Università C. Cattaneo, Corso di Laurea in Economia Aziendale, A.A. 2016-2017
STATISTICA — 15.06.17 - PROVA GENERALE
FREQUENTANTE (SVOLTO ASSIGNMENT): SI' □ NO □

(A) ai fini della valutazione verranno considerate **solo** le risposte riportate dallo studente **negli appositi riquadri** bianchi: in caso di necessità **si può anche andare fuori dai margini che delimitano i riquadri**.
(B) **nello svolgimento del compito si utilizzino tre cifre decimali.**

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

ESERCIZIO 1 (punti 7). Una variabile X ha distribuzione Binomiale con parametri 6 e 0.4; una variabile Y ha distribuzione di Poisson con media 2. X e Y sono indipendenti.
a) Si calcoli la probabilità che X sia maggiore di 3.
b) Si calcoli la probabilità che Y sia minore di 2.
c) Si calcoli la probabilità che S=X+Y sia uguale a 0.
d) Si calcoli il valore atteso di T=2-2X-Y.
e) Si calcoli lo scarto quadratico medio di T.
f) Si calcoli la probabilità che, congiuntamente, X sia minore di 2 e Y sia maggiore di 3.
g) Quanto vale il coefficiente di correlazione lineare tra X e Y?

$X \sim Bin(6, 0.4)$; $Y \sim Poisson(2)$; X e Y indipendenti

a) $P(X>3) = P(X=4)+P(X=5)+P(X=6) = \frac{6!}{4!\,2!}(0.4)^4\cdot(0.6)^2 + \frac{6!}{5!\,1!}(0.4)^5\cdot(0.6) + \frac{6!}{6!\,0!}(0.4)^6 = 0.138+0.037+0.004 = \boxed{0.179}$

b) $P(Y<2) = P(Y \leq 1) = \boxed{0.406}$

c) $P(S=0) = P(X+Y=0) = P(X=0, Y=0) = P(X=0)\cdot P(Y=0) = \left[\frac{6!}{0!\,6!}(0.4)^0(0.6)^6\right]\cdot 0.135 = (0.047)(0.135) = \boxed{0.006}$

d) $E(T) = E(2-2X-Y) = 2-2E(X)-E(Y) = 2-2\cdot(6\cdot 0.4)-2 = \boxed{-4.8}$

e) $Var(T) = 4\,Var(X)+Var(Y) = 4\cdot(6\cdot 0.4\cdot 0.6)+2 = 7.76$
$\sigma(T) = \sqrt{Var(T)} = \sqrt{7.76} = \boxed{2.786}$

f) $P(X<2, Y>3) = P(X<2)\cdot P(Y>3) = \left\{\frac{6!}{1!\,5!}(0.4)^1\cdot(0.6)^5 + \frac{6!}{0!\,6!}(0.4)^0(0.6)^6\right\}\cdot\left\{1-P(Y\leq 3)\right\} = (0.233)\cdot(0.143) = \boxed{0.033}$

g) $\rho(X,Y) = 0$ in quanto X e Y sono indipend.

Università C. Cattaneo, Corso di Laurea in Economia Aziendale, A.A. 2016-2017

ESERCIZIO 2 (punti 7.5). Il rendimento di un portafoglio finanziario è determinato per il 30% da un titolo A e per il 70% da un titolo B. Il rendimento mensile di A ha distribuzione gaussiana con media 0.03 e scarto quadratico medio 0.02. Il rendimento mensile di B ha media 0.04 e scarto quadratico medio 0.03. Il coefficiente di correlazione lineare tra i due rendimenti è 0.5.

a) Si calcoli la probabilità che il rendimento di A sia positivo.
b) Si determini il quantile di ordine 0.75 del rendimento di A.
c) Si determini il valore atteso del rendimento del portafoglio.
d) Si determini la varianza del rendimento del portafoglio.
e) Assumendo che il rendimento del portafoglio abbia distribuzione gaussiana, si calcoli la probabilità che esso sia compreso tra 0.02 e 0.04.

R = rendimento portaf. ; $R = (0.3)\cdot X + (0.7) Y$, dove $X \sim N(0.03, 0.02^2)$ e Y ha $E(Y) = 0.04$, $\sigma(Y) = 0.03$, $\rho(X,Y) = 0.5$ [X è il rend. di A, Y quello di B].

a) $P(X>0) = P\left(Z > \frac{0-0.03}{0.02}\right) = P(Z>-1.5) = \boxed{0.9332}$

b) $P(X \leq q_{0.75}) = 0.75$, cioè $P\left(Z \leq \frac{q_{0.75}-0.03}{0.02}\right) = 0.75$, da cui $\frac{q_{0.75}-0.03}{0.02} = 0.67$, ovvero $q_{0.75} = \boxed{0.043}$

c) $E(R) = (0.3)(0.03) + (0.7)\cdot(0.04) = \boxed{0.037}$

d) $Var(R) = (0.3)^2\cdot(0.02)^2 + (0.7)^2\cdot(0.03)^2 + 2\cdot(0.3)(0.7)(0.5)(0.02)(0.03) =$
$= \boxed{0.0006}$

e) $R \sim N(0.037, 0.0006)$ $\quad \sigma_R = \sqrt{0.0006} = 0.024$

$P(0.02 < R < 0.04) = P\left(\frac{0.02-0.037}{0.024} < Z < \frac{0.04-0.037}{0.024}\right) =$
$= P(-0.708 < Z < 0.125) = P(Z<0.13) - [1-P(Z<0.71)] =$
$= 0.5517 - [1-0.7611] = \boxed{0.3128}$

Università C. Cattaneo, Corso di Laurea in Economia Aziendale, A.A. 2016-2017

ESERCIZIO 3 (punti 6) Si vuole stimare la spesa media (pro-capite) sostenuta dai clienti di un ristorante per la cena. Si estrae un campione di 120 clienti; la spesa media dei 120 clienti è stata pari a 32 euro, lo scarto quadratico medio delle spese dei 120 clienti è stato pari a 4 euro.
a) Si determini un intervallo di confidenza di livello 0.95 per la spesa media per la cena nel ristorante in esame.
b) L'intervallo al 90%, basato sugli stessi dati campionari, risulta più o meno lungo del precedente? Si giustifichi la risposta.
c) Si vuole verificare se la spesa media per la cena è inferiore a 35 euro. Si proceda alla verifica calcolando il p-value e traendo la conclusione (a livello 0.05) sulla base di questo.

$X_1, \dots, X_{120}$ campione da X = spesa cena
$\bar{x} = 32$, $S = 4$, $n = 120$. $\mu = E(X)$

a) $\left(\bar{x} - 1.96 \cdot \frac{S}{\sqrt{n}}, \bar{x} + 1.96 \cdot \frac{S}{\sqrt{n}}\right) = (32 \pm 0.716) = (31.284, 32.716)$

b) L'intervallo al 90% risulta più corto del precedente: se $1-\alpha$ diminuisce, la lunghezza diminuisce (ovvero, se aumenta la confidenza, l'intervallo è meno preciso).

c) $\begin{cases} H_0: \mu = 35 \\ H_1: \mu < 35 \end{cases}$ ~~sotto H_0~~ $Z = \frac{\bar{X} - 35}{S/\sqrt{n}}$ p-value $= P(Z \leq z_{oss}) =$

$= P\left(Z \leq \frac{32-35}{4/\sqrt{120}}\right) = P(Z \leq -8.216) \simeq 0$

Poiché p-value $< \alpha = 0.05$, si rifiuta H_0 = c'è evidenza che la spesa è inferiore a 35.

ESERCIZIO 4 (punti 5). Attraverso un'indagine basata su un campione di 120 clienti di un ristorante si vuole stimare la proporzione di quelli che si ritengono soddisfatti della cena. Dei 120 clienti del campione, 72 si dicono soddisfatti.
a) Si determini una stima della proporzione di clienti soddisfatti della cena.
b) Si determini lo standard error (ovvero deviazione standard, o scarto quadratico medio) della stima calcolata al punto a).
c) Si verifichi, a livello 0.01, se la proporzione di clienti soddisfatti della cena è diversa da 0.65.

$X_1, \dots, X_{120}$ campione da $X \sim Bern(p)$
$\bar{x} = \frac{72}{120} = 0.6$

a) $\bar{x} = 0.6$ è la stima di p

b) $\sqrt{\frac{\bar{x}(1-\bar{x})}{n}} = \sqrt{\frac{0.6 \cdot 0.4}{120}} = 0.045$

c) $\begin{cases} H_0: p = 0.65 \\ H_1: p \neq 0.65 \end{cases}$ $R_{0.01} = |Z| = \left|\frac{\bar{X} - 0.65}{\sqrt{\frac{0.65(1-0.65)}{120}}}\right| \geq 2.576$

$|z_{oss}| = \left|\frac{0.6 - 0.65}{\sqrt{\frac{0.65 \cdot 0.35}{120}}}\right| = 1.148 \not\geq 2.576$ quindi non si rifiuta H_0

Non c'è evidenza che la prop. di clienti soddisfatti sia diversa da 0.65

ESERCIZIO 5 (1.5 punti)

Per i valori x della variabile X (riportati nella prima colonna a sinistra nella tabella qui sotto) ed i valori y della variabile Y (riportati nella prima riga in alto della stessa tabella) si sono rilevate le frequenze assolute congiunte riportate nella stessa tabella.

X \ Y	4	5	6	7
0	20	5		5
1				5
2	5			10
3	5		15	30

Mostrando i calcoli principali, si risponda alle seguenti domande:
a) si determini la mediana della variabile Y.
b) si determini la frequenza relativa delle coppie (x, y) di valori delle due variabili con x compreso fra 2 e 3 (inclusi) e con y compreso fra 5 e 6 (inclusi).
c) si determinino media e varianza della variabile Y.
d) si determini la covarianza delle variabili X e Y.

X/ Y	4	5	6	7	
0	0,2	0,05		0,05	0,3
1				0,05	0,05
2	0,05			0,1	0,15
3	0,05		0,15	0,3	0,5
	0,3	0,05	0,15	0,5	1

Domanda (a)

cumulate					
Y	0,3	0,35	0,5	-->	**mediana Y = 6**

Domanda (b)

C'è una sola coppia (x,y)=(3,6) che verifica le condizioni, dunque
la frequenza relativa richiesta è **0,15**

Domanda (c)

y	f(y)	y*f(y)	y^2	(y^2)*f(y)
4	0,3	1,2	16	4,8
5	0,05	0,25	25	1,25
6	0,15	0,9	36	5,4
7	0,5	3,5	49	24,5
	1	**5,85**		35,95
		media Y		

quadrato di media Y 34,2225

varianza di Y = 35,95 meno 34,2225 **1,7275**

Domanda (d)

E' necessario calcolare la media di X.

Con le prime tre colonne come in (c), la media di X è 1,85

Inoltre serve il prodotto delle due medie cioè 10,8225

Cov(X,Y)= totale ultima colonna qui sotto meno tale prodotto= **0,9275**

x	y	f(x,y)	x*y*f(x,y)
1	7	0,05	0,35
2	4	0,05	0,4
2	7	0,1	1,4
3	4	0,05	0,6
3	6	0,15	2,7
3	7	0,3	6,3
			11,75

ESERCIZIO 6 (punti 3) **[SOLO PER STUDENTI NON FREQUENTANTI, CHE NON HANNO SVOLTO CIOE' L'ASSIGNMENT NELL'ANNO 2016-2017]**

Si calcoli la probabilità che la media campionaria relativa a un campione di dimensione 100 da una distribuzione con valore atteso 1 e varianza 25 sia negativa. Sulla base di quali proprietà/teoremi viene calcolata la probabilità richiesta?

$X_1, \ldots, X_{100}$ i.i.d con $E(X_i)=1$
e $Var(X_i) = 25$.

$$\bar{X} = \frac{1}{n}\sum X_i = \frac{1}{100}\sum_{i=1}^{100} X_i$$

$$\bar{X} \approx N(1, 25/100) = N(1, 0.25)$$

$$P(\bar{X} < 0) = P\left(Z < \frac{0-1}{\sqrt{0.25}}\right) = P(Z < -2) =$$

$$= 1 - 0.9772 = \boxed{0.0228}$$

Per calcolare la probabilità è stato usato il teorema centrale del limite, per cui $\bar{X}$ ha distribuzione approssimat. normale se n è suffic. grande.