativo con modalità non raggruppate in classi.

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.

Esercizio 5.



La media aritmetica

$$M(Y|x_1) = \frac{1}{n_{1\bullet}} \sum_{j=1}^k y_j n_{1j}$$

risulta 1.8286.

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.

Esercizio 5.



La media aritmetica

$$M(Y|x_1) = \frac{1}{n_{1\bullet}} \sum_{j=1}^k y_j n_{1j}$$

risulta 1.8286.

Si ottiene

y_j	n_{1j}	$y_j n_{1j}$
1	11	11
2	19	38
3	5	15
35		64

da cui segue

$$M(Y|x_1) = \frac{1}{n_{1\bullet}} \sum_{j=1}^k y_j n_{1j} = \frac{64}{35} = 1.8286.$$

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.

Esercizio 5.



La varianza

$$\text{Var}(Y|x_1) = M(Y^2|x_1) - [M(Y|x_1)]^2$$

risulta 0.4278.

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.

Esercizio 5.



La varianza

$$\text{Var}(Y|x_1) = M(Y^2|x_1) - [M(Y|x_1)]^2$$

risulta 0.4278. Si ottiene

y_j	n_{1j}	$y_j n_{1j}$	$y_j^2 n_{1j}$
1	11	11	11
2	19	38	76
3	5	15	45
	35	64	132

da cui segue

$$\text{Var}(Y|x_1) = \frac{132}{35} - 1.8286^2 = 3.7714 - 3.3437 = 0.4277.$$

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.

Esercizio 5.



Per effettuare confronti in termini di variabilità tra caratteri quantitativi occorre utilizzare una misura relativa, ad esempio il coefficiente di variazione

$$CV = \frac{\sigma}{\mu} = \frac{\sqrt{\text{Var}(Y|x_1)}}{M(Y|x_1)} = \frac{\sqrt{0.4277}}{1.8286} = \frac{0.654}{1.8286} = 0.3577.$$

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.

Esercizio 5.



Si considera la distribuzione condizionata $Y|x_2$

y_j	n_{2j}
1	3
2	7
3	5

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.

Esercizio 5.



Si tratta di un carattere quantitativo con modalità non raggruppate in classi.

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.

Esercizio 5.



La media aritmetica

$$M(Y|x_2) = \frac{1}{n_{2\bullet}} \sum_{j=1}^k y_j n_{2j}$$

risulta 2.1333.

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.

Esercizio 5.



La media aritmetica

$$M(Y|x_2) = \frac{1}{n_{2\bullet}} \sum_{j=1}^k y_j n_{2j}$$

risulta 2.1333.

Si ottiene

y_j	n_{2j}	$y_j n_{2j}$
1	3	3
2	7	14
3	5	15
	15	32

da cui segue

$$M(Y|x_2) = \frac{1}{n_{2\bullet}} \sum_{j=1}^k y_j n_{2j} = \frac{32}{15} = 2.1333.$$

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.

Esercizio 5.



La varianza

$$\text{Var}(Y|x_2) = M(Y^2|x_2) - [M(Y|x_2)]^2$$

risulta 0.5156.

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.

Esercizio 5.



Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.

Esercizio 5.

La varianza

$$\text{Var}(Y|x_2) = M(Y^2|x_2) - [M(Y|x_2)]^2$$

risulta 0.5156. Si ottiene

y_j	n_{2j}	$y_j n_{2j}$	$y_j^2 n_{2j}$
1	3	3	3
2	7	14	28
3	5	15	45
	15	32	76

da cui segue

$$\text{Var}(Y|x_2) = \frac{76}{15} - 2.1333^2 = 5.0667 - 4.5511 = 0.5156.$$



Per effettuare confronti in termini di variabilità tra caratteri quantitativi occorre utilizzare una misura relativa, ad esempio il coefficiente di variazione

$$CV = \frac{\sigma}{\mu} = \frac{\sqrt{\text{Var}(Y|x_2)}}{M(Y|x_2)} = \frac{\sqrt{0.5156}}{2.1333} = \frac{0.7181}{2.1333} = 0.3366.$$

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.

Esercizio 5.



$Y|X = I$:

Indice di posizione: $M(Y|X = I) = 1.8286$

Indice di variabilità: $CV(Y|X = I) = 0.3577$

$Y|X = S$:

Indice di posizione: $M(Y|X = S) = 2.1333$

Indice di variabilità: $CV(Y|X = S) = 0.3366$

$Y|X = S$ presenta un valore medio più elevato ma un minore livello di variabilità rispetto a $Y|X = I$.

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.

Esercizio 5.



3. Si calcoli un opportuno indice di variabilità di X .

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.

Esercizio 5.



Si considera la distribuzione marginale della variabile X , ottenuta sommando le frequenze di riga

x_i	$n_{i\bullet}$
I	35
S	15

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.

Esercizio 5.



In presenza di un carattere qualitativo la variabilità può essere misurata mediante gli indici di Gini e di Frosini normalizzati

$$G_N = \frac{1 - \sum_{i=1}^h f_{i\bullet}^2}{1 - \frac{1}{h}}$$

e

$$F_N = 1 - \sqrt{1 - G_N}.$$

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.

Esercizio 5.



In presenza di un carattere qualitativo la variabilità può essere misurata mediante gli indici di Gini e di Frosini normalizzati

$$G_N = \frac{1 - \sum_{i=1}^h f_{i\bullet}^2}{1 - \frac{1}{h}}$$

e

$$F_N = 1 - \sqrt{1 - G_N}.$$

Si ottiene:

x_i	$n_{i\bullet}$	$f_{i\bullet}$	$f_{i\bullet}^2$
I	35	0.7	0.49
S	15	0.3	0.09
	50	1	0.58

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.

Esercizio 5.



In presenza di un carattere qualitativo la variabilità può essere misurata mediante gli indici di Gini e di Frosini normalizzati

$$G_N = \frac{1 - \sum_{i=1}^h f_{i\bullet}^2}{1 - \frac{1}{h}}$$

e

$$F_N = 1 - \sqrt{1 - G_N}.$$

Si ottiene:

x_i	$n_{i\bullet}$	$f_{i\bullet}$	$f_{i\bullet}^2$
I	35	0.7	0.49
S	15	0.3	0.09
	50	1	0.58

Da cui segue:

$$G_N = \frac{1 - 0.58}{1 - \frac{1}{2}} = 0.84$$

e

$$F_N = 1 - \sqrt{1 - 0.84} = 0.6.$$

Sussiste, quindi, un livello medio di eterogeneità.



Dati i seguenti *NIBM* delle vendite di un'azienda tra il 2001 e il 2005:

t	<i>NIBM</i>
2001	
2002	0.94
2003	1.12
2004	1.06
2005	1.11

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.

Esercizio 5.



1. Si calcolino i numeri indice a base fissa (base 2001).

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.

Esercizio 5.



1. Si calcolino i numeri indice a base fissa (base 2001).

Si ricorda che ${}_{2001}I_{2001} = 1$ e che per valori di $t > 2001$ vale

$${}_{2001}I_t = {}_{2001}I_{2002} \cdot {}_{2002}I_{2003} \cdot \dots \cdot {}_{t-1}I_t.$$

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.

Esercizio 5.



Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.

Esercizio 5.

1. Si calcolino i numeri indice a base fissa (base 2001).

Si ricorda che ${}_{2001}I_{2001} = 1$ e che per valori di $t > 2001$ vale

$${}_{2001}I_t = {}_{2001}I_{2002} \cdot {}_{2002}I_{2003} \cdot \dots \cdot {}_{t-1}I_t.$$

Si ottiene quindi

t	$NIBM$	$NIBF (2001)$
2001		1.0000
2002	0.94	0.9400
2003	1.12	1.0528
2004	1.06	1.1160
2005	1.11	1.2387



2. Si calcoli il tasso medio di variazione tra il 2001 e il 2005 esprimendolo in valore percentuale.

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.

Esercizio 5.



2. Si calcoli il tasso medio di variazione tra il 2001 e il 2005 esprimendolo in valore percentuale.

Si ricorda che

$${}_{2001}I_{2005} = {}_{2001}I_{2002} \cdot {}_{2002}I_{2003} \cdot {}_{2003}I_{2004} \cdot {}_{2004}I_{2005}.$$

Si deve trovare il numero indice a base mobile medio.

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.

Esercizio 5.



2. Si calcoli il tasso medio di variazione tra il 2001 e il 2005 esprimendolo in valore percentuale.

Si ricorda che

$${}_{2001}I_{2005} = {}_{2001}I_{2002} \cdot {}_{2002}I_{2003} \cdot {}_{2003}I_{2004} \cdot {}_{2004}I_{2005}.$$

Si deve trovare il numero indice a base mobile medio. Applicando il criterio di scelta della media secondo Chisini si ottiene

$${}_{2001}I_{2005} = {}_{2001}I_{2002} \cdot {}_{2002}I_{2003} \cdot {}_{2003}I_{2004} \cdot {}_{2004}I_{2005} = \alpha \cdot \alpha \cdot \alpha \cdot \alpha = \alpha^4.$$

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.

Esercizio 5.



Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.

Esercizio 5.

2. Si calcoli il tasso medio di variazione tra il 2001 e il 2005 esprimendolo in valore percentuale.

Si ricorda che

$${}_{2001}I_{2005} = {}_{2001}I_{2002} \cdot {}_{2002}I_{2003} \cdot {}_{2003}I_{2004} \cdot {}_{2004}I_{2005}.$$

Si deve trovare il numero indice a base mobile medio. Applicando il criterio di scelta della media secondo Chisini si ottiene

$${}_{2001}I_{2005} = {}_{2001}I_{2002} \cdot {}_{2002}I_{2003} \cdot {}_{2003}I_{2004} \cdot {}_{2004}I_{2005} = \alpha \cdot \alpha \cdot \alpha \cdot \alpha = \alpha^4.$$

Quindi

$$\alpha = \sqrt[4]{{}_{2001}I_{2005}} = \sqrt[4]{{}_{2001}I_{2002} \cdot {}_{2002}I_{2003} \cdot {}_{2003}I_{2004} \cdot {}_{2004}I_{2005}} = 1.055.$$

Da cui segue una variazione relativa media pari a 0.055 , corrispondente a un incremento medio annuo tra il 2001 e il 2005 del 5.5%.



3. Per ottenere un tasso medio di variazione tra il 2001 e il 2006 del 7% che valore dovrebbe avere ${}_{2001}I_{2006}$?

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.

Esercizio 5.



Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.

Esercizio 5.

3. Per ottenere un tasso medio di variazione tra il 2001 e il 2006 del 7% che valore dovrebbe avere ${}_{2001}I_{2006}$?

$${}_{2001}I_{2006} = 1.07^5 = 1.4026$$



Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.

Esercizio 5.

Il seguente prospetto riporta i valori rilevati in corrispondenza a 5 unità statistiche con riferimento a due caratteri quantitativi X e Y

x_i	1	1	2	3	3
y_i	2	2	7	5	3

1. Si rappresentino graficamente le coppie di dati e la funzione di regressione.



La funzione di regressione associa a ogni valore x appartenente all'insieme di definizione della variabile X la media di Y condizionata a x , $M(Y|x)$

x	$M(Y x)$
1	2
2	7
3	$\frac{1}{2}(3 + 5) = 4$

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.

Esercizio 5.



Esercizio 1.

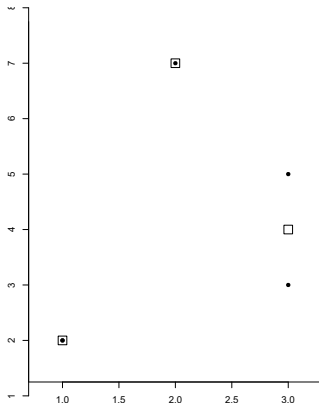
Esercizio 2.

Esercizio 3.

Esercizio 4.

Esercizio 5.

Le coppie di dati possono essere rappresentate mediante un grafico a dispersione.





2. Si calcolino, secondo il metodo dei minimi quadrati, i parametri e l'indice di adattamento dei seguenti modelli di regressione:

$$I) Y^* = a + bX \quad II) Y^* = c$$

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.

Esercizio 5.



Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.

Esercizio 5.

Si riportano i calcoli necessari¹ per ottenere le stime dei parametri e gli indici di adattamento del modello

$$Y = a + bX + E$$

x_i	y_i	$x_i y_i$	x_i^2	y_i^2
1	2	2	1	4
1	2	2	1	4
2	7	14	4	49
3	5	15	9	25
3	3	9	9	9
10	19	42	24	91
2	3.8	8.4	4.8	18.2
$M(X)$	$M(Y)$	$M(XY)$	$M(X^2)$	$M(Y^2)$

$$\hat{b} =$$

¹Si osserva che per ragioni di stampa alcuni calcoli potrebbero essere duplicati con riferimento ai diversi modelli oggetto di studio



Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.

Esercizio 5.

Si riportano i calcoli necessari¹ per ottenere le stime dei parametri e gli indici di adattamento del modello

$$Y = a + bX + E$$

x_i	y_i	$x_i y_i$	x_i^2	y_i^2
1	2	2	1	4
1	2	2	1	4
2	7	14	4	49
3	5	15	9	25
3	3	9	9	9
10	19	42	24	91
2	3.8	8.4	4.8	18.2
$M(X)$	$M(Y)$	$M(XY)$	$M(X^2)$	$M(Y^2)$

$$\begin{aligned}\hat{b} &= \frac{\text{Cov}(X, Y)}{\text{Var}(X)} = \frac{M(XY) - M(X)M(Y)}{M(X^2) - [M(X)^2]} = \\ &= \frac{8.4 - 2 \cdot 3.8}{4.8 - 2^2} = \frac{0.8}{0.8} = 1 \\ \hat{a} &= \end{aligned}$$

¹Si osserva che per ragioni di stampa alcuni calcoli potrebbero essere duplicati con riferimento ai diversi modelli oggetto di studio



Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.

Esercizio 5.

Si riportano i calcoli necessari¹ per ottenere le stime dei parametri e gli indici di adattamento del modello

$$Y = a + bX + E$$

x_i	y_i	$x_i y_i$	x_i^2	y_i^2
1	2	2	1	4
1	2	2	1	4
2	7	14	4	49
3	5	15	9	25
3	3	9	9	9
10	19	42	24	91
2	3.8	8.4	4.8	18.2
$M(X)$	$M(Y)$	$M(XY)$	$M(X^2)$	$M(Y^2)$

$$\hat{b} = \frac{\text{Cov}(X, Y)}{\text{Var}(X)} = \frac{M(XY) - M(X)M(Y)}{M(X^2) - [M(X)]^2} =$$

$$= \frac{8.4 - 2 \cdot 3.8}{4.8 - 2^2} = \frac{0.8}{0.8} = 1$$

$$\hat{a} = M(Y) - \hat{b}M(X) = 1.8$$

$$\rho^2 =$$

¹Si osserva che per ragioni di stampa alcuni calcoli potrebbero essere duplicati con riferimento ai diversi modelli oggetto di studio



Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.

Esercizio 5.

Si riportano i calcoli necessari¹ per ottenere le stime dei parametri e gli indici di adattamento del modello

$$Y = a + bX + E$$

x_i	y_i	$x_i y_i$	x_i^2	y_i^2
1	2	2	1	4
1	2	2	1	4
2	7	14	4	49
3	5	15	9	25
3	3	9	9	9
10	19	42	24	91
2	3.8	8.4	4.8	18.2
$M(X)$	$M(Y)$	$M(XY)$	$M(X^2)$	$M(Y^2)$

$$\begin{aligned}\hat{b} &= \frac{\text{Cov}(X, Y)}{\text{Var}(X)} = \frac{M(XY) - M(X)M(Y)}{M(X^2) - [M(X)]^2} = \\ &= \frac{8.4 - 2 \cdot 3.8}{4.8 - 2^2} = \frac{0.8}{0.8} = 1 \\ \hat{a} &= M(Y) - \hat{b}M(X) = 1.8 \\ \rho^2 &= \frac{[\text{Cov}(X, Y)]^2}{\text{Var}(X)\text{Var}(Y)} = \frac{0.8^2}{0.8 \cdot 3.76} = 0.2128\end{aligned}$$

¹Si osserva che per ragioni di stampa alcuni calcoli potrebbero essere duplicati con riferimento ai diversi modelli oggetto di studio



Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.

Esercizio 5.

Si riportano i calcoli necessari² per ottenere le stime dei parametri e gli indici di adattamento del modello

$$Y = c + E$$

x_i	y_i	y_i^2
1	2	4
1	2	4
2	7	49
3	5	25
3	3	9
19		91
3.8		18.2
$M(Y)$		$M(Y^2)$

$$\hat{c} =$$

²Si osserva che per ragioni di stampa alcuni calcoli potrebbero essere duplicati con riferimento ai diversi modelli oggetto di studio



Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.

Esercizio 5.

Si riportano i calcoli necessari² per ottenere le stime dei parametri e gli indici di adattamento del modello

$$Y = c + E$$

x_i	y_i	y_i^2
1	2	4
1	2	4
2	7	49
3	5	25
3	3	9
19		91
3.8		18.2
$M(Y)$		$M(Y^2)$

$$\hat{c} = M(Y) = 3.8$$

$$R^2 =$$

²Si osserva che per ragioni di stampa alcuni calcoli potrebbero essere duplicati con riferimento ai diversi modelli oggetto di studio



Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.

Esercizio 5.

Si riportano i calcoli necessari² per ottenere le stime dei parametri e gli indici di adattamento del modello

$$Y = c + E$$

x_i	y_i	y_i^2
1	2	4
1	2	4
2	7	49
3	5	25
3	3	9
19		91
3.8		18.2
$M(Y)$		$M(Y^2)$

$$\hat{c} = M(Y) = 3.8$$

$$R^2 = 0$$

²Si osserva che per ragioni di stampa alcuni calcoli potrebbero essere duplicati con riferimento ai diversi modelli oggetto di studio



Il modello migliore è $Y^* = 1.8 + X$.

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.

Esercizio 5.



Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.

Esercizio 5.

3. Utilizzando il modello ritenuto migliore si determini y quando $x = 5$.



Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.

Esercizio 5.

3. Utilizzando il modello ritenuto migliore si determini y quando $x = 5$.

$$Y^*(5) = 1.8 + 5 = 6.8$$



4. Si dica, giustificando teoricamente la risposta, se è verificata la relazione $\rho^2 = \eta_{Y|X}^2$.

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.

Esercizio 5.



4. Si dica, giustificando teoricamente la risposta, se è verificata la relazione $\rho^2 = \eta_{Y|X}^2$.
La relazione non è verificata in quanto le medie condizionate non sono allineate.

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.

Esercizio 5.

Esercizio 4

Prova scritta di
STATISTICA del
07.02.2013

Giuseppe Boari
Gabriele Cantaluppi



UNIVERSITÀ
CATTOLICA
del Sacro Cuore

Si completino le seguenti tabelle, giustificando opportunamente la risposta, in modo tale che :

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.

Esercizio 5.



Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.

Esercizio 5.

	a) $\chi^2 = 62$	e	$M(X) = 2$	
	$y_1 = 2$	$y_2 = 4$	$y_3 = 6$	
$x_1 = 1$	7	0	0	7
$x_2 = 2$	0	0	17	17
$x_3 = 3$	0	7	0	7
				$n = 31$

Se $\chi^2 = 62$ allora $\chi^2_N = 1$ sussiste dipendenza funzionale (biettiva) X deve essere simmetrica affinché $M(X) = 2$



Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.

Esercizio 5.

	b) $\chi^2 = 0$	e	$M(Y) = 4$	
	$y_1 = 2$	$y_2 = 4$	$y_3 = 6$	
$x_1 = 1$	7	n_{12}	7	
$x_2 = 2$	7c	$n_{12}c$	7c	
$x_3 = 3$	7d	$n_{12}d$	7d	
				$n =$

Se $\chi^2 = 0$ allora sussiste indipendenza stocastica

→ distribuzioni condizionate somigianti

n_{12} , c , d numeri interi positivi

Y deve essere simmetrica affinché $M(Y) = 2$

Esercizio 5

Prova scritta di
STATISTICA del
07.02.2013

Giuseppe Boari
Gabriele Cantaluppi



UNIVERSITÀ
CATTOLICA
del Sacro Cuore

Siano A , B , C e D quattro eventi che costituiscono una partizione dello spazio campionario, tali che $P(A) = P(B) = 0.2$ e $P(D) = 2 \cdot P(C)$.

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.

Esercizio 5.



1. Si calcolino $P(A|B)$, $P(B \cup D)$ e $P(\overline{A \cup B})$.

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.

Esercizio 5.



1. Si calcolino $P(A|B)$, $P(B \cup D)$ e $P(\overline{A \cup B})$.

A, B, C, D costituiscono una partizione di Ω se $A \cup B \cup C \cup D = \Omega$ e A, B, C e D non hanno elementi in comune.

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.

Esercizio 5.



1. Si calcolino $P(A|B)$, $P(B \cup D)$ e $P(\overline{A \cup B})$.

A, B, C, D costituiscono una partizione di Ω se $A \cup B \cup C \cup D = \Omega$ e A, B, C e D non hanno elementi in comune.

$$1 = P(A) + P(B) + P(C) + P(D) = 0.2 + 0.2 + P(C) + 2 \cdot P(C) \rightarrow P(C) = 0.2, P(D) = 0.4$$

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.

Esercizio 5.



1. Si calcolino $P(A|B)$, $P(B \cup D)$ e $P(\overline{A \cup B})$.

A, B, C, D costituiscono una partizione di Ω se $A \cup B \cup C \cup D = \Omega$ e A, B, C e D non hanno elementi in comune.

$$1 = P(A) + P(B) + P(C) + P(D) = 0.2 + 0.2 + P(C) + 2 \cdot P(C) \rightarrow P(C) = 0.2, P(D) = 0.4$$

$$P(A|B) = 0 \text{ in quanto } A \cap B = \emptyset,$$

$$P(B \cup D) = P(B) + P(D) = 0.2 + 0.4 = 0.6$$

$$P(\overline{A \cup B}) = 1 - P(A \cup B) = 1 - \{P(A) + P(B)\} = 1 - 0.2 - 0.2 = 0.4 = P(C \cup D)$$

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.

Esercizio 5.



2. Si dica se A e C possono ritenersi indipendenti.

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.

Esercizio 5.



2. Si dica se A e C possono ritenersi indipendenti.

A e C non possono ritenersi indipendenti in quanto $A \cap C = \emptyset$, quindi

$$P(A \cap C) = 0 \neq P(A) \cdot P(C)$$

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.

Esercizio 5.



Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.

Esercizio 5.

3. Si calcoli la probabilità di ottenere meno di 6 successi in 80 prove indipendenti con probabilità di successo $p = P(B)$.



3. Si calcoli la probabilità di ottenere meno di 6 successi in 80 prove indipendenti con probabilità di successo $p = P(B)$.

Indicata con X la variabile casuale 'numero di successi' nelle 80 prove indipendenti risulta

$$X \sim$$

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.

Esercizio 5.



3. Si calcoli la probabilità di ottenere meno di 6 successi in 80 prove indipendenti con probabilità di successo $p = P(B)$.

Indicata con X la variabile casuale 'numero di successi' nelle 80 prove indipendenti risulta

$$X \sim \text{Bin}(n = 80, p = 0.2).$$

$$M(X) = 16; \text{Var}(X) = 12.8;$$

utilizzando l'approssimazione della variabile casuale Binomiale con la variabile casuale Normale

$$X \sim$$

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.

Esercizio 5.



3. Si calcoli la probabilità di ottenere meno di 6 successi in 80 prove indipendenti con probabilità di successo $p = P(B)$.

Indicata con X la variabile casuale 'numero di successi' nelle 80 prove indipendenti risulta

$$X \sim \text{Bin}(n = 80, p = 0.2).$$

$$M(X) = 16; \text{Var}(X) = 12.8;$$

utilizzando l'approssimazione della variabile casuale Binomiale con la variabile casuale Normale

$$X \sim N(\mu = 16, \sigma^2 = 12.8)$$

si ottiene

$$\begin{aligned} P(X < 6) &= P\left(\frac{X - \mu}{\sigma} < \frac{6 - \mu}{\sigma}\right) = P\left(Z < \frac{6 - 16}{3.5777}\right) = \\ &= P(Z < -2.80) = 1 - P(Z \leq 2.80) = 1 - 0.9974 = 0.0026. \end{aligned}$$

Tema T274-1

Prova scritta di STATISTICA del 06.06.2013

Prova scritta di
STATISTICA del
06.06.2013

Giuseppe Boari
Gabriele Cantaluppi



UNIVERSITÀ
CATTOLICA
del Sacro Cuore

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.

Esercizio 5.

Giuseppe Boari
Gabriele Cantaluppi
Università Cattolica del Sacro Cuore, Milano



Su 50 famiglie si sono rilevate le variabili X 'nazionalità capofamiglia' (I =italiana, S =straniera) e Y 'n. figli'.

$X \ Y$	1	2	3
I	11	19	5
S	3	7	5

1. Si rappresenti graficamente la funzione di ripartizione del carattere Y e se ne calcolino media e coefficiente di variazione.

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.

Esercizio 5.



Si considera la distribuzione marginale della variabile Y , ottenuta sommando le frequenze di colonna.

y_j	$n_{\bullet j}$	$N_{\bullet j}$	$F_{\bullet j}$
1	14	14	0.28
2	26	40	0.8
3	10	50	1

Esercizio 1.

Esercizio 2.

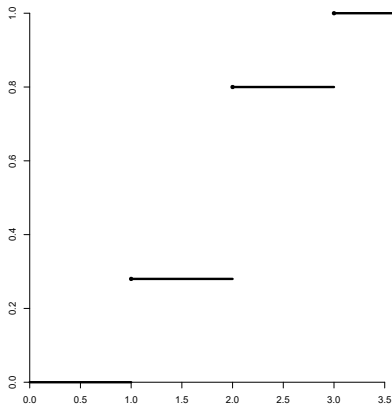
Esercizio 3.

Esercizio 4.

Esercizio 5.



Si tratta di un carattere quantitativo con modalità non raggruppate in classi. La funzione di ripartizione ha l'aspetto di una funzione a gradini.



Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.

Esercizio 5.



Si tratta di un carattere quantitativo con modalità non raggruppate in classi.

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.

Esercizio 5.



La media aritmetica

$$M(Y) = \frac{1}{n} \sum_{j=1}^k y_j n_{\bullet j}$$

risulta 1.92.

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.

Esercizio 5.



La media aritmetica

$$M(Y) = \frac{1}{n} \sum_{j=1}^k y_j n_{\bullet j}$$

risulta 1.92.

Si ottiene

y_j	$n_{\bullet j}$	$y_j n_{\bullet j}$
1	14	14
2	26	52
3	10	30
	50	96

da cui segue

$$M(Y) = \frac{1}{n} \sum_{i=1}^k y_j n_{\bullet j} = \frac{96}{50} = 1.92.$$

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.

Esercizio 5.



La varianza

$$\text{Var}(Y) = M[(Y - \mu_Y)^2] = M(Y^2) - \mu_Y^2 = \frac{1}{n} \sum_{j=1}^k y_j^2 n_{\bullet j} - \mu_Y^2$$

risulta 0.4736.

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.

Esercizio 5.



La varianza

$$\text{Var}(Y) = M[(Y - \mu_Y)^2] = M(Y^2) - \mu_Y^2 = \frac{1}{n} \sum_{j=1}^k y_j^2 n_{\bullet j} - \mu_Y^2$$

risulta 0.4736.

Si ottiene

y_j	$n_{\bullet j}$	$y_j n_{\bullet j}$	$y_j^2 n_{\bullet j}$
1	14	14	14
2	26	52	104
3	10	30	90
	50	96	208

da cui segue

$$\begin{aligned} \text{Var}(Y) = \sigma_Y^2 &= M(Y^2) - \mu_Y^2 = \\ &= \frac{208}{50} - 1.92^2 = 4.16 - 3.6864 = 0.4736. \end{aligned}$$

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.

Esercizio 5.



Per effettuare confronti in termini di variabilità tra caratteri quantitativi occorre utilizzare una misura relativa, ad esempio il coefficiente di variazione

$$CV = \frac{\sigma}{\mu} = \frac{\sqrt{\text{Var}(Y)}}{M(Y)} = \frac{\sqrt{0.4736}}{1.92} = \frac{0.6882}{1.92} = 0.3584.$$

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.

Esercizio 5.



2. Si calcoli un opportuno indice di connessione tra X e Y commentando il risultato ottenuto.

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.

Esercizio 5.



Dalla distribuzione delle frequenze congiunte n_{ij} :

	1	2	3	$n_{i\bullet}$
I	11	19	5	35
S	3	7	5	15
$n_{\bullet j}$	14	26	10	50

è possibile ottenere le frequenze teoriche nella situazione di indipendenza stocastica

$$\hat{n}_{ij} = \frac{n_{i\bullet} n_{\bullet j}}{n}:$$

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.

Esercizio 5.



Dalla distribuzione delle frequenze congiunte n_{ij} :

	1	2	3	$n_{i\bullet}$
I	11	19	5	35
S	3	7	5	15
$n_{\bullet j}$	14	26	10	50

è possibile ottenere le frequenze teoriche nella situazione di indipendenza stocastica

$$\hat{n}_{ij} = \frac{n_{i\bullet} \cdot n_{\bullet j}}{n}:$$

	1	2	3	$n_{i\bullet}$
I	9.8	18.2	7	35
S	4.2	7.8	3	15
$n_{\bullet j}$	14.0	26.0	10	50

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.

Esercizio 5.



le contingenze assolute $n_{ij} - \hat{n}_{ij}$:

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.

Esercizio 5.



le contingenze assolute $n_{ij} - \hat{n}_{ij}$:

	1	2	3
I	1.2	0.8	-2
S	-1.2	-0.8	2

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.

Esercizio 5.



i rapporti di contingenza $\frac{n_{ij}}{\hat{n}_{ij}}$:

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.

Esercizio 5.



i rapporti di contingenza $\frac{n_{ij}}{\hat{n}_{ij}}$:

	1	2	3
I	1.1224	1.0440	0.7143
S	0.7143	0.8974	1.6667

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.

Esercizio 5.



e le contingenze relative, espresse in termini percentuali, $100 \left(\frac{n_{ij} - \hat{n}_{ij}}{\hat{n}_{ij}} \right) \% = 100 \left(\frac{n_{ij}}{\hat{n}_{ij}} - 1 \right) \%$:

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.

Esercizio 5.



e le contingenze relative, espresse in termini percentuali, $100 \left(\frac{n_{ij} - \hat{n}_{ij}}{\hat{n}_{ij}} \right) \% = 100 \left(\frac{n_{ij}}{\hat{n}_{ij}} - 1 \right) \%$:

	1	2	3
I	12.24	4.40	-28.57
S	-28.57	-10.26	66.67

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.

Esercizio 5.



Risultati intermedi $\frac{(n_{ij} - \hat{n}_{ij})^2}{\hat{n}_{ij}}$ per il calcolo del χ^2

	1	2	3
I	0.1469	0.0352	0.5714
S	0.3429	0.0821	1.3333

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.

Esercizio 5.



Risultati intermedi $\frac{n_{ij}^2}{n_{i\bullet} \cdot n_{\bullet j}}$ per il calcolo del χ^2 (formula operativa)

	1	2	3
I	0.2469	0.3967	0.0714
S	0.0429	0.1256	0.1667

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.

Esercizio 5.



Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.

Esercizio 5.

$$\chi^2 = \sum_{i=1}^h \sum_{j=1}^k \frac{(n_{ij} - \hat{n}_{ij})^2}{\hat{n}_{ij}} = 2.5118,$$

$$\chi_N^2 = \frac{\chi^2}{n \cdot \min(h-1, k-1)} = 0.0502$$

$$\chi_N = \sqrt{\chi_N^2} = 0.2241.$$

Sussiste, quindi, un livello di connessione molto basso.



3. Si calcoli la mediana di X .

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.

Esercizio 5.



Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.

Esercizio 5.

3. Si calcoli la mediana di X .

Impossibile calcolare la mediana in quanto il carattere X è qualitativo sconnesso.



Dati i seguenti *NIBF* (base 2001) delle vendite di un'azienda tra il 2001 e il 2005:

t	$NIBF$
2001	1.00
2002	0.94
2003	1.12
2004	1.06
2005	1.11

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.

Esercizio 5.



1. Si calcolino i numeri indice a base mobile.

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.

Esercizio 5.



1. Si calcolino i numeri indice a base mobile.

Si ricorda che ${}_{t-1}I_t = \frac{x_t}{x_{t-1}} = \frac{{}_{2001}I_t}{{}_{2001}I_{t-1}}$.

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.

Esercizio 5.



1. Si calcolino i numeri indice a base mobile.

Si ricorda che ${}_{t-1}I_t = \frac{x_t}{x_{t-1}} = \frac{{}_{2001}I_t}{{}_{2001}I_{t-1}}$. Si ottiene quindi

<i>Anno</i>	<i>NIBF</i>	<i>NIBM</i>
2001	1.00	
2002	0.94	0.9400
2003	1.12	1.1915
2004	1.06	0.9464
2005	1.11	1.0472

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.

Esercizio 5.



2. Si calcoli il tasso medio di variazione tra il 2003 e il 2005 esprimendolo in valore percentuale.

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.

Esercizio 5.



2. Si calcoli il tasso medio di variazione tra il 2003 e il 2005 esprimendolo in valore percentuale.

Si ricorda che

$${}_{2003}I_{2005} = {}_{2003}I_{2004} \cdot {}_{2004}I_{2005}$$

Si deve trovare il numero indice a base mobile medio.

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.

Esercizio 5.



Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.

Esercizio 5.

2. Si calcoli il tasso medio di variazione tra il 2003 e il 2005 esprimendolo in valore percentuale.

Si ricorda che

$${}_{2003}I_{2005} = {}_{2003}I_{2004} \cdot {}_{2004}I_{2005}$$

Si deve trovare il numero indice a base mobile medio. Applicando il criterio di scelta della media secondo Chisini si ottiene

$${}_{2003}I_{2005} = {}_{2003}I_{2004} \cdot {}_{2004}I_{2005} = \alpha \cdot \alpha = \alpha^2.$$



Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.

Esercizio 5.

2. Si calcoli il tasso medio di variazione tra il 2003 e il 2005 esprimendolo in valore percentuale.

Si ricorda che

$${}_{2003}I_{2005} = {}_{2003}I_{2004} \cdot {}_{2004}I_{2005}$$

Si deve trovare il numero indice a base mobile medio. Applicando il criterio di scelta della media secondo Chisini si ottiene

$${}_{2003}I_{2005} = {}_{2003}I_{2004} \cdot {}_{2004}I_{2005} = \alpha \cdot \alpha = \alpha^2.$$

Quindi

$$\alpha = \sqrt[2]{{}_{2003}I_{2005}} = \sqrt[2]{{}_{2003}I_{2004} \cdot {}_{2004}I_{2005}} = 0.9955.$$

Da cui segue una variazione relativa media pari a -0.0045 , corrispondente a un decremento medio annuo tra il 2003 e il 2005 del 0.45%.



3. Sapendo che tra il 2005 e il 2006 vi è stato un calo delle vendite del 10% si determinino $I_{2005/2006}$ e il valore delle vendite nel 2006.

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.

Esercizio 5.



Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.

Esercizio 5.

3. Sapendo che tra il 2005 e il 2006 vi è stato un calo delle vendite del 10% si determinino ${}_{2005}I_{2006}$ e il valore delle vendite nel 2006.

$${}_{2005}I_{2006} = \frac{x_{2006}}{x_{2005}} = 0.90.$$

Non è possibile determinare il valore x_{2006} in quanto non si conosce x_{2005} .



Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.

Esercizio 5.

Il seguente prospetto riporta i valori rilevati in corrispondenza a 5 unità statistiche con riferimento a due caratteri quantitativi X e Y

x_i	2.7183	2.7183	7.3891	20.0855	20.0855
y_i	2	2	7	5	3

1. Si rappresentino graficamente le coppie di dati e la funzione di regressione.



La funzione di regressione associa a ogni valore x appartenente all'insieme di definizione della variabile X la media di Y condizionata a x , $M(Y|x)$

x	$M(Y x)$
2.7183	2
7.3891	7
20.0855	$\frac{1}{2}(3 + 5) = 4$

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.

Esercizio 5.



Esercizio 1.

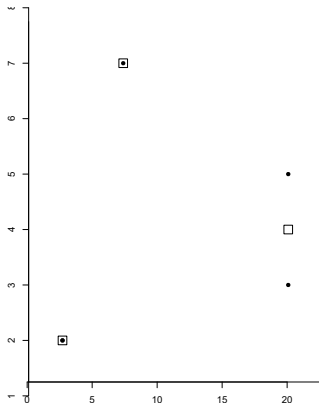
Esercizio 2.

Esercizio 3.

Esercizio 4.

Esercizio 5.

Le coppie di dati possono essere rappresentate mediante un grafico a dispersione.





2. Si calcolino, secondo il metodo dei minimi quadrati, i parametri e l'indice di adattamento dei seguenti modelli di regressione:

$$I) Y^* = a + b \cdot \ln X \qquad II) Y^* = c + dX$$

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.

Esercizio 5.



Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.

Esercizio 5.

Si riportano i calcoli necessari¹ per ottenere le stime dei parametri e gli indici di adattamento del modello

$$Y = a + b \ln(X) + E$$

$\ln(x)_i$	y_i	$\ln(x)_i y_i$	$\ln(x)_i^2$	y_i^2
1	2	2	1	4
1	2	2	1	4
2	7	14	4	49
3	5	15	9	25
3	3	9	9	9
10	19	42	24	91
2	3.8	8.4	4.8	18.2
$M(\ln(X))$	$M(Y)$	$M(\ln(X)Y)$	$M(\ln(X)^2)$	$M(Y^2)$

$$\hat{b} =$$

¹Si osserva che per ragioni di stampa alcuni calcoli potrebbero essere duplicati con riferimento ai diversi modelli oggetto di studio



Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.

Esercizio 5.

Si riportano i calcoli necessari¹ per ottenere le stime dei parametri e gli indici di adattamento del modello

$$Y = a + b \ln(X) + E$$

$\ln(x)_i$	y_i	$\ln(x)_i y_i$	$\ln(x)_i^2$	y_i^2
1	2	2	1	4
1	2	2	1	4
2	7	14	4	49
3	5	15	9	25
3	3	9	9	9
10	19	42	24	91
2	3.8	8.4	4.8	18.2
$M(\ln(X))$	$M(Y)$	$M(\ln(X)Y)$	$M(\ln(X)^2)$	$M(Y^2)$

$$\begin{aligned} \hat{b} &= \frac{\text{Cov}(\ln(X), Y)}{\text{Var}(\ln(X))} = \frac{M(\ln(X)Y) - M(\ln(X))M(Y)}{M(\ln(X)^2) - [M(\ln(X))]^2} = \\ &= \frac{8.4 - 2 \cdot 3.8}{4.8 - 2^2} = \frac{0.8}{0.8} = 1 \end{aligned}$$

$$\hat{a} =$$

¹Si osserva che per ragioni di stampa alcuni calcoli potrebbero essere duplicati con riferimento ai diversi modelli oggetto di studio



Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.

Esercizio 5.

Si riportano i calcoli necessari¹ per ottenere le stime dei parametri e gli indici di adattamento del modello

$$Y = a + b \ln(X) + E$$

$\ln(x)_i$	y_i	$\ln(x)_i y_i$	$\ln(x)_i^2$	y_i^2
1	2	2	1	4
1	2	2	1	4
2	7	14	4	49
3	5	15	9	25
3	3	9	9	9
10	19	42	24	91
2	3.8	8.4	4.8	18.2
$M(\ln(X))$	$M(Y)$	$M(\ln(X)Y)$	$M(\ln(X)^2)$	$M(Y^2)$

$$\begin{aligned} \hat{b} &= \frac{\text{Cov}(\ln(X), Y)}{\text{Var}(\ln(X))} = \frac{M(\ln(X)Y) - M(\ln(X))M(Y)}{M(\ln(X)^2) - [M(\ln(X))]^2} = \\ &= \frac{8.4 - 2 \cdot 3.8}{4.8 - 2^2} = \frac{0.8}{0.8} = 1 \end{aligned}$$

$$\hat{a} = M(Y) - \hat{b} M(\ln(X)) = 1.8$$

$$\rho^2 =$$

¹Si osserva che per ragioni di stampa alcuni calcoli potrebbero essere duplicati con riferimento ai diversi modelli oggetto di studio



Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.

Esercizio 5.

Si riportano i calcoli necessari¹ per ottenere le stime dei parametri e gli indici di adattamento del modello

$$Y = a + b \ln(X) + E$$

$\ln(x)_i$	y_i	$\ln(x)_i y_i$	$\ln(x)_i^2$	y_i^2
1	2	2	1	4
1	2	2	1	4
2	7	14	4	49
3	5	15	9	25
3	3	9	9	9
10	19	42	24	91
2	3.8	8.4	4.8	18.2
$M(\ln(X))$	$M(Y)$	$M(\ln(X)Y)$	$M(\ln(X)^2)$	$M(Y^2)$

$$\begin{aligned}\hat{b} &= \frac{\text{Cov}(\ln(X), Y)}{\text{Var}(\ln(X))} = \frac{M(\ln(X)Y) - M(\ln(X))M(Y)}{M(\ln(X)^2) - [M(\ln(X))]^2} = \\ &= \frac{8.4 - 2 \cdot 3.8}{4.8 - 2^2} = \frac{0.8}{0.8} = 1\end{aligned}$$

$$\hat{a} = M(Y) - \hat{b} M(\ln(X)) = 1.8$$

$$\rho^2 = \frac{[\text{Cov}(\ln(X), Y)]^2}{\text{Var}(\ln(X)) \text{Var}(Y)} = \frac{0.8^2}{0.8 \cdot 3.76} = 0.2128$$

¹Si osserva che per ragioni di stampa alcuni calcoli potrebbero essere duplicati con riferimento ai diversi modelli oggetto di studio



Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.

Esercizio 5.

Si riportano i calcoli necessari² per ottenere le stime dei parametri e gli indici di adattamento del modello

$$Y = c + dX + E$$

x_i	y_i	$x_i y_i$	x_i^2	y_i^2
2.7183	2	5.4366	7.3892	4
2.7183	2	5.4366	7.3892	4
7.3891	7	51.7237	54.5988	49
20.0855	5	100.4275	403.4273	25
20.0855	3	60.2565	403.4273	9
52.9967	19	223.2809	876.2317	91
10.5993	3.8	44.6562	175.2463	18.2
$M(X)$	$M(Y)$	$M(XY)$	$M(X^2)$	$M(Y^2)$

$$\hat{d} =$$

²Si osserva che per ragioni di stampa alcuni calcoli potrebbero essere duplicati con riferimento ai diversi modelli oggetto di studio



Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.

Esercizio 5.

Si riportano i calcoli necessari² per ottenere le stime dei parametri e gli indici di adattamento del modello

$$Y = c + dX + E$$

x_i	y_i	$x_i y_i$	x_i^2	y_i^2
2.7183	2	5.4366	7.3892	4
2.7183	2	5.4366	7.3892	4
7.3891	7	51.7237	54.5988	49
20.0855	5	100.4275	403.4273	25
20.0855	3	60.2565	403.4273	9
52.9967	19	223.2809	876.2317	91
10.5993	3.8	44.6562	175.2463	18.2
$M(X)$	$M(Y)$	$M(XY)$	$M(X^2)$	$M(Y^2)$

$$\begin{aligned}\hat{d} &= \frac{\text{Cov}(X, Y)}{\text{Var}(X)} = \frac{M(XY) - M(X)M(Y)}{M(X^2) - [M(X)]^2} = \\ &= \frac{44.6562 - 10.5993 \cdot 3.8}{175.2463 - 10.5993^2} = \frac{4.3787}{62.9003} = 0.0696 \\ \hat{c} &= \end{aligned}$$

²Si osserva che per ragioni di stampa alcuni calcoli potrebbero essere duplicati con riferimento ai diversi modelli oggetto di studio



Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.

Esercizio 5.

Si riportano i calcoli necessari² per ottenere le stime dei parametri e gli indici di adattamento del modello

$$Y = c + dX + E$$

x_i	y_i	$x_i y_i$	x_i^2	y_i^2
2.7183	2	5.4366	7.3892	4
2.7183	2	5.4366	7.3892	4
7.3891	7	51.7237	54.5988	49
20.0855	5	100.4275	403.4273	25
20.0855	3	60.2565	403.4273	9
52.9967	19	223.2809	876.2317	91
10.5993	3.8	44.6562	175.2463	18.2
$M(X)$	$M(Y)$	$M(XY)$	$M(X^2)$	$M(Y^2)$

$$\begin{aligned}\hat{d} &= \frac{\text{Cov}(X, Y)}{\text{Var}(X)} = \frac{M(XY) - M(X)M(Y)}{M(X^2) - [M(X)]^2} = \\ &= \frac{44.6562 - 10.5993 \cdot 3.8}{175.2463 - 10.5993^2} = \frac{4.3787}{62.9003} = 0.0696\end{aligned}$$

$$\hat{c} = M(Y) - \hat{d}M(X) = 3.0621$$

$$\rho^2 =$$

²Si osserva che per ragioni di stampa alcuni calcoli potrebbero essere duplicati con riferimento ai diversi modelli oggetto di studio



Si riportano i calcoli necessari² per ottenere le stime dei parametri e gli indici di adattamento del modello

$$Y = c + dX + E$$

x_i	y_i	$x_i y_i$	x_i^2	y_i^2
2.7183	2	5.4366	7.3892	4
2.7183	2	5.4366	7.3892	4
7.3891	7	51.7237	54.5988	49
20.0855	5	100.4275	403.4273	25
20.0855	3	60.2565	403.4273	9
52.9967	19	223.2809	876.2317	91
10.5993	3.8	44.6562	175.2463	18.2
$M(X)$	$M(Y)$	$M(XY)$	$M(X^2)$	$M(Y^2)$

$$\begin{aligned}\hat{d} &= \frac{\text{Cov}(X, Y)}{\text{Var}(X)} = \frac{M(XY) - M(X)M(Y)}{M(X^2) - [M(X)]^2} = \\ &= \frac{44.6562 - 10.5993 \cdot 3.8}{175.2463 - 10.5993^2} = \frac{4.3787}{62.9003} = 0.0696\end{aligned}$$

$$\hat{c} = M(Y) - \hat{d}M(X) = 3.0621$$

$$\rho^2 = \frac{[\text{Cov}(X, Y)]^2}{\text{Var}(X) \text{Var}(Y)} = \frac{4.3787^2}{62.9003 \cdot 3.76} = 0.0811$$

²Si osserva che per ragioni di stampa alcuni calcoli potrebbero essere duplicati con riferimento ai diversi modelli oggetto di studio



Il modello migliore è $Y^* = 1.8 + \ln X$.

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.

Esercizio 5.

Esercizio 4



Si completi la seguente tabella, giustificando opportunamente la risposta, in modo tale che :

$$\text{a) } \chi_N^2 = \rho^2 = \eta_{Y|X}^2$$

	$y_1 = 2$	$y_2 = 4$	$y_3 = 6$	
$x_1 = 1$	7			
$x_2 = 2$				
$x_3 = 3$				
				$n = 120$

[Esercizio 1.](#)

[Esercizio 2.](#)

[Esercizio 3.](#)

[Esercizio 4.](#)

[Esercizio 5.](#)

È possibile realizzare (senza effettuare calcoli) la situazione $\chi_N^2 = \rho^2 = \eta_{Y|X}^2$ se:



Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.

Esercizio 5.

È possibile realizzare (senza effettuare calcoli) la situazione $\chi_N^2 = \rho^2 = \eta_{Y|X}^2$ se:

$$\chi_N^2 = \rho^2 = \eta_{Y|X}^2 = 0$$

situazione di indipendenza stocastica

	$y_1 = 2$	$y_2 = 4$	$y_3 = 6$	
$x_1 = 1$	7	8	9	
$x_2 = 2$	14	16	18	
$x_2 = 2$	14	16	18	
				$n = 120$

oppure



Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.

Esercizio 5.

È possibile realizzare (senza effettuare calcoli) la situazione $\chi_N^2 = \rho^2 = \eta_{Y|X}^2$ se:

$$\chi_N^2 = \rho^2 = \eta_{Y|X}^2 = 0$$

situazione di indipendenza stocastica

	$y_1 = 2$	$y_2 = 4$	$y_3 = 6$	
$x_1 = 1$	7	8	9	
$x_2 = 2$	14	16	18	
$x_2 = 2$	14	16	18	
				$n = 120$

oppure

$$\chi_N^2 = \rho^2 = \eta_{Y|X}^2 = 1$$

situazione di dipendenza funzionale lineare

	$y_1 = 2$	$y_2 = 4$	$y_3 = 6$	
$x_1 = 1$	7	0	0	
$x_2 = 2$	0	n₂₂	0	
$x_3 = 3$	0	0	120 - (7 + n₂₂)	
				$n = 120$



Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.

Esercizio 5.

Esercizio 5

Prova scritta di
STATISTICA del
06.06.2013

Giuseppe Boari
Gabriele Cantaluppi



UNIVERSITÀ
CATTOLICA
del Sacro Cuore

L'urna U_1 contiene 3 palline nere e 1 pallina rossa. L'urna U_2 contiene 4 palline nere e 5 palline rosse.

1. Si calcoli la probabilità che estraendo con reimmissione 5 palline dall'urna U_1 ce ne siano almeno 4 nere.

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.

Esercizio 5.

Esercizio 5



L'urna $U1$ contiene 3 palline nere e 1 pallina rossa. L'urna $U2$ contiene 4 palline nere e 5 palline rosse.

1. Si calcoli la probabilità che estraendo con reimmissione 5 palline dall'urna $U1$ ce ne siano almeno 4 nere.

Sia X la variabile casuale 'n. di palline nere nelle 5 estrazioni con reimmissione'.

$$X \sim$$

[Esercizio 1.](#)

[Esercizio 2.](#)

[Esercizio 3.](#)

[Esercizio 4.](#)

[Esercizio 5.](#)

Esercizio 5



L'urna $U1$ contiene 3 palline nere e 1 pallina rossa. L'urna $U2$ contiene 4 palline nere e 5 palline rosse.

1. Si calcoli la probabilità che estraendo con reimmissione 5 palline dall'urna $U1$ ce ne siano almeno 4 nere.

Sia X la variabile casuale 'n. di palline nere nelle 5 estrazioni con reimmissione'.

$$X \sim \text{Bin}(n = 5, p = 0.75).$$

$$\begin{aligned} P(X \geq 4) &= P(X = 4) + P(X = 5) = \\ &= \binom{5}{4} 0.75^4 0.25^1 + \binom{5}{5} 0.75^5 0.25^0 = \\ &= 0.3955 + 0.2373 = 0.6328. \end{aligned}$$

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.

Esercizio 5.



2. Si estragga una pallina dall'urna $U1$ e la si metta nell'urna $U2$. Si calcoli la probabilità che estraendo senza reimmissione 3 palline dall'urna $U2$ si ottengano 2 palline nere.

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.

Esercizio 5.



Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.

Esercizio 5.

2. Si estragga una pallina dall'urna $U1$ e la si metta nell'urna $U2$. Si calcoli la probabilità che estraendo senza reimmissione 3 palline dall'urna $U2$ si ottengano 2 palline nere.

Siano N l'evento 'estrazione di una pallina nera dall'urna $U1$ ', $X|N$ e $X|\bar{N}$ le variabile casuale 'n. di palline nere nelle 3 estrazioni senza reimmissione dall'urna $U2$ ' condizionate al verificarsi degli eventi N e $\bar{N}$.

$$\begin{aligned} P(2 \text{ nere in } 3 \text{ estraz. senza reimmissione}) &= \\ &= P(N) \cdot P\{(X|N) = 2\} + P(\bar{N}) \cdot P\{(X|\bar{N}) = 2\} = \\ &= \frac{3}{4} \cdot \frac{\binom{5}{2} \binom{5}{1}}{\binom{10}{3}} + \frac{1}{4} \cdot \frac{\binom{4}{2} \binom{6}{1}}{\binom{10}{3}} = \\ &= 0.75 \cdot 0.4167 + 0.25 \cdot 0.3 = 0.3875. \end{aligned}$$

Tema T274

Prova scritta di STATISTICA del 27.06.2013

Prova scritta di
STATISTICA del
27.06.2013

Giuseppe Boari
Gabriele Cantaluppi



UNIVERSITÀ
CATTOLICA
del Sacro Cuore

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.

Giuseppe Boari
Gabriele Cantaluppi
Università Cattolica del Sacro Cuore, Milano



In un'indagine su 1000 cittadini sono stati rilevati l'area di residenza X (N = Nord; C = Centro; S = Sud Isole) e la spesa media mensile per libri Y . Si riportano la distribuzione di frequenze relative di X e la densità di frequenza di Y .

x_i	f_i
N	0.60
C	0.21
S	0.19

y_i	d_i
$5 \rightarrow 10$	90
$10 \rightarrow 20$	21
$20 \rightarrow 40$	17

1. Si costruiscano le distribuzioni delle frequenze assolute di X e di Y .

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.



Moltiplicando le frequenze relative f_i per il numero totale di osservazioni n si ottengono le frequenze assolute n_i della variabile X .

x_i	f_i	n_i
N	0.6	600
C	0.21	210
S	0.19	190

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.



Moltiplicando le densità di frequenza, d_i , per le ampiezze di classe, a_i , si ottengono le frequenze assolute n_i della variabile Y ,

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.



Moltiplicando le densità di frequenza, d_i , per le ampiezze di classe, a_i , si ottengono le frequenze assolute n_i della variabile Y ,

y_i	d_i	a_i	n_i
5 $\rightarrow$ 10	90	10 $-$ 5 = 5	450
10 $\rightarrow$ 20	21	20 $-$ 10 = 10	210
20 $\rightarrow$ 40	17	40 $-$ 20 = 20	340

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.



2. Si rappresentino graficamente le due variabili e si riassumano le stesse con opportuni indici di posizione e di variabilità.

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.



Si considera la distribuzione della variabile X ,

x_i	n_i
N	600
C	210
S	190

Si tratta di un carattere qualitativo sconnesso che può essere rappresentato graficamente mediante un grafico a bastoncini o a canne d'organo.

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.

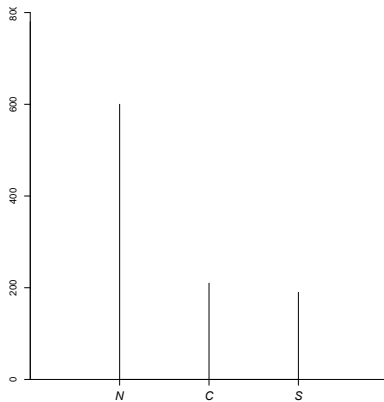


Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.





Si considera la distribuzione della variabile X ,

x_i	n_i
N	600
C	210
S	190

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.



Si tratta di un carattere qualitativo sconnesso.

L'unico indice di posizione che può essere definito è la moda (categoria alla quale è associata la massima frequenza) che risulta N .

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.

In presenza di un carattere qualitativo la variabilità può essere misurata mediante gli indici di Gini e di Frosini normalizzati

$$G_N = \frac{1 - \sum_{i=1}^k f_i^2}{1 - \frac{1}{k}}$$

e

$$F_N = 1 - \sqrt{1 - G_N}.$$



Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.



In presenza di un carattere qualitativo la variabilità può essere misurata mediante gli indici di Gini e di Frosini normalizzati

$$G_N = \frac{1 - \sum_{i=1}^k f_i^2}{1 - \frac{1}{k}}$$

e

$$F_N = 1 - \sqrt{1 - G_N}.$$

Si ottiene:

x_i	n_i	f_i	f_i^2
N	600	0.6	0.36
C	210	0.21	0.0441
S	190	0.19	0.0361
	1000	1	0.4402

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.



In presenza di un carattere qualitativo la variabilità può essere misurata mediante gli indici di Gini e di Frosini normalizzati

$$G_N = \frac{1 - \sum_{i=1}^k f_i^2}{1 - \frac{1}{k}}$$

e

$$F_N = 1 - \sqrt{1 - G_N}.$$

Si ottiene:

x_i	n_i	f_i	f_i^2
N	600	0.6	0.36
C	210	0.21	0.0441
S	190	0.19	0.0361
	1000	1	0.4402

Da cui segue:

$$G_N = \frac{1 - 0.4402}{1 - \frac{1}{3}} = 0.8397$$

e

$$F_N = 1 - \sqrt{1 - 0.8397} = 0.5996.$$

Sussiste, quindi, un livello medio di eterogeneità.



Si considera la distribuzione della variabile X ,

x_i	n_i
5 $\vdash$ 10	450
10 $\vdash$ 20	210
20 $\vdash$ 40	340

Esercizio 1.

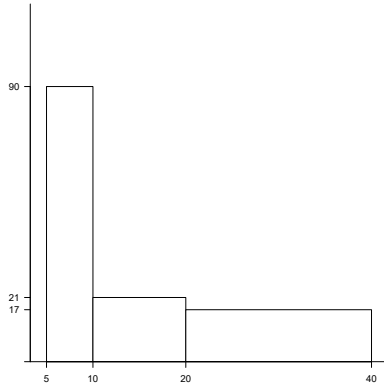
Esercizio 2.

Esercizio 3.

Esercizio 4.



Si tratta di un carattere quantitativo con modalità raggruppate in classi che può essere rappresentato graficamente mediante un istogramma.



Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.



Si considera la distribuzione della variabile Y ,

y_i	n_i
5 $\vdash$ 10	450
10 $\vdash$ 20	210
20 $\vdash$ 40	340

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.



Si tratta di un carattere quantitativo con modalità raggruppate in classi.

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.



La media aritmetica

$$M(Y) = \frac{1}{n} \sum_{i=1}^k y_i n_i$$

risulta 16.725.

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.



La media aritmetica

$$M(Y) = \frac{1}{n} \sum_{i=1}^k y_i n_i$$

risulta 16.725.

Si ottiene (considerando come modalità i valori centrali delle classi)

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.



La media aritmetica

$$M(Y) = \frac{1}{n} \sum_{i=1}^k y_i n_i$$

risulta 16.725.

Si ottiene (considerando come modalità i valori centrali delle classi)

y_i	n_i	$y_i n_i$
7.5	450	3375
15.0	210	3150
30.0	340	10200
	1000	16725

da cui segue

$$M(Y) = \frac{1}{n} \sum_{i=1}^k y_i n_i = \frac{16725}{1000} = 16.725.$$

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.



La varianza

$$\text{Var}(Y) = M[(Y - \mu_Y)^2] = M(Y^2) - \mu_Y^2 = \frac{1}{n} \sum_{i=1}^k y_i^2 n_i - \mu_Y^2$$

risulta 98.8369.

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.



La varianza

$$\text{Var}(Y) = M[(Y - \mu_Y)^2] = M(Y^2) - \mu_Y^2 = \frac{1}{n} \sum_{i=1}^k y_i^2 n_i - \mu_Y^2$$

risulta 98.8369.

Si ottiene (considerando come modalità i valori centrali delle classi)

y_i	n_i	$y_i n_i$	$y_i^2 n_i$
7.5	450	3375	25312.5
15.0	210	3150	47250.0
30.0	340	10200	306000.0
	1000	16725	378562.5

da cui segue

$$\begin{aligned} \text{Var}(Y) = \sigma_Y^2 &= M(Y^2) - \mu_Y^2 = \\ &= \frac{378562.5}{1000} - 16.725^2 = 378.5625 - 279.7256 = 98.8369. \end{aligned}$$

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.



Per effettuare confronti in termini di variabilità tra caratteri quantitativi occorre utilizzare una misura relativa, ad esempio il coefficiente di variazione

$$CV = \frac{\sigma}{\mu} = \frac{\sqrt{\text{Var}(Y)}}{M(Y)} = \frac{\sqrt{98.8369}}{16.725} = \frac{9.9417}{16.725} = 0.5944.$$

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.



3. Si calcoli la media geometrica e si individui su un opportuno grafico la mediana della variabile Y .

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.



Si tratta di un carattere quantitativo con modalità raggruppate in classi.

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.



La mediana (percentile di ordine 0.50) può essere ottenuta considerando l'inversa della Funzione di Ripartizione in corrispondenza di 0.50.

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.



La mediana (percentile di ordine 0.50) può essere ottenuta considerando l'inversa della Funzione di Ripartizione in corrispondenza di 0.50.

y_i	n_i	N_i	F_i
5 - 10	450	450	0.45
10 - 20	210	660	0.66
20 - 40	340	1000	1

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.

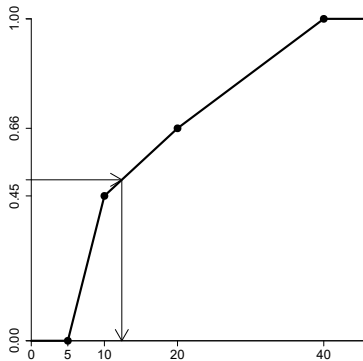


Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.



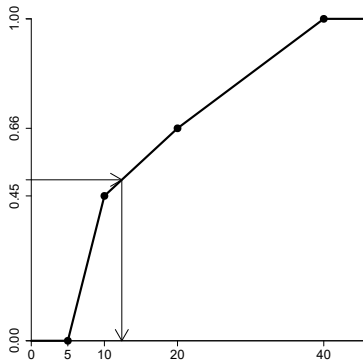


Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.



Si ottiene

$$y_{0.5} = 10 + \frac{0.5 - 0.45}{0.66 - 0.45} \cdot (20 - 10) = 12.381.$$



La media geometrica

$$\mu^{(0)}(Y) = \prod_{i=1}^k y_i^{f_i}$$

risulta 13.8988.

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.



La media geometrica

$$\mu^{(0)}(Y) = \prod_{i=1}^k y_i^{f_i}$$

risulta 13.8988. Si ottiene (considerando come modalità i valori centrali delle classi)

y_i	f_i	$y_i^{f_i}$
7.5	0.45	2.4762
15.0	0.21	1.7660
30.0	0.34	3.1785
		13.8988

da cui segue

$$\mu^{(0)} = \prod_{j=1}^k y_j^{f_j} = 13.8988.$$

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.

Esercizio 2

Prova scritta di
STATISTICA del
27.06.2013

Giuseppe Boari
Gabriele Cantaluppi



UNIVERSITÀ
CATTOLICA
del Sacro Cuore

Si consideri la serie NIBF (base 2006) relativa alla spesa delle Amministrazioni Pubbliche rilevata tra il 2005 e il 2009:

t	$NIBF$
2005	0.9705
2006	1.0000
2007	1.0164
2008	0.7607
2009	1.0951

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.



1. Si calcolino i numeri indice a base mobile.

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.



1. Si calcolino i numeri indice a base mobile.

Si ricorda che ${}_{t-1}I_t = \frac{x_t}{x_{t-1}} = \frac{{}_{2006}I_t}{{}_{2006}I_{t-1}}$.

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.



1. Si calcolino i numeri indice a base mobile.

Si ricorda che ${}_{t-1}I_t = \frac{x_t}{x_{t-1}} = \frac{{}_{2006}I_t}{{}_{2006}I_{t-1}}$. Si ottiene quindi

Anno	NIBF	NIBM
2005	0.9705	
2006	1.0000	1.0304
2007	1.0164	1.0164
2008	0.7607	0.7484
2009	1.0951	1.4396

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.



2. Si calcoli il tasso medio di variazione tra il 2006 e il 2009 esprimendolo in valore percentuale.

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.



2. Si calcoli il tasso medio di variazione tra il 2006 e il 2009 esprimendolo in valore percentuale.

Si ricorda che

$${}_{2006}I_{2009} = {}_{2006}I_{2007} \cdot {}_{2007}I_{2008} \cdot {}_{2008}I_{2009}$$

Si deve trovare il numero indice a base mobile medio.

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.



2. Si calcoli il tasso medio di variazione tra il 2006 e il 2009 esprimendolo in valore percentuale.

Si ricorda che

$${}_{2006}I_{2009} = {}_{2006}I_{2007} \cdot {}_{2007}I_{2008} \cdot {}_{2008}I_{2009}$$

Si deve trovare il numero indice a base mobile medio. Applicando il criterio di scelta della media secondo Chisini si ottiene

$${}_{2006}I_{2009} = {}_{2006}I_{2007} \cdot {}_{2007}I_{2008} \cdot {}_{2008}I_{2009} = \alpha \cdot \alpha \cdot \alpha = \alpha^3.$$

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.



Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.

2. Si calcoli il tasso medio di variazione tra il 2006 e il 2009 esprimendolo in valore percentuale.

Si ricorda che

$${}_{2006}I_{2009} = {}_{2006}I_{2007} \cdot {}_{2007}I_{2008} \cdot {}_{2008}I_{2009}$$

Si deve trovare il numero indice a base mobile medio. Applicando il criterio di scelta della media secondo Chisini si ottiene

$${}_{2006}I_{2009} = {}_{2006}I_{2007} \cdot {}_{2007}I_{2008} \cdot {}_{2008}I_{2009} = \alpha \cdot \alpha \cdot \alpha = \alpha^3.$$

Quindi

$$\alpha = \sqrt[3]{{}_{2006}I_{2009}} = \sqrt[3]{{}_{2006}I_{2007} \cdot {}_{2007}I_{2008} \cdot {}_{2008}I_{2009}} = 1.0307.$$

Da cui segue una variazione relativa media pari a 0.0307, corrispondente a un incremento medio annuo tra il 2006 e il 2009 del 3.07%.



3. Si calcoli il Numero Indice del 2006 con base l'anno 2009 interpretando il valore ottenuto.

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.



3. Si calcoli il Numero Indice del 2006 con base l'anno 2009 interpretando il valore ottenuto.

$${}_{2009}I_{2006} = \frac{X_{2006}}{X_{2009}} = \frac{{}_{2006}I_{2006}}{{}_{2006}I_{2009}} = \frac{1}{1.0951} = 0.9132$$

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.



Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.

3. Si calcoli il Numero Indice del 2006 con base l'anno 2009 interpretando il valore ottenuto.

$${}_{2009}I_{2006} = \frac{X_{2006}}{X_{2009}} = \frac{{}_{2006}I_{2006}}{{}_{2006}I_{2009}} = \frac{1}{1.0951} = 0.9132$$

La grandezza al tempo 2006 era inferiore del 8.68% rispetto alla grandezza al tempo 2009. Ovvero il valore della grandezza nel 2006 era il 91.32% di quello registrato nel 2009.



Il seguente prospetto riporta i valori rilevati in corrispondenza a 6 unità statistiche con riferimento a due caratteri quantitativi X e Y

x_i	5	5	6	6	7	7
y_i	102	98	62	58	102	98

1. Si rappresentino graficamente le coppie di dati e la funzione di regressione.

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.



La funzione di regressione associa a ogni valore x appartenente all'insieme di definizione della variabile X la media di Y condizionata a x , $M(Y|x)$

x	$M(Y x)$
5	$\frac{1}{2}(98 + 102) = 100$
6	$\frac{1}{2}(58 + 62) = 60$
7	$\frac{1}{2}(98 + 102) = 100$

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.



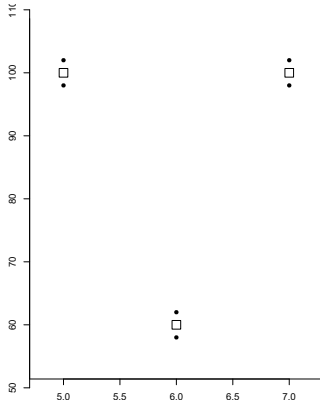
Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.

Le coppie di dati possono essere rappresentate mediante un grafico a dispersione.





2. Si calcolino, secondo il metodo dei minimi quadrati, i parametri e l'indice di adattamento dei seguenti modelli di regressione:

$$I) Y^* = a + b \cdot X \qquad II) Y^* = c \cdot X$$

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.

Si riportano i calcoli necessari¹ per ottenere le stime dei parametri e gli indici di adattamento del modello

$$Y = a + bX + E$$

x_i	y_i	$x_i y_i$	x_i^2	y_i^2
5	102	510	25	10404
5	98	490	25	9604
6	62	372	36	3844
6	58	348	36	3364
7	102	714	49	10404
7	98	686	49	9604
36	520	3120	220	47224
6	86.6667	520	36.6667	7870.6667
$M(X)$	$M(Y)$	$M(XY)$	$M(X^2)$	$M(Y^2)$

$$\hat{b} =$$



Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.

¹Si osserva che per ragioni di stampa alcuni calcoli potrebbero essere duplicati con riferimento ai diversi modelli oggetto di studio

Si riportano i calcoli necessari¹ per ottenere le stime dei parametri e gli indici di adattamento del modello

$$Y = a + bX + E$$

x_i	y_i	$x_i y_i$	x_i^2	y_i^2
5	102	510	25	10404
5	98	490	25	9604
6	62	372	36	3844
6	58	348	36	3364
7	102	714	49	10404
7	98	686	49	9604
36	520	3120	220	47224
6	86.6667	520	36.6667	7870.6667
$M(X)$	$M(Y)$	$M(XY)$	$M(X^2)$	$M(Y^2)$

$$\begin{aligned}\hat{b} &= \frac{\text{Cov}(X, Y)}{\text{Var}(X)} = \frac{M(XY) - M(X)M(Y)}{M(X^2) - [M(X)]^2} = \\ &= \frac{520 - 6 \cdot 86.6667}{36.6667 - 6^2} = \frac{0}{0.6667} = 0 \\ \hat{a} &= \end{aligned}$$

¹Si osserva che per ragioni di stampa alcuni calcoli potrebbero essere duplicati con riferimento ai diversi modelli oggetto di studio



Si riportano i calcoli necessari¹ per ottenere le stime dei parametri e gli indici di adattamento del modello

$$Y = a + bX + E$$

x_i	y_i	$x_i y_i$	x_i^2	y_i^2
5	102	510	25	10404
5	98	490	25	9604
6	62	372	36	3844
6	58	348	36	3364
7	102	714	49	10404
7	98	686	49	9604
36	520	3120	220	47224
6	86.6667	520	36.6667	7870.6667
$M(X)$	$M(Y)$	$M(XY)$	$M(X^2)$	$M(Y^2)$

$$\hat{b} = \frac{\text{Cov}(X, Y)}{\text{Var}(X)} = \frac{M(XY) - M(X)M(Y)}{M(X^2) - [M(X)^2]} =$$

$$= \frac{520 - 6 \cdot 86.6667}{36.6667 - 6^2} = \frac{0}{0.6667} = 0$$

$$\hat{a} = M(Y) - \hat{b}M(X) = 86.6667$$

$$\rho^2 =$$

¹Si osserva che per ragioni di stampa alcuni calcoli potrebbero essere duplicati con riferimento ai diversi modelli oggetto di studio



Si riportano i calcoli necessari¹ per ottenere le stime dei parametri e gli indici di adattamento del modello

$$Y = a + bX + E$$

x_i	y_i	$x_i y_i$	x_i^2	y_i^2
5	102	510	25	10404
5	98	490	25	9604
6	62	372	36	3844
6	58	348	36	3364
7	102	714	49	10404
7	98	686	49	9604
36	520	3120	220	47224
6	86.6667	520	36.6667	7870.6667
$M(X)$	$M(Y)$	$M(XY)$	$M(X^2)$	$M(Y^2)$

$$\hat{b} = \frac{\text{Cov}(X, Y)}{\text{Var}(X)} = \frac{M(XY) - M(X)M(Y)}{M(X^2) - [M(X)]^2} =$$

$$= \frac{520 - 6 \cdot 86.6667}{36.6667 - 6^2} = \frac{0}{0.6667} = 0$$

$$\hat{a} = M(Y) - \hat{b}M(X) = 86.6667$$

$$\rho^2 = \frac{[\text{Cov}(X, Y)]^2}{\text{Var}(X) \text{Var}(Y)} = \frac{0^2}{0.6667 \cdot 359.5556} = 0$$

$$M(\hat{E}^2) = \text{Var}(\hat{E}) = \text{varianza residua} =$$

¹Si osserva che per ragioni di stampa alcuni calcoli potrebbero essere duplicati con riferimento ai diversi modelli oggetto di studio



Si riportano i calcoli necessari¹ per ottenere le stime dei parametri e gli indici di adattamento del modello

$$Y = a + bX + E$$

x_i	y_i	$x_i y_i$	x_i^2	y_i^2
5	102	510	25	10404
5	98	490	25	9604
6	62	372	36	3844
6	58	348	36	3364
7	102	714	49	10404
7	98	686	49	9604
36	520	3120	220	47224
6	86.6667	520	36.6667	7870.6667
$M(X)$	$M(Y)$	$M(XY)$	$M(X^2)$	$M(Y^2)$

$$\hat{b} = \frac{\text{Cov}(X, Y)}{\text{Var}(X)} = \frac{M(XY) - M(X)M(Y)}{M(X^2) - [M(X)]^2} =$$

$$= \frac{520 - 6 \cdot 86.6667}{36.6667 - 6^2} = \frac{0}{0.6667} = 0$$

$$\hat{a} = M(Y) - \hat{b}M(X) = 86.6667$$

$$\rho^2 = \frac{[\text{Cov}(X, Y)]^2}{\text{Var}(X) \text{Var}(Y)} = \frac{0^2}{0.6667 \cdot 359.5556} = 0$$

$$M(\hat{E}^2) = \text{Var}(\hat{E}) = \text{varianza residua} = \text{Var}(Y) \cdot (1 - \rho^2) = 359.5556$$

¹Si osserva che per ragioni di stampa alcuni calcoli potrebbero essere duplicati con riferimento ai diversi modelli oggetto di studio





Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.

Si riportano i calcoli necessari² per ottenere le stime dei parametri e la misura di adattamento del modello

$$Y = cX + E$$

x_i	y_i	$x_i y_i$	x_i^2	y_i^2
5	102	510	25	10404
5	98	490	25	9604
6	62	372	36	3844
6	58	348	36	3364
7	102	714	49	10404
7	98	686	49	9604
		3120	220	47224
		520	36.6667	7870.6667
		$M(XY)$	$M(X^2)$	$M(Y^2)$

$$\hat{c} =$$

²Si osserva che per ragioni di stampa alcuni calcoli potrebbero essere duplicati con riferimento ai diversi modelli oggetto di studio



Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.

Si riportano i calcoli necessari² per ottenere le stime dei parametri e la misura di adattamento del modello

$$Y = cX + E$$

x_i	y_i	$x_i y_i$	x_i^2	y_i^2
5	102	510	25	10404
5	98	490	25	9604
6	62	372	36	3844
6	58	348	36	3364
7	102	714	49	10404
7	98	686	49	9604
		3120	220	47224
		520	36.6667	7870.6667
		$M(XY)$	$M(X^2)$	$M(Y^2)$

$$\hat{c} = \frac{M(XY)}{M(X^2)} = \frac{520}{36.6667} = 14.1818$$

$$M(\hat{E}^2) =$$

²Si osserva che per ragioni di stampa alcuni calcoli potrebbero essere duplicati con riferimento ai diversi modelli oggetto di studio



Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.

Si riportano i calcoli necessari² per ottenere le stime dei parametri e la misura di adattamento del modello

$$Y = cX + E$$

x_i	y_i	$x_i y_i$	x_i^2	y_i^2
5	102	510	25	10404
5	98	490	25	9604
6	62	372	36	3844
6	58	348	36	3364
7	102	714	49	10404
7	98	686	49	9604
		3120	220	47224
		520	36.6667	7870.6667
		$M(XY)$	$M(X^2)$	$M(Y^2)$

$$\hat{c} = \frac{M(XY)}{M(X^2)} = \frac{520}{36.6667} = 14.1818$$

$$M(\hat{E}^2) = M[Y \cdot (Y - \hat{c}X)] = M(Y^2) - \hat{c}M(XY) = 7870.6667 - 14.1818 \cdot 520 = 496.1279$$

²Si osserva che per ragioni di stampa alcuni calcoli potrebbero essere duplicati con riferimento ai diversi modelli oggetto di studio



Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.

Il modello migliore è $Y^* = 86.6667$.



3. Si fornisca la previsione in corrispondenza di $x = 8$

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.



Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.

3. Si fornisca la previsione in corrispondenza di $x = 8$
 $Y^*(8) = 86.6667$.



4. Si determini il valore del rapporto di correlazione $\eta^2_{Y|X}$.

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.



Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.

Il rapporto di correlazione

$$\eta_{Y|X}^2 = \frac{\text{Var}_X[M(Y|X)]}{\text{Var}(Y)}$$

può essere ottenuto considerando il rapporto tra la varianza spiegata $\text{Var}_X[M(Y|X)]$,
varianza della variabile statistica medie condizionate

x	$M(Y x)$
5	100
6	60
7	100

e la varianza della variabile dipendente $\text{Var}(Y) = 359.5556$.

La varianza spiegata risulta $\text{Var}_X[M(Y|X)] = 355.5556$.



Si ottiene

x	$M(Y x)$	$n(x)$	$M(Y x) \cdot n(x)$	$M(Y^2 x) \cdot n(x)$
5	100	2	200	20000
6	60	2	120	7200
7	100	2	200	20000
		6	520	47200

da cui segue

$$Var_X[M(Y|X)] = \frac{47200}{6} - 86.6667^2 = 7866.6667 - 7511.1111 = 355.5556.$$

Il rapporto di correlazione $\eta_{Y|X}^2$ risulta:

$$\eta_{Y|X}^2 = \frac{Var_X[M(Y|X)]}{Var(Y)} = \frac{355.5556}{359.5556} = 0.9889.$$

Sussiste, quindi, un livello di dipendenza molto alto.



Il rapporto di correlazione $\eta^2_{Y|X}$ può anche essere ottenuto nel modo seguente.

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.



Il rapporto di correlazione

$$\eta_{Y|X}^2 = 1 - \frac{M_X[\text{Var}(Y|X)]}{\text{Var}(Y)}$$

può essere ottenuto considerando il complemento a 1 del rapporto tra la varianza residua $M_X[\text{Var}(Y|X)]$, media della variabile statistica varianze condizionate

x	$\text{Var}(Y x)$
5	4
6	4
7	4

e la varianza della variabile dipendente $\text{Var}(Y) = 359.5556$.

La varianza residua risulta $M_X[\text{Var}(Y|X)] = 4$. Si osserva che la distribuzione è degenere. Il rapporto di correlazione $\eta_{Y|X}^2$ risulta:

$$\eta_{Y|X}^2 = 1 - \frac{M_X[\text{Var}(Y|X)]}{\text{Var}(Y)} = 1 - \frac{4}{359.5556} = 0.9889.$$

Sussiste, quindi, un livello di dipendenza molto alto.

Esercizio 4

Prova scritta di
STATISTICA del
27.06.2013

Giuseppe Boari
Gabriele Cantaluppi



UNIVERSITÀ
CATTOLICA
del Sacro Cuore

Un macchinario presenta un tasso di difettosità del 6 %.

1. Scelti a caso 4 pezzi (con reimmissione) dal flusso produttivo si calcoli la probabilità che nessuno sia difettoso.

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.

Esercizio 4

Prova scritta di
STATISTICA del
27.06.2013

Giuseppe Boari
Gabriele Cantaluppi



UNIVERSITÀ
CATTOLICA
del Sacro Cuore

Un macchinario presenta un tasso di difettosità del 6 %.

1. Scelti a caso 4 pezzi (con reimmissione) dal flusso produttivo si calcoli la probabilità che nessuno sia difettoso.

Sia X la variabile casuale 'n. di pezzi difettosi nelle 4 prove indipendenti'.

$$X \sim$$

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.

Esercizio 4



Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.

Un macchinario presenta un tasso di difettosità del 6 %.

1. Scelti a caso 4 pezzi (con reimmissione) dal flusso produttivo si calcoli la probabilità che nessuno sia difettoso.

Sia X la variabile casuale 'n. di pezzi difettosi nelle 4 prove indipendenti'.

$$X \sim \text{Bin}(n = 4, p = 0.06).$$

$$P(X = 0) =$$



Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.

Un macchinario presenta un tasso di difettosità del 6 %.

1. Scelti a caso 4 pezzi (con reimmissione) dal flusso produttivo si calcoli la probabilità che nessuno sia difettoso.

Sia X la variabile casuale 'n. di pezzi difettosi nelle 4 prove indipendenti'.

$$X \sim \text{Bin}(n = 4, p = 0.06).$$

$$P(X = 0) = \binom{4}{0} 0.06^0 0.94^4 = 0.7807.$$



2. Nell'ipotesi in cui siano effettuate 60 estrazioni si calcoli la probabilità che vi sia almeno un pezzo difettoso.

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.



2. Nell'ipotesi in cui siano effettuate 60 estrazioni si calcoli la probabilità che vi sia almeno un pezzo difettoso.

Indicata con X la variabile casuale 'numero pezzi difettosi nelle 60 prove indipendenti' risulta

$$X \sim$$

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.



Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.

2. Nell'ipotesi in cui siano effettuate 60 estrazioni si calcoli la probabilità che vi sia almeno un pezzo difettoso.

Indicata con X la variabile casuale 'numero pezzi difettosi nelle 60 prove indipendenti' risulta

$$X \sim \text{Bin}(n = 60, p = 0.06).$$

quindi

$$P(X \geq 1) =$$



2. Nell'ipotesi in cui siano effettuate 60 estrazioni si calcoli la probabilità che vi sia almeno un pezzo difettoso.

Indicata con X la variabile casuale 'numero pezzi difettosi nelle 60 prove indipendenti' risulta

$$X \sim \text{Bin}(n = 60, p = 0.06).$$

quindi

$$\begin{aligned} P(X \geq 1) &= 1 - P(X = 0) = \\ &= 1 - \binom{60}{0} 0.06^0 0.94^{60} = 1 - 0.02442 = 0.97558. \end{aligned}$$



Si osserva come non valgano le condizioni per ottenere un risultato affidabile utilizzando l'approssimazione della variabile casuale Binomiale con la variabile casuale Normale, infatti

$$np = 60 \cdot 0.06 = 3.6 \not\geq 5.$$

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.



Si riportano, per completezza, i risultati che si avrebbero utilizzando tale approssimazione:

- Ricordando che $M(X) = 3.6$ e $Var(X) = 3.3840$ abbiamo

$$X \sim N(\mu = 3.6, \sigma^2 = 3.3840)$$

quindi

$$\begin{aligned} P(X \geq 1) &= P\left(\frac{X - \mu}{\sigma} \geq \frac{1 - \mu}{\sigma}\right) = P\left(Z \geq \frac{1 - 3.6}{1.8396}\right) = \\ &= P(Z \geq -1.41) = P(Z \leq 1.41) = 0.9207. \end{aligned}$$

- anche tenendo conto della correzione per la continuità si ottiene

$$\begin{aligned} P(X \geq 1 - 0.5) &= P\left(\frac{X - \mu}{\sigma} \geq \frac{1 - 0.5 - \mu}{\sigma}\right) = \\ &= P\left(Z \geq \frac{1 - 0.5 - 3.6}{1.8396}\right) = \\ &= P(Z \geq -1.69) = P(Z \leq 1.69) = 0.9545. \end{aligned}$$

Tema T274-2

Prova scritta di STATISTICA del 11.07.2013

Prova scritta di
STATISTICA del
11.07.2013

Giuseppe Boari
Gabriele Cantaluppi



UNIVERSITÀ
CATTOLICA
del Sacro Cuore

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.

Esercizio 5.

Giuseppe Boari
Gabriele Cantaluppi
Università Cattolica del Sacro Cuore, Milano



Nel seguente prospetto sono riportati i valori assunti dalla variabile X 'stipendio mensile' riferiti a 16 soggetti, appartenenti alle categorie $i =$ 'impiegato' e $d =$ 'dirigente'.

d	3.6	4.1	3.1	4.6	4.7	4.3					
i	1.6	2.1	2.6	1.3	2.1	2.5	1.1	1.4	2.2	1.5	

1. Si rappresenti graficamente, mediante un grafico a scatola e baffi (*box & whiskers plot*), la variabile X (per tutte le unità statistiche), commentando i risultati ottenuti.

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.

Esercizio 5.



Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.

Esercizio 5.

Si considera la distribuzione della variabile X ,

x_i	n_i
1.1	1
1.3	1
1.4	1
1.5	1
1.6	1
2.1	2
2.2	1
2.5	1
2.6	1
3.1	1
3.6	1
4.1	1
4.3	1
4.6	1
4.7	1

Si tratta di un carattere quantitativo con modalità non raggruppate in classi.



Per costruire il grafico Box & Whiskers plot occorre determinare in primo luogo i valori dei tre quartili $x_{0.25}$, $x_{0.5}$ e $x_{0.75}$ considerando l'inversa della Funzione di Ripartizione in corrispondenza rispettivamente di 0.25, 0.50 e 0.75.

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.

Esercizio 5.



Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.

Esercizio 5.

x_i	n_i	N_i	F_i
1.1	1	1	0.0625
1.3	1	2	0.1250
1.4	1	3	0.1875
1.5	1	4	0.2500
1.6	1	5	0.3125
2.1	2	7	0.4375
2.2	1	8	0.5000
2.5	1	9	0.5625
2.6	1	10	0.6250
3.1	1	11	0.6875
3.6	1	12	0.7500
4.1	1	13	0.8125
4.3	1	14	0.8750
4.6	1	15	0.9375
4.7	1	16	1.0000



Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.

Esercizio 5.

Si ottiene

$$x_{0.25} = 1.55$$

$$x_{0.5} = 2.35$$

$$x_{0.75} = 3.85$$

La lunghezza teorica dei baffi risulta

$$1.5 \cdot (x_{0.75} - x_{0.25}) = 3.45.$$

Il baffo inferiore è posizionato in

$$\max(x_{\min}, x_{0.25} - 3.45) = \max(1.1, -1.9) = 1.1.$$

Il baffo superiore è posizionato in

$$\min(x_{\max}, x_{0.75} + 3.45) = \min(4.7, 7.3) = 4.7.$$



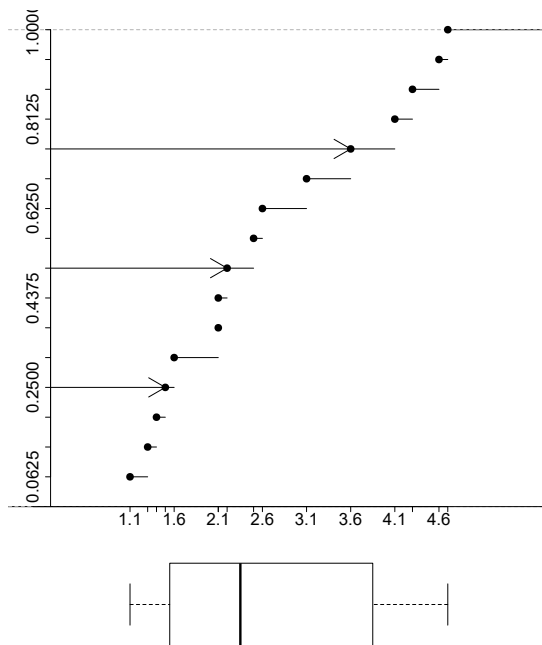
Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.

Esercizio 5.





2. Si confrontino con opportuni indici di posizione e di variabilità le distribuzioni delle variabili $X|i$ e $X|d$ per le due categorie lavorative. Si commentino i risultati ottenuti.

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.

Esercizio 5.



Per effettuare un confronto in termini di posizione della variabile X nei 2 gruppi *dirigenti* e *impiegati* si considerano le due medie aritmetiche di gruppo $\mu_{dirigenti}$ e $\mu_{impiegati}$:

$$\mu_{dirigenti} = \frac{1}{6}(3.6 + 4.1 + 3.1 + 4.6 + 4.7 + 4.3) = 4.0667$$

e

$$\mu_{impiegati} = \frac{1}{10}(1.6 + 2.1 + 2.6 + 1.3 + 2.1 + 2.5 + 1.1 + 1.4 + 2.2 + 1.5) = 1.84$$

Il livello medio nel gruppo *dirigenti* è superiore a quello osservato nel gruppo *impiegati*.

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.

Esercizio 5.

Per effettuare un confronto in termini di variabilità della variabile X nei 2 gruppi *dirigenti* e *impiegati* si considerano i coefficienti di variazione calcolati per i due gruppi $CV(X_{\text{dirigenti}})$ e $CV(X_{\text{impiegati}})$.



Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.

Esercizio 5.



Per effettuare un confronto in termini di variabilità della variabile X nei 2 gruppi *dirigenti* e *impiegati* si considerano i coefficienti di variazione calcolati per i due gruppi $CV(X_{dirigenti})$ e $CV(X_{impiegati})$. A tal fine occorre in primo luogo calcolare le varianze nei due gruppi

$$\sigma_{dirigenti}^2 = \frac{1}{6}(3.6^2 + 4.1^2 + 3.1^2 + 4.6^2 + 4.7^2 + 4.3^2) - \mu_{dirigenti}^2 = 0.3156$$

e

$$\sigma_{impiegati}^2 = \frac{1}{10}(1.6^2 + 2.1^2 + 2.6^2 + 1.3^2 + 2.1^2 + 2.5^2 + 1.1^2 + 1.4^2 + 2.2^2 + 1.5^2) - \mu_{impiegati}^2 = 0.2484$$

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.

Esercizio 5.



Per effettuare un confronto in termini di variabilità della variabile X nei 2 gruppi *dirigenti* e *impiegati* si considerano i coefficienti di variazione calcolati per i due gruppi $CV(X_{dirigenti})$ e $CV(X_{impiegati})$. A tal fine occorre in primo luogo calcolare le varianze nei due gruppi

$$\sigma_{dirigenti}^2 = \frac{1}{6}(3.6^2 + 4.1^2 + 3.1^2 + 4.6^2 + 4.7^2 + 4.3^2) - \mu_{dirigenti}^2 = 0.3156$$

e

$$\sigma_{impiegati}^2 = \frac{1}{10}(1.6^2 + 2.1^2 + 2.6^2 + 1.3^2 + 2.1^2 + 2.5^2 + 1.1^2 + 1.4^2 + 2.2^2 + 1.5^2) - \mu_{impiegati}^2 = 0.2484$$

Ricordando che il coefficiente di variazione è definito dal rapporto tra scarto quadratico medio e media aritmetica risulta

$$CV(X_{dirigenti}) = \frac{\sigma_{dirigenti}}{\mu_{dirigenti}} = \frac{0.5617433182}{4.066666667} = 0.1381$$

e

$$CV(X_{impiegati}) = \frac{\sigma_{impiegati}}{\mu_{impiegati}} = \frac{0.4983974318}{1.84} = 0.2709$$

Il livello di variabilità presente nel gruppo *dirigenti* è inferiore a quello osservato nel gruppo *impiegati*.



3. Sulla base dei valori delle medie e varianze parziali, ottenute al punto precedente, si ricavano i valori della media e varianza della variabile X riferiti all'intera popolazione.

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.

Esercizio 5.



Applicando la proprietà associativa della media aritmetica si ottiene:

$$\begin{aligned}\mu &= \frac{1}{16} (\mu_{\text{dirigenti}} \cdot 6 + \mu_{\text{impiegati}} \cdot 10) \\ &= \frac{1}{16} (4.0667 \cdot 6 + 1.84 \cdot 10) = 2.675\end{aligned}$$

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.

Esercizio 5.



Applicando il risultato di scomposizione della varianza si ottiene:

$$\sigma_{Between}^2 =$$

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.

Esercizio 5.



Applicando il risultato di scomposizione della varianza si ottiene:

$$\begin{aligned}\sigma_{Between}^2 &= \frac{1}{16} \left\{ (\mu_{dirigenti} - \mu)^2 \cdot 6 + (\mu_{impiegati} - \mu)^2 \cdot 10 \right\} \\ &= \frac{1}{16} \left(\mu_{dirigenti}^2 \cdot 6 + \mu_{impiegati}^2 \cdot 10 \right) - \mu^2 \\ &= \frac{1}{16} \left(4.0667^2 \cdot 6 + 1.84^2 \cdot 10 \right) - 2.675^2 = 1.162\end{aligned}$$

$$\sigma_{Within}^2 =$$

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.

Esercizio 5.



Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.

Esercizio 5.

Applicando il risultato di scomposizione della varianza si ottiene:

$$\begin{aligned}\sigma_{Between}^2 &= \frac{1}{16} \left\{ (\mu_{dirigenti} - \mu)^2 \cdot 6 + (\mu_{impiegati} - \mu)^2 \cdot 10 \right\} \\ &= \frac{1}{16} \left(\mu_{dirigenti}^2 \cdot 6 + \mu_{impiegati}^2 \cdot 10 \right) - \mu^2 \\ &= \frac{1}{16} \left(4.0667^2 \cdot 6 + 1.84^2 \cdot 10 \right) - 2.675^2 = 1.162\end{aligned}$$

$$\begin{aligned}\sigma_{Within}^2 &= \frac{1}{16} \left(\sigma_{dirigenti}^2 \cdot 6 + \sigma_{impiegati}^2 \cdot 10 \right) \\ &= \frac{1}{16} (0.3156 \cdot 6 + 0.2484 \cdot 10) = 0.2736.\end{aligned}$$

e

$$\sigma^2 =$$



Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.

Esercizio 5.

Applicando il risultato di scomposizione della varianza si ottiene:

$$\begin{aligned}\sigma_{Between}^2 &= \frac{1}{16} \left\{ (\mu_{dirigenti} - \mu)^2 \cdot 6 + (\mu_{impiegati} - \mu)^2 \cdot 10 \right\} \\ &= \frac{1}{16} \left(\mu_{dirigenti}^2 \cdot 6 + \mu_{impiegati}^2 \cdot 10 \right) - \mu^2 \\ &= \frac{1}{16} \left(4.0667^2 \cdot 6 + 1.84^2 \cdot 10 \right) - 2.675^2 = 1.162\end{aligned}$$

$$\begin{aligned}\sigma_{Within}^2 &= \frac{1}{16} \left(\sigma_{dirigenti}^2 \cdot 6 + \sigma_{impiegati}^2 \cdot 10 \right) \\ &= \frac{1}{16} (0.3156 \cdot 6 + 0.2484 \cdot 10) = 0.2736.\end{aligned}$$

e

$$\sigma^2 = \sigma_{Between}^2 + \sigma_{Within}^2 = 1.162 + 0.2736 = 1.4356.$$



Si consideri la serie NIBM relativa alla spesa delle Amministrazioni Pubbliche rilevata tra il 2005 e il 2009:

t	$NIBM$
2005	
2006	1.0304
2007	1.0164
2008	0.7484
2009	1.4397

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.

Esercizio 5.



1. Sapendo che $x_{2009} = 334$ si calcoli il valore di x_{2005} .

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.

Esercizio 5.



Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.

Esercizio 5.

1. Sapendo che $x_{2009} = 334$ si calcoli il valore di x_{2005} .

$$\begin{aligned} {}_{2005}I_{2009} &= {}_{2005}I_{2006} \cdot {}_{2006}I_{2007} \cdot {}_{2007}I_{2008} \cdot {}_{2008}I_{2009} = \\ &= 1.0304 \cdot 1.0164 \cdot 0.7484 \cdot 1.4397 = 1.1284 \end{aligned}$$

$${}_{2005}I_{2009} = \frac{x_{2009}}{x_{2005}} = 1.1284$$

da cui

$$x_{2005} = \frac{334}{1.1284} = 296$$



Il seguente prospetto riporta i valori rilevati in corrispondenza a 6 unità statistiche con riferimento a due caratteri quantitativi X e Y

x_i	5	5	6	6	7	7
y_i	102	98	62	58	102	98

1. Si rappresentino graficamente le coppie di dati.

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.

Esercizio 5.



Esercizio 1.

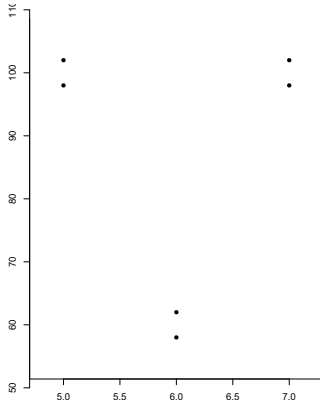
Esercizio 2.

Esercizio 3.

Esercizio 4.

Esercizio 5.

Le coppie di dati possono essere rappresentate mediante un grafico a dispersione.





2. Si calcolino, secondo il metodo dei minimi quadrati, i parametri e l'indice di adattamento dei seguenti modelli di regressione:

$$I) Y^* = a + b \cdot (X - 6)^2 \qquad II) Y^* = c$$

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.

Esercizio 5.

Si riportano i calcoli necessari¹ per ottenere le stime dei parametri e gli indici di adattamento del modello

$$Y = a + b(X - 6)^2 + E$$

$(x - 6)^2_i$	y_i	$(x - 6)^2_i y_i$	$(x - 6)^{2^2}_i$	y_i^2
1	102	102	1	10404
1	98	98	1	9604
0	62	0	0	3844
0	58	0	0	3364
1	102	102	1	10404
1	98	98	1	9604
4	520	400	4	47224
0.6667	86.6667	66.6667	0.6667	7870.6667
$M((X - 6)^2)$	$M(Y)$	$M((X - 6)^2 Y)$	$M((X - 6)^{2^2})$	$M(Y^2)$

$$\hat{b} =$$



Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.

Esercizio 5.

¹Si osserva che per ragioni di stampa alcuni calcoli potrebbero essere duplicati con riferimento ai diversi modelli oggetto di studio

Si riportano i calcoli necessari¹ per ottenere le stime dei parametri e gli indici di adattamento del modello

$$Y = a + b(X - 6)^2 + E$$

$(x - 6)^2_i$	y_i	$(x - 6)^2_i y_i$	$(x - 6)^{2^2}_i$	y_i^2
1	102	102	1	10404
1	98	98	1	9604
0	62	0	0	3844
0	58	0	0	3364
1	102	102	1	10404
1	98	98	1	9604
4	520	400	4	47224
0.6667	86.6667	66.6667	0.6667	7870.6667
$M((X - 6)^2)$	$M(Y)$	$M((X - 6)^2 Y)$	$M((X - 6)^{2^2})$	$M(Y^2)$

$$\begin{aligned}\hat{b} &= \frac{\text{Cov}((X - 6)^2, Y)}{\text{Var}((X - 6)^2)} = \frac{M((X - 6)^2 Y) - M((X - 6)^2) M(Y)}{M((X - 6)^{2^2}) - [M((X - 6)^2)]^2} = \\ &= \frac{66.6667 - 0.6667 \cdot 86.6667}{0.6667 - 0.6667^2} = \frac{8.8889}{0.2222} = 40.0041 \\ \hat{a} &= \end{aligned}$$

¹Si osserva che per ragioni di stampa alcuni calcoli potrebbero essere duplicati con riferimento ai diversi modelli oggetto di studio



Si riportano i calcoli necessari¹ per ottenere le stime dei parametri e gli indici di adattamento del modello

$$Y = a + b(X - 6)^2 + E$$

$(x - 6)^2_i$	y_i	$(x - 6)^2_i y_i$	$(x - 6)^{2^2}_i$	y_i^2
1	102	102	1	10404
1	98	98	1	9604
0	62	0	0	3844
0	58	0	0	3364
1	102	102	1	10404
1	98	98	1	9604
4	520	400	4	47224
0.6667	86.6667	66.6667	0.6667	7870.6667
$M((X - 6)^2)$	$M(Y)$	$M((X - 6)^2 Y)$	$M((X - 6)^{2^2})$	$M(Y^2)$

$$\begin{aligned}\hat{b} &= \frac{\text{Cov}((X - 6)^2, Y)}{\text{Var}((X - 6)^2)} = \frac{M((X - 6)^2 Y) - M((X - 6)^2) M(Y)}{M((X - 6)^{2^2}) - [M((X - 6)^2)]^2} = \\ &= \frac{66.6667 - 0.6667 \cdot 86.6667}{0.6667 - 0.6667^2} = \frac{8.8889}{0.2222} = 40.0041 \\ \hat{a} &= M(Y) - \hat{b} M((X - 6)^2) = 59.9973 \\ \rho^2 &= \end{aligned}$$

¹Si osserva che per ragioni di stampa alcuni calcoli potrebbero essere duplicati con riferimento ai diversi modelli oggetto di studio



Si riportano i calcoli necessari¹ per ottenere le stime dei parametri e gli indici di adattamento del modello

$$Y = a + b(X - 6)^2 + E$$

$(x - 6)^2_i$	y_i	$(x - 6)^2_i y_i$	$(x - 6)^{2^2}_i$	y_i^2
1	102	102	1	10404
1	98	98	1	9604
0	62	0	0	3844
0	58	0	0	3364
1	102	102	1	10404
1	98	98	1	9604
4	520	400	4	47224
0.6667	86.6667	66.6667	0.6667	7870.6667
$M((X - 6)^2)$	$M(Y)$	$M((X - 6)^2 Y)$	$M((X - 6)^{2^2})$	$M(Y^2)$

$$\hat{b} = \frac{\text{Cov}((X - 6)^2, Y)}{\text{Var}((X - 6)^2)} = \frac{M((X - 6)^2 Y) - M((X - 6)^2) M(Y)}{M((X - 6)^{2^2}) - [M((X - 6)^2)]^2} =$$

$$= \frac{66.6667 - 0.6667 \cdot 86.6667}{0.6667 - 0.6667^2} = \frac{8.8889}{0.2222} = 40.0041$$

$$\hat{a} = M(Y) - \hat{b} M((X - 6)^2) = 59.9973$$

$$\rho^2 = \frac{[\text{Cov}((X - 6)^2, Y)]^2}{\text{Var}((X - 6)^2) \text{Var}(Y)} = \frac{8.8889^2}{0.2222 \cdot 359.5556} = 0.989$$

¹Si osserva che per ragioni di stampa alcuni calcoli potrebbero essere duplicati con riferimento ai diversi modelli oggetto di studio





Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.

Esercizio 5.

Si riportano i calcoli necessari² per ottenere le stime dei parametri e gli indici di adattamento del modello

$$Y = c + E$$

x_i	y_i	y_i^2
5	102	10404
5	98	9604
6	62	3844
6	58	3364
7	102	10404
7	98	9604
	520	47224
	86.6667	7870.6667
	$M(Y)$	$M(Y^2)$

$$\hat{c} =$$

²Si osserva che per ragioni di stampa alcuni calcoli potrebbero essere duplicati con riferimento ai diversi modelli oggetto di studio



Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.

Esercizio 5.

Si riportano i calcoli necessari² per ottenere le stime dei parametri e gli indici di adattamento del modello

$$Y = c + E$$

x_i	y_i	y_i^2
5	102	10404
5	98	9604
6	62	3844
6	58	3364
7	102	10404
7	98	9604
	520	47224
	86.6667	7870.6667
	$M(Y)$	$M(Y^2)$

$$\hat{c} = M(Y) = 86.6667$$

$$R^2 =$$

²Si osserva che per ragioni di stampa alcuni calcoli potrebbero essere duplicati con riferimento ai diversi modelli oggetto di studio



Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.

Esercizio 5.

Si riportano i calcoli necessari² per ottenere le stime dei parametri e gli indici di adattamento del modello

$$Y = c + E$$

x_i	y_i	y_i^2
5	102	10404
5	98	9604
6	62	3844
6	58	3364
7	102	10404
7	98	9604
	520	47224
	86.6667	7870.6667
	$M(Y)$	$M(Y^2)$

$$\hat{c} = M(Y) = 86.6667$$

$$R^2 = 0$$

²Si osserva che per ragioni di stampa alcuni calcoli potrebbero essere duplicati con riferimento ai diversi modelli oggetto di studio



Il modello migliore è $Y^* = 60 + 40X$.

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.

Esercizio 5.

Esercizio 4

Prova scritta di
STATISTICA del
11.07.2013

Giuseppe Boari
Gabriele Cantaluppi



UNIVERSITÀ
CATTOLICA
del Sacro Cuore

Si completino le seguenti tabelle, giustificando opportunamente la risposta, in modo tale che:

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.

Esercizio 5.



Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.

Esercizio 5.

a) $\rho^2 = 1$ ed $\eta_{Y|X}^2 = 0$

	$y_1 = 2$	$y_2 = 4$	$y_3 = 6$	
$x_1 = 1$	7			
$x_2 = 2$				
				$n = 20$



Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.

Esercizio 5.

a) $\rho^2 = 1$ ed $\eta_{Y|X}^2 = 0$

	$y_1 = 2$	$y_2 = 4$	$y_3 = 6$	
$x_1 = 1$	7			
$x_2 = 2$				
				$n = 20$

Impossibile in quanto $\rho^2 \leq \eta_{Y|X}^2$



Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.

Esercizio 5.

b) $\eta_{X|Y}^2 = 1$ ed $\eta_{Y|X}^2 = 0$

	$y_1 = 2$	$y_2 = 4$	$y_3 = 6$	
$x_1 = 1$	7	0	7	
$x_2 = 2$	0	6	0	
				$n = 20$



Esercizio 1.

Esercizio 2.

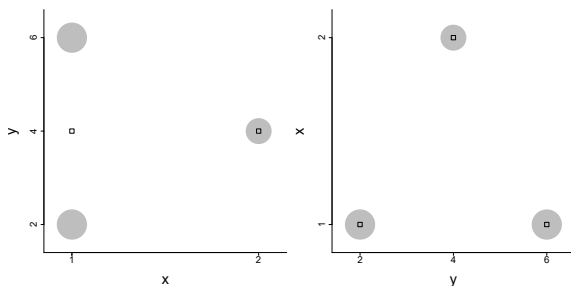
Esercizio 3.

Esercizio 4.

Esercizio 5.

b) $\eta_{X|Y}^2 = 1$ ed $\eta_{Y|X}^2 = 0$

	$y_1 = 2$	$y_2 = 4$	$y_3 = 6$	
$x_1 = 1$	7	0	7	
$x_2 = 2$	0	6	0	
				$n = 20$



Dipendenza funzionale di X da $Y \rightarrow X|y_j$ degeneri Indipendenza in media di Y da $X \rightarrow$
 $M(Y|x_1) = M(Y|x_2) = M(Y) = y_2$

Esercizio 5

Prova scritta di
STATISTICA del
11.07.2013

Giuseppe Boari
Gabriele Cantaluppi



UNIVERSITÀ
CATTOLICA
del Sacro Cuore

Per il seguente gioco è prevista una posta di 2 Euro: si lanciano due dadi e se la somma dei numeri ottenuti è almeno pari a 9 si vincono 6 Euro.

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.

Esercizio 5.

1. Si stabilisca se il gioco è equo.



Esercizio 1.

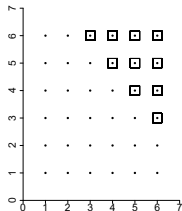
Esercizio 2.

Esercizio 3.

Esercizio 4.

Esercizio 5.

1. Si stabilisca se il gioco è equo.



Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.

Esercizio 5.



Esercizio 1.

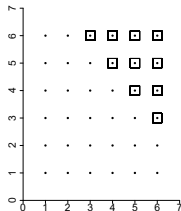
Esercizio 2.

Esercizio 3.

Esercizio 4.

Esercizio 5.

1. Si stabilisca se il gioco è equo.



$$P(\text{vincita}) = \frac{10}{36} = 0.2778$$



Esercizio 1.

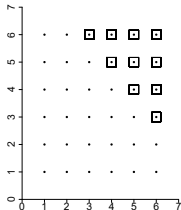
Esercizio 2.

Esercizio 3.

Esercizio 4.

Esercizio 5.

1. Si stabilisca se il gioco è equo.



$$P(\text{vincita}) = \frac{10}{36} = 0.2778$$

	x_i	$P(X = x_i)$	$x_i \cdot P(X = x_i)$
<i>posta</i>	-2	0.7222	-1.4444
<i>vincita - posta</i>	4	0.2778	1.1111
	1		-0.3333



Esercizio 1.

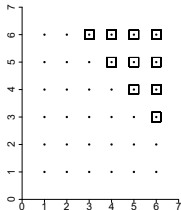
Esercizio 2.

Esercizio 3.

Esercizio 4.

Esercizio 5.

1. Si stabilisca se il gioco è equo.



$$P(\text{vincita}) = \frac{10}{36} = 0.2778$$

	x_i	$P(X = x_i)$	$x_i \cdot P(X = x_i)$
<i>posta</i>	-2	0.7222	-1.4444
<i>vincita - posta</i>	4	0.2778	1.1111
	1		-0.3333

Il valore atteso della variabile casuale guadagno è negativo, quindi il gioco non è equo.



2. Nell'ipotesi di ripetere il gioco 5 volte si calcoli la probabilità di vincere almeno 2 volte.

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.

Esercizio 5.



2. Nell'ipotesi di ripetere il gioco 5 volte si calcoli la probabilità di vincere almeno 2 volte. Sia X la variabile casuale 'n. di vincite nelle 5 prove indipendenti'.

$$X \sim$$

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.

Esercizio 5.



2. Nell'ipotesi di ripetere il gioco 5 volte si calcoli la probabilità di vincere almeno 2 volte. Sia X la variabile casuale 'n. di vincite nelle 5 prove indipendenti'.

$$X \sim \text{Bin}(n = 5, p = 0.2778).$$

$$\begin{aligned} P(X \geq 2) &= 1 - P(X < 2) = 1 - \{P(X = 0) + P(X = 1)\} = \\ &= 1 - \binom{5}{0} 0.2778^0 0.7222^5 - \binom{5}{1} 0.2778^1 0.7222^4 = \\ &= 1 - 0.1965 - 0.3779 = 0.4256. \end{aligned}$$

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.

Esercizio 5.

Tema T275

Prova scritta di STATISTICA del 05.09.2013

Prova scritta di
STATISTICA del
05.09.2013

Giuseppe Boari
Gabriele Cantaluppi



UNIVERSITÀ
CATTOLICA
del Sacro Cuore

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.

Esercizio 5.

Giuseppe Boari
Gabriele Cantaluppi
Università Cattolica del Sacro Cuore, Milano



Nella tabella si riportano i risultati di un'indagine sull'uso dei mezzi di trasporto da parte dei giovani nella città di Milano.

Mezzo di trasporto (T) \ Età (E)	12 - 16	16 - 20	20 - 30
bicicletta	3	6	9
mezzo pubblico	10	8	9
motorino	3	11	8

1. Si rappresenti graficamente la variabile marginale E .

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.

Esercizio 5.



Si considera la distribuzione marginale della variabile E , ottenuta sommando le frequenze di colonna

e_j	$n_{\bullet j}$
12 $\rightarrow$ 16	16
16 $\rightarrow$ 20	25
20 $\rightarrow$ 30	26

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.

Esercizio 5.



Si tratta di un carattere quantitativo con modalità raggruppate in classi che può essere rappresentato graficamente mediante un istogramma.

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.

Esercizio 5.



L'istogramma può essere costruito dopo avere calcolato le densità $d_{\bullet j} = \frac{n_{\bullet j}}{a_j}$, (dove a_j sono le ampiezze di classe).

e_j	$n_{\bullet j}$	$a_{\bullet j}$	$d_{\bullet j}$
12 - 16	16	16 - 12 = 4	4
16 - 20	25	20 - 16 = 4	6.25
20 - 30	26	30 - 20 = 10	2.6

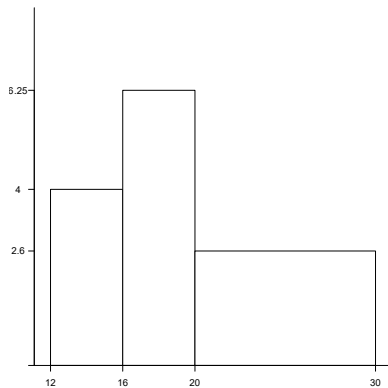
Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.

Esercizio 5.



Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.

Esercizio 5.

2. Si calcolino, se possibile, moda, media aritmetica e mediana per la variabile T . Si calcolino, se possibile, moda, media aritmetica e si individui su un opportuno grafico la mediana per la variabile E .



Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.

Esercizio 5.



Si considera la distribuzione marginale della variabile T , ottenuta sommando le frequenze di riga

t_i	$n_{i\bullet}$
<i>bicicletta</i>	18
<i>mezzo pubblico</i>	27
<i>motorino</i>	22

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.

Esercizio 5.



Si tratta di un carattere qualitativo sconnesso.

L'unico indice di posizione che può essere definito è la moda (categoria alla quale è associata la massima frequenza) che risulta *mezzo pubblico*.

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.

Esercizio 5.



Si considera la distribuzione marginale della variabile E , ottenuta sommando le frequenze di colonna

e_j	$n_{\bullet j}$
12 $\rightarrow$ 16	16
16 $\rightarrow$ 20	25
20 $\rightarrow$ 30	26

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.

Esercizio 5.



Si tratta di un carattere quantitativo con modalità raggruppate in classi.

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.

Esercizio 5.



La moda corrisponde alla classe alla quale è associata la massima densità di frequenza

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.

Esercizio 5.



La moda corrisponde alla classe alla quale è associata la massima densità di frequenza

e_j	$n_{\bullet j}$	$a_{\bullet j}$	$d_{\bullet j}$
12 $\vdash$ 16	16	$16 - 12 = 4$	4
16 $\vdash$ 20	25	$20 - 16 = 4$	6.25
20 $\vdash$ 30	26	$30 - 20 = 10$	2.6

e risulta 16 $\vdash$ 20.

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.

Esercizio 5.



La mediana (percentile di ordine 0.50) può essere ottenuta considerando l'inversa della Funzione di Ripartizione in corrispondenza di 0.50.

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.

Esercizio 5.



La mediana (percentile di ordine 0.50) può essere ottenuta considerando l'inversa della Funzione di Ripartizione in corrispondenza di 0.50.

e_j	$n_{\bullet j}$	$N_{\bullet j}$	$F_{\bullet j}$
12 - 16	16	16	0.2388
16 - 20	25	41	0.6119
20 - 30	26	67	1

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.

Esercizio 5.



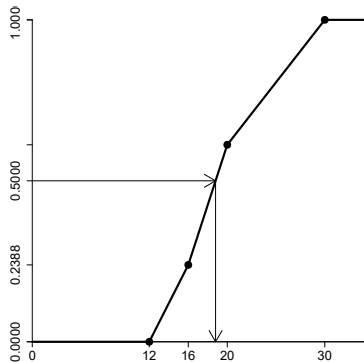
Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.

Esercizio 5.





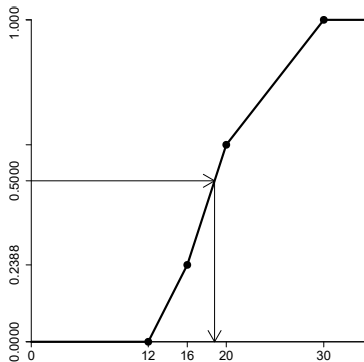
Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.

Esercizio 5.



Si ottiene

$$e_{0.5} = 16 + \frac{0.5 - 0.2388}{0.6119 - 0.2388} \cdot (20 - 16) = 18.8.$$



La media aritmetica

$$M(E) = \frac{1}{n} \sum_{j=1}^k e_j n_{\bullet j}$$

risulta 19.7612.

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.

Esercizio 5.



La media aritmetica

$$M(E) = \frac{1}{n} \sum_{j=1}^k e_j n_{\bullet j}$$

risulta 19.7612.

Si ottiene (considerando come modalità i valori centrali delle classi)

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.

Esercizio 5.



La media aritmetica

$$M(E) = \frac{1}{n} \sum_{j=1}^k e_j n_{\bullet j}$$

risulta 19.7612.

Si ottiene (considerando come modalità i valori centrali delle classi)

e_j	$n_{\bullet j}$	$e_j n_{\bullet j}$
14	16	224
18	25	450
25	26	650
	67	1324

da cui segue

$$M(E) = \frac{1}{n} \sum_{i=1}^k e_j n_{\bullet j} = \frac{1324}{67} = 19.7612.$$

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.

Esercizio 5.

3. Si calcoli per la variabile T un opportuno indice di variabilità e se ne commenti il risultato.

Prova scritta di
STATISTICA del
05.09.2013

Giuseppe Boari
Gabriele Cantaluppi



UNIVERSITÀ
CATTOLICA
del Sacro Cuore

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.

Esercizio 5.



Si considera la distribuzione marginale della variabile T , ottenuta sommando le frequenze di riga

t_i	$n_{i\bullet}$
<i>bicicletta</i>	18
<i>mezzo pubblico</i>	27
<i>motorino</i>	22

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.

Esercizio 5.

In presenza di un carattere qualitativo la variabilità può essere misurata mediante gli indici di Gini e di Frosini normalizzati

$$G_N = \frac{1 - \sum_{i=1}^h f_{i\bullet}^2}{1 - \frac{1}{h}}$$

e

$$F_N = 1 - \sqrt{1 - G_N}.$$



Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.

Esercizio 5.



Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.

Esercizio 5.

In presenza di un carattere qualitativo la variabilità può essere misurata mediante gli indici di Gini e di Frosini normalizzati

$$G_N = \frac{1 - \sum_{i=1}^h f_{i\bullet}^2}{1 - \frac{1}{h}}$$

e

$$F_N = 1 - \sqrt{1 - G_N}.$$

Si ottiene:

t_i	$n_{i\bullet}$	$f_{i\bullet}$	$f_{i\bullet}^2$
<i>bicicletta</i>	18	0.2687	0.0722
<i>mezzo pubblico</i>	27	0.403	0.1624
<i>motorino</i>	22	0.3284	0.1078
	67	1	0.3424



In presenza di un carattere qualitativo la variabilità può essere misurata mediante gli indici di Gini e di Frosini normalizzati

$$G_N = \frac{1 - \sum_{i=1}^h f_{i\bullet}^2}{1 - \frac{1}{h}}$$

e

$$F_N = 1 - \sqrt{1 - G_N}.$$

Si ottiene:

t_i	$n_{i\bullet}$	$f_{i\bullet}$	$f_{i\bullet}^2$
<i>bicicletta</i>	18	0.2687	0.0722
<i>mezzo pubblico</i>	27	0.403	0.1624
<i>motorino</i>	22	0.3284	0.1078
	67	1	0.3424

Da cui segue:

$$G_N = \frac{1 - 0.3424}{1 - \frac{1}{3}} = 0.9864$$

e

$$F_N = 1 - \sqrt{1 - 0.9864} = 0.8834.$$

Sussiste, quindi, un livello alto di eterogeneità.



4. Si calcoli un opportuno indice di connessione.

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.

Esercizio 5.



Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.

Esercizio 5.

Dalla distribuzione delle frequenze congiunte n_{ij} :

	12 ÷ 16	16 ÷ 20	20 ÷ 30	$n_{i\bullet}$
<i>bicicletta</i>	3	6	9	18
<i>mezzo pubblico</i>	10	8	9	27
<i>motorino</i>	3	11	8	22
$n_{\bullet j}$	16	25	26	67

è possibile ottenere le frequenze teoriche nella situazione di indipendenza stocastica

$$\hat{n}_{ij} = \frac{n_{i\bullet} n_{\bullet j}}{n}:$$



Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.

Esercizio 5.

Dalla distribuzione delle frequenze congiunte n_{ij} :

	12 ÷ 16	16 ÷ 20	20 ÷ 30	$n_{i\bullet}$
<i>bicicletta</i>	3	6	9	18
<i>mezzo pubblico</i>	10	8	9	27
<i>motorino</i>	3	11	8	22
$n_{\bullet j}$	16	25	26	67

è possibile ottenere le frequenze teoriche nella situazione di indipendenza stocastica

$$\hat{n}_{ij} = \frac{n_{i\bullet} \cdot n_{\bullet j}}{n}:$$

	12 ÷ 16	16 ÷ 20	20 ÷ 30	$n_{i\bullet}$
<i>bicicletta</i>	4.2985	6.7164	6.9851	18
<i>mezzo pubblico</i>	6.4478	10.0746	10.4776	27
<i>motorino</i>	5.2537	8.2090	8.5373	22
$n_{\bullet j}$	16.0000	25.0000	26.0000	67



le contingenze assolute $n_{ij} - \hat{n}_{ij}$:

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.

Esercizio 5.



le contingenze assolute $n_{ij} - \hat{n}_{ij}$:

	12 ÷ 16	16 ÷ 20	20 ÷ 30
<i>bicicletta</i>	-1.2985	-0.7164	2.0149
<i>mezzo pubblico</i>	3.5522	-2.0746	-1.4776
<i>motorino</i>	-2.2537	2.7910	-0.5373

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.

Esercizio 5.



i rapporti di contingenza $\frac{n_{ij}}{\hat{n}_{ij}}$:

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.

Esercizio 5.



i rapporti di contingenza $\frac{n_{ij}}{\hat{n}_{ij}}$:

	12 ÷ 16	16 ÷ 20	20 ÷ 30
<i>bicicletta</i>	0.6979	0.8933	1.2885
<i>mezzo pubblico</i>	1.5509	0.7941	0.8590
<i>motorino</i>	0.5710	1.3400	0.9371

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.

Esercizio 5.



e le contingenze relative, espresse in termini percentuali, $100 \left(\frac{n_{ij} - \hat{n}_{ij}}{\hat{n}_{ij}} \right) \% = 100 \left(\frac{n_{ij}}{\hat{n}_{ij}} - 1 \right) \%$:

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.

Esercizio 5.



e le contingenze relative, espresse in termini percentuali, $100 \left(\frac{n_{ij} - \hat{n}_{ij}}{\hat{n}_{ij}} \right) \% = 100 \left(\frac{n_{ij}}{\hat{n}_{ij}} - 1 \right) \%$:

	12 ÷ 16	16 ÷ 20	20 ÷ 30
<i>bicicletta</i>	-30.21	-10.67	28.85
<i>mezzo pubblico</i>	55.09	-20.59	-14.10
<i>motorino</i>	-42.90	34.00	-6.29

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.

Esercizio 5.



Risultati intermedi $\frac{(n_{ij} - \hat{n}_{ij})^2}{\hat{n}_{ij}}$ per il calcolo del χ^2

	12 ÷ 16	16 ÷ 20	20 ÷ 30
<i>bicicletta</i>	0.3923	0.0764	0.5812
<i>mezzo pubblico</i>	1.9570	0.4272	0.2084
<i>motorino</i>	0.9668	0.9490	0.0338

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.

Esercizio 5.



Risultati intermedi $\frac{n_{ij}^2}{n_{i\bullet} \cdot n_{\bullet j}}$ per il calcolo del χ^2 (formula operativa)

	12 ÷ 16	16 ÷ 20	20 ÷ 30
<i>bicicletta</i>	0.0312	0.0800	0.1731
<i>mezzo pubblico</i>	0.2315	0.0948	0.1154
<i>motorino</i>	0.0256	0.2200	0.1119

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.

Esercizio 5.



Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.

Esercizio 5.

$$\chi^2 = \sum_{i=1}^h \sum_{j=1}^k \frac{(n_{ij} - \hat{n}_{ij})^2}{\hat{n}_{ij}} = 5.5921,$$

$$\chi_N^2 = \frac{\chi^2}{n \cdot \min(h-1, k-1)} = 0.0417$$

$$\chi_N = \sqrt{\chi_N^2} = 0.2042.$$

Sussiste, quindi, un livello di connessione molto basso.



Si consideri la seguente serie storica degli indici a base mobile relativa ai costi di un'azienda tra il 2008 e il 2012:

t	$NIBM$
2008	
2009	1.20
2010	0.55
2011	0.85
2012	1.05

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.

Esercizio 5.



1. Si costruisca la serie dei numeri indici a base fissa con base 2008.

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.

Esercizio 5.



1. Si costruisca la serie dei numeri indici a base fissa con base 2008.
Si ricorda che ${}_{2008}I_{2008} = 1$ e che per valori di $t > 2008$ vale

$${}_{2008}I_t = {}_{2008}I_{2009} \cdot {}_{2009}I_{2010} \cdot \dots \cdot {}_{t-1}I_t.$$

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.

Esercizio 5.



Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.

Esercizio 5.

1. Si costruisca la serie dei numeri indici a base fissa con base 2008.
Si ricorda che ${}_{2008}I_{2008} = 1$ e che per valori di $t > 2008$ vale

$${}_{2008}I_t = {}_{2008}I_{2009} \cdot {}_{2009}I_{2010} \cdot \dots \cdot {}_{t-1}I_t.$$

Si ottiene quindi

t	$NIBM$	$NIBF (2008)$
2008		1.0000
2009	1.20	1.2000
2010	0.55	0.6600
2011	0.85	0.5610
2012	1.05	0.5891



2. Si calcoli l'indice a base fissa: $_{2010}I_{2012}$ e si commenti il risultato.

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.

Esercizio 5.



2. Si calcoli l'indice a base fissa: ${}_{2010}I_{2012}$ e si commenti il risultato.
Si ricorda che

$${}_{2010}I_{2012} = {}_{2010}I_{2011} \cdot {}_{2011}I_{2012} = 0.8925.$$

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.

Esercizio 5.



2. Si calcoli l'indice a base fissa: ${}_{2010}I_{2012}$ e si commenti il risultato.

Si ricorda che

$${}_{2010}I_{2012} = {}_{2010}I_{2011} \cdot {}_{2011}I_{2012} = 0.8925.$$

Tra il tempo 2010 e il tempo 2012 si è registrato un decremento complessivo del 10.75% della grandezza oggetto di studio.

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.

Esercizio 5.



3. Si calcoli il tasso medio di variazione tra il 2010 e il 2012 e lo si confronti con l'indice calcolato al precedente punto.

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.

Esercizio 5.



3. Si calcoli il tasso medio di variazione tra il 2010 e il 2012 e lo si confronti con l'indice calcolato al precedente punto.

Si ricorda che

$${}_{2010}I_{2012} = {}_{2010}I_{2011} \cdot {}_{2011}I_{2012}.$$

Si deve trovare il numero indice a base mobile medio.

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.

Esercizio 5.



3. Si calcoli il tasso medio di variazione tra il 2010 e il 2012 e lo si confronti con l'indice calcolato al precedente punto.

Si ricorda che

$${}_{2010}I_{2012} = {}_{2010}I_{2011} \cdot {}_{2011}I_{2012}.$$

Si deve trovare il numero indice a base mobile medio. Applicando il criterio di scelta della media secondo Chisini si ottiene

$${}_{2010}I_{2012} = {}_{2010}I_{2011} \cdot {}_{2011}I_{2012} = \alpha \cdot \alpha = \alpha^2.$$

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.

Esercizio 5.



Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.

Esercizio 5.

3. Si calcoli il tasso medio di variazione tra il 2010 e il 2012 e lo si confronti con l'indice calcolato al precedente punto.

Si ricorda che

$${}_{2010}I_{2012} = {}_{2010}I_{2011} \cdot {}_{2011}I_{2012}.$$

Si deve trovare il numero indice a base mobile medio. Applicando il criterio di scelta della media secondo Chisini si ottiene

$${}_{2010}I_{2012} = {}_{2010}I_{2011} \cdot {}_{2011}I_{2012} = \alpha \cdot \alpha = \alpha^2.$$

Quindi

$$\alpha = \sqrt[2]{{}_{2010}I_{2012}} = \sqrt[2]{{}_{2010}I_{2011} \cdot {}_{2011}I_{2012}} = 0.9447.$$

Da cui segue una variazione relativa media pari a -0.0553 , corrispondente a un decremento medio annuo tra il 2010 e il 2012 del 5.53%.



Il seguente prospetto riporta i valori rilevati in corrispondenza a 5 unità statistiche con riferimento a due caratteri quantitativi X e Y

x_i	2	2	4	6	6
y_i	22	26	24	27	21

1. Si rappresentino graficamente le coppie di dati e la funzione di regressione.

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.

Esercizio 5.



La funzione di regressione associa a ogni valore x appartenente all'insieme di definizione della variabile X la media di Y condizionata a x , $M(Y|x)$

x	$M(Y x)$
2	$\frac{1}{2}(22 + 26) = 24$
4	24
6	$\frac{1}{2}(21 + 27) = 24$

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.

Esercizio 5.



Esercizio 1.

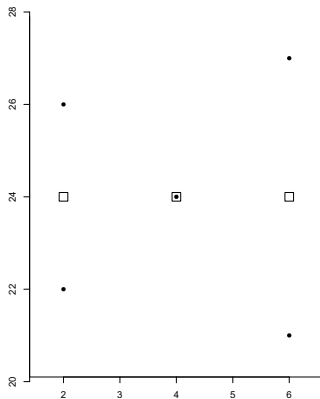
Esercizio 2.

Esercizio 3.

Esercizio 4.

Esercizio 5.

Le coppie di dati possono essere rappresentate mediante un grafico a dispersione.





2. Si calcolino, secondo il metodo dei minimi quadrati, i parametri e l'indice di adattamento dei seguenti modelli di regressione:

$$I) Y^* = a + b \cdot X \qquad II) Y^* = c$$

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.

Esercizio 5.



Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.

Esercizio 5.

Si riportano i calcoli necessari¹ per ottenere le stime dei parametri e gli indici di adattamento del modello

$$Y = a + bX + E$$

x_i	y_i	$x_i y_i$	x_i^2	y_i^2
2	22	44	4	484
2	26	52	4	676
4	24	96	16	576
6	27	162	36	729
6	21	126	36	441
20	120	480	96	2906
4	24	96	19.2	581.2
$M(X)$	$M(Y)$	$M(XY)$	$M(X^2)$	$M(Y^2)$

$$\hat{b} =$$

¹Si osserva che per ragioni di stampa alcuni calcoli potrebbero essere duplicati con riferimento ai diversi modelli oggetto di studio



Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.

Esercizio 5.

Si riportano i calcoli necessari¹ per ottenere le stime dei parametri e gli indici di adattamento del modello

$$Y = a + bX + E$$

x_i	y_i	$x_i y_i$	x_i^2	y_i^2
2	22	44	4	484
2	26	52	4	676
4	24	96	16	576
6	27	162	36	729
6	21	126	36	441
20	120	480	96	2906
4	24	96	19.2	581.2
$M(X)$	$M(Y)$	$M(XY)$	$M(X^2)$	$M(Y^2)$

$$\begin{aligned}\hat{b} &= \frac{\text{Cov}(X, Y)}{\text{Var}(X)} = \frac{M(XY) - M(X)M(Y)}{M(X^2) - [M(X)]^2} = \\ &= \frac{96 - 4 \cdot 24}{19.2 - 4^2} = \frac{0}{3.2} = 0\end{aligned}$$

$$\hat{a} =$$

¹Si osserva che per ragioni di stampa alcuni calcoli potrebbero essere duplicati con riferimento ai diversi modelli oggetto di studio



Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.

Esercizio 5.

Si riportano i calcoli necessari¹ per ottenere le stime dei parametri e gli indici di adattamento del modello

$$Y = a + bX + E$$

x_i	y_i	$x_i y_i$	x_i^2	y_i^2
2	22	44	4	484
2	26	52	4	676
4	24	96	16	576
6	27	162	36	729
6	21	126	36	441
20	120	480	96	2906
4	24	96	19.2	581.2
$M(X)$	$M(Y)$	$M(XY)$	$M(X^2)$	$M(Y^2)$

$$\hat{b} = \frac{\text{Cov}(X, Y)}{\text{Var}(X)} = \frac{M(XY) - M(X)M(Y)}{M(X^2) - [M(X)]^2} =$$

$$= \frac{96 - 4 \cdot 24}{19.2 - 4^2} = \frac{0}{3.2} = 0$$

$$\hat{a} = M(Y) - \hat{b}M(X) = 24$$

$$\rho^2 =$$

¹Si osserva che per ragioni di stampa alcuni calcoli potrebbero essere duplicati con riferimento ai diversi modelli oggetto di studio



Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.

Esercizio 5.

Si riportano i calcoli necessari¹ per ottenere le stime dei parametri e gli indici di adattamento del modello

$$Y = a + bX + E$$

x_i	y_i	$x_i y_i$	x_i^2	y_i^2
2	22	44	4	484
2	26	52	4	676
4	24	96	16	576
6	27	162	36	729
6	21	126	36	441
20	120	480	96	2906
4	24	96	19.2	581.2
$M(X)$	$M(Y)$	$M(XY)$	$M(X^2)$	$M(Y^2)$

$$\hat{b} = \frac{\text{Cov}(X, Y)}{\text{Var}(X)} = \frac{M(XY) - M(X)M(Y)}{M(X^2) - [M(X)]^2} =$$

$$= \frac{96 - 4 \cdot 24}{19.2 - 4^2} = \frac{0}{3.2} = 0$$

$$\hat{a} = M(Y) - \hat{b}M(X) = 24$$

$$\rho^2 = \frac{[\text{Cov}(X, Y)]^2}{\text{Var}(X) \text{Var}(Y)} = \frac{0^2}{3.2 \cdot 5.2} = 0$$

¹Si osserva che per ragioni di stampa alcuni calcoli potrebbero essere duplicati con riferimento ai diversi modelli oggetto di studio



Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.

Esercizio 5.

Si riportano i calcoli necessari² per ottenere le stime dei parametri e gli indici di adattamento del modello

$$Y = c + E$$

x_i	y_i	y_i^2
2	22	484
2	26	676
4	24	576
6	27	729
6	21	441
120		2906
24		581.2
$M(Y)$		$M(Y^2)$

$$\hat{c} =$$

²Si osserva che per ragioni di stampa alcuni calcoli potrebbero essere duplicati con riferimento ai diversi modelli oggetto di studio



Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.

Esercizio 5.

Si riportano i calcoli necessari² per ottenere le stime dei parametri e gli indici di adattamento del modello

$$Y = c + E$$

x_i	y_i	y_i^2
2	22	484
2	26	676
4	24	576
6	27	729
6	21	441
120		2906
24		581.2
$M(Y)$		$M(Y^2)$

$$\hat{c} = M(Y) = 24$$

$$R^2 =$$

²Si osserva che per ragioni di stampa alcuni calcoli potrebbero essere duplicati con riferimento ai diversi modelli oggetto di studio



Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.

Esercizio 5.

Si riportano i calcoli necessari² per ottenere le stime dei parametri e gli indici di adattamento del modello

$$Y = c + E$$

x_i	y_i	y_i^2
2	22	484
2	26	676
4	24	576
6	27	729
6	21	441
120		2906
24		581.2
$M(Y)$		$M(Y^2)$

$$\hat{c} = M(Y) = 24$$

$$R^2 = 0$$

²Si osserva che per ragioni di stampa alcuni calcoli potrebbero essere duplicati con riferimento ai diversi modelli oggetto di studio



I due modelli coincidono.

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.

Esercizio 5.



3. Si fornisca la previsione in corrispondenza di $x = 10$

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.

Esercizio 5.



3. Si fornisca la previsione in corrispondenza di $x = 10$
 $Y^*(10) = 24$.

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.

Esercizio 5.



4. Si determini il valore dei rapporti di correlazione $\eta_{Y|X}^2$ e $\eta_{X|Y}^2$.

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.

Esercizio 5.



4. Si determini il valore dei rapporti di correlazione $\eta^2_{Y|X}$ e $\eta^2_{X|Y}$.

$\eta^2_{Y|X} = 0$ sussiste indipendenza in media in quanto le medie condizionate coincidono.

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.

Esercizio 5.



4. Si determini il valore dei rapporti di correlazione $\eta_{Y|X}^2$ e $\eta_{X|Y}^2$.

$\eta_{Y|X}^2 = 0$ sussiste indipendenza in media in quanto le medie condizionate coincidono.

$\eta_{X|Y}^2 = 1$ sussiste dipendenza funzionale di X da Y (a ogni valore di Y corrisponde un solo valore distinto di X).

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.

Esercizio 5.

Esercizio 4

Prova scritta di
STATISTICA del
05.09.2013

Giuseppe Boari
Gabriele Cantaluppi



UNIVERSITÀ
CATTOLICA
del Sacro Cuore

1. Si completi la seguente tabella, giustificando opportunamente la risposta, in modo tale che $\chi^2 = 44$ e $\rho^2 = 0$

	$x_1 = 12$	$x_2 = \_\_\_\_\_\_$	$x_3 = 18$	
$y_1 = 2$		20		
$y_2 = 4$				
				$n = 44$

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.

Esercizio 5.



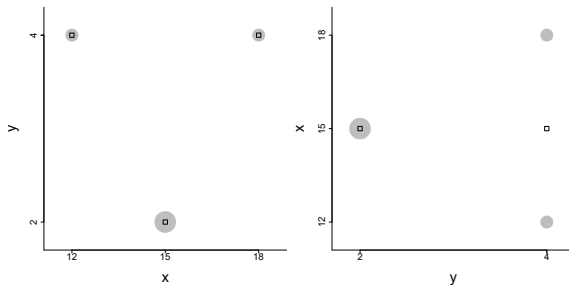
Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.

Esercizio 5.





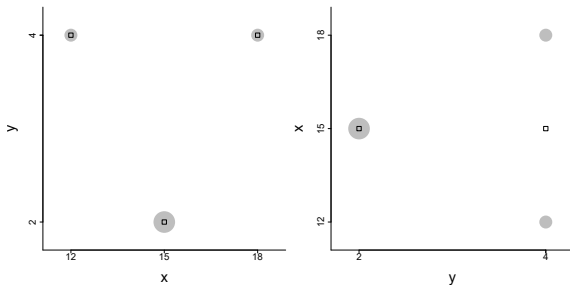
Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.

Esercizio 5.



Dal momento che $\chi^2 = \chi^2_{\max}$, si richiede una situazione di dipendenza funzionale e assenza di correlazione lineare.



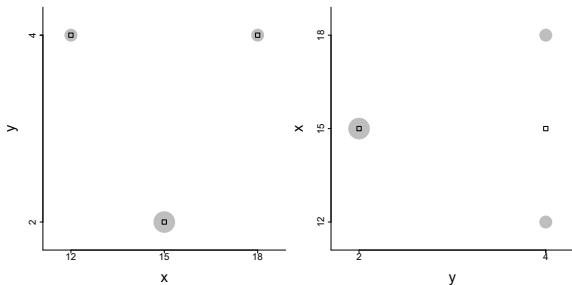
Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.

Esercizio 5.



Dal momento che $\chi^2 = \chi^2_{\max}$, si richiede una situazione di dipendenza funzionale e assenza di correlazione lineare.

Ora, con riferimento alla tabella in oggetto, la situazione può essere realizzata imponendo $\eta^2_{Y|X} = 1$, distribuzioni $Y|X$ degeneri, ed $\eta^2_{X|Y} = 0$, indipendenza in media di Y da X ($\eta^2_{X|Y} = 0 \rightarrow \rho^2 = 0$). $n_{11} = n_{13} = n_{22} = 0$ e $n_{21} = n_{23} = 12$.



2. Si calcoli il valore di $\eta^2_{Y|X}$.

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.

Esercizio 5.



Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.

Esercizio 5.

2. Si calcoli il valore di $\eta_{Y|X}^2$.

$\eta_{Y|X}^2 = 1$ in quanto le distribuzioni condizionate $Y|x_i$ sono degeneri

Esercizio 5

Prova scritta di
STATISTICA del
05.09.2013

Giuseppe Boari
Gabriele Cantaluppi



UNIVERSITÀ
CATTOLICA
del Sacro Cuore

Si consideri l'esperimento di estrazione di 3 palline da un'urna contenente 16 palline rosse, 19 gialle e 5 blu.

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.

Esercizio 5.



1. Si calcoli la probabilità di estrarre zero palline rosse nell'ipotesi che le estrazioni vengano fatte senza reimmissione.

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.

Esercizio 5.



1. Si calcoli la probabilità di estrarre zero palline rosse nell'ipotesi che le estrazioni vengano fatte senza reimmissione.

$$P(0 R) =$$

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.

Esercizio 5.



1. Si calcoli la probabilità di estrarre zero palline rosse nell'ipotesi che le estrazioni vengano fatte senza reimmissione.

$$P(0 R) = P(\bar{R}_1 \cap \bar{R}_2 \cap \bar{R}_3) =$$

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.

Esercizio 5.



Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.

Esercizio 5.

1. Si calcoli la probabilità di estrarre zero palline rosse nell'ipotesi che le estrazioni vengano fatte senza reimmissione.

$$\begin{aligned}P(0 R) &= P(\bar{R}_1 \cap \bar{R}_2 \cap \bar{R}_3) = P(\bar{R}_1) \cdot P(\bar{R}_2 | \bar{R}_1) \cdot P(\bar{R}_3 | \bar{R}_1 \cap \bar{R}_2) \\&= \frac{24}{40} \frac{23}{39} \frac{22}{38} = 0.2049\end{aligned}$$

oppure indicando con X la v.c. ipergeometrica 'n. palline rosse nelle 3 estrazioni senza reimmissione'



Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.

Esercizio 5.

1. Si calcoli la probabilità di estrarre zero palline rosse nell'ipotesi che le estrazioni vengano fatte senza reimmissione.

$$\begin{aligned}P(0 R) &= P(\bar{R}_1 \cap \bar{R}_2 \cap \bar{R}_3) = P(\bar{R}_1) \cdot P(\bar{R}_2 | \bar{R}_1) \cdot P(\bar{R}_3 | \bar{R}_1 \cap \bar{R}_2) \\&= \frac{24}{40} \frac{23}{39} \frac{22}{38} = 0.2049\end{aligned}$$

oppure indicando con X la v.c. ipergeometrica 'n. palline rosse nelle 3 estrazioni senza reimmissione'

$$P(X = 0) = \frac{\binom{24}{3} \binom{16}{0}}{\binom{40}{3}} = 0.2049.$$



Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.

Esercizio 5.

2. Si consideri, ora, l'esperimento di 50 estrazioni con reinserimento. Si calcoli la probabilità di ottenere almeno 31 palline di colore blu.



2. Si consideri, ora, l'esperimento di 50 estrazioni con reinserimento. Si calcoli la probabilità di ottenere almeno 31 palline di colore blu.

Indicata con X la variabile casuale 'numero di palline blu' nelle 50 estrazioni con reimmissione (prove indipendenti) risulta

$$X \sim$$

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.

Esercizio 5.



Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.

Esercizio 5.

2. Si consideri, ora, l'esperimento di 50 estrazioni con reinserimento. Si calcoli la probabilità di ottenere almeno 31 palline di colore blu.

Indicata con X la variabile casuale 'numero di palline blu' nelle 50 estrazioni con reimmissione (prove indipendenti) risulta

$$X \sim \text{Bin}(n = 50, p = 0.125).$$

$$M(X) = 6.25; \text{Var}(X) = 5.4688;$$

utilizzando l'approssimazione della variabile casuale Binomiale con la variabile casuale Normale

$$X \sim$$



2. Si consideri, ora, l'esperimento di 50 estrazioni con reinserimento. Si calcoli la probabilità di ottenere almeno 31 palline di colore blu.

Indicata con X la variabile casuale 'numero di palline blu' nelle 50 estrazioni con reimmissione (prove indipendenti) risulta

$$X \sim \text{Bin}(n = 50, p = 0.125).$$

$$M(X) = 6.25; \text{Var}(X) = 5.4688;$$

utilizzando l'approssimazione della variabile casuale Binomiale con la variabile casuale Normale

$$X \sim N(\mu = 6.25, \sigma^2 = 5.4688)$$

si ottiene

$$\begin{aligned} P(X \geq 31) &= P\left(\frac{X - \mu}{\sigma} \geq \frac{31 - \mu}{\sigma}\right) \\ &= P\left(Z \geq \frac{31 - 6.25}{2.3385}\right) = P(Z \geq 10.58) = 0. \end{aligned}$$

Tema T275-1

Prova scritta di STATISTICA del 16.09.2013

Prova scritta di
STATISTICA del
16.09.2013

Giuseppe Boari
Gabriele Cantaluppi



UNIVERSITÀ
CATTOLICA
del Sacro Cuore

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.

Esercizio 5.

Giuseppe Boari
Gabriele Cantaluppi
Università Cattolica del Sacro Cuore, Milano



Nella tabella sottostante vengono riportati i risultati di un'indagine sull'uso dei mezzi di trasporto da parte dei giovani nella città di Milano.

Mezzo di trasporto (T) \ Età (E)	12 - 16	16 - 20	20 - 30
bicicletta	3	6	9
mezzo pubblico	10	8	9
motorino	3	11	8

1. Si rappresenti graficamente la variabile marginale T .

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.

Esercizio 5.



Si considera la distribuzione marginale della variabile T , ottenuta sommando le frequenze di riga

t_i	$n_{i\bullet}$
<i>bicicletta</i>	18
<i>mezzo pubblico</i>	27
<i>motorino</i>	22

Si tratta di un carattere qualitativo sconnesso che può essere rappresentato graficamente mediante un grafico a bastoncini o a canne d'organo.

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.

Esercizio 5.



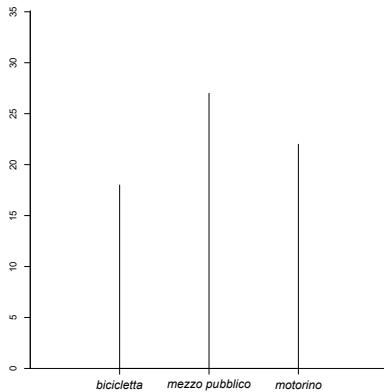
Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.

Esercizio 5.





2. Si confrontino con opportune misure di posizione e variabilità le distribuzioni condizionate $T|E = 12 \div 16$ e $T|E = 20 \div 30$.

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.

Esercizio 5.



Si considera la distribuzione condizionata $T|e_1$

t_i	n_{i1}
<i>bicicletta</i>	3
<i>mezzo pubblico</i>	10
<i>motorino</i>	3

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.

Esercizio 5.



Si tratta di un carattere qualitativo sconnesso.

L'unico indice di posizione che può essere definito è la moda (categoria alla quale è associata la massima frequenza) che risulta *mezzo pubblico*.

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.

Esercizio 5.



In presenza di un carattere qualitativo la variabilità può essere misurata mediante gli indici di Gini e di Frosini normalizzati

$$G_N = \frac{1 - \sum_{i=1}^h f_{i1}^2}{1 - \frac{1}{h}}$$

e

$$F_N = 1 - \sqrt{1 - G_N}.$$

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.

Esercizio 5.



In presenza di un carattere qualitativo la variabilità può essere misurata mediante gli indici di Gini e di Frosini normalizzati

$$G_N = \frac{1 - \sum_{i=1}^h f_{i1}^2}{1 - \frac{1}{h}}$$

e

$$F_N = 1 - \sqrt{1 - G_N}.$$

Si ottiene:

t_i	n_{i1}	f_{i1}	f_{i1}^2
<i>bicicletta</i>	3	0.1875	0.0352
<i>mezzo pubblico</i>	10	0.625	0.3906
<i>motorino</i>	3	0.1875	0.0352
	16	1	0.461

Da cui segue: $G_N = 0.8086$ e $F_N = 0.5625$. Sussiste, quindi, un livello medio di eterogeneità.

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.

Esercizio 5.



Si considera la distribuzione condizionata $T|e_3$

t_i	n_{i3}
<i>bicicletta</i>	9
<i>mezzo pubblico</i>	9
<i>motorino</i>	8

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.

Esercizio 5.



Si tratta di un carattere qualitativo sconnesso.

L'unico indice di posizione che può essere definito è la moda (categoria alla quale è associata la massima frequenza) che risulta *bicicletta*, *mezzo pubblico* (situazione bimodale).

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.

Esercizio 5.



In presenza di un carattere qualitativo la variabilità può essere misurata mediante gli indici di Gini e di Frosini normalizzati

$$G_N = \frac{1 - \sum_{i=1}^h f_{i3}^2}{1 - \frac{1}{h}}$$

e

$$F_N = 1 - \sqrt{1 - G_N}.$$

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.

Esercizio 5.



In presenza di un carattere qualitativo la variabilità può essere misurata mediante gli indici di Gini e di Frosini normalizzati

$$G_N = \frac{1 - \sum_{i=1}^h f_{i3}^2}{1 - \frac{1}{h}}$$

e

$$F_N = 1 - \sqrt{1 - G_N}.$$

Si ottiene:

t_i	n_{i3}	f_{i3}	f_{i3}^2
<i>bicicletta</i>	9	0.3462	0.1198
<i>mezzo pubblico</i>	9	0.3462	0.1198
<i>motorino</i>	8	0.3077	0.0947
	26	1	0.3343

Da cui segue: $G_N = 0.9985$ e $F_N = 0.9615$. Sussiste, quindi, un livello molto alto di eterogeneità.



$T|_{E=12-16}$ Indice di posizione: $\text{Moda}(T|_{E=12-16}) = \text{'mezzo pubblico'}$ Indice di variabilità:
 $F_N = 0.5625$

$T|_{E=20-30}$ Indice di posizione: $\text{Moda}(T|_{E=20-30})$ è bimodale Indice di variabilità: $F_N = 0.9613$
La fascia di età 20 - 30 presenta un livello di eterogeneità nell'utilizzo dei mezzi di trasporto molto più elevato di quello della fascia di età 12 - 16.

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.

Esercizio 5.



3. Si calcoli il rapporto di correlazione $\eta^2_{E|T}$ commentando il risultato.

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.

Esercizio 5.



Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.

Esercizio 5.

3. Si calcoli il rapporto di correlazione $\eta_{E|T}^2$ commentando il risultato.
Occorre, in primo luogo, determinare le medie condizionate $M(E|t_i)$, $i = 1, 2, 3$ e la varianza della variabile dipendente E .



Si considera la distribuzione condizionata $E|t_1$

e_j	n_{1j}
12 $\rightarrow$ 16	3
16 $\rightarrow$ 20	6
20 $\rightarrow$ 30	9

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.

Esercizio 5.



Si tratta di un carattere quantitativo con modalità raggruppate in classi.

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.

Esercizio 5.



La media aritmetica

$$M(E|t_1) = \frac{1}{n_{1\bullet}} \sum_{j=1}^k e_j n_{1j}$$

risulta 20.8333.

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.

Esercizio 5.



La media aritmetica

$$M(E|t_1) = \frac{1}{n_{1\bullet}} \sum_{j=1}^k e_j n_{1j}$$

risulta 20.8333.

Si ottiene (considerando come modalità i valori centrali delle classi)

e_j	n_{1j}	$e_j n_{1j}$
14	3	42
18	6	108
25	9	225
	18	375

da cui segue

$$M(E|t_1) = \frac{1}{n_{1\bullet}} \sum_{j=1}^k e_j n_{1j} = \frac{375}{18} = 20.8333.$$

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.

Esercizio 5.



Si considera la distribuzione condizionata $E|t_2$

e_j	n_{2j}
12 $\dashv$ 16	10
16 $\dashv$ 20	8
20 $\dashv$ 30	9

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.

Esercizio 5.



Si tratta di un carattere quantitativo con modalità raggruppate in classi.

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.

Esercizio 5.



La media aritmetica

$$M(E|t_2) = \frac{1}{n_{2\bullet}} \sum_{j=1}^k e_j n_{2j}$$

risulta 18.8519.

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.

Esercizio 5.



La media aritmetica

$$M(E|t_2) = \frac{1}{n_{2\bullet}} \sum_{j=1}^k e_j n_{2j}$$

risulta 18.8519.

Si ottiene (considerando come modalità i valori centrali delle classi)

e_j	n_{2j}	$e_j n_{2j}$
14	10	140
18	8	144
25	9	225
	27	509

da cui segue

$$M(E|t_2) = \frac{1}{n_{2\bullet}} \sum_{j=1}^k e_j n_{2j} = \frac{509}{27} = 18.8519.$$

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.

Esercizio 5.



Si considera la distribuzione condizionata $E|t_3$

e_j	n_{3j}
12 $\rightarrow$ 16	3
16 $\rightarrow$ 20	11
20 $\rightarrow$ 30	8

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.

Esercizio 5.



Si tratta di un carattere quantitativo con modalità raggruppate in classi.

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.

Esercizio 5.



La media aritmetica

$$M(E|t_3) = \frac{1}{n_{3\bullet}} \sum_{j=1}^k e_j n_{3j}$$

risulta 20.

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.

Esercizio 5.



La media aritmetica

$$M(E|t_3) = \frac{1}{n_{3\bullet}} \sum_{j=1}^k e_j n_{3j}$$

risulta 20.

Si ottiene (considerando come modalità i valori centrali delle classi)

e_j	n_{3j}	$e_j n_{3j}$
14	3	42
18	11	198
25	8	200
	22	440

da cui segue

$$M(E|t_3) = \frac{1}{n_{3\bullet}} \sum_{j=1}^k e_j n_{3j} = \frac{440}{22} = 20.$$

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.

Esercizio 5.



Si considera la distribuzione marginale della variabile E , ottenuta sommando le frequenze di colonna

e_j	$n_{\bullet j}$
12 $\rightarrow$ 16	16
16 $\rightarrow$ 20	25
20 $\rightarrow$ 30	26

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.

Esercizio 5.



Si tratta di un carattere quantitativo con modalità raggruppate in classi.

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.

Esercizio 5.



La media aritmetica

$$M(E) = \frac{1}{n} \sum_{j=1}^k e_j n_{\bullet j}$$

risulta 19.7612.

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.

Esercizio 5.



La media aritmetica

$$M(E) = \frac{1}{n} \sum_{j=1}^k e_j n_{\bullet j}$$

risulta 19.7612.

Si ottiene (considerando come modalità i valori centrali delle classi)

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.

Esercizio 5.



La media aritmetica

$$M(E) = \frac{1}{n} \sum_{j=1}^k e_j n_{\bullet j}$$

risulta 19.7612.

Si ottiene (considerando come modalità i valori centrali delle classi)

e_j	$n_{\bullet j}$	$e_j n_{\bullet j}$
14	16	224
18	25	450
25	26	650
	67	1324

da cui segue

$$M(E) = \frac{1}{n} \sum_{i=1}^k e_j n_{\bullet j} = \frac{1324}{67} = 19.7612.$$

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.

Esercizio 5.



La varianza

$$\text{Var}(E) = M[(E - \mu_E)^2] = M(E^2) - \mu_E^2 = \frac{1}{n} \sum_{j=1}^k e_j^2 n_{\bullet j} - \mu_E^2$$

risulta 19.734.

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.

Esercizio 5.



La varianza

$$\text{Var}(E) = M[(E - \mu_E)^2] = M(E^2) - \mu_E^2 = \frac{1}{n} \sum_{j=1}^k e_j^2 n_{\bullet j} - \mu_E^2$$

risulta 19.734.

Si ottiene (considerando come modalità i valori centrali delle classi)

e_j	$n_{\bullet j}$	$e_j n_{\bullet j}$	$e_j^2 n_{\bullet j}$
14	16	224	3136
18	25	450	8100
25	26	650	16250
	67	1324	27486

da cui segue

$$\begin{aligned} \text{Var}(E) = \sigma_E^2 &= M(E^2) - \mu_E^2 = \\ &= \frac{27486}{67} - 19.7612^2 = 410.2388 - 390.5048 = 19.734. \end{aligned}$$

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.

Esercizio 5.



Il rapporto di correlazione

$$\eta_{E|T}^2 = \frac{\text{Var}_T[M(E|T)]}{\text{Var}(E)}$$

può essere ottenuto considerando il rapporto tra la varianza spiegata $\text{Var}_T[M(E|T)]$, varianza della variabile statistica medie condizionate

i	$M(T e_i)$	$n_{i\bullet}$
1	20.8333	18
2	18.8519	27
3	20.0000	22

e la varianza della variabile dipendente $\text{Var}(E) = 19.734$.

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.

Esercizio 5.



La varianza spiegata risulta $\text{Var}_T[M(E|T)] = 0.6608$.

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.

Esercizio 5.



La varianza spiegata risulta $\text{Var}_T[M(E|T)] = 0.6608$. Si ottiene

i	$M(T e_i)$	$n_{i\bullet}$	$M(T e_i) \cdot n_{i\bullet}$	$[M(T e_i)]^2 \cdot n_{i\bullet}$
1	20.8333	18	375	7812.5000
2	18.8519	27	509	9595.5926
3	20.0000	22	440	8800.0000
		67	1324	26208.0926

da cui segue

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.

Esercizio 5.



La varianza spiegata risulta $Var_T[M(E|T)] = 0.6608$. Si ottiene

i	$M(T e_i)$	$n_{i\bullet}$	$M(T e_i) \cdot n_{i\bullet}$	$[M(T e_i)]^2 \cdot n_{i\bullet}$
1	20.8333	18	375	7812.5000
2	18.8519	27	509	9595.5926
3	20.0000	22	440	8800.0000
		67	1324	26208.0926

da cui segue

$$Var_T[M(E|T)] = \frac{26208.0926}{67} - 19.7612^2 = 391.1656 - 390.5048 = 0.6608.$$

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.

Esercizio 5.



Il rapporto di correlazione $\eta_{E|T}^2$ risulta:

$$\eta_{E|T}^2 = \frac{\text{Var}_T[M(E|T)]}{\text{Var}(E)} = \frac{0.6608}{19.734} = 0.0335.$$

Sussiste, quindi, un livello di dipendenza molto basso.

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.

Esercizio 5.



Si ricorda che, in base al risultato di scomposizione della varianza, la varianza della variabile dipendente può essere ottenuta come somma della varianza spiegata $Var_T[M(E|T)]$ e della varianza residua $M_T[Var(E|T)]$.

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.

Esercizio 5.



Si consideri la seguente serie storica dei costi di un'azienda tra il 2008 e il 2012:

t	x_t
2008	200
2009	240
2010	110
2011	170
2012	210

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.

Esercizio 5.



1. Si costruisca la serie dei numeri indici a base mobile.

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.

Esercizio 5.



1. Si costruisca la serie dei numeri indici a base mobile.

Si ricorda che ${}_{t-1}I_t = \frac{x_t}{x_{t-1}}$.

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.

Esercizio 5.



1. Si costruisca la serie dei numeri indici a base mobile.

Si ricorda che ${}_{t-1}I_t = \frac{x_t}{x_{t-1}}$. Si ottiene quindi

t	x_t	$NIBM$
2008	200	
2009	240	1.2000
2010	110	0.4583
2011	170	1.5455
2012	210	1.2353

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.

Esercizio 5.



2. Si calcoli l'indice a base fissa: ${}_{2012}I_{2010}$ e si commenti il risultato.

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.

Esercizio 5.



2. Si calcoli l'indice a base fissa: ${}_{2012}I_{2010}$ e si commenti il risultato.

$${}_{2012}I_{2010} = \frac{x_{2010}}{x_{2012}} = \frac{110}{210} = 0.5238$$

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.

Esercizio 5.



Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.

Esercizio 5.

2. Si calcoli l'indice a base fissa: ${}_{2012}I_{2010}$ e si commenti il risultato.

$${}_{2012}I_{2010} = \frac{x_{2010}}{x_{2012}} = \frac{110}{210} = 0.5238$$

La grandezza al tempo 2010 era inferiore del 47.62% rispetto alla grandezza al tempo 2012. Ovvero il valore della grandezza nel 2010 era il 52.38% del valore registrato nel 2012.



3. Si calcoli il tasso medio di variazione tra il 2009 e il 2012 commentando il risultato ottenuto.

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.

Esercizio 5.



3. Si calcoli il tasso medio di variazione tra il 2009 e il 2012 commentando il risultato ottenuto.
Si ricorda che

$${}_{2009}I_{2012} = {}_{2009}I_{2010} \cdot {}_{2010}I_{2011} \cdot {}_{2011}I_{2012}$$

Si deve trovare il numero indice a base mobile medio.

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.

Esercizio 5.



Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.

Esercizio 5.

3. Si calcoli il tasso medio di variazione tra il 2009 e il 2012 commentando il risultato ottenuto. Si ricorda che

$${}_{2009}I_{2012} = {}_{2009}I_{2010} \cdot {}_{2010}I_{2011} \cdot {}_{2011}I_{2012}$$

Si deve trovare il numero indice a base mobile medio. Applicando il criterio di scelta della media secondo Chisini si ottiene

$${}_{2009}I_{2012} = {}_{2009}I_{2010} \cdot {}_{2010}I_{2011} \cdot {}_{2011}I_{2012} = \alpha \cdot \alpha \cdot \alpha = \alpha^3.$$



Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.

Esercizio 5.

3. Si calcoli il tasso medio di variazione tra il 2009 e il 2012 commentando il risultato ottenuto. Si ricorda che

$${}_{2009}I_{2012} = {}_{2009}I_{2010} \cdot {}_{2010}I_{2011} \cdot {}_{2011}I_{2012}$$

Si deve trovare il numero indice a base mobile medio. Applicando il criterio di scelta della media secondo Chisini si ottiene

$${}_{2009}I_{2012} = {}_{2009}I_{2010} \cdot {}_{2010}I_{2011} \cdot {}_{2011}I_{2012} = \alpha \cdot \alpha \cdot \alpha = \alpha^3.$$

Quindi

$$\alpha = \sqrt[3]{{}_{2009}I_{2012}} = \sqrt[3]{{}_{2009}I_{2010} \cdot {}_{2010}I_{2011} \cdot {}_{2011}I_{2012}} = 0.9565.$$

Da cui segue una variazione relativa media pari a -0.0435 , corrispondente a un decremento medio annuo tra il 2009 e il 2012 del 4.35%.



4. Sapendo che ${}_{2012}l_{2013} = 1$ si calcolino i costi del 2013.

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.

Esercizio 5.



Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.

Esercizio 5.

4. Sapendo che ${}_{2012}l_{2013} = 1$ si calcolino i costi del 2013.

$x_{2013} = x_{2012} = 210$ dal momento che non vi è stata alcuna variazione tra il 2012 e il 2013.



Il seguente prospetto riporta i valori rilevati in corrispondenza a 5 unità statistiche con riferimento a due caratteri quantitativi X e Y

x_i	2	2	4	6	6
y_i	22	26	25	29	23

1. Si rappresentino graficamente le coppie di dati e la funzione di regressione.

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.

Esercizio 5.



La funzione di regressione associa a ogni valore x appartenente all'insieme di definizione della variabile X la media di Y condizionata a x , $M(Y|x)$

x	$M(Y x)$
2	$\frac{1}{2}(22 + 26) = 24$
4	25
6	$\frac{1}{2}(23 + 29) = 26$

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.

Esercizio 5.



Esercizio 1.

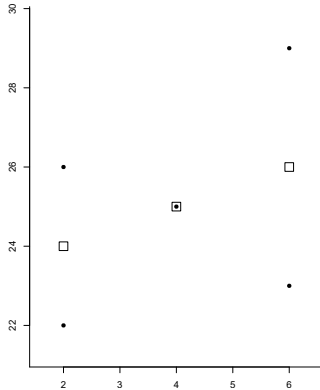
Esercizio 2.

Esercizio 3.

Esercizio 4.

Esercizio 5.

Le coppie di dati possono essere rappresentate mediante un grafico a dispersione.





2. Si calcolino, secondo il metodo dei minimi quadrati, i parametri e l'indice di adattamento dei seguenti modelli di regressione:

$$I) Y^* = a + b \cdot X \qquad II) Y^* = cX$$

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.

Esercizio 5.

Si riportano i calcoli necessari¹ per ottenere le stime dei parametri e gli indici di adattamento del modello

$$Y = a + bX + E$$

x_i	y_i	$x_i y_i$	x_i^2	y_i^2
2	22	44	4	484
2	26	52	4	676
4	25	100	16	625
6	29	174	36	841
6	23	138	36	529
20	125	508	96	3155
4	25	101.6	19.2	631
$M(X)$	$M(Y)$	$M(XY)$	$M(X^2)$	$M(Y^2)$

$$\hat{b} =$$



Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.

Esercizio 5.

¹Si osserva che per ragioni di stampa alcuni calcoli potrebbero essere duplicati con riferimento ai diversi modelli oggetto di studio

Si riportano i calcoli necessari¹ per ottenere le stime dei parametri e gli indici di adattamento del modello

$$Y = a + bX + E$$

x_i	y_i	$x_i y_i$	x_i^2	y_i^2
2	22	44	4	484
2	26	52	4	676
4	25	100	16	625
6	29	174	36	841
6	23	138	36	529
20	125	508	96	3155
4	25	101.6	19.2	631
$M(X)$	$M(Y)$	$M(XY)$	$M(X^2)$	$M(Y^2)$

$$\begin{aligned}\hat{b} &= \frac{\text{Cov}(X, Y)}{\text{Var}(X)} = \frac{M(XY) - M(X)M(Y)}{M(X^2) - [M(X)]^2} = \\ &= \frac{101.6 - 4 \cdot 25}{19.2 - 4^2} = \frac{1.6}{3.2} = 0.5 \\ \hat{a} &= \end{aligned}$$



Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.

Esercizio 5.

¹Si osserva che per ragioni di stampa alcuni calcoli potrebbero essere duplicati con riferimento ai diversi modelli oggetto di studio

Si riportano i calcoli necessari¹ per ottenere le stime dei parametri e gli indici di adattamento del modello

$$Y = a + bX + E$$

x_i	y_i	$x_i y_i$	x_i^2	y_i^2
2	22	44	4	484
2	26	52	4	676
4	25	100	16	625
6	29	174	36	841
6	23	138	36	529
20	125	508	96	3155
4	25	101.6	19.2	631
$M(X)$	$M(Y)$	$M(XY)$	$M(X^2)$	$M(Y^2)$

$$\begin{aligned}\hat{b} &= \frac{\text{Cov}(X, Y)}{\text{Var}(X)} = \frac{M(XY) - M(X)M(Y)}{M(X^2) - [M(X)]^2} = \\ &= \frac{101.6 - 4 \cdot 25}{19.2 - 4^2} = \frac{1.6}{3.2} = 0.5 \\ \hat{a} &= M(Y) - \hat{b}M(X) = 23 \\ \rho^2 &= \end{aligned}$$



Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.

Esercizio 5.

¹Si osserva che per ragioni di stampa alcuni calcoli potrebbero essere duplicati con riferimento ai diversi modelli oggetto di studio

Si riportano i calcoli necessari¹ per ottenere le stime dei parametri e gli indici di adattamento del modello

$$Y = a + bX + E$$

x_i	y_i	$x_i y_i$	x_i^2	y_i^2
2	22	44	4	484
2	26	52	4	676
4	25	100	16	625
6	29	174	36	841
6	23	138	36	529
20	125	508	96	3155
4	25	101.6	19.2	631
$M(X)$	$M(Y)$	$M(XY)$	$M(X^2)$	$M(Y^2)$

$$\begin{aligned}\hat{b} &= \frac{\text{Cov}(X, Y)}{\text{Var}(X)} = \frac{M(XY) - M(X)M(Y)}{M(X^2) - [M(X)]^2} = \\ &= \frac{101.6 - 4 \cdot 25}{19.2 - 4^2} = \frac{1.6}{3.2} = 0.5 \\ \hat{a} &= M(Y) - \hat{b}M(X) = 23 \\ \rho^2 &= \frac{[\text{Cov}(X, Y)]^2}{\text{Var}(X) \text{Var}(Y)} = \frac{1.6^2}{3.2 \cdot 6} = 0.1333\end{aligned}$$

$$M(\hat{E}^2) = \text{Var}(\hat{E}) = \text{varianza residua} =$$

¹Si osserva che per ragioni di stampa alcuni calcoli potrebbero essere duplicati con riferimento ai diversi modelli oggetto di studio



Si riportano i calcoli necessari¹ per ottenere le stime dei parametri e gli indici di adattamento del modello

$$Y = a + bX + E$$

x_i	y_i	$x_i y_i$	x_i^2	y_i^2
2	22	44	4	484
2	26	52	4	676
4	25	100	16	625
6	29	174	36	841
6	23	138	36	529
20	125	508	96	3155
4	25	101.6	19.2	631
$M(X)$	$M(Y)$	$M(XY)$	$M(X^2)$	$M(Y^2)$

$$\hat{b} = \frac{\text{Cov}(X, Y)}{\text{Var}(X)} = \frac{M(XY) - M(X)M(Y)}{M(X^2) - [M(X)]^2} =$$

$$= \frac{101.6 - 4 \cdot 25}{19.2 - 4^2} = \frac{1.6}{3.2} = 0.5$$

$$\hat{a} = M(Y) - \hat{b}M(X) = 23$$

$$\rho^2 = \frac{[\text{Cov}(X, Y)]^2}{\text{Var}(X) \text{Var}(Y)} = \frac{1.6^2}{3.2 \cdot 6} = 0.1333$$

$$M(\hat{E}^2) = \text{Var}(\hat{E}) = \text{varianza residua} = \text{Var}(Y) \cdot (1 - \rho^2) = 5.2$$

¹Si osserva che per ragioni di stampa alcuni calcoli potrebbero essere duplicati con riferimento ai diversi modelli oggetto di studio





Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.

Esercizio 5.

Si riportano i calcoli necessari² per ottenere le stime dei parametri e la misura di adattamento del modello

$$Y = cX + E$$

x_i	y_i	$x_i y_i$	x_i^2	y_i^2
2	22	44	4	484
2	26	52	4	676
4	25	100	16	625
6	29	174	36	841
6	23	138	36	529
		508	96	3155
		101.6	19.2	631
		$M(XY)$	$M(X^2)$	$M(Y^2)$

$$\hat{c} =$$

²Si osserva che per ragioni di stampa alcuni calcoli potrebbero essere duplicati con riferimento ai diversi modelli oggetto di studio



Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.

Esercizio 5.

Si riportano i calcoli necessari² per ottenere le stime dei parametri e la misura di adattamento del modello

$$Y = cX + E$$

x_i	y_i	$x_i y_i$	x_i^2	y_i^2
2	22	44	4	484
2	26	52	4	676
4	25	100	16	625
6	29	174	36	841
6	23	138	36	529
		508	96	3155
		101.6	19.2	631
		$M(XY)$	$M(X^2)$	$M(Y^2)$

$$\hat{c} = \frac{M(XY)}{M(X^2)} = \frac{101.6}{19.2} = 5.2917$$

$$M(\hat{E}^2) =$$

²Si osserva che per ragioni di stampa alcuni calcoli potrebbero essere duplicati con riferimento ai diversi modelli oggetto di studio



Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.

Esercizio 5.

Si riportano i calcoli necessari² per ottenere le stime dei parametri e la misura di adattamento del modello

$$Y = cX + E$$

x_i	y_i	$x_i y_i$	x_i^2	y_i^2
2	22	44	4	484
2	26	52	4	676
4	25	100	16	625
6	29	174	36	841
6	23	138	36	529
		508	96	3155
		101.6	19.2	631
		$M(XY)$	$M(X^2)$	$M(Y^2)$

$$\hat{c} = \frac{M(XY)}{M(X^2)} = \frac{101.6}{19.2} = 5.2917$$

$$M(\hat{E}^2) = M[Y \cdot (Y - \hat{c}X)] = M(Y^2) - \hat{c}M(XY) = 631 - 5.2917 \cdot 101.6 = 93.3667$$

²Si osserva che per ragioni di stampa alcuni calcoli potrebbero essere duplicati con riferimento ai diversi modelli oggetto di studio



Il modello migliore è $Y^* = 23 + 0.5X$.

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.

Esercizio 5.



3. Si fornisca la previsione in corrispondenza di $x = 10$

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.

Esercizio 5.



3. Si fornisca la previsione in corrispondenza di $x = 10$
 $Y^*(10) = 28$

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.

Esercizio 5.

Esercizio 4

Prova scritta di
STATISTICA del
16.09.2013

Giuseppe Boari
Gabriele Cantaluppi



UNIVERSITÀ
CATTOLICA
del Sacro Cuore

1. Si completi la seguente tabella, giustificando opportunamente la risposta, in modo tale che $\chi^2 = 88$ e $\rho^2 = 1$

	$x_1 = 12$	$x_2 = 15$	$x_3 = 18$	
$y_1 = 2$				
$y_2 = 4$		20		
$y_2 = 5$				
				$n = 44$

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.

Esercizio 5.



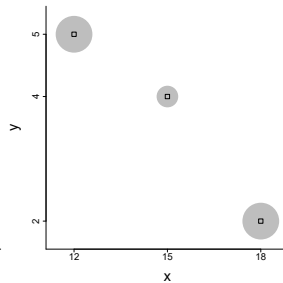
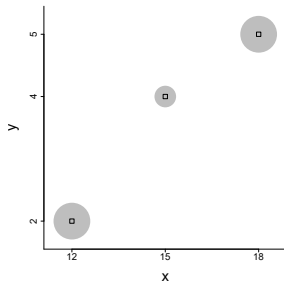
Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.

Esercizio 5.





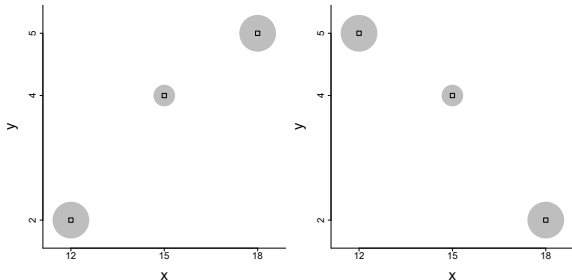
Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.

Esercizio 5.



Impossibile, in quanto $\rho^2 = 1$ implica una situazione di dipendenza funzionale lineare, compatibile con $\chi^2 = 88$ ma che non può realizzarsi in quanto non è possibile disporre le medie condizionate su una retta.

Esercizio 5

Prova scritta di
STATISTICA del
16.09.2013

Giuseppe Boari
Gabriele Cantaluppi



UNIVERSITÀ
CATTOLICA
del Sacro Cuore

Si consideri l'esperimento di estrazione di 9 palline da un'urna contenente 16 palline rosse, 19 gialle e 5 blu.

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.

Esercizio 5.



1. Si calcoli la probabilità di estrarre 2 palline rosse, 1 gialla e 6 blu nell'ipotesi che le estrazioni vengano fatte senza reimmissione.

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.

Esercizio 5.



1. Si calcoli la probabilità di estrarre 2 palline rosse, 1 gialla e 6 blu nell'ipotesi che le estrazioni vengano fatte senza reimmissione.

$P(2R \cap 1G \cap 6B) = 0$ in quanto non è possibile estrarre 6 palline blu se le estrazioni vengono effettuate senza reimmissione; l'evento è quindi impossibile.

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.

Esercizio 5.



2. Si calcoli la probabilità di estrarre 3 palline rosse, 4 gialle e 2 blu nell'ipotesi che le estrazioni vengano fatte con reimmissione.

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.

Esercizio 5.



2. Si calcoli la probabilità di estrarre 3 palline rosse, 4 gialle e 2 blu nell'ipotesi che le estrazioni vengano fatte con reimmissione.

$$P(3R \cap 4G \cap 2B) =$$

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.

Esercizio 5.



2. Si calcoli la probabilità di estrarre 3 palline rosse, 4 gialle e 2 blu nell'ipotesi che le estrazioni vengano fatte con reimmissione.

$$P(3R \cap 4G \cap 2B) = \text{n. sequenze} \cdot P(\text{generica sequenza}) =$$
$$=$$

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.

Esercizio 5.



2. Si calcoli la probabilità di estrarre 3 palline rosse, 4 gialle e 2 blu nell'ipotesi che le estrazioni vengano fatte con reimmissione.

$$\begin{aligned} P(3R \cap 4G \cap 2B) &= \text{n. sequenze} \cdot P(\text{generica sequenza}) = \\ &= \frac{9!}{3! \cdot 4! \cdot 2!} \left(\frac{16}{40}\right)^3 \left(\frac{19}{40}\right)^4 \left(\frac{5}{40}\right)^2 = 0.0641. \end{aligned}$$

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.

Esercizio 5.



3. Si calcoli la probabilità di ottenere almeno 12 palline di colore blu in 100 estrazioni con reimmissione.

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.

Esercizio 5.



3. Si calcoli la probabilità di ottenere almeno 12 palline di colore blu in 100 estrazioni con reimmissione.

Indicata con X la variabile casuale 'numero di palline blu' nelle 100 estrazioni con reimmissione (prove indipendenti) risulta

$$X \sim$$

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.

Esercizio 5.



Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.

Esercizio 5.

3. Si calcoli la probabilità di ottenere almeno 12 palline di colore blu in 100 estrazioni con reimmissione.

Indicata con X la variabile casuale 'numero di palline blu' nelle 100 estrazioni con reimmissione (prove indipendenti) risulta

$$X \sim \text{Bin}(n = 100, p = 0.125).$$

$$M(X) = 12.5; \text{Var}(X) = 10.9375;$$

utilizzando l'approssimazione della variabile casuale Binomiale con la variabile casuale Normale

$$X \sim$$



3. Si calcoli la probabilità di ottenere almeno 12 palline di colore blu in 100 estrazioni con reimmissione.

Indicata con X la variabile casuale 'numero di palline blu' nelle 100 estrazioni con reimmissione (prove indipendenti) risulta

$$X \sim \text{Bin}(n = 100, p = 0.125).$$

$$M(X) = 12.5; \text{Var}(X) = 10.9375;$$

utilizzando l'approssimazione della variabile casuale Binomiale con la variabile casuale Normale

$$X \sim N(\mu = 12.5, \sigma^2 = 10.9375)$$

si ottiene

$$\begin{aligned} P(X \geq 12) &= P\left(\frac{X - \mu}{\sigma} \geq \frac{12 - \mu}{\sigma}\right) = P\left(Z \geq \frac{12 - 12.5}{3.3072}\right) = \\ &= P(Z \geq -0.15) = P(Z \leq 0.15) = 0.5596. \end{aligned}$$

Tema T276

Prova scritta di STATISTICA del 16.01.2014

Prova scritta di
STATISTICA del
16.01.2014

Giuseppe Boari
Gabriele Cantaluppi



UNIVERSITÀ
CATTOLICA
del Sacro Cuore

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.

Esercizio 5.

Giuseppe Boari
Gabriele Cantaluppi
Università Cattolica del Sacro Cuore, Milano



La tabella seguente riporta le misure delle altezze al garrese (H) di 20 renne adulte, distinte per genere ($G = m, f$).

m	m	m	m	m	m	m	m	m	m	f	f	f	f	f	f	f	f	f	f
158	162	165	170	180	182	184	184	188	190	153	155	158	160	160	168	170	180	180	182

1. Si costruisca una tabella a doppia entrata, usando per le altezze le seguenti classi (150 ÷ 160, 160 ÷ 180, 180 ÷ 190).

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.

Esercizio 5.



Occorre effettuare lo spoglio dei dati (ricordando che le classi sono chiuse a destra)

	150 – 160	160 – 180	180 – 190
<i>m</i>	1	4	5
<i>f</i>	5	4	1

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.

Esercizio 5.



2. Si rappresenti graficamente la variabile marginale G .

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.

Esercizio 5.



Si considera la distribuzione marginale della variabile G , ottenuta sommando le frequenze di riga

g_i	$n_{i\bullet}$
m	10
f	10

Si tratta di un carattere qualitativo sconnesso che può essere rappresentato graficamente mediante un grafico a bastoncini o a canne d'organo.

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.

Esercizio 5.



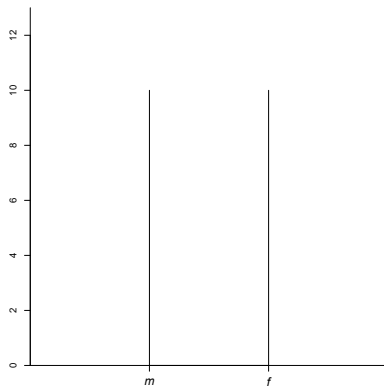
Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.

Esercizio 5.





3. Si individui il valore della mediana della marginale H .

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.

Esercizio 5.



Si considera la distribuzione marginale della variabile G , ottenuta sommando le frequenze di riga

g_i	$n_{i\bullet}$
150 + 160	6
160 + 180	8
180 + 190	6

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.

Esercizio 5.



Si tratta di un carattere quantitativo con modalità raggruppate in classi.

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.

Esercizio 5.



Si osserva che la distribuzione è simmetrica rispetto a 170; la mediana (percentile di ordine 0.50) risulta $g_{0.5} = 170$.

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.

Esercizio 5.



4. Si fornisca un indice di variabilità di G .

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.

Esercizio 5.



La mutabile G presenta massima eterogeneità, quindi gli indici di Gini e di Frosini assumono valore massimo, $F_N = G_N = 1$.

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.

Esercizio 5.



5. Si calcoli un opportuno indice di connessione tra H e G .

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.

Esercizio 5.



Dalla distribuzione delle frequenze congiunte n_{ij} :

	150 ÷ 160	160 ÷ 180	180 ÷ 190	$n_{i\bullet}$
m	1	4	5	10
f	5	4	1	10
$n_{\bullet j}$	6	8	6	20

è possibile ottenere le frequenze teoriche nella situazione di indipendenza stocastica

$$\hat{n}_{ij} = \frac{n_{i\bullet} n_{\bullet j}}{n}:$$

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.

Esercizio 5.



Dalla distribuzione delle frequenze congiunte n_{ij} :

	150 ÷ 160	160 ÷ 180	180 ÷ 190	$n_{i\bullet}$
m	1	4	5	10
f	5	4	1	10
$n_{\bullet j}$	6	8	6	20

è possibile ottenere le frequenze teoriche nella situazione di indipendenza stocastica

$$\hat{n}_{ij} = \frac{n_{i\bullet} \cdot n_{\bullet j}}{n}:$$

	150 ÷ 160	160 ÷ 180	180 ÷ 190	$n_{i\bullet}$
m	3	4	3	10
f	3	4	3	10
$n_{\bullet j}$	6	8	6	20

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.

Esercizio 5.



le contingenze assolute $n_{ij} - \hat{n}_{ij}$:

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.

Esercizio 5.



le contingenze assolute $n_{ij} - \hat{n}_{ij}$:

	150 ÷ 160	160 ÷ 180	180 ÷ 190
m	-2	0	2
f	2	0	-2

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.

Esercizio 5.



i rapporti di contingenza $\frac{n_{ij}}{\hat{n}_{ij}}$:

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.

Esercizio 5.



i rapporti di contingenza $\frac{n_{ij}}{\hat{n}_{ij}}$:

	150 + 160	160 + 180	180 + 190
<i>m</i>	0.3333	1	1.6667
<i>f</i>	1.6667	1	0.3333

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.

Esercizio 5.



e le contingenze relative, espresse in termini percentuali, $100 \left(\frac{n_{ij} - \hat{n}_{ij}}{\hat{n}_{ij}} \right) \% = 100 \left(\frac{n_{ij}}{\hat{n}_{ij}} - 1 \right) \%$:

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.

Esercizio 5.



e le contingenze relative, espresse in termini percentuali, $100 \left(\frac{n_{ij} - \hat{n}_{ij}}{\hat{n}_{ij}} \right) \% = 100 \left(\frac{n_{ij}}{\hat{n}_{ij}} - 1 \right) \%$:

	150 ÷ 160	160 ÷ 180	180 ÷ 190
<i>m</i>	-66.67	0	66.67
<i>f</i>	66.67	0	-66.67

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.

Esercizio 5.



Risultati intermedi $\frac{(n_{ij} - \hat{n}_{ij})^2}{\hat{n}_{ij}}$ per il calcolo del χ^2

	150 + 160	160 + 180	180 + 190
<i>m</i>	1.3333	0	1.3333
<i>f</i>	1.3333	0	1.3333

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.

Esercizio 5.



Risultati intermedi $\frac{n_{ij}^2}{n_{i\bullet} \cdot n_{\bullet j}}$ per il calcolo del χ^2 (formula operativa)

	150 + 160	160 + 180	180 + 190
<i>m</i>	0.0167	0.2	0.4167
<i>f</i>	0.4167	0.2	0.0167

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.

Esercizio 5.



Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.

Esercizio 5.

$$\chi^2 = \sum_{i=1}^h \sum_{j=1}^k \frac{(n_{ij} - \hat{n}_{ij})^2}{\hat{n}_{ij}} = 5.3333,$$

$$\chi_N^2 = \frac{\chi^2}{n \cdot \min(h-1, k-1)} = 0.2667$$

$$\chi_N = \sqrt{\chi_N^2} = 0.5164.$$

Sussiste, quindi, un livello di connessione medio.



Con riferimento alla seguente serie di Numeri Indice a Base Mobile:

t	$NIBM$
2008	
2009	0.92
2010	0.95
2011	1.05
2012	1.08

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.

Esercizio 5.



1. Si ricostruisca la serie storica dei dati x_t sapendo che $x_{2010} = 90$.

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.

Esercizio 5.



1. Si ricostruisca la serie storica dei dati x_t sapendo che $x_{2010} = 90$.
Si ricorda che $x_{2010} = 90$, che per valori di $t > 2010$ vale

$$x_t = x_{2010} \cdot {}_{2010}I_{2011} \cdot {}_{2011}I_{2012} \cdot \dots \cdot {}_{t-1}I_t$$

e che per valori di $t < 2010$ vale

$$x_t = \frac{x_{2010}}{{}_tI_{t+1} \cdot \dots \cdot {}_{2008}I_{2009} \cdot {}_{2009}I_{2010}}.$$

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.

Esercizio 5.



1. Si ricostruisca la serie storica dei dati x_t sapendo che $x_{2010} = 90$.

Si ricorda che $x_{2010} = 90$, che per valori di $t > 2010$ vale

$$x_t = x_{2010} \cdot {}_{2010}l_{2011} \cdot {}_{2011}l_{2012} \cdot \dots \cdot {}_{t-1}l_t$$

e che per valori di $t < 2010$ vale

$$x_t = \frac{x_{2010}}{{}_t l_{t+1} \cdot \dots \cdot {}_{2008}l_{2009} \cdot {}_{2009}l_{2010}}.$$

Si ottiene quindi

t	$NIBM$	x_t
2008		102.9748
2009	0.92	94.7368
2010	0.95	90.0000
2011	1.05	94.5000
2012	1.08	102.0600

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.

Esercizio 5.



2. Si calcoli il tasso medio di variazione nel periodo 2008-2012 interpretando il risultato ottenuto.

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.

Esercizio 5.



2. Si calcoli il tasso medio di variazione nel periodo 2008-2012 interpretando il risultato ottenuto.

Si ricorda che

$${}_{2008}I_{2012} = {}_{2008}I_{2009} \cdot {}_{2009}I_{2010} \cdot {}_{2010}I_{2011} \cdot {}_{2011}I_{2012}.$$

Si deve trovare il numero indice a base mobile medio.

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.

Esercizio 5.



2. Si calcoli il tasso medio di variazione nel periodo 2008-2012 interpretando il risultato ottenuto.

Si ricorda che

$${}_{2008}I_{2012} = {}_{2008}I_{2009} \cdot {}_{2009}I_{2010} \cdot {}_{2010}I_{2011} \cdot {}_{2011}I_{2012}.$$

Si deve trovare il numero indice a base mobile medio. Applicando il criterio di scelta della media secondo Chisini si ottiene

$${}_{2008}I_{2012} = {}_{2008}I_{2009} \cdot {}_{2009}I_{2010} \cdot {}_{2010}I_{2011} \cdot {}_{2011}I_{2012} = \alpha \cdot \alpha \cdot \alpha \cdot \alpha = \alpha^4.$$

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.

Esercizio 5.



Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.

Esercizio 5.

2. Si calcoli il tasso medio di variazione nel periodo 2008-2012 interpretando il risultato ottenuto.

Si ricorda che

$${}_{2008}I_{2012} = {}_{2008}I_{2009} \cdot {}_{2009}I_{2010} \cdot {}_{2010}I_{2011} \cdot {}_{2011}I_{2012}.$$

Si deve trovare il numero indice a base mobile medio. Applicando il criterio di scelta della media secondo Chisini si ottiene

$${}_{2008}I_{2012} = {}_{2008}I_{2009} \cdot {}_{2009}I_{2010} \cdot {}_{2010}I_{2011} \cdot {}_{2011}I_{2012} = \alpha \cdot \alpha \cdot \alpha \cdot \alpha = \alpha^4.$$

Quindi

$$\alpha = \sqrt[4]{{}_{2008}I_{2012}} = \sqrt[4]{{}_{2008}I_{2009} \cdot {}_{2009}I_{2010} \cdot {}_{2010}I_{2011} \cdot {}_{2011}I_{2012}} = 0.9978.$$

Da cui segue una variazione relativa media pari a -0.0022 , corrispondente a un decremento medio annuo tra il 2008 e il 2012 del 0.22%.



3. Si calcoli il numero indice al tempo 2012 con base 2010 e si commenti il risultato.

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.

Esercizio 5.



3. Si calcoli il numero indice al tempo 2012 con base 2010 e si commenti il risultato.

$${}_{2010}I_{2012} = \frac{x_{2012}}{x_{2010}} = \frac{102.06}{90} = 1.134$$

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.

Esercizio 5.



3. Si calcoli il numero indice al tempo 2012 con base 2010 e si commenti il risultato.

$${}_{2010}I_{2012} = \frac{x_{2012}}{x_{2010}} = \frac{102.06}{90} = 1.134$$

Tra il tempo 2010 e il tempo 2012 si è registrato un incremento complessivo del 13.4% della grandezza oggetto di studio.

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.

Esercizio 5.



Il seguente prospetto riporta i valori rilevati in corrispondenza a 8 unità statistiche con riferimento a due caratteri quantitativi X e Y

x_i	1	1	2	2	3	4	4	5
y_i	10	10	13	13	18	20	20	20

1. Si rappresenti graficamente Y in funzione di X .

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.

Esercizio 5.



Esercizio 1.

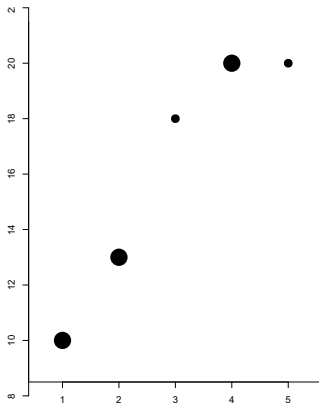
Esercizio 2.

Esercizio 3.

Esercizio 4.

Esercizio 5.

Le coppie di dati possono essere rappresentate mediante un *bubble plot diagram*.





2. Si calcolino, secondo il metodo dei minimi quadrati, i parametri e l'indice di adattamento dei seguenti modelli di regressione:

$$I) Y^* = a + b \cdot X \qquad II) Y^* = \alpha + \beta \sqrt{X}$$

[Esercizio 1.](#)

[Esercizio 2.](#)

[Esercizio 3.](#)

[Esercizio 4.](#)

[Esercizio 5.](#)

Si riportano i calcoli necessari¹ per ottenere le stime dei parametri e gli indici di adattamento del modello

$$Y = a + bX + E$$

x_i	y_i	$x_i y_i$	x_i^2	y_i^2
1	10	10	1	100
1	10	10	1	100
2	13	26	4	169
2	13	26	4	169
3	18	54	9	324
4	20	80	16	400
4	20	80	16	400
5	20	100	25	400
22	124	386	76	2062
2.75	15.5	48.25	9.5	257.75
$M(X)$	$M(Y)$	$M(XY)$	$M(X^2)$	$M(Y^2)$

$$\hat{b} =$$

¹Si osserva che per ragioni di stampa alcuni calcoli potrebbero essere duplicati con riferimento ai diversi modelli oggetto di studio



Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.

Esercizio 5.

Si riportano i calcoli necessari¹ per ottenere le stime dei parametri e gli indici di adattamento del modello

$$Y = a + bX + E$$

x_i	y_i	$x_i y_i$	x_i^2	y_i^2
1	10	10	1	100
1	10	10	1	100
2	13	26	4	169
2	13	26	4	169
3	18	54	9	324
4	20	80	16	400
4	20	80	16	400
5	20	100	25	400
22	124	386	76	2062
2.75	15.5	48.25	9.5	257.75
$M(X)$	$M(Y)$	$M(XY)$	$M(X^2)$	$M(Y^2)$

$$\begin{aligned}\hat{b} &= \frac{\text{Cov}(X, Y)}{\text{Var}(X)} = \frac{M(XY) - M(X)M(Y)}{M(X^2) - [M(X)]^2} = \\ &= \frac{48.25 - 2.75 \cdot 15.5}{9.5 - 2.75^2} = \frac{5.625}{1.9375} = 2.9032 \\ \hat{a} &= \end{aligned}$$

¹Si osserva che per ragioni di stampa alcuni calcoli potrebbero essere duplicati con riferimento ai diversi modelli oggetto di studio



Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.

Esercizio 5.

Si riportano i calcoli necessari¹ per ottenere le stime dei parametri e gli indici di adattamento del modello

$$Y = a + bX + E$$

x_i	y_i	$x_i y_i$	x_i^2	y_i^2
1	10	10	1	100
1	10	10	1	100
2	13	26	4	169
2	13	26	4	169
3	18	54	9	324
4	20	80	16	400
4	20	80	16	400
5	20	100	25	400
22	124	386	76	2062
2.75	15.5	48.25	9.5	257.75
$M(X)$	$M(Y)$	$M(XY)$	$M(X^2)$	$M(Y^2)$

$$\begin{aligned}\hat{b} &= \frac{\text{Cov}(X, Y)}{\text{Var}(X)} = \frac{M(XY) - M(X)M(Y)}{M(X^2) - [M(X)]^2} = \\ &= \frac{48.25 - 2.75 \cdot 15.5}{9.5 - 2.75^2} = \frac{5.625}{1.9375} = 2.9032 \\ \hat{a} &= M(Y) - \hat{b} M(X) = 7.5161 \\ \rho^2 &= \end{aligned}$$

¹Si osserva che per ragioni di stampa alcuni calcoli potrebbero essere duplicati con riferimento ai diversi modelli oggetto di studio



Si riportano i calcoli necessari¹ per ottenere le stime dei parametri e gli indici di adattamento del modello

$$Y = a + bX + E$$

x_i	y_i	$x_i y_i$	x_i^2	y_i^2
1	10	10	1	100
1	10	10	1	100
2	13	26	4	169
2	13	26	4	169
3	18	54	9	324
4	20	80	16	400
4	20	80	16	400
5	20	100	25	400
22	124	386	76	2062
2.75	15.5	48.25	9.5	257.75
$M(X)$	$M(Y)$	$M(XY)$	$M(X^2)$	$M(Y^2)$

$$\hat{b} = \frac{\text{Cov}(X, Y)}{\text{Var}(X)} = \frac{M(XY) - M(X)M(Y)}{M(X^2) - [M(X)]^2} =$$

$$= \frac{48.25 - 2.75 \cdot 15.5}{9.5 - 2.75^2} = \frac{5.625}{1.9375} = 2.9032$$

$$\hat{a} = M(Y) - \hat{b} M(X) = 7.5161$$

$$\rho^2 = \frac{[\text{Cov}(X, Y)]^2}{\text{Var}(X) \text{Var}(Y)} = \frac{5.625^2}{1.9375 \cdot 17.5} = 0.9332$$

¹Si osserva che per ragioni di stampa alcuni calcoli potrebbero essere duplicati con riferimento ai diversi modelli oggetto di studio



Si riportano i calcoli necessari² per ottenere le stime dei parametri e gli indici di adattamento del modello

$$Y = \alpha + \beta X^{0.5} + E$$

$x_i^{0.5}$	y_i	$x_i^{0.5} y_i$	$x_i^{0.5^2}$	y_i^2
1	10	10	1	100
1	10	10	1	100
1.4142	13	18.3848	2	169
1.4142	13	18.3848	2	169
1.7321	18	31.1769	3	324
2	20	40	4	400
2	20	40	4	400
2.2361	20	44.7214	5	400
12.7965	124	212.6678	22	2062
1.5996	15.5	26.5835	2.75	257.75
$M(X^{0.5})$	$M(Y)$	$M(X^{0.5} Y)$	$M(X^{0.5^2})$	$M(Y^2)$

$$\hat{\beta} =$$



Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.

Esercizio 5.

²Si osserva che per ragioni di stampa alcuni calcoli potrebbero essere duplicati con riferimento ai diversi modelli oggetto di studio

Si riportano i calcoli necessari² per ottenere le stime dei parametri e gli indici di adattamento del modello

$$Y = \alpha + \beta X^{0.5} + E$$

$x_i^{0.5}$	y_i	$x_i^{0.5} y_i$	$x_i^{0.5^2}$	y_i^2
1	10	10	1	100
1	10	10	1	100
1.4142	13	18.3848	2	169
1.4142	13	18.3848	2	169
1.7321	18	31.1769	3	324
2	20	40	4	400
2	20	40	4	400
2.2361	20	44.7214	5	400
12.7965	124	212.6678	22	2062
1.5996	15.5	26.5835	2.75	257.75
$M(X^{0.5})$	$M(Y)$	$M(X^{0.5} Y)$	$M(X^{0.5^2})$	$M(Y^2)$

$$\begin{aligned}\hat{\beta} &= \frac{\text{Cov}(X^{0.5}, Y)}{\text{Var}(X^{0.5})} = \frac{M(X^{0.5} Y) - M(X^{0.5}) M(Y)}{M(X^{0.5^2}) - [M(X^{0.5})^2]} = \\ &= \frac{26.5835 - 1.5996 \cdot 15.5}{2.75 - 1.5996^2} = \frac{1.7902}{0.1914} = 9.3532 \\ \hat{\alpha} &= \end{aligned}$$



Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.

Esercizio 5.

²Si osserva che per ragioni di stampa alcuni calcoli potrebbero essere duplicati con riferimento ai diversi modelli oggetto di studio

Si riportano i calcoli necessari² per ottenere le stime dei parametri e gli indici di adattamento del modello

$$Y = \alpha + \beta X^{0.5} + E$$

$x_i^{0.5}$	y_i	$x_i^{0.5} y_i$	$x_i^{0.5^2}$	y_i^2
1	10	10	1	100
1	10	10	1	100
1.4142	13	18.3848	2	169
1.4142	13	18.3848	2	169
1.7321	18	31.1769	3	324
2	20	40	4	400
2	20	40	4	400
2.2361	20	44.7214	5	400
12.7965	124	212.6678	22	2062
1.5996	15.5	26.5835	2.75	257.75
$M(X^{0.5})$	$M(Y)$	$M(X^{0.5} Y)$	$M(X^{0.5^2})$	$M(Y^2)$

$$\begin{aligned}\hat{\beta} &= \frac{\text{Cov}(X^{0.5}, Y)}{\text{Var}(X^{0.5})} = \frac{M(X^{0.5} Y) - M(X^{0.5}) M(Y)}{M(X^{0.5^2}) - [M(X^{0.5})^2]} = \\ &= \frac{26.5835 - 1.5996 \cdot 15.5}{2.75 - 1.5996^2} = \frac{1.7902}{0.1914} = 9.3532 \\ \hat{\alpha} &= M(Y) - \hat{\beta} M(X^{0.5}) = 0.5389 \\ \rho^2 &= \end{aligned}$$

²Si osserva che per ragioni di stampa alcuni calcoli potrebbero essere duplicati con riferimento ai diversi modelli oggetto di studio



Si riportano i calcoli necessari² per ottenere le stime dei parametri e gli indici di adattamento del modello

$$Y = \alpha + \beta X^{0.5} + E$$

$x_i^{0.5}$	y_i	$x_i^{0.5} y_i$	$x_i^{0.5^2}$	y_i^2
1	10	10	1	100
1	10	10	1	100
1.4142	13	18.3848	2	169
1.4142	13	18.3848	2	169
1.7321	18	31.1769	3	324
2	20	40	4	400
2	20	40	4	400
2.2361	20	44.7214	5	400
12.7965	124	212.6678	22	2062
1.5996	15.5	26.5835	2.75	257.75
$M(X^{0.5})$	$M(Y)$	$M(X^{0.5} Y)$	$M(X^{0.5^2})$	$M(Y^2)$

$$\begin{aligned}\hat{\beta} &= \frac{\text{Cov}(X^{0.5}, Y)}{\text{Var}(X^{0.5})} = \frac{M(X^{0.5} Y) - M(X^{0.5}) M(Y)}{M(X^{0.5^2}) - [M(X^{0.5})]^2} = \\ &= \frac{26.5835 - 1.5996 \cdot 15.5}{2.75 - 1.5996^2} = \frac{1.7902}{0.1914} = 9.3532 \\ \hat{\alpha} &= M(Y) - \hat{\beta} M(X^{0.5}) = 0.5389 \\ \rho^2 &= \frac{[\text{Cov}(X^{0.5}, Y)]^2}{\text{Var}(X^{0.5}) \text{Var}(Y)} = \frac{1.7902^2}{0.1914 \cdot 17.5} = 0.9568\end{aligned}$$

²Si osserva che per ragioni di stampa alcuni calcoli potrebbero essere duplicati con riferimento ai diversi modelli oggetto di studio





3. Si individui il modello migliore

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.

Esercizio 5.



Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.

Esercizio 5.

3. Si individui il modello migliore

Il modello migliore è $Y^* = 0.5377 + 9.3539\sqrt{X}$.



4. Si motivi perché $\eta_{Y|X}^2 > \eta_{X|Y}^2$

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.

Esercizio 5.



4. Si motivi perché $\eta_{Y|X}^2 > \eta_{X|Y}^2$

Sussiste dipendenza funzionale di Y da X , quindi $\eta_{Y|X}^2 = 1$.

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.

Esercizio 5.



Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.

Esercizio 5.

4. Si motivi perché $\eta_{Y|X}^2 > \eta_{X|Y}^2$

Sussiste dipendenza funzionale di Y da X , quindi $\eta_{Y|X}^2 = 1$.

$0 < \eta_{X|Y}^2 < 1$ in quanto non sussiste indipendenza in media e nemmeno dipendenza funzionale di X da Y (X assume entrambi i valori 4 e 5, in corrispondenza di $Y = 20$).

Esercizio 4

Prova scritta di
STATISTICA del
16.01.2014

Giuseppe Boari
Gabriele Cantaluppi



UNIVERSITÀ
CATTOLICA
del Sacro Cuore

Si completino le seguenti tabelle, giustificando la risposta, in modo che siano soddisfatte le condizioni indicate.

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.

Esercizio 5.



	$y_1 = 10$	$y_2 = 15$
$x_1 = 1$	11	
$x_2 = 2$		10

a) $\eta_{Y|X}^2 = \rho^2 > 0$

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.

Esercizio 5.



Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.

Esercizio 5.

	$y_1 = 10$	$y_2 = 15$
$x_1 = 1$	11	
$x_2 = 2$		10

$$\text{a) } \eta_{Y|X}^2 = \rho^2 > 0$$

$\eta_{Y|X}^2 = \rho^2 \rightarrow$ medie condiz. allineate:

$M(Y|x_1)$ e $M(Y|x_2)$ sono allineate qualsiasi siano le frequenze assegnate.

Occorre escludere la combinazione $n_{12} = 11, n_{21} = 10$, in quanto si avrebbe

$$M(Y|x_1) = M(Y|x_2)$$

con

$$\eta_{Y|X}^2 = \rho^2 = 0.$$



Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.

Esercizio 5.

	$y_1 = 10$	$y_2 = 15$
$x_1 = 1$	11	
$x_2 = 2$	5	

b) $\chi^2 = 0$



Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.

Esercizio 5.

	$y_1 = 10$	$y_2 = 15$
$x_1 = 1$	11	
$x_2 = 2$	5	

b) $\chi^2 = 0$

$\chi^2 = 0$ indep. stocastica

$n_{12} = 11c, n_{22} = 5c$ con $c > 0$ intero



	$y_1 = 10$	$y_2 = 15$
$x_1 = 1$	11	
$x_2 = 2$		

c) $\rho = -1$

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.

Esercizio 5.



Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.

Esercizio 5.

	$y_1 = 10$	$y_2 = 15$
$x_1 = 1$	11	
$x_2 = 2$		

c) $\rho = -1$

$\rho = -1$ dip. funz. lineare inversa tra X e Y
impossibile;

si verificherebbe con $n_{11} = 0, n_{12} > 0, n_{21} > 0, n_{22} = 0$

Esercizio 5

Vengono lanciati 3 dadi regolari.



Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.

Esercizio 5.



1. Si calcoli la probabilità che, essendosi presentati 3 numeri uguali, questi siano i numeri da 1 a 4.

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.

Esercizio 5.



Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.

Esercizio 5.

1. Si calcoli la probabilità che, essendosi presentati 3 numeri uguali, questi siano i numeri da 1 a 4.

Si indichino con A e B gli eventi $A =$ 'si presentano i numeri da 1 a 4' e $B =$ 'si presentano 3 numeri uguali'.

L'evento $(A \cap B)$ corrisponde a 3 numeri 1 oppure 3 numeri 2, ..., 3 numeri 4 e ha probabilità

$$P(A \cap B) = 4 \left(\frac{1}{6}\right)^3.$$



Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.

Esercizio 5.

1. Si calcoli la probabilità che, essendosi presentati 3 numeri uguali, questi siano i numeri da 1 a 4.

Si indichino con A e B gli eventi $A =$ 'si presentano i numeri da 1 a 4' e $B =$ 'si presentano 3 numeri uguali'.

L'evento $(A \cap B)$ corrisponde a 3 numeri 1 oppure 3 numeri 2, ..., 3 numeri 4 e ha probabilità

$$P(A \cap B) = 4 \left(\frac{1}{6}\right)^3.$$

$$\text{Invece } P(B) = 6 \left(\frac{1}{6}\right)^3.$$



[Esercizio 1.](#)

[Esercizio 2.](#)

[Esercizio 3.](#)

[Esercizio 4.](#)

[Esercizio 5.](#)

1. Si calcoli la probabilità che, essendosi presentati 3 numeri uguali, questi siano i numeri da 1 a 4.

Si indichino con A e B gli eventi $A =$ 'si presentano i numeri da 1 a 4' e $B =$ 'si presentano 3 numeri uguali'.

L'evento $(A \cap B)$ corrisponde a 3 numeri 1 oppure 3 numeri 2, ..., 3 numeri 4 e ha probabilità

$$P(A \cap B) = 4 \left(\frac{1}{6}\right)^3.$$

$$\text{Invece } P(B) = 6 \left(\frac{1}{6}\right)^3.$$

Applicando, ora, la formula della probabilità condizionata si ottiene

$$P(A|B) =$$



[Esercizio 1.](#)

[Esercizio 2.](#)

[Esercizio 3.](#)

[Esercizio 4.](#)

[Esercizio 5.](#)

1. Si calcoli la probabilità che, essendosi presentati 3 numeri uguali, questi siano i numeri da 1 a 4.

Si indichino con A e B gli eventi $A =$ 'si presentano i numeri da 1 a 4' e $B =$ 'si presentano 3 numeri uguali'.

L'evento $(A \cap B)$ corrisponde a 3 numeri 1 oppure 3 numeri 2, ..., 3 numeri 4 e ha probabilità

$$P(A \cap B) = 4 \left(\frac{1}{6}\right)^3.$$

$$\text{Invece } P(B) = 6 \left(\frac{1}{6}\right)^3.$$

Applicando, ora, la formula della probabilità condizionata si ottiene

$$P(A|B) = \frac{P(A \cap B)}{P(B)} = \frac{4 \left(\frac{1}{6}\right)^3}{6 \left(\frac{1}{6}\right)^3} = \frac{4}{6} = 0.6667$$

2. Si calcoli la probabilità che su 51 lanci in almeno 10 tutti e 3 i dadi presentino un numero pari.



Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.

Esercizio 5.

2. Si calcoli la probabilità che su 51 lanci in almeno 10 tutti e 3 i dadi presentino un numero pari.

L'evento tutti e 3 i dadi presentino un numero pari in un singolo lancio ha probabilità

$$p = \left(\frac{3}{6}\right)^3 = 0.5^3 = 0.125.$$



Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.

Esercizio 5.

2. Si calcoli la probabilità che su 51 lanci in almeno 10 tutti e 3 i dadi presentino un numero pari.

L'evento tutti e 3 i dadi presentino un numero pari in un singolo lancio ha probabilità

$$p = \left(\frac{3}{6}\right)^3 = 0.5^3 = 0.125.$$

Indicata con X la variabile casuale 'numero di volte che tutti e 3 i dadi presentino un numero pari' nei 51 lanci (prove indipendenti) risulta

$$X \sim$$



[Esercizio 1.](#)

[Esercizio 2.](#)

[Esercizio 3.](#)

[Esercizio 4.](#)

[Esercizio 5.](#)

2. Si calcoli la probabilità che su 51 lanci in almeno 10 tutti e 3 i dadi presentino un numero pari.

L'evento tutti e 3 i dadi presentino un numero pari in un singolo lancio ha probabilità

$$p = \left(\frac{3}{6}\right)^3 = 0.5^3 = 0.125.$$

Indicata con X la variabile casuale 'numero di volte che tutti e 3 i dadi presentino un numero pari' nei 51 lanci (prove indipendenti) risulta

$$X \sim \text{Bin}(n = 51, p = 0.125).$$

$$M(X) = 6.375; \text{Var}(X) = 5.5781.$$

Utilizzando l'approssimazione della variabile casuale Binomiale con la variabile casuale Normale

$$X \sim$$



Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.

Esercizio 5.

2. Si calcoli la probabilità che su 51 lanci in almeno 10 tutti e 3 i dadi presentino un numero pari.

L'evento tutti e 3 i dadi presentino un numero pari in un singolo lancio ha probabilità

$$p = \left(\frac{3}{6}\right)^3 = 0.5^3 = 0.125.$$

Indicata con X la variabile casuale 'numero di volte che tutti e 3 i dadi presentino un numero pari' nei 51 lanci (prove indipendenti) risulta

$$X \sim \text{Bin}(n = 51, p = 0.125).$$

$$M(X) = 6.375; \text{Var}(X) = 5.5781.$$

Utilizzando l'approssimazione della variabile casuale Binomiale con la variabile casuale Normale

$$X \sim N(\mu = 6.375, \sigma^2 = 5.5781)$$

si ottiene

$$\begin{aligned} P(X \geq 10) &= P\left(\frac{X - \mu}{\sigma} \geq \frac{10 - \mu}{\sigma}\right) = \\ &= P\left(Z \geq \frac{10 - 6.375}{2.3618}\right) = P(Z \geq 1.53) = \\ &= 1 - P(Z < 1.53) = 1 - P(Z \leq 1.53) = 0.063. \end{aligned}$$



Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.

Esercizio 5.

Tema T277

Prova scritta di STATISTICA del 30.01.2014

Prova scritta di
STATISTICA del
30.01.2014

Giuseppe Boari
Gabriele Cantaluppi



UNIVERSITÀ
CATTOLICA
del Sacro Cuore

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.

Esercizio 5.

Giuseppe Boari
Gabriele Cantaluppi
Università Cattolica del Sacro Cuore, Milano



Si riporta la valutazione della qualità (Q) e il prezzo (P) di uno stesso piatto proposto da 30 ristoranti italiani:

$Q \setminus P$	10 – 15	15 – 18	18 – 22
<i>discreta</i>	9	2	0
<i>buona</i>	5	7	5
<i>eccellente</i>	0	0	2

1. Fornire la rappresentazione grafica della distribuzione marginale P .

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.

Esercizio 5.



Si considera la distribuzione marginale della variabile P , ottenuta sommando le frequenze di colonna

p_j	$n_{\bullet j}$
10 $\dashv$ 15	14
15 $\dashv$ 18	9
18 $\dashv$ 22	7

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.

Esercizio 5.



Si tratta di un carattere quantitativo con modalità raggruppate in classi che può essere rappresentato graficamente mediante un istogramma.

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.

Esercizio 5.



L'istogramma può essere costruito dopo avere calcolato le densità $d_{\bullet j} = \frac{n_{\bullet j}}{a_j}$, (dove a_j sono le ampiezze di classe).

p_j	$n_{\bullet j}$	$a_{\bullet j}$	$d_{\bullet j}$
10 - 15	14	$15 - 10 = 5$	2.8
15 - 18	9	$18 - 15 = 3$	3
18 - 22	7	$22 - 18 = 4$	1.75

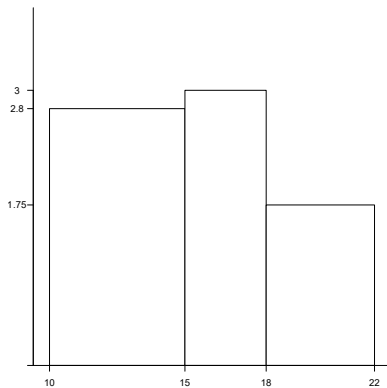
Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.

Esercizio 5.



Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.

Esercizio 5.



2. Calcolare la mediana e una misura di variabilità della marginale Q .

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.

Esercizio 5.



Si considera la distribuzione marginale della variabile Q , ottenuta sommando le frequenze di riga

q_i	$n_{i\bullet}$
<i>discreta</i>	11
<i>buona</i>	17
<i>eccellente</i>	2

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.

Esercizio 5.



Si tratta di un carattere qualitativo ordinato.

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.

Esercizio 5.



La mediana (percentile di ordine 0.50) in presenza di un numero pari di osservazioni corrisponde a uno degli elementi che occupano le posizioni $\frac{n}{2} = 15$ ed $\frac{n}{2} + 1 = 16$ nella serie dei dati ordinati. Dalla distribuzione delle frequenze cumulate

q_i	$n_{i\bullet}$	$N_{i\bullet}$
<i>discreta</i>	11	11
<i>buona</i>	17	28
<i>eccellente</i>	2	30

si trova che i due elementi coincidono e risulta $q_{0.5} = \textit{buona}$.

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.

Esercizio 5.

In presenza di un carattere qualitativo la variabilità può essere misurata mediante gli indici di Gini e di Frosini normalizzati

$$G_N = \frac{1 - \sum_{i=1}^h f_{i\bullet}^2}{1 - \frac{1}{h}}$$

e

$$F_N = 1 - \sqrt{1 - G_N}.$$



Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.

Esercizio 5.



In presenza di un carattere qualitativo la variabilità può essere misurata mediante gli indici di Gini e di Frosini normalizzati

$$G_N = \frac{1 - \sum_{i=1}^h f_{i\bullet}^2}{1 - \frac{1}{h}}$$

e

$$F_N = 1 - \sqrt{1 - G_N}.$$

Si ottiene:

q_i	$n_{i\bullet}$	$f_{i\bullet}$	$f_{i\bullet}^2$
<i>discreta</i>	11	0.3667	0.1344
<i>buona</i>	17	0.5667	0.3211
<i>eccellente</i>	2	0.0667	0.0044
	30	1	0.4599

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.

Esercizio 5.



In presenza di un carattere qualitativo la variabilità può essere misurata mediante gli indici di Gini e di Frosini normalizzati

$$G_N = \frac{1 - \sum_{i=1}^h f_{i\bullet}^2}{1 - \frac{1}{h}}$$

e

$$F_N = 1 - \sqrt{1 - G_N}.$$

Si ottiene:

q_i	$n_{i\bullet}$	$f_{i\bullet}$	$f_{i\bullet}^2$
<i>discreta</i>	11	0.3667	0.1344
<i>buona</i>	17	0.5667	0.3211
<i>eccellente</i>	2	0.0667	0.0044
	30	1	0.4599

Da cui segue:

$$G_N = \frac{1 - 0.46}{1 - \frac{1}{3}} = 0.81$$

e

$$F_N = 1 - \sqrt{1 - 0.81} = 0.5641.$$

Sussiste, quindi, un livello medio di eterogeneità.



3. Dopo aver valutato, senza fare calcoli, se Q e P sono stocasticamente indipendenti, calcolare un indice di connessione.

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.

Esercizio 5.



3. Dopo aver valutato, senza fare calcoli, se Q e P sono stocasticamente indipendenti, calcolare un indice di connessione.

Q e P non sono stocasticamente indipendenti in quanto la presenza di frequenze congiunte nulle esclude che le distribuzioni condizionate siano somiglianti.

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.

Esercizio 5.



Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.

Esercizio 5.

Dalla distribuzione delle frequenze congiunte n_{ij} :

	10 ÷ 15	15 ÷ 18	18 ÷ 22	$n_{i\bullet}$
<i>discreta</i>	9	2	0	11
<i>buona</i>	5	7	5	17
<i>eccellente</i>	0	0	2	2
$n_{\bullet j}$	14	9	7	30

è possibile ottenere le frequenze teoriche nella situazione di indipendenza stocastica

$$\hat{n}_{ij} = \frac{n_{i\bullet} \cdot n_{\bullet j}}{n}:$$



Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.

Esercizio 5.

Dalla distribuzione delle frequenze congiunte n_{ij} :

	10 ÷ 15	15 ÷ 18	18 ÷ 22	$n_{i\bullet}$
<i>discreta</i>	9	2	0	11
<i>buona</i>	5	7	5	17
<i>eccellente</i>	0	0	2	2
$n_{\bullet j}$	14	9	7	30

è possibile ottenere le frequenze teoriche nella situazione di indipendenza stocastica

$$\hat{n}_{ij} = \frac{n_{i\bullet} \cdot n_{\bullet j}}{n}:$$

	10 ÷ 15	15 ÷ 18	18 ÷ 22	$n_{i\bullet}$
<i>discreta</i>	5.1333	3.3	2.5667	11
<i>buona</i>	7.9333	5.1	3.9667	17
<i>eccellente</i>	0.9333	0.6	0.4667	2
$n_{\bullet j}$	14.0000	9.0	7.0000	30



le contingenze assolute $n_{ij} - \hat{n}_{ij}$:

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.

Esercizio 5.



le contingenze assolute $n_{ij} - \hat{n}_{ij}$:

	10 ÷ 15	15 ÷ 18	18 ÷ 22
<i>discreta</i>	3.8667	-1.3	-2.5667
<i>buona</i>	-2.9333	1.9	1.0333
<i>eccellente</i>	-0.9333	-0.6	1.5333

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.

Esercizio 5.



i rapporti di contingenza $\frac{n_{ij}}{\hat{n}_{ij}}$:

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.

Esercizio 5.



i rapporti di contingenza $\frac{n_{ij}}{\hat{n}_{ij}}$:

	10 ÷ 15	15 ÷ 18	18 ÷ 22
<i>discreta</i>	1.7532	0.6061	0.0000
<i>buona</i>	0.6303	1.3725	1.2605
<i>eccellente</i>	0.0000	0.0000	4.2857

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.

Esercizio 5.



e le contingenze relative, espresse in termini percentuali, $100 \left(\frac{n_{ij} - \hat{n}_{ij}}{\hat{n}_{ij}} \right) \% = 100 \left(\frac{n_{ij}}{\hat{n}_{ij}} - 1 \right) \%$:

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.

Esercizio 5.



e le contingenze relative, espresse in termini percentuali, $100 \left(\frac{n_{ij} - \hat{n}_{ij}}{\hat{n}_{ij}} \right) \% = 100 \left(\frac{n_{ij}}{\hat{n}_{ij}} - 1 \right) \%$:

	10 ÷ 15	15 ÷ 18	18 ÷ 22
<i>discreta</i>	75.32	-39.39	-100.00
<i>buona</i>	-36.97	37.25	26.05
<i>eccellente</i>	-100.00	-100.00	328.57

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.

Esercizio 5.



Risultati intermedi $\frac{(n_{ij} - \hat{n}_{ij})^2}{\hat{n}_{ij}}$ per il calcolo del χ^2

	10 - 15	15 - 18	18 - 22
<i>discreta</i>	2.9126	0.5121	2.5667
<i>buona</i>	1.0846	0.7078	0.2692
<i>eccellente</i>	0.9333	0.6000	5.0381

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.

Esercizio 5.



Risultati intermedi $\frac{n_{ij}^2}{n_{i\bullet} \cdot n_{\bullet j}}$ per il calcolo del χ^2 (formula operativa)

	10 ÷ 15	15 ÷ 18	18 ÷ 22
<i>discreta</i>	0.526	0.0404	0.0000
<i>buona</i>	0.105	0.3203	0.2101
<i>eccellente</i>	0.000	0.0000	0.2857

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.

Esercizio 5.



Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.

Esercizio 5.

$$\chi^2 = \sum_{i=1}^h \sum_{j=1}^k \frac{(n_{ij} - \hat{n}_{ij})^2}{\hat{n}_{ij}} = 14.6244,$$

$$\chi_N^2 = \frac{\chi^2}{n \cdot \min(h-1, k-1)} = 0.2437$$

$$\chi_N = \sqrt{\chi_N^2} = 0.4937.$$

Sussiste, quindi, un livello di connessione medio.



Il prospetto seguente riporta i prezzi medi annui di 2 beni di uso quotidiano nel periodo 2002 – 2006:

t	$bene A$	$bene B$
2002	99.4	77.0
2003	117.3	65.8
2004	129.8	47.2
2005	149.3	35.8
2006	158.1	33.8

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.

Esercizio 5.



1. Costruire la serie dei numeri indici a base fissa con base 2003 per il bene A.

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.

Esercizio 5.



1. Costruire la serie dei numeri indici a base fissa con base 2003 per il bene A.

Si ricorda che ${}_{2003}I_t = \frac{x_t}{x_{2003}}$.

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.

Esercizio 5.



1. Costruire la serie dei numeri indici a base fissa con base 2003 per il bene A.
Si ricorda che ${}_{2003}I_t = \frac{x_t}{x_{2003}}$. Si ottiene quindi

t	x_t	$NIBF (2003)$
2002	99.4	0.8474
2003	117.3	1.0000
2004	129.8	1.1066
2005	149.3	1.2728
2006	158.1	1.3478

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.

Esercizio 5.



2. Calcolare il tasso medio di variazione per il prezzo del bene B tra il 2002 e il 2005, interpretando il risultato.

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.

Esercizio 5.



2. Calcolare il tasso medio di variazione per il prezzo del bene B tra il 2002 e il 2005, interpretando il risultato.

Si ricorda che

$${}_{2002}I_{2005} = {}_{2002}I_{2003} \cdot {}_{2003}I_{2004} \cdot {}_{2004}I_{2005}$$

Si deve trovare il numero indice a base mobile medio.

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.

Esercizio 5.



2. Calcolare il tasso medio di variazione per il prezzo del bene B tra il 2002 e il 2005, interpretando il risultato.

Si ricorda che

$${}_{2002}I_{2005} = {}_{2002}I_{2003} \cdot {}_{2003}I_{2004} \cdot {}_{2004}I_{2005}$$

Si deve trovare il numero indice a base mobile medio. Applicando il criterio di scelta della media secondo Chisini si ottiene

$${}_{2002}I_{2005} = {}_{2002}I_{2003} \cdot {}_{2003}I_{2004} \cdot {}_{2004}I_{2005} = \alpha \cdot \alpha \cdot \alpha = \alpha^3.$$

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.

Esercizio 5.



Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.

Esercizio 5.

2. Calcolare il tasso medio di variazione per il prezzo del bene B tra il 2002 e il 2005, interpretando il risultato.

Si ricorda che

$${}_{2002}I_{2005} = {}_{2002}I_{2003} \cdot {}_{2003}I_{2004} \cdot {}_{2004}I_{2005}$$

Si deve trovare il numero indice a base mobile medio. Applicando il criterio di scelta della media secondo Chisini si ottiene

$${}_{2002}I_{2005} = {}_{2002}I_{2003} \cdot {}_{2003}I_{2004} \cdot {}_{2004}I_{2005} = \alpha \cdot \alpha \cdot \alpha = \alpha^3.$$

Quindi

$$\alpha = \sqrt[3]{{}_{2002}I_{2005}} = \sqrt[3]{{}_{2002}I_{2003} \cdot {}_{2003}I_{2004} \cdot {}_{2004}I_{2005}} = 0.7747.$$

Da cui segue una variazione relativa media pari a -0.2253 , corrispondente a un decremento medio 2002 e il 2005 del 22.53%.



Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.

Esercizio 5.

3. Note le quantità ${}_A q_{2003} = 1500$ e ${}_B q_{2003} = 8500$, calcolare l'indice di Laspeyres dei prezzi al 2006 con base 2003.



L'indice dei prezzi di Laspeyres si ottiene considerando la media aritmetica dei 'prezzi relativi' riferiti ai singoli prodotti $x_i = \frac{i p_{2006}}{i p_{2003}}$ calcolata utilizzando come frequenze i valori

$i v_{2003} = i p_{2003} \cdot i q_{2003}$ scambiati al tempo 2003

$${}^p_{2003} I^L_{2006} = \frac{1}{\sum_{i=1}^h i p_{2003} i q_{2003}} \sum_{i=1}^h \frac{i p_{2006}}{i p_{2003}} i p_{2003} i q_{2003}$$

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.

Esercizio 5.



Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.

Esercizio 5.

$\frac{i p_{2006}}{i p_{2003}}$	$i p_{2003} \cdot i q_{2003}$	$i p_{2006} \cdot i q_{2003}$
$\frac{158.1}{117.3} = 1.3478$	$117.3 \cdot 1500 = 175950$	237150
$\frac{33.8}{65.8} = 0.5137$	$65.8 \cdot 8500 = 559300$	287300
	735250	524450

$$p_{2003} I_{2006}^L = \frac{1}{735250} 524450 = 0.7133.$$

La variazione relativa nel livello generale dei prezzi, secondo l'Indice di Laspeyres, è del -28.67% .



Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.

Esercizio 5.

La tabella seguente riporta le osservazioni della variabile Y in corrispondenza di 5 osservazioni della X :

x_i	1	1	3	5	5
y_i	30	40	45	50	60

1. Si rappresentino graficamente le coppie dei dati e la funzione di regressione ipotizzando Y in funzione di X .



La funzione di regressione associa a ogni valore x appartenente all'insieme di definizione della variabile X la media di Y condizionata a x , $M(Y|x)$

x	$M(Y x)$
1	$\frac{1}{2}(30 + 40) = 35$
3	45
5	$\frac{1}{2}(50 + 60) = 55$

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.

Esercizio 5.



Esercizio 1.

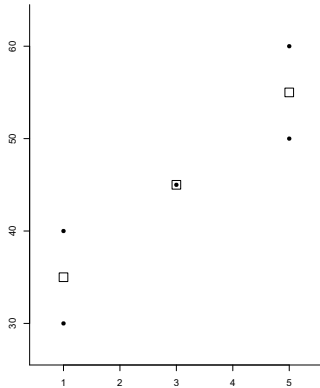
Esercizio 2.

Esercizio 3.

Esercizio 4.

Esercizio 5.

Le coppie di dati possono essere rappresentate mediante un grafico a dispersione.





2. Si calcolino, secondo il metodo dei minimi quadrati, i parametri e l'indice di adattamento dei seguenti modelli di regressione, individuando il modello migliore:

$$I) Y^* = a + b \cdot X \qquad II) Y^* = \alpha \cdot X^2$$

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.

Esercizio 5.

Si riportano i calcoli necessari¹ per ottenere le stime dei parametri e gli indici di adattamento del modello

$$Y = a + bX + E$$

x_i	y_i	$x_i y_i$	x_i^2	y_i^2
1	30	30	1	900
1	40	40	1	1600
3	45	135	9	2025
5	50	250	25	2500
5	60	300	25	3600
15	225	755	61	10625
3	45	151	12.2	2125
$M(X)$	$M(Y)$	$M(XY)$	$M(X^2)$	$M(Y^2)$

$$\hat{b} =$$



Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.

Esercizio 5.

¹Si osserva che per ragioni di stampa alcuni calcoli potrebbero essere duplicati con riferimento ai diversi modelli oggetto di studio

Si riportano i calcoli necessari¹ per ottenere le stime dei parametri e gli indici di adattamento del modello

$$Y = a + bX + E$$

x_i	y_i	$x_i y_i$	x_i^2	y_i^2
1	30	30	1	900
1	40	40	1	1600
3	45	135	9	2025
5	50	250	25	2500
5	60	300	25	3600
15	225	755	61	10625
3	45	151	12.2	2125
$M(X)$	$M(Y)$	$M(XY)$	$M(X^2)$	$M(Y^2)$

$$\begin{aligned}\hat{b} &= \frac{\text{Cov}(X, Y)}{\text{Var}(X)} = \frac{M(XY) - M(X)M(Y)}{M(X^2) - [M(X)]^2} = \\ &= \frac{151 - 3 \cdot 45}{12.2 - 3^2} = \frac{16}{3.2} = 5 \\ \hat{a} &= \end{aligned}$$

¹Si osserva che per ragioni di stampa alcuni calcoli potrebbero essere duplicati con riferimento ai diversi modelli oggetto di studio



Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.

Esercizio 5.

Si riportano i calcoli necessari¹ per ottenere le stime dei parametri e gli indici di adattamento del modello

$$Y = a + bX + E$$

x_i	y_i	$x_i y_i$	x_i^2	y_i^2
1	30	30	1	900
1	40	40	1	1600
3	45	135	9	2025
5	50	250	25	2500
5	60	300	25	3600
15	225	755	61	10625
3	45	151	12.2	2125
$M(X)$	$M(Y)$	$M(XY)$	$M(X^2)$	$M(Y^2)$

$$\hat{b} = \frac{\text{Cov}(X, Y)}{\text{Var}(X)} = \frac{M(XY) - M(X)M(Y)}{M(X^2) - [M(X)]^2} =$$

$$= \frac{151 - 3 \cdot 45}{12.2 - 3^2} = \frac{16}{3.2} = 5$$

$$\hat{a} = M(Y) - \hat{b}M(X) = 30$$

$$\rho^2 =$$



Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.

Esercizio 5.

¹Si osserva che per ragioni di stampa alcuni calcoli potrebbero essere duplicati con riferimento ai diversi modelli oggetto di studio

Si riportano i calcoli necessari¹ per ottenere le stime dei parametri e gli indici di adattamento del modello

$$Y = a + bX + E$$

x_i	y_i	$x_i y_i$	x_i^2	y_i^2
1	30	30	1	900
1	40	40	1	1600
3	45	135	9	2025
5	50	250	25	2500
5	60	300	25	3600
15	225	755	61	10625
3	45	151	12.2	2125
$M(X)$	$M(Y)$	$M(XY)$	$M(X^2)$	$M(Y^2)$

$$\hat{b} = \frac{\text{Cov}(X, Y)}{\text{Var}(X)} = \frac{M(XY) - M(X)M(Y)}{M(X^2) - [M(X)]^2} =$$

$$= \frac{151 - 3 \cdot 45}{12.2 - 3^2} = \frac{16}{3.2} = 5$$

$$\hat{a} = M(Y) - \hat{b}M(X) = 30$$

$$\rho^2 = \frac{[\text{Cov}(X, Y)]^2}{\text{Var}(X) \text{Var}(Y)} = \frac{16^2}{3.2 \cdot 100} = 0.8$$

$$M(\hat{E}^2) = \text{Var}(\hat{E}) = \text{varianza residua} =$$

¹Si osserva che per ragioni di stampa alcuni calcoli potrebbero essere duplicati con riferimento ai diversi modelli oggetto di studio



Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.

Esercizio 5.

Si riportano i calcoli necessari¹ per ottenere le stime dei parametri e gli indici di adattamento del modello

$$Y = a + bX + E$$

x_i	y_i	$x_i y_i$	x_i^2	y_i^2
1	30	30	1	900
1	40	40	1	1600
3	45	135	9	2025
5	50	250	25	2500
5	60	300	25	3600
15	225	755	61	10625
3	45	151	12.2	2125
$M(X)$	$M(Y)$	$M(XY)$	$M(X^2)$	$M(Y^2)$

$$\hat{b} = \frac{\text{Cov}(X, Y)}{\text{Var}(X)} = \frac{M(XY) - M(X)M(Y)}{M(X^2) - [M(X)]^2} =$$

$$= \frac{151 - 3 \cdot 45}{12.2 - 3^2} = \frac{16}{3.2} = 5$$

$$\hat{a} = M(Y) - \hat{b}M(X) = 30$$

$$\rho^2 = \frac{[\text{Cov}(X, Y)]^2}{\text{Var}(X) \text{Var}(Y)} = \frac{16^2}{3.2 \cdot 100} = 0.8$$

$$M(\hat{E}^2) = \text{Var}(\hat{E}) = \text{varianza residua} = \text{Var}(Y) \cdot (1 - \rho^2) = 20$$

¹Si osserva che per ragioni di stampa alcuni calcoli potrebbero essere duplicati con riferimento ai diversi modelli oggetto di studio





Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.

Esercizio 5.

Si riportano i calcoli necessari² per ottenere le stime dei parametri e la misura di adattamento del modello

$$Y = cX^2 + E$$

x_i^2	y_i	$x_i^2 y_i$	x_i^4	y_i^2
1	30	30	1	900
1	40	40	1	1600
9	45	405	81	2025
25	50	1250	625	2500
25	60	1500	625	3600
		3225	1333	10625
		645	266.6	2125
		$M(X^2 Y)$	$M(X^4)$	$M(Y^2)$

$$\hat{c} =$$

²Si osserva che per ragioni di stampa alcuni calcoli potrebbero essere duplicati con riferimento ai diversi modelli oggetto di studio



Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.

Esercizio 5.

Si riportano i calcoli necessari² per ottenere le stime dei parametri e la misura di adattamento del modello

$$Y = cX^2 + E$$

x_i^2	y_i	$x_i^2 y_i$	$x_i^{2^2}$	y_i^2
1	30	30	1	900
1	40	40	1	1600
9	45	405	81	2025
25	50	1250	625	2500
25	60	1500	625	3600
		3225	1333	10625
		645	266.6	2125
		$M(X^2 Y)$	$M(X^{2^2})$	$M(Y^2)$

$$\hat{c} = \frac{M(X^2 Y)}{M(X^{2^2})} = \frac{645}{266.6} = 2.4194$$

$$M(\hat{E}^2) =$$

²Si osserva che per ragioni di stampa alcuni calcoli potrebbero essere duplicati con riferimento ai diversi modelli oggetto di studio



Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.

Esercizio 5.

Si riportano i calcoli necessari² per ottenere le stime dei parametri e la misura di adattamento del modello

$$Y = cX^2 + E$$

x_i^2	y_i	$x_i^2 y_i$	$x_i^{2^2}$	y_i^2
1	30	30	1	900
1	40	40	1	1600
9	45	405	81	2025
25	50	1250	625	2500
25	60	1500	625	3600
		3225	1333	10625
		645	266.6	2125
		$M(X^2 Y)$	$M(X^{2^2})$	$M(Y^2)$

$$\hat{c} = \frac{M(X^2 Y)}{M(X^{2^2})} = \frac{645}{266.6} = 2.4194$$

$$M(\hat{E}^2) = M[Y \cdot (Y - \hat{c}X^2)] = M(Y^2) - \hat{c}M(X^2 Y) = 2125 - 2.4194 \cdot 645 = 564.5161$$

²Si osserva che per ragioni di stampa alcuni calcoli potrebbero essere duplicati con riferimento ai diversi modelli oggetto di studio



Il modello migliore è $Y^* = 30 + 5X$.

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.

Esercizio 5.



3. Si indichi il valore assunto dal rapporto di correlazione $\eta^2_{Y|X}$, commentando il risultato.

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.

Esercizio 5.



Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.

Esercizio 5.

3. Si indichi il valore assunto dal rapporto di correlazione $\eta_{Y|X}^2$, commentando il risultato. Si osserva che le medie condizionate $M(Y|x_i)$ sono allineate; quindi $\eta_{Y|X}^2 = \rho^2 = 0.8$.

Esercizio 4

Prova scritta di
STATISTICA del
30.01.2014

Giuseppe Boari
Gabriele Cantaluppi



UNIVERSITÀ
CATTOLICA
del Sacro Cuore

Si completino le seguenti tabelle, giustificando la risposta, in modo che siano soddisfatte le condizioni indicate.

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.

Esercizio 5.



	$y_1 = 10$	$y_2 = 20$	$y_3 = 30$
$x_1 = 1$	0	11	
$x_2 = 2$		0	

a) $\eta_{X|Y}^2 = 1, \rho^2 = 0$

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.

Esercizio 5.



	$y_1 = 10$	$y_2 = 20$	$y_3 = 30$
$x_1 = 1$	0	11	
$x_2 = 2$		0	

a) $\eta_{X|Y}^2 = 1, \rho^2 = 0$

$\eta_{X|Y}^2 = 1, \rho^2 = 0$ dip. funz. di X da Y e indep. lin.

ricordando che $\eta_{Y|X}^2 = 0 \rightarrow \rho^2 = 0$

$n_{13} = 0, n_{21} = n_{23} > 0$

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.

Esercizio 5.



	$y_1 = 10$	$y_2 = 20$	$y_3 = 30$
$x_1 = 1$		11	
$x_2 = 2$		11	

b) $\eta_{Y|X}^2 = \rho^2 \neq 0$

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.

Esercizio 5.



Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.

Esercizio 5.

	$y_1 = 10$	$y_2 = 20$	$y_3 = 30$
$x_1 = 1$		11	
$x_2 = 2$		11	

$$\text{b) } \eta_{Y|X}^2 = \rho^2 \neq 0$$

$\eta_{Y|X}^2 = \rho^2 \rightarrow$ medie condiz. allineate.

Le due medie condizionate saranno allineate qualsiasi siano le frequenze assegnate.

Occorre escludere una delle combinazioni $n_{11} = n_{13}$, $n_{21} = n_{23}$.

Nel caso fossero entrambe verificate, $Y|x_1$ e $Y|x_2$ risulterebbero simmetriche con $M(Y|x_1) = M(Y|x_2)$ ed $\eta_{Y|X}^2 = \rho^2 = 0$.



Date due differenti urne, U_1 e U_2 , contenenti palline numerate, un esperimento casuale consiste nell'estrarre una pallina da ciascuna urna. Sapendo che nell'urna U_1 vi sono 5 palline con numero pari (2,4,6,8,10) e che nell'urna U_2 vi sono 2 palline con il numero 1, 2 con numero 3 e 2 con il numero 5:

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.

Esercizio 5.



1. costruire la distribuzione di probabilità delle variabili casuali $X = \text{"risultato estrazione urna } U_1\text{"}$ e $Y = \text{"risultato estrazione urna } U_2\text{"}$;

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.

Esercizio 5.



1. costruire la distribuzione di probabilità delle variabili casuali $X = \text{“risultato estrazione urna } U_1\text{”}$ e $Y = \text{“risultato estrazione urna } U_2\text{”}$;

x	$P(X = x)$
2	0.2
4	0.2
6	0.2
8	0.2
10	0.2

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.

Esercizio 5.



1. costruire la distribuzione di probabilità delle variabili casuali $X =$ “risultato estrazione urna U_1 ” e $Y =$ “risultato estrazione urna U_2 ”;

x	$P(X = x)$
2	0.2
4	0.2
6	0.2
8	0.2
10	0.2

y	$P(Y = y)$
1	$1/3$
3	$1/3$
5	$1/3$

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.

Esercizio 5.



2. calcolare $P(\{X = 10\} \cap \{Y = 1\})$, ovvero la probabilità che sia estratto il numero 10 da U_1 e il numero 1 da U_2 ;

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.

Esercizio 5.



Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.

Esercizio 5.

2. calcolare $P(\{X = 10\} \cap \{Y = 1\})$, ovvero la probabilità che sia estratto il numero 10 da U_1 e il numero 1 da U_2 ;

Dal momento che le estrazioni dalle due urne avvengono in maniera indipendente abbiamo che

$$\begin{aligned} P(\{X = 10\} \cap \{Y = 1\}) &= P(\{X = 10\}) \cdot P(\{Y = 1\}) = \\ &= 0.2 \cdot 0.3333 = 0.06667. \end{aligned}$$



3. calcolare la probabilità che, su 100 estrazioni con reinserimento dalla sola urna U_2 , almeno 21 presentino il numero 3.

Indicata con X la variabile casuale 'numero di volte che si presenta il numero 3' nelle 100 estrazioni con reinserimento (prove indipendenti) risulta

$$X \sim$$

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.

Esercizio 5.



Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.

Esercizio 5.

3. calcolare la probabilità che, su 100 estrazioni con reinserimento dalla sola urna U_2 , almeno 21 presentino il numero 3.

Indicata con X la variabile casuale 'numero di volte che si presenta il numero 3' nelle 100 estrazioni con reinserimento (prove indipendenti) risulta

$$X \sim \text{Bin}(n = 100, p = 0.3333).$$

$$M(X) = 33.3333; \text{Var}(X) = 22.2222;$$

utilizzando l'approssimazione della variabile casuale Binomiale con la variabile casuale Normale

$$X \sim$$



3. calcolare la probabilità che, su 100 estrazioni con reinserimento dalla sola urna U_2 , almeno 21 presentino il numero 3.

Indicata con X la variabile casuale 'numero di volte che si presenta il numero 3' nelle 100 estrazioni con reinserimento (prove indipendenti) risulta

$$X \sim \text{Bin}(n = 100, p = 0.3333).$$

$$M(X) = 33.3333; \text{Var}(X) = 22.2222;$$

utilizzando l'approssimazione della variabile casuale Binomiale con la variabile casuale Normale

$$X \sim N(\mu = 33.3333, \sigma^2 = 22.2222)$$

si ottiene

$$\begin{aligned} P(X \geq 21) &= P\left(\frac{X - \mu}{\sigma} \geq \frac{21 - \mu}{\sigma}\right) = P\left(Z \geq \frac{21 - 33.3333}{4.714}\right) \\ &= P(Z \geq -2.62) = P(Z \leq 2.62) = 0.9956. \end{aligned}$$

Tema T278

Prova scritta di STATISTICA del 13.02.2014

Prova scritta di
STATISTICA del
13.02.2014

Giuseppe Boari
Gabriele Cantaluppi



UNIVERSITÀ
CATTOLICA
del Sacro Cuore

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.

Giuseppe Boari
Gabriele Cantaluppi
Università Cattolica del Sacro Cuore, Milano



Con riferimento alla distribuzione congiunta delle variabili X e Y sono note le sole frequenze congiunte riportate nella seguente tabella a doppia entrata:

$X \setminus Y$	y_1	y_2	y_3
x_1		11	
x_2		11	

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.



1. Si completi la tabella in modo tale che le distribuzioni condizionate $Y|x_1$ e $Y|x_2$ siano caratterizzate rispettivamente da minima e massima eterogeneità. Si indichi il numero totale delle unità statistiche.

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.



1. Si completi la tabella in modo tale che le distribuzioni condizionate $Y|x_1$ e $Y|x_2$ siano caratterizzate rispettivamente da minima e massima eterogeneità. Si indichi il numero totale delle unità statistiche.

Affinché sussista minima eterogeneità, la distribuzione condizionata $Y|x_1$ deve essere degenerare.

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.



1. Si completi la tabella in modo tale che le distribuzioni condizionate $Y|x_1$ e $Y|x_2$ siano caratterizzate rispettivamente da minima e massima eterogeneità. Si indichi il numero totale delle unità statistiche.

Affinché sussista minima eterogeneità, la distribuzione condizionata $Y|x_1$ deve essere degenerare.

Affinché $Y|x_2$ sia caratterizzata da massima eterogeneità si deve avere equidistribuzione delle frequenze.

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.



Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.

1. Si completi la tabella in modo tale che le distribuzioni condizionate $Y|x_1$ e $Y|x_2$ siano caratterizzate rispettivamente da minima e massima eterogeneità. Si indichi il numero totale delle unità statistiche.

Affinché sussista minima eterogeneità, la distribuzione condizionata $Y|x_1$ deve essere degenerare.

Affinché $Y|x_2$ sia caratterizzata da massima eterogeneità si deve avere equidistribuzione delle frequenze.

$X \setminus Y$	y_1	y_2	y_3	
x_1	0	11	0	11
x_2	11	11	11	33
	11	22	11	44



2. Sapendo che $y_1 = 10$, $y_2 = 15$ e $y_3 = 20$, si confrontino le distribuzioni condizionate $Y|x_1$ e $Y|x_2$ con opportuni indici di posizione, variabilità e asimmetria, commentando i risultati ottenuti.

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.



Si considera la distribuzione condizionata $Y|x_1$

y_j	n_{1j}
10	0
15	11
20	0

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.



Si tratta di un carattere quantitativo con modalità non raggruppate in classi.

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.



Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.

Si osserva che la distribuzione è degenere.

$M(Y|x_1) = 15$. La varianza e il coefficiente di variazione sono nulli.



Anche l'indice di asimmetria assume valore nullo.

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.



Si considera la distribuzione condizionata $Y|x_2$

y_j	n_{2j}
10	11
15	11
20	11

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.



Si tratta di un carattere quantitativo con modalità non raggruppate in classi.

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.



Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.

Si osserva che la distribuzione è simmetrica rispetto a 15; la mediana (percentile di ordine 0.50) risulta $y_{0.5} = 15$.



La media aritmetica

$$M(Y|x_2) = \frac{1}{n_{2\bullet}} \sum_{j=1}^k y_j n_{2j}$$

risulta 15.

Si osserva, infatti, che la distribuzione è simmetrica.

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.



La varianza

$$\text{Var}(Y|x_2) = M(Y^2|x_2) - [M(Y|x_2)]^2$$

risulta 16.6667.

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.



La varianza

$$\text{Var}(Y|x_2) = M(Y^2|x_2) - [M(Y|x_2)]^2$$

risulta 16.6667. Si ottiene

y_j	n_{2j}	$y_j n_{2j}$	$y_j^2 n_{2j}$
10	11	110	1100
15	11	165	2475
20	11	220	4400
	33	495	7975

da cui segue

$$\text{Var}(Y|x_2) = \frac{7975}{33} - 15^2 = 241.6667 - 225 = 16.6667.$$

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.



Per effettuare confronti in termini di variabilità tra caratteri quantitativi occorre utilizzare una misura relativa, ad esempio il coefficiente di variazione

$$CV = \frac{\sigma}{\mu} = \frac{\sqrt{\text{Var}(Y|x_2)}}{M(Y|x_2)} = \frac{\sqrt{16.6667}}{15} = \frac{4.0825}{15} = 0.2722.$$

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.



La distribuzione è simmetrica rispetto al valore 15; quindi l'indice di asimmetria assume valore nullo.

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.



Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.

È possibile concludere che entrambe le distribuzioni sono simmetriche e hanno eguale media; $Y|x_2$ presenta un livello di variabilità superiore a $Y|x_1$, distribuzione degenera.



3. Si rappresenti graficamente la distribuzione marginale Y .

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.



Si considera la distribuzione marginale della variabile Y , ottenuta sommando le frequenze di colonna

y_j	$n_{\bullet j}$
10	11
15	22
20	11

Esercizio 1.

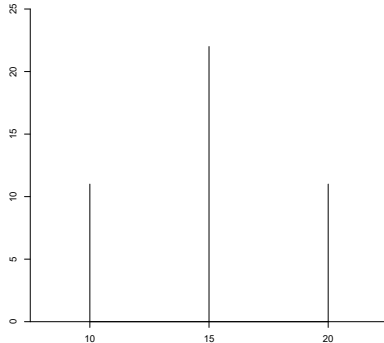
Esercizio 2.

Esercizio 3.

Esercizio 4.



Si tratta di un carattere quantitativo con modalità non raggruppate in classi che può essere rappresentato graficamente mediante un grafico a bastoncini.



Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.

Esercizio 2



Della seguente serie storica sono note solo alcune informazioni.

t	x_t
1	200
2	x_2
3	180
4	x_4
5	160

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.



1. Si ricostruisca la serie storica sapendo che tra il tempo 1 e il tempo 2 vi è stato un incremento di 21 e che tra $t = 3$ e $t = 4$ vi è stata una diminuzione del 5%.

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.



1. Si ricostruisca la serie storica sapendo che tra il tempo 1 e il tempo 2 vi è stato un incremento di 21 e che tra $t = 3$ e $t = 4$ vi è stata una diminuzione del 5%.

$$x_2 = x_1 + 21 = 200 + 21 = 221$$

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.



1. Si ricostruisca la serie storica sapendo che tra il tempo 1 e il tempo 2 vi è stato un incremento di 21 e che tra $t = 3$ e $t = 4$ vi è stata una diminuzione del 5%.

$$x_2 = x_1 + 21 = 200 + 21 = 221$$

$$x_4 = 0.95 \cdot x_3 = 0.95 \cdot 180 = 171$$

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.



2. Si costruiscano le serie dei numeri indici a base fissa, con base $t = 1$, e a base mobile.

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.



2. Si costruiscano le serie dei numeri indici a base fissa, con base $t = 1$, e a base mobile.
Si ricorda che ${}_1I_t = \frac{x_t}{x_1}$.

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.



2. Si costruiscano le serie dei numeri indici a base fissa, con base $t = 1$, e a base mobile. Si ricorda che ${}_1I_t = \frac{x_t}{x_1}$. Si ottiene quindi

t	x_t	$NIBF(1)$
1	200	1.000
2	221	1.105
3	180	0.900
4	171	0.855
5	160	0.800

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.



Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.

2. Si costruiscano le serie dei numeri indici a base fissa, con base $t = 1$, e a base mobile. Si ricorda che ${}_1I_t = \frac{x_t}{x_1}$. Si ottiene quindi

t	x_t	$NIBF(1)$
1	200	1.000
2	221	1.105
3	180	0.900
4	171	0.855
5	160	0.800

Si ricorda che ${}_{t-1}I_t = \frac{x_t}{x_{t-1}}$.



Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.

2. Si costruiscano le serie dei numeri indici a base fissa, con base $t = 1$, e a base mobile. Si ricorda che ${}_1l_t = \frac{x_t}{x_1}$. Si ottiene quindi

t	x_t	$NIBF(1)$
1	200	1.000
2	221	1.105
3	180	0.900
4	171	0.855
5	160	0.800

Si ricorda che ${}_{t-1}l_t = \frac{x_t}{x_{t-1}}$. Si ottiene quindi

t	x_t	$NIBM$
1	200	
2	221	1.1050
3	180	0.8145
4	171	0.9500
5	160	0.9357



3. Si calcoli il tasso medio di variazione di x_t tra l'istante temporale 1 e l'istante 3, interpretando il risultato.

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.



3. Si calcoli il tasso medio di variazione di x_t tra l'istante temporale 1 e l'istante 3, interpretando il risultato.

Si ricorda che

$${}_1l_3 = {}_1l_2 \cdot {}_2l_3$$

Si deve trovare il numero indice a base mobile medio.

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.



Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.

3. Si calcoli il tasso medio di variazione di x_t tra l'istante temporale 1 e l'istante 3, interpretando il risultato.

Si ricorda che

$${}_1l_3 = {}_1l_2 \cdot {}_2l_3$$

Si deve trovare il numero indice a base mobile medio. Applicando il criterio di scelta della media secondo Chisini si ottiene

$${}_1l_3 = {}_1l_2 \cdot {}_2l_3 = \alpha \cdot \alpha = \alpha^2.$$



Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.

3. Si calcoli il tasso medio di variazione di x_t tra l'istante temporale 1 e l'istante 3, interpretando il risultato.

Si ricorda che

$${}_1l_3 = {}_1l_2 \cdot {}_2l_3$$

Si deve trovare il numero indice a base mobile medio. Applicando il criterio di scelta della media secondo Chisini si ottiene

$${}_1l_3 = {}_1l_2 \cdot {}_2l_3 = \alpha \cdot \alpha = \alpha^2.$$

Quindi

$$\alpha = \sqrt[2]{{}_1l_3} = \sqrt[2]{{}_1l_2 \cdot {}_2l_3} = 0.9487.$$

Da cui segue una variazione relativa media pari a -0.0513 , corrispondente a un decremento medio 1 e il 3 del 5.13%.



Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.

Il seguente prospetto riporta i valori rilevati in corrispondenza a 5 unità statistiche con riferimento a due caratteri quantitativi X e Y

x_i	1	1	3	5	5
y_i	20	40	40	40	20

1. Si rappresentino graficamente le coppie dei dati e la funzione di regressione ipotizzando Y in funzione di X .



La funzione di regressione associa a ogni valore x appartenente all'insieme di definizione della variabile X la media di Y condizionata a x , $M(Y|x)$

x	$M(Y x)$
1	$\frac{1}{2}(20 + 40) = 30$
3	40
5	$\frac{1}{2}(20 + 40) = 30$

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.



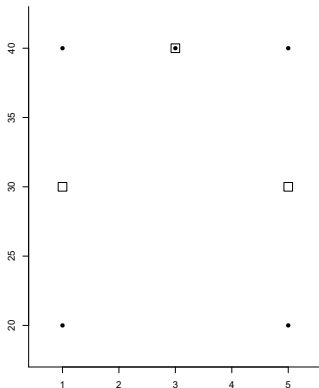
Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.

Le coppie di dati possono essere rappresentate mediante un grafico a dispersione.





2. Si calcolino, secondo il metodo dei minimi quadrati, i parametri e l'indice di adattamento dei seguenti modelli di regressione, individuando il modello migliore:

$$I) Y^* = a + b \cdot X \qquad II) Y^* = c$$

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.



Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.

Si riportano i calcoli necessari¹ per ottenere le stime dei parametri e gli indici di adattamento del modello

$$Y = a + bX + E$$

x_i	y_i	$x_i y_i$	x_i^2	y_i^2
1	20	20	1	400
1	40	40	1	1600
3	40	120	9	1600
5	40	200	25	1600
5	20	100	25	400
15	160	480	61	5600
3	32	96	12.2	1120
$M(X)$	$M(Y)$	$M(XY)$	$M(X^2)$	$M(Y^2)$

$$\hat{b} =$$

¹Si osserva che per ragioni di stampa alcuni calcoli potrebbero essere duplicati con riferimento ai diversi modelli oggetto di studio



Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.

Si riportano i calcoli necessari¹ per ottenere le stime dei parametri e gli indici di adattamento del modello

$$Y = a + bX + E$$

x_i	y_i	$x_i y_i$	x_i^2	y_i^2
1	20	20	1	400
1	40	40	1	1600
3	40	120	9	1600
5	40	200	25	1600
5	20	100	25	400
15	160	480	61	5600
3	32	96	12.2	1120
$M(X)$	$M(Y)$	$M(XY)$	$M(X^2)$	$M(Y^2)$

$$\begin{aligned}\hat{b} &= \frac{\text{Cov}(X, Y)}{\text{Var}(X)} = \frac{M(XY) - M(X)M(Y)}{M(X^2) - [M(X)]^2} = \\ &= \frac{96 - 3 \cdot 32}{12.2 - 3^2} = \frac{0}{3.2} = 0 \\ \hat{a} &= \end{aligned}$$

¹Si osserva che per ragioni di stampa alcuni calcoli potrebbero essere duplicati con riferimento ai diversi modelli oggetto di studio



Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.

Si riportano i calcoli necessari¹ per ottenere le stime dei parametri e gli indici di adattamento del modello

$$Y = a + bX + E$$

x_i	y_i	$x_i y_i$	x_i^2	y_i^2
1	20	20	1	400
1	40	40	1	1600
3	40	120	9	1600
5	40	200	25	1600
5	20	100	25	400
15	160	480	61	5600
3	32	96	12.2	1120
$M(X)$	$M(Y)$	$M(XY)$	$M(X^2)$	$M(Y^2)$

$$\hat{b} = \frac{\text{Cov}(X, Y)}{\text{Var}(X)} = \frac{M(XY) - M(X)M(Y)}{M(X^2) - [M(X)]^2} =$$

$$= \frac{96 - 3 \cdot 32}{12.2 - 3^2} = \frac{0}{3.2} = 0$$

$$\hat{a} = M(Y) - \hat{b}M(X) = 32$$

$$\rho^2 =$$

¹Si osserva che per ragioni di stampa alcuni calcoli potrebbero essere duplicati con riferimento ai diversi modelli oggetto di studio



Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.

Si riportano i calcoli necessari¹ per ottenere le stime dei parametri e gli indici di adattamento del modello

$$Y = a + bX + E$$

x_i	y_i	$x_i y_i$	x_i^2	y_i^2
1	20	20	1	400
1	40	40	1	1600
3	40	120	9	1600
5	40	200	25	1600
5	20	100	25	400
15	160	480	61	5600
3	32	96	12.2	1120
$M(X)$	$M(Y)$	$M(XY)$	$M(X^2)$	$M(Y^2)$

$$\hat{b} = \frac{\text{Cov}(X, Y)}{\text{Var}(X)} = \frac{M(XY) - M(X)M(Y)}{M(X^2) - [M(X)]^2} =$$

$$= \frac{96 - 3 \cdot 32}{12.2 - 3^2} = \frac{0}{3.2} = 0$$

$$\hat{a} = M(Y) - \hat{b}M(X) = 32$$

$$\rho^2 = \frac{[\text{Cov}(X, Y)]^2}{\text{Var}(X) \text{Var}(Y)} = \frac{0^2}{3.2 \cdot 96} = 0$$

¹Si osserva che per ragioni di stampa alcuni calcoli potrebbero essere duplicati con riferimento ai diversi modelli oggetto di studio



Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.

Si riportano i calcoli necessari² per ottenere le stime dei parametri e gli indici di adattamento del modello

$$Y = c + E$$

x_i	y_i	y_i^2
1	20	400
1	40	1600
3	40	1600
5	40	1600
5	20	400
160		5600
32		1120
$M(Y)$		$M(Y^2)$

$$\hat{c} =$$

²Si osserva che per ragioni di stampa alcuni calcoli potrebbero essere duplicati con riferimento ai diversi modelli oggetto di studio



Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.

Si riportano i calcoli necessari² per ottenere le stime dei parametri e gli indici di adattamento del modello

$$Y = c + E$$

x_i	y_i	y_i^2
1	20	400
1	40	1600
3	40	1600
5	40	1600
5	20	400
160		5600
32		1120
$M(Y)$		$M(Y^2)$

$$\hat{c} = M(Y) = 32$$

$$R^2 =$$

²Si osserva che per ragioni di stampa alcuni calcoli potrebbero essere duplicati con riferimento ai diversi modelli oggetto di studio



Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.

Si riportano i calcoli necessari² per ottenere le stime dei parametri e gli indici di adattamento del modello

$$Y = c + E$$

x_i	y_i	y_i^2
1	20	400
1	40	1600
3	40	1600
5	40	1600
5	20	400
160		5600
32		1120
$M(Y)$		$M(Y^2)$

$$\hat{c} = M(Y) = 32$$

$$R^2 = 0$$

²Si osserva che per ragioni di stampa alcuni calcoli potrebbero essere duplicati con riferimento ai diversi modelli oggetto di studio



I due modelli sono equivalenti.

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.



3. Si motivi perché $\eta_{Y|X}^2 > \eta_{X|Y}^2$.

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.



Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.

3. Si motivi perché $\eta_{Y|X}^2 > \eta_{X|Y}^2$.

Dal momento che $M(X|y = 20) = M(X|y = 40) = 3$, sussiste indipendenza in media di X da Y , quindi $\eta_{X|Y}^2 = 0$. $0 < \eta_{Y|X}^2 < 1$ in quanto non sussiste indipendenza in media e nemmeno dipendenza funzionale di Y da X .

Esercizio 4

Prova scritta di
STATISTICA del
13.02.2014

Giuseppe Boari
Gabriele Cantaluppi



UNIVERSITÀ
CATTOLICA
del Sacro Cuore

Un'urna è composta da gettoni rossi e neri in proporzione 0.15 e 0.85. Ogni gettone rosso reca impresso il numero 1, i neri il numero 2. Si consideri l'esperimento casuale $\mathcal{G}_3 =$ 'estrazione con reimmissione di 3 gettoni dall'urna'.

1. Si costruisca lo spazio probabilistico associato all'esperimento $\mathcal{G}_3$.

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.



Un'urna è composta da gettoni rossi e neri in proporzione 0.15 e 0.85. Ogni gettone rosso reca impresso il numero 1, i neri il numero 2. Si consideri l'esperimento casuale $\mathcal{G}_3 =$ 'estrazione con reimmissione di 3 gettoni dall'urna'.

1. Si costruisca lo spazio probabilistico associato all'esperimento $\mathcal{G}_3$.

RRR		111
RRN, RNR, NRR	<i>oppure</i>	$112, 121, 211$
NNR, NRN, RNN		$221, 212, 122$
NNN		222



2. Si costruisca la variabile casuale $X =$ 'somma dei valori impressi sui gettoni estratti' e si calcoli la probabilità che X assuma valori maggiori o uguali a 5.

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.



2. Si costruisca la variabile casuale $X =$ 'somma dei valori impressi sui gettoni estratti' e si calcoli la probabilità che X assuma valori maggiori o uguali a 5.

x	$P_{\Omega}(X^{-1}(x))$	$P(X = x)$
3	$P(RRR)$	0.003375
4	$P(RRN \cup RNR \cup NRR)$	0.057375
5	$P(NNR \cup NRN \cup RNN)$	0.325125
6	$P(NNN)$	0.614125

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.



Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.

2. Si costruisca la variabile casuale $X =$ 'somma dei valori impressi sui gettoni estratti' e si calcoli la probabilità che X assuma valori maggiori o uguali a 5.

x	$P_{\Omega}(X^{-1}(x))$	$P(X = x)$
3	$P(RRR)$	0.003375
4	$P(RRN \cup RNR \cup NRR)$	0.057375
5	$P(NNR \cup NRN \cup RNN)$	0.325125
6	$P(NNN)$	0.614125

$$P(X \geq 5) = 0.325125 + 0.614125 = 0.93925.$$



3. Si calcoli la probabilità di ottenere almeno due gettoni rossi nelle 3 estrazioni.
Sia X la variabile casuale 'n. di gettoni rossi nelle 3 estrazioni (con reimmissione) (prove indipendenti)'.

$$X \sim$$

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.



3. Si calcoli la probabilità di ottenere almeno due gettoni rossi nelle 3 estrazioni.
Sia X la variabile casuale 'n. di gettoni rossi nelle 3 estrazioni (con reimmissione) (prove indipendenti)'.

$$X \sim \text{Bin}(n = 3, p = 0.15).$$

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.



Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.

3. Si calcoli la probabilità di ottenere almeno due gettoni rossi nelle 3 estrazioni.
Sia X la variabile casuale 'n. di gettoni rossi nelle 3 estrazioni (con reimmissione) (prove indipendenti)'.

$$X \sim \text{Bin}(n = 3, p = 0.15).$$

$$\begin{aligned} P(X \geq 2) &= \binom{3}{2} 0.15^2 0.85^1 + \binom{3}{3} 0.15^3 0.85^0 = \\ &= 0.003375 + 0.003375 = 0.00675. \end{aligned}$$



Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.

4. Si calcoli la probabilità di ottenere non più di un gettone rosso in 26 estrazioni con reimmissione.

Sia X la variabile casuale 'n. di gettoni rossi nelle 26 estrazioni (con reimmissione) (prove indipendenti)'.

$$X \sim$$



4. Si calcoli la probabilità di ottenere non più di un gettone rosso in 26 estrazioni con reimmissione.

Sia X la variabile casuale 'n. di gettoni rossi nelle 26 estrazioni (con reimmissione) (prove indipendenti)'.

$$X \sim \text{Bin}(n = 26, p = 0.15).$$

$$\begin{aligned} P(X \leq 1) &= \binom{26}{0} 0.15^0 0.85^{26} + \binom{26}{1} 0.15^1 0.85^{25} = \\ &= 0.01462 + 0.06707 = 0.08169. \end{aligned}$$

Si osserva come non valga l'approssimazione della variabile casuale Binomiale con la variabile casuale Normale in quanto $np = 3.9 < 5$.

Tema T279-1

Prova scritta di STATISTICA del 05.06.2014

Prova scritta di
STATISTICA del
05.06.2014

Giuseppe Boari
Gabriele Cantaluppi



UNIVERSITÀ
CATTOLICA
del Sacro Cuore

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.

Giuseppe Boari
Gabriele Cantaluppi
Università Cattolica del Sacro Cuore, Milano



Si riportano le classi e le densità di frequenza della variabile statistica X :

$h_{i-1} \vdash h_i$	d_i
10 $\vdash$ 22	0.025
22 $\vdash$ 31	0.05
31 $\vdash$ 51	0.0125

1. Si ricostruisca la distribuzione delle frequenze cumulate relative di X e se ne dia una rappresentazione grafica.

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.



Moltiplicando le densità di frequenza, d_i , per le ampiezze di classe, a_i , si ottengono le frequenze relative f_i della variabile X ,

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.



Moltiplicando le densità di frequenza, d_i , per le ampiezze di classe, a_i , si ottengono le frequenze relative f_i della variabile X ,

x_i	d_i	a_i	f_i	F_i
10 + 22	0.025	$22 - 10 = 12$	0.3	0.3
22 + 31	0.05	$31 - 22 = 9$	0.45	0.75
31 + 51	0.0125	$51 - 31 = 20$	0.25	1

Esercizio 1.

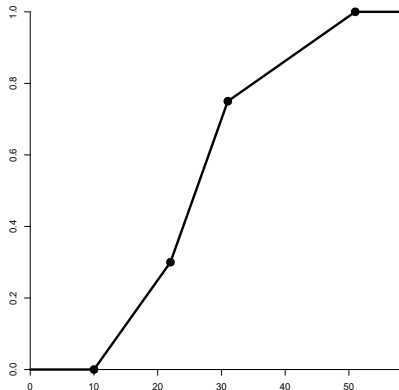
Esercizio 2.

Esercizio 3.

Esercizio 4.



Si tratta di un carattere quantitativo con modalità raggruppate in classi. La funzione di ripartizione ha l'aspetto di una spezzata.



Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.



2. Si rappresenti X mediante un grafico Box & Whiskers Plot.

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.



Si considera la distribuzione della variabile X ,

x_i	f_i
10 – 22	0.3
22 – 31	0.45
31 – 51	0.25

Si tratta di un carattere quantitativo con modalità raggruppate in classi.

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.



Per costruire il grafico Box & Whiskers plot occorre determinare in primo luogo i valori dei tre quartili $x_{0.25}$, $x_{0.5}$ e $x_{0.75}$ considerando l'inversa della Funzione di Ripartizione in corrispondenza rispettivamente di 0.25, 0.50 e 0.75.

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.

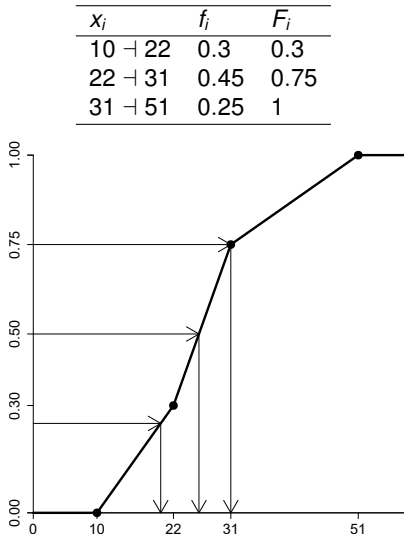


Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.





Si ottiene

$$x_{0.25} = 10 + \frac{0.25 - 0}{0.3 - 0} \cdot (22 - 10) = 20$$

$$x_{0.5} = 22 + \frac{0.5 - 0.3}{0.75 - 0.3} \cdot (31 - 22) = 26$$

$$x_{0.75} = 22 + \frac{0.75 - 0.3}{0.75 - 0.3} \cdot (31 - 22) = 31$$

La lunghezza teorica dei baffi risulta

$$1.5 \cdot (x_{0.75} - x_{0.25}) = 16.5.$$

Il baffo inferiore è posizionato in

$$\max(x_{\min}, x_{0.25} - 16.5) = \max(10, 3.5) = 10.$$

Il baffo superiore è posizionato in

$$\min(x_{\max}, x_{0.75} + 16.5) = \min(51, 47.5) = 47.5.$$

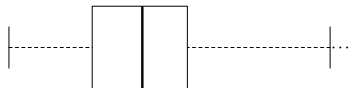
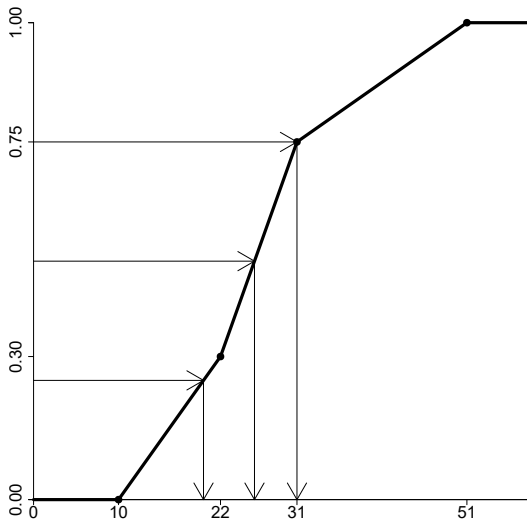


Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.





Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.

3. Si calcolino opportuni indici di posizione e di variabilità per la variabile statistica X , al fine di confrontarla con la variabile Y della quale sono noti solo i seguenti indici sintetici $M(Y) = 50$ e $Var(Y) = 100$.



Si considera la distribuzione della variabile X ,

x_i	f_i
10 $\rightarrow$ 22	0.3
22 $\rightarrow$ 31	0.45
31 $\rightarrow$ 51	0.25

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.



Si tratta di un carattere quantitativo con modalità raggruppate in classi.

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.



La media aritmetica

$$M(X) = \sum_{i=1}^k x_i f_i$$

risulta 26.975.

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.



La media aritmetica

$$M(X) = \sum_{i=1}^k x_i f_i$$

risulta 26.975.

Si ottiene (considerando come modalità i valori centrali delle classi)

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.



La media aritmetica

$$M(X) = \sum_{i=1}^k x_i f_i$$

risulta 26.975.

Si ottiene (considerando come modalità i valori centrali delle classi)

x_i	f_i	$x_i f_i$
16.0	0.30	4.800
26.5	0.45	11.925
41.0	0.25	10.250
	1.00	26.975

da cui segue

$$M(X) = \sum_{i=1}^k x_i f_i = 26.975.$$

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.



La varianza

$$Var(X) = M[(X - \mu_X)^2] = M(X^2) - \mu_X^2 = \sum_{i=1}^k x_i^2 f_i - \mu_X^2$$

risulta 85.4119.

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.



La varianza

$$Var(X) = M[(X - \mu_X)^2] = M(X^2) - \mu_X^2 = \sum_{i=1}^k x_i^2 f_i - \mu_X^2$$

risulta 85.4119.

Si ottiene (considerando come modalità i valori centrali delle classi)

x_i	f_i	$x_i f_i$	$x_i^2 f_i$
16.0	0.30	4.800	76.8000
26.5	0.45	11.925	316.0125
41.0	0.25	10.250	420.2500
	1.00	26.975	813.0625

da cui segue

$$\begin{aligned} Var(X) = \sigma_X^2 &= M(X^2) - \mu_X^2 = \\ &= 813.0625 - 26.975^2 = 813.0625 - 727.6506 = 85.4119. \end{aligned}$$

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.



Per effettuare confronti in termini di variabilità tra caratteri quantitativi occorre utilizzare una misura relativa, ad esempio il coefficiente di variazione

$$CV = \frac{\sigma}{\mu} = \frac{\sqrt{\text{Var}(X)}}{M(X)} = \frac{\sqrt{85.4119}}{26.975} = \frac{9.2419}{26.975} = 0.3426.$$

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.



Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.

indice di posizione: $M(X) = 26.975$ indice di variabilità $CV(X) = 0.3426$

indice di posizione: $M(Y) = 50$ indice di variabilità $CV(Y) = 0.2$

Il livello medio di X è inferiore a quello di Y . X presenta un livello di variabilità superiore a Y .

Esercizio 2

Prova scritta di
STATISTICA del
05.06.2014

Giuseppe Boari
Gabriele Cantaluppi



UNIVERSITÀ
CATTOLICA
del Sacro Cuore

Sia data la seguente serie storica.

t	x_t
1	200
2	221
3	180
4	171
5	160

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.



1. Si calcoli il numero indice al tempo 4 con base 3, commentando il risultato ottenuto.

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.



1. Si calcoli il numero indice al tempo 4 con base 3, commentando il risultato ottenuto.

$${}_3I_4 = \frac{x_4}{x_3} = \frac{171}{180} = 0.95$$

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.



1. Si calcoli il numero indice al tempo 4 con base 3, commentando il risultato ottenuto.

$${}_3I_4 = \frac{x_4}{x_3} = \frac{171}{180} = 0.95$$

Tra il tempo 3 e il tempo 4 si è registrato un decremento complessivo del 5% della grandezza oggetto di studio.

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.



2. Si calcoli il tasso medio di variazione di x_t tra l'istante temporale 2 e l'istante 5, interpretando il risultato.

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.



2. Si calcoli il tasso medio di variazione di x_t tra l'istante temporale 2 e l'istante 5, interpretando il risultato.

Si ricorda che

$${}_2I_5 = {}_2I_3 \cdot {}_3I_4 \cdot {}_4I_5$$

Si deve trovare il numero indice a base mobile medio.

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.



Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.

2. Si calcoli il tasso medio di variazione di x_t tra l'istante temporale 2 e l'istante 5, interpretando il risultato.

Si ricorda che

$${}_2l_5 = {}_2l_3 \cdot {}_3l_4 \cdot {}_4l_5$$

Si deve trovare il numero indice a base mobile medio. Applicando il criterio di scelta della media secondo Chisini si ottiene

$${}_2l_5 = {}_2l_3 \cdot {}_3l_4 \cdot {}_4l_5 = \alpha \cdot \alpha \cdot \alpha = \alpha^3.$$



Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.

2. Si calcoli il tasso medio di variazione di x_t tra l'istante temporale 2 e l'istante 5, interpretando il risultato.

Si ricorda che

$${}_2I_5 = {}_2I_3 \cdot {}_3I_4 \cdot {}_4I_5$$

Si deve trovare il numero indice a base mobile medio. Applicando il criterio di scelta della media secondo Chisini si ottiene

$${}_2I_5 = {}_2I_3 \cdot {}_3I_4 \cdot {}_4I_5 = \alpha \cdot \alpha \cdot \alpha = \alpha^3.$$

Quindi

$$\alpha = \sqrt[3]{{}_2I_5} = \sqrt[3]{{}_2I_3 \cdot {}_3I_4 \cdot {}_4I_5} = 0.8979.$$

Da cui segue una variazione relativa media pari a -0.1021 , corrispondente a un decremento medio 2 e il 5 del 10.21%.



[Esercizio 1.](#)

[Esercizio 2.](#)

[Esercizio 3.](#)

[Esercizio 4.](#)

Il seguente prospetto riporta i valori rilevati in corrispondenza a 5 unità statistiche con riferimento a due caratteri quantitativi X e Y

x_i	1	1	3	5	5
y_i	10	20	30	40	50

1. Si rappresentino graficamente le coppie dei dati e la funzione di regressione ipotizzando Y in funzione di X .



La funzione di regressione associa a ogni valore x appartenente all'insieme di definizione della variabile X la media di Y condizionata a x , $M(Y|x)$

x	$M(Y x)$
1	$\frac{1}{2}(10 + 20) = 15$
3	30
5	$\frac{1}{2}(40 + 50) = 45$

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.



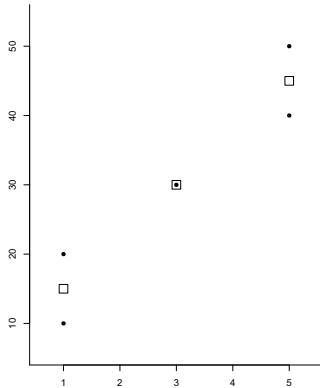
Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.

Le coppie di dati possono essere rappresentate mediante un grafico a dispersione.





2. Si calcolino, secondo il metodo dei minimi quadrati, i parametri e l'indice di adattamento dei seguenti modelli di regressione, individuando il modello migliore:

$$I) Y^* = a + b \cdot X \qquad II) Y^* = cX$$

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.

Si riportano i calcoli necessari¹ per ottenere le stime dei parametri e gli indici di adattamento del modello

$$Y = a + bX + E$$

x_i	y_i	$x_i y_i$	x_i^2	y_i^2
1	10	10	1	100
1	20	20	1	400
3	30	90	9	900
5	40	200	25	1600
5	50	250	25	2500
15	150	570	61	5500
3	30	114	12.2	1100
$M(X)$	$M(Y)$	$M(XY)$	$M(X^2)$	$M(Y^2)$

$$\hat{b} =$$



Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.

¹Si osserva che per ragioni di stampa alcuni calcoli potrebbero essere duplicati con riferimento ai diversi modelli oggetto di studio

Si riportano i calcoli necessari¹ per ottenere le stime dei parametri e gli indici di adattamento del modello

$$Y = a + bX + E$$

x_i	y_i	$x_i y_i$	x_i^2	y_i^2
1	10	10	1	100
1	20	20	1	400
3	30	90	9	900
5	40	200	25	1600
5	50	250	25	2500
15	150	570	61	5500
3	30	114	12.2	1100
$M(X)$	$M(Y)$	$M(XY)$	$M(X^2)$	$M(Y^2)$

$$\begin{aligned}\hat{b} &= \frac{\text{Cov}(X, Y)}{\text{Var}(X)} = \frac{M(XY) - M(X)M(Y)}{M(X^2) - [M(X)]^2} = \\ &= \frac{114 - 3 \cdot 30}{12.2 - 3^2} = \frac{24}{3.2} = 7.5 \\ \hat{a} &= \end{aligned}$$

¹Si osserva che per ragioni di stampa alcuni calcoli potrebbero essere duplicati con riferimento ai diversi modelli oggetto di studio



Si riportano i calcoli necessari¹ per ottenere le stime dei parametri e gli indici di adattamento del modello

$$Y = a + bX + E$$

x_i	y_i	$x_i y_i$	x_i^2	y_i^2
1	10	10	1	100
1	20	20	1	400
3	30	90	9	900
5	40	200	25	1600
5	50	250	25	2500
15	150	570	61	5500
3	30	114	12.2	1100
$M(X)$	$M(Y)$	$M(XY)$	$M(X^2)$	$M(Y^2)$

$$\hat{b} = \frac{\text{Cov}(X, Y)}{\text{Var}(X)} = \frac{M(XY) - M(X)M(Y)}{M(X^2) - [M(X)]^2} =$$

$$= \frac{114 - 3 \cdot 30}{12.2 - 3^2} = \frac{24}{3.2} = 7.5$$

$$\hat{a} = M(Y) - \hat{b}M(X) = 7.5$$

$$\rho^2 =$$



Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.

¹Si osserva che per ragioni di stampa alcuni calcoli potrebbero essere duplicati con riferimento ai diversi modelli oggetto di studio

Si riportano i calcoli necessari¹ per ottenere le stime dei parametri e gli indici di adattamento del modello

$$Y = a + bX + E$$

x_i	y_i	$x_i y_i$	x_i^2	y_i^2
1	10	10	1	100
1	20	20	1	400
3	30	90	9	900
5	40	200	25	1600
5	50	250	25	2500
15	150	570	61	5500
3	30	114	12.2	1100
$M(X)$	$M(Y)$	$M(XY)$	$M(X^2)$	$M(Y^2)$

$$\hat{b} = \frac{\text{Cov}(X, Y)}{\text{Var}(X)} = \frac{M(XY) - M(X)M(Y)}{M(X^2) - [M(X)]^2} =$$

$$= \frac{114 - 3 \cdot 30}{12.2 - 3^2} = \frac{24}{3.2} = 7.5$$

$$\hat{a} = M(Y) - \hat{b}M(X) = 7.5$$

$$\rho^2 = \frac{[\text{Cov}(X, Y)]^2}{\text{Var}(X) \text{Var}(Y)} = \frac{24^2}{3.2 \cdot 200} = 0.9$$

$$M(\hat{E}^2) = \text{Var}(\hat{E}) = \text{varianza residua} =$$

¹Si osserva che per ragioni di stampa alcuni calcoli potrebbero essere duplicati con riferimento ai diversi modelli oggetto di studio



Si riportano i calcoli necessari¹ per ottenere le stime dei parametri e gli indici di adattamento del modello

$$Y = a + bX + E$$

x_i	y_i	$x_i y_i$	x_i^2	y_i^2
1	10	10	1	100
1	20	20	1	400
3	30	90	9	900
5	40	200	25	1600
5	50	250	25	2500
15	150	570	61	5500
3	30	114	12.2	1100
$M(X)$	$M(Y)$	$M(XY)$	$M(X^2)$	$M(Y^2)$

$$\begin{aligned}\hat{b} &= \frac{\text{Cov}(X, Y)}{\text{Var}(X)} = \frac{M(XY) - M(X)M(Y)}{M(X^2) - [M(X)]^2} = \\ &= \frac{114 - 3 \cdot 30}{12.2 - 3^2} = \frac{24}{3.2} = 7.5 \\ \hat{a} &= M(Y) - \hat{b}M(X) = 7.5 \\ \rho^2 &= \frac{[\text{Cov}(X, Y)]^2}{\text{Var}(X) \text{Var}(Y)} = \frac{24^2}{3.2 \cdot 200} = 0.9\end{aligned}$$

$$M(\hat{E}^2) = \text{Var}(\hat{E}) = \text{varianza residua} = \text{Var}(Y) \cdot (1 - \rho^2) = 20$$

¹Si osserva che per ragioni di stampa alcuni calcoli potrebbero essere duplicati con riferimento ai diversi modelli oggetto di studio





Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.

Si riportano i calcoli necessari² per ottenere le stime dei parametri e la misura di adattamento del modello

$$Y = cX + E$$

x_i	y_i	$x_i y_i$	x_i^2	y_i^2
1	10	10	1	100
1	20	20	1	400
3	30	90	9	900
5	40	200	25	1600
5	50	250	25	2500
		570	61	5500
		114	12.2	1100
		$M(XY)$	$M(X^2)$	$M(Y^2)$

$$\hat{c} =$$

²Si osserva che per ragioni di stampa alcuni calcoli potrebbero essere duplicati con riferimento ai diversi modelli oggetto di studio



Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.

Si riportano i calcoli necessari² per ottenere le stime dei parametri e la misura di adattamento del modello

$$Y = cX + E$$

x_i	y_i	$x_i y_i$	x_i^2	y_i^2
1	10	10	1	100
1	20	20	1	400
3	30	90	9	900
5	40	200	25	1600
5	50	250	25	2500
		570	61	5500
		114	12.2	1100
		$M(XY)$	$M(X^2)$	$M(Y^2)$

$$\hat{c} = \frac{M(XY)}{M(X^2)} = \frac{114}{12.2} = 9.3443$$

$$M(\hat{E}^2) =$$

²Si osserva che per ragioni di stampa alcuni calcoli potrebbero essere duplicati con riferimento ai diversi modelli oggetto di studio



Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.

Si riportano i calcoli necessari² per ottenere le stime dei parametri e la misura di adattamento del modello

$$Y = cX + E$$

x_i	y_i	$x_i y_i$	x_i^2	y_i^2
1	10	10	1	100
1	20	20	1	400
3	30	90	9	900
5	40	200	25	1600
5	50	250	25	2500
		570	61	5500
		114	12.2	1100
		$M(XY)$	$M(X^2)$	$M(Y^2)$

$$\hat{c} = \frac{M(XY)}{M(X^2)} = \frac{114}{12.2} = 9.3443$$

$$M(\hat{E}^2) = M[Y \cdot (Y - \hat{c}X)] = M(Y^2) - \hat{c}M(XY) = 1100 - 9.3443 \cdot 114 = 34.7541$$

²Si osserva che per ragioni di stampa alcuni calcoli potrebbero essere duplicati con riferimento ai diversi modelli oggetto di studio



Il modello migliore è $Y^* = 7.5 + 7.5X$.

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.



3. Si indichino i valori assunti da $\eta_{Y|X}^2$ ed $\eta_{X|Y}^2$.

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.



3. Si indichino i valori assunti da $\eta_{Y|X}^2$ ed $\eta_{X|Y}^2$.

$\eta_{Y|X}^2 = \rho^2 = 0.9$ in quanto le medie condizionate $M(Y|X)$ sono allineate

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.



3. Si indichino i valori assunti da $\eta_{Y|X}^2$ ed $\eta_{X|Y}^2$.

$\eta_{Y|X}^2 = \rho^2 = 0.9$ in quanto le medie condizionate $M(Y|X)$ sono allineate

$\eta_{X|Y}^2 = 1$ in quanto sussiste dipendenza funzionale di X da Y .

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.

Esercizio 4

Per il seguente gioco è prevista una posta di 2 Euro: si lanciano due dadi e se la somma dei numeri ottenuti è pari a 6 si vincono 6 Euro.

1. Si stabilisca se il gioco è equo.



Esercizio 1.

Esercizio 2.

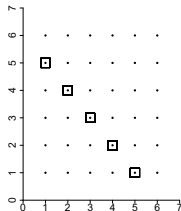
Esercizio 3.

Esercizio 4.

Esercizio 4

Per il seguente gioco è prevista una posta di 2 Euro: si lanciano due dadi e se la somma dei numeri ottenuti è pari a 6 si vincono 6 Euro.

1. Si stabilisca se il gioco è equo.



Esercizio 1.

Esercizio 2.

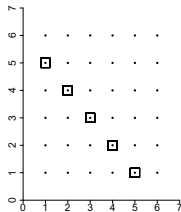
Esercizio 3.

Esercizio 4.

Esercizio 4

Per il seguente gioco è prevista una posta di 2 Euro: si lanciano due dadi e se la somma dei numeri ottenuti è pari a 6 si vincono 6 Euro.

1. Si stabilisca se il gioco è equo.



$$P(\text{vincita}) = \frac{5}{36} = 0.1389$$



Esercizio 1.

Esercizio 2.

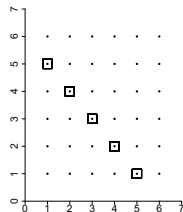
Esercizio 3.

Esercizio 4.

Esercizio 4

Per il seguente gioco è prevista una posta di 2 Euro: si lanciano due dadi e se la somma dei numeri ottenuti è pari a 6 si vincono 6 Euro.

1. Si stabilisca se il gioco è equo.



$$P(\text{vincita}) = \frac{5}{36} = 0.1389$$

	x_i	$P(X = x_i)$	$x_i \cdot P(X = x_i)$
<i>posta</i>	-2	0.8611	-1.7222
<i>vincita - posta</i>	4	0.1389	0.5556
		1	-1.1667



Esercizio 1.

Esercizio 2.

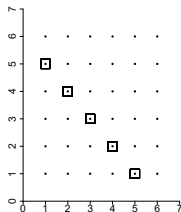
Esercizio 3.

Esercizio 4.

Esercizio 4

Per il seguente gioco è prevista una posta di 2 Euro: si lanciano due dadi e se la somma dei numeri ottenuti è pari a 6 si vincono 6 Euro.

1. Si stabilisca se il gioco è equo.



$$P(\text{vincita}) = \frac{5}{36} = 0.1389$$

	x_i	$P(X = x_i)$	$x_i \cdot P(X = x_i)$
<i>posta</i>	-2	0.8611	-1.7222
<i>vincita - posta</i>	4	0.1389	0.5556
	1		-1.1667

Il valore atteso della variabile casuale guadagno è negativo, quindi il gioco non è equo.



Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.



2. Nell'ipotesi di ripetere il gioco 8 volte si calcoli la probabilità di vincere almeno 3 volte. Sia X la variabile casuale 'n. di vincite nelle 8 prove indipendenti'.

$$X \sim$$

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.



2. Nell'ipotesi di ripetere il gioco 8 volte si calcoli la probabilità di vincere almeno 3 volte. Sia X la variabile casuale 'n. di vincite nelle 8 prove indipendenti'.

$$X \sim \text{Bin}(n = 8, p = 0.1389).$$

$$P(X \geq 3) =$$

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.



Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.

2. Nell'ipotesi di ripetere il gioco 8 volte si calcoli la probabilità di vincere almeno 3 volte. Sia X la variabile casuale 'n. di vincite nelle 8 prove indipendenti'.

$$X \sim \text{Bin}(n = 8, p = 0.1389).$$

$$\begin{aligned} P(X \geq 3) &= 1 - P(X < 3) = 1 - \{P(X = 0) + P(X = 1) + P(X = 2)\} = \\ &= \end{aligned}$$



2. Nell'ipotesi di ripetere il gioco 8 volte si calcoli la probabilità di vincere almeno 3 volte. Sia X la variabile casuale 'n. di vincite nelle 8 prove indipendenti'.

$$X \sim \text{Bin}(n = 8, p = 0.1389).$$

$$\begin{aligned} P(X \geq 3) &= 1 - P(X < 3) = 1 - \{P(X = 0) + P(X = 1) + P(X = 2)\} = \\ &= 1 - \binom{8}{0} 0.1389^0 0.8611^8 - \binom{8}{1} 0.1389^1 0.8611^7 - \binom{8}{2} 0.1389^2 0.8611^6 = \\ &= 1 - 0.3023 - 0.3901 - 0.2202 = 0.0874. \end{aligned}$$

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.



3. Nell'ipotesi di ripetere il gioco 200 volte si calcoli la probabilità di vincere almeno 30 volte. Indicata con X la variabile casuale 'numero di volte che si presenta l'evento vincita nelle 200 ripetizioni del gioco (prove indipendenti) risulta

$$X \sim$$

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.



Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.

3. Nell'ipotesi di ripetere il gioco 200 volte si calcoli la probabilità di vincere almeno 30 volte. Indicata con X la variabile casuale 'numero di volte che si presenta l'evento vincita nelle 200 ripetizioni del gioco (prove indipendenti) risulta

$$X \sim \text{Bin}(n = 200, p = 0.1389).$$

$M(X) = 27.7778$; $\text{Var}(X) = 23.9198$; utilizzando l'approssimazione della variabile casuale Binomiale con la variabile casuale Normale

$$X \sim$$



3. Nell'ipotesi di ripetere il gioco 200 volte si calcoli la probabilità di vincere almeno 30 volte. Indicata con X la variabile casuale 'numero di volte che si presenta l'evento vincita nelle 200 ripetizioni del gioco (prove indipendenti) risulta

$$X \sim \text{Bin}(n = 200, p = 0.1389).$$

$M(X) = 27.7778$; $\text{Var}(X) = 23.9198$; utilizzando l'approssimazione della variabile casuale Binomiale con la variabile casuale Normale

$$X \sim N(\mu = 27.7778, \sigma^2 = 23.9198)$$

si ottiene

$$\begin{aligned} P(X \geq 30) &= P\left(\frac{X - \mu}{\sigma} \geq \frac{30 - \mu}{\sigma}\right) = P\left(Z \geq \frac{30 - 27.7778}{4.8908}\right) = \\ &= P(Z \geq 0.45) = 1 - P(Z \leq 0.45) = 1 - 0.6736 = 0.3264. \end{aligned}$$

Tema T279

Prova scritta di STATISTICA del 26.06.2014

Prova scritta di
STATISTICA del
26.06.2014

Giuseppe Boari
Gabriele Cantaluppi



UNIVERSITÀ
CATTOLICA
del Sacro Cuore

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.

Giuseppe Boari
Gabriele Cantaluppi
Università Cattolica del Sacro Cuore, Milano



La seguente tabella riporta i dati, relativi ad un anno solare, per un campione di 9 Stati UE classificati sulla base delle variabili X : area geografica, con modalità CN = Centro Nord e CS = Centro Sud e Y : PIL pro capite in migliaia di euro.

X	CN	CN	CN	CN	CN	CS	CS	CS	CS
Y	41.9	30.5	32.1	49.1	28.8	25.3	22.2	32.7	21.8

1. Si confronti la variabilità della variabile Y nei gruppi definiti dalla variabile X (CN e CS).

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.



Per effettuare un confronto in termini di posizione della variabile Y nei 2 gruppi CN e CS si considerano le due medie aritmetiche di gruppo μ_{CN} e μ_{CS} :

$$\mu_{CN} = \frac{1}{5}(41.9 + 30.5 + 32.1 + 49.1 + 28.8) = 36.48$$

e

$$\mu_{CS} = \frac{1}{4}(25.3 + 22.2 + 32.7 + 21.8) = 25.5$$

Il livello medio nel gruppo CN è superiore a quello osservato nel gruppo CS .

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.



Per effettuare un confronto in termini di variabilità della variabile Y nei 2 gruppi CN e CS si considerano i coefficienti di variazione calcolati per i due gruppi $CV(Y_{CN})$ e $CV(Y_{CS})$.

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.



Per effettuare un confronto in termini di variabilità della variabile Y nei 2 gruppi CN e CS si considerano i coefficienti di variazione calcolati per i due gruppi $CV(Y_{CN})$ e $CV(Y_{CS})$. A tal fine occorre in primo luogo calcolare le varianze nei due gruppi

$$\sigma_{CN}^2 = \frac{1}{5}(41.9^2 + 30.5^2 + 32.1^2 + 49.1^2 + 28.8^2) - \mu_{CN}^2 = 60.5136$$

e

$$\sigma_{CS}^2 = \frac{1}{4}(25.3^2 + 22.2^2 + 32.7^2 + 21.8^2) - \mu_{CS}^2 = 19.115$$

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.



Per effettuare un confronto in termini di variabilità della variabile Y nei 2 gruppi CN e CS si considerano i coefficienti di variazione calcolati per i due gruppi $CV(Y_{CN})$ e $CV(Y_{CS})$. A tal fine occorre in primo luogo calcolare le varianze nei due gruppi

$$\sigma_{CN}^2 = \frac{1}{5}(41.9^2 + 30.5^2 + 32.1^2 + 49.1^2 + 28.8^2) - \mu_{CN}^2 = 60.5136$$

e

$$\sigma_{CS}^2 = \frac{1}{4}(25.3^2 + 22.2^2 + 32.7^2 + 21.8^2) - \mu_{CS}^2 = 19.115$$

Ricordando che il coefficiente di variazione è definito dal rapporto tra scarto quadratico medio e media aritmetica risulta

$$CV(Y_{CN}) = \frac{\sigma_{CN}}{\mu_{CN}} = \frac{7.779048785}{36.48} = 0.2132$$

e

$$CV(Y_{CS}) = \frac{\sigma_{CS}}{\mu_{CS}} = \frac{4.372070448}{25.5} = 0.1715$$

Il livello di variabilità presente nel gruppo CN è superiore a quello osservato nel gruppo CS .



Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.

2. Si calcolino la varianza between e la varianza within. Si commentino i risultati ottenuti.



Applicando il risultato di scomposizione della varianza si ottiene:

$$\sigma_{Between}^2 =$$

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.



Applicando il risultato di scomposizione della varianza si ottiene:

$$\begin{aligned}\sigma_{Between}^2 &= \frac{1}{9} \left\{ (\mu_{CN} - \mu)^2 \cdot 5 + (\mu_{CS} - \mu)^2 \cdot 4 \right\} \\ &= \frac{1}{9} \left(\mu_{CN}^2 \cdot 5 + \mu_{CS}^2 \cdot 4 \right) - \mu^2 \\ &= \frac{1}{9} \left(36.48^2 \cdot 5 + 25.5^2 \cdot 4 \right) - 31.6^2 = 29.768\end{aligned}$$

$$\sigma_{Within}^2 =$$

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.



Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.

Applicando il risultato di scomposizione della varianza si ottiene:

$$\begin{aligned}\sigma_{Between}^2 &= \frac{1}{9} \left\{ (\mu_{CN} - \mu)^2 \cdot 5 + (\mu_{CS} - \mu)^2 \cdot 4 \right\} \\ &= \frac{1}{9} \left(\mu_{CN}^2 \cdot 5 + \mu_{CS}^2 \cdot 4 \right) - \mu^2 \\ &= \frac{1}{9} \left(36.48^2 \cdot 5 + 25.5^2 \cdot 4 \right) - 31.6^2 = 29.768\end{aligned}$$

$$\begin{aligned}\sigma_{Within}^2 &= \frac{1}{9} \left(\sigma_{CN}^2 \cdot 5 + \sigma_{CS}^2 \cdot 4 \right) \\ &= \frac{1}{9} (60.5136 \cdot 5 + 19.115 \cdot 4) = 42.1142.\end{aligned}$$

e

$$\sigma^2 =$$



Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.

Applicando il risultato di scomposizione della varianza si ottiene:

$$\begin{aligned}\sigma_{Between}^2 &= \frac{1}{9} \left\{ (\mu_{CN} - \mu)^2 \cdot 5 + (\mu_{CS} - \mu)^2 \cdot 4 \right\} \\ &= \frac{1}{9} \left(\mu_{CN}^2 \cdot 5 + \mu_{CS}^2 \cdot 4 \right) - \mu^2 \\ &= \frac{1}{9} \left(36.48^2 \cdot 5 + 25.5^2 \cdot 4 \right) - 31.6^2 = 29.768\end{aligned}$$

$$\begin{aligned}\sigma_{Within}^2 &= \frac{1}{9} \left(\sigma_{CN}^2 \cdot 5 + \sigma_{CS}^2 \cdot 4 \right) \\ &= \frac{1}{9} (60.5136 \cdot 5 + 19.115 \cdot 4) = 42.1142.\end{aligned}$$

e

$$\sigma^2 = \sigma_{Between}^2 + \sigma_{Within}^2 = 29.768 + 42.1142 = 71.8822.$$



Il rapporto di correlazione $\eta^2_{X|Y}$ risulta:

$$\eta^2_{X|Y} = \frac{\sigma^2_{Between}}{\sigma^2_{Between} + \sigma^2_{Within}} = 0.4141.$$

Sussiste, quindi, un livello di dipendenza basso.

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.



Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.

3. Si costruisca la distribuzione di frequenze della variabile Y con riferimento alle classi $(20 \rightarrow 30, 30 \rightarrow 40, 40 \rightarrow 50)$ e si rappresenti graficamente tale distribuzione.



Occorre effettuare lo spoglio dei dati (ricordando che le classi sono chiuse a destra)

y_i	n_i
20 - 30	4
30 - 40	3
40 - 50	2

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.



Si tratta di un carattere quantitativo con modalità raggruppate in classi che può essere rappresentato graficamente mediante un istogramma.

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.



L'istogramma può essere costruito dopo avere calcolato le densità $d_i = \frac{n_i}{a_i}$, (dove a_i sono le ampiezze di classe).

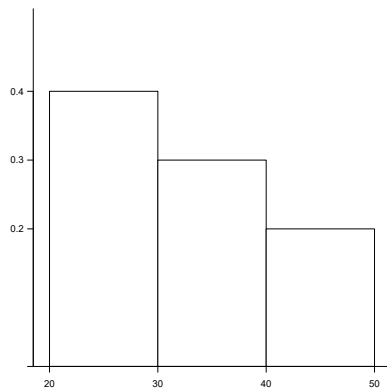
y_i	n_i	a_i	d_i
20 $\vdash$ 30	4	30 - 20 = 10	0.4
30 $\vdash$ 40	3	40 - 30 = 10	0.3
40 $\vdash$ 50	2	50 - 40 = 10	0.2

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.



Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.



4. Costruita la tabella a doppia entrata (X, Y) , con Y riclassificata come nel punto 3, si calcoli un opportuno indice di connessione tra X e Y .

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.



Dalla distribuzione delle frequenze congiunte n_{ij} :

	20 ÷ 30	30 ÷ 40	40 ÷ 50	$n_{i\bullet}$
CN	1	2	2	5
CS	3	1	0	4
$n_{\bullet j}$	4	3	2	9

è possibile ottenere le frequenze teoriche nella situazione di indipendenza stocastica

$$\hat{n}_{ij} = \frac{n_{i\bullet} n_{\bullet j}}{n}:$$

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.



Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.

Dalla distribuzione delle frequenze congiunte n_{ij} :

	20 ÷ 30	30 ÷ 40	40 ÷ 50	$n_{i\bullet}$
CN	1	2	2	5
CS	3	1	0	4
$n_{\bullet j}$	4	3	2	9

è possibile ottenere le frequenze teoriche nella situazione di indipendenza stocastica

$$\hat{n}_{ij} = \frac{n_{i\bullet} \cdot n_{\bullet j}}{n}:$$

	20 ÷ 30	30 ÷ 40	40 ÷ 50	$n_{i\bullet}$
CN	2.2222	1.6667	1.1111	5
CS	1.7778	1.3333	0.8889	4
$n_{\bullet j}$	4.0000	3.0000	2.0000	9



le contingenze assolute $n_{ij} - \hat{n}_{ij}$:

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.



le contingenze assolute $n_{ij} - \hat{n}_{ij}$:

	20 ÷ 30	30 ÷ 40	40 ÷ 50
CN	-1.2222	0.3333	0.8889
CS	1.2222	-0.3333	-0.8889

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.



i rapporti di contingenza $\frac{n_{ij}}{\hat{n}_{ij}}$:

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.



i rapporti di contingenza $\frac{n_{ij}}{\hat{n}_{ij}}$:

	20 ÷ 30	30 ÷ 40	40 ÷ 50
CN	0.4500	1.20	1.8
CS	1.6875	0.75	0.0

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.



e le contingenze relative, espresse in termini percentuali, $100 \left(\frac{n_{ij} - \hat{n}_{ij}}{\hat{n}_{ij}} \right) \% = 100 \left(\frac{n_{ij}}{\hat{n}_{ij}} - 1 \right) \%$:

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.



e le contingenze relative, espresse in termini percentuali, $100 \left(\frac{n_{ij} - \hat{n}_{ij}}{\hat{n}_{ij}} \right) \% = 100 \left(\frac{n_{ij}}{\hat{n}_{ij}} - 1 \right) \%$:

	20 - 30	30 - 40	40 - 50
CN	-55.00	20	80
CS	68.75	-25	-100

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.



Risultati intermedi $\frac{(n_{ij} - \hat{n}_{ij})^2}{\hat{n}_{ij}}$ per il calcolo del χ^2

	20 ÷ 30	30 ÷ 40	40 ÷ 50
CN	0.6722	0.0667	0.7111
CS	0.8403	0.0833	0.8889

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.



Risultati intermedi $\frac{n_{ij}^2}{n_{i\bullet} \cdot n_{\bullet j}}$ per il calcolo del χ^2 (formula operativa)

	20 ÷ 30	30 ÷ 40	40 ÷ 50
CN	0.0500	0.2667	0.4
CS	0.5625	0.0833	0.0

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.



Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.

$$\chi^2 = \sum_{i=1}^h \sum_{j=1}^k \frac{(n_{ij} - \hat{n}_{ij})^2}{\hat{n}_{ij}} = 3.2625,$$

$$\chi_N^2 = \frac{\chi^2}{n \cdot \min(h-1, k-1)} = 0.3625$$

$$\chi_N = \sqrt{\chi_N^2} = 0.6021.$$

Sussiste, quindi, un livello di connessione medio.



Si consideri la seguente serie dei numeri indici a base fissa, con base 2009:

<i>anno</i>	<i>NIBF</i>
2009	1
2010	0.90
2011	0.93
2012	1.02
2013	1.11

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.



1. Si costruisca la corrispondente serie dei numeri indici a base mobile.

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.



1. Si costruisca la corrispondente serie dei numeri indici a base mobile.

Si ricorda che ${}_{t-1}I_t = \frac{x_t}{x_{t-1}} = \frac{{}_{2009}I_t}{{}_{2009}I_{t-1}}$.

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.



1. Si costruisca la corrispondente serie dei numeri indici a base mobile.

Si ricorda che ${}_t I_t = \frac{x_t}{x_{t-1}} = \frac{{}_{2009}I_t}{{}_{2009}I_{t-1}}$. Si ottiene quindi

<i>Anno</i>	<i>NIBF</i>	<i>NIBM</i>
2009	1.00	
2010	0.90	0.9000
2011	0.93	1.0333
2012	1.02	1.0968
2013	1.11	1.0882

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.



Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.

2. Sapendo che $x_{2012} = 265$, si calcoli x_{2013} .



Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.

2. Sapendo che $x_{2012} = 265$, si calcoli x_{2013} .

$$x_{2013} = x_{2012} \cdot l_{2013} = 265 \cdot 1.0882 = 288.3824$$



3. Si calcoli il tasso medio di variazione dal 2009 al 2013 e si commenti il risultato ottenuto.

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.



3. Si calcoli il tasso medio di variazione dal 2009 al 2013 e si commenti il risultato ottenuto.
Si ricorda che

$${}_{2009}I_{2013} = {}_{2009}I_{2010} \cdot {}_{2010}I_{2011} \cdot {}_{2011}I_{2012} \cdot {}_{2012}I_{2013}$$

Si deve trovare il numero indice a base mobile medio.

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.



Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.

3. Si calcoli il tasso medio di variazione dal 2009 al 2013 e si commenti il risultato ottenuto. Si ricorda che

$${}_{2009}I_{2013} = {}_{2009}I_{2010} \cdot {}_{2010}I_{2011} \cdot {}_{2011}I_{2012} \cdot {}_{2012}I_{2013}$$

Si deve trovare il numero indice a base mobile medio. Applicando il criterio di scelta della media secondo Chisini si ottiene

$${}_{2009}I_{2013} = {}_{2009}I_{2010} \cdot {}_{2010}I_{2011} \cdot {}_{2011}I_{2012} \cdot {}_{2012}I_{2013} = \alpha \cdot \alpha \cdot \alpha \cdot \alpha = \alpha^4.$$



Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.

3. Si calcoli il tasso medio di variazione dal 2009 al 2013 e si commenti il risultato ottenuto. Si ricorda che

$${}_{2009}I_{2013} = {}_{2009}I_{2010} \cdot {}_{2010}I_{2011} \cdot {}_{2011}I_{2012} \cdot {}_{2012}I_{2013}$$

Si deve trovare il numero indice a base mobile medio. Applicando il criterio di scelta della media secondo Chisini si ottiene

$${}_{2009}I_{2013} = {}_{2009}I_{2010} \cdot {}_{2010}I_{2011} \cdot {}_{2011}I_{2012} \cdot {}_{2012}I_{2013} = \alpha \cdot \alpha \cdot \alpha \cdot \alpha = \alpha^4.$$

Quindi

$$\alpha = \sqrt[4]{{}_{2009}I_{2013}} = \sqrt[4]{{}_{2009}I_{2010} \cdot {}_{2010}I_{2011} \cdot {}_{2011}I_{2012} \cdot {}_{2012}I_{2013}} = 1.0264.$$

Da cui segue una variazione relativa media pari a 0.0264, corrispondente a un incremento medio annuo tra il 2009 e il 2013 del 2.64%.



[Esercizio 1.](#)

[Esercizio 2.](#)

[Esercizio 3.](#)

[Esercizio 4.](#)

Il seguente prospetto riporta i valori rilevati in corrispondenza a 8 unità statistiche con riferimento a due caratteri quantitativi X e Y

x_i	0	-1	0	1	-1	0	-1	-1
y_i	15	8	15	8	8	15	8	8

1. Si rappresentino graficamente le coppie dei dati e la funzione di regressione ipotizzando Y in funzione di X .



La funzione di regressione associa a ogni valore x appartenente all'insieme di definizione della variabile X la media di Y condizionata a x , $M(Y|x)$

x	$M(Y x)$
-1	8
0	15
1	8

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.



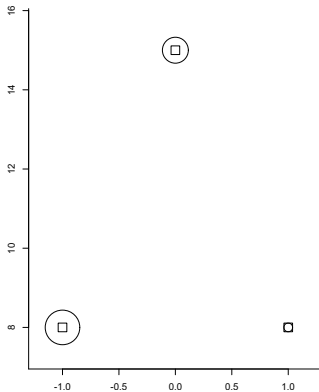
Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.

Le coppie di dati possono essere rappresentate mediante un *bubble plot diagram*.





2. Si calcolino, secondo il metodo dei minimi quadrati, i parametri e l'indice di adattamento dei seguenti modelli di regressione, individuando il modello migliore:

$$I) Y^* = a + bX^2$$

$$II) Y^* = \alpha + \beta X^3$$

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.

Si riportano i calcoli necessari¹ per ottenere le stime dei parametri e gli indici di adattamento del modello

$$Y = a + bX^2 + E$$

x_i^2	y_i	$x_i^2 y_i$	$x_i^{2^2}$	y_i^2
0	15	0	0	225
1	8	8	1	64
0	15	0	0	225
1	8	8	1	64
1	8	8	1	64
0	15	0	0	225
1	8	8	1	64
1	8	8	1	64
5	85	40	5	995
0.625	10.625	5	0.625	124.375
$M(X^2)$	$M(Y)$	$M(X^2 Y)$	$M(X^{2^2})$	$M(Y^2)$

$$\hat{b} =$$

¹Si osserva che per ragioni di stampa alcuni calcoli potrebbero essere duplicati con riferimento ai diversi modelli oggetto di studio



Si riportano i calcoli necessari¹ per ottenere le stime dei parametri e gli indici di adattamento del modello

$$Y = a + bX^2 + E$$

x_i^2	y_i	$x_i^2 y_i$	$x_i^{2^2}$	y_i^2
0	15	0	0	225
1	8	8	1	64
0	15	0	0	225
1	8	8	1	64
1	8	8	1	64
0	15	0	0	225
1	8	8	1	64
1	8	8	1	64
5	85	40	5	995
0.625	10.625	5	0.625	124.375
$M(X^2)$	$M(Y)$	$M(X^2 Y)$	$M(X^{2^2})$	$M(Y^2)$

$$\begin{aligned}\hat{b} &= \frac{\text{Cov}(X^2, Y)}{\text{Var}(X^2)} = \frac{M(X^2 Y) - M(X^2) M(Y)}{M(X^{2^2}) - [M(X^2)]^2} = \\ &= \frac{5 - 0.625 \cdot 10.625}{0.625 - 0.625^2} = \frac{-1.6406}{0.2344} = -6.9991 \\ \hat{a} &= \end{aligned}$$

¹Si osserva che per ragioni di stampa alcuni calcoli potrebbero essere duplicati con riferimento ai diversi modelli oggetto di studio



Si riportano i calcoli necessari¹ per ottenere le stime dei parametri e gli indici di adattamento del modello

$$Y = a + bX^2 + E$$

x_i^2	y_i	$x_i^2 y_i$	$x_i^{2^2}$	y_i^2
0	15	0	0	225
1	8	8	1	64
0	15	0	0	225
1	8	8	1	64
1	8	8	1	64
0	15	0	0	225
1	8	8	1	64
1	8	8	1	64
5	85	40	5	995
0.625	10.625	5	0.625	124.375
$M(X^2)$	$M(Y)$	$M(X^2 Y)$	$M(X^{2^2})$	$M(Y^2)$

$$\hat{b} = \frac{\text{Cov}(X^2, Y)}{\text{Var}(X^2)} = \frac{M(X^2 Y) - M(X^2) M(Y)}{M(X^{2^2}) - [M(X^2)]^2} =$$

$$= \frac{5 - 0.625 \cdot 10.625}{0.625 - 0.625^2} = \frac{-1.6406}{0.2344} = -6.9991$$

$$\hat{a} = M(Y) - \hat{b} M(X^2) = 14.9995$$

$$\rho^2 =$$

¹Si osserva che per ragioni di stampa alcuni calcoli potrebbero essere duplicati con riferimento ai diversi modelli oggetto di studio



Si riportano i calcoli necessari¹ per ottenere le stime dei parametri e gli indici di adattamento del modello

$$Y = a + bX^2 + E$$

x_i^2	y_i	$x_i^2 y_i$	$x_i^{2^2}$	y_i^2
0	15	0	0	225
1	8	8	1	64
0	15	0	0	225
1	8	8	1	64
1	8	8	1	64
0	15	0	0	225
1	8	8	1	64
1	8	8	1	64
5	85	40	5	995
0.625	10.625	5	0.625	124.375
$M(X^2)$	$M(Y)$	$M(X^2 Y)$	$M(X^{2^2})$	$M(Y^2)$

$$\hat{b} = \frac{\text{Cov}(X^2, Y)}{\text{Var}(X^2)} = \frac{M(X^2 Y) - M(X^2) M(Y)}{M(X^{2^2}) - [M(X^2)]^2} =$$

$$= \frac{5 - 0.625 \cdot 10.625}{0.625 - 0.625^2} = \frac{-1.6406}{0.2344} = -6.9991$$

$$\hat{a} = M(Y) - \hat{b} M(X^2) = 14.9995$$

$$\rho^2 = \frac{[\text{Cov}(X^2, Y)]^2}{\text{Var}(X^2) \text{Var}(Y)} = \frac{(-1.6406)^2}{0.2344 \cdot 11.4844} = 0.9999$$

¹Si osserva che per ragioni di stampa alcuni calcoli potrebbero essere duplicati con riferimento ai diversi modelli oggetto di studio



Si riportano i calcoli necessari² per ottenere le stime dei parametri e gli indici di adattamento del modello

$$Y = \alpha + \beta X^3 + E$$

x_i^3	y_i	$x_i^3 y_i$	$x_i^{3^2}$	y_i^2
0	15	0	0	225
-1	8	-8	1	64
0	15	0	0	225
1	8	8	1	64
-1	8	-8	1	64
0	15	0	0	225
-1	8	-8	1	64
-1	8	-8	1	64
-3	85	-24	5	995
-0.375	10.625	-3	0.625	124.375
$M(X^3)$	$M(Y)$	$M(X^3 Y)$	$M(X^{3^2})$	$M(Y^2)$

$$\hat{\beta} =$$



Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.

²Si osserva che per ragioni di stampa alcuni calcoli potrebbero essere duplicati con riferimento ai diversi modelli oggetto di studio

Si riportano i calcoli necessari² per ottenere le stime dei parametri e gli indici di adattamento del modello

$$Y = \alpha + \beta X^3 + E$$

x_i^3	y_i	$x_i^3 y_i$	$x_i^{3^2}$	y_i^2
0	15	0	0	225
-1	8	-8	1	64
0	15	0	0	225
1	8	8	1	64
-1	8	-8	1	64
0	15	0	0	225
-1	8	-8	1	64
-1	8	-8	1	64
-3	85	-24	5	995
-0.375	10.625	-3	0.625	124.375
$M(X^3)$	$M(Y)$	$M(X^3 Y)$	$M(X^{3^2})$	$M(Y^2)$

$$\begin{aligned}\hat{\beta} &= \frac{\text{Cov}(X^3, Y)}{\text{Var}(X^3)} = \frac{M(X^3 Y) - M(X^3) M(Y)}{M(X^{3^2}) - [M(X^3)]^2} = \\ &= \frac{-3 - (-0.375) \cdot 10.625}{0.625 - (-0.375)^2} = \frac{0.9844}{0.4844} = 2.0322 \\ \hat{\alpha} &= \end{aligned}$$

²Si osserva che per ragioni di stampa alcuni calcoli potrebbero essere duplicati con riferimento ai diversi modelli oggetto di studio



Si riportano i calcoli necessari² per ottenere le stime dei parametri e gli indici di adattamento del modello

$$Y = \alpha + \beta X^3 + E$$

x_i^3	y_i	$x_i^3 y_i$	$x_i^{3^2}$	y_i^2
0	15	0	0	225
-1	8	-8	1	64
0	15	0	0	225
1	8	8	1	64
-1	8	-8	1	64
0	15	0	0	225
-1	8	-8	1	64
-1	8	-8	1	64
-3	85	-24	5	995
-0.375	10.625	-3	0.625	124.375
$M(X^3)$	$M(Y)$	$M(X^3 Y)$	$M(X^{3^2})$	$M(Y^2)$

$$\hat{\beta} = \frac{\text{Cov}(X^3, Y)}{\text{Var}(X^3)} = \frac{M(X^3 Y) - M(X^3) M(Y)}{M(X^{3^2}) - [M(X^3)]^2} =$$

$$= \frac{-3 - (-0.375) \cdot 10.625}{0.625 - (-0.375)^2} = \frac{0.9844}{0.4844} = 2.0322$$

$$\hat{\alpha} = M(Y) - \hat{\beta} M(X^3) = 11.3871$$

$$\rho^2 =$$

²Si osserva che per ragioni di stampa alcuni calcoli potrebbero essere duplicati con riferimento ai diversi modelli oggetto di studio



Si riportano i calcoli necessari² per ottenere le stime dei parametri e gli indici di adattamento del modello

$$Y = \alpha + \beta X^3 + E$$

x_i^3	y_i	$x_i^3 y_i$	$x_i^{3^2}$	y_i^2
0	15	0	0	225
-1	8	-8	1	64
0	15	0	0	225
1	8	8	1	64
-1	8	-8	1	64
0	15	0	0	225
-1	8	-8	1	64
-1	8	-8	1	64
-3	85	-24	5	995
-0.375	10.625	-3	0.625	124.375
$M(X^3)$	$M(Y)$	$M(X^3 Y)$	$M(X^{3^2})$	$M(Y^2)$

$$\hat{\beta} = \frac{\text{Cov}(X^3, Y)}{\text{Var}(X^3)} = \frac{M(X^3 Y) - M(X^3) M(Y)}{M(X^{3^2}) - [M(X^3)]^2} =$$

$$= \frac{-3 - (-0.375) \cdot 10.625}{0.625 - (-0.375)^2} = \frac{0.9844}{0.4844} = 2.0322$$

$$\hat{\alpha} = M(Y) - \hat{\beta} M(X^3) = 11.3871$$

$$\rho^2 = \frac{[\text{Cov}(X^3, Y)]^2}{\text{Var}(X^3) \text{Var}(Y)} = \frac{0.9844^2}{0.4844 \cdot 11.4844} = 0.1742$$

²Si osserva che per ragioni di stampa alcuni calcoli potrebbero essere duplicati con riferimento ai diversi modelli oggetto di studio





Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.

Il modello migliore è $Y^* = 15 - 7X^2$.



3. Per il modello migliore, si fornisca la previsione in corrispondenza di $X = 2$.

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.



3. Per il modello migliore, si fornisca la previsione in corrispondenza di $X = 2$.
 $Y^*(2) = 15 - 7 \cdot 2^2 = 15 - 7 \cdot 4 = 15 - 28 = -13$.

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.



Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.

4. Si determini il valore del rapporto di correlazione $\eta^2_{Y|X}$ motivando la risposta.



4. Si determini il valore del rapporto di correlazione $\eta_{Y|X}^2$ motivando la risposta.
 $\eta_{Y|X}^2 = 1$ in quanto sussiste dipendenza funzionale di Y da X .

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.

Esercizio 1



Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.

Si considerino 500 studenti di una Facoltà ripartiti secondo tre diversi corsi di studio (A, B, C) e secondo lo stato (in corso = IC, fuori corso = FC):

	A	B	C
IC	200	130	100
FC	35	25	10

1. Si calcoli la probabilità che, estraendo uno studente a caso, sia fuori corso.

Esercizio 1



Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.

Si considerino 500 studenti di una Facoltà ripartiti secondo tre diversi corsi di studio (A, B, C) e secondo lo stato (in corso = IC, fuori corso = FC):

	A	B	C
IC	200	130	100
FC	35	25	10

1. Si calcoli la probabilità che, estraendo uno studente a caso, sia fuori corso.

$$P(FC) = \frac{35+25+10}{500} = \frac{70}{500} = 0.14.$$

Esercizio 1



Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.

Si considerino 500 studenti di una Facoltà ripartiti secondo tre diversi corsi di studio (A, B, C) e secondo lo stato (in corso = IC, fuori corso = FC):

	A	B	C
IC	200	130	100
FC	35	25	10

1. Si calcoli la probabilità che, estraendo uno studente a caso, sia fuori corso.

$$P(FC) = \frac{35+25+10}{500} = \frac{70}{500} = 0.14.$$

2. Si calcoli la probabilità che, estraendo uno studente a caso, sia iscritto al corso B.



Si considerino 500 studenti di una Facoltà ripartiti secondo tre diversi corsi di studio (A, B, C) e secondo lo stato (in corso = IC, fuori corso = FC):

	A	B	C
IC	200	130	100
FC	35	25	10

1. Si calcoli la probabilità che, estraendo uno studente a caso, sia fuori corso.

$$P(FC) = \frac{35+25+10}{500} = \frac{70}{500} = 0.14.$$

2. Si calcoli la probabilità che, estraendo uno studente a caso, sia iscritto al corso B.

$$P(B) = \frac{130+25}{500} = \frac{155}{500} = 0.31.$$



Si considerino 500 studenti di una Facoltà ripartiti secondo tre diversi corsi di studio (A, B, C) e secondo lo stato (in corso = IC, fuori corso = FC):

	A	B	C
IC	200	130	100
FC	35	25	10

1. Si calcoli la probabilità che, estraendo uno studente a caso, sia fuori corso.

$$P(FC) = \frac{35+25+10}{500} = \frac{70}{500} = 0.14.$$

2. Si calcoli la probabilità che, estraendo uno studente a caso, sia iscritto al corso B.

$$P(B) = \frac{130+25}{500} = \frac{155}{500} = 0.31.$$

3. Si determini la probabilità che, estraendo uno studente a caso, sia fuori corso oppure sia iscritto al corso B.



Si considerino 500 studenti di una Facoltà ripartiti secondo tre diversi corsi di studio (A, B, C) e secondo lo stato (in corso = IC, fuori corso = FC):

	A	B	C
IC	200	130	100
FC	35	25	10

1. Si calcoli la probabilità che, estraendo uno studente a caso, sia fuori corso.

$$P(FC) = \frac{35+25+10}{500} = \frac{70}{500} = 0.14.$$

2. Si calcoli la probabilità che, estraendo uno studente a caso, sia iscritto al corso B.

$$P(B) = \frac{130+25}{500} = \frac{155}{500} = 0.31.$$

3. Si determini la probabilità che, estraendo uno studente a caso, sia fuori corso oppure sia iscritto al corso B.

$$P(FC \cup B) =$$



Si considerino 500 studenti di una Facoltà ripartiti secondo tre diversi corsi di studio (A, B, C) e secondo lo stato (in corso = IC, fuori corso = FC):

	A	B	C
IC	200	130	100
FC	35	25	10

1. Si calcoli la probabilità che, estraendo uno studente a caso, sia fuori corso.

$$P(FC) = \frac{35+25+10}{500} = \frac{70}{500} = 0.14.$$

2. Si calcoli la probabilità che, estraendo uno studente a caso, sia iscritto al corso B.

$$P(B) = \frac{130+25}{500} = \frac{155}{500} = 0.31.$$

3. Si determini la probabilità che, estraendo uno studente a caso, sia fuori corso oppure sia iscritto al corso B.

$$P(FC \cup B) = P(FC) + P(B) - P(FC \cap B) =$$



Si considerino 500 studenti di una Facoltà ripartiti secondo tre diversi corsi di studio (A, B, C) e secondo lo stato (in corso = IC, fuori corso = FC):

	A	B	C
IC	200	130	100
FC	35	25	10

1. Si calcoli la probabilità che, estraendo uno studente a caso, sia fuori corso.

$$P(FC) = \frac{35+25+10}{500} = \frac{70}{500} = 0.14.$$

2. Si calcoli la probabilità che, estraendo uno studente a caso, sia iscritto al corso B.

$$P(B) = \frac{130+25}{500} = \frac{155}{500} = 0.31.$$

3. Si determini la probabilità che, estraendo uno studente a caso, sia fuori corso oppure sia iscritto al corso B.

$$P(FC \cup B) = P(FC) + P(B) - P(FC \cap B) = \frac{70}{500} + \frac{155}{500} - \frac{25}{500} = 0.14 + 0.31 - .05 = 0.4 \left(= \frac{130+35+25+10}{500} = \frac{200}{500} \right).$$



Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.

4. Si calcoli la probabilità che uno studente, estratto a caso, sia fuori corso sapendo che è iscritto al corso C .

$$P(FC|C) =$$



4. Si calcoli la probabilità che uno studente, estratto a caso, sia fuori corso sapendo che è iscritto al corso C .

$$P(FC|C) = \frac{P(FC \cap C)}{P(C)} = \frac{\frac{10}{500}}{\frac{110}{500}} = \frac{10}{110} = 0.0909.$$

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.



4. Si calcoli la probabilità che uno studente, estratto a caso, sia fuori corso sapendo che è iscritto al corso C .

$$P(FC|C) = \frac{P(FC \cap C)}{P(C)} = \frac{\frac{10}{500}}{\frac{110}{500}} = \frac{10}{110} = 0.0909.$$

5. Si calcoli la probabilità che, estratti senza reinserimento 12 studenti, la metà sia iscritta al corso A .

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.



4. Si calcoli la probabilità che uno studente, estratto a caso, sia fuori corso sapendo che è iscritto al corso C .

$$P(FC|C) = \frac{P(FC \cap C)}{P(C)} = \frac{\frac{10}{500}}{\frac{110}{500}} = \frac{10}{110} = 0.0909.$$

5. Si calcoli la probabilità che, estratti senza reinserimento 12 studenti, la metà sia iscritta al corso A .

$$P(6 A) =$$

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.



4. Si calcoli la probabilità che uno studente, estratto a caso, sia fuori corso sapendo che è iscritto al corso C .

$$P(FC|C) = \frac{P(FC \cap C)}{P(C)} = \frac{\frac{10}{500}}{\frac{110}{500}} = \frac{10}{110} = 0.0909.$$

5. Si calcoli la probabilità che, estratti senza reinserimento 12 studenti, la metà sia iscritta al corso A .

$$\begin{aligned} P(6 A) &= \binom{12}{6} P(A_1 \cap A_2 \cap \dots \cap A_6 \cap \bar{A}_7 \cap \bar{A}_8 \cap \dots \cap \bar{A}_{12}) = \\ &= \end{aligned}$$

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.



4. Si calcoli la probabilità che uno studente, estratto a caso, sia fuori corso sapendo che è iscritto al corso C.

$$P(FC|C) = \frac{P(FC \cap C)}{P(C)} = \frac{\frac{10}{500}}{\frac{110}{500}} = \frac{10}{110} = 0.0909.$$

5. Si calcoli la probabilità che, estratti senza reinserimento 12 studenti, la metà sia iscritta al corso A.

$$\begin{aligned} P(6 A) &= \binom{12}{6} P(A_1 \cap A_2 \cap \dots \cap A_6 \cap \bar{A}_7 \cap \bar{A}_8 \cap \dots \cap \bar{A}_{12}) = \\ &= \binom{12}{6} P(A_1) \cdot P(A_2|A_1) \cdot \dots \cdot P(A_6|A_1 \cap \dots \cap A_5) \cdot P(\bar{A}_7|A_1 \cap \dots \cap A_6) \cdot P(\bar{A}_8|A_1 \cap \dots \cap A_6 \cap \bar{A}_7) \cdot \dots \cdot P(\bar{A}_{12}|A_1 \cap \dots \cap A_6 \cap \bar{A}_7 \cap \dots \cap \bar{A}_{11}) = \\ &= \end{aligned}$$

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.



Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.

4. Si calcoli la probabilità che uno studente, estratto a caso, sia fuori corso sapendo che è iscritto al corso C .

$$P(FC|C) = \frac{P(FC \cap C)}{P(C)} = \frac{\frac{10}{500}}{\frac{110}{500}} = \frac{10}{110} = 0.0909.$$

5. Si calcoli la probabilità che, estratti senza reinserimento 12 studenti, la metà sia iscritta al corso A .

$$\begin{aligned} P(6 A) &= \binom{12}{6} P(A_1 \cap A_2 \cap \dots \cap A_6 \cap \bar{A}_7 \cap \bar{A}_8 \cap \dots \cap \bar{A}_{12}) = \\ &= \binom{12}{6} P(A_1) \cdot P(A_2|A_1) \cdot \dots \cdot P(A_6|A_1 \cap \dots \cap A_5) \cdot P(\bar{A}_7|A_1 \cap \dots \cap A_6) \cdot P(\bar{A}_8|A_1 \cap \dots \cap A_6 \cap \bar{A}_7) \cdot \dots \cdot P(\bar{A}_{12}|A_1 \cap \dots \cap A_6 \cap \bar{A}_7 \cap \dots \cap \bar{A}_{11}) = \\ &= \frac{12!}{6! \cdot 6!} \cdot \frac{235}{500} \cdot \frac{234}{499} \cdot \frac{233}{498} \cdot \frac{232}{497} \cdot \frac{231}{496} \cdot \frac{230}{495} \cdot \frac{265}{494} \cdot \frac{264}{493} \cdot \frac{263}{492} \cdot \frac{262}{491} \cdot \frac{261}{490} \cdot \frac{260}{489} = \\ &= 0.2234 \end{aligned}$$

oppure



Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.

4. Si calcoli la probabilità che uno studente, estratto a caso, sia fuori corso sapendo che è iscritto al corso C .

$$P(FC|C) = \frac{P(FC \cap C)}{P(C)} = \frac{\frac{10}{500}}{\frac{110}{500}} = \frac{10}{110} = 0.0909.$$

5. Si calcoli la probabilità che, estratti senza reinserimento 12 studenti, la metà sia iscritta al corso A .

$$\begin{aligned} P(6 A) &= \binom{12}{6} P(A_1 \cap A_2 \cap \dots \cap A_6 \cap \bar{A}_7 \cap \bar{A}_8 \cap \dots \cap \bar{A}_{12}) = \\ &= \binom{12}{6} P(A_1) \cdot P(A_2|A_1) \cdot \dots \cdot P(A_6|A_1 \cap \dots \cap A_5) \cdot P(\bar{A}_7|A_1 \cap \dots \cap A_6) \cdot P(\bar{A}_8|A_1 \cap \dots \cap A_6 \cap \bar{A}_7) \cdot \dots \cdot P(\bar{A}_{12}|A_1 \cap \dots \cap A_6 \cap \bar{A}_7 \cap \dots \cap \bar{A}_{11}) = \\ &= \frac{12!}{6! \cdot 6!} \cdot \frac{235}{500} \cdot \frac{234}{499} \cdot \frac{233}{498} \cdot \frac{232}{497} \cdot \frac{231}{496} \cdot \frac{230}{495} \cdot \frac{265}{494} \cdot \frac{264}{493} \cdot \frac{263}{492} \cdot \frac{262}{491} \cdot \frac{261}{490} \cdot \frac{260}{489} = \\ &= 0.2234 \end{aligned}$$

oppure indicando con X la v.c. ipergeometrica 'n. studenti iscritti al corso A nelle 12 estrazioni senza reimmissione'

$$P(X = 6) =$$



Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.

4. Si calcoli la probabilità che uno studente, estratto a caso, sia fuori corso sapendo che è iscritto al corso C .

$$P(FC|C) = \frac{P(FC \cap C)}{P(C)} = \frac{\frac{10}{500}}{\frac{110}{500}} = \frac{10}{110} = 0.0909.$$

5. Si calcoli la probabilità che, estratti senza reinserimento 12 studenti, la metà sia iscritta al corso A .

$$\begin{aligned} P(6 A) &= \binom{12}{6} P(A_1 \cap A_2 \cap \dots \cap A_6 \cap \bar{A}_7 \cap \bar{A}_8 \cap \dots \cap \bar{A}_{12}) = \\ &= \binom{12}{6} P(A_1) \cdot P(A_2|A_1) \cdot \dots \cdot P(A_6|A_1 \cap \dots \cap A_5) \cdot P(\bar{A}_7|A_1 \cap \dots \cap A_6) \cdot P(\bar{A}_8|A_1 \cap \dots \cap A_6 \cap \bar{A}_7) \cdot \dots \cdot P(\bar{A}_{12}|A_1 \cap \dots \cap A_6 \cap \bar{A}_7 \cap \dots \cap \bar{A}_{11}) = \\ &= \frac{12!}{6! \cdot 6!} \cdot \frac{235}{500} \cdot \frac{234}{499} \cdot \frac{233}{498} \cdot \frac{232}{497} \cdot \frac{231}{496} \cdot \frac{230}{495} \cdot \frac{265}{494} \cdot \frac{264}{493} \cdot \frac{263}{492} \cdot \frac{262}{491} \cdot \frac{261}{490} \cdot \frac{260}{489} = \\ &= 0.2234 \end{aligned}$$

oppure indicando con X la v.c. ipergeometrica 'n. studenti iscritti al corso A nelle 12 estrazioni senza reimmissione'

$$P(X = 6) = \frac{\binom{235}{6} \binom{265}{6}}{\binom{500}{12}} = 0.2234.$$

Tema T279-2

Prova scritta di STATISTICA del 10.07.2014

Prova scritta di
STATISTICA del
10.07.2014

Giuseppe Boari
Gabriele Cantaluppi



UNIVERSITÀ
CATTOLICA
del Sacro Cuore

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.

Esercizio 5.

Giuseppe Boari
Gabriele Cantaluppi
Università Cattolica del Sacro Cuore, Milano

Esercizio 1



La seguente tabella riporta la distribuzione congiunta della variabile statistica doppia (X, Y) per un campione di 150 unità statistiche.

$X \setminus Y$	10 - 15	15 - 25	25 - 30	$n_{i\bullet}$
x_1	22	60	18	100
x_2	11	30	9	50
$n_{\bullet j}$	33	90	27	150

1. Si confrontino mediante opportuni indici di posizione e di variabilità le distribuzioni condizionate $Y|x_1$ e $Y|x_2$.

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.

Esercizio 5.



La seguente tabella riporta la distribuzione congiunta della variabile statistica doppia (X, Y) per un campione di 150 unità statistiche.

$X \setminus Y$	10 - 15	15 - 25	25 - 30	$n_{i\bullet}$
x_1	22	60	18	100
x_2	11	30	9	50
$n_{\bullet j}$	33	90	27	150

1. Si confrontino mediante opportuni indici di posizione e di variabilità le distribuzioni condizionate $Y|x_1$ e $Y|x_2$.

Osservando la proporzionalità delle frequenze condizionate si evince che sussiste indipendenza stocastica; quindi le distribuzioni condizionate sono caratterizzate dalla stessa media e dallo stesso livello di variabilità.

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.

Esercizio 5.



Si considera la distribuzione condizionata $Y|x_1$

y_j	n_{1j}
10 $\rightarrow$ 15	22
15 $\rightarrow$ 25	60
25 $\rightarrow$ 30	18

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.

Esercizio 5.



Si tratta di un carattere quantitativo con modalità raggruppate in classi.

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.

Esercizio 5.



La media aritmetica

$$M(Y|x_1) = \frac{1}{n_{1\bullet}} \sum_{j=1}^k y_j n_{1j}$$

risulta 19.7.

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.

Esercizio 5.



La media aritmetica

$$M(Y|x_1) = \frac{1}{n_{1\bullet}} \sum_{j=1}^k y_j n_{1j}$$

risulta 19.7.

Si ottiene (considerando come modalità i valori centrali delle classi)

y_j	n_{1j}	$y_j n_{1j}$
12.5	22	275
20.0	60	1200
27.5	18	495
	100	1970

da cui segue

$$M(Y|x_1) = \frac{1}{n_{1\bullet}} \sum_{j=1}^k y_j n_{1j} = \frac{1970}{100} = 19.7.$$

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.

Esercizio 5.



La varianza

$$\text{Var}(Y|x_1) = M(Y^2|x_1) - [M(Y|x_1)]^2$$

risulta 22.41.

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.

Esercizio 5.



Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.

Esercizio 5.

La varianza

$$\text{Var}(Y|x_1) = M(Y^2|x_1) - [M(Y|x_1)]^2$$

risulta 22.41. Si ottiene (considerando come modalità i valori centrali delle classi)

y_j	n_{1j}	$y_j n_{1j}$	$y_j^2 n_{1j}$
12.5	22	275	3437.5
20.0	60	1200	24000.0
27.5	18	495	13612.5
	100	1970	41050.0

da cui segue

$$\text{Var}(Y|x_1) = \frac{41050}{100} - 19.7^2 = 410.5 - 388.09 = 22.41.$$



Per effettuare confronti in termini di variabilità tra caratteri quantitativi occorre utilizzare una misura relativa, ad esempio il coefficiente di variazione

$$CV = \frac{\sigma}{\mu} = \frac{\sqrt{\text{Var}(Y|x_1)}}{M(Y|x_1)} = \frac{\sqrt{22.41}}{19.7} = \frac{4.7339}{19.7} = 0.2403.$$

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.

Esercizio 5.



2. Si rappresentino graficamente le distribuzioni marginali X e Y .

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.

Esercizio 5.



Si considera la distribuzione marginale della variabile X , ottenuta sommando le frequenze di riga

x_i	$n_{i\bullet}$
x_1	100
x_2	50

Si tratta di un carattere qualitativo sconnesso che può essere rappresentato graficamente mediante un grafico a bastoncini o a canne d'organo.

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.

Esercizio 5.



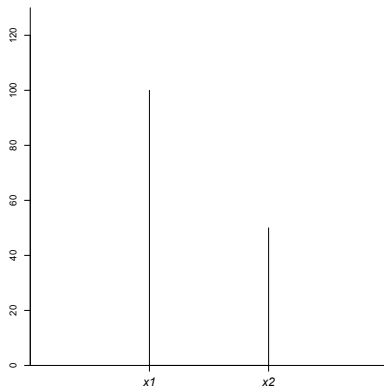
Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.

Esercizio 5.





Si considera la distribuzione marginale della variabile Y , ottenuta sommando le frequenze di colonna

y_j	$n_{\bullet j}$
10 $\rightarrow$ 15	33
15 $\rightarrow$ 25	90
25 $\rightarrow$ 30	27

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.

Esercizio 5.



Si tratta di un carattere quantitativo con modalità raggruppate in classi che può essere rappresentato graficamente mediante un istogramma.

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.

Esercizio 5.



L'istogramma può essere costruito dopo avere calcolato le densità $d_{\bullet j} = \frac{n_{\bullet j}}{a_j}$, (dove a_j sono le ampiezze di classe).

y_j	$n_{\bullet j}$	$a_{\bullet j}$	$d_{\bullet j}$
10 - 15	33	$15 - 10 = 5$	6.6
15 - 25	90	$25 - 15 = 10$	9
25 - 30	27	$30 - 25 = 5$	5.4

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.

Esercizio 5.



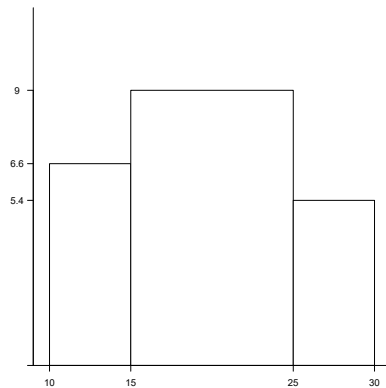
Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.

Esercizio 5.





Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.

Esercizio 5.

3. Si calcoli un opportuno indice di connessione tra X e Y .
Si osserva che sussiste indipendenza stocastica, quindi $\chi_N = 0$.

Esercizio 2



Si consideri la seguente serie storica:

<i>anno</i>	x_t
2009	200
2010	180
2011	186
2012	204
2013	222

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.

Esercizio 5.



1. Si costruiscano le serie dei numeri indici a base fissa (2009) e a base mobile.

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.

Esercizio 5.

1. Si costruiscano le serie dei numeri indici a base fissa (2009) e a base mobile.
Si ricorda che ${}_{2009}I_t = \frac{x_t}{x_{2009}}$.



Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.

Esercizio 5.



1. Si costruiscano le serie dei numeri indici a base fissa (2009) e a base mobile.
Si ricorda che ${}_{2009}I_t = \frac{x_t}{x_{2009}}$. Si ottiene quindi

Anno	x_t	NIBF (2009)
2009	200	1.00
2010	180	0.90
2011	186	0.93
2012	204	1.02
2013	222	1.11

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.

Esercizio 5.



Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.

Esercizio 5.

1. Si costruiscano le serie dei numeri indici a base fissa (2009) e a base mobile.
Si ricorda che ${}_{2009}I_t = \frac{x_t}{x_{2009}}$. Si ottiene quindi

Anno	x_t	NIBF (2009)
2009	200	1.00
2010	180	0.90
2011	186	0.93
2012	204	1.02
2013	222	1.11

Si ricorda che ${}_{t-1}I_t = \frac{x_t}{x_{t-1}}$.



Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.

Esercizio 5.

1. Si costruiscano le serie dei numeri indici a base fissa (2009) e a base mobile.
Si ricorda che ${}_{2009}I_t = \frac{x_t}{x_{2009}}$. Si ottiene quindi

Anno	x_t	NIBF (2009)
2009	200	1.00
2010	180	0.90
2011	186	0.93
2012	204	1.02
2013	222	1.11

Si ricorda che ${}_{t-1}I_t = \frac{x_t}{x_{t-1}}$. Si ottiene quindi

Anno	x_t	NIBM
2009	200	
2010	180	0.9000
2011	186	1.0333
2012	204	1.0968
2013	222	1.0882



2. Si calcoli ${}_{2010}I_{2013}$ commentando il risultato ottenuto.

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.

Esercizio 5.



2. Si calcoli ${}_{2010}I_{2013}$ commentando il risultato ottenuto.

$${}_{2010}I_{2013} = \frac{x_{2013}}{x_{2010}} = \frac{222}{180} = 1.2333$$

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.

Esercizio 5.



Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.

Esercizio 5.

2. Si calcoli ${}_{2010}I_{2013}$ commentando il risultato ottenuto.

$${}_{2010}I_{2013} = \frac{x_{2013}}{x_{2010}} = \frac{222}{180} = 1.2333$$

Tra il tempo 2010 e il tempo 2013 si è registrato un incremento complessivo del 23.33% della grandezza oggetto di studio.



3. Si calcoli il tasso medio di variazione dal 2010 al 2013 e si commenti il risultato ottenuto.

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.

Esercizio 5.



3. Si calcoli il tasso medio di variazione dal 2010 al 2013 e si commenti il risultato ottenuto.
Si ricorda che

$${}_{2010}I_{2013} = {}_{2010}I_{2011} \cdot {}_{2011}I_{2012} \cdot {}_{2012}I_{2013}$$

Si deve trovare il numero indice a base mobile medio.

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.

Esercizio 5.



Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.

Esercizio 5.

3. Si calcoli il tasso medio di variazione dal 2010 al 2013 e si commenti il risultato ottenuto. Si ricorda che

$${}_{2010}I_{2013} = {}_{2010}I_{2011} \cdot {}_{2011}I_{2012} \cdot {}_{2012}I_{2013}$$

Si deve trovare il numero indice a base mobile medio. Applicando il criterio di scelta della media secondo Chisini si ottiene

$${}_{2010}I_{2013} = {}_{2010}I_{2011} \cdot {}_{2011}I_{2012} \cdot {}_{2012}I_{2013} = \alpha \cdot \alpha \cdot \alpha = \alpha^3.$$



Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.

Esercizio 5.

3. Si calcoli il tasso medio di variazione dal 2010 al 2013 e si commenti il risultato ottenuto. Si ricorda che

$${}_{2010}l_{2013} = {}_{2010}l_{2011} \cdot {}_{2011}l_{2012} \cdot {}_{2012}l_{2013}$$

Si deve trovare il numero indice a base mobile medio. Applicando il criterio di scelta della media secondo Chisini si ottiene

$${}_{2010}l_{2013} = {}_{2010}l_{2011} \cdot {}_{2011}l_{2012} \cdot {}_{2012}l_{2013} = \alpha \cdot \alpha \cdot \alpha = \alpha^3.$$

Quindi

$$\alpha = \sqrt[3]{{}_{2010}l_{2013}} = \sqrt[3]{{}_{2010}l_{2011} \cdot {}_{2011}l_{2012} \cdot {}_{2012}l_{2013}} = 1.0724.$$

Da cui segue una variazione relativa media pari a 0.0724, corrispondente a un incremento medio annuo tra il 2010 e il 2013 del 7.24%.



[Esercizio 1.](#)

[Esercizio 2.](#)

[Esercizio 3.](#)

[Esercizio 4.](#)

[Esercizio 5.](#)

Il seguente prospetto riporta i valori rilevati in corrispondenza a 5 unità statistiche con riferimento a due caratteri quantitativi X e Y

x_i	5	5	3	1	1
y_i	11	25	25	25	39

1. Si rappresentino graficamente le coppie dei dati e la funzione di regressione ipotizzando Y in funzione di X .



La funzione di regressione associa a ogni valore x appartenente all'insieme di definizione della variabile X la media di Y condizionata a x , $M(Y|x)$

x	$M(Y x)$
1	$\frac{1}{2}(25 + 39) = 32$
3	25
5	$\frac{1}{2}(11 + 25) = 18$

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.

Esercizio 5.



Esercizio 1.

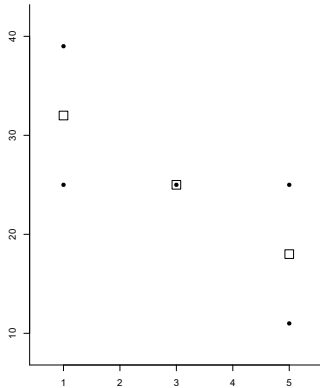
Esercizio 2.

Esercizio 3.

Esercizio 4.

Esercizio 5.

Le coppie di dati possono essere rappresentate mediante un grafico a dispersione.





2. Si calcolino, secondo il metodo dei minimi quadrati, i parametri e l'indice di adattamento dei seguenti modelli di regressione, individuando il modello migliore:

$$I) Y^* = a + bX \qquad II) Y^* = c$$

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.

Esercizio 5.



Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.

Esercizio 5.

Si riportano i calcoli necessari¹ per ottenere le stime dei parametri e gli indici di adattamento del modello

$$Y = a + bX + E$$

x_i	y_i	$x_i y_i$	x_i^2	y_i^2
5	11	55	25	121
5	25	125	25	625
3	25	75	9	625
1	25	25	1	625
1	39	39	1	1521
15	125	319	61	3517
3	25	63.8	12.2	703.4
$M(X)$	$M(Y)$	$M(XY)$	$M(X^2)$	$M(Y^2)$

$$\hat{b} =$$

¹Si osserva che per ragioni di stampa alcuni calcoli potrebbero essere duplicati con riferimento ai diversi modelli oggetto di studio



Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.

Esercizio 5.

Si riportano i calcoli necessari¹ per ottenere le stime dei parametri e gli indici di adattamento del modello

$$Y = a + bX + E$$

x_i	y_i	$x_i y_i$	x_i^2	y_i^2
5	11	55	25	121
5	25	125	25	625
3	25	75	9	625
1	25	25	1	625
1	39	39	1	1521
15	125	319	61	3517
3	25	63.8	12.2	703.4
$M(X)$	$M(Y)$	$M(XY)$	$M(X^2)$	$M(Y^2)$

$$\begin{aligned}\hat{b} &= \frac{\text{Cov}(X, Y)}{\text{Var}(X)} = \frac{M(XY) - M(X)M(Y)}{M(X^2) - [M(X)]^2} = \\ &= \frac{63.8 - 3 \cdot 25}{12.2 - 3^2} = \frac{-11.2}{3.2} = -3.5\end{aligned}$$

$$\hat{a} =$$

¹Si osserva che per ragioni di stampa alcuni calcoli potrebbero essere duplicati con riferimento ai diversi modelli oggetto di studio



Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.

Esercizio 5.

Si riportano i calcoli necessari¹ per ottenere le stime dei parametri e gli indici di adattamento del modello

$$Y = a + bX + E$$

x_i	y_i	$x_i y_i$	x_i^2	y_i^2
5	11	55	25	121
5	25	125	25	625
3	25	75	9	625
1	25	25	1	625
1	39	39	1	1521
15	125	319	61	3517
3	25	63.8	12.2	703.4
$M(X)$	$M(Y)$	$M(XY)$	$M(X^2)$	$M(Y^2)$

$$\begin{aligned}\hat{b} &= \frac{\text{Cov}(X, Y)}{\text{Var}(X)} = \frac{M(XY) - M(X)M(Y)}{M(X^2) - [M(X)]^2} = \\ &= \frac{63.8 - 3 \cdot 25}{12.2 - 3^2} = \frac{-11.2}{3.2} = -3.5 \\ \hat{a} &= M(Y) - \hat{b}M(X) = 35.5 \\ \rho^2 &= \end{aligned}$$

¹Si osserva che per ragioni di stampa alcuni calcoli potrebbero essere duplicati con riferimento ai diversi modelli oggetto di studio



Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.

Esercizio 5.

Si riportano i calcoli necessari¹ per ottenere le stime dei parametri e gli indici di adattamento del modello

$$Y = a + bX + E$$

x_i	y_i	$x_i y_i$	x_i^2	y_i^2
5	11	55	25	121
5	25	125	25	625
3	25	75	9	625
1	25	25	1	625
1	39	39	1	1521
15	125	319	61	3517
3	25	63.8	12.2	703.4
$M(X)$	$M(Y)$	$M(XY)$	$M(X^2)$	$M(Y^2)$

$$\begin{aligned}\hat{b} &= \frac{\text{Cov}(X, Y)}{\text{Var}(X)} = \frac{M(XY) - M(X)M(Y)}{M(X^2) - [M(X)]^2} = \\ &= \frac{63.8 - 3 \cdot 25}{12.2 - 3^2} = \frac{-11.2}{3.2} = -3.5 \\ \hat{a} &= M(Y) - \hat{b}M(X) = 35.5 \\ \rho^2 &= \frac{[\text{Cov}(X, Y)]^2}{\text{Var}(X)\text{Var}(Y)} = \frac{(-11.2)^2}{3.2 \cdot 78.4} = 0.5\end{aligned}$$

¹Si osserva che per ragioni di stampa alcuni calcoli potrebbero essere duplicati con riferimento ai diversi modelli oggetto di studio



Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.

Esercizio 5.

Si riportano i calcoli necessari² per ottenere le stime dei parametri e gli indici di adattamento del modello

$$Y = c + E$$

x_i	y_i	y_i^2
5	11	121
5	25	625
3	25	625
1	25	625
1	39	1521
125		3517
25		703.4
$M(Y)$		$M(Y^2)$

$$\hat{c} =$$

²Si osserva che per ragioni di stampa alcuni calcoli potrebbero essere duplicati con riferimento ai diversi modelli oggetto di studio



Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.

Esercizio 5.

Si riportano i calcoli necessari² per ottenere le stime dei parametri e gli indici di adattamento del modello

$$Y = c + E$$

x_i	y_i	y_i^2
5	11	121
5	25	625
3	25	625
1	25	625
1	39	1521
125		3517
25		703.4
$M(Y)$		$M(Y^2)$

$$\hat{c} = M(Y) = 25$$

$$R^2 =$$

²Si osserva che per ragioni di stampa alcuni calcoli potrebbero essere duplicati con riferimento ai diversi modelli oggetto di studio



Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.

Esercizio 5.

Si riportano i calcoli necessari² per ottenere le stime dei parametri e gli indici di adattamento del modello

$$Y = c + E$$

x_i	y_i	y_i^2
5	11	121
5	25	625
3	25	625
1	25	625
1	39	1521
125		3517
25		703.4
$M(Y)$		$M(Y^2)$

$$\hat{c} = M(Y) = 25$$

$$R^2 = 0$$

²Si osserva che per ragioni di stampa alcuni calcoli potrebbero essere duplicati con riferimento ai diversi modelli oggetto di studio



Il modello migliore è $Y^* = 35.5 - 3.5X$.

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.

Esercizio 5.



Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.

Esercizio 5.

3. Per il modello migliore, si fornisca la previsione in corrispondenza di $X = 3$.



3. Per il modello migliore, si fornisca la previsione in corrispondenza di $X = 3$.

$$Y^*(3) = 35.5 - 3.5 \cdot 3 = 25$$

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.

Esercizio 5.

Esercizio 4

Prova scritta di
STATISTICA del
10.07.2014

Giuseppe Boari
Gabriele Cantaluppi



UNIVERSITÀ
CATTOLICA
del Sacro Cuore

Si completino le seguenti tabelle, giustificando la risposta, in modo che siano soddisfatte le condizioni indicate.

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.

Esercizio 5.



	$y_1 = 10$	$y_2 = 20$	$y_3 = 30$
$x_1 = 1$	0	11	
$x_2 = 2$		0	

a) $\eta_{Y|X}^2 = \rho^2$ con $\rho > 0$

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.

Esercizio 5.



Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.

Esercizio 5.

	$y_1 = 10$	$y_2 = 20$	$y_3 = 30$
$x_1 = 1$	0	11	
$x_2 = 2$		0	

a) $\eta_{Y|X}^2 = \rho^2$ con $\rho > 0$

$\eta_{Y|X}^2$ risulta sempre uguale a ρ^2 in quanto abbiamo due sole medie condizionate che saranno sempre allineate.

Affinché $\rho > 0$ la retta deve avere pendenza crescente, vale a dire $M(Y|x_2) > M(Y|x_1)$.

Una possibile soluzione è $n_{13} = 0, 0 < n_{21} < n_{23}$.



	$y_1 = 10$	$y_2 = 20$	$y_3 = 30$
$x_1 = 1$		11	
$x_2 = 2$		11	

b) $\rho^2 = 0$

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.

Esercizio 5.



Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.

Esercizio 5.

	$y_1 = 10$	$y_2 = 20$	$y_3 = 30$
$x_1 = 1$		11	
$x_2 = 2$		11	

b) $\rho^2 = 0$

Se $\eta_{Y|X}^2 = 0$ anche $\rho^2 = 0$ in quanto l'indipendenza in media implica l'indipendenza lineare.
 $n_{11} = n_{13}$, $n_{21} = n_{23}$ rappresenta una possibile soluzione.

Esercizio 5

Prova scritta di
STATISTICA del
10.07.2014

Giuseppe Boari
Gabriele Cantaluppi



UNIVERSITÀ
CATTOLICA
del Sacro Cuore

Un'urna contiene 10 palline nere e 20 palline rosse.

Si consideri il seguente esperimento: si toglie, in primo luogo, una pallina dall'urna; si estraggono, poi, 7 palline.

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.

Esercizio 5.



1. Si calcoli la probabilità che tra le 7 palline estratte ve ne siano 6 rosse nell'ipotesi che le 7 estrazioni avvengano con reimmissione.

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.

Esercizio 5.



1. Si calcoli la probabilità che tra le 7 palline estratte ve ne siano 6 rosse nell'ipotesi che le 7 estrazioni avvengano con reimmissione.

Indicati con R l'evento 'la pallina tolta è rossa' e con A l'evento '6 palline rosse tra le 7 estratte con reimmissione' abbiamo:

$$P(A) =$$

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.

Esercizio 5.



1. Si calcoli la probabilità che tra le 7 palline estratte ve ne siano 6 rosse nell'ipotesi che le 7 estrazioni avvengano con reimmissione.

Indicati con R l'evento 'la pallina tolta è rossa' e con A l'evento '6 palline rosse tra le 7 estratte con reimmissione' abbiamo:

$$\begin{aligned} P(A) &= P(R) \cdot P(A|R) + P(\bar{R}) \cdot P(A|\bar{R}) = \\ &= \end{aligned}$$

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.

Esercizio 5.



1. Si calcoli la probabilità che tra le 7 palline estratte ve ne siano 6 rosse nell'ipotesi che le 7 estrazioni avvengano con reimmissione.

Indicati con R l'evento 'la pallina tolta è rossa' e con A l'evento '6 palline rosse tra le 7 estratte con reimmissione' abbiamo:

$$\begin{aligned} P(A) &= P(R) \cdot P(A|R) + P(\bar{R}) \cdot P(A|\bar{R}) = \\ &= \frac{20}{30} \cdot \binom{7}{6} \left(\frac{19}{29}\right)^6 \left(\frac{10}{29}\right)^1 + \frac{10}{30} \cdot \binom{7}{6} \left(\frac{20}{29}\right)^6 \left(\frac{9}{29}\right)^1 = \\ &= \end{aligned}$$

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.

Esercizio 5.



Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.

Esercizio 5.

1. Si calcoli la probabilità che tra le 7 palline estratte ve ne siano 6 rosse nell'ipotesi che le 7 estrazioni avvengano con reimmissione.

Indicati con R l'evento 'la pallina tolta è rossa' e con A l'evento '6 palline rosse tra le 7 estratte con reimmissione' abbiamo:

$$\begin{aligned} P(A) &= P(R) \cdot P(A|R) + P(\bar{R}) \cdot P(A|\bar{R}) = \\ &= \frac{20}{30} \cdot \binom{7}{6} \left(\frac{19}{29}\right)^6 \left(\frac{10}{29}\right)^1 + \frac{10}{30} \cdot \binom{7}{6} \left(\frac{20}{29}\right)^6 \left(\frac{9}{29}\right)^1 = \\ &= \frac{2}{3} \cdot 0.1909 + \frac{1}{3} \cdot 0.2337 = 0.2052. \end{aligned}$$



2. Si calcoli la probabilità che tra le 7 palline estratte ve ne siano 6 rosse nell'ipotesi che le 7 estrazioni avvengano senza reimmissione.

Indicati con R l'evento 'la pallina tolta è rossa' e con A l'evento '6 palline rosse tra le 7 estratte senza reimmissione' abbiamo:

$$P(A) =$$

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.

Esercizio 5.



2. Si calcoli la probabilità che tra le 7 palline estratte ve ne siano 6 rosse nell'ipotesi che le 7 estrazioni avvengano senza reimmissione.

Indicati con R l'evento 'la pallina tolta è rossa' e con A l'evento '6 palline rosse tra le 7 estratte senza reimmissione' abbiamo:

$$\begin{aligned} P(A) &= P(R) \cdot P(A|R) + P(\bar{R}) \cdot P(A|\bar{R}) = \\ &= \end{aligned}$$

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.

Esercizio 5.



Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.

Esercizio 5.

2. Si calcoli la probabilità che tra le 7 palline estratte ve ne siano 6 rosse nell'ipotesi che le 7 estrazioni avvengano senza reimmissione.

Indicati con R l'evento 'la pallina tolta è rossa' e con A l'evento '6 palline rosse tra le 7 estratte senza reimmissione' abbiamo:

$$\begin{aligned} P(A) &= P(R) \cdot P(A|R) + P(\bar{R}) \cdot P(A|\bar{R}) = \\ &= \frac{20}{30} \cdot \frac{\binom{19}{6} \binom{10}{1}}{\binom{29}{7}} + \frac{10}{30} \cdot \frac{\binom{20}{6} \binom{9}{1}}{\binom{29}{7}} = \\ &= \end{aligned}$$



Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.

Esercizio 5.

2. Si calcoli la probabilità che tra le 7 palline estratte ve ne siano 6 rosse nell'ipotesi che le 7 estrazioni avvengano senza reimmissione.

Indicati con R l'evento 'la pallina tolta è rossa' e con A l'evento '6 palline rosse tra le 7 estratte senza reimmissione' abbiamo:

$$\begin{aligned} P(A) &= P(R) \cdot P(A|R) + P(\bar{R}) \cdot P(A|\bar{R}) = \\ &= \frac{20}{30} \cdot \frac{\binom{19}{6} \binom{10}{1}}{\binom{29}{7}} + \frac{10}{30} \cdot \frac{\binom{20}{6} \binom{9}{1}}{\binom{29}{7}} = \\ &= \frac{2}{3} \cdot 0.1738 + \frac{1}{3} \cdot 0.2235 = 0.1904. \end{aligned}$$

Tema T280-1

Prova scritta di STATISTICA del 04.09.2014

Prova scritta di
STATISTICA del
04.09.2014

Giuseppe Boari
Gabriele Cantaluppi



UNIVERSITÀ
CATTOLICA
del Sacro Cuore

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.

Esercizio 5.

Giuseppe Boari
Gabriele Cantaluppi
Università Cattolica del Sacro Cuore, Milano



La seguente tabella riporta la distribuzione congiunta della variabile statistica doppia (X, Y) , con X carattere qualitativo, per un campione di 150 unità statistiche.

$X \setminus Y$	10 – 15	15 – 25	25 – 30	$n_{i\bullet}$
x_1	22	60	18	100
x_2	11	30	9	50
$n_{\bullet j}$	33	90	27	150

1. Si riassumano mediante opportuni indici di posizione e di variabilità le distribuzioni marginali X e Y .

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.

Esercizio 5.



Si considera la distribuzione marginale della variabile X , ottenuta sommando le frequenze di riga

x_i	$n_{i\bullet}$
x_1	100
x_2	50

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.

Esercizio 5.



Si tratta di un carattere qualitativo ordinato.

La moda corrisponde alla categoria alla quale è associata la massima frequenza) che risulta x_1 .

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.

Esercizio 5.



La mediana (percentile di ordine 0.50) in presenza di un numero pari di osservazioni corrisponde a uno degli elementi che occupano le posizioni $\frac{n}{2} = 75$ ed $\frac{n}{2} + 1 = 76$ nella serie dei dati ordinati. Dalla distribuzione delle frequenze cumulate

x_i	$n_{i\bullet}$	$N_{i\bullet}$
x_1	100	100
x_2	50	150

si trova che i due elementi coincidono e risulta $x_{0.5} = x_1$.

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.

Esercizio 5.



In presenza di un carattere qualitativo la variabilità può essere misurata mediante gli indici di Gini e di Frosini normalizzati

$$G_N = \frac{1 - \sum_{i=1}^h f_{i\bullet}^2}{1 - \frac{1}{h}}$$

e

$$F_N = 1 - \sqrt{1 - G_N}.$$

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.

Esercizio 5.



In presenza di un carattere qualitativo la variabilità può essere misurata mediante gli indici di Gini e di Frosini normalizzati

$$G_N = \frac{1 - \sum_{i=1}^h f_{i\bullet}^2}{1 - \frac{1}{h}}$$

e

$$F_N = 1 - \sqrt{1 - G_N}.$$

Si ottiene:

x_i	$n_{i\bullet}$	$f_{i\bullet}$	$f_{i\bullet}^2$
x_1	100	0.6667	0.4444
x_2	50	0.3333	0.1111
	150	1	0.5555

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.

Esercizio 5.



In presenza di un carattere qualitativo la variabilità può essere misurata mediante gli indici di Gini e di Frosini normalizzati

$$G_N = \frac{1 - \sum_{i=1}^h f_{i\bullet}^2}{1 - \frac{1}{h}}$$

e

$$F_N = 1 - \sqrt{1 - G_N}.$$

Si ottiene:

x_i	$n_{i\bullet}$	$f_{i\bullet}$	$f_{i\bullet}^2$
x_1	100	0.6667	0.4444
x_2	50	0.3333	0.1111
	150	1	0.5555

Da cui segue:

$$G_N = \frac{1 - 0.5556}{1 - \frac{1}{2}} = 0.8889$$

e

$$F_N = 1 - \sqrt{1 - 0.8889} = 0.6667.$$

Sussiste, quindi, un livello medio alto di eterogeneità.



Si considera la distribuzione marginale della variabile Y , ottenuta sommando le frequenze di colonna

y_j	$n_{\bullet j}$
10 $\rightarrow$ 15	33
15 $\rightarrow$ 25	90
25 $\rightarrow$ 30	27

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.

Esercizio 5.



Si tratta di un carattere quantitativo con modalità raggruppate in classi.

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.

Esercizio 5.



La media aritmetica

$$M(Y) = \frac{1}{n} \sum_{j=1}^k y_j n_{\bullet j}$$

risulta 19.7.

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.

Esercizio 5.



La media aritmetica

$$M(Y) = \frac{1}{n} \sum_{j=1}^k y_j n_{\bullet j}$$

risulta 19.7.

Si ottiene (considerando come modalità i valori centrali delle classi)

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.

Esercizio 5.



La media aritmetica

$$M(Y) = \frac{1}{n} \sum_{j=1}^k y_j n_{\bullet j}$$

risulta 19.7.

Si ottiene (considerando come modalità i valori centrali delle classi)

y_j	$n_{\bullet j}$	$y_j n_{\bullet j}$
12.5	33	412.5
20.0	90	1800.0
27.5	27	742.5
	150	2955.0

da cui segue

$$M(Y) = \frac{1}{n} \sum_{i=1}^k y_i n_{\bullet i} = \frac{2955}{150} = 19.7.$$

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.

Esercizio 5.



La varianza

$$\text{Var}(Y) = M[(Y - \mu_Y)^2] = M(Y^2) - \mu_Y^2 = \frac{1}{n} \sum_{j=1}^k y_j^2 n_{\bullet j} - \mu_Y^2$$

risulta 22.41.

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.

Esercizio 5.



La varianza

$$Var(Y) = M[(Y - \mu_Y)^2] = M(Y^2) - \mu_Y^2 = \frac{1}{n} \sum_{j=1}^k y_j^2 n_{\bullet j} - \mu_Y^2$$

risulta 22.41.

Si ottiene (considerando come modalità i valori centrali delle classi)

y_j	$n_{\bullet j}$	$y_j n_{\bullet j}$	$y_j^2 n_{\bullet j}$
12.5	33	412.5	5156.25
20.0	90	1800.0	36000.00
27.5	27	742.5	20418.75
	150	2955.0	61575.00

da cui segue

$$\begin{aligned} Var(Y) = \sigma_Y^2 &= M(Y^2) - \mu_Y^2 = \\ &= \frac{61575}{150} - 19.7^2 = 410.5 - 388.09 = 22.41. \end{aligned}$$

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.

Esercizio 5.



Per effettuare confronti in termini di variabilità tra caratteri quantitativi occorre utilizzare una misura relativa, ad esempio il coefficiente di variazione

$$CV = \frac{\sigma}{\mu} = \frac{\sqrt{\text{Var}(Y)}}{M(Y)} = \frac{\sqrt{22.41}}{19.7} = \frac{4.7339}{19.7} = 0.2403.$$

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.

Esercizio 5.



Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.

Esercizio 5.

2. Si rappresenti graficamente la distribuzione condizionata $Y|x_2$.



Si considera la distribuzione condizionata $Y|x_2$

y_j	n_{2j}
10 + 15	11
15 + 25	30
25 + 30	9

Esercizio 1.

Esercizio 2.

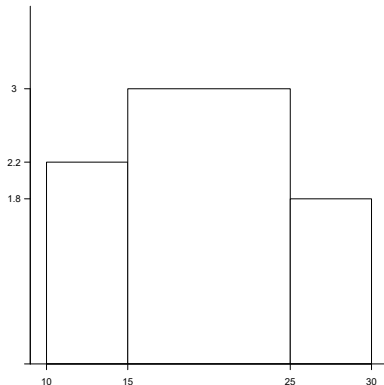
Esercizio 3.

Esercizio 4.

Esercizio 5.

Si tratta di un carattere quantitativo con modalità raggruppate in classi che può essere rappresentato graficamente mediante un istogramma. L'istogramma può essere costruito dopo avere calcolato le densità $d_{2j} = \frac{n_{2j}}{a_j}$, (dove a_j sono le ampiezze di classe).

y_j	n_{2j}	a_{2j}	d_{2j}
10 - 15	11	$15 - 10 = 5$	2.2
15 - 25	30	$25 - 15 = 10$	3
25 - 30	9	$30 - 25 = 5$	1.8



Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.

Esercizio 5.



3. Si costruisca il grafico Box & Whiskers plot per la marginale Y .

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.

Esercizio 5.



Si considera la distribuzione marginale della variabile Y , ottenuta sommando le frequenze di colonna

y_j	$n_{\bullet j}$
10 - 15	33
15 - 25	90
25 - 30	27

Si tratta di un carattere quantitativo con modalità raggruppate in classi.

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.

Esercizio 5.



Per costruire il grafico Box & Whiskers plot occorre determinare in primo luogo i valori dei tre quartili $y_{0.25}$, $y_{0.5}$ e $y_{0.75}$ considerando l'inversa della Funzione di Ripartizione in corrispondenza rispettivamente di 0.25, 0.50 e 0.75.

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.

Esercizio 5.



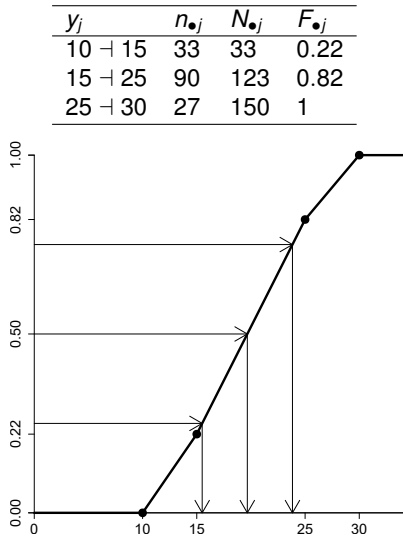
Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.

Esercizio 5.





Si ottiene

$$y_{0.25} = 15 + \frac{0.25 - 0.22}{0.82 - 0.22} \cdot (25 - 15) = 15.5$$

$$y_{0.5} = 15 + \frac{0.5 - 0.22}{0.82 - 0.22} \cdot (25 - 15) = 19.6667$$

$$y_{0.75} = 15 + \frac{0.75 - 0.22}{0.82 - 0.22} \cdot (25 - 15) = 23.8333$$

La lunghezza teorica dei baffi risulta

$$1.5 \cdot (y_{0.75} - y_{0.25}) = 12.5.$$

Il baffo inferiore è posizionato in

$$\max(y_{\min}, y_{0.25} - 12.5) = \max(10, 3) = 10.$$

Il baffo superiore è posizionato in

$$\min(y_{\max}, y_{0.75} + 12.5) = \min(30, 36.3333) = 30.$$



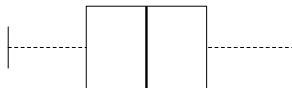
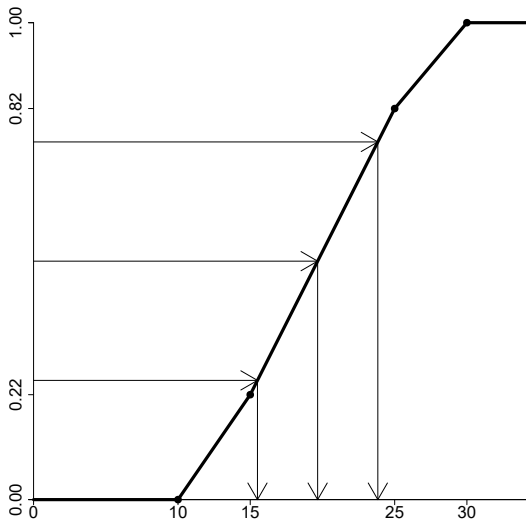
Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.

Esercizio 5.



Esercizio 2

Prova scritta di
STATISTICA del
04.09.2014

Giuseppe Boari
Gabriele Cantaluppi



UNIVERSITÀ
CATTOLICA
del Sacro Cuore

Si consideri la seguente serie di numeri indici a base mobile:

<i>anno</i>	<i>NIBM</i>
2009	
2010	0.9000
2011	1.0333
2012	1.0968
2013	1.0882

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.

Esercizio 5.



1. Si costruiscano la serie dei numeri indici a base fissa (2009) e la serie dei dati iniziali sapendo che $x_{2009} = 200$.

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.

Esercizio 5.



1. Si costruiscano la serie dei numeri indici a base fissa (2009) e la serie dei dati iniziali sapendo che $x_{2009} = 200$.

Si ricorda che ${}_{2009}I_{2009} = 1$ e che per valori di $t > 2009$ vale

$${}_{2009}I_t = {}_{2009}I_{2010} \cdot {}_{2010}I_{2011} \cdot \dots \cdot {}_{t-1}I_t.$$

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.

Esercizio 5.



Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.

Esercizio 5.

1. Si costruiscano la serie dei numeri indici a base fissa (2009) e la serie dei dati iniziali sapendo che $x_{2009} = 200$.

Si ricorda che ${}_{2009}l_{2009} = 1$ e che per valori di $t > 2009$ vale

$${}_{2009}l_t = {}_{2009}l_{2010} \cdot {}_{2010}l_{2011} \cdot \dots \cdot {}_{t-1}l_t.$$

Si ottiene quindi

Anno	NIBM	NIBF (2009)
2009		1.00
2010	0.9000	0.90
2011	1.0333	0.93
2012	1.0968	1.02
2013	1.0882	1.11



Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.

Esercizio 5.



Sapendo che $x_{2009} = 200$ e che per valori di $t > 2009$ vale

$$x_t = x_{2009} \cdot {}_{2009}I_{2010} \cdot {}_{2010}I_{2011} \cdot \dots \cdot {}_{t-1}I_t.$$

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.

Esercizio 5.



Sapendo che $x_{2009} = 200$ e che per valori di $t > 2009$ vale

$$x_t = x_{2009} \cdot {}_{2009}l_{2010} \cdot {}_{2010}l_{2011} \cdot \dots \cdot {}_{t-1}l_t.$$

Si ottiene quindi

Anno	NIBM	x_t
2009		200
2010	0.9000	180
2011	1.0333	186
2012	1.0968	204
2013	1.0882	222

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.

Esercizio 5.



2. Si calcoli il tasso medio di variazione dal 2011 al 2013 e si commenti il risultato ottenuto.

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.

Esercizio 5.



2. Si calcoli il tasso medio di variazione dal 2011 al 2013 e si commenti il risultato ottenuto.
Si ricorda che

$${}_{2011}I_{2013} = {}_{2011}I_{2012} \cdot {}_{2012}I_{2013}$$

Si deve trovare il numero indice a base mobile medio.

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.

Esercizio 5.



Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.

Esercizio 5.

2. Si calcoli il tasso medio di variazione dal 2011 al 2013 e si commenti il risultato ottenuto. Si ricorda che

$${}_{2011}I_{2013} = {}_{2011}I_{2012} \cdot {}_{2012}I_{2013}$$

Si deve trovare il numero indice a base mobile medio. Applicando il criterio di scelta della media secondo Chisini si ottiene

$${}_{2011}I_{2013} = {}_{2011}I_{2012} \cdot {}_{2012}I_{2013} = \alpha \cdot \alpha = \alpha^2.$$



Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.

Esercizio 5.

2. Si calcoli il tasso medio di variazione dal 2011 al 2013 e si commenti il risultato ottenuto. Si ricorda che

$${}_{2011}l_{2013} = {}_{2011}l_{2012} \cdot {}_{2012}l_{2013}$$

Si deve trovare il numero indice a base mobile medio. Applicando il criterio di scelta della media secondo Chisini si ottiene

$${}_{2011}l_{2013} = {}_{2011}l_{2012} \cdot {}_{2012}l_{2013} = \alpha \cdot \alpha = \alpha^2.$$

Quindi

$$\alpha = \sqrt[2]{{}_{2011}l_{2013}} = \sqrt[2]{{}_{2011}l_{2012} \cdot {}_{2012}l_{2013}} = 1.0925.$$

Da cui segue una variazione relativa media pari a 0.0925, corrispondente a un incremento medio annuo tra il 2011 e il 2013 del 9.25%.



[Esercizio 1.](#)

[Esercizio 2.](#)

[Esercizio 3.](#)

[Esercizio 4.](#)

[Esercizio 5.](#)

Il seguente prospetto riporta i valori rilevati in corrispondenza a 5 unità statistiche con riferimento a due caratteri quantitativi X e Y

x_i	1	1	3	5	5
y_i	11	25	25	25	39

1. Si rappresentino graficamente le coppie dei dati e la funzione di regressione ipotizzando Y in funzione di X .



La funzione di regressione associa a ogni valore x appartenente all'insieme di definizione della variabile X la media di Y condizionata a x , $M(Y|x)$

x	$M(Y x)$
1	$\frac{1}{2}(11 + 25) = 18$
3	25
5	$\frac{1}{2}(25 + 39) = 32$

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.

Esercizio 5.



Esercizio 1.

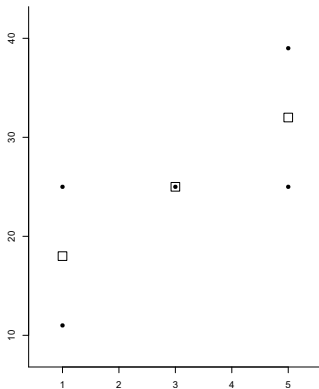
Esercizio 2.

Esercizio 3.

Esercizio 4.

Esercizio 5.

Le coppie di dati possono essere rappresentate mediante un grafico a dispersione.





2. Si calcolino, secondo il metodo dei minimi quadrati, i parametri e l'indice di adattamento dei seguenti modelli di regressione, individuando il modello migliore:

$$I) Y^* = a + bX \quad II) Y^* = cX$$

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.

Esercizio 5.

Si riportano i calcoli necessari¹ per ottenere le stime dei parametri e gli indici di adattamento del modello

$$Y = a + bX + E$$

x_i	y_i	$x_i y_i$	x_i^2	y_i^2
1	11	11	1	121
1	25	25	1	625
3	25	75	9	625
5	25	125	25	625
5	39	195	25	1521
15	125	431	61	3517
3	25	86.2	12.2	703.4
$M(X)$	$M(Y)$	$M(XY)$	$M(X^2)$	$M(Y^2)$

$$\hat{b} =$$



Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.

Esercizio 5.

¹Si osserva che per ragioni di stampa alcuni calcoli potrebbero essere duplicati con riferimento ai diversi modelli oggetto di studio

Si riportano i calcoli necessari¹ per ottenere le stime dei parametri e gli indici di adattamento del modello

$$Y = a + bX + E$$

x_i	y_i	$x_i y_i$	x_i^2	y_i^2
1	11	11	1	121
1	25	25	1	625
3	25	75	9	625
5	25	125	25	625
5	39	195	25	1521
15	125	431	61	3517
3	25	86.2	12.2	703.4
$M(X)$	$M(Y)$	$M(XY)$	$M(X^2)$	$M(Y^2)$

$$\begin{aligned}\hat{b} &= \frac{\text{Cov}(X, Y)}{\text{Var}(X)} = \frac{M(XY) - M(X)M(Y)}{M(X^2) - [M(X)]^2} = \\ &= \frac{86.2 - 3 \cdot 25}{12.2 - 3^2} = \frac{11.2}{3.2} = 3.5 \\ \hat{a} &= \end{aligned}$$



Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.

Esercizio 5.

¹Si osserva che per ragioni di stampa alcuni calcoli potrebbero essere duplicati con riferimento ai diversi modelli oggetto di studio

Si riportano i calcoli necessari¹ per ottenere le stime dei parametri e gli indici di adattamento del modello

$$Y = a + bX + E$$

x_i	y_i	$x_i y_i$	x_i^2	y_i^2
1	11	11	1	121
1	25	25	1	625
3	25	75	9	625
5	25	125	25	625
5	39	195	25	1521
15	125	431	61	3517
3	25	86.2	12.2	703.4
$M(X)$	$M(Y)$	$M(XY)$	$M(X^2)$	$M(Y^2)$

$$\hat{b} = \frac{\text{Cov}(X, Y)}{\text{Var}(X)} = \frac{M(XY) - M(X)M(Y)}{M(X^2) - [M(X)]^2} =$$

$$= \frac{86.2 - 3 \cdot 25}{12.2 - 3^2} = \frac{11.2}{3.2} = 3.5$$

$$\hat{a} = M(Y) - \hat{b}M(X) = 14.5$$

$$\rho^2 =$$



Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.

Esercizio 5.

¹Si osserva che per ragioni di stampa alcuni calcoli potrebbero essere duplicati con riferimento ai diversi modelli oggetto di studio

Si riportano i calcoli necessari¹ per ottenere le stime dei parametri e gli indici di adattamento del modello

$$Y = a + bX + E$$

x_i	y_i	$x_i y_i$	x_i^2	y_i^2
1	11	11	1	121
1	25	25	1	625
3	25	75	9	625
5	25	125	25	625
5	39	195	25	1521
15	125	431	61	3517
3	25	86.2	12.2	703.4
$M(X)$	$M(Y)$	$M(XY)$	$M(X^2)$	$M(Y^2)$

$$\begin{aligned}\hat{b} &= \frac{\text{Cov}(X, Y)}{\text{Var}(X)} = \frac{M(XY) - M(X)M(Y)}{M(X^2) - [M(X)]^2} = \\ &= \frac{86.2 - 3 \cdot 25}{12.2 - 3^2} = \frac{11.2}{3.2} = 3.5 \\ \hat{a} &= M(Y) - \hat{b}M(X) = 14.5 \\ \rho^2 &= \frac{[\text{Cov}(X, Y)]^2}{\text{Var}(X) \text{Var}(Y)} = \frac{11.2^2}{3.2 \cdot 78.4} = 0.5\end{aligned}$$

$$M(\hat{E}^2) = \text{Var}(\hat{E}) = \text{varianza residua} =$$

¹Si osserva che per ragioni di stampa alcuni calcoli potrebbero essere duplicati con riferimento ai diversi modelli oggetto di studio



Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.

Esercizio 5.

Si riportano i calcoli necessari¹ per ottenere le stime dei parametri e gli indici di adattamento del modello

$$Y = a + bX + E$$

x_i	y_i	$x_i y_i$	x_i^2	y_i^2
1	11	11	1	121
1	25	25	1	625
3	25	75	9	625
5	25	125	25	625
5	39	195	25	1521
15	125	431	61	3517
3	25	86.2	12.2	703.4
$M(X)$	$M(Y)$	$M(XY)$	$M(X^2)$	$M(Y^2)$

$$\hat{b} = \frac{\text{Cov}(X, Y)}{\text{Var}(X)} = \frac{M(XY) - M(X)M(Y)}{M(X^2) - [M(X)]^2} =$$

$$= \frac{86.2 - 3 \cdot 25}{12.2 - 3^2} = \frac{11.2}{3.2} = 3.5$$

$$\hat{a} = M(Y) - \hat{b}M(X) = 14.5$$

$$\rho^2 = \frac{[\text{Cov}(X, Y)]^2}{\text{Var}(X) \text{Var}(Y)} = \frac{11.2^2}{3.2 \cdot 78.4} = 0.5$$

$$M(\hat{E}^2) = \text{Var}(\hat{E}) = \text{varianza residua} = \text{Var}(Y) \cdot (1 - \rho^2) = 39.2$$

¹Si osserva che per ragioni di stampa alcuni calcoli potrebbero essere duplicati con riferimento ai diversi modelli oggetto di studio





Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.

Esercizio 5.

Si riportano i calcoli necessari² per ottenere le stime dei parametri e la misura di adattamento del modello

$$Y = cX + E$$

x_i	y_i	$x_i y_i$	x_i^2	y_i^2
1	11	11	1	121
1	25	25	1	625
3	25	75	9	625
5	25	125	25	625
5	39	195	25	1521
		431	61	3517
		86.2	12.2	703.4
		$M(XY)$	$M(X^2)$	$M(Y^2)$

$$\hat{c} =$$

²Si osserva che per ragioni di stampa alcuni calcoli potrebbero essere duplicati con riferimento ai diversi modelli oggetto di studio



Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.

Esercizio 5.

Si riportano i calcoli necessari² per ottenere le stime dei parametri e la misura di adattamento del modello

$$Y = cX + E$$

x_i	y_i	$x_i y_i$	x_i^2	y_i^2
1	11	11	1	121
1	25	25	1	625
3	25	75	9	625
5	25	125	25	625
5	39	195	25	1521
		431	61	3517
		86.2	12.2	703.4
		$M(XY)$	$M(X^2)$	$M(Y^2)$

$$\hat{c} = \frac{M(XY)}{M(X^2)} = \frac{86.2}{12.2} = 7.0656$$

$$M(\hat{E}^2) =$$

²Si osserva che per ragioni di stampa alcuni calcoli potrebbero essere duplicati con riferimento ai diversi modelli oggetto di studio



Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.

Esercizio 5.

Si riportano i calcoli necessari² per ottenere le stime dei parametri e la misura di adattamento del modello

$$Y = cX + E$$

x_i	y_i	$x_i y_i$	x_i^2	y_i^2
1	11	11	1	121
1	25	25	1	625
3	25	75	9	625
5	25	125	25	625
5	39	195	25	1521
		431	61	3517
		86.2	12.2	703.4
		$M(XY)$	$M(X^2)$	$M(Y^2)$

$$\hat{c} = \frac{M(XY)}{M(X^2)} = \frac{86.2}{12.2} = 7.0656$$

$$M(\hat{E}^2) = M[Y \cdot (Y - \hat{c}X)] = M(Y^2) - \hat{c}M(XY) = 703.4 - 7.0656 \cdot 86.2 = 94.3475$$

²Si osserva che per ragioni di stampa alcuni calcoli potrebbero essere duplicati con riferimento ai diversi modelli oggetto di studio



Il modello migliore è $Y^* = 14.5 + 3.5X$.

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.

Esercizio 5.



3. Per il modello migliore, si fornisca la previsione in corrispondenza di $X = 3$.

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.

Esercizio 5.



3. Per il modello migliore, si fornisca la previsione in corrispondenza di $X = 3$.

$$Y^*(3) = 14.5 + 3.5 \cdot 3 = 25$$

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.

Esercizio 5.

Esercizio 4

Prova scritta di
STATISTICA del
04.09.2014

Giuseppe Boari
Gabriele Cantaluppi



UNIVERSITÀ
CATTOLICA
del Sacro Cuore

Si completino le seguenti tabelle, giustificando la risposta, in modo che siano soddisfatte le condizioni indicate.

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.

Esercizio 5.



	$y_1 = 10$	$y_2 = 20$	$y_3 = 30$
$x_1 = 1$	0	11	
$x_2 = 2$		0	

a) $\eta_{Y|X}^2 > \rho^2$, con $\rho < 0$

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.

Esercizio 5.



	$y_1 = 10$	$y_2 = 20$	$y_3 = 30$
$x_1 = 1$	0	11	
$x_2 = 2$		0	

a) $\eta_{Y|X}^2 > \rho^2$, con $\rho < 0$

Impossibile: $\eta_{Y|X}^2$ risulta sempre uguale a ρ^2
in quanto figurano due sole medie condizionate
che saranno sempre allineate.

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.

Esercizio 5.



	$y_1 = 10$	$y_2 = 20$	$y_3 = 30$
$x_1 = 1$		11	
$x_2 = 2$		11	

b) $\eta_{Y|X}^2 = 1$

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.

Esercizio 5.



Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.

Esercizio 5.

	$y_1 = 10$	$y_2 = 20$	$y_3 = 30$
$x_1 = 1$		11	
$x_2 = 2$		11	

b) $\eta_{Y|X}^2 = 1$

Impossibile: $\eta_{Y|X}^2 = 1$ quando

sussiste dipendenza funzionale di Y da X

→ distribuzioni condizionate $Y|x_1$ e $Y|x_2$ degeneri.

Se si ponesse $n_{11} = n_{13} = n_{21} = n_{23} = 0$ le medie risulterebbero tutte identiche ($\eta_{Y|X}^2 = 0$) e la tabella non sarebbe comunque ammissibile.

Esercizio 5

Prova scritta di
STATISTICA del
04.09.2014

Giuseppe Boari
Gabriele Cantaluppi



UNIVERSITÀ
CATTOLICA
del Sacro Cuore

Per il seguente gioco è prevista una posta di 2 Euro: si lanciano due dadi e se il prodotto dei numeri ottenuti è pari a 4 si vincono 6 Euro.

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.

Esercizio 5.

1. Si stabilisca se il gioco è equo.



Esercizio 1.

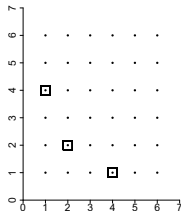
Esercizio 2.

Esercizio 3.

Esercizio 4.

Esercizio 5.

1. Si stabilisca se il gioco è equo.
Con riferimento allo spazio probabilistico



abbiamo

$$P(\text{vincita}) =$$



Esercizio 1.

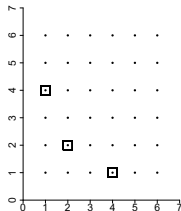
Esercizio 2.

Esercizio 3.

Esercizio 4.

Esercizio 5.

1. Si stabilisca se il gioco è equo.
Con riferimento allo spazio probabilistico



abbiamo

$$P(\text{vincita}) = \frac{3}{36} = 0.08333$$



Esercizio 1.

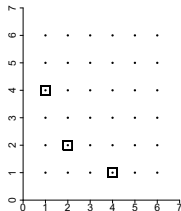
Esercizio 2.

Esercizio 3.

Esercizio 4.

Esercizio 5.

1. Si stabilisca se il gioco è equo.
Con riferimento allo spazio probabilistico



abbiamo

$$P(\text{vincita}) = \frac{3}{36} = 0.08333$$

La variabile casuale guadagno risulta, quindi

	x_i	$P(X = x_i)$	$x_i \cdot P(X = x_i)$
<i>posta</i>	-2	0.91667	-1.83333
<i>vincita - posta</i>	4	0.08333	0.33333
	1		-1.50000



Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.

Esercizio 5.



Esercizio 1.

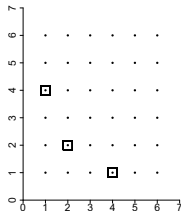
Esercizio 2.

Esercizio 3.

Esercizio 4.

Esercizio 5.

1. Si stabilisca se il gioco è equo.
Con riferimento allo spazio probabilistico



abbiamo

$$P(\text{vincita}) = \frac{3}{36} = 0.08333$$

La variabile casuale guadagno risulta, quindi

	x_i	$P(X = x_i)$	$x_i \cdot P(X = x_i)$
<i>posta</i>	-2	0.91667	-1.83333
<i>vincita - posta</i>	4	0.08333	0.33333
		1	-1.50000

Il valore atteso della variabile casuale guadagno è negativo, quindi il gioco non è equo.



2. Nell'ipotesi di ripetere il gioco 4 volte si calcoli la probabilità di vincere al massimo 1 volta.

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.

Esercizio 5.



2. Nell'ipotesi di ripetere il gioco 4 volte si calcoli la probabilità di vincere al massimo 1 volta. Sia X la variabile casuale 'n. di vincite nelle 4 prove indipendenti'.

$$X \sim$$

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.

Esercizio 5.



2. Nell'ipotesi di ripetere il gioco 4 volte si calcoli la probabilità di vincere al massimo 1 volta. Sia X la variabile casuale 'n. di vincite nelle 4 prove indipendenti'.

$$X \sim \text{Bin}(n = 4, p = 0.08333).$$

$$P(X < 1) =$$

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.

Esercizio 5.



Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.

Esercizio 5.

2. Nell'ipotesi di ripetere il gioco 4 volte si calcoli la probabilità di vincere al massimo 1 volta. Sia X la variabile casuale 'n. di vincite nelle 4 prove indipendenti'.

$$X \sim \text{Bin}(n = 4, p = 0.08333).$$

$$\begin{aligned} P(X < 1) &= P(X = 0) + P(X = 1) = \\ &= \binom{4}{0} 0.08333^0 \cdot 0.91667^4 + \binom{4}{1} 0.08333^1 \cdot 0.91667^3 = \\ &= 0.7061 + 0.2568 = 0.9628. \end{aligned}$$

Tema T280

Prova scritta di STATISTICA del 18.09.2014

Prova scritta di
STATISTICA del
18.09.2014

Giuseppe Boari
Gabriele Cantaluppi



UNIVERSITÀ
CATTOLICA
del Sacro Cuore

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.

Giuseppe Boari
Gabriele Cantaluppi
Università Cattolica del Sacro Cuore, Milano



Nella tabella sottostante vengono riportati i risultati di un'indagine sull'uso dei pesticidi nei terreni agricoli in Italia.

Regione (R) / Pesticidi in mg/ha (P)	0 – 25	25 – 50	50 – 100
nord	10	6	9
centro	5	8	9
sud	5	11	12

1. Si rappresenti graficamente la variabile marginale P .

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.



Si considera la distribuzione marginale della variabile P , ottenuta sommando le frequenze di colonna

p_j	$n_{\bullet j}$
0 $\dashv$ 25	20
25 $\dashv$ 50	25
50 $\dashv$ 100	30

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.



Si tratta di un carattere quantitativo con modalità raggruppate in classi che può essere rappresentato graficamente mediante un istogramma.

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.



L'istogramma può essere costruito dopo avere calcolato le densità $d_{\bullet j} = \frac{n_{\bullet j}}{a_j}$, (dove a_j sono le ampiezze di classe).

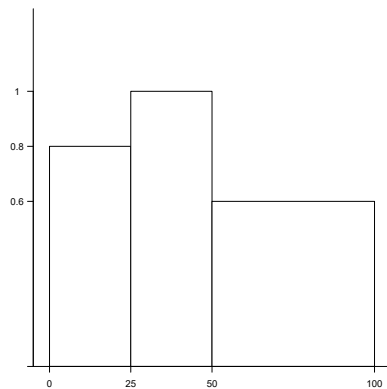
p_j	$n_{\bullet j}$	$a_{\bullet j}$	$d_{\bullet j}$
0 - 25	20	$25 - 0 = 25$	0.8
25 - 50	25	$50 - 25 = 25$	1
50 - 100	30	$100 - 50 = 50$	0.6

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.



Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.



2. Si calcolino, se possibile, moda, media aritmetica e mediana per la variabile R . Si calcolino, se possibile, moda, media aritmetica e si individui su un opportuno grafico la mediana per la variabile P .

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.



Si considera la distribuzione marginale della variabile R , ottenuta sommando le frequenze di riga

r_i	$n_{i\bullet}$
<i>nord</i>	25
<i>centro</i>	22
<i>sud</i>	28

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.



Si tratta di un carattere qualitativo sconnesso.

L'unico indice di posizione che può essere definito è la moda (categoria alla quale è associata la massima frequenza) che risulta *sud*.

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.



Si considera la distribuzione marginale della variabile P , ottenuta sommando le frequenze di colonna

p_j	$n_{\bullet j}$
0 $\vdash$ 25	20
25 $\vdash$ 50	25
50 $\vdash$ 100	30

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.



Si tratta di un carattere quantitativo con modalità raggruppate in classi.

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.



La moda corrisponde alla classe alla quale è associata la massima densità di frequenza

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.



La moda corrisponde alla classe alla quale è associata la massima densità di frequenza

p_j	$n_{\bullet j}$	$a_{\bullet j}$	$d_{\bullet j}$
$0 \vdash 25$	20	$25 - 0 = 25$	0.8
$25 \vdash 50$	25	$50 - 25 = 25$	1
$50 \vdash 100$	30	$100 - 50 = 50$	0.6

e risulta $25 \vdash 50$.

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.



La mediana (percentile di ordine 0.50) può essere ottenuta considerando l'inversa della Funzione di Ripartizione in corrispondenza di 0.50.

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.



La mediana (percentile di ordine 0.50) può essere ottenuta considerando l'inversa della Funzione di Ripartizione in corrispondenza di 0.50.

p_j	$n_{\bullet j}$	$N_{\bullet j}$	$F_{\bullet j}$
0 - 25	20	20	0.2667
25 - 50	25	45	0.6
50 - 100	30	75	1

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.

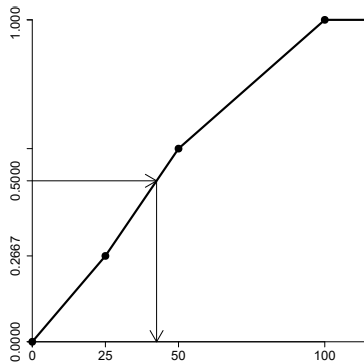


Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.



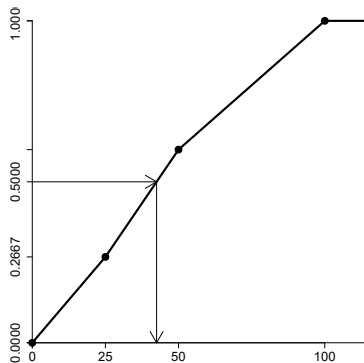


Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.



Si ottiene

$$p_{0.5} = 25 + \frac{0.5 - 0.2667}{0.6 - 0.2667} \cdot (50 - 25) = 42.5.$$



La media aritmetica

$$M(P) = \frac{1}{n} \sum_{j=1}^k p_j n_{\bullet j}$$

risulta 45.8333.

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.



La media aritmetica

$$M(P) = \frac{1}{n} \sum_{j=1}^k p_j n_{\bullet j}$$

risulta 45.8333.

Si ottiene (considerando come modalità i valori centrali delle classi)

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.



La media aritmetica

$$M(P) = \frac{1}{n} \sum_{j=1}^k p_j n_{\bullet j}$$

risulta 45.8333.

Si ottiene (considerando come modalità i valori centrali delle classi)

p_j	$n_{\bullet j}$	$p_j n_{\bullet j}$
12.5	20	250.0
37.5	25	937.5
75.0	30	2250.0
	75	3437.5

da cui segue

$$M(P) = \frac{1}{n} \sum_{i=1}^k p_j n_{\bullet j} = \frac{3437.5}{75} = 45.8333.$$

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.



Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.

3. Si calcoli per la variabile R un opportuno indice di variabilità e se ne commenti il risultato.



Si considera la distribuzione marginale della variabile R , ottenuta sommando le frequenze di riga

r_i	$n_{i\bullet}$
<i>nord</i>	25
<i>centro</i>	22
<i>sud</i>	28

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.

In presenza di un carattere qualitativo la variabilità può essere misurata mediante gli indici di Gini e di Frosini normalizzati

$$G_N = \frac{1 - \sum_{i=1}^h f_{i\bullet}^2}{1 - \frac{1}{h}}$$

e

$$F_N = 1 - \sqrt{1 - G_N}.$$



Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.



In presenza di un carattere qualitativo la variabilità può essere misurata mediante gli indici di Gini e di Frosini normalizzati

$$G_N = \frac{1 - \sum_{i=1}^h f_{i\bullet}^2}{1 - \frac{1}{h}}$$

e

$$F_N = 1 - \sqrt{1 - G_N}.$$

Si ottiene:

r_i	$n_{i\bullet}$	$f_{i\bullet}$	$f_{i\bullet}^2$
<i>nord</i>	25	0.3333	0.1111
<i>centro</i>	22	0.2933	0.086
<i>sud</i>	28	0.3733	0.1394
	75	1	0.3365

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.



In presenza di un carattere qualitativo la variabilità può essere misurata mediante gli indici di Gini e di Frosini normalizzati

$$G_N = \frac{1 - \sum_{i=1}^h f_{i\bullet}^2}{1 - \frac{1}{h}}$$

e

$$F_N = 1 - \sqrt{1 - G_N}.$$

Si ottiene:

r_i	$n_{i\bullet}$	$f_{i\bullet}$	$f_{i\bullet}^2$
<i>nord</i>	25	0.3333	0.1111
<i>centro</i>	22	0.2933	0.086
<i>sud</i>	28	0.3733	0.1394
	75	1	0.3365

Da cui segue:

$$G_N = \frac{1 - 0.3365}{1 - \frac{1}{3}} = 0.9952$$

e

$$F_N = 1 - \sqrt{1 - 0.9952} = 0.9307.$$

Sussiste, quindi, un livello molto alto di eterogeneità.



4. Si calcoli un opportuno indice di connessione.

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.



Dalla distribuzione delle frequenze congiunte n_{ij} :

	0 + 25	25 + 50	50 + 100	$n_{i\bullet}$
<i>nord</i>	10	6	9	25
<i>centro</i>	5	8	9	22
<i>sud</i>	5	11	12	28
$n_{\bullet j}$	20	25	30	75

è possibile ottenere le frequenze teoriche nella situazione di indipendenza stocastica

$$\hat{n}_{ij} = \frac{n_{i\bullet} n_{\bullet j}}{n}:$$

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.



Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.

Dalla distribuzione delle frequenze congiunte n_{ij} :

	0 - 25	25 - 50	50 - 100	$n_{i\bullet}$
<i>nord</i>	10	6	9	25
<i>centro</i>	5	8	9	22
<i>sud</i>	5	11	12	28
$n_{\bullet j}$	20	25	30	75

è possibile ottenere le frequenze teoriche nella situazione di indipendenza stocastica

$$\hat{n}_{ij} = \frac{n_{i\bullet} \cdot n_{\bullet j}}{n}:$$

	0 - 25	25 - 50	50 - 100	$n_{i\bullet}$
<i>nord</i>	6.6667	8.3333	10.0	25
<i>centro</i>	5.8667	7.3333	8.8	22
<i>sud</i>	7.4667	9.3333	11.2	28
$n_{\bullet j}$	20.0000	25.0000	30.0	75



le contingenze assolute $n_{ij} - \hat{n}_{ij}$:

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.



le contingenze assolute $n_{ij} - \hat{n}_{ij}$:

	0 - 25	25 - 50	50 - 100
<i>nord</i>	3.3333	-2.3333	-1.0
<i>centro</i>	-0.8667	0.6667	0.2
<i>sud</i>	-2.4667	1.6667	0.8

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.



i rapporti di contingenza $\frac{n_{ij}}{\hat{n}_{ij}}$:

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.



i rapporti di contingenza $\frac{n_{ij}}{\hat{n}_{ij}}$:

	0 - 25	25 - 50	50 - 100
<i>nord</i>	1.5000	0.7200	0.9000
<i>centro</i>	0.8523	1.0909	1.0227
<i>sud</i>	0.6696	1.1786	1.0714

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.



e le contingenze relative, espresse in termini percentuali, $100 \left(\frac{n_{ij} - \hat{n}_{ij}}{\hat{n}_{ij}} \right) \% = 100 \left(\frac{n_{ij}}{\hat{n}_{ij}} - 1 \right) \%$:

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.



e le contingenze relative, espresse in termini percentuali, $100 \left(\frac{n_{ij} - \hat{n}_{ij}}{\hat{n}_{ij}} \right) \% = 100 \left(\frac{n_{ij}}{\hat{n}_{ij}} - 1 \right) \%$:

	0 ÷ 25	25 ÷ 50	50 ÷ 100
<i>nord</i>	50.00	-28.00	-10.00
<i>centro</i>	-14.77	9.09	2.27
<i>sud</i>	-33.04	17.86	7.14

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.



Risultati intermedi $\frac{(n_{ij} - \hat{n}_{ij})^2}{\hat{n}_{ij}}$ per il calcolo del χ^2

	0 - 25	25 - 50	50 - 100
<i>nord</i>	1.6667	0.6533	0.1000
<i>centro</i>	0.1280	0.0606	0.0045
<i>sud</i>	0.8149	0.2976	0.0571

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.



Risultati intermedi $\frac{n_{ij}^2}{n_{i\bullet} \cdot n_{\bullet j}}$ per il calcolo del χ^2 (formula operativa)

	0 ÷ 25	25 ÷ 50	50 ÷ 100
<i>nord</i>	0.2000	0.0576	0.1080
<i>centro</i>	0.0568	0.1164	0.1227
<i>sud</i>	0.0446	0.1729	0.1714

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.



Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.

$$\chi^2 = \sum_{i=1}^h \sum_{j=1}^k \frac{(n_{ij} - \hat{n}_{ij})^2}{\hat{n}_{ij}} = 3.7828,$$

$$\chi_N^2 = \frac{\chi^2}{n \cdot \min(h-1, k-1)} = 0.0252$$

$$\chi_N = \sqrt{\chi_N^2} = 0.1587.$$

Sussiste, quindi, un livello di connessione molto basso.



Si consideri la seguente serie di numeri indici a base mobile dei prezzi di un titolo azionario:

<i>anno</i>	<i>NIBM</i>
2008	0.8
2009	1.5
2010	1.5
2011	1.5
2012	1.5

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.



1. Si costruisca la corrispondente serie dei numeri indici a base fissa (anno base 2008).

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.



1. Si costruisca la corrispondente serie dei numeri indici a base fissa (anno base 2008).
Si ricorda che ${}_{2008}I_{2008} = 1$ e che per valori di $t > 2008$ vale

$${}_{2008}I_t = {}_{2008}I_{2009} \cdot {}_{2009}I_{2010} \cdot \dots \cdot {}_{t-1}I_t.$$

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.



Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.

1. Si costruisca la corrispondente serie dei numeri indici a base fissa (anno base 2008).
Si ricorda che ${}_{2008}I_{2008} = 1$ e che per valori di $t > 2008$ vale

$${}_{2008}I_t = {}_{2008}I_{2009} \cdot {}_{2009}I_{2010} \cdot \dots \cdot {}_{t-1}I_t.$$

Si ottiene quindi

Anno	NIBM	NIBF (2008)
2008	0.8	1.0000
2009	1.5	1.5000
2010	1.5	2.2500
2011	1.5	3.3750
2012	1.5	5.0625



2. Se nel 2008 fossero state acquistate 100 azioni a un prezzo di 10 Euro per azione, si calcoli il valore totale di queste azioni nel 2012.

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.



Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.

2. Se nel 2008 fossero state acquistate 100 azioni a un prezzo di 10 Euro per azione, si calcoli il valore totale di queste azioni nel 2012.

Da $x_{2008} = 100 \cdot 10 = 1000$ e da

$${}_{2008}I_{2012} = \frac{x_{2012}}{x_{2008}} = 5.0625$$

si ottiene

$$x_{2012} = x_{2008} \cdot 5.0625 = 1000 \cdot 5.0625 = 5062.50.$$



3. Si calcoli il tasso medio di variazione dal 2009 al 2012 e si commenti il risultato ottenuto.

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.



3. Si calcoli il tasso medio di variazione dal 2009 al 2012 e si commenti il risultato ottenuto.
Si ricorda che

$${}_{2009}I_{2012} = {}_{2009}I_{2010} \cdot {}_{2010}I_{2011} \cdot {}_{2011}I_{2012}.$$

Si deve trovare il numero indice a base mobile medio.

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.



3. Si calcoli il tasso medio di variazione dal 2009 al 2012 e si commenti il risultato ottenuto. Si ricorda che

$${}_{2009}I_{2012} = {}_{2009}I_{2010} \cdot {}_{2010}I_{2011} \cdot {}_{2011}I_{2012}.$$

Si deve trovare il numero indice a base mobile medio. Applicando il criterio di scelta della media secondo Chisini si ottiene

$${}_{2009}I_{2012} = {}_{2009}I_{2010} \cdot {}_{2010}I_{2011} \cdot {}_{2011}I_{2012} = \alpha \cdot \alpha \cdot \alpha = \alpha^3.$$

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.



Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.

3. Si calcoli il tasso medio di variazione dal 2009 al 2012 e si commenti il risultato ottenuto. Si ricorda che

$${}_{2009}I_{2012} = {}_{2009}I_{2010} \cdot {}_{2010}I_{2011} \cdot {}_{2011}I_{2012}.$$

Si deve trovare il numero indice a base mobile medio. Applicando il criterio di scelta della media secondo Chisini si ottiene

$${}_{2009}I_{2012} = {}_{2009}I_{2010} \cdot {}_{2010}I_{2011} \cdot {}_{2011}I_{2012} = \alpha \cdot \alpha \cdot \alpha = \alpha^3.$$

Quindi

$$\alpha = \sqrt[3]{{}_{2009}I_{2012}} = \sqrt[3]{{}_{2009}I_{2010} \cdot {}_{2010}I_{2011} \cdot {}_{2011}I_{2012}} = 1.5.$$

Da cui segue una variazione relativa media pari a 0.5 , corrispondente a un incremento medio annuo tra il 2009 e il 2012 del 50%.

Si osserva dalla serie dei numeri indici a base mobile che l'incremento nel periodo considerato è stato costante (la serie dei NIBM tra il 2009 e il 2012 è degenera).



Il seguente prospetto riporta i valori rilevati in corrispondenza del numero di bagnanti (Y) presenti in una piscina in 5 domeniche in funzione della situazione meteo rilevata con una variabile dicotomica X ($= 0$ se nuvoloso e $= 1$ se soleggiato)

x_i	0	0	0	1	1
y_i	22	20	24	32	34

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.



1. Si rappresentino graficamente le coppie dei dati e la funzione di regressione ipotizzando Y in funzione di X .

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.



La funzione di regressione associa a ogni valore x appartenente all'insieme di definizione della variabile X la media di Y condizionata a x , $M(Y|x)$

x	$M(Y x)$
0	$\frac{1}{3}(20 + 22 + 24) = 22$
1	$\frac{1}{2}(32 + 34) = 33$

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.



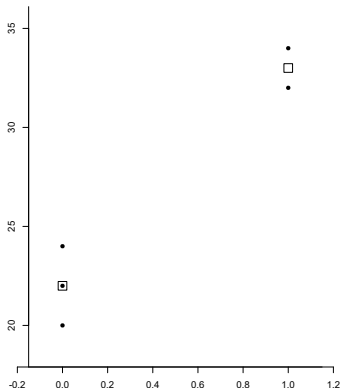
Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.

Le coppie di dati possono essere rappresentate mediante un grafico a dispersione.





2. Si calcolino, secondo il metodo dei minimi quadrati, i parametri e l'indice di adattamento dei seguenti modelli di regressione, individuando il modello migliore:

$$I) Y^* = a + bX \qquad II) Y^* = c$$

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.



Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.

Si riportano i calcoli necessari¹ per ottenere le stime dei parametri e gli indici di adattamento del modello

$$Y = a + bX + E$$

x_i	y_i	$x_i y_i$	x_i^2	y_i^2
0	22	0	0	484
0	20	0	0	400
0	24	0	0	576
1	32	32	1	1024
1	34	34	1	1156
2	132	66	2	3640
0.4	26.4	13.2	0.4	728
$M(X)$	$M(Y)$	$M(XY)$	$M(X^2)$	$M(Y^2)$

$$\hat{b} =$$

¹Si osserva che per ragioni di stampa alcuni calcoli potrebbero essere duplicati con riferimento ai diversi modelli oggetto di studio



Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.

Si riportano i calcoli necessari¹ per ottenere le stime dei parametri e gli indici di adattamento del modello

$$Y = a + bX + E$$

x_i	y_i	$x_i y_i$	x_i^2	y_i^2
0	22	0	0	484
0	20	0	0	400
0	24	0	0	576
1	32	32	1	1024
1	34	34	1	1156
2	132	66	2	3640
0.4	26.4	13.2	0.4	728
$M(X)$	$M(Y)$	$M(XY)$	$M(X^2)$	$M(Y^2)$

$$\begin{aligned}\hat{b} &= \frac{\text{Cov}(X, Y)}{\text{Var}(X)} = \frac{M(XY) - M(X)M(Y)}{M(X^2) - [M(X)]^2} = \\ &= \frac{13.2 - 0.4 \cdot 26.4}{0.4 - 0.4^2} = \frac{2.64}{0.24} = 11 \\ \hat{a} &= \end{aligned}$$

¹Si osserva che per ragioni di stampa alcuni calcoli potrebbero essere duplicati con riferimento ai diversi modelli oggetto di studio



Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.

Si riportano i calcoli necessari¹ per ottenere le stime dei parametri e gli indici di adattamento del modello

$$Y = a + bX + E$$

x_i	y_i	$x_i y_i$	x_i^2	y_i^2
0	22	0	0	484
0	20	0	0	400
0	24	0	0	576
1	32	32	1	1024
1	34	34	1	1156
2	132	66	2	3640
0.4	26.4	13.2	0.4	728
$M(X)$	$M(Y)$	$M(XY)$	$M(X^2)$	$M(Y^2)$

$$\begin{aligned}\hat{b} &= \frac{\text{Cov}(X, Y)}{\text{Var}(X)} = \frac{M(XY) - M(X)M(Y)}{M(X^2) - [M(X)]^2} = \\ &= \frac{13.2 - 0.4 \cdot 26.4}{0.4 - 0.4^2} = \frac{2.64}{0.24} = 11\end{aligned}$$

$$\hat{a} = M(Y) - \hat{b}M(X) = 22$$

$$\rho^2 =$$

¹Si osserva che per ragioni di stampa alcuni calcoli potrebbero essere duplicati con riferimento ai diversi modelli oggetto di studio



Si riportano i calcoli necessari¹ per ottenere le stime dei parametri e gli indici di adattamento del modello

$$Y = a + bX + E$$

x_i	y_i	$x_i y_i$	x_i^2	y_i^2
0	22	0	0	484
0	20	0	0	400
0	24	0	0	576
1	32	32	1	1024
1	34	34	1	1156
2	132	66	2	3640
0.4	26.4	13.2	0.4	728
$M(X)$	$M(Y)$	$M(XY)$	$M(X^2)$	$M(Y^2)$

$$\begin{aligned}\hat{b} &= \frac{\text{Cov}(X, Y)}{\text{Var}(X)} = \frac{M(XY) - M(X)M(Y)}{M(X^2) - [M(X)]^2} = \\ &= \frac{13.2 - 0.4 \cdot 26.4}{0.4 - 0.4^2} = \frac{2.64}{0.24} = 11 \\ \hat{a} &= M(Y) - \hat{b}M(X) = 22 \\ \rho^2 &= \frac{[\text{Cov}(X, Y)]^2}{\text{Var}(X) \text{Var}(Y)} = \frac{2.64^2}{0.24 \cdot 31.04} = 0.9356\end{aligned}$$

¹Si osserva che per ragioni di stampa alcuni calcoli potrebbero essere duplicati con riferimento ai diversi modelli oggetto di studio



Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.

Si riportano i calcoli necessari² per ottenere le stime dei parametri e gli indici di adattamento del modello

$$Y = c + E$$

x_i	y_i	y_i^2
0	22	484
0	20	400
0	24	576
1	32	1024
1	34	1156
132		3640
26.4		728
$M(Y)$		$M(Y^2)$

$$\hat{c} =$$

²Si osserva che per ragioni di stampa alcuni calcoli potrebbero essere duplicati con riferimento ai diversi modelli oggetto di studio



Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.

Si riportano i calcoli necessari² per ottenere le stime dei parametri e gli indici di adattamento del modello

$$Y = c + E$$

x_i	y_i	y_i^2
0	22	484
0	20	400
0	24	576
1	32	1024
1	34	1156
132		3640
26.4		728
$M(Y)$		$M(Y^2)$

$$\hat{c} = M(Y) = 26.4$$

$$R^2 =$$

²Si osserva che per ragioni di stampa alcuni calcoli potrebbero essere duplicati con riferimento ai diversi modelli oggetto di studio



Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.

Si riportano i calcoli necessari² per ottenere le stime dei parametri e gli indici di adattamento del modello

$$Y = c + E$$

x_i	y_i	y_i^2
0	22	484
0	20	400
0	24	576
1	32	1024
1	34	1156
132		3640
26.4		728
$M(Y)$		$M(Y^2)$

$$\hat{c} = M(Y) = 26.4$$

$$R^2 = 0$$

²Si osserva che per ragioni di stampa alcuni calcoli potrebbero essere duplicati con riferimento ai diversi modelli oggetto di studio



Il modello migliore è $Y^* = 22 + 11X$.

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.



3. Si determini, indicando il ragionamento, il valore dei rapporti di correlazione $\eta_{Y|X}^2$ e $\eta_{X|Y}^2$.

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.



3. Si determini, indicando il ragionamento, il valore dei rapporti di correlazione $\eta_{Y|X}^2$ e $\eta_{X|Y}^2$.
 $\eta_{Y|X}^2 = \rho^2 = 0.936$ in quanto le medie condizionate sono allineate.

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.



Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.

- 3.** Si determini, indicando il ragionamento, il valore dei rapporti di correlazione $\eta_{Y|X}^2$ e $\eta_{X|Y}^2$.
 $\eta_{Y|X}^2 = \rho^2 = 0.936$ in quanto le medie condizionate sono allineate.
 $\eta_{X|Y}^2 = 1$ sussiste dipendenza funzionale di X da Y (a ogni valore di Y corrisponde un solo valore distinto di X).

Esercizio 4

Prova scritta di
STATISTICA del
18.09.2014

Giuseppe Boari
Gabriele Cantaluppi



UNIVERSITÀ
CATTOLICA
del Sacro Cuore

Sia 0.06 la probabilità di colpire il centro del bersaglio lanciando una freccetta.

1. Si calcoli la probabilità di colpire il centro almeno una volta su 20 lanci.

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.

Esercizio 4

Prova scritta di
STATISTICA del
18.09.2014

Giuseppe Boari
Gabriele Cantaluppi



UNIVERSITÀ
CATTOLICA
del Sacro Cuore

Sia 0.06 la probabilità di colpire il centro del bersaglio lanciando una freccetta.

1. Si calcoli la probabilità di colpire il centro almeno una volta su 20 lanci.

Sia X la variabile casuale 'numero di volte che si colpisce il centro nelle 20 prove indipendenti'.

$$X \sim$$

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.



Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.

Sia 0.06 la probabilità di colpire il centro del bersaglio lanciando una freccetta.

1. Si calcoli la probabilità di colpire il centro almeno una volta su 20 lanci.

Sia X la variabile casuale 'numero di volte che si colpisce il centro nelle 20 prove indipendenti'.

$$X \sim \text{Bin}(n = 20, p = 0.06).$$

$$P(X \geq 1) =$$



Sia 0.06 la probabilità di colpire il centro del bersaglio lanciando una freccetta.

1. Si calcoli la probabilità di colpire il centro almeno una volta su 20 lanci.

Sia X la variabile casuale 'numero di volte che si colpisce il centro nelle 20 prove indipendenti'.

$$X \sim \text{Bin}(n = 20, p = 0.06).$$

$$P(X \geq 1) = 1 - P(X = 0) = 1 - \binom{20}{0} 0.06^0 0.94^{20} = 1 - 0.2901 = 0.7099.$$

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.



2. Si calcoli la probabilità che, su 100 lanci, si colpisca almeno 6 volte il centro.

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.



2. Si calcoli la probabilità che, su 100 lanci, si colpisca almeno 6 volte il centro.
Indicata con X la variabile casuale 'numero di volte che si colpisce il centro' nelle 100 prove indipendenti risulta

$$X \sim$$

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.



Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.

2. Si calcoli la probabilità che, su 100 lanci, si colpisca almeno 6 volte il centro. Indicata con X la variabile casuale 'numero di volte che si colpisce il centro' nelle 100 prove indipendenti risulta

$$X \sim \text{Bin}(n = 100, p = 0.06).$$

$$M(X) = 6; \text{Var}(X) = 5.64;$$

utilizzando l'approssimazione della variabile casuale Binomiale con la variabile casuale Normale

$$X \sim$$



2. Si calcoli la probabilità che, su 100 lanci, si colpisca almeno 6 volte il centro. Indicata con X la variabile casuale 'numero di volte che si colpisce il centro' nelle 100 prove indipendenti risulta

$$X \sim \text{Bin}(n = 100, p = 0.06).$$

$$M(X) = 6; \text{Var}(X) = 5.64;$$

utilizzando l'approssimazione della variabile casuale Binomiale con la variabile casuale Normale

$$X \sim N(\mu = 6, \sigma^2 = 5.64)$$

si ottiene

$$P(X \geq 6) = P(X \geq \mu) =$$



2. Si calcoli la probabilità che, su 100 lanci, si colpisca almeno 6 volte il centro. Indicata con X la variabile casuale 'numero di volte che si colpisce il centro' nelle 100 prove indipendenti risulta

$$X \sim \text{Bin}(n = 100, p = 0.06).$$

$$M(X) = 6; \text{Var}(X) = 5.64;$$

utilizzando l'approssimazione della variabile casuale Binomiale con la variabile casuale Normale

$$X \sim N(\mu = 6, \sigma^2 = 5.64)$$

si ottiene

$$P(X \geq 6) = P(X \geq \mu) = 0.5.$$

Tema T281

Prova scritta di STATISTICA del 15.01.2015

Prova scritta di
STATISTICA del
15.01.2015

Giuseppe Boari
Gabriele Cantaluppi



UNIVERSITÀ
CATTOLICA
del Sacro Cuore

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.

Giuseppe Boari
Gabriele Cantaluppi
Università Cattolica del Sacro Cuore, Milano



Con riferimento alla distribuzione congiunta delle variabili quantitative discrete X e Y si consideri la seguente tabella a doppia entrata:

$X \setminus Y$	y_1	y_2	y_3
x_1	2	2	2
x_2	0	6	0

1. Dopo avere rappresentato graficamente le distribuzioni **marginali** X e Y , si stabilisca, senza effettuare calcoli e motivando teoricamente la risposta, quale delle due distribuzioni presenta una maggiore eterogeneità.

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.



Si considera la distribuzione marginale della variabile X , ottenuta sommando le frequenze di riga

x_i	$n_{i\bullet}$
x_1	6
x_2	6

Si tratta di un carattere qualitativo sconnesso che può essere rappresentato graficamente mediante un grafico a bastoncini o a canne d'organo.

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.

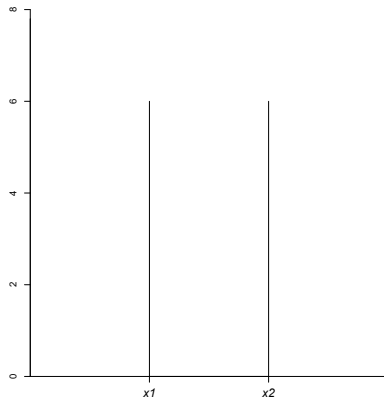


Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.





Si considera la distribuzione marginale della variabile Y , ottenuta sommando le frequenze di colonna

y_j	$n_{\bullet j}$
1	2
2	8
3	2

Esercizio 1.

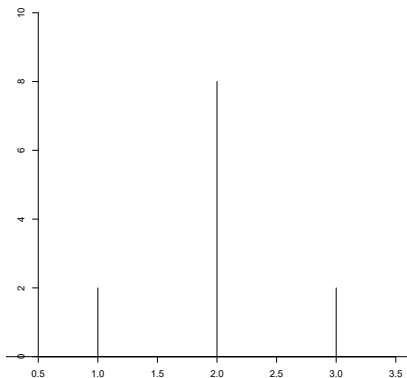
Esercizio 2.

Esercizio 3.

Esercizio 4.



Si tratta di un carattere quantitativo con modalità non raggruppate in classi che può essere rappresentato graficamente mediante un grafico a bastoncini.



Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.



La distribuzione che presenta maggiore eterogeneità è la X in quanto caratterizzata da equidistribuzione delle frequenze (situazione di massima eterogeneità).

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.



Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.

2. Sapendo che $y_1 = 1$, $y_2 = 2$ e $y_3 = 3$, si confrontino le distribuzioni condizionate $Y|x_1$ e $Y|x_2$ con opportuni indici di posizione e variabilità, commentando i risultati ottenuti.



Si considera la distribuzione condizionata $Y|x_1$

y_j	n_{1j}
1	2
2	2
3	2

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.



Si tratta di un carattere quantitativo con modalità non raggruppate in classi.

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.



Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.

Si osserva che la distribuzione è simmetrica rispetto a 2; la mediana (percentile di ordine 0.50) risulta $y_{0.5} = 2$.



La media aritmetica

$$M(Y|x_1) = \frac{1}{n_{1\bullet}} \sum_{j=1}^k y_j n_{1j}$$

risulta 2.

Si osserva, infatti, che la distribuzione è simmetrica.

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.



La varianza

$$\text{Var}(Y|x_1) = M(Y^2|x_1) - [M(Y|x_1)]^2$$

risulta 0.6667.

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.



La varianza

$$\text{Var}(Y|x_1) = M(Y^2|x_1) - [M(Y|x_1)]^2$$

risulta 0.6667. Si ottiene

y_j	n_{1j}	$y_j n_{1j}$	$y_j^2 n_{1j}$
1	2	2	2
2	2	4	8
3	2	6	18
	6	12	28

da cui segue

$$\text{Var}(Y|x_1) = \frac{28}{6} - 2^2 = 4.6667 - 4 = 0.6667.$$

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.



Per effettuare confronti in termini di variabilità tra caratteri quantitativi occorre utilizzare una misura relativa, ad esempio il coefficiente di variazione

$$CV = \frac{\sigma}{\mu} = \frac{\sqrt{\text{Var}(Y|x_1)}}{M(Y|x_1)} = \frac{\sqrt{0.6667}}{2} = \frac{0.8165}{2} = 0.4083.$$

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.



Si considera la distribuzione condizionata $Y|x_2$

y_j	n_{2j}
1	0
2	6
3	0

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.



Si tratta di un carattere quantitativo con modalità non raggruppate in classi.

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.



Si osserva che la distribuzione è degenera.

$M(Y|x_2) = 2$. La varianza e il coefficiente di variazione sono nulli.

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.



Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.

$Y|x_1$: Indice di posizione: $M(Y|x_1) = 2$ Indice di variabilità: $CV(Y|x_1) = 0.4083$

$Y|x_2$: Indice di posizione: $M(Y|x_2) = 2$ Indice di variabilità: $CV(Y|x_2) = 0$

Le distribuzioni sono caratterizzate dallo stesso livello medio; $Y|x_2$ non presenta variabilità.



3. Si determini il grado di connessione tra le variabili X e Y con un opportuno indice normalizzato.

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.



Dalla distribuzione delle frequenze congiunte n_{ij} :

	y_1	y_2	y_3	$n_{i\bullet}$
x_1	2	2	2	6
x_2	0	6	0	6
$n_{\bullet j}$	2	8	2	12

è possibile ottenere le frequenze teoriche nella situazione di indipendenza stocastica

$$\hat{n}_{ij} = \frac{n_{i\bullet} n_{\bullet j}}{n}:$$

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.



Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.

Dalla distribuzione delle frequenze congiunte n_{ij} :

	y_1	y_2	y_3	$n_{i\bullet}$
x_1	2	2	2	6
x_2	0	6	0	6
$n_{\bullet j}$	2	8	2	12

è possibile ottenere le frequenze teoriche nella situazione di indipendenza stocastica

$$\hat{n}_{ij} = \frac{n_{i\bullet} n_{\bullet j}}{n}:$$

	y_1	y_2	y_3	$n_{i\bullet}$
x_1	1	4	1	6
x_2	1	4	1	6
$n_{\bullet j}$	2	8	2	12



le contingenze assolute $n_{ij} - \hat{n}_{ij}$:

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.



le contingenze assolute $n_{ij} - \hat{n}_{ij}$:

	y_1	y_2	y_3
x_1	1	-2	1
x_2	-1	2	-1

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.



i rapporti di contingenza $\frac{n_{ij}}{\hat{n}_{ij}}$:

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.



i rapporti di contingenza $\frac{n_{ij}}{\hat{n}_{ij}}$:

	y_1	y_2	y_3
x_1	2	0.5	2
x_2	0	1.5	0

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.



e le contingenze relative, espresse in termini percentuali, $100 \left(\frac{n_{ij} - \hat{n}_{ij}}{\hat{n}_{ij}} \right) \% = 100 \left(\frac{n_{ij}}{\hat{n}_{ij}} - 1 \right) \%$:

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.



e le contingenze relative, espresse in termini percentuali, $100 \left(\frac{n_{ij} - \hat{n}_{ij}}{\hat{n}_{ij}} \right) \% = 100 \left(\frac{n_{ij}}{\hat{n}_{ij}} - 1 \right) \%$:

	y_1	y_2	y_3
x_1	100	-50	100
x_2	-100	50	-100

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.



Risultati intermedi $\frac{(n_{ij} - \hat{n}_{ij})^2}{\hat{n}_{ij}}$ per il calcolo del χ^2

	y_1	y_2	y_3
x_1	1	1	1
x_2	1	1	1

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.



Risultati intermedi $\frac{n_{ij}^2}{n_{i\bullet} \cdot n_{\bullet j}}$ per il calcolo del χ^2 (formula operativa)

	y_1	y_2	y_3
x_1	0.3333	0.0833	0.3333
x_2	0.0000	0.7500	0.0000

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.



Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.

$$\chi^2 = \sum_{i=1}^h \sum_{j=1}^k \frac{(n_{ij} - \hat{n}_{ij})^2}{\hat{n}_{ij}} = 6,$$

$$\chi_N^2 = \frac{\chi^2}{n \cdot \min(h-1, k-1)} = 0.5$$

$$\chi_N = \sqrt{\chi_N^2} = 0.7071.$$

Sussiste, quindi, un livello di connessione medio alto.



4. Si indichi inoltre se e come potrebbero essere modificate le frequenze congiunte, a parità di numerosità totale, per ottenere una situazione di massima connessione.

$X \setminus Y$	y_1	y_2	y_3	
x_1	2	0	2	4
x_2	0	8	0	8
	2	8	2	12

Nel caso in esame è possibile realizzare una situazione di dipendenza funzionale di X da Y (distribuzioni condizionate $X|y_j$ degeneri).

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.



Il prospetto seguente riporta i prezzi medi annui di 2 beni di uso quotidiano nel periodo 2006 – 2010:

t	$bene A$	$bene B$
2006	100	460
2007	98	460
2008	103	460
2009	105	460
2010	105	480

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.



1. Costruire la serie dei numeri indici a base fissa con base 2007 per il bene A.

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.



1. Costruire la serie dei numeri indici a base fissa con base 2007 per il bene A.

Si ricorda che ${}_{2007}I_t = \frac{x_t}{x_{2007}}$.

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.



1. Costruire la serie dei numeri indici a base fissa con base 2007 per il bene A.
Si ricorda che ${}_{2007}I_t = \frac{x_t}{x_{2007}}$. Si ottiene quindi

Anno	x_t	NIBF (2007)
2006	100	1.0204
2007	98	1.0000
2008	103	1.0510
2009	105	1.0714
2010	105	1.0714

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.



2. Calcolare il tasso medio di variazione per il prezzo del bene B tra il 2006 e il 2009, interpretando il risultato.

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.



2. Calcolare il tasso medio di variazione per il prezzo del bene B tra il 2006 e il 2009, interpretando il risultato.

Si ricorda che

$${}_{2006}I_{2009} = {}_{2006}I_{2007} \cdot {}_{2007}I_{2008} \cdot {}_{2008}I_{2009}$$

Si deve trovare il numero indice a base mobile medio.

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.



Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.

2. Calcolare il tasso medio di variazione per il prezzo del bene B tra il 2006 e il 2009, interpretando il risultato.

Si ricorda che

$${}_{2006}I_{2009} = {}_{2006}I_{2007} \cdot {}_{2007}I_{2008} \cdot {}_{2008}I_{2009}$$

Si deve trovare il numero indice a base mobile medio. Applicando il criterio di scelta della media secondo Chisini si ottiene

$${}_{2006}I_{2009} = {}_{2006}I_{2007} \cdot {}_{2007}I_{2008} \cdot {}_{2008}I_{2009} = \alpha \cdot \alpha \cdot \alpha = \alpha^3.$$



Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.

2. Calcolare il tasso medio di variazione per il prezzo del bene B tra il 2006 e il 2009, interpretando il risultato.

Si ricorda che

$${}_{2006}I_{2009} = {}_{2006}I_{2007} \cdot {}_{2007}I_{2008} \cdot {}_{2008}I_{2009}$$

Si deve trovare il numero indice a base mobile medio. Applicando il criterio di scelta della media secondo Chisini si ottiene

$${}_{2006}I_{2009} = {}_{2006}I_{2007} \cdot {}_{2007}I_{2008} \cdot {}_{2008}I_{2009} = \alpha \cdot \alpha \cdot \alpha = \alpha^3.$$

Quindi

$$\alpha = \sqrt[3]{{}_{2006}I_{2009}} = \sqrt[3]{{}_{2006}I_{2007} \cdot {}_{2007}I_{2008} \cdot {}_{2008}I_{2009}} = 1.$$

Nel periodo considerato la grandezza in oggetto è rimasta immutata



Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.

3. Note le quantità ${}_A q_{2007} = 1000$ e ${}_B q_{2007} = 8500$, calcolare l'indice di Laspeyres dei prezzi al 2010 con base 2007.



L'indice dei prezzi di Laspeyres si ottiene considerando la media aritmetica dei 'prezzi relativi' riferiti ai singoli prodotti $x_i = \frac{i p_{2010}}{i p_{2007}}$ calcolata utilizzando come frequenze i valori $i v_{2007} = i p_{2007} \cdot i q_{2007}$ scambiati al tempo 2007

$${}^p_{2007}I^L_{2010} = \frac{1}{\sum_{i=1}^h i p_{2007} i q_{2007}} \sum_{i=1}^h \frac{i p_{2010}}{i p_{2007}} i p_{2007} i q_{2007}$$

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.



Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.

$\frac{i p_{2010}}{i p_{2007}}$	$i p_{2007} \cdot i q_{2007}$	$i p_{2010} \cdot i q_{2007}$
$\frac{105}{98} = 1.0714$	$98 \cdot 1000 = 98000$	105000
$\frac{480}{460} = 1.0435$	$460 \cdot 8500 = 3910000$	4080000
	4008000	4185000

$$p_{2007}^L /_{2010}^L = \frac{1}{4008000} 4185000 = 1.0442.$$

La variazione relativa nel livello generale dei prezzi, secondo l'Indice di Laspeyres, è del 4.42%.



Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.

Il seguente prospetto riporta i valori rilevati in corrispondenza a 5 unità statistiche con riferimento a due caratteri quantitativi X e Y

x_i	18	18	20	22	22
y_i	17	19	20	21	23

1. Si rappresentino graficamente le coppie dei dati e la funzione di regressione ipotizzando Y in funzione di X .



La funzione di regressione associa a ogni valore x appartenente all'insieme di definizione della variabile X la media di Y condizionata a x , $M(Y|x)$

x	$M(Y x)$
18	$\frac{1}{2}(17 + 19) = 18$
20	20
22	$\frac{1}{2}(21 + 23) = 22$

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.



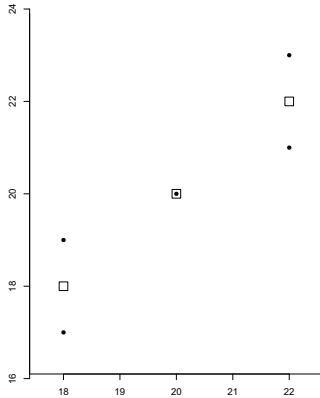
Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.

Le coppie di dati possono essere rappresentate mediante un grafico a dispersione.





2. Si calcolino, secondo il metodo dei minimi quadrati, i parametri e l'indice di adattamento dei seguenti modelli di regressione, individuando il modello migliore:

$$I) Y^* = a + b \cdot X \qquad II) Y^* = c \cdot X$$

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.

Si riportano i calcoli necessari¹ per ottenere le stime dei parametri e gli indici di adattamento del modello

$$Y = a + bX + E$$

x_i	y_i	$x_i y_i$	x_i^2	y_i^2
18	17	306	324	289
18	19	342	324	361
20	20	400	400	400
22	21	462	484	441
22	23	506	484	529
100	100	2016	2016	2020
20	20	403.2	403.2	404
$M(X)$	$M(Y)$	$M(XY)$	$M(X^2)$	$M(Y^2)$

$$\hat{b} =$$



Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.

¹Si osserva che per ragioni di stampa alcuni calcoli potrebbero essere duplicati con riferimento ai diversi modelli oggetto di studio

Si riportano i calcoli necessari¹ per ottenere le stime dei parametri e gli indici di adattamento del modello

$$Y = a + bX + E$$

x_i	y_i	$x_i y_i$	x_i^2	y_i^2
18	17	306	324	289
18	19	342	324	361
20	20	400	400	400
22	21	462	484	441
22	23	506	484	529
100	100	2016	2016	2020
20	20	403.2	403.2	404
$M(X)$	$M(Y)$	$M(XY)$	$M(X^2)$	$M(Y^2)$

$$\begin{aligned}\hat{b} &= \frac{\text{Cov}(X, Y)}{\text{Var}(X)} = \frac{M(XY) - M(X)M(Y)}{M(X^2) - [M(X)^2]} = \\ &= \frac{403.2 - 20 \cdot 20}{403.2 - 20^2} = \frac{3.2}{3.2} = 1 \\ \hat{a} &= \end{aligned}$$

¹Si osserva che per ragioni di stampa alcuni calcoli potrebbero essere duplicati con riferimento ai diversi modelli oggetto di studio



Si riportano i calcoli necessari¹ per ottenere le stime dei parametri e gli indici di adattamento del modello

$$Y = a + bX + E$$

x_i	y_i	$x_i y_i$	x_i^2	y_i^2
18	17	306	324	289
18	19	342	324	361
20	20	400	400	400
22	21	462	484	441
22	23	506	484	529
100	100	2016	2016	2020
20	20	403.2	403.2	404
$M(X)$	$M(Y)$	$M(XY)$	$M(X^2)$	$M(Y^2)$

$$\hat{b} = \frac{\text{Cov}(X, Y)}{\text{Var}(X)} = \frac{M(XY) - M(X)M(Y)}{M(X^2) - [M(X)]^2} =$$

$$= \frac{403.2 - 20 \cdot 20}{403.2 - 20^2} = \frac{3.2}{3.2} = 1$$

$$\hat{a} = M(Y) - \hat{b}M(X) = 0$$

$$\rho^2 =$$



Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.

¹Si osserva che per ragioni di stampa alcuni calcoli potrebbero essere duplicati con riferimento ai diversi modelli oggetto di studio

Si riportano i calcoli necessari¹ per ottenere le stime dei parametri e gli indici di adattamento del modello

$$Y = a + bX + E$$

x_i	y_i	$x_i y_i$	x_i^2	y_i^2
18	17	306	324	289
18	19	342	324	361
20	20	400	400	400
22	21	462	484	441
22	23	506	484	529
100	100	2016	2016	2020
20	20	403.2	403.2	404
$M(X)$	$M(Y)$	$M(XY)$	$M(X^2)$	$M(Y^2)$

$$\hat{b} = \frac{\text{Cov}(X, Y)}{\text{Var}(X)} = \frac{M(XY) - M(X)M(Y)}{M(X^2) - [M(X)]^2} =$$

$$= \frac{403.2 - 20 \cdot 20}{403.2 - 20^2} = \frac{3.2}{3.2} = 1$$

$$\hat{a} = M(Y) - \hat{b}M(X) = 0$$

$$\rho^2 = \frac{[\text{Cov}(X, Y)]^2}{\text{Var}(X) \text{Var}(Y)} = \frac{3.2^2}{3.2 \cdot 4} = 0.8$$

$$M(\hat{E}^2) = \text{Var}(\hat{E}) = \text{varianza residua} =$$

¹Si osserva che per ragioni di stampa alcuni calcoli potrebbero essere duplicati con riferimento ai diversi modelli oggetto di studio



Si riportano i calcoli necessari¹ per ottenere le stime dei parametri e gli indici di adattamento del modello

$$Y = a + bX + E$$

x_i	y_i	$x_i y_i$	x_i^2	y_i^2
18	17	306	324	289
18	19	342	324	361
20	20	400	400	400
22	21	462	484	441
22	23	506	484	529
100	100	2016	2016	2020
20	20	403.2	403.2	404
$M(X)$	$M(Y)$	$M(XY)$	$M(X^2)$	$M(Y^2)$

$$\hat{b} = \frac{\text{Cov}(X, Y)}{\text{Var}(X)} = \frac{M(XY) - M(X)M(Y)}{M(X^2) - [M(X)]^2} =$$

$$= \frac{403.2 - 20 \cdot 20}{403.2 - 20^2} = \frac{3.2}{3.2} = 1$$

$$\hat{a} = M(Y) - \hat{b}M(X) = 0$$

$$\rho^2 = \frac{[\text{Cov}(X, Y)]^2}{\text{Var}(X) \text{Var}(Y)} = \frac{3.2^2}{3.2 \cdot 4} = 0.8$$

$$M(\hat{E}^2) = \text{Var}(\hat{E}) = \text{varianza residua} = \text{Var}(Y) \cdot (1 - \rho^2) = 0.8$$

¹Si osserva che per ragioni di stampa alcuni calcoli potrebbero essere duplicati con riferimento ai diversi modelli oggetto di studio





Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.

Si riportano i calcoli necessari² per ottenere le stime dei parametri e la misura di adattamento del modello

$$Y = cX + E$$

x_i	y_i	$x_i y_i$	x_i^2	y_i^2
18	17	306	324	289
18	19	342	324	361
20	20	400	400	400
22	21	462	484	441
22	23	506	484	529
		2016	2016	2020
		403.2	403.2	404
		$M(XY)$	$M(X^2)$	$M(Y^2)$

$$\hat{c} =$$

²Si osserva che per ragioni di stampa alcuni calcoli potrebbero essere duplicati con riferimento ai diversi modelli oggetto di studio



Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.

Si riportano i calcoli necessari² per ottenere le stime dei parametri e la misura di adattamento del modello

$$Y = cX + E$$

x_i	y_i	$x_i y_i$	x_i^2	y_i^2
18	17	306	324	289
18	19	342	324	361
20	20	400	400	400
22	21	462	484	441
22	23	506	484	529
		2016	2016	2020
		403.2	403.2	404
		$M(XY)$	$M(X^2)$	$M(Y^2)$

$$\hat{c} = \frac{M(XY)}{M(X^2)} = \frac{403.2}{403.2} = 1$$

$$M(\hat{E}^2) =$$

²Si osserva che per ragioni di stampa alcuni calcoli potrebbero essere duplicati con riferimento ai diversi modelli oggetto di studio



Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.

Si riportano i calcoli necessari² per ottenere le stime dei parametri e la misura di adattamento del modello

$$Y = cX + E$$

x_i	y_i	$x_i y_i$	x_i^2	y_i^2
18	17	306	324	289
18	19	342	324	361
20	20	400	400	400
22	21	462	484	441
22	23	506	484	529
		2016	2016	2020
		403.2	403.2	404
		$M(XY)$	$M(X^2)$	$M(Y^2)$

$$\hat{c} = \frac{M(XY)}{M(X^2)} = \frac{403.2}{403.2} = 1$$

$$M(\hat{E}^2) = M[Y \cdot (Y - \hat{c}X)] = M(Y^2) - \hat{c}M(XY) = 404 - 1 \cdot 403.2 = 0.8$$

²Si osserva che per ragioni di stampa alcuni calcoli potrebbero essere duplicati con riferimento ai diversi modelli oggetto di studio



Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.

I due modelli sono equivalenti.
Abbiamo

$$Y = X + E$$



3. Senza effettuare calcoli si individui il valore di $\eta_{Y|X}^2$. Si motivi perché $\eta_{Y|X}^2 < \eta_{X|Y}^2$.

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.



3. Senza effettuare calcoli si individui il valore di $\eta_{Y|X}^2$. Si motivi perché $\eta_{Y|X}^2 < \eta_{X|Y}^2$.
 $\eta_{Y|X}^2 = \rho^2$ Dalla rappresentazione grafica si osserva che le medie condizionate sono allineate.

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.



Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.

3. Senza effettuare calcoli si individui il valore di $\eta_{Y|X}^2$. Si motivi perché $\eta_{Y|X}^2 < \eta_{X|Y}^2$.
 $\eta_{Y|X}^2 = \rho^2$ Dalla rappresentazione grafica si osserva che le medie condizionate sono allineate.
 $\eta_{X|Y}^2 = 1$ in quanto sussiste dipendenza funzionale di X da Y .



Un'urna è composta da 26 gettoni rossi e 74 gettoni neri. Si consideri l'esperimento casuale $\mathcal{G}$ = 'estrazione **senza reimmissione** di 2 gettoni dall'urna, dei quali viene rilevato il colore'.
1. Si costruisca lo spazio probabilistico associato all'esperimento $\mathcal{G}$ e si calcoli la probabilità di ciascun evento elementare.

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.

Esercizio 4

Prova scritta di
STATISTICA del
15.01.2015

Giuseppe Boari
Gabriele Cantaluppi



UNIVERSITÀ
CATTOLICA
del Sacro Cuore

Un'urna è composta da 26 gettoni rossi e 74 gettoni neri. Si consideri l'esperimento casuale $\mathcal{G}$ = 'estrazione **senza reimmissione** di 2 gettoni dall'urna, dei quali viene rilevato il colore'.

1. Si costruisca lo spazio probabilistico associato all'esperimento $\mathcal{G}$ e si calcoli la probabilità di ciascun evento elementare.

RR
RN
NR
NN

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.



Le probabilità degli eventi elementari risultano

$$P(R_1 \cap R_2) = P(R_1)P(R_2|R_1) = \frac{26}{100} \frac{25}{99} = 0.0657,$$

$$P(R_1 \cap N_2) = P(R_1)P(N_2|R_1) = \frac{26}{100} \frac{74}{99} = 0.1943,$$

$$P(N_1 \cap R_2) = P(N_1)P(R_2|N_1) = \frac{74}{100} \frac{26}{99} = 0.1943,$$

$$P(N_1 \cap N_2) = P(N_1)P(N_2|N_1) = \frac{74}{100} \frac{73}{99} = 0.5457.$$

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.



2. Si calcoli la probabilità che nelle due estrazioni si ottengano due gettoni del medesimo colore.

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.



2. Si calcoli la probabilità che nelle due estrazioni si ottengano due gettoni del medesimo colore.

$$P(RR \cup NN) = P(RR) + P(NN) = 0.0657 + 0.5457 = 0.6113.$$

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.



3. Un giocatore puntando 1 Euro, ne vince 3 qualora i due gettoni estratti siano del medesimo colore. In caso contrario perde la somma puntata. Si stabilisca se il gioco è equo (guadagno atteso nullo).

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.



Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.

3. Un giocatore puntando 1 Euro, ne vince 3 qualora i due gettoni estratti siano del medesimo colore. In caso contrario perde la somma puntata. Si stabilisca se il gioco è equo (guadagno atteso nullo).

Ricordando che la probabilità di vincita risulta

$$P(\text{vincita}) = 0.6113,$$

è possibile costruire la variabile casuale guadagno



3. Un giocatore puntando 1 Euro, ne vince 3 qualora i due gettoni estratti siano del medesimo colore. In caso contrario perde la somma puntata. Si stabilisca se il gioco è equo (guadagno atteso nullo).

Ricordando che la probabilità di vincita risulta

$$P(\text{vincita}) = 0.6113,$$

è possibile costruire la variabile casuale guadagno

	x_i	$P(X = x_i)$	$x_i \cdot P(X = x_i)$
<i>posta</i>	-1	0.3887	-0.3887
<i>vincita - posta</i>	2	0.6113	1.2226
		1	0.8339



3. Un giocatore puntando 1 Euro, ne vince 3 qualora i due gettoni estratti siano del medesimo colore. In caso contrario perde la somma puntata. Si stabilisca se il gioco è equo (guadagno atteso nullo).

Ricordando che la probabilità di vincita risulta

$$P(\text{vincita}) = 0.6113,$$

è possibile costruire la variabile casuale guadagno

	x_i	$P(X = x_i)$	$x_i \cdot P(X = x_i)$
<i>posta</i>	-1	0.3887	-0.3887
<i>vincita - posta</i>	2	0.6113	1.2226
		1	0.8339

Il valore atteso della variabile casuale guadagno è positivo, quindi il gioco non è equo ma favorevole al giocatore.



4. Si calcoli la probabilità di ottenere non più di 10 gettoni rossi in 51 estrazioni **con reimmissione**.

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.



4. Si calcoli la probabilità di ottenere non più di 10 gettoni rossi in 51 estrazioni **con reimmissione**.

Indicata con X la variabile casuale 'numero di gettoni rossi' nelle 51 estrazioni con reimmissione (prove indipendenti) risulta

$$X \sim$$

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.



4. Si calcoli la probabilità di ottenere non più di 10 gettoni rossi in 51 estrazioni **con reimmissione**.

Indicata con X la variabile casuale 'numero di gettoni rossi' nelle 51 estrazioni con reimmissione (prove indipendenti) risulta

$$X \sim \text{Bin}(n = 51, p = 0.26).$$

$$M(X) = 13.26; \text{Var}(X) = 9.8124;$$

utilizzando l'approssimazione della variabile casuale Binomiale con la variabile casuale Normale

$$X \sim$$



4. Si calcoli la probabilità di ottenere non più di 10 gettoni rossi in 51 estrazioni **con reimmissione**.

Indicata con X la variabile casuale 'numero di gettoni rossi' nelle 51 estrazioni con reimmissione (prove indipendenti) risulta

$$X \sim \text{Bin}(n = 51, p = 0.26).$$

$$M(X) = 13.26; \text{Var}(X) = 9.8124;$$

utilizzando l'approssimazione della variabile casuale Binomiale con la variabile casuale Normale

$$X \sim N(\mu = 13.26, \sigma^2 = 9.8124)$$

si ottiene

$$\begin{aligned} P(X \leq 10) &= P\left(\frac{X - \mu}{\sigma} \leq \frac{10 - \mu}{\sigma}\right) = P\left(Z \leq \frac{10 - 13.26}{3.1325}\right) = \\ &= P(Z \leq -1.04) = P(Z \geq 1.04) = \\ &= 1 - P(Z < 1.04) = 1 - 0.8508 = 0.1492. \end{aligned}$$

Tema T282

Prova scritta di STATISTICA del 29.01.2015

Prova scritta di
STATISTICA del
29.01.2015

Giuseppe Boari
Gabriele Cantaluppi



UNIVERSITÀ
CATTOLICA
del Sacro Cuore

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.

Giuseppe Boari
Gabriele Cantaluppi
Università Cattolica del Sacro Cuore, Milano



La tabella seguente contiene la distribuzione delle scadenze in anni (A) dei 108 titoli di stato detenuti in portafoglio da 2 banche italiane (B):

$B \setminus A$	0 + 1	1 + 5	5 + 15	
$b_1 = \text{Intesa Sanpaolo}$	21	18	19	
$b_2 = \text{Unicredit}$	28	13	9	

1. Fornire la rappresentazione grafica della distribuzione $A|Unicredit$.

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.



Si considera la distribuzione condizionata $A|b_2$

a_j	n_{2j}
0 + 1	28
1 + 5	13
5 + 15	9

Esercizio 1.

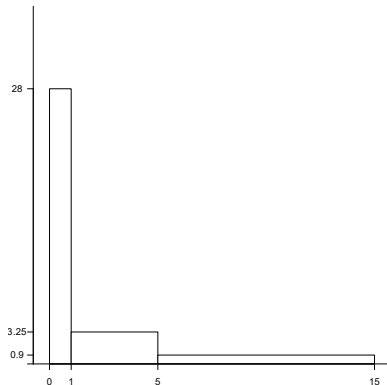
Esercizio 2.

Esercizio 3.

Esercizio 4.

Si tratta di un carattere quantitativo con modalità raggruppate in classi che può essere rappresentato graficamente mediante un istogramma. L'istogramma può essere costruito dopo avere calcolato le densità $d_{2j} = \frac{n_{2j}}{a_j}$, (dove a_j sono le ampiezze di classe).

a_j	n_{2j}	a_{2j}	d_{2j}
0 - 1	28	1 - 0 = 1	28
1 - 5	13	5 - 1 = 4	3.25
5 - 15	9	15 - 5 = 10	0.9



Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.



2. Confrontare opportunamente la variabilità delle due distribuzioni $A|b_i$ ($i = 1, 2$).

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.



Si considera la distribuzione condizionata $A|b_1$

a_j	n_{1j}
0 + 1	21
1 + 5	18
5 + 15	19

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.



Si tratta di un carattere quantitativo con modalità raggruppate in classi.

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.



La media aritmetica

$$M(A|b_1) = \frac{1}{n_{1\bullet}} \sum_{j=1}^k a_j n_{1j}$$

risulta 4.3879.

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.



La media aritmetica

$$M(A|b_1) = \frac{1}{n_{1\bullet}} \sum_{j=1}^k a_j n_{1j}$$

risulta 4.3879.

Si ottiene (considerando come modalità i valori centrali delle classi)

a_j	n_{1j}	$a_j n_{1j}$
0.5	21	10.5
3.0	18	54.0
10.0	19	190.0
	58	254.5

da cui segue

$$M(A|b_1) = \frac{1}{n_{1\bullet}} \sum_{j=1}^k a_j n_{1j} = \frac{254.5}{58} = 4.3879.$$

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.



La varianza

$$\text{Var}(A|b_1) = M(A^2|b_1) - [M(A|b_1)]^2$$

risulta 16.3883.

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.



La varianza

$$Var(A|b_1) = M(A^2|b_1) - [M(A|b_1)]^2$$

risulta 16.3883. Si ottiene (considerando come modalità i valori centrali delle classi)

a_j	n_{1j}	$a_j n_{1j}$	$a_j^2 n_{1j}$
0.5	21	10.5	5.25
3.0	18	54.0	162.00
10.0	19	190.0	1900.00
	58	254.5	2067.25

da cui segue

$$Var(A|b_1) = \frac{2067.25}{58} - 4.3879^2 = 35.6422 - 19.2539 = 16.3883.$$

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.



Per effettuare confronti in termini di variabilità tra caratteri quantitativi occorre utilizzare una misura relativa, ad esempio il coefficiente di variazione

$$CV = \frac{\sigma}{\mu} = \frac{\sqrt{\text{Var}(A|b_1)}}{M(A|b_1)} = \frac{\sqrt{16.3883}}{4.3879} = \frac{4.0482}{4.3879} = 0.9226.$$

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.



Si considera la distribuzione condizionata $A|b_2$

a_j	n_{2j}
0 + 1	28
1 + 5	13
5 + 15	9

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.



Si tratta di un carattere quantitativo con modalità raggruppate in classi.

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.



La media aritmetica

$$M(A|b_2) = \frac{1}{n_{2\bullet}} \sum_{j=1}^k a_j n_{2j}$$

risulta 2.86.

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.



La media aritmetica

$$M(A|b_2) = \frac{1}{n_{2\bullet}} \sum_{j=1}^k a_j n_{2j}$$

risulta 2.86.

Si ottiene (considerando come modalità i valori centrali delle classi)

a_j	n_{2j}	$a_j n_{2j}$
0.5	28	14
3.0	13	39
10.0	9	90
	50	143

da cui segue

$$M(A|b_2) = \frac{1}{n_{2\bullet}} \sum_{j=1}^k a_j n_{2j} = \frac{143}{50} = 2.86.$$

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.



La varianza

$$\text{Var}(A|b_2) = M(A^2|b_2) - [M(A|b_2)]^2$$

risulta 12.3004.

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.



Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.

La varianza

$$\text{Var}(A|b_2) = M(A^2|b_2) - [M(A|b_2)]^2$$

risulta 12.3004. Si ottiene (considerando come modalità i valori centrali delle classi)

a_j	n_{2j}	$a_j n_{2j}$	$a_j^2 n_{2j}$
0.5	28	14	7
3.0	13	39	117
10.0	9	90	900
	50	143	1024

da cui segue

$$\text{Var}(A|b_2) = \frac{1024}{50} - 2.86^2 = 20.48 - 8.1796 = 12.3004.$$



Per effettuare confronti in termini di variabilità tra caratteri quantitativi occorre utilizzare una misura relativa, ad esempio il coefficiente di variazione

$$CV = \frac{\sigma}{\mu} = \frac{\sqrt{\text{Var}(A|b_2)}}{M(A|b_2)} = \frac{\sqrt{12.3004}}{2.86} = \frac{3.5072}{2.86} = 1.2263.$$

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.



Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.

$A|b_1$: Indice di posizione: $M(A|b_1) = 4.3879$ Indice di variabilità: $CV(A|b_1) = 0.9226$

$A|b_2$: Indice di posizione: $M(A|b_2) = 2.86$ Indice di variabilità: $CV(A|b_1) = 1.2263$

La scadenza media dei titoli posseduti da *Intesa Sanpaolo* è superiore a quella dei titoli posseduti da *Unicredit*. La distribuzione delle scadenze dei titoli posseduti da *Unicredit* presenta un livello di variabilità più elevato.



Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.

3. Misurare la dipendenza di A da B con un opportuno indice di connessione.



Dalla distribuzione delle frequenze congiunte n_{ij} :

	a_1	a_2	a_3	$n_{i\bullet}$
b_1	21	18	19	58
b_2	28	13	9	50
$n_{\bullet j}$	49	31	28	108

è possibile ottenere le frequenze teoriche nella situazione di indipendenza stocastica

$$\hat{n}_{ij} = \frac{n_{i\bullet} n_{\bullet j}}{n}:$$

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.



Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.

Dalla distribuzione delle frequenze congiunte n_{ij} :

	a_1	a_2	a_3	$n_{i\bullet}$
b_1	21	18	19	58
b_2	28	13	9	50
$n_{\bullet j}$	49	31	28	108

è possibile ottenere le frequenze teoriche nella situazione di indipendenza stocastica

$$\hat{n}_{ij} = \frac{n_{i\bullet} n_{\bullet j}}{n}:$$

	a_1	a_2	a_3	$n_{i\bullet}$
b_1	26.3148	16.6481	15.037	58
b_2	22.6852	14.3519	12.963	50
$n_{\bullet j}$	49.0000	31.0000	28.000	108



le contingenze assolute $n_{ij} - \hat{n}_{ij}$:

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.



le contingenze assolute $n_{ij} - \hat{n}_{ij}$:

	a_1	a_2	a_3
b_1	-5.3148	1.3519	3.963
b_2	5.3148	-1.3519	-3.963

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.



i rapporti di contingenza $\frac{n_{ij}}{\hat{n}_{ij}}$:

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.



i rapporti di contingenza $\frac{n_{ij}}{\hat{n}_{ij}}$:

	a_1	a_2	a_3
b_1	0.7980	1.0812	1.2635
b_2	1.2343	0.9058	0.6943

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.



e le contingenze relative, espresse in termini percentuali, $100 \left(\frac{n_{ij} - \hat{n}_{ij}}{\hat{n}_{ij}} \right) \% = 100 \left(\frac{n_{ij}}{\hat{n}_{ij}} - 1 \right) \%$:

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.



e le contingenze relative, espresse in termini percentuali, $100 \left(\frac{n_{ij} - \hat{n}_{ij}}{\hat{n}_{ij}} \right) \% = 100 \left(\frac{n_{ij}}{\hat{n}_{ij}} - 1 \right) \%$:

	a_1	a_2	a_3
b_1	-20.20	8.12	26.35
b_2	23.43	-9.42	-30.57

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.



Risultati intermedi $\frac{(n_{ij} - \hat{n}_{ij})^2}{\hat{n}_{ij}}$ per il calcolo del χ^2

	a_1	a_2	a_3
b_1	1.0734	0.1098	1.0444
b_2	1.2452	0.1273	1.2115

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.



Risultati intermedi $\frac{n_{ij}^2}{n_{i\bullet} \cdot n_{\bullet j}}$ per il calcolo del χ^2 (formula operativa)

	a_1	a_2	a_3
b_1	0.1552	0.1802	0.2223
b_2	0.3200	0.1090	0.0579

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.



Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.

$$\chi^2 = \sum_{i=1}^h \sum_{j=1}^k \frac{(n_{ij} - \hat{n}_{ij})^2}{\hat{n}_{ij}} = 4.8117,$$

$$\chi_N^2 = \frac{\chi^2}{n \cdot \min(h-1, k-1)} = 0.0446$$

$$\chi_N = \sqrt{\chi_N^2} = 0.2112.$$

Sussiste, quindi, un livello di connessione molto basso.



4. Indicare le condizioni nelle quali $\eta_{A|B}^2$ e χ^2 risultano nulli.

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.



4. Indicare le condizioni nelle quali $\eta_{A|B}^2$ e χ^2 risultano nulli.

$\eta_{A|B}^2$ risulta nullo quando sussiste indipendenza in media di A da B , vale a dire le medie condizionate coincidono

$$M(A|b_1) = M(A|b_2).$$

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.



Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.

4. Indicare le condizioni nelle quali $\eta_{A|B}^2$ e χ^2 risultano nulli.

$\eta_{A|B}^2$ risulta nullo quando sussiste indipendenza in media di A da B , vale a dire le medie condizionate coincidono

$$M(A|b_1) = M(A|b_2).$$

χ^2 risulta nullo quando sussiste indipendenza stocastica tra A e B , vale a dire le distribuzioni condizionate sono somiglianti:

$$A|b_1 \sim A|b_2 \sim A$$

e

$$B|A_1 \sim B|a_2 \sim B|a_3 \sim B$$

.



Si consideri la variabile X = vendite annuali di un prodotto, in migliaia di pezzi, misurata tra il 2008 e il 2012:

t	x_t	$NIBM$
2008	888	
2009	619	
2010	568	
2011	509	
2012	528	

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.



1. Costruire la serie dei numeri indici a base mobile.

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.



1. Costruire la serie dei numeri indici a base mobile.

Si ricorda che ${}_{t-1}I_t = \frac{x_t}{x_{t-1}}$.

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.



1. Costruire la serie dei numeri indici a base mobile.

Si ricorda che ${}_{t-1}I_t = \frac{x_t}{x_{t-1}}$. Si ottiene quindi

Anno	x_t	NIBM
2008	888	
2009	619	0.6971
2010	568	0.9176
2011	509	0.8961
2012	528	1.0373

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.



Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.

2. Interpretare il valore dell'ultimo degli indici sopra calcolati.

${}_{2011}I_{2012} = 1.0373$ Tra il 2011 e il 2012 si è registrato un incremento delle vendite del 3.73%.



3. Calcolare il tasso medio di variazione delle vendite registrato tra il 2008 e il 2012, interpretando il risultato.

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.



3. Calcolare il tasso medio di variazione delle vendite registrato tra il 2008 e il 2012, interpretando il risultato.

Si ricorda che

$${}_{2008}I_{2012} = {}_{2008}I_{2009} \cdot {}_{2009}I_{2010} \cdot {}_{2010}I_{2011} \cdot {}_{2011}I_{2012}$$

Si deve trovare il numero indice a base mobile medio.

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.



Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.

3. Calcolare il tasso medio di variazione delle vendite registrato tra il 2008 e il 2012, interpretando il risultato.

Si ricorda che

$${}_{2008}I_{2012} = {}_{2008}I_{2009} \cdot {}_{2009}I_{2010} \cdot {}_{2010}I_{2011} \cdot {}_{2011}I_{2012}$$

Si deve trovare il numero indice a base mobile medio. Applicando il criterio di scelta della media secondo Chisini si ottiene

$${}_{2008}I_{2012} = {}_{2008}I_{2009} \cdot {}_{2009}I_{2010} \cdot {}_{2010}I_{2011} \cdot {}_{2011}I_{2012} = \alpha \cdot \alpha \cdot \alpha \cdot \alpha = \alpha^4.$$



Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.

3. Calcolare il tasso medio di variazione delle vendite registrato tra il 2008 e il 2012, interpretando il risultato.

Si ricorda che

$${}_{2008}I_{2012} = {}_{2008}I_{2009} \cdot {}_{2009}I_{2010} \cdot {}_{2010}I_{2011} \cdot {}_{2011}I_{2012}$$

Si deve trovare il numero indice a base mobile medio. Applicando il criterio di scelta della media secondo Chisini si ottiene

$${}_{2008}I_{2012} = {}_{2008}I_{2009} \cdot {}_{2009}I_{2010} \cdot {}_{2010}I_{2011} \cdot {}_{2011}I_{2012} = \alpha \cdot \alpha \cdot \alpha \cdot \alpha = \alpha^4.$$

Quindi

$$\alpha = \sqrt[4]{{}_{2008}I_{2012}} = \sqrt[4]{{}_{2008}I_{2009} \cdot {}_{2009}I_{2010} \cdot {}_{2010}I_{2011} \cdot {}_{2011}I_{2012}} = 0.8781.$$

Da cui segue una variazione relativa media pari a -0.1219 , corrispondente a un decremento medio annuo tra il 2008 e il 2012 del 12.19%.



Con riferimento allo studio della dipendenza tra due caratteri quantitativi X e Y , si sono raccolte le seguenti 7 coppie di osservazioni (x_i, y_i) $i = 1, 2, \dots, 7$

x_i	1	1	2	2	5	5	5
y_i	4	5	4	5	11	12	13

1. Fornire la rappresentazione grafica delle coppie di punti, insieme alla funzione di regressione.

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.



La funzione di regressione associa a ogni valore x appartenente all'insieme di definizione della variabile X la media di Y condizionata a x , $M(Y|x)$

x	$M(Y x)$
1	$\frac{1}{2}(4 + 5) = 4.5$
2	$\frac{1}{2}(4 + 5) = 4.5$
5	$\frac{1}{3}(11 + 12 + 13) = 12$

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.



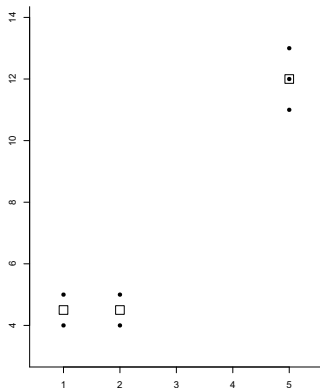
Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.

Le coppie di dati possono essere rappresentate mediante un grafico a dispersione.





2. Si calcolino, secondo il metodo dei minimi quadrati, i parametri e l'indice di adattamento dei seguenti modelli di regressione, individuando il modello migliore:

$$I) Y^* = a + b \cdot X \quad II) Y^* = \alpha + \beta \cdot X^2$$

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.



Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.

Si riportano i calcoli necessari¹ per ottenere le stime dei parametri e gli indici di adattamento del modello

$$Y = a + bX + E$$

x_i	y_i	$x_i y_i$	x_i^2	y_i^2
1	4	4	1	16
1	5	5	1	25
2	4	8	4	16
2	5	10	4	25
5	11	55	25	121
5	12	60	25	144
5	13	65	25	169
21	54	207	85	516
3	7.7143	29.5714	12.1429	73.7143
$M(X)$	$M(Y)$	$M(XY)$	$M(X^2)$	$M(Y^2)$

$$\hat{b} =$$

¹Si osserva che per ragioni di stampa alcuni calcoli potrebbero essere duplicati con riferimento ai diversi modelli oggetto di studio



Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.

Si riportano i calcoli necessari¹ per ottenere le stime dei parametri e gli indici di adattamento del modello

$$Y = a + bX + E$$

x_i	y_i	$x_i y_i$	x_i^2	y_i^2
1	4	4	1	16
1	5	5	1	25
2	4	8	4	16
2	5	10	4	25
5	11	55	25	121
5	12	60	25	144
5	13	65	25	169
21	54	207	85	516
3	7.7143	29.5714	12.1429	73.7143
$M(X)$	$M(Y)$	$M(XY)$	$M(X^2)$	$M(Y^2)$

$$\begin{aligned}\hat{b} &= \frac{\text{Cov}(X, Y)}{\text{Var}(X)} = \frac{M(XY) - M(X)M(Y)}{M(X^2) - [M(X)]^2} = \\ &= \frac{29.5714 - 3 \cdot 7.7143}{12.1429 - 3^2} = \frac{6.4286}{3.1429} = 2.0454 \\ \hat{a} &= \end{aligned}$$

¹Si osserva che per ragioni di stampa alcuni calcoli potrebbero essere duplicati con riferimento ai diversi modelli oggetto di studio



Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.

Si riportano i calcoli necessari¹ per ottenere le stime dei parametri e gli indici di adattamento del modello

$$Y = a + bX + E$$

x_i	y_i	$x_i y_i$	x_i^2	y_i^2
1	4	4	1	16
1	5	5	1	25
2	4	8	4	16
2	5	10	4	25
5	11	55	25	121
5	12	60	25	144
5	13	65	25	169
21	54	207	85	516
3	7.7143	29.5714	12.1429	73.7143
$M(X)$	$M(Y)$	$M(XY)$	$M(X^2)$	$M(Y^2)$

$$\begin{aligned}\hat{b} &= \frac{\text{Cov}(X, Y)}{\text{Var}(X)} = \frac{M(XY) - M(X)M(Y)}{M(X^2) - [M(X)]^2} = \\ &= \frac{29.5714 - 3 \cdot 7.7143}{12.1429 - 3^2} = \frac{6.4286}{3.1429} = 2.0454 \\ \hat{a} &= M(Y) - \hat{b}M(X) = 1.578 \\ \rho^2 &= \end{aligned}$$

¹Si osserva che per ragioni di stampa alcuni calcoli potrebbero essere duplicati con riferimento ai diversi modelli oggetto di studio



Si riportano i calcoli necessari¹ per ottenere le stime dei parametri e gli indici di adattamento del modello

$$Y = a + bX + E$$

x_i	y_i	$x_i y_i$	x_i^2	y_i^2
1	4	4	1	16
1	5	5	1	25
2	4	8	4	16
2	5	10	4	25
5	11	55	25	121
5	12	60	25	144
5	13	65	25	169
21	54	207	85	516
3	7.7143	29.5714	12.1429	73.7143
$M(X)$	$M(Y)$	$M(XY)$	$M(X^2)$	$M(Y^2)$

$$\begin{aligned}\hat{b} &= \frac{\text{Cov}(X, Y)}{\text{Var}(X)} = \frac{M(XY) - M(X)M(Y)}{M(X^2) - [M(X)]^2} = \\ &= \frac{29.5714 - 3 \cdot 7.7143}{12.1429 - 3^2} = \frac{6.4286}{3.1429} = 2.0454\end{aligned}$$

$$\hat{a} = M(Y) - \hat{b}M(X) = 1.578$$

$$\rho^2 = \frac{[\text{Cov}(X, Y)]^2}{\text{Var}(X) \text{Var}(Y)} = \frac{6.4286^2}{3.1429 \cdot 14.2041} = 0.9257$$

¹Si osserva che per ragioni di stampa alcuni calcoli potrebbero essere duplicati con riferimento ai diversi modelli oggetto di studio

Si riportano i calcoli necessari² per ottenere le stime dei parametri e gli indici di adattamento del modello

$$Y = \alpha + \beta X^2 + E$$

x_i^2	y_i	$x_i^2 y_i$	$x_i^{2^2}$	y_i^2
1	4	4	1	16
1	5	5	1	25
4	4	16	16	16
4	5	20	16	25
25	11	275	625	121
25	12	300	625	144
25	13	325	625	169
85	54	945	1909	516
12.1429	7.7143	135	272.7143	73.7143
$M(X^2)$	$M(Y)$	$M(X^2 Y)$	$M(X^{2^2})$	$M(Y^2)$

$$\hat{\beta} =$$



Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.

²Si osserva che per ragioni di stampa alcuni calcoli potrebbero essere duplicati con riferimento ai diversi modelli oggetto di studio

Si riportano i calcoli necessari² per ottenere le stime dei parametri e gli indici di adattamento del modello

$$Y = \alpha + \beta X^2 + E$$

x_i^2	y_i	$x_i^2 y_i$	$x_i^{2^2}$	y_i^2
1	4	4	1	16
1	5	5	1	25
4	4	16	16	16
4	5	20	16	25
25	11	275	625	121
25	12	300	625	144
25	13	325	625	169
85	54	945	1909	516
12.1429	7.7143	135	272.7143	73.7143
$M(X^2)$	$M(Y)$	$M(X^2 Y)$	$M(X^{2^2})$	$M(Y^2)$

$$\begin{aligned}\hat{\beta} &= \frac{\text{Cov}(X^2, Y)}{\text{Var}(X^2)} = \frac{M(X^2 Y) - M(X^2) M(Y)}{M(X^{2^2}) - [M(X^2)]^2} = \\ &= \frac{135 - 12.1429 \cdot 7.7143}{272.7143 - 12.1429^2} = \frac{41.3265}{125.2653} = 0.3299 \\ \hat{\alpha} &= \end{aligned}$$

²Si osserva che per ragioni di stampa alcuni calcoli potrebbero essere duplicati con riferimento ai diversi modelli oggetto di studio



Si riportano i calcoli necessari² per ottenere le stime dei parametri e gli indici di adattamento del modello

$$Y = \alpha + \beta X^2 + E$$

x_i^2	y_i	$x_i^2 y_i$	$x_i^{2^2}$	y_i^2
1	4	4	1	16
1	5	5	1	25
4	4	16	16	16
4	5	20	16	25
25	11	275	625	121
25	12	300	625	144
25	13	325	625	169
85	54	945	1909	516
12.1429	7.7143	135	272.7143	73.7143
$M(X^2)$	$M(Y)$	$M(X^2 Y)$	$M(X^{2^2})$	$M(Y^2)$

$$\hat{\beta} = \frac{\text{Cov}(X^2, Y)}{\text{Var}(X^2)} = \frac{M(X^2 Y) - M(X^2) M(Y)}{M(X^{2^2}) - [M(X^2)^2]} =$$

$$= \frac{135 - 12.1429 \cdot 7.7143}{272.7143 - 12.1429^2} = \frac{41.3265}{125.2653} = 0.3299$$

$$\hat{\alpha} = M(Y) - \hat{\beta} M(X^2) = 3.7082$$

$$\rho^2 =$$

²Si osserva che per ragioni di stampa alcuni calcoli potrebbero essere duplicati con riferimento ai diversi modelli oggetto di studio



Si riportano i calcoli necessari² per ottenere le stime dei parametri e gli indici di adattamento del modello

$$Y = \alpha + \beta X^2 + E$$

x_i^2	y_i	$x_i^2 y_i$	$x_i^{2^2}$	y_i^2
1	4	4	1	16
1	5	5	1	25
4	4	16	16	16
4	5	20	16	25
25	11	275	625	121
25	12	300	625	144
25	13	325	625	169
85	54	945	1909	516
12.1429	7.7143	135	272.7143	73.7143
$M(X^2)$	$M(Y)$	$M(X^2 Y)$	$M(X^{2^2})$	$M(Y^2)$

$$\begin{aligned}\hat{\beta} &= \frac{\text{Cov}(X^2, Y)}{\text{Var}(X^2)} = \frac{M(X^2 Y) - M(X^2) M(Y)}{M(X^{2^2}) - [M(X^2)^2]} = \\ &= \frac{135 - 12.1429 \cdot 7.7143}{272.7143 - 12.1429^2} = \frac{41.3265}{125.2653} = 0.3299 \\ \hat{\alpha} &= M(Y) - \hat{\beta} M(X^2) = 3.7082 \\ \rho^2 &= \frac{[\text{Cov}(X^2, Y)]^2}{\text{Var}(X^2) \text{Var}(Y)} = \frac{41.3265^2}{125.2653 \cdot 14.2041} = 0.9599\end{aligned}$$

²Si osserva che per ragioni di stampa alcuni calcoli potrebbero essere duplicati con riferimento ai diversi modelli oggetto di studio





Il modello migliore è

$$Y^* = 0.3299 + 3.708X^2$$

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.



3. Indicare l'intervallo dei valori possibili per $\eta^2_{Y|X}$.

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.



3. Indicare l'intervallo dei valori possibili per $\eta_{Y|X}^2$.

$$0.9599 < \eta_{Y|X}^2 < 1.$$

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.



Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.

3. Indicare l'intervallo dei valori possibili per $\eta_{Y|X}^2$.

$$0.9599 < \eta_{Y|X}^2 < 1.$$

Si osserva che le medie condizionate non stanno su una parabola dal momento che $M(Y|x=1) = M(Y|x=2) = 4.5$.



Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.

3. Indicare l'intervallo dei valori possibili per $\eta_{Y|X}^2$.

$$0.9599 < \eta_{Y|X}^2 < 1.$$

Si osserva che le medie condizionate non stanno su una parabola dal momento che $M(Y|x=1) = M(Y|x=2) = 4.5$. Non sussiste nemmeno dipendenza funzionale di Y da X .

Esercizio 4

Una linea di produzione confeziona pacchetti di caramelle il cui peso complessivo è fissato al valore nominale di $250g$. Sapendo che detti pesi seguono una v.c. $N(\mu = 250, \sigma^2 = 21)$ e che si ritengono difettosi i pacchetti con peso inferiore del 3% rispetto al valore nominale:

1. calcolare la probabilità che la linea produca pezzi difettosi;



Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.

Esercizio 4

Una linea di produzione confeziona pacchetti di caramelle il cui peso complessivo è fissato al valore nominale di $250g$. Sapendo che detti pesi seguono una v.c. $N(\mu = 250, \sigma^2 = 21)$ e che si ritengono difettosi i pacchetti con peso inferiore del 3% rispetto al valore nominale:

1. calcolare la probabilità che la linea produca pezzi difettosi;

Si ritengono difettosi i pacchetti con peso inferiore a



Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.

Esercizio 4



Una linea di produzione confeziona pacchetti di caramelle il cui peso complessivo è fissato al valore nominale di $250g$. Sapendo che detti pesi seguono una v.c. $N(\mu = 250, \sigma^2 = 21)$ e che si ritengono difettosi i pacchetti con peso inferiore del 3% rispetto al valore nominale:

1. calcolare la probabilità che la linea produca pezzi difettosi;

Si ritengono difettosi i pacchetti con peso inferiore a

$$242.5 = 0.97 \cdot 250$$

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.



Una linea di produzione confeziona pacchetti di caramelle il cui peso complessivo è fissato al valore nominale di $250g$. Sapendo che detti pesi seguono una v.c. $N(\mu = 250, \sigma^2 = 21)$ e che si ritengono difettosi i pacchetti con peso inferiore del 3% rispetto al valore nominale:

1. calcolare la probabilità che la linea produca pezzi difettosi;

Si ritengono difettosi i pacchetti con peso inferiore a

$$242.5 = 0.97 \cdot 250$$

Indicata con $X \sim N(\mu = 250, \sigma^2 = 21)$ la variabile casuale che descrive il peso di un pacchetto di caramelle, la probabilità che il peso sia inferiore a $242.5g$ risulta

Esercizio 4



Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.

Una linea di produzione confeziona pacchetti di caramelle il cui peso complessivo è fissato al valore nominale di $250g$. Sapendo che detti pesi seguono una v.c. $N(\mu = 250, \sigma^2 = 21)$ e che si ritengono difettosi i pacchetti con peso inferiore del 3% rispetto al valore nominale:

1. calcolare la probabilità che la linea produca pezzi difettosi;

Si ritengono difettosi i pacchetti con peso inferiore a

$$242.5 = 0.97 \cdot 250$$

Indicata con $X \sim N(\mu = 250, \sigma^2 = 21)$ la variabile casuale che descrive il peso di un pacchetto di caramelle, la probabilità che il peso sia inferiore a $242.5g$ risulta

$$P(X \leq 242.5) =$$

Esercizio 4



Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.

Una linea di produzione confeziona pacchetti di caramelle il cui peso complessivo è fissato al valore nominale di $250g$. Sapendo che detti pesi seguono una v.c. $N(\mu = 250, \sigma^2 = 21)$ e che si ritengono difettosi i pacchetti con peso inferiore del 3% rispetto al valore nominale:

1. calcolare la probabilità che la linea produca pezzi difettosi;

Si ritengono difettosi i pacchetti con peso inferiore a

$$242.5 = 0.97 \cdot 250$$

Indicata con $X \sim N(\mu = 250, \sigma^2 = 21)$ la variabile casuale che descrive il peso di un pacchetto di caramelle, la probabilità che il peso sia inferiore a $242.5g$ risulta

$$P(X \leq 242.5) = P\left(\frac{X - \mu}{\sigma} \leq \frac{242.5 - \mu}{\sigma}\right) =$$



Una linea di produzione confeziona pacchetti di caramelle il cui peso complessivo è fissato al valore nominale di 250g. Sapendo che detti pesi seguono una v.c. $N(\mu = 250, \sigma^2 = 21)$ e che si ritengono difettosi i pacchetti con peso inferiore del 3% rispetto al valore nominale:

1. calcolare la probabilità che la linea produca pezzi difettosi;

Si ritengono difettosi i pacchetti con peso inferiore a

$$242.5 = 0.97 \cdot 250$$

Indicata con $X \sim N(\mu = 250, \sigma^2 = 21)$ la variabile casuale che descrive il peso di un pacchetto di caramelle, la probabilità che il peso sia inferiore a 242.5g risulta

$$\begin{aligned} P(X \leq 242.5) &= P\left(\frac{X - \mu}{\sigma} \leq \frac{242.5 - \mu}{\sigma}\right) = P\left(Z \leq \frac{242.5 - 250}{4.5826}\right) = \\ &= \end{aligned}$$



Una linea di produzione confeziona pacchetti di caramelle il cui peso complessivo è fissato al valore nominale di 250g. Sapendo che detti pesi seguono una v.c. $N(\mu = 250, \sigma^2 = 21)$ e che si ritengono difettosi i pacchetti con peso inferiore del 3% rispetto al valore nominale:

1. calcolare la probabilità che la linea produca pezzi difettosi;

Si ritengono difettosi i pacchetti con peso inferiore a

$$242.5 = 0.97 \cdot 250$$

Indicata con $X \sim N(\mu = 250, \sigma^2 = 21)$ la variabile casuale che descrive il peso di un pacchetto di caramelle, la probabilità che il peso sia inferiore a 242.5g risulta

$$\begin{aligned} P(X \leq 242.5) &= P\left(\frac{X - \mu}{\sigma} \leq \frac{242.5 - \mu}{\sigma}\right) = P\left(Z \leq \frac{242.5 - 250}{4.5826}\right) = \\ &= P(Z \leq -1.64) = \end{aligned}$$



Una linea di produzione confeziona pacchetti di caramelle il cui peso complessivo è fissato al valore nominale di 250g. Sapendo che detti pesi seguono una v.c. $N(\mu = 250, \sigma^2 = 21)$ e che si ritengono difettosi i pacchetti con peso inferiore del 3% rispetto al valore nominale:

1. calcolare la probabilità che la linea produca pezzi difettosi;

Si ritengono difettosi i pacchetti con peso inferiore a

$$242.5 = 0.97 \cdot 250$$

Indicata con $X \sim N(\mu = 250, \sigma^2 = 21)$ la variabile casuale che descrive il peso di un pacchetto di caramelle, la probabilità che il peso sia inferiore a 242.5g risulta

$$\begin{aligned} P(X \leq 242.5) &= P\left(\frac{X - \mu}{\sigma} \leq \frac{242.5 - \mu}{\sigma}\right) = P\left(Z \leq \frac{242.5 - 250}{4.5826}\right) = \\ &= P(Z \leq -1.64) = P(Z \geq 1.64) = \\ &= \end{aligned}$$



Una linea di produzione confeziona pacchetti di caramelle il cui peso complessivo è fissato al valore nominale di 250g. Sapendo che detti pesi seguono una v.c. $N(\mu = 250, \sigma^2 = 21)$ e che si ritengono difettosi i pacchetti con peso inferiore del 3% rispetto al valore nominale:

1. calcolare la probabilità che la linea produca pezzi difettosi;

Si ritengono difettosi i pacchetti con peso inferiore a

$$242.5 = 0.97 \cdot 250$$

Indicata con $X \sim N(\mu = 250, \sigma^2 = 21)$ la variabile casuale che descrive il peso di un pacchetto di caramelle, la probabilità che il peso sia inferiore a 242.5g risulta

$$\begin{aligned} P(X \leq 242.5) &= P\left(\frac{X - \mu}{\sigma} \leq \frac{242.5 - \mu}{\sigma}\right) = P\left(Z \leq \frac{242.5 - 250}{4.5826}\right) = \\ &= P(Z \leq -1.64) = P(Z \geq 1.64) = \\ &= 1 - P(Z < 1.64) = \end{aligned}$$



Una linea di produzione confeziona pacchetti di caramelle il cui peso complessivo è fissato al valore nominale di 250g. Sapendo che detti pesi seguono una v.c. $N(\mu = 250, \sigma^2 = 21)$ e che si ritengono difettosi i pacchetti con peso inferiore del 3% rispetto al valore nominale:

1. calcolare la probabilità che la linea produca pezzi difettosi;

Si ritengono difettosi i pacchetti con peso inferiore a

$$242.5 = 0.97 \cdot 250$$

Indicata con $X \sim N(\mu = 250, \sigma^2 = 21)$ la variabile casuale che descrive il peso di un pacchetto di caramelle, la probabilità che il peso sia inferiore a 242.5g risulta

$$\begin{aligned} P(X \leq 242.5) &= P\left(\frac{X - \mu}{\sigma} \leq \frac{242.5 - \mu}{\sigma}\right) = P\left(Z \leq \frac{242.5 - 250}{4.5826}\right) = \\ &= P(Z \leq -1.64) = P(Z \geq 1.64) = \\ &= 1 - P(Z < 1.64) = 1 - 0.9495 = 0.0505. \end{aligned}$$



Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.

2. calcolare la probabilità che su 100 pacchetti estratti (con re-immissione) dal flusso produttivo, almeno il 5% risulti fuori norma.



2. calcolare la probabilità che su 100 pacchetti estratti (con re-immissione) dal flusso produttivo, almeno il 5% risulti fuori norma.

Indicata ora con X la variabile casuale 'numero di pacchetti fuori norma' tra i 100 estratti con reimmissione (prove indipendenti) risulta

$$X \sim$$

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.



Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.

2. calcolare la probabilità che su 100 pacchetti estratti (con re-immissione) dal flusso produttivo, almeno il 5% risulti fuori norma.

Indicata ora con X la variabile casuale 'numero di pacchetti fuori norma' tra i 100 estratti con reimmissione (prove indipendenti) risulta

$$X \sim \text{Bin}(n = 100, p = 0.0505).$$

$$M(X) = 5.05; \text{Var}(X) = 4.795.$$

Dal momento che $np = 5.05 > 5$ ed $n(1 - p) = 94.95 > 5$ è possibile utilizzare l'approssimazione della variabile casuale Binomiale con la variabile casuale Normale

$$X \sim$$



2. calcolare la probabilità che su 100 pacchetti estratti (con re-immissione) dal flusso produttivo, almeno il 5% risulti fuori norma.

Indicata ora con X la variabile casuale 'numero di pacchetti fuori norma' tra i 100 estratti con reimmissione (prove indipendenti) risulta

$$X \sim \text{Bin}(n = 100, p = 0.0505).$$

$$M(X) = 5.05; \text{Var}(X) = 4.795.$$

Dal momento che $np = 5.05 > 5$ ed $n(1 - p) = 94.95 > 5$ è possibile utilizzare l'approssimazione della variabile casuale Binomiale con la variabile casuale Normale

$$X \sim N(\mu = 5.05, \sigma^2 = 4.795).$$

Si ottiene

$$\begin{aligned} P(X \geq 5) &= P\left(\frac{X - \mu}{\sigma} \geq \frac{5 - \mu}{\sigma}\right) = P\left(Z \geq \frac{5 - 5.05}{2.1897}\right) = \\ &= P(Z \geq -0.02) = P(Z \leq 0.02) = 0.5080, \end{aligned}$$

essendo 5 il 5% di 100.

Si osserva come la condizione $np = 5.05 > 5$ sia marginalmente verificata. Applicando direttamente la variabile casuale binomiale si ottiene



Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.



Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.

Si osserva come la condizione $np = 5.05 > 5$ sia marginalmente verificata. Applicando direttamente la variabile casuale binomiale si ottiene

$$\begin{aligned} P(X \geq 5) &= 1 - P(X < 5) = 1 - P(X \leq 4) = \\ &= 1 - \left[\binom{100}{0} 0.0505^0 \cdot 0.9495^{100} + \binom{100}{1} 0.0505^1 \cdot 0.9495^{99} + \binom{100}{2} 0.0505^2 \cdot 0.9495^{98} + \right. \\ &\quad \left. + \binom{100}{3} 0.0505^3 \cdot 0.9495^{97} + \binom{100}{4} 0.0505^4 \cdot 0.9495^{96} \right] = \\ &= 1 - (0.0056 + 0.0299 + 0.0786 + 0.1366 + 0.1762) = 1 - 0.4269 = 0.5731. \end{aligned}$$



Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.

Si osserva come la condizione $np = 5.05 > 5$ sia marginalmente verificata. Applicando direttamente la variabile casuale binomiale si ottiene

$$\begin{aligned} P(X \geq 5) &= 1 - P(X < 5) = 1 - P(X \leq 4) = \\ &= 1 - \left[\binom{100}{0} 0.0505^0 \cdot 0.9495^{100} + \binom{100}{1} 0.0505^1 \cdot 0.9495^{99} + \binom{100}{2} 0.0505^2 \cdot 0.9495^{98} + \right. \\ &\quad \left. + \binom{100}{3} 0.0505^3 \cdot 0.9495^{97} + \binom{100}{4} 0.0505^4 \cdot 0.9495^{96} \right] = \\ &= 1 - (0.0056 + 0.0299 + 0.0786 + 0.1366 + 0.1762) = 1 - 0.4269 = 0.5731. \end{aligned}$$

Con riferimento all'approssimazione con la variabile casuale normale si ottiene un risultato più preciso avvalendosi della correzione per continuità



Si osserva come la condizione $np = 5.05 > 5$ sia marginalmente verificata. Applicando direttamente la variabile casuale binomiale si ottiene

$$\begin{aligned} P(X \geq 5) &= 1 - P(X < 5) = 1 - P(X \leq 4) = \\ &= 1 - \left[\binom{100}{0} 0.0505^0 \cdot 0.9495^{100} + \binom{100}{1} 0.0505^1 \cdot 0.9495^{99} + \binom{100}{2} 0.0505^2 \cdot 0.9495^{98} + \right. \\ &\quad \left. + \binom{100}{3} 0.0505^3 \cdot 0.9495^{97} + \binom{100}{4} 0.0505^4 \cdot 0.9495^{96} \right] = \\ &= 1 - (0.0056 + 0.0299 + 0.0786 + 0.1366 + 0.1762) = 1 - 0.4269 = 0.5731. \end{aligned}$$

Con riferimento all'approssimazione con la variabile casuale normale si ottiene un risultato più preciso avvalendosi della correzione per continuità

$$\begin{aligned} P(X \geq 5) &= P\left(\frac{X - .5 - \mu}{\sigma} \geq \frac{5 - .5 - \mu}{\sigma}\right) = P\left(Z \geq \frac{4.5 - 5.05}{2.1897}\right) = \\ &= P(Z \geq -0.25) = P(Z \leq 0.25) = 0.5987. \end{aligned}$$

Si osserva, infatti, che lo scostamento dell'ultimo risultato da quello esatto ottenuto con l'applicazione diretta della variabile casuale binomiale è inferiore rispetto a quello precedentemente ottenuto non considerando la correzione per continuità.

Tema T283

Prova scritta di STATISTICA del 12.02.2015

Prova scritta di
STATISTICA del
12.02.2015

Giuseppe Boari
Gabriele Cantaluppi



UNIVERSITÀ
CATTOLICA
del Sacro Cuore

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.

Giuseppe Boari
Gabriele Cantaluppi
Università Cattolica del Sacro Cuore, Milano



Di una v.s. doppia (X, Y) si dispone della seguente distribuzione delle densità di frequenza della condizionata $Y|X_1$:

$h_{j-1} \vdash h_j$	d_{1j}	n_{1j}
0 $\vdash$ 2	4	
2 $\vdash$ 6	8	
6 $\vdash$ 8	4	

1. Si ricostruisca la distribuzione delle frequenze assolute, n_{1j} , di $Y|X_1$ e se ne dia una rappresentazione grafica.

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.



Moltiplicando le densità di frequenza, d_{1j} , per le ampiezze di classe, a_j , si ottengono le frequenze assolute n_{1j} della variabile Y ,

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.



Moltiplicando le densità di frequenza, d_{1j} , per le ampiezze di classe, a_j , si ottengono le frequenze assolute n_{1j} della variabile Y ,

y_j	d_{1j}	a_j	n_{1j}
0 - 2	4	$2 - 0 = 2$	8
2 - 6	8	$6 - 2 = 4$	32
6 - 8	4	$8 - 6 = 2$	8

Esercizio 1.

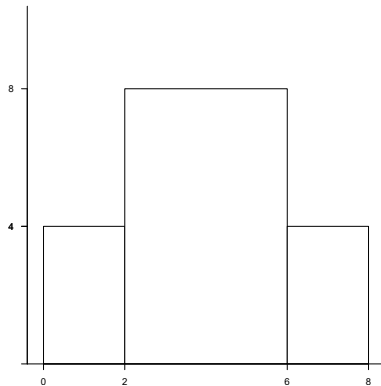
Esercizio 2.

Esercizio 3.

Esercizio 4.



Si tratta di un carattere quantitativo con modalità raggruppate in classi che può essere rappresentato graficamente mediante un istogramma. L'istogramma può essere costruito dopo avere calcolato le densità $d_{1j} = \frac{n_{1j}}{a_j}$, (dove a_j sono le ampiezze di classe).



Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.



Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.

2. Si calcolino la media aritmetica, la mediana e la varianza della variabile $Y|x_1$.



Si tratta di un carattere quantitativo con modalità raggruppate in classi.

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.



Si osserva che la distribuzione è simmetrica rispetto a 4; la mediana (percentile di ordine 0.50) risulta $y_{0.5} = 4$.

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.



La media aritmetica

$$M(Y|x_1) = \frac{1}{n_{1\bullet}} \sum_{j=1}^k y_j n_{1j}$$

risulta 4.

Si osserva, infatti, che la distribuzione è simmetrica.

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.



La varianza

$$\text{Var}(Y|x_1) = M(Y^2|x_1) - [M(Y|x_1)]^2$$

risulta 3.

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.



Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.

La varianza

$$\text{Var}(Y|x_1) = M(Y^2|x_1) - [M(Y|x_1)]^2$$

risulta 3. Si ottiene (considerando come modalità i valori centrali delle classi)

y_j	n_{1j}	$y_j n_{1j}$	$y_j^2 n_{1j}$
1	8	8	8
4	32	128	512
7	8	56	392
	48	192	912

da cui segue

$$\text{Var}(Y|x_1) = \frac{912}{48} - 4^2 = 19 - 16 = 3.$$



Per effettuare confronti in termini di variabilità tra caratteri quantitativi occorre utilizzare una misura relativa, ad esempio il coefficiente di variazione

$$CV = \frac{\sigma}{\mu} = \frac{\sqrt{\text{Var}(Y|x_1)}}{M(Y|x_1)} = \frac{\sqrt{3}}{4} = \frac{1.7321}{4} = 0.433.$$

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.



3. Sapendo che la distribuzione condizionata $Y|x_2$, riferita a $n_{2\bullet} = 16$ unità statistiche, è caratterizzata da media $M(Y|x_2) = 4$ e varianza $Var(Y|x_2) = 0$ si ricostruisca la distribuzione congiunta di X e Y e si misuri la dipendenza tra X e Y con un opportuno indice di connessione.

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.



Dal momento che $Var(Y|x_2) = 0$ la distribuzione condizionata $Y|x_2$ risulta degenerare.

$X \setminus Y$	0 + 2	2 + 6	6 + 8	
x_1	8	32	8	48
x_2	0	16	0	16
	8	48	8	64

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.



Dalla distribuzione delle frequenze congiunte n_{ij} :

	0 ÷ 2	2 ÷ 6	6 ÷ 8	$n_{i\bullet}$
x_1	8	32	8	48
x_2	0	16	0	16
$n_{\bullet j}$	8	48	8	64

è possibile ottenere le frequenze teoriche nella situazione di indipendenza stocastica

$$\hat{n}_{ij} = \frac{n_{i\bullet} n_{\bullet j}}{n}:$$

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.



Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.

Dalla distribuzione delle frequenze congiunte n_{ij} :

	0 + 2	2 + 6	6 + 8	$n_{i\bullet}$
x_1	8	32	8	48
x_2	0	16	0	16
$n_{\bullet j}$	8	48	8	64

è possibile ottenere le frequenze teoriche nella situazione di indipendenza stocastica

$$\hat{n}_{ij} = \frac{n_{i\bullet} n_{\bullet j}}{n}:$$

	0 + 2	2 + 6	6 + 8	$n_{i\bullet}$
x_1	6	36	6	48
x_2	2	12	2	16
$n_{\bullet j}$	8	48	8	64



le contingenze assolute $n_{ij} - \hat{n}_{ij}$:

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.



le contingenze assolute $n_{ij} - \hat{n}_{ij}$:

	0 - 2	2 - 6	6 - 8
x_1	2	-4	2
x_2	-2	4	-2

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.



i rapporti di contingenza $\frac{n_{ij}}{\hat{n}_{ij}}$:

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.



i rapporti di contingenza $\frac{n_{ij}}{\hat{n}_{ij}}$:

	0 ÷ 2	2 ÷ 6	6 ÷ 8
x_1	1.3333	0.8889	1.3333
x_2	0.0000	1.3333	0.0000

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.



e le contingenze relative, espresse in termini percentuali, $100 \left(\frac{n_{ij} - \hat{n}_{ij}}{\hat{n}_{ij}} \right) \% = 100 \left(\frac{n_{ij}}{\hat{n}_{ij}} - 1 \right) \%$:

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.



e le contingenze relative, espresse in termini percentuali, $100 \left(\frac{n_{ij} - \hat{n}_{ij}}{\hat{n}_{ij}} \right) \% = 100 \left(\frac{n_{ij}}{\hat{n}_{ij}} - 1 \right) \%$:

	0 ÷ 2	2 ÷ 6	6 ÷ 8
x_1	33.33	-11.11	33.33
x_2	-100.00	33.33	-100.00

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.



Risultati intermedi $\frac{(n_{ij} - \hat{n}_{ij})^2}{\hat{n}_{ij}}$ per il calcolo del χ^2

	0 ÷ 2	2 ÷ 6	6 ÷ 8
x_1	0.6667	0.4444	0.6667
x_2	2.0000	1.3333	2.0000

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.



Risultati intermedi $\frac{n_{ij}^2}{n_{i\bullet} \cdot n_{\bullet j}}$ per il calcolo del χ^2 (formula operativa)

	0 ÷ 2	2 ÷ 6	6 ÷ 8
x_1	0.1667	0.4444	0.1667
x_2	0.0000	0.3333	0.0000

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.



Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.

$$\chi^2 = \sum_{i=1}^h \sum_{j=1}^k \frac{(n_{ij} - \hat{n}_{ij})^2}{\hat{n}_{ij}} = 7.1111,$$

$$\chi_N^2 = \frac{\chi^2}{n \cdot \min(h-1, k-1)} = 0.1111$$

$$\chi_N = \sqrt{\chi_N^2} = 0.3333.$$

Sussiste, quindi, un livello di connessione basso.



4. Ponendo ora $n_{2\bullet} = 48$, si indichi quali dovrebbero essere le frequenze della distribuzione condizionata $Y|x_2$ affinché sussista indipendenza stocastica tra X e Y .

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.



4. Ponendo ora $n_{2\bullet} = 48$, si indichi quali dovrebbero essere le frequenze della distribuzione condizionata $Y|_{x_2}$ affinché sussista indipendenza stocastica tra X e Y .

Affinché sussista indipendenza stocastica tra X e Y le distribuzioni condizionate $Y|_{x_i}$, $i = 1, 2$ devono essere somiglianti

$$Y|_{x_1} \sim Y|_{x_2}$$

$X \setminus Y$	0 + 2	2 + 6	6 + 8	
x_1	8	32	8	48
x_2	8	32	8	48
	16	64	16	96

Nel caso in esame $Y|_{x_1}$ e $Y|_{x_2}$ risultano identiche.

Esercizio 2



Il prospetto seguente riporta i prezzi di chiusura di 2 titoli azionari nei giorni dal 7 al 10 gennaio:

t	<i>titolo A</i>	<i>titolo B</i>
7	100	100
8	120	100
9	144	110
10	150	121

1. Si calcoli e interpreti il numero indice al tempo 7 con base 9 per il *titolo A*.

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.



Il prospetto seguente riporta i prezzi di chiusura di 2 titoli azionari nei giorni dal 7 al 10 gennaio:

t	<i>titolo A</i>	<i>titolo B</i>
7	100	100
8	120	100
9	144	110
10	150	121

1. Si calcoli e interpreti il numero indice al tempo 7 con base 9 per il *titolo A*.

$${}_9I_7 = \frac{x_7}{x_9} = \frac{100}{144} = 0.6944$$

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.



Il prospetto seguente riporta i prezzi di chiusura di 2 titoli azionari nei giorni dal 7 al 10 gennaio:

t	<i>titolo A</i>	<i>titolo B</i>
7	100	100
8	120	100
9	144	110
10	150	121

1. Si calcoli e interpreti il numero indice al tempo 7 con base 9 per il *titolo A*.

$${}_9I_7 = \frac{x_7}{x_9} = \frac{100}{144} = 0.6944$$

La grandezza al tempo 7 era inferiore del 30.56% rispetto alla grandezza al tempo 9.
Ovvero il valore della grandezza al tempo 7 era il 69.44% del valore registrato al tempo 9.

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.



2. Si calcoli il tasso medio di variazione registrato tra il giorno 7 e il giorno 9 per il *titolo A*, interpretando il risultato.

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.



2. Si calcoli il tasso medio di variazione registrato tra il giorno 7 e il giorno 9 per il *titolo A*, interpretando il risultato.

Si ricorda che

$${}_7I_9 = {}_7I_8 \cdot {}_8I_9$$

Si deve trovare il numero indice a base mobile medio.

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.



2. Si calcoli il tasso medio di variazione registrato tra il giorno 7 e il giorno 9 per il *titolo A*, interpretando il risultato.

Si ricorda che

$${}_7l_9 = {}_7l_8 \cdot {}_8l_9$$

Si deve trovare il numero indice a base mobile medio. Applicando il criterio di scelta della media secondo Chisini si ottiene

$${}_7l_9 = {}_7l_8 \cdot {}_8l_9 = \alpha \cdot \alpha = \alpha^2.$$

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.



Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.

2. Si calcoli il tasso medio di variazione registrato tra il giorno 7 e il giorno 9 per il *titolo A*, interpretando il risultato.

Si ricorda che

$${}_7I_9 = {}_7I_8 \cdot {}_8I_9$$

Si deve trovare il numero indice a base mobile medio. Applicando il criterio di scelta della media secondo Chisini si ottiene

$${}_7I_9 = {}_7I_8 \cdot {}_8I_9 = \alpha \cdot \alpha = \alpha^2.$$

Quindi

$$\alpha = \sqrt[2]{{}_7I_9} = \sqrt[2]{{}_7I_8 \cdot {}_8I_9} = 1.2.$$

Da cui segue una variazione relativa media pari a 0.2, corrispondente a un incremento medio giornaliero tra il giorno 7 e il giorno 9 del 20%.



3. Supponendo che la media aritmetica calcolata a partire dai valori degli indici di Laspeyres, $\frac{P}{I}_{10}^L$, e di Paasche, $\frac{P}{I}_{10}^P$, sia risultata pari a 1.31, si indichi se l'indice di Fisher può assumere un valore maggiore di 1.31.

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.



3. Supponendo che la media aritmetica calcolata a partire dai valori degli indici di Laspeyres, $\frac{P}{7}I_{10}^L$, e di Paasche, $\frac{P}{7}I_{10}^P$, sia risultata pari a 1.31, si indichi se l'indice di Fisher può assumere un valore maggiore di 1.31.

L'indice dei prezzi di Fisher è definito come media geometrica degli indici dei prezzi di Laspeyres e di Paasche.

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.



3. Supponendo che la media aritmetica calcolata a partire dai valori degli indici di Laspeyres, $\frac{P}{I}_{10}^L$, e di Paasche, $\frac{P}{I}_{10}^P$, sia risultata pari a 1.31, si indichi se l'indice di Fisher può assumere un valore maggiore di 1.31.

L'indice dei prezzi di Fisher è definito come media geometrica degli indici dei prezzi di Laspeyres e di Paasche.

La media geometrica dei due indici risulterà inferiore o eguale alla media aritmetica degli stessi; l'eguaglianza si può realizzare solo nel caso in cui l'indice di Laspeyres coincide con quello di Paasche.

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.



Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.

Il seguente prospetto riporta i valori rilevati in corrispondenza a 5 unità statistiche con riferimento a due caratteri quantitativi X e Y

x_i	1	2	5	5	5
y_i	4	5	11	12	13

1. Fornire la rappresentazione grafica delle coppie di punti, insieme alla funzione di regressione di Y dato X .



La funzione di regressione associa a ogni valore x appartenente all'insieme di definizione della variabile X la media di Y condizionata a x , $M(Y|x)$

x	$M(Y x)$
1	4
2	5
5	$\frac{1}{3}(11 + 12 + 13) = 12$

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.



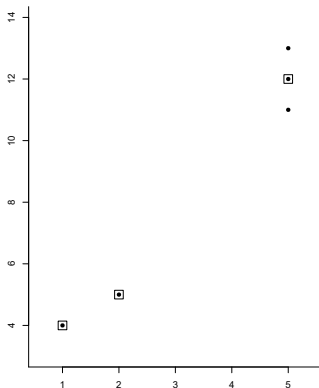
Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.

Le coppie di dati possono essere rappresentate mediante un grafico a dispersione.





2. Si calcolino, secondo il metodo dei minimi quadrati, i parametri e l'indice di adattamento dei seguenti modelli di regressione, individuando il modello migliore:

$$I) Y^* = a + b \cdot X \qquad II) Y^* = c$$

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.



Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.

Si riportano i calcoli necessari¹ per ottenere le stime dei parametri e gli indici di adattamento del modello

$$Y = a + bX + E$$

x_i	y_i	$x_i y_i$	x_i^2	y_i^2
1	4	4	1	16
2	5	10	4	25
5	11	55	25	121
5	12	60	25	144
5	13	65	25	169
18	45	194	80	475
3.6	9	38.8	16	95
$M(X)$	$M(Y)$	$M(XY)$	$M(X^2)$	$M(Y^2)$

$$\hat{b} =$$

¹Si osserva che per ragioni di stampa alcuni calcoli potrebbero essere duplicati con riferimento ai diversi modelli oggetto di studio



Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.

Si riportano i calcoli necessari¹ per ottenere le stime dei parametri e gli indici di adattamento del modello

$$Y = a + bX + E$$

x_i	y_i	$x_i y_i$	x_i^2	y_i^2
1	4	4	1	16
2	5	10	4	25
5	11	55	25	121
5	12	60	25	144
5	13	65	25	169
18	45	194	80	475
3.6	9	38.8	16	95
$M(X)$	$M(Y)$	$M(XY)$	$M(X^2)$	$M(Y^2)$

$$\begin{aligned}\hat{b} &= \frac{\text{Cov}(X, Y)}{\text{Var}(X)} = \frac{M(XY) - M(X)M(Y)}{M(X^2) - [M(X)]^2} = \\ &= \frac{38.8 - 3.6 \cdot 9}{16 - 3.6^2} = \frac{6.4}{3.04} = 2.1053 \\ \hat{a} &= \end{aligned}$$

¹Si osserva che per ragioni di stampa alcuni calcoli potrebbero essere duplicati con riferimento ai diversi modelli oggetto di studio



Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.

Si riportano i calcoli necessari¹ per ottenere le stime dei parametri e gli indici di adattamento del modello

$$Y = a + bX + E$$

x_i	y_i	$x_i y_i$	x_i^2	y_i^2
1	4	4	1	16
2	5	10	4	25
5	11	55	25	121
5	12	60	25	144
5	13	65	25	169
18	45	194	80	475
3.6	9	38.8	16	95
$M(X)$	$M(Y)$	$M(XY)$	$M(X^2)$	$M(Y^2)$

$$\begin{aligned}\hat{b} &= \frac{\text{Cov}(X, Y)}{\text{Var}(X)} = \frac{M(XY) - M(X)M(Y)}{M(X^2) - [M(X)]^2} = \\ &= \frac{38.8 - 3.6 \cdot 9}{16 - 3.6^2} = \frac{6.4}{3.04} = 2.1053 \\ \hat{a} &= M(Y) - \hat{b}M(X) = 1.4211 \\ \rho^2 &= \end{aligned}$$

¹Si osserva che per ragioni di stampa alcuni calcoli potrebbero essere duplicati con riferimento ai diversi modelli oggetto di studio



Si riportano i calcoli necessari¹ per ottenere le stime dei parametri e gli indici di adattamento del modello

$$Y = a + bX + E$$

x_i	y_i	$x_i y_i$	x_i^2	y_i^2
1	4	4	1	16
2	5	10	4	25
5	11	55	25	121
5	12	60	25	144
5	13	65	25	169
18	45	194	80	475
3.6	9	38.8	16	95
$M(X)$	$M(Y)$	$M(XY)$	$M(X^2)$	$M(Y^2)$

$$\begin{aligned}\hat{b} &= \frac{\text{Cov}(X, Y)}{\text{Var}(X)} = \frac{M(XY) - M(X)M(Y)}{M(X^2) - [M(X)]^2} = \\ &= \frac{38.8 - 3.6 \cdot 9}{16 - 3.6^2} = \frac{6.4}{3.04} = 2.1053 \\ \hat{a} &= M(Y) - \hat{b}M(X) = 1.4211 \\ \rho^2 &= \frac{[\text{Cov}(X, Y)]^2}{\text{Var}(X)\text{Var}(Y)} = \frac{6.4^2}{3.04 \cdot 14} = 0.9624\end{aligned}$$

¹Si osserva che per ragioni di stampa alcuni calcoli potrebbero essere duplicati con riferimento ai diversi modelli oggetto di studio



Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.

Si riportano i calcoli necessari² per ottenere le stime dei parametri e gli indici di adattamento del modello

$$Y = c + E$$

x_i	y_i	y_i^2
1	4	16
2	5	25
5	11	121
5	12	144
5	13	169
		475
		95
		$M(Y)$ $M(Y^2)$

$$\hat{c} =$$

²Si osserva che per ragioni di stampa alcuni calcoli potrebbero essere duplicati con riferimento ai diversi modelli oggetto di studio



Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.

Si riportano i calcoli necessari² per ottenere le stime dei parametri e gli indici di adattamento del modello

$$Y = c + E$$

x_i	y_i	y_i^2
1	4	16
2	5	25
5	11	121
5	12	144
5	13	169
		475
		95
		$M(Y)$ $M(Y^2)$

$$\hat{c} = M(Y) = 9$$

$$R^2 =$$

²Si osserva che per ragioni di stampa alcuni calcoli potrebbero essere duplicati con riferimento ai diversi modelli oggetto di studio



Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.

Si riportano i calcoli necessari² per ottenere le stime dei parametri e gli indici di adattamento del modello

$$Y = c + E$$

x_i	y_i	y_i^2
1	4	16
2	5	25
5	11	121
5	12	144
5	13	169
45		475
9		95
$M(Y)$		$M(Y^2)$

$$\hat{c} = M(Y) = 9$$

$$R^2 = 0$$

²Si osserva che per ragioni di stampa alcuni calcoli potrebbero essere duplicati con riferimento ai diversi modelli oggetto di studio



Il modello migliore è $Y^* = 1.4211 + 2.1053X$.

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.



Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.

3. Si calcoli $\eta^2_{Y|X}$.



Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.

Il rapporto di correlazione

$$\eta_{Y|X}^2 = \frac{\text{Var}_X[M(Y|X)]}{\text{Var}(Y)}$$

può essere ottenuto considerando il rapporto tra la varianza spiegata $\text{Var}_X[M(Y|X)]$,
varianza della variabile statistica medie condizionate

x	$M(Y x)$
1	4
2	5
5	12

e la varianza della variabile dipendente $\text{Var}(Y) = 14$.

La varianza spiegata risulta $\text{Var}_X[M(Y|X)] = 13.6$.



Si ottiene

x	$M(Y x)$	$n(x)$	$M(Y x) \cdot n(x)$	$M(Y^2 x) \cdot n(x)$
1	4	1	4	16
2	5	1	5	25
5	12	3	36	432
		5	45	473

da cui segue

$$Var_X[M(Y|X)] = \frac{473}{5} - 9^2 = 94.6 - 81 = 13.6.$$

Il rapporto di correlazione $\eta_{Y|X}^2$ risulta:

$$\eta_{Y|X}^2 = \frac{Var_X[M(Y|X)]}{Var(Y)} = \frac{13.6}{14} = 0.9714.$$

Sussiste, quindi, un livello di dipendenza molto alto.



Il rapporto di correlazione $\eta^2_{Y|X}$ può anche essere ottenuto nel modo seguente.

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.



Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.

Il rapporto di correlazione

$$\eta_{Y|X}^2 = 1 - \frac{M_X[\text{Var}(Y|X)]}{\text{Var}(Y)}$$

può essere ottenuto considerando il complemento a 1 del rapporto tra la varianza residua $M_X[\text{Var}(Y|X)]$, media della variabile statistica varianze condizionate

x	$\text{Var}(Y x)$
1	0.0000
2	0.0000
5	0.6667

e la varianza della variabile dipendente $\text{Var}(Y) = 14$.

La varianza residua risulta $M_X[\text{Var}(Y|X)] = 0.4$.



Si ottiene

x	$Var(Y x)$	$n(x)$	$Var(Y x) \cdot n(x)$
1	0.0000	1	0
2	0.0000	1	0
5	0.6667	3	2
		5	2

da cui segue

$$M_X[Var(Y|X)] = \frac{2}{5} = 0.4.$$

Il rapporto di correlazione $\eta_{Y|X}^2$ risulta:

$$\eta_{Y|X}^2 = 1 - \frac{M_X[Var(Y|X)]}{Var(Y)} = 1 - \frac{0.4}{14} = 0.9714.$$

Sussiste, quindi, un livello di dipendenza molto alto.

Esercizio 4

Prova scritta di
STATISTICA del
12.02.2015

Giuseppe Boari
Gabriele Cantaluppi



UNIVERSITÀ
CATTOLICA
del Sacro Cuore

Un'urna contiene 5 palline nere e 15 palline rosse.

1. Si calcoli la probabilità che, estraendo **senza reimmissione** 5 palline dall'urna, si ottengano 3 palline nere.

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.

Esercizio 4



Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.

Un'urna contiene 5 palline nere e 15 palline rosse.

1. Si calcoli la probabilità che, estraendo **senza reimmissione** 5 palline dall'urna, si ottengano 3 palline nere.

Risulta

$$P(3N) = \binom{5}{3} \frac{5}{20} \frac{4}{19} \frac{3}{18} \frac{15}{17} \frac{14}{16} = 0.0677$$

oppure, indicata con X la variabile casuale 'numero di palline nere' nelle 5 estrazioni senza reimmissione, è possibile avvalersi della v.c. ipergeometrica

$$P(X = 3) = \frac{\binom{5}{3} \binom{15}{2}}{\binom{20}{5}} = 0.0677.$$



2. Indicata con Y la variabile casuale 'numero di palline nere' ottenute in 100 estrazioni **con reimmissione** dall'urna, si calcoli $P(Y \leq 6)$.

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.



2. Indicata con Y la variabile casuale 'numero di palline nere' ottenute in 100 estrazioni **con reimmissione** dall'urna, si calcoli $P(Y \leq 6)$.

Risulta

$$Y \sim$$

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.



2. Indicata con Y la variabile casuale 'numero di palline nere' ottenute in 100 estrazioni **con reimmissione** dall'urna, si calcoli $P(Y \leq 6)$.

Risulta

$$Y \sim \text{Bin}(n = 100, p = 0.25).$$

$$M(X) = 25; \text{Var}(X) = 18.75.$$

Dal momento che $np = 25 > 5$ ed $n(1 - p) = 75 > 5$ è possibile utilizzare l'approssimazione della variabile casuale Binomiale con la variabile casuale Normale

$$X \sim$$



2. Indicata con Y la variabile casuale 'numero di palline nere' ottenute in 100 estrazioni **con reimmissione** dall'urna, si calcoli $P(Y \leq 6)$.

Risulta

$$Y \sim \text{Bin}(n = 100, p = 0.25).$$

$$M(X) = 25; \text{Var}(X) = 18.75.$$

Dal momento che $np = 25 > 5$ ed $n(1 - p) = 75 > 5$ è possibile utilizzare l'approssimazione della variabile casuale Binomiale con la variabile casuale Normale

$$X \sim N(\mu = 25, \sigma^2 = 18.75).$$

Si ottiene

$$\begin{aligned} P(X \leq 6) &= P\left(\frac{X - \mu}{\sigma} \leq \frac{6 - \mu}{\sigma}\right) = P\left(Z \leq \frac{6 - 25}{4.3301}\right) = \\ &= P(Z \leq -4.39) \simeq 0. \end{aligned}$$



3. Sia ora $p = 0.8729$. Si determini il valore del percentile di ordine p della v.c. normale $X \sim N(\mu = 2, \sigma^2 = 1)$, vale a dire il valore x_p tale che $P(X \leq x_p) = p$.

$$P(X \leq x_p) = 0.8729$$

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.



Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.

3. Sia ora $p = 0.8729$. Si determini il valore del percentile di ordine p della v.c. normale $X \sim N(\mu = 2, \sigma^2 = 1)$, vale a dire il valore x_p tale che $P(X \leq x_p) = p$.

$$P(X \leq x_p) = 0.8729$$

$$P\left(\frac{X - \mu}{\sigma} \leq \frac{x_p - \mu}{\sigma}\right) = 0.8729$$



Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.

3. Sia ora $p = 0.8729$. Si determini il valore del percentile di ordine p della v.c. normale $X \sim N(\mu = 2, \sigma^2 = 1)$, vale a dire il valore x_p tale che $P(X \leq x_p) = p$.

$$P(X \leq x_p) = 0.8729$$

$$P\left(\frac{X - \mu}{\sigma} \leq \frac{x_p - \mu}{\sigma}\right) = 0.8729$$

$$P\left(Z \leq \frac{x_p - 2}{1}\right) = P(Z \leq z_p) = 0.8729$$

$$\frac{x_p - 2}{1} = z_p =$$



Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.

3. Sia ora $p = 0.8729$. Si determini il valore del percentile di ordine p della v.c. normale $X \sim N(\mu = 2, \sigma^2 = 1)$, vale a dire il valore x_p tale che $P(X \leq x_p) = p$.

$$P(X \leq x_p) = 0.8729$$

$$P\left(\frac{X - \mu}{\sigma} \leq \frac{x_p - \mu}{\sigma}\right) = 0.8729$$

$$P\left(Z \leq \frac{x_p - 2}{1}\right) = P(Z \leq z_p) = 0.8729$$

$$\frac{x_p - 2}{1} = z_p = 1.14$$



Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.

3. Sia ora $p = 0.8729$. Si determini il valore del percentile di ordine p della v.c. normale $X \sim N(\mu = 2, \sigma^2 = 1)$, vale a dire il valore x_p tale che $P(X \leq x_p) = p$.

$$P(X \leq x_p) = 0.8729$$

$$P\left(\frac{X - \mu}{\sigma} \leq \frac{x_p - \mu}{\sigma}\right) = 0.8729$$

$$P\left(Z \leq \frac{x_p - 2}{1}\right) = P(Z \leq z_p) = 0.8729$$

$$\frac{x_p - 2}{1} = z_p = 1.14$$

$$x_p =$$



Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.

3. Sia ora $p = 0.8729$. Si determini il valore del percentile di ordine p della v.c. normale $X \sim N(\mu = 2, \sigma^2 = 1)$, vale a dire il valore x_p tale che $P(X \leq x_p) = p$.

$$P(X \leq x_p) = 0.8729$$

$$P\left(\frac{X - \mu}{\sigma} \leq \frac{x_p - \mu}{\sigma}\right) = 0.8729$$

$$P\left(Z \leq \frac{x_p - 2}{1}\right) = P(Z \leq z_p) = 0.8729$$

$$\frac{x_p - 2}{1} = z_p = 1.14$$

$$x_p = 2 + 1.14 \cdot 1 = 3.14.$$

Tema T284-1

Prova scritta di STATISTICA del 04.06.2015

Prova scritta di
STATISTICA del
04.06.2015

Giuseppe Boari
Gabriele Cantaluppi



UNIVERSITÀ
CATTOLICA
del Sacro Cuore

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.

Giuseppe Boari
Gabriele Cantaluppi
Università Cattolica del Sacro Cuore, Milano



Si completino le frequenze nella seguente tabella sapendo che il numero totale di unità statistiche sulle quali sono stati congiuntamente rilevati i caratteri X e Y è 11 e che $n_{\bullet 2} = 1$.

$X \setminus Y$	10 + 15	15 + 25	25 + 30	$n_{i\bullet}$
x_1	1	0	9	10
x_2	0	1	0	1
$n_{\bullet j}$	1	1	9	11

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.



1. Si confrontino le distribuzioni condizionate $Y|x_1$ e $Y|x_2$ con opportuni indici di posizione e di variabilità.

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.



Si considera la distribuzione condizionata $Y|x_1$

y_j	n_{1j}
10 ÷ 15	1
15 ÷ 25	0
25 ÷ 30	9

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.



Si tratta di un carattere quantitativo con modalità raggruppate in classi.

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.



La media aritmetica

$$M(Y|x_1) = \frac{1}{n_{1\bullet}} \sum_{j=1}^k y_j n_{1j}$$

risulta 26.

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.



La media aritmetica

$$M(Y|x_1) = \frac{1}{n_{1\bullet}} \sum_{j=1}^k y_j n_{1j}$$

risulta 26.

Si ottiene (considerando come modalità i valori centrali delle classi)

y_j	n_{1j}	$y_j n_{1j}$
12.5	1	12.5
20.0	0	0.0
27.5	9	247.5
	10	260.0

da cui segue

$$M(Y|x_1) = \frac{1}{n_{1\bullet}} \sum_{j=1}^k y_j n_{1j} = \frac{260}{10} = 26.$$

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.



La varianza

risulta 20.25.

$$\text{Var}(Y|x_1) = M(Y^2|x_1) - [M(Y|x_1)]^2$$

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.



Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.

La varianza

$$\text{Var}(Y|x_1) = M(Y^2|x_1) - [M(Y|x_1)]^2$$

risulta 20.25. Si ottiene (considerando come modalità i valori centrali delle classi)

y_j	n_{1j}	$y_j n_{1j}$	$y_j^2 n_{1j}$
12.5	1	12.5	156.25
20.0	0	0.0	0.00
27.5	9	247.5	6806.25
	10	260.0	6962.50

da cui segue

$$\text{Var}(Y|x_1) = \frac{6962.5}{10} - 26^2 = 696.25 - 676 = 20.25.$$



Per effettuare confronti in termini di variabilità tra caratteri quantitativi occorre utilizzare una misura relativa, ad esempio il coefficiente di variazione

$$CV = \frac{\sigma}{\mu} = \frac{\sqrt{\text{Var}(Y|x_1)}}{M(Y|x_1)} = \frac{\sqrt{20.25}}{26} = \frac{4.5}{26} = 0.1731.$$

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.



Si considera la distribuzione condizionata $Y|x_2$

y_j	n_{2j}
10 $\rightarrow$ 15	0
15 $\rightarrow$ 25	1
25 $\rightarrow$ 30	0

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.



Si tratta di un carattere quantitativo con modalità raggruppate in classi.

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.



Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.

Si osserva che la distribuzione è degenere.

$M(Y|x_2) = 20$. La varianza e il coefficiente di variazione sono nulli.



Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.

$Y|x_1$ Indice di posizione: $M(Y|x_1) = 26$ Indice di variabilità: $CV(Y|x_1) = 0.1731$

$Y|x_2$ Indice di posizione: $M(Y|x_2) = 20$ Indice di variabilità: $CV(Y|x_2) = 0$

La distribuzione $Y|x_1$ presenta un livello medio superiore a quello della distribuzione $Y|x_2$. La distribuzione condizionata $Y|x_2$ presenta variabilità nulla.



2. Si rappresenti graficamente la marginale X .

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.



Si considera la distribuzione marginale della variabile X , ottenuta sommando le frequenze di riga

x_i	$n_{i\bullet}$
x_1	10
x_2	1

Si tratta di un carattere qualitativo sconnesso che può essere rappresentato graficamente mediante un grafico a bastoncini o a canne d'organo.

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.

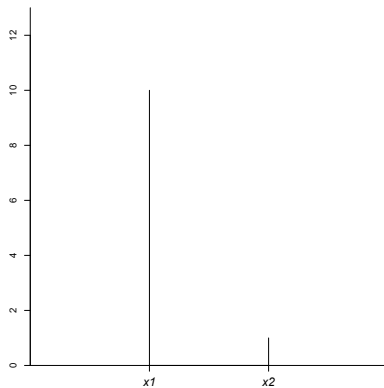


Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.





3. Si riassume la marginale qualitativa X con opportuni indici di posizione e di variabilità.

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.



Si considera la distribuzione marginale della variabile X , ottenuta sommando le frequenze di riga

x_i	$n_{i\bullet}$
x_1	10
x_2	1

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.



Si tratta di un carattere qualitativo sconnesso.

L'unico indice di posizione che può essere definito è la moda (categoria alla quale è associata la massima frequenza) che risulta x_1 .

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.



In presenza di un carattere qualitativo la variabilità può essere misurata mediante gli indici di Gini e di Frosini normalizzati

$$G_N = \frac{1 - \sum_{i=1}^h f_{i\bullet}^2}{1 - \frac{1}{h}}$$

e

$$F_N = 1 - \sqrt{1 - G_N}.$$

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.



In presenza di un carattere qualitativo la variabilità può essere misurata mediante gli indici di Gini e di Frosini normalizzati

$$G_N = \frac{1 - \sum_{i=1}^h f_{i\bullet}^2}{1 - \frac{1}{h}}$$

e

$$F_N = 1 - \sqrt{1 - G_N}.$$

Si ottiene:

x_i	$n_{i\bullet}$	$f_{i\bullet}$	$f_{i\bullet}^2$
x_1	10	0.9091	0.8264
x_2	1	0.0909	0.0083
	11	1	0.8347

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.



In presenza di un carattere qualitativo la variabilità può essere misurata mediante gli indici di Gini e di Frosini normalizzati

$$G_N = \frac{1 - \sum_{i=1}^h f_{i\bullet}^2}{1 - \frac{1}{h}}$$

e

$$F_N = 1 - \sqrt{1 - G_N}.$$

Si ottiene:

x_i	$n_{i\bullet}$	$f_{i\bullet}$	$f_{i\bullet}^2$
x_1	10	0.9091	0.8264
x_2	1	0.0909	0.0083
	11	1	0.8347

Da cui segue:

$$G_N = \frac{1 - 0.8347}{1 - \frac{1}{2}} = 0.3306$$

e

$$F_N = 1 - \sqrt{1 - 0.3306} = 0.1818.$$

Sussiste, quindi, un livello molto basso di eterogeneità.



Il prospetto seguente riporta i prezzi di chiusura di 2 titoli azionari nei giorni dal 7 al 10 gennaio:

t	<i>titolo A</i>	<i>titolo B</i>
7	100	100
8	120	100
9	144	110
10	150	121

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.



1. Si calcoli e interpreti il numero indice al tempo 10 con base 7 per il *titolo A* e si costruisca la serie dei numeri indici a base fissa (al tempo 7) per il *titolo B*.

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.



1. Si calcoli e interpreti il numero indice al tempo 10 con base 7 per il *titolo A* e si costruisca la serie dei numeri indici a base fissa (al tempo 7) per il *titolo B*.

$${}_7I_{10} = \frac{x_{10}}{x_7} = \frac{150}{100} = 1.5$$

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.



1. Si calcoli e interpreti il numero indice al tempo 10 con base 7 per il *titolo A* e si costruisca la serie dei numeri indici a base fissa (al tempo 7) per il *titolo B*.

$${}_7I_{10} = \frac{x_{10}}{x_7} = \frac{150}{100} = 1.5$$

Tra il tempo 7 e il tempo 10 si è registrato un incremento complessivo del 50% della grandezza oggetto di studio.

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.



1. Si calcoli e interpreti il numero indice al tempo 10 con base 7 per il *titolo A* e si costruisca la serie dei numeri indici a base fissa (al tempo 7) per il *titolo B*.

$${}_7I_{10} = \frac{x_{10}}{x_7} = \frac{150}{100} = 1.5$$

Tra il tempo 7 e il tempo 10 si è registrato un incremento complessivo del 50% della grandezza oggetto di studio.

Si ricorda che ${}_7I_t = \frac{x_t}{x_7}$.

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.



1. Si calcoli e interpreti il numero indice al tempo 10 con base 7 per il *titolo A* e si costruisca la serie dei numeri indici a base fissa (al tempo 7) per il *titolo B*.

$${}_7I_{10} = \frac{x_{10}}{x_7} = \frac{150}{100} = 1.5$$

Tra il tempo 7 e il tempo 10 si è registrato un incremento complessivo del 50% della grandezza oggetto di studio.

Si ricorda che ${}_7I_t = \frac{x_t}{x_7}$. Si ottiene quindi

<i>giorno</i>	x_t	<i>NIBF (7)</i>
7	100	1.00
8	100	1.00
9	110	1.10
10	121	1.21

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.



2. Si calcoli il tasso medio di variazione registrato tra il giorno 7 e il giorno 10 per il *titolo B*, interpretando il risultato.

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.



2. Si calcoli il tasso medio di variazione registrato tra il giorno 7 e il giorno 10 per il *titolo B*, interpretando il risultato.

Si ricorda che

$${}_7I_{10} = {}_7I_8 \cdot {}_8I_9 \cdot {}_9I_{10}$$

Si deve trovare il numero indice a base mobile medio.

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.



Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.

2. Si calcoli il tasso medio di variazione registrato tra il giorno 7 e il giorno 10 per il *titolo B*, interpretando il risultato.

Si ricorda che

$${}_7I_{10} = {}_7I_8 \cdot {}_8I_9 \cdot {}_9I_{10}$$

Si deve trovare il numero indice a base mobile medio. Applicando il criterio di scelta della media secondo Chisini si ottiene

$${}_7I_{10} = {}_7I_8 \cdot {}_8I_9 \cdot {}_9I_{10} = \alpha \cdot \alpha \cdot \alpha = \alpha^3.$$



Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.

2. Si calcoli il tasso medio di variazione registrato tra il giorno 7 e il giorno 10 per il *titolo B*, interpretando il risultato.

Si ricorda che

$${}_7I_{10} = {}_7I_8 \cdot {}_8I_9 \cdot {}_9I_{10}$$

Si deve trovare il numero indice a base mobile medio. Applicando il criterio di scelta della media secondo Chisini si ottiene

$${}_7I_{10} = {}_7I_8 \cdot {}_8I_9 \cdot {}_9I_{10} = \alpha \cdot \alpha \cdot \alpha = \alpha^3.$$

Quindi

$$\alpha = \sqrt[3]{{}_7I_{10}} = \sqrt[3]{{}_7I_8 \cdot {}_8I_9 \cdot {}_9I_{10}} = 1.0656.$$

Da cui segue una variazione relativa media pari a 0.0656, corrispondente a un incremento medio giornaliero tra il giorno 7 e il giorno 10 del 6.56%.



3. Supponendo che la mediana calcolata a partire dai valori degli indici di Laspeyres, $\frac{p}{7}I_{10}^L$, e di Paasche, $\frac{p}{7}I_{10}^P$, sia risultata pari a 1.31, si indichi l'intervallo dei valori possibili per l'indice di Fisher.

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.



Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.

3. Supponendo che la mediana calcolata a partire dai valori degli indici di Laspeyres, ${}_7^P I_{10}^L$, e di Paasche, ${}_7^P I_{10}^P$, sia risultata pari a 1.31, si indichi l'intervallo dei valori possibili per l'indice di Fisher.

Gli indici dei prezzi di Laspeyres e di Paasche sono medie potenziate dei prezzi relativi; saranno entrambi compresi tra il prezzo relativo minimo $1.21 = \frac{121}{100}$ e quello massimo $1.5 = \frac{150}{100}$.



Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.

3. Supponendo che la mediana calcolata a partire dai valori degli indici di Laspeyres, ${}_7^P I_{10}^L$, e di Paasche, ${}_7^P I_{10}^P$, sia risultata pari a 1.31, si indichi l'intervallo dei valori possibili per l'indice di Fisher.

Gli indici dei prezzi di Laspeyres e di Paasche sono medie potenziate dei prezzi relativi; saranno entrambi compresi tra il prezzo relativo minimo $1.21 = \frac{121}{100}$ e quello massimo $1.5 = \frac{150}{100}$.

L'indice dei prezzi di Fisher è definito come media geometrica degli indici dei prezzi di Laspeyres e di Paasche.



Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.

3. Supponendo che la mediana calcolata a partire dai valori degli indici di Laspeyres, $\frac{p}{7}I_{10}^L$, e di Paasche, $\frac{p}{7}I_{10}^P$, sia risultata pari a 1.31, si indichi l'intervallo dei valori possibili per l'indice di Fisher.

Gli indici dei prezzi di Laspeyres e di Paasche sono medie potenziate dei prezzi relativi; saranno entrambi compresi tra il prezzo relativo minimo $1.21 = \frac{121}{100}$ e quello massimo $1.5 = \frac{150}{100}$.

L'indice dei prezzi di Fisher è definito come media geometrica degli indici dei prezzi di Laspeyres e di Paasche.

La media geometrica dei due indici risulterà inferiore o eguale alla media aritmetica degli stessi che coincide con la mediana; l'eguaglianza si può realizzare solo nel caso in cui l'indice di Laspeyres coincida con quello di Paasche.



3. Supponendo che la mediana calcolata a partire dai valori degli indici di Laspeyres, $\frac{p}{7}I_{10}^L$, e di Paasche, $\frac{p}{7}I_{10}^P$, sia risultata pari a 1.31, si indichi l'intervallo dei valori possibili per l'indice di Fisher.

Gli indici dei prezzi di Laspeyres e di Paasche sono medie potenziate dei prezzi relativi; saranno entrambi compresi tra il prezzo relativo minimo $1.21 = \frac{121}{100}$ e quello massimo $1.5 = \frac{150}{100}$.

L'indice dei prezzi di Fisher è definito come media geometrica degli indici dei prezzi di Laspeyres e di Paasche.

La media geometrica dei due indici risulterà inferiore o eguale alla media aritmetica degli stessi che coincide con la mediana; l'eguaglianza si può realizzare solo nel caso in cui l'indice di Laspeyres coincida con quello di Paasche.

Vale, quindi

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.



3. Supponendo che la mediana calcolata a partire dai valori degli indici di Laspeyres, $\frac{p}{7}I_{10}^L$, e di Paasche, $\frac{p}{7}I_{10}^P$, sia risultata pari a 1.31, si indichi l'intervallo dei valori possibili per l'indice di Fisher.

Gli indici dei prezzi di Laspeyres e di Paasche sono medie potenziate dei prezzi relativi; saranno entrambi compresi tra il prezzo relativo minimo $1.21 = \frac{121}{100}$ e quello massimo $1.5 = \frac{150}{100}$.

L'indice dei prezzi di Fisher è definito come media geometrica degli indici dei prezzi di Laspeyres e di Paasche.

La media geometrica dei due indici risulterà inferiore o eguale alla media aritmetica degli stessi che coincide con la mediana; l'eguaglianza si può realizzare solo nel caso in cui l'indice di Laspeyres coincida con quello di Paasche.

Vale, quindi

$$1.21 < \frac{p}{7}I_{10}^F \leq 1.31.$$



Il seguente prospetto riporta i valori rilevati in corrispondenza a 5 unità statistiche con riferimento a due caratteri quantitativi X e Y

x_i	1	2	3	4	5
y_i	22	20	24	32	34

1. Si rappresentino graficamente le coppie dei dati ipotizzando Y in funzione di X .

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.



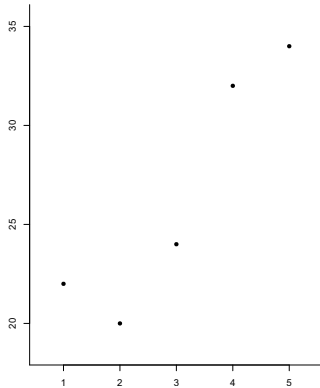
Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.

Le coppie di dati possono essere rappresentate mediante un grafico a dispersione.





2. Si calcolino, secondo il metodo dei minimi quadrati, i parametri e l'indice di adattamento dei seguenti modelli di regressione, individuando il modello migliore:

$$I) Y^* = a + bX \quad II) Y^* = cX$$

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.

Si riportano i calcoli necessari¹ per ottenere le stime dei parametri e gli indici di adattamento del modello

$$Y = a + bX + E$$

x_i	y_i	$x_i y_i$	x_i^2	y_i^2
1	22	22	1	484
2	20	40	4	400
3	24	72	9	576
4	32	128	16	1024
5	34	170	25	1156
15	132	432	55	3640
3	26.4	86.4	11	728
$M(X)$	$M(Y)$	$M(XY)$	$M(X^2)$	$M(Y^2)$

$$\hat{b} =$$



Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.

¹Si osserva che per ragioni di stampa alcuni calcoli potrebbero essere duplicati con riferimento ai diversi modelli oggetto di studio

Si riportano i calcoli necessari¹ per ottenere le stime dei parametri e gli indici di adattamento del modello

$$Y = a + bX + E$$

x_i	y_i	$x_i y_i$	x_i^2	y_i^2
1	22	22	1	484
2	20	40	4	400
3	24	72	9	576
4	32	128	16	1024
5	34	170	25	1156
15	132	432	55	3640
3	26.4	86.4	11	728
$M(X)$	$M(Y)$	$M(XY)$	$M(X^2)$	$M(Y^2)$

$$\begin{aligned}\hat{b} &= \frac{\text{Cov}(X, Y)}{\text{Var}(X)} = \frac{M(XY) - M(X)M(Y)}{M(X^2) - [M(X)]^2} = \\ &= \frac{86.4 - 3 \cdot 26.4}{11 - 3^2} = \frac{7.2}{2} = 3.6 \\ \hat{a} &= \end{aligned}$$

¹Si osserva che per ragioni di stampa alcuni calcoli potrebbero essere duplicati con riferimento ai diversi modelli oggetto di studio



Si riportano i calcoli necessari¹ per ottenere le stime dei parametri e gli indici di adattamento del modello

$$Y = a + bX + E$$

x_i	y_i	$x_i y_i$	x_i^2	y_i^2
1	22	22	1	484
2	20	40	4	400
3	24	72	9	576
4	32	128	16	1024
5	34	170	25	1156
15	132	432	55	3640
3	26.4	86.4	11	728
$M(X)$	$M(Y)$	$M(XY)$	$M(X^2)$	$M(Y^2)$

$$\begin{aligned}\hat{b} &= \frac{\text{Cov}(X, Y)}{\text{Var}(X)} = \frac{M(XY) - M(X)M(Y)}{M(X^2) - [M(X)]^2} = \\ &= \frac{86.4 - 3 \cdot 26.4}{11 - 3^2} = \frac{7.2}{2} = 3.6 \\ \hat{a} &= M(Y) - \hat{b}M(X) = 15.6 \\ \rho^2 &= \end{aligned}$$

¹Si osserva che per ragioni di stampa alcuni calcoli potrebbero essere duplicati con riferimento ai diversi modelli oggetto di studio



Si riportano i calcoli necessari¹ per ottenere le stime dei parametri e gli indici di adattamento del modello

$$Y = a + bX + E$$

x_i	y_i	$x_i y_i$	x_i^2	y_i^2
1	22	22	1	484
2	20	40	4	400
3	24	72	9	576
4	32	128	16	1024
5	34	170	25	1156
15	132	432	55	3640
3	26.4	86.4	11	728
$M(X)$	$M(Y)$	$M(XY)$	$M(X^2)$	$M(Y^2)$

$$\begin{aligned}\hat{b} &= \frac{\text{Cov}(X, Y)}{\text{Var}(X)} = \frac{M(XY) - M(X)M(Y)}{M(X^2) - [M(X)]^2} = \\ &= \frac{86.4 - 3 \cdot 26.4}{11 - 3^2} = \frac{7.2}{2} = 3.6 \\ \hat{a} &= M(Y) - \hat{b}M(X) = 15.6 \\ \rho^2 &= \frac{[\text{Cov}(X, Y)]^2}{\text{Var}(X) \text{Var}(Y)} = \frac{7.2^2}{2 \cdot 31.04} = 0.8351\end{aligned}$$

$$M(\hat{E}^2) = \text{Var}(\hat{E}) = \text{varianza residua} =$$

¹Si osserva che per ragioni di stampa alcuni calcoli potrebbero essere duplicati con riferimento ai diversi modelli oggetto di studio



Si riportano i calcoli necessari¹ per ottenere le stime dei parametri e gli indici di adattamento del modello

$$Y = a + bX + E$$

x_i	y_i	$x_i y_i$	x_i^2	y_i^2
1	22	22	1	484
2	20	40	4	400
3	24	72	9	576
4	32	128	16	1024
5	34	170	25	1156
15	132	432	55	3640
3	26.4	86.4	11	728
$M(X)$	$M(Y)$	$M(XY)$	$M(X^2)$	$M(Y^2)$

$$\begin{aligned}\hat{b} &= \frac{\text{Cov}(X, Y)}{\text{Var}(X)} = \frac{M(XY) - M(X)M(Y)}{M(X^2) - [M(X)]^2} = \\ &= \frac{86.4 - 3 \cdot 26.4}{11 - 3^2} = \frac{7.2}{2} = 3.6 \\ \hat{a} &= M(Y) - \hat{b}M(X) = 15.6 \\ \rho^2 &= \frac{[\text{Cov}(X, Y)]^2}{\text{Var}(X) \text{Var}(Y)} = \frac{7.2^2}{2 \cdot 31.04} = 0.8351\end{aligned}$$

$$M(\hat{E}^2) = \text{Var}(\hat{E}) = \text{varianza residua} = \text{Var}(Y) \cdot (1 - \rho^2) = 5.12$$

¹Si osserva che per ragioni di stampa alcuni calcoli potrebbero essere duplicati con riferimento ai diversi modelli oggetto di studio





Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.

Si riportano i calcoli necessari² per ottenere le stime dei parametri e la misura di adattamento del modello

$$Y = cX + E$$

x_i	y_i	$x_i y_i$	x_i^2	y_i^2
1	22	22	1	484
2	20	40	4	400
3	24	72	9	576
4	32	128	16	1024
5	34	170	25	1156
		432	55	3640
		86.4	11	728
		$M(XY)$	$M(X^2)$	$M(Y^2)$

$$\hat{c} =$$

²Si osserva che per ragioni di stampa alcuni calcoli potrebbero essere duplicati con riferimento ai diversi modelli oggetto di studio



Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.

Si riportano i calcoli necessari² per ottenere le stime dei parametri e la misura di adattamento del modello

$$Y = cX + E$$

x_i	y_i	$x_i y_i$	x_i^2	y_i^2
1	22	22	1	484
2	20	40	4	400
3	24	72	9	576
4	32	128	16	1024
5	34	170	25	1156
		432	55	3640
		86.4	11	728
		$M(XY)$	$M(X^2)$	$M(Y^2)$

$$\hat{c} = \frac{M(XY)}{M(X^2)} = \frac{86.4}{11} = 7.8545$$

$$M(\hat{E}^2) =$$

²Si osserva che per ragioni di stampa alcuni calcoli potrebbero essere duplicati con riferimento ai diversi modelli oggetto di studio



Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.

Si riportano i calcoli necessari² per ottenere le stime dei parametri e la misura di adattamento del modello

$$Y = cX + E$$

x_i	y_i	$x_i y_i$	x_i^2	y_i^2
1	22	22	1	484
2	20	40	4	400
3	24	72	9	576
4	32	128	16	1024
5	34	170	25	1156
		432	55	3640
		86.4	11	728
		$M(XY)$	$M(X^2)$	$M(Y^2)$

$$\hat{c} = \frac{M(XY)}{M(X^2)} = \frac{86.4}{11} = 7.8545$$

$$M(\hat{E}^2) = M[Y \cdot (Y - \hat{c}X)] = M(Y^2) - \hat{c}M(XY) = 728 - 7.8545 \cdot 86.4 = 49.3673$$

²Si osserva che per ragioni di stampa alcuni calcoli potrebbero essere duplicati con riferimento ai diversi modelli oggetto di studio



$a = 15.6$, $b = 3.6$, indice adattamento $M(\hat{E}) = 5.12$ $c = 7.8545$, indice adattamento

$M(\hat{E}^2) = 49.367$

Il modello migliore è $Y^* = 15.6 + 3.6X$.

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.



3. Per il modello migliore, si fornisca la previsione in corrispondenza di $X = 3$.
 $Y^*(3) = 15.6 + 3.6 \cdot 3 = 26.4$.

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.



Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.

4. Si determini il valore del rapporto di correlazione $\eta_{Y|X}^2$ motivando la risposta.
 $\eta_{Y|X}^2 = 1$ in quanto sussiste dipendenza funzionale di Y da X .

Esercizio 4

Si consideri l'esperimento di estrazione di 4 palline da un'urna contenente 21 palline rosse e 29 palline blu.

1. Si calcoli la probabilità di ottenere almeno 3 palline blu nell'ipotesi che le 4 estrazioni siano effettuate senza reimmissione.



Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.

Esercizio 4

Si consideri l'esperimento di estrazione di 4 palline da un'urna contenente 21 palline rosse e 29 palline blu.

1. Si calcoli la probabilità di ottenere almeno 3 palline blu nell'ipotesi che le 4 estrazioni siano effettuate senza reimmissione.

Risulta

$$P(\text{almeno } 3 B) =$$



Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.

Esercizio 4



Si consideri l'esperimento di estrazione di 4 palline da un'urna contenente 21 palline rosse e 29 palline blu.

1. Si calcoli la probabilità di ottenere almeno 3 palline blu nell'ipotesi che le 4 estrazioni siano effettuate senza reimmissione.

Risulta

$$P(\text{almeno } 3 B) = P(3B) + P(4B) =$$

=

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.

Esercizio 4



Si consideri l'esperimento di estrazione di 4 palline da un'urna contenente 21 palline rosse e 29 palline blu.

1. Si calcoli la probabilità di ottenere almeno 3 palline blu nell'ipotesi che le 4 estrazioni siano effettuate senza reimmissione.

Risulta

$$\begin{aligned} P(\text{almeno } 3 B) &= P(3B) + P(4B) = \\ &= \binom{4}{3} \frac{29}{50} \frac{28}{49} \frac{27}{48} \frac{21}{47} + \binom{4}{4} \frac{29}{50} \frac{28}{49} \frac{27}{48} \frac{26}{47} = \\ &= \end{aligned}$$

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.

Esercizio 4



Si consideri l'esperimento di estrazione di 4 palline da un'urna contenente 21 palline rosse e 29 palline blu.

1. Si calcoli la probabilità di ottenere almeno 3 palline blu nell'ipotesi che le 4 estrazioni siano effettuate senza reimmissione.

Risulta

$$\begin{aligned} P(\text{almeno } 3 B) &= P(3B) + P(4B) = \\ &= \binom{4}{3} \frac{29}{50} \frac{28}{49} \frac{27}{48} \frac{21}{47} + \binom{4}{4} \frac{29}{50} \frac{28}{49} \frac{27}{48} \frac{26}{47} = \\ &= 0.3332 + 0.1031 = \end{aligned}$$

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.

Esercizio 4



Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.

Si consideri l'esperimento di estrazione di 4 palline da un'urna contenente 21 palline rosse e 29 palline blu.

1. Si calcoli la probabilità di ottenere almeno 3 palline blu nell'ipotesi che le 4 estrazioni siano effettuate senza reimmissione.

Risulta

$$\begin{aligned} P(\text{almeno } 3 B) &= P(3B) + P(4B) = \\ &= \binom{4}{3} \frac{29}{50} \frac{28}{49} \frac{27}{48} \frac{21}{47} + \binom{4}{4} \frac{29}{50} \frac{28}{49} \frac{27}{48} \frac{26}{47} = \\ &= 0.3332 + 0.1031 = 0.4363 \end{aligned}$$

o mediante la v.c. ipergeometrica

Esercizio 4



Si consideri l'esperimento di estrazione di 4 palline da un'urna contenente 21 palline rosse e 29 palline blu.

1. Si calcoli la probabilità di ottenere almeno 3 palline blu nell'ipotesi che le 4 estrazioni siano effettuate senza reimmissione.

Risulta

$$\begin{aligned}P(\text{almeno } 3 B) &= P(3B) + P(4B) = \\&= \binom{4}{3} \frac{29}{50} \frac{28}{49} \frac{27}{48} \frac{21}{47} + \binom{4}{4} \frac{29}{50} \frac{28}{49} \frac{27}{48} \frac{26}{47} = \\&= 0.3332 + 0.1031 = 0.4363\end{aligned}$$

o mediante la v.c. ipergeometrica

$$P(X = 3) + P(X = 4) = \frac{\binom{29}{3} \binom{21}{1}}{\binom{50}{4}} + \frac{\binom{29}{4} \binom{21}{0}}{\binom{50}{4}} = 0.4363$$

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.



2. Si calcoli la probabilità di ottenere almeno 3 palline blu nell'ipotesi che le 4 estrazioni siano effettuate con reimmissione.

Sia X la variabile casuale 'n. di palline blu nelle 4 estrazioni (con reimmissione) (prove indipendenti)'. $X \sim \text{Bin}(n = 4, p = 0.58)$.

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.



Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.

2. Si calcoli la probabilità di ottenere almeno 3 palline blu nell'ipotesi che le 4 estrazioni siano effettuate con reimmissione.

Sia X la variabile casuale 'n. di palline blu nelle 4 estrazioni (con reimmissione) (prove indipendenti)'. $X \sim \text{Bin}(n = 4, p = 0.58)$.

$$P(X \geq 3) = \binom{4}{3} 0.58^3 0.42^1 + \binom{4}{4} 0.58^4 0.42^0 = 0.3278 + 0.1132 = 0.4410.$$



3. Si calcoli la probabilità di ottenere almeno 36 palline blu nell'ipotesi che siano effettuate 60 estrazioni con reimmissione.

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.



3. Si calcoli la probabilità di ottenere almeno 36 palline blu nell'ipotesi che siano effettuate 60 estrazioni con reimmissione.

Indicata con X la variabile casuale 'numero di successi' nelle 60 prove indipendenti risulta

$$X \sim$$

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.



3. Si calcoli la probabilità di ottenere almeno 36 palline blu nell'ipotesi che siano effettuate 60 estrazioni con reimmissione.

Indicata con X la variabile casuale 'numero di successi' nelle 60 prove indipendenti risulta

$$X \sim \text{Bin}(n = 60, p = 0.58).$$

$$M(X) = 34.8; \text{Var}(X) = 14.6160.$$

Dal momento che $np = 34.8 > 5$ ed $n(1 - p) = 25.2 > 5$ è possibile utilizzare l'approssimazione della variabile casuale Binomiale con la variabile casuale Normale

$$X \sim$$



3. Si calcoli la probabilità di ottenere almeno 36 palline blu nell'ipotesi che siano effettuate 60 estrazioni con reimmissione.

Indicata con X la variabile casuale 'numero di successi' nelle 60 prove indipendenti risulta

$$X \sim \text{Bin}(n = 60, p = 0.58).$$

$$M(X) = 34.8; \text{Var}(X) = 14.6160.$$

Dal momento che $np = 34.8 > 5$ ed $n(1 - p) = 25.2 > 5$ è possibile utilizzare l'approssimazione della variabile casuale Binomiale con la variabile casuale Normale

$$X \sim N(\mu = 34.8, \sigma^2 = 14.6160).$$

Si ottiene



3. Si calcoli la probabilità di ottenere almeno 36 palline blu nell'ipotesi che siano effettuate 60 estrazioni con reimmissione.

Indicata con X la variabile casuale 'numero di successi' nelle 60 prove indipendenti risulta

$$X \sim \text{Bin}(n = 60, p = 0.58).$$

$$M(X) = 34.8; \text{Var}(X) = 14.6160.$$

Dal momento che $np = 34.8 > 5$ ed $n(1 - p) = 25.2 > 5$ è possibile utilizzare l'approssimazione della variabile casuale Binomiale con la variabile casuale Normale

$$X \sim N(\mu = 34.8, \sigma^2 = 14.6160).$$

Si ottiene

$$P(X > 36) =$$



3. Si calcoli la probabilità di ottenere almeno 36 palline blu nell'ipotesi che siano effettuate 60 estrazioni con reimmissione.

Indicata con X la variabile casuale 'numero di successi' nelle 60 prove indipendenti risulta

$$X \sim \text{Bin}(n = 60, p = 0.58).$$

$$M(X) = 34.8; \text{Var}(X) = 14.6160.$$

Dal momento che $np = 34.8 > 5$ ed $n(1 - p) = 25.2 > 5$ è possibile utilizzare l'approssimazione della variabile casuale Binomiale con la variabile casuale Normale

$$X \sim N(\mu = 34.8, \sigma^2 = 14.6160).$$

Si ottiene

$$P(X > 36) = P\left(\frac{X - \mu}{\sigma} > \frac{36 - \mu}{\sigma}\right) =$$



3. Si calcoli la probabilità di ottenere almeno 36 palline blu nell'ipotesi che siano effettuate 60 estrazioni con reimmissione.

Indicata con X la variabile casuale 'numero di successi' nelle 60 prove indipendenti risulta

$$X \sim \text{Bin}(n = 60, p = 0.58).$$

$$M(X) = 34.8; \text{Var}(X) = 14.6160.$$

Dal momento che $np = 34.8 > 5$ ed $n(1 - p) = 25.2 > 5$ è possibile utilizzare l'approssimazione della variabile casuale Binomiale con la variabile casuale Normale

$$X \sim N(\mu = 34.8, \sigma^2 = 14.6160).$$

Si ottiene

$$P(X > 36) = P\left(\frac{X - \mu}{\sigma} > \frac{36 - \mu}{\sigma}\right) = P\left(Z > \frac{36 - 34.8}{3.8231}\right) =$$

=



3. Si calcoli la probabilità di ottenere almeno 36 palline blu nell'ipotesi che siano effettuate 60 estrazioni con reimmissione.

Indicata con X la variabile casuale 'numero di successi' nelle 60 prove indipendenti risulta

$$X \sim \text{Bin}(n = 60, p = 0.58).$$

$$M(X) = 34.8; \text{Var}(X) = 14.6160.$$

Dal momento che $np = 34.8 > 5$ ed $n(1 - p) = 25.2 > 5$ è possibile utilizzare l'approssimazione della variabile casuale Binomiale con la variabile casuale Normale

$$X \sim N(\mu = 34.8, \sigma^2 = 14.6160).$$

Si ottiene

$$\begin{aligned} P(X > 36) &= P\left(\frac{X - \mu}{\sigma} > \frac{36 - \mu}{\sigma}\right) = P\left(Z > \frac{36 - 34.8}{3.8231}\right) = \\ &= P(Z > 0.31) = \end{aligned}$$



3. Si calcoli la probabilità di ottenere almeno 36 palline blu nell'ipotesi che siano effettuate 60 estrazioni con reimmissione.

Indicata con X la variabile casuale 'numero di successi' nelle 60 prove indipendenti risulta

$$X \sim \text{Bin}(n = 60, p = 0.58).$$

$$M(X) = 34.8; \text{Var}(X) = 14.6160.$$

Dal momento che $np = 34.8 > 5$ ed $n(1 - p) = 25.2 > 5$ è possibile utilizzare l'approssimazione della variabile casuale Binomiale con la variabile casuale Normale

$$X \sim N(\mu = 34.8, \sigma^2 = 14.6160).$$

Si ottiene

$$\begin{aligned} P(X > 36) &= P\left(\frac{X - \mu}{\sigma} > \frac{36 - \mu}{\sigma}\right) = P\left(Z > \frac{36 - 34.8}{3.8231}\right) = \\ &= P(Z > 0.31) = 1 - P(Z \leq 0.31) = \end{aligned}$$



3. Si calcoli la probabilità di ottenere almeno 36 palline blu nell'ipotesi che siano effettuate 60 estrazioni con reimmissione.

Indicata con X la variabile casuale 'numero di successi' nelle 60 prove indipendenti risulta

$$X \sim \text{Bin}(n = 60, p = 0.58).$$

$$M(X) = 34.8; \text{Var}(X) = 14.6160.$$

Dal momento che $np = 34.8 > 5$ ed $n(1 - p) = 25.2 > 5$ è possibile utilizzare l'approssimazione della variabile casuale Binomiale con la variabile casuale Normale

$$X \sim N(\mu = 34.8, \sigma^2 = 14.6160).$$

Si ottiene

$$\begin{aligned} P(X > 36) &= P\left(\frac{X - \mu}{\sigma} > \frac{36 - \mu}{\sigma}\right) = P\left(Z > \frac{36 - 34.8}{3.8231}\right) = \\ &= P(Z > 0.31) = 1 - P(Z \leq 0.31) = 1 - 0.6217 = \end{aligned}$$



3. Si calcoli la probabilità di ottenere almeno 36 palline blu nell'ipotesi che siano effettuate 60 estrazioni con reimmissione.

Indicata con X la variabile casuale 'numero di successi' nelle 60 prove indipendenti risulta

$$X \sim \text{Bin}(n = 60, p = 0.58).$$

$$M(X) = 34.8; \text{Var}(X) = 14.6160.$$

Dal momento che $np = 34.8 > 5$ ed $n(1 - p) = 25.2 > 5$ è possibile utilizzare l'approssimazione della variabile casuale Binomiale con la variabile casuale Normale

$$X \sim N(\mu = 34.8, \sigma^2 = 14.6160).$$

Si ottiene

$$\begin{aligned} P(X > 36) &= P\left(\frac{X - \mu}{\sigma} > \frac{36 - \mu}{\sigma}\right) = P\left(Z > \frac{36 - 34.8}{3.8231}\right) = \\ &= P(Z > 0.31) = 1 - P(Z \leq 0.31) = 1 - 0.6217 = 0.3783. \end{aligned}$$

Tema T284

Prova scritta di STATISTICA del 25.06.2015

Prova scritta di
STATISTICA del
25.06.2015

Giuseppe Boari
Gabriele Cantaluppi



UNIVERSITÀ
CATTOLICA
del Sacro Cuore

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.

Giuseppe Boari
Gabriele Cantaluppi
Università Cattolica del Sacro Cuore, Milano



La Direzione di un sito destinato alla vendita online di libri ha estratto un campione di opere attualmente in vendita, rilevando le variabili G = giudizio dei clienti sull'opera (b = basso, m = medio, a = alto) e P = prezzo di vendita (in euro). I dati raccolti sono riportati nella seguente tabella a doppia entrata:

$P \setminus G$	b	m	a	
5 - 15	1	0	2	
15 - 30	3	7	8	

1. Si rappresentino graficamente le distribuzioni marginali P e G .

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.



Si considera la distribuzione marginale della variabile P , ottenuta sommando le frequenze di riga

p_i	$n_{i\bullet}$
5 $\rightarrow$ 15	3
15 $\rightarrow$ 30	18

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.



Si tratta di un carattere quantitativo con modalità raggruppate in classi che può essere rappresentato graficamente mediante un istogramma.

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.



L'istogramma può essere costruito dopo avere calcolato le densità $d_{i\bullet} = \frac{n_{i\bullet}}{a_i}$, (dove a_i sono le ampiezze di classe).

p_i	$n_{i\bullet}$	$a_{i\bullet}$	$d_{i\bullet}$
5 - 15	3	15 - 5 = 10	0.3
15 - 30	18	30 - 15 = 15	1.2

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.

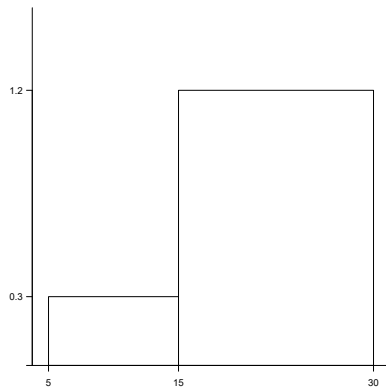


Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.





Si considera la distribuzione marginale della variabile G , ottenuta sommando le frequenze di colonna

g_j	$n_{\bullet j}$
b	4
m	7
a	10

Si tratta di un carattere qualitativo ordinato che può essere rappresentato graficamente mediante un grafico a bastoncini o a canne d'organo.

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.

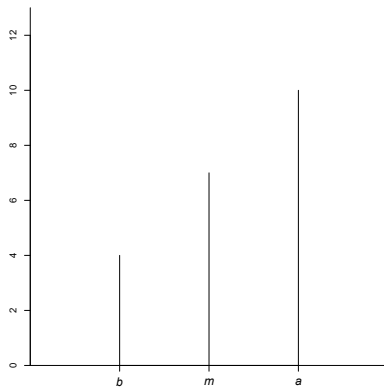


Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.





2. Si calcolino la mediana e un opportuno indice di variabilità della marginale G .

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.



Si tratta di un carattere qualitativo ordinato.

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.



La mediana (percentile di ordine 0.50) in presenza di un numero dispari di osservazioni corrisponde all'elemento che occupa la posizione $\frac{n+1}{2} = 11$ nella serie dei dati ordinati. Dalla distribuzione delle frequenze cumulate

g_j	$n_{\bullet j}$	$N_{\bullet j}$
b	4	4
m	7	11
a	10	21

si ottiene $g_{0.5} = m$.

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.

In presenza di un carattere qualitativo la variabilità può essere misurata mediante gli indici di Gini e di Frosini normalizzati

$$G_N = \frac{1 - \sum_{i=1}^k f_{\bullet j}^2}{1 - \frac{1}{k}}$$

e

$$F_N = 1 - \sqrt{1 - G_N}.$$



Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.



In presenza di un carattere qualitativo la variabilità può essere misurata mediante gli indici di Gini e di Frosini normalizzati

$$G_N = \frac{1 - \sum_{i=1}^k f_{\bullet j}^2}{1 - \frac{1}{k}}$$

e

$$F_N = 1 - \sqrt{1 - G_N}.$$

Si ottiene:

g_j	$n_{\bullet j}$	$f_{\bullet j}$	$f_{\bullet j}^2$
b	4	0.1905	0.0363
m	7	0.3333	0.1111
a	10	0.4762	0.2268
	21	1	0.3742

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.



In presenza di un carattere qualitativo la variabilità può essere misurata mediante gli indici di Gini e di Frosini normalizzati

$$G_N = \frac{1 - \sum_{i=1}^k f_{\bullet j}^2}{1 - \frac{1}{k}}$$

e

$$F_N = 1 - \sqrt{1 - G_N}.$$

Si ottiene:

g_j	$n_{\bullet j}$	$f_{\bullet j}$	$f_{\bullet j}^2$
b	4	0.1905	0.0363
m	7	0.3333	0.1111
a	10	0.4762	0.2268
	21	1	0.3742

Da cui segue:

$$G_N = \frac{1 - 0.3741}{1 - \frac{1}{3}} = 0.9388$$

e

$$F_N = 1 - \sqrt{1 - 0.9388} = 0.7526.$$

Sussiste, quindi, un livello medio alto di eterogeneità.



Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.

3. Si confronti la variabilità delle distribuzioni condizionate $P|G = m$ e $P|G = a$.



Si considera la distribuzione condizionata $P|g_2$

p_i	n_{i2}
5 → 15	0
15 → 30	7

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.



Si tratta di un carattere quantitativo con modalità raggruppate in classi.

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.



Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.

Si osserva che la distribuzione è degenere.

$M(P|g_2) = 22.5$. La varianza e il coefficiente di variazione sono nulli.



Si considera la distribuzione condizionata $P|g_3$

p_i	n_{i3}
5 → 15	2
15 → 30	8

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.



Si tratta di un carattere quantitativo con modalità raggruppate in classi.

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.



La media aritmetica

$$M(P|g_3) = \frac{1}{n_{\bullet 3}} \sum_{i=1}^h p_i n_{i3}$$

risulta 20.

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.



La media aritmetica

$$M(P|g_3) = \frac{1}{n_{\bullet 3}} \sum_{i=1}^h p_i n_{i3}$$

risulta 20.

Si ottiene (considerando come modalità i valori centrali delle classi)

p_i	n_{i3}	$p_i n_{i3}$
10.0	2	20
22.5	8	180
	10	200

da cui segue

$$M(P|g_3) = \frac{1}{n_{\bullet 3}} \sum_{i=1}^h p_i n_{i3} = \frac{200}{10} = 20.$$

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.



La varianza

$$\text{Var}(P|g_3) = M(P^2|g_3) - [M(P|g_3)]^2$$

risulta 25.

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.



Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.

La varianza

$$\text{Var}(P|g_3) = M(P^2|g_3) - [M(P|g_3)]^2$$

risulta 25. Si ottiene (considerando come modalità i valori centrali delle classi)

p_i	n_{i3}	$p_i n_{i3}$	$p_i^2 n_{i3}$
10.0	2	20	200
22.5	8	180	4050
	10	200	4250

da cui segue

$$\text{Var}(P|g_3) = \frac{4250}{10} - 20^2 = 425 - 400 = 25.$$



Per effettuare confronti in termini di variabilità tra caratteri quantitativi occorre utilizzare una misura relativa, ad esempio il coefficiente di variazione

$$CV = \frac{\sigma}{\mu} = \frac{\sqrt{\text{Var}(P|g_3)}}{M(P|g_3)} = \frac{\sqrt{25}}{20} = \frac{5}{20} = 0.25.$$

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.



Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.

$P|g = m$: Indice di posizione: $M(P|g = m) = 22.5$ Indice di variabilità: $CV(P|g = m) = 0$

$P|g = a$: Indice di posizione: $M(P|g = a) = 20$ Indice di variabilità: $CV(P|g = a) = 0.25$

La distribuzione $P|g = m$ presenta un livello medio più elevato di $P|g = a$, e variabilità nulla.



4. Si determini il grado di connessione tra le variabili P e G mediante il calcolo di un opportuno indice.

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.



Dalla distribuzione delle frequenze congiunte n_{ij} :

	b	m	a	$n_{i\bullet}$
$5 \vdash 15$	1	0	2	3
$15 \vdash 30$	3	7	8	18
$n_{\bullet j}$	4	7	10	21

è possibile ottenere le frequenze teoriche nella situazione di indipendenza stocastica

$$\hat{n}_{ij} = \frac{n_{i\bullet} n_{\bullet j}}{n}:$$

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.



Dalla distribuzione delle frequenze congiunte n_{ij} :

	b	m	a	$n_{i\bullet}$
5 + 15	1	0	2	3
15 + 30	3	7	8	18
$n_{\bullet j}$	4	7	10	21

è possibile ottenere le frequenze teoriche nella situazione di indipendenza stocastica

$$\hat{n}_{ij} = \frac{n_{i\bullet} \cdot n_{\bullet j}}{n}:$$

	b	m	a	$n_{i\bullet}$
5 + 15	0.5714	1	1.4286	3
15 + 30	3.4286	6	8.5714	18
$n_{\bullet j}$	4.0000	7	10.0000	21

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.



le contingenze assolute $n_{ij} - \hat{n}_{ij}$:

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.



le contingenze assolute $n_{ij} - \hat{n}_{ij}$:

	b	m	a
$5 \vdash 15$	0.4286	-1	0.5714
$15 \vdash 30$	-0.4286	1	-0.5714

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.



i rapporti di contingenza $\frac{n_{ij}}{\hat{n}_{ij}}$:

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.



i rapporti di contingenza $\frac{n_{ij}}{\hat{n}_{ij}}$:

	<i>b</i>	<i>m</i>	<i>a</i>
5 $\dashv$ 15	1.750	0.0000	1.4000
15 $\dashv$ 30	0.875	1.1667	0.9333

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.



e le contingenze relative, espresse in termini percentuali, $100 \left(\frac{n_{ij} - \hat{n}_{ij}}{\hat{n}_{ij}} \right) \% = 100 \left(\frac{n_{ij}}{\hat{n}_{ij}} - 1 \right) \%$:

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.



e le contingenze relative, espresse in termini percentuali, $100 \left(\frac{n_{ij} - \hat{n}_{ij}}{\hat{n}_{ij}} \right) \% = 100 \left(\frac{n_{ij}}{\hat{n}_{ij}} - 1 \right) \%$:

	<i>b</i>	<i>m</i>	<i>a</i>
5 $\vdash$ 15	75.0	-100.00	40.00
15 $\vdash$ 30	-12.5	16.67	-6.67

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.



Risultati intermedi $\frac{(n_{ij} - \hat{n}_{ij})^2}{\hat{n}_{ij}}$ per il calcolo del χ^2

	<i>b</i>	<i>m</i>	<i>a</i>
5 $\vdash$ 15	0.3214	1.0000	0.2286
15 $\vdash$ 30	0.0536	0.1667	0.0381

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.



Risultati intermedi $\frac{n_{ij}^2}{n_{i\bullet} \cdot n_{\bullet j}}$ per il calcolo del χ^2 (formula operativa)

	<i>b</i>	<i>m</i>	<i>a</i>
5 $\vdash$ 15	0.0833	0.0000	0.1333
15 $\vdash$ 30	0.1250	0.3889	0.3556

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.



Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.

$$\chi^2 = \sum_{i=1}^h \sum_{j=1}^k \frac{(n_{ij} - \hat{n}_{ij})^2}{\hat{n}_{ij}} = 1.8083,$$

$$\chi_N^2 = \frac{\chi^2}{n \cdot \min(h-1, k-1)} = 0.0861$$

$$\chi_N = \sqrt{\chi_N^2} = 0.2934.$$

Sussiste, quindi, un livello di connessione molto basso.

Esercizio 2

Prova scritta di
STATISTICA del
25.06.2015

Giuseppe Boari
Gabriele Cantaluppi



UNIVERSITÀ
CATTOLICA
del Sacro Cuore

Si consideri la seguente serie dei numeri indici a base fissa, con base $2010 = 1$, della serie X_t dei prezzi del prodotto A:

t	$NIBF$
2010	1
2011	0.92
2012	0.98
2013	1.03
2014	1.01

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.



1. Si costruisca la corrispondente serie dei numeri indici a base mobile.

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.



1. Si costruisca la corrispondente serie dei numeri indici a base mobile.

Si ricorda che ${}_{t-1}I_t = \frac{x_t}{x_{t-1}} = \frac{{}_{2010}I_t}{{}_{2010}I_{t-1}}$.

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.



1. Si costruisca la corrispondente serie dei numeri indici a base mobile.

Si ricorda che ${}_{t-1}I_t = \frac{x_t}{x_{t-1}} = \frac{{}_{2010}I_t}{{}_{2010}I_{t-1}}$. Si ottiene quindi

<i>Anno</i>	<i>NIBF</i>	<i>NIBM</i>
2010	1.00	
2011	0.92	0.9200
2012	0.98	1.0652
2013	1.03	1.0510
2014	1.01	0.9806

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.



2. Sapendo che $x_{2013} = 261$, si calcoli x_{2014} .

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.



2. Sapendo che $x_{2013} = 261$, si calcoli x_{2014} .

Si ricorda che

$${}_{2013}I_{2014} = \frac{x_{2014}}{x_{2013}} = \frac{{}_{2010}I_{2014}}{{}_{2010}I_{2013}} = 0.9806.$$

Abbiamo, quindi,

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.



Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.

2. Sapendo che $x_{2013} = 261$, si calcoli x_{2014} .

Si ricorda che

$${}_{2013}I_{2014} = \frac{x_{2014}}{x_{2013}} = \frac{{}_{2010}I_{2014}}{{}_{2010}I_{2013}} = 0.9806.$$

Abbiamo, quindi,

$$x_{2014} = x_{2013} \cdot {}_{2013}I_{2014} = 261 \cdot 0.9806 = 255.93$$



3. Si calcoli il tasso medio di variazione dal 2010 al 2014 e si commenti il risultato ottenuto.

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.



3. Si calcoli il tasso medio di variazione dal 2010 al 2014 e si commenti il risultato ottenuto.
Si ricorda che

$${}_{2010}I_{2014} = {}_{2010}I_{2011} \cdot {}_{2011}I_{2012} \cdot {}_{2012}I_{2013} \cdot {}_{2013}I_{2014}$$

Si deve trovare il numero indice a base mobile medio.

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.



Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.

3. Si calcoli il tasso medio di variazione dal 2010 al 2014 e si commenti il risultato ottenuto. Si ricorda che

$${}_{2010}I_{2014} = {}_{2010}I_{2011} \cdot {}_{2011}I_{2012} \cdot {}_{2012}I_{2013} \cdot {}_{2013}I_{2014}$$

Si deve trovare il numero indice a base mobile medio. Applicando il criterio di scelta della media secondo Chisini si ottiene

$${}_{2010}I_{2014} = {}_{2010}I_{2011} \cdot {}_{2011}I_{2012} \cdot {}_{2012}I_{2013} \cdot {}_{2013}I_{2014} = \alpha \cdot \alpha \cdot \alpha \cdot \alpha = \alpha^4.$$



Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.

3. Si calcoli il tasso medio di variazione dal 2010 al 2014 e si commenti il risultato ottenuto. Si ricorda che

$${}_{2010}I_{2014} = {}_{2010}I_{2011} \cdot {}_{2011}I_{2012} \cdot {}_{2012}I_{2013} \cdot {}_{2013}I_{2014}$$

Si deve trovare il numero indice a base mobile medio. Applicando il criterio di scelta della media secondo Chisini si ottiene

$${}_{2010}I_{2014} = {}_{2010}I_{2011} \cdot {}_{2011}I_{2012} \cdot {}_{2012}I_{2013} \cdot {}_{2013}I_{2014} = \alpha \cdot \alpha \cdot \alpha \cdot \alpha = \alpha^4.$$

Quindi

$$\alpha = \sqrt[4]{{}_{2010}I_{2014}} = \sqrt[4]{{}_{2010}I_{2011} \cdot {}_{2011}I_{2012} \cdot {}_{2012}I_{2013} \cdot {}_{2013}I_{2014}} = 1.0025.$$

Da cui segue una variazione relativa media pari a 0.0025, corrispondente a un incremento medio annuo tra il 2010 e il 2014 del 0.25%.



[Esercizio 1.](#)

[Esercizio 2.](#)

[Esercizio 3.](#)

[Esercizio 4.](#)

Il seguente prospetto riporta i valori rilevati in corrispondenza a 5 unità statistiche con riferimento a due caratteri quantitativi X e Y

x_i	18	18	20	20	20
y_i	17	19	20	22	24

1. Si rappresentino graficamente le coppie dei dati e la funzione di regressione ipotizzando Y in funzione di X .



Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.

La funzione di regressione associa a ogni valore x appartenente all'insieme di definizione della variabile X la media di Y condizionata a x , $M(Y|x)$

x	$M(Y x)$
18	$\frac{1}{2}(17 + 19) = 18$
20	$\frac{1}{3}(20 + 22 + 24) = 22$



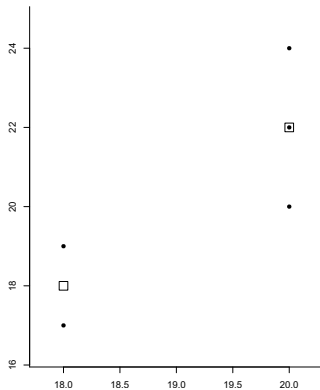
Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.

Le coppie di dati possono essere rappresentate mediante un grafico a dispersione.





2. Si calcolino, secondo il metodo dei minimi quadrati, i parametri e l'indice di adattamento dei seguenti modelli di regressione, individuando il modello migliore:

$$I) Y^* = a + b \cdot X \qquad II) Y^* = c$$

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.



Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.

Si riportano i calcoli necessari¹ per ottenere le stime dei parametri e gli indici di adattamento del modello

$$Y = a + bX + E$$

x_i	y_i	$x_i y_i$	x_i^2	y_i^2
18	17	306	324	289
18	19	342	324	361
20	20	400	400	400
20	22	440	400	484
20	24	480	400	576
96	102	1968	1848	2110
19.2	20.4	393.6	369.6	422
$M(X)$	$M(Y)$	$M(XY)$	$M(X^2)$	$M(Y^2)$

$$\hat{b} =$$

¹Si osserva che per ragioni di stampa alcuni calcoli potrebbero essere duplicati con riferimento ai diversi modelli oggetto di studio



Si riportano i calcoli necessari¹ per ottenere le stime dei parametri e gli indici di adattamento del modello

$$Y = a + bX + E$$

x_i	y_i	$x_i y_i$	x_i^2	y_i^2
18	17	306	324	289
18	19	342	324	361
20	20	400	400	400
20	22	440	400	484
20	24	480	400	576
96	102	1968	1848	2110
19.2	20.4	393.6	369.6	422
$M(X)$	$M(Y)$	$M(XY)$	$M(X^2)$	$M(Y^2)$

$$\begin{aligned}\hat{b} &= \frac{\text{Cov}(X, Y)}{\text{Var}(X)} = \frac{M(XY) - M(X)M(Y)}{M(X^2) - [M(X)]^2} = \\ &= \frac{393.6 - 19.2 \cdot 20.4}{369.6 - 19.2^2} = \frac{1.92}{0.96} = 2 \\ \hat{a} &= \end{aligned}$$

¹Si osserva che per ragioni di stampa alcuni calcoli potrebbero essere duplicati con riferimento ai diversi modelli oggetto di studio



Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.

Si riportano i calcoli necessari¹ per ottenere le stime dei parametri e gli indici di adattamento del modello

$$Y = a + bX + E$$

x_i	y_i	$x_i y_i$	x_i^2	y_i^2
18	17	306	324	289
18	19	342	324	361
20	20	400	400	400
20	22	440	400	484
20	24	480	400	576
96	102	1968	1848	2110
19.2	20.4	393.6	369.6	422
$M(X)$	$M(Y)$	$M(XY)$	$M(X^2)$	$M(Y^2)$

$$\hat{b} = \frac{\text{Cov}(X, Y)}{\text{Var}(X)} = \frac{M(XY) - M(X)M(Y)}{M(X^2) - [M(X)^2]} =$$

$$= \frac{393.6 - 19.2 \cdot 20.4}{369.6 - 19.2^2} = \frac{1.92}{0.96} = 2$$

$$\hat{a} = M(Y) - \hat{b}M(X) = -18$$

$$\rho^2 =$$

¹Si osserva che per ragioni di stampa alcuni calcoli potrebbero essere duplicati con riferimento ai diversi modelli oggetto di studio



Si riportano i calcoli necessari¹ per ottenere le stime dei parametri e gli indici di adattamento del modello

$$Y = a + bX + E$$

x_i	y_i	$x_i y_i$	x_i^2	y_i^2
18	17	306	324	289
18	19	342	324	361
20	20	400	400	400
20	22	440	400	484
20	24	480	400	576
96	102	1968	1848	2110
19.2	20.4	393.6	369.6	422
$M(X)$	$M(Y)$	$M(XY)$	$M(X^2)$	$M(Y^2)$

$$\begin{aligned}\hat{b} &= \frac{\text{Cov}(X, Y)}{\text{Var}(X)} = \frac{M(XY) - M(X)M(Y)}{M(X^2) - [M(X)]^2} = \\ &= \frac{393.6 - 19.2 \cdot 20.4}{369.6 - 19.2^2} = \frac{1.92}{0.96} = 2 \\ \hat{a} &= M(Y) - \hat{b}M(X) = -18 \\ \rho^2 &= \frac{[\text{Cov}(X, Y)]^2}{\text{Var}(X) \text{Var}(Y)} = \frac{1.92^2}{0.96 \cdot 5.84} = 0.6575\end{aligned}$$

¹Si osserva che per ragioni di stampa alcuni calcoli potrebbero essere duplicati con riferimento ai diversi modelli oggetto di studio



Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.

Si riportano i calcoli necessari² per ottenere le stime dei parametri e gli indici di adattamento del modello

$$Y = c + E$$

x_i	y_i	y_i^2
18	17	289
18	19	361
20	20	400
20	22	484
20	24	576
102		2110
20.4		422
$M(Y)$		$M(Y^2)$

$$\hat{c} =$$

²Si osserva che per ragioni di stampa alcuni calcoli potrebbero essere duplicati con riferimento ai diversi modelli oggetto di studio



Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.

Si riportano i calcoli necessari² per ottenere le stime dei parametri e gli indici di adattamento del modello

$$Y = c + E$$

x_i	y_i	y_i^2
18	17	289
18	19	361
20	20	400
20	22	484
20	24	576
102		2110
20.4		422
$M(Y)$		$M(Y^2)$

$$\hat{c} = M(Y) = 20.4$$

$$R^2 =$$

²Si osserva che per ragioni di stampa alcuni calcoli potrebbero essere duplicati con riferimento ai diversi modelli oggetto di studio



Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.

Si riportano i calcoli necessari² per ottenere le stime dei parametri e gli indici di adattamento del modello

$$Y = c + E$$

x_i	y_i	y_i^2
18	17	289
18	19	361
20	20	400
20	22	484
20	24	576
102		2110
20.4		422
$M(Y)$		$M(Y^2)$

$$\hat{c} = M(Y) = 20.4$$

$$R^2 = 0$$

²Si osserva che per ragioni di stampa alcuni calcoli potrebbero essere duplicati con riferimento ai diversi modelli oggetto di studio



$a = -18$, $b = 2$, indice adattamento $\rho^2 = 0.6575$ $c = 20.4$, $\rho^2 = 0$
Il modello migliore è $Y^* = -18 + 2X$.

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.



Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.

3. Si calcoli $\eta^2_{Y|X}$.



Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.

3. Si calcoli $\eta_{Y|X}^2$.

$\eta_{Y|X}^2 = \rho^2 = 0.6575$ in quanto le medie condizionate sono allineate.

Esercizio 4

Prova scritta di
STATISTICA del
25.06.2015

Giuseppe Boari
Gabriele Cantaluppi



UNIVERSITÀ
CATTOLICA
del Sacro Cuore

Una compagnia assicurativa valuta che la probabilità che un suo assicurato richieda un risarcimento per danni provocati da eventi naturali sia pari all'1.1%.

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.



Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.

1. Si calcoli la probabilità che, estratti in modo indipendente 10 clienti che hanno stipulato una polizza per danni provocati da eventi naturali, almeno due richiedano il risarcimento assicurativo.

Sia X la variabile casuale 'n. di clienti che richiedono il risarcimento assicurativo nelle 10 estrazioni indipendenti'. $X \sim$



Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.

1. Si calcoli la probabilità che, estratti in modo indipendente 10 clienti che hanno stipulato una polizza per danni provocati da eventi naturali, almeno due richiedano il risarcimento assicurativo.

Sia X la variabile casuale 'n. di clienti che richiedono il risarcimento assicurativo nelle 10 estrazioni indipendenti'. $X \sim \text{Bin}(n = 10, p = 0.011)$.

$$P(X \geq 2) =$$



Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.

1. Si calcoli la probabilità che, estratti in modo indipendente 10 clienti che hanno stipulato una polizza per danni provocati da eventi naturali, almeno due richiedano il risarcimento assicurativo.

Sia X la variabile casuale 'n. di clienti che richiedono il risarcimento assicurativo nelle 10 estrazioni indipendenti'. $X \sim \text{Bin}(n = 10, p = 0.011)$.

$$P(X \geq 2) = 1 - P(X < 2) =$$



Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.

1. Si calcoli la probabilità che, estratti in modo indipendente 10 clienti che hanno stipulato una polizza per danni provocati da eventi naturali, almeno due richiedano il risarcimento assicurativo.

Sia X la variabile casuale 'n. di clienti che richiedono il risarcimento assicurativo nelle 10 estrazioni indipendenti'. $X \sim \text{Bin}(n = 10, p = 0.011)$.

$$P(X \geq 2) = 1 - P(X < 2) = 1 - (P(X = 0) + P(X = 1)) =$$

=



Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.

1. Si calcoli la probabilità che, estratti in modo indipendente 10 clienti che hanno stipulato una polizza per danni provocati da eventi naturali, almeno due richiedano il risarcimento assicurativo.

Sia X la variabile casuale 'n. di clienti che richiedono il risarcimento assicurativo nelle 10 estrazioni indipendenti'. $X \sim \text{Bin}(n = 10, p = 0.011)$.

$$\begin{aligned} P(X \geq 2) &= 1 - P(X < 2) = 1 - (P(X = 0) + P(X = 1)) = \\ &= 1 - \left\{ \binom{10}{0} 0.011^0 0.989^{10} + \binom{10}{1} 0.011^1 0.989^9 \right\} = \\ &= \end{aligned}$$



Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.

1. Si calcoli la probabilità che, estratti in modo indipendente 10 clienti che hanno stipulato una polizza per danni provocati da eventi naturali, almeno due richiedano il risarcimento assicurativo.

Sia X la variabile casuale 'n. di clienti che richiedono il risarcimento assicurativo nelle 10 estrazioni indipendenti'. $X \sim \text{Bin}(n = 10, p = 0.011)$.

$$\begin{aligned} P(X \geq 2) &= 1 - P(X < 2) = 1 - (P(X = 0) + P(X = 1)) = \\ &= 1 - \left\{ \binom{10}{0} 0.011^0 0.989^{10} + \binom{10}{1} 0.011^1 0.989^9 \right\} = \\ &= 1 - (0.8953 + 0.0996) = 1 - 0.9949 = 0.0051. \end{aligned}$$



2. Si calcolino media e varianza della variabile casuale descritta dall'esperimento al punto precedente.

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.



2. Si calcolino media e varianza della variabile casuale descritta dall'esperimento al punto precedente.

Essendo $X \sim \text{Bin}(n = 10, p = 0.011)$ abbiamo

$$M(X) =$$

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.



2. Si calcolino media e varianza della variabile casuale descritta dall'esperimento al punto precedente.

Essendo $X \sim \text{Bin}(n = 10, p = 0.011)$ abbiamo

$$M(X) = np = 10 \cdot 0.011 = 0.11$$

$$\text{Var}(X) =$$

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.



2. Si calcolino media e varianza della variabile casuale descritta dall'esperimento al punto precedente.

Essendo $X \sim \text{Bin}(n = 10, p = 0.011)$ abbiamo

$$\begin{aligned} M(X) &= np = 10 \cdot 0.011 = 0.11 \\ \text{Var}(X) &= np(1 - p) = 10 \cdot 0.011 \cdot (1 - 0.011) = \\ &= 10 \cdot 0.011 \cdot 0.989 = 0.1088 \end{aligned}$$

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.



Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.

3. Si calcoli la probabilità che, estratti in modo indipendente 1000 clienti che hanno stipulato una polizza per danni provocati da eventi naturali, al massimo 20 richiedano il risarcimento assicurativo.

Indicata con X la variabile casuale 'n. di clienti che richiedono il risarcimento assicurativo nelle 1000 estrazioni indipendenti' risulta

$$X \sim$$



3. Si calcoli la probabilità che, estratti in modo indipendente 1000 clienti che hanno stipulato una polizza per danni provocati da eventi naturali, al massimo 20 richiedano il risarcimento assicurativo.

Indicata con X la variabile casuale 'n. di clienti che richiedono il risarcimento assicurativo nelle 1000 estrazioni indipendenti' risulta

$$X \sim \text{Bin}(n = 1000, p = 0.011).$$

$M(X) = 11$; $\text{Var}(X) = 10.879$. Dal momento che $np = 11 > 5$ ed $n(1 - p) = 989 > 5$ è possibile utilizzare l'approssimazione della variabile casuale Binomiale con la variabile casuale Normale

$$X \sim N(\mu = 11, \sigma^2 = 10.879).$$

Si ottiene

$$P(X \leq 20) =$$



3. Si calcoli la probabilità che, estratti in modo indipendente 1000 clienti che hanno stipulato una polizza per danni provocati da eventi naturali, al massimo 20 richiedano il risarcimento assicurativo.

Indicata con X la variabile casuale 'n. di clienti che richiedono il risarcimento assicurativo nelle 1000 estrazioni indipendenti' risulta

$$X \sim \text{Bin}(n = 1000, p = 0.011).$$

$M(X) = 11$; $\text{Var}(X) = 10.879$. Dal momento che $np = 11 > 5$ ed $n(1 - p) = 989 > 5$ è possibile utilizzare l'approssimazione della variabile casuale Binomiale con la variabile casuale Normale

$$X \sim N(\mu = 11, \sigma^2 = 10.879).$$

Si ottiene

$$P(X \leq 20) = P\left(\frac{X - \mu}{\sigma} \leq \frac{20 - \mu}{\sigma}\right) =$$

$$=$$



3. Si calcoli la probabilità che, estratti in modo indipendente 1000 clienti che hanno stipulato una polizza per danni provocati da eventi naturali, al massimo 20 richiedano il risarcimento assicurativo.

Indicata con X la variabile casuale 'n. di clienti che richiedono il risarcimento assicurativo nelle 1000 estrazioni indipendenti' risulta

$$X \sim \text{Bin}(n = 1000, p = 0.011).$$

$M(X) = 11$; $\text{Var}(X) = 10.879$. Dal momento che $np = 11 > 5$ ed $n(1 - p) = 989 > 5$ è possibile utilizzare l'approssimazione della variabile casuale Binomiale con la variabile casuale Normale

$$X \sim N(\mu = 11, \sigma^2 = 10.879).$$

Si ottiene

$$\begin{aligned} P(X \leq 20) &= P\left(\frac{X - \mu}{\sigma} \leq \frac{20 - \mu}{\sigma}\right) = \\ &= P\left(Z \leq \frac{20 - 11}{3.2983}\right) = \end{aligned}$$



3. Si calcoli la probabilità che, estratti in modo indipendente 1000 clienti che hanno stipulato una polizza per danni provocati da eventi naturali, al massimo 20 richiedano il risarcimento assicurativo.

Indicata con X la variabile casuale 'n. di clienti che richiedono il risarcimento assicurativo nelle 1000 estrazioni indipendenti' risulta

$$X \sim \text{Bin}(n = 1000, p = 0.011).$$

$M(X) = 11$; $\text{Var}(X) = 10.879$. Dal momento che $np = 11 > 5$ ed $n(1 - p) = 989 > 5$ è possibile utilizzare l'approssimazione della variabile casuale Binomiale con la variabile casuale Normale

$$X \sim N(\mu = 11, \sigma^2 = 10.879).$$

Si ottiene

$$\begin{aligned} P(X \leq 20) &= P\left(\frac{X - \mu}{\sigma} \leq \frac{20 - \mu}{\sigma}\right) = \\ &= P\left(Z \leq \frac{20 - 11}{3.2983}\right) = P(Z \leq 2.73) = \end{aligned}$$



3. Si calcoli la probabilità che, estratti in modo indipendente 1000 clienti che hanno stipulato una polizza per danni provocati da eventi naturali, al massimo 20 richiedano il risarcimento assicurativo.

Indicata con X la variabile casuale 'n. di clienti che richiedono il risarcimento assicurativo nelle 1000 estrazioni indipendenti' risulta

$$X \sim \text{Bin}(n = 1000, p = 0.011).$$

$M(X) = 11$; $\text{Var}(X) = 10.879$. Dal momento che $np = 11 > 5$ ed $n(1 - p) = 989 > 5$ è possibile utilizzare l'approssimazione della variabile casuale Binomiale con la variabile casuale Normale

$$X \sim N(\mu = 11, \sigma^2 = 10.879).$$

Si ottiene

$$\begin{aligned} P(X \leq 20) &= P\left(\frac{X - \mu}{\sigma} \leq \frac{20 - \mu}{\sigma}\right) = \\ &= P\left(Z \leq \frac{20 - 11}{3.2983}\right) = P(Z \leq 2.73) = 0.9968. \end{aligned}$$

Tema T284-2

Prova scritta di STATISTICA del 09.07.2015

Prova scritta di
STATISTICA del
09.07.2015

Giuseppe Boari
Gabriele Cantaluppi



UNIVERSITÀ
CATTOLICA
del Sacro Cuore

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.

Giuseppe Boari
Gabriele Cantaluppi
Università Cattolica del Sacro Cuore, Milano



Il seguente prospetto riporta la distribuzione dei clienti presenti su un mercato dove opera l'agenzia A, riclassificati mediante le variabili X , con categorie x_1 = 'clienti agenzia A' e x_2 = 'clienti della concorrenza', e Y = 'tipologia di viaggio', con categorie y_1 = 'low cost', y_2 = 'medium cost' e y_3 = 'high cost'.

Nella soluzione di questo esercizio si tenga presente che la clientela *target* dell'agenzia A è del tipo 'high cost'.

$X \setminus Y$	y_1	y_2	y_3
x_1	200	500	800
x_2	5000	2000	1500

1. Si rappresentino graficamente le distribuzioni condizionate $Y|x_1$ e $Y|x_2$.

Esercizio 1.

Esercizio 2.

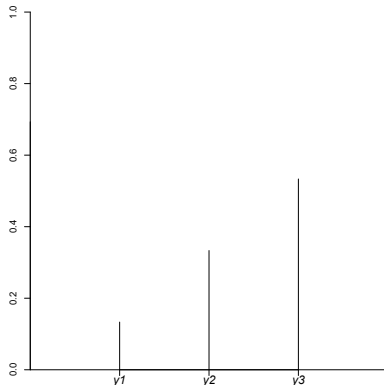
Esercizio 3.

Esercizio 4.

Si considera la distribuzione condizionata $Y|x_1$

y_j	n_{1j}
y_1	200
y_2	500
y_3	800

Si tratta di un carattere qualitativo ordinato che può essere rappresentato graficamente mediante un grafico a bastoncini o a canne d'organo.



Esercizio 1.

Esercizio 2.

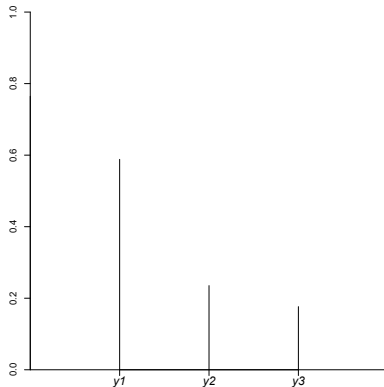
Esercizio 3.

Esercizio 4.

Si considera la distribuzione condizionata $Y|x_2$

y_j	n_{2j}
y_1	5000
y_2	2000
y_3	1500

Si tratta di un carattere qualitativo ordinato che può essere rappresentato graficamente mediante un grafico a bastoncini o a canne d'organo.



[Esercizio 1.](#)

[Esercizio 2.](#)

[Esercizio 3.](#)

[Esercizio 4.](#)



2. Si confrontino le distribuzioni condizionate $Y|x_1$ e $Y|x_2$ con opportuni indici di posizione e di variabilità.

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.



Si considera la distribuzione condizionata $Y|x_1$

y_j	n_{1j}
y_1	200
y_2	500
y_3	800

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.



Si tratta di un carattere qualitativo ordinato.

La moda corrisponde alla categoria alla quale è associata la massima frequenza) che risulta y_3 .

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.



La mediana (percentile di ordine 0.50) in presenza di un numero pari di osservazioni corrisponde a uno degli elementi che occupano le posizioni $\frac{n}{2} = 750$ ed $\frac{n}{2} + 1 = 751$ nella serie dei dati ordinati. Dalla distribuzione delle frequenze cumulate

y_j	n_{1j}	N_{1j}
y_1	200	200
y_2	500	700
y_3	800	1500

i due elementi coincidono e risulta $y_{0.5} = y_3$.

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.



In presenza di un carattere qualitativo la variabilità può essere misurata mediante gli indici di Gini e di Frosini normalizzati

$$G_N = \frac{1 - \sum_{i=1}^k f_{ij}^2}{1 - \frac{1}{k}}$$

e

$$F_N = 1 - \sqrt{1 - G_N}.$$

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.



In presenza di un carattere qualitativo la variabilità può essere misurata mediante gli indici di Gini e di Frosini normalizzati

$$G_N = \frac{1 - \sum_{i=1}^k f_{ij}^2}{1 - \frac{1}{k}}$$

e

$$F_N = 1 - \sqrt{1 - G_N}.$$

Si ottiene:

y_j	n_{1j}	f_{1j}	f_{1j}^2
y_1	200	0.1333	0.0178
y_2	500	0.3333	0.1111
y_3	800	0.5333	0.2844
	1500	1	0.4133

Da cui segue: $G_N = 0.88$ e $F_N = 0.6536$. Sussiste, quindi, un livello medio alto di eterogeneità.



Si considera la distribuzione condizionata $Y|x_2$

y_j	n_{2j}
y_1	5000
y_2	2000
y_3	1500

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.



Si tratta di un carattere qualitativo ordinato.

La moda corrisponde alla categoria alla quale è associata la massima frequenza) che risulta y_1 .

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.



La mediana (percentile di ordine 0.50) in presenza di un numero pari di osservazioni corrisponde a uno degli elementi che occupano le posizioni $\frac{n}{2} = 4250$ ed $\frac{n}{2} + 1 = 4251$ nella serie dei dati ordinati. Dalla distribuzione delle frequenze cumulate

y_j	n_{2j}	N_{2j}
y_1	5000	5000
y_2	2000	7000
y_3	1500	8500

i due elementi coincidono e risulta $y_{0.5} = y_1$.

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.



In presenza di un carattere qualitativo la variabilità può essere misurata mediante gli indici di Gini e di Frosini normalizzati

$$G_N = \frac{1 - \sum_{i=1}^k f_{2j}^2}{1 - \frac{1}{k}}$$

e

$$F_N = 1 - \sqrt{1 - G_N}.$$

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.



In presenza di un carattere qualitativo la variabilità può essere misurata mediante gli indici di Gini e di Frosini normalizzati

$$G_N = \frac{1 - \sum_{i=1}^k f_{2j}^2}{1 - \frac{1}{k}}$$

e

$$F_N = 1 - \sqrt{1 - G_N}.$$

Si ottiene:

y_j	n_{2j}	f_{2j}	f_{2j}^2
y_1	5000	0.5882	0.346
y_2	2000	0.2353	0.0554
y_3	1500	0.1765	0.0311
	8500	1	0.4325

Da cui segue: $G_N = 0.8512$ e $F_N = 0.6143$. Sussiste, quindi, un livello medio di eterogeneità.



$Y|x_1$: Indice di posizione: $\text{Moda}(Y|x_1) = y_3$ Indice di variabilità: $F_N(Y|x_1) = 0.6536$

$Y|x_2$: Indice di posizione: $\text{Moda}(Y|x_2) = y_1$ Indice di variabilità: $F_N(Y|x_2) = 0.6143$

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.



$Y|x_1$: Indice di posizione: $\text{Moda}(Y|x_1) = y_3$ Indice di variabilità: $F_N(Y|x_1) = 0.6536$

$Y|x_2$: Indice di posizione: $\text{Moda}(Y|x_2) = y_1$ Indice di variabilità: $F_N(Y|x_2) = 0.6143$

La maggior parte delle tipologie di viaggio scelte dai clienti dell'agenzia A (53%) sono di tipo 'high cost'. La concorrenza è, invece, focalizzata su tipologie di viaggio 'low cost' (59%).

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.



$Y|x_1$: Indice di posizione: $\text{Moda}(Y|x_1) = y_3$ Indice di variabilità: $F_N(Y|x_1) = 0.6536$

$Y|x_2$: Indice di posizione: $\text{Moda}(Y|x_2) = y_1$ Indice di variabilità: $F_N(Y|x_2) = 0.6143$

La maggior parte delle tipologie di viaggio scelte dai clienti dell'agenzia A (53%) sono di tipo 'high cost'. La concorrenza è, invece, focalizzata su tipologie di viaggio 'low cost' (59%).

Entrambe le distribuzioni presentano un livello medio di eterogeneità, confermando la presenza di tutte e 3 le tipologie di viaggio nell'offerta sia dell'agenzia A che della concorrenza.

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.



3. Si studi il posizionamento dell'agenzia A sul mercato utilizzando le contingenze assolute e le contingenze relative.

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.



3. Si studi il posizionamento dell'agenzia A sul mercato utilizzando le contingenze assolute e le contingenze relative.

Si veda la risposta al punto successivo.

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.



4. Si determini il grado di connessione tra le variabili X e Y mediante il calcolo di un opportuno indice.

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.



Dalla distribuzione delle frequenze congiunte n_{ij} :

	y_1	y_2	y_3	$n_{i\bullet}$
x_1	200	500	800	1500
x_2	5000	2000	1500	8500
$n_{\bullet j}$	5200	2500	2300	10000

è possibile ottenere le frequenze teoriche nella situazione di indipendenza stocastica

$$\hat{n}_{ij} = \frac{n_{i\bullet} n_{\bullet j}}{n}:$$

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.



Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.

Dalla distribuzione delle frequenze congiunte n_{ij} :

	y_1	y_2	y_3	$n_{i\bullet}$
x_1	200	500	800	1500
x_2	5000	2000	1500	8500
$n_{\bullet j}$	5200	2500	2300	10000

è possibile ottenere le frequenze teoriche nella situazione di indipendenza stocastica

$$\hat{n}_{ij} = \frac{n_{i\bullet} \cdot n_{\bullet j}}{n}:$$

	y_1	y_2	y_3	$n_{i\bullet}$
x_1	780	375	345	1500
x_2	4420	2125	1955	8500
$n_{\bullet j}$	5200	2500	2300	10000



le contingenze assolute $n_{ij} - \hat{n}_{ij}$:

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.



le contingenze assolute $n_{ij} - \hat{n}_{ij}$:

	y_1	y_2	y_3
x_1	-580	125	455
x_2	580	-125	-455

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.



i rapporti di contingenza $\frac{n_{ij}}{\hat{n}_{ij}}$:

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.



i rapporti di contingenza $\frac{n_{ij}}{\hat{n}_{ij}}$:

	y_1	y_2	y_3
x_1	0.2564	1.3333	2.3188
x_2	1.1312	0.9412	0.7673

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.



e le contingenze relative, espresse in termini percentuali, $100 \left(\frac{n_{ij} - \hat{n}_{ij}}{\hat{n}_{ij}} \right) \% = 100 \left(\frac{n_{ij}}{\hat{n}_{ij}} - 1 \right) \%$:

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.



e le contingenze relative, espresse in termini percentuali, $100 \left(\frac{n_{ij} - \hat{n}_{ij}}{\hat{n}_{ij}} \right) \% = 100 \left(\frac{n_{ij}}{\hat{n}_{ij}} - 1 \right) \%$:

	y_1	y_2	y_3
x_1	-74.36	33.33	131.88
x_2	13.12	-5.88	-23.27

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.



Risultati intermedi $\frac{(n_{ij} - \hat{n}_{ij})^2}{\hat{n}_{ij}}$ per il calcolo del χ^2

	y_1	y_2	y_3
x_1	431.2821	41.6667	600.0725
x_2	76.1086	7.3529	105.8951

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.



Risultati intermedi $\frac{n_{ij}^2}{n_{i\bullet} \cdot n_{\bullet j}}$ per il calcolo del χ^2 (formula operativa)

	y_1	y_2	y_3
x_1	0.0051	0.0667	0.1855
x_2	0.5656	0.1882	0.1151

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.



Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.

$$\chi^2 = \sum_{i=1}^h \sum_{j=1}^k \frac{(n_{ij} - \hat{n}_{ij})^2}{\hat{n}_{ij}} = 1262.3779,$$

$$\chi_N^2 = \frac{\chi^2}{n \cdot \min(h-1, k-1)} = 0.1262$$

$$\chi_N = \sqrt{\chi_N^2} = 0.3552.$$

Sussiste, quindi, un livello di connessione basso.



Si rileva un ottimo posizionamento dell'agenzia *A* con riferimento alle mete di viaggio '*high cost*' con un *surplus* di clientela rispetto alla situazione di indipendenza stocastica di 455 unità in termini assoluti e del 131.88% in termini relativi.

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.



Si rileva un ottimo posizionamento dell'agenzia *A* con riferimento alle mete di viaggio '*high cost*' con un *surplus* di clientela rispetto alla situazione di indipendenza stocastica di 455 unità in termini assoluti e del 131.88% in termini relativi.

Con riferimento alle mete '*low cost*' il *deficit* di clientela è di 580 unità, circa il 74.36% in meno dei 780 clienti che si avrebbero nella situazione di indipendenza stocastica. In tale situazione il posizionamento dell'agenzia *A* sarebbe identico a quello della concorrenza.

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.



Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.

Si rileva un ottimo posizionamento dell'agenzia *A* con riferimento alle mete di viaggio '*high cost*' con un *surplus* di clientela rispetto alla situazione di indipendenza stocastica di 455 unità in termini assoluti e del 131.88% in termini relativi.

Con riferimento alle mete '*low cost*' il *deficit* di clientela è di 580 unità, circa il 74.36% in meno dei 780 clienti che si avrebbero nella situazione di indipendenza stocastica. In tale situazione il posizionamento dell'agenzia *A* sarebbe identico a quello della concorrenza.

I surplus (131.88%) e deficit (−74.36%) di clientela rappresentano delle componenti distintive dell'agenzia *A* rispetto al comportamento medio generale del mercato.

Esercizio 2

Prova scritta di
STATISTICA del
09.07.2015

Giuseppe Boari
Gabriele Cantaluppi



UNIVERSITÀ
CATTOLICA
del Sacro Cuore

Si consideri la seguente serie dei numeri indici a base mobile della serie X_t dei prezzi del prodotto A:

t	$NIBM$
2010	--
2011	0.92
2012	0.98
2013	1.03
2014	1.01

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.



1. Si costruisca la corrispondente serie dei numeri indici a base fissa, con base $2010 = 1$.

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.



1. Si costruisca la corrispondente serie dei numeri indici a base fissa, con base $2010 = 1$.
Si ricorda che ${}_{2010}I_{2010} = 1$ e che per valori di $t > 2010$ vale

$${}_{2010}I_t = {}_{2010}I_{2011} \cdot {}_{2011}I_{2012} \cdot \dots \cdot {}_{t-1}I_t.$$

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.



Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.

1. Si costruisca la corrispondente serie dei numeri indici a base fissa, con base $2010 = 1$.
Si ricorda che ${}_{2010}I_{2010} = 1$ e che per valori di $t > 2010$ vale

$${}_{2010}I_t = {}_{2010}I_{2011} \cdot {}_{2011}I_{2012} \cdot \dots \cdot {}_{t-1}I_t.$$

Si ottiene quindi

Anno	NIBM	NIBF (2010)
2010		1.0000
2011	0.92	0.9200
2012	0.98	0.9016
2013	1.03	0.9286
2014	1.01	0.9379



2. Sapendo che $x_{2011} = 261$, si calcoli x_{2010} .

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.



Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.

2. Sapendo che $x_{2011} = 261$, si calcoli x_{2010} .

Tenendo presente che

$${}_{2010}I_{2011} = \frac{x_{2011}}{x_{2010}} = 0.92$$

si ottiene

$$x_{2010} = \frac{x_{2011}}{{}_{2010}I_{2011}} = \frac{261}{0.92} = 283.6957.$$



3. Si calcoli il tasso medio di variazione dal 2010 al 2014 e si commenti il risultato ottenuto.

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.



3. Si calcoli il tasso medio di variazione dal 2010 al 2014 e si commenti il risultato ottenuto.
Si ricorda che

$${}_{2010}I_{2014} = {}_{2010}I_{2011} \cdot {}_{2011}I_{2012} \cdot {}_{2012}I_{2013} \cdot {}_{2013}I_{2014}.$$

Si deve trovare il numero indice a base mobile medio.

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.



Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.

3. Si calcoli il tasso medio di variazione dal 2010 al 2014 e si commenti il risultato ottenuto. Si ricorda che

$${}_{2010}I_{2014} = {}_{2010}I_{2011} \cdot {}_{2011}I_{2012} \cdot {}_{2012}I_{2013} \cdot {}_{2013}I_{2014}.$$

Si deve trovare il numero indice a base mobile medio. Applicando il criterio di scelta della media secondo Chisini si ottiene

$${}_{2010}I_{2014} = {}_{2010}I_{2011} \cdot {}_{2011}I_{2012} \cdot {}_{2012}I_{2013} \cdot {}_{2013}I_{2014} = \alpha \cdot \alpha \cdot \alpha \cdot \alpha = \alpha^4.$$



Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.

3. Si calcoli il tasso medio di variazione dal 2010 al 2014 e si commenti il risultato ottenuto. Si ricorda che

$${}_{2010}I_{2014} = {}_{2010}I_{2011} \cdot {}_{2011}I_{2012} \cdot {}_{2012}I_{2013} \cdot {}_{2013}I_{2014}.$$

Si deve trovare il numero indice a base mobile medio. Applicando il criterio di scelta della media secondo Chisini si ottiene

$${}_{2010}I_{2014} = {}_{2010}I_{2011} \cdot {}_{2011}I_{2012} \cdot {}_{2012}I_{2013} \cdot {}_{2013}I_{2014} = \alpha \cdot \alpha \cdot \alpha \cdot \alpha = \alpha^4.$$

Quindi

$$\alpha = \sqrt[4]{{}_{2010}I_{2014}} = \sqrt[4]{{}_{2010}I_{2011} \cdot {}_{2011}I_{2012} \cdot {}_{2012}I_{2013} \cdot {}_{2013}I_{2014}} = 0.9841.$$

Da cui segue una variazione relativa media pari a -0.0159 , corrispondente a un decremento medio annuo tra il 2010 e il 2014 del 1.59%.



Il seguente prospetto riporta i valori rilevati in corrispondenza a 5 unità statistiche con riferimento a due caratteri quantitativi X e Y

x_i	18	18	20	20	20
y_i	17	19	20	22	24

1. Si rappresentino graficamente le coppie dei dati ipotizzando Y in funzione di X .

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.



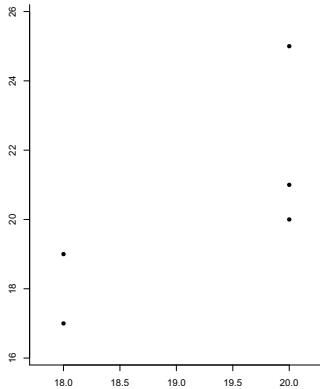
Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.

Le coppie di dati possono essere rappresentate mediante un grafico a dispersione.





2. Si calcolino, secondo il metodo dei minimi quadrati, i parametri e l'indice di adattamento dei seguenti modelli di regressione, individuando il modello migliore:

$$I) Y^* = a + b \cdot X \qquad II) Y^* = c \cdot X$$

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.

Si riportano i calcoli necessari¹ per ottenere le stime dei parametri e gli indici di adattamento del modello

$$Y = a + bX + E$$

x_i	y_i	$x_i y_i$	x_i^2	y_i^2
18	17	306	324	289
18	19	342	324	361
20	20	400	400	400
20	22	440	400	484
20	24	480	400	576
96	102	1968	1848	2110
19.2	20.4	393.6	369.6	422
$M(X)$	$M(Y)$	$M(XY)$	$M(X^2)$	$M(Y^2)$

$$\hat{b} =$$



Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.

¹Si osserva che per ragioni di stampa alcuni calcoli potrebbero essere duplicati con riferimento ai diversi modelli oggetto di studio

Si riportano i calcoli necessari¹ per ottenere le stime dei parametri e gli indici di adattamento del modello

$$Y = a + bX + E$$

x_i	y_i	$x_i y_i$	x_i^2	y_i^2
18	17	306	324	289
18	19	342	324	361
20	20	400	400	400
20	22	440	400	484
20	24	480	400	576
96	102	1968	1848	2110
19.2	20.4	393.6	369.6	422
$M(X)$	$M(Y)$	$M(XY)$	$M(X^2)$	$M(Y^2)$

$$\begin{aligned}\hat{b} &= \frac{\text{Cov}(X, Y)}{\text{Var}(X)} = \frac{M(XY) - M(X)M(Y)}{M(X^2) - [M(X)]^2} = \\ &= \frac{393.6 - 19.2 \cdot 20.4}{369.6 - 19.2^2} = \frac{1.92}{0.96} = 2 \\ \hat{a} &= \end{aligned}$$



Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.

¹Si osserva che per ragioni di stampa alcuni calcoli potrebbero essere duplicati con riferimento ai diversi modelli oggetto di studio

Si riportano i calcoli necessari¹ per ottenere le stime dei parametri e gli indici di adattamento del modello

$$Y = a + bX + E$$

x_i	y_i	$x_i y_i$	x_i^2	y_i^2
18	17	306	324	289
18	19	342	324	361
20	20	400	400	400
20	22	440	400	484
20	24	480	400	576
96	102	1968	1848	2110
19.2	20.4	393.6	369.6	422
$M(X)$	$M(Y)$	$M(XY)$	$M(X^2)$	$M(Y^2)$

$$\hat{b} = \frac{\text{Cov}(X, Y)}{\text{Var}(X)} = \frac{M(XY) - M(X)M(Y)}{M(X^2) - [M(X)]^2} =$$

$$= \frac{393.6 - 19.2 \cdot 20.4}{369.6 - 19.2^2} = \frac{1.92}{0.96} = 2$$

$$\hat{a} = M(Y) - \hat{b}M(X) = -18$$

$$\rho^2 =$$



Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.

¹Si osserva che per ragioni di stampa alcuni calcoli potrebbero essere duplicati con riferimento ai diversi modelli oggetto di studio

Si riportano i calcoli necessari¹ per ottenere le stime dei parametri e gli indici di adattamento del modello

$$Y = a + bX + E$$

x_i	y_i	$x_i y_i$	x_i^2	y_i^2
18	17	306	324	289
18	19	342	324	361
20	20	400	400	400
20	22	440	400	484
20	24	480	400	576
96	102	1968	1848	2110
19.2	20.4	393.6	369.6	422
$M(X)$	$M(Y)$	$M(XY)$	$M(X^2)$	$M(Y^2)$

$$\hat{b} = \frac{\text{Cov}(X, Y)}{\text{Var}(X)} = \frac{M(XY) - M(X)M(Y)}{M(X^2) - [M(X)]^2} =$$

$$= \frac{393.6 - 19.2 \cdot 20.4}{369.6 - 19.2^2} = \frac{1.92}{0.96} = 2$$

$$\hat{a} = M(Y) - \hat{b}M(X) = -18$$

$$\rho^2 = \frac{[\text{Cov}(X, Y)]^2}{\text{Var}(X) \text{Var}(Y)} = \frac{1.92^2}{0.96 \cdot 5.84} = 0.6575$$

$$M(\hat{E}^2) = \text{Var}(\hat{E}) = \text{varianza residua} =$$

¹Si osserva che per ragioni di stampa alcuni calcoli potrebbero essere duplicati con riferimento ai diversi modelli oggetto di studio



Si riportano i calcoli necessari¹ per ottenere le stime dei parametri e gli indici di adattamento del modello

$$Y = a + bX + E$$

x_i	y_i	$x_i y_i$	x_i^2	y_i^2
18	17	306	324	289
18	19	342	324	361
20	20	400	400	400
20	22	440	400	484
20	24	480	400	576
96	102	1968	1848	2110
19.2	20.4	393.6	369.6	422
$M(X)$	$M(Y)$	$M(XY)$	$M(X^2)$	$M(Y^2)$

$$\hat{b} = \frac{\text{Cov}(X, Y)}{\text{Var}(X)} = \frac{M(XY) - M(X)M(Y)}{M(X^2) - [M(X)]^2} =$$

$$= \frac{393.6 - 19.2 \cdot 20.4}{369.6 - 19.2^2} = \frac{1.92}{0.96} = 2$$

$$\hat{a} = M(Y) - \hat{b}M(X) = -18$$

$$\rho^2 = \frac{[\text{Cov}(X, Y)]^2}{\text{Var}(X) \text{Var}(Y)} = \frac{1.92^2}{0.96 \cdot 5.84} = 0.6575$$

$$M(\hat{E}^2) = \text{Var}(\hat{E}) = \text{varianza residua} = \text{Var}(Y) \cdot (1 - \rho^2) = 2$$

¹Si osserva che per ragioni di stampa alcuni calcoli potrebbero essere duplicati con riferimento ai diversi modelli oggetto di studio





Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.

Si riportano i calcoli necessari² per ottenere le stime dei parametri e la misura di adattamento del modello

$$Y = cX + E$$

x_i	y_i	$x_i y_i$	x_i^2	y_i^2
18	17	306	324	289
18	19	342	324	361
20	20	400	400	400
20	22	440	400	484
20	24	480	400	576
		1968	1848	2110
		393.6	369.6	422
		$M(XY)$	$M(X^2)$	$M(Y^2)$

$$\hat{c} =$$

²Si osserva che per ragioni di stampa alcuni calcoli potrebbero essere duplicati con riferimento ai diversi modelli oggetto di studio



Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.

Si riportano i calcoli necessari² per ottenere le stime dei parametri e la misura di adattamento del modello

$$Y = cX + E$$

x_i	y_i	$x_i y_i$	x_i^2	y_i^2
18	17	306	324	289
18	19	342	324	361
20	20	400	400	400
20	22	440	400	484
20	24	480	400	576
		1968	1848	2110
		393.6	369.6	422
		$M(XY)$	$M(X^2)$	$M(Y^2)$

$$\hat{c} = \frac{M(XY)}{M(X^2)} = \frac{393.6}{369.6} = 1.0649$$

$$M(\hat{E}^2) =$$

²Si osserva che per ragioni di stampa alcuni calcoli potrebbero essere duplicati con riferimento ai diversi modelli oggetto di studio



Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.

Si riportano i calcoli necessari² per ottenere le stime dei parametri e la misura di adattamento del modello

$$Y = cX + E$$

x_i	y_i	$x_i y_i$	x_i^2	y_i^2
18	17	306	324	289
18	19	342	324	361
20	20	400	400	400
20	22	440	400	484
20	24	480	400	576
		1968	1848	2110
		393.6	369.6	422
		$M(XY)$	$M(X^2)$	$M(Y^2)$

$$\hat{c} = \frac{M(XY)}{M(X^2)} = \frac{393.6}{369.6} = 1.0649$$

$$M(\hat{E}^2) = M[Y \cdot (Y - \hat{c}X)] = M(Y^2) - \hat{c}M(XY) = 422 - 1.0649 \cdot 393.6 = 2.8416$$

²Si osserva che per ragioni di stampa alcuni calcoli potrebbero essere duplicati con riferimento ai diversi modelli oggetto di studio



$a = -18, b = 2$, indice adattamento $M(\hat{E}) = 2$ $c = 1.0649$, indice adattamento

$M(\hat{E}^2) = 2.8416$

Il modello migliore è $Y^* = -18 + 2X$.

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.

Esercizio 4

Prova scritta di
STATISTICA del
09.07.2015

Giuseppe Boari
Gabriele Cantaluppi



UNIVERSITÀ
CATTOLICA
del Sacro Cuore

Si consideri il seguente esperimento. Si lancia una moneta (equilibrata). Se esce testa si effettuano 2 ulteriori lanci della moneta; nel caso esca croce si effettua 1 solo ulteriore lancio.

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.

1. Si costruisca la variabile casuale $T =$ 'n. di teste ottenute alla fine dell'esperimento'.

Prova scritta di
STATISTICA del
09.07.2015

Giuseppe Boari
Gabriele Cantaluppi



UNIVERSITÀ
CATTOLICA
del Sacro Cuore

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.

1. Si costruisca la variabile casuale $T =$ 'n. di teste ottenute alla fine dell'esperimento'. Indicate con T_1 e T_2 le variabili casuali che descrivono il numero di teste ottenute nel primo lancio della moneta e nella seconda fase dell'esperimento abbiamo

$$P(T = 0) =$$



Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.

1. Si costruisca la variabile casuale $T =$ 'n. di teste ottenute alla fine dell'esperimento'. Indicate con T_1 e T_2 le variabili casuali che descrivono il numero di teste ottenute nel primo lancio della moneta e nella seconda fase dell'esperimento abbiamo

$$P(T = 0) = P(T_1 = 0) \cdot P(T_2 = 0 | T_1 = 0) = 0.5 \cdot 0.5$$

$$P(T = 1) =$$



Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.

1. Si costruisca la variabile casuale $T =$ 'n. di teste ottenute alla fine dell'esperimento'. Indicate con T_1 e T_2 le variabili casuali che descrivono il numero di teste ottenute nel primo lancio della moneta e nella seconda fase dell'esperimento abbiamo

$$P(T = 0) = P(T_1 = 0) \cdot P(T_2 = 0 | T_1 = 0) = 0.5 \cdot 0.5$$

$$P(T = 1) = P(T_1 = 0) \cdot P(T_2 = 1 | T_1 = 0) + P(T_1 = 1) \cdot P(T_2 = 0 | T_1 = 1) =$$
$$=$$



1. Si costruisca la variabile casuale $T =$ 'n. di teste ottenute alla fine dell'esperimento'. Indicate con T_1 e T_2 le variabili casuali che descrivono il numero di teste ottenute nel primo lancio della moneta e nella seconda fase dell'esperimento abbiamo

$$P(T = 0) = P(T_1 = 0) \cdot P(T_2 = 0 | T_1 = 0) = 0.5 \cdot 0.5$$

$$\begin{aligned} P(T = 1) &= P(T_1 = 0) \cdot P(T_2 = 1 | T_1 = 0) + P(T_1 = 1) \cdot P(T_2 = 0 | T_1 = 1) = \\ &= 0.5 \cdot 0.5 + 0.5 \cdot \binom{2}{0} 0.5^0 0.5^2 = \\ &= \end{aligned}$$



1. Si costruisca la variabile casuale $T =$ 'n. di teste ottenute alla fine dell'esperimento'. Indicate con T_1 e T_2 le variabili casuali che descrivono il numero di teste ottenute nel primo lancio della moneta e nella seconda fase dell'esperimento abbiamo

$$P(T = 0) = P(T_1 = 0) \cdot P(T_2 = 0 | T_1 = 0) = 0.5 \cdot 0.5$$

$$\begin{aligned} P(T = 1) &= P(T_1 = 0) \cdot P(T_2 = 1 | T_1 = 0) + P(T_1 = 1) \cdot P(T_2 = 0 | T_1 = 1) = \\ &= 0.5 \cdot 0.5 + 0.5 \cdot \binom{2}{0} 0.5^0 0.5^2 = \\ &= 0.25 + 0.5 \cdot 0.25 = \end{aligned}$$



1. Si costruisca la variabile casuale $T =$ 'n. di teste ottenute alla fine dell'esperimento'. Indicate con T_1 e T_2 le variabili casuali che descrivono il numero di teste ottenute nel primo lancio della moneta e nella seconda fase dell'esperimento abbiamo

$$P(T = 0) = P(T_1 = 0) \cdot P(T_2 = 0 | T_1 = 0) = 0.5 \cdot 0.5$$

$$\begin{aligned} P(T = 1) &= P(T_1 = 0) \cdot P(T_2 = 1 | T_1 = 0) + P(T_1 = 1) \cdot P(T_2 = 0 | T_1 = 1) = \\ &= 0.5 \cdot 0.5 + 0.5 \cdot \binom{2}{0} 0.5^0 0.5^2 = \\ &= 0.25 + 0.5 \cdot 0.25 = 0.25 + 0.125 = 0.375 \end{aligned}$$

$$P(T = 2) =$$



1. Si costruisca la variabile casuale $T =$ 'n. di teste ottenute alla fine dell'esperimento'. Indicate con T_1 e T_2 le variabili casuali che descrivono il numero di teste ottenute nel primo lancio della moneta e nella seconda fase dell'esperimento abbiamo

$$P(T = 0) = P(T_1 = 0) \cdot P(T_2 = 0 | T_1 = 0) = 0.5 \cdot 0.5$$

$$\begin{aligned} P(T = 1) &= P(T_1 = 0) \cdot P(T_2 = 1 | T_1 = 0) + P(T_1 = 1) \cdot P(T_2 = 0 | T_1 = 1) = \\ &= 0.5 \cdot 0.5 + 0.5 \cdot \binom{2}{0} 0.5^0 0.5^2 = \\ &= 0.25 + 0.5 \cdot 0.25 = 0.25 + 0.125 = 0.375 \end{aligned}$$

$$\begin{aligned} P(T = 2) &= P(T_1 = 1) \cdot P(T_2 = 1 | T_1 = 1) = \\ &= \end{aligned}$$



1. Si costruisca la variabile casuale $T =$ 'n. di teste ottenute alla fine dell'esperimento'. Indicate con T_1 e T_2 le variabili casuali che descrivono il numero di teste ottenute nel primo lancio della moneta e nella seconda fase dell'esperimento abbiamo

$$P(T = 0) = P(T_1 = 0) \cdot P(T_2 = 0 | T_1 = 0) = 0.5 \cdot 0.5$$

$$\begin{aligned} P(T = 1) &= P(T_1 = 0) \cdot P(T_2 = 1 | T_1 = 0) + P(T_1 = 1) \cdot P(T_2 = 0 | T_1 = 1) = \\ &= 0.5 \cdot 0.5 + 0.5 \cdot \binom{2}{0} 0.5^0 0.5^2 = \\ &= 0.25 + 0.5 \cdot 0.25 = 0.25 + 0.125 = 0.375 \end{aligned}$$

$$\begin{aligned} P(T = 2) &= P(T_1 = 1) \cdot P(T_2 = 1 | T_1 = 1) = \\ &= 0.5 \cdot \binom{2}{1} 0.5^1 0.5^1 = \end{aligned}$$



1. Si costruisca la variabile casuale $T =$ 'n. di teste ottenute alla fine dell'esperimento'. Indicate con T_1 e T_2 le variabili casuali che descrivono il numero di teste ottenute nel primo lancio della moneta e nella seconda fase dell'esperimento abbiamo

$$P(T = 0) = P(T_1 = 0) \cdot P(T_2 = 0 | T_1 = 0) = 0.5 \cdot 0.5$$

$$\begin{aligned} P(T = 1) &= P(T_1 = 0) \cdot P(T_2 = 1 | T_1 = 0) + P(T_1 = 1) \cdot P(T_2 = 0 | T_1 = 1) = \\ &= 0.5 \cdot 0.5 + 0.5 \cdot \binom{2}{0} 0.5^0 0.5^2 = \\ &= 0.25 + 0.5 \cdot 0.25 = 0.25 + 0.125 = 0.375 \end{aligned}$$

$$\begin{aligned} P(T = 2) &= P(T_1 = 1) \cdot P(T_2 = 1 | T_1 = 1) = \\ &= 0.5 \cdot \binom{2}{1} 0.5^1 0.5^1 = 0.5 \cdot 2 \cdot 0.25 = 0.5 \cdot 0.5 = 0.25 \end{aligned}$$

$$P(T = 3) =$$



1. Si costruisca la variabile casuale $T =$ 'n. di teste ottenute alla fine dell'esperimento'. Indicate con T_1 e T_2 le variabili casuali che descrivono il numero di teste ottenute nel primo lancio della moneta e nella seconda fase dell'esperimento abbiamo

$$P(T = 0) = P(T_1 = 0) \cdot P(T_2 = 0 | T_1 = 0) = 0.5 \cdot 0.5$$

$$\begin{aligned} P(T = 1) &= P(T_1 = 0) \cdot P(T_2 = 1 | T_1 = 0) + P(T_1 = 1) \cdot P(T_2 = 0 | T_1 = 1) = \\ &= 0.5 \cdot 0.5 + 0.5 \cdot \binom{2}{0} 0.5^0 0.5^2 = \\ &= 0.25 + 0.5 \cdot 0.25 = 0.25 + 0.125 = 0.375 \end{aligned}$$

$$\begin{aligned} P(T = 2) &= P(T_1 = 1) \cdot P(T_2 = 1 | T_1 = 1) = \\ &= 0.5 \cdot \binom{2}{1} 0.5^1 0.5^1 = 0.5 \cdot 2 \cdot 0.25 = 0.5 \cdot 0.5 = 0.25 \end{aligned}$$

$$\begin{aligned} P(T = 3) &= P(T_1 = 1) \cdot P(T_2 = 2 | T_1 = 1) \\ &= \end{aligned}$$



1. Si costruisca la variabile casuale $T =$ 'n. di teste ottenute alla fine dell'esperimento'. Indicate con T_1 e T_2 le variabili casuali che descrivono il numero di teste ottenute nel primo lancio della moneta e nella seconda fase dell'esperimento abbiamo

$$P(T = 0) = P(T_1 = 0) \cdot P(T_2 = 0 | T_1 = 0) = 0.5 \cdot 0.5$$

$$\begin{aligned} P(T = 1) &= P(T_1 = 0) \cdot P(T_2 = 1 | T_1 = 0) + P(T_1 = 1) \cdot P(T_2 = 0 | T_1 = 1) = \\ &= 0.5 \cdot 0.5 + 0.5 \cdot \binom{2}{0} 0.5^0 0.5^2 = \\ &= 0.25 + 0.5 \cdot 0.25 = 0.25 + 0.125 = 0.375 \end{aligned}$$

$$\begin{aligned} P(T = 2) &= P(T_1 = 1) \cdot P(T_2 = 1 | T_1 = 1) = \\ &= 0.5 \cdot \binom{2}{1} 0.5^1 0.5^1 = 0.5 \cdot 2 \cdot 0.25 = 0.5 \cdot 0.5 = 0.25 \end{aligned}$$

$$\begin{aligned} P(T = 3) &= P(T_1 = 1) \cdot P(T_2 = 2 | T_1 = 1) \\ &= 0.5 \cdot \binom{2}{2} 0.5^2 0.5^0 = 0.5 \cdot 0.25 = 0.125 \end{aligned}$$



e, quindi,



Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.



e, quindi,

t	$P(T = t)$
0	0.25
1	0.375
2	0.25
3	0.125
	1

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.



2. Si calcolino media e varianza della variabile casuale descritta dall'esperimento al punto precedente.

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.



2. Si calcolino media e varianza della variabile casuale descritta dall'esperimento al punto precedente.

$$M(T) = \sum_{t=0}^3 t \cdot P(T = t) = 1.25$$

$$\text{Var}(T) = \sum_{t=0}^3 \{t - M(t)\}^2 \cdot P(T = t) = \sum_{t=0}^3 t^2 \cdot P(T = t) - \{M(t)\}^2 = 0.9375$$

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.



3. Si calcoli la probabilità che, lanciando 1000 volte la moneta si ottengano almeno 500 teste.

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.



Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.

3. Si calcoli la probabilità che, lanciando 1000 volte la moneta si ottengano almeno 500 teste.

Indicata con X la variabile casuale 'n. di teste nei 1000 lanci (prove indipendenti)' risulta

$$X \sim$$



3. Si calcoli la probabilità che, lanciando 1000 volte la moneta si ottengano almeno 500 teste.

Indicata con X la variabile casuale 'n. di teste nei 1000 lanci (prove indipendenti)' risulta

$$X \sim \text{Bin}(n = 1000, p = 0.5).$$

$M(X) = 500$; $\text{Var}(X) = 250$. Dal momento che $np = 500 > 5$ ed $n(1 - p) = 500 > 5$ è possibile utilizzare l'approssimazione della variabile casuale Binomiale con la variabile casuale Normale

$$X \sim N(\mu = 500, \sigma^2 = 250).$$

Si ottiene

$$P(X \geq 500) =$$



3. Si calcoli la probabilità che, lanciando 1000 volte la moneta si ottengano almeno 500 teste.

Indicata con X la variabile casuale 'n. di teste nei 1000 lanci (prove indipendenti)' risulta

$$X \sim \text{Bin}(n = 1000, p = 0.5).$$

$M(X) = 500$; $\text{Var}(X) = 250$. Dal momento che $np = 500 > 5$ ed $n(1 - p) = 500 > 5$ è possibile utilizzare l'approssimazione della variabile casuale Binomiale con la variabile casuale Normale

$$X \sim N(\mu = 500, \sigma^2 = 250).$$

Si ottiene

$$P(X \geq 500) = P(X \geq \mu) =$$



3. Si calcoli la probabilità che, lanciando 1000 volte la moneta si ottengano almeno 500 teste.

Indicata con X la variabile casuale 'n. di teste nei 1000 lanci (prove indipendenti)' risulta

$$X \sim \text{Bin}(n = 1000, p = 0.5).$$

$M(X) = 500$; $\text{Var}(X) = 250$. Dal momento che $np = 500 > 5$ ed $n(1 - p) = 500 > 5$ è possibile utilizzare l'approssimazione della variabile casuale Binomiale con la variabile casuale Normale

$$X \sim N(\mu = 500, \sigma^2 = 250).$$

Si ottiene

$$P(X \geq 500) = P(X \geq \mu) = 0.5$$

essendo la distribuzione Normale simmetrica.

Tema T285-1

Prova scritta di STATISTICA del 01.09.2015

Prova scritta di
STATISTICA del
01.09.2015

Giuseppe Boari
Gabriele Cantaluppi



UNIVERSITÀ
CATTOLICA
del Sacro Cuore

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.

Giuseppe Boari
Gabriele Cantaluppi
Università Cattolica del Sacro Cuore, Milano



Il seguente prospetto riporta la distribuzione dei clienti presenti su un mercato dove opera l'agenzia A, riclassificati mediante le variabili X , con categorie x_1 = 'clienti agenzia A' e x_2 = 'clienti della concorrenza', e Y = 'tipologia di viaggio', con categorie y_1 = 'low cost', y_2 = 'medium cost' e y_3 = 'high cost'.

Nella soluzione di questo esercizio si tenga presente che la clientela *target* dell'agenzia A è del tipo 'high cost'.

$X \setminus Y$	y_1	y_2	y_3
x_1	780	375	345
x_2	4420	2125	1955

1. Si rappresentino graficamente le distribuzioni marginali X e Y .

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.



Si considera la distribuzione marginale della variabile X , ottenuta sommando le frequenze di riga

x_i	$n_{i\bullet}$
x_1	1500
x_2	8500

Si tratta di un carattere qualitativo sconnesso che può essere rappresentato graficamente mediante un grafico a bastoncini o a canne d'organo.

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.

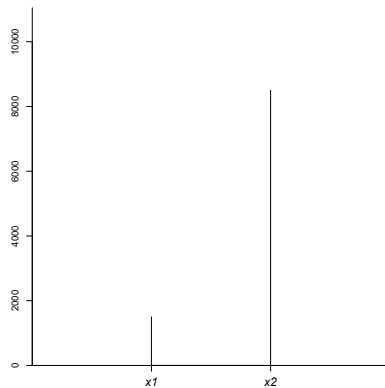


Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.





Si considera la distribuzione marginale della variabile Y , ottenuta sommando le frequenze di colonna

y_j	$n_{\bullet j}$
y_1	5200
y_2	2500
y_3	2300

Si tratta di un carattere qualitativo ordinato che può essere rappresentato graficamente mediante un grafico a bastoncini o a canne d'organo.

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.

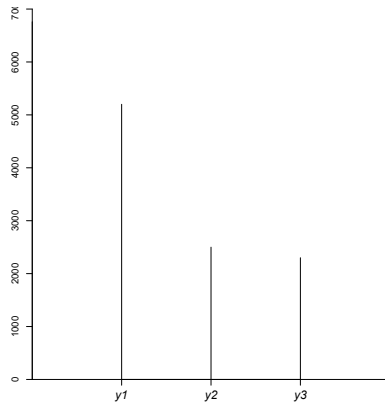


Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.





2. Si confrontino le distribuzioni condizionate $Y|x_1$ e $Y|x_2$, calcolando moda, mediana, media aritmetica, e una opportuna misura di variabilità.

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.



Si considera la distribuzione condizionata $Y|x_1$

y_j	n_{1j}
y_1	780
y_2	375
y_3	345

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.



Si tratta di un carattere qualitativo ordinato.

La moda corrisponde alla categoria alla quale è associata la massima frequenza) che risulta y_1 .

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.



La mediana (percentile di ordine 0.50) in presenza di un numero pari di osservazioni corrisponde a uno degli elementi che occupano le posizioni $\frac{n}{2} = 750$ ed $\frac{n}{2} + 1 = 751$ nella serie dei dati ordinati. Dalla distribuzione delle frequenze cumulate

y_j	n_{1j}	N_{1j}
y_1	780	780
y_2	375	1155
y_3	345	1500

i due elementi coincidono e risulta $y_{0.5} = y_1$.

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.



In presenza di un carattere qualitativo la variabilità può essere misurata mediante gli indici di Gini e di Frosini normalizzati

$$G_N = \frac{1 - \sum_{i=1}^k f_{ij}^2}{1 - \frac{1}{k}}$$

e

$$F_N = 1 - \sqrt{1 - G_N}.$$

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.



In presenza di un carattere qualitativo la variabilità può essere misurata mediante gli indici di Gini e di Frosini normalizzati

$$G_N = \frac{1 - \sum_{i=1}^k f_{ij}^2}{1 - \frac{1}{k}}$$

e

$$F_N = 1 - \sqrt{1 - G_N}.$$

Si ottiene:

y_j	n_{1j}	f_{1j}	f_{1j}^2
y_1	780	0.52	0.2704
y_2	375	0.25	0.0625
y_3	345	0.23	0.0529
	1500	1	0.3858

Da cui segue: $G_N = 0.9213$ e $F_N = 0.7195$. Sussiste, quindi, un livello medio alto di eterogeneità.



Si considera la distribuzione condizionata $Y|x_2$

y_j	n_{2j}
y_1	4420
y_2	2125
y_3	1955

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.



Si tratta di un carattere qualitativo ordinato.

La moda corrisponde alla categoria alla quale è associata la massima frequenza) che risulta y_1 .

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.



La mediana (percentile di ordine 0.50) in presenza di un numero pari di osservazioni corrisponde a uno degli elementi che occupano le posizioni $\frac{n}{2} = 4250$ ed $\frac{n}{2} + 1 = 4251$ nella serie dei dati ordinati. Dalla distribuzione delle frequenze cumulate

y_j	n_{2j}	N_{2j}
y_1	4420	4420
y_2	2125	6545
y_3	1955	8500

i due elementi coincidono e risulta $y_{0.5} = y_1$.

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.



In presenza di un carattere qualitativo la variabilità può essere misurata mediante gli indici di Gini e di Frosini normalizzati

$$G_N = \frac{1 - \sum_{i=1}^k f_{2j}^2}{1 - \frac{1}{k}}$$

e

$$F_N = 1 - \sqrt{1 - G_N}.$$

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.



In presenza di un carattere qualitativo la variabilità può essere misurata mediante gli indici di Gini e di Frosini normalizzati

$$G_N = \frac{1 - \sum_{i=1}^k f_{2j}^2}{1 - \frac{1}{k}}$$

e

$$F_N = 1 - \sqrt{1 - G_N}.$$

Si ottiene:

y_j	n_{2j}	f_{2j}	f_{2j}^2
y_1	4420	0.52	0.2704
y_2	2125	0.25	0.0625
y_3	1955	0.23	0.0529
	8500	1	0.3858

Da cui segue: $G_N = 0.9213$ e $F_N = 0.7195$. Sussiste, quindi, un livello medio alto di eterogeneità.



$Y|x_1$: moda: $y_1 = 'low\ cost'$ mediana: $y_1 = 'low\ cost'$ media aritmetica: non è definita per un carattere qualitativo
Indice di variabilità: $F_N = 0.7195$

$Y|x_2$: moda: $y_1 = 'low\ cost'$ mediana: $y_1 = 'low\ cost'$ media aritmetica: non è definita per un carattere qualitativo
Indice di variabilità: $F_N = 0.7195$

Sia l'agenzia A che la concorrenza sono principalmente focalizzate su tipologie di viaggio *'low cost'*. Le distribuzioni delle frequenze condizionate relative, utilizzate per il calcolo dell'indice di Gini, sono identiche; si evince, quindi, che tra le variabili X e Y sussiste indipendenza stocastica. La struttura della clientela dell'agenzia A coincide con quella della concorrenza.

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.



3. Si studi il posizionamento dell'agenzia *A* sul mercato utilizzando le contingenze assolute e le contingenze relative.

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.



3. Si studi il posizionamento dell'agenzia A sul mercato utilizzando le contingenze assolute e le contingenze relative.

contingenze assolute	$X \setminus Y$	y_1	y_2	y_3
	x_1	0	0	0
	x_2	0	0	0

contingenze relative	$X \setminus Y$	y_1	y_2	y_3
	x_1	0	0	0
	x_2	0	0	0

Dal momento che sussiste indipendenza stocastica le frequenze congiunte n_{ij} coincidono con le frequenze teoriche $\hat{n}_{ij}$; quindi le contingenze assolute $n_{ij} - \hat{n}_{ij}$ sono nulle, i rapporti di contingenza $\frac{n_{ij}}{\hat{n}_{ij}}$ assumono valore unitario e le contingenze relative $\frac{n_{ij} - \hat{n}_{ij}}{\hat{n}_{ij}} = \frac{n_{ij}}{\hat{n}_{ij}} - 1$ sono pure nulle.

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.



4. Si indichi se esiste connessione tra le variabili X e Y .

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.



4. Si indichi se esiste connessione tra le variabili X e Y .

Dal momento che sussiste indipendenza stocastica non è presente connessione tra le variabili X e Y .

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.

Esercizio 2

Prova scritta di
STATISTICA del
01.09.2015

Giuseppe Boari
Gabriele Cantaluppi



UNIVERSITÀ
CATTOLICA
del Sacro Cuore

Si consideri la seguente serie X_t dei prezzi del prodotto A:

t	X_t
2010	90
2011	92
2012	98
2013	103
2014	101

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.

1. Si costruiscano le corrispondenti serie dei numeri indici a base fissa, con base $2010 = 1$, e a base mobile.



Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.

1. Si costruiscano le corrispondenti serie dei numeri indici a base fissa, con base $2010 = 1$, e a base mobile.

Si ricorda che ${}_{2010}I_t = \frac{x_t}{x_{2010}}$.



Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.



Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.

1. Si costruiscano le corrispondenti serie dei numeri indici a base fissa, con base $2010 = 1$, e a base mobile.

Si ricorda che ${}_{2010}I_t = \frac{x_t}{x_{2010}}$. Si ottiene quindi

Anno	x_t	NIBF (2010)
2010	90	1.0000
2011	92	1.0222
2012	98	1.0889
2013	103	1.1444
2014	101	1.1222



Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.

1. Si costruiscano le corrispondenti serie dei numeri indici a base fissa, con base $2010 = 1$, e a base mobile.

Si ricorda che ${}_{2010}I_t = \frac{x_t}{x_{2010}}$. Si ottiene quindi

Anno	x_t	NIBF (2010)
2010	90	1.0000
2011	92	1.0222
2012	98	1.0889
2013	103	1.1444
2014	101	1.1222

Si ricorda che ${}_{t-1}I_t = \frac{x_t}{x_{t-1}}$.



Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.

1. Si costruiscano le corrispondenti serie dei numeri indici a base fissa, con base 2010 = 1, e a base mobile.

Si ricorda che ${}_{2010}I_t = \frac{x_t}{x_{2010}}$. Si ottiene quindi

Anno	x_t	NIBF (2010)
2010	90	1.0000
2011	92	1.0222
2012	98	1.0889
2013	103	1.1444
2014	101	1.1222

Si ricorda che ${}_{t-1}I_t = \frac{x_t}{x_{t-1}}$. Si ottiene quindi

Anno	x_t	NIBM
2010	90	
2011	92	1.0222
2012	98	1.0652
2013	103	1.0510
2014	101	0.9806

2. Sapendo che $x_{2016} = 261$, si calcoli il tasso medio di variazione tra il 2010 e il 2016 e si commenti il risultato ottenuto.



Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.



2. Sapendo che $x_{2016} = 261$, si calcoli il tasso medio di variazione tra il 2010 e il 2016 e si commenti il risultato ottenuto.

Si ricorda che

$${}_{2010}I_{2016} = {}_{2010}I_{2011} \cdot {}_{2011}I_{2012} \cdot {}_{2012}I_{2013} \cdot {}_{2013}I_{2014} \cdot {}_{2014}I_{2015} \cdot {}_{2015}I_{2016}$$

Si deve trovare il numero indice a base mobile medio.

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.



Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.

2. Sapendo che $x_{2016} = 261$, si calcoli il tasso medio di variazione tra il 2010 e il 2016 e si commenti il risultato ottenuto.

Si ricorda che

$${}_{2010}I_{2016} = {}_{2010}I_{2011} \cdot {}_{2011}I_{2012} \cdot {}_{2012}I_{2013} \cdot {}_{2013}I_{2014} \cdot {}_{2014}I_{2015} \cdot {}_{2015}I_{2016}$$

Si deve trovare il numero indice a base mobile medio. Applicando il criterio di scelta della media secondo Chisini si ottiene

$${}_{2010}I_{2016} = {}_{2010}I_{2011} \cdot {}_{2011}I_{2012} \cdot {}_{2012}I_{2013} \cdot {}_{2013}I_{2014} \cdot {}_{2014}I_{2015} \cdot {}_{2015}I_{2016} = \alpha \cdot \alpha \cdot \alpha \cdot \alpha \cdot \alpha \cdot \alpha = \alpha^6.$$



Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.

2. Sapendo che $x_{2016} = 261$, si calcoli il tasso medio di variazione tra il 2010 e il 2016 e si commenti il risultato ottenuto.

Si ricorda che

$${}_{2010}I_{2016} = {}_{2010}I_{2011} \cdot {}_{2011}I_{2012} \cdot {}_{2012}I_{2013} \cdot {}_{2013}I_{2014} \cdot {}_{2014}I_{2015} \cdot {}_{2015}I_{2016}$$

Si deve trovare il numero indice a base mobile medio. Applicando il criterio di scelta della media secondo Chisini si ottiene

$${}_{2010}I_{2016} = {}_{2010}I_{2011} \cdot {}_{2011}I_{2012} \cdot {}_{2012}I_{2013} \cdot {}_{2013}I_{2014} \cdot {}_{2014}I_{2015} \cdot {}_{2015}I_{2016} = \alpha \cdot \alpha \cdot \alpha \cdot \alpha \cdot \alpha \cdot \alpha = \alpha^6.$$

Quindi

$$\alpha = \sqrt[6]{{}_{2010}I_{2016}} = \sqrt[6]{{}_{2010}I_{2011} \cdot {}_{2011}I_{2012} \cdot {}_{2012}I_{2013} \cdot {}_{2013}I_{2014} \cdot {}_{2014}I_{2015} \cdot {}_{2015}I_{2016}} = 1.1942.$$

Da cui segue una variazione relativa media pari a 0.1942, corrispondente a un incremento medio annuo tra il 2010 e il 2016 del 19.42%.



Il seguente prospetto riporta i valori rilevati in corrispondenza a 5 unità statistiche con riferimento a due caratteri quantitativi X e Y

x_i	18	18	20	20	20
y_i	24	22	20	18	16

1. Si rappresentino graficamente le coppie dei dati ipotizzando Y in funzione di X .

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.



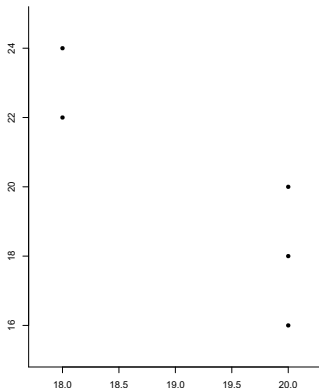
Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.

Le coppie di dati possono essere rappresentate mediante un grafico a dispersione.





2. Si calcolino, secondo il metodo dei minimi quadrati, i parametri e l'indice di adattamento dei seguenti modelli di regressione, individuando il modello migliore:

$$I) Y^* = a + b \cdot X \qquad II) Y^* = c \cdot X$$

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.

Si riportano i calcoli necessari¹ per ottenere le stime dei parametri e gli indici di adattamento del modello

$$Y = a + bX + E$$

x_i	y_i	$x_i y_i$	x_i^2	y_i^2
18	24	432	324	576
18	22	396	324	484
20	20	400	400	400
20	18	360	400	324
20	16	320	400	256
96	100	1908	1848	2040
19.2	20	381.6	369.6	408
$M(X)$	$M(Y)$	$M(XY)$	$M(X^2)$	$M(Y^2)$

$$\hat{b} =$$



Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.

¹Si osserva che per ragioni di stampa alcuni calcoli potrebbero essere duplicati con riferimento ai diversi modelli oggetto di studio

Si riportano i calcoli necessari¹ per ottenere le stime dei parametri e gli indici di adattamento del modello

$$Y = a + bX + E$$

x_i	y_i	$x_i y_i$	x_i^2	y_i^2
18	24	432	324	576
18	22	396	324	484
20	20	400	400	400
20	18	360	400	324
20	16	320	400	256
96	100	1908	1848	2040
19.2	20	381.6	369.6	408
$M(X)$	$M(Y)$	$M(XY)$	$M(X^2)$	$M(Y^2)$

$$\begin{aligned}\hat{b} &= \frac{\text{Cov}(X, Y)}{\text{Var}(X)} = \frac{M(XY) - M(X)M(Y)}{M(X^2) - [M(X)]^2} = \\ &= \frac{381.6 - 19.2 \cdot 20}{369.6 - 19.2^2} = \frac{-2.4}{0.96} = -2.5 \\ \hat{a} &= \end{aligned}$$



Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.

¹Si osserva che per ragioni di stampa alcuni calcoli potrebbero essere duplicati con riferimento ai diversi modelli oggetto di studio

Si riportano i calcoli necessari¹ per ottenere le stime dei parametri e gli indici di adattamento del modello

$$Y = a + bX + E$$

x_i	y_i	$x_i y_i$	x_i^2	y_i^2
18	24	432	324	576
18	22	396	324	484
20	20	400	400	400
20	18	360	400	324
20	16	320	400	256
96	100	1908	1848	2040
19.2	20	381.6	369.6	408
$M(X)$	$M(Y)$	$M(XY)$	$M(X^2)$	$M(Y^2)$

$$\hat{b} = \frac{\text{Cov}(X, Y)}{\text{Var}(X)} = \frac{M(XY) - M(X)M(Y)}{M(X^2) - [M(X)]^2} =$$

$$= \frac{381.6 - 19.2 \cdot 20}{369.6 - 19.2^2} = \frac{-2.4}{0.96} = -2.5$$

$$\hat{a} = M(Y) - \hat{b}M(X) = 68$$

$$\rho^2 =$$

¹Si osserva che per ragioni di stampa alcuni calcoli potrebbero essere duplicati con riferimento ai diversi modelli oggetto di studio



Si riportano i calcoli necessari¹ per ottenere le stime dei parametri e gli indici di adattamento del modello

$$Y = a + bX + E$$

x_i	y_i	$x_i y_i$	x_i^2	y_i^2
18	24	432	324	576
18	22	396	324	484
20	20	400	400	400
20	18	360	400	324
20	16	320	400	256
96	100	1908	1848	2040
19.2	20	381.6	369.6	408
$M(X)$	$M(Y)$	$M(XY)$	$M(X^2)$	$M(Y^2)$

$$\hat{b} = \frac{\text{Cov}(X, Y)}{\text{Var}(X)} = \frac{M(XY) - M(X)M(Y)}{M(X^2) - [M(X)]^2} =$$

$$= \frac{381.6 - 19.2 \cdot 20}{369.6 - 19.2^2} = \frac{-2.4}{0.96} = -2.5$$

$$\hat{a} = M(Y) - \hat{b}M(X) = 68$$

$$\rho^2 = \frac{[\text{Cov}(X, Y)]^2}{\text{Var}(X) \text{Var}(Y)} = \frac{(-2.4)^2}{0.96 \cdot 8} = 0.75$$

$$M(\hat{E}^2) = \text{Var}(\hat{E}) = \text{varianza residua} =$$

¹Si osserva che per ragioni di stampa alcuni calcoli potrebbero essere duplicati con riferimento ai diversi modelli oggetto di studio



Si riportano i calcoli necessari¹ per ottenere le stime dei parametri e gli indici di adattamento del modello

$$Y = a + bX + E$$

x_i	y_i	$x_i y_i$	x_i^2	y_i^2
18	24	432	324	576
18	22	396	324	484
20	20	400	400	400
20	18	360	400	324
20	16	320	400	256
96	100	1908	1848	2040
19.2	20	381.6	369.6	408
$M(X)$	$M(Y)$	$M(XY)$	$M(X^2)$	$M(Y^2)$

$$\begin{aligned}\hat{b} &= \frac{\text{Cov}(X, Y)}{\text{Var}(X)} = \frac{M(XY) - M(X)M(Y)}{M(X^2) - [M(X)]^2} = \\ &= \frac{381.6 - 19.2 \cdot 20}{369.6 - 19.2^2} = \frac{-2.4}{0.96} = -2.5 \\ \hat{a} &= M(Y) - \hat{b}M(X) = 68 \\ \rho^2 &= \frac{[\text{Cov}(X, Y)]^2}{\text{Var}(X) \text{Var}(Y)} = \frac{(-2.4)^2}{0.96 \cdot 8} = 0.75\end{aligned}$$

$$M(\hat{E}^2) = \text{Var}(\hat{E}) = \text{varianza residua} = \text{Var}(Y) \cdot (1 - \rho^2) = 2$$

¹Si osserva che per ragioni di stampa alcuni calcoli potrebbero essere duplicati con riferimento ai diversi modelli oggetto di studio





Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.

Si riportano i calcoli necessari² per ottenere le stime dei parametri e la misura di adattamento del modello

$$Y = cX + E$$

x_i	y_i	$x_i y_i$	x_i^2	y_i^2
18	24	432	324	576
18	22	396	324	484
20	20	400	400	400
20	18	360	400	324
20	16	320	400	256
		1908	1848	2040
		381.6	369.6	408
		$M(XY)$	$M(X^2)$	$M(Y^2)$

$$\hat{c} =$$

²Si osserva che per ragioni di stampa alcuni calcoli potrebbero essere duplicati con riferimento ai diversi modelli oggetto di studio



Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.

Si riportano i calcoli necessari² per ottenere le stime dei parametri e la misura di adattamento del modello

$$Y = cX + E$$

x_i	y_i	$x_i y_i$	x_i^2	y_i^2
18	24	432	324	576
18	22	396	324	484
20	20	400	400	400
20	18	360	400	324
20	16	320	400	256
		1908	1848	2040
		381.6	369.6	408
		$M(XY)$	$M(X^2)$	$M(Y^2)$

$$\hat{c} = \frac{M(XY)}{M(X^2)} = \frac{381.6}{369.6} = 1.0325$$

$$M(\hat{E}^2) =$$

²Si osserva che per ragioni di stampa alcuni calcoli potrebbero essere duplicati con riferimento ai diversi modelli oggetto di studio



Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.

Si riportano i calcoli necessari² per ottenere le stime dei parametri e la misura di adattamento del modello

$$Y = cX + E$$

x_i	y_i	$x_i y_i$	x_i^2	y_i^2
18	24	432	324	576
18	22	396	324	484
20	20	400	400	400
20	18	360	400	324
20	16	320	400	256
		1908	1848	2040
		381.6	369.6	408
		$M(XY)$	$M(X^2)$	$M(Y^2)$

$$\hat{c} = \frac{M(XY)}{M(X^2)} = \frac{381.6}{369.6} = 1.0325$$

$$M(\hat{E}^2) = M[Y \cdot (Y - \hat{c}X)] = M(Y^2) - \hat{c}M(XY) = 408 - 1.0325 \cdot 381.6 = 14.0104$$

²Si osserva che per ragioni di stampa alcuni calcoli potrebbero essere duplicati con riferimento ai diversi modelli oggetto di studio



Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.

$a = 68$, $b = -2.5$, indice adattamento $M(\hat{E}^2) = 2$ $c = 1.0325$, indice adattamento

$M(\hat{E}^2) = 14.0104$

Il modello migliore è $Y^* = 68 - 2.5X$.

Esercizio 4

Per il seguente gioco è prevista una posta di 3 Euro.

Si lancia una moneta (equilibrata).

Se esce testa si lancia un dado con 6 facce; nel caso esca croce si lancia un dado con 4 facce.

Si vince un ammontare in Euro pari al numero che si è presentato sulla base del dado lanciato.

1. Si stabilisca se il gioco è equo.



Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.



Per il seguente gioco è prevista una posta di 3 Euro.

Si lancia una moneta (equilibrata).

Se esce testa si lancia un dado con 6 facce; nel caso esca croce si lancia un dado con 4 facce.

Si vince un ammontare in Euro pari al numero che si è presentato sulla base del dado lanciato.

1. Si stabilisca se il gioco è equo.

Indicati con M , $D6$ e $D4$ i risultati del lancio della moneta e dei dadi a 6 e 4 facce, la variabile casuale V vinta al gioco risulta

v_j	$P(V = v_j)$
1	$P(M = T) \cdot P(D6 = 1 M = T) + P(M = C) \cdot P(D4 = 1 M = C) = \frac{1}{2} \cdot \frac{1}{6} + \frac{1}{2} \cdot \frac{1}{4} = \frac{1}{12} + \frac{1}{8} = \frac{5}{24}$
2	$P(M = T) \cdot P(D6 = 2 M = T) + P(M = C) \cdot P(D4 = 2 M = C) = \frac{1}{2} \cdot \frac{1}{6} + \frac{1}{2} \cdot \frac{1}{4} = \frac{1}{12} + \frac{1}{8} = \frac{5}{24}$
3	$P(M = T) \cdot P(D6 = 3 M = T) + P(M = C) \cdot P(D4 = 3 M = C) = \frac{1}{2} \cdot \frac{1}{6} + \frac{1}{2} \cdot \frac{1}{4} = \frac{1}{12} + \frac{1}{8} = \frac{5}{24}$
4	$P(M = T) \cdot P(D6 = 4 M = T) + P(M = C) \cdot P(D4 = 4 M = C) = \frac{1}{2} \cdot \frac{1}{6} + \frac{1}{2} \cdot \frac{1}{4} = \frac{1}{12} + \frac{1}{8} = \frac{5}{24}$
5	$P(M = T) \cdot P(D6 = 5 M = T) = \frac{1}{2} \cdot \frac{1}{6} = \frac{1}{12} = \frac{2}{24}$
6	$P(M = T) \cdot P(D6 = 6 M = T) = \frac{1}{2} \cdot \frac{1}{6} = \frac{1}{12} = \frac{2}{24}$



La variabile casuale guadagno $X = V - 3$ risulta

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.



La variabile casuale guadagno $X = V - 3$ risulta

x_i	$P(X = x_i)$	$x_i \cdot P(X = x_i)$
-2	$\frac{5}{24}$	$-\frac{10}{24}$
-1	$\frac{5}{24}$	$-\frac{5}{24}$
0	$\frac{5}{24}$	0
1	$\frac{5}{24}$	$\frac{5}{24}$
2	$\frac{2}{24}$	$\frac{4}{24}$
3	$\frac{2}{24}$	$\frac{6}{24}$
	1	0

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.



Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.

La variabile casuale guadagno $X = V - 3$ risulta

x_i	$P(X = x_i)$	$x_i \cdot P(X = x_i)$
-2	$\frac{5}{24}$	$-\frac{10}{24}$
-1	$\frac{5}{24}$	$-\frac{5}{24}$
0	$\frac{5}{24}$	0
1	$\frac{5}{24}$	$\frac{5}{24}$
2	$\frac{2}{24}$	$\frac{4}{24}$
3	$\frac{2}{24}$	$\frac{6}{24}$
	1	0

Il valore atteso della variabile casuale guadagno è nullo, quindi il gioco è equo.



2. Si calcoli la probabilità che, lanciando 9 volte la moneta si ottengano non più di 7 teste.

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.



Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.

2. Si calcoli la probabilità che, lanciando 9 volte la moneta si ottengano non più di 7 teste.

Sia X la variabile casuale 'n. di teste nelle 9 prove indipendenti'.

$$X \sim$$



Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.

2. Si calcoli la probabilità che, lanciando 9 volte la moneta si ottengano non più di 7 teste.

Sia X la variabile casuale 'n. di teste nelle 9 prove indipendenti'.

$$X \sim \text{Bin}(n = 9, p = 0.5).$$

$$P(X \leq 7) = 1 - P(X > 7) = 1 - [$$



2. Si calcoli la probabilità che, lanciando 9 volte la moneta si ottengano non più di 7 teste.

Sia X la variabile casuale 'n. di teste nelle 9 prove indipendenti'.

$$X \sim \text{Bin}(n = 9, p = 0.5).$$

$$\begin{aligned} P(X \leq 7) &= 1 - P(X > 7) = 1 - [P(X = 8) + P(X = 9)] = \\ &= 1 - \left[\binom{9}{8} 0.5^8 \cdot 0.5^1 + \binom{9}{9} 0.5^9 \cdot 0.5^0 \right] = \\ &= 1 - (0.01758 + 0.001953) = 1 - 0.019533 = 0.980467. \end{aligned}$$



Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.

3. Si calcoli la probabilità che, lanciando 900 volte la moneta si ottengano non più di 450 teste.



3. Si calcoli la probabilità che, lanciando 900 volte la moneta si ottengano non più di 450 teste.

Indicata con X la variabile casuale 'n. di teste nei 900 lanci (prove indipendenti)' risulta

$$X \sim$$

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.



3. Si calcoli la probabilità che, lanciando 900 volte la moneta si ottengano non più di 450 teste.

Indicata con X la variabile casuale 'n. di teste nei 900 lanci (prove indipendenti)' risulta

$$X \sim \text{Bin}(n = 900, p = 0.5).$$

$M(X) = 450$; $\text{Var}(X) = 225$. Dal momento che $np = 450 > 5$ ed $n(1 - p) = 450 > 5$ è possibile utilizzare l'approssimazione della variabile casuale Binomiale con la variabile casuale Normale

$$X \sim N(\mu = 450, \sigma^2 = 225).$$

Si ottiene

$$P(X \leq 450) =$$



3. Si calcoli la probabilità che, lanciando 900 volte la moneta si ottengano non più di 450 teste.

Indicata con X la variabile casuale 'n. di teste nei 900 lanci (prove indipendenti)' risulta

$$X \sim \text{Bin}(n = 900, p = 0.5).$$

$M(X) = 450$; $\text{Var}(X) = 225$. Dal momento che $np = 450 > 5$ ed $n(1 - p) = 450 > 5$ è possibile utilizzare l'approssimazione della variabile casuale Binomiale con la variabile casuale Normale

$$X \sim N(\mu = 450, \sigma^2 = 225).$$

Si ottiene

$$P(X \leq 450) = P(X \leq \mu) =$$



3. Si calcoli la probabilità che, lanciando 900 volte la moneta si ottengano non più di 450 teste.

Indicata con X la variabile casuale 'n. di teste nei 900 lanci (prove indipendenti)' risulta

$$X \sim \text{Bin}(n = 900, p = 0.5).$$

$M(X) = 450$; $\text{Var}(X) = 225$. Dal momento che $np = 450 > 5$ ed $n(1 - p) = 450 > 5$ è possibile utilizzare l'approssimazione della variabile casuale Binomiale con la variabile casuale Normale

$$X \sim N(\mu = 450, \sigma^2 = 225).$$

Si ottiene

$$P(X \leq 450) = P(X \leq \mu) = 0.5$$

essendo la distribuzione Normale simmetrica.

Tema T285

Prova scritta di STATISTICA del 15.09.2015

Prova scritta di
STATISTICA del
15.09.2015

Giuseppe Boari
Gabriele Cantaluppi



UNIVERSITÀ
CATTOLICA
del Sacro Cuore

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.

Giuseppe Boari
Gabriele Cantaluppi
Università Cattolica del Sacro Cuore, Milano



Per 8 sale cinematografiche vengono rilevate le seguenti variabili: F = fatturato medio settimanale (in migliaia di euro), S = spesa mensile per manutenzione sala e noleggio film (in migliaia di euro) e L = localizzazione della sala (C = centrale, P = periferica). I dati sono riportati di seguito:

F	15	6	11	16	18	14	16	12
S	8.2	3.4	6.2	12.5	13.2	7.2	11.2	5.2
L	C	C	P	C	P	C	C	P

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.



1. Si rappresenti graficamente la distribuzione della mutabile L .

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.



1. Si rappresenti graficamente la distribuzione della mutabile L .

l_i	n_i
C	5
P	3

Si tratta di un carattere qualitativo sconnesso che può essere rappresentato graficamente mediante un grafico a bastoncini.

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.

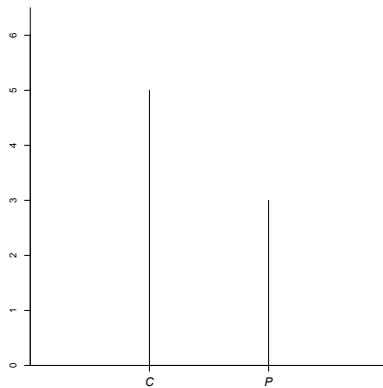


Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.





2. Si valuti, confrontando opportuni indici, la variabilità delle Spese e del Fatturato.

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.



2. Si valuti, confrontando opportuni indici, la variabilità delle Spese e del Fatturato. Per effettuare un confronto in termini di variabilità delle variabili Spese e Fatturato si considerano i coefficienti di variazione $CV(S)$ e $CV(F)$.

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.



Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.

2. Si valuti, confrontando opportuni indici, la variabilità delle Spese e del Fatturato. Per effettuare un confronto in termini di variabilità delle variabili Spese e Fatturato si considerano i coefficienti di variazione $CV(S)$ e $CV(F)$. Ricordando che il coefficiente di variazione è definito dal rapporto tra scarto quadratico medio e media aritmetica risulta

$$CV(S) = \frac{\sigma_S}{\mu_S} = \frac{\sqrt{11.1561}}{8.3875} = 0.3982$$

e

$$CV(F) = \frac{\sigma_F}{\mu_F} = \frac{\sqrt{12.5}}{13.5} = 0.2619$$

Come ci si poteva attendere, le spese presentano un livello medio inferiore al fatturato. Le spese presentano una maggiore variabilità rispetto al fatturato. Si rimanda per il dettaglio dei calcoli alla soluzione dell'esercizio 3, modello retta.



Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.

S Indice di posizione: $M(S) = 8.3875$ Indice di variabilità: $CV(S) = 0.3982$

F Indice di posizione: $M(F) = 13.5$ Indice di variabilità: $CV(F) = 0.2619$

Le spese presentano un livello medio inferiore al fatturato, ma una maggiore variabilità.



3. Per determinare il grado di connessione tra le variabili F e L si costruisca la corrispondente tabella a doppia entrata raggruppando le modalità di F in 3 classi, chiuse a destra, di ampiezza 7 a partire da 0 (limite inf della prima classe) e si calcoli un opportuno indice.

$L \setminus F$	$0 \div 7$	$7 \div 14$	$14 \div 21$
C	1	1	3
P	0	2	1

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.



Dalla distribuzione delle frequenze congiunte n_{ij} :

	0 ÷ 7	7 ÷ 14	14 ÷ 21	$n_{i\bullet}$
C	1	1	3	5
P	0	2	1	3
$n_{\bullet j}$	1	3	4	8

è possibile ottenere le frequenze teoriche nella situazione di indipendenza stocastica

$$\hat{n}_{ij} = \frac{n_{i\bullet} n_{\bullet j}}{n}:$$

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.



Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.

Dalla distribuzione delle frequenze congiunte n_{ij} :

	0 ÷ 7	7 ÷ 14	14 ÷ 21	$n_{i\bullet}$
C	1	1	3	5
P	0	2	1	3
$n_{\bullet j}$	1	3	4	8

è possibile ottenere le frequenze teoriche nella situazione di indipendenza stocastica

$$\hat{n}_{ij} = \frac{n_{i\bullet} n_{\bullet j}}{n}:$$

	0 ÷ 7	7 ÷ 14	14 ÷ 21	$n_{i\bullet}$
C	0.625	1.875	2.5	5
P	0.375	1.125	1.5	3
$n_{\bullet j}$	1.000	3.000	4.0	8



le contingenze assolute $n_{ij} - \hat{n}_{ij}$:

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.



le contingenze assolute $n_{ij} - \hat{n}_{ij}$:

	0 ÷ 7	7 ÷ 14	14 ÷ 21
<i>C</i>	0.375	-0.875	0.5
<i>P</i>	-0.375	0.875	-0.5

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.



i rapporti di contingenza $\frac{n_{ij}}{\hat{n}_{ij}}$:

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.



i rapporti di contingenza $\frac{n_{ij}}{\hat{n}_{ij}}$:

	0 ÷ 7	7 ÷ 14	14 ÷ 21
<i>C</i>	1.6	0.5333	1.2000
<i>P</i>	0.0	1.7778	0.6667

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.



e le contingenze relative, espresse in termini percentuali, $100 \left(\frac{n_{ij} - \hat{n}_{ij}}{\hat{n}_{ij}} \right) \% = 100 \left(\frac{n_{ij}}{\hat{n}_{ij}} - 1 \right) \%$:

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.



e le contingenze relative, espresse in termini percentuali, $100 \left(\frac{n_{ij} - \hat{n}_{ij}}{\hat{n}_{ij}} \right) \% = 100 \left(\frac{n_{ij}}{\hat{n}_{ij}} - 1 \right) \%$:

	0 ÷ 7	7 ÷ 14	14 ÷ 21
<i>C</i>	60	-46.67	20.00
<i>P</i>	-100	77.78	-33.33

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.



Risultati intermedi $\frac{(n_{ij} - \hat{n}_{ij})^2}{\hat{n}_{ij}}$ per il calcolo del χ^2

	0 ÷ 7	7 ÷ 14	14 ÷ 21
C	0.225	0.4083	0.1000
P	0.375	0.6806	0.1667

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.



Risultati intermedi $\frac{n_{ij}^2}{n_{i\bullet} \cdot n_{\bullet j}}$ per il calcolo del χ^2 (formula operativa)

	0 ÷ 7	7 ÷ 14	14 ÷ 21
<i>C</i>	0.2	0.0667	0.4500
<i>P</i>	0.0	0.4444	0.0833

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.



Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.

$$\chi^2 = \sum_{i=1}^h \sum_{j=1}^k \frac{(n_{ij} - \hat{n}_{ij})^2}{\hat{n}_{ij}} = 1.9556,$$

$$\chi_N^2 = \frac{\chi^2}{n \cdot \min(h-1, k-1)} = 0.2444$$

$$\chi_N = \sqrt{\chi_N^2} = 0.4944.$$

Sussiste, quindi, un livello di connessione medio.



Della serie storica X_t del fatturato medio nazionale delle sale cinematografiche tra il 2008 e il 2012 si hanno le seguenti serie incomplete di numeri indici espressi su scala centesimale (base 100):

t	$NIBM$	$NIBF$
2008		100
2009	114	
2010		116
2011	106	
2012		131

1. Si completino le due serie di numeri indici in tabella.

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.



Occorre lavorare con le grandezze relative e, quindi, dividere i valori presenti nella tabella per 100.

$$2009/2010 =$$

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.



Occorre lavorare con le grandezze relative e, quindi, dividere i valori presenti nella tabella per 100.

$$\begin{aligned} 2009 I_{2010} &= \frac{2008 I_{2010}}{2008 I_{2009}} = \frac{1.16}{1.14} = 1.0175 \\ 2008 I_{2011} &= \end{aligned}$$

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.



Occorre lavorare con le grandezze relative e, quindi, dividere i valori presenti nella tabella per 100.

$${}_{2009}I_{2010} = \frac{{}_{2008}I_{2010}}{{}_{2008}I_{2009}} = \frac{1.16}{1.14} = 1.0175$$

$${}_{2008}I_{2011} = {}_{2008}I_{2010} \cdot {}_{2010}I_{2011} = 1.16 \cdot 1.06 = 1.2296$$

$${}_{2011}I_{2012} =$$

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.



Occorre lavorare con le grandezze relative e, quindi, dividere i valori presenti nella tabella per 100.

$${}_{2009}I_{2010} = \frac{{}_{2008}I_{2010}}{{}_{2008}I_{2009}} = \frac{1.16}{1.14} = 1.0175$$

$${}_{2008}I_{2011} = {}_{2008}I_{2010} \cdot {}_{2010}I_{2011} = 1.16 \cdot 1.06 = 1.2296$$

$${}_{2011}I_{2012} = \frac{{}_{2008}I_{2012}}{{}_{2008}I_{2011}} = \frac{1.31}{1.2296} = 1.0654$$

t	$NIBM$	$NIBF$
2008	—	100
2009	114	114
2010	101.75	116
2011	106	122.96
2012	106.54	131

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.



Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.

2. Sapendo che $x_{2010} = 261$, si calcoli x_{2012} .

$$x_{2012} = x_{2010} \cdot {}_{2010}l_{2011} \cdot {}_{2011}l_{2012} = 261 \cdot 1.06 \cdot 1.0654 = 294.75.$$



3. Si calcoli il tasso medio di variazione dal 2010 al 2012 e si commenti il risultato ottenuto.

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.



3. Si calcoli il tasso medio di variazione dal 2010 al 2012 e si commenti il risultato ottenuto.

Si ricorda che

$${}_{2010}I_{2012} = {}_{2010}I_{2011} \cdot {}_{2011}I_{2012}$$

Si deve trovare il numero indice a base mobile medio.

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.



Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.

3. Si calcoli il tasso medio di variazione dal 2010 al 2012 e si commenti il risultato ottenuto.

Si ricorda che

$${}_{2010}I_{2012} = {}_{2010}I_{2011} \cdot {}_{2011}I_{2012}$$

Si deve trovare il numero indice a base mobile medio. Applicando il criterio di scelta della media secondo Chisini si ottiene

$${}_{2010}I_{2012} = {}_{2010}I_{2011} \cdot {}_{2011}I_{2012} = \alpha \cdot \alpha = \alpha^2.$$



Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.

3. Si calcoli il tasso medio di variazione dal 2010 al 2012 e si commenti il risultato ottenuto.

Si ricorda che

$${}_{2010}I_{2012} = {}_{2010}I_{2011} \cdot {}_{2011}I_{2012}$$

Si deve trovare il numero indice a base mobile medio. Applicando il criterio di scelta della media secondo Chisini si ottiene

$${}_{2010}I_{2012} = {}_{2010}I_{2011} \cdot {}_{2011}I_{2012} = \alpha \cdot \alpha = \alpha^2.$$

Quindi

$$\alpha = \sqrt[2]{{}_{2010}I_{2012}} = \sqrt[2]{{}_{2010}I_{2011} \cdot {}_{2011}I_{2012}} = 1.0627.$$

Da cui segue una variazione relativa media pari a 0.0627, corrispondente a un incremento medio annuo tra il 2010 e il 2012 del 6.27%.



Con riferimento ai singoli dati relativi alle 8 sale cinematografiche (già utilizzati al punto 2 dell'Esercizio 1)

F	15	6	11	16	18	14	16	12
S	8.2	3.4	6.2	12.5	13.2	7.2	11.2	5.2
L	C	C	P	C	P	C	C	P

si vuole studiare la relazione tra Fatturato (variabile dipendente) e Spesa.

1. Si rappresentino graficamente le coppie di dati e la funzione di regressione di F dato S .

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.



La funzione di regressione associa a ogni valore s appartenente all'insieme di definizione della variabile S la media di F condizionata a s , $M(F|s)$

s	$M(F s)$
3.4	6
5.2	12
6.2	11
7.2	14
8.2	15
11.2	16
12.5	16
13.2	18

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.



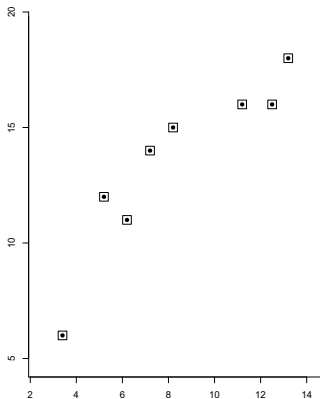
Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.

Le coppie di dati possono essere rappresentate mediante un grafico a dispersione.





2. Si calcolino, secondo il metodo dei minimi quadrati, i parametri e l'indice di adattamento dei seguenti modelli di regressione, individuando il modello migliore:

$$I) F^* = a + b \cdot S \qquad II) F^* = c \cdot \sqrt{S}$$

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.

Si riportano i calcoli necessari¹ per ottenere le stime dei parametri e gli indici di adattamento del modello

$$F = a + bS + E$$

s_i	f_i	$s_i f_i$	s_i^2	f_i^2
8.2	15	123	67.24	225
3.4	6	20.4	11.56	36
6.2	11	68.2	38.44	121
12.5	16	200	156.25	256
13.2	18	237.6	174.24	324
7.2	14	100.8	51.84	196
11.2	16	179.2	125.44	256
5.2	12	62.4	27.04	144
67.1	108	991.6	652.05	1558
8.3875	13.5	123.95	81.5063	194.75
$M(S)$	$M(F)$	$M(SF)$	$M(S^2)$	$M(F^2)$

$$\hat{b} =$$



Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.

¹Si osserva che per ragioni di stampa alcuni calcoli potrebbero essere duplicati con riferimento ai diversi modelli oggetto di studio

Si riportano i calcoli necessari¹ per ottenere le stime dei parametri e gli indici di adattamento del modello

$$F = a + bS + E$$

s_i	f_i	$s_i f_i$	s_i^2	f_i^2
8.2	15	123	67.24	225
3.4	6	20.4	11.56	36
6.2	11	68.2	38.44	121
12.5	16	200	156.25	256
13.2	18	237.6	174.24	324
7.2	14	100.8	51.84	196
11.2	16	179.2	125.44	256
5.2	12	62.4	27.04	144
67.1	108	991.6	652.05	1558
8.3875	13.5	123.95	81.5063	194.75
$M(S)$	$M(F)$	$M(SF)$	$M(S^2)$	$M(F^2)$

$$\begin{aligned}\hat{b} &= \frac{\text{Cov}(S, F)}{\text{Var}(S)} = \frac{M(SF) - M(S)M(F)}{M(S^2) - [M(S)]^2} = \\ &= \frac{123.95 - 8.3875 \cdot 13.5}{81.5063 - 8.3875^2} = \frac{10.7188}{11.1561} = 0.9608 \\ \hat{a} &= \end{aligned}$$

¹Si osserva che per ragioni di stampa alcuni calcoli potrebbero essere duplicati con riferimento ai diversi modelli oggetto di studio



Si riportano i calcoli necessari¹ per ottenere le stime dei parametri e gli indici di adattamento del modello

$$F = a + bS + E$$

s_i	f_i	$s_i f_i$	s_i^2	f_i^2
8.2	15	123	67.24	225
3.4	6	20.4	11.56	36
6.2	11	68.2	38.44	121
12.5	16	200	156.25	256
13.2	18	237.6	174.24	324
7.2	14	100.8	51.84	196
11.2	16	179.2	125.44	256
5.2	12	62.4	27.04	144
67.1	108	991.6	652.05	1558
8.3875	13.5	123.95	81.5063	194.75
$M(S)$	$M(F)$	$M(SF)$	$M(S^2)$	$M(F^2)$

$$\begin{aligned}\hat{b} &= \frac{\text{Cov}(S, F)}{\text{Var}(S)} = \frac{M(SF) - M(S)M(F)}{M(S^2) - [M(S)]^2} = \\ &= \frac{123.95 - 8.3875 \cdot 13.5}{81.5063 - 8.3875^2} = \frac{10.7188}{11.1561} = 0.9608\end{aligned}$$

$$\hat{a} = M(F) - \hat{b}M(S) = 5.4413$$

$$\rho^2 =$$

¹Si osserva che per ragioni di stampa alcuni calcoli potrebbero essere duplicati con riferimento ai diversi modelli oggetto di studio



Si riportano i calcoli necessari¹ per ottenere le stime dei parametri e gli indici di adattamento del modello

$$F = a + bS + E$$

s_i	f_i	$s_i f_i$	s_i^2	f_i^2
8.2	15	123	67.24	225
3.4	6	20.4	11.56	36
6.2	11	68.2	38.44	121
12.5	16	200	156.25	256
13.2	18	237.6	174.24	324
7.2	14	100.8	51.84	196
11.2	16	179.2	125.44	256
5.2	12	62.4	27.04	144
67.1	108	991.6	652.05	1558
8.3875	13.5	123.95	81.5063	194.75
$M(S)$	$M(F)$	$M(SF)$	$M(S^2)$	$M(F^2)$

$$\begin{aligned}\hat{b} &= \frac{\text{Cov}(S, F)}{\text{Var}(S)} = \frac{M(SF) - M(S)M(F)}{M(S^2) - [M(S)]^2} = \\ &= \frac{123.95 - 8.3875 \cdot 13.5}{81.5063 - 8.3875^2} = \frac{10.7188}{11.1561} = 0.9608\end{aligned}$$

$$\hat{a} = M(F) - \hat{b} M(S) = 5.4413$$

$$\rho^2 = \frac{[\text{Cov}(S, F)]^2}{\text{Var}(S) \text{Var}(F)} = \frac{10.7188^2}{11.1561 \cdot 12.5} = 0.8239$$

$$M(\hat{E}^2) = \text{Var}(\hat{E}) = \text{varianza residua} =$$

¹Si osserva che per ragioni di stampa alcuni calcoli potrebbero essere duplicati con riferimento ai diversi modelli oggetto di studio



Si riportano i calcoli necessari¹ per ottenere le stime dei parametri e gli indici di adattamento del modello

$$F = a + bS + E$$

s_i	f_i	$s_i f_i$	s_i^2	f_i^2
8.2	15	123	67.24	225
3.4	6	20.4	11.56	36
6.2	11	68.2	38.44	121
12.5	16	200	156.25	256
13.2	18	237.6	174.24	324
7.2	14	100.8	51.84	196
11.2	16	179.2	125.44	256
5.2	12	62.4	27.04	144
67.1	108	991.6	652.05	1558
8.3875	13.5	123.95	81.5063	194.75
$M(S)$	$M(F)$	$M(SF)$	$M(S^2)$	$M(F^2)$

$$\begin{aligned}\hat{b} &= \frac{\text{Cov}(S, F)}{\text{Var}(S)} = \frac{M(SF) - M(S)M(F)}{M(S^2) - [M(S)]^2} = \\ &= \frac{123.95 - 8.3875 \cdot 13.5}{81.5063 - 8.3875^2} = \frac{10.7188}{11.1561} = 0.9608\end{aligned}$$

$$\hat{a} = M(F) - \hat{b}M(S) = 5.4413$$

$$\rho^2 = \frac{[\text{Cov}(S, F)]^2}{\text{Var}(S) \text{Var}(F)} = \frac{10.7188^2}{11.1561 \cdot 12.5} = 0.8239$$

$$M(\hat{E}^2) = \text{Var}(\hat{E}) = \text{varianza residua} = \text{Var}(F) \cdot (1 - \rho^2) = 2.2015$$

¹Si osserva che per ragioni di stampa alcuni calcoli potrebbero essere duplicati con riferimento ai diversi modelli oggetto di studio





Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.

Si riportano i calcoli necessari² per ottenere le stime dei parametri e la misura di adattamento del modello

$$F = c\sqrt{S} + E$$

$\sqrt{s_i}$	f_i	$\sqrt{s_i} f_i$	$\sqrt{s_i^2}$	f_i^2
2.8636	15	42.954	8.2002	225
1.8439	6	11.0634	3.4	36
2.49	11	27.39	6.2001	121
3.5355	16	56.568	12.4998	256
3.6332	18	65.3976	13.2001	324
2.6833	14	37.5662	7.2001	196
3.3466	16	53.5456	11.1997	256
2.2804	12	27.3648	5.2002	144
		321.8496	67.1002	1558
		40.2312	8.3875	194.75
		$M(\sqrt{S}F)$	$M(\sqrt{S^2})$	$M(F^2)$

$\hat{c} =$

²Si osserva che per ragioni di stampa alcuni calcoli potrebbero essere duplicati con riferimento ai diversi modelli oggetto di studio



Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.

Si riportano i calcoli necessari² per ottenere le stime dei parametri e la misura di adattamento del modello

$$F = c\sqrt{S} + E$$

$\sqrt{s_i}$	f_i	$\sqrt{s_i} f_i$	$\sqrt{s_i^2}$	f_i^2
2.8636	15	42.954	8.2002	225
1.8439	6	11.0634	3.4	36
2.49	11	27.39	6.2001	121
3.5355	16	56.568	12.4998	256
3.6332	18	65.3976	13.2001	324
2.6833	14	37.5662	7.2001	196
3.3466	16	53.5456	11.1997	256
2.2804	12	27.3648	5.2002	144
		321.8496	67.1002	1558
		40.2312	8.3875	194.75
		$M(\sqrt{S}F)$	$M(\sqrt{S^2})$	$M(F^2)$

$$\hat{c} = \frac{M(\sqrt{S}F)}{M(\sqrt{S^2})} = \frac{40.2312}{8.3875} = 4.7966$$

$$M(\hat{E}^2) =$$

²Si osserva che per ragioni di stampa alcuni calcoli potrebbero essere duplicati con riferimento ai diversi modelli oggetto di studio



Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.

Si riportano i calcoli necessari² per ottenere le stime dei parametri e la misura di adattamento del modello

$$F = c\sqrt{S} + E$$

$\sqrt{s_i}$	f_i	$\sqrt{s_i} f_i$	$\sqrt{s_i^2}$	f_i^2
2.8636	15	42.954	8.2002	225
1.8439	6	11.0634	3.4	36
2.49	11	27.39	6.2001	121
3.5355	16	56.568	12.4998	256
3.6332	18	65.3976	13.2001	324
2.6833	14	37.5662	7.2001	196
3.3466	16	53.5456	11.1997	256
2.2804	12	27.3648	5.2002	144
		321.8496	67.1002	1558
		40.2312	8.3875	194.75
		$M(\sqrt{S}F)$	$M(\sqrt{S^2})$	$M(F^2)$

$$\hat{c} = \frac{M(\sqrt{S}F)}{M(\sqrt{S^2})} = \frac{40.2312}{8.3875} = 4.7966$$

$$M(\hat{E}^2) = M[F \cdot (F - \hat{c}\sqrt{S})] = M(F^2) - \hat{c}M(\sqrt{S}F) = 194.75 - 4.7966 \cdot 40.2312 = 1.7784$$

²Si osserva che per ragioni di stampa alcuni calcoli potrebbero essere duplicati con riferimento ai diversi modelli oggetto di studio



$a = 5.4413$, $b = 0.9608$, indice adattamento $M(\hat{E}^2) = 2.2014$ $c = 4.7966$, indice
adattamento $M(\hat{E}^2) = 1.7784$
Il modello migliore è $Y^* = 4.7966\sqrt{X}$.

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.



3. Sulla base del miglior modello si preveda il fatturato di una sala cinematografica con spese pari a 5 mila euro.

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.



Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.

3. Sulla base del miglior modello si preveda il fatturato di una sala cinematografica con spese pari a 5 mila euro.

$$Y^*(5) = 4.7966\sqrt{5} = 10.7255.$$



Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.

4. Si determini il valore del rapporto di correlazione tra le due variabili $\eta_{F|S}^2$, commentando il valore trovato.

$\eta_{F|S}^2 = 1$ in quanto sussiste dipendenza funzionale di Y da X .



Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.

5. Descrivere come si potrebbe trattare la mutabile L per inserirla nel modello di regressione l).

Occorre creare una variabile indicatrice, o variabile *dummy*, X , definita a partire da L , ponendo in corrispondenza della i -ma unità statistica:

$$x_i = \begin{cases} 1 & \text{se } l_i = C \\ 0 & \text{se } l_i = P \end{cases} \quad \text{oppure} \quad x_i = \begin{cases} 1 & \text{se } l_i = P \\ 0 & \text{se } l_i = C \end{cases}$$



Date due differenti urne, U_1 e U_2 , contenenti palline numerate, un esperimento casuale consiste nell'estrarre una pallina da ciascuna urna. Sapendo che nell'urna U_1 vi sono 5 palline con numeri da 1 a 5 e che nell'urna U_2 vi sono 1 pallina con il numero 1, 1 pallina con il numero 3 e 8 palline con il numero 5:

1. costruire la distribuzione di probabilità della variabile casuale $X =$ "risultato estrazione urna U_2 ".

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.



Date due differenti urne, U_1 e U_2 , contenenti palline numerate, un esperimento casuale consiste nell'estrarre una pallina da ciascuna urna. Sapendo che nell'urna U_1 vi sono 5 palline con numeri da 1 a 5 e che nell'urna U_2 vi sono 1 pallina con il numero 1, 1 pallina con il numero 3 e 8 palline con il numero 5:

1. costruire la distribuzione di probabilità della variabile casuale $X = \text{"risultato estrazione urna } U_2\text{"}$.

x_i	$P(X = x_i)$
1	0.1
3	0.1
5	0.8

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.



2. Calcolare la probabilità che la somma dei numeri delle 2 palline estratte dia 6.

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.



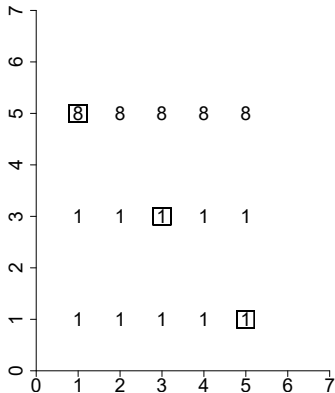
Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.

2. Calcolare la probabilità che la somma dei numeri delle 2 palline estratte dia 6.



$$P(\text{somma dei numeri} = 6) = \frac{8+1+1}{50} = \frac{10}{50} = 0.2$$



oppure

$$= P([(U_1 = 1) \cap (U_2 = 5)] \cup [(U_1 = 3) \cap (U_2 = 3)] \cup [(U_1 = 5) \cap (U_2 = 1)]) =$$

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.



oppure

$$\begin{aligned} &= P([(U_1 = 1) \cap (U_2 = 5)] \cup [(U_1 = 3) \cap (U_2 = 3)] \cup [(U_1 = 5) \cap (U_2 = 1)]) = \\ &= P(U_1 = 1) \cdot P(U_2 = 5) + P(U_1 = 3) \cdot P(U_2 = 3) + P(U_1 = 5) \cdot P(U_2 = 1) = \end{aligned}$$

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.



oppure

$$\begin{aligned} &= P([(U_1 = 1) \cap (U_2 = 5)] \cup [(U_1 = 3) \cap (U_2 = 3)] \cup [(U_1 = 5) \cap (U_2 = 1)]) = \\ &= P(U_1 = 1) \cdot P(U_2 = 5) + P(U_1 = 3) \cdot P(U_2 = 3) + P(U_1 = 5) \cdot P(U_2 = 1) = \\ &= \frac{1}{5} \cdot \frac{8}{10} + \frac{1}{5} \cdot \frac{1}{10} + \frac{1}{5} \cdot \frac{1}{10} = \frac{10}{50} = 0.2. \end{aligned}$$

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.



3. Calcolare la probabilità che, su 10 estrazioni con reinserimento dall'urna U_2 , almeno 2 palline presentino il numero 1.

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.



3. Calcolare la probabilità che, su 10 estrazioni con reinserimento dall'urna U_2 , almeno 2 palline presentino il numero 1.

Sia X la variabile casuale 'n. di palline con il numero 1 nelle 10 estrazioni con reinserimento (prove indipendenti)', $X \sim$

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.



3. Calcolare la probabilità che, su 10 estrazioni con reinserimento dall'urna U_2 , almeno 2 palline presentino il numero 1.

Sia X la variabile casuale 'n. di palline con il numero 1 nelle 10 estrazioni con reinserimento (prove indipendenti)', $X \sim \text{Bin}(n = 10, p = 0.1)$.

$$P(X \geq 2) = 1 - P(X < 2) = 1 - [$$

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.



Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.

3. Calcolare la probabilità che, su 10 estrazioni con reinserimento dall'urna U_2 , almeno 2 palline presentino il numero 1.

Sia X la variabile casuale 'n. di palline con il numero 1 nelle 10 estrazioni con reinserimento (prove indipendenti)', $X \sim \text{Bin}(n = 10, p = 0.1)$.

$$\begin{aligned} P(X \geq 2) &= 1 - P(X < 2) = 1 - [P(X = 0) + P(X = 1)] = \\ &= 1 - \left[\binom{10}{0} 0.1^0 \cdot 0.9^{10} + \binom{10}{1} 0.1^1 \cdot 0.9^9 \right] = \\ &= 1 - (0.3487 + 0.3874) = 1 - 0.7361 = 0.2639. \end{aligned}$$

Tema T286

Prova scritta di STATISTICA del 14.01.2016

Prova scritta di
STATISTICA del
14.01.2016

Giuseppe Boari
Gabriele Cantaluppi



UNIVERSITÀ
CATTOLICA
del Sacro Cuore

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.

Giuseppe Boari
Gabriele Cantaluppi
Università Cattolica del Sacro Cuore, Milano



Data la seguente tabella a doppia entrata della distribuzione congiunta dei caratteri $X =$ 'quantità di neve caduta (in cm)' in un mese e $Y =$ 'qualità della neve (scadente o buona)' in alcune località di montagna:

$Y \setminus X$	0 - 10	10 - 50	50 - 100
<i>scadente</i>	0	2	0
<i>buona</i>	2	2	2

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.



1. Fornire la rappresentazione grafica delle distribuzioni *marginali*.

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.



Si considera la distribuzione marginale della variabile Y , ottenuta sommando le frequenze di riga

y_i	$n_{i\bullet}$
<i>scadente</i>	2
<i>buona</i>	6

Si tratta di un carattere qualitativo ordinato che può essere rappresentato graficamente mediante un grafico a bastoncini o a canne d'organo.

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.

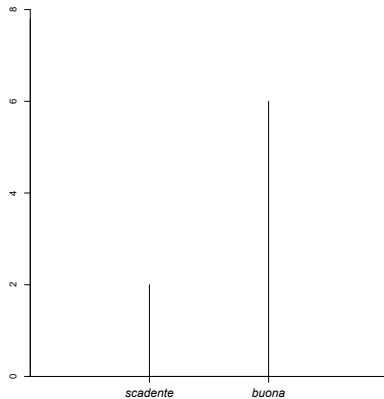


Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.





Si considera la distribuzione marginale della variabile X , ottenuta sommando le frequenze di colonna

x_j	$n_{\bullet j}$
0 $\vdash$ 10	2
10 $\vdash$ 50	4
50 $\vdash$ 100	2

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.



Si tratta di un carattere quantitativo con modalità raggruppate in classi che può essere rappresentato graficamente mediante un istogramma.

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.



L'istogramma può essere costruito dopo avere calcolato le densità $d_{\bullet j} = \frac{n_{\bullet j}}{a_j}$, (dove a_j sono le ampiezze di classe).

x_j	$n_{\bullet j}$	$a_{\bullet j}$	$d_{\bullet j}$
0 $\vdash$ 10	2	10 - 0 = 10	0.2
10 $\vdash$ 50	4	50 - 10 = 40	0.1
50 $\vdash$ 100	2	100 - 50 = 50	0.04

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.

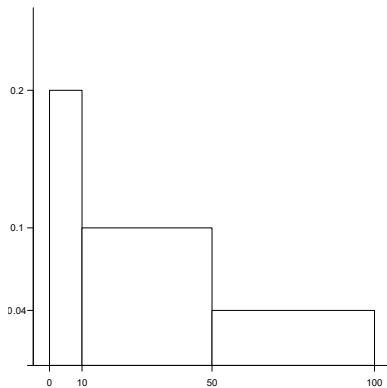


Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.





2. Confrontare con un opportuno indice la variabilità delle distribuzioni condizionate della *qualità* della neve (variabile Y) rispetto alla *quantità* di neve caduta. Commentare i risultati.

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.



Si considera la distribuzione condizionata $Y|x_1$

y_i	n_{i1}
<i>scadente</i>	0
<i>buona</i>	2

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.



Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.

La mutabile Y è degenere ed è, quindi, caratterizzata da assenza di eterogeneità. Gli indici di Gini e di Frosini assumono valore minimo, $F_N = G_N = 0$.



Si considera la distribuzione condizionata $Y|x_1$

y_i	n_{i1}
<i>scadente</i>	2
<i>buona</i>	2

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.



La mutabile Y presenta massima eterogeneità, quindi gli indici di Gini e di Frosini assumono valore massimo, $F_N = G_N = 1$.

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.



Si considera la distribuzione condizionata $Y|x_1$

y_i	n_{i1}
<i>scadente</i>	2
<i>buona</i>	0

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.



La mutabile Y è degenere ed è, quindi, caratterizzata da assenza di eterogeneità. Gli indici di Gini e di Frosini assumono valore minimo, $F_N = G_N = 0$.

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.



Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.

Indice per confronto variabilità: Indice di Frosini normalizzato

valore $Y|x_1$ $F_N = 0$ valore $Y|x_2$ $F_N = 1$ valore $Y|x_3$ $F_N = 0$

Le distribuzioni condizionate $Y|x_1$ e $Y|x_3$ sono degeneri e quindi caratterizzate da assenza di eterogeneità. La distribuzione condizionata $Y|x_2$ è caratterizzata da equidistribuzione delle frequenze; presenta, quindi, massima eterogeneità.



3. Calcolare moda e mediana per la marginale X .

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.



Si tratta di un carattere quantitativo con modalità raggruppate in classi.

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.



La moda corrisponde alla classe alla quale è associata la massima densità di frequenza

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.



La moda corrisponde alla classe alla quale è associata la massima densità di frequenza

y_i	$n_{i\bullet}$	$a_{i\bullet}$	$d_{i\bullet}$
0 $\vdash$ 10	2	$10 - 0 = 10$	0.2
10 $\vdash$ 50	4	$50 - 10 = 40$	0.1
50 $\vdash$ 100	2	$100 - 50 = 50$	0.04

e risulta 0 $\vdash$ 10.

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.



La mediana (percentile di ordine 0.50) può essere ottenuta considerando l'inversa della Funzione di Ripartizione in corrispondenza di 0.50.

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.



La mediana (percentile di ordine 0.50) può essere ottenuta considerando l'inversa della Funzione di Ripartizione in corrispondenza di 0.50.

y_i	$n_{i\bullet}$	$N_{i\bullet}$	$F_{i\bullet}$
0 $\rightarrow$ 10	2	2	0.25
10 $\rightarrow$ 50	4	6	0.75
50 $\rightarrow$ 100	2	8	1

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.

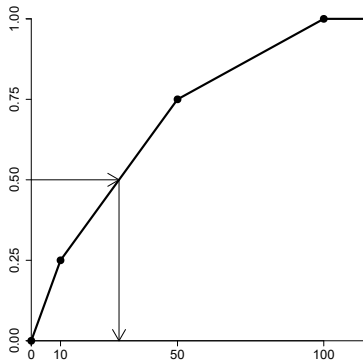


Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.



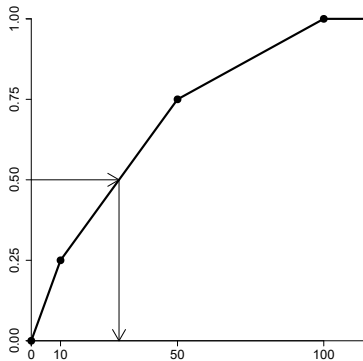


Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.



Si ottiene

$$y_{0.5} = 10 + \frac{0.5 - 0.25}{0.75 - 0.25} \cdot (50 - 10) = 30.$$



Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.

4. Misurare la dipendenza tra le due variabili con un opportuno indice di connessione.



Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.

Dalla distribuzione delle frequenze congiunte n_{ij} :

	0 ÷ 10	10 ÷ 50	50 ÷ 100	$n_{i\bullet}$
<i>scadente</i>	0	2	0	2
<i>buona</i>	2	2	2	6
$n_{\bullet j}$	2	4	2	8

è possibile ottenere le frequenze teoriche nella situazione di indipendenza stocastica

$$\hat{n}_{ij} = \frac{n_{i\bullet} n_{\bullet j}}{n}:$$



Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.

Dalla distribuzione delle frequenze congiunte n_{ij} :

	0 ÷ 10	10 ÷ 50	50 ÷ 100	$n_{i\bullet}$
<i>scadente</i>	0	2	0	2
<i>buona</i>	2	2	2	6
$n_{\bullet j}$	2	4	2	8

è possibile ottenere le frequenze teoriche nella situazione di indipendenza stocastica

$$\hat{n}_{ij} = \frac{n_{i\bullet} n_{\bullet j}}{n}:$$

	0 ÷ 10	10 ÷ 50	50 ÷ 100	$n_{i\bullet}$
<i>scadente</i>	0.5	1	0.5	2
<i>buona</i>	1.5	3	1.5	6
$n_{\bullet j}$	2.0	4	2.0	8



le contingenze assolute $n_{ij} - \hat{n}_{ij}$:

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.



le contingenze assolute $n_{ij} - \hat{n}_{ij}$:

	0 - 10	10 - 50	50 - 100
<i>scadente</i>	-0.5	1	-0.5
<i>buona</i>	0.5	-1	0.5

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.



i rapporti di contingenza $\frac{n_{ij}}{\hat{n}_{ij}}$:

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.



i rapporti di contingenza $\frac{n_{ij}}{\hat{n}_{ij}}$:

	0 - 10	10 - 50	50 - 100
<i>scadente</i>	0.0000	2.0000	0.0000
<i>buona</i>	1.3333	0.6667	1.3333

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.



e le contingenze relative, espresse in termini percentuali, $100 \left(\frac{n_{ij} - \hat{n}_{ij}}{\hat{n}_{ij}} \right) \% = 100 \left(\frac{n_{ij}}{\hat{n}_{ij}} - 1 \right) \%$:

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.



e le contingenze relative, espresse in termini percentuali, $100 \left(\frac{n_{ij} - \hat{n}_{ij}}{\hat{n}_{ij}} \right) \% = 100 \left(\frac{n_{ij}}{\hat{n}_{ij}} - 1 \right) \%$:

	0 - 10	10 - 50	50 - 100
<i>scadente</i>	-100.00	100.00	-100.00
<i>buona</i>	33.33	-33.33	33.33

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.



Risultati intermedi $\frac{(n_{ij} - \hat{n}_{ij})^2}{\hat{n}_{ij}}$ per il calcolo del χ^2

	0 - 10	10 - 50	50 - 100
<i>scadente</i>	0.5000	1.0000	0.5000
<i>buona</i>	0.1667	0.3333	0.1667

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.



Risultati intermedi $\frac{n_{ij}^2}{n_{i\bullet} \cdot n_{\bullet j}}$ per il calcolo del χ^2 (formula operativa)

	0 - 10	10 - 50	50 - 100
<i>scadente</i>	0.0000	0.5000	0.0000
<i>buona</i>	0.3333	0.1667	0.3333

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.



Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.

$$\chi^2 = \sum_{i=1}^h \sum_{j=1}^k \frac{(n_{ij} - \hat{n}_{ij})^2}{\hat{n}_{ij}} = 2.6667,$$

$$\chi_N^2 = \frac{\chi^2}{n \cdot \min(h-1, k-1)} = 0.3333$$

$$\chi_N = \sqrt{\chi_N^2} = 0.5773.$$

Sussiste, quindi, un livello di connessione medio.



Il seguente prospetto riporta i dati dei prezzi medi annui di 2 beni di largo consumo per 5 anni:

t	$bene A$	$bene B$
2005	10	2
2006	11	3
2007	10	4
2008	12	4
2009	14	3

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.



1. Calcolare, giustificando la risposta, la variazione percentuale del prezzo medio del bene A passando dal 2006 al 2009.

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.



1. Calcolare, giustificando la risposta, la variazione percentuale del prezzo medio del bene A passando dal 2006 al 2009.

$${}_{2006}I_{2009} = \frac{X_{2009}}{X_{2006}} = \frac{14}{11} = 1.2727$$

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.



1. Calcolare, giustificando la risposta, la variazione percentuale del prezzo medio del bene A passando dal 2006 al 2009.

$${}_{2006}I_{2009} = \frac{x_{2009}}{x_{2006}} = \frac{14}{11} = 1.2727$$

Tra il tempo 2006 e il tempo 2009 si è registrato un incremento complessivo del 27.27% della grandezza oggetto di studio.

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.



2. Calcolare, mostrandone il procedimento, la serie dei NIBM del bene B .

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.



2. Calcolare, mostrandone il procedimento, la serie dei NIBM del bene B .

Si ricorda che ${}_{t-1}I_t = \frac{x_t}{x_{t-1}}$.

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.



2. Calcolare, mostrandone il procedimento, la serie dei NIBM del bene B .
Si ricorda che ${}_{t-1}I_t = \frac{x_t}{x_{t-1}}$. Si ottiene quindi

<i>Anno</i>	x_t	<i>NIBM</i>
2005	2	
2006	3	1.5000
2007	4	1.3333
2008	4	1.0000
2009	3	0.7500

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.



3. Calcolare il tasso medio di variazione del prezzo del bene B nell'intervallo 2005-2008, commentandone il risultato.

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.



3. Calcolare il tasso medio di variazione del prezzo del bene B nell'intervallo 2005-2008, commentandone il risultato.

Si ricorda che

$${}_{2005}I_{2008} = {}_{2005}I_{2006} \cdot {}_{2006}I_{2007} \cdot {}_{2007}I_{2008}$$

Si deve trovare il numero indice a base mobile medio.

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.



Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.

3. Calcolare il tasso medio di variazione del prezzo del bene B nell'intervallo 2005-2008, commentandone il risultato.

Si ricorda che

$${}_{2005}I_{2008} = {}_{2005}I_{2006} \cdot {}_{2006}I_{2007} \cdot {}_{2007}I_{2008}$$

Si deve trovare il numero indice a base mobile medio. Applicando il criterio di scelta della media secondo Chisini si ottiene

$${}_{2005}I_{2008} = {}_{2005}I_{2006} \cdot {}_{2006}I_{2007} \cdot {}_{2007}I_{2008} = \alpha \cdot \alpha \cdot \alpha = \alpha^3.$$



Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.

3. Calcolare il tasso medio di variazione del prezzo del bene B nell'intervallo 2005-2008, commentandone il risultato.

Si ricorda che

$${}_{2005}I_{2008} = {}_{2005}I_{2006} \cdot {}_{2006}I_{2007} \cdot {}_{2007}I_{2008}$$

Si deve trovare il numero indice a base mobile medio. Applicando il criterio di scelta della media secondo Chisini si ottiene

$${}_{2005}I_{2008} = {}_{2005}I_{2006} \cdot {}_{2006}I_{2007} \cdot {}_{2007}I_{2008} = \alpha \cdot \alpha \cdot \alpha = \alpha^3.$$

Quindi

$$\alpha = \sqrt[3]{{}_{2005}I_{2008}} = \sqrt[3]{{}_{2005}I_{2006} \cdot {}_{2006}I_{2007} \cdot {}_{2007}I_{2008}} = 1.2599.$$

Da cui segue una variazione relativa media pari a 0.2599, corrispondente a un incremento medio annuo tra il 2005 e il 2008 del 25.99%.



Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.

4. Sapendo che le quantità dei due beni vendute nel 2005 sono state rispettivamente 100 e 300, calcolare, se possibile, l'indice di Laspeyres dei prezzi al 2009 con base 2005.



L'indice dei prezzi di Laspeyres si ottiene considerando la media aritmetica dei 'prezzi relativi' riferiti ai singoli prodotti $x_i = \frac{i p_{2008}}{i p_{2005}}$ calcolata utilizzando come frequenze i valori

$i v_{2005} = i p_{2005} \cdot i q_{2005}$ scambiati al tempo 2005

$${}^p_{2005}I^L_{2008} = \frac{1}{\sum_{i=1}^h i p_{2005} i q_{2005}} \sum_{i=1}^h \frac{i p_{2008}}{i p_{2005}} i p_{2005} i q_{2005}$$

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.



Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.

$\frac{i p_{2008}}{i p_{2005}}$	$i p_{2005} \cdot i q_{2005}$	$i p_{2008} \cdot i q_{2005}$
$\frac{12}{10} = 1.2$	$10 \cdot 100 = 1000$	1200
$\frac{4}{2} = 2$	$2 \cdot 300 = 600$	1200
	1600	2400

$$^p_{2005} I^L_{2008} = \frac{1}{1600} 2400 = 1.5.$$

La variazione relativa nel livello generale dei prezzi, secondo l'Indice di Laspeyres, è del 50%.



Con riferimento allo studio della dipendenza tra due caratteri quantitativi X e Y , si sono raccolte le seguenti 5 coppie di osservazioni (x_i, y_i) $i = 1, 2, \dots, 5$

x_i	2	3	3	3	4
y_i	20	13	13	15	20

1. Fornire la rappresentazione grafica delle coppie di punti e della funzione di regressione di Y dato X .

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.



La funzione di regressione associa a ogni valore x appartenente all'insieme di definizione della variabile X la media di Y condizionata a x , $M(Y|x)$

x	$M(Y x)$
2	20
3	$\frac{1}{3}(13 + 13 + 15) = 13.6667$
4	20

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.



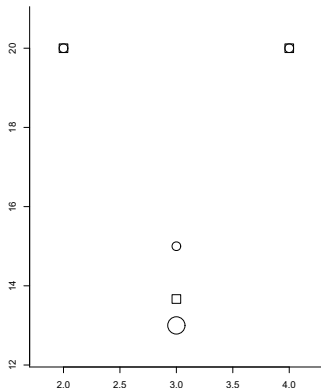
Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.

Le coppie di dati possono essere rappresentate mediante un *bubble plot diagram*.





Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.

2. Si calcolino, secondo il metodo dei minimi quadrati, i parametri dei seguenti modelli di regressione.

$$I) Y^* = a + b \cdot X \qquad II) Y^* = \alpha + \beta \cdot (X - 3)^2$$



Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.

Si riportano i calcoli necessari¹ per ottenere le stime dei parametri e gli indici di adattamento del modello

$$Y = a + bX + E$$

x_i	y_i	$x_i y_i$	x_i^2	y_i^2
2	20	40	4	400
3	13	39	9	169
3	13	39	9	169
3	15	45	9	225
4	20	80	16	400
15	81	243	47	1363
3	16.2	48.6	9.4	272.6
$M(X)$	$M(Y)$	$M(XY)$	$M(X^2)$	$M(Y^2)$

$$\hat{b} =$$

¹Si osserva che per ragioni di stampa alcuni calcoli potrebbero essere duplicati con riferimento ai diversi modelli oggetto di studio



Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.

Si riportano i calcoli necessari¹ per ottenere le stime dei parametri e gli indici di adattamento del modello

$$Y = a + bX + E$$

x_i	y_i	$x_i y_i$	x_i^2	y_i^2
2	20	40	4	400
3	13	39	9	169
3	13	39	9	169
3	15	45	9	225
4	20	80	16	400
15	81	243	47	1363
3	16.2	48.6	9.4	272.6
$M(X)$	$M(Y)$	$M(XY)$	$M(X^2)$	$M(Y^2)$

$$\begin{aligned}\hat{b} &= \frac{\text{Cov}(X, Y)}{\text{Var}(X)} = \frac{M(XY) - M(X)M(Y)}{M(X^2) - [M(X)]^2} = \\ &= \frac{48.6 - 3 \cdot 16.2}{9.4 - 3^2} = \frac{0}{0.4} = 0\end{aligned}$$

$$\hat{a} =$$

¹Si osserva che per ragioni di stampa alcuni calcoli potrebbero essere duplicati con riferimento ai diversi modelli oggetto di studio



Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.

Si riportano i calcoli necessari¹ per ottenere le stime dei parametri e gli indici di adattamento del modello

$$Y = a + bX + E$$

x_i	y_i	$x_i y_i$	x_i^2	y_i^2
2	20	40	4	400
3	13	39	9	169
3	13	39	9	169
3	15	45	9	225
4	20	80	16	400
15	81	243	47	1363
3	16.2	48.6	9.4	272.6
$M(X)$	$M(Y)$	$M(XY)$	$M(X^2)$	$M(Y^2)$

$$\hat{b} = \frac{\text{Cov}(X, Y)}{\text{Var}(X)} = \frac{M(XY) - M(X)M(Y)}{M(X^2) - [M(X)]^2} =$$

$$= \frac{48.6 - 3 \cdot 16.2}{9.4 - 3^2} = \frac{0}{0.4} = 0$$

$$\hat{a} = M(Y) - \hat{b}M(X) = 16.2$$

$$\rho^2 =$$

¹Si osserva che per ragioni di stampa alcuni calcoli potrebbero essere duplicati con riferimento ai diversi modelli oggetto di studio



Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.

Si riportano i calcoli necessari¹ per ottenere le stime dei parametri e gli indici di adattamento del modello

$$Y = a + bX + E$$

x_i	y_i	$x_i y_i$	x_i^2	y_i^2
2	20	40	4	400
3	13	39	9	169
3	13	39	9	169
3	15	45	9	225
4	20	80	16	400
15	81	243	47	1363
3	16.2	48.6	9.4	272.6
$M(X)$	$M(Y)$	$M(XY)$	$M(X^2)$	$M(Y^2)$

$$\hat{b} = \frac{\text{Cov}(X, Y)}{\text{Var}(X)} = \frac{M(XY) - M(X)M(Y)}{M(X^2) - [M(X)]^2} =$$

$$= \frac{48.6 - 3 \cdot 16.2}{9.4 - 3^2} = \frac{0}{0.4} = 0$$

$$\hat{a} = M(Y) - \hat{b}M(X) = 16.2$$

$$\rho^2 = \frac{[\text{Cov}(X, Y)]^2}{\text{Var}(X) \text{Var}(Y)} = \frac{0^2}{0.4 \cdot 10.16} = 0$$

¹Si osserva che per ragioni di stampa alcuni calcoli potrebbero essere duplicati con riferimento ai diversi modelli oggetto di studio

Si riportano i calcoli necessari² per ottenere le stime dei parametri e gli indici di adattamento del modello

$$Y = \alpha + \beta(X - 3)^2 + E$$

$(x - 3)^2_i$	y_i	$(x - 3)^2_i y_i$	$(x - 3)^{2^2}_i$	y_i^2
1	20	20	1	400
0	13	0	0	169
0	13	0	0	169
0	15	0	0	225
1	20	20	1	400
2	81	40	2	1363
0.4	16.2	8	0.4	272.6
$M((X - 3)^2)$	$M(Y)$	$M((X - 3)^2 Y)$	$M((X - 3)^{2^2})$	$M(Y^2)$

$$\hat{\beta} =$$



Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.

²Si osserva che per ragioni di stampa alcuni calcoli potrebbero essere duplicati con riferimento ai diversi modelli oggetto di studio



Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.

Si riportano i calcoli necessari² per ottenere le stime dei parametri e gli indici di adattamento del modello

$$Y = \alpha + \beta(X - 3)^2 + E$$

$(x - 3)^2_i$	y_i	$(x - 3)^2_i y_i$	$(x - 3)^{2^2}_i$	y_i^2
1	20	20	1	400
0	13	0	0	169
0	13	0	0	169
0	15	0	0	225
1	20	20	1	400
2	81	40	2	1363
0.4	16.2	8	0.4	272.6
$M((X - 3)^2)$	$M(Y)$	$M((X - 3)^2 Y)$	$M((X - 3)^{2^2})$	$M(Y^2)$

$$\begin{aligned}\hat{\beta} &= \frac{\text{Cov}((X - 3)^2, Y)}{\text{Var}((X - 3)^2)} = \frac{M((X - 3)^2 Y) - M((X - 3)^2) M(Y)}{M((X - 3)^{2^2}) - [M((X - 3)^2)^2]} = \\ &= \frac{8 - 0.4 \cdot 16.2}{0.4 - 0.4^2} = \frac{1.52}{0.24} = 6.3333 \\ \hat{\alpha} &= \end{aligned}$$

²Si osserva che per ragioni di stampa alcuni calcoli potrebbero essere duplicati con riferimento ai diversi modelli oggetto di studio



Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.

Si riportano i calcoli necessari² per ottenere le stime dei parametri e gli indici di adattamento del modello

$$Y = \alpha + \beta(X - 3)^2 + E$$

$(x - 3)^2_i$	y_i	$(x - 3)^2_i y_i$	$(x - 3)^{2^2}_i$	y_i^2
1	20	20	1	400
0	13	0	0	169
0	13	0	0	169
0	15	0	0	225
1	20	20	1	400
2	81	40	2	1363
0.4	16.2	8	0.4	272.6
$M((X - 3)^2)$	$M(Y)$	$M((X - 3)^2 Y)$	$M((X - 3)^{2^2})$	$M(Y^2)$

$$\hat{\beta} = \frac{\text{Cov}((X - 3)^2, Y)}{\text{Var}((X - 3)^2)} = \frac{M((X - 3)^2 Y) - M((X - 3)^2) M(Y)}{M((X - 3)^{2^2}) - [M((X - 3)^2)^2]} =$$

$$= \frac{8 - 0.4 \cdot 16.2}{0.4 - 0.4^2} = \frac{1.52}{0.24} = 6.3333$$

$$\hat{\alpha} = M(Y) - \hat{\beta} M((X - 3)^2) = 13.6667$$

$$\rho^2 =$$

²Si osserva che per ragioni di stampa alcuni calcoli potrebbero essere duplicati con riferimento ai diversi modelli oggetto di studio



Si riportano i calcoli necessari² per ottenere le stime dei parametri e gli indici di adattamento del modello

$$Y = \alpha + \beta(X - 3)^2 + E$$

$(x - 3)^2_i$	y_i	$(x - 3)^2_i y_i$	$(x - 3)^{2^2}_i$	y_i^2
1	20	20	1	400
0	13	0	0	169
0	13	0	0	169
0	15	0	0	225
1	20	20	1	400
2	81	40	2	1363
0.4	16.2	8	0.4	272.6
$M((X - 3)^2)$	$M(Y)$	$M((X - 3)^2 Y)$	$M((X - 3)^{2^2})$	$M(Y^2)$

$$\begin{aligned}\hat{\beta} &= \frac{\text{Cov}((X - 3)^2, Y)}{\text{Var}((X - 3)^2)} = \frac{M((X - 3)^2 Y) - M((X - 3)^2) M(Y)}{M((X - 3)^{2^2}) - [M((X - 3)^2)^2]} = \\ &= \frac{8 - 0.4 \cdot 16.2}{0.4 - 0.4^2} = \frac{1.52}{0.24} = 6.3333 \\ \hat{\alpha} &= M(Y) - \hat{\beta} M((X - 3)^2) = 13.6667 \\ \rho^2 &= \frac{[\text{Cov}((X - 3)^2, Y)]^2}{\text{Var}((X - 3)^2) \text{Var}(Y)} = \frac{1.52^2}{0.24 \cdot 10.16} = 0.9475\end{aligned}$$

²Si osserva che per ragioni di stampa alcuni calcoli potrebbero essere duplicati con riferimento ai diversi modelli oggetto di studio



Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.

$a = 16.2$, $b = 0$, indice adattamento $\rho^2 = 0$ $\alpha = 13.6667$, $\beta = 6.3333$, indice adattamento $\rho^2 = 0.9475$

Il modello migliore è $Y^* = 13.6667 + 6.3333(X - 3)^2$.



3. Confrontare gli indici di adattamento, commentando le eventuali differenze.

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.



3. Confrontare gli indici di adattamento, commentando le eventuali differenze.

Cfr. punto precedente per indici adattamento.

Si osserva che il coefficiente angolare b nel modello I) è nullo; non sussiste, quindi, correlazione lineare tra le variabili in gioco.

Il modello II) è una parabola con asse di simmetria in $x = 3$ e passa esattamente per le tre medie condizionate. L'indice di adattamento del secondo modello coincide, quindi, con il rapporto di correlazione $\eta^2_{Y|X}$.

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.



4. Sulla base del miglior modello si preveda il valore di Y in corrispondenza di $X = 4$.

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.



Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.

4. Sulla base del miglior modello si preveda il valore di Y in corrispondenza di $X = 4$.
 $Y^*(4) = 13.6667 + 6.3333(4 - 3)^2 = 20.$

Esercizio 4

Un dado Δ viene truccato in modo che la faccia 4 non esca mai:

1. Descrivere (valori e probabilità) e rappresentare graficamente la v.c. $X =$ 'risultato del lancio del dado Δ '.



Esercizio 1.

Esercizio 2.

Esercizio 3.

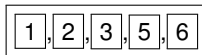
Esercizio 4.

Esercizio 4

Un dado Δ viene truccato in modo che la faccia 4 non esca mai:

1. Descrivere (valori e probabilità) e rappresentare graficamente la v.c. $X =$ 'risultato del lancio del dado Δ '.

Lo spazio probabilistico Ω è composto dai seguenti 5 elementi



caratterizzati da probabilità positiva. La variabile casuale 'risultato del lancio del dado Δ ' assume la seguente forma

x_i	$P(X = x_i)$
1	0.2
2	0.2
3	0.2
4	0
5	0.2
6	0.2

Nel prospetto si è riportato il valore 4, anche se caratterizzato da probabilità pari a 0, in quanto valore di interesse, dal momento che la corrispondente faccia è fisicamente presente sul dado.





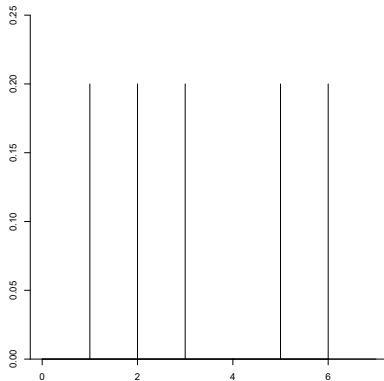
Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.

La variabile casuale può essere rappresentata graficamente mediante un grafico a bastoncini.





Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.

2. Calcolare la probabilità che, su 120 lanci del dado Δ , almeno la metà dia risultato pari.



2. Calcolare la probabilità che, su 120 lanci del dado Δ , almeno la metà dia risultato pari.

Indicata con X la variabile casuale 'numero di successi' nelle 120 prove indipendenti risulta

$$X \sim$$

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.



2. Calcolare la probabilità che, su 120 lanci del dado Δ , almeno la metà dia risultato pari.

Indicata con X la variabile casuale 'numero di successi' nelle 120 prove indipendenti risulta

$$X \sim \text{Bin}(n = 120, p = 0.4).$$

$$M(X) = 48; \text{Var}(X) = 28.8.$$

Dal momento che $np = 48 > 5$ ed $n(1 - p) = 72 > 5$ è possibile utilizzare l'approssimazione della variabile casuale Binomiale con la variabile casuale Normale

$$X \sim$$



2. Calcolare la probabilità che, su 120 lanci del dado Δ , almeno la metà dia risultato pari.

Indicata con X la variabile casuale 'numero di successi' nelle 120 prove indipendenti risulta

$$X \sim \text{Bin}(n = 120, p = 0.4).$$

$$M(X) = 48; \text{Var}(X) = 28.8.$$

Dal momento che $np = 48 > 5$ ed $n(1 - p) = 72 > 5$ è possibile utilizzare l'approssimazione della variabile casuale Binomiale con la variabile casuale Normale

$$X \sim N(\mu = 48, \sigma^2 = 28.8).$$

Si ottiene



2. Calcolare la probabilità che, su 120 lanci del dado Δ , almeno la metà dia risultato pari.

Indicata con X la variabile casuale 'numero di successi' nelle 120 prove indipendenti risulta

$$X \sim \text{Bin}(n = 120, p = 0.4).$$

$$M(X) = 48; \text{Var}(X) = 28.8.$$

Dal momento che $np = 48 > 5$ ed $n(1 - p) = 72 > 5$ è possibile utilizzare l'approssimazione della variabile casuale Binomiale con la variabile casuale Normale

$$X \sim N(\mu = 48, \sigma^2 = 28.8).$$

Si ottiene

$$P(X > 60) =$$



2. Calcolare la probabilità che, su 120 lanci del dado Δ , almeno la metà dia risultato pari.

Indicata con X la variabile casuale 'numero di successi' nelle 120 prove indipendenti risulta

$$X \sim \text{Bin}(n = 120, p = 0.4).$$

$$M(X) = 48; \text{Var}(X) = 28.8.$$

Dal momento che $np = 48 > 5$ ed $n(1 - p) = 72 > 5$ è possibile utilizzare l'approssimazione della variabile casuale Binomiale con la variabile casuale Normale

$$X \sim N(\mu = 48, \sigma^2 = 28.8).$$

Si ottiene

$$P(X > 60) = P\left(\frac{X - \mu}{\sigma} > \frac{60 - \mu}{\sigma}\right) =$$



2. Calcolare la probabilità che, su 120 lanci del dado Δ , almeno la metà dia risultato pari.

Indicata con X la variabile casuale 'numero di successi' nelle 120 prove indipendenti risulta

$$X \sim \text{Bin}(n = 120, p = 0.4).$$

$$M(X) = 48; \text{Var}(X) = 28.8.$$

Dal momento che $np = 48 > 5$ ed $n(1 - p) = 72 > 5$ è possibile utilizzare l'approssimazione della variabile casuale Binomiale con la variabile casuale Normale

$$X \sim N(\mu = 48, \sigma^2 = 28.8).$$

Si ottiene

$$P(X > 60) = P\left(\frac{X - \mu}{\sigma} > \frac{60 - \mu}{\sigma}\right) = P\left(Z > \frac{60 - 48}{5.3666}\right) =$$

=



2. Calcolare la probabilità che, su 120 lanci del dado Δ , almeno la metà dia risultato pari.

Indicata con X la variabile casuale 'numero di successi' nelle 120 prove indipendenti risulta

$$X \sim \text{Bin}(n = 120, p = 0.4).$$

$$M(X) = 48; \text{Var}(X) = 28.8.$$

Dal momento che $np = 48 > 5$ ed $n(1 - p) = 72 > 5$ è possibile utilizzare l'approssimazione della variabile casuale Binomiale con la variabile casuale Normale

$$X \sim N(\mu = 48, \sigma^2 = 28.8).$$

Si ottiene

$$\begin{aligned} P(X > 60) &= P\left(\frac{X - \mu}{\sigma} > \frac{60 - \mu}{\sigma}\right) = P\left(Z > \frac{60 - 48}{5.3666}\right) = \\ &= P(Z > 2.24) = \end{aligned}$$



2. Calcolare la probabilità che, su 120 lanci del dado Δ , almeno la metà dia risultato pari.

Indicata con X la variabile casuale 'numero di successi' nelle 120 prove indipendenti risulta

$$X \sim \text{Bin}(n = 120, p = 0.4).$$

$$M(X) = 48; \text{Var}(X) = 28.8.$$

Dal momento che $np = 48 > 5$ ed $n(1 - p) = 72 > 5$ è possibile utilizzare l'approssimazione della variabile casuale Binomiale con la variabile casuale Normale

$$X \sim N(\mu = 48, \sigma^2 = 28.8).$$

Si ottiene

$$\begin{aligned} P(X > 60) &= P\left(\frac{X - \mu}{\sigma} > \frac{60 - \mu}{\sigma}\right) = P\left(Z > \frac{60 - 48}{5.3666}\right) = \\ &= P(Z > 2.24) = 1 - P(Z \leq 2.24) = \end{aligned}$$



2. Calcolare la probabilità che, su 120 lanci del dado Δ , almeno la metà dia risultato pari.

Indicata con X la variabile casuale 'numero di successi' nelle 120 prove indipendenti risulta

$$X \sim \text{Bin}(n = 120, p = 0.4).$$

$$M(X) = 48; \text{Var}(X) = 28.8.$$

Dal momento che $np = 48 > 5$ ed $n(1 - p) = 72 > 5$ è possibile utilizzare l'approssimazione della variabile casuale Binomiale con la variabile casuale Normale

$$X \sim N(\mu = 48, \sigma^2 = 28.8).$$

Si ottiene

$$\begin{aligned} P(X > 60) &= P\left(\frac{X - \mu}{\sigma} > \frac{60 - \mu}{\sigma}\right) = P\left(Z > \frac{60 - 48}{5.3666}\right) = \\ &= P(Z > 2.24) = 1 - P(Z \leq 2.24) = 1 - 0.9875 = \end{aligned}$$



2. Calcolare la probabilità che, su 120 lanci del dado Δ , almeno la metà dia risultato pari.

Indicata con X la variabile casuale 'numero di successi' nelle 120 prove indipendenti risulta

$$X \sim \text{Bin}(n = 120, p = 0.4).$$

$$M(X) = 48; \text{Var}(X) = 28.8.$$

Dal momento che $np = 48 > 5$ ed $n(1 - p) = 72 > 5$ è possibile utilizzare l'approssimazione della variabile casuale Binomiale con la variabile casuale Normale

$$X \sim N(\mu = 48, \sigma^2 = 28.8).$$

Si ottiene

$$\begin{aligned} P(X > 60) &= P\left(\frac{X - \mu}{\sigma} > \frac{60 - \mu}{\sigma}\right) = P\left(Z > \frac{60 - 48}{5.3666}\right) = \\ &= P(Z > 2.24) = 1 - P(Z \leq 2.24) = 1 - 0.9875 = 0.0125. \end{aligned}$$



3. Calcolare la probabilità che in una generica estrazione del LOTTO (5 numeri estratti da 90, senza reimmissione) i primi 3 siano dispari e gli altri siano pari.

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.



3. Calcolare la probabilità che in una generica estrazione del LOTTO (5 numeri estratti da 90, senza reimmissione) i primi 3 siano dispari e gli altri siano pari.

$$P(d_1 \cap d_2 \cap d_3 \cap p_4 \cap p_5) =$$

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.



3. Calcolare la probabilità che in una generica estrazione del LOTTO (5 numeri estratti da 90, senza reimmissione) i primi 3 siano dispari e gli altri siano pari.

$$P(d_1 \cap d_2 \cap d_3 \cap p_4 \cap p_5) = P(d_1) \cdot P(d_2|d_1) \cdot P(d_3|d_1 \cap d_2) \cdot P(p_4|d_1 \cap d_2 \cap d_3) \cdot P(p_5|d_1 \cap d_2 \cap d_3 \cap p_4) =$$
$$=$$

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.



3. Calcolare la probabilità che in una generica estrazione del LOTTO (5 numeri estratti da 90, senza reimmissione) i primi 3 siano dispari e gli altri siano pari.

$$\begin{aligned} P(d_1 \cap d_2 \cap d_3 \cap p_4 \cap p_5) &= P(d_1) \cdot P(d_2|d_1) \cdot P(d_3|d_1 \cap d_2) \cdot P(p_4|d_1 \cap d_2 \cap d_3) \cdot P(p_5|d_1 \cap d_2 \cap d_3 \cap p_4) = \\ &= \frac{45}{90} \cdot \frac{44}{89} \cdot \frac{43}{88} \cdot \frac{45}{87} \cdot \frac{44}{86} = 0.03196 \end{aligned}$$

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.

Tema T287

Prova scritta di STATISTICA del 28.01.2016

Prova scritta di
STATISTICA del
28.01.2016

Giuseppe Boari
Gabriele Cantaluppi



UNIVERSITÀ
CATTOLICA
del Sacro Cuore

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.

Giuseppe Boari
Gabriele Cantaluppi
Università Cattolica del Sacro Cuore, Milano



La seguente tabella riporta alcune informazioni sulla distribuzione congiunta della variabile statistica doppia (X, Y) , con X carattere qualitativo.

$X \setminus Y$	y_1	y_2	y_3	
x_1		2		
x_2				

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.



1. Si completino le frequenze nella tabella in modo tale che siano verificate contemporaneamente le tre seguenti condizioni: a) la distribuzione $Y|x_1$ sia caratterizzata da massima eterogeneità; b) $\chi^2_{\max} = 9$; c) sussista indipendenza stocastica tra le variabili X e Y .

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.



1. Si completino le frequenze nella tabella in modo tale che siano verificate contemporaneamente le tre seguenti condizioni: a) la distribuzione $Y|x_1$ sia caratterizzata da massima eterogeneità; b) $\chi^2_{\max} = 9$; c) sussista indipendenza stocastica tra le variabili X e Y . La situazione di massima eterogeneità si realizza in presenza di equidistribuzione delle frequenze, quindi $n_{11} = n_{13} = 2$ e $n_{1\bullet} = 6$.

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.



1. Si completino le frequenze nella tabella in modo tale che siano verificate contemporaneamente le tre seguenti condizioni: a) la distribuzione $Y|x_1$ sia caratterizzata da massima eterogeneità; b) $\chi^2_{\max} = 9$; c) sussista indipendenza stocastica tra le variabili X e Y . La situazione di massima eterogeneità si realizza in presenza di equidistribuzione delle frequenze, quindi $n_{11} = n_{13} = 2$ e $n_{1\bullet} = 6$. Ricordando che $\chi^2_{\max} = n \cdot \min(h-1, k-1) = n \cdot \min(2-1, 3-1) = n \cdot \min(1, 2) = n \cdot 1 = 9$ si ottiene $n = 9$ e $n_{2\bullet} = n - n_{1\bullet} = 9 - 6 = 3$.

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.



Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.

1. Si completino le frequenze nella tabella in modo tale che siano verificate contemporaneamente le tre seguenti condizioni: a) la distribuzione $Y|x_1$ sia caratterizzata da massima eterogeneità; b) $\chi^2_{\max} = 9$; c) sussista indipendenza stocastica tra le variabili X e Y . La situazione di massima eterogeneità si realizza in presenza di equidistribuzione delle frequenze, quindi $n_{11} = n_{13} = 2$ e $n_{1\bullet} = 6$.

Ricordando che $\chi^2_{\max} = n \cdot \min(h-1, k-1) = n \cdot \min(2-1, 3-1) = n \cdot \min(1, 2) = n \cdot 1 = 9$ si ottiene $n = 9$ e $n_{2\bullet} = n - n_{1\bullet} = 9 - 6 = 3$.

Dal momento che sussiste indipendenza stocastica, le distribuzioni condizionate sono somiglianti (proporzionali) e, quindi, $n_{21} = n_{22} = n_{23} = 1$.

$X \setminus Y$	y_1	y_2	y_3	
x_1	2	2	2	6
x_2	1	1	1	3
	3	3	3	9



2. Si sostituiscano alle modalità y_1 , y_2 e y_3 di Y le classi $2 \div 4$, $4 \div 6$ e $6 \div 10$, e si rappresentino graficamente le distribuzioni marginali X e Y .

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.



2. Si sostituiscano alle modalità y_1 , y_2 e y_3 di Y le classi $2 \div 4$, $4 \div 6$ e $6 \div 10$, e si rappresentino graficamente le distribuzioni marginali X e Y .

$X \setminus Y$	$2 \div 4$	$4 \div 6$	$6 \div 10$	
x_1	2	2	2	6
x_2	1	1	1	3
	3	3	3	9

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.



Si considera la distribuzione marginale della variabile X , ottenuta sommando le frequenze di riga

x_i	$n_{i\bullet}$
x_1	6
x_2	3

Si tratta di un carattere qualitativo ordinato che può essere rappresentato graficamente mediante un grafico a bastoncini o a canne d'organo.

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.

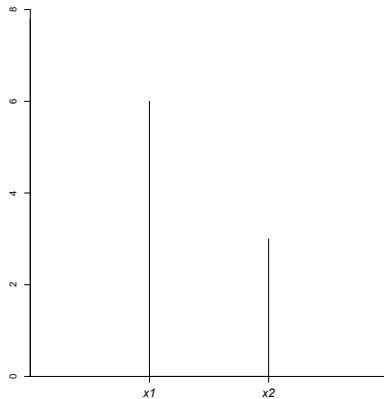


Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.





Si considera la distribuzione marginale della variabile Y , ottenuta sommando le frequenze di colonna

y_j	$n_{\bullet j}$
2 $\rightarrow$ 4	3
4 $\rightarrow$ 6	3
6 $\rightarrow$ 10	3

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.



Si tratta di un carattere quantitativo con modalità raggruppate in classi che può essere rappresentato graficamente mediante un istogramma.

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.



L'istogramma può essere costruito dopo avere calcolato le densità $d_{\bullet j} = \frac{n_{\bullet j}}{a_j}$, (dove a_j sono le ampiezze di classe).

y_j	$n_{\bullet j}$	$a_{\bullet j}$	$d_{\bullet j}$
2 ÷ 4	3	4 - 2 = 2	1.5
4 ÷ 6	3	6 - 4 = 2	1.5
6 ÷ 10	3	10 - 6 = 4	0.75

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.

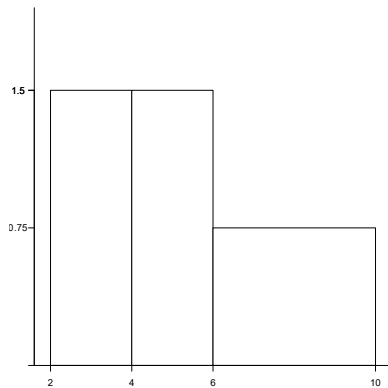


Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.





3. Posto ora $n_{11} = n_{13} = 2$ e $n_{21} = n_{22} = n_{23} = 11$, si calcolino le medie e le varianze delle distribuzioni condizionate $Y|x_1$ e $Y|x_2$ e si confronti la variabilità delle due distribuzioni condizionate.

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.



3. Posto ora $n_{11} = n_{13} = 2$ e $n_{21} = n_{22} = n_{23} = 11$, si calcolino le medie e le varianze delle distribuzioni condizionate $Y|x_1$ e $Y|x_2$ e si confronti la variabilità delle due distribuzioni condizionate.

Dal momento che sussiste indipendenza stocastica le distribuzioni condizionate $Y|x_1$ e $Y|x_2$ sono somiglianti e, quindi, caratterizzate da eguali valori di media aritmetica, varianza e coefficiente di variazione.

Si riporta il calcolo di tali indici per la distribuzione condizionata $Y|x_1$.

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.



Si considera la distribuzione condizionata $Y|x_1$

y_j	n_{1j}
2 $\vdash$ 4	2
4 $\vdash$ 6	2
6 $\vdash$ 10	2

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.



Si tratta di un carattere quantitativo con modalità raggruppate in classi.

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.



La media aritmetica

$$M(Y|x_1) = \frac{1}{n_{1\bullet}} \sum_{j=1}^k y_j n_{1j}$$

risulta 5.3333.

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.



La media aritmetica

$$M(Y|x_1) = \frac{1}{n_{1\bullet}} \sum_{j=1}^k y_j n_{1j}$$

risulta 5.3333.

Si ottiene (considerando come modalità i valori centrali delle classi)

y_j	n_{1j}	$y_j n_{1j}$
3	2	6
5	2	10
8	2	16
	6	32

da cui segue

$$M(Y|x_1) = \frac{1}{n_{1\bullet}} \sum_{j=1}^k y_j n_{1j} = \frac{32}{6} = 5.3333.$$

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.



La varianza

$$\text{Var}(Y|x_1) = M(Y^2|x_1) - [M(Y|x_1)]^2$$

risulta 4.2222.

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.



La varianza

$$\text{Var}(Y|x_1) = M(Y^2|x_1) - [M(Y|x_1)]^2$$

risulta 4.2222. Si ottiene (considerando come modalità i valori centrali delle classi)

y_j	n_{1j}	$y_j n_{1j}$	$y_j^2 n_{1j}$
3	2	6	18
5	2	10	50
8	2	16	128
	6	32	196

da cui segue

$$\text{Var}(Y|x_1) = \frac{196}{6} - 5.3333^2 = 32.6667 - 28.4444 = 4.2223.$$

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.



Per effettuare confronti in termini di variabilità tra caratteri quantitativi occorre utilizzare una misura relativa, ad esempio il coefficiente di variazione

$$CV = \frac{\sigma}{\mu} = \frac{\sqrt{\text{Var}(Y|x_1)}}{M(Y|x_1)} = \frac{\sqrt{4.2223}}{5.3333} = \frac{2.0548}{5.3333} = 0.3853.$$

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.



Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.

$M(Y|x_1) = 5.3333$ $Var(Y|x_1) = 4.2222$ Indice per confronto variabilità $CV(Y|x_1) = 0.3853$
 $M(Y|x_2) = 5.3333$ $Var(Y|x_2) = 4.2222$ Indice per confronto variabilità $CV(Y|x_2) = 0.3853$

Dal momento che sussiste indipendenza stocastica le distribuzioni condizionate $Y|x_1$ e $Y|x_2$ sono somiglianti e, quindi, caratterizzate da eguali valori di media aritmetica, varianza e coefficiente di variazione.



4. Si ricostruisca la varianza della marginale Y a partire dai valori delle medie e delle varianze condizionate determinati al punto precedente.

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.



4. Si ricostruisca la varianza della marginale Y a partire dai valori delle medie e delle varianze condizionate determinati al punto precedente.

$$\text{Var}(Y) = 4.2222$$

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.



4. Si ricostruisca la varianza della marginale Y a partire dai valori delle medie e delle varianze condizionate determinati al punto precedente.

$$\text{Var}(Y) = 4.2222$$

Dal momento che sussiste indipendenza stocastica le distribuzioni condizionate $Y|x_1$ e $Y|x_2$ sono somiglianti alla marginale Y e, quindi, Y avrà lo stesso valore della varianza delle due distribuzioni condizionate.

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.



4. Si ricostruisca la varianza della marginale Y a partire dai valori delle medie e delle varianze condizionate determinati al punto precedente.

$$\text{Var}(Y) = 4.2222$$

Dal momento che sussiste indipendenza stocastica le distribuzioni condizionate $Y|x_1$ e $Y|x_2$ sono somiglianti alla marginale Y e, quindi, Y avrà lo stesso valore della varianza delle due distribuzioni condizionate.

Con riferimento al risultato di scomposizione della varianza

$$\text{Var}(Y) = \text{Var}_X\{M(Y|X)\} + M_X\{\text{Var}(Y|X)\}$$

- le due distribuzioni condizionate hanno la stessa media e, quindi, la varianza spiegata $\text{Var}_X\{M(Y|X)\}$ assume valore nullo;
- le due distribuzioni condizionate hanno la stessa varianza e, quindi, la varianza residua $M_X\{\text{Var}(Y|X)\}$ assume lo stesso valore delle varianze condizionate.

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.

Esercizio 2



Si consideri la seguente serie storica:

t	x_t
2008	500
2009	250
2010	125
2011	250
2012	500

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.



1. Si ricostruiscano le serie dei numeri indici a base fissa, con base 2009, e a base mobile.

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.



1. Si ricostruiscano le serie dei numeri indici a base fissa, con base 2009, e a base mobile.
Si ricorda che ${}_{2009}I_t = \frac{x_t}{x_{2009}}$.

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.



Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.

1. Si ricostruiscano le serie dei numeri indici a base fissa, con base 2009, e a base mobile. Si ricorda che ${}_{2009}I_t = \frac{x_t}{x_{2009}}$. Si ottiene quindi

Anno	x_t	NIBF (2009)
2008	500	2.0
2009	250	1.0
2010	125	0.5
2011	250	1.0
2012	500	2.0



Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.

1. Si ricostruiscano le serie dei numeri indici a base fissa, con base 2009, e a base mobile. Si ricorda che ${}_{2009}I_t = \frac{x_t}{x_{2009}}$. Si ottiene quindi

Anno	x_t	NIBF (2009)
2008	500	2.0
2009	250	1.0
2010	125	0.5
2011	250	1.0
2012	500	2.0

Si ricorda che ${}_{t-1}I_t = \frac{x_t}{x_{t-1}}$.



Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.

1. Si ricostruiscano le serie dei numeri indici a base fissa, con base 2009, e a base mobile. Si ricorda che ${}_{2009}I_t = \frac{x_t}{x_{2009}}$. Si ottiene quindi

Anno	x_t	NIBF (2009)
2008	500	2.0
2009	250	1.0
2010	125	0.5
2011	250	1.0
2012	500	2.0

Si ricorda che ${}_{t-1}I_t = \frac{x_t}{x_{t-1}}$. Si ottiene quindi

Anno	x_t	NIBM
2008	500	
2009	250	0.5
2010	125	0.5
2011	250	2.0
2012	500	2.0



2. Si calcolino e si commentino i seguenti numeri indici ${}_{2010}I_{2008}$ e ${}_{2010}I_{2012}$.

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.



2. Si calcolino e si commentino i seguenti numeri indici ${}_{2010}I_{2008}$ e ${}_{2010}I_{2012}$.

$${}_{2010}I_{2008} = \frac{x_{2008}}{x_{2010}} = \frac{500}{125} = 4$$

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.



Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.

2. Si calcolino e si commentino i seguenti numeri indici ${}_{2010}I_{2008}$ e ${}_{2010}I_{2012}$.

$${}_{2010}I_{2008} = \frac{x_{2008}}{x_{2010}} = \frac{500}{125} = 4$$

La grandezza al tempo 2008 era superiore del 300% rispetto alla grandezza al tempo 2010.
Ovvero il valore della grandezza nel 2008 era il 400% del valore registrato nel 2010.



Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.

2. Si calcolino e si commentino i seguenti numeri indici ${}_{2010}I_{2008}$ e ${}_{2010}I_{2012}$.

$${}_{2010}I_{2008} = \frac{x_{2008}}{x_{2010}} = \frac{500}{125} = 4$$

La grandezza al tempo 2008 era superiore del 300% rispetto alla grandezza al tempo 2010.
Ovvero il valore della grandezza nel 2008 era il 400% del valore registrato nel 2010.

$${}_{2010}I_{2012} = \frac{x_{2012}}{x_{2010}} = \frac{500}{125} = 4$$



Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.

2. Si calcolino e si commentino i seguenti numeri indici ${}_{2010}I_{2008}$ e ${}_{2010}I_{2012}$.

$${}_{2010}I_{2008} = \frac{x_{2008}}{x_{2010}} = \frac{500}{125} = 4$$

La grandezza al tempo 2008 era superiore del 300% rispetto alla grandezza al tempo 2010. Ovvero il valore della grandezza nel 2008 era il 400% del valore registrato nel 2010.

$${}_{2010}I_{2012} = \frac{x_{2012}}{x_{2010}} = \frac{500}{125} = 4$$

Tra il tempo 2010 e il tempo 2012 si è registrato un incremento complessivo del 300% della grandezza oggetto di studio.



3. Si calcoli il tasso medio di variazione dal 2008 al 2011 e si commenti il risultato ottenuto.

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.



3. Si calcoli il tasso medio di variazione dal 2008 al 2011 e si commenti il risultato ottenuto.
Si ricorda che

$${}_{2008}I_{2011} = {}_{2008}I_{2009} \cdot {}_{2009}I_{2010} \cdot {}_{2010}I_{2011}$$

Si deve trovare il numero indice a base mobile medio.

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.



Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.

3. Si calcoli il tasso medio di variazione dal 2008 al 2011 e si commenti il risultato ottenuto. Si ricorda che

$${}_{2008}I_{2011} = {}_{2008}I_{2009} \cdot {}_{2009}I_{2010} \cdot {}_{2010}I_{2011}$$

Si deve trovare il numero indice a base mobile medio. Applicando il criterio di scelta della media secondo Chisini si ottiene

$${}_{2008}I_{2011} = {}_{2008}I_{2009} \cdot {}_{2009}I_{2010} \cdot {}_{2010}I_{2011} = \alpha \cdot \alpha \cdot \alpha = \alpha^3.$$



Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.

3. Si calcoli il tasso medio di variazione dal 2008 al 2011 e si commenti il risultato ottenuto. Si ricorda che

$${}_{2008}I_{2011} = {}_{2008}I_{2009} \cdot {}_{2009}I_{2010} \cdot {}_{2010}I_{2011}$$

Si deve trovare il numero indice a base mobile medio. Applicando il criterio di scelta della media secondo Chisini si ottiene

$${}_{2008}I_{2011} = {}_{2008}I_{2009} \cdot {}_{2009}I_{2010} \cdot {}_{2010}I_{2011} = \alpha \cdot \alpha \cdot \alpha = \alpha^3.$$

Quindi

$$\alpha = \sqrt[3]{{}_{2008}I_{2011}} = \sqrt[3]{{}_{2008}I_{2009} \cdot {}_{2009}I_{2010} \cdot {}_{2010}I_{2011}} = 0.7937.$$

Da cui segue una variazione relativa media pari a -0.2063 , corrispondente a un decremento medio annuo tra il 2008 e il 2011 del 20.63%.



Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.

Il seguente prospetto riporta i valori rilevati in corrispondenza a 5 unità statistiche con riferimento a due caratteri quantitativi X e Y

x_i	54.6	148.41	403.43	1096.63	2980.96
y_i	6	8	9	10	12

1. Si rappresentino graficamente le coppie dei dati ipotizzando Y in funzione di X .



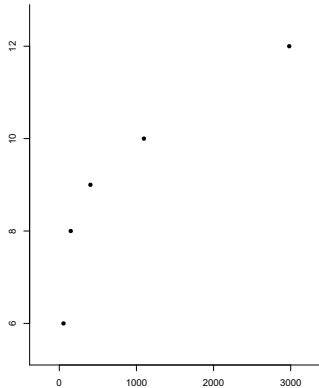
Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.

Le coppie di dati possono essere rappresentate mediante un grafico a dispersione.





2. Si calcolino, secondo il metodo dei minimi quadrati, i parametri e l'indice di adattamento dei seguenti modelli di regressione, individuando il modello migliore:

$$I) Y^* = a + b \cdot \ln(X) \qquad II) Y^* = c$$

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.



Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.

Si riportano i calcoli necessari¹ per ottenere le stime dei parametri e gli indici di adattamento del modello

$$Y = a + b \ln(X) + E$$

$\ln(x)_i$	y_i	$\ln(x)_i y_i$	$\ln(x)_i^2$	y_i^2
4	6	24	16	36
5	8	40	25	64
6	9	54	36	81
7	10	70	49	100
8	12	96	64	144
30	45	284	190	425
6	9	56.8	38	85
$M(\ln(X))$	$M(Y)$	$M(\ln(X)Y)$	$M(\ln(X)^2)$	$M(Y^2)$

$$\hat{b} =$$

¹Si osserva che per ragioni di stampa alcuni calcoli potrebbero essere duplicati con riferimento ai diversi modelli oggetto di studio



Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.

Si riportano i calcoli necessari¹ per ottenere le stime dei parametri e gli indici di adattamento del modello

$$Y = a + b \ln(X) + E$$

$\ln(x)_i$	y_i	$\ln(x)_i y_i$	$\ln(x)_i^2$	y_i^2
4	6	24	16	36
5	8	40	25	64
6	9	54	36	81
7	10	70	49	100
8	12	96	64	144
30	45	284	190	425
6	9	56.8	38	85
$M(\ln(X))$	$M(Y)$	$M(\ln(X)Y)$	$M(\ln(X)^2)$	$M(Y^2)$

$$\begin{aligned}\hat{b} &= \frac{\text{Cov}(\ln(X), Y)}{\text{Var}(\ln(X))} = \frac{M(\ln(X)Y) - M(\ln(X))M(Y)}{M(\ln(X)^2) - [M(\ln(X))]^2} = \\ &= \frac{56.8 - 6 \cdot 9}{38 - 6^2} = \frac{2.8}{2} = 1.4 \\ \hat{a} &= \end{aligned}$$

¹Si osserva che per ragioni di stampa alcuni calcoli potrebbero essere duplicati con riferimento ai diversi modelli oggetto di studio



Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.

Si riportano i calcoli necessari¹ per ottenere le stime dei parametri e gli indici di adattamento del modello

$$Y = a + b \ln(X) + E$$

$\ln(x)_i$	y_i	$\ln(x)_i y_i$	$\ln(x)_i^2$	y_i^2
4	6	24	16	36
5	8	40	25	64
6	9	54	36	81
7	10	70	49	100
8	12	96	64	144
30	45	284	190	425
6	9	56.8	38	85
$M(\ln(X))$	$M(Y)$	$M(\ln(X)Y)$	$M(\ln(X)^2)$	$M(Y^2)$

$$\hat{b} = \frac{\text{Cov}(\ln(X), Y)}{\text{Var}(\ln(X))} = \frac{M(\ln(X)Y) - M(\ln(X))M(Y)}{M(\ln(X)^2) - [M(\ln(X))]^2} =$$

$$= \frac{56.8 - 6 \cdot 9}{38 - 6^2} = \frac{2.8}{2} = 1.4$$

$$\hat{a} = M(Y) - \hat{b} M(\ln(X)) = 0.6$$

$$\rho^2 =$$

¹Si osserva che per ragioni di stampa alcuni calcoli potrebbero essere duplicati con riferimento ai diversi modelli oggetto di studio



Si riportano i calcoli necessari¹ per ottenere le stime dei parametri e gli indici di adattamento del modello

$$Y = a + b \ln(X) + E$$

$\ln(x)_i$	y_i	$\ln(x)_i y_i$	$\ln(x)_i^2$	y_i^2
4	6	24	16	36
5	8	40	25	64
6	9	54	36	81
7	10	70	49	100
8	12	96	64	144
30	45	284	190	425
6	9	56.8	38	85
$M(\ln(X))$	$M(Y)$	$M(\ln(X)Y)$	$M(\ln(X)^2)$	$M(Y^2)$

$$\begin{aligned}\hat{b} &= \frac{\text{Cov}(\ln(X), Y)}{\text{Var}(\ln(X))} = \frac{M(\ln(X)Y) - M(\ln(X))M(Y)}{M(\ln(X)^2) - [M(\ln(X))]^2} = \\ &= \frac{56.8 - 6 \cdot 9}{38 - 6^2} = \frac{2.8}{2} = 1.4\end{aligned}$$

$$\hat{a} = M(Y) - \hat{b} M(\ln(X)) = 0.6$$

$$\rho^2 = \frac{[\text{Cov}(\ln(X), Y)]^2}{\text{Var}(\ln(X)) \text{Var}(Y)} = \frac{2.8^2}{2 \cdot 4} = 0.98$$

¹Si osserva che per ragioni di stampa alcuni calcoli potrebbero essere duplicati con riferimento ai diversi modelli oggetto di studio



Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.

Si riportano i calcoli necessari² per ottenere le stime dei parametri e gli indici di adattamento del modello

$$Y = c + E$$

x_i	y_i	y_i^2
4	6	36
5	8	64
6	9	81
7	10	100
8	12	144
		45
		425
		9
		85
		$M(Y)$
		$M(Y^2)$

$$\hat{c} =$$

²Si osserva che per ragioni di stampa alcuni calcoli potrebbero essere duplicati con riferimento ai diversi modelli oggetto di studio



Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.

Si riportano i calcoli necessari² per ottenere le stime dei parametri e gli indici di adattamento del modello

$$Y = c + E$$

x_i	y_i	y_i^2
4	6	36
5	8	64
6	9	81
7	10	100
8	12	144
		425
		9
		$M(Y)$ $M(Y^2)$

$$\hat{c} = M(Y) = 9$$

$$R^2 =$$

²Si osserva che per ragioni di stampa alcuni calcoli potrebbero essere duplicati con riferimento ai diversi modelli oggetto di studio



Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.

Si riportano i calcoli necessari² per ottenere le stime dei parametri e gli indici di adattamento del modello

$$Y = c + E$$

x_i	y_i	y_i^2
4	6	36
5	8	64
6	9	81
7	10	100
8	12	144
45		425
9		85
$M(Y)$		$M(Y^2)$

$$\hat{c} = M(Y) = 9$$

$$R^2 = 0$$

²Si osserva che per ragioni di stampa alcuni calcoli potrebbero essere duplicati con riferimento ai diversi modelli oggetto di studio



$\hat{a} = 0.6$, $\hat{b} = 1.4$, indice adattamento $\rho^2 = 0.98$ $\hat{c} = 9$, indice adattamento $\rho^2 = 0$
Il modello migliore è $Y^* = 0.6 + 1.4 \cdot \ln(X)$.

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.



3. Si indichino i valori assunti da $\eta_{Y|X}^2$ ed $\eta_{X|Y}^2$.

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.



Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.

3. Si indichino i valori assunti da $\eta_{Y|X}^2$ ed $\eta_{X|Y}^2$.

$\eta_{Y|X}^2 = 1$ ed $\eta_{X|Y}^2 = 1$ in quanto sussiste dipendenza funzionale bijectiva.

Esercizio 4

Prova scritta di
STATISTICA del
28.01.2016

Giuseppe Boari
Gabriele Cantaluppi



UNIVERSITÀ
CATTOLICA
del Sacro Cuore

Per il seguente gioco è prevista una posta di 4 Euro: si lancia un dado regolare; se esce la faccia 3 oppure la 4 si vince 1 Euro, se esce la faccia 5 oppure la 6 si vincono 8 Euro, altrimenti si perde.

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.



1. Costruita la variabile casuale $Y =$ 'guadagno del giocatore' si stabilisca se il gioco è equo.

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.



1. Costruita la variabile casuale $Y =$ 'guadagno del giocatore' si stabilisca se il gioco è equo.

	x_i	$P(X = x_i)$	$x_i \cdot P(X = x_i)$
<i>posta</i>	-4	0.3333	-1.3332
<i>vincita se esce 3 o 4 – posta</i>	-3	0.3333	-1
<i>vincita se esce 5 o 6 – posta</i>	4	0.3333	1.3332
		1	-1

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.



Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.

1. Costruita la variabile casuale $Y =$ 'guadagno del giocatore' si stabilisca se il gioco è equo.

	x_i	$P(X = x_i)$	$x_i \cdot P(X = x_i)$
<i>posta</i>	-4	0.3333	-1.3332
<i>vincita se esce 3 o 4 – posta</i>	-3	0.3333	-1
<i>vincita se esce 5 o 6 – posta</i>	4	0.3333	1.3332
		1	-1

Il valore atteso della variabile casuale guadagno è negativo, quindi il gioco non è equo.



2. Nell'ipotesi di ripetere il gioco 3 volte si calcoli la probabilità di vincere almeno 2 volte.

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.



Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.

2. Nell'ipotesi di ripetere il gioco 3 volte si calcoli la probabilità di vincere almeno 2 volte.

Sia X la variabile casuale 'n. di vincite nelle 3 prove indipendenti'.

$$X \sim$$



Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.

2. Nell'ipotesi di ripetere il gioco 3 volte si calcoli la probabilità di vincere almeno 2 volte.

Sia X la variabile casuale 'n. di vincite nelle 3 prove indipendenti'.

$$X \sim \text{Bin}(n = 3, p = 0.6667).$$

$$P(X \geq 2) =$$



Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.

2. Nell'ipotesi di ripetere il gioco 3 volte si calcoli la probabilità di vincere almeno 2 volte.

Sia X la variabile casuale 'n. di vittorie nelle 3 prove indipendenti'.

$$X \sim \text{Bin}(n = 3, p = 0.6667).$$

$$\begin{aligned} P(X \geq 2) &= P(X = 2) + P(X = 3) = \\ &= \binom{3}{2} 0.6667^2 \cdot 0.3333^1 + \binom{3}{3} 0.6667^3 \cdot 0.3333^0 = \\ &= 0.4444 + 0.2963 = 0.7407. \end{aligned}$$



Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.

3. Nell'ipotesi di ripetere il gioco 100 volte si calcoli la probabilità di vincere almeno 59 volte.



3. Nell'ipotesi di ripetere il gioco 100 volte si calcoli la probabilità di vincere almeno 59 volte.

Indicata con X la variabile casuale 'numero di successi' nelle 100 prove indipendenti risulta

$$X \sim$$

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.



3. Nell'ipotesi di ripetere il gioco 100 volte si calcoli la probabilità di vincere almeno 59 volte.

Indicata con X la variabile casuale 'numero di successi' nelle 100 prove indipendenti risulta

$$X \sim \text{Bin}(n = 100, p = 0.6667).$$

$$M(X) = 66.67; \text{Var}(X) = 22.22.$$

Dal momento che $np = 66.67 > 5$ ed $n(1 - p) = 33.33 > 5$ è possibile utilizzare l'approssimazione della variabile casuale Binomiale con la variabile casuale Normale

$$X \sim$$



3. Nell'ipotesi di ripetere il gioco 100 volte si calcoli la probabilità di vincere almeno 59 volte.

Indicata con X la variabile casuale 'numero di successi' nelle 100 prove indipendenti risulta

$$X \sim \text{Bin}(n = 100, p = 0.6667).$$

$$M(X) = 66.67; \text{Var}(X) = 22.22.$$

Dal momento che $np = 66.67 > 5$ ed $n(1 - p) = 33.33 > 5$ è possibile utilizzare l'approssimazione della variabile casuale Binomiale con la variabile casuale Normale

$$X \sim N(\mu = 66.67, \sigma^2 = 22.22).$$

Si ottiene



3. Nell'ipotesi di ripetere il gioco 100 volte si calcoli la probabilità di vincere almeno 59 volte.

Indicata con X la variabile casuale 'numero di successi' nelle 100 prove indipendenti risulta

$$X \sim \text{Bin}(n = 100, p = 0.6667).$$

$$M(X) = 66.67; \text{Var}(X) = 22.22.$$

Dal momento che $np = 66.67 > 5$ ed $n(1 - p) = 33.33 > 5$ è possibile utilizzare l'approssimazione della variabile casuale Binomiale con la variabile casuale Normale

$$X \sim N(\mu = 66.67, \sigma^2 = 22.22).$$

Si ottiene

$$P(X \geq 59) =$$



3. Nell'ipotesi di ripetere il gioco 100 volte si calcoli la probabilità di vincere almeno 59 volte.

Indicata con X la variabile casuale 'numero di successi' nelle 100 prove indipendenti risulta

$$X \sim \text{Bin}(n = 100, p = 0.6667).$$

$$M(X) = 66.67; \text{Var}(X) = 22.22.$$

Dal momento che $np = 66.67 > 5$ ed $n(1 - p) = 33.33 > 5$ è possibile utilizzare l'approssimazione della variabile casuale Binomiale con la variabile casuale Normale

$$X \sim N(\mu = 66.67, \sigma^2 = 22.22).$$

Si ottiene

$$P(X \geq 59) = P\left(\frac{X - \mu}{\sigma} \geq \frac{59 - \mu}{\sigma}\right) =$$



3. Nell'ipotesi di ripetere il gioco 100 volte si calcoli la probabilità di vincere almeno 59 volte.

Indicata con X la variabile casuale 'numero di successi' nelle 100 prove indipendenti risulta

$$X \sim \text{Bin}(n = 100, p = 0.6667).$$

$$M(X) = 66.67; \text{Var}(X) = 22.22.$$

Dal momento che $np = 66.67 > 5$ ed $n(1 - p) = 33.33 > 5$ è possibile utilizzare l'approssimazione della variabile casuale Binomiale con la variabile casuale Normale

$$X \sim N(\mu = 66.67, \sigma^2 = 22.22).$$

Si ottiene

$$P(X \geq 59) = P\left(\frac{X - \mu}{\sigma} \geq \frac{59 - \mu}{\sigma}\right) = P\left(Z \geq \frac{59 - 66.67}{4.7148}\right) =$$

=



3. Nell'ipotesi di ripetere il gioco 100 volte si calcoli la probabilità di vincere almeno 59 volte.

Indicata con X la variabile casuale 'numero di successi' nelle 100 prove indipendenti risulta

$$X \sim \text{Bin}(n = 100, p = 0.6667).$$

$$M(X) = 66.67; \text{Var}(X) = 22.22.$$

Dal momento che $np = 66.67 > 5$ ed $n(1 - p) = 33.33 > 5$ è possibile utilizzare l'approssimazione della variabile casuale Binomiale con la variabile casuale Normale

$$X \sim N(\mu = 66.67, \sigma^2 = 22.22).$$

Si ottiene

$$\begin{aligned} P(X \geq 59) &= P\left(\frac{X - \mu}{\sigma} \geq \frac{59 - \mu}{\sigma}\right) = P\left(Z \geq \frac{59 - 66.67}{4.7148}\right) = \\ &= P(Z \geq -1.63) = \end{aligned}$$



3. Nell'ipotesi di ripetere il gioco 100 volte si calcoli la probabilità di vincere almeno 59 volte.

Indicata con X la variabile casuale 'numero di successi' nelle 100 prove indipendenti risulta

$$X \sim \text{Bin}(n = 100, p = 0.6667).$$

$$M(X) = 66.67; \text{Var}(X) = 22.22.$$

Dal momento che $np = 66.67 > 5$ ed $n(1 - p) = 33.33 > 5$ è possibile utilizzare l'approssimazione della variabile casuale Binomiale con la variabile casuale Normale

$$X \sim N(\mu = 66.67, \sigma^2 = 22.22).$$

Si ottiene

$$\begin{aligned} P(X \geq 59) &= P\left(\frac{X - \mu}{\sigma} \geq \frac{59 - \mu}{\sigma}\right) = P\left(Z \geq \frac{59 - 66.67}{4.7148}\right) = \\ &= P(Z \geq -1.63) = P(Z \leq 1.63) = \end{aligned}$$



3. Nell'ipotesi di ripetere il gioco 100 volte si calcoli la probabilità di vincere almeno 59 volte.

Indicata con X la variabile casuale 'numero di successi' nelle 100 prove indipendenti risulta

$$X \sim \text{Bin}(n = 100, p = 0.6667).$$

$$M(X) = 66.67; \text{Var}(X) = 22.22.$$

Dal momento che $np = 66.67 > 5$ ed $n(1 - p) = 33.33 > 5$ è possibile utilizzare l'approssimazione della variabile casuale Binomiale con la variabile casuale Normale

$$X \sim N(\mu = 66.67, \sigma^2 = 22.22).$$

Si ottiene

$$\begin{aligned} P(X \geq 59) &= P\left(\frac{X - \mu}{\sigma} \geq \frac{59 - \mu}{\sigma}\right) = P\left(Z \geq \frac{59 - 66.67}{4.7148}\right) = \\ &= P(Z \geq -1.63) = P(Z \leq 1.63) = 0.9484. \end{aligned}$$

Tema T288

Prova scritta di STATISTICA del 11.02.2016

Prova scritta di
STATISTICA del
11.02.2016

Giuseppe Boari
Gabriele Cantaluppi



UNIVERSITÀ
CATTOLICA
del Sacro Cuore

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.

Giuseppe Boari
Gabriele Cantaluppi
Università Cattolica del Sacro Cuore, Milano



Al fine di studiare la relazione esistente tra il prezzo di un determinato prodotto e l'andamento economico, vengono rilevate le variabili X = 'andamento economico' e Y = 'prezzo del prodotto'. I dati sono raccolti nella seguente tabella:

$X \setminus Y$	40 - 50	50 - 70	70 - 80
<i>recessione</i>	4	4	4
<i>espansione</i>	4	4	0

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.



1. Fornire la rappresentazione grafica della distribuzione *marginale* di X e della *condizionata* $Y|X = \text{recessione}$.

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.



Si considera la distribuzione marginale della variabile X , ottenuta sommando le frequenze di riga

X_i	$n_{i\bullet}$
<i>recessione</i>	12
<i>espansione</i>	8

Si tratta di un carattere qualitativo ordinato che può essere rappresentato graficamente mediante un grafico a bastoncini o a canne d'organo.

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.

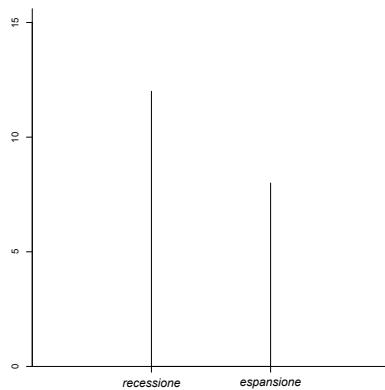


Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.





Si considera la distribuzione condizionata $Y|x_1$

y_j	n_{1j}
40 ÷ 50	4
50 ÷ 70	4
70 ÷ 80	4

Esercizio 1.

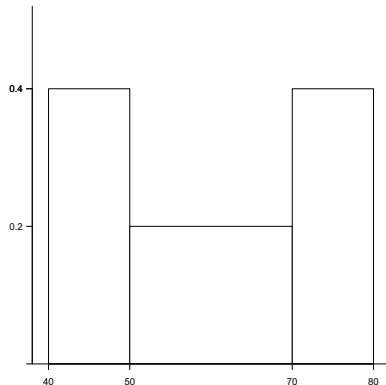
Esercizio 2.

Esercizio 3.

Esercizio 4.

Si tratta di un carattere quantitativo con modalità raggruppate in classi che può essere rappresentato graficamente mediante un istogramma. L'istogramma può essere costruito dopo avere calcolato le densità $d_{1j} = \frac{n_{1j}}{a_j}$, (dove a_j sono le ampiezze di classe).

y_j	n_{1j}	a_{1j}	d_{1j}
40 - 50	4	$50 - 40 = 10$	0.4
50 - 70	4	$70 - 50 = 20$	0.2
70 - 80	4	$80 - 70 = 10$	0.4



Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.



Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.

2. Si calcoli un opportuno indice di posizione per la variabile X e la media geometrica di Y .



Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.

Si tratta di un carattere qualitativo ordinato.

La moda corrisponde alla categoria alla quale è associata la massima frequenza) che risulta *recessione*.



La mediana (percentile di ordine 0.50) in presenza di un numero pari di osservazioni corrisponde a uno degli elementi che occupano le posizioni $\frac{n}{2} = 10$ ed $\frac{n}{2} + 1 = 11$ nella serie dei dati ordinati. Dalla distribuzione delle frequenze cumulate

x_i	$n_{i\bullet}$	$N_{i\bullet}$
<i>recessione</i>	12	12
<i>espansione</i>	8	20

si trova che i due elementi coincidono e risulta $x_{0.5} = \textit{recessione}$.

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.



Si considera la distribuzione marginale della variabile Y , ottenuta sommando le frequenze di colonna

y_j	$n_{\bullet j}$
40 $\rightarrow$ 50	8
50 $\rightarrow$ 70	8
70 $\rightarrow$ 80	4

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.



Si tratta di un carattere quantitativo con modalità raggruppate in classi.

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.



La media geometrica

$$\mu^{(0)}(Y) = \prod_{i=1}^k y_j^{f \cdot j}$$

risulta 55.9188.

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.



La media geometrica

$$\mu^{(0)}(Y) = \prod_{i=1}^k y_j^{f_{\bullet j}}$$

risulta 55.9188. Si ottiene (considerando come modalità i valori centrali delle classi)

y_j	$f_{\bullet j}$	$y_j^{f_{\bullet j}}$
45	0.4	4.5844
60	0.4	5.1435
75	0.2	2.3714
		55.9188

da cui segue

$$\mu^{(0)} = \prod_{j=1}^k y_j^{f_{\bullet j}} = 55.9188.$$

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.



Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.

3. Si calcoli con un opportuno indice la mutabilità di X , commentandone il risultato.



In presenza di un carattere qualitativo la variabilità può essere misurata mediante gli indici di Gini e di Frosini normalizzati

$$G_N = \frac{1 - \sum_{i=1}^h f_{i\bullet}^2}{1 - \frac{1}{h}}$$

e

$$F_N = 1 - \sqrt{1 - G_N}.$$

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.



In presenza di un carattere qualitativo la variabilità può essere misurata mediante gli indici di Gini e di Frosini normalizzati

$$G_N = \frac{1 - \sum_{i=1}^h f_{i\bullet}^2}{1 - \frac{1}{h}}$$

e

$$F_N = 1 - \sqrt{1 - G_N}.$$

Si ottiene:

x_i	$n_{i\bullet}$	$f_{i\bullet}$	$f_{i\bullet}^2$
<i>recessione</i>	12	0.6	0.36
<i>espansione</i>	8	0.4	0.16
	20	1	0.52

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.



In presenza di un carattere qualitativo la variabilità può essere misurata mediante gli indici di Gini e di Frosini normalizzati

$$G_N = \frac{1 - \sum_{i=1}^h f_{i\bullet}^2}{1 - \frac{1}{h}}$$

e

$$F_N = 1 - \sqrt{1 - G_N}.$$

Si ottiene:

x_i	$n_{i\bullet}$	$f_{i\bullet}$	$f_{i\bullet}^2$
<i>recessione</i>	12	0.6	0.36
<i>espansione</i>	8	0.4	0.16
	20	1	0.52

Da cui segue:

$$G_N = \frac{1 - 0.52}{1 - \frac{1}{2}} = 0.96$$

e

$$F_N = 1 - \sqrt{1 - 0.96} = 0.8.$$

Sussiste, quindi, un livello alto di eterogeneità.



4. Misurare la connessione tra le variabili X e Y e commentare il risultato.

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.



Dalla distribuzione delle frequenze congiunte n_{ij} :

	40 ÷ 50	50 ÷ 70	70 ÷ 80	$n_{i\bullet}$
<i>recessione</i>	4	4	4	12
<i>espansione</i>	4	4	0	8
$n_{\bullet j}$	8	8	4	20

è possibile ottenere le frequenze teoriche nella situazione di indipendenza stocastica

$$\hat{n}_{ij} = \frac{n_{i\bullet} n_{\bullet j}}{n}:$$

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.



Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.

Dalla distribuzione delle frequenze congiunte n_{ij} :

	40 ÷ 50	50 ÷ 70	70 ÷ 80	$n_{i\bullet}$
<i>recessione</i>	4	4	4	12
<i>espansione</i>	4	4	0	8
$n_{\bullet j}$	8	8	4	20

è possibile ottenere le frequenze teoriche nella situazione di indipendenza stocastica

$$\hat{n}_{ij} = \frac{n_{i\bullet} \cdot n_{\bullet j}}{n}:$$

	40 ÷ 50	50 ÷ 70	70 ÷ 80	$n_{i\bullet}$
<i>recessione</i>	4.8	4.8	2.4	12
<i>espansione</i>	3.2	3.2	1.6	8
$n_{\bullet j}$	8.0	8.0	4.0	20



le contingenze assolute $n_{ij} - \hat{n}_{ij}$:

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.



le contingenze assolute $n_{ij} - \hat{n}_{ij}$:

	40 ÷ 50	50 ÷ 70	70 ÷ 80
<i>recessione</i>	-0.8	-0.8	1.6
<i>espansione</i>	0.8	0.8	-1.6

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.



i rapporti di contingenza $\frac{n_{ij}}{\hat{n}_{ij}}$:

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.



i rapporti di contingenza $\frac{n_{ij}}{\hat{n}_{ij}}$:

	40 ÷ 50	50 ÷ 70	70 ÷ 80
<i>recessione</i>	0.8333	0.8333	1.6667
<i>espansione</i>	1.2500	1.2500	0.0000

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.



e le contingenze relative, espresse in termini percentuali, $100 \left(\frac{n_{ij} - \hat{n}_{ij}}{\hat{n}_{ij}} \right) \% = 100 \left(\frac{n_{ij}}{\hat{n}_{ij}} - 1 \right) \%$:

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.



e le contingenze relative, espresse in termini percentuali, $100 \left(\frac{n_{ij} - \hat{n}_{ij}}{\hat{n}_{ij}} \right) \% = 100 \left(\frac{n_{ij}}{\hat{n}_{ij}} - 1 \right) \%$:

	40 ÷ 50	50 ÷ 70	70 ÷ 80
<i>recessione</i>	-16.67	-16.67	66.67
<i>espansione</i>	25.00	25.00	-100.00

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.



Risultati intermedi $\frac{(n_{ij} - \hat{n}_{ij})^2}{\hat{n}_{ij}}$ per il calcolo del χ^2

	40 ÷ 50	50 ÷ 70	70 ÷ 80
<i>recessione</i>	0.1333	0.1333	1.0667
<i>espansione</i>	0.2000	0.2000	1.6000

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.



Risultati intermedi $\frac{n_{ij}^2}{n_{i\bullet} \cdot n_{\bullet j}}$ per il calcolo del χ^2 (formula operativa)

	40 ÷ 50	50 ÷ 70	70 ÷ 80
<i>recessione</i>	0.1667	0.1667	0.3333
<i>espansione</i>	0.2500	0.2500	0.0000

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.



Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.

$$\chi^2 = \sum_{i=1}^h \sum_{j=1}^k \frac{(n_{ij} - \hat{n}_{ij})^2}{\hat{n}_{ij}} = 3.3333,$$

$$\chi_N^2 = \frac{\chi^2}{n \cdot \min(h-1, k-1)} = 0.1667$$

$$\chi_N = \sqrt{\chi_N^2} = 0.4083.$$

Sussiste, quindi, un livello di connessione medio/basso.

Esercizio 2

Prova scritta di
STATISTICA del
11.02.2016

Giuseppe Boari
Gabriele Cantaluppi



UNIVERSITÀ
CATTOLICA
del Sacro Cuore

La seguente tabella riporta la serie dei NIBF di un titolo azionario quotato sul mercato finanziario:

t	$NIBF$
2011	0.95
2012	1.05
2013	1
2014	0.95
2015	0.9

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.



1. Si identifichi l'anno base e si interpreti il significato del numero indice a base fissa riferito al tempo 2015.

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.



1. Si identifichi l'anno base e si interpreti il significato del numero indice a base fissa riferito al tempo 2015.

Tra l'anno base 2013 e l'anno 2015 si osserva una riduzione del 10% del prezzo del titolo.

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.



1. Si identifichi l'anno base e si interpreti il significato del numero indice a base fissa riferito al tempo 2015.

Tra l'anno base 2013 e l'anno 2015 si osserva una riduzione del 10% del prezzo del titolo.
Abbiamo infatti $100 \cdot (0.9 - 1)\% = -10\%$

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.



2. Calcolare, mostrandone il procedimento, la serie dei NIBM.

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.



2. Calcolare, mostrandone il procedimento, la serie dei NIBM.

Si ricorda che ${}_{t-1}l_t = \frac{x_t}{x_{t-1}} = \frac{{}_{2013}l_t}{{}_{2013}l_{t-1}}$.

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.



2. Calcolare, mostrandone il procedimento, la serie dei NIBM.

Si ricorda che ${}_t l_t = \frac{x_t}{x_{t-1}} = \frac{{}_{2013}l_t}{{}_{2013}l_{t-1}}$. Si ottiene quindi

Anno	NIBF	NIBM
2011	0.95	
2012	1.05	1.1053
2013	1.00	0.9524
2014	0.90	0.9000
2015	0.90	1.0000

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.



3. Sapendo che il prezzo nel 2015 era 100, calcolare il prezzo del titolo nel 2011.

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.



3. Sapendo che il prezzo nel 2015 era 100, calcolare il prezzo del titolo nel 2011.

Ricordando che la serie dei NIBF è proporzionale a quella dei dati abbiamo

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.



Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.

3. Sapendo che il prezzo nel 2015 era 100, calcolare il prezzo del titolo nel 2011.

Ricordando che la serie dei NIBF è proporzionale a quella dei dati abbiamo

$$x_{2015} : 0.9 = x_{2011} : 0.95 \quad \text{ovvero} \quad \frac{x_{2015}}{x_{2011}} = \frac{0.9}{0.95}$$

da cui si ottiene



Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.

3. Sapendo che il prezzo nel 2015 era 100, calcolare il prezzo del titolo nel 2011.

Ricordando che la serie dei NIBF è proporzionale a quella dei dati abbiamo

$$x_{2015} : 0.9 = x_{2011} : 0.95 \quad \text{ovvero} \quad \frac{x_{2015}}{x_{2011}} = \frac{0.9}{0.95}$$

da cui si ottiene

$$x_{2011} = x_{2015} \cdot \frac{0.95}{0.9} = 105.5556.$$

oppure



Ricordando la definizione di NIBF

$$2013 I_t =$$

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.



Ricordando la definizione di NIBF

$${}_{2013}I_t = \frac{x_t}{x_{2013}}$$

abbiamo

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.



Ricordando la definizione di NIBF

$${}_{2013}I_t = \frac{X_t}{X_{2013}}$$

abbiamo

$${}_{2013}I_{2015} = \frac{X_{2015}}{X_{2013}} = 0.9$$

da cui

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.



Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.

Ricordando la definizione di NIBF

$${}_{2013}I_t = \frac{X_t}{X_{2013}}$$

abbiamo

$${}_{2013}I_{2015} = \frac{X_{2015}}{X_{2013}} = 0.9$$

da cui

$$X_{2013} = \frac{X_{2015}}{0.9} = \frac{100}{0.9} = 111.1111$$

e

$$X_{2011} = X_{2013} \cdot$$



Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.

Ricordando la definizione di NIBF

$${}_{2013}I_t = \frac{X_t}{X_{2013}}$$

abbiamo

$${}_{2013}I_{2015} = \frac{X_{2015}}{X_{2013}} = 0.9$$

da cui

$$X_{2013} = \frac{X_{2015}}{0.9} = \frac{100}{0.9} = 111.1111$$

e

$$\begin{aligned} X_{2011} &= X_{2013} \cdot {}_{2015}I_{2013} = X_{2013} \cdot \frac{X_{2015}}{X_{2013}} = \\ &= \end{aligned}$$



Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.

Ricordando la definizione di NIBF

$${}_{2013}I_t = \frac{X_t}{X_{2013}}$$

abbiamo

$${}_{2013}I_{2015} = \frac{X_{2015}}{X_{2013}} = 0.9$$

da cui

$$X_{2013} = \frac{X_{2015}}{0.9} = \frac{100}{0.9} = 111.1111$$

e

$$\begin{aligned} X_{2011} &= X_{2013} \cdot {}_{2015}I_{2013} = X_{2013} \cdot \frac{X_{2015}}{X_{2013}} = \\ &= X_{2013} \cdot 0.95 = \end{aligned}$$



Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.

Ricordando la definizione di NIBF

$${}_{2013}I_t = \frac{X_t}{X_{2013}}$$

abbiamo

$${}_{2013}I_{2015} = \frac{X_{2015}}{X_{2013}} = 0.9$$

da cui

$$X_{2013} = \frac{X_{2015}}{0.9} = \frac{100}{0.9} = 111.1111$$

e

$$\begin{aligned} X_{2011} &= X_{2013} \cdot {}_{2015}I_{2013} = X_{2013} \cdot \frac{X_{2015}}{X_{2013}} = \\ &= X_{2013} \cdot 0.95 = 111.1111 \cdot 0.95 = 105.5556. \end{aligned}$$

oppure



Dalla serie dei NIBF possiamo ricavare

$$2011 / 2015 =$$

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.



Dalla serie dei NIBF possiamo ricavare

$${}_{2011}l_{2015} = \frac{{}_{2013}l_{2015}}{{}_{2013}l_{2011}} = \frac{0.90}{0.95} = 0.9474$$

e dal momento che, con riferimento ai dati iniziali vale

$${}_{2011}l_{2015} =$$

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.



Dalla serie dei NIBF possiamo ricavare

$${}_{2011}l_{2015} = \frac{{}_{2013}l_{2015}}{{}_{2013}l_{2011}} = \frac{0.90}{0.95} = 0.9474$$

e dal momento che, con riferimento ai dati iniziali vale

$${}_{2011}l_{2015} = \frac{x_{2015}}{x_{2011}}$$

abbiamo

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.



Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.

Dalla serie dei NIBF possiamo ricavare

$${}_{2011}I_{2015} = \frac{{}_{2013}I_{2015}}{{}_{2013}I_{2011}} = \frac{0.90}{0.95} = 0.9474$$

e dal momento che, con riferimento ai dati iniziali vale

$${}_{2011}I_{2015} = \frac{X_{2015}}{X_{2011}}$$

abbiamo

$$X_{2011} = \frac{X_{2015}}{{}_{2011}I_{2015}} =$$



Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.

Dalla serie dei NIBF possiamo ricavare

$${}_{2011}I_{2015} = \frac{{}_{2013}I_{2015}}{{}_{2013}I_{2011}} = \frac{0.90}{0.95} = 0.9474$$

e dal momento che, con riferimento ai dati iniziali vale

$${}_{2011}I_{2015} = \frac{x_{2015}}{x_{2011}}$$

abbiamo

$$x_{2011} = \frac{x_{2015}}{{}_{2011}I_{2015}} = \frac{100}{0.9474} =$$



Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.

Dalla serie dei NIBF possiamo ricavare

$${}_{2011}l_{2015} = \frac{{}_{2013}l_{2015}}{{}_{2013}l_{2011}} = \frac{0.90}{0.95} = 0.9474$$

e dal momento che, con riferimento ai dati iniziali vale

$${}_{2011}l_{2015} = \frac{x_{2015}}{x_{2011}}$$

abbiamo

$$x_{2011} = \frac{x_{2015}}{{}_{2011}l_{2015}} = \frac{100}{0.9474} = 105.552$$



4. Calcolare il tasso medio di variazione annuo nell'intervallo 2012-2015, commentandone il risultato.

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.



4. Calcolare il tasso medio di variazione annuo nell'intervallo 2012-2015, commentandone il risultato.

Si ricorda che

$${}_{2012}I_{2015} = {}_{2012}I_{2013} \cdot {}_{2013}I_{2014} \cdot {}_{2014}I_{2015}$$

Si deve trovare il numero indice a base mobile medio.

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.



Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.

4. Calcolare il tasso medio di variazione annuo nell'intervallo 2012-2015, commentandone il risultato.

Si ricorda che

$${}_{2012}I_{2015} = {}_{2012}I_{2013} \cdot {}_{2013}I_{2014} \cdot {}_{2014}I_{2015}$$

Si deve trovare il numero indice a base mobile medio. Applicando il criterio di scelta della media secondo Chisini si ottiene

$${}_{2012}I_{2015} = {}_{2012}I_{2013} \cdot {}_{2013}I_{2014} \cdot {}_{2014}I_{2015} = \alpha \cdot \alpha \cdot \alpha = \alpha^3.$$



Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.

4. Calcolare il tasso medio di variazione annuo nell'intervallo 2012-2015, commentandone il risultato.

Si ricorda che

$${}_{2012}l_{2015} = {}_{2012}l_{2013} \cdot {}_{2013}l_{2014} \cdot {}_{2014}l_{2015}$$

Si deve trovare il numero indice a base mobile medio. Applicando il criterio di scelta della media secondo Chisini si ottiene

$${}_{2012}l_{2015} = {}_{2012}l_{2013} \cdot {}_{2013}l_{2014} \cdot {}_{2014}l_{2015} = \alpha \cdot \alpha \cdot \alpha = \alpha^3.$$

Quindi

$$\alpha = \sqrt[3]{{}_{2012}l_{2015}} = \sqrt[3]{{}_{2012}l_{2013} \cdot {}_{2013}l_{2014} \cdot {}_{2014}l_{2015}} = 0.9499.$$

Da cui segue una variazione relativa media pari a -0.0501 , corrispondente a un decremento medio annuo tra il 2012 e il 2015 del 5.01%.



Il seguente prospetto riporta i valori rilevati in corrispondenza a 5 unità statistiche con riferimento a due caratteri quantitativi X e Y

x_i	7	8	9	10	11
y_i	2	2	4	3	2

1. Si rappresentino graficamente le coppie dei dati ipotizzando Y in funzione di X .

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.



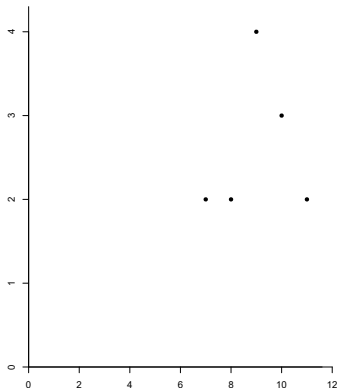
Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.

Le coppie di dati possono essere rappresentate mediante un grafico a dispersione.





2. Si calcolino, secondo il metodo dei minimi quadrati, i parametri dei seguenti modelli di regressione:

$$I) Y^* = a + b \cdot X \quad II) Y^* = \beta \cdot X$$

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.

Si riportano i calcoli necessari¹ per ottenere le stime dei parametri e gli indici di adattamento del modello

$$Y = a + bX + E$$

x_i	y_i	$x_i y_i$	x_i^2	y_i^2
7	2	14	49	4
8	2	16	64	4
9	4	36	81	16
10	3	30	100	9
11	2	22	121	4
45	13	118	415	37
9	2.6	23.6	83	7.4
$M(X)$	$M(Y)$	$M(XY)$	$M(X^2)$	$M(Y^2)$

$$\hat{b} =$$



Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.

¹Si osserva che per ragioni di stampa alcuni calcoli potrebbero essere duplicati con riferimento ai diversi modelli oggetto di studio

Si riportano i calcoli necessari¹ per ottenere le stime dei parametri e gli indici di adattamento del modello

$$Y = a + bX + E$$

x_i	y_i	$x_i y_i$	x_i^2	y_i^2
7	2	14	49	4
8	2	16	64	4
9	4	36	81	16
10	3	30	100	9
11	2	22	121	4
45	13	118	415	37
9	2.6	23.6	83	7.4
$M(X)$	$M(Y)$	$M(XY)$	$M(X^2)$	$M(Y^2)$

$$\begin{aligned}\hat{b} &= \frac{\text{Cov}(X, Y)}{\text{Var}(X)} = \frac{M(XY) - M(X)M(Y)}{M(X^2) - [M(X)]^2} = \\ &= \frac{23.6 - 9 \cdot 2.6}{83 - 9^2} = \frac{0.2}{2} = 0.1 \\ \hat{a} &= \end{aligned}$$

¹Si osserva che per ragioni di stampa alcuni calcoli potrebbero essere duplicati con riferimento ai diversi modelli oggetto di studio



Si riportano i calcoli necessari¹ per ottenere le stime dei parametri e gli indici di adattamento del modello

$$Y = a + bX + E$$

x_i	y_i	$x_i y_i$	x_i^2	y_i^2
7	2	14	49	4
8	2	16	64	4
9	4	36	81	16
10	3	30	100	9
11	2	22	121	4
45	13	118	415	37
9	2.6	23.6	83	7.4
$M(X)$	$M(Y)$	$M(XY)$	$M(X^2)$	$M(Y^2)$

$$\begin{aligned}\hat{b} &= \frac{\text{Cov}(X, Y)}{\text{Var}(X)} = \frac{M(XY) - M(X)M(Y)}{M(X^2) - [M(X)]^2} = \\ &= \frac{23.6 - 9 \cdot 2.6}{83 - 9^2} = \frac{0.2}{2} = 0.1 \\ \hat{a} &= M(Y) - \hat{b}M(X) = 1.7 \\ \rho^2 &= \end{aligned}$$

¹Si osserva che per ragioni di stampa alcuni calcoli potrebbero essere duplicati con riferimento ai diversi modelli oggetto di studio



Si riportano i calcoli necessari¹ per ottenere le stime dei parametri e gli indici di adattamento del modello

$$Y = a + bX + E$$

x_i	y_i	$x_i y_i$	x_i^2	y_i^2
7	2	14	49	4
8	2	16	64	4
9	4	36	81	16
10	3	30	100	9
11	2	22	121	4
45	13	118	415	37
9	2.6	23.6	83	7.4
$M(X)$	$M(Y)$	$M(XY)$	$M(X^2)$	$M(Y^2)$

$$\hat{b} = \frac{\text{Cov}(X, Y)}{\text{Var}(X)} = \frac{M(XY) - M(X)M(Y)}{M(X^2) - [M(X)]^2} =$$

$$= \frac{23.6 - 9 \cdot 2.6}{83 - 9^2} = \frac{0.2}{2} = 0.1$$

$$\hat{a} = M(Y) - \hat{b}M(X) = 1.7$$

$$\rho^2 = \frac{[\text{Cov}(X, Y)]^2}{\text{Var}(X) \text{Var}(Y)} = \frac{0.2^2}{2 \cdot 0.64} = 0.0313$$

$$M(\hat{E}^2) = \text{Var}(\hat{E}) = \text{varianza residua} =$$

¹Si osserva che per ragioni di stampa alcuni calcoli potrebbero essere duplicati con riferimento ai diversi modelli oggetto di studio



Si riportano i calcoli necessari¹ per ottenere le stime dei parametri e gli indici di adattamento del modello

$$Y = a + bX + E$$

x_i	y_i	$x_i y_i$	x_i^2	y_i^2
7	2	14	49	4
8	2	16	64	4
9	4	36	81	16
10	3	30	100	9
11	2	22	121	4
45	13	118	415	37
9	2.6	23.6	83	7.4
$M(X)$	$M(Y)$	$M(XY)$	$M(X^2)$	$M(Y^2)$

$$\begin{aligned}\hat{b} &= \frac{\text{Cov}(X, Y)}{\text{Var}(X)} = \frac{M(XY) - M(X)M(Y)}{M(X^2) - [M(X)]^2} = \\ &= \frac{23.6 - 9 \cdot 2.6}{83 - 9^2} = \frac{0.2}{2} = 0.1 \\ \hat{a} &= M(Y) - \hat{b}M(X) = 1.7 \\ \rho^2 &= \frac{[\text{Cov}(X, Y)]^2}{\text{Var}(X) \text{Var}(Y)} = \frac{0.2^2}{2 \cdot 0.64} = 0.0313\end{aligned}$$

$$M(\hat{E}^2) = \text{Var}(\hat{E}) = \text{varianza residua} = \text{Var}(Y) \cdot (1 - \rho^2) = 0.62$$

¹Si osserva che per ragioni di stampa alcuni calcoli potrebbero essere duplicati con riferimento ai diversi modelli oggetto di studio





Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.

Si riportano i calcoli necessari² per ottenere le stime dei parametri e la misura di adattamento del modello

$$Y = \beta X + E$$

x_i	y_i	$x_i y_i$	x_i^2	y_i^2
7	2	14	49	4
8	2	16	64	4
9	4	36	81	16
10	3	30	100	9
11	2	22	121	4
		118	415	37
		23.6	83	7.4
		$M(XY)$	$M(X^2)$	$M(Y^2)$

$$\hat{\beta} =$$

²Si osserva che per ragioni di stampa alcuni calcoli potrebbero essere duplicati con riferimento ai diversi modelli oggetto di studio



Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.

Si riportano i calcoli necessari² per ottenere le stime dei parametri e la misura di adattamento del modello

$$Y = \beta X + E$$

x_i	y_i	$x_i y_i$	x_i^2	y_i^2
7	2	14	49	4
8	2	16	64	4
9	4	36	81	16
10	3	30	100	9
11	2	22	121	4
		118	415	37
		23.6	83	7.4
		$M(XY)$	$M(X^2)$	$M(Y^2)$

$$\hat{\beta} = \frac{M(XY)}{M(X^2)} = \frac{23.6}{83} = 0.2843$$

$$M(\hat{E}^2) =$$

²Si osserva che per ragioni di stampa alcuni calcoli potrebbero essere duplicati con riferimento ai diversi modelli oggetto di studio



Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.

Si riportano i calcoli necessari² per ottenere le stime dei parametri e la misura di adattamento del modello

$$Y = \beta X + E$$

x_i	y_i	$x_i y_i$	x_i^2	y_i^2
7	2	14	49	4
8	2	16	64	4
9	4	36	81	16
10	3	30	100	9
11	2	22	121	4
		118	415	37
		23.6	83	7.4
		$M(XY)$	$M(X^2)$	$M(Y^2)$

$$\hat{\beta} = \frac{M(XY)}{M(X^2)} = \frac{23.6}{83} = 0.2843$$

$$M(\hat{E}^2) = M[Y \cdot (Y - \hat{\beta}X)] = M(Y^2) - \hat{\beta}M(XY) = 7.4 - 0.2843 \cdot 23.6 = 0.6896$$

²Si osserva che per ragioni di stampa alcuni calcoli potrebbero essere duplicati con riferimento ai diversi modelli oggetto di studio



Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.

$\hat{a} = 1.7$, $\hat{b} = 0.1$, indice adattamento $M(\hat{E}^2) = 0.62$ $\hat{\beta} = 0.2843$, indice adattamento
 $M(\hat{E}^2) = 0.6896$



3. Si indichi, motivando, quale modello presenta il migliore adattamento.

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.



3. Si indichi, motivando, quale modello presenta il migliore adattamento.

Il modello migliore è $Y^* = 1.7 + 0.1 \cdot X$ in quanto caratterizzato da residuo quadratico medio (varianza residua) inferiore.

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.



Nell'urna U_1 vengono inserite 2 palline blu e 2 palline rosse e si esegue la seguente procedura:

- si estrae una prima pallina;
- la pallina estratta viene reinserita nell'urna, insieme a 2 ulteriori palline del suo stesso colore;
- si estrae una seconda pallina dall'urna così modificata.

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.



1. Calcolare la probabilità che la prima pallina estratta sia blu.

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.



Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.

1. Calcolare la probabilità che la prima pallina estratta sia blu.

$$P(\text{prima pallina estratta blu}) = \frac{n. \text{ palline blu}}{n. \text{ palline}} = \frac{2}{4} = 0.5.$$



2. Usando la legge delle probabilità totali, calcolare la probabilità che la seconda pallina estratta sia blu.

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.



2. Usando la legge delle probabilità totali, calcolare la probabilità che la seconda pallina estratta sia blu.

Indicati con B_i e R_i gli eventi pallina blu alla i -esima estrazione e pallina rossa alla i -esima estrazione ($i = 1, 2$) abbiamo

$$P(B_2) = P(B_1) \cdot$$

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.



2. Usando la legge delle probabilità totali, calcolare la probabilità che la seconda pallina estratta sia blu.

Indicati con B_i e R_i gli eventi pallina blu alla i -esima estrazione e pallina rossa alla i -esima estrazione ($i = 1, 2$) abbiamo

$$P(B_2) = P(B_1) \cdot P(B_2|B_1) +$$

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.



2. Usando la legge delle probabilità totali, calcolare la probabilità che la seconda pallina estratta sia blu.

Indicati con B_i e R_i gli eventi pallina blu alla i -esima estrazione e pallina rossa alla i -esima estrazione ($i = 1, 2$) abbiamo

$$P(B_2) = P(B_1) \cdot P(B_2|B_1) + P(R_1) \cdot$$

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.



2. Usando la legge delle probabilità totali, calcolare la probabilità che la seconda pallina estratta sia blu.

Indicati con B_i e R_i gli eventi pallina blu alla i -esima estrazione e pallina rossa alla i -esima estrazione ($i = 1, 2$) abbiamo

$$P(B_2) = P(B_1) \cdot P(B_2|B_1) + P(R_1) \cdot P(B_2|R_1) =$$

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.



Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.

2. Usando la legge delle probabilità totali, calcolare la probabilità che la seconda pallina estratta sia blu.

Indicati con B_i e R_i gli eventi pallina blu alla i -esima estrazione e pallina rossa alla i -esima estrazione ($i = 1, 2$) abbiamo

$$\begin{aligned} P(B_2) &= P(B_1) \cdot P(B_2|B_1) + P(R_1) \cdot P(B_2|R_1) = \\ &= \frac{2}{4} \cdot \end{aligned}$$



Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.

2. Usando la legge delle probabilità totali, calcolare la probabilità che la seconda pallina estratta sia blu.

Indicati con B_i e R_i gli eventi pallina blu alla i -esima estrazione e pallina rossa alla i -esima estrazione ($i = 1, 2$) abbiamo

$$\begin{aligned} P(B_2) &= P(B_1) \cdot P(B_2|B_1) + P(R_1) \cdot P(B_2|R_1) = \\ &= \frac{2}{4} \cdot \frac{4}{6} + \end{aligned}$$



Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.

2. Usando la legge delle probabilità totali, calcolare la probabilità che la seconda pallina estratta sia blu.

Indicati con B_i e R_i gli eventi pallina blu alla i -esima estrazione e pallina rossa alla i -esima estrazione ($i = 1, 2$) abbiamo

$$\begin{aligned} P(B_2) &= P(B_1) \cdot P(B_2|B_1) + P(R_1) \cdot P(B_2|R_1) = \\ &= \frac{2}{4} \cdot \frac{4}{6} + \frac{2}{4} \cdot \end{aligned}$$



Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.

2. Usando la legge delle probabilità totali, calcolare la probabilità che la seconda pallina estratta sia blu.

Indicati con B_i e R_i gli eventi pallina blu alla i -esima estrazione e pallina rossa alla i -esima estrazione ($i = 1, 2$) abbiamo

$$\begin{aligned} P(B_2) &= P(B_1) \cdot P(B_2|B_1) + P(R_1) \cdot P(B_2|R_1) = \\ &= \frac{2}{4} \cdot \frac{4}{6} + \frac{2}{4} \cdot \frac{2}{6} = \end{aligned}$$



2. Usando la legge delle probabilità totali, calcolare la probabilità che la seconda pallina estratta sia blu.

Indicati con B_i e R_i gli eventi pallina blu alla i -esima estrazione e pallina rossa alla i -esima estrazione ($i = 1, 2$) abbiamo

$$\begin{aligned} P(B_2) &= P(B_1) \cdot P(B_2|B_1) + P(R_1) \cdot P(B_2|R_1) = \\ &= \frac{2}{4} \cdot \frac{4}{6} + \frac{2}{4} \cdot \frac{2}{6} = 0.3333 + 0.1667 = \end{aligned}$$



Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.

2. Usando la legge delle probabilità totali, calcolare la probabilità che la seconda pallina estratta sia blu.

Indicati con B_i e R_i gli eventi pallina blu alla i -esima estrazione e pallina rossa alla i -esima estrazione ($i = 1, 2$) abbiamo

$$\begin{aligned} P(B_2) &= P(B_1) \cdot P(B_2|B_1) + P(R_1) \cdot P(B_2|R_1) = \\ &= \frac{2}{4} \cdot \frac{4}{6} + \frac{2}{4} \cdot \frac{2}{6} = 0.3333 + 0.1667 = 0.5. \end{aligned}$$



3. Gli eventi $B_1 = \{\text{prima pallina estratta è blu}\}$ e $B_2 = \{\text{seconda pallina estratta è blu}\}$ sono indipendenti?

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.



Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.

3. Gli eventi $B_1 = \{\text{prima pallina estratta è blu}\}$ e $B_2 = \{\text{seconda pallina estratta è blu}\}$ sono indipendenti?

Gli eventi B_1 e B_2 non sono indipendenti, abbiamo infatti

$$P(B_2) = 0.5 \neq P(B_2|B_1) = \frac{4}{6}$$

vale inoltre



Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.

3. Gli eventi $B_1 = \{\text{prima pallina estratta è blu}\}$ e $B_2 = \{\text{seconda pallina estratta è blu}\}$ sono indipendenti?

Gli eventi B_1 e B_2 non sono indipendenti, abbiamo infatti

$$P(B_2) = 0.5 \neq P(B_2|B_1) = \frac{4}{6}$$

vale inoltre

$$P(B_1 \cap B_2) = 0.3333 \neq P(B_1) \cdot P(B_2) = 0.5 \cdot 0.5 = 0.25.$$



4. Trovare il percentile di ordine 0.025 di X Gaussiana con media 1 e varianza 4 , cioè x_p tale che $P(X < x_p) = 0.025$.

$$P(X \leq x_p) = 0.025$$

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.



Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.

4. Trovare il percentile di ordine 0.025 di X Gaussiana con media 1 e varianza 4 , cioè x_p tale che $P(X < x_p) = 0.025$.

$$P(X \leq x_p) = 0.025$$
$$P\left(\frac{X - \mu}{\sigma} \leq \frac{x_p - \mu}{\sigma}\right) = 0.025$$



Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.

4. Trovare il percentile di ordine 0.025 di X Gaussiana con media 1 e varianza 4 , cioè x_p tale che $P(X < x_p) = 0.025$.

$$P(X \leq x_p) = 0.025$$

$$P\left(\frac{X - \mu}{\sigma} \leq \frac{x_p - \mu}{\sigma}\right) = 0.025$$

$$P\left(Z \leq \frac{x_p - 1}{2}\right) = P(Z \leq z_p) = 0.025$$

$$\frac{x_p - 1}{2} = z_p =$$



4. Trovare il percentile di ordine 0.025 di X Gaussiana con media 1 e varianza 4 , cioè x_p tale che $P(X < x_p) = 0.025$.

$$P(X \leq x_p) = 0.025$$

$$P\left(\frac{X - \mu}{\sigma} \leq \frac{x_p - \mu}{\sigma}\right) = 0.025$$

$$P\left(Z \leq \frac{x_p - 1}{2}\right) = P(Z \leq z_p) = 0.025$$

$$\frac{x_p - 1}{2} = z_p = -1.96$$



Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.

4. Trovare il percentile di ordine 0.025 di X Gaussiana con media 1 e varianza 4 , cioè x_p tale che $P(X < x_p) = 0.025$.

$$P(X \leq x_p) = 0.025$$

$$P\left(\frac{X - \mu}{\sigma} \leq \frac{x_p - \mu}{\sigma}\right) = 0.025$$

$$P\left(Z \leq \frac{x_p - 1}{2}\right) = P(Z \leq z_p) = 0.025$$

$$\frac{x_p - 1}{2} = z_p = -1.96$$

$$x_p =$$



4. Trovare il percentile di ordine 0.025 di X Gaussiana con media 1 e varianza 4 , cioè x_p tale che $P(X < x_p) = 0.025$.

$$P(X \leq x_p) = 0.025$$

$$P\left(\frac{X - \mu}{\sigma} \leq \frac{x_p - \mu}{\sigma}\right) = 0.025$$

$$P\left(Z \leq \frac{x_p - 1}{2}\right) = P(Z \leq z_p) = 0.025$$

$$\frac{x_p - 1}{2} = z_p = -1.96$$

$$x_p = 1 - 1.96 \cdot 2 = -2.92.$$

Tema T289-1

Prova scritta di STATISTICA del 09.06.2016

Prova scritta di
STATISTICA del
09.06.2016

Giuseppe Boari
Gabriele Cantaluppi



UNIVERSITÀ
CATTOLICA
del Sacro Cuore

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.

Giuseppe Boari
Gabriele Cantaluppi
Università Cattolica del Sacro Cuore, Milano

Esercizio 1

Prova scritta di
STATISTICA del
09.06.2016

Giuseppe Boari
Gabriele Cantaluppi



UNIVERSITÀ
CATTOLICA
del Sacro Cuore

La seguente tabella riporta alcune informazioni sulla distribuzione congiunta della variabile statistica doppia (X, Y) .

$X \setminus Y$	y_1	y_2	y_3
2		2	
4			

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.



1. Si completino le frequenze nella tabella in modo tale che siano verificate contemporaneamente le due seguenti condizioni: a) la distribuzione $Y|x_1$ sia caratterizzata da assenza di variabilità; b) $n_{21} = n_{23}$; c) $\chi^2 = \chi_{\max}^2 = 6$.

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.



1. Si completino le frequenze nella tabella in modo tale che siano verificate contemporaneamente le due seguenti condizioni: a) la distribuzione $Y|x_1$ sia caratterizzata da assenza di variabilità; b) $n_{21} = n_{23}$; c) $\chi^2 = \chi_{\max}^2 = 6$.

La situazione di assenza di variabilità si realizza in presenza di una variabile statistica degenerare, quindi $n_{11} = n_{13} = 0$ e $n_{1\bullet} = 2$.

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.



Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.

1. Si completino le frequenze nella tabella in modo tale che siano verificate contemporaneamente le due seguenti condizioni: a) la distribuzione $Y|X_1$ sia caratterizzata da assenza di variabilità; b) $n_{21} = n_{23}$; c) $\chi^2 = \chi_{\max}^2 = 6$.

La situazione di assenza di variabilità si realizza in presenza di una variabile statistica degenerare, quindi $n_{11} = n_{13} = 0$ e $n_{1\bullet} = 2$.

$\chi^2 = \chi_{\max}^2$ si realizza solo in presenza di dipendenza funzionale, nel caso in esame di X da Y . Le distribuzioni condizionate $X|y_j, j = 1, 2, 3$ sono degeneri, quindi $n_{22} = 0$. Ricordando che $\chi_{\max}^2 = n \cdot \min(h - 1, k - 1) = n \cdot \min(2 - 1, 3 - 1) = n \cdot \min(1, 2) = n \cdot 1 = n = 6$ segue che $n_{21} = n_{23} = 3$.

$X \setminus Y$	y_1	y_2	y_3	
2	0	2	0	2
4	2	0	2	4
	2	2	2	6



2. Si sostituiscano alle modalità y_1 , y_2 e y_3 di Y le classi $2 \div 4$, $4 \div 6$ e $6 \div 10$, e si rappresentino graficamente le distribuzioni marginali X e Y .

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.



2. Si sostituiscano alle modalità y_1 , y_2 e y_3 di Y le classi $2 \div 4$, $4 \div 6$ e $6 \div 10$, e si rappresentino graficamente le distribuzioni marginali X e Y .

$X \setminus Y$	$2 \div 4$	$4 \div 6$	$6 \div 10$	
2	0	2	0	2
4	2	0	2	4
	2	2	2	6

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.



Si considera la distribuzione marginale della variabile X , ottenuta sommando le frequenze di riga

X_i	$n_{i\bullet}$
2	2
4	4

Esercizio 1.

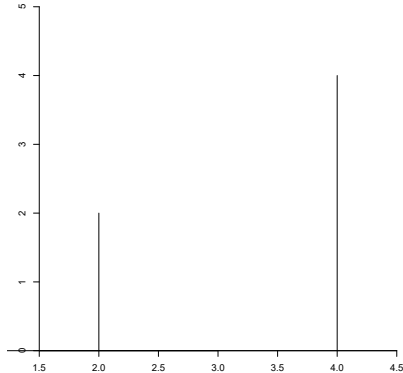
Esercizio 2.

Esercizio 3.

Esercizio 4.



Si tratta di un carattere quantitativo con modalità non raggruppate in classi che può essere rappresentato graficamente mediante un grafico a bastoncini.



Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.



Si considera la distribuzione marginale della variabile Y , ottenuta sommando le frequenze di colonna

y_j	$n_{\bullet j}$
2 $\rightarrow$ 4	2
4 $\rightarrow$ 6	2
6 $\rightarrow$ 10	2

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.



Si tratta di un carattere quantitativo con modalità raggruppate in classi che può essere rappresentato graficamente mediante un istogramma.

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.



L'istogramma può essere costruito dopo avere calcolato le densità $d_{\bullet j} = \frac{n_{\bullet j}}{a_j}$, (dove a_j sono le ampiezze di classe).

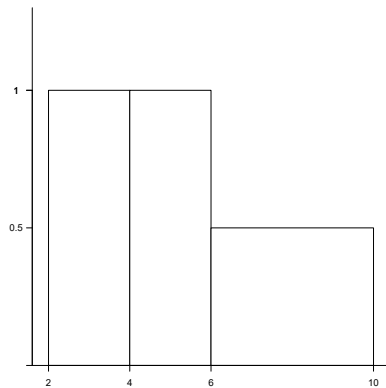
y_j	$n_{\bullet j}$	$a_{\bullet j}$	$d_{\bullet j}$
2 ÷ 4	2	4 - 2 = 2	1
4 ÷ 6	2	6 - 4 = 2	1
6 ÷ 10	2	10 - 6 = 4	0.5

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.



Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.



3. Si riassumano le distribuzioni marginali X e Y con opportuni indici di posizione e variabilità.

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.



Si considera la distribuzione marginale della variabile X , ottenuta sommando le frequenze di riga

X_i	$n_{i\bullet}$
2	2
4	4

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.



Si tratta di un carattere quantitativo con modalità non raggruppate in classi.

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.



La media aritmetica

$$M(X) = \frac{1}{n} \sum_{i=1}^h x_i n_i.$$

risulta 3.3333.

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.



La media aritmetica

$$M(X) = \frac{1}{n} \sum_{i=1}^h x_i n_{i\bullet}$$

risulta 3.3333.

Si ottiene

x_i	$n_{i\bullet}$	$x_i n_{i\bullet}$
2	2	4
4	4	16
	6	20

da cui segue

$$M(X) = \frac{1}{n} \sum_{i=1}^h x_i n_{i\bullet} = \frac{20}{6} = 3.3333.$$

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.



La varianza

$$\text{Var}(X) = M[(X - \mu_X)^2] = M(X^2) - \mu_X^2 = \frac{1}{n} \sum_{i=1}^h x_i^2 n_{i\bullet} - \mu_X^2$$

risulta 0.8889.

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.



La varianza

$$\text{Var}(X) = M[(X - \mu_X)^2] = M(X^2) - \mu_X^2 = \frac{1}{n} \sum_{i=1}^h x_i^2 n_{i\bullet} - \mu_X^2$$

risulta 0.8889.

Si ottiene

x_i	$n_{i\bullet}$	$x_i n_{i\bullet}$	$x_i^2 n_{i\bullet}$
2	2	4	8
4	4	16	64
	6	20	72

da cui segue

$$\begin{aligned} \text{Var}(X) = \sigma_X^2 &= M(X^2) - \mu_X^2 = \\ &= \frac{72}{6} - 3.3333^2 = 12 - 11.1111 = 0.8889. \end{aligned}$$

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.



Per effettuare confronti in termini di variabilità tra caratteri quantitativi occorre utilizzare una misura relativa, ad esempio il coefficiente di variazione

$$CV = \frac{\sigma}{\mu} = \frac{\sqrt{\text{Var}(X)}}{M(X)} = \frac{\sqrt{0.8889}}{3.3333} = \frac{0.9428}{3.3333} = 0.2828.$$

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.



Si considera la distribuzione marginale della variabile Y , ottenuta sommando le frequenze di colonna

y_j	$n_{\bullet j}$
2 $\rightarrow$ 4	2
4 $\rightarrow$ 6	2
6 $\rightarrow$ 10	2

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.



Si tratta di un carattere quantitativo con modalità raggruppate in classi.

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.



La media aritmetica

$$M(Y) = \frac{1}{n} \sum_{j=1}^k y_j n_{\bullet j}$$

risulta 5.3333.

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.



La media aritmetica

$$M(Y) = \frac{1}{n} \sum_{j=1}^k y_j n_{\bullet j}$$

risulta 5.3333.

Si ottiene (considerando come modalità i valori centrali delle classi)

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.



La media aritmetica

$$M(Y) = \frac{1}{n} \sum_{j=1}^k y_j n_{\bullet j}$$

risulta 5.3333.

Si ottiene (considerando come modalità i valori centrali delle classi)

y_j	$n_{\bullet j}$	$y_j n_{\bullet j}$
3	2	6
5	2	10
8	2	16
	6	32

da cui segue

$$M(Y) = \frac{1}{n} \sum_{i=1}^k y_i n_{\bullet i} = \frac{32}{6} = 5.3333.$$

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.



La varianza

$$\text{Var}(Y) = M[(Y - \mu_Y)^2] = M(Y^2) - \mu_Y^2 = \frac{1}{n} \sum_{j=1}^k y_j^2 n_{\bullet j} - \mu_Y^2$$

risulta 4.2222.

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.



La varianza

$$\text{Var}(Y) = M[(Y - \mu_Y)^2] = M(Y^2) - \mu_Y^2 = \frac{1}{n} \sum_{j=1}^k y_j^2 n_{\bullet j} - \mu_Y^2$$

risulta 4.2222.

Si ottiene (considerando come modalità i valori centrali delle classi)

y_j	$n_{\bullet j}$	$y_j n_{\bullet j}$	$y_j^2 n_{\bullet j}$
3	2	6	18
5	2	10	50
8	2	16	128
	6	32	196

da cui segue

$$\begin{aligned} \text{Var}(Y) = \sigma_Y^2 &= M(Y^2) - \mu_Y^2 = \\ &= \frac{196}{6} - 5.3333^2 = 32.6667 - 28.4444 = 4.2223. \end{aligned}$$

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.



Per effettuare confronti in termini di variabilità tra caratteri quantitativi occorre utilizzare una misura relativa, ad esempio il coefficiente di variazione

$$CV = \frac{\sigma}{\mu} = \frac{\sqrt{\text{Var}(Y)}}{M(Y)} = \frac{\sqrt{4.2223}}{5.3333} = \frac{2.0548}{5.3333} = 0.3853.$$

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.



Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.

X Indice di posizione: $M(X) = 3.3333$ Indice di variabilità: $CV(X) = 0.2828$

Y Indice di posizione: $M(Y) = 5.3333$ Indice di variabilità: $CV(Y) = 0.3853$

La variabile Y è caratterizzata da un livello medio e da una variabilità superiori a X .



4. Si determini il grado di connessione tra le variabili X e Y mediante il calcolo di un opportuno indice.

Sussiste dipendenza funzionale di X da Y , quindi $\chi_N = 1$.

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.

Esercizio 2



Si consideri il seguente prospetto:

t	x_t	$NIBF(2010 = 1)$	$NIBM$
2008		4	
2009	100		
2010		1	0.5
2011	100		
2012			2

1. Si completi il prospetto con i dati mancanti.

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.



Considerando i numeri indici a base mobile assegnati è possibile ricostruire la serie dei dati

$${}_{2009}I_{2010} =$$

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.



Considerando i numeri indici a base mobile assegnati è possibile ricostruire la serie dei dati

$${}_{2009}I_{2010} = \frac{x_{2010}}{x_{2009}} = \frac{x_{2010}}{100} = 0.5 \rightarrow x_{2010} =$$

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.



Considerando i numeri indici a base mobile assegnati è possibile ricostruire la serie dei dati

$${}_{2009}I_{2010} = \frac{x_{2010}}{x_{2009}} = \frac{x_{2010}}{100} = 0.5 \rightarrow x_{2010} = 50$$

$${}_{2011}I_{2012} =$$

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.



Considerando i numeri indici a base mobile assegnati è possibile ricostruire la serie dei dati

$${}_{2009}I_{2010} = \frac{x_{2010}}{x_{2009}} = \frac{x_{2010}}{100} = 0.5 \rightarrow x_{2010} = 50$$

$${}_{2011}I_{2012} = \frac{x_{2012}}{x_{2011}} = \frac{x_{2012}}{100} = 2 \rightarrow x_{2012} =$$

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.



Considerando i numeri indici a base mobile assegnati è possibile ricostruire la serie dei dati

$${}_{2009}I_{2010} = \frac{x_{2010}}{x_{2009}} = \frac{x_{2010}}{100} = 0.5 \rightarrow x_{2010} = 50$$

$${}_{2011}I_{2012} = \frac{x_{2012}}{x_{2011}} = \frac{x_{2012}}{100} = 2 \rightarrow x_{2012} = 200$$

$${}_{2010}I_{2008} =$$

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.



Considerando i numeri indici a base mobile assegnati è possibile ricostruire la serie dei dati

$${}_{2009}I_{2010} = \frac{X_{2010}}{X_{2009}} = \frac{X_{2010}}{100} = 0.5 \rightarrow x_{2010} = 50$$

$${}_{2011}I_{2012} = \frac{X_{2012}}{X_{2011}} = \frac{X_{2012}}{100} = 2 \rightarrow x_{2012} = 200$$

$${}_{2010}I_{2008} = \frac{X_{2008}}{X_{2010}} = \frac{X_{2008}}{50} = 4 \rightarrow x_{2008} =$$

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.



Considerando i numeri indici a base mobile assegnati è possibile ricostruire la serie dei dati

$${}_{2009}I_{2010} = \frac{X_{2010}}{X_{2009}} = \frac{X_{2010}}{100} = 0.5 \rightarrow x_{2010} = 50$$

$${}_{2011}I_{2012} = \frac{X_{2012}}{X_{2011}} = \frac{X_{2012}}{100} = 2 \rightarrow x_{2012} = 200$$

$${}_{2010}I_{2008} = \frac{X_{2008}}{X_{2010}} = \frac{X_{2008}}{50} = 4 \rightarrow x_{2008} = 200$$

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.



Considerando la serie dei dati ricostruita abbiamo

$$2010/2009 =$$

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.



Considerando la serie dei dati ricostruita abbiamo

$${}_{2010}I_{2009} = \frac{x_{2009}}{x_{2010}} = \frac{100}{50} = 2$$

$${}_{2010}I_{2011} =$$

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.



Considerando la serie dei dati ricostruita abbiamo

$${}_{2010}I_{2009} = \frac{X_{2009}}{X_{2010}} = \frac{100}{50} = 2$$

$${}_{2010}I_{2011} = \frac{X_{2011}}{X_{2010}} = \frac{100}{50} = 2$$

$${}_{2010}I_{2012} =$$

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.



Considerando la serie dei dati ricostruita abbiamo

$${}_{2010}I_{2009} = \frac{X_{2009}}{X_{2010}} = \frac{100}{50} = 2$$

$${}_{2010}I_{2011} = \frac{X_{2011}}{X_{2010}} = \frac{100}{50} = 2$$

$${}_{2010}I_{2012} = \frac{X_{2012}}{X_{2010}} = \frac{200}{50} = 4$$

$${}_{2008}I_{2009} =$$

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.



Considerando la serie dei dati ricostruita abbiamo

$${}_{2010}I_{2009} = \frac{X_{2009}}{X_{2010}} = \frac{100}{50} = 2$$

$${}_{2010}I_{2011} = \frac{X_{2011}}{X_{2010}} = \frac{100}{50} = 2$$

$${}_{2010}I_{2012} = \frac{X_{2012}}{X_{2010}} = \frac{200}{50} = 4$$

$${}_{2008}I_{2009} = \frac{X_{2009}}{X_{2008}} = \frac{100}{200} = 0.5$$

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.



Si ottiene, quindi

t	x_t	$NIBF(t = 2010)$	$NIBM$
2008	200	4	—
2009	100	2	0.5
2010	50	1	0.5
2011	100	2	2
2012	200	4	2

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.



2. Si calcoli e si commenti il numero indice $_{2008}I_{2010}$.

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.



2. Si calcoli e si commenti il numero indice ${}_{2008}I_{2010}$.

$${}_{2008}I_{2010} = \frac{X_{2010}}{X_{2008}} = \frac{{}_{2010}I_{2010}}{{}_{2010}I_{2008}} = \frac{1}{4} = 0.25$$

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.



Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.

2. Si calcoli e si commenti il numero indice ${}_{2008}I_{2010}$.

$${}_{2008}I_{2010} = \frac{X_{2010}}{X_{2008}} = \frac{{}_{2010}I_{2010}}{{}_{2010}I_{2008}} = \frac{1}{4} = 0.25$$

Tra il tempo 2008 e il tempo 2010 si è registrato un decremento complessivo del 75% della grandezza oggetto di studio.



3. Si calcoli il tasso medio di variazione dal 2008 al 2012 e si commenti il risultato ottenuto.

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.



Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.

3. Si calcoli il tasso medio di variazione dal 2008 al 2012 e si commenti il risultato ottenuto.

La grandezza esaminata non ha registrato variazione tra l'anno 2008 e l'anno 2012.



Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.

Il seguente prospetto riporta i valori rilevati in corrispondenza a 5 unità statistiche con riferimento a due caratteri quantitativi X e Y

x_i	4	5	6	7	8
y_i	6	8	9	10	12

1. Si rappresentino graficamente le coppie dei dati ipotizzando Y in funzione di X .



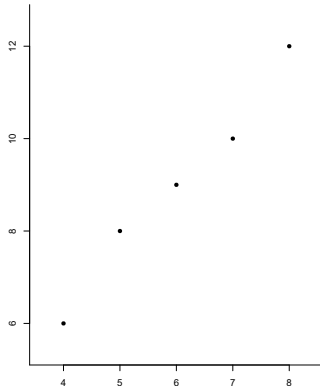
Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.

Le coppie di dati possono essere rappresentate mediante un grafico a dispersione.





2. Si calcolino, secondo il metodo dei minimi quadrati, i parametri e l'indice di adattamento dei seguenti modelli di regressione, individuando il modello migliore:

$$I) Y^* = a + b \cdot X \qquad II) Y^* = c$$

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.

Si riportano i calcoli necessari¹ per ottenere le stime dei parametri e gli indici di adattamento del modello

$$Y = a + bX + E$$

x_i	y_i	$x_i y_i$	x_i^2	y_i^2
4	6	24	16	36
5	8	40	25	64
6	9	54	36	81
7	10	70	49	100
8	12	96	64	144
30	45	284	190	425
6	9	56.8	38	85
$M(X)$	$M(Y)$	$M(XY)$	$M(X^2)$	$M(Y^2)$

$$\hat{b} =$$



Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.

¹Si osserva che per ragioni di stampa alcuni calcoli potrebbero essere duplicati con riferimento ai diversi modelli oggetto di studio

Si riportano i calcoli necessari¹ per ottenere le stime dei parametri e gli indici di adattamento del modello

$$Y = a + bX + E$$

x_i	y_i	$x_i y_i$	x_i^2	y_i^2
4	6	24	16	36
5	8	40	25	64
6	9	54	36	81
7	10	70	49	100
8	12	96	64	144
30	45	284	190	425
6	9	56.8	38	85
$M(X)$	$M(Y)$	$M(XY)$	$M(X^2)$	$M(Y^2)$

$$\begin{aligned}\hat{b} &= \frac{\text{Cov}(X, Y)}{\text{Var}(X)} = \frac{M(XY) - M(X)M(Y)}{M(X^2) - [M(X)]^2} = \\ &= \frac{56.8 - 6 \cdot 9}{38 - 6^2} = \frac{2.8}{2} = 1.4 \\ \hat{a} &= \end{aligned}$$

¹Si osserva che per ragioni di stampa alcuni calcoli potrebbero essere duplicati con riferimento ai diversi modelli oggetto di studio



Si riportano i calcoli necessari¹ per ottenere le stime dei parametri e gli indici di adattamento del modello

$$Y = a + bX + E$$

x_i	y_i	$x_i y_i$	x_i^2	y_i^2
4	6	24	16	36
5	8	40	25	64
6	9	54	36	81
7	10	70	49	100
8	12	96	64	144
30	45	284	190	425
6	9	56.8	38	85
$M(X)$	$M(Y)$	$M(XY)$	$M(X^2)$	$M(Y^2)$

$$\hat{b} = \frac{\text{Cov}(X, Y)}{\text{Var}(X)} = \frac{M(XY) - M(X)M(Y)}{M(X^2) - [M(X)]^2} =$$

$$= \frac{56.8 - 6 \cdot 9}{38 - 6^2} = \frac{2.8}{2} = 1.4$$

$$\hat{a} = M(Y) - \hat{b}M(X) = 0.6$$

$$\rho^2 =$$



Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.

¹Si osserva che per ragioni di stampa alcuni calcoli potrebbero essere duplicati con riferimento ai diversi modelli oggetto di studio

Si riportano i calcoli necessari¹ per ottenere le stime dei parametri e gli indici di adattamento del modello

$$Y = a + bX + E$$

x_i	y_i	$x_i y_i$	x_i^2	y_i^2
4	6	24	16	36
5	8	40	25	64
6	9	54	36	81
7	10	70	49	100
8	12	96	64	144
30	45	284	190	425
6	9	56.8	38	85
$M(X)$	$M(Y)$	$M(XY)$	$M(X^2)$	$M(Y^2)$

$$\hat{b} = \frac{\text{Cov}(X, Y)}{\text{Var}(X)} = \frac{M(XY) - M(X)M(Y)}{M(X^2) - [M(X)]^2} =$$

$$= \frac{56.8 - 6 \cdot 9}{38 - 6^2} = \frac{2.8}{2} = 1.4$$

$$\hat{a} = M(Y) - \hat{b}M(X) = 0.6$$

$$\rho^2 = \frac{[\text{Cov}(X, Y)]^2}{\text{Var}(X) \text{Var}(Y)} = \frac{2.8^2}{2 \cdot 4} = 0.98$$

$$M(\hat{E}^2) = \text{Var}(\hat{E}) = \text{varianza residua} =$$

¹Si osserva che per ragioni di stampa alcuni calcoli potrebbero essere duplicati con riferimento ai diversi modelli oggetto di studio



Si riportano i calcoli necessari¹ per ottenere le stime dei parametri e gli indici di adattamento del modello

$$Y = a + bX + E$$

x_i	y_i	$x_i y_i$	x_i^2	y_i^2
4	6	24	16	36
5	8	40	25	64
6	9	54	36	81
7	10	70	49	100
8	12	96	64	144
30	45	284	190	425
6	9	56.8	38	85
$M(X)$	$M(Y)$	$M(XY)$	$M(X^2)$	$M(Y^2)$

$$\begin{aligned}\hat{b} &= \frac{\text{Cov}(X, Y)}{\text{Var}(X)} = \frac{M(XY) - M(X)M(Y)}{M(X^2) - [M(X)]^2} = \\ &= \frac{56.8 - 6 \cdot 9}{38 - 6^2} = \frac{2.8}{2} = 1.4 \\ \hat{a} &= M(Y) - \hat{b}M(X) = 0.6 \\ \rho^2 &= \frac{[\text{Cov}(X, Y)]^2}{\text{Var}(X) \text{Var}(Y)} = \frac{2.8^2}{2 \cdot 4} = 0.98\end{aligned}$$

$$M(\hat{E}^2) = \text{Var}(\hat{E}) = \text{varianza residua} = \text{Var}(Y) \cdot (1 - \rho^2) = 0.08$$

¹Si osserva che per ragioni di stampa alcuni calcoli potrebbero essere duplicati con riferimento ai diversi modelli oggetto di studio





Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.

Si riportano i calcoli necessari² per ottenere le stime dei parametri e gli indici di adattamento del modello

$$Y = c + E$$

x_i	y_i	y_i^2
4	6	36
5	8	64
6	9	81
7	10	100
8	12	144
45		425
9		85
$M(Y)$		$M(Y^2)$

$$\hat{c} =$$

²Si osserva che per ragioni di stampa alcuni calcoli potrebbero essere duplicati con riferimento ai diversi modelli oggetto di studio



Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.

Si riportano i calcoli necessari² per ottenere le stime dei parametri e gli indici di adattamento del modello

$$Y = c + E$$

x_i	y_i	y_i^2
4	6	36
5	8	64
6	9	81
7	10	100
8	12	144
	45	425
	9	85
	$M(Y)$	$M(Y^2)$

$$\hat{c} = M(Y) = 9$$

$$R^2 =$$

²Si osserva che per ragioni di stampa alcuni calcoli potrebbero essere duplicati con riferimento ai diversi modelli oggetto di studio



Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.

Si riportano i calcoli necessari² per ottenere le stime dei parametri e gli indici di adattamento del modello

$$Y = c + E$$

x_i	y_i	y_i^2
4	6	36
5	8	64
6	9	81
7	10	100
8	12	144
45		425
9		85
$M(Y)$		$M(Y^2)$

$$\hat{c} = M(Y) = 9$$

$$R^2 = 0$$

$$M(\hat{E}^2) = \text{Var}(\hat{E}) = \text{varianza residua} =$$

²Si osserva che per ragioni di stampa alcuni calcoli potrebbero essere duplicati con riferimento ai diversi modelli oggetto di studio



Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.

Si riportano i calcoli necessari² per ottenere le stime dei parametri e gli indici di adattamento del modello

$$Y = c + E$$

x_i	y_i	y_i^2
4	6	36
5	8	64
6	9	81
7	10	100
8	12	144
		45
		425
		9
		85
		$M(Y)$
		$M(Y^2)$

$$\hat{c} = M(Y) = 9$$

$$R^2 = 0$$

$$M(\hat{E}^2) = \text{Var}(\hat{E}) = \text{varianza residua} = \text{Var}(Y) = 4$$

²Si osserva che per ragioni di stampa alcuni calcoli potrebbero essere duplicati con riferimento ai diversi modelli oggetto di studio



Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.

$\hat{a} = 0.6$, $\hat{b} = 1.4$, indice adattamento $\rho^2 = 0.98$ $\hat{c} = 9$, indice adattamento $R^2 = 0$



3. Si giustifichi se vale la relazione $\eta^2_{Y|X} < \eta^2_{X|Y}$.

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.



Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.

3. Si giustifichi se vale la relazione $\eta_{Y|X}^2 < \eta_{X|Y}^2$.

Nel caso in esame sussiste dipendenza funzionale bijectiva tra X e Y , quindi vale

$$\eta_{Y|X}^2 = \eta_{X|Y}^2 = 1.$$



Nell'urna U_1 vengono inserite 2 palline blu e 2 palline rosse e si esegue la seguente procedura:

- si estrae una prima pallina;
- la pallina estratta viene reinserita nell'urna, insieme a 4 ulteriori palline: 2 del suo stesso colore e 2 dell'altro colore;
- si estrae una seconda pallina dall'urna così modificata.

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.



1. Calcolare la probabilità che la prima pallina estratta sia blu.

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.



Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.

1. Calcolare la probabilità che la prima pallina estratta sia blu.

$$P(\text{prima pallina estratta blu}) = \frac{n. \text{ palline blu}}{n. \text{ palline}} = \frac{2}{4} = 0.5.$$



2. Usando la legge delle probabilità totali, calcolare la probabilità che la seconda pallina estratta sia blu.

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.



2. Usando la legge delle probabilità totali, calcolare la probabilità che la seconda pallina estratta sia blu.

Indicati con B_i e R_i gli eventi pallina blu alla i -esima estrazione e pallina rossa alla i -esima estrazione ($i = 1, 2$) abbiamo

$$P(B_2) = P(B_1) \cdot$$

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.



2. Usando la legge delle probabilità totali, calcolare la probabilità che la seconda pallina estratta sia blu.

Indicati con B_i e R_i gli eventi pallina blu alla i -esima estrazione e pallina rossa alla i -esima estrazione ($i = 1, 2$) abbiamo

$$P(B_2) = P(B_1) \cdot P(B_2|B_1) +$$

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.



2. Usando la legge delle probabilità totali, calcolare la probabilità che la seconda pallina estratta sia blu.

Indicati con B_i e R_i gli eventi pallina blu alla i -esima estrazione e pallina rossa alla i -esima estrazione ($i = 1, 2$) abbiamo

$$P(B_2) = P(B_1) \cdot P(B_2|B_1) + P(R_1) \cdot$$

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.



2. Usando la legge delle probabilità totali, calcolare la probabilità che la seconda pallina estratta sia blu.

Indicati con B_i e R_i gli eventi pallina blu alla i -esima estrazione e pallina rossa alla i -esima estrazione ($i = 1, 2$) abbiamo

$$P(B_2) = P(B_1) \cdot P(B_2|B_1) + P(R_1) \cdot P(B_2|R_1) =$$

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.



2. Usando la legge delle probabilità totali, calcolare la probabilità che la seconda pallina estratta sia blu.

Indicati con B_i e R_i gli eventi pallina blu alla i -esima estrazione e pallina rossa alla i -esima estrazione ($i = 1, 2$) abbiamo

$$\begin{aligned} P(B_2) &= P(B_1) \cdot P(B_2|B_1) + P(R_1) \cdot P(B_2|R_1) = \\ &= \frac{2}{4} . \end{aligned}$$

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.



2. Usando la legge delle probabilità totali, calcolare la probabilità che la seconda pallina estratta sia blu.

Indicati con B_i e R_i gli eventi pallina blu alla i -esima estrazione e pallina rossa alla i -esima estrazione ($i = 1, 2$) abbiamo

$$\begin{aligned} P(B_2) &= P(B_1) \cdot P(B_2|B_1) + P(R_1) \cdot P(B_2|R_1) = \\ &= \frac{2}{4} \cdot \frac{4}{8} + \end{aligned}$$

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.



2. Usando la legge delle probabilità totali, calcolare la probabilità che la seconda pallina estratta sia blu.

Indicati con B_i e R_i gli eventi pallina blu alla i -esima estrazione e pallina rossa alla i -esima estrazione ($i = 1, 2$) abbiamo

$$\begin{aligned} P(B_2) &= P(B_1) \cdot P(B_2|B_1) + P(R_1) \cdot P(B_2|R_1) = \\ &= \frac{2}{4} \cdot \frac{4}{8} + \frac{2}{4} \cdot \end{aligned}$$

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.



Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.

2. Usando la legge delle probabilità totali, calcolare la probabilità che la seconda pallina estratta sia blu.

Indicati con B_i e R_i gli eventi pallina blu alla i -esima estrazione e pallina rossa alla i -esima estrazione ($i = 1, 2$) abbiamo

$$\begin{aligned} P(B_2) &= P(B_1) \cdot P(B_2|B_1) + P(R_1) \cdot P(B_2|R_1) = \\ &= \frac{2}{4} \cdot \frac{4}{8} + \frac{2}{4} \cdot \frac{4}{8} = \end{aligned}$$



Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.

2. Usando la legge delle probabilità totali, calcolare la probabilità che la seconda pallina estratta sia blu.

Indicati con B_i e R_i gli eventi pallina blu alla i -esima estrazione e pallina rossa alla i -esima estrazione ($i = 1, 2$) abbiamo

$$\begin{aligned} P(B_2) &= P(B_1) \cdot P(B_2|B_1) + P(R_1) \cdot P(B_2|R_1) = \\ &= \frac{2}{4} \cdot \frac{4}{8} + \frac{2}{4} \cdot \frac{4}{8} = 0.25 + 0.25 = \end{aligned}$$



Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.

2. Usando la legge delle probabilità totali, calcolare la probabilità che la seconda pallina estratta sia blu.

Indicati con B_i e R_i gli eventi pallina blu alla i -esima estrazione e pallina rossa alla i -esima estrazione ($i = 1, 2$) abbiamo

$$\begin{aligned} P(B_2) &= P(B_1) \cdot P(B_2|B_1) + P(R_1) \cdot P(B_2|R_1) = \\ &= \frac{2}{4} \cdot \frac{4}{8} + \frac{2}{4} \cdot \frac{4}{8} = 0.25 + 0.25 = 0.5. \end{aligned}$$



3. Gli eventi $A_1 = \{\text{prima pallina estratta è blu}\}$ e $A_2 = \{\text{seconda pallina estratta è blu}\}$ sono indipendenti?

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.



3. Gli eventi $A_1 = \{\text{prima pallina estratta è blu}\}$ e $A_2 = \{\text{seconda pallina estratta è blu}\}$ sono indipendenti?

Gli eventi B_1 e B_2 sono indipendenti, abbiamo infatti

$$P(B_2) = 0.5 = P(B_2|B_1) = \frac{4}{8}$$

vale inoltre

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.



Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.

3. Gli eventi $A_1 = \{\text{prima pallina estratta è blu}\}$ e $A_2 = \{\text{seconda pallina estratta è blu}\}$ sono indipendenti?

Gli eventi B_1 e B_2 sono indipendenti, abbiamo infatti

$$P(B_2) = 0.5 = P(B_2|B_1) = \frac{4}{8}$$

vale inoltre

$$P(B_1 \cap B_2) = 0.25 = P(B_1) \cdot P(B_2) = 0.5 \cdot 0.5.$$

Tema T289

Prova scritta di STATISTICA del 23.06.2016

Prova scritta di
STATISTICA del
23.06.2016

Giuseppe Boari
Gabriele Cantaluppi



UNIVERSITÀ
CATTOLICA
del Sacro Cuore

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.

Giuseppe Boari
Gabriele Cantaluppi
Università Cattolica del Sacro Cuore, Milano



Data la seguente tabella a doppia entrata della distribuzione congiunta dei caratteri $X =$ 'tipo di contratto di lavoro' e $Y =$ 'salario lordo orario (in euro)' in una società industriale:

$X \setminus Y$	0 - 10	10 - 20	20 - 40
<i>dipendente</i>	1	2	0
<i>autonomo</i>	0	0	1

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.



1. Fornire la rappresentazione grafica delle distribuzioni marginali.

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.



Si considera la distribuzione marginale della variabile X , ottenuta sommando le frequenze di riga

x_i	$n_{i\bullet}$
<i>dipendente</i>	3
<i>autonomo</i>	1

Si tratta di un carattere qualitativo sconnesso che può essere rappresentato graficamente mediante un grafico a bastoncini o a canne d'organo.

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.

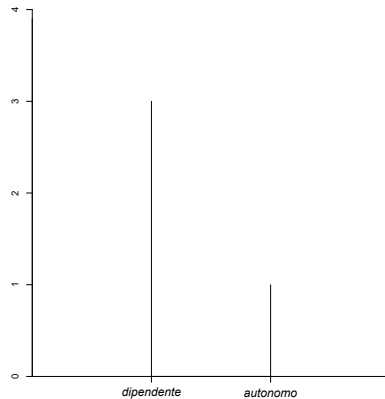


Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.





Si considera la distribuzione marginale della variabile Y , ottenuta sommando le frequenze di colonna

y_j	$n_{\bullet j}$
0 $\rightarrow$ 10	1
10 $\rightarrow$ 20	2
20 $\rightarrow$ 40	1

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.



Si tratta di un carattere quantitativo con modalità raggruppate in classi che può essere rappresentato graficamente mediante un istogramma.

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.



L'istogramma può essere costruito dopo avere calcolato le densità $d_{\bullet j} = \frac{n_{\bullet j}}{a_j}$, (dove a_j sono le ampiezze di classe).

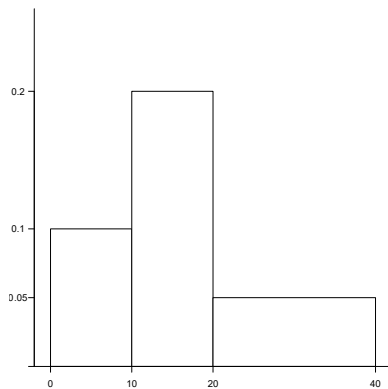
y_j	$n_{\bullet j}$	$a_{\bullet j}$	$d_{\bullet j}$
0 $\vdash$ 10	1	10 - 0 = 10	0.1
10 $\vdash$ 20	2	20 - 10 = 10	0.2
20 $\vdash$ 40	1	40 - 20 = 20	0.05

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.



Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.



Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.

2. Confrontare con un opportuno indice la variabilità delle distribuzioni condizionate ($Y|X$) del salario rispetto al tipo di contratto. Commentare i risultati.



Si considera la distribuzione condizionata $Y|x_1$

y_j	n_{1j}
0 $\rightarrow$ 10	1
10 $\rightarrow$ 20	2
20 $\rightarrow$ 40	0

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.



Si tratta di un carattere quantitativo con modalità raggruppate in classi.

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.



La media aritmetica

$$M(Y|x_1) = \frac{1}{n_{1\bullet}} \sum_{j=1}^k y_j n_{1j}$$

risulta 11.6667.

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.



La media aritmetica

$$M(Y|x_1) = \frac{1}{n_{1\bullet}} \sum_{j=1}^k y_j n_{1j}$$

risulta 11.6667.

Si ottiene (considerando come modalità i valori centrali delle classi)

y_j	n_{1j}	$y_j n_{1j}$
5	1	5
15	2	30
30	0	0
		3
		35

da cui segue

$$M(Y|x_1) = \frac{1}{n_{1\bullet}} \sum_{j=1}^k y_j n_{1j} = \frac{35}{3} = 11.6667.$$

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.



La varianza

$$\text{Var}(Y|x_1) = M(Y^2|x_1) - [M(Y|x_1)]^2$$

risulta 22.2222.

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.



Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.

La varianza

$$Var(Y|x_1) = M(Y^2|x_1) - [M(Y|x_1)]^2$$

risulta 22.2222. Si ottiene (considerando come modalità i valori centrali delle classi)

y_j	n_{1j}	$y_j n_{1j}$	$y_j^2 n_{1j}$
5	1	5	25
15	2	30	450
30	0	0	0
	3	35	475

da cui segue

$$Var(Y|x_1) = \frac{475}{3} - 11.6667^2 = 158.3333 - 136.1111 = 22.2222.$$



Per effettuare confronti in termini di variabilità tra caratteri quantitativi occorre utilizzare una misura relativa, ad esempio il coefficiente di variazione

$$CV = \frac{\sigma}{\mu} = \frac{\sqrt{\text{Var}(Y|x_1)}}{M(Y|x_1)} = \frac{\sqrt{22.2222}}{11.6667} = \frac{4.714}{11.6667} = 0.4041.$$

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.



Si considera la distribuzione condizionata $Y|x_2$

y_j	n_{2j}
0 $\rightarrow$ 10	0
10 $\rightarrow$ 20	0
20 $\rightarrow$ 40	1

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.



Si tratta di un carattere quantitativo con modalità raggruppate in classi.

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.



Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.

Si osserva che la distribuzione è degenere.

$M(Y|x_2) = 30$. La varianza e il coefficiente di variazione sono nulli.



Indice per confronto variabilità: Coefficiente di variazione

$$CV(Y|x_1) = 0.4041 \quad CV(Y|x_2) = 0$$

La distribuzione del salario dei lavoratori dipendenti presenta variabilità superiore a quella dei lavoratori autonomi

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.



3. Calcolare moda, media e mediana per la marginale Y .

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.



Si tratta di un carattere quantitativo con modalità raggruppate in classi.

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.



La moda corrisponde alla classe alla quale è associata la massima densità di frequenza

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.



La moda corrisponde alla classe alla quale è associata la massima densità di frequenza

y_j	$n_{\bullet j}$	$a_{\bullet j}$	$d_{\bullet j}$
0 $\vdash$ 10	1	$10 - 0 = 10$	0.1
10 $\vdash$ 20	2	$20 - 10 = 10$	0.2
20 $\vdash$ 40	1	$40 - 20 = 20$	0.05

e risulta 10 $\vdash$ 20.

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.



La mediana (percentile di ordine 0.50) può essere ottenuta considerando l'inversa della Funzione di Ripartizione in corrispondenza di 0.50.

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.



La mediana (percentile di ordine 0.50) può essere ottenuta considerando l'inversa della Funzione di Ripartizione in corrispondenza di 0.50.

y_j	$n_{\bullet j}$	$N_{\bullet j}$	$F_{\bullet j}$
0 $\vdash$ 10	1	1	0.25
10 $\vdash$ 20	2	3	0.75
20 $\vdash$ 40	1	4	1

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.

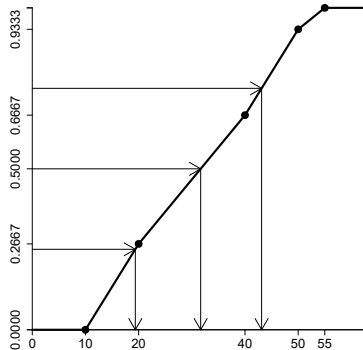


Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.



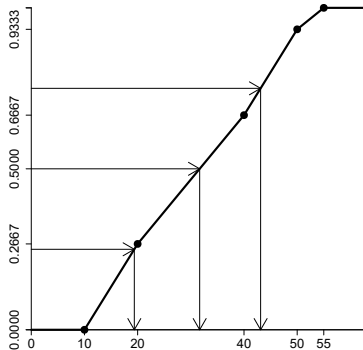


Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.



Si ottiene

$$y_{0.5} = 10 + \frac{0.5 - 0.25}{0.75 - 0.25} \cdot (20 - 10) = 15.$$



La media aritmetica

$$M(Y) = \frac{1}{n} \sum_{j=1}^k y_j n_{\bullet j}$$

risulta 16.25.

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.



La media aritmetica

$$M(Y) = \frac{1}{n} \sum_{j=1}^k y_j n_{\bullet j}$$

risulta 16.25.

Si ottiene (considerando come modalità i valori centrali delle classi)

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.



La media aritmetica

$$M(Y) = \frac{1}{n} \sum_{j=1}^k y_j n_{\bullet j}$$

risulta 16.25.

Si ottiene (considerando come modalità i valori centrali delle classi)

y_j	$n_{\bullet j}$	$y_j n_{\bullet j}$
5	1	5
15	2	30
30	1	30
		4
		65

da cui segue

$$M(Y) = \frac{1}{n} \sum_{i=1}^k y_j n_{\bullet j} = \frac{65}{4} = 16.25.$$

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.



Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.

4. Misurare la dipendenza tra le due variabili con un opportuno indice di connessione.

Sussiste dipendenza funzionale di X da Y , quindi $\chi_N = 1$.



La seguente tabella riporta la serie dei NIBM di un titolo azionario quotato in borsa:

t	$NIBM$
2010	-
2011	1.25
2012	0.64
2013	1.25
2014	0.75

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.



1. Si interpreti il significato di $_{2010}I_{2011}$, numero indice riferito al tempo 2011 con base 2010.

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.



Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.

1. Si interpreti il significato di $_{2010}I_{2011}$, numero indice riferito al tempo 2011 con base 2010.

$_{2010}I_{2011} = 1.25$. Tra il 2010 e il 2011 si è registrato un incremento del 25% del prezzo del titolo azionario.



2. Calcolare, mostrandone il procedimento, la serie dei NIBF con base 2010.

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.



2. Calcolare, mostrandone il procedimento, la serie dei NIBF con base 2010.
Si ricorda che ${}_{2010}I_{2010} = 1$ e che per valori di $t > 2010$ vale

$${}_{2010}I_t = {}_{2010}I_{2011} \cdot {}_{2011}I_{2012} \cdot \dots \cdot {}_{t-1}I_t.$$

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.



Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.

2. Calcolare, mostrandone il procedimento, la serie dei NIBF con base 2010.
Si ricorda che ${}_{2010}I_{2010} = 1$ e che per valori di $t > 2010$ vale

$${}_{2010}I_t = {}_{2010}I_{2011} \cdot {}_{2011}I_{2012} \cdot \dots \cdot {}_{t-1}I_t.$$

Si ottiene quindi

Anno	NIBM	NIBF (2010)
2010	1.00	1.00
2011	1.25	1.25
2012	0.64	0.80
2013	1.25	1.00
2014	0.75	0.75



3. Sapendo che il prezzo del titolo era pari a 200 nel 2013, calcolare il prezzo nel 2011.

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.



Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.

3. Sapendo che il prezzo del titolo era pari a 200 nel 2013, calcolare il prezzo nel 2011.

$$x_{2011} = 250$$

$$\text{Infatti } {}_{2011}I_{2013} = {}_{2011}I_{2012} \cdot {}_{2012}I_{2013} = 0.64 \cdot 1.25 = 0.80 = \frac{x_{2013}}{x_{2011}} = \frac{200}{x_{2011}}, \text{ quindi}$$

$$x_{2011} = \frac{200}{0.80} = 250$$



4. Calcolare il tasso medio di variazione annuo nell'intervallo 2011-2014, commentandone il risultato.

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.



4. Calcolare il tasso medio di variazione annuo nell'intervallo 2011-2014, commentandone il risultato.

Si ricorda che

$${}_{2011}I_{2014} = {}_{2011}I_{2012} \cdot {}_{2012}I_{2013} \cdot {}_{2013}I_{2014}.$$

Si deve trovare il numero indice a base mobile medio.

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.



Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.

4. Calcolare il tasso medio di variazione annuo nell'intervallo 2011-2014, commentandone il risultato.

Si ricorda che

$${}_{2011}I_{2014} = {}_{2011}I_{2012} \cdot {}_{2012}I_{2013} \cdot {}_{2013}I_{2014}.$$

Si deve trovare il numero indice a base mobile medio. Applicando il criterio di scelta della media secondo Chisini si ottiene

$${}_{2011}I_{2014} = {}_{2011}I_{2012} \cdot {}_{2012}I_{2013} \cdot {}_{2013}I_{2014} = \alpha \cdot \alpha \cdot \alpha = \alpha^3.$$



Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.

4. Calcolare il tasso medio di variazione annuo nell'intervallo 2011-2014, commentandone il risultato.

Si ricorda che

$${}_{2011}l_{2014} = {}_{2011}l_{2012} \cdot {}_{2012}l_{2013} \cdot {}_{2013}l_{2014}.$$

Si deve trovare il numero indice a base mobile medio. Applicando il criterio di scelta della media secondo Chisini si ottiene

$${}_{2011}l_{2014} = {}_{2011}l_{2012} \cdot {}_{2012}l_{2013} \cdot {}_{2013}l_{2014} = \alpha \cdot \alpha \cdot \alpha = \alpha^3.$$

Quindi

$$\alpha = \sqrt[3]{{}_{2011}l_{2014}} = \sqrt[3]{{}_{2011}l_{2012} \cdot {}_{2012}l_{2013} \cdot {}_{2013}l_{2014}} = 0.8434.$$

Da cui segue una variazione relativa media pari a -0.1566 , corrispondente a un decremento medio annuo tra il 2011 e il 2014 del 15.66%.



Con riferimento allo studio della dipendenza tra due caratteri quantitativi X e Y , si sono raccolte le seguenti 5 coppie di osservazioni (x_i, y_i) $i = 1, 2, \dots, 5$

x_i	-7	-7	-6	-5	-5
y_i	2	4	4	4	6

1. Fornire la rappresentazione grafica delle coppie di punti, insieme alla funzione di regressione di Y dato X .

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.



La funzione di regressione associa a ogni valore x appartenente all'insieme di definizione della variabile X la media di Y condizionata a x , $M(Y|x)$

x	$M(Y x)$
-7	$\frac{1}{2}(2 + 4) = 3$
-6	4
-5	$\frac{1}{2}(4 + 6) = 5$

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.



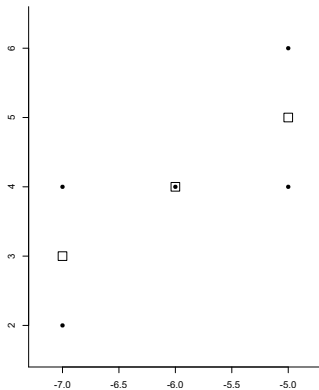
Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.

Le coppie di dati possono essere rappresentate mediante un grafico a dispersione.





2. Si calcolino, secondo il metodo dei minimi quadrati, i parametri dei seguenti modelli di regressione:

$$I) Y^* = a + b \cdot X \qquad II) Y^* = c \cdot (X - 3)$$

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.

Si riportano i calcoli necessari¹ per ottenere le stime dei parametri e gli indici di adattamento del modello

$$Y = a + bX + E$$

x_i	y_i	$x_i y_i$	x_i^2	y_i^2
-7	2	-14	49	4
-7	4	-28	49	16
-6	4	-24	36	16
-5	4	-20	25	16
-5	6	-30	25	36
-30	20	-116	184	88
-6	4	-23.2	36.8	17.6
$M(X)$	$M(Y)$	$M(XY)$	$M(X^2)$	$M(Y^2)$

$$\hat{b} =$$



Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.

¹Si osserva che per ragioni di stampa alcuni calcoli potrebbero essere duplicati con riferimento ai diversi modelli oggetto di studio

Si riportano i calcoli necessari¹ per ottenere le stime dei parametri e gli indici di adattamento del modello

$$Y = a + bX + E$$

x_i	y_i	$x_i y_i$	x_i^2	y_i^2
-7	2	-14	49	4
-7	4	-28	49	16
-6	4	-24	36	16
-5	4	-20	25	16
-5	6	-30	25	36
-30	20	-116	184	88
-6	4	-23.2	36.8	17.6
$M(X)$	$M(Y)$	$M(XY)$	$M(X^2)$	$M(Y^2)$

$$\begin{aligned}\hat{b} &= \frac{\text{Cov}(X, Y)}{\text{Var}(X)} = \frac{M(XY) - M(X)M(Y)}{M(X^2) - [M(X)]^2} = \\ &= \frac{-23.2 - (-6) \cdot 4}{36.8 - (-6)^2} = \frac{0.8}{0.8} = 1 \\ \hat{a} &= \end{aligned}$$



Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.

¹Si osserva che per ragioni di stampa alcuni calcoli potrebbero essere duplicati con riferimento ai diversi modelli oggetto di studio

Si riportano i calcoli necessari¹ per ottenere le stime dei parametri e gli indici di adattamento del modello

$$Y = a + bX + E$$

x_i	y_i	$x_i y_i$	x_i^2	y_i^2
-7	2	-14	49	4
-7	4	-28	49	16
-6	4	-24	36	16
-5	4	-20	25	16
-5	6	-30	25	36
-30	20	-116	184	88
-6	4	-23.2	36.8	17.6
$M(X)$	$M(Y)$	$M(XY)$	$M(X^2)$	$M(Y^2)$

$$\hat{b} = \frac{\text{Cov}(X, Y)}{\text{Var}(X)} = \frac{M(XY) - M(X)M(Y)}{M(X^2) - [M(X)]^2} =$$

$$= \frac{-23.2 - (-6) \cdot 4}{36.8 - (-6)^2} = \frac{0.8}{0.8} = 1$$

$$\hat{a} = M(Y) - \hat{b}M(X) = 10$$

$$\rho^2 =$$



Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.

¹Si osserva che per ragioni di stampa alcuni calcoli potrebbero essere duplicati con riferimento ai diversi modelli oggetto di studio

Si riportano i calcoli necessari¹ per ottenere le stime dei parametri e gli indici di adattamento del modello

$$Y = a + bX + E$$

x_i	y_i	$x_i y_i$	x_i^2	y_i^2
-7	2	-14	49	4
-7	4	-28	49	16
-6	4	-24	36	16
-5	4	-20	25	16
-5	6	-30	25	36
-30	20	-116	184	88
-6	4	-23.2	36.8	17.6
$M(X)$	$M(Y)$	$M(XY)$	$M(X^2)$	$M(Y^2)$

$$\hat{b} = \frac{\text{Cov}(X, Y)}{\text{Var}(X)} = \frac{M(XY) - M(X)M(Y)}{M(X^2) - [M(X)]^2} =$$

$$= \frac{-23.2 - (-6) \cdot 4}{36.8 - (-6)^2} = \frac{0.8}{0.8} = 1$$

$$\hat{a} = M(Y) - \hat{b}M(X) = 10$$

$$\rho^2 = \frac{[\text{Cov}(X, Y)]^2}{\text{Var}(X) \text{Var}(Y)} = \frac{0.8^2}{0.8 \cdot 1.6} = 0.5$$

$$M(\hat{E}^2) = \text{Var}(\hat{E}) = \text{varianza residua} =$$

¹Si osserva che per ragioni di stampa alcuni calcoli potrebbero essere duplicati con riferimento ai diversi modelli oggetto di studio



Si riportano i calcoli necessari¹ per ottenere le stime dei parametri e gli indici di adattamento del modello

$$Y = a + bX + E$$

x_i	y_i	$x_i y_i$	x_i^2	y_i^2
-7	2	-14	49	4
-7	4	-28	49	16
-6	4	-24	36	16
-5	4	-20	25	16
-5	6	-30	25	36
-30	20	-116	184	88
-6	4	-23.2	36.8	17.6
$M(X)$	$M(Y)$	$M(XY)$	$M(X^2)$	$M(Y^2)$

$$\hat{b} = \frac{\text{Cov}(X, Y)}{\text{Var}(X)} = \frac{M(XY) - M(X)M(Y)}{M(X^2) - [M(X)]^2} =$$

$$= \frac{-23.2 - (-6) \cdot 4}{36.8 - (-6)^2} = \frac{0.8}{0.8} = 1$$

$$\hat{a} = M(Y) - \hat{b}M(X) = 10$$

$$\rho^2 = \frac{[\text{Cov}(X, Y)]^2}{\text{Var}(X) \text{Var}(Y)} = \frac{0.8^2}{0.8 \cdot 1.6} = 0.5$$

$$M(\hat{E}^2) = \text{Var}(\hat{E}) = \text{varianza residua} = \text{Var}(Y) \cdot (1 - \rho^2) = 0.8$$

¹Si osserva che per ragioni di stampa alcuni calcoli potrebbero essere duplicati con riferimento ai diversi modelli oggetto di studio





Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.

Si riportano i calcoli necessari² per ottenere le stime dei parametri e la misura di adattamento del modello

$$Y = c(X - 3) + E$$

$(x - 3)_i$	y_i	$(x - 3)_i y_i$	$(x - 3)_i^2$	y_i^2
-10	2	-20	100	4
-10	4	-40	100	16
-9	4	-36	81	16
-8	4	-32	64	16
-8	6	-48	64	36
		-176	409	88
		-35.2	81.8	17.6
		$M((X - 3)Y)$	$M((X - 3)^2)$	$M(Y^2)$

$$\hat{c} =$$

²Si osserva che per ragioni di stampa alcuni calcoli potrebbero essere duplicati con riferimento ai diversi modelli oggetto di studio



Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.

Si riportano i calcoli necessari² per ottenere le stime dei parametri e la misura di adattamento del modello

$$Y = c(X - 3) + E$$

$(x - 3)_i$	y_i	$(x - 3)_i y_i$	$(x - 3)_i^2$	y_i^2
-10	2	-20	100	4
-10	4	-40	100	16
-9	4	-36	81	16
-8	4	-32	64	16
-8	6	-48	64	36
		-176	409	88
		-35.2	81.8	17.6
		$M((X - 3)Y)$	$M((X - 3)^2)$	$M(Y^2)$

$$\hat{c} = \frac{M((X - 3)Y)}{M((X - 3)^2)} = \frac{-35.2}{81.8} = -0.4303$$

$$M(\hat{E}^2) =$$

²Si osserva che per ragioni di stampa alcuni calcoli potrebbero essere duplicati con riferimento ai diversi modelli oggetto di studio



Si riportano i calcoli necessari² per ottenere le stime dei parametri e la misura di adattamento del modello

$$Y = c(X - 3) + E$$

$(x - 3)_i$	y_i	$(x - 3)_i y_i$	$(x - 3)_i^2$	y_i^2
-10	2	-20	100	4
-10	4	-40	100	16
-9	4	-36	81	16
-8	4	-32	64	16
-8	6	-48	64	36
		-176	409	88
		-35.2	81.8	17.6
		$M((X - 3)Y)$	$M((X - 3)^2)$	$M(Y^2)$

$$\hat{c} = \frac{M((X - 3)Y)}{M((X - 3)^2)} = \frac{-35.2}{81.8} = -0.4303$$

$$M(\hat{E}^2) = M[Y \cdot (Y - \hat{c}(X - 3))] = M(Y^2) - \hat{c}M((X - 3)Y) = 17.6 - (-0.4303 \cdot -35.2) = 2.4528$$

²Si osserva che per ragioni di stampa alcuni calcoli potrebbero essere duplicati con riferimento ai diversi modelli oggetto di studio



Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.

$\hat{a} = 10$, $\hat{b} = 1$, indice adattamento $M(\hat{E}^2) = 0.8$ $\hat{c} = -0.4303$, indice adattamento
 $M(\hat{E}^2) = 2.4528$



3. Confrontare gli indici di adattamento, commentando i risultati.

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.



3. Confrontare gli indici di adattamento, commentando i risultati.

Il modello retta presenta sicuramente un adattamento superiore al secondo modello in quanto coincide con il modello funzione di regressione, essendo le medie condizionate allineate.

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.



4. Senza svolgere i conti, cosa è possibile dire sul valore di $\eta^2_{Y|X}$.

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.



4. Senza svolgere i conti, cosa è possibile dire sul valore di $\eta_{Y|X}^2$.
 $\eta_{Y|X}^2 = \rho^2 = 0.5$ in quanto le medie condizionate sono allineate.

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.



5. Sulla base del miglior modello si preveda il valore di Y in corrispondenza di $X = 10$.

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.



Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.

5. Sulla base del miglior modello si preveda il valore di Y in corrispondenza di $X = 10$.

$$\hat{Y}(10) = \hat{a} + \hat{b} \cdot 10 = 10 + 1 \cdot 10 = 20.$$

Esercizio 4

Prova scritta di
STATISTICA del
23.06.2016

Giuseppe Boari
Gabriele Cantaluppi



UNIVERSITÀ
CATTOLICA
del Sacro Cuore

La durata (in mesi) delle batterie prodotte da un'azienda segue una distribuzione normale di media 21 e scarto quadratico medio 9. Sono considerate difettose le batterie che durano meno di 17 mesi.

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.



1. Qual è la probabilità che una batteria estratta a caso dal flusso produttivo sia difettosa?

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.



1. Qual è la probabilità che una batteria estratta a caso dal flusso produttivo sia difettosa?

Indicata con X la variabile casuale 'durata delle batterie' abbiamo $X \sim$

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.



1. Qual è la probabilità che una batteria estratta a caso dal flusso produttivo sia difettosa?

Indicata con X la variabile casuale 'durata delle batterie' abbiamo $X \sim N(\mu = 21, \sigma^2 = 81)$.

$$P(X \leq 17) =$$

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.



Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.

1. Qual è la probabilità che una batteria estratta a caso dal flusso produttivo sia difettosa?

Indicata con X la variabile casuale 'durata delle batterie' abbiamo $X \sim N(\mu = 21, \sigma^2 = 81)$.

$$P(X \leq 17) = P\left(\frac{X - \mu}{\sigma} \leq \frac{17 - \mu}{\sigma}\right)$$
$$=$$



Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.

1. Qual è la probabilità che una batteria estratta a caso dal flusso produttivo sia difettosa?

Indicata con X la variabile casuale 'durata delle batterie' abbiamo $X \sim N(\mu = 21, \sigma^2 = 81)$.

$$\begin{aligned} P(X \leq 17) &= P\left(\frac{X - \mu}{\sigma} \leq \frac{17 - \mu}{\sigma}\right) \\ &= P\left(Z \leq \frac{17 - 21}{9}\right) \\ &= \end{aligned}$$



Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.

1. Qual è la probabilità che una batteria estratta a caso dal flusso produttivo sia difettosa?

Indicata con X la variabile casuale 'durata delle batterie' abbiamo $X \sim N(\mu = 21, \sigma^2 = 81)$.

$$\begin{aligned} P(X \leq 17) &= P\left(\frac{X - \mu}{\sigma} \leq \frac{17 - \mu}{\sigma}\right) \\ &= P\left(Z \leq \frac{17 - 21}{9}\right) \\ &= P(Z \leq -0.44) = P(Z \geq 0.44) \\ &= \end{aligned}$$



Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.

1. Qual è la probabilità che una batteria estratta a caso dal flusso produttivo sia difettosa?

Indicata con X la variabile casuale 'durata delle batterie' abbiamo $X \sim N(\mu = 21, \sigma^2 = 81)$.

$$\begin{aligned} P(X \leq 17) &= P\left(\frac{X - \mu}{\sigma} \leq \frac{17 - \mu}{\sigma}\right) \\ &= P\left(Z \leq \frac{17 - 21}{9}\right) \\ &= P(Z \leq -0.44) = P(Z \geq 0.44) \\ &= 1 - P(Z \leq 0.44) = 1 - 0.67 = \end{aligned}$$



1. Qual è la probabilità che una batteria estratta a caso dal flusso produttivo sia difettosa?

Indicata con X la variabile casuale 'durata delle batterie' abbiamo $X \sim N(\mu = 21, \sigma^2 = 81)$.

$$\begin{aligned} P(X \leq 17) &= P\left(\frac{X - \mu}{\sigma} \leq \frac{17 - \mu}{\sigma}\right) \\ &= P\left(Z \leq \frac{17 - 21}{9}\right) \\ &= P(Z \leq -0.44) = P(Z \geq 0.44) \\ &= 1 - P(Z \leq 0.44) = 1 - 0.67 = 0.33. \end{aligned}$$



2. Estraendo con reimmissione un campione di 10 batterie, quale è la probabilità che meno di due siano difettose?

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.



Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.

2. Estraendo con reimmissione un campione di 10 batterie, quale è la probabilità che meno di due siano difettose?

Sia X la variabile casuale 'n. di difettose nelle 10 prove indipendenti'.

$$X \sim$$



Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.

2. Estraendo con reimmissione un campione di 10 batterie, quale è la probabilità che meno di due siano difettose?

Sia X la variabile casuale 'n. di difettose nelle 10 prove indipendenti'.

$$X \sim \text{Bin}(n = 10, p = 0.33).$$

$$P(X < 2) =$$



Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.

2. Estraendo con reimmissione un campione di 10 batterie, quale è la probabilità che meno di due siano difettose?

Sia X la variabile casuale 'n. di difettose nelle 10 prove indipendenti'.

$$X \sim \text{Bin}(n = 10, p = 0.33).$$

$$P(X < 2) = P(X = 0) + P(X = 1)$$

=



Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.

2. Estraendo con reimmissione un campione di 10 batterie, quale è la probabilità che meno di due siano difettose?

Sia X la variabile casuale 'n. di difettose nelle 10 prove indipendenti'.

$$X \sim \text{Bin}(n = 10, p = 0.33).$$

$$\begin{aligned} P(X < 2) &= P(X = 0) + P(X = 1) \\ &= \binom{10}{0} 0.33^0 0.67^{10} + \binom{10}{1} 0.33^1 0.67^9 \\ &= \end{aligned}$$



Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.

2. Estraendo con reimmissione un campione di 10 batterie, quale è la probabilità che meno di due siano difettose?

Sia X la variabile casuale 'n. di difettose nelle 10 prove indipendenti'.

$$X \sim \text{Bin}(n = 10, p = 0.33).$$

$$\begin{aligned} P(X < 2) &= P(X = 0) + P(X = 1) \\ &= \binom{10}{0} 0.33^0 0.67^{10} + \binom{10}{1} 0.33^1 0.67^9 \\ &= 0.01823 + 0.08978 = \end{aligned}$$



Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.

2. Estraendo con reimmissione un campione di 10 batterie, quale è la probabilità che meno di due siano difettose?

Sia X la variabile casuale 'n. di difettose nelle 10 prove indipendenti'.

$$X \sim \text{Bin}(n = 10, p = 0.33).$$

$$\begin{aligned} P(X < 2) &= P(X = 0) + P(X = 1) \\ &= \binom{10}{0} 0.33^0 0.67^{10} + \binom{10}{1} 0.33^1 0.67^9 \\ &= 0.01823 + 0.08978 = 0.10801. \end{aligned}$$

3. Considerato un campione di 4400 batterie estratte con reimmissione, indicare qual è la probabilità che non più di 400 siano difettose.



Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.



Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.

3. Considerato un campione di 4400 batterie estratte con reimmissione, indicare qual è la probabilità che non più di 400 siano difettose.

Indicata con X la variabile casuale 'numero di difettose' nelle 4400 prove indipendenti risulta

$$X \sim$$



3. Considerato un campione di 4400 batterie estratte con reimmissione, indicare qual è la probabilità che non più di 400 siano difettose.

Indicata con X la variabile casuale 'numero di difettose' nelle 4400 prove indipendenti risulta

$$X \sim \text{Bin}(n = 4400, p = 0.33).$$

$$M(X) = 1452; \text{Var}(X) = 972.84.$$

Dal momento che $np = 1452 > 5$ ed $n(1 - p) = 2948 > 5$ è possibile utilizzare l'approssimazione della variabile casuale Binomiale con la variabile casuale Normale

$$X \sim$$



3. Considerato un campione di 4400 batterie estratte con reimmissione, indicare qual è la probabilità che non più di 400 siano difettose.

Indicata con X la variabile casuale 'numero di difettose' nelle 4400 prove indipendenti risulta

$$X \sim \text{Bin}(n = 4400, p = 0.33).$$

$$M(X) = 1452; \text{Var}(X) = 972.84.$$

Dal momento che $np = 1452 > 5$ ed $n(1 - p) = 2948 > 5$ è possibile utilizzare l'approssimazione della variabile casuale Binomiale con la variabile casuale Normale

$$X \sim N(\mu = 1452, \sigma^2 = 972.84).$$

Si ottiene

$$P(X \leq 400) =$$



3. Considerato un campione di 4400 batterie estratte con reimmissione, indicare qual è la probabilità che non più di 400 siano difettose.

Indicata con X la variabile casuale 'numero di difettose' nelle 4400 prove indipendenti risulta

$$X \sim \text{Bin}(n = 4400, p = 0.33).$$

$$M(X) = 1452; \text{Var}(X) = 972.84.$$

Dal momento che $np = 1452 > 5$ ed $n(1 - p) = 2948 > 5$ è possibile utilizzare l'approssimazione della variabile casuale Binomiale con la variabile casuale Normale

$$X \sim N(\mu = 1452, \sigma^2 = 972.84).$$

Si ottiene

$$P(X \leq 400) = P\left(\frac{X - \mu}{\sigma} \leq \frac{400 - \mu}{\sigma}\right)$$
$$=$$



3. Considerato un campione di 4400 batterie estratte con reimmissione, indicare qual è la probabilità che non più di 400 siano difettose.

Indicata con X la variabile casuale 'numero di difettose' nelle 4400 prove indipendenti risulta

$$X \sim \text{Bin}(n = 4400, p = 0.33).$$

$$M(X) = 1452; \text{Var}(X) = 972.84.$$

Dal momento che $np = 1452 > 5$ ed $n(1 - p) = 2948 > 5$ è possibile utilizzare l'approssimazione della variabile casuale Binomiale con la variabile casuale Normale

$$X \sim N(\mu = 1452, \sigma^2 = 972.84).$$

Si ottiene

$$\begin{aligned} P(X \leq 400) &= P\left(\frac{X - \mu}{\sigma} \leq \frac{400 - \mu}{\sigma}\right) \\ &= P\left(Z \leq \frac{400 - 1452}{31.1904}\right) \\ &= \end{aligned}$$



3. Considerato un campione di 4400 batterie estratte con reimmissione, indicare qual è la probabilità che non più di 400 siano difettose.

Indicata con X la variabile casuale 'numero di difettose' nelle 4400 prove indipendenti risulta

$$X \sim \text{Bin}(n = 4400, p = 0.33).$$

$$M(X) = 1452; \text{Var}(X) = 972.84.$$

Dal momento che $np = 1452 > 5$ ed $n(1 - p) = 2948 > 5$ è possibile utilizzare l'approssimazione della variabile casuale Binomiale con la variabile casuale Normale

$$X \sim N(\mu = 1452, \sigma^2 = 972.84).$$

Si ottiene

$$\begin{aligned} P(X \leq 400) &= P\left(\frac{X - \mu}{\sigma} \leq \frac{400 - \mu}{\sigma}\right) \\ &= P\left(Z \leq \frac{400 - 1452}{31.1904}\right) \\ &= P(Z \leq -33.73) = 0. \end{aligned}$$

Tema T289-2

Prova scritta di STATISTICA del 14.07.2016

Prova scritta di
STATISTICA del
14.07.2016

Giuseppe Boari
Gabriele Cantaluppi



UNIVERSITÀ
CATTOLICA
del Sacro Cuore

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.

Giuseppe Boari
Gabriele Cantaluppi
Università Cattolica del Sacro Cuore, Milano

Esercizio 1

Prova scritta di
STATISTICA del
14.07.2016

Giuseppe Boari
Gabriele Cantaluppi



UNIVERSITÀ
CATTOLICA
del Sacro Cuore

La seguente tabella riporta la distribuzione congiunta dei caratteri quantitativi X e Y .

$X \backslash Y$	$0 \div 2$	$2 \div 4$	$4 \div 6$	
2	1	2	1	
4	1	0	1	

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.



Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.

2. Si rappresentino graficamente le distribuzioni condizionate $X|y_1$ e $Y|x_1$.



Si considera la distribuzione condizionata $X|y_1$

x_i	n_{i1}
2	1
4	1

Esercizio 1.

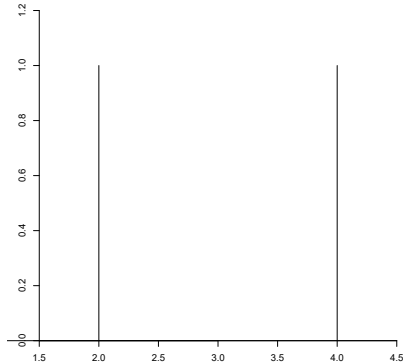
Esercizio 2.

Esercizio 3.

Esercizio 4.



Si tratta di un carattere quantitativo con modalità non raggruppate in classi che può essere rappresentato graficamente mediante un grafico a bastoncini.



Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.



Si considera la distribuzione condizionata $Y|x_1$

y_j	n_{1j}
0 + 2	1
2 + 4	2
4 + 6	1

Esercizio 1.

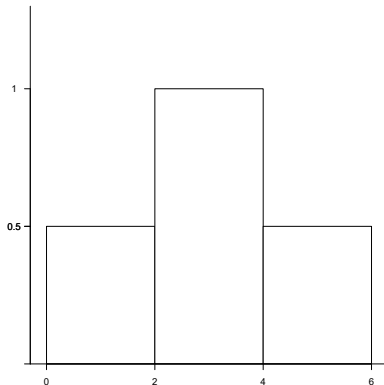
Esercizio 2.

Esercizio 3.

Esercizio 4.

Si tratta di un carattere quantitativo con modalità raggruppate in classi che può essere rappresentato graficamente mediante un istogramma. L'istogramma può essere costruito dopo avere calcolato le densità $d_{1j} = \frac{n_{1j}}{a_j}$, (dove a_j sono le ampiezze di classe).

y_j	n_{1j}	a_{1j}	d_{1j}
0 $\vdash$ 2	1	$2 - 0 = 2$	0.5
2 $\vdash$ 4	2	$4 - 2 = 2$	1
4 $\vdash$ 6	1	$6 - 4 = 2$	0.5



Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.



3. Si riassumano le distribuzioni condizionate $Y|x_1$ e $Y|x_2$ con opportuni indici di posizione e variabilità.

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.



Si considera la distribuzione condizionata $Y|x_1$

y_j	n_{1j}
0 + 2	1
2 + 4	2
4 + 6	1

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.



Si tratta di un carattere quantitativo con modalità raggruppate in classi.

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.



Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.

Si osserva che la distribuzione è simmetrica rispetto a 3; la mediana (percentile di ordine 0.50) risulta $y_{0.5} = 3$.



La media aritmetica

$$M(Y|x_1) = \frac{1}{n_{1\bullet}} \sum_{j=1}^k y_j n_{1j}$$

risulta 3.

Si osserva, infatti, che la distribuzione è simmetrica.

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.



La varianza
risulta 2.

$$\text{Var}(Y|x_1) = M(Y^2|x_1) - [M(Y|x_1)]^2$$

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.



La varianza

$$\text{Var}(Y|x_1) = M(Y^2|x_1) - [M(Y|x_1)]^2$$

risulta 2. Si ottiene (considerando come modalità i valori centrali delle classi)

y_j	n_{1j}	$y_j n_{1j}$	$y_j^2 n_{1j}$
1	1	1	1
3	2	6	18
5	1	5	25
	4	12	44

da cui segue

$$\text{Var}(Y|x_1) = \frac{44}{4} - 3^2 = 11 - 9 = 2.$$

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.



Per effettuare confronti in termini di variabilità tra caratteri quantitativi occorre utilizzare una misura relativa, ad esempio il coefficiente di variazione

$$CV = \frac{\sigma}{\mu} = \frac{\sqrt{\text{Var}(Y|x_1)}}{M(Y|x_1)} = \frac{\sqrt{2}}{3} = \frac{1.4142}{3} = 0.4714.$$

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.



Si considera la distribuzione condizionata $Y|x_2$

y_j	n_{2j}
0 + 2	1
2 + 4	0
4 + 6	1

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.



Si tratta di un carattere quantitativo con modalità raggruppate in classi.

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.



Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.

Si osserva che la distribuzione è simmetrica rispetto a 3; la mediana (percentile di ordine 0.50) risulta $y_{0.5} = 3$.



La media aritmetica

$$M(Y|x_2) = \frac{1}{n_{2\bullet}} \sum_{j=1}^k y_j n_{2j}$$

risulta 3.

Si osserva, infatti, che la distribuzione è simmetrica.

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.



La varianza
risulta 4.

$$\text{Var}(Y|x_2) = M(Y^2|x_2) - [M(Y|x_2)]^2$$

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.



La varianza

$$\text{Var}(Y|x_2) = M(Y^2|x_2) - [M(Y|x_2)]^2$$

risulta 4. Si ottiene (considerando come modalità i valori centrali delle classi)

y_j	n_{2j}	$y_j n_{2j}$	$y_j^2 n_{2j}$
1	1	1	1
3	0	0	0
5	1	5	25
2		6	26

da cui segue

$$\text{Var}(Y|x_2) = \frac{26}{2} - 3^2 = 13 - 9 = 4.$$

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.



Per effettuare confronti in termini di variabilità tra caratteri quantitativi occorre utilizzare una misura relativa, ad esempio il coefficiente di variazione

$$CV = \frac{\sigma}{\mu} = \frac{\sqrt{\text{Var}(Y|x_2)}}{M(Y|x_2)} = \frac{\sqrt{4}}{3} = \frac{2}{3} = 0.6667.$$

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.



Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.

$Y|_{x_1}$ Indice di posizione: $M(Y|_{x_1}) = 3$ Indice di variabilità: $CV(Y|_{x_1}) = 0.4714$

$Y|_{x_2}$ Indice di posizione: $M(Y|_{x_2}) = 3$ Indice di variabilità: $CV(Y|_{x_2}) = 0.6667$

Le distribuzioni condizionate $Y|_{x_1}$ e $Y|_{x_2}$ sono caratterizzate dallo stesso livello medio; $Y|_{x_2}$ presenta variabilità superiore a $Y|_{x_1}$.



3. Si determini il grado di connessione tra le variabili X e Y mediante il calcolo di un opportuno indice.

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.



Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.

Dalla distribuzione delle frequenze congiunte n_{ij} :

	0 + 2	2 + 4	4 + 6	$n_{i\bullet}$
2	1	2	1	4
4	1	0	1	2
$n_{\bullet j}$	2	2	2	6

è possibile ottenere le frequenze teoriche nella situazione di indipendenza stocastica

$$\hat{n}_{ij} = \frac{n_{i\bullet} n_{\bullet j}}{n}:$$



Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.

Dalla distribuzione delle frequenze congiunte n_{ij} :

	0 ÷ 2	2 ÷ 4	4 ÷ 6	$n_{i\bullet}$
2	1	2	1	4
4	1	0	1	2
$n_{\bullet j}$	2	2	2	6

è possibile ottenere le frequenze teoriche nella situazione di indipendenza stocastica

$$\hat{n}_{ij} = \frac{n_{i\bullet} \cdot n_{\bullet j}}{n}:$$

	0 ÷ 2	2 ÷ 4	4 ÷ 6	$n_{i\bullet}$
2	1.3333	1.3333	1.3333	4
4	0.6667	0.6667	0.6667	2
$n_{\bullet j}$	2.0000	2.0000	2.0000	6



le contingenze assolute $n_{ij} - \hat{n}_{ij}$:

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.



le contingenze assolute $n_{ij} - \hat{n}_{ij}$:

	0 ÷ 2	2 ÷ 4	4 ÷ 6
2	-0.3333	0.6667	-0.3333
4	0.3333	-0.6667	0.3333

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.



i rapporti di contingenza $\frac{n_{ij}}{\hat{n}_{ij}}$:

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.



i rapporti di contingenza $\frac{n_{ij}}{\hat{n}_{ij}}$:

	0 - 2	2 - 4	4 - 6
2	0.75	1.5	0.75
4	1.50	0.0	1.50

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.



e le contingenze relative, espresse in termini percentuali, $100 \left(\frac{n_{ij} - \hat{n}_{ij}}{\hat{n}_{ij}} \right) \% = 100 \left(\frac{n_{ij}}{\hat{n}_{ij}} - 1 \right) \%$:

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.



e le contingenze relative, espresse in termini percentuali, $100 \left(\frac{n_{ij} - \hat{n}_{ij}}{\hat{n}_{ij}} \right) \% = 100 \left(\frac{n_{ij}}{\hat{n}_{ij}} - 1 \right) \%$:

	0 + 2	2 + 4	4 + 6
2	-25	50	-25
4	50	-100	50

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.



Risultati intermedi $\frac{(n_{ij} - \hat{n}_{ij})^2}{\hat{n}_{ij}}$ per il calcolo del χ^2

	0 ÷ 2	2 ÷ 4	4 ÷ 6
2	0.0833	0.3333	0.0833
4	0.1667	0.6667	0.1667

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.



Risultati intermedi $\frac{n_{ij}^2}{n_{i\bullet} \cdot n_{\bullet j}}$ per il calcolo del χ^2 (formula operativa)

	0 ÷ 2	2 ÷ 4	4 ÷ 6
2	0.125	0.5	0.125
4	0.250	0.0	0.250

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.



Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.

$$\chi^2 = \sum_{i=1}^h \sum_{j=1}^k \frac{(n_{ij} - \hat{n}_{ij})^2}{\hat{n}_{ij}} = 1.5,$$

$$\chi_N^2 = \frac{\chi^2}{n \cdot \min(h-1, k-1)} = 0.25$$

$$\chi_N = \sqrt{\chi_N^2} = 0.5.$$

Sussiste, quindi, un livello di connessione medio.

Esercizio 2

Prova scritta di
STATISTICA del
14.07.2016

Giuseppe Boari
Gabriele Cantaluppi



UNIVERSITÀ
CATTOLICA
del Sacro Cuore

Si consideri il seguente prospetto:

t	x_t	$NIBF(2010 = 1)$	$NIBM$
2008			
2009	100		.5
2010		1	
2011	100		2
2012		4	

1. Si completi il prospetto con i dati mancanti.

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.



Considerando i numeri indici a base mobile e l'ultimo numero indice a base fissa assegnati è possibile ricostruire la serie dei dati

$$2008/2009 =$$

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.



Considerando i numeri indici a base mobile e l'ultimo numero indice a base fissa assegnati è possibile ricostruire la serie dei dati

$${}_{2008}I_{2009} = \frac{x_{2009}}{x_{2008}} = \frac{100}{x_{2008}} = 0.5 \rightarrow x_{2009} =$$

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.



Considerando i numeri indici a base mobile e l'ultimo numero indice a base fissa assegnati è possibile ricostruire la serie dei dati

$${}_{2008}I_{2009} = \frac{x_{2009}}{x_{2008}} = \frac{100}{x_{2008}} = 0.5 \rightarrow x_{2009} = 200$$

$${}_{2010}I_{2011} =$$

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.



Considerando i numeri indici a base mobile e l'ultimo numero indice a base fissa assegnati è possibile ricostruire la serie dei dati

$${}_{2008}I_{2009} = \frac{x_{2009}}{x_{2008}} = \frac{100}{x_{2008}} = 0.5 \rightarrow x_{2009} = 200$$

$${}_{2010}I_{2011} = \frac{x_{2011}}{x_{2010}} = \frac{100}{x_{2010}} = 2 \rightarrow x_{2010} =$$

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.



Considerando i numeri indici a base mobile e l'ultimo numero indice a base fissa assegnati è possibile ricostruire la serie dei dati

$${}_{2008}I_{2009} = \frac{x_{2009}}{x_{2008}} = \frac{100}{x_{2008}} = 0.5 \rightarrow x_{2009} = 200$$

$${}_{2010}I_{2011} = \frac{x_{2011}}{x_{2010}} = \frac{100}{x_{2010}} = 2 \rightarrow x_{2010} = 50$$

$${}_{2010}I_{2012} =$$

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.



Considerando i numeri indici a base mobile e l'ultimo numero indice a base fissa assegnati è possibile ricostruire la serie dei dati

$${}_{2008}I_{2009} = \frac{x_{2009}}{x_{2008}} = \frac{100}{x_{2008}} = 0.5 \rightarrow x_{2009} = 200$$

$${}_{2010}I_{2011} = \frac{x_{2011}}{x_{2010}} = \frac{100}{x_{2010}} = 2 \rightarrow x_{2010} = 50$$

$${}_{2010}I_{2012} = \frac{x_{2012}}{x_{2010}} = \frac{x_{2012}}{50} = 4 \rightarrow x_{2012} =$$

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.



Considerando i numeri indici a base mobile e l'ultimo numero indice a base fissa assegnati è possibile ricostruire la serie dei dati

$${}_{2008}I_{2009} = \frac{x_{2009}}{x_{2008}} = \frac{100}{x_{2008}} = 0.5 \rightarrow x_{2009} = 200$$

$${}_{2010}I_{2011} = \frac{x_{2011}}{x_{2010}} = \frac{100}{x_{2010}} = 2 \rightarrow x_{2010} = 50$$

$${}_{2010}I_{2012} = \frac{x_{2012}}{x_{2010}} = \frac{x_{2012}}{50} = 4 \rightarrow x_{2012} = 200$$

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.



Considerando la serie dei dati ricostruita abbiamo

$$2010/2008 =$$

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.



Considerando la serie dei dati ricostruita abbiamo

$${}_{2010}I_{2008} = \frac{x_{2008}}{x_{2010}} = \frac{200}{50} = 4$$

$${}_{2010}I_{2009} =$$

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.



Considerando la serie dei dati ricostruita abbiamo

$${}_{2010}I_{2008} = \frac{x_{2008}}{x_{2010}} = \frac{200}{50} = 4$$

$${}_{2010}I_{2009} = \frac{x_{2009}}{x_{2010}} = \frac{100}{50} = 2$$

$${}_{2010}I_{2011} =$$

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.



Considerando la serie dei dati ricostruita abbiamo

$${}_{2010}I_{2008} = \frac{x_{2008}}{x_{2010}} = \frac{200}{50} = 4$$

$${}_{2010}I_{2009} = \frac{x_{2009}}{x_{2010}} = \frac{100}{50} = 2$$

$${}_{2010}I_{2011} = \frac{x_{2011}}{x_{2010}} = \frac{100}{50} = 2$$

$${}_{2009}I_{2010} =$$

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.



Considerando la serie dei dati ricostruita abbiamo

$${}_{2010}I_{2008} = \frac{x_{2008}}{x_{2010}} = \frac{200}{50} = 4$$

$${}_{2010}I_{2009} = \frac{x_{2009}}{x_{2010}} = \frac{100}{50} = 2$$

$${}_{2010}I_{2011} = \frac{x_{2011}}{x_{2010}} = \frac{100}{50} = 2$$

$${}_{2009}I_{2010} = \frac{x_{2010}}{x_{2009}} = \frac{50}{100} = 0.5$$

$${}_{2011}I_{2012} =$$

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.



Considerando la serie dei dati ricostruita abbiamo

$${}_{2010}I_{2008} = \frac{X_{2008}}{X_{2010}} = \frac{200}{50} = 4$$

$${}_{2010}I_{2009} = \frac{X_{2009}}{X_{2010}} = \frac{100}{50} = 2$$

$${}_{2010}I_{2011} = \frac{X_{2011}}{X_{2010}} = \frac{100}{50} = 2$$

$${}_{2009}I_{2010} = \frac{X_{2010}}{X_{2009}} = \frac{50}{100} = 0.5$$

$${}_{2011}I_{2012} = \frac{X_{2012}}{X_{2011}} = \frac{200}{100} = 2$$

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.



Si ottiene, quindi

t	x_t	$NIBF(t = 2010)$	$NIBM$
2008	200	4	—
2009	100	2	0.5
2010	50	1	0.5
2011	100	2	2
2012	200	4	2

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.



2. Si calcoli e si commenti il numero indice $_{2012}I_{2010}$.

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.



2. Si calcoli e si commenti il numero indice ${}_{2012}I_{2010}$.

$${}_{2012}I_{2010} = \frac{x_{2010}}{x_{2012}} = \frac{50}{200} = 0.25$$

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.



Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.

2. Si calcoli e si commenti il numero indice ${}_{2012}I_{2010}$.

$${}_{2012}I_{2010} = \frac{x_{2010}}{x_{2012}} = \frac{50}{200} = 0.25$$

La grandezza al tempo 2010 era inferiore del 75% rispetto alla grandezza al tempo 2012.
Ovvero il valore della grandezza nel 2010 era il 25% del valore registrato nel 2012.



Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.

3. Si calcoli il tasso medio di variazione dal 2008 al 2012 e si commenti il risultato ottenuto.



Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.

3. Si calcoli il tasso medio di variazione dal 2008 al 2012 e si commenti il risultato ottenuto.

La grandezza esaminata non ha registrato variazione tra l'anno 2008 e l'anno 2012.



Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.

Il seguente prospetto riporta i valori rilevati in corrispondenza a 5 unità statistiche con riferimento a due caratteri quantitativi X e Y

x_i	2	2	3	4	4
y_i	2	4	3	2	4

1. Si rappresentino graficamente le coppie dei dati ipotizzando Y in funzione di X .



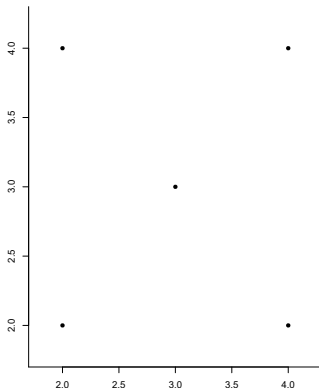
Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.

Le coppie di dati possono essere rappresentate mediante un grafico a dispersione.





2. Si calcolino, secondo il metodo dei minimi quadrati, i parametri e l'indice di adattamento dei seguenti modelli di regressione, individuando il modello migliore:

$$I) Y^* = a + b \cdot X \qquad II) Y^* = c \cdot X$$

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.

Si riportano i calcoli necessari¹ per ottenere le stime dei parametri e gli indici di adattamento del modello

$$Y = a + bX + E$$

x_i	y_i	$x_i y_i$	x_i^2	y_i^2
2	2	4	4	4
2	4	8	4	16
3	3	9	9	9
4	2	8	16	4
4	4	16	16	16
15	15	45	49	49
3	3	9	9.8	9.8
$M(X)$	$M(Y)$	$M(XY)$	$M(X^2)$	$M(Y^2)$

$$\hat{b} =$$



Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.

¹Si osserva che per ragioni di stampa alcuni calcoli potrebbero essere duplicati con riferimento ai diversi modelli oggetto di studio

Si riportano i calcoli necessari¹ per ottenere le stime dei parametri e gli indici di adattamento del modello

$$Y = a + bX + E$$

x_i	y_i	$x_i y_i$	x_i^2	y_i^2
2	2	4	4	4
2	4	8	4	16
3	3	9	9	9
4	2	8	16	4
4	4	16	16	16
15	15	45	49	49
3	3	9	9.8	9.8
$M(X)$	$M(Y)$	$M(XY)$	$M(X^2)$	$M(Y^2)$

$$\begin{aligned}\hat{b} &= \frac{\text{Cov}(X, Y)}{\text{Var}(X)} = \frac{M(XY) - M(X)M(Y)}{M(X^2) - [M(X)]^2} = \\ &= \frac{9 - 3 \cdot 3}{9.8 - 3^2} = \frac{0}{0.8} = 0 \\ \hat{a} &= \end{aligned}$$

¹Si osserva che per ragioni di stampa alcuni calcoli potrebbero essere duplicati con riferimento ai diversi modelli oggetto di studio



Si riportano i calcoli necessari¹ per ottenere le stime dei parametri e gli indici di adattamento del modello

$$Y = a + bX + E$$

x_i	y_i	$x_i y_i$	x_i^2	y_i^2
2	2	4	4	4
2	4	8	4	16
3	3	9	9	9
4	2	8	16	4
4	4	16	16	16
15	15	45	49	49
3	3	9	9.8	9.8
$M(X)$	$M(Y)$	$M(XY)$	$M(X^2)$	$M(Y^2)$

$$\hat{b} = \frac{\text{Cov}(X, Y)}{\text{Var}(X)} = \frac{M(XY) - M(X)M(Y)}{M(X^2) - [M(X)]^2} =$$

$$= \frac{9 - 3 \cdot 3}{9.8 - 3^2} = \frac{0}{0.8} = 0$$

$$\hat{a} = M(Y) - \hat{b}M(X) = 3$$

$$\rho^2 =$$



Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.

¹Si osserva che per ragioni di stampa alcuni calcoli potrebbero essere duplicati con riferimento ai diversi modelli oggetto di studio

Si riportano i calcoli necessari¹ per ottenere le stime dei parametri e gli indici di adattamento del modello

$$Y = a + bX + E$$

x_i	y_i	$x_i y_i$	x_i^2	y_i^2
2	2	4	4	4
2	4	8	4	16
3	3	9	9	9
4	2	8	16	4
4	4	16	16	16
15	15	45	49	49
3	3	9	9.8	9.8
$M(X)$	$M(Y)$	$M(XY)$	$M(X^2)$	$M(Y^2)$

$$\hat{b} = \frac{\text{Cov}(X, Y)}{\text{Var}(X)} = \frac{M(XY) - M(X)M(Y)}{M(X^2) - [M(X)]^2} =$$

$$= \frac{9 - 3 \cdot 3}{9.8 - 3^2} = \frac{0}{0.8} = 0$$

$$\hat{a} = M(Y) - \hat{b}M(X) = 3$$

$$\rho^2 = \frac{[\text{Cov}(X, Y)]^2}{\text{Var}(X) \text{Var}(Y)} = \frac{0^2}{0.8 \cdot 0.8} = 0$$

$$M(\hat{E}^2) = \text{Var}(\hat{E}) = \text{varianza residua} =$$

¹Si osserva che per ragioni di stampa alcuni calcoli potrebbero essere duplicati con riferimento ai diversi modelli oggetto di studio



Si riportano i calcoli necessari¹ per ottenere le stime dei parametri e gli indici di adattamento del modello

$$Y = a + bX + E$$

x_i	y_i	$x_i y_i$	x_i^2	y_i^2
2	2	4	4	4
2	4	8	4	16
3	3	9	9	9
4	2	8	16	4
4	4	16	16	16
15	15	45	49	49
3	3	9	9.8	9.8
$M(X)$	$M(Y)$	$M(XY)$	$M(X^2)$	$M(Y^2)$

$$\hat{b} = \frac{\text{Cov}(X, Y)}{\text{Var}(X)} = \frac{M(XY) - M(X)M(Y)}{M(X^2) - [M(X)]^2} =$$

$$= \frac{9 - 3 \cdot 3}{9.8 - 3^2} = \frac{0}{0.8} = 0$$

$$\hat{a} = M(Y) - \hat{b}M(X) = 3$$

$$\rho^2 = \frac{[\text{Cov}(X, Y)]^2}{\text{Var}(X) \text{Var}(Y)} = \frac{0^2}{0.8 \cdot 0.8} = 0$$

$$M(\hat{E}^2) = \text{Var}(\hat{E}) = \text{varianza residua} = \text{Var}(Y) \cdot (1 - \rho^2) = 0.8$$

¹Si osserva che per ragioni di stampa alcuni calcoli potrebbero essere duplicati con riferimento ai diversi modelli oggetto di studio





Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.

Si riportano i calcoli necessari² per ottenere le stime dei parametri e la misura di adattamento del modello

$$Y = cX + E$$

x_i	y_i	$x_i y_i$	x_i^2	y_i^2
2	2	4	4	4
2	4	8	4	16
3	3	9	9	9
4	2	8	16	4
4	4	16	16	16
		45	49	49
		9	9.8	9.8
		$M(XY)$	$M(X^2)$	$M(Y^2)$

$$\hat{c} =$$

²Si osserva che per ragioni di stampa alcuni calcoli potrebbero essere duplicati con riferimento ai diversi modelli oggetto di studio



Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.

Si riportano i calcoli necessari² per ottenere le stime dei parametri e la misura di adattamento del modello

$$Y = cX + E$$

x_i	y_i	$x_i y_i$	x_i^2	y_i^2
2	2	4	4	4
2	4	8	4	16
3	3	9	9	9
4	2	8	16	4
4	4	16	16	16
		45	49	49
		9	9.8	9.8
		$M(XY)$	$M(X^2)$	$M(Y^2)$

$$\hat{c} = \frac{M(XY)}{M(X^2)} = \frac{9}{9.8} = 0.9184$$

$$M(\hat{E}^2) =$$

²Si osserva che per ragioni di stampa alcuni calcoli potrebbero essere duplicati con riferimento ai diversi modelli oggetto di studio



Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.

Si riportano i calcoli necessari² per ottenere le stime dei parametri e la misura di adattamento del modello

$$Y = cX + E$$

x_i	y_i	$x_i y_i$	x_i^2	y_i^2
2	2	4	4	4
2	4	8	4	16
3	3	9	9	9
4	2	8	16	4
4	4	16	16	16
		45	49	49
		9	9.8	9.8
		$M(XY)$	$M(X^2)$	$M(Y^2)$

$$\hat{c} = \frac{M(XY)}{M(X^2)} = \frac{9}{9.8} = 0.9184$$

$$M(\hat{E}^2) = M[Y \cdot (Y - \hat{c}X)] = M(Y^2) - \hat{c}M(XY) = 9.8 - 0.9184 \cdot 9 = 1.5347$$

²Si osserva che per ragioni di stampa alcuni calcoli potrebbero essere duplicati con riferimento ai diversi modelli oggetto di studio



Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.

$\hat{a} = 3$, $\hat{b} = 0$, indice adattamento $M(\hat{E}^2) = 0.8$ $\hat{c} = 0.9184$, indice adattamento
 $M(\hat{E}^2) = 1.5347$



Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.

3. Si giustifichi se vale la relazione $\eta_{Y|X}^2 \geq \eta_{X|Y}^2$.



Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.

3. Si giustifichi se vale la relazione $\eta_{Y|X}^2 \geq \eta_{X|Y}^2$.

Dalla rappresentazione grafica si evince che nel caso in esame sussistono contemporaneamente indipendenza in media di Y da X , abbiamo infatti

$$M(Y|X=2) = M(Y|X=3) = M(Y|X=4) = 3$$

e indipendenza in media di X da Y , abbiamo infatti

$$M(X|Y=2) = M(X|Y=3) = M(X|Y=4) = 3$$

Vale, quindi

$$\eta_{Y|X}^2 = \eta_{X|Y}^2 = 0.$$

Esercizio 4

Prova scritta di
STATISTICA del
14.07.2016

Giuseppe Boari
Gabriele Cantaluppi



UNIVERSITÀ
CATTOLICA
del Sacro Cuore

In un gioco si possono vincere, con un biglietto del costo di 1 Euro, fino a due premi: il premio a di 1 Euro con probabilità 0.11 e il premio b di 2 Euro con probabilità 0.15. La probabilità complessiva di vincere o uno o l'altro dei due premi è pari a 0.20.

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.



1. Si calcoli la probabilità di vincere entrambi i premi.

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.



1. Si calcoli la probabilità di vincere entrambi i premi.

Abbiamo $P(a \cap b) = 0.06$

Vale infatti

$$P(a \cup b) =$$

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.



1. Si calcoli la probabilità di vincere entrambi i premi.

Abbiamo $P(a \cap b) = 0.06$

Vale infatti

$$\begin{aligned} P(a \cup b) &= P(a) + P(b) - P(a \cap b) \\ 0.20 &= \end{aligned}$$

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.



Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.

1. Si calcoli la probabilità di vincere entrambi i premi.

Abbiamo $P(a \cap b) = 0.06$

Vale infatti

$$P(a \cup b) = P(a) + P(b) - P(a \cap b)$$

$$0.20 = 0.11 + 0.15 - P(a \cap b)$$

$$P(a \cap b) =$$



Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.

1. Si calcoli la probabilità di vincere entrambi i premi.

Abbiamo $P(a \cap b) = 0.06$

Vale infatti

$$P(a \cup b) = P(a) + P(b) - P(a \cap b)$$

$$0.20 = 0.11 + 0.15 - P(a \cap b)$$

$$P(a \cap b) = 0.11 + 0.15 - 0.20 = 0.06$$



1. Si stabilisca se il gioco è equo.

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.



1. Si stabilisca se il gioco è equo.
La variabile casuale guadagno risulta

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.



1. Si stabilisca se il gioco è equo.
La variabile casuale guadagno risulta

x_i	$P(X = x_i)$	$x_i \cdot P(X = x_i)$
-1		

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.



1. Si stabilisca se il gioco è equo.

La variabile casuale guadagno risulta

x_i	$P(X = x_i)$	$x_i \cdot P(X = x_i)$
-1	$1 - P(a \cup b) = 1 - .20 = 0.80$	-0.80
0		

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.



1. Si stabilisca se il gioco è equo.

La variabile casuale guadagno risulta

x_i	$P(X = x_i)$	$x_i \cdot P(X = x_i)$
-1	$1 - P(a \cup b) = 1 - .20 = 0.80$	-0.80
0	$P(a - b) = P(a) - P(a \cap b) = 0.11 - 0.06 = 0.05$	0.00
1		

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.



1. Si stabilisca se il gioco è equo.

La variabile casuale guadagno risulta

x_i	$P(X = x_i)$	$x_i \cdot P(X = x_i)$
-1	$1 - P(a \cup b) = 1 - .20 = 0.80$	-0.80
0	$P(a - b) = P(a) - P(a \cap b) = 0.11 - 0.06 = 0.05$	0.00
1	$P(b - a) = P(b) - P(a \cap b) = 0.15 - 0.06 = 0.09$	0.09
2		

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.



1. Si stabilisca se il gioco è equo.

La variabile casuale guadagno risulta

x_i	$P(X = x_i)$	$x_i \cdot P(X = x_i)$
-1	$1 - P(a \cup b) = 1 - .20 = 0.80$	-0.80
0	$P(a - b) = P(a) - P(a \cap b) = 0.11 - 0.06 = 0.05$	0.00
1	$P(b - a) = P(b) - P(a \cap b) = 0.15 - 0.06 = 0.09$	0.09
2	$P(a \cap b) = 0.06$	0.12
	1	-0.59

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.



Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.

1. Si stabilisca se il gioco è equo.

La variabile casuale guadagno risulta

x_i	$P(X = x_i)$	$x_i \cdot P(X = x_i)$
-1	$1 - P(a \cup b) = 1 - .20 = 0.80$	-0.80
0	$P(a - b) = P(a) - P(a \cap b) = 0.11 - 0.06 = 0.05$	0.00
1	$P(b - a) = P(b) - P(a \cap b) = 0.15 - 0.06 = 0.09$	0.09
2	$P(a \cap b) = 0.06$	0.12
	1	-0.59

Il valore atteso della variabile casuale guadagno è negativo, quindi il gioco è sfavorevole al giocatore.



2. Nell'ipotesi di ripetere il gioco 3 volte si calcoli la probabilità di vincere il premio a oppure il premio b almeno 2 volte.

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.



Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.

2. Nell'ipotesi di ripetere il gioco 3 volte si calcoli la probabilità di vincere il premio a oppure il premio b almeno 2 volte.

Sia X la variabile casuale 'n. di vincite nelle 3 prove indipendenti'.

$$X \sim$$



Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.

2. Nell'ipotesi di ripetere il gioco 3 volte si calcoli la probabilità di vincere il premio a oppure il premio b almeno 2 volte.

Sia X la variabile casuale 'n. di vincite nelle 3 prove indipendenti'.

$$X \sim \text{Bin}(n = 3, p = 0.20).$$

$$P(X \geq 2) =$$



Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.

2. Nell'ipotesi di ripetere il gioco 3 volte si calcoli la probabilità di vincere il premio a oppure il premio b almeno 2 volte.

Sia X la variabile casuale 'n. di vincite nelle 3 prove indipendenti'.

$$X \sim \text{Bin}(n = 3, p = 0.20).$$

$$\begin{aligned} P(X \geq 2) &= P(X = 2) + P(X = 3) = \\ &= \binom{3}{2} 0.20^2 0.80^1 + \binom{3}{3} 0.20^3 0.80^0 \\ &= 0.096 + 0.008 = 0.104. \end{aligned}$$



3. Nell'ipotesi di ripetere il gioco 100 volte si calcoli la probabilità di vincere il premio a oppure il premio b almeno 12 volte.

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.



3. Nell'ipotesi di ripetere il gioco 100 volte si calcoli la probabilità di vincere il premio a oppure il premio b almeno 12 volte.

Indicata con X la variabile casuale 'numero di successi' nelle 100 prove indipendenti risulta $X \sim$

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.



Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.

3. Nell'ipotesi di ripetere il gioco 100 volte si calcoli la probabilità di vincere il premio a oppure il premio b almeno 12 volte.

Indicata con X la variabile casuale 'numero di successi' nelle 100 prove indipendenti risulta $X \sim \text{Bin}(n = 100, p = 0.2)$. $M(X) = 20$; $\text{Var}(X) = 16$.

Dal momento che $np = 20 > 5$ ed $n(1 - p) = 80 > 5$ è possibile utilizzare l'approssimazione della variabile casuale Binomiale con la variabile casuale Normale $X \sim N(\mu = 20, \sigma^2 = 16)$. Si ottiene

$$\begin{aligned} P(X \geq 12) &= P\left(\frac{X - \mu}{\sigma} \geq \frac{12 - \mu}{\sigma}\right) \\ &= P\left(Z \geq \frac{12 - 20}{4}\right) = P(Z \geq -2) = P(Z \leq 2) = 0.9772. \end{aligned}$$

Tema T2901

Prova scritta di STATISTICA del 23.08.2016

Prova scritta di
STATISTICA del
23.08.2016

Giuseppe Boari
Gabriele Cantaluppi



UNIVERSITÀ
CATTOLICA
del Sacro Cuore

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.

Giuseppe Boari
Gabriele Cantaluppi
Università Cattolica del Sacro Cuore, Milano



La seguente tabella riporta alcune informazioni sulla distribuzione congiunta della variabile statistica doppia (X, Y) , con X carattere qualitativo sconnesso.

$X \setminus Y$	y_1	y_2	y_3	y_4	
x_1		2			
x_2					

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.



1. Si completino le frequenze nella tabella in modo tale che siano verificate contemporaneamente le tre seguenti condizioni: a) la distribuzione $Y|x_1$ sia caratterizzata da massima eterogeneità; b) $\chi^2_{\max} = 12$; c) sussista indipendenza stocastica tra le variabili X e Y .

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.



1. Si completino le frequenze nella tabella in modo tale che siano verificate contemporaneamente le tre seguenti condizioni: a) la distribuzione $Y|x_1$ sia caratterizzata da massima eterogeneità; b) $\chi^2_{\max} = 12$; c) sussista indipendenza stocastica tra le variabili X e Y .

La situazione di massima eterogeneità si realizza in presenza di equidistribuzione delle frequenze, quindi $n_{11} = n_{13} = n_{14} = 2$ e $n_{1\bullet} = 8$.

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.



1. Si completino le frequenze nella tabella in modo tale che siano verificate contemporaneamente le tre seguenti condizioni: a) la distribuzione $Y|x_1$ sia caratterizzata da massima eterogeneità; b) $\chi^2_{\max} = 12$; c) sussista indipendenza stocastica tra le variabili X e Y .

La situazione di massima eterogeneità si realizza in presenza di equidistribuzione delle frequenze, quindi $n_{11} = n_{13} = n_{14} = 2$ e $n_{1\bullet} = 8$.

Ricordando che $\chi^2_{\max} = n \cdot \min(h-1, k-1) = n \cdot \min(2-1, 4-1) = n \cdot \min(1, 3) = n \cdot 1 = 12$ si ottiene $n = 12$ e $n_{2\bullet} = n - n_{1\bullet} = 12 - 8 = 4$.

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.



1. Si completino le frequenze nella tabella in modo tale che siano verificate contemporaneamente le tre seguenti condizioni: a) la distribuzione $Y|x_1$ sia caratterizzata da massima eterogeneità; b) $\chi^2_{\max} = 12$; c) sussista indipendenza stocastica tra le variabili X e Y .

La situazione di massima eterogeneità si realizza in presenza di equidistribuzione delle frequenze, quindi $n_{11} = n_{13} = n_{14} = 2$ e $n_{1\bullet} = 8$.

Ricordando che $\chi^2_{\max} = n \cdot \min(h-1, k-1) = n \cdot \min(2-1, 4-1) = n \cdot \min(1, 3) = n \cdot 1 = 12$ si ottiene $n = 12$ e $n_{2\bullet} = n - n_{1\bullet} = 12 - 8 = 4$.

Dal momento che sussiste indipendenza stocastica, le distribuzioni condizionate sono somiglianti (proporzionali) e, quindi, $n_{21} = n_{22} = n_{23} = n_{24} = 1$.

$X \setminus Y$	y_1	y_2	y_3	y_4	
x_1	2	2	2	2	8
x_2	1	1	1	1	4
	3	3	3	3	12



2. Si sostituiscano ora alle modalità y_1, y_2, y_3 , e y_4 di Y le classi $2 \dashv 4, 4 \dashv 6, 6 \dashv 10$ e $10 \dashv 12$ e le frequenze congiunte $n_{11} = n_{13} = n_{14} = 2$ e $n_{21} = n_{22} = n_{23} = n_{24} = 11$.

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.



2. Si sostituiscano ora alle modalità y_1, y_2, y_3 , e y_4 di Y le classi $2 \vdash 4$, $4 \vdash 6$, $6 \vdash 10$ e $10 \vdash 12$ e le frequenze congiunte $n_{11} = n_{13} = n_{14} = 2$ e $n_{21} = n_{22} = n_{23} = n_{24} = 11$.

$X \setminus Y$	$2 \vdash 4$	$4 \vdash 6$	$6 \vdash 10$	$10 \vdash 12$	
x_1	2	2	2	2	8
x_2	11	11	11	11	44
	13	13	13	13	52

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.



Si rappresentino graficamente le funzioni di ripartizione delle distribuzioni marginali X e Y .

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.



Si rappresentino graficamente le funzioni di ripartizione delle distribuzioni marginali X e Y .
Dal momento che il carattere qualitativo X è sconnesso, le frequenze cumulate e la funzione di ripartizione non sono definite.

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.



Si rappresentino graficamente le funzioni di ripartizione delle distribuzioni marginali X e Y .
Dal momento che il carattere qualitativo X è sconnesso, le frequenze cumulate e la funzione di ripartizione non sono definite.

Si veda al punto successivo per la funzione di ripartizione di Y .

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.



3. Si costruisca il grafico Box & Whiskers plot della marginale Y .

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.



Si considera la distribuzione marginale della variabile Y , ottenuta sommando le frequenze di colonna

y_j	$n_{\bullet j}$
2 $\rightarrow$ 4	13
4 $\rightarrow$ 6	13
6 $\rightarrow$ 10	13
10 $\rightarrow$ 12	13

Si tratta di un carattere quantitativo con modalità raggruppate in classi.

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.



Per costruire il grafico Box & Whiskers plot occorre determinare in primo luogo i valori dei tre quartili $y_{0.25}$, $y_{0.5}$ e $y_{0.75}$ considerando l'inversa della Funzione di Ripartizione in corrispondenza rispettivamente di 0.25, 0.50 e 0.75.

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.



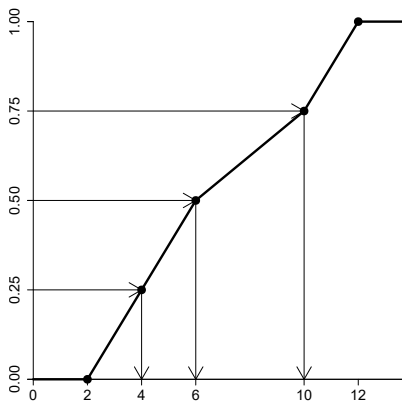
Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.

y_j	$n_{\bullet j}$	$N_{\bullet j}$	$F_{\bullet j}$
2 $\vdash$ 4	13	13	0.25
4 $\vdash$ 6	13	26	0.5
6 $\vdash$ 10	13	39	0.75
10 $\vdash$ 12	13	52	1





Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.

Si ottiene

$$y_{0.25} = 2 + \frac{0.25 - 0}{0.25 - 0} \cdot (4 - 2) = 4$$

$$y_{0.5} = 4 + \frac{0.5 - 0.25}{0.5 - 0.25} \cdot (6 - 4) = 6$$

$$y_{0.75} = 6 + \frac{0.75 - 0.5}{0.75 - 0.5} \cdot (10 - 6) = 10$$

La lunghezza teorica dei baffi risulta

$$1.5 \cdot (y_{0.75} - y_{0.25}) = 9.$$

Il baffo inferiore è posizionato in

$$\max(y_{\min}, y_{0.25} - 9) = \max(2, -5) = 2.$$

Il baffo superiore è posizionato in

$$\min(y_{\max}, y_{0.75} + 9) = \min(12, 19) = 12.$$

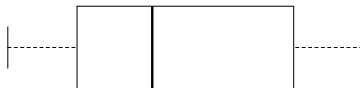
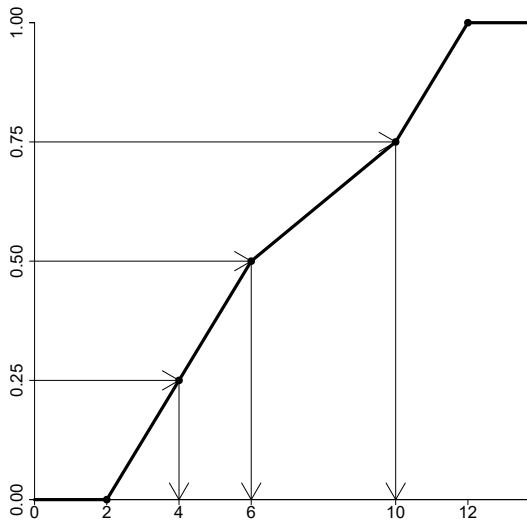


Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.





4. Si determini il grado di connessione tra le variabili X e Y mediante il calcolo di un opportuno indice.

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.



4. Si determini il grado di connessione tra le variabili X e Y mediante il calcolo di un opportuno indice.

Sussiste indipendenza stocastica, quindi $\chi^2 = \chi_N^2 = \chi_N = 0$

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.

Esercizio 2



Si consideri la seguente serie storica:

t	x_t
2011	300
2012	150
2013	75
2014	150
2015	300

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.



1. Si ricostruiscano le serie dei numeri indici a base fissa, con base 2012, e a base mobile.

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.



1. Si ricostruiscano le serie dei numeri indici a base fissa, con base 2012, e a base mobile.
Si ricorda che ${}_{2012}I_t = \frac{x_t}{x_{2012}}$.

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.



1. Si ricostruiscano le serie dei numeri indici a base fissa, con base 2012, e a base mobile. Si ricorda che ${}_{2012}I_t = \frac{x_t}{x_{2012}}$. Si ottiene quindi

Anno	x_t	NIBF (2012)
2011	300	2.0
2012	150	1.0
2013	75	0.5
2014	150	1.0
2015	300	2.0

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.



Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.

1. Si ricostruiscano le serie dei numeri indici a base fissa, con base 2012, e a base mobile. Si ricorda che ${}_{2012}I_t = \frac{x_t}{x_{2012}}$. Si ottiene quindi

Anno	x_t	NIBF (2012)
2011	300	2.0
2012	150	1.0
2013	75	0.5
2014	150	1.0
2015	300	2.0

Si ricorda che ${}_{t-1}I_t = \frac{x_t}{x_{t-1}}$.



Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.

1. Si ricostruiscano le serie dei numeri indici a base fissa, con base 2012, e a base mobile. Si ricorda che ${}_{2012}I_t = \frac{x_t}{x_{2012}}$. Si ottiene quindi

Anno	x_t	NIBF (2012)
2011	300	2.0
2012	150	1.0
2013	75	0.5
2014	150	1.0
2015	300	2.0

Si ricorda che ${}_{t-1}I_t = \frac{x_t}{x_{t-1}}$. Si ottiene quindi

Anno	x_t	NIBM
2011	300	
2012	150	0.5
2013	75	0.5
2014	150	2.0
2015	300	2.0



2. Si calcolino e si commentino i seguenti numeri indici ${}_{2011}I_{2013}$ e ${}_{2013}I_{2015}$.

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.



2. Si calcolino e si commentino i seguenti numeri indici ${}_{2011}I_{2013}$ e ${}_{2013}I_{2015}$.

$${}_{2011}I_{2013} = \frac{x_{2013}}{x_{2011}} = \frac{75}{300} = 0.25$$

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.



2. Si calcolino e si commentino i seguenti numeri indici ${}_{2011}I_{2013}$ e ${}_{2013}I_{2015}$.

$${}_{2011}I_{2013} = \frac{x_{2013}}{x_{2011}} = \frac{75}{300} = 0.25$$

Tra il tempo 2011 e il tempo 2013 si è registrato un decremento complessivo del 75% della grandezza oggetto di studio.

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.



Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.

2. Si calcolino e si commentino i seguenti numeri indici ${}_{2011}I_{2013}$ e ${}_{2013}I_{2015}$.

$${}_{2011}I_{2013} = \frac{x_{2013}}{x_{2011}} = \frac{75}{300} = 0.25$$

Tra il tempo 2011 e il tempo 2013 si è registrato un decremento complessivo del 75% della grandezza oggetto di studio.

$${}_{2013}I_{2015} = \frac{x_{2015}}{x_{2013}} = \frac{300}{75} = 4$$



Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.

2. Si calcolino e si commentino i seguenti numeri indici ${}_{2011}I_{2013}$ e ${}_{2013}I_{2015}$.

$${}_{2011}I_{2013} = \frac{x_{2013}}{x_{2011}} = \frac{75}{300} = 0.25$$

Tra il tempo 2011 e il tempo 2013 si è registrato un decremento complessivo del 75% della grandezza oggetto di studio.

$${}_{2013}I_{2015} = \frac{x_{2015}}{x_{2013}} = \frac{300}{75} = 4$$

Tra il tempo 2013 e il tempo 2015 si è registrato un incremento complessivo del 300% della grandezza oggetto di studio.



3. Si calcoli il tasso medio di variazione dal 2012 al 2015 e si commenti il risultato ottenuto.

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.



3. Si calcoli il tasso medio di variazione dal 2012 al 2015 e si commenti il risultato ottenuto.
Si ricorda che

$${}_{2012}I_{2015} = {}_{2012}I_{2013} \cdot {}_{2013}I_{2014} \cdot {}_{2014}I_{2015}$$

Si deve trovare il numero indice a base mobile medio.

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.



Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.

3. Si calcoli il tasso medio di variazione dal 2012 al 2015 e si commenti il risultato ottenuto. Si ricorda che

$${}_{2012}I_{2015} = {}_{2012}I_{2013} \cdot {}_{2013}I_{2014} \cdot {}_{2014}I_{2015}$$

Si deve trovare il numero indice a base mobile medio. Applicando il criterio di scelta della media secondo Chisini si ottiene

$${}_{2012}I_{2015} = {}_{2012}I_{2013} \cdot {}_{2013}I_{2014} \cdot {}_{2014}I_{2015} = \alpha \cdot \alpha \cdot \alpha = \alpha^3.$$



Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.

3. Si calcoli il tasso medio di variazione dal 2012 al 2015 e si commenti il risultato ottenuto. Si ricorda che

$${}_{2012}I_{2015} = {}_{2012}I_{2013} \cdot {}_{2013}I_{2014} \cdot {}_{2014}I_{2015}$$

Si deve trovare il numero indice a base mobile medio. Applicando il criterio di scelta della media secondo Chisini si ottiene

$${}_{2012}I_{2015} = {}_{2012}I_{2013} \cdot {}_{2013}I_{2014} \cdot {}_{2014}I_{2015} = \alpha \cdot \alpha \cdot \alpha = \alpha^3.$$

Quindi

$$\alpha = \sqrt[3]{{}_{2012}I_{2015}} = \sqrt[3]{{}_{2012}I_{2013} \cdot {}_{2013}I_{2014} \cdot {}_{2014}I_{2015}} = 1.2599.$$

Da cui segue una variazione relativa media pari a 0.2599, corrispondente a un incremento medio annuo tra il 2012 e il 2015 del 25.99%.



Il seguente prospetto riporta i valori rilevati in corrispondenza a 5 unità statistiche con riferimento a due caratteri quantitativi X e Y

x_i	16	25	36	49	64
y_i	6	8	9	10	12

1. Si rappresentino graficamente le coppie dei dati ipotizzando Y in funzione di X .

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.



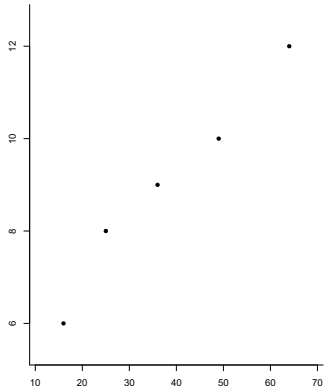
Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.

Le coppie di dati possono essere rappresentate mediante un grafico a dispersione.





Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.

2. Si calcolino, secondo il metodo dei minimi quadrati, i parametri e l'indice di adattamento dei seguenti modelli di regressione, individuando il modello migliore:

$$I) Y^* = a + b \cdot \sqrt{X} \qquad II) Y^* = c$$

Si riportano i calcoli necessari¹ per ottenere le stime dei parametri e gli indici di adattamento del modello

$$Y = a + b\sqrt{X} + E$$

$\sqrt{x_i}$	y_i	$\sqrt{x_i}y_i$	$\sqrt{x_i}^2$	y_i^2
4	6	24	16	36
5	8	40	25	64
6	9	54	36	81
7	10	70	49	100
8	12	96	64	144
30	45	284	190	425
6	9	56.8	38	85
$M(\sqrt{X})$	$M(Y)$	$M(\sqrt{X}Y)$	$M(\sqrt{X}^2)$	$M(Y^2)$

$$\hat{b} =$$



Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.

¹Si osserva che per ragioni di stampa alcuni calcoli potrebbero essere duplicati con riferimento ai diversi modelli oggetto di studio

Si riportano i calcoli necessari¹ per ottenere le stime dei parametri e gli indici di adattamento del modello

$$Y = a + b\sqrt{X} + E$$

$\sqrt{x_i}$	y_i	$\sqrt{x_i}y_i$	$\sqrt{x_i}^2$	y_i^2
4	6	24	16	36
5	8	40	25	64
6	9	54	36	81
7	10	70	49	100
8	12	96	64	144
30	45	284	190	425
6	9	56.8	38	85
$M(\sqrt{X})$	$M(Y)$	$M(\sqrt{X}Y)$	$M(\sqrt{X}^2)$	$M(Y^2)$

$$\hat{b} = \frac{\text{Cov}(\sqrt{X}, Y)}{\text{Var}(\sqrt{X})} = \frac{M(\sqrt{X}Y) - M(\sqrt{X})M(Y)}{M(\sqrt{X}^2) - [M(\sqrt{X})]^2} =$$

$$= \frac{56.8 - 6 \cdot 9}{38 - 6^2} = \frac{2.8}{2} = 1.4$$

$$\hat{a} =$$



¹Si osserva che per ragioni di stampa alcuni calcoli potrebbero essere duplicati con riferimento ai diversi modelli oggetto di studio

Si riportano i calcoli necessari¹ per ottenere le stime dei parametri e gli indici di adattamento del modello

$$Y = a + b\sqrt{X} + E$$

$\sqrt{x_i}$	y_i	$\sqrt{x_i}y_i$	$\sqrt{x_i}^2$	y_i^2
4	6	24	16	36
5	8	40	25	64
6	9	54	36	81
7	10	70	49	100
8	12	96	64	144
30	45	284	190	425
6	9	56.8	38	85
$M(\sqrt{X})$	$M(Y)$	$M(\sqrt{X}Y)$	$M(\sqrt{X}^2)$	$M(Y^2)$

$$\hat{b} = \frac{\text{Cov}(\sqrt{X}, Y)}{\text{Var}(\sqrt{X})} = \frac{M(\sqrt{X}Y) - M(\sqrt{X})M(Y)}{M(\sqrt{X}^2) - [M(\sqrt{X})]^2} =$$

$$= \frac{56.8 - 6 \cdot 9}{38 - 6^2} = \frac{2.8}{2} = 1.4$$

$$\hat{a} = M(Y) - \hat{b}M(\sqrt{X}) = 0.6$$

$$\rho^2 =$$



¹Si osserva che per ragioni di stampa alcuni calcoli potrebbero essere duplicati con riferimento ai diversi modelli oggetto di studio

Si riportano i calcoli necessari¹ per ottenere le stime dei parametri e gli indici di adattamento del modello

$$Y = a + b\sqrt{X} + E$$

$\sqrt{x_i}$	y_i	$\sqrt{x_i}y_i$	$\sqrt{x_i}^2$	y_i^2
4	6	24	16	36
5	8	40	25	64
6	9	54	36	81
7	10	70	49	100
8	12	96	64	144
30	45	284	190	425
6	9	56.8	38	85
$M(\sqrt{X})$	$M(Y)$	$M(\sqrt{X}Y)$	$M(\sqrt{X}^2)$	$M(Y^2)$

$$\begin{aligned}\hat{b} &= \frac{\text{Cov}(\sqrt{X}, Y)}{\text{Var}(\sqrt{X})} = \frac{M(\sqrt{X}Y) - M(\sqrt{X})M(Y)}{M(\sqrt{X}^2) - [M(\sqrt{X})]^2} = \\ &= \frac{56.8 - 6 \cdot 9}{38 - 6^2} = \frac{2.8}{2} = 1.4 \\ \hat{a} &= M(Y) - \hat{b}M(\sqrt{X}) = 0.6 \\ \rho^2 &= \frac{[\text{Cov}(\sqrt{X}, Y)]^2}{\text{Var}(\sqrt{X}) \text{Var}(Y)} = \frac{2.8^2}{2 \cdot 4} = 0.98\end{aligned}$$

¹Si osserva che per ragioni di stampa alcuni calcoli potrebbero essere duplicati con riferimento ai diversi modelli oggetto di studio





Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.

Si riportano i calcoli necessari² per ottenere le stime dei parametri e gli indici di adattamento del modello

$$Y = c + E$$

x_i	y_i	y_i^2
16	6	36
25	8	64
36	9	81
49	10	100
64	12	144
		45
		425
		9
		85
		$M(Y)$
		$M(Y^2)$

$$\hat{c} =$$

²Si osserva che per ragioni di stampa alcuni calcoli potrebbero essere duplicati con riferimento ai diversi modelli oggetto di studio



Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.

Si riportano i calcoli necessari² per ottenere le stime dei parametri e gli indici di adattamento del modello

$$Y = c + E$$

x_i	y_i	y_i^2
16	6	36
25	8	64
36	9	81
49	10	100
64	12	144
		45
		85
		$M(Y)$ $M(Y^2)$

$$\hat{c} = M(Y) = 9$$

$$R^2 =$$

²Si osserva che per ragioni di stampa alcuni calcoli potrebbero essere duplicati con riferimento ai diversi modelli oggetto di studio



Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.

Si riportano i calcoli necessari² per ottenere le stime dei parametri e gli indici di adattamento del modello

$$Y = c + E$$

x_i	y_i	y_i^2
16	6	36
25	8	64
36	9	81
49	10	100
64	12	144
		45
		85
		$M(Y)$ $M(Y^2)$

$$\hat{c} = M(Y) = 9$$

$$R^2 = 0$$

²Si osserva che per ragioni di stampa alcuni calcoli potrebbero essere duplicati con riferimento ai diversi modelli oggetto di studio



$\hat{a} = 0.6$, $\hat{b} = 1.4$, indice adattamento $\rho^2 = 0.98$ $\hat{c} = 9$, indice adattamento $R^2 = 0$
Il modello migliore è $Y^* = 0.6 + 1.4 \cdot \sqrt{X}$.

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.



3. Si giustifichi se vale la relazione $\eta^2_{Y|X} > \eta^2_{X|Y}$.

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.



3. Si giustifichi se vale la relazione $\eta_{Y|X}^2 > \eta_{X|Y}^2$.

La relazione non vale; infatti, $\eta_{Y|X}^2 = 1$ ed $\eta_{X|Y}^2 = 1$ in quanto sussiste dipendenza funzionale bijectiva.

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.

Esercizio 4

Prova scritta di
STATISTICA del
23.08.2016

Giuseppe Boari
Gabriele Cantaluppi



UNIVERSITÀ
CATTOLICA
del Sacro Cuore

Per il seguente gioco è prevista una posta di 1 Euro: si lanciano due dadi; se la somma dei punteggi ottenuti con i due dadi è inferiore a 5 si vincono 10 Euro, se invece è superiore a 9 si vincono 15 Euro.

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.



1. Si stabilisca se il gioco è equo.

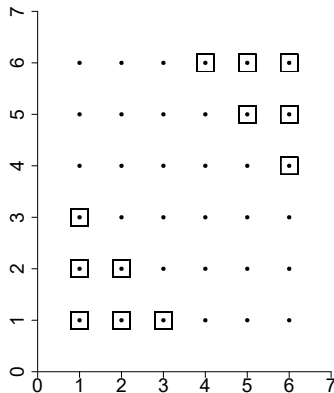
Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.

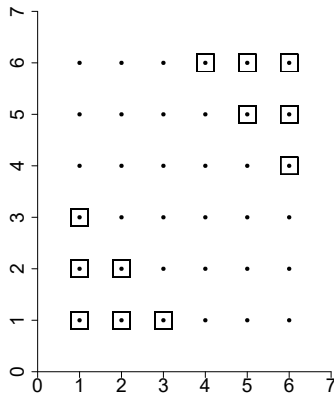
Occorre innanzitutto determinare le probabilità di vincere rispettivamente 10 e 15 Euro.
Con riferimento allo spazio probabilistico



si evince che entrambe le probabilità in oggetto sono entrambe pari a



Occorre innanzitutto determinare le probabilità di vincere rispettivamente 10 e 15 Euro.
Con riferimento allo spazio probabilistico



si evince che entrambe le probabilità in oggetto sono entrambe pari a

$$P(\text{vincere 10 Euro}) = P(\text{vincere 15 Euro}) = \frac{6}{36}.$$





Abbiamo quindi

	x_i	$P(X = x_i)$	$x_i \cdot P(X = x_i)$
<i>posta</i>	-1	$\frac{24}{36}$	$-\frac{24}{36}$
<i>vincita se somma < 5 – posta</i>	9	$\frac{6}{36}$	$\frac{54}{36}$
<i>vincita se somma > 9 – posta</i>	14	$\frac{6}{36}$	$\frac{84}{36}$
		1	$\frac{114}{36}$

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.



Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.

Abbiamo quindi

	x_i	$P(X = x_i)$	$x_i \cdot P(X = x_i)$
<i>posta</i>	-1	$\frac{24}{36}$	$-\frac{24}{36}$
<i>vincita se somma < 5 – posta</i>	9	$\frac{6}{36}$	$\frac{54}{36}$
<i>vincita se somma > 9 – posta</i>	14	$\frac{6}{36}$	$\frac{84}{36}$
		1	$\frac{114}{36}$

Il valore atteso della variabile casuale guadagno è positivo, quindi il gioco non è equo, ma favorevole al giocatore.



2. Nell'ipotesi di ripetere il gioco 10 volte si calcoli la probabilità di vincere almeno 2 volte.

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.



2. Nell'ipotesi di ripetere il gioco 10 volte si calcoli la probabilità di vincere almeno 2 volte.

Sia X la variabile casuale 'n. di vincite nelle 10 prove indipendenti'.

$$X \sim$$

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.



Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.

2. Nell'ipotesi di ripetere il gioco 10 volte si calcoli la probabilità di vincere almeno 2 volte.

Sia X la variabile casuale 'n. di vincite nelle 10 prove indipendenti'.

$$X \sim \text{Bin}(n = 10, p = \frac{12}{36} = 0.3333).$$

$$P(X \geq 2) =$$



Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.

2. Nell'ipotesi di ripetere il gioco 10 volte si calcoli la probabilità di vincere almeno 2 volte.

Sia X la variabile casuale 'n. di vincite nelle 10 prove indipendenti'.

$$X \sim \text{Bin}(n = 10, p = \frac{12}{36} = 0.3333).$$

$$P(X \geq 2) = 1 - P(X < 2) =$$



2. Nell'ipotesi di ripetere il gioco 10 volte si calcoli la probabilità di vincere almeno 2 volte.

Sia X la variabile casuale 'n. di vincite nelle 10 prove indipendenti'.

$$X \sim \text{Bin}(n = 10, p = \frac{12}{36} = 0.3333).$$

$$\begin{aligned} P(X \geq 2) &= 1 - P(X < 2) = 1 - \{P(X = 0) + P(X = 1)\} = \\ &= 1 - \left\{ \binom{10}{0} 0.3333^0 \cdot 0.6667^{10} + \binom{10}{1} 0.3333^1 \cdot 0.6667^9 \right\} = \\ &= 1 - (0.0173 + 0.0867) = 1 - 0.1040 = 0.8960. \end{aligned}$$



Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.

3. Nell'ipotesi di ripetere il gioco 100 volte si calcoli la probabilità di vincere almeno 40 volte.



3. Nell'ipotesi di ripetere il gioco 100 volte si calcoli la probabilità di vincere almeno 40 volte.

Indicata con X la variabile casuale 'numero di successi' nelle 100 prove indipendenti risulta

$$X \sim$$

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.



3. Nell'ipotesi di ripetere il gioco 100 volte si calcoli la probabilità di vincere almeno 40 volte.

Indicata con X la variabile casuale 'numero di successi' nelle 100 prove indipendenti risulta

$$X \sim \text{Bin}(n = 100, p = 0.3333).$$

$$M(X) = 33.33; \text{Var}(X) = 22.22.$$

Dal momento che $np = 33.33 > 5$ ed $n(1 - p) = 66.67 > 5$ è possibile utilizzare l'approssimazione della variabile casuale Binomiale con la variabile casuale Normale

$$X \sim$$



3. Nell'ipotesi di ripetere il gioco 100 volte si calcoli la probabilità di vincere almeno 40 volte.

Indicata con X la variabile casuale 'numero di successi' nelle 100 prove indipendenti risulta

$$X \sim \text{Bin}(n = 100, p = 0.3333).$$

$$M(X) = 33.33; \text{Var}(X) = 22.22.$$

Dal momento che $np = 33.33 > 5$ ed $n(1 - p) = 66.67 > 5$ è possibile utilizzare l'approssimazione della variabile casuale Binomiale con la variabile casuale Normale

$$X \sim N(\mu = 33.33, \sigma^2 = 22.22).$$

Si ottiene



3. Nell'ipotesi di ripetere il gioco 100 volte si calcoli la probabilità di vincere almeno 40 volte.

Indicata con X la variabile casuale 'numero di successi' nelle 100 prove indipendenti risulta

$$X \sim \text{Bin}(n = 100, p = 0.3333).$$

$$M(X) = 33.33; \text{Var}(X) = 22.22.$$

Dal momento che $np = 33.33 > 5$ ed $n(1 - p) = 66.67 > 5$ è possibile utilizzare l'approssimazione della variabile casuale Binomiale con la variabile casuale Normale

$$X \sim N(\mu = 33.33, \sigma^2 = 22.22).$$

Si ottiene

$$P(X \geq 40) =$$



3. Nell'ipotesi di ripetere il gioco 100 volte si calcoli la probabilità di vincere almeno 40 volte.

Indicata con X la variabile casuale 'numero di successi' nelle 100 prove indipendenti risulta

$$X \sim \text{Bin}(n = 100, p = 0.3333).$$

$$M(X) = 33.33; \text{Var}(X) = 22.22.$$

Dal momento che $np = 33.33 > 5$ ed $n(1 - p) = 66.67 > 5$ è possibile utilizzare l'approssimazione della variabile casuale Binomiale con la variabile casuale Normale

$$X \sim N(\mu = 33.33, \sigma^2 = 22.22).$$

Si ottiene

$$P(X \geq 40) = P\left(\frac{X - \mu}{\sigma} \geq \frac{40 - \mu}{\sigma}\right) =$$



3. Nell'ipotesi di ripetere il gioco 100 volte si calcoli la probabilità di vincere almeno 40 volte.

Indicata con X la variabile casuale 'numero di successi' nelle 100 prove indipendenti risulta

$$X \sim \text{Bin}(n = 100, p = 0.3333).$$

$$M(X) = 33.33; \text{Var}(X) = 22.22.$$

Dal momento che $np = 33.33 > 5$ ed $n(1 - p) = 66.67 > 5$ è possibile utilizzare l'approssimazione della variabile casuale Binomiale con la variabile casuale Normale

$$X \sim N(\mu = 33.33, \sigma^2 = 22.22).$$

Si ottiene

$$\begin{aligned} P(X \geq 40) &= P\left(\frac{X - \mu}{\sigma} \geq \frac{40 - \mu}{\sigma}\right) = P\left(Z \geq \frac{40 - 33.33}{4.7148}\right) = \\ &= \end{aligned}$$



3. Nell'ipotesi di ripetere il gioco 100 volte si calcoli la probabilità di vincere almeno 40 volte.

Indicata con X la variabile casuale 'numero di successi' nelle 100 prove indipendenti risulta

$$X \sim \text{Bin}(n = 100, p = 0.3333).$$

$$M(X) = 33.33; \text{Var}(X) = 22.22.$$

Dal momento che $np = 33.33 > 5$ ed $n(1 - p) = 66.67 > 5$ è possibile utilizzare l'approssimazione della variabile casuale Binomiale con la variabile casuale Normale

$$X \sim N(\mu = 33.33, \sigma^2 = 22.22).$$

Si ottiene

$$\begin{aligned} P(X \geq 40) &= P\left(\frac{X - \mu}{\sigma} \geq \frac{40 - \mu}{\sigma}\right) = P\left(Z \geq \frac{40 - 33.33}{4.7148}\right) = \\ &= P(Z \geq 1.41) = \end{aligned}$$



3. Nell'ipotesi di ripetere il gioco 100 volte si calcoli la probabilità di vincere almeno 40 volte.

Indicata con X la variabile casuale 'numero di successi' nelle 100 prove indipendenti risulta

$$X \sim \text{Bin}(n = 100, p = 0.3333).$$

$$M(X) = 33.33; \text{Var}(X) = 22.22.$$

Dal momento che $np = 33.33 > 5$ ed $n(1 - p) = 66.67 > 5$ è possibile utilizzare l'approssimazione della variabile casuale Binomiale con la variabile casuale Normale

$$X \sim N(\mu = 33.33, \sigma^2 = 22.22).$$

Si ottiene

$$\begin{aligned} P(X \geq 40) &= P\left(\frac{X - \mu}{\sigma} \geq \frac{40 - \mu}{\sigma}\right) = P\left(Z \geq \frac{40 - 33.33}{4.7148}\right) = \\ &= P(Z \geq 1.41) = 1 - P(Z \leq 1.41) = \end{aligned}$$



3. Nell'ipotesi di ripetere il gioco 100 volte si calcoli la probabilità di vincere almeno 40 volte.

Indicata con X la variabile casuale 'numero di successi' nelle 100 prove indipendenti risulta

$$X \sim \text{Bin}(n = 100, p = 0.3333).$$

$$M(X) = 33.33; \text{Var}(X) = 22.22.$$

Dal momento che $np = 33.33 > 5$ ed $n(1 - p) = 66.67 > 5$ è possibile utilizzare l'approssimazione della variabile casuale Binomiale con la variabile casuale Normale

$$X \sim N(\mu = 33.33, \sigma^2 = 22.22).$$

Si ottiene

$$\begin{aligned} P(X \geq 40) &= P\left(\frac{X - \mu}{\sigma} \geq \frac{40 - \mu}{\sigma}\right) = P\left(Z \geq \frac{40 - 33.33}{4.7148}\right) = \\ &= P(Z \geq 1.41) = 1 - P(Z \leq 1.41) = 1 - 0.9207 = 0.0793. \end{aligned}$$

Tema T290

Prova scritta di STATISTICA del 08.09.2016

Prova scritta di
STATISTICA del
08.09.2016

Giuseppe Boari
Gabriele Cantaluppi



UNIVERSITÀ
CATTOLICA
del Sacro Cuore

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.

Giuseppe Boari
Gabriele Cantaluppi
Università Cattolica del Sacro Cuore, Milano

Esercizio 1

Prova scritta di
STATISTICA del
08.09.2016

Giuseppe Boari
Gabriele Cantaluppi



UNIVERSITÀ
CATTOLICA
del Sacro Cuore

La tabella seguente raccoglie i dati relativi a 91 nazioni, sul reddito pro-capite X (in migliaia di dollari) e tasso di mortalità infantile Y (n. decessi ogni 1000 nati vivi):

$X \setminus Y$	0 - 10	10 - 40	40 - 100
0 - 10	1	16	26
10 - 30	10	18	3
30 - 100	15	2	0

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.



1. Rappresentare graficamente la marginale X .

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.



Si considera la distribuzione marginale della variabile X , ottenuta sommando le frequenze di riga

x_i	$n_{i\bullet}$
0 $\vdash$ 10	43
10 $\vdash$ 30	31
30 $\vdash$ 100	17

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.



Si tratta di un carattere quantitativo con modalità raggruppate in classi che può essere rappresentato graficamente mediante un istogramma.

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.



L'istogramma può essere costruito dopo avere calcolato le densità $d_{i\bullet} = \frac{n_{i\bullet}}{a_i}$, (dove a_i sono le ampiezze di classe).

x_i	$n_{i\bullet}$	$a_{i\bullet}$	$d_{i\bullet}$
0 ÷ 10	43	10 - 0 = 10	4.3
10 ÷ 30	31	30 - 10 = 20	1.55
30 ÷ 100	17	100 - 30 = 70	0.2429

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.

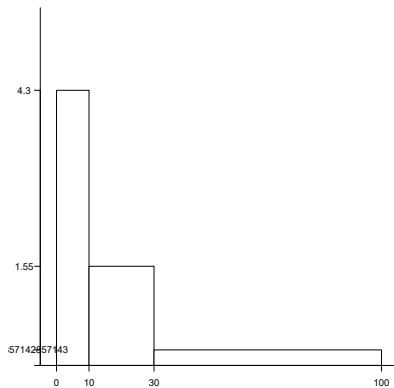


Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.





2. Confrontare la variabilità delle marginali X e Y .

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.



Si considera la distribuzione marginale della variabile Y , ottenuta sommando le frequenze di colonna

y_j	$n_{\bullet j}$
0 - 10	43
10 - 30	31
30 - 100	17

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.



Si tratta di un carattere quantitativo con modalità raggruppate in classi.

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.



La media aritmetica

$$M(Y) = \frac{1}{n} \sum_{j=1}^k y_j n_{\bullet j}$$

risulta 21.3187.

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.



La media aritmetica

$$M(Y) = \frac{1}{n} \sum_{j=1}^k y_j n_{\bullet j}$$

risulta 21.3187.

Si ottiene (considerando come modalità i valori centrali delle classi)

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.



La media aritmetica

$$M(Y) = \frac{1}{n} \sum_{j=1}^k y_j n_{\bullet j}$$

risulta 21.3187.

Si ottiene (considerando come modalità i valori centrali delle classi)

y_j	$n_{\bullet j}$	$y_j n_{\bullet j}$
5	43	215
20	31	620
65	17	1105
	91	1940

da cui segue

$$M(Y) = \frac{1}{n} \sum_{i=1}^k y_j n_{\bullet j} = \frac{1940}{91} = 21.3187.$$

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.



La varianza

$$\text{Var}(Y) = M[(Y - \mu_Y)^2] = M(Y^2) - \mu_Y^2 = \frac{1}{n} \sum_{j=1}^k y_j^2 n_{\bullet j} - \mu_Y^2$$

risulta 482.8765.

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.



La varianza

$$\text{Var}(Y) = M[(Y - \mu_Y)^2] = M(Y^2) - \mu_Y^2 = \frac{1}{n} \sum_{j=1}^k y_j^2 n_{\bullet j} - \mu_Y^2$$

risulta 482.8765.

Si ottiene (considerando come modalità i valori centrali delle classi)

y_j	$n_{\bullet j}$	$y_j n_{\bullet j}$	$y_j^2 n_{\bullet j}$
5	43	215	1075
20	31	620	12400
65	17	1105	71825
91	1940	85300	

da cui segue

$$\begin{aligned} \text{Var}(Y) = \sigma_Y^2 &= M(Y^2) - \mu_Y^2 = \\ &= \frac{85300}{91} - 21.3187^2 = 937.3626 - 454.4862 = 482.8764. \end{aligned}$$

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.



Per effettuare confronti in termini di variabilità tra caratteri quantitativi occorre utilizzare una misura relativa, ad esempio il coefficiente di variazione

$$CV = \frac{\sigma}{\mu} = \frac{\sqrt{\text{Var}(Y)}}{M(Y)} = \frac{\sqrt{482.8764}}{21.3187} = \frac{21.9744}{21.3187} = 1.0308.$$

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.



Si considera la distribuzione marginale della variabile Y , ottenuta sommando le frequenze di colonna

y_j	$n_{\bullet j}$
0 $\vdash$ 10	26
10 $\vdash$ 40	36
40 $\vdash$ 100	29

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.



Si tratta di un carattere quantitativo con modalità raggruppate in classi.

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.



La media aritmetica

$$M(Y) = \frac{1}{n} \sum_{j=1}^k y_j n_{\bullet j}$$

risulta 33.6264.

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.



La media aritmetica

$$M(Y) = \frac{1}{n} \sum_{j=1}^k y_j n_{\bullet j}$$

risulta 33.6264.

Si ottiene (considerando come modalità i valori centrali delle classi)

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.



La media aritmetica

$$M(Y) = \frac{1}{n} \sum_{j=1}^k y_j n_{\bullet j}$$

risulta 33.6264.

Si ottiene (considerando come modalità i valori centrali delle classi)

y_j	$n_{\bullet j}$	$y_j n_{\bullet j}$
5	26	130
25	36	900
70	29	2030
	91	3060

da cui segue

$$M(Y) = \frac{1}{n} \sum_{i=1}^k y_j n_{\bullet j} = \frac{3060}{91} = 33.6264.$$

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.



La varianza

$$\text{Var}(Y) = M[(Y - \mu_Y)^2] = M(Y^2) - \mu_Y^2 = \frac{1}{n} \sum_{j=1}^k y_j^2 n_{\bullet j} - \mu_Y^2$$

risulta 685.2011.

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.



La varianza

$$Var(Y) = M[(Y - \mu_Y)^2] = M(Y^2) - \mu_Y^2 = \frac{1}{n} \sum_{j=1}^k y_j^2 n_{\bullet j} - \mu_Y^2$$

risulta 685.2011.

Si ottiene (considerando come modalità i valori centrali delle classi)

y_j	$n_{\bullet j}$	$y_j n_{\bullet j}$	$y_j^2 n_{\bullet j}$
5	26	130	650
25	36	900	22500
70	29	2030	142100
	91	3060	165250

da cui segue

$$\begin{aligned} Var(Y) = \sigma_Y^2 &= M(Y^2) - \mu_Y^2 = \\ &= \frac{165250}{91} - 33.6264^2 = 1815.9341 - 1130.733 = 685.2011. \end{aligned}$$

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.



Per effettuare confronti in termini di variabilità tra caratteri quantitativi occorre utilizzare una misura relativa, ad esempio il coefficiente di variazione

$$CV = \frac{\sigma}{\mu} = \frac{\sqrt{\text{Var}(Y)}}{M(Y)} = \frac{\sqrt{685.2011}}{33.6264} = \frac{26.1763}{33.6264} = 0.7784.$$

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.



Riepilogo

Indice per confronto variabilità: coefficiente di variazione

valore X 1.0308 valore Y 0.7784

Il reddito pro-capite presenta un livello di variabilità superiore a quello del tasso di mortalità infantile.

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.



3. Valutare la connessione tra X e Y con un opportuno indice.

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.



Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.

Dalla distribuzione delle frequenze congiunte n_{ij} :

	0 - 10	10 - 40	40 - 100	$n_{i\bullet}$
0 - 10	1	16	26	43
10 - 30	10	18	3	31
30 - 100	15	2	0	17
$n_{\bullet j}$	26	36	29	91

è possibile ottenere le frequenze teoriche nella situazione di indipendenza stocastica

$$\hat{n}_{ij} = \frac{n_{i\bullet} n_{\bullet j}}{n}:$$



Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.

Dalla distribuzione delle frequenze congiunte n_{ij} :

	0 ÷ 10	10 ÷ 40	40 ÷ 100	$n_{i\bullet}$
0 ÷ 10	1	16	26	43
10 ÷ 30	10	18	3	31
30 ÷ 100	15	2	0	17
$n_{\bullet j}$	26	36	29	91

è possibile ottenere le frequenze teoriche nella situazione di indipendenza stocastica

$$\hat{n}_{ij} = \frac{n_{i\bullet} \cdot n_{\bullet j}}{n}:$$

	0 ÷ 10	10 ÷ 40	40 ÷ 100	$n_{i\bullet}$
0 ÷ 10	12.2857	17.0110	13.7033	43
10 ÷ 30	8.8571	12.2637	9.8791	31
30 ÷ 100	4.8571	6.7253	5.4176	17
$n_{\bullet j}$	26.0000	36.0000	29.0000	91



le contingenze assolute $n_{ij} - \hat{n}_{ij}$:

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.



le contingenze assolute $n_{ij} - \hat{n}_{ij}$:

	0 ÷ 10	10 ÷ 40	40 ÷ 100
0 ÷ 10	-11.2857	-1.0110	12.2967
10 ÷ 30	1.1429	5.7363	-6.8791
30 ÷ 100	10.1429	-4.7253	-5.4176

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.



i rapporti di contingenza $\frac{n_{ij}}{\hat{n}_{ij}}$:

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.



i rapporti di contingenza $\frac{n_{ij}}{\hat{n}_{ij}}$:

	0 - 10	10 - 40	40 - 100
0 - 10	0.0814	0.9406	1.8974
10 - 30	1.1290	1.4677	0.3037
30 - 100	3.0882	0.2974	0.0000

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.



e le contingenze relative, espresse in termini percentuali, $100 \left(\frac{n_{ij} - \hat{n}_{ij}}{\hat{n}_{ij}} \right) \% = 100 \left(\frac{n_{ij}}{\hat{n}_{ij}} - 1 \right) \%$:

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.



e le contingenze relative, espresse in termini percentuali, $100 \left(\frac{n_{ij} - \hat{n}_{ij}}{\hat{n}_{ij}} \right) \% = 100 \left(\frac{n_{ij}}{\hat{n}_{ij}} - 1 \right) \%$:

	0 ÷ 10	10 ÷ 40	40 ÷ 100
0 ÷ 10	-91.86	-5.94	89.74
10 ÷ 30	12.90	46.77	-69.63
30 ÷ 100	208.82	-70.26	-100.00

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.



Risultati intermedi $\frac{(n_{ij} - \hat{n}_{ij})^2}{\hat{n}_{ij}}$ per il calcolo del χ^2

	0 ÷ 10	10 ÷ 40	40 ÷ 100
0 ÷ 10	10.3671	0.0601	11.0345
10 ÷ 30	0.1475	2.6831	4.7901
30 ÷ 100	21.1807	3.3200	5.4176

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.



Risultati intermedi $\frac{n_{ij}^2}{n_{i\bullet} \cdot n_{\bullet j}}$ per il calcolo del χ^2 (formula operativa)

	0 ÷ 10	10 ÷ 40	40 ÷ 100
0 ÷ 10	0.0009	0.1654	0.5421
10 ÷ 30	0.1241	0.2903	0.0100
30 ÷ 100	0.5090	0.0065	0.0000

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.



Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.

$$\chi^2 = \sum_{i=1}^h \sum_{j=1}^k \frac{(n_{ij} - \hat{n}_{ij})^2}{\hat{n}_{ij}} = 59.0007,$$

$$\chi_N^2 = \frac{\chi^2}{n \cdot \min(h-1, k-1)} = 0.3242$$

$$\chi_N = \sqrt{\chi_N^2} = 0.5694.$$

Sussiste, quindi, un livello di connessione medio.

Esercizio 2



A partire dalla serie storica delle vendite, in n. pezzi, $Y(i)$, $i = 1, 2, \dots, 5$, di un certo prodotto in 5 successivi anni, si è costruita la seguente tabella, dove $\Delta_i = Y(i) - Y(i - 1)$, mentre ${}_1I_i$ sono i numeri indici con base $i = 1$:

i	Δ_i	${}_1I_i$
1	-	1
2	697	1.043
3	696	1.086
4	211	1.099
5	600	1.136

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.



1. Fornire il significato di ${}_1I_2$, numero indice riferito al tempo 2 con base 1.

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.



1. Fornire il significato di ${}_1I_2$, numero indice riferito al tempo 2 con base 1.

${}_1I_2 = 1.043$ Tra il tempo 1 e il tempo 2 si è registrato un incremento del

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.



1. Fornire il significato di ${}_1I_2$, numero indice riferito al tempo 2 con base 1.

${}_1I_2 = 1.043$ Tra il tempo 1 e il tempo 2 si è registrato un incremento del 4.3%.

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.



Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.

1. Fornire il significato di ${}_1I_2$, numero indice riferito al tempo 2 con base 1.

${}_1I_2 = 1.043$ Tra il tempo 1 e il tempo 2 si è registrato un incremento del 4.3%.

Dalla prima colonna si evince che all'incremento percentuale in oggetto corrisponde una variazione assoluta pari a 697.



2. Calcolare, mostrandone il procedimento, la serie dei NIBM.

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.



2. Calcolare, mostrandone il procedimento, la serie dei NIBM.

Si ricorda che ${}_t I_t = \frac{x_t}{x_{t-1}} = \frac{{}_1 I_t}{{}_1 I_{t-1}}$.

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.



2. Calcolare, mostrandone il procedimento, la serie dei NIBM.

Si ricorda che ${}_t I_t = \frac{x_t}{x_{t-1}} = \frac{{}_1 I_t}{{}_1 I_{t-1}}$. Si ottiene quindi

<i>tempo</i>	<i>NIBF</i>	<i>NIBM</i>
1	1.000	
2	1.043	1.0430
3	1.086	1.0412
4	1.099	1.0120
5	1.136	1.0337

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.



3. Calcolare il valore di $Y(1)$.

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.



Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.

3. Calcolare il valore di $Y(1)$.

La variazione relativa e la variazione assoluta tra il tempo 1 e il tempo 2 risulta:

$$\frac{Y_2}{Y_1} - 1 = \frac{Y_2 - Y_1}{Y_1} = 0.043$$

Considerando il dato relativo alla variazione assoluta (presente nella prima colonna)

$$Y_2 - Y_1 = 697$$

abbiamo

$$Y_1 =$$



Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.

3. Calcolare il valore di $Y(1)$.

La variazione relativa e la variazione assoluta tra il tempo 1 e il tempo 2 risulta:

$$\frac{Y_2}{Y_1} - 1 = \frac{Y_2 - Y_1}{Y_1} = 0.043$$

Considerando il dato relativo alla variazione assoluta (presente nella prima colonna)

$$Y_2 - Y_1 = 697$$

abbiamo

$$Y_1 = \frac{Y_2 - Y_1}{0.043} = \frac{697}{0.043} = 16209.3023.$$



4. Calcolare il tasso medio di variazione annuo tra il secondo ($i = 2$) e il quinto anno ($i = 5$), interpretando il risultato.

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.



4. Calcolare il tasso medio di variazione annuo tra il secondo ($i = 2$) e il quinto anno ($i = 5$), interpretando il risultato.

Si ricorda che

$${}_2I_5 = {}_2I_3 \cdot {}_3I_4 \cdot {}_4I_5$$

Si deve trovare il numero indice a base mobile medio.

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.



Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.

4. Calcolare il tasso medio di variazione annuo tra il secondo ($i = 2$) e il quinto anno ($i = 5$), interpretando il risultato.

Si ricorda che

$${}_2l_5 = {}_2l_3 \cdot {}_3l_4 \cdot {}_4l_5$$

Si deve trovare il numero indice a base mobile medio. Applicando il criterio di scelta della media secondo Chisini si ottiene

$${}_2l_5 = {}_2l_3 \cdot {}_3l_4 \cdot {}_4l_5 = \alpha \cdot \alpha \cdot \alpha = \alpha^3.$$



Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.

4. Calcolare il tasso medio di variazione annuo tra il secondo ($i = 2$) e il quinto anno ($i = 5$), interpretando il risultato.

Si ricorda che

$${}_2I_5 = {}_2I_3 \cdot {}_3I_4 \cdot {}_4I_5$$

Si deve trovare il numero indice a base mobile medio. Applicando il criterio di scelta della media secondo Chisini si ottiene

$${}_2I_5 = {}_2I_3 \cdot {}_3I_4 \cdot {}_4I_5 = \alpha \cdot \alpha \cdot \alpha = \alpha^3.$$

Quindi

$$\alpha = \sqrt[3]{{}_2I_5} = \sqrt[3]{{}_2I_3 \cdot {}_3I_4 \cdot {}_4I_5} = 1.0289.$$

Da cui segue una variazione relativa media pari a 0.0289, corrispondente a un incremento medio 2 e il 5 del 2.89%.



Con riferimento allo studio della dipendenza tra due caratteri quantitativi X e Y , si sono raccolte le seguenti 8 coppie di osservazioni (x_i, y_i) $i = 1, 2, \dots, 8$

x_i	5	5	10	10.0	15	15.0	20	20
y_i	2	4	1	2.8	1	0.5	1	2

1. Si rappresentino graficamente le coppie dei dati ipotizzando Y in funzione di X .



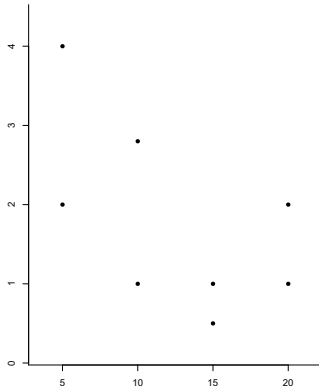
Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.

Le coppie di dati possono essere rappresentate mediante un grafico a dispersione.





2. Si calcolino, secondo il metodo dei minimi quadrati, i parametri dei seguenti modelli di regressione:

$$I) Y^* = a + b \cdot X$$

$$II) Y^* = c + d/X$$

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.

Si riportano i calcoli necessari¹ per ottenere le stime dei parametri e gli indici di adattamento del modello

$$Y = a + bX + E$$

x_i	y_i	$x_i y_i$	x_i^2	y_i^2
5	2	10	25	4
5	4	20	25	16
10	1	10	100	1
10	2.8	28	100	7.84
15	1	15	225	1
15	0.5	7.5	225	0.25
20	1	20	400	1
20	2	40	400	4
100	14.3	150.5	1500	35.09
12.5	1.7875	18.8125	187.5	4.3862
$M(X)$	$M(Y)$	$M(XY)$	$M(X^2)$	$M(Y^2)$

$$\hat{b} =$$



Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.

¹Si osserva che per ragioni di stampa alcuni calcoli potrebbero essere duplicati con riferimento ai diversi modelli oggetto di studio

Si riportano i calcoli necessari¹ per ottenere le stime dei parametri e gli indici di adattamento del modello

$$Y = a + bX + E$$

x_i	y_i	$x_i y_i$	x_i^2	y_i^2
5	2	10	25	4
5	4	20	25	16
10	1	10	100	1
10	2.8	28	100	7.84
15	1	15	225	1
15	0.5	7.5	225	0.25
20	1	20	400	1
20	2	40	400	4
100	14.3	150.5	1500	35.09
12.5	1.7875	18.8125	187.5	4.3862
$M(X)$	$M(Y)$	$M(XY)$	$M(X^2)$	$M(Y^2)$

$$\begin{aligned}\hat{b} &= \frac{\text{Cov}(X, Y)}{\text{Var}(X)} = \frac{M(XY) - M(X)M(Y)}{M(X^2) - [M(X)]^2} = \\ &= \frac{18.8125 - 12.5 \cdot 1.7875}{187.5 - 12.5^2} = \frac{-3.5313}{31.25} = -0.113 \\ \hat{a} &= \end{aligned}$$

¹Si osserva che per ragioni di stampa alcuni calcoli potrebbero essere duplicati con riferimento ai diversi modelli oggetto di studio



Si riportano i calcoli necessari¹ per ottenere le stime dei parametri e gli indici di adattamento del modello

$$Y = a + bX + E$$

x_i	y_i	$x_i y_i$	x_i^2	y_i^2
5	2	10	25	4
5	4	20	25	16
10	1	10	100	1
10	2.8	28	100	7.84
15	1	15	225	1
15	0.5	7.5	225	0.25
20	1	20	400	1
20	2	40	400	4
100	14.3	150.5	1500	35.09
12.5	1.7875	18.8125	187.5	4.3862
$M(X)$	$M(Y)$	$M(XY)$	$M(X^2)$	$M(Y^2)$

$$\begin{aligned}\hat{b} &= \frac{\text{Cov}(X, Y)}{\text{Var}(X)} = \frac{M(XY) - M(X)M(Y)}{M(X^2) - [M(X)]^2} = \\ &= \frac{18.8125 - 12.5 \cdot 1.7875}{187.5 - 12.5^2} = \frac{-3.5313}{31.25} = -0.113\end{aligned}$$

$$\hat{a} = M(Y) - \hat{b}M(X) = 3.2$$

$$\rho^2 =$$

¹Si osserva che per ragioni di stampa alcuni calcoli potrebbero essere duplicati con riferimento ai diversi modelli oggetto di studio



Si riportano i calcoli necessari¹ per ottenere le stime dei parametri e gli indici di adattamento del modello

$$Y = a + bX + E$$

x_i	y_i	$x_i y_i$	x_i^2	y_i^2
5	2	10	25	4
5	4	20	25	16
10	1	10	100	1
10	2.8	28	100	7.84
15	1	15	225	1
15	0.5	7.5	225	0.25
20	1	20	400	1
20	2	40	400	4
100	14.3	150.5	1500	35.09
12.5	1.7875	18.8125	187.5	4.3862
$M(X)$	$M(Y)$	$M(XY)$	$M(X^2)$	$M(Y^2)$

$$\begin{aligned}\hat{b} &= \frac{\text{Cov}(X, Y)}{\text{Var}(X)} = \frac{M(XY) - M(X)M(Y)}{M(X^2) - [M(X)]^2} = \\ &= \frac{18.8125 - 12.5 \cdot 1.7875}{187.5 - 12.5^2} = \frac{-3.5313}{31.25} = -0.113\end{aligned}$$

$$\hat{a} = M(Y) - \hat{b} M(X) = 3.2$$

$$\rho^2 = \frac{[\text{Cov}(X, Y)]^2}{\text{Var}(X) \text{Var}(Y)} = \frac{(-3.5313)^2}{31.25 \cdot 1.1911} = 0.335$$

$$M(\hat{E}^2) = \text{Var}(\hat{E}) = \text{varianza residua} =$$

¹Si osserva che per ragioni di stampa alcuni calcoli potrebbero essere duplicati con riferimento ai diversi modelli oggetto di studio



Si riportano i calcoli necessari¹ per ottenere le stime dei parametri e gli indici di adattamento del modello

$$Y = a + bX + E$$

x_i	y_i	$x_i y_i$	x_i^2	y_i^2
5	2	10	25	4
5	4	20	25	16
10	1	10	100	1
10	2.8	28	100	7.84
15	1	15	225	1
15	0.5	7.5	225	0.25
20	1	20	400	1
20	2	40	400	4
100	14.3	150.5	1500	35.09
12.5	1.7875	18.8125	187.5	4.3862
$M(X)$	$M(Y)$	$M(XY)$	$M(X^2)$	$M(Y^2)$

$$\begin{aligned}\hat{b} &= \frac{\text{Cov}(X, Y)}{\text{Var}(X)} = \frac{M(XY) - M(X)M(Y)}{M(X^2) - [M(X)]^2} = \\ &= \frac{18.8125 - 12.5 \cdot 1.7875}{187.5 - 12.5^2} = \frac{-3.5313}{31.25} = -0.113\end{aligned}$$

$$\hat{a} = M(Y) - \hat{b} M(X) = 3.2$$

$$\rho^2 = \frac{[\text{Cov}(X, Y)]^2}{\text{Var}(X) \text{Var}(Y)} = \frac{(-3.5313)^2}{31.25 \cdot 1.1911} = 0.335$$

$$M(\hat{E}^2) = \text{Var}(\hat{E}) = \text{varianza residua} = \text{Var}(Y) \cdot (1 - \rho^2) = 0.7921$$

¹Si osserva che per ragioni di stampa alcuni calcoli potrebbero essere duplicati con riferimento ai diversi modelli oggetto di studio



Si riportano i calcoli necessari² per ottenere le stime dei parametri e gli indici di adattamento del modello

$$Y = c + d1/X + E$$

$1/x_i$	y_i	$1/x_i y_i$	$1/x_i^2$	y_i^2
0.2	2	0.4	0.04	4
0.2	4	0.8	0.04	16
0.1	1	0.1	0.01	1
0.1	2.8	0.28	0.01	7.84
0.0667	1	0.0667	0.0044	1
0.0667	0.5	0.0333	0.0044	0.25
0.05	1	0.05	0.0025	1
0.05	2	0.1	0.0025	4
0.8333	14.3	1.83	0.1139	35.09
0.1042	1.7875	0.2288	0.0142	4.3862
$M(1/X)$	$M(Y)$	$M(1/XY)$	$M(1/X^2)$	$M(Y^2)$

$$\hat{d} =$$



Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.

²Si osserva che per ragioni di stampa alcuni calcoli potrebbero essere duplicati con riferimento ai diversi modelli oggetto di studio

Si riportano i calcoli necessari² per ottenere le stime dei parametri e gli indici di adattamento del modello

$$Y = c + d1/X + E$$

$1/x_i$	y_i	$1/x_i y_i$	$1/x_i^2$	y_i^2
0.2	2	0.4	0.04	4
0.2	4	0.8	0.04	16
0.1	1	0.1	0.01	1
0.1	2.8	0.28	0.01	7.84
0.0667	1	0.0667	0.0044	1
0.0667	0.5	0.0333	0.0044	0.25
0.05	1	0.05	0.0025	1
0.05	2	0.1	0.0025	4
0.8333	14.3	1.83	0.1139	35.09
0.1042	1.7875	0.2288	0.0142	4.3862
$M(1/X)$	$M(Y)$	$M(1/XY)$	$M(1/X^2)$	$M(Y^2)$

$$\begin{aligned}\hat{d} &= \frac{\text{Cov}(1/X, Y)}{\text{Var}(1/X)} = \frac{M(1/XY) - M(1/X)M(Y)}{M(1/X^2) - [M(1/X)]^2} = \\ &= \frac{0.2288 - 0.1042 \cdot 1.7875}{0.0142 - 0.1042^2} = \frac{0.0426}{0.0034} = 12.5702 \\ \hat{c} &= \end{aligned}$$

²Si osserva che per ragioni di stampa alcuni calcoli potrebbero essere duplicati con riferimento ai diversi modelli oggetto di studio



Si riportano i calcoli necessari² per ottenere le stime dei parametri e gli indici di adattamento del modello

$$Y = c + d1/X + E$$

$1/x_i$	y_i	$1/x_i y_i$	$1/x_i^2$	y_i^2
0.2	2	0.4	0.04	4
0.2	4	0.8	0.04	16
0.1	1	0.1	0.01	1
0.1	2.8	0.28	0.01	7.84
0.0667	1	0.0667	0.0044	1
0.0667	0.5	0.0333	0.0044	0.25
0.05	1	0.05	0.0025	1
0.05	2	0.1	0.0025	4
0.8333	14.3	1.83	0.1139	35.09
0.1042	1.7875	0.2288	0.0142	4.3862
$M(1/X)$	$M(Y)$	$M(1/XY)$	$M(1/X^2)$	$M(Y^2)$

$$\begin{aligned}\hat{d} &= \frac{\text{Cov}(1/X, Y)}{\text{Var}(1/X)} = \frac{M(1/XY) - M(1/X)M(Y)}{M(1/X^2) - [M(1/X)]^2} = \\ &= \frac{0.2288 - 0.1042 \cdot 1.7875}{0.0142 - 0.1042^2} = \frac{0.0426}{0.0034} = 12.5702\end{aligned}$$

$$\hat{c} = M(Y) - \hat{d} M(1/X) = 0.4781$$

$$\rho^2 =$$



²Si osserva che per ragioni di stampa alcuni calcoli potrebbero essere duplicati con riferimento ai diversi modelli oggetto di studio

Si riportano i calcoli necessari² per ottenere le stime dei parametri e gli indici di adattamento del modello

$$Y = c + d1/X + E$$

$1/x_i$	y_i	$1/x_i y_i$	$1/x_i^2$	y_i^2
0.2	2	0.4	0.04	4
0.2	4	0.8	0.04	16
0.1	1	0.1	0.01	1
0.1	2.8	0.28	0.01	7.84
0.0667	1	0.0667	0.0044	1
0.0667	0.5	0.0333	0.0044	0.25
0.05	1	0.05	0.0025	1
0.05	2	0.1	0.0025	4
0.8333	14.3	1.83	0.1139	35.09
0.1042	1.7875	0.2288	0.0142	4.3862
$M(1/X)$	$M(Y)$	$M(1/XY)$	$M(1/X^2)$	$M(Y^2)$

$$\begin{aligned}\hat{d} &= \frac{\text{Cov}(1/X, Y)}{\text{Var}(1/X)} = \frac{M(1/XY) - M(1/X)M(Y)}{M(1/X^2) - [M(1/X)]^2} = \\ &= \frac{0.2288 - 0.1042 \cdot 1.7875}{0.0142 - 0.1042^2} = \frac{0.0426}{0.0034} = 12.5702\end{aligned}$$

$$\hat{c} = M(Y) - \hat{d} M(1/X) = 0.4781$$

$$\rho^2 = \frac{[\text{Cov}(1/X, Y)]^2}{\text{Var}(1/X) \text{Var}(Y)} = \frac{0.0426^2}{0.0034 \cdot 1.1911} = 0.4492$$

$$M(\hat{E}^2) = \text{Var}(\hat{E}) = \text{varianza residua} =$$

²Si osserva che per ragioni di stampa alcuni calcoli potrebbero essere duplicati con riferimento ai diversi modelli oggetto di studio



Si riportano i calcoli necessari² per ottenere le stime dei parametri e gli indici di adattamento del modello

$$Y = c + d1/X + E$$

$1/x_i$	y_i	$1/x_i y_i$	$1/x_i^2$	y_i^2
0.2	2	0.4	0.04	4
0.2	4	0.8	0.04	16
0.1	1	0.1	0.01	1
0.1	2.8	0.28	0.01	7.84
0.0667	1	0.0667	0.0044	1
0.0667	0.5	0.0333	0.0044	0.25
0.05	1	0.05	0.0025	1
0.05	2	0.1	0.0025	4
0.8333	14.3	1.83	0.1139	35.09
0.1042	1.7875	0.2288	0.0142	4.3862
$M(1/X)$	$M(Y)$	$M(1/XY)$	$M(1/X^2)$	$M(Y^2)$

$$\begin{aligned}\hat{d} &= \frac{\text{Cov}(1/X, Y)}{\text{Var}(1/X)} = \frac{M(1/XY) - M(1/X)M(Y)}{M(1/X^2) - [M(1/X)]^2} = \\ &= \frac{0.2288 - 0.1042 \cdot 1.7875}{0.0142 - 0.1042^2} = \frac{0.0426}{0.0034} = 12.5702\end{aligned}$$

$$\hat{c} = M(Y) - \hat{d} M(1/X) = 0.4781$$

$$\rho^2 = \frac{[\text{Cov}(1/X, Y)]^2}{\text{Var}(1/X) \text{Var}(Y)} = \frac{0.0426^2}{0.0034 \cdot 1.1911} = 0.4492$$

$$M(\hat{E}^2) = \text{Var}(\hat{E}) = \text{varianza residua} = \text{Var}(Y) \cdot (1 - \rho^2) = 0.6561$$

²Si osserva che per ragioni di stampa alcuni calcoli potrebbero essere duplicati con riferimento ai diversi modelli oggetto di studio





Riepilogo

$\hat{a} = 3.2$, $\hat{b} = -0.1130$, indice adattamento $\rho^2 = 0.3350$ $\hat{c} = 0.4782$, $\hat{d} = 12.5692$, indice adattamento $R^2 = 0.4490$

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.



3. Confrontare gli indici di adattamento, commentando i risultati.

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.



3. Confrontare gli indici di adattamento, commentando i risultati.

Il modello migliore è quello caratterizzato dal massimo valore dell'indice di adattamento:

$$Y^* = 0.4782 + 12.5692/X.$$

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.



Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.

4. In base al modello migliore fornire il valore teorico di Y per $X = 30$.



Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.

4. In base al modello migliore fornire il valore teorico di Y per $X = 30$.

$$Y^*(30) = 0.4782 + 12.5692/30 = 0.8972$$

Esercizio 4

Prova scritta di
STATISTICA del
08.09.2016

Giuseppe Boari
Gabriele Cantaluppi



UNIVERSITÀ
CATTOLICA
del Sacro Cuore

Si lanciano 2 dadi equilibrati, il primo regolare (con simboli 1, 2, 3, 4, 5, 6) e il secondo con simboli di valore 1, 2, 0, 0, 0, 0.

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.



1. Descrivere (specificando valori e probabilità) le variabili casuali $X_1 =$ 'risultato del dado 1' e $X_2 =$ 'risultato del dado 2'.

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.



1. Descrivere (specificando valori e probabilità) le variabili casuali $X_1 = \text{'risultato del dado 1'}$ e $X_2 = \text{'risultato del dado 2'}$.

Abbiamo

$$\frac{x_{1i} \quad P(X_1 = x_{1i})}{1}$$

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.



1. Descrivere (specificando valori e probabilità) le variabili casuali $X_1 = \text{'risultato del dado 1'}$ e $X_2 = \text{'risultato del dado 2'}$.

Abbiamo

x_{1i}	$P(X_1 = x_{1i})$
1	1/6
2	

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.



1. Descrivere (specificando valori e probabilità) le variabili casuali $X_1 = \text{'risultato del dado 1'}$ e $X_2 = \text{'risultato del dado 2'}$.

Abbiamo

x_{1i}	$P(X_1 = x_{1i})$
1	1/6
2	1/6
3	

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.



1. Descrivere (specificando valori e probabilità) le variabili casuali $X_1 = \text{'risultato del dado 1'}$ e $X_2 = \text{'risultato del dado 2'}$.

Abbiamo

x_{1i}	$P(X_1 = x_{1i})$
1	1/6
2	1/6
3	1/6
4	

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.



1. Descrivere (specificando valori e probabilità) le variabili casuali $X_1 =$ 'risultato del dado 1' e $X_2 =$ 'risultato del dado 2'.

Abbiamo

x_{1i}	$P(X_1 = x_{1i})$
1	1/6
2	1/6
3	1/6
4	1/6
5	

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.



1. Descrivere (specificando valori e probabilità) le variabili casuali $X_1 =$ 'risultato del dado 1' e $X_2 =$ 'risultato del dado 2'.

Abbiamo

x_{1i}	$P(X_1 = x_{1i})$
1	1/6
2	1/6
3	1/6
4	1/6
5	1/6
6	

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.



Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.

1. Descrivere (specificando valori e probabilità) le variabili casuali $X_1 = \text{'risultato del dado 1'}$ e $X_2 = \text{'risultato del dado 2'}$.

Abbiamo

x_{1i}	$P(X_1 = x_{1i})$
1	1/6
2	1/6
3	1/6
4	1/6
5	1/6
6	1/6
	1

e

x_{2i}	$P(X_2 = x_{2i})$
0	



1. Descrivere (specificando valori e probabilità) le variabili casuali $X_1 = \text{'risultato del dado 1'}$ e $X_2 = \text{'risultato del dado 2'}$.

Abbiamo

x_{1i}	$P(X_1 = x_{1i})$
1	1/6
2	1/6
3	1/6
4	1/6
5	1/6
6	1/6
	1

e

x_{2i}	$P(X_2 = x_{2i})$
0	4/6
1	

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.



1. Descrivere (specificando valori e probabilità) le variabili casuali $X_1 = \text{'risultato del dado 1'}$ e $X_2 = \text{'risultato del dado 2'}$.

Abbiamo

x_{1i}	$P(X_1 = x_{1i})$
1	1/6
2	1/6
3	1/6
4	1/6
5	1/6
6	1/6
	1

e

x_{2i}	$P(X_2 = x_{2i})$
0	4/6
1	1/6
2	

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.



1. Descrivere (specificando valori e probabilità) le variabili casuali $X_1 = \text{'risultato del dado 1'}$ e $X_2 = \text{'risultato del dado 2'}$.

Abbiamo

x_{1i}	$P(X_1 = x_{1i})$
1	1/6
2	1/6
3	1/6
4	1/6
5	1/6
6	1/6
	1

e

x_{2i}	$P(X_2 = x_{2i})$
0	4/6
1	1/6
2	1/6
	1

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.

2. Calcolare $M(X_1 + X_2)$.



Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.

2. Calcolare $M(X_1 + X_2)$.

Vale

$$M(X_1) =$$



Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.

2. Calcolare $M(X_1 + X_2)$.

Vale

$$M(X_1) = 3.5 \quad M(X_2) =$$



Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.

2. Calcolare $M(X_1 + X_2)$.

Vale

$$M(X_1) = 3.5 \quad M(X_2) = 0.5$$

Ricordando la definizione di valore atteso di una variabile casuale abbiamo, infatti



Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.

2. Calcolare $M(X_1 + X_2)$.

Vale

$$M(X_1) = 3.5 \quad M(X_2) = 0.5$$

Ricordando la definizione di valore atteso di una variabile casuale abbiamo, infatti

x_{1i}	$P(X_1 = x_{1i})$	$x_{1i} \cdot P(X_1 = x_{1i})$
1	1/6	1/6
2	1/6	2/6
3	1/6	3/6
4	1/6	4/6
5	1/6	5/6
6	1/6	6/6
1		21/6

e

x_{2i}	$P(X_2 = x_{2i})$	$x_{2i} \cdot P(X_2 = x_{2i})$
0	4/6	0
1	1/6	1/6
2	1/6	2/6
1		3/6



Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.

2. Calcolare $M(X_1 + X_2)$.

Vale

$$M(X_1) = 3.5 \quad M(X_2) = 0.5$$

Ricordando la definizione di valore atteso di una variabile casuale abbiamo, infatti

x_{1i}	$P(X_1 = x_{1i})$	$x_{1i} \cdot P(X_1 = x_{1i})$
1	1/6	1/6
2	1/6	2/6
3	1/6	3/6
4	1/6	4/6
5	1/6	5/6
6	1/6	6/6
		1 21/6

e

x_{2i}	$P(X_2 = x_{2i})$	$x_{2i} \cdot P(X_2 = x_{2i})$
0	4/6	0
1	1/6	1/6
2	1/6	2/6
		1 3/6

Considerando, ora, la media di una combinazione lineare si ottiene

$$M(X_1 + X_2) =$$



Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.

2. Calcolare $M(X_1 + X_2)$.

Vale

$$M(X_1) = 3.5 \quad M(X_2) = 0.5$$

Ricordando la definizione di valore atteso di una variabile casuale abbiamo, infatti

x_{1i}	$P(X_1 = x_{1i})$	$x_{1i} \cdot P(X_1 = x_{1i})$
1	1/6	1/6
2	1/6	2/6
3	1/6	3/6
4	1/6	4/6
5	1/6	5/6
6	1/6	6/6
1		21/6

e

x_{2i}	$P(X_2 = x_{2i})$	$x_{2i} \cdot P(X_2 = x_{2i})$
0	4/6	0
1	1/6	1/6
2	1/6	2/6
1		3/6

Considerando, ora, la media di una combinazione lineare si ottiene

$$M(X_1 + X_2) = M(X_1) +$$



Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.

2. Calcolare $M(X_1 + X_2)$.

Vale

$$M(X_1) = 3.5 \quad M(X_2) = 0.5$$

Ricordando la definizione di valore atteso di una variabile casuale abbiamo, infatti

x_{1i}	$P(X_1 = x_{1i})$	$x_{1i} \cdot P(X_1 = x_{1i})$
1	1/6	1/6
2	1/6	2/6
3	1/6	3/6
4	1/6	4/6
5	1/6	5/6
6	1/6	6/6
1		21/6

e

x_{2i}	$P(X_2 = x_{2i})$	$x_{2i} \cdot P(X_2 = x_{2i})$
0	4/6	0
1	1/6	1/6
2	1/6	2/6
1		3/6

Considerando, ora, la media di una combinazione lineare si ottiene

$$M(X_1 + X_2) = M(X_1) + M(X_2) = 3.5 + 0.5 = 4$$



Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.



3. Calcolare la probabilità che, su 6 lanci, il secondo dado presenti 3 volte valore 0.

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.



3. Calcolare la probabilità che, su 6 lanci, il secondo dado presenti 3 volte valore 0.

Sia X la variabile casuale 'n. di volte che si presenta 0 nelle 6 prove indipendenti'.
Abbiamo

$$X \sim$$

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.



Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.

3. Calcolare la probabilità che, su 6 lanci, il secondo dado presenti 3 volte valore 0.

Sia X la variabile casuale 'n. di volte che si presenta 0 nelle 6 prove indipendenti'.
Abbiamo

$$X \sim \text{Bin}(n = 6, p = \frac{4}{6} = 0.6667)$$

e

$$P(X = 3) =$$



Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.

3. Calcolare la probabilità che, su 6 lanci, il secondo dado presenti 3 volte valore 0.

Sia X la variabile casuale 'n. di volte che si presenta 0 nelle 6 prove indipendenti'.
Abbiamo

$$X \sim \text{Bin}(n = 6, p = \frac{4}{6} = 0.6667)$$

e

$$P(X = 3) = \binom{6}{3} 0.6667^3 \cdot 0.3333^3 = 0.2195.$$



Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.

4. Calcolare la probabilità che su $n = 200$ lanci l'evento 'secondo dado di valore 0' si presenti non più di 127 volte.



Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.

4. Calcolare la probabilità che su $n = 200$ lanci l'evento 'secondo dado di valore 0' si presenti non più di 127 volte.

Indicata con X la variabile casuale 'numero di volte che si presenta 0' nelle 200 prove indipendenti risulta



Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.

4. Calcolare la probabilità che su $n = 200$ lanci l'evento 'secondo dado di valore 0' si presenti non più di 127 volte.

Indicata con X la variabile casuale 'numero di volte che si presenta 0' nelle 200 prove indipendenti risulta

$$X \sim \text{Bin}(n = 200, p = 0.6667).$$

$$M(X) = 133.33; \text{Var}(X) = 44.44.$$



4. Calcolare la probabilità che su $n = 200$ lanci l'evento 'secondo dado di valore 0' si presenti non più di 127 volte.

Indicata con X la variabile casuale 'numero di volte che si presenta 0' nelle 200 prove indipendenti risulta

$$X \sim \text{Bin}(n = 200, p = 0.6667).$$

$$M(X) = 133.33; \text{Var}(X) = 44.44.$$

Dal momento che $np = 133.33 > 5$ ed $n(1 - p) = 66.67 > 5$ è possibile utilizzare l'approssimazione della variabile casuale Binomiale con la variabile casuale Normale

$$X \sim N(\mu = 133.33, \sigma^2 = 44.44).$$



4. Calcolare la probabilità che su $n = 200$ lanci l'evento 'secondo dado di valore 0' si presenti non più di 127 volte.

Indicata con X la variabile casuale 'numero di volte che si presenta 0' nelle 200 prove indipendenti risulta

$$X \sim \text{Bin}(n = 200, p = 0.6667).$$

$$M(X) = 133.33; \text{Var}(X) = 44.44.$$

Dal momento che $np = 133.33 > 5$ ed $n(1 - p) = 66.67 > 5$ è possibile utilizzare l'approssimazione della variabile casuale Binomiale con la variabile casuale Normale

$$X \sim N(\mu = 133.33, \sigma^2 = 44.44).$$

Si ottiene

$$P(X \leq 127) =$$



4. Calcolare la probabilità che su $n = 200$ lanci l'evento 'secondo dado di valore 0' si presenti non più di 127 volte.

Indicata con X la variabile casuale 'numero di volte che si presenta 0' nelle 200 prove indipendenti risulta

$$X \sim \text{Bin}(n = 200, p = 0.6667).$$

$$M(X) = 133.33; \text{Var}(X) = 44.44.$$

Dal momento che $np = 133.33 > 5$ ed $n(1 - p) = 66.67 > 5$ è possibile utilizzare l'approssimazione della variabile casuale Binomiale con la variabile casuale Normale

$$X \sim N(\mu = 133.33, \sigma^2 = 44.44).$$

Si ottiene

$$P(X \leq 127) = P\left(\frac{X - \mu}{\sigma} \leq \frac{127 - \mu}{\sigma}\right) =$$



4. Calcolare la probabilità che su $n = 200$ lanci l'evento 'secondo dado di valore 0' si presenti non più di 127 volte.

Indicata con X la variabile casuale 'numero di volte che si presenta 0' nelle 200 prove indipendenti risulta

$$X \sim \text{Bin}(n = 200, p = 0.6667).$$

$$M(X) = 133.33; \text{Var}(X) = 44.44.$$

Dal momento che $np = 133.33 > 5$ ed $n(1 - p) = 66.67 > 5$ è possibile utilizzare l'approssimazione della variabile casuale Binomiale con la variabile casuale Normale

$$X \sim N(\mu = 133.33, \sigma^2 = 44.44).$$

Si ottiene

$$P(X \leq 127) = P\left(\frac{X - \mu}{\sigma} \leq \frac{127 - \mu}{\sigma}\right) = P\left(Z \leq \frac{127 - 133.33}{6.6667}\right) =$$

=



4. Calcolare la probabilità che su $n = 200$ lanci l'evento 'secondo dado di valore 0' si presenti non più di 127 volte.

Indicata con X la variabile casuale 'numero di volte che si presenta 0' nelle 200 prove indipendenti risulta

$$X \sim \text{Bin}(n = 200, p = 0.6667).$$

$$M(X) = 133.33; \text{Var}(X) = 44.44.$$

Dal momento che $np = 133.33 > 5$ ed $n(1 - p) = 66.67 > 5$ è possibile utilizzare l'approssimazione della variabile casuale Binomiale con la variabile casuale Normale

$$X \sim N(\mu = 133.33, \sigma^2 = 44.44).$$

Si ottiene

$$\begin{aligned} P(X \leq 127) &= P\left(\frac{X - \mu}{\sigma} \leq \frac{127 - \mu}{\sigma}\right) = P\left(Z \leq \frac{127 - 133.33}{6.6667}\right) = \\ &= P(Z \leq -0.95) = \end{aligned}$$



4. Calcolare la probabilità che su $n = 200$ lanci l'evento 'secondo dado di valore 0' si presenti non più di 127 volte.

Indicata con X la variabile casuale 'numero di volte che si presenta 0' nelle 200 prove indipendenti risulta

$$X \sim \text{Bin}(n = 200, p = 0.6667).$$

$$M(X) = 133.33; \text{Var}(X) = 44.44.$$

Dal momento che $np = 133.33 > 5$ ed $n(1 - p) = 66.67 > 5$ è possibile utilizzare l'approssimazione della variabile casuale Binomiale con la variabile casuale Normale

$$X \sim N(\mu = 133.33, \sigma^2 = 44.44).$$

Si ottiene

$$\begin{aligned} P(X \leq 127) &= P\left(\frac{X - \mu}{\sigma} \leq \frac{127 - \mu}{\sigma}\right) = P\left(Z \leq \frac{127 - 133.33}{6.6667}\right) = \\ &= P(Z \leq -0.95) = P(Z \geq 0.95) = \\ &= \end{aligned}$$



4. Calcolare la probabilità che su $n = 200$ lanci l'evento 'secondo dado di valore 0' si presenti non più di 127 volte.

Indicata con X la variabile casuale 'numero di volte che si presenta 0' nelle 200 prove indipendenti risulta

$$X \sim \text{Bin}(n = 200, p = 0.6667).$$

$$M(X) = 133.33; \text{Var}(X) = 44.44.$$

Dal momento che $np = 133.33 > 5$ ed $n(1 - p) = 66.67 > 5$ è possibile utilizzare l'approssimazione della variabile casuale Binomiale con la variabile casuale Normale

$$X \sim N(\mu = 133.33, \sigma^2 = 44.44).$$

Si ottiene

$$\begin{aligned} P(X \leq 127) &= P\left(\frac{X - \mu}{\sigma} \leq \frac{127 - \mu}{\sigma}\right) = P\left(Z \leq \frac{127 - 133.33}{6.6667}\right) = \\ &= P(Z \leq -0.95) = P(Z \geq 0.95) = \\ &= 1 - P(Z < 0.95) = 1 - P(Z \leq 0.95) = \end{aligned}$$



4. Calcolare la probabilità che su $n = 200$ lanci l'evento 'secondo dado di valore 0' si presenti non più di 127 volte.

Indicata con X la variabile casuale 'numero di volte che si presenta 0' nelle 200 prove indipendenti risulta

$$X \sim \text{Bin}(n = 200, p = 0.6667).$$

$$M(X) = 133.33; \text{Var}(X) = 44.44.$$

Dal momento che $np = 133.33 > 5$ ed $n(1 - p) = 66.67 > 5$ è possibile utilizzare l'approssimazione della variabile casuale Binomiale con la variabile casuale Normale

$$X \sim N(\mu = 133.33, \sigma^2 = 44.44).$$

Si ottiene

$$\begin{aligned} P(X \leq 127) &= P\left(\frac{X - \mu}{\sigma} \leq \frac{127 - \mu}{\sigma}\right) = P\left(Z \leq \frac{127 - 133.33}{6.6667}\right) = \\ &= P(Z \leq -0.95) = P(Z \geq 0.95) = \\ &= 1 - P(Z < 0.95) = 1 - P(Z \leq 0.95) = 1 - 0.8289 = 0.1711. \end{aligned}$$

Tema T291

Prova scritta di STATISTICA del 20.01.2017

Prova scritta di
STATISTICA del
20.01.2017

Giuseppe Boari
Gabriele Cantaluppi



UNIVERSITÀ
CATTOLICA
del Sacro Cuore

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.

Giuseppe Boari
Gabriele Cantaluppi
Università Cattolica del Sacro Cuore, Milano



Al fine di analizzare la relazione tra la variabile $X = \text{'Nazione'}$ e $Y = \text{'Salario annuo (in migliaia di euro)'}$ rilevate su 20 unità statistiche, si sono raccolti i seguenti dati:

$X \setminus Y$	10 – 20	20 – 25	25 – 30
<i>Francia</i>	2	7	0
<i>Italia</i>	8	2	1

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.



1. Si rappresentino graficamente la distribuzione *marginale* di X e la *condizionata* di $Y|X = Italia$.

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.



Si considera la distribuzione marginale della variabile X , ottenuta sommando le frequenze di riga

x_i	$n_{i\bullet}$
<i>Francia</i>	9
<i>Italia</i>	11

Si tratta di un carattere qualitativo sconnesso che può essere rappresentato graficamente mediante un grafico a bastoncini o a canne d'organo.

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.

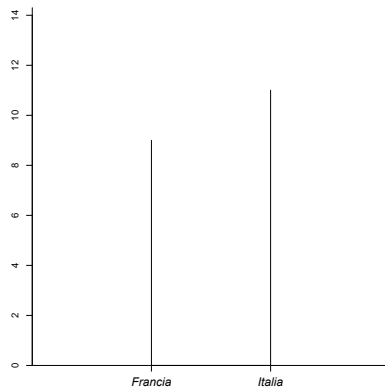


Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.





Si considera la distribuzione marginale della variabile Y , ottenuta sommando le frequenze di colonna

y_j	$n_{\bullet j}$
10 $\rightarrow$ 20	8
20 $\rightarrow$ 25	2
25 $\rightarrow$ 30	1

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.



Si tratta di un carattere quantitativo con modalità raggruppate in classi che può essere rappresentato graficamente mediante un istogramma.

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.



L'istogramma può essere costruito dopo avere calcolato le densità $d_{\bullet j} = \frac{n_{\bullet j}}{a_j}$, (dove a_j sono le ampiezze di classe).

y_j	$n_{\bullet j}$	$a_{\bullet j}$	$d_{\bullet j}$
10 + 20	8	$20 - 10 = 10$	0.8
20 + 25	2	$25 - 20 = 5$	0.4
25 + 30	1	$30 - 25 = 5$	0.2

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.

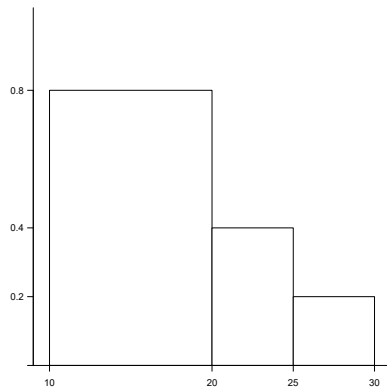


Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.





2. Si calcoli un opportuno indice di posizione per la variabile X . Si calcolino inoltre la media aritmetica e la mediana di $Y|X = \textit{Francia}$.

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.



Si tratta di un carattere qualitativo sconnesso.

L'unico indice di posizione che può essere definito è la moda (categoria alla quale è associata la massima frequenza) che risulta *Italia*.

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.



Si tratta di un carattere quantitativo con modalità raggruppate in classi.

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.



La mediana (percentile di ordine 0.50) può essere ottenuta considerando l'inversa della Funzione di Ripartizione in corrispondenza di 0.50.

y_j	n_{1j}	N_{1j}	F_{1j}
10 - 20	2	2	0.2222
20 - 25	7	9	1
25 - 30	0	9	1

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.

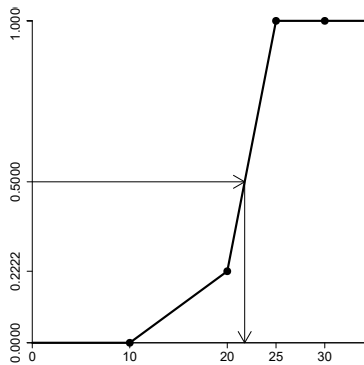


Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.



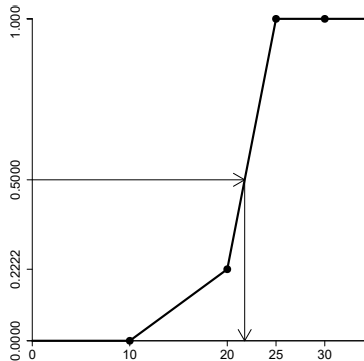


Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.



Si ottiene

$$y_{0.5} = 20 + \frac{0.5 - 0.2222}{1 - 0.2222} \cdot (25 - 20) = 21.7857.$$



La media aritmetica

$$M(Y|x_1) = \frac{1}{n_{1\bullet}} \sum_{j=1}^k y_j n_{1j}$$

risulta 20.8333.

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.



La media aritmetica

$$M(Y|x_1) = \frac{1}{n_{1\bullet}} \sum_{j=1}^k y_j n_{1j}$$

risulta 20.8333.

Si ottiene (considerando come modalità i valori centrali delle classi)

y_j	n_{1j}	$y_j n_{1j}$
15.0	2	30.0
22.5	7	157.5
27.5	0	0.0
	9	187.5

da cui segue

$$M(Y|x_1) = \frac{1}{n_{1\bullet}} \sum_{j=1}^k y_j n_{1j} = \frac{187.5}{9} = 20.8333.$$

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.



Riepilogo

indice posizione X : $\text{Moda}(X) = \text{Italia}$

$$M(Y|X = \text{Francia}) = 20.8333$$

$$y_{0.5}|X = \text{Francia} = 21.7857$$

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.



Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.

3. Si calcoli l'indice di mutabilità normalizzato di X , commentandone il risultato.



In presenza di un carattere qualitativo la variabilità può essere misurata mediante gli indici di Gini e di Frosini normalizzati

$$G_N = \frac{1 - \sum_{i=1}^h f_{i\bullet}^2}{1 - \frac{1}{h}}$$

e

$$F_N = 1 - \sqrt{1 - G_N}.$$

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.



In presenza di un carattere qualitativo la variabilità può essere misurata mediante gli indici di Gini e di Frosini normalizzati

$$G_N = \frac{1 - \sum_{i=1}^h f_{i\bullet}^2}{1 - \frac{1}{h}}$$

e

$$F_N = 1 - \sqrt{1 - G_N}.$$

Si ottiene:

x_i	$n_{i\bullet}$	$f_{i\bullet}$	$f_{i\bullet}^2$
<i>Francia</i>	9	0.45	0.2025
<i>Italia</i>	11	0.55	0.3025
	20	1	0.505

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.



In presenza di un carattere qualitativo la variabilità può essere misurata mediante gli indici di Gini e di Frosini normalizzati

$$G_N = \frac{1 - \sum_{i=1}^h f_{i\bullet}^2}{1 - \frac{1}{h}}$$

e

$$F_N = 1 - \sqrt{1 - G_N}.$$

Si ottiene:

x_i	$n_{i\bullet}$	$f_{i\bullet}$	$f_{i\bullet}^2$
<i>Francia</i>	9	0.45	0.2025
<i>Italia</i>	11	0.55	0.3025
	20	1	0.505

Da cui segue:

$$G_N = \frac{1 - 0.505}{1 - \frac{1}{2}} = 0.99$$

e

$$F_N = 1 - \sqrt{1 - 0.99} = 0.9.$$

Sussiste, quindi, un livello alto di eterogeneità.



4. Misurare e commentare la connessione tra le variabili X e Y .

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.



Dalla distribuzione delle frequenze congiunte n_{ij} :

	10 ÷ 20	20 ÷ 25	25 ÷ 30	$n_{i\bullet}$
<i>Francia</i>	2	7	0	9
<i>Italia</i>	8	2	1	11
$n_{\bullet j}$	10	9	1	20

è possibile ottenere le frequenze teoriche nella situazione di indipendenza stocastica

$$\hat{n}_{ij} = \frac{n_{i\bullet} n_{\bullet j}}{n}:$$

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.



Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.

Dalla distribuzione delle frequenze congiunte n_{ij} :

	10 ÷ 20	20 ÷ 25	25 ÷ 30	$n_{i\bullet}$
<i>Francia</i>	2	7	0	9
<i>Italia</i>	8	2	1	11
$n_{\bullet j}$	10	9	1	20

è possibile ottenere le frequenze teoriche nella situazione di indipendenza stocastica

$$\hat{n}_{ij} = \frac{n_{i\bullet} \cdot n_{\bullet j}}{n}:$$

	10 ÷ 20	20 ÷ 25	25 ÷ 30	$n_{i\bullet}$
<i>Francia</i>	4.5	4.05	0.45	9
<i>Italia</i>	5.5	4.95	0.55	11
$n_{\bullet j}$	10.0	9.00	1.00	20



le contingenze assolute $n_{ij} - \hat{n}_{ij}$:

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.



le contingenze assolute $n_{ij} - \hat{n}_{ij}$:

	10 ÷ 20	20 ÷ 25	25 ÷ 30
<i>Francia</i>	-2.5	2.95	-0.45
<i>Italia</i>	2.5	-2.95	0.45

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.



i rapporti di contingenza $\frac{n_{ij}}{\hat{n}_{ij}}$:

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.



i rapporti di contingenza $\frac{n_{ij}}{\hat{n}_{ij}}$:

	10 - 20	20 - 25	25 - 30
<i>Francia</i>	0.4444	1.7284	0.0000
<i>Italia</i>	1.4545	0.4040	1.8182

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.



e le contingenze relative, espresse in termini percentuali, $100 \left(\frac{n_{ij} - \hat{n}_{ij}}{\hat{n}_{ij}} \right) \% = 100 \left(\frac{n_{ij}}{\hat{n}_{ij}} - 1 \right) \%$:

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.



e le contingenze relative, espresse in termini percentuali, $100 \left(\frac{n_{ij} - \hat{n}_{ij}}{\hat{n}_{ij}} \right) \% = 100 \left(\frac{n_{ij}}{\hat{n}_{ij}} - 1 \right) \%$:

	10 ÷ 20	20 ÷ 25	25 ÷ 30
<i>Francia</i>	-55.56	72.84	-100.00
<i>Italia</i>	45.45	-59.60	81.82

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.



Risultati intermedi $\frac{(n_{ij} - \hat{n}_{ij})^2}{\hat{n}_{ij}}$ per il calcolo del χ^2

	10 ÷ 20	20 ÷ 25	25 ÷ 30
<i>Francia</i>	1.3889	2.1488	0.4500
<i>Italia</i>	1.1364	1.7581	0.3682

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.



Risultati intermedi $\frac{n_{ij}^2}{n_{i\bullet} \cdot n_{\bullet j}}$ per il calcolo del χ^2 (formula operativa)

	10 ÷ 20	20 ÷ 25	25 ÷ 30
<i>Francia</i>	0.0444	0.6049	0.0000
<i>Italia</i>	0.5818	0.0404	0.0909

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.



Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.

$$\chi^2 = \sum_{i=1}^h \sum_{j=1}^k \frac{(n_{ij} - \hat{n}_{ij})^2}{\hat{n}_{ij}} = 7.2503,$$

$$\chi_N^2 = \frac{\chi^2}{n \cdot \min(h-1, k-1)} = 0.3625$$

$$\chi_N = \sqrt{\chi_N^2} = 0.6021.$$

Sussiste, quindi, un livello di connessione medio.

Esercizio 2



Si consideri il seguente prospetto:

t	x_t	$NIBF(t = 2010)$	$NIBM$
2008			
2009			0.25
2010	150	1	0.25
2011		1	
2012			0.5

1. Si completi il prospetto con i dati mancanti.

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.



Considerando i numeri indici a base fissa e a base mobile assegnati è possibile ricostruire la serie dei dati

$${}_{2010}I_{2011} =$$

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.



Considerando i numeri indici a base fissa e a base mobile assegnati è possibile ricostruire la serie dei dati

$${}_{2010}I_{2011} = \frac{x_{2011}}{x_{2010}} = \frac{x_{2011}}{150} = 1 \rightarrow x_{2011} =$$

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.



Considerando i numeri indici a base fissa e a base mobile assegnati è possibile ricostruire la serie dei dati

$${}_{2010}I_{2011} = \frac{X_{2011}}{X_{2010}} = \frac{X_{2011}}{150} = 1 \rightarrow x_{2011} = 150$$

$${}_{2009}I_{2010} =$$

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.



Considerando i numeri indici a base fissa e a base mobile assegnati è possibile ricostruire la serie dei dati

$${}_{2010}I_{2011} = \frac{X_{2011}}{X_{2010}} = \frac{X_{2011}}{150} = 1 \rightarrow x_{2011} = 150$$

$${}_{2009}I_{2010} = \frac{X_{2010}}{X_{2009}} = \frac{150}{X_{2009}} = 0.25 \rightarrow x_{2009} =$$

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.



Considerando i numeri indici a base fissa e a base mobile assegnati è possibile ricostruire la serie dei dati

$${}_{2010}I_{2011} = \frac{X_{2011}}{X_{2010}} = \frac{X_{2011}}{150} = 1 \rightarrow x_{2011} = 150$$

$${}_{2009}I_{2010} = \frac{X_{2010}}{X_{2009}} = \frac{150}{X_{2009}} = 0.25 \rightarrow x_{2009} = 600$$

$${}_{2008}I_{2009} =$$

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.



Considerando i numeri indici a base fissa e a base mobile assegnati è possibile ricostruire la serie dei dati

$${}_{2010}I_{2011} = \frac{X_{2011}}{X_{2010}} = \frac{X_{2011}}{150} = 1 \rightarrow x_{2011} = 150$$

$${}_{2009}I_{2010} = \frac{X_{2010}}{X_{2009}} = \frac{150}{X_{2009}} = 0.25 \rightarrow x_{2009} = 600$$

$${}_{2008}I_{2009} = \frac{X_{2009}}{X_{2008}} = \frac{600}{X_{2008}} = 0.25 \rightarrow x_{2008} =$$

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.



Considerando i numeri indici a base fissa e a base mobile assegnati è possibile ricostruire la serie dei dati

$${}_{2010}I_{2011} = \frac{X_{2011}}{X_{2010}} = \frac{X_{2011}}{150} = 1 \rightarrow x_{2011} = 150$$

$${}_{2009}I_{2010} = \frac{X_{2010}}{X_{2009}} = \frac{150}{X_{2009}} = 0.25 \rightarrow x_{2009} = 600$$

$${}_{2008}I_{2009} = \frac{X_{2009}}{X_{2008}} = \frac{600}{X_{2008}} = 0.25 \rightarrow x_{2008} = 2400$$

$${}_{2011}I_{2012} =$$

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.



Considerando i numeri indici a base fissa e a base mobile assegnati è possibile ricostruire la serie dei dati

$${}_{2010}I_{2011} = \frac{x_{2011}}{x_{2010}} = \frac{x_{2011}}{150} = 1 \rightarrow x_{2011} = 150$$

$${}_{2009}I_{2010} = \frac{x_{2010}}{x_{2009}} = \frac{150}{x_{2009}} = 0.25 \rightarrow x_{2009} = 600$$

$${}_{2008}I_{2009} = \frac{x_{2009}}{x_{2008}} = \frac{600}{x_{2008}} = 0.25 \rightarrow x_{2008} = 2400$$

$${}_{2011}I_{2012} = \frac{x_{2012}}{x_{2011}} = \frac{x_{2012}}{150} = 0.5 \rightarrow x_{2012} =$$

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.



Considerando i numeri indici a base fissa e a base mobile assegnati è possibile ricostruire la serie dei dati

$${}_{2010}I_{2011} = \frac{X_{2011}}{X_{2010}} = \frac{X_{2011}}{150} = 1 \rightarrow x_{2011} = 150$$

$${}_{2009}I_{2010} = \frac{X_{2010}}{X_{2009}} = \frac{150}{X_{2009}} = 0.25 \rightarrow x_{2009} = 600$$

$${}_{2008}I_{2009} = \frac{X_{2009}}{X_{2008}} = \frac{600}{X_{2008}} = 0.25 \rightarrow x_{2008} = 2400$$

$${}_{2011}I_{2012} = \frac{X_{2012}}{X_{2011}} = \frac{X_{2012}}{150} = 0.5 \rightarrow x_{2012} = 75$$

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.



Considerando la serie dei dati ricostruita abbiamo

$$2010 I_{2008} =$$

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.



Considerando la serie dei dati ricostruita abbiamo

$${}_{2010}I_{2008} = \frac{x_{2008}}{x_{2010}} = \frac{2400}{150} = 16$$

$${}_{2010}I_{2009} =$$

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.



Considerando la serie dei dati ricostruita abbiamo

$${}_{2010}I_{2008} = \frac{x_{2008}}{x_{2010}} = \frac{2400}{150} = 16$$

$${}_{2010}I_{2009} = \frac{x_{2009}}{x_{2010}} = \frac{600}{150} = 4$$

$${}_{2010}I_{2012} =$$

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.



Considerando la serie dei dati ricostruita abbiamo

$${}_{2010}I_{2008} = \frac{x_{2008}}{x_{2010}} = \frac{2400}{150} = 16$$

$${}_{2010}I_{2009} = \frac{x_{2009}}{x_{2010}} = \frac{600}{150} = 4$$

$${}_{2010}I_{2012} = \frac{x_{2012}}{x_{2010}} = \frac{75}{150} = 0.5$$

$${}_{2010}I_{2011} =$$

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.



Considerando la serie dei dati ricostruita abbiamo

$${}_{2010}I_{2008} = \frac{x_{2008}}{x_{2010}} = \frac{2400}{150} = 16$$

$${}_{2010}I_{2009} = \frac{x_{2009}}{x_{2010}} = \frac{600}{150} = 4$$

$${}_{2010}I_{2012} = \frac{x_{2012}}{x_{2010}} = \frac{75}{150} = 0.5$$

$${}_{2010}I_{2011} = \frac{x_{2011}}{x_{2010}} = \frac{150}{150} = 1$$

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.



Si ottiene, quindi

t	x_t	$NIBF(t = 2010)$	$NIBM$
2008	2400	16	
2009	600	4	0.25
2010	150	1	0.25
2011	150	1	1
2012	75	0.5	0.5

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.



2. Si calcoli e si commenti $_{2008/2010}$, numero indice riferito al tempo 2010 con base 2008.

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.



2. Si calcoli e si commenti ${}_{2008}I_{2010}$, numero indice riferito al tempo 2010 con base 2008.
Si ricorda che

$${}_{2008}I_{2010} = {}_{2008}I_{2009} \cdot {}_{2009}I_{2010} = 0.0625.$$

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.



Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.

2. Si calcoli e si commenti ${}_{2008}I_{2010}$, numero indice riferito al tempo 2010 con base 2008. Si ricorda che

$${}_{2008}I_{2010} = {}_{2008}I_{2009} \cdot {}_{2009}I_{2010} = 0.0625.$$

Tra il tempo 2008 e il tempo 2010 si è registrato un decremento complessivo del 93.75% della grandezza oggetto di studio.



3. Si calcoli il tasso medio di variazione dal 2008 al 2010 e si commenti il risultato ottenuto.

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.



3. Si calcoli il tasso medio di variazione dal 2008 al 2010 e si commenti il risultato ottenuto.
Si ricorda che

$${}_{2008}I_{2010} = {}_{2008}I_{2009} \cdot {}_{2009}I_{2010}.$$

Si deve trovare il numero indice a base mobile medio.

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.



Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.

3. Si calcoli il tasso medio di variazione dal 2008 al 2010 e si commenti il risultato ottenuto. Si ricorda che

$${}_{2008}I_{2010} = {}_{2008}I_{2009} \cdot {}_{2009}I_{2010}.$$

Si deve trovare il numero indice a base mobile medio. Applicando il criterio di scelta della media secondo Chisini si ottiene

$${}_{2008}I_{2010} = {}_{2008}I_{2009} \cdot {}_{2009}I_{2010} = \alpha \cdot \alpha = \alpha^2.$$



Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.

3. Si calcoli il tasso medio di variazione dal 2008 al 2010 e si commenti il risultato ottenuto. Si ricorda che

$${}_{2008}I_{2010} = {}_{2008}I_{2009} \cdot {}_{2009}I_{2010}.$$

Si deve trovare il numero indice a base mobile medio. Applicando il criterio di scelta della media secondo Chisini si ottiene

$${}_{2008}I_{2010} = {}_{2008}I_{2009} \cdot {}_{2009}I_{2010} = \alpha \cdot \alpha = \alpha^2.$$

Quindi

$$\alpha = \sqrt[2]{{}_{2008}I_{2010}} = \sqrt[2]{{}_{2008}I_{2009} \cdot {}_{2009}I_{2010}} = 0.25.$$

Da cui segue una variazione relativa media pari a -0.75 , corrispondente a un decremento medio annuo tra il 2008 e il 2010 del 75%.



Il seguente prospetto riporta i valori rilevati in corrispondenza a 8 capofamiglia con riferimento a due caratteri quantitativi X = numero dei componenti della famiglia e Y = spesa per un vacanza di due giorni (in centinaia di euro).

x_i	1	2	2	2	3	4	4	4
y_i	1	1	2	3	3	3	4	5

1. Si rappresentino graficamente le coppie dei dati e la funzione di regressione ipotizzando la spesa in funzione del numero dei componenti della famiglia.



La funzione di regressione associa a ogni valore x appartenente all'insieme di definizione della variabile X la media di Y condizionata a x , $M(Y|x)$

x	$M(Y x)$
1	1
2	$\frac{1}{3}(1 + 2 + 3) = 2$
3	3
4	$\frac{1}{3}(3 + 4 + 5) = 4$

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.



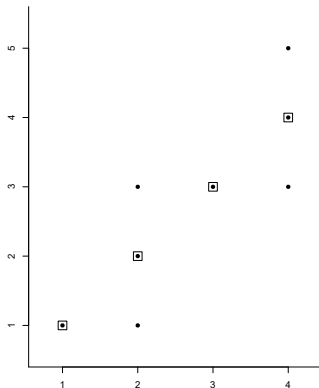
Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.

Le coppie di dati possono essere rappresentate mediante un grafico a dispersione.





2. Si calcolino, secondo il metodo dei minimi quadrati, i parametri dei seguenti modelli di regressione:

$$I) Y^* = a + b \cdot X \quad II) Y^* = c \cdot X$$

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.

Si riportano i calcoli necessari¹ per ottenere le stime dei parametri e gli indici di adattamento del modello

$$Y = a + bX + E$$

x_i	y_i	$x_i y_i$	x_i^2	y_i^2
1	1	1	1	1
2	1	2	4	1
2	2	4	4	4
2	3	6	4	9
3	3	9	9	9
4	3	12	16	9
4	4	16	16	16
4	5	20	16	25
22	22	70	70	74
2.75	2.75	8.75	8.75	9.25
$M(X)$	$M(Y)$	$M(XY)$	$M(X^2)$	$M(Y^2)$

$$\hat{b} =$$



Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.

¹Si osserva che per ragioni di stampa alcuni calcoli potrebbero essere duplicati con riferimento ai diversi modelli oggetto di studio

Si riportano i calcoli necessari¹ per ottenere le stime dei parametri e gli indici di adattamento del modello

$$Y = a + bX + E$$

x_i	y_i	$x_i y_i$	x_i^2	y_i^2
1	1	1	1	1
2	1	2	4	1
2	2	4	4	4
2	3	6	4	9
3	3	9	9	9
4	3	12	16	9
4	4	16	16	16
4	5	20	16	25
22	22	70	70	74
2.75	2.75	8.75	8.75	9.25
$M(X)$	$M(Y)$	$M(XY)$	$M(X^2)$	$M(Y^2)$

$$\begin{aligned}\hat{b} &= \frac{\text{Cov}(X, Y)}{\text{Var}(X)} = \frac{M(XY) - M(X)M(Y)}{M(X^2) - [M(X)]^2} = \\ &= \frac{8.75 - 2.75 \cdot 2.75}{8.75 - 2.75^2} = \frac{1.1875}{1.1875} = 1 \\ \hat{a} &= \end{aligned}$$

¹Si osserva che per ragioni di stampa alcuni calcoli potrebbero essere duplicati con riferimento ai diversi modelli oggetto di studio



Si riportano i calcoli necessari¹ per ottenere le stime dei parametri e gli indici di adattamento del modello

$$Y = a + bX + E$$

x_i	y_i	$x_i y_i$	x_i^2	y_i^2
1	1	1	1	1
2	1	2	4	1
2	2	4	4	4
2	3	6	4	9
3	3	9	9	9
4	3	12	16	9
4	4	16	16	16
4	5	20	16	25
22	22	70	70	74
2.75	2.75	8.75	8.75	9.25
$M(X)$	$M(Y)$	$M(XY)$	$M(X^2)$	$M(Y^2)$

$$\begin{aligned}\hat{b} &= \frac{\text{Cov}(X, Y)}{\text{Var}(X)} = \frac{M(XY) - M(X)M(Y)}{M(X^2) - [M(X)]^2} = \\ &= \frac{8.75 - 2.75 \cdot 2.75}{8.75 - 2.75^2} = \frac{1.1875}{1.1875} = 1\end{aligned}$$

$$\hat{a} = M(Y) - \hat{b}M(X) = 0$$

$$\rho^2 =$$

¹Si osserva che per ragioni di stampa alcuni calcoli potrebbero essere duplicati con riferimento ai diversi modelli oggetto di studio



Si riportano i calcoli necessari¹ per ottenere le stime dei parametri e gli indici di adattamento del modello

$$Y = a + bX + E$$

x_i	y_i	$x_i y_i$	x_i^2	y_i^2
1	1	1	1	1
2	1	2	4	1
2	2	4	4	4
2	3	6	4	9
3	3	9	9	9
4	3	12	16	9
4	4	16	16	16
4	5	20	16	25
22	22	70	70	74
2.75	2.75	8.75	8.75	9.25
$M(X)$	$M(Y)$	$M(XY)$	$M(X^2)$	$M(Y^2)$

$$\hat{b} = \frac{\text{Cov}(X, Y)}{\text{Var}(X)} = \frac{M(XY) - M(X)M(Y)}{M(X^2) - [M(X)]^2} =$$

$$= \frac{8.75 - 2.75 \cdot 2.75}{8.75 - 2.75^2} = \frac{1.1875}{1.1875} = 1$$

$$\hat{a} = M(Y) - \hat{b} M(X) = 0$$

$$\rho^2 = \frac{[\text{Cov}(X, Y)]^2}{\text{Var}(X) \text{Var}(Y)} = \frac{1.1875^2}{1.1875 \cdot 1.6875} = 0.7037$$

$$M(\hat{E}^2) = \text{Var}(\hat{E}) = \text{varianza residua} =$$

¹Si osserva che per ragioni di stampa alcuni calcoli potrebbero essere duplicati con riferimento ai diversi modelli oggetto di studio



Si riportano i calcoli necessari¹ per ottenere le stime dei parametri e gli indici di adattamento del modello

$$Y = a + bX + E$$

x_i	y_i	$x_i y_i$	x_i^2	y_i^2
1	1	1	1	1
2	1	2	4	1
2	2	4	4	4
2	3	6	4	9
3	3	9	9	9
4	3	12	16	9
4	4	16	16	16
4	5	20	16	25
22	22	70	70	74
2.75	2.75	8.75	8.75	9.25
$M(X)$	$M(Y)$	$M(XY)$	$M(X^2)$	$M(Y^2)$

$$\begin{aligned}\hat{b} &= \frac{\text{Cov}(X, Y)}{\text{Var}(X)} = \frac{M(XY) - M(X)M(Y)}{M(X^2) - [M(X)]^2} = \\ &= \frac{8.75 - 2.75 \cdot 2.75}{8.75 - 2.75^2} = \frac{1.1875}{1.1875} = 1\end{aligned}$$

$$\hat{a} = M(Y) - \hat{b} M(X) = 0$$

$$\rho^2 = \frac{[\text{Cov}(X, Y)]^2}{\text{Var}(X) \text{Var}(Y)} = \frac{1.1875^2}{1.1875 \cdot 1.6875} = 0.7037$$

$$M(\hat{E}^2) = \text{Var}(\hat{E}) = \text{varianza residua} = \text{Var}(Y) \cdot (1 - \rho^2) = 0.5$$

¹Si osserva che per ragioni di stampa alcuni calcoli potrebbero essere duplicati con riferimento ai diversi modelli oggetto di studio





Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.

Si riportano i calcoli necessari² per ottenere le stime dei parametri e la misura di adattamento del modello

$$Y = cX + E$$

x_i	y_i	$x_i y_i$	x_i^2	y_i^2
1	1	1	1	1
2	1	2	4	1
2	2	4	4	4
2	3	6	4	9
3	3	9	9	9
4	3	12	16	9
4	4	16	16	16
4	5	20	16	25
		70	70	74
		8.75	8.75	9.25
		$M(XY)$	$M(X^2)$	$M(Y^2)$

$$\hat{c} =$$

²Si osserva che per ragioni di stampa alcuni calcoli potrebbero essere duplicati con riferimento ai diversi modelli oggetto di studio



Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.

Si riportano i calcoli necessari² per ottenere le stime dei parametri e la misura di adattamento del modello

$$Y = cX + E$$

x_i	y_i	$x_i y_i$	x_i^2	y_i^2
1	1	1	1	1
2	1	2	4	1
2	2	4	4	4
2	3	6	4	9
3	3	9	9	9
4	3	12	16	9
4	4	16	16	16
4	5	20	16	25
		70	70	74
		8.75	8.75	9.25
		$M(XY)$	$M(X^2)$	$M(Y^2)$

$$\hat{c} = \frac{M(XY)}{M(X^2)} = \frac{8.75}{8.75} = 1$$

$$M(\hat{E}^2) =$$

²Si osserva che per ragioni di stampa alcuni calcoli potrebbero essere duplicati con riferimento ai diversi modelli oggetto di studio



Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.

Si riportano i calcoli necessari² per ottenere le stime dei parametri e la misura di adattamento del modello

$$Y = cX + E$$

x_i	y_i	$x_i y_i$	x_i^2	y_i^2
1	1	1	1	1
2	1	2	4	1
2	2	4	4	4
2	3	6	4	9
3	3	9	9	9
4	3	12	16	9
4	4	16	16	16
4	5	20	16	25
		70	70	74
		8.75	8.75	9.25
		$M(XY)$	$M(X^2)$	$M(Y^2)$

$$\hat{c} = \frac{M(XY)}{M(X^2)} = \frac{8.75}{8.75} = 1$$

$$M(\hat{E}^2) = M[Y \cdot (Y - \hat{c}X)] = M(Y^2) - \hat{c}M(XY) = 9.25 - 1 \cdot 8.75 = 0.5$$

²Si osserva che per ragioni di stampa alcuni calcoli potrebbero essere duplicati con riferimento ai diversi modelli oggetto di studio



Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.

$\hat{a} = 0$, $\hat{b} = 1$, indice adattamento $M(\hat{E}^2) = 0.5$ $\hat{c} = 1$, indice adattamento $M(\hat{E}^2) = 0.5$



3. Confrontare gli indici di adattamento, commentando i risultati.

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.



3. Confrontare gli indici di adattamento, commentando i risultati.
I due modelli sono equivalenti.

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.



4. Si determini il valore di $\eta^2_{Y|X}$.

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.



Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.

4. Si determini il valore di $\eta_{Y|X}^2$.

$\eta_{Y|X}^2 = \rho^2 = 0.7037$ in quanto le medie condizionate sono allineate.



5. Si dica, senza effettuare calcoli e motivando la risposta, se $\eta^2_{X|Y} = 1$.

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.



Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.

5. Si dica, senza effettuare calcoli e motivando la risposta, se $\eta^2_{X|Y} = 1$.

$\eta^2_{X|Y} < 1$ in quanto le distribuzioni $X|y_j$ non sono degeneri.

Esercizio 4

Prova scritta di
STATISTICA del
20.01.2017

Giuseppe Boari
Gabriele Cantaluppi



UNIVERSITÀ
CATTOLICA
del Sacro Cuore

Una indagine di una compagnia telefonica ha stabilito che la durata (in secondi) delle chiamate dei propri utenti è distribuita come una Normale di media μ e varianza σ^2 .

1. Sapendo che il 50% degli utenti effettua chiamate inferiori a 190 secondi, stabilire il valore della media μ .

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.



Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.

Una indagine di una compagnia telefonica ha stabilito che la durata (in secondi) delle chiamate dei propri utenti è distribuita come una Normale di media μ e varianza σ^2 .

1. Sapendo che il 50% degli utenti effettua chiamate inferiori a 190 secondi, stabilire il valore della media μ .

$\mu = 190$ in quanto la funzione di densità della variabile casuale Normale è simmetrica, quindi la media aritmetica coincide con la mediana 190.



Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.

2. Supponendo che $\mu = 290$ e $\sigma = 80$, si calcoli la probabilità che una telefonata duri più di 370 secondi.

Sia X la variabile casuale 'durata di una telefonata'.

$$X \sim N(\mu = 290, \sigma^2 = 80^2 = 6400).$$

$$P(X > 370) =$$



Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.

2. Supponendo che $\mu = 290$ e $\sigma = 80$, si calcoli la probabilità che una telefonata duri più di 370 secondi.

Sia X la variabile casuale 'durata di una telefonata'.

$$X \sim N(\mu = 290, \sigma^2 = 80^2 = 6400).$$

$$P(X > 370) = P\left(\frac{X - \mu}{\sigma} > \frac{370 - \mu}{\sigma}\right) =$$



Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.

2. Supponendo che $\mu = 290$ e $\sigma = 80$, si calcoli la probabilità che una telefonata duri più di 370 secondi.

Sia X la variabile casuale 'durata di una telefonata'.

$$X \sim N(\mu = 290, \sigma^2 = 80^2 = 6400).$$

$$P(X > 370) = P\left(\frac{X - \mu}{\sigma} > \frac{370 - \mu}{\sigma}\right) = P\left(Z > \frac{370 - 290}{80}\right) =$$

=



Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.

2. Supponendo che $\mu = 290$ e $\sigma = 80$, si calcoli la probabilità che una telefonata duri più di 370 secondi.

Sia X la variabile casuale 'durata di una telefonata'.

$$X \sim N(\mu = 290, \sigma^2 = 80^2 = 6400).$$

$$\begin{aligned} P(X > 370) &= P\left(\frac{X - \mu}{\sigma} > \frac{370 - \mu}{\sigma}\right) = P\left(Z > \frac{370 - 290}{80}\right) = \\ &= P(Z > 1) = \end{aligned}$$



Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.

2. Supponendo che $\mu = 290$ e $\sigma = 80$, si calcoli la probabilità che una telefonata duri più di 370 secondi.

Sia X la variabile casuale 'durata di una telefonata'.

$$X \sim N(\mu = 290, \sigma^2 = 80^2 = 6400).$$

$$\begin{aligned} P(X > 370) &= P\left(\frac{X - \mu}{\sigma} > \frac{370 - \mu}{\sigma}\right) = P\left(Z > \frac{370 - 290}{80}\right) = \\ &= P(Z > 1) = 1 - P(Z \leq 1) = \end{aligned}$$



Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.

2. Supponendo che $\mu = 290$ e $\sigma = 80$, si calcoli la probabilità che una telefonata duri più di 370 secondi.

Sia X la variabile casuale 'durata di una telefonata'.

$$X \sim N(\mu = 290, \sigma^2 = 80^2 = 6400).$$

$$\begin{aligned} P(X > 370) &= P\left(\frac{X - \mu}{\sigma} > \frac{370 - \mu}{\sigma}\right) = P\left(Z > \frac{370 - 290}{80}\right) = \\ &= P(Z > 1) = 1 - P(Z \leq 1) = 1 - 0.8413 = \end{aligned}$$



Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.

2. Supponendo che $\mu = 290$ e $\sigma = 80$, si calcoli la probabilità che una telefonata duri più di 370 secondi.

Sia X la variabile casuale 'durata di una telefonata'.

$$X \sim N(\mu = 290, \sigma^2 = 80^2 = 6400).$$

$$\begin{aligned} P(X > 370) &= P\left(\frac{X - \mu}{\sigma} > \frac{370 - \mu}{\sigma}\right) = P\left(Z > \frac{370 - 290}{80}\right) = \\ &= P(Z > 1) = 1 - P(Z \leq 1) = 1 - 0.8413 = 0.1587. \end{aligned}$$



3. Si calcoli la probabilità che su 10 telefonate almeno una duri più di 370 secondi.

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.



Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.

3. Si calcoli la probabilità che su 10 telefonate almeno una duri più di 370 secondi. Sia Y la variabile casuale 'n. di telefonate con durata superiore a 370 secondi nelle 10 prove indipendenti'.

$$Y \sim \text{Bin}(n = 10, p = 0.1587).$$

$$P(Y \geq 1) =$$



Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.

3. Si calcoli la probabilità che su 10 telefonate almeno una duri più di 370 secondi. Sia Y la variabile casuale 'n. di telefonate con durata superiore a 370 secondi nelle 10 prove indipendenti'.

$$Y \sim \text{Bin}(n = 10, p = 0.1587).$$

$$P(Y \geq 1) = 1 - P(Y = 0) =$$

=



Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.

3. Si calcoli la probabilità che su 10 telefonate almeno una duri più di 370 secondi. Sia Y la variabile casuale 'n. di telefonate con durata superiore a 370 secondi nelle 10 prove indipendenti'.

$$Y \sim \text{Bin}(n = 10, p = 0.1587).$$

$$\begin{aligned} P(Y \geq 1) &= 1 - P(Y = 0) = \\ &= 1 - \binom{10}{0} 0.1587^0 0.8413^{10} = \\ &= 1 - 0.8413^{10} = \end{aligned}$$



Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.

3. Si calcoli la probabilità che su 10 telefonate almeno una duri più di 370 secondi.
Sia Y la variabile casuale 'n. di telefonate con durata superiore a 370 secondi nelle 10 prove indipendenti'.

$$Y \sim \text{Bin}(n = 10, p = 0.1587).$$

$$\begin{aligned} P(Y \geq 1) &= 1 - P(Y = 0) = \\ &= 1 - \binom{10}{0} 0.1587^0 0.8413^{10} = \\ &= 1 - 0.8413^{10} = 1 - 0.1776 = \end{aligned}$$



Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.

3. Si calcoli la probabilità che su 10 telefonate almeno una duri più di 370 secondi. Sia Y la variabile casuale 'n. di telefonate con durata superiore a 370 secondi nelle 10 prove indipendenti'.

$$Y \sim \text{Bin}(n = 10, p = 0.1587).$$

$$\begin{aligned} P(Y \geq 1) &= 1 - P(Y = 0) = \\ &= 1 - \binom{10}{0} 0.1587^0 0.8413^{10} = \\ &= 1 - 0.8413^{10} = 1 - 0.1776 = 0.8224. \end{aligned}$$

Tema T292

Prova scritta di STATISTICA del 06.02.2017

Prova scritta di
STATISTICA del
06.02.2017

Giuseppe Boari
Gabriele Cantaluppi



UNIVERSITÀ
CATTOLICA
del Sacro Cuore

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.

Giuseppe Boari
Gabriele Cantaluppi
Università Cattolica del Sacro Cuore, Milano



Il prospetto seguente riporta le distribuzioni relative del carattere quantitativo Y *condizionate* rispetto alle categorie del carattere qualitativo X .

$Y x_i$	1	2	3	
x_1	1	0	0	1
x_2	0	0.5	0.5	1

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.



1. Sapendo che le frequenze marginali di X sono $n_{1\bullet} = 10$ e $n_{2\bullet} = 20$ si ricostruiscano le frequenze congiunte della variabile doppia (X, Y) .

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.



Moltiplicando le frequenze condizionate relative $f.cond.r_{1j} = \frac{n_{1j}}{n_{1\bullet}}$ per il numero totale di osservazioni $n_{1\bullet}$ della distribuzione condizionata, si ottengono le frequenze assolute congiunte n_{1j} della distribuzione condizionata $Y|X_1$.

y_j	$f.cond.r_{1j}$	n_{1j}
1	1	10
2	0	0
3	0	0

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.



Moltiplicando le frequenze condizionate relative $f.cond.r_{2j} = \frac{n_{2j}}{n_{2\bullet}}$ per il numero totale di osservazioni $n_{2\bullet}$ della distribuzione condizionata, si ottengono le frequenze assolute congiunte n_{2j} della distribuzione condizionata $Y|x_2$.

y_j	$f.cond.r_{2j}$	n_{2j}
1	0.0	0
2	0.5	10
3	0.5	10

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.



La distribuzione delle frequenze congiunte risulta

$X \setminus Y$	1	2	3	$n_{i\bullet}$
x_1	10	0	0	10
x_2	0	10	10	20
$n_{\bullet j}$	10	10	10	30

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.



2. Si rappresentino graficamente le frequenze cumulate della marginale Y e se ne determini il valore della mediana.

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.



Si considera la distribuzione marginale della variabile Y , ottenuta sommando le frequenze di colonna.

y_j	$n_{\bullet j}$	$N_{\bullet j}$	$F_{\bullet j}$
1	10	10	0.3333
2	10	20	0.6667
3	10	30	1

Esercizio 1.

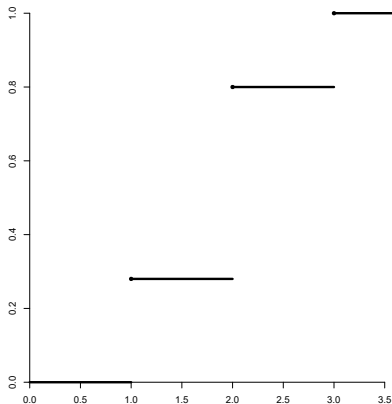
Esercizio 2.

Esercizio 3.

Esercizio 4.



Si tratta di un carattere quantitativo con modalità non raggruppate in classi. La funzione di ripartizione ha l'aspetto di una funzione a gradini.



Esercizio 1.

Esercizio 2.

Esercizio 3.

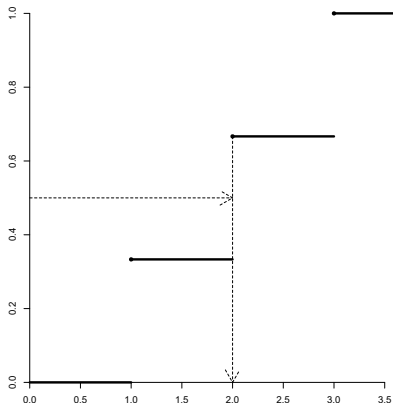
Esercizio 4.

La mediana risulta

$$y_{0.5} = \frac{1}{2} (y_{(15)} + y_{(16)}) = 2$$

Si osserva che la distribuzione è simmetrica.

È anche possibile ricavare il valore della mediana dal grafico della funzione di ripartizione



Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.



3. Si confrontino le distribuzioni condizionate $Y|x_1$ e $Y|x_2$ mediante opportuni indici di posizione e di variabilità.

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.



Si considera la distribuzione condizionata $Y|x_1$

y_j	n_{1j}
1	10
2	0
3	0

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.



Si tratta di un carattere quantitativo con modalità non raggruppate in classi.

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.



Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.

Si osserva che la distribuzione è degenere.

La moda risulta 1. $y_{0.5} = 1$. $M(Y|x_1) = 1$. La varianza e il coefficiente di variazione sono nulli.



Si considera la distribuzione condizionata $Y|x_1$

y_j	n_{1j}
1	0
2	10
3	10

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.



Si tratta di un carattere quantitativo con modalità non raggruppate in classi.

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.



La moda corrisponde alla modalità alla quale è associata la massima frequenza e risulta $\{2, 3\}$ (situazione bimodale).

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.



La mediana (percentile di ordine 0.50) in presenza di un numero pari di osservazioni corrisponde a uno degli elementi che occupano le posizioni $\frac{n}{2} = 10$ ed $\frac{n}{2} + 1 = 11$ nella serie dei dati ordinati. Dalla distribuzione delle frequenze cumulate

y_j	n_{1j}	N_{1j}
1	0	0
2	10	10
3	10	20

i due elementi risultano $y_{0.5} = \{2, 3\}$

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.



La media aritmetica

$$M(Y|x_1) = \frac{1}{n_{1\bullet}} \sum_{j=1}^k y_j n_{1j}$$

risulta 2.5.

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.



La media aritmetica

$$M(Y|x_1) = \frac{1}{n_{1\bullet}} \sum_{j=1}^k y_j n_{1j}$$

risulta 2.5.

Si ottiene

y_j	n_{1j}	$y_j n_{1j}$
1	0	0
2	10	20
3	10	30
20		50

da cui segue

$$M(Y|x_1) = \frac{1}{n_{1\bullet}} \sum_{j=1}^k y_j n_{1j} = \frac{50}{20} = 2.5.$$

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.



La varianza

$$\text{Var}(Y|x_1) = M(Y^2|x_1) - [M(Y|x_1)]^2$$

risulta 0.25.

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.



La varianza

$$\text{Var}(Y|x_1) = M(Y^2|x_1) - [M(Y|x_1)]^2$$

risulta 0.25. Si ottiene

y_j	n_{1j}	$y_j n_{1j}$	$y_j^2 n_{1j}$
1	0	0	0
2	10	20	40
3	10	30	90
	20	50	130

da cui segue

$$\text{Var}(Y|x_1) = \frac{130}{20} - 2.5^2 = 6.5 - 6.25 = 0.25.$$

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.



Per effettuare confronti in termini di variabilità tra caratteri quantitativi occorre utilizzare una misura relativa, ad esempio il coefficiente di variazione

$$CV = \frac{\sigma}{\mu} = \frac{\sqrt{\text{Var}(Y|x_1)}}{M(Y|x_1)} = \frac{\sqrt{0.25}}{2.5} = \frac{0.5}{2.5} = 0.2.$$

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.



Riepilogo

$Y|x_1$ Indice di posizione: $M(Y|x_1) = 1$ Indice di variabilità: $CV(Y|x_1) = 0$.

$Y|x_2$ Indice di posizione: $M(Y|x_2) = 2.5$ Indice di variabilità: $CV(Y|x_2) = 0.2$.

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.



4. Si determini il grado di connessione tra le variabili X e Y mediante il calcolo di un opportuno indice.

$\chi_N = 1$ dal momento che sussiste dipendenza funzionale di X da Y (le distribuzioni condizionate $X|y_j$ sono degeneri).

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.



La seguente tabella riporta la serie dei NIBF del fatturato annuo di un'ipotetica azienda:

t	$NIBF$
2012	1
2013	1.1
2014	1.045
2015	1.254
2016	1.254

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.



1. Calcolare, mostrandone il procedimento, la serie dei NIBM e interpretare il valore di $2012/I_{2013}$, numero indice al tempo 2013 con base 2012.

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.



1. Calcolare, mostrandone il procedimento, la serie dei NIBM e interpretare il valore di ${}_{2012}I_{2013}$, numero indice al tempo 2013 con base 2012.

Si ricorda che ${}_{t-1}I_t = \frac{x_t}{x_{t-1}} = \frac{{}_{2012}I_t}{{}_{2012}I_{t-1}}$.

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.



Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.

1. Calcolare, mostrandone il procedimento, la serie dei NIBM e interpretare il valore di ${}_{2012}I_{2013}$, numero indice al tempo 2013 con base 2012.

Si ricorda che ${}_{t-1}I_t = \frac{x_t}{x_{t-1}} = \frac{{}_{2012}I_t}{{}_{2012}I_{t-1}}$. Si ottiene quindi

<i>tempo</i>	<i>NIBF</i>	<i>NIBM</i>
2012	1.000	
2013	1.100	1.10
2014	1.045	0.95
2015	1.254	1.20
2016	1.254	1.00

${}_{2012}I_{2013} = 1.1$. Tra il 2012 e il 2013 il fatturato ha registrato un incremento del 10%.



Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.

2. Sapendo che il fatturato nel 2016 era 100000 Euro, calcolare il fatturato nel 2013.

$$x_{2013} = 87719.30.$$

Si può arrivare al risultato in diversi modi

- da

$${}_{2012}l_{2013} = \frac{x_{2013}}{x_{2012}} = 1.1$$

e

$${}_{2012}l_{2016} = \frac{x_{2016}}{x_{2012}} = 1.254$$

si ottiene

$$x_{2013} = x_{2016} \cdot \frac{{}_{2012}l_{2013}}{{}_{2012}l_{2016}} = 100000 \cdot \frac{1.1}{1.254}$$

oppure



Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.

- da

$${}_{2013}I_{2016} = \frac{{}_{2012}I_{2016}}{{}_{2012}I_{2013}} = \frac{1.254}{1.1}$$

e

$${}_{2013}I_{2016} = \frac{x_{2016}}{x_{2013}} = 1.14$$

abbiamo

$$x_{2013} = \frac{x_{2016}}{{}_{2013}I_{2016}} = \frac{100000}{1.14}$$

oppure



- da

$$x_{2013} = \frac{\left(\frac{\left(\frac{x_{2016}}{2015} \right)_{2016}}{2014} \right)_{2015}}{2013} \bigg/_{2014} = \frac{\left(\frac{\left(\frac{100000}{1.254} \right)}{1.254} \right)}{1.045}$$

oppure

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.



Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.

- da

abbiamo ancora

$$X_{2013} : X_{2016} = {}_{2012}I_{2013} : {}_{2012}I_{2016}$$

$$X_{2013} = \frac{X_{2016} \cdot {}_{2012}I_{2013}}{{}_{2012}I_{2016}}$$



3. Si calcoli il tasso medio di variazione dal 2013 al 2016 e si commenti il risultato ottenuto.

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.



3. Si calcoli il tasso medio di variazione dal 2013 al 2016 e si commenti il risultato ottenuto.
Si ricorda che

$${}_{2013}I_{2016} = {}_{2013}I_{2014} \cdot {}_{2014}I_{2015} \cdot {}_{2015}I_{2016}$$

Si deve trovare il numero indice a base mobile medio.

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.



Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.

3. Si calcoli il tasso medio di variazione dal 2013 al 2016 e si commenti il risultato ottenuto. Si ricorda che

$${}_{2013}l_{2016} = {}_{2013}l_{2014} \cdot {}_{2014}l_{2015} \cdot {}_{2015}l_{2016}$$

Si deve trovare il numero indice a base mobile medio. Applicando il criterio di scelta della media secondo Chisini si ottiene

$${}_{2013}l_{2016} = {}_{2013}l_{2014} \cdot {}_{2014}l_{2015} \cdot {}_{2015}l_{2016} = \alpha \cdot \alpha \cdot \alpha = \alpha^3.$$



Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.

3. Si calcoli il tasso medio di variazione dal 2013 al 2016 e si commenti il risultato ottenuto. Si ricorda che

$${}_{2013}I_{2016} = {}_{2013}I_{2014} \cdot {}_{2014}I_{2015} \cdot {}_{2015}I_{2016}$$

Si deve trovare il numero indice a base mobile medio. Applicando il criterio di scelta della media secondo Chisini si ottiene

$${}_{2013}I_{2016} = {}_{2013}I_{2014} \cdot {}_{2014}I_{2015} \cdot {}_{2015}I_{2016} = \alpha \cdot \alpha \cdot \alpha = \alpha^3.$$

Quindi

$$\alpha = \sqrt[3]{{}_{2013}I_{2016}} = \sqrt[3]{{}_{2013}I_{2014} \cdot {}_{2014}I_{2015} \cdot {}_{2015}I_{2016}} = 1.0446.$$

Da cui segue una variazione relativa media pari a 0.0446, corrispondente a un incremento medio annuo tra il 2013 e il 2016 del 4.46%.



Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.

Il seguente prospetto riporta i valori rilevati in corrispondenza a 5 unità statistiche con riferimento a due caratteri quantitativi X e Y

x_i	2	2	4	6	6
y_i	1	2	3	1	2

1. Si rappresentino graficamente le coppie dei dati e la funzione di regressione ipotizzando Y in funzione di X .



La funzione di regressione associa a ogni valore x appartenente all'insieme di definizione della variabile X la media di Y condizionata a x , $M(Y|x)$

x	$M(Y x)$
2	$\frac{1}{2}(1 + 2) = 1.5$
4	3
6	$\frac{1}{2}(1 + 2) = 1.5$

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.



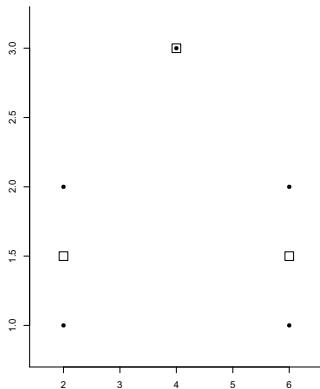
Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.

Le coppie di dati possono essere rappresentate mediante un grafico a dispersione.





2. Si calcolino, secondo il metodo dei minimi quadrati, i parametri dei seguenti modelli di regressione:

$$I) Y^* = a + b \cdot X \qquad II) Y^* = c$$

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.



Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.

Si riportano i calcoli necessari¹ per ottenere le stime dei parametri e gli indici di adattamento del modello

$$Y = a + bX + E$$

x_i	y_i	$x_i y_i$	x_i^2	y_i^2
2	1	2	4	1
2	2	4	4	4
4	3	12	16	9
6	1	6	36	1
6	2	12	36	4
20	9	36	96	19
4	1.8	7.2	19.2	3.8
$M(X)$	$M(Y)$	$M(XY)$	$M(X^2)$	$M(Y^2)$

$$\hat{b} =$$

¹Si osserva che per ragioni di stampa alcuni calcoli potrebbero essere duplicati con riferimento ai diversi modelli oggetto di studio



Si riportano i calcoli necessari¹ per ottenere le stime dei parametri e gli indici di adattamento del modello

$$Y = a + bX + E$$

x_i	y_i	$x_i y_i$	x_i^2	y_i^2
2	1	2	4	1
2	2	4	4	4
4	3	12	16	9
6	1	6	36	1
6	2	12	36	4
20	9	36	96	19
4	1.8	7.2	19.2	3.8
$M(X)$	$M(Y)$	$M(XY)$	$M(X^2)$	$M(Y^2)$

$$\begin{aligned}\hat{b} &= \frac{\text{Cov}(X, Y)}{\text{Var}(X)} = \frac{M(XY) - M(X)M(Y)}{M(X^2) - [M(X)^2]} = \\ &= \frac{7.2 - 4 \cdot 1.8}{19.2 - 4^2} = \frac{0}{3.2} = 0 \\ \hat{a} &= \end{aligned}$$

¹Si osserva che per ragioni di stampa alcuni calcoli potrebbero essere duplicati con riferimento ai diversi modelli oggetto di studio



Si riportano i calcoli necessari¹ per ottenere le stime dei parametri e gli indici di adattamento del modello

$$Y = a + bX + E$$

x_i	y_i	$x_i y_i$	x_i^2	y_i^2
2	1	2	4	1
2	2	4	4	4
4	3	12	16	9
6	1	6	36	1
6	2	12	36	4
20	9	36	96	19
4	1.8	7.2	19.2	3.8
$M(X)$	$M(Y)$	$M(XY)$	$M(X^2)$	$M(Y^2)$

$$\begin{aligned}\hat{b} &= \frac{\text{Cov}(X, Y)}{\text{Var}(X)} = \frac{M(XY) - M(X)M(Y)}{M(X^2) - [M(X)]^2} = \\ &= \frac{7.2 - 4 \cdot 1.8}{19.2 - 4^2} = \frac{0}{3.2} = 0 \\ \hat{a} &= M(Y) - \hat{b}M(X) = 1.8 \\ \rho^2 &= \end{aligned}$$

¹Si osserva che per ragioni di stampa alcuni calcoli potrebbero essere duplicati con riferimento ai diversi modelli oggetto di studio



Si riportano i calcoli necessari¹ per ottenere le stime dei parametri e gli indici di adattamento del modello

$$Y = a + bX + E$$

x_i	y_i	$x_i y_i$	x_i^2	y_i^2
2	1	2	4	1
2	2	4	4	4
4	3	12	16	9
6	1	6	36	1
6	2	12	36	4
20	9	36	96	19
4	1.8	7.2	19.2	3.8
$M(X)$	$M(Y)$	$M(XY)$	$M(X^2)$	$M(Y^2)$

$$\begin{aligned}\hat{b} &= \frac{\text{Cov}(X, Y)}{\text{Var}(X)} = \frac{M(XY) - M(X)M(Y)}{M(X^2) - [M(X)]^2} = \\ &= \frac{7.2 - 4 \cdot 1.8}{19.2 - 4^2} = \frac{0}{3.2} = 0 \\ \hat{a} &= M(Y) - \hat{b}M(X) = 1.8 \\ \rho^2 &= \frac{[\text{Cov}(X, Y)]^2}{\text{Var}(X)\text{Var}(Y)} = \frac{0^2}{3.2 \cdot 0.56} = 0\end{aligned}$$

¹Si osserva che per ragioni di stampa alcuni calcoli potrebbero essere duplicati con riferimento ai diversi modelli oggetto di studio



Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.

Si riportano i calcoli necessari² per ottenere le stime dei parametri e gli indici di adattamento del modello

$$Y = c + E$$

$1/x_i$	y_i	y_i^2
2	1	1
2	2	4
4	3	9
6	1	1
6	2	4
9		19
1.8		3.8
$M(Y)$		$M(Y^2)$

$$\hat{c} =$$

²Si osserva che per ragioni di stampa alcuni calcoli potrebbero essere duplicati con riferimento ai diversi modelli oggetto di studio



Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.

Si riportano i calcoli necessari² per ottenere le stime dei parametri e gli indici di adattamento del modello

$$Y = c + E$$

$1/x_i$	y_i	y_i^2
2	1	1
2	2	4
4	3	9
6	1	1
6	2	4
9		19
1.8		3.8
$M(Y)$		$M(Y^2)$

$$\hat{c} = M(Y) = 1.8$$

$$R^2 =$$

²Si osserva che per ragioni di stampa alcuni calcoli potrebbero essere duplicati con riferimento ai diversi modelli oggetto di studio



Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.

Si riportano i calcoli necessari² per ottenere le stime dei parametri e gli indici di adattamento del modello

$$Y = c + E$$

$1/x_i$	y_i	y_i^2
2	1	1
2	2	4
4	3	9
6	1	1
6	2	4
9		19
1.8		3.8
$M(Y)$		$M(Y^2)$

$$\hat{c} = M(Y) = 1.8$$

$$R^2 = 0$$

²Si osserva che per ragioni di stampa alcuni calcoli potrebbero essere duplicati con riferimento ai diversi modelli oggetto di studio



Riepilogo

$\hat{a} = 1.8$, $\hat{b} = 0$, indice adattamento $M(\hat{E}^2) = 0.56$ $\hat{c} = 1.8$, indice adattamento
 $M(\hat{E}^2) = 0.56$

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.



3. Confrontare gli indici di adattamento, commentando i risultati.

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.



Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.

3. Confrontare gli indici di adattamento, commentando i risultati.
I due modelli sono equivalenti.



4. Si motivi perché $\eta_{X|Y}^2 < \eta_{Y|X}^2$.

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.



4. Si motivi perché $\eta_{X|Y}^2 < \eta_{Y|X}^2$.

$\eta_{X|Y}^2 < \eta_{Y|X}^2$ perché $0 = \eta_{X|Y}^2 < \eta_{Y|X}^2 < 1$.

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.



Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.

4. Si motivi perché $\eta_{X|Y}^2 < \eta_{Y|X}^2$.

$\eta_{X|Y}^2 < \eta_{Y|X}^2$ perché $0 = \eta_{X|Y}^2 < \eta_{Y|X}^2 < 1$.

Abbiamo, infatti,

$$M(X|y = 1) = M(X|y = 2) = M(X|y = 3) = 4,$$

mentre vale

$$M(Y|x = 2) = M(Y|x = 6) = 1.5 \neq M(Y|x = 4) = 3.$$

Esercizio 4

Prova scritta di
STATISTICA del
06.02.2017

Giuseppe Boari
Gabriele Cantaluppi



UNIVERSITÀ
CATTOLICA
del Sacro Cuore

L'urna $U1$ contiene 3 palline nere e 1 pallina rossa. L'urna $U2$ contiene 9 palline nere e 10 palline rosse.

1. Si calcoli la probabilità che estraendo una pallina dall'urna $U1$ e una pallina dall'urna $U2$ solo una delle due palline sia nera (si ottengano una pallina nera e una pallina rossa).

[Esercizio 1.](#)

[Esercizio 2.](#)

[Esercizio 3.](#)

[Esercizio 4.](#)

Esercizio 4



L'urna $U1$ contiene 3 palline nere e 1 pallina rossa. L'urna $U2$ contiene 9 palline nere e 10 palline rosse.

1. Si calcoli la probabilità che estraendo una pallina dall'urna $U1$ e una pallina dall'urna $U2$ solo una delle due palline sia nera (si ottengano una pallina nera e una pallina rossa).

$$P(1 \text{ rossa e } 1 \text{ nera}) =$$

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.



L'urna U_1 contiene 3 palline nere e 1 pallina rossa. L'urna U_2 contiene 9 palline nere e 10 palline rosse.

1. Si calcoli la probabilità che estraendo una pallina dall'urna U_1 e una pallina dall'urna U_2 solo una delle due palline sia nera (si ottengano una pallina nera e una pallina rossa).

$$P(1 \text{ rossa e } 1 \text{ nera}) = P\{(N_1 \cap R_2) \cup (R_1 \cap N_2)\} =$$
$$=$$

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.



L'urna $U1$ contiene 3 palline nere e 1 pallina rossa. L'urna $U2$ contiene 9 palline nere e 10 palline rosse.

1. Si calcoli la probabilità che estraendo una pallina dall'urna $U1$ e una pallina dall'urna $U2$ solo una delle due palline sia nera (si ottengano una pallina nera e una pallina rossa).

$$\begin{aligned} P(1 \text{ rossa e } 1 \text{ nera}) &= P\{(N_1 \cap R_2) \cup (R_1 \cap N_2)\} = \\ &= P(N_1 \cap R_2) + P(R_1 \cap N_2) = \\ &= \end{aligned}$$

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.



L'urna $U1$ contiene 3 palline nere e 1 pallina rossa. L'urna $U2$ contiene 9 palline nere e 10 palline rosse.

1. Si calcoli la probabilità che estraendo una pallina dall'urna $U1$ e una pallina dall'urna $U2$ solo una delle due palline sia nera (si ottengano una pallina nera e una pallina rossa).

$$\begin{aligned}P(1 \text{ rossa e } 1 \text{ nera}) &= P\{(N_1 \cap R_2) \cup (R_1 \cap N_2)\} = \\&= P(N_1 \cap R_2) + P(R_1 \cap N_2) = \\&= 0.3947 + 0.1184 = 0.5132.\end{aligned}$$

[Esercizio 1.](#)

[Esercizio 2.](#)

[Esercizio 3.](#)

[Esercizio 4.](#)



2. Si calcoli la probabilità che estraendo senza reimmissione 4 palline dall'urna U_2 ce ne siano almeno 3 nere.

$$P(\text{almeno } 3N) =$$

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.



2. Si calcoli la probabilità che estraendo senza reimmissione 4 palline dall'urna U_2 ce ne siano almeno 3 nere.

$$P(\text{almeno } 3N) = \binom{4}{3} \frac{9}{19} \frac{8}{18} \frac{7}{17} \frac{10}{16} +$$

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.



2. Si calcoli la probabilità che estraendo senza reimmissione 4 palline dall'urna U_2 ce ne siano almeno 3 nere.

$$P(\text{almeno } 3N) = \binom{4}{3} \frac{9}{19} \frac{8}{18} \frac{7}{17} \frac{10}{16} + \frac{9}{19} \frac{8}{18} \frac{7}{17} \frac{6}{16} =$$
$$=$$

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.



Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.

2. Si calcoli la probabilità che estraendo senza reimmissione 4 palline dall'urna U_2 ce ne siano almeno 3 nere.

$$\begin{aligned} P(\text{almeno } 3N) &= \binom{4}{3} \frac{9}{19} \frac{8}{18} \frac{7}{17} \frac{10}{16} + \frac{9}{19} \frac{8}{18} \frac{7}{17} \frac{6}{16} = \\ &= 0.2167 + 0.0325 = \end{aligned}$$



Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.

2. Si calcoli la probabilità che estraendo senza reimmissione 4 palline dall'urna U_2 ce ne siano almeno 3 nere.

$$\begin{aligned} P(\text{almeno } 3N) &= \binom{4}{3} \frac{9}{19} \frac{8}{18} \frac{7}{17} \frac{10}{16} + \frac{9}{19} \frac{8}{18} \frac{7}{17} \frac{6}{16} = \\ &= 0.2167 + 0.0325 = 0.2492 \end{aligned}$$

o mediante la v.c. ipergeometrica

$$P(\text{almeno } 3N) =$$



Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.

2. Si calcoli la probabilità che estraendo senza reimmissione 4 palline dall'urna U_2 ce ne siano almeno 3 nere.

$$\begin{aligned} P(\text{almeno } 3N) &= \binom{4}{3} \frac{9}{19} \frac{8}{18} \frac{7}{17} \frac{10}{16} + \frac{9}{19} \frac{8}{18} \frac{7}{17} \frac{6}{16} = \\ &= 0.2167 + 0.0325 = 0.2492 \end{aligned}$$

o mediante la v.c. ipergeometrica

$$P(\text{almeno } 3N) = \frac{\binom{9}{3} \binom{10}{1}}{\binom{19}{4}} + \frac{\binom{9}{4} \binom{10}{0}}{\binom{19}{4}} =$$



Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.

2. Si calcoli la probabilità che estraendo senza reimmissione 4 palline dall'urna $U2$ ce ne siano almeno 3 nere.

$$\begin{aligned} P(\text{almeno } 3N) &= \binom{4}{3} \frac{9}{19} \frac{8}{18} \frac{7}{17} \frac{10}{16} + \frac{9}{19} \frac{8}{18} \frac{7}{17} \frac{6}{16} = \\ &= 0.2167 + 0.0325 = 0.2492 \end{aligned}$$

o mediante la v.c. ipergeometrica

$$P(\text{almeno } 3N) = \frac{\binom{9}{3} \binom{10}{1}}{\binom{19}{4}} + \frac{\binom{9}{4} \binom{10}{0}}{\binom{19}{4}} = 0.2167 + 0.0325 =$$



Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.

2. Si calcoli la probabilità che estraendo senza reimmissione 4 palline dall'urna U_2 ce ne siano almeno 3 nere.

$$\begin{aligned} P(\text{almeno } 3N) &= \binom{4}{3} \frac{9}{19} \frac{8}{18} \frac{7}{17} \frac{10}{16} + \frac{9}{19} \frac{8}{18} \frac{7}{17} \frac{6}{16} = \\ &= 0.2167 + 0.0325 = 0.2492 \end{aligned}$$

o mediante la v.c. ipergeometrica

$$P(\text{almeno } 3N) = \frac{\binom{9}{3} \binom{10}{1}}{\binom{19}{4}} + \frac{\binom{9}{4} \binom{10}{0}}{\binom{19}{4}} = 0.2167 + 0.0325 = 0.2492.$$



3. Si estragga una pallina dall'urna $U1$ e la si metta nell'urna $U2$. Si calcoli la probabilità che estraendo con reimmissione 2 palline dall'urna $U2$ si ottengano 2 palline nere.

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.



3. Si estragga una pallina dall'urna $U1$ e la si metta nell'urna $U2$. Si calcoli la probabilità che estraendo con reimmissione 2 palline dall'urna $U2$ si ottengano 2 palline nere.
In base alla legge delle probabilità totali abbiamo

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.



3. Si estragga una pallina dall'urna $U1$ e la si metta nell'urna $U2$. Si calcoli la probabilità che estraendo con reimmissione 2 palline dall'urna $U2$ si ottengano 2 palline nere.

In base alla legge delle probabilità totali abbiamo

$$P(2N) = P((R_1 \cap 2N) \cup (N_1 \cap 2N)) =$$
$$=$$

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.



3. Si estragga una pallina dall'urna U_1 e la si metta nell'urna U_2 . Si calcoli la probabilità che estraendo con reimmissione 2 palline dall'urna U_2 si ottengano 2 palline nere.

In base alla legge delle probabilità totali abbiamo

$$\begin{aligned} P(2N) &= P((R_1 \cap 2N) \cup (N_1 \cap 2N)) = \\ &= P(R_1 \cap 2N) + P(N_1 \cap 2N) = \\ &= \end{aligned}$$

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.



Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.

3. Si estragga una pallina dall'urna $U1$ e la si metta nell'urna $U2$. Si calcoli la probabilità che estraendo con reimmissione 2 palline dall'urna $U2$ si ottengano 2 palline nere.

In base alla legge delle probabilità totali abbiamo

$$\begin{aligned} P(2N) &= P((R_1 \cap 2N) \cup (N_1 \cap 2N)) = \\ &= P(R_1 \cap 2N) + P(N_1 \cap 2N) = \\ &= P(R_1)P(2N|R_1) + P(N_1)P(2N|N_1) = \\ &= \end{aligned}$$



Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.

3. Si estragga una pallina dall'urna $U1$ e la si metta nell'urna $U2$. Si calcoli la probabilità che estraendo con reimmissione 2 palline dall'urna $U2$ si ottengano 2 palline nere.

In base alla legge delle probabilità totali abbiamo

$$\begin{aligned} P(2N) &= P((R_1 \cap 2N) \cup (N_1 \cap 2N)) = \\ &= P(R_1 \cap 2N) + P(N_1 \cap 2N) = \\ &= P(R_1)P(2N|R_1) + P(N_1)P(2N|N_1) = \\ &= \frac{1}{4} \left(\frac{9}{20} \right)^2 + \frac{3}{4} \left(\frac{10}{20} \right)^2 = \\ &= \end{aligned}$$



Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.

3. Si estragga una pallina dall'urna $U1$ e la si metta nell'urna $U2$. Si calcoli la probabilità che estraendo con reimmissione 2 palline dall'urna $U2$ si ottengano 2 palline nere.

In base alla legge delle probabilità totali abbiamo

$$\begin{aligned}P(2N) &= P((R_1 \cap 2N) \cup (N_1 \cap 2N)) = \\&= P(R_1 \cap 2N) + P(N_1 \cap 2N) = \\&= P(R_1)P(2N|R_1) + P(N_1)P(2N|N_1) = \\&= \frac{1}{4} \left(\frac{9}{20} \right)^2 + \frac{3}{4} \left(\frac{10}{20} \right)^2 = \\&= 0.25 \cdot 0.2025 + 0.75 \cdot 0.25 = \\&= \end{aligned}$$



Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.

3. Si estragga una pallina dall'urna $U1$ e la si metta nell'urna $U2$. Si calcoli la probabilità che estraendo con reimmissione 2 palline dall'urna $U2$ si ottengano 2 palline nere.

In base alla legge delle probabilità totali abbiamo

$$\begin{aligned} P(2N) &= P((R_1 \cap 2N) \cup (N_1 \cap 2N)) = \\ &= P(R_1 \cap 2N) + P(N_1 \cap 2N) = \\ &= P(R_1)P(2N|R_1) + P(N_1)P(2N|N_1) = \\ &= \frac{1}{4} \left(\frac{9}{20} \right)^2 + \frac{3}{4} \left(\frac{10}{20} \right)^2 = \\ &= 0.25 \cdot 0.2025 + 0.75 \cdot 0.25 = \\ &= 0.050625 + 0.1875 = 0.238125 \end{aligned}$$

Tema T293

Prova scritta di STATISTICA del 21.02.2017

Prova scritta di
STATISTICA del
21.02.2017

Giuseppe Boari
Gabriele Cantaluppi



UNIVERSITÀ
CATTOLICA
del Sacro Cuore

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.

Giuseppe Boari
Gabriele Cantaluppi
Università Cattolica del Sacro Cuore, Milano



Una indagine su 210 cittadini riporta la categoria di età X e la spesa media mensile per l'attività sportiva Y . Le distribuzioni delle *frequenze relative* di X e delle *densità di frequenza* di Y sono le seguenti:

x_i	f_i
<i>giovani</i>	0.50
<i>adulti</i>	0.30
<i>senior</i>	0.20

$h_{j-1} \vdash h_j$	d_j
5 $\vdash$ 10	10
10 $\vdash$ 20	11
20 $\vdash$ 25	10

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.



1. Si costruiscano le distribuzioni di frequenza assolute delle due variabili.

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.



Moltiplicando le frequenze relative f_i per il numero totale di osservazioni n si ottengono le frequenze assolute n_i della variabile X .

x_i	f_i	n_i
<i>giovani</i>	0.5	105
<i>adulti</i>	0.3	63
<i>senior</i>	0.2	42

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.



Moltiplicando le densità di frequenza, d_i , per le ampiezze di classe, a_i , si ottengono le frequenze assolute n_i della variabile X ,

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.



Moltiplicando le densità di frequenza, d_i , per le ampiezze di classe, a_i , si ottengono le frequenze assolute n_i della variabile X ,

x_i	d_i	a_i	n_i
5 $\rightarrow$ 10	10	10 - 5 = 5	50
10 $\rightarrow$ 20	11	20 - 10 = 10	110
20 $\rightarrow$ 25	10	25 - 20 = 5	50

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.



2. Si rappresentino graficamente le due variabili e si riassumano le stesse con opportuni indici di posizione e di variabilità.

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.



2. Si rappresentino graficamente le due variabili e si riassumano le stesse con opportuni indici di posizione e di variabilità.

Si tratta di un carattere qualitativo sconnesso che può essere rappresentato graficamente mediante un grafico a bastoncini o a canne d'organo.

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.

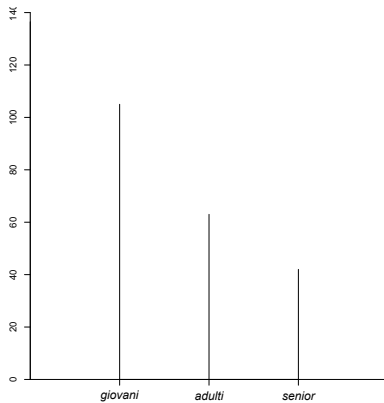


Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.





Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.

Si tratta di un carattere qualitativo sconnesso.

L'unico indice di posizione che può essere definito è la moda (categoria alla quale è associata la massima frequenza) che risulta *giovani*.

In presenza di un carattere qualitativo la variabilità può essere misurata mediante gli indici di Gini e di Frosini normalizzati

$$G_N = \frac{1 - \sum_{i=1}^k f_i^2}{1 - \frac{1}{k}}$$

e

$$F_N = 1 - \sqrt{1 - G_N}.$$



Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.



In presenza di un carattere qualitativo la variabilità può essere misurata mediante gli indici di Gini e di Frosini normalizzati

$$G_N = \frac{1 - \sum_{i=1}^k f_i^2}{1 - \frac{1}{k}}$$

e

$$F_N = 1 - \sqrt{1 - G_N}.$$

Si ottiene:

x_i	n_i	f_i	f_i^2
<i>giovani</i>	105	0.5	0.25
<i>adulti</i>	63	0.3	0.09
<i>senior</i>	42	0.2	0.04
	210	1	0.38

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.



In presenza di un carattere qualitativo la variabilità può essere misurata mediante gli indici di Gini e di Frosini normalizzati

$$G_N = \frac{1 - \sum_{i=1}^k f_i^2}{1 - \frac{1}{k}}$$

e

$$F_N = 1 - \sqrt{1 - G_N}.$$

Si ottiene:

x_i	n_i	f_i	f_i^2
<i>giovani</i>	105	0.5	0.25
<i>adulti</i>	63	0.3	0.09
<i>senior</i>	42	0.2	0.04
	210	1	0.38

Da cui segue:

$$G_N = \frac{1 - 0.38}{1 - \frac{1}{3}} = 0.93$$

e

$$F_N = 1 - \sqrt{1 - 0.93} = 0.7354.$$

Sussiste, quindi, un livello medio alto di eterogeneità.



Si considera la distribuzione della variabile X ,

x_i	n_i
5 $\vdash$ 10	50
10 $\vdash$ 20	110
20 $\vdash$ 25	50

Esercizio 1.

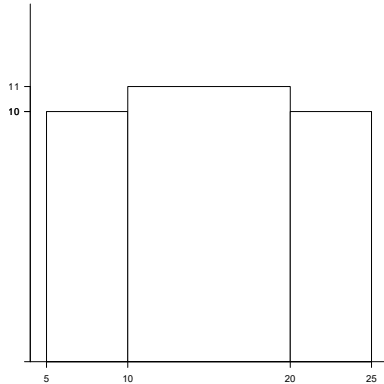
Esercizio 2.

Esercizio 3.

Esercizio 4.



Si tratta di un carattere quantitativo con modalità raggruppate in classi che può essere rappresentato graficamente mediante un istogramma.



Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.



Si tratta di un carattere quantitativo con modalità raggruppate in classi.

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.



La media aritmetica

$$M(X) = \frac{1}{n} \sum_{i=1}^k x_i n_i$$

risulta 15.

Si osserva, infatti, che la distribuzione è simmetrica.

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.



La varianza

$$\text{Var}(X) = M[(X - \mu_X)^2] = M(X^2) - \mu_X^2 = \frac{1}{n} \sum_{i=1}^k x_i^2 n_i - \mu_X^2$$

risulta 26.7857.

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.



La varianza

$$\text{Var}(X) = M[(X - \mu_X)^2] = M(X^2) - \mu_X^2 = \frac{1}{n} \sum_{i=1}^k x_i^2 n_i - \mu_X^2$$

risulta 26.7857.

Si ottiene (considerando come modalità i valori centrali delle classi)

x_i	n_i	$x_i n_i$	$x_i^2 n_i$
7.5	50	375	2812.5
15.0	110	1650	24750.0
22.5	50	1125	25312.5
	210	3150	52875.0

da cui segue

$$\begin{aligned} \text{Var}(X) = \sigma_X^2 &= M(X^2) - \mu_X^2 = \\ &= \frac{52875}{210} - 15^2 = 251.7857 - 225 = 26.7857. \end{aligned}$$

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.



Per effettuare confronti in termini di variabilità tra caratteri quantitativi occorre utilizzare una misura relativa, ad esempio il coefficiente di variazione

$$CV = \frac{\sigma}{\mu} = \frac{\sqrt{\text{Var}(X)}}{M(X)} = \frac{\sqrt{26.7857}}{15} = \frac{5.1755}{15} = 0.345.$$

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.



Riepilogo

X Indice di posizione: $\text{Moda}(X) = \text{giovani}$ Indice di variabilità: $F_N = 0.7354$.

Y Indice di posizione: $M(Y) = 15$ Indice di variabilità: $CV(Y) = 0.3450$.

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.



3. Si calcoli l'indice di asimmetria della variabile Y .

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.



Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.

3. Si calcoli l'indice di asimmetria della variabile Y .

La distribuzione è simmetrica, quindi gli indici di asimmetria assumono valore nullo



4. Si costruisca la tabella delle frequenze congiunte delle due variabili nel caso di indipendenza stocastica e si dica invece quanto varrebbe l'indice di connessione χ^2 nel caso di massima dipendenza.

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.



4. Si costruisca la tabella delle frequenze congiunte delle due variabili nel caso di indipendenza stocastica e si dica invece quanto varrebbe l'indice di connessione χ^2 nel caso di massima dipendenza.

Le frequenze congiunte nella situazione di indipendenza stocastica possono essere ottenute come

$$\hat{n}_{ij} = \frac{n_{i\bullet} n_{\bullet j}}{n}$$

Abbiamo, quindi,

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.



4. Si costruisca la tabella delle frequenze congiunte delle due variabili nel caso di indipendenza stocastica e si dica invece quanto varrebbe l'indice di connessione χ^2 nel caso di massima dipendenza.

Le frequenze congiunte nella situazione di indipendenza stocastica possono essere ottenute come

$$\hat{n}_{ij} = \frac{n_{i\bullet} n_{\bullet j}}{n}$$

Abbiamo, quindi,

$X \setminus Y$	5 - 10	10 - 20	20 - 25	$n_{i\bullet}$
<i>giovani</i>	25	55	25	105
<i>adulti</i>	15	33	15	63
<i>senior</i>	10	22	10	42
$n_{\bullet j}$	50	110	50	210

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.



4. Si costruisca la tabella delle frequenze congiunte delle due variabili nel caso di indipendenza stocastica e si dica invece quanto varrebbe l'indice di connessione χ^2 nel caso di massima dipendenza.

Le frequenze congiunte nella situazione di indipendenza stocastica possono essere ottenute come

$$\hat{n}_{ij} = \frac{n_{i\bullet} n_{\bullet j}}{n}$$

Abbiamo, quindi,

$X \setminus Y$	5 - 10	10 - 20	20 - 25	$n_{i\bullet}$
<i>giovani</i>	25	55	25	105
<i>adulti</i>	15	33	15	63
<i>senior</i>	10	22	10	42
$n_{\bullet j}$	50	110	50	210

In caso di massima dipendenza abbiamo

$$\chi^2 = n \cdot \min(h - 1, k - 1) = 420$$



Con riferimento alla seguente serie di Numeri Indici a Base Mobile:

t	$NIBM$
2008	
2009	0.92
2010	0.95
2011	1.05
2012	1.08

1. Si ricostruisca, mostrandone il procedimento, la serie storica dei dati x_t sapendo che $x_{2010} = 91$.

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.



Si ricorda che $x_{2010} = 91$, che per valori di $t > 2010$ vale

$$x_t = x_{2010} \cdot {}_{2010}l_{2011} \cdot {}_{2011}l_{2012} \cdot \dots \cdot {}_{t-1}l_t$$

e che per valori di $t < 2010$ vale

$$x_t = \frac{x_{2010}}{{}_t l_{t+1} \cdot \dots \cdot {}_{2008}l_{2009} \cdot {}_{2009}l_{2010}}.$$

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.



Si ricorda che $x_{2010} = 91$, che per valori di $t > 2010$ vale

$$x_t = x_{2010} \cdot {}_{2010}l_{2011} \cdot {}_{2011}l_{2012} \cdot \dots \cdot {}_{t-1}l_t$$

e che per valori di $t < 2010$ vale

$$x_t = \frac{x_{2010}}{{}_t l_{t+1} \cdot \dots \cdot {}_{2008}l_{2009} \cdot {}_{2009}l_{2010}}.$$

Si ottiene quindi

Anno	NIBM	x_t
2008	0.80	104.1190
2009	0.92	95.7895
2010	0.95	91.0000
2011	1.05	95.5500
2012	1.08	103.1940

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.



2. Si calcoli il numero indice al tempo 2012 con base 2010 e si commenti il risultato ottenuto.

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.



2. Si calcoli il numero indice al tempo 2012 con base 2010 e si commenti il risultato ottenuto.
Si ricorda che

$${}_{2010}I_{2012} = {}_{2010}I_{2011} \cdot {}_{2011}I_{2012} = 1.134.$$

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.



Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.

2. Si calcoli il numero indice al tempo 2012 con base 2010 e si commenti il risultato ottenuto. Si ricorda che

$${}_{2010}I_{2012} = {}_{2010}I_{2011} \cdot {}_{2011}I_{2012} = 1.134.$$

Tra il tempo 2010 e il tempo 2012 si è registrato un incremento complessivo del 13.4% della grandezza oggetto di studio.



3. Si calcoli il tasso medio di variazione nel periodo 2008-2010 interpretando il risultato ottenuto.

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.



Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.

3. Si calcoli il tasso medio di variazione nel periodo 2008-2010 interpretando il risultato ottenuto.

Si ricorda che

$${}_{2008}I_{2010} = {}_{2008}I_{2009} \cdot {}_{2009}I_{2010}.$$

Si deve trovare il numero indice a base mobile medio.



Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.

3. Si calcoli il tasso medio di variazione nel periodo 2008-2010 interpretando il risultato ottenuto.

Si ricorda che

$${}_{2008}I_{2010} = {}_{2008}I_{2009} \cdot {}_{2009}I_{2010}.$$

Si deve trovare il numero indice a base mobile medio. Applicando il criterio di scelta della media secondo Chisini si ottiene

$${}_{2008}I_{2010} = {}_{2008}I_{2009} \cdot {}_{2009}I_{2010} = \alpha \cdot \alpha = \alpha^2.$$



Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.

3. Si calcoli il tasso medio di variazione nel periodo 2008-2010 interpretando il risultato ottenuto.

Si ricorda che

$${}_{2008}I_{2010} = {}_{2008}I_{2009} \cdot {}_{2009}I_{2010}.$$

Si deve trovare il numero indice a base mobile medio. Applicando il criterio di scelta della media secondo Chisini si ottiene

$${}_{2008}I_{2010} = {}_{2008}I_{2009} \cdot {}_{2009}I_{2010} = \alpha \cdot \alpha = \alpha^2.$$

Quindi

$$\alpha = \sqrt[2]{{}_{2008}I_{2010}} = \sqrt[2]{{}_{2008}I_{2009} \cdot {}_{2009}I_{2010}} = 0.9349.$$

Da cui segue una variazione relativa media pari a -0.0651 , corrispondente a un decremento medio annuo tra il 2008 e il 2010 del 6.51%.



Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.

Il seguente prospetto riporta i valori rilevati in corrispondenza a 8 unità statistiche con riferimento a due caratteri quantitativi X e Y

x_i	1	1	2	2	3	4	4	5
y_i	10	10	13	13	18	20	20	20

1. Si rappresentino graficamente le coppie dei dati e la funzione di regressione ipotizzando Y in funzione di X .



La funzione di regressione associa a ogni valore x appartenente all'insieme di definizione della variabile X la media di Y condizionata a x , $M(Y|x)$

x	$M(Y x)$
1	10
2	13
3	18
4	20
5	20

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.



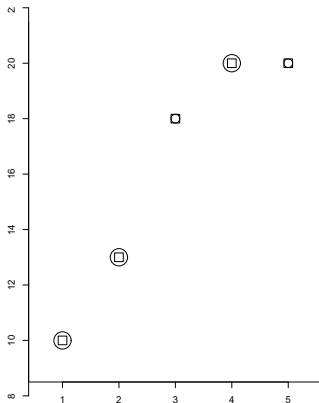
Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.

Le coppie di dati possono essere rappresentate mediante un *bubble plot diagram*.





2. Si calcolino, secondo il metodo dei minimi quadrati, i parametri dei seguenti modelli di regressione:

$$I) Y^* = a + b \cdot X \quad II) Y^* = \exp(cX)$$

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.

Si riportano i calcoli necessari¹ per ottenere le stime dei parametri e gli indici di adattamento del modello

$$Y = a + bX + E$$

x_i	y_i	$x_i y_i$	x_i^2	y_i^2
1	10	10	1	100
1	10	10	1	100
2	13	26	4	169
2	13	26	4	169
3	18	54	9	324
4	20	80	16	400
4	20	80	16	400
5	20	100	25	400
22	124	386	76	2062
2.75	15.5	48.25	9.5	257.75
$M(X)$	$M(Y)$	$M(XY)$	$M(X^2)$	$M(Y^2)$

$$\hat{b} =$$

¹Si osserva che per ragioni di stampa alcuni calcoli potrebbero essere duplicati con riferimento ai diversi modelli oggetto di studio



Si riportano i calcoli necessari¹ per ottenere le stime dei parametri e gli indici di adattamento del modello

$$Y = a + bX + E$$

x_i	y_i	$x_i y_i$	x_i^2	y_i^2
1	10	10	1	100
1	10	10	1	100
2	13	26	4	169
2	13	26	4	169
3	18	54	9	324
4	20	80	16	400
4	20	80	16	400
5	20	100	25	400
22	124	386	76	2062
2.75	15.5	48.25	9.5	257.75
$M(X)$	$M(Y)$	$M(XY)$	$M(X^2)$	$M(Y^2)$

$$\begin{aligned}\hat{b} &= \frac{\text{Cov}(X, Y)}{\text{Var}(X)} = \frac{M(XY) - M(X)M(Y)}{M(X^2) - [M(X)]^2} = \\ &= \frac{48.25 - 2.75 \cdot 15.5}{9.5 - 2.75^2} = \frac{5.625}{1.9375} = 2.9032\end{aligned}$$

$$\hat{a} =$$



¹Si osserva che per ragioni di stampa alcuni calcoli potrebbero essere duplicati con riferimento ai diversi modelli oggetto di studio

Si riportano i calcoli necessari¹ per ottenere le stime dei parametri e gli indici di adattamento del modello

$$Y = a + bX + E$$

x_i	y_i	$x_i y_i$	x_i^2	y_i^2
1	10	10	1	100
1	10	10	1	100
2	13	26	4	169
2	13	26	4	169
3	18	54	9	324
4	20	80	16	400
4	20	80	16	400
5	20	100	25	400
22	124	386	76	2062
2.75	15.5	48.25	9.5	257.75
$M(X)$	$M(Y)$	$M(XY)$	$M(X^2)$	$M(Y^2)$

$$\begin{aligned}\hat{b} &= \frac{\text{Cov}(X, Y)}{\text{Var}(X)} = \frac{M(XY) - M(X)M(Y)}{M(X^2) - [M(X)]^2} = \\ &= \frac{48.25 - 2.75 \cdot 15.5}{9.5 - 2.75^2} = \frac{5.625}{1.9375} = 2.9032\end{aligned}$$

$$\hat{a} = M(Y) - \hat{b} M(X) = 7.5161$$

$$\rho^2 =$$

¹Si osserva che per ragioni di stampa alcuni calcoli potrebbero essere duplicati con riferimento ai diversi modelli oggetto di studio



Si riportano i calcoli necessari¹ per ottenere le stime dei parametri e gli indici di adattamento del modello

$$Y = a + bX + E$$

x_i	y_i	$x_i y_i$	x_i^2	y_i^2
1	10	10	1	100
1	10	10	1	100
2	13	26	4	169
2	13	26	4	169
3	18	54	9	324
4	20	80	16	400
4	20	80	16	400
5	20	100	25	400
22	124	386	76	2062
2.75	15.5	48.25	9.5	257.75
$M(X)$	$M(Y)$	$M(XY)$	$M(X^2)$	$M(Y^2)$

$$\begin{aligned}\hat{b} &= \frac{\text{Cov}(X, Y)}{\text{Var}(X)} = \frac{M(XY) - M(X)M(Y)}{M(X^2) - [M(X)]^2} = \\ &= \frac{48.25 - 2.75 \cdot 15.5}{9.5 - 2.75^2} = \frac{5.625}{1.9375} = 2.9032\end{aligned}$$

$$\hat{a} = M(Y) - \hat{b} M(X) = 7.5161$$

$$\rho^2 = \frac{[\text{Cov}(X, Y)]^2}{\text{Var}(X) \text{Var}(Y)} = \frac{5.625^2}{1.9375 \cdot 17.5} = 0.9332$$

$$M(\hat{E}^2) = \text{Var}(\hat{E}) = \text{varianza residua} =$$

¹Si osserva che per ragioni di stampa alcuni calcoli potrebbero essere duplicati con riferimento ai diversi modelli oggetto di studio



Si riportano i calcoli necessari¹ per ottenere le stime dei parametri e gli indici di adattamento del modello

$$Y = a + bX + E$$

x_i	y_i	$x_i y_i$	x_i^2	y_i^2
1	10	10	1	100
1	10	10	1	100
2	13	26	4	169
2	13	26	4	169
3	18	54	9	324
4	20	80	16	400
4	20	80	16	400
5	20	100	25	400
22	124	386	76	2062
2.75	15.5	48.25	9.5	257.75
$M(X)$	$M(Y)$	$M(XY)$	$M(X^2)$	$M(Y^2)$

$$\begin{aligned}\hat{b} &= \frac{\text{Cov}(X, Y)}{\text{Var}(X)} = \frac{M(XY) - M(X)M(Y)}{M(X^2) - [M(X)]^2} = \\ &= \frac{48.25 - 2.75 \cdot 15.5}{9.5 - 2.75^2} = \frac{5.625}{1.9375} = 2.9032\end{aligned}$$

$$\hat{a} = M(Y) - \hat{b} M(X) = 7.5161$$

$$\rho^2 = \frac{[\text{Cov}(X, Y)]^2}{\text{Var}(X) \text{Var}(Y)} = \frac{5.625^2}{1.9375 \cdot 17.5} = 0.9332$$

$$M(\hat{E}^2) = \text{Var}(\hat{E}) = \text{varianza residua} = \text{Var}(Y) \cdot (1 - \rho^2) = 1.1694$$

¹Si osserva che per ragioni di stampa alcuni calcoli potrebbero essere duplicati con riferimento ai diversi modelli oggetto di studio





Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.

Si riportano i calcoli necessari² per ottenere le stime approssimate dei parametri e la misura di adattamento del modello

$$Y^* = \exp\{cX\}$$

stimato avvalendosi della trasformata logaritmica per ricondursi alla forma 'linearizzata'

$$\ln(Y) = cX + E$$

x_i	y_i	$\ln y_i$	$x_i \ln y_i$	x_i^2
1	10	2.3026	2.3026	1
1	10	2.3026	2.3026	1
2	13	2.5649	5.1298	4
2	13	2.5649	5.1298	4
3	18	2.8904	8.6712	9
4	20	2.9957	11.9828	16
4	20	2.9957	11.9828	16
5	20	2.9957	14.9785	25
		21.6125	62.4801	76
		2.7016	7.81	9.5
		$M(\ln Y)$	$M(X \ln Y)$	$M(X^2)$

$$\hat{c} =$$

²Si osserva che per ragioni di stampa alcuni calcoli potrebbero essere duplicati con riferimento ai diversi modelli oggetto di studio



Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.

Si riportano i calcoli necessari² per ottenere le stime approssimate dei parametri e la misura di adattamento del modello

$$Y^* = \exp\{cX\}$$

stimato avvalendosi della trasformata logaritmica per ricondursi alla forma 'linearizzata'

$$\ln(Y) = cX + E$$

x_i	y_i	$\ln y_i$	$x_i \ln y_i$	x_i^2
1	10	2.3026	2.3026	1
1	10	2.3026	2.3026	1
2	13	2.5649	5.1298	4
2	13	2.5649	5.1298	4
3	18	2.8904	8.6712	9
4	20	2.9957	11.9828	16
4	20	2.9957	11.9828	16
5	20	2.9957	14.9785	25
		21.6125	62.4801	76
		2.7016	7.81	9.5
		$M(\ln Y)$	$M(X \ln Y)$	$M(X^2)$

$$\hat{c} = \frac{M(X \ln Y)}{M(X^2)} = \frac{7.81}{9.5} = 0.8221$$

²Si osserva che per ragioni di stampa alcuni calcoli potrebbero essere duplicati con riferimento ai diversi modelli oggetto di studio



Il residuo quadratico medio $M(\hat{E}^2)$ può essere ottenuto calcolando la media del quadrato della variabile residuo $\hat{E} = Y - \hat{Y} = Y - \exp\{\hat{c}X\}$, costruita utilizzando la stima approssimata di c ricavata avvalendosi della forma linearizzata del modello.

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.



Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.

Il residuo quadratico medio $M(\hat{E}^2)$ può essere ottenuto calcolando la media del quadrato della variabile residuo $\hat{E} = Y - \hat{Y} = Y - \exp\{\hat{c}X\}$, costruita utilizzando la stima approssimata di c ricavata avvalendosi della forma linearizzata del modello. Abbiamo

y_i	$\hat{y}_i$	e_i	e_i^2
10	2.2753	7.7247	59.671
10	2.2753	7.7247	59.671
13	5.1769	7.8231	61.2009
13	5.1769	7.8231	61.2009
18	11.7788	6.2212	38.7033
20	26.7999	-6.7999	46.2386
20	26.7999	-6.7999	46.2386
20	60.9772	-40.9772	1679.1309
		-17.2602	2052.0552
		-2.1575	256.5069
		$M(E)$	$M(E^2)$

e

$$M(\hat{E}^2) = 256.5069$$



Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.

$\hat{a} = 7.5161$, $\hat{b} = 2.9032$, indice adattamento $M(\hat{E}^2) = 1.1694$ $\hat{c} = 0.8221$, indice adattamento $M(\hat{E}^2) = 256.5069$



3. Si individui il modello migliore.
Il modello migliore è

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.



3. Si individui il modello migliore.

Il modello migliore è $Y^* = 7.5161 + 2.9032 \cdot X$ in quanto caratterizzato da residuo quadratico medio inferiore.

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.



4. Si dica, motivando la risposta ma senza fare calcoli, se $\eta_{Y|X}^2 \neq \eta_{X|Y}^2$.
La relazione risulta verificata; abbiamo, infatti:

- $\eta_{Y|X}^2$

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.



4. Si dica, motivando la risposta ma senza fare calcoli, se $\eta_{Y|X}^2 \neq \eta_{X|Y}^2$.
La relazione risulta verificata; abbiamo, infatti:

- $\eta_{Y|X}^2 = 1$ in quanto

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.



Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.

4. Si dica, motivando la risposta ma senza fare calcoli, se $\eta_{Y|X}^2 \neq \eta_{X|Y}^2$.
La relazione risulta verificata; abbiamo, infatti:

- $\eta_{Y|X}^2 = 1$ in quanto le distribuzioni $Y|x_i$ sono degeneri
- $0 < \eta_{X|Y}^2$



Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.

4. Si dica, motivando la risposta ma senza fare calcoli, se $\eta_{Y|X}^2 \neq \eta_{X|Y}^2$.
La relazione risulta verificata; abbiamo, infatti:

- $\eta_{Y|X}^2 = 1$ in quanto le distribuzioni $Y|x_i$ sono degeneri
- $0 < \eta_{X|Y}^2 < 1$ in quanto



Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.

4. Si dica, motivando la risposta ma senza fare calcoli, se $\eta_{Y|X}^2 \neq \eta_{X|Y}^2$.

La relazione risulta verificata; abbiamo, infatti:

- $\eta_{Y|X}^2 = 1$ in quanto le distribuzioni $Y|x_i$ sono degeneri
- $0 < \eta_{X|Y}^2 < 1$ in quanto le distribuzioni condizionate $X|y_j$ non sono degeneri e le medie condizionate $M(Y|x_i)$ non sono tutte uguali.

Esercizio 4



Si considerano due urne contenenti palline numerate. L'urna U_1 contiene palline coi numeri 1, 2 e 3, in eguale numerosità, mentre l'urna U_2 è così composta: $\{4, 4, 4, 4, 5, 6, 7, 8, 8, 8\}$.

1. Si costruisca la distribuzione di probabilità delle variabili casuali $X =$ 'risultato di una estrazione dall'urna U_1 ' e $Y =$ 'risultato di una estrazione dall'urna U_2 '.

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.



[Esercizio 1.](#)

[Esercizio 2.](#)

[Esercizio 3.](#)

[Esercizio 4.](#)

Si considerano due urne contenenti palline numerate. L'urna U_1 contiene palline coi numeri 1, 2 e 3, in eguale numerosità, mentre l'urna U_2 è così composta: $\{4, 4, 4, 4, 5, 6, 7, 8, 8, 8\}$.

1. Si costruisca la distribuzione di probabilità delle variabili casuali $X =$ 'risultato di una estrazione dall'urna U_1 ' e $Y =$ 'risultato di una estrazione dall'urna U_2 '.

Abbiamo

x	$P(X = x)$
1	1/3
2	1/3
3	1/3



Si considerano due urne contenenti palline numerate. L'urna U_1 contiene palline coi numeri 1, 2 e 3, in eguale numerosità, mentre l'urna U_2 è così composta: $\{4, 4, 4, 4, 5, 6, 7, 8, 8, 8\}$.

1. Si costruisca la distribuzione di probabilità delle variabili casuali $X =$ 'risultato di una estrazione dall'urna U_1 ' e $Y =$ 'risultato di una estrazione dall'urna U_2 '.

Abbiamo

x	$P(X = x)$	y	$P(Y = y)$
1	1/3	4	0.4
2	1/3	5	0.1
3	1/3	6	0.1
		7	0.1
		8	0.3



Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.

2. Si estrae una pallina da ciascuna urna. Si calcoli la probabilità che la somma dei numeri estratti sia minore o uguale a 6.

$$P(\text{somma} \leq 6) = 0.3.$$

Abbiamo

$$P(\text{somma} \leq 6) =$$



Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.

2. Si estrae una pallina da ciascuna urna. Si calcoli la probabilità che la somma dei numeri estratti sia minore o uguale a 6.

$$P(\text{somma} \leq 6) = 0.3.$$

Abbiamo

$$P(\text{somma} \leq 6) = P((U_1 = 1) \cap (U_2 = 4) \cup (U_1 = 1) \cap (U_2 = 5) \cup (U_1 = 2) \cap (U_2 = 4)) =$$

=



Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.

2. Si estrae una pallina da ciascuna urna. Si calcoli la probabilità che la somma dei numeri estratti sia minore o uguale a 6.

$$P(\text{somma} \leq 6) = 0.3.$$

Abbiamo

$$\begin{aligned} P(\text{somma} \leq 6) &= P((U_1 = 1) \cap (U_2 = 4) \cup (U_1 = 1) \cap (U_2 = 5) \cup (U_1 = 2) \cap (U_2 = 4)) = \\ &= P((U_1 = 1) \cap (U_2 = 4)) + P((U_1 = 1) \cap (U_2 = 5)) + P((U_1 = 2) \cap (U_2 = 4)) = \\ &= \end{aligned}$$



Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.

2. Si estrae una pallina da ciascuna urna. Si calcoli la probabilità che la somma dei numeri estratti sia minore o uguale a 6.

$$P(\text{somma} \leq 6) = 0.3.$$

Abbiamo

$$\begin{aligned} P(\text{somma} \leq 6) &= P((U_1 = 1) \cap (U_2 = 4) \cup (U_1 = 1) \cap (U_2 = 5) \cup (U_1 = 2) \cap (U_2 = 4)) = \\ &= P((U_1 = 1) \cap (U_2 = 4)) + P((U_1 = 1) \cap (U_2 = 5)) + P((U_1 = 2) \cap (U_2 = 4)) = \\ &= P(U_1 = 1) \cdot P(U_2 = 4) + P(U_1 = 1) \cdot P(U_2 = 5) + P(U_1 = 2) \cdot P(U_2 = 4) = \\ &= \end{aligned}$$



Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.

2. Si estrae una pallina da ciascuna urna. Si calcoli la probabilità che la somma dei numeri estratti sia minore o uguale a 6.

$$P(\text{somma} \leq 6) = 0.3.$$

Abbiamo

$$\begin{aligned} P(\text{somma} \leq 6) &= P((U_1 = 1) \cap (U_2 = 4) \cup (U_1 = 1) \cap (U_2 = 5) \cup (U_1 = 2) \cap (U_2 = 4)) = \\ &= P((U_1 = 1) \cap (U_2 = 4)) + P((U_1 = 1) \cap (U_2 = 5)) + P((U_1 = 2) \cap (U_2 = 4)) = \\ &= P(U_1 = 1) \cdot P(U_2 = 4) + P(U_1 = 1) \cdot P(U_2 = 5) + P(U_1 = 2) \cdot P(U_2 = 4) = \\ &= \frac{1}{3} \cdot 0.4 + \frac{1}{3} \cdot 0.1 + \frac{1}{3} \cdot 0.4 = \frac{1}{3} \cdot 0.9 = \end{aligned}$$



Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.

2. Si estrae una pallina da ciascuna urna. Si calcoli la probabilità che la somma dei numeri estratti sia minore o uguale a 6.

$$P(\text{somma} \leq 6) = 0.3.$$

Abbiamo

$$\begin{aligned} P(\text{somma} \leq 6) &= P((U_1 = 1) \cap (U_2 = 4) \cup (U_1 = 1) \cap (U_2 = 5) \cup (U_1 = 2) \cap (U_2 = 4)) = \\ &= P((U_1 = 1) \cap (U_2 = 4)) + P((U_1 = 1) \cap (U_2 = 5)) + P((U_1 = 2) \cap (U_2 = 4)) = \\ &= P(U_1 = 1) \cdot P(U_2 = 4) + P(U_1 = 1) \cdot P(U_2 = 5) + P(U_1 = 2) \cdot P(U_2 = 4) = \\ &= \frac{1}{3} \cdot 0.4 + \frac{1}{3} \cdot 0.1 + \frac{1}{3} \cdot 0.4 = \frac{1}{3} \cdot 0.9 = 0.3 \end{aligned}$$



3. Considerando solo l'urna U_2 , si calcoli la probabilità che su 200 estrazioni con reimmissione il numero 8 esca almeno 50 volte.

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.



3. Considerando solo l'urna U_2 , si calcoli la probabilità che su 200 estrazioni con reimmissione il numero 8 esca almeno 50 volte.

Indicata con X la variabile casuale 'numero di volte che si presenta il numero 8' nelle 200 estrazioni con reimmissione risulta $X \sim$

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.



3. Considerando solo l'urna U_2 , si calcoli la probabilità che su 200 estrazioni con reimmissione il numero 8 esca almeno 50 volte.

Indicata con X la variabile casuale 'numero di volte che si presenta il numero 8' nelle 200 estrazioni con reimmissione risulta $X \sim \text{Bin}(n = 200, p = 0.3)$. $M(X) = 60$; $\text{Var}(X) = 42$. Dal momento che $np = 60 > 5$ ed $n(1 - p) = 140 > 5$ è possibile utilizzare l'approssimazione della variabile casuale Binomiale con la variabile casuale Normale $X \sim N(\mu = 60, \sigma^2 = 42)$.

Si ottiene

$$\begin{aligned} P(X \geq 50) &= P\left(\frac{X - \mu}{\sigma} \geq \frac{50 - \mu}{\sigma}\right) = P\left(Z \geq \frac{50 - 60}{6.4807}\right) = \\ &= P(Z \geq -1.54) = P(Z \leq 1.54) = 0.9382. \end{aligned}$$

Tema T294-1

Prova scritta di STATISTICA del 08.06.2017

Prova scritta di
STATISTICA del
08.06.2017

Giuseppe Boari
Gabriele Cantaluppi



UNIVERSITÀ
CATTOLICA
del Sacro Cuore

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.

Giuseppe Boari
Gabriele Cantaluppi
Università Cattolica del Sacro Cuore, Milano

Esercizio 1

Prova scritta di
STATISTICA del
08.06.2017

Giuseppe Boari
Gabriele Cantaluppi



UNIVERSITÀ
CATTOLICA
del Sacro Cuore

Il prospetto seguente riporta la distribuzione congiunta del carattere qualitativo X e del carattere quantitativo Y .

$X \setminus Y$	1	2	3	
x_1	8	16	16	
x_2	16	32	32	

[Esercizio 1.](#)

[Esercizio 2.](#)

[Esercizio 3.](#)

[Esercizio 4.](#)



Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.

1. Si rappresentino graficamente le distribuzioni di frequenza delle marginali X e Y .



Si considera la distribuzione marginale della variabile X , ottenuta sommando le frequenze di riga

x_i	$n_{i\bullet}$
x_1	40
x_2	80

Si tratta di un carattere qualitativo sconnesso che può essere rappresentato graficamente mediante un grafico a bastoncini o a canne d'organo.

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.

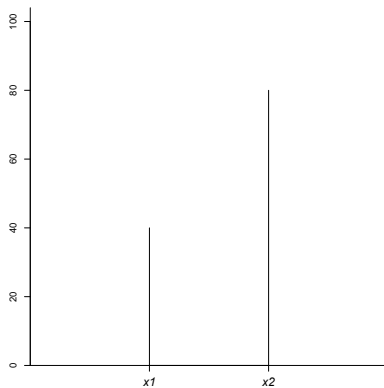


Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.





Si considera la distribuzione marginale della variabile Y , ottenuta sommando le frequenze di colonna

y_j	$n_{\bullet j}$
1	24
2	48
3	48

Esercizio 1.

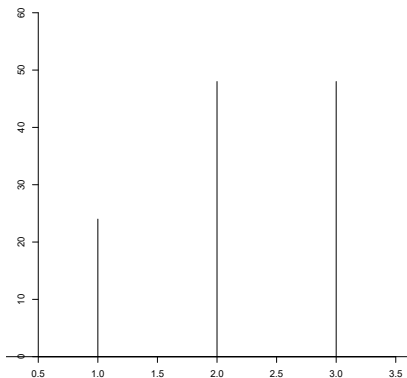
Esercizio 2.

Esercizio 3.

Esercizio 4.



Si tratta di un carattere quantitativo con modalità non raggruppate in classi che può essere rappresentato graficamente mediante un grafico a bastoncini.



Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.



2. Si confrontino le distribuzioni condizionate $Y|x_1$ e $Y|x_2$ mediante opportuni indici di posizione e di variabilità.

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.



2. Si confrontino le distribuzioni condizionate $Y|x_1$ e $Y|x_2$ mediante opportuni indici di posizione e di variabilità.

Si osserva che sussiste indipendenza stocastica in quanto le distribuzioni condizionate $Y|x_1$ e $Y|x_2$ sono somiglianti. Le due distribuzioni saranno caratterizzate dallo stesso livello medio e dalla stessa variabilità.

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.



2. Si confrontino le distribuzioni condizionate $Y|x_1$ e $Y|x_2$ mediante opportuni indici di posizione e di variabilità.

Si osserva che sussiste indipendenza stocastica in quanto le distribuzioni condizionate $Y|x_1$ e $Y|x_2$ sono somiglianti. Le due distribuzioni saranno caratterizzate dallo stesso livello medio e dalla stessa variabilità.

È, quindi, sufficiente procedere al calcolo degli indici per una sola delle due distribuzioni.

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.



Si considera la distribuzione condizionata $Y|x_1$

y_j	n_{1j}
1	8
2	16
3	16

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.



Si tratta di un carattere quantitativo con modalità non raggruppate in classi.

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.



La media aritmetica

$$M(Y|x_1) = \frac{1}{n_{1\bullet}} \sum_{j=1}^k y_j n_{1j}$$

risulta 2.2.

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.



La media aritmetica

$$M(Y|x_1) = \frac{1}{n_{1\bullet}} \sum_{j=1}^k y_j n_{1j}$$

risulta 2.2.

Si ottiene

y_j	n_{1j}	$y_j n_{1j}$
1	8	8
2	16	32
3	16	48
	40	88

da cui segue

$$M(Y|x_1) = \frac{1}{n_{1\bullet}} \sum_{j=1}^k y_j n_{1j} = \frac{88}{40} = 2.2.$$

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.



La varianza

$$\text{Var}(Y|x_1) = M(Y^2|x_1) - [M(Y|x_1)]^2$$

risulta 0.56.

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.



La varianza

$$\text{Var}(Y|x_1) = M(Y^2|x_1) - [M(Y|x_1)]^2$$

risulta 0.56. Si ottiene

y_j	n_{1j}	$y_j n_{1j}$	$y_j^2 n_{1j}$
1	8	8	8
2	16	32	64
3	16	48	144
	40	88	216

da cui segue

$$\text{Var}(Y|x_1) = \frac{216}{40} - 2.2^2 = 5.4 - 4.84 = 0.56.$$

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.



Per effettuare confronti in termini di variabilità tra caratteri quantitativi occorre utilizzare una misura relativa, ad esempio il coefficiente di variazione

$$CV = \frac{\sigma}{\mu} = \frac{\sqrt{\text{Var}(Y|x_1)}}{M(Y|x_1)} = \frac{\sqrt{0.56}}{2.2} = \frac{0.7483}{2.2} = 0.3401.$$

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.



$Y|x_1$ Indice di posizione: $M(Y|x_1) = M(Y|x_2) = 2.2$ Indice di variabilità:
 $CV(Y|x_1) = CV(Y|x_2) = 0.3401$

Le distribuzioni condizionate $Y|x_1$ e $Y|x_2$ sono caratterizzate dallo stesso livello medio e dalla stessa variabilità in quanto sussiste indipendenza stocastica.

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.



3. Si determini il grado di connessione tra le variabili X e Y mediante il calcolo di un opportuno indice.

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.



3. Si determini il grado di connessione tra le variabili X e Y mediante il calcolo di un opportuno indice.

L'indice di connessione χ_N^2 assume valore nullo in quanto sussiste indipendenza stocastica.

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.



La seguente tabella riporta la serie storica del fatturato annuo di un'azienda:

t	x_t	$NIBM$
2012	50	
2013	55	
2014	52.25	
2015	62.7	
2016	62.7	

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.



1. Si costruisca la serie dei numeri indici a base mobile e si calcoli, commentando il risultato, il valore di ${}_{2013}I_{2015}$, numero indice al tempo 2015 con base 2013.

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.



1. Si costruisca la serie dei numeri indici a base mobile e si calcoli, commentando il risultato, il valore di ${}_{2013}I_{2015}$, numero indice al tempo 2015 con base 2013.

Si ricorda che ${}_{t-1}I_t = \frac{x_t}{x_{t-1}}$.

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.



1. Si costruisca la serie dei numeri indici a base mobile e si calcoli, commentando il risultato, il valore di ${}_{2013}I_{2015}$, numero indice al tempo 2015 con base 2013.

Si ricorda che ${}_{t-1}I_t = \frac{x_t}{x_{t-1}}$. Si ottiene quindi

Anno	x_t	NIBM
2012	50.00	
2013	55.00	1.10
2014	52.25	0.95
2015	62.70	1.20
2016	62.70	1.00

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.



Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.



$${}_{2013}I_{2015} = \frac{X_{2015}}{X_{2013}} = \frac{62.7}{55} = 1.14$$

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.



Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.

$${}_{2013}I_{2015} = \frac{X_{2015}}{X_{2013}} = \frac{62.7}{55} = 1.14$$

Tra il tempo 2013 e il tempo 2015 si è registrato un incremento complessivo del 14% della grandezza oggetto di studio.



2. Si calcoli il tasso medio di variazione dal 2013 al 2015 e si commenti il risultato ottenuto.

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.



2. Si calcoli il tasso medio di variazione dal 2013 al 2015 e si commenti il risultato ottenuto.
Si ricorda che

$${}_{2013}I_{2015} = {}_{2013}I_{2014} \cdot {}_{2014}I_{2015}$$

Si deve trovare il numero indice a base mobile medio.

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.



Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.

2. Si calcoli il tasso medio di variazione dal 2013 al 2015 e si commenti il risultato ottenuto. Si ricorda che

$${}_{2013}I_{2015} = {}_{2013}I_{2014} \cdot {}_{2014}I_{2015}$$

Si deve trovare il numero indice a base mobile medio. Applicando il criterio di scelta della media secondo Chisini si ottiene

$${}_{2013}I_{2015} = {}_{2013}I_{2014} \cdot {}_{2014}I_{2015} = \alpha \cdot \alpha = \alpha^2.$$



Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.

2. Si calcoli il tasso medio di variazione dal 2013 al 2015 e si commenti il risultato ottenuto. Si ricorda che

$${}_{2013}I_{2015} = {}_{2013}I_{2014} \cdot {}_{2014}I_{2015}$$

Si deve trovare il numero indice a base mobile medio. Applicando il criterio di scelta della media secondo Chisini si ottiene

$${}_{2013}I_{2015} = {}_{2013}I_{2014} \cdot {}_{2014}I_{2015} = \alpha \cdot \alpha = \alpha^2.$$

Quindi

$$\alpha = \sqrt[2]{{}_{2013}I_{2015}} = \sqrt[2]{{}_{2013}I_{2014} \cdot {}_{2014}I_{2015}} = 1.0677.$$

Da cui segue una variazione relativa media pari a 0.0677, corrispondente a un incremento medio annuo tra il 2013 e il 2015 del 6.77%.



Il seguente prospetto riporta i valori rilevati in corrispondenza a 5 unità statistiche con riferimento a due caratteri quantitativi X e Y

x_i	2	2	4	6	6
y_i	2	2	3	2	2

1. Si rappresentino graficamente le coppie dei dati e la funzione di regressione ipotizzando Y in funzione di X .

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.



La funzione di regressione associa a ogni valore x appartenente all'insieme di definizione della variabile X la media di Y condizionata a x , $M(Y|x)$

x	$M(Y x)$
2	2
4	3
6	2

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.



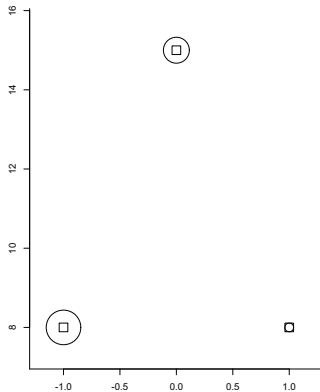
Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.

Le coppie di dati possono essere rappresentate mediante un *bubble plot diagram*.





2. Si calcolino, secondo il metodo dei minimi quadrati, i parametri dei seguenti modelli di regressione:

$$I) Y^* = a + b \cdot X \quad II) Y^* = c \cdot X$$

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.

Si riportano i calcoli necessari¹ per ottenere le stime dei parametri e gli indici di adattamento del modello

$$Y = a + bX + E$$

x_i	y_i	$x_i y_i$	x_i^2	y_i^2
2	2	4	4	4
2	2	4	4	4
4	3	12	16	9
6	2	12	36	4
6	2	12	36	4
20	11	44	96	25
4	2.2	8.8	19.2	5
$M(X)$	$M(Y)$	$M(XY)$	$M(X^2)$	$M(Y^2)$

$$\hat{b} =$$



Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.

¹Si osserva che per ragioni di stampa alcuni calcoli potrebbero essere duplicati con riferimento ai diversi modelli oggetto di studio

Si riportano i calcoli necessari¹ per ottenere le stime dei parametri e gli indici di adattamento del modello

$$Y = a + bX + E$$

x_i	y_i	$x_i y_i$	x_i^2	y_i^2
2	2	4	4	4
2	2	4	4	4
4	3	12	16	9
6	2	12	36	4
6	2	12	36	4
20	11	44	96	25
4	2.2	8.8	19.2	5
$M(X)$	$M(Y)$	$M(XY)$	$M(X^2)$	$M(Y^2)$

$$\begin{aligned}\hat{b} &= \frac{\text{Cov}(X, Y)}{\text{Var}(X)} = \frac{M(XY) - M(X)M(Y)}{M(X^2) - [M(X)]^2} = \\ &= \frac{8.8 - 4 \cdot 2.2}{19.2 - 4^2} = \frac{0}{3.2} = 0 \\ \hat{a} &= \end{aligned}$$

¹Si osserva che per ragioni di stampa alcuni calcoli potrebbero essere duplicati con riferimento ai diversi modelli oggetto di studio



Si riportano i calcoli necessari¹ per ottenere le stime dei parametri e gli indici di adattamento del modello

$$Y = a + bX + E$$

x_i	y_i	$x_i y_i$	x_i^2	y_i^2
2	2	4	4	4
2	2	4	4	4
4	3	12	16	9
6	2	12	36	4
6	2	12	36	4
20	11	44	96	25
4	2.2	8.8	19.2	5
$M(X)$	$M(Y)$	$M(XY)$	$M(X^2)$	$M(Y^2)$

$$\begin{aligned}\hat{b} &= \frac{\text{Cov}(X, Y)}{\text{Var}(X)} = \frac{M(XY) - M(X)M(Y)}{M(X^2) - [M(X)]^2} = \\ &= \frac{8.8 - 4 \cdot 2.2}{19.2 - 4^2} = \frac{0}{3.2} = 0 \\ \hat{a} &= M(Y) - \hat{b}M(X) = 2.2 \\ \rho^2 &= \end{aligned}$$

¹Si osserva che per ragioni di stampa alcuni calcoli potrebbero essere duplicati con riferimento ai diversi modelli oggetto di studio



Si riportano i calcoli necessari¹ per ottenere le stime dei parametri e gli indici di adattamento del modello

$$Y = a + bX + E$$

x_i	y_i	$x_i y_i$	x_i^2	y_i^2
2	2	4	4	4
2	2	4	4	4
4	3	12	16	9
6	2	12	36	4
6	2	12	36	4
20	11	44	96	25
4	2.2	8.8	19.2	5
$M(X)$	$M(Y)$	$M(XY)$	$M(X^2)$	$M(Y^2)$

$$\hat{b} = \frac{\text{Cov}(X, Y)}{\text{Var}(X)} = \frac{M(XY) - M(X)M(Y)}{M(X^2) - [M(X)]^2} =$$

$$= \frac{8.8 - 4 \cdot 2.2}{19.2 - 4^2} = \frac{0}{3.2} = 0$$

$$\hat{a} = M(Y) - \hat{b}M(X) = 2.2$$

$$\rho^2 = \frac{[\text{Cov}(X, Y)]^2}{\text{Var}(X) \text{Var}(Y)} = \frac{0^2}{3.2 \cdot 0.16} = 0$$

$$M(\hat{E}^2) = \text{Var}(\hat{E}) = \text{varianza residua} =$$

¹Si osserva che per ragioni di stampa alcuni calcoli potrebbero essere duplicati con riferimento ai diversi modelli oggetto di studio



Si riportano i calcoli necessari¹ per ottenere le stime dei parametri e gli indici di adattamento del modello

$$Y = a + bX + E$$

x_i	y_i	$x_i y_i$	x_i^2	y_i^2
2	2	4	4	4
2	2	4	4	4
4	3	12	16	9
6	2	12	36	4
6	2	12	36	4
20	11	44	96	25
4	2.2	8.8	19.2	5
$M(X)$	$M(Y)$	$M(XY)$	$M(X^2)$	$M(Y^2)$

$$\hat{b} = \frac{Cov(X, Y)}{Var(X)} = \frac{M(XY) - M(X)M(Y)}{M(X^2) - [M(X)]^2} =$$

$$= \frac{8.8 - 4 \cdot 2.2}{19.2 - 4^2} = \frac{0}{3.2} = 0$$

$$\hat{a} = M(Y) - \hat{b}M(X) = 2.2$$

$$\rho^2 = \frac{[Cov(X, Y)]^2}{Var(X) Var(Y)} = \frac{0^2}{3.2 \cdot 0.16} = 0$$

$$M(\hat{E}^2) = Var(\hat{E}) = \text{varianza residua} = Var(Y) \cdot (1 - \rho^2) = 0.16$$

¹Si osserva che per ragioni di stampa alcuni calcoli potrebbero essere duplicati con riferimento ai diversi modelli oggetto di studio





Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.

Si riportano i calcoli necessari² per ottenere le stime dei parametri e la misura di adattamento del modello

$$Y = cX + E$$

x_i	y_i	$x_i y_i$	x_i^2	y_i^2
2	2	4	4	4
2	2	4	4	4
4	3	12	16	9
6	2	12	36	4
6	2	12	36	4
		44	96	25
		8.8	19.2	5
		$M(XY)$	$M(X^2)$	$M(Y^2)$

$$\hat{c} =$$

²Si osserva che per ragioni di stampa alcuni calcoli potrebbero essere duplicati con riferimento ai diversi modelli oggetto di studio



Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.

Si riportano i calcoli necessari² per ottenere le stime dei parametri e la misura di adattamento del modello

$$Y = cX + E$$

x_i	y_i	$x_i y_i$	x_i^2	y_i^2
2	2	4	4	4
2	2	4	4	4
4	3	12	16	9
6	2	12	36	4
6	2	12	36	4
		44	96	25
		8.8	19.2	5
		$M(XY)$	$M(X^2)$	$M(Y^2)$

$$\hat{c} = \frac{M(XY)}{M(X^2)} = \frac{8.8}{19.2} = 0.4583$$

$$M(\hat{E}^2) =$$

²Si osserva che per ragioni di stampa alcuni calcoli potrebbero essere duplicati con riferimento ai diversi modelli oggetto di studio



Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.

Si riportano i calcoli necessari² per ottenere le stime dei parametri e la misura di adattamento del modello

$$Y = cX + E$$

x_i	y_i	$x_i y_i$	x_i^2	y_i^2
2	2	4	4	4
2	2	4	4	4
4	3	12	16	9
6	2	12	36	4
6	2	12	36	4
		44	96	25
		8.8	19.2	5
		$M(XY)$	$M(X^2)$	$M(Y^2)$

$$\hat{c} = \frac{M(XY)}{M(X^2)} = \frac{8.8}{19.2} = 0.4583$$

$$M(\hat{E}^2) = M[Y \cdot (Y - \hat{c}X)] = M(Y^2) - \hat{c}M(XY) = 5 - 0.4583 \cdot 8.8 = 0.9667$$

²Si osserva che per ragioni di stampa alcuni calcoli potrebbero essere duplicati con riferimento ai diversi modelli oggetto di studio



Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.

$\hat{a} = 2.2$, $\hat{b} = 0$, indice adattamento $M(\hat{E}^2) = 0.16$ $\hat{c} = 0.4583$, indice adattamento
 $M(\hat{E}^2) = 0.9667$



Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.

3. Confrontare gli indici di adattamento, commentando i risultati.



Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.

3. Confrontare gli indici di adattamento, commentando i risultati.

Il modello migliore è



3. Confrontare gli indici di adattamento, commentando i risultati.

Il modello migliore è $Y^* = 2.2$ in quanto caratterizzato da residuo quadratico medio inferiore.

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.

Esercizio 4



L'urna $U1$ contiene 3 palline nere e 1 pallina rossa. L'urna $U2$ contiene 9 palline nere e 10 palline rosse.

1. Si calcoli la probabilità che estraendo una pallina dall'urna $U1$ e una pallina dall'urna $U2$ si ottengano due palline dello stesso colore (si ottengano due palline nere oppure due palline rosse).

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.

Esercizio 4



L'urna $U1$ contiene 3 palline nere e 1 pallina rossa. L'urna $U2$ contiene 9 palline nere e 10 palline rosse.

1. Si calcoli la probabilità che estraendo una pallina dall'urna $U1$ e una pallina dall'urna $U2$ si ottengano due palline dello stesso colore (si ottengano due palline nere oppure due palline rosse).

$$P(2 \text{ nere o } 2 \text{ rosse}) =$$

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.

Esercizio 4



L'urna $U1$ contiene 3 palline nere e 1 pallina rossa. L'urna $U2$ contiene 9 palline nere e 10 palline rosse.

1. Si calcoli la probabilità che estraendo una pallina dall'urna $U1$ e una pallina dall'urna $U2$ si ottengano due palline dello stesso colore (si ottengano due palline nere oppure due palline rosse).

$$P(2 \text{ nere o } 2 \text{ rosse}) = P\{(N_1 \cap N_2) \cup (R_1 \cap R_2)\} =$$
$$=$$

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.



L'urna $U1$ contiene 3 palline nere e 1 pallina rossa. L'urna $U2$ contiene 9 palline nere e 10 palline rosse.

1. Si calcoli la probabilità che estraendo una pallina dall'urna $U1$ e una pallina dall'urna $U2$ si ottengano due palline dello stesso colore (si ottengano due palline nere oppure due palline rosse).

$$\begin{aligned} P(2 \text{ nere o } 2 \text{ rosse}) &= P\{(N_1 \cap N_2) \cup (R_1 \cap R_2)\} = \\ &= P(N_1 \cap N_2) + P(R_1 \cap R_2) = \\ &= \end{aligned}$$

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.

Esercizio 4



L'urna $U1$ contiene 3 palline nere e 1 pallina rossa. L'urna $U2$ contiene 9 palline nere e 10 palline rosse.

1. Si calcoli la probabilità che estraendo una pallina dall'urna $U1$ e una pallina dall'urna $U2$ si ottengano due palline dello stesso colore (si ottengano due palline nere oppure due palline rosse).

$$\begin{aligned} P(2 \text{ nere o } 2 \text{ rosse}) &= P\{(N_1 \cap N_2) \cup (R_1 \cap R_2)\} = \\ &= P(N_1 \cap N_2) + P(R_1 \cap R_2) = \\ &= P(N_1) \cdot P(N_2) + P(R_1) \cdot P(R_2) = \\ &= \end{aligned}$$

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.



L'urna $U1$ contiene 3 palline nere e 1 pallina rossa. L'urna $U2$ contiene 9 palline nere e 10 palline rosse.

1. Si calcoli la probabilità che estraendo una pallina dall'urna $U1$ e una pallina dall'urna $U2$ si ottengano due palline dello stesso colore (si ottengano due palline nere oppure due palline rosse).

$$\begin{aligned} P(2 \text{ nere o } 2 \text{ rosse}) &= P\{(N_1 \cap N_2) \cup (R_1 \cap R_2)\} = \\ &= P(N_1 \cap N_2) + P(R_1 \cap R_2) = \\ &= P(N_1) \cdot P(N_2) + P(R_1) \cdot P(R_2) = \\ &= \frac{3}{4} \cdot \frac{9}{19} + \frac{1}{4} \cdot \frac{10}{19} = \\ &= \end{aligned}$$

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.



L'urna $U1$ contiene 3 palline nere e 1 pallina rossa. L'urna $U2$ contiene 9 palline nere e 10 palline rosse.

1. Si calcoli la probabilità che estraendo una pallina dall'urna $U1$ e una pallina dall'urna $U2$ si ottengano due palline dello stesso colore (si ottengano due palline nere oppure due palline rosse).

$$\begin{aligned} P(2 \text{ nere o } 2 \text{ rosse}) &= P\{(N_1 \cap N_2) \cup (R_1 \cap R_2)\} = \\ &= P(N_1 \cap N_2) + P(R_1 \cap R_2) = \\ &= P(N_1) \cdot P(N_2) + P(R_1) \cdot P(R_2) = \\ &= \frac{3}{4} \cdot \frac{9}{19} + \frac{1}{4} \cdot \frac{10}{19} = \\ &= 0.3553 + 0.1316 = 0.4868. \end{aligned}$$

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.



2. Si calcoli la probabilità che estraendo con reimmissione 4 palline dall'urna U_2 ce ne siano almeno 3 nere.

$$P(\text{almeno } 3N) =$$

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.



2. Si calcoli la probabilità che estraendo con reimmissione 4 palline dall'urna $U2$ ce ne siano almeno 3 nere.

$$P(\text{almeno } 3N) = \binom{4}{3} \left(\frac{9}{19}\right)^3 \left(\frac{10}{19}\right) +$$

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.



2. Si calcoli la probabilità che estraendo con reimmissione 4 palline dall'urna $U2$ ce ne siano almeno 3 nere.

$$P(\text{almeno } 3N) = \binom{4}{3} \left(\frac{9}{19}\right)^3 \left(\frac{10}{19}\right) + \left(\frac{9}{19}\right)^4 =$$
$$=$$

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.



2. Si calcoli la probabilità che estraendo con reimmissione 4 palline dall'urna $U2$ ce ne siano almeno 3 nere.

$$\begin{aligned} P(\text{almeno } 3N) &= \binom{4}{3} \left(\frac{9}{19}\right)^3 \left(\frac{10}{19}\right) + \left(\frac{9}{19}\right)^4 = \\ &= 0.2238 + 0.0503 = \end{aligned}$$

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.



2. Si calcoli la probabilità che estraendo con reimmissione 4 palline dall'urna $U2$ ce ne siano almeno 3 nere.

$$\begin{aligned} P(\text{almeno } 3N) &= \binom{4}{3} \left(\frac{9}{19}\right)^3 \left(\frac{10}{19}\right) + \left(\frac{9}{19}\right)^4 = \\ &= 0.2238 + 0.0503 = 0.2741 \end{aligned}$$

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.



3. Si estraggano contemporaneamente due palline dall'urna $U1$ e le si mettano nell'urna $U2$. Si calcoli la probabilità che estraendo con reimmissione 2 palline dall'urna $U2$ (modificata) si ottengano 2 palline nere.

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.



3. Si estraggano contemporaneamente due palline dall'urna $U1$ e le si mettano nell'urna $U2$. Si calcoli la probabilità che estraendo con reimmissione 2 palline dall'urna $U2$ (modificata) si ottengano 2 palline nere.

In base alla legge delle probabilità totali abbiamo

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.



Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.

3. Si estraggano contemporaneamente due palline dall'urna $U1$ e le si mettano nell'urna $U2$. Si calcoli la probabilità che estraendo con reimmissione 2 palline dall'urna $U2$ (modificata) si ottengano 2 palline nere.

In base alla legge delle probabilità totali abbiamo

$$P(2N) = P(((2R)_1 \cap 2N) \cup ((2N)_1 \cap 2N) \cup ((1N1R)_1 \cap 2N)) =$$
$$=$$



Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.

3. Si estraggano contemporaneamente due palline dall'urna $U1$ e le si mettano nell'urna $U2$. Si calcoli la probabilità che estraendo con reimmissione 2 palline dall'urna $U2$ (modificata) si ottengano 2 palline nere.

In base alla legge delle probabilità totali abbiamo

$$\begin{aligned} P(2N) &= P(((2R)_1 \cap 2N) \cup ((2N)_1 \cap 2N) \cup ((1N1R)_1 \cap 2N)) = \\ &= P((2R)_1 \cap 2N) + P((2N)_1 \cap 2N) + P((1N1R)_1 \cap 2N) = \\ &= \end{aligned}$$



3. Si estraggano contemporaneamente due palline dall'urna $U1$ e le si mettano nell'urna $U2$. Si calcoli la probabilità che estraendo con reimmissione 2 palline dall'urna $U2$ (modificata) si ottengano 2 palline nere.

In base alla legge delle probabilità totali abbiamo

$$\begin{aligned}
 P(2N) &= P(((2R)_1 \cap 2N) \cup ((2N)_1 \cap 2N) \cup ((1N1R)_1 \cap 2N)) = \\
 &= P((2R)_1 \cap 2N) + P((2N)_1 \cap 2N) + P((1N1R)_1 \cap 2N) = \\
 &= P((2R)_1)P(2N|(2R)_1) + P((2N)_1)P(2N|(2N)_1) + P((1N1R)_1)P(2N|(1N1R)_1) = \\
 &= 0 + \left(\frac{3}{4} \cdot \frac{2}{3}\right) \left(\frac{11}{21}\right)^2 + \left(2 \cdot \frac{3}{4} \cdot \frac{1}{3}\right) \left(\frac{10}{21}\right)^2 = \\
 &=
 \end{aligned}$$



3. Si estraggano contemporaneamente due palline dall'urna $U1$ e le si mettano nell'urna $U2$. Si calcoli la probabilità che estraendo con reimmissione 2 palline dall'urna $U2$ (modificata) si ottengano 2 palline nere.

In base alla legge delle probabilità totali abbiamo

$$\begin{aligned}
 P(2N) &= P(((2R)_1 \cap 2N) \cup ((2N)_1 \cap 2N) \cup ((1N1R)_1 \cap 2N)) = \\
 &= P((2R)_1 \cap 2N) + P((2N)_1 \cap 2N) + P((1N1R)_1 \cap 2N) = \\
 &= P((2R)_1)P(2N|(2R)_1) + P((2N)_1)P(2N|(2N)_1) + P((1N1R)_1)P(2N|(1N1R)_1) = \\
 &= 0 + \left(\frac{3}{4} \cdot \frac{2}{3}\right) \left(\frac{11}{21}\right)^2 + \left(2 \cdot \frac{3}{4} \cdot \frac{1}{3}\right) \left(\frac{10}{21}\right)^2 = \\
 &= 0.5 \cdot 0.2744 + 0.5 \cdot 0.2268 = \\
 &=
 \end{aligned}$$



3. Si estraggano contemporaneamente due palline dall'urna $U1$ e le si mettano nell'urna $U2$. Si calcoli la probabilità che estraendo con reimmissione 2 palline dall'urna $U2$ (modificata) si ottengano 2 palline nere.

In base alla legge delle probabilità totali abbiamo

$$\begin{aligned}
 P(2N) &= P(((2R)_1 \cap 2N) \cup ((2N)_1 \cap 2N) \cup ((1N1R)_1 \cap 2N)) = \\
 &= P((2R)_1 \cap 2N) + P((2N)_1 \cap 2N) + P((1N1R)_1 \cap 2N) = \\
 &= P((2R)_1)P(2N|(2R)_1) + P((2N)_1)P(2N|(2N)_1) + P((1N1R)_1)P(2N|(1N1R)_1) = \\
 &= 0 + \left(\frac{3}{4} \cdot \frac{2}{3}\right) \left(\frac{11}{21}\right)^2 + \left(2 \cdot \frac{3}{4} \cdot \frac{1}{3}\right) \left(\frac{10}{21}\right)^2 = \\
 &= 0.5 \cdot 0.2744 + 0.5 \cdot 0.2268 = \\
 &= 0.1372 + 0.1134 = 0.2506
 \end{aligned}$$

Tema T294

Prova scritta di STATISTICA del 22.06.2017

Prova scritta di
STATISTICA del
22.06.2017

Giuseppe Boari
Gabriele Cantaluppi



UNIVERSITÀ
CATTOLICA
del Sacro Cuore

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.

Giuseppe Boari
Gabriele Cantaluppi
Università Cattolica del Sacro Cuore, Milano



Data la seguente tabella a doppia entrata che riporta la distribuzione congiunta dei caratteri $X = \text{'genere'}$ e $Y = \text{'tempo dedicato settimanalmente ai lavori domestici (in ore)'}$ per un collettivo di 10 persone:

$X \setminus Y$	0 - 2	2 - 4	4 - 8	
<i>maschio</i>	6	0	0	
<i>femmina</i>	0	2	2	

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.



Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.

1. Fornire una opportuna rappresentazione grafica delle distribuzioni marginali di X e Y .



Si considera la distribuzione marginale della variabile X , ottenuta sommando le frequenze di riga

x_i	$n_{i\bullet}$
<i>maschio</i>	6
<i>femmina</i>	4

Si tratta di un carattere qualitativo sconnesso che può essere rappresentato graficamente mediante un grafico a bastoncini o a canne d'organo.

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.

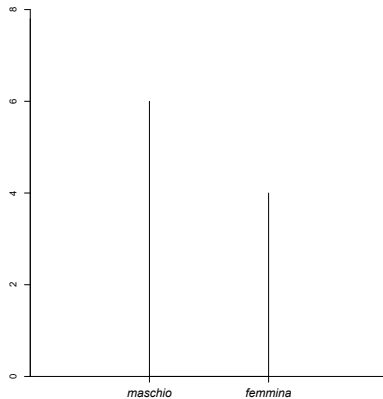


Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.





Si considera la distribuzione marginale della variabile Y , ottenuta sommando le frequenze di colonna

y_j	$n_{\bullet j}$
0 $\rightarrow$ 2	6
2 $\rightarrow$ 4	2
4 $\rightarrow$ 8	2

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.



Si tratta di un carattere quantitativo con modalità raggruppate in classi che può essere rappresentato graficamente mediante un istogramma.

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.



L'istogramma può essere costruito dopo avere calcolato le densità $d_{\bullet j} = \frac{n_{\bullet j}}{a_j}$, (dove a_j sono le ampiezze di classe).

y_j	$n_{\bullet j}$	$a_{\bullet j}$	$d_{\bullet j}$
0 $\rightarrow$ 2	6	2 $-$ 0 = 2	3
2 $\rightarrow$ 4	2	4 $-$ 2 = 2	1
4 $\rightarrow$ 8	2	8 $-$ 4 = 4	0.5

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.

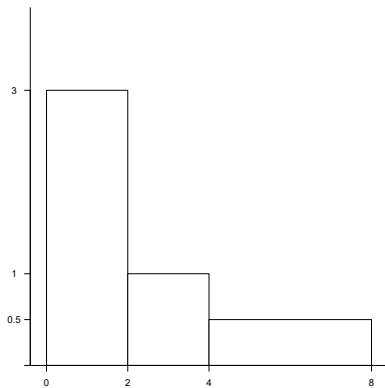


Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.





2. Calcolare moda, media e mediana per la marginale Y .

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.



Si tratta di un carattere quantitativo con modalità raggruppate in classi.

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.



La moda corrisponde alla classe alla quale è associata la massima densità di frequenza

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.



La moda corrisponde alla classe alla quale è associata la massima densità di frequenza

y_j	$n_{\bullet j}$	$a_{\bullet j}$	$d_{\bullet j}$
0 $\rightarrow$ 2	6	$2 - 0 = 2$	3
2 $\rightarrow$ 4	2	$4 - 2 = 2$	1
4 $\rightarrow$ 8	2	$8 - 4 = 4$	0.5

e risulta 0 $\rightarrow$ 2.

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.



La mediana (percentile di ordine 0.50) può essere ottenuta considerando l'inversa della Funzione di Ripartizione in corrispondenza di 0.50.

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.



La mediana (percentile di ordine 0.50) può essere ottenuta considerando l'inversa della Funzione di Ripartizione in corrispondenza di 0.50.

y_j	$n_{\bullet j}$	$N_{\bullet j}$	$F_{\bullet j}$
0 - 2	6	6	0.6
2 - 4	2	8	0.8
4 - 8	2	10	1

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.

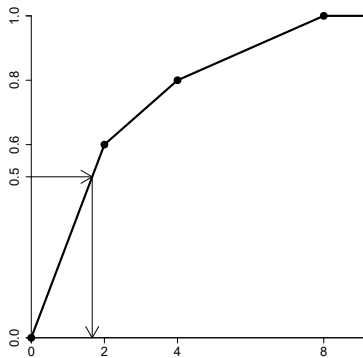


Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.



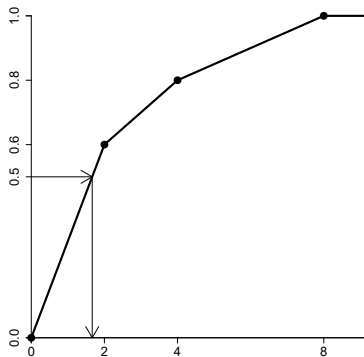


Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.



Si ottiene

$$y_{0.5} = 0 + \frac{0.5 - 0}{0.6 - 0} \cdot (2 - 0) = 1.6667.$$



La media aritmetica

$$M(Y) = \frac{1}{n} \sum_{j=1}^k y_j n_{\bullet j}$$

risulta 2.4.

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.



La media aritmetica

$$M(Y) = \frac{1}{n} \sum_{j=1}^k y_j n_{\bullet j}$$

risulta 2.4.

Si ottiene (considerando come modalità i valori centrali delle classi)

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.



La media aritmetica

$$M(Y) = \frac{1}{n} \sum_{j=1}^k y_j n_{\bullet j}$$

risulta 2.4.

Si ottiene (considerando come modalità i valori centrali delle classi)

y_j	$n_{\bullet j}$	$y_j n_{\bullet j}$
1	6	6
3	2	6
6	2	12
	10	24

da cui segue

$$M(Y) = \frac{1}{n} \sum_{i=1}^k y_i n_{\bullet i} = \frac{24}{10} = 2.4.$$

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.



$$\text{Moda}(Y) = 0 \div 2 \quad M(Y) = 2.4 \quad y_{0.5} = 1.6667$$

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.



3. Si calcoli un opportuno indice di eterogeneità per la variabile X e se ne commenti il valore.

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.



In presenza di un carattere qualitativo la variabilità può essere misurata mediante gli indici di Gini e di Frosini normalizzati

$$G_N = \frac{1 - \sum_{i=1}^h f_{i\bullet}^2}{1 - \frac{1}{h}}$$

e

$$F_N = 1 - \sqrt{1 - G_N}.$$

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.



In presenza di un carattere qualitativo la variabilità può essere misurata mediante gli indici di Gini e di Frosini normalizzati

$$G_N = \frac{1 - \sum_{i=1}^h f_{i\bullet}^2}{1 - \frac{1}{h}}$$

e

$$F_N = 1 - \sqrt{1 - G_N}.$$

Si ottiene:

x_i	$n_{i\bullet}$	$f_{i\bullet}$	$f_{i\bullet}^2$
<i>maschio</i>	6	0.6	0.36
<i>femmina</i>	4	0.4	0.16
	10	1	0.52

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.



In presenza di un carattere qualitativo la variabilità può essere misurata mediante gli indici di Gini e di Frosini normalizzati

$$G_N = \frac{1 - \sum_{i=1}^h f_{i\bullet}^2}{1 - \frac{1}{h}}$$

e

$$F_N = 1 - \sqrt{1 - G_N}.$$

Si ottiene:

x_i	$n_{i\bullet}$	$f_{i\bullet}$	$f_{i\bullet}^2$
<i>maschio</i>	6	0.6	0.36
<i>femmina</i>	4	0.4	0.16
	10	1	0.52

Da cui segue:

$$G_N = \frac{1 - 0.52}{1 - \frac{1}{2}} = 0.96$$

e

$$F_N = 1 - \sqrt{1 - 0.96} = 0.8.$$

Sussiste, quindi, un livello alto di eterogeneità.



4. Si calcolino la varianza within e la varianza between del tempo dedicato ai lavori domestici considerando come gruppi quelli definiti dal genere e si commentino i risultati ottenuti.

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.



Si considera la distribuzione condizionata $Y|x_1$

y_j	n_{1j}
0 $\rightarrow$ 2	6
2 $\rightarrow$ 4	0
4 $\rightarrow$ 8	0

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.



Si tratta di un carattere quantitativo con modalità raggruppate in classi.

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.



Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.

Si osserva che la distribuzione è degenere.

$M(Y|x_1) = 1$. La varianza e il coefficiente di variazione sono nulli.



Si considera la distribuzione condizionata $Y|x_2$

y_j	n_{2j}
0 $\dashv$ 2	0
2 $\dashv$ 4	2
4 $\dashv$ 8	2

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.



Si tratta di un carattere quantitativo con modalità raggruppate in classi.

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.



La media aritmetica

$$M(Y|x_2) = \frac{1}{n_{2\bullet}} \sum_{j=1}^k y_j n_{2j}$$

risulta 4.5.

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.



La media aritmetica

$$M(Y|x_2) = \frac{1}{n_{2\bullet}} \sum_{j=1}^k y_j n_{2j}$$

risulta 4.5.

Si ottiene (considerando come modalità i valori centrali delle classi)

y_j	n_{2j}	$y_j n_{2j}$
1	0	0
3	2	6
6	2	12
	4	18

da cui segue

$$M(Y|x_2) = \frac{1}{n_{2\bullet}} \sum_{j=1}^k y_j n_{2j} = \frac{18}{4} = 4.5.$$

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.



La varianza

risulta 2.25.

$$\text{Var}(Y|x_2) = M(Y^2|x_2) - [M(Y|x_2)]^2$$

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.



Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.

La varianza

$$\text{Var}(Y|x_2) = M(Y^2|x_2) - [M(Y|x_2)]^2$$

risulta 2.25. Si ottiene (considerando come modalità i valori centrali delle classi)

y_j	n_{2j}	$y_j n_{2j}$	$y_j^2 n_{2j}$
1	0	0	0
3	2	6	18
6	2	12	72
	4	18	90

da cui segue

$$\text{Var}(Y|x_2) = \frac{90}{4} - 4.5^2 = 22.5 - 20.25 = 2.25.$$



Si considera la distribuzione marginale della variabile Y , ottenuta sommando le frequenze di colonna

y_j	$n_{\bullet j}$
0 $\rightarrow$ 2	6
2 $\rightarrow$ 4	2
4 $\rightarrow$ 8	2

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.



Si tratta di un carattere quantitativo con modalità raggruppate in classi.

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.



La media aritmetica

$$M(Y) = \frac{1}{n} \sum_{j=1}^k y_j n_{\bullet j}$$

risulta 2.4.

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.



La media aritmetica

$$M(Y) = \frac{1}{n} \sum_{j=1}^k y_j n_{\bullet j}$$

risulta 2.4.

Si ottiene (considerando come modalità i valori centrali delle classi)

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.



La media aritmetica

$$M(Y) = \frac{1}{n} \sum_{j=1}^k y_j n_{\bullet j}$$

risulta 2.4.

Si ottiene (considerando come modalità i valori centrali delle classi)

y_j	$n_{\bullet j}$	$y_j n_{\bullet j}$
1	6	6
3	2	6
6	2	12
	10	24

da cui segue

$$M(Y) = \frac{1}{n} \sum_{i=1}^k y_i n_{\bullet i} = \frac{24}{10} = 2.4.$$

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.



La varianza

$$\text{Var}(Y) = M[(Y - \mu_Y)^2] = M(Y^2) - \mu_Y^2 = \frac{1}{n} \sum_{j=1}^k y_j^2 n_{\bullet j} - \mu_Y^2$$

risulta 3.84.

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.



La varianza

$$\text{Var}(Y) = M[(Y - \mu_Y)^2] = M(Y^2) - \mu_Y^2 = \frac{1}{n} \sum_{j=1}^k y_j^2 n_{\bullet j} - \mu_Y^2$$

risulta 3.84.

Si ottiene (considerando come modalità i valori centrali delle classi)

y_j	$n_{\bullet j}$	$y_j n_{\bullet j}$	$y_j^2 n_{\bullet j}$
1	6	6	6
3	2	6	18
6	2	12	72
	10	24	96

da cui segue

$$\begin{aligned} \text{Var}(Y) = \sigma_Y^2 &= M(Y^2) - \mu_Y^2 = \\ &= \frac{96}{10} - 2.4^2 = 9.6 - 5.76 = 3.84. \end{aligned}$$

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.



Il rapporto di correlazione

$$\eta_{Y|X}^2 = \frac{\text{Var}_X[M(Y|X)]}{\text{Var}(Y)}$$

può essere ottenuto considerando il rapporto tra la varianza spiegata $\text{Var}_X[M(Y|X)]$,
varianza della variabile statistica medie condizionate

i	$M(X y_i)$	$n_{i\bullet}$
1	1.0	6
2	4.5	4

e la varianza della variabile dipendente $\text{Var}(Y) = 3.84$.

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.



La varianza spiegata risulta $Var_X[M(Y|X)] = 2.94$.

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.



La varianza spiegata risulta $Var_X[M(Y|X)] = 2.94$. Si ottiene

i	$M(X y_i)$	$n_{i\bullet}$	$M(X y_i) \cdot n_{i\bullet}$	$[M(X y_i)]^2 \cdot n_{i\bullet}$
1	1.0	6	6	6
2	4.5	4	18	81
		10	24	87

da cui segue

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.



La varianza spiegata risulta $Var_X[M(Y|X)] = 2.94$. Si ottiene

i	$M(X y_i)$	$n_{i\bullet}$	$M(X y_i) \cdot n_{i\bullet}$	$[M(X y_i)]^2 \cdot n_{i\bullet}$
1	1.0	6	6	6
2	4.5	4	18	81
		10	24	87

da cui segue

$$Var_X[M(Y|X)] = \frac{87}{10} - 2.4^2 = 8.7 - 5.76 = 2.94.$$

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.



Il rapporto di correlazione $\eta^2_{Y|X}$ risulta:

$$\eta^2_{Y|X} = \frac{\text{Var}_X[M(Y|X)]}{\text{Var}(Y)} = \frac{2.94}{3.84} = 0.7656.$$

Sussiste, quindi, un livello di dipendenza medio alto.

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.



Si ricorda che, in base al risultato di scomposizione della varianza, la varianza della variabile dipendente può essere ottenuta come somma della varianza spiegata $Var_X[M(Y|X)]$ e della varianza residua $M_X[Var(Y|X)]$.

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.



La varianza residua $M_X[\text{Var}(Y|X)]$, media della variabile statistica varianze condizionate

i	$\text{Var}(X y_i)$	$n_{i\bullet}$
1	0.00	6
2	2.25	4

risulta $M_X[\text{Var}(Y|X)] = 0.9$.

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.



Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.

La varianza residua $M_X[Var(Y|X)]$, media della variabile statistica varianze condizionate

i	$Var(X y_i)$	$n_{i\bullet}$
1	0.00	6
2	2.25	4

risulta $M_X[Var(Y|X)] = 0.9$. Si ottiene

i	$Var(X y_i)$	$n_{i\bullet}$	$Var(X y_i) \cdot n_{i\bullet}$
1	0.00	6	0
2	2.25	4	9
		10	9

da cui segue



Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.

La varianza residua $M_X[Var(Y|X)]$, media della variabile statistica varianze condizionate

i	$Var(X y_i)$	$n_{i\bullet}$
1	0.00	6
2	2.25	4

risulta $M_X[Var(Y|X)] = 0.9$. Si ottiene

i	$Var(X y_i)$	$n_{i\bullet}$	$Var(X y_i) \cdot n_{i\bullet}$
1	0.00	6	0
2	2.25	4	9
		10	9

da cui segue

$$M_X[Var(Y|X)] = \frac{9}{10} = 0.9.$$



Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.

$\sigma_{Between}^2 = 2.94$ $\sigma_{Within}^2 = 0.9$ da cui abbiamo $\eta_{Y|X}^2 = \frac{2.94}{2.94+0.9} = 0.7656$.
Sussiste, quindi, un livello di dipendenza medio alto.



5. Senza fare calcoli, si dica che valore assume l'indice χ^2 normalizzato.

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.



Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.

5. Senza fare calcoli, si dica che valore assume l'indice χ^2 normalizzato.
 $\chi_N^2 = 1$ in quanto sussiste dipendenza funzionale di X da Y essendo le distribuzioni condizionate $x|y_j$ degeneri.



Il prospetto seguente riporta la serie storica dei prezzi di due beni A e B fra il 2012 e il 2016:

t	${}_Ap_t$	${}_Bp_t$
2012	200	400
2013	210	410
2014	220	390
2015	240	390
2016	260	410

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.



1. Costruire la serie dei numeri indice a base fissa con base nel 2012 per il bene A.

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.



1. Costruire la serie dei numeri indice a base fissa con base nel 2012 per il bene A.

Si ricorda che ${}_{2012}I_t = \frac{x_t}{x_{2012}}$.

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.



1. Costruire la serie dei numeri indice a base fissa con base nel 2012 per il bene A.
Si ricorda che ${}_{2012}I_t = \frac{x_t}{x_{2012}}$. Si ottiene quindi

Anno	x_t	NIBF (2012)
2012	200	1.00
2013	210	1.05
2014	220	1.10
2015	240	1.20
2016	260	1.30

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.



Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.

2. Sapendo che le quantità ${}_Aq_{2012} = 100$ e ${}_Bq_{2012} = 200$, calcolare l'indice di Laspeyres dei prezzi al 2013 con base al 2012.



L'indice dei prezzi di Laspeyres si ottiene considerando la media aritmetica dei 'prezzi relativi' riferiti ai singoli prodotti $x_i = \frac{i p_{2013}}{i p_{2012}}$ calcolata utilizzando come frequenze i valori

$i v_{2012} = i p_{2012} \cdot i q_{2012}$ scambiati al tempo 2012

$${}^p_{2012} I^L_{2013} = \frac{1}{\sum_{i=1}^h i p_{2012} i q_{2012}} \sum_{i=1}^h \frac{i p_{2013}}{i p_{2012}} i p_{2012} i q_{2012}$$

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.



Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.

$\frac{i p_{2013}}{i p_{2012}}$	$i p_{2012} \cdot i q_{2012}$	$i p_{2013} \cdot i q_{2012}$
$\frac{210}{200} = 1.05$	$200 \cdot 100 = 20000$	21000
$\frac{410}{400} = 1.025$	$400 \cdot 200 = 80000$	82000
	$1e + 05$	103000

$${}^p_{2012}I^L_{2013} = \frac{1}{1e + 05} 103000 = 1.03.$$

La variazione relativa nel livello generale dei prezzi, secondo l'Indice di Laspeyres, è del 3%.



3. Si calcoli il tasso medio di variazione del prezzo del prodotto A nel periodo 2014-2016 interpretando il risultato ottenuto.

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.



3. Si calcoli il tasso medio di variazione del prezzo del prodotto *A* nel periodo 2014-2016 interpretando il risultato ottenuto.

Si ricorda che

$${}_{2014}I_{2016} = {}_{2014}I_{2015} \cdot {}_{2015}I_{2016}$$

Si deve trovare il numero indice a base mobile medio.

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.



Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.

3. Si calcoli il tasso medio di variazione del prezzo del prodotto A nel periodo 2014-2016 interpretando il risultato ottenuto.

Si ricorda che

$${}_{2014}I_{2016} = {}_{2014}I_{2015} \cdot {}_{2015}I_{2016}$$

Si deve trovare il numero indice a base mobile medio. Applicando il criterio di scelta della media secondo Chisini si ottiene

$${}_{2014}I_{2016} = {}_{2014}I_{2015} \cdot {}_{2015}I_{2016} = \alpha \cdot \alpha = \alpha^2.$$



Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.

3. Si calcoli il tasso medio di variazione del prezzo del prodotto *A* nel periodo 2014-2016 interpretando il risultato ottenuto.

Si ricorda che

$${}_{2014}I_{2016} = {}_{2014}I_{2015} \cdot {}_{2015}I_{2016}$$

Si deve trovare il numero indice a base mobile medio. Applicando il criterio di scelta della media secondo Chisini si ottiene

$${}_{2014}I_{2016} = {}_{2014}I_{2015} \cdot {}_{2015}I_{2016} = \alpha \cdot \alpha = \alpha^2.$$

Quindi

$$\alpha = \sqrt[2]{{}_{2014}I_{2016}} = \sqrt[2]{{}_{2014}I_{2015} \cdot {}_{2015}I_{2016}} = 1.0871.$$

Da cui segue una variazione relativa media pari a 0.0871, corrispondente a un incremento medio annuo tra il 2014 e il 2016 del 8.71%.



Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.

Con riferimento allo studio della dipendenza tra due caratteri quantitativi X e Y , si sono raccolti i dati relativi alle seguenti 5 coppie di osservazioni (x_i, y_i) $i = 1, 2, \dots, 5$

x_i	2	3	4	5	2
y_i	-1	0	1	2	3

1. Si rappresentino graficamente le coppie dei dati e la funzione di regressione ipotizzando Y in funzione di X .



La funzione di regressione associa a ogni valore x appartenente all'insieme di definizione della variabile X la media di Y condizionata a x , $M(Y|x)$

x	$M(Y x)$
2	$\frac{1}{2}(-1 + 3) = 1$
3	0
4	1
5	2

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.



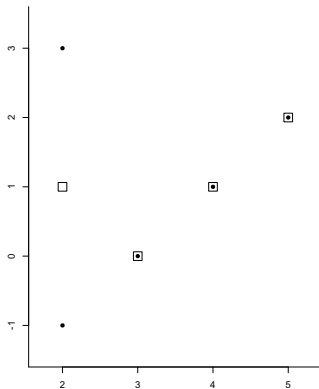
Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.

Le coppie di dati possono essere rappresentate mediante un grafico a dispersione.





2. Si calcolino, secondo il metodo dei minimi quadrati, i parametri dei seguenti modelli di regressione:

$$I) Y^* = a + b \cdot X$$

$$II) Y^* = c \cdot X$$

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.

Si riportano i calcoli necessari¹ per ottenere le stime dei parametri e gli indici di adattamento del modello

$$Y = a + bX + E$$

x_i	y_i	$x_i y_i$	x_i^2	y_i^2
2	-1	-2	4	1
3	0	0	9	0
4	1	4	16	1
5	2	10	25	4
2	3	6	4	9
16	5	18	58	15
3.2	1	3.6	11.6	3
$M(X)$	$M(Y)$	$M(XY)$	$M(X^2)$	$M(Y^2)$

$$\hat{b} =$$



Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.

¹Si osserva che per ragioni di stampa alcuni calcoli potrebbero essere duplicati con riferimento ai diversi modelli oggetto di studio

Si riportano i calcoli necessari¹ per ottenere le stime dei parametri e gli indici di adattamento del modello

$$Y = a + bX + E$$

x_i	y_i	$x_i y_i$	x_i^2	y_i^2
2	-1	-2	4	1
3	0	0	9	0
4	1	4	16	1
5	2	10	25	4
2	3	6	4	9
16	5	18	58	15
3.2	1	3.6	11.6	3
$M(X)$	$M(Y)$	$M(XY)$	$M(X^2)$	$M(Y^2)$

$$\begin{aligned}\hat{b} &= \frac{\text{Cov}(X, Y)}{\text{Var}(X)} = \frac{M(XY) - M(X)M(Y)}{M(X^2) - [M(X)]^2} = \\ &= \frac{3.6 - 3.2 \cdot 1}{11.6 - 3.2^2} = \frac{0.4}{1.36} = 0.2941 \\ \hat{a} &= \end{aligned}$$



Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.

¹Si osserva che per ragioni di stampa alcuni calcoli potrebbero essere duplicati con riferimento ai diversi modelli oggetto di studio

Si riportano i calcoli necessari¹ per ottenere le stime dei parametri e gli indici di adattamento del modello

$$Y = a + bX + E$$

x_i	y_i	$x_i y_i$	x_i^2	y_i^2
2	-1	-2	4	1
3	0	0	9	0
4	1	4	16	1
5	2	10	25	4
2	3	6	4	9
16	5	18	58	15
3.2	1	3.6	11.6	3
$M(X)$	$M(Y)$	$M(XY)$	$M(X^2)$	$M(Y^2)$

$$\begin{aligned}\hat{b} &= \frac{\text{Cov}(X, Y)}{\text{Var}(X)} = \frac{M(XY) - M(X)M(Y)}{M(X^2) - [M(X)]^2} = \\ &= \frac{3.6 - 3.2 \cdot 1}{11.6 - 3.2^2} = \frac{0.4}{1.36} = 0.2941 \\ \hat{a} &= M(Y) - \hat{b}M(X) = 0.0588 \\ \rho^2 &= \end{aligned}$$

¹Si osserva che per ragioni di stampa alcuni calcoli potrebbero essere duplicati con riferimento ai diversi modelli oggetto di studio



Si riportano i calcoli necessari¹ per ottenere le stime dei parametri e gli indici di adattamento del modello

$$Y = a + bX + E$$

x_i	y_i	$x_i y_i$	x_i^2	y_i^2
2	-1	-2	4	1
3	0	0	9	0
4	1	4	16	1
5	2	10	25	4
2	3	6	4	9
16	5	18	58	15
3.2	1	3.6	11.6	3
$M(X)$	$M(Y)$	$M(XY)$	$M(X^2)$	$M(Y^2)$

$$\hat{b} = \frac{\text{Cov}(X, Y)}{\text{Var}(X)} = \frac{M(XY) - M(X)M(Y)}{M(X^2) - [M(X)]^2} =$$

$$= \frac{3.6 - 3.2 \cdot 1}{11.6 - 3.2^2} = \frac{0.4}{1.36} = 0.2941$$

$$\hat{a} = M(Y) - \hat{b}M(X) = 0.0588$$

$$\rho^2 = \frac{[\text{Cov}(X, Y)]^2}{\text{Var}(X) \text{Var}(Y)} = \frac{0.4^2}{1.36 \cdot 2} = 0.0588$$

$$M(\hat{E}^2) = \text{Var}(\hat{E}) = \text{varianza residua} =$$

¹Si osserva che per ragioni di stampa alcuni calcoli potrebbero essere duplicati con riferimento ai diversi modelli oggetto di studio



Si riportano i calcoli necessari¹ per ottenere le stime dei parametri e gli indici di adattamento del modello

$$Y = a + bX + E$$

x_i	y_i	$x_i y_i$	x_i^2	y_i^2
2	-1	-2	4	1
3	0	0	9	0
4	1	4	16	1
5	2	10	25	4
2	3	6	4	9
16	5	18	58	15
3.2	1	3.6	11.6	3
$M(X)$	$M(Y)$	$M(XY)$	$M(X^2)$	$M(Y^2)$

$$\begin{aligned}\hat{b} &= \frac{\text{Cov}(X, Y)}{\text{Var}(X)} = \frac{M(XY) - M(X)M(Y)}{M(X^2) - [M(X)]^2} = \\ &= \frac{3.6 - 3.2 \cdot 1}{11.6 - 3.2^2} = \frac{0.4}{1.36} = 0.2941 \\ \hat{a} &= M(Y) - \hat{b}M(X) = 0.0588 \\ \rho^2 &= \frac{[\text{Cov}(X, Y)]^2}{\text{Var}(X) \text{Var}(Y)} = \frac{0.4^2}{1.36 \cdot 2} = 0.0588\end{aligned}$$

$$M(\hat{E}^2) = \text{Var}(\hat{E}) = \text{varianza residua} = \text{Var}(Y) \cdot (1 - \rho^2) = 1.8824$$

¹Si osserva che per ragioni di stampa alcuni calcoli potrebbero essere duplicati con riferimento ai diversi modelli oggetto di studio





Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.

Si riportano i calcoli necessari² per ottenere le stime dei parametri e la misura di adattamento del modello

$$Y = cX + E$$

x_i	y_i	$x_i y_i$	x_i^2	y_i^2
2	-1	-2	4	1
3	0	0	9	0
4	1	4	16	1
5	2	10	25	4
2	3	6	4	9
		18	58	15
		3.6	11.6	3
		$M(XY)$	$M(X^2)$	$M(Y^2)$

$$\hat{c} =$$

²Si osserva che per ragioni di stampa alcuni calcoli potrebbero essere duplicati con riferimento ai diversi modelli oggetto di studio



Si riportano i calcoli necessari² per ottenere le stime dei parametri e la misura di adattamento del modello

$$Y = cX + E$$

x_i	y_i	$x_i y_i$	x_i^2	y_i^2
2	-1	-2	4	1
3	0	0	9	0
4	1	4	16	1
5	2	10	25	4
2	3	6	4	9
		18	58	15
		3.6	11.6	3
		$M(XY)$	$M(X^2)$	$M(Y^2)$

$$\hat{c} = \frac{M(XY)}{M(X^2)} = \frac{3.6}{11.6} = 0.3103$$

$$M(\hat{E}^2) =$$

²Si osserva che per ragioni di stampa alcuni calcoli potrebbero essere duplicati con riferimento ai diversi modelli oggetto di studio



Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.

Si riportano i calcoli necessari² per ottenere le stime dei parametri e la misura di adattamento del modello

$$Y = cX + E$$

x_i	y_i	$x_i y_i$	x_i^2	y_i^2
2	-1	-2	4	1
3	0	0	9	0
4	1	4	16	1
5	2	10	25	4
2	3	6	4	9
		18	58	15
		3.6	11.6	3
		$M(XY)$	$M(X^2)$	$M(Y^2)$

$$\hat{c} = \frac{M(XY)}{M(X^2)} = \frac{3.6}{11.6} = 0.3103$$

$$M(\hat{E}^2) = M[Y \cdot (Y - \hat{c}X)] = M(Y^2) - \hat{c}M(XY) = 3 - 0.3103 \cdot 3.6 = 1.8828$$

²Si osserva che per ragioni di stampa alcuni calcoli potrebbero essere duplicati con riferimento ai diversi modelli oggetto di studio



Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.

$\hat{a} = 0.0588$, $\hat{b} = 0.2941$, indice adattamento $M(\hat{E}^2) = 1.8824$ $\hat{c} = 0.3103$, indice adattamento $M(\hat{E}^2) = 1.8828$



3. Confrontare gli indici di adattamento, commentando le eventuali differenze.

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.



3. Confrontare gli indici di adattamento, commentando le eventuali differenze.

I due modelli sono quasi equivalenti, è marginalmente migliore il modello

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.



Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.

3. Confrontare gli indici di adattamento, commentando le eventuali differenze.

I due modelli sono quasi equivalenti, è marginalmente migliore il modello

$Y^* = 0.0588 + 0.2941 \cdot X$ in quanto caratterizzato da residuo quadratico medio leggermente inferiore.



4. Senza svolgere i conti, cosa è possibile dire sui valori di $\eta_{Y|X}^2$ e $\eta_{X|Y}^2$.

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.



Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.

4. Senza svolgere i conti, cosa è possibile dire sui valori di $\eta_{Y|X}^2$ e $\eta_{X|Y}^2$.

Vale la relazione

$$\rho^2 (= 0.0588) < \eta_{Y|X}^2 < \eta_{X|Y}^2 = 1$$

sussiste, infatti, dipendenza funzionale di X da Y in quanto a ogni y_j corrisponde una sola x_i ; $0 < \eta_{Y|X}^2 < 1$ in quanto la distribuzione condizionata $Y|X = 2$ non è degenere.



Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.

5. Sulla base del miglior modello si preveda il valore di Y in corrispondenza di $X = 5$.



Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.

5. Sulla base del miglior modello si preveda il valore di Y in corrispondenza di $X = 5$.

$$Y^*(5) = 0.0588 + 0.2941 \cdot 5 = 1.5294$$



Nel corso della sperimentazione dei metodi di indagine per l'intolleranza al glutine viene messo a punto un nuovo test. I possibili esiti che il test può dare sono due: POSITIVO se il test dichiara l'individuo come intollerante o NEGATIVO se nell'esame non emerge una intolleranza al glutine.

Tuttavia il metodo di screening non è infallibile e la positività al test non predice perfettamente se l'individuo è davvero intollerante (e quindi MALATO) oppure se è SANO.

Sapendo che la probabilità che un individuo sia malato nella popolazione di riferimento è pari a 0.2. È inoltre noto che la diagnosi attraverso il test clinico ha le seguenti caratteristiche: la probabilità che un soggetto risulti positivo dato che nella realtà è malato è pari a 0.85 mentre la probabilità che un individuo risulti positivo dato che è sano è pari 0.1.

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.



Si risponda alle seguenti domande:

1. Quale è la probabilità che un generico individuo risulti positivo al test?

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.



Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.

Si risponda alle seguenti domande:

1. Quale è la probabilità che un generico individuo risulti positivo al test?

Indicati con M ed S gli eventi 'soggetto malato' e 'soggetto sano' e con $+$ e $-$ gli eventi 'positivo al test' e 'negativo al test' sono disponibili le seguenti informazioni:

$$0.2 =$$



Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.

Si risponda alle seguenti domande:

1. Quale è la probabilità che un generico individuo risulti positivo al test?

Indicati con M ed S gli eventi 'soggetto malato' e 'soggetto sano' e con $+$ e $-$ gli eventi 'positivo al test' e 'negativo al test' sono disponibili le seguenti informazioni:

$$0.2 = P(M)$$

$$0.85 =$$



Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.

Si risponda alle seguenti domande:

1. Quale è la probabilità che un generico individuo risulti positivo al test?

Indicati con M ed S gli eventi 'soggetto malato' e 'soggetto sano' e con $+$ e $-$ gli eventi 'positivo al test' e 'negativo al test' sono disponibili le seguenti informazioni:

$$0.2 = P(M)$$

$$0.85 = P(+|M)$$

$$0.1 =$$



Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.

Si risponda alle seguenti domande:

1. Quale è la probabilità che un generico individuo risulti positivo al test?

Indicati con M ed S gli eventi 'soggetto malato' e 'soggetto sano' e con $+$ e $-$ gli eventi 'positivo al test' e 'negativo al test' sono disponibili le seguenti informazioni:

$$0.2 = P(M)$$

$$0.85 = P(+|M)$$

$$0.1 = P(+|S)$$

Applicando la legge delle probabilità totali abbiamo:

$$P(+) =$$



Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.

Si risponda alle seguenti domande:

1. Quale è la probabilità che un generico individuo risulti positivo al test?

Indicati con M ed S gli eventi 'soggetto malato' e 'soggetto sano' e con $+$ e $-$ gli eventi 'positivo al test' e 'negativo al test' sono disponibili le seguenti informazioni:

$$0.2 = P(M)$$

$$0.85 = P(+|M)$$

$$0.1 = P(+|S)$$

Applicando la legge delle probabilità totali abbiamo:

$$P(+) = P(+ \cap (M \cup S)) =$$



Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.

Si risponda alle seguenti domande:

1. Quale è la probabilità che un generico individuo risulti positivo al test?

Indicati con M ed S gli eventi 'soggetto malato' e 'soggetto sano' e con $+$ e $-$ gli eventi 'positivo al test' e 'negativo al test' sono disponibili le seguenti informazioni:

$$0.2 = P(M)$$

$$0.85 = P(+|M)$$

$$0.1 = P(+|S)$$

Applicando la legge delle probabilità totali abbiamo:

$$\begin{aligned} P(+) &= P(+ \cap (M \cup S)) = P((+ \cap M) \cup (+ \cap S)) = \\ &= \end{aligned}$$



Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.

Si risponda alle seguenti domande:

1. Quale è la probabilità che un generico individuo risulti positivo al test?

Indicati con M ed S gli eventi 'soggetto malato' e 'soggetto sano' e con $+$ e $-$ gli eventi 'positivo al test' e 'negativo al test' sono disponibili le seguenti informazioni:

$$0.2 = P(M)$$

$$0.85 = P(+|M)$$

$$0.1 = P(+|S)$$

Applicando la legge delle probabilità totali abbiamo:

$$\begin{aligned} P(+) &= P(+ \cap (M \cup S)) = P((+ \cap M) \cup (+ \cap S)) = \\ &= P(+ \cap M) + P(+ \cap S) = \end{aligned}$$



Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.

Si risponda alle seguenti domande:

1. Quale è la probabilità che un generico individuo risulti positivo al test?

Indicati con M ed S gli eventi 'soggetto malato' e 'soggetto sano' e con $+$ e $-$ gli eventi 'positivo al test' e 'negativo al test' sono disponibili le seguenti informazioni:

$$0.2 = P(M)$$

$$0.85 = P(+|M)$$

$$0.1 = P(+|S)$$

Applicando la legge delle probabilità totali abbiamo:

$$\begin{aligned} P(+) &= P(+ \cap (M \cup S)) = P((+ \cap M) \cup (+ \cap S)) = \\ &= P(+ \cap M) + P(+ \cap S) = P(M) \cdot P(+|M) + P(S) \cdot P(+|S) = \\ &= \end{aligned}$$



Si risponda alle seguenti domande:

1. Quale è la probabilità che un generico individuo risulti positivo al test?

Indicati con M ed S gli eventi 'soggetto malato' e 'soggetto sano' e con $+$ e $-$ gli eventi 'positivo al test' e 'negativo al test' sono disponibili le seguenti informazioni:

$$0.2 = P(M)$$

$$0.85 = P(+|M)$$

$$0.1 = P(+|S)$$

Applicando la legge delle probabilità totali abbiamo:

$$\begin{aligned} P(+) &= P(+ \cap (M \cup S)) = P((+ \cap M) \cup (+ \cap S)) = \\ &= P(+ \cap M) + P(+ \cap S) = P(M) \cdot P(+|M) + P(S) \cdot P(+|S) = \\ &= P(M) \cdot P(+|M) + (1 - P(M)) \cdot P(+|S) = \\ &= \end{aligned}$$



Si risponda alle seguenti domande:

1. Quale è la probabilità che un generico individuo risulti positivo al test?

Indicati con M ed S gli eventi 'soggetto malato' e 'soggetto sano' e con $+$ e $-$ gli eventi 'positivo al test' e 'negativo al test' sono disponibili le seguenti informazioni:

$$0.2 = P(M)$$

$$0.85 = P(+|M)$$

$$0.1 = P(+|S)$$

Applicando la legge delle probabilità totali abbiamo:

$$\begin{aligned} P(+) &= P(+ \cap (M \cup S)) = P((+ \cap M) \cup (+ \cap S)) = \\ &= P(+ \cap M) + P(+ \cap S) = P(M) \cdot P(+|M) + P(S) \cdot P(+|S) = \\ &= P(M) \cdot P(+|M) + (1 - P(M)) \cdot P(+|S) = \\ &= 0.2 \cdot 0.85 + 0.8 \cdot 0.1 = \end{aligned}$$



Si risponda alle seguenti domande:

1. Quale è la probabilità che un generico individuo risulti positivo al test?

Indicati con M ed S gli eventi 'soggetto malato' e 'soggetto sano' e con $+$ e $-$ gli eventi 'positivo al test' e 'negativo al test' sono disponibili le seguenti informazioni:

$$0.2 = P(M)$$

$$0.85 = P(+|M)$$

$$0.1 = P(+|S)$$

Applicando la legge delle probabilità totali abbiamo:

$$\begin{aligned} P(+) &= P(+ \cap (M \cup S)) = P((+ \cap M) \cup (+ \cap S)) = \\ &= P(+ \cap M) + P(+ \cap S) = P(M) \cdot P(+|M) + P(S) \cdot P(+|S) = \\ &= P(M) \cdot P(+|M) + (1 - P(M)) \cdot P(+|S) = \\ &= 0.2 \cdot 0.85 + 0.8 \cdot 0.1 = 0.17 + 0.08 = \end{aligned}$$



Si risponda alle seguenti domande:

1. Quale è la probabilità che un generico individuo risulti positivo al test?

Indicati con M ed S gli eventi 'soggetto malato' e 'soggetto sano' e con $+$ e $-$ gli eventi 'positivo al test' e 'negativo al test' sono disponibili le seguenti informazioni:

$$0.2 = P(M)$$

$$0.85 = P(+|M)$$

$$0.1 = P(+|S)$$

Applicando la legge delle probabilità totali abbiamo:

$$\begin{aligned} P(+) &= P(+ \cap (M \cup S)) = P((+ \cap M) \cup (+ \cap S)) = \\ &= P(+ \cap M) + P(+ \cap S) = P(M) \cdot P(+|M) + P(S) \cdot P(+|S) = \\ &= P(M) \cdot P(+|M) + (1 - P(M)) \cdot P(+|S) = \\ &= 0.2 \cdot 0.85 + 0.8 \cdot 0.1 = 0.17 + 0.08 = 0.25 \end{aligned}$$



2. Quale è la probabilità che un individuo sia davvero malato dato che è risultato positivo al test?

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.



2. Quale è la probabilità che un individuo sia davvero malato dato che è risultato positivo al test?

$$P(M|+) =$$

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.



2. Quale è la probabilità che un individuo sia davvero malato dato che è risultato positivo al test?

$$P(M|+) = \frac{P(M \cap +)}{P(+)} =$$

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.



Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.

2. Quale è la probabilità che un individuo sia davvero malato dato che è risultato positivo al test?

$$P(M|+) = \frac{P(M \cap +)}{P(+)} = \frac{0.17}{0.25} =$$



Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.

2. Quale è la probabilità che un individuo sia davvero malato dato che è risultato positivo al test?

$$P(M|+) = \frac{P(M \cap +)}{P(+)} = \frac{0.17}{0.25} = 0.68$$



3. Estratti con reinserimento 10 individui dalla popolazione, dire quale è la probabilità che almeno uno di essi sia malato.

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.



Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.

3. Estratti con reinserimento 10 individui dalla popolazione, dire quale è la probabilità che almeno uno di essi sia malato.

Sia X la variabile casuale 'n. di individui malati tra i 10 estratti'. Supponendo che la popolazione sia composta da un numero elevato di soggetti le estrazioni 'effettuate in maniera casuale' possono ritenersi indipendenti.

Abbiamo

$$X \sim$$



Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.

3. Estratti con reinserimento 10 individui dalla popolazione, dire quale è la probabilità che almeno uno di essi sia malato.

Sia X la variabile casuale 'n. di individui malati tra i 10 estratti'. Supponendo che la popolazione sia composta da un numero elevato di soggetti le estrazioni 'effettuate in maniera casuale' possono ritenersi indipendenti.

Abbiamo

$$X \sim \text{Bin}(n = 10, p = 0.2)$$

e

$$P(X \geq 1) = 1 - P(X = 0) =$$



Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.

3. Estratti con reinserimento 10 individui dalla popolazione, dire quale è la probabilità che almeno uno di essi sia malato.

Sia X la variabile casuale 'n. di individui malati tra i 10 estratti'. Supponendo che la popolazione sia composta da un numero elevato di soggetti le estrazioni 'effettuate in maniera casuale' possono ritenersi indipendenti.

Abbiamo

$$X \sim \text{Bin}(n = 10, p = 0.2)$$

e

$$P(X \geq 1) = 1 - P(X = 0) = 1 - \binom{10}{0} 0.2^0 \cdot 0.8^{10} =$$



Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.

3. Estratti con reinserimento 10 individui dalla popolazione, dire quale è la probabilità che almeno uno di essi sia malato.

Sia X la variabile casuale 'n. di individui malati tra i 10 estratti'. Supponendo che la popolazione sia composta da un numero elevato di soggetti le estrazioni 'effettuate in maniera casuale' possono ritenersi indipendenti.

Abbiamo

$$X \sim \text{Bin}(n = 10, p = 0.2)$$

e

$$P(X \geq 1) = 1 - P(X = 0) = 1 - \binom{10}{0} 0.2^0 \cdot 0.8^{10} = 1 - 0.1074 =$$



Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.

3. Estratti con reinserimento 10 individui dalla popolazione, dire quale è la probabilità che almeno uno di essi sia malato.

Sia X la variabile casuale 'n. di individui malati tra i 10 estratti'. Supponendo che la popolazione sia composta da un numero elevato di soggetti le estrazioni 'effettuate in maniera casuale' possono ritenersi indipendenti.

Abbiamo

$$X \sim \text{Bin}(n = 10, p = 0.2)$$

e

$$P(X \geq 1) = 1 - P(X = 0) = 1 - \binom{10}{0} 0.2^0 \cdot 0.8^{10} = 1 - 0.1074 = 0.8926.$$

Tema T294-2

Prova scritta di STATISTICA del 13.07.2017

Prova scritta di
STATISTICA del
13.07.2017

Giuseppe Boari
Gabriele Cantaluppi



UNIVERSITÀ
CATTOLICA
del Sacro Cuore

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.

Giuseppe Boari
Gabriele Cantaluppi
Università Cattolica del Sacro Cuore, Milano

Esercizio 1



Il prospetto seguente riporta la distribuzione congiunta del carattere qualitativo X e del carattere quantitativo Y .

$X \setminus Y$	1	2	3
x_1	2	2	0
x_2	0	2	2

[Esercizio 1.](#)

[Esercizio 2.](#)

[Esercizio 3.](#)

[Esercizio 4.](#)



1. Si rappresenti graficamente la distribuzione di frequenza della condizionata $Y|x_1$ e si misuri l'eterogeneità delle distribuzioni $X|y_2$ e $X|y_3$.

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.



Si considera la distribuzione condizionata $Y|x_1$

y_j	n_{1j}
1	2
2	2
3	0

Esercizio 1.

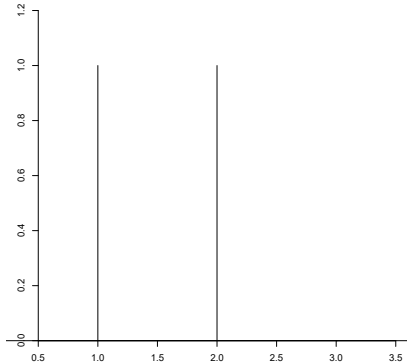
Esercizio 2.

Esercizio 3.

Esercizio 4.



Si tratta di un carattere quantitativo con modalità non raggruppate in classi che può essere rappresentato graficamente mediante un grafico a bastoncini.



Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.



Si considera la distribuzione condizionata $X|y_2$

x_i	n_{i2}
x_1	2
x_2	2

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.



Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.

La mutabile X presenta massima eterogeneità, quindi gli indici di Gini e di Frosini assumono valore massimo, $F_N = G_N = 1$.



Si considera la distribuzione condizionata $X|y_3$

x_i	n_{i3}
x_1	0
x_2	2

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.



La mutabile X è degenere ed è, quindi, caratterizzata da assenza di eterogeneità. Gli indici di Gini e di Frosini assumono valore minimo, $F_N = G_N = 0$.

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.



2. Si confrontino le distribuzioni condizionate $Y|x_1$ e $Y|x_2$ mediante opportuni indici di posizione e di variabilità.

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.



Si considera la distribuzione condizionata $Y|x_1$

y_j	n_{1j}
1	2
2	2
3	0

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.



Si tratta di un carattere quantitativo con modalità non raggruppate in classi.

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.



La media aritmetica

$$M(Y|x_1) = \frac{1}{n_{1\bullet}} \sum_{j=1}^k y_j n_{1j}$$

risulta 1.5.

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.



La media aritmetica

$$M(Y|x_1) = \frac{1}{n_{1\bullet}} \sum_{j=1}^k y_j n_{1j}$$

risulta 1.5.
Si ottiene

y_j	n_{1j}	$y_j n_{1j}$
1	2	2
2	2	4
3	0	0
		4
		6

da cui segue

$$M(Y|x_1) = \frac{1}{n_{1\bullet}} \sum_{j=1}^k y_j n_{1j} = \frac{6}{4} = 1.5.$$

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.



La varianza

$$\text{Var}(Y|x_1) = M(Y^2|x_1) - [M(Y|x_1)]^2$$

risulta 0.25.

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.



La varianza

$$\text{Var}(Y|x_1) = M(Y^2|x_1) - [M(Y|x_1)]^2$$

risulta 0.25. Si ottiene

y_j	n_{1j}	$y_j n_{1j}$	$y_j^2 n_{1j}$
1	2	2	2
2	2	4	8
3	0	0	0
4		6	10

da cui segue

$$\text{Var}(Y|x_1) = \frac{10}{4} - 1.5^2 = 2.5 - 2.25 = 0.25.$$

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.



Per effettuare confronti in termini di variabilità tra caratteri quantitativi occorre utilizzare una misura relativa, ad esempio il coefficiente di variazione

$$CV = \frac{\sigma}{\mu} = \frac{\sqrt{\text{Var}(Y|x_1)}}{M(Y|x_1)} = \frac{\sqrt{0.25}}{1.5} = \frac{0.5}{1.5} = 0.3333.$$

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.



Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.

Si osserva che per la distribuzione condizionata $Y|x_2$ vale:

$$Y|x_2 = Y|x_1 + 1$$

vale, quindi,

$$\begin{aligned} M(Y|x_2) &= M(Y|x_1) + 1 = 1.5 + 1 = 2.5, \\ \text{Var}(Y|x_2) &= \text{Var}(Y|x_1) = 0.25. \end{aligned}$$

Si riporta, comunque, anche il calcolo esteso (non necessario) di tali indici



Si considera la distribuzione condizionata $Y|x_2$

y_j	n_{2j}
1	0
2	2
3	2

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.



Si tratta di un carattere quantitativo con modalità non raggruppate in classi.

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.



La media aritmetica

$$M(Y|x_2) = \frac{1}{n_{2\bullet}} \sum_{j=1}^k y_j n_{2j}$$

risulta 2.5.

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.



La media aritmetica

$$M(Y|x_2) = \frac{1}{n_{2\bullet}} \sum_{j=1}^k y_j n_{2j}$$

risulta 2.5.

Si ottiene

y_j	n_{2j}	$y_j n_{2j}$
1	0	0
2	2	4
3	2	6
	4	10

da cui segue

$$M(Y|x_2) = \frac{1}{n_{2\bullet}} \sum_{j=1}^k y_j n_{2j} = \frac{10}{4} = 2.5.$$

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.



La varianza

$$\text{Var}(Y|x_2) = M(Y^2|x_2) - [M(Y|x_2)]^2$$

risulta 0.25.

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.



La varianza

$$\text{Var}(Y|x_2) = M(Y^2|x_2) - [M(Y|x_2)]^2$$

risulta 0.25. Si ottiene

y_j	n_{2j}	$y_j n_{2j}$	$y_j^2 n_{2j}$
1	0	0	0
2	2	4	8
3	2	6	18
4		10	26

da cui segue

$$\text{Var}(Y|x_2) = \frac{26}{4} - 2.5^2 = 6.5 - 6.25 = 0.25.$$

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.



Per effettuare confronti in termini di variabilità tra caratteri quantitativi occorre utilizzare una misura relativa, ad esempio il coefficiente di variazione

$$CV = \frac{\sigma}{\mu} = \frac{\sqrt{\text{Var}(Y|x_2)}}{M(Y|x_2)} = \frac{\sqrt{0.25}}{2.5} = \frac{0.5}{2.5} = 0.2.$$

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.



Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.

$Y|_{x_1}$ Indice di posizione: $M(Y|_{x_1}) = 1.5$ Indice di variabilità: $CV(Y|_{x_1}) = 0.3333$

$Y|_{x_2}$ Indice di posizione: $M(Y|_{x_2}) = 2.5$ Indice di variabilità: $CV(Y|_{x_2}) = 0.25$

La distribuzione condizionata $Y|_{x_2}$ presenta un livello medio superiore, ma una variabilità inferiore a $Y|_{x_1}$.



3. Si determini il grado di connessione tra le variabili X e Y mediante il calcolo di un opportuno indice.

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.



Dalla distribuzione delle frequenze congiunte n_{ij} :

	1	2	3	$n_{i\bullet}$
x_1	2	2	0	4
x_2	0	2	2	4
$n_{\bullet j}$	2	4	2	8

è possibile ottenere le frequenze teoriche nella situazione di indipendenza stocastica

$$\hat{n}_{ij} = \frac{n_{i\bullet} n_{\bullet j}}{n}:$$

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.



Dalla distribuzione delle frequenze congiunte n_{ij} :

	1	2	3	$n_{i\bullet}$
x_1	2	2	0	4
x_2	0	2	2	4
$n_{\bullet j}$	2	4	2	8

è possibile ottenere le frequenze teoriche nella situazione di indipendenza stocastica

$$\hat{n}_{ij} = \frac{n_{i\bullet} n_{\bullet j}}{n}:$$

	1	2	3	$n_{i\bullet}$
x_1	1	2	1	4
x_2	1	2	1	4
$n_{\bullet j}$	2	4	2	8

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.



le contingenze assolute $n_{ij} - \hat{n}_{ij}$:

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.



le contingenze assolute $n_{ij} - \hat{n}_{ij}$:

	1	2	3
x_1	1	0	-1
x_2	-1	0	1

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.



i rapporti di contingenza $\frac{n_{ij}}{\hat{n}_{ij}}$:

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.



i rapporti di contingenza $\frac{n_{ij}}{\hat{n}_{ij}}$:

	1	2	3
x_1	2	1	0
x_2	0	1	2

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.



e le contingenze relative, espresse in termini percentuali, $100 \left(\frac{n_{ij} - \hat{n}_{ij}}{\hat{n}_{ij}} \right) \% = 100 \left(\frac{n_{ij}}{\hat{n}_{ij}} - 1 \right) \%$:

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.



e le contingenze relative, espresse in termini percentuali, $100 \left(\frac{n_{ij} - \hat{n}_{ij}}{\hat{n}_{ij}} \right) \% = 100 \left(\frac{n_{ij}}{\hat{n}_{ij}} - 1 \right) \%$:

	1	2	3
x_1	100	0	-100
x_2	-100	0	100

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.



Risultati intermedi $\frac{(n_{ij} - \hat{n}_{ij})^2}{\hat{n}_{ij}}$ per il calcolo del χ^2

	1	2	3
x_1	1	0	1
x_2	1	0	1

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.



Risultati intermedi $\frac{n_{ij}^2}{n_{i\bullet} \cdot n_{\bullet j}}$ per il calcolo del χ^2 (formula operativa)

	1	2	3
x_1	0.5	0.25	0.0
x_2	0.0	0.25	0.5

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.



Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.

$$\chi^2 = \sum_{i=1}^h \sum_{j=1}^k \frac{(n_{ij} - \hat{n}_{ij})^2}{\hat{n}_{ij}} = 4,$$

$$\chi_N^2 = \frac{\chi^2}{n \cdot \min(h-1, k-1)} = 0.5$$

$$\chi_N = \sqrt{\chi_N^2} = 0.7071.$$

Sussiste, quindi, un livello di connessione medio alto.



La seguente tabella riporta la serie storica del fatturato annuo di un'azienda:

t	x_t
2012	50
2013	55
2014	52.25
2015	62.7
2016	62.7

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.



1. Si costruisca la serie dei numeri indici a base fissa con base 2013 e si calcoli, commentando il risultato, il valore di ${}_{2015}I_{2012}$, numero indice al tempo 2012 con base 2015.

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.



1. Si costruisca la serie dei numeri indici a base fissa con base 2013 e si calcoli, commentando il risultato, il valore di ${}_{2015}I_{2012}$, numero indice al tempo 2012 con base 2015. Si ricorda che ${}_{2013}I_t = \frac{x_t}{x_{2013}}$.

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.



1. Si costruisca la serie dei numeri indici a base fissa con base 2013 e si calcoli, commentando il risultato, il valore di ${}_{2015}I_{2012}$, numero indice al tempo 2012 con base 2015. Si ricorda che ${}_{2013}I_t = \frac{x_t}{x_{2013}}$. Si ottiene quindi

Anno	x_t	NIBF (2013)
2012	50.00	0.9091
2013	55.00	1.0000
2014	52.25	0.9500
2015	62.70	1.1400
2016	62.70	1.1400

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.



Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.



$${}_{2015}I_{2012} = \frac{x_{2012}}{x_{2015}} = \frac{50}{62.7} = 0.7974$$

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.



Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.

$${}_{2015}I_{2012} = \frac{x_{2012}}{x_{2015}} = \frac{50}{62.7} = 0.7974$$

La grandezza al tempo 2012 era inferiore del 20.26% rispetto alla grandezza al tempo 2015. Ovvero il valore della grandezza nel 2012 era il 79.74% del valore registrato nel 2015.



2. Si calcoli il tasso medio di variazione dal 2012 al 2016 e si commenti il risultato ottenuto.

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.



2. Si calcoli il tasso medio di variazione dal 2012 al 2016 e si commenti il risultato ottenuto.
Si ricorda che

$${}_{2012}I_{2016} = {}_{2012}I_{2013} \cdot {}_{2013}I_{2014} \cdot {}_{2014}I_{2015} \cdot {}_{2015}I_{2016}$$

Si deve trovare il numero indice a base mobile medio.

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.



Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.

2. Si calcoli il tasso medio di variazione dal 2012 al 2016 e si commenti il risultato ottenuto. Si ricorda che

$${}_{2012}I_{2016} = {}_{2012}I_{2013} \cdot {}_{2013}I_{2014} \cdot {}_{2014}I_{2015} \cdot {}_{2015}I_{2016}$$

Si deve trovare il numero indice a base mobile medio. Applicando il criterio di scelta della media secondo Chisini si ottiene

$${}_{2012}I_{2016} = {}_{2012}I_{2013} \cdot {}_{2013}I_{2014} \cdot {}_{2014}I_{2015} \cdot {}_{2015}I_{2016} = \alpha \cdot \alpha \cdot \alpha \cdot \alpha = \alpha^4.$$



Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.

2. Si calcoli il tasso medio di variazione dal 2012 al 2016 e si commenti il risultato ottenuto. Si ricorda che

$${}_{2012}I_{2016} = {}_{2012}I_{2013} \cdot {}_{2013}I_{2014} \cdot {}_{2014}I_{2015} \cdot {}_{2015}I_{2016}$$

Si deve trovare il numero indice a base mobile medio. Applicando il criterio di scelta della media secondo Chisini si ottiene

$${}_{2012}I_{2016} = {}_{2012}I_{2013} \cdot {}_{2013}I_{2014} \cdot {}_{2014}I_{2015} \cdot {}_{2015}I_{2016} = \alpha \cdot \alpha \cdot \alpha \cdot \alpha = \alpha^4.$$

Quindi

$$\alpha = \sqrt[4]{{}_{2012}I_{2016}} = \sqrt[4]{{}_{2012}I_{2013} \cdot {}_{2013}I_{2014} \cdot {}_{2014}I_{2015} \cdot {}_{2015}I_{2016}} = 1.0582.$$

Da cui segue una variazione relativa media pari a 0.0582, corrispondente a un incremento medio annuo tra il 2012 e il 2016 del 5.82%.



Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.

Il seguente prospetto riporta i valori rilevati in corrispondenza a 5 unità statistiche con riferimento a due caratteri quantitativi X e Y

x_i	1	2	3	4	5
y_i	2	4	4	4	6

1. Si rappresentino graficamente le coppie dei dati ipotizzando Y in funzione di X .



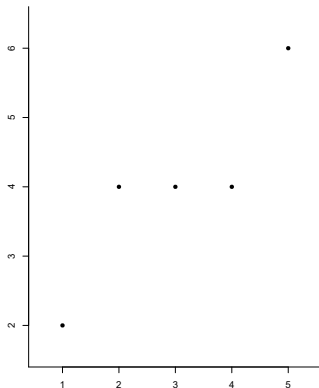
Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.

Le coppie di dati possono essere rappresentate mediante un grafico a dispersione.





2. Si calcolino, secondo il metodo dei minimi quadrati, i parametri dei seguenti modelli di regressione:

$$I) Y^* = a + b \cdot X \qquad II) Y^* = c$$

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.



Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.

Si riportano i calcoli necessari¹ per ottenere le stime dei parametri e gli indici di adattamento del modello

$$Y = a + bX + E$$

x_i	y_i	$x_i y_i$	x_i^2	y_i^2
1	2	2	1	4
2	4	8	4	16
3	4	12	9	16
4	4	16	16	16
5	6	30	25	36
15	20	68	55	88
3	4	13.6	11	17.6
$M(X)$	$M(Y)$	$M(XY)$	$M(X^2)$	$M(Y^2)$

$$\hat{b} =$$

¹Si osserva che per ragioni di stampa alcuni calcoli potrebbero essere duplicati con riferimento ai diversi modelli oggetto di studio



Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.

Si riportano i calcoli necessari¹ per ottenere le stime dei parametri e gli indici di adattamento del modello

$$Y = a + bX + E$$

x_i	y_i	$x_i y_i$	x_i^2	y_i^2
1	2	2	1	4
2	4	8	4	16
3	4	12	9	16
4	4	16	16	16
5	6	30	25	36
15	20	68	55	88
3	4	13.6	11	17.6
$M(X)$	$M(Y)$	$M(XY)$	$M(X^2)$	$M(Y^2)$

$$\begin{aligned}\hat{b} &= \frac{\text{Cov}(X, Y)}{\text{Var}(X)} = \frac{M(XY) - M(X)M(Y)}{M(X^2) - [M(X)]^2} = \\ &= \frac{13.6 - 3 \cdot 4}{11 - 3^2} = \frac{1.6}{2} = 0.8 \\ \hat{a} &= \end{aligned}$$

¹Si osserva che per ragioni di stampa alcuni calcoli potrebbero essere duplicati con riferimento ai diversi modelli oggetto di studio



Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.

Si riportano i calcoli necessari¹ per ottenere le stime dei parametri e gli indici di adattamento del modello

$$Y = a + bX + E$$

x_i	y_i	$x_i y_i$	x_i^2	y_i^2
1	2	2	1	4
2	4	8	4	16
3	4	12	9	16
4	4	16	16	16
5	6	30	25	36
15	20	68	55	88
3	4	13.6	11	17.6
$M(X)$	$M(Y)$	$M(XY)$	$M(X^2)$	$M(Y^2)$

$$\hat{b} = \frac{\text{Cov}(X, Y)}{\text{Var}(X)} = \frac{M(XY) - M(X)M(Y)}{M(X^2) - [M(X)]^2} =$$

$$= \frac{13.6 - 3 \cdot 4}{11 - 3^2} = \frac{1.6}{2} = 0.8$$

$$\hat{a} = M(Y) - \hat{b}M(X) = 1.6$$

$$\rho^2 =$$

¹Si osserva che per ragioni di stampa alcuni calcoli potrebbero essere duplicati con riferimento ai diversi modelli oggetto di studio



Si riportano i calcoli necessari¹ per ottenere le stime dei parametri e gli indici di adattamento del modello

$$Y = a + bX + E$$

x_i	y_i	$x_i y_i$	x_i^2	y_i^2
1	2	2	1	4
2	4	8	4	16
3	4	12	9	16
4	4	16	16	16
5	6	30	25	36
15	20	68	55	88
3	4	13.6	11	17.6
$M(X)$	$M(Y)$	$M(XY)$	$M(X^2)$	$M(Y^2)$

$$\begin{aligned}\hat{b} &= \frac{\text{Cov}(X, Y)}{\text{Var}(X)} = \frac{M(XY) - M(X)M(Y)}{M(X^2) - [M(X)]^2} = \\ &= \frac{13.6 - 3 \cdot 4}{11 - 3^2} = \frac{1.6}{2} = 0.8\end{aligned}$$

$$\hat{a} = M(Y) - \hat{b}M(X) = 1.6$$

$$\rho^2 = \frac{[\text{Cov}(X, Y)]^2}{\text{Var}(X) \text{Var}(Y)} = \frac{1.6^2}{2 \cdot 1.6} = 0.8$$

¹Si osserva che per ragioni di stampa alcuni calcoli potrebbero essere duplicati con riferimento ai diversi modelli oggetto di studio



Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.

Si riportano i calcoli necessari² per ottenere le stime dei parametri e gli indici di adattamento del modello

$$Y = c + E$$

x_i	y_i	y_i^2
1	2	4
2	4	16
3	4	16
4	4	16
5	6	36
20		88
4		17.6
$M(Y)$		$M(Y^2)$

$$\hat{c} =$$

²Si osserva che per ragioni di stampa alcuni calcoli potrebbero essere duplicati con riferimento ai diversi modelli oggetto di studio



Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.

Si riportano i calcoli necessari² per ottenere le stime dei parametri e gli indici di adattamento del modello

$$Y = c + E$$

x_i	y_i	y_i^2
1	2	4
2	4	16
3	4	16
4	4	16
5	6	36
20		88
4		17.6
$M(Y)$		$M(Y^2)$

$$\hat{c} = M(Y) = 4$$

$$R^2 =$$

²Si osserva che per ragioni di stampa alcuni calcoli potrebbero essere duplicati con riferimento ai diversi modelli oggetto di studio



Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.

Si riportano i calcoli necessari² per ottenere le stime dei parametri e gli indici di adattamento del modello

$$Y = c + E$$

x_i	y_i	y_i^2
1	2	4
2	4	16
3	4	16
4	4	16
5	6	36
20		88
4		17.6
$M(Y)$		$M(Y^2)$

$$\hat{c} = M(Y) = 4$$

$$R^2 = 0$$

²Si osserva che per ragioni di stampa alcuni calcoli potrebbero essere duplicati con riferimento ai diversi modelli oggetto di studio



Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.

$\hat{a} = 1.6$, $\hat{b} = 0.8$, indice adattamento $\rho^2 = 0.8$ $\hat{c} = 4$, indice adattamento $R^2 = 0$



Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.

3. Confrontare gli indici di adattamento, commentando i risultati.



Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.

3. Confrontare gli indici di adattamento, commentando i risultati.

Il modello migliore è



Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.

3. Confrontare gli indici di adattamento, commentando i risultati.

Il modello migliore è $Y^* = 1.6 + 0.8 \cdot X$ in quanto caratterizzato da indice di adattamento più elevato.



4. Senza svolgere i conti, cosa è possibile dire con riferimento ai dati in esame riguardo alla relazione tra ρ^2 ed $\eta_{Y|X}^2$.

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.



Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.

4. Senza svolgere i conti, cosa è possibile dire con riferimento ai dati in esame riguardo alla relazione tra ρ^2 ed $\eta_{Y|X}^2$.

Vale la relazione

$$\rho^2 (= 0.8) < \eta_{Y|X}^2 = 1$$

sussiste, infatti, dipendenza funzionale di Y da X in quanto le distribuzioni condizionate $Y|X_i$ sono degeneri.



5. Si motivi perché $\eta_{X|Y}^2 < \eta_{Y|X}^2$.

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.



Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.

5. Si motivi perché $\eta_{X|Y}^2 < \eta_{Y|X}^2$.

Vale la relazione

$$\eta_{X|Y}^2 = \rho^2 (= 0.8) < \eta_{Y|X}^2 = 1$$

in quanto le medie condizionate $M(X|y_j)$ sono allineate.

Vale, infatti, quanto segue.



La funzione di regressione associa a ogni valore y appartenente all'insieme di definizione della variabile Y la media di X condizionata a y , $M(X|y)$

y	$M(X y)$
2	1
4	$\frac{1}{3}(2 + 3 + 4) = 3$
6	5

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.



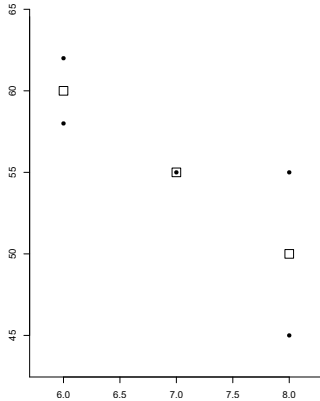
Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.

Le coppie di dati possono essere rappresentate mediante un grafico a dispersione.





L'urna $U1$ contiene 6 palline nere e 2 palline rosse. L'urna $U2$ contiene 18 palline nere e 1 pallina rossa.

1. Si calcoli la probabilità che estraendo prima una pallina dall'urna $U1$ e poi una pallina dall'urna $U2$ la prima sia rossa e la seconda sia nera.

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.



L'urna $U1$ contiene 6 palline nere e 2 palline rosse. L'urna $U2$ contiene 18 palline nere e 1 pallina rossa.

1. Si calcoli la probabilità che estraendo prima una pallina dall'urna $U1$ e poi una pallina dall'urna $U2$ la prima sia rossa e la seconda sia nera.

$$P(R_1 \cap N_2) =$$

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.



L'urna $U1$ contiene 6 palline nere e 2 palline rosse. L'urna $U2$ contiene 18 palline nere e 1 pallina rossa.

1. Si calcoli la probabilità che estraendo prima una pallina dall'urna $U1$ e poi una pallina dall'urna $U2$ la prima sia rossa e la seconda sia nera.

$$P(R_1 \cap N_2) = P(R_1) \cdot P(N_2) =$$

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.



L'urna $U1$ contiene 6 palline nere e 2 palline rosse. L'urna $U2$ contiene 18 palline nere e 1 pallina rossa.

1. Si calcoli la probabilità che estraendo prima una pallina dall'urna $U1$ e poi una pallina dall'urna $U2$ la prima sia rossa e la seconda sia nera.

$$P(R_1 \cap N_2) = P(R_1) \cdot P(N_2) = \frac{2}{8} \cdot \frac{18}{19} =$$

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.



L'urna $U1$ contiene 6 palline nere e 2 palline rosse. L'urna $U2$ contiene 18 palline nere e 1 pallina rossa.

1. Si calcoli la probabilità che estraendo prima una pallina dall'urna $U1$ e poi una pallina dall'urna $U2$ la prima sia rossa e la seconda sia nera.

$$P(R_1 \cap N_2) = P(R_1) \cdot P(N_2) = \frac{2}{8} \cdot \frac{18}{19} = 0.2368.$$

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.



2. Si calcoli la probabilità che estraendo con reimmissione 40 palline dall'urna U_2 almeno una sia rossa.

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.



2. Si calcoli la probabilità che estraendo con reimmissione 40 palline dall'urna U_2 almeno una sia rossa.

Sia X la variabile casuale 'n. di palline rosse tra le 40 estratte con reimmissione'.
Abbiamo

$$X \sim$$

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.



Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.

2. Si calcoli la probabilità che estraendo con reimmissione 40 palline dall'urna $U2$ almeno una sia rossa.

Sia X la variabile casuale 'n. di palline rosse tra le 40 estratte con reimmissione'.
Abbiamo

$$X \sim \text{Bin}(n = 40, p = \frac{1}{19} = 0.0526)$$

e

$$P(X \geq 1) = 1 - P(X = 0) =$$



Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.

2. Si calcoli la probabilità che estraendo con reimmissione 40 palline dall'urna $U2$ almeno una sia rossa.

Sia X la variabile casuale 'n. di palline rosse tra le 40 estratte con reimmissione'.
Abbiamo

$$X \sim \text{Bin}(n = 40, p = \frac{1}{19} = 0.0526)$$

e

$$P(X \geq 1) = 1 - P(X = 0) = 1 - \binom{40}{0} 0.0526^0 \cdot 0.9474^{40} =$$



Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.

2. Si calcoli la probabilità che estraendo con reimmissione 40 palline dall'urna $U2$ almeno una sia rossa.

Sia X la variabile casuale 'n. di palline rosse tra le 40 estratte con reimmissione'.
Abbiamo

$$X \sim \text{Bin}(n = 40, p = \frac{1}{19} = 0.0526)$$

e

$$P(X \geq 1) = 1 - P(X = 0) = 1 - \binom{40}{0} 0.0526^0 \cdot 0.9474^{40} = 1 - 0.1150 =$$



2. Si calcoli la probabilità che estraendo con reimmissione 40 palline dall'urna U_2 almeno una sia rossa.

Sia X la variabile casuale 'n. di palline rosse tra le 40 estratte con reimmissione'.
Abbiamo

$$X \sim \text{Bin}(n = 40, p = \frac{1}{19} = 0.0526)$$

e

$$P(X \geq 1) = 1 - P(X = 0) = 1 - \binom{40}{0} 0.0526^0 \cdot 0.9474^{40} = 1 - 0.1150 = 0.8850.$$

Si osserva che non sussistono le condizioni per l'approssimazione della variabile casuale binomiale con la normale in quanto $np = 2.1053 < 5$, pur essendo $n(1 - p) = 37.8947 > 5$.



Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.

3. Si estraggano, con reimmissione, due palline dall'urna $U1$ e le si mettano nell'urna $U2$. Si calcoli la probabilità che estraendo con reimmissione 2 palline dall'urna $U2$ (modificata) si ottengano 2 palline nere.



Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.

3. Si estraggano, con reimmissione, due palline dall'urna U_1 e le si mettano nell'urna U_2 . Si calcoli la probabilità che estraendo con reimmissione 2 palline dall'urna U_2 (modificata) si ottengano 2 palline nere.

Ricordando che la frazione di palline nere nell'urna U_1 è pari a 0.75, in base alla legge delle probabilità totali abbiamo



Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.

3. Si estraggano, con reimmissione, due palline dall'urna U_1 e le si mettano nell'urna U_2 . Si calcoli la probabilità che estraendo con reimmissione 2 palline dall'urna U_2 (modificata) si ottengano 2 palline nere.

Ricordando che la frazione di palline nere nell'urna U_1 è pari a 0.75, in base alla legge delle probabilità totali abbiamo

$$P(2N) = P(((2R)_1 \cap 2N) \cup ((2N)_1 \cap 2N) \cup ((1N1R)_1 \cap 2N)) =$$
$$=$$



Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.

3. Si estraggano, con reimmissione, due palline dall'urna U_1 e le si mettano nell'urna U_2 . Si calcoli la probabilità che estraendo con reimmissione 2 palline dall'urna U_2 (modificata) si ottengano 2 palline nere.

Ricordando che la frazione di palline nere nell'urna U_1 è pari a 0.75, in base alla legge delle probabilità totali abbiamo

$$\begin{aligned} P(2N) &= P(((2R)_1 \cap 2N) \cup ((2N)_1 \cap 2N) \cup ((1N1R)_1 \cap 2N)) = \\ &= P((2R)_1 \cap 2N) + P((2N)_1 \cap 2N) + P((1N1R)_1 \cap 2N) = \\ &= \end{aligned}$$



Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.

3. Si estraggano, con reimmissione, due palline dall'urna U_1 e le si mettano nell'urna U_2 . Si calcoli la probabilità che estraendo con reimmissione 2 palline dall'urna U_2 (modificata) si ottengano 2 palline nere.

Ricordando che la frazione di palline nere nell'urna U_1 è pari a 0.75, in base alla legge delle probabilità totali abbiamo

$$\begin{aligned} P(2N) &= P(((2R)_1 \cap 2N) \cup ((2N)_1 \cap 2N) \cup ((1N1R)_1 \cap 2N)) = \\ &= P((2R)_1 \cap 2N) + P((2N)_1 \cap 2N) + P((1N1R)_1 \cap 2N) = \\ &= P((2R)_1)P(2N|(2R)_1) + P((2N)_1)P(2N|(2N)_1) + P((1N1R)_1)P(2N|(1N1R)_1) = \\ &= \end{aligned}$$



Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.

3. Si estraggano, con reimmissione, due palline dall'urna U_1 e le si mettano nell'urna U_2 . Si calcoli la probabilità che estraendo con reimmissione 2 palline dall'urna U_2 (modificata) si ottengano 2 palline nere.

Ricordando che la frazione di palline nere nell'urna U_1 è pari a 0.75, in base alla legge delle probabilità totali abbiamo

$$\begin{aligned} P(2N) &= P(((2R)_1 \cap 2N) \cup ((2N)_1 \cap 2N) \cup ((1N1R)_1 \cap 2N)) = \\ &= P((2R)_1 \cap 2N) + P((2N)_1 \cap 2N) + P((1N1R)_1 \cap 2N) = \\ &= P((2R)_1)P(2N|(2R)_1) + P((2N)_1)P(2N|(2N)_1) + P((1N1R)_1)P(2N|(1N1R)_1) = \\ &= \left\{ \binom{2}{0} 0.75^0 \cdot 0.25^2 \right\} \cdot \left\{ \binom{2}{2} \frac{18^2}{21} \cdot \frac{3}{21}^0 \right\} + \end{aligned}$$



Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.

3. Si estraggano, con reimmissione, due palline dall'urna U_1 e le si mettano nell'urna U_2 . Si calcoli la probabilità che estraendo con reimmissione 2 palline dall'urna U_2 (modificata) si ottengano 2 palline nere.

Ricordando che la frazione di palline nere nell'urna U_1 è pari a 0.75, in base alla legge delle probabilità totali abbiamo

$$\begin{aligned} P(2N) &= P(((2R)_1 \cap 2N) \cup ((2N)_1 \cap 2N) \cup ((1N1R)_1 \cap 2N)) = \\ &= P((2R)_1 \cap 2N) + P((2N)_1 \cap 2N) + P((1N1R)_1 \cap 2N) = \\ &= P((2R)_1)P(2N|(2R)_1) + P((2N)_1)P(2N|(2N)_1) + P((1N1R)_1)P(2N|(1N1R)_1) = \\ &= \left\{ \binom{2}{0} 0.75^0 \cdot 0.25^2 \right\} \cdot \left\{ \binom{2}{2} \frac{18}{21}^2 \cdot \frac{3}{21}^0 \right\} + \left\{ \binom{2}{2} 0.75^2 \cdot 0.25^0 \right\} \cdot \left\{ \binom{2}{2} \frac{20}{21}^2 \cdot \frac{1}{21}^0 \right\} + \end{aligned}$$



Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.

3. Si estraggano, con reimmissione, due palline dall'urna U_1 e le si mettano nell'urna U_2 . Si calcoli la probabilità che estraendo con reimmissione 2 palline dall'urna U_2 (modificata) si ottengano 2 palline nere.

Ricordando che la frazione di palline nere nell'urna U_1 è pari a 0.75, in base alla legge delle probabilità totali abbiamo

$$\begin{aligned}
 P(2N) &= P(((2R)_1 \cap 2N) \cup ((2N)_1 \cap 2N) \cup ((1N1R)_1 \cap 2N)) = \\
 &= P((2R)_1 \cap 2N) + P((2N)_1 \cap 2N) + P((1N1R)_1 \cap 2N) = \\
 &= P((2R)_1)P(2N|(2R)_1) + P((2N)_1)P(2N|(2N)_1) + P((1N1R)_1)P(2N|(1N1R)_1) = \\
 &= \left\{ \binom{2}{0} 0.75^0 \cdot 0.25^2 \right\} \cdot \left\{ \binom{2}{2} \frac{18^2}{21} \cdot \frac{3}{21} \right\} + \left\{ \binom{2}{2} 0.75^2 \cdot 0.25^0 \right\} \cdot \left\{ \binom{2}{2} \frac{20^2}{21} \cdot \frac{1}{21} \right\} + \\
 &\quad + \left\{ \binom{2}{1} 0.75^1 \cdot 0.25^1 \right\} \cdot \left\{ \binom{2}{2} \frac{19^2}{21} \cdot \frac{2}{21} \right\} \\
 &=
 \end{aligned}$$



Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.

3. Si estraggano, con reimmissione, due palline dall'urna U_1 e le si mettano nell'urna U_2 . Si calcoli la probabilità che estraendo con reimmissione 2 palline dall'urna U_2 (modificata) si ottengano 2 palline nere.

Ricordando che la frazione di palline nere nell'urna U_1 è pari a 0.75, in base alla legge delle probabilità totali abbiamo

$$\begin{aligned}
 P(2N) &= P(((2R)_1 \cap 2N) \cup ((2N)_1 \cap 2N) \cup ((1N1R)_1 \cap 2N)) = \\
 &= P((2R)_1 \cap 2N) + P((2N)_1 \cap 2N) + P((1N1R)_1 \cap 2N) = \\
 &= P((2R)_1)P(2N|(2R)_1) + P((2N)_1)P(2N|(2N)_1) + P((1N1R)_1)P(2N|(1N1R)_1) = \\
 &= \left\{ \binom{2}{0} 0.75^0 \cdot 0.25^2 \right\} \cdot \left\{ \binom{2}{2} \frac{18^2}{21} \cdot \frac{3}{21} \right\} + \left\{ \binom{2}{2} 0.75^2 \cdot 0.25^0 \right\} \cdot \left\{ \binom{2}{2} \frac{20^2}{21} \cdot \frac{1}{21} \right\} + \\
 &\quad + \left\{ \binom{2}{1} 0.75^1 \cdot 0.25^1 \right\} \cdot \left\{ \binom{2}{2} \frac{19^2}{21} \cdot \frac{2}{21} \right\} \\
 &= \left\{ \binom{2}{0} 0.75^0 \cdot 0.25^2 \right\} \cdot \frac{18^2}{21} +
 \end{aligned}$$



Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.

3. Si estraggano, con reimmissione, due palline dall'urna U_1 e le si mettano nell'urna U_2 . Si calcoli la probabilità che estraendo con reimmissione 2 palline dall'urna U_2 (modificata) si ottengano 2 palline nere.

Ricordando che la frazione di palline nere nell'urna U_1 è pari a 0.75, in base alla legge delle probabilità totali abbiamo

$$\begin{aligned}
 P(2N) &= P(((2R)_1 \cap 2N) \cup ((2N)_1 \cap 2N) \cup ((1N1R)_1 \cap 2N)) = \\
 &= P((2R)_1 \cap 2N) + P((2N)_1 \cap 2N) + P((1N1R)_1 \cap 2N) = \\
 &= P((2R)_1)P(2N|(2R)_1) + P((2N)_1)P(2N|(2N)_1) + P((1N1R)_1)P(2N|(1N1R)_1) = \\
 &= \left\{ \binom{2}{0} 0.75^0 \cdot 0.25^2 \right\} \cdot \left\{ \binom{2}{2} \frac{18^2}{21} \cdot \frac{3}{21}^0 \right\} + \left\{ \binom{2}{2} 0.75^2 \cdot 0.25^0 \right\} \cdot \left\{ \binom{2}{2} \frac{20^2}{21} \cdot \frac{1}{21}^0 \right\} + \\
 &\quad + \left\{ \binom{2}{1} 0.75^1 \cdot 0.25^1 \right\} \cdot \left\{ \binom{2}{2} \frac{19^2}{21} \cdot \frac{2}{21}^0 \right\} \\
 &= \left\{ \binom{2}{0} 0.75^0 \cdot 0.25^2 \right\} \cdot \frac{18^2}{21} + \left\{ \binom{2}{2} 0.75^2 \cdot 0.25^0 \right\} \cdot \frac{20^2}{21} +
 \end{aligned}$$



Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.

3. Si estraggano, con reimmissione, due palline dall'urna U_1 e le si mettano nell'urna U_2 . Si calcoli la probabilità che estraendo con reimmissione 2 palline dall'urna U_2 (modificata) si ottengano 2 palline nere.

Ricordando che la frazione di palline nere nell'urna U_1 è pari a 0.75, in base alla legge delle probabilità totali abbiamo

$$\begin{aligned}
 P(2N) &= P(((2R)_1 \cap 2N) \cup ((2N)_1 \cap 2N) \cup ((1N1R)_1 \cap 2N)) = \\
 &= P((2R)_1 \cap 2N) + P((2N)_1 \cap 2N) + P((1N1R)_1 \cap 2N) = \\
 &= P((2R)_1)P(2N|(2R)_1) + P((2N)_1)P(2N|(2N)_1) + P((1N1R)_1)P(2N|(1N1R)_1) = \\
 &= \left\{ \binom{2}{0} 0.75^0 \cdot 0.25^2 \right\} \cdot \left\{ \binom{2}{2} \frac{18^2}{21} \cdot \frac{3^0}{21} \right\} + \left\{ \binom{2}{2} 0.75^2 \cdot 0.25^0 \right\} \cdot \left\{ \binom{2}{2} \frac{20^2}{21} \cdot \frac{1^0}{21} \right\} + \\
 &\quad + \left\{ \binom{2}{1} 0.75^1 \cdot 0.25^1 \right\} \cdot \left\{ \binom{2}{2} \frac{19^2}{21} \cdot \frac{2^0}{21} \right\} \\
 &= \left\{ \binom{2}{0} 0.75^0 \cdot 0.25^2 \right\} \cdot \frac{18^2}{21} + \left\{ \binom{2}{2} 0.75^2 \cdot 0.25^0 \right\} \cdot \frac{20^2}{21} + \left\{ \binom{2}{1} 0.75^1 \cdot 0.25^1 \right\} \cdot \frac{19^2}{21} = \\
 &=
 \end{aligned}$$



Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.

3. Si estraggano, con reimmissione, due palline dall'urna U_1 e le si mettano nell'urna U_2 . Si calcoli la probabilità che estraendo con reimmissione 2 palline dall'urna U_2 (modificata) si ottengano 2 palline nere.

Ricordando che la frazione di palline nere nell'urna U_1 è pari a 0.75, in base alla legge delle probabilità totali abbiamo

$$\begin{aligned}
 P(2N) &= P(((2R)_1 \cap 2N) \cup ((2N)_1 \cap 2N) \cup ((1N1R)_1 \cap 2N)) = \\
 &= P((2R)_1 \cap 2N) + P((2N)_1 \cap 2N) + P((1N1R)_1 \cap 2N) = \\
 &= P((2R)_1)P(2N|(2R)_1) + P((2N)_1)P(2N|(2N)_1) + P((1N1R)_1)P(2N|(1N1R)_1) = \\
 &= \left\{ \binom{2}{0} 0.75^0 \cdot 0.25^2 \right\} \cdot \left\{ \binom{2}{2} \frac{18^2}{21} \cdot \frac{3^0}{21} \right\} + \left\{ \binom{2}{2} 0.75^2 \cdot 0.25^0 \right\} \cdot \left\{ \binom{2}{2} \frac{20^2}{21} \cdot \frac{1^0}{21} \right\} + \\
 &\quad + \left\{ \binom{2}{1} 0.75^1 \cdot 0.25^1 \right\} \cdot \left\{ \binom{2}{2} \frac{19^2}{21} \cdot \frac{2^0}{21} \right\} \\
 &= \left\{ \binom{2}{0} 0.75^0 \cdot 0.25^2 \right\} \cdot \frac{18^2}{21} + \left\{ \binom{2}{2} 0.75^2 \cdot 0.25^0 \right\} \cdot \frac{20^2}{21} + \left\{ \binom{2}{1} 0.75^1 \cdot 0.25^1 \right\} \cdot \frac{19^2}{21} = \\
 &= 0.0625 \cdot 0.7347 + 0.5625 \cdot 0.9070 + 0.375 \cdot 0.8186 = \\
 &=
 \end{aligned}$$



Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.

3. Si estraggano, con reimmissione, due palline dall'urna U_1 e le si mettano nell'urna U_2 . Si calcoli la probabilità che estraendo con reimmissione 2 palline dall'urna U_2 (modificata) si ottengano 2 palline nere.

Ricordando che la frazione di palline nere nell'urna U_1 è pari a 0.75, in base alla legge delle probabilità totali abbiamo

$$\begin{aligned}
 P(2N) &= P(((2R)_1 \cap 2N) \cup ((2N)_1 \cap 2N) \cup ((1N1R)_1 \cap 2N)) = \\
 &= P((2R)_1 \cap 2N) + P((2N)_1 \cap 2N) + P((1N1R)_1 \cap 2N) = \\
 &= P((2R)_1)P(2N|(2R)_1) + P((2N)_1)P(2N|(2N)_1) + P((1N1R)_1)P(2N|(1N1R)_1) = \\
 &= \left\{ \binom{2}{0} 0.75^0 \cdot 0.25^2 \right\} \cdot \left\{ \binom{2}{2} \frac{18}{21}^2 \cdot \frac{3}{21}^0 \right\} + \left\{ \binom{2}{2} 0.75^2 \cdot 0.25^0 \right\} \cdot \left\{ \binom{2}{2} \frac{20}{21}^2 \cdot \frac{1}{21}^0 \right\} + \\
 &\quad + \left\{ \binom{2}{1} 0.75^1 \cdot 0.25^1 \right\} \cdot \left\{ \binom{2}{2} \frac{19}{21}^2 \cdot \frac{2}{21}^0 \right\} \\
 &= \left\{ \binom{2}{0} 0.75^0 \cdot 0.25^2 \right\} \cdot \frac{18^2}{21} + \left\{ \binom{2}{2} 0.75^2 \cdot 0.25^0 \right\} \cdot \frac{20^2}{21} + \left\{ \binom{2}{1} 0.75^1 \cdot 0.25^1 \right\} \cdot \frac{19^2}{21} = \\
 &= 0.0625 \cdot 0.7347 + 0.5625 \cdot 0.9070 + 0.375 \cdot 0.8186 = \\
 &= 0.0459 + 0.5102 + 0.3070 = 0.8631
 \end{aligned}$$

Tema T295-1

Prova scritta di STATISTICA del 21.08.2017

Prova scritta di
STATISTICA del
21.08.2017

Giuseppe Boari
Gabriele Cantaluppi



UNIVERSITÀ
CATTOLICA
del Sacro Cuore

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.

Giuseppe Boari
Gabriele Cantaluppi
Università Cattolica del Sacro Cuore, Milano



Il prospetto seguente riporta alcune informazioni sulla distribuzione congiunta del carattere qualitativo X e del carattere quantitativo Y .

$X \setminus Y$	2	3	4	
x_1		4		
x_2		2	2	6
				18

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.



1. Sapendo che la varianza between di Y rispetto ai gruppi definiti dalle categorie della variabile X è nulla, si costruisca la distribuzione delle frequenze congiunte delle due variabili.

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.



Si osserva in primo luogo che $n_{1\bullet} = 12$ e $n_{21} = 2$.

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.



Si osserva in primo luogo che $n_{1\bullet} = 12$ e $n_{21} = 2$.

Vale $M(Y|x_2) = 3$ dal momento che la distribuzione $Y|x_2$ risulta simmetrica.

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.



Si osserva in primo luogo che $n_{1\bullet} = 12$ e $n_{21} = 2$.

Vale $M(Y|x_2) = 3$ dal momento che la distribuzione $Y|x_2$ risulta simmetrica.

Affinché la varianza between di Y rispetto ai gruppi definiti dalle categorie della variabile X sia nulla le due medie di gruppo devono essere identiche, $M(Y|x_1) = M(Y|x_2) = 3$ e, quindi, anche la distribuzione $Y|x_1$ deve essere simmetrica, con $n_{21} = n_{23} =$

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.



Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.

Si osserva in primo luogo che $n_{1\bullet} = 12$ e $n_{21} = 2$.

Vale $M(Y|x_2) = 3$ dal momento che la distribuzione $Y|x_2$ risulta simmetrica.

Affinché la varianza between di Y rispetto ai gruppi definiti dalle categorie della variabile X sia nulla le due medie di gruppo devono essere identiche, $M(Y|x_1) = M(Y|x_2) = 3$ e, quindi, anche la distribuzione $Y|x_1$ deve essere simmetrica, con $n_{21} = n_{23} = 4$.

La distribuzione congiunta risulta, quindi:



Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.

Si osserva in primo luogo che $n_{1\bullet} = 12$ e $n_{21} = 2$.

Vale $M(Y|x_2) = 3$ dal momento che la distribuzione $Y|x_2$ risulta simmetrica.

Affinché la varianza between di Y rispetto ai gruppi definiti dalle categorie della variabile X sia nulla le due medie di gruppo devono essere identiche, $M(Y|x_1) = M(Y|x_2) = 3$ e, quindi, anche la distribuzione $Y|x_1$ deve essere simmetrica, con $n_{21} = n_{23} = 4$.

La distribuzione congiunta risulta, quindi:

$X \setminus Y$	2	3	4	
x_1	4	4	4	12
x_2	2	2	2	6
	6	6	6	18

Le variabili X e Y sono stocasticamente indipendenti in quanto le distribuzioni condizionate $Y|x_1$ e $Y|x_2$ sono somiglianti.



2. Si rappresentino graficamente le distribuzioni marginali e si confrontino le distribuzioni condizionate $Y|X_1$ e $Y|X_2$ con opportuni indici di posizione e di variabilità.

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.



Si considera la distribuzione marginale della variabile X , ottenuta sommando le frequenze di riga

x_i	$n_{i\bullet}$
x_1	12
x_2	6

Si tratta di un carattere qualitativo sconnesso che può essere rappresentato graficamente mediante un grafico a bastoncini o a canne d'organo.

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.

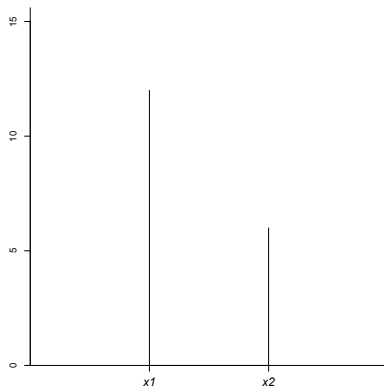


Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.





Si considera la distribuzione marginale della variabile Y , ottenuta sommando le frequenze di colonna

y_j	$n_{\bullet j}$
2	6
3	6
4	6

Esercizio 1.

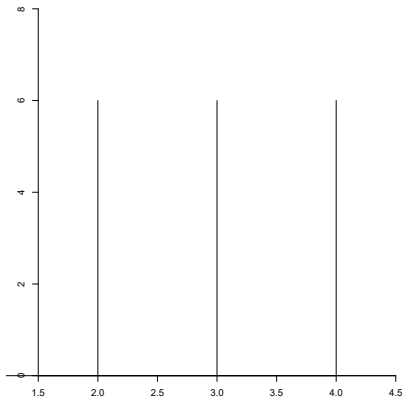
Esercizio 2.

Esercizio 3.

Esercizio 4.



Si tratta di un carattere quantitativo con modalità non raggruppate in classi che può essere rappresentato graficamente mediante un grafico a bastoncini.



Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.



Si osserva che sussiste indipendenza stocastica in quanto le distribuzioni condizionate $Y|x_1$ e $Y|x_2$ sono somiglianti. Le due distribuzioni saranno caratterizzate dallo stesso livello medio e dalla stessa variabilità.

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.



Si osserva che sussiste indipendenza stocastica in quanto le distribuzioni condizionate $Y|x_1$ e $Y|x_2$ sono somiglianti. Le due distribuzioni saranno caratterizzate dallo stesso livello medio e dalla stessa variabilità.

È, quindi, sufficiente procedere al calcolo degli indici per una sola delle due distribuzioni.

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.



Si considera la distribuzione condizionata $Y|x_1$

y_j	n_{1j}
2	4
3	4
4	4

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.



Si tratta di un carattere quantitativo con modalità non raggruppate in classi.

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.



Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.

Si osserva che la distribuzione è simmetrica rispetto a 3; la mediana (percentile di ordine 0.50) risulta $y_{0.5} = 3$.



La media aritmetica

$$M(Y|x_1) = \frac{1}{n_{1\bullet}} \sum_{j=1}^k y_j n_{1j}$$

risulta 3.

Si osserva, infatti, che la distribuzione è simmetrica.

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.



La varianza

$$\text{Var}(Y|x_1) = M(Y^2|x_1) - [M(Y|x_1)]^2$$

risulta 0.6667.

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.



La varianza

$$\text{Var}(Y|x_1) = M(Y^2|x_1) - [M(Y|x_1)]^2$$

risulta 0.6667. Si ottiene

y_j	n_{1j}	$y_j n_{1j}$	$y_j^2 n_{1j}$
2	4	8	16
3	4	12	36
4	4	16	64
	12	36	116

da cui segue

$$\text{Var}(Y|x_1) = \frac{116}{12} - 3^2 = 9.6667 - 9 = 0.6667.$$

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.



Per effettuare confronti in termini di variabilità tra caratteri quantitativi occorre utilizzare una misura relativa, ad esempio il coefficiente di variazione

$$CV = \frac{\sigma}{\mu} = \frac{\sqrt{\text{Var}(Y|x_1)}}{M(Y|x_1)} = \frac{\sqrt{0.6667}}{3} = \frac{0.8165}{3} = 0.2722.$$

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.



Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.

$Y|x_1$ Indice di posizione: $M(Y|x_1) = 3$ Indice di variabilità: $CV(Y|x_1) = 0.2722$

$Y|x_2$ Indice di posizione: $M(Y|x_2) = 3$ Indice di variabilità: $CV(Y|x_2) = 0.2722$

Le distribuzioni condizionate $Y|x_1$ e $Y|x_2$ sono somiglienti e, quindi, caratterizzate da identici valori degli indici di posizione e di variabilità.



Della serie storica X_t sono noti i valori riportati nel seguente prospetto :

t	x_t
2010	20
2011	
2012	24
2013	
2014	

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.



1. Sapendo che nel periodo considerato la grandezza X_t ha subito variazioni assolute costanti, si ricostruiscano, mostrandone il procedimento, la serie storica dei dati x_t e la serie dei numeri indici a base fissa con base 2010.

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.

Dal momento che

$$x_t - x_{t-1} = \text{cost}, \quad t = 2011, \dots, 2014$$

e che

$$x_{2012} - x_{2010} = 24 - 20 = 4,$$

abbiamo $\text{cost} = 2$ e

t	x_t
2010	20
2011	22
2012	24
2013	26
2014	28



Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.

Dal momento che

$$x_t - x_{t-1} = \text{cost}, \quad t = 2011, \dots, 2014$$

e che

$$x_{2012} - x_{2010} = 24 - 20 = 4,$$

abbiamo $\text{cost} = 2$ e

t	x_t
2010	20
2011	22
2012	24
2013	26
2014	28

Si ricorda che ${}_{2010}I_t = \frac{x_t}{x_{2010}}$.



Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.

Dal momento che

$$x_t - x_{t-1} = \text{cost}, \quad t = 2011, \dots, 2014$$

e che

$$x_{2012} - x_{2010} = 24 - 20 = 4,$$

abbiamo $\text{cost} = 2$ e

t	x_t
2010	20
2011	22
2012	24
2013	26
2014	28

Si ricorda che ${}_{2010}I_t = \frac{x_t}{x_{2010}}$. Si ottiene quindi

Anno	x_t	NIBF (2010)
2010	20	1.0
2011	22	1.1
2012	24	1.2
2013	26	1.3
2014	28	1.4



Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.



2. Si calcoli il numero indice al tempo 2013 con base 2011 e si commenti il risultato ottenuto.

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.



2. Si calcoli il numero indice al tempo 2013 con base 2011 e si commenti il risultato ottenuto.

$${}_{2011}I_{2013} = \frac{X_{2013}}{X_{2011}} = \frac{26}{22} = 1.1818$$

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.



2. Si calcoli il numero indice al tempo 2013 con base 2011 e si commenti il risultato ottenuto.

$${}_{2011}I_{2013} = \frac{X_{2013}}{X_{2011}} = \frac{26}{22} = 1.1818$$

Tra il tempo 2011 e il tempo 2013 si è registrato un incremento complessivo del 18.18% della grandezza oggetto di studio.

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.



3. Si calcoli il tasso medio di variazione tra il 2012 e il 2014, interpretando il risultato ottenuto.

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.



3. Si calcoli il tasso medio di variazione tra il 2012 e il 2014, interpretando il risultato ottenuto. Si ricorda che

$${}_{2012}I_{2014} = {}_{2012}I_{2013} \cdot {}_{2013}I_{2014}$$

Si deve trovare il numero indice a base mobile medio.

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.



Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.

3. Si calcoli il tasso medio di variazione tra il 2012 e il 2014, interpretando il risultato ottenuto. Si ricorda che

$${}_{2012}l_{2014} = {}_{2012}l_{2013} \cdot {}_{2013}l_{2014}$$

Si deve trovare il numero indice a base mobile medio. Applicando il criterio di scelta della media secondo Chisini si ottiene

$${}_{2012}l_{2014} = {}_{2012}l_{2013} \cdot {}_{2013}l_{2014} = \alpha \cdot \alpha = \alpha^2.$$



Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.

3. Si calcoli il tasso medio di variazione tra il 2012 e il 2014, interpretando il risultato ottenuto. Si ricorda che

$${}_{2012}l_{2014} = {}_{2012}l_{2013} \cdot {}_{2013}l_{2014}$$

Si deve trovare il numero indice a base mobile medio. Applicando il criterio di scelta della media secondo Chisini si ottiene

$${}_{2012}l_{2014} = {}_{2012}l_{2013} \cdot {}_{2013}l_{2014} = \alpha \cdot \alpha = \alpha^2.$$

Quindi

$$\alpha = \sqrt[2]{{}_{2012}l_{2014}} = \sqrt[2]{{}_{2012}l_{2013} \cdot {}_{2013}l_{2014}} = 1.0801.$$

Da cui segue una variazione relativa media pari a 0.0801, corrispondente a un incremento medio annuo tra il 2012 e il 2014 del 8.01%.



Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.

Il seguente prospetto riporta i valori rilevati in corrispondenza a 5 unità statistiche con riferimento a due caratteri quantitativi X e Y

x_i	2	3	3	3	5
y_i	1	1	2	3	4

1. Si rappresentino graficamente le coppie dei dati e la funzione di regressione ipotizzando Y in funzione di X .



La funzione di regressione associa a ogni valore x appartenente all'insieme di definizione della variabile X la media di Y condizionata a x , $M(Y|x)$

x	$M(Y x)$
2	1
3	$\frac{1}{3}(1 + 2 + 3) = 2$
5	4

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.



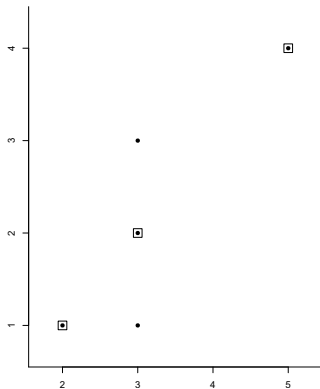
Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.

Le coppie di dati possono essere rappresentate mediante un grafico a dispersione.





2. Si calcolino, secondo il metodo dei minimi quadrati, i parametri dei seguenti modelli di regressione:

$$I) Y^* = a + b \cdot X \qquad II) Y^* = c$$

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.

Si riportano i calcoli necessari¹ per ottenere le stime dei parametri e gli indici di adattamento del modello

$$Y = a + bX + E$$

x_i	y_i	$x_i y_i$	x_i^2	y_i^2
2	1	2	4	1
3	1	3	9	1
3	2	6	9	4
3	3	9	9	9
5	4	20	25	16
16	11	40	56	31
3.2	2.2	8	11.2	6.2
$M(X)$	$M(Y)$	$M(XY)$	$M(X^2)$	$M(Y^2)$

$$\hat{b} =$$



Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.

¹Si osserva che per ragioni di stampa alcuni calcoli potrebbero essere duplicati con riferimento ai diversi modelli oggetto di studio

Si riportano i calcoli necessari¹ per ottenere le stime dei parametri e gli indici di adattamento del modello

$$Y = a + bX + E$$

x_i	y_i	$x_i y_i$	x_i^2	y_i^2
2	1	2	4	1
3	1	3	9	1
3	2	6	9	4
3	3	9	9	9
5	4	20	25	16
16	11	40	56	31
3.2	2.2	8	11.2	6.2
$M(X)$	$M(Y)$	$M(XY)$	$M(X^2)$	$M(Y^2)$

$$\begin{aligned}\hat{b} &= \frac{\text{Cov}(X, Y)}{\text{Var}(X)} = \frac{M(XY) - M(X)M(Y)}{M(X^2) - [M(X)]^2} = \\ &= \frac{8 - 3.2 \cdot 2.2}{11.2 - 3.2^2} = \frac{0.96}{0.96} = 1 \\ \hat{a} &= \end{aligned}$$



Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.

¹Si osserva che per ragioni di stampa alcuni calcoli potrebbero essere duplicati con riferimento ai diversi modelli oggetto di studio

Si riportano i calcoli necessari¹ per ottenere le stime dei parametri e gli indici di adattamento del modello

$$Y = a + bX + E$$

x_i	y_i	$x_i y_i$	x_i^2	y_i^2
2	1	2	4	1
3	1	3	9	1
3	2	6	9	4
3	3	9	9	9
5	4	20	25	16
16	11	40	56	31
3.2	2.2	8	11.2	6.2
$M(X)$	$M(Y)$	$M(XY)$	$M(X^2)$	$M(Y^2)$

$$\hat{b} = \frac{\text{Cov}(X, Y)}{\text{Var}(X)} = \frac{M(XY) - M(X)M(Y)}{M(X^2) - [M(X)]^2} =$$

$$= \frac{8 - 3.2 \cdot 2.2}{11.2 - 3.2^2} = \frac{0.96}{0.96} = 1$$

$$\hat{a} = M(Y) - \hat{b}M(X) = -1$$

$$\rho^2 =$$



Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.

¹Si osserva che per ragioni di stampa alcuni calcoli potrebbero essere duplicati con riferimento ai diversi modelli oggetto di studio

Si riportano i calcoli necessari¹ per ottenere le stime dei parametri e gli indici di adattamento del modello

$$Y = a + bX + E$$

x_i	y_i	$x_i y_i$	x_i^2	y_i^2
2	1	2	4	1
3	1	3	9	1
3	2	6	9	4
3	3	9	9	9
5	4	20	25	16
16	11	40	56	31
3.2	2.2	8	11.2	6.2
$M(X)$	$M(Y)$	$M(XY)$	$M(X^2)$	$M(Y^2)$

$$\hat{b} = \frac{\text{Cov}(X, Y)}{\text{Var}(X)} = \frac{M(XY) - M(X)M(Y)}{M(X^2) - [M(X)]^2} =$$

$$= \frac{8 - 3.2 \cdot 2.2}{11.2 - 3.2^2} = \frac{0.96}{0.96} = 1$$

$$\hat{a} = M(Y) - \hat{b}M(X) = -1$$

$$\rho^2 = \frac{[\text{Cov}(X, Y)]^2}{\text{Var}(X) \text{Var}(Y)} = \frac{0.96^2}{0.96 \cdot 1.36} = 0.7059$$

$$M(\hat{E}^2) = \text{Var}(\hat{E}) = \text{varianza residua} =$$

¹Si osserva che per ragioni di stampa alcuni calcoli potrebbero essere duplicati con riferimento ai diversi modelli oggetto di studio



Si riportano i calcoli necessari¹ per ottenere le stime dei parametri e gli indici di adattamento del modello

$$Y = a + bX + E$$

x_i	y_i	$x_i y_i$	x_i^2	y_i^2
2	1	2	4	1
3	1	3	9	1
3	2	6	9	4
3	3	9	9	9
5	4	20	25	16
16	11	40	56	31
3.2	2.2	8	11.2	6.2
$M(X)$	$M(Y)$	$M(XY)$	$M(X^2)$	$M(Y^2)$

$$\hat{b} = \frac{\text{Cov}(X, Y)}{\text{Var}(X)} = \frac{M(XY) - M(X)M(Y)}{M(X^2) - [M(X)]^2} =$$

$$= \frac{8 - 3.2 \cdot 2.2}{11.2 - 3.2^2} = \frac{0.96}{0.96} = 1$$

$$\hat{a} = M(Y) - \hat{b}M(X) = -1$$

$$\rho^2 = \frac{[\text{Cov}(X, Y)]^2}{\text{Var}(X) \text{Var}(Y)} = \frac{0.96^2}{0.96 \cdot 1.36} = 0.7059$$

$$M(\hat{E}^2) = \text{Var}(\hat{E}) = \text{varianza residua} = \text{Var}(Y) \cdot (1 - \rho^2) = 0.4$$

¹Si osserva che per ragioni di stampa alcuni calcoli potrebbero essere duplicati con riferimento ai diversi modelli oggetto di studio





Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.

Si riportano i calcoli necessari² per ottenere le stime dei parametri e gli indici di adattamento del modello

$$Y = c + E$$

x_i	y_i	y_i^2
2	1	1
3	1	1
3	2	4
3	3	9
5	4	16
11		31
2.2		6.2
$M(Y)$		$M(Y^2)$

$$\hat{c} =$$

²Si osserva che per ragioni di stampa alcuni calcoli potrebbero essere duplicati con riferimento ai diversi modelli oggetto di studio



Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.

Si riportano i calcoli necessari² per ottenere le stime dei parametri e gli indici di adattamento del modello

$$Y = c + E$$

x_i	y_i	y_i^2
2	1	1
3	1	1
3	2	4
3	3	9
5	4	16
	11	31
	2.2	6.2
	$M(Y)$	$M(Y^2)$

$$\hat{c} = M(Y) = 2.2$$

$$R^2 =$$

²Si osserva che per ragioni di stampa alcuni calcoli potrebbero essere duplicati con riferimento ai diversi modelli oggetto di studio



Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.

Si riportano i calcoli necessari² per ottenere le stime dei parametri e gli indici di adattamento del modello

$$Y = c + E$$

x_i	y_i	y_i^2
2	1	1
3	1	1
3	2	4
3	3	9
5	4	16
11		31
2.2		6.2
$M(Y)$		$M(Y^2)$

$$\hat{c} = M(Y) = 2.2$$

$$R^2 = 0$$

$$M(\hat{E}^2) = \text{Var}(\hat{E}) = \text{varianza residua} =$$

²Si osserva che per ragioni di stampa alcuni calcoli potrebbero essere duplicati con riferimento ai diversi modelli oggetto di studio



Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.

Si riportano i calcoli necessari² per ottenere le stime dei parametri e gli indici di adattamento del modello

$$Y = c + E$$

x_i	y_i	y_i^2
2	1	1
3	1	1
3	2	4
3	3	9
5	4	16
11		31
2.2		6.2
$M(Y)$		$M(Y^2)$

$$\hat{c} = M(Y) = 2.2$$

$$R^2 = 0$$

$$M(\hat{E}^2) = \text{Var}(\hat{E}) = \text{varianza residua} = \text{Var}(Y) = 1.36$$

²Si osserva che per ragioni di stampa alcuni calcoli potrebbero essere duplicati con riferimento ai diversi modelli oggetto di studio



Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.

$\hat{a} = -1$, $\hat{b} = 1$, indice adattamento $\rho^2 = 0.7059$ $\hat{c} = 2.2$, indice adattamento $R^2 = 0$



3. Si individui il modello migliore.

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.



3. Si individui il modello migliore.

Il modello migliore è $Y^* = -1 + X$ in quanto caratterizzato da indice di adattamento più elevato.

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.



4. Si indichi, motivando la risposta ma senza fare calcoli, l'intervallo dei valori che può assumere $\eta_{Y|X}^2$.

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.



Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.

4. Si indichi, motivando la risposta ma senza fare calcoli, l'intervallo dei valori che può assumere $\eta_{Y|X}^2$.

$\eta_{Y|X}^2 = \rho^2 = 0.7059$ essendo le medie condizionate allineate.



5. Sulla base del miglior modello si preveda il valore di Y in corrispondenza di $X = 10$.

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.



Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.

5. Sulla base del miglior modello si preveda il valore di Y in corrispondenza di $X = 10$.

$$\hat{Y}(x = 10) = -1 + 10 = 9.$$

Esercizio 4



Si consideri il seguente gioco. Si lancia in primo luogo una moneta. Se esce testa si puntano 2 Euro sul risultato del lancio di un dado e si vincono 3 Euro se esce un numero pari. Se, invece, esce croce si puntano 4 Euro sul risultato del lancio di un dado e si vincono 9 Euro se esce un numero dispari.

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.



Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.

1. Si stabilisca se il gioco è equo.



Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.

1. Si stabilisca se il gioco è equo.

Abbiamo

$$P(T \cap \text{Pari}) = \frac{1}{2} \cdot \frac{3}{6} = 0.25 \quad P(T \cap \text{Dispari}) = \frac{1}{2} \cdot \frac{3}{6} = 0.25$$

$$P(C \cap \text{Pari}) = \frac{1}{2} \cdot \frac{3}{6} = 0.25 \quad P(C \cap \text{Dispari}) = \frac{1}{2} \cdot \frac{3}{6} = 0.25$$

La variabile casuale guadagno



1. Si stabilisca se il gioco è equo.

Abbiamo

$$P(T \cap \text{Pari}) = \frac{1}{2} \cdot \frac{3}{6} = 0.25 \quad P(T \cap \text{Dispari}) = \frac{1}{2} \cdot \frac{3}{6} = 0.25$$

$$P(C \cap \text{Pari}) = \frac{1}{2} \cdot \frac{3}{6} = 0.25 \quad P(C \cap \text{Dispari}) = \frac{1}{2} \cdot \frac{3}{6} = 0.25$$

La variabile casuale guadagno

	x_i	$P(X = x_i)$	$x_i \cdot P(X = x_i)$
<i>posta se T</i>	-2	0.25	-0.5
<i>posta se C</i>	-4	0.25	-1.0
<i>vincita se T ∩ Pari – posta</i>	1	0.25	0.25
<i>vincita se C ∩ Dispari – posta</i>	5	0.25	1.25
		1	0

ha valore atteso



1. Si stabilisca se il gioco è equo.

Abbiamo

$$P(T \cap \text{Pari}) = \frac{1}{2} \cdot \frac{3}{6} = 0.25 \quad P(T \cap \text{Dispari}) = \frac{1}{2} \cdot \frac{3}{6} = 0.25$$

$$P(C \cap \text{Pari}) = \frac{1}{2} \cdot \frac{3}{6} = 0.25 \quad P(C \cap \text{Dispari}) = \frac{1}{2} \cdot \frac{3}{6} = 0.25$$

La variabile casuale guadagno

	x_i	$P(X = x_i)$	$x_i \cdot P(X = x_i)$
<i>posta se T</i>	-2	0.25	-0.5
<i>posta se C</i>	-4	0.25	-1.0
<i>vincita se T ∩ Pari – posta</i>	1	0.25	0.25
<i>vincita se C ∩ Dispari – posta</i>	5	0.25	1.25
		1	0

ha valore atteso nullo e, quindi, il gioco è equo.



Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.

2. Si lanciano 100 volte 2 dadi. Calcolare la probabilità che almeno 2 volte entrambi diano risultato 1.



Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.

2. Si lanciano 100 volte 2 dadi. Calcolare la probabilità che almeno 2 volte entrambi diano risultato 1.

Sia X la variabile casuale 'n. di volte che entrambi i dadi danno risultato 1 nelle 100 prove indipendenti'.

Abbiamo

$$X \sim \text{Bin} \left(n = 100, p = \left(\frac{1}{6} \right)^2 = 0.02778 \right).$$

e

$$P(X \geq 2) =$$



Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.

2. Si lanciano 100 volte 2 dadi. Calcolare la probabilità che almeno 2 volte entrambi diano risultato 1.

Sia X la variabile casuale 'n. di volte che entrambi i dadi danno risultato 1 nelle 100 prove indipendenti'.

Abbiamo

$$X \sim \text{Bin} \left(n = 100, p = \left(\frac{1}{6} \right)^2 = 0.02778 \right).$$

e

$$P(X \geq 2) = 1 - P(X < 2) =$$



Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.

2. Si lanciano 100 volte 2 dadi. Calcolare la probabilità che almeno 2 volte entrambi diano risultato 1.

Sia X la variabile casuale 'n. di volte che entrambi i dadi danno risultato 1 nelle 100 prove indipendenti'.

Abbiamo

$$X \sim \text{Bin} \left(n = 100, p = \left(\frac{1}{6} \right)^2 = 0.02778 \right).$$

e

$$P(X \geq 2) = 1 - P(X < 2) = 1 - \{P(X = 0) + P(X = 1)\} =$$

=



Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.

2. Si lanciano 100 volte 2 dadi. Calcolare la probabilità che almeno 2 volte entrambi diano risultato 1.

Sia X la variabile casuale 'n. di volte che entrambi i dadi danno risultato 1 nelle 100 prove indipendenti'.

Abbiamo

$$X \sim \text{Bin} \left(n = 100, p = \left(\frac{1}{6} \right)^2 = 0.02778 \right).$$

e

$$\begin{aligned} P(X \geq 2) &= 1 - P(X < 2) = 1 - \{P(X = 0) + P(X = 1)\} = \\ &= 1 - \left\{ \binom{100}{0} 0.02778^0 \cdot 0.97222^{100} + \binom{100}{1} 0.02778^1 \cdot 0.97222^{99} \right\} = \\ &= \end{aligned}$$



2. Si lanciano 100 volte 2 dadi. Calcolare la probabilità che almeno 2 volte entrambi diano risultato 1.

Sia X la variabile casuale 'n. di volte che entrambi i dadi danno risultato 1 nelle 100 prove indipendenti'.

Abbiamo

$$X \sim \text{Bin} \left(n = 100, p = \left(\frac{1}{6} \right)^2 = 0.02778 \right).$$

e

$$\begin{aligned} P(X \geq 2) &= 1 - P(X < 2) = 1 - \{P(X = 0) + P(X = 1)\} = \\ &= 1 - \left\{ \binom{100}{0} 0.02778^0 \cdot 0.97222^{100} + \binom{100}{1} 0.02778^1 \cdot 0.97222^{99} \right\} = \\ &= 1 - (0.05978 + 0.1708) = 1 - 0.23058 = 0.76942. \end{aligned}$$



3. Si lanciano 100 volte 2 dadi. Calcolare la probabilità che almeno 17 volte entrambi diano un risultato dispari.

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.



3. Si lanciano 100 volte 2 dadi. Calcolare la probabilità che almeno 17 volte entrambi diano un risultato dispari.

Indicata con X la variabile casuale 'n. di volte che entrambi i dadi danno risultato dispari nelle 100 prove indipendenti' risulta

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.



3. Si lanciano 100 volte 2 dadi. Calcolare la probabilità che almeno 17 volte entrambi diano un risultato dispari.

Indicata con X la variabile casuale 'n. di volte che entrambi i dadi danno risultato dispari nelle 100 prove indipendenti' risulta $X \sim \text{Bin}(n = 100, p = (\frac{1}{2})^2 = 0.25)$. $M(X) = 25$;
 $\text{Var}(X) = 18.75$.

[Esercizio 1.](#)

[Esercizio 2.](#)

[Esercizio 3.](#)

[Esercizio 4.](#)



3. Si lanciano 100 volte 2 dadi. Calcolare la probabilità che almeno 17 volte entrambi diano un risultato dispari.

Indicata con X la variabile casuale 'n. di volte che entrambi i dadi danno risultato dispari nelle 100 prove indipendenti' risulta $X \sim \text{Bin}(n = 100, p = (\frac{1}{2})^2 = 0.25)$. $M(X) = 25$;
 $\text{Var}(X) = 18.75$.

Dal momento che $np = 25 > 5$ ed $n(1 - p) = 75 > 5$ è possibile utilizzare l'approssimazione della variabile casuale Binomiale con la variabile casuale Normale
 $X \sim N(\mu = 25, \sigma^2 = 18.75)$.



3. Si lanciano 100 volte 2 dadi. Calcolare la probabilità che almeno 17 volte entrambi diano un risultato dispari.

Indicata con X la variabile casuale 'n. di volte che entrambi i dadi danno risultato dispari nelle 100 prove indipendenti' risulta $X \sim \text{Bin}(n = 100, p = (\frac{1}{2})^2 = 0.25)$. $M(X) = 25$; $\text{Var}(X) = 18.75$.

Dal momento che $np = 25 > 5$ ed $n(1 - p) = 75 > 5$ è possibile utilizzare l'approssimazione della variabile casuale Binomiale con la variabile casuale Normale $X \sim N(\mu = 25, \sigma^2 = 18.75)$. Si ottiene

$$P(X \geq 17) =$$



3. Si lanciano 100 volte 2 dadi. Calcolare la probabilità che almeno 17 volte entrambi diano un risultato dispari.

Indicata con X la variabile casuale 'n. di volte che entrambi i dadi danno risultato dispari nelle 100 prove indipendenti' risulta $X \sim \text{Bin}(n = 100, p = (\frac{1}{2})^2 = 0.25)$. $M(X) = 25$; $\text{Var}(X) = 18.75$.

Dal momento che $np = 25 > 5$ ed $n(1 - p) = 75 > 5$ è possibile utilizzare l'approssimazione della variabile casuale Binomiale con la variabile casuale Normale $X \sim N(\mu = 25, \sigma^2 = 18.75)$. Si ottiene

$$P(X \geq 17) = P\left(\frac{X - \mu}{\sigma} \geq \frac{17 - \mu}{\sigma}\right) =$$



3. Si lanciano 100 volte 2 dadi. Calcolare la probabilità che almeno 17 volte entrambi diano un risultato dispari.

Indicata con X la variabile casuale 'n. di volte che entrambi i dadi danno risultato dispari nelle 100 prove indipendenti' risulta $X \sim \text{Bin}(n = 100, p = (\frac{1}{2})^2 = 0.25)$. $M(X) = 25$; $\text{Var}(X) = 18.75$.

Dal momento che $np = 25 > 5$ ed $n(1 - p) = 75 > 5$ è possibile utilizzare l'approssimazione della variabile casuale Binomiale con la variabile casuale Normale $X \sim N(\mu = 25, \sigma^2 = 18.75)$. Si ottiene

$$P(X \geq 17) = P\left(\frac{X - \mu}{\sigma} \geq \frac{17 - \mu}{\sigma}\right) = P\left(Z \geq \frac{17 - 25}{4.3301}\right) =$$

=



3. Si lanciano 100 volte 2 dadi. Calcolare la probabilità che almeno 17 volte entrambi diano un risultato dispari.

Indicata con X la variabile casuale 'n. di volte che entrambi i dadi danno risultato dispari nelle 100 prove indipendenti' risulta $X \sim \text{Bin}(n = 100, p = (\frac{1}{2})^2 = 0.25)$. $M(X) = 25$; $\text{Var}(X) = 18.75$.

Dal momento che $np = 25 > 5$ ed $n(1 - p) = 75 > 5$ è possibile utilizzare l'approssimazione della variabile casuale Binomiale con la variabile casuale Normale $X \sim N(\mu = 25, \sigma^2 = 18.75)$. Si ottiene

$$\begin{aligned} P(X \geq 17) &= P\left(\frac{X - \mu}{\sigma} \geq \frac{17 - \mu}{\sigma}\right) = P\left(Z \geq \frac{17 - 25}{4.3301}\right) = \\ &= P(Z \geq -1.85) = \end{aligned}$$



3. Si lanciano 100 volte 2 dadi. Calcolare la probabilità che almeno 17 volte entrambi diano un risultato dispari.

Indicata con X la variabile casuale 'n. di volte che entrambi i dadi danno risultato dispari nelle 100 prove indipendenti' risulta $X \sim \text{Bin}(n = 100, p = (\frac{1}{2})^2 = 0.25)$. $M(X) = 25$; $\text{Var}(X) = 18.75$.

Dal momento che $np = 25 > 5$ ed $n(1 - p) = 75 > 5$ è possibile utilizzare l'approssimazione della variabile casuale Binomiale con la variabile casuale Normale $X \sim N(\mu = 25, \sigma^2 = 18.75)$. Si ottiene

$$\begin{aligned} P(X \geq 17) &= P\left(\frac{X - \mu}{\sigma} \geq \frac{17 - \mu}{\sigma}\right) = P\left(Z \geq \frac{17 - 25}{4.3301}\right) = \\ &= P(Z \geq -1.85) = P(Z \leq 1.85) = \end{aligned}$$



3. Si lanciano 100 volte 2 dadi. Calcolare la probabilità che almeno 17 volte entrambi diano un risultato dispari.

Indicata con X la variabile casuale 'n. di volte che entrambi i dadi danno risultato dispari nelle 100 prove indipendenti' risulta $X \sim \text{Bin}(n = 100, p = (\frac{1}{2})^2 = 0.25)$. $M(X) = 25$; $\text{Var}(X) = 18.75$.

Dal momento che $np = 25 > 5$ ed $n(1 - p) = 75 > 5$ è possibile utilizzare l'approssimazione della variabile casuale Binomiale con la variabile casuale Normale $X \sim N(\mu = 25, \sigma^2 = 18.75)$. Si ottiene

$$\begin{aligned} P(X \geq 17) &= P\left(\frac{X - \mu}{\sigma} \geq \frac{17 - \mu}{\sigma}\right) = P\left(Z \geq \frac{17 - 25}{4.3301}\right) = \\ &= P(Z \geq -1.85) = P(Z \leq 1.85) = 0.9678. \end{aligned}$$

Tema T295

Prova scritta di STATISTICA del 07.09.2017

Prova scritta di
STATISTICA del
07.09.2017

Giuseppe Boari
Gabriele Cantaluppi



UNIVERSITÀ
CATTOLICA
del Sacro Cuore

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.

Giuseppe Boari
Gabriele Cantaluppi
Università Cattolica del Sacro Cuore, Milano



Il prospetto seguente riporta alcune informazioni sulla distribuzione congiunta del carattere qualitativo X e del carattere quantitativo Y .

$X \setminus Y$	2	3	4	
x_1	0		0	
x_2		2		6

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.



1. Sapendo che la distribuzione marginale di X è caratterizzata da massima eterogeneità e che $M(Y|x_1) = M(Y|x_2)$ si costruisca la distribuzione delle frequenze congiunte.

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.



Dal momento che la distribuzione marginale di X è caratterizzata da massima eterogeneità, abbiamo una situazione di equidistribuzione delle frequenze, quindi $n_{1\bullet} =$

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.



Dal momento che la distribuzione marginale di X è caratterizzata da massima eterogeneità, abbiamo una situazione di equidistribuzione delle frequenze, quindi $n_{1\bullet} = n_{2\bullet} = 6$.

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.



Dal momento che la distribuzione marginale di X è caratterizzata da massima eterogeneità, abbiamo una situazione di equidistribuzione delle frequenze, quindi $n_{1\bullet} = n_{2\bullet} = 6$.

La distribuzione condizionata $Y|x_1$ è degenere dal momento che solo $n_{22} = 6 \neq 0$, e quindi $M(Y|x_1) = 3$.

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.



Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.

Dal momento che la distribuzione marginale di X è caratterizzata da massima eterogeneità, abbiamo una situazione di equidistribuzione delle frequenze, quindi $n_{1\bullet} = n_{2\bullet} = 6$.

La distribuzione condizionata $Y|x_1$ è degenera dal momento che solo $n_{22} = 6 \neq 0$, e quindi $M(Y|x_1) = 3$.

Affinché anche $M(Y|x_2) = 3$ la distribuzione $Y|x_2$ deve essere simmetrica, e quindi $n_{21} =$



Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.

Dal momento che la distribuzione marginale di X è caratterizzata da massima eterogeneità, abbiamo una situazione di equidistribuzione delle frequenze, quindi $n_{1\bullet} = n_{2\bullet} = 6$.

La distribuzione condizionata $Y|x_1$ è degenera dal momento che solo $n_{22} = 6 \neq 0$, e quindi $M(Y|x_1) = 3$.

Affinché anche $M(Y|x_2) = 3$ la distribuzione $Y|x_2$ deve essere simmetrica, e quindi $n_{21} = n_{23} = 2$

La distribuzione congiunta risulta, quindi:



Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.

Dal momento che la distribuzione marginale di X è caratterizzata da massima eterogeneità, abbiamo una situazione di equidistribuzione delle frequenze, quindi $n_{1\bullet} = n_{2\bullet} = 6$.

La distribuzione condizionata $Y|X_1$ è degenera dal momento che solo $n_{22} = 6 \neq 0$, e quindi $M(Y|X_1) = 3$.

Affinché anche $M(Y|X_2) = 3$ la distribuzione $Y|X_2$ deve essere simmetrica, e quindi $n_{21} = n_{23} = 2$

La distribuzione congiunta risulta, quindi:

$X \setminus Y$	2	3	4	
x_1	0	6	0	6
x_2	2	2	2	6
	2	8	2	12



2. Si calcolino la media e la varianza della Y .

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.



Si considera la distribuzione marginale della variabile Y , ottenuta sommando le frequenze di colonna

y_j	$n_{\bullet j}$
2	2
3	8
4	2

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.



Si tratta di un carattere quantitativo con modalità non raggruppate in classi.

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.



La media aritmetica

$$M(Y) = \frac{1}{n} \sum_{j=1}^k y_j n_{\bullet j}$$

risulta 3.

Si osserva, infatti, che la distribuzione è simmetrica.

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.



La varianza

$$\text{Var}(Y) = M[(Y - \mu_Y)^2] = M(Y^2) - \mu_Y^2 = \frac{1}{n} \sum_{j=1}^k y_j^2 n_{\bullet j} - \mu_Y^2$$

risulta 0.3333.

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.



La varianza

$$Var(Y) = M[(Y - \mu_Y)^2] = M(Y^2) - \mu_Y^2 = \frac{1}{n} \sum_{j=1}^k y_j^2 n_{\bullet j} - \mu_Y^2$$

risulta 0.3333.

Si ottiene

y_j	$n_{\bullet j}$	$y_j n_{\bullet j}$	$y_j^2 n_{\bullet j}$
2	2	4	8
3	8	24	72
4	2	8	32
	12	36	112

da cui segue

$$\begin{aligned} Var(Y) = \sigma_Y^2 &= M(Y^2) - \mu_Y^2 = \\ &= \frac{112}{12} - 3^2 = 9.3333 - 9 = 0.3333. \end{aligned}$$

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.



3. Considerati i due gruppi di osservazioni relativi alle condizionate $Y|x_1$ e $Y|x_2$ si indichi il valore della varianza within.

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.



Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.

3. Considerati i due gruppi di osservazioni relativi alle condizionate $Y|x_1$ e $Y|x_2$ si indichi il valore della varianza within.

Dal momento che $M(Y|x_1) = M(Y|x_2)$ (le medie di gruppo sono identiche) abbiamo

$$\sigma_{Between}^2 = 0 \text{ e } \sigma_{Within}^2 = \sigma_Y^2.$$



4. Senza effettuare calcoli, indicare se, e perché, le marginali X e Y sono indipendenti.

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.



4. Senza effettuare calcoli, indicare se, e perché, le marginali X e Y sono indipendenti. Le variabili X e Y non sono stocasticamente indipendenti in quanto le distribuzioni condizionate $Y|x_1$ e $Y|x_2$ non sono somiglianti.

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.



Della serie storica X_t sono noti i valori riportati nel seguente prospetto:

t	x_t	$NIBM$
2010	20	
2011		
2012	24.2	
2013		
2014		

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.



1. Sapendo che nel periodo considerato la grandezza X_t ha subito variazioni relative costanti, si ricostruiscano, mostrandone il procedimento, la serie storica dei dati x_t e la serie dei numeri indici a base mobile.

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.



1. Sapendo che nel periodo considerato la grandezza X_t ha subito variazioni relative costanti, si ricostruiscano, mostrandone il procedimento, la serie storica dei dati x_t e la serie dei numeri indici a base mobile.

Dal momento che le variazioni relative $\frac{x_t - x_{t-1}}{x_{t-1}} = cost, t = 2011, \dots, 2014$, abbiamo

$$\frac{x_t}{x_{t-1}} = 1 + cost, t = 2011, \dots, 2014.$$

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.



Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.

1. Sapendo che nel periodo considerato la grandezza X_t ha subito variazioni relative costanti, si ricostruiscano, mostrandone il procedimento, la serie storica dei dati x_t e la serie dei numeri indici a base mobile.

Dal momento che le variazioni relative $\frac{x_t - x_{t-1}}{x_{t-1}} = cost$, $t = 2011, \dots, 2014$, abbiamo

$$\frac{x_t}{x_{t-1}} = 1 + cost, t = 2011, \dots, 2014.$$

Essendo $\frac{x_{2012}}{x_{2010}} = \frac{x_{2011}}{x_{2010}} \cdot \frac{x_{2012}}{x_{2011}} = (1 + cost)^2 = 1.21$, abbiamo $(1 + cost) = 1.1$ e $cost = 0.1$.



Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.

1. Sapendo che nel periodo considerato la grandezza X_t ha subito variazioni relative costanti, si ricostruiscano, mostrandone il procedimento, la serie storica dei dati x_t e la serie dei numeri indici a base mobile.

Dal momento che le variazioni relative $\frac{x_t - x_{t-1}}{x_{t-1}} = cost$, $t = 2011, \dots, 2014$, abbiamo

$$\frac{x_t}{x_{t-1}} = 1 + cost, \quad t = 2011, \dots, 2014.$$

Essendo $\frac{x_{2012}}{x_{2010}} = \frac{x_{2011}}{x_{2010}} \cdot \frac{x_{2012}}{x_{2011}} = (1 + cost)^2 = 1.21$, abbiamo $(1 + cost) = 1.1$ e $cost = 0.1$.

Nel periodo considerato la grandezza oggetto di studio ha subito un incremento del 10% tra un anno e il successivo e, quindi

- $x_t = 1.1 \cdot x_{t-1}, \quad t = 2011, \dots, 2014.$
- $\frac{x_t}{x_{t-1}} = 1.1, \quad t = 2011, \dots, 2014.$



La serie storica dei dati x_t e la serie dei numeri indici a base mobile risultano

t	x_t	$NIBM$
2010	20	
2011	22	1.1
2012	24.2	1.1
2013	26.62	1.1
2014	29.282	1.1

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.



2. Si calcoli il numero indice al tempo 2013 con base 2011 e si commenti il risultato ottenuto.

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.



2. Si calcoli il numero indice al tempo 2013 con base 2011 e si commenti il risultato ottenuto.

$${}_{2011}I_{2013} = \frac{x_{2013}}{x_{2011}} = \frac{26.62}{22} = 1.21$$

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.



2. Si calcoli il numero indice al tempo 2013 con base 2011 e si commenti il risultato ottenuto.

$${}_{2011}I_{2013} = \frac{x_{2013}}{x_{2011}} = \frac{26.62}{22} = 1.21$$

Tra il tempo 2011 e il tempo 2013 si è registrato un incremento complessivo del 21% della grandezza oggetto di studio.

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.



3. Si calcoli il tasso medio di variazione tra il 2012 e il 2014, interpretando il risultato ottenuto.

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.



Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.

3. Si calcoli il tasso medio di variazione tra il 2012 e il 2014, interpretando il risultato ottenuto. Dal momento che la serie ha subito variazioni relative costanti nel periodo, abbiamo che la variazione relativa media coincide con $cost = 0.1$. Tra ogni anno e il successivo la serie ha subito un incremento del 10%.



Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.

Il seguente prospetto riporta i valori rilevati in corrispondenza a 8 unità statistiche con riferimento a due caratteri quantitativi X e Y

x_i	1	2	2	2	4
y_i	2	2	3	4	4

1. Si rappresentino graficamente le coppie dei dati e la funzione di regressione ipotizzando Y in funzione di X .



La funzione di regressione associa a ogni valore x appartenente all'insieme di definizione della variabile X la media di Y condizionata a x , $M(Y|x)$

x	$M(Y x)$
1	2
2	$\frac{1}{3}(2 + 3 + 4) = 3$
4	4

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.



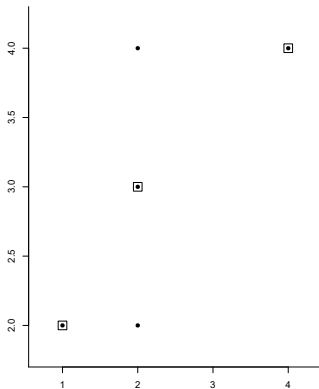
Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.

Le coppie di dati possono essere rappresentate mediante un grafico a dispersione.





Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.

2. Si calcolino, secondo il metodo dei minimi quadrati, i parametri dei seguenti modelli di regressione:

$$I) Y^* = a + b \cdot X \qquad II) Y^* = c + d \cdot \sqrt{X}$$



Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.

Si riportano i calcoli necessari¹ per ottenere le stime dei parametri e gli indici di adattamento del modello

$$Y = a + bX + E$$

x_i	y_i	$x_i y_i$	x_i^2	y_i^2
1	2	2	1	4
2	2	4	4	4
2	3	6	4	9
2	4	8	4	16
4	4	16	16	16
11	15	36	29	49
2.2	3	7.2	5.8	9.8
$M(X)$	$M(Y)$	$M(XY)$	$M(X^2)$	$M(Y^2)$

$$\hat{b} =$$

¹Si osserva che per ragioni di stampa alcuni calcoli potrebbero essere duplicati con riferimento ai diversi modelli oggetto di studio



Si riportano i calcoli necessari¹ per ottenere le stime dei parametri e gli indici di adattamento del modello

$$Y = a + bX + E$$

x_i	y_i	$x_i y_i$	x_i^2	y_i^2
1	2	2	1	4
2	2	4	4	4
2	3	6	4	9
2	4	8	4	16
4	4	16	16	16
11	15	36	29	49
2.2	3	7.2	5.8	9.8
$M(X)$	$M(Y)$	$M(XY)$	$M(X^2)$	$M(Y^2)$

$$\begin{aligned}\hat{b} &= \frac{\text{Cov}(X, Y)}{\text{Var}(X)} = \frac{M(XY) - M(X)M(Y)}{M(X^2) - [M(X)]^2} = \\ &= \frac{7.2 - 2.2 \cdot 3}{5.8 - 2.2^2} = \frac{0.6}{0.96} = 0.625 \\ \hat{a} &= \end{aligned}$$

¹Si osserva che per ragioni di stampa alcuni calcoli potrebbero essere duplicati con riferimento ai diversi modelli oggetto di studio



Si riportano i calcoli necessari¹ per ottenere le stime dei parametri e gli indici di adattamento del modello

$$Y = a + bX + E$$

x_i	y_i	$x_i y_i$	x_i^2	y_i^2
1	2	2	1	4
2	2	4	4	4
2	3	6	4	9
2	4	8	4	16
4	4	16	16	16
11	15	36	29	49
2.2	3	7.2	5.8	9.8
$M(X)$	$M(Y)$	$M(XY)$	$M(X^2)$	$M(Y^2)$

$$\begin{aligned}\hat{b} &= \frac{\text{Cov}(X, Y)}{\text{Var}(X)} = \frac{M(XY) - M(X)M(Y)}{M(X^2) - [M(X)]^2} = \\ &= \frac{7.2 - 2.2 \cdot 3}{5.8 - 2.2^2} = \frac{0.6}{0.96} = 0.625 \\ \hat{a} &= M(Y) - \hat{b}M(X) = 1.625 \\ \rho^2 &= \end{aligned}$$

¹Si osserva che per ragioni di stampa alcuni calcoli potrebbero essere duplicati con riferimento ai diversi modelli oggetto di studio



Si riportano i calcoli necessari¹ per ottenere le stime dei parametri e gli indici di adattamento del modello

$$Y = a + bX + E$$

x_i	y_i	$x_i y_i$	x_i^2	y_i^2
1	2	2	1	4
2	2	4	4	4
2	3	6	4	9
2	4	8	4	16
4	4	16	16	16
11	15	36	29	49
2.2	3	7.2	5.8	9.8
$M(X)$	$M(Y)$	$M(XY)$	$M(X^2)$	$M(Y^2)$

$$\begin{aligned}\hat{b} &= \frac{\text{Cov}(X, Y)}{\text{Var}(X)} = \frac{M(XY) - M(X)M(Y)}{M(X^2) - [M(X)]^2} = \\ &= \frac{7.2 - 2.2 \cdot 3}{5.8 - 2.2^2} = \frac{0.6}{0.96} = 0.625 \\ \hat{a} &= M(Y) - \hat{b}M(X) = 1.625 \\ \rho^2 &= \frac{[\text{Cov}(X, Y)]^2}{\text{Var}(X)\text{Var}(Y)} = \frac{0.6^2}{0.96 \cdot 0.8} = 0.4687\end{aligned}$$

¹Si osserva che per ragioni di stampa alcuni calcoli potrebbero essere duplicati con riferimento ai diversi modelli oggetto di studio

Si riportano i calcoli necessari² per ottenere le stime dei parametri e gli indici di adattamento del modello

$$Y = a + b\sqrt{X} + E$$

$\sqrt{x_i}$	y_i	$\sqrt{x_i}y_i$	$\sqrt{x_i}^2$	y_i^2
1	2	2	1	4
1.4142	2	2.8284	2	4
1.4142	3	4.2426	2	9
1.4142	4	5.6569	2	16
2	4	8	4	16
7.2426	15	22.7279	11	49
1.4485	3	4.5456	2.2	9.8
$M(\sqrt{X})$	$M(Y)$	$M(\sqrt{X}Y)$	$M(\sqrt{X}^2)$	$M(Y^2)$

$$\hat{b} =$$



Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.

²Si osserva che per ragioni di stampa alcuni calcoli potrebbero essere duplicati con riferimento ai diversi modelli oggetto di studio

Si riportano i calcoli necessari² per ottenere le stime dei parametri e gli indici di adattamento del modello

$$Y = a + b\sqrt{X} + E$$

$\sqrt{x_i}$	y_i	$\sqrt{x_i}y_i$	$\sqrt{x_i}^2$	y_i^2
1	2	2	1	4
1.4142	2	2.8284	2	4
1.4142	3	4.2426	2	9
1.4142	4	5.6569	2	16
2	4	8	4	16
7.2426	15	22.7279	11	49
1.4485	3	4.5456	2.2	9.8
$M(\sqrt{X})$	$M(Y)$	$M(\sqrt{X}Y)$	$M(\sqrt{X}^2)$	$M(Y^2)$

$$\begin{aligned}\hat{b} &= \frac{\text{Cov}(\sqrt{X}, Y)}{\text{Var}(\sqrt{X})} = \frac{M(\sqrt{X}Y) - M(\sqrt{X})M(Y)}{M(\sqrt{X}^2) - [M(\sqrt{X})]^2} = \\ &= \frac{4.5456 - 1.4485 \cdot 3}{2.2 - 1.4485^2} = \frac{0.2}{0.1018} = 1.9646\end{aligned}$$

$$\hat{a} =$$



²Si osserva che per ragioni di stampa alcuni calcoli potrebbero essere duplicati con riferimento ai diversi modelli oggetto di studio

Si riportano i calcoli necessari² per ottenere le stime dei parametri e gli indici di adattamento del modello

$$Y = a + b\sqrt{X} + E$$

$\sqrt{x_i}$	y_i	$\sqrt{x_i}y_i$	$\sqrt{x_i}^2$	y_i^2
1	2	2	1	4
1.4142	2	2.8284	2	4
1.4142	3	4.2426	2	9
1.4142	4	5.6569	2	16
2	4	8	4	16
7.2426	15	22.7279	11	49
1.4485	3	4.5456	2.2	9.8
$M(\sqrt{X})$	$M(Y)$	$M(\sqrt{X}Y)$	$M(\sqrt{X}^2)$	$M(Y^2)$

$$\hat{b} = \frac{\text{Cov}(\sqrt{X}, Y)}{\text{Var}(\sqrt{X})} = \frac{M(\sqrt{X}Y) - M(\sqrt{X})M(Y)}{M(\sqrt{X}^2) - [M(\sqrt{X})]^2} =$$

$$= \frac{4.5456 - 1.4485 \cdot 3}{2.2 - 1.4485^2} = \frac{0.2}{0.1018} = 1.9646$$

$$\hat{a} = M(Y) - \hat{b}M(\sqrt{X}) = 0.1542$$

$$\rho^2 =$$



²Si osserva che per ragioni di stampa alcuni calcoli potrebbero essere duplicati con riferimento ai diversi modelli oggetto di studio

Si riportano i calcoli necessari² per ottenere le stime dei parametri e gli indici di adattamento del modello

$$Y = a + b\sqrt{X} + E$$

$\sqrt{x_i}$	y_i	$\sqrt{x_i}y_i$	$\sqrt{x_i}^2$	y_i^2
1	2	2	1	4
1.4142	2	2.8284	2	4
1.4142	3	4.2426	2	9
1.4142	4	5.6569	2	16
2	4	8	4	16
7.2426	15	22.7279	11	49
1.4485	3	4.5456	2.2	9.8
$M(\sqrt{X})$	$M(Y)$	$M(\sqrt{X}Y)$	$M(\sqrt{X}^2)$	$M(Y^2)$

$$\begin{aligned}\hat{b} &= \frac{\text{Cov}(\sqrt{X}, Y)}{\text{Var}(\sqrt{X})} = \frac{M(\sqrt{X}Y) - M(\sqrt{X})M(Y)}{M(\sqrt{X}^2) - [M(\sqrt{X})]^2} = \\ &= \frac{4.5456 - 1.4485 \cdot 3}{2.2 - 1.4485^2} = \frac{0.2}{0.1018} = 1.9646 \\ \hat{a} &= M(Y) - \hat{b}M(\sqrt{X}) = 0.1542 \\ \rho^2 &= \frac{[\text{Cov}(\sqrt{X}, Y)]^2}{\text{Var}(\sqrt{X}) \text{Var}(Y)} = \frac{0.2^2}{0.1018 \cdot 0.8} = 0.4912\end{aligned}$$

²Si osserva che per ragioni di stampa alcuni calcoli potrebbero essere duplicati con riferimento ai diversi modelli oggetto di studio





Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.

$\hat{a} = 1.625$, $\hat{b} = 0.625$, indice adattamento $\rho^2 = 0.4688$ $\hat{c} = 0.1532$, $\hat{d} = 1.9653$, indice adattamento $R^2 = 0.4913$



3. Si individui il modello migliore.

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.



3. Si individui il modello migliore.

Il modello migliore è $Y^* = 0.1532 + 1.9653 \cdot \sqrt{X}$ in quanto caratterizzato da indice di adattamento più elevato.

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.



4. Si indichi, **motivando** la risposta ma senza fare calcoli, quale relazione è vera:

- $\eta_{X|Y}^2 > \rho^2$
- $\eta_{X|Y}^2 = \rho^2$
- $\eta_{X|Y}^2 < \rho^2$

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.



Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.

4. Si indichi, **motivando** la risposta ma senza fare calcoli, quale relazione è vera:

- $\eta_{X|Y}^2 > \rho^2$
- $\eta_{X|Y}^2 = \rho^2$
- $\eta_{X|Y}^2 < \rho^2$

Dal momento che le medie condizionate non sono allineate abbiamo che vale

$$\eta_{X|Y}^2 > \rho^2 = 0.4688$$

Esercizio 4

Prova scritta di
STATISTICA del
07.09.2017

Giuseppe Boari
Gabriele Cantaluppi



UNIVERSITÀ
CATTOLICA
del Sacro Cuore

Da un mazzo di carte da bridge (52 carte) vengono tolte le 13 carte di picche.

1. Considerando l'esperimento di estrazione casuale, con reimmissione, di 3 carte costruire la variabile casuale $X = \text{'n. di carte con seme di colore nero'}$ (valori e probabilità).

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.



Da un mazzo di carte da bridge (52 carte) vengono tolte le 13 carte di picche.

1. Considerando l'esperimento di estrazione casuale, con reimmissione, di 3 carte costruire la variabile casuale $X =$ 'n. di carte con seme di colore nero' (valori e probabilità).

Una volta tolte le 13 carte di picche, il mazzo di carte da bridge risulta composto da 39 carte, 13 cuori, 13 quadri e 13 fiori, e quindi da 26 carte rosse (R) e 13 carte nere (N).

La variabile casuale $X =$ 'n. di carte con seme di colore nero' tra le 3 carte estratte con reimmissione risulta una variabile casuale Binomiale $X \sim \text{Bin}(n = 3, p = \frac{13}{39} = \frac{1}{3})$ con funzione di probabilità

$$P(X = x) = \binom{3}{x} \left(\frac{1}{3}\right)^x \left(\frac{2}{3}\right)^{3-x}, \quad x = 0, 1, 2, 3.$$



2. Calcolare $M(X)$.

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.



Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.

2. Calcolare $M(X)$.

La media di una variabile casuale $X \sim \text{Bin}(n, p)$ risulta $M(X) = n \cdot p$; nel caso in esame abbiamo $M(X) = 1$.



3. Considerando ora $n = 30$ successive estrazioni (sempre con reimmissione) calcolare la probabilità che almeno 15 siano di colore nero.

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.



3. Considerando ora $n = 30$ successive estrazioni (sempre con reimmissione) calcolare la probabilità che almeno 15 siano di colore nero.

Indicata con X la variabile casuale 'n. di carte con seme nero tra le 30 estratte con reimmissione' (prove indipendenti) risulta

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.



Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.

3. Considerando ora $n = 30$ successive estrazioni (sempre con reimmissione) calcolare la probabilità che almeno 15 siano di colore nero.

Indicata con X la variabile casuale 'n. di carte con seme nero tra le 30 estratte con reimmissione' (prove indipendenti) risulta $X \sim \text{Bin}(n = 30, p = \frac{1}{3})$. $M(X) = 10$; $\text{Var}(X) = 6.6667$.



Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.

3. Considerando ora $n = 30$ successive estrazioni (sempre con reimmissione) calcolare la probabilità che almeno 15 siano di colore nero.

Indicata con X la variabile casuale 'n. di carte con seme nero tra le 30 estratte con reimmissione' (prove indipendenti) risulta $X \sim \text{Bin}(n = 30, p = \frac{1}{3})$. $M(X) = 10$; $\text{Var}(X) = 6.6667$.

Dal momento che $np = 10 > 5$ ed $n(1 - p) = 20 > 5$ è possibile utilizzare l'approssimazione della variabile casuale Binomiale con la variabile casuale Normale $X \sim N(\mu = 10, \sigma^2 = 6.6667)$.



3. Considerando ora $n = 30$ successive estrazioni (sempre con reimmissione) calcolare la probabilità che almeno 15 siano di colore nero.

Indicata con X la variabile casuale 'n. di carte con seme nero tra le 30 estratte con reimmissione' (prove indipendenti) risulta $X \sim \text{Bin}(n = 30, p = \frac{1}{3})$. $M(X) = 10$; $\text{Var}(X) = 6.6667$.

Dal momento che $np = 10 > 5$ ed $n(1 - p) = 20 > 5$ è possibile utilizzare l'approssimazione della variabile casuale Binomiale con la variabile casuale Normale $X \sim N(\mu = 10, \sigma^2 = 6.6667)$. Si ottiene

$$P(X \geq 15) =$$



3. Considerando ora $n = 30$ successive estrazioni (sempre con reimmissione) calcolare la probabilità che almeno 15 siano di colore nero.

Indicata con X la variabile casuale 'n. di carte con seme nero tra le 30 estratte con reimmissione' (prove indipendenti) risulta $X \sim \text{Bin}(n = 30, p = \frac{1}{3})$. $M(X) = 10$; $\text{Var}(X) = 6.6667$.

Dal momento che $np = 10 > 5$ ed $n(1 - p) = 20 > 5$ è possibile utilizzare l'approssimazione della variabile casuale Binomiale con la variabile casuale Normale $X \sim N(\mu = 10, \sigma^2 = 6.6667)$. Si ottiene

$$P(X \geq 15) = P\left(\frac{X - \mu}{\sigma} \geq \frac{15 - \mu}{\sigma}\right) =$$



3. Considerando ora $n = 30$ successive estrazioni (sempre con reimmissione) calcolare la probabilità che almeno 15 siano di colore nero.

Indicata con X la variabile casuale 'n. di carte con seme nero tra le 30 estratte con reimmissione' (prove indipendenti) risulta $X \sim \text{Bin}(n = 30, p = \frac{1}{3})$. $M(X) = 10$; $\text{Var}(X) = 6.6667$.

Dal momento che $np = 10 > 5$ ed $n(1 - p) = 20 > 5$ è possibile utilizzare l'approssimazione della variabile casuale Binomiale con la variabile casuale Normale $X \sim N(\mu = 10, \sigma^2 = 6.6667)$. Si ottiene

$$P(X \geq 15) = P\left(\frac{X - \mu}{\sigma} \geq \frac{15 - \mu}{\sigma}\right) = P\left(Z \geq \frac{15 - 10}{2.582}\right) =$$

=



3. Considerando ora $n = 30$ successive estrazioni (sempre con reimmissione) calcolare la probabilità che almeno 15 siano di colore nero.

Indicata con X la variabile casuale 'n. di carte con seme nero tra le 30 estratte con reimmissione' (prove indipendenti) risulta $X \sim \text{Bin}(n = 30, p = \frac{1}{3})$. $M(X) = 10$; $\text{Var}(X) = 6.6667$.

Dal momento che $np = 10 > 5$ ed $n(1 - p) = 20 > 5$ è possibile utilizzare l'approssimazione della variabile casuale Binomiale con la variabile casuale Normale $X \sim N(\mu = 10, \sigma^2 = 6.6667)$. Si ottiene

$$\begin{aligned} P(X \geq 15) &= P\left(\frac{X - \mu}{\sigma} \geq \frac{15 - \mu}{\sigma}\right) = P\left(Z \geq \frac{15 - 10}{2.582}\right) = \\ &= P(Z \geq 1.94) = \end{aligned}$$



3. Considerando ora $n = 30$ successive estrazioni (sempre con reimmissione) calcolare la probabilità che almeno 15 siano di colore nero.

Indicata con X la variabile casuale 'n. di carte con seme nero tra le 30 estratte con reimmissione' (prove indipendenti) risulta $X \sim \text{Bin}(n = 30, p = \frac{1}{3})$. $M(X) = 10$; $\text{Var}(X) = 6.6667$.

Dal momento che $np = 10 > 5$ ed $n(1 - p) = 20 > 5$ è possibile utilizzare l'approssimazione della variabile casuale Binomiale con la variabile casuale Normale $X \sim N(\mu = 10, \sigma^2 = 6.6667)$. Si ottiene

$$\begin{aligned} P(X \geq 15) &= P\left(\frac{X - \mu}{\sigma} \geq \frac{15 - \mu}{\sigma}\right) = P\left(Z \geq \frac{15 - 10}{2.582}\right) = \\ &= P(Z \geq 1.94) = 1 - P(Z \leq 1.94) = \end{aligned}$$



3. Considerando ora $n = 30$ successive estrazioni (sempre con reimmissione) calcolare la probabilità che almeno 15 siano di colore nero.

Indicata con X la variabile casuale 'n. di carte con seme nero tra le 30 estratte con reimmissione' (prove indipendenti) risulta $X \sim \text{Bin}(n = 30, p = \frac{1}{3})$. $M(X) = 10$; $\text{Var}(X) = 6.6667$.

Dal momento che $np = 10 > 5$ ed $n(1 - p) = 20 > 5$ è possibile utilizzare l'approssimazione della variabile casuale Binomiale con la variabile casuale Normale $X \sim N(\mu = 10, \sigma^2 = 6.6667)$. Si ottiene

$$\begin{aligned} P(X \geq 15) &= P\left(\frac{X - \mu}{\sigma} \geq \frac{15 - \mu}{\sigma}\right) = P\left(Z \geq \frac{15 - 10}{2.582}\right) = \\ &= P(Z \geq 1.94) = 1 - P(Z \leq 1.94) = 1 - 0.9738 = 0.0262. \end{aligned}$$

Tema T296

Prova scritta di STATISTICA del 17.01.2018

Prova scritta di
STATISTICA del
17.01.2018

Giuseppe Boari
Gabriele Cantaluppi



UNIVERSITÀ
CATTOLICA
del Sacro Cuore

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.

Giuseppe Boari
Gabriele Cantaluppi
Università Cattolica del Sacro Cuore, Milano



La seguente tabella riporta le distribuzioni relative condizionate della variabile $Y = \text{Numero di ingressi al museo nell'ultimo mese}$ rispetto alla variabile $X = \text{Zona di provenienza del visitatore}$, ricavata a seguito di una indagine campionaria:

$Y X$	0 + 2	2 + 7	7 + 10	
<i>NORD</i>	0.3	0.4	0.3	1
<i>CENTRO</i>	0.2	0.6	0.2	1
<i>SUD</i>	0.3	0.4	0.3	1

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.



1. Senza effettuare calcoli, si dica se è possibile concludere che esiste indipendenza stocastica tra X e Y . Motivare la risposta.

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.



1. Senza effettuare calcoli, si dica se è possibile concludere che esiste indipendenza stocastica tra X e Y . Motivare la risposta.

Non sussiste indipendenza stocastica tra le variabili X e Y in quanto le distribuzioni condizionate $Y|x_i$ non sono tra loro somiglianti. Abbiamo, infatti:

$$Y|x_1 \sim Y|x_3 \approx Y|x_2.$$

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.



2. Costruire la tabella delle frequenze congiunte sapendo che 50 dei visitatori intervistati provengono dal *NORD*, 50 dal *CENTRO* e 50 dal *SUD*.

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.



2. Costruire la tabella delle frequenze congiunte sapendo che 50 dei visitatori intervistati provengono dal *NORD*, 50 dal *CENTRO* e 50 dal *SUD*.

Ricordando che per la generica frequenza condizionata relativa vale $f(y_j|x_i) = \frac{n_{ij}}{n_{i\bullet}}$ abbiamo che $n_{ij} = f(y_j|x_i) \cdot n_{i\bullet}$ e

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.



2. Costruire la tabella delle frequenze congiunte sapendo che 50 dei visitatori intervistati provengono dal *NORD*, 50 dal *CENTRO* e 50 dal *SUD*.

Ricordando che per la generica frequenza condizionata relativa vale $f(y_j|x_i) = \frac{n_{ij}}{n_{i\bullet}}$ abbiamo che $n_{ij} = f(y_j|x_i) \cdot n_{i\bullet}$ e

$X \setminus Y$	0 - 2	2 - 7	7 - 10	$n_{i\bullet}$
<i>NORD</i>	15	20	15	50
<i>CENTRO</i>	10	30	10	50
<i>SUD</i>	15	20	15	50
$n_{\bullet j}$	40	70	40	150

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.



Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.

3. Si rappresenti graficamente la variabile Y e se ne calcoli moda e varianza.



Si considera la distribuzione marginale della variabile Y , ottenuta sommando le frequenze di colonna

y_j	$n_{\bullet j}$
0 $\rightarrow$ 2	40
2 $\rightarrow$ 7	70
7 $\rightarrow$ 10	40

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.



Si tratta di un carattere quantitativo con modalità raggruppate in classi che può essere rappresentato graficamente mediante un istogramma.

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.



L'istogramma può essere costruito dopo avere calcolato le densità $d_{\bullet j} = \frac{n_{\bullet j}}{a_j}$, (dove a_j sono le ampiezze di classe).

y_j	$n_{\bullet j}$	$a_{\bullet j}$	$d_{\bullet j}$
0 - 2	40	$2 - 0 = 2$	20
2 - 7	70	$7 - 2 = 5$	14
7 - 10	40	$10 - 7 = 3$	13.3333

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.

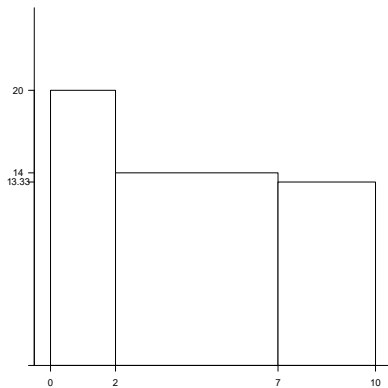


Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.





Si tratta di un carattere quantitativo con modalità raggruppate in classi.

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.



La moda corrisponde alla classe alla quale è associata la massima densità di frequenza

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.



La moda corrisponde alla classe alla quale è associata la massima densità di frequenza

y_j	$n_{\bullet j}$	$a_{\bullet j}$	$d_{\bullet j}$
0 - 2	40	$2 - 0 = 2$	20
2 - 7	70	$7 - 2 = 5$	14
7 - 10	40	$10 - 7 = 3$	13.3333

e risulta 0 - 2.

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.



La media aritmetica

$$M(Y) = \frac{1}{n} \sum_{j=1}^k y_j n_{\bullet j}$$

risulta 4.6333.

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.



La media aritmetica

$$M(Y) = \frac{1}{n} \sum_{j=1}^k y_j n_{\bullet j}$$

risulta 4.6333.

Si ottiene (considerando come modalità i valori centrali delle classi)

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.



La media aritmetica

$$M(Y) = \frac{1}{n} \sum_{j=1}^k y_j n_{\bullet j}$$

risulta 4.6333.

Si ottiene (considerando come modalità i valori centrali delle classi)

y_j	$n_{\bullet j}$	$y_j n_{\bullet j}$
1.0	40	40
4.5	70	315
8.5	40	340
	150	695

da cui segue

$$M(Y) = \frac{1}{n} \sum_{i=1}^k y_i n_{\bullet i} = \frac{695}{150} = 4.6333.$$

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.



La varianza

$$\text{Var}(Y) = M[(Y - \mu_Y)^2] = M(Y^2) - \mu_Y^2 = \frac{1}{n} \sum_{j=1}^k y_j^2 n_{\bullet j} - \mu_Y^2$$

risulta 7.5156.

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.



La varianza

$$Var(Y) = M[(Y - \mu_Y)^2] = M(Y^2) - \mu_Y^2 = \frac{1}{n} \sum_{j=1}^k y_j^2 n_{\bullet j} - \mu_Y^2$$

risulta 7.5156.

Si ottiene (considerando come modalità i valori centrali delle classi)

y_j	$n_{\bullet j}$	$y_j n_{\bullet j}$	$y_j^2 n_{\bullet j}$
1.0	40	40	40.0
4.5	70	315	1417.5
8.5	40	340	2890.0
	150	695	4347.5

da cui segue

$$\begin{aligned} Var(Y) = \sigma_Y^2 &= M(Y^2) - \mu_Y^2 = \\ &= \frac{4347.5}{150} - 4.6333^2 = 28.9833 - 21.4678 = 7.5155. \end{aligned}$$

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.



$$\text{Moda}(Y) = 0 \div 2 \quad \text{Var}(Y) = 7.5156$$

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.



Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.

4. Si calcoli l'indice di mutabilità normalizzato di X e se ne commenti il risultato.



Si considera la distribuzione marginale della variabile X , ottenuta sommando le frequenze di riga

X_i	$n_{i\bullet}$
<i>NORD</i>	50
<i>CENTRO</i>	50
<i>SUD</i>	50

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.



La mutabile X presenta massima eterogeneità, quindi gli indici di Gini e di Frosini assumono valore massimo, $F_N = G_N = 1$.

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.



Si consideri il seguente prospetto:

t	$NIBF(base = 2014)$	$NIBM$
2012		
2013		0.8
2014		1
2015		1
2016	0.6	

1. Si completi il prospetto con i dati mancanti.

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.



Dal momento che ${}_{2014}I_{2014} = 1$, considerando la serie dei numeri indici a base mobile assegnati si ricava

$${}_{2014}I_{2015} =$$

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.



Dal momento che ${}_{2014}I_{2014} = 1$, considerando la serie dei numeri indici a base mobile assegnati si ricava

$${}_{2014}I_{2015} = 1$$

$${}_{2014}I_{2013} =$$

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.



Dal momento che ${}_{2014}I_{2014} = 1$, considerando la serie dei numeri indici a base mobile assegnati si ricava

$${}_{2014}I_{2015} = 1$$

$${}_{2014}I_{2013} = 1$$

$${}_{2012}I_{2013} =$$

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.



Dal momento che ${}_{2014}I_{2014} = 1$, considerando la serie dei numeri indici a base mobile assegnati si ricava

$${}_{2014}I_{2015} = 1$$

$${}_{2014}I_{2013} = 1$$

$${}_{2012}I_{2013} = \frac{{}_{2014}I_{2013}}{{}_{2014}I_{2012}} = \frac{1}{{}_{2014}I_{2012}} = 0.8 \rightarrow {}_{2014}I_{2012} =$$

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.



Dal momento che ${}_{2014}I_{2014} = 1$, considerando la serie dei numeri indici a base mobile assegnati si ricava

$${}_{2014}I_{2015} = 1$$

$${}_{2014}I_{2013} = 1$$

$${}_{2012}I_{2013} = \frac{{}_{2014}I_{2013}}{{}_{2014}I_{2012}} = \frac{1}{{}_{2014}I_{2012}} = 0.8 \rightarrow {}_{2014}I_{2012} = \frac{1}{0.8} = 1.25$$

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.



Dal momento che ${}_{2014}I_{2014} = 1$, considerando la serie dei numeri indici a base mobile assegnati si ricava

$${}_{2014}I_{2015} = 1$$

$${}_{2014}I_{2013} = 1$$

$${}_{2012}I_{2013} = \frac{{}_{2014}I_{2013}}{{}_{2014}I_{2012}} = \frac{1}{{}_{2014}I_{2012}} = 0.8 \rightarrow {}_{2014}I_{2012} = \frac{1}{0.8} = 1.25$$

Abbiamo infine

$${}_{2015}I_{2016} =$$

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.



Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.

Dal momento che ${}_{2014}I_{2014} = 1$, considerando la serie dei numeri indici a base mobile assegnati si ricava

$${}_{2014}I_{2015} = 1$$

$${}_{2014}I_{2013} = 1$$

$${}_{2012}I_{2013} = \frac{{}_{2014}I_{2013}}{{}_{2014}I_{2012}} = \frac{1}{{}_{2014}I_{2012}} = 0.8 \rightarrow {}_{2014}I_{2012} = \frac{1}{0.8} = 1.25$$

Abbiamo infine

$${}_{2015}I_{2016} = \frac{{}_{2014}I_{2016}}{{}_{2014}I_{2015}} = \frac{0.6}{1} = 0.6$$



Si ottiene, quindi

t	$NIBF(base = 2014)$	$NIBM$
2012	1.25	
2013	1	0.8
2014	1	1
2015	1	1
2016	0.6	0.6

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.



2. Si calcoli il numero indice al tempo 2016 con base 2013 e si commenti il risultato ottenuto.

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.



2. Si calcoli il numero indice al tempo 2016 con base 2013 e si commenti il risultato ottenuto.

$${}_{2013}I_{2016} = \frac{X_{2016}}{X_{2013}} = \frac{{}_{2014}I_{2016}}{{}_{2014}I_{2013}} = \frac{0.6}{1} = 0.6$$

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.



2. Si calcoli il numero indice al tempo 2016 con base 2013 e si commenti il risultato ottenuto.

$${}_{2013}I_{2016} = \frac{X_{2016}}{X_{2013}} = \frac{{}_{2014}I_{2016}}{{}_{2014}I_{2013}} = \frac{0.6}{1} = 0.6$$

Tra il tempo 2013 e il tempo 2016 si è registrato un decremento complessivo del 40% della grandezza oggetto di studio.

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.



3. Si calcoli il tasso medio di variazione dal 2013 al 2016 e si commenti il risultato ottenuto.

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.



3. Si calcoli il tasso medio di variazione dal 2013 al 2016 e si commenti il risultato ottenuto.
Si ricorda che

$${}_{2013}I_{2016} = {}_{2013}I_{2014} \cdot {}_{2014}I_{2015} \cdot {}_{2015}I_{2016}$$

Si deve trovare il numero indice a base mobile medio.

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.



Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.

3. Si calcoli il tasso medio di variazione dal 2013 al 2016 e si commenti il risultato ottenuto. Si ricorda che

$${}_{2013}I_{2016} = {}_{2013}I_{2014} \cdot {}_{2014}I_{2015} \cdot {}_{2015}I_{2016}$$

Si deve trovare il numero indice a base mobile medio. Applicando il criterio di scelta della media secondo Chisini si ottiene

$${}_{2013}I_{2016} = {}_{2013}I_{2014} \cdot {}_{2014}I_{2015} \cdot {}_{2015}I_{2016} = \alpha \cdot \alpha \cdot \alpha = \alpha^3.$$



Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.

3. Si calcoli il tasso medio di variazione dal 2013 al 2016 e si commenti il risultato ottenuto. Si ricorda che

$${}_{2013}I_{2016} = {}_{2013}I_{2014} \cdot {}_{2014}I_{2015} \cdot {}_{2015}I_{2016}$$

Si deve trovare il numero indice a base mobile medio. Applicando il criterio di scelta della media secondo Chisini si ottiene

$${}_{2013}I_{2016} = {}_{2013}I_{2014} \cdot {}_{2014}I_{2015} \cdot {}_{2015}I_{2016} = \alpha \cdot \alpha \cdot \alpha = \alpha^3.$$

Quindi

$$\alpha = \sqrt[3]{{}_{2013}I_{2016}} = \sqrt[3]{{}_{2013}I_{2014} \cdot {}_{2014}I_{2015} \cdot {}_{2015}I_{2016}} = 0.8434.$$

Da cui segue una variazione relativa media pari a -0.1566 , corrispondente a un decremento medio annuo tra il 2013 e il 2016 del 15.66%.



Il seguente prospetto riporta i valori rilevati in corrispondenza a 8 capofamiglia con riferimento a due caratteri quantitativi X = numero dei componenti della famiglia e Y = spesa mensile per musei e mostre (in euro).

x_i	1	2	2	3	3	4	4	5
y_i	20	29	31	40	50	21	19	10

1. Si rappresentino graficamente le coppie dei dati e la funzione di regressione ipotizzando la spesa in funzione del numero dei componenti della famiglia.



La funzione di regressione associa a ogni valore x appartenente all'insieme di definizione della variabile X la media di Y condizionata a x , $M(Y|x)$

x	$M(Y x)$
1	20
2	$\frac{1}{2}(29 + 31) = 30$
3	$\frac{1}{2}(40 + 50) = 45$
4	$\frac{1}{2}(19 + 21) = 20$
5	10

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.



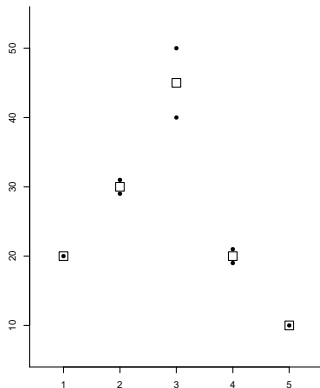
Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.

Le coppie di dati possono essere rappresentate mediante un grafico a dispersione.





Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.

2. Si calcolino, secondo il metodo dei minimi quadrati, i parametri dei seguenti modelli di regressione:

$$I) Y^* = a + b \cdot X \quad II) Y^* = c \cdot (X - 3)^2$$

Si riportano i calcoli necessari¹ per ottenere le stime dei parametri e gli indici di adattamento del modello

$$Y = a + bX + E$$

x_i	y_i	$x_i y_i$	x_i^2	y_i^2
1	20	20	1	400
2	29	58	4	841
2	31	62	4	961
3	40	120	9	1600
3	50	150	9	2500
4	21	84	16	441
4	19	76	16	361
5	10	50	25	100
24	220	620	84	7204
3	27.5	77.5	10.5	900.5
$M(X)$	$M(Y)$	$M(XY)$	$M(X^2)$	$M(Y^2)$

$$\hat{b} =$$



Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.

¹Si osserva che per ragioni di stampa alcuni calcoli potrebbero essere duplicati con riferimento ai diversi modelli oggetto di studio

Si riportano i calcoli necessari¹ per ottenere le stime dei parametri e gli indici di adattamento del modello

$$Y = a + bX + E$$

x_i	y_i	$x_i y_i$	x_i^2	y_i^2
1	20	20	1	400
2	29	58	4	841
2	31	62	4	961
3	40	120	9	1600
3	50	150	9	2500
4	21	84	16	441
4	19	76	16	361
5	10	50	25	100
24	220	620	84	7204
3	27.5	77.5	10.5	900.5
$M(X)$	$M(Y)$	$M(XY)$	$M(X^2)$	$M(Y^2)$

$$\begin{aligned}\hat{b} &= \frac{\text{Cov}(X, Y)}{\text{Var}(X)} = \frac{M(XY) - M(X)M(Y)}{M(X^2) - [M(X)]^2} = \\ &= \frac{77.5 - 3 \cdot 27.5}{10.5 - 3^2} = \frac{-5}{1.5} = -3.3333\end{aligned}$$

$$\hat{a} =$$

¹Si osserva che per ragioni di stampa alcuni calcoli potrebbero essere duplicati con riferimento ai diversi modelli oggetto di studio



Si riportano i calcoli necessari¹ per ottenere le stime dei parametri e gli indici di adattamento del modello

$$Y = a + bX + E$$

x_i	y_i	$x_i y_i$	x_i^2	y_i^2
1	20	20	1	400
2	29	58	4	841
2	31	62	4	961
3	40	120	9	1600
3	50	150	9	2500
4	21	84	16	441
4	19	76	16	361
5	10	50	25	100
24	220	620	84	7204
3	27.5	77.5	10.5	900.5
$M(X)$	$M(Y)$	$M(XY)$	$M(X^2)$	$M(Y^2)$

$$\hat{b} = \frac{\text{Cov}(X, Y)}{\text{Var}(X)} = \frac{M(XY) - M(X)M(Y)}{M(X^2) - [M(X)]^2} =$$

$$= \frac{77.5 - 3 \cdot 27.5}{10.5 - 3^2} = \frac{-5}{1.5} = -3.3333$$

$$\hat{a} = M(Y) - \hat{b}M(X) = 37.5$$

$$\rho^2 =$$

¹Si osserva che per ragioni di stampa alcuni calcoli potrebbero essere duplicati con riferimento ai diversi modelli oggetto di studio



Si riportano i calcoli necessari¹ per ottenere le stime dei parametri e gli indici di adattamento del modello

$$Y = a + bX + E$$

x_i	y_i	$x_i y_i$	x_i^2	y_i^2
1	20	20	1	400
2	29	58	4	841
2	31	62	4	961
3	40	120	9	1600
3	50	150	9	2500
4	21	84	16	441
4	19	76	16	361
5	10	50	25	100
24	220	620	84	7204
3	27.5	77.5	10.5	900.5
$M(X)$	$M(Y)$	$M(XY)$	$M(X^2)$	$M(Y^2)$

$$\begin{aligned}\hat{b} &= \frac{\text{Cov}(X, Y)}{\text{Var}(X)} = \frac{M(XY) - M(X)M(Y)}{M(X^2) - [M(X)]^2} = \\ &= \frac{77.5 - 3 \cdot 27.5}{10.5 - 3^2} = \frac{-5}{1.5} = -3.3333 \\ \hat{a} &= M(Y) - \hat{b}M(X) = 37.5 \\ \rho^2 &= \frac{[\text{Cov}(X, Y)]^2}{\text{Var}(X) \text{Var}(Y)} = \frac{(-5)^2}{1.5 \cdot 144.25} = 0.1155\end{aligned}$$

$$M(\hat{E}^2) = \text{Var}(\hat{E}) = \text{varianza residua} =$$

¹Si osserva che per ragioni di stampa alcuni calcoli potrebbero essere duplicati con riferimento ai diversi modelli oggetto di studio



Si riportano i calcoli necessari¹ per ottenere le stime dei parametri e gli indici di adattamento del modello

$$Y = a + bX + E$$

x_i	y_i	$x_i y_i$	x_i^2	y_i^2
1	20	20	1	400
2	29	58	4	841
2	31	62	4	961
3	40	120	9	1600
3	50	150	9	2500
4	21	84	16	441
4	19	76	16	361
5	10	50	25	100
24	220	620	84	7204
3	27.5	77.5	10.5	900.5
$M(X)$	$M(Y)$	$M(XY)$	$M(X^2)$	$M(Y^2)$

$$\begin{aligned}\hat{b} &= \frac{\text{Cov}(X, Y)}{\text{Var}(X)} = \frac{M(XY) - M(X)M(Y)}{M(X^2) - [M(X)]^2} = \\ &= \frac{77.5 - 3 \cdot 27.5}{10.5 - 3^2} = \frac{-5}{1.5} = -3.3333 \\ \hat{a} &= M(Y) - \hat{b}M(X) = 37.5 \\ \rho^2 &= \frac{[\text{Cov}(X, Y)]^2}{\text{Var}(X) \text{Var}(Y)} = \frac{(-5)^2}{1.5 \cdot 144.25} = 0.1155\end{aligned}$$

$$M(\hat{E}^2) = \text{Var}(\hat{E}) = \text{varianza residua} = \text{Var}(Y) \cdot (1 - \rho^2) = 127.5833$$

¹Si osserva che per ragioni di stampa alcuni calcoli potrebbero essere duplicati con riferimento ai diversi modelli oggetto di studio





Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.

Si riportano i calcoli necessari² per ottenere le stime dei parametri e la misura di adattamento del modello

$$Y = c(X - 3)^2 + E$$

$(x - 3)^2_j$	y_i	$(x - 3)^2_j y_i$	$(x - 3)^{22}_j$	y_i^2
4	20	80	16	400
1	29	29	1	841
1	31	31	1	961
0	40	0	0	1600
0	50	0	0	2500
1	21	21	1	441
1	19	19	1	361
4	10	40	16	100
		220	36	7204
		27.5	4.5	900.5
		$M((X - 3)^2 Y)$	$M((X - 3)^{22})$	$M(Y^2)$

$$\hat{c} =$$

²Si osserva che per ragioni di stampa alcuni calcoli potrebbero essere duplicati con riferimento ai diversi modelli oggetto di studio



Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.

Si riportano i calcoli necessari² per ottenere le stime dei parametri e la misura di adattamento del modello

$$Y = c(X - 3)^2 + E$$

$(x - 3)^2_i$	y_i	$(x - 3)^2_i y_i$	$(x - 3)^{22}_i$	y_i^2
4	20	80	16	400
1	29	29	1	841
1	31	31	1	961
0	40	0	0	1600
0	50	0	0	2500
1	21	21	1	441
1	19	19	1	361
4	10	40	16	100
		220	36	7204
		27.5	4.5	900.5
		$M((X - 3)^2 Y)$	$M((X - 3)^{22})$	$M(Y^2)$

$$\hat{c} = \frac{M((X - 3)^2 Y)}{M((X - 3)^{22})} = \frac{27.5}{4.5} = 6.1111$$

$$M(\hat{E}^2) =$$

²Si osserva che per ragioni di stampa alcuni calcoli potrebbero essere duplicati con riferimento ai diversi modelli oggetto di studio



Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.

Si riportano i calcoli necessari² per ottenere le stime dei parametri e la misura di adattamento del modello

$$Y = c(X - 3)^2 + E$$

$(x - 3)^2_i$	y_i	$(x - 3)^2_i y_i$	$(x - 3)^{22}_i$	y_i^2
4	20	80	16	400
1	29	29	1	841
1	31	31	1	961
0	40	0	0	1600
0	50	0	0	2500
1	21	21	1	441
1	19	19	1	361
4	10	40	16	100
		220	36	7204
		27.5	4.5	900.5
		$M((X - 3)^2 Y)$	$M((X - 3)^{22})$	$M(Y^2)$

$$\hat{c} = \frac{M((X - 3)^2 Y)}{M((X - 3)^{22})} = \frac{27.5}{4.5} = 6.1111$$

$$M(\hat{E}^2) = M[Y \cdot (Y - \hat{c}(X - 3)^2)] = M(Y^2) - \hat{c}M((X - 3)^2 Y) = 900.5 - 6.1111 \cdot 27.5 = 732.4444$$

²Si osserva che per ragioni di stampa alcuni calcoli potrebbero essere duplicati con riferimento ai diversi modelli oggetto di studio



Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.

$\hat{a} = 37.5$, $\hat{b} = -3.3333$, indice adattamento $M(\hat{E}^2) = 127.583$ $\hat{c} = 6.1111$, indice adattamento $M(\hat{E}^2) = 732.444$



3. Confrontare gli indici di adattamento, commentando i risultati.

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.



3. Confrontare gli indici di adattamento, commentando i risultati.

Il modello migliore è $Y^* = 37.5 - 3.3333 \cdot X$ in quanto caratterizzato da residuo quadratico medio più basso.

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.



4. Senza effettuare calcoli e motivando la risposta, si stabilisca: a) l'intervallo di possibili valori assumibili da $\eta_{Y|X}^2$; b) se $\eta_{X|Y}^2 < 1$.

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.



Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.

4. Senza effettuare calcoli e motivando la risposta, si stabilisca: a) l'intervallo di possibili valori assumibili da $\eta_{Y|X}^2$; b) se $\eta_{X|Y}^2 < 1$.

$\rho^2 < \eta_{Y|X}^2 < 1$ in quanto le medie condizionate non sono allineate e non sussiste dipendenza funzionale di Y da X dal momento che le distribuzioni condizionate $Y|X, x \in \{2, 3, 4\}$ non sono degeneri.



Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.

4. Senza effettuare calcoli e motivando la risposta, si stabilisca: a) l'intervallo di possibili valori assumibili da $\eta_{Y|X}^2$; b) se $\eta_{X|Y}^2 < 1$.

$\rho^2 < \eta_{Y|X}^2 < 1$ in quanto le medie condizionate non sono allineate e non sussiste dipendenza funzionale di Y da X dal momento che le distribuzioni condizionate $Y|X, x \in \{2, 3, 4\}$ non sono degeneri.

$\eta_{X|Y}^2 = 1$ in quanto le distribuzioni condizionate $X|y$ sono tutte degeneri.

Esercizio 4

Prova scritta di
STATISTICA del
17.01.2018

Giuseppe Boari
Gabriele Cantaluppi



UNIVERSITÀ
CATTOLICA
del Sacro Cuore

La retribuzione media mensile lorda di un laureato in Economia è distribuita come una Normale di media $\mu = 1600$ euro e varianza $\sigma^2 = 10\,000$ euro.

1. Quale percentuale di laureati guadagna più di 1700 euro al mese?

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.

Esercizio 4



Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.

La retribuzione media mensile lorda di un laureato in Economia è distribuita come una Normale di media $\mu = 1600$ euro e varianza $\sigma^2 = 10\,000$ euro.

1. Quale percentuale di laureati guadagna più di 1700 euro al mese?

Indicata con $X \sim N(\mu = 1600, \sigma^2 = 10000)$ la variabile casuale che descrive la retribuzione media mensile lorda, occorre, in primo luogo, determinare la seguente probabilità

$$P(X > 1700).$$

Abbiamo

$$P(X > 1700) =$$

Esercizio 4



Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.

La retribuzione media mensile lorda di un laureato in Economia è distribuita come una Normale di media $\mu = 1600$ euro e varianza $\sigma^2 = 10\,000$ euro.

1. Quale percentuale di laureati guadagna più di 1700 euro al mese?

Indicata con $X \sim N(\mu = 1600, \sigma^2 = 10000)$ la variabile casuale che descrive la retribuzione media mensile lorda, occorre, in primo luogo, determinare la seguente probabilità

$$P(X > 1700).$$

Abbiamo

$$P(X > 1700) = P\left(\frac{X - \mu}{\sigma} > \frac{1700 - \mu}{\sigma}\right) =$$

Esercizio 4



[Esercizio 1.](#)

[Esercizio 2.](#)

[Esercizio 3.](#)

[Esercizio 4.](#)

La retribuzione media mensile lorda di un laureato in Economia è distribuita come una Normale di media $\mu = 1600$ euro e varianza $\sigma^2 = 10\,000$ euro.

1. Quale percentuale di laureati guadagna più di 1700 euro al mese?

Indicata con $X \sim N(\mu = 1600, \sigma^2 = 10000)$ la variabile casuale che descrive la retribuzione media mensile lorda, occorre, in primo luogo, determinare la seguente probabilità

$$P(X > 1700).$$

Abbiamo

$$\begin{aligned} P(X > 1700) &= P\left(\frac{X - \mu}{\sigma} > \frac{1700 - \mu}{\sigma}\right) = P\left(Z > \frac{1700 - 1600}{100}\right) = \\ &= \end{aligned}$$

Esercizio 4



Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.

La retribuzione media mensile lorda di un laureato in Economia è distribuita come una Normale di media $\mu = 1600$ euro e varianza $\sigma^2 = 10\,000$ euro.

1. Quale percentuale di laureati guadagna più di 1700 euro al mese?

Indicata con $X \sim N(\mu = 1600, \sigma^2 = 10000)$ la variabile casuale che descrive la retribuzione media mensile lorda, occorre, in primo luogo, determinare la seguente probabilità

$$P(X > 1700).$$

Abbiamo

$$\begin{aligned} P(X > 1700) &= P\left(\frac{X - \mu}{\sigma} > \frac{1700 - \mu}{\sigma}\right) = P\left(Z > \frac{1700 - 1600}{100}\right) = \\ &= P(Z > 1) = \end{aligned}$$



La retribuzione media mensile lorda di un laureato in Economia è distribuita come una Normale di media $\mu = 1600$ euro e varianza $\sigma^2 = 10\,000$ euro.

1. Quale percentuale di laureati guadagna più di 1700 euro al mese?

Indicata con $X \sim N(\mu = 1600, \sigma^2 = 10000)$ la variabile casuale che descrive la retribuzione media mensile lorda, occorre, in primo luogo, determinare la seguente probabilità

$$P(X > 1700).$$

Abbiamo

$$\begin{aligned} P(X > 1700) &= P\left(\frac{X - \mu}{\sigma} > \frac{1700 - \mu}{\sigma}\right) = P\left(Z > \frac{1700 - 1600}{100}\right) = \\ &= P(Z > 1) = 1 - P(Z \leq 1) = \end{aligned}$$



La retribuzione media mensile lorda di un laureato in Economia è distribuita come una Normale di media $\mu = 1600$ euro e varianza $\sigma^2 = 10\,000$ euro.

1. Quale percentuale di laureati guadagna più di 1700 euro al mese?

Indicata con $X \sim N(\mu = 1600, \sigma^2 = 10000)$ la variabile casuale che descrive la retribuzione media mensile lorda, occorre, in primo luogo, determinare la seguente probabilità

$$P(X > 1700).$$

Abbiamo

$$\begin{aligned} P(X > 1700) &= P\left(\frac{X - \mu}{\sigma} > \frac{1700 - \mu}{\sigma}\right) = P\left(Z > \frac{1700 - 1600}{100}\right) = \\ &= P(Z > 1) = 1 - P(Z \leq 1) = 1 - 0.8413 = 0.1587. \end{aligned}$$

La percentuale di laureati che guadagna più di 1700 euro al mese risulta 15.87%.



2. Quale è il reddito limite mensile che definisce il 10% superiore della retribuzione dei laureati in Economia?

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.



Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.

2. Quale è il reddito limite mensile che definisce il 10% superiore della retribuzione dei laureati in Economia?

Il reddito limite mensile che definisce il 10% superiore della retribuzione dei laureati in Economia corrisponde al percentile di ordine 0.90 della distribuzione di X .

$$\{x : P(X \leq x) = 0.90\} =$$



Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.

2. Quale è il reddito limite mensile che definisce il 10% superiore della retribuzione dei laureati in Economia?

Il reddito limite mensile che definisce il 10% superiore della retribuzione dei laureati in Economia corrisponde al percentile di ordine 0.90 della distribuzione di X .

$$\{x : P(X \leq x) = 0.90\} = \left\{x : P\left(\frac{X - \mu}{\sigma} \leq \frac{x - \mu}{\sigma}\right) = 0.90\right\} =$$
$$=$$



Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.

2. Quale è il reddito limite mensile che definisce il 10% superiore della retribuzione dei laureati in Economia?

Il reddito limite mensile che definisce il 10% superiore della retribuzione dei laureati in Economia corrisponde al percentile di ordine 0.90 della distribuzione di X .

$$\begin{aligned}\{x : P(X \leq x) = 0.90\} &= \left\{x : P\left(\frac{X - \mu}{\sigma} \leq \frac{x - \mu}{\sigma}\right) = 0.90\right\} = \\ &= \left\{x : P\left(Z \leq \frac{x - 1600}{100}\right) = 0.90\right\}.\end{aligned}$$

Dalla lettura inversa della Tavola abbiamo $z =$



Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.

2. Quale è il reddito limite mensile che definisce il 10% superiore della retribuzione dei laureati in Economia?

Il reddito limite mensile che definisce il 10% superiore della retribuzione dei laureati in Economia corrisponde al percentile di ordine 0.90 della distribuzione di X .

$$\begin{aligned}\{x : P(X \leq x) = 0.90\} &= \left\{x : P\left(\frac{X - \mu}{\sigma} \leq \frac{x - \mu}{\sigma}\right) = 0.90\right\} = \\ &= \left\{x : P\left(Z \leq \frac{x - 1600}{100}\right) = 0.90\right\}.\end{aligned}$$

Dalla lettura inversa della Tavola abbiamo $z = \frac{x - 1600}{100} = 1.28$ da cui segue $x =$



Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.

2. Quale è il reddito limite mensile che definisce il 10% superiore della retribuzione dei laureati in Economia?

Il reddito limite mensile che definisce il 10% superiore della retribuzione dei laureati in Economia corrisponde al percentile di ordine 0.90 della distribuzione di X .

$$\begin{aligned}\{x : P(X \leq x) = 0.90\} &= \left\{x : P\left(\frac{X - \mu}{\sigma} \leq \frac{x - \mu}{\sigma}\right) = 0.90\right\} = \\ &= \left\{x : P\left(Z \leq \frac{x - 1600}{100}\right) = 0.90\right\}.\end{aligned}$$

Dalla lettura inversa della Tavola abbiamo $z = \frac{x - 1600}{100} = 1.28$ da cui segue $x = 1728$.



3. Si calcoli la probabilità che su 10 laureati selezionati casualmente almeno uno guadagni più di 1700 euro al mese.

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.



Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.

3. Si calcoli la probabilità che su 10 laureati selezionati casualmente almeno uno guadagni più di 1700 euro al mese.

Sia X la variabile casuale 'n. di laureati con guadagno superiore a 1700 euro tra i 10 selezionati'. Dal momento che la popolazione dei laureati ha dimensione molto elevata, abbiamo $X \sim \text{Bin}(n = 10, p = 0.1587)$.

$$P(X \geq 1) =$$



Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.

3. Si calcoli la probabilità che su 10 laureati selezionati casualmente almeno uno guadagni più di 1700 euro al mese.

Sia X la variabile casuale 'n. di laureati con guadagno superiore a 1700 euro tra i 10 selezionati'. Dal momento che la popolazione dei laureati ha dimensione molto elevata, abbiamo $X \sim \text{Bin}(n = 10, p = 0.1587)$.

$$P(X \geq 1) = 1 - P(X = 0) =$$

=



Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.

3. Si calcoli la probabilità che su 10 laureati selezionati casualmente almeno uno guadagni più di 1700 euro al mese.

Sia X la variabile casuale 'n. di laureati con guadagno superiore a 1700 euro tra i 10 selezionati'. Dal momento che la popolazione dei laureati ha dimensione molto elevata, abbiamo $X \sim \text{Bin}(n = 10, p = 0.1587)$.

$$\begin{aligned} P(X \geq 1) &= 1 - P(X = 0) = \\ &= 1 - \binom{10}{0} 0.1587^0 \cdot 0.8513^{10} = \\ &= \end{aligned}$$



Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.

3. Si calcoli la probabilità che su 10 laureati selezionati casualmente almeno uno guadagni più di 1700 euro al mese.

Sia X la variabile casuale 'n. di laureati con guadagno superiore a 1700 euro tra i 10 selezionati'. Dal momento che la popolazione dei laureati ha dimensione molto elevata, abbiamo $X \sim \text{Bin}(n = 10, p = 0.1587)$.

$$\begin{aligned} P(X \geq 1) &= 1 - P(X = 0) = \\ &= 1 - \binom{10}{0} 0.1587^0 \cdot 0.8513^{10} = \\ &= 1 - 0.1777 = 0.8223. \end{aligned}$$

Tema T297

Prova scritta di STATISTICA del 31.01.2018

Prova scritta di
STATISTICA del
31.01.2018

Giuseppe Boari
Gabriele Cantaluppi



UNIVERSITÀ
CATTOLICA
del Sacro Cuore

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.

Giuseppe Boari
Gabriele Cantaluppi
Università Cattolica del Sacro Cuore, Milano



Con riferimento a 10 famiglie sono disponibili le distribuzioni marginali dei due caratteri $X =$ 'numero di visite effettuate a musei nello scorso anno' e $Y =$ 'durata massima della visita'.

x_i	n_i
10	4
11	2
12	4

y_j	n_j
1 + 3	4
3 + 7	6

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.



1. Si rappresentino graficamente le due variabili X e Y .

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.



Si considera la distribuzione della variabile X ,

x_i	n_i
10	4
11	2
12	4

Esercizio 1.

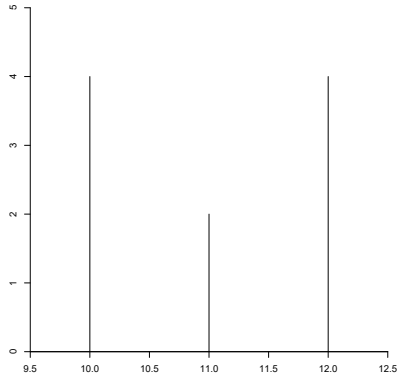
Esercizio 2.

Esercizio 3.

Esercizio 4.



Si tratta di un carattere quantitativo con modalità non raggruppate in classi che può essere rappresentato graficamente mediante un grafico a bastoncini.



Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.



Si considera la distribuzione della variabile X ,

x_i	n_i
1 + 3	4
3 + 7	6

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.



Si tratta di un carattere quantitativo con modalità raggruppate in classi che può essere rappresentato graficamente mediante un istogramma.

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.



L'istogramma può essere costruito dopo avere calcolato le densità $d_i = \frac{n_i}{a_i}$, (dove a_i sono le ampiezze di classe).

x_i	n_i	a_i	d_i
1 - 3	4	3 - 1 = 2	2
3 - 7	6	7 - 3 = 4	1.5

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.

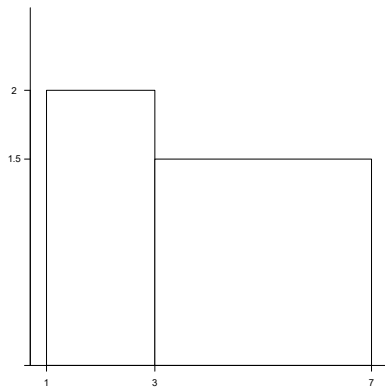


Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.





2. Si confrontino le due distribuzioni mediante opportuni indici di posizione e di variabilità.

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.



Si considera la distribuzione della variabile X ,

x_i	n_i
10	4
11	2
12	4

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.



Si tratta di un carattere quantitativo con modalità non raggruppate in classi.

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.



La media aritmetica

$$M(X) = \frac{1}{n} \sum_{i=1}^k x_i n_i$$

risulta 11.

Si osserva, infatti, che la distribuzione è simmetrica.

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.



La varianza

$$\text{Var}(X) = M[(X - \mu_X)^2] = M(X^2) - \mu_X^2 = \frac{1}{n} \sum_{i=1}^k x_i^2 n_i - \mu_X^2$$

risulta 0.8.

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.



La varianza

$$\text{Var}(X) = M[(X - \mu_X)^2] = M(X^2) - \mu_X^2 = \frac{1}{n} \sum_{i=1}^k x_i^2 n_i - \mu_X^2$$

risulta 0.8.

Si ottiene

x_i	n_i	$x_i n_i$	$x_i^2 n_i$
10	4	40	400
11	2	22	242
12	4	48	576
	10	110	1218

da cui segue

$$\begin{aligned}\text{Var}(X) = \sigma_X^2 &= M(X^2) - \mu_X^2 = \\ &= \frac{1218}{10} - 11^2 = 121.8 - 121 = 0.8.\end{aligned}$$

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.



Per effettuare confronti in termini di variabilità tra caratteri quantitativi occorre utilizzare una misura relativa, ad esempio il coefficiente di variazione

$$CV = \frac{\sigma}{\mu} = \frac{\sqrt{\text{Var}(X)}}{M(X)} = \frac{\sqrt{0.8}}{11} = \frac{0.8944}{11} = 0.0813.$$

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.



Si considera la distribuzione della variabile Y ,

y_i	n_i
1 + 3	4
3 + 7	6

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.



Si tratta di un carattere quantitativo con modalità raggruppate in classi.

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.



La media aritmetica

$$M(Y) = \frac{1}{n} \sum_{i=1}^k y_i n_i$$

risulta 3.8.

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.



La media aritmetica

$$M(Y) = \frac{1}{n} \sum_{i=1}^k y_i n_i$$

risulta 3.8.

Si ottiene (considerando come modalità i valori centrali delle classi)

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.



La media aritmetica

$$M(Y) = \frac{1}{n} \sum_{i=1}^k y_i n_i$$

risulta 3.8.

Si ottiene (considerando come modalità i valori centrali delle classi)

y_i	n_i	$y_i n_i$
2	4	8
5	6	30
	10	38

da cui segue

$$M(Y) = \frac{1}{n} \sum_{i=1}^k y_i n_i = \frac{38}{10} = 3.8.$$

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.



La varianza

$$\text{Var}(Y) = M[(Y - \mu_Y)^2] = M(Y^2) - \mu_Y^2 = \frac{1}{n} \sum_{i=1}^k y_i^2 n_i - \mu_Y^2$$

risulta 2.16.

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.



La varianza

$$\text{Var}(Y) = M[(Y - \mu_Y)^2] = M(Y^2) - \mu_Y^2 = \frac{1}{n} \sum_{i=1}^k y_i^2 n_i - \mu_Y^2$$

risulta 2.16.

Si ottiene (considerando come modalità i valori centrali delle classi)

y_i	n_i	$y_i n_i$	$y_i^2 n_i$
2	4	8	16
5	6	30	150
	10	38	166

da cui segue

$$\begin{aligned} \text{Var}(Y) = \sigma_Y^2 &= M(Y^2) - \mu_Y^2 = \\ &= \frac{166}{10} - 3.8^2 = 16.6 - 14.44 = 2.16. \end{aligned}$$

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.



Per effettuare confronti in termini di variabilità tra caratteri quantitativi occorre utilizzare una misura relativa, ad esempio il coefficiente di variazione

$$CV = \frac{\sigma}{\mu} = \frac{\sqrt{\text{Var}(Y)}}{M(Y)} = \frac{\sqrt{2.16}}{3.8} = \frac{1.4697}{3.8} = 0.3868.$$

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.



X Indice di posizione: $M(X) = 11$ Indice di variabilità: $CV(X) = 0.0813$

Y Indice di posizione: $M(Y) = 3.8$ Indice di variabilità: $CV(Y) = 0.3868$

Non ha senso il confronto mediante indici di posizione tra le due variabili.

La durata media della visita presenta un livello di variabilità superiore rispetto al numero medio delle visite.

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.



3. Si costruisca la distribuzione congiunta di X e Y nell'ipotesi che tra i due caratteri sussista dipendenza funzionale.

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.



3. Si costruisca la distribuzione congiunta di X e Y nell'ipotesi che tra i due caratteri sussista dipendenza funzionale.

Occorre costruire una tabella nella quale le distribuzioni condizionate $Y|x_i$ sono degeneri. Si riportano due possibili soluzioni.

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.



3. Si costruisca la distribuzione congiunta di X e Y nell'ipotesi che tra i due caratteri sussista dipendenza funzionale.

Occorre costruire una tabella nella quale le distribuzioni condizionate $Y|x_i$ sono degeneri. Si riportano due possibili soluzioni.

$X \setminus Y$	1	3	
10	4	0	4
11	0	2	2
12	0	4	4
	4	6	10

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.



3. Si costruisca la distribuzione congiunta di X e Y nell'ipotesi che tra i due caratteri sussista dipendenza funzionale.

Occorre costruire una tabella nella quale le distribuzioni condizionate $Y|x_i$ sono degeneri. Si riportano due possibili soluzioni.

$X \setminus Y$	1	3	7
10	4	0	4
11	0	2	2
12	0	4	4
	4	6	10

$X \setminus Y$	1	3	7
10	0	4	4
11	0	2	2
12	4	0	4
	4	6	10

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.



4. Si costruisca la distribuzione congiunta di X e Y nell'ipotesi di indipendenza stocastica.

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.



4. Si costruisca la distribuzione congiunta di X e Y nell'ipotesi di indipendenza stocastica.

Occorre costruire una tabella con distribuzioni condizionate somiglianti: $Y|x_i \sim Y$ e $X|y_j \sim X$.

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.



4. Si costruisca la distribuzione congiunta di X e Y nell'ipotesi di indipendenza stocastica.

Occorre costruire una tabella con distribuzioni condizionate somiglianti: $Y|x_i \sim Y$ e $X|y_j \sim X$.

$X \backslash Y$	1-3	3-7	
10	1.6	2.4	4
11	0.8	1.2	2
12	1.6	2.4	4
	4	6	10

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.



Si consideri il seguente prospetto:

t	x_t	$NIBF(base = 2013)$	$NIBM$
2012			
2013	24	1	0.5
2014	24		
2015		0.5	
2016			0.5

1. Si completi il prospetto con i dati mancanti.

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.

Abbiamo:

$$2012/2013 =$$



Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.



Abbiamo:

$${}_{2012}I_{2013} = \frac{{}_{2013}I_{2013}}{{}_{2013}I_{2012}} = \frac{x_{2013}}{x_{2012}} = 0.5 \rightarrow x_{2012} =$$

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.



Abbiamo:

$${}_{2012}I_{2013} = \frac{{}_{2013}I_{2013}}{{}_{2013}I_{2012}} = \frac{x_{2013}}{x_{2012}} = 0.5 \rightarrow x_{2012} = \frac{24}{0.5} = 48 \text{ e}$$

$${}_{2013}I_{2012} =$$

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.



Abbiamo:

$${}_{2012}I_{2013} = \frac{{}_{2013}I_{2013}}{{}_{2013}I_{2012}} = \frac{x_{2013}}{x_{2012}} = 0.5 \rightarrow x_{2012} = \frac{24}{0.5} = 48 \text{ e}$$

$${}_{2013}I_{2012} = \frac{1}{{}_{2012}I_{2013}} = \frac{1}{0.5} = 2$$

$${}_{2013}I_{2014} =$$

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.



Abbiamo:

$${}_{2012}I_{2013} = \frac{{}_{2013}I_{2013}}{{}_{2013}I_{2012}} = \frac{x_{2013}}{x_{2012}} = 0.5 \rightarrow x_{2012} = \frac{24}{0.5} = 48 \text{ e}$$

$${}_{2013}I_{2012} = \frac{1}{{}_{2012}I_{2013}} = \frac{1}{0.5} = 2$$

$${}_{2013}I_{2014} = \frac{x_{2014}}{x_{2013}} = \frac{24}{24} = 1$$

$${}_{2013}I_{2015} =$$

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.



Abbiamo:

$${}_{2012}I_{2013} = \frac{{}_{2013}I_{2013}}{{}_{2013}I_{2012}} = \frac{x_{2013}}{x_{2012}} = 0.5 \rightarrow x_{2012} = \frac{24}{0.5} = 48 \text{ e}$$

$${}_{2013}I_{2012} = \frac{1}{{}_{2012}I_{2013}} = \frac{1}{0.5} = 2$$

$${}_{2013}I_{2014} = \frac{x_{2014}}{x_{2013}} = \frac{24}{24} = 1$$

$${}_{2013}I_{2015} = \frac{x_{2015}}{x_{2013}} = 0.5 \rightarrow x_{2015} =$$

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.



Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.

Abbiamo:

$${}_{2012}l_{2013} = \frac{{}_{2013}l_{2013}}{{}_{2013}l_{2012}} = \frac{x_{2013}}{x_{2012}} = 0.5 \rightarrow x_{2012} = \frac{24}{0.5} = 48 \text{ e}$$

$${}_{2013}l_{2012} = \frac{1}{{}_{2012}l_{2013}} = \frac{1}{0.5} = 2$$

$${}_{2013}l_{2014} = \frac{x_{2014}}{x_{2013}} = \frac{24}{24} = 1$$

$${}_{2013}l_{2015} = \frac{x_{2015}}{x_{2013}} = 0.5 \rightarrow x_{2015} = x_{2013} \cdot 0.5 = 12$$

$${}_{2014}l_{2015} =$$



Abbiamo:

$${}_{2012}I_{2013} = \frac{{}_{2013}I_{2013}}{{}_{2013}I_{2012}} = \frac{x_{2013}}{x_{2012}} = 0.5 \rightarrow x_{2012} = \frac{24}{0.5} = 48 \text{ e}$$

$${}_{2013}I_{2012} = \frac{1}{{}_{2012}I_{2013}} = \frac{1}{0.5} = 2$$

$${}_{2013}I_{2014} = \frac{x_{2014}}{x_{2013}} = \frac{24}{24} = 1$$

$${}_{2013}I_{2015} = \frac{x_{2015}}{x_{2013}} = 0.5 \rightarrow x_{2015} = x_{2013} \cdot 0.5 = 12$$

$${}_{2014}I_{2015} = \frac{x_{2015}}{x_{2014}} =$$

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.



Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.

Abbiamo:

$${}_{2012}I_{2013} = \frac{{}_{2013}I_{2013}}{{}_{2013}I_{2012}} = \frac{x_{2013}}{x_{2012}} = 0.5 \rightarrow x_{2012} = \frac{24}{0.5} = 48 \text{ e}$$

$${}_{2013}I_{2012} = \frac{1}{{}_{2012}I_{2013}} = \frac{1}{0.5} = 2$$

$${}_{2013}I_{2014} = \frac{x_{2014}}{x_{2013}} = \frac{24}{24} = 1$$

$${}_{2013}I_{2015} = \frac{x_{2015}}{x_{2013}} = 0.5 \rightarrow x_{2015} = x_{2013} \cdot 0.5 = 12$$

$${}_{2014}I_{2015} = \frac{x_{2015}}{x_{2014}} = \frac{12}{24} = 0.5$$

$${}_{2015}I_{2016} =$$



Abbiamo:

$${}_{2012}I_{2013} = \frac{{}_{2013}I_{2013}}{{}_{2013}I_{2012}} = \frac{x_{2013}}{x_{2012}} = 0.5 \rightarrow x_{2012} = \frac{24}{0.5} = 48 \text{ e}$$

$${}_{2013}I_{2012} = \frac{1}{{}_{2012}I_{2013}} = \frac{1}{0.5} = 2$$

$${}_{2013}I_{2014} = \frac{x_{2014}}{x_{2013}} = \frac{24}{24} = 1$$

$${}_{2013}I_{2015} = \frac{x_{2015}}{x_{2013}} = 0.5 \rightarrow x_{2015} = x_{2013} \cdot 0.5 = 12$$

$${}_{2014}I_{2015} = \frac{x_{2015}}{x_{2014}} = \frac{12}{24} = 0.5$$

$${}_{2015}I_{2016} = \frac{{}_{2013}I_{2016}}{{}_{2013}I_{2015}} = \frac{x_{2016}}{x_{2015}} = 0.5 \rightarrow x_{2016} =$$

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.



Abbiamo:

$${}_{2012}I_{2013} = \frac{{}_{2013}I_{2013}}{{}_{2013}I_{2012}} = \frac{x_{2013}}{x_{2012}} = 0.5 \rightarrow x_{2012} = \frac{24}{0.5} = 48 \text{ e}$$

$${}_{2013}I_{2012} = \frac{1}{{}_{2012}I_{2013}} = \frac{1}{0.5} = 2$$

$${}_{2013}I_{2014} = \frac{x_{2014}}{x_{2013}} = \frac{24}{24} = 1$$

$${}_{2013}I_{2015} = \frac{x_{2015}}{x_{2013}} = 0.5 \rightarrow x_{2015} = x_{2013} \cdot 0.5 = 12$$

$${}_{2014}I_{2015} = \frac{x_{2015}}{x_{2014}} = \frac{12}{24} = 0.5$$

$${}_{2015}I_{2016} = \frac{{}_{2013}I_{2016}}{{}_{2013}I_{2015}} = \frac{x_{2016}}{x_{2015}} = 0.5 \rightarrow x_{2016} = 12 \cdot 0.5 = 6 \text{ e}$$

$${}_{2013}I_{2016} =$$

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.



Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.

Abbiamo:

$${}_{2012}l_{2013} = \frac{{}_{2013}l_{2013}}{{}_{2013}l_{2012}} = \frac{x_{2013}}{x_{2012}} = 0.5 \rightarrow x_{2012} = \frac{24}{0.5} = 48 \text{ e}$$

$${}_{2013}l_{2012} = \frac{1}{{}_{2012}l_{2013}} = \frac{1}{0.5} = 2$$

$${}_{2013}l_{2014} = \frac{x_{2014}}{x_{2013}} = \frac{24}{24} = 1$$

$${}_{2013}l_{2015} = \frac{x_{2015}}{x_{2013}} = 0.5 \rightarrow x_{2015} = x_{2013} \cdot 0.5 = 12$$

$${}_{2014}l_{2015} = \frac{x_{2015}}{x_{2014}} = \frac{12}{24} = 0.5$$

$${}_{2015}l_{2016} = \frac{{}_{2013}l_{2016}}{{}_{2013}l_{2015}} = \frac{x_{2016}}{x_{2015}} = 0.5 \rightarrow x_{2016} = 12 \cdot 0.5 = 6 \text{ e}$$

$${}_{2013}l_{2016} = {}_{2013}l_{2015} \cdot 0.5 = 0.5 \cdot 0.5 = 0.25$$



Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.

Si ottiene, quindi

t	x_t	$NIBF(base = 2013)$	$NIBM$
2012	48	2	
2013	24	1	0.5
2014	24	1	1
2015	12	0.5	0.5
2016	6	0.25	0.5

2. Si calcoli il numero indice al tempo 2013 con base 2015 e si commenti il risultato ottenuto.



Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.

Si ottiene, quindi

t	x_t	$NIBF(base = 2013)$	$NIBM$
2012	48	2	
2013	24	1	0.5
2014	24	1	1
2015	12	0.5	0.5
2016	6	0.25	0.5

2. Si calcoli il numero indice al tempo 2013 con base 2015 e si commenti il risultato ottenuto.

$${}_{2015}I_{2013} = \frac{x_{2013}}{x_{2015}} = \frac{24}{12} = 2$$



Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.

Si ottiene, quindi

t	x_t	$NIBF(base = 2013)$	$NIBM$
2012	48	2	
2013	24	1	0.5
2014	24	1	1
2015	12	0.5	0.5
2016	6	0.25	0.5

2. Si calcoli il numero indice al tempo 2013 con base 2015 e si commenti il risultato ottenuto.

$${}_{2015}I_{2013} = \frac{x_{2013}}{x_{2015}} = \frac{24}{12} = 2$$

La grandezza al tempo 2013 era superiore del 100% rispetto alla grandezza al tempo 2015. Ovvero il valore della grandezza nel 2013 era il 200% del valore registrato nel 2015.



3. Si calcoli il tasso medio di variazione dal 2013 al 2015 e si commenti il risultato ottenuto.

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.



3. Si calcoli il tasso medio di variazione dal 2013 al 2015 e si commenti il risultato ottenuto.
Si ricorda che

$${}_{2013}I_{2015} = {}_{2013}I_{2014} \cdot {}_{2014}I_{2015}$$

Si deve trovare il numero indice a base mobile medio.

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.



Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.

3. Si calcoli il tasso medio di variazione dal 2013 al 2015 e si commenti il risultato ottenuto. Si ricorda che

$${}_{2013}I_{2015} = {}_{2013}I_{2014} \cdot {}_{2014}I_{2015}$$

Si deve trovare il numero indice a base mobile medio. Applicando il criterio di scelta della media secondo Chisini si ottiene

$${}_{2013}I_{2015} = {}_{2013}I_{2014} \cdot {}_{2014}I_{2015} = \alpha \cdot \alpha = \alpha^2.$$



Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.

3. Si calcoli il tasso medio di variazione dal 2013 al 2015 e si commenti il risultato ottenuto. Si ricorda che

$${}_{2013}I_{2015} = {}_{2013}I_{2014} \cdot {}_{2014}I_{2015}$$

Si deve trovare il numero indice a base mobile medio. Applicando il criterio di scelta della media secondo Chisini si ottiene

$${}_{2013}I_{2015} = {}_{2013}I_{2014} \cdot {}_{2014}I_{2015} = \alpha \cdot \alpha = \alpha^2.$$

Quindi

$$\alpha = \sqrt[2]{{}_{2013}I_{2015}} = \sqrt[2]{{}_{2013}I_{2014} \cdot {}_{2014}I_{2015}} = 0.7071.$$

Da cui segue una variazione relativa media pari a -0.2929 , corrispondente a un decremento medio annuo tra il 2013 e il 2015 del 29.29%.



Il seguente prospetto riporta i valori rilevati in corrispondenza a 5 capifamiglia con riferimento ai tre caratteri X = 'numero dei componenti della famiglia', Y = 'spesa mensile per musei e mostre (in euro)' e W = 'livello di gradimento registrato nelle visite dello scorso anno' con categorie b = 'basso' e a = 'alto'.

x_i	2	2	3	4	4
y_i	41	45	45	45	49
w_i	a	a	b	b	a

1. Si rappresentino graficamente le coppie dei dati e la funzione di regressione ipotizzando la spesa in funzione del numero dei componenti della famiglia.

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.



La funzione di regressione associa a ogni valore x appartenente all'insieme di definizione della variabile X la media di Y condizionata a x , $M(Y|x)$

x	$M(Y x)$
2	$\frac{1}{2}(41 + 45) = 43$
3	45
4	$\frac{1}{2}(45 + 49) = 47$

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.



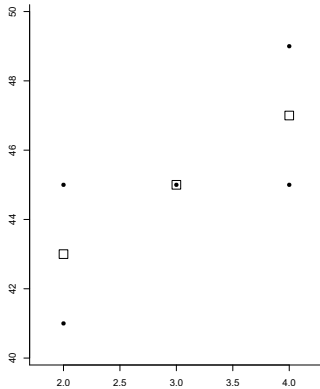
Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.

Le coppie di dati possono essere rappresentate mediante un grafico a dispersione.





Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.

2. Si calcolino, secondo il metodo dei minimi quadrati, i parametri dei seguenti modelli di regressione:

$$I) Y^* = a + b \cdot X \quad II) Y^* = c + d \cdot V$$

con V variabile dummy che ricodifica le categorie di W nel modo seguente

$$v_i = \begin{cases} 0 & \text{se } w_i = b \\ 1 & \text{se } w_i = a \end{cases}$$



Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.

Si riportano i calcoli necessari¹ per ottenere le stime dei parametri e gli indici di adattamento del modello

$$Y = a + bX + E$$

x_i	y_i	$x_i y_i$	x_i^2	y_i^2
2	41	82	4	1681
2	45	90	4	2025
3	45	135	9	2025
4	45	180	16	2025
4	49	196	16	2401
15	225	683	49	10157
3	45	136.6	9.8	2031.4
$M(X)$	$M(Y)$	$M(XY)$	$M(X^2)$	$M(Y^2)$

$$\hat{b} =$$

¹Si osserva che per ragioni di stampa alcuni calcoli potrebbero essere duplicati con riferimento ai diversi modelli oggetto di studio



Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.

Si riportano i calcoli necessari¹ per ottenere le stime dei parametri e gli indici di adattamento del modello

$$Y = a + bX + E$$

x_i	y_i	$x_i y_i$	x_i^2	y_i^2
2	41	82	4	1681
2	45	90	4	2025
3	45	135	9	2025
4	45	180	16	2025
4	49	196	16	2401
15	225	683	49	10157
3	45	136.6	9.8	2031.4
$M(X)$	$M(Y)$	$M(XY)$	$M(X^2)$	$M(Y^2)$

$$\begin{aligned}\hat{b} &= \frac{\text{Cov}(X, Y)}{\text{Var}(X)} = \frac{M(XY) - M(X)M(Y)}{M(X^2) - [M(X)]^2} = \\ &= \frac{136.6 - 3 \cdot 45}{9.8 - 3^2} = \frac{1.6}{0.8} = 2 \\ \hat{a} &= \end{aligned}$$

¹Si osserva che per ragioni di stampa alcuni calcoli potrebbero essere duplicati con riferimento ai diversi modelli oggetto di studio



Si riportano i calcoli necessari¹ per ottenere le stime dei parametri e gli indici di adattamento del modello

$$Y = a + bX + E$$

x_i	y_i	$x_i y_i$	x_i^2	y_i^2
2	41	82	4	1681
2	45	90	4	2025
3	45	135	9	2025
4	45	180	16	2025
4	49	196	16	2401
15	225	683	49	10157
3	45	136.6	9.8	2031.4
$M(X)$	$M(Y)$	$M(XY)$	$M(X^2)$	$M(Y^2)$

$$\hat{b} = \frac{\text{Cov}(X, Y)}{\text{Var}(X)} = \frac{M(XY) - M(X)M(Y)}{M(X^2) - [M(X)]^2} =$$

$$= \frac{136.6 - 3 \cdot 45}{9.8 - 3^2} = \frac{1.6}{0.8} = 2$$

$$\hat{a} = M(Y) - \hat{b}M(X) = 39$$

$$\rho^2 =$$

¹Si osserva che per ragioni di stampa alcuni calcoli potrebbero essere duplicati con riferimento ai diversi modelli oggetto di studio



Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.

Si riportano i calcoli necessari¹ per ottenere le stime dei parametri e gli indici di adattamento del modello

$$Y = a + bX + E$$

x_i	y_i	$x_i y_i$	x_i^2	y_i^2
2	41	82	4	1681
2	45	90	4	2025
3	45	135	9	2025
4	45	180	16	2025
4	49	196	16	2401
15	225	683	49	10157
3	45	136.6	9.8	2031.4
$M(X)$	$M(Y)$	$M(XY)$	$M(X^2)$	$M(Y^2)$

$$\begin{aligned}\hat{b} &= \frac{\text{Cov}(X, Y)}{\text{Var}(X)} = \frac{M(XY) - M(X)M(Y)}{M(X^2) - [M(X)]^2} = \\ &= \frac{136.6 - 3 \cdot 45}{9.8 - 3^2} = \frac{1.6}{0.8} = 2 \\ \hat{a} &= M(Y) - \hat{b}M(X) = 39 \\ \rho^2 &= \frac{[\text{Cov}(X, Y)]^2}{\text{Var}(X)\text{Var}(Y)} = \frac{1.6^2}{0.8 \cdot 6.4} = 0.5\end{aligned}$$

¹Si osserva che per ragioni di stampa alcuni calcoli potrebbero essere duplicati con riferimento ai diversi modelli oggetto di studio



Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.

Si riportano i calcoli necessari² per ottenere le stime dei parametri e gli indici di adattamento del modello

$$Y = c + dV + E$$

v_i	y_i	$v_i y_i$	v_i^2	y_i^2
1	41	41	1	1681
1	45	45	1	2025
0	45	0	0	2025
0	45	0	0	2025
1	49	49	1	2401
3	225	135	3	10157
0.6	45	27	0.6	2031.4
$M(V)$	$M(Y)$	$M(VY)$	$M(V^2)$	$M(Y^2)$

$$\hat{d} =$$

²Si osserva che per ragioni di stampa alcuni calcoli potrebbero essere duplicati con riferimento ai diversi modelli oggetto di studio



Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.

Si riportano i calcoli necessari² per ottenere le stime dei parametri e gli indici di adattamento del modello

$$Y = c + dV + E$$

v_i	y_i	$v_i y_i$	v_i^2	y_i^2
1	41	41	1	1681
1	45	45	1	2025
0	45	0	0	2025
0	45	0	0	2025
1	49	49	1	2401
3	225	135	3	10157
0.6	45	27	0.6	2031.4
$M(V)$	$M(Y)$	$M(VY)$	$M(V^2)$	$M(Y^2)$

$$\begin{aligned}\hat{d} &= \frac{\text{Cov}(V, Y)}{\text{Var}(V)} = \frac{M(VY) - M(V)M(Y)}{M(V^2) - [M(V)]^2} = \\ &= \frac{27 - 0.6 \cdot 45}{0.6 - 0.6^2} = \frac{0}{0.24} = 0 \\ \hat{c} &= \end{aligned}$$

²Si osserva che per ragioni di stampa alcuni calcoli potrebbero essere duplicati con riferimento ai diversi modelli oggetto di studio



Si riportano i calcoli necessari² per ottenere le stime dei parametri e gli indici di adattamento del modello

$$Y = c + dV + E$$

v_i	y_i	$v_i y_i$	v_i^2	y_i^2
1	41	41	1	1681
1	45	45	1	2025
0	45	0	0	2025
0	45	0	0	2025
1	49	49	1	2401
3	225	135	3	10157
0.6	45	27	0.6	2031.4
$M(V)$	$M(Y)$	$M(VY)$	$M(V^2)$	$M(Y^2)$

$$\begin{aligned}\hat{d} &= \frac{\text{Cov}(V, Y)}{\text{Var}(V)} = \frac{M(VY) - M(V)M(Y)}{M(V^2) - [M(V)]^2} = \\ &= \frac{27 - 0.6 \cdot 45}{0.6 - 0.6^2} = \frac{0}{0.24} = 0 \\ \hat{c} &= M(Y) - \hat{d}M(V) = 45 \\ \rho^2 &= \end{aligned}$$

²Si osserva che per ragioni di stampa alcuni calcoli potrebbero essere duplicati con riferimento ai diversi modelli oggetto di studio



Si riportano i calcoli necessari² per ottenere le stime dei parametri e gli indici di adattamento del modello

$$Y = c + dV + E$$

v_i	y_i	$v_i y_i$	v_i^2	y_i^2
1	41	41	1	1681
1	45	45	1	2025
0	45	0	0	2025
0	45	0	0	2025
1	49	49	1	2401
3	225	135	3	10157
0.6	45	27	0.6	2031.4
$M(V)$	$M(Y)$	$M(VY)$	$M(V^2)$	$M(Y^2)$

$$\hat{d} = \frac{\text{Cov}(V, Y)}{\text{Var}(V)} = \frac{M(VY) - M(V)M(Y)}{M(V^2) - [M(V)]^2} =$$

$$= \frac{27 - 0.6 \cdot 45}{0.6 - 0.6^2} = \frac{0}{0.24} = 0$$

$$\hat{c} = M(Y) - \hat{d}M(V) = 45$$

$$\rho^2 = \frac{[\text{Cov}(V, Y)]^2}{\text{Var}(V) \text{Var}(Y)} = \frac{0^2}{0.24 \cdot 6.4} = 0$$

²Si osserva che per ragioni di stampa alcuni calcoli potrebbero essere duplicati con riferimento ai diversi modelli oggetto di studio



$\hat{a} = 39$, $\hat{b} = 2$, indice adattamento $\rho^2 = 0.5$ $\hat{c} = 45$, $\hat{d} = 0$, indice adattamento $\rho^2 = 0$
Si sono confrontati gli indici di adattamento normalizzati in quanto i due modelli sono caratterizzati dalla stessa variabile dipendente e dallo stesso numero di regressori.

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.



3. Si indichi, motivando la risposta, quale modello presenta il migliore adattamento.

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.



3. Si indichi, motivando la risposta, quale modello presenta il migliore adattamento.

Il modello migliore è $Y^* = 39 + 2 \cdot X$ in quanto caratterizzato da indice di adattamento ρ^2 più elevato.

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.



4. Si indichi l'intervallo dei valori possibili per $\eta_{Y|X}^2$ e $\eta_{Y|W}^2$.

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.



Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.

4. Si indichi l'intervallo dei valori possibili per $\eta_{Y|X}^2$ e $\eta_{Y|W}^2$.

$\eta_{Y|X}^2 = \rho_{XY}^2 = 0.5$ in quanto le medie condizionate $M(Y|x_i)$ sono allineate. $\eta_{Y|W}^2 = \rho_{XV}^2 = 0$ in quanto il modello II) include solo variabili indicatrici tra i regressori, e risulta quindi equivalente alla funzione di regressione.



L'urna A contiene 4 elementi contrassegnati con i simboli 1, 2, 3, 4, l'urna B 3 elementi contrassegnati con 2, 4, 6.

1. Si paga una posta di 3€ per partecipare al seguente gioco. Si lancia una moneta con $P(\text{testa}) = 0.1$ e $P(\text{croce}) = 0.9$. Se esce testa si estrae un elemento dall'urna A altrimenti uno dall'urna B . Si vince una somma in € pari al simbolo che contrassegna l'elemento estratto. Si costruisca la variabile casuale guadagno del giocatore e si stabilisca se il gioco è equo.

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.



Lo spazio probabilistico è composto dai seguenti eventi elementari

$$\Omega = \{T1, T2, T3, T4, C2, C4, C6\},$$

a cui corrispondono le probabilità

$$\{0.1 \cdot 0.25 = 0.025, 0.025, 0.025, 0.025, 0.9 \cdot 0.3333 = 0.3, 0.3, 0.3\}.$$

La variabile casuale guadagno

$$\underline{g_i \quad P(G = g_i) \quad g_i P(G = g_i)}$$



Lo spazio probabilistico è composto dai seguenti eventi elementari

$$\Omega = \{T1, T2, T3, T4, C2, C4, C6\},$$

a cui corrispondono le probabilità

$$\{0.1 \cdot 0.25 = 0.025, 0.025, 0.025, 0.025, 0.9 \cdot 0.3333 = 0.3, 0.3, 0.3\}.$$

La variabile casuale guadagno

g_i	$P(G = g_i)$	$g_i P(G = g_i)$
-2	0.025	-0.050
-1	0.325	-0.325
0	0.025	0.000
1	0.325	0.325
3	0.300	0.900
	1.000	0.850

ha valore atteso pari a $0.85 \neq 0$. Il gioco non è equo ma favorevole al giocatore.



2. Si consideri il seguente esperimento. Si estrae un elemento dall'urna B e lo si inserisce nell'urna A ; si estrae poi un elemento dall'urna A (modificata). Si calcoli la probabilità che l'elemento estratto dall'urna A rechi un simbolo dispari.

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.



2. Si consideri il seguente esperimento. Si estrae un elemento dall'urna B e lo si inserisce nell'urna A ; si estrae poi un elemento dall'urna A (modificata). Si calcoli la probabilità che l'elemento estratto dall'urna A rechi un simbolo dispari.

Indicati con D_B e D_A gli eventi estrazione di un elemento con simbolo dispari dall'urna B e A modificata, occorre calcolare la seguente probabilità

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.



Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.

2. Si consideri il seguente esperimento. Si estrae un elemento dall'urna B e lo si inserisce nell'urna A ; si estrae poi un elemento dall'urna A (modificata). Si calcoli la probabilità che l'elemento estratto dall'urna A rechi un simbolo dispari.

Indicati con D_B e D_A gli eventi estrazione di un elemento con simbolo dispari dall'urna B e A modificata, occorre calcolare la seguente probabilità

$$P(D_B) \cdot P(D_A|D_B) + P(\bar{D}_B) \cdot P(D_A|\bar{D}_B) =$$



Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.

2. Si consideri il seguente esperimento. Si estrae un elemento dall'urna B e lo si inserisce nell'urna A ; si estrae poi un elemento dall'urna A (modificata). Si calcoli la probabilità che l'elemento estratto dall'urna A rechi un simbolo dispari.

Indicati con D_B e D_A gli eventi estrazione di un elemento con simbolo dispari dall'urna B e A modificata, occorre calcolare la seguente probabilità

$$P(D_B) \cdot P(D_A|D_B) + P(\bar{D}_B) \cdot P(D_A|\bar{D}_B) = 0 \cdot \frac{2}{4} + 1 \cdot \frac{2}{5} = 0.4.$$



3. Si calcoli la probabilità che in 576 estrazioni con reimmissione dall'urna B ve ne siano almeno 181 con il simbolo 2.

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.



3. Si calcoli la probabilità che in 576 estrazioni con reimmissione dall'urna B ve ne siano almeno 181 con il simbolo 2.

Sia X la variabile casuale 'n. elementi con il simbolo 2 nelle 576 estrazioni con reimmissione dall'urna B '. Abbiamo $X \sim$

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.



Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.

3. Si calcoli la probabilità che in 576 estrazioni con reimmissione dall'urna B ve ne siano almeno 181 con il simbolo 2.

Sia X la variabile casuale 'n. elementi con il simbolo 2 nelle 576 estrazioni con reimmissione dall'urna B '. Abbiamo $X \sim \text{Bin}(n = 576, p = \frac{1}{3})$ con $E(X) =$



Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.

3. Si calcoli la probabilità che in 576 estrazioni con reimmissione dall'urna B ve ne siano almeno 181 con il simbolo 2.

Sia X la variabile casuale 'n. elementi con il simbolo 2 nelle 576 estrazioni con reimmissione dall'urna B '. Abbiamo $X \sim \text{Bin}(n = 576, p = \frac{1}{3})$ con $E(X) = n \cdot p = 576 \cdot \frac{1}{3} = 192$ e $\text{Var}(X) =$



3. Si calcoli la probabilità che in 576 estrazioni con reimmissione dall'urna B ve ne siano almeno 181 con il simbolo 2.

Sia X la variabile casuale 'n. elementi con il simbolo 2 nelle 576 estrazioni con reimmissione dall'urna B '. Abbiamo $X \sim \text{Bin}(n = 576, p = \frac{1}{3})$ con $E(X) = n \cdot p = 576 \cdot \frac{1}{3} = 192$ e $\text{Var}(X) = n \cdot p \cdot (1 - p) = 576 \cdot \frac{1}{3} \cdot \frac{2}{3} = 128$.

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.



3. Si calcoli la probabilità che in 576 estrazioni con reimmissione dall'urna B ve ne siano almeno 181 con il simbolo 2.

Sia X la variabile casuale 'n. elementi con il simbolo 2 nelle 576 estrazioni con reimmissione dall'urna B '. Abbiamo $X \sim \text{Bin}(n = 576, p = \frac{1}{3})$ con $E(X) = n \cdot p = 576 \cdot \frac{1}{3} = 192$ e $\text{Var}(X) = n \cdot p \cdot (1 - p) = 576 \cdot \frac{1}{3} \cdot \frac{2}{3} = 128$.

Dal momento che $np = 192 > 5$ ed $n(1 - p) = 384 > 5$ è possibile utilizzare l'approssimazione della variabile casuale Binomiale con la variabile casuale Normale $X \sim N(\mu =$



3. Si calcoli la probabilità che in 576 estrazioni con reimmissione dall'urna B ve ne siano almeno 181 con il simbolo 2.

Sia X la variabile casuale 'n. elementi con il simbolo 2 nelle 576 estrazioni con reimmissione dall'urna B '. Abbiamo $X \sim \text{Bin}(n = 576, p = \frac{1}{3})$ con $E(X) = n \cdot p = 576 \cdot \frac{1}{3} = 192$ e $\text{Var}(X) = n \cdot p \cdot (1 - p) = 576 \cdot \frac{1}{3} \cdot \frac{2}{3} = 128$.

Dal momento che $np = 192 > 5$ ed $n(1 - p) = 384 > 5$ è possibile utilizzare l'approssimazione della variabile casuale Binomiale con la variabile casuale Normale $X \sim N(\mu = 192, \sigma^2 =$



3. Si calcoli la probabilità che in 576 estrazioni con reimmissione dall'urna B ve ne siano almeno 181 con il simbolo 2.

Sia X la variabile casuale 'n. elementi con il simbolo 2 nelle 576 estrazioni con reimmissione dall'urna B '. Abbiamo $X \sim \text{Bin}(n = 576, p = \frac{1}{3})$ con $E(X) = n \cdot p = 576 \cdot \frac{1}{3} = 192$ e $\text{Var}(X) = n \cdot p \cdot (1 - p) = 576 \cdot \frac{1}{3} \cdot \frac{2}{3} = 128$.

Dal momento che $np = 192 > 5$ ed $n(1 - p) = 384 > 5$ è possibile utilizzare l'approssimazione della variabile casuale Binomiale con la variabile casuale Normale $X \sim N(\mu = 192, \sigma^2 = 128)$.

Si ottiene

$$P(X \geq 181) =$$



3. Si calcoli la probabilità che in 576 estrazioni con reimmissione dall'urna B ve ne siano almeno 181 con il simbolo 2.

Sia X la variabile casuale 'n. elementi con il simbolo 2 nelle 576 estrazioni con reimmissione dall'urna B '. Abbiamo $X \sim \text{Bin}(n = 576, p = \frac{1}{3})$ con $E(X) = n \cdot p = 576 \cdot \frac{1}{3} = 192$ e $\text{Var}(X) = n \cdot p \cdot (1 - p) = 576 \cdot \frac{1}{3} \cdot \frac{2}{3} = 128$.

Dal momento che $np = 192 > 5$ ed $n(1 - p) = 384 > 5$ è possibile utilizzare l'approssimazione della variabile casuale Binomiale con la variabile casuale Normale $X \sim N(\mu = 192, \sigma^2 = 128)$.

Si ottiene

$$P(X \geq 181) = P\left(\frac{X - \mu}{\sigma} \geq \frac{181 - \mu}{\sigma}\right) =$$



3. Si calcoli la probabilità che in 576 estrazioni con reimmissione dall'urna B ve ne siano almeno 181 con il simbolo 2.

Sia X la variabile casuale 'n. elementi con il simbolo 2 nelle 576 estrazioni con reimmissione dall'urna B '. Abbiamo $X \sim \text{Bin}(n = 576, p = \frac{1}{3})$ con $E(X) = n \cdot p = 576 \cdot \frac{1}{3} = 192$ e $\text{Var}(X) = n \cdot p \cdot (1 - p) = 576 \cdot \frac{1}{3} \cdot \frac{2}{3} = 128$.

Dal momento che $np = 192 > 5$ ed $n(1 - p) = 384 > 5$ è possibile utilizzare l'approssimazione della variabile casuale Binomiale con la variabile casuale Normale $X \sim N(\mu = 192, \sigma^2 = 128)$.

Si ottiene

$$P(X \geq 181) = P\left(\frac{X - \mu}{\sigma} \geq \frac{181 - \mu}{\sigma}\right) = P\left(Z \geq \frac{181 - 192}{11.3137}\right) =$$

=



3. Si calcoli la probabilità che in 576 estrazioni con reimmissione dall'urna B ve ne siano almeno 181 con il simbolo 2.

Sia X la variabile casuale 'n. elementi con il simbolo 2 nelle 576 estrazioni con reimmissione dall'urna B '. Abbiamo $X \sim \text{Bin}(n = 576, p = \frac{1}{3})$ con $E(X) = n \cdot p = 576 \cdot \frac{1}{3} = 192$ e $\text{Var}(X) = n \cdot p \cdot (1 - p) = 576 \cdot \frac{1}{3} \cdot \frac{2}{3} = 128$.

Dal momento che $np = 192 > 5$ ed $n(1 - p) = 384 > 5$ è possibile utilizzare l'approssimazione della variabile casuale Binomiale con la variabile casuale Normale $X \sim N(\mu = 192, \sigma^2 = 128)$.

Si ottiene

$$\begin{aligned} P(X \geq 181) &= P\left(\frac{X - \mu}{\sigma} \geq \frac{181 - \mu}{\sigma}\right) = P\left(Z \geq \frac{181 - 192}{11.3137}\right) = \\ &= P(Z \geq -0.97) = \end{aligned}$$



3. Si calcoli la probabilità che in 576 estrazioni con reimmissione dall'urna B ve ne siano almeno 181 con il simbolo 2.

Sia X la variabile casuale 'n. elementi con il simbolo 2 nelle 576 estrazioni con reimmissione dall'urna B '. Abbiamo $X \sim \text{Bin}(n = 576, p = \frac{1}{3})$ con $E(X) = n \cdot p = 576 \cdot \frac{1}{3} = 192$ e $\text{Var}(X) = n \cdot p \cdot (1 - p) = 576 \cdot \frac{1}{3} \cdot \frac{2}{3} = 128$.

Dal momento che $np = 192 > 5$ ed $n(1 - p) = 384 > 5$ è possibile utilizzare l'approssimazione della variabile casuale Binomiale con la variabile casuale Normale $X \sim N(\mu = 192, \sigma^2 = 128)$.

Si ottiene

$$\begin{aligned} P(X \geq 181) &= P\left(\frac{X - \mu}{\sigma} \geq \frac{181 - \mu}{\sigma}\right) = P\left(Z \geq \frac{181 - 192}{11.3137}\right) = \\ &= P(Z \geq -0.97) = P(Z \leq 0.97) = \end{aligned}$$



3. Si calcoli la probabilità che in 576 estrazioni con reimmissione dall'urna B ve ne siano almeno 181 con il simbolo 2.

Sia X la variabile casuale 'n. elementi con il simbolo 2 nelle 576 estrazioni con reimmissione dall'urna B '. Abbiamo $X \sim \text{Bin}(n = 576, p = \frac{1}{3})$ con $E(X) = n \cdot p = 576 \cdot \frac{1}{3} = 192$ e $\text{Var}(X) = n \cdot p \cdot (1 - p) = 576 \cdot \frac{1}{3} \cdot \frac{2}{3} = 128$.

Dal momento che $np = 192 > 5$ ed $n(1 - p) = 384 > 5$ è possibile utilizzare l'approssimazione della variabile casuale Binomiale con la variabile casuale Normale $X \sim N(\mu = 192, \sigma^2 = 128)$.

Si ottiene

$$\begin{aligned} P(X \geq 181) &= P\left(\frac{X - \mu}{\sigma} \geq \frac{181 - \mu}{\sigma}\right) = P\left(Z \geq \frac{181 - 192}{11.3137}\right) = \\ &= P(Z \geq -0.97) = P(Z \leq 0.97) = 0.8340. \end{aligned}$$

Tema T298

Prova scritta di STATISTICA del 20.02.2018

Prova scritta di
STATISTICA del
20.02.2018

Giuseppe Boari
Gabriele Cantaluppi



UNIVERSITÀ
CATTOLICA
del Sacro Cuore

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.

Giuseppe Boari
Gabriele Cantaluppi
Università Cattolica del Sacro Cuore, Milano



Data la seguente tabella a doppia entrata che riporta la distribuzione congiunta dei caratteri $X = \text{'genere'}$ e $Y = \text{'numero di visite mediche specialistiche fatte in un anno'}$ di 9 persone:

$X \setminus Y$	0 + 2	2 + 3	3 + 8
<i>maschio</i>	0	0	5
<i>femmina</i>	2	2	0

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.



1. Rappresentare graficamente le distribuzioni marginali di X e di Y .

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.



Si considera la distribuzione marginale della variabile X , ottenuta sommando le frequenze di riga

x_i	$n_{i\bullet}$
<i>maschio</i>	5
<i>femmina</i>	4

Si tratta di un carattere qualitativo sconnesso che può essere rappresentato graficamente mediante un grafico a bastoncini o a canne d'organo.

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.

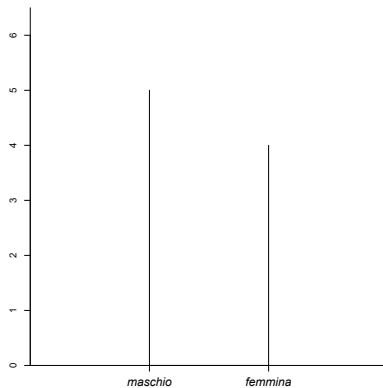


Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.





Si considera la distribuzione marginale della variabile Y , ottenuta sommando le frequenze di colonna

y_j	$n_{\bullet j}$
0 + 2	2
2 + 3	2
3 + 8	5

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.



Si tratta di un carattere quantitativo con modalità raggruppate in classi che può essere rappresentato graficamente mediante un istogramma.

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.



L'istogramma può essere costruito dopo avere calcolato le densità $d_{\bullet j} = \frac{n_{\bullet j}}{a_j}$, (dove a_j sono le ampiezze di classe).

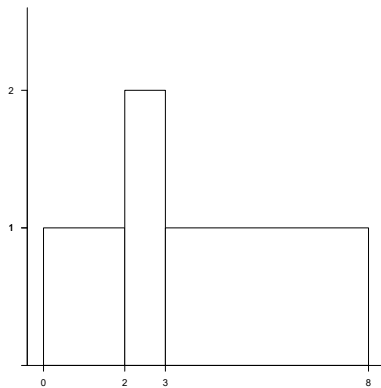
y_j	$n_{\bullet j}$	$a_{\bullet j}$	$d_{\bullet j}$
0 - 2	2	$2 - 0 = 2$	1
2 - 3	2	$3 - 2 = 1$	2
3 - 8	5	$8 - 3 = 5$	1

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.



Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.



Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.

2. Calcolare media aritmetica e media geometrica della marginale Y .



Si tratta di un carattere quantitativo con modalità raggruppate in classi.

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.



La media aritmetica

$$M(Y) = \frac{1}{n} \sum_{j=1}^k y_j n_{\bullet j}$$

risulta 3.8333.

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.



La media aritmetica

$$M(Y) = \frac{1}{n} \sum_{j=1}^k y_j n_{\bullet j}$$

risulta 3.8333.

Si ottiene (considerando come modalità i valori centrali delle classi)

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.



La media aritmetica

$$M(Y) = \frac{1}{n} \sum_{j=1}^k y_j n_{\bullet j}$$

risulta 3.8333.

Si ottiene (considerando come modalità i valori centrali delle classi)

y_j	$n_{\bullet j}$	$y_j n_{\bullet j}$
1.0	2	2.0
2.5	2	5.0
5.5	5	27.5
	9	34.5

da cui segue

$$M(Y) = \frac{1}{n} \sum_{i=1}^k y_j n_{\bullet j} = \frac{34.5}{9} = 3.8333.$$

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.



La media geometrica

$$\mu^{(0)}(Y) = \prod_{j=1}^k y_j^{f_j}$$

risulta 3.1604.

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.



La media geometrica

$$\mu^{(0)}(Y) = \prod_{j=1}^k y_j^{f_{\bullet j}}$$

risulta 3.1604. Si ottiene

y_j	$f_{\bullet j}$	$y_j^{f_{\bullet j}}$
1.0	0.2222	1.0000
2.5	0.2222	1.2258
5.5	0.5556	2.5782
		3.1604

da cui segue

$$\mu^{(0)} = \prod_{j=1}^k y_j^{f_{\bullet j}} = 3.1604.$$

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.



3. Calcolare la varianza within e la varianza between di Y considerando come gruppi quelli definiti dal genere (X) e commentare i risultati ottenuti.

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.



Si considera la distribuzione condizionata $Y|x_1$

y_j	n_{1j}
0 + 2	0
2 + 3	0
3 + 8	5

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.



Si tratta di un carattere quantitativo con modalità raggruppate in classi.

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.



Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.

Si osserva che la distribuzione è degenere.

$M(Y|x_1) = 5.5$. La varianza e il coefficiente di variazione sono nulli.



Si considera la distribuzione condizionata $Y|x_2$

y_j	n_{2j}
0 + 2	2
2 + 3	2
3 + 8	0

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.



Si tratta di un carattere quantitativo con modalità raggruppate in classi.

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.



La media aritmetica

$$M(Y|x_2) = \frac{1}{n_{2\bullet}} \sum_{j=1}^k y_j n_{2j}$$

risulta 1.75.

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.



La media aritmetica

$$M(Y|x_2) = \frac{1}{n_{2\bullet}} \sum_{j=1}^k y_j n_{2j}$$

risulta 1.75.

Si ottiene (considerando come modalità i valori centrali delle classi)

y_j	n_{2j}	$y_j n_{2j}$
1.0	2	2
2.5	2	5
5.5	0	0
	4	7

da cui segue

$$M(Y|x_2) = \frac{1}{n_{2\bullet}} \sum_{j=1}^k y_j n_{2j} = \frac{7}{4} = 1.75.$$

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.



La varianza

$$\text{Var}(Y|x_2) = M(Y^2|x_2) - [M(Y|x_2)]^2$$

risulta 0.5625.

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.



Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.

La varianza

$$\text{Var}(Y|x_2) = M(Y^2|x_2) - [M(Y|x_2)]^2$$

risulta 0.5625. Si ottiene (considerando come modalità i valori centrali delle classi)

y_j	n_{2j}	$y_j n_{2j}$	$y_j^2 n_{2j}$
1.0	2	2	2.0
2.5	2	5	12.5
5.5	0	0	0.0
	4	7	14.5

da cui segue

$$\text{Var}(Y|x_2) = \frac{14.5}{4} - 1.75^2 = 3.625 - 3.0625 = 0.5625.$$



Si considera, ora, la distribuzione marginale Y .

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.



Si tratta di un carattere quantitativo con modalità raggruppate in classi.

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.



La media aritmetica

$$M(Y) = \frac{1}{n} \sum_{j=1}^k y_j n_{\bullet j}$$

risulta 3.8333.

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.



La media aritmetica

$$M(Y) = \frac{1}{n} \sum_{j=1}^k y_j n_{\bullet j}$$

risulta 3.8333.

Si ottiene (considerando come modalità i valori centrali delle classi)

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.



La media aritmetica

$$M(Y) = \frac{1}{n} \sum_{j=1}^k y_j n_{\bullet j}$$

risulta 3.8333.

Si ottiene (considerando come modalità i valori centrali delle classi)

y_j	$n_{\bullet j}$	$y_j n_{\bullet j}$
1.0	2	2.0
2.5	2	5.0
5.5	5	27.5
	9	34.5

da cui segue

$$M(Y) = \frac{1}{n} \sum_{i=1}^k y_j n_{\bullet j} = \frac{34.5}{9} = 3.8333.$$

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.



La varianza

$$\text{Var}(Y) = M[(Y - \mu_Y)^2] = M(Y^2) - \mu_Y^2 = \frac{1}{n} \sum_{j=1}^k y_j^2 n_{\bullet j} - \mu_Y^2$$

risulta 3.7222.

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.



La varianza

$$Var(Y) = M[(Y - \mu_Y)^2] = M(Y^2) - \mu_Y^2 = \frac{1}{n} \sum_{j=1}^k y_j^2 n_{\bullet j} - \mu_Y^2$$

risulta 3.7222.

Si ottiene (considerando come modalità i valori centrali delle classi)

y_j	$n_{\bullet j}$	$y_j n_{\bullet j}$	$y_j^2 n_{\bullet j}$
1.0	2	2.0	2.00
2.5	2	5.0	12.50
5.5	5	27.5	151.25
	9	34.5	165.75

da cui segue

$$\begin{aligned} Var(Y) = \sigma_Y^2 &= M(Y^2) - \mu_Y^2 = \\ &= \frac{165.75}{9} - 3.8333^2 = 18.4167 - 14.6944 = 3.7223. \end{aligned}$$

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.



Il rapporto di correlazione

$$\eta_{Y|X}^2 = \frac{\text{Var}_X[M(Y|X)]}{\text{Var}(Y)}$$

può essere ottenuto considerando il rapporto tra la varianza spiegata $\text{Var}_X[M(Y|X)]$,
varianza della variabile statistica medie condizionate

i	$M(X y_i)$	$n_{i\bullet}$
1	5.50	5
2	1.75	4

e la varianza della variabile dipendente $\text{Var}(Y) = 3.7222$.

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.



La varianza spiegata risulta $Var_X[M(Y|X)] = 3.4722$.

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.



La varianza spiegata risulta $\text{Var}_X[M(Y|X)] = 3.4722$. Si ottiene

i	$M(X y_i)$	$n_{i\bullet}$	$M(X y_i) \cdot n_{i\bullet}$	$[M(X y_i)]^2 \cdot n_{i\bullet}$
1	5.50	5	27.5	151.25
2	1.75	4	7.0	12.25
		9	34.5	163.50

da cui segue

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.



La varianza spiegata risulta $Var_X[M(Y|X)] = 3.4722$. Si ottiene

i	$M(X y_i)$	$n_{i\bullet}$	$M(X y_i) \cdot n_{i\bullet}$	$[M(X y_i)]^2 \cdot n_{i\bullet}$
1	5.50	5	27.5	151.25
2	1.75	4	7.0	12.25
		9	34.5	163.50

da cui segue

$$Var_X[M(Y|X)] = \frac{163.5}{9} - 3.8333^2 = 18.1667 - 14.6944 = 3.4722.$$

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.



Il rapporto di correlazione $\eta^2_{Y|X}$ risulta:

$$\eta^2_{Y|X} = \frac{\text{Var}_X[M(Y|X)]}{\text{Var}(Y)} = \frac{3.4722}{3.7222} = 0.9328.$$

Sussiste, quindi, un livello di dipendenza molto alto.

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.



Si ricorda che, in base al risultato di scomposizione della varianza, la varianza della variabile dipendente può essere ottenuta come somma della varianza spiegata $Var_X[M(Y|X)]$ e della varianza residua $M_X[Var(Y|X)]$.

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.



La varianza residua $M_X[\text{Var}(Y|X)]$, media della variabile statistica varianze condizionate

i	$\text{Var}(X y_i)$	$n_{i\bullet}$
1	0.0000	5
2	0.5625	4

risulta $M_X[\text{Var}(Y|X)] = 0.25$.

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.



Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.

La varianza residua $M_X[Var(Y|X)]$, media della variabile statistica varianze condizionate

i	$Var(X y_i)$	$n_{i\bullet}$
1	0.0000	5
2	0.5625	4

risulta $M_X[Var(Y|X)] = 0.25$. Si ottiene

i	$Var(X y_i)$	$n_{i\bullet}$	$Var(X y_i) \cdot n_{i\bullet}$
1	0.0000	5	0.00
2	0.5625	4	2.25
		9	2.25

da cui segue



La varianza residua $M_X[Var(Y|X)]$, media della variabile statistica varianze condizionate

i	$Var(X y_i)$	$n_{i\bullet}$
1	0.0000	5
2	0.5625	4

risulta $M_X[Var(Y|X)] = 0.25$. Si ottiene

i	$Var(X y_i)$	$n_{i\bullet}$	$Var(X y_i) \cdot n_{i\bullet}$
1	0.0000	5	0.00
2	0.5625	4	2.25
		9	2.25

da cui segue

$$M_X[Var(Y|X)] = \frac{2.25}{9} = 0.25.$$



Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.

$$\begin{aligned}\sigma_{within}^2 &= M_X[Var(Y|X)] = 0.25, \\ \sigma_{between}^2 &= Var_X[M(Y|X)] = 3.4722 = \sigma_Y^2 - \sigma_{within}^2 = 3.7223 - 0.25 = 3.4723\end{aligned}$$



4. Senza fare calcoli e giustificando la risposta, si indichi il valore dell'indice χ^2 normalizzato.

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.



Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.

4. Senza fare calcoli e giustificando la risposta, si indichi il valore dell'indice χ^2 normalizzato.
 $\chi_N^2 = 1$ in quanto sussiste dipendenza funzionale di X da Y

Esercizio 2



Il prospetto seguente riporta i prezzi medi annui di 2 beni di uso quotidiano nel periodo 2012-2016:

t	$bene A$	$bene B$
2012	9	77
2013	11	65
2014	12	47
2015	14	35
2016	15	33

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.

1. Costruire la serie dei numeri indici a base fissa dei prezzi con base 2013 per il bene A e la serie dei numeri indici a base mobile dei prezzi per il bene B .



Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.



1. Costruire la serie dei numeri indici a base fissa dei prezzi con base 2013 per il bene A e la serie dei numeri indici a base mobile dei prezzi per il bene B .

Si ricorda che ${}_{2013}I_t = \frac{x_t}{x_{2013}}$.

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.



Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.

1. Costruire la serie dei numeri indici a base fissa dei prezzi con base 2013 per il bene *A* e la serie dei numeri indici a base mobile dei prezzi per il bene *B*.

Si ricorda che ${}_{2013}I_t = \frac{x_t}{x_{2013}}$. Si ottiene quindi

Anno	x_t	NIBF (2013)
2012	9	0.8182
2013	11	1.0000
2014	12	1.0909
2015	14	1.2727
2016	15	1.3636



Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.

1. Costruire la serie dei numeri indici a base fissa dei prezzi con base 2013 per il bene A e la serie dei numeri indici a base mobile dei prezzi per il bene B .

Si ricorda che ${}_{2013}I_t = \frac{x_t}{x_{2013}}$. Si ottiene quindi

Anno	x_t	NIBF (2013)
2012	9	0.8182
2013	11	1.0000
2014	12	1.0909
2015	14	1.2727
2016	15	1.3636

Si ricorda che ${}_{t-1}I_t = \frac{x_t}{x_{t-1}}$.



Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.

1. Costruire la serie dei numeri indici a base fissa dei prezzi con base 2013 per il bene A e la serie dei numeri indici a base mobile dei prezzi per il bene B .

Si ricorda che ${}_{2013}I_t = \frac{x_t}{x_{2013}}$. Si ottiene quindi

Anno	x_t	NIBF (2013)
2012	9	0.8182
2013	11	1.0000
2014	12	1.0909
2015	14	1.2727
2016	15	1.3636

Si ricorda che ${}_{t-1}I_t = \frac{x_t}{x_{t-1}}$. Si ottiene quindi

Anno	x_t	NIBM
2012	77	
2013	65	0.8442
2014	47	0.7231
2015	35	0.7447
2016	33	0.9429



2. Calcolare il tasso medio di variazione per il prezzo del bene B tra il 2013 e il 2015, interpretando il risultato.

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.



2. Calcolare il tasso medio di variazione per il prezzo del bene B tra il 2013 e il 2015, interpretando il risultato.

Si ricorda che

$${}_{2013}I_{2015} = {}_{2013}I_{2014} \cdot {}_{2014}I_{2015}$$

Si deve trovare il numero indice a base mobile medio.

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.



Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.

2. Calcolare il tasso medio di variazione per il prezzo del bene B tra il 2013 e il 2015, interpretando il risultato.

Si ricorda che

$${}_{2013}I_{2015} = {}_{2013}I_{2014} \cdot {}_{2014}I_{2015}$$

Si deve trovare il numero indice a base mobile medio. Applicando il criterio di scelta della media secondo Chisini si ottiene

$${}_{2013}I_{2015} = {}_{2013}I_{2014} \cdot {}_{2014}I_{2015} = \alpha \cdot \alpha = \alpha^2.$$



Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.

2. Calcolare il tasso medio di variazione per il prezzo del bene B tra il 2013 e il 2015, interpretando il risultato.

Si ricorda che

$${}_{2013}I_{2015} = {}_{2013}I_{2014} \cdot {}_{2014}I_{2015}$$

Si deve trovare il numero indice a base mobile medio. Applicando il criterio di scelta della media secondo Chisini si ottiene

$${}_{2013}I_{2015} = {}_{2013}I_{2014} \cdot {}_{2014}I_{2015} = \alpha \cdot \alpha = \alpha^2.$$

Quindi

$$\alpha = \sqrt[2]{{}_{2013}I_{2015}} = \sqrt[2]{{}_{2013}I_{2014} \cdot {}_{2014}I_{2015}} = 0.7338.$$

Da cui segue una variazione relativa media pari a -0.2662 , corrispondente a un decremento medio annuo tra il 2013 e il 2015 del 26.62%.



Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.

3. Note le quantità ${}_Aq_{2013} = 150$ e ${}_Bq_{2013} = 850$, calcolare l'indice di Laspeyres dei prezzi al 2016 con base 2013.



L'indice dei prezzi di Laspeyres si ottiene considerando la media aritmetica dei 'prezzi relativi' riferiti ai singoli prodotti $x_i = \frac{i p_{2016}}{i p_{2013}}$ calcolata utilizzando come frequenze i valori

$i v_{2013} = i p_{2013} \cdot i q_{2013}$ scambiati al tempo 2013

$${}^p_{2013} I^L_{2016} = \frac{1}{\sum_{i=1}^h i p_{2013} i q_{2013}} \sum_{i=1}^h \frac{i p_{2016}}{i p_{2013}} i p_{2013} i q_{2013}$$

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.



Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.

$\frac{i p_{2016}}{i p_{2013}}$	$i p_{2013} \cdot i q_{2013}$	$i p_{2016} \cdot i q_{2013}$
$\frac{15}{11} = 1.3636$	$11 \cdot 150 = 1650$	2250
$\frac{33}{65} = 0.5077$	$65 \cdot 850 = 55250$	28050
	56900	30300

$$p_{2013} I_{2016}^L = \frac{1}{56900} 30300 = 0.5325.$$

La variazione relativa nel livello generale dei prezzi, secondo l'Indice di Laspeyres, è del -46.75% .



Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.

Il seguente prospetto riporta i valori rilevati in corrispondenza a 7 unità statistiche con riferimento a due caratteri quantitativi X e Y per i quali si vuole studiare la dipendenza di Y da X :

x_i	1	1	3	3	5	5	10
y_i	10	10	21	21	30	30	50

1. Si rappresentino graficamente le coppie dei dati e la funzione di regressione di Y in funzione di X .



La funzione di regressione associa a ogni valore x appartenente all'insieme di definizione della variabile X la media di Y condizionata a x , $M(Y|x)$

x	$M(Y x)$
1	10
3	21
5	30
10	50

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.



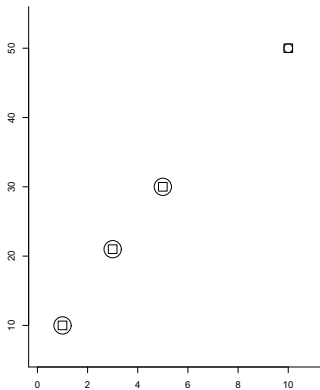
Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.

Le coppie di dati possono essere rappresentate mediante un *bubble plot diagram*.





2. Si calcolino, secondo il metodo dei minimi quadrati, i parametri dei seguenti modelli di regressione:

$$I) Y^* = a + b \cdot X \quad II) Y^* = c \cdot \ln(X)$$

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.

Si riportano i calcoli necessari¹ per ottenere le stime dei parametri e gli indici di adattamento del modello

$$Y = a + bX + E$$

x_i	y_i	$x_i y_i$	x_i^2	y_i^2
1	10	10	1	100
1	10	10	1	100
3	21	63	9	441
3	21	63	9	441
5	30	150	25	900
5	30	150	25	900
10	50	500	100	2500
28	172	946	170	5382
4	24.5714	135.1429	24.2857	768.8571
$M(X)$	$M(Y)$	$M(XY)$	$M(X^2)$	$M(Y^2)$

$$\hat{b} =$$



Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.

¹Si osserva che per ragioni di stampa alcuni calcoli potrebbero essere duplicati con riferimento ai diversi modelli oggetto di studio

Si riportano i calcoli necessari¹ per ottenere le stime dei parametri e gli indici di adattamento del modello

$$Y = a + bX + E$$

x_i	y_i	$x_i y_i$	x_i^2	y_i^2
1	10	10	1	100
1	10	10	1	100
3	21	63	9	441
3	21	63	9	441
5	30	150	25	900
5	30	150	25	900
10	50	500	100	2500
28	172	946	170	5382
4	24.5714	135.1429	24.2857	768.8571
$M(X)$	$M(Y)$	$M(XY)$	$M(X^2)$	$M(Y^2)$

$$\begin{aligned}\hat{b} &= \frac{\text{Cov}(X, Y)}{\text{Var}(X)} = \frac{M(XY) - M(X)M(Y)}{M(X^2) - [M(X)]^2} = \\ &= \frac{135.1429 - 4 \cdot 24.5714}{24.2857 - 4^2} = \frac{36.8571}{8.2857} = 4.4483\end{aligned}$$

$$\hat{a} =$$



Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.

¹Si osserva che per ragioni di stampa alcuni calcoli potrebbero essere duplicati con riferimento ai diversi modelli oggetto di studio

Si riportano i calcoli necessari¹ per ottenere le stime dei parametri e gli indici di adattamento del modello

$$Y = a + bX + E$$

x_i	y_i	$x_i y_i$	x_i^2	y_i^2
1	10	10	1	100
1	10	10	1	100
3	21	63	9	441
3	21	63	9	441
5	30	150	25	900
5	30	150	25	900
10	50	500	100	2500
28	172	946	170	5382
4	24.5714	135.1429	24.2857	768.8571
$M(X)$	$M(Y)$	$M(XY)$	$M(X^2)$	$M(Y^2)$

$$\begin{aligned}\hat{b} &= \frac{\text{Cov}(X, Y)}{\text{Var}(X)} = \frac{M(XY) - M(X)M(Y)}{M(X^2) - [M(X)]^2} = \\ &= \frac{135.1429 - 4 \cdot 24.5714}{24.2857 - 4^2} = \frac{36.8571}{8.2857} = 4.4483\end{aligned}$$

$$\hat{a} = M(Y) - \hat{b} M(X) = 6.7783$$

$$\rho^2 =$$

¹Si osserva che per ragioni di stampa alcuni calcoli potrebbero essere duplicati con riferimento ai diversi modelli oggetto di studio



Si riportano i calcoli necessari¹ per ottenere le stime dei parametri e gli indici di adattamento del modello

$$Y = a + bX + E$$

x_i	y_i	$x_i y_i$	x_i^2	y_i^2
1	10	10	1	100
1	10	10	1	100
3	21	63	9	441
3	21	63	9	441
5	30	150	25	900
5	30	150	25	900
10	50	500	100	2500
28	172	946	170	5382
4	24.5714	135.1429	24.2857	768.8571
$M(X)$	$M(Y)$	$M(XY)$	$M(X^2)$	$M(Y^2)$

$$\begin{aligned}\hat{b} &= \frac{\text{Cov}(X, Y)}{\text{Var}(X)} = \frac{M(XY) - M(X)M(Y)}{M(X^2) - [M(X)]^2} = \\ &= \frac{135.1429 - 4 \cdot 24.5714}{24.2857 - 4^2} = \frac{36.8571}{8.2857} = 4.4483\end{aligned}$$

$$\hat{a} = M(Y) - \hat{b} M(X) = 6.7783$$

$$\rho^2 = \frac{[\text{Cov}(X, Y)]^2}{\text{Var}(X) \text{Var}(Y)} = \frac{36.8571^2}{8.2857 \cdot 165.102} = 0.993$$

$$M(\hat{E}^2) = \text{Var}(\hat{E}) = \text{varianza residua} =$$

¹Si osserva che per ragioni di stampa alcuni calcoli potrebbero essere duplicati con riferimento ai diversi modelli oggetto di studio



Si riportano i calcoli necessari¹ per ottenere le stime dei parametri e gli indici di adattamento del modello

$$Y = a + bX + E$$

x_i	y_i	$x_i y_i$	x_i^2	y_i^2
1	10	10	1	100
1	10	10	1	100
3	21	63	9	441
3	21	63	9	441
5	30	150	25	900
5	30	150	25	900
10	50	500	100	2500
28	172	946	170	5382
4	24.5714	135.1429	24.2857	768.8571
$M(X)$	$M(Y)$	$M(XY)$	$M(X^2)$	$M(Y^2)$

$$\begin{aligned}\hat{b} &= \frac{\text{Cov}(X, Y)}{\text{Var}(X)} = \frac{M(XY) - M(X)M(Y)}{M(X^2) - [M(X)]^2} = \\ &= \frac{135.1429 - 4 \cdot 24.5714}{24.2857 - 4^2} = \frac{36.8571}{8.2857} = 4.4483\end{aligned}$$

$$\hat{a} = M(Y) - \hat{b} M(X) = 6.7783$$

$$\rho^2 = \frac{[\text{Cov}(X, Y)]^2}{\text{Var}(X) \text{Var}(Y)} = \frac{36.8571^2}{8.2857 \cdot 165.102} = 0.993$$

$$M(\hat{E}^2) = \text{Var}(\hat{E}) = \text{varianza residua} = \text{Var}(Y) \cdot (1 - \rho^2) = 1.151$$

¹Si osserva che per ragioni di stampa alcuni calcoli potrebbero essere duplicati con riferimento ai diversi modelli oggetto di studio





Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.

Si riportano i calcoli necessari² per ottenere le stime dei parametri e la misura di adattamento del modello

$$Y = c \ln(X) + E$$

$\ln(x)_i$	y_i	$\ln(x)_i y_i$	$\ln(x)_i^2$	y_i^2
0	10	0	0	100
0	10	0	0	100
1.0986	21	23.0706	1.2069	441
1.0986	21	23.0706	1.2069	441
1.6094	30	48.282	2.5902	900
1.6094	30	48.282	2.5902	900
2.3026	50	115.13	5.302	2500
		257.8352	12.8962	5382
		36.8336	1.8423	768.8571
		$M(\ln(X)Y)$	$M(\ln(X)^2)$	$M(Y^2)$

$\hat{c} =$

²Si osserva che per ragioni di stampa alcuni calcoli potrebbero essere duplicati con riferimento ai diversi modelli oggetto di studio



Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.

Si riportano i calcoli necessari² per ottenere le stime dei parametri e la misura di adattamento del modello

$$Y = c \ln(X) + E$$

$\ln(x)_i$	y_i	$\ln(x)_i y_i$	$\ln(x)_i^2$	y_i^2
0	10	0	0	100
0	10	0	0	100
1.0986	21	23.0706	1.2069	441
1.0986	21	23.0706	1.2069	441
1.6094	30	48.282	2.5902	900
1.6094	30	48.282	2.5902	900
2.3026	50	115.13	5.302	2500
		257.8352	12.8962	5382
		36.8336	1.8423	768.8571
		$M(\ln(X)Y)$	$M(\ln(X)^2)$	$M(Y^2)$

$$\hat{c} = \frac{M(\ln(X)Y)}{M(\ln(X)^2)} = \frac{36.8336}{1.8423} = 19.9933$$

$$M(\hat{E}^2) =$$

²Si osserva che per ragioni di stampa alcuni calcoli potrebbero essere duplicati con riferimento ai diversi modelli oggetto di studio



Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.

Si riportano i calcoli necessari² per ottenere le stime dei parametri e la misura di adattamento del modello

$$Y = c \ln(X) + E$$

$\ln(x)_i$	y_i	$\ln(x)_i y_i$	$\ln(x)_i^2$	y_i^2
0	10	0	0	100
0	10	0	0	100
1.0986	21	23.0706	1.2069	441
1.0986	21	23.0706	1.2069	441
1.6094	30	48.282	2.5902	900
1.6094	30	48.282	2.5902	900
2.3026	50	115.13	5.302	2500
		257.8352	12.8962	5382
		36.8336	1.8423	768.8571
		$M(\ln(X)Y)$	$M(\ln(X)^2)$	$M(Y^2)$

$$\hat{c} = \frac{M(\ln(X)Y)}{M(\ln(X)^2)} = \frac{36.8336}{1.8423} = 19.9933$$

$$M(\hat{E}^2) = M[Y \cdot (Y - \hat{c} \ln(X))] = M(Y^2) - \hat{c} M(\ln(X)Y) = 768.8571 - 19.9933 \cdot 36.8336 = 32.433$$

²Si osserva che per ragioni di stampa alcuni calcoli potrebbero essere duplicati con riferimento ai diversi modelli oggetto di studio



$\hat{a} = 6.7783$, $\hat{b} = 4.4483$, indice adattamento $M(\hat{E}^2) = 1.1513$ $\hat{c} = 19.9933$, indice adattamento $M(\hat{E}^2) = 32.433$

3. Si indichi quale modello presenta il migliore adattamento ai dati.

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.



$\hat{a} = 6.7783$, $\hat{b} = 4.4483$, indice adattamento $M(\hat{E}^2) = 1.1513$ $\hat{c} = 19.9933$, indice adattamento $M(\hat{E}^2) = 32.433$

3. Si indichi quale modello presenta il migliore adattamento ai dati.

Il modello migliore è $Y^* = 6.7783 + 4.4483 \cdot X$ in quanto caratterizzato da residuo quadratico medio più basso.

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.



Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.

4. Si dica, motivando teoricamente la risposta, se $\eta_{Y|X}^2 \neq \eta_{X|Y}^2$.



Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.

4. Si dica, motivando teoricamente la risposta, se $\eta_{Y|X}^2 \neq \eta_{X|Y}^2$.

$\eta_{Y|X}^2 = \eta_{X|Y}^2$ in quanto sussiste dipendenza funzionale bijectiva.

Esercizio 4

Prova scritta di
STATISTICA del
20.02.2018

Giuseppe Boari
Gabriele Cantaluppi



UNIVERSITÀ
CATTOLICA
del Sacro Cuore

Un'urna è composta da 4 gettoni gialli e 16 gettoni blu. Si consideri l'esperimento casuale 'estrazione con reimmissione di 3 gettoni dall'urna'.

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.



Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.

1. Si costruisca lo spazio probabilistico Ω associato all'esperimento.



1. Si costruisca lo spazio probabilistico Ω associato all'esperimento.

Indicati con G_i e B_i gli eventi estrazione di gettone giallo e blu nella i -esima estrazione ($i = 1, 2, 3$), lo spazio probabilistico risulta composto dai seguenti eventi elementari:

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.



Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.

1. Si costruisca lo spazio probabilistico Ω associato all'esperimento.

Indicati con G_i e B_i gli eventi estrazione di gettone giallo e blu nella i -esima estrazione ($i = 1, 2, 3$), lo spazio probabilistico risulta composto dai seguenti eventi elementari:

$$\Omega = \left\{ \begin{array}{ccc} G_1 G_2 G_3, & G_1 G_2 G_3, & B_1 G_2 G_3, \\ G_1 G_2 B_3, & G_1 B_2 G_3, & B_1 G_2 G_3, \\ G_1 B_2 B_3, & B_1 G_2 B_3, & B_1 B_2 G_3, \\ & B_1 B_2 B_3 & \end{array} \right\},$$

a cui corrispondono le probabilità



1. Si costruisca lo spazio probabilistico Ω associato all'esperimento.

Indicati con G_i e B_i gli eventi estrazione di gettone giallo e blu nella i -esima estrazione ($i = 1, 2, 3$), lo spazio probabilistico risulta composto dai seguenti eventi elementari:

$$\Omega = \left\{ \begin{array}{ccc} G_1 G_2 G_3, & & \\ G_1 G_2 B_3, & G_1 B_2 G_3, & B_1 G_2 G_3, \\ G_1 B_2 B_3, & B_1 G_2 B_3, & B_1 B_2 G_3, \\ & B_1 B_2 B_3 & \end{array} \right\},$$

a cui corrispondono le probabilità

$$\left\{ \begin{array}{ccc} & 0.2^3 = 0.008, & \\ 0.2^2 \cdot 0.8 = 0.032, & 0.032, & 0.032, \\ 0.2 \cdot 0.8^2 = 0.128, & 0.128, & 0.128, \\ & 0.8^3 = 0.512 & \end{array} \right\}$$

tenendo conto che le estrazioni sono con reimmissione (prove indipendenti) e che

$$P(G_i) = P(G) = 0.2, \quad P(B_i) = P(B) = 0.8, \quad i = 1, 2, 3.$$

2. Sapendo che ogni gettone giallo reca impresso il numero 2, ogni gettone blu il numero 3, si costruisca la variabile casuale $X =$ 'somma dei valori impressi sui gettoni estratti' e si calcoli la probabilità che X assuma valori maggiori o uguali a 8.



Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.



Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.

2. Sapendo che ogni gettone giallo reca impresso il numero 2, ogni gettone blu il numero 3, si costruisca la variabile casuale $X =$ 'somma dei valori impressi sui gettoni estratti' e si calcoli la probabilità che X assuma valori maggiori o uguali a 8.

Gli eventi elementari di Ω hanno le seguenti immagini in $\mathcal{R}$



2. Sapendo che ogni gettone giallo reca impresso il numero 2, ogni gettone blu il numero 3, si costruisca la variabile casuale X = 'somma dei valori impressi sui gettoni estratti' e si calcoli la probabilità che X assuma valori maggiori o uguali a 8.

Gli eventi elementari di Ω hanno le seguenti immagini in $\mathcal{R}$

$$\begin{aligned} X(G_1 G_2 G_3) &= 6, \\ X(G_1 G_2 B_3) &= 7, & X(G_1 B_2 G_3) &= 7, & X(B_1 G_2 G_3) &= 7, \\ X(G_1 B_2 B_3) &= 8, & X(B_1 G_2 B_3) &= 8, & X(B_1 B_2 G_3) &= 8, \\ X(B_1 B_2 B_3) &= 9. \end{aligned}$$

La variabile casuale X è caratterizzata dalla seguente distribuzione di probabilità



2. Sapendo che ogni gettone giallo reca impresso il numero 2, ogni gettone blu il numero 3, si costruisca la variabile casuale X = 'somma dei valori impressi sui gettoni estratti' e si calcoli la probabilità che X assuma valori maggiori o uguali a 8.

Gli eventi elementari di Ω hanno le seguenti immagini in $\mathcal{R}$

$$\begin{aligned} X(G_1 G_2 G_3) &= 6, \\ X(G_1 G_2 B_3) &= 7, & X(G_1 B_2 G_3) &= 7, & X(B_1 G_2 G_3) &= 7, \\ X(G_1 B_2 B_3) &= 8, & X(B_1 G_2 B_3) &= 8, & X(B_1 B_2 G_3) &= 8, \\ X(B_1 B_2 B_3) &= 9. \end{aligned}$$

La variabile casuale X è caratterizzata dalla seguente distribuzione di probabilità

x	$P(X = x)$
6	0.008
7	0.096
8	0.384
9	0.512
	1.000

Abbiamo, infine, $P(X \geq 8) =$



2. Sapendo che ogni gettone giallo reca impresso il numero 2, ogni gettone blu il numero 3, si costruisca la variabile casuale X = 'somma dei valori impressi sui gettoni estratti' e si calcoli la probabilità che X assuma valori maggiori o uguali a 8.

Gli eventi elementari di Ω hanno le seguenti immagini in $\mathcal{R}$

$$\begin{aligned} X(G_1 G_2 G_3) &= 6, \\ X(G_1 G_2 B_3) &= 7, & X(G_1 B_2 G_3) &= 7, & X(B_1 G_2 G_3) &= 7, \\ X(G_1 B_2 B_3) &= 8, & X(B_1 G_2 B_3) &= 8, & X(B_1 B_2 G_3) &= 8, \\ X(B_1 B_2 B_3) &= 9. \end{aligned}$$

La variabile casuale X è caratterizzata dalla seguente distribuzione di probabilità

x	$P(X = x)$
6	0.008
7	0.096
8	0.384
9	0.512
	1.000

Abbiamo, infine, $P(X \geq 8) = P(X = 8) + P(X = 9) =$



2. Sapendo che ogni gettone giallo reca impresso il numero 2, ogni gettone blu il numero 3, si costruisca la variabile casuale X = 'somma dei valori impressi sui gettoni estratti' e si calcoli la probabilità che X assuma valori maggiori o uguali a 8.

Gli eventi elementari di Ω hanno le seguenti immagini in $\mathcal{R}$

$$\begin{aligned} X(G_1 G_2 G_3) &= 6, \\ X(G_1 G_2 B_3) &= 7, & X(G_1 B_2 G_3) &= 7, & X(B_1 G_2 G_3) &= 7, \\ X(G_1 B_2 B_3) &= 8, & X(B_1 G_2 B_3) &= 8, & X(B_1 B_2 G_3) &= 8, \\ X(B_1 B_2 B_3) &= 9. \end{aligned}$$

La variabile casuale X è caratterizzata dalla seguente distribuzione di probabilità

x	$P(X = x)$
6	0.008
7	0.096
8	0.384
9	0.512
	1.000

Abbiamo, infine, $P(X \geq 8) = P(X = 8) + P(X = 9) = 0.384 + 0.512 = 0.896$.



3. Facendo ora 4 estrazioni **senza reimmissione**, si calcoli la probabilità di ottenere almeno 1 gettone giallo.

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.



Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.

3. Facendo ora 4 estrazioni **senza reimmissione**, si calcoli la probabilità di ottenere almeno 1 gettone giallo.

$$P(\text{almeno 1 gettone giallo}) =$$



Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.

3. Facendo ora 4 estrazioni **senza reimmissione**, si calcoli la probabilità di ottenere almeno 1 gettone giallo.

$$P(\text{almeno 1 gettone giallo}) = 1 - P(\text{nessun gettone giallo}) =$$
$$=$$



Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.

3. Facendo ora 4 estrazioni **senza reimmissione**, si calcoli la probabilità di ottenere almeno 1 gettone giallo.

$$\begin{aligned} P(\text{almeno 1 gettone giallo}) &= 1 - P(\text{nessun gettone giallo}) = \\ &= 1 - P(B_1 \cap B_2 \cap B_3 \cap B_4) = \\ &= \end{aligned}$$



Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.

3. Facendo ora 4 estrazioni **senza reimmissione**, si calcoli la probabilità di ottenere almeno 1 gettone giallo.

$$\begin{aligned}P(\text{almeno 1 gettone giallo}) &= 1 - P(\text{nessun gettone giallo}) = \\&= 1 - P(B_1 \cap B_2 \cap B_3 \cap B_4) = \\&= 1 - P(B_1) \cdot P(B_2|B_1) \cdot P(B_3|B_1 \cap B_2) \cdot P(B_4|B_1 \cap B_2 \cap B_3) = \\&= \end{aligned}$$



Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.

3. Facendo ora 4 estrazioni **senza reimmissione**, si calcoli la probabilità di ottenere almeno 1 gettone giallo.

$$\begin{aligned}P(\text{almeno 1 gettone giallo}) &= 1 - P(\text{nessun gettone giallo}) = \\&= 1 - P(B_1 \cap B_2 \cap B_3 \cap B_4) = \\&= 1 - P(B_1) \cdot P(B_2|B_1) \cdot P(B_3|B_1 \cap B_2) \cdot P(B_4|B_1 \cap B_2 \cap B_3) = \\&= 1 - \frac{16}{20} \cdot \frac{15}{19} \cdot \frac{14}{18} \cdot \frac{13}{17} = \\&= \end{aligned}$$



3. Facendo ora 4 estrazioni **senza reimmissione**, si calcoli la probabilità di ottenere almeno 1 gettone giallo.

$$\begin{aligned}P(\text{almeno 1 gettone giallo}) &= 1 - P(\text{nessun gettone giallo}) = \\&= 1 - P(B_1 \cap B_2 \cap B_3 \cap B_4) = \\&= 1 - P(B_1) \cdot P(B_2|B_1) \cdot P(B_3|B_1 \cap B_2) \cdot P(B_4|B_1 \cap B_2 \cap B_3) = \\&= 1 - \frac{16}{20} \cdot \frac{15}{19} \cdot \frac{14}{18} \cdot \frac{13}{17} = \\&= 1 - 0.3756 = 0.6244.\end{aligned}$$

Tema T2991

Prova scritta di STATISTICA del 04.06.2018

Prova scritta di
STATISTICA del
04.06.2018

Giuseppe Boari
Gabriele Cantaluppi



UNIVERSITÀ
CATTOLICA
del Sacro Cuore

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.

Giuseppe Boari
Gabriele Cantaluppi
Università Cattolica del Sacro Cuore, Milano



Si sono rilevati la dimensione X (p = piccola; m = media; g = grande) e il fatturato annuo in centinaia di migliaia di euro Y di 1000 imprese. Si riportano la distribuzione delle frequenze relative di X e le densità di frequenza di Y .

x_i	f_i
p	0.60
m	0.21
g	0.19

y_i	d_i
5 $\rightarrow$ 10	90
10 $\rightarrow$ 20	21
20 $\rightarrow$ 40	17

1. Si costruiscano le distribuzioni delle frequenze assolute di X e di Y .

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.



Moltiplicando le frequenze relative f_i per il numero totale di osservazioni n si ottengono le frequenze assolute n_i della variabile X .

x_i	f_i	n_i
p	0.6	600
m	0.21	210
g	0.19	190

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.



Moltiplicando le densità di frequenza, d_i , per le ampiezze di classe, a_i , si ottengono le frequenze assolute n_i della variabile Y ,

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.



Moltiplicando le densità di frequenza, d_i , per le ampiezze di classe, a_i , si ottengono le frequenze assolute n_i della variabile Y ,

y_i	d_i	a_i	n_i
5 ÷ 10	90	$10 - 5 = 5$	450
10 ÷ 20	21	$20 - 10 = 10$	210
20 ÷ 40	17	$40 - 20 = 20$	340

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.



2. Si rappresentino graficamente le due variabili e si riassumano le stesse con opportuni indici di posizione e di variabilità.

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.



Si considera la distribuzione della variabile X ,

x_i	n_i
p	600
m	210
g	190

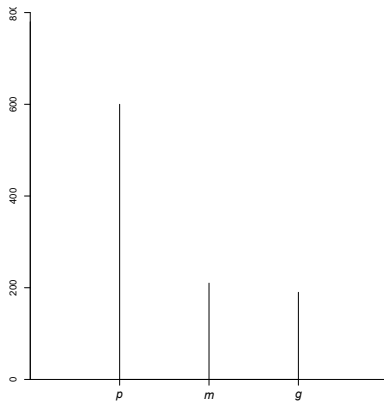
Si tratta di un carattere qualitativo sconnesso che può essere rappresentato graficamente mediante un grafico a bastoncini o a canne d'organo.

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.



Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.



Si considera la distribuzione della variabile X ,

x_i	n_i
p	600
m	210
g	190

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.



Si tratta di un carattere qualitativo sconnesso.

L'unico indice di posizione che può essere definito è la moda (categoria alla quale è associata la massima frequenza) che risulta p .

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.

In presenza di un carattere qualitativo la variabilità può essere misurata mediante gli indici di Gini e di Frosini normalizzati

$$G_N = \frac{1 - \sum_{i=1}^k f_i^2}{1 - \frac{1}{k}}$$

e

$$F_N = 1 - \sqrt{1 - G_N}.$$



Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.



In presenza di un carattere qualitativo la variabilità può essere misurata mediante gli indici di Gini e di Frosini normalizzati

$$G_N = \frac{1 - \sum_{i=1}^k f_i^2}{1 - \frac{1}{k}}$$

e

$$F_N = 1 - \sqrt{1 - G_N}.$$

Si ottiene:

x_i	n_i	f_i	f_i^2
p	600	0.6	0.36
m	210	0.21	0.0441
g	190	0.19	0.0361
	1000	1	0.4402

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.



In presenza di un carattere qualitativo la variabilità può essere misurata mediante gli indici di Gini e di Frosini normalizzati

$$G_N = \frac{1 - \sum_{i=1}^k f_i^2}{1 - \frac{1}{k}}$$

e

$$F_N = 1 - \sqrt{1 - G_N}.$$

Si ottiene:

x_i	n_i	f_i	f_i^2
p	600	0.6	0.36
m	210	0.21	0.0441
g	190	0.19	0.0361
	1000	1	0.4402

Da cui segue:

$$G_N = \frac{1 - 0.4402}{1 - \frac{1}{3}} = 0.8397$$

e

$$F_N = 1 - \sqrt{1 - 0.8397} = 0.5996.$$

Sussiste, quindi, un livello medio di eterogeneità.



Si considera la distribuzione della variabile X ,

x_i	n_i
5 $\vdash$ 10	450
10 $\vdash$ 20	210
20 $\vdash$ 40	340

Esercizio 1.

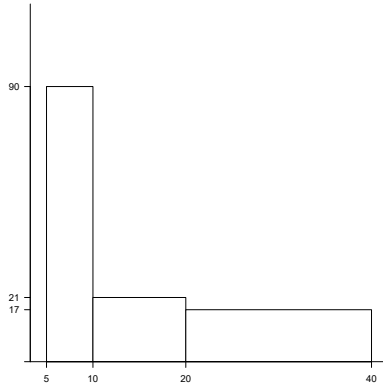
Esercizio 2.

Esercizio 3.

Esercizio 4.



Si tratta di un carattere quantitativo con modalità raggruppate in classi che può essere rappresentato graficamente mediante un istogramma.



Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.



Si considera la distribuzione della variabile Y ,

y_i	n_i
5 $\vdash$ 10	450
10 $\vdash$ 20	210
20 $\vdash$ 40	340

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.



Si tratta di un carattere quantitativo con modalità raggruppate in classi.

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.



La media aritmetica

$$M(Y) = \frac{1}{n} \sum_{i=1}^k y_i n_i$$

risulta 16.725.

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.



La media aritmetica

$$M(Y) = \frac{1}{n} \sum_{i=1}^k y_i n_i$$

risulta 16.725.

Si ottiene (considerando come modalità i valori centrali delle classi)

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.



La media aritmetica

$$M(Y) = \frac{1}{n} \sum_{i=1}^k y_i n_i$$

risulta 16.725.

Si ottiene (considerando come modalità i valori centrali delle classi)

y_i	n_i	$y_i n_i$
7.5	450	3375
15.0	210	3150
30.0	340	10200
	1000	16725

da cui segue

$$M(Y) = \frac{1}{n} \sum_{i=1}^k y_i n_i = \frac{16725}{1000} = 16.725.$$

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.



La varianza

$$\text{Var}(Y) = M[(Y - \mu_Y)^2] = M(Y^2) - \mu_Y^2 = \frac{1}{n} \sum_{i=1}^k y_i^2 n_i - \mu_Y^2$$

risulta 98.8369.

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.



La varianza

$$\text{Var}(Y) = M[(Y - \mu_Y)^2] = M(Y^2) - \mu_Y^2 = \frac{1}{n} \sum_{i=1}^k y_i^2 n_i - \mu_Y^2$$

risulta 98.8369.

Si ottiene (considerando come modalità i valori centrali delle classi)

y_i	n_i	$y_i n_i$	$y_i^2 n_i$
7.5	450	3375	25312.5
15.0	210	3150	47250.0
30.0	340	10200	306000.0
	1000	16725	378562.5

da cui segue

$$\begin{aligned} \text{Var}(Y) = \sigma_Y^2 &= M(Y^2) - \mu_Y^2 = \\ &= \frac{378562.5}{1000} - 16.725^2 = 378.5625 - 279.7256 = 98.8369. \end{aligned}$$

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.



Per effettuare confronti in termini di variabilità tra caratteri quantitativi occorre utilizzare una misura relativa, ad esempio il coefficiente di variazione

$$CV = \frac{\sigma}{\mu} = \frac{\sqrt{\text{Var}(Y)}}{M(Y)} = \frac{\sqrt{98.8369}}{16.725} = \frac{9.9417}{16.725} = 0.5944.$$

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.



3. Si calcoli la media geometrica e si individui su un opportuno grafico la mediana della variabile Y .

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.



Si tratta di un carattere quantitativo con modalità raggruppate in classi.

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.



La mediana (percentile di ordine 0.50) può essere ottenuta considerando l'inversa della Funzione di Ripartizione in corrispondenza di 0.50.

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.



La mediana (percentile di ordine 0.50) può essere ottenuta considerando l'inversa della Funzione di Ripartizione in corrispondenza di 0.50.

y_i	n_i	N_i	F_i
5 - 10	450	450	0.45
10 - 20	210	660	0.66
20 - 40	340	1000	1

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.

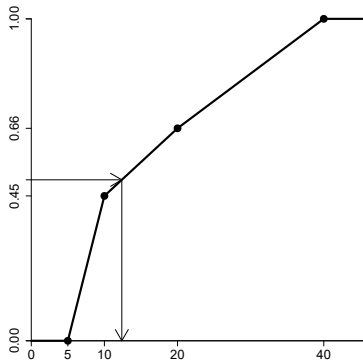


Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.



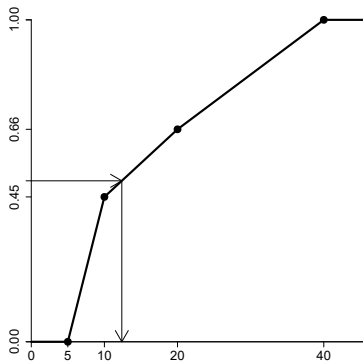


Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.



Si ottiene

$$y_{0.5} = 10 + \frac{0.5 - 0.45}{0.66 - 0.45} \cdot (20 - 10) = 12.381.$$



La media geometrica

$$\mu^{(0)}(Y) = \prod_{i=1}^k y_i^{f_i}$$

risulta 13.8988.

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.



La media geometrica

$$\mu^{(0)}(Y) = \prod_{i=1}^k y_i^{f_i}$$

risulta 13.8988. Si ottiene (considerando come modalità i valori centrali delle classi)

y_i	f_i	$y_i^{f_i}$
7.5	0.45	2.4762
15.0	0.21	1.7660
30.0	0.34	3.1785
		13.8988

da cui segue

$$\mu^{(0)} = \prod_{j=1}^k y_j^{f_j} = 13.8988.$$

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.



Il prospetto seguente contiene le serie storiche dei prezzi dei prodotti A , B e C :

t	$A p_t$	$B p_t$	$C p_t$
2011	200	400	600
2012	220	445	650
2013	210	420	$C p_{2013}$
2014	230	460	$C p_{2014}$
2015	250	490	700

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.



1. Sapendo che nel 2011 sono state scambiate le seguenti quantità per i 3 prodotti:
 ${}_Aq_{2011} = 100$, ${}_Bq_{2011} = 200$ e ${}_Cq_{2011} = 100$ si determini, giustificando la risposta, il prezzo del prodotto C nel 2013 tale che la variazione del livello generale dei prezzi dei 3 prodotti al tempo 2013 con base 2011, determinata secondo l'indice di Laspeyres, risulti pari al 5%.

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.



Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.

1. Sapendo che nel 2011 sono state scambiate le seguenti quantità per i 3 prodotti:

${}_A q_{2011} = 100$, ${}_B q_{2011} = 200$ e ${}_C q_{2011} = 100$ si determini, giustificando la risposta, il prezzo del prodotto C nel 2013 tale che la variazione del livello generale dei prezzi dei 3 prodotti al tempo 2013 con base 2011, determinata secondo l'indice di Laspeyres, risulti pari al 5%.

Dal momento che i prezzi relativi riferiti ai prodotti A e B risultano

$$\frac{{}_A p_{2013}}{{}_A p_{2011}} = \frac{210}{200} = 1.05 \quad \text{e} \quad \frac{{}_B p_{2013}}{{}_B p_{2011}} = \frac{420}{400} = 1.05$$

al fine di ottenere ${}_P_{2011}^L I_{2013}^L = 1.05$ deve valere anche $\frac{{}_C p_{2013}}{{}_C p_{2011}} = 1.05$ da cui segue ${}_C p_{2013} = 630$.



2. Si calcoli il numero indice del prezzo del prodotto A al tempo 2013 con base 2011 e si commenti il risultato ottenuto.

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.



2. Si calcoli il numero indice del prezzo del prodotto A al tempo 2013 con base 2011 e si commenti il risultato ottenuto.

$${}_{2011}I_{2013} = \frac{x_{2013}}{x_{2011}} = \frac{210}{200} = 1.05$$

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.



Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.

2. Si calcoli il numero indice del prezzo del prodotto A al tempo 2013 con base 2011 e si commenti il risultato ottenuto.

$${}_{2011}I_{2013} = \frac{x_{2013}}{x_{2011}} = \frac{210}{200} = 1.05$$

Tra il tempo 2011 e il tempo 2013 si è registrato un incremento complessivo del 5% della grandezza oggetto di studio.



3. Si calcoli il tasso medio di variazione del prezzo del prodotto *A* tra il 2012 e il 2014, interpretando il risultato ottenuto.

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.



Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.

3. Si calcoli il tasso medio di variazione del prezzo del prodotto *A* tra il 2012 e il 2014, interpretando il risultato ottenuto.

Si ricorda che

$${}_{2012}I_{2014} = {}_{2012}I_{2013} \cdot {}_{2013}I_{2014}$$

Si deve trovare il numero indice a base mobile medio.



Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.

3. Si calcoli il tasso medio di variazione del prezzo del prodotto *A* tra il 2012 e il 2014, interpretando il risultato ottenuto.

Si ricorda che

$${}_{2012}l_{2014} = {}_{2012}l_{2013} \cdot {}_{2013}l_{2014}$$

Si deve trovare il numero indice a base mobile medio. Applicando il criterio di scelta della media secondo Chisini si ottiene

$${}_{2012}l_{2014} = {}_{2012}l_{2013} \cdot {}_{2013}l_{2014} = \alpha \cdot \alpha = \alpha^2.$$



Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.

3. Si calcoli il tasso medio di variazione del prezzo del prodotto A tra il 2012 e il 2014, interpretando il risultato ottenuto.

Si ricorda che

$${}_{2012}I_{2014} = {}_{2012}I_{2013} \cdot {}_{2013}I_{2014}$$

Si deve trovare il numero indice a base mobile medio. Applicando il criterio di scelta della media secondo Chisini si ottiene

$${}_{2012}I_{2014} = {}_{2012}I_{2013} \cdot {}_{2013}I_{2014} = \alpha \cdot \alpha = \alpha^2.$$

Quindi

$$\alpha = \sqrt[2]{{}_{2012}I_{2014}} = \sqrt[2]{{}_{2012}I_{2013} \cdot {}_{2013}I_{2014}} = 1.0225.$$

Da cui segue una variazione relativa media pari a 0.0225, corrispondente a un incremento medio annuo tra il 2012 e il 2014 del 2.25%.



Il seguente prospetto riporta i valori rilevati in corrispondenza a 5 unità statistiche con riferimento a due caratteri quantitativi X e Y

x_i	2	3	3	3	5
y_i	1	2	2	3	4

1. Si rappresentino graficamente le coppie dei dati e la funzione di regressione ipotizzando Y in funzione di X .

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.



La funzione di regressione associa a ogni valore x appartenente all'insieme di definizione della variabile X la media di Y condizionata a x , $M(Y|x)$

x	$M(Y x)$
2	1
3	$\frac{1}{3}(2 + 2 + 3) = 2.3333$
5	4

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.



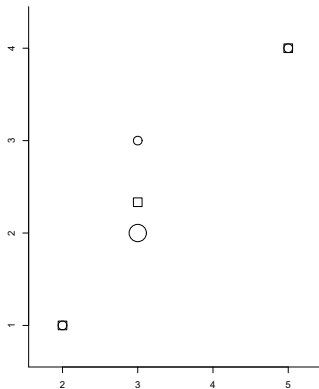
Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.

Le coppie di dati possono essere rappresentate mediante un *bubble plot diagram*.





2. Si calcolino, secondo il metodo dei minimi quadrati, i parametri dei seguenti modelli di regressione:

$$I) Y^* = a + b \cdot X$$

$$II) Y^* = c \cdot X$$

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.

Si riportano i calcoli necessari¹ per ottenere le stime dei parametri e gli indici di adattamento del modello

$$Y = a + bX + E$$

x_i	y_i	$x_i y_i$	x_i^2	y_i^2
2	1	2	4	1
3	2	6	9	4
3	2	6	9	4
3	3	9	9	9
5	4	20	25	16
16	12	43	56	34
3.2	2.4	8.6	11.2	6.8
$M(X)$	$M(Y)$	$M(XY)$	$M(X^2)$	$M(Y^2)$

$$\hat{b} =$$



Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.

¹Si osserva che per ragioni di stampa alcuni calcoli potrebbero essere duplicati con riferimento ai diversi modelli oggetto di studio

Si riportano i calcoli necessari¹ per ottenere le stime dei parametri e gli indici di adattamento del modello

$$Y = a + bX + E$$

x_i	y_i	$x_i y_i$	x_i^2	y_i^2
2	1	2	4	1
3	2	6	9	4
3	2	6	9	4
3	3	9	9	9
5	4	20	25	16
16	12	43	56	34
3.2	2.4	8.6	11.2	6.8
$M(X)$	$M(Y)$	$M(XY)$	$M(X^2)$	$M(Y^2)$

$$\begin{aligned}\hat{b} &= \frac{\text{Cov}(X, Y)}{\text{Var}(X)} = \frac{M(XY) - M(X)M(Y)}{M(X^2) - [M(X)]^2} = \\ &= \frac{8.6 - 3.2 \cdot 2.4}{11.2 - 3.2^2} = \frac{0.92}{0.96} = 0.9583 \\ \hat{a} &= \end{aligned}$$

¹Si osserva che per ragioni di stampa alcuni calcoli potrebbero essere duplicati con riferimento ai diversi modelli oggetto di studio



Si riportano i calcoli necessari¹ per ottenere le stime dei parametri e gli indici di adattamento del modello

$$Y = a + bX + E$$

x_i	y_i	$x_i y_i$	x_i^2	y_i^2
2	1	2	4	1
3	2	6	9	4
3	2	6	9	4
3	3	9	9	9
5	4	20	25	16
16	12	43	56	34
3.2	2.4	8.6	11.2	6.8
$M(X)$	$M(Y)$	$M(XY)$	$M(X^2)$	$M(Y^2)$

$$\hat{b} = \frac{\text{Cov}(X, Y)}{\text{Var}(X)} = \frac{M(XY) - M(X)M(Y)}{M(X^2) - [M(X)]^2} =$$

$$= \frac{8.6 - 3.2 \cdot 2.4}{11.2 - 3.2^2} = \frac{0.92}{0.96} = 0.9583$$

$$\hat{a} = M(Y) - \hat{b}M(X) = -0.6667$$

$$\rho^2 =$$



Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.

¹Si osserva che per ragioni di stampa alcuni calcoli potrebbero essere duplicati con riferimento ai diversi modelli oggetto di studio

Si riportano i calcoli necessari¹ per ottenere le stime dei parametri e gli indici di adattamento del modello

$$Y = a + bX + E$$

x_i	y_i	$x_i y_i$	x_i^2	y_i^2
2	1	2	4	1
3	2	6	9	4
3	2	6	9	4
3	3	9	9	9
5	4	20	25	16
16	12	43	56	34
3.2	2.4	8.6	11.2	6.8
$M(X)$	$M(Y)$	$M(XY)$	$M(X^2)$	$M(Y^2)$

$$\hat{b} = \frac{\text{Cov}(X, Y)}{\text{Var}(X)} = \frac{M(XY) - M(X)M(Y)}{M(X^2) - [M(X)]^2} =$$

$$= \frac{8.6 - 3.2 \cdot 2.4}{11.2 - 3.2^2} = \frac{0.92}{0.96} = 0.9583$$

$$\hat{a} = M(Y) - \hat{b}M(X) = -0.6667$$

$$\rho^2 = \frac{[\text{Cov}(X, Y)]^2}{\text{Var}(X) \text{Var}(Y)} = \frac{0.92^2}{0.96 \cdot 1.04} = 0.8478$$

$$M(\hat{E}^2) = \text{Var}(\hat{E}) = \text{varianza residua} =$$

¹Si osserva che per ragioni di stampa alcuni calcoli potrebbero essere duplicati con riferimento ai diversi modelli oggetto di studio



Si riportano i calcoli necessari¹ per ottenere le stime dei parametri e gli indici di adattamento del modello

$$Y = a + bX + E$$

x_i	y_i	$x_i y_i$	x_i^2	y_i^2
2	1	2	4	1
3	2	6	9	4
3	2	6	9	4
3	3	9	9	9
5	4	20	25	16
16	12	43	56	34
3.2	2.4	8.6	11.2	6.8
$M(X)$	$M(Y)$	$M(XY)$	$M(X^2)$	$M(Y^2)$

$$\begin{aligned}\hat{b} &= \frac{\text{Cov}(X, Y)}{\text{Var}(X)} = \frac{M(XY) - M(X)M(Y)}{M(X^2) - [M(X)]^2} = \\ &= \frac{8.6 - 3.2 \cdot 2.4}{11.2 - 3.2^2} = \frac{0.92}{0.96} = 0.9583 \\ \hat{a} &= M(Y) - \hat{b}M(X) = -0.6667 \\ \rho^2 &= \frac{[\text{Cov}(X, Y)]^2}{\text{Var}(X) \text{Var}(Y)} = \frac{0.92^2}{0.96 \cdot 1.04} = 0.8478\end{aligned}$$

$$M(\hat{E}^2) = \text{Var}(\hat{E}) = \text{varianza residua} = \text{Var}(Y) \cdot (1 - \rho^2) = 0.1583$$

¹Si osserva che per ragioni di stampa alcuni calcoli potrebbero essere duplicati con riferimento ai diversi modelli oggetto di studio





Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.

Si riportano i calcoli necessari² per ottenere le stime dei parametri e la misura di adattamento del modello

$$Y = cX + E$$

x_i	y_i	$x_i y_i$	x_i^2	y_i^2
2	1	2	4	1
3	2	6	9	4
3	2	6	9	4
3	3	9	9	9
5	4	20	25	16
		43	56	34
		8.6	11.2	6.8
		$M(XY)$	$M(X^2)$	$M(Y^2)$

$$\hat{c} =$$

²Si osserva che per ragioni di stampa alcuni calcoli potrebbero essere duplicati con riferimento ai diversi modelli oggetto di studio



Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.

Si riportano i calcoli necessari² per ottenere le stime dei parametri e la misura di adattamento del modello

$$Y = cX + E$$

x_i	y_i	$x_i y_i$	x_i^2	y_i^2
2	1	2	4	1
3	2	6	9	4
3	2	6	9	4
3	3	9	9	9
5	4	20	25	16
		43	56	34
		8.6	11.2	6.8
		$M(XY)$	$M(X^2)$	$M(Y^2)$

$$\hat{c} = \frac{M(XY)}{M(X^2)} = \frac{8.6}{11.2} = 0.7679$$

$$M(\hat{E}^2) =$$

²Si osserva che per ragioni di stampa alcuni calcoli potrebbero essere duplicati con riferimento ai diversi modelli oggetto di studio



Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.

Si riportano i calcoli necessari² per ottenere le stime dei parametri e la misura di adattamento del modello

$$Y = cX + E$$

x_i	y_i	$x_i y_i$	x_i^2	y_i^2
2	1	2	4	1
3	2	6	9	4
3	2	6	9	4
3	3	9	9	9
5	4	20	25	16
		43	56	34
		8.6	11.2	6.8
		$M(XY)$	$M(X^2)$	$M(Y^2)$

$$\hat{c} = \frac{M(XY)}{M(X^2)} = \frac{8.6}{11.2} = 0.7679$$

$$M(\hat{E}^2) = M[Y \cdot (Y - \hat{c}X)] = M(Y^2) - \hat{c}M(XY) = 6.8 - 0.7679 \cdot 8.6 = 0.1964$$

²Si osserva che per ragioni di stampa alcuni calcoli potrebbero essere duplicati con riferimento ai diversi modelli oggetto di studio



$\hat{a} = -0.6667$, $\hat{b} = 0.9583$, indice adattamento $M(\hat{E}^2) = 0.1583$ $\hat{c} = 0.7679$, indice adattamento $M(\hat{E}^2) = 0.1964$

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.



3. Si individui il modello migliore.

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.



Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.

3. Si individui il modello migliore.

Il modello migliore è $Y^* = -0.6667 + 0.9583 \cdot X$ in quanto caratterizzato da residuo quadratico medio più basso.



4. Si indichi, motivando la risposta ma senza fare calcoli, l'intervallo dei valori che può assumere $\eta_{X|Y}^2$.

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.



4. Si indichi, motivando la risposta ma senza fare calcoli, l'intervallo dei valori che può assumere $\eta_{X|Y}^2$.

$\eta_{X|Y}^2 = 1$ in quanto sussiste dipendenza funzionale di X da Y dal momento che le distribuzioni condizionate $X|y_j$ sono degeneri.

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.

Esercizio 4

Si consideri l'esperimento di estrazione di 4 palline da un'urna contenente 21 palline rosse e 29 palline blu.

1. Si calcoli la probabilità di ottenere almeno 3 palline blu nell'ipotesi che le 4 estrazioni siano effettuate senza reimmissione.



Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.

Esercizio 4

Si consideri l'esperimento di estrazione di 4 palline da un'urna contenente 21 palline rosse e 29 palline blu.

1. Si calcoli la probabilità di ottenere almeno 3 palline blu nell'ipotesi che le 4 estrazioni siano effettuate senza reimmissione.

Risulta

$$P(\text{almeno 3 } B) =$$



Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.

Esercizio 4



Si consideri l'esperimento di estrazione di 4 palline da un'urna contenente 21 palline rosse e 29 palline blu.

1. Si calcoli la probabilità di ottenere almeno 3 palline blu nell'ipotesi che le 4 estrazioni siano effettuate senza reimmissione.

Risulta

$$P(\text{almeno } 3 B) = P(3B) + P(4B) =$$

=

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.

Esercizio 4



Si consideri l'esperimento di estrazione di 4 palline da un'urna contenente 21 palline rosse e 29 palline blu.

1. Si calcoli la probabilità di ottenere almeno 3 palline blu nell'ipotesi che le 4 estrazioni siano effettuate senza reimmissione.

Risulta

$$\begin{aligned} P(\text{almeno } 3 B) &= P(3B) + P(4B) = \\ &= \binom{4}{3} \frac{29}{50} \frac{28}{49} \frac{27}{48} \frac{21}{47} + \binom{4}{4} \frac{29}{50} \frac{28}{49} \frac{27}{48} \frac{26}{47} = \\ &= \end{aligned}$$

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.

Esercizio 4



Si consideri l'esperimento di estrazione di 4 palline da un'urna contenente 21 palline rosse e 29 palline blu.

1. Si calcoli la probabilità di ottenere almeno 3 palline blu nell'ipotesi che le 4 estrazioni siano effettuate senza reimmissione.

Risulta

$$\begin{aligned} P(\text{almeno } 3 B) &= P(3B) + P(4B) = \\ &= \binom{4}{3} \frac{29}{50} \frac{28}{49} \frac{27}{48} \frac{21}{47} + \binom{4}{4} \frac{29}{50} \frac{28}{49} \frac{27}{48} \frac{26}{47} = \\ &= 0.3332 + 0.1031 = \end{aligned}$$

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.

Esercizio 4



Si consideri l'esperimento di estrazione di 4 palline da un'urna contenente 21 palline rosse e 29 palline blu.

1. Si calcoli la probabilità di ottenere almeno 3 palline blu nell'ipotesi che le 4 estrazioni siano effettuate senza reimmissione.

Risulta

$$\begin{aligned} P(\text{almeno } 3 B) &= P(3B) + P(4B) = \\ &= \binom{4}{3} \frac{29}{50} \frac{28}{49} \frac{27}{48} \frac{21}{47} + \binom{4}{4} \frac{29}{50} \frac{28}{49} \frac{27}{48} \frac{26}{47} = \\ &= 0.3332 + 0.1031 = 0.4363 \end{aligned}$$

o mediante la v.c. ipergeometrica

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.

Esercizio 4



Si consideri l'esperimento di estrazione di 4 palline da un'urna contenente 21 palline rosse e 29 palline blu.

1. Si calcoli la probabilità di ottenere almeno 3 palline blu nell'ipotesi che le 4 estrazioni siano effettuate senza reimmissione.

Risulta

$$\begin{aligned}P(\text{almeno } 3 B) &= P(3B) + P(4B) = \\&= \binom{4}{3} \frac{29}{50} \frac{28}{49} \frac{27}{48} \frac{21}{47} + \binom{4}{4} \frac{29}{50} \frac{28}{49} \frac{27}{48} \frac{26}{47} = \\&= 0.3332 + 0.1031 = 0.4363\end{aligned}$$

o mediante la v.c. ipergeometrica

$$P(X = 3) + P(X = 4) = \frac{\binom{29}{3} \binom{21}{1}}{\binom{50}{4}} + \frac{\binom{29}{4} \binom{21}{0}}{\binom{50}{4}} = 0.4363$$

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.



2. Si calcoli la probabilità di ottenere almeno 3 palline blu nell'ipotesi che le 4 estrazioni siano effettuate con reimmissione.

Sia X la variabile casuale 'n. di palline blu nelle 4 estrazioni (con reimmissione) (prove indipendenti)'. $X \sim \text{Bin}(n = 4, p = 0.58)$.

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.



Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.

2. Si calcoli la probabilità di ottenere almeno 3 palline blu nell'ipotesi che le 4 estrazioni siano effettuate con reimmissione.

Sia X la variabile casuale 'n. di palline blu nelle 4 estrazioni (con reimmissione) (prove indipendenti)'. $X \sim \text{Bin}(n = 4, p = 0.58)$.

$$P(X \geq 3) = \binom{4}{3} 0.58^3 0.42^1 + \binom{4}{4} 0.58^4 0.42^0 = 0.3278 + 0.1132 = 0.4410.$$



3. Si calcoli la probabilità di ottenere almeno 36 palline blu nell'ipotesi che siano effettuate 60 estrazioni con reimmissione.

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.



3. Si calcoli la probabilità di ottenere almeno 36 palline blu nell'ipotesi che siano effettuate 60 estrazioni con reimmissione.

Indicata con X la variabile casuale 'numero di successi' nelle 60 prove indipendenti risulta

$$X \sim$$

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.



3. Si calcoli la probabilità di ottenere almeno 36 palline blu nell'ipotesi che siano effettuate 60 estrazioni con reimmissione.

Indicata con X la variabile casuale 'numero di successi' nelle 60 prove indipendenti risulta

$$X \sim \text{Bin}(n = 60, p = 0.58).$$

$$M(X) = 34.8; \text{Var}(X) = 14.6160.$$

Dal momento che $np = 34.8 > 5$ ed $n(1 - p) = 25.2 > 5$ è possibile utilizzare l'approssimazione della variabile casuale Binomiale con la variabile casuale Normale

$$X \sim$$



3. Si calcoli la probabilità di ottenere almeno 36 palline blu nell'ipotesi che siano effettuate 60 estrazioni con reimmissione.

Indicata con X la variabile casuale 'numero di successi' nelle 60 prove indipendenti risulta

$$X \sim \text{Bin}(n = 60, p = 0.58).$$

$$M(X) = 34.8; \text{Var}(X) = 14.6160.$$

Dal momento che $np = 34.8 > 5$ ed $n(1 - p) = 25.2 > 5$ è possibile utilizzare l'approssimazione della variabile casuale Binomiale con la variabile casuale Normale

$$X \sim N(\mu = 34.8, \sigma^2 = 14.6160).$$

Si ottiene



3. Si calcoli la probabilità di ottenere almeno 36 palline blu nell'ipotesi che siano effettuate 60 estrazioni con reimmissione.

Indicata con X la variabile casuale 'numero di successi' nelle 60 prove indipendenti risulta

$$X \sim \text{Bin}(n = 60, p = 0.58).$$

$$M(X) = 34.8; \text{Var}(X) = 14.6160.$$

Dal momento che $np = 34.8 > 5$ ed $n(1 - p) = 25.2 > 5$ è possibile utilizzare l'approssimazione della variabile casuale Binomiale con la variabile casuale Normale

$$X \sim N(\mu = 34.8, \sigma^2 = 14.6160).$$

Si ottiene

$$P(X > 36) =$$



3. Si calcoli la probabilità di ottenere almeno 36 palline blu nell'ipotesi che siano effettuate 60 estrazioni con reimmissione.

Indicata con X la variabile casuale 'numero di successi' nelle 60 prove indipendenti risulta

$$X \sim \text{Bin}(n = 60, p = 0.58).$$

$$M(X) = 34.8; \text{Var}(X) = 14.6160.$$

Dal momento che $np = 34.8 > 5$ ed $n(1 - p) = 25.2 > 5$ è possibile utilizzare l'approssimazione della variabile casuale Binomiale con la variabile casuale Normale

$$X \sim N(\mu = 34.8, \sigma^2 = 14.6160).$$

Si ottiene

$$P(X > 36) = P\left(\frac{X - \mu}{\sigma} > \frac{36 - \mu}{\sigma}\right) =$$



3. Si calcoli la probabilità di ottenere almeno 36 palline blu nell'ipotesi che siano effettuate 60 estrazioni con reimmissione.

Indicata con X la variabile casuale 'numero di successi' nelle 60 prove indipendenti risulta

$$X \sim \text{Bin}(n = 60, p = 0.58).$$

$$M(X) = 34.8; \text{Var}(X) = 14.6160.$$

Dal momento che $np = 34.8 > 5$ ed $n(1 - p) = 25.2 > 5$ è possibile utilizzare l'approssimazione della variabile casuale Binomiale con la variabile casuale Normale

$$X \sim N(\mu = 34.8, \sigma^2 = 14.6160).$$

Si ottiene

$$P(X > 36) = P\left(\frac{X - \mu}{\sigma} > \frac{36 - \mu}{\sigma}\right) = P\left(Z > \frac{36 - 34.8}{3.8231}\right) =$$

=



3. Si calcoli la probabilità di ottenere almeno 36 palline blu nell'ipotesi che siano effettuate 60 estrazioni con reimmissione.

Indicata con X la variabile casuale 'numero di successi' nelle 60 prove indipendenti risulta

$$X \sim \text{Bin}(n = 60, p = 0.58).$$

$$M(X) = 34.8; \text{Var}(X) = 14.6160.$$

Dal momento che $np = 34.8 > 5$ ed $n(1 - p) = 25.2 > 5$ è possibile utilizzare l'approssimazione della variabile casuale Binomiale con la variabile casuale Normale

$$X \sim N(\mu = 34.8, \sigma^2 = 14.6160).$$

Si ottiene

$$\begin{aligned} P(X > 36) &= P\left(\frac{X - \mu}{\sigma} > \frac{36 - \mu}{\sigma}\right) = P\left(Z > \frac{36 - 34.8}{3.8231}\right) = \\ &= P(Z > 0.31) = \end{aligned}$$



3. Si calcoli la probabilità di ottenere almeno 36 palline blu nell'ipotesi che siano effettuate 60 estrazioni con reimmissione.

Indicata con X la variabile casuale 'numero di successi' nelle 60 prove indipendenti risulta

$$X \sim \text{Bin}(n = 60, p = 0.58).$$

$$M(X) = 34.8; \text{Var}(X) = 14.6160.$$

Dal momento che $np = 34.8 > 5$ ed $n(1 - p) = 25.2 > 5$ è possibile utilizzare l'approssimazione della variabile casuale Binomiale con la variabile casuale Normale

$$X \sim N(\mu = 34.8, \sigma^2 = 14.6160).$$

Si ottiene

$$\begin{aligned} P(X > 36) &= P\left(\frac{X - \mu}{\sigma} > \frac{36 - \mu}{\sigma}\right) = P\left(Z > \frac{36 - 34.8}{3.8231}\right) = \\ &= P(Z > 0.31) = 1 - P(Z \leq 0.31) = \end{aligned}$$



3. Si calcoli la probabilità di ottenere almeno 36 palline blu nell'ipotesi che siano effettuate 60 estrazioni con reimmissione.

Indicata con X la variabile casuale 'numero di successi' nelle 60 prove indipendenti risulta

$$X \sim \text{Bin}(n = 60, p = 0.58).$$

$$M(X) = 34.8; \text{Var}(X) = 14.6160.$$

Dal momento che $np = 34.8 > 5$ ed $n(1 - p) = 25.2 > 5$ è possibile utilizzare l'approssimazione della variabile casuale Binomiale con la variabile casuale Normale

$$X \sim N(\mu = 34.8, \sigma^2 = 14.6160).$$

Si ottiene

$$\begin{aligned} P(X > 36) &= P\left(\frac{X - \mu}{\sigma} > \frac{36 - \mu}{\sigma}\right) = P\left(Z > \frac{36 - 34.8}{3.8231}\right) = \\ &= P(Z > 0.31) = 1 - P(Z \leq 0.31) = 1 - 0.6217 = \end{aligned}$$



3. Si calcoli la probabilità di ottenere almeno 36 palline blu nell'ipotesi che siano effettuate 60 estrazioni con reimmissione.

Indicata con X la variabile casuale 'numero di successi' nelle 60 prove indipendenti risulta

$$X \sim \text{Bin}(n = 60, p = 0.58).$$

$$M(X) = 34.8; \text{Var}(X) = 14.6160.$$

Dal momento che $np = 34.8 > 5$ ed $n(1 - p) = 25.2 > 5$ è possibile utilizzare l'approssimazione della variabile casuale Binomiale con la variabile casuale Normale

$$X \sim N(\mu = 34.8, \sigma^2 = 14.6160).$$

Si ottiene

$$\begin{aligned} P(X > 36) &= P\left(\frac{X - \mu}{\sigma} > \frac{36 - \mu}{\sigma}\right) = P\left(Z > \frac{36 - 34.8}{3.8231}\right) = \\ &= P(Z > 0.31) = 1 - P(Z \leq 0.31) = 1 - 0.6217 = 0.3783. \end{aligned}$$

Tema T299

Prova scritta di STATISTICA del 18.06.2018

Prova scritta di
STATISTICA del
18.06.2018

Giuseppe Boari
Gabriele Cantaluppi



UNIVERSITÀ
CATTOLICA
del Sacro Cuore

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.

Giuseppe Boari
Gabriele Cantaluppi
Università Cattolica del Sacro Cuore, Milano



Un'azienda che produce cibo e prodotti per animali ha condotto una indagine campionaria su 40 individui residenti in una città e proprietari di animali domestici al fine di valutare il luogo ideale per aprire un nuovo punto vendita. Nella tabella sottostante sono state rilevate le frequenze congiunte di due variabili $X = \text{'Luogo di abitazione dell'intervistato'}$ e $Y = \text{'Spesa settimanale in cibo e prodotti per animali'}$ (in euro):

$X \setminus Y$	0 - 10	10 - 20	20 - 100
<i>centro</i>	7	9	0
<i>periferia</i>	8	8	8

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.



1. Fornire una opportuna rappresentazione grafica delle distribuzioni marginali di X e di Y .

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.



Si considera la distribuzione marginale della variabile X , ottenuta sommando le frequenze di riga

x_i	$n_{i\bullet}$
<i>centro</i>	16
<i>periferia</i>	24

Si tratta di un carattere qualitativo sconnesso che può essere rappresentato graficamente mediante un grafico a bastoncini o a canne d'organo.

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.

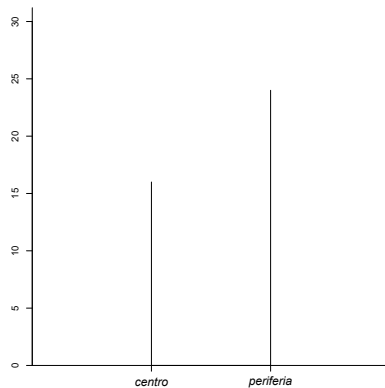


Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.





Si considera la distribuzione marginale della variabile Y , ottenuta sommando le frequenze di colonna

y_j	$n_{\bullet j}$
0 $\vdash$ 10	15
10 $\vdash$ 20	17
20 $\vdash$ 100	8

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.



Si tratta di un carattere quantitativo con modalità raggruppate in classi che può essere rappresentato graficamente mediante un istogramma.

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.



L'istogramma può essere costruito dopo avere calcolato le densità $d_{\bullet j} = \frac{n_{\bullet j}}{a_j}$, (dove a_j sono le ampiezze di classe).

y_j	$n_{\bullet j}$	$a_{\bullet j}$	$d_{\bullet j}$
0 ÷ 10	15	10 - 0 = 10	1.5
10 ÷ 20	17	20 - 10 = 10	1.7
20 ÷ 100	8	100 - 20 = 80	0.1

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.

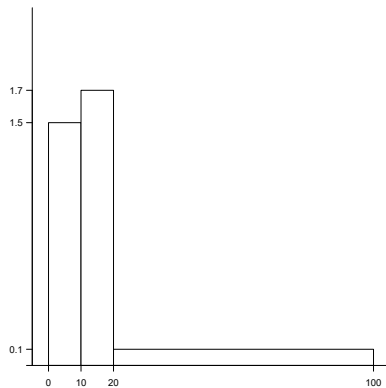


Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.





2. Calcolare moda, media e mediana per la marginale Y .

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.



Si considera la distribuzione marginale della variabile Y , ottenuta sommando le frequenze di colonna

y_j	$n_{\bullet j}$
0 $\vdash$ 10	15
10 $\vdash$ 20	17
20 $\vdash$ 100	8

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.



Si tratta di un carattere quantitativo con modalità raggruppate in classi.

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.



La moda corrisponde alla classe alla quale è associata la massima densità di frequenza

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.



La moda corrisponde alla classe alla quale è associata la massima densità di frequenza

y_j	$n_{\bullet j}$	$a_{\bullet j}$	$d_{\bullet j}$
0 $\vdash$ 10	15	$10 - 0 = 10$	1.5
10 $\vdash$ 20	17	$20 - 10 = 10$	1.7
20 $\vdash$ 100	8	$100 - 20 = 80$	0.1

e risulta 10 $\vdash$ 20.

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.



La mediana (percentile di ordine 0.50) può essere ottenuta considerando l'inversa della Funzione di Ripartizione in corrispondenza di 0.50.

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.



La mediana (percentile di ordine 0.50) può essere ottenuta considerando l'inversa della Funzione di Ripartizione in corrispondenza di 0.50.

y_j	$n_{\bullet j}$	$N_{\bullet j}$	$F_{\bullet j}$
0 - 10	15	15	0.375
10 - 20	17	32	0.8
20 - 100	8	40	1

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.

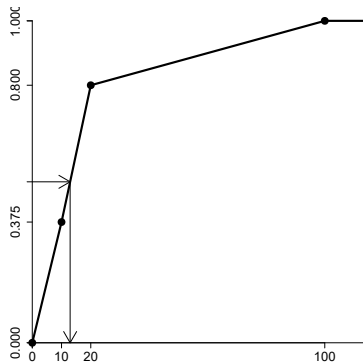


Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.



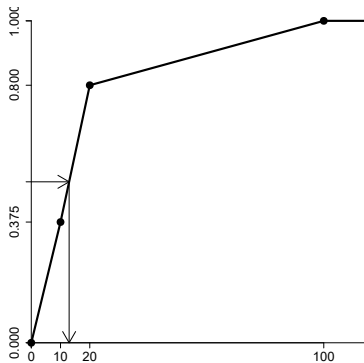


Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.



Si ottiene

$$y_{0.5} = 10 + \frac{0.5 - 0.375}{0.8 - 0.375} \cdot (20 - 10) = 12.9412.$$



La media aritmetica

$$M(Y) = \frac{1}{n} \sum_{j=1}^k y_j n_{\bullet j}$$

risulta 20.25.

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.



La media aritmetica

$$M(Y) = \frac{1}{n} \sum_{j=1}^k y_j n_{\bullet j}$$

risulta 20.25.

Si ottiene (considerando come modalità i valori centrali delle classi)

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.



La media aritmetica

$$M(Y) = \frac{1}{n} \sum_{j=1}^k y_j n_{\bullet j}$$

risulta 20.25.

Si ottiene (considerando come modalità i valori centrali delle classi)

y_j	$n_{\bullet j}$	$y_j n_{\bullet j}$
5	15	75
15	17	255
60	8	480
	40	810

da cui segue

$$M(Y) = \frac{1}{n} \sum_{i=1}^k y_i n_{\bullet i} = \frac{810}{40} = 20.25.$$

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.



moda = 10 + 20 media aritmetica $M(Y) = 20.25$

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.



Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.

3. Calcolare un opportuno indice di eterogeneità per la variabile X e commentarne il valore.



Si considera la distribuzione marginale della variabile X , ottenuta sommando le frequenze di riga

X_i	$n_{i\bullet}$
<i>centro</i>	16
<i>periferia</i>	24

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.



In presenza di un carattere qualitativo la variabilità può essere misurata mediante gli indici di Gini e di Frosini normalizzati

$$G_N = \frac{1 - \sum_{i=1}^h f_{i\bullet}^2}{1 - \frac{1}{h}}$$

e

$$F_N = 1 - \sqrt{1 - G_N}.$$

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.



In presenza di un carattere qualitativo la variabilità può essere misurata mediante gli indici di Gini e di Frosini normalizzati

$$G_N = \frac{1 - \sum_{i=1}^h f_{i\bullet}^2}{1 - \frac{1}{h}}$$

e

$$F_N = 1 - \sqrt{1 - G_N}.$$

Si ottiene:

x_i	$n_{i\bullet}$	$f_{i\bullet}$	$f_{i\bullet}^2$
<i>centro</i>	16	0.4	0.16
<i>periferia</i>	24	0.6	0.36
	40	1	0.52

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.



In presenza di un carattere qualitativo la variabilità può essere misurata mediante gli indici di Gini e di Frosini normalizzati

$$G_N = \frac{1 - \sum_{i=1}^h f_{i\bullet}^2}{1 - \frac{1}{h}}$$

e

$$F_N = 1 - \sqrt{1 - G_N}.$$

Si ottiene:

x_i	$n_{i\bullet}$	$f_{i\bullet}$	$f_{i\bullet}^2$
<i>centro</i>	16	0.4	0.16
<i>periferia</i>	24	0.6	0.36
	40	1	0.52

Da cui segue:

$$G_N = \frac{1 - 0.52}{1 - \frac{1}{2}} = 0.96$$

e

$$F_N = 1 - \sqrt{1 - 0.96} = 0.8.$$

Sussiste, quindi, un livello alto di eterogeneità.



4. Valutare il grado di connessione fra i caratteri X e Y utilizzando un opportuno indice normalizzato.

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.



Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.

Dalla distribuzione delle frequenze congiunte n_{ij} :

	0 ÷ 10	10 ÷ 20	20 ÷ 100	$n_{i\bullet}$
<i>centro</i>	7	9	0	16
<i>periferia</i>	8	8	8	24
$n_{\bullet j}$	15	17	8	40

è possibile ottenere le frequenze teoriche nella situazione di indipendenza stocastica

$$\hat{n}_{ij} = \frac{n_{i\bullet} n_{\bullet j}}{n}:$$



Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.

Dalla distribuzione delle frequenze congiunte n_{ij} :

	0 - 10	10 - 20	20 - 100	$n_{i\bullet}$
<i>centro</i>	7	9	0	16
<i>periferia</i>	8	8	8	24
$n_{\bullet j}$	15	17	8	40

è possibile ottenere le frequenze teoriche nella situazione di indipendenza stocastica

$$\hat{n}_{ij} = \frac{n_{i\bullet} \cdot n_{\bullet j}}{n}:$$

	0 - 10	10 - 20	20 - 100	$n_{i\bullet}$
<i>centro</i>	6	6.8	3.2	16
<i>periferia</i>	9	10.2	4.8	24
$n_{\bullet j}$	15	17.0	8.0	40



le contingenze assolute $n_{ij} - \hat{n}_{ij}$:

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.



le contingenze assolute $n_{ij} - \hat{n}_{ij}$:

	0 ÷ 10	10 ÷ 20	20 ÷ 100
<i>centro</i>	1	2.2	-3.2
<i>periferia</i>	-1	-2.2	3.2

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.



i rapporti di contingenza $\frac{n_{ij}}{\hat{n}_{ij}}$:

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.



i rapporti di contingenza $\frac{n_{ij}}{\hat{n}_{ij}}$:

	0 - 10	10 - 20	20 - 100
<i>centro</i>	1.1667	1.3235	0.0000
<i>periferia</i>	0.8889	0.7843	1.6667

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.



e le contingenze relative, espresse in termini percentuali, $100 \left(\frac{n_{ij} - \hat{n}_{ij}}{\hat{n}_{ij}} \right) \% = 100 \left(\frac{n_{ij}}{\hat{n}_{ij}} - 1 \right) \%$:

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.



e le contingenze relative, espresse in termini percentuali, $100 \left(\frac{n_{ij} - \hat{n}_{ij}}{\hat{n}_{ij}} \right) \% = 100 \left(\frac{n_{ij}}{\hat{n}_{ij}} - 1 \right) \%$:

	0 - 10	10 - 20	20 - 100
<i>centro</i>	16.67	32.35	-100.00
<i>periferia</i>	-11.11	-21.57	66.67

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.



Risultati intermedi $\frac{(n_{ij} - \hat{n}_{ij})^2}{\hat{n}_{ij}}$ per il calcolo del χ^2

	0 ÷ 10	10 ÷ 20	20 ÷ 100
<i>centro</i>	0.1667	0.7118	3.2000
<i>periferia</i>	0.1111	0.4745	2.1333

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.



Risultati intermedi $\frac{n_{ij}^2}{n_{i\bullet} \cdot n_{\bullet j}}$ per il calcolo del χ^2 (formula operativa)

	0 ÷ 10	10 ÷ 20	20 ÷ 100
<i>centro</i>	0.2042	0.2978	0.0000
<i>periferia</i>	0.1778	0.1569	0.3333

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.



Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.

$$\chi^2 = \sum_{i=1}^h \sum_{j=1}^k \frac{(n_{ij} - \hat{n}_{ij})^2}{\hat{n}_{ij}} = 6.7974,$$

$$\chi_N^2 = \frac{\chi^2}{n \cdot \min(h-1, k-1)} = 0.1699$$

$$\chi_N = \sqrt{\chi_N^2} = 0.4122.$$

Sussiste, quindi, un livello di connessione medio/basso.



Il prospetto seguente riporta la serie storica dei prezzi di due beni A e B fra il 2014 e il 2017:

$anno$	p_A	p_B
2014	100	200
2015	105	205
2016	110	220
2017	120	190

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.



1. Costruire la serie dei numeri indici a base fissa dei prezzi con base 2014 per il bene A.

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.



1. Costruire la serie dei numeri indici a base fissa dei prezzi con base 2014 per il bene A.
Si ricorda che ${}_{2014}I_t = \frac{x_t}{x_{2014}}$.

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.



1. Costruire la serie dei numeri indici a base fissa dei prezzi con base 2014 per il bene A. Si ricorda che ${}_{2014}I_t = \frac{x_t}{x_{2014}}$. Si ottiene quindi

Anno	x_t	NIBF (2014)
2014	100	1.00
2015	105	1.05
2016	110	1.10
2017	120	1.20

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.



2. Calcolare il tasso medio di variazione per il prezzo del bene A nell'intervallo 2014-2017 e commentare il risultato ottenuto.

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.



2. Calcolare il tasso medio di variazione per il prezzo del bene *A* nell'intervallo 2014-2017 e commentare il risultato ottenuto.

Si ricorda che

$${}_{2014}I_{2017} = {}_{2014}I_{2015} \cdot {}_{2015}I_{2016} \cdot {}_{2016}I_{2017}$$

Si deve trovare il numero indice a base mobile medio.

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.



Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.

2. Calcolare il tasso medio di variazione per il prezzo del bene A nell'intervallo 2014-2017 e commentare il risultato ottenuto.

Si ricorda che

$${}_{2014}I_{2017} = {}_{2014}I_{2015} \cdot {}_{2015}I_{2016} \cdot {}_{2016}I_{2017}$$

Si deve trovare il numero indice a base mobile medio. Applicando il criterio di scelta della media secondo Chisini si ottiene

$${}_{2014}I_{2017} = {}_{2014}I_{2015} \cdot {}_{2015}I_{2016} \cdot {}_{2016}I_{2017} = \alpha \cdot \alpha \cdot \alpha = \alpha^3.$$



Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.

2. Calcolare il tasso medio di variazione per il prezzo del bene A nell'intervallo 2014-2017 e commentare il risultato ottenuto.

Si ricorda che

$${}_{2014}I_{2017} = {}_{2014}I_{2015} \cdot {}_{2015}I_{2016} \cdot {}_{2016}I_{2017}$$

Si deve trovare il numero indice a base mobile medio. Applicando il criterio di scelta della media secondo Chisini si ottiene

$${}_{2014}I_{2017} = {}_{2014}I_{2015} \cdot {}_{2015}I_{2016} \cdot {}_{2016}I_{2017} = \alpha \cdot \alpha \cdot \alpha = \alpha^3.$$

Quindi

$$\alpha = \sqrt[3]{{}_{2014}I_{2017}} = \sqrt[3]{{}_{2014}I_{2015} \cdot {}_{2015}I_{2016} \cdot {}_{2016}I_{2017}} = 1.0627.$$

Da cui segue una variazione relativa media pari a 0.0627, corrispondente a un incremento medio annuo tra il 2014 e il 2017 del 6.27%.



3. Sapendo che le quantità ${}_{2014}q_A = 100$ e ${}_{2014}q_B = 200$ e che ${}_{2016}q_A = 150$ e ${}_{2016}q_B = 180$, calcolare l'indice di Fisher dei prezzi al 2016 con base al 2014.

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.



Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.

3. Sapendo che le quantità ${}_{2014}q_A = 100$ e ${}_{2014}q_B = 200$ e che ${}_{2016}q_A = 150$ e ${}_{2016}q_B = 180$, calcolare l'indice di Fisher dei prezzi al 2016 con base al 2014.

Si osserva che i prezzi di entrambi i beni hanno subito un incremento del 10% nel periodo considerato.

Abbiamo, quindi, ${}^P_{2014}I^F_{2016} = 1.10$. La variazione relativa nel livello generale dei prezzi, secondo l'Indice di Fisher, è del 10%.



Con riferimento allo studio della dipendenza tra due caratteri quantitativi X e Y , si sono raccolte le seguenti 5 coppie di osservazioni (x_i, y_i) $i = 1, 2, \dots, 5$

x_i	-8	-4	-3	-4	1
y_i	2	1	0	1	-2

1. Fornire la rappresentazione grafica delle coppie di punti e della funzione di regressione di Y dato X .

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.



La funzione di regressione associa a ogni valore x appartenente all'insieme di definizione della variabile X la media di Y condizionata a x , $M(Y|x)$

x	$M(Y x)$
-8	2
-4	1
-3	0
1	-2

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.



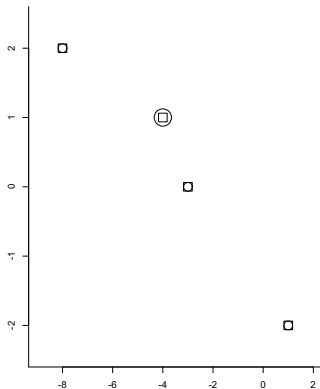
Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.

Le coppie di dati possono essere rappresentate mediante un *bubble plot diagram*.





2. Calcolare, secondo il metodo dei minimi quadrati, i parametri dei seguenti modelli di regressione:

$$I) Y^* = a + b \cdot X \quad II) Y^* = c \cdot X^2$$

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.

Si riportano i calcoli necessari¹ per ottenere le stime dei parametri e gli indici di adattamento del modello

$$Y = a + bX + E$$

x_i	y_i	$x_i y_i$	x_i^2	y_i^2
-8	2	-16	64	4
-4	1	-4	16	1
-3	0	0	9	0
-4	1	-4	16	1
1	-2	-2	1	4
-18	2	-26	106	10
-3.6	0.4	-5.2	21.2	2
$M(X)$	$M(Y)$	$M(XY)$	$M(X^2)$	$M(Y^2)$

$$\hat{b} =$$



Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.

¹Si osserva che per ragioni di stampa alcuni calcoli potrebbero essere duplicati con riferimento ai diversi modelli oggetto di studio

Si riportano i calcoli necessari¹ per ottenere le stime dei parametri e gli indici di adattamento del modello

$$Y = a + bX + E$$

x_i	y_i	$x_i y_i$	x_i^2	y_i^2
-8	2	-16	64	4
-4	1	-4	16	1
-3	0	0	9	0
-4	1	-4	16	1
1	-2	-2	1	4
-18	2	-26	106	10
-3.6	0.4	-5.2	21.2	2
$M(X)$	$M(Y)$	$M(XY)$	$M(X^2)$	$M(Y^2)$

$$\begin{aligned}\hat{b} &= \frac{\text{Cov}(X, Y)}{\text{Var}(X)} = \frac{M(XY) - M(X)M(Y)}{M(X^2) - [M(X)]^2} = \\ &= \frac{-5.2 - (-3.6) \cdot 0.4}{21.2 - (-3.6)^2} = \frac{-3.76}{8.24} = -0.4563 \\ \hat{a} &= \end{aligned}$$



Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.

¹Si osserva che per ragioni di stampa alcuni calcoli potrebbero essere duplicati con riferimento ai diversi modelli oggetto di studio

Si riportano i calcoli necessari¹ per ottenere le stime dei parametri e gli indici di adattamento del modello

$$Y = a + bX + E$$

x_i	y_i	$x_i y_i$	x_i^2	y_i^2
-8	2	-16	64	4
-4	1	-4	16	1
-3	0	0	9	0
-4	1	-4	16	1
1	-2	-2	1	4
-18	2	-26	106	10
-3.6	0.4	-5.2	21.2	2
$M(X)$	$M(Y)$	$M(XY)$	$M(X^2)$	$M(Y^2)$

$$\begin{aligned}\hat{b} &= \frac{\text{Cov}(X, Y)}{\text{Var}(X)} = \frac{M(XY) - M(X)M(Y)}{M(X^2) - [M(X)]^2} = \\ &= \frac{-5.2 - (-3.6) \cdot 0.4}{21.2 - (-3.6)^2} = \frac{-3.76}{8.24} = -0.4563 \\ \hat{a} &= M(Y) - \hat{b}M(X) = -1.2427 \\ \rho^2 &= \end{aligned}$$

¹Si osserva che per ragioni di stampa alcuni calcoli potrebbero essere duplicati con riferimento ai diversi modelli oggetto di studio



Si riportano i calcoli necessari¹ per ottenere le stime dei parametri e gli indici di adattamento del modello

$$Y = a + bX + E$$

x_i	y_i	$x_i y_i$	x_i^2	y_i^2
-8	2	-16	64	4
-4	1	-4	16	1
-3	0	0	9	0
-4	1	-4	16	1
1	-2	-2	1	4
-18	2	-26	106	10
-3.6	0.4	-5.2	21.2	2
$M(X)$	$M(Y)$	$M(XY)$	$M(X^2)$	$M(Y^2)$

$$\begin{aligned}\hat{b} &= \frac{\text{Cov}(X, Y)}{\text{Var}(X)} = \frac{M(XY) - M(X)M(Y)}{M(X^2) - [M(X)]^2} = \\ &= \frac{-5.2 - (-3.6) \cdot 0.4}{21.2 - (-3.6)^2} = \frac{-3.76}{8.24} = -0.4563 \\ \hat{a} &= M(Y) - \hat{b}M(X) = -1.2427 \\ \rho^2 &= \frac{[\text{Cov}(X, Y)]^2}{\text{Var}(X) \text{Var}(Y)} = \frac{(-3.76)^2}{8.24 \cdot 1.84} = 0.9325\end{aligned}$$

$$M(\hat{E}^2) = \text{Var}(\hat{E}) = \text{varianza residua} =$$

¹Si osserva che per ragioni di stampa alcuni calcoli potrebbero essere duplicati con riferimento ai diversi modelli oggetto di studio



Si riportano i calcoli necessari¹ per ottenere le stime dei parametri e gli indici di adattamento del modello

$$Y = a + bX + E$$

x_i	y_i	$x_i y_i$	x_i^2	y_i^2
-8	2	-16	64	4
-4	1	-4	16	1
-3	0	0	9	0
-4	1	-4	16	1
1	-2	-2	1	4
-18	2	-26	106	10
-3.6	0.4	-5.2	21.2	2
$M(X)$	$M(Y)$	$M(XY)$	$M(X^2)$	$M(Y^2)$

$$\begin{aligned}\hat{b} &= \frac{\text{Cov}(X, Y)}{\text{Var}(X)} = \frac{M(XY) - M(X)M(Y)}{M(X^2) - [M(X)]^2} = \\ &= \frac{-5.2 - (-3.6) \cdot 0.4}{21.2 - (-3.6)^2} = \frac{-3.76}{8.24} = -0.4563 \\ \hat{a} &= M(Y) - \hat{b}M(X) = -1.2427 \\ \rho^2 &= \frac{[\text{Cov}(X, Y)]^2}{\text{Var}(X) \text{Var}(Y)} = \frac{(-3.76)^2}{8.24 \cdot 1.84} = 0.9325\end{aligned}$$

$$M(\hat{E}^2) = \text{Var}(\hat{E}) = \text{varianza residua} = \text{Var}(Y) \cdot (1 - \rho^2) = 0.1243$$

¹Si osserva che per ragioni di stampa alcuni calcoli potrebbero essere duplicati con riferimento ai diversi modelli oggetto di studio





Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.

Si riportano i calcoli necessari² per ottenere le stime dei parametri e la misura di adattamento del modello

$$Y = cX^2 + E$$

x_i^2	y_i	$x_i^2 y_i$	$x_i^{2^2}$	y_i^2
64	2	128	4096	4
16	1	16	256	1
9	0	0	81	0
16	1	16	256	1
1	-2	-2	1	4
		158	4690	10
		31.6	938	2
		$M(X^2 Y)$	$M(X^{2^2})$	$M(Y^2)$

$$\hat{c} =$$

²Si osserva che per ragioni di stampa alcuni calcoli potrebbero essere duplicati con riferimento ai diversi modelli oggetto di studio



Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.

Si riportano i calcoli necessari² per ottenere le stime dei parametri e la misura di adattamento del modello

$$Y = cX^2 + E$$

x_i^2	y_i	$x_i^2 y_i$	$x_i^{2^2}$	y_i^2
64	2	128	4096	4
16	1	16	256	1
9	0	0	81	0
16	1	16	256	1
1	-2	-2	1	4
		158	4690	10
		31.6	938	2
		$M(X^2 Y)$	$M(X^{2^2})$	$M(Y^2)$

$$\hat{c} = \frac{M(X^2 Y)}{M(X^{2^2})} = \frac{31.6}{938} = 0.0337$$

$$M(\hat{E}^2) =$$

²Si osserva che per ragioni di stampa alcuni calcoli potrebbero essere duplicati con riferimento ai diversi modelli oggetto di studio



Si riportano i calcoli necessari² per ottenere le stime dei parametri e la misura di adattamento del modello

$$Y = cX^2 + E$$

x_i^2	y_i	$x_i^2 y_i$	$x_i^{2^2}$	y_i^2
64	2	128	4096	4
16	1	16	256	1
9	0	0	81	0
16	1	16	256	1
1	-2	-2	1	4
		158	4690	10
		31.6	938	2
		$M(X^2 Y)$	$M(X^{2^2})$	$M(Y^2)$

$$\hat{c} = \frac{M(X^2 Y)}{M(X^{2^2})} = \frac{31.6}{938} = 0.0337$$

$$M(\hat{E}^2) = M[Y \cdot (Y - \hat{c}X^2)] = M(Y^2) - \hat{c}M(X^2 Y) = 2 - 0.0337 \cdot 31.6 = 0.9354$$

²Si osserva che per ragioni di stampa alcuni calcoli potrebbero essere duplicati con riferimento ai diversi modelli oggetto di studio



$\hat{a} = -1.2427$, $\hat{b} = -0.4563$, indice adattamento $M(\hat{E}^2) = 0.1243$ $\hat{c} = 0.03369$, indice adattamento $M(\hat{E}^2) = 0.9354$

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.



3. Confrontare gli indici di adattamento, commentando le eventuali differenze.

Il modello migliore è $Y^* = -1.2427 - 0.4563 \cdot X$ in quanto caratterizzato da residuo quadratico medio più basso.

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.



4. Sulla base dei risultati precedenti e senza svolgere ulteriori conti, cosa è possibile dire sul valore di $\eta^2_{Y|X}$.

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.



Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.

4. Sulla base dei risultati precedenti e senza svolgere ulteriori conti, cosa è possibile dire sul valore di $\eta_{Y|X}^2$.

$\eta_{Y|X}^2 = 1$ in quanto sussiste dipendenza funzionale di Y da X .



5. Sulla base del miglior modello individuato al punto 3. prevedere il valore di Y in corrispondenza di $X = 5$.

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.



Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.

5. Sulla base del miglior modello individuato al punto 3. prevedere il valore di Y in corrispondenza di $X = 5$.

$$\hat{Y}(x = 5) = -1.2427 - 0.4563 \cdot 5 = -3.5243.$$

Esercizio 4

Prova scritta di
STATISTICA del
18.06.2018

Giuseppe Boari
Gabriele Cantaluppi



UNIVERSITÀ
CATTOLICA
del Sacro Cuore

Con riferimento a un processo produttivo è noto che la proporzione di componenti difettose nella popolazione è pari a 0.12.

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.



1. Dopo avere estratto con reimmissione un campione di 4 componenti, quale è la probabilità che non più di una componente risulti difettosa?

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.



Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.

1. Dopo avere estratto con reimmissione un campione di 4 componenti, quale è la probabilità che non più di una componente risulti difettosa?

Sia X la variabile casuale 'n. di difettosi nelle 4 estrazioni con reimmissione'. Abbiamo $X \sim \text{Bin}(n = 4, p = 0.12)$.

$$P(X \leq 1) =$$



Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.

1. Dopo avere estratto con reimmissione un campione di 4 componenti, quale è la probabilità che non più di una componente risulti difettosa?

Sia X la variabile casuale 'n. di difettosi nelle 4 estrazioni con reimmissione'. Abbiamo $X \sim \text{Bin}(n = 4, p = 0.12)$.

$$P(X \leq 1) = P(X = 0) + P(X = 1) =$$

=



Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.

1. Dopo avere estratto con reimmissione un campione di 4 componenti, quale è la probabilità che non più di una componente risulti difettosa?

Sia X la variabile casuale 'n. di difettosi nelle 4 estrazioni con reimmissione'. Abbiamo $X \sim \text{Bin}(n = 4, p = 0.12)$.

$$\begin{aligned} P(X \leq 1) &= P(X = 0) + P(X = 1) = \\ &= \binom{4}{0} 0.12^0 \cdot 0.88^4 + \binom{4}{1} 0.12^1 \cdot 0.88^3 = \\ &= \end{aligned}$$



Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.

1. Dopo avere estratto con reimmissione un campione di 4 componenti, quale è la probabilità che non più di una componente risulti difettosa?

Sia X la variabile casuale 'n. di difettosi nelle 4 estrazioni con reimmissione'. Abbiamo $X \sim \text{Bin}(n = 4, p = 0.12)$.

$$\begin{aligned} P(X \leq 1) &= P(X = 0) + P(X = 1) = \\ &= \binom{4}{0} 0.12^0 \cdot 0.88^4 + \binom{4}{1} 0.12^1 \cdot 0.88^3 = \\ &= 0.5997 + 0.3271 = \end{aligned}$$



Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.

1. Dopo avere estratto con reimmissione un campione di 4 componenti, quale è la probabilità che non più di una componente risulti difettosa?

Sia X la variabile casuale 'n. di difettosi nelle 4 estrazioni con reimmissione'. Abbiamo $X \sim \text{Bin}(n = 4, p = 0.12)$.

$$\begin{aligned} P(X \leq 1) &= P(X = 0) + P(X = 1) = \\ &= \binom{4}{0} 0.12^0 \cdot 0.88^4 + \binom{4}{1} 0.12^1 \cdot 0.88^3 = \\ &= 0.5997 + 0.3271 = 0.9268. \end{aligned}$$



2. Calcolare il valore atteso e la varianza del numero di pezzi difettosi per un campione di 100 componenti.

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.



Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.

2. Calcolare il valore atteso e la varianza del numero di pezzi difettosi per un campione di 100 componenti.

Sia X la variabile casuale 'n. di difettosi nelle 100 estrazioni con reimmissione'. Abbiamo $X \sim \text{Bin}(n = 100, p = 0.12)$.

$$E(X) = n \cdot p = 100 \cdot 0.12 = 12 \quad \text{Var}(X) = n \cdot p \cdot (1 - p) = 100 \cdot 0.12 \cdot 0.88 = 10.56$$



3. Con riferimento al campione di 100 componenti del punto precedente, calcolare la probabilità che la quota di pezzi difettosi sia inferiore al 19%.

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.



Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.

3. Con riferimento al campione di 100 componenti del punto precedente, calcolare la probabilità che la quota di pezzi difettosi sia inferiore al 19%.

Dal momento che $np = 12 > 5$ ed $n(1 - p) = 88 > 5$ è possibile utilizzare l'approssimazione della variabile casuale Binomiale con la variabile casuale Normale

$X \sim N(\mu = 12, \sigma^2 = 10.56)$.

Si ottiene

$$P(X \leq 19) =$$



Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.

3. Con riferimento al campione di 100 componenti del punto precedente, calcolare la probabilità che la quota di pezzi difettosi sia inferiore al 19%.

Dal momento che $np = 12 > 5$ ed $n(1 - p) = 88 > 5$ è possibile utilizzare l'approssimazione della variabile casuale Binomiale con la variabile casuale Normale

$X \sim N(\mu = 12, \sigma^2 = 10.56)$.

Si ottiene

$$P(X \leq 19) = P\left(\frac{X - \mu}{\sigma} \leq \frac{19 - \mu}{\sigma}\right) =$$



Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.

3. Con riferimento al campione di 100 componenti del punto precedente, calcolare la probabilità che la quota di pezzi difettosi sia inferiore al 19%.

Dal momento che $np = 12 > 5$ ed $n(1 - p) = 88 > 5$ è possibile utilizzare l'approssimazione della variabile casuale Binomiale con la variabile casuale Normale

$X \sim N(\mu = 12, \sigma^2 = 10.56)$.

Si ottiene

$$P(X \leq 19) = P\left(\frac{X - \mu}{\sigma} \leq \frac{19 - \mu}{\sigma}\right) = P\left(Z \leq \frac{19 - 12}{3.2496}\right) =$$

=



Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.

3. Con riferimento al campione di 100 componenti del punto precedente, calcolare la probabilità che la quota di pezzi difettosi sia inferiore al 19%.

Dal momento che $np = 12 > 5$ ed $n(1 - p) = 88 > 5$ è possibile utilizzare l'approssimazione della variabile casuale Binomiale con la variabile casuale Normale

$X \sim N(\mu = 12, \sigma^2 = 10.56)$.

Si ottiene

$$\begin{aligned} P(X \leq 19) &= P\left(\frac{X - \mu}{\sigma} \leq \frac{19 - \mu}{\sigma}\right) = P\left(Z \leq \frac{19 - 12}{3.2496}\right) = \\ &= P(Z \leq 2.15) = \end{aligned}$$



Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.

3. Con riferimento al campione di 100 componenti del punto precedente, calcolare la probabilità che la quota di pezzi difettosi sia inferiore al 19%.

Dal momento che $np = 12 > 5$ ed $n(1 - p) = 88 > 5$ è possibile utilizzare l'approssimazione della variabile casuale Binomiale con la variabile casuale Normale

$X \sim N(\mu = 12, \sigma^2 = 10.56)$.

Si ottiene

$$\begin{aligned} P(X \leq 19) &= P\left(\frac{X - \mu}{\sigma} \leq \frac{19 - \mu}{\sigma}\right) = P\left(Z \leq \frac{19 - 12}{3.2496}\right) = \\ &= P(Z \leq 2.15) = 0.9842. \end{aligned}$$

Tema T2992

Prova scritta di STATISTICA del 10.07.2018

Prova scritta di
STATISTICA del
10.07.2018

Giuseppe Boari
Gabriele Cantaluppi



UNIVERSITÀ
CATTOLICA
del Sacro Cuore

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.

Giuseppe Boari
Gabriele Cantaluppi
Università Cattolica del Sacro Cuore, Milano



Data la seguente tabella a doppia entrata che riporta la distribuzione congiunta dei caratteri $X = \text{'genere'}$ e $Y = \text{'numero di visite mediche specialistiche fatte in un anno'}$ riferite a 9 persone:

$X \setminus Y$	0 + 2	2 + 3	3 + 8
<i>maschio</i>	0	0	5
<i>femmina</i>	2	2	0

1. Rappresentare graficamente le distribuzioni condizionate $Y|x_1$ e $Y|x_2$.

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.



Si considera la distribuzione condizionata $Y|x_1$

y_j	n_{1j}
0 + 2	0
2 + 3	0
3 + 8	5

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.

Si tratta di un carattere quantitativo con modalità raggruppate in classi che può essere rappresentato graficamente mediante un istogramma. L'istogramma può essere costruito dopo avere calcolato le densità $d_{1j} = \frac{n_{1j}}{a_j}$, (dove a_j sono le ampiezze di classe).

y_j	n_{1j}	a_{1j}	d_{1j}
0 - 2	0	$2 - 0 = 2$	0
2 - 3	0	$3 - 2 = 1$	0
3 - 8	5	$8 - 3 = 5$	1



Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.



Si considera la distribuzione condizionata $Y|x_1$

y_j	n_{1j}
0 + 2	2
2 + 3	2
3 + 8	0

Esercizio 1.

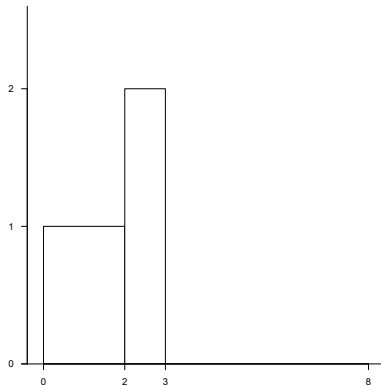
Esercizio 2.

Esercizio 3.

Esercizio 4.

Si tratta di un carattere quantitativo con modalità raggruppate in classi che può essere rappresentato graficamente mediante un istogramma. L'istogramma può essere costruito dopo avere calcolato le densità $d_{1j} = \frac{n_{1j}}{a_j}$, (dove a_j sono le ampiezze di classe).

y_j	n_{1j}	a_{1j}	d_{1j}
0 - 2	2	$2 - 0 = 2$	1
2 - 3	2	$3 - 2 = 1$	2
3 - 8	0	$8 - 3 = 5$	0



Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.



Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.

2. Si riassume la marginale X con opportuni indici di posizione e di variabilità.



Si considera la distribuzione marginale della variabile X , ottenuta sommando le frequenze di riga

x_i	$n_{i\bullet}$
<i>maschio</i>	5
<i>femmina</i>	4

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.



Si tratta di un carattere qualitativo sconnesso.

L'unico indice di posizione che può essere definito è la moda (categoria alla quale è associata la massima frequenza) che risulta *maschio*.

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.



In presenza di un carattere qualitativo la variabilità può essere misurata mediante gli indici di Gini e di Frosini normalizzati

$$G_N = \frac{1 - \sum_{i=1}^h f_{i\bullet}^2}{1 - \frac{1}{h}}$$

e

$$F_N = 1 - \sqrt{1 - G_N}.$$

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.



In presenza di un carattere qualitativo la variabilità può essere misurata mediante gli indici di Gini e di Frosini normalizzati

$$G_N = \frac{1 - \sum_{i=1}^h f_{i\bullet}^2}{1 - \frac{1}{h}}$$

e

$$F_N = 1 - \sqrt{1 - G_N}.$$

Si ottiene:

x_i	$n_{i\bullet}$	$f_{i\bullet}$	$f_{i\bullet}^2$
<i>maschio</i>	5	0.5556	0.3086
<i>femmina</i>	4	0.4444	0.1975
	9	1	0.5061

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.



In presenza di un carattere qualitativo la variabilità può essere misurata mediante gli indici di Gini e di Frosini normalizzati

$$G_N = \frac{1 - \sum_{i=1}^h f_{i\bullet}^2}{1 - \frac{1}{h}}$$

e

$$F_N = 1 - \sqrt{1 - G_N}.$$

Si ottiene:

x_i	$n_{i\bullet}$	$f_{i\bullet}$	$f_{i\bullet}^2$
<i>maschio</i>	5	0.5556	0.3086
<i>femmina</i>	4	0.4444	0.1975
	9	1	0.5061

Da cui segue:

$$G_N = \frac{1 - 0.5062}{1 - \frac{1}{2}} = 0.9877$$

e

$$F_N = 1 - \sqrt{1 - 0.9877} = 0.8889.$$

Sussiste, quindi, un livello alto di eterogeneità.



$\text{Moda}(X) = \text{maschio}$; $F_N(X) = 0.8889$. Sussiste un elevato livello di eterogeneità. La moda non è rappresentativa della distribuzione.

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.



3. Si confrontino le distribuzioni condizionate $Y|x_1$ e $Y|x_2$ con opportuni indici di posizione e di variabilità.

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.



Si considera la distribuzione condizionata $Y|x_1$

y_j	n_{1j}
0 + 2	0
2 + 3	0
3 + 8	5

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.



Si tratta di un carattere quantitativo con modalità raggruppate in classi.

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.



Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.

Si osserva che la distribuzione è degenere.

$M(Y|x_1) = 5.5$. La varianza e il coefficiente di variazione sono nulli.



Si considera la distribuzione condizionata $Y|x_2$

y_j	n_{2j}
0 + 2	2
2 + 3	2
3 + 8	0

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.



Si tratta di un carattere quantitativo con modalità raggruppate in classi.

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.



La media aritmetica

$$M(Y|x_2) = \frac{1}{n_{2\bullet}} \sum_{j=1}^k y_j n_{2j}$$

risulta 1.75.

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.



La media aritmetica

$$M(Y|x_2) = \frac{1}{n_{2\bullet}} \sum_{j=1}^k y_j n_{2j}$$

risulta 1.75.

Si ottiene (considerando come modalità i valori centrali delle classi)

y_j	n_{2j}	$y_j n_{2j}$
1.0	2	2
2.5	2	5
5.5	0	0
	4	7

da cui segue

$$M(Y|x_2) = \frac{1}{n_{2\bullet}} \sum_{j=1}^k y_j n_{2j} = \frac{7}{4} = 1.75.$$

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.



La varianza

$$\text{Var}(Y|x_2) = M(Y^2|x_2) - [M(Y|x_2)]^2$$

risulta 0.5625.

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.



Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.

La varianza

$$\text{Var}(Y|x_2) = M(Y^2|x_2) - [M(Y|x_2)]^2$$

risulta 0.5625. Si ottiene (considerando come modalità i valori centrali delle classi)

y_j	n_{2j}	$y_j n_{2j}$	$y_j^2 n_{2j}$
1.0	2	2	2.0
2.5	2	5	12.5
5.5	0	0	0.0
	4	7	14.5

da cui segue

$$\text{Var}(Y|x_2) = \frac{14.5}{4} - 1.75^2 = 3.625 - 3.0625 = 0.5625.$$



Per effettuare confronti in termini di variabilità tra caratteri quantitativi occorre utilizzare una misura relativa, ad esempio il coefficiente di variazione

$$CV = \frac{\sigma}{\mu} = \frac{\sqrt{\text{Var}(Y|x_2)}}{M(Y|x_2)} = \frac{\sqrt{0.5625}}{1.75} = \frac{0.75}{1.75} = 0.4286.$$

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.



$Y|x_1$ Indice di posizione: $M(Y|x_1) = 5.5$ Indice di variabilità: $CV(Y|x_1) = 0$

$Y|x_2$ Indice di posizione: $M(Y|x_2) = 1.75$ Indice di variabilità: $CV(Y|x_2) = 0.4286$

Nel campione considerato si osserva un numero medio di visite specialistiche per i maschi superiore a quello delle femmine. La distribuzione del numero medio di visite specialistiche per le femmine presenta un livello di variabilità superiore a quello dei maschi.

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.



4. Senza fare calcoli e giustificando la risposta, si indichi il valore dell'indice χ^2 (assoluto) e dell'indice χ_N normalizzato.

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.



4. Senza fare calcoli e giustificando la risposta, si indichi il valore dell'indice χ^2 (assoluto) e dell'indice χ_N normalizzato.

$\chi^2 = \chi_{\max}^2 = 9$ e $\chi_N = 1$. Sussiste dipendenza funzionale di X da Y essendo le distribuzioni condizionate $X|y_j$ degeneri.

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.

Esercizio 2



Il prospetto seguente riporta i prezzi medi annui di 2 beni di uso quotidiano nel periodo 2012-2016:

t	$bene A$	$bene B$
2012	9	27
2013	10	20
2014	12	27
2015	14	35
2016	15	30

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.

1. Costruire la serie dei numeri indici a base fissa dei prezzi con base 2013 per il bene A e la serie dei numeri indici a base mobile dei prezzi per il bene B .



Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.



1. Costruire la serie dei numeri indici a base fissa dei prezzi con base 2013 per il bene A e la serie dei numeri indici a base mobile dei prezzi per il bene B .

Si ricorda che ${}_{2013}I_t = \frac{x_t}{x_{2013}}$.

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.



Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.

1. Costruire la serie dei numeri indici a base fissa dei prezzi con base 2013 per il bene *A* e la serie dei numeri indici a base mobile dei prezzi per il bene *B*.

Si ricorda che ${}_{2013}I_t = \frac{x_t}{x_{2013}}$. Si ottiene quindi

Anno	x_t	NIBF (2013)
2012	9	0.9
2013	10	1.0
2014	12	1.2
2015	14	1.4
2016	15	1.5



Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.

1. Costruire la serie dei numeri indici a base fissa dei prezzi con base 2013 per il bene A e la serie dei numeri indici a base mobile dei prezzi per il bene B .

Si ricorda che ${}_{2013}I_t = \frac{x_t}{x_{2013}}$. Si ottiene quindi

Anno	x_t	NIBF (2013)
2012	9	0.9
2013	10	1.0
2014	12	1.2
2015	14	1.4
2016	15	1.5

Si ricorda che ${}_{t-1}I_t = \frac{x_t}{x_{t-1}}$.



Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.

1. Costruire la serie dei numeri indici a base fissa dei prezzi con base 2013 per il bene A e la serie dei numeri indici a base mobile dei prezzi per il bene B .

Si ricorda che ${}_{2013}I_t = \frac{x_t}{x_{2013}}$. Si ottiene quindi

Anno	x_t	NIBF (2013)
2012	9	0.9
2013	10	1.0
2014	12	1.2
2015	14	1.4
2016	15	1.5

Si ricorda che ${}_{t-1}I_t = \frac{x_t}{x_{t-1}}$. Si ottiene quindi

Anno	x_t	NIBM
2012	27	
2013	20	0.7407
2014	27	1.3500
2015	35	1.2963
2016	30	0.8571



2. Calcolare il tasso medio di variazione per il prezzo del bene A tra il 2013 e il 2016, interpretando il risultato.

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.



2. Calcolare il tasso medio di variazione per il prezzo del bene A tra il 2013 e il 2016, interpretando il risultato.

Si ricorda che

$${}_{2013}I_{2016} = {}_{2013}I_{2014} \cdot {}_{2014}I_{2015} \cdot {}_{2015}I_{2016}$$

Si deve trovare il numero indice a base mobile medio.

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.



Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.

2. Calcolare il tasso medio di variazione per il prezzo del bene A tra il 2013 e il 2016, interpretando il risultato.

Si ricorda che

$${}_{2013}I_{2016} = {}_{2013}I_{2014} \cdot {}_{2014}I_{2015} \cdot {}_{2015}I_{2016}$$

Si deve trovare il numero indice a base mobile medio. Applicando il criterio di scelta della media secondo Chisini si ottiene

$${}_{2013}I_{2016} = {}_{2013}I_{2014} \cdot {}_{2014}I_{2015} \cdot {}_{2015}I_{2016} = \alpha \cdot \alpha \cdot \alpha = \alpha^3.$$



Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.

2. Calcolare il tasso medio di variazione per il prezzo del bene A tra il 2013 e il 2016, interpretando il risultato.

Si ricorda che

$${}_{2013}I_{2016} = {}_{2013}I_{2014} \cdot {}_{2014}I_{2015} \cdot {}_{2015}I_{2016}$$

Si deve trovare il numero indice a base mobile medio. Applicando il criterio di scelta della media secondo Chisini si ottiene

$${}_{2013}I_{2016} = {}_{2013}I_{2014} \cdot {}_{2014}I_{2015} \cdot {}_{2015}I_{2016} = \alpha \cdot \alpha \cdot \alpha = \alpha^3.$$

Quindi

$$\alpha = \sqrt[3]{{}_{2013}I_{2016}} = \sqrt[3]{{}_{2013}I_{2014} \cdot {}_{2014}I_{2015} \cdot {}_{2015}I_{2016}} = 1.1447.$$

Da cui segue una variazione relativa media pari a 0.1447, corrispondente a un incremento medio annuo tra il 2013 e il 2016 del 14.47%.



3. Nota la quantità ${}_Aq_{2013} = 150$, si determini la quantità ${}_Bq_{2013}$ in modo che l'indice dei prezzi di Laspeyres al tempo 2016 con base 2013 risulti pari a 1.5.

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.



Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.

3. Nota la quantità ${}_Aq_{2013} = 150$, si determini la quantità ${}_Bq_{2013}$ in modo che l'indice dei prezzi di Laspeyres al tempo 2016 con base 2013 risulti pari a 1.5.

Dal momento che i prezzi relativi riferiti ai prodotti A e B risultano

$$\frac{{}_Ap_{2016}}{{}_Ap_{2013}} = \frac{15}{10} = 1.5 \quad \text{e} \quad \frac{{}_Bp_{2016}}{{}_Bp_{2013}} = \frac{30}{20} = 1.5$$

la corrispondente variabile statistica risulta degenerare e, quindi, ${}^p_{2013}I^L_{2016} = 1.5$ per qualsiasi quantità ${}_Bq_{2013} \geq 0$.



Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.

Il seguente prospetto riporta i valori rilevati in corrispondenza a 5 unità statistiche con riferimento a due caratteri quantitativi X e Y per i quali si vuole studiare la dipendenza di Y da X :

x_i	1	1	3	5	10
y_i	10	10	21	30	50

1. Si rappresentino graficamente le coppie dei dati e la funzione di regressione di Y in funzione di X .



La funzione di regressione associa a ogni valore x appartenente all'insieme di definizione della variabile X la media di Y condizionata a x , $M(Y|x)$

x	$M(Y x)$
1	10
3	21
5	30
10	50

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.



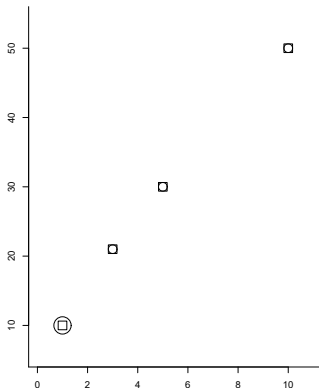
Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.

Le coppie di dati possono essere rappresentate mediante un *bubble plot diagram*.





2. Si calcolino, secondo il metodo dei minimi quadrati, i parametri dei seguenti modelli di regressione:

$$I) Y^* = a + b \cdot X \quad II) Y^* = c \cdot X$$

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.

Si riportano i calcoli necessari¹ per ottenere le stime dei parametri e gli indici di adattamento del modello

$$Y = a + bX + E$$

x_i	y_i	$x_i y_i$	x_i^2	y_i^2
1	10	10	1	100
1	10	10	1	100
3	21	63	9	441
5	30	150	25	900
10	50	500	100	2500
20	121	733	136	4041
4	24.2	146.6	27.2	808.2
$M(X)$	$M(Y)$	$M(XY)$	$M(X^2)$	$M(Y^2)$

$$\hat{b} =$$



Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.

¹Si osserva che per ragioni di stampa alcuni calcoli potrebbero essere duplicati con riferimento ai diversi modelli oggetto di studio

Si riportano i calcoli necessari¹ per ottenere le stime dei parametri e gli indici di adattamento del modello

$$Y = a + bX + E$$

x_i	y_i	$x_i y_i$	x_i^2	y_i^2
1	10	10	1	100
1	10	10	1	100
3	21	63	9	441
5	30	150	25	900
10	50	500	100	2500
20	121	733	136	4041
4	24.2	146.6	27.2	808.2
$M(X)$	$M(Y)$	$M(XY)$	$M(X^2)$	$M(Y^2)$

$$\begin{aligned}\hat{b} &= \frac{\text{Cov}(X, Y)}{\text{Var}(X)} = \frac{M(XY) - M(X)M(Y)}{M(X^2) - [M(X)]^2} = \\ &= \frac{146.6 - 4 \cdot 24.2}{27.2 - 4^2} = \frac{49.8}{11.2} = 4.4464 \\ \hat{a} &= \end{aligned}$$



Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.

¹Si osserva che per ragioni di stampa alcuni calcoli potrebbero essere duplicati con riferimento ai diversi modelli oggetto di studio

Si riportano i calcoli necessari¹ per ottenere le stime dei parametri e gli indici di adattamento del modello

$$Y = a + bX + E$$

x_i	y_i	$x_i y_i$	x_i^2	y_i^2
1	10	10	1	100
1	10	10	1	100
3	21	63	9	441
5	30	150	25	900
10	50	500	100	2500
20	121	733	136	4041
4	24.2	146.6	27.2	808.2
$M(X)$	$M(Y)$	$M(XY)$	$M(X^2)$	$M(Y^2)$

$$\hat{b} = \frac{\text{Cov}(X, Y)}{\text{Var}(X)} = \frac{M(XY) - M(X)M(Y)}{M(X^2) - [M(X)]^2} =$$

$$= \frac{146.6 - 4 \cdot 24.2}{27.2 - 4^2} = \frac{49.8}{11.2} = 4.4464$$

$$\hat{a} = M(Y) - \hat{b} M(X) = 6.4143$$

$$\rho^2 =$$



Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.

¹Si osserva che per ragioni di stampa alcuni calcoli potrebbero essere duplicati con riferimento ai diversi modelli oggetto di studio

Si riportano i calcoli necessari¹ per ottenere le stime dei parametri e gli indici di adattamento del modello

$$Y = a + bX + E$$

x_i	y_i	$x_i y_i$	x_i^2	y_i^2
1	10	10	1	100
1	10	10	1	100
3	21	63	9	441
5	30	150	25	900
10	50	500	100	2500
20	121	733	136	4041
4	24.2	146.6	27.2	808.2
$M(X)$	$M(Y)$	$M(XY)$	$M(X^2)$	$M(Y^2)$

$$\hat{b} = \frac{\text{Cov}(X, Y)}{\text{Var}(X)} = \frac{M(XY) - M(X)M(Y)}{M(X^2) - [M(X)]^2} =$$

$$= \frac{146.6 - 4 \cdot 24.2}{27.2 - 4^2} = \frac{49.8}{11.2} = 4.4464$$

$$\hat{a} = M(Y) - \hat{b} M(X) = 6.4143$$

$$\rho^2 = \frac{[\text{Cov}(X, Y)]^2}{\text{Var}(X) \text{Var}(Y)} = \frac{49.8^2}{11.2 \cdot 222.56} = 0.9949$$

$$M(\hat{E}^2) = \text{Var}(\hat{E}) = \text{varianza residua} =$$

¹Si osserva che per ragioni di stampa alcuni calcoli potrebbero essere duplicati con riferimento ai diversi modelli oggetto di studio



Si riportano i calcoli necessari¹ per ottenere le stime dei parametri e gli indici di adattamento del modello

$$Y = a + bX + E$$

x_i	y_i	$x_i y_i$	x_i^2	y_i^2
1	10	10	1	100
1	10	10	1	100
3	21	63	9	441
5	30	150	25	900
10	50	500	100	2500
20	121	733	136	4041
4	24.2	146.6	27.2	808.2
$M(X)$	$M(Y)$	$M(XY)$	$M(X^2)$	$M(Y^2)$

$$\hat{b} = \frac{\text{Cov}(X, Y)}{\text{Var}(X)} = \frac{M(XY) - M(X)M(Y)}{M(X^2) - [M(X)]^2} =$$

$$= \frac{146.6 - 4 \cdot 24.2}{27.2 - 4^2} = \frac{49.8}{11.2} = 4.4464$$

$$\hat{a} = M(Y) - \hat{b} M(X) = 6.4143$$

$$\rho^2 = \frac{[\text{Cov}(X, Y)]^2}{\text{Var}(X) \text{Var}(Y)} = \frac{49.8^2}{11.2 \cdot 222.56} = 0.9949$$

$$M(\hat{E}^2) = \text{Var}(\hat{E}) = \text{varianza residua} = \text{Var}(Y) \cdot (1 - \rho^2) = 1.1279$$

¹Si osserva che per ragioni di stampa alcuni calcoli potrebbero essere duplicati con riferimento ai diversi modelli oggetto di studio





Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.

Si riportano i calcoli necessari² per ottenere le stime dei parametri e la misura di adattamento del modello

$$Y = cX + E$$

x_i	y_i	$x_i y_i$	x_i^2	y_i^2
1	10	10	1	100
1	10	10	1	100
3	21	63	9	441
5	30	150	25	900
10	50	500	100	2500
		733	136	4041
		146.6	27.2	808.2
		$M(XY)$	$M(X^2)$	$M(Y^2)$

$$\hat{c} =$$

²Si osserva che per ragioni di stampa alcuni calcoli potrebbero essere duplicati con riferimento ai diversi modelli oggetto di studio



Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.

Si riportano i calcoli necessari² per ottenere le stime dei parametri e la misura di adattamento del modello

$$Y = cX + E$$

x_i	y_i	$x_i y_i$	x_i^2	y_i^2
1	10	10	1	100
1	10	10	1	100
3	21	63	9	441
5	30	150	25	900
10	50	500	100	2500
		733	136	4041
		146.6	27.2	808.2
		$M(XY)$	$M(X^2)$	$M(Y^2)$

$$\hat{c} = \frac{M(XY)}{M(X^2)} = \frac{146.6}{27.2} = 5.3897$$

$$M(\hat{E}^2) =$$

²Si osserva che per ragioni di stampa alcuni calcoli potrebbero essere duplicati con riferimento ai diversi modelli oggetto di studio



Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.

Si riportano i calcoli necessari² per ottenere le stime dei parametri e la misura di adattamento del modello

$$Y = cX + E$$

x_i	y_i	$x_i y_i$	x_i^2	y_i^2
1	10	10	1	100
1	10	10	1	100
3	21	63	9	441
5	30	150	25	900
10	50	500	100	2500
		733	136	4041
		146.6	27.2	808.2
		$M(XY)$	$M(X^2)$	$M(Y^2)$

$$\hat{c} = \frac{M(XY)}{M(X^2)} = \frac{146.6}{27.2} = 5.3897$$

$$M(\hat{E}^2) = M[Y \cdot (Y - \hat{c}X)] = M(Y^2) - \hat{c}M(XY) = 808.2 - 5.3897 \cdot 146.6 = 18.0691$$

²Si osserva che per ragioni di stampa alcuni calcoli potrebbero essere duplicati con riferimento ai diversi modelli oggetto di studio



$\hat{a} = 6.4143$, $\hat{b} = 4.4464$, indice adattamento $M(\hat{E}^2) = 1.1279$ $\hat{c} = 5.3897$, indice adattamento $M(\hat{E}^2) = 18.0691$

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.



3. Si indichi quale modello presenta il migliore adattamento ai dati.

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.



Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.

3. Si indichi quale modello presenta il migliore adattamento ai dati.

Il modello migliore è $Y^* = 6.4143 + 4.4464 \cdot X$ in quanto caratterizzato da residuo quadratico medio più basso.



4. Si dica, motivando teoricamente la risposta, se $\eta_{Y|X}^2 \neq \eta_{X|Y}^2$.

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.



Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.

4. Si dica, motivando teoricamente la risposta, se $\eta_{Y|X}^2 \neq \eta_{X|Y}^2$.

$\eta_{Y|X}^2 = \eta_{X|Y}^2 = 1$ in quanto sussiste dipendenza funzionale bigettiva.

Esercizio 4

Prova scritta di
STATISTICA del
10.07.2018

Giuseppe Boari
Gabriele Cantaluppi



UNIVERSITÀ
CATTOLICA
del Sacro Cuore

Un'urna è composta da 6 gettoni gialli e 4 gettoni blu. Si consideri l'esperimento casuale 'estrazione con reimmissione di 2 gettoni dall'urna'.

1. Si costruisca lo spazio probabilistico Ω associato all'esperimento.

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.

Esercizio 4

Prova scritta di
STATISTICA del
10.07.2018

Giuseppe Boari
Gabriele Cantaluppi



UNIVERSITÀ
CATTOLICA
del Sacro Cuore

Un'urna è composta da 6 gettoni gialli e 4 gettoni blu. Si consideri l'esperimento casuale 'estrazione con reimmissione di 2 gettoni dall'urna'.

1. Si costruisca lo spazio probabilistico Ω associato all'esperimento.

Indicati con G_i e B_i gli eventi estrazione di gettone giallo e blu nella i -esima estrazione ($i = 1, 2$), lo spazio probabilistico risulta composto dai seguenti eventi elementari:

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.



Un'urna è composta da 6 gettoni gialli e 4 gettoni blu. Si consideri l'esperimento casuale 'estrazione con reimmissione di 2 gettoni dall'urna'.

1. Si costruisca lo spazio probabilistico Ω associato all'esperimento.

Indicati con G_i e B_i gli eventi estrazione di gettone giallo e blu nella i -esima estrazione ($i = 1, 2$), lo spazio probabilistico risulta composto dai seguenti eventi elementari:

$$\Omega = \{G_1 G_2, G_1 B_2, B_1 G_2, B_1 B_2\},$$

a cui corrispondono le probabilità



Un'urna è composta da 6 gettoni gialli e 4 gettoni blu. Si consideri l'esperimento casuale 'estrazione con reimmissione di 2 gettoni dall'urna'.

1. Si costruisca lo spazio probabilistico Ω associato all'esperimento.

Indicati con G_i e B_i gli eventi estrazione di gettone giallo e blu nella i -esima estrazione ($i = 1, 2$), lo spazio probabilistico risulta composto dai seguenti eventi elementari:

$$\Omega = \{G_1 G_2, G_1 B_2, B_1 G_2, B_1 B_2\},$$

a cui corrispondono le probabilità

$$\{0.6^2 = 0.36, 0.6 \cdot 0.4 = 0.24, 0.24, 0.4^2 = 0.16\}$$

tenendo conto che le estrazioni sono con reimmissione (prove indipendenti) e che $P(G_i) = P(G) = 0.6$, $P(B_i) = P(B) = 0.4$, $i = 1, 2$.



2. Un gioco, per cui è prevista una posta di 1 euro, consiste nell' 'estrazione con reimmissione di 2 gettoni dall'urna' e comporta l'obbligo di pagare altri 4 euro per ogni gettone giallo che si presenta, mentre si vincono 8 euro per ogni gettone blu che si presenta. Si costruisca la variabile casuale G = 'guadagno del giocatore' e si stabilisca se il gioco è equo.

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.



Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.

2. Un gioco, per cui è prevista una posta di 1 euro, consiste nell' 'estrazione con reimmissione di 2 gettoni dall'urna' e comporta l'obbligo di pagare altri 4 euro per ogni gettone giallo che si presenta, mentre si vincono 8 euro per ogni gettone blu che si presenta. Si costruisca la variabile casuale $G =$ 'guadagno del giocatore' e si stabilisca se il gioco è equo.

La variabile casuale guadagno



2. Un gioco, per cui è prevista una posta di 1 euro, consiste nell' 'estrazione con reimmissione di 2 gettoni dall'urna' e comporta l'obbligo di pagare altri 4 euro per ogni gettone giallo che si presenta, mentre si vincono 8 euro per ogni gettone blu che si presenta. Si costruisca la variabile casuale G = 'guadagno del giocatore' e si stabilisca se il gioco è equo.

La variabile casuale guadagno

<i>evento</i>	x_i	$P(X = x_i)$	$x_i \cdot P(X = x_i)$
$2G$	-9	0.36	-3.24
$1G1B$	3	0.48	1.44
$2B$	15	0.16	2.40
		1.00	0.60

ha valore atteso positivo pari a 0.60



2. Un gioco, per cui è prevista una posta di 1 euro, consiste nell' 'estrazione con reimmissione di 2 gettoni dall'urna' e comporta l'obbligo di pagare altri 4 euro per ogni gettone giallo che si presenta, mentre si vincono 8 euro per ogni gettone blu che si presenta. Si costruisca la variabile casuale G = 'guadagno del giocatore' e si stabilisca se il gioco è equo.

La variabile casuale guadagno

<i>evento</i>	x_i	$P(X = x_i)$	$x_i \cdot P(X = x_i)$
$2G$	-9	0.36	-3.24
$1G1B$	3	0.48	1.44
$2B$	15	0.16	2.40
		1.00	0.60

ha valore atteso positivo pari a 0.60 e, quindi, il gioco è favorevole al giocatore.



3. Si calcoli la probabilità di ottenere almeno 52 gettoni gialli in 100 estrazioni con reimmissione dall'urna.

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.



3. Si calcoli la probabilità di ottenere almeno 52 gettoni gialli in 100 estrazioni con reimmissione dall'urna.

Indicata con X la variabile casuale 'numero di gettoni gialli' nelle 100 estrazioni con reimmissione dall'urna risulta $X \sim$

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.



Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.

3. Si calcoli la probabilità di ottenere almeno 52 gettoni gialli in 100 estrazioni con reimmissione dall'urna.

Indicata con X la variabile casuale 'numero di gettoni gialli' nelle 100 estrazioni con reimmissione dall'urna risulta $X \sim \text{Bin}(n = 100, p = 0.6)$. $M(X) = 60$; $\text{Var}(X) = 24$. Dal momento che $np = 60 > 5$ ed $n(1 - p) = 40 > 5$ è possibile utilizzare l'approssimazione della variabile casuale Binomiale con la variabile casuale Normale

$$X \sim N(\mu = 60, \sigma^2 = 24).$$

Si ottiene



3. Si calcoli la probabilità di ottenere almeno 52 gettoni gialli in 100 estrazioni con reimmissione dall'urna.

Indicata con X la variabile casuale 'numero di gettoni gialli' nelle 100 estrazioni con reimmissione dall'urna risulta $X \sim \text{Bin}(n = 100, p = 0.6)$. $M(X) = 60$; $\text{Var}(X) = 24$. Dal momento che $np = 60 > 5$ ed $n(1 - p) = 40 > 5$ è possibile utilizzare l'approssimazione della variabile casuale Binomiale con la variabile casuale Normale

$$X \sim N(\mu = 60, \sigma^2 = 24).$$

Si ottiene

$$P(X > 52) =$$



3. Si calcoli la probabilità di ottenere almeno 52 gettoni gialli in 100 estrazioni con reimmissione dall'urna.

Indicata con X la variabile casuale 'numero di gettoni gialli' nelle 100 estrazioni con reimmissione dall'urna risulta $X \sim \text{Bin}(n = 100, p = 0.6)$. $M(X) = 60$; $\text{Var}(X) = 24$. Dal momento che $np = 60 > 5$ ed $n(1 - p) = 40 > 5$ è possibile utilizzare l'approssimazione della variabile casuale Binomiale con la variabile casuale Normale

$$X \sim N(\mu = 60, \sigma^2 = 24).$$

Si ottiene

$$P(X > 52) = P\left(\frac{X - \mu}{\sigma} > \frac{52 - \mu}{\sigma}\right) =$$



3. Si calcoli la probabilità di ottenere almeno 52 gettoni gialli in 100 estrazioni con reimmissione dall'urna.

Indicata con X la variabile casuale 'numero di gettoni gialli' nelle 100 estrazioni con reimmissione dall'urna risulta $X \sim \text{Bin}(n = 100, p = 0.6)$. $M(X) = 60$; $\text{Var}(X) = 24$. Dal momento che $np = 60 > 5$ ed $n(1 - p) = 40 > 5$ è possibile utilizzare l'approssimazione della variabile casuale Binomiale con la variabile casuale Normale

$$X \sim N(\mu = 60, \sigma^2 = 24).$$

Si ottiene

$$P(X > 52) = P\left(\frac{X - \mu}{\sigma} > \frac{52 - \mu}{\sigma}\right) = P\left(Z > \frac{52 - 60}{4.899}\right) =$$

=



3. Si calcoli la probabilità di ottenere almeno 52 gettoni gialli in 100 estrazioni con reimmissione dall'urna.

Indicata con X la variabile casuale 'numero di gettoni gialli' nelle 100 estrazioni con reimmissione dall'urna risulta $X \sim \text{Bin}(n = 100, p = 0.6)$. $M(X) = 60$; $\text{Var}(X) = 24$. Dal momento che $np = 60 > 5$ ed $n(1 - p) = 40 > 5$ è possibile utilizzare l'approssimazione della variabile casuale Binomiale con la variabile casuale Normale

$$X \sim N(\mu = 60, \sigma^2 = 24).$$

Si ottiene

$$\begin{aligned} P(X > 52) &= P\left(\frac{X - \mu}{\sigma} > \frac{52 - \mu}{\sigma}\right) = P\left(Z > \frac{52 - 60}{4.899}\right) = \\ &= P(Z > -1.63) = \end{aligned}$$



3. Si calcoli la probabilità di ottenere almeno 52 gettoni gialli in 100 estrazioni con reimmissione dall'urna.

Indicata con X la variabile casuale 'numero di gettoni gialli' nelle 100 estrazioni con reimmissione dall'urna risulta $X \sim \text{Bin}(n = 100, p = 0.6)$. $M(X) = 60$; $\text{Var}(X) = 24$. Dal momento che $np = 60 > 5$ ed $n(1 - p) = 40 > 5$ è possibile utilizzare l'approssimazione della variabile casuale Binomiale con la variabile casuale Normale

$$X \sim N(\mu = 60, \sigma^2 = 24).$$

Si ottiene

$$\begin{aligned} P(X > 52) &= P\left(\frac{X - \mu}{\sigma} > \frac{52 - \mu}{\sigma}\right) = P\left(Z > \frac{52 - 60}{4.899}\right) = \\ &= P(Z > -1.63) = P(Z < 1.63) = \end{aligned}$$



3. Si calcoli la probabilità di ottenere almeno 52 gettoni gialli in 100 estrazioni con reimmissione dall'urna.

Indicata con X la variabile casuale 'numero di gettoni gialli' nelle 100 estrazioni con reimmissione dall'urna risulta $X \sim \text{Bin}(n = 100, p = 0.6)$. $M(X) = 60$; $\text{Var}(X) = 24$. Dal momento che $np = 60 > 5$ ed $n(1 - p) = 40 > 5$ è possibile utilizzare l'approssimazione della variabile casuale Binomiale con la variabile casuale Normale

$$X \sim N(\mu = 60, \sigma^2 = 24).$$

Si ottiene

$$\begin{aligned} P(X > 52) &= P\left(\frac{X - \mu}{\sigma} > \frac{52 - \mu}{\sigma}\right) = P\left(Z > \frac{52 - 60}{4.899}\right) = \\ &= P(Z > -1.63) = P(Z < 1.63) = P(Z \leq 1.63) = \end{aligned}$$



3. Si calcoli la probabilità di ottenere almeno 52 gettoni gialli in 100 estrazioni con reimmissione dall'urna.

Indicata con X la variabile casuale 'numero di gettoni gialli' nelle 100 estrazioni con reimmissione dall'urna risulta $X \sim \text{Bin}(n = 100, p = 0.6)$. $M(X) = 60$; $\text{Var}(X) = 24$. Dal momento che $np = 60 > 5$ ed $n(1 - p) = 40 > 5$ è possibile utilizzare l'approssimazione della variabile casuale Binomiale con la variabile casuale Normale

$$X \sim N(\mu = 60, \sigma^2 = 24).$$

Si ottiene

$$\begin{aligned} P(X > 52) &= P\left(\frac{X - \mu}{\sigma} > \frac{52 - \mu}{\sigma}\right) = P\left(Z > \frac{52 - 60}{4.899}\right) = \\ &= P(Z > -1.63) = P(Z < 1.63) = P(Z \leq 1.63) = 0.9484. \end{aligned}$$

Tema T3001

Prova scritta di STATISTICA del 23.08.2018

Prova scritta di
STATISTICA del
23.08.2018

Giuseppe Boari
Gabriele Cantaluppi



UNIVERSITÀ
CATTOLICA
del Sacro Cuore

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.

Giuseppe Boari
Gabriele Cantaluppi
Università Cattolica del Sacro Cuore, Milano



La seguente tabella riporta le distribuzioni relative condizionate della variabile $Y = \text{Numero di ingressi al museo nell'ultimo mese}$ rispetto alla variabile $X = \text{Zona di provenienza del visitatore}$, ricavata a seguito di una indagine campionaria:

$Y X$	0 + 2	2 + 7	7 + 10	
<i>NORD</i>	0.2	0.6	0.2	1
<i>CENTRO</i>	0.2	0.6	0.2	1
<i>SUD</i>	0.2	0.6	0.2	1

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.



1. Senza effettuare calcoli, si dica se è possibile concludere che esiste indipendenza stocastica tra X e Y . Motivare la risposta.

Tra le variabili X e Y sussiste indipendenza stocastica in quanto le distribuzioni condizionate $Y|x_i$ sono tra loro somiglianti. Abbiamo, infatti, $Y|x_1 \sim Y|x_2 \sim Y|x_3$.

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.



2. Costruire la tabella delle frequenze congiunte sapendo che 50 dei visitatori intervistati provengono dal *NORD*, 100 dal *CENTRO* e 150 dal *SUD*.

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.



2. Costruire la tabella delle frequenze congiunte sapendo che 50 dei visitatori intervistati provengono dal *NORD*, 100 dal *CENTRO* e 150 dal *SUD*.

Ricordando che per la generica frequenza condizionata relativa vale $f(y_j|x_i) = \frac{n_{ij}}{n_{i\bullet}}$ abbiamo che $n_{ij} = f(y_j|x_i) \cdot n_{i\bullet}$ e

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.



2. Costruire la tabella delle frequenze congiunte sapendo che 50 dei visitatori intervistati provengono dal *NORD*, 100 dal *CENTRO* e 150 dal *SUD*.

Ricordando che per la generica frequenza condizionata relativa vale $f(y_j|x_i) = \frac{n_{ij}}{n_{i\bullet}}$ abbiamo che $n_{ij} = f(y_j|x_i) \cdot n_{i\bullet}$ e

$X \setminus Y$	0 - 2	2 - 7	7 - 10	$n_{i\bullet}$
<i>NORD</i>	10	30	10	50
<i>CENTRO</i>	20	60	20	100
<i>SUD</i>	30	90	30	150
$n_{\bullet j}$	60	180	60	300

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.



Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.

3. Si rappresenti graficamente la variabile X e se ne calcoli moda ed eterogeneità.



Si considera la distribuzione marginale della variabile X , ottenuta sommando le frequenze di riga

x_i	$n_{i\bullet}$
<i>NORD</i>	50
<i>CENTRO</i>	100
<i>SUD</i>	150

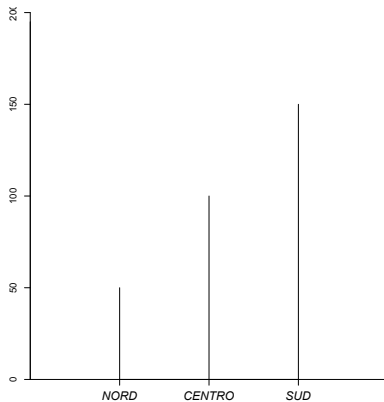
Si tratta di un carattere qualitativo sconnesso che può essere rappresentato graficamente mediante un grafico a bastoncini o a canne d'organo.

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.



Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.



Si considera la distribuzione marginale della variabile X , ottenuta sommando le frequenze di riga

x_i	$n_{i\bullet}$
<i>NORD</i>	50
<i>CENTRO</i>	100
<i>SUD</i>	150

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.



Si tratta di un carattere qualitativo sconnesso.

L'unico indice di posizione che può essere definito è la moda (categoria alla quale è associata la massima frequenza) che risulta *SUD*.

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.

In presenza di un carattere qualitativo la variabilità può essere misurata mediante gli indici di Gini e di Frosini normalizzati

$$G_N = \frac{1 - \sum_{i=1}^h f_{i\bullet}^2}{1 - \frac{1}{h}}$$

e

$$F_N = 1 - \sqrt{1 - G_N}.$$



Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.



In presenza di un carattere qualitativo la variabilità può essere misurata mediante gli indici di Gini e di Frosini normalizzati

$$G_N = \frac{1 - \sum_{i=1}^h f_{i\bullet}^2}{1 - \frac{1}{h}}$$

e

$$F_N = 1 - \sqrt{1 - G_N}.$$

Si ottiene:

x_i	$n_{i\bullet}$	$f_{i\bullet}$	$f_{i\bullet}^2$
<i>NORD</i>	50	0.1667	0.0278
<i>CENTRO</i>	100	0.3333	0.1111
<i>SUD</i>	150	0.5	0.25
	300	1	0.3889



In presenza di un carattere qualitativo la variabilità può essere misurata mediante gli indici di Gini e di Frosini normalizzati

$$G_N = \frac{1 - \sum_{i=1}^h f_{i\bullet}^2}{1 - \frac{1}{h}}$$

e

$$F_N = 1 - \sqrt{1 - G_N}.$$

Si ottiene:

x_i	$n_{i\bullet}$	$f_{i\bullet}$	$f_{i\bullet}^2$
<i>NORD</i>	50	0.1667	0.0278
<i>CENTRO</i>	100	0.3333	0.1111
<i>SUD</i>	150	0.5	0.25
	300	1	0.3889

Da cui segue:

$$G_N = \frac{1 - 0.3889}{1 - \frac{1}{3}} = 0.9167$$

e

$$F_N = 1 - \sqrt{1 - 0.9167} = 0.7113.$$

Sussiste, quindi, un livello medio alto di eterogeneità.



Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.

$$\text{Moda}(X) = SUD \quad F_N = 0.7113$$



4. Si confrontino le distribuzioni condizionate $Y|X = \text{NORD}$ e $Y|X = \text{SUD}$ con opportuni indici di posizione e di variabilità.

Dal momento che sussiste indipendenza stocastica le distribuzioni sono caratterizzate dal medesimo valore medio e dallo stesso livello di variabilità. Si riportano solo per eventuale verifica del lettore i calcoli per entrambe le distribuzioni condizionate.

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.



Si considera la distribuzione condizionata $Y|x_1$

y_j	n_{1j}
0 + 2	10
2 + 7	30
7 + 10	10

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.



Si tratta di un carattere quantitativo con modalità raggruppate in classi.

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.



La media aritmetica

$$M(Y|x_1) = \frac{1}{n_{1\bullet}} \sum_{j=1}^k y_j n_{1j}$$

risulta 4.6.

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.



La media aritmetica

$$M(Y|x_1) = \frac{1}{n_{1\bullet}} \sum_{j=1}^k y_j n_{1j}$$

risulta 4.6.

Si ottiene (considerando come modalità i valori centrali delle classi)

y_j	n_{1j}	$y_j n_{1j}$
1.0	10	10
4.5	30	135
8.5	10	85
	50	230

da cui segue

$$M(Y|x_1) = \frac{1}{n_{1\bullet}} \sum_{j=1}^k y_j n_{1j} = \frac{230}{50} = 4.6.$$

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.



La varianza

risulta 5.64.

$$\text{Var}(Y|x_1) = M(Y^2|x_1) - [M(Y|x_1)]^2$$

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.



La varianza

$$Var(Y|x_1) = M(Y^2|x_1) - [M(Y|x_1)]^2$$

risulta 5.64. Si ottiene (considerando come modalità i valori centrali delle classi)

y_j	n_{1j}	$y_j n_{1j}$	$y_j^2 n_{1j}$
1.0	10	10	10.0
4.5	30	135	607.5
8.5	10	85	722.5
	50	230	1340.0

da cui segue

$$Var(Y|x_1) = \frac{1340}{50} - 4.6^2 = 26.8 - 21.16 = 5.64.$$



Per effettuare confronti in termini di variabilità tra caratteri quantitativi occorre utilizzare una misura relativa, ad esempio il coefficiente di variazione

$$CV = \frac{\sigma}{\mu} = \frac{\sqrt{\text{Var}(Y|x_1)}}{M(Y|x_1)} = \frac{\sqrt{5.64}}{4.6} = \frac{2.3749}{4.6} = 0.5163.$$

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.



Si considera la distribuzione condizionata $Y|x_3$

y_j	n_{3j}
0 + 2	30
2 + 7	90
7 + 10	30

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.



Si tratta di un carattere quantitativo con modalità raggruppate in classi.

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.



La media aritmetica

$$M(Y|x_3) = \frac{1}{n_{3\bullet}} \sum_{j=1}^k y_j n_{3j}$$

risulta 4.6.

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.



La media aritmetica

$$M(Y|x_3) = \frac{1}{n_{3\bullet}} \sum_{j=1}^k y_j n_{3j}$$

risulta 4.6.

Si ottiene (considerando come modalità i valori centrali delle classi)

y_j	n_{3j}	$y_j n_{3j}$
1.0	30	30
4.5	90	405
8.5	30	255
	150	690

da cui segue

$$M(Y|x_3) = \frac{1}{n_{3\bullet}} \sum_{j=1}^k y_j n_{3j} = \frac{690}{150} = 4.6.$$

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.



La varianza

$$\text{Var}(Y|x_3) = M(Y^2|x_3) - [M(Y|x_3)]^2$$

risulta 5.64.

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.



La varianza

$$Var(Y|x_3) = M(Y^2|x_3) - [M(Y|x_3)]^2$$

risulta 5.64. Si ottiene (considerando come modalità i valori centrali delle classi)

y_j	n_{3j}	$y_j n_{3j}$	$y_j^2 n_{3j}$
1.0	30	30	30.0
4.5	90	405	1822.5
8.5	30	255	2167.5
	150	690	4020.0

da cui segue

$$Var(Y|x_3) = \frac{4020}{150} - 4.6^2 = 26.8 - 21.16 = 5.64.$$



Per effettuare confronti in termini di variabilità tra caratteri quantitativi occorre utilizzare una misura relativa, ad esempio il coefficiente di variazione

$$CV = \frac{\sigma}{\mu} = \frac{\sqrt{\text{Var}(Y|x_3)}}{M(Y|x_3)} = \frac{\sqrt{5.64}}{4.6} = \frac{2.3749}{4.6} = 0.5163.$$

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.



$Y|X = NORD$ Indice di posizione: $M(Y|X = NORD) = 4.6$ Indice di variabilità:

$$CV(Y|X = NORD) = 0.5163$$

$Y|X = SUD$ Indice di posizione: $M(Y|X = SUD) = 4.6$ Indice di variabilità:

$$CV(Y|X = SUD) = 0.5163$$

Dal momento che sussiste indipendenza stocastica le distribuzioni sono caratterizzate dal medesimo valore medio e dallo stesso livello di variabilità.

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.

Esercizio 2



Si consideri il seguente prospetto:

t	$NIBF(base = 2014)$	$NIBM$
2012		
2013	0.89	1
2014		
2015	0.5	
2016		1

1. Si completi il prospetto con i dati mancanti.

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.



In primo luogo vale ${}_{2014}I_{2014} = 1$ e ${}_{2014}I_{2015} = 0.5$.

Considerando la serie dei numeri indici a base mobile assegnati si ricava

$${}_{2014}I_{2012} =$$

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.



In primo luogo vale ${}_{2014}I_{2014} = 1$ e ${}_{2014}I_{2015} = 0.5$.

Considerando la serie dei numeri indici a base mobile assegnati si ricava

$${}_{2014}I_{2012} = 0.89$$

$${}_{2014}I_{2016} =$$

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.



In primo luogo vale ${}_{2014}I_{2014} = 1$ e ${}_{2014}I_{2015} = 0.5$.

Considerando la serie dei numeri indici a base mobile assegnati si ricava

$${}_{2014}I_{2012} = 0.89$$

$${}_{2014}I_{2016} = 0.5$$

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.



Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.

In primo luogo vale ${}_{2014}I_{2014} = 1$ e ${}_{2014}I_{2015} = 0.5$.

Considerando la serie dei numeri indici a base mobile assegnati si ricava

$${}_{2014}I_{2012} = 0.89$$

$${}_{2014}I_{2016} = 0.5$$

Abbiamo infine

$${}_{2013}I_{2014} =$$



Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.

In primo luogo vale ${}_{2014}I_{2014} = 1$ e ${}_{2014}I_{2015} = 0.5$.

Considerando la serie dei numeri indici a base mobile assegnati si ricava

$${}_{2014}I_{2012} = 0.89$$

$${}_{2014}I_{2016} = 0.5$$

Abbiamo infine

$${}_{2013}I_{2014} = \frac{{}_{2014}I_{2014}}{{}_{2014}I_{2013}} = \frac{1}{{}_{2014}I_{2013}} = \frac{1}{0.89} = 1.1236$$



Si ottiene, quindi

t	$NIBF(base = 2014)$	$NIBM$
2012	0.89	
2013	0.89	1
2014	1	1.1236
2015	0.5	0.5
2016	0.5	1

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.



2. Si calcoli il numero indice al tempo 2013 con base 2016 e si commenti il risultato ottenuto.

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.



2. Si calcoli il numero indice al tempo 2013 con base 2016 e si commenti il risultato ottenuto.

$${}_{2016}I_{2013} = \frac{x_{2013}}{x_{2016}} = \frac{{}_{2014}I_{2013}}{{}_{2014}I_{2016}} = \frac{0.89}{0.5} = 1.78$$

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.



Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.

2. Si calcoli il numero indice al tempo 2013 con base 2016 e si commenti il risultato ottenuto.

$${}_{2016}I_{2013} = \frac{x_{2013}}{x_{2016}} = \frac{{}_{2014}I_{2013}}{{}_{2014}I_{2016}} = \frac{0.89}{0.5} = 1.78$$

La grandezza al tempo 2013 era superiore del 78% rispetto alla grandezza al tempo 2016. Ovvero il valore della grandezza nel 2013 era il 178% di quello registrato nel 2016.



3. Si calcoli il tasso medio di variazione dal 2012 al 2016 e si commenti il risultato ottenuto.

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.



3. Si calcoli il tasso medio di variazione dal 2012 al 2016 e si commenti il risultato ottenuto.
Si ricorda che

$${}_{2012}I_{2016} = {}_{2012}I_{2013} \cdot {}_{2013}I_{2014} \cdot {}_{2014}I_{2015} \cdot {}_{2015}I_{2016}$$

Si deve trovare il numero indice a base mobile medio.

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.



Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.

3. Si calcoli il tasso medio di variazione dal 2012 al 2016 e si commenti il risultato ottenuto. Si ricorda che

$${}_{2012}I_{2016} = {}_{2012}I_{2013} \cdot {}_{2013}I_{2014} \cdot {}_{2014}I_{2015} \cdot {}_{2015}I_{2016}$$

Si deve trovare il numero indice a base mobile medio. Applicando il criterio di scelta della media secondo Chisini si ottiene

$${}_{2012}I_{2016} = {}_{2012}I_{2013} \cdot {}_{2013}I_{2014} \cdot {}_{2014}I_{2015} \cdot {}_{2015}I_{2016} = \alpha \cdot \alpha \cdot \alpha \cdot \alpha = \alpha^4.$$



Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.

3. Si calcoli il tasso medio di variazione dal 2012 al 2016 e si commenti il risultato ottenuto. Si ricorda che

$${}_{2012}I_{2016} = {}_{2012}I_{2013} \cdot {}_{2013}I_{2014} \cdot {}_{2014}I_{2015} \cdot {}_{2015}I_{2016}$$

Si deve trovare il numero indice a base mobile medio. Applicando il criterio di scelta della media secondo Chisini si ottiene

$${}_{2012}I_{2016} = {}_{2012}I_{2013} \cdot {}_{2013}I_{2014} \cdot {}_{2014}I_{2015} \cdot {}_{2015}I_{2016} = \alpha \cdot \alpha \cdot \alpha \cdot \alpha = \alpha^4.$$

Quindi

$$\alpha = \sqrt[4]{{}_{2012}I_{2016}} = \sqrt[4]{{}_{2012}I_{2013} \cdot {}_{2013}I_{2014} \cdot {}_{2014}I_{2015} \cdot {}_{2015}I_{2016}} = 0.8658.$$

Da cui segue una variazione relativa media pari a -0.1342 , corrispondente a un decremento medio annuo tra il 2012 e il 2016 del 13.42%.



Il seguente prospetto riporta i valori rilevati in corrispondenza a 5 capofamiglia con riferimento a due caratteri quantitativi X = numero dei componenti della famiglia e Y = spesa mensile per musei e mostre (in euro).

x_i	1	3	3	3	5
y_i	21	31	44	21	11

1. Si rappresentino graficamente le coppie dei dati e la funzione di regressione ipotizzando la spesa in funzione del numero dei componenti della famiglia.

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.



La funzione di regressione associa a ogni valore x appartenente all'insieme di definizione della variabile X la media di Y condizionata a x , $M(Y|x)$

x	$M(Y x)$
1	21
3	$\frac{1}{3}(21 + 31 + 44) = 32$
5	11

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.



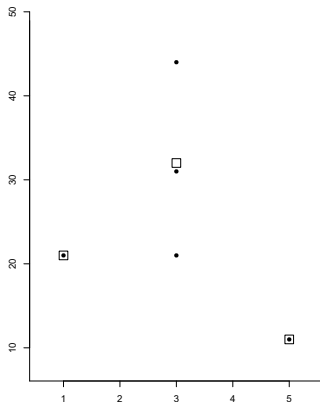
Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.

Le coppie di dati possono essere rappresentate mediante un grafico a dispersione.





2. Si calcolino, secondo il metodo dei minimi quadrati, i parametri dei seguenti modelli di regressione:

$$I) Y^* = a + b \cdot X \qquad II) Y^* = c$$

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.



Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.

Si riportano i calcoli necessari¹ per ottenere le stime dei parametri e gli indici di adattamento del modello

$$Y = a + bX + E$$

x_i	y_i	$x_i y_i$	x_i^2	y_i^2
1	21	21	1	441
3	31	93	9	961
3	44	132	9	1936
3	21	63	9	441
5	11	55	25	121
15	128	364	53	3900
3	25.6	72.8	10.6	780
$M(X)$	$M(Y)$	$M(XY)$	$M(X^2)$	$M(Y^2)$

$$\hat{b} =$$

¹Si osserva che per ragioni di stampa alcuni calcoli potrebbero essere duplicati con riferimento ai diversi modelli oggetto di studio



Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.

Si riportano i calcoli necessari¹ per ottenere le stime dei parametri e gli indici di adattamento del modello

$$Y = a + bX + E$$

x_i	y_i	$x_i y_i$	x_i^2	y_i^2
1	21	21	1	441
3	31	93	9	961
3	44	132	9	1936
3	21	63	9	441
5	11	55	25	121
15	128	364	53	3900
3	25.6	72.8	10.6	780
$M(X)$	$M(Y)$	$M(XY)$	$M(X^2)$	$M(Y^2)$

$$\begin{aligned}\hat{b} &= \frac{\text{Cov}(X, Y)}{\text{Var}(X)} = \frac{M(XY) - M(X)M(Y)}{M(X^2) - [M(X)]^2} = \\ &= \frac{72.8 - 3 \cdot 25.6}{10.6 - 3^2} = \frac{-4}{1.6} = -2.5 \\ \hat{a} &= \end{aligned}$$

¹Si osserva che per ragioni di stampa alcuni calcoli potrebbero essere duplicati con riferimento ai diversi modelli oggetto di studio



Si riportano i calcoli necessari¹ per ottenere le stime dei parametri e gli indici di adattamento del modello

$$Y = a + bX + E$$

x_i	y_i	$x_i y_i$	x_i^2	y_i^2
1	21	21	1	441
3	31	93	9	961
3	44	132	9	1936
3	21	63	9	441
5	11	55	25	121
15	128	364	53	3900
3	25.6	72.8	10.6	780
$M(X)$	$M(Y)$	$M(XY)$	$M(X^2)$	$M(Y^2)$

$$\begin{aligned}\hat{b} &= \frac{\text{Cov}(X, Y)}{\text{Var}(X)} = \frac{M(XY) - M(X)M(Y)}{M(X^2) - [M(X)]^2} = \\ &= \frac{72.8 - 3 \cdot 25.6}{10.6 - 3^2} = \frac{-4}{1.6} = -2.5 \\ \hat{a} &= M(Y) - \hat{b}M(X) = 33.1 \\ \rho^2 &= \end{aligned}$$

¹Si osserva che per ragioni di stampa alcuni calcoli potrebbero essere duplicati con riferimento ai diversi modelli oggetto di studio



Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.

Si riportano i calcoli necessari¹ per ottenere le stime dei parametri e gli indici di adattamento del modello

$$Y = a + bX + E$$

x_i	y_i	$x_i y_i$	x_i^2	y_i^2
1	21	21	1	441
3	31	93	9	961
3	44	132	9	1936
3	21	63	9	441
5	11	55	25	121
15	128	364	53	3900
3	25.6	72.8	10.6	780
$M(X)$	$M(Y)$	$M(XY)$	$M(X^2)$	$M(Y^2)$

$$\begin{aligned}\hat{b} &= \frac{\text{Cov}(X, Y)}{\text{Var}(X)} = \frac{M(XY) - M(X)M(Y)}{M(X^2) - [M(X)]^2} = \\ &= \frac{72.8 - 3 \cdot 25.6}{10.6 - 3^2} = \frac{-4}{1.6} = -2.5 \\ \hat{a} &= M(Y) - \hat{b}M(X) = 33.1 \\ \rho^2 &= \frac{[\text{Cov}(X, Y)]^2}{\text{Var}(X) \text{Var}(Y)} = \frac{(-4)^2}{1.6 \cdot 124.64} = 0.0802\end{aligned}$$

¹Si osserva che per ragioni di stampa alcuni calcoli potrebbero essere duplicati con riferimento ai diversi modelli oggetto di studio



Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.

Si riportano i calcoli necessari² per ottenere le stime dei parametri e gli indici di adattamento del modello

$$Y = c + E$$

x_i	y_i	y_i^2
1	21	441
3	31	961
3	44	1936
3	21	441
5	11	121
128		3900
25.6		780
$M(Y)$		$M(Y^2)$

$$\hat{c} =$$

²Si osserva che per ragioni di stampa alcuni calcoli potrebbero essere duplicati con riferimento ai diversi modelli oggetto di studio



Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.

Si riportano i calcoli necessari² per ottenere le stime dei parametri e gli indici di adattamento del modello

$$Y = c + E$$

x_i	y_i	y_i^2
1	21	441
3	31	961
3	44	1936
3	21	441
5	11	121
128		3900
25.6		780
$M(Y)$		$M(Y^2)$

$$\hat{c} = M(Y) = 25.6$$

$$R^2 =$$

²Si osserva che per ragioni di stampa alcuni calcoli potrebbero essere duplicati con riferimento ai diversi modelli oggetto di studio



Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.

Si riportano i calcoli necessari² per ottenere le stime dei parametri e gli indici di adattamento del modello

$$Y = c + E$$

x_i	y_i	y_i^2
1	21	441
3	31	961
3	44	1936
3	21	441
5	11	121
128		3900
25.6		780
$M(Y)$		$M(Y^2)$

$$\hat{c} = M(Y) = 25.6$$

$$R^2 = 0$$

²Si osserva che per ragioni di stampa alcuni calcoli potrebbero essere duplicati con riferimento ai diversi modelli oggetto di studio



Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.

$\hat{a} = 33.1$, $\hat{b} = -2.5$, indice adattamento $\rho^2 = 0.0802$ $\hat{c} = 25.6$, indice adattamento $R^2 = 0$



3. Confrontare gli indici di adattamento, commentando i risultati.

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.



3. Confrontare gli indici di adattamento, commentando i risultati.

Il modello migliore è $Y^* = 33.1 - 2.5 \cdot X$ in quanto caratterizzato da indice di adattamento più elevato.

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.



4. Senza effettuare calcoli e motivando la risposta, si stabilisca: a) l'intervallo di possibili valori assumibili da $\eta_{Y|X}^2$; b) se $\eta_{X|Y}^2 < 1$.

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.



Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.

4. Senza effettuare calcoli e motivando la risposta, si stabilisca: a) l'intervallo di possibili valori assumibili da $\eta_{Y|X}^2$; b) se $\eta_{X|Y}^2 < 1$.

$\rho^2 < \eta_{Y|X}^2 < 1$ in quanto le medie condizionate non sono allineate e non sussiste dipendenza funzionale di Y da X dal momento che la distribuzione condizionata $Y|x = 3$ non è degenera.



Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.

4. Senza effettuare calcoli e motivando la risposta, si stabilisca: a) l'intervallo di possibili valori assumibili da $\eta_{Y|X}^2$; b) se $\eta_{X|Y}^2 < 1$.

$\rho^2 < \eta_{Y|X}^2 < 1$ in quanto le medie condizionate non sono allineate e non sussiste dipendenza funzionale di Y da X dal momento che la distribuzione condizionata $Y|x = 3$ non è degenerare.

$\rho^2 < \eta_{X|Y}^2 < 1$ in quanto le medie condizionate non sono allineate e non sussiste dipendenza funzionale di X da Y dal momento che le distribuzioni condizionate $X|y = 21$ non è degenerare.

Esercizio 4

La retribuzione media mensile lorda di un laureato in Economia è distribuita come una Normale di media $\mu = 2300$ euro e varianza $\sigma^2 = 10\,000$ euro.

1. Quale percentuale di laureati guadagna meno di 2200 euro al mese?



Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.

Esercizio 4



Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.

La retribuzione media mensile lorda di un laureato in Economia è distribuita come una Normale di media $\mu = 2300$ euro e varianza $\sigma^2 = 10\,000$ euro.

1. Quale percentuale di laureati guadagna meno di 2200 euro al mese?

Indicata con $X \sim N(\mu = 2300, \sigma^2 = 10000)$ la variabile casuale che descrive la retribuzione media mensile lorda, occorre, in primo luogo, determinare la seguente probabilità:

$$P(X < 2200).$$

Abbiamo

$$P(X < 2200) =$$



La retribuzione media mensile lorda di un laureato in Economia è distribuita come una Normale di media $\mu = 2300$ euro e varianza $\sigma^2 = 10\,000$ euro.

1. Quale percentuale di laureati guadagna meno di 2200 euro al mese?

Indicata con $X \sim N(\mu = 2300, \sigma^2 = 10000)$ la variabile casuale che descrive la retribuzione media mensile lorda, occorre, in primo luogo, determinare la seguente probabilità:

$$P(X < 2200).$$

Abbiamo

$$P(X < 2200) = P\left(\frac{X - \mu}{\sigma} < \frac{2200 - \mu}{\sigma}\right) =$$



La retribuzione media mensile lorda di un laureato in Economia è distribuita come una Normale di media $\mu = 2300$ euro e varianza $\sigma^2 = 10\,000$ euro.

1. Quale percentuale di laureati guadagna meno di 2200 euro al mese?

Indicata con $X \sim N(\mu = 2300, \sigma^2 = 10000)$ la variabile casuale che descrive la retribuzione media mensile lorda, occorre, in primo luogo, determinare la seguente probabilità:

$$P(X < 2200).$$

Abbiamo

$$\begin{aligned} P(X < 2200) &= P\left(\frac{X - \mu}{\sigma} < \frac{2200 - \mu}{\sigma}\right) = P\left(Z < \frac{2200 - 2300}{100}\right) = \\ &= \end{aligned}$$



La retribuzione media mensile lorda di un laureato in Economia è distribuita come una Normale di media $\mu = 2300$ euro e varianza $\sigma^2 = 10\,000$ euro.

1. Quale percentuale di laureati guadagna meno di 2200 euro al mese?

Indicata con $X \sim N(\mu = 2300, \sigma^2 = 10000)$ la variabile casuale che descrive la retribuzione media mensile lorda, occorre, in primo luogo, determinare la seguente probabilità:

$$P(X < 2200).$$

Abbiamo

$$\begin{aligned} P(X < 2200) &= P\left(\frac{X - \mu}{\sigma} < \frac{2200 - \mu}{\sigma}\right) = P\left(Z < \frac{2200 - 2300}{100}\right) = \\ &= P(Z < -1) = \end{aligned}$$



La retribuzione media mensile lorda di un laureato in Economia è distribuita come una Normale di media $\mu = 2300$ euro e varianza $\sigma^2 = 10\,000$ euro.

1. Quale percentuale di laureati guadagna meno di 2200 euro al mese?

Indicata con $X \sim N(\mu = 2300, \sigma^2 = 10000)$ la variabile casuale che descrive la retribuzione media mensile lorda, occorre, in primo luogo, determinare la seguente probabilità:

$$P(X < 2200).$$

Abbiamo

$$\begin{aligned} P(X < 2200) &= P\left(\frac{X - \mu}{\sigma} < \frac{2200 - \mu}{\sigma}\right) = P\left(Z < \frac{2200 - 2300}{100}\right) = \\ &= P(Z < -1) = P(Z > 1) = \end{aligned}$$



La retribuzione media mensile lorda di un laureato in Economia è distribuita come una Normale di media $\mu = 2300$ euro e varianza $\sigma^2 = 10\,000$ euro.

1. Quale percentuale di laureati guadagna meno di 2200 euro al mese?

Indicata con $X \sim N(\mu = 2300, \sigma^2 = 10000)$ la variabile casuale che descrive la retribuzione media mensile lorda, occorre, in primo luogo, determinare la seguente probabilità:

$$P(X < 2200).$$

Abbiamo

$$\begin{aligned} P(X < 2200) &= P\left(\frac{X - \mu}{\sigma} < \frac{2200 - \mu}{\sigma}\right) = P\left(Z < \frac{2200 - 2300}{100}\right) = \\ &= P(Z < -1) = P(Z > 1) = 1 - P(Z \leq 1) = \\ &= \end{aligned}$$



La retribuzione media mensile lorda di un laureato in Economia è distribuita come una Normale di media $\mu = 2300$ euro e varianza $\sigma^2 = 10\,000$ euro.

1. Quale percentuale di laureati guadagna meno di 2200 euro al mese?

Indicata con $X \sim N(\mu = 2300, \sigma^2 = 10000)$ la variabile casuale che descrive la retribuzione media mensile lorda, occorre, in primo luogo, determinare la seguente probabilità:

$$P(X < 2200).$$

Abbiamo

$$\begin{aligned} P(X < 2200) &= P\left(\frac{X - \mu}{\sigma} < \frac{2200 - \mu}{\sigma}\right) = P\left(Z < \frac{2200 - 2300}{100}\right) = \\ &= P(Z < -1) = P(Z > 1) = 1 - P(Z \leq 1) = \\ &= 1 - 0.8413 = 0.1587. \end{aligned}$$

La percentuale di laureati che guadagna meno di 2200 euro al mese risulta 15.87%.



2. Quale è il reddito limite mensile che definisce il 10% inferiore della retribuzione dei laureati in Economia?

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.



2. Quale è il reddito limite mensile che definisce il 10% inferiore della retribuzione dei laureati in Economia?

Il reddito limite mensile che definisce il 10% inferiore della retribuzione dei laureati in Economia corrisponde al percentile di ordine 0.10 della distribuzione di X .

$$\{x : P(X \leq x) = 0.10\} =$$

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.



Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.

2. Quale è il reddito limite mensile che definisce il 10% inferiore della retribuzione dei laureati in Economia?

Il reddito limite mensile che definisce il 10% inferiore della retribuzione dei laureati in Economia corrisponde al percentile di ordine 0.10 della distribuzione di X .

$$\{x : P(X \leq x) = 0.10\} = \left\{x : P\left(\frac{X - \mu}{\sigma} \leq \frac{x - \mu}{\sigma}\right) = 0.10\right\} =$$
$$=$$



2. Quale è il reddito limite mensile che definisce il 10% inferiore della retribuzione dei laureati in Economia?

Il reddito limite mensile che definisce il 10% inferiore della retribuzione dei laureati in Economia corrisponde al percentile di ordine 0.10 della distribuzione di X .

$$\begin{aligned}\{x : P(X \leq x) = 0.10\} &= \left\{x : P\left(\frac{X - \mu}{\sigma} \leq \frac{x - \mu}{\sigma}\right) = 0.10\right\} = \\ &= \left\{x : P\left(Z \leq \frac{x - 2300}{100}\right) = 0.10\right\}.\end{aligned}$$

Dalla lettura inversa della Tavola possiamo ricavare il percentile di ordine 0.90, che risulta



2. Quale è il reddito limite mensile che definisce il 10% inferiore della retribuzione dei laureati in Economia?

Il reddito limite mensile che definisce il 10% inferiore della retribuzione dei laureati in Economia corrisponde al percentile di ordine 0.10 della distribuzione di X .

$$\begin{aligned}\{x : P(X \leq x) = 0.10\} &= \left\{x : P\left(\frac{X - \mu}{\sigma} \leq \frac{x - \mu}{\sigma}\right) = 0.10\right\} = \\ &= \left\{x : P\left(Z \leq \frac{x - 2300}{100}\right) = 0.10\right\}.\end{aligned}$$

Dalla lettura inversa della Tavola possiamo ricavare il percentile di ordine 0.90, che risulta 1.28 (vale, infatti $P(Z \leq 1.28) = 0.90$)).



2. Quale è il reddito limite mensile che definisce il 10% inferiore della retribuzione dei laureati in Economia?

Il reddito limite mensile che definisce il 10% inferiore della retribuzione dei laureati in Economia corrisponde al percentile di ordine 0.10 della distribuzione di X .

$$\begin{aligned}\{x : P(X \leq x) = 0.10\} &= \left\{x : P\left(\frac{X - \mu}{\sigma} \leq \frac{x - \mu}{\sigma}\right) = 0.10\right\} = \\ &= \left\{x : P\left(Z \leq \frac{x - 2300}{100}\right) = 0.10\right\}.\end{aligned}$$

Dalla lettura inversa della Tavola possiamo ricavare il percentile di ordine 0.90, che risulta 1.28 (vale, infatti $P(Z \leq 1.28) = 0.90$)).

Tenendo ora conto che la distribuzione della variabile casuale Normale è simmetrica, abbiamo, con riferimento al percentile di ordine 0.10, $z = \frac{x - 2300}{100} =$



2. Quale è il reddito limite mensile che definisce il 10% inferiore della retribuzione dei laureati in Economia?

Il reddito limite mensile che definisce il 10% inferiore della retribuzione dei laureati in Economia corrisponde al percentile di ordine 0.10 della distribuzione di X .

$$\begin{aligned}\{x : P(X \leq x) = 0.10\} &= \left\{x : P\left(\frac{X - \mu}{\sigma} \leq \frac{x - \mu}{\sigma}\right) = 0.10\right\} = \\ &= \left\{x : P\left(Z \leq \frac{x - 2300}{100}\right) = 0.10\right\}.\end{aligned}$$

Dalla lettura inversa della Tavola possiamo ricavare il percentile di ordine 0.90, che risulta 1.28 (vale, infatti $P(Z \leq 1.28) = 0.90$)).

Tenendo ora conto che la distribuzione della variabile casuale Normale è simmetrica, abbiamo, con riferimento al percentile di ordine 0.10, $z = \frac{x - 2300}{100} = -1.28$ da cui segue $x = 2172$.



3. Si calcoli la probabilità che su 10 laureati selezionati casualmente almeno uno guadagni meno di 2200 euro al mese.

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.



Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.

3. Si calcoli la probabilità che su 10 laureati selezionati casualmente almeno uno guadagni meno di 2200 euro al mese.

Sia X la variabile casuale 'n. di laureati con guadagno inferiore a 2200 euro tra i 10 selezionati'. Dal momento che la popolazione dei laureati ha dimensione molto elevata, abbiamo $X \sim \text{Bin}(n = 10, p = 0.1587)$.

$$P(X \geq 1) =$$



Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.

3. Si calcoli la probabilità che su 10 laureati selezionati casualmente almeno uno guadagni meno di 2200 euro al mese.

Sia X la variabile casuale 'n. di laureati con guadagno inferiore a 2200 euro tra i 10 selezionati'. Dal momento che la popolazione dei laureati ha dimensione molto elevata, abbiamo $X \sim \text{Bin}(n = 10, p = 0.1587)$.

$$P(X \geq 1) = 1 - P(X = 0) =$$

=



Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.

3. Si calcoli la probabilità che su 10 laureati selezionati casualmente almeno uno guadagni meno di 2200 euro al mese.

Sia X la variabile casuale 'n. di laureati con guadagno inferiore a 2200 euro tra i 10 selezionati'. Dal momento che la popolazione dei laureati ha dimensione molto elevata, abbiamo $X \sim \text{Bin}(n = 10, p = 0.1587)$.

$$\begin{aligned} P(X \geq 1) &= 1 - P(X = 0) = \\ &= 1 - \binom{10}{0} 0.1587^0 \cdot 0.8513^{10} = \\ &= \end{aligned}$$



Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.

3. Si calcoli la probabilità che su 10 laureati selezionati casualmente almeno uno guadagni meno di 2200 euro al mese.

Sia X la variabile casuale 'n. di laureati con guadagno inferiore a 2200 euro tra i 10 selezionati'. Dal momento che la popolazione dei laureati ha dimensione molto elevata, abbiamo $X \sim \text{Bin}(n = 10, p = 0.1587)$.

$$\begin{aligned} P(X \geq 1) &= 1 - P(X = 0) = \\ &= 1 - \binom{10}{0} 0.1587^0 \cdot 0.8513^{10} = \\ &= 1 - 0.1777 = 0.8223. \end{aligned}$$

Tema T300

Prova scritta di STATISTICA del 06.09.2018

Prova scritta di
STATISTICA del
06.09.2018

Giuseppe Boari
Gabriele Cantaluppi



UNIVERSITÀ
CATTOLICA
del Sacro Cuore

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.

Giuseppe Boari
Gabriele Cantaluppi
Università Cattolica del Sacro Cuore, Milano



Con riferimento a due teatri italiani (x_1 = Teatro alla Scala, x_2 = Teatro Regio di Torino) si riporta nella seguente tabella la distribuzione della variabile Y 'età dello spettatore':

$X \setminus Y$	15 - 35	35 - 65	65 - 95
x_1	0	10	5
x_2	5	5	10

1. Si rappresentino graficamente le distribuzioni marginali di X e di Y .

[Esercizio 1.](#)

[Esercizio 2.](#)

[Esercizio 3.](#)

[Esercizio 4.](#)



Si considera la distribuzione marginale della variabile X , ottenuta sommando le frequenze di riga

x_i	$n_{i\bullet}$
x_1	15
x_2	20

Si tratta di un carattere qualitativo sconnesso che può essere rappresentato graficamente mediante un grafico a bastoncini o a canne d'organo.

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.

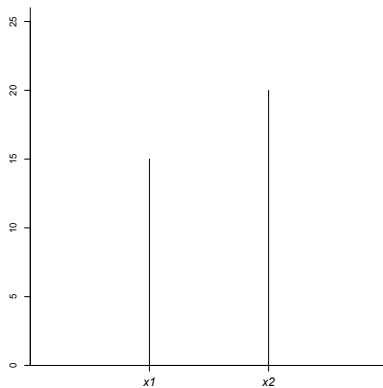


Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.





Si considera la distribuzione marginale della variabile Y , ottenuta sommando le frequenze di colonna

y_j	$n_{\bullet j}$
15 + 35	5
35 + 65	15
65 + 95	15

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.



Si tratta di un carattere quantitativo con modalità raggruppate in classi che può essere rappresentato graficamente mediante un istogramma.

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.



L'istogramma può essere costruito dopo avere calcolato le densità $d_{\bullet j} = \frac{n_{\bullet j}}{a_j}$, (dove a_j sono le ampiezze di classe).

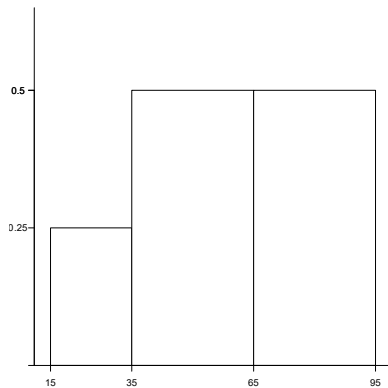
y_j	$n_{\bullet j}$	$a_{\bullet j}$	$d_{\bullet j}$
15 - 35	5	35 - 15 = 20	0.25
35 - 65	15	65 - 35 = 30	0.5
65 - 95	15	95 - 65 = 30	0.5

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.



Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.



2. Si confronti la distribuzione delle età degli spettatori nei due teatri attraverso opportuni indici di posizione e di variabilità, commentando i risultati ottenuti.

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.



Si considera la distribuzione condizionata $Y|x_1$

y_j	n_{1j}
15 → 35	0
35 → 65	10
65 → 95	5

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.



Si tratta di un carattere quantitativo con modalità raggruppate in classi.

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.



La media aritmetica

$$M(Y|x_1) = \frac{1}{n_{1\bullet}} \sum_{j=1}^k y_j n_{1j}$$

risulta 60.

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.



La media aritmetica

$$M(Y|x_1) = \frac{1}{n_{1\bullet}} \sum_{j=1}^k y_j n_{1j}$$

risulta 60.

Si ottiene (considerando come modalità i valori centrali delle classi)

y_j	n_{1j}	$y_j n_{1j}$
25	0	0
50	10	500
80	5	400
	15	900

da cui segue

$$M(Y|x_1) = \frac{1}{n_{1\bullet}} \sum_{j=1}^k y_j n_{1j} = \frac{900}{15} = 60.$$

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.



La varianza
risulta 200.

$$\text{Var}(Y|x_1) = M(Y^2|x_1) - [M(Y|x_1)]^2$$

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.



Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.

La varianza

$$Var(Y|x_1) = M(Y^2|x_1) - [M(Y|x_1)]^2$$

risulta 200. Si ottiene (considerando come modalità i valori centrali delle classi)

y_j	n_{1j}	$y_j n_{1j}$	$y_j^2 n_{1j}$
25	0	0	0
50	10	500	25000
80	5	400	32000
	15	900	57000

da cui segue

$$Var(Y|x_1) = \frac{57000}{15} - 60^2 = 3800 - 3600 = 200.$$



Per effettuare confronti in termini di variabilità tra caratteri quantitativi occorre utilizzare una misura relativa, ad esempio il coefficiente di variazione

$$CV = \frac{\sigma}{\mu} = \frac{\sqrt{\text{Var}(Y|x_1)}}{M(Y|x_1)} = \frac{\sqrt{200}}{60} = \frac{14.1421}{60} = 0.2357.$$

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.



Si considera la distribuzione condizionata $Y|x_2$

y_j	n_{2j}
15 $\rightarrow$ 35	5
35 $\rightarrow$ 65	5
65 $\rightarrow$ 95	10

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.



Si tratta di un carattere quantitativo con modalità raggruppate in classi.

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.



La media aritmetica

$$M(Y|x_2) = \frac{1}{n_{2\bullet}} \sum_{j=1}^k y_j n_{2j}$$

risulta 58.75.

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.



La media aritmetica

$$M(Y|x_2) = \frac{1}{n_{2\bullet}} \sum_{j=1}^k y_j n_{2j}$$

risulta 58.75.

Si ottiene (considerando come modalità i valori centrali delle classi)

y_j	n_{2j}	$y_j n_{2j}$
25	5	125
50	5	250
80	10	800
	20	1175

da cui segue

$$M(Y|x_2) = \frac{1}{n_{2\bullet}} \sum_{j=1}^k y_j n_{2j} = \frac{1175}{20} = 58.75.$$

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.



La varianza

$$\text{Var}(Y|x_2) = M(Y^2|x_2) - [M(Y|x_2)]^2$$

risulta 529.6875.

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.



Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.

La varianza

$$Var(Y|x_2) = M(Y^2|x_2) - [M(Y|x_2)]^2$$

risulta 529.6875. Si ottiene (considerando come modalità i valori centrali delle classi)

y_j	n_{2j}	$y_j n_{2j}$	$y_j^2 n_{2j}$
25	5	125	3125
50	5	250	12500
80	10	800	64000
	20	1175	79625

da cui segue

$$Var(Y|x_2) = \frac{79625}{20} - 58.75^2 = 3981.25 - 3451.5625 = 529.6875.$$



Per effettuare confronti in termini di variabilità tra caratteri quantitativi occorre utilizzare una misura relativa, ad esempio il coefficiente di variazione

$$CV = \frac{\sigma}{\mu} = \frac{\sqrt{\text{Var}(Y|x_2)}}{M(Y|x_2)} = \frac{\sqrt{529.6875}}{58.75} = \frac{23.0149}{58.75} = 0.3917.$$

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.



Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.

$Y|x_1$ Indice di posizione: $M(Y|x_1) = 60$ Indice di variabilità: $CV(Y|x_1) = 0.2357$

$Y|x_2$ Indice di posizione: $M(Y|x_2) = 58.75$ Indice di variabilità: $CV(Y|x_2) = 0.3917$



Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.

$Y|x_1$ Indice di posizione: $M(Y|x_1) = 60$ Indice di variabilità: $CV(Y|x_1) = 0.2357$

$Y|x_2$ Indice di posizione: $M(Y|x_2) = 58.75$ Indice di variabilità: $CV(Y|x_2) = 0.3917$

L'età degli spettatori del Teatro alla Scala risulta in media leggermente più elevata e meno variabile di quella degli spettatori del Teatro Regio di Torino.



Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.

3. Si determini il valore dell'indice χ^2 normalizzato e si commenti il risultato ottenuto.



Dalla distribuzione delle frequenze congiunte n_{ij} :

	15 ÷ 35	35 ÷ 65	65 ÷ 95	$n_{i\bullet}$
x_1	0	10	5	15
x_2	5	5	10	20
$n_{\bullet j}$	5	15	15	35

è possibile ottenere le frequenze teoriche nella situazione di indipendenza stocastica

$$\hat{n}_{ij} = \frac{n_{i\bullet} n_{\bullet j}}{n}:$$

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.



Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.

Dalla distribuzione delle frequenze congiunte n_{ij} :

	15 ÷ 35	35 ÷ 65	65 ÷ 95	$n_{i\bullet}$
x_1	0	10	5	15
x_2	5	5	10	20
$n_{\bullet j}$	5	15	15	35

è possibile ottenere le frequenze teoriche nella situazione di indipendenza stocastica

$$\hat{n}_{ij} = \frac{n_{i\bullet} \cdot n_{\bullet j}}{n}:$$

	15 ÷ 35	35 ÷ 65	65 ÷ 95	$n_{i\bullet}$
x_1	2.1429	6.4286	6.4286	15
x_2	2.8571	8.5714	8.5714	20
$n_{\bullet j}$	5.0000	15.0000	15.0000	35



le contingenze assolute $n_{ij} - \hat{n}_{ij}$:

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.



le contingenze assolute $n_{ij} - \hat{n}_{ij}$:

	15 ÷ 35	35 ÷ 65	65 ÷ 95
x_1	-2.1429	3.5714	-1.4286
x_2	2.1429	-3.5714	1.4286

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.



i rapporti di contingenza $\frac{n_{ij}}{\hat{n}_{ij}}$:

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.



i rapporti di contingenza $\frac{n_{ij}}{\hat{n}_{ij}}$:

	15 - 35	35 - 65	65 - 95
x_1	0.00	1.5556	0.7778
x_2	1.75	0.5833	1.1667

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.



e le contingenze relative, espresse in termini percentuali, $100 \left(\frac{n_{ij} - \hat{n}_{ij}}{\hat{n}_{ij}} \right) \% = 100 \left(\frac{n_{ij}}{\hat{n}_{ij}} - 1 \right) \%$:

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.



e le contingenze relative, espresse in termini percentuali, $100 \left(\frac{n_{ij} - \hat{n}_{ij}}{\hat{n}_{ij}} \right) \% = 100 \left(\frac{n_{ij}}{\hat{n}_{ij}} - 1 \right) \%$:

	15 ÷ 35	35 ÷ 65	65 ÷ 95
x_1	-100	55.56	-22.22
x_2	75	-41.67	16.67

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.



Risultati intermedi $\frac{(n_{ij} - \hat{n}_{ij})^2}{\hat{n}_{ij}}$ per il calcolo del χ^2

	15 ÷ 35	35 ÷ 65	65 ÷ 95
x_1	2.1429	1.9841	0.3175
x_2	1.6071	1.4881	0.2381

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.



Risultati intermedi $\frac{n_{ij}^2}{n_{i\bullet} \cdot n_{\bullet j}}$ per il calcolo del χ^2 (formula operativa)

	15 + 35	35 + 65	65 + 95
x_1	0.00	0.4444	0.1111
x_2	0.25	0.0833	0.3333

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.



Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.

$$\chi^2 = \sum_{i=1}^h \sum_{j=1}^k \frac{(n_{ij} - \hat{n}_{ij})^2}{\hat{n}_{ij}} = 7.7778,$$

$$\chi_N^2 = \frac{\chi^2}{n \cdot \min(h-1, k-1)} = 0.2222$$

$$\chi_N = \sqrt{\chi_N^2} = 0.4714.$$

Sussiste, quindi, un livello di connessione medio.



Il prospetto seguente riporta la serie storica delle quotazioni dei titoli azionari x_t di una società calcistica italiana in cinque giornate del mese di luglio 2018:

t	x_t
12	65
13	70
14	80
15	85
16	85

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.



1. Si costruiscano la serie dei numeri indici a base fissa con base $t = 12$ e la serie dei numeri indici a base mobile.

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.



Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.

1. Si costruiscano la serie dei numeri indici a base fissa con base $t = 12$ e la serie dei numeri indici a base mobile.

Si ricorda che ${}_{12}I_t = \frac{x_t}{x_{12}}$.



1. Si costruiscano la serie dei numeri indici a base fissa con base $t = 12$ e la serie dei numeri indici a base mobile.

Si ricorda che ${}_{12}I_t = \frac{x_t}{x_{12}}$. Si ottiene quindi

Anno	x_t	NIBF (12)
12	65	1.0000
13	70	1.0769
14	80	1.2308
15	85	1.3077
16	85	1.3077

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.



Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.

1. Si costruiscano la serie dei numeri indici a base fissa con base $t = 12$ e la serie dei numeri indici a base mobile.

Si ricorda che ${}_{12}I_t = \frac{x_t}{x_{12}}$. Si ottiene quindi

Anno	x_t	NIBF (12)
12	65	1.0000
13	70	1.0769
14	80	1.2308
15	85	1.3077
16	85	1.3077

Si ricorda che ${}_{t-1}I_t = \frac{x_t}{x_{t-1}}$.



Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.

1. Si costruiscano la serie dei numeri indici a base fissa con base $t = 12$ e la serie dei numeri indici a base mobile.

Si ricorda che ${}_{12}I_t = \frac{x_t}{x_{12}}$. Si ottiene quindi

Anno	x_t	NIBF (12)
12	65	1.0000
13	70	1.0769
14	80	1.2308
15	85	1.3077
16	85	1.3077

Si ricorda che ${}_{t-1}I_t = \frac{x_t}{x_{t-1}}$. Si ottiene quindi

Anno	x_t	NIBM
12	65	
13	70	1.0769
14	80	1.1429
15	85	1.0625
16	85	1.0000



2. Si calcoli il tasso medio di variazione delle quotazioni tra $t = 12$ e $t = 16$ e si interpreti il valore ottenuto.

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.



2. Si calcoli il tasso medio di variazione delle quotazioni tra $t = 12$ e $t = 16$ e si interpreti il valore ottenuto.

Si ricorda che

$${}_{12}l_{16} = {}_{12}l_{13} \cdot {}_{13}l_{14} \cdot {}_{14}l_{15} \cdot {}_{15}l_{16}$$

Si deve trovare il numero indice a base mobile medio.

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.



Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.

2. Si calcoli il tasso medio di variazione delle quotazioni tra $t = 12$ e $t = 16$ e si interpreti il valore ottenuto.

Si ricorda che

$${}_{12}l_{16} = {}_{12}l_{13} \cdot {}_{13}l_{14} \cdot {}_{14}l_{15} \cdot {}_{15}l_{16}$$

Si deve trovare il numero indice a base mobile medio. Applicando il criterio di scelta della media secondo Chisini si ottiene

$${}_{12}l_{16} = {}_{12}l_{13} \cdot {}_{13}l_{14} \cdot {}_{14}l_{15} \cdot {}_{15}l_{16} = \alpha \cdot \alpha \cdot \alpha \cdot \alpha = \alpha^4.$$



Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.

2. Si calcoli il tasso medio di variazione delle quotazioni tra $t = 12$ e $t = 16$ e si interpreti il valore ottenuto.

Si ricorda che

$${}_{12}l_{16} = {}_{12}l_{13} \cdot {}_{13}l_{14} \cdot {}_{14}l_{15} \cdot {}_{15}l_{16}$$

Si deve trovare il numero indice a base mobile medio. Applicando il criterio di scelta della media secondo Chisini si ottiene

$${}_{12}l_{16} = {}_{12}l_{13} \cdot {}_{13}l_{14} \cdot {}_{14}l_{15} \cdot {}_{15}l_{16} = \alpha \cdot \alpha \cdot \alpha \cdot \alpha = \alpha^4.$$

Quindi

$$\alpha = \sqrt[4]{{}_{12}l_{16}} = \sqrt[4]{{}_{12}l_{13} \cdot {}_{13}l_{14} \cdot {}_{14}l_{15} \cdot {}_{15}l_{16}} = 1.0694.$$

Da cui segue una variazione relativa media pari a 0.0694, corrispondente a un incremento medio annuo tra il 12 e il 16 del 6.94%.



3. Sapendo che il numero indice a base mobile per il tempo $t = 17$ è risultato pari a 0.95, si determini il valore del titolo x_{17} .

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.



Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.

3. Sapendo che il numero indice a base mobile per il tempo $t = 17$ è risultato pari a 0.95, si determini il valore del titolo x_{17} .

Abbiamo

$$x_{17} = x_{16} \cdot {}_{16}I_{17} = 85 \cdot 0.95 = 80.75.$$



Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.

Il seguente prospetto riporta i valori rilevati in corrispondenza a 5 unità statistiche con riferimento a due caratteri quantitativi X e Y

x_i	1	2	3	4	5
y_i	1	1	3	5	5

1. Si rappresentino graficamente le coppie dei dati e la funzione di regressione di Y in funzione di X .



La funzione di regressione associa a ogni valore x appartenente all'insieme di definizione della variabile X la media di Y condizionata a x , $M(Y|x)$

x	$M(Y x)$
1	1
2	1
3	3
4	5
5	5

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.



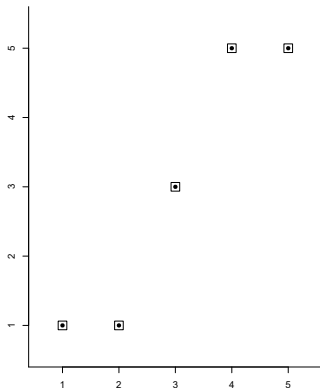
Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.

Le coppie di dati possono essere rappresentate mediante un grafico a dispersione.





2. Si calcolino, secondo il metodo dei minimi quadrati, i parametri dei seguenti modelli di regressione:

$$I) Y^* = a + b \cdot \ln(X) \qquad II) Y^* = c$$

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.



Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.

Si riportano i calcoli necessari¹ per ottenere le stime dei parametri e gli indici di adattamento del modello

$$Y = a + b \ln(X) + E$$

$\ln(x)_i$	y_i	$\ln(x)_i y_i$	$\ln(x)_i^2$	y_i^2
0	1	0	0	1
0.6931	1	0.6931	0.4805	1
1.0986	3	3.2958	1.2069	9
1.3863	5	6.9315	1.9218	25
1.6094	5	8.0472	2.5903	25
4.7875	15	18.9676	6.1995	61
0.9575	3	3.7935	1.2399	12.2
$M(\ln(X))$	$M(Y)$	$M(\ln(X)Y)$	$M(\ln(X)^2)$	$M(Y^2)$

$$\hat{b} =$$

¹Si osserva che per ragioni di stampa alcuni calcoli potrebbero essere duplicati con riferimento ai diversi modelli oggetto di studio



Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.

Si riportano i calcoli necessari¹ per ottenere le stime dei parametri e gli indici di adattamento del modello

$$Y = a + b \ln(X) + E$$

$\ln(x)_i$	y_i	$\ln(x)_i y_i$	$\ln(x)_i^2$	y_i^2
0	1	0	0	1
0.6931	1	0.6931	0.4805	1
1.0986	3	3.2958	1.2069	9
1.3863	5	6.9315	1.9218	25
1.6094	5	8.0472	2.5903	25
4.7875	15	18.9676	6.1995	61
0.9575	3	3.7935	1.2399	12.2
$M(\ln(X))$	$M(Y)$	$M(\ln(X)Y)$	$M(\ln(X)^2)$	$M(Y^2)$

$$\begin{aligned}\hat{b} &= \frac{\text{Cov}(\ln(X), Y)}{\text{Var}(\ln(X))} = \frac{M(\ln(X)Y) - M(\ln(X))M(Y)}{M(\ln(X)^2) - [M(\ln(X))]^2} = \\ &= \frac{3.7935 - 0.9575 \cdot 3}{1.2399 - 0.9575^2} = \frac{0.921}{0.3231} = 2.8505 \\ \hat{a} &= \end{aligned}$$

¹Si osserva che per ragioni di stampa alcuni calcoli potrebbero essere duplicati con riferimento ai diversi modelli oggetto di studio



Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.

Si riportano i calcoli necessari¹ per ottenere le stime dei parametri e gli indici di adattamento del modello

$$Y = a + b \ln(X) + E$$

$\ln(x)_i$	y_i	$\ln(x)_i y_i$	$\ln(x)_i^2$	y_i^2
0	1	0	0	1
0.6931	1	0.6931	0.4805	1
1.0986	3	3.2958	1.2069	9
1.3863	5	6.9315	1.9218	25
1.6094	5	8.0472	2.5903	25
4.7875	15	18.9676	6.1995	61
0.9575	3	3.7935	1.2399	12.2
$M(\ln(X))$	$M(Y)$	$M(\ln(X)Y)$	$M(\ln(X)^2)$	$M(Y^2)$

$$\begin{aligned}\hat{b} &= \frac{\text{Cov}(\ln(X), Y)}{\text{Var}(\ln(X))} = \frac{M(\ln(X)Y) - M(\ln(X))M(Y)}{M(\ln(X)^2) - [M(\ln(X))]^2} = \\ &= \frac{3.7935 - 0.9575 \cdot 3}{1.2399 - 0.9575^2} = \frac{0.921}{0.3231} = 2.8505\end{aligned}$$

$$\hat{a} = M(Y) - \hat{b} M(\ln(X)) = 0.2706$$

$$\rho^2 =$$

¹Si osserva che per ragioni di stampa alcuni calcoli potrebbero essere duplicati con riferimento ai diversi modelli oggetto di studio



Si riportano i calcoli necessari¹ per ottenere le stime dei parametri e gli indici di adattamento del modello

$$Y = a + b \ln(X) + E$$

$\ln(x)_i$	y_i	$\ln(x)_i y_i$	$\ln(x)_i^2$	y_i^2
0	1	0	0	1
0.6931	1	0.6931	0.4805	1
1.0986	3	3.2958	1.2069	9
1.3863	5	6.9315	1.9218	25
1.6094	5	8.0472	2.5903	25
4.7875	15	18.9676	6.1995	61
0.9575	3	3.7935	1.2399	12.2
$M(\ln(X))$	$M(Y)$	$M(\ln(X)Y)$	$M(\ln(X)^2)$	$M(Y^2)$

$$\begin{aligned}\hat{b} &= \frac{\text{Cov}(\ln(X), Y)}{\text{Var}(\ln(X))} = \frac{M(\ln(X)Y) - M(\ln(X))M(Y)}{M(\ln(X)^2) - [M(\ln(X))]^2} = \\ &= \frac{3.7935 - 0.9575 \cdot 3}{1.2399 - 0.9575^2} = \frac{0.921}{0.3231} = 2.8505 \\ \hat{a} &= M(Y) - \hat{b} M(\ln(X)) = 0.2706 \\ \rho^2 &= \frac{[\text{Cov}(\ln(X), Y)]^2}{\text{Var}(\ln(X)) \text{Var}(Y)} = \frac{0.921^2}{0.3231 \cdot 3.2} = 0.8205\end{aligned}$$

¹Si osserva che per ragioni di stampa alcuni calcoli potrebbero essere duplicati con riferimento ai diversi modelli oggetto di studio



Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.

Si riportano i calcoli necessari² per ottenere le stime dei parametri e gli indici di adattamento del modello

$$Y = c + E$$

x_i	y_i	y_i^2
1	1	1
2	1	1
3	3	9
4	5	25
5	5	25
15		61
3		12.2
$M(Y)$		$M(Y^2)$

$$\hat{c} =$$

²Si osserva che per ragioni di stampa alcuni calcoli potrebbero essere duplicati con riferimento ai diversi modelli oggetto di studio



Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.

Si riportano i calcoli necessari² per ottenere le stime dei parametri e gli indici di adattamento del modello

$$Y = c + E$$

x_i	y_i	y_i^2
1	1	1
2	1	1
3	3	9
4	5	25
5	5	25
15		61
3		12.2
$M(Y)$		$M(Y^2)$

$$\hat{c} = M(Y) = 3$$

$$R^2 =$$

²Si osserva che per ragioni di stampa alcuni calcoli potrebbero essere duplicati con riferimento ai diversi modelli oggetto di studio



Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.

Si riportano i calcoli necessari² per ottenere le stime dei parametri e gli indici di adattamento del modello

$$Y = c + E$$

x_i	y_i	y_i^2
1	1	1
2	1	1
3	3	9
4	5	25
5	5	25
15		61
3		12.2
$M(Y)$		$M(Y^2)$

$$\hat{c} = M(Y) = 3$$

$$R^2 = 0$$

$$M(\hat{E}^2) = \text{Var}(\hat{E}) = \text{varianza residua} =$$

²Si osserva che per ragioni di stampa alcuni calcoli potrebbero essere duplicati con riferimento ai diversi modelli oggetto di studio



Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.

Si riportano i calcoli necessari² per ottenere le stime dei parametri e gli indici di adattamento del modello

$$Y = c + E$$

x_i	y_i	y_i^2
1	1	1
2	1	1
3	3	9
4	5	25
5	5	25
15		61
3		12.2
$M(Y)$		$M(Y^2)$

$$\hat{c} = M(Y) = 3$$

$$R^2 = 0$$

$$M(\hat{E}^2) = \text{Var}(\hat{E}) = \text{varianza residua} = \text{Var}(Y) = 3.2$$

²Si osserva che per ragioni di stampa alcuni calcoli potrebbero essere duplicati con riferimento ai diversi modelli oggetto di studio



Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.

$\hat{a} = 0.2705$, $\hat{b} = 2.8506$, indice adattamento $\rho^2 = 0.8205$ $\hat{c} = 3$, indice adattamento $R^2 = 0$



Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.

3. Si indichi quale modello presenta il migliore adattamento ai dati.

Il modello migliore è $Y^* = 0.2705 + 2.8506 \cdot \ln X$ in quanto caratterizzato da indice di adattamento più elevato.



4. Si stabilisca, senza effettuare calcoli e giustificando teoricamente la risposta, il valore assunto dagli indici $\eta_{Y|X}^2$ e $\eta_{X|Y}^2$.

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.



4. Si stabilisca, senza effettuare calcoli e giustificando teoricamente la risposta, il valore assunto dagli indici $\eta^2_{Y|X}$ e $\eta^2_{X|Y}$.

$\eta^2_{Y|X} = 1$ in quanto sussiste dipendenza funzionale di Y da X .

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.



Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.

4. Si stabilisca, senza effettuare calcoli e giustificando teoricamente la risposta, il valore assunto dagli indici $\eta_{Y|X}^2$ e $\eta_{X|Y}^2$.

$\eta_{Y|X}^2 = 1$ in quanto sussiste dipendenza funzionale di Y da X .

$0 < \eta_{X|Y}^2 < 1$ in quanto non sussistono né indipendenza in media né dipendenza funzionale di X da Y .



Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.

4. Si stabilisca, senza effettuare calcoli e giustificando teoricamente la risposta, il valore assunto dagli indici $\eta_{Y|X}^2$ e $\eta_{X|Y}^2$.

$\eta_{Y|X}^2 = 1$ in quanto sussiste dipendenza funzionale di Y da X .

$0 < \eta_{X|Y}^2 < 1$ in quanto non sussistono né indipendenza in media né dipendenza funzionale di X da Y .

Si osserva come non sia possibile stabilire la relazione esistente tra $\eta_{X|Y}^2$ e $\rho_{\ln(X), Y}^2$ non essendo $\ln(X)$ trasformazione lineare di X .

Esercizio 4

Prova scritta di
STATISTICA del
06.09.2018

Giuseppe Boari
Gabriele Cantaluppi



UNIVERSITÀ
CATTOLICA
del Sacro Cuore

Una pasticceria produce Sachertorte il cui peso in grammi X si distribuisce secondo una variabile aleatoria Normale con media $\mu = 500$ e scarto quadratico medio $\sigma = 16$.

1. Si calcoli la probabilità che una torta pesi meno di 460 grammi.

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.

Esercizio 4



[Esercizio 1.](#)

[Esercizio 2.](#)

[Esercizio 3.](#)

[Esercizio 4.](#)

Una pasticceria produce Sachertorte il cui peso in grammi X si distribuisce secondo una variabile aleatoria Normale con media $\mu = 500$ e scarto quadratico medio $\sigma = 16$.

1. Si calcoli la probabilità che una torta pesi meno di 460 grammi.

Indicata con $X \sim N(\mu = 500, \sigma^2 = 256)$ la variabile casuale che descrive il peso in grammi delle Sacherorte, occorre determinare la seguente probabilità:

$$P(X < 460).$$

Abbiamo

$$P(X < 460) =$$

Esercizio 4



[Esercizio 1.](#)

[Esercizio 2.](#)

[Esercizio 3.](#)

[Esercizio 4.](#)

Una pasticceria produce Sachertorte il cui peso in grammi X si distribuisce secondo una variabile aleatoria Normale con media $\mu = 500$ e scarto quadratico medio $\sigma = 16$.

1. Si calcoli la probabilità che una torta pesi meno di 460 grammi.

Indicata con $X \sim N(\mu = 500, \sigma^2 = 256)$ la variabile casuale che descrive il peso in grammi delle Sacherorte, occorre determinare la seguente probabilità:

$$P(X < 460).$$

Abbiamo

$$P(X < 460) = P\left(\frac{X - \mu}{\sigma} < \frac{460 - \mu}{\sigma}\right) =$$



Una pasticceria produce Sachertorte il cui peso in grammi X si distribuisce secondo una variabile aleatoria Normale con media $\mu = 500$ e scarto quadratico medio $\sigma = 16$.

1. Si calcoli la probabilità che una torta pesi meno di 460 grammi.

Indicata con $X \sim N(\mu = 500, \sigma^2 = 256)$ la variabile casuale che descrive il peso in grammi delle Sacherorte, occorre determinare la seguente probabilità:

$$P(X < 460).$$

Abbiamo

$$\begin{aligned} P(X < 460) &= P\left(\frac{X - \mu}{\sigma} < \frac{460 - \mu}{\sigma}\right) = P\left(Z < \frac{460 - 500}{16}\right) = \\ &= \end{aligned}$$



Una pasticceria produce Sachertorte il cui peso in grammi X si distribuisce secondo una variabile aleatoria Normale con media $\mu = 500$ e scarto quadratico medio $\sigma = 16$.

1. Si calcoli la probabilità che una torta pesi meno di 460 grammi.

Indicata con $X \sim N(\mu = 500, \sigma^2 = 256)$ la variabile casuale che descrive il peso in grammi delle Sacherorte, occorre determinare la seguente probabilità:

$$P(X < 460).$$

Abbiamo

$$\begin{aligned} P(X < 460) &= P\left(\frac{X - \mu}{\sigma} < \frac{460 - \mu}{\sigma}\right) = P\left(Z < \frac{460 - 500}{16}\right) = \\ &= P(Z < -2.5) = \end{aligned}$$



Una pasticceria produce Sachertorte il cui peso in grammi X si distribuisce secondo una variabile aleatoria Normale con media $\mu = 500$ e scarto quadratico medio $\sigma = 16$.

1. Si calcoli la probabilità che una torta pesi meno di 460 grammi.

Indicata con $X \sim N(\mu = 500, \sigma^2 = 256)$ la variabile casuale che descrive il peso in grammi delle Sacherorte, occorre determinare la seguente probabilità:

$$P(X < 460).$$

Abbiamo

$$\begin{aligned} P(X < 460) &= P\left(\frac{X - \mu}{\sigma} < \frac{460 - \mu}{\sigma}\right) = P\left(Z < \frac{460 - 500}{16}\right) = \\ &= P(Z < -2.5) = P(Z > 2.5) = \end{aligned}$$



Una pasticceria produce Sachertorte il cui peso in grammi X si distribuisce secondo una variabile aleatoria Normale con media $\mu = 500$ e scarto quadratico medio $\sigma = 16$.

1. Si calcoli la probabilità che una torta pesi meno di 460 grammi.

Indicata con $X \sim N(\mu = 500, \sigma^2 = 256)$ la variabile casuale che descrive il peso in grammi delle Sacherorte, occorre determinare la seguente probabilità:

$$P(X < 460).$$

Abbiamo

$$\begin{aligned} P(X < 460) &= P\left(\frac{X - \mu}{\sigma} < \frac{460 - \mu}{\sigma}\right) = P\left(Z < \frac{460 - 500}{16}\right) = \\ &= P(Z < -2.5) = P(Z > 2.5) = 1 - P(Z \leq 2.5) = \\ &= \end{aligned}$$



Una pasticceria produce Sachertorte il cui peso in grammi X si distribuisce secondo una variabile aleatoria Normale con media $\mu = 500$ e scarto quadratico medio $\sigma = 16$.

1. Si calcoli la probabilità che una torta pesi meno di 460 grammi.

Indicata con $X \sim N(\mu = 500, \sigma^2 = 256)$ la variabile casuale che descrive il peso in grammi delle Sacherorte, occorre determinare la seguente probabilità:

$$P(X < 460).$$

Abbiamo

$$\begin{aligned} P(X < 460) &= P\left(\frac{X - \mu}{\sigma} < \frac{460 - \mu}{\sigma}\right) = P\left(Z < \frac{460 - 500}{16}\right) = \\ &= P(Z < -2.5) = P(Z > 2.5) = 1 - P(Z \leq 2.5) = \\ &= 1 - 0.9938 = 0.0062. \end{aligned}$$



2. Sapendo che vengono esclusi dalla vendita i prodotti con peso inferiore a 460 grammi e che in una giornata sono state prodotte 10 torte, si calcoli la probabilità che tutte vengano messe in vendita, descrivendo la variabile aleatoria utilizzata.

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.



Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.

2. Sapendo che vengono esclusi dalla vendita i prodotti con peso inferiore a 460 grammi e che in una giornata sono state prodotte 10 torte, si calcoli la probabilità che tutte vengano messe in vendita, descrivendo la variabile aleatoria utilizzata.

La probabilità che una singola Sachertorte venga messa in vendita risulta pari a

$$p = 1 - P(X < 460) = 1 - 0.0062 = 0.9938$$

Ipotizzando, l'indipendenza nella produzione, il numero di Sachertorte messe in vendita risulta distribuito secondo una variabile casuale



Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.

2. Sapendo che vengono esclusi dalla vendita i prodotti con peso inferiore a 460 grammi e che in una giornata sono state prodotte 10 torte, si calcoli la probabilità che tutte vengano messe in vendita, descrivendo la variabile aleatoria utilizzata.

La probabilità che una singola Sachertorte venga messa in vendita risulta pari a

$$p = 1 - P(X < 460) = 1 - 0.0062 = 0.9938$$

Ipotizzando, l'indipendenza nella produzione, il numero di Sachertorte messe in vendita risulta distribuito secondo una variabile casuale Binomiale con parametri $n = 10$ e $p = 0.9938$. Abbiamo, quindi

$$P(X = 10) =$$



Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.

2. Sapendo che vengono esclusi dalla vendita i prodotti con peso inferiore a 460 grammi e che in una giornata sono state prodotte 10 torte, si calcoli la probabilità che tutte vengano messe in vendita, descrivendo la variabile aleatoria utilizzata.

La probabilità che una singola Sachertorte venga messa in vendita risulta pari a

$$p = 1 - P(X < 460) = 1 - 0.0062 = 0.9938$$

Ipotizzando, l'indipendenza nella produzione, il numero di Sachertorte messe in vendita risulta distribuito secondo una variabile casuale Binomiale con parametri $n = 10$ e $p = 0.9938$. Abbiamo, quindi

$$P(X = 10) = 0.9938^{10} =$$



Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.

2. Sapendo che vengono esclusi dalla vendita i prodotti con peso inferiore a 460 grammi e che in una giornata sono state prodotte 10 torte, si calcoli la probabilità che tutte vengano messe in vendita, descrivendo la variabile aleatoria utilizzata.

La probabilità che una singola Sachertorte venga messa in vendita risulta pari a

$$p = 1 - P(X < 460) = 1 - 0.0062 = 0.9938$$

Ipotizzando, l'indipendenza nella produzione, il numero di Sachertorte messe in vendita risulta distribuito secondo una variabile casuale Binomiale con parametri $n = 10$ e $p = 0.9938$. Abbiamo, quindi

$$P(X = 10) = 0.9938^{10} = 0.9397.$$



3. Il guadagno sul prezzo di vendita di ciascuna torta è di 10 euro. Si calcoli il guadagno medio atteso per la giornata (si ipotizzi che tutte le Sachertorte messe in vendita vengano acquistate).

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.



Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.

3. Il guadagno sul prezzo di vendita di ciascuna torta è di 10 euro. Si calcoli il guadagno medio atteso per la giornata (si ipotizzi che tutte le Sachertorte messe in vendita vengano acquistate).

Dal momento che il numero Y di Sachertorte messe in vendita risulta distribuito secondo una variabile casuale Binomiale $Y \sim \text{Bin}(n = 10, p = 0.9938)$, abbiamo

$$E(Y) =$$



3. Il guadagno sul prezzo di vendita di ciascuna torta è di 10 euro. Si calcoli il guadagno medio atteso per la giornata (si ipotizzi che tutte le Sachertorte messe in vendita vengano acquistate).

Dal momento che il numero Y di Sachertorte messe in vendita risulta distribuito secondo una variabile casuale Binomiale $Y \sim \text{Bin}(n = 10, p = 0.9938)$, abbiamo

$$E(Y) = n \cdot p = 9.938,$$

numero atteso di Sachertorte vendute a cui corrisponde un guadagno G atteso pari a

$$E(G) = 10 \cdot 9.938 = 99.38,$$

tenendo conto della relazione $G = 10 \cdot Y$ tra guadagno e numero di torte vendute e che la media (valore atteso) è un operatore lineare.

Tema T301

Prova scritta di STATISTICA del 16.01.2019

Prova scritta di
STATISTICA del
16.01.2019

Giuseppe Boari
Gabriele Cantaluppi



UNIVERSITÀ
CATTOLICA
del Sacro Cuore

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.

Giuseppe Boari
Gabriele Cantaluppi
Università Cattolica del Sacro Cuore, Milano



La seguente tabella riporta le distribuzioni relative condizionate - ricavate a seguito di una indagine campionaria condotta su $n = 20$ lavoratori autonomi - della variabile $Y = \text{Reddito mensile}$, espressa in migliaia di Euro, rispetto alla variabile $X = \text{Genere}$:

$Y X$	0 - 2	2 - 6	6 - 8	
m	0.1	0.4	0.5	1
f	0.5	0.4	0.1	1

1. Dopo avere rappresentato graficamente la variabile $Y|X = m$, si dica se è possibile concludere che la distribuzione è simmetrica.

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.



Si considera la distribuzione condizionata $Y|x_1$

y_j	$f.cond.r._{1j}$
0 $\rightarrow$ 2	0.1
2 $\rightarrow$ 6	0.4
6 $\rightarrow$ 8	0.5

Esercizio 1.

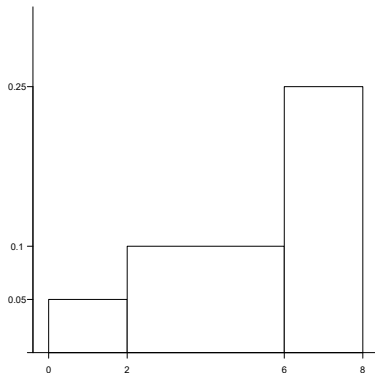
Esercizio 2.

Esercizio 3.

Esercizio 4.

Si tratta di un carattere quantitativo con modalità raggruppate in classi che può essere rappresentato graficamente mediante un istogramma. L'istogramma può essere costruito dopo avere calcolato le densità $d_{1j} = \frac{f.\text{cond}.r.1j}{a_j}$, (dove a_j sono le ampiezze di classe).

y_j	$f.\text{cond}.r.1j$	a_{1j}	d_{1j}
0 ↗ 2	0.1	$2 - 0 = 2$	0.05
2 ↗ 6	0.4	$6 - 2 = 4$	0.1
6 ↗ 8	0.5	$8 - 6 = 2$	0.25



Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.



La distribuzione non è simmetrica perché le densità di frequenza della prima e terza classe non sono uguali tra loro.

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.



Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.

2. Costruire la tabella delle frequenze congiunte sapendo che $n/2$ dei lavoratori intervistati sono *maschi* e $n/2$ sono *femmine*.



Moltiplicando le frequenze condizionate relative $f.cond.r_{1j} = \frac{n_{1j}}{n_{1\bullet}}$ per il numero totale di osservazioni $n_{1\bullet}$ della distribuzione condizionata, si ottengono le frequenze assolute congiunte n_{1j} della distribuzione condizionata $Y|X_1$.

y_j	$f.cond.r_{1j}$	n_{1j}
0 ÷ 2	0.1	1
2 ÷ 6	0.4	4
6 ÷ 8	0.5	5

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.



Moltiplicando le frequenze condizionate relative $f.cond.r_{2j} = \frac{n_{2j}}{n_{2\bullet}}$ per il numero totale di osservazioni $n_{2\bullet}$ della distribuzione condizionata, si ottengono le frequenze assolute congiunte n_{2j} della distribuzione condizionata $Y|X_2$.

y_j	$f.cond.r_{2j}$	n_{2j}
0 $\dashv$ 2	0.5	5
2 $\dashv$ 6	0.4	4
6 $\dashv$ 8	0.1	1

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.



La distribuzione delle frequenze congiunte risulta

$X \setminus Y$	0 - 2	2 - 6	6 - 8	
m	1	4	5	10
f	5	4	1	10
	6	8	6	20

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.



3. Confrontare, con opportuni indici di posizione e di variabilità, la distribuzione della variabile reddito nei *maschi* (m) e nelle *femmine* (f). Commentare il risultato ottenuto.

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.



Si considera la distribuzione condizionata $Y|x_1$

y_j	n_{1j}
0 + 2	1
2 + 6	4
6 + 8	5

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.



Si tratta di un carattere quantitativo con modalità raggruppate in classi.

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.



La media aritmetica

$$M(Y|x_1) = \frac{1}{n_{1\bullet}} \sum_{j=1}^k y_j n_{1j}$$

risulta 5.2.

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.



La media aritmetica

$$M(Y|x_1) = \frac{1}{n_{1\bullet}} \sum_{j=1}^k y_j n_{1j}$$

risulta 5.2.

Si ottiene (considerando come modalità i valori centrali delle classi)

y_j	n_{1j}	$y_j n_{1j}$
1	1	1
4	4	16
7	5	35
	10	52

da cui segue

$$M(Y|x_1) = \frac{1}{n_{1\bullet}} \sum_{j=1}^k y_j n_{1j} = \frac{52}{10} = 5.2.$$

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.



La varianza

$$\text{Var}(Y|x_1) = M(Y^2|x_1) - [M(Y|x_1)]^2$$

risulta 3.96.

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.



Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.

La varianza

$$Var(Y|x_1) = M(Y^2|x_1) - [M(Y|x_1)]^2$$

risulta 3.96. Si ottiene (considerando come modalità i valori centrali delle classi)

y_j	n_{1j}	$y_j n_{1j}$	$y_j^2 n_{1j}$
1	1	1	1
4	4	16	64
7	5	35	245
	10	52	310

da cui segue

$$Var(Y|x_1) = \frac{310}{10} - 5.2^2 = 31 - 27.04 = 3.96.$$



Per effettuare confronti in termini di variabilità tra caratteri quantitativi occorre utilizzare una misura relativa, ad esempio il coefficiente di variazione

$$CV = \frac{\sigma}{\mu} = \frac{\sqrt{\text{Var}(Y|x_1)}}{M(Y|x_1)} = \frac{\sqrt{3.96}}{5.2} = \frac{1.99}{5.2} = 0.3827.$$

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.



Si considera la distribuzione condizionata $Y|x_2$

y_j	n_{2j}
0 + 2	5
2 + 6	4
6 + 8	1

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.



Si tratta di un carattere quantitativo con modalità raggruppate in classi.

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.



La media aritmetica

$$M(Y|x_2) = \frac{1}{n_{2\bullet}} \sum_{j=1}^k y_j n_{2j}$$

risulta 2.8.

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.



La media aritmetica

$$M(Y|x_2) = \frac{1}{n_{2\bullet}} \sum_{j=1}^k y_j n_{2j}$$

risulta 2.8.

Si ottiene (considerando come modalità i valori centrali delle classi)

y_j	n_{2j}	$y_j n_{2j}$
1	5	5
4	4	16
7	1	7
	10	28

da cui segue

$$M(Y|x_2) = \frac{1}{n_{2\bullet}} \sum_{j=1}^k y_j n_{2j} = \frac{28}{10} = 2.8.$$

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.



La varianza

$$\text{Var}(Y|x_2) = M(Y^2|x_2) - [M(Y|x_2)]^2$$

risulta 3.96.

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.



La varianza

$$\text{Var}(Y|x_2) = M(Y^2|x_2) - [M(Y|x_2)]^2$$

risulta 3.96. Si ottiene (considerando come modalità i valori centrali delle classi)

y_j	n_{2j}	$y_j n_{2j}$	$y_j^2 n_{2j}$
1	5	5	5
4	4	16	64
7	1	7	49
	10	28	118

da cui segue

$$\text{Var}(Y|x_2) = \frac{118}{10} - 2.8^2 = 11.8 - 7.84 = 3.96.$$



Per effettuare confronti in termini di variabilità tra caratteri quantitativi occorre utilizzare una misura relativa, ad esempio il coefficiente di variazione

$$CV = \frac{\sigma}{\mu} = \frac{\sqrt{\text{Var}(Y|x_2)}}{M(Y|x_2)} = \frac{\sqrt{3.96}}{2.8} = \frac{1.99}{2.8} = 0.7107.$$

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.



Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.

$Y|X = m$ Indice di posizione: $M(Y|X = m)$ valore 5.2 Indice di variabilità: $CV(Y|X = m)$ valore 0.3827

$Y|X = f$ Indice di posizione: $M(Y|X = f)$ valore 2.8 Indice di variabilità: $CV(Y|X = f)$ valore 0.7107

Il reddito medio mensile dei maschi è superiore a quello delle femmine. Il reddito delle femmine presenta un livello di variabilità superiore a quello dei maschi.



Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.

4. Calcolare l'indice di mutabilità normalizzato di X e commentare il risultato ottenuto.



Si considera la distribuzione marginale della variabile X , ottenuta sommando le frequenze di riga

x_j	$n_{j\bullet}$
x_1	10
x_2	10

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.



La mutabile X presenta massima eterogeneità, quindi gli indici di Gini e di Frosini assumono valore massimo, $F_N = G_N = 1$.

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.



5. Senza effettuare calcoli, dire se è possibile concludere che esiste indipendenza stocastica tra X e Y .

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.



Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.

Le variabili non sono stocasticamente indipendenti in quanto le distribuzioni condizionate $Y|X = m$ e $Y|X = f$ non sono somiglianti.

Esercizio 2

Prova scritta di
STATISTICA del
16.01.2019

Giuseppe Boari
Gabriele Cantaluppi



UNIVERSITÀ
CATTOLICA
del Sacro Cuore

Si consideri il seguente prospetto:

t	$NIBM$
2014	—
2015	1
2016	1
2017	0.9
2018	

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.



1. Sapendo che la grandezza oggetto di studio ha subito un decremento medio annuo del 30% dal 2014 al 2018, trovare il valore del *NIBM* del 2018.

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.



Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.

1. Sapendo che la grandezza oggetto di studio ha subito un decremento medio annuo del 30% dal 2014 al 2018, trovare il valore del *NIBM* del 2018.

A un decremento medio annuo del 30% corrisponde una media geometrica dei numeri indici a base mobile pari a 0.7.

Occorre, quindi, risolvere l'equazione $0.7 = \sqrt[4]{1 \cdot 1 \cdot 0.9 \cdot {}_{2017}I_{2018}}$ da cui segue

$${}_{2017}I_{2018} = \frac{0.7^4}{0.9} = 0.2668.$$



2. Ricostruire la serie dei *NIBF* con base 2014.

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.



2. Ricostruire la serie dei *NIBF* con base 2014.

Si ricorda che ${}_{2014}I_{2014} = 1$ e che per valori di $t > 2014$ vale

$${}_{2014}I_t = {}_{2014}I_{2015} \cdot {}_{2015}I_{2016} \cdot \dots \cdot {}_{t-1}I_t.$$

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.



Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.

2. Ricostruire la serie dei *NIBF* con base 2014.

Si ricorda che ${}_{2014}I_{2014} = 1$ e che per valori di $t > 2014$ vale

$${}_{2014}I_t = {}_{2014}I_{2015} \cdot {}_{2015}I_{2016} \cdot \dots \cdot {}_{t-1}I_t.$$

Si ottiene quindi

Anno	<i>NIBM</i>	<i>NIBF</i> (2014)
2014	1.0000	1.0000
2015	1.0000	1.0000
2016	1.0000	1.0000
2017	0.9000	0.9000
2018	0.2668	0.2401



Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.

3. Calcolare il valore del bene nel 2015, sapendo che nel 2017 è risultato pari a 33.



Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.

3. Calcolare il valore del bene nel 2015, sapendo che nel 2017 è risultato pari a 33.

Dal momento che ${}_{2015}I_{2017} = \frac{x_{2017}}{x_{2015}} = 0.9$ segue che $x_{2015} = \frac{33}{0.9} = 36.6667$.



Su un insieme di 8 negozi vengono rilevate le seguenti variabili: Y = fatturato medio giornaliero (in migliaia di Euro), X = spesa annua di riscaldamento (in migliaia di Euro) e L = localizzazione del negozio (c = centrale, p = periferico). I dati sono riportati di seguito:

x_i	0.6	0.6	2	1.4	4.8	0.4	0.6	2
y_i	1.2	1.8	2.2	1.5	3.2	1.4	1.7	3.4
l_i	c	c	p	c	p	c	c	p

1. Rappresentare graficamente le coppie dei dati (x_i, y_i) e la funzione di regressione ipotizzando il fatturato in funzione della spesa di riscaldamento.



La funzione di regressione associa a ogni valore x appartenente all'insieme di definizione della variabile X la media di Y condizionata a x , $M(Y|x)$

x	$M(Y x)$
0.4	1.4
0.6	$\frac{1}{3}(1.2 + 1.7 + 1.8) = 1.5667$
1.4	1.5
2	$\frac{1}{2}(2.2 + 3.4) = 2.8$
4.8	3.2

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.



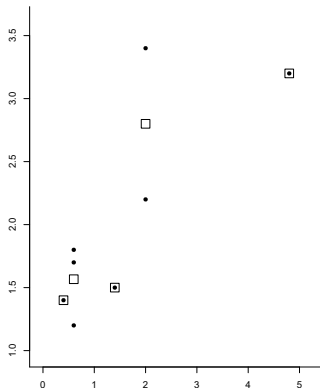
Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.

Le coppie di dati possono essere rappresentate mediante un grafico a dispersione.





Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.

2. Calcolare, secondo il metodo dei minimi quadrati, i parametri dei seguenti modelli di regressione:

$$I) Y^* = a + b \cdot X \quad II) Y^* = c + d \cdot W$$

con W variabile dummy, indicatrice della localizzazione periferica del negozio

Si riportano i calcoli necessari¹ per ottenere le stime dei parametri e gli indici di adattamento del modello

$$Y = a + bX + E$$

x_i	y_i	$x_i y_i$	x_i^2	y_i^2
0.6	1.2	0.72	0.36	1.44
0.6	1.8	1.08	0.36	3.24
2	2.2	4.4	4	4.84
1.4	1.5	2.1	1.96	2.25
4.8	3.2	15.36	23.04	10.24
0.4	1.4	0.56	0.16	1.96
0.6	1.7	1.02	0.36	2.89
2	3.4	6.8	4	11.56
12.4	16.4	32.04	34.24	38.42
1.55	2.05	4.005	4.28	4.8025
$M(X)$	$M(Y)$	$M(XY)$	$M(X^2)$	$M(Y^2)$

$$\hat{b} =$$



Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.

¹Si osserva che per ragioni di stampa alcuni calcoli potrebbero essere duplicati con riferimento ai diversi modelli oggetto di studio

Si riportano i calcoli necessari¹ per ottenere le stime dei parametri e gli indici di adattamento del modello

$$Y = a + bX + E$$

x_i	y_i	$x_i y_i$	x_i^2	y_i^2
0.6	1.2	0.72	0.36	1.44
0.6	1.8	1.08	0.36	3.24
2	2.2	4.4	4	4.84
1.4	1.5	2.1	1.96	2.25
4.8	3.2	15.36	23.04	10.24
0.4	1.4	0.56	0.16	1.96
0.6	1.7	1.02	0.36	2.89
2	3.4	6.8	4	11.56
12.4	16.4	32.04	34.24	38.42
1.55	2.05	4.005	4.28	4.8025
$M(X)$	$M(Y)$	$M(XY)$	$M(X^2)$	$M(Y^2)$

$$\begin{aligned}\hat{b} &= \frac{\text{Cov}(X, Y)}{\text{Var}(X)} = \frac{M(XY) - M(X)M(Y)}{M(X^2) - [M(X)]^2} = \\ &= \frac{4.005 - 1.55 \cdot 2.05}{4.28 - 1.55^2} = \frac{0.8275}{1.8775} = 0.4407 \\ \hat{a} &= \end{aligned}$$

¹Si osserva che per ragioni di stampa alcuni calcoli potrebbero essere duplicati con riferimento ai diversi modelli oggetto di studio



Si riportano i calcoli necessari¹ per ottenere le stime dei parametri e gli indici di adattamento del modello

$$Y = a + bX + E$$

x_i	y_i	$x_i y_i$	x_i^2	y_i^2
0.6	1.2	0.72	0.36	1.44
0.6	1.8	1.08	0.36	3.24
2	2.2	4.4	4	4.84
1.4	1.5	2.1	1.96	2.25
4.8	3.2	15.36	23.04	10.24
0.4	1.4	0.56	0.16	1.96
0.6	1.7	1.02	0.36	2.89
2	3.4	6.8	4	11.56
12.4	16.4	32.04	34.24	38.42
1.55	2.05	4.005	4.28	4.8025
$M(X)$	$M(Y)$	$M(XY)$	$M(X^2)$	$M(Y^2)$

$$\begin{aligned}\hat{b} &= \frac{\text{Cov}(X, Y)}{\text{Var}(X)} = \frac{M(XY) - M(X)M(Y)}{M(X^2) - [M(X)]^2} = \\ &= \frac{4.005 - 1.55 \cdot 2.05}{4.28 - 1.55^2} = \frac{0.8275}{1.8775} = 0.4407\end{aligned}$$

$$\hat{a} = M(Y) - \hat{b}M(X) = 1.3668$$

$$\rho^2 =$$

¹Si osserva che per ragioni di stampa alcuni calcoli potrebbero essere duplicati con riferimento ai diversi modelli oggetto di studio



Si riportano i calcoli necessari¹ per ottenere le stime dei parametri e gli indici di adattamento del modello

$$Y = a + bX + E$$

x_i	y_i	$x_i y_i$	x_i^2	y_i^2
0.6	1.2	0.72	0.36	1.44
0.6	1.8	1.08	0.36	3.24
2	2.2	4.4	4	4.84
1.4	1.5	2.1	1.96	2.25
4.8	3.2	15.36	23.04	10.24
0.4	1.4	0.56	0.16	1.96
0.6	1.7	1.02	0.36	2.89
2	3.4	6.8	4	11.56
12.4	16.4	32.04	34.24	38.42
1.55	2.05	4.005	4.28	4.8025
$M(X)$	$M(Y)$	$M(XY)$	$M(X^2)$	$M(Y^2)$

$$\begin{aligned}\hat{b} &= \frac{\text{Cov}(X, Y)}{\text{Var}(X)} = \frac{M(XY) - M(X)M(Y)}{M(X^2) - [M(X)]^2} = \\ &= \frac{4.005 - 1.55 \cdot 2.05}{4.28 - 1.55^2} = \frac{0.8275}{1.8775} = 0.4407 \\ \hat{a} &= M(Y) - \hat{b}M(X) = 1.3668 \\ \rho^2 &= \frac{[\text{Cov}(X, Y)]^2}{\text{Var}(X) \text{Var}(Y)} = \frac{0.8275^2}{1.8775 \cdot 0.6} = 0.6079\end{aligned}$$

¹Si osserva che per ragioni di stampa alcuni calcoli potrebbero essere duplicati con riferimento ai diversi modelli oggetto di studio



Si riportano i calcoli necessari² per ottenere le stime dei parametri e gli indici di adattamento del modello

$$Y = c + dW + E$$

w_i	y_i	$w_i y_i$	w_i^2	y_i^2
0	1.2	0	0	1.44
0	1.8	0	0	3.24
1	2.2	2.2	1	4.84
0	1.5	0	0	2.25
1	3.2	3.2	1	10.24
0	1.4	0	0	1.96
0	1.7	0	0	2.89
1	3.4	3.4	1	11.56
3	16.4	8.8	3	38.42
0.375	2.05	1.1	0.375	4.8025
$M(W)$	$M(Y)$	$M(WY)$	$M(W^2)$	$M(Y^2)$

$$\hat{d} =$$



Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.

²Si osserva che per ragioni di stampa alcuni calcoli potrebbero essere duplicati con riferimento ai diversi modelli oggetto di studio

Si riportano i calcoli necessari² per ottenere le stime dei parametri e gli indici di adattamento del modello

$$Y = c + dW + E$$

w_i	y_i	$w_i y_i$	w_i^2	y_i^2
0	1.2	0	0	1.44
0	1.8	0	0	3.24
1	2.2	2.2	1	4.84
0	1.5	0	0	2.25
1	3.2	3.2	1	10.24
0	1.4	0	0	1.96
0	1.7	0	0	2.89
1	3.4	3.4	1	11.56
3	16.4	8.8	3	38.42
0.375	2.05	1.1	0.375	4.8025
$M(W)$	$M(Y)$	$M(WY)$	$M(W^2)$	$M(Y^2)$

$$\begin{aligned}\hat{d} &= \frac{\text{Cov}(W, Y)}{\text{Var}(W)} = \frac{M(WY) - M(W)M(Y)}{M(W^2) - [M(W)]^2} = \\ &= \frac{1.1 - 0.375 \cdot 2.05}{0.375 - 0.375^2} = \frac{0.3313}{0.2344} = 1.4134 \\ \hat{c} &= \end{aligned}$$

²Si osserva che per ragioni di stampa alcuni calcoli potrebbero essere duplicati con riferimento ai diversi modelli oggetto di studio



Si riportano i calcoli necessari² per ottenere le stime dei parametri e gli indici di adattamento del modello

$$Y = c + dW + E$$

w_i	y_i	$w_i y_i$	w_i^2	y_i^2
0	1.2	0	0	1.44
0	1.8	0	0	3.24
1	2.2	2.2	1	4.84
0	1.5	0	0	2.25
1	3.2	3.2	1	10.24
0	1.4	0	0	1.96
0	1.7	0	0	2.89
1	3.4	3.4	1	11.56
3	16.4	8.8	3	38.42
0.375	2.05	1.1	0.375	4.8025
$M(W)$	$M(Y)$	$M(WY)$	$M(W^2)$	$M(Y^2)$

$$\begin{aligned}\hat{d} &= \frac{\text{Cov}(W, Y)}{\text{Var}(W)} = \frac{M(WY) - M(W)M(Y)}{M(W^2) - [M(W)]^2} = \\ &= \frac{1.1 - 0.375 \cdot 2.05}{0.375 - 0.375^2} = \frac{0.3313}{0.2344} = 1.4134 \\ \hat{c} &= M(Y) - \hat{d} M(W) = 1.52 \\ \rho^2 &= \end{aligned}$$

²Si osserva che per ragioni di stampa alcuni calcoli potrebbero essere duplicati con riferimento ai diversi modelli oggetto di studio



Si riportano i calcoli necessari² per ottenere le stime dei parametri e gli indici di adattamento del modello

$$Y = c + dW + E$$

w_i	y_i	$w_i y_i$	w_i^2	y_i^2
0	1.2	0	0	1.44
0	1.8	0	0	3.24
1	2.2	2.2	1	4.84
0	1.5	0	0	2.25
1	3.2	3.2	1	10.24
0	1.4	0	0	1.96
0	1.7	0	0	2.89
1	3.4	3.4	1	11.56
3	16.4	8.8	3	38.42
0.375	2.05	1.1	0.375	4.8025
$M(W)$	$M(Y)$	$M(WY)$	$M(W^2)$	$M(Y^2)$

$$\begin{aligned}\hat{d} &= \frac{\text{Cov}(W, Y)}{\text{Var}(W)} = \frac{M(WY) - M(W)M(Y)}{M(W^2) - [M(W)]^2} = \\ &= \frac{1.1 - 0.375 \cdot 2.05}{0.375 - 0.375^2} = \frac{0.3313}{0.2344} = 1.4134 \\ \hat{c} &= M(Y) - \hat{d}M(W) = 1.52 \\ \rho^2 &= \frac{[\text{Cov}(W, Y)]^2}{\text{Var}(W)\text{Var}(Y)} = \frac{0.3313^2}{0.2344 \cdot 0.6} = 0.78\end{aligned}$$

²Si osserva che per ragioni di stampa alcuni calcoli potrebbero essere duplicati con riferimento ai diversi modelli oggetto di studio





Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.

$\hat{a} = 1.3668$, $\hat{b} = 0.4407$, indice adattamento $\rho^2 = 0.6079$ $\hat{c} = 1.52$, $\hat{d} = 1.4133$, indice adattamento $\rho^2 = 0.7803$



3. Confrontare gli indici di adattamento, commentando i risultati.

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.



Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.

3. Confrontare gli indici di adattamento, commentando i risultati.

Il modello migliore è $Y^* = 1.52 + 1.4133 \cdot W$ in quanto caratterizzato da indice di adattamento più elevato.



4. Fornire una interpretazione delle stime dei coefficienti del modello II).

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.



Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.

4. Fornire una interpretazione delle stime dei coefficienti del modello II).

Abbiamo:

$$\hat{c} = M(Y|W = c) = 1.52,$$

e

$$\hat{d} = M(Y|W = p) - M(Y|W = c) = 1.4133$$



5. Senza effettuare calcoli e motivando la risposta, stabilire l'intervallo di possibili valori assumibili da $\eta^2_{Y|X}$.

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.



Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.

5. Senza effettuare calcoli e motivando la risposta, stabilire l'intervallo di possibili valori assumibili da $\eta_{Y|X}^2$.

$0.6079 < \eta_{Y|X}^2 < 1$ in quanto le medie condizionate non sono allineate e le distribuzioni condizionate $Y|X$ non sono degeneri.

Esercizio 4

Prova scritta di
STATISTICA del
16.01.2019

Giuseppe Boari
Gabriele Cantaluppi



UNIVERSITÀ
CATTOLICA
del Sacro Cuore

Il valore di un portafoglio titoli, costituito da un paniere di azioni, si distribuisce come una variabile aleatoria X con distribuzione Normale, di media $\mu = 100$ euro e varianza $\sigma^2 = 100$ euro². Il portafoglio titoli è considerato *svantaggioso* quando il suo valore è inferiore a 82 euro.

1. Calcolare la probabilità che il valore del portafoglio diventi *svantaggioso*.

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.

Esercizio 4



Il valore di un portafoglio titoli, costituito da un paniere di azioni, si distribuisce come una variabile aleatoria X con distribuzione Normale, di media $\mu = 100$ euro e varianza $\sigma^2 = 100$ euro². Il portafoglio titoli è considerato *svantaggioso* quando il suo valore è inferiore a 82 euro.

1. Calcolare la probabilità che il valore del portafoglio diventi *svantaggioso*.

Indicata con $X \sim N(\mu = 100, \sigma^2 = 100)$ la variabile casuale che descrive il valore del portafoglio titoli, occorre determinare la seguente probabilità: $P(X < 82)$.

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.



Il valore di un portafoglio titoli, costituito da un paniere di azioni, si distribuisce come una variabile aleatoria X con distribuzione Normale, di media $\mu = 100$ euro e varianza $\sigma^2 = 100$ euro². Il portafoglio titoli è considerato *svantaggioso* quando il suo valore è inferiore a 82 euro.

1. Calcolare la probabilità che il valore del portafoglio diventi *svantaggioso*.

Indicata con $X \sim N(\mu = 100, \sigma^2 = 100)$ la variabile casuale che descrive il valore del portafoglio titoli, occorre determinare la seguente probabilità: $P(X < 82)$. Abbiamo

$$P(X < 82) =$$

Esercizio 4



Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.

Il valore di un portafoglio titoli, costituito da un paniere di azioni, si distribuisce come una variabile aleatoria X con distribuzione Normale, di media $\mu = 100$ euro e varianza $\sigma^2 = 100$ euro². Il portafoglio titoli è considerato *svantaggioso* quando il suo valore è inferiore a 82 euro.

1. Calcolare la probabilità che il valore del portafoglio diventi *svantaggioso*.

Indicata con $X \sim N(\mu = 100, \sigma^2 = 100)$ la variabile casuale che descrive il valore del portafoglio titoli, occorre determinare la seguente probabilità: $P(X < 82)$. Abbiamo

$$P(X < 82) = P\left(\frac{X - \mu}{\sigma} < \frac{82 - \mu}{\sigma}\right) =$$



Il valore di un portafoglio titoli, costituito da un paniere di azioni, si distribuisce come una variabile aleatoria X con distribuzione Normale, di media $\mu = 100$ euro e varianza $\sigma^2 = 100$ euro². Il portafoglio titoli è considerato *svantaggioso* quando il suo valore è inferiore a 82 euro.

1. Calcolare la probabilità che il valore del portafoglio diventi *svantaggioso*.

Indicata con $X \sim N(\mu = 100, \sigma^2 = 100)$ la variabile casuale che descrive il valore del portafoglio titoli, occorre determinare la seguente probabilità: $P(X < 82)$. Abbiamo

$$P(X < 82) = P\left(\frac{X - \mu}{\sigma} < \frac{82 - \mu}{\sigma}\right) = P\left(Z < \frac{82 - 100}{10}\right) =$$

=



Il valore di un portafoglio titoli, costituito da un paniere di azioni, si distribuisce come una variabile aleatoria X con distribuzione Normale, di media $\mu = 100$ euro e varianza $\sigma^2 = 100$ euro². Il portafoglio titoli è considerato *svantaggioso* quando il suo valore è inferiore a 82 euro.

1. Calcolare la probabilità che il valore del portafoglio diventi *svantaggioso*.

Indicata con $X \sim N(\mu = 100, \sigma^2 = 100)$ la variabile casuale che descrive il valore del portafoglio titoli, occorre determinare la seguente probabilità: $P(X < 82)$. Abbiamo

$$\begin{aligned} P(X < 82) &= P\left(\frac{X - \mu}{\sigma} < \frac{82 - \mu}{\sigma}\right) = P\left(Z < \frac{82 - 100}{10}\right) = \\ &= P(Z < -1.8) = \end{aligned}$$



Il valore di un portafoglio titoli, costituito da un paniere di azioni, si distribuisce come una variabile aleatoria X con distribuzione Normale, di media $\mu = 100$ euro e varianza $\sigma^2 = 100$ euro². Il portafoglio titoli è considerato *svantaggioso* quando il suo valore è inferiore a 82 euro.

1. Calcolare la probabilità che il valore del portafoglio diventi *svantaggioso*.

Indicata con $X \sim N(\mu = 100, \sigma^2 = 100)$ la variabile casuale che descrive il valore del portafoglio titoli, occorre determinare la seguente probabilità: $P(X < 82)$. Abbiamo

$$\begin{aligned} P(X < 82) &= P\left(\frac{X - \mu}{\sigma} < \frac{82 - \mu}{\sigma}\right) = P\left(Z < \frac{82 - 100}{10}\right) = \\ &= P(Z < -1.8) = P(Z > 1.8) = \end{aligned}$$



Il valore di un portafoglio titoli, costituito da un paniere di azioni, si distribuisce come una variabile aleatoria X con distribuzione Normale, di media $\mu = 100$ euro e varianza $\sigma^2 = 100$ euro². Il portafoglio titoli è considerato *svantaggioso* quando il suo valore è inferiore a 82 euro.

1. Calcolare la probabilità che il valore del portafoglio diventi *svantaggioso*.

Indicata con $X \sim N(\mu = 100, \sigma^2 = 100)$ la variabile casuale che descrive il valore del portafoglio titoli, occorre determinare la seguente probabilità: $P(X < 82)$. Abbiamo

$$\begin{aligned} P(X < 82) &= P\left(\frac{X - \mu}{\sigma} < \frac{82 - \mu}{\sigma}\right) = P\left(Z < \frac{82 - 100}{10}\right) = \\ &= P(Z < -1.8) = P(Z > 1.8) = 1 - P(Z \leq 1.8) = \\ &= \end{aligned}$$



Il valore di un portafoglio titoli, costituito da un paniere di azioni, si distribuisce come una variabile aleatoria X con distribuzione Normale, di media $\mu = 100$ euro e varianza $\sigma^2 = 100$ euro². Il portafoglio titoli è considerato *svantaggioso* quando il suo valore è inferiore a 82 euro.

1. Calcolare la probabilità che il valore del portafoglio diventi *svantaggioso*.

Indicata con $X \sim N(\mu = 100, \sigma^2 = 100)$ la variabile casuale che descrive il valore del portafoglio titoli, occorre determinare la seguente probabilità: $P(X < 82)$. Abbiamo

$$\begin{aligned} P(X < 82) &= P\left(\frac{X - \mu}{\sigma} < \frac{82 - \mu}{\sigma}\right) = P\left(Z < \frac{82 - 100}{10}\right) = \\ &= P(Z < -1.8) = P(Z > 1.8) = 1 - P(Z \leq 1.8) = \\ &= 1 - 0.9641 = 0.0359. \end{aligned}$$



2. Indicare quale è il valore limite che definisce il 2% superiore dei valori del portafoglio titoli.

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.



Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.

2. Indicare quale è il valore limite che definisce il 2% superiore dei valori del portafoglio titoli. Il reddito limite mensile che definisce il 2% superiore dei valori del portafoglio titoli corrisponde al percentile di ordine 0.98 della distribuzione di X .

$$\begin{aligned}\{x : P(X \leq x) = 0.98\} &= \left\{x : P\left(\frac{X - \mu}{\sigma} \leq \frac{x - \mu}{\sigma}\right) = 0.98\right\} = \\ &= \left\{x : P\left(Z \leq \frac{x - 100}{10}\right) = 0.98\right\}.\end{aligned}$$



Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.

2. Indicare quale è il valore limite che definisce il 2% superiore dei valori del portafoglio titoli. Il reddito limite mensile che definisce il 2% superiore dei valori del portafoglio titoli corrisponde al percentile di ordine 0.98 della distribuzione di X .

$$\begin{aligned}\{x : P(X \leq x) = 0.98\} &= \left\{x : P\left(\frac{X - \mu}{\sigma} \leq \frac{x - \mu}{\sigma}\right) = 0.98\right\} = \\ &= \left\{x : P\left(Z \leq \frac{x - 100}{10}\right) = 0.98\right\}.\end{aligned}$$

Dalla lettura inversa della Tavola possiamo ricavare il percentile di ordine 0.98, che risulta 2.05 (infatti $P(Z \leq 2.05) = 0.98$). Abbiamo, quindi, $z = \frac{x - 100}{10} = 2.05$ da cui segue $x = 120.5$.



3. Assumendo di avere investito in 5 diversi portafogli titoli, tra loro indipendenti ma con le stesse caratteristiche del portafoglio titoli definito sopra, calcolare la probabilità che al più uno sia *svantaggioso*.

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.



Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.

3. Assumendo di avere investito in 5 diversi portafogli titoli, tra loro indipendenti ma con le stesse caratteristiche del portafoglio titoli definito sopra, calcolare la probabilità che al più uno sia *svantaggioso*.

Sia X la variabile casuale 'n. di portafogli svantaggiosi tra i 5 in cui si è investito'. Abbiamo $X \sim \text{Bin}(n = 5, p = 0.0359)$.

$$P(X \leq 1) =$$



Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.

3. Assumendo di avere investito in 5 diversi portafogli titoli, tra loro indipendenti ma con le stesse caratteristiche del portafoglio titoli definito sopra, calcolare la probabilità che al più uno sia *svantaggioso*.

Sia X la variabile casuale 'n. di portafogli svantaggiosi tra i 5 in cui si è investito'. Abbiamo $X \sim \text{Bin}(n = 5, p = 0.0359)$.

$$P(X \leq 1) = P(X = 0) + P(X = 1) =$$

=



Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.

3. Assumendo di avere investito in 5 diversi portafogli titoli, tra loro indipendenti ma con le stesse caratteristiche del portafoglio titoli definito sopra, calcolare la probabilità che al più uno sia *svantaggioso*.

Sia X la variabile casuale 'n. di portafogli svantaggiosi tra i 5 in cui si è investito'. Abbiamo $X \sim \text{Bin}(n = 5, p = 0.0359)$.

$$\begin{aligned} P(X \leq 1) &= P(X = 0) + P(X = 1) = \\ &= \binom{5}{0} 0.0359^0 \cdot 0.9641^5 + \binom{5}{1} 0.0359^1 \cdot 0.9641^4 = \\ &= \end{aligned}$$



3. Assumendo di avere investito in 5 diversi portafogli titoli, tra loro indipendenti ma con le stesse caratteristiche del portafoglio titoli definito sopra, calcolare la probabilità che al più uno sia *svantaggioso*.

Sia X la variabile casuale 'n. di portafogli svantaggiosi tra i 5 in cui si è investito'. Abbiamo $X \sim \text{Bin}(n = 5, p = 0.0359)$.

$$\begin{aligned} P(X \leq 1) &= P(X = 0) + P(X = 1) = \\ &= \binom{5}{0} 0.0359^0 \cdot 0.9641^5 + \binom{5}{1} 0.0359^1 \cdot 0.9641^4 = \\ &= 0.8328 + 0.1552 = 0.9880. \end{aligned}$$

Tema T302

Prova scritta di STATISTICA del 30.01.2019

Prova scritta di
STATISTICA del
30.01.2019

Giuseppe Boari
Gabriele Cantaluppi



UNIVERSITÀ
CATTOLICA
del Sacro Cuore

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.

Giuseppe Boari
Gabriele Cantaluppi
Università Cattolica del Sacro Cuore, Milano



La seguente tabella riporta i dati di 9 località sciistiche italiane per le variabili X : tipologia di neve (con modalità a = artificiale e f = fresca) e Y : % di impianti aperti, nella settimana di Natale.

x_i	a	a	a	a	a	f	f	f	f
y_i	99	89	94	84	94	100	81	89	91

1. Si confronti la variabilità della variabile Y nei gruppi definiti dalla variabile X .

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.



Per effettuare un confronto in termini di posizione della variabile Y nei 2 gruppi a e f si considerano le due medie aritmetiche di gruppo μ_a e μ_f :

$$\mu_a = \frac{1}{5}(99 + 89 + 94 + 84 + 94) = 92$$

e

$$\mu_f = \frac{1}{4}(100 + 81 + 89 + 91) = 90.25$$

Il livello medio nel gruppo a è superiore a quello osservato nel gruppo f .

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.

Per effettuare un confronto in termini di variabilità della variabile Y nei 2 gruppi a e f si considerano i coefficienti di variazione calcolati per i due gruppi $CV(Y_a)$ e $CV(Y_f)$.



Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.



Per effettuare un confronto in termini di variabilità della variabile Y nei 2 gruppi a e f si considerano i coefficienti di variazione calcolati per i due gruppi $CV(Y_a)$ e $CV(Y_f)$. A tal fine occorre in primo luogo calcolare le varianze nei due gruppi

$$\sigma_a^2 = \frac{1}{5}(99^2 + 89^2 + 94^2 + 84^2 + 94^2) - \mu_a^2 = 26$$

e

$$\sigma_f^2 = \frac{1}{4}(100^2 + 81^2 + 89^2 + 91^2) - \mu_f^2 = 45.6875$$

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.



Per effettuare un confronto in termini di variabilità della variabile Y nei 2 gruppi a e f si considerano i coefficienti di variazione calcolati per i due gruppi $CV(Y_a)$ e $CV(Y_f)$. A tal fine occorre in primo luogo calcolare le varianze nei due gruppi

$$\sigma_a^2 = \frac{1}{5}(99^2 + 89^2 + 94^2 + 84^2 + 94^2) - \mu_a^2 = 26$$

e

$$\sigma_f^2 = \frac{1}{4}(100^2 + 81^2 + 89^2 + 91^2) - \mu_f^2 = 45.6875$$

Ricordando che il coefficiente di variazione è definito dal rapporto tra scarto quadratico medio e media aritmetica risulta

$$CV(Y_a) = \frac{\sigma_a}{\mu_a} = \frac{5.099019514}{92} = 0.0554$$

e

$$CV(Y_f) = \frac{\sigma_f}{\mu_f} = \frac{6.759252917}{90.25} = 0.0749$$

Il livello di variabilità presente nel gruppo a è inferiore a quello osservato nel gruppo f .



2. Si calcolino la varianza between e la varianza within di Y nei gruppi definiti dalla variabile X . Si commentino i risultati ottenuti.

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.



Applicando il risultato di scomposizione della varianza si ottiene:

$$\sigma_{Between}^2 =$$

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.



Applicando il risultato di scomposizione della varianza si ottiene:

$$\begin{aligned}\sigma_{Between}^2 &= \frac{1}{9} \left\{ (\mu_a - \mu)^2 \cdot 5 + (\mu_f - \mu)^2 \cdot 4 \right\} \\ &= \frac{1}{9} \left(\mu_a^2 \cdot 5 + \mu_f^2 \cdot 4 \right) - \mu^2 \\ &= \frac{1}{9} \left(92^2 \cdot 5 + 90.25^2 \cdot 4 \right) - 91.2222^2 = 0.7562\end{aligned}$$

$$\sigma_{Within}^2 =$$

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.



Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.

Applicando il risultato di scomposizione della varianza si ottiene:

$$\begin{aligned}\sigma_{Between}^2 &= \frac{1}{9} \left\{ (\mu_a - \mu)^2 \cdot 5 + (\mu_f - \mu)^2 \cdot 4 \right\} \\ &= \frac{1}{9} \left(\mu_a^2 \cdot 5 + \mu_f^2 \cdot 4 \right) - \mu^2 \\ &= \frac{1}{9} \left(92^2 \cdot 5 + 90.25^2 \cdot 4 \right) - 91.2222^2 = 0.7562\end{aligned}$$

$$\begin{aligned}\sigma_{Within}^2 &= \frac{1}{9} \left(\sigma_a^2 \cdot 5 + \sigma_f^2 \cdot 4 \right) \\ &= \frac{1}{9} (26 \cdot 5 + 45.6875 \cdot 4) = 34.75.\end{aligned}$$

e

$$\sigma^2 =$$



Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.

Applicando il risultato di scomposizione della varianza si ottiene:

$$\begin{aligned}\sigma_{Between}^2 &= \frac{1}{9} \left\{ (\mu_a - \mu)^2 \cdot 5 + (\mu_f - \mu)^2 \cdot 4 \right\} \\ &= \frac{1}{9} \left(\mu_a^2 \cdot 5 + \mu_f^2 \cdot 4 \right) - \mu^2 \\ &= \frac{1}{9} \left(92^2 \cdot 5 + 90.25^2 \cdot 4 \right) - 91.2222^2 = 0.7562\end{aligned}$$

$$\begin{aligned}\sigma_{Within}^2 &= \frac{1}{9} \left(\sigma_a^2 \cdot 5 + \sigma_f^2 \cdot 4 \right) \\ &= \frac{1}{9} (26 \cdot 5 + 45.6875 \cdot 4) = 34.75.\end{aligned}$$

e

$$\sigma^2 = \sigma_{Between}^2 + \sigma_{Within}^2 = 0.7562 + 34.75 = 35.5062.$$



Il rapporto di correlazione $\eta^2_{Y|X}$ risulta:

$$\eta^2_{Y|X} = \frac{\sigma^2_{Between}}{\sigma^2_{Between} + \sigma^2_{Within}} = 0.0213.$$

Sussiste, quindi, un livello di dipendenza molto basso.

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.



3. Si rappresenti graficamente la distribuzione di frequenze della variabile X e si calcolino un indice di posizione e uno di variabilità.

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.



Si considera la distribuzione marginale della variabile X , ottenuta sommando le frequenze di riga

x_i	$n_{i\bullet}$
a	5
f	4

Si tratta di un carattere qualitativo sconnesso che può essere rappresentato graficamente mediante un grafico a bastoncini o a canne d'organo.

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.

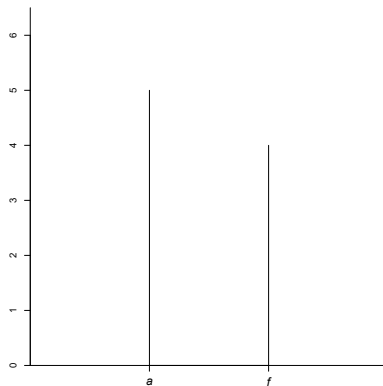


Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.





Si tratta di un carattere qualitativo sconnesso.

L'unico indice di posizione che può essere definito è la moda (categoria alla quale è associata la massima frequenza) che risulta a .

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.



In presenza di un carattere qualitativo la variabilità può essere misurata mediante gli indici di Gini e di Frosini normalizzati

$$G_N = \frac{1 - \sum_{i=1}^h f_{i\bullet}^2}{1 - \frac{1}{h}}$$

e

$$F_N = 1 - \sqrt{1 - G_N}.$$

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.



In presenza di un carattere qualitativo la variabilità può essere misurata mediante gli indici di Gini e di Frosini normalizzati

$$G_N = \frac{1 - \sum_{i=1}^h f_{i\bullet}^2}{1 - \frac{1}{h}}$$

e

$$F_N = 1 - \sqrt{1 - G_N}.$$

Si ottiene:

x_i	$n_{i\bullet}$	$f_{i\bullet}$	$f_{i\bullet}^2$
a	5	0.5556	0.3086
f	4	0.4444	0.1975
	9	1	0.5061

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.



In presenza di un carattere qualitativo la variabilità può essere misurata mediante gli indici di Gini e di Frosini normalizzati

$$G_N = \frac{1 - \sum_{i=1}^h f_{i\bullet}^2}{1 - \frac{1}{h}}$$

e

$$F_N = 1 - \sqrt{1 - G_N}.$$

Si ottiene:

x_i	$n_{i\bullet}$	$f_{i\bullet}$	$f_{i\bullet}^2$
a	5	0.5556	0.3086
f	4	0.4444	0.1975
	9	1	0.5061

Da cui segue:

$$G_N = \frac{1 - 0.5062}{1 - \frac{1}{2}} = 0.9877$$

e

$$F_N = 1 - \sqrt{1 - 0.9877} = 0.8889.$$

Sussiste, quindi, un livello alto di eterogeneità.



4. Costruita la tabella a doppia entrata (X, Y) , con Y riclassificata in 2 classi di uguale ampiezza (pari a 10), a partire da 80, si indichi se le variabili possono ritenersi stocasticamente indipendenti.

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.



4. Costruita la tabella a doppia entrata (X, Y) , con Y riclassificata in 2 classi di uguale ampiezza (pari a 10), a partire da 80, si indichi se le variabili possono ritenersi stocasticamente indipendenti.

La distribuzione congiunta risulta

$X \backslash Y$	80 - 90	90 - 100
a	2	3
f	2	2

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.



Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.

4. Costruita la tabella a doppia entrata (X, Y) , con Y riclassificata in 2 classi di uguale ampiezza (pari a 10), a partire da 80, si indichi se le variabili possono ritenersi stocasticamente indipendenti.

La distribuzione congiunta risulta

$X \backslash Y$	80 - 90	90 - 100
a	2	3
f	2	2

Le due variabili non possono ritenersi stocasticamente indipendenti dal momento che le distribuzioni condizionate non sono tra loro somiglianti.



Si consideri la seguente serie dei numeri indici a base fissa, con base 2009:

<i>anno</i>	<i>NIBF</i>
2009	1
2010	0.90
2011	0.93
2012	1.02
2013	1.09

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.



1. Si costruisca la corrispondente serie dei numeri indici a base mobile.

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.



1. Si costruisca la corrispondente serie dei numeri indici a base mobile.

Si ricorda che ${}_{t-1}I_t = \frac{x_t}{x_{t-1}} = \frac{{}_{2009}I_t}{{}_{2009}I_{t-1}}$.

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.



1. Si costruisca la corrispondente serie dei numeri indici a base mobile.

Si ricorda che ${}_t I_t = \frac{x_t}{x_{t-1}} = \frac{{}_{2009}I_t}{{}_{2009}I_{t-1}}$. Si ottiene quindi

<i>Anno</i>	<i>NIBF</i>	<i>NIBM</i>
2009	1.00	
2010	0.90	0.9000
2011	0.93	1.0333
2012	1.02	1.0968
2013	1.09	1.0686

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.



Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.

2. Sapendo che $x_{2011} = 26$, si calcoli x_{2013} .



Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.

2. Sapendo che $x_{2011} = 26$, si calcoli x_{2013} .

$$x_{2013} = x_{2011} \cdot {}_{2011}I_{2013} = 26 \cdot \frac{1.09}{0.93} = 30.4731$$



3. Si calcoli il tasso medio di variazione dal 2011 al 2013 e si commenti il risultato ottenuto.

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.



3. Si calcoli il tasso medio di variazione dal 2011 al 2013 e si commenti il risultato ottenuto.
Si ricorda che

$${}_{2011}I_{2013} = {}_{2011}I_{2012} \cdot {}_{2012}I_{2013}$$

Si deve trovare il numero indice a base mobile medio.

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.



Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.

3. Si calcoli il tasso medio di variazione dal 2011 al 2013 e si commenti il risultato ottenuto. Si ricorda che

$${}_{2011}l_{2013} = {}_{2011}l_{2012} \cdot {}_{2012}l_{2013}$$

Si deve trovare il numero indice a base mobile medio. Applicando il criterio di scelta della media secondo Chisini si ottiene

$${}_{2011}l_{2013} = {}_{2011}l_{2012} \cdot {}_{2012}l_{2013} = \alpha \cdot \alpha = \alpha^2.$$



Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.

3. Si calcoli il tasso medio di variazione dal 2011 al 2013 e si commenti il risultato ottenuto. Si ricorda che

$${}_{2011}l_{2013} = {}_{2011}l_{2012} \cdot {}_{2012}l_{2013}$$

Si deve trovare il numero indice a base mobile medio. Applicando il criterio di scelta della media secondo Chisini si ottiene

$${}_{2011}l_{2013} = {}_{2011}l_{2012} \cdot {}_{2012}l_{2013} = \alpha \cdot \alpha = \alpha^2.$$

Quindi

$$\alpha = \sqrt[2]{{}_{2011}l_{2013}} = \sqrt[2]{{}_{2011}l_{2012} \cdot {}_{2012}l_{2013}} = 1.0826.$$

Da cui segue una variazione relativa media pari a 0.0826, corrispondente a un incremento medio annuo tra il 2011 e il 2013 del 8.26%.



Il seguente prospetto riporta i valori rilevati su 8 unità statistiche con riferimento a due caratteri quantitativi X e Y

x_i	1	3	1	2	3	2	3	1
y_i	6	5	5	1	5	1	5	4

1. Si rappresentino graficamente le coppie dei dati e la funzione di regressione ipotizzando Y in funzione di X .

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.



La funzione di regressione associa a ogni valore x appartenente all'insieme di definizione della variabile X la media di Y condizionata a x , $M(Y|x)$

x	$M(Y x)$
1	$\frac{1}{3}(4 + 5 + 6) = 5$
2	1
3	5

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.



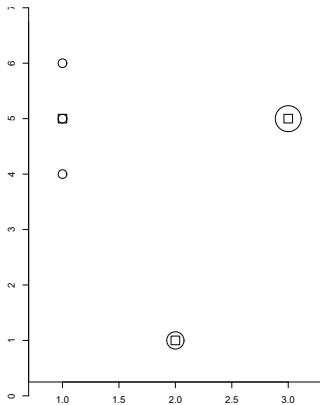
Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.

Le coppie di dati possono essere rappresentate mediante un *bubble plot diagram*.





2. Si calcolino, secondo il metodo dei minimi quadrati, i parametri e l'indice di adattamento dei seguenti modelli di regressione, individuando il modello migliore:

$$I) Y^* = a + b \cdot X \qquad II) Y^* = c + d \cdot X^2$$

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.

Si riportano i calcoli necessari¹ per ottenere le stime dei parametri e gli indici di adattamento del modello

$$Y = a + bX + E$$

x_i	y_i	$x_i y_i$	x_i^2	y_i^2
1	6	6	1	36
3	5	15	9	25
1	5	5	1	25
2	1	2	4	1
3	5	15	9	25
2	1	2	4	1
3	5	15	9	25
1	4	4	1	16
16	32	64	38	154
2	4	8	4.75	19.25
$M(X)$	$M(Y)$	$M(XY)$	$M(X^2)$	$M(Y^2)$

$$\hat{b} =$$



Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.

¹Si osserva che per ragioni di stampa alcuni calcoli potrebbero essere duplicati con riferimento ai diversi modelli oggetto di studio

Si riportano i calcoli necessari¹ per ottenere le stime dei parametri e gli indici di adattamento del modello

$$Y = a + bX + E$$

x_i	y_i	$x_i y_i$	x_i^2	y_i^2
1	6	6	1	36
3	5	15	9	25
1	5	5	1	25
2	1	2	4	1
3	5	15	9	25
2	1	2	4	1
3	5	15	9	25
1	4	4	1	16
16	32	64	38	154
2	4	8	4.75	19.25
$M(X)$	$M(Y)$	$M(XY)$	$M(X^2)$	$M(Y^2)$

$$\begin{aligned}\hat{b} &= \frac{\text{Cov}(X, Y)}{\text{Var}(X)} = \frac{M(XY) - M(X)M(Y)}{M(X^2) - [M(X)]^2} = \\ &= \frac{8 - 2 \cdot 4}{4.75 - 2^2} = \frac{0}{0.75} = 0 \\ \hat{a} &= \end{aligned}$$

¹Si osserva che per ragioni di stampa alcuni calcoli potrebbero essere duplicati con riferimento ai diversi modelli oggetto di studio



Si riportano i calcoli necessari¹ per ottenere le stime dei parametri e gli indici di adattamento del modello

$$Y = a + bX + E$$

x_i	y_i	$x_i y_i$	x_i^2	y_i^2
1	6	6	1	36
3	5	15	9	25
1	5	5	1	25
2	1	2	4	1
3	5	15	9	25
2	1	2	4	1
3	5	15	9	25
1	4	4	1	16
16	32	64	38	154
2	4	8	4.75	19.25
$M(X)$	$M(Y)$	$M(XY)$	$M(X^2)$	$M(Y^2)$

$$\begin{aligned}\hat{b} &= \frac{\text{Cov}(X, Y)}{\text{Var}(X)} = \frac{M(XY) - M(X)M(Y)}{M(X^2) - [M(X)]^2} = \\ &= \frac{8 - 2 \cdot 4}{4.75 - 2^2} = \frac{0}{0.75} = 0 \\ \hat{a} &= M(Y) - \hat{b}M(X) = 4 \\ \rho^2 &= \end{aligned}$$

¹Si osserva che per ragioni di stampa alcuni calcoli potrebbero essere duplicati con riferimento ai diversi modelli oggetto di studio



Si riportano i calcoli necessari¹ per ottenere le stime dei parametri e gli indici di adattamento del modello

$$Y = a + bX + E$$

x_i	y_i	$x_i y_i$	x_i^2	y_i^2
1	6	6	1	36
3	5	15	9	25
1	5	5	1	25
2	1	2	4	1
3	5	15	9	25
2	1	2	4	1
3	5	15	9	25
1	4	4	1	16
16	32	64	38	154
2	4	8	4.75	19.25
$M(X)$	$M(Y)$	$M(XY)$	$M(X^2)$	$M(Y^2)$

$$\begin{aligned}\hat{b} &= \frac{\text{Cov}(X, Y)}{\text{Var}(X)} = \frac{M(XY) - M(X)M(Y)}{M(X^2) - [M(X)]^2} = \\ &= \frac{8 - 2 \cdot 4}{4.75 - 2^2} = \frac{0}{0.75} = 0 \\ \hat{a} &= M(Y) - \hat{b}M(X) = 4 \\ \rho^2 &= \frac{[\text{Cov}(X, Y)]^2}{\text{Var}(X) \text{Var}(Y)} = \frac{0^2}{0.75 \cdot 3.25} = 0\end{aligned}$$

¹Si osserva che per ragioni di stampa alcuni calcoli potrebbero essere duplicati con riferimento ai diversi modelli oggetto di studio



Si riportano i calcoli necessari² per ottenere le stime dei parametri e gli indici di adattamento del modello

$$Y = \alpha + \beta X^2 + E$$

x_i^2	y_i	$x_i^2 y_i$	$x_i^{2^2}$	y_i^2
1	6	6	1	36
9	5	45	81	25
1	5	5	1	25
4	1	4	16	1
9	5	45	81	25
4	1	4	16	1
9	5	45	81	25
1	4	4	1	16
38	32	158	278	154
4.75	4	19.75	34.75	19.25
$M(X^2)$	$M(Y)$	$M(X^2 Y)$	$M(X^{2^2})$	$M(Y^2)$

$$\hat{\beta} =$$



Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.

²Si osserva che per ragioni di stampa alcuni calcoli potrebbero essere duplicati con riferimento ai diversi modelli oggetto di studio

Si riportano i calcoli necessari² per ottenere le stime dei parametri e gli indici di adattamento del modello

$$Y = \alpha + \beta X^2 + E$$

x_i^2	y_i	$x_i^2 y_i$	x_i^{22}	y_i^2
1	6	6	1	36
9	5	45	81	25
1	5	5	1	25
4	1	4	16	1
9	5	45	81	25
4	1	4	16	1
9	5	45	81	25
1	4	4	1	16
38	32	158	278	154
4.75	4	19.75	34.75	19.25
$M(X^2)$	$M(Y)$	$M(X^2 Y)$	$M(X^{22})$	$M(Y^2)$

$$\begin{aligned}\hat{\beta} &= \frac{\text{Cov}(X^2, Y)}{\text{Var}(X^2)} = \frac{M(X^2 Y) - M(X^2) M(Y)}{M(X^{22}) - [M(X^2)]^2} = \\ &= \frac{19.75 - 4.75 \cdot 4}{34.75 - 4.75^2} = \frac{0.75}{12.1875} = 0.0615\end{aligned}$$

$$\hat{\alpha} =$$

²Si osserva che per ragioni di stampa alcuni calcoli potrebbero essere duplicati con riferimento ai diversi modelli oggetto di studio



Si riportano i calcoli necessari² per ottenere le stime dei parametri e gli indici di adattamento del modello

$$Y = \alpha + \beta X^2 + E$$

x_i^2	y_i	$x_i^2 y_i$	x_i^{22}	y_i^2
1	6	6	1	36
9	5	45	81	25
1	5	5	1	25
4	1	4	16	1
9	5	45	81	25
4	1	4	16	1
9	5	45	81	25
1	4	4	1	16
38	32	158	278	154
4.75	4	19.75	34.75	19.25
$M(X^2)$	$M(Y)$	$M(X^2 Y)$	$M(X^{22})$	$M(Y^2)$

$$\hat{\beta} = \frac{\text{Cov}(X^2, Y)}{\text{Var}(X^2)} = \frac{M(X^2 Y) - M(X^2) M(Y)}{M(X^{22}) - [M(X^2)]^2} =$$

$$= \frac{19.75 - 4.75 \cdot 4}{34.75 - 4.75^2} = \frac{0.75}{12.1875} = 0.0615$$

$$\hat{\alpha} = M(Y) - \hat{\beta} M(X^2) = 3.7077$$

$$\rho^2 =$$

²Si osserva che per ragioni di stampa alcuni calcoli potrebbero essere duplicati con riferimento ai diversi modelli oggetto di studio



Si riportano i calcoli necessari² per ottenere le stime dei parametri e gli indici di adattamento del modello

$$Y = \alpha + \beta X^2 + E$$

x_i^2	y_i	$x_i^2 y_i$	$x_i^{2^2}$	y_i^2
1	6	6	1	36
9	5	45	81	25
1	5	5	1	25
4	1	4	16	1
9	5	45	81	25
4	1	4	16	1
9	5	45	81	25
1	4	4	1	16
38	32	158	278	154
4.75	4	19.75	34.75	19.25
$M(X^2)$	$M(Y)$	$M(X^2 Y)$	$M(X^{2^2})$	$M(Y^2)$

$$\hat{\beta} = \frac{\text{Cov}(X^2, Y)}{\text{Var}(X^2)} = \frac{M(X^2 Y) - M(X^2) M(Y)}{M(X^{2^2}) - [M(X^2)]^2} =$$

$$= \frac{19.75 - 4.75 \cdot 4}{34.75 - 4.75^2} = \frac{0.75}{12.1875} = 0.0615$$

$$\hat{\alpha} = M(Y) - \hat{\beta} M(X^2) = 3.7077$$

$$\rho^2 = \frac{[\text{Cov}(X^2, Y)]^2}{\text{Var}(X^2) \text{Var}(Y)} = \frac{0.75^2}{12.1875 \cdot 3.25} = 0.0142$$

²Si osserva che per ragioni di stampa alcuni calcoli potrebbero essere duplicati con riferimento ai diversi modelli oggetto di studio





Il modello migliore è $Y^* = 3.7077 + 0.0615X^2$.

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.



3. Si interpretino i coefficienti del modello I).

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.



3. Si interpretino i coefficienti del modello I).

In base al primo modello il valore atteso della Y non subisce variazioni al variare del livello della X e risulta pari a 4.

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.



4. Si determinino i possibili valori del rapporto di correlazione $\eta^2_{Y|X}$ motivando la risposta.

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.



4. Si determinino i possibili valori del rapporto di correlazione $\eta^2_{Y|X}$ motivando la risposta.
 $0.0142 < \eta^2_{Y|X} < 1$ in quanto le medie condizionate non sono una parabola avente come
asse l'asse delle ordinate e le distribuzioni condizionate $Y|X$ non sono degeneri.

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.



5. Per il modello migliore, si fornisca la previsione in corrispondenza di $X = 5$.

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.



Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.

5. Per il modello migliore, si fornisca la previsione in corrispondenza di $X = 5$.

$$Y^*(5) = 3.7077 + 0.0615 \cdot 5^2 = 3.7077 + 0.0615 \cdot 25 = 5.2462.$$



Da una ricerca condotta da una rivista automobilistica è risultato che le auto fuoristrada (tipo SUV, Van e Pickup) sono più soggette a ribaltamento se coinvolte in incidenti stradali rispetto alle altre auto. Nel 1997 il 24% degli incidenti ha comportato un ribaltamento di auto; di questi il 15 % ha coinvolto dei fuoristrada, mentre negli incidenti senza ribaltamento la percentuale di fuoristrada coinvolti scende al 5 %.

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.



1. Si calcoli la probabilità che in un incidente, estratto a caso tra i sinistri del 1997, sia coinvolto un fuoristrada.

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.



Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.

1. Si calcoli la probabilità che in un incidente, estratto a caso tra i sinistri del 1997, sia coinvolto un fuoristrada.

Indicati con F ed $\bar{F}$ gli eventi 'auto fuoristrada' e 'altre auto' e con R e $\bar{R}$ gli eventi 'incidente con ribaltamento' e 'incidente senza ribaltamento' sono disponibili le seguenti informazioni:

$$P(R) = 0.24, P(F|R) = 0.15, P(F|\bar{R}) = 0.05.$$

Applicando la legge delle probabilità totali abbiamo:



Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.

1. Si calcoli la probabilità che in un incidente, estratto a caso tra i sinistri del 1997, sia coinvolto un fuoristrada.

Indicati con F ed $\bar{F}$ gli eventi 'auto fuoristrada' e 'altre auto' e con R e $\bar{R}$ gli eventi 'incidente con ribaltamento' e 'incidente senza ribaltamento' sono disponibili le seguenti informazioni:

$$P(R) = 0.24, P(F|R) = 0.15, P(F|\bar{R}) = 0.05.$$

Applicando la legge delle probabilità totali abbiamo:

$$P(F) =$$



Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.

1. Si calcoli la probabilità che in un incidente, estratto a caso tra i sinistri del 1997, sia coinvolto un fuoristrada.

Indicati con F ed $\bar{F}$ gli eventi 'auto fuoristrada' e 'altre auto' e con R e $\bar{R}$ gli eventi 'incidente con ribaltamento' e 'incidente senza ribaltamento' sono disponibili le seguenti informazioni:

$$P(R) = 0.24, P(F|R) = 0.15, P(F|\bar{R}) = 0.05.$$

Applicando la legge delle probabilità totali abbiamo:

$$P(F) = P(F \cap (R \cup \bar{R})) = P((F \cap R) \cup (F \cap \bar{R})) =$$
$$=$$



Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.

1. Si calcoli la probabilità che in un incidente, estratto a caso tra i sinistri del 1997, sia coinvolto un fuoristrada.

Indicati con F ed $\bar{F}$ gli eventi 'auto fuoristrada' e 'altre auto' e con R e $\bar{R}$ gli eventi 'incidente con ribaltamento' e 'incidente senza ribaltamento' sono disponibili le seguenti informazioni:

$$P(R) = 0.24, P(F|R) = 0.15, P(F|\bar{R}) = 0.05.$$

Applicando la legge delle probabilità totali abbiamo:

$$\begin{aligned} P(F) &= P(F \cap (R \cup \bar{R})) = P((F \cap R) \cup (F \cap \bar{R})) = \\ &= P(F \cap R) + P(F \cap \bar{R}) = \end{aligned}$$



1. Si calcoli la probabilità che in un incidente, estratto a caso tra i sinistri del 1997, sia coinvolto un fuoristrada.

Indicati con F ed $\bar{F}$ gli eventi 'auto fuoristrada' e 'altre auto' e con R e $\bar{R}$ gli eventi 'incidente con ribaltamento' e 'incidente senza ribaltamento' sono disponibili le seguenti informazioni:

$$P(R) = 0.24, P(F|R) = 0.15, P(F|\bar{R}) = 0.05.$$

Applicando la legge delle probabilità totali abbiamo:

$$\begin{aligned} P(F) &= P(F \cap (R \cup \bar{R})) = P((F \cap R) \cup (F \cap \bar{R})) = \\ &= P(F \cap R) + P(F \cap \bar{R}) = P(R) \cdot P(F|R) + P(\bar{R}) \cdot P(F|\bar{R}) = \\ &= \end{aligned}$$



1. Si calcoli la probabilità che in un incidente, estratto a caso tra i sinistri del 1997, sia coinvolto un fuoristrada.

Indicati con F ed $\bar{F}$ gli eventi 'auto fuoristrada' e 'altre auto' e con R e $\bar{R}$ gli eventi 'incidente con ribaltamento' e 'incidente senza ribaltamento' sono disponibili le seguenti informazioni:

$$P(R) = 0.24, P(F|R) = 0.15, P(F|\bar{R}) = 0.05.$$

Applicando la legge delle probabilità totali abbiamo:

$$\begin{aligned} P(F) &= P(F \cap (R \cup \bar{R})) = P((F \cap R) \cup (F \cap \bar{R})) = \\ &= P(F \cap R) + P(F \cap \bar{R}) = P(R) \cdot P(F|R) + P(\bar{R}) \cdot P(F|\bar{R}) = \\ &= P(R) \cdot P(F|R) + (1 - P(R)) \cdot P(F|\bar{R}) = \\ &= \end{aligned}$$



1. Si calcoli la probabilità che in un incidente, estratto a caso tra i sinistri del 1997, sia coinvolto un fuoristrada.

Indicati con F ed $\bar{F}$ gli eventi 'auto fuoristrada' e 'altre auto' e con R e $\bar{R}$ gli eventi 'incidente con ribaltamento' e 'incidente senza ribaltamento' sono disponibili le seguenti informazioni:

$$P(R) = 0.24, P(F|R) = 0.15, P(F|\bar{R}) = 0.05.$$

Applicando la legge delle probabilità totali abbiamo:

$$\begin{aligned} P(F) &= P(F \cap (R \cup \bar{R})) = P((F \cap R) \cup (F \cap \bar{R})) = \\ &= P(F \cap R) + P(F \cap \bar{R}) = P(R) \cdot P(F|R) + P(\bar{R}) \cdot P(F|\bar{R}) = \\ &= P(R) \cdot P(F|R) + (1 - P(R)) \cdot P(F|\bar{R}) = \\ &= 0.24 \cdot 0.15 + 0.76 \cdot 0.05 = 0.036 + 0.038 = 0.074 \end{aligned}$$



2. Si calcoli la probabilità che un incidente, estratto a caso tra i sinistri del 1997, sia con ribaltamento sapendo che in esso è coinvolto un fuoristrada.

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.



Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.

2. Si calcoli la probabilità che un incidente, estratto a caso tra i sinistri del 1997, sia con ribaltamento sapendo che in esso è coinvolto un fuoristrada.

$$P(R|F) = \frac{P(R \cap F)}{P(F)} = \frac{0.036}{0.074} = 0.4865$$



Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.

3. Si calcoli la probabilità che su 10 incidenti, estratti a caso tra i sinistri del 1997, in al più 2 ci sia un ribaltamento di un fuoristrada.



Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.

3. Si calcoli la probabilità che su 10 incidenti, estratti a caso tra i sinistri del 1997, in al più 2 ci sia un ribaltamento di un fuoristrada.

Sia X la variabile casuale 'n. di incidenti con ribaltamento di un fuoristrada tra i 10 estratti'.

Supponendo che la popolazione sia composta da un numero elevato di soggetti le estrazioni 'effettuate in maniera casuale' possono ritenersi indipendenti. Abbiamo

$$X \sim \text{Bin}(n = 10, p = 0.036)$$

e

$$P(X \leq 2) =$$



Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.

3. Si calcoli la probabilità che su 10 incidenti, estratti a caso tra i sinistri del 1997, in al più 2 ci sia un ribaltamento di un fuoristrada.

Sia X la variabile casuale 'n. di incidenti con ribaltamento di un fuoristrada tra i 10 estratti'.

Supponendo che la popolazione sia composta da un numero elevato di soggetti le estrazioni 'effettuate in maniera casuale' possono ritenersi indipendenti. Abbiamo

$$X \sim \text{Bin}(n = 10, p = 0.036)$$

e

$$P(X \leq 2) = P(X = 0) + P(X = 1) + P(X = 2) =$$



3. Si calcoli la probabilità che su 10 incidenti, estratti a caso tra i sinistri del 1997, in al più 2 ci sia un ribaltamento di un fuoristrada.

Sia X la variabile casuale 'n. di incidenti con ribaltamento di un fuoristrada tra i 10 estratti'.

Supponendo che la popolazione sia composta da un numero elevato di soggetti le estrazioni 'effettuate in maniera casuale' possono ritenersi indipendenti. Abbiamo

$$X \sim \text{Bin}(n = 10, p = 0.036)$$

e

$$\begin{aligned} P(X \leq 2) &= P(X = 0) + P(X = 1) + P(X = 2) = \\ &= \binom{10}{0} 0.036^0 \cdot 0.964^{10} + \binom{10}{1} 0.036^1 \cdot 0.964^9 + \\ &\quad + \binom{10}{2} 0.036^2 \cdot 0.964^8 = \end{aligned}$$



3. Si calcoli la probabilità che su 10 incidenti, estratti a caso tra i sinistri del 1997, in al più 2 ci sia un ribaltamento di un fuoristrada.

Sia X la variabile casuale 'n. di incidenti con ribaltamento di un fuoristrada tra i 10 estratti'.

Supponendo che la popolazione sia composta da un numero elevato di soggetti le estrazioni 'effettuate in maniera casuale' possono ritenersi indipendenti. Abbiamo

$$X \sim \text{Bin}(n = 10, p = 0.036)$$

e

$$\begin{aligned} P(X \leq 2) &= P(X = 0) + P(X = 1) + P(X = 2) = \\ &= \binom{10}{0} 0.036^0 \cdot 0.964^{10} + \binom{10}{1} 0.036^1 \cdot 0.964^9 + \\ &\quad + \binom{10}{2} 0.036^2 \cdot 0.964^8 = \\ &= \end{aligned}$$



3. Si calcoli la probabilità che su 10 incidenti, estratti a caso tra i sinistri del 1997, in al più 2 ci sia un ribaltamento di un fuoristrada.

Sia X la variabile casuale 'n. di incidenti con ribaltamento di un fuoristrada tra i 10 estratti'. Supponendo che la popolazione sia composta da un numero elevato di soggetti le estrazioni 'effettuate in maniera casuale' possono ritenersi indipendenti. Abbiamo

$$X \sim \text{Bin}(n = 10, p = 0.036)$$

e

$$\begin{aligned} P(X \leq 2) &= P(X = 0) + P(X = 1) + P(X = 2) = \\ &= \binom{10}{0} 0.036^0 \cdot 0.964^{10} + \binom{10}{1} 0.036^1 \cdot 0.964^9 + \\ &\quad + \binom{10}{2} 0.036^2 \cdot 0.964^8 = \\ &= 0.6931 + 0.2588 + 0.0435 = 0.9954. \end{aligned}$$

Tema T303

Prova scritta di STATISTICA del 19.02.2019

Prova scritta di
STATISTICA del
19.02.2019

Giuseppe Boari
Gabriele Cantaluppi



UNIVERSITÀ
CATTOLICA
del Sacro Cuore

Esercizio 1.

Esercizio 2

Esercizio 3.

Esercizio 4.

Giuseppe Boari
Gabriele Cantaluppi
Università Cattolica del Sacro Cuore, Milano



Il prospetto seguente riporta alcune informazioni sulla distribuzione congiunta del carattere qualitativo X e del carattere quantitativo Y .

$X \setminus Y$	2	3	4	
x_1		8		
x_2		4	2	8
				24

1. Sapendo che la varianza within di Y rispetto ai gruppi definiti dalle categorie della variabile X è uguale alla varianza della distribuzione marginale di Y , si costruisca la distribuzione delle frequenze congiunte delle due variabili.

Esercizio 1.

Esercizio 2

Esercizio 3.

Esercizio 4.



Se la varianza within è uguale alla varianza della distribuzione marginale, allora la varianza between è nulla e le medie condizionate sono tra loro uguali. Dal momento che la frequenza $n_{21} = 2$, $M(Y|x_2) = 3$ (con $Y|x_2$ distribuzione simmetrica). Affinché $M(Y|x_1) = 3$ occorre fissare $n_{11} = n_{13}$.

$X \setminus Y$	2	3	4	
x_1	4	8	4	16
x_2	2	4	2	8
	6	12	6	24

Esercizio 1.

Esercizio 2

Esercizio 3.

Esercizio 4.



Esercizio 1.

Esercizio 2

Esercizio 3.

Esercizio 4.

2. Si rappresentino graficamente le distribuzioni marginali e si confrontino le distribuzioni condizionate $Y|X_1$ e $Y|X_2$ con opportuni indici di posizione e di variabilità.



Si considera la distribuzione marginale della variabile X , ottenuta sommando le frequenze di riga

x_i	$n_{i\bullet}$
x_1	16
x_2	8

Si tratta di un carattere qualitativo sconnesso che può essere rappresentato graficamente mediante un grafico a bastoncini o a canne d'organo.

Esercizio 1.

Esercizio 2

Esercizio 3.

Esercizio 4.

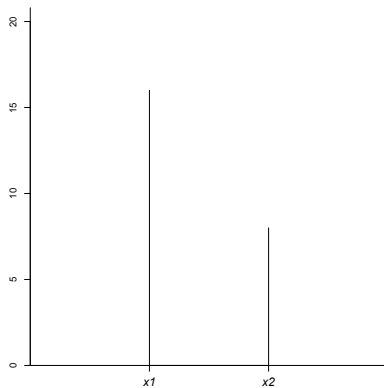


Esercizio 1.

Esercizio 2

Esercizio 3.

Esercizio 4.





Si considera la distribuzione marginale della variabile Y , ottenuta sommando le frequenze di colonna

y_j	$n_{\bullet j}$
2	6
3	12
4	6

Esercizio 1.

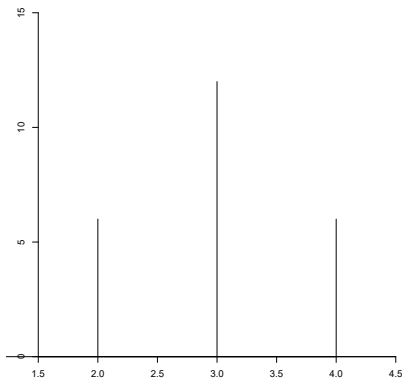
Esercizio 2

Esercizio 3.

Esercizio 4.



Si tratta di un carattere quantitativo con modalità non raggruppate in classi che può essere rappresentato graficamente mediante un grafico a bastoncini.



Esercizio 1.

Esercizio 2

Esercizio 3.

Esercizio 4.



Si osserva che le variabili X e Y sono stocasticamente indipendenti; quindi, le distribuzioni condizionate saranno caratterizzate dallo stesso livello medio e da eguale variabilità. È, quindi, sufficiente procedere al calcolo degli indici per una sola delle distribuzioni condizionate.

[Esercizio 1.](#)

[Esercizio 2](#)

[Esercizio 3.](#)

[Esercizio 4.](#)



Si considera la distribuzione condizionata $Y|x_1$

y_j	n_{1j}
2	4
3	8
4	4

Esercizio 1.

Esercizio 2

Esercizio 3.

Esercizio 4.



Si tratta di un carattere quantitativo con modalità non raggruppate in classi.

Esercizio 1.

Esercizio 2

Esercizio 3.

Esercizio 4.



Esercizio 1.

Esercizio 2

Esercizio 3.

Esercizio 4.

Si osserva che la distribuzione è simmetrica rispetto a 3; la mediana (percentile di ordine 0.50) risulta $y_{0.5} = 3$.



La media aritmetica

$$M(Y|x_1) = \frac{1}{n_{1\bullet}} \sum_{j=1}^k y_j n_{1j}$$

risulta 3.

Si osserva, infatti, che la distribuzione è simmetrica.

Esercizio 1.

Esercizio 2

Esercizio 3.

Esercizio 4.



La varianza

$$\text{Var}(Y|x_1) = M(Y^2|x_1) - [M(Y|x_1)]^2$$

risulta 0.5.

Esercizio 1.

Esercizio 2

Esercizio 3.

Esercizio 4.



La varianza

$$\text{Var}(Y|x_1) = M(Y^2|x_1) - [M(Y|x_1)]^2$$

risulta 0.5. Si ottiene

y_j	n_{1j}	$y_j n_{1j}$	$y_j^2 n_{1j}$
2	4	8	16
3	8	24	72
4	4	16	64
	16	48	152

da cui segue

$$\text{Var}(Y|x_1) = \frac{152}{16} - 3^2 = 9.5 - 9 = 0.5.$$

Esercizio 1.

Esercizio 2

Esercizio 3.

Esercizio 4.



Per effettuare confronti in termini di variabilità tra caratteri quantitativi occorre utilizzare una misura relativa, ad esempio il coefficiente di variazione

$$CV = \frac{\sigma}{\mu} = \frac{\sqrt{\text{Var}(Y|x_1)}}{M(Y|x_1)} = \frac{\sqrt{0.5}}{3} = \frac{0.7071}{3} = 0.2357.$$

Esercizio 1.

Esercizio 2

Esercizio 3.

Esercizio 4.



$Y|X = x_1$ Indice di posizione: $M(Y|X = x_1)$ valore 3 Indice di variabilità: $CV(Y|X = x_1)$ valore 0.2357

$Y|X = x_2$ Indice di posizione: $M(Y|X = x_2)$ valore 3 Indice di variabilità: $CV(Y|X = x_2)$ valore 0.2357

Sussiste indipendenza stocastica. Le distribuzioni condizionate sono caratterizzate dallo stesso livello medio e da eguale variabilità.

Esercizio 1.

Esercizio 2

Esercizio 3.

Esercizio 4.



3. Si determini il grado di connessione tra X e Y mediante il calcolo di un opportuno indice.

Esercizio 1.

Esercizio 2

Esercizio 3.

Esercizio 4.



3. Si determini il grado di connessione tra X e Y mediante il calcolo di un opportuno indice. Dal momento che sussiste indipendenza stocastica vi è assenza di connessione

Esercizio 1.

Esercizio 2

Esercizio 3.

Esercizio 4.



La quotazione del Bitcoin ha registrato negli ultimi 5 giorni le seguenti variazioni percentuali:

t	<i>variaz.relativa</i> ⁰ %
1	—
2	+11%
3	−11%
4	+12%
5	−12%

1. Si costruiscano la serie dei numeri indici a base mobile e la serie dei numeri indici a base fissa con base al tempo 3 per la quotazione del Bitcoin.

Esercizio 1.

Esercizio 2

Esercizio 3.

Esercizio 4.



Dal momento che la variazione relativa si ottiene sottraendo 1 al corrispondente numero indice (misura relativa), è sufficiente, al fine di ottenere la serie dei numeri indici a base mobile, aggiungere una unità alle variazioni relative.

t	<i>variaz.relativa%</i>	<i>NIBM</i>
1	—	
2	+11%	1.11
3	−11%	0.89
4	+12%	1.12
5	−12%	0.88

Esercizio 1.

Esercizio 2

Esercizio 3.

Esercizio 4.



Esercizio 1.

Esercizio 2

Esercizio 3.

Esercizio 4.



Si ricorda che ${}_3l_3 = 1$, che per valori di $t > 3$ vale

$${}_3l_t = {}_3l_4 \cdot {}_4l_5 \cdot \dots \cdot {}_{t-1}l_t$$

e che per valori di $t < 3$ vale

$${}_3l_t = \frac{1}{{}_tl_{t+1} \cdot \dots \cdot {}_1l_2 \cdot {}_2l_3}.$$

Esercizio 1.

Esercizio 2

Esercizio 3.

Esercizio 4.



Esercizio 1.

Esercizio 2

Esercizio 3.

Esercizio 4.

Si ricorda che ${}_3l_3 = 1$, che per valori di $t > 3$ vale

$${}_3l_t = {}_3l_4 \cdot {}_4l_5 \cdot \dots \cdot {}_{t-1}l_t$$

e che per valori di $t < 3$ vale

$${}_3l_t = \frac{1}{{}_t l_{t+1} \cdot \dots \cdot {}_1 l_2 \cdot {}_2 l_3}.$$

Si ottiene quindi

<i>giorno</i>	<i>NIBM</i>	<i>NIBF (3)</i>
1	1.00	1.0122
2	1.11	1.1236
3	0.89	1.0000
4	1.12	1.1200
5	0.88	0.9856



2. Si calcoli il numero indice al tempo 4 con base 2, interpretando il valore ottenuto.

Esercizio 1.

Esercizio 2

Esercizio 3.

Esercizio 4.



2. Si calcoli il numero indice al tempo 4 con base 2, interpretando il valore ottenuto.
Si ricorda che

$${}_2I_4 = {}_2I_3 \cdot {}_3I_4 = 0.9968.$$

Esercizio 1.

Esercizio 2

Esercizio 3.

Esercizio 4.



Esercizio 1.

Esercizio 2

Esercizio 3.

Esercizio 4.

2. Si calcoli il numero indice al tempo 4 con base 2, interpretando il valore ottenuto.
Si ricorda che

$${}_2I_4 = {}_2I_3 \cdot {}_3I_4 = 0.9968.$$

Tra il tempo 2 e il tempo 4 si è registrato un decremento complessivo del 0.32% della grandezza oggetto di studio.



3. Si calcoli la media aritmetica dei 4 numeri indici a base mobile ottenuti al punto 1 e si fornisca un intervallo di possibili valori per il tasso medio di variazione della quotazione del Bitcoin tra il tempo 1 e il tempo 5.

Esercizio 1.

Esercizio 2

Esercizio 3.

Esercizio 4.



3. Si calcoli la media aritmetica dei 4 numeri indici a base mobile ottenuti al punto 1 e si fornisca un intervallo di possibili valori per il tasso medio di variazione della quotazione del Bitcoin tra il tempo 1 e il tempo 5.

$$\frac{1}{4}(1.11 + 0.89 + 1.12 + 0.88) = 1.$$

Esercizio 1.

Esercizio 2

Esercizio 3.

Esercizio 4.



Esercizio 1.

Esercizio 2

Esercizio 3.

Esercizio 4.

3. Si calcoli la media aritmetica dei 4 numeri indici a base mobile ottenuti al punto 1 e si fornisca un intervallo di possibili valori per il tasso medio di variazione della quotazione del Bitcoin tra il tempo 1 e il tempo 5.

$$\frac{1}{4}(1.11 + 0.89 + 1.12 + 0.88) = 1.$$

In base al criterio di scelta della media secondo Chisini la media obiettivo dei numeri indici a base mobile ai fini del calcolo della tasso medio di variazione è la media geometrica. Abbiamo quindi, dal momento che la variabile statistica costituita dai numeri indici a base mobile non è degenerare:

$$I_{min} = 0.88 < \mu^{(0)} < \mu = 1$$

e

$$-12\% < \text{variazione relativa media \%} = \text{tasso medio di variazione} < 0\%$$



[Esercizio 1.](#)

[Esercizio 2](#)

[Esercizio 3.](#)

[Esercizio 4.](#)

Il seguente prospetto riporta i valori rilevati in corrispondenza a 5 unità statistiche con riferimento a due caratteri quantitativi X e Y

x_i	1	2	2	2	3
y_i	2	2	3	4	4

1. Si rappresentino graficamente le coppie dei dati e la funzione di regressione ipotizzando Y in funzione di X .



La funzione di regressione associa a ogni valore x appartenente all'insieme di definizione della variabile X la media di Y condizionata a x , $M(Y|x)$

x	$M(Y x)$
1	2
2	$\frac{1}{3}(2 + 3 + 4) = 3$
3	4

Esercizio 1.

Esercizio 2

Esercizio 3.

Esercizio 4.



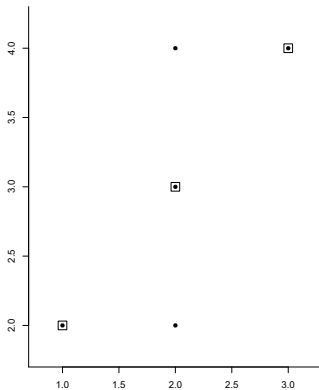
Esercizio 1.

Esercizio 2

Esercizio 3.

Esercizio 4.

Le coppie di dati possono essere rappresentate mediante un grafico a dispersione.





2. Si calcolino, secondo il metodo dei minimi quadrati, i parametri e l'indice di adattamento dei seguenti modelli di regressione:

$$I) Y^* = a + b \cdot X \qquad II) Y^* = c$$

[Esercizio 1.](#)

[Esercizio 2](#)

[Esercizio 3.](#)

[Esercizio 4.](#)



Si riportano i calcoli necessari¹ per ottenere le stime dei parametri e gli indici di adattamento del modello

$$Y = a + bX + E$$

x_i	y_i	$x_i y_i$	x_i^2	y_i^2
1	2	2	1	4
2	2	4	4	4
2	3	6	4	9
2	4	8	4	16
3	4	12	9	16
10	15	32	22	49
2	3	6.4	4.4	9.8
$M(X)$	$M(Y)$	$M(XY)$	$M(X^2)$	$M(Y^2)$

$$\hat{b} =$$

¹Si osserva che per ragioni di stampa alcuni calcoli potrebbero essere duplicati con riferimento ai diversi modelli oggetto di studio



Si riportano i calcoli necessari¹ per ottenere le stime dei parametri e gli indici di adattamento del modello

$$Y = a + bX + E$$

x_i	y_i	$x_i y_i$	x_i^2	y_i^2
1	2	2	1	4
2	2	4	4	4
2	3	6	4	9
2	4	8	4	16
3	4	12	9	16
10	15	32	22	49
2	3	6.4	4.4	9.8
$M(X)$	$M(Y)$	$M(XY)$	$M(X^2)$	$M(Y^2)$

$$\begin{aligned}\hat{b} &= \frac{\text{Cov}(X, Y)}{\text{Var}(X)} = \frac{M(XY) - M(X)M(Y)}{M(X^2) - [M(X)^2]} = \\ &= \frac{6.4 - 2 \cdot 3}{4.4 - 2^2} = \frac{0.4}{0.4} = 1 \\ \hat{a} &= \end{aligned}$$

¹Si osserva che per ragioni di stampa alcuni calcoli potrebbero essere duplicati con riferimento ai diversi modelli oggetto di studio



Si riportano i calcoli necessari¹ per ottenere le stime dei parametri e gli indici di adattamento del modello

$$Y = a + bX + E$$

x_i	y_i	$x_i y_i$	x_i^2	y_i^2
1	2	2	1	4
2	2	4	4	4
2	3	6	4	9
2	4	8	4	16
3	4	12	9	16
10	15	32	22	49
2	3	6.4	4.4	9.8
$M(X)$	$M(Y)$	$M(XY)$	$M(X^2)$	$M(Y^2)$

$$\hat{b} = \frac{\text{Cov}(X, Y)}{\text{Var}(X)} = \frac{M(XY) - M(X)M(Y)}{M(X^2) - [M(X)]^2} =$$

$$= \frac{6.4 - 2 \cdot 3}{4.4 - 2^2} = \frac{0.4}{0.4} = 1$$

$$\hat{a} = M(Y) - \hat{b}M(X) = 1$$

$$\rho^2 =$$

¹Si osserva che per ragioni di stampa alcuni calcoli potrebbero essere duplicati con riferimento ai diversi modelli oggetto di studio



Si riportano i calcoli necessari¹ per ottenere le stime dei parametri e gli indici di adattamento del modello

$$Y = a + bX + E$$

x_i	y_i	$x_i y_i$	x_i^2	y_i^2
1	2	2	1	4
2	2	4	4	4
2	3	6	4	9
2	4	8	4	16
3	4	12	9	16
10	15	32	22	49
2	3	6.4	4.4	9.8
$M(X)$	$M(Y)$	$M(XY)$	$M(X^2)$	$M(Y^2)$

$$\hat{b} = \frac{\text{Cov}(X, Y)}{\text{Var}(X)} = \frac{M(XY) - M(X)M(Y)}{M(X^2) - [M(X)]^2} =$$

$$= \frac{6.4 - 2 \cdot 3}{4.4 - 2^2} = \frac{0.4}{0.4} = 1$$

$$\hat{a} = M(Y) - \hat{b}M(X) = 1$$

$$\rho^2 = \frac{[\text{Cov}(X, Y)]^2}{\text{Var}(X) \text{Var}(Y)} = \frac{0.4^2}{0.4 \cdot 0.8} = 0.5$$

¹Si osserva che per ragioni di stampa alcuni calcoli potrebbero essere duplicati con riferimento ai diversi modelli oggetto di studio



Esercizio 1.

Esercizio 2

Esercizio 3.

Esercizio 4.

Si riportano i calcoli necessari² per ottenere le stime dei parametri e gli indici di adattamento del modello

$$Y = c + E$$

x_i	y_i	y_i^2
1	2	4
2	2	4
2	3	9
2	4	16
3	4	16
15		49
3		9.8
$M(Y)$		$M(Y^2)$

$$\hat{c} =$$

²Si osserva che per ragioni di stampa alcuni calcoli potrebbero essere duplicati con riferimento ai diversi modelli oggetto di studio



Esercizio 1.

Esercizio 2

Esercizio 3.

Esercizio 4.

Si riportano i calcoli necessari² per ottenere le stime dei parametri e gli indici di adattamento del modello

$$Y = c + E$$

x_i	y_i	y_i^2
1	2	4
2	2	4
2	3	9
2	4	16
3	4	16
15		49
3		9.8
$M(Y)$		$M(Y^2)$

$$\hat{c} = M(Y) = 3$$

$$R^2 =$$

²Si osserva che per ragioni di stampa alcuni calcoli potrebbero essere duplicati con riferimento ai diversi modelli oggetto di studio



Esercizio 1.

Esercizio 2

Esercizio 3.

Esercizio 4.

Si riportano i calcoli necessari² per ottenere le stime dei parametri e gli indici di adattamento del modello

$$Y = c + E$$

x_i	y_i	y_i^2
1	2	4
2	2	4
2	3	9
2	4	16
3	4	16
15		49
3		9.8
$M(Y)$		$M(Y^2)$

$$\hat{c} = M(Y) = 3$$

$$R^2 = 0$$

²Si osserva che per ragioni di stampa alcuni calcoli potrebbero essere duplicati con riferimento ai diversi modelli oggetto di studio



Esercizio 1.

Esercizio 2

Esercizio 3.

Esercizio 4.

$\hat{a} = 1$, $\hat{b} = 1$, indice adattamento $\rho^2 = 0.5$ $\hat{c} = 3$, indice adattamento $R^2 = 0$



3. Si individui il modello migliore.

Esercizio 1.

Esercizio 2

Esercizio 3.

Esercizio 4.



Esercizio 1.

Esercizio 2

Esercizio 3.

Esercizio 4.

3. Si individui il modello migliore.

Il modello migliore è $Y^* = 1 + 1 \cdot X$ in quanto caratterizzato da indice di adattamento più elevato.



4. Si indichino i valori assunti da $\eta^2_{Y|X}$ ed $\eta^2_{X|Y}$.

Esercizio 1.

Esercizio 2

Esercizio 3.

Esercizio 4.



Esercizio 1.

Esercizio 2

Esercizio 3.

Esercizio 4.

4. Si indichino i valori assunti da $\eta_{Y|X}^2$ ed $\eta_{X|Y}^2$.

$\eta_{Y|X}^2 = \eta_{X|Y}^2 = \rho^2 = 0.5$ in quanto sia le medie condizionate $M(Y|X)$ che le medie condizionate $M(X|Y)$ sono allineate.



Esercizio 1.

Esercizio 2

Esercizio 3.

Esercizio 4.

5. Sulla base del miglior modello si preveda il valore di Y in corrispondenza di $X = 4$.



Esercizio 1.

Esercizio 2

Esercizio 3.

Esercizio 4.

5. Sulla base del miglior modello si preveda il valore di Y in corrispondenza di $X = 4$.
 $Y^*(4) = 1 + 1 \cdot 4 = 5.$

Esercizio 4

L'urna U_1 contiene 71 palline bianche e 9 palline rosse; l'urna U_2 contiene 21 palline bianche e 29 palline rosse. Si consideri il seguente esperimento. Si lanciano 3 dadi regolari; se questi mostrano tutti la stessa faccia si effettuano 2 estrazioni senza reimmissione dall'urna U_1 altrimenti le 2 estrazioni vengono effettuate dall'urna U_2 .

1. Si calcoli la probabilità che le due palline estratte siano entrambe bianche.



Esercizio 1.

Esercizio 2

Esercizio 3.

Esercizio 4.

Esercizio 4

L'urna U_1 contiene 71 palline bianche e 9 palline rosse; l'urna U_2 contiene 21 palline bianche e 29 palline rosse. Si consideri il seguente esperimento. Si lanciano 3 dadi regolari; se questi mostrano tutti la stessa faccia si effettuano 2 estrazioni senza reimmissione dall'urna U_1 altrimenti le 2 estrazioni vengono effettuate dall'urna U_2 .

1. Si calcoli la probabilità che le due palline estratte siano entrambe bianche.

Si osserva, in primo luogo, che la probabilità che i 3 dadi regolari mostrino tutti la stessa faccia risulta pari a



Esercizio 1.

Esercizio 2

Esercizio 3.

Esercizio 4.

Esercizio 4



Esercizio 1.

Esercizio 2

Esercizio 3.

Esercizio 4.

L'urna U_1 contiene 71 palline bianche e 9 palline rosse; l'urna U_2 contiene 21 palline bianche e 29 palline rosse. Si consideri il seguente esperimento. Si lanciano 3 dadi regolari; se questi mostrano tutti la stessa faccia si effettuano 2 estrazioni senza reimmissione dall'urna U_1 altrimenti le 2 estrazioni vengono effettuate dall'urna U_2 .

1. Si calcoli la probabilità che le due palline estratte siano entrambe bianche.

Si osserva, in primo luogo, che la probabilità che i 3 dadi regolari mostrino tutti la stessa faccia risulta pari a

$$6 \cdot \frac{1}{6} \frac{1}{6} \frac{1}{6} = \frac{6}{216} = \frac{1}{36}.$$

Indicato con A l'evento 'i 3 dadi mostrano tutti la stessa faccia', abbiamo:

$$\begin{aligned} P(2B) &= P(A) \cdot P(2B|A) + P(\bar{A}) \cdot P(2B|\bar{A}) = \\ &= \frac{1}{36} \cdot \frac{71}{80} \frac{70}{79} + \frac{35}{36} \cdot \frac{21}{50} \frac{20}{49} = \\ &= 0.02778 \cdot 0.7864 + 0.9722 \cdot 0.1714 = \\ &= 0.02185 + 0.1666 = 0.1885. \end{aligned}$$



2. Si consideri un gioco per cui si paga una posta di 2Euro e si vincono 6Euro nel caso le 2 palline di cui all'esperimento sopra descritto siano entrambe bianche. Si stabilisca se il gioco è equo.

Esercizio 1.

Esercizio 2

Esercizio 3.

Esercizio 4.



2. Si consideri un gioco per cui si paga una posta di 2Euro e si vincono 6Euro nel caso le 2 palline di cui all'esperimento sopra descritto siano entrambe bianche. Si stabilisca se il gioco è equo.

La variabile casuale guadagno

<i>evento</i>	x_i	$P(X = x_i)$	$x_i \cdot P(X = x_i)$
<i>no 2B</i>	-2	0.8115	-1.623
<i>2B</i>	4	0.1885	0.754
		1.0000	-0.869

ha valore atteso negativo pari a -0.869 e, quindi, il gioco è favorevole al banco.

3. Si calcoli la probabilità che effettuando 200 estrazioni con reimmissione dall'urna U_2 si ottengano almeno 86 palline bianche.



Esercizio 1.

Esercizio 2

Esercizio 3.

Esercizio 4.



Esercizio 1.

Esercizio 2

Esercizio 3.

Esercizio 4.

3. Si calcoli la probabilità che effettuando 200 estrazioni con reimmissione dall'urna U_2 si ottengano almeno 86 palline bianche.

Indicata con X la variabile casuale 'numero palline bianche nelle 200 estrazioni con reimmissione dall'urna U_2 ' risulta $X \sim \text{Bin}(n = 200, p = 0.42)$. $M(X) = 84$;
 $\text{Var}(X) = 48.72$.



Esercizio 1.

Esercizio 2

Esercizio 3.

Esercizio 4.

3. Si calcoli la probabilità che effettuando 200 estrazioni con reimmissione dall'urna U_2 si ottengano almeno 86 palline bianche.

Indicata con X la variabile casuale 'numero palline bianche nelle 200 estrazioni con reimmissione dall'urna U_2 ' risulta $X \sim \text{Bin}(n = 200, p = 0.42)$. $M(X) = 84$; $\text{Var}(X) = 48.72$. Dal momento che $np = 84 > 5$ ed $n(1 - p) = 116 > 5$ è possibile utilizzare l'approssimazione della variabile casuale Binomiale con la variabile casuale Normale

$$X \sim$$



3. Si calcoli la probabilità che effettuando 200 estrazioni con reimmissione dall'urna U_2 si ottengano almeno 86 palline bianche.

Indicata con X la variabile casuale 'numero palline bianche nelle 200 estrazioni con reimmissione dall'urna U_2 ' risulta $X \sim \text{Bin}(n = 200, p = 0.42)$. $M(X) = 84$; $\text{Var}(X) = 48.72$. Dal momento che $np = 84 > 5$ ed $n(1 - p) = 116 > 5$ è possibile utilizzare l'approssimazione della variabile casuale Binomiale con la variabile casuale Normale

$$X \sim N(\mu = 84, \sigma^2 = 48.72).$$

Si ottiene

$$P(X > 86) =$$



3. Si calcoli la probabilità che effettuando 200 estrazioni con reimmissione dall'urna U_2 si ottengano almeno 86 palline bianche.

Indicata con X la variabile casuale 'numero palline bianche nelle 200 estrazioni con reimmissione dall'urna U_2 ' risulta $X \sim \text{Bin}(n = 200, p = 0.42)$. $M(X) = 84$; $\text{Var}(X) = 48.72$. Dal momento che $np = 84 > 5$ ed $n(1 - p) = 116 > 5$ è possibile utilizzare l'approssimazione della variabile casuale Binomiale con la variabile casuale Normale

$$X \sim N(\mu = 84, \sigma^2 = 48.72).$$

Si ottiene

$$P(X > 86) = P\left(\frac{X - \mu}{\sigma} > \frac{86 - \mu}{\sigma}\right) =$$

$$=$$



3. Si calcoli la probabilità che effettuando 200 estrazioni con reimmissione dall'urna U_2 si ottengano almeno 86 palline bianche.

Indicata con X la variabile casuale 'numero palline bianche nelle 200 estrazioni con reimmissione dall'urna U_2 ' risulta $X \sim \text{Bin}(n = 200, p = 0.42)$. $M(X) = 84$; $\text{Var}(X) = 48.72$. Dal momento che $np = 84 > 5$ ed $n(1 - p) = 116 > 5$ è possibile utilizzare l'approssimazione della variabile casuale Binomiale con la variabile casuale Normale

$$X \sim N(\mu = 84, \sigma^2 = 48.72).$$

Si ottiene

$$\begin{aligned} P(X > 86) &= P\left(\frac{X - \mu}{\sigma} > \frac{86 - \mu}{\sigma}\right) = \\ &= P\left(Z > \frac{86 - 84}{6.98}\right) = \\ &= \end{aligned}$$



3. Si calcoli la probabilità che effettuando 200 estrazioni con reimmissione dall'urna U_2 si ottengano almeno 86 palline bianche.

Indicata con X la variabile casuale 'numero palline bianche nelle 200 estrazioni con reimmissione dall'urna U_2 ' risulta $X \sim \text{Bin}(n = 200, p = 0.42)$. $M(X) = 84$; $\text{Var}(X) = 48.72$. Dal momento che $np = 84 > 5$ ed $n(1 - p) = 116 > 5$ è possibile utilizzare l'approssimazione della variabile casuale Binomiale con la variabile casuale Normale

$$X \sim N(\mu = 84, \sigma^2 = 48.72).$$

Si ottiene

$$\begin{aligned} P(X > 86) &= P\left(\frac{X - \mu}{\sigma} > \frac{86 - \mu}{\sigma}\right) = \\ &= P\left(Z > \frac{86 - 84}{6.98}\right) = \\ &= P(Z > 0.29) = \\ &= 1 - P(Z \leq 0.29) = \\ &= 1 - 0.6141 = 0.3859. \end{aligned}$$

Tema T3041

Prova scritta di STATISTICA del 03.06.2019

Prova scritta di
STATISTICA del
03.06.2019

Giuseppe Boari
Gabriele Cantaluppi



UNIVERSITÀ
CATTOLICA
del Sacro Cuore

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.

Giuseppe Boari
Gabriele Cantaluppi
Università Cattolica del Sacro Cuore, Milano

Esercizio 1

Prova scritta di
STATISTICA del
03.06.2019

Giuseppe Boari
Gabriele Cantaluppi



UNIVERSITÀ
CATTOLICA
del Sacro Cuore

Il prospetto seguente riporta la distribuzione congiunta del carattere qualitativo X e del carattere quantitativo Y .

$X \setminus Y$	1	2	3
x_1	1	1	0
x_2	0	1	1

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.



1. Si rappresenti graficamente la distribuzione di frequenza della condizionata $Y|X_1$ e si confronti il livello di eterogeneità della distribuzione $X|y_2$ con quello della marginale X .

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.



Si considera la distribuzione condizionata $Y|x_1$

y_j	n_{1j}
1	1
2	1
3	0

Esercizio 1.

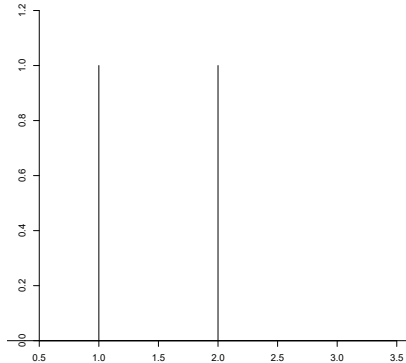
Esercizio 2.

Esercizio 3.

Esercizio 4.



Si tratta di un carattere quantitativo con modalità non raggruppate in classi che può essere rappresentato graficamente mediante un grafico a bastoncini.



Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.



Si considera la distribuzione condizionata $X|y_2$

x_i	n_{i2}
x_1	1
x_2	1

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.



La mutabile X presenta massima eterogeneità, quindi gli indici di Gini e di Frosini assumono valore massimo, $F_N = G_N = 1$.

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.



Si considera la distribuzione marginale della variabile X , ottenuta sommando le frequenze di riga

x_j	$n_{j\bullet}$
x_1	2
x_2	2

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.



La mutabile X presenta massima eterogeneità, quindi gli indici di Gini e di Frosini assumono valore massimo, $F_N = G_N = 1$.

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.



eterogeneità $X|y_2$: $G_N(X) = F_N(X) = 1$

eterogeneità X : $G_N(X) = F_N(X) = 1$

Entrambe le distribuzioni sono caratterizzate da massima eterogeneità.

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.



2. Si confrontino la distribuzione condizionata $Y|x_1$ e la marginale Y mediante opportuni indici di posizione e di variabilità.

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.



Si tratta di un carattere quantitativo con modalità non raggruppate in classi.

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.



La media aritmetica

$$M(Y|x_1) = \frac{1}{n_{1\bullet}} \sum_{j=1}^k y_j n_{1j}$$

risulta 1.5.

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.



La media aritmetica

$$M(Y|x_1) = \frac{1}{n_{1\bullet}} \sum_{j=1}^k y_j n_{1j}$$

risulta 1.5.
Si ottiene

y_j	n_{1j}	$y_j n_{1j}$
1	1	1
2	1	2
3	0	0
2		3

da cui segue

$$M(Y|x_1) = \frac{1}{n_{1\bullet}} \sum_{j=1}^k y_j n_{1j} = \frac{3}{2} = 1.5.$$

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.



La varianza

$$\text{Var}(Y|x_1) = M(Y^2|x_1) - [M(Y|x_1)]^2$$

risulta 0.25.

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.



La varianza

$$\text{Var}(Y|x_1) = M(Y^2|x_1) - [M(Y|x_1)]^2$$

risulta 0.25. Si ottiene

y_j	n_{1j}	$y_j n_{1j}$	$y_j^2 n_{1j}$
1	1	1	1
2	1	2	4
3	0	0	0
2		3	5

da cui segue

$$\text{Var}(Y|x_1) = \frac{5}{2} - 1.5^2 = 2.5 - 2.25 = 0.25.$$

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.



Per effettuare confronti in termini di variabilità tra caratteri quantitativi occorre utilizzare una misura relativa, ad esempio il coefficiente di variazione

$$CV = \frac{\sigma}{\mu} = \frac{\sqrt{\text{Var}(Y|x_1)}}{M(Y|x_1)} = \frac{\sqrt{0.25}}{1.5} = \frac{0.5}{1.5} = 0.3333.$$

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.



Si considera la distribuzione marginale della variabile Y , ottenuta sommando le frequenze di colonna

y_j	$n_{\bullet j}$
1	1
2	2
3	1

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.



Si tratta di un carattere quantitativo con modalità non raggruppate in classi.

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.



La media aritmetica

$$M(Y) = \frac{1}{n} \sum_{j=1}^k y_j n_{\bullet j}$$

risulta 2.

Si osserva, infatti, che la distribuzione è simmetrica.

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.



La varianza

$$\text{Var}(Y) = M[(Y - \mu_Y)^2] = M(Y^2) - \mu_Y^2 = \frac{1}{n} \sum_{j=1}^k y_j^2 n_{\bullet j} - \mu_Y^2$$

risulta 0.5.

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.



Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.

La varianza

$$\text{Var}(Y) = M[(Y - \mu_Y)^2] = M(Y^2) - \mu_Y^2 = \frac{1}{n} \sum_{j=1}^k y_j^2 n_{\bullet j} - \mu_Y^2$$

risulta 0.5.

Si ottiene

y_j	$n_{\bullet j}$	$y_j n_{\bullet j}$	$y_j^2 n_{\bullet j}$
1	1	1	1
2	2	4	8
3	1	3	9
	4	8	18

da cui segue

$$\begin{aligned} \text{Var}(Y) = \sigma_Y^2 &= M(Y^2) - \mu_Y^2 = \\ &= \frac{18}{4} - 2^2 = 4.5 - 4 = 0.5. \end{aligned}$$



Per effettuare confronti in termini di variabilità tra caratteri quantitativi occorre utilizzare una misura relativa, ad esempio il coefficiente di variazione

$$CV = \frac{\sigma}{\mu} = \frac{\sqrt{\text{Var}(Y)}}{M(Y)} = \frac{\sqrt{0.5}}{2} = \frac{0.7071}{2} = 0.3535.$$

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.



$Y|x_1$ Indice di posizione: $M(Y|x_1)$ valore 1.5 Indice di variabilità: $CV(Y|x_1)$ valore 0.3333
 Y Indice di posizione: $M(Y)$ valore 2 Indice di variabilità: $CV(Y)$ valore 0.3536
La distribuzione condizionata $Y|x_1$ è caratterizzata da un livello medio e da variabilità inferiori rispetto alla marginale Y .

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.



3. Si determini il grado di connessione tra le variabili X e Y mediante il calcolo di un opportuno indice.

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.



Dalla distribuzione delle frequenze congiunte n_{ij} :

	1	2	3	$n_{i\bullet}$
x_1	1	1	0	2
x_2	0	1	1	2
$n_{\bullet j}$	1	2	1	4

è possibile ottenere le frequenze teoriche nella situazione di indipendenza stocastica

$$\hat{n}_{ij} = \frac{n_{i\bullet} n_{\bullet j}}{n}:$$

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.



Dalla distribuzione delle frequenze congiunte n_{ij} :

	1	2	3	$n_{i\bullet}$
x1	1	1	0	2
x2	0	1	1	2
$n_{\bullet j}$	1	2	1	4

è possibile ottenere le frequenze teoriche nella situazione di indipendenza stocastica

$$\hat{n}_{ij} = \frac{n_{i\bullet} n_{\bullet j}}{n}:$$

	1	2	3	$n_{i\bullet}$
x1	0.5	1	0.5	2
x2	0.5	1	0.5	2
$n_{\bullet j}$	1.0	2	1.0	4

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.



le contingenze assolute $n_{ij} - \hat{n}_{ij}$:

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.



le contingenze assolute $n_{ij} - \hat{n}_{ij}$:

	1	2	3
x1	0.5	0	-0.5
x2	-0.5	0	0.5

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.



i rapporti di contingenza $\frac{n_{ij}}{\hat{n}_{ij}}$:

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.



i rapporti di contingenza $\frac{n_{ij}}{\hat{n}_{ij}}$:

	1	2	3
x1	2	1	0
x2	0	1	2

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.



e le contingenze relative, espresse in termini percentuali, $100 \left(\frac{n_{ij} - \hat{n}_{ij}}{\hat{n}_{ij}} \right) \% = 100 \left(\frac{n_{ij}}{\hat{n}_{ij}} - 1 \right) \%$:

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.



e le contingenze relative, espresse in termini percentuali, $100 \left(\frac{n_{ij} - \hat{n}_{ij}}{\hat{n}_{ij}} \right) \% = 100 \left(\frac{n_{ij}}{\hat{n}_{ij}} - 1 \right) \%$:

	1	2	3
x1	100	0	-100
x2	-100	0	100

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.



Risultati intermedi $\frac{(n_{ij} - \hat{n}_{ij})^2}{\hat{n}_{ij}}$ per il calcolo del χ^2

	1	2	3
x1	0.5	0	0.5
x2	0.5	0	0.5

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.



Risultati intermedi $\frac{n_{ij}^2}{n_{i\bullet} \cdot n_{\bullet j}}$ per il calcolo del χ^2 (formula operativa)

	1	2	3
x1	0.5	0.25	0.0
x2	0.0	0.25	0.5

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.



Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.

$$\chi^2 = \sum_{i=1}^h \sum_{j=1}^k \frac{(n_{ij} - \hat{n}_{ij})^2}{\hat{n}_{ij}} = 2,$$

$$\chi_N^2 = \frac{\chi^2}{n \cdot \min(h-1, k-1)} = 0.5$$

$$\chi_N = \sqrt{\chi_N^2} = 0.7071.$$

Sussiste, quindi, un livello di connessione medio alto.



La seguente tabella riporta la serie storica del fatturato annuo di un'azienda:

t	x_t
2012	50
2013	55
2014	52.25
2015	62.7
2016	62.7

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.



1. Si costruisca la serie dei numeri indici a base mobile e si calcoli, commentando il risultato, il valore di $_{2014}I_{2016}$, numero indice al tempo 2016 con base 2014.

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.



1. Si costruisca la serie dei numeri indici a base mobile e si calcoli, commentando il risultato, il valore di ${}_{2014}I_{2016}$, numero indice al tempo 2016 con base 2014.

Si ricorda che ${}_{t-1}I_t = \frac{x_t}{x_{t-1}}$.

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.



1. Si costruisca la serie dei numeri indici a base mobile e si calcoli, commentando il risultato, il valore di ${}_{2014}I_{2016}$, numero indice al tempo 2016 con base 2014.

Si ricorda che ${}_{t-1}I_t = \frac{x_t}{x_{t-1}}$. Si ottiene quindi

Anno	x_t	NIBM
2012	50.00	
2013	55.00	1.10
2014	52.25	0.95
2015	62.70	1.20
2016	62.70	1.00

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.



1. Si costruisca la serie dei numeri indici a base mobile e si calcoli, commentando il risultato, il valore di ${}_{2014}I_{2016}$, numero indice al tempo 2016 con base 2014.

Si ricorda che ${}_{t-1}I_t = \frac{x_t}{x_{t-1}}$. Si ottiene quindi

Anno	x_t	NIBM
2012	50.00	
2013	55.00	1.10
2014	52.25	0.95
2015	62.70	1.20
2016	62.70	1.00

$${}_{2014}I_{2016} = \frac{x_{2016}}{x_{2014}} = \frac{62.7}{52.25} = 1.2$$

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.



Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.

1. Si costruisca la serie dei numeri indici a base mobile e si calcoli, commentando il risultato, il valore di ${}_{2014}I_{2016}$, numero indice al tempo 2016 con base 2014.

Si ricorda che ${}_{t-1}I_t = \frac{x_t}{x_{t-1}}$. Si ottiene quindi

Anno	x_t	NIBM
2012	50.00	
2013	55.00	1.10
2014	52.25	0.95
2015	62.70	1.20
2016	62.70	1.00

$${}_{2014}I_{2016} = \frac{x_{2016}}{x_{2014}} = \frac{62.7}{52.25} = 1.2$$

Tra il tempo 2014 e il tempo 2016 si è registrato un incremento complessivo del 20% della grandezza oggetto di studio.



2. Si calcoli il tasso medio di variazione dal 2012 al 2014 e si commenti il risultato ottenuto.

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.



2. Si calcoli il tasso medio di variazione dal 2012 al 2014 e si commenti il risultato ottenuto.
Si ricorda che

$${}_{2012}I_{2014} = {}_{2012}I_{2013} \cdot {}_{2013}I_{2014}$$

Si deve trovare il numero indice a base mobile medio.

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.



Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.

2. Si calcoli il tasso medio di variazione dal 2012 al 2014 e si commenti il risultato ottenuto. Si ricorda che

$${}_{2012}l_{2014} = {}_{2012}l_{2013} \cdot {}_{2013}l_{2014}$$

Si deve trovare il numero indice a base mobile medio. Applicando il criterio di scelta della media secondo Chisini si ottiene

$${}_{2012}l_{2014} = {}_{2012}l_{2013} \cdot {}_{2013}l_{2014} = \alpha \cdot \alpha = \alpha^2.$$



Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.

2. Si calcoli il tasso medio di variazione dal 2012 al 2014 e si commenti il risultato ottenuto. Si ricorda che

$${}_{2012}l_{2014} = {}_{2012}l_{2013} \cdot {}_{2013}l_{2014}$$

Si deve trovare il numero indice a base mobile medio. Applicando il criterio di scelta della media secondo Chisini si ottiene

$${}_{2012}l_{2014} = {}_{2012}l_{2013} \cdot {}_{2013}l_{2014} = \alpha \cdot \alpha = \alpha^2.$$

Quindi

$$\alpha = \sqrt[2]{{}_{2012}l_{2014}} = \sqrt[2]{{}_{2012}l_{2013} \cdot {}_{2013}l_{2014}} = 1.0223.$$

Da cui segue una variazione relativa media pari a 0.0223, corrispondente a un incremento medio annuo tra il 2012 e il 2014 del 2.23%.



Il seguente prospetto riporta i valori rilevati in corrispondenza a 5 unità statistiche con riferimento a due caratteri quantitativi X e Y

x_i	1	2	3	4	5
y_i	2	4	4	4	6

1. Si rappresentino graficamente le coppie dei dati ipotizzando Y in funzione di X .

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.



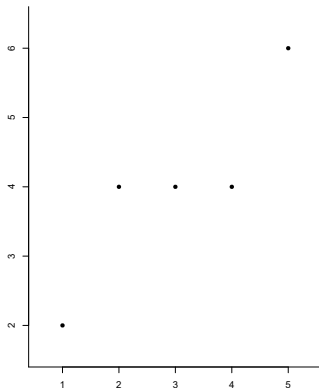
Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.

Le coppie di dati possono essere rappresentate mediante un grafico a dispersione.





2. Si calcolino, secondo il metodo dei minimi quadrati, i parametri dei seguenti modelli di regressione:

$$I) Y^* = a + b \cdot X \quad II) Y^* = c \cdot X$$

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.

Si riportano i calcoli necessari¹ per ottenere le stime dei parametri e gli indici di adattamento del modello

$$Y = a + bX + E$$

x_i	y_i	$x_i y_i$	x_i^2	y_i^2
1	2	2	1	4
2	4	8	4	16
3	4	12	9	16
4	4	16	16	16
5	6	30	25	36
15	20	68	55	88
3	4	13.6	11	17.6
$M(X)$	$M(Y)$	$M(XY)$	$M(X^2)$	$M(Y^2)$

$$\hat{b} =$$



Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.

¹Si osserva che per ragioni di stampa alcuni calcoli potrebbero essere duplicati con riferimento ai diversi modelli oggetto di studio

Si riportano i calcoli necessari¹ per ottenere le stime dei parametri e gli indici di adattamento del modello

$$Y = a + bX + E$$

x_i	y_i	$x_i y_i$	x_i^2	y_i^2
1	2	2	1	4
2	4	8	4	16
3	4	12	9	16
4	4	16	16	16
5	6	30	25	36
15	20	68	55	88
3	4	13.6	11	17.6
$M(X)$	$M(Y)$	$M(XY)$	$M(X^2)$	$M(Y^2)$

$$\begin{aligned}\hat{b} &= \frac{\text{Cov}(X, Y)}{\text{Var}(X)} = \frac{M(XY) - M(X)M(Y)}{M(X^2) - [M(X)]^2} = \\ &= \frac{13.6 - 3 \cdot 4}{11 - 3^2} = \frac{1.6}{2} = 0.8 \\ \hat{a} &= \end{aligned}$$



Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.

¹Si osserva che per ragioni di stampa alcuni calcoli potrebbero essere duplicati con riferimento ai diversi modelli oggetto di studio

Si riportano i calcoli necessari¹ per ottenere le stime dei parametri e gli indici di adattamento del modello

$$Y = a + bX + E$$

x_i	y_i	$x_i y_i$	x_i^2	y_i^2
1	2	2	1	4
2	4	8	4	16
3	4	12	9	16
4	4	16	16	16
5	6	30	25	36
15	20	68	55	88
3	4	13.6	11	17.6
$M(X)$	$M(Y)$	$M(XY)$	$M(X^2)$	$M(Y^2)$

$$\hat{b} = \frac{\text{Cov}(X, Y)}{\text{Var}(X)} = \frac{M(XY) - M(X)M(Y)}{M(X^2) - [M(X)]^2} =$$

$$= \frac{13.6 - 3 \cdot 4}{11 - 3^2} = \frac{1.6}{2} = 0.8$$

$$\hat{a} = M(Y) - \hat{b}M(X) = 1.6$$

$$\rho^2 =$$



Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.

¹Si osserva che per ragioni di stampa alcuni calcoli potrebbero essere duplicati con riferimento ai diversi modelli oggetto di studio

Si riportano i calcoli necessari¹ per ottenere le stime dei parametri e gli indici di adattamento del modello

$$Y = a + bX + E$$

x_i	y_i	$x_i y_i$	x_i^2	y_i^2
1	2	2	1	4
2	4	8	4	16
3	4	12	9	16
4	4	16	16	16
5	6	30	25	36
15	20	68	55	88
3	4	13.6	11	17.6
$M(X)$	$M(Y)$	$M(XY)$	$M(X^2)$	$M(Y^2)$

$$\hat{b} = \frac{\text{Cov}(X, Y)}{\text{Var}(X)} = \frac{M(XY) - M(X)M(Y)}{M(X^2) - [M(X)]^2} =$$

$$= \frac{13.6 - 3 \cdot 4}{11 - 3^2} = \frac{1.6}{2} = 0.8$$

$$\hat{a} = M(Y) - \hat{b}M(X) = 1.6$$

$$\rho^2 = \frac{[\text{Cov}(X, Y)]^2}{\text{Var}(X) \text{Var}(Y)} = \frac{1.6^2}{2 \cdot 1.6} = 0.8$$

$$M(\hat{E}^2) = \text{Var}(\hat{E}) = \text{varianza residua} =$$

¹Si osserva che per ragioni di stampa alcuni calcoli potrebbero essere duplicati con riferimento ai diversi modelli oggetto di studio



Si riportano i calcoli necessari¹ per ottenere le stime dei parametri e gli indici di adattamento del modello

$$Y = a + bX + E$$

x_i	y_i	$x_i y_i$	x_i^2	y_i^2
1	2	2	1	4
2	4	8	4	16
3	4	12	9	16
4	4	16	16	16
5	6	30	25	36
15	20	68	55	88
3	4	13.6	11	17.6
$M(X)$	$M(Y)$	$M(XY)$	$M(X^2)$	$M(Y^2)$

$$\begin{aligned}\hat{b} &= \frac{\text{Cov}(X, Y)}{\text{Var}(X)} = \frac{M(XY) - M(X)M(Y)}{M(X^2) - [M(X)]^2} = \\ &= \frac{13.6 - 3 \cdot 4}{11 - 3^2} = \frac{1.6}{2} = 0.8 \\ \hat{a} &= M(Y) - \hat{b}M(X) = 1.6 \\ \rho^2 &= \frac{[\text{Cov}(X, Y)]^2}{\text{Var}(X) \text{Var}(Y)} = \frac{1.6^2}{2 \cdot 1.6} = 0.8\end{aligned}$$

$$M(\hat{E}^2) = \text{Var}(\hat{E}) = \text{varianza residua} = \text{Var}(Y) \cdot (1 - \rho^2) = 0.32$$

¹Si osserva che per ragioni di stampa alcuni calcoli potrebbero essere duplicati con riferimento ai diversi modelli oggetto di studio





Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.

Si riportano i calcoli necessari² per ottenere le stime dei parametri e la misura di adattamento del modello

$$Y = cX + E$$

x_i	y_i	$x_i y_i$	x_i^2	y_i^2
1	2	2	1	4
2	4	8	4	16
3	4	12	9	16
4	4	16	16	16
5	6	30	25	36
		68	55	88
		13.6	11	17.6
		$M(XY)$	$M(X^2)$	$M(Y^2)$

$$\hat{c} =$$

²Si osserva che per ragioni di stampa alcuni calcoli potrebbero essere duplicati con riferimento ai diversi modelli oggetto di studio



Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.

Si riportano i calcoli necessari² per ottenere le stime dei parametri e la misura di adattamento del modello

$$Y = cX + E$$

x_i	y_i	$x_i y_i$	x_i^2	y_i^2
1	2	2	1	4
2	4	8	4	16
3	4	12	9	16
4	4	16	16	16
5	6	30	25	36
		68	55	88
		13.6	11	17.6
		$M(XY)$	$M(X^2)$	$M(Y^2)$

$$\hat{c} = \frac{M(XY)}{M(X^2)} = \frac{13.6}{11} = 1.2364$$

$$M(\hat{E}^2) =$$

²Si osserva che per ragioni di stampa alcuni calcoli potrebbero essere duplicati con riferimento ai diversi modelli oggetto di studio



Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.

Si riportano i calcoli necessari² per ottenere le stime dei parametri e la misura di adattamento del modello

$$Y = cX + E$$

x_i	y_i	$x_i y_i$	x_i^2	y_i^2
1	2	2	1	4
2	4	8	4	16
3	4	12	9	16
4	4	16	16	16
5	6	30	25	36
		68	55	88
		13.6	11	17.6
		$M(XY)$	$M(X^2)$	$M(Y^2)$

$$\hat{c} = \frac{M(XY)}{M(X^2)} = \frac{13.6}{11} = 1.2364$$

$$M(\hat{E}^2) = M[Y \cdot (Y - \hat{c}X)] = M(Y^2) - \hat{c}M(XY) = 17.6 - 1.2364 \cdot 13.6 = 0.7855$$

²Si osserva che per ragioni di stampa alcuni calcoli potrebbero essere duplicati con riferimento ai diversi modelli oggetto di studio



Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.

$\hat{a} = 1.6$, $\hat{b} = 0.8$, indice adattamento $M(\hat{E}^2) = 0.32$
 $M(\hat{E}^2) = 0.7855$

$\hat{c} = 1.2364$, indice adattamento



3. Confrontare gli indici di adattamento, commentando i risultati.

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.



Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.

3. Confrontare gli indici di adattamento, commentando i risultati.
Il modello migliore è $Y^* = 1.6 + 0.8X$.



4. Senza svolgere i conti, cosa è possibile dire con riferimento ai dati in esame riguardo alla relazione tra ρ^2 ed $\eta_{Y|X}^2$.

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.



Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.

4. Senza svolgere i conti, cosa è possibile dire con riferimento ai dati in esame riguardo alla relazione tra ρ^2 ed $\eta_{Y|X}^2$.

Vale la relazione $\rho^2 = 0.8 < \eta_{Y|X}^2 = 1$, in quanto sussiste dipendenza funzionale di Y da X .



5. Senza svolgere i conti, cosa è possibile dire con riferimento ai dati in esame riguardo alla relazione tra ρ^2 ed $\eta_{X|Y}^2$.

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.



Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.

5. Senza svolgere i conti, cosa è possibile dire con riferimento ai dati in esame riguardo alla relazione tra ρ^2 ed $\eta_{X|Y}^2$.

Vale la relazione $\rho^2 = \eta_{X|Y}^2 = 0.8$, in quanto le medie condizionate $M(X|Y)$ sono allineate.

Esercizio 4

Prova scritta di
STATISTICA del
03.06.2019

Giuseppe Boari
Gabriele Cantaluppi



UNIVERSITÀ
CATTOLICA
del Sacro Cuore

L'urna $U1$ contiene 3 palline nere e 1 pallina rossa. L'urna $U2$ contiene 18 palline nere e 1 pallina rossa.

1. Si calcoli la probabilità che estraendo con reimmissione 800 palline dall'urna $U1$ almeno 200 siano rosse.

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.

Esercizio 4

Prova scritta di
STATISTICA del
03.06.2019

Giuseppe Boari
Gabriele Cantaluppi



UNIVERSITÀ
CATTOLICA
del Sacro Cuore

L'urna $U1$ contiene 3 palline nere e 1 pallina rossa. L'urna $U2$ contiene 18 palline nere e 1 pallina rossa.

1. Si calcoli la probabilità che estraendo con reimmissione 800 palline dall'urna $U1$ almeno 200 siano rosse.

Indicata con X la variabile casuale 'numero palline rosse nelle 800 estrazioni con reimmissione dall'urna U_1 ' risulta $X \sim$

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.



Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.

L'urna U_1 contiene 3 palline nere e 1 pallina rossa. L'urna U_2 contiene 18 palline nere e 1 pallina rossa.

1. Si calcoli la probabilità che estraendo con reimmissione 800 palline dall'urna U_1 almeno 200 siano rosse.

Indicata con X la variabile casuale 'numero palline rosse nelle 800 estrazioni con reimmissione dall'urna U_1 ' risulta $X \sim \text{Bin}(n = 800, p = 0.25)$. $M(X) = 200$; $\text{Var}(X) = 150$. Dal momento che $np = 200 > 5$ ed $n(1 - p) = 600 > 5$ è possibile utilizzare l'approssimazione della variabile casuale Binomiale con la variabile casuale Normale

$$X \sim$$



L'urna U_1 contiene 3 palline nere e 1 pallina rossa. L'urna U_2 contiene 18 palline nere e 1 pallina rossa.

1. Si calcoli la probabilità che estraendo con reimmissione 800 palline dall'urna U_1 almeno 200 siano rosse.

Indicata con X la variabile casuale 'numero palline rosse nelle 800 estrazioni con reimmissione dall'urna U_1 ' risulta $X \sim \text{Bin}(n = 800, p = 0.25)$. $M(X) = 200$; $\text{Var}(X) = 150$. Dal momento che $np = 200 > 5$ ed $n(1 - p) = 600 > 5$ è possibile utilizzare l'approssimazione della variabile casuale Binomiale con la variabile casuale Normale

$$X \sim N(\mu = 200, \sigma^2 = 150).$$

Dal momento che la variabile casuale Normale è simmetrica rispetto alla media, si ottiene

$$P(X > 200) = 0.5.$$



2. Si calcoli la probabilità che estraendo con reimmissione 40 palline dall'urna U_2 almeno una sia rossa.

Sia X la variabile casuale 'numero palline rosse nelle 40 estrazioni con reimmissione dall'urna U_2 '.

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.



2. Si calcoli la probabilità che estraendo con reimmissione 40 palline dall'urna U_2 almeno una sia rossa.

Sia X la variabile casuale 'numero palline rosse nelle 40 estrazioni con reimmissione dall'urna U_2 '.

Abbiamo $X \sim$

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.



Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.

2. Si calcoli la probabilità che estraendo con reimmissione 40 palline dall'urna U_2 almeno una sia rossa.

Sia X la variabile casuale 'numero palline rosse nelle 40 estrazioni con reimmissione dall'urna U_2 '.

Abbiamo $X \sim \text{Bin}(n = 40, p = 0.0526)$.

$$\begin{aligned} P(X \geq 1) &= 1 - P(X = 0) = \\ &= 1 - \binom{40}{0} 0.0526^0 \cdot 0.9474^{40} = \\ &= 1 - 0.1150 = 0.8850. \end{aligned}$$



3. Si estragga una pallina dall'urna $U1$ e la si metta nell'urna $U2$. Si calcoli la probabilità che estraendo con reimmissione 2 palline dall'urna $U2$ (modificata) si ottengano 2 palline nere.

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.



Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.

3. Si estragga una pallina dall'urna $U1$ e la si metta nell'urna $U2$. Si calcoli la probabilità che estraendo con reimmissione 2 palline dall'urna $U2$ (modificata) si ottengano 2 palline nere.

Siano N l'evento 'estrazione di una pallina nera dall'urna $U1$ ', $X|N$ e $X|\bar{N}$ le variabile casuale 'n. di palline nere nelle 2 estrazioni con reimmissione dall'urna $U2$ ' condizionate al verificarsi degli eventi N e $\bar{N}$.

$$P(2N) =$$



Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.

3. Si estragga una pallina dall'urna $U1$ e la si metta nell'urna $U2$. Si calcoli la probabilità che estraendo con reimmissione 2 palline dall'urna $U2$ (modificata) si ottengano 2 palline nere.

Siano N l'evento 'estrazione di una pallina nera dall'urna $U1$ ', $X|N$ e $X|\bar{N}$ le variabile casuale 'n. di palline nere nelle 2 estrazioni con reimmissione dall'urna $U2$ ' condizionate al verificarsi degli eventi N e $\bar{N}$.

$$P(2N) = P(N) \cdot P\{(X|N) = 2\} + P(\bar{N}) \cdot P\{(X|\bar{N}) = 2\} =$$

=



Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.

3. Si estragga una pallina dall'urna $U1$ e la si metta nell'urna $U2$. Si calcoli la probabilità che estraendo con reimmissione 2 palline dall'urna $U2$ (modificata) si ottengano 2 palline nere.

Siano N l'evento 'estrazione di una pallina nera dall'urna $U1$ ', $X|N$ e $X|\bar{N}$ le variabile casuale 'n. di palline nere nelle 2 estrazioni con reimmissione dall'urna $U2$ ' condizionate al verificarsi degli eventi N e $\bar{N}$.

$$\begin{aligned} P(2N) &= P(N) \cdot P\{(X|N) = 2\} + P(\bar{N}) \cdot P\{(X|\bar{N}) = 2\} = \\ &= \frac{3}{4} \cdot \left(\frac{19}{20}\right)^2 + \frac{1}{4} \cdot \left(\frac{18}{20}\right)^2 = 0.75 \cdot 0.9025 + 0.25 \cdot 0.81 = 0.8794. \end{aligned}$$

Tema T304

Prova scritta di STATISTICA del 17.06.2019

Prova scritta di
STATISTICA del
17.06.2019

Giuseppe Boari
Gabriele Cantaluppi



UNIVERSITÀ
CATTOLICA
del Sacro Cuore

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.

Giuseppe Boari
Gabriele Cantaluppi
Università Cattolica del Sacro Cuore, Milano



Per valutare la dipendenza fra costo orario (in euro) di una prestazione professionale (Y) e giorno di erogazione (X) è stata svolta una indagine di mercato su un campione di 80 committenze. Nella tabella sottostante sono riportate le frequenze congiunte di (X, Y) .

$X \setminus Y$	0 - 10	10 - 30
<i>feriale</i>	40	0
<i>festivo</i>	20	20

1. Fornire una opportuna rappresentazione grafica delle distribuzioni marginali di X e Y .

[Esercizio 1.](#)

[Esercizio 2.](#)

[Esercizio 3.](#)

[Esercizio 4.](#)



Si considera la distribuzione marginale della variabile X , ottenuta sommando le frequenze di riga

x_i	$n_{i\bullet}$
<i>feriale</i>	40
<i>festivo</i>	40

Si tratta di un carattere qualitativo sconnesso che può essere rappresentato graficamente mediante un grafico a bastoncini o a canne d'organo.

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.

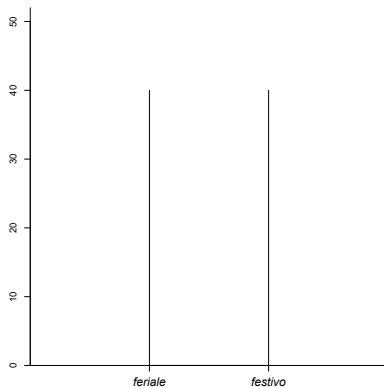


Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.





Si considera la distribuzione marginale della variabile Y , ottenuta sommando le frequenze di colonna

y_j	$n_{\bullet j}$
0 $\rightarrow$ 10	60
10 $\rightarrow$ 30	20

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.



Si tratta di un carattere quantitativo con modalità raggruppate in classi che può essere rappresentato graficamente mediante un istogramma.

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.



L'istogramma può essere costruito dopo avere calcolato le densità $d_{\bullet j} = \frac{n_{\bullet j}}{a_j}$, (dove a_j sono le ampiezze di classe).

y_j	$n_{\bullet j}$	$a_{\bullet j}$	$d_{\bullet j}$
0 - 10	60	10 - 0 = 10	6
10 - 30	20	30 - 10 = 20	1

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.

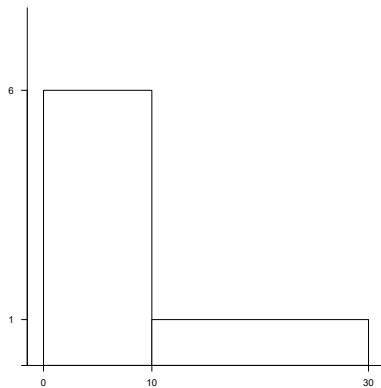


Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.





2. Calcolare il valore della varianza between e della varianza within per il carattere Y rispetto ai gruppi definiti dal carattere X . Commentare il risultato ottenuto.

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.



Si considera la distribuzione condizionata $Y|x_1$

y_j	n_{1j}
0 $\rightarrow$ 10	40
10 $\rightarrow$ 30	0

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.



Si tratta di un carattere quantitativo con modalità raggruppate in classi.

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.



Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.

Si osserva che la distribuzione è degenere.

$M(Y|x_1) = 5$. La varianza e il coefficiente di variazione sono nulli.



Si considera la distribuzione condizionata $Y|x_2$

y_j	n_{2j}
0 $\rightarrow$ 10	20
10 $\rightarrow$ 30	20

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.



Si tratta di un carattere quantitativo con modalità raggruppate in classi.

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.



La media aritmetica

$$M(Y|x_2) = \frac{1}{n_{2\bullet}} \sum_{j=1}^k y_j n_{2j}$$

risulta 12.5.

Si osserva, infatti, che la distribuzione è simmetrica.

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.



La varianza

$$\text{Var}(Y|x_2) = M(Y^2|x_2) - [M(Y|x_2)]^2$$

risulta 56.25.

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.



Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.

La varianza

$$Var(Y|x_2) = M(Y^2|x_2) - [M(Y|x_2)]^2$$

risulta 56.25. Si ottiene (considerando come modalità i valori centrali delle classi)

y_j	n_{2j}	$y_j n_{2j}$	$y_j^2 n_{2j}$
5	20	100	500
20	20	400	8000
	40	500	8500

da cui segue

$$Var(Y|x_2) = \frac{8500}{40} - 12.5^2 = 212.5 - 156.25 = 56.25.$$



Si considera la distribuzione marginale della variabile Y , ottenuta sommando le frequenze di colonna

y_j	$n_{\bullet j}$
0 $\rightarrow$ 10	60
10 $\rightarrow$ 30	20

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.



Si tratta di un carattere quantitativo con modalità raggruppate in classi.

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.



La media aritmetica

$$M(Y) = \frac{1}{n} \sum_{j=1}^k y_j n_{\bullet j}$$

risulta 8.75.

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.



La media aritmetica

$$M(Y) = \frac{1}{n} \sum_{j=1}^k y_j n_{\bullet j}$$

risulta 8.75.

Si ottiene (considerando come modalità i valori centrali delle classi)

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.



La media aritmetica

$$M(Y) = \frac{1}{n} \sum_{j=1}^k y_j n_{\bullet j}$$

risulta 8.75.

Si ottiene (considerando come modalità i valori centrali delle classi)

y_j	$n_{\bullet j}$	$y_j n_{\bullet j}$
5	60	300
20	20	400
	80	700

da cui segue

$$M(Y) = \frac{1}{n} \sum_{i=1}^k y_i n_{\bullet i} = \frac{700}{80} = 8.75.$$

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.



La varianza

$$\text{Var}(Y) = M[(Y - \mu_Y)^2] = M(Y^2) - \mu_Y^2 = \frac{1}{n} \sum_{j=1}^k y_j^2 n_{\bullet j} - \mu_Y^2$$

risulta 42.1875.

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.



La varianza

$$Var(Y) = M[(Y - \mu_Y)^2] = M(Y^2) - \mu_Y^2 = \frac{1}{n} \sum_{j=1}^k y_j^2 n_{\bullet j} - \mu_Y^2$$

risulta 42.1875.

Si ottiene (considerando come modalità i valori centrali delle classi)

y_j	$n_{\bullet j}$	$y_j n_{\bullet j}$	$y_j^2 n_{\bullet j}$
5	60	300	1500
20	20	400	8000
	80	700	9500

da cui segue

$$\begin{aligned} Var(Y) = \sigma_Y^2 &= M(Y^2) - \mu_Y^2 = \\ &= \frac{9500}{80} - 8.75^2 = 118.75 - 76.5625 = 42.1875. \end{aligned}$$



Il rapporto di correlazione

$$\eta_{Y|X}^2 = \frac{\text{Var}_X[M(Y|X)]}{\text{Var}(Y)}$$

può essere ottenuto considerando il rapporto tra la varianza spiegata $\text{Var}_X[M(Y|X)]$,
varianza della variabile statistica medie condizionate

i	$M(X y_i)$	$n_{i\bullet}$
1	5.0	40
2	12.5	40

e la varianza della variabile dipendente $\text{Var}(Y) = 42.1875$.

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.



La varianza spiegata risulta $Var_X[M(Y|X)] = 14.0625$.

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.



La varianza spiegata risulta $\text{Var}_X[M(Y|X)] = 14.0625$. Si ottiene

i	$M(X y_i)$	$n_{i\bullet}$	$M(X y_i) \cdot n_{i\bullet}$	$[M(X y_i)]^2 \cdot n_{i\bullet}$
1	5.0	40	200	1000
2	12.5	40	500	6250
		80	700	7250

da cui segue

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.



La varianza spiegata risulta $Var_X[M(Y|X)] = 14.0625$. Si ottiene

i	$M(X y_i)$	$n_{i\bullet}$	$M(X y_i) \cdot n_{i\bullet}$	$[M(X y_i)]^2 \cdot n_{i\bullet}$
1	5.0	40	200	1000
2	12.5	40	500	6250
		80	700	7250

da cui segue

$$Var_X[M(Y|X)] = \frac{7250}{80} - 8.75^2 = 90.625 - 76.5625 = 14.0625.$$

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.



Il rapporto di correlazione $\eta_{Y|X}^2$ risulta:

$$\eta_{Y|X}^2 = \frac{\text{Var}_X[M(Y|X)]}{\text{Var}(Y)} = \frac{14.0625}{42.1875} = 0.3333.$$

Sussiste, quindi, un livello di dipendenza basso.

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.



Si ricorda che, in base al risultato di scomposizione della varianza, la varianza della variabile dipendente può essere ottenuta come somma della varianza spiegata $Var_X[M(Y|X)]$ e della varianza residua $M_X[Var(Y|X)]$.

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.



La varianza residua $M_X[Var(Y|X)]$, media della variabile statistica varianze condizionate

i	$Var(X y_i)$	$n_{i\bullet}$
1	0.00	40
2	56.25	40

risulta $M_X[Var(Y|X)] = 28.125$.

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.



Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.

La varianza residua $M_X[Var(Y|X)]$, media della variabile statistica varianze condizionate

i	$Var(X y_i)$	$n_{i\bullet}$
1	0.00	40
2	56.25	40

risulta $M_X[Var(Y|X)] = 28.125$. Si ottiene

i	$Var(X y_i)$	$n_{i\bullet}$	$Var(X y_i) \cdot n_{i\bullet}$
1	0.00	40	0
2	56.25	40	2250
		80	2250

da cui segue



La varianza residua $M_X[Var(Y|X)]$, media della variabile statistica varianze condizionate

i	$Var(X y_i)$	$n_{i\bullet}$
1	0.00	40
2	56.25	40

risulta $M_X[Var(Y|X)] = 28.125$. Si ottiene

i	$Var(X y_i)$	$n_{i\bullet}$	$Var(X y_i) \cdot n_{i\bullet}$
1	0.00	40	0
2	56.25	40	2250
		80	2250

da cui segue

$$M_X[Var(Y|X)] = \frac{2250}{80} = 28.125.$$



$\sigma_{Between}^2 = 14.0625$ $\sigma_{Within}^2 = 28.125$ da cui abbiamo

$$\eta_{Y|\text{giorno erogazione}}^2 = \frac{14.0625}{14.0625 + 28.125} = 0.3333.$$

Sussiste, quindi, un livello di dipendenza medio basso del costo della prestazione professionale rispetto al giorno di erogazione.

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.



Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.

3. Calcolare un opportuno indice di eterogeneità per la variabile X e commentarne il valore.



Si considera la distribuzione marginale della variabile X , ottenuta sommando le frequenze di riga

x_i	$n_{i\bullet}$
<i>feriale</i>	40
<i>festivo</i>	40

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.



La mutabile X presenta massima eterogeneità, quindi gli indici di Gini e di Frosini assumono valore massimo, $F_N = G_N = 1$.

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.



4. Valutare il grado di connessione fra i caratteri X e Y utilizzando un opportuno indice normalizzato.

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.



Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.

Dalla distribuzione delle frequenze congiunte n_{ij} :

	0 - 10	10 - 30	$n_{i\bullet}$
<i>feriale</i>	40	0	40
<i>festivo</i>	20	20	40
$n_{\bullet j}$	60	20	80

è possibile ottenere le frequenze teoriche nella situazione di indipendenza stocastica

$$\hat{n}_{ij} = \frac{n_{i\bullet} n_{\bullet j}}{n}:$$



Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.

Dalla distribuzione delle frequenze congiunte n_{ij} :

	0 - 10	10 - 30	$n_{i\bullet}$
<i>feriale</i>	40	0	40
<i>festivo</i>	20	20	40
$n_{\bullet j}$	60	20	80

è possibile ottenere le frequenze teoriche nella situazione di indipendenza stocastica

$$\hat{n}_{ij} = \frac{n_{i\bullet} n_{\bullet j}}{n}:$$

	0 - 10	10 - 30	$n_{i\bullet}$
<i>feriale</i>	30	10	40
<i>festivo</i>	30	10	40
$n_{\bullet j}$	60	20	80



le contingenze assolute $n_{ij} - \hat{n}_{ij}$:

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.



le contingenze assolute $n_{ij} - \hat{n}_{ij}$:

	0 - 10	10 - 30
<i>feriale</i>	10	-10
<i>festivo</i>	-10	10

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.



i rapporti di contingenza $\frac{n_{ij}}{\hat{n}_{ij}}$:

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.



i rapporti di contingenza $\frac{n_{ij}}{\hat{n}_{ij}}$:

	0 - 10	10 - 30
<i>feriale</i>	1.3333	0
<i>festivo</i>	0.6667	2

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.



e le contingenze relative, espresse in termini percentuali, $100 \left(\frac{n_{ij} - \hat{n}_{ij}}{\hat{n}_{ij}} \right) \% = 100 \left(\frac{n_{ij}}{\hat{n}_{ij}} - 1 \right) \%$:

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.



e le contingenze relative, espresse in termini percentuali, $100 \left(\frac{n_{ij} - \hat{n}_{ij}}{\hat{n}_{ij}} \right) \% = 100 \left(\frac{n_{ij}}{\hat{n}_{ij}} - 1 \right) \%$:

	0 - 10	10 - 30
<i>feriale</i>	33.33	-100
<i>festivo</i>	-33.33	100

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.



Risultati intermedi $\frac{(n_{ij} - \hat{n}_{ij})^2}{\hat{n}_{ij}}$ per il calcolo del χ^2

	0 - 10	10 - 30
<i>feriale</i>	3.3333	10
<i>festivo</i>	3.3333	10

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.



Risultati intermedi $\frac{n_{ij}^2}{n_{i\bullet} \cdot n_{\bullet j}}$ per il calcolo del χ^2 (formula operativa)

	0 - 10	10 - 30
<i>feriale</i>	0.6667	0.0
<i>festivo</i>	0.1667	0.5

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.



Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.

$$\chi^2 = \sum_{i=1}^h \sum_{j=1}^k \frac{(n_{ij} - \hat{n}_{ij})^2}{\hat{n}_{ij}} = 26.6667,$$

$$\chi_N^2 = \frac{\chi^2}{n \cdot \min(h-1, k-1)} = 0.3333$$

$$\chi_N = \sqrt{\chi_N^2} = 0.5773.$$

Sussiste, quindi, un livello di connessione medio.



La tabella seguente riporta la serie dei numeri indice dei prezzi a base fissa (con anno base 2015) per due beni A e B registrati fra il 2015 e il 2018:

t	$NIBF_A$	$NIBF_B$
2015	1	1
2016	1.05	1
2017	1.05	1.1
2018	1.1	1.1

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.



1. Costruire la serie dei numeri indice a base mobile per il bene A.

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.



1. Costruire la serie dei numeri indice a base mobile per il bene A.

Si ricorda che ${}_{t-1}I_t = \frac{x_t}{x_{t-1}} = \frac{{}_{2015}I_t}{{}_{2015}I_{t-1}}$.

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.



1. Costruire la serie dei numeri indice a base mobile per il bene A.

Si ricorda che ${}_{t-1}I_t = \frac{x_t}{x_{t-1}} = \frac{{}_{2015}I_t}{{}_{2015}I_{t-1}}$. Si ottiene quindi

Anno	NIBF	NIBM
2015	1.00	
2016	1.05	1.0500
2017	1.05	1.0000
2018	1.10	1.0476

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.



2. Calcolare il tasso medio di variazione annuo del bene A nell'intervallo 2015-2018 e commentare il risultato ottenuto.

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.



2. Calcolare il tasso medio di variazione annuo del bene A nell'intervallo 2015-2018 e commentare il risultato ottenuto.

Si ricorda che

$${}_{2015}I_{2018} = {}_{2015}I_{2016} \cdot {}_{2016}I_{2017} \cdot {}_{2017}I_{2018}$$

Si deve trovare il numero indice a base mobile medio.

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.



Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.

2. Calcolare il tasso medio di variazione annuo del bene A nell'intervallo 2015-2018 e commentare il risultato ottenuto.

Si ricorda che

$${}_{2015}I_{2018} = {}_{2015}I_{2016} \cdot {}_{2016}I_{2017} \cdot {}_{2017}I_{2018}$$

Si deve trovare il numero indice a base mobile medio. Applicando il criterio di scelta della media secondo Chisini si ottiene

$${}_{2015}I_{2018} = {}_{2015}I_{2016} \cdot {}_{2016}I_{2017} \cdot {}_{2017}I_{2018} = \alpha \cdot \alpha \cdot \alpha = \alpha^3.$$



Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.

2. Calcolare il tasso medio di variazione annuo del bene A nell'intervallo 2015-2018 e commentare il risultato ottenuto.

Si ricorda che

$${}_{2015}I_{2018} = {}_{2015}I_{2016} \cdot {}_{2016}I_{2017} \cdot {}_{2017}I_{2018}$$

Si deve trovare il numero indice a base mobile medio. Applicando il criterio di scelta della media secondo Chisini si ottiene

$${}_{2015}I_{2018} = {}_{2015}I_{2016} \cdot {}_{2016}I_{2017} \cdot {}_{2017}I_{2018} = \alpha \cdot \alpha \cdot \alpha = \alpha^3.$$

Quindi

$$\alpha = \sqrt[3]{{}_{2015}I_{2018}} = \sqrt[3]{{}_{2015}I_{2016} \cdot {}_{2016}I_{2017} \cdot {}_{2017}I_{2018}} = 1.0323.$$

Da cui segue una variazione relativa media pari a 0.0323, corrispondente a un incremento medio annuo tra il 2015 e il 2018 del 3.23%.



3. Sapendo che ${}_{2015}q_A = 100$ e ${}_{2015}q_B = 100$ sono le quantità vendute per i beni A e B nel 2015, calcolare l'indice di Laspeyres dei prezzi al 2016 con base al 2015 nell'ipotesi che il prezzo del bene B nell'anno base sia doppio rispetto a quello del bene A .

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.



Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.

3. Sapendo che ${}_{2015}q_A = 100$ e ${}_{2015}q_B = 100$ sono le quantità vendute per i beni A e B nel 2015, calcolare l'indice di Laspeyres dei prezzi al 2016 con base al 2015 nell'ipotesi che il prezzo del bene B nell'anno base sia doppio rispetto a quello del bene A .

L'indice dei prezzi di Laspeyres si ottiene considerando la media aritmetica dei 'prezzi relativi' riferiti ai singoli prodotti $x_i = \frac{{}_i p_{2016}}{{}_i p_{2015}}$ calcolata utilizzando come frequenze i valori

${}_i v_{2015} = {}_i p_{2015} \cdot {}_i q_{2015}$ scambiati al tempo 2015.



3. Sapendo che ${}_{2015}q_A = 100$ e ${}_{2015}q_B = 100$ sono le quantità vendute per i beni A e B nel 2015, calcolare l'indice di Laspeyres dei prezzi al 2016 con base al 2015 nell'ipotesi che il prezzo del bene B nell'anno base sia doppio rispetto a quello del bene A .

L'indice dei prezzi di Laspeyres si ottiene considerando la media aritmetica dei 'prezzi relativi' riferiti ai singoli prodotti $x_i = \frac{{}_ip_{2016}}{{}_ip_{2015}}$ calcolata utilizzando come frequenze i valori

${}_iv_{2015} = {}_ip_{2015} \cdot {}_iq_{2015}$ scambiati al tempo 2015.

Dal momento che le quantità scambiate nel 2015 sono le stesse per i due prodotti, ma ${}_Bp_{2015} = 2 \cdot {}_Ap_{2015}$ abbiamo che il valore del prodotto B scambiato nel 2015 è il doppio di quello del prodotto A : ${}_Bv_{2015} = 2 \cdot {}_Av_{2015}$.

$$\begin{aligned} {}^p_{2015}I^L_{2016} &= \frac{1}{\sum_{i=1}^h {}_ip_{2015} {}_iq_{2015}} \sum_{i=1}^h \frac{{}_ip_{2016}}{{}_ip_{2015}} {}_ip_{2015} {}_iq_{2015} = \\ &= \frac{1}{\sum_{i=1}^h {}_iv_{2015}} \sum_{i=1}^h \frac{{}_ip_{2016}}{{}_ip_{2015}} {}_iv_{2015} \end{aligned}$$



Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.

$\frac{i p_{2016}}{i p_{2015}}$	$i V_{2015}$	$\frac{i p_{2016}}{i p_{2015}} \cdot i V_{2015}$
1.05	1	1.05
1	2	2
	3	3.05

$$p_{2015}^p /_{2016}^L = \frac{1}{3} 3.05 = 1.0167.$$

La variazione relativa nel livello generale dei prezzi, secondo l'Indice di Laspeyres, è del 1.67%.



Con riferimento allo studio della dipendenza tra due caratteri quantitativi X e Y , si sono raccolti i dati relativi alle seguenti 5 coppie di osservazioni (x_i, y_i) $i = 1, 2, \dots, 5$

x_i	0	0	1	2	2
y_i	1	5	4	4	6

1. Fornire la rappresentazione grafica delle coppie di punti e della funzione di regressione di Y dato X .

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.



La funzione di regressione associa a ogni valore x appartenente all'insieme di definizione della variabile X la media di Y condizionata a x , $M(Y|x)$

x	$M(Y x)$
0	$\frac{1}{2}(1 + 5) = 3$
1	4
2	$\frac{1}{2}(4 + 6) = 5$

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.



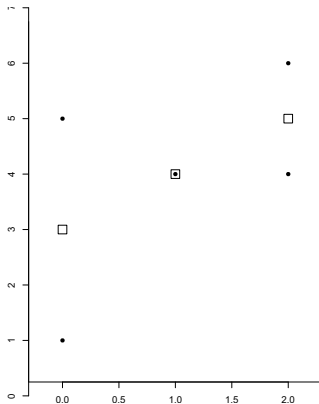
Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.

Le coppie di dati possono essere rappresentate mediante un grafico a dispersione.





2. Si calcolino, secondo il metodo dei minimi quadrati, i parametri dei seguenti modelli di regressione:

$$I) Y^* = a + b \cdot X$$

$$II) Y^* = c \cdot X$$

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.

Si riportano i calcoli necessari¹ per ottenere le stime dei parametri e gli indici di adattamento del modello

$$Y = a + bX + E$$

x_i	y_i	$x_i y_i$	x_i^2	y_i^2
0	1	0	0	1
0	5	0	0	25
1	4	4	1	16
2	4	8	4	16
2	6	12	4	36
5	20	24	9	94
1	4	4.8	1.8	18.8
$M(X)$	$M(Y)$	$M(XY)$	$M(X^2)$	$M(Y^2)$

$$\hat{b} =$$



Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.

¹Si osserva che per ragioni di stampa alcuni calcoli potrebbero essere duplicati con riferimento ai diversi modelli oggetto di studio

Si riportano i calcoli necessari¹ per ottenere le stime dei parametri e gli indici di adattamento del modello

$$Y = a + bX + E$$

x_i	y_i	$x_i y_i$	x_i^2	y_i^2
0	1	0	0	1
0	5	0	0	25
1	4	4	1	16
2	4	8	4	16
2	6	12	4	36
5	20	24	9	94
1	4	4.8	1.8	18.8
$M(X)$	$M(Y)$	$M(XY)$	$M(X^2)$	$M(Y^2)$

$$\begin{aligned}\hat{b} &= \frac{\text{Cov}(X, Y)}{\text{Var}(X)} = \frac{M(XY) - M(X)M(Y)}{M(X^2) - [M(X)]^2} = \\ &= \frac{4.8 - 1 \cdot 4}{1.8 - 1^2} = \frac{0.8}{0.8} = 1 \\ \hat{a} &= \end{aligned}$$

¹Si osserva che per ragioni di stampa alcuni calcoli potrebbero essere duplicati con riferimento ai diversi modelli oggetto di studio



Si riportano i calcoli necessari¹ per ottenere le stime dei parametri e gli indici di adattamento del modello

$$Y = a + bX + E$$

x_i	y_i	$x_i y_i$	x_i^2	y_i^2
0	1	0	0	1
0	5	0	0	25
1	4	4	1	16
2	4	8	4	16
2	6	12	4	36
5	20	24	9	94
1	4	4.8	1.8	18.8
$M(X)$	$M(Y)$	$M(XY)$	$M(X^2)$	$M(Y^2)$

$$\hat{b} = \frac{\text{Cov}(X, Y)}{\text{Var}(X)} = \frac{M(XY) - M(X)M(Y)}{M(X^2) - [M(X)]^2} =$$

$$= \frac{4.8 - 1 \cdot 4}{1.8 - 1^2} = \frac{0.8}{0.8} = 1$$

$$\hat{a} = M(Y) - \hat{b}M(X) = 3$$

$$\rho^2 =$$



Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.

¹Si osserva che per ragioni di stampa alcuni calcoli potrebbero essere duplicati con riferimento ai diversi modelli oggetto di studio

Si riportano i calcoli necessari¹ per ottenere le stime dei parametri e gli indici di adattamento del modello

$$Y = a + bX + E$$

x_i	y_i	$x_i y_i$	x_i^2	y_i^2
0	1	0	0	1
0	5	0	0	25
1	4	4	1	16
2	4	8	4	16
2	6	12	4	36
5	20	24	9	94
1	4	4.8	1.8	18.8
$M(X)$	$M(Y)$	$M(XY)$	$M(X^2)$	$M(Y^2)$

$$\hat{b} = \frac{Cov(X, Y)}{Var(X)} = \frac{M(XY) - M(X)M(Y)}{M(X^2) - [M(X)]^2} =$$

$$= \frac{4.8 - 1 \cdot 4}{1.8 - 1^2} = \frac{0.8}{0.8} = 1$$

$$\hat{a} = M(Y) - \hat{b}M(X) = 3$$

$$\rho^2 = \frac{[Cov(X, Y)]^2}{Var(X) Var(Y)} = \frac{0.8^2}{0.8 \cdot 2.8} = 0.2857$$

$$M(\hat{E}^2) = Var(\hat{E}) = \text{varianza residua} =$$

¹Si osserva che per ragioni di stampa alcuni calcoli potrebbero essere duplicati con riferimento ai diversi modelli oggetto di studio



Si riportano i calcoli necessari¹ per ottenere le stime dei parametri e gli indici di adattamento del modello

$$Y = a + bX + E$$

x_i	y_i	$x_i y_i$	x_i^2	y_i^2
0	1	0	0	1
0	5	0	0	25
1	4	4	1	16
2	4	8	4	16
2	6	12	4	36
5	20	24	9	94
1	4	4.8	1.8	18.8
$M(X)$	$M(Y)$	$M(XY)$	$M(X^2)$	$M(Y^2)$

$$\hat{b} = \frac{\text{Cov}(X, Y)}{\text{Var}(X)} = \frac{M(XY) - M(X)M(Y)}{M(X^2) - [M(X)]^2} =$$

$$= \frac{4.8 - 1 \cdot 4}{1.8 - 1^2} = \frac{0.8}{0.8} = 1$$

$$\hat{a} = M(Y) - \hat{b}M(X) = 3$$

$$\rho^2 = \frac{[\text{Cov}(X, Y)]^2}{\text{Var}(X) \text{Var}(Y)} = \frac{0.8^2}{0.8 \cdot 2.8} = 0.2857$$

$$M(\hat{E}^2) = \text{Var}(\hat{E}) = \text{varianza residua} = \text{Var}(Y) \cdot (1 - \rho^2) = 2$$

¹Si osserva che per ragioni di stampa alcuni calcoli potrebbero essere duplicati con riferimento ai diversi modelli oggetto di studio





Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.

Si riportano i calcoli necessari² per ottenere le stime dei parametri e la misura di adattamento del modello

$$Y = cX + E$$

x_i	y_i	$x_i y_i$	x_i^2	y_i^2
0	1	0	0	1
0	5	0	0	25
1	4	4	1	16
2	4	8	4	16
2	6	12	4	36
		24	9	94
		4.8	1.8	18.8
		$M(XY)$	$M(X^2)$	$M(Y^2)$

$$\hat{c} =$$

²Si osserva che per ragioni di stampa alcuni calcoli potrebbero essere duplicati con riferimento ai diversi modelli oggetto di studio



Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.

Si riportano i calcoli necessari² per ottenere le stime dei parametri e la misura di adattamento del modello

$$Y = cX + E$$

x_i	y_i	$x_i y_i$	x_i^2	y_i^2
0	1	0	0	1
0	5	0	0	25
1	4	4	1	16
2	4	8	4	16
2	6	12	4	36
		24	9	94
		4.8	1.8	18.8
		$M(XY)$	$M(X^2)$	$M(Y^2)$

$$\hat{c} = \frac{M(XY)}{M(X^2)} = \frac{4.8}{1.8} = 2.6667$$

$$M(\hat{E}^2) =$$

²Si osserva che per ragioni di stampa alcuni calcoli potrebbero essere duplicati con riferimento ai diversi modelli oggetto di studio



Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.

Si riportano i calcoli necessari² per ottenere le stime dei parametri e la misura di adattamento del modello

$$Y = cX + E$$

x_i	y_i	$x_i y_i$	x_i^2	y_i^2
0	1	0	0	1
0	5	0	0	25
1	4	4	1	16
2	4	8	4	16
2	6	12	4	36
		24	9	94
		4.8	1.8	18.8
		$M(XY)$	$M(X^2)$	$M(Y^2)$

$$\hat{c} = \frac{M(XY)}{M(X^2)} = \frac{4.8}{1.8} = 2.6667$$

$$M(\hat{E}^2) = M[Y \cdot (Y - \hat{c}X)] = M(Y^2) - \hat{c}M(XY) = 18.8 - 2.6667 \cdot 4.8 = 6$$

²Si osserva che per ragioni di stampa alcuni calcoli potrebbero essere duplicati con riferimento ai diversi modelli oggetto di studio



Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.

$\hat{a} = 3$, $\hat{b} = 1$, indice adattamento $M(\hat{E}^2) = 2$ $\hat{c} = 2.6667$, indice adattamento $M(\hat{E}^2) = 6$



Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.

3. Confrontare gli indici di adattamento, commentando le eventuali differenze.



Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.

3. Confrontare gli indici di adattamento, commentando le eventuali differenze.
Il modello migliore è $Y^* = 3 + X$.



4. Sulla base dei risultati precedenti e senza svolgere ulteriori conti, cosa è possibile dire sui valori di $\eta^2_{Y|X}$?

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.



4. Sulla base dei risultati precedenti e senza svolgere ulteriori conti, cosa è possibile dire sui valori di $\eta^2_{Y|X}$?

$\eta^2_{Y|X} = \rho^2$ in quanto le medie condizionate sono allineate.

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.



Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.

5. Sulla base del miglior modello prevedere il valore di Y in corrispondenza di $X = 3$.



Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.

5. Sulla base del miglior modello prevedere il valore di Y in corrispondenza di $X = 3$.
 $Y^*(3) = 3 + 1 \cdot 3 = 6$.



In una comunità la percentuale di pazienti allergici a pollini e graminacee è nota essere pari a 0.2. Sapendo che la probabilità di risultare positivo ad un test se un individuo è allergico è del 0.8 mentre la probabilità di risultare negativo se un individuo non è allergico è pari a 0.6 si risponda alle seguenti domande:

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.



1. Calcolare la probabilità di risultare positivo al test per un individuo selezionato casualmente.

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.



Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.

1. Calcolare la probabilità di risultare positivo al test per un individuo selezionato casualmente.

Indicati con A ed $\bar{A}$ gli eventi 'allergico' e 'non allergico' e con $+$ e $-$ gli eventi 'positivo al test' e 'negativo al test' sono disponibili le seguenti informazioni:

$$P(A) = 0.2, P(+|A) = 0.8, P(-|\bar{A}) = 0.6.$$

Applicando la legge delle probabilità totali abbiamo:

$$P(+) =$$



1. Calcolare la probabilità di risultare positivo al test per un individuo selezionato casualmente.

Indicati con A ed $\bar{A}$ gli eventi 'allergico' e 'non allergico' e con $+$ e $-$ gli eventi 'positivo al test' e 'negativo al test' sono disponibili le seguenti informazioni:

$$P(A) = 0.2, P(+|A) = 0.8, P(-|\bar{A}) = 0.6.$$

Applicando la legge delle probabilità totali abbiamo:

$$\begin{aligned} P(+) &= P(+ \cap (A \cup \bar{A})) = \\ &= \end{aligned}$$

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.



1. Calcolare la probabilità di risultare positivo al test per un individuo selezionato casualmente.

Indicati con A ed $\bar{A}$ gli eventi 'allergico' e 'non allergico' e con $+$ e $-$ gli eventi 'positivo al test' e 'negativo al test' sono disponibili le seguenti informazioni:

$$P(A) = 0.2, P(+|A) = 0.8, P(-|\bar{A}) = 0.6.$$

Applicando la legge delle probabilità totali abbiamo:

$$\begin{aligned} P(+) &= P(+ \cap (A \cup \bar{A})) = \\ &= P((+ \cap A) \cup (+ \cap \bar{A})) = \\ &= \end{aligned}$$

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.



Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.

1. Calcolare la probabilità di risultare positivo al test per un individuo selezionato casualmente.

Indicati con A ed $\bar{A}$ gli eventi 'allergico' e 'non allergico' e con $+$ e $-$ gli eventi 'positivo al test' e 'negativo al test' sono disponibili le seguenti informazioni:

$$P(A) = 0.2, P(+|A) = 0.8, P(-|\bar{A}) = 0.6.$$

Applicando la legge delle probabilità totali abbiamo:

$$\begin{aligned} P(+) &= P(+ \cap (A \cup \bar{A})) = \\ &= P((+ \cap A) \cup (+ \cap \bar{A})) = \\ &= P(+ \cap A) + P(+ \cap \bar{A}) = \\ &= \end{aligned}$$



Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.

1. Calcolare la probabilità di risultare positivo al test per un individuo selezionato casualmente.

Indicati con A ed $\bar{A}$ gli eventi 'allergico' e 'non allergico' e con $+$ e $-$ gli eventi 'positivo al test' e 'negativo al test' sono disponibili le seguenti informazioni:

$$P(A) = 0.2, P(+|A) = 0.8, P(-|\bar{A}) = 0.6.$$

Applicando la legge delle probabilità totali abbiamo:

$$\begin{aligned} P(+) &= P(+ \cap (A \cup \bar{A})) = \\ &= P((+ \cap A) \cup (+ \cap \bar{A})) = \\ &= P(+ \cap A) + P(+ \cap \bar{A}) = \\ &= P(A) \cdot P(+|A) + P(\bar{A}) \cdot P(+|\bar{A}) = \\ &= \end{aligned}$$



Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.

1. Calcolare la probabilità di risultare positivo al test per un individuo selezionato casualmente.

Indicati con A ed $\bar{A}$ gli eventi 'allergico' e 'non allergico' e con $+$ e $-$ gli eventi 'positivo al test' e 'negativo al test' sono disponibili le seguenti informazioni:

$$P(A) = 0.2, P(+|A) = 0.8, P(-|\bar{A}) = 0.6.$$

Applicando la legge delle probabilità totali abbiamo:

$$\begin{aligned} P(+) &= P(+ \cap (A \cup \bar{A})) = \\ &= P((+ \cap A) \cup (+ \cap \bar{A})) = \\ &= P(+ \cap A) + P(+ \cap \bar{A}) = \\ &= P(A) \cdot P(+|A) + P(\bar{A}) \cdot P(+|\bar{A}) = \\ &= P(A) \cdot P(+|A) + (1 - P(A)) \cdot (1 - P(-|\bar{A})) = \\ &= \end{aligned}$$



Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.

1. Calcolare la probabilità di risultare positivo al test per un individuo selezionato casualmente.

Indicati con A ed $\bar{A}$ gli eventi 'allergico' e 'non allergico' e con $+$ e $-$ gli eventi 'positivo al test' e 'negativo al test' sono disponibili le seguenti informazioni:

$$P(A) = 0.2, P(+|A) = 0.8, P(-|\bar{A}) = 0.6.$$

Applicando la legge delle probabilità totali abbiamo:

$$\begin{aligned} P(+) &= P(+ \cap (A \cup \bar{A})) = \\ &= P((+ \cap A) \cup (+ \cap \bar{A})) = \\ &= P(+ \cap A) + P(+ \cap \bar{A}) = \\ &= P(A) \cdot P(+|A) + P(\bar{A}) \cdot P(+|\bar{A}) = \\ &= P(A) \cdot P(+|A) + (1 - P(A)) \cdot (1 - P(-|\bar{A})) = \\ &= 0.2 \cdot 0.8 + 0.8 \cdot 0.4 = \end{aligned}$$



Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.

1. Calcolare la probabilità di risultare positivo al test per un individuo selezionato casualmente.

Indicati con A ed $\bar{A}$ gli eventi 'allergico' e 'non allergico' e con $+$ e $-$ gli eventi 'positivo al test' e 'negativo al test' sono disponibili le seguenti informazioni:

$$P(A) = 0.2, P(+|A) = 0.8, P(-|\bar{A}) = 0.6.$$

Applicando la legge delle probabilità totali abbiamo:

$$\begin{aligned} P(+) &= P(+ \cap (A \cup \bar{A})) = \\ &= P((+ \cap A) \cup (+ \cap \bar{A})) = \\ &= P(+ \cap A) + P(+ \cap \bar{A}) = \\ &= P(A) \cdot P(+|A) + P(\bar{A}) \cdot P(+|\bar{A}) = \\ &= P(A) \cdot P(+|A) + (1 - P(A)) \cdot (1 - P(-|\bar{A})) = \\ &= 0.2 \cdot 0.8 + 0.8 \cdot 0.4 = 0.16 + 0.32 = 0.48 \end{aligned}$$



2. Calcolare la probabilità di essere allergico, dato che un paziente è risultato positivo al test.

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.



Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.

2. Calcolare la probabilità di essere allergico, dato che un paziente è risultato positivo al test.

$$P(A|+) = \frac{P(A \cap +)}{P(+)} = \frac{0.16}{0.48} = 0.3333$$

3. Estraendo quattro individui con reimmissione a caso dalla comunità, calcolare che almeno uno risulti allergico.



2. Calcolare la probabilità di essere allergico, dato che un paziente è risultato positivo al test.

$$P(A|+) = \frac{P(A \cap +)}{P(+)} = \frac{0.16}{0.48} = 0.3333$$

3. Estraendo quattro individui con reimmissione a caso dalla comunità, calcolare che almeno uno risulti allergico.

Sia X la variabile casuale 'n. di individui allergici tra i 4 estratti'. Abbiamo

$$X \sim \text{Bin}(n = 4, p = 0.2)$$

e

$$P(X \geq 1) = 1 - P(X = 0) = 1 - \binom{4}{0} 0.2^0 \cdot 0.8^4 = 1 - 0.4096 = 0.5904.$$

Tema T3042

Prova scritta di STATISTICA del 11.07.2019

Prova scritta di
STATISTICA del
11.07.2019

Giuseppe Boari
Gabriele Cantaluppi



UNIVERSITÀ
CATTOLICA
del Sacro Cuore

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.

Giuseppe Boari
Gabriele Cantaluppi
Università Cattolica del Sacro Cuore, Milano

Esercizio 1

Il prospetto seguente riporta la distribuzione congiunta del carattere qualitativo X e del carattere quantitativo Y .

$X \setminus Y$	1 + 3	3 + 7	7 + 9
x_1	12		
x_2		4	2

1. Si completi la distribuzione delle frequenze congiunte, sapendo che il numero totale di unità statistiche è 20, che $M(Y|x_1) = 2$ e che $M(Y|x_2) = 5$.



Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.

Esercizio 1

Il prospetto seguente riporta la distribuzione congiunta del carattere qualitativo X e del carattere quantitativo Y .

$X \setminus Y$	1 + 3	3 + 7	7 + 9	
x_1	12			
x_2		4	2	

1. Si completi la distribuzione delle frequenze congiunte, sapendo che il numero totale di unità statistiche è 20, che $M(Y|x_1) = 2$ e che $M(Y|x_2) = 5$.

$M(Y|x_1) = 2$ implica che la distribuzione condizionata $Y|x_1$ è degenere e, quindi,

$$n_{12} = n_{13} = 0.$$

$$n_{21} = 2 \text{ realizza } M(Y|x_2) = 5$$

$X \setminus Y$	1 + 3	3 + 7	7 + 9	
x_1	12	0	0	12
x_2	2	4	2	8
	14	4	2	20



Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.



2. Si rappresentino graficamente le distribuzioni marginali X e Y .

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.



Si considera la distribuzione marginale della variabile X , ottenuta sommando le frequenze di riga

x_i	$n_{i\bullet}$
x_1	12
x_2	8

Si tratta di un carattere qualitativo sconnesso che può essere rappresentato graficamente mediante un grafico a bastoncini o a canne d'organo.

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.

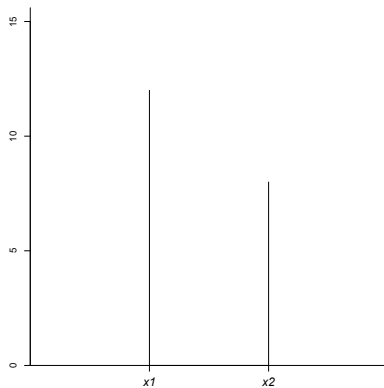


Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.





Si considera la distribuzione marginale della variabile Y , ottenuta sommando le frequenze di colonna

y_j	$n_{\bullet j}$
1 + 3	14
3 + 7	4
7 + 9	2

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.



Si tratta di un carattere quantitativo con modalità raggruppate in classi che può essere rappresentato graficamente mediante un istogramma.

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.



L'istogramma può essere costruito dopo avere calcolato le densità $d_{\bullet j} = \frac{n_{\bullet j}}{a_j}$, (dove a_j sono le ampiezze di classe).

y_j	$n_{\bullet j}$	$a_{\bullet j}$	$d_{\bullet j}$
1 - 3	14	3 - 1 = 2	7
3 - 7	4	7 - 3 = 4	1
7 - 9	2	9 - 7 = 2	1

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.

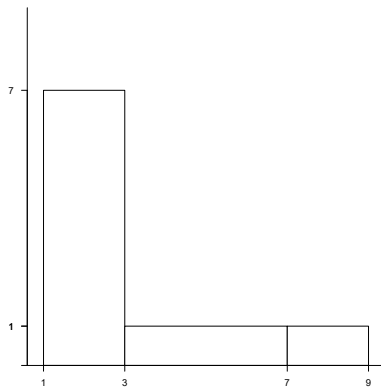


Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.





Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.

2. Si calcolino media, varianza e coefficiente di variazione per la distribuzione marginale Y .



Si tratta di un carattere quantitativo con modalità raggruppate in classi.

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.



La media aritmetica

$$M(Y) = \frac{1}{n} \sum_{j=1}^k y_j n_{\bullet j}$$

risulta 3.2.

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.



La media aritmetica

$$M(Y) = \frac{1}{n} \sum_{j=1}^k y_j n_{\bullet j}$$

risulta 3.2.

Si ottiene (considerando come modalità i valori centrali delle classi)

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.



La media aritmetica

$$M(Y) = \frac{1}{n} \sum_{j=1}^k y_j n_{\bullet j}$$

risulta 3.2.

Si ottiene (considerando come modalità i valori centrali delle classi)

y_j	$n_{\bullet j}$	$y_j n_{\bullet j}$
2	14	28
5	4	20
8	2	16
	20	64

da cui segue

$$M(Y) = \frac{1}{n} \sum_{i=1}^k y_i n_{\bullet i} = \frac{64}{20} = 3.2.$$

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.



La varianza

$$\text{Var}(Y) = M[(Y - \mu_Y)^2] = M(Y^2) - \mu_Y^2 = \frac{1}{n} \sum_{j=1}^k y_j^2 n_{\bullet j} - \mu_Y^2$$

risulta 3.96.

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.



La varianza

$$Var(Y) = M[(Y - \mu_Y)^2] = M(Y^2) - \mu_Y^2 = \frac{1}{n} \sum_{j=1}^k y_j^2 n_{\bullet j} - \mu_Y^2$$

risulta 3.96.

Si ottiene (considerando come modalità i valori centrali delle classi)

y_j	$n_{\bullet j}$	$y_j n_{\bullet j}$	$y_j^2 n_{\bullet j}$
2	14	28	56
5	4	20	100
8	2	16	128
	20	64	284

da cui segue

$$\begin{aligned} Var(Y) = \sigma_Y^2 &= M(Y^2) - \mu_Y^2 = \\ &= \frac{284}{20} - 3.2^2 = 14.2 - 10.24 = 3.96. \end{aligned}$$

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.



Per effettuare confronti in termini di variabilità tra caratteri quantitativi occorre utilizzare una misura relativa, ad esempio il coefficiente di variazione

$$CV = \frac{\sigma}{\mu} = \frac{\sqrt{\text{Var}(Y)}}{M(Y)} = \frac{\sqrt{3.96}}{3.2} = \frac{1.99}{3.2} = 0.6219.$$

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.



Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.

3. Si costruisca il grafico Box & Whiskers Plot per la distribuzione condizionata $Y|x_2$.



Si considera la distribuzione marginale della variabile Y , ottenuta sommando le frequenze di colonna

y_j	$n_{\bullet j}$
1 $\rightarrow$ 3	2
3 $\rightarrow$ 7	4
7 $\rightarrow$ 9	2

Si tratta di un carattere quantitativo con modalità raggruppate in classi.

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.



Per costruire il grafico Box & Whiskers plot occorre determinare in primo luogo i valori dei tre quartili $y_{0.25}$, $y_{0.5}$ e $y_{0.75}$ considerando l'inversa della Funzione di Ripartizione in corrispondenza rispettivamente di 0.25, 0.50 e 0.75.

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.

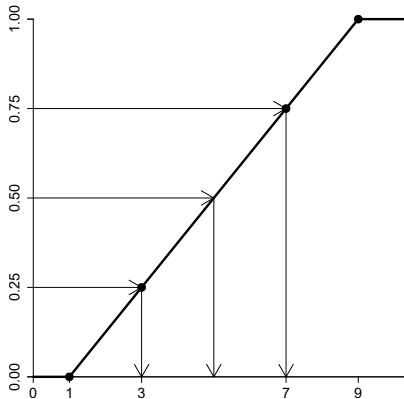


Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.





Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.

Si ottiene

$$y_{0.25} = 1 + \frac{0.25 - 0}{0.25 - 0} \cdot (3 - 1) = 3$$

$$y_{0.5} = 3 + \frac{0.5 - 0.25}{0.75 - 0.25} \cdot (7 - 3) = 5$$

$$y_{0.75} = 3 + \frac{0.75 - 0.25}{0.75 - 0.25} \cdot (7 - 3) = 7$$

La lunghezza teorica dei baffi risulta

$$1.5 \cdot (y_{0.75} - y_{0.25}) = 6.$$

Il baffo inferiore è posizionato in

$$\max(y_{\min}, y_{0.25} - 6) = \max(1, -3) = 1.$$

Il baffo superiore è posizionato in

$$\min(y_{\max}, y_{0.75} + 6) = \min(9, 13) = 9.$$

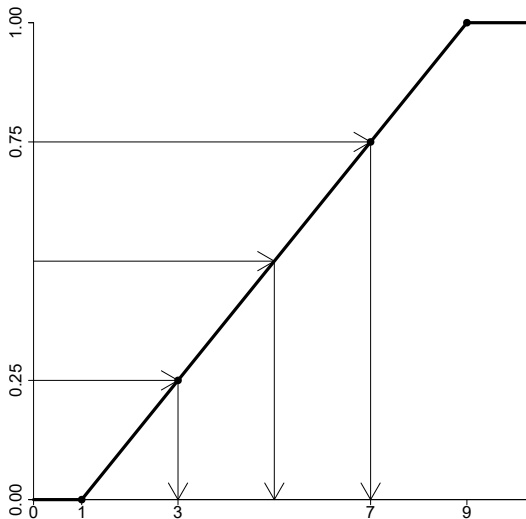


Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.



Esercizio 2

Prova scritta di
STATISTICA del
11.07.2019

Giuseppe Boari
Gabriele Cantaluppi



UNIVERSITÀ
CATTOLICA
del Sacro Cuore

Si riportano i prezzi di un prodotto registrati fra il 2015 e il 2018:

t	x_t
2015	20
2016	22
2017	25
2018	27.1

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.

1. Si costruiscano le serie dei numeri indice a base mobile e a base fissa con base 2015.



Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.



1. Si costruiscano le serie dei numeri indice a base mobile e a base fissa con base 2015.
Si ricorda che ${}_{t-1}I_t = \frac{x_t}{x_{t-1}}$.

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.



1. Si costruiscano le serie dei numeri indice a base mobile e a base fissa con base 2015. Si ricorda che ${}_{t-1}I_t = \frac{x_t}{x_{t-1}}$. Si ottiene quindi

Anno	x_t	NIBM
2015	20.0	
2016	22.0	1.1000
2017	25.0	1.1364
2018	27.1	1.0840

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.



Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.

1. Si costruiscano le serie dei numeri indice a base mobile e a base fissa con base 2015. Si ricorda che ${}_{t-1}I_t = \frac{x_t}{x_{t-1}}$. Si ottiene quindi

Anno	x_t	NIBM
2015	20.0	
2016	22.0	1.1000
2017	25.0	1.1364
2018	27.1	1.0840

Si ricorda che ${}_{2015}I_t = \frac{x_t}{x_{2015}}$.



Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.

1. Si costruiscano le serie dei numeri indice a base mobile e a base fissa con base 2015. Si ricorda che ${}_{t-1}I_t = \frac{x_t}{x_{t-1}}$. Si ottiene quindi

Anno	x_t	NIBM
2015	20.0	
2016	22.0	1.1000
2017	25.0	1.1364
2018	27.1	1.0840

Si ricorda che ${}_{2015}I_t = \frac{x_t}{x_{2015}}$. Si ottiene quindi

Anno	x_t	NIBF (2015)
2015	20.0	1.000
2016	22.0	1.100
2017	25.0	1.250
2018	27.1	1.355



2. Si calcoli il numero indice al tempo 2016 con base 2017 interpretando il valore ottenuto.

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.



2. Si calcoli il numero indice al tempo 2016 con base 2017 interpretando il valore ottenuto.

$${}_{2017}I_{2016} = \frac{X_{2016}}{X_{2017}} = \frac{22}{25} = 0.88$$

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.



Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.

2. Si calcoli il numero indice al tempo 2016 con base 2017 interpretando il valore ottenuto.

$${}_{2017}I_{2016} = \frac{X_{2016}}{X_{2017}} = \frac{22}{25} = 0.88$$

La grandezza al tempo 2016 era inferiore del 12% rispetto alla grandezza al tempo 2017. Ovvero il valore della grandezza nel 2016 era il 88% del valore registrato nel 2017.



3. Si calcoli il tasso medio di variazione tra il 2016 e il 2018.

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.



Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.

3. Si calcoli il tasso medio di variazione tra il 2016 e il 2018.
Si ricorda che

$${}_{2016}I_{2018} = {}_{2016}I_{2017} \cdot {}_{2017}I_{2018}$$

Si deve trovare il numero indice a base mobile medio.



Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.

3. Si calcoli il tasso medio di variazione tra il 2016 e il 2018.

Si ricorda che

$${}_{2016}I_{2018} = {}_{2016}I_{2017} \cdot {}_{2017}I_{2018}$$

Si deve trovare il numero indice a base mobile medio. Applicando il criterio di scelta della media secondo Chisini si ottiene

$${}_{2016}I_{2018} = {}_{2016}I_{2017} \cdot {}_{2017}I_{2018} = \alpha \cdot \alpha = \alpha^2.$$



Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.

3. Si calcoli il tasso medio di variazione tra il 2016 e il 2018.

Si ricorda che

$${}_{2016}l_{2018} = {}_{2016}l_{2017} \cdot {}_{2017}l_{2018}$$

Si deve trovare il numero indice a base mobile medio. Applicando il criterio di scelta della media secondo Chisini si ottiene

$${}_{2016}l_{2018} = {}_{2016}l_{2017} \cdot {}_{2017}l_{2018} = \alpha \cdot \alpha = \alpha^2.$$

Quindi

$$\alpha = \sqrt[2]{{}_{2016}l_{2018}} = \sqrt[2]{{}_{2016}l_{2017} \cdot {}_{2017}l_{2018}} = 1.1099.$$

Da cui segue una variazione relativa media pari a 0.1099, corrispondente a un incremento medio annuo tra il 2016 e il 2018 del 10.99%.



Il seguente prospetto riporta i valori rilevati in corrispondenza a 5 unità statistiche con riferimento a due caratteri quantitativi X e Y tranne che per l'osservazione y_2

x_i	2	3	3	3	5
y_i	2		3	4	5

1. Si rappresentino graficamente le coppie dei dati, con anche il valore di y_2 tale per cui $\eta_{Y|X}^2 = \rho^2$. Si completi il grafico con la funzione di regressione ipotizzando Y in funzione di X .

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.



Il seguente prospetto riporta i valori rilevati in corrispondenza a 5 unità statistiche con riferimento a due caratteri quantitativi X e Y tranne che per l'osservazione y_2

x_i	2	3	3	3	5
y_i	2		3	4	5

1. Si rappresentino graficamente le coppie dei dati, con anche il valore di y_2 tale per cui $\eta_{Y|X}^2 = \rho^2$. Si completi il grafico con la funzione di regressione ipotizzando Y in funzione di X . $\eta_{Y|X}^2 = \rho^2$ se le medie condizionate $M(Y|X)$ sono allineate e con $y_2 = 2$ abbiamo $M(Y|X = 3) = 3$.

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.



La funzione di regressione associa a ogni valore x appartenente all'insieme di definizione della variabile X la media di Y condizionata a x , $M(Y|x)$

x	$M(Y x)$
2	2
3	$\frac{1}{3}(2 + 3 + 4) = 3$
5	5

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.



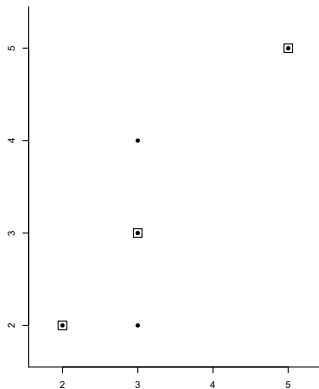
Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.

Le coppie di dati possono essere rappresentate mediante un grafico a dispersione.





2. Si calcolino, secondo il metodo dei minimi quadrati, i parametri dei seguenti modelli di regressione:

$$I) Y^* = a + b \cdot X \qquad II) Y^* = c$$

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.



Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.

Si riportano i calcoli necessari¹ per ottenere le stime dei parametri e gli indici di adattamento del modello

$$Y = a + bX + E$$

x_i	y_i	$x_i y_i$	x_i^2	y_i^2
2	2	4	4	4
3	2	6	9	4
3	3	9	9	9
3	4	12	9	16
5	5	25	25	25
16	16	56	56	58
3.2	3.2	11.2	11.2	11.6
$M(X)$	$M(Y)$	$M(XY)$	$M(X^2)$	$M(Y^2)$

$$\hat{b} =$$

¹Si osserva che per ragioni di stampa alcuni calcoli potrebbero essere duplicati con riferimento ai diversi modelli oggetto di studio



Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.

Si riportano i calcoli necessari¹ per ottenere le stime dei parametri e gli indici di adattamento del modello

$$Y = a + bX + E$$

x_i	y_i	$x_i y_i$	x_i^2	y_i^2
2	2	4	4	4
3	2	6	9	4
3	3	9	9	9
3	4	12	9	16
5	5	25	25	25
16	16	56	56	58
3.2	3.2	11.2	11.2	11.6
$M(X)$	$M(Y)$	$M(XY)$	$M(X^2)$	$M(Y^2)$

$$\begin{aligned}\hat{b} &= \frac{\text{Cov}(X, Y)}{\text{Var}(X)} = \frac{M(XY) - M(X)M(Y)}{M(X^2) - [M(X)^2]} = \\ &= \frac{11.2 - 3.2 \cdot 3.2}{11.2 - 3.2^2} = \frac{0.96}{0.96} = 1 \\ \hat{a} &= \end{aligned}$$

¹Si osserva che per ragioni di stampa alcuni calcoli potrebbero essere duplicati con riferimento ai diversi modelli oggetto di studio



Si riportano i calcoli necessari¹ per ottenere le stime dei parametri e gli indici di adattamento del modello

$$Y = a + bX + E$$

x_i	y_i	$x_i y_i$	x_i^2	y_i^2
2	2	4	4	4
3	2	6	9	4
3	3	9	9	9
3	4	12	9	16
5	5	25	25	25
16	16	56	56	58
3.2	3.2	11.2	11.2	11.6
$M(X)$	$M(Y)$	$M(XY)$	$M(X^2)$	$M(Y^2)$

$$\hat{b} = \frac{\text{Cov}(X, Y)}{\text{Var}(X)} = \frac{M(XY) - M(X)M(Y)}{M(X^2) - [M(X)]^2} =$$

$$= \frac{11.2 - 3.2 \cdot 3.2}{11.2 - 3.2^2} = \frac{0.96}{0.96} = 1$$

$$\hat{a} = M(Y) - \hat{b}M(X) = 0$$

$$\rho^2 =$$

¹Si osserva che per ragioni di stampa alcuni calcoli potrebbero essere duplicati con riferimento ai diversi modelli oggetto di studio



Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.

Si riportano i calcoli necessari¹ per ottenere le stime dei parametri e gli indici di adattamento del modello

$$Y = a + bX + E$$

x_i	y_i	$x_i y_i$	x_i^2	y_i^2
2	2	4	4	4
3	2	6	9	4
3	3	9	9	9
3	4	12	9	16
5	5	25	25	25
16	16	56	56	58
3.2	3.2	11.2	11.2	11.6
$M(X)$	$M(Y)$	$M(XY)$	$M(X^2)$	$M(Y^2)$

$$\begin{aligned}\hat{b} &= \frac{\text{Cov}(X, Y)}{\text{Var}(X)} = \frac{M(XY) - M(X)M(Y)}{M(X^2) - [M(X)]^2} = \\ &= \frac{11.2 - 3.2 \cdot 3.2}{11.2 - 3.2^2} = \frac{0.96}{0.96} = 1 \\ \hat{a} &= M(Y) - \hat{b}M(X) = 0 \\ \rho^2 &= \frac{[\text{Cov}(X, Y)]^2}{\text{Var}(X) \text{Var}(Y)} = \frac{0.96^2}{0.96 \cdot 1.36} = 0.7059\end{aligned}$$

¹Si osserva che per ragioni di stampa alcuni calcoli potrebbero essere duplicati con riferimento ai diversi modelli oggetto di studio



Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.

Si riportano i calcoli necessari² per ottenere le stime dei parametri e gli indici di adattamento del modello

$$Y = c + E$$

x_i	y_i	y_i^2
2	2	4
3	2	4
3	3	9
3	4	16
5	5	25
16		58
3.2		11.6
$M(Y)$		$M(Y^2)$

$$\hat{c} =$$

²Si osserva che per ragioni di stampa alcuni calcoli potrebbero essere duplicati con riferimento ai diversi modelli oggetto di studio



Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.

Si riportano i calcoli necessari² per ottenere le stime dei parametri e gli indici di adattamento del modello

$$Y = c + E$$

x_i	y_i	y_i^2
2	2	4
3	2	4
3	3	9
3	4	16
5	5	25
16		58
3.2		11.6
$M(Y)$		$M(Y^2)$

$$\hat{c} = M(Y) = 3.2$$

$$R^2 =$$

²Si osserva che per ragioni di stampa alcuni calcoli potrebbero essere duplicati con riferimento ai diversi modelli oggetto di studio



Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.

Si riportano i calcoli necessari² per ottenere le stime dei parametri e gli indici di adattamento del modello

$$Y = c + E$$

x_i	y_i	y_i^2
2	2	4
3	2	4
3	3	9
3	4	16
5	5	25
16		58
3.2		11.6
$M(Y)$		$M(Y^2)$

$$\hat{c} = M(Y) = 3.2$$

$$R^2 = 0$$

²Si osserva che per ragioni di stampa alcuni calcoli potrebbero essere duplicati con riferimento ai diversi modelli oggetto di studio



Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.

$\hat{a} = 0$, $\hat{b} = 1$, indice adattamento $\rho^2 = 0.7059$ $\hat{c} = 3.2$, indice adattamento $R^2 = 0$



3. Si indichi, motivando, quale modello presenta il migliore adattamento.

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.



Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.

3. Si indichi, motivando, quale modello presenta il migliore adattamento.

Il modello migliore è $Y^* = X$ in quanto caratterizzato da indice di adattamento più elevato.



4. Indicare l'intervallo dei valori possibili per $\eta_{X|Y}^2$.

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.



Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.

4. Indicare l'intervallo dei valori possibili per $\eta_{X|Y}^2$.

$\rho^2 = 0.7059 < \eta_{X|Y}^2 < 1$ in quanto le medie condizionate $M(X|Y)$ non sono allineate e la distribuzione condizionata $X|Y = 2$ non è degenere.



Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.

5. Sulla base del miglior modello prevedere il valore di Y in corrispondenza di $X = 3$.



Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.

5. Sulla base del miglior modello prevedere il valore di Y in corrispondenza di $X = 3$.
 $Y^*(3) = 1 \cdot 3 = 3$.

Esercizio 4

Prova scritta di
STATISTICA del
11.07.2019

Giuseppe Boari
Gabriele Cantaluppi



UNIVERSITÀ
CATTOLICA
del Sacro Cuore

In una sala di un museo sono presenti le 4 opere $\{a, b, c, d\}$. Ogni visitatore può visionare le opere secondo un ordine arbitrario.

1. Si calcoli la probabilità che un visitatore visioni le opere secondo la sequenza d, b, a, c oppure secondo la sequenza b, d, a, c .

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.



Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.

In una sala di un museo sono presenti le 4 opere $\{a, b, c, d\}$. Ogni visitatore può visionare le opere secondo un ordine arbitrario.

1. Si calcoli la probabilità che un visitatore visioni le opere secondo la sequenza d, b, a, c oppure secondo la sequenza b, d, a, c .

Le due sequenze hanno la stessa probabilità di verificarsi:

$$P(d \cap b \cap a \cap c) = P(d) \cdot P(b|d) \cdot P(a|d \cap b) \cdot P(c|d \cap b \cap a) = \frac{1}{4} \cdot \frac{1}{3} \cdot \frac{1}{2} \cdot \frac{1}{1} = \frac{1}{24}$$

e



Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.

In una sala di un museo sono presenti le 4 opere $\{a, b, c, d\}$. Ogni visitatore può visionare le opere secondo un ordine arbitrario.

1. Si calcoli la probabilità che un visitatore visioni le opere secondo la sequenza d, b, a, c oppure secondo la sequenza b, d, a, c .

Le due sequenze hanno la stessa probabilità di verificarsi:

$$P(d \cap b \cap a \cap c) = P(d) \cdot P(b|d) \cdot P(a|d \cap b) \cdot P(c|d \cap b \cap a) = \frac{1}{4} \cdot \frac{1}{3} \cdot \frac{1}{2} \cdot \frac{1}{1} = \frac{1}{24}$$

e

$$P(\{d \cap b \cap a \cap c\} \cup \{b \cap d \cap a \cap c\}) = P(\{d \cap b \cap a \cap c\}) + P(\{b \cap d \cap a \cap c\}) = \frac{1}{24} + \frac{1}{24} = \frac{1}{12} = 0.08333.$$



2. Si calcoli la probabilità che scelti a caso 3 visitatori, che hanno comportamento indipendente, questi adottino la stessa sequenza di visione delle opere.

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.



2. Si calcoli la probabilità che scelti a caso 3 visitatori, che hanno comportamento indipendente, questi adottino la stessa sequenza di visione delle opere.

$$24 \cdot \frac{1}{24} \cdot \frac{1}{24} \cdot \frac{1}{24} = \frac{24}{24^3} = \frac{1}{24^2} = 0.001736$$

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.



3. In prossimità dell'orario di chiusura sono presenti 10 visitatori. Viene data loro la possibilità di visionare una sola opera. Si calcoli la probabilità che almeno 3 visitatori scelgano l'opera a , assumendo che la scelta dell'opera venga effettuata dai visitatori in maniera casuale e indipendente.

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.



Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.

3. In prossimità dell'orario di chiusura sono presenti 10 visitatori. Viene data loro la possibilità di visionare una sola opera. Si calcoli la probabilità che almeno 3 visitatori scelgano l'opera a , assumendo che la scelta dell'opera venga effettuata dai visitatori in maniera casuale e indipendente.

Sia X la variabile casuale 'n. di visitatori che scelgono l'opera a tra i 10 presenti'. Abbiamo

$$X \sim$$



Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.

3. In prossimità dell'orario di chiusura sono presenti 10 visitatori. Viene data loro la possibilità di visionare una sola opera. Si calcoli la probabilità che almeno 3 visitatori scelgano l'opera a , assumendo che la scelta dell'opera venga effettuata dai visitatori in maniera casuale e indipendente.

Sia X la variabile casuale 'n. di visitatori che scelgono l'opera a tra i 10 presenti'. Abbiamo

$$X \sim \text{Bin}(n = 10, p = 0.25)$$

e

$$P(X \geq 3) =$$



Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.

3. In prossimità dell'orario di chiusura sono presenti 10 visitatori. Viene data loro la possibilità di visionare una sola opera. Si calcoli la probabilità che almeno 3 visitatori scelgano l'opera a , assumendo che la scelta dell'opera venga effettuata dai visitatori in maniera casuale e indipendente.

Sia X la variabile casuale 'n. di visitatori che scelgono l'opera a tra i 10 presenti'. Abbiamo

$$X \sim \text{Bin}(n = 10, p = 0.25)$$

e

$$P(X \geq 3) = 1 - \{P(X = 0) + P(X = 1) + P(X = 2)\} =$$

=



3. In prossimità dell'orario di chiusura sono presenti 10 visitatori. Viene data loro la possibilità di visionare una sola opera. Si calcoli la probabilità che almeno 3 visitatori scelgano l'opera a , assumendo che la scelta dell'opera venga effettuata dai visitatori in maniera casuale e indipendente.

Sia X la variabile casuale 'n. di visitatori che scelgono l'opera a tra i 10 presenti'. Abbiamo

$$X \sim \text{Bin}(n = 10, p = 0.25)$$

e

$$\begin{aligned} P(X \geq 3) &= 1 - \{P(X = 0) + P(X = 1) + P(X = 2)\} = \\ &= 1 - \left\{ \binom{10}{0} 0.25^0 \cdot 0.75^{10} + \binom{10}{1} 0.25^1 \cdot 0.75^9 + \right. \\ &= \qquad \qquad \qquad \left. + \binom{10}{2} 0.25^2 \cdot 0.75^8 \right\} = \\ &= \end{aligned}$$



3. In prossimità dell'orario di chiusura sono presenti 10 visitatori. Viene data loro la possibilità di visionare una sola opera. Si calcoli la probabilità che almeno 3 visitatori scelgano l'opera a , assumendo che la scelta dell'opera venga effettuata dai visitatori in maniera casuale e indipendente.

Sia X la variabile casuale 'n. di visitatori che scelgono l'opera a tra i 10 presenti'. Abbiamo

$$X \sim \text{Bin}(n = 10, p = 0.25)$$

e

$$\begin{aligned} P(X \geq 3) &= 1 - \{P(X = 0) + P(X = 1) + P(X = 2)\} = \\ &= 1 - \left\{ \binom{10}{0} 0.25^0 \cdot 0.75^{10} + \binom{10}{1} 0.25^1 \cdot 0.75^9 + \right. \\ &= \qquad \qquad \qquad \left. + \binom{10}{2} 0.25^2 \cdot 0.75^8 \right\} = \\ &= 1 - (0.05631 + 0.18771 + 0.28157) = \\ &= \end{aligned}$$



3. In prossimità dell'orario di chiusura sono presenti 10 visitatori. Viene data loro la possibilità di visionare una sola opera. Si calcoli la probabilità che almeno 3 visitatori scelgano l'opera a , assumendo che la scelta dell'opera venga effettuata dai visitatori in maniera casuale e indipendente.

Sia X la variabile casuale 'n. di visitatori che scelgono l'opera a tra i 10 presenti'. Abbiamo

$$X \sim \text{Bin}(n = 10, p = 0.25)$$

e

$$\begin{aligned} P(X \geq 3) &= 1 - \{P(X = 0) + P(X = 1) + P(X = 2)\} = \\ &= 1 - \left\{ \binom{10}{0} 0.25^0 \cdot 0.75^{10} + \binom{10}{1} 0.25^1 \cdot 0.75^9 + \right. \\ &= \qquad \qquad \qquad \left. + \binom{10}{2} 0.25^2 \cdot 0.75^8 \right\} = \\ &= 1 - (0.05631 + 0.18771 + 0.28157) = \\ &= 1 - 0.52559 = 0.47441. \end{aligned}$$

Tema T3051

Prova scritta di STATISTICA del 29.08.2019

Prova scritta di
STATISTICA del
29.08.2019

Giuseppe Boari
Gabriele Cantaluppi



UNIVERSITÀ
CATTOLICA
del Sacro Cuore

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.

Giuseppe Boari
Gabriele Cantaluppi
Università Cattolica del Sacro Cuore, Milano

Esercizio 1

Il prospetto seguente riporta la distribuzione congiunta del carattere qualitativo X e del carattere quantitativo Y .

$X \setminus Y$	$1 \div 3$	$3 \div 7$	$7 \div 9$
x_1	12		
x_2		4	2

1. Si completi la distribuzione delle frequenze congiunte, sapendo che il numero totale di unità statistiche è 20, che $\text{Var}(Y|x_1) = 0$ e che la distribuzione $Y|x_2$ è simmetrica.



Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.

Esercizio 1

Il prospetto seguente riporta la distribuzione congiunta del carattere qualitativo X e del carattere quantitativo Y .

$X \setminus Y$	1 + 3	3 + 7	7 + 9	
x_1	12			
x_2		4	2	

1. Si completi la distribuzione delle frequenze congiunte, sapendo che il numero totale di unità statistiche è 20, che $\text{Var}(Y|x_1) = 0$ e che la distribuzione $Y|x_2$ è simmetrica.

$\text{Var}(Y|x_1) = 0$ implica che la distribuzione condizionata $Y|x_1$ è degenera e, quindi, $n_{12} = n_{13} = 0$.

$n_{21} = 2$ realizza la simmetria della distribuzione $Y|x_2$.

$X \setminus Y$	1 + 3	3 + 7	7 + 9	
x_1	12	0	0	12
x_2	2	4	2	8
	14	4	2	20



Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.



Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.

2. Si rappresentino graficamente la distribuzione marginale X e la distribuzione condizionata $Y|_{x_2}$.



Si considera la distribuzione marginale della variabile X , ottenuta sommando le frequenze di riga

x_i	$n_{i\bullet}$
x_1	12
x_2	8

Si tratta di un carattere qualitativo sconnesso che può essere rappresentato graficamente mediante un grafico a bastoncini o a canne d'organo.

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.

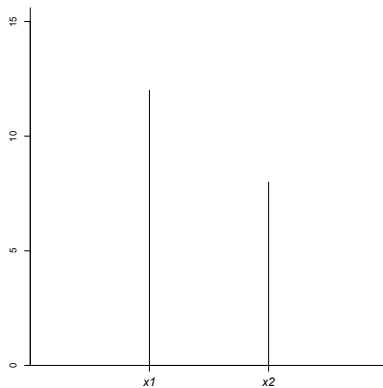


Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.





Si considera la distribuzione condizionata $Y|x_2$

y_j	n_{2j}
1 + 3	2
3 + 7	4
7 + 9	2

Esercizio 1.

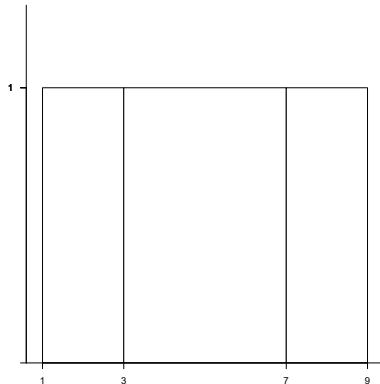
Esercizio 2.

Esercizio 3.

Esercizio 4.

Si tratta di un carattere quantitativo con modalità raggruppate in classi che può essere rappresentato graficamente mediante un istogramma. L'istogramma può essere costruito dopo avere calcolato le densità $d_{2j} = \frac{n_{2j}}{a_j}$, (dove a_j sono le ampiezze di classe).

y_j	n_{2j}	a_{2j}	d_{2j}
1 - 3	2	3 - 1 = 2	1
3 - 7	4	7 - 3 = 4	1
7 - 9	2	9 - 7 = 2	1



Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.



3. Si confrontino con opportuni indici di posizione e variabilità le distribuzioni condizionate $Y|x_1$ e $Y|x_2$.

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.



Si considera la distribuzione condizionata $Y|x_1$

y_j	n_{1j}
1 $\rightarrow$ 3	12
3 $\rightarrow$ 7	0
7 $\rightarrow$ 9	0

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.



Si tratta di un carattere quantitativo con modalità raggruppate in classi.

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.



Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.

Si osserva che la distribuzione è degenera.

$M(Y|x_1) = 2$. La varianza e il coefficiente di variazione sono nulli.



Si considera la distribuzione condizionata $Y|x_2$

y_j	n_{2j}
1 + 3	2
3 + 7	4
7 + 9	2

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.



Si tratta di un carattere quantitativo con modalità raggruppate in classi.

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.



Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.

Si osserva che la distribuzione è simmetrica rispetto a 5; la mediana (percentile di ordine 0.50) risulta $y_{0.5} = 5$.



La media aritmetica

$$M(Y|x_2) = \frac{1}{n_{2\bullet}} \sum_{j=1}^k y_j n_{2j}$$

risulta 5.

Si osserva, infatti, che la distribuzione è simmetrica.

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.



La varianza

$$\text{Var}(Y|x_2) = M(Y^2|x_2) - [M(Y|x_2)]^2$$

risulta 4.5.

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.



La varianza

$$Var(Y|x_2) = M(Y^2|x_2) - [M(Y|x_2)]^2$$

risulta 4.5. Si ottiene (considerando come modalità i valori centrali delle classi)

y_j	n_{2j}	$y_j n_{2j}$	$y_j^2 n_{2j}$
2	2	4	8
5	4	20	100
8	2	16	128
	8	40	236

da cui segue

$$Var(Y|x_2) = \frac{236}{8} - 5^2 = 29.5 - 25 = 4.5.$$



Per effettuare confronti in termini di variabilità tra caratteri quantitativi occorre utilizzare una misura relativa, ad esempio il coefficiente di variazione

$$CV = \frac{\sigma}{\mu} = \frac{\sqrt{\text{Var}(Y|x_2)}}{M(Y|x_2)} = \frac{\sqrt{4.5}}{5} = \frac{2.1213}{5} = 0.4243.$$

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.



$Y|x_1$ Indice di posizione: $M(Y|x_1)$ valore 2 Indice di variabilità: $CV(Y|x_1)$ valore 0
 $Y|x_2$ Indice di posizione: $M(Y|x_2)$ valore 5 Indice di variabilità: $CV(Y|x_2)$ valore 0.4243
La distribuzione condizionata $Y|x_1$ è caratterizzata da un livello medio e da variabilità inferiori rispetto alla condizionata $Y|x_2$.

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.



4. Si determini il grado di connessione tra le variabili X e Y mediante il calcolo di un opportuno indice.

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.



Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.

Dalla distribuzione delle frequenze congiunte n_{ij} :

	1 + 3	3 + 7	7 + 9	$n_{i\bullet}$
x1	12	0	0	12
x2	2	4	2	8
$n_{\bullet j}$	14	4	2	20

è possibile ottenere le frequenze teoriche nella situazione di indipendenza stocastica

$$\hat{n}_{ij} = \frac{n_{i\bullet} n_{\bullet j}}{n}:$$



Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.

Dalla distribuzione delle frequenze congiunte n_{ij} :

	1 + 3	3 + 7	7 + 9	$n_{i\bullet}$
x1	12	0	0	12
x2	2	4	2	8
$n_{\bullet j}$	14	4	2	20

è possibile ottenere le frequenze teoriche nella situazione di indipendenza stocastica

$$\hat{n}_{ij} = \frac{n_{i\bullet} \cdot n_{\bullet j}}{n}:$$

	1 + 3	3 + 7	7 + 9	$n_{i\bullet}$
x1	8.4	2.4	1.2	12
x2	5.6	1.6	0.8	8
$n_{\bullet j}$	14.0	4.0	2.0	20



le contingenze assolute $n_{ij} - \hat{n}_{ij}$:

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.



le contingenze assolute $n_{ij} - \hat{n}_{ij}$:

	1 + 3	3 + 7	7 + 9
x1	3.6	-2.4	-1.2
x2	-3.6	2.4	1.2

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.



i rapporti di contingenza $\frac{n_{ij}}{\hat{n}_{ij}}$:

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.



i rapporti di contingenza $\frac{n_{ij}}{\hat{n}_{ij}}$:

	1 + 3	3 + 7	7 + 9
x1	1.4286	0.0	0.0
x2	0.3571	2.5	2.5

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.



e le contingenze relative, espresse in termini percentuali, $100 \left(\frac{n_{ij} - \hat{n}_{ij}}{\hat{n}_{ij}} \right) \% = 100 \left(\frac{n_{ij}}{\hat{n}_{ij}} - 1 \right) \%$:

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.



e le contingenze relative, espresse in termini percentuali, $100 \left(\frac{n_{ij} - \hat{n}_{ij}}{\hat{n}_{ij}} \right) \% = 100 \left(\frac{n_{ij}}{\hat{n}_{ij}} - 1 \right) \%$:

	1 + 3	3 + 7	7 + 9
x1	42.86	-100	-100
x2	-64.29	150	150

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.



Risultati intermedi $\frac{(n_{ij} - \hat{n}_{ij})^2}{\hat{n}_{ij}}$ per il calcolo del χ^2

	1 + 3	3 + 7	7 + 9
x1	1.5429	2.4	1.2
x2	2.3143	3.6	1.8

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.



Risultati intermedi $\frac{n_{ij}^2}{n_{i\bullet} \cdot n_{\bullet j}}$ per il calcolo del χ^2 (formula operativa)

	1 + 3	3 + 7	7 + 9
x1	0.8571	0.0	0.00
x2	0.0357	0.5	0.25

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.



Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.

$$\chi^2 = \sum_{i=1}^h \sum_{j=1}^k \frac{(n_{ij} - \hat{n}_{ij})^2}{\hat{n}_{ij}} = 12.8571,$$

$$\chi_N^2 = \frac{\chi^2}{n \cdot \min(h-1, k-1)} = 0.6429$$

$$\chi_N = \sqrt{\chi_N^2} = 0.8018.$$

Sussiste, quindi, un livello di connessione alto.



Si consideri il seguente prospetto:

t	x_t	$NIBF(base = 2013)$	$NIBM$
2012			
2013	24	1	0.5
2014	12		
2015			0.5
2016		0.5	

1. Si completi il prospetto con i dati mancanti.

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.

Abbiamo:

$$2012/2013 =$$



Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.



Abbiamo:

$${}_{2012}I_{2013} = \frac{{}_{2013}I_{2013}}{{}_{2013}I_{2012}} = \frac{x_{2013}}{x_{2012}} = 0.5 \rightarrow x_{2012} =$$

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.



Abbiamo:

$${}_{2012}I_{2013} = \frac{{}_{2013}I_{2013}}{{}_{2013}I_{2012}} = \frac{x_{2013}}{x_{2012}} = 0.5 \rightarrow x_{2012} = \frac{24}{0.5} = 48 \text{ e}$$

$${}_{2013}I_{2012} =$$

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.



Abbiamo:

$${}_{2012}I_{2013} = \frac{{}_{2013}I_{2013}}{{}_{2013}I_{2012}} = \frac{x_{2013}}{x_{2012}} = 0.5 \rightarrow x_{2012} = \frac{24}{0.5} = 48 \text{ e}$$

$${}_{2013}I_{2012} = \frac{1}{{}_{2012}I_{2013}} = \frac{1}{0.5} = 2$$

$${}_{2013}I_{2014} =$$

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.



Abbiamo:

$${}_{2012}I_{2013} = \frac{{}_{2013}I_{2013}}{{}_{2013}I_{2012}} = \frac{x_{2013}}{x_{2012}} = 0.5 \rightarrow x_{2012} = \frac{24}{0.5} = 48 \text{ e}$$

$${}_{2013}I_{2012} = \frac{1}{{}_{2012}I_{2013}} = \frac{1}{0.5} = 2$$

$${}_{2013}I_{2014} = \frac{x_{2014}}{x_{2013}} = \frac{12}{24} = 0.5$$

$${}_{2014}I_{2015} =$$

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.



Abbiamo:

$${}_{2012}l_{2013} = \frac{{}_{2013}l_{2013}}{{}_{2013}l_{2012}} = \frac{x_{2013}}{x_{2012}} = 0.5 \rightarrow x_{2012} = \frac{24}{0.5} = 48 \text{ e}$$

$${}_{2013}l_{2012} = \frac{1}{{}_{2012}l_{2013}} = \frac{1}{0.5} = 2$$

$${}_{2013}l_{2014} = \frac{x_{2014}}{x_{2013}} = \frac{12}{24} = 0.5$$

$${}_{2014}l_{2015} = \frac{{}_{2013}l_{2015}}{{}_{2013}l_{2014}} = \frac{x_{2015}}{x_{2014}} = 0.5 \rightarrow x_{2015} =$$

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.



Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.

Abbiamo:

$${}_{2012}l_{2013} = \frac{{}_{2013}l_{2013}}{{}_{2013}l_{2012}} = \frac{x_{2013}}{x_{2012}} = 0.5 \rightarrow x_{2012} = \frac{24}{0.5} = 48 \text{ e}$$

$${}_{2013}l_{2012} = \frac{1}{{}_{2012}l_{2013}} = \frac{1}{0.5} = 2$$

$${}_{2013}l_{2014} = \frac{x_{2014}}{x_{2013}} = \frac{12}{24} = 0.5$$

$${}_{2014}l_{2015} = \frac{{}_{2013}l_{2015}}{{}_{2013}l_{2014}} = \frac{x_{2015}}{x_{2014}} = 0.5 \rightarrow x_{2015} = 12 \cdot 0.5 = 6 \text{ e}$$

$${}_{2013}l_{2015} =$$



Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.

Abbiamo:

$${}_{2012}l_{2013} = \frac{{}_{2013}l_{2013}}{{}_{2013}l_{2012}} = \frac{x_{2013}}{x_{2012}} = 0.5 \rightarrow x_{2012} = \frac{24}{0.5} = 48 \text{ e}$$

$${}_{2013}l_{2012} = \frac{1}{{}_{2012}l_{2013}} = \frac{1}{0.5} = 2$$

$${}_{2013}l_{2014} = \frac{x_{2014}}{x_{2013}} = \frac{12}{24} = 0.5$$

$${}_{2014}l_{2015} = \frac{{}_{2013}l_{2015}}{{}_{2013}l_{2014}} = \frac{x_{2015}}{x_{2014}} = 0.5 \rightarrow x_{2015} = 12 \cdot 0.5 = 6 \text{ e}$$

$${}_{2013}l_{2015} = {}_{2013}l_{2014} \cdot 0.5 = 0.5 \cdot 0.5 = 0.25$$

$${}_{2013}l_{2016} =$$



Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.

Abbiamo:

$${}_{2012}l_{2013} = \frac{{}_{2013}l_{2013}}{{}_{2013}l_{2012}} = \frac{x_{2013}}{x_{2012}} = 0.5 \rightarrow x_{2012} = \frac{24}{0.5} = 48 \text{ e}$$

$${}_{2013}l_{2012} = \frac{1}{{}_{2012}l_{2013}} = \frac{1}{0.5} = 2$$

$${}_{2013}l_{2014} = \frac{x_{2014}}{x_{2013}} = \frac{12}{24} = 0.5$$

$${}_{2014}l_{2015} = \frac{{}_{2013}l_{2015}}{{}_{2013}l_{2014}} = \frac{x_{2015}}{x_{2014}} = 0.5 \rightarrow x_{2015} = 12 \cdot 0.5 = 6 \text{ e}$$

$${}_{2013}l_{2015} = {}_{2013}l_{2014} \cdot 0.5 = 0.5 \cdot 0.5 = 0.25$$

$${}_{2013}l_{2016} = \frac{x_{2016}}{x_{2013}} = 0.5 \rightarrow x_{2016} =$$



Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.

Abbiamo:

$${}_{2012}l_{2013} = \frac{{}_{2013}l_{2013}}{{}_{2013}l_{2012}} = \frac{x_{2013}}{x_{2012}} = 0.5 \rightarrow x_{2012} = \frac{24}{0.5} = 48 \text{ e}$$

$${}_{2013}l_{2012} = \frac{1}{{}_{2012}l_{2013}} = \frac{1}{0.5} = 2$$

$${}_{2013}l_{2014} = \frac{x_{2014}}{x_{2013}} = \frac{12}{24} = 0.5$$

$${}_{2014}l_{2015} = \frac{{}_{2013}l_{2015}}{{}_{2013}l_{2014}} = \frac{x_{2015}}{x_{2014}} = 0.5 \rightarrow x_{2015} = 12 \cdot 0.5 = 6 \text{ e}$$

$${}_{2013}l_{2015} = {}_{2013}l_{2014} \cdot 0.5 = 0.5 \cdot 0.5 = 0.25$$

$${}_{2013}l_{2016} = \frac{x_{2016}}{x_{2013}} = 0.5 \rightarrow x_{2016} = x_{2013} \cdot 0.5 = 12$$

$${}_{2015}l_{2016} =$$



Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.

Abbiamo:

$${}_{2012}l_{2013} = \frac{{}_{2013}l_{2013}}{{}_{2013}l_{2012}} = \frac{x_{2013}}{x_{2012}} = 0.5 \rightarrow x_{2012} = \frac{24}{0.5} = 48 \text{ e}$$

$${}_{2013}l_{2012} = \frac{1}{{}_{2012}l_{2013}} = \frac{1}{0.5} = 2$$

$${}_{2013}l_{2014} = \frac{x_{2014}}{x_{2013}} = \frac{12}{24} = 0.5$$

$${}_{2014}l_{2015} = \frac{{}_{2013}l_{2015}}{{}_{2013}l_{2014}} = \frac{x_{2015}}{x_{2014}} = 0.5 \rightarrow x_{2015} = 12 \cdot 0.5 = 6 \text{ e}$$

$${}_{2013}l_{2015} = {}_{2013}l_{2014} \cdot 0.5 = 0.5 \cdot 0.5 = 0.25$$

$${}_{2013}l_{2016} = \frac{x_{2016}}{x_{2013}} = 0.5 \rightarrow x_{2016} = x_{2013} \cdot 0.5 = 12$$

$${}_{2015}l_{2016} = \frac{x_{2016}}{x_{2015}} =$$



Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.

Abbiamo:

$${}_{2012}l_{2013} = \frac{{}_{2013}l_{2013}}{{}_{2013}l_{2012}} = \frac{x_{2013}}{x_{2012}} = 0.5 \rightarrow x_{2012} = \frac{24}{0.5} = 48 \text{ e}$$

$${}_{2013}l_{2012} = \frac{1}{{}_{2012}l_{2013}} = \frac{1}{0.5} = 2$$

$${}_{2013}l_{2014} = \frac{x_{2014}}{x_{2013}} = \frac{12}{24} = 0.5$$

$${}_{2014}l_{2015} = \frac{{}_{2013}l_{2015}}{{}_{2013}l_{2014}} = \frac{x_{2015}}{x_{2014}} = 0.5 \rightarrow x_{2015} = 12 \cdot 0.5 = 6 \text{ e}$$

$${}_{2013}l_{2015} = {}_{2013}l_{2014} \cdot 0.5 = 0.5 \cdot 0.5 = 0.25$$

$${}_{2013}l_{2016} = \frac{x_{2016}}{x_{2013}} = 0.5 \rightarrow x_{2016} = x_{2013} \cdot 0.5 = 12$$

$${}_{2015}l_{2016} = \frac{x_{2016}}{x_{2015}} = \frac{12}{6} = 2$$



Si ottiene, quindi

t	x_t	$NIBF(base = 2013)$	$NIBM$
2012	48	2	—
2013	24	1	0.5
2014	12	0.5	0.5
2015	6	0.25	0.5
2016	12	0.5	2

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.



2. Si calcoli il numero indice al tempo 2013 con base 2015 e si commenti il risultato ottenuto.

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.



2. Si calcoli il numero indice al tempo 2013 con base 2015 e si commenti il risultato ottenuto.

$${}_{2015}I_{2013} = \frac{x_{2013}}{x_{2015}} = \frac{24}{6} = 4$$

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.



2. Si calcoli il numero indice al tempo 2013 con base 2015 e si commenti il risultato ottenuto.

$${}_{2015}I_{2013} = \frac{x_{2013}}{x_{2015}} = \frac{24}{6} = 4$$

La grandezza al tempo 2013 era superiore del 300% rispetto alla grandezza al tempo 2015. Ovvero il valore della grandezza nel 2013 era il 400% del valore registrato nel 2015.

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.



3. Si calcoli il tasso medio di variazione dal 2013 al 2015 e si commenti il risultato ottenuto.

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.



3. Si calcoli il tasso medio di variazione dal 2013 al 2015 e si commenti il risultato ottenuto.
Si ricorda che

$${}_{2013}I_{2015} = {}_{2013}I_{2014} \cdot {}_{2014}I_{2015}$$

Si deve trovare il numero indice a base mobile medio.

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.



3. Si calcoli il tasso medio di variazione dal 2013 al 2015 e si commenti il risultato ottenuto. Si ricorda che

$${}_{2013}I_{2015} = {}_{2013}I_{2014} \cdot {}_{2014}I_{2015}$$

Si deve trovare il numero indice a base mobile medio. Applicando il criterio di scelta della media secondo Chisini si ottiene

$${}_{2013}I_{2015} = {}_{2013}I_{2014} \cdot {}_{2014}I_{2015} = \alpha \cdot \alpha = \alpha^2.$$

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.



Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.

3. Si calcoli il tasso medio di variazione dal 2013 al 2015 e si commenti il risultato ottenuto. Si ricorda che

$${}_{2013}I_{2015} = {}_{2013}I_{2014} \cdot {}_{2014}I_{2015}$$

Si deve trovare il numero indice a base mobile medio. Applicando il criterio di scelta della media secondo Chisini si ottiene

$${}_{2013}I_{2015} = {}_{2013}I_{2014} \cdot {}_{2014}I_{2015} = \alpha \cdot \alpha = \alpha^2.$$

Quindi

$$\alpha = \sqrt[2]{{}_{2013}I_{2015}} = \sqrt[2]{{}_{2013}I_{2014} \cdot {}_{2014}I_{2015}} = 0.5.$$

Da cui segue una variazione relativa media pari a -0.5 , corrispondente a un decremento medio annuo tra il 2013 e il 2015 del 50%.



Il seguente prospetto riporta i valori rilevati in corrispondenza a 5 unità statistiche con riferimento a due caratteri quantitativi X e Y tranne che per l'osservazione y_3

x_i	3	4	4	4	5
y_i	4	2		6	4

1. Si rappresentino graficamente le coppie dei dati, con anche il valore di y_3 tale per cui $\eta_{Y|X}^2 = \eta_{X|Y}^2$. Si completi il grafico con la funzione di regressione ipotizzando Y in funzione di X .

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.



Il seguente prospetto riporta i valori rilevati in corrispondenza a 5 unità statistiche con riferimento a due caratteri quantitativi X e Y tranne che per l'osservazione y_3

x_i	3	4	4	4	5
y_i	4	2		6	4

1. Si rappresentino graficamente le coppie dei dati, con anche il valore di y_3 tale per cui $\eta_{Y|X}^2 = \eta_{X|Y}^2$. Si completi il grafico con la funzione di regressione ipotizzando Y in funzione di X .

È agevole realizzare la condizione $\eta_{Y|X}^2 = \eta_{X|Y}^2 = 0$ fissando $y_3 = 4$.

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.



La funzione di regressione associa a ogni valore x appartenente all'insieme di definizione della variabile X la media di Y condizionata a x , $M(Y|x)$

x	$M(Y x)$
3	4
4	$\frac{1}{3}(2 + 4 + 6) = 4$
5	4

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.



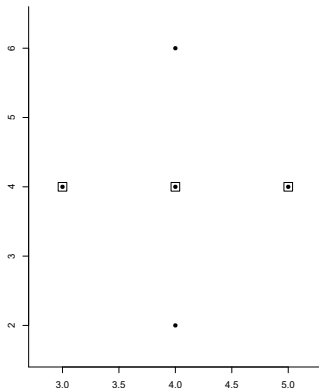
Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.

Le coppie di dati possono essere rappresentate mediante un grafico a dispersione.





2. Si calcolino, secondo il metodo dei minimi quadrati, i parametri e l'indice di adattamento dei seguenti modelli di regressione:

$$I) Y^* = a + b \cdot X \qquad II) Y^* = c$$

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.



Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.

Si riportano i calcoli necessari¹ per ottenere le stime dei parametri e gli indici di adattamento del modello

$$Y = a + bX + E$$

x_i	y_i	$x_i y_i$	x_i^2	y_i^2
3	4	12	9	16
4	2	8	16	4
4	4	16	16	16
4	6	24	16	36
5	4	20	25	16
20	20	80	82	88
4	4	16	16.4	17.6
$M(X)$	$M(Y)$	$M(XY)$	$M(X^2)$	$M(Y^2)$

$$\hat{b} =$$

¹Si osserva che per ragioni di stampa alcuni calcoli potrebbero essere duplicati con riferimento ai diversi modelli oggetto di studio



Si riportano i calcoli necessari¹ per ottenere le stime dei parametri e gli indici di adattamento del modello

$$Y = a + bX + E$$

x_i	y_i	$x_i y_i$	x_i^2	y_i^2
3	4	12	9	16
4	2	8	16	4
4	4	16	16	16
4	6	24	16	36
5	4	20	25	16
20	20	80	82	88
4	4	16	16.4	17.6
$M(X)$	$M(Y)$	$M(XY)$	$M(X^2)$	$M(Y^2)$

$$\begin{aligned}\hat{b} &= \frac{\text{Cov}(X, Y)}{\text{Var}(X)} = \frac{M(XY) - M(X)M(Y)}{M(X^2) - [M(X)]^2} = \\ &= \frac{16 - 4 \cdot 4}{16.4 - 4^2} = \frac{0}{0.4} = 0 \\ \hat{a} &= \end{aligned}$$

¹Si osserva che per ragioni di stampa alcuni calcoli potrebbero essere duplicati con riferimento ai diversi modelli oggetto di studio



Si riportano i calcoli necessari¹ per ottenere le stime dei parametri e gli indici di adattamento del modello

$$Y = a + bX + E$$

x_i	y_i	$x_i y_i$	x_i^2	y_i^2
3	4	12	9	16
4	2	8	16	4
4	4	16	16	16
4	6	24	16	36
5	4	20	25	16
20	20	80	82	88
4	4	16	16.4	17.6
$M(X)$	$M(Y)$	$M(XY)$	$M(X^2)$	$M(Y^2)$

$$\hat{b} = \frac{\text{Cov}(X, Y)}{\text{Var}(X)} = \frac{M(XY) - M(X)M(Y)}{M(X^2) - [M(X)]^2} =$$

$$= \frac{16 - 4 \cdot 4}{16.4 - 4^2} = \frac{0}{0.4} = 0$$

$$\hat{a} = M(Y) - \hat{b}M(X) = 4$$

$$\rho^2 =$$

¹Si osserva che per ragioni di stampa alcuni calcoli potrebbero essere duplicati con riferimento ai diversi modelli oggetto di studio



Si riportano i calcoli necessari¹ per ottenere le stime dei parametri e gli indici di adattamento del modello

$$Y = a + bX + E$$

x_i	y_i	$x_i y_i$	x_i^2	y_i^2
3	4	12	9	16
4	2	8	16	4
4	4	16	16	16
4	6	24	16	36
5	4	20	25	16
20	20	80	82	88
4	4	16	16.4	17.6
$M(X)$	$M(Y)$	$M(XY)$	$M(X^2)$	$M(Y^2)$

$$\hat{b} = \frac{\text{Cov}(X, Y)}{\text{Var}(X)} = \frac{M(XY) - M(X)M(Y)}{M(X^2) - [M(X)]^2} =$$

$$= \frac{16 - 4 \cdot 4}{16.4 - 4^2} = \frac{0}{0.4} = 0$$

$$\hat{a} = M(Y) - \hat{b}M(X) = 4$$

$$\rho^2 = \frac{[\text{Cov}(X, Y)]^2}{\text{Var}(X) \text{Var}(Y)} = \frac{0^2}{0.4 \cdot 1.6} = 0$$

¹Si osserva che per ragioni di stampa alcuni calcoli potrebbero essere duplicati con riferimento ai diversi modelli oggetto di studio



Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.

Si riportano i calcoli necessari² per ottenere le stime dei parametri e gli indici di adattamento del modello

$$Y = c + E$$

x_i	y_i	y_i^2
3	4	16
4	2	4
4	4	16
4	6	36
5	4	16
20		88
4		17.6
$M(Y)$		$M(Y^2)$

$$\hat{c} =$$

²Si osserva che per ragioni di stampa alcuni calcoli potrebbero essere duplicati con riferimento ai diversi modelli oggetto di studio



Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.

Si riportano i calcoli necessari² per ottenere le stime dei parametri e gli indici di adattamento del modello

$$Y = c + E$$

x_i	y_i	y_i^2
3	4	16
4	2	4
4	4	16
4	6	36
5	4	16
20		88
4		17.6
$M(Y)$		$M(Y^2)$

$$\hat{c} = M(Y) = 4$$

$$R^2 =$$

²Si osserva che per ragioni di stampa alcuni calcoli potrebbero essere duplicati con riferimento ai diversi modelli oggetto di studio



Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.

Si riportano i calcoli necessari² per ottenere le stime dei parametri e gli indici di adattamento del modello

$$Y = c + E$$

x_i	y_i	y_i^2
3	4	16
4	2	4
4	4	16
4	6	36
5	4	16
20		88
4		17.6
$M(Y)$		$M(Y^2)$

$$\hat{c} = M(Y) = 4$$

$$R^2 = 0$$

²Si osserva che per ragioni di stampa alcuni calcoli potrebbero essere duplicati con riferimento ai diversi modelli oggetto di studio



$\hat{a} = 4$, $\hat{b} = 0$, indice adattamento $\rho^2 = 0$ $\hat{c} = 4$, indice adattamento $R^2 = 0$

Dal momento che sussiste indipendenza in media di Y da X , essendo $\eta_{Y|X}^2 = 0$, abbiamo anche indipendenza lineare, $\rho^2 = \text{Cov}(X, Y) = b = 0$ e i due modelli risultano equivalenti con $a = c = M(Y)$.

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.

Esercizio 4

Prova scritta di
STATISTICA del
29.08.2019

Giuseppe Boari
Gabriele Cantaluppi



UNIVERSITÀ
CATTOLICA
del Sacro Cuore

1. Si consideri l'allestimento di una sala quadrata (con 4 pareti) con 4 quadri: 2 di Monet e 2 di Chagall.

Si calcoli la probabilità che, disponendo in maniera casuale un quadro su ciascuna delle 4 pareti, i due quadri di Monet figurino su 2 pareti opposte.

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.

Esercizio 4

Prova scritta di
STATISTICA del
29.08.2019

Giuseppe Boari
Gabriele Cantaluppi



UNIVERSITÀ
CATTOLICA
del Sacro Cuore

1. Si consideri l'allestimento di una sala quadrata (con 4 pareti) con 4 quadri: 2 di Monet e 2 di Chagall.

Si calcoli la probabilità che, disponendo in maniera casuale un quadro su ciascuna delle 4 pareti, i due quadri di Monet figurino su 2 pareti opposte.

Il problema può essere risolto in maniera intuitiva mediante la seguente riformulazione.

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.

Esercizio 4



Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.

1. Si consideri l'allestimento di una sala quadrata (con 4 pareti) con 4 quadri: 2 di Monet e 2 di Chagall.

Si calcoli la probabilità che, disponendo in maniera casuale un quadro su ciascuna delle 4 pareti, i due quadri di Monet figurino su 2 pareti opposte.

Il problema può essere risolto in maniera intuitiva mediante la seguente riformulazione. Scelta in maniera casuale una parete per una delle quattro opere, ad esempio una di Monet, qual è la probabilità che la seconda opera di Monet figuri sulla parete opposta?



1. Si consideri l'allestimento di una sala quadrata (con 4 pareti) con 4 quadri: 2 di Monet e 2 di Chagall.

Si calcoli la probabilità che, disponendo in maniera casuale un quadro su ciascuna delle 4 pareti, i due quadri di Monet figurino su 2 pareti opposte.

Il problema può essere risolto in maniera intuitiva mediante la seguente riformulazione. Scelta in maniera casuale una parete per una delle quattro opere, ad esempio una di Monet, qual è la probabilità che la seconda opera di Monet figuri sulla parete opposta?

La seconda opera può essere posizionata su una delle 3 pareti rimanenti, e solo una di queste risulta di fronte a quella inizialmente scelta. Abbiamo, quindi:



1. Si consideri l'allestimento di una sala quadrata (con 4 pareti) con 4 quadri: 2 di Monet e 2 di Chagall.

Si calcoli la probabilità che, disponendo in maniera casuale un quadro su ciascuna delle 4 pareti, i due quadri di Monet figurino su 2 pareti opposte.

Il problema può essere risolto in maniera intuitiva mediante la seguente riformulazione. Scelta in maniera casuale una parete per una delle quattro opere, ad esempio una di Monet, qual è la probabilità che la seconda opera di Monet figuri sulla parete opposta?

La seconda opera può essere posizionata su una delle 3 pareti rimanenti, e solo una di queste risulta di fronte a quella inizialmente scelta. Abbiamo, quindi:

$$P(\text{i due quadri di Monet figurino su 2 pareti opposte}) = \frac{1}{3}.$$

Il problema può essere risolto anche facendo riferimento alla probabilità condizionata o al calcolo combinatorio.



Se si considerano le 4 opere distinte, indicati con M_1 e M_2 i quadri di Monet e con C_1 e C_2 i quadri di Chagall, vi sono 8 possibili sequenze di opere che soddisfano la richiesta:

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.



Se si considerano le 4 opere distinte, indicati con M_1 e M_2 i quadri di Monet e con C_1 e C_2 i quadri di Chagall, vi sono 8 possibili sequenze di opere che soddisfano la richiesta:

$$\{M_1, C_1, M_2, C_2\}$$

$$\{M_2, C_1, M_1, C_2\}$$

$$\{M_1, C_2, M_2, C_1\}$$

$$\{M_2, C_2, M_1, C_1\}$$

$$\{C_1, M_1, C_2, M_2\}$$

$$\{C_2, M_1, C_1, M_2\}$$

$$\{C_1, M_2, C_2, M_1\}$$

$$\{C_2, M_2, C_1, M_1\}$$

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.



Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.

La probabilità della generica sequenza risulta:

$$\begin{aligned} P(M_1 \cap C_1 \cap M_2 \cap C_2) &= P(M_1) \cdot P(C_1|M_1) \cdot P(M_2|M_1 \cap C_1) \cdot \\ &\quad \cdot P(C_2|M_1 \cap C_1 \cap M_2) = \\ &= \frac{1}{4} \cdot \frac{1}{3} \cdot \frac{1}{2} \cdot \frac{1}{1} = \frac{1}{24}. \end{aligned}$$

Abbiamo quindi:

$$8 \cdot \frac{1}{4} \cdot \frac{1}{3} \cdot \frac{1}{2} \cdot \frac{1}{1} = \frac{2 \cdot 2 \cdot 2}{4!} = \frac{8}{24} = \frac{1}{3} = 0.3333,$$

dove $4!$ indica il numero delle sequenze che si possono costruire con i 4 quadri.



Se si considerano 2 opere di Monet e 2 di Chagall, indicate con M e C , vi sono 2 possibili sequenze di opere che soddisfano la richiesta

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.



Se si considerano 2 opere di Monet e 2 di Chagall, indicate con M e C , vi sono 2 possibili sequenze di opere che soddisfano la richiesta

$$\begin{aligned} &\{M, C, M, C\} \\ &\{C, M, C, M\} \end{aligned}$$

La probabilità della generica sequenza risulta:

$$\begin{aligned} P(M \cap C \cap M \cap C) &= P(M) \cdot P(C|M) \cdot P(M|M \cap C) \cdot P(C|M \cap C \cap M) = \\ &= \frac{2}{4} \cdot \frac{2}{3} \cdot \frac{1}{2} \cdot \frac{1}{1} = \frac{4}{24}. \end{aligned}$$

Abbiamo quindi:

$$2 \cdot \frac{2}{4} \cdot \frac{2}{3} \cdot \frac{1}{2} \cdot \frac{1}{1} = \frac{2}{\binom{4}{2}} = \frac{2}{6} = \frac{1}{3} = 0.3333,$$

dove $\binom{4}{2} = \frac{4!}{2!2!}$ indica il numero delle sequenze che si possono costruire con i 4 quadri di cui 2 di Monet e 2 di Chagall.

2. Si scelgono in maniera casuale 5 opere tra 8 quadri di Manet, 8 di Degas e 4 di Sisley per la composizione di un art visual tour. Si calcoli la probabilità che almeno 3 tra le 5 opere scelte siano di Manet.



Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.



2. Si scelgono in maniera casuale 5 opere tra 8 quadri di Manet, 8 di Degas e 4 di Sisley per la composizione di un art visual tour. Si calcoli la probabilità che almeno 3 tra le 5 opere scelte siano di Manet.

$$\begin{aligned}
 P(\text{almeno } 3 M) &= P(3M) + P(4M) + P(5M) = \\
 &= \binom{5}{3} \frac{8}{20} \frac{7}{19} \frac{6}{18} \frac{12}{17} \frac{11}{16} + \binom{5}{4} \frac{8}{20} \frac{7}{19} \frac{6}{18} \frac{5}{17} \frac{12}{16} + \\
 &\quad + \binom{5}{5} \frac{8}{20} \frac{7}{19} \frac{6}{18} \frac{5}{17} \frac{4}{16} = \\
 &= 0.23839 + 0.05418 + 0.003612 = 0.2962
 \end{aligned}$$



2. Si scelgono in maniera casuale 5 opere tra 8 quadri di Manet, 8 di Degas e 4 di Sisley per la composizione di un art visual tour. Si calcoli la probabilità che almeno 3 tra le 5 opere scelte siano di Manet.

$$\begin{aligned}
 P(\text{almeno } 3 M) &= P(3M) + P(4M) + P(5M) = \\
 &= \binom{5}{3} \frac{8}{20} \frac{7}{19} \frac{6}{18} \frac{12}{17} \frac{11}{16} + \binom{5}{4} \frac{8}{20} \frac{7}{19} \frac{6}{18} \frac{5}{17} \frac{12}{16} + \\
 &\quad + \binom{5}{5} \frac{8}{20} \frac{7}{19} \frac{6}{18} \frac{5}{17} \frac{4}{16} = \\
 &= 0.23839 + 0.05418 + 0.003612 = 0.2962
 \end{aligned}$$

o mediante la v.c. ipergeometrica



2. Si scelgono in maniera casuale 5 opere tra 8 quadri di Manet, 8 di Degas e 4 di Sisley per la composizione di un art visual tour. Si calcoli la probabilità che almeno 3 tra le 5 opere scelte siano di Manet.

$$\begin{aligned}
 P(\text{almeno } 3 M) &= P(3M) + P(4M) + P(5M) = \\
 &= \binom{5}{3} \frac{8}{20} \frac{7}{19} \frac{6}{18} \frac{12}{17} \frac{11}{16} + \binom{5}{4} \frac{8}{20} \frac{7}{19} \frac{6}{18} \frac{5}{17} \frac{12}{16} + \\
 &\quad + \binom{5}{5} \frac{8}{20} \frac{7}{19} \frac{6}{18} \frac{5}{17} \frac{4}{16} = \\
 &= 0.23839 + 0.05418 + 0.003612 = 0.2962
 \end{aligned}$$

o mediante la v.c. ipergeometrica

$$\begin{aligned}
 P(X \geq 3) &= P(X = 3) + P(X = 4) + P(X = 5) = \\
 &= \frac{\binom{8}{3} \binom{12}{2}}{\binom{20}{5}} + \frac{\binom{8}{4} \binom{12}{1}}{\binom{20}{5}} + \frac{\binom{8}{5} \binom{12}{0}}{\binom{20}{5}} = 0.2962.
 \end{aligned}$$



3. Il visual tour dura 3 minuti. La probabilità che un generico visitatore completi la visione del tour è pari a 0.98. Si calcoli la probabilità che tra 200 visitatori con comportamento indipendente almeno 198 completino la visione del tour.

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.



Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.

3. Il visual tour dura 3 minuti. La probabilità che un generico visitatore completi la visione del tour è pari a 0.98. Si calcoli la probabilità che tra 200 visitatori con comportamento indipendente almeno 198 completino la visione del tour.

Sia X la variabile casuale 'n. di visitatori che completano il tour tra i 200 considerati'.



Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.

3. Il visual tour dura 3 minuti. La probabilità che un generico visitatore completi la visione del tour è pari a 0.98. Si calcoli la probabilità che tra 200 visitatori con comportamento indipendente almeno 198 completino la visione del tour.

Sia X la variabile casuale 'n. di visitatori che completano il tour tra i 200 considerati'.
Abbiamo

$$X \sim \text{Bin}(n = 200, p = 0.98)$$

e

$$P(X \geq 198) =$$



3. Il visual tour dura 3 minuti. La probabilità che un generico visitatore completi la visione del tour è pari a 0.98. Si calcoli la probabilità che tra 200 visitatori con comportamento indipendente almeno 198 completino la visione del tour.

Sia X la variabile casuale 'n. di visitatori che completano il tour tra i 200 considerati'.
Abbiamo

$$X \sim \text{Bin}(n = 200, p = 0.98)$$

e

$$P(X \geq 198) = \{P(X = 198) + P(X = 199) + P(X = 200)\} =$$
$$=$$



Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.

3. Il visual tour dura 3 minuti. La probabilità che un generico visitatore completi la visione del tour è pari a 0.98. Si calcoli la probabilità che tra 200 visitatori con comportamento indipendente almeno 198 completino la visione del tour.

Sia X la variabile casuale 'n. di visitatori che completano il tour tra i 200 considerati'.
Abbiamo

$$X \sim \text{Bin}(n = 200, p = 0.98)$$

e

$$\begin{aligned} P(X \geq 198) &= \{P(X = 198) + P(X = 199) + P(X = 200)\} = \\ &= \binom{200}{198} 0.98^{198} \cdot 0.02^2 + \binom{200}{199} 0.98^{199} \cdot 0.02^1 + \\ &\quad + \binom{200}{200} 0.98^{200} \cdot 0.02^0 = \\ &= \end{aligned}$$



3. Il visual tour dura 3 minuti. La probabilità che un generico visitatore completi la visione del tour è pari a 0.98. Si calcoli la probabilità che tra 200 visitatori con comportamento indipendente almeno 198 completino la visione del tour.

Sia X la variabile casuale 'n. di visitatori che completano il tour tra i 200 considerati'.
Abbiamo

$$X \sim \text{Bin}(n = 200, p = 0.98)$$

e

$$\begin{aligned} P(X \geq 198) &= \{P(X = 198) + P(X = 199) + P(X = 200)\} = \\ &= \binom{200}{198} 0.98^{198} \cdot 0.02^2 + \binom{200}{199} 0.98^{199} \cdot 0.02^1 + \\ &\quad + \binom{200}{200} 0.98^{200} \cdot 0.02^0 = \\ &= 0.14577 + 0.07179 + 0.01759 = 0.2351. \end{aligned}$$



Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.

Si osserva che non sussistono le condizioni per approssimare la variabile casuale binomiale con la Normale.

Abbiamo, infatti, $np = 196 > 5$ ma $n(1 - p) = 4 < 5$.

Nel caso si utilizzasse l'approssimazione della variabile casuale Binomiale con la variabile casuale Normale

$$X \sim$$



Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.

Si osserva che non sussistono le condizioni per approssimare la variabile casuale binomiale con la Normale.

Abbiamo, infatti, $np = 196 > 5$ ma $n(1 - p) = 4 < 5$.

Nel caso si utilizzasse l'approssimazione della variabile casuale Binomiale con la variabile casuale Normale

$$X \sim N(\mu = 196, \sigma^2 = 3.92),$$

si otterrebbe

$$P(X > 198) =$$



Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.

Si osserva che non sussistono le condizioni per approssimare la variabile casuale binomiale con la Normale.

Abbiamo, infatti, $np = 196 > 5$ ma $n(1 - p) = 4 < 5$.

Nel caso si utilizzasse l'approssimazione della variabile casuale Binomiale con la variabile casuale Normale

$$X \sim N(\mu = 196, \sigma^2 = 3.92),$$

si otterrebbe

$$P(X > 198) = P\left(\frac{X - \mu}{\sigma} > \frac{198 - \mu}{\sigma}\right) =$$
$$=$$



Si osserva che non sussistono le condizioni per approssimare la variabile casuale binomiale con la Normale.

Abbiamo, infatti, $np = 196 > 5$ ma $n(1 - p) = 4 < 5$.

Nel caso si utilizzasse l'approssimazione della variabile casuale Binomiale con la variabile casuale Normale

$$X \sim N(\mu = 196, \sigma^2 = 3.92),$$

si otterrebbe

$$\begin{aligned} P(X > 198) &= P\left(\frac{X - \mu}{\sigma} > \frac{198 - \mu}{\sigma}\right) = \\ &= P\left(Z > \frac{198 - 196}{1.9799}\right) = \\ &= \end{aligned}$$



Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.

Si osserva che non sussistono le condizioni per approssimare la variabile casuale binomiale con la Normale.

Abbiamo, infatti, $np = 196 > 5$ ma $n(1 - p) = 4 < 5$.

Nel caso si utilizzasse l'approssimazione della variabile casuale Binomiale con la variabile casuale Normale

$$X \sim N(\mu = 196, \sigma^2 = 3.92),$$

si otterrebbe

$$\begin{aligned} P(X > 198) &= P\left(\frac{X - \mu}{\sigma} > \frac{198 - \mu}{\sigma}\right) = \\ &= P\left(Z > \frac{198 - 196}{1.9799}\right) = \\ &= P(Z > 1.0102) = \\ &= 1 - P(Z \leq 1.0102) = \\ &= 1 - 0.8438 = 0.1562. \end{aligned}$$

sottostimando di circa il 33% la probabilità corretta 0.2351.

Tema T305

Prova scritta di STATISTICA del 05.09.2019

Prova scritta di
STATISTICA del
05.09.2019

Giuseppe Boari
Gabriele Cantaluppi



UNIVERSITÀ
CATTOLICA
del Sacro Cuore

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.

Giuseppe Boari
Gabriele Cantaluppi
Università Cattolica del Sacro Cuore, Milano



Con riferimento a due classi di studenti ($x_1 = A$, $x_2 = B$) viene riportata nel seguente prospetto la distribuzione del numero di esami Y sostenuti durante il primo anno accademico:

$X \setminus Y$	4	5	6	7
x_1	30	40	40	0
x_2	0	40	50	20

[Esercizio 1.](#)

[Esercizio 2.](#)

[Esercizio 3.](#)

[Esercizio 4.](#)



Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.

1. Si rappresentino graficamente le distribuzioni condizionate di Y dato X .



Si considera la distribuzione condizionata $Y|x_1$

y_j	n_{1j}
4	30
5	40
6	40
7	0

Esercizio 1.

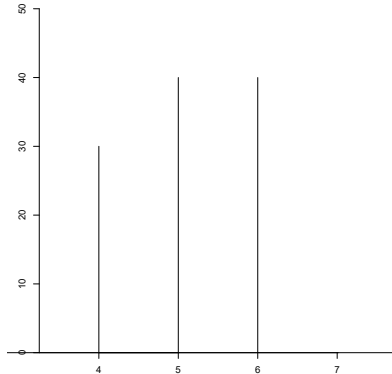
Esercizio 2.

Esercizio 3.

Esercizio 4.



Si tratta di un carattere quantitativo con modalità non raggruppate in classi che può essere rappresentato graficamente mediante un grafico a bastoncini.



Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.



Si considera la distribuzione condizionata $Y|x_2$

y_j	n_{2j}
4	0
5	40
6	50
7	20

Esercizio 1.

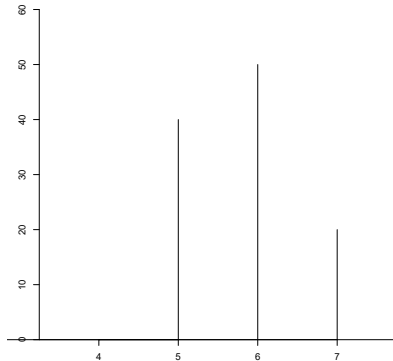
Esercizio 2.

Esercizio 3.

Esercizio 4.



Si tratta di un carattere quantitativo con modalità non raggruppate in classi che può essere rappresentato graficamente mediante un grafico a bastoncini.



Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.



Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.

2. Si confrontino le distribuzioni condizionate di Y dato X calcolando medie e varianze.



Si considera la distribuzione condizionata $Y|x_1$

y_j	n_{1j}
4	30
5	40
6	40
7	0

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.



Si tratta di un carattere quantitativo con modalità non raggruppate in classi.

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.



La media aritmetica

$$M(Y|x_1) = \frac{1}{n_{1\bullet}} \sum_{j=1}^k y_j n_{1j}$$

risulta 5.0909.

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.



La media aritmetica

$$M(Y|x_1) = \frac{1}{n_{1\bullet}} \sum_{j=1}^k y_j n_{1j}$$

risulta 5.0909.
Si ottiene

y_j	n_{1j}	$y_j n_{1j}$
4	30	120
5	40	200
6	40	240
7	0	0
110		560

da cui segue

$$M(Y|x_1) = \frac{1}{n_{1\bullet}} \sum_{j=1}^k y_j n_{1j} = \frac{560}{110} = 5.0909.$$

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.



La varianza

$$\text{Var}(Y|x_1) = M(Y^2|x_1) - [M(Y|x_1)]^2$$

risulta 0.6281.

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.



Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.

La varianza

$$\text{Var}(Y|x_1) = M(Y^2|x_1) - [M(Y|x_1)]^2$$

risulta 0.6281. Si ottiene

y_j	n_{1j}	$y_j n_{1j}$	$y_j^2 n_{1j}$
4	30	120	480
5	40	200	1000
6	40	240	1440
7	0	0	0
	110	560	2920

da cui segue

$$\text{Var}(Y|x_1) = \frac{2920}{110} - 5.0909^2 = 26.5455 - 25.9174 = 0.6281.$$



Per effettuare confronti in termini di variabilità tra caratteri quantitativi occorre utilizzare una misura relativa, ad esempio il coefficiente di variazione

$$CV = \frac{\sigma}{\mu} = \frac{\sqrt{\text{Var}(Y|x_1)}}{M(Y|x_1)} = \frac{\sqrt{0.6281}}{5.0909} = \frac{0.7925}{5.0909} = 0.1557.$$

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.



Si considera la distribuzione condizionata $Y|x_2$

y_j	n_{2j}
4	0
5	40
6	50
7	20

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.



Si tratta di un carattere quantitativo con modalità non raggruppate in classi.

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.



La media aritmetica

$$M(Y|x_2) = \frac{1}{n_{2\bullet}} \sum_{j=1}^k y_j n_{2j}$$

risulta 5.8182.

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.



La media aritmetica

$$M(Y|x_2) = \frac{1}{n_{2\bullet}} \sum_{j=1}^k y_j n_{2j}$$

risulta 5.8182.
Si ottiene

y_j	n_{2j}	$y_j n_{2j}$
4	0	0
5	40	200
6	50	300
7	20	140
110		640

da cui segue

$$M(Y|x_2) = \frac{1}{n_{2\bullet}} \sum_{j=1}^k y_j n_{2j} = \frac{640}{110} = 5.8182.$$

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.



La varianza

$$\text{Var}(Y|x_2) = M(Y^2|x_2) - [M(Y|x_2)]^2$$

risulta 0.5124.

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.



La varianza

$$Var(Y|x_2) = M(Y^2|x_2) - [M(Y|x_2)]^2$$

risulta 0.5124. Si ottiene

y_j	n_{2j}	$y_j n_{2j}$	$y_j^2 n_{2j}$
4	0	0	0
5	40	200	1000
6	50	300	1800
7	20	140	980
	110	640	3780

da cui segue

$$Var(Y|x_2) = \frac{3780}{110} - 5.8182^2 = 34.3636 - 33.8512 = 0.5124.$$

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.



Per effettuare confronti in termini di variabilità tra caratteri quantitativi occorre utilizzare una misura relativa, ad esempio il coefficiente di variazione

$$CV = \frac{\sigma}{\mu} = \frac{\sqrt{\text{Var}(Y|x_2)}}{M(Y|x_2)} = \frac{\sqrt{0.5124}}{5.8182} = \frac{0.7158}{5.8182} = 0.123.$$

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.



$Y|x_1$ Indice di posizione: $M(Y|x_1)$ valore 5.0909 Indice di variabilità: $CV(Y|x_1)$ valore 0.1557
 $Y|x_2$ Indice di posizione: $M(Y|x_2)$ valore 5.8182 Indice di variabilità: $CV(Y|x_2)$ valore 0.1230
La distribuzione condizionata $Y|x_1$ è caratterizzata da un livello medio leggermente inferiore a quello della condizionata $Y|x_2$ ma da variabilità superiore. Per il confronto in termini di variabilità non è possibile utilizzare le varianze $Var(Y|x_1) = 0.6281$, $Var(Y|x_2) = 0.5124$, ma occorre calcolare i coefficienti di variazione.

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.



3. Utilizzando i risultati al punto precedente si calcoli un indice di dipendenza in media di Y dato X .

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.



Il rapporto di correlazione

$$\eta_{Y|X}^2 = \frac{\text{Var}_X[M(Y|X)]}{\text{Var}(Y)}$$

può essere ottenuto considerando il rapporto tra la varianza spiegata $\text{Var}_X[M(Y|X)]$,
varianza della variabile statistica medie condizionate

i	$M(X y_i)$	$n_{i\bullet}$
1	5.0909	110
2	5.8182	110

e la varianza della variabile dipendente $\text{Var}(Y) = 0.7024$.

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.



La varianza spiegata risulta $Var_X[M(Y|X)] = 0.1322$.

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.



La varianza spiegata risulta $\text{Var}_X[M(Y|X)] = 0.1322$. Si ottiene

i	$M(X y_i)$	$n_{i\bullet}$	$M(X y_i) \cdot n_{i\bullet}$	$[M(X y_i)]^2 \cdot n_{i\bullet}$
1	5.0909	110	559.999	2850.8989
2	5.8182	110	640.002	3723.6596
		220	1200.001	6574.5585

da cui segue

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.



La varianza spiegata risulta $Var_X[M(Y|X)] = 0.1322$. Si ottiene

i	$M(X y_i)$	$n_{i\bullet}$	$M(X y_i) \cdot n_{i\bullet}$	$[M(X y_i)]^2 \cdot n_{i\bullet}$
1	5.0909	110	559.999	2850.8989
2	5.8182	110	640.002	3723.6596
		220	1200.001	6574.5585

da cui segue

$$Var_X[M(Y|X)] = \frac{6574.5585}{220} - 5.4546^2 = 29.8844 - 29.7521 = 0.1322.$$

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.



Il rapporto di correlazione $\eta^2_{Y|X}$ risulta:

$$\eta^2_{Y|X} = \frac{\text{Var}_X[M(Y|X)]}{\text{Var}(Y)} = \frac{0.1322}{0.7024} = 0.1882.$$

Sussiste, quindi, un livello di dipendenza molto basso.

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.



Si ricorda che, in base al risultato di scomposizione della varianza, la varianza della variabile dipendente può essere ottenuta come somma della varianza spiegata $Var_X[M(Y|X)]$ e della varianza residua $M_X[Var(Y|X)]$.

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.



La varianza residua $M_X[Var(Y|X)]$, media della variabile statistica varianze condizionate

i	$Var(X y_i)$	$n_{i\bullet}$
1	0.6281	110
2	0.5124	110

risulta $M_X[Var(Y|X)] = 0.5702$.

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.



Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.

La varianza residua $M_X[Var(Y|X)]$, media della variabile statistica varianze condizionate

i	$Var(X y_i)$	$n_{i\bullet}$
1	0.6281	110
2	0.5124	110

risulta $M_X[Var(Y|X)] = 0.5702$. Si ottiene

i	$Var(X y_i)$	$n_{i\bullet}$	$Var(X y_i) \cdot n_{i\bullet}$
1	0.6281	110	69.091
2	0.5124	110	56.364
		220	125.455

da cui segue



La varianza residua $M_X[Var(Y|X)]$, media della variabile statistica varianze condizionate

i	$Var(X y_i)$	$n_{i\bullet}$
1	0.6281	110
2	0.5124	110

risulta $M_X[Var(Y|X)] = 0.5702$. Si ottiene

i	$Var(X y_i)$	$n_{i\bullet}$	$Var(X y_i) \cdot n_{i\bullet}$
1	0.6281	110	69.091
2	0.5124	110	56.364
		220	125.455

da cui segue

$$M_X[Var(Y|X)] = \frac{125.455}{220} = 0.5702.$$



$\sigma_{Between}^2 = 0.1322$ $\sigma_{Within}^2 = 0.5702$. da cui abbiamo $\eta_{Y|X}^2 = 0.1882$.

Sussiste, quindi, un livello di dipendenza basso del numero di esami sostenuti rispetto alla classe.

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.



4. Senza effettuare calcoli, si dica se può sussistere indipendenza stocastica tra le due variabili motivando la risposta.

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.



4. Senza effettuare calcoli, si dica se può sussistere indipendenza stocastica tra le due variabili motivando la risposta.

Non sussiste indipendenza stocastica in quanto le distribuzioni condizionate non sono somiglianti.

[Esercizio 1.](#)

[Esercizio 2.](#)

[Esercizio 3.](#)

[Esercizio 4.](#)

Esercizio 2



Con riferimento alla serie storica di un titolo azionario x_t sono note le informazioni riportate nel seguente prospetto:

t	x_t	$NIBM_t$
1		—
2		1.05
3	100	1.00
4		0.95
5		0.90

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.



1. Si ricostruisca la serie delle osservazioni x_t a partire dalle informazioni disponibili. Si ricorda che $x_3 = 100$, che per valori di $t > 3$ vale

$$x_t = x_3 \cdot {}_3l_4 \cdot {}_4l_5 \cdot \dots \cdot {}_{t-1}l_t$$

e che per valori di $t < 3$ vale

$$x_t = \frac{x_3}{{}_t l_{t+1} \cdot \dots \cdot {}_1 l_2 \cdot {}_2 l_3}.$$

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.



Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.

1. Si ricostruisca la serie delle osservazioni x_t a partire dalle informazioni disponibili. Si ricorda che $x_3 = 100$, che per valori di $t > 3$ vale

$$x_t = x_3 \cdot {}_3l_4 \cdot {}_4l_5 \cdot \dots \cdot {}_{t-1}l_t$$

e che per valori di $t < 3$ vale

$$x_t = \frac{x_3}{{}_t l_{t+1} \cdot \dots \cdot {}_1 l_2 \cdot {}_2 l_3}.$$

Si ottiene quindi

t	$NIBM$	x_t
1	0.80	95.2381
2	1.05	100.0000
3	1.00	100.0000
4	0.95	95.0000
5	0.90	85.5000



2. Si calcoli il tasso medio di variazione tra $t = 1$ e $t = 5$ e si interpreti il valore ottenuto.

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.



2. Si calcoli il tasso medio di variazione tra $t = 1$ e $t = 5$ e si interpreti il valore ottenuto.
Si ricorda che

$${}_1I_5 = {}_1I_2 \cdot {}_2I_3 \cdot {}_3I_4 \cdot {}_4I_5.$$

Si deve trovare il numero indice a base mobile medio.

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.



Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.

2. Si calcoli il tasso medio di variazione tra $t = 1$ e $t = 5$ e si interpreti il valore ottenuto. Si ricorda che

$${}_1l_5 = {}_1l_2 \cdot {}_2l_3 \cdot {}_3l_4 \cdot {}_4l_5.$$

Si deve trovare il numero indice a base mobile medio. Applicando il criterio di scelta della media secondo Chisini si ottiene

$${}_1l_5 = {}_1l_2 \cdot {}_2l_3 \cdot {}_3l_4 \cdot {}_4l_5 = \alpha \cdot \alpha \cdot \alpha \cdot \alpha = \alpha^4.$$



Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.

2. Si calcoli il tasso medio di variazione tra $t = 1$ e $t = 5$ e si interpreti il valore ottenuto. Si ricorda che

$${}_1I_5 = {}_1I_2 \cdot {}_2I_3 \cdot {}_3I_4 \cdot {}_4I_5.$$

Si deve trovare il numero indice a base mobile medio. Applicando il criterio di scelta della media secondo Chisini si ottiene

$${}_1I_5 = {}_1I_2 \cdot {}_2I_3 \cdot {}_3I_4 \cdot {}_4I_5 = \alpha \cdot \alpha \cdot \alpha \cdot \alpha = \alpha^4.$$

Quindi

$$\alpha = \sqrt[4]{{}_1I_5} = \sqrt[4]{{}_1I_2 \cdot {}_2I_3 \cdot {}_3I_4 \cdot {}_4I_5} = 0.9734.$$

Da cui segue una variazione relativa media pari a -0.0266 , corrispondente a un decremento medio 1 e il 5 del 2.66%.



3. Si determini il valore del titolo al tempo $t = 7$ assumendo tra $t = 5$ e $t = 7$ la stessa variazione (media annua) osservata nei periodi precedenti.

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.



3. Si determini il valore del titolo al tempo $t = 7$ assumendo tra $t = 5$ e $t = 7$ la stessa variazione (media annua) osservata nei periodi precedenti.

Il numero indice a base mobile medio è pari a $\alpha = 0.9734$. Abbiamo, quindi:

$$x_7 = x_5 \cdot \alpha \cdot \alpha = x_5 \cdot 0.9734 \cdot 0.9734 = 81.0110$$

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.



Il seguente prospetto riporta i valori rilevati in corrispondenza a 5 unità statistiche con riferimento a due caratteri quantitativi X e Y :

x_i	1	1	3	5	5
y_i	1	2	3	4	5

1. Si rappresentino graficamente le coppie dei dati e la funzione di regressione di Y in funzione di X .

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.



La funzione di regressione associa a ogni valore x appartenente all'insieme di definizione della variabile X la media di Y condizionata a x , $M(Y|x)$

x	$M(Y x)$
1	$\frac{1}{2}(1 + 2) = 1.5$
3	3
5	$\frac{1}{2}(4 + 5) = 4.5$

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.



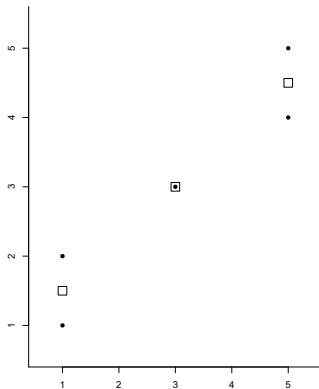
Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.

Le coppie di dati possono essere rappresentate mediante un grafico a dispersione.





2. Si calcolino, secondo il metodo dei minimi quadrati, i parametri dei seguenti modelli di regressione:

$$I) Y^* = a + b \cdot X \quad II) Y^* = c \cdot X$$

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.

Si riportano i calcoli necessari¹ per ottenere le stime dei parametri e gli indici di adattamento del modello

$$Y = a + bX + E$$

x_i	y_i	$x_i y_i$	x_i^2	y_i^2
1	1	1	1	1
1	2	2	1	4
3	3	9	9	9
5	4	20	25	16
5	5	25	25	25
15	15	57	61	55
3	3	11.4	12.2	11
$M(X)$	$M(Y)$	$M(XY)$	$M(X^2)$	$M(Y^2)$

$$\hat{b} =$$



Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.

¹Si osserva che per ragioni di stampa alcuni calcoli potrebbero essere duplicati con riferimento ai diversi modelli oggetto di studio

Si riportano i calcoli necessari¹ per ottenere le stime dei parametri e gli indici di adattamento del modello

$$Y = a + bX + E$$

x_i	y_i	$x_i y_i$	x_i^2	y_i^2
1	1	1	1	1
1	2	2	1	4
3	3	9	9	9
5	4	20	25	16
5	5	25	25	25
15	15	57	61	55
3	3	11.4	12.2	11
$M(X)$	$M(Y)$	$M(XY)$	$M(X^2)$	$M(Y^2)$

$$\begin{aligned}\hat{b} &= \frac{\text{Cov}(X, Y)}{\text{Var}(X)} = \frac{M(XY) - M(X)M(Y)}{M(X^2) - [M(X)]^2} = \\ &= \frac{11.4 - 3 \cdot 3}{12.2 - 3^2} = \frac{2.4}{3.2} = 0.75 \\ \hat{a} &= \end{aligned}$$

¹Si osserva che per ragioni di stampa alcuni calcoli potrebbero essere duplicati con riferimento ai diversi modelli oggetto di studio



Si riportano i calcoli necessari¹ per ottenere le stime dei parametri e gli indici di adattamento del modello

$$Y = a + bX + E$$

x_i	y_i	$x_i y_i$	x_i^2	y_i^2
1	1	1	1	1
1	2	2	1	4
3	3	9	9	9
5	4	20	25	16
5	5	25	25	25
15	15	57	61	55
3	3	11.4	12.2	11
$M(X)$	$M(Y)$	$M(XY)$	$M(X^2)$	$M(Y^2)$

$$\hat{b} = \frac{\text{Cov}(X, Y)}{\text{Var}(X)} = \frac{M(XY) - M(X)M(Y)}{M(X^2) - [M(X)]^2} =$$

$$= \frac{11.4 - 3 \cdot 3}{12.2 - 3^2} = \frac{2.4}{3.2} = 0.75$$

$$\hat{a} = M(Y) - \hat{b}M(X) = 0.75$$

$$\rho^2 =$$



Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.

¹Si osserva che per ragioni di stampa alcuni calcoli potrebbero essere duplicati con riferimento ai diversi modelli oggetto di studio

Si riportano i calcoli necessari¹ per ottenere le stime dei parametri e gli indici di adattamento del modello

$$Y = a + bX + E$$

x_i	y_i	$x_i y_i$	x_i^2	y_i^2
1	1	1	1	1
1	2	2	1	4
3	3	9	9	9
5	4	20	25	16
5	5	25	25	25
15	15	57	61	55
3	3	11.4	12.2	11
$M(X)$	$M(Y)$	$M(XY)$	$M(X^2)$	$M(Y^2)$

$$\hat{b} = \frac{\text{Cov}(X, Y)}{\text{Var}(X)} = \frac{M(XY) - M(X)M(Y)}{M(X^2) - [M(X)]^2} =$$

$$= \frac{11.4 - 3 \cdot 3}{12.2 - 3^2} = \frac{2.4}{3.2} = 0.75$$

$$\hat{a} = M(Y) - \hat{b}M(X) = 0.75$$

$$\rho^2 = \frac{[\text{Cov}(X, Y)]^2}{\text{Var}(X) \text{Var}(Y)} = \frac{2.4^2}{3.2 \cdot 2} = 0.9$$

$$M(\hat{E}^2) = \text{Var}(\hat{E}) = \text{varianza residua} =$$

¹Si osserva che per ragioni di stampa alcuni calcoli potrebbero essere duplicati con riferimento ai diversi modelli oggetto di studio



Si riportano i calcoli necessari¹ per ottenere le stime dei parametri e gli indici di adattamento del modello

$$Y = a + bX + E$$

x_i	y_i	$x_i y_i$	x_i^2	y_i^2
1	1	1	1	1
1	2	2	1	4
3	3	9	9	9
5	4	20	25	16
5	5	25	25	25
15	15	57	61	55
3	3	11.4	12.2	11
$M(X)$	$M(Y)$	$M(XY)$	$M(X^2)$	$M(Y^2)$

$$\hat{b} = \frac{\text{Cov}(X, Y)}{\text{Var}(X)} = \frac{M(XY) - M(X)M(Y)}{M(X^2) - [M(X)]^2} =$$

$$= \frac{11.4 - 3 \cdot 3}{12.2 - 3^2} = \frac{2.4}{3.2} = 0.75$$

$$\hat{a} = M(Y) - \hat{b}M(X) = 0.75$$

$$\rho^2 = \frac{[\text{Cov}(X, Y)]^2}{\text{Var}(X) \text{Var}(Y)} = \frac{2.4^2}{3.2 \cdot 2} = 0.9$$

$$M(\hat{E}^2) = \text{Var}(\hat{E}) = \text{varianza residua} = \text{Var}(Y) \cdot (1 - \rho^2) = 0.2$$

¹Si osserva che per ragioni di stampa alcuni calcoli potrebbero essere duplicati con riferimento ai diversi modelli oggetto di studio





Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.

Si riportano i calcoli necessari² per ottenere le stime dei parametri e la misura di adattamento del modello

$$Y = cX + E$$

x_i	y_i	$x_i y_i$	x_i^2	y_i^2
1	1	1	1	1
1	2	2	1	4
3	3	9	9	9
5	4	20	25	16
5	5	25	25	25
		57	61	55
		11.4	12.2	11
		$M(XY)$	$M(X^2)$	$M(Y^2)$

$$\hat{c} =$$

²Si osserva che per ragioni di stampa alcuni calcoli potrebbero essere duplicati con riferimento ai diversi modelli oggetto di studio



Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.

Si riportano i calcoli necessari² per ottenere le stime dei parametri e la misura di adattamento del modello

$$Y = cX + E$$

x_i	y_i	$x_i y_i$	x_i^2	y_i^2
1	1	1	1	1
1	2	2	1	4
3	3	9	9	9
5	4	20	25	16
5	5	25	25	25
		57	61	55
		11.4	12.2	11
		$M(XY)$	$M(X^2)$	$M(Y^2)$

$$\hat{c} = \frac{M(XY)}{M(X^2)} = \frac{11.4}{12.2} = 0.9344$$

$$M(\hat{E}^2) =$$

²Si osserva che per ragioni di stampa alcuni calcoli potrebbero essere duplicati con riferimento ai diversi modelli oggetto di studio



Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.

Si riportano i calcoli necessari² per ottenere le stime dei parametri e la misura di adattamento del modello

$$Y = cX + E$$

x_i	y_i	$x_i y_i$	x_i^2	y_i^2
1	1	1	1	1
1	2	2	1	4
3	3	9	9	9
5	4	20	25	16
5	5	25	25	25
		57	61	55
		11.4	12.2	11
		$M(XY)$	$M(X^2)$	$M(Y^2)$

$$\hat{c} = \frac{M(XY)}{M(X^2)} = \frac{11.4}{12.2} = 0.9344$$

$$M(\hat{E}^2) = M[Y \cdot (Y - \hat{c}X)] = M(Y^2) - \hat{c}M(XY) = 11 - 0.9344 \cdot 11.4 = 0.3475$$

²Si osserva che per ragioni di stampa alcuni calcoli potrebbero essere duplicati con riferimento ai diversi modelli oggetto di studio



$\hat{a} = 0.75$, $\hat{b} = 0.75$, indice adattamento $M(\hat{E}^2) = 0.2$
 $M(\hat{E}^2) = 0.3475$

$\hat{c} = 0.9344$, indice adattamento

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.



Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.

$\hat{a} = 0.75$, $\hat{b} = 0.75$, indice adattamento $M(\hat{E}^2) = 0.2$
 $M(\hat{E}^2) = 0.3475$

Il modello migliore è $Y^* = 0.75 + 0.75X$.

$\hat{c} = 0.9344$, indice adattamento



Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.

3. In corrispondenza del modello migliore, si fornisca la previsione per $x = 7$.



Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.

3. In corrispondenza del modello migliore, si fornisca la previsione per $x = 7$.
 $Y^*(7) = 0.75 + 0.75 \cdot 7 = 6.$



4. Si stabilisca, senza effettuare calcoli e giustificando teoricamente la risposta, il valore assunto dagli indici $\eta^2_{Y|X}$ e $\eta^2_{X|Y}$.

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.



4. Si stabilisca, senza effettuare calcoli e giustificando teoricamente la risposta, il valore assunto dagli indici $\eta^2_{Y|X}$ e $\eta^2_{X|Y}$.

$$\eta^2_{Y|X} =$$

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.



4. Si stabilisca, senza effettuare calcoli e giustificando teoricamente la risposta, il valore assunto dagli indici $\eta_{Y|X}^2$ e $\eta_{X|Y}^2$.

$\eta_{Y|X}^2 = \rho^2 = 0.9$ in quanto le medie condizionate $M(Y|X)$ sono allineate

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.



Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.

4. Si stabilisca, senza effettuare calcoli e giustificando teoricamente la risposta, il valore assunto dagli indici $\eta_{Y|X}^2$ e $\eta_{X|Y}^2$.

$\eta_{Y|X}^2 = \rho^2 = 0.9$ in quanto le medie condizionate $M(Y|X)$ sono allineate
 $\eta_{X|Y}^2 = 1$ in quanto sussiste dipendenza funzionale di X da Y .

Esercizio 4

Prova scritta di
STATISTICA del
05.09.2019

Giuseppe Boari
Gabriele Cantaluppi



UNIVERSITÀ
CATTOLICA
del Sacro Cuore

Un'urna contiene 50 palline, di cui 10 rosse, mentre si sa che le restanti possono essere blu o verdi.

1. Si effettuano 5 estrazioni con reimmissione. Si calcoli la probabilità che nessuna delle palline estratte sia rossa.

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.

Esercizio 4

Prova scritta di
STATISTICA del
05.09.2019

Giuseppe Boari
Gabriele Cantaluppi



UNIVERSITÀ
CATTOLICA
del Sacro Cuore

Un'urna contiene 50 palline, di cui 10 rosse, mentre si sa che le restanti possono essere blu o verdi.

1. Si effettuano 5 estrazioni con reimmissione. Si calcoli la probabilità che nessuna delle palline estratte sia rossa.

Sia X la variabile casuale 'n. palline rosse tra le 5 estratti'. Abbiamo

[Esercizio 1.](#)

[Esercizio 2.](#)

[Esercizio 3.](#)

[Esercizio 4.](#)



[Esercizio 1.](#)

[Esercizio 2.](#)

[Esercizio 3.](#)

[Esercizio 4.](#)

Un'urna contiene 50 palline, di cui 10 rosse, mentre si sa che le restanti possono essere blu o verdi.

1. Si effettuano 5 estrazioni con reimmissione. Si calcoli la probabilità che nessuna delle palline estratte sia rossa.

Sia X la variabile casuale 'n. palline rosse tra le 5 estratti'. Abbiamo

$$X \sim \text{Bin}(n = 5, p = 0.2)$$

e

$$P(X = 0) =$$



[Esercizio 1.](#)

[Esercizio 2.](#)

[Esercizio 3.](#)

[Esercizio 4.](#)

Un'urna contiene 50 palline, di cui 10 rosse, mentre si sa che le restanti possono essere blu o verdi.

1. Si effettuano 5 estrazioni con reimmissione. Si calcoli la probabilità che nessuna delle palline estratte sia rossa.

Sia X la variabile casuale 'n. palline rosse tra le 5 estratti'. Abbiamo

$$X \sim \text{Bin}(n = 5, p = 0.2)$$

e

$$P(X = 0) = \binom{5}{0} 0.2^0 \cdot 0.8^5 = 0.32768.$$



2. Si ipotizzi ora di effettuare 100 estrazioni con reimmissione. Si calcoli la probabilità di ottenere almeno 20 palline rosse.

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.



2. Si ipotizzi ora di effettuare 100 estrazioni con reimmissione. Si calcoli la probabilità di ottenere almeno 20 palline rosse.

Indicata con X la variabile casuale 'numero palline rosse nelle 100 estrazioni con reimmissione dall'urna U_1 ', risulta $X \sim$

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.



Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.

2. Si ipotizzi ora di effettuare 100 estrazioni con reimmissione. Si calcoli la probabilità di ottenere almeno 20 palline rosse.

Indicata con X la variabile casuale 'numero palline rosse nelle 100 estrazioni con reimmissione dall'urna U_1 ' risulta $X \sim \text{Bin}(n = 100, p = 0.2)$. $M(X) = 20$; $\text{Var}(X) = 16$. Dal momento che $np = 20 > 5$ ed $n(1 - p) = 80 > 5$ è possibile utilizzare l'approssimazione della variabile casuale Binomiale con la variabile casuale Normale

$$X \sim$$



2. Si ipotizzi ora di effettuare 100 estrazioni con reimmissione. Si calcoli la probabilità di ottenere almeno 20 palline rosse.

Indicata con X la variabile casuale 'numero palline rosse nelle 100 estrazioni con reimmissione dall'urna U_1 ' risulta $X \sim \text{Bin}(n = 100, p = 0.2)$. $M(X) = 20$; $\text{Var}(X) = 16$. Dal momento che $np = 20 > 5$ ed $n(1 - p) = 80 > 5$ è possibile utilizzare l'approssimazione della variabile casuale Binomiale con la variabile casuale Normale

$$X \sim N(\mu = 20, \sigma^2 = 16).$$

Dal momento che la variabile casuale Normale è simmetrica rispetto alla media, si ottiene

$$P(X > 20) =$$



2. Si ipotizzi ora di effettuare 100 estrazioni con reimmissione. Si calcoli la probabilità di ottenere almeno 20 palline rosse.

Indicata con X la variabile casuale 'numero palline rosse nelle 100 estrazioni con reimmissione dall'urna U_1 ' risulta $X \sim \text{Bin}(n = 100, p = 0.2)$. $M(X) = 20$; $\text{Var}(X) = 16$. Dal momento che $np = 20 > 5$ ed $n(1 - p) = 80 > 5$ è possibile utilizzare l'approssimazione della variabile casuale Binomiale con la variabile casuale Normale

$$X \sim N(\mu = 20, \sigma^2 = 16).$$

Dal momento che la variabile casuale Normale è simmetrica rispetto alla media, si ottiene

$$P(X > 20) = 0.5.$$



3. Sapendo che la probabilità di estrarre 3 palline blu in 3 estrazioni con reimmissione è pari a 0.125, si determini il numero di palline blu contenute nell'urna.

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.



3. Sapendo che la probabilità di estrarre 3 palline blu in 3 estrazioni con reimmissione è pari a 0.125, si determini il numero di palline blu contenute nell'urna.

Indicata con p la frazione di palline blu contenute nell'urna, che coincide con la probabilità di ottenere una pallina blu in una generica estrazione Bernoulliana, la probabilità di ottenere palline blu in 3 estrazioni con reimmissione risulta

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.



Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.

3. Sapendo che la probabilità di estrarre 3 palline blu in 3 estrazioni con reimmissione è pari a 0.125, si determini il numero di palline blu contenute nell'urna.

Indicata con p la frazione di palline blu contenute nell'urna, che coincide con la probabilità di ottenere una pallina blu in una generica estrazione Bernoulliana, la probabilità di ottenere palline blu in 3 estrazioni con reimmissione risulta

$$p^3 = 0.125$$

da cui segue $p =$



Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.

3. Sapendo che la probabilità di estrarre 3 palline blu in 3 estrazioni con reimmissione è pari a 0.125, si determini il numero di palline blu contenute nell'urna.

Indicata con p la frazione di palline blu contenute nell'urna, che coincide con la probabilità di ottenere una pallina blu in una generica estrazione Bernoulliana, la probabilità di ottenere palline blu in 3 estrazioni con reimmissione risulta

$$p^3 = 0.125$$

da cui segue $p = 0.5$. Il numero di palline blu contenute nell'urna risulta, quindi,



Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.

3. Sapendo che la probabilità di estrarre 3 palline blu in 3 estrazioni con reimmissione è pari a 0.125, si determini il numero di palline blu contenute nell'urna.

Indicata con p la frazione di palline blu contenute nell'urna, che coincide con la probabilità di ottenere una pallina blu in una generica estrazione Bernoulliana, la probabilità di ottenere palline blu in 3 estrazioni con reimmissione risulta

$$p^3 = 0.125$$

da cui segue $p = 0.5$. Il numero di palline blu contenute nell'urna risulta, quindi,
 $50 \cdot p = 50 \cdot 0.5 = 25$.

Tema T3052

Prova scritta di STATISTICA del 12.09.2019

Prova scritta di
STATISTICA del
12.09.2019

Giuseppe Boari
Gabriele Cantaluppi



UNIVERSITÀ
CATTOLICA
del Sacro Cuore

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.

Giuseppe Boari
Gabriele Cantaluppi
Università Cattolica del Sacro Cuore, Milano



Il prospetto seguente riporta la distribuzione delle frequenze relative del carattere quantitativo Y condizionate rispetto alle categorie del carattere qualitativo X .

$Y X$	2	5	8	
x_1	0.80	0.10	0.10	1
x_2	0.10	0.10	0.80	1

1. Si completi la distribuzione delle frequenze congiunte sapendo che il numero di unità statistiche con categoria x_1 è 20, e che numero di unità statistiche con categoria x_2 è 80.

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.



Moltiplicando le frequenze condizionate relative $f.cond.r_{1j} = \frac{n_{1j}}{n_{1\bullet}}$ per il numero totale di osservazioni $n_{1\bullet}$ della distribuzione condizionata, si ottengono le frequenze assolute congiunte n_{1j} della distribuzione condizionata $Y|X_1$.

y_j	$f.cond.r_{1j}$	n_{1j}
2	0.8	16
5	0.1	2
8	0.1	2

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.



Moltiplicando le frequenze condizionate relative $f.cond.r_{2j} = \frac{n_{2j}}{n_{2\bullet}}$ per il numero totale di osservazioni $n_{2\bullet}$ della distribuzione condizionata, si ottengono le frequenze assolute congiunte n_{2j} della distribuzione condizionata $Y|X_2$.

y_j	$f.cond.r_{2j}$	n_{2j}
2	0.1	8
5	0.1	8
8	0.8	64

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.



La distribuzione delle frequenze congiunte risulta

$X \backslash Y$	2	5	8	
x_1	16	2	2	20
x_2	8	8	64	80
	24	10	66	100

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.



Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.

2. Si rappresentino graficamente le distribuzioni condizionate $Y|x_1$ e $Y|x_2$.



Si considera la distribuzione condizionata $Y|x_1$

y_j	f_{1j}
2	0.8
5	0.1
8	0.1

Esercizio 1.

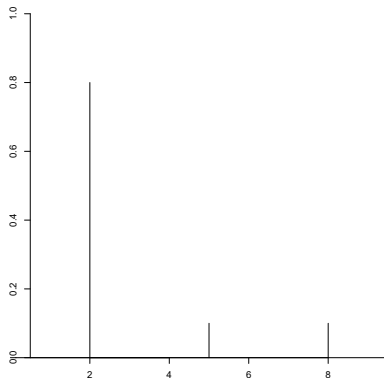
Esercizio 2.

Esercizio 3.

Esercizio 4.



Si tratta di un carattere quantitativo con modalità non raggruppate in classi che può essere rappresentato graficamente mediante un grafico a bastoncini.



Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.



Si considera la distribuzione condizionata $Y|x_2$

y_j	f_{2j}
2	0.1
5	0.1
8	0.8

Esercizio 1.

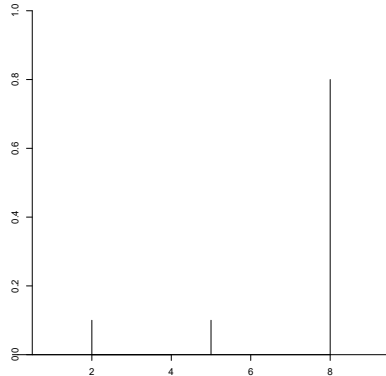
Esercizio 2.

Esercizio 3.

Esercizio 4.



Si tratta di un carattere quantitativo con modalità non raggruppate in classi che può essere rappresentato graficamente mediante un grafico a bastoncini.



Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.



3. Si confrontino con opportuni indici di posizione e variabilità le distribuzioni condizionate $Y|x_1$ e $Y|x_2$.

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.



Si considera la distribuzione condizionata $Y|x_1$

y_j	n_{1j}
2	0.8
5	0.1
8	0.1

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.



Si tratta di un carattere quantitativo con modalità non raggruppate in classi.

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.



La media aritmetica

$$M(Y|x_1) = \frac{1}{n_{1\bullet}} \sum_{j=1}^k y_j n_{1j}$$

risulta 2.9.

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.



La media aritmetica

$$M(Y|x_1) = \frac{1}{n_{1\bullet}} \sum_{j=1}^k y_j n_{1j}$$

risulta 2.9.

Si ottiene

y_j	n_{1j}	$y_j n_{1j}$
2	0.8	1.6
5	0.1	0.5
8	0.1	0.8
	1.0	2.9

da cui segue

$$M(Y|x_1) = \frac{1}{n_{1\bullet}} \sum_{j=1}^k y_j n_{1j} = \frac{2.9}{1} = 2.9.$$

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.



La varianza

$$\text{Var}(Y|x_1) = M(Y^2|x_1) - [M(Y|x_1)]^2$$

risulta 3.69.

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.



La varianza

$$\text{Var}(Y|x_1) = M(Y^2|x_1) - [M(Y|x_1)]^2$$

risulta 3.69. Si ottiene

y_j	n_{1j}	$y_j n_{1j}$	$y_j^2 n_{1j}$
2	0.8	1.6	3.2
5	0.1	0.5	2.5
8	0.1	0.8	6.4
	1.0	2.9	12.1

da cui segue

$$\text{Var}(Y|x_1) = \frac{12.1}{1} - 2.9^2 = 12.1 - 8.41 = 3.69.$$

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.



Per effettuare confronti in termini di variabilità tra caratteri quantitativi occorre utilizzare una misura relativa, ad esempio il coefficiente di variazione

$$CV = \frac{\sigma}{\mu} = \frac{\sqrt{\text{Var}(Y|x_1)}}{M(Y|x_1)} = \frac{\sqrt{3.69}}{2.9} = \frac{1.9209}{2.9} = 0.6624.$$

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.



Si considera la distribuzione condizionata $Y|x_2$

y_j	n_{2j}
2	0.1
5	0.1
8	0.8

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.



Si tratta di un carattere quantitativo con modalità non raggruppate in classi.

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.



La media aritmetica

$$M(Y|x_2) = \frac{1}{n_{2\bullet}} \sum_{j=1}^k y_j n_{2j}$$

risulta 7.1.

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.



La media aritmetica

$$M(Y|x_2) = \frac{1}{n_{2\bullet}} \sum_{j=1}^k y_j n_{2j}$$

risulta 7.1.
Si ottiene

y_j	n_{2j}	$y_j n_{2j}$
2	0.1	0.2
5	0.1	0.5
8	0.8	6.4
	1.0	7.1

da cui segue

$$M(Y|x_2) = \frac{1}{n_{2\bullet}} \sum_{j=1}^k y_j n_{2j} = \frac{7.1}{1} = 7.1.$$

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.



La varianza

$$\text{Var}(Y|x_2) = M(Y^2|x_2) - [M(Y|x_2)]^2$$

risulta 3.69.

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.



La varianza

$$\text{Var}(Y|x_2) = M(Y^2|x_2) - [M(Y|x_2)]^2$$

risulta 3.69. Si ottiene

y_j	n_{2j}	$y_j n_{2j}$	$y_j^2 n_{2j}$
2	0.1	0.2	0.4
5	0.1	0.5	2.5
8	0.8	6.4	51.2
	1.0	7.1	54.1

da cui segue

$$\text{Var}(Y|x_2) = \frac{54.1}{1} - 7.1^2 = 54.1 - 50.41 = 3.69.$$

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.



Per effettuare confronti in termini di variabilità tra caratteri quantitativi occorre utilizzare una misura relativa, ad esempio il coefficiente di variazione

$$CV = \frac{\sigma}{\mu} = \frac{\sqrt{\text{Var}(Y|x_2)}}{M(Y|x_2)} = \frac{\sqrt{3.69}}{7.1} = \frac{1.9209}{7.1} = 0.2705.$$

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.



Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.

$Y|x_1$ Indice di posizione: $M(Y|x_1) = 2.9$ Indice di variabilità: $CV(Y|x_1) = 0.6624$

$Y|x_2$ Indice di posizione: $M(Y|x_2) = 7.1$ Indice di variabilità: $CV(Y|x_2) = 0.2705$

La distribuzione condizionata $Y|x_1$ è caratterizzata da un livello medio inferiore a quello della condizionata $Y|x_2$ ma da variabilità superiore.



4. Si determini il grado di connessione tra le variabili X e Y mediante il calcolo di un opportuno indice.

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.



Dalla distribuzione delle frequenze congiunte n_{ij} :

	2	5	8	$n_{i\bullet}$
x_1	16	2	2	20
x_2	8	8	64	80
$n_{\bullet j}$	24	10	66	100

è possibile ottenere le frequenze teoriche nella situazione di indipendenza stocastica

$$\hat{n}_{ij} = \frac{n_{i\bullet} n_{\bullet j}}{n}:$$

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.



Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.

Dalla distribuzione delle frequenze congiunte n_{ij} :

	2	5	8	$n_{i\bullet}$
x_1	16	2	2	20
x_2	8	8	64	80
$n_{\bullet j}$	24	10	66	100

è possibile ottenere le frequenze teoriche nella situazione di indipendenza stocastica

$$\hat{n}_{ij} = \frac{n_{i\bullet} n_{\bullet j}}{n}:$$

	2	5	8	$n_{i\bullet}$
x_1	4.8	2	13.2	20
x_2	19.2	8	52.8	80
$n_{\bullet j}$	24.0	10	66.0	100



le contingenze assolute $n_{ij} - \hat{n}_{ij}$:

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.



le contingenze assolute $n_{ij} - \hat{n}_{ij}$:

	2	5	8
x_1	11.2	0	-11.2
x_2	-11.2	0	11.2

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.



i rapporti di contingenza $\frac{n_{ij}}{\hat{n}_{ij}}$:

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.



i rapporti di contingenza $\frac{n_{ij}}{\hat{n}_{ij}}$:

	2	5	8
x_1	3.3333	1	0.1515
x_2	0.4167	1	1.2121

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.



e le contingenze relative, espresse in termini percentuali, $100 \left(\frac{n_{ij} - \hat{n}_{ij}}{\hat{n}_{ij}} \right) \% = 100 \left(\frac{n_{ij}}{\hat{n}_{ij}} - 1 \right) \%$:

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.



e le contingenze relative, espresse in termini percentuali, $100 \left(\frac{n_{ij} - \hat{n}_{ij}}{\hat{n}_{ij}} \right) \% = 100 \left(\frac{n_{ij}}{\hat{n}_{ij}} - 1 \right) \%$:

	2	5	8
x_1	233.33	0	-84.85
x_2	-58.33	0	21.21

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.



Risultati intermedi $\frac{(n_{ij} - \hat{n}_{ij})^2}{\hat{n}_{ij}}$ per il calcolo del χ^2

	2	5	8
x_1	26.1333	0	9.5030
x_2	6.5333	0	2.3758

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.



Risultati intermedi $\frac{n_{ij}^2}{n_{i\bullet} \cdot n_{\bullet j}}$ per il calcolo del χ^2 (formula operativa)

	2	5	8
x_1	0.5333	0.02	0.0030
x_2	0.0333	0.08	0.7758

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.



Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.

$$\chi^2 = \sum_{i=1}^h \sum_{j=1}^k \frac{(n_{ij} - \hat{n}_{ij})^2}{\hat{n}_{ij}} = 44.5455,$$

$$\chi_N^2 = \frac{\chi^2}{n \cdot \min(h-1, k-1)} = 0.4455$$

$$\chi_N = \sqrt{\chi_N^2} = 0.6675.$$

Sussiste, quindi, un livello di connessione medio alto.

Esercizio 2

Il seguente prospetto riporta i valori del fatturato di un'azienda tra il 2012 e il 2016:

t	x_t
2012	10
2013	30
2014	50
2015	30
2016	10

1. Si costruiscano le serie dei numeri indici a base fissa, con base 2014, e a base mobile.



Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.

Esercizio 2

Il seguente prospetto riporta i valori del fatturato di un'azienda tra il 2012 e il 2016:

t	x_t
2012	10
2013	30
2014	50
2015	30
2016	10

1. Si costruiscano le serie dei numeri indici a base fissa, con base 2014, e a base mobile.
Si ricorda che ${}_{2014}I_t = \frac{x_t}{x_{2014}}$.



Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.

Esercizio 2



Il seguente prospetto riporta i valori del fatturato di un'azienda tra il 2012 e il 2016:

t	x_t
2012	10
2013	30
2014	50
2015	30
2016	10

1. Si costruiscano le serie dei numeri indici a base fissa, con base 2014, e a base mobile. Si ricorda che ${}_{2014}I_t = \frac{x_t}{x_{2014}}$. Si ottiene quindi

Anno	x_t	NIBF (2014)
2012	10	0.2
2013	30	0.6
2014	50	1.0
2015	30	0.6
2016	10	0.2

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.



Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.



Si ricorda che ${}_{t-1}l_t = \frac{x_t}{x_{t-1}}$.

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.



Si ricorda che $_{t-1}l_t = \frac{x_t}{x_{t-1}}$. Si ottiene quindi

Anno	x_t	NIBM
2012	10	
2013	30	3.0000
2014	50	1.6667
2015	30	0.6000
2016	10	0.3333

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.



2. Si calcoli il numero indice al tempo 2014 con base 2013 e si commenti il risultato ottenuto.

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.



2. Si calcoli il numero indice al tempo 2014 con base 2013 e si commenti il risultato ottenuto.

$${}_{2013}I_{2014} = \frac{x_{2014}}{x_{2013}} = \frac{50}{30} = 1.6667$$

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.



2. Si calcoli il numero indice al tempo 2014 con base 2013 e si commenti il risultato ottenuto.

$${}_{2013}I_{2014} = \frac{X_{2014}}{X_{2013}} = \frac{50}{30} = 1.6667$$

Tra il tempo 2013 e il tempo 2014 si è registrato un incremento complessivo del 66.67% della grandezza oggetto di studio.

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.



3. Si calcoli il tasso medio di variazione tra il 2012 e il 2015 e si commenti il risultato ottenuto.

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.



3. Si calcoli il tasso medio di variazione tra il 2012 e il 2015 e si commenti il risultato ottenuto.
Si ricorda che

$${}_{2012}I_{2015} = {}_{2012}I_{2013} \cdot {}_{2013}I_{2014} \cdot {}_{2014}I_{2015}$$

Si deve trovare il numero indice a base mobile medio.

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.



Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.

3. Si calcoli il tasso medio di variazione tra il 2012 e il 2015 e si commenti il risultato ottenuto. Si ricorda che

$${}_{2012}I_{2015} = {}_{2012}I_{2013} \cdot {}_{2013}I_{2014} \cdot {}_{2014}I_{2015}$$

Si deve trovare il numero indice a base mobile medio. Applicando il criterio di scelta della media secondo Chisini si ottiene

$${}_{2012}I_{2015} = {}_{2012}I_{2013} \cdot {}_{2013}I_{2014} \cdot {}_{2014}I_{2015} = \alpha \cdot \alpha \cdot \alpha = \alpha^3.$$



Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.

3. Si calcoli il tasso medio di variazione tra il 2012 e il 2015 e si commenti il risultato ottenuto. Si ricorda che

$${}_{2012}I_{2015} = {}_{2012}I_{2013} \cdot {}_{2013}I_{2014} \cdot {}_{2014}I_{2015}$$

Si deve trovare il numero indice a base mobile medio. Applicando il criterio di scelta della media secondo Chisini si ottiene

$${}_{2012}I_{2015} = {}_{2012}I_{2013} \cdot {}_{2013}I_{2014} \cdot {}_{2014}I_{2015} = \alpha \cdot \alpha \cdot \alpha = \alpha^3.$$

Quindi

$$\alpha = \sqrt[3]{{}_{2012}I_{2015}} = \sqrt[3]{{}_{2012}I_{2013} \cdot {}_{2013}I_{2014} \cdot {}_{2014}I_{2015}} = 1.4422.$$

Da cui segue una variazione relativa media pari a 0.4422, corrispondente a un incremento medio annuo tra il 2012 e il 2015 del 44.22%.

Esercizio 3



Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.

Il seguente prospetto riporta i valori rilevati in corrispondenza a 5 unità statistiche con riferimento a due caratteri quantitativi X e Y tranne che per l'osservazione y_3

x_i	1	1	2	3	3
y_i	2	8		3	7

1. Si rappresentino graficamente le coppie dei dati, con anche il valore di y_3 tale per cui $\eta_{Y|X}^2 = 0$ ed $\eta_{X|Y}^2 = 1$. Si completi il grafico con la funzione di regressione ipotizzando Y in funzione di X .



Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.

Il seguente prospetto riporta i valori rilevati in corrispondenza a 5 unità statistiche con riferimento a due caratteri quantitativi X e Y tranne che per l'osservazione y_3

x_i	1	1	2	3	3
y_i	2	8		3	7

1. Si rappresentino graficamente le coppie dei dati, con anche il valore di y_3 tale per cui $\eta_{Y|X}^2 = 0$ ed $\eta_{X|Y}^2 = 1$. Si completi il grafico con la funzione di regressione ipotizzando Y in funzione di X .

$\eta_{Y|X}^2 = 0$ solo se sussiste indipendenza in media di Y da X , vale a dire se $M(Y|x = 1) = M(Y|x = 2) = M(Y|x = 3) = 5$.



Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.

Il seguente prospetto riporta i valori rilevati in corrispondenza a 5 unità statistiche con riferimento a due caratteri quantitativi X e Y tranne che per l'osservazione y_3

x_i	1	1	2	3	3
y_i	2	8		3	7

1. Si rappresentino graficamente le coppie dei dati, con anche il valore di y_3 tale per cui $\eta_{Y|X}^2 = 0$ ed $\eta_{X|Y}^2 = 1$. Si completi il grafico con la funzione di regressione ipotizzando Y in funzione di X .

$\eta_{Y|X}^2 = 0$ solo se sussiste indipendenza in media di Y da X , vale a dire se $M(Y|x=1) = M(Y|x=2) = M(Y|x=3) = 5$.

Segue che $y_3 =$



Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.

Il seguente prospetto riporta i valori rilevati in corrispondenza a 5 unità statistiche con riferimento a due caratteri quantitativi X e Y tranne che per l'osservazione y_3

x_i	1	1	2	3	3
y_i	2	8		3	7

1. Si rappresentino graficamente le coppie dei dati, con anche il valore di y_3 tale per cui $\eta_{Y|X}^2 = 0$ ed $\eta_{X|Y}^2 = 1$. Si completi il grafico con la funzione di regressione ipotizzando Y in funzione di X .

$\eta_{Y|X}^2 = 0$ solo se sussiste indipendenza in media di Y da X , vale a dire se $M(Y|x=1) = M(Y|x=2) = M(Y|x=3) = 5$.

Segue che $y_3 = 5$.



Il seguente prospetto riporta i valori rilevati in corrispondenza a 5 unità statistiche con riferimento a due caratteri quantitativi X e Y tranne che per l'osservazione y_3

x_i	1	1	2	3	3
y_i	2	8		3	7

1. Si rappresentino graficamente le coppie dei dati, con anche il valore di y_3 tale per cui $\eta_{Y|X}^2 = 0$ ed $\eta_{X|Y}^2 = 1$. Si completi il grafico con la funzione di regressione ipotizzando Y in funzione di X .

$\eta_{Y|X}^2 = 0$ solo se sussiste indipendenza in media di Y da X , vale a dire se $M(Y|x=1) = M(Y|x=2) = M(Y|x=3) = 5$.

Segue che $y_3 = 5$.

$\eta_{X|Y}^2 = 1$ in quanto sussiste dipendenza funzionale di Y da X qualsiasi sia il valore assegnato a y_3 .



La funzione di regressione associa a ogni valore x appartenente all'insieme di definizione della variabile X la media di Y condizionata a x , $M(Y|x)$

x	$M(Y x)$
1	$\frac{1}{2}(2 + 8) = 5$
2	5
3	$\frac{1}{2}(3 + 7) = 5$

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.



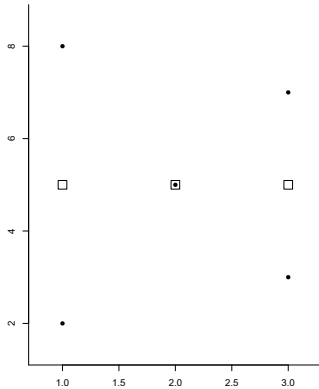
Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.

Le coppie di dati possono essere rappresentate mediante un grafico a dispersione.





2. Si calcolino, secondo il metodo dei minimi quadrati, i parametri e l'indice di adattamento dei seguenti modelli di regressione:

$$I) Y^* = a + b \cdot X \qquad II) Y^* = c \cdot X$$

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.

Si riportano i calcoli necessari¹ per ottenere le stime dei parametri e gli indici di adattamento del modello

$$Y = a + bX + E$$

x_i	y_i	$x_i y_i$	x_i^2	y_i^2
1	2	2	1	4
1	8	8	1	64
2	5	10	4	25
3	3	9	9	9
3	7	21	9	49
10	25	50	24	151
2	5	10	4.8	30.2
$M(X)$	$M(Y)$	$M(XY)$	$M(X^2)$	$M(Y^2)$

$$\hat{b} =$$



Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.

¹Si osserva che per ragioni di stampa alcuni calcoli potrebbero essere duplicati con riferimento ai diversi modelli oggetto di studio

Si riportano i calcoli necessari¹ per ottenere le stime dei parametri e gli indici di adattamento del modello

$$Y = a + bX + E$$

x_i	y_i	$x_i y_i$	x_i^2	y_i^2
1	2	2	1	4
1	8	8	1	64
2	5	10	4	25
3	3	9	9	9
3	7	21	9	49
10	25	50	24	151
2	5	10	4.8	30.2
$M(X)$	$M(Y)$	$M(XY)$	$M(X^2)$	$M(Y^2)$

$$\begin{aligned}\hat{b} &= \frac{\text{Cov}(X, Y)}{\text{Var}(X)} = \frac{M(XY) - M(X)M(Y)}{M(X^2) - [M(X)]^2} = \\ &= \frac{10 - 2 \cdot 5}{4.8 - 2^2} = \frac{0}{0.8} = 0 \\ \hat{a} &= \end{aligned}$$

¹Si osserva che per ragioni di stampa alcuni calcoli potrebbero essere duplicati con riferimento ai diversi modelli oggetto di studio



Si riportano i calcoli necessari¹ per ottenere le stime dei parametri e gli indici di adattamento del modello

$$Y = a + bX + E$$

x_i	y_i	$x_i y_i$	x_i^2	y_i^2
1	2	2	1	4
1	8	8	1	64
2	5	10	4	25
3	3	9	9	9
3	7	21	9	49
10	25	50	24	151
2	5	10	4.8	30.2
$M(X)$	$M(Y)$	$M(XY)$	$M(X^2)$	$M(Y^2)$

$$\hat{b} = \frac{\text{Cov}(X, Y)}{\text{Var}(X)} = \frac{M(XY) - M(X)M(Y)}{M(X^2) - [M(X)]^2} =$$

$$= \frac{10 - 2 \cdot 5}{4.8 - 2^2} = \frac{0}{0.8} = 0$$

$$\hat{a} = M(Y) - \hat{b}M(X) = 5$$

$$\rho^2 =$$



Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.

¹Si osserva che per ragioni di stampa alcuni calcoli potrebbero essere duplicati con riferimento ai diversi modelli oggetto di studio

Si riportano i calcoli necessari¹ per ottenere le stime dei parametri e gli indici di adattamento del modello

$$Y = a + bX + E$$

x_i	y_i	$x_i y_i$	x_i^2	y_i^2
1	2	2	1	4
1	8	8	1	64
2	5	10	4	25
3	3	9	9	9
3	7	21	9	49
10	25	50	24	151
2	5	10	4.8	30.2
$M(X)$	$M(Y)$	$M(XY)$	$M(X^2)$	$M(Y^2)$

$$\hat{b} = \frac{\text{Cov}(X, Y)}{\text{Var}(X)} = \frac{M(XY) - M(X)M(Y)}{M(X^2) - [M(X)]^2} =$$

$$= \frac{10 - 2 \cdot 5}{4.8 - 2^2} = \frac{0}{0.8} = 0$$

$$\hat{a} = M(Y) - \hat{b}M(X) = 5$$

$$\rho^2 = \frac{[\text{Cov}(X, Y)]^2}{\text{Var}(X) \text{Var}(Y)} = \frac{0^2}{0.8 \cdot 5.2} = 0$$

$$M(\hat{E}^2) = \text{Var}(\hat{E}) = \text{varianza residua} =$$

¹Si osserva che per ragioni di stampa alcuni calcoli potrebbero essere duplicati con riferimento ai diversi modelli oggetto di studio



Si riportano i calcoli necessari¹ per ottenere le stime dei parametri e gli indici di adattamento del modello

$$Y = a + bX + E$$

x_i	y_i	$x_i y_i$	x_i^2	y_i^2
1	2	2	1	4
1	8	8	1	64
2	5	10	4	25
3	3	9	9	9
3	7	21	9	49
10	25	50	24	151
2	5	10	4.8	30.2
$M(X)$	$M(Y)$	$M(XY)$	$M(X^2)$	$M(Y^2)$

$$\hat{b} = \frac{\text{Cov}(X, Y)}{\text{Var}(X)} = \frac{M(XY) - M(X)M(Y)}{M(X^2) - [M(X)]^2} =$$

$$= \frac{10 - 2 \cdot 5}{4.8 - 2^2} = \frac{0}{0.8} = 0$$

$$\hat{a} = M(Y) - \hat{b}M(X) = 5$$

$$\rho^2 = \frac{[\text{Cov}(X, Y)]^2}{\text{Var}(X) \text{Var}(Y)} = \frac{0^2}{0.8 \cdot 5.2} = 0$$

$$M(\hat{E}^2) = \text{Var}(\hat{E}) = \text{varianza residua} = \text{Var}(Y) \cdot (1 - \rho^2) = 5.2$$

¹Si osserva che per ragioni di stampa alcuni calcoli potrebbero essere duplicati con riferimento ai diversi modelli oggetto di studio





Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.

Si riportano i calcoli necessari² per ottenere le stime dei parametri e la misura di adattamento del modello

$$Y = cX + E$$

x_i	y_i	$x_i y_i$	x_i^2	y_i^2
1	2	2	1	4
1	8	8	1	64
2	5	10	4	25
3	3	9	9	9
3	7	21	9	49
		50	24	151
		10	4.8	30.2
		$M(XY)$	$M(X^2)$	$M(Y^2)$

$$\hat{c} =$$

²Si osserva che per ragioni di stampa alcuni calcoli potrebbero essere duplicati con riferimento ai diversi modelli oggetto di studio



Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.

Si riportano i calcoli necessari² per ottenere le stime dei parametri e la misura di adattamento del modello

$$Y = cX + E$$

x_i	y_i	$x_i y_i$	x_i^2	y_i^2
1	2	2	1	4
1	8	8	1	64
2	5	10	4	25
3	3	9	9	9
3	7	21	9	49
		50	24	151
		10	4.8	30.2
		$M(XY)$	$M(X^2)$	$M(Y^2)$

$$\hat{c} = \frac{M(XY)}{M(X^2)} = \frac{10}{4.8} = 2.0833$$

$$M(\hat{E}^2) =$$

²Si osserva che per ragioni di stampa alcuni calcoli potrebbero essere duplicati con riferimento ai diversi modelli oggetto di studio



Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.

Si riportano i calcoli necessari² per ottenere le stime dei parametri e la misura di adattamento del modello

$$Y = cX + E$$

x_i	y_i	$x_i y_i$	x_i^2	y_i^2
1	2	2	1	4
1	8	8	1	64
2	5	10	4	25
3	3	9	9	9
3	7	21	9	49
		50	24	151
		10	4.8	30.2
		$M(XY)$	$M(X^2)$	$M(Y^2)$

$$\hat{c} = \frac{M(XY)}{M(X^2)} = \frac{10}{4.8} = 2.0833$$

$$M(\hat{E}^2) = M[Y \cdot (Y - \hat{c}X)] = M(Y^2) - \hat{c}M(XY) = 30.2 - 2.0833 \cdot 10 = 9.3667$$

²Si osserva che per ragioni di stampa alcuni calcoli potrebbero essere duplicati con riferimento ai diversi modelli oggetto di studio



$\hat{a} = 5$, $\hat{b} = 0$, indice adattamento $M(\hat{E}) = 5.2$ $\hat{c} = 2.0833$, indice adattamento
 $M(\hat{E}) = 9.3667$

Il modello migliore è $Y^* = 5$.

Dal momento che sussiste indipendenza in media di Y da X , essendo $\eta_{Y|X}^2 = 0$, abbiamo anche indipendenza lineare, $\rho^2 = \text{Cov}(X, Y) = b = 0$ e $a = M(Y)$.

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.



Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.

3. Sulla base del miglior modello si preveda il valore di Y in corrispondenza di $X = 3$.



3. Sulla base del miglior modello si preveda il valore di Y in corrispondenza di $X = 3$.

$$Y^*(3) = 5 + 0 \cdot 3 = 5.$$

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.

Esercizio 4

Lo sfondo di una pagina web viene ridefinito a ogni visualizzazione estraendo in maniera casuale e indipendente un'opera tra 8 quadri di Kandinskij, 8 di Mondrian e 4 di Delaunay.

1. Si calcoli la probabilità che un soggetto che effettua 5 visualizzazioni trovi come sfondo un quadro di Kandinskij in 3 visualizzazioni consecutive.



Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.

Esercizio 4



Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.

Lo sfondo di una pagina web viene ridefinito a ogni visualizzazione estraendo in maniera casuale e indipendente un'opera tra 8 quadri di Kandinskij, 8 di Mondrian e 4 di Delaunay.

1. Si calcoli la probabilità che un soggetto che effettua 5 visualizzazioni trovi come sfondo un quadro di Kandinskij in 3 visualizzazioni consecutive.

La probabilità della generica sequenza contenente 3 quadri di Kandinskij in 5 visualizzazioni risulta:

$$P(K \cap K \cap K \cap \bar{K} \cap \bar{K}) =$$

Esercizio 4



Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.

Lo sfondo di una pagina web viene ridefinito a ogni visualizzazione estraendo in maniera casuale e indipendente un'opera tra 8 quadri di Kandinskij, 8 di Mondrian e 4 di Delaunay.

1. Si calcoli la probabilità che un soggetto che effettua 5 visualizzazioni trovi come sfondo un quadro di Kandinskij in 3 visualizzazioni consecutive.

La probabilità della generica sequenza contenente 3 quadri di Kandinskij in 5 visualizzazioni risulta:

$$P(K \cap K \cap K \cap \bar{K} \cap \bar{K}) = 0.4 \cdot 0.4 \cdot 0.4 \cdot 0.6 \cdot 0.6 =$$

Esercizio 4



Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.

Lo sfondo di una pagina web viene ridefinito a ogni visualizzazione estraendo in maniera casuale e indipendente un'opera tra 8 quadri di Kandinskij, 8 di Mondrian e 4 di Delaunay.

1. Si calcoli la probabilità che un soggetto che effettua 5 visualizzazioni trovi come sfondo un quadro di Kandinskij in 3 visualizzazioni consecutive.

La probabilità della generica sequenza contenente 3 quadri di Kandinskij in 5 visualizzazioni risulta:

$$P(K \cap K \cap K \cap \bar{K} \cap \bar{K}) = 0.4 \cdot 0.4 \cdot 0.4 \cdot 0.6 \cdot 0.6 = 0.02304$$

e

$$P(3K \text{ consecutive in } 5 \text{ visual.}) =$$

Esercizio 4



Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.

Lo sfondo di una pagina web viene ridefinito a ogni visualizzazione estraendo in maniera casuale e indipendente un'opera tra 8 quadri di Kandinskij, 8 di Mondrian e 4 di Delaunay.

1. Si calcoli la probabilità che un soggetto che effettua 5 visualizzazioni trovi come sfondo un quadro di Kandinskij in 3 visualizzazioni consecutive.

La probabilità della generica sequenza contenente 3 quadri di Kandinskij in 5 visualizzazioni risulta:

$$P(K \cap K \cap K \cap \bar{K} \cap \bar{K}) = 0.4 \cdot 0.4 \cdot 0.4 \cdot 0.6 \cdot 0.6 = 0.02304$$

e

$$\begin{aligned} P(3K \text{ consecutive in 5 visual.}) &= P(K \cap K \cap K \cap \bar{K} \cap \bar{K}) + \\ &+ P(\bar{K} \cap K \cap K \cap K \cap \bar{K}) + \\ &+ P(\bar{K} \cap \bar{K} \cap K \cap K \cap K) = \\ &= \end{aligned}$$

Esercizio 4



Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.

Lo sfondo di una pagina web viene ridefinito a ogni visualizzazione estraendo in maniera casuale e indipendente un'opera tra 8 quadri di Kandinskij, 8 di Mondrian e 4 di Delaunay.

1. Si calcoli la probabilità che un soggetto che effettua 5 visualizzazioni trovi come sfondo un quadro di Kandinskij in 3 visualizzazioni consecutive.

La probabilità della generica sequenza contenente 3 quadri di Kandinskij in 5 visualizzazioni risulta:

$$P(K \cap K \cap K \cap \bar{K} \cap \bar{K}) = 0.4 \cdot 0.4 \cdot 0.4 \cdot 0.6 \cdot 0.6 = 0.02304$$

e

$$\begin{aligned} P(3K \text{ consecutive in 5 visual.}) &= P(K \cap K \cap K \cap \bar{K} \cap \bar{K}) + \\ &+ P(\bar{K} \cap K \cap K \cap K \cap \bar{K}) + \\ &+ P(\bar{K} \cap \bar{K} \cap K \cap K \cap K) = \\ &= 3 \cdot 0.4^3 \cdot 0.6^2 = \end{aligned}$$

Esercizio 4



Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.

Lo sfondo di una pagina web viene ridefinito a ogni visualizzazione estraendo in maniera casuale e indipendente un'opera tra 8 quadri di Kandinskij, 8 di Mondrian e 4 di Delaunay.

1. Si calcoli la probabilità che un soggetto che effettua 5 visualizzazioni trovi come sfondo un quadro di Kandinskij in 3 visualizzazioni consecutive.

La probabilità della generica sequenza contenente 3 quadri di Kandinskij in 5 visualizzazioni risulta:

$$P(K \cap K \cap K \cap \bar{K} \cap \bar{K}) = 0.4 \cdot 0.4 \cdot 0.4 \cdot 0.6 \cdot 0.6 = 0.02304$$

e

$$\begin{aligned} P(3K \text{ consecutive in 5 visual.}) &= P(K \cap K \cap K \cap \bar{K} \cap \bar{K}) + \\ &+ P(\bar{K} \cap K \cap K \cap K \cap \bar{K}) + \\ &+ P(\bar{K} \cap \bar{K} \cap K \cap K \cap K) = \\ &= 3 \cdot 0.4^3 \cdot 0.6^2 = 3 \cdot 0.02304 = 0.06912. \end{aligned}$$



2. Si calcoli la probabilità che un soggetto che effettua 5 visualizzazioni trovi come sfondo un quadro di Kandinskij in almeno 3 visualizzazioni.

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.



Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.

2. Si calcoli la probabilità che un soggetto che effettua 5 visualizzazioni trovi come sfondo un quadro di Kandinskij in almeno 3 visualizzazioni.

Sia X la variabile casuale 'n. di visualizzazioni con un quadro di Kandinskij nelle 5 prove indipendenti'. Abbiamo

$$X \sim$$



Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.

2. Si calcoli la probabilità che un soggetto che effettua 5 visualizzazioni trovi come sfondo un quadro di Kandinskij in almeno 3 visualizzazioni.

Sia X la variabile casuale 'n. di visualizzazioni con un quadro di Kandinskij nelle 5 prove indipendenti'. Abbiamo

$$X \sim \text{Bin}(n = 5, p = 0.4)$$

e

$$P(X \geq 3) =$$



2. Si calcoli la probabilità che un soggetto che effettua 5 visualizzazioni trovi come sfondo un quadro di Kandinskij in almeno 3 visualizzazioni.

Sia X la variabile casuale 'n. di visualizzazioni con un quadro di Kandinskij nelle 5 prove indipendenti'. Abbiamo

$$X \sim \text{Bin}(n = 5, p = 0.4)$$

e

$$P(X \geq 3) = \{P(X = 3) + P(X = 4) + P(X = 5)\} =$$

=



2. Si calcoli la probabilità che un soggetto che effettua 5 visualizzazioni trovi come sfondo un quadro di Kandinskij in almeno 3 visualizzazioni.

Sia X la variabile casuale 'n. di visualizzazioni con un quadro di Kandinskij nelle 5 prove indipendenti'. Abbiamo

$$X \sim \text{Bin}(n = 5, p = 0.4)$$

e

$$\begin{aligned} P(X \geq 3) &= \{P(X = 3) + P(X = 4) + P(X = 5)\} = \\ &= \binom{5}{3} 0.4^3 \cdot 0.6^2 + \binom{5}{4} 0.4^4 \cdot 0.6^1 + \binom{5}{5} 0.4^5 \cdot 0.6^0 = \\ &= \end{aligned}$$



2. Si calcoli la probabilità che un soggetto che effettua 5 visualizzazioni trovi come sfondo un quadro di Kandinskij in almeno 3 visualizzazioni.

Sia X la variabile casuale 'n. di visualizzazioni con un quadro di Kandinskij nelle 5 prove indipendenti'. Abbiamo

$$X \sim \text{Bin}(n = 5, p = 0.4)$$

e

$$\begin{aligned} P(X \geq 3) &= \{P(X = 3) + P(X = 4) + P(X = 5)\} = \\ &= \binom{5}{3} 0.4^3 \cdot 0.6^2 + \binom{5}{4} 0.4^4 \cdot 0.6^1 + \binom{5}{5} 0.4^5 \cdot 0.6^0 = \\ &= 0.2304 + 0.0768 + 0.0102 = \end{aligned}$$



2. Si calcoli la probabilità che un soggetto che effettua 5 visualizzazioni trovi come sfondo un quadro di Kandinskij in almeno 3 visualizzazioni.

Sia X la variabile casuale 'n. di visualizzazioni con un quadro di Kandinskij nelle 5 prove indipendenti'. Abbiamo

$$X \sim \text{Bin}(n = 5, p = 0.4)$$

e

$$\begin{aligned} P(X \geq 3) &= \{P(X = 3) + P(X = 4) + P(X = 5)\} = \\ &= \binom{5}{3} 0.4^3 \cdot 0.6^2 + \binom{5}{4} 0.4^4 \cdot 0.6^1 + \binom{5}{5} 0.4^5 \cdot 0.6^0 = \\ &= 0.2304 + 0.0768 + 0.0102 = 0.3174. \end{aligned}$$



3. Si calcoli la probabilità che un quadro di Kandinskij figuri come sfondo in almeno 31 visualizzazioni su 100.

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.



Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.

3. Si calcoli la probabilità che un quadro di Kandinskij figuri come sfondo in almeno 31 visualizzazioni su 100.

Indicata con X la variabile casuale 'n. di visualizzazioni con un quadro di Kandinskij nelle 100 prove indipendenti' risulta

$$X \sim$$



Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.

3. Si calcoli la probabilità che un quadro di Kandinskij figuri come sfondo in almeno 31 visualizzazioni su 100.

Indicata con X la variabile casuale 'n. di visualizzazioni con un quadro di Kandinskij nelle 100 prove indipendenti' risulta

$$X \sim \text{Bin}(n = 100, p = 0.4).$$

$$M(X) = 40; \text{Var}(X) = 24.$$



Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.

3. Si calcoli la probabilità che un quadro di Kandinskij figuri come sfondo in almeno 31 visualizzazioni su 100.

Indicata con X la variabile casuale 'n. di visualizzazioni con un quadro di Kandinskij nelle 100 prove indipendenti' risulta

$$X \sim \text{Bin}(n = 100, p = 0.4).$$

$$M(X) = 40; \text{Var}(X) = 24.$$

Dal momento che $np = 40 > 5$ ed $n(1 - p) = 60 > 5$ è possibile utilizzare l'approssimazione della variabile casuale Binomiale con la variabile casuale Normale

$$X \sim$$



Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.

3. Si calcoli la probabilità che un quadro di Kandinskij figuri come sfondo in almeno 31 visualizzazioni su 100.

Indicata con X la variabile casuale 'n. di visualizzazioni con un quadro di Kandinskij nelle 100 prove indipendenti' risulta

$$X \sim \text{Bin}(n = 100, p = 0.4).$$

$$M(X) = 40; \text{Var}(X) = 24.$$

Dal momento che $np = 40 > 5$ ed $n(1 - p) = 60 > 5$ è possibile utilizzare l'approssimazione della variabile casuale Binomiale con la variabile casuale Normale

$$X \sim N(\mu = 40, \sigma^2 = 24).$$

Si ottiene

$$P(X \geq 31) =$$



3. Si calcoli la probabilità che un quadro di Kandinskij figuri come sfondo in almeno 31 visualizzazioni su 100.

Indicata con X la variabile casuale 'n. di visualizzazioni con un quadro di Kandinskij nelle 100 prove indipendenti' risulta

$$X \sim \text{Bin}(n = 100, p = 0.4).$$

$$M(X) = 40; \text{Var}(X) = 24.$$

Dal momento che $np = 40 > 5$ ed $n(1 - p) = 60 > 5$ è possibile utilizzare l'approssimazione della variabile casuale Binomiale con la variabile casuale Normale

$$X \sim N(\mu = 40, \sigma^2 = 24).$$

Si ottiene

$$P(X \geq 31) = P\left(\frac{X - \mu}{\sigma} \geq \frac{31 - \mu}{\sigma}\right) =$$



3. Si calcoli la probabilità che un quadro di Kandinskij figuri come sfondo in almeno 31 visualizzazioni su 100.

Indicata con X la variabile casuale 'n. di visualizzazioni con un quadro di Kandinskij nelle 100 prove indipendenti' risulta

$$X \sim \text{Bin}(n = 100, p = 0.4).$$

$$M(X) = 40; \text{Var}(X) = 24.$$

Dal momento che $np = 40 > 5$ ed $n(1 - p) = 60 > 5$ è possibile utilizzare l'approssimazione della variabile casuale Binomiale con la variabile casuale Normale

$$X \sim N(\mu = 40, \sigma^2 = 24).$$

Si ottiene

$$P(X \geq 31) = P\left(\frac{X - \mu}{\sigma} \geq \frac{31 - \mu}{\sigma}\right) = P\left(Z \geq \frac{31 - 40}{4.899}\right) =$$

$$=$$



3. Si calcoli la probabilità che un quadro di Kandinskij figuri come sfondo in almeno 31 visualizzazioni su 100.

Indicata con X la variabile casuale 'n. di visualizzazioni con un quadro di Kandinskij nelle 100 prove indipendenti' risulta

$$X \sim \text{Bin}(n = 100, p = 0.4).$$

$$M(X) = 40; \text{Var}(X) = 24.$$

Dal momento che $np = 40 > 5$ ed $n(1 - p) = 60 > 5$ è possibile utilizzare l'approssimazione della variabile casuale Binomiale con la variabile casuale Normale

$$X \sim N(\mu = 40, \sigma^2 = 24).$$

Si ottiene

$$\begin{aligned} P(X \geq 31) &= P\left(\frac{X - \mu}{\sigma} \geq \frac{31 - \mu}{\sigma}\right) = P\left(Z \geq \frac{31 - 40}{4.899}\right) = \\ &= P(Z \geq -1.84) = P(Z \leq 1.84) = 0.9671. \end{aligned}$$

Tema T306

Prova scritta di STATISTICA del 15.01.2020

Prova scritta di
STATISTICA del
15.01.2020

Giuseppe Boari
Gabriele Cantaluppi



UNIVERSITÀ
CATTOLICA
del Sacro Cuore

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.

Giuseppe Boari
Gabriele Cantaluppi
Università Cattolica del Sacro Cuore, Milano



La seguente tabella riporta le distribuzioni relative condizionate – ricavate a seguito di una indagine campionaria condotta su $n = 150$ automobilisti – della variabile $Y = \text{Ore al giorno trascorse alla guida della macchina}$ rispetto alla variabile $X = \text{Utilizzo delle cinture di sicurezza}$:

$Y X$	0 ÷ 2	2 ÷ 5	5 ÷ 7	
$sì$	0.2	0.3	0.5	1
no	0.6	0.3	0.1	1

1. Dopo avere rappresentato graficamente la variabile $Y|X = sì$, se ne calcoli la moda. È possibile dire che la moda coincide con la mediana di $Y|X = sì$?

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.



Si considera la distribuzione condizionata $Y|x_1$

y_j	$f.cond.r.1j$
0 $\rightarrow$ 2	0.2
2 $\rightarrow$ 5	0.3
5 $\rightarrow$ 7	0.5

Esercizio 1.

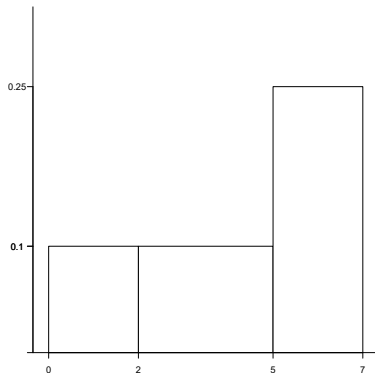
Esercizio 2.

Esercizio 3.

Esercizio 4.

Si tratta di un carattere quantitativo con modalità raggruppate in classi che può essere rappresentato graficamente mediante un istogramma. L'istogramma può essere costruito dopo avere calcolato le densità $d_{1j} = \frac{f.\text{cond}.r.1j}{a_j}$, (dove a_j sono le ampiezze di classe).

y_j	$f.\text{cond}.r.1j$	a_{1j}	d_{1j}
0 - 2	0.2	$2 - 0 = 2$	0.1
2 - 5	0.3	$5 - 2 = 3$	0.1
5 - 7	0.5	$7 - 5 = 2$	0.25



Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.



Si tratta di un carattere quantitativo con modalità raggruppate in classi.

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.



La moda corrisponde alla classe alla quale è associata la massima densità di frequenza

y_j	n_{1j}	a_j	d_{1j}
$0 \vdash 2$	0.2	$2 - 0 = 2$	0.1
$2 \vdash 5$	0.3	$5 - 2 = 3$	0.1
$5 \vdash 7$	0.5	$7 - 5 = 2$	0.25

e risulta $5 \vdash 7$.

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.



La mediana (percentile di ordine 0.50) può essere ottenuta considerando l'inversa della Funzione di Ripartizione in corrispondenza di 0.50.

y_j	$f.cond.r._{1j}$	$F.cond.r._{1j}$
0 $\rightarrow$ 2	0.2	0.2
2 $\rightarrow$ 5	0.3	0.5
5 $\rightarrow$ 7	0.5	1

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.

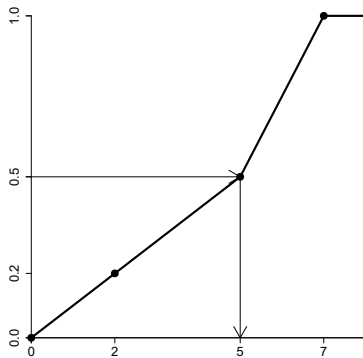


Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.



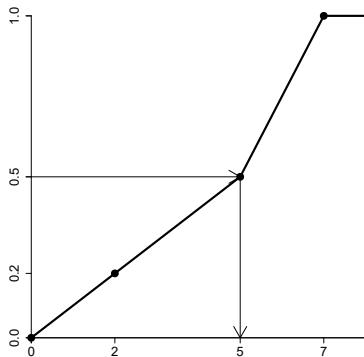


Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.



Si ottiene

$$y_{0.5} = 2 + \frac{0.5 - 0.2}{0.5 - 0.2} \cdot (5 - 2) = 5.$$



La mediana $y_{0.5} = 5$ coincide, quindi, con l'estremo inferiore della classe modale $5 \div 7$.

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.



2. Costruire la tabella delle frequenze congiunte sapendo che il 60% dei lavoratori intervistati usa la cintura di sicurezza.

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.



Moltiplicando le frequenze condizionate relative $f.cond.r_{1j} = \frac{n_{1j}}{n_{1\bullet}}$ per il numero totale di osservazioni $n_{1\bullet}$ della distribuzione condizionata, si ottengono le frequenze assolute congiunte n_{1j} della distribuzione condizionata $Y|X_1$.

y_j	$f.cond.r_{1j}$	n_{1j}
0 + 2	0.2	18
2 + 5	0.3	27
5 + 7	0.5	45

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.



Moltiplicando le frequenze condizionate relative $f.cond.r_{2j} = \frac{n_{2j}}{n_{2\bullet}}$ per il numero totale di osservazioni $n_{2\bullet}$ della distribuzione condizionata, si ottengono le frequenze assolute congiunte n_{2j} della distribuzione condizionata $Y|X_2$.

y_j	$f.cond.r_{2j}$	n_{2j}
0 + 2	0.6	36
2 + 5	0.3	18
5 + 7	0.1	6

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.



La distribuzione delle frequenze congiunte risulta

$X \setminus Y$	0 - 2	2 - 5	5 - 7	$n_{i\bullet}$
$sì$	18	27	45	90
no	36	18	6	60
$n_{\bullet j}$	54	45	51	150

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.



3. Confrontare, con opportuni indici di posizione e di variabilità, la distribuzione della variabile Y (ore al giorno trascorse alla guida) nei due gruppi di utilizzatori ($X = sì$) e non utilizzatori ($X = no$) della cintura di sicurezza. Commentare il risultato ottenuto.

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.



Si considera la distribuzione condizionata $Y|x_1$

y_j	n_{1j}
0 + 2	18
2 + 5	27
5 + 7	45

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.



Si tratta di un carattere quantitativo con modalità raggruppate in classi.

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.



La media aritmetica

$$M(Y|x_1) = \frac{1}{n_{1\bullet}} \sum_{j=1}^k y_j n_{1j}$$

risulta 4.25.

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.



La media aritmetica

$$M(Y|x_1) = \frac{1}{n_{1\bullet}} \sum_{j=1}^k y_j n_{1j}$$

risulta 4.25.

Si ottiene (considerando come modalità i valori centrali delle classi)

y_j	n_{1j}	$y_j n_{1j}$
1.0	18	18.0
3.5	27	94.5
6.0	45	270.0
	90	382.5

da cui segue

$$M(Y|x_1) = \frac{1}{n_{1\bullet}} \sum_{j=1}^k y_j n_{1j} = \frac{382.5}{90} = 4.25.$$

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.



La varianza

$$\text{Var}(Y|x_1) = M(Y^2|x_1) - [M(Y|x_1)]^2$$

risulta 3.8125.

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.



Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.

La varianza

$$Var(Y|x_1) = M(Y^2|x_1) - [M(Y|x_1)]^2$$

risulta 3.8125. Si ottiene (considerando come modalità i valori centrali delle classi)

y_j	n_{1j}	$y_j n_{1j}$	$y_j^2 n_{1j}$
1.0	18	18.0	18.00
3.5	27	94.5	330.75
6.0	45	270.0	1620.00
	90	382.5	1968.75

da cui segue

$$Var(Y|x_1) = \frac{1968.75}{90} - 4.25^2 = 21.875 - 18.0625 = 3.8125.$$



Per effettuare confronti in termini di variabilità tra caratteri quantitativi occorre utilizzare una misura relativa, ad esempio il coefficiente di variazione

$$CV = \frac{\sigma}{\mu} = \frac{\sqrt{\text{Var}(Y|x_1)}}{M(Y|x_1)} = \frac{\sqrt{3.8125}}{4.25} = \frac{1.9526}{4.25} = 0.4594.$$

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.



Si considera la distribuzione condizionata $Y|x_2$

y_j	n_{2j}
0 + 2	36
2 + 5	18
5 + 7	6

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.



Si tratta di un carattere quantitativo con modalità raggruppate in classi.

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.



La media aritmetica

$$M(Y|x_2) = \frac{1}{n_{2\bullet}} \sum_{j=1}^k y_j n_{2j}$$

risulta 2.25.

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.



La media aritmetica

$$M(Y|x_2) = \frac{1}{n_{2\bullet}} \sum_{j=1}^k y_j n_{2j}$$

risulta 2.25.

Si ottiene (considerando come modalità i valori centrali delle classi)

y_j	n_{2j}	$y_j n_{2j}$
1.0	36	36
3.5	18	63
6.0	6	36
	60	135

da cui segue

$$M(Y|x_2) = \frac{1}{n_{2\bullet}} \sum_{j=1}^k y_j n_{2j} = \frac{135}{60} = 2.25.$$

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.



La varianza

$$\text{Var}(Y|x_2) = M(Y^2|x_2) - [M(Y|x_2)]^2$$

risulta 2.8125.

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.



La varianza

$$Var(Y|x_2) = M(Y^2|x_2) - [M(Y|x_2)]^2$$

risulta 2.8125. Si ottiene (considerando come modalità i valori centrali delle classi)

y_j	n_{2j}	$y_j n_{2j}$	$y_j^2 n_{2j}$
1.0	36	36	36.0
3.5	18	63	220.5
6.0	6	36	216.0
	60	135	472.5

da cui segue

$$Var(Y|x_2) = \frac{472.5}{60} - 2.25^2 = 7.875 - 5.0625 = 2.8125.$$



Per effettuare confronti in termini di variabilità tra caratteri quantitativi occorre utilizzare una misura relativa, ad esempio il coefficiente di variazione

$$CV = \frac{\sigma}{\mu} = \frac{\sqrt{\text{Var}(Y|x_2)}}{M(Y|x_2)} = \frac{\sqrt{2.8125}}{2.25} = \frac{1.6771}{2.25} = 0.7454.$$

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.



$Y|X = sì$ Indice di posizione: $M(Y|X = sì)$ valore 4.25 Indice di variabilità: $CV(Y|X = sì)$ valore 0.4594

$Y|X = no$ Indice di posizione: $M(Y|X = no)$ valore 2.25 Indice di variabilità: $CV(Y|X = no)$ valore 0.7454

Il numero medio di ore al giorno trascorse alla guida di coloro che utilizzano la cintura di sicurezza è superiore a quello di coloro che non la utilizzano. Il numero medio di ore al giorno trascorse alla guida di coloro che non utilizzano la cintura di sicurezza presenta un livello di variabilità superiore a quello di coloro che la utilizzano.

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.



4. Senza effettuare calcoli e motivando la risposta, dire se l'indice di mutabilità normalizzato di X assume un valore più vicino a 0 o a 1.

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.



F_N è più vicino a 1 essendo le frequenze marginali

$$f_{1\bullet} = \frac{90}{150} = 0.6$$

e

$$f_{2\bullet} = \frac{60}{150} = 0.4$$

prossime alla situazione di equidistribuzione

$$f_{1\bullet} = f_{2\bullet} = 0.50.$$

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.



Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.

5. Calcolare un indice di connessione tra X e Y e commentare il risultato ottenuto.



Dalla distribuzione delle frequenze congiunte n_{ij} :

	0 + 2	2 + 5	5 + 7	$n_{i\bullet}$
x_1	18	27	45	90
x_2	36	18	6	60
$n_{\bullet j}$	54	45	51	150

è possibile ottenere le frequenze teoriche nella situazione di indipendenza stocastica

$$\hat{n}_{ij} = \frac{n_{i\bullet} n_{\bullet j}}{n}:$$

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.



Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.

Dalla distribuzione delle frequenze congiunte n_{ij} :

	0 + 2	2 + 5	5 + 7	$n_{i\bullet}$
x_1	18	27	45	90
x_2	36	18	6	60
$n_{\bullet j}$	54	45	51	150

è possibile ottenere le frequenze teoriche nella situazione di indipendenza stocastica

$$\hat{n}_{ij} = \frac{n_{i\bullet} n_{\bullet j}}{n}:$$

	0 + 2	2 + 5	5 + 7	$n_{i\bullet}$
x_1	32.4	27	30.6	90
x_2	21.6	18	20.4	60
$n_{\bullet j}$	54.0	45	51.0	150



le contingenze assolute $n_{ij} - \hat{n}_{ij}$:

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.



le contingenze assolute $n_{ij} - \hat{n}_{ij}$:

	0 ÷ 2	2 ÷ 5	5 ÷ 7
x_1	-14.4	0	14.4
x_2	14.4	0	-14.4

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.



i rapporti di contingenza $\frac{n_{ij}}{\hat{n}_{ij}}$:

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.



i rapporti di contingenza $\frac{n_{ij}}{\hat{n}_{ij}}$:

	0 + 2	2 + 5	5 + 7
x_1	0.5556	1	1.4706
x_2	1.6667	1	0.2941

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.



e le contingenze relative, espresse in termini percentuali, $100 \left(\frac{n_{ij} - \hat{n}_{ij}}{\hat{n}_{ij}} \right) \% = 100 \left(\frac{n_{ij}}{\hat{n}_{ij}} - 1 \right) \%$:

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.



e le contingenze relative, espresse in termini percentuali, $100 \left(\frac{n_{ij} - \hat{n}_{ij}}{\hat{n}_{ij}} \right) \% = 100 \left(\frac{n_{ij}}{\hat{n}_{ij}} - 1 \right) \%$:

	0 ÷ 2	2 ÷ 5	5 ÷ 7
x_1	-44.44	0	47.06
x_2	66.67	0	-70.59

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.



Risultati intermedi $\frac{(n_{ij} - \hat{n}_{ij})^2}{\hat{n}_{ij}}$ per il calcolo del χ^2

	0 ÷ 2	2 ÷ 5	5 ÷ 7
x_1	6.4	0	6.7765
x_2	9.6	0	10.1647

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.



Risultati intermedi $\frac{n_{ij}^2}{n_{i\bullet} \cdot n_{\bullet j}}$ per il calcolo del χ^2 (formula operativa)

	0 ÷ 2	2 ÷ 5	5 ÷ 7
x_1	0.0667	0.18	0.4412
x_2	0.4000	0.12	0.0118

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.



Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.

$$\chi^2 = \sum_{i=1}^h \sum_{j=1}^k \frac{(n_{ij} - \hat{n}_{ij})^2}{\hat{n}_{ij}} = 32.9412,$$

$$\chi_N^2 = \frac{\chi^2}{n \cdot \min(h-1, k-1)} = 0.2196$$

$$\chi_N = \sqrt{\chi_N^2} = 0.4686.$$

Sussiste, quindi, un livello di connessione medio.



Si consideri il seguente prospetto che riporta i $NIBF$ (base=2015) di un bene dal 2015 al 2018:

t	$NIBF(base = 2015)$	$NIBM$
2015	1	
2016	1	
2017	1.1	
2018	0.96	
2019		

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.



1. Ricostruire la serie dei *NIBM* dal 2015 al 2018.

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.



1. Ricostruire la serie dei *NIBM* dal 2015 al 2018.

Si ricorda che ${}_t l_t = \frac{x_t}{x_{t-1}} = \frac{{}_{2015}l_t}{{}_{2015}l_{t-1}}$.

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.



1. Ricostruire la serie dei *NIBM* dal 2015 al 2018.

Si ricorda che $_{t-1}I_t = \frac{x_t}{x_{t-1}} = \frac{_{2015}I_t}{_{2015}I_{t-1}}$. Si ottiene quindi

Anno	NIBF	NIBM
2015	1.00	
2016	1.00	1.0000
2017	1.10	1.1000
2018	0.96	0.8727

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.



2. Sapendo che il bene ha subito un decremento medio annuo del 10% dal 2015 al 2019, completare la tabella con il valore dei numeri indici a base fissa e a base mobile al tempo 2019.

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.



2. Sapendo che il bene ha subito un decremento medio annuo del 10% dal 2015 al 2019, completare la tabella con il valore dei numeri indici a base fissa e a base mobile al tempo 2019.

Un decremento medio del 10% corrisponde a una misura relativa media dello 0.90; quindi:

$${}_{2015}I_{2019} =$$

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.



2. Sapendo che il bene ha subito un decremento medio annuo del 10% dal 2015 al 2019, completare la tabella con il valore dei numeri indici a base fissa e a base mobile al tempo 2019.

Un decremento medio del 10% corrisponde a una misura relativa media dello 0.90; quindi:

$${}_{2015}I_{2019} = 0.90^4 =$$

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.



2. Sapendo che il bene ha subito un decremento medio annuo del 10% dal 2015 al 2019, completare la tabella con il valore dei numeri indici a base fissa e a base mobile al tempo 2019.

Un decremento medio del 10% corrisponde a una misura relativa media dello 0.90; quindi:

$${}_{2015}I_{2019} = 0.90^4 = 0.6561.$$

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.



3. Calcolare il valore del bene nel 2019, sapendo che nel 2018 è risultato pari a 21.

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.



Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.

3. Calcolare il valore del bene nel 2019, sapendo che nel 2018 è risultato pari a 21.
Vale

$${}_{2018}I_{2019} =$$



3. Calcolare il valore del bene nel 2019, sapendo che nel 2018 è risultato pari a 21.
Vale

$${}_{2018}I_{2019} = \frac{{}_{2015}I_{2019}}{{}_{2015}I_{2018}} =$$

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.



Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.

3. Calcolare il valore del bene nel 2019, sapendo che nel 2018 è risultato pari a 21.
Vale

$${}_{2018}I_{2019} = \frac{{}_{2015}I_{2019}}{{}_{2015}I_{2018}} = \frac{0.6561}{0.96} =$$



Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.

3. Calcolare il valore del bene nel 2019, sapendo che nel 2018 è risultato pari a 21.
Vale

$${}_{2018}I_{2019} = \frac{{}_{2015}I_{2019}}{{}_{2015}I_{2018}} = \frac{0.6561}{0.96} = \frac{x_{2019}}{x_{2018}} =$$



Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.

3. Calcolare il valore del bene nel 2019, sapendo che nel 2018 è risultato pari a 21.
Vale

$${}_{2018}I_{2019} = \frac{{}_{2015}I_{2019}}{{}_{2015}I_{2018}} = \frac{0.6561}{0.96} = \frac{x_{2019}}{x_{2018}} = \frac{x_{2019}}{21} = 0.6834.$$



Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.

3. Calcolare il valore del bene nel 2019, sapendo che nel 2018 è risultato pari a 21.
Vale

$${}_{2018}I_{2019} = \frac{{}_{2015}I_{2019}}{{}_{2015}I_{2018}} = \frac{0.6561}{0.96} = \frac{x_{2019}}{x_{2018}} = \frac{x_{2019}}{21} = 0.6834.$$

Da cui segue $x_{2019} = 14.3522$.



Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.

Su un insieme di 5 individui vengono rilevate le seguenti variabili: X = numero di regali di Natale, Y = importo speso per i regali (in centinaia di Euro) e G = genere (f = femmina, m = maschio). I dati sono riportati di seguito:

x_i	2	6	4	2	4
y_i	1	6	5	3	3
g_i	f	f	f	m	m

1. Costruire il grafico delle coppie dei dati (x_i, y_i) e della funzione di regressione della spesa in funzione del numero di regali.



La funzione di regressione associa a ogni valore x appartenente all'insieme di definizione della variabile X la media di Y condizionata a x , $M(Y|x)$

x	$M(Y x)$
2	$\frac{1}{2}(1 + 3) = 2$
4	$\frac{1}{2}(3 + 5) = 4$
6	6

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.



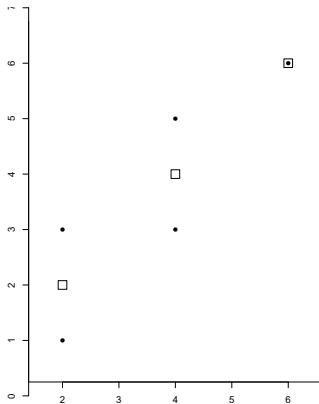
Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.

Le coppie di dati possono essere rappresentate mediante un grafico a dispersione.





Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.

2. Calcolare, secondo il metodo dei minimi quadrati, i parametri dei seguenti modelli di regressione:

$$I) Y^* = a + b \cdot X \quad II) Y^* = c + d \cdot W$$

con W variabile dummy, indicatrice del genere femmina, ($w_i = 0$ se $g_i = m$, $w_i = 1$ se $g_i = f$).



Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.

Si riportano i calcoli necessari¹ per ottenere le stime dei parametri e gli indici di adattamento del modello

$$Y = a + bX + E$$

x_i	y_i	$x_i y_i$	x_i^2	y_i^2
2	1	2	4	1
6	6	36	36	36
4	5	20	16	25
2	3	6	4	9
4	3	12	16	9
18	18	76	76	80
3.6	3.6	15.2	15.2	16
$M(X)$	$M(Y)$	$M(XY)$	$M(X^2)$	$M(Y^2)$

$$\hat{b} =$$

¹Si osserva che per ragioni di stampa alcuni calcoli potrebbero essere duplicati con riferimento ai diversi modelli oggetto di studio



Si riportano i calcoli necessari¹ per ottenere le stime dei parametri e gli indici di adattamento del modello

$$Y = a + bX + E$$

x_i	y_i	$x_i y_i$	x_i^2	y_i^2
2	1	2	4	1
6	6	36	36	36
4	5	20	16	25
2	3	6	4	9
4	3	12	16	9
18	18	76	76	80
3.6	3.6	15.2	15.2	16
$M(X)$	$M(Y)$	$M(XY)$	$M(X^2)$	$M(Y^2)$

$$\begin{aligned}\hat{b} &= \frac{\text{Cov}(X, Y)}{\text{Var}(X)} = \frac{M(XY) - M(X)M(Y)}{M(X^2) - [M(X)^2]} = \\ &= \frac{15.2 - 3.6 \cdot 3.6}{15.2 - 3.6^2} = \frac{2.24}{2.24} = 1 \\ \hat{a} &= \end{aligned}$$

¹Si osserva che per ragioni di stampa alcuni calcoli potrebbero essere duplicati con riferimento ai diversi modelli oggetto di studio



Si riportano i calcoli necessari¹ per ottenere le stime dei parametri e gli indici di adattamento del modello

$$Y = a + bX + E$$

x_i	y_i	$x_i y_i$	x_i^2	y_i^2
2	1	2	4	1
6	6	36	36	36
4	5	20	16	25
2	3	6	4	9
4	3	12	16	9
18	18	76	76	80
3.6	3.6	15.2	15.2	16
$M(X)$	$M(Y)$	$M(XY)$	$M(X^2)$	$M(Y^2)$

$$\hat{b} = \frac{\text{Cov}(X, Y)}{\text{Var}(X)} = \frac{M(XY) - M(X)M(Y)}{M(X^2) - [M(X)^2]} =$$

$$= \frac{15.2 - 3.6 \cdot 3.6}{15.2 - 3.6^2} = \frac{2.24}{2.24} = 1$$

$$\hat{a} = M(Y) - \hat{b}M(X) = 0$$

$$\rho^2 =$$

¹Si osserva che per ragioni di stampa alcuni calcoli potrebbero essere duplicati con riferimento ai diversi modelli oggetto di studio



Si riportano i calcoli necessari¹ per ottenere le stime dei parametri e gli indici di adattamento del modello

$$Y = a + bX + E$$

x_i	y_i	$x_i y_i$	x_i^2	y_i^2
2	1	2	4	1
6	6	36	36	36
4	5	20	16	25
2	3	6	4	9
4	3	12	16	9
18	18	76	76	80
3.6	3.6	15.2	15.2	16
$M(X)$	$M(Y)$	$M(XY)$	$M(X^2)$	$M(Y^2)$

$$\hat{b} = \frac{\text{Cov}(X, Y)}{\text{Var}(X)} = \frac{M(XY) - M(X)M(Y)}{M(X^2) - [M(X)]^2} =$$

$$= \frac{15.2 - 3.6 \cdot 3.6}{15.2 - 3.6^2} = \frac{2.24}{2.24} = 1$$

$$\hat{a} = M(Y) - \hat{b}M(X) = 0$$

$$\rho^2 = \frac{[\text{Cov}(X, Y)]^2}{\text{Var}(X) \text{Var}(Y)} = \frac{2.24^2}{2.24 \cdot 3.04} = 0.7368$$

¹Si osserva che per ragioni di stampa alcuni calcoli potrebbero essere duplicati con riferimento ai diversi modelli oggetto di studio



Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.

Si riportano i calcoli necessari² per ottenere le stime dei parametri e gli indici di adattamento del modello

$$Y = c + dW + E$$

w_i	y_i	$w_i y_i$	w_i^2	y_i^2
1	1	1	1	1
1	6	6	1	36
1	5	5	1	25
0	3	0	0	9
0	3	0	0	9
3	18	12	3	80
0.6	3.6	2.4	0.6	16
$M(W)$	$M(Y)$	$M(WY)$	$M(W^2)$	$M(Y^2)$

$$\hat{d} =$$

²Si osserva che per ragioni di stampa alcuni calcoli potrebbero essere duplicati con riferimento ai diversi modelli oggetto di studio



Si riportano i calcoli necessari² per ottenere le stime dei parametri e gli indici di adattamento del modello

$$Y = c + dW + E$$

w_i	y_i	$w_i y_i$	w_i^2	y_i^2
1	1	1	1	1
1	6	6	1	36
1	5	5	1	25
0	3	0	0	9
0	3	0	0	9
3	18	12	3	80
0.6	3.6	2.4	0.6	16
$M(W)$	$M(Y)$	$M(WY)$	$M(W^2)$	$M(Y^2)$

$$\begin{aligned}\hat{d} &= \frac{\text{Cov}(W, Y)}{\text{Var}(W)} = \frac{M(WY) - M(W)M(Y)}{M(W^2) - [M(W)]^2} = \\ &= \frac{2.4 - 0.6 \cdot 3.6}{0.6 - 0.6^2} = \frac{0.24}{0.24} = 1\end{aligned}$$

$$\hat{c} =$$

²Si osserva che per ragioni di stampa alcuni calcoli potrebbero essere duplicati con riferimento ai diversi modelli oggetto di studio



Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.

Si riportano i calcoli necessari² per ottenere le stime dei parametri e gli indici di adattamento del modello

$$Y = c + dW + E$$

w_i	y_i	$w_i y_i$	w_i^2	y_i^2
1	1	1	1	1
1	6	6	1	36
1	5	5	1	25
0	3	0	0	9
0	3	0	0	9
3	18	12	3	80
0.6	3.6	2.4	0.6	16
$M(W)$	$M(Y)$	$M(WY)$	$M(W^2)$	$M(Y^2)$

$$\hat{d} = \frac{\text{Cov}(W, Y)}{\text{Var}(W)} = \frac{M(WY) - M(W)M(Y)}{M(W^2) - [M(W)]^2} =$$

$$= \frac{2.4 - 0.6 \cdot 3.6}{0.6 - 0.6^2} = \frac{0.24}{0.24} = 1$$

$$\hat{c} = M(Y) - \hat{d} M(W) = 3$$

$$\rho^2 =$$

²Si osserva che per ragioni di stampa alcuni calcoli potrebbero essere duplicati con riferimento ai diversi modelli oggetto di studio



Si riportano i calcoli necessari² per ottenere le stime dei parametri e gli indici di adattamento del modello

$$Y = c + dW + E$$

w_i	y_i	$w_i y_i$	w_i^2	y_i^2
1	1	1	1	1
1	6	6	1	36
1	5	5	1	25
0	3	0	0	9
0	3	0	0	9
3	18	12	3	80
0.6	3.6	2.4	0.6	16
$M(W)$	$M(Y)$	$M(WY)$	$M(W^2)$	$M(Y^2)$

$$\begin{aligned}\hat{d} &= \frac{\text{Cov}(W, Y)}{\text{Var}(W)} = \frac{M(WY) - M(W)M(Y)}{M(W^2) - [M(W)]^2} = \\ &= \frac{2.4 - 0.6 \cdot 3.6}{0.6 - 0.6^2} = \frac{0.24}{0.24} = 1 \\ \hat{c} &= M(Y) - \hat{d} M(W) = 3 \\ \rho^2 &= \frac{[\text{Cov}(W, Y)]^2}{\text{Var}(W) \text{Var}(Y)} = \frac{0.24^2}{0.24 \cdot 3.04} = 0.0789\end{aligned}$$

²Si osserva che per ragioni di stampa alcuni calcoli potrebbero essere duplicati con riferimento ai diversi modelli oggetto di studio



$\hat{a} = 0, \hat{b} = 1$, indice adattamento $\rho^2 = 0.7368$

$\hat{c} = 3, \hat{d} = 1$, indice adattamento $\rho^2 = 0.0789$

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.



3. Confrontare gli indici di adattamento, commentando i risultati.

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.



Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.

3. Confrontare gli indici di adattamento, commentando i risultati.

Il modello migliore è $Y^* = X$ in quanto caratterizzato da indice di adattamento più elevato.



4. Fornire una interpretazione delle stime dei parametri del modello II).

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.



Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.

4. Fornire una interpretazione delle stime dei parametri del modello II).

Abbiamo:

$$\hat{c} = M(Y|W = m) = 3,$$

e

$$d = M(Y|W = f) - M(Y|W = m) = 1 (= 4 - 3)$$



5. Senza effettuare calcoli e motivando la risposta, indicare il valore di $\eta_{Y|X}^2$.

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.



Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.

5. Senza effettuare calcoli e motivando la risposta, indicare il valore di $\eta^2_{Y|X}$.
 $\eta^2_{Y|X} = 0.7368$ in quanto le medie condizionate $M(Y|x = 2)$, $M(Y|x = 4)$ e $M(Y|x = 6)$ sono allineate.

Esercizio 4

Prova scritta di
STATISTICA del
15.01.2020

Giuseppe Boari
Gabriele Cantaluppi



UNIVERSITÀ
CATTOLICA
del Sacro Cuore

Una ricerca ha evidenziato che i tempi di gestazione delle donne in gravidanza seguono una distribuzione Normale con media $\mu = 270$ giorni e varianza $\sigma^2 = 900$.

1. Indicare quale è la probabilità che un bambino nasca dopo 274 giorni.

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.

Esercizio 4



Una ricerca ha evidenziato che i tempi di gestazione delle donne in gravidanza seguono una distribuzione Normale con media $\mu = 270$ giorni e varianza $\sigma^2 = 900$.

1. Indicare quale è la probabilità che un bambino nasca dopo 274 giorni.

Indicata con $X \sim N(\mu = 270, \sigma^2 = 900)$ la variabile casuale che descrive i tempi di gestazione delle donne in gravidanza, occorre determinare la seguente probabilità: $P(X > 274)$.

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.



Una ricerca ha evidenziato che i tempi di gestazione delle donne in gravidanza seguono una distribuzione Normale con media $\mu = 270$ giorni e varianza $\sigma^2 = 900$.

1. Indicare quale è la probabilità che un bambino nasca dopo 274 giorni.

Indicata con $X \sim N(\mu = 270, \sigma^2 = 900)$ la variabile casuale che descrive i tempi di gestazione delle donne in gravidanza, occorre determinare la seguente probabilità: $P(X > 274)$.

Abbiamo

$$P(X > 274) =$$



Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.

Una ricerca ha evidenziato che i tempi di gestazione delle donne in gravidanza seguono una distribuzione Normale con media $\mu = 270$ giorni e varianza $\sigma^2 = 900$.

1. Indicare quale è la probabilità che un bambino nasca dopo 274 giorni.

Indicata con $X \sim N(\mu = 270, \sigma^2 = 900)$ la variabile casuale che descrive i tempi di gestazione delle donne in gravidanza, occorre determinare la seguente probabilità: $P(X > 274)$.

Abbiamo

$$P(X > 274) = P\left(\frac{X - \mu}{\sigma} > \frac{274 - \mu}{\sigma}\right) =$$



Una ricerca ha evidenziato che i tempi di gestazione delle donne in gravidanza seguono una distribuzione Normale con media $\mu = 270$ giorni e varianza $\sigma^2 = 900$.

1. Indicare quale è la probabilità che un bambino nasca dopo 274 giorni.

Indicata con $X \sim N(\mu = 270, \sigma^2 = 900)$ la variabile casuale che descrive i tempi di gestazione delle donne in gravidanza, occorre determinare la seguente probabilità: $P(X > 274)$.

Abbiamo

$$P(X > 274) = P\left(\frac{X - \mu}{\sigma} > \frac{274 - \mu}{\sigma}\right) = P\left(Z > \frac{274 - 270}{30}\right) =$$
$$=$$



Una ricerca ha evidenziato che i tempi di gestazione delle donne in gravidanza seguono una distribuzione Normale con media $\mu = 270$ giorni e varianza $\sigma^2 = 900$.

1. Indicare quale è la probabilità che un bambino nasca dopo 274 giorni.

Indicata con $X \sim N(\mu = 270, \sigma^2 = 900)$ la variabile casuale che descrive i tempi di gestazione delle donne in gravidanza, occorre determinare la seguente probabilità: $P(X > 274)$.

Abbiamo

$$\begin{aligned} P(X > 274) &= P\left(\frac{X - \mu}{\sigma} > \frac{274 - \mu}{\sigma}\right) = P\left(Z > \frac{274 - 270}{30}\right) = \\ &= P(Z > 0.13) = \end{aligned}$$



Una ricerca ha evidenziato che i tempi di gestazione delle donne in gravidanza seguono una distribuzione Normale con media $\mu = 270$ giorni e varianza $\sigma^2 = 900$.

1. Indicare quale è la probabilità che un bambino nasca dopo 274 giorni.

Indicata con $X \sim N(\mu = 270, \sigma^2 = 900)$ la variabile casuale che descrive i tempi di gestazione delle donne in gravidanza, occorre determinare la seguente probabilità: $P(X > 274)$.

Abbiamo

$$\begin{aligned} P(X > 274) &= P\left(\frac{X - \mu}{\sigma} > \frac{274 - \mu}{\sigma}\right) = P\left(Z > \frac{274 - 270}{30}\right) = \\ &= P(Z > 0.13) = 1 - P(Z \leq 0.13) = \\ &= \end{aligned}$$



Una ricerca ha evidenziato che i tempi di gestazione delle donne in gravidanza seguono una distribuzione Normale con media $\mu = 270$ giorni e varianza $\sigma^2 = 900$.

1. Indicare quale è la probabilità che un bambino nasca dopo 274 giorni.

Indicata con $X \sim N(\mu = 270, \sigma^2 = 900)$ la variabile casuale che descrive i tempi di gestazione delle donne in gravidanza, occorre determinare la seguente probabilità: $P(X > 274)$.

Abbiamo

$$\begin{aligned} P(X > 274) &= P\left(\frac{X - \mu}{\sigma} > \frac{274 - \mu}{\sigma}\right) = P\left(Z > \frac{274 - 270}{30}\right) = \\ &= P(Z > 0.13) = 1 - P(Z \leq 0.13) = \\ &= 1 - 0.5517 = 0.4483. \end{aligned}$$



2. Sapendo che la probabilità che un bambino nasca prematuro è del 2%, stabilire quale è il tempo gestazionale minimo (numero di giorni) affinché un bambino non sia considerato prematuro.

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.



Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.

2. Sapendo che la probabilità che un bambino nasca prematuro è del 2%, stabilire quale è il tempo gestazionale minimo (numero di giorni) affinché un bambino non sia considerato prematuro.

Il tempo gestazionale minimo (numero di giorni) affinché un bambino non sia considerato prematuro corrisponde al percentile di ordine 0.02 della distribuzione di X .



Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.

2. Sapendo che la probabilità che un bambino nasca prematuro è del 2%, stabilire quale è il tempo gestazionale minimo (numero di giorni) affinché un bambino non sia considerato prematuro.

Il tempo gestazionale minimo (numero di giorni) affinché un bambino non sia considerato prematuro corrisponde al percentile di ordine 0.02 della distribuzione di X .

$$\begin{aligned}\{x : P(X \leq x) = 0.02\} &= \left\{x : P\left(\frac{X - \mu}{\sigma} \leq \frac{x - \mu}{\sigma}\right) = 0.02\right\} = \\ &= \left\{x : P\left(Z \leq \frac{x - 270}{30}\right) = 0.02\right\}.\end{aligned}$$



2. Sapendo che la probabilità che un bambino nasca prematuro è del 2%, stabilire quale è il tempo gestazionale minimo (numero di giorni) affinché un bambino non sia considerato prematuro.

Il tempo gestazionale minimo (numero di giorni) affinché un bambino non sia considerato prematuro corrisponde al percentile di ordine 0.02 della distribuzione di X .

$$\begin{aligned}\{x : P(X \leq x) = 0.02\} &= \left\{x : P\left(\frac{X - \mu}{\sigma} \leq \frac{x - \mu}{\sigma}\right) = 0.02\right\} = \\ &= \left\{x : P\left(Z \leq \frac{x - 270}{30}\right) = 0.02\right\}.\end{aligned}$$

Dalla lettura inversa della Tavola possiamo ricavare il percentile di ordine 0.02, che risulta -2.05 (infatti, $P(Z \leq -2.05) = 0.02$ dal momento che la distribuzione della variabile casuale Normale è simmetrica e $P(Z \leq 2.05) = 0.98$).



2. Sapendo che la probabilità che un bambino nasca prematuro è del 2%, stabilire quale è il tempo gestazionale minimo (numero di giorni) affinché un bambino non sia considerato prematuro.

Il tempo gestazionale minimo (numero di giorni) affinché un bambino non sia considerato prematuro corrisponde al percentile di ordine 0.02 della distribuzione di X .

$$\begin{aligned}\{x : P(X \leq x) = 0.02\} &= \left\{x : P\left(\frac{X - \mu}{\sigma} \leq \frac{x - \mu}{\sigma}\right) = 0.02\right\} = \\ &= \left\{x : P\left(Z \leq \frac{x - 270}{30}\right) = 0.02\right\}.\end{aligned}$$

Dalla lettura inversa della Tavola possiamo ricavare il percentile di ordine 0.02, che risulta -2.05 (infatti, $P(Z \leq -2.05) = 0.02$ dal momento che la distribuzione della variabile casuale Normale è simmetrica e $P(Z \leq 2.05) = 0.98$).

Abbiamo, quindi, $z = \frac{x - 270}{30} = -2.05$ da cui segue $x = 208.5 \simeq 209$.



3. Un ginecologo ha in carico 10 donne in gravidanza: quale è la probabilità che almeno una di loro abbia un parto prematuro?

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.



3. Un ginecologo ha in carico 10 donne in gravidanza: quale è la probabilità che almeno una di loro abbia un parto prematuro?

Sia X la variabile casuale 'n. pazienti in carico con parto prematuro'. Dal momento che la popolazione è composta da un numero elevato di soggetti si può ritenere che l'evento parto prematuro si presenti in maniera indipendente in diversi soggetti. Abbiamo

$$X \sim \text{Bin}(n = 10, p = 0.02)$$

e

$$P(X \geq 1) =$$



3. Un ginecologo ha in carico 10 donne in gravidanza: quale è la probabilità che almeno una di loro abbia un parto prematuro?

Sia X la variabile casuale 'n. pazienti in carico con parto prematuro'. Dal momento che la popolazione è composta da un numero elevato di soggetti si può ritenere che l'evento parto prematuro si presenti in maniera indipendente in diversi soggetti. Abbiamo

$$X \sim \text{Bin}(n = 10, p = 0.02)$$

e

$$P(X \geq 1) = 1 - P(X = 0) =$$



3. Un ginecologo ha in carico 10 donne in gravidanza: quale è la probabilità che almeno una di loro abbia un parto prematuro?

Sia X la variabile casuale 'n. pazienti in carico con parto prematuro'. Dal momento che la popolazione è composta da un numero elevato di soggetti si può ritenere che l'evento parto prematuro si presenti in maniera indipendente in diversi soggetti. Abbiamo

$$X \sim \text{Bin}(n = 10, p = 0.02)$$

e

$$\begin{aligned} P(X \geq 1) &= 1 - P(X = 0) = \\ &= 1 - \binom{10}{0} 0.02^0 \cdot 0.98^{10} = 1 - 0.8171 = 0.1829. \end{aligned}$$

Tema T307

Prova scritta di STATISTICA del 29.01.2020

Prova scritta di
STATISTICA del
29.01.2020

Giuseppe Boari
Gabriele Cantaluppi



UNIVERSITÀ
CATTOLICA
del Sacro Cuore

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.

Giuseppe Boari
Gabriele Cantaluppi
Università Cattolica del Sacro Cuore, Milano



Il prospetto seguente riporta la distribuzione delle frequenze congiunte del carattere qualitativo X e del carattere quantitativo Y .

$X \backslash Y$	9	12	15
x_1	1	2	
x_2		1	

1. Si completi la distribuzione delle frequenze congiunte in modo che la mediana e la media aritmetica della distribuzione condizionata $Y|x_1$ risultino 12 e che il coefficiente di variazione della distribuzione condizionata $Y|x_2$ sia nullo.

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.



Con $n_{13} = 1$ la distribuzione $Y|x_1$ risulta

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.



Con $n_{13} = 1$ la distribuzione $Y|x_1$ risulta simmetrica e vale $M(Y|x_1) = y_{0.5|x_1} = 12$.
 $CV(Y|x_2) = 0$ implica che

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.



Con $n_{13} = 1$ la distribuzione $Y|x_1$ risulta simmetrica e vale $M(Y|x_1) = y_{0.5|x_1} = 12$.
 $CV(Y|x_2) = 0$ implica che la distribuzione condizionata $Y|x_2$ è degenere e, quindi,
 $n_{21} = n_{23} = 0$.

Si ottiene, quindi,

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.



Con $n_{13} = 1$ la distribuzione $Y|x_1$ risulta simmetrica e vale $M(Y|x_1) = y_{0.5|x_1} = 12$.
 $CV(Y|x_2) = 0$ implica che la distribuzione condizionata $Y|x_2$ è degenera e, quindi,
 $n_{21} = n_{23} = 0$.

Si ottiene, quindi,

$X \backslash Y$	9	12	15	$n_{i\bullet}$
x_1	1	2	1	4
x_2	0	1	0	1
$n_{\bullet j}$	1	3	1	5

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.



2. Si rappresentino graficamente le distribuzioni marginali X e Y .

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.



Si considera la distribuzione marginale della variabile X , ottenuta sommando le frequenze di riga

x_i	$n_{i\bullet}$
x_1	4
x_2	1

Si tratta di un carattere qualitativo sconnesso che può essere rappresentato graficamente mediante un grafico a bastoncini o a canne d'organo.

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.

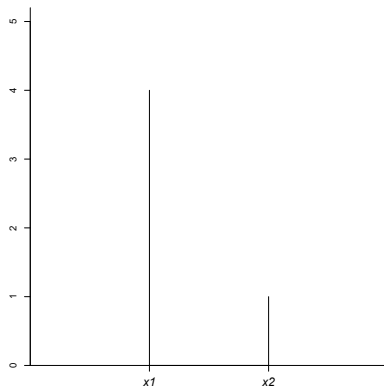


Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.





Si considera la distribuzione marginale della variabile Y , ottenuta sommando le frequenze di colonna

y_j	$n_{\bullet j}$
9	1
12	3
15	1

Esercizio 1.

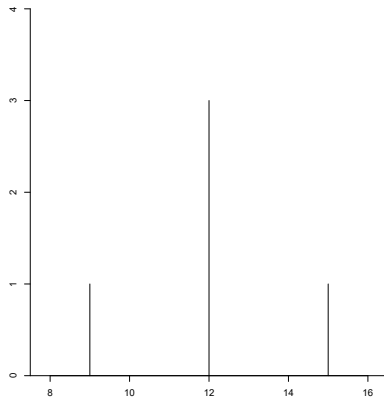
Esercizio 2.

Esercizio 3.

Esercizio 4.



Si tratta di un carattere quantitativo con modalità non raggruppate in classi che può essere rappresentato graficamente mediante un grafico a bastoncini.



Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.



3. Si ricavino i valori della media e della varianza della distribuzione marginale Y a partire dalle medie e varianze delle distribuzioni condizionate $Y|X_1$ e $Y|X_2$ e si indichi il valore del rapporto di correlazione $\eta_{Y|X}^2$.

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.



Si considera la distribuzione condizionata $Y|x_1$

y_j	n_{1j}
9	1
12	2
15	1

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.



Si tratta di un carattere quantitativo con modalità non raggruppate in classi.

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.



Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.

Si osserva che la distribuzione è simmetrica rispetto a 12; la mediana (percentile di ordine 0.50) risulta $y_{0.5} = 12$.



La media aritmetica

$$M(Y|x_1) = \frac{1}{n_{1\bullet}} \sum_{j=1}^k y_j n_{1j}$$

risulta 12.

Si osserva, infatti, che la distribuzione è simmetrica.

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.



La varianza

$$\text{Var}(Y|x_1) = M(Y^2|x_1) - [M(Y|x_1)]^2$$

risulta 4.5.

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.



La varianza

$$\text{Var}(Y|x_1) = M(Y^2|x_1) - [M(Y|x_1)]^2$$

risulta 4.5. Si ottiene

y_j	n_{1j}	$y_j n_{1j}$	$y_j^2 n_{1j}$
9	1	9	81
12	2	24	288
15	1	15	225
	4	48	594

da cui segue

$$\text{Var}(Y|x_1) = \frac{594}{4} - 12^2 = 148.5 - 144 = 4.5.$$

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.



Si considera la distribuzione condizionata $Y|x_2$

y_j	n_{2j}
9	0
12	1
15	0

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.



Si tratta di un carattere quantitativo con modalità non raggruppate in classi.

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.



Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.

Si osserva che la distribuzione è degenere.

$M(Y|x_2) = 12$. La varianza e il coefficiente di variazione sono nulli.



Da

$$M(Y|x_1) = M(Y|x_2) = 12$$

e

$$\text{Var}(Y|x_1) = 4.5 \text{ e } \text{Var}(Y|x_2) = 0$$

segue:

$$M(Y) =$$

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.



Da

$$M(Y|x_1) = M(Y|x_2) = 12$$

e

$$\text{Var}(Y|x_1) = 4.5 \text{ e } \text{Var}(Y|x_2) = 0$$

segue:

$$M(Y) = M_X\{M(Y|X)\} =$$

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.



Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.

Da

$$M(Y|x_1) = M(Y|x_2) = 12$$

e

$$Var(Y|x_1) = 4.5 \text{ e } Var(Y|x_2) = 0$$

segue:

$$M(Y) = M_X\{M(Y|X)\} = 12$$

e

$$Var(Y) =$$



Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.

Da

$$M(Y|x_1) = M(Y|x_2) = 12$$

e

$$Var(Y|x_1) = 4.5 \text{ e } Var(Y|x_2) = 0$$

segue:

$$M(Y) = M_X\{M(Y|X)\} = 12$$

e

$$Var(Y) = Var_X\{M(Y|X)\} +$$



Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.

Da

$$M(Y|x_1) = M(Y|x_2) = 12$$

e

$$Var(Y|x_1) = 4.5 \text{ e } Var(Y|x_2) = 0$$

segue:

$$M(Y) = M_X\{M(Y|X)\} = 12$$

e

$$Var(Y) = Var_X\{M(Y|X)\} + M_X\{Var(Y|X)\} =$$



Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.

Da

$$M(Y|x_1) = M(Y|x_2) = 12$$

e

$$Var(Y|x_1) = 4.5 \text{ e } Var(Y|x_2) = 0$$

segue:

$$M(Y) = M_X\{M(Y|X)\} = 12$$

e

$$Var(Y) = Var_X\{M(Y|X)\} + M_X\{Var(Y|X)\} = 0 +$$



Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.

Da

$$M(Y|x_1) = M(Y|x_2) = 12$$

e

$$Var(Y|x_1) = 4.5 \text{ e } Var(Y|x_2) = 0$$

segue:

$$M(Y) = M_X\{M(Y|X)\} = 12$$

e

$$Var(Y) = Var_X\{M(Y|X)\} + M_X\{Var(Y|X)\} = 0 + 3.6 =$$



Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.

Da

$$M(Y|x_1) = M(Y|x_2) = 12$$

e

$$Var(Y|x_1) = 4.5 \text{ e } Var(Y|x_2) = 0$$

segue:

$$M(Y) = M_X\{M(Y|X)\} = 12$$

e

$$Var(Y) = Var_X\{M(Y|X)\} + M_X\{Var(Y|X)\} = 0 + 3.6 = 3.6$$

Vale, infine:

$$\eta_{Y|X}^2 =$$



Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.

Da

$$M(Y|x_1) = M(Y|x_2) = 12$$

e

$$Var(Y|x_1) = 4.5 \text{ e } Var(Y|x_2) = 0$$

segue:

$$M(Y) = M_X\{M(Y|X)\} = 12$$

e

$$Var(Y) = Var_X\{M(Y|X)\} + M_X\{Var(Y|X)\} = 0 + 3.6 = 3.6$$

Vale, infine:

$$\eta_{Y|X}^2 = \frac{Var_X\{M(Y|X)\}}{Var(Y)} =$$



Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.

Da

$$M(Y|x_1) = M(Y|x_2) = 12$$

e

$$Var(Y|x_1) = 4.5 \text{ e } Var(Y|x_2) = 0$$

segue:

$$M(Y) = M_X\{M(Y|X)\} = 12$$

e

$$Var(Y) = Var_X\{M(Y|X)\} + M_X\{Var(Y|X)\} = 0 + 3.6 = 3.6$$

Vale, infine:

$$\eta_{Y|X}^2 = \frac{Var_X\{M(Y|X)\}}{Var(Y)} = \frac{0}{3.6} = 0$$



4. Si confrontino con opportuni indici di posizione e variabilità la distribuzione condizionata $Y|x_1$ e la distribuzione marginale Y .

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.



Si considera la distribuzione marginale della variabile Y , ottenuta sommando le frequenze di colonna

y_j	$n_{\bullet j}$
9	1
12	3
15	1

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.



Si tratta di un carattere quantitativo con modalità non raggruppate in classi.

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.



La media aritmetica

$$M(Y) = \frac{1}{n} \sum_{j=1}^k y_j n_{\bullet j}$$

risulta 12.

Si osserva, infatti, che la distribuzione è simmetrica.

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.



La varianza

$$\text{Var}(Y) = M[(Y - \mu_Y)^2] = M(Y^2) - \mu_Y^2 = \frac{1}{n} \sum_{j=1}^k y_j^2 n_{\bullet j} - \mu_Y^2$$

risulta 3.6.

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.



La varianza

$$\text{Var}(Y) = M[(Y - \mu_Y)^2] = M(Y^2) - \mu_Y^2 = \frac{1}{n} \sum_{j=1}^k y_j^2 n_{\bullet j} - \mu_Y^2$$

risulta 3.6.

Si ottiene

y_j	$n_{\bullet j}$	$y_j n_{\bullet j}$	$y_j^2 n_{\bullet j}$
9	1	9	81
12	3	36	432
15	1	15	225
	5	60	738

da cui segue

$$\begin{aligned} \text{Var}(Y) = \sigma_Y^2 &= M(Y^2) - \mu_Y^2 = \\ &= \frac{738}{5} - 12^2 = 147.6 - 144 = 3.6. \end{aligned}$$



Per effettuare confronti in termini di variabilità tra caratteri quantitativi occorre utilizzare una misura relativa, ad esempio il coefficiente di variazione

$$CV = \frac{\sigma}{\mu} = \frac{\sqrt{\text{Var}(Y)}}{M(Y)} = \frac{\sqrt{3.6}}{12} = \frac{1.8974}{12} = 0.1581.$$

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.



Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.

$Y|x_1$ Indice di posizione: $M(Y|x_1) = 12$ Indice di variabilità: $CV(Y|x_1) = 0.1768$

Y Indice di posizione: $M(Y) = 12$ Indice di variabilità: $CV(Y) = 0.1581$

Le due distribuzioni hanno lo stesso livello medio e la distribuzione condizionata $Y|x_1$ è caratterizzata da variabilità superiore rispetto alla distribuzione marginale Y .



Il seguente prospetto riporta alcune informazioni sui numeri indici con base fissa 2012 ($NIBF(base = 2012)$), sui numeri indici a base mobile ($NIBM$) e sulle variazioni percentuali rispetto all'anno precedente ($variaz.relative\%$) riferiti ai valori del fatturato di un'azienda tra il 2012 e il 2016:

t	$NIBF(base = 2012)$	$NIBM$	$variaz.relative\%$
2012			
2013	1.11		
2014			+11%
2015		1.11	
2016	0.78		

1. Si completi il prospetto con i dati mancanti.

La variazione relativa si ottiene sottraendo 1 al corrispondente numero indice (misura relativa).

Si ottiene, quindi

$$2012 I_{2012} =$$



Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.

La variazione relativa si ottiene sottraendo 1 al corrispondente numero indice (misura relativa).

Si ottiene, quindi

$$2012 I_{2012} = 1$$

$$2013 I_{2014} =$$



Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.

La variazione relativa si ottiene sottraendo 1 al corrispondente numero indice (misura relativa).

Si ottiene, quindi

$$2012 I_{2012} = 1$$

$$2013 I_{2014} = 1.11$$

$$2013 I_{2014} =$$



Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.

La variazione relativa si ottiene sottraendo 1 al corrispondente numero indice (misura relativa).

Si ottiene, quindi

$${}_{2012}I_{2012} = 1$$

$${}_{2013}I_{2014} = 1.11$$

$${}_{2013}I_{2014} = \frac{x_{2014}}{x_{2013}} = \frac{12}{24} = 0.5$$

$${}_{2012}I_{2014} =$$



Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.

La variazione relativa si ottiene sottraendo 1 al corrispondente numero indice (misura relativa).

Si ottiene, quindi

$${}_{2012}I_{2012} = 1$$

$${}_{2013}I_{2014} = 1.11$$

$${}_{2013}I_{2014} = \frac{x_{2014}}{x_{2013}} = \frac{12}{24} = 0.5$$

$${}_{2012}I_{2014} = {}_{2012}I_{2013} \cdot {}_{2013}I_{2014} = 1.11 \cdot 1.11 = 1.2321$$

$${}_{2012}I_{2015} =$$



Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.

La variazione relativa si ottiene sottraendo 1 al corrispondente numero indice (misura relativa).

Si ottiene, quindi

$${}_{2012}I_{2012} = 1$$

$${}_{2013}I_{2014} = 1.11$$

$${}_{2013}I_{2014} = \frac{x_{2014}}{x_{2013}} = \frac{12}{24} = 0.5$$

$${}_{2012}I_{2014} = {}_{2012}I_{2013} \cdot {}_{2013}I_{2014} = 1.11 \cdot 1.11 = 1.2321$$

$${}_{2012}I_{2015} = {}_{2012}I_{2014} \cdot {}_{2014}I_{2015} = 1.2321 \cdot 1.11 = 1.3676$$

$${}_{2015}I_{2016} =$$



Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.

La variazione relativa si ottiene sottraendo 1 al corrispondente numero indice (misura relativa).

Si ottiene, quindi

$${}_{2012}I_{2012} = 1$$

$${}_{2013}I_{2014} = 1.11$$

$${}_{2013}I_{2014} = \frac{x_{2014}}{x_{2013}} = \frac{12}{24} = 0.5$$

$${}_{2012}I_{2014} = {}_{2012}I_{2013} \cdot {}_{2013}I_{2014} = 1.11 \cdot 1.11 = 1.2321$$

$${}_{2012}I_{2015} = {}_{2012}I_{2014} \cdot {}_{2014}I_{2015} = 1.2321 \cdot 1.11 = 1.3676$$

$${}_{2015}I_{2016} = \frac{{}_{2012}I_{2016}}{{}_{2012}I_{2015}} = \frac{0.78}{1.3676} = 0.5703$$

e

$$variaz.relativa\%(2012 \rightarrow 2013) =$$



Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.

La variazione relativa si ottiene sottraendo 1 al corrispondente numero indice (misura relativa).

Si ottiene, quindi

$${}_{2012}I_{2012} = 1$$

$${}_{2013}I_{2014} = 1.11$$

$${}_{2013}I_{2014} = \frac{x_{2014}}{x_{2013}} = \frac{12}{24} = 0.5$$

$${}_{2012}I_{2014} = {}_{2012}I_{2013} \cdot {}_{2013}I_{2014} = 1.11 \cdot 1.11 = 1.2321$$

$${}_{2012}I_{2015} = {}_{2012}I_{2014} \cdot {}_{2014}I_{2015} = 1.2321 \cdot 1.11 = 1.3676$$

$${}_{2015}I_{2016} = \frac{{}_{2012}I_{2016}}{{}_{2012}I_{2015}} = \frac{0.78}{1.3676} = 0.5703$$

e

$$\text{variaz.relativa}\%(2012 \rightarrow 2013) = 100 \cdot (1.11 - 1) = +11\%$$

$$\text{variaz.relativa}\%(2014 \rightarrow 2015) =$$



Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.

La variazione relativa si ottiene sottraendo 1 al corrispondente numero indice (misura relativa).

Si ottiene, quindi

$${}_{2012}I_{2012} = 1$$

$${}_{2013}I_{2014} = 1.11$$

$${}_{2013}I_{2014} = \frac{x_{2014}}{x_{2013}} = \frac{12}{24} = 0.5$$

$${}_{2012}I_{2014} = {}_{2012}I_{2013} \cdot {}_{2013}I_{2014} = 1.11 \cdot 1.11 = 1.2321$$

$${}_{2012}I_{2015} = {}_{2012}I_{2014} \cdot {}_{2014}I_{2015} = 1.2321 \cdot 1.11 = 1.3676$$

$${}_{2015}I_{2016} = \frac{{}_{2012}I_{2016}}{{}_{2012}I_{2015}} = \frac{0.78}{1.3676} = 0.5703$$

e

$$\text{variaz.relativa}\%(2012 \rightarrow 2013) = 100 \cdot (1.11 - 1) = +11\%$$

$$\text{variaz.relativa}\%(2014 \rightarrow 2015) = 100 \cdot (1.11 - 1) = +11\%$$

$$\text{variaz.relativa}\%(2015 \rightarrow 2016) =$$



Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.

La variazione relativa si ottiene sottraendo 1 al corrispondente numero indice (misura relativa).

Si ottiene, quindi

$${}_{2012}I_{2012} = 1$$

$${}_{2013}I_{2014} = 1.11$$

$${}_{2013}I_{2014} = \frac{x_{2014}}{x_{2013}} = \frac{12}{24} = 0.5$$

$${}_{2012}I_{2014} = {}_{2012}I_{2013} \cdot {}_{2013}I_{2014} = 1.11 \cdot 1.11 = 1.2321$$

$${}_{2012}I_{2015} = {}_{2012}I_{2014} \cdot {}_{2014}I_{2015} = 1.2321 \cdot 1.11 = 1.3676$$

$${}_{2015}I_{2016} = \frac{{}_{2012}I_{2016}}{{}_{2012}I_{2015}} = \frac{0.78}{1.3676} = 0.5703$$

e

$$\text{variaz.relativa}\%(2012 \rightarrow 2013) = 100 \cdot (1.11 - 1) = +11\%$$

$$\text{variaz.relativa}\%(2014 \rightarrow 2015) = 100 \cdot (1.11 - 1) = +11\%$$

$$\text{variaz.relativa}\%(2015 \rightarrow 2016) = 100 \cdot (0.5703 - 1) = -42.97\%$$



Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.



e, infine

t	$NIBF(base = 2012)$	$NIBM$	$variaz.relative\%$
2012	1	—	
2013	1.11	1.11	+11%
2014	1.2321	1.11	+11%
2015	1.3676	1.11	+11%
2016	0.78	0.5703	-42.97%

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.



2. Si calcoli il numero indice al tempo 2013 con base 2016 e si commenti il risultato ottenuto.

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.



2. Si calcoli il numero indice al tempo 2013 con base 2016 e si commenti il risultato ottenuto.

$${}_{2016}I_{2013} = \frac{X_{2013}}{X_{2016}} = \frac{{}_{2012}I_{2013}}{{}_{2012}I_{2016}} = \frac{1.11}{0.78} = 1.4231$$

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.



Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.

2. Si calcoli il numero indice al tempo 2013 con base 2016 e si commenti il risultato ottenuto.

$${}_{2016}I_{2013} = \frac{X_{2013}}{X_{2016}} = \frac{{}_{2012}I_{2013}}{{}_{2012}I_{2016}} = \frac{1.11}{0.78} = 1.4231$$

La grandezza al tempo 2013 era superiore del 42.31% rispetto alla grandezza al tempo 2016. Ovvero il valore della grandezza nel 2013 era il 142.31% di quello registrato nel 2016.



3. Si calcoli il tasso medio di variazione tra il 2013 e il 2016 e si commenti il risultato ottenuto.

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.



3. Si calcoli il tasso medio di variazione tra il 2013 e il 2016 e si commenti il risultato ottenuto.
Si ricorda che

$${}_{2013}I_{2016} = {}_{2013}I_{2014} \cdot {}_{2014}I_{2015} \cdot {}_{2015}I_{2016}$$

Si deve trovare il numero indice a base mobile medio.

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.



Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.

3. Si calcoli il tasso medio di variazione tra il 2013 e il 2016 e si commenti il risultato ottenuto. Si ricorda che

$${}_{2013}I_{2016} = {}_{2013}I_{2014} \cdot {}_{2014}I_{2015} \cdot {}_{2015}I_{2016}$$

Si deve trovare il numero indice a base mobile medio. Applicando il criterio di scelta della media secondo Chisini si ottiene

$${}_{2013}I_{2016} = {}_{2013}I_{2014} \cdot {}_{2014}I_{2015} \cdot {}_{2015}I_{2016} = \alpha \cdot \alpha \cdot \alpha = \alpha^3.$$



Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.

3. Si calcoli il tasso medio di variazione tra il 2013 e il 2016 e si commenti il risultato ottenuto. Si ricorda che

$${}_{2013}I_{2016} = {}_{2013}I_{2014} \cdot {}_{2014}I_{2015} \cdot {}_{2015}I_{2016}$$

Si deve trovare il numero indice a base mobile medio. Applicando il criterio di scelta della media secondo Chisini si ottiene

$${}_{2013}I_{2016} = {}_{2013}I_{2014} \cdot {}_{2014}I_{2015} \cdot {}_{2015}I_{2016} = \alpha \cdot \alpha \cdot \alpha = \alpha^3.$$

Quindi

$$\alpha = \sqrt[3]{{}_{2013}I_{2016}} = \sqrt[3]{{}_{2013}I_{2014} \cdot {}_{2014}I_{2015} \cdot {}_{2015}I_{2016}} = 0.889.$$

Da cui segue una variazione relativa media pari a -0.111 , corrispondente a un decremento medio annuo tra il 2013 e il 2016 del 11.1%.



Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.

Il seguente prospetto riporta i valori rilevati in corrispondenza a 5 unità statistiche con riferimento a due caratteri quantitativi X e Y

x_i	1	1	2	2	3
y_i	2	4	2	6	5

1. Si rappresentino graficamente le coppie dei dati e la funzione di regressione ipotizzando Y in funzione di X .



La funzione di regressione associa a ogni valore x appartenente all'insieme di definizione della variabile X la media di Y condizionata a x , $M(Y|x)$

x	$M(Y x)$
1	$\frac{1}{2}(2 + 4) = 3$
2	$\frac{1}{2}(2 + 6) = 4$
3	5

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.



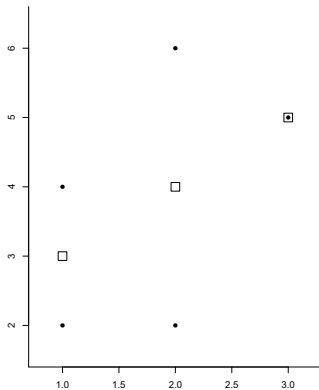
Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.

Le coppie di dati possono essere rappresentate mediante un grafico a dispersione.





2. Si calcolino, secondo il metodo dei minimi quadrati, i parametri e l'indice di adattamento dei seguenti modelli di regressione:

$$I) Y^* = a + b \cdot X \qquad II) Y^* = c \cdot X$$

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.

Si riportano i calcoli necessari¹ per ottenere le stime dei parametri e gli indici di adattamento del modello

$$Y = a + bX + E$$

x_i	y_i	$x_i y_i$	x_i^2	y_i^2
1	2	2	1	4
1	4	4	1	16
2	2	4	4	4
2	6	12	4	36
3	5	15	9	25
9	19	37	19	85
1.8	3.8	7.4	3.8	17
$M(X)$	$M(Y)$	$M(XY)$	$M(X^2)$	$M(Y^2)$

$$\hat{b} =$$



Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.

¹Si osserva che per ragioni di stampa alcuni calcoli potrebbero essere duplicati con riferimento ai diversi modelli oggetto di studio

Si riportano i calcoli necessari¹ per ottenere le stime dei parametri e gli indici di adattamento del modello

$$Y = a + bX + E$$

x_i	y_i	$x_i y_i$	x_i^2	y_i^2
1	2	2	1	4
1	4	4	1	16
2	2	4	4	4
2	6	12	4	36
3	5	15	9	25
9	19	37	19	85
1.8	3.8	7.4	3.8	17
$M(X)$	$M(Y)$	$M(XY)$	$M(X^2)$	$M(Y^2)$

$$\begin{aligned}\hat{b} &= \frac{\text{Cov}(X, Y)}{\text{Var}(X)} = \frac{M(XY) - M(X)M(Y)}{M(X^2) - [M(X)]^2} = \\ &= \frac{7.4 - 1.8 \cdot 3.8}{3.8 - 1.8^2} = \frac{0.56}{0.56} = 1 \\ \hat{a} &= \end{aligned}$$



Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.

¹Si osserva che per ragioni di stampa alcuni calcoli potrebbero essere duplicati con riferimento ai diversi modelli oggetto di studio

Si riportano i calcoli necessari¹ per ottenere le stime dei parametri e gli indici di adattamento del modello

$$Y = a + bX + E$$

x_i	y_i	$x_i y_i$	x_i^2	y_i^2
1	2	2	1	4
1	4	4	1	16
2	2	4	4	4
2	6	12	4	36
3	5	15	9	25
9	19	37	19	85
1.8	3.8	7.4	3.8	17
$M(X)$	$M(Y)$	$M(XY)$	$M(X^2)$	$M(Y^2)$

$$\hat{b} = \frac{\text{Cov}(X, Y)}{\text{Var}(X)} = \frac{M(XY) - M(X)M(Y)}{M(X^2) - [M(X)]^2} =$$

$$= \frac{7.4 - 1.8 \cdot 3.8}{3.8 - 1.8^2} = \frac{0.56}{0.56} = 1$$

$$\hat{a} = M(Y) - \hat{b}M(X) = 2$$

$$\rho^2 =$$



Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.

¹Si osserva che per ragioni di stampa alcuni calcoli potrebbero essere duplicati con riferimento ai diversi modelli oggetto di studio

Si riportano i calcoli necessari¹ per ottenere le stime dei parametri e gli indici di adattamento del modello

$$Y = a + bX + E$$

x_i	y_i	$x_i y_i$	x_i^2	y_i^2
1	2	2	1	4
1	4	4	1	16
2	2	4	4	4
2	6	12	4	36
3	5	15	9	25
9	19	37	19	85
1.8	3.8	7.4	3.8	17
$M(X)$	$M(Y)$	$M(XY)$	$M(X^2)$	$M(Y^2)$

$$\hat{b} = \frac{\text{Cov}(X, Y)}{\text{Var}(X)} = \frac{M(XY) - M(X)M(Y)}{M(X^2) - [M(X)]^2} =$$

$$= \frac{7.4 - 1.8 \cdot 3.8}{3.8 - 1.8^2} = \frac{0.56}{0.56} = 1$$

$$\hat{a} = M(Y) - \hat{b}M(X) = 2$$

$$\rho^2 = \frac{[\text{Cov}(X, Y)]^2}{\text{Var}(X) \text{Var}(Y)} = \frac{0.56^2}{0.56 \cdot 2.56} = 0.2187$$

$$M(\hat{E}^2) = \text{Var}(\hat{E}) = \text{varianza residua} =$$

¹Si osserva che per ragioni di stampa alcuni calcoli potrebbero essere duplicati con riferimento ai diversi modelli oggetto di studio



Si riportano i calcoli necessari¹ per ottenere le stime dei parametri e gli indici di adattamento del modello

$$Y = a + bX + E$$

x_i	y_i	$x_i y_i$	x_i^2	y_i^2
1	2	2	1	4
1	4	4	1	16
2	2	4	4	4
2	6	12	4	36
3	5	15	9	25
9	19	37	19	85
1.8	3.8	7.4	3.8	17
$M(X)$	$M(Y)$	$M(XY)$	$M(X^2)$	$M(Y^2)$

$$\hat{b} = \frac{\text{Cov}(X, Y)}{\text{Var}(X)} = \frac{M(XY) - M(X)M(Y)}{M(X^2) - [M(X)]^2} =$$

$$= \frac{7.4 - 1.8 \cdot 3.8}{3.8 - 1.8^2} = \frac{0.56}{0.56} = 1$$

$$\hat{a} = M(Y) - \hat{b}M(X) = 2$$

$$\rho^2 = \frac{[\text{Cov}(X, Y)]^2}{\text{Var}(X) \text{Var}(Y)} = \frac{0.56^2}{0.56 \cdot 2.56} = 0.2187$$

$$M(\hat{E}^2) = \text{Var}(\hat{E}) = \text{varianza residua} = \text{Var}(Y) \cdot (1 - \rho^2) = 2$$

¹Si osserva che per ragioni di stampa alcuni calcoli potrebbero essere duplicati con riferimento ai diversi modelli oggetto di studio





Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.

Si riportano i calcoli necessari² per ottenere le stime dei parametri e la misura di adattamento del modello

$$Y = cX + E$$

x_i	y_i	$x_i y_i$	x_i^2	y_i^2
1	2	2	1	4
1	4	4	1	16
2	2	4	4	4
2	6	12	4	36
3	5	15	9	25
		37	19	85
		7.4	3.8	17
		$M(XY)$	$M(X^2)$	$M(Y^2)$

$$\hat{c} =$$

²Si osserva che per ragioni di stampa alcuni calcoli potrebbero essere duplicati con riferimento ai diversi modelli oggetto di studio



Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.

Si riportano i calcoli necessari² per ottenere le stime dei parametri e la misura di adattamento del modello

$$Y = cX + E$$

x_i	y_i	$x_i y_i$	x_i^2	y_i^2
1	2	2	1	4
1	4	4	1	16
2	2	4	4	4
2	6	12	4	36
3	5	15	9	25
		37	19	85
		7.4	3.8	17
		$M(XY)$	$M(X^2)$	$M(Y^2)$

$$\hat{c} = \frac{M(XY)}{M(X^2)} = \frac{7.4}{3.8} = 1.9474$$

$$M(\hat{E}^2) =$$

²Si osserva che per ragioni di stampa alcuni calcoli potrebbero essere duplicati con riferimento ai diversi modelli oggetto di studio



Si riportano i calcoli necessari² per ottenere le stime dei parametri e la misura di adattamento del modello

$$Y = cX + E$$

x_i	y_i	$x_i y_i$	x_i^2	y_i^2
1	2	2	1	4
1	4	4	1	16
2	2	4	4	4
2	6	12	4	36
3	5	15	9	25
		37	19	85
		7.4	3.8	17
		$M(XY)$	$M(X^2)$	$M(Y^2)$

$$\hat{c} = \frac{M(XY)}{M(X^2)} = \frac{7.4}{3.8} = 1.9474$$

$$M(\hat{E}^2) = M[Y \cdot (Y - \hat{c}X)] = M(Y^2) - \hat{c}M(XY) = 17 - 1.9474 \cdot 7.4 = 2.5895$$

²Si osserva che per ragioni di stampa alcuni calcoli potrebbero essere duplicati con riferimento ai diversi modelli oggetto di studio



$\hat{a} = 2$, $\hat{b} = 1$, indice adattamento $M(\hat{E}^2) = 2$
 $\hat{c} = 1.9474$, indice adattamento $M(\hat{E}^2) = 2.5895$
Il modello migliore è $Y^* = 2 + X$.

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.



3. Senza svolgere i conti, cosa è possibile dire con riferimento ai dati in esame riguardo alla relazione tra ρ^2 ed $\eta_{X|Y}^2$?

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.



3. Senza svolgere i conti, cosa è possibile dire con riferimento ai dati in esame riguardo alla relazione tra ρ^2 ed $\eta_{X|Y}^2$?

$\rho^2 = \eta_{X|Y}^2 = 0.2188$ in quanto le medie condizionate $M(Y|X)$ sono allineate.

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.



Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.

4. Sulla base del miglior modello si preveda il valore di Y in corrispondenza di $X = 3$.



Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.

4. Sulla base del miglior modello si preveda il valore di Y in corrispondenza di $X = 3$.

$$Y^*(3) = 2 + 1 \cdot 3 = 5.$$

Esercizio 4

Prova scritta di
STATISTICA del
29.01.2020

Giuseppe Boari
Gabriele Cantaluppi



UNIVERSITÀ
CATTOLICA
del Sacro Cuore

Un albero di Natale è addobbato con led di diversi colori: 20 bianchi, 15 rossi e 15 verdi. A ogni accensione 3 diversi led si illuminano mentre i restanti 47 restano spenti.

1. Si dica quale v.c. descrive il numero di led bianchi che si illuminano in una accensione e si calcoli il suo valore atteso.

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.

Esercizio 4



Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.

Un albero di Natale è addobbato con led di diversi colori: 20 bianchi, 15 rossi e 15 verdi. A ogni accensione 3 diversi led si illuminano mentre i restanti 47 restano spenti.

1. Si dica quale v.c. descrive il numero di led bianchi che si illuminano in una accensione e si calcoli il suo valore atteso.

L'esperimento corrisponde alla scelta, senza reimmissione, di 3 led tra i 50 presenti sull'albero. Il numero di led bianchi tra i 3 led che si illuminano in una accensione può essere descritto mediante una variabile casuale ipergeometrica

$$P(X = x) =$$

Esercizio 4



Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.

Un albero di Natale è addobbato con led di diversi colori: 20 bianchi, 15 rossi e 15 verdi. A ogni accensione 3 diversi led si illuminano mentre i restanti 47 restano spenti.

1. Si dica quale v.c. descrive il numero di led bianchi che si illuminano in una accensione e si calcoli il suo valore atteso.

L'esperimento corrisponde alla scelta, senza reimmissione, di 3 led tra i 50 presenti sull'albero. Il numero di led bianchi tra i 3 led che si illuminano in una accensione può essere descritto mediante una variabile casuale ipergeometrica

$$P(X = x) = \frac{\binom{20}{x} \binom{30}{3-x}}{\binom{50}{3}}, \quad x = 0, 1, 2, 3$$

Il valore atteso risulta pari a

Esercizio 4



Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.

Un albero di Natale è addobbato con led di diversi colori: 20 bianchi, 15 rossi e 15 verdi. A ogni accensione 3 diversi led si illuminano mentre i restanti 47 restano spenti.

1. Si dica quale v.c. descrive il numero di led bianchi che si illuminano in una accensione e si calcoli il suo valore atteso.

L'esperimento corrisponde alla scelta, senza reimmissione, di 3 led tra i 50 presenti sull'albero. Il numero di led bianchi tra i 3 led che si illuminano in una accensione può essere descritto mediante una variabile casuale ipergeometrica

$$P(X = x) = \frac{\binom{20}{x} \binom{30}{3-x}}{\binom{50}{3}}, \quad x = 0, 1, 2, 3$$

Il valore atteso risulta pari a

$$E(X) = 3 \cdot \frac{20}{50} = 1.2$$



2. Si calcoli la probabilità che in una accensione sia illuminato almeno 1 led bianco.

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.



2. Si calcoli la probabilità che in una accensione sia illuminato almeno 1 led bianco.

$$P(X \geq 1) =$$

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.



2. Si calcoli la probabilità che in una accensione sia illuminato almeno 1 led bianco.

$$P(X \geq 1) = 1 - P(X = 0) =$$

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.



2. Si calcoli la probabilità che in una accensione sia illuminato almeno 1 led bianco.

$$P(X \geq 1) = 1 - P(X = 0) = 1 - \frac{\binom{20}{0} \binom{30}{3}}{\binom{50}{3}} =$$
$$=$$

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.



Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.

2. Si calcoli la probabilità che in una accensione sia illuminato almeno 1 led bianco.

$$\begin{aligned} P(X \geq 1) &= 1 - P(X = 0) = 1 - \frac{\binom{20}{0} \binom{30}{3}}{\binom{50}{3}} = \\ &= 1 - P(\bar{B}_1 \cap \bar{B}_2 \cap \bar{B}_3) = \\ &= \end{aligned}$$



2. Si calcoli la probabilità che in una accensione sia illuminato almeno 1 led bianco.

$$\begin{aligned}P(X \geq 1) &= 1 - P(X = 0) = 1 - \frac{\binom{20}{0} \binom{30}{3}}{\binom{50}{3}} = \\&= 1 - P(\bar{B}_1 \cap \bar{B}_2 \cap \bar{B}_3) = \\&= 1 - P(\bar{B}_1) \cdot P(\bar{B}_2 | \bar{B}_1) \cdot P(\bar{B}_3 | \bar{B}_1 \cap \bar{B}_2) = \\&= \end{aligned}$$

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.



2. Si calcoli la probabilità che in una accensione sia illuminato almeno 1 led bianco.

$$\begin{aligned} P(X \geq 1) &= 1 - P(X = 0) = 1 - \frac{\binom{20}{0} \binom{30}{3}}{\binom{50}{3}} = \\ &= 1 - P(\bar{B}_1 \cap \bar{B}_2 \cap \bar{B}_3) = \\ &= 1 - P(\bar{B}_1) \cdot P(\bar{B}_2 | \bar{B}_1) \cdot P(\bar{B}_3 | \bar{B}_1 \cap \bar{B}_2) = \\ &= 1 - \frac{30}{50} \frac{29}{49} \frac{28}{48} = \end{aligned}$$

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.



2. Si calcoli la probabilità che in una accensione sia illuminato almeno 1 led bianco.

$$\begin{aligned}P(X \geq 1) &= 1 - P(X = 0) = 1 - \frac{\binom{20}{0} \binom{30}{3}}{\binom{50}{3}} = \\&= 1 - P(\bar{B}_1 \cap \bar{B}_2 \cap \bar{B}_3) = \\&= 1 - P(\bar{B}_1) \cdot P(\bar{B}_2 | \bar{B}_1) \cdot P(\bar{B}_3 | \bar{B}_1 \cap \bar{B}_2) = \\&= 1 - \frac{30}{50} \frac{29}{49} \frac{28}{48} = 1 - 0.2071 =\end{aligned}$$

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.



2. Si calcoli la probabilità che in una accensione sia illuminato almeno 1 led bianco.

$$\begin{aligned}P(X \geq 1) &= 1 - P(X = 0) = 1 - \frac{\binom{20}{0} \binom{30}{3}}{\binom{50}{3}} = \\&= 1 - P(\bar{B}_1 \cap \bar{B}_2 \cap \bar{B}_3) = \\&= 1 - P(\bar{B}_1) \cdot P(\bar{B}_2 | \bar{B}_1) \cdot P(\bar{B}_3 | \bar{B}_1 \cap \bar{B}_2) = \\&= 1 - \frac{30}{50} \frac{29}{49} \frac{28}{48} = 1 - 0.2071 = 0.7929.\end{aligned}$$

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.



3. Si calcoli la probabilità che in una accensione siano illuminati 3 led tutti dello stesso colore.

$$P((B_1 \cap B_2 \cap B_3) \cup (R_1 \cap R_2 \cap R_3) \cup (V_1 \cap V_2 \cap V_3)) =$$
$$=$$

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.



Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.

3. Si calcoli la probabilità che in una accensione siano illuminati 3 led tutti dello stesso colore.

$$\begin{aligned} &P((B_1 \cap B_2 \cap B_3) \cup (R_1 \cap R_2 \cap R_3) \cup (V_1 \cap V_2 \cap V_3)) = \\ &= P(B_1) \cdot P(B_2|B_1) \cdot P(B_3|B_1 \cap B_2) + \\ &\quad + P(R_1) \cdot P(R_2|R_1) \cdot P(R_3|R_1 \cap R_2) + \\ &\quad + P(V_1) \cdot P(V_2|V_1) \cdot P(V_3|V_1 \cap V_2) = \\ &= \end{aligned}$$



Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.

3. Si calcoli la probabilità che in una accensione siano illuminati 3 led tutti dello stesso colore.

$$\begin{aligned} &P((B_1 \cap B_2 \cap B_3) \cup (R_1 \cap R_2 \cap R_3) \cup (V_1 \cap V_2 \cap V_3)) = \\ &= P(B_1) \cdot P(B_2|B_1) \cdot P(B_3|B_1 \cap B_2) + \\ &\quad + P(R_1) \cdot P(R_2|R_1) \cdot P(R_3|R_1 \cap R_2) + \\ &\quad + P(V_1) \cdot P(V_2|V_1) \cdot P(V_3|V_1 \cap V_2) = \\ &= \frac{20}{50} \frac{19}{49} \frac{18}{48} + \frac{15}{50} \frac{14}{49} \frac{13}{48} + \frac{15}{50} \frac{14}{49} \frac{13}{48} = \\ &= \end{aligned}$$



3. Si calcoli la probabilità che in una accensione siano illuminati 3 led tutti dello stesso colore.

$$\begin{aligned} &P((B_1 \cap B_2 \cap B_3) \cup (R_1 \cap R_2 \cap R_3) \cup (V_1 \cap V_2 \cap V_3)) = \\ &= P(B_1) \cdot P(B_2|B_1) \cdot P(B_3|B_1 \cap B_2) + \\ &\quad + P(R_1) \cdot P(R_2|R_1) \cdot P(R_3|R_1 \cap R_2) + \\ &\quad + P(V_1) \cdot P(V_2|V_1) \cdot P(V_3|V_1 \cap V_2) = \\ &= \frac{20}{50} \frac{19}{49} \frac{18}{48} + \frac{15}{50} \frac{14}{49} \frac{13}{48} + \frac{15}{50} \frac{14}{49} \frac{13}{48} = \\ &= 0.05816 + 0.02321 + 0.02321 = \end{aligned}$$



3. Si calcoli la probabilità che in una accensione siano illuminati 3 led tutti dello stesso colore.

$$\begin{aligned} &P((B_1 \cap B_2 \cap B_3) \cup (R_1 \cap R_2 \cap R_3) \cup (V_1 \cap V_2 \cap V_3)) = \\ &= P(B_1) \cdot P(B_2|B_1) \cdot P(B_3|B_1 \cap B_2) + \\ &\quad + P(R_1) \cdot P(R_2|R_1) \cdot P(R_3|R_1 \cap R_2) + \\ &\quad + P(V_1) \cdot P(V_2|V_1) \cdot P(V_3|V_1 \cap V_2) = \\ &= \frac{20}{50} \frac{19}{49} \frac{18}{48} + \frac{15}{50} \frac{14}{49} \frac{13}{48} + \frac{15}{50} \frac{14}{49} \frac{13}{48} = \\ &= 0.05816 + 0.02321 + 0.02321 = 0.10458 \end{aligned}$$

Tema T308

Prova scritta di STATISTICA del 19.02.2020

Prova scritta di
STATISTICA del
19.02.2020

Giuseppe Boari
Gabriele Cantaluppi



UNIVERSITÀ
CATTOLICA
del Sacro Cuore

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.

Giuseppe Boari
Gabriele Cantaluppi
Università Cattolica del Sacro Cuore, Milano



La seguente tabella riporta le informazioni sul peso, variabile X (misurato in Kg), e sul tipo di sospensione, variabile Y (con modalità fu =integrale e fe =frontale), di 8 tipi di mountain bike:

x_i	7	12	12	12	12	12	13	14
y_i	fu	fe	fe	fu	fu	fe	fu	fu

1. Si rappresenti graficamente la distribuzione di frequenza della variabile Y e si calcolino opportuni indici di posizione e di variabilità. Si commentino i risultati.

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.



Si considera la distribuzione marginale della variabile Y , ottenuta sommando le frequenze di colonna

y_j	$n_{\bullet j}$
fe	3
fu	5

Si tratta di un carattere qualitativo sconnesso che può essere rappresentato graficamente mediante un grafico a bastoncini o a canne d'organo.

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.

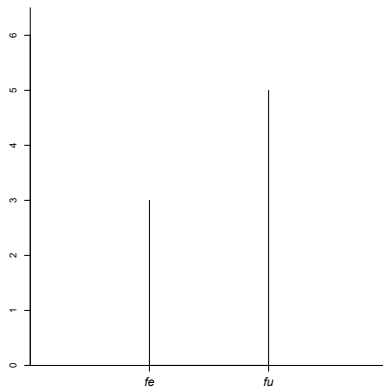


Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.





Si tratta di un carattere qualitativo sconnesso.

L'unico indice di posizione che può essere definito è la moda (categoria alla quale è associata la massima frequenza) che risulta *fu*.

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.



In presenza di un carattere qualitativo la variabilità può essere misurata mediante gli indici di Gini e di Frosini normalizzati

$$G_N = \frac{1 - \sum_{i=1}^k f_{\bullet j}^2}{1 - \frac{1}{k}}$$

e

$$F_N = 1 - \sqrt{1 - G_N}.$$

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.



In presenza di un carattere qualitativo la variabilità può essere misurata mediante gli indici di Gini e di Frosini normalizzati

$$G_N = \frac{1 - \sum_{i=1}^k f_{\bullet j}^2}{1 - \frac{1}{k}}$$

e

$$F_N = 1 - \sqrt{1 - G_N}.$$

Si ottiene:

y_j	$n_{\bullet j}$	$f_{\bullet j}$	$f_{\bullet j}^2$
<i>fe</i>	3	0.375	0.1406
<i>fu</i>	5	0.625	0.3906
	8	1	0.5312

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.



In presenza di un carattere qualitativo la variabilità può essere misurata mediante gli indici di Gini e di Frosini normalizzati

$$G_N = \frac{1 - \sum_{i=1}^k f_{\bullet j}^2}{1 - \frac{1}{k}}$$

e

$$F_N = 1 - \sqrt{1 - G_N}.$$

Si ottiene:

y_j	$n_{\bullet j}$	$f_{\bullet j}$	$f_{\bullet j}^2$
<i>fe</i>	3	0.375	0.1406
<i>fu</i>	5	0.625	0.3906
	8	1	0.5312

Da cui segue:

$$G_N = \frac{1 - 0.5313}{1 - \frac{1}{2}} = 0.9375$$

e

$$F_N = 1 - \sqrt{1 - 0.9375} = 0.75.$$

Sussiste, quindi, un livello medio alto di eterogeneità.



Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.

2. Si confronti la variabilità di X nei gruppi definiti dalla variabile Y .



Per effettuare un confronto in termini di posizione della variabile X nei 2 gruppi fu e fe si considerano le due medie aritmetiche di gruppo μ_{fu} e μ_{fe} :

$$\mu_{fu} = \frac{1}{5}(7 + 12 + 12 + 13 + 14) = 11.6$$

e

$$\mu_{fe} = \frac{1}{3}(12 + 12 + 12) = 12$$

Il livello medio nel gruppo fu è inferiore a quello osservato nel gruppo fe .

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.



Per effettuare un confronto in termini di variabilità della variabile X nei 2 gruppi fu e fe si considerano i coefficienti di variazione calcolati per i due gruppi $CV(X_{fu})$ e $CV(X_{fe})$.

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.



Per effettuare un confronto in termini di variabilità della variabile X nei 2 gruppi fu e fe si considerano i coefficienti di variazione calcolati per i due gruppi $CV(X_{fu})$ e $CV(X_{fe})$. A tal fine occorre in primo luogo calcolare le varianze nei due gruppi

$$\sigma_{fu}^2 = \frac{1}{5}(7^2 + 12^2 + 12^2 + 13^2 + 14^2) - \mu_{fu}^2 = 5.84$$

e

$$\sigma_{fe}^2 = \frac{1}{3}(12^2 + 12^2 + 12^2) - \mu_{fe}^2 = 0$$

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.



Per effettuare un confronto in termini di variabilità della variabile X nei 2 gruppi fu e fe si considerano i coefficienti di variazione calcolati per i due gruppi $CV(X_{fu})$ e $CV(X_{fe})$. A tal fine occorre in primo luogo calcolare le varianze nei due gruppi

$$\sigma_{fu}^2 = \frac{1}{5}(7^2 + 12^2 + 12^2 + 13^2 + 14^2) - \mu_{fu}^2 = 5.84$$

e

$$\sigma_{fe}^2 = \frac{1}{3}(12^2 + 12^2 + 12^2) - \mu_{fe}^2 = 0$$

Ricordando che il coefficiente di variazione è definito dal rapporto tra scarto quadratico medio e media aritmetica risulta

$$CV(X_{fu}) = \frac{\sigma_{fu}}{\mu_{fu}} = \frac{2.416609195}{11.6} = 0.2083$$

e

$$CV(X_{fe}) = \frac{\sigma_{fe}}{\mu_{fe}} = \frac{0}{12} = 0$$

Il livello di variabilità presente nel gruppo fu è superiore a quello osservato nel gruppo fe .



3. Si calcolino la varianza between e la varianza within di X nei gruppi definiti dalla variabile Y . Si commentino i risultati ottenuti.

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.



Applicando il risultato di scomposizione della varianza si ottiene:

$$\sigma_{Between}^2 =$$

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.



Applicando il risultato di scomposizione della varianza si ottiene:

$$\begin{aligned}\sigma_{Between}^2 &= \frac{1}{8} \left\{ (\mu_{fu} - \mu)^2 \cdot 5 + (\mu_{fe} - \mu)^2 \cdot 3 \right\} \\ &= \frac{1}{8} \left(\mu_{fu}^2 \cdot 5 + \mu_{fe}^2 \cdot 3 \right) - \mu^2 \\ &= \frac{1}{8} \left(11.6^2 \cdot 5 + 12^2 \cdot 3 \right) - 11.75^2 = 0.0375\end{aligned}$$

$$\sigma_{Within}^2 =$$

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.



Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.

Applicando il risultato di scomposizione della varianza si ottiene:

$$\begin{aligned}\sigma_{Between}^2 &= \frac{1}{8} \left\{ (\mu_{fu} - \mu)^2 \cdot 5 + (\mu_{fe} - \mu)^2 \cdot 3 \right\} \\ &= \frac{1}{8} \left(\mu_{fu}^2 \cdot 5 + \mu_{fe}^2 \cdot 3 \right) - \mu^2 \\ &= \frac{1}{8} \left(11.6^2 \cdot 5 + 12^2 \cdot 3 \right) - 11.75^2 = 0.0375\end{aligned}$$

$$\begin{aligned}\sigma_{Within}^2 &= \frac{1}{8} \left(\sigma_{fu}^2 \cdot 5 + \sigma_{fe}^2 \cdot 3 \right) \\ &= \frac{1}{8} (5.84 \cdot 5 + 0 \cdot 3) = 3.65.\end{aligned}$$

e

$$\sigma^2 =$$



Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.

Applicando il risultato di scomposizione della varianza si ottiene:

$$\begin{aligned}\sigma_{Between}^2 &= \frac{1}{8} \left\{ (\mu_{fu} - \mu)^2 \cdot 5 + (\mu_{fe} - \mu)^2 \cdot 3 \right\} \\ &= \frac{1}{8} \left(\mu_{fu}^2 \cdot 5 + \mu_{fe}^2 \cdot 3 \right) - \mu^2 \\ &= \frac{1}{8} \left(11.6^2 \cdot 5 + 12^2 \cdot 3 \right) - 11.75^2 = 0.0375\end{aligned}$$

$$\begin{aligned}\sigma_{Within}^2 &= \frac{1}{8} \left(\sigma_{fu}^2 \cdot 5 + \sigma_{fe}^2 \cdot 3 \right) \\ &= \frac{1}{8} (5.84 \cdot 5 + 0 \cdot 3) = 3.65.\end{aligned}$$

e

$$\sigma^2 = \sigma_{Between}^2 + \sigma_{Within}^2 = 0.0375 + 3.65 = 3.6875.$$



Il rapporto di correlazione $\eta^2_{X|Y}$ risulta:

$$\eta^2_{X|Y} = \frac{\sigma^2_{Between}}{\sigma^2_{Between} + \sigma^2_{Within}} = 0.0102.$$

Sussiste, quindi, un livello di dipendenza molto basso.

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.



4. Si costruisca la tabella a doppia entrata (W, Y) , con W variabile qualitativa con categorie: $w_1 = \text{leggera}$ se il peso X è minore o UGUALE a 12 e $w_2 = \text{pesante}$ se il peso X è maggiore di 12. Si indichi se le variabili possono ritenersi stocasticamente indipendenti.

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.



4. Si costruisca la tabella a doppia entrata (W, Y) , con W variabile qualitativa con categorie: $w_1 = \text{leggera}$ se il peso X è minore o UGUALE a 12 e $w_2 = \text{pesante}$ se il peso X è maggiore di 12. Si indichi se le variabili possono ritenersi stocasticamente indipendenti.

La distribuzione congiunta risulta

$W \backslash Y$	$f_{e.}$	$f_{.u}$
<i>leggera</i>	3	3
<i>pesante</i>	0	2

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.



4. Si costruisca la tabella a doppia entrata (W, Y) , con W variabile qualitativa con categorie: $w_1 = \text{leggera}$ se il peso X è minore o UGUALE a 12 e $w_2 = \text{pesante}$ se il peso X è maggiore di 12. Si indichi se le variabili possono ritenersi stocasticamente indipendenti.

La distribuzione congiunta risulta

$W \backslash Y$	fe	fu
<i>leggera</i>	3	3
<i>pesante</i>	0	2

Le variabili W e Y non sono stocasticamente indipendenti in quanto le distribuzioni condizionate non sono somiglianti.

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.



Per i prezzi di due beni A e B sono state calcolate le seguenti serie dei numeri indici negli anni 2012-2016, con base fissa 2012 ($NIBF_A$) per il bene A e con base mobile per il bene B ($NIBM_B$):

t	$NIBF_A$	$NIBM_B$
2012	1	—
2013	1.01	1.01
2014	1.05	1.1
2015	1.16	1.2
2016	1.15	1.3

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.

1. Costruire le serie dei *NIBM* per il bene *A* e dei *NIBF* con base 2012 per il bene *B*.



Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.

1. Costruire le serie dei *NIBM* per il bene *A* e dei *NIBF* con base 2012 per il bene *B*.

Si ricorda che ${}_{t-1}I_t = \frac{x_t}{x_{t-1}} = \frac{{}_{2012}I_t}{{}_{2012}I_{t-1}}$.



Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.



Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.

1. Costruire le serie dei *NIBM* per il bene *A* e dei *NIBF* con base 2012 per il bene *B*.
Si ricorda che ${}_{t-1}I_t = \frac{x_t}{x_{t-1}} = \frac{{}_{2012}I_t}{{}_{2012}I_{t-1}}$. Si ottiene quindi

Anno	NIBF	NIBM
2012	1.00	
2013	1.01	1.0100
2014	1.05	1.0396
2015	1.16	1.1048
2016	1.15	0.9914



Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.

1. Costruire le serie dei *NIBM* per il bene *A* e dei *NIBF* con base 2012 per il bene *B*.
Si ricorda che ${}_{t-1}I_t = \frac{x_t}{x_{t-1}} = \frac{{}_{2012}I_t}{{}_{2012}I_{t-1}}$. Si ottiene quindi

Anno	NIBF	NIBM
2012	1.00	
2013	1.01	1.0100
2014	1.05	1.0396
2015	1.16	1.1048
2016	1.15	0.9914

Si ricorda che ${}_{2012}I_{2012} = 1$ e che per valori di $t > 2012$ vale

$${}_{2012}I_t = {}_{2012}I_{2013} \cdot {}_{2013}I_{2014} \cdot \dots \cdot {}_{t-1}I_t.$$



1. Costruire le serie dei *NIBM* per il bene *A* e dei *NIBF* con base 2012 per il bene *B*.
Si ricorda che ${}_{t-1}I_t = \frac{x_t}{x_{t-1}} = \frac{{}_{2012}I_t}{{}_{2012}I_{t-1}}$. Si ottiene quindi

Anno	NIBF	NIBM
2012	1.00	
2013	1.01	1.0100
2014	1.05	1.0396
2015	1.16	1.1048
2016	1.15	0.9914

Si ricorda che ${}_{2012}I_{2012} = 1$ e che per valori di $t > 2012$ vale

$${}_{2012}I_t = {}_{2012}I_{2013} \cdot {}_{2013}I_{2014} \cdot \dots \cdot {}_{t-1}I_t.$$

Si ottiene quindi

Anno	NIBM	NIBF (2012)
2012	0.80	1.0000
2013	1.01	1.0100
2014	1.10	1.1110
2015	1.20	1.3332
2016	1.30	1.7332



2. Dire, giustificando la risposta, quale dei due beni ha avuto maggior variazione media annua negli anni considerati.

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.



2. Dire, giustificando la risposta, quale dei due beni ha avuto maggior variazione media annua negli anni considerati.

Si ricorda che

$${}_{2012}I_{2016} = {}_{2012}I_{2013} \cdot {}_{2013}I_{2014} \cdot {}_{2014}I_{2015} \cdot {}_{2015}I_{2016}$$

Si deve trovare il numero indice a base mobile medio.

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.



Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.

2. Dire, giustificando la risposta, quale dei due beni ha avuto maggior variazione media annua negli anni considerati.

Si ricorda che

$${}_{2012}I_{2016} = {}_{2012}I_{2013} \cdot {}_{2013}I_{2014} \cdot {}_{2014}I_{2015} \cdot {}_{2015}I_{2016}$$

Si deve trovare il numero indice a base mobile medio. Applicando il criterio di scelta della media secondo Chisini si ottiene

$${}_{2012}I_{2016} = {}_{2012}I_{2013} \cdot {}_{2013}I_{2014} \cdot {}_{2014}I_{2015} \cdot {}_{2015}I_{2016} = \alpha \cdot \alpha \cdot \alpha \cdot \alpha = \alpha^4.$$



Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.

2. Dire, giustificando la risposta, quale dei due beni ha avuto maggior variazione media annua negli anni considerati.

Si ricorda che

$${}_{2012}l_{2016} = {}_{2012}l_{2013} \cdot {}_{2013}l_{2014} \cdot {}_{2014}l_{2015} \cdot {}_{2015}l_{2016}$$

Si deve trovare il numero indice a base mobile medio. Applicando il criterio di scelta della media secondo Chisini si ottiene

$${}_{2012}l_{2016} = {}_{2012}l_{2013} \cdot {}_{2013}l_{2014} \cdot {}_{2014}l_{2015} \cdot {}_{2015}l_{2016} = \alpha \cdot \alpha \cdot \alpha \cdot \alpha = \alpha^4.$$

Quindi

$$\alpha = \sqrt[4]{{}_{2012}l_{2016}} = \sqrt[4]{{}_{2012}l_{2013} \cdot {}_{2013}l_{2014} \cdot {}_{2014}l_{2015} \cdot {}_{2015}l_{2016}} = 1.0356.$$

Da cui segue una variazione relativa media pari a 0.0356, corrispondente a un incremento medio annuo tra il 2012 e il 2016 del 3.56%.



Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.



Si ricorda che

$${}_{2012}I_{2016} = {}_{2012}I_{2013} \cdot {}_{2013}I_{2014} \cdot {}_{2014}I_{2015} \cdot {}_{2015}I_{2016}.$$

Si deve trovare il numero indice a base mobile medio.

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.



Si ricorda che

$${}_{2012}I_{2016} = {}_{2012}I_{2013} \cdot {}_{2013}I_{2014} \cdot {}_{2014}I_{2015} \cdot {}_{2015}I_{2016}.$$

Si deve trovare il numero indice a base mobile medio. Applicando il criterio di scelta della media secondo Chisini si ottiene

$${}_{2012}I_{2016} = {}_{2012}I_{2013} \cdot {}_{2013}I_{2014} \cdot {}_{2014}I_{2015} \cdot {}_{2015}I_{2016} = \alpha \cdot \alpha \cdot \alpha \cdot \alpha = \alpha^4.$$

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.



Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.

Si ricorda che

$${}_{2012}I_{2016} = {}_{2012}I_{2013} \cdot {}_{2013}I_{2014} \cdot {}_{2014}I_{2015} \cdot {}_{2015}I_{2016}.$$

Si deve trovare il numero indice a base mobile medio. Applicando il criterio di scelta della media secondo Chisini si ottiene

$${}_{2012}I_{2016} = {}_{2012}I_{2013} \cdot {}_{2013}I_{2014} \cdot {}_{2014}I_{2015} \cdot {}_{2015}I_{2016} = \alpha \cdot \alpha \cdot \alpha \cdot \alpha = \alpha^4.$$

Quindi

$$\alpha = \sqrt[4]{{}_{2012}I_{2016}} = \sqrt[4]{{}_{2012}I_{2013} \cdot {}_{2013}I_{2014} \cdot {}_{2014}I_{2015} \cdot {}_{2015}I_{2016}} = 1.1474.$$

Da cui segue una variazione relativa media pari a 0.1474 , corrispondente a un incremento medio annuo tra il 2012 e il 2016 del 14.74%.



Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.

Si ricorda che

$${}_{2012}l_{2016} = {}_{2012}l_{2013} \cdot {}_{2013}l_{2014} \cdot {}_{2014}l_{2015} \cdot {}_{2015}l_{2016}.$$

Si deve trovare il numero indice a base mobile medio. Applicando il criterio di scelta della media secondo Chisini si ottiene

$${}_{2012}l_{2016} = {}_{2012}l_{2013} \cdot {}_{2013}l_{2014} \cdot {}_{2014}l_{2015} \cdot {}_{2015}l_{2016} = \alpha \cdot \alpha \cdot \alpha \cdot \alpha = \alpha^4.$$

Quindi

$$\alpha = \sqrt[4]{{}_{2012}l_{2016}} = \sqrt[4]{{}_{2012}l_{2013} \cdot {}_{2013}l_{2014} \cdot {}_{2014}l_{2015} \cdot {}_{2015}l_{2016}} = 1.1474.$$

Da cui segue una variazione relativa media pari a 0.1474 , corrispondente a un incremento medio annuo tra il 2012 e il 2016 del 14.74%.

Tra il 2012 e il 2016 i beni *A* e *B* hanno registrato incrementi medi annui del 3.56% e 14.74%, rispettivamente.

Il bene *A* ha registrato una variazione relativa più



Si ricorda che

$${}_{2012}I_{2016} = {}_{2012}I_{2013} \cdot {}_{2013}I_{2014} \cdot {}_{2014}I_{2015} \cdot {}_{2015}I_{2016}.$$

Si deve trovare il numero indice a base mobile medio. Applicando il criterio di scelta della media secondo Chisini si ottiene

$${}_{2012}I_{2016} = {}_{2012}I_{2013} \cdot {}_{2013}I_{2014} \cdot {}_{2014}I_{2015} \cdot {}_{2015}I_{2016} = \alpha \cdot \alpha \cdot \alpha \cdot \alpha = \alpha^4.$$

Quindi

$$\alpha = \sqrt[4]{{}_{2012}I_{2016}} = \sqrt[4]{{}_{2012}I_{2013} \cdot {}_{2013}I_{2014} \cdot {}_{2014}I_{2015} \cdot {}_{2015}I_{2016}} = 1.1474.$$

Da cui segue una variazione relativa media pari a 0.1474 , corrispondente a un incremento medio annuo tra il 2012 e il 2016 del 14.74%.

Tra il 2012 e il 2016 i beni *A* e *B* hanno registrato incrementi medi annui del 3.56% e 14.74%, rispettivamente.

Il bene *A* ha registrato una variazione relativa più elevata del bene *B*.



3. Sapendo che le quantità vendute dei due beni nel 2012 sono rispettivamente $q_A = 1.2$ e $q_B = 0.1$ calcolare l'indice di Laspeyres dei prezzi al 2013, con base 2012 (ipotizzando, se necessario, che i prezzi dei due beni nell'anno base siano uguali ai corrispondenti *NIBF* dell'anno 2016 con base 2012).

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.



Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.

3. Sapendo che le quantità vendute dei due beni nel 2012 sono rispettivamente $q_A = 1.2$ e $q_B = 0.1$ calcolare l'indice di Laspeyres dei prezzi al 2013, con base 2012 (ipotizzando, se necessario, che i prezzi dei due beni nell'anno base siano uguali ai corrispondenti *NIBF* dell'anno 2016 con base 2012).

$${}^p_{2012}I^L_{2013} = 1.01$$

dal momento che

$${}_{2012}I^{\text{bene } A}_{2013} = {}_{2012}I^{\text{bene } B}_{2013} = 1.01.$$



Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.

3. Sapendo che le quantità vendute dei due beni nel 2012 sono rispettivamente $q_A = 1.2$ e $q_B = 0.1$ calcolare l'indice di Laspeyres dei prezzi al 2013, con base 2012 (ipotizzando, se necessario, che i prezzi dei due beni nell'anno base siano uguali ai corrispondenti *NIBF* dell'anno 2016 con base 2012).

$${}_{2012}^p I_{2013}^L = 1.01$$

dal momento che

$${}_{2012}^p I_{2013}^{\text{bene } A} = {}_{2012}^p I_{2013}^{\text{bene } B} = 1.01.$$

La variazione relativa nel livello generale dei prezzi, secondo l'Indice di Laspeyres, è dell'1%.



Il seguente prospetto riporta i valori rilevati in corrispondenza a 8 unità statistiche con riferimento a due caratteri quantitativi X e Y

x_i	1	1	2	3	3	4	5	5
y_i	3	3	5	6	6	7	10	10

1. Si rappresentino graficamente le coppie dei dati e le medie condizionate di Y in funzione di X e si indichi il grado del polinomio mediante il quale è possibile rappresentare la funzione di regressione.

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.



La funzione di regressione associa a ogni valore x appartenente all'insieme di definizione della variabile X la media di Y condizionata a x , $M(Y|x)$

x	$M(Y x)$
1	3
2	5
3	6
4	7
5	10

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.



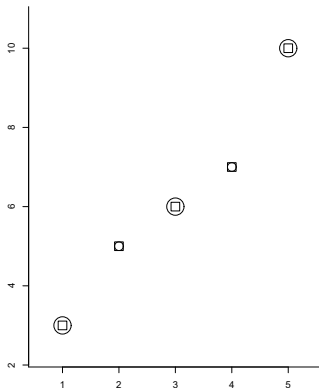
Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.

Le coppie di dati possono essere rappresentate mediante un *bubble plot diagram*.





È possibile rappresentare la funzione di regressione mediante un polinomio di grado al massimo pari a 4.

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.



2. Si calcolino, secondo il metodo dei minimi quadrati, i parametri e l'indice di adattamento dei seguenti modelli di regressione:

$$I) Y^* = a + b \cdot X$$

$$II) Y^* = e^{cX}$$

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.

Si riportano i calcoli necessari¹ per ottenere le stime dei parametri e gli indici di adattamento del modello

$$Y = a + bX + E$$

x_i	y_i	$x_i y_i$	x_i^2	y_i^2
1	3	3	1	9
1	3	3	1	9
2	5	10	4	25
3	6	18	9	36
3	6	18	9	36
4	7	28	16	49
5	10	50	25	100
5	10	50	25	100
24	50	180	90	364
3	6.25	22.5	11.25	45.5
$M(X)$	$M(Y)$	$M(XY)$	$M(X^2)$	$M(Y^2)$

$$\hat{b} =$$



Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.

¹Si osserva che per ragioni di stampa alcuni calcoli potrebbero essere duplicati con riferimento ai diversi modelli oggetto di studio

Si riportano i calcoli necessari¹ per ottenere le stime dei parametri e gli indici di adattamento del modello

$$Y = a + bX + E$$

x_i	y_i	$x_i y_i$	x_i^2	y_i^2
1	3	3	1	9
1	3	3	1	9
2	5	10	4	25
3	6	18	9	36
3	6	18	9	36
4	7	28	16	49
5	10	50	25	100
5	10	50	25	100
24	50	180	90	364
3	6.25	22.5	11.25	45.5
$M(X)$	$M(Y)$	$M(XY)$	$M(X^2)$	$M(Y^2)$

$$\begin{aligned}\hat{b} &= \frac{\text{Cov}(X, Y)}{\text{Var}(X)} = \frac{M(XY) - M(X)M(Y)}{M(X^2) - [M(X)]^2} = \\ &= \frac{22.5 - 3 \cdot 6.25}{11.25 - 3^2} = \frac{3.75}{2.25} = 1.6667\end{aligned}$$

$$\hat{a} =$$

¹Si osserva che per ragioni di stampa alcuni calcoli potrebbero essere duplicati con riferimento ai diversi modelli oggetto di studio



Si riportano i calcoli necessari¹ per ottenere le stime dei parametri e gli indici di adattamento del modello

$$Y = a + bX + E$$

x_i	y_i	$x_i y_i$	x_i^2	y_i^2
1	3	3	1	9
1	3	3	1	9
2	5	10	4	25
3	6	18	9	36
3	6	18	9	36
4	7	28	16	49
5	10	50	25	100
5	10	50	25	100
24	50	180	90	364
3	6.25	22.5	11.25	45.5
$M(X)$	$M(Y)$	$M(XY)$	$M(X^2)$	$M(Y^2)$

$$\hat{b} = \frac{\text{Cov}(X, Y)}{\text{Var}(X)} = \frac{M(XY) - M(X)M(Y)}{M(X^2) - [M(X)]^2} =$$

$$= \frac{22.5 - 3 \cdot 6.25}{11.25 - 3^2} = \frac{3.75}{2.25} = 1.6667$$

$$\hat{a} = M(Y) - \hat{b} M(X) = 1.25$$

$$\rho^2 =$$

¹Si osserva che per ragioni di stampa alcuni calcoli potrebbero essere duplicati con riferimento ai diversi modelli oggetto di studio



Si riportano i calcoli necessari¹ per ottenere le stime dei parametri e gli indici di adattamento del modello

$$Y = a + bX + E$$

x_i	y_i	$x_i y_i$	x_i^2	y_i^2
1	3	3	1	9
1	3	3	1	9
2	5	10	4	25
3	6	18	9	36
3	6	18	9	36
4	7	28	16	49
5	10	50	25	100
5	10	50	25	100
24	50	180	90	364
3	6.25	22.5	11.25	45.5
$M(X)$	$M(Y)$	$M(XY)$	$M(X^2)$	$M(Y^2)$

$$\hat{b} = \frac{\text{Cov}(X, Y)}{\text{Var}(X)} = \frac{M(XY) - M(X)M(Y)}{M(X^2) - [M(X)]^2} =$$

$$= \frac{22.5 - 3 \cdot 6.25}{11.25 - 3^2} = \frac{3.75}{2.25} = 1.6667$$

$$\hat{a} = M(Y) - \hat{b} M(X) = 1.25$$

$$\rho^2 = \frac{[\text{Cov}(X, Y)]^2}{\text{Var}(X) \text{Var}(Y)} = \frac{3.75^2}{2.25 \cdot 6.4375} = 0.9709$$

$$M(\hat{E}^2) = \text{Var}(\hat{E}) = \text{varianza residua} =$$

¹Si osserva che per ragioni di stampa alcuni calcoli potrebbero essere duplicati con riferimento ai diversi modelli oggetto di studio



Si riportano i calcoli necessari¹ per ottenere le stime dei parametri e gli indici di adattamento del modello

$$Y = a + bX + E$$

x_i	y_i	$x_i y_i$	x_i^2	y_i^2
1	3	3	1	9
1	3	3	1	9
2	5	10	4	25
3	6	18	9	36
3	6	18	9	36
4	7	28	16	49
5	10	50	25	100
5	10	50	25	100
24	50	180	90	364
3	6.25	22.5	11.25	45.5
$M(X)$	$M(Y)$	$M(XY)$	$M(X^2)$	$M(Y^2)$

$$\hat{b} = \frac{\text{Cov}(X, Y)}{\text{Var}(X)} = \frac{M(XY) - M(X)M(Y)}{M(X^2) - [M(X)]^2} =$$

$$= \frac{22.5 - 3 \cdot 6.25}{11.25 - 3^2} = \frac{3.75}{2.25} = 1.6667$$

$$\hat{a} = M(Y) - \hat{b} M(X) = 1.25$$

$$\rho^2 = \frac{[\text{Cov}(X, Y)]^2}{\text{Var}(X) \text{Var}(Y)} = \frac{3.75^2}{2.25 \cdot 6.4375} = 0.9709$$

$$M(\hat{E}^2) = \text{Var}(\hat{E}) = \text{varianza residua} = \text{Var}(Y) \cdot (1 - \rho^2) = 0.1875$$

¹Si osserva che per ragioni di stampa alcuni calcoli potrebbero essere duplicati con riferimento ai diversi modelli oggetto di studio





Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.

Si riportano i calcoli necessari² per ottenere le stime approssimate dei parametri e la misura di adattamento del modello

$$Y^* = \exp\{cX\}$$

stimato avvalendosi della trasformata logaritmica per ricondursi alla forma 'linearizzata'

$$\ln(Y) = cX + E$$

x_i	y_i	$\ln y_i$	$x_i \ln y_i$	x_i^2
1	3	1.0986	1.0986	1
1	3	1.0986	1.0986	1
2	5	1.6094	3.2188	4
3	6	1.7918	5.3754	9
3	6	1.7918	5.3754	9
4	7	1.9459	7.7836	16
5	10	2.3026	11.513	25
5	10	2.3026	11.513	25
		13.9413	46.9764	90
		1.7427	5.8721	11.25
		$M(\ln Y)$	$M(X \ln Y)$	$M(X^2)$

$$\hat{c} =$$

²Si osserva che per ragioni di stampa alcuni calcoli potrebbero essere duplicati con riferimento ai diversi modelli oggetto di studio



Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.

Si riportano i calcoli necessari² per ottenere le stime approssimate dei parametri e la misura di adattamento del modello

$$Y^* = \exp\{cX\}$$

stimato avvalendosi della trasformata logaritmica per ricondursi alla forma 'linearizzata'

$$\ln(Y) = cX + E$$

x_i	y_i	$\ln y_i$	$x_i \ln y_i$	x_i^2
1	3	1.0986	1.0986	1
1	3	1.0986	1.0986	1
2	5	1.6094	3.2188	4
3	6	1.7918	5.3754	9
3	6	1.7918	5.3754	9
4	7	1.9459	7.7836	16
5	10	2.3026	11.513	25
5	10	2.3026	11.513	25
		13.9413	46.9764	90
		1.7427	5.8721	11.25
		$M(\ln Y)$	$M(X \ln Y)$	$M(X^2)$

$$\hat{c} = \frac{M(X \ln Y)}{M(X^2)} = \frac{5.8721}{11.25} = 0.522$$

²Si osserva che per ragioni di stampa alcuni calcoli potrebbero essere duplicati con riferimento ai diversi modelli oggetto di studio

Il residuo quadratico medio $M(\hat{E}^2)$ può essere ottenuto calcolando la media del quadrato della variabile residuo $\hat{E} = Y - \hat{Y} = Y - \exp\{\hat{c}X\}$, costruita utilizzando la stima approssimata di c ricavata avvalendosi della forma linearizzata del modello.



Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.



Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.

Il residuo quadratico medio $M(\hat{E}^2)$ può essere ottenuto calcolando la media del quadrato della variabile residuo $\hat{E} = Y - \hat{Y} = Y - \exp\{\hat{c}X\}$, costruita utilizzando la stima approssimata di c ricavata avvalendosi della forma linearizzata del modello. Abbiamo

y_i	$\hat{y}_i$	e_i	e_i^2
3	1.6854	1.3146	1.7282
3	1.6854	1.3146	1.7282
5	2.8406	2.1594	4.663
6	4.7875	1.2125	1.4702
6	4.7875	1.2125	1.4702
7	8.0688	-1.0688	1.1423
10	13.5991	-3.5991	12.9535
10	13.5991	-3.5991	12.9535
		-1.0534	38.1091
		-0.1317	4.7636
		$M(E)$	$M(E^2)$

e

$$M(\hat{E}^2) = 4.7636$$



Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.

$\hat{a} = 1.25$, $\hat{b} = 1.6667$, indice adattamento $M(\hat{E}^2) = 0.1875$ $\hat{c} = 0.5220$, indice
adattamento $M(\hat{E}^2) = 4.7636$

Il modello migliore è $Y^* = 1.25 + 1.6667 \cdot X$.



3. Senza svolgere i calcoli ma motivando la risposta si dica quali valori possono assumere $\eta_{X|Y}^2$ ed $\eta_{Y|X}^2$.

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.



Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.

3. Senza svolgere i calcoli ma motivando la risposta si dica quali valori possono assumere $\eta_{X|Y}^2$ ed $\eta_{Y|X}^2$.
 $\eta_{X|Y}^2 = \eta_{Y|X}^2 = 1$ in quanto le distribuzioni condizionate $Y|X$ e $X|Y$ sono degeneri.



4. Sulla base del miglior modello si preveda il valore di Y in corrispondenza di $X = 6$.

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.



4. Sulla base del miglior modello si preveda il valore di Y in corrispondenza di $X = 6$.
 $Y^*(6) = 1.25 + 1.6667 \cdot 6 = 11.25$.

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.



Un giocatore professionista di basket in un certo momento di una partita deve tirare due tiri liberi in sequenza. Si supponga di sapere che: la probabilità che metta a segno il primo tiro sia 0.1; la probabilità che realizzi il secondo tiro, dato che il primo sia andato a segno, sia pari a 0.2 e che la probabilità che il giocatore realizzi il secondo tiro, dato che il primo NON sia andato a segno sia 0.45.

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.



1. Calcolare la probabilità che realizzi il canestro in entrambi i tiri.

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.



Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.

1. Calcolare la probabilità che realizzi il canestro in entrambi i tiri.

Indicati con S_1 e S_2 gli eventi 'il primo tiro a segno' e 'il secondo tiro a segno' sono disponibili le seguenti informazioni:

$$P(S_1) = 0.1, P(S_2|S_1) = 0.2, P(S_2|\bar{S}_1) = 0.45.$$

Applicando la probabilità composta abbiamo:



Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.

1. Calcolare la probabilità che realizzi il canestro in entrambi i tiri.

Indicati con S_1 e S_2 gli eventi 'il primo tiro a segno' e 'il secondo tiro a segno' sono disponibili le seguenti informazioni:

$$P(S_1) = 0.1, P(S_2|S_1) = 0.2, P(S_2|\bar{S}_1) = 0.45.$$

Applicando la probabilità composta abbiamo:

$$P(S_1 \cap S_2) =$$



Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.

1. Calcolare la probabilità che realizzi il canestro in entrambi i tiri.

Indicati con S_1 e S_2 gli eventi 'il primo tiro a segno' e 'il secondo tiro a segno' sono disponibili le seguenti informazioni:

$$P(S_1) = 0.1, P(S_2|S_1) = 0.2, P(S_2|\bar{S}_1) = 0.45.$$

Applicando la probabilità composta abbiamo:

$$\begin{aligned} P(S_1 \cap S_2) &= P(S_1) \cdot P(S_2|S_1) = \\ &= \end{aligned}$$



Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.

1. Calcolare la probabilità che realizzi il canestro in entrambi i tiri.

Indicati con S_1 e S_2 gli eventi 'il primo tiro a segno' e 'il secondo tiro a segno' sono disponibili le seguenti informazioni:

$$P(S_1) = 0.1, P(S_2|S_1) = 0.2, P(S_2|\bar{S}_1) = 0.45.$$

Applicando la probabilità composta abbiamo:

$$\begin{aligned} P(S_1 \cap S_2) &= P(S_1) \cdot P(S_2|S_1) = \\ &= 0.1 \cdot 0.2 = 0.02 \end{aligned}$$



2. Dire, motivando la risposta, se i risultati dei due tiri sono indipendenti.

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.



2. Dire, motivando la risposta, se i risultati dei due tiri sono indipendenti.

Dal momento che $P(S_2|S_1) = 0.2 \neq P(S_2|\bar{S}_1) = 0.45$ i risultati dei due tiri non sono indipendenti.

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.



3. Calcolare la probabilità che realizzi il canestro in uno solo dei due tiri (considerare tutte le possibili coppie di tiri).

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.



3. Calcolare la probabilità che realizzi il canestro in uno solo dei due tiri (considerare tutte le possibili coppie di tiri).

Abbiamo

$$P((S_1 \cap \bar{S}_2) \cup (\bar{S}_1 \cap S_2)) =$$

Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.



Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.

3. Calcolare la probabilità che realizzi il canestro in uno solo dei due tiri (considerare tutte le possibili coppie di tiri).

Abbiamo

$$P((S_1 \cap \bar{S}_2) \cup (\bar{S}_1 \cap S_2)) = P(S_1 \cap \bar{S}_2) + P(\bar{S}_1 \cap S_2) =$$
$$=$$



Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.

3. Calcolare la probabilità che realizzi il canestro in uno solo dei due tiri (considerare tutte le possibili coppie di tiri).

Abbiamo

$$\begin{aligned} P((S_1 \cap \bar{S}_2) \cup (\bar{S}_1 \cap S_2)) &= P(S_1 \cap \bar{S}_2) + P(\bar{S}_1 \cap S_2) = \\ &= P(S_1) \cdot P(\bar{S}_2|S_1) + P(\bar{S}_1) \cdot P(S_2|\bar{S}_1) = \\ &= \end{aligned}$$



Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.

3. Calcolare la probabilità che realizzi il canestro in uno solo dei due tiri (considerare tutte le possibili coppie di tiri).

Abbiamo

$$\begin{aligned} P((S_1 \cap \bar{S}_2) \cup (\bar{S}_1 \cap S_2)) &= P(S_1 \cap \bar{S}_2) + P(\bar{S}_1 \cap S_2) = \\ &= P(S_1) \cdot P(\bar{S}_2|S_1) + P(\bar{S}_1) \cdot P(S_2|\bar{S}_1) = \\ &= P(S_1) \cdot (1 - P(S_2|S_1)) + (1 - P(S_1)) \cdot P(S_2|\bar{S}_1) = \\ &= \end{aligned}$$



3. Calcolare la probabilità che realizzi il canestro in uno solo dei due tiri (considerare tutte le possibili coppie di tiri).

Abbiamo

$$\begin{aligned} P((S_1 \cap \bar{S}_2) \cup (\bar{S}_1 \cap S_2)) &= P(S_1 \cap \bar{S}_2) + P(\bar{S}_1 \cap S_2) = \\ &= P(S_1) \cdot P(\bar{S}_2|S_1) + P(\bar{S}_1) \cdot P(S_2|\bar{S}_1) = \\ &= P(S_1) \cdot (1 - P(S_2|S_1)) + (1 - P(S_1)) \cdot P(S_2|\bar{S}_1) = \\ &= 0.1 \cdot 0.8 + 0.9 \cdot 0.45 \\ &= \end{aligned}$$



Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.

3. Calcolare la probabilità che realizzi il canestro in uno solo dei due tiri (considerare tutte le possibili coppie di tiri).

Abbiamo

$$\begin{aligned} P((S_1 \cap \bar{S}_2) \cup (\bar{S}_1 \cap S_2)) &= P(S_1 \cap \bar{S}_2) + P(\bar{S}_1 \cap S_2) = \\ &= P(S_1) \cdot P(\bar{S}_2|S_1) + P(\bar{S}_1) \cdot P(S_2|\bar{S}_1) = \\ &= P(S_1) \cdot (1 - P(S_2|S_1)) + (1 - P(S_1)) \cdot P(S_2|\bar{S}_1) = \\ &= 0.1 \cdot 0.8 + 0.9 \cdot 0.45 \\ &= 0.08 + 0.405 = 0.485 \end{aligned}$$



Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.

4. Sapendo che alla fine della stagione il giocatore ha fatto 150 tiri liberi singoli, si calcoli la probabilità che abbia messo a segno almeno 20 tiri, ipotizzando che i risultati dei tiri liberi singoli siano tra loro indipendenti.



Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.

4. Sapendo che alla fine della stagione il giocatore ha fatto 150 tiri liberi singoli, si calcoli la probabilità che abbia messo a segno almeno 20 tiri, ipotizzando che i risultati dei tiri liberi singoli siano tra loro indipendenti.

Indicata con X la variabile casuale 'n. di tiri a segno tra i 150 tiri liberi singoli indipendenti' risulta

$$X \sim$$



Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.

4. Sapendo che alla fine della stagione il giocatore ha fatto 150 tiri liberi singoli, si calcoli la probabilità che abbia messo a segno almeno 20 tiri, ipotizzando che i risultati dei tiri liberi singoli siano tra loro indipendenti.

Indicata con X la variabile casuale 'n. di tiri a segno tra i 150 tiri liberi singoli indipendenti' risulta

$$X \sim \text{Bin}(n = 150, p = 0.1).$$

$$M(X) = 15; \text{Var}(X) = 13.5.$$



Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.

4. Sapendo che alla fine della stagione il giocatore ha fatto 150 tiri liberi singoli, si calcoli la probabilità che abbia messo a segno almeno 20 tiri, ipotizzando che i risultati dei tiri liberi singoli siano tra loro indipendenti.

Indicata con X la variabile casuale 'n. di tiri a segno tra i 150 tiri liberi singoli indipendenti' risulta

$$X \sim \text{Bin}(n = 150, p = 0.1).$$

$$M(X) = 15; \text{Var}(X) = 13.5.$$

Dal momento che $np = 15 > 5$ ed $n(1 - p) = 135 > 5$ è possibile utilizzare l'approssimazione della variabile casuale Binomiale con la variabile casuale Normale

$$X \sim$$



Esercizio 1.

Esercizio 2.

Esercizio 3.

Esercizio 4.

4. Sapendo che alla fine della stagione il giocatore ha fatto 150 tiri liberi singoli, si calcoli la probabilità che abbia messo a segno almeno 20 tiri, ipotizzando che i risultati dei tiri liberi singoli siano tra loro indipendenti.

Indicata con X la variabile casuale 'n. di tiri a segno tra i 150 tiri liberi singoli indipendenti' risulta

$$X \sim \text{Bin}(n = 150, p = 0.1).$$

$$M(X) = 15; \text{Var}(X) = 13.5.$$

Dal momento che $np = 15 > 5$ ed $n(1 - p) = 135 > 5$ è possibile utilizzare l'approssimazione della variabile casuale Binomiale con la variabile casuale Normale

$$X \sim N(\mu = 15, \sigma^2 = 13.5).$$

Si ottiene



4. Sapendo che alla fine della stagione il giocatore ha fatto 150 tiri liberi singoli, si calcoli la probabilità che abbia messo a segno almeno 20 tiri, ipotizzando che i risultati dei tiri liberi singoli siano tra loro indipendenti.

Indicata con X la variabile casuale 'n. di tiri a segno tra i 150 tiri liberi singoli indipendenti' risulta

$$X \sim \text{Bin}(n = 150, p = 0.1).$$

$$M(X) = 15; \text{Var}(X) = 13.5.$$

Dal momento che $np = 15 > 5$ ed $n(1 - p) = 135 > 5$ è possibile utilizzare l'approssimazione della variabile casuale Binomiale con la variabile casuale Normale

$$X \sim N(\mu = 15, \sigma^2 = 13.5).$$

Si ottiene

$$P(X \geq 20) =$$



4. Sapendo che alla fine della stagione il giocatore ha fatto 150 tiri liberi singoli, si calcoli la probabilità che abbia messo a segno almeno 20 tiri, ipotizzando che i risultati dei tiri liberi singoli siano tra loro indipendenti.

Indicata con X la variabile casuale 'n. di tiri a segno tra i 150 tiri liberi singoli indipendenti' risulta

$$X \sim \text{Bin}(n = 150, p = 0.1).$$

$$M(X) = 15; \text{Var}(X) = 13.5.$$

Dal momento che $np = 15 > 5$ ed $n(1 - p) = 135 > 5$ è possibile utilizzare l'approssimazione della variabile casuale Binomiale con la variabile casuale Normale

$$X \sim N(\mu = 15, \sigma^2 = 13.5).$$

Si ottiene

$$P(X \geq 20) = P\left(\frac{X - \mu}{\sigma} \geq \frac{20 - \mu}{\sigma}\right) =$$



4. Sapendo che alla fine della stagione il giocatore ha fatto 150 tiri liberi singoli, si calcoli la probabilità che abbia messo a segno almeno 20 tiri, ipotizzando che i risultati dei tiri liberi singoli siano tra loro indipendenti.

Indicata con X la variabile casuale 'n. di tiri a segno tra i 150 tiri liberi singoli indipendenti' risulta

$$X \sim \text{Bin}(n = 150, p = 0.1).$$

$$M(X) = 15; \text{Var}(X) = 13.5.$$

Dal momento che $np = 15 > 5$ ed $n(1 - p) = 135 > 5$ è possibile utilizzare l'approssimazione della variabile casuale Binomiale con la variabile casuale Normale

$$X \sim N(\mu = 15, \sigma^2 = 13.5).$$

Si ottiene

$$P(X \geq 20) = P\left(\frac{X - \mu}{\sigma} \geq \frac{20 - \mu}{\sigma}\right) = P\left(Z \geq \frac{20 - 15}{3.6742}\right) =$$

$$=$$



4. Sapendo che alla fine della stagione il giocatore ha fatto 150 tiri liberi singoli, si calcoli la probabilità che abbia messo a segno almeno 20 tiri, ipotizzando che i risultati dei tiri liberi singoli siano tra loro indipendenti.

Indicata con X la variabile casuale 'n. di tiri a segno tra i 150 tiri liberi singoli indipendenti' risulta

$$X \sim \text{Bin}(n = 150, p = 0.1).$$

$$M(X) = 15; \text{Var}(X) = 13.5.$$

Dal momento che $np = 15 > 5$ ed $n(1 - p) = 135 > 5$ è possibile utilizzare l'approssimazione della variabile casuale Binomiale con la variabile casuale Normale

$$X \sim N(\mu = 15, \sigma^2 = 13.5).$$

Si ottiene

$$\begin{aligned} P(X \geq 20) &= P\left(\frac{X - \mu}{\sigma} \geq \frac{20 - \mu}{\sigma}\right) = P\left(Z \geq \frac{20 - 15}{3.6742}\right) = \\ &= P(Z \geq 1.36) = 1 - P(Z \leq 1.36) = 1 - 0.9132 = 0.0868. \end{aligned}$$