

Informazione e stima – 02/09/2023

- Ogni esercizio viene valutato da 0 a 6 punti
 - Verranno valutate solo le parti scritte in penna – non usare correttori
 - Non riportare solo il risultato, ma cerca di argomentare sinteticamente la risposta
 - Riportare il proprio nome, cognome, e codice persona su ogni foglio consegnato
 - Esercizio da escludere dal punteggio finale:
-
- ① Supponiamo di avere una scatola con 8 palline rosse, 5 palline verdi, e 7 palline blu. Estraiamo casualmente 3 palline senza reinserimento. Qual è la probabilità che ci siano almeno 2 palline rosse tra quelle estratte?
 - ② Siano $X \sim \text{Exp}(2)$ e $Y = aX + Z$ con $Z \sim \mathcal{N}(0, 1)$ e indipendente da X .
 - (a) Si calcoli $\rho[X, Y]$.
 - (b) Si dica se è possibile avere $\rho[X, Y] = 1$ per qualche valore di a . Giustificare la risposta.
 - ③ Sia $X \sim \text{Exp}(\lambda)$ e $Y = \sqrt{X}$. Trovare la legge di Y .
 - ④ In una fabbrica, il processo di produzione di un certo componente è descritto da un processo di Poisson con un tasso di produzione medio di 2 componenti all'ora. Supponendo che la probabilità che un componente sia difettoso sia 0.1, indipendentemente dagli altri componenti
 - (a) qual è la probabilità che il primo componente difettoso sia prodotto non prima di 3 ore?
 - (b) Mediamente in quante ore saranno prodotti 10 componenti difettosi?
 - ⑤ Sia $X_n = \max\{L_1, L_2, \dots, L_n\}$ il massimo tra n lanci indipendenti di un dado a 6 facce ben bilanciato. Determinare se la successione di v.a. $\{X_n\}_n$ converge in probabilità e, se sì, a quale valore.
 - ⑥ Avendo a disposizione un generatore di campioni uniformi $U \sim \mathcal{U}[0, 1]$, descrivere un algoritmo (con pseudocodice) che con la massima efficienza possibile generi campioni distribuiti con legge $f_X(x) = \frac{1}{x^2}e^{-\frac{1}{x}}$ per $x > 0$ e $f_X(x) = 0$ altrimenti.

Soluzioni

Problema 1

Lo spazio di probabilità è uniforme discreto. La probabilità di avere almeno 2 palline rosse è la probabilità di averne esattamente 2 rosse oppure esattamente 3 rosse. Entrambe queste probabilità sono ipergeometriche e sommandole si trova il risultato finale:

$$p = \frac{\binom{8}{2}\binom{12}{1} + \binom{8}{3}\binom{12}{0}}{\binom{20}{3}} = \frac{336 + 56}{1140} = \frac{392}{1140} \approx 0.3439. \quad (1)$$

Problema 2

1. Il coefficiente di correlazione lineare è

$$\rho[X, Y] = \frac{\text{Cov}[X, Y]}{\sigma_X \sigma_Y} \quad (2)$$

dove

$$\sigma_X^2 = \text{Var}[X] = \frac{2}{2^2} - \frac{1}{2^2} = \frac{1}{4} \quad (3)$$

$$\sigma_Y^2 = \text{Var}[Y] = \text{Var}[aX] + \text{Var}[Z] = \frac{a^2}{4} + 1 = \frac{a^2 + 4}{4} \quad (4)$$

$$\text{Cov}[X, Y] = \text{E}[XY] - \text{E}[X]\text{E}[Y] \quad (5)$$

$$= \text{E}[X(aX + Z)] - \text{E}[X]\text{E}[aX + Z] \quad (6)$$

$$= a\text{E}[X^2] + \text{E}[XZ] - a\text{E}[X]^2 - \text{E}[X]\text{E}[Z] \quad (7)$$

$$= a\text{Var}[X] \quad (8)$$

$$= \frac{a}{4}. \quad (9)$$

Il risultato finale è

$$\rho[X, Y] = \frac{\frac{a}{4}}{\frac{\sqrt{a^2+4}}{4}} = \frac{a}{\sqrt{a^2+4}}. \quad (10)$$

2. Il coefficiente di correlazione lineare non può essere pari a 1 per nessun valore di a , perché, a causa della presenza della variabile indipendente Z , non ci potrà mai essere un legame lineare deterministico tra X e Y .

Problema 3

Innanzitutto si noti che $Y \geq 0$. Applicando il metodo della cumulata si ha che

$$F_Y(y) = \Pr(Y \leq y) = \Pr(\sqrt{X} \leq y) \quad (11)$$

$$= \Pr(X \leq y^2) \quad (12)$$

$$= F_X(y^2), \quad y \geq 0. \quad (13)$$

Calcolando la derivata, si ottiene

$$f_Y(y) = \frac{d}{dy} F_Y(y) = 2y f_X(y^2) = 2y \lambda e^{-\lambda y^2}, \quad y \geq 0, \quad (14)$$

e $f_Y(y) = 0$ altrimenti.

Problema 4

Siccome i componenti sono difettosi in maniera indipendente con la stessa probabilità 0.1, il processo dei componenti difettosi è di Poisson con tasso medio di $2 \cdot 0.1 = 0.2$ componenti difettosi all'ora.

1. L'evento di interesse è equivalente all'avere 0 componenti difettosi in 3 ore di produzione:

$$\Pr(N_{[0,3]} = 0) = e^{-3 \cdot 0.2} = e^{-0.6} \approx 0.5488. \quad (15)$$

2. Il tempo alla produzione del decimo componente difettoso Y_{10} è una somma di 10 tempi di interarrivo indipendenti e identicamente distribuiti come $T_i \sim \text{Exp}(0.2)$. Pertanto, si ha

$$\text{E}[Y_{10}] = 10\text{E}[T_i] = \frac{10}{0.2} = 50 \text{ ore}. \quad (16)$$

Problema 5

Intuitivamente, facendo tanti lanci indipendenti di un dado a 6 facce ben bilanciato, si ottiene un valore massimo pari a 6 con alta probabilità. Dunque, ci aspettiamo una convergenza in probabilità al valore 6. Dimostriamolo:

$$\lim_{n \rightarrow \infty} \Pr(6 - X_n > \varepsilon) = \lim_{n \rightarrow \infty} \Pr(X_n < 6 - \varepsilon) \quad (17)$$

$$\leq \lim_{n \rightarrow \infty} \Pr(X_n \leq 5) \quad (18)$$

$$= \lim_{n \rightarrow \infty} \Pr(L_1 \leq 5, L_2 \leq 5, \dots, L_n \leq 5) \quad (19)$$

$$= \lim_{n \rightarrow \infty} (\Pr(L_i \leq 5))^n \quad (20)$$

$$= \lim_{n \rightarrow \infty} \left(\frac{5}{6}\right)^n \quad (21)$$

$$= 0, \quad \forall \varepsilon > 0. \quad (22)$$

Pertanto, rimane dimostrato che $X_n \xrightarrow{P} 6$

Problema 6

Con il metodo della cumulata inversa è possibile avere un algoritmo con efficienza del 100%. La funzione cumulata di probabilità si calcola come

$$F_X(x) = \int_0^x f_X(t)dt = \int_0^x \frac{1}{t^2} e^{-\frac{1}{t}} dt = \left[e^{-\frac{1}{t}} \right]_0^x = e^{-\frac{1}{x}}, \quad x > 0. \quad (23)$$

La funzione inversa della cumulata si calcola come:

$$u = e^{-\frac{1}{x}} \quad (24)$$

$$\log(u) = -\frac{1}{x} \quad (25)$$

$$-\log(u) = \frac{1}{x} \quad (26)$$

$$-\frac{1}{\log(u)} = x. \quad (27)$$

Pertanto, l'algoritmo sarà come segue:

1. Genero un campione $U \sim \mathcal{U}[0, 1]$
2. Calcolo $X = -\frac{1}{\log(U)}$