Prova scritta di

Calcolo delle probabilità e Statistica (12 crediti)

17 gennaio 2007

Esercizio 1.

Le due simpatiche formichine α ed ω sono costrette a vivere in una zona quadrata del piano Q, di vertici $P = (2,0)$, $Q = (4,0)$, $R = (2,2)$ e $S = (4,2)$. La loro posizione è distribuita in modo uniforme all'interno del suddetto quadrato. Sotto l'azione di un potente anestetico α e ω cessano, nel medesimo istante T, di muoversi. Definite le v.a. α_X, α_Y e ω_X, ω_Y come l'ascissa e l'ordinata che determinano la posizione di α (risp. ω), all'interno di Q, determinare (al tempo T):

a. La distribuzione di probabilità di $\alpha_X, \alpha_Y, \omega_X, \omega_Y$

b. La distribuzione di probabilità della distanza orizzontale (risp. verticale) definita come:

$$D_X \equiv \mid \alpha_X - \omega_X \mid \qquad (\text{risp. } D_Y \equiv \mid \alpha_Y - \omega_Y \mid)$$

c. La distribuzione di probabilità della distanza tra le due formichine, definita come:

$$D \equiv D_X + D_Y.$$

Soluzione. Le due formiche determinano due vettori aleatori indipendenti ed uniformemente distribuiti, aventi marginali

$$f_{\alpha_X}(t) = f_{\omega_X}(t) = \frac{1}{2}\mathbf{1}_{[2,4]}(t), \qquad f_{\alpha_Y}(t) = f_{\omega_Y}(t) = \frac{1}{2}\mathbf{1}_{[0,2]}(t).$$

Indichiamo con $D_X = |\alpha_X - \omega_X|$: allora D_X prende valori in $[0,2]$ e per $t \in [0,2]$ si ha

$$\Pr(D_X > t) = \frac{1}{4}(2-t)^2.$$

Si osserva inoltre che D_Y ha la stessa distribuzione di D_X e che le due variabili aleatorie sono indipendenti. Dalla indipendenza segue che $D = D_X + D_Y$ ha distribuzione in $[0,4]$ che è pari alla convoluzione delle due leggi; si ha

$$\Pr(D \le t) = \int_0^t f_{\alpha_X}(s) F_{\alpha_Y}(t-s)\,\mathrm{d}s = \dots$$

Esercizio 2. Nella trasmissione di una immagine il colore di un pixel è descritto da un vettore di lunghezza 8 bit (ognuno dei quali assume il valore 0 oppure 1). Durante la decodifica dell'immagine, ogni bit può essere distorto con probabilità $p = 2 \cdot 10^{-4}$, indipendentemente da uno all'altro.

a. Qual è la probabilità α che un singolo pixel sia trasmesso correttamente?

b. Un'immagine è composta da 1306×728 pixel. Qual è il numero medio di pixel distorti nell'immagine?

b. Un' ulteriore stima per il numero di pixel distorti può essere ricavata utilizzando il *Teorema del Limite Centrale*. Se l'immagine è composta da 1306×728 pixel, qual è la probabilità che vi siano più di 1500 pixel distorti?

Soluzione. Indichiamo con π la probabilità di trasmissione corretta di un pixel. Risulta $\pi = 1 - (1-p)^8 = 0,9984$; il numero medio di pixel distorti è $N\pi$, dove $N = 1306 \cdot 728$, da cui $N\pi = 1520,16$. Usando l'approssimazione normale, indicando con S_N il numero totale di pixel distorti, si ottiene

$$\Pr(S_N > 1500) \simeq \Phi(\frac{N\pi - 1500}{\sqrt{N\pi(1-\pi)}}) \simeq \Phi(0,52) = 0,698$$

Esercizio 3.

Risolvere, giustificando opportunamente ogni risposta, i seguenti quesiti:

a. In un'urna sono contenute 3 palline verdi, 2 rosse, 4 nere ed 1 bianca. Effettuo due estrazioni (senza reimmissione) in serie. Calcolare la probabilità di avere una pallina nera alla prima estrazione, sapendo che ho estratto una pallina verde alla seconda.

b. Giochiamo con un dado, a sei facce, non truccato. È più facile ottenere almeno un sei su 4 lanci, oppure almeno un doppio sei su 24 lanci?

c. Un' urna contiene 5 monete. Due di queste sono *oneste* e presentano i due simboli **T** e **C** sulle opposte facce. Le rimanenti tre monete sono invece truccate. Di queste tre, infatti, due hanno il simbolo **T** su entrambe i lati, l'ultima presenta invece due simboli **C** sulle due facce. Si sceglie, a caso, una delle cinque monete e la si lancia osservando, come risultato, il simbolo **T**. Qual è la probabilità di aver pescato una moneta onesta?

Soluzione. a. $\Pr(1_N \mid 2_V) = \frac{\Pr(1_N \cap 2_V)}{\Pr(2_V)} = \frac{\frac{4}{10}\frac{3}{9}}{\frac{3}{10}} = \frac{4}{9}$.

b. La probabilità di ottenere un 6 lanciando 4 volte un dado è $1 - (1 - \frac{1}{6})^4 \simeq 0,5177$, la probabilità di ottenere una coppia di 6 lanciando 24 volte 2 dadi è $1 - (1 - \frac{1}{36})^{24} \simeq 0,4914$ quindi è inferiore.

c. La probabilità cercata si ottiene da una applicazione della formula di Bayes ed è

$$\Pr(\text{onesta} \mid \mathbf{T}) = \frac{\Pr(\mathbf{T} \mid \text{onesta}) \Pr(\text{onesta})}{\Pr(\mathbf{T})} = \frac{\frac{1}{2}\frac{2}{5}}{\frac{1}{2}\frac{2}{5} + 1 \cdot \frac{2}{5} + 0 \cdot \frac{1}{5}} = \frac{1}{3}.$$

Esercizio 4.

Un'urna contiene inizialmente 2 palline rosse e 3 palline nere. Si effettuano ripetutamente estrazioni dall'urna; ad ogni estrazione, se esce una pallina nera essa viene messa "in panchina", se esce una pallina rossa viene reintrodotta insieme ad una nera (se ve ne sono "in panchina").

a. Mostrare che il sistema può essere modellato attraverso una catena di Markov.

b. Determinare la matrice di transizione e descrivere gli stati del sistema.

c. Determinare la probabilità che dopo l'estrazione n-esima non vi siano palline nere nell'urna, con $n = 1, 2, 3$

d. Mostrare che il sistema è regolare. Determinare la misura invariante e calcolare il numero medio di palline presenti nell'urna a regime.

Soluzione. In ogni istante, lo stato del sistema è definito dal numero di palline nere nell'urna (le palline rosse rimangono 2, le altre nere sono in panchina) e la variazione del numero di palline nell'urna dipende solo dallo stato attuale e non dalla storia (le probabilità di estrarre rosso, risp. nero, dipendono solo dal numero di palline nere nell'urna). La matrice di transizione è

$$\begin{pmatrix} 0 & 1 & 0 & 0 \\ 1/3 & 0 & 2/3 & 0 \\ 0 & 1/2 & 0 & 1/2 \\ 0 & 0 & 3/5 & 2/5 \end{pmatrix}$$

e lo stato iniziale è $\pi_0 = \begin{pmatrix} 0 & 0 & 0 & 1 \end{pmatrix}$, da cui risulta $\Pr(X_1 = 0 \mid \pi_0) = \Pr(X_2 = 0 \mid \pi_0) = 0$, $\Pr(X_3 = 0 \mid \pi_0) = 1/10$. La misura invariante è $\pi = \{3/34, 9/34, 12/34, 10/34\}$ e la media a regime vale $\mathbb{E}[2 + \pi] = 131/34 \simeq 3,85$.

Esercizio 5.
Un carattere è distribuito in una popolazione secondo la legge

$$f(x \mid \theta) = \begin{cases} \frac{\theta}{2} & x = \pm 1 \\ 1 - \theta & x = 0. \end{cases}$$

Si effettua un campionamento di ampiezza $n = 10$, da cui si ottengono 3 valori $+1$, 4 valori -1 e i rimanenti 0. Determinare la stima di massima verosimiglianza di θ.

Soluzione. Indichiamo con $\bar{X} = (1, 1, 1, -1, -1, -1, -1, 0, 0, 0)$ il vettore delle osservazioni. Scriviamo la funzione di massima verosimiglianza

$$L(\theta \mid \bar{X}) = \left(\tfrac{\theta}{2}\right)^7 (1 - \theta)^3$$

e cerchiamo il massimo di questa funzione che si ottiene per $\theta = \frac{7}{10}$.