CAPITOLO 5 – Variabili aleatorie continue

5.1 Introduzione

Le variabili aleatorie discrete assumono un numero finito o un'infinità numerabile di valori.

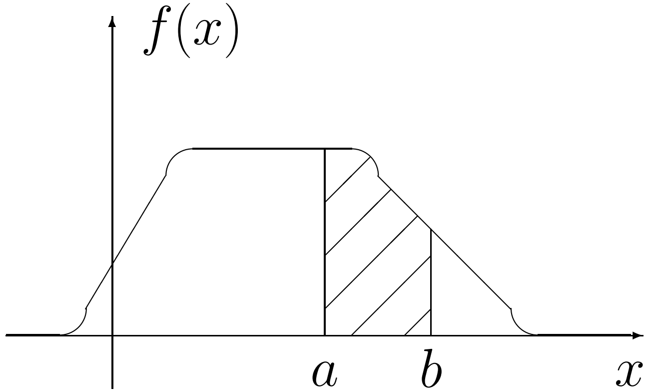
Esistono comunque variabili aleatorie il cui insieme dei valori è non numerabile, come ad esempio l'ora di arrivo di un treno o il tempo di vita di un dispositivo elettronico.

Definizione. Una variabile aleatoria X è detta continua (o, anche, assolutamente continua) se esiste una funzione $f\colon \mathbb{R} \to [0, +\infty)$ tale che per ogni sottoinsieme B di numeri reali risulta

$$P(X \in B) = \int_B f(x)\, dx.$$

La funzione f è chiamata *funzione di densità* della variabile aleatoria X, ed è di fatto caratterizzata dalle seguenti proprietà:

$$f(x) \geq 0 \quad \text{per } x \in \mathbb{R}, \qquad \int_{-\infty}^{\infty} f(x)\, dx = 1.$$



$$P(a \leq X \leq b) = \text{ area della regione tratteggiata } = \int_a^b f(x)\,dx.$$

Tale relazione ricalca la seguente, già vista in passato, che sussiste se X è una variabile aleatoria discreta:

$$P(a \leq X \leq b) = \sum_{k:\, a \leq x_k \leq b} p(x_k).$$

Abbiamo visto che se X è una variabile aleatoria continua, si ha

$$P(a \leq X \leq b) = \int_a^b f(x)\,dx.$$

Quindi, se $b = a$ si ricava

$$P(X = a) = \int_a^a f(x)\,dx = 0.$$

Pertanto, a differenza del caso discreto, la probabilità che una variabile aleatoria continua X assuma un singolo valore è uguale a zero, cosicché risulta

$$P(X < a) = P(X \leq a) = F(a) = \int_{-\infty}^a f(x)\,dx$$

e quindi

$$F(b) - F(a) = P(a < X \leq b) = \int_a^b f(x)\,dx.$$

Ne segue che la funzione di distribuzione $F(x)$ di una variabile aleatoria X continua è una funzione continua per ogni x reale.

Esempio. Sia X una variabile aleatoria continua di densità

$$f(x) = \begin{cases} c(4x - 2x^2) & 0 < x < 2 \\ 0 & \text{altrimenti.} \end{cases}$$

(a) Calcolare c.

(b) Determinare $P(X > 1)$.

Soluzione. (a) Per determinare c, imponendo che sia $\displaystyle\int_{-\infty}^{\infty} f(x)\,dx = 1$ si ha

$$1 = c\int_0^2 (4x - 2x^2)\,dx = c\left[2x^2 - \frac{2x^3}{3}\right]_{x=0}^{x=2} = c\left[8 - \frac{16}{3}\right] = c\,\frac{8}{3}$$

da cui segue $c = 3/8$.

(b) Risulta quindi

$$P(X > 1) = \int_1^{\infty} f(x)\,dx = \frac{3}{8}\int_1^2 (4x - 2x^2)\,dx = \frac{3}{8}\left[2x^2 - \frac{2x^3}{3}\right]_{x=1}^{x=2} = 1 - \frac{1}{2} = \frac{1}{2}.$$

Esempio. Il tempo (in ore) che un computer funzioni prima di bloccarsi è una variabile aleatoria continua X di densità

$$f(x) = \begin{cases} \lambda e^{-x/100} & x \geq 0 \\ 0 & x < 0. \end{cases}$$

(a) Calcolare la probabilità che il computer funzioni tra 50 e 150 ore senza bloccarsi.

(b) Calcolare la probabilità che il computer funzioni per meno di 100 ore senza bloccarsi.

Soluzione. (a) Per determinare λ, imponendo che risulti $\displaystyle\int_{-\infty}^{\infty} f(x)\,dx = 1$ otteniamo

$$1 = \lambda \int_0^{\infty} e^{-x/100}\,dx = \lambda \Big[-100\, e^{-x/100}\Big]_{x=0}^{x=\infty} = \lambda\, 100 \qquad \Longrightarrow \qquad \lambda = \frac{1}{100}$$

e pertanto

$$P(50 < X < 150) = \int_{50}^{150} \frac{1}{100}\, e^{-x/100}\,dx = \Big[-e^{-x/100}\Big]_{50}^{150} = e^{-1/2} - e^{-3/2} \approx 0{,}383.$$

(b) Analogamente si ha

$$P(X < 100) = \int_0^{100} \frac{1}{100}\, e^{-x/100}\,dx = \Big[-e^{-x/100}\Big]_0^{100} = 1 - e^{-1} \approx 0{,}632.$$

Esempio. Il tempo (in ore) di vita di certe pile per la radio è una variabile aleatoria continua X di densità

$$f(x) = \begin{cases} 0 & x \leq 100 \\ 100/x^2 & x > 100. \end{cases}$$

Determinare la probabilità che esattamente 2 pile della radio su 5 debbano essere sostituite entro le 150 ore di attività, supponendo che gli eventi $E_i = \{$l'i-esima pila va rimpiazzata entro 150 ore d'uso$\}$, $1 \leq i \leq 5$, siano indipendenti.

Soluzione. Risulta

$$P(E_i) = \int_0^{150} f(x)\,dx = 100 \int_{100}^{150} \frac{1}{x^2}\,dx = 100 \left[-\frac{1}{x}\right]_{x=100}^{x=150} = \frac{1}{3}.$$

Quindi, per l'indipendenza degli eventi E_i, la probabilità richiesta è

$$\binom{5}{2} \left(\frac{1}{3}\right)^2 \left(\frac{2}{3}\right)^3 = 10 \cdot \frac{1}{9} \cdot \frac{8}{27} = \frac{80}{243} \approx 0{,}329.$$

La relazione tra la funzione di distribuzione F di una variabile aleatoria continua e la densità f è data da

$$F(x) = P\{X \in (-\infty, x]\} = \int_{-\infty}^{x} f(t)\, dt.$$

Quindi, se x è un reale in cui F è derivabile, derivando ambo i membri si ottiene

$$\frac{d}{dx} F(x) = f(x),$$

cosicché la densità è la derivata della funzione di distribuzione.

Un'interpretazione intuitiva della densità segue da:

$$P\left\{x - \frac{\varepsilon}{2} \leq X \leq x + \frac{\varepsilon}{2}\right\} = \int_{x-\varepsilon/2}^{x+\varepsilon/2} f(x)\, dx \approx \varepsilon\, f(x),$$

dove l'approssimazione vale quando ε è prossimo a 0 e se la densità f è continua in x. Pertanto, la probabilità che X assuma valori in un intorno di x avente ampiezza ε è approssimativamente uguale a $\varepsilon\, f(x)$.

Esempio. Sia X una variabile aleatoria continua con funzione di distribuzione F_X e densità f_X. Determinare la densità di $Y = aX + b$, con $a \neq 0$.

Soluzione. Se $a > 0$, la funzione di distribuzione di Y è così esprimibile:

$$F_Y(x) = P(Y \leq x) = P(aX + b \leq x) = P\left(X \leq \frac{x-b}{a}\right) = F_X\left(\frac{x-b}{a}\right),$$

mentre per $a < 0$ risulta:

$$F_Y(x) = P\left(X \geq \frac{x-b}{a}\right) = 1 - F_X\left(\frac{x-b}{a}\right).$$

Derivando rispetto ad x si ottiene infine:

$$f_Y(x) = \begin{cases} \dfrac{1}{a} f_X\left(\dfrac{x-b}{a}\right) & \text{se } a > 0 \\ -\dfrac{1}{a} f_X\left(\dfrac{x-b}{a}\right) & \text{se } a < 0 \end{cases} = \frac{1}{|a|} f_X\left(\frac{x-b}{a}\right).$$

Esempio. Sia X una variabile aleatoria continua con densità

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

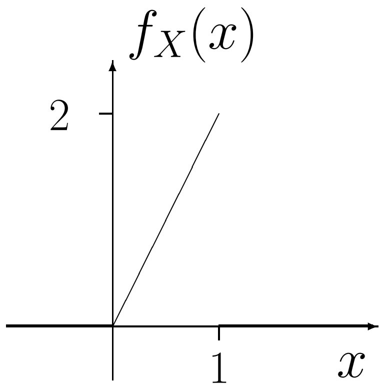
Determinare la densità di $Y = 1 - X$.

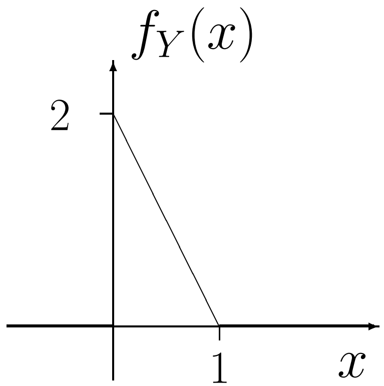
Soluzione. Ricordando che la densità di $aX + b$ è data da

$$f_{aX+b}(x) = \frac{1}{|a|} f_X\left(\frac{x-b}{a}\right),$$

per $a = -1$ e $b = 1$ segue che la densità di $Y = 1 - X$ è:

$$f_Y(x) = f_X(1-x) = \begin{cases} 2(1-x) & 0 \leq x \leq 1 \\ 0 & \text{altrimenti.} \end{cases}$$





Alternativamente, è possibile ricavare la densità di Y notando che la funzione di distribuzione di $Y = 1 - X$ è data da

$$F_Y(x) = P(Y \leq x) = P(1 - X \leq x) = P(X \geq 1 - x) = 1 - P(X < 1 - x).$$

Poiché X è una variabile aleatoria continua, risulta

$$P(X < 1 - x) = P(X \leq 1 - x) = F_X(1 - x).$$

Pertanto la densità di probabilità di Y si può ricavare come segue:

$$f_Y(x) = \frac{d}{dx} F_Y(x) = \frac{d}{dx} [1 - F_X(1 - x)] = f_X(1 - x).$$

Usando tale indentità e ricordando che la densità di probabilità di X è

$$f_X(x) = \begin{cases} 2x & 0 \leq x \leq 1 \\ 0 & \text{altrimenti} \end{cases}$$

si ottiene infine la densità di Y:

$$f_Y(x) = \begin{cases} 2(1 - x) & 0 \leq x \leq 1 \\ 0 & \text{altrimenti.} \end{cases}$$

5.2 Valore atteso e varianza di una variabile aleatoria continua

Come visto in precedenza, il valore atteso di una variabile aleatoria discreta X è

$$E[X] = \sum_x x\, P(X = x).$$

Analogamente, se X è una variabile aleatoria continua con densità $f(x)$, poiché

$$f(x)\, dx \approx P(x \leq X \leq x + dx) \qquad \text{per } dx \text{ piccolo},$$

si definisce il valore atteso di X come

$$E[X] = \int_{-\infty}^{\infty} x\, f(x)\, dx.$$

Esempio. Determinare $E[X]$ sapendo che la densità di X è

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

Soluzione. Si ha

$$E[X] = \int_{-\infty}^{\infty} x\, f(x)\, dx = \int_0^1 2x^2\, dx = \left[\frac{2x^3}{3}\right]_0^1 = \frac{2}{3}.$$

Esempio. Determinare $E\left[e^X\right]$ sapendo che la densità di X è

$$f_X(x)=\begin{cases}1 & 0\le x\le 1\\ 0 & \text{altrimenti.}\end{cases}$$

Soluzione. Posto $Y=e^X$, determiniamo innanzitutto la funzione di distribuzione e la densità di Y. Per $1\le x\le e$ risulta

$$F_Y(x)=P(Y\le x)=P\left(e^X\le x\right)=P(X\le\log x)=\int_0^{\log x} f_X(y)\,dy=\log x.$$

Derivando $F_Y(x)$ si ha

$$f_Y(x)=\frac{d}{dx}F_Y(x)=\frac{d}{dx}\log x=\frac{1}{x},\qquad 1\le x\le e,$$

da cui segue

$$E\left[e^X\right]=E[Y]=\int_{-\infty}^{\infty}x\,f_Y(x)\,dx=\int_1^e dx=e-1=1{,}71828.$$

Se X è una variabile aleatoria discreta, sappiamo che

$$E[g(X)] = \sum_i g(x_i)\, p(x_i).$$

È possibile dimostrare il seguente analogo risultato:

Proposizione. Se X è una variabile aleatoria continua con densità $f(x)$, allora per ogni funzione g a valori reali,

$$E[g(X)] = \int_{-\infty}^{\infty} g(x)\, f(x)\, dx.$$

Ad esempio, applicando tale Proposizione alla variabile aleatoria continua di densità

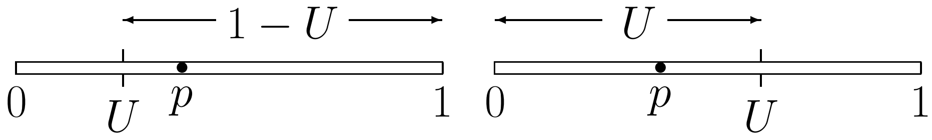
$$f_X(x) = \begin{cases} 1 & 0 \le x \le 1 \\ 0 & \text{altrimenti} \end{cases}$$

si ottiene

$$E\left[e^X\right] = \int_{-\infty}^{\infty} e^x\, f_X(x)\, dx = \int_0^1 e^x\, dx = [e^x]_0^1 = e - 1 = 1{,}71828.$$

Esempio. Un bastoncino di lunghezza 1 è spezzato in un punto U scelto a caso su di esso, e quindi distribuito uniformemente su $(0,1)$. Determinare il valore atteso della lunghezza del pezzo di bastoncino che contiene il punto p, $0 \le p \le 1$.

Soluzione. Sia $L(U)$ la lunghezza del pezzo di bastoncino contenente p. Si ha



$$L(U) = \begin{cases} 1-U & \text{se } 0 < U < p \\ U & \text{se } p < U < 1. \end{cases}$$

Quindi, essendo $f_U(x) = 1$, $0 \le x \le 1$, segue che

$$\begin{aligned} E[L(U)] &= \int_{-\infty}^{\infty} L(x)\, f_U(x)\, dx = \int_0^p (1-x)\, dx + \int_p^1 x\, dx \\ &= \left[x - \frac{x^2}{2}\right]_0^p + \left[\frac{x^2}{2}\right]_p^1 = p - \frac{p^2}{2} + \frac{1}{2} - \frac{p^2}{2} = \frac{1}{2} + p(1-p). \end{aligned}$$

Il valore atteso della lunghezza del pezzo di bastoncino contente p è massimo se p è il punto centrale, ossia se $p = 1/2$. Notiamo inoltre che $E[L(U)] \ge E[U] = 1/2$.

Corollario. Come nel caso discreto, se X è una variabile continua, con a e b costanti, si ha

$$E[aX + b] = aE[X] + b.$$

Se X è una variabile aleatoria continua di valore atteso μ, in analogia col caso discreto la varianza di X è definita da

$$\mathrm{Var}(X) = E[(X - \mu)^2] = \int_{-\infty}^{\infty} (x - \mu)^2 f(x)\, dx.$$

La formula alternativa,

$$\mathrm{Var}(X) = E[X^2] - (E[X])^2,$$

si dimostra imitando quanto fatto nel caso discreto. Analogamente, se a e b sono delle costanti, allora

$$\mathrm{Var}(aX + b) = a^2 Var(X).$$

Esempio. Sia X una variabile aleatoria continua avente densità

$$f_X(x) = \begin{cases} 2x & 0 \le x \le 1 \\ 0 & \text{altrimenti.} \end{cases}$$

Determinare la funzione di distribuzione e la varianza di X.

Soluzione. La funzione di distribuzione di X è data da

$$F_X(x) = \int_{-\infty}^{x} f_X(t)\,dt = \begin{cases} 0, & x < 0 \\ \int_0^x 2t\,dt = x^2, & 0 \le x < 1 \\ 1, & x \ge 1. \end{cases}$$

Ricaviamo la varianza di X ricordando che $\mu = E[X] = 2/3$; pertanto:

$$\mathrm{Var}(X) = \int_{-\infty}^{\infty} (x-\mu)^2 f_X(x)\,dx = \int_0^1 \left(x - \frac{2}{3}\right)^2 2x\,dx$$

$$= 2\int_0^1 \left(x^3 - \frac{4}{3}x^2 + \frac{4}{9}x\right) dx = 2\left[\frac{x^4}{4} - \frac{4}{9}x^3 + \frac{2}{9}x^2\right]_0^1 = \frac{1}{18}.$$

Si perviene allo stesso risultato anche notando che $E(X^2) = \int_0^1 2t^3\,dt = 1/2$.

Esempio. Sia X una variabile aleatoria continua avente funzione di distribuzione

$$F(x) = \begin{cases} 0 & x < 0 \\ \sqrt{x} & 0 \leq x < 1 \\ 1 & x \geq 1. \end{cases}$$

Determinare la densità, il valore atteso e la varianza di X.

Soluzione. Derivando $F(x)$ si ottiene la densità di X:

$$f(x) = \frac{d}{dx}F(x) = \begin{cases} \dfrac{1}{2\sqrt{x}} & 0 < x < 1 \\ 0 & \text{altrimenti.} \end{cases}$$

Da questa si ricavano i momenti di X:

$$E\left[X^n\right] = \int_{-\infty}^{\infty} x^n\, f(x)\, dx = \int_0^1 \frac{x^{n-1/2}}{2}\, dx = \frac{1}{2}\left[\frac{x^{n+1/2}}{n+1/2}\right]_0^1 = \frac{1}{2n+1},$$

da cui seguono valore atteso e la varianza di X:

$$E[X] = \frac{1}{3}, \qquad \mathrm{Var}(X) = E[X^2] - (E[X])^2 = \frac{1}{5} - \frac{1}{9} = \frac{4}{45}.$$