Capitolo 5

Principali distribuzioni di probabilità continue

In questo capitolo presentiamo alcune distribuzioni di probabilità assolutamente continue.

5.1 La distribuzione uniforme

La distribuzione uniforme è la distribuzione di eventi equiprobabili. Nel caso continuo, una variabile aleatoria X si dice che segue una *distribuzione uniforme* nell'intervallo $[a, b]$, con $-\infty < a < b < \infty$, se ha la seguente funzione di densità:

$$f(x) := \begin{cases} \dfrac{1}{b-a} & \text{se } x \in [a, b] \\ 0 & \text{altrimenti.} \end{cases}$$

Scriveremo allora $X \sim U(a, b)$. La speranza matematica e la varianza di una variabile aleatoria $X \sim U(a, b)$ sono date rispettivamente da:

$$\mathbb{E}(X) := \frac{a+b}{2}$$
$$\text{Var}(X) := \frac{(b-a)^2}{12} .$$

Figura 5.1: Funzione di densità di distribuzione uniforme continua su (a, b).

Il paradosso di Bertrand. Si sceglie a caso una corda di un cerchio di raggio R e centro C. Qual è la probabilità p che sia maggiore del lato del triangolo equilatero inscritto?

Il problema ha più di una soluzione in quanto non risulta chiaro il significato della frase "scegliere una corda a caso". Indichiamo con l la lunghezza del lato del triangolo equilatero inscritto.

1. Considerata una corda, sia A il punto intermedio. La corda è più lunga del lato del triangolo equilatero se il punto A ricade all'interno del cerchio di raggio $R/2$ e centro C. Il rapporto fra tali aree è 1/4 e pertanto risulta $p = 1/4$.

2. In base alla simmetria, un estremo della corda può essere scelto sulla circonferenza del cerchio in corrispondenza di uno dei vertici del triangolo. I vertici del triangolo dividono la circonferenza in tre archi di uguale lunghezza e pertanto si ha $p = 1/3$.

3. Scelto un punto B a caso, uniformemente sul raggio del cerchio, consideriamo la corda perpendicolare a tale raggio passante per B. Allora la corda casuale ha lunghezza maggiore di di l se il punto B appartiene alla metà del raggio più vicina al centro. In base alla simmetria, non importa quale raggio viene scelto e pertanto si ha $p = 1/2$.

Spiegazione . Ciascuno dei tre metodi utilizza una diversa distribuzione uniforme: 1) sul cerchio; 2) sulla circonferenza; 3) sul raggio del cerchio. ♣

5.2 La distribuzione esponenziale negativa

Abbiamo visto in precedenza che, sotto opportune ipotesi, il numero di volte X in cui un certo evento A si verifica in un intervallo di tempo di lunghezza T può descriversi mediante una distribuzione di Poisson:

$$P(X = x) := \mathrm{e}^{-\lambda T} \frac{(\lambda T)^x}{x!} \qquad x = 0, 1, 2, \ldots$$

Se si indica con X la v.a. che rappresenta il tempo necessario affinchè l'evento A si manifesti per la prima volta, allora X è una v.a. continua. La distribuzione di X si chiama *distribuzione esponenziale negativa,* e scriveremo $X \sim \mathcal{E}(\lambda)$; la sua funzione di densità è data da:

$$f(x) := \begin{cases} \lambda\, \mathrm{e}^{-\lambda x} & \text{per } x > 0 \\ 0 & \text{altrove.} \end{cases} \tag{5.1}$$

Si noti che la (5.1) può essere scritta sinteticamente come:

$$f(x) \;=\; \lambda\, \mathrm{e}^{-\lambda x}\, I_{(0,+\infty)}(x)\ .$$

La verifica della (4.3) porge:

$$\int_{-\infty}^{+\infty} \lambda\, \mathrm{e}^{-\lambda x}\, I_{(0,+\infty)}(x)\, \mathrm{d}x = \int_{0}^{+\infty} \lambda\, \mathrm{e}^{-\lambda x}\, \mathrm{d}x = \left[-\mathrm{e}^{-\lambda x}\right]_0^{+\infty} = 1\ .$$

La funzione di ripartizione è data da:

$$F(x) := \int_{-\infty}^{x} \lambda \mathrm{e}^{-\lambda t} I_{(0,+\infty)}(t)\, \mathrm{d}t = \int_{0}^{x} \mathrm{e}^{-\lambda t}\, d(\lambda t) = \left[-\mathrm{e}^{-\lambda t}\right]_0^x = 1 - \mathrm{e}^{-\lambda x} .$$

La funzione:

$$S(x) := P(X > x) = 1 - F(x) = \mathrm{e}^{-\lambda x}$$

viene chiamata *funzione di sopravvivenza*. Essa esprime la probabilità che la v.a. X assuma un valore superiore a x.

La speranza matematica, integrando per parti, è data da:

$$\begin{aligned} \mathbb{E}[X] &= \int_{-\infty}^{+\infty} x\lambda\, \mathrm{e}^{-\lambda x} I_{(0,+\infty)}(x)\, \mathrm{d}x = \int_{0}^{+\infty} \lambda x\, \mathrm{e}^{-\lambda x}\, \mathrm{d}x \\ &= \left[x(-\mathrm{e}^{-\lambda x})\right]_0^{+\infty} + \int_0^{+\infty} \mathrm{e}^{-\lambda x}\, dx = \frac{1}{\lambda} . \end{aligned}$$

In maniera analoga si ottiene:

$$\mathrm{Var}(X) = \frac{1}{\lambda^2}$$

e quindi la distribuzione esponenziale negativa è caratterizzata dal fatto che la speranza matematica è uguale allo scarto quadratico medio.

Al pari della distribuuzione geometrica, la distribuzione esponenziale negativa possiede la proprietà di *assenza di memoria*, infatti:

$$\begin{aligned} P(X \leq t + \Delta t | X > t) &= \frac{P(X \leq t + \Delta t, X > t)}{P(X > t)} = \frac{P(t < X \leq t + \Delta t)}{P(X > t)} \\ &= \frac{\mathrm{e}^{-\lambda t} - \mathrm{e}^{-\lambda(t+\Delta t)}}{\mathrm{e}^{-\lambda t}} = 1 - \mathrm{e}^{-\lambda \Delta t} = P(X \leq \Delta t) \end{aligned}$$

essendo $\Delta t \geq 0$.

Si noti che la distribuzione esponenziale negativa può essere anche interpretata, invece che come tempo d'attesa di un evento, anche come legge di densità dell'intervallo di tempo che separa gli istanti in cui si verificano due eventi aleatori in accordo ad una legge di Poisson.

Esercizio 5.1 Un sistema S è composto da due componenti in serie A e B, i cui tempi di durata fino al guasto sono quantità aleatorie indipendenti, denotate rispettivamente con X e Y. Il vettore aleatorio (X, Y) ha una densità di probabilità $f(x, y) = k\mathrm{e}^{-x-2y}$, per $x > 0$ e $y > 0$. Calcolare

1. Il valore della costante k;
2. La funzione di sopravvivenza $S(z) = P(Z > z)$ del tempo di guasto Z di S;
3. la previsione $\mathbb{E}(Z)$ di Z.

Soluzione. In base alle ipotesi, la funzione di densità di (X, Y) è data da:

$$f(x,y) = \begin{cases} k\mathrm{e}^{-x-2y} & \text{per } x > 0, y > 0 \\ 0 & \text{altrimenti.} \end{cases}$$

In base alle ipotesi deve risultare:

$$\int_0^\infty \int_0^\infty f(x,y)\, dxdy = 1$$

cioè:

$$\int_0^\infty \int_0^\infty \mathrm{e}^{-x-2y}\, dxdy = \frac{1}{k}$$

da cui

$$\int_0^\infty \mathrm{e}^{-x}\, dx \int_0^\infty \mathrm{e}^{-2y}\, dy = 1 \cdot \frac{1}{2} = \frac{1}{k}$$

e quindi $k = 2$. La v.a. Z è data da $Z = \min(X, Y)$ e quindi:

$$P\{Z > z\} = P\{X > z\}P\{Y > z\} = \int_z^\infty \mathrm{e}^{-x}\, dx \int_z^\infty \mathrm{e}^{-2y}\, dy = \mathrm{e}^{-z} \cdot \mathrm{e}^{-2z} = \mathrm{e}^{-3z} \ ,$$

e quindi Z segue una distribuzione esponenziale di parametro $\lambda = 3$, cioè $Z \sim \mathcal{E}(3)$, e quindi $\mathbb{E}[Z] = \frac{1}{3}$. ♣

Proposizione 5.2 Siano X e Y due v.a. indipendenti aventi distribuzione esponenziale di parametro λ. Allora $Z = X/(X + Y)$ segue una distribuzione $U(0, 1)$.

5.3 La distribuzione normale

La distribuzione normale o gaussiana è la distribuzione che più di ogni altra trova applicazione in statistica. La ragione principale di ciò risiede nel fatto che essa costituisce un modello che approssima numerose altre distribuzioni e possiede notevoli proprietà matematiche.

Distribuzione normale standard. Diremo che una variabile aleatoria Z segue una *distribuzione normale standard* o gaussiana standard, e scriveremo $Z \sim N(1, 0)$, se ha la seguente funzione di densità:

$$f(z) := \frac{1}{\sqrt{2\pi}}\, \mathrm{e}^{-z^2/2} \qquad z \in \mathbb{R} \ . \tag{5.2}$$

Per dimostrare che effettivamente la (5.2) è effettivamente una funzione di densità, dobbiamo dimostrare che vale la proprietà (4.3). A tal scopo, dimostriamo la seguente relazione:

$$\int_{-\infty}^{+\infty} \mathrm{e}^{-z^2/2}\, dz = \sqrt{2\pi} \ . \tag{5.3}$$

Consideriamo la quantità:

$$\left(\int_{-\infty}^{+\infty} \mathrm{e}^{-z^2/2}\,dz\right)^2 = \int_{-\infty}^{+\infty} \mathrm{e}^{-x^2/2}\,dx \cdot \int_{-\infty}^{+\infty} \mathrm{e}^{-y^2/2}\,dy = \int_{-\infty}^{+\infty}\int_{-\infty}^{+\infty} \mathrm{e}^{(x^2+y^2)/2}\,dx\,dy\ .$$

Considerata la trasformazione in coordinate polari $(x, y) \to (\rho, \theta)$ definita da $x = \rho\cos\theta$ e $y = \rho\sin\theta$ e considerato che il modulo dello jacobiano[1] è uguale a ρ, si ottiene:

$$\int_{-\infty}^{+\infty}\int_{-\infty}^{+\infty} \mathrm{e}^{(x^2+y^2)/2}\,dx\,dy = \int_{0}^{+\infty}\int_{0}^{2\pi} \mathrm{e}^{-\rho^2/2}\rho\,d\rho\,d\theta = 2\pi\int_{0}^{+\infty} \mathrm{e}^{-\rho^2/2}\rho\,d\rho = 2\pi\ .$$

Si ha quindi la (5.3), da cui segue:

$$\int_{-\infty}^{+\infty} \frac{1}{\sqrt{2\pi}} e^{-z^2/2}\,dz = 1\ .$$

La speranza $\mathbb{E}[Z]$ è data da:

$$\begin{aligned}\mathbb{E}[Z] &= \int_{-\infty}^{\infty} z f(z)\,dz = \frac{1}{\sqrt{2\pi}}\int_{-\infty}^{\infty} z e^{-z^2/2}\,dz \\ &= \frac{1}{\sqrt{2\pi}}\left(-\int_{-\infty}^{0} e^{-z^2/2} d(\frac{z^2}{2}) - \int_{0}^{+\infty} e^{-z^2/2} d(\frac{z^2}{2})\right) \\ &= \frac{1}{\sqrt{2\pi}}\left(\left[-e^{-z^2/2}\right]_{-\infty}^{0} - \left[-e^{-z^2/2}\right]_{0}^{+\infty}\right) = \frac{1}{\sqrt{2\pi}}(-1+1) = 0\ .\end{aligned}$$

Per quanto riguarda la varianza, poichè $\mathbb{E}[Z] = 0$, allora segue Var$(Z) = \mathbb{E}[Z^2]$ e, in base al teorema di integrazione per parti si ha:

$$\begin{aligned}\mathrm{Var}(Z) = \mathbb{E}[Z^2] &= \frac{1}{\sqrt{2\pi}}\int_{-\infty}^{\infty} z^2 e^{-z^2/2}\,dz = \frac{1}{\sqrt{2\pi}}\int_{-\infty}^{\infty} z\cdot(z e^{-z^2/2})\,dz \\ &= \left[-z\cdot e^{-z^2/2}\right]_{-\infty}^{\infty} + \int_{-\infty}^{\infty} e^{-z^2/2}\,dz = 1\ .\end{aligned}$$

La funzione di ripartizione della distribuzione normale standard viene usualmente denotata col simbolo $\Phi(z)$, cioè

$$\Phi(z) := P(Z \le z) = \frac{1}{\sqrt{2\pi}}\int_{-\infty}^{z} \mathrm{e}^{-t^2/2}\,dt\ .$$

Si noti che, per la simmetria della distribuzione normale standard, si ha:

$$\Phi(-z) = 1 - \Phi(z) \qquad \text{per ogni } z \in \mathbb{R}\ .$$

In particolare, dalla relazione precedente si ottiene la seguente relazione fra i quantili α e $(1-\alpha)$, con $0 < \alpha < 1$ di una distribuzione normale standard:

$$z_\alpha = -z_{1-\alpha} \qquad 0 < \alpha < 1\ , \tag{5.4}$$

che risulta molto utile nel calcolo pratico dei quantili di una distribuzione normale standard.

[1] Per il calcolo dello Jacobiano, si rimanda ai testi di analisi matematica.

Famiglia delle distribuzioni normali. Sia $Z \sim N(0,1)$ e siano $\mu \in \mathbb{R}$ e $\sigma > 0$. Consideriamo la v.a.

$$X := \sigma Z + \mu \, . \tag{5.5}$$

In base alla (4.24) la v.a. X ha funzione di densità:

$$f(x) = \frac{1}{\sqrt{2\pi\sigma^2}} \exp\left(-\frac{(x-\mu)^2}{2\sigma^2}\right) \qquad x \in \mathbb{R} \, . \tag{5.6}$$

In questo caso diremo che la v.a. X segue una *distribuzione normale* di parametri μ e σ e scriveremo $X \sim N(\mu, \sigma^2)$. Osserviamo che per $\mu = 0$ e $\sigma^2 = 1$ dalla (5.6) si ottiene la densità della normale standard (5.2)

In base alla (4.24), dalla (5.5), si ottiene la funzione di ripartizione di X:

$$F_X(x) := P(X \leq x) = \Phi\left(\frac{x-\mu}{\sigma}\right) \, . \tag{5.7}$$

Con ragionamenti analoghi al caso standard, otteniamo la speranza matematica e la varianza di una variabile aleatoria $X \sim N(\mu, \sigma^2)$:

$$\begin{aligned} \mathbb{E}(X) &= \mu \, , \\ \text{Var}(X) &= \sigma^2 \, . \end{aligned}$$

Proposizione 5.3 Siano $X_1, \ldots, X_n$ v.a. indipendenti con $X_k \sim N(\mu_k, \sigma_k^2)$, per $k = 1, \ldots, n$. Allora il v.a. $X_1 + \cdots + X_n$ segue una distribuzione $N(\sum_{k=1}^n \mu_k, \sum_{k=1}^n \sigma_k^2)$.

Più in generale se $X_1, \ldots, X_n$ sono v.a. indipendenti aventi distribuzione rispettivamente $N(\mu_k, \sigma_k^2)$, con $k = 1, \ldots, n$ e $\alpha_1, \ldots, \alpha_n$ sono numeri reali, allora la variabile aleatoria $\sum_{k=1}^n \alpha_k X_k$ ha distribuzione $N(\sum_{k=1}^n \alpha_k \mu_k, \sum_{k=1}^n \alpha_k^2 \sigma_k^2)$.

Proposizione 5.4 Siano X, Y v.a. indipendenti con distribuzione $N(\mu_1, \sigma_1^2)$ e $N(\mu_2, \sigma_2^2)$. Allora $X + Y$ e $X - Y$ sono indipendenti.

Esercizio 5.5 Sia X una v.a. avente distribuzione normale con media $\mu = 10$ e varianza $\sigma^2 = 16$. Calcolare

a) $P(X > 12)$;

b) $P(7.5 \leq X \leq 11.2)$;

c) $P(|X - 10| < 2)$;

d) $P(X > 12 | X > 9)$.

Soluzione. Essendo $X \sim N(10, 10)$, per i punti a) e b) si ha rispettivamente;

$$\begin{aligned}
P(X > 12) &= P\left(\frac{X-10}{4} > \frac{12-10}{4}\right) = P(Z > 0.5) = 1 - \Phi(0.5) \\
&= 1 - 0.6915 = 0.3085 \\
P(7.5 \leq X \leq 11.2) &= P\left(\frac{7.5-10}{4} \leq \frac{X-10}{4} \leq \frac{11.2-10}{4}\right) \\
&= P(-0.625 \leq Z \leq 0.3) = \Phi(0.3) - \Phi(-0.625) \\
&= \Phi(0.3) - 1 + \Phi(0.625) = 0.6179 - 1 + 0.7340 = 0.3519\,.
\end{aligned}$$

dove Z denota la v.a. normale standard. Per quanto riguarda il terzo punto, osserviamo preliminarmente che la relazione $|X - 10| < 2$ è equivalente a

$$-2 < X - 10 < 2 \qquad \text{da cui segue:} \quad 8 < X < 12$$

e quindi:

$$\begin{aligned}
P(|X-10| < 2) &= P(8 < X < 12) P\left(\frac{8-10}{4} \leq \frac{X-10}{4} \leq \frac{12-10}{4}\right) \\
&= P(-0.5 \leq Z \leq 0.5) = \Phi(0.5) - \Phi(-0.5) \\
&= \Phi(0.5) - 1 + \Phi(0.5) = 0.6915 - 1 + 0.6915 = 0.3829\,.
\end{aligned}$$

Infine, tenendo conto che

$$P(X > 12 | X > 9) = \frac{P(\{X > 12\} \cap \{X > 9\})}{P(X > 9)} = \frac{P(X > 12)}{P(X > 9)}$$

ed essendo, in base al punto a) $P(X > 12) = 0.3085$, dobbiamo ricavare preliminarmente solo $P(X > 9)$

$$\begin{aligned}
P(X > 9) &= P\left(\frac{9-10}{4} > \frac{9-10}{4}\right) = P(Z > -0.25) = 1 - \Phi(0.25) \\
&= 1 - 0.5987 = 0.4013
\end{aligned}$$

da cui segue

$$P(X > 12 | X > 9) = \frac{P(X > 12)}{P(X > 9)}) = \frac{0.3085}{0.4013} = 0.7688\,.$$

♣

Esercizio 5.6 Sia X una v.a. normale avente speranza matematica $\mu = 4$. Calcolare:

1. la varianza σ^2 di X sapendo che $P(X > 5) = 0.3707$;
2. $P(|X - 5| < 2)$;
3. $P(X > 7 | X > 4)$.

Soluzione. Per quanto riguarda il primo punto, dall'equazione $P(X > 5) = 0.3707$ segue:

$$P(X > 5) = P\left(\frac{X-4}{\sigma} > \frac{5-4}{\sigma}\right) = P\left(Z > \frac{1}{\sigma}\right) = 0.3707$$

da cui:

$$\Phi\left(\frac{1}{\sigma}\right) = 1 - \Phi\left(\frac{1}{\sigma}\right) = 1 - 0.3707 = 0.6293\,.$$

Indicato con $z_{0.6293}$ il quantile 0.6293 (cioè il valore per cui la funzione di ripartizione assume valore 0.6293), dalle tavole si ottiene $z_{0.6293} = 0.33$ da cui:

$$z_{0.6293} = \frac{1}{\sigma} = 0.33$$

e quindi $\sigma = 3$. Pertanto $X \sim N(4, 9)$.

Per quanto concerne il secondo punto, preliminarmente conviene scrivere l'insieme $\{|X - 5| < 2\}$ come $\{-2 < X - 5 < 2\}$ cioè $\{3 < X < 7\}$. Si ha quindi:

$$\begin{aligned}
P(3 < X < 7) &= P\left(\frac{3-4}{3} < \frac{X-4}{3} < \frac{7-4}{3}\right) = P\left(-\frac{1}{3} < Z < 1\right) = \Phi(1) - \Phi\left(-\frac{1}{3}\right) \\
&= \Phi(1) - [1 - \Phi\left(\frac{1}{3}\right)] = \Phi(1) + \Phi\left(\frac{1}{3}\right) - 1 \\
&= 0.8413 + 0.6293 - 1 = 0.4706\,.
\end{aligned}$$

Infine si ha:

$$\begin{aligned}
P(X > 7|X > 4) &= \frac{P(\{X > 7\}, \{X > 4\})}{P(X > 4)} = \frac{P(X > 7)}{0.5} = 2 \cdot P\left(\frac{X-4}{3} > \frac{7-4}{3}\right) \\
&= 2[1 - \Phi(1)] = 2[1 - 0.8413] = 0.3174
\end{aligned}$$

♣

Esercizio 5.7 Un'industria produce su commissione delle sbarre di acciaio cilindriche, il cui diametro deve essere di 4 cm, ma che tuttavia sono accettabili se hanno un diametro compreso fra 3.995 cm e 4.005 cm. Il cliente, nel controllare le sbarre fornitegli, constata che il 5% sono di diametro inferiore a quello tollerato ed il 12% sono di diametro superiore. Supponendo che le misure x dei diametri seguano una distribuzione normale, calcolarne i parametri μ e σ^2.

Determinare inoltre il valore di σ affinchè la probabilità che le sbarre abbiano un diametro superiore a quello tollerato sia inferiore al 2%.

Soluzione. Valutando la probabilità con la frequenza, in base ai dati del problema possiamo impostare il sistema:

$$\begin{aligned}
P\{X < 3.995\} = 0.05 \\
P\{X > 4.005\} = 0.12
\end{aligned}$$

cioè:

$$P\left(\frac{X-\mu}{\sigma} < \frac{3.995-\mu}{\sigma}\right) = \Phi\left(\frac{3.995-\mu}{\sigma}\right) = 0.05$$
$$P\left(\frac{X-\mu}{\sigma} > \frac{4.005-\mu}{\sigma}\right) = 1-\Phi\left(\frac{4.005-\mu}{\sigma}\right) = 0.12\ .$$

Da tali equazioni si ottiene:

$$\begin{array}{rcl} \dfrac{3.995-\mu}{\sigma} & = & z_{0.05} \\ \dfrac{4.005-\mu}{\sigma} & = & z_{0.88} \end{array}$$

dove $z_{0.05} = -1.65$ e $z_{0.88} = 1.18$ sono rispettivamente i quantili 0.05 e 0.88 della distribuzione normale standard, ottenuti dalle tavole della stessa distribuzione. Otteniamo pertanto il sistema:

$$\begin{cases} -1.65\sigma & + & \mu & = & 3.995 \\ 1.18\sigma & + & \mu & = & 4.005 \end{cases}$$

da cui si ottengono i valori: $\mu = 4.001$ e $\sigma = 0.0035$.

Per quanto riguarda l'ultimo punto, si richiede di calcolare il valore dello scarto σ tale che:

$$P\{X > 4.005\} < 0.02$$

cioè:

$$P\left(\frac{X-4.001}{\sigma} > \frac{4.005-4.001}{\sigma}\right) = 1-\Phi\left(\frac{0,004}{\sigma}\right) < 0.02$$

ovvero:

$$\Phi\left(\frac{0.004}{\sigma}\right) > 0.98 \qquad \text{da cui} \quad \frac{0.004}{\sigma} > z_{0.98}\ .$$

Dalle tavole della normale si ricava il quantile 0.98, cioè $z_{0.98} = 2.06$ da cui si ottiene infine $\sigma < 0.0019$. ♣

5.4 Altri esercizi

Esercizio 5.8 Siano X_1, X_2, X_3, X_4 quattro variabili aleatorie indipendenti ed identicamente distribuite con legge uniforme nell'intervallo $[0, 2]$. Considerata la variabile aleatoria

$$Z = \max\{X_1, X_2, X_3, X_4\}$$

calcolare la densità di probabilità ed il valore atteso della v.a. $Y = 2 - Z$.

Soluzione. Preliminarmente osserviamo che, per ipotesi, ciascuna delle variabili X_1, X_2, X_3, X_4 ha funzione di densità:

$$f(x) = \begin{cases} \dfrac{1}{2} & \text{per } 0 \leq x \leq 2 \\ \\ 0 & \text{altrimenti.} \end{cases}$$

Per calcolare la funzione di densità di Z, calcoliamo dapprima la funzione di ripartizione:

$$\begin{aligned} F_Z(z) &= P\{Z \leq z\} = P\{\max(X_1, X_2, X_3, X_4) \leq z\} \\ &= P\{X_1 \leq z, X_2 \leq z, X_3 \leq z, X_4 \leq z\} \end{aligned}$$

Per l'ipotesi di indipendenza delle v.a. e per l'ipotesi di legge uniforme in $[0, 2]$ si ha:

$$\begin{aligned} F_Z(z) &= P\{X_1 \leq z, X_2 \leq z, X_3 \leq z, X_4 \leq z\} \\ &= P\{X_1 \leq z\}P\{X_2 \leq z\}P\{X_3 \leq z\}P\{X_4 \leq z\} = \frac{z}{2} \cdot \frac{z}{2} \cdot \frac{z}{2} \cdot \frac{z}{2} = \frac{z^4}{16} \,. \end{aligned}$$

La funzione di ripartizione di Y è data da:

$$F_Y(y) = P\{Y \leq y\} = P\{2 - Z \leq y\} = P\{Z \geq 2 - y\} = 1 - F_Z(2 - y)$$

e si ha pertanto:

$$F_Y(y) = \begin{cases} 0 & \text{per } y \leq 0 \\ 1 - \left(\dfrac{2-y}{2}\right)^4 & \text{per } 0 \leq y \leq 2 \\ 1 & \text{per } y \geq 2\,. \end{cases}$$

Segue allora la funzione di densità di Y:

$$f_Y(y) = \begin{cases} 0 & \text{per } y \notin [0, 2] \\ \\ \dfrac{1}{4}(2-y)^3 & \text{per } 0 \leq y \leq 2\,. \end{cases}$$

Nota la funzione di densità, possiamo calcolare il valore atteso di Y:

$$\begin{aligned} \mathbb{E}[Y] &= \int_0^2 y\, f_Y(y)\, dy = \int_0^2 y \cdot \frac{1}{4}(2-y)^3\, dy = \frac{1}{4}\int_0^2 y(8 - 12y + 6y^2 - y^3)\, dy \\ &= \frac{1}{4}\int_0^2 (8y - 12y^2 + 6y^3 - y^4)\, dy = \frac{1}{4}\left[4y^2 - 4y^3 + \frac{3}{2}y^4 - \frac{1}{5}y^5\right]_0^2 \\ &= \frac{1}{4}\left(16 - 32 + 24 - \frac{32}{5}\right) = 0,4\,. \end{aligned}$$

♣

Esercizio 5.9 Un sistema S è composto da 4 componenti A, B, C e D. Il componente A è in serie con B, mentre C è in serie con D. Inoltre A e B sono in parallelo con C e D. I tempi (aleatori) fino al guasto dei quattro componenti sono rispettivamente X, Y, Z e W. Il vettore aleatorio (X, Y, Z, W) ha come densità la funzione:

$$f(x,y,z,w) = \begin{cases} \mathrm{e}^{-(x+y+z+w)} & \text{per } x \geq 0, y \geq 0, z \geq 0, w \geq 0 \\ 0 & \text{altrove.} \end{cases}$$

Calcolare:

1. la densità congiunta $g(x, y)$ del vettore (X, Y) per ogni $x \geq 0$, $y \geq 0$;
2. il valore atteso μ del tempo di guasto del sottosistema costituito dai componenti A e B in serie;
3. la probabilità τ che il sistema S non si guasti fino ad un determinato tempo t.

Soluzione. Osserviamo immediatamente che poichè:

$$\mathrm{e}^{-(x+y+z+w)} = \mathrm{e}^{-x}\mathrm{e}^{-y}\mathrm{e}^{-z}\mathrm{e}^{-w}$$

le v.a. X, Y, Z e W sono indipendenti e seguono una distribuzione esponenziale di parametro $\lambda = 1$. Si ha inoltre:

$$\begin{aligned} g(x,y) &= \int_0^{+\infty}\int_0^{+\infty} f(x,y,z,w)\, dz\mathrm{d}w = \int_0^{+\infty}\int_0^{+\infty} \mathrm{e}^{-(x+y+z+w)}\, dzdw \\ &= \mathrm{e}^{-(x+y)}\int_0^{+\infty}\int_0^{+\infty} \mathrm{e}^{-(z+w)}\, dzdw = \mathrm{e}^{-(x+y)} \qquad x \geq 0\, y \geq 0 \ . \end{aligned}$$

Indicato con T il tempo di guasto del sottosistema costituito dai componenti A e B, si ha:

$$T = \min(X, Y)$$

ed essendo X, Y v.a. indipendenti, si ha:

$$P(T > t) = P(\min(X,Y) > t) = P(X > t)P(Y > t) = \mathrm{e}^{-t} \cdot \mathrm{e}^{-t} = \mathrm{e}^{-2t}$$

e quindi $T \sim \mathcal{E}(2)$, da cui segue:

$$\mu = \mathbb{E}[T] = \frac{1}{2} \ .$$

Per lo stesso motivo, posto $V = \min(Z, W)$ segue $V \sim \mathcal{E}(2)$. Si ha quindi:

$$\begin{aligned} \alpha &= P[\{T > t\} \cup \{V > t\}] = P\{T > t\} + P\{V > t\} - P[\{T > t\}, \{V > t\}] \\ &= 2\mathrm{e}^{-2t} - \mathrm{e}^{-4t} \ . \end{aligned}$$

♣

Esercizio 5.10 Si considerino una variabile aleatoria avente distribuzione uniforme sull'insieme $\{-2, 2\}$ ed una v.a. Y, indipendente da X, avente distribuzione normale standard. Posto $Z = XY + 1$, determinare:

1. il coefficiente di correlazione ρ di X e Z;

2. la funzione di ripartizione $F(z)$;

3. il valore k tale che $P(Z > 2k + 1) = 0,1587$.

Soluzione. In base alle ipotesi, essendo X uniforme su $\{-2, 2\}$, si ha:

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

e quindi $\mathbb{E}[X] = 0$. Pertanto in base alle proprietà del coefficiente di correlazione, risulta:

$$\begin{aligned}\text{Cov}(X, Z) &= \text{Cov}(X, XY + 1) = \text{Cov}(X, XY) = \mathbb{E}[X^2Y] - \mathbb{E}[X]\,\mathbb{E}[XY] \\ &= \mathbb{E}[X^2Y]\ .\end{aligned}$$

Inoltre, poichè X e Y sono indipendenti, anche X^2 e Y sono indipendenti e quindi:

$$\text{Cov}(X, Z) = \mathbb{E}[X^2Y] = \mathbb{E}[X^2]\,\mathbb{E}[Y] = 0$$

in quanto Y segue una distribuzione normale standard e quindi $\mathbb{E}[Y] = 0$. Pertanto segue:

$$\rho = 0\ .$$

Per quanto concerne la funzione di ripartizione di Z, si ha:

$$\begin{aligned}F(z) &= P\{Z \le z\} = P\{XY + 1 \le z\} = P\{XY \le z - 1\} \\ &= P\{-2Y \le z - 1 | X = -2\}P\{X = -2\} + P\{2Y \le z - 1 | X = 2\}P\{X = 2\} \\ &= \frac{1}{2}\left(P\{-2Y \le z - 1 | X = -2\} + P\{2Y \le z - 1 | X = 2\}\right) \\ &= \frac{1}{2}\left(P\left\{Y \ge -\frac{z-1}{2}\right\} + P\left\{Y \le \frac{z-1}{2}\right\}\right) = P\left\{Y \le \frac{z-1}{2}\right\} = \Phi\left(\frac{z-1}{2}\right)\ .\end{aligned}$$

Pertanto $Z \sim N(1, 4)$ e quindi il valore k tale che

$$P(Z > 2k + 1) = 0,1587$$

è dato dal valore k tale che

$$1 - \Phi(k) = 0,1587$$

cioè $k = z_{0,8413} = 1$. ♣

Esercizio 5.11 Un impiegato abita all'estrema periferia della città e la strada più breve per raggiungere il suo ufficio passa per il centro, in questo caso il tempo richiesto per percorrere tragitto può essere descritto mediante una variabile aleatoria T_1 di media $\mu_1 = 28$ minuti e scarto $\sigma_1 = 7$ minuti e 20 secondi. Sono possibili altre due strade alternative il cui tragitto può essere descritto mediante rispettivamente una v.a. T_2 con media $\mu_2 = 32$ minuti e scarto $\sigma_2 = 4$ minuti e 30 secondi ed una v.a. T_3 con media $\mu_3 = 33$ minuti e scarto $\sigma_3 = 1$ minuto e 10 secondi. Assumendo che le tre v.a. abbiano tutte distribuzione normale, calcolare quale sia tragitto più opportuno da percorrere nel caso in cui: *i*.) l'impiegato debba essere in ufficio entro 30 minuti; *ii*.) l'impiegato debba essere in ufficio entro 35 minuti?

Soluzione. Preliminarmente, scriviamo gli scarti quadratici σ_i in termini di unità di misura omogenee: 30 secondi equivalgono a 0.5 minuti, 20 secondi equivalgono a 0.3333 minuti e 10 secondi equivale a 0,167 minuti, per cui in base alle ipotesi del problema si ha:

$$T_1 \sim N(28, 7.5^2) \qquad T_2 \sim N(32, 4.333^2) \qquad T_3 \sim N(33, 1.167^2)$$

Nel primo caso il tragitto da scegliere è quello per cui risulta maggiore la quantità $P(T_i \leq 30)$, con $i = 1, 2, 3$, mentre nel secondo caso si sceglierà il tragitto per cui risulta maggiore òln base alle ipotesi del problema si ha $P(T_i \leq 35)$, con $i = 1, 2, 3$. Valutiamo pertanto:

$$\begin{aligned}
P(T_1 \leq 30) &= \Phi\left(\frac{30-28}{7.5}\right) = \Phi(0.27) = 0.6064 \\
P(T_2 \leq 30) &= \Phi\left(\frac{30-32}{4.333}\right) = \Phi(-0.46) = 1 - \Phi(0.46) \ = \ 1 - 0.6772 = 0.3228 \\
P(T_1 \leq 30) &= \Phi\left(\frac{30-33}{1.167}\right) = \Phi(-2.57) = 1 - \Phi(2.57) = 1 - 0.9949 = 0.0051 \,.
\end{aligned}$$

In questo caso risulta opportuno percorrere il tragitto più breve. Nel secondo caso si ha:

$$\begin{aligned}
P(T_1 \leq 35) &= \Phi\left(\frac{35-28}{7.5}\right) = \Phi(0.93) = 0.8247 \\
P(T_2 \leq 35) &= \Phi\left(\frac{35-32}{4.333}\right) = \Phi(0.69) = 0.7485 \\
P(T_1 \leq 35) &= \Phi\left(\frac{35-33}{1.167}\right) = \Phi(1.71) = 1 - 0.9567 \,.
\end{aligned}$$

In questo caso è più opportuno percorrere il terzo tragitto. ♣

Esercizio 5.12 Sia (X, Y) una v.a. doppia a valori in $\{(x, y) \in \mathbb{R}^2 \ : \ 0 \leq x \leq 1 \,, 0 \leq y \leq x/3\}$ ed avente densità:

$$f(x, y) = \begin{cases} c(x - 3y) & \text{per } 0 \leq x \leq 1 \,, 0 \leq y \leq x/3 \\ 0 & \text{altrove.} \end{cases}$$

Calcolare il valore della costante c e le funzioni di densità $f_1(x)$ e $f(y|x)$.

Soluzione. Essendo la funzione $f(x, y)$ una funzione di densità, deve risultare:

$$c \int_0^1 dx \int_0^{x/3} (x - 3y)\, dy = 1 \ .$$

Calcoliamo la funzione di densità $f_1(x)$:

$$f_1(x) = \int_0^{x/3} (x - 3y) dy = c \left[xy - 3\frac{y^2}{2} \right]_0^{x/3} = c \left(\frac{x^2}{3} - 3\frac{x^2}{18} \right) = c\frac{x^2}{6} \ .$$

e quindi

$$c \int_0^1 dx \int_0^{x/3} (x - 3y)\, dy = c \int_0^1 \frac{x^2}{6} dx = c \left[\frac{x^3}{18} \right]_0^1 = c\frac{1}{18} \ ,$$

da cui $c = 18$ e quindi

$$f_1(x) = c\frac{x^2}{6} = 3x^2 \ .$$

Ed infine

$$f(y|x) = \frac{f(x, y)}{f_1(x)} = \frac{18(x - 3y)}{3x^2} = 6\frac{x - 3y}{x^2} \ .$$

♣

Esercizio 5.13 Sia X una v.a. continua avente funzione di distribuzione F_X e funzione di densità f_X. Dimostrare che la variabile aleatoria $Y = X^2$ è ancora continua ed esprimere la sua funzione di distribuzione F_Y e la sua funzione di densità f_Y in termini di F_X e f_X. Inoltre calcolare la funzione di densità di $Y = X^2$ nei casi in cui:

1. X segue una distribuzione normale $N(\mu, \sigma^2)$;
2. X segue una distribuzione di Laplace avente densità $f(x) = \frac{1}{2}\mathrm{e}^{-|x|}$;
3. X segue una distribuzione di Cauchy avente densità $f(x) = \frac{1}{\pi}\frac{1}{1+x^2}$.

Soluzione. La funzione di distribuzione di Y èp data da:

$$F_Y(y) = P(Y \leq y) = P(X^2 \leq y) = P(-\sqrt{y} \leq X \leq \sqrt{y}) = F_X(\sqrt{y}) - F_X(-\sqrt{y}) \ .$$

Poichè la funzione di distribuzione F_X è per ipotesi differenziabile (infatti esiste la funzione di densità f_X), allora anche F_Y è differenziabile e Y è una v.a. continua con funzione di densità

$$f_Y(y) = \frac{d}{dy}F_Y(y) = \frac{d}{dy}F_X(\sqrt{y}) - \frac{d}{dy}F_X(-\sqrt{y}) = \frac{1}{2\sqrt{y}}[f_X(\sqrt{y}) + f_X(-\sqrt{y})] \ .$$

Applicando la formula sopra ricavata, nei tre casi richiesti si ricava:

1. $f_Y(y) = \dfrac{1}{2\sigma\sqrt{2\pi y}} \exp\left[-\dfrac{y + \mu^2}{\sigma^2} \right]$ per $y > 0$;

2. $f_Y(y) = \dfrac{1}{2\sqrt{y}} \mathrm{e}^{-\sqrt{y}}$ per $y > 0$;

3. $\dfrac{1}{\pi} \dfrac{1}{\sqrt{y}(1+y)}$ per $y > 0$.

♣

Esercizio 5.14 Siano X_1 e X_2 due v.a. indipendenti aventi distribuzione normale con medie rispettivamente $\mu_1 = 10$, $\mu_2 = 4$ ed aventi la stessa varianza $\sigma^2 = 9$.

a) $P(X_1 + X_2 > 16)$;

b) $P(2X_1 > 16)$;

c) $P(7 < X_1 - X_2 < 10)$;

d) $P(X_2 - X_1 > -5)$.

Soluzione. Poichè per ipotesi $X_1 \sim N(10, 9)$ e $X_2 \sim N(4, 9)$ sono v.a. normali indipendenti, allora posto $U = X_1 + X_2$ si ha $U \sim N(14, 18)$ e quindi segue:

$$\begin{aligned} P(U > 16) &= P\left(\frac{U - 14}{\sqrt{18}} > \frac{16 - 14}{\sqrt{18}}\right) = P(Z > 0.4714) = 1 - \Phi(0.47) \\ &= 1 - 0.6813 = 0.3187 \,, \end{aligned}$$

dove Z denota la v.a. normale standard. Per quanto riguarda il punto successivo, si ha:

$$\begin{aligned} P(2X_1 > 16) &= P(X_1 > 8) = P\left(\frac{X_1 - 10}{3} > \frac{8 - 10}{3}\right) = P(Z > -0.6667) \\ &= 1 - \Phi(-0.6667) = 1 - 1 + \Phi(0.6667) = 0.7475 \,. \end{aligned}$$

Per quanto riguarda il terzo punto, posto $V = X_1 - X_2$, si ha $V \sim N(6, 18)$ e quindi

$$\begin{aligned} P(7 < V < 10) &= P\left(\frac{7 - 6}{\sqrt{18}} < \frac{V - 6}{\sqrt{18}} < \frac{10 - 6}{\sqrt{18}}\right) = P(0.2357 < Z < 0.9428) \\ &= \Phi(0.9428) - \Phi(0.2357) = 0.8271 - 0.5932 = 0.2339 \,. \end{aligned}$$

Infine, posto $W = X_2 - X_1$, si ha $W \sim N(-6, 18)$ e quindi:

$$P(W > -5) = P\left(\frac{W + 6}{\sqrt{18}} > \frac{-5 + 6}{\sqrt{18}}\right) = P(Z > 0.2357) = 0.5932 \,.$$

♣

Esercizio 5.15 Si consideri un negozio in cui arrivano, a caso, mediamente 20 clienti l'ora.

1. Calcolare la probabilità che gli intervalli di tempo fra due arrivi consecutivi siano inferiori a tre minuti;
2. Calcolare la probabilità che gli intervalli di tempo fra due arrivi consecutivi siano superiori a quattro minuti;
3. Supponendo che il 10% dei clienti compri un certo oggetto, ricavare la distribuzione di probabilità del numero di cliento che acquistano un oggetto in un'ora.

Soluzione. Sia T l'intervallo di tempo (in minuti) che intercorre fra l'arrivo di due clienti consecutivi. Allora T segue una distribuzione esponenziale di parametro 3, ed ha quindi densità:

$$f(t) = \begin{cases} \frac{1}{3}\mathrm{e}^{-t/3} & \text{per } t > 0 \\ 0 & \text{altrove.} \end{cases}$$

Indicata con $F(t)$ la funzione di ripartizione, segue pertanto:

1. $P(T < 3) = F(3) = 1 - \mathrm{e}^{-1} \approx 0,37$.
2. $P(T > 4) = 1 - P(T \leq 4) = 1 - F(4) = \mathrm{e}^{-4/3} \approx 0,26$.

Per quanto riguarda il terzo punto, indichiamo con X il numero di clienti che entrano nel negozio in un 'ora e con Y il numero di quelli, fra di essi, che effettuano acquisti. La variabile X segue allora una distribuzone di Poisson di parametro 20:

$$p(x) = \begin{cases} \mathrm{e}^{-20}\dfrac{20^x}{x!} & \text{per } x = 0, 1, 2 \ldots \\ 0 & \text{altrove.} \end{cases}$$

In base alle ipotesi del problema, si ha poi che la variabile Y, dato $X = n$, segue una distribuzione binomiale di parametri n e $p = 0,1$:

$$p(y|X = n) = \begin{cases} \binom{n}{y}(0,1)^y(0,9)^{n-y} & \text{per } y = 0, 1, 2 \ldots, n \\ 0 & \text{altrove.} \end{cases}$$

Si ha pertanto:

$$\begin{aligned} P(Y = k) &= \sum_{n=k}^{\infty} P(Y = k|X = n)P(X = n) = \sum_{n=k}^{\infty} \binom{n}{k}(0,1)^k(0,9)^{n-k}\mathrm{e}^{-20}\frac{20^n}{n!} \\ &= \sum_{n=k}^{\infty} \frac{n!}{k!(n-k)!}(0,1)^k(0,9)^{-k}(0,9)^n\mathrm{e}^{-20}\frac{20^n}{n!} = \frac{1}{k!}\left(\frac{0,1}{0,9}\right)^k \mathrm{e}^{-20}\sum_{n=k}^{\infty}\frac{18^n}{(n-k)!} \\ &= \frac{1}{k!}\left(\frac{1}{9}\right)^k \mathrm{e}^{-20}\sum_{n=k}^{\infty} 18^k\frac{18^{n-k}}{(n-k)!} = \mathrm{e}^{-2}\frac{2^k}{k!} \end{aligned}$$

tenendo conto che

$$\sum_{z=0}^{\infty} \frac{18^z}{z!} = \mathrm{e}^{18} .$$

Pertanto, la variabile Y segue una distribuzione di Poisson di parametro 2. ♣

Esercizio 5.16 Siano $f_1(x)$ e $f_2(x)$ due funzioni di densità esponenziali di parametro rispettivamente $\lambda_1 = 1$ e $\lambda_2 = 2$. Si consideri la funzione $f(x)$ definita come segue:

$$f(x) \;=\; \alpha f_1(x) + (1-\alpha) f_2(x) \qquad\qquad 0 < \alpha < 1 .$$

1. Dimostrare che la funzione $f(x)$ è anch'essa una funzione di densità.
2. Calcolare la speranza matematica di $f(x)$.
3. Denotata con X una v.a. avente legge $f(x)$, calcolare il valore della costante α tale che $P(X > 2) = 0,1$.

Soluzione. In base alle ipotesi del problema, le funzioni $f_1(x)$ e $f_2(x)$ sono funzioni di densità esponenziali di parametri $\lambda_1 = 1$ e $\lambda_2 = 2$:

$$f_1(x) = \begin{cases} \mathrm{e}^{-x} & \text{per } x > 0 \\ 0 & \text{altrove.} \end{cases} \qquad\qquad f_2(x) = \begin{cases} 2\mathrm{e}^{-2x} & \text{per } x > 0 \\ 0 & \text{altrove.} \end{cases}$$

Ne segue che $f(x)$ è definita su tutto l'asse reale essendo una combinazione lineare di $f_1(x)$ e $f_2(x)$. Per quanto riguarda il primo punto bisogna verificare che $f(x)$ soddisfa le due ipotesi richieste per la funzione di densità:

1. $f(x) \geq 0$ per ogni $x \in R$,
2. $\int_R f(x)\, dx = 1$.

La verifica della prima condizione è immediata in quanto $f(x)$ è una combinazione lineare di funzioni di densità rispetto ai pesi α e $(1-\alpha)$ che sono entrambi positivi. Anche la seconda condizione è verificata poichè:

$$\begin{aligned} \int_{\mathbb{R}} f(x)\, dx &= \int_{\mathbb{R}} [\alpha f_1(x) + (1-\alpha) f_2(x)]\, dx = \alpha \int_{\mathbb{R}} f_1(x) dx + (1-\alpha) \int_{\mathbb{R}} f_2(x)\, dx \\ &= \alpha + (1-\alpha) = 1 . \end{aligned}$$

La speranza matematica di X si ottiene da un'applicazione della proprietà di linearità, ricordando che la speranza matematica di una v.a. esponenziale di parametro λ è uguale a $\frac{1}{\lambda}$

$$\begin{aligned} \mathbb{E}[X] &= \int_{\mathbb{R}} x[\alpha f_1(x) + (1-\alpha) f_2(x)]\, dx = \alpha \int_{\mathbb{R}} x f_1(x) dx + (1-\alpha) \int_{R} x f_2(x)\, dx \\ &= \alpha + \frac{1-\alpha}{2} = \frac{\alpha+1}{2} . \end{aligned}$$

Per quanto riguarda l'ultimo punto, calcoliamo dapprima $P(X > 2)$. Analogamente a quanto visto in precedenza, ricordando che la funzione di sopravvivenza di una densità esponenziale di parametro λ è $S(x) = \mathrm{e}^{-\lambda x}$, si ha:

$$\begin{aligned} P(X > 2) &= \int_2^{+\infty} f(x)dx = \int_2^{+\infty} [\alpha f_1(x) + (1-\alpha) f_2(x)]\, dx \\ &= \alpha \int_2^{+\infty} f_1(x)dx + (1-\alpha) \int_2^{+\infty} f_2(x)\, dx = \alpha \mathrm{e}^{-2} + (1-\alpha)\mathrm{e}^{-4} \ . \end{aligned}$$

Pertanto il valore di α tale che $P(X > 2) = 0,1$ si ricava come soluzione dell'equazione:

$$\alpha \mathrm{e}^{-2} + (1-\alpha)\mathrm{e}^{-4} = 0,1 \ ,$$

da cui

$$\alpha = \frac{0,1 - \mathrm{e}^{-4}}{\mathrm{e}^{-2} - \mathrm{e}^{-4}} \approx 0,6980 \ .$$

♣

Esercizio 5.17 Sia X la vita in ore di un certo tipo di lampadine. Da osservazioni passate, è noto che X segue una distribuzione normale $N(200, \sigma^2)$. Se un acquirente vuole che almeno il 95% di una partita di tali lampadine abbia una vita superiore alle 150 ore, calcolare quale può essere la massima varianza σ^2 della popolazione di lampadine che il compratore sia disposto ad accettare.

Soluzione. Dai dati del problema si assume che $X \sim N(200, \sigma^2)$. Si richiede di valutare la varianza σ^2 in modo tale che

$$P(X > 150) \geq 0,95 \ ,$$

in particolare, la varianza massima che l'acquirente è disposto a tollerare è quella per cui

$$P(X > 150) = 0,95 \ . \tag{5.8}$$

Denotiamo al solito con $Z \sim N(0,1)$ la variabile aleatoria normale standard, allora la (5.8) è equivalente a

$$\begin{aligned} P(X > 150) &= P\left(\frac{X-\mu}{\sigma} > \frac{150-200}{\sigma}\right) \\ &= P(Z > -\frac{50}{\sigma}) = 0,95 \ . \end{aligned}$$

Indicato con z_α il quantile α di una distribuzione normale standard, cioè il quantile per cui risulta $P(Z < z_\alpha) = \alpha$, in particolare dalle tavole dei valori delle aree della distribuzione normale standard si ha che $z_{0,05} = -1,645$. Il valore di σ richiesto si ottiene pertanto da:

$$z_{0,05} = -\frac{50}{\sigma} \quad = \quad -1,645$$

da cui segue $\sigma = 30,39$ e quindi $\sigma^2 = 923,86$. □

♣

Esercizio 5.18 Determinare i parametri della distribuzione di una variabile aleatoria $X \sim N(\mu, \sigma^2)$ sapendo che il primo quartile di X è uguale a 40 ed il terzo quartile è uguale a 60. Successivamente calcolare $P(X < 55|X > 40)$.

Soluzione. Il problema richiede la determinazione dei parametri di una distribuzione normale $N(\mu, \sigma^2)$, noti i valori del primo e del terzo quartile, rispettivamente $Q_1(X)$ e $Q_3(X)$. Si tratta di impostare un sistema di due equazioni nelle due incognite μ e σ in dipendenza dei parametri noti $Q_1(X)$ e $Q_3(X)$:

$$\begin{cases} \mu & = \dfrac{Q_1(X) + Q_3(X)}{2} \\ \dfrac{Q_3(X) - \mu}{\sigma} & = Q_3(Z) \end{cases} \tag{5.9}$$

dove $Q_3(Z)$ denota il terzo quartile di una variabile normale standard $Z \sim N(0, 1)$. In tale sistema la prima equazione si ottiene considerando che la distribuzione normale è simmetrica rispetto alla media e pertanto questa è data dalla semisomma del primo e terzo quantile; la seconda equazione esprime la relazione fra una variabile $X \sim N(\mu, \sigma^2)$ e la corrispondente variabile standard $Z \sim N(0, 1)$. Indicato con z_α il quantile α di una distribuzione normale standard, cioè il quantile per cui si risulta $P(Z < z_\alpha) = \alpha$, in base alle tavole si ottiene:

$$Q_3(Z) \;=\; z_{0.75} = 0.6745\ .$$

inoltre in base ai dati del problema si ha $Q_1(X) = 40$ e $Q_3(X) = 60$, pertanto il sistema (5.9) si scrive

$$\begin{cases} \mu & = \dfrac{40 + 60}{2} \\ \dfrac{60 - \mu}{\sigma} & = 0.6745 \end{cases}$$

da cui si ottiene la soluzione

$$\begin{aligned} \mu &= 50 \\ \sigma &= 14,83\ . \end{aligned}$$

Si ha poi:

$$P(X < 55|X > 40) \;=\; \frac{P(X < 55, X > 40)}{P(X > 40} = \frac{P(40 < X < 55)}{P(X > 40)}\ .$$

Poichè

$$\begin{aligned} P(40 < X < 55) &= P\left(\frac{40 - 50}{14,83} \le Z \le \frac{55 - 50}{14,83}\right) \\ &= \Phi(0,3372) - \Phi(-0,6743) \;=\; \Phi(0,3372) - 1 + \Phi(0,6742) \\ &= 0,6331 - 1 + 0,7486 \;=\; 0,3817\ , \end{aligned}$$

segue:

$$P(X < 55|X > 40) = \frac{P(40 < X < 55)}{P(X > 40)} = \frac{0,3817}{0,75} = 0,5089\ .$$

♣

Tavola dei quantili della distribuzione normale standard: $\Phi(z) = P(Z \leq z)$

	0. 00	0. 01	0. 02	0. 03	0. 04	0. 05	0. 06	0. 07	0. 08	0. 09
0. 0	0. 5000	0. 5040	0. 5080	0. 5120	0. 5160	0. 5199	0. 5239	0. 5279	0. 5319	0. 5359
0. 1	0. 5398	0. 5438	0. 5478	0. 5517	0. 5557	0. 5596	0. 5636	0. 5675	0. 5714	0. 5753
0. 2	0. 5793	0. 5832	0. 5871	0. 5910	0. 5948	0. 5987	0. 6026	0. 6064	0. 6103	0. 6141
0. 3	0. 6179	0. 6217	0. 6255	0. 6293	0. 6331	0. 6368	0. 6406	0. 6443	0. 6480	0. 6517
0. 4	0. 6554	0. 6591	0. 6628	0. 6664	0. 6700	0. 6736	0. 6772	0. 6808	0. 6844	0. 6879
0. 5	0. 6915	0. 6950	0. 6985	0. 7019	0. 7054	0. 7088	0. 7123	0. 7157	0. 7190	0. 7224
0. 6	0. 7257	0. 7291	0. 7324	0. 7357	0. 7389	0. 7422	0. 7454	0. 7486	0. 7517	0. 7549
0. 7	0. 7580	0. 7611	0. 7642	0. 7673	0. 7704	0. 7734	0. 7764	0. 7794	0. 7823	0. 7852
0. 8	0. 7881	0. 7910	0. 7939	0. 7967	0. 7995	0. 8023	0. 8051	0. 8078	0. 8106	0. 8133
0. 9	0. 8159	0. 8186	0. 8212	0. 8238	0. 8264	0. 8289	0. 8315	0. 8340	0. 8365	0. 8389
1. 0	0. 8413	0. 8438	0. 8461	0. 8485	0. 8508	0. 8531	0. 8554	0. 8577	0. 8599	0. 8621
1. 1	0. 8643	0. 8665	0. 8686	0. 8708	0. 8729	0. 8749	0. 8770	0. 8790	0. 8810	0. 8830
1. 2	0. 8849	0. 8869	0. 8888	0. 8907	0. 8925	0. 8944	0. 8962	0. 8980	0. 8997	0. 9015
1. 3	0. 9032	0. 9049	0. 9066	0. 9082	0. 9099	0. 9115	0. 9131	0. 9147	0. 9162	0. 9177
1. 4	0. 9192	0. 9207	0. 9222	0. 9236	0. 9251	0. 9265	0. 9279	0. 9292	0. 9306	0. 9319
1. 5	0. 9332	0. 9345	0. 9357	0. 9370	0. 9382	0. 9394	0. 9406	0. 9418	0. 9429	0. 9441
1. 6	0. 9452	0. 9463	0. 9474	0. 9484	0. 9495	0. 9505	0. 9515	0. 9525	0. 9535	0. 9545
1. 7	0. 9554	0. 9564	0. 9573	0. 9582	0. 9591	0. 9599	0. 9608	0. 9616	0. 9625	0. 9633
1. 8	0. 9641	0. 9649	0. 9656	0. 9664	0. 9671	0. 9678	0. 9686	0. 9693	0. 9699	0. 9706
1. 9	0. 9713	0. 9719	0. 9726	0. 9732	0. 9738	0. 9744	0. 9750	0. 9756	0. 9761	0. 9767
2. 0	0. 9772	0. 9778	0. 9783	0. 9788	0. 9793	0. 9798	0. 9803	0. 9808	0. 9812	0. 9817
2. 1	0. 9821	0. 9826	0. 9830	0. 9834	0. 9838	0. 9842	0. 9846	0. 9850	0. 9854	0. 9857
2. 2	0. 9861	0. 9864	0. 9868	0. 9871	0. 9875	0. 9878	0. 9881	0. 9884	0. 9887	0. 9890
2. 3	0. 9893	0. 9896	0. 9898	0. 9901	0. 9904	0. 9906	0. 9909	0. 9911	0. 9913	0. 9916
2. 4	0. 9918	0. 9920	0. 9922	0. 9925	0. 9927	0. 9929	0. 9931	0. 9932	0. 9934	0. 9936
2. 5	0. 9938	0. 9940	0. 9941	0. 9943	0. 9945	0. 9946	0. 9948	0. 9949	0. 9951	0. 9952
2. 6	0. 9953	0. 9955	0. 9956	0. 9957	0. 9959	0. 9960	0. 9961	0. 9962	0. 9963	0. 9964
2. 7	0. 9965	0. 9966	0. 9967	0. 9968	0. 9969	0. 9970	0. 9971	0. 9972	0. 9973	0. 9974
2. 8	0. 9974	0. 9975	0. 9976	0. 9977	0. 9977	0. 9978	0. 9979	0. 9979	0. 9980	0. 9981
2. 9	0. 9981	0. 9982	0. 9982	0. 9983	0. 9984	0. 9984	0. 9985	0. 9985	0. 9986	0. 9986
3. 0	0. 9987	0. 9987	0. 9987	0. 9988	0. 9988	0. 9989	0. 9989	0. 9989	0. 9990	0. 9990
3. 1	0. 9990	0. 9991	0. 9991	0. 9991	0. 9992	0. 9992	0. 9992	0. 9992	0. 9993	0. 9993
3. 2	0. 9993	0. 9993	0. 9994	0. 9994	0. 9994	0. 9994	0. 9994	0. 9995	0. 9995	0. 9995
3. 3	0. 9995	0. 9995	0. 9995	0. 9996	0. 9996	0. 9996	0. 9996	0. 9996	0. 9996	0. 9997
3. 4	0. 9997	0. 9997	0. 9997	0. 9997	0. 9997	0. 9997	0. 9997	0. 9997	0. 9997	0. 9998
3. 5	0. 9998	0. 9998	0. 9998	0. 9998	0. 9998	0. 9998	0. 9998	0. 9998	0. 9998	0. 9998
3. 6	0. 9998	0. 9998	0. 9999	0. 9999	0. 9999	0. 9999	0. 9999	0. 9999	0. 9999	0. 9999
3. 7	0. 9999	0. 9999	0. 9999	0. 9999	0. 9999	0. 9999	0. 9999	0. 9999	0. 9999	0. 9999
3. 8	0. 9999	0. 9999	0. 9999	0. 9999	0. 9999	0. 9999	0. 9999	0. 9999	0. 9999	0. 9999