8994

Corso di Statistica - Prof. Fabio Zucca
I prova in itinere - 5 maggio 2014

Nome e cognome: .... **Matricola:** ....

Esercizio 1 Tre monete vengono messe in fila su un tavolo: due di esse sono equilibrate, mentre per la rimanente la probabilità che esca testa è $2/3$. La posizione della moneta non equilibrata nella fila è casuale e tutte le possibili posizioni sono equiprobabili. Una volta messe in fila vengono prese una ad una e lanciate in maniera indipendente. Si osserva la sequenza (T, T, C).

1. Qual è la probabilità che si osservi la sequenza uscita dato che la moneta non equilibrata è la prima della fila? E se fosse la seconda della fila? Oppure la terza?

2. Qual è la probabilità che che la moneta non equilibrata sia la prima della fila sapendo che si osserva la sequenza (T, T, C)? Analogamente qual è la probabilità che che la moneta non equilibrata sia la seconda data la sequenza osservata? E la terza?

3. Se doveste scommettere sulla posizione di una moneta equilibrata, quale posizione scegliereste? Con che probabilità avreste ragione? Giustificare la risposta.

Soluzione. Chiamiamo per semplicità p la probabilità che esca testa per la moneta non equilibrata, in questo caso $p = 2/3$. Poichè le due monete equilibrate sono intercambiabili in questo problema, quello che interessa è la sola posizione della moneta non equilibrata nella fila; sia essa X. Chiaramente $\mathbb{P}(X = i) = 1/3$ per ogni $i = 1, 2, 3$. Chiamiamo Y la sequenza casuale dei lanci dove $Y = (Y_1, Y_2, Y_3)$.

1. Dobbiamo calcolare $\mathbb{P}(Y = (T, T, C)|X = i) \equiv \mathbb{P}(Y_1 = T|X = i)\mathbb{P}(Y_2 = T|X = i)\mathbb{P}(Y_3 = C|X = i)$ (dove si è utilizzata l'indipendenza dei lanci condizionata alla scelta della posizione Y). Notiamo che per simmetria si ha $\mathbb{P}(Y = (T, T, C)|X = 1) = \mathbb{P}(Y = (T, T, C)|X = 2)$. Dai dati del problema
$$\mathbb{P}(Y = (T, T, C)|X = i) = \begin{cases} p \cdot \dfrac{1}{2} \cdot \dfrac{1}{2} = \dfrac{p}{4} = \dfrac{1}{6} & \text{se } i = 1, 2; \\ \\ (1-p) \cdot \dfrac{1}{2} \cdot \dfrac{1}{2} = \dfrac{1-p}{4} = \dfrac{1}{12} & \text{se } i = 3. \end{cases}$$

2. Dobbiamo calcolare $\mathbb{P}(X = i|Y = (T, T, C))$ al variare di $i = 1, 2, 3$. Notiamo che per simmetria si ha $\mathbb{P}(X = 1|Y = (T, T, C)) = \mathbb{P}(X = 2|Y = (T, T, C))$. Applichiamo la formula di Bayes
$$\mathbb{P}(X = i|Y = (T, T, C)) = \frac{\mathbb{P}(Y = (T, T, C)|X = i)}{\sum_{j=1}^{3} \mathbb{P}(Y = (T, T, C)|X = j)}$$
(si ricordi che $\mathbb{P}(X = i) = 1/3$ per ogni $i = 1, 2, 3$), da cui
$$\mathbb{P}(X = 1|Y = (T, T, C)) = \begin{cases} \dfrac{p}{1+p} = \dfrac{2}{5} & \text{se } i = 1, 2; \\ \\ \dfrac{1-p}{1+p} = \dfrac{1}{5} & \text{se } i = 3. \end{cases}$$

3. Dai risultati del punto precedente si ha che la posizione meno probabile per la moneta non equilibrata è la terza e la probabilità è $\frac{1}{5}$. Pertanto scommettendo sulla terza posizione avrei ragione con probabilità $1 - \frac{1}{5} = \frac{4}{5}$.

Una soluzione alternativa parte dalla definizione esplicita dello spazio Ω e della misura $\mathbb{P}$. Diamo qualche cenno (provate a terminare l'esercizio). Non è difficile vedere che una definizione possibile è $\Omega = \{1, 2, 3\} \times \{T, C\}^3$ dove interpretiamo (i, x, y, z) come l'evento elementare "la moneta truccata è in posizione i e il risultato dei tre lanci é il vettore $(x, y, z) \in \{T, C\}^3$". La misura $\mathbb{P}$ è definita come

$$\mathbb{P}(\{(1, T, y, z)\}) = \mathbb{P}(\{(2, x, T, z)\}) = \mathbb{P}(\{(3, x, y, T)\}) = (2/3) \cdot (1/2) \cdot (1/2)$$
$$\mathbb{P}(\{(1, C, y, z)\}) = \mathbb{P}(\{(2, x, C, z)\}) = \mathbb{P}(\{(3, x, y, C)\}) = (1/3) \cdot (1/2) \cdot (1/2)$$

per ogni $x, y, z \in \{T, C\}$.

$\boxed{8994}$

Nome e cognome: .... **Matricola:** ....

Esercizio 2 Si consideri la seguente famiglia di funzioni reali di variabili reali ad un parametro $c \in \mathbb{R}$

$$f_c(x) := \begin{cases} x(c + \frac{6}{5}x) & \text{se } x \in [0,1], \\ 0 & \text{altrimenti.} \end{cases}$$

1. Per quali valori di c la funzione f_c rappresenta la densità di una variabile assolutamente continua?

2. Utilizzando uno dei valori di c calcolati in precedenza e denotato con X la variabile, se ne calcoli il valore atteso e la varianza.

3. Siano $\{X_i\}_{i=1}^n$ n v.a. *i.i.d.* con densità f_c dove c è uno dei valori determinati al punto 1. Calcolare, eventualmente ricorrendo ad una approssimazione, il minimo valore di n tale che $\mathbb{P}(\sum_{i=1}^n X_i > 72) \geq 0.15865$.

Soluzione.

1. La condizione $f_c \geq 0$ (quasi ovunque) è equivalente a $c \geq 0$. Inoltre

$$1 = \int_0^1 f_c(x)\mathrm{d}x = \int_0^1 x\left(c + \frac{6}{5}x\right)\mathrm{d}x = \left.\frac{c}{2}x^2 + \frac{2}{5}x^3\right|_0^1 = \frac{c}{2} + \frac{2}{5}$$

da cui $c = 6/5$.

2. Calcoliamo il valore atteso,

$$\mathbb{E}[X] = \int_{\mathbb{R}} x f_c(x)\mathrm{d}x = \frac{6}{5}\int_0^1 x^2(1+x)\,\mathrm{d}x = \left.\frac{6}{5}\left(\frac{x^3}{3} + \frac{x^4}{4}\right)\right|_0^1 = \frac{7}{10}$$

Calcoliamo ora

$$\mathbb{E}[X^2] = \int_{\mathbb{R}} x^2 f_c(x)\mathrm{d}x = \frac{6}{5}\int_0^1 x^3(1+x)\,\mathrm{d}x = \left.\frac{6}{5}\left(\frac{x^4}{4} + \frac{x^5}{5}\right)\right|_0^1 = \frac{27}{50}$$

da cui

$$\mathrm{Var}(X) = \mathbb{E}[X^2] - \mathbb{E}[X]^2 = \frac{27}{50} - \left(\frac{7}{10}\right)^2 = \frac{5}{100} = \frac{1}{20}.$$

3. Per $n > 30$ possiamo applicare il TCL:

$$\mathbb{P}\left(\sum_{i=1}^n X_i > 72\right) \approx 1 - \Phi\left(\frac{72 - n\mu}{\sqrt{n}\sigma}\right) \geq 0.15865$$

da cui

$$\frac{72 - n\mu}{\sqrt{n}\sigma} \leq \Phi^{-1}(0.84135)$$

ed essendo $\Phi^{-1}(0.84135) = 1$, $\mu = 7/10$ e $\sigma = \sqrt{5}/10$ ricaviamo:

$$n + \frac{\sqrt{5}}{7}\sqrt{n} - \frac{10}{7}72 \geq 0$$

che, risolta, offre $(\sqrt{n} \leq -10.30282779) \vee (\sqrt{n} \geq 9.983389508)$ da cui l'unica soluzione accettabile:

$$n \geq 99.66806606$$

cioè, essendo n intero, $n \geq 100$.

8994

Nome e cognome: .... **Matricola:** ....

Esercizio 3 Il numero di particelle che riducono la purezza della superficie di un DVD segue la distribuzione di Poisson e il numero medio di particelle di impurità per cm^2 di superficie è 0.1.

1. Calcolare la probabilità di trovare esattamente 12 particelle su una superficie di 100 cm^2.
2. Calcolare la probabilità di non trovare particelle su una superficie di 20 cm^2.
3. Prendiamo 10 DVD e controlliamo $20cm^2$ di superficie di ciascuno. Calcolare la probabilità che si trovino particelle esattamente in 8 DVD.
4. Calcolare, eventualmente ricorrendo ad una approssimazione, la probabilità di trovare al più 12 particelle su una superficie di 100 cm^2.

Soluzione. Sia s la misura in centimetri della superficie oggetto di test. Sia $X_s \sim P(\lambda_s)$ la v.a. poissoniana di parametro λ_s che conta le particelle presenti sulla superficie di area s.
Poiché $\lambda_s = \lambda_1 s$ e dal testo sappiamo che $\lambda_1 = 0.1$, ricaviamo $\lambda_s = 0.1s$

1. Abbiamo $s = 100$ da cui $\lambda_{100} = 10$ e $X_{100} \sim P(10)$.
   Quindi:
   $$\mathbb{P}\left(X_{100} = 12\right) = \frac{\lambda_{100}^{12}}{12!} e^{-\lambda_{100}} = \frac{10^{12}}{12!} e^{-10} \approx 0.095$$

2. Abbiamo $s = 20$ da cui $\lambda_{20} = 2$ e $X_{20} \sim P(2)$.
   Quindi:
   $$\mathbb{P}\left(X_{20} = 0\right) = \frac{\lambda_{20}^{0}}{0!} e^{-\lambda_{20}} = e^{-2} \approx 0.135$$

3. Sia Z_i la variabile che vale 1 se sull'i-esimo DVD si trova almeno una particella di polvere ($i = 1, \ldots, 10$). Le variabili $\{Z_i\}_{i=1}^{10}$ sono i.i.d. con distribuzione di Bernoulli di parametro $p = \mathbb{P}(Z_i = 1) = \mathbb{P}(X_{20} > 0) = 1 - e^{-2} \approx 0.865$. La variabile che conta il numero di DVD in cui si sono rilevate impurità è $Z = \sum_{i=1}^{10} Z_i \sim B(10, p)$. Pertanto
   $$\mathbb{P}(Z = 8) = \binom{10}{8} p^8 (1-p)^2 = 45(1 - e^{-2})^8 e^{-4} \approx 0.257523.$$

4. Di nuovo $s = 100$. Quindi, detto p il valore cercato:
   $$p = \mathbb{P}\left(X_{100} \leq 12\right) = \sum_{k=0}^{12} \frac{10^k}{k!} e^{-10}$$
   Con l'uso del calcolatore si ricava: $p \approx 0.792$.
   Oppure, osservando che la condizione di applicabilità del Teorema del Limite Centrale ($\lambda_{100} > 5$) risulta soddisfatta e ricordando la correzione di continuità, possiamo ricavare la seguente approssimazione normale:
   $$p \approx \Phi\left(\frac{12.5 - 10}{\sqrt{10}}\right) \approx \Phi\left(0.791\right) \approx 0.786$$