

PROVA SCRITTA DI STATISTICA (COD 4038)

27 Marzo 2003

MODALITÀ A

APPROSSIMARE TUTTI I CALCOLI ALLA QUARTA CIFRA DECIMALE

ESERCIZIO 1 (8 punti)

Il quotidiano "Repubblica" ha pubblicato i 10 libri più venduti nella settimana (dal 3 al 9 Febbraio).

Nella seguente tabella, per ciascun libro è indicato:

TITOLO

AUTORE

EDITORE

PREZZO

PAGINE

numero pagine

GEN

Genere del libro (C=comico, G=giallo, R=romanzo, T=thriller)

TITOLO	Autore	Editore	Prezzo	Pagine	Gen	Prezzo ²	Pagine ²
Il volo del calabrone	Ken Follett	Mondadori	19.00	444	G	361	197136
Io uccido	G. Faletti	Baldini	17.20	681	T	295.84	463761
Famiglia	Mario Puzo	Sonzogno	17.50	393	R	306.25	154449
La principessa sul pisello	L. Littizzetto	Mondadori	13.00	240	C	169	57600
Buick 8	Stephen King	Sperling & K.	18.00	476	R	324	226576
La città delle bestie	I. Allende	Feltrinelli	14.00	224	R	196	50176
L'amore dura tre anni	F. Beigbeder	Feltrinelli	8.00	160	R	64	25600
Non ti muovere	M. Mazzantini	Mondadori	16.50	304	R	272.25	92416
Harry Potter e il prigioniero di Azkaban	J.K. Rowling	Salani	15.00	376	R	225	141376
Il sogno più dolce	Doris Lesing	Feltrinelli	18.00	456	R	324	207936
SOMMA			156.2	3754		2537.34	1617026

- 1) Fornire un'opportuna rappresentazione grafica del carattere GEN (1 punti)
- 2) Calcolare un'opportuna misura di posizione per descrivere il carattere GEN (1 punti)
- 3) Considerando solo i libri editi da Feltrinelli e Mondadori, possiamo affermare che il PREZZO è regressivamente indipendente dall'EDITORE? Giustificare la risposta (2 punti)
- 4) Quale tra i due caratteri PREZZO e PAGINE presenta maggiore variabilità? (2 punti)
- 5) Prevedere il prezzo di un nuovo libro di 500 pagine che uscirà tra breve, sulla base della relazione lineare tra i caratteri PREZZO e PAGINE presenti nella precedente tabella. (la covarianza tra PREZZO e PAGINE risulta 335.722) (2 punti)

ESERCIZIO 2 (3 punti)

Siano X e Y due variabili statistiche, rispondere con VERO o FALSO alle seguenti affermazioni:

- a) Se X è regressivamente indipendente da Y, allora Y è regressivamente indipendente da X.
- b) $\rho(X,Y)=0$ implica $f^{\sim}(X,Y) < 1$
- c) $f^{\sim 2}(X,Y)=0$ implica $\rho(X,Y)=0$

ESERCIZIO 3 (4 punti)

Sia X una variabile aleatoria con funzione di ripartizione $F(x)$ data dalla seguente espressione:

$$F(x) = \begin{cases} 0 & x < 2 \\ \frac{x-2}{3} & 2 \leq x < 5 \\ 1 & x \geq 5 \end{cases}$$

- Calcolare la funzione di densità di probabilità della variabile aleatoria X . **(2 punti)**
- Calcolare il valore atteso, $E(X)$ **(2 punti)**

ESERCIZIO 4 (4 punti)

Sia X una variabile aleatoria con $E(X)=\mu+3$ e $V(X)=1$.

Dato un campione Bernoulliano di ampiezza n :

- proporre uno stimatore non distorto per μ ; **(2 punti)**
- calcolare la varianza dello stimatore trovato al punto precedente. **(2 punti)**

ESERCIZIO 5 (8 punti)

Il direttore dell'Ospedale XYZ, situato in un quartiere molto povero nei sobborghi di New York, sospetta che i neonati che nascono lì abbiano un peso inferiore rispetto alla media nazionale (pari a 3.2 Kg), tale da dover richiedere un intervento di prevenzione sulla malnutrizione delle donne del quartiere.

Possiamo supporre che la variabile aleatoria X ="peso alla nascita" si distribuisca (nel quartiere) come una normale di valore atteso non noto μ e varianza non nota σ^2 . Il direttore decide allora di verificare l'ipotesi $H_0: \mu=3.2$ in alternativa all'ipotesi $H_1: \mu < 3.2$, misura quindi il peso $(x_1, x_2, \dots, x_{81})$ di 81 bambini scelti casualmente tra i neonati nell'ultimo anno e riscontra che

$$\sum_{i=1}^{81} x_i = 234.9 \quad \sum_{i=1}^{81} x_i^2 = 1001.21$$

- Proporre un opportuno stimatore non distorto per il parametro σ^2 e fornirne una stima. **(2 punti)**
- Calcolare l'intervallo di confidenza a livello $1-\alpha=0.9$ per il parametro μ . **(2 punti)**
- Scrivere l'espressione analitica del p-value corrispondente alla realizzazione campionaria osservata e calcolarne il valore. Decidere inoltre se il direttore riterrà necessario un intervento di sanità pubblica contro la malnutrizione, avendo fissato α , il livello di significatività del test uguale a 0.06 (motivare la risposta). **(4 punti)**