PROVA SCRITTA DI STATISTICA

CLEA/CLEFIN/CLEMIT (cod. 5047/4038/371/377)

3 Novembre 2004

MOD. A

Esercizio N. 1 (3 punti).

Data la v.s. X avente funzione di densità:

$$f(x) = \begin{cases} 1/20 & 0 \le x < 1 \\ 9/40 & 1 \le x < 3 \\ 1/10 & 3 \le x < 7 \\ 1/20 & 7 \le x < 9 \\ 0 & \text{altrove} \end{cases}$$

a) Determinare la distribuzione di frequenza.
b) Calcolarne media, moda e mediana.

Esercizio N. 2 (2 punti).

Enunciare la disuguaglianza di Chebyshev e verificarla per la v.a.

$$X = \begin{cases} 0 & 4 & 5 & 6 & 10 \\ \\ 0.05 & 0.1 & 0.7 & 0.1 & 0.05 \end{cases}$$

ponendo $\lambda = 2$. Riportare dettagliatamente tutti i passaggi.

Esercizio N. 3 (3 punti).

E' data una v.s. a due dimensioni (X,Y) con E(X) = 5, E(Y) = 4, Var(X) = 36, Var(Y) = 9, COV(X,Y) = 18.
a) Quanto vale il coefficiente quadratico medio relativo di contingenza $\tilde{\varphi}^2$? (Suggerimento: calcolare prima il coefficiente di correlazione lineare ρ)
b) Determinare l'equazione della retta di regressione di Y su X.
c) Prevedere il valore assunto da Y quando X = 8.

Esercizio N. 4 (4 punti).

Dati i numeri aleatori indipendenti X ed Y distribuiti secondo la legge bernoulliana con parametri rispettivi $p_X = 0.2$ e $p_Y = 0.3$
a) Trovare la distribuzione del numero aleatorio Z = X – Y.
b) Determinare E(Z) e Var(Z) e scrivere le relazioni di E(X) ed E(Y) con E(Z) e di Var(X) e Var(Y) con Var(Z).

Esercizio N. 5 (4 punti).

Un'urna contiene 4 palline bianche e 6 rosse. Si estraggono 3 palline.
a) Verificare se è più probabile estrarre 3 palline bianche con la tecnica di estrazione con reimmissione oppure con quella di estrazione senza reimmissione.
b) Trovare la probabilità che le palline bianche siano più delle palline rosse, nel caso di reimmissione.
c) Trovare la probabilità che fra le palline estratte ce ne sia almeno una rossa, nel caso di reimmissione.

Esercizio N. 6 (4 punti).

Estraiamo dalla v.a. X dell'esercizio n. 2 un campione bernoulliano di 100 elementi
a) Trovare media e varianza della variabile media campionaria.
b) Determinare inoltre la probabilità approssimata che la media campionaria cada nell'intervallo (5 , 5.2) .
c) E' possibile che la media campionaria superi il valore 15? Giustificare la risposta.

Esercizio N. 7 (2 punti).

Data una v.a. normale con varianza nota pari a 16 si estrae un campione bernoulliano di n elementi. L'intervallo di confidenza per la media, determinato con un livello di confidenza del 95%, ha lunghezza 3.1360. Determinare n.

Esercizio N. 8 (5 punti).

Per testare l'efficacia di un farmaco antinfluenzale questo viene somministrato ad un campione di 100 pazienti. Di questi 80 guariscono entro la prima settimana di cure. Sia p la percentuale di soggetti che potrebbero guarire entro la prima settimana nell'intera popolazione.
a) Definire una regione critica per verificare l'ipotesi $H_0 : p \geq 0.9$ contro l'alternativa $H_1 : p < 0.9$. al livello $\alpha = 0.01$
b) Dare la definizione di p-value.
c) Determinare il p-value relativo alla realizzazione campionaria osservata.
d) E' ragionevole accettare l'ipotesi H_0 ? Giustificare la decisione.

PROVA SCRITTA DI STATISTICA

CLEA/CLEFIN/CLEMIT (cod. 5047/4038/371/377)

3 Novembre 2004

MOD. B

Esercizio N. 1 (6 punti)

Si è fatta un'indagine per stabilire l'efficacia di messaggi pubblicitari su un quotidiano per favorire la vendita di un certo prodotto nei supermercati di una città, per 20 settimane consecutive.
Sia X il numero di messaggi pubblicati in una settimana ed Y l'incasso, in migliaia di euro, per i prodotti venduti nello stesso periodo, riportati nella tabella seguente.

Y \ X	0	1	2	3	4	
[0 , 10]	**1**	**1**				**2**
(10,20]		**1**	**1**			**2**
(20,30]		**1**	**3**	**1**		**5**
(30,40]			**1**	**4**		**5**
(40,50]				**1**	**5**	**6**
	1	**3**	**5**	**6**	**5**	**20**

a) Rappresentare graficamente la distribuzione di frequenze di Y.
b) Calcolare media e moda di Y.
c) Trovare la funzione di regressione di Y su X.
d) Trovare mediana e varianza della distribuzione di Y condizionata ad X = 1.
e) Dare la definizione di indipendenza statistica e di indipendenza regressiva.

Esercizio N. 2 (3 punti)

Nella seguente tabella a doppia entrata

Y \ X	0	1	
0			0.6
1			0.4
	0.4	0.6	1.0

si conosce il coefficiente di correlazione lineare $\rho = -1$.

a) Determinare il valore del coefficiente quadratico medio di contingenza φ^2 .
b) Riportare sul foglio la tabella completata con le frequenze congiunte.

Esercizio N. 3 (3 punti)

Un'urna contiene 6 palline bianche, 3 nere ed una pallina rossa.
Si estrae una pallina. Se questa è bianca o nera il gioco termina. Se questa è rossa essa viene rimessa nell'urna e viene estratta un'altra pallina (In ogni caso non si procede ad altra estrazione).

a) Definire lo spazio Ω degli eventi elementari e calcolare la probabilità di ogni evento.
b) Se ogni volta che si estrae una pallina bianca si guadagna 1 euro e per ogni pallina nera si perde un euro, calcolare il guadagno atteso al termine del gioco.

Esercizio N. 4 (3 punti)

Sia X una v.a. discreta tale che:
$Pr(X = 0) = q$
$Pr(X = 1) = 3q$
$Pr(X \neq 0,1) = 0$

a) Determinare media e varianza di X.
b) Data $S = X_1 + X_2 + \ldots + X_{50}$ somma di 50 variabili indipendenti tutte uguali ad X e $T = 4S + 2$ calcolare la media e la varianza di T.

Esercizio N. 5 (7 punti)

Sia X una v.a. normale con media μ e varianza 16. Si estrae da X un campione bernoulliano di n elementi.

a) Determinare il minimo valore di n tale che la probabilità che la media campionaria assuma valori compresi fra μ - 5 e μ + 5 sia pari almeno al 90%.

N.B. Se non si è risolto il punto a) porre n = 20 negli esercizi b) c) d).

b) Avendo trovato che la media del campione definito in precedenza è pari a 48 determinare l'intervallo di confidenza di μ al livello α = 0.05.
c) Indicare la regione di rifiuto dell'ipotesi $H_0 : \mu = 50$ contro l'alternativa $H_1 : \mu \neq 50$ fissato α = 0.05.
d) Stabilire, sempre sapendo che $\bar{x} = 48$ e che α =0.05, se possiamo rifiutare l'ipotesi H_0 giustificando la decisione presa.

Esercizio N. 6 (5 punti)

E' stata fatta una rilevazione su 10 nuclei famigliari, in particolare sui redditi x e sul numero Y di figli, ottenendo i risultati seguenti:

$$\sum_{i=1}^{10} x_i = 250 \qquad \sum_{i=1}^{10} Y_i = 26 \qquad \sum_{i=1}^{10} x_i Y_i = 900 \qquad \sum_{i=1}^{10} x^2{}_i = 9000 \qquad \sum_{i=1}^{10} Y^2{}_i = 80$$

a) Determinare i parametri del modello lineare di Y in x.
b) Conoscendo la somma dei quadrati degli errori $\sum_{i=1}^{10} \hat{\varepsilon}_i^{\,2} = 5.03$ fornire una stima corretta della varianza di ε_i.
c) Utilizzando ancora l'informazione data al punto b) verificare l'ipotesi $H_0 : \beta_1 = 0$ contro l'alternativa $H_0 : \beta_1 \neq 0$ fissato α = 0.05 (se non avete svolto il punto b) scrivere la regione di rifiuto dell'ipotesi $H_0 : \beta_1 = 0$ contro l'alternativa $H_0 : \beta_1 \neq 0$ fissato α = 0.05).
d) Enunciare le ipotesi del modello lineare distinguendo quelle deboli e quelle forti.