CdL in Ing. Robotica e dell'Automazione, Prob. e Proc. Stoc. (455AA)
A.A. 2024/25 - Prova 2025-07-16

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

Problema 1

In un laboratorio scientifico ci sono 30 provette di tipo A, 20 di tipo B e 10 provette di tipo C (60 in tutto), tutte indistinguibili dall'esterno. Si sa però che ogni provetta di tipo A contiene una sostanza reattiva con probabilità 10%, ogni provetta di tipo B contiene una sostanza reattiva con probabilità 60%, mentre ogni provetta di tipo C contiene una sostanza reattiva con probabilità 96%, ciascuna indipendentemente dalle altre.

1. Qual è la probabilità che una provetta scelta a caso contenga una sostanza reattiva?
2. Avendo osservato che una provetta scelta a caso contiene una sostanza reattiva, qual è il suo tipo più probabile?
3. Avendo osservato che una provetta scelta a caso contiene una sostanza reattiva, è più probabile che sia di tipo B oppure che non lo sia?
4. Avendo osservato che una provetta scelta a caso contiene una sostanza reattiva, è più probabile che una ulteriore provetta selta a caso sia di tipo A oppure non lo sia?

Una soluzione:

Poniamo $R =$ 'provetta estratta è reattiva'. 1. $P(R) = \frac{1}{2} \cdot 10\% + \frac{1}{3} \cdot 60\% + \frac{1}{6} \cdot 96\% = 41\%$.

2. Per Bayes, $P(tipo|R) = P(R|tipo)P(tipo)/P(reattiva)$ per ogni $tipo \in \{A, B, C\}$, quindi basta confrontare i termini $P(R|tipo)P(tipo)$:

$$P(R|A)P(A) = 5\%, P(R|B)P(B) = 20\%, P(R|C)P(C) = 16\%$$

e quindi il tipo B è il più probabile.

3. Abbiamo calcolato le probabilità a posteriori con Bayes al punto precedente. Ora le alternative da confrontare sono due: B contro non $B = A \cup C$, per cui

$$P(B|reattiva) = \frac{20\%}{P(reattiva)}, \quad P(A \cup C|reattiva) = \frac{5\% + 16\%}{P(reattiva)} = \frac{21\%}{P(reattiva)}$$

quindi è più probabile che non sia di tipo B.

4. Usando Bayes, troviamo che la probabilità che la prima provetta estratta sia del tipo A (evento A_1) è $\propto 5\%$, mentre che non lo sia è $\propto 36\%$, quindi precisamente $P(A_1|R) = 5/41$, $P(\text{non } A_1|R) = 36/41$. Nel caso sia vero A_1, la probabilità che una nuova provetta sia del tipo A (evento A_2) è $29/59$, mentre nel secondo caso è $30/59$, quindi concludiamo che $P(A_2|R) = \frac{5}{41} \cdot \frac{29}{59} + \frac{36}{41} \cdot \frac{30}{59} \approx 50,6\%$ ossia è (leggermente) più probabile che sia del tipo A.

Problema 2

Per $\theta > 0$, si consideri la densità di probabilità continua

$$f_\theta(t) := \begin{cases} ct \exp(-t/\theta) & \text{se } t \geq 0. \\ 0 & \text{altrimenti,} \end{cases}$$

dove $c = c(\theta)$ è una opportuna costante che può dipendere da θ, ed $n \geq 1$ variabili aleatorie indipendenti $(T_1, \ldots, T_n)$, ciascuna con densità f_θ.

1. Determinare c in modo che f_θ sia una densità di probabilità.
2. Calcolare il valor medio di T_1 e la varianza di T_1 (in funzione di θ).
3. Scrivere la densità (se esiste) della variabile $\sqrt{T_1}$ (in funzione di θ).
4. Per $n = 8$, si osservano i valori $(T_i)_{i=1}^8 = (45, 50, 48, 47, 52, 49, 60, 49)$. Fornire una stima di massima verosimiglianza per θ.

Una soluzione:

1. Si tratta di imporre che l'integrale della densità valga sempre 1 (per ogni θ). Troviamo con un cambio di variabile che

$$\int_0^\infty t \exp(-t/\theta) dt \stackrel{x=t/\theta}{=} \theta^2 \int_0^\infty x \exp(-x) dx = 1$$

, quindi basta imporre $c(\theta) = \theta^{-2}$. 2. Per il valor medio usiamo la formula

$$\int_0^\infty t f_\theta(t) dt \stackrel{x=t/\theta}{=} \theta \int_0^\infty x^2 \exp(-x) dx = 2\theta$$

avendo integrato per parti (saltando un passaggio). Per la varianza, calcoliamo similmente il momento secondo

$$\int_0^\infty t^2 f_\theta(t) dt \stackrel{x=t/\theta}{=} \theta^2 \int_0^\infty x^3 \exp(-x) dx = 6\theta^2$$

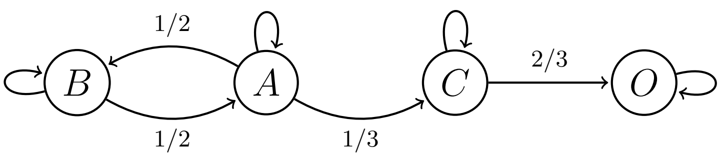
e poi otteviamo $\sigma^2 = \theta^2(6-4) = 2\theta^2$. 3. T_1 è positiva, pertanto per $x \geq 0$, $P(\sqrt{T_1} \leq x) = P(T_1 \leq x^2)$ e quindi derivando si trova

$$p(\sqrt{T_1} = x) = 2x f_\theta(x^2) = 2cx^3 \exp\left(-x^2/\theta\right).$$

(calcoli analoghi si possono fare con il cambio di variabile). 4. Impostando al solito la verosimiglianza e derivando in θ per massimizzarla, si trova lo stimatore $\theta_{MLE} = \frac{1}{2n} \sum_{i=1}^n x_i$ dove $(x_i)_{i=1}^n$ sono le osservazioni. Pertanto si trova in questo caso che $\theta_{MLE} = 25$.

Problema 3

Si consideri una catena di Markov $(X_n)_n$ con probabilità di transizione rappresentate in figura (completare con le probabilità mancanti) e tale che $P(X_0 = A) = P(X_0 = B) = P(X_0 = C) = 1/3$.



1. Classificare gli stati e calcolare tutte le distribuzioni invarianti.
2. Calcolare la probabilità di osservare la parola ABACO (ossia $X_0 = A$, $X_1 = B$, $X_2 = A$, ecc.).
3. Avendo osservato $X_4 = O$, quale tra i tre stati A, B, C è il più probabile per la variabile X_0?

4. Avendo osservato $X_4 = X_0$, quale tra i tre stati A, B, C è il più probabile per la variabile X_0?

Una soluzione:

1. Gli stati $\{A, B, C\}$ sono transienti, lo stato $\{O\}$ è ricorrente (assorbente). Le distribuzioni invarianti valgono 0 sui transienti, e quindi è solo il vettore $\pi = (0, 0, 0, 1)$ (nell'ordine (B, A, C, O)).

2. Il peso del cammino $ABACO$ è $\frac{1}{2} \cdot \frac{1}{2} \cdot \frac{1}{3} \cdot \frac{2}{3} = \frac{1}{18}$, mentre la probabilità che $X_0 = A$ è pure $1/3$, quindi troviamo che la probabilità di osservare la parola in questione è $1/54$.

3. Essendo gli stati A, B, C inizialmente equiprobabili, si tratta solo di stabilire quello con la maggiore verosimiglianza $L(i) = P(X_4 = O|X_0 = i)$ per $i \in \{A, B, C\}$. È chiaro tuttavia che lo stato con maggiore verosimiglianza è C, perché i cammini che partono da A o B al tempo $t = 0$ e al tempo $t = 4$ sono in O devono necessariamente passare per C. Più precisamente, i contribuiti del tipo $\gamma \to C \to O$ con γ cammino solo sugli stati $\{A, B\}$, avranno verosimiglianza minore del cammino $C \to O \to O \to O \to O$. Similmente per i cammini è della forma $\gamma \to C \to C \to O$, avranno verosimiglianza minore del singolo cammino $C \to C \to O \to O \to O$ ecc. Quindi il valore più probabile a posteriori è $X_0 = C$.

4. Di nuovo, essendo gli stati A, B, C inizialmente equiprobabili, si tratta di stabilire quello con la maggiore verosimiglianza $L(i) = P(X_4 = A|X_0 = i)$ per $i \in \{A, B, C\}$. Osserviamo che nel caso $i = C$ abbiamo solo il cammino $C \to C \to C \to C \to C \to C$ che ha probabilità $1/3^4$. Per lo stato $i = B$ osserviamo che già il cammino banale $B \to B \to B \to B \to B$ darebbe un contribuito $1/2^4 > 1/3^4$, e similmente per lo stato $i = A$. Quindi $i = C$ non massimizza la verosimiglianza. Per capire quale tra A e B è massimo, osserviamo che ciascun cammino che parte da B al tempo $t = 0$ e al tempo $t = 4$ è in B può essere messo in corrispondenza con un cammino che parte da A e arriva in A al tempo $t = 4$ (e viceversa): essi saranno sequenze di sole lettere B, A e equindi basta scambiare ciascuna A con B. Ora però i pesi di questi cammini saranno diversi, poiché $Q_{B \to B} = 1/2$, mentre $Q_{A \to A} = 1/6$, e invece $Q_{A \to B} = Q_{B \to A} = 1/2$. Quindi ciascun cammino che parte da B e arriva in B avrà sempre peso maggiore del corrispondente cammino che parte da A e arriva A. In conclusione quindi abbiamo che la verosimiglianza è massima per $i = B$.