Probabilità e Statistica

9 settembre 2013

AVVERTENZE:

1. La prova dura 2 ore.
2. E' ammesso il solo utilizzo delle tavole presenti nel sito del corso.
3. Alla fine della prova si dovranno consegnare SOLO i fogli con il testo del compito e le soluzioni riportate in modo sintetico negli appositi spazi. NON si accetteranno fogli di brutta copia.
4. Il compito è considerato insufficiente se vi sono meno di 6 risposte esatte ai quesiti a risposta multipla.

COGNOME .... NOME .... MATRICOLA ....

Quesiti a risposta multipla

1. Se la distribuzione di una variabile è simmetrica allora necessariamente

   A lo scarto quadratico medio è pari a 1

   B la media aritmetica è pari a 0

   ~~C~~ $Q_2 = (Q_1 + Q_3)/2$, dove Q_i indica l'i-esimo quartile

2. La funzione di ripartizione di una v.c. continua a valori in $(-\infty, \infty)$

   A può assumere qualsiasi valore in $(-\infty, \infty)$

   ~~B~~ può assumere qualsiasi valore in $[0, 1]$

   C può non essere definita

3. Quale delle seguenti relazioni è vera?

   A $A \cup B = A^c \cap B^c$

   ~~B~~ $(A \cup B)^c = A^c \cap B^c$

   C $A \cap B = A^c \cup B^c$

4. Quale delle seguenti espressioni è sempre zero?

   ~~A~~ $\sum_{i=1}^n y_i/n - \overline{y}$

   B $n \sum_{i=1}^n y_i - \overline{y}$

   C $\sum_{i=1}^n y_i - \overline{y}$

5. Siano $x_1, \ldots, x_n$ e $z_1, \ldots, z_n$ due insiemi di dati tali che $z_i = a + b \cdot x_i$, $i = 1, \ldots, n$. Allora

   A $\text{var}(z_1, \ldots, z_n) = b \cdot \text{var}(x_1, \ldots, x_n)$

   ~~B~~ $\text{var}(z_1, \ldots, z_n) = b^2 \cdot \text{var}(x_1, \ldots, x_n)$

   C $\text{var}(z_1, \ldots, z_n) = a + b^2 \cdot \text{var}(x_1, \ldots, x_n)$

6. Se due v.c. $X \sim$ Bernoulli(0.1) e $Y \sim$ Bernoulli(0.2) sono indipendenti, allora:

   A $P(X=1) = P(Y=1)$
   B $P(X = 1 \cap Y = 0) = 0.08$
   C $P(Y = 1|X = 0) = 0.8$

7. Si associ al comando `pexp(0.5, 2)` il corrispondente risultato:

   A 0.6321
   B -0.5
   C 0.3679

8. Quale comando si utilizza in R per calcolare $P(X > 2.3)$ se $X \sim Poisson(\lambda = 1)$?

   A `dpois(2.3, 1)`
   B `1-ppois(2.3, 1)`
   C `ppois(1, 2.3)`

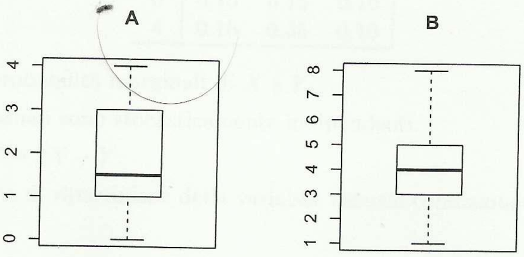
9. Se $Y_1, \ldots, Y_n$ sono variabili casuali indipendenti, tutte di media μ e varianza σ^2 e n è sufficientemente grande, allora

   A $\sum_i Y_i$ ha distribuzione $\mathcal{N}(n\mu, \sigma^2)$
   B $\overline{Y}$ ha distribuzione $\mathcal{N}(\mu, \sigma^2/n)$
   C $\overline{Y}$ ha distribuzione approssimata $\mathcal{N}(\mu, \sigma^2/n)$

10. Se $X \sim \mathcal{N}(2, 4)$ e $Y \sim \mathcal{N}(3, 1)$ sono variabili casuali indipendenti, allora $Z = X/2 - 3Y$ è tale che

    A $Z \sim \mathcal{N}(-8, 10)$
    B $Z \sim \mathcal{N}(8, 10)$
    C $X \sim \mathcal{N}(8, 8)$

1. I grafici qui sotto riportati si riferiscono ad un'indagine su un campione di 100 dipendenti, 50 dell'azienda A (sinistra) e 50 dell'azienda B (destra). A ciascun dipendente è stato chiesto di indicare il numero di ore di lavoro straordinario effettuate nel mese di dicembre.



(a) Determinare un indice di posizione per le due distribuzioni.

(b) Determinare un indice di variabilità per le due distribuzioni.

(c) Dare un giudizio sulla simmetria della distribuzione per ciascuna delle due aziende.

(d) In quale azienda si fanno più straordinari?

a) INDICE POSIZIONE – A: $X_{0,25} = 1$, $X_{0,5}$ = MEDIANA = 1,5, $X_{0,75} = 3$

– B: $X_{0,25} = 3$, $X_{0,5}$ = MEDIANA = 4, $X_{0,75} = 5$

b) INDICE DI VARIABILITÀ: – A: CAMPO DI VARIAZIONE = 4 − 0 = 4

SCARTO INTERQUARTILE = 3 − 1 = 2

– B: CAMPO DI VARIAZIONE = 8 − 1 = 7

SCARTO INTERQUARTILE = 5 − 3 = 2

c) LA PRIMA È PIÙ RISCOSTATA DELLA SECONDA.

LA SECONDA È SIMMETRICA TRA $X_{0,25}$ E $X_{0,75}$.

d) NELLA SECONDA.

2. Sia data la seguente funzione di probabilità congiunta per le variabili casuali discrete X e Y:

	Y		
X	0	1	2
0	0.15	0.15	0.10
4	0.15	0.35	0.10

(a) Calcolare le funzioni di probabilità marginali di X e Y.

(b) Dire se le due variabili casuali sono stocasticamente indipendenti.

(c) Calcolare $\text{Var}(Z)$, dove $Z = 2X - Y$.

(d) Calcolare media e funzione di ripartizione della variabile casuale condizionata $X|Y = 2$.

(e) Calcolare $\Pr(Y \cdot X > 0)$.

$X \backslash Y$	0	1	2	$P_X(x)$
0	0,15	0,15	0,1	0,4
4	0,15	0,35	0,1	0,6
$P_Y(y)$	0,3	0,5	0,2	1

$X \perp\!\!\!\perp Y \iff \forall_{x,y}\ f(x,y) = f_X(x) \cdot f_Y(y)$

$\Rightarrow X \not\perp\!\!\!\perp Y \Rightarrow 0,3 \cdot 0,4 \neq 0,15$

X	$P_X(x)$
0	0,4
4	0,6

Y	$P_Y(y)$
0	0,3
1	0,5
2	0,2

$E(x) = 4 \cdot 0,6 = 2,4$

$E(y) = 0,5 + 0,2 \cdot 2 = 0,9$

$E(x^2) = 4^2 \cdot 0,6 = 9,6$

$E(y^2) = 0,5 + 0,2 \cdot 4 = 1,3$

$VAR(x) = E(x^2) - E(x)^2 = 9,6 - 2,4^2 = 3,48$

$VAR(y) = E(y^2) - E(y)^2 = 1,3 - 0,9^2 = 0,49$

$VAR(2x - y) = 4VAR(x) + VAR(y) - COV(x,y)$

$\overset{!}{=} 4 \cdot 3,48 + 0,49 - E(2x \cdot y) + E(x) \cdot E(y)$

$\overset{!}{=} 4 \cdot 3,48 + 0,49 + 2E(xy) + 2E(x) \cdot E(y) =$

$\overset{!}{=} 4 \cdot 3,48 + 0,49 + 2(4 \cdot 0,35 + 8 \cdot 0,1) + 2 \cdot 0,9 \cdot 2,4 =$

$\overset{!}{=}$

x	$P_{X\|Y}(x\|y=2)$
0	$0,1/0,2 = 1/2$
4	$0,1/0,2 = 1/2$

$E(x) = 2$

$$F(x) = \begin{cases} 0 & x < 0 \\ 1/2 & 0 \leq x \leq 4 \\ 1 & x > 4 \end{cases}$$

$P(x \cdot y > 0) = P(4,1) + P(4,2) = 0,35 + 0,1 = 0,45$

3. Siano $X_1, \dots, X_n$ variabili casuali indipendenti e identicamente distribuite con distribuzione uniforme nell'intervallo $(1, 2)$. Utilizzando il teorema limite centrale, approssimare $P\{\overline{X}_n < 1.45\}$ per $n = 100$, dove $\overline{X}_n = \sum_{i=1}^{n} X_i / n$.

$$\overline{X}_m \underset{\text{APPROSSIMATO}}{\sim} N\left(\underset{\substack{\shortparallel \\ 1{,}5}}{\frac{a+b}{2}} , \underset{\substack{\shortparallel \\ 1/12}}{\frac{(a-b)^2}{12}} \Big/ m \right)$$

$$P(\overline{X}_m < 1{,}45) = \Phi\left(\frac{1{,}45 - 1{,}5}{\frac{1}{10\sqrt{12}}} \right) = \Phi(-1{,}44) = 1 - \Phi(1{,}44) = 1 - 0{,}92507 = 0{,}07493$$