Esercizi di Calcolo delle Probabilità e Statistica

a.a.2023/2024
31/02/2024

Esercizio 1 *Sapendo che la probabilità che un uomo nasca affetto da daltonismo è del 7% mentre che ciò accada ad una donna è dello 0,5%, considerato un gruppo di N persone costituito da M uomini e F donne scelti a caso, determinare:*

1. *la probabilità che nel gruppo non ci siano persone daltoniche.*
2. *la probabilità che nel gruppo ci siano esattamente un uomo daltonico e nessuna donna daltonica.*
3. *Se Z indica il numero totale di persone daltoniche del gruppo. Calcolare il valore atteso di Z.*
4. *Scelta a caso una persona nel gruppo, qual e la probabilità che sia un uomo sapendo che si tratta di una persona daltonica.*

Soluzione: Sia $\{X_i\}_{i=1}^F$ la collezione di vv.aa. i.i.d. con distribuzione di Bernoulli di parametro f ciascuna delle quali indica se l'i-esima donna del gruppo è afflitta o meno da daltonismo, ovvero, per ogni $i = 1, .., F, X_i$ è la funzione indicatrice dell'evento $\{$l'i-esima donna del gruppo è daltonica$\}$. Allo stesso modo, indichiamo con $\{Y_j\}_{j=1}^M$ la collezione di vv.aa. i.i.d. con distribuzione di Bernoulli di parametro m, ciascuna delle quali indipendente da quelle della collezzione $\{X_i\}_{i=1}^F$, tali che Y_j indica se il j-esimo uomo del gruppo è afflitto o meno da daltonismo, ovvero, per ogni $j = 1, .., M$, Y_j è la funzione indicatrice dell'evento $\{$il j-esimo uomo del gruppo è daltonico$\}$.

1. L'evento

$$\{\text{nel gruppo non ci sono persone daltoniche}\} = \left(\bigcap_{i=1}^{F} \{X_i = 0\}\right) \cap \left(\bigcap_{j=1}^{M} \{Y_j = 0\}\right) ,$$

perciò la sua probabilità è $(1 - f)^F (1 - m)^M$.

2. Sia U la v.a. che conta il numero di uomini daltonici nel gruppo, cioè $U = \sum_{j=1}^M Y_i$. Poiché le vv.aa. $Y_1, .., Y_M$ sono i.i.d. $Ber(m), U \overset{d}{=} Bin(m, M)$. Analogamente, $V := \sum_{i=1}^F X_i \overset{d}{=} Bin(f, F)$ è la v.a. che conta il numero di donne daltoniche nel gruppo. Inoltre, poiché ogni

elemento di $\{X_i\}_{i=1}^F$ è stocasticamente indipendente da ogni elemento di $\{Y_j\}_{j=1}^M$, U e V sono tra loro stocasticamente indipendenti. Perciò,

$$\begin{aligned}&\{\text{nel gruppo c'è solo un uomo daltonico e nessuna donna daltonica}\}\\ &= \{U = 1\} \cap \{V = 0\}\end{aligned}$$

la cui probabilità è $Mm^1 (1-m)^{M-1} (1-f)^F$.

3. $Z = U + V$. Perciò, $\mathbb{E}[U+V] = \mathbb{E}[U] + \mathbb{E}[V] = Mm + Ff$.

4. Sia α la v.a. che indica che la persona scelta a caso tra quelle del gruppo sia uomo, ovvero α è la funzione indicatrice dell'evento $\{\text{la persona scelta a caso è uomo}\}$, che una v.a. di Bernoulli di parametro $p = \frac{M}{N}$. Analogamente, W la v.a. che indica se la persona scelta a caso dal gruppo sia o meno daltonica, ovvero W è la funzione indicatrice dell'evento $\{\text{la persona scelta a caso è daltonica}\}$. W è una v.a con distribuzione di Bernoulli di parametro q dove, per la formula delle probabilità totali,

$$q = \mathbb{P}\{W=1\} = \mathbb{P}(\{W=1\} \mid \{\alpha=1\})\,\mathbb{P}\{\alpha=1\} + \mathbb{P}(\{W=1\} \mid \{\alpha=0\})\,\mathbb{P}\{\alpha=0\} \ .$$

Ma poiché $\mathbb{P}(\{W=1\} \mid \{\alpha=1\})$ è la probabilità che la persona scelta sia daltonica sapendo che è donna e $\mathbb{P}(\{W=1\} \mid \{\alpha=0\})$ è la probabilità che la persona scelta sia daltonica sapendo che è uomo, otteniamo $q = fp + m(1-p)$. Allora, per la formula di Bayes,

$$\begin{aligned}\mathbb{P}(\{\alpha=0\} \mid \{W=1\}) &= \mathbb{P}(\{W=1\} \mid \{\alpha=0\})\,\frac{\mathbb{P}\{\alpha=0\}}{\mathbb{P}\{W=1\}}\\ &= \frac{m(1-p)}{fp + m(1-p)} \ .\end{aligned}$$

■

Esercizio 2 *Sia (X,Y) un vettore aleatorio uniformemente distribuito sull'insieme*

$$D := \left\{(x,y) \in \mathbb{R}^2 : |x| \le 1, |y| \le 1, x^2 + y^2 \ge 1\right\} \ .$$

1. *Calcolare la costante di normalizzazione della densità di probabilità congiunta.*

2. *Verificare se le componenti del vettore sono indipendenti e in caso contrario calcolare la densità della probabilità condizionata della componente Y rispetto alla componente X.*

3. *Calcolare la matrice di covarianza di (X,Y), ovvero*

$$\begin{pmatrix} Cov(X,X) & Cov(X,Y) \\ Cov(Y,X) & Cov(Y,Y) \end{pmatrix} \ .$$

Soluzione:

1. La costante di normalizzazione c della densità di probabilità di (X,Y) è pari all'inverso della differenza tra l'area del quadrato $Q := \left\{(x,y) \in \mathbb{R}^2 : |x| \le 1, |y| \le 1\right\}$ e quella del disco $C := \left\{(x,y) \in \mathbb{R}^2 : x^2 + y^2 \le 1\right\}$ ovvero $c = \frac{1}{4-\pi}$.

2. Le densità di probabilità marginali di (X,Y) sono

$$\begin{aligned} f_X(x) &= \mathbf{1}_{[-1,1]}(x)\left(\int_{-1}^{-\sqrt{1-x^2}} dy + \int_{\sqrt{1-x^2}}^{1} dy\right)\frac{1}{4-\pi} \\ &= \frac{2}{4-\pi}\left(1-\sqrt{1-x^2}\right) \end{aligned}$$

e, analogamente, per simmetria rispetto allo scambio degli assi coordinati,

$$f_Y(y) = \frac{2}{4-\pi}\left(1-\sqrt{1-y^2}\right)\mathbf{1}_{[-1,1]}(y) \ .$$

Da cui segie che le vv.aa. X e Y non sono tra loro stocasticamente indipendenti. Infatti, scegliendo per esempio il punto $\left(\frac{1}{2},\frac{1}{2}\right)$ si ottiene

$$f_X\left(\frac{1}{2}\right) f_Y\left(\frac{1}{2}\right) = \frac{4}{(4-\pi)^2}\left(1-\frac{\sqrt{3}}{2}\right)^2 \neq f_{(X,Y)}\left(\frac{1}{2},\frac{1}{2}\right) = 0 \ .$$

Allora,

$$f_{Y|X}(y|x) = \frac{f_{(X,Y)}(x,y)}{f_X(x)} = \frac{1}{2}\frac{\mathbf{1}_{\left[-1,-\sqrt{1-x^2}\right]}(y) + \mathbf{1}_{\left[\sqrt{1-x^2},1\right]}(y)}{\left(1-\sqrt{1-x^2}\right)} \ .$$

3. Poiché f_X è una funzione pari supportata su un intervallo simmetrico, mentre la funzione identità $(\mathbb{R} \ni x \longmapsto x \in \mathbb{R})$ è dispari, $\mathbb{E}[X]=0$.

$$\begin{aligned} \mathbb{E}\left[X^2\right] &= \frac{2}{4-\pi}\int_{-1}^{1} dx x^2\left(1-\sqrt{1-x^2}\right) = \\ &= \frac{2}{4-\pi}\left[\frac{2}{3} - \int_{-1}^{1} dx x^2\sqrt{1-x^2}\right] \\ &= \frac{2}{4-\pi}\left[\frac{2}{3}+\frac{\pi}{8}\right] \ , \end{aligned}$$

poiché

$$\begin{aligned} -\int_{-1}^{1} dx^2\sqrt{1-x^2} &= \int_{\pi}^{2\pi} dt \sin^2 t\cos^2 t = \frac{1}{4}\int_{\pi}^{2\pi} dt \sin^2(2t) \\ &= \frac{1}{8}\int_{2\pi}^{4\pi} du \sin^2 u = \frac{1}{16}\left(x - \sin x\cos x\right)\Big|_{2\pi}^{4\pi} = \frac{\pi}{8} \ . \end{aligned}$$

Analoghe considerazioni valgono per la v.a. Y.

$$\begin{aligned} Cov(X,Y) &= Cov(Y,X) = \mathbb{E}[XY] = \frac{1}{4-\pi}\int_D dxdyxy \\ &= \frac{1}{4-\pi}\left[\int_Q dxdyxy - \int_C dxdyxy\right] \\ &= \frac{1}{4-\pi}\left[\left(\int_{-1}^{1} dxx\right)^2 - \int_0^1 d\rho\rho^3 \int_0^{2\pi} d\theta \cos\theta \sin\theta\right] \\ &= -\frac{1}{4-\pi}\frac{1}{8}\int_0^{2\pi} d\theta \sin 2\theta = -\frac{1}{4-\pi}\frac{1}{16}\int_0^{4\pi} dt \sin t \\ &= \frac{1}{4-\pi}\frac{1}{16}\cos t|_0^{4\pi} = 0 \ . \end{aligned}$$

Perciò

$$\left(\begin{array}{cc} Cov(X,X) & Cov(X,Y) \\ Cov(Y,X) & Cov(Y,Y) \end{array}\right) = \left(\begin{array}{cc} \frac{\pi}{8} & 0 \\ 0 & \frac{\pi}{8} \end{array}\right) .$$

■

Esercizio 3 *Supponiamo di avere una successione di vv.aa.* $\{X_n\}_{n\geq 0}$ *sullo spazio di probabilità* $(\Omega,\mathcal{F},\mathbb{P})$ *che descrive una catena di Markov con 4 stati,* $S=\{1,2,3,4\}$ *tale che: dallo stato* 1 *si transisce con probabilità* $\frac{1}{2}$ *agli stati* 2 *e* 3. *Dallo stato* 2, *si transisce con probabilità* $\frac{1}{2}$ *agli stati* 1 *e* 4. *Dallo stato* 3 *si transisce con probabilità* $\frac{1}{2}$ *agli stati* 1 *e* 4. *Dallo stato* 4 *si transisce con probabilità* $\frac{1}{2}$ *agli stati* 2 *e* 3.

1. *Elencare le classi di equivalenza degli stati e calcolarne il periodo.*
2. *Calcolare* $\mathbb{P}(\{X_5=2\}\,|\,\{X_2=3\})$ *e, assumendo come distribuzione iniziale quella uniforme, calcolare* $\mathbb{E}[X_2]$.
3. *Se* $\{P_{i,j}\}_{i,j=1,..,4}$ *indica la matrice delle probabilità di transizione tra gli stati della catena, calcolare se esistono:* $\lim_{n\to\infty}\mathbb{P}\{\omega\in\Omega : X_n(\omega)=4\}$, $\lim_{n\to\infty}\left(P_{1,2}^n+P_{3,4}^n\right)$, $\lim_{n\to\infty}P_{2,3}^{2n}$.

Soluzione:

1. La matrice delle probabilità di transizione tra gli stati della catena è

$$P=\left(\begin{array}{cccc} 0 & \frac{1}{2} & \frac{1}{2} & 0 \\ \frac{1}{2} & 0 & 0 & \frac{1}{2} \\ \frac{1}{2} & 0 & 0 & \frac{1}{2} \\ 0 & \frac{1}{2} & \frac{1}{2} & 0 \end{array}\right)$$

perciò la catena può essere rappresentata mediante il grafo diretto

$$\begin{array}{ccc} 1 & \leftrightarrow & 2 \\ \updownarrow & & \updownarrow \\ 3 & \leftrightarrow & 4 \end{array}$$

da cui si deduce che esiste un unica classe di stati comunicanti. Inoltre, partendo dallo stato 1 si ritorna su di esso sempre con un numero pari di passi e quindi il periodo della classe è 2. Le sottoclassi cicliche dell'unica classe di stati comunicanti corrispondono quindi alle classi di stati comunicati della matrice delle probabilità di transizione

$$P^2 = \begin{pmatrix} \frac{1}{2} & 0 & 0 & \frac{1}{2} \\ 0 & \frac{1}{2} & \frac{1}{2} & 0 \\ 0 & \frac{1}{2} & \frac{1}{2} & 0 \\ \frac{1}{2} & 0 & 0 & \frac{1}{2} \end{pmatrix}$$

il cui grafo diretto associato è

$$\begin{array}{ccc} 1 & \leftrightarrows & 4 \\ \circlearrowright & & \circlearrowleft \\ 2 & \leftrightarrows & 3 \\ \circlearrowright & & \circlearrowleft \end{array}$$

da cui segue che la catena di Markov descritta da P^2 è decomponibile e ha due classi di stati comunicanti aperiodiche $\{1,4\}$ e $\{2.3\}$.

2. Poiché la catena di Markov è omogenea

$$\begin{aligned} \mathbb{P}(\{X_5 = 2\} \mid \{X_2 = 3\}) &= \mathbb{P}(\{X_3 = 2\} \mid \{X_0 = 3\}) \\ &= P^3_{3,2} = \sum_{i=1}^{4} P_{3,i} P^2_{i,2} = 0 \ . \end{aligned}$$

$$\begin{aligned} \mathbb{E}[X_2] &= \sum_{i,j=1}^{4} j \mathbb{P}(\{X_2 = j\} \mid \{X_0 = i\}) \mathbb{P}\{X_0 = i\} = \\ &= \sum_{i,j=1}^{4} \frac{1}{4} P^2_{i,j} j = \frac{1}{4} \sum_{j=1}^{4} j = \frac{5}{2} \ , \end{aligned}$$

dove nell'ultimo passaggio si è usato che, in questo caso, P^2 è bistocastica, cioè anche la somma sull'indice delle righe degli elementi di matrice di P^2 è pari a 1.

3. Poiché $\{2.3\}$ è una sottoclasse ciclica di periodo 2 di P considerando la catena di Markov indecomponibile e aperiodica costruita sugli stati 2 e 3 per mezzo degl elementi di matrice di P^2 che rappresentato le probabilità di transizione tra questi stati, ovvero considerando la catena di Markov definita dalla matrice delle probabilità di transizione

$$Q = \begin{pmatrix} Q_{22} & Q_{23} \\ Q_{32} & Q_{33} \end{pmatrix} = \begin{pmatrix} P^2_{22} & P^2_{23} \\ P^2_{32} & P^2_{33} \end{pmatrix} = \begin{pmatrix} \frac{1}{2} & \frac{1}{2} \\ \frac{1}{2} & \frac{1}{2} \end{pmatrix} ,$$

applicando il teorema ergodico a questa catena otteniamo che $\lim_{n\to\infty} P^{2n}_{2,3} = \lim_{n\to\infty} Q^n_{2,3} = \pi_3 = \frac{1}{2}$, dove $\pi := (\pi_2, \pi_3)$ è la distribuzione di probabilità invariante per Q, ovvero $\pi Q = \pi$. D'altra parte, poiché $P^n_{1,2}$ e $P^n_{3,4}$ sono diverse da zero soltanto se n è dispari poiché 1 e 2, come 3 e 4 appartengono a sottoclassi cicliche distinte, $\lim_{n\to\infty} \left(P^n_{1,2} + P^n_{3,4}\right)$ non esiste. Le stesse considerazioni valgono per $\lim_{n\to\infty}\mathbb{P}\{\omega\in\Omega : X_n(\omega)=4\}$ visto che

$$\begin{aligned}\mathbb{P}\{\omega\in\Omega : X_n(\omega)=4\} &= \sum_{i=1}^{4}\mathbb{P}\left(\{X_n(\omega)=4\}\,|\,\{X_0(\omega)=i\}\right)\mathbb{P}\{X_0(\omega)=i\} \\ &= \sum_{i=1}^{4}\mathbb{P}\{X_0(\omega)=i\}\,P^n_{i,4}\ .\end{aligned}$$

■

Esercizio 4 *Sia* $\{E_i\}_{i\geq 1}$ *una successione di eventi stocasticamenti indipendenti subordinatamente alla conoscenza di un parametro* Θ, *che si assume essere una v.a. la cui distribuzione di probabilità a priori è descritta dalla funzione di ripartizione (distribuzione)*

$$F_\Theta(\theta) := \theta^3\mathbf{1}_{[0,1]}(\theta)\ \ ,$$

tale che $\mathbb{P}(E_i|\Theta)=\Theta$. *Dopo le prime quattro osservazioni si registra che soltanto il primo evento non è accaduto.*

1. *Calcolare la densità di probabilità a posteriori di* Θ.
2. *Calcolare la probabilità a priori che* $\Theta\in\left[\frac{1}{2},1\right]$.
3. *Calcolare la probabilità a posteriori che* $\Theta\in\left[\frac{1}{2},1\right]$.
4. *Calcolare il valore massimo della densità di probabilità a posteriori di* Θ.

Soluzione:

1. Consideriamo le vv.aa. $\xi_i := \mathbf{1}_{E_i}, i=1,..,4$. La densità a posteriori di Θ dato il vettore aleatorio $(\xi_1,\xi_2,\xi_3,\xi_4)$ sarà

$$\begin{gathered}\pi_4\left(\Theta=\theta|E_1^c\cap E_2\cap E_3\cap E_4\right)= \\ \pi_4\left(\Theta=\theta|\left\{\omega\in\Omega:(\xi_1,\xi_2,\xi_3,\xi_4)=(0,1,1,1)\right\}\right)= \\ K\mathbb{P}\left(\left\{\omega\in\Omega:(\xi_1,\xi_2,\xi_3,\xi_4)=(0,1,1,1)\right\}|\Theta=\theta\right)\pi_0(\theta) \\ =K\mathbb{P}\left(E_1^c\cap E_2\cap E_3\cap E_4|\Theta=\theta\right)\pi_0(\theta) \\ =K\mathbb{P}\left(E_1^c|\Theta=\theta\right)\mathbb{P}\left(E_2|\Theta=\theta\right)\times \\ \times\mathbb{P}\left(E_3|\Theta=\theta\right)\mathbb{P}\left(E_4|\Theta=\theta\right)\pi_0(\theta)= \\ K\left(1-\mathbb{P}\left(E_1|\Theta=\theta\right)\right)\mathbb{P}\left(E_2|\Theta=\theta\right)\mathbb{P}\left(E_3|\Theta=\theta\right)\times \\ \times\mathbb{P}\left(E_4|\Theta=\theta\right)\pi_0(\theta)=K\theta^3(1-\theta)\pi_0(\theta) \\ =3K\theta^3(1-\theta)\end{gathered}$$

dove si è usato che $\pi_0(\theta)=\frac{d}{d\theta}F_\Theta(\theta)=3\theta^2$. Poiché

$$\int_0^1 d\theta 3K\theta^3(1-\theta)=\frac{3}{5\cdot 4}K=1\implies K=\frac{20}{3}\ .$$

2. La probabilità *a priori* che $\Theta \in \left[\frac{1}{2}, 1\right]$ per definizione è

$$\int_{\frac{1}{2}}^{1} d\theta \pi_0 \left(\theta\right) = F_\Theta \left(1\right) - F_\Theta \left(\frac{1}{2}\right) = 1 - \frac{1}{8} = \frac{7}{8} \ .$$

3. La probabilità *a posteriori* che $\Theta \in \left[\frac{1}{2}, 1\right]$ per definizione sarà

$$\begin{aligned} \int_{\frac{1}{2}}^{1} d\theta \pi_4 \left(\Theta = \theta | E_1^c \cap E_2 \cap E_3 \cap E_4\right) &= \int_{\frac{1}{2}}^{1} d\theta 20\theta^3 \left(1 - \theta\right) \\ = 20 \left(\frac{\theta^4}{4}|_{\frac{1}{2}}^{1} - \frac{\theta^5}{5}|_{\frac{1}{2}}^{1}\right) &= 5 \left(1 - \frac{1}{16}\right) - 4 \left(1 - \frac{1}{32}\right) \\ &= \frac{150 - 124}{32} = \frac{13}{16} \ . \end{aligned}$$

4. Differenziando $\pi_4 \left(\Theta = \theta | E_1^c \cap E_2 \cap E_3 \cap E_4\right) = 20\theta^3 \left(1 - \theta\right)$ otteniamo

$$3\theta^2 \left(1 - \theta\right) - \theta^3 = 0$$

ovvero che il valore massimo della densità di probabilità a posteriori di Θ è raggiunto per $\theta = \frac{3}{4}$.

■