CdL in Ing. Robotica e dell'Automazione, Prob. e Proc. Stoc. (455AA)
A.A. 2023/24 - Appello straordinario 2024-04-04

La durata della prova è di **120 minuti**. *Fornire risposte dettagliate.*

Problema 1

Un sistema di telecomunicazioni ha due dispositivi per la trasmissione dati, A e B. A volte, per errori di trasmissione, i dati inviati da ciascuno dei dispositivi possono essere corrotti. Si stima che il dispositivo A abbia una probabilità del 50% di inviare dati corrotti, mentre il dispositivo B è più affidabile, ossia corrompe i dati con probabilità 5%, ad ogni invio, indipendentemente dalle volte precedenti.

L'ingegner Luca è un po' (troppo) sbadato e ha messo nella stessa scatola i due dispositivi, che dall'esterno sono indistinguibili. Perciò decide di condurre un esperimento in cui trasmette gli stessi dati per 5 volte consecutive. Tuttavia, ogni volta sceglie completamente a caso uno dei due dispositivi, indipendentemente dalle volte precedenti. Ad ogni trasmissione, Luca verifica se i dati sono corrotti o meno.

1. Calcolare la probabilità che Luca osservi dati corrotti in 4 trasmissioni su 5 (in qualsiasi ordine).
2. Sapendo che Luca ha osservato dati corrotti in 4 trasmissioni su 5, calcolare la probabilità che non abbia mai utilizzato il dispositivo B per la trasmissione.
3. Sapendo che Luca ha osservato dati corrotti in 4 trasmissioni su 5, calcolare la probabilità che abbia utilizzato il dispositivo B esattamente 1 volta.

Una soluzione:

1. Dal testo deduciamo che si tratta di 5 esperimenti indipendenti, ciascuno con probabilità di "successo" (ossia di inviare dati corrotti)

$$\begin{aligned} P(\text{“osserva dati corrotti”}) &= P(\text{'sceglie } A\text{'})P(\text{'}A \text{ corrompe i dati'}) + P(\text{'sceglie } B\text{'})P(\text{'}B \text{ corrompe i dati'}) \\ &= \frac{1}{2}\cdot\frac{1}{2} + \frac{1}{2}\cdot\frac{5}{100} = 0.275. \end{aligned} \tag{1}$$

Possiamo dedurne quindi che la probabilità di osservare 4 "successi" su 5 esperimenti è data dalla densità binomiale

$$P(\text{'osserva 4 corrotte su 5'}) = \binom{5}{4}(0.275)^4(1-0.275) = 5\cdot(0.275)^4\cdot 0.725 \approx 2\%. \tag{2}$$

2. Poniamo C_4 = 'osserva 4 corrotte su 5' e A_5 = 'Luca ha sempre usato il dispositivo A'. Per Bayes,

$$P(A_5|C_4) = \frac{P(A_5)\cdot P(C_4|A_5)}{P(C_4)}. \tag{3}$$

Il denominatore è stato calcolato nel punto precedente. Per il numeratore, troviamo

$$P(A_5) = \left(\frac{1}{2}\right)^5 \tag{4}$$

e

$$P(C_4|A_5) = \binom{5}{4}\left(\frac{1}{2}\right)^5. \tag{5}$$

Si trova $P(A_5|C_4) \approx 24\%$.

3. Poniamo $B_1 =$ "Luca ha usato B una volta sola". Troviamo

$$P(B_1|C_4) = P(B_1 \text{ e } C_4)/P(C_4). \tag{6}$$

Il denominatore è già stato calcolato. Per il numeratore, supponiamo ad esempio che la sequenza sia $AAAAB$, la cui probabilità è $1/2^5$. Allora la probabilità che Luca osservi i dati corrotti 4 volte (ossia i dati corretti una volta sola) è $4 \cdot (50\%)^4 \cdot 5\% + (50\%)^4 \cdot 95\%$ (si tratta di determinare quale dei 5 tentativi ha trasmesso correttamente). Notiamo che tale probabilità non dipende dall'ordine e quindi (ad esempio usando la formula di disintegrazione sulle 5 possibili sequenze $AAAAB$, $AAABA$, $AABAA$, $ABAAA$, $BAAAA$) la probabilità cercata è

$$P(C_4 \text{ e } B_1) = \frac{5}{2^5} \cdot \left[4 \cdot (50\%)^4 \cdot 5\% + (50\%)^4 \cdot 95\%\right] \tag{7}$$

Si trova infine

$$P(B_1|C_4) = \frac{\frac{5}{2^5} \cdot \left[4 \cdot (50\%)^4 \cdot 5\% + (50\%)^4 \cdot 95\%\right]}{5 \cdot (0.275)^4 \cdot 0.725} \approx 54\%. \tag{8}$$

Problema 2

Sia X una variabile aleatoria continua a valori reali avente densità data dalla funzione

$$p(X = x) = c \cdot \begin{cases} e^{-x} & \text{per } x \geq 0, \\ \frac{1}{|x-1|^3} & \text{per } x < 0. \end{cases}$$

dove c è una opportuna costante.

1. Determinare c.
2. Calcolare la mediana di X. Dire in particolare se è positiva o negativa (o nulla).
3. Calcolare la densità di $Y = X^2$ (se esiste).

Una soluzione:

1. La costante c si trova imponendo

$$\int_{-\infty}^{\infty} p(X = x) dx = 1. \tag{9}$$

Si trova

$$\begin{aligned} \int_{-\infty}^{\infty} p(X = x) dx &= c \left[\int_0^{\infty} e^{-x} dx + \int_{-\infty}^{0} |x-1|^{-3} dx\right] \\ &= c \left[1 + \int_0^{\infty} |1-x|^{-3} dx\right] \\ &= c\,[1 + 1/2] = 3c/2 \end{aligned} \tag{10}$$

concludiamo che $c = 2/3$.

2. Per la mediana dobbiamo prima calcolare la $\mathrm{CDF}_X(t)$. Se $t \leq 0$, troviamo

$$\begin{aligned} P(X \leq t) &= \int_{-\infty}^{t} p(X = x)dx = c \int_{-\infty}^{t} |x-1|^{-3} dx \\ &= c \int_{-t}^{\infty} |1-x|^{-3} dx = c\frac{(1-y)^{-2}}{-2}|_{-t}^{\infty} = \frac{(1+t)^{-2}}{3}. \end{aligned} \tag{11}$$

dove nell'ultima equazione abbiamo usato che $c = 2/3$. Troviamo in particolare che $P(X \leq 0) = 1/3$, quindi la mediana è positiva. Per $t \geq 0$ troviamo che

$$P(X \leq t) = P(X \leq 0) + P(0 \leq X \leq t) = \frac{1}{3} + c\int_0^t e^{-x} dx = \frac{1}{3} + \frac{2}{3}(1 - e^{-t}). \tag{12}$$

Dobbiamo quindi trovare $t \geq 0$ in modo che l'espressione sopra valga $1/2$. Si ha pertanto

$$\frac{1}{2} = \frac{1}{3} + \frac{2}{3}(1 - e^{-t}) \quad \leftrightarrow \quad e^{-t} = 3/4 \tag{13}$$

e quindi la mediana vale $\log(4/3)$ (logaritmo in base naturale).

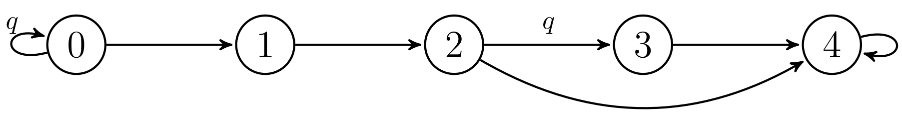
3. Poiché $y = g(x) = x^2$ soddisfa le ipotesi del cambio di variabile (generalizzato), con derivata $|g'(x)| = 2|x|$ e inversa $g^{-1}(y) = \{-\sqrt{y}, \sqrt{y}\}$ per $y > 0$, $g^{-1}(0) = \{0\}$ e $g^{-1}(y) = \emptyset$ se $y < 0$, troviamo (per $y > 0$)

$$p(Y = y) = \frac{1}{3\sqrt{y}} \left[\frac{1}{(1+\sqrt{y})^3} + e^{-\sqrt{y}} \right], \tag{14}$$

mentre $p(Y = y) = 0$ se $y \leq 0$.

Problema 3

Si consideri la catena di Markov $(X_n)_{n \geq 0}$ con matrice di transizione Q rappresentata graficamente in figura, dove $q \in [0, 1]$ è un parametro, e $X_0 = 0$.



1. Al variare del parametro $q \in [0, 1]$ classificare gli stati della catena e calcolare tutte le distribuzioni invarianti.
2. Al variare del parametro $q \in [0, 1]$, calcolare $P(X_5 = 4)$.
3. Si supponga che il parametro q sia anch'esso una variabile aleatoria, distribuito a priori con una densità continua uniforme su $[0, 1]$. Avendo osservato $X_5 = 4$, calcolare la densità a posteriori di q. Dire se è più probabile $q > 1/2$ o $q < 1/2$.

Una soluzione:

1. Se $0 < q < 1$ allora tutti gli archi hanno peso strettamente positivo, quindi l'unico stato ricorrente (assorbente) è 4 e i rimanenti sono transitori. L'unica distribuzione invariante è $\pi = (0, 0, 0, 0, 1)$ (ordinando gli stati nel modo naturale). Lo stesso discorso vale nel caso $q = 0$. Se $q = 1$ allora gli stati 0 e 4 sono ricorrenti (ciascuno costituisce una

classe chiusa irriducibile) e i rimanenti sono transitori, quindi le distribuzioni invarianti sono infinite e tutte del tipo $\pi = (\alpha, 0, 0, 0, 1-\alpha)$ per $\alpha \in [0, 1]$.

2. Poiché sappiamo che $X_0 = 0$, dobbiamo calcolare i pesi dei cammini che partono da 0 e arrivano in 4 in 5 passi. Troviamo solo i cammini

- $0 \to 1 \to 2 \to 4 \to 4 \to 4$ con peso $(1-q)^2$,
- $0 \to 1 \to 2 \to 3 \to 4 \to 4$, con peso $q(1-q)$,
- $0 \to 0 \to 1 \to 2 \to 4 \to 4$ con peso $q(1-q)^2$,
- $0 \to 0 \to 1 \to 2 \to 3 \to 4$ con peso $q \cdot (1-q) \cdot 1 \cdot q \cdot 1 = q^2(1-q)$,
- $0 \to 0 \to 0 \to 1 \to 2 \to 4$ con peso $q \cdot q \cdot (1-q) \cdot 1 \cdot (1-q) = q^2(1-q)^2$.

Concludiamo che la probabilità cercata è

$$\begin{aligned} P(X_5 = 4) &= (1-q)^2 + q(1-q) + q(1-q)^2 + q^2(1-q) + q^2(1-q)^2 \\ &= (1-q)\left[1 - q + q + q(1-q) + q^2 + q^2(1-q)\right] \\ &= (1-q)(1 + q + q^2 - q^3) \\ &= 1 - 2q^3 + q^4. \end{aligned} \tag{15}$$

3. Avendo osservato $P(X_5 = 4)$, dobbiamo calcolare la densità a posteriori per q, data dalla formula di Bayes (caso continuo/discreto): per $x \in [0, 1]$,

$$p(q = x | X_5 = 4) \propto p(q = x) \cdot P(X_5 = 4 | q = x) \propto P(X_5 = 4 | q = x) \tag{16}$$

dove $P(X_5 = 4|q = x)$ è stata calcolata al punto precedente. Troviamo quindi

$$p(q = x|X_5 = 4) = c\left(1 - 2x^3 + x^4\right) \tag{17}$$

dove c è da determinare. Integrando tra 0 e 1, imponendo che l'integrale faccia 1 e ricordando cho $\int_0^1 x^n dx = 1/(n+1)$, si trova

$$c = \left(1 - 2 \cdot \frac{1}{4} + \frac{1}{5}\right)^{-1} = \frac{10}{7}. \tag{18}$$

Per capire invece se sia più probabile che $q < 1/2$ o $q > 1/2$, basta integrare la densità trovata tra 0 e 1/2, per trovare

$$P(q < 1/2|X_5 = 4) = \frac{10}{7} \cdot \left[\frac{1}{2} - \frac{2}{2^4 \cdot 4} + \frac{1}{5 \cdot 2^5}\right] \approx 67\% > 1/2, \tag{19}$$

quindi è più probabile che sia $q < 1/2$.