CdL in Ing. Robotica e dell'Automazione, Prob. e Proc. Stoc. (455AA)
A.A. 2023/24 - Appello 2024-07-17

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

Problema 1

Un'azienda di consegne si occupa di prodotti alimentari e per la cura della persona. Ciascuna consegna contiene un numero di prodotti alimentari rappresentato da una variabile Poisson di parametro λ_a, e un numero di prodotti per la cura della persona rappresentato da una variabile Poisson di parametro λ_c: i due numeri sono indipendenti tra loro e indipendenti per ciascuna consegna.

1. Calcolare (in funzione di λ_a e λ_c) il valor medio e la varianza del numero complessivo di prodotti in 10 consegne.
2. Sapendo che in 10 consegne sono stati portati 300 prodotti, di cui 200 alimentari e 100 per la cura della persona, fornire una stima di λ_a e λ_c.
3. Sapendo invece solamente che in 10 consegne sono stati portati 300 prodotti, è possibile fornire una stima separatamente per λ_a e λ_c?

Una soluzione:

1. Il valor medio è dato dalla somma dei valor medi, quindi $\mathbb{E}[N] = 10 * (\lambda_a + \lambda_c)$. Per indipendenza, la varianza si somma e ricordando che la varianza di una Poisson è data dal parametro, troviamo $\operatorname{Var} N = 10 * (\lambda_a + \lambda_c)$.

2. Usiamo il fatto che la somma di Poisson indipendenti è pure Poisson (si sommano i parametri): troviamo la verosimiglianza

$$L(\lambda_a, \lambda_c) \propto e^{-10(\lambda_a+\lambda_c)} \lambda_a^{200} \lambda_c^{100}. \tag{1}$$

che è massima quando $\lambda_a = 20$, $\lambda_c = 10$.

3. Scrivendo la verosimiglianza in questo caso si ottiene

$$L(\lambda_a, \lambda_c) \propto e^{-10(\lambda_a+\lambda_c)} (\lambda_a + \lambda_c)^{300}. \tag{2}$$

Non avendo informazioni a priori su λ_a, λ_c, possiamo solo trovare una stima per $\lambda_a + \lambda_c$ (in particolare otteniamo la stima $\lambda_a + \lambda_c = 30$). Ogni scelta di parametri λ_a, λ_c che somma a 30 è una possibile stima di massima verosimiglianza (però non è unica).

Problema 2

Sia X una variabile aleatoria che segue una distribuzione esponenziale di parametro $\lambda = \ln(10)$. Poniamo $Y = \log_{10} X$.

1. Determinare la densità di Y.
2. Calcolare la funzione quantile di Y.
3. Supponendo di osservare che $Y > 4$, calcolare la densità della variabile $X - 10^4$.

Una soluzione:

1. Per ogni $y \in \mathbb{R}$, scriviamo

$$P(Y > y) = P(\log_{10} X > y) = P(X > 10^y) = e^{-\lambda 10^y} = 10^{-10^y}. \tag{3}$$

Derivando e cambiando segno si ottiene quindi la densità:

$$p(Y = y) = (\ln(10)^2 10^{-10^y} \cdot 10^y. \tag{4}$$

2. Avendo trovato la funzione di sopravvivenza, la funzione quantile si ottiene invertendola (siamo nel caso continuo): dato $\alpha \in (0, 1)$, si ha

$$10^{-10^{q(\alpha)}} = 1 - \alpha \tag{5}$$

da cui

$$q(\alpha) = \log_{10}(-\log_{10}(1 - \alpha)). \tag{6}$$

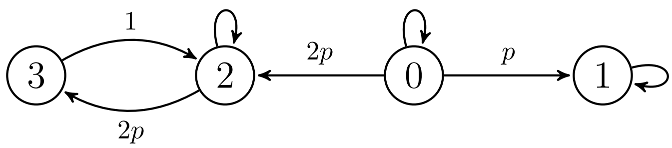
3. L'evento $Y = \log_{10}(X) > 4$ coincide con $X > 10^4$. Ne segue che $X - 10^4$ è una variabile non negativa. Calcolando la funzione di sopravvivenza, per $t > 0$ si trova

$$P(X - 10^4 > t | X > 10^4) = \frac{P(X > 10^4 + t)}{P(X > 10^4)} = \frac{\exp\left(-\lambda(10^4 + t)\right)}{\exp\left(-\lambda 10^4\right)} = \exp\left(-\lambda t\right) = 10^{-t}, \tag{7}$$

ossia $X - 10^4$ ha densità esponenziale di parametro $\lambda = \ln(10)$.

Problema 3

Si consideri una catena di Markov $(X_n)_{n \geq 0}$ sull'insieme degli stati $\{0, 1, 2, 3\}$ con matrice di transizione rappresentata in figura



dove $p \in [0, 1/3]$ è un parametro.

1. Al variare di p, classificare gli stati della catena e calcolarne tutte le distribuzioni invarianti.
2. Supponendo di sapere che $X_0 = 0$, si osserva poi l'evento

$$A = \{\text{“}X_1 = 0 \text{ e } X_2 = 1\text{” oppure “}X_1 = X_2 = 2\text{”}\}.$$

   Fornire una stima di massima verosimiglianza per p.
3. Calcolare (in funzione di p) la probabilità $\lim_{n \to \infty} P(X_n = 1 | X_0 = 0)$

Una soluzione:

1. Se $p = 0$ lo stato 3 è transitorio, mentre 0, 1, 2 sono assorbenti: tutte le distribuzioni invarianti sono del tipo $(\alpha_0, \alpha_1, \alpha_2, 0)$ con $\alpha_i \in [0, 1]$, $\alpha_0 + \alpha_1 + \alpha_2 = 0$.

In tutti gli altri casi $p \in (0, 1/3]$ solo lo stato 0 è transitorio gli altri sono ricorrenti e ci sono due classi chiuse irriducibili $\{1\}$, $\{2, 3\}$. Le distribuzioni invarianti sono quindi (impostando il bilancio di flusso per la seconda classe)

$$(0, \alpha, (1-\alpha) \cdot \frac{1}{1+2p}, (1-\alpha) \cdot \frac{2p}{1+2p}), \tag{8}$$

dove $\alpha \in [0, 1]$.

2. Fissato p, la probabilità dell'evento A si calcola pesando i due cammini $0 \to 0 \to 1$ con probabilità $(1-3p)p$, e $0 \to 2 \to 2$ con probabilità $2p(1-2p)$. Pertanto la verosimiglianza è

$$L(p; A) = (1-3p)p + 2p(1-2p). \tag{9}$$

Osserviamo che per $p = 0$ è nulla e pure per $p = 1/3$, quindi il massimo sarà in un punto interno all'intervallo $[0, 1/3]$. Derivando rispetto a p troviamo l'equazione

$$-3p + 1 - 3p + 2 - 4p - 4p = 0 \quad \to \quad p = \frac{3}{14}. \tag{10}$$

3. Se $p = 0$ avremo $P(X_n = 0 | X_0 = 0) = 1$ per ogni n, e quindi il limite richiesto vale 0. Altrimenti, per $p > 0$, si tratta di calcolare la probabilità con cui la catena entrerà nella classe chiusa irriducibile $\{1\}$ (invece di entrare in quella alternative $\{2, 3\}$. Posto T il primo istante in cui $X_n \neq 0$, si trova che $P(X_T = 1 | X_0 = 0) = p/(p + 2p) = 1/3$.