

Esercizi di Calcolo delle Probabilità e Statistica

a.a.2023/2024
16/09/2024

Esercizio 1 *Un'azienda produce componenti per autovetture. Un'indagine statistica sulla produzione indica che lo 0,2% dei componenti presenta un difetto di gravità 1, lo 0,1% presenta un difetto di gravità 2 e lo 0,05% presenta un difetto di gravità 3. La percentuale rimanente dei componenti prodotti non presenta difetti. Inoltre l'indagine mostra che, durante la fase di rodaggio, un componente con un difetto di gravità 1 ha una probabilità di rompersi del 65%, questa stessa probabilità sale all'80% per i difetti di gravità 2 e al 90% per quelli di gravità 3.*

1. Calcolare la probabilità che un componente scelto a caso si rompa durante la fase di rodaggio.
2. Calcolare la probabilità che un componente che si rompe nella fase di rodaggio abbia un difetto di gravità 1.
3. Calcolare valore atteso e varianza del numero di componenti che possono rompersi in fase di rodaggio se la produzione è di 50000 unità.
4. Calcolare la distribuzione di probabilità del numero di componenti che possono rompersi in fase di rodaggio che hanno difetti di gravità al più 2.

Soluzione:

1. Il problema si può riformulare nel modo seguente:

Per ogni $i = 1, 2, 3$ consideriamo un'urna contenente un numero totale N_i di palline bianche e nere pari al numero di componenti difettosi di gravità i di cui il numero di palline bianche H_i è pari al numero di componenti difettosi di gravità i che si rompono.

Pertanto la probabilità che un componente estratto a caso si rompa è pari alla probabilità di estrarre una pallina bianca, ovvero, indicando con R questo evento e, per ogni $i = 1, 2, 3$, con U_i l'evento che indica la scelta dell'urna i ,

$$\mathbb{P}(R) = \sum_{i=1}^3 \mathbb{P}(R|U_i) \mathbb{P}(U_i) .$$

Ma, $\mathbb{P}(U_i)$, la probabilità di scegliere l'urna i , è pari al rapporto tra il numero di componenti prodotti con difetto di gravità i e il numero totale di componenti prodotti, mentre $\mathbb{P}(R|U_i)$ è la probabilità di estrarre una pallina bianca dall'urna i , ovvero la probabilità che il componente con difetto di gravità i si rompa. Perciò,

$$\begin{aligned} \mathbb{P}(R) &= 2 \cdot 10^{-3} \cdot 65 \cdot 10^{-2} + 10^{-3} \cdot 8 \cdot 10^{-1} + 5 \cdot 10^{-4} \cdot 9 \cdot 10^{-1} \\ &= (130 + 80 + 45) 10^{-5} = 255 \cdot 10^{-5} = 0,00255 . \end{aligned}$$

2. Segendo lo schema dato al punto precedente, la probabilità che un componente che si rompe nella fase di rodaggio abbia un difetto di gravità 1 è

$$\mathbb{P}(U_1|R) = \mathbb{P}(R|U_1) \frac{\mathbb{P}(U_1)}{\mathbb{P}(R)} = \frac{130 \cdot 10^{-5}}{255 \cdot 10^{-5}} \simeq 0,51 .$$

3. Per ogni $i = 1, 2, 3$, posto $N_i = (5 \cdot 10^4) \mathbb{P}(U_i)$ il numero di componenti difettosi di gravità i , sia $X_k^{(i)}$ la v.a. che indica che il k -simo componente con difetto di gravità i tra gli N_i si rompa in fase di rodaggio. Allora, la v.a. che rappresenta in numero di componenti con difetto di gravità i che possono rompersi $T_i := \sum_{k=0}^{N_i} X_k^{(i)}$, essendo una somma di vv.aa. bernoulliane di parametro $\mathbb{P}(R|U_i)$, ha distribuzione binomiale di parametri $\mathbb{P}(R|U_i)$ e N_i . Il numero di componenti che possono rompersi N è quindi dato dalla somma delle vv.aa. N_i . Perciò,

$$\begin{aligned} \mathbb{E}[N] &= \sum_{i=1}^3 \mathbb{E}[N_i] = \sum_{i=1}^3 N_i \mathbb{P}(R|U_i) \\ &= (5 \cdot 10^4) \sum_{i=1}^3 \mathbb{P}(R|U_i) \mathbb{P}(U_i) \\ &= (5 \cdot 10^4) (255 \cdot 10^{-5}) = 5 \cdot 255 \cdot 10^{-1} = 1275 . \end{aligned}$$

Poiché le vv.aa. N_i sono tra loro stocasticamente indipendenti, la varianza di N è pari alla somma delle varianze delle N_i che a loro volta ciascuna di loro è pari alla somma delle varianze delle $X_k^{(i)}$. Dunque,

$$\begin{aligned} Var[N] &= \sum_{i=1}^3 Var[N_i] = \sum_{i=1}^3 N_i \mathbb{P}(R|U_i) (1 - \mathbb{P}(R|U_i)) \\ &= (5 \cdot 10^4) \sum_{i=1}^3 \mathbb{P}(R|U_i) \mathbb{P}(U_i) (1 - \mathbb{P}(R|U_i)) \\ &= 5 (13 \cdot 35 + 80 \cdot 2 + 45) 10^{-2} = 33 . \end{aligned}$$

4. Ponendo $p_i := \mathbb{P}(R|U_i)$, $i = 1, 2$,

$$\mathbb{P}(T_1 + T_2 = n) = \sum_{\{(k_1, k_2) \in \{0, \dots, N_1\} \times \{0, \dots, N_2\} : k_1 + k_2 = n\}} \binom{N_1}{k_1} p_1^{k_1} (1 - p_1)^{N_1 - k_1} \binom{N_2}{k_2} p_2^{k_2} (1 - p_2)^{N_2 - k_2} ,$$

per $n \in \{0, \dots, N_1 + N_2\}$.

■

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

$$D := \left\{ (x, y) \in \mathbb{R}^2 : |x| \vee |y| \geq 1, \frac{x^2}{4} + 4y^2 \leq 1 \right\} .$$

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

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

Soluzione:

1. L'area di D è pari a $4c$ dove

$$\begin{aligned} c &= \int_1^2 dx \int_0^{\frac{1}{2}\sqrt{1-\frac{x^2}{4}}} dy = \frac{1}{2} \int_1^2 dx \sqrt{1-\frac{x^2}{4}} = \int_{\frac{1}{2}}^1 dz \sqrt{1-z^2} \\ &= \int_{\frac{\pi}{3}}^{\frac{\pi}{2}} dt \cos^2 t = \frac{1}{2} (t + \sin t \cos t) \Big|_{\frac{\pi}{3}}^{\frac{\pi}{2}} = \frac{1}{2} \left(\frac{\pi}{6} - \frac{\sqrt{3}}{4} \right) \\ &= \frac{1}{4} \left(\frac{\pi}{3} - \frac{\sqrt{3}}{2} \right) \end{aligned}$$

Perciò la costante di normalizzazione della distribuzione di probabilità congiunta del vettore aleatorio (X, Y) è pari a $K := \left(\frac{\pi}{3} - \frac{\sqrt{3}}{2} \right)^{-1}$.

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

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

e

$$\begin{aligned} f_Y(y) &= \mathbf{1}_{[-\frac{1}{2}, \frac{1}{2}]}(y) K \left[\int_{-2\sqrt{1-4y^2}}^{-1} dx + \int_1^{2\sqrt{1-4y^2}} dx \right] \\ &= \mathbf{1}_{[-\frac{1}{2}, \frac{1}{2}]}(y) 2K (2\sqrt{1-4y^2} - 1). \end{aligned}$$

Da cui segue che le vv.aa. X e Y non sono tra loro stocasticamente indipendenti. Infatti, scegliendo per esempio il punto $(\frac{3}{2}, 0)$ si ottiene

$$f_X\left(\frac{3}{2}\right) f_Y(0) = K^2 \sqrt{5} \neq f_{(X, Y)}\left(\frac{3}{2}, 0\right) = K.$$

Allora,

$$f_{Y|X}(y|x) = \frac{f_{(X, Y)}(x, y)}{f_X(x)} = \frac{\mathbf{1}_{\{(u, v) \in [-1, 1]^2 : |x| \leq 1, -1 \leq y \leq x^2\}}(x, y)}{\sqrt{1-\frac{x^2}{4}}}.$$

3. Poiché f_X è una funzione pari, quindi $\mathbb{R} \ni x \mapsto xf_X(x) \in \mathbb{R}$ è dispari, e $[-2, -1] \cup [1, 2]$ è un dominio d'integrazione simmetrico si ha che

$$\mathbb{E}[X] = \int_{\mathbb{R}} dx f_X(x) x = K \left(\int_{-2}^{-1} dx x \sqrt{1 - \frac{x^2}{4}} + \int_1^2 dx x \sqrt{1 - \frac{x^2}{4}} \right) = 0 .$$

Le stesse considerazioni valgono per f_Y , perciò

$$\mathbb{E}[Y] = \int_{\mathbb{R}} dy f_Y(y) y = 2K \int_{-\frac{1}{2}}^{\frac{1}{2}} dy y \left(2\sqrt{1 - 4y^2} - 1 \right) = 0 .$$

Inoltre, per lo stesso motivo, $\forall x \in [-2, 2], \int_{-\frac{1}{2}\sqrt{1-\frac{x^2}{4}}}^{\frac{1}{2}\sqrt{1-\frac{x^2}{4}}} dy y = 0$, dunque,

$$\mathbb{E}[XY] = K \int_D dx dy xy = K \int_{-2}^{-1} dx x \int_{-\frac{1}{2}\sqrt{1-\frac{x^2}{4}}}^{\frac{1}{2}\sqrt{1-\frac{x^2}{4}}} dy y = 0 ,$$

quindi

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

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

$$\begin{aligned} \mathbb{E}[X^2] &= K \left(\int_{-2}^{-1} dx x^2 \sqrt{1 - \frac{x^2}{4}} + \int_1^2 dx x^2 \sqrt{1 - \frac{x^2}{4}} \right) = \\ &= 2K \int_1^2 dx x^2 \sqrt{1 - \frac{x^2}{4}} = 4K \int_{\frac{\pi}{3}}^{\frac{\pi}{2}} du \sin^2 u = \\ &= 2K (u - \sin u \cos u) \Big|_{\frac{\pi}{3}}^{\frac{\pi}{2}} = K \left(\frac{\pi}{3} + \frac{\sqrt{3}}{2} \right) \\ \mathbb{E}[Y^2] &= 2K \int_{-\frac{1}{2}}^{\frac{1}{2}} dy y^2 \left(2\sqrt{1 - 4y^2} - 1 \right) = -\frac{K}{6} + \\ &\quad + \frac{K}{2} \int_{-1}^1 dt t^2 \sqrt{1 - t^2} \\ &= -\frac{K}{6} + \frac{K}{16} \int_{\pi}^{\pi} du \sin^2 u = \frac{K}{2} \left(\frac{\pi}{4} - \frac{1}{3} \right) \end{aligned}$$

Quindi

$$\begin{aligned} \begin{pmatrix} \text{Cov}(X, X) & \text{Cov}(X, Y) \\ \text{Cov}(Y, X) & \text{Cov}(Y, Y) \end{pmatrix} &= \begin{pmatrix} \mathbb{E}[X^2] & 0 \\ 0 & \mathbb{E}[Y^2] \end{pmatrix} \\ &= \begin{pmatrix} \frac{2\pi+3\sqrt{3}}{2\pi-3\sqrt{3}} & 0 \\ 0 & \frac{3\pi-4}{4(2\pi-3\sqrt{3})} \end{pmatrix} . \end{aligned}$$

■

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

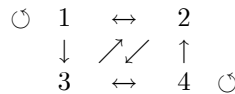
1. Elencare le classi di equivalenza degli stati e calcolarne il periodo.
2. Calcolare $\mathbb{P}(\{X_3 = 2\} | \{X_1 = 3\})$ e, assumendo come distribuzione iniziale quella uniforme, calcolare $\mathbb{E}[X_2]$.
3. Se $\{P_{i,j}\}_{i,j=1,\dots,4}$ indica la matrice delle probabilità di transizione tra gli stati della catena, calcolare se esistono: $\lim_{n \rightarrow \infty} \mathbb{P}\{\omega \in \Omega : X_{2n+1}(\omega) = 4\}$, $\lim_{n \rightarrow \infty} (P_{3,2}^n + P_{3,4}^n)$, $\lim_{n \rightarrow \infty} P_{1,3}^{2n}$.

Soluzione:

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

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

perciò la catena può essere rappresentata mediante il grafo diretto



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

2. Poiché la catena di Markov è omogenea

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

Ma

$$P^2 = \begin{pmatrix} \frac{9}{18} & \frac{4}{18} & \frac{4}{18} & \frac{1}{18} \\ \frac{18}{18} & \frac{18}{18} & \frac{18}{18} & \frac{18}{18} \\ \frac{18}{18} & \frac{18}{18} & \frac{18}{18} & \frac{18}{18} \\ \frac{1}{18} & \frac{4}{18} & \frac{4}{18} & \frac{1}{18} \end{pmatrix},$$

Quindi, $P_{3,2}^3 = \frac{1}{18}$.

$$\begin{aligned} \mathbb{E}[X_2] &= \sum_{i,j=1}^4 j \mathbb{P}(\{X_2 = j\} | \{X_0 = i\}) \mathbb{P}\{X_0 = i\} = \\ &= \sum_{i,j=1}^4 \frac{1}{4} P_{i,j}^2 j = \frac{1}{4} \left(\frac{11}{6} + \frac{27}{12} + \frac{33}{12} + \frac{57}{18} \right) = \\ &= \frac{1}{4} \frac{180}{18} = \frac{5}{2}. \end{aligned}$$

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

$$\begin{aligned}\lim_{n \rightarrow \infty} P_{1,3}^{2n} &= \pi_3, \quad \lim_{n \rightarrow \infty} P_{3,2}^n = \pi_2, \\ \lim_{n \rightarrow \infty} P_{3,4}^n &= \pi_4 = \sum_{i=1}^4 \mathbb{P}\{\omega \in \Omega : X_0(\omega) = i\} P_{i,4}^n = \lim_{n \rightarrow \infty} \mathbb{P}\{\omega \in \Omega : X_n(\omega) = 4\},\end{aligned}$$

e pertanto $\lim_{n \rightarrow \infty} (P_{1,3}^n + P_{3,4}^n) = \pi_3 + \pi_4$, dove $\pi := (\pi_1, \pi_2, \pi_3, \pi_4)$ è la distribuzione di probabilità invariante per P , ovvero $\pi P = \pi$. Pertanto, poiché P è bistocastica, cioè anche la somma degli elementi di ciascuna colonna di P è pari a 1, la distribuzione di probabilità stazionaria è quella uniforme, come si può anche vedere risolvendo

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

Perciò si ha $\pi_1 = \pi_2 = \pi_3 = \pi_4 = \frac{1}{4}$.

■

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

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

Supponendo che la densità a priori di Θ sia un esponenziale di parametro 1, se si osservano i primi valori $x_1 = 1, x_2 = 2, x_3 = \frac{1}{2}, x_4 = \frac{3}{2}$, 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 $\{\Theta \leq 3\}$;
3. calcolare valore atteso e varianza a posteriori della v.a. $Z = \sqrt{\Theta}$.

Soluzione:

1. La densità a posteriori di Θ sarà pari a

$$\begin{aligned}\pi\left(\Theta = \theta | (x_1, x_2, x_3, x_4) = \left(1, 2, \frac{1}{2}, \frac{3}{2}\right)\right) &= \\ &= K f_{\Theta}\left(x_1 = 1, x_2 = 2, x_3 = \frac{1}{2}, x_4 = \frac{3}{2}\right) \pi_{\Theta}(\theta) \\ &= K f(x_1 = 1|\theta) f(x_2 = 2|\theta) f\left(x_3 = \frac{1}{2}|\theta\right) f\left(x_4 = \frac{3}{2}|\theta\right) e^{-\theta} \mathbf{1}_{\mathbb{R}_+}(\theta) \\ &= \frac{K}{(2\pi)^2} \theta^2 e^{-\theta \left[\frac{(1+4+\frac{1}{4}+\frac{9}{4})}{2} + 1\right]} = \frac{K}{(2\pi)^2} \theta^2 e^{-\theta \left[\frac{19}{4}\right]},\end{aligned}$$

e quindi è una $\Gamma\left(3, \frac{19}{4}\right)$. Perciò $K = (2\pi)^2 \left(\frac{19}{4}\right)^3 \frac{1}{2} = \pi^2 \frac{(19)^3}{2^3}$.

Il punto di massimo si ha per $\theta = \frac{8}{19}$.

2. Ponendo $\lambda = \frac{19}{4}$, integrando due volte per parti, la probabilità a posteriori dell'evento $\{\Theta \leq 3\}$ risulta allora pari a

$$\frac{\lambda^3}{2} \int_0^3 d\theta \theta^2 e^{-\theta\lambda} = 1 - \left(\frac{9}{2}\lambda^2 + 3\lambda + 1\right) e^{-3\lambda}.$$

3. Il valore atteso a posteriori di $\sqrt{\Theta}$ è pari a

$$\frac{\lambda^3}{2} \int_0^{+\infty} d\theta \theta^{\frac{5}{2}} e^{-\theta\lambda} = \frac{\Gamma\left(\frac{7}{2}\right)}{2\sqrt{\lambda}} = \frac{5 \cdot 3\sqrt{\pi}}{2^3\sqrt{19}};$$

il valore atteso a posteriori di Θ è pari a

$$\frac{\lambda^3}{2} \int_0^{+\infty} d\theta \theta^3 e^{-\theta\lambda} = \frac{3!}{2\lambda} = \frac{2^2 3}{19}.$$

Perciò la varianza a posteriori di Θ è uguale a $\frac{2^2 3}{19} - \frac{5^2 3^2 \pi}{2^6 19} = \frac{3}{19} \left(2^2 - \frac{5^2 3 \pi}{2^6}\right)$.

■