Prova d'esame

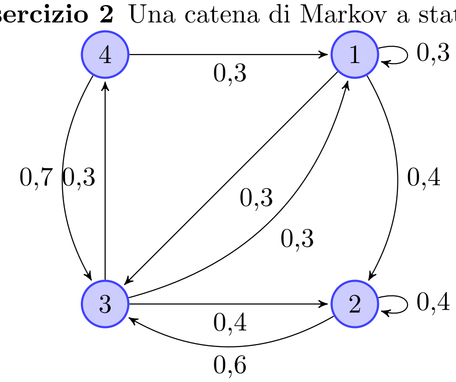
Calcolo delle Probabilità e Statistica (12 crediti)

7 luglio 2014

Esercizio 1 Un tipografo prepara un libro di $N = 200$ pagine, contenenti ognuna $n = 800$ caratteri. Ogni carattere, indipendentemente dagli altri, può essere sbagliato con probabilità $p = 0.001$. **(a)** Utilizzando l'approssimazione alla binomiale con leggi di Poisson, calcolare la probabilità π_0 che una pagina fissata sia senza errori. **(b)** Utilizzando il valore π_0 determinato in precedenza, calcolare il numero medio di pagine del libro che contengono errori.

Soluzione. Indichiamo con $X \sim Po(\lambda)$ la variabile aleatoria che conta il numero di errori in una pagina fissata. Si sceglie $\lambda = np = 0.8$ e si ottiene $\pi_0 = \mathbb{P}(X = 0) = e^{-\lambda} = 0.449$. Indichiamo con Y_k la variabile aleatoria che vale 0 se la pagina k è corretta, vale 1 se la pagina k contiene almeno un errore. Risulta che il numero di pagine con errori è pari alla somma $Y_1 + \cdots + Y_N$; dato che $\mathbb{E}[Y_k] = 1 - \pi_0$, si ottiene che il numero medio di pagine con errori è $N(1 - \pi_0) = 110.2$.

Esercizio 2 Una catena di Markov a stati discreti ha il seguente grafo



(a) Scrivere la matrice di transizione del sistema.
(b) Determinare le caratteristiche del sistema: è irriducibile? periodico? regolare? Esiste una misura invariante? è unica? nel caso, determinarla.
(c) Se al tempo 0 siamo in 1, qual è la probabilità di trovarsi in 3 al tempo $n = 2$?
(d) Se al tempo 0 siamo in 1, qual è la probabilità di trovarsi in 3 al tempo $n = 1234$?

Soluzione. La matrice di transizione del sistema è

$$T = \begin{pmatrix} 0.3 & 0.4 & 0.3 & 0 \\ 0 & 0.4 & 0.6 & 0 \\ 0.3 & 0.4 & 0 & 0.3 \\ 0.3 & 0 & 0.7 & 0 \end{pmatrix}$$

Il sistema è irriducibile, aperiodico, regolare, ergodico; tutti gli stati sono ricorrenti positivi; la distribuzione invariante è anche limite e vale $\pi^* = (\frac{117}{608}, \frac{218}{608}, \frac{210}{608}, \frac{63}{608})$. Si ottiene $p^{(2)}(1,3) = 0,33$; inoltre, dato che il sistema è regolare, $p^{(1234)}(1,3) \simeq \pi^*(3) = \frac{210}{608}$.

Esercizio 3 Il test Elisa per verificare la presenza di anticorpi Hiv Ab ha un livello di sensibilità (probabilità di individuare la presenza del virus se presente) pari a 99% e una attendibilità (probabilità di rivelare la presenza del virus se non presente) pari al 0.5%. Nella popolazione di età superiore ai 15 anni (dati: Regione Piemonte) la prevalenza dell'infezione risulta pari a circa 2 casi ogni 1.000 abitanti.

1. Determinare la probabilità che una persona abbia effettivamente contratto il virus sapendo che il test ne ha indicato la presenza.

2. Determinare la probabilità che una persona non abbia contratto il virus sapendo che il test ne ha indicato l'assenza.

Soluzione. I dati del problema affermano che la probabilità a priori di essere sani è $\mathbb{P}(S) = 0.998$; la probabilità che una persona sana risulti negativa al test è pari a $\mathbb{P}(T- \mid S) = 0.995$ mentre la probabilità che una persona malata risulti positiva è pari a $\mathbb{P}(T+ \mid M) = 0.99$; possiamo allora calcolare la probabilità di essere positivi al test

$$\mathbb{P}(T+) = \mathbb{P}(T+ \mid S)\mathbb{P}(S) + \mathbb{P}(T+ \mid M)\mathbb{P}(M) = 0.005 \cdot 0.998 + 0.99 \cdot 0.002 = 0,697\%$$

la probabilità di essere malati sapendo di essere risultati positivi al test è

$$\mathbb{P}(M \mid T+) = \frac{\mathbb{P}(T+ \mid M)\mathbb{P}(M)}{\mathbb{P}(T+)} = \frac{198}{697} \simeq 0.284$$

la probabilità di essere sani sapendo di essere risultati negativi al test è

$$\mathbb{P}(S \mid T-) = \frac{\mathbb{P}(T- \mid S)\mathbb{P}(S)}{\mathbb{P}(T-)} = \frac{99301}{99303} \simeq 99.998\%$$

Esercizio 4 Sia data la distribuzione congiunta delle variabili aleatorie (X, Y)

$$f(x,y) = \begin{cases} 6e^{-2x-3y} & \text{per } x > 0 \text{ e } y > 0 \\ 0 & \text{altrove} \end{cases}$$

Si determinino le leggi marginali di X e Y, le loro media e varianza. X e Y sono indipendenti? Si determini infine la legge della variabile $T = \max(X, Y)$.

Soluzione. Si nota che

$$f(x,y) = 2e^{-2x} \cdot 3e^{-3y} = f_X(x) f_Y(y)$$

si fattorizza come prodotto di due densità di probabilità rispettivamente nella variabile x e y; ne risulta allora che $f_X(x)$ e $f_Y(y)$ sono le densità di probabilità delle due marginali. Riconosciamo in esse delle leggi esponenziali di parametro $\lambda_X = 2$ e $\lambda_Y = 3$ da cui

$$\mathbb{E}[X] = \frac{1}{2}, \qquad \mathbb{E}[Y] = \frac{1}{3}, \qquad \mathrm{Var}(X) = \frac{1}{4}, \qquad \mathrm{Var}(Y) = \frac{1}{9}.$$

Inoltre, siccome la legge congiunta è il prodotto delle leggi marginali, le variabili X e Y risultano indipendenti.

Infine, osserviamo che $T = \max(X, Y)$ verifica

$$(T \leq \alpha) = (X \leq \alpha, Y \leq \alpha)$$

da cui

$$\mathbb{P}(T \leq \alpha) = \mathbb{P}(X \leq \alpha, Y \leq \alpha) = \mathbb{P}(X \leq \alpha)\mathbb{P}(Y \leq \alpha) = (1 - e^{-2\alpha})(1 - e^{-3\alpha})$$

la densità di T è espressa dalla derivata della formula precedente, ossia

$$f_T(t) = 2e^{-2t} + 3e^{-3t} - 5e^{-5t}$$

da cui si ottiene $\mathbb{E}[T] = \frac{19}{30}$, $\mathbb{E}[T^2] = \frac{578}{900}$ e $\mathrm{Var}(T) = \frac{217}{900}$.

Esercizio 5 Consideriamo una classe di studenti. Essi vengono divisi in cinque gruppi di 11, 15, 18, 20 e 16 studenti che hanno altezza compresa, rispettivamente, negli intervalli $(155, 165)$, $(165, 171)$, $(171, 177)$, $(177, 183)$ e $(183, 191)$. **(a)** Tracciare un istogramma dei dati. **(b)** Determinare media e varianza campionaria dell'altezza degli studenti. **(c)** Supponendo che l'altezza sia una variabile aleatoria con distribuzione gaussiana, determinare un intervallo di confidenza al livello $\alpha = 95\%$ per la media della classe (utilizzare, come stima della varianza, quanto trovato nel punto (b)).

Soluzione. Media e varianza campionaria:

$$\mu = 175.05$$
$$s^2 = 76.3519$$

Con questi dati, si ottiene l'intervallo di confidenza

$$\left(\mu - t_{0.975}^{(79)}\sqrt{\frac{s^2}{n}},\ \mu + t_{0.975}^{(79)}\sqrt{\frac{s^2}{n}}\right) = \big(173.102,\ 176.998\big).$$