Prova scritta (Appello 25.06.2024)

Tempo a disposizione: 120 minuti

Nome, Cognome e #matricola: ____________________

☐ Vecchio ordinamento (senza algebra lineare) ☐ Nuovo ordinamento (con algebra lineare)

Riportare nome e cognome su ogni foglio consegnato. 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. Il risultato di ogni esercizio deve essere riportato nell'apposito spazio su questo foglio (o in forma di risposta o in forma di crocetta), in caso contrario l'esercizio non verrà corretto. Per le domande a crocette, è necessario dare anche una spiegazione della propria risposta (calcoli o argomento) sul foglio a parte. In assenza di tale spiegazione la risposta viene considerata errata. 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 una scatola contenente, per ogni numero tra 1 e 10, tre biglie, una di colore rosso, una di colore bianco e una di colore verde[1].

   (a) Supponiamo di estrarre due biglie a caso dalla scatola senza reinserimento. Siano X_1, X_2 il numero sulla prima e sulla seconda biglia pescata, rispettivamente. Queste variabili aleatorie sono indipendenti?

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

   (b) Supponiamo di estrarre due biglie dalla scatola senza reinserimento. Qual è la probabilità che la somma dei numeri sulle biglie estratte faccia esattamente 18?

   a) 0.01 b) 0.017 c) 0.019 d) 0.028 e)Altro: ______

   (c) Supponiamo ora di pescare con reinserimento. Quale è la probabilità di pescare almeno 4 biglie rosse nei primi 5 tentativi?

   a) 0.8 b) 0.041 c) 0.045 d) 0.165 e)Altro: ______

   (d) Supponiamo ora di pescare 100 volte con reinserimento. Sia X_i il numero indicato sull'i-esima biglia pescata. Calcolare $\mathbb{E}(X_1 + 3 \cdot X_2)$ e $\mathrm{Var}(2 \cdot X_1 + 9)$.

   Risposta: __________________

   (e) Calcolare, usando un'approssimazione normale, la probabilità che la somma $X_1 + X_2 + \cdots + X_{100}$ sia inferiore a 500.

   Risposta: __________________

2. Una ditta vuole confrontare la performance di due dei suoi negozi situati in due città diverse (Parigi e Londra) sulla base delle vendite giornaliere. I totali delle vendite giornaliere (in euro) dei due negozi sull'arco di una settimana lavorativa sono i seguenti.

Parigi :	2659.50	1882.50	2165.00	1949.00	2453.50	2050.00
Londra :	2963.00	2685.00	3083.50	2151.50	2466.50	2020.00

   Si assuma che in entrambi i casi la distribuzione delle vendite giornaliere sia normale (Parigi $\mathcal{N}(\mu_1, \sigma_1^2)$, Londra $\mathcal{N}(\mu_2, \sigma_2^2)$).

   (a) Calcolare la media campionaria e la deviazione standard campionaria delle vendite giornaliere di entrambi i negozi.

   Risposta: __________________

[1] quindi: una biglia rossa con il numero 1, una biglia bianca con il numero 1, una biglia verde con il numero 1, una biglia rossa con il numero 2, e così via.

(b) Assumiamo che σ_1 sia sconosciuta. Si calcoli l'intervallo di confidenza bilaterale al 90% per lo stimatore di μ_1 dato dalla media empirica.

Risposta: ___________________

(c) Si definiscano l'ipotesi nulla $\mathcal{H}_0$ e l'ipotesi alternativa $\mathcal{H}_1$ per il test che ha come obbiettivo quello di dimostrare che il negozio di Londra ha una performance migliore di quello di Parigi.

Risposta: ___________________

(d) Assumiamo ora di conoscere $\sigma_1 = 300$, $\sigma_2 = 400$. Si testi l'ipotesi nulla formulata al punto precendente al livello di significatività del 5%. Che tipo di conclusione ci permette di trarre il risultato del test?

Risposta: ___________________

3. Un test diagnostico ha sensibilità (probabilità che il test sia negativo se la persona che fa il test è sana) di 0.97 e specificità (probabilità che il test sia positivo se la persona che fa il test è malata) di 0.99. Per ridurre la probabilità di errore, il test viene eseguito 3 volte sulla stessa persona. Supponiamo che, fissata la persona su cui viene eseguito il test, gli esiti del test siano indipendenti. Le persone malate sono il 2% della popolazione. Si estrae una persona a caso nella popolazione e si esegue il test 3 volte su tale persona.

   (a) Calcolare la probabilità che il primo test sia positivo.

   Risposta: ___________________

   (b) Calcolare la probabilità che, se la persona è sana, il test sia positivo in almeno due esecuzioni.

   Risposta: ___________________

   (c) Se il test è positivo in almeno due esecuzioni, calcolare la probabilità che la persona sia malata.

   Risposta: ___________________

4. Definiamo X come una variabile aleatoria discreta con funzione di massa con parametro $\theta \in (0,1)$ definita come

$$p_X(k|\theta) = \theta^k(1-\theta)^{1-k} \qquad \text{per } k \in \{0,1\}$$

e che vale 0 per qualsiasi altro $k \in \mathbb{N}$.

   (a) Si verifichi che f è una funzione di massa per ogni $\theta \in (0,1)$

   Risposta: ___________________

   (b) Si determini lo stimatore di massima verosimiglianza $\hat{\theta}$ di θ per un campione $\{x_1, \ldots, x_n\}$

   Risposta: ___________________

   (c) Si tratta di uno stimatore corretto e/o consistente?

   Risposta: ___________________

5. Consideriamo l'integrale

$$\int_{-7}^{10} \frac{\sin(7 + \log(x^2+1))}{2 + e^x}\, dx$$

   (a) Discutere in generale il metodo di integrazione di Monte Carlo (il suo fondamento teorico e la sua implementazione) e come può essere applicato per l'approssimazione di questo integrale in particolare.

   Risposta su foglio a parte: ☐

   (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 sin), `log(·)` (il logaritmo naturale) e `exp(·)` (la funzione esponenziale).

   Risposta su foglio a parte: ☐