

FIRMA DELLO STUDENTE

**PRIMA PROVA INTERMEDIA DI STATISTICA  
CLEA (COD. 5047/4038)  
5 Novembre 2003**

*Cognome*

*Nome*

*Numero di matricola*

**COMPITO C2**

**Ai fini della valutazione si terrà conto solo ed esclusivamente di quanto riportato negli appositi spazi. Al termine della prova, è OBBLIGATORIO consegnare il presente foglio ed il foglio di brutta (DI CUI NON SI TERRÀ CONTO AI FINI DELLA VALUTAZIONE).**

**APPROSSIMARE TUTTI I CALCOLI ALLA QUARTA CIFRA DECIMALE**

Supponiamo che, al termine della stagione 2002-03, la Federazione Italiana Pallacanestro abbia reso noti i seguenti dati relativi a sei caratteri rilevati su dodici società (squadre) che hanno partecipato al campionato di serie A1.

GIOCATORI            componenti la rosa dei giocatori a disposizione degli allenatori.  
 STAFF TECNICO      componenti lo staff tecnico: allenatori, preparatori atletici, massaggiatori.  
 BUDGET              entità del budget a disposizione della società all'inizio del campionato.  
 PUNTI                punti totalizzati al termine del campionato.  
 ABBONATI            numero di spettatori abbonati alla stagione 2002-03.  
 INCASSI              incasso medio per partita della stagione 2002- 03, (in migliaia di Euro).

I dati sono riportati nella tabella seguente:

	Squadre	Giocatori <b>G</b>	Staff <b>S</b>	Budget <b>B</b>	Punti <b>P</b>	Incassi <b>I</b>	Abbonati <b>A</b>	<b>P<sup>2</sup></b>	<b>P</b>	<b>A<sup>2</sup></b>	<b>P × I</b>
1.	Benetton	14	8	Cospicuo	827	21,8	1427	683929	475,24	2036329	18028,6
2	Skipper	12	6	Medio	631	16,3	1014	398161	265,69	1028196	10285,3
3	Oregon	14	7	Cospicuo	759	19,9	1402	576081	396,01	1965604	15104,1
4	Lottomatica	10	5	Esiguo	567	15,2	792	321489	231,04	627264	8618,4
5	Monte Paschi	12	5	Medio	604	16,3	873	364816	265,69	762129	9845,2
6	Air	14	6	Medio	648	17,6	1116	419904	309,76	1245456	11404,8
7	Mabo	11	5	Esiguo	518	15,8	746	268324	249,64	556516	8184,4
8	Trieste	14	7	Cospicuo	741	19,7	1368	549081	388,09	1871424	14597,7
9	Breil	12	7	Medio	673	18,1	1215	452929	327,61	1476225	12181,3
10	Pompea	11	6	Esiguo	589	16,1	834	346921	259,21	695556	9482,9
11	Metis	14	8	Cospicuo	729	19,4	1309	531441	376,36	1713481	14142,6
12	Snaidero	14	5	Medio	682	18,5	1278	465124	342,25	1633284	12617,0
<b>tot</b>		<b>152</b>	<b>75</b>		<b>7968</b>	<b>214,70</b>	<b>13374</b>	<b>5378200</b>	<b>3886,59</b>	<b>15611464</b>	<b>144492,3</b>

**1. (1 punto)** Determinare la moda del carattere GIOCATORI.

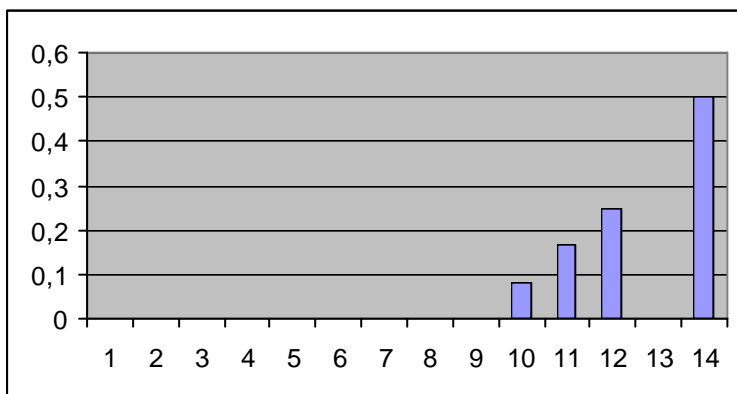
$$G = \left\{ \begin{array}{cccc} 10 & 11 & 12 & 14 \\ \frac{1}{12} & \frac{2}{12} & \frac{3}{12} & \frac{6}{12} \end{array} \right. \quad \text{Moda}=14$$

2. (3 punti) Costruire la distribuzione delle frequenze del carattere PUNTI ricodificato nelle classi [500, 600), [600, 730), [730, 830]. Calcolare la frequenza relativa del carattere nell'intervallo (540, 620].

a) Distribuzione delle frequenze			b) $Fr(540 < P \leq 620) =$
$x_i; x_{i+1}$	frequenze assolute	frequenze relative	
[500;600)	3	$\frac{3}{12} = \frac{1}{4} = 0.25$	$\frac{60}{100} \cdot \frac{1}{4} + \frac{20}{130} \cdot \frac{1}{2} = \frac{3}{20} + \frac{1}{13} = \frac{59}{260} = 0.2269$
[600;730)	6	$\frac{6}{12} = \frac{1}{2} = 0.5$	
[730;830]	3	$\frac{3}{12} = \frac{1}{4} = 0.25$	
(2 punti)			(1 punto)

3. (3 punti) Rappresentare graficamente la distribuzione delle frequenze del carattere GIOCATORI e calcolare media e mediana. Determinare la forma della distribuzione motivando opportunamente la conclusione raggiunta.

a) Grafico:



b) Media = 12.6666      Mediana = 13

c) Forma : Il grafico è obliquo a sinistra. Lo si vede dal grafico e dal fatto che la media è minore della mediana.

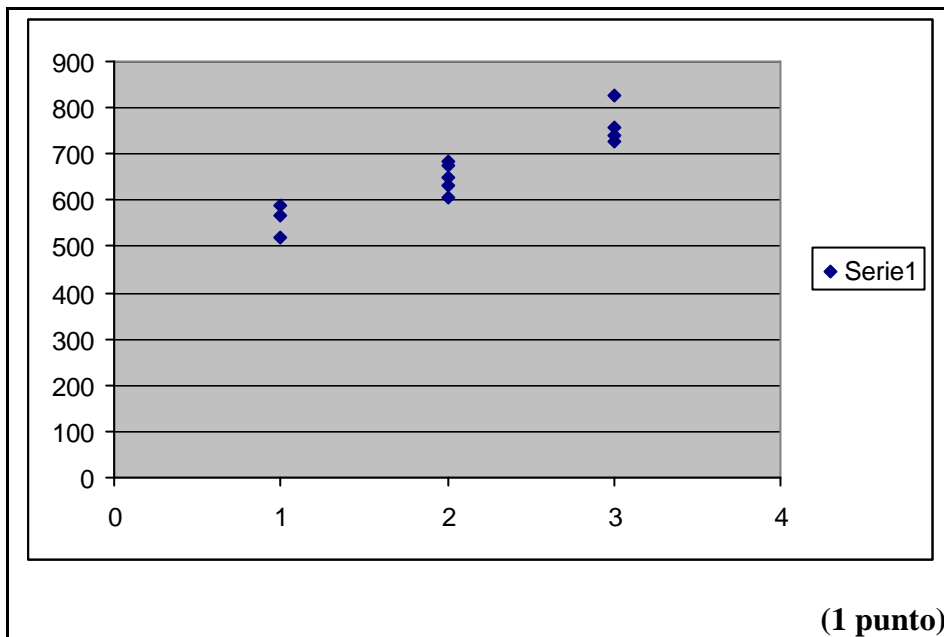
4. (3 punti) A partire dalla distribuzione delle frequenze del carattere STAFF di seguito riportata, calcolare l'indice di concentrazione e commentare il risultato ottenuto. (Completare la tabella con le quantità necessarie per calcolare l'indice.)

$x_i^*$	$n_i^*$	$F_i^*$	$Q_i^*$
5	4	$\frac{4}{12} = 0.3333$	$\frac{20}{75} = 0.2666$
6	3	$\frac{7}{12} = 0.5833$	$\frac{38}{75} = 0.5066$
7	3	$\frac{10}{12} = 0.8333$	$\frac{59}{75} = 0.7866$
8	2	1	1
tot	12		

$$R^* = \frac{\sum_{i=1}^{k-1} F_i^* - Q_i^*}{\sum_{i=1}^{k-1} F_i^*} = \frac{\sum_{i=1}^3 F_i^* - \sum_{i=1}^3 Q_i^*}{\sum_{i=1}^3 F_i^*} = \frac{\frac{21}{12} - \frac{117}{75}}{\frac{300}{12}} = \frac{\frac{57}{525}}{\frac{7}{4}} = 0.1086$$

Commento: la concentrazione è bassa. L'indice è vicino a 0.

5. (3 punti) a) Rappresentare graficamente la distribuzione congiunta dei caratteri B = BUDGET e P = PUNTI.



(b) Calcolare le medie subordinate di P | B. Il carattere PUNTI è regressivamente indipendente dal carattere BUDGET? Giustificare la risposta.

$$m(P | B = Esiguo) = \frac{1674}{3} = 558$$

$$m(P | B = Medio) = \frac{3238}{5} = 647.6$$

$$m(P | B = Cospicuo) = \frac{3056}{4} = 764$$

Il carattere Punti (P) non è regressivamente indipendente dal carattere BUDGET (B) perché la funzione di regressione non è costante (e pari alla media di P).

(2 punti)

6. (2 punti) Senza fare riferimento al *dataset* di pagina 1, in generale:

- Tra caratteri di quali tipologie è possibile calcolare l'indice  $\phi^2$  ?
- Cosa misura tale indice e quando risulta pari a zero?
- Dare la definizione di contingenza.

- a) qualitativi (nominali e ordinali) e quantitativi discreti con un numero non elevato di modalità differenti (oppure raggruppate) e continui se raggruppati in classi
- b) è un indice assoluto di connessione che vale 0 in caso di indipendenza statistica tra i due caratteri
- c) la contingenza è la differenza tra la frequenza congiunta osservata e quella che si avrebbe in caso di indipendenza statistica, cioè:  

$$c_{XY}(x_i, y_j) = p_{XY}(x_i, y_j) - p_X(x_i)p_Y(y_j) \quad \forall i, j$$

**7. (3 punti)** Calcolare il coefficiente di correlazione lineare tra i caratteri PUNTI e INCASSI. Commentare il risultato ottenuto.

$$r(P, I) = \frac{Cov(P, I)}{s_P s_I} = \frac{160.9362}{85.3659 \cdot 1.9415} = 0.9710$$

Il coefficiente di correlazione lineare è prossimo a 1 (suo valore massimo).  
 Il legame lineare tra i due caratteri è positivo (all'aumentare di uno dei due caratteri aumenta anche l'altro) e molto forte.

**8. (4 punti)** Nell'urna A ci sono tre palline bianche e due palline nere, nell'urna B ci sono una pallina nera e tre palline bianche. Si estraggono due palline, una da ciascuna urna. Costruire la distribuzione di probabilità della variabile aleatoria X = "numero di palline bianche estratte" e calcolarne il valore atteso.

$$p(x) = \begin{cases} \frac{2}{20} & x = 0 \\ \frac{9}{20} & x = 1 \\ \frac{9}{20} & x = 2 \\ 0 & \text{altrove} \end{cases} \quad E(X) = \frac{27}{20} = 1.35$$

**9. (2 punti)** Si assuma che un'azione abbia la medesima probabilità di salire di prezzo (pari a 4/5) e di scendere di prezzo (pari a 1/5) nei cinque giorni della settimana lavorativa e che le variazioni avvengano in modo indipendente. Detta X la variabile aleatoria "numero di volte che il prezzo scende nei cinque giorni", scrivere la funzione di probabilità di X.

$$p(x) = \begin{cases} \binom{5}{x} \left(\frac{1}{5}\right)^x \left(\frac{4}{5}\right)^{5-x} & x = 0, 1, 2, 3, 4, 5 \\ 0 & \text{altrove} \end{cases}$$