Esercizi di Calcolo delle Probabilità e Statistica

a.a.2023/2024
15/07/2024

Esercizio 1 *Il numero di bottiglie di vino bianco vendute in un giorno dall'enoteca di quartiere è una v.a. di Poisson di parametro* $\lambda = 5$, *mentre il numero di bottiglie di vino rosso vendute in un giorno è una v.a. di Poisson di parametro* $\lambda = 3$. *Si considerino il numero di bottiglie di vino rosso e il numero di bottiglie di vino bianco come vv.aa. indipendenti così come il numero di bottiglie di vino dello stesso colore vendute in giorni diversi.*

1. *Calcolare la probabilità che in un qualsiasi giorno di apertura venga venduta una sola bottiglia di rosso.*
2. *Calcolare la probabilità che in un qualsiasi giorno di apertura vengano vendute almeno* 3 *bottiglie.*
3. *Se l'esercente mettesse da parte* 2 *euro per ciascuna bottiglia di vino venduta, stimare la probabilità che in un anno lavorativo di* 255 *giorni l'esercente accantoni almeno* 5000 *euro sapendo che se* $\xi \stackrel{d}{=} Poi(\lambda)$ *e* $t > \lambda$,

$$\mathbb{P}\{\xi \geq t\} \leq e^{-\lambda}\left(\frac{e\lambda}{t}\right)^t \ . \tag{1}$$

4. *Dimostrare la (1) usando la disuguaglianza di Markov per la v.a.* $e^{s\xi}, s > 0$ *e successivamente minimizzando rispetto ad* s.

Soluzione:

1. Sia X la v.a. che indica il numero di bottiglie di rosso vendute in un giorno. Allora,
$$\mathbb{P}\{X = 1\} = e^{-3}3 \ .$$
2. Sia Y la v.a. che il numero di bottiglie di bianco vendute in un giorno. Allora, $Z := X + Y$ è la v.a. che indica il numero totale di bottiglie vendute in un giorno. Poiché X e Y sono vv.aa. con distribuzione di Poisson rispettivamente di parametri $\lambda_r = 3$ e $\lambda_b = 5, Z$ è una v.a. con distribuzione di Poisson di parametro $\lambda = \lambda_r + \lambda_b = 8$ *(dimostrarlo!)*. Perciò, se $\{Z \geq 3\}$ è l'evento che vengano vendute almeno 3 bottiglie al giorno,
$$\begin{aligned}\mathbb{P}\{Z \geq 3\} &= 1 - \mathbb{P}\{Z \leq 2\} = 1 - \sum_{k=0}^{2} e^{-8}\frac{8^k}{k!} \\ &= 1 - e^{-8}(1 + 8 + 32) = 1 - 41e^{-8} \ .\end{aligned}$$

3. Poniamo $N := 255$. Se Z_i indica il numero di bottiglie vendute il giorno i, la collezione di vv.aa. $\{Z_i\}_{i=1}^N$ è composta da vv.aa. idipendenti identicamente distribuite tali che ciascuna è distribuita secondo la distribuzione di probabilità di Poisson di parametro $\lambda = 8$. Perciò, indicando $S := \sum_{i=1}^N Z_i$ la v.a. che rappresenta il numero di bottiglie vendute in un anno di esercizio e con $W := 2S$ la v.a. che rappresenta la somma di denaro complessivamente accantonata per la vendita delle bottiglie in un anno di esercizio, poiché $S \stackrel{d}{=} Poi(8N)$ *(dimostrarlo!)* e $8N = 2040 < 2500$, per la (1) si ha

$$\mathbb{P}\{W \geq 5000\} = \mathbb{P}\{S \geq 2500\} \leq e^{-2040}\left(\frac{e2040}{2500}\right)^{2500} .$$

4. Per ogni $s > 0$, poiché $\mathbb{R}_+ \ni x \longmapsto e^{sx} \in \mathbb{R}_+$ è una funzione monotona non decrescente, $\mathbb{P}\{\xi \geq t\} = \mathbb{P}\left\{e^{s\xi} \geq e^{st}\right\}$. Allora per la disuguaglenza di Markov si ottiene

$$\begin{aligned}\mathbb{P}\left\{e^{s\xi} \geq e^{st}\right\} &\leq e^{-st}\mathbb{E}\left[e^{s\xi}\right] = e^{-st}\sum_{k\geq 0} e^{sk}e^{-\lambda}\frac{\lambda^k}{k!} \\ &= e^{-st-\lambda}\sum_{k\geq 0}\frac{\left(e^s\lambda\right)^k}{k!} = \exp\left\{-\left[st+\lambda\left(1-e^s\right)\right]\right\}\end{aligned}$$

per tutti i valori di $s > 0$ e in particolare per quello che rende minima questa espressione, cioè per $s = \log\frac{t}{\lambda}$, da cui la tesi.

■

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

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

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 $Q := \left\{(x, y) \in \mathbb{R}^2 : |x| \vee |y| \leq 1\right\}$ e qiella della regione del piano $C := \left\{(x, y) \in \mathbb{R}^2 : |x| \leq 1, x^2 \leq y \leq 1\right\}$. Ovvero, poiché

$$\int_{-1}^1 dx \int_{x^2}^1 dy = \int_{-1}^1 dx\left(1 - x^2\right) = 2 - \frac{2}{3} = \frac{4}{3} ,$$

$c = \left(4 - \frac{4}{3}\right)^{-1} = \frac{3}{8}$.

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

$$f_X(x) = \mathbf{1}_{[-1,1]}(x)\frac{3}{8}\int_{-1}^{x^2} dy = \frac{3}{8}\left(x^2+1\right)\mathbf{1}_{[-1,1]}(x)$$

e

$$\begin{aligned} f_Y(y) &= \frac{3}{8}\left[2\mathbf{1}_{[-1,0]}(y) + \left(\int_{-1}^{-\sqrt{y}} dx + \int_{\sqrt{y}}^{1} dy\right)\mathbf{1}_{[0,1]}(y)\right] \\ &= \frac{3}{4}\left(\mathbf{1}_{[-1,0]}(y) + (1-\sqrt{y})\,\mathbf{1}_{[0,1]}(y)\right) . \end{aligned}$$

Da cui segie che le vv.aa. X e Y non sono tra loro stocasticamente indipendenti. Infatti, scegliendo per esempio il punto $\left(0,\frac{1}{2}\right)$ si ottiene

$$f_X(0) f_Y\left(\frac{1}{2}\right) = \frac{3}{8}\left(1-\frac{1}{\sqrt{2}}\right) \neq f_{(X,Y)}\left(0,\frac{1}{2}\right) = 0 .$$

Allora,

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

3. Poiché f_X è una funzione pari, quindi $\mathbb{R} \ni x \longmapsto x f_X(x) \in \mathbb{R}$ è dispari, e $[-1,1]$ è un intervallo simmetrico si ha che

$$\mathbb{E}[X] = \int_{\mathbb{R}} dx f_X(x)\, x = \frac{3}{8}\int_{-1}^{1} dx\left(x^2+1\right)x = 0 ,$$

mentre

$$\begin{aligned} \mathbb{E}[Y] &= \int_{\mathbb{R}} dy f_Y(y)\, y = \frac{3}{4}\left(-\frac{1}{2} + \int_0^1 dy y\left(1-\sqrt{y}\right)\right) \\ &= \frac{3}{4}\left(-\frac{1}{2}+\frac{1}{2}-\int_0^1 dy y^{\frac{3}{2}}\right) = \frac{3}{4}\left(-\frac{2}{5}\right) = -\frac{3}{10} . \end{aligned}$$

Inoltre,

$$\begin{aligned} \mathbb{E}[XY] &= \frac{3}{8}\int_D dxdyxy \\ &= \frac{3}{8}\left[\int_Q dxdyxy - \int_C dxdyxy\right] . \end{aligned}$$

Notiamo che, integrando funzioni dispari su un intervallo pari,

$$\int_Q dxdyxy = \int_{-1}^{-1} dxx \int_{-1}^{1} dyy = 0 .$$

Inoltre, per lo stesso motivo,

$$\int_C dxdyxy = \int_{-1}^{1} dxx \int_{x^2}^{1} dyy = \int_{-1}^{1} dxx \left(1 - x^2\right) = 0 \ .$$

Perciò $\mathbb{E}[XY] = 0$, quindi

$$Cov(X,Y) = Cov(Y,X) = \mathbb{E}[XY] - \mathbb{E}[X]\,\mathbb{E}[Y] = 0 \ ,$$

ovvero X e Y sono vv.aa. scorrelate sebbene non indipendenti. Non resta che calcolare $\mathbb{E}\left[X^2\right]$ e $\mathbb{E}\left[Y^2\right]$.

$$\begin{aligned} \mathbb{E}\left[X^2\right] &= \frac{3}{8}\int_{-1}^{1} dx \left(x^2+1\right)x^2 = \frac{3}{8}\left(\frac{2}{5}+\frac{2}{3}\right) = \frac{2}{5} \\ \mathbb{E}\left[Y^2\right] &= \frac{3}{4}\left(\frac{1}{3} + \int_0^1 dy\,(1-\sqrt{y})\,y^2\right) = \frac{3}{4}\left(\frac{2}{3}-\frac{2}{7}\right) = \frac{2}{7} \end{aligned}$$

Quindi

$$\begin{aligned} \begin{pmatrix} Cov(X,X) & Cov(X,Y) \\ Cov(Y,X) & Cov(Y,Y) \end{pmatrix} &= \begin{pmatrix} \mathbb{E}\left[X^2\right] & 0 \\ 0 & \mathbb{E}\left[Y^2\right] - \mathbb{E}^2[Y] \end{pmatrix} \\ &= \begin{pmatrix} \frac{2}{5} & 0 \\ 0 & \frac{2}{7} - \frac{9}{100} \end{pmatrix} \ . \end{aligned}$$

■

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* $\frac{1}{6}$ *agli stati* 2 *e* 3, *e si resta nello stesso stato con probabilità* $\frac{2}{3}$. *Dallo stato* 2, *si transisce con probabilità* $\frac{1}{2}$ *agli stati* 1 *e* 3. *Dallo stato* 3 *si transisce con probabilità* $\frac{1}{2}$ *agli stati* 2 *e* 4. *Dallo stato* 4 *si transisce con probabilità* $\frac{1}{6}$ *agli stati* 2 *e* 3, *e si resta nello stesso stato con probabilità* $\frac{2}{3}$.

1. *Elencare le classi di equivalenza degli stati e calcolarne il periodo.*
2. *Calcolare* $\mathbb{P}(\{X_3 = 2\} \,|\, \{X_1 = 3\})$ *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:* $\lim_{n\to\infty} \mathbb{P}\{\omega \in \Omega : X_{2n+1}(\omega) = 4\}, \lim_{n\to\infty}\left(P^n_{3,2} + P^n_{3,4}\right), \lim_{n\to\infty} P^{2n}_{1,3}$.

Soluzione:

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

$$P = \begin{pmatrix} \frac{2}{3} & \frac{1}{6} & \frac{1}{6} & 0 \\ \frac{1}{2} & 0 & \frac{1}{2} & 0 \\ 0 & \frac{1}{2} & 0 & \frac{1}{2} \\ 0 & \frac{1}{6} & \frac{1}{6} & \frac{2}{3} \end{pmatrix}$$

perciò la catena può essere rappresentata mediante il grafo diretto

$$\circlearrowleft\ 1 \quad \leftrightarrow \quad 2$$
$$\downarrow \quad \nearrow\swarrow \quad \uparrow$$
$$3 \quad \leftrightarrow \quad 4 \ \circlearrowleft$$

da cui si deduce che esiste un unica classe di stati comunicanti aperiodica.

2. Poiché la catena di Markov è omogenea

$$\mathbb{P}\left(\{X_3 = 2\} \,|\, \{X_1 = 3\}\right) = \mathbb{P}\left(\{X_2 = 2\} \,|\, \{X_0 = 3\}\right) = P^2_{3,2} \ .$$

Ma

$$P^2 = \begin{pmatrix} \frac{19}{36} & \frac{7}{36} & \frac{7}{36} & \frac{3}{36} \\ \frac{4}{12} & \frac{4}{12} & \frac{1}{12} & \frac{3}{12} \\ \frac{3}{12} & \frac{1}{12} & \frac{4}{12} & \frac{4}{12} \\ \frac{3}{36} & \frac{7}{36} & \frac{7}{36} & \frac{19}{36} \end{pmatrix} \ ,$$

Quindi, $P^3_{3,2} = \frac{1}{12}$.

$$\mathbb{E}\left[X_2\right] = \sum_{i,j=1}^{4} j\mathbb{P}\left(\{X_2 = j\} \,|\, \{X_0 = i\}\right)\mathbb{P}\left\{X_0 = i\right\} =$$
$$= \sum_{i,j=1}^{4} \frac{1}{4} P^2_{i,j} j = \frac{1}{4}\left(\frac{11}{6} + \frac{27}{12} + \frac{33}{12} + \frac{57}{18}\right) =$$
$$= \frac{1}{4}\frac{360}{36} = \frac{5}{2} \ .$$

3. Poiché la catena di Markov descritta dalla matrice delle probabilità di transizione P è indecomponibile e aperiodica vale il teorema ergodico. Perciò,

$$\lim_{n\to\infty} P^{2n}_{1,3} = \pi_3, \ \lim_{n\to\infty} P^n_{3,2} = \pi_2,$$
$$\lim_{n\to\infty} P^n_{3,4} = \pi_4 = \sum_{i=1}^{4} \mathbb{P}\left\{\omega \in \Omega : X_0\left(\omega\right) = i\right\} P^n_{i,4} = \lim_{n\to\infty} \mathbb{P}\left\{\omega \in \Omega : X_n\left(\omega\right) = 4\right\},$$

e pertanto $\lim_{n\to\infty}\left(P^n_{1,3} + P^n_{3,4}\right) = \pi_3 + \pi_4$, dove $\pi := (\pi_1, \pi_2, \pi_3, \pi_4)$ è la distribuzione di probabilità invariante per P, ovvero $\pi P = \pi$. Pertanto, risolvendo

$$\begin{cases} \frac{2}{3}\pi_1 + \frac{1}{2}\pi_2 = \pi_1 \\ \frac{1}{6}\pi_1 + \frac{1}{2}\pi_3 + \frac{1}{6}\pi_4 = \pi_2 \\ \frac{1}{6}\pi_1 + \frac{1}{2}\pi_2 + \frac{1}{6}\pi_4 = \pi_3 \\ \frac{2}{3}\pi_4 + \frac{1}{2}\pi_3 = \pi_4 \end{cases}$$
$$\pi_1 + \pi_2 + \pi_3 + \pi_4 = 1$$

si ha $\pi_1 = \pi_4 = \frac{3}{10}, \pi_2 = \pi_3 = \frac{1}{5}$.

■

Esercizio 4 *Sia $\{X_i\}_{i\geq 1}$ una successione di vv.aa. stocasticamenti indipendenti, subordinatamente alla conoscenza di un parametro Θ, con densità di probabilità condizionata marginale*

$$f(x_i|\theta) = \sqrt{\frac{\theta}{2\pi}} \exp\left\{-\frac{\theta}{2}x_i^2\right\} \ , \ x_i \in \mathbb{R} \ .$$

Supponendo che la densità a priori di Θ sia un esponenziale di parametro 1, se si osservano i primi valori $x_1 = 2, x_2 = \sqrt{2}, x_3 = 1, x_4 = \frac{3}{2}$, della successione,

1. *calcolare la densità di probabilità a posteriori di Θ e il suo punto di massimo;*
2. *calcolare la probabilità a posteriori dell'evento $\{\Theta > 2\}$;*
3. *calcolare valore atteso e varianza a posteriori della v.a. $Z = \Theta^2$.*

Soluzione:

1. La densità a posteriori di Θ sarà pari a

$$\pi\left(\Theta = \theta|\,(x_1, x_2, x_3, x_4) = \left(2, \sqrt{2}, 1, \frac{3}{2}\right)\right) =$$
$$= Kf_\Theta\left(x_1 = 2, x_2 = \sqrt{2}, x_3 = 1, x_4 = \frac{3}{2}\right)\pi_\Theta(\theta)$$
$$= Kf(x_1 = 2|\theta)\, f\left(x_2 = \sqrt{2}|\theta\right) f(x_3 = 1|\theta)\, f\left(x_4 = \frac{3}{2}|\theta\right) e^{-\theta}\mathbf{1}_{\mathbb{R}_+}(\theta)$$
$$= \frac{K}{(2\pi)^2}\theta^2 e^{-\theta\left[\frac{\left(4+2+1+\frac{9}{4}\right)}{2}+1\right]} = \frac{K}{(2\pi)^2}\theta^2 e^{-\theta\left[\frac{45}{8}\right]} \ ,$$

   e quindi è una $\Gamma\left(3, \frac{45}{8}\right)$. Perciò $K = (2\pi)^2\left(\frac{45}{8}\right)^3\frac{1}{2} = \pi^2\frac{3^6 5^3}{2^8}$.
   Il punto di massimo si ha per $\theta = \frac{16}{45}$.

2. Ponendo $\lambda = \frac{45}{8}$, integrando due volte per parti, la probabilità a posteriori dell'evento $\{\Theta > 2\}$ risulta allora pari a

$$\frac{\lambda^3}{2}\int_2^{+\infty} d\theta\theta^2 e^{-\theta\lambda} = \frac{1}{2}\left(4\lambda^2 + 4\lambda + 2\right)e^{-2\lambda}$$
$$= \left(2\lambda^2 + 2\lambda + 1\right)e^{-2\lambda} \ .$$

3. Il valore atteso a posteriori di Θ^2 è pari a

$$\frac{\lambda^3}{2}\int_0^{+\infty} d\theta\theta^4 e^{-\theta\lambda} = \frac{4!}{2\lambda^2} = \frac{2^8}{3^3 5^2} \ ;$$

   il valore atteso a posteriori di Θ^4 è pari a

$$\frac{\lambda^3}{2}\int_0^{+\infty} d\theta\theta^6 e^{-\theta\lambda} = \frac{6!}{2\lambda^4} = \frac{2^{15}}{3^6 5^3} \ .$$

   Perciò la varianza a posteriori di Θ è uguale a $\frac{2^{15}}{3^6 5^3} - \frac{2^{16}}{3^6 5^4} = \frac{2^{15}}{3^6 5^3}\left(1 - \frac{2}{5}\right)$.

■