Università di Bologna - Scuola di Scienze

Esame scritto di Calcolo delle Probabilità e Statistica

12 gennaio 2024

Esercizio 1

Si consideri un'urna contenente 6 palline, di cui 4 sono di color rosso e 2 di color nero. Viene effettuata una serie di 3 estrazioni con reimbussolamento, con la seguente particolarità: quando una pallina viene estratta, essa viene reimbussolata assieme ad una nuova pallina dello stesso colore. Tale meccanismo è detto *Urna di Polya*.

Si determini:

1) la probabilità dell'evento

$$R_3 = \text{ “viene estratta una pallina rossa alla terza estrazione”};$$

2) la probabilità dell'evento

$$A = \text{ “dopo la terza estrazione l'urna contiene 6 palline rosse”};$$

3) la probabilità dell'evento

$$R_1 = \text{ “viene estratta una pallina rossa alla prima estrazione”}$$

condizionata all'evento

$$N_3 = \text{ “viene estratta una pallina nera alla terza estrazione”};$$

4) un evento B tale che $\mathbb{P}(N_3|B) = \mathbb{P}(R_3|B)$.

Soluzione

Denotiamo generalmente

$$\begin{aligned} R_i &= \text{ “viene estratta una pallina rossa alla } i\text{-esima estrazione”}, \\ N_i &= \text{ “viene estratta una pallina nera alla } i\text{-esima estrazione”}. \end{aligned}$$

1. Esaminando l'albero relativo all'esperimento, decomponiamo l'evento R_3 come

$$R_3 = \{R_1 \cap R_2 \cap R_3\} \cup \{R_1 \cap N_2 \cap R_3\} \cup \{N_1 \cap R_2 \cap R_3\} \cup \{N_1 \cap N_2 \cap R_3\}.$$

Perciò, utilizzando la regola della catena otteniamo

$$\begin{aligned} \mathbb{P}(R_3) &= \mathbb{P}(R_1) \times \mathbb{P}(R_2|R_1) \times \mathbb{P}(R_3|R_1 \cap R_2) \\ &\quad + \mathbb{P}(R_1) \times \mathbb{P}(N_2|R_1) \times \mathbb{P}(R_3|R_1 \cap N_2) \\ &\quad + \mathbb{P}(N_1) \times \mathbb{P}(R_2|N_1) \times \mathbb{P}(R_3|N_1 \cap R_2) \\ &\quad + \mathbb{P}(N_1) \times \mathbb{P}(N_2|N_1) \times \mathbb{P}(R_3|N_1 \cap N_2) \\ &= \frac{4}{6} \times \frac{5}{7} \times \frac{6}{8} + \frac{4}{6} \times \frac{2}{7} \times \frac{5}{8} + \frac{2}{6} \times \frac{4}{7} \times \frac{5}{8} + \frac{2}{6} \times \frac{3}{7} \times \frac{4}{8} = \frac{2}{3}. \end{aligned}$$

2. Si noti che possiamo scrivere

$$A = \{R_1 \cap R_2 \cap N_3\} \cup \{R_1 \cap N_2 \cap R_3\} \cup \{N_1 \cap R_2 \cap R_3\}$$

Perciò, utilizzando la regola della catena otteniamo

$$\begin{aligned}\mathbb{P}(A) &= \mathbb{P}(R_1) \times \mathbb{P}(R_2|R_1) \times \mathbb{P}(N_3|R_1 \cap R_2)\\ &\quad + \mathbb{P}(R_1) \times \mathbb{P}(N_2|R_1) \times \mathbb{P}(R_3|R_1 \cap N_2)\\ &\quad + \mathbb{P}(N_1) \times \mathbb{P}(R_2|N_1) \times \mathbb{P}(R_3|N_1 \cap R_2)\\ &= \frac{4}{6} \times \frac{5}{7} \times \frac{2}{8} + \frac{4}{6} \times \frac{2}{7} \times \frac{5}{8} + \frac{2}{6} \times \frac{4}{7} \times \frac{5}{8} = \frac{5}{14}.\end{aligned}$$

3. Dal punto 1 si ottiene $\mathbb{P}(N_3) = 1 - \mathbb{P}(R_3) = 1/3$. Inoltre, decomponendo e utilizzando sempre la regola della catena, si ha

$$\begin{aligned}\mathbb{P}(N_3|R_1) &= \mathbb{P}(R_2|R_1)\mathbb{P}(N_3|R_1 \cap R_2) + \mathbb{P}(N_2|R_1)\mathbb{P}(N_3|R_1 \cap N_2)\\ &= \frac{5}{7} \times \frac{2}{8} + \frac{2}{7} \times \frac{3}{8} = \frac{2}{7}.\end{aligned}$$

Quindi, utilizzando la formula di Bayes, si ottiene:

$$\mathbb{P}(R_1|N_3) = \frac{\mathbb{P}(N_3|R_1)\mathbb{P}(R_1)}{\mathbb{P}(N_3)} = \frac{\frac{2}{7} \times \frac{2}{3}}{\frac{1}{3}} = \frac{4}{7}.$$

4. Poniamo $B = N_1 \cap N_2$. Tale evento coincide con l'evento "dopo la seconda estrazione l'urna contiene 4 palline rosse e 4 palline nere". Pertanto

$$\mathbb{P}(R_3|B) = 1/2 = \mathbb{P}(N_3|B).$$

Esercizio 3

Siano X e Y due variabili aleatorie discrete con distribuzione discreta congiunta parzialmente data da:

$X \backslash Y$	-2	0	3	p_X
-1		0.2		
3	0.1			
p_Y		0.4		1

1) Si completi la tabella in modo che X e Y risultino indipendenti;

2) Calcolare $\mathbb{E}[XY]$;

3) Calcolare $\mathrm{var}(X^2 + 3Y)$;

4) sia Z una variabile aleatoria distribuita come X, indipendente sia da X che da Y. Si definisca la distribuzione congiunta del vettore (X, Y, Z) in modo che le tre variabili aleatorie X, Y, Z non siano indipendenti.

Soluzione

1. Si può ragionare come segue:

   - per complementarità, vale
   $$p_{(X,Y)}(3, 0) = p_Y(0) - p_{(X,Y)}(-1, 0) = 0.4 - 0.2 = 0.2;$$
   - per l'indipendenza, deve anche valere
   $$p_{(X,Y)}(3, 0) = p_X(3) \times p_Y(0)$$
   e quindi
   $$p_X(3) = 0.2/0.4 = 0.5;$$
   - di nuovo per l'indipendenza, deve valere
   $$p_{(X,Y)}(3, -2) = p_X(3) \times p_Y(-2)$$
   e quindi
   $$p_Y(-2) = 0.1/0.5 = 0.2;$$
   - per complementarità, vale
   $$p_Y(3) = 1 - p_Y(-2) - p_Y(0) = 1 - 0.2 - 0.4 = 0.4;$$
   - infine, ancora per l'indipendenza, vale
   $$\begin{aligned} p_{(X,Y)}(-1, -2) &= p_X(-1) \times p_Y(-2) = 0.2 \times 0.5 = 0.1, \\ p_{(X,Y)}(-1, 3) &= p_X(-1) \times p_Y(-3) = 0.4 \times 0.5 = 0.2. \end{aligned}$$

   Ricapitolando, si ottiene:

$X \backslash Y$	-2	0	3	p_X
-1	0.1	0.2	0.2	0.5
3	0.1	0.2	0.2	0.5
p_Y	0.2	0.4	0.4	1

.

2. Date le distribuzioni di X e Y, otteniamo

$$\begin{aligned}\mathbb{E}[X] &= -1 \times 0.5 + 3 \times 0.5 = 1,\\ \mathbb{E}[Y] &= -2 \times 0.2 + 0 \times 0.4 + 3 \times 0.4 = 0.8.\end{aligned}$$

Per l'indipendenza,

$$\mathbb{E}[XY] = \mathbb{E}[X] \times \mathbb{E}[Y] = 1 \times 0.8 = 0.8.$$

3. Sempre per l'indipendenza di X e Y, e per la proprietà di bilinearità della covarianza, otteniamo

$$\mathrm{var}(X^2 + 3Y) = \mathrm{var}(X^2) + 9\,\mathrm{var}(Y) = \mathbb{E}[X^4] - \mathbb{E}[X^2]^2 + 9\mathbb{E}[Y^2] - 9\mathbb{E}[Y]^2.$$

Quindi, calcolando

$$\begin{aligned}\mathbb{E}[X^2] &= 1 \times 0.5 + 9 \times 0.5 = 5,\\ \mathbb{E}[X^4] &= 1 \times 0.5 + 81 \times 0.5 = 41,\\ \mathbb{E}[Y^2] &= 4 \times 0.2 + 0 \times 0.4 + 9 \times 0.4 = 4.4.\end{aligned}$$

otteniamo

$$\mathrm{var}(X^2 + 3Y) = 41 - 25 + 9 \times 4.4 - 9 \times 0.64 = 49.84.$$

4. Notiamo innanzitutto che determinare la distribuzione di (X, Y, Z) è equivalente a determinare la distribuzione di (X, Z), condizionata agli eventi $\{Y = -2\}$, $\{Y = 0\}$ and $\{Y = 3\}$, ovvero a determinare le matrici

$$\begin{pmatrix} \mathbb{P}(X = -1, Z = -1 | Y = -2) & \mathbb{P}(X = -1, Z = 3 | Y = -2) \\ \mathbb{P}(X = 3, Z = -1 | Y = -2) & \mathbb{P}(X = 3, Z = 3 | Y = -2) \end{pmatrix} = \begin{pmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{pmatrix} = A,$$

$$\begin{pmatrix} \mathbb{P}(X = -1, Z = -1 | Y = 0) & \mathbb{P}(X = -1, Z = 3 | Y = 0) \\ \mathbb{P}(X = 3, Z = -1 | Y = 0) & \mathbb{P}(X = 3, Z = 3 | Y = 0) \end{pmatrix} = \begin{pmatrix} b_{11} & b_{12} \\ b_{21} & b_{22} \end{pmatrix} = B,$$

$$\begin{pmatrix} \mathbb{P}(X = -1, Z = -1 | Y = 3) & \mathbb{P}(X = -1, Z = 3 | Y = 3) \\ \mathbb{P}(X = 3, Z = -1 | Y = 3) & \mathbb{P}(X = 3, Z = 3 | Y = 3) \end{pmatrix} = \begin{pmatrix} c_{11} & c_{12} \\ c_{21} & c_{22} \end{pmatrix} = C.$$

Per la formula delle probabilità totali, vale

$$\begin{pmatrix} p_{(X,Z)}(-1,-1) & p_{(X,Z)}(-1,3) \\ p_{(X,Z)}(3,-1) & p_{(X,Z)}(3,3) \end{pmatrix} = \mathbb{P}(Y = -2)A + \mathbb{P}(Y = 0)B + \mathbb{P}(Y = 3)C.$$

D'altra parte, dato che Z deve essere distribuito come X ed essere indipendente da X, deve valere

$$\begin{pmatrix} p_{(X,Z)}(-1,-1) & p_{(X,Z)}(-1,3) \\ p_{(X,Z)}(3,-1) & p_{(X,Z)}(3,3) \end{pmatrix} = \begin{pmatrix} 0.25 & 0.25 \\ 0.25 & 0.25 \end{pmatrix}.$$

Perciò, deve essere soddisfatto il seguente sistema

$$\begin{cases} 0.2\,a_{11} + 0.4\,b_{11} + 0.4\,c_{11} = 0.25 \\ 0.2\,a_{12} + 0.4\,b_{12} + 0.4\,c_{12} = 0.25 \\ 0.2\,a_{21} + 0.4\,b_{21} + 0.4\,c_{21} = 0.25 \\ 0.2\,a_{22} + 0.4\,b_{22} + 0.4\,c_{22} = 0.25 \end{cases} . \tag{0.1}$$

Inoltre, dato che sia X deve essere indipendente da Y, deve valere

$$p_X(-1) = \mathbb{P}(X = -1|Y = -2) = \mathbb{P}(X = -1|Y = 0) = \mathbb{P}(X = -1|Y = 3),$$
$$p_X(3) = \mathbb{P}(X = 3|Y = -2) = \mathbb{P}(X = 3|Y = 0) = \mathbb{P}(X = 3|Y = 3),$$

che è equivalente ai sistemi

$$\begin{cases} a_{11} + a_{12} = 0.5 \\ a_{21} + a_{22} = 0.5 \end{cases}, \quad \begin{cases} b_{11} + b_{12} = 0.5 \\ b_{21} + b_{22} = 0.5 \end{cases}, \quad \begin{cases} b_{11} + b_{12} = 0.5 \\ b_{21} + b_{22} = 0.5 \end{cases}. \tag{0.2}$$

Ragionando allo stesso, per l'indipendenza di Z da Y, si ottiene

$$\begin{cases} a_{11} + a_{12} = 0.5 \\ a_{21} + a_{22} = 0.5 \end{cases}, \quad \begin{cases} b_{11} + b_{12} = 0.5 \\ b_{21} + b_{22} = 0.5 \end{cases}, \quad \begin{cases} b_{11} + b_{12} = 0.5 \\ b_{21} + b_{22} = 0.5 \end{cases}. \tag{0.3}$$

Risolvendo i sistemi (0.2)-(0.3), si ottiene

$$\begin{cases} a_{11} = a_{11} = 0.5 - a_{12} \\ a_{21} = a_{12} \end{cases}, \quad \begin{cases} b_{11} = b_{11} = 0.5 - b_{12} \\ b_{21} = b_{12} \end{cases}, \quad \begin{cases} c_{11} = c_{11} = 0.5 - c_{12} \\ c_{21} = c_{12} \end{cases}. \tag{0.4}$$

Ora, sostituendo queste relazioni nel sistema (0.2), si ottiene

$$a_{12} = \frac{0.25 - 0.4(b_{12} + c_{12})}{0.2}. \tag{0.5}$$

Abbiamo quindi una certa libertà nello scegliere $b_{1,2}$ e c_{12}. Poniamo ad esempio

$$b_{1,2} = 0, \qquad c_{12} = 0.5.$$

Notiamo che, con questa scelta, si ha

$$\mathbb{P}(X = -1, Z = 3|Y = 0) \neq \mathbb{P}(X = -1, Z = 3|Y = 3)$$

e quindi X, Y, Z non possono essere indipendenti[1]. Inoltre, da (0.5) si ottiene $a_{12} = 0.25$, e da (0.4) si ottiene infine

$$A = \begin{pmatrix} 0.25 & 0.25 \\ 0.25 & 0.25 \end{pmatrix}, \quad B = \begin{pmatrix} 0.5 & 0 \\ 0 & 0.5 \end{pmatrix}, \quad C = \begin{pmatrix} 0 & 0.5 \\ 0.5 & 0 \end{pmatrix}.$$

[1]Semplice esercizio

Esercizio 4

Un cliente arriva alle casse di un supermercato e può scegliere tra cassa A e cassa B. Scegliendo la cassa A, il suo tempo d'attesa T, misurato in minuti, avrà distribuzione distribuzione esponenziale con parametro 0.2, mentre scegliendo la cassa B, avrà sempre distribuzione esponenziale, ma con parametro 0.5.

Si assume che il cliente scelga la cassa in maniera casuale.

1) Determinare la funzione di ripartizione di T.

2) Calcolare $\mathbb{E}[T]$ e $\mathrm{var}(T)$.

3) Si determini se T gode della proprietà di perdita di memoria, della quale gode la distribuzione esponenziale. Ovvero, determinare se

$$\mathbb{P}(T > t + h | T > t) = \mathbb{P}(T > h), \qquad t, h > 0. \tag{0.6}$$

4) Se si assume che, in maniera indipendente da ciò che succede alle casse, si può verificare un black-out di 1 minuto con una probabilità dello 0.1%, durante il quale l'attività si ferma completamente, come cambia la funzione di ripartizione del tempo di attesa?

Soluzione

1. Definiamo gli eventi

$$A := \text{“ il cliente sceglie la cassa A”}, \quad B := A^c = \text{“ il cliente sceglie la cassa B”}.$$

Per la formula delle probabilità totali, otteniamo

$$\begin{aligned} F_T(t) = \mathbb{P}(T \le t) &= \mathbb{P}(T \le t | A) \times \mathbb{P}(A) + \mathbb{P}(T \le t | B) \times \mathbb{P}(B) \\ &= \frac{1}{2}\big(1 - e^{-0.2t} + 1 - e^{-0.5t}\big) = 1 - \frac{e^{-0.2t} + e^{-0.5t}}{2}, \qquad t \ge 0, \end{aligned}$$

mentre $F_T(t) = 0$ per ogni $t < 0$.

2. Calcoliamo prima la densità di T, derivando la funzione di ripartizione. Si ottiene

$$f_T(t) = \begin{cases} \frac{1}{2}\big(0.2\, e^{-0.2t} + 0.5\, e^{-0.5t}\big) & \text{se } t \ge 0, \\ 0 & \text{se } t < 0. \end{cases}$$

Ora, per definizione di valore atteso, abbiamo

$$\mathbb{E}[T] = \int_{\mathbb{R}} t\, f_T(t) dt = \frac{1}{2}\Big(\int_{\mathbb{R}} t\, 0.2\, e^{-0.2t} dt + \int_{\mathbb{R}} t\, 0.5\, e^{-0.5t} dt \Big) = \frac{1}{2}\Big(\frac{1}{0.2} + \frac{1}{0.5}\Big) = 3.5.$$

Notiamo che i due integrali sopra sono i valori attesi di due variabili aleatorie con distribuzione esponenziale di parametri uguali a 0.2 e 0.5 rispettivamente. Per calcolare la varianza, calcoliamo

$$\mathbb{E}[T^2] = \int_{\mathbb{R}} t^2\, f_T(t) dt = \frac{1}{2}\Big(\int_{\mathbb{R}} t^2\, 0.2\, e^{-0.2t} dt + \int_{\mathbb{R}} t^2\, 0.5\, e^{-0.5t} dt \Big) = \frac{1}{2}\Big(\frac{2}{0.04} + \frac{2}{0.25}\Big) = 29.$$

Quindi otteniamo

$$\mathrm{var}(T) = \mathbb{E}[T^2] - \mathbb{E}[T]^2 = 29 - 3.5^2 = 16.75.$$

3. Dal Punto 1. si ottiene

$$\mathbb{P}(T > h) = 1 - F_T(h) = \frac{e^{-0.2h} + e^{-0.5h}}{2}.$$

Analogamente,

$$\mathbb{P}(T > t + h, T > t) = \mathbb{P}(T > t + h) = 1 - F_T(t + h) = \frac{e^{-0.2(t+h)} + e^{-0.5(t+h)}}{2},$$

e quindi

$$\mathbb{P}(T > t + h | T > t) = \frac{e^{-0.2(t+h)} + e^{-0.5(t+h)}}{e^{-0.2t} + e^{-0.5t}}.$$

Si vede quindi che non vale la proprietà (0.6).

4. Introduciamo l'evento

$$E := \text{“Si verifica il black-out”}, \tag{0.7}$$

e la variabile aleatoria $T' = T + \mathbf{1}_E$, che denota il tempo di attesa con eventuale ritardo dovuto al black-out. Di nuovo per la formula delle probabilità totali otteniamo

$$\begin{aligned} F_{T'}(t) &= \mathbb{P}(T' \leq t | E) \times \mathbb{P}(E) + \mathbb{P}(T' \leq t | E^c) \times \mathbb{P}(E^c) \\ &= \mathbb{P}(T + \mathbf{1}_E \leq t | E) \times 10^{-4} + \mathbb{P}(T + \mathbf{1}_E \leq t | E^c) \times (1 - 10^{-4}) \\ &= \mathbb{P}(T + 1 \leq t | E) \times 10^{-4} + \mathbb{P}(T \leq t | E^c) \times (1 - 10^{-4}) \end{aligned}$$

(per l'ipotesi di indipendenza di E da T)

$$\begin{aligned} &= \mathbb{P}(T + 1 \leq t) \times 10^{-4} + \mathbb{P}(T \leq t) \times (1 - 10^{-4}) \\ &= F_T(t - 1) \times 10^{-4} + F_T(t) \times (1 - 10^{-4}). \end{aligned}$$

Quindi, per il Punto 1, otteniamo

$$\begin{cases} \frac{10^{-4}}{2}\left(2 - e^{-0.2(t-1)} - e^{-0.5(t-1)}\right) + \frac{1-10^{-4}}{2}\left(2 - e^{-0.2t} - e^{-0.5t}\right) & \text{se } t \geq 1, \\ \frac{1-10^{-4}}{2}\left(2 - e^{-0.2t} - e^{-0.5t}\right) & \text{se } 0 \leq t < 1, \\ 0 & \text{se } t < 0. \end{cases}$$