Università Carlo Cattaneo

Corso di laurea in Economia Aziendale

STATISTICA -- 13.02.2008

NB: **(A)** Ai fini della valutazione verranno considerate solo le risposte riportate dallo studente negli appositi riquadri bianchi del testo d'esame. **(B)** Nello svolgimento del compito, si utilizzino almeno quattro cifre decimali. **(C)** Allo studente/studentessa che consulti foglietti, appunti, libri , ecc o che parli con altri sarà annullata la prova d'esame.

COGNOME:______________________________ NOME:______________________ MATR:_____

ESERCIZIO 1.1 (punti 4). Dall'esperienza passata si sa che la probabilità che la scorta alla fine del mese di un certo prodotto in magazzino risulti esaurita è pari a 0.4.

a) **(punti 2)** Mostrando gli opportuni calcoli, si determini la probabilità che la scorta alla fine del mese risulti esaurita per 2 mesi su 5 mesi.

$$\Pr(Y=2)=\binom{5}{2}0.4^2(1-0.4)^{5-2}=10\cdot 0.16\cdot 0.216=0.3456$$

b) **(punti 2)** Mostrando gli opportuni calcoli, si determini la probabilità che la scorta alla fine del mese sia esaurita per almeno due mesi su 5 mesi.

$$\Pr(Y\geq 2)=1-\Pr(Y=0)+\Pr(Y=1)=1-\left[\binom{5}{0}0.4^0 0.6^5+\binom{5}{1}0.4^1 0.6^4\right]=1-\left(0.07776+2\cdot 0.1296\right)=$$

$$=1-\left(0.07776+0.2592\right)=1-0.33696=0.66304$$

ESERCIZIO 1.2 (punti 4). Il rendimento di un titolo nel prossimo mese è una variabile aleatoria normale standardizzata.

Mostrando i calcoli, si determini la probabilità che il rendimento sia maggiore di 0.07.

$$P\left(N(0;1)>0.07\right)=1-P\left(N(0;1)\leq 0.07\right)=1-0.5279=0.4721$$

ESERCIZIO 1.3 (punti 2). Si determinino, mostrando i calcoli effettuati, il valore atteso e la varianza di una variabile aleatoria X bernoulliana di parametro p .

$$E(X)=0\cdot p+1\cdot(1-p)=1-p$$

$$(1)\ V(X)=\left[0-E(X)\right]^2 p+\left[1-E(X)\right]^2(1-p)=(1-p)^2 p+p^2(1-p)=p(1-p)$$

$$(2)\ V(X)=E\left(X^2\right)-E(X)^2=(1-p)-(1-p)^2=(1-p)\,p$$

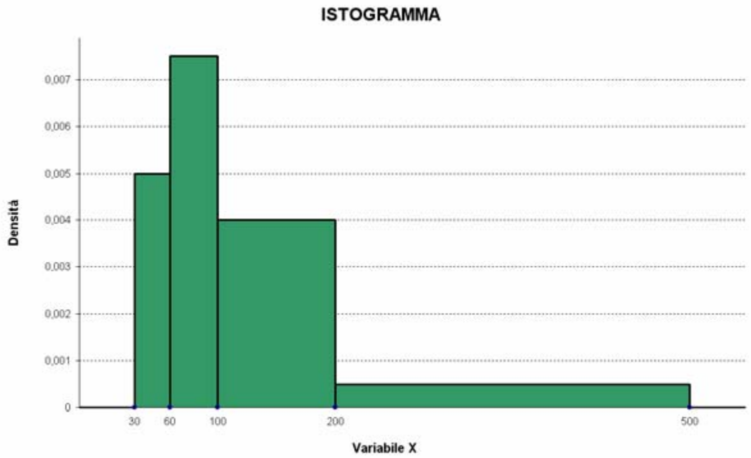
$$E\left(X^2\right)=0^2\cdot p+1^2\cdot(1-p)=1-p$$

ESERCIZIO 1.4 (punti 12). I prezzi (in euro) dei telefoni cellulari venduti in un negozio hanno dato luogo alla seguente distribuzione per intervalli:

Intervallo	Frequenza relativa
[30,60)	0.15
[60,100)	0.3
[100,200)	0.4
[200,500)	0.15

a) **(punti 3)** Si rappresenti mediante un grafico (istogramma) la distribuzione dei prezzi dei telefoni.

Le densità delle quattro classi sono rispettivamente 0.005, 0.0075, 0.004 e 0.0005.



b) (punti 4) Si determinino, mostrando i principali calcoli, media e varianza dei prezzi dei telefoni cellulari.

Indicando con m_i *i valori centrali delle classi e con* $f(x_i)$ *le frequenze relative, la media è:*

$$M(X)=\sum_{i=1}^{4} m_i f(x_i)=45\cdot 0.15+80\cdot 0.3+150\cdot 0.4+350\cdot 0.15=6.75+24+60+52.5=143.25$$

Mentre la varianza, utilizzando la formula di calcolo:

$$M\left(X^2\right)=\sum_{i=1}^{4} m_i^2 f(x_i)=2025\cdot 0.15+6400\cdot 0.3+22500\cdot 0.4+122500\cdot 0.15=$$

$$=303.75+1920+9000+18375=29598.75$$

$$V(X)=M\left(X^2\right)-M\left(X\right)^2=29598.75-143.25^2=29598.75-20520.56=9078.188$$

c) (punti 3) Si determini, mostrando i principali calcoli, la percentuale di telefoni cellulari che hanno prezzo compreso tra 50 e 150 euro.

$$Fr(50\le X\le 150)=Fr(50\le X\le 60)+Fr(60\le X\le 100)+Fr(100\le X\le 150)=$$

$$=(60-50)\cdot 0.005+0.3+(150-100)\cdot 0.004=0.05+0.3+0.2=0.55$$

d) (punti 2) Si determini, mostrando i principali calcoli, il terzo quartile dei prezzi dei telefoni.

Il terzo quartile appartiene alla terza classe, quindi:

$$0.45+(Q3-100)\cdot 0.004=0.75$$

da cui

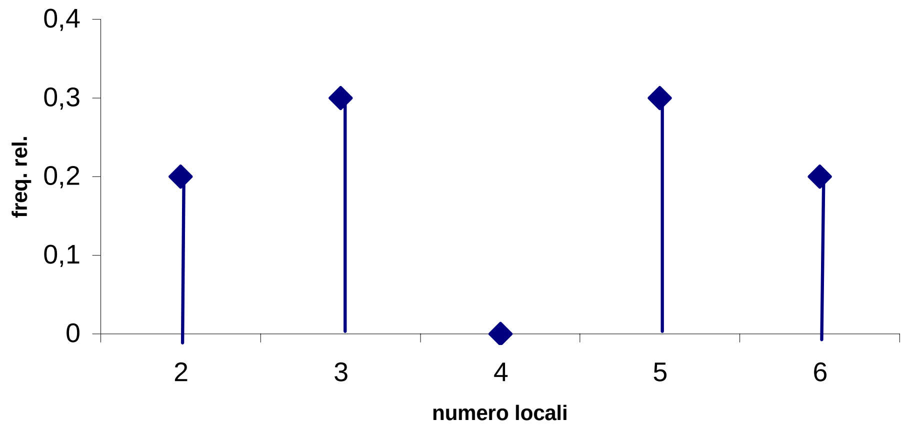
$$Q3=\frac{0.75-0.45}{0.004}+100=175$$

ESERCIZIO 1.5 (punti 10) Viene rilevato il numero di locali di cui sono composti 10 appartamenti di un certo quartiere. I dati così ottenuti sono riportati di seguito: 2,3,3,5,2,6,5,6,5,3.

a) **(punti 2)** Si scriva la distribuzione di frequenze del carattere “numero di locali”.

$$X=\begin{cases} 2 & 3 & 5 & 6 \\ 0.2 & 0.3 & 0.3 & 0.2 \end{cases}$$

b) **(punti 2)** Si rappresenti graficamente il carattere “numero di locali”.



c) **(punti 4)** Si determinino media e mediana del carattere “numero di locali”.

$$M(X)=\sum_{x=2}^{6} xf(x)=2\cdot 0.2+3\cdot 0.3+5\cdot 0.3+6\cdot 0.2=4$$

(è possibile arrivare alla stessa conclusione, semplicemente osservando la simmetria del grafico)

MEDIANA(X) = 3

Infatti, x=3 è il più piccolo valore in corrispondenza del quale la frequenza cumulata è maggiore o uguale a 0.5.

d) **(punti 2)** Sulla base di quanto ottenuto ai punti c) e d) precedenti, si dica se la distribuzione del carattere in esame è approssimativamente simmetrica, obliqua a destra o obliqua a sinistra, motivando la risposta.

Il grafico al punto b mostra chiaramente che la distribuzione del carattere X=”numeri di locali” è simmetrica. In questo caso, non è possibile osservare che M(X) = MEDIANA(X) poiché il valore centrale x=4 non presenta alcuna frequenza.

ESERCIZIO 2.1 (punti 3). Il profitto di un'azienda nel prossimo anno è una variabile aleatoria con distribuzione normale di media 40 e varianza 1600.

a) (punti 3) Mostrando i calcoli, si determini la probabilità che il profitto sia compreso tra 30 e 70.

Sia X il profitto dell'azienda; la probabilità richiesta è pari a:

$$P(30 < X < 70) = P\left(\frac{30-40}{\sqrt{1600}} < N(0;1) < \frac{70-40}{\sqrt{1600}}\right) = P(-0.25 < N(0;1) < 0.75) =$$

$$= P(N(0;1) < 0.75) - P(N(0;1) \leq -0.25) = P(Z < 0.75) - [1 - P(Z < 0.25)] =$$

$$= 0.7734 - (1 - 0.5987) = 0.3721$$

ESERCIZIO 2.2 (punti 4) Le variabili aleatorie X e Y rappresentano i costi in euro di un'impresa edile per il prossimo mese relativi rispettivamente al personale ed all'acquisto di materiali.

Considerato il costo totale *T=X+Y* e supponendo che costi per personale e per acquisto materiali siano indipendenti, si calcoli il costo totale atteso e la varianza del costo totale.

$$E(T) = E(X+Y) = E(X) + E(Y)$$

$$V(T) = V(X+Y) = V(X) + V(Y)$$

ESERCIZIO 2.3 (punti 8) Data una popolazione statistica X con valore atteso μ e varianza σ^2, si considerino i due seguenti stimatori del parametro μ:

$$T_1(X_1, X_2, X_3) = \frac{2X_1 + X_2 - X_3}{3}, \quad T_2(X_1, X_2, X_3) = \frac{X_1 + X_2 + X_3}{3}$$

(**a**) (punti 4) Si determinino mostrando i calcoli valore atteso, varianza ed eventuale distorsione dei due stimatori

$$E(T_1) = E\left(\frac{2X_1 + X_2 - X_3}{3}\right) = \frac{2E(X_1) + E(X_2) - E(X_3)}{3} = \frac{2}{3}\mu$$

$$D_{T_1}(\mu) = E(T_1) - \mu = \frac{2}{3}\mu - \mu = -\frac{1}{3}\mu$$

$$V(T_1) = V\left(\frac{2X_1 + X_2 - X_3}{3}\right) = \frac{4V(X_1) + V(X_2) + V(X_3)}{9} = \frac{6}{9}\sigma^2 = \frac{2}{3}\sigma^2$$

$$E(T_2) = E\left(\frac{X_1 + X_2 + X_3}{3}\right) = \frac{E(X_1) + E(X_2) + E(X_3)}{3} = \frac{3}{3}\mu = \mu$$

$$D_{T_2}(\mu) = E(T_2) - \mu = \mu - \mu = 0$$

$$V(T_2) = V\left(\frac{X_1 + X_2 + X_3}{3}\right) = \frac{V(X_1) + V(X_2) + V(X_3)}{9} = \frac{3}{9}\sigma^2$$

(**b**) (punti 4) In base ai risultati ottenuti sopra si determini mostrando i calcoli l'errore quadratico medio dei due stimatori

$$EQM_{T_1}(\theta) = V(T_1) + D_{T_1}(\mu)^2 = \frac{2}{3}\sigma^2 + \left(-\frac{1}{3}\mu\right)^2 = \frac{2}{3}\sigma^2 + \frac{1}{9}\mu^2$$

$$EQM_{T_2}(\theta) = V(T_2) + D_{T_2}(\mu)^2 = \frac{3}{9}\sigma^2 - 0^2 = \frac{1}{3}\sigma^2$$

ESERCIZIO 2.4 (8 punti) Un'indagine campionaria eseguita nel mese di ottobre di quest'anno con un campione di 400 persone sul consumo mensile medio pro-capite di pane ha fornito $\bar{x} = 2.1$ chilogrammi di pane ed $s_c^2 = 16$.

(**a**) (punti 3) Si indichi uno stimatore non distorto per il consumo medio pro-capite mensile μ_X e la corrispondente stima puntuale $\hat{\mu}_X$ per il consumo medio pro-capite del mese di ottobre di quest'anno.

Uno stimatore non distorto del consumo medio pro-capite (valore atteso di X) è la media campionaria, la corrispondente stima è pari a $\bar{x} = 2.1$.

(**b**) (punti 5) Si determini l'intervallo di confidenza, al 95%, per il consumo medio mensile pro-capite del mese di ottobre di quest'anno.

$$(\bar{x} \pm z_{1-\alpha/2}\frac{s_C}{\sqrt{n}}) = (2.1 \pm z_{0.975}\sqrt{\frac{16}{400}}) = (2.1 \pm 1.96 \cdot 0.2) = (2.1 \pm 0.392) = (1.708, 2.492)$$

ESERCIZIO 2.5 **(9 punti)** Si consideri il modello di regressione lineare $Y_i = \beta_0 + \beta_1 x_i + \varepsilon_i$ ($i = 1, ..., n = 21$). Le coppie osservate (x_i, y_i) sono state elaborate con il software statistico SPSS che ha dato i risultati riportati nel tabulato di cui alla pagina successiva. Sulla base di tale tabulato si risponda alle seguenti domande.

(a) (punti 3) Si scriva la stima puntuale di β_0.

$\hat{\beta}_0 = 80.038$

(b) (punti 3) Si preveda il valore di Y in corrispondenza ad un valore della variabile indipendente X pari a 20.

$\hat{Y} = \hat{\beta}_0 + \hat{\beta}_1 20 = 80.038 - 1.661 \cdot 20 = 46.818$

(c) (punti 3) Si dica, sulla base del valore di un opportuno indicatore, se la previsione effettuata al punto precedente si può considerare oppure no affidabile oppure no.

Il coefficiente di determinazione (R^2) è pari a 0.913, quindi la previsione è da considerarsi affidabile.

TABULATO SPSS

Model Summary

Model	R	R Square	Adjusted R Square	Std. Error of the Estimate
1	.956(a)	.913	.909	4.10447

a Predictors: (Constant), X

ANOVA(b)

Model		Sum of Squares	df	Mean Square	F	Sig.
1	Regression	3379.723	1	3379.723	200.617	.000(a)
	Residual	320.086	19	16.847		
	Total	3699.810	20			

a Predictors: (Constant), X
b Dependent Variable: Y

Coefficients(a)

Model		Unstandardized Coefficients		Standardized Coefficients	t	Sig.
		B	Std. Error	Beta		
1	(Constant)	80.038	3.242		24.687	.000
	X	-1.661	.117	-.956	-14.164	.000

a Dependent Variable: Y