NUMERO: 2239A ANNO: 2017

A P P U N T I

STUDENTE: Tosti Michela

MATERIA: Statistica - Esercizi + Schemi - Prof. Vicario

STATISTICA

ESERCITAZIONI + SCHEMA X ESAME

FOGLIO 1

ESERCIZIO 1 (F1)

50 b
7 c } 1 E
5 M
:

a) $\mathbb{P}[C] = 7/50$
b) $\mathbb{P}[M] = 5/50$
c) $\mathbb{P}[E] = 1/50$
d) $\mathbb{P}[C \cup M] = \mathbb{P}[C] + \mathbb{P}[M] - \mathbb{P}[C \cap M] = \frac{7}{50} + \frac{5}{50} - \frac{1}{50} = 11/50$
e) $\mathbb{P}[\bar{M} \cap \bar{C}] = 1 - \mathbb{P}[C \cup M] = 39/50$
f) $\mathbb{P}[[M \cap \bar{C}] \cup [C \cap \bar{M}]] = \mathbb{P}[C] + \mathbb{P}[M] - 2\mathbb{P}[E] = \frac{12}{50} - \frac{2}{50} = \frac{10}{50} = \frac{1}{5}$

ESERCIZIO 2 (F1)

$\mathbb{P}[A] = 0,2$ $\quad \mathbb{P}[AB] = 0,08$ $\quad \mathbb{P}[ABC] = 0,02$
$\mathbb{P}[B] = 0,16$ $\quad \mathbb{P}[AC] = 0,05$
$\mathbb{P}[C] = 0,14$ $\quad \mathbb{P}[BC] = 0,04$

a) $\mathbb{P}[\text{almeno } 1] = \mathbb{P}(A] \cup \mathbb{P}(B] \cup \mathbb{P}(C] - \mathbb{P}[AB] - \mathbb{P}[BC] - \mathbb{P}[AC] + \mathbb{P}[ABC] = 0,2 + 0,16 + 0,14 - 0,08 - 0,05 - 0,04 + 0,02 = 0,35$

b) $\mathbb{P}[\text{nessuno}] = 1 - \mathbb{P}[\text{almeno } 1] = 1 - 0,35 = 0,65$

c) $\mathbb{P}[\text{un solo quotidiano}] = \mathbb{P}(A \cap \bar{B} \cap \bar{C}) + \mathbb{P}(\bar{A} \cap B \cap \bar{C}) + \mathbb{P}(\bar{A} \cap \bar{B} \cap C)$
$= \mathbb{P}[A] + \mathbb{P}[B] + \mathbb{P}[C] - 2\mathbb{P}[AC] - \mathbb{P}[BC] - 2\mathbb{P}[AC] + 3\mathbb{P}[ABC] = 0,22$

ESERCIZIO 3 (F1) dado lanciato

$\mathbb{P}[4] = \mathbb{P}[6] =$ doppio delle altre

a) $S\{\ \}$

$$S \begin{cases} \mathbb{P}[4] = \mathbb{P}[6] = 2/8 \\ \mathbb{P}[1,2,3,5] = 1/8 \end{cases}$$

b) $\mathbb{P}[\text{dispari}]$

$\mathbb{P}[D] = \frac{1}{8} + \frac{1}{8} + \frac{1}{8} = 3/8$

c) $\mathbb{P}[N > 3] =$

$\mathbb{P}[>3] = \frac{2}{8} + \frac{1}{8} + \frac{2}{8} = 5/8$

ESERCIZIO 4

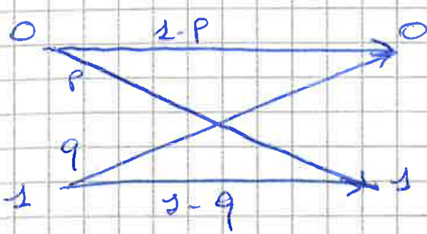
lancio 3 dadi e annoto punteggio minimo

a) $\mathbb{P}[\text{min} = 6] = 1/216$ $\quad 6^3 = 216$ combo

b) $\mathbb{P}[\text{min} > 1] = 5^3/216 = 125/216$

c) $\mathbb{P}[\text{min} = 1] = 1 - \frac{125}{216} = 91/216$

Esercizio 5 (F_2)



$p = 0{,}07$
$q = 0{,}05$
$\mathbb{P}[T_1] = 0{,}6$

a) $\mathbb{P}[\text{Errore}] = \mathbb{P}[R_0 / T_1]\, T_1 + \mathbb{P}[R_1 / T_0]\, T_0 = 0{,}05 \cdot 0{,}6 + 0{,}07 \cdot 0{,}4 = 0{,}058$

b) successive trasmissioni indipendenti: $\mathbb{P}[3 \text{ trasmissioni senza errore}]$

$\left(\mathbb{P}[R_1 / T_1]\, \mathbb{P}[T_1] + \mathbb{P}[R_0 / T_0]\, \mathbb{P}[T_0]\right) = (1 - 0{,}05)\, 0{,}6 + (1 - 0{,}07)\, 0{,}4$
$= 0{,}942$

$0{,}942 \quad 0{,}942 \cdot 0{,}942 = 0{,}8359$

c) $A = 00 \quad \mathbb{P}[\text{Errore} / A]$?

$1 - \mathbb{P}[R_0 / T_0]\, \mathbb{P}[R_0 / T_0] = 1 - (1 - 0{,}07)^2 = 0{,}1351$

Esercizio 6 (F_2)

Dati $P(A) = 0{,}5 \quad P(B) = 0{,}6 \quad \mathbb{P}(A \cup B)$ sapendo $P(B/A) = 0{,}4$

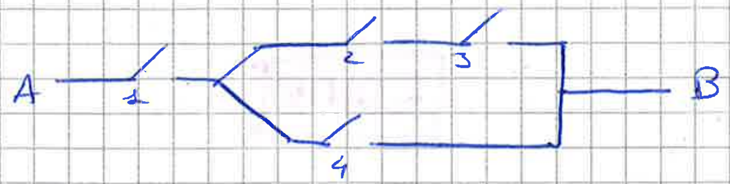
$\mathbb{P}(A \cup B) = \mathbb{P}[A] + \mathbb{P}[B] - \mathbb{P}[A \cap B]$

$\mathbb{P}[A \cap B] = \mathbb{P}[B/A] \cdot \mathbb{P}[A]$

$\mathbb{P}[A \cap B] = 0{,}4 \cdot 0{,}5 = 0{,}2$

$\mathbb{P}(A \cup B) = 0{,}5 + 0{,}6 - 0{,}2 = 0{,}9$

Esercizio 7 (F_2)



$\mathbb{P}[E] = 2p$ per $1, 2, 3$ indipendenti

$\mathbb{P}[E] = p$ per 4 $\qquad \mathbb{P}[\text{conduttore}]$?

$$\mathbb{P}[\bar{E}_1 \cap \bar{E}_2 \cap \bar{E}_3] + \mathbb{P}[\bar{E}_1 \cap \bar{E}_4] - \mathbb{P}[\bar{E}_1 \cap \bar{E}_2 \cap \bar{E}_3 \cap \bar{E}_4]$$

$$= (1 - 2p)(1 - 4p^2 + 4p^3)$$

Esercizio 9 (F2)

$\mathbb{P}(A) = \frac{1}{10}$

$\mathbb{P}(B) = \frac{8}{10}$

$\mathbb{P}(C) = \frac{9}{10}$

a) Vero che $\mathbb{P}$ che si verifichi almeno 1 tra A, B^c, C^c è $< 3/5$?

$$\mathbb{P}(A) + \mathbb{P}(\bar{B}) + \mathbb{P}(\bar{C}) - \mathbb{P}(A \cdot B^c) - \mathbb{P}(A C^c) - \mathbb{P}(B^c C^c) + \mathbb{P}(A\, B^c\, C^c) = \frac{1}{10} + \frac{2}{10} + \frac{1}{10} - \frac{2}{100} - \frac{1}{100} - \frac{2}{100} + \frac{2}{1000} = \frac{44}{125} \text{ vero } < 3/5$$

b) INDIPENDENTI $\mathbb{P}(A \cup B^c)$ e $\mathbb{P}(A \cup B^c \cup C^c)$?

$$\mathbb{P}(A) + \mathbb{P}(B^c) - \mathbb{P}(A \cap B^c) = \frac{1}{10} + \frac{2}{10} - \frac{2}{100} = \frac{7}{25}$$

$\mathbb{P}(A \cup B^c \cup C^c) = 44/125$

CAPITOLO 1 - ESERCIZI SVOLTI -

Esercizio 1

$1/1000$ Affetto da AIDS $P[\text{Affetto}] = \frac{1}{1000}$ $\quad \mathbb{P}[A_R] = 0{,}999$ affetto realmente

Ho AIDS indicato giusto al $0{,}999$

se non ce l'ho può darsi che lo indichi al $0{,}002$

Persona scelta a caso, test indica che è affetto, $\mathbb{P}$ che lo sia realmente?

Bayes: $$P\left[A_R / \text{Affetto}\right] = \frac{1/1000 \cdot 0{,}999}{\frac{1}{1000} \cdot 0{,}999 + \frac{999}{1000} \cdot 0{,}002} = \frac{1}{3} = 0{,}333$$

Esercizio 2

A: 1500 bottiglie $\mapsto$ 10 % difettose

B: 2500 " $\longmapsto$ 5 % "

} 2 bottiglie estratte senza reinmissione da 1 scatola casualmente

1) $$P[\text{entrambe difettose}] = \frac{150}{1500} \cdot \frac{149}{1499} \cdot \frac{1}{2} + \frac{125}{2500} \cdot \frac{124}{2499} \cdot \frac{1}{2} = 0{,}00621$$

$$\mathbb{P}[D_1] \cdot \mathbb{P}[D_2] \cdot \mathbb{P}[A] + \mathbb{P}[D_1]\, \mathbb{P}[D_2] \cdot \mathbb{P}[B]$$

2) $\mathbb{P}\left[A / \text{entrambe difettose}\right]$

Bayes = $$\frac{\mathbb{P}\left[\text{entrambe dif} / A\right] \mathbb{P}[A]}{\mathbb{P}\left[\text{entrambe dif} / A\right] \mathbb{P}[A] + \mathbb{P}\left[\text{entrambe dif} / B\right] \mathbb{P}[B]} =$$

$$= \frac{\frac{150}{1500} \cdot \frac{149}{1499} \cdot \frac{1}{2}}{\frac{150}{1500} \cdot \frac{149}{1499} \cdot \frac{1}{2} + \frac{125}{2500} \cdot \frac{124}{2499} \cdot \frac{1}{2}} = 0{,}8$$

FOGLIO 3

ESERCIZIO 1 (F3)

100 dadi: 80 comuni; 20: $P[6] = 0,5$, $P[1,2,3,4,5] = 0,1$

lancio 1 dado
X : va indica il numero apparso

a) calcolare la funzione di densità, valore atteso e varianza della variabile casuale X

x	1	2	3	4	5	6
$f_X(x)$	$\frac{23}{150}$	$\frac{23}{150}$	$\frac{23}{150}$	$\frac{23}{150}$	$\frac{23}{150}$	$\frac{35}{150}$

$0,8 \cdot \frac{1}{6} + 0,2 \cdot 0,1$

$$\mu_x = \frac{23}{150}\cdot 1 + \frac{23}{150}\cdot 2 + \frac{23}{150}\cdot 3 + \frac{23}{150}\cdot 4 + \frac{23}{150}\, 5 + \frac{35}{150}\, 6 = 3,7$$

$$Var_x = \frac{23}{150}\cdot 1 + \frac{23}{150}\, 2^2 + \frac{23}{150}\cdot 3^2 + \frac{23}{150}\cdot 4^2 + \frac{23}{150}\, 5^2 + \frac{35}{150}\, 6^2 - 3,7^2 = 3,14$$

b) $P[x \leq 2]$ e $P[x \geq 5]$

$$P[x \leq 2] = \frac{23}{150} + \frac{23}{150} = \frac{46}{150} = 0,307$$

$$P[x \geq 5] = \frac{23}{150} + \frac{35}{150} = \frac{58}{150} = 0,387$$

ESERCIZIO 2 (F3)

S_1 S_2 S_3 S_{4A} S_{4B}

estraggo senza reintroduzione 2 sfere
X: somma numeri corrispondenti alle sfere

a) spazio campioni

$S: \{(1,2), (1,3), (1,4_A), (1,4_B), (2,1), (2,3), (2,4_A), (2,4_B), (3,1), (3,2), (3,4_A), (3,4_B), (4_A,1), (4_A,2), (4_A,3), (4_A,4_B), (4_B,1), (4_B,2), (4_B,3), (4_B,4_A)\}$

b) $f_X(x)$

x	3	4	5	6	7	8
$f_X(x)$	$\frac{2}{20}$	$\frac{2}{20}$	$\frac{6}{20}$	$\frac{4}{20}$	$\frac{4}{20}$	$\frac{2}{20}$

Esercizio 3 (F3)

$f_X(x) = Ke^{-|x|}$

Per il punto ⓐ
se $x>0 \mapsto Ke^{-x}$
$x<0 \mapsto Ke^{x}$

Per il punto ⓑ
$\int_{-\infty}^{0} e^{u}$ $x<0$
$\int_{-\infty}^{0} e^{u} + \int_0^x e^{-u}$ $x>0$

in 0 ho discontinuità dello spettro $x>0$

ⓐ K affinché sia funzione di densità

$$Ke^{-|x|} > 0 \mapsto K>0$$

$$\int_{-\infty}^{+\infty} Ke^{-|x|} = 1 \mapsto \int_0^{+\infty} Ke^{-x} + \int_{-\infty}^{0} Ke^{x} = 1$$

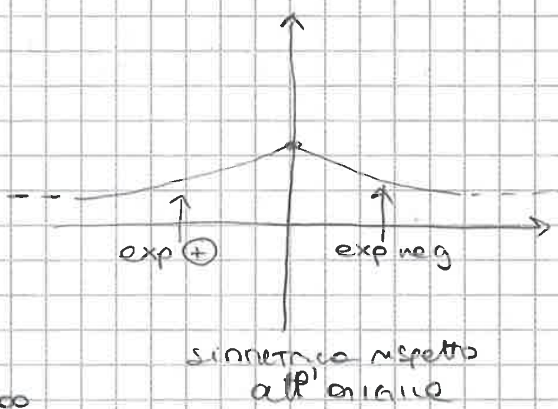
$$K\int_0^{+\infty} e^{-x} + K\int_{-\infty}^{0} e^{x} = 1 \mapsto K\left(-e^{-x}\right)_0^{+\infty} + K\left(e^{x}\right)_{-\infty}^{0}$$

$$= K\left(-e^{-\infty}+e^{0}\right) + K\left(e^{0}-e^{-\infty}\right) = 2K = 1 \qquad K = \frac{1}{2}$$

ⓑ determinare $F_X(x) \mapsto \int_{-\infty}^{x} f_X(u)\,du$

$$\begin{cases} \int_{-\infty}^{0}\frac{1}{2}e^{u} + \int_0^x \frac{1}{2}e^{-u} & x>0 \quad \frac{1}{2} + \frac{1}{2}\left(-e^{-x}+e^{0}\right) = \frac{1}{2} - \frac{1}{2}e^{-x} + \frac{1}{2} = 1 - \frac{1}{2}e^{-x} \\ \int_{-\infty}^{x}\frac{1}{2}e^{u}\,du & x<0 \Rightarrow \frac{1}{2}e^{x} \end{cases}$$

poiché in zero ho una discontinuità



ⓒ valore atteso e varianza di X

$$\mu_X = \frac{1}{2}\int_{-\infty}^{+\infty} e^{-|x|}x\,dx = \frac{1}{2}\int_{-\infty}^{0} e^{x}x + \frac{1}{2}\int_0^{+\infty} e^{-x}x = 0$$

$$\mathrm{Var}[X] = E[X^2] - \left(E[X]\right)^2 = \int_{-\infty}^{+\infty} x^2 f_X(x)\,dx = \frac{1}{2}\int_{-\infty}^{0} x^2e^{x}\,dx + \frac{1}{2}\int_0^{+\infty} x^2e^{-x}\,dx$$

$E[X^2] \to g(x)$; $E[X] \to 0$

$$= \; ? \; = 2$$

$k=0 \mapsto \binom{4}{0} 0{,}5^0 (1-0{,}5)^{4-0} = 0{,}5^4 = 1/16$

$k=1 \mapsto \binom{4}{1} 0{,}5^1 (1-0{,}5)^3 = 1/4$

$k=2 \Rightarrow \binom{4}{2} 0{,}5^2 (1-0{,}5)^2 = \frac{4!}{2!\,2!} 0{,}5^2 \cdot 0{,}5^2 = 3/8$

$k=3 \mapsto \binom{4}{3} 0{,}5^3 (0{,}5)^1 = \frac{4!}{3!\,1!} 0{,}5^3\, 0{,}5 = 1/4$

$k=4 \mapsto \binom{4}{4} 0{,}5^4 (0{,}5)^0 = \frac{4!}{4!\,0!} = 1 \cdot 0{,}5^4\, 1 = \frac{1}{16}$

Sx
$\mapsto \frac{1}{16} + \frac{1}{4} = \frac{1+4}{16} = \frac{5}{16} + \frac{3}{8} = \frac{5+6}{16} = \frac{11}{16} > \frac{1}{2}$ $x=2$

Dx
$\mapsto \frac{1}{16} + \frac{1}{4} + \frac{3}{8} = \frac{11}{16} > \frac{1}{2}$ $x=2$

$x = 2$ mediana

b) $n=5$ $p=0{,}5$

k da 0 a 5

0	1	2	3	4	5
$\frac{1}{32}$	$\frac{5}{32}$	$\frac{5}{16}$	$\frac{5}{16}$	$\frac{5}{32}$	$\frac{1}{32}$

$k=0 \mapsto \binom{5}{0} 0{,}5^0\, 0{,}5^5 = \frac{1}{32}$

$k=1 \mapsto \binom{5}{1} 0{,}5^1\, 0{,}5^4 = 5/32$

$k=2 \mapsto \binom{5}{2} 0{,}5^2\, 0{,}5^3 = \frac{5!}{2!\,3!} 0{,}5^2 \cdot 0{,}5^3 = 5/16$

$k=3 \mapsto \binom{5}{3} 0{,}5^3\, 0{,}5^2 = \frac{5!}{3!\,2!} 0{,}5^3\, 0{,}5^2 = \frac{5}{16}$

$k=4 \longmapsto \binom{5}{4} 0{,}5^4\, 0{,}5^1 = 5/32$

$k=5 \longmapsto \binom{5}{5} 0{,}5^5\, 0{,}5^0 = \frac{1}{32}$

Da Sx: $\frac{1}{32} + \frac{5}{32} + \frac{5}{16} = \frac{1+5+10}{32} = \frac{16}{32} = \frac{1}{2}$ stop $x=2$

Da Dx $\frac{1}{32} + \frac{5}{32} + \frac{5}{16} = \frac{16}{32} = \frac{1}{2}$
Stop $x=3$

↓

Scelgo il minore
mediana $x=2$

© Proprietà riservata dell'autore - Digitalizzazione e distribuzione a cura del CENTRO APPUNTI - Corso Luigi Einaudi, 55 - Torino / Pagina 13 di 74

FOGLIO 4

Esercizio 1 (F4)

20 passeggeri ↦ 3 hanno droghe
Ispeziono 4 passeggeri. ℙ(almeno 1 ha sostanze non consentite)?

X: VC: indica ha droghe

No reimmissione: ipergeometrica

$$f_X(x) = f_X(x, N, K, n) = f_X(0, 20, 4, 3)$$

$$= \frac{\binom{4}{0}\binom{20-4}{3-0}}{\binom{20}{3}} = \frac{1 \cdot \frac{16!}{3!\,13!}}{\frac{20!}{3!\,17!}} \;\; \frac{14 \cdot 15 \cdot 16}{20 \cdot 19 \cdot 18} = 0{,}491 \mapsto \mathbb{P}\ \text{nex ha droghe}$$

$$\mathbb{P}[\text{almeno 1 a droghe}] = 1 - \mathbb{P}[\text{NEX}] = 1 - 0{,}491 = 0{,}509$$

Esercizio 2 (F4)

ℙ[difettose] = 0,1

1a) confezioni da $C = 8$ pezzi
ℙ[al più due difettose in ogni C] X: lampadino difettoso

No reimmissione: ipergeometrica ↦ no perché prove indipendenti, uso binomiale

con x da 0 a 2

$$(x, 8, 0{,}1) = \binom{8}{0} 0{,}1^0 (0{,}9)^8 + \binom{8}{1} 0{,}1^1 (0{,}9)^7 + \binom{8}{2} 0{,}1^2 (0{,}9)^6$$

$$= 0{,}430 + 0{,}383 + 0{,}1488 = 0{,}9619$$

1b) $f = 150$ confezioni: se 2 o più confezioni contengono più di 2 difettose è rifiutato, ℙ che sia accettato?

$$\text{Binomiale} \sim (\underset{\substack{\uparrow \\ 0,1}}{x}, 150, \underset{\substack{\uparrow \\ 0{,}0381}}{(1 - 0{,}9619)}) = \binom{150}{0} 0{,}0381^0 (0{,}9619)^{150}$$

$$+ \binom{150}{1} 0{,}0381^1 (0{,}9619)^{149} = 0{,}0203$$

ESERCIZIO 3 (F4)

Urna contiene: 3R 4B 5N estrazioni con reimmissione

a) $\mathbb{P}$[5 estrazioni] ↦ 3 palline nere

$\mathbb{P}[N] = 5/12$ Binomiale $f_x(x) = f_x(\underset{3}{x}, 5, 5/12)$

$$\binom{5}{3} \left(5/12\right)^3 \left(1 - 5/12\right)^{5-3} = \frac{5!}{3!\,2!} \left(\frac{5}{12}\right)^3 \left(\frac{7}{12}\right)^2 = 0{,}246$$

b) $\mathbb{P}$[1 pallina nera al 3° tentativo] ? INDIPENDENTI

$$\mathbb{P}[N] = 5/12 \qquad \frac{7}{12} \cdot \frac{7}{12} \cdot \frac{5}{12} = \frac{245}{1728} = 0{,}1418$$

c) N° estrazioni necessarie affinché la frequenza di 1 pallina nera si discosti dalla $\mathbb{P}$ di uscita in una singola estrazione meno di $\varepsilon = 0{,}01$ con $\mathbb{P}$ non $< 0{,}9$

$$\mathbb{P}[|X - \mu_x| < \underbrace{t\sigma_x}_{\varepsilon}] \geq \boxed{1 - 1/t^2} \qquad \mathbb{P} = 0{,}9$$

p = alla probabilità della pallina nera

$$\varepsilon = t\sqrt{\frac{p(1-p)}{n}} = 0{,}01$$

$$\begin{cases} t\sqrt{\frac{p(1-p)}{n}} = 0{,}01 \mapsto 3{,}16\sqrt{\frac{5/12\,(7/12)}{n}} = 0{,}01 \mapsto \frac{5/12\,(7/12)}{(3{,}16\cdot 10^{-3})^2} \mapsto n = 24300 \\ 1 - 1/t^2 = 0{,}9 \mapsto t^2 - 1 = 0{,}9\,t^2 \quad 0{,}1\,t^2 = 1 \quad t^2 = 10 \quad t = 3{,}16 \end{cases}$$

d) Ripetere il punto (c) (e confrontare risultati) se $\varepsilon = 0{,}1$

$$\mathbb{P}[|X - \mu_x| < t\sigma_x] \geq 1 - 1/t^2 \quad \mathbb{P} = 0{,}9$$

$$\varepsilon = t\sqrt{\frac{p(1-p)}{n}}$$

$$\begin{cases} 0{,}1 = 3{,}16\sqrt{\frac{5/12 \cdot 7/12}{n}} \mapsto 1{,}0014 \cdot 10^{-3} = \frac{0{,}243}{n} \mapsto n = 244 \\ t = 3{,}16 \end{cases}$$

meno estrazioni del precedente

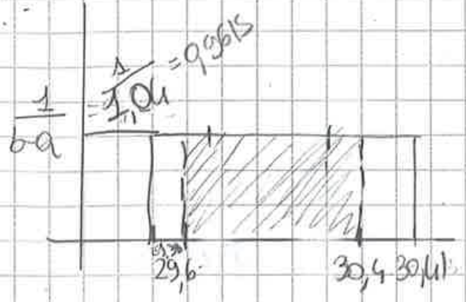
FOGLIO 5

ESERCIZIO 1 (F5)

$F_1 \mapsto 70\%$ UNF $\mu = 29{,}9$ $\sigma^2 = 0{,}09$ } limiti specifico $(29{,}6\ ;\ 30{,}4)$
$F_2 \mapsto 30\%$ UNIF $(29{,}5 - 30{,}5)$ }

% di prodotto ricevuto dal supermercato che non soddisfa le prescrizioni?

(F1)



$$\mathbb{E}[x] \begin{cases} \dfrac{a+b}{2} = 29{,}9 \Rightarrow a = 59{,}8 - b \\ \mathrm{var}[x] = \dfrac{(b-a)^2}{12} = 0{,}09 \Rightarrow b - a = 1{,}039 \end{cases}$$

$$\Rightarrow \begin{cases} a = 59{,}8 - 30{,}41 = 29{,}38 \\ b = 59{,}8 + 1{,}039 - b \mapsto b = 30{,}41 \end{cases}$$

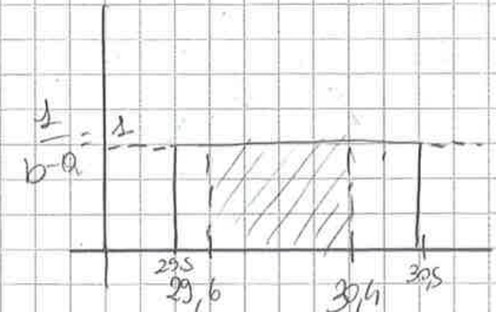
$$\int_{29{,}6}^{29{,}4} 0{,}9615\, dx = 0{,}9615\,(30{,}40 - 29{,}60) = 0{,}769$$

$\mathbb{P}_{\text{soddisfi}}$

$\downarrow$

$\mathbb{P}_{\text{non sodd}} = 1 - 0{,}769 = 0{,}231$

(F2)



$$\int_{29{,}6}^{30{,}4} 1\, dx = 30{,}4 - 29{,}6 = 0{,}8$$

$\uparrow$

$\mathbb{P}_{\text{soddisfi}}$

$\downarrow$

$\mathbb{P}_{\text{non sodd}} = 1 - 0{,}8 = 0{,}2$

$$\mathbb{P}_{F_1} \cdot \mathbb{P}_{\text{non sodd}} + \mathbb{P}_{F_2} \mathbb{P}_{\text{non sodd}}$$

$$0{,}7 \cdot 0{,}231 + 0{,}3 \cdot 0{,}2$$

$$= 0{,}22$$

Esercizio 3 (FS)

. Diametro in mm di un certo tipo di valvole distribuito con $N(65, \sigma^2)$
84% ha $d > 59,5$

a) determina % di valvole il cui diametro differisce dal diametro medio per più di 1 volta e mezzo la deviazione standard

$$\mathbb{P}[d > 65 \pm \sigma \cdot 1,5]$$

$$\mapsto 1 - \left[\frac{65 - 1,5\sigma - \mu}{\sigma} < z < \frac{65 + 1,5\sigma - \mu}{\sigma}\right] = 1 - \left[-1,5 < z < 1,5\right]$$

perché voglio quelle esterne

$$1 - z < 1,5 - \left[1 - \left[z < 1,5\right]\right] = 1 - \left[0,9332 - \left[1 - 0,9332\right]\right] = 0,1336$$

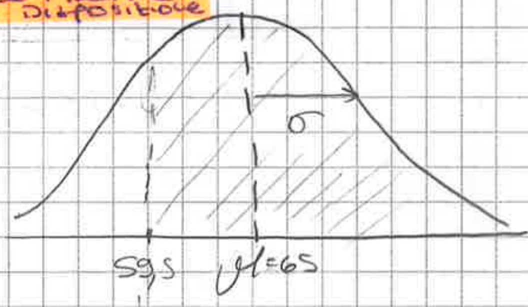
b) % di valvole con diametro > 68 tra quelle con diametro > dello $\mu + \frac{1}{2}\sigma$

RICORDA: trova σ analizzando bene i dati e disposizione

Mi ricavo prima di tutto la σ

$$d > \mu + \frac{1}{2}\sigma$$

$$\frac{d > 68}{d > \mu + \frac{1}{2}\sigma} = \frac{d > 68 \cap d > \mu + \frac{1}{2}\sigma}{d > \mu + \frac{1}{2}\sigma}$$



Devo trovare σ

$$\mathbb{P}[d > 59,5] = 0,84$$

$$z > \frac{59,5 - 65}{\sigma} = 0,84$$

$$z > -5,5/\sigma = 0,84$$

$$z_{0,84} = \frac{-5,5}{\sigma}$$

$$0,995 = 5,5/\sigma$$

$$\sigma = 5,527$$

$$\frac{d > 68}{d > 65 + 0,5 \cdot 5,527} = \frac{z > \frac{68 - 65}{5,527}}{z > \frac{67,7635 - 65}{5,527}} = \frac{z > 0,54}{z > 0,5}$$

$$= \frac{1 - z < 0,54}{1 - z < 0,5} = \frac{1 - 0,7054}{1 - 0,6915} = 0,9549$$

c) avendo acquistato 12 valvole, P che almeno 2 abbiano diametro compreso tra 63,5 e 66,5

FAI SEMPRE MOLTA ATTENZIONE ALLE PAROLE USATE

$$63,5 < d < 66,5$$

$$\frac{63,5 - 65}{5,527} < z < \frac{66,5 - 65}{5,527} = -0,27 < z < 0,27$$

$$z < 0,27 - (1 - z_{0,27}) = 0,2128$$

Binomiale $(X, 12, 0,2128)$: $1 - \binom{12}{0} 0,2128^0 (0,7872)^{12} + \binom{12}{1} 0,2128^1 (0,7872)^{11} = 1 - 0,2403 = 0,7597$
$0,1$

ESERCIZIO 5 (FS)

Tempo di durata in h di un componente elettronico segue esponenz.
$\mu = 2$ h
$\mathbb{P}$[su 10 componenti almeno 2 abbiano durata > di 3 h?]

$$\mu = \frac{1}{\lambda} \mapsto \lambda = \frac{1}{2} = 0{,}5 \mapsto \text{var} = \frac{1}{0{,}5^2} = 4$$

$$f_X(x) = f_X(x, \lambda) = \lambda e^{-\lambda x} = 0{,}5\, e^{-0{,}5x}$$

$$\text{durata} \geq 3h \mapsto \int_3^{+\infty} \frac{1}{2} e^{-\frac{1}{2}t}\, dt = e^{-3/2}$$

L'ESPONENZIALE devo integrare nel tempo che mi serve ovvero tra le 3 ore e $+\infty$

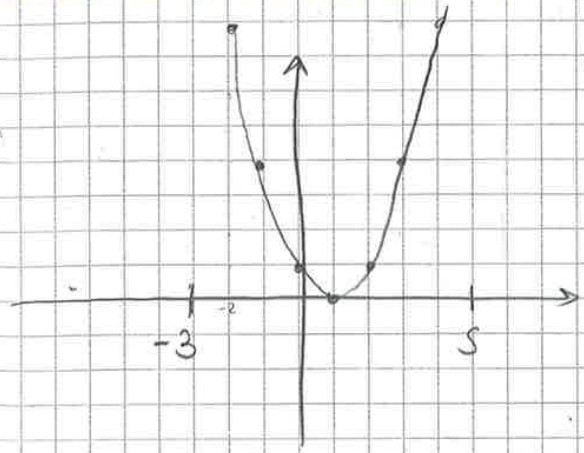
ottengo binomiale $(x, 10, e^{-3/2})$

$$\mathbb{P}[x \geq 2] = 1 - \mathbb{P}[x < 2] = 1 - \binom{10}{0} (e^{-3/2})^0 (1 - e^{-3/2})^{10} + \binom{10}{1} (e^{-3/2}) (1 - e^{-3/2})^9$$

$= 0{,}69$

ESERCIZIO 3 (F6)

Sia data la variabile casuale x con distribuzione uniforme sull'intervallo $(-3,5)$ calcolare la distribuzione e la media di $Y=(x-1)^2$



$y=(x-1)^2$

$x=-2$	$y=9$
$x=-1$	$y=4$
$x=0$	$y=1$
$x=1$	$y=0$
$x=2$	$y=1$
$x=3$	$y=4$
$x=4$	$y=9$

$D_x(-3,5) \mapsto$ SPEZZO:

$D_x(-3,1) \cup D_x(-1,5)$

$D_y(0,+\infty) \hookrightarrow (0,16)$

$\Downarrow$

$x=\pm\sqrt{y}+1$

$$f_y(y)=\begin{cases}\left|\pm\dfrac{\sqrt{y}+1}{dy}\right|\cdot\dfrac{1}{5+3} \\ 0 \quad \text{altrove}\end{cases}$$

Uniforme

$f_x(x,a,b)=\dfrac{1}{b-a}$

$\oplus \mapsto f_y(y)=\left|\dfrac{1}{2\sqrt{y}}\cdot\dfrac{1}{8}\right|$

$\ominus \mapsto f_y(y)=\left|-\dfrac{1}{2\sqrt{y}}\,\dfrac{1}{8}\right|$

$\mapsto \dfrac{1}{2\sqrt{y}}\cdot\dfrac{1}{8}+\dfrac{1}{2\sqrt{y}}\,\dfrac{1}{8}=\dfrac{2}{16\sqrt{y}}=\dfrac{1}{8\sqrt{y}}$

RICORDA

- Se lo spezzi in 2 poi le due parti le devi sommare (in valore assoluto)
- la media, se usi l'integrale della y devi usare gli estremi della y e dentro devi mettere la funzione con la y

$\mu[x]=\dfrac{a+b}{2}=\dfrac{2}{2}=1$

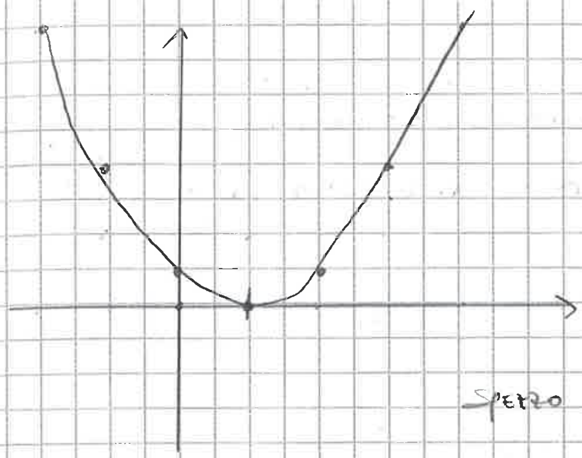
$\mu[y]=\displaystyle\int_0^{16} y\,\frac{1}{8\sqrt{y}}=\frac{1}{8}\int_0^{16} y\cdot y^{-\frac12}=\frac{1}{8}\int_0^{16} y^{\frac12}\,dy=\frac{1}{8}\left(\frac{2}{3}y^{\frac32}\right)_0^{16}=$

$=\dfrac{1}{8}\left(\dfrac{2}{3}16^{\frac32}\right)=\dfrac{16}{3}$

© Proprietà riservata dell'autore - Digitalizzazione e distribuzione a cura del CENTRO APPUNTI - Corso Luigi Einaudi, 55 - Torino / Pagina 25 di 74

Esercizio 5 (F6)

X v.c Gaussiana $\mathcal{N}(2,4)$ calcolare la legge della variabile $Y=(\frac{X}{2}-1)^2$ e la sua media



$x=0 \quad y=1$
$x=2 \quad y=0$
$x=4 \quad y=1$
$x=-2 \quad y=4$
$x=-4 \quad y=9$
$x=6 \quad y=4$
$x=8 \quad y=9$

spezzo $D(x)\ (-\infty,2) \cup D_x\ (2,+\infty)$

ovunque $D_y\ (0,+\infty)$

$$x=(\pm\sqrt{y}+1)\,2$$

$$\begin{cases} \left|\dfrac{d\,\pm2\sqrt{y}+2}{dy}\right| \cdot \dfrac{1}{\sqrt{2\pi 4}} \cdot e^{\left(-\frac{1}{2}\frac{[(\pm2\sqrt{y}+2)-2]^2}{4}\right)} & (0,+\infty) \\ 0 & \text{altrove} \end{cases}$$

NOTA BENE
se $g(x)=(x-\mu_x)^2$ $\mathbb{E}[g(x)]=\text{var}\,x$
se $g(x)=x$ $\mathbb{E}[g(x)]=\mathbb{E}[x]$

$$\oplus \rightarrow \frac{\cancel{2}\,1}{\cancel{2}\sqrt{y}} \cdot \frac{1}{\sqrt{8\pi}}\, e^{\left(-\frac{1}{2}y\right)}$$

$$\ominus \rightarrow \frac{1}{\sqrt{y}} \cdot \frac{1}{\sqrt{8\pi}}\, e^{\left(-\frac{1}{2}y\right)}$$

sommo: $\dfrac{\cancel{2}}{\cancel{2}\sqrt{2\pi y}}\, e^{-\frac{1}{2}y}$

perché è $(x-2)^2$

e la media è: $\mathbb{E}[y]=\mathbb{E}\left[\left(\frac{x}{2}-1\right)^2\right]=\frac{1}{4}\mathbb{E}[(x-2)^2]=\frac{1}{4}\,4=1$

se $g(x)=(x-\mu_x)^2$
si ha che $\mathbb{E}[g(x)]=\text{var}[x]$

$$= \frac{\cancel{e^{-1}}\frac{1^1}{1!}\cdot \cancel{e^{-1,6}}\frac{1,6^2}{2!}}{\cancel{e^{-2,6}}\frac{2,6^3}{3!}} = \frac{1,6^2}{2}\cdot\frac{3!}{2,6^3} = 0,437$$

ESERCIZIO 2 (F7)

N° mezzi che circolano in autostrada durante le ore centrali di una giornata d'inverno sono distribuite in un processo di Poisson con media 3 passaggi al minuto

1) So che in 15 min passano 35 veicoli, P che nei primi 5 min non ne siano passati

$$\frac{P[N_v(5)=0]\times P[N_v(10)=35]}{P[N_v(15)=35]} =$$

$$\left.\begin{array}{l} N_v(5)=0 \mapsto \frac{3}{1}\cdot 5 = 15 \\ N_v(10)=35 \mapsto \frac{3}{1}\cdot 10 = 30 \\ N_v(15)=35 \mapsto 30+15 = 45 \end{array}\right\} \mapsto \frac{\cancel{e^{-15}}\frac{15^0}{0!}\cdot \cancel{e^{-30}}\frac{30^{35}}{\cancel{35!}}}{\cancel{e^{-45}}\frac{45^{35}}{\cancel{35!}}} = \frac{30^{35}}{45^{35}} = 6,87\cdot 10^{-7}$$

2) N° camper approssimabile con Poisson indipendente dai veicoli di altro tipo, 0,2 passaggi al minuto, P che passi un camper in 10 min se in tot sono passati 10 veicoli

$$\frac{P[N_c(10)=1]\cdot P[N_{v-c}(10)=9]}{N_v(10)=10} =$$

$$\left.\begin{array}{l} N_c(10)=1 \mapsto \frac{0,2}{1}\cdot 10 = 2 \\ N_{v-c}(10)=9 \mapsto 30-2 = 28 \\ N_v(10)=10 \mapsto \frac{3}{1}\cdot 10 = 30 \end{array}\right\} \mapsto \frac{\cancel{e^{-2}}\frac{2^1}{1!}\cdot \cancel{e^{-28}}\frac{28^9}{9!}}{\cancel{e^{-30}}\frac{30^{10}}{10!}}$$

$$= 2\cdot\frac{28^9}{9!}\cdot\frac{10!}{30^{10}} = 0,3583$$

ESERCIZIO 4 (F7)

Il n° di aperture della porta automatica di una azienda nell'orario compreso tra le 17.30 e le 20 sono distribuite secondo Poisson di intensità pari a 3 aperture ogni 2 min.
Il giorno 4 febbraio 2010 la porta si è aperta 5 volte tra le 18.25 e le 18.29 calcola la P:

1) tra le 18.29 e le 18.27 si è aperta 3 volte

$$\frac{\mathbb{P}(N_A(2)=3)\cdot\mathbb{P}(N_A(2)=2)}{\mathbb{P}(N_A(4)=5)} = \frac{\cancel{e^{-3}}\frac{3^3}{3!}\cdot\cancel{e^{-3}}\frac{3^2}{2!}}{\cancel{e^{-6}}\frac{6^5}{5!}} = 0{,}3125$$

$N_A(2)=3 \longmapsto \frac{3}{2}\cdot 2 = 3$

$N_A(2)=2 \longmapsto \frac{3}{2}\cdot 2 = 3$

$N_A(4)=5 \longmapsto \frac{3}{2}\cdot 4 = 6$

2) tra le 18.29 e le 18.31 si sia aperta 3 volte

$$\mathbb{P}[N_A(2)=2] = e^{-3}\frac{3^2}{2!} = 0{,}224$$

© Proprietà riservata dell'autore - Digitalizzazione e distribuzione a cura del CENTRO APPUNTI - Corso Luigi Einaudi, 55 - Torino / Pagina 31 di 74

$$Z > \frac{1,21 - 13,916}{23,62} = -0,538 \mapsto \mathbb{P}(z < 0,39) = 0,6517$$

b) $\mathbb{P}$ che le precipitazioni del 2013 superino quelle del 2012 per più di 14 mm

$$P_{2013} + 14 < P_{2014}$$

$$\underbrace{P_{2014} - P_{2013}}_{Y} > 14$$

$$\mathbb{E}[Y] = 0$$

$$\sigma_Y = \sqrt{278,89 + 278,89} = 23,62$$

$$z > \frac{14 - 0}{23,62} \mapsto z > 0,59$$

$$\mathbb{P}[z > 0,59] = 1 - \mathbb{P}[z < 0,59] = 0,2776$$

ESERCIZIO 4 (F8)

Considerare 3 variabili casuali X_1, X_2, X_3 con distribuzioni esponenziali con parametro rispettivamente 1, 2, 4 mutuamente non correlate.

1) μ e varianza di $Y = X_1 + X_2 - X_3$

$$\mu[Y] = \frac{1}{1} + \frac{1}{2} - \frac{1}{4} = \frac{3}{2} - \frac{1}{4} = \frac{5}{4}$$

$$\text{var}[Y] = \frac{1}{1^2} + \frac{1}{2^2} + \frac{1}{4^2} = 1 + \frac{1}{4} + \frac{1}{16} = \frac{16 + 4 + 1}{16} = \frac{21}{16}$$

$$\sigma_Y = \sqrt{21/16} = 1,14$$

2) $U = 2X_1 - X_3$ $\quad$ Sono non correlate?
$V = X_1 + X_2 + X_3$ $\quad$ calcolo coeff. di correlazione!

$$\text{cov}[U,V] = 2(X_1, X_1) + \cancel{2(X_1 X_2)} + \cancel{2(X_1 X_3)} - \cancel{(X_3 X_1)} - \cancel{(X_3 X_2)} - (X_3 X_3)$$

perché partendo dal presupposto che sono indipendenti

$$= 2\,\text{var}(X_1) - \text{var}(X_3)$$

$$= 2 - \frac{1}{16} = \frac{31}{16}$$

COEFF DI CORRELAZ.

$$\rho = \frac{\text{cov}[U,V]}{\sigma_U \sigma_V}$$

$$\sigma_U = \sqrt{4 \cdot 1^2 + \left(\frac{1}{16}\right)} = 2,01 \qquad \sigma_V = \sqrt{1 + \frac{1}{4} + \frac{1}{16}} = 1,14 \qquad \rho = \frac{31/16}{2,01 \cdot 1,16} = 0,84$$

correlate

b) usando μ e var del punto 1 calcolo $P(\bar{X} \geq 55)$ e $P(58 \leq \bar{X} \leq 62)$

$\bar{X} \mapsto X/3$

$\mu[\bar{x}] = \frac{180}{3} = 60$

$var[\bar{x}] = \frac{6,7}{3} = 2,23$

$z \geq \frac{55-60}{2,23} \mapsto z \geq -2,24$

$\hookrightarrow z \leq 2,24 = 0,9875$

$58 \leq z \leq 62$

$\frac{58-60}{2,23} \leq z \leq \frac{62-60}{2,23}$

$-0,897 \leq z \leq 0,897$

$z \leq 0,897 - (1 - z \leq 0,897)$

$0,8159 - (1 - 0,8159) = 0,63$

$\bar{X} \Rightarrow$ DEVO DIVIDERE TUTTO PER LA NUMEROSITÀ

c) Usando μ e var del p. 1 calcolare

$P(-10 \leq \underbrace{x_1 - \frac{1}{2}x_2 - \frac{1}{2}x_3}_{y} \leq 5)$

$\mu_y = 60 - \frac{1}{2}60 - \frac{1}{2}60 = 0$

$\sigma_y = \sqrt{15 + \frac{1}{4}15 + \frac{1}{4}15} = 4,74$

$\frac{-10-0}{4,74} \leq z \leq \frac{5-0}{4,74}$

$-2,11 \leq z \leq 1,055$

$P(z < 1,05) - 1 - P(z < 2,11)$

$0,8531 - (1 - 0,9826) = 0,836$

d) $\mu_1 = 40$ $\quad \mu_2 = 50$ $\quad \mu_3 = 60$

$\sigma_1^2 = 10$ $\quad \sigma_2^2 = 12$ $\quad \sigma_3^2 = 14$

$P(\underline{x_1 + x_2 \geq 2x_3})$

$\underbrace{x_1 + x_2 - 2x_3}_{v} \geq 0$

calcolare

$P(\underbrace{x_1 + x_2 + x_3}_{y} \leq 160)$

$\mu[y] = 40 + 50 + 60 = 150$

$\sigma_y = \sqrt{10+12+14} = 6$

$z \leq \frac{160-150}{6} = 1,67$

$z \leq 1,67 \Rightarrow 0,9525$

$\mu[v] = 40 + 50 - 120 = -30$

$\sigma_v = \sqrt{10 + 12 + 4 \cdot 14} = 8,83$

$z \geq \frac{-(-30)}{8,83} = 3,39$

$1 - z < 3,39 = 1 - 0,9997 = 0,0003$

ESERCIZIO 2 (F9)

Stimare con il metodo dei momenti il parametro θ di una Uniforme$(-\theta,\theta)$ attraverso il campione $\{x_1, x_2 \dots x_n\}$ da esso estratto

$$f_X(x) = \frac{1}{\theta+\theta} = \frac{1}{2\theta}$$

$$f_X(x,\theta) = \begin{cases} \frac{1}{b-a} & x \in (-\theta,\theta) \\ 0 & \text{altrove} \end{cases}$$

Calcola prima la media; se $\theta = 0$ calcola media di x^2 poi ricorda: $M_2' = \frac{1}{2}\sum_{i=1}^{n} x_i$

$$\mu_1 = E[x] = \frac{a+b}{2} = \frac{-\theta+\theta}{2} = 0$$

$$\mu_2 = E[x^2] = \frac{(b-a)^2}{12} = \frac{\theta^2}{3} \mapsto \theta = \sqrt{3\mu_2}$$

$$M_2' = \frac{1}{2}\sum_{i=1}^{n} x_i \mapsto \Theta_{mm} = \sqrt{3M_2} = \sqrt{\frac{3}{n}\sum x_i^2}$$

ESERCIZIO 3 (F9)

Dato il campione $\{x_1, x_2 \dots x_n\}$ estratto da una popolazione con funzione di densità

$$f(x,\theta) = \begin{cases} (3/2 + 4\theta)x^{4\theta+1/2} & 0 < x < 1 \\ 0 & \text{altrove} \end{cases}$$

a) stimatore di massima verosimiglianza per $\theta > -3/8$
b) stimatore con metodo dei momenti

$$\mathcal{L}(\theta, x_1 \dots x_n) = \prod_{i=1}^{n} 3/2 + 4\theta x_i^{4\theta+1/2}$$

$$\ln \mathcal{L}(\theta, x_1 \dots x_n) = \ln\left((3/2+4\theta)^n \prod_{i=1}^{n} x_i^{4\theta+1/2}\right) =$$

$$n\ln\left(\tfrac{3}{2}+4\theta\right) + \sum_{h=1}^{n} \ln x_i^{4\theta+1/2} = n\ln\left(\tfrac{3}{2}+4\theta\right) + \left(4\theta+\tfrac{1}{2}\right)\sum_{i=0}^{n} \ln x_i$$

$$\frac{d\ell}{d\theta} = n\frac{1}{(3/2+4\theta)}4 + 4\sum_{i=1}^{n}\ln x_i = \frac{4n}{3/2+4\theta} + 4\sum_{i=1}^{n}\ln x_i = 0$$

$$\frac{2n}{3+8\theta} + \sum_{i=1}^{n}\ln x_i = 0$$

$$2n + 3\sum_{i=1}^{n}\ln x_i + 8\theta\sum_{i=1}^{n}\ln x_i = 0 \qquad 3 + \theta 8 \neq 0 \quad \text{OK}$$

$$\Theta_{MV} = \frac{-2n - 3\sum_i^n \ln x_i}{8\sum_i^n \ln x} = \frac{\left(\frac{-2n}{\sum_i \ln x_i} + 3\right)\cancel{\sum \ln x_i}}{8\cancel{\sum \ln x_i}} = -\frac{1}{4\sum_i^n \ln x_i} + \frac{3}{8}$$

(raccolgo)

STIMATORE CORRETTO se:

$E[\Theta] = \theta$ → fondamentale è una semplice perché lo calcolo come voglio

↳ dove questo lo calcolo:

$E[\Theta] = Var[x] + (E[x])^2$

errore quadratico medio ↳

ESERCIZIO 5 (F9)

Dato un campione $\{X_1 \dots X_n\}$ estratto da una popolazione con funzione di densità

$$f_x(x,\theta) \begin{cases} \frac{1}{\sqrt{\theta}} \cdot e^{-x/\sqrt{\theta}} & x>0 \\ 0 & \text{altrove} \end{cases}$$

Determinare stimatori Θ_{MV} Θ_{MM}, sono corretti?

Θ_{MV}: $\mathcal{L}(\theta, x_1 \dots x_n) = \frac{1}{\sqrt{\theta}} e^{-x_1/\sqrt{\theta}} \cdot \frac{1}{\sqrt{\theta}} e^{-\frac{x_2}{\sqrt{\theta}}} \dots = \left(\frac{1}{\sqrt{\theta}}\right)^n e^{\sum x_i/\sqrt{\theta}}$

$= \left(\frac{1}{\sqrt{\theta}}\right)^n e^{\frac{1}{\sqrt{\theta}}\sum_{i=1}^n x_i}$

applico il logaritmo: $\ln\left(\frac{1}{\sqrt{\theta}}\right)^n + \ln e^{\frac{1}{\sqrt{\theta}}\sum x_i} = -\frac{n}{2}\ln\theta + \frac{1}{\sqrt{\theta}}\sum_{i=1}^n x_i =$

derivo rispetto a θ e pongo $=0$: $-\frac{n}{2\theta} + \frac{1}{2}\theta^{-3/2}\sum_i x_i = 0$

$$\frac{-n + \theta^{-1/2}\sum^n x_i}{2\theta} = 0 \qquad -n + \theta^{-1/2}\sum_i x_i = 0$$

$$n / \sum x_i = \theta^{-1/2} \mapsto \underbrace{\frac{\sum x_i}{n}}_{\bar{X}_n} = \theta^{1/2} \mapsto \Theta_{MV} = \bar{X}_n^2$$

Θ_{MM}

Mi accorgo che la f_x è una esponenziale con parametro $\frac{1}{\sqrt{\theta}}$ per cui applico la media x l'exp

$E[x] = \frac{1}{\lambda} = \sqrt{\theta} \Rightarrow \bar{X} \qquad \theta = \bar{X}_n^2$

$\sigma^2 = \frac{1}{\lambda^2} = (\sqrt{\theta})^2 = \frac{\theta}{n}$

$\frac{1}{\lambda} = (\sqrt{\theta})^2 = \theta$

Sono corretti se: $E(\Theta) = \theta$

$E([\Theta]) = E[(\bar{X}_n)^2] = Var[\bar{x}] + (E[x])^2 = \frac{\sigma_x^2}{n} + \mu_x^2 = \frac{\theta}{n} + \theta$

$= \frac{\theta + n\theta}{n} = \frac{\theta(n+1)}{n} \neq \theta$

↳ posso modificarlo x renderlo corretto

$$\Theta = \frac{n}{n+1} X_n^2$$

ESERCIZIO 6 (F9)

Dato un campione, estratto da popolazione con distribuzione uniforme $U(0,\theta)$

Θ_{MV} per θ è corretto? se no, come lo modifico

Θ_{MM} per θ è corretto? se no come lo modifico

Uniforme perciò → $f_x = \frac{1}{b-a}$

Max verosimiglianza → è il $\max\{X_1 \dots X_n\} \Rightarrow Y_n$

(*) pag dop

FOGLIO 10

ESERCIZIO 1 (F10)

STRUMENTO DI MISURA impiegato per DETERMINARE GRANDEZZA INCOGNITA μ FORNISCE $X = \mu + \varepsilon$ con $\varepsilon \sim N(0, \sigma^2)$
7 MISURE INDIPENDENTI CON I RISULTATI: 2.1, 2.2, 1.9, 1.8, 2.3, 2.2, 1.7

a) intervallo di fiducia al 95% per μ

$0,95 = 1 - \alpha \qquad \alpha = 0,05$

$$\left(\bar{X}_n - t_{n-1, 1-\alpha/2} \frac{\sqrt{S_n^2}}{\sqrt{n}}, \; \bar{X}_n + t_{n-1, 1-\alpha/2} \frac{\sqrt{S_n^2}}{\sqrt{n}} \right)$$

con $n = 7$

$$S_n^2 = \frac{1}{n-1} \sum_i (X_i - \bar{X}_n)^2 = \frac{1}{7-1} \cdot \left[(0,07)^2 + (0,17)^2 + (-0,13)^2 + (-0,23)^2 + (0,27)^2 + (0,17)^2 + (0,33)^2 \right] = 0,052$$

$$\frac{2,1 + 2,2 + 1,9 + 1,8 + 2,3 + 2,2 + 1,7}{7} = 2,03$$

$S_n^2 = 19,69$

$$2 \pm \overbrace{t_{6; 0,975}}^{2,447} \frac{\sqrt{0,052}}{\sqrt{n}} \mapsto (l_i, l_s) \quad (2 \pm 0,21) = (\overset{l_i}{1,78} - \overset{l_s}{2,21})$$

$2 \pm 2,11 < \begin{matrix} 6,1 \\ -2,1 \end{matrix}$

b) intervallo di fiducia per σ^2 al 95%

$$\left(\frac{(n-1) S_n^2}{\chi^2_{n-1, \underset{0,975}{1-\alpha/2}}}, \; \frac{(n-1) S_n^2}{\chi^2_{n-1, \underset{0,025}{\alpha/2}}} \right)$$

$$\left(\frac{6 \cdot 0,052}{\chi^2_{6, 0,975}}, \; \frac{6 \cdot 0,052}{\chi^2_{6, 0,025}} \right) =$$

$$\left(\frac{6 \cdot 0,052}{14,449}, \; \frac{6 \cdot 0,052}{1,237} \right) = (\overset{l_i}{0,02} \; ; \; \overset{l_s}{0,254})$$

ESERCIZIO 3 (F10)

In una ricerca condotta da un centro studi sull'utilizzo internet è stato rilevato che su 540 intervistati, 248 ricevono 15 SMS al gg. Stima della proporzione nella popolazione degli utilizzatori di internet con le caratteristiche sopra indicate e si determini intervallo di fiducia al 98%.

$P[\text{usare}] = \frac{248}{540} = 0,46$ $\qquad 1-\alpha = 0,98 \qquad \alpha = 0,02$

uso proporzione : $(l_i\ l_s) = \left(\hat{p} \pm z_{1-\alpha/2}\sqrt{\frac{\hat{p}(1-\hat{p})}{n}}\right)$

$(l_i\ l_s) = \left(0,46 \pm z_{0,99}\sqrt{\frac{0,46(0,56)}{540}}\right) = 0,46 \pm 2,33 \cdot 0,022 =$

$= 0,46 \pm 0,05$

$(l_i\ l_s) = (0,41,\ 0,51)$

ESERCIZIO 4 (F10)

Il 14 ottobre 1997 il NY Times riportò un sondaggio che indicava che il 52% della popolazione era soddisfatta di Clinton.
Fornì anche l'intervallo di fiducia al 95% relativo alla stima effettuata
$I = (48\%\ \ 56\%)$
n° persone intervistate?

$I = \frac{56-48}{2} = \frac{8}{2} = 4 \mapsto 0,04$

$0,04 = z_{0,975}\sqrt{\frac{0,52(0,48)}{n}}$ $\qquad 0,04 = 1,96\sqrt{\frac{0,2496}{n}}$

$0,02 = \sqrt{\frac{0,2496}{n}}$ $\qquad 4,16 \cdot 10^{-4} = \frac{0,2496}{n}$ $\qquad n = 599,28 \simeq 600$

ESERCIZIO 5 (F10)

In un'azienda addetta all'imbollaggio di una certa merce sono in uso 2 macchine diverse A e B. Ci si chiede se i tempi di imbollaggio sono diversi. Si osservano i tempi di 10 operazioni in secondi, con i risultati:

Descriptive statistics: A; B

Variable	Count	Mean	StDev	Variance
A	10	55	1,4	1,96
B	10	53	1,5	2,25

Per provenire dalla stessa popolazione devono essere uguali medie e varianze

Assumendo che i tempi di operazione dell'imbollaggio seguono normale si costruisca l'intervallo di fiducia per $\mu_A - \mu_B$ con livello di fiducia 95%. Si può affermare allo stesso livello che i 2 campioni provengono dalla stessa popolazione?

Per capire se le varianze possono ritenersi uguali o no

$$\frac{\sigma_2^2}{\sigma_1^2} = \left(\frac{S_2^2}{S_1^2} f_{n_1-1,\,n_2-1,\,\alpha/2}\ ,\ \frac{S_2^2}{S_1^2} f_{n_1-1,\,n_2-1,\,1-\alpha/2}\right)$$

$\hookrightarrow$ deve essere > 1

devo calcolare $(l_i\ l_s)\ \frac{\sigma_2^2}{\sigma_1^2}$ e vedere se 1 è all'intervallo, se sì le varianze sono uguali

FOGLIO 11

ESERCIZIO 1 (F11)

Nel testare l'Hp nulla $H_0: \mu = 10$ con test monolaterale e P errore di prima specie 5% (sapendo che la varianza della popolazione, che si suppone distribuita normalmente è = 4), si determini la numerosità campionaria necessaria per avere una P di errore di 2^a specie del 5% nel caso in cui sia vera l'Hp alternativa $H_A: \mu = 7$. Cosa cambia se test bilaterale?

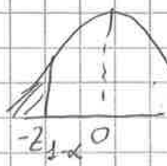
$\alpha = 0{,}05 \qquad \beta = 0{,}05$

$$\begin{cases} H_0: \mu = 10 \\ H_A: \mu = 7 \end{cases}$$

$7 < 10$ rifiuto in $(-\infty, 0)$ (monolaterale inferiore)

Monolaterale $1-\alpha$
Bilaterale $1-\alpha/2$

$P[\text{rifiutare } H_0 / H_0 \text{ vera}] = \alpha$



$\rightarrow \left[\frac{\bar{X}_{min} - \mu_0}{\sigma}\sqrt{n} < -z_{1-\alpha}\right] = 0{,}05 \mapsto \left[\frac{\bar{X}_{min} - 10}{2}\sqrt{n} < -1{,}65\right] = 0{,}05$

$\left[\bar{X}_{min} = -3{,}33 + 10\sqrt{n}\right] = 0{,}05$

$(P \text{ accettare } H_0 / H_A \text{ vera}) = \beta$

$P(\bar{X}_{min} / \mu = 7) = \beta$

$$P\left(z > \frac{10 - 1{,}645 \cdot \frac{2}{\sqrt{n}} - 7}{2/\sqrt{n}}\right) = 0{,}05$$

$P\left(z > 1{,}645 - 1{,}5\sqrt{n}\right) = 0{,}05$

$1 - P\left(z < 1{,}645 - 1{,}5\sqrt{n}\right) = 0{,}05 \rightarrow 0{,}95 = P\left(z < 1{,}645 - 1{,}5\sqrt{n}\right)$

$z_{0,95} = 1{,}645 - 1{,}5\sqrt{n} \qquad 1{,}65 + 1{,}645 = -1{,}5\sqrt{n}$

$n \simeq 5$

Se bilaterale:

$$\begin{cases} H_0 = \mu = 10 \\ H_A: \mu \neq 10 \end{cases}$$

rifiuto $(-\infty, 0) \cup (0 + \infty)$

ESERCIZIO 2 (F11)

Nel testare l'HP nulla $H_0: \mu = 11$ con test monolaterale superiore ed una P di errore di primo specie del 5% (stima della varianza su 20 elementi è 4) determinare n per avere $P\beta = 10\%$ nel caso sia vera $H_A: \mu = 12$

$\alpha = 0{,}05$

$$\begin{cases} H_0 & \mu = 11 \\ H_A & \mu > 11 \to \mu = 12 \end{cases}$$

Rifiuto se $z > (0, +\infty)$

$P[\text{rifiutare } H_0 / H_0 \text{ vero}]$

$$0{,}05 = \left[\frac{(X_{max} - \mu_0)}{\sqrt{S_n^2}}\sqrt{n} > T_{n-1, 1-\alpha}\right] \to X_{max} = \frac{1{,}729 \cdot 2}{\sqrt{n}} + 11\sqrt{n}$$

$$\to = \frac{3{,}458}{\sqrt{n}} + 11$$

se anche una + stessa quindi usare + bar come le stime

$$P(\text{accettare } H_0 / H_0 \text{ vero}) = \beta$$

$$P\left(11 + 3{,}458/\sqrt{n} \;/\; \mu = 7\right) = 0{,}05 \to P\left(t_{19} < \frac{11 + \frac{3{,}458}{\sqrt{n}} - 7}{2/\sqrt{n}}\right) = 0{,}1$$

$$\left(t_{19} < 1{,}729 + 2\sqrt{n}\right) = 0{,}1 \qquad t_{19, 0{,}1} = 1{,}729 + 2\sqrt{n}$$

$$\downarrow$$

$$-t_{19, 99} \to -1{,}328 = 1{,}729 + 2\sqrt{n}$$

$$\sqrt{n} = 3{,}037 \to n \approx 38$$

(1 + se non ho il valore lo prendo (-) il suo complementare

ESERCIZIO 3 F(11)

Campione con ampiezza 100 con $\bar{x} = 2{,}7$ e $\sum_{i=1}^{100}(x_i - \bar{x})^2 = 225$, sottoporre a test le $H_0: \mu = 3$ e $H_0: \sigma^2 = 2{,}5$ con livello di fiducia del 99%. (Popolazione normale). P di accettare erroneamente l'HP $H_0: \mu = 3$ quando è vero $\mu = 2{,}5$.

$\alpha = 0{,}01$

$$\begin{cases} H_0: & \mu_0 = 3 \\ H_A: & \mu_0 \neq 3 \end{cases} \quad \text{rigetto } (-\infty, 0)\ (0, +\infty)$$

$$S_n^2 = \frac{1}{n-1}\sum_i^n (x_i - \bar{X}_n)^2$$

$$S_n^2 = \frac{1}{99}\sum_i^{100}(x_i - \bar{X}_n)^2 = \frac{1}{99} \cdot 225 = 2{,}273$$

$$P[\text{rifiutare } H_0 / H_0 \text{ vero}] = 0{,}01$$

$$P\left[\frac{X_{max} - 3}{\sqrt{2{,}273}}\sqrt{100} > T_{n-1, 1-\alpha/2}\right] \cup \left[\frac{X_{min}}{\sqrt{2{,}273}} \cdot 100 < -T_{n-1, 1-\alpha/2}\right]$$

$99, 0{,}995$

$$X_{max} \approx (2{,}576 \cdot 1{,}5)/10 + 3 = 3{,}864 \qquad X_{min} = 2{,}61$$

ESERCIZIO 4. (F11)

Da una compagnia aerea è stata sviluppata una proposta di adesione ad un programma fedeltà - prevedendo che il 5% dei clienti abituali vi avrebbero aderito. Un campione casuale di 500 clienti ha fornito 34 adesioni.

a) Testare al livello di significatività 0,01 l'Hp: H_0 che la previsione sia corretta contro H_A: che non lo sia

$\mathbb{P}[\text{rifiutare } H_0 / H_0 \text{ vero}]$

$$\mathbb{P}_{reale} = \frac{34}{500} = 0{,}068$$

$$\begin{cases} H_0 & p = 0{,}05 \\ H_A & p \neq 0{,}05 \end{cases}$$

$$\mathbb{P}\left(\left(\frac{\hat{p}-p_0}{\sqrt{\frac{\hat{p}(1-\hat{p})}{n}}} > z_{1-\alpha/2}\right) \cup \left(\frac{\hat{p}-p_0}{\sqrt{\frac{\hat{p}(1-\hat{p})}{n}}} < -z_{1-\alpha/2}\right)\right)$$

$$\left(\frac{0{,}068-0{,}05}{\sqrt{\frac{0{,}068(1-0{,}068)}{500}}} > 2{,}58\right) \cup \left(\frac{0{,}068-0{,}05}{\sqrt{\frac{0{,}068(1-0{,}068)}{500}}} < -2{,}58\right)$$

$1{,}59 < -2{,}58 \quad \cup \quad 1{,}59 > 2{,}58 \longmapsto$ NON APPARTENGONO all'intervallo perciò non posso rifiutare

Hp che sia corretta (avendo non rifiuto H_A allora la faccio così)

Se trovo da testare delle Hp sulle probabilità uso $z_{calc} = \frac{\hat{p}-p_0}{\sqrt{\frac{\hat{p}(1-\hat{p})}{n}}}$ accetto se $|z_{calc}| < z$

b) Test precedente: qual è la $\mathbb{P}$ che la previsione sia corretta quando il 10% aderisce?

$$\begin{cases} H_0: & p = 0{,}05 \\ H_A & p = 0{,}1 \end{cases}$$

rifiuto $(0,+\infty)$ e $(-\infty,0)$

Ho da Ha quindi ricomincio e impostare come prima

$\mathbb{P}[\text{accetto } H_0 / H_A \text{ corretto}] \Rightarrow \hat{p} = \pm z_{1-\frac{\alpha}{2}} \sqrt{\frac{\hat{p}(1-\hat{p})}{n}} + p_0 \begin{cases} 0{,}021 \\ 0{,}079 \end{cases}$

$$\beta = \mathbb{P}(p_{MIN} < p < p_{MAX} / 0{,}1) = \mathbb{P}(0{,}021 < p < 0{,}079 / p = 0{,}1)$$

Calcolo prima la $\mathbb{P}$ che il 10% aderisca e poi β.

$$\frac{0{,}021-0{,}1}{\sqrt{\frac{0{,}1(0{,}9)}{500}}} < Z < \frac{0{,}075-0{,}1}{\sqrt{\frac{0{,}1 \cdot 0{,}9}{500}}}$$

$$-5{,}59 < Z < -1{,}86$$

$$1-(Z<1{,}86) - \underbrace{(1-(Z<5{,}59))}_{\simeq 0} = 1 - 0{,}5686 = 0{,}03$$

3%

ESERCIZIO 6 (F 11)

Due misuratori A e B di pressione su 16 persone hanno dato i risultati:

A) 79, 75, 93, 84, 93, 56, 87, 86
B) 66, 74, 82, 77, 91, 58, 75, 80

a) $\alpha = 0{,}05$, differenza significativa tra le medie?

$S_A = 12{,}02$

$n_A = 8$ $\quad \bar{X}_A = 81{,}625$ $\quad S_A^2 = 145{,}696$ → ricavato da $S_n^2 = \frac{1}{n-1} \sum_i (x_i - \bar{x}_n)^2$

$n_B = 8$ $\quad \bar{X}_B = 75{,}875$ $\quad S_B^2 = 102{,}125$ $\quad S_B = 10{,}106$

$\alpha = 0{,}05$ $\quad H_0: \mu_A = \mu_B$
$\quad H_A: \mu_A \neq \mu_B$

Ma faccio prima il TEST sulle varianze:

$\mathbb{P}[H_0 \text{ sia rifiut} \mid H_0 \text{ vera}]$

$$T = \frac{(\bar{X}_1 - \bar{X}_2) - d}{\sqrt{S^2_{pool}} \sqrt{\frac{1}{n_1} + \frac{1}{n_2}}} \mapsto$$

$\mu_A - \mu_B = 0$ per H_0

$$S^2_{pool} = \frac{(n_1 - 1) S_1^2 + (n_2 - 1) S_2^2}{n_1 + n_2 - 2}$$

$$\mapsto \frac{81{,}625 - 75{,}875 - 0}{\sqrt{123{,}91} \sqrt{1/4}} \; > t_{n_1+n_2-2;\, 1-\alpha/2} \text{ e } < -t_{n_1+n_2-2;\, 1-\alpha/2}$$

14 $\quad$ 0,975

$$\hookrightarrow \frac{7 \cdot 145{,}696 + 7 \cdot 102{,}125}{14} = 123{,}91$$

$1{,}033 > 2{,}145$
e
$1{,}033 < -2{,}145$

non appartiene perciò non posso rifiutare H_0

b) Calcolare errore di 2^ specie per le Hp alternative
$\mu_A - \mu_B = 3$
$\mu_A - \mu_B = 0$
$\mu_A - \mu_B = -3$

$$-1{,}033 < \frac{81{,}625 - 75{,}875 - 3}{5{,}56}$$

$-1{,}033 < 0{,}49$

(?)

B) calcolare intervallo di fiducia (livello 98%) per il contrasto $\Gamma = \mu_4 - \mu_3$

$\alpha = 0{,}02$

$$I(l_i, l_s) = \left(c \pm t_{n, 1-\frac{\alpha}{2}} \sqrt{\frac{MSERR}{n} \sum_{i=1}^{k} c_i^2} \right)$$

$1^2 + 1^2 = 2$

$$I(l_i, l_s) = \left(-1{,}51 \pm \underbrace{t_{8, 0{,}99}}_{2{,}896} \sqrt{\frac{0{,}1182}{3} \cdot 2} \right)$$

$$c \longrightarrow \sum c\,\bar{y} = 252{,}85 \cdot 1 - 254{,}36 \cdot 1 = -1{,}51$$

$$l_i\, l_s = (-2{,}323\, ,\, -0{,}697)$$

c) Si può dire a tale livello che le medie sono uguali?

No, non contengono lo 0 !!

Intervalli nella Anova

se conosco σ^2

$$\begin{matrix} l_i \\ l_s \end{matrix} \quad c \pm z_{1-\frac{\alpha}{2}} \sqrt{\frac{\sigma^2}{n} \sum c_i^2}$$

se non conosco σ^2 ma ne ho una stima

$$\begin{matrix} l_i \\ l_s \end{matrix} \quad c \pm t_{n, 1-\frac{\alpha}{2}} \sqrt{\frac{MSERR}{n} \sum c_i^2}$$

$\mapsto$ Al solito, si ritengono uguali le medie se l'intervallo contiene lo 0

FOGLIO 13

ESERCIZIO 1 (F13)

Regressione lineare semplice $Y = \beta_0 + \beta_1 X + \varepsilon$ con $\varepsilon \sim IIND(0, \sigma^2)$ e si suppone di osservare $n = 14$ coppie di dati (x_i, y_i), $i = 1 \dots n$

Sapendo che: $\sum_1 x_i = 890$ $\sum_1 x_i^2 = 67182$ $\sum_1 y_i = 37,6$

$\sum_1 y_i^2 = 103,54$ $\sum_1 x_i y_i = 2234,3$

a) calcolare b_0 e b_1

$$b_1 = \frac{\sum_1 x_i y_i - n\bar{x}\bar{y}}{\sum_1 x_i^2 - n\bar{x}^2} = \frac{2234,3 - 14 \cdot \frac{890}{14} \cdot \frac{37,6}{14}}{67182 - 14 \cdot \left(\frac{890}{14}\right)^2}$$

$= -0,0147$

$$b_0 = \bar{Y} - B_1 \bar{X} = \frac{37,6}{14} + 0,0147 \, \frac{890}{14} = 3,62$$

b) calcolare una stima di σ^2

$$S^2 = b_1^2 \left(\sum X_i^2 - n\bar{x}^2\right) = (-0,0147)^2 \cdot 67182 - 14\left(\frac{890}{14}\right)^2$$

$S^2 = 2,29 \simeq 0,02$

$b_1 = \dfrac{\sum x_i y_i - n\bar{x}\bar{y}}{\sum x_i^2 - n\bar{x}^2}$

$b_0 =$ formulario

$S^2 = b_1^2 \left(\sum X_i^2 - n\bar{x}^2\right)$

ESERCIZIO 2 (F13)

Si considera regressione lineare semplice $Y = \beta_0 + \beta_1 x + \varepsilon$ con $\varepsilon \sim IIND(0, \sigma^2)$ e si suppone di osservare $n = 18$ coppie di dati (x_i, y_i) $i = 1 \dots n$

Dato che:

$b_0 = 27,18$ $b_1 = -0,2976$ $\sum_1 x_i = 659$ $\sum (x_i - \bar{x})^2 = 4840,7778$

$s = 2,8640$

a) calcolare intervallo di fiducia x il valor medio previsto corrispondente a $x = 45$ $(1-\alpha = 95\%)$

$$(l_i \; l_s) = b_0 + b_1 x_k \pm t_{n-2, 1-\frac{\alpha}{2}} \sqrt{var[\hat{Y}_k]}$$

$$\sigma^2 \left[\frac{1}{n} + \frac{(x_k - \bar{x})^2}{\sum (x_i - \bar{x})^2}\right]$$

$$= 27,18 - (0,2976 \cdot 45) \pm t_{16, 0,975} \sqrt{2,864^2 \left(\frac{1}{18} + \frac{\left(45 - \frac{659}{18}\right)^2}{4840,7778}\right)}$$

$= (12,18 - 15,40)$

$(l_i \; l_s) = b_0 + b_1 x_k \pm t_{n-2, 1-\frac{\alpha}{2}} \sqrt{var[\hat{y}]}$

dove: $var\,\hat{y} = \sigma^2 \left[\frac{1}{n} + \frac{(x_k' - \bar{x})^2}{\sum (x_i - \bar{x})^2}\right]$

ESERCIZIO 4 (F13)

Nella Tabella (vedi foglio) sono riportati i dati di peso, lunghezza e tipologie di trazione di 54 veicoli. In fondo alla tabella sono disponibili alcune statistiche.
Calcolare la matrice di correlazione, valutando la significatività dei valori ottenuti

Matrice di correlazione

$$\begin{pmatrix} 1 & r_{12} & r_{13} \\ r_{21} & 1 & \\ r_{31} & & 1 \end{pmatrix}$$

li calcolo con la formula della correlazione

nello specifico sono

$$\begin{pmatrix} 1 & r_{wp} \\ r_{ew} & 1 \end{pmatrix}$$

uguali

$$r_{we} = \frac{\sum x_{ie} x_{iw} - n\bar{x}_e \bar{x}_w}{\sqrt{\left(\sum x_{ie}^2 - \bar{x}_e^2 n\right)\left(\sum x_{iw}^2 - \bar{x}_w^2 n\right)}}$$

MATRICE DI CORRELAZIONE

$$\begin{pmatrix} 1 & r_{12} & r_{13} \dots \\ r_{21} & 1 & 1 \\ r_{31} & \dots & \end{pmatrix}$$

Li calcolo con il coeff di correlazione

$$= \frac{35906345 - 54 \cdot \frac{9740}{54} \cdot \frac{197487}{54}}{\sqrt{\left(1764276 - \left(\frac{9740}{54}\right)^2 54\right)\left(740397563 - \left(\frac{197487}{54}\right)^2 54\right)}} = 0{,}7755$$

significatività valori: $t_{calc} = \frac{r}{\sqrt{1-r^2}}\sqrt{n-2}$

$$t_{calc} = \frac{0{,}7755}{\sqrt{1-0{,}7755^2}}\sqrt{54-2} = 8{,}86$$

$$t_{n-2,\,1-\frac{\alpha}{2}} \Rightarrow t_{52;\,0{,}975}$$

$\to$ 30 approx Pvalue a 0

Livello significatività
$t_{calc} = \frac{r}{\sqrt{1-r^2}}\sqrt{n-2}$
Pvalue $\mapsto$ se $n>30$ $\simeq 0$

Intervallo di fiducia per la previsione per la singola osservazione

$$(l_i, l_s) = \hat{y} \pm t_{n-2,\, 1-\alpha/2} \sqrt{\sigma^2 \left(1 + \frac{1}{n}\left(\frac{(x_* - \bar{x})^2}{\Sigma(x_i - \bar{x})^2}\right)\right)}$$

$$l_i l_s = 8{,}135 \pm 2{,}131 \sqrt{0{,}68\left(1 + \frac{1}{17}\left(\frac{(4{,}5-9)^2}{408}\right)\right)}$$

STESSE FORMULE DI PRIMA

$$= 8{,}135 \pm 2{,}131 \cdot 0{,}8258 \longmapsto (6{,}28\,,\; 9{,}9)$$

ESERCIZIO 6 (F13)

Output con regressione ottenuta con $n = 25$

The regression equation is

Diretto = 7,7 + 1,42 ona − 0,290 Opact

MSE 14,6304

Matrix XPXI1

0,188690	−0,025000	−0,017857
−0,025000	0,007143	−0,000000
−0,017857	−0,000000	0,003968

$$\begin{pmatrix} \mathrm{var}\,B_0 & \mathrm{cov}(B_0 B_1) & \mathrm{cov}\,B_0B_2 \\ & \mathrm{var}\,B_1 & \\ & & \mathrm{var}\,B_2 \end{pmatrix}$$

a) calcola le stime delle covarianze tra i 3 stimatori dei parametri del modello di regressione stimato. Testare l'Hp nulla che i 3 stimatori sono mutuamente non correlati al livello di significatività 5%.

Per ottenere le stime delle covarianze devo moltiplicare i valori nei posti delle covarianze per σ^2

$$\mathrm{cov}(B_0, B_1) = \mathrm{cov}(B_1 B_0) = -0{,}025 \cdot 14{,}6304 = -0{,}36576$$

$$\mathrm{cov}(B_0, B_2) = \mathrm{cov}(B_2 B_0) = -0{,}017857 \cdot 14{,}6304 = -0{,}261255$$

$$\mathrm{cov}(B_1 B_3) = \mathrm{cov}(B_3 B_1) = 0$$

3° test con $B_1 B_2$

$$r_{B_1 B_2} = \frac{cov[B_1 B_2]}{\sqrt{var[B_1]\ var[B_2]}} = 0$$

$$t_{calc} = 0$$

$|0| < 2{,}069 \mapsto$ NON Rifiuto $\mapsto$ accetto H_0

Stima covarianza: quella nella matrice $\cdot \sigma^2$

TEST D'HP:

$$r_{B_\alpha B_\beta} = \frac{cov[B_\alpha \beta_\beta] \cdot \cancel{\sigma^2}}{\sqrt{var[\beta_\alpha] var[B_\beta]} \cdot \cancel{\sigma^2}}$$

t_{calc}

$\hookrightarrow t_{n-2,\, 1-\alpha/2}$

se è vera accetto H_0

$|t_{calc}| < t_{n-2, 1-\alpha/2}$

ESERCIZIO 7 (F13)

Predictor	Coef	SE Coef	T	P
Constant	-48,972	7,152	-6,85	0,000
X_1	0,3861	0,1302	2,96	0,006
X_2	2,0711	0,2791	7,42	0,000
X_3	-0,01424	0,01219	-1,17	0,253
X_4	0,144976	0,007405	19,25	0,000

Residual standard Error: 3,05929 on 27 Degrees of Freedom

a) intervallo di fiducia al 98% per β_2

$$(l_i\ l_s) = \underbrace{b_0 + b_1 x_k}_{\hat{y} \to coef} \pm t_{DF,\, 1-\frac{\alpha}{2}} \sqrt{Var[B_2]}$$

$\sqrt{Var[\beta]} = SE\ coef$

$$2{,}0711 \pm t_{27,\, 0{,}99}\ 0{,}2791$$

$$= (2{,}7613\ ,\ 1{,}3808)$$

ESERCIZIO 9 (F13)

Allo scopo di costruire un modello lineare polinomiale che esplichi il peso in kg in funzione dell'altezza in cm, è stato estratto un campione di 20 persone, con età tra i 30 e i 40 di sesso maschile.

h: 177 176 173 174 176 165 172 172 177 176 180 175 178
183 191 174 172 172 181 193

p: 78 66 82 76 60 61 72 68 73 80 70 77 78
81 72 78 74 63 82 110

+ VEDI OUTPUT

a) sulla base dei dati stimare i parametri del modello

$$b_1 = \frac{\sum x_i y_i - n\bar{x}\bar{y}}{\sum x_i^2 - n\bar{x}^2} = \frac{\overset{\text{sum}}{266280} - 20 \cdot \overset{\text{sum p}}{\frac{1501}{20}} \cdot \overset{\text{sum h}}{\frac{3537}{20}}}{626317 - 20 \cdot \left(\frac{3537}{20}\right)^2} = 1{,}037$$

↑ SUM OF squares X_i^2

$$b_1 = \frac{\sum x_i y_i - n\bar{x}\bar{y}}{\sum x_i^2 - n\bar{x}^2}$$

$x_i y_i$ = sum
n = gradi
$\bar{y}$ = sum p/n
$\bar{x}$ = sum h/n
$\bar{x}_i^2$ = SUM OF squares X_i

perché nell'output ci dice che peso è la y

$$b_0 = \bar{y} - B_1\bar{x} = \frac{1501}{20} - 1{,}037 \cdot \frac{3537}{20} = -108{,}3$$

b_0 = FORMULARIO!!

b) Testare H_0 che $\beta_1 = 1$ contro $H_a: \beta_1 \neq 1$

$H_0 \Rightarrow \beta_1 = 1$
$H_A \Rightarrow \beta_1 \neq 1$

$$t_{calc} = \frac{b_1 - 1}{\sqrt{var[B_1]}} = \frac{1{,}037 - 1}{}$$

Probabilità:

→ se sono indipendenti: $\mathbb{P}[A \cap B] = \mathbb{P}[A] \cdot \mathbb{P}[B]$

(almeno uno) $\mathbb{P}[A \cup B] = \mathbb{P}[A] + \mathbb{P}[B] - \mathbb{P}[A \cap B]$

(sommo dispari, sottraggo pari)

→ $\mathbb{P}[B_3 / B_1 \cap B_2] = \frac{\mathbb{P}(B_1 \cap B_2 \cap B_3)}{\mathbb{P}(B_1 \cap B_2)}$ $\quad \mathbb{P}[A/B] = \mathbb{P}[A \cap B] / \mathbb{P}[B] = \mathbb{P}[A]$ se vero che sono indip.

→ se sono mutuamente esclusivi: $\mathbb{P}[A \cap B] = \emptyset$

$\mathbb{P}[A \cup B] = \mathbb{P}[A] + \mathbb{P}[B] - \cancel{\mathbb{P}[A \cap B]}$ $\emptyset$

($\mathbb{P}$ totali) ↦ $\mathbb{P}[A\bar{B}] : \mapsto \mathbb{P}[A] = \mathbb{P}[AB] + \mathbb{P}[A\bar{B}]$

→ se NON sono esclusivi

→ $\mathbb{P}[\text{almeno uno}] = \mathbb{P}[A] + \mathbb{P}[B] + \mathbb{P}[C] - \mathbb{P}[AB] - \mathbb{P}[AC] - \mathbb{P}[BC] + \mathbb{P}[ABC]$

→ $\mathbb{P}[\text{nex}] = 1 - \mathbb{P}[\text{almeno } 1]$

→ $\mathbb{P}[\text{uno solo}] = \mathbb{P}[A] + \mathbb{P}[B] + \mathbb{P}[C] - 2\mathbb{P}[AB] - 2\mathbb{P}[BC] - 2\mathbb{P}[AC] + 3\mathbb{P}[ABC]$

se indipendenti

$(C_1 \cap C_2 \cap C_3) \cup (C_1 \cap C_4) = C_1 \cdot C_2 \cdot C_3 + C_1 \cdot C_4 - C_1 \cdot C_2 \cdot C_3 \cdot C_4$

↦ lanciando 3 dadi possibili combo 6^3

↦ quando un tot di lettere possibili combo n°lettere!

↦ se ho una probabilità totale ↦ no Bayes

© Proprietà riservata dell'autore - Digitalizzazione e distribuzione a cura del CENTRO APPUNTI - Corso Luigi Einaudi, 55 - Torino / Pagina 65 di 74

Trasformazioni di variabili casuali

- Se ho x discreta con valori $\{a, b \dots\}$ e $f(a)$: $f(b)$ ... per calcolare la $Y =$ ___ sostituisco le x che ho, per calcolare poi $f(Y)$ ⇝ prendo i valori che ho ottenuto della y una sola volta e vedo a quali valori di f_x corrispondono. vedi es 1 foglio 6

- Trasformazioni: mi da: $\circ$ $f(x)$ o distribuzione

$\underset{g(x)}{Y} = \dots$ ↳ estraggo la $X = \underset{g^{-1}(y)}{\dots}$

verifico dominio x e y

~ se x non biunivoca spezzo

~ se y sale di più de una parte — 2° intervallo, 1° intervallo

a questo punto faccio:

$f_Y(y) =$ Formulario

↳ se la $x = \pm \dots$ devo fare separatamente la parte + e la −

↳ se D_y è spezzato devo farlo per entrambi i domini

⇨ Poi faccio $f_Y(y)^{TOT}$

Medie:

- $\int_{-\infty}^{+\infty} x \, f_x \, dx$ ← dominio x
- $\int_{-\infty}^{+\infty} y \, f_y \, dy$ ← dominio y
- $\mathbb{E}[g[x]] = \int g(x) f(x) dx$ ← dominio x
- $\mathrm{var}[\mathbb{E}(x)] = \mathbb{E}[(x - \mathbb{E}(x))^2]$

↳ se $g(x) = (x - \mu_x)^2$ ⇝ $\mathbb{E}[g(x)] = \mathrm{var}[x]$

STIMA PUNTUALE

⇨ METODO DEI MOMENTI:

↝ Si calcola la media, con una delle varie formule; se ∄ $E(x) = 0$ si fa: $E(x^2) = \dots$ ecc.

una volta ottenuta la $\bar{x}$ ↝ si ricava la Θ^{MM}

↑ ricorda che è uguale a: $\frac{1}{n}\sum x_i$

⇨ METODO DI MAX VEROSIMIGLIANZA:

→ Si scrive $\mathcal{L}(x_1 \dots x_n, \Theta) = $ (usando $\sum$ e $\prod$) ← sommatorie ← produttorie

→ applico il logaritmo

→ derivo rispetto a Θ ...

→ estraggo Θ^{MV}

⇩

STIMATORI CORRETTI SE

$E(\Theta^{MV} \text{ o } \Theta^{MM}) = \Theta$

↑ calcolo la media delle funzioni ottenute

⇨ se ho: $E[x^2] = \text{var}[x] + E[x]^2$ (Tabelle)

INTERVALLI DI FIDUCIA

↦ livello di fiducia ⇒ $1-\alpha$

↦ Ampiezza: $\ell_s - \ell_i = I$

↝ in base a quello che ho sfrutto il formulario

↝ se mi da "%" sfrutto quelle per le "proporzioni"

↝ se do numero di misurazioni mi servono anche per calcolarmi $\bar{X}$

⇨ STIMATORE CORRETTO PER LA MEDIA se:

$E(T_n) = \mu$

NOTA:

$\sum_i^n i = \frac{n(n+1)}{2}$ $\quad \sum_i i^2 = \frac{n(n+1)(2n+1)}{6}$ $\quad \sum_i (-1)^i <^{-1 \text{ dispari}}_{0 \text{ pari}}$

NOTA 2:

$\sum^n E[x_i] = n\mu$

$\sum i E[x_i] = \mu \sum i$

↦ idem per le varianze solo che nel caso $\text{var} \sum i x_i = \sum i^2 \sigma^2$ formula

⇨ STIMATORE PIÙ EFFICIENTE: quello con la varianza minore

⇨ STIMATORE CONSISTENTE: $\lim \underbrace{\text{var}[T_n] + [E[T_n] - \mu]^2}_{= MSE(T_n)} = 0$

© Proprietà riservata dell'autore - Digitalizzazione e distribuzione a cura del CENTRO APPUNTI - Corso Luigi Einaudi, 55 - Torino / Pagina 69 di 74

Pure errore

X_i	Y
1	10, 10, 6
2	14, 12, 10
3	⋮
4	
5	

$$\frac{10+10+6}{3} = 9 \qquad (10-9)^2+(10-9)^2+(6-9)^2 = 6$$

$\dots \quad 0$

Sommi tutti i valori: Pure errore

- Impara formule
- Leggi e impara roba lab
- rileggi foglio 13

© Proprietà riservata dell'autore - Digitalizzazione e distribuzione a cura del CENTRO APPUNTI - Corso Luigi Einaudi, 55 - Torino / Pagina 71 di 74

Stimatore della risposta SINGOLA PREVISIONE

$$(l_i\ l_s) = \hat{Y}_0 \pm t_{n-2,1-\alpha/2}\left(\sqrt{\mathrm{var}[Y_k]}\right)$$

$\hat{Y}_0 = B_0 + B_1 x_0$ (x_0 valore dato)

$$\mathrm{var}[Y_k] = \sigma^2\left(1+\frac{1}{n}+\frac{(x_0-\bar{x})^2}{\sum(x_i-\bar{x})^2}\right)$$

↦ se lo chiede per β_2 o β_1 calcolo usando le formule adatte e comp β_2 = COEFF X_2, β_1 = COEFF X_1 ....

Valore: $S^2_{x_n} = \dfrac{X^2_{nn} - n x^2_{n}}{n}$

$= \left((X_n^2) - n(\bar{x}_n)^2\right)\frac{1}{n}$ (coriognale / non diogonale)

correlazione: $\rho_{B_0B_1} = \dfrac{\mathrm{cov}[B_0B_1]}{\sqrt{\sigma^2_{B_0}\sigma^2_{B_1}}}$

$$r_{x_n x_m} = \frac{\mathrm{cov}[X_n X_m]}{\sqrt{\frac{S x_n^2}{n}\frac{S x_m^2}{n}}} = \frac{\sum(x_n x_m - n\bar{x}_n\bar{x}_m)}{\sqrt{\sum(x_{in}^2 - n\bar{x}_n^2)(\sum x_{im}^2 - n\bar{x}_m^2)}}$$

valore nella n+1 riga m+1 colonna

Matrix:

$$\begin{pmatrix} 1 & r_{12} & r_{13} \\ r_{21} & 1 & r_{23} \\ r_{31} & r_{32} & 1 \end{pmatrix}$$

TEST D' HP

↦ STIMATORI (TESTORE) mutuamente non correlati calcolo:

$$\rho_{B_0B_1} \rightsquigarrow T_{calc} = \frac{\rho}{\sqrt{1-\rho^2}}\sqrt{n-2}$$

voluto se T_{calc} appartiene o meno a:

$$-T_{n-2,1-\alpha/2} < T_{calc} < T_{n-2,1-\alpha/2}$$

OUTPUT:
- MSE 1 = σ^2
- P = pvalue
- T = T_{calc}
- S = $\hat{\sigma}$ ⇝ elevo alla seconda e ottengo $\hat{\sigma}^2$
- Se fit = $\sqrt{\mathrm{var}[Y_k]}$
- $SE_{B_2} = \sqrt{\widehat{\mathrm{var}}[B_2]}$
- $SE_{B_1} = \sqrt{\widehat{\mathrm{var}}[B_1]}$

→ per i singoli parametri guardo il T dell'output e lo confronto con $\pm T_{n-2,1-\alpha/2}$

↦ → testare Hp: $\beta_1 = \gamma$ contro HA: $\beta_1 \neq \gamma$

$$t_{calc} = \frac{B_1 - \gamma}{\sqrt{\mathrm{var}[B_1]}}$$ e confronto con $\pm t_{n-2,1-\alpha/2}$

→ Pericolo di II specie

$$-(t_{n-2,1-\alpha/2}) = \frac{b_{MIN} - \gamma_{H_0}}{\sqrt{\mathrm{var}[B_{1,\sigma^2}]}}$$ ricavo b_{MIN}

© Proprietà riservata dell'autore - Digitalizzazione e distribuzione a cura del CENTRO APPUNTI - Corso Luigi Einaudi, 55 - Torino / Pagina 73 di 74