CdL in Ing. Robotica e dell'Automazione, Prob. e Proc. Stoc. (455AA)
A.A. 2022/23 - Prova scritta 2023-02-15

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

Problema 1

Valentina per recarsi al lavoro a Pisa può scegliere se usare l'auto oppure andare a piedi. La sua decisione è solamente influenzata dalle condizioni atmosferiche del giorno. Se piove, allora prende l'auto con probabilità 95%. Se non piove, la stessa probabilità scende al 25%. Introduciamo un parametro $p \in (0,1)$ che indica la probabilità che in un dato giorno piova a Pisa. Supponiamo per semplicità che sia lo stesso per tutti i giorni considerati, e che il meteo in un dato giorno sia indipendente da quello degli altri giorni. Si osserva il comportamento di Valentina per una *settimana lavorativa* di 5 giorni consecutivi.

1. Supponendo noto p, calcolare la probabilità che Valentina abbia usato l'auto esattamente per tre giorni consecutivi e i rimanenti due sia andata a piedi (dare il risultato in funzione di p).
2. Si osserva che Valentina ha usato l'auto esattamente per 3 giorni (non necessariamente consecutivi). Fornire la stima di massima verosimiglianza per p.
3. Come cambia la risposta del punto precedente se si osserva più precisamente che Valentina ha usato l'auto il lunedì, il martedì e il mercoledì (e i rimanenti giorni è andata a piedi)?

Una soluzione:

1. La probabilità che Valentina usi l'auto in un dato giorno è

$$q = p \cdot 95\% + (1-p) \cdot 25\% = 25\% + p70\%.$$

Data l'indipendenza tra i vari giorni, si ha che la probabilità che prenda l'auto tre volte consecutive è

$$3q^3(1-q)^2 = 3(25\% + p70\%)^3(75\% - p70\%)^2,$$

dove il fattore 3 è dovuto ai tre casi possibili in cui ha preso l'auto: Lun-Mar-Mer, Mar-Mer-Gio, Mer-Gio-Ven.

2. Scriviamo la verosimiglianza, ossia la probabilità (noto p) che abbia usato l'auto tre volte su 5. La densità è binomiale, quindi

$$L(p) = \binom{5}{3} q^3(1-q)^2,$$

dove $q = p \cdot 95\% + (1-p) \cdot 25\%$ come al punto precedente. Passando al logaritmo e deriviamo:

$$\partial_p \log L(p) = 3\frac{\partial_p q}{q} - 2\frac{\partial_p q}{1-q} = \left(\frac{3}{q} - \frac{2}{1-q}\right)\partial_p q.$$

imponendo che la derivata si annulli, si trova che poiché $\partial_p q = 95\% - 25\% \neq 0$, deve valere

$$\frac{3}{q} - \frac{2}{1-q} = 0,$$

ossia $q = 2/5 = 40\%$, da cui

$$25\% + p70\% = 40\%, \quad \to \quad p = 15/70 \approx 22\%.$$

3. La verosimiglianza questa volta è

$$L(p) = q^3(1-q)^2,$$

quindi i calcoli sono gli stessi e si ottiene la stessa stima.

Problema 2

La durata di un dispositivo è rappresentata da una variabile esponenziale avente media 12 mesi se funzionante, mentre la media scende a 1 mese se difettoso. Si prendono due dispositivi, uno funzionante e uno difettoso, e si mettono a confronto installandoli contemporaneamente. Poniamo T_F la durata del dispositivo funzionante, T_D quella del dispositivo difettoso (si sa quale è difettoso e quale funzionante). Supponiamo che i due tempi siano indipendenti e poniamo

$$M = \max\{T_F, T_D\}, \quad N = \min\{T_F, T_D\}.$$

1. Determinare la densità della variabile N, calcolarne il valor medio e la deviazione standard.
2. Determinare la densità della variabile M, calcolarne il valor medio e la deviazione standard.
3. Le due variabili M, N sono indipendenti?

Una soluzione:

1. Sappiamo dal corso che il minimo di variabili esponenziali indipendenti è una esponenziale di parametro avente la somma dei parametri, quindi in questo caso N è esponenziale di parametro $1/1 + 1/12 = 13/12$ (le medie sono l'inverso del parametro). Quindi il valor medio è $12/13$, la varianza è $12/13$, la deviazione standard è $\sqrt{12/13}$.

2. Ricordiamo che il massimo di variabili indipendenti la CDF data dal prodotto delle due CDF. Quindi in questo caso otteniamo (per $t > 0$)

$$\mathrm{CDF}_M(t) = \mathrm{CDF}_{T_F}(t)\,\mathrm{CDF}_{T_D}(t) = (1-e^{-t})(1-e^{-t/12}).$$

La densità si ottiene derivando (per $t > 0$, altrimenti è nulla)

$$\begin{aligned} p(M = t) &= \frac{d}{dt}\mathrm{CDF}_M(t) = e^{-t}(1-e^{-t/12}) + \frac{1}{12}e^{-t/12}(1-e^{-t}) \\ &= e^{-t} + \frac{1}{12}e^{-t/12} - \frac{13}{12}e^{-t13/12}. \end{aligned}$$

Per calcolare il valor medio, scriviamo l'integrale

$$\begin{aligned}\mathbb{E}[M] &= \int_0^\infty tp(M=t)dt = \int_0^\infty t\left[e^{-t} + \frac{1}{12}e^{-t/12} - \frac{13}{12}e^{-t13/12}\right]dt \\ &= 1 + 12 - \frac{12}{13} = 13 - \frac{12}{13} \approx 12,08\end{aligned}$$

Per la varianza,

$$\begin{aligned}\mathbb{E}\left[M^2\right] &= \int_0^\infty t^2 p(M=t)dt = \int_0^\infty t^2\left[e^{-t} + \frac{1}{12}e^{-t/12} - \frac{13}{12}e^{-t13/12}\right]dt \\ &= 2\cdot\left(1^2 + 12^2 - \left(\frac{12}{13}\right)^2\right) \approx 2\cdot 169,14.\end{aligned}$$

Per la varianza, otteniamo quindi

$$\mathrm{Var}(M) = \mathbb{E}\left[M^2\right] - \mathbb{E}[M]^2 \approx 192,$$

da cui $\sigma_M \approx 13,9$.

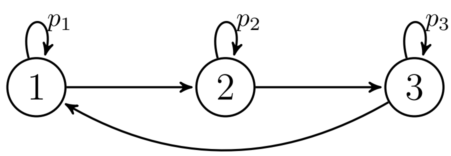
3. No, le variabili non sono indipendenti. Basta notare ad esempio che

$$P(M<1, N>1) = 0,$$

mentre $P(M<1) = \int_0^1 p(M=t)dt > 0$, $P(N>1) = \int_1^\infty p(N=t)dt > 0$, perché le densità sono strettamente positive per $t>0$.

Problema 3

Si consideri una catena di Markov $(X_n)_{n=0}^\infty$ con matrice di transizione rappresentata mediante il seguente grafo, dove $p_1, p_2, p_3 \in [0,1]$ sono parametri.



1. Al variare dei parametri p_1, p_2, p_3, classificare gli stati (transitori/ricorrenti), determinare tutte le classi chiuse irriducibili, e dire se la catena è regolare.
2. Mostrare che la distribuzione uniforme sui tre stati è invariante se e solo se $p_1 = p_2 = p_3$. In quali casi è l'unica distribuzione invariante?
3. Supponendo $p_1 = p_2 = 1/2$, $p_3 = 1$, calcolare $P(X_3 = 2 | X_0 = 1, X_6 = 3)$.

Una soluzione:

1. Se tutti i parametri sono strettamente minori di 1, allora la catena è irriducibile (tutti gli stati comunicano tra loro). Se inoltre almeno uno è strettamente positivo, allora la

catena è anche regolare per il criterio. Se sono tutti nulli, la catena non è regolare (perché $Q^3 = I$ la matrice identica). Se alcuni parametri valgono 1, allora gli stati corrispondenti sono assorbenti, quindi ricorrenti e ciascuno (preso da solo) costituisce una classe chiusa irriducibile. I rimanenti stati sono transitori.

2. Se $p_1 = p_2 = p_3$ allora la distribuzione uniforme è invariante, ad esempio soddisfa il bilancio di flusso. Se viceversa, supponiamo che la distribuzione uniforme sia invariante. Scrivendo il bilancio di flusso, si ottiene

$$\frac{1}{3}(1-p_1) = \frac{1}{3}(1-p_3), \quad \frac{1}{3}(1-p_2) = \frac{1}{3}(1-p_1) \quad \frac{1}{3}(1-p_3) = \frac{1}{3}(1-p_2).$$

Si trova

$$1-p_1 = 1-p_3, \quad 1-p_2 = 1-p_1 \quad 1-p_3 = 1-p_2,$$

da cui $p_1 = p_2 = p_3$. Questa è l'unica distribuzione invariante se e solo se $p_1 = p_2 = p_3 < 1$ (la catena è irriducibile), altrimenti se $p_1 = p_2 = p_3$ ci sono infinite distribuzioni invarianti (tutte le distribuzioni sono invarianti).

3. Scriviamo in modo equivalente,

$$\begin{aligned} P(X_3 = 2 | X_0 = 1, X_6 = 3) &= \frac{P(X_3 = 2, X_6 = 3 | X_0 = 1)}{P(X_6 = 3 | X_0 = 1)} \\ &= \frac{P(X_6 = 3 | X_3 = 2) P(X_3 = 2 | X_0 = 1)}{P(X_6 = 3 | X_0 = 1)}, \end{aligned}$$

avendo usato la proprietà di Markov. Per il denominatore, si tratta al solito di elencare i cammini che partono da 1 e in 6 passi (o meno) raggiungono lo stato 3 :

$$1 \to 2 \to 3 \quad \to \frac{1}{2^2}$$

$$1 \to 1 \to 2 \to 3 \quad \to \frac{1}{2^3}$$

$$1 \to 2 \to 2 \to 3 \quad \to \frac{1}{2^3}$$

$$1 \to 1 \to 1 \to 2 \to 3 \quad \to \frac{1}{2^4}$$

$$1 \to 1 \to 2 \to 2 \to 3 \quad \to \frac{1}{2^4}$$

$$1 \to 2 \to 2 \to 2 \to 3 \quad \to \frac{1}{2^4}$$

$$1 \to 1 \to 1 \to 1 \to 2 \to 3 \quad \to \frac{1}{2^5}$$

$$1 \to 1 \to 1 \to 2 \to 2 \to 3 \quad \to \frac{1}{2^5}$$

$$1 \to 1 \to 2 \to 2 \to 2 \to 3 \quad \to \frac{1}{2^5}$$

$$1 \to 2 \to 2 \to 2 \to 2 \to 3 \quad \to \frac{1}{2^5}$$

$$1 \to 1 \to 1 \to 1 \to 1 \to 2 \to 3 \quad \to \frac{1}{2^6}$$

$$1 \to 1 \to 1 \to 1 \to 2 \to 2 \to 3 \quad \to \frac{1}{2^6}$$

$$1 \to 1 \to 1 \to 2 \to 2 \to 2 \to 3 \quad \to \frac{1}{2^6}$$

$$1 \to 1 \to 2 \to 2 \to 2 \to 2 \to 3 \quad \to \frac{1}{2^6}$$

$$1 \to 2 \to 2 \to 2 \to 2 \to 2 \to 3 \quad \to \frac{1}{2^6}$$

Sommando tutte le probabilità, si trova

$$P(X_6 = 3 | X_0 = 1) = \frac{1}{2^2} + \frac{2}{2^3} + \frac{3}{2^4} + \frac{4}{2^5} + \frac{5}{2^6}.$$

Per il numeratore, calcoliamo i cammini

$$1 \to 1 \to 1 \to 2$$

$$1 \to 1 \to 2 \to 2$$

$$1 \to 2 \to 2 \to 2$$

da cui

$$P(X_3 = 2 | X_0 = 1) = \frac{3}{2^3}.$$

Per l'altro termine, usando l'omogeneità si tratta di calcolare

$$P(X_6 = 3 | X_3 = 2) = P(X_3 = 3 | X_0 = 2),$$

che corrisponde ai cammini

$$2 \to 2 \to 2 \to 3$$

$$2 \to 2 \to 3$$

$$2 \to 3$$

e quindi

$$P(X_6 = 3 | X_3 = 2) = \frac{1}{2^3} + \frac{1}{2^2} + \frac{1}{2}.$$

In conclusione

$$P(X_3 = 2 | X_0 = 1, X_6 = 3) = \frac{\frac{3}{2^3} \cdot \left(\frac{1}{2^3} + \frac{1}{2^2} + \frac{1}{2}\right)}{\frac{1}{2^2} + \frac{2}{2^3} + \frac{3}{2^4} + \frac{4}{2^5} + \frac{5}{2^6}} \approx 37\%.$$