

**PROVA SCRITTA DI STATISTICA  
CLEA-CLEFIN-CLELI (COD. 4038)  
6 novembre 2002**

L'Associazione Industriali di una provincia lombarda ha commissionato un'indagine per studiare le esigenze del mercato del lavoro locale. Allo scopo sono stati somministrati 2000 questionari alle industrie della provincia. I dati per 15 di queste aziende sono raccolti nella seguente tabella:

- L Livello di studi prevalente del personale impiegato a tempo indeterminato  
 A Ampiezza misurata attraverso il numero di dipendenti  
 S Settore prevalente di attività  
 CA Numero di contratti di lavoro diversi da quelli a tempo indeterminato  
 V Volume di affari per lavoratore a tempo indeterminato (in migliaia di euro)

L	A	S	CA	V
laurea	25	alta tecnologia	6	228
media superiore	40	abbigliamento	4	62
media superiore	15	agroalimentare	10	80
laurea	20	alta tecnologia	3	600
media inferiore	45	agroalimentare	5	78
media inferiore	34	abbigliamento	0	44
laurea	150	agroalimentare	130	62
media inferiore	76	agroalimentare	9	224
media superiore	60	abbigliamento	4	36
media inferiore	35	agroalimentare	5	85
media superiore	21	alta tecnologia	5	548
media inferiore	62	agroalimentare	50	112
laurea	45	agroalimentare	40	83
laurea	10	alta tecnologia	0	600
media inferiore	30	agroalimentare	15	95

**ESERCIZIO 1. (2 punti)** Si dica, per ciascuna delle variabili considerate, quali tra gli strumenti indicati è possibile utilizzare:

	L	A	S	CA	V
Media					
Mediana					
Q1 (I quartile)					
Box-plot (grafico a scatola e baffi)					

**ESERCIZIO 2. (2 punti)** Si fornisca una opportuna rappresentazione grafica per la variabile S.

**ESERCIZIO 3. (3 punti)** Si costruisca la funzione di regressione di  $V$  su  $S$  e si dica se  $V$  è regressivamente indipendente da  $S$ .

**ESERCIZIO 4. (2 punti)** Si costruisca la tabella a doppia entrata relativa alle variabili  $L$  ed  $S$ . Si dica se esse sono oppure no statisticamente indipendenti.

**ESERCIZIO 5. (4 punti)** Siete spirati. Vi trovate di fronte a due porte identiche, una delle quali per l'Inferno ( $Inf$ ) e l'altra per il Paradiso ( $Par$ ). Siete destinati ad andare all'Inferno -- chi non lo è -- ma  $S$ . Pietro, per il vostro impegno profuso nel corso di Statistica, decide di proporvi comunque una via di salvezza per il Paradiso e vi lascia guardare dietro una delle due porte.  $S$ . Pietro vi informa che dietro la porta per l'Inferno vi sono un angelo e un diavolo. Dietro la porta per il Paradiso ci sono due angeli. Scegliendo a caso una delle due porte ( $P(Inf)=P(Par)=0.5$ ), la aprite e riuscite a scorgere solo una sagoma.

- Con quale probabilità scorgete un angelo?
- Se avete scorto un angelo, con quale probabilità avete aperto la porta per l'Inferno?

**ESERCIZIO 6. (4 punti)** Si consideri una variabile aleatoria  $X$  la cui funzione di ripartizione è:

$$F(x) = \begin{cases} 0 & x < -2 \\ 1/4 & -2 \leq x < 0 \\ 2/4 & 0 \leq x < 2 \\ 1 & x \geq 2 \end{cases}$$

- Si determini la funzione di probabilità di  $X$ .
- Si calcoli il valore atteso della variabile aleatoria  $Y=3X-2$ .
- Si calcoli  $P(X \geq 1)$ .

**ESERCIZIO 7. (7 punti)** E' noto che il prezzo medio di vendita di un determinato articolo nei negozi della periferia di una città è pari a 8. Indicato con  $m$  il prezzo medio (incognito) dello stesso articolo nei negozi del centro della città, si vuole verificare se  $m$  è differente da 8. Allo scopo si rileva il prezzo dell'articolo in un campione di 7 negozi del centro. Indicati con  $x_1, x_2, \dots, x_7$  i prezzi di vendita rilevati nei negozi del centro, si ha:

$$\sum_{i=1}^7 x_i = 63, \quad \sum_{i=1}^7 x_i^2 = 581,$$

- Si fornisca una stima del prezzo medio di vendita dell'articolo nei negozi del centro.
- Si fornisca una stima della varianza del prezzo nei negozi del centro.
- Supponendo normale la distribuzione del prezzo di vendita dell'articolo nei negozi del centro si stabilisca se, al livello  $\alpha = 0.05$ , è da rifiutare oppure no l'ipotesi  $m = 8$  contro l'alternativa  $m \neq 8$ .

**ESERCIZIO 8. (4 punti)** Si consideri una popolazione statistica  $X$  con distribuzione gaussiana, media incognita  $m$  e varianza nota  $s^2$ .

- a) Si scriva l'espressione dell'intervallo di confidenza per  $m$ , di coefficiente di confidenza  $1-\alpha$ , basato su un campione di numerosità  $n$  la cui realizzazione ha fornito il valore  $\bar{x}$  per la media campionaria.
- b) Si dica se, con lo stesso valore  $\bar{x}$  e con  $\alpha' > \alpha$ , l'ampiezza dell'intervallo risulta maggiore o minore della precedente; si giustifichi la risposta.

**ESERCIZIO 9. (2 punti)** Si dica quali operazioni è necessario effettuare per ottenere, con EXCEL, gli indicatori mediana e varianza di una  $n$ -upla di dati numerici.