

Università C. Cattaneo, Corso di Laurea in Economia Aziendale, A.A. 2014-2015
STATISTICA—04.11.14 - I PROVA PARZIALE
SOLUZIONI Modalità B

(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 **non meno di due cifre decimali**.

ESERCIZIO 1 (punti 8). Un portafoglio è costituito da due titoli, A e B; il suo rendimento è determinato per il 40% dal rendimento di A e per il resto da quello di B. Il titolo A ha rendimento distribuito secondo una gaussiana (o normale) di media 0.11 e varianza 0.04; il rendimento di B ha distribuzione gaussiana con media 0.05 e varianza 0.09. I due rendimenti hanno coefficiente di correlazione lineare pari a 0.4.

- a) Si calcoli la probabilità che il rendimento di B sia negativo.
 b) Si determini il III quartile (ovvero, 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) Supponendo che anche il rendimento del portafoglio abbia distribuzione gaussiana, si calcoli la probabilità che esso sia compreso tra 0.05 e 0.1.

a) rendimento titolo B = X gaussiana $N(0,05;0,09)$, e dunque si ha

$$P(X < 0) = P\left(Z < \frac{0-0,05}{0,3}\right) = P\left(Z < -\frac{0,05}{0,3}\right) = P(Z < -0,166) = 1 - P(Z < 0,17) = 1 - 0,5675 = 0,4325$$

b) rendimento titolo A = Y gaussiana $N(0,11;0,04)$, e dunque si ha

$$y_{0,75} = z_{0,75}\sigma_Y + \mu_Y = z_{0,75} \cdot 0,2 + 0,11 = 0,68 \cdot 0,2 + 0,11 = 0,246$$

dove $P(Z < z_{0,75}) = 0,75 \Rightarrow z_{0,75} = 0,68$, o per difetto $z_{0,75} = 0,67$ ed allora: $y_{0,75} = 0,244$

c) rendimento portafoglio $R = 0,4Y + 0,6X \Rightarrow E(R) = 0,4E(Y) + 0,6E(X) = 0,4 \cdot 0,11 + 0,6 \cdot 0,05 = 0,074$

$$\begin{aligned} d) V(R) &= 0,4^2 V(Y) + 0,6^2 V(X) + 2 \cdot 0,4 \cdot 0,6 \cdot Cov(X, Y) = \\ &= 0,4^2 \cdot 0,04 + 0,6^2 \cdot 0,09 + 0,48 \cdot \sigma_X \cdot \sigma_Y \cdot \rho(X, Y) = \\ &= 0,0388 + 0,48 \cdot 0,3 \cdot 0,2 \cdot 0,4 = 0,0388 + 0,48 \cdot 0,024 = 0,05032 \end{aligned}$$

e) rendimento portafoglio R = gaussiana $N(0,074;0,05032)$, $\sigma_R = \sqrt{0,05032} = 0,22432$

$$\begin{aligned} P(0,05 < R < 0,1) &= P\left(\frac{0,05-0,074}{0,22432} < Z < \frac{0,1-0,074}{0,22432}\right) = P(-0,10699 < Z < 0,11591) = \\ &= P(Z < 0,11591) - [1 - P(Z < 0,10699)] = P(Z < 0,11591) + P(Z < 0,10699) - 1 = \\ &= P(Z < 0,12) + P(Z < 0,11) - 1 = 0,5478 + 0,5438 - 1 = 1,0916 - 1 = 0,0916 \end{aligned}$$

ESERCIZIO 2 (punti 3). Il numero mensile di richieste di assenza per malattia arrivano all'ufficio personale di un'azienda in accordo ad una distribuzione di Poisson. Il numero medio di richieste mensili è 4.

- a) Si calcoli la probabilità che in un mese arrivino più di 2 richieste.
 b) Si calcoli la probabilità che in un mese arrivino esattamente 3 richieste.
 c) Si calcoli lo scarto quadratico medio (o deviazione standard) del numero di richieste mensili.

a) X e' Poissoniana con media 4 e varianza 4.

$$P(X > 2) = 1 - P(X \leq 2) = 1 - 0,238 = 0,762$$

$$b) P(X = 3) = P(X \leq 3) - P(X \leq 2) = 0,433 - 0,238 = 0,195$$

$$c) \sigma_X = \sqrt{\sigma_X^2} = \sqrt{4} = 2$$

ESERCIZIO 3(punti 3).In un supermercato, la probabilità che un generico cliente acquisti un determinato prodotto A è pari a 0.3. Si considerano 6 clienti.

- a) Si calcoli la probabilità che nessuno dei 6 clienti acquisti il prodotto A.
 b) Si calcoli la probabilità che almeno 3 dei 6 clienti acquistino il prodotto A.
 c) Si determini il numero medio di clienti che acquistano il prodotto A, tra i 6 considerati.

$$a) p = 0,3, n = 6, \text{ numero dei clienti che acquistano X e' } Bi(0,3;6)$$

$$P(X=0) = p_X(0) = \binom{6}{0} 0,3^0 \cdot 0,7^{6-0} = 0,7^6 = 0,11765$$

$$\begin{aligned} \text{b) } P(X \geq 3) &= 1 - P(X \leq 2) = 1 - (p_X(0) + p_X(1) + p_X(2)) = \\ &= 1 - (0,11765 + p_X(1) + p_X(2)) = 1 - 0,74431 = 0,25569 \end{aligned}$$

dove

$$P(X=1) = p_X(1) = \binom{6}{1} 0,3^1 \cdot 0,7^{6-1} = 6 \cdot 0,3 \cdot 0,7^5 = 0,30253$$

$$P(X=2) = p_X(2) = \binom{6}{2} 0,3^2 \cdot 0,7^{6-2} = 15 \cdot 0,3^2 \cdot 0,7^4 = 0,32414$$

$$\text{c) } E(X) = np = 6 \cdot 0,3 = 1,8$$

ESERCIZIO 4 (punti3) Il peso (in grammi) di ciascuno di 250 blocchi di calcestruzzo prodotti da un impianto è una variabile aleatoria con media 4000 e varianza 940. Si assuma che tali variabili aleatorie siano I.I.D., ovvero formino un campione.

- a) Si determini il valore atteso della media dei pesi dei 200 blocchi.
 b) Si determini la varianza della media dei pesi dei 200 blocchi.
 c) Si calcoli la probabilità che la media dei pesi dei 200 blocchi sia maggiore di 3995.

SOLUZIONI CON 250 BLOCCHI

a) Il peso di ciascun blocco è una variabile aleatoria $X_i, i=1, 2, \dots, 250$. Tali variabili sono I.I.D. con $\mu_i = \mu = 4000$ e $\sigma_i^2 = \sigma^2 = 940$. Il peso medio dei 250 blocchi è la variabile aleatoria media campionaria \bar{X}_{250} che ha valore medio atteso $\mu = 4000$.

Infatti per la definizione di media campionaria e per le proprietà delle combinazioni lineari si ha

$$E(\bar{X}_{250}) = E\left(\sum_{i=1}^{250} \frac{1}{250} X_i\right) = \sum_{i=1}^{250} \frac{1}{250} E(X_i) = 250 \cdot \left(\frac{1}{250} 4000\right) = 4000$$

b) Per la definizione di media campionaria, per le proprietà delle combinazioni lineari e poiché le variabili sono stocasticamente indipendenti si ha

$$V(\bar{X}_{250}) = V\left(\sum_{i=1}^{250} \frac{1}{250} X_i\right) = \sum_{i=1}^{250} \left(\frac{1}{250}\right)^2 V(X_i) = 250 \cdot \left(\frac{1}{250^2} 940\right) = \frac{940}{250} = 3,76$$

$$\begin{aligned} \text{c) } P(\bar{X}_{250} > 3995) &= P\left(Z > \frac{3995 - 4000}{\sqrt{3,76}}\right) = P\left(Z > -\frac{5}{1,93907}\right) = 1 - P(Z < -2,57855) = \\ &= 1 - [1 - P(Z < 2,58)] = P(Z < 2,58) = 0,9951 \end{aligned}$$

SOLUZIONI CON 200 BLOCCHI

a) Il peso di ciascun blocco è una variabile aleatoria $X_i, i=1, 2, \dots, 200$. Tali variabili sono I.I.D. con $\mu_i = \mu = 4000$ e $\sigma_i^2 = \sigma^2 = 940$. Il peso medio dei 200 blocchi è la variabile aleatoria media campionaria \bar{X}_{200} che ha valore medio atteso $\mu = 4000$.

Infatti per la definizione di media campionaria e per le proprietà delle combinazioni lineari si ha

$$E(\bar{X}_{200}) = E\left(\sum_{i=1}^{200} \frac{1}{200} X_i\right) = \sum_{i=1}^{200} \frac{1}{200} E(X_i) = 200 \cdot \left(\frac{1}{200} 4000\right) = 4000$$

b) Per la definizione di media campionaria, per le proprietà delle combinazioni lineari e poiché le variabili sono stocasticamente indipendenti si ha

$$V(\bar{X}_{200}) = V\left(\sum_{i=1}^{200} \frac{1}{200} X_i\right) = \sum_{i=1}^{200} \left(\frac{1}{200}\right)^2 V(X_i) = 200 \cdot \left(\frac{1}{200^2} 940\right) = \frac{940}{200} = 4,7$$

$$\begin{aligned} \text{c) } P(\bar{X}_{200} > 3995) &= P\left(Z > \frac{3995 - 4000}{\sqrt{4,7}}\right) = P\left(Z > -\frac{5}{2,16975}\right) = 1 - P(Z < -2,30633) = \\ &= 1 - [1 - P(Z < 2,31)] = P(Z < 2,31) = 0,9896 \end{aligned}$$

Università C. Cattaneo, Corso di Laurea in Economia Aziendale, A.A. 2014-2015

ESERCIZIO 5 (punti 5) Siano X e Y variabili aleatorie la cui distribuzione congiunta è assegnata dalla seguente tabella a doppia entrata.

X \ Y	0	1	2
0	0	0.08	0.32
1	0.09	0.34	0.17

- a) Si dica, giustificando la risposta, se X e Y sono indipendenti.
- b) Si calcoli la covarianza di X e Y.
- c) Si verifichi che X ha distribuzione bernoulliana, determinandone il parametro.
- d) Si verifichi che Y ha distribuzione binomiale di parametri 2 e 0.7.
- e) Si calcoli la probabilità (congiunta) che X sia uguale a 1 e Y minore o uguale a 1.

a) $p_{XY}(0,0) = 0 \neq p_X(0) p_Y(0) = 0,4 \cdot 0,09 = 0,036 \Rightarrow X$ e Y non sono stocasticamente indipendenti

b) Servono le medie delle due marginali X e Y:
 $E(X) = 0 \cdot 0,4 + 1 \cdot 0,6 = 0,6$
 $E(Y) = 0 \cdot 0,09 + 1 \cdot 0,42 + 2 \cdot 0,49 = 1,4$
 $Cov(X, Y) = 1 \cdot 1 \cdot 0,34 + 1 \cdot 2 \cdot 0,17 = 0,68 - 0,6 \cdot 1,4 = 0,68 - 0,84 = -0,16$

c) Si ha

$$p_X(x) = \begin{cases} 0,4 & x=0 \\ 0,6 & x=1 \\ 0 & \text{altrove} \end{cases} \Rightarrow \begin{cases} S_X = \{0,1\} \\ p_X(1) = 0,6 = p \end{cases} \Rightarrow X \text{ e' } Be(0,6)$$

d) [I modo] Affinche' Y sia $Bi(2; 0,7)$, e' sufficiente (e necessario) verificare solo due probabilita' di Y (infatti la terza probabilita' e' data in ogni caso da 1 meno la somma delle altre due). La verifica da' dunque

$$p_Y(y) = \begin{cases} 0,09 & y=0 \Rightarrow (2!/0!2!) \cdot 0,7^0 \cdot 0,3^2 = 0,09 \\ 0,42 & y=1 \Rightarrow (2!/1!1!) \cdot 0,7 \cdot 0,3 = 0,42 \\ 0,49 & y=2 \Rightarrow (2!/2!0!) \cdot 0,7^2 \cdot 0,3^0 = 0,49 \\ 0 & \text{altrove} \end{cases} \Rightarrow X \text{ e' } Bi(2; 0,7)$$

[II modo] Affinche' Y sia una $Bi(2; p)$ e' necessario che valga l'equazione data da una probabilita' di Y e la corrispondente probabilita' della binomiale con p incognita, ovvero, nel caso i della prima probabilita' $p_Y(0) = 0,09$, l'equazione qui sotto da cui risulta $p = 0,7$

$$(2!/0!2!) \cdot p^0 \cdot (1-p)^2 = 0,09 \Rightarrow (1-p)^2 = 0,09 \Rightarrow (1-p) = 0,3 \Rightarrow p = 0,7$$

Poi, con $p = 0,7$, va comunque verificata un'altra probabilita' di Y come fatto nel I modo, oppure va impostata un'altra equazione simile alla precedente con incognita p, cioe' l'equazione qui sotto da cui risulta ancora $p = 0,7$:

$$(2!/2!0!) \cdot p^2 \cdot (1-p)^0 = 0,49 \Rightarrow p^2 = 0,49 \Rightarrow p = 0,7$$

E' cosi' completata la verifica che Y e' $Bi(2; 0,7)$.

e) $P(X = 1, Y \leq 1) = p_{XY}(1,0) + p_{XY}(1,1) = 0,09 + 0,34 = 0,43$

ESERCIZIO 6 (punti 8) – Analisi dei dati Excel

Per ciascuna delle ultime 200 fatture emesse dall'azienda Az Srl si è rilevato il numero delle unità del prodotto A (variabile X) e del prodotto B (variabile Y) che sono state fatturate. A partire da tali dati grezzi, con i comandi Tabella-pivot di Excel si è ottenuta la seguente tabella a doppia entrata delle frequenze relative (percentuali) congiunte e marginali riguardanti le dette unità fatturate dei due prodotti.

	Y				
X	1	3	4	5	
1	10,00%				10,00%
2		10,00%		30,00%	40,00%
3				10,00%	10,00%
4			10,00%	30,00%	40,00%
	10,00%	10,00%	10,00%	70,00%	100,00%

Università C. Cattaneo, Corso di Laurea in Economia Aziendale, A.A. 2014-2015

Sulla base di tale tabella si risponda alle seguenti domande:

- a) Per la variabile X si specifichino la colonna dei valori di detta variabile e la colonna delle corrispondenti frequenze relative % cumulate. Cio' fatto, si determini il sesto decile della variabile X e si specifichi il significato del valore ottenuto.
- b) Mostrando i principali calcoli si determini il valore medio e lo scarto quadratico medio della variabile X.
- c) Sapendo che lo scarto quadratico medio della variabile Y e' 1,96 si specifichi quale delle due variabili Y e X ha minore variabilita'.
- d) Tenendo presente che il valore medio della variabile Y e' 4,3, si determini la covarianza ed il coefficiente di correlazione lineare delle due variabili X e Y specificando inoltre cosa indica il valore ottenuto di detto coefficiente di correlazione.
- e) Se si fosse calcolato il coefficiente di correlazione della variabile Y con se' stessa quale sarebbe stato il valore ottenuto?

(a) Il sesto decile e' dato da 3 unita' fatturate del prodotto A e tale valore indica che nel 60,00% delle fatture emesse sono state fatturate al massimo 3 unita' di detto prodotto (e quindi nel restante 40,00% sono state fatturate piu' di tre unita' di prodotto).

q.prod.A	fr.r.%cum
1	10,00%
2	50,00%
3	60,00%
4	100,00%

(b) La media della variabile X e' 2,8 e il suo s.q.m. e' 1,08. Uno schema di calcolo con Excel di media e s.q.m e' il seguente:

x_k	f_k	$x_k f_k$	$(x_k - \bar{x})^2$	$(x_k - \bar{x})^2 f_k$
1	0,10	0,1	3,24	0,324
2	0,40	0,8	0,64	0,256
3	0,10	0,3	0,04	0,004
4	0,40	1,6	1,44	0,576
	1,00	2,8		1,16
	\bar{x} media			varianza X
				s.q.m. X

(c) Poiche' si ha: $s.q.m.(X) = 1,08 < s.q.m.(Y) = 1,96$, si conclude che ha minore variabilita' la variabile X. La stessa conclusione si ottiene confrontando le due varianze (poiche' le media di X e la media di Y hanno lo stesso ordine di grandezza non si usano i coefficienti di variazione per il confronto della variabilita' di X e Y).

(d) $Cov(X,Y) = 12,8 - 2,8 \cdot 4,3 = 0,76$. $Rho(X,Y) = 0,76 / (1,08 \cdot 1,96) = 0,36$ che indica un basso grado di correlazione lineare di tipo diretto o positivo) tra le due variabili X e Y. Uno schema per il calcolo con Excel della covarianza e' il seguente:

x_k	y_k	$f(x_k, y_k)$	$x_k y_k f(x_k, y_k)$
1	1	0,1	0,1
2	3	0,1	0,6
2	5	0,3	3
3	5	0,1	1,5
4	4	0,1	1,6
4	5	0,3	6
			12,8

e) $\rho(Y,Y) = 1$. (il coefficiente di correlazione lineare di qualsiasi variabile con se' stessa e' sempre pari a 1 poiche' qualsiasi variabile e' perfettamente correlata linearmente in modo diretto con se' stessa).

[soluzioni redatte da R. D'Angio', Castellanza, 10/11/2014]