Probabilità e Statistica
Scritto: 27 Giugno 2005 ore 10.00
Consegna-orali: 29 giugno ore 10.00

Cognome e Nome ______________________________

Numero matricola ____________________Firma ______________________

ESERCIZIO 1

Un'agenzia di autonoleggio possiede due tipi di auto per la classe D (60 del tipoA e 40 del tipoB). Viene rilevato il numero di guasti nei primi 50.000 km e si ottengono i seguenti risultati :

numero guasti	tipoA	tipoB
0	33	9
1	20	13
2	6	10
3	1	5
4	0	3

a) Calcolare il numero medio di guasti per i due tipi di auto.

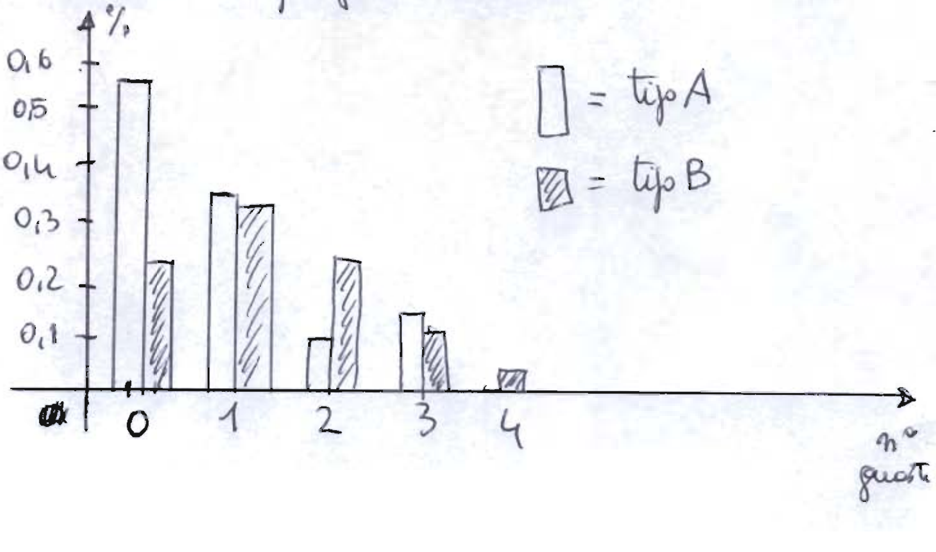
$$\text{n° medio guasti tipo A} = \frac{33\cdot 0 + 20\cdot 1 + 6\cdot 2 + 1\cdot 3 + 0\cdot 4}{60} = \frac{35}{60} = 0{,}5$$

$$\text{n° medio guasti tipo B} = \frac{9\cdot 0 + 13\cdot 1 + 10\cdot 2 + 5\cdot 3 + 3\cdot 4}{40} = \frac{50}{40} = 1{,}2$$

b) Disegnare il grafico a barre del numero di guasti per i due tipi di auto.

Per disegnare un grafico a barre per confrontare le due distribuzioni è necessario costruire i profili colonne.

n° guasti	tipo A	tipo B
0	0,55	0,225
1	0,333	0,325
2	0,10	0,25
3	0,167	0,125
4	0	0,075
tot	100	100



ESERCIZIO 2

Una malattia virale ha infettato il 25% dei polli di un allevamento. Un test per la presenza del virus fornisce esito positivo (virus presente) l'84% delle volte in cui il pollo è malato e fornisce esito negativo (virus assente) l'80% delle volte in cui il pollo non ha la malattia.

Si sceglie a caso un pollo e lo si sottopone ad analisi.

a) Completare la tabella seguente percentuale

	TEST+	TEST -	totale
Virus presente	21	4	25
Virus assente	15	60	75
totale	36	64	100

b) Se il test è positivo, qual è la probabilità che il pollo sia stato infettato dal virus?

$$p = \frac{21}{36} \cong 0,58$$

c) Se il test è negativo, qual è la probabilità che il pollo sia stato infettato dal virus?

$$p = \frac{4}{64} \cong 0,02$$

ESERCIZIO 3

Un vaccino è efficace per immunizzare contro una certa malattia il 99.4% delle persone trattate. Si determini, per 125 persone vaccinate, quanto vale la probabilità che il vaccino non sia efficace per tre persone.

La probabilità che il vaccino non sia efficace vale $p = 1 - 0,994 = 0,006$. La variabile che modella l'esperimento è $X \sim \mathcal{B}(125; 0,006)$ e si deve calcolare

$$\mathbb{P}(X = 3) = \binom{125}{3} (0,006)^3 (0,994)^{122}.$$

ESERCIZIO 4

Utilizzare i dati riportati in tabella per stabilire se la distribuzione della variabile X ha legge uniforme (effettuare un test chiquadrato a livello 1%).

Variabile X	A	B	C
Valori osservati	25	50	25
Valori attesi	33.3	33,3	33,3

$$C = \frac{(25-33,3)^2}{33,3} + \frac{(50-33,3)^2}{33,3} + \frac{(25-33,3)^2}{33,3} =$$

$$= 2,07 + 24,05 + 2,07 = 28,19$$

$$\chi^2_{1\%}(2) = 9,2104$$

$\chi^2_{1\%}(2) < C \Rightarrow$ si rifiuta l'ipotesi che la variabile X abbia distribuzione (legge) uniforme.

ESERCIZIO 5

I dati riportati nella tabella seguente sono misure che mettono in relazione un particolare parametro di funzionalità epatica (SGOT) con il livello di colesterolo HDL nel sangue.

SGOT [x]	9.5	11	13.5	15.5	17.5	19.5	20.5
HDL (mg/dL) [y]	40	41.2	42.3	42.8	43.8	43.6	46.5

$$\sum x_i = 107 \quad \sum x_i^2 = 1740.5 \quad \sum x_i y_i = 4637.6$$

$$\sum y_i = 300.2 \quad \sum y_i^2 = 12900.2$$

Si utilizzino i dati riportati per:

a) Calcolare media e varianza delle variabili SGOT e HDL.

$$\bar{x} = \frac{\sum x_i}{7} = 15.29 \qquad VAR(x) = \frac{\sum x_i^2}{7} - \bar{x}^2 = 14.86$$

$$\bar{y} = \frac{\sum y_i}{7} = 42.89 \qquad VAR(y) = \frac{\sum y^2}{7} - \bar{y}^2 = 3.33$$

b) Calcolare la covarianza fra le variabili SGOT e HDL.

$$cov(x,y) = \left(\frac{1}{n} \sum x_i y_i\right) - \bar{x}\bar{y} = \frac{4637,6}{7} - 15.29 \cdot 42.89 =$$

$$= 6,726$$

c) Determinare l'espressione della retta di regressione fra il valore di SGOT (x) e HDL (y).

$$y = \frac{cov(x,y)}{VAR(x)} (x - \bar{x}) + \bar{y} =$$

$$= \frac{6,726}{14.86} (x - 15,29) + 42.89 = 0,45x + 35,97$$

ESERCIZIO 6

Qui a fianco sono riportati i risultati di una rilevazione quantitativa, ordinati in modo crescente.

percentile di ordine p: $x_{(n+1)p}$ se $(n+1)p$ non intero

$x_{(n+1)p} + x_{(n+2)p}$ se $(n+1)p$ è intero

a) Indicare il primo quartile, la mediana, il terzo quartile e la distanza interquartile

Q1: 2

Mediana 3

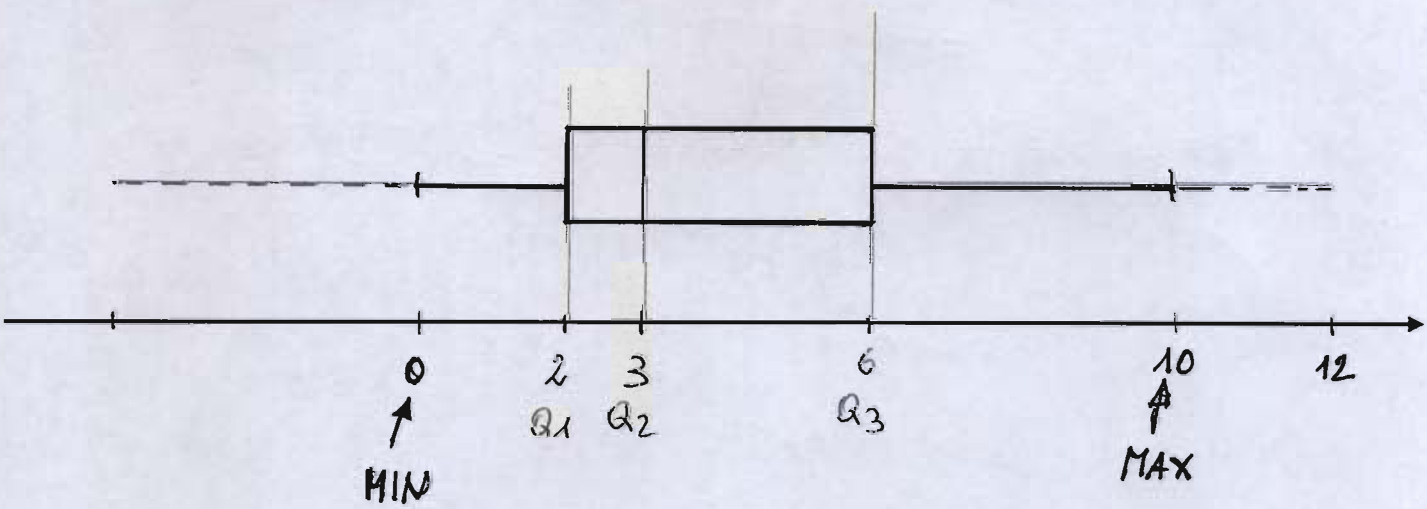
Q3: 6

IQR $Q_3 - Q_1 = 4$

$1{,}5\,(IQR) = 6$: lunghezza "teorica" dei baffi.

i	*X*	*i*	*X*
1	0	21	3
2	0	22	4
3	1	23	4
4	1	24	4
5	2	25	4
6	2	26	5
7	2	27	5
8	2	28	5
9	2	29	6
10	2	30	6
11	2	31	6
12	2	32	7
13	2	33	7
14	2	34	7
15	3	35	7
16	3	36	9
17	3	37	10
18	3	38	10
19	3	39	10
20	3	40	10

b) Costruire il boxplot della variabile X.



ESERCIZIO 7

Sia X la variabile aleatoria che modella la caratteristica di una popolazione e sia $X_1, X_2, \ldots, X_{16}$ un campione estratto da X.

a) Indicare uno stimatore non distorto per la media μ di X.

$$S = \overline{X} = \frac{X_1 + \cdots + X_{16}}{16}$$

b) Indicare lo stimatore non distorto per la varianza σ^2 di X.

$$S^2 = \frac{1}{15} \sum_{i=1}^{16} (x_i - \overline{x})^2$$

ESERCIZIO 8

Scrivere, in modo conciso, cosa significa che:" [2.3, 4.5] è un intervallo di confidenza per la media a livello 90%".

La popolazione X ha media μ incognita.

Un Int di Conf per μ è [2.3; 4.5] se la media μ appartiene all'intervallo con probabilità p = 0.90.

ESERCIZIO 9

Su un campione $X_1, X_2, \dots, X_n$, estratto da una popolazione con legge normale; determinare l'intervallo di confidenza per la media μ a livello 95% :

a) Con $n=30$, $\sigma^2=25$ e $\sum_{i=1}^{30} x_i = 88.2$.

$$\bar{x}_{30} = \frac{\sum x_i}{30} = \frac{88.2}{30} = 2{,}94$$

$$z_{1-\frac{\alpha}{2}} = z_{0.975} = 1{,}96$$

$$I = \left[2{,}94 - 1{,}96 \frac{\sqrt{25}}{\sqrt{30}} \, ; 2{,}94 + 1{,}96 \frac{\sqrt{25}}{\sqrt{30}} \right] =$$

$$= [1{,}15 \, ; \, 4{,}73]$$

b) Con $n=30$, σ^2 ignoto, $\sum_{i=1}^{30} x_i = 88.2$, $\sum_{i=1}^{30} (x_i - 2.94)^2 = 148.5$

$$S^2 = \frac{1}{n-1} \sum_{i=1}^{n} (x_i - \bar{x})^2 \quad ; \quad S^2 = \frac{1}{29} \cdot 148.5 = 5.12$$

$$t^{29}_{0{,}05} = 1{,}6991 \quad ; \quad t^{29}_{0{,}025} = 2.0452$$

$$I = \left(2.94 - 2.0452 \cdot \frac{\sqrt{5.12}}{\sqrt{30}} \, ; 2.94 + 2.0452 \cdot \frac{\sqrt{5{,}12}}{\sqrt{30}} \right) =$$

$$= (2.095 \, ; \, 3.78)$$

ESERCIZIO 10

Un'industria alimentare che produce bibite dichiara che il contenuto nominale di una lattina è 0.33 ml . In realtà il contenuto è una variabile aleatoria di legge normale con media e varianza sconosciute.

Si scelgono n=20 lattine e si ottengono i seguenti risultati :

$$\sum_{i=1}^{20} x_i = 6560\,,\ \sum_{i=1}^{20} (x_i - 328)^2 = 194.56$$

a) Calcolare la media e la varianza campionaria.

$$\overline{X}_{20} = \frac{6560}{20} = 328$$

$$S^2 = \frac{1}{19} \cdot 194{,}56 = 10{,}24$$

b) Effettuare un test a livello 5% con le seguenti ipotesi : $\begin{cases} H_0 : \mu=330 \\ H_1 : \mu<330 \end{cases}$

Regione di rifiuto :

$$\left(-\infty, \mu - t^{19}_{0{,}05} \frac{S}{\sqrt{20}}\right) \cong \left(-\infty\,, 330 - 1{,}7291\, \frac{3{,}2}{4{,}5}\right) =$$

$$= \left(-\infty; 328{,}8\right)$$ $\overline{x}_{20} = 328 \in R$, quindi non si accetta H_0

c) Mantenendo le stesse ipotesi del punto **b)**, se si accetta H_0 a livello 5%, a quale livello si è certi di accettare ancora l'ipotesi H_0? (motivare ogni affermazione)

- 10%
- 3%

Se si accetta H_0 a livello α, si accetta H_0 per ogni valore $< \alpha$.

Quindi se si accetta a livello $\alpha = 5\%$, si accetta a livello 3% e non si hanno informazioni sul livello 10%.

ESERCIZIO 11

Una scatola contiene n=120 dadi da gioco. 40 dadi sono equilibrati, i restanti la probabilità di ottenere 3 vale 0.3, di ottenere 5 vale 0.3 e i restanti numeri escono con probabilità 0.1 .

a) Scrivere la legge del dado non equilibrato.

x	1	2	3	4	5	6	tot
$p(x)$	0,1	0,1	0,3	0,1	0,3	0,1	1

b) Si estrae a caso un dado e lo si lancia. Calcolare la probabilità di ottenere 4.

$$\mathbb{P}(\text{ottenere } 4) = \mathbb{P}(\text{ottenere } 4 / \text{dado eq.})\, \mathbb{P}(\text{dado eq}) + \mathbb{P}(\text{ottenere } 4 / \text{dado non eq})\, \mathbb{P}(\text{dado non eq}) =$$

$$= \frac{1}{10} \cdot \frac{40}{120} + \frac{1}{6} \cdot \frac{80}{120} = 0{,}034 + 0{,}11 = 0{,}144$$

c) Quale risultato si otterà, in media, per ciascuno dei due dadi ?

Dado equilibrato : media = 3,5

Dado non equilibrato : media = 3.7