PRIMA PROVA INTERMEDIA DI STATISTICA

CLEA, CLEFIN (COD. 5047/4038/371/377)

3 Novembre 2004

Cognome ***Nome***

Numero di matricola

COMPITO B1

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

Un osservatore sportivo del basket ha raccolto alcuni dati relativi a **9** giocatori appartenenti a tre diverse squadre: Dallas, LA Lakers e G State. Ecco la tabella e la spiegazione delle variabili utilizzate (*notare che sono anche state riportate le somme dei valori e le somme dei quadrati dei valori*).

Giocatore: nome del giocatore | **Squadra**: squadra del giocatore

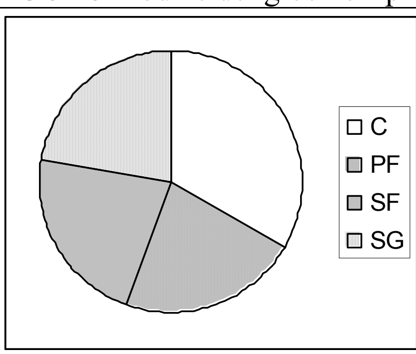
RUOLO: ruolo del giocatore (C= centro, SF = ala piccola, PF = ala grande, PG = playmaker)

PG: numero di partite giocate | **PI**: partite giocate dal primo minuto

MIN: media dei minuti giocati per partita | **PTS**: media stagionale dei punti del giocatore

Giocatore	Squadra	RUOLO	PG	PI	MIN	PTS
La Frentz	Dallas	C	69	43	23,3	9,3
Finley	Dallas	SF	69	69	38,3	19,3
Nowitzki	Dallas	PF	80	80	39,0	25,1
O' Neal	LA Lakers	C	68	66	37,8	27,5
Gorge	LA Lakers	PF	71	7	22,7	6,9
Bryant	LA Lakers	PG	82	82	41,5	30
Fox	LA Lakers	?	76	75	28,7	9
Jamison	G State	C	67	0	19,4	8,8
Boykins	G State	PG	82	82	32,9	15,6
		Somme	664	504	283,6	151,5
		Somme dei quadrati	49300	36488	9479,62	3194,65

1. (2 punti) Osservando il grafico sotto riportato, ricavare il ruolo in cui gioca Fox e indicare qual è il RUOLO modale dei giocatori presi in considerazione dall'osservatore sportivo.



Dato mancante (1 punto)

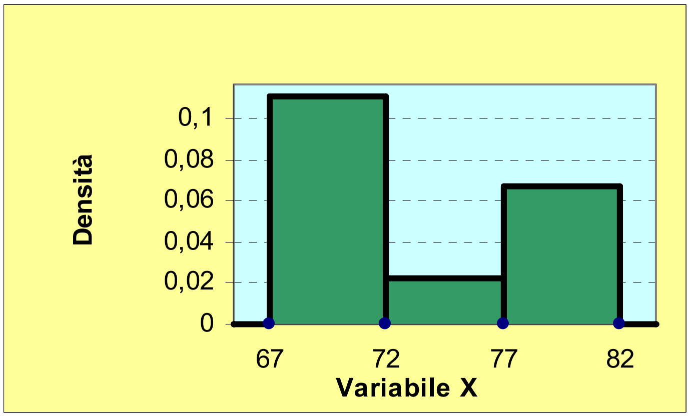
Ruolo(Fox) = **SF**

Moda di RUOLO (1 punto)

C (centro)

2. (4 punti) Rappresentare *graficamente* la variabile numero di partite giocate (PG), dopo averla riclassificata utilizzando tre classi di uguale ampiezza. Trovare inoltre la mediana *a partire dalla nuova classificazione della variabile*. Infine, dare una definizione formale di mediana nel caso generale.

Grafico (2 punti)



Valore della mediana (1 punto)

$$67 + 5 \cdot \frac{(0{,}5 - 0)}{0{,}5556} = \mathbf{71,\ 5}$$

Definizione di mediana (1 punto)

La mediana (Me) del carattere X è quel valore tale che:

$$\begin{cases} Fr(X \le Me) \ge 0{,}5 \\ Fr(X \ge Me) \ge 0{,}5 \end{cases}$$

3. (2 punti) Quale carattere tra PG (numero di partite giocate) e PI (numero di partite giocate dal primo minuto) è più variabile? Riportare anche i calcoli effettuati.

(Calcoli e commento, 2 punti)

Per confrontare la variabilità, devo utilizzare il coefficiente di variazione.

Media(PG)= 664/9 = 73,7778 Media(PI)= 504/9 = 56

Var(PG)= 5477,7778 – 5443,1605 = 34,6173 Var(PI)= 4054,222 – 3136 = 918,2222

Sqm(PG)= 5,8836 Sqm(PI)= 30,3022

Coeff. variaz.(PG)= 0,0797 Coeff. variaz.(PI)= 0,5411

PI è più variabile di PG, in quanto ha il coefficiente di variazione più elevato.

4. (3 punti)

Calcolare la frequenza relativa di giocatori che hanno effettuato un numero di partite dal primo minuto (variabile di riferimento nella tabella PI) comprese tra 26 e 86.

Frequenza (1 punto)

Fr(26 ≤ X ≤ 86)= 7/9 = **0,7778**

Un altro osservatore, invece, ha riportato, nel gruppo di atleti che sta osservando, una media di partite giocate dal primo minuto per cestista pari a 56, e una varianza pari a 100. Utilizzando Chebyshev, cosa si può dire della frequenza del medesimo intervallo (numero di partite comprese tra 26 e 86)?

Frequenza Chebyshev (2 punti)

Scriviamo l'intervallo in termini di distanza dalla media ed esprimiamo tale distanza tramite σ.

$Fr(26 \le X \le 86) = (56 - 30 \le X \le 56 + 30) = (\mu - 30 \le X \le \mu + 30) = (\mu - 3\sigma \le X \le \mu + 3\sigma) \ge$

$1 - \left(\frac{1}{3}\right)^2$ = **0,8889**

5. (3 punti) Sapendo che la somma dei prodotti tra MIN e PTS è pari a 5318,17, trovare la covarianza tra le due variabili. Inoltre, determinare ρ e commentarne il valore. Specificare la differenza tra Cov(MIN,PTS) e ρ.

Covarianza (1 punto)

Cov(MIN, PTS) = M(MIN · PTS) – M(MIN) M(PTS) = (5318,17)/9 – 31,5111 · 16,8333 = 590,9078 – 530,437= **60,4707**

ρ + commento (1 punto)

Var(MIN)= 1053,2911- 992,9501= 60,3410 Var(PTS)= 354,9611- 283,3611 =71,6

Sqm(MIN)= 7,7679 Sqm(PTS)= 8,4617

ρ= 60,4707/(7,7679 · 8,4617) = **0,9200**

Il valore di ρ rivela una fortissima relazione lineare positiva tra le due variabili.

Differenza covarianza e ρ (1 punto)

La covarianza è un indice assoluto di concordanza, ρ è un indice relativo che misura l'intensità del legame lineare.

6. (4 punti) Parliamo ora di indipendenza regressiva (indipendenza in media) tra due variabili.

a. Cosa significa in generale dire che W è regressivamente indipendente da Z? Darne una definizione.

(1 punto)

W è regressivamente indipendente da Z se tutte le medie delle distribuzioni subordinate W|Z=z sono uguali tra loro e uguali alla media di W.

b. Valutare se esiste indipendenza regressiva di PG da SQUADRA riportando i valori eventualmente calcolati.

(1 punto)

Media(PG|Squadra= Dallas)	72,6667
Media(PG|Squadra= LA Lakers)	74,25
Media(PG|Squadra= G State)	74,5

Dato che le medie subordinate non sono uguali tra loro, PG non è regressivamente indipendente da SQUADRA.

c. Pensando a due variabili quantitative, dire se le affermazioni seguenti sono vere o false.

- L'indipendenza regressiva implica indipendenza correlativa.

(1 punto)

Vero.

- L'indipendenza regressiva implica indipendenza statistica.

(1 punto)

Falso. E' invece vero il contrario.
[Per giustificarlo basta un esempio.]

X Y	1	2
0	0,3	0
1	0	0,4
2	0,3	0

7. (2 punti) Dalla Tabella presentata, si vede che: dei 3 giocatori del Dallas, 2 hanno una media punti inferiore a 20, mentre su 4 giocatori del LA Lakers, 2 conquistano, in media, meno di 20 punti a partita. Scelto a caso un giocatore del Dallas (tra i 3 presentati in tabella) *e* un giocatore del LA Lakers (tra i 4), in modo indipendente, costruire la funzione di probabilità della variabile aleatoria che descrive il numero di giocatori che hanno una media punti minore di 20 (sui 2 estratti).

Funzione di probabilità (2 punti)

Per costruire la funzione di probabilità, dobbiamo vedere che possibili valori assume la variabile casuale e associare ad ogni valore la sua probabilità. I possibili valori sono 0, 1, 2. E le probabilità associate si calcoleranno come (somme di) prodotti di probabilità.

X	Pr(X=x)
0	$\frac{1}{3}\cdot\frac{2}{4}=$ **0,1667**
1	$\frac{1}{3}\cdot\frac{2}{4}+\frac{2}{3}\cdot\frac{2}{4}=$ **0,5**
2	$\frac{2}{3}\cdot\frac{2}{4}=$ **0,3333**
	1

8. (2 punti) Calcolare valore atteso e varianza della funzione di probabilità appena costruita. (Se non si è costruita, usare la funzione di probabilità di una variabile aleatoria X che assume solo valori 0, 1, 2, con probabilità P(X=0) = 0,3; P(X=1) = 0,5; P(X=2) = 0,2).

Valore atteso (1 punto)

E(X)= $0{,}1667\cdot 0+0{,}5\cdot 1+0{,}3333\cdot 2=$ **1,1667**

[Se usati valori fittizi: E(X)= **0,9**]

Varianza (1 punto)

Var(X) = $E(X^2) - E(X)^2 = 1{,}8333 - (1{,}1667)^2 =$ **0,4722**

[se usati valori fittizi: Var(X)= $1{,}3 - (0{,}9)^2 =$ **0,49**]

9. (2 punti) Si hanno due eventi, A e B, definiti sullo stesso spazio Ω. Si sa che P(A)=0,5 e P(B)=0,8.

1. Quanto vale $P(A\cap B)$ nel caso in cui i due eventi siano indipendenti? Mostrare le relazioni usate.

(1 punto)

$P(A\cap B) = P(A)\cdot P(B) = 0{,}5\cdot 0{,}8 =$ **0,4**

2. A e B possono essere eventi incompatibili? Motivare la risposta.

(1 punto)

Se A e B fossero incompatibili, $P(A\cup B) = P(A) + P(B)$ (infatti, $A\cap B = \varnothing$ per definizione di incompatibilità, e $P(\varnothing) = 0$). In questo caso avremmo perciò che $P(A\cup B)= 1{,}3$. Ma questo non è possibile. Quindi, A e B **non** possono essere eventi incompatibili.

PRIMA PROVA INTERMEDIA DI STATISTICA

CLEA, CLEFIN (COD. 5047/4038/371/377)

3 Novembre 2004

Cognome ***Nome***

Numero di matricola

COMPITO B2

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

Un osservatore sportivo del basket ha raccolto alcuni dati relativi a **9** giocatori appartenenti a tre diverse squadre: Menphis, Phoenix e Seattle. Ecco la tabella e la spiegazione delle variabili utilizzate (*notare che sono anche state riportate le somme dei valori e le somme dei quadrati dei valori*).

Giocatore: nome del giocatore **Squadra**: squadra del giocatore

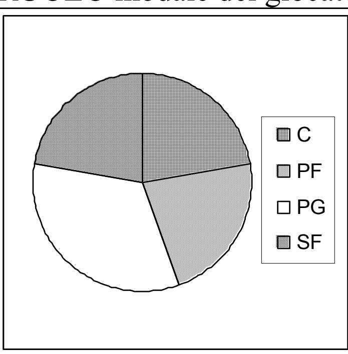
RUOLO: ruolo del giocatore (C= centro, SF = ala piccola, PF = ala grande, PG = playmaker)

PG: numero di partite giocate **PI**: partite giocate dal primo minuto

MIN: media dei minuti giocati per partita **PTS**: media stagionale dei punti del giocatore

Giocatore	Squadra	RUOLO	PG	PI	MIN	PTS
Wright	Menphis	C	70	50	28,3	11,4
Gasol	Menphis	PF	82	82	36	19
Williams	Menphis	PG	76	76	31,7	12,1
Marbury	Phoenix	PG	81	81	40	22,3
Marion	Phoenix	PF	81	81	41,6	21,2
Johnson	Phoenix	SF	82	34	27,5	9,8
Payton	Seattle	PG	76	72	40,8	20,8
Barry	Seattle	?	75	68	33,1	10,3
Allen	Seattle	SF	77	77	39,5	18,1
		Somme	700	621	318,5	145,0
		Somme dei quadrati	54576	45015	11509,09	2546,48

1. (2 punti) Osservando il grafico sotto riportato, ricavare il ruolo in cui gioca Barry e indicare qual è il RUOLO modale dei giocatori presi in considerazione dall'osservatore sportivo.



Dato mancante (1 punto)

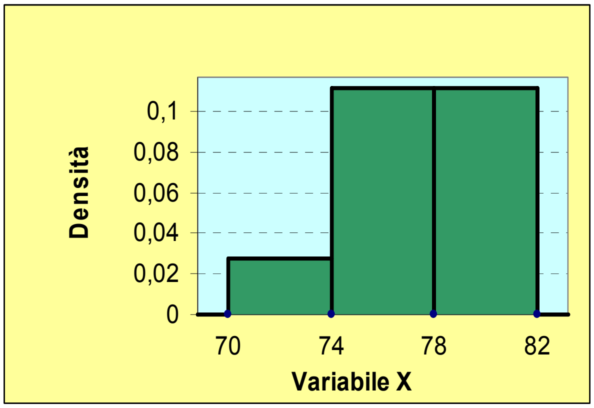
Ruolo(Barry) = **C**

Moda di RUOLO (1 punto)

PG (playmaker)

2. (4 punti) Rappresentare *graficamente* la variabile numero di partite giocate (PG), dopo averla riclassificata utilizzando queste classi: [70, 74); [74, 78); [78, 82]. Trovare inoltre la mediana *a partire dalla nuova classificazione della variabile*. Infine, dare una definizione formale di mediana nel caso generale.

Grafico (2 punti)



Valore di mediana (1 punto)

$$74 + 4 \cdot \frac{(0{,}5 - 0{,}1111)}{0{,}4444} = \mathbf{77{,}5}$$

Definizione di mediana (1 punto)

La mediana (Me) del carattere X è quel valore tale che:

$$\begin{cases} Fr(X \leq Me) \geq 0{,}5 \\ Fr(X \geq Me) \geq 0{,}5 \end{cases}$$

3. (2 punti) Quale carattere tra PG (numero di partite giocate) e PI (numero di partite giocate dal primo minuto) è più variabile? Riportare anche i calcoli effettuati.

(Calcoli e commento, 2 punti)

Per confrontare la variabilità, devo utilizzare il coefficiente di variazione.

Media(PG)= 700/9 = 77,7778	Media(PI)= 621/9 = 69
Var(PG)= 6064 – 6049,3827= 14,6173	Var(PI)= 5001,667 – 4761= 240,6667
Sqm(PG)= 3,8233	Sqm(PI)= 15,5134
Coeff. variaz.(PG)= 0,0492	Coeff. variaz.(PI)= 0,2248

PI è più variabile di PG, in quanto ha il coefficiente di variazione più elevato.

4. (3 punti)

Calcolare la frequenza relativa di giocatori che hanno effettuato un numero di partite dal primo minuto (variabile di riferimento nella tabella PI) comprese tra 49 e 89.

Frequenza (1 punto)

$Fr(49 \leq X \leq 89) = 8/9 =$ **0,8889**

Un altro osservatore, invece, ha riportato, nel gruppo di atleti che sta osservando, una media delle partite giocate dal primo minuto per cestista pari a 69 e una varianza pari a 25. Utilizzando Chebyshev, cosa si può dire della frequenza del medesimo intervallo (numero di partite comprese tra 49 e 89)?

Frequenza Chebyshev (2 punti)

Scriviamo l'intervallo in termini di distanza dalla media e esprimiamo tale distanza tramite σ.

$Fr(49 \leq X \leq 89) = (69 - 20 \leq X \leq 69 + 20) = (\mu - 20 \leq X \leq \mu + 20) = (\mu - 4\sigma \leq X \leq \mu + 4\sigma) \geq$

$1 - \left(\frac{1}{4}\right)^2$ = **0,9375**

5. (3 punti) Sapendo che la somma dei prodotti tra MIN e PTS è pari a 5338,13, trovare la covarianza tra le due variabili e commentare il valore trovato. Inoltre, determinare ρ e commentarne il valore. Specificare la differenza tra Cov(MIN,PTS) e ρ.

Covarianza (1 punto)

Cov(MIN, PTS) = M($\text{MIN} \cdot \text{PTS}$) – M(MIN) M(PTS) = (5338,13)/9 – 35,3889 · 16,1111 = 593,1256 – 570,1543= **22,9712**

ρ + commento (1 punto)

Var(MIN)= 1278,7878 – 1252,3734= 26,4144 Var(PTS)= 282,9422 – 259,5679 =23,3743

Sqm(MIN)= 5,1395 Sqm(PTS)= 4,8347

ρ= 22,9712/(5,1395 · 4,8347) = **0,9245**

Il valore di ρ rivela una fortissima relazione lineare tra le due variabili.

Differenza covarianza e ρ (1 punto)

La covarianza è un indice assoluto di concordanza, ρ è un indice relativo che misura l'intensità del legame lineare.

6. (4 punti) Parliamo ora di indipendenza regressiva (indipendenza in media) tra due variabili.

a. Cosa significa in generale dire che W è regressivamente indipendente da Z? Darne una definizione.

(1 punto)

W è regressivamente indipendente da Z se tutte le medie delle distribuzioni subordinate W|Z=z sono uguali tra loro e uguali alla media di W.

b. Valutare se esiste indipendenza regressiva di PG da SQUADRA riportando i valori eventualmente calcolati.

(1 punto)

Media(PG|Squadra= Menphis)	76
Media(PG|Squadra= Phoenix)	81,3333
Media(PG|Squadra= Seattle)	76

Dato che le medie subordinate non sono tutte uguali tra loro, PG non è regressivamente indipendente da SQUADRA.

c. Pensando a due variabili quantitative, dire se le affermazioni seguenti sono vere o false.

- L'indipendenza correlativa implica indipendenza regressiva.

(1 punto)

Falso. E' invece vero il contrario.
[Per giustificarlo basta un esempio.]

X Y	1	2	3	4
-1	0,1	0	0,1	0
0	0,1	0,2	0,1	0,1
1	0	0,2	0	0,1

- L'indipendenza statistica implica indipendenza regressiva.

(1 punto)

Vero.

7. (2 punti) Dalla Tabella presentata, si vede che: dei 3 giocatori del Phoenix, 2 hanno una media punti superiore a 20, mentre su 3 giocatori del Seattle, 1 solo conquista, in media, almeno 20 punti a partita. Scelto a caso un giocatore del Phoenix (tra i 3 presentati in tabella) *e* un giocatore del Seattle (tra i 3), in modo indipendente, costruire la funzione di probabilità della variabile aleatoria che descrive il numero di giocatori che hanno una media punti maggiore di 20 (sui 2 estratti).

Funzione di probabilità (2 punti)

Per costruire la funzione di probabilità, dobbiamo vedere che possibili valori assume la variabile casuale e associare ad ogni valore la sua probabilità. I possibili valori sono 0, 1, 2. E le probabilità associate si calcoleranno come (somme di) prodotti di probabilità.

X	Pr(X=x)
0	$\frac{1}{3}\cdot\frac{2}{3}=$ **0,2222**
1	$\frac{1}{3}\cdot\frac{1}{3}+\frac{2}{3}\cdot\frac{2}{3}=\frac{5}{9}=$ **0,5556**
2	$\frac{2}{3}\cdot\frac{1}{3}=$ **0,2222**
	1

8. (2 punti) Calcolare valore atteso e varianza della funzione di probabilità appena costruita. (Se non si è costruita, usare la funzione di probabilità di una variabile aleatoria X che assume solo valori 0, 1, 2, con probabilità P(X=0) = 0,2; P(X=1) = 0,5; P(X=2) = 0,3).

Valore atteso (1 punto)

E(X)= $0{,}2222\cdot 0+0{,}5556\cdot 1+0{,}2222\cdot 2=1$

[Se usati valori fittizi: E(X)= **1,1**]

Varianza (1 punto)

Var(X) = $E(X^2) - E(X)^2 = 1{,}4444 - (1)^2 =$ **0,4444**

[se usati valori fittizi: Var(X)= $1{,}7 - (1{,}1)^2 =$ **0,49**]

9. (2 punti) Si hanno due eventi, A e B, definiti sullo stesso spazio Ω. Si sa che P(A)=0,5 e P(B)=0,8.

1. Quanto vale $P(A\cap B)$ nel caso in cui $P(A\cup B)=1$? Mostrare le relazioni usate.

(1 punto)

$P(A\cup B) = P(A) + P(B) - P(A\cap B)$. In questo caso abbiamo quindi $1= 0{,}5 + 0{,}8 - P(A\cap B)$.

Perciò, $P(A\cap B)=$ **0,3**.

2. A e B possono essere eventi complementari? Motivare la risposta.

(1 punto)

Se A e B fossero complementari, avremmo che $A\cup B = \Omega$ e nello stesso tempo $A\cap B = \varnothing$. Ma $P(\varnothing) = 0$, mentre noi abbiamo $P(A\cap B)=0{,}3$ (diverso da 0). Quindi A e B **non** possono essere eventi **complementari**!