Università C. Cattaneo, Corso di Laurea in Economia Aziendale, A.A. 2018-2019
STATISTICA — 30.10.18 - I PROVA PARZIALE
Modalità A

(A) ai fini della valutazione verranno considerate **solo** le risposte riportate dallo studente **negli appositi riquadri** bianchi.
(B) nello svolgimento del compito si utilizzino almeno tre cifre decimali.

COGNOME.............................NOME...........................MATR...................

ESERCIZIO 1 (punti 6) Si considerino due variabili aleatorie X e Y con distribuzione congiunta riportata nella seguente tabella a doppia entrata:

X \ Y	0	1	3	
-1	0.1	0.15	0.35	0.6
1	0	0.2	0	0.2
2	0.2	0	0	0.2
	0.3	0.35	0.35	

a) Si determini la funzione di probabilità della variabile aleatoria $T=4-X^2$.
b) Si dica, giustificando la risposta, se X e Y sono indipendenti.
c) Si calcoli $P(X<Y)$.
d) Si determini la varianza di $W=Y-3X+1$.

a) $T = 4 - X^2$ $\quad X \rightarrow$

x_i	$p_X(x_i)$
-1	0.6
1	0.2
2	0.2

$S_T = \{0, 3\}$ $\quad p_T(t) = \begin{cases} 0.2 & t=0 \\ 0.8 & t=3 \\ 0 & \text{altrove} \end{cases}$

b) No $\quad 0.1 \neq 0.6 \cdot 0.3$ (ad esempio)

c) $P(X<Y) = P(X-Y<0) = P(X=-1, Y=0) + P(X=-1, Y=1) + P(X=-1, Y=3) + P(X=1, Y=3) + P(X=2, Y=3) =$
$= \boxed{0.6}$

d) $Var(W) = Var(Y) + 9 Var(X) - 6 cov(X,Y) =$
$= 1.54 + 9 \cdot 1.6 - 6 \cdot (-1) = 1.54 + 14.4 + 6 = \boxed{21.94}$

$Var(X) = [0.6 + 0.2 + 0.8] - 0 = 1.6$ $\qquad cov(X,Y) = [-0.15 - 1.05 + 0.2] =$
$Var(Y) = [0.35 + 3.15] - 1.96 = 1.54$ $\qquad - 0 = -1$

ESERCIZIO 2 (punti 4). Il numero di email classificate come SPAM che arrivano in 24 ore ad una casella di posta elettronica ha distribuzione di Poisson di media 6.
a) Si calcoli la probabilità che in 12 ore arrivino più di 3 SPAM.
b) Si calcoli il numero atteso di SPAM che arrivano in 8 ore.
c) Si calcoli lo scarto quadratico medio (deviazione standard) del numero di SPAM che arrivano in 8 ore.

$E(X) = -0.6 + 0.2 + 0.4 = 0$
$E(Y) = 0.35 + 1.05 = 1.4$

X = num email in 12 ore Y = num. email in 8 ore

a) $X \sim Poiss(3)$ $P(X>3) = 1 - P(X \leq 3) = 1 - 0.647 = \boxed{0.353}$

b) $Y \sim Poiss(2)$ $E(Y) = 2$

c) $\sigma(Y) = \sqrt{Var(Y)} = \sqrt{2} \doteq \boxed{1.4142}$

ESERCIZIO 3 (punti 4). Nell'ultimo anno il 20% delle aziende di un Paese ha subito trasformazioni societarie.

a) Si considerino 5 aziende del Paese. Si calcoli la probabilità che non più di 3 di esse abbiano subito trasformazioni societarie.

b) Si considerino 90 aziende. Si calcoli la probabilità che almeno un terzo di esse abbiano subito trasformazioni societarie.

a) $X \sim Bin(5, 0.2)$ [num. aziende tra le 5 che hanno subito trasformaz.]

$$P(X \leq 3) = 1 - P(X=4) - P(X=5) =$$

$$= 1 - \frac{5!}{4!\,1!}(0.2)^4(0.8) - \frac{5!}{5!}(0.2)^5 = 1 - 5\cdot(0.0016)(0.8) -$$

$$- 0.00032 = 1 - 0.0064 - 0.00032 = \boxed{0.9933}$$

TEO. CENTR. DEL LIMITE

b) $Y \sim Bin(90, 0.2)$ e $Y \approx N(90\cdot 0.2, 90\cdot 0.2\cdot 0.8)$ $= N(18, 14.4)$

[num. aziende tra le 90 che hanno subito trasform.]

$$P(Y \geq 30) = P\left(Z \geq \frac{30-18}{\sqrt{14.4}}\right) = P(Z \geq 3.16) = 1 - 0.9992 =$$

$$= 0.0008$$

oppure (equivalentemente)

$$P(\hat{P} \geq 0.33) = P\left(Z \geq \frac{0.33 - 0.2}{\sqrt{0.0018}}\right) = P(Z \geq 3.18) = 0.0008$$

$$\hat{P} \approx N\left(0.2, \frac{0.2\cdot 0.8}{90}\right) = N(0.2, 0.0018)$$

ESERCIZIO 4 (punti 8). Il rendimento annuo di un portafogli è determinato per il 40% da un titolo A, per il 50% da un titolo B e per il resto da un titolo C. Il rendimento di A ha distribuzione normale con media 0.02 e scarto quadratico medio 0.01, il rendimento di B ha media 0.04 e scarto quadratico medio 0.02, il rendimento di C ha media 0.03 e scarto quadratico medio 0.01. Inoltre, i rendimenti di A e B hanno coefficiente di correlazione lineare 0.09, i rendimenti di A e C sono indipendenti ed i rendimenti di B e C sono indipendenti.

a) Si calcoli la probabilità che il rendimento di A sia compreso tra 0 e 0.03.

b) Si determini il quantile di ordine 0.18 del rendimento di A.

c) Si determini il rendimento atteso del portafoglio ~~portafoglio~~

d) Si determini la varianza del rendimento del portafoglio ~~portafoglio~~

e) Si dica, giustificando la risposta, se i rendimenti di A e B sono indipendenti.

$X \sim N(0.02, 0.01^2)$ $\quad \rho(X,Y) = 0.08$

$Y \sim \quad 0.04 / 0.02^2$ $\quad cov(Y,W) = 0$

$W \sim \quad 0.03 / 0.01^2$ $\quad cov(X,W) = 0$

a) $P(X \in (0, 0.03)) = P\left(\frac{0-0.02}{0.01} < Z < \frac{0.03-0.02}{0.01}\right) =$

$= P(-2 < Z < 1) = 0.8413 - 1 + 0.9772 =$

$= \boxed{0.8185}$

b) $q_{0.18}(X) = k$

$P(X \leq k) = 0.18$

$P\left(Z \leq \frac{k-0.02}{0.01}\right) = 0.18$

$P\left(Z \leq -\frac{k-0.02}{0.01}\right) = 0.82$

$-\frac{k-0.02}{0.01} = 0.92 \quad \frac{k-0.02}{0.01} = -0.92$

$k - 0.02 = -0.0092 \quad k = 0.02 - 0.0092 =$

$= 0.0108$

c) $T = 0.4X + 0.5Y + 0.1W$

$E(T) = (0.4)(0.02) + (0.5)(0.04) + (0.1)(0.03) = 0.031$

d) $Var(T) = (0.16)(0.01)^2 + (0.25)(0.02)^2 + (0.01)(0.01)^2 + 2\cdot(0.08)(0.01)\cdot$
$(0.02)(0.4)(0.5) = 0.00012$

e) No (se fossero indipend., $\rho(X,Y)$ sarebbe uguale a 0)

ESERCIZIO 5 (punti 2) Il tempo necessario per produrre un'unità di un prodotto è in media pari a 13 minuti, con uno scarto quadratico medio pari a 1 minuto. Si consideri [redacted] 100 unità di questo prodotto. Si calcoli la probabilità che il tempo totale necessario per produrre questi 100 prodotti sia maggiore di 1290 minuti. [redacted] "Si considerino 100 unità ...", etc.

$X \quad E(X) = 13 \quad \sigma(X) = 1 \quad X_1 — X_{100} \quad T = \sum X_i$

$T \approx N(1300, 100) \quad P(T > 1290) = P\left(Z > \frac{1290-1300}{10}\right) = P(Z > -1) =$

$= \boxed{0.8413}$

$cov(X,Y) = (0.08)\cdot(0.01)(0.02) =$

ESERCIZIO 6 (punti 3) Il Data Base delle "Paghe e Contributi" relativo ai 1000 dipendenti dell'azienda AZ S.p.a. contiene, tra le altre, la colonna, o variabile X, "numero di giorni di assenza per malattia" e la colonna, o variabile Y, "numero di ore di straordinario" per ciascun dipendente. Per fini di analisi gestionale si sono considerati i valori di tali variabili in una data settimana ottenendo la seguente tabella a doppia entrata:

X\Y	0	6	10
0	40,00%		
3	5,00%	25,00%	
5		5,00%	15,00%
8			10,00%

a) Si specifichi l'unità statistica (o unità osservativa) cui si riferiscono i valori delle variabili X e Y registrati nel Data Base.
b) Si specifichi la moda della variabile Y, e quanti sono i dipendenti con un numero di ore di straordinario pari a tale moda.
c) Si calcoli il numero medio di giorni di assenza per malattia dei dipendenti dell'azienda.
d) Si specifichi la frequenza relativa dei dipendenti con numero di giorni di assenza per malattia compreso fra 3 e 5 (inclusi) e con numero di ore di straordinario compreso fra 6 e 10 (inclusi).

$n = 1000$

a) DIPENDENTE = UNITÀ STATISTICA

b) moda$(Y) = 0$ 45% di 1000 $\rightarrow$ 450
(modalità con frequenza più elevata)

c) media$(X) = 0 \cdot 0.4 + 3 \cdot 0.3 + 5 \cdot 0.2 + 8 \cdot 0.1 = 2.7$

d) $\underbrace{Fr(3 \leq X \leq 5, 6 \leq Y \leq 10)}_{\text{frequenza relativa cercata (congiunta)}} = 0.45 \; (0.25 + 0.05 + 0.15)$