

PROVA SCRITTA DI STATISTICA
(COD. 5047 - COD. 4038 - 371-377)

7 luglio 2005

APPROSSIMARE TUTTI I CALCOLI ALLA QUARTA CIFRA DECIMALE

MODALITÀ A

Esercizio 1 (9 punti)

Supponiamo di aver osservato la seguente distribuzione doppia (frequenze assolute congiunte) relativa ai caratteri $X = \text{voto di diploma}$ e $Y = \text{voto medio esami sostenuti}$, su un collettivo di studenti.

Y	[18, 25)	[25, 27)	[27, 30]	
X				
[70, 80)	10	3	17	
[80, 90)	13	10	27	
[90, 100]	3	7	10	

- a) Ricavare la distribuzione marginale di $Y = \text{voto medio}$.
- b) Rappresentarla attraverso l'opportuna rappresentazione grafica.
- c) Calcolare la media di Y .
- d) I due caratteri X e Y sono indipendenti?
- e) È possibile calcolare un indice di correlazione per i caratteri in questione? Se sì, calcolarlo e commentare il valore ottenuto.
- f) A questo punto è utile calcolare anche l'indice di connessione relativo $\tilde{\varphi}^2$ per tali dati? Sì, no, perchè?

Esercizio 2 (3 punti)

Si consideri il seguente esperimento aleatorio: si lanciano 2 monete regolari, e si osservano le facce risultanti. Se compare lo stesso simbolo su entrambe le facce, si lancia una terza moneta regolare, altrimenti l'esperimento è concluso.

- a) Costruire lo spazio Ω dei possibili risultati dell'esperimento e calcolare la probabilità di ciascun evento elementare.
- b) Costruire la distribuzione di probabilità della variabile aleatoria $X = \text{Numero di teste osservate al termine dell'esperimento}$.

Esercizio 3 (3 punti)

Assumiamo che il punteggio X di un test d'ammissione ad un corso di dottorato di ricerca in una prestigiosa università italiana si distribuisca come una v.a. X Uniforme continua nell'intervallo $[36, 60]$.

- Calcolare il valore atteso di X .
- Qual è la probabilità che X risulti maggiore di 57?
- Uno studente viene ammesso soltanto nel caso raggiunga un punteggio superiore a 57. Si presenta alla selezione un campione casuale semplice di 10 studenti. Qual è la probabilità che almeno due studenti vengano ammessi?

Esercizio 4 (2 punti)

Siano $(x_1, x_2, \dots, x_{64})$ i risultati riportati da un campione casuale semplice di 64 aspiranti presentatori ad un test di selezione per una trasmissione televisiva. Il risultato del test, per ogni partecipante alla selezione, è una v.a. X_i con distribuzione incognita di media $\mu = 20$ e varianza $\sigma^2 = 256$.

- Determinare il valore atteso e la varianza della v.a. $\bar{X}_{64} = \text{punteggio medio}$
- Qual è la probabilità (approssimata) che \bar{X}_{64} risulti inferiore a 22.6?

Esercizio 5 (4 punti)

Estratto un campione casuale semplice di ampiezza $n = 25$ da una popolazione Normale, di media e varianza entrambe incognite, si ottiene $\sum_{i=1}^{25} x_i = 480$, $\sum_{i=1}^{25} x_i^2 = 10400$. Si vuole testare l'ipotesi nulla $H_0 : \mu \leq 15$ contro l'ipotesi alternativa $H_1 : \mu > 15$.

- Per un generico livello α dell'errore di prima specie, riportare l'espressione analitica della regione di rifiuto del test.
- Qual è il p -value relativo alla realizzazione campionaria assegnata? Se α fosse fissato pari a 0.01 rifiutereste o no l'ipotesi nulla?

Esercizio 6 (2 punti)

Siano T_n e U_n due stimatori per un parametro incognito θ di una popolazione. Assumiamo che $E(T_n) = \theta$, ovvero che T_n sia non distorto per θ , e che $Var_\theta(U_n) \leq Var_\theta(T_n)$, per ogni θ .

- È possibile stabilire quale stimatore sia più efficiente? Se sì, fornire la risposta.
- Se fosse anche $E(U_n) = \theta$ la risposta precedente cambierebbe? Se sì, come?

Esercizio 7 (2 punti)

Sia (X_1, \dots, X_n) un campione casuale semplice estratto da una popolazione Normale $N(\mu, 9)$. Assumiamo che per il campione osservato risulti $\sum_{i=1}^n x_i = 4680$.

- Determinare l'ampiezza campionaria minima necessaria affinché l'intervallo di confidenza a livello $(1 - \alpha) = 0.90$ risulti di lunghezza inferiore a 0.5.
- Calcolare esplicitamente l'intervallo di confidenza in questione utilizzando il valore di n trovato al punto a). (Se non si è risolto il punto a) assumere $n = 36$).

Esercizio 8 (2 punti)

Riportare l'espressione analitica dell'indice per la valutazione della capacità di adattamento del modello lineare, noto come *coefficiente di determinazione* esplicitando tutte le quantità coinvolte.