

**PROVA SCRITTA DI STATISTICA**  
**CLEA-CLEFIN-CLELI (COD. 5047-4038)**  
**6 febbraio 2003 -- A**

**ESERCIZIO 1 (8 punti)**

Uno studio legale ha 7 uffici. Per ciascuno di questi uffici, vengono rilevati i caratteri:

X: numero di personal computer presenti nell'ufficio;

Y: numero di impiegati che lavorano nell'ufficio;

Z: superficie (in metri quadrati) dell'ufficio;

W: piano in cui l'ufficio si trova.

I dati rilevati sono presentati nella tabella che segue:

UFFICIO	X	Y	Z	W
I	3	3	16	I
II	2	3	22	II
III	2	2	13	I
IV	1	1	18	I
V	3	6	34	II
VI	2	3	20	I
VII	1	2	12	I

- Si fornisca una opportuna rappresentazione grafica del carattere W.
- Si dica, sulla base del valore di un opportuno indice numerico, se il carattere X può ritenersi molto, poco o mediamente concentrato.
- Si determini la tabella a doppia entrata relativa ai caratteri X e Y.
- Si determini la funzione di regressione di Y su X.

**ESERCIZIO 2 (6 punti)**

Si lanciano due dadi regolari a 4 facce e si considerano le variabili aleatorie X: "somma dei risultati dei due dadi" ed Y: "minimo tra i risultati dei due dadi".

- Si calcoli il valore atteso di X.
- Si calcoli  $P(X < 5)$ .
- Si dica se X e Y sono indipendenti.

**ESERCIZIO 3 (4 punti)**

Un quiz si compone di due parti. Il tempo (in minuti) necessario ad un individuo per risolvere la I parte è una variabile aleatoria con distribuzione normale con media 12 e varianza 9, mentre il tempo necessario alla risoluzione della seconda parte è una variabile aleatoria con distribuzione normale di media 9 e varianza 7. Si possono assumere indipendenti i tempi necessari alla risoluzione delle due parti.

- Si calcoli la probabilità che l'individuo risolva la prima parte del quiz in più di 13 minuti.
- Si calcoli la probabilità che l'individuo risolva l'intero quiz in meno di 17 minuti.

#### **ESERCIZIO 4 (4 punti)**

Si vuole verificare se la durata media  $m$  delle comunicazioni effettuate con telefono cellulare in una determinata fascia oraria è inferiore a 100 secondi. A questo scopo, si rilevano le durate  $x_1, x_2, \dots, x_{12}$  in un campione di 12 comunicazioni, ottenendo  $\bar{x} = 98$  e

$s^2 = \frac{1}{12} \sum (x_i - \bar{x})^2 = 10$ . Si assuma, per la durata di una telefonata, una distribuzione normale.

- Si fornisca una stima, ottenuta da uno stimatore non distorto, per la varianza  $\mathbf{s}^2$  della durata delle telefonate.
- Si determini una regione di rifiuto, relativa ad un test di dimensione 0.1, per l'ipotesi nulla  $H_0 : m \geq 100$  contro l'alternativa  $H_1 : m < 100$ .
- Si decida, sulla base dei dati del campione, se rifiutare oppure no l'ipotesi nulla utilizzando il test di cui al punto precedente.

#### **ESERCIZIO 5 (5 punti)**

Si consideri un modello lineare  $Y = \mathbf{b}_1 x + \mathbf{b}_0 + \mathbf{e}$  per i dati  $(x_1, y_1), \dots, (x_n, y_n)$ , con

$$\sum_{i=1}^n (x_i - \bar{x})^2 = 2, \sum_{i=1}^n (x_i - \bar{x})(y_i - \bar{y}) = 1, \bar{x} = 3, \bar{y} = 2.$$

- Si determinino stime, ottenute da stimatori non distorti, per i parametri  $\mathbf{b}_0$  e  $\mathbf{b}_1$  del modello.
- Sapendo che  $(0.32, 0.68)$  è un intervallo di confidenza per  $\mathbf{b}_1$  con coefficiente di confidenza 0.95, si decida se rifiutare oppure no, al livello 0.05, l'ipotesi nulla  $\mathbf{b}_1 = 0$  contro l'alternativa  $\mathbf{b}_1 \neq 0$ . Motivare la risposta.
- Si dica quale significato assumerebbe, con riferimento al modello lineare, il mancato rifiuto dell'ipotesi nulla  $\mathbf{b}_1 = 0$ .