Probabilità e Statistica – Corso di Laurea in Informatica
A.A. 2019/2020

ESERCITAZIONE 6

E6.1. In un canale di trasmissione vengono trasmessi simboli binari. I disturbi sul canale fanno sì che ogni simbolo trasmesso abbia la probabilità del 2% di essere ricevuto errato, indipendentemente dagli altri simboli. I messaggi vengono trasmessi in stringhe di 50 simboli. Qual è la probabilità che una stringa venga ricevuta con almeno due simboli errati?

Soluzione. Costruiamo uno schema di Bernoulli. Per $i = 1, \dots, 50$, definiamo le variabili aleatorie

$$X_i = \begin{cases} 1 & \text{se l'}i\text{-esimo simbolo viene ricevuto errato} \\ 0 & \text{altrimenti} \end{cases}$$

Le variabili $X_1, \dots, X_{50}$ sono identicamente distribuite con distribuzione $\mathrm{Be}(0.02)$; inoltre, gli eventi $\{X_1 = 1\}, \dots, \{X_{50} = 1\}$ sono indipendenti (è specificato nel testo). Il numero di simboli errati in una stringa è dato da $X = \sum_{i=1}^{50} X_i \sim \mathrm{Bin}(50, 0.02)$. Pertanto, si ha

$$\begin{aligned} P(X \geq 2) &= 1 - P(X \leq 1) \\ &= 1 - P(X = 0) - P(X = 1) \\ &= 1 - (0.98)^{50} - (50)(0.02)(0.98)^{49} \\ &\simeq 0.264\,. \end{aligned}$$

E6.2. Sia X il numero di successi nelle prime quattro prove di uno schema di Bernoulli di parametro $p = 0.4$. Qual è la probabilità che X sia un numero pari?

Soluzione. Possiamo scrivere $X = \sum_{i=1}^{4} X_i$, dove $X_i \sim \mathrm{Be}(0.4)$ rappresenta l'esito della prova i. La variabile aleatoria X ha distribuzione binomiale $\mathrm{Bin}(4, 0.4)$. Quindi,

$$\begin{aligned} P(X \text{ pari}) &= P\big(\{X = 0\} \cup \{X = 2\} \cup \{X = 4\}\big) \\ &\overset{\text{(eventi disgiunti)}}{=} P(X = 0) + P(X = 2) + P(X = 4) \\ &= (0.6)^4 + 6\,(0.4)^2(0.6)^2 + (0.4)^4 \\ &= 0.5008\,. \end{aligned}$$

E6.3. I semi di una pianta possono essere gialli o verdi. Un certo incrocio tra piante produce una progenie nel rapporto 3 *gialli* : 1 *verde* (cioè la probabilità che un seme prodotto sia giallo è $\frac{3}{4}$). Quattro semi così prodotti vengono scelti a caso. Qual è la probabilità che tre siano gialli e uno verde? Che siano tutti gialli? E che siano tutti dello stesso colore?

Soluzione. Costruiamo uno schema di Bernoulli. Per $i = 1, \dots, 4$, definiamo le variabili aleatorie

$$X_i = \begin{cases} 1 & \text{se l'}i\text{-esimo seme scelto è giallo} \\ 0 & \text{altrimenti} \end{cases}$$

Le variabili $X_1, \dots, X_4$ sono identicamente distribuite con distribuzione Be $\left(\frac{3}{4}\right)$; inoltre, gli eventi $\{X_1 = 1\}, \dots, \{X_4 = 1\}$ sono indipendenti (è ragionevole supporlo, visto dal testo non si evince un'eventuale dipendenza). Il numero di semi gialli tra i quattro scelti è dato da $X = \sum_{i=1}^{4} X_i \sim$ Bin $\left(4, \frac{3}{4}\right)$. Pertanto, si ha

- $P(X = 3) = 4 \left(\frac{3}{4}\right)^3 \left(\frac{1}{4}\right) \simeq 0.422$;
- $P(X = 4) = \left(\frac{3}{4}\right)^4 \simeq 0.316$;
- $P(\{X = 0\} \cup \{X = 4\}) \overset{\text{(eventi disgiunti)}}{=} P(X = 0) + P(X = 4) = \left(\frac{1}{4}\right)^4 + \left(\frac{3}{4}\right)^4 \simeq 0.32.$

E6.4. Si lancia un dado non truccato 10 volte e si indica con X il numero di volte in cui è uscito un numero minore o uguale a 2. Determinare (a) il valor medio e la varianza di X; (b) la probabilità dell'evento $\{X \leq 4\}$ condizionata alla conoscenza dell'evento $\{X \geq 3\}$.

Soluzione. Costruiamo uno schema di Bernoulli. Per $i = 1, \dots, 10$, definiamo le variabili aleatorie

$$X_i = \begin{cases} 1 & \text{se all'}i\text{-esimo lancio di dado esce un numero minore o uguale a 2} \\ 0 & \text{altrimenti} \end{cases}$$

Le variabili $X_1, \dots, X_{10}$ sono identicamente distribuite con distribuzione Be $\left(\frac{1}{3}\right)$; inoltre, gli eventi $\{X_1 = 1\}, \dots, \{X_{10} = 1\}$ sono indipendenti. Abbiamo $X = \sum_{i=1}^{10} X_i \sim$ Bin $\left(10, \frac{1}{3}\right)$. Ora possiamo determinare quanto richiesto.

(a) Usando quanto sappiamo sui parametri riassuntivi delle variabili aleatorie binomiali, otteniamo $E(X) = 10 \cdot \frac{1}{3} = \frac{10}{3}$ e $Var(X) = 10 \cdot \frac{1}{3} \cdot \frac{2}{3} = \frac{20}{9}$.

(b) Usando la definizione di probabilità condizionata, calcoliamo

$$\begin{aligned} P(X \leq 4 | X \geq 3) &= \frac{P(\{X \leq 4\} \cap \{X \geq 3\})}{P(X \geq 3)} \\ &= \frac{P(\{X = 3\} \cup \{X = 4\})}{P(X \geq 3)} \\ &\overset{\text{(eventi disgiunti)}}{=} \frac{P(X = 3) + P(X = 4)}{P(X \geq 3)} \end{aligned}$$

con

$$P(X = 3) = \binom{10}{3} \left(\frac{1}{3}\right)^3 \left(\frac{2}{3}\right)^7 \simeq 0.26$$

$$P(X = 4) = \binom{10}{4} \left(\frac{1}{3}\right)^4 \left(\frac{2}{3}\right)^6 \simeq 0.228$$

$$\begin{aligned} P(X \geq 3) &= 1 - P(X \leq 2) = 1 - P(X = 0) - P(X = 1) - P(X = 2) \\ &= 1 - \left(\frac{2}{3}\right)^{10} - 10 \left(\frac{1}{3}\right) \left(\frac{2}{3}\right)^9 - 45 \left(\frac{1}{3}\right)^2 \left(\frac{2}{3}\right)^8 \simeq 0.701\,. \end{aligned}$$

Quindi otteniamo $P(X \leq 4 | X \geq 3) \simeq 0.696$.

E6.5. Da un'urna contenente 10 palline bianche e 15 palline nere, si eseguono estrazioni *con reinserimento* fino all'estrazione di una pallina nera. Calcolare la probabilità che (a) servano 20 estrazioni; (b) servano almeno 10 estrazioni.

Soluzione. Costruiamo uno schema di Bernoulli. Per $i \geq 1$, definiamo le variabili aleatorie

$$X_i = \begin{cases} 1 & \text{se l'}i\text{-esima pallina estratta è nera} \\ 0 & \text{altrimenti} \end{cases}$$

Le variabili $X_1, X_2, \dots$ sono identicamente distribuite con distribuzione $\mathrm{Be}\left(\frac{3}{5}\right)$; inoltre, gli eventi $\{X_1 = 1\}, \{X_2 = 1\}, \dots$ sono indipendenti (in quanto le estrazioni vengono fatte *con reinserimento*). Siamo interessati all'istante di prima estrazione di una pallina nera, cioè alla variabile aleatoria $T = \inf\{i \geq 1 : X_i = 1\} \sim \mathrm{Ge}\left(\frac{3}{5}\right)$. Pertanto, otteniamo

(a) $P(T = 20) = \left(\dfrac{2}{5}\right)^{19} \left(\dfrac{3}{5}\right)$;

(b) $P(T \geq 10) = P(T > 9) = \left(\dfrac{2}{5}\right)^{9}$, dove abbiamo usato la formula generale per la probabilità di lunga attesa.

E6.6. Da un'urna contenente 5 palline rosse, 4 palline verdi e 6 palline gialle, si eseguono 5 estrazioni *senza reinserimento.* Sia X il numero di palline verdi estratte. Si calcolino valor medio di X, varianza di X e valor medio di $(X + 1)^2$.

Ora supponiamo invece di procedere ad estrazioni *con reinserimento* dalla stessa urna. Denotiamo con T il numero di estrazioni necessarie all'estrazione di una pallina verde. Si calcolino valor medio e varianza di T.

Soluzione. Estrazioni senza reinserimento. La variabile aleatoria X prende valori nell'insieme $\mathcal{X} = \{0, 1, 2, 3, 4\}$. Determiniamo la sua densità discreta. Calcoliamo le probabilità come "casi favorevoli su casi possibili". Ci sono $\binom{15}{5}$ modi di scegliere 5 palline dalle 15 nell'urna (casi possibili). Inoltre, per ogni $k \in \mathcal{X}$,

> $X = k$ se estraiamo k palline verdi e $5 - k$ palline che non lo sono. Scegliamo k palline verdi tra le 4 verdi presenti nell'urna in $\binom{4}{k}$ modi e scegliamo $5 - k$ palline non verdi tra le 11 che non sono verdi in $\binom{11}{5-k}$ modi. Quindi, si ha
>
> $$P(X = k) = \frac{\binom{4}{k}\binom{11}{5-k}}{\binom{15}{5}}.$$

Pertanto la densità discreta di X è data da $p_X(0) = \frac{2}{13}$, $p_X(1) = \frac{40}{91}$, $p_X(2) = \frac{30}{91}$, $p_X(3) = \frac{20}{273}$ e $p_X(4) = \frac{1}{273}$. Calcoliamo

$$E(X) = 1 \cdot \frac{40}{91} + 2 \cdot \frac{30}{91} + 3 \cdot \frac{20}{273} + 4 \cdot \frac{1}{273} = \frac{4}{3}$$

e

$$Var(X) = \underbrace{1^2 \cdot \frac{40}{91} + 2^2 \cdot \frac{30}{91} + 3^2 \cdot \frac{20}{273} + 4^2 \cdot \frac{1}{273}}_{E(X^2)} - \left(\frac{4}{3}\right)^2 = \frac{52}{21} - \frac{16}{9} = \frac{44}{63}.$$

Inoltre, sfruttando la linearità del valor medio, otteniamo

$$E[(X + 1)^2] = E[X^2 + 2X + 1] = E(X^2) + 2E(X) + 1 = \frac{52}{21} + 2 \cdot \frac{4}{3} + 1 = \frac{43}{7}.$$

<u>Estrazioni con reinserimento</u>. In questo caso, *poiché con reinserimento, le estrazioni successive sono tra loro indipendenti* e possiamo metterci nel contesto dello schema di Bernoulli. Consideriamo quindi, per $i \geq 1$, le variabili aleatorie

$$X_i = \begin{cases} 1 & \text{se l'}i\text{-esima pallina estratta è verde} \\ 0 & \text{altrimenti} \end{cases}$$

Si ha $X_i \sim \text{Be}\left(\frac{4}{15}\right)$, per ogni $i \geq 1$. La variabile aleatoria T è l'istante di prima estrazione di una pallina verde, pertanto $T = \inf\{i \geq 1 : X_i = 1\} \sim \text{Ge}\left(\frac{4}{15}\right)$.
Usando quanto sappiamo sui parametri riassuntivi delle variabili aleatorie geometriche, otteniamo $E(T) = \frac{15}{4}$ e $Var(T) = \frac{11}{15} \cdot \left(\frac{15}{4}\right)^2 = \frac{165}{16}$.

E6.7. Sia $X \sim \text{Ge}(p)$ e tale che $P(X = 1) = 2P(X = 2)$. Determinare il valore del parametro p e calcolare poi la probabilità dell'evento $\{X > E(X)\}$.

Soluzione. Usando la definizione della densità discreta per una variabile aleatoria geometrica otteniamo

$$P(X = 1) = 2P(X = 2) \quad \Leftrightarrow \quad p = 2(1-p)p \quad \Leftrightarrow \quad p = \frac{1}{2}$$

(abbiamo scartato il caso $p = 0$). Di conseguenza $E(X) = 2$ e, per la probabilità di lunga attesa, otteniamo $P(X > E(X)) = P(X > 2) = \left(1 - \frac{1}{2}\right)^2 = \frac{1}{4}$.

E6.8. Nella nostra dotazione di 1200 microchip, 500 unità sono fornite dalla ditta A e 700 dalla ditta B. Da controlli di qualità risulta che la produzione della ditta A è difettosa con probabilità $p_{\mathsf{A}} = 0.005$ e quella della ditta B con probabilità $p_{\mathsf{B}} = 0.002$. In media, quanti microchip possiamo scegliere (casualmente) prima di trovarne uno difettoso ?

Soluzione. Poiché dal testo non si evince il contrario, assumiamo che non ci sia dipendenza nella difettosità di microchip distinti e costruiamo uno schema di Bernoulli. Per $i = 1, \ldots, 1200$, definiamo le variabili aleatorie

$$X_i = \begin{cases} 1 & \text{se l'}i\text{-esimo microchip scelto è difettoso} \\ 0 & \text{altrimenti} \end{cases}$$

Le variabili aleatorie $X_1, \ldots, X_{1200}$ sono identicamente distribuite con distribuzione $\text{Be}(p)$. Dobbiamo determinare il parametro p.
Consideriamo gli eventi $F_i =$"il microchip scelto è stato fornito dalla ditta i" (con $i = \mathsf{A}, \mathsf{B}$) e $D =$"il microchip scelto è difettoso". Dal testo ricaviamo $P(F_{\mathsf{A}}) = \frac{5}{12}$, $P(F_{\mathsf{B}}) = \frac{7}{12}$, $P(D|F_{\mathsf{A}}) = p_{\mathsf{A}}$ e $P(D|F_{\mathsf{B}}) = p_{\mathsf{B}}$. Pertanto, si ottiene

$$p = P(D) \overset{\text{(prob. totali)}}{=} P(D|F_{\mathsf{A}})P(F_{\mathsf{A}}) + P(D|F_{\mathsf{B}})P(F_{\mathsf{B}}) = \frac{5}{12}(0.005) + \frac{7}{12}(0.002) = 0.00325.$$

Se $T = \inf\{i = 1, \ldots, 1200 : X_i = 1\} \sim \text{Ge}(0.00325)$ rappresenta l'istante in cui scegliamo il primo microchip difettoso, otteniamo $E(T) = \frac{1}{0.00325} \simeq 307.7$. In media, potremo quindi scegliere 307 microchip prima di trovarne uno difettoso.

E6.9 (⚑ tratto da appello; difficile). Sul tavolo ci sono due dadi indistinguibili. Entrambi i dadi sono truccati: il primo in modo che la probabilità di ottenere 6 sia il doppio della probabilità di ottenere ogni altro punteggio; il secondo in modo che la probabilità di ottenere 6 sia la metà della probabilità di ottenere ogni altro punteggio. Qual è la probabilità di ottenere 6 lanciando il primo dado? E lanciando il secondo?

Si sceglie a caso uno dei due dadi e con il dado scelto si effettuano dei lanci ripetuti. Qual è la probabilità di ottenere il punteggio 6, per la prima volta, al terzo lancio?

Supponiamo ora che, con il dado scelto, vengano eseguiti 10 lanci in totale. Qual è la probabilità di ottenere 3 volte un 6? Sapendo che il punteggio 6 è stato ottenuto 3 volte, qual è la probabilità che i lanci siano stati effettuati con il primo dado?

Soluzione.

- Ricaviamo la descrizione probabilistica dei due dadi.

  Dado 1 – la probabilità di ottenere 6 è il doppio della probabilità di ottenere ogni altro punteggio.
  Se poniamo $x = P(\text{ottenere } i)$, con $i \in \{1, \ldots, 5\}$, allora $P(\text{ottenere } 6) = 2x$. Per la condizione di normalizzazione deve valere $5x + 2x = 1$, da cui segue $x = \frac{1}{7}$.

  Dado 2 – la probabilità di ottenere 6 è la metà della probabilità di ottenere ogni altro punteggio.
  Se poniamo $x = P(\text{ottenere } i)$, con $i \in \{1, \ldots, 5\}$, allora $P(\text{ottenere } 6) = \frac{x}{2}$. Per la condizione di normalizzazione deve valere $5x + \frac{x}{2} = 1$, da cui segue $x = \frac{2}{11}$.

  Pertanto, concludendo, la probabilità di ottenere 6 lanciando il primo dado è pari a $\frac{2}{7}$, mentre lanciando il secondo è $\frac{1}{11}$.

- Consideriamo gli eventi $E =$"si ottiene 6, per la prima volta, al terzo lancio" e $D_i =$"viene lanciato il dado i", con $i = 1, 2$. Per la formula delle probabilità totali abbiamo

$$P(E) = P(E|D_1)P(D_1) + P(E|D_2)P(D_2).$$

  Poiché il dado lanciato viene scelto a caso, si ha $P(D_1) = P(D_2) = \frac{1}{2}$. Per calcolare $P(E|D_1)$ e $P(E|D_2)$ costruiamo due schemi di Bernoulli, ciascuno associato ai lanci consecutivi di uno specifico dado.

  Dado 1 – Per $i \geq 1$, definiamo le variabili aleatorie

$$X_i = \begin{cases} 1 & \text{se all'}i\text{-esimo lancio del dado 1 esce un 6} \\ 0 & \text{altrimenti.} \end{cases}$$

  Le variabili aleatorie $X_1, X_2, \ldots$ sono indipendenti ed identicamente distribuite con distribuzione $\mathrm{Be}\left(\frac{2}{7}\right)$. Siamo interessati al lancio in cui esce un 6 per la prima volta (istante di primo successo), cioè alla variabile aleatoria $T_1 = \inf\{i \geq 1 : X_i = 1\} \sim \mathrm{Ge}\left(\frac{2}{7}\right)$. Abbiamo

$$P(E|D_1) = P(T_1 = 3) = \frac{2}{7}\left(1 - \frac{2}{7}\right)^2 = \frac{50}{343}.$$

  Dado 2 – Per $i \geq 1$, definiamo le variabili aleatorie

$$Y_i = \begin{cases} 1 & \text{se all'}i\text{-esimo lancio del dado 2 esce un 6} \\ 0 & \text{altrimenti.} \end{cases}$$

  Le variabili aleatorie $Y_1, Y_2, \ldots$ sono indipendenti ed identicamente distribuite con distribuzione $\mathrm{Be}\left(\frac{1}{11}\right)$. Siamo interessati al lancio in cui esce un 6 per la prima volta (istante di primo successo), cioè alla variabile aleatoria $T_2 = \inf\{i \geq 1 : Y_i = 1\} \sim \mathrm{Ge}\left(\frac{1}{11}\right)$. Abbiamo

$$P(E|D_2) = P(T_2 = 3) = \frac{1}{11}\left(1 - \frac{1}{11}\right)^2 = \frac{100}{1331}.$$

  Pertanto, concludendo, si ottiene $P(E) = \frac{50}{343} \cdot \frac{1}{2} + \frac{100}{1331} \cdot \frac{1}{2} \approx 0.11$.

- Consideriamo gli eventi $F =$"il punteggio 6 viene ottenuto 3 volte nella sequenza di 10 lanci" e $D_i =$"viene lanciato il dado i", con $i = 1, 2$. Per la formula delle probabilità totali abbiamo

$$P(F) = P(F|D_1)P(D_1) + P(F|D_2)P(D_2).$$

Poiché il dado lanciato viene scelto a caso, si ha $P(D_1) = P(D_2) = \frac{1}{2}$. Per calcolare $P(F|D_1)$ e $P(F|D_2)$ usiamo i due schemi di Bernoulli introdotti al punto (b). Consideriamo però solo le prime 10 variabili.

Dado 1 – Le variabili $X_1, \ldots, X_{10}$ sono indipendenti ed identicamente distribuite con distribuzione Be $\left(\frac{2}{7}\right)$. Il numero totale di punteggi 6 ottenuti nella sequenza di 10 lanci è dato dalla variabile aleatoria $X = \sum_{i=1}^{10} X_i \sim$ Bin $\left(10, \frac{2}{7}\right)$. Quindi, si ha

$$P(F|D_1) = P(X = 3) = \binom{10}{3}\left(\frac{2}{7}\right)^3\left(1 - \frac{2}{7}\right)^7 \approx 0.266.$$

Dado 2 – Le variabili $Y_1, \ldots, Y_{10}$ sono indipendenti ed identicamente distribuite con distribuzione Be $\left(\frac{1}{11}\right)$. Il numero totale di punteggi 6 ottenuti nella sequenza di 10 lanci è dato dalla variabile aleatoria $Y = \sum_{i=1}^{10} Y_i \sim$ Bin $\left(10, \frac{1}{11}\right)$. Quindi, si ha

$$P(F|D_2) = P(Y = 3) = \binom{10}{3}\left(\frac{1}{11}\right)^3\left(1 - \frac{1}{11}\right)^7 \approx 0.046.$$

Pertanto, concludendo, si ottiene $P(F) = (0.266)\frac{1}{2} + (0.046)\frac{1}{2} = 0.156$.

- Usando la formula di Bayes, ricaviamo

$$P(D_1|F) = \frac{P(F|D_1)P(D_1)}{P(F)} \approx 0.853.$$

E6.10. Il numero di ciclisti che in un'ora riescono a raggiungere la cima del famigerato *Colle della Probabilità* è distribuito come una variabile aleatoria di Poisson di media 2. Calcolare la probabilità che (a) più di tre ciclisti arrivino alla cima in un'ora; (b) nessun ciclista arrivi alla cima in un'ora.
Soluzione. Se indichiamo con $X \sim$ Po(2) la variabile aleatoria che corrisponde al numero di ciclisti che raggiungono la cima del Colle della Probabilità in un'ora, abbiamo

(a) $P(X > 3) = 1 - P(X \leq 3) = 1 - \sum_{k=0}^{3} P(X = k) = 1 - e^{-2}\,\frac{19}{3}$;

(b) $P(X = 0) = e^{-2}$.

E6.11. La probabilità che un individuo sia allergico al farmaco XXX è $p = 0.002$. Si sceglie un campione casuale di 1500 persone.

(a) Determinare un'espressione *esatta* (non cercare di calcolare il valore numerico) per la probabilità che 3 di queste persone siano allergiche al farmaco, mentre le altre 1497 non lo siano.

(b) Determinare, usando un'opportuna approssimazione, il valore numerico della probabilità richiesta nel punto (a).

Soluzione. Costruiamo uno schema di Bernoulli. Per $i = 1, \dots, 1500$, definiamo le variabili aleatorie

$$X_i = \begin{cases} 1 & \text{se l'}i\text{-esimo individuo è allergico al farmaco XXX} \\ 0 & \text{altrimenti} \end{cases}$$

Le variabili $X_1, \dots, X_{1500}$ sono identicamente distribuite con distribuzione $\mathrm{Be}\,(0.002)$; inoltre gli eventi $\{X_1 = 1\}, \dots, \{X_{1500} = 1\}$ sono indipendenti. Il numero di persone (nel campione) allergiche al farmaco è dato dalla variabile aleatoria $X = \sum_{i=1}^{1500} X_i \sim \mathrm{Bin}\,(1500, 0.002)$.

(a) Si ha $P(X = 3) = \binom{1500}{3}(0.002)^3(0.998)^{1497}$.

(b) Per l'approssimazione di Poisson, X ha approssimativamente distribuzione di Poisson di parametro $(1500)(0.002) = 3$. Perciò abbiamo

$$P(X = 3) \simeq e^{-3}\,\frac{3^3}{3!} \simeq 0.224\,.$$

E6.12. Durante uno sciopero dei trasporti, il centralino del numero verde delle ferrovie risponde in media a 1.7 telefonate al minuto. Si supponga che il numero di risposte segua una distribuzione di Poisson. Calcolare: (a) la probabilità che tra le 11.00 e le 11.02 dal centralino rispondano ad una chiamata; (b) la probabilità che tra le 16.00 e le 16.03 giungano al più due chiamate; (c) il numero medio di telefonate che giungono tra le 09.00 e le 18.00.

Soluzione.

(a) Dato che l'intervallo di tempo da considerare è di 2 minuti, il centralino risponde in media a $2 \cdot (1.7) = 3.4$ chiamate. Pertanto, il numero di risposte è descritto da $X \sim \mathrm{Po}(3.4)$ e la probabilità richiesta è $P(X = 1) = e^{-3.4} \cdot (3.4) \simeq 0.1135$.

(b) In questo caso, l'intervallo di tempo da considerare è di 3 minuti e quindi il centralino risponde in media a $3 \cdot (1.7) = 5.1$ chiamate. Pertanto, il numero di risposte è descritto da $Y \sim \mathrm{Po}(5.1)$ e si ha

$$P(Y \le 2) = \sum_{k=0}^{2} P(Y = k) = e^{-5.1}\left(1 + 5.1 + \frac{(5.1)^2}{2}\right) \simeq 0.1165.$$

(c) Poiché 9 ore corrispondono a 540 minuti, il numero medio di chiamate che giungono al centralino tra le 09.00 e le 18.00 è $540 \cdot (1.7) = 918$.

E6.13 (Calcolare l'affidabilità di un software). Quando si sviluppa un pacchetto software, prima di renderlo disponibile, si effettuano dei test per trovare ed eliminare eventuali bug. Tipicamente si usa il pacchetto per risolvere un insieme di problemi noti, così da vedere se si verificano degli errori, che eventualmente vengono annotati. Quando il periodo di prova termina, il pacchetto viene revisionato in modo da eliminare tutti i bug presenti.

È usuale modellare il numero di errori che si verificano a causa di un bug con la distribuzione di Poisson. Supponiamo che il numero medio di errori al minuto dovuto ad un particolare bug sia pari a 0.0001. (a) Qual è la probabilità che in 20 minuti non si verifichi alcun errore? (b) Per quanto tempo devo far girare il programma se voglio che si verifichi almeno un errore con probabilità 99.95% (così da evidenziare la presenza di questo bug)?

Soluzione.

(a) Dato che l'intervallo di tempo da considerare è di 20 minuti, il numero medio di errori verificatisi in tale lasso di tempo è pari a $20 \cdot (0.0001) = 0.002$. Pertanto, il numero di errori verificatisi in 20 minuti è descritto dalla variabile aleatoria $X \sim \mathrm{Po}(0.002)$ e la probabilità richiesta è $P(X = 0) = e^{-0.002} \simeq 0.998$.

(b) Consideriamo un intervallo di k minuti, dove k è l'incognita del nostro problema ed è quindi da determinare. Il numero medio di errori verificatisi in un lasso di tempo di k minuti è pari a $k \cdot (0.0001)$. Pertanto, il numero di errori verificatisi in k minuti è descritto dalla variabile aleatoria $X_k \sim \mathrm{Po}((0.0001)k)$. Dobbiamo determinare per quali valori di k si ottiene $P(X_k \geq 1) \geq 0.9995$. Si ha

$$\begin{aligned} P(X_k \geq 1) \geq 0.9995 \quad &\Leftrightarrow \quad 1 - P(X_k = 0) \geq 0.9995 \\ &\Leftrightarrow \quad P(X_k = 0) \leq 0.0005 \\ &\Leftrightarrow \quad e^{-(0.0001)k} \leq 0.0005, \end{aligned}$$

da cui si ricava (circa) $k \geq 76\,009$. Quindi, per riuscire ad evidenziare la presenza del bug con una probabilità del 99.95%, dobbiamo far girare il programma per circa 76 009 minuti, che corrispondono a circa 1267 ore, cioè circa 53 giorni.