

Calcolo delle Probabilità e Statistica

Prova scritta del IV appello - 16/6/2004

Il candidato risolva i tre problemi proposti, motivando le proprie risposte.

1. Un'urna contiene 5 palline bianche e 2 rosse.
 - (a) Si eseguono 3 estrazioni con rimpiazzo. Sapendo che si sono ottenute 2 palline bianche, calcolare la probabilità di aver ottenuto una pallina bianca nella seconda estrazione.
 - (b) Calcolare la stessa probabilità nel caso che le estrazioni siano state eseguite senza rimpiazzo.
2. Per motivi di lavoro, un tale si sposta in auto tutte le settimane dalla città A alla città B , impiegando per il tragitto un tempo aleatorio X , misurato in ore, dove X è una v. a. avente densità

$$f(x) = \begin{cases} 2\alpha x e^{-\alpha x^2} & \text{per } x > 0 \\ 0 & \text{altrove} \end{cases}$$

dove $\alpha > 0$ è una costante. Inoltre, per l'esperienza acquisita, il tale ha calcolato che la probabilità di impiegare più di due ore per il viaggio è pari a e^{-4} .

- (a) Calcolare α .
 - (b) Calcolare la media di X .
 - (c) Supponendo che il tizio compia un solo viaggio alla settimana, quanto vale la probabilità che, in un mese, ogni volta egli impieghi almeno 1 ora e mezzo per andare da A a B ? (Si suppone che le durate dei vari viaggi siano l'una indipendente dall'altra; si considera un mese composto da 4 settimane).
3. Sia $X_1, X_2, \dots, X_n$ un campione di taglia n di legge

$$f(x; \theta) = \begin{cases} \theta x^{\theta-1} & \text{per } 0 < x < 1 \\ 0 & \text{altrimenti} \end{cases}$$

dove $\theta > 0$ è un parametro non noto. Si vuole effettuare il test

$$H_0 : \theta \leq 1, \quad H_1 : \theta > 1,$$

utilizzando la statistica $T = \min(X_1, \dots, X_n)$.

- (a) Calcolare la funzione di ripartizione di T (in funzione di θ).
 - (b) Determinare (in funzione di n) il numero reale $c \in (0, 1)$ in modo che $\{T > c\}$ sia una regione critica per il test assegnato di livello $\alpha = 0.01$.
 - (c) Supponiamo che T abbia preso il valore $t = 0.2$; quanto era grande la taglia del campione se l'ipotesi nulla è stata respinta (ancora al livello $\alpha = 0.01$)?

Calcolo delle Probabilità e Statistica

Soluzioni della prova scritta del IV appello - 16/6/2004

1. Definiamo per $i = 1, 2, 3$, le v.a.

$$X_i = \begin{cases} 1 & \text{se esce una pallina bianca all'i-esima estrazione} \\ 0 & \text{altrimenti} \end{cases}$$

e gli eventi $A_1 = \{X_1 = 0, X_2 = X_3 = 1\}$, $A_2 = \{X_2 = 0, X_1 = X_3 = 1\}$, $A_3 = \{X_3 = 0, X_1 = X_2 = 1\}$. Si tratta allora di calcolare

$$\begin{aligned} P(X_2 = 1 | X_1 + X_2 + X_3 = 2) &= \frac{P(X_2 = 1, X_1 + X_2 + X_3 = 2)}{P(X_1 + X_2 + X_3 = 2)} \\ &= \frac{P(X_1 = 1, X_2 = 1, X_3 = 0) + P(X_1 = 0, X_2 = 1, X_3 = 1)}{P(X_1 + X_2 + X_3 = 2)} \\ &= \frac{P(A_1) + P(A_3)}{P(A_1) + P(A_2) + P(A_3)} \end{aligned}$$

Basta allora osservare che in entrambi i casi si ha $P(A_1) = P(A_2) = P(A_3)$ per concludere che in entrambi i casi la probabilità cercata è $\frac{2}{3}$.

Analizziamo i due casi più nel dettaglio.

- (a) In questo caso le v. a. X_1, X_2, X_3 sono indipendenti e tutte di legge $B(1, p)$, con $p = 5/7$. Si può anche osservare che la v. a. $X_1 + X_2 + X_3$ ha legge $B(3, p)$. Dunque

$$\begin{aligned} P(X_1 = 1, X_2 = 1, X_3 = 0) &= P(X_1 = 0, X_2 = 1, X_3 = 1) = p^2(1-p); \\ P(X_1 + X_2 + X_3 = 2) &= \binom{3}{2} p^2(1-p) = 3p^2(1-p). \end{aligned}$$

$$\text{Quindi } P(X_2 = 1 | X_1 + X_2 + X_3 = 2) = \frac{2p^2(1-p)}{3p^2(1-p)} = \frac{2}{3}.$$

- (b) Se le estrazioni sono senza rimpiazzo si ha

$$\begin{aligned} P(X_1 = 1, X_2 = 1, X_3 = 0) &= P(X_3 = 0 | X_2 = 1, X_1 = 1) P(X_2 = 1 | X_1 = 1) P(X_1 = 1) \\ &= \frac{2}{5} \frac{4}{6} \frac{5}{7} = \frac{4}{21} \end{aligned}$$

e analogamente $P(X_1 = 0, X_2 = 1, X_3 = 1) = \frac{4}{5} \frac{5}{6} \frac{2}{7} = \frac{4}{21}$. Inoltre, poiché $X_1 + X_2 + X_3$ ha densità ipergeometrica di parametri 5, 2, 3, si trova

$$P(X_1 + X_2 + X_3 = 2) = \frac{\binom{5}{2} \binom{2}{1}}{\binom{7}{3}} = \frac{4}{7}.$$

$$\text{Quindi } P(X_2 = 1 | X_1 + X_2 + X_3 = 2) = \frac{\frac{8}{21}}{\frac{4}{7}} = \frac{2}{3}$$

2. (a) La costante α si determina in base alla richiesta che sia $P(X > 2) = e^{-4}$. Si ha infatti

$$P(X > 2) = \int_2^{+\infty} 2\alpha x e^{-\alpha x^2} dx = \left[-e^{-\alpha x^2}\right]_2^{+\infty} = e^{-4\alpha}$$

e dall'uguaglianza $e^{-4} = e^{-4\alpha}$ si deduce $\alpha = 1$.

- (b) Si ha

$$\mathbf{E}[|X|] = \mathbf{E}[X] = \int_0^{+\infty} 2x^2 e^{-x^2} dx$$

L'integrale qui sopra non è calcolabile per via elementare, tuttavia, per la simmetria della funzione $x \mapsto x^2 e^{-x^2}$, si ha

$$\int_0^{+\infty} 2x^2 e^{-x^2} dx = \int_{-\infty}^{+\infty} x^2 e^{-x^2} dx = \sqrt{\pi} \left(\int_{-\infty}^{+\infty} \frac{1}{\sqrt{\pi}} x^2 e^{-x^2} dx \right);$$

Osservando ora che l'ultimo integrale ottenuto coincide con la varianza di una v. a. avente legge $\mathcal{N}(0, 1/2)$, e quindi vale $1/2$, si conclude che X ha speranza matematica finita e $\mathbf{E}[X] = \sqrt{\pi}/2$.

- (c) Indichiamo con X_1, X_2, X_3, X_4 le durate dei 4 viaggi che il tizio deve fare in un mese. Si tratta di 4 v. a. indipendenti, aventi tutte la legge assegnata (con $\alpha = 1$). Dobbiamo allora calcolare

$$\begin{aligned} P(X_1 \geq 1.5, X_2 \geq 1.5, X_3 \geq 1.5, X_4 \geq 1.5) \\ &= P(X_1 \geq 1.5)P(X_2 \geq 1.5)P(X_3 \geq 1.5)P(X_4 \geq 1.5) = (P(X_1 \geq 1.5))^4 \\ &= \left(\int_{1.5}^{+\infty} 2x e^{-x^2} dx \right)^4 = \left(\left[-e^{-x^2}\right]_{1.5}^{+\infty} \right)^4 \\ &= e^{-2.25 \times 4} = e^{-9} \end{aligned}$$

3. (a) È immediato vedere che $P(T \leq t) = 0$ per $t \leq 0$ e $P(T \leq t) = 1$ per $t \geq 1$. Per $0 < t < 1$ si ha $P(T \leq t) = 1 - P(T > t)$, dove

$$\begin{aligned} P(T > t) &= P(X_1 > t, \dots, X_n > t) = (P(X_1 > t))^n \\ &= \left(\int_t^1 \theta x^{\theta-1} dx \right)^n = (1 - t^\theta)^n. \end{aligned}$$

Dunque la funzione di ripartizione di T è

$$P(T \leq t) = \begin{cases} 0 & \text{per } t < 0 \\ 1 - (1 - t^\theta)^n & \text{per } 0 \leq t < 1 \\ 1 & \text{per } t \geq 1 \end{cases}$$

- (b) Supponendo di avere determinato c , il livello del test è dato da

$$\sup_{\theta \leq 1} P(T > c) = \sup_{\theta \leq 1} (1 - c^\theta)^n.$$

Osserviamo ora che la funzione $(1 - c^\theta)^n$ è crescente rispetto a θ (si ha infatti $c^{\theta_2} \leq c^{\theta_1}$ per $\theta_1 \leq \theta_2$ per $0 \leq c \leq 1$). Quindi il livello del test è $(1 - c)^n$. Deve allora essere $(1 - c)^n = 0.01$, da cui si ricava $c = 1 - (0.01)^{1/n}$.

- (c) Poiché l'ipotesi nulla viene respinta, il valore $t = 0.2$ appartiene alla regione critica. In altre parole vale la disuguaglianza $0.2 > 1 - (0.01)^{1/n}$, che, risolta rispetto a n , dà $n > 20.61$. Dunque il campione è di taglia almeno 21.