Prova scritta

Appello 06.06.2023

Tempo a disposizione: 120 minuti

La soluzione di ogni esercizio deve essere scritta in maniera pulita e leggibile sugli appositi fogli contrassegnati forniti a inizio prova, spiegando dettagliatamente il ragionamento con argomenti matematicamente validi. I risultati numerici devono essere dati con 4 cifre significative. Punti potranno essere sottratti per inadempienza a questi criteri. Esercizi consegnati su fogli esterni non saranno ritenuti validi. Sono benvenute le frasi che forniscono una spiegazione euristica della strategia che si intende seguire, ma non sono sufficienti, ci si aspetta calcoli espliciti. È permesso l'uso di calcolatrici non programmabili, e di **un foglio formato A4** (scritto fronte-retro, anche al computer) contenente appunti/formule. Un foglio supplementare è permesso per le tavole necessarie.

1. Si consideri un mazzo di carte italiane tradizionali composto da 40 carte (4 semi: denari, coppe, spade, bastoni; e 10 numeri: dall'1 al 7, e tre figure: fante (F), cavallo (C), re (R)).

   (a) Supponiamo di estrarre due carte a caso dal mazzo senza reinserimento. Siano X_1, X_2 il valore della prima e della seconda carta pescata, rispettivamente (le figure valgono 10). Queste variabili aleatorie sono indipendenti?

   a) si b) no c) dati insuff. d) correlazione positiva e)Altro: ______

   (b) Supponiamo di estrarre due carte a caso dal mazzo senza reinserimento. Qual è la probabilità che la loro somma faccia esattamente 4 (le figure valgono 10)?

   a) 0.01 b) 0.018 c) 0.02 d) 0.028 e)Altro: ______

   (c) Supponiamo ora di pescare con reinserimento. Quale è la probabilità di pescare almeno 3 carte di coppe nei primi 4 tentativi?

   a) 0.262 b) 0.051 c) 0.046 d) 0.291 e)Altro: ______

   (d) Sempre pescando con reinserimento, sia X il numero di estrazioni necessarie per osservare il primo fante dopo il primo 4.[1] Calcolare $\mathbb{E}(X)$ e $\mathrm{Var}(X)$.

   Risposta: ________________

2. Supponiamo di essere interessati a confrontare la durata di due tipi di batterie alcaline di marche diverse, chiamate "Marca A" e "Marca B". Per fare ciò, selezioniamo casualmente 7 batterie di ciascuna marca e misuriamo la durata delle batterie in ore. I risultati ottenuti sono i seguenti:

Marca A:	720,	730,	740,	710,	705,	715,	725
Marca B:	708,	703,	693,	718,	713,	698	

   Assumiamo che la durata delle batterie di entrambe le marche segua una distribuzione normale, con medie μ_A e μ_B sconosciute e varianze $\sigma_A^2 = 121$ e $\sigma_B^2 = 81$ conosciute.

   (a) Calcola l'intervallo di confidenza bilaterale al 95% per la media μ_A.

   Risposta: ________________

   (b) Definire l'ipotesi nulla $\mathcal{H}_0$ e l'ipotesi alternativa $\mathcal{H}_1$ per il test che ha come obiettivo quello di verificare che c'è una differenza significativa nella durata media delle batterie tra le due marche.

   Risposta: ________________

[1] Per esempio, pescando $1, R, 4, 7, F, \ldots$ abbiamo $X = 5$, mentre per $F, 3, 6, C, 4, 4, F, \ldots$ abbiamo $X = 7$

(c) Testare l'ipotesi nulla formulata al punto precedente al livello di significatività del 5%. Qual è la conclusione che possiamo trarre dal risultato del test?

Risposta: ________________

(d) Calcola approssimativamente il valore p per il test discusso ai punti b) e c).

Risposta: ________________

3. Un call center riceve chiamate da due regioni, che chiamiamo A e B. Il numero di chiamate al minuto dalla regione A ha distribuzione di Poisson di parametro $\lambda_A = 2$, mentre il numero di chiamate al minuto dalla regione B ha distribuzione di Poisson di parametro $\lambda_B = 3$. Supponiamo che il numero di chiamate da A e il numero di chiamate da B siano indipendenti.

(a) Calcolare la probabilità che, in un dato minuto, il call center riceva almento 2 chiamate da A.

a) 0.594 b) 0.323 c) 0.271 d) 0.729 e)Altro: ______

(b) Calcolare la probabilità che, in un dato minuto, il call center riceva almeno 2 chiamate *in totale*.

Risposta: ________________

(c) Calcolare la probabilità che se sono arrivate esattamente due chiamate al call center durante l'ultimo minuto, entrambe siano arrivate dalla regione B.

Risposta: ________________

(d) Supponiamo che il numero di chiamate dalla regione A in un dato minuto sia indipendente dal numero di chiamate da A in altri minuti. Calcolare (almeno in modo approssimato) la probabilità che, in un dato intervallo di 2 ore, il call center riceva almeno 260 chiamate da A.

Risposta: ________________

4. Si consideri la seguente funzione,

$$f_\theta(x) = \frac{2}{\theta} x \exp\left(-\frac{x^2}{\theta}\right) \mathbb{1}_{(0,\infty)}(x),$$

dipendente dal parametro $\theta > 0$.

(a) Si mostri che, per ogni $\theta > 0$, f_θ è una densità di probabilità di una variabile aleatoria X, e si calcoli $\mathbb{P}(X \leq x)$ in funzione di x. **Suggerimento:** Sostituzione $x^2/\theta = u$

Risposta: ________________

(b) Dato un campione $\{X_1, \ldots, X_n\}$ estratto iid da $f_\theta(x)$, calcolare lo stimatore di massima verosimiglianza del parametro θ.

Risposta: ________________

(c) Si tratta di uno stimatore corretto?

Risposta: ________________

5. Consideriamo l'integrale

$$\int_{-3}^{5} \frac{\log(7+e^x)}{(2+\sin(x))}\, dx$$

(a) Discutere brevemente il metodo di integrazione di Monte Carlo e come può essere applicato per l'approssimazione di questo integrale.

(b) Scrivere i comandi in R la cui esecuzione restituisce, per il metodo discusso sopra, un valore approssimativo di questo integrale.
Nota: potrebbero essere utili i comandi `sin(·)` (il seno), `log(·)` (il logaritmo naturale) e `exp(·)` (la funzione esponenziale).