1 Variabili aleatorie continue

Definizione 1 Sia $X : \Omega \longrightarrow \mathbb{R}$ una v.a.; la funzione $F : \mathbb{R} \longrightarrow \mathbb{R}$ definita da:

$$F(x) = P(\{\omega \in \Omega : X(\omega) \leq x\}$$

o, più semplicemente $F(x) = P(X \leq x)$, si chiama *funzione di distribuzione (o ripartizione)* della v.a. X.
Si ha:

$$P(X \in (a, b]) = P(a < X \leq b) = P(X \leq b) - P(X \leq a) = F(b) - F(a).$$

Se X è discreta, l'andamento della sua f.d.d. è molto semplice da descrivere: se, per semplicità, supponiamo che X sia a valori interi, allora $F(x)$ è costante nell'intervallo tra due interi successivi, mentre può presentare una discontinuità in corrispondenza dei valori interi. La funzione di ripartizione è importante poiché la sua conoscenza è equivalente a quella della densità di X. Infatti, se $p(x)$ è la densità discreta di X, che assume valori $x_1, x_2, \ldots, x_n$, allora si ha:

$$F(x) = P(X \leq x) = \sum_{k:\ x_k \leq x} p(x_k),$$

che esprime la f.d.d. in termini della densità. Viceversa, se per esempio X prende solo valori interi, allora

$$p(k) = P(X = k) = P(k - 1 < X \leq k) = F(k) - F(k - 1). \tag{1.1}$$

Naturalmente, $F(x)$ è una funzione non decrescente. Infatti, sia $x_1 < x_2$, occorre mostrare che $F(x_1) \leq F(x_2)$; siccome l'evento $\{X \leq x_1\} \subset \{X \leq x_2\}$, abbiamo $P(\{X \leq x_1\}) \leq P(\{X \leq x_2\})$, ovvero $F(x_1) \leq F(x_2)$. Inoltre, risulta $F(x) = 0$ se $x < \min\{x_1, x_2, \ldots, x_n\}$ e $F(x) = 1$ se $x \geq \max\{x_1, x_2, \ldots, x_n\}$. A volte, per calcolare la densità di una v.a. può essere più facile calcolare prima la funzione di ripartizione $F(x)$, e poi da questa ricavare la densità tramite la (1.1). Questa è la procedura che abbiamo seguito, per esempio, per trovare la densità del tempo T di primo successo. Se k è intero:

$$F(k) = P(T \leq k) = 1 - P(T > k) = 1 - (1 - p)^k,$$

e quindi, se $k \leq x < k + 1$, ovvero $[x] = k$ ($[x]$ denota la parte intera di x), si ha

$$F(x) = P(X \leq k) = 1 - (1 - p)^{[x]}.$$

In Fig. 1 è riportato il grafico della Funzione di ripartizione di una v.a. uniforme su $(0, \ldots, 6)$, in Fig.2 quella di una v.a. Binomiale $B(6, 1/2)$. Osserviamo che, per una v.a. discreta X che assume valori $x_1, x_2, \ldots, x_n$, risulta

$$p_k = P(X = x_k) = F(x_k) - F(x_k^-),$$

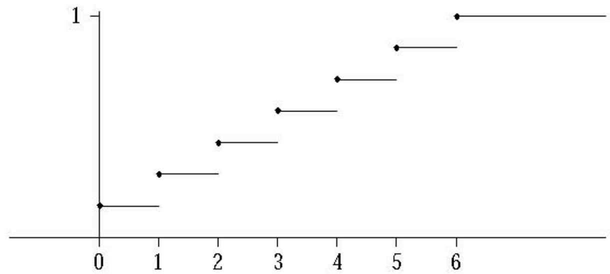

Fig. 1 Funzione di ripartizione di una v.a. uniforme su $0, \ldots, 6$.

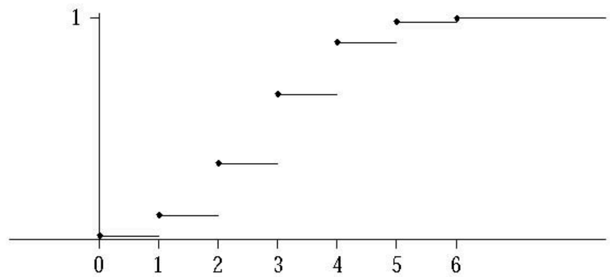

Fig. 2 Funzione di ripartizione di una v.a. di legge $B(6, 0.5)$.

Figure 1:

dove $F(x_0^-) = \lim_{x \to x_0^-} F(x)$; quindi, p_k è uguale all'ampiezza del salto attraverso il punto di discontinuità di F. Per $X \sim Uni(0, \ldots, 6)$, l'ampiezza del salto è costante ed uguale a $1/6$, per $X \sim B(6, 1/2)$, l'ampiezza dei salti va via via diminuendo.

Una v.a. X può assumere valori in un insieme continuo, e si dice v.a. continua. Per esempio, è questo il caso se X denota un numero aleatorio nell'intervallo $(0, 1)$: il range di X è l'intervallo $(0, 1)$, che è un s.i. continuo di $\mathbb{R}$.
La definizione di f.f.d. rimane invariata per v.a. continue. Riassumiamone le proprietà.

Proprietà della funzione di distribuzione $F(x) = P(X \leq x)$

1. $\forall x \in \mathbb{R}$ risulta $0 \leq F(x) \leq 1$;
2. $F(x)$ è una funzione non decrescente;
3. $\lim_{x \to -\infty} F(x) = 0$ e $\lim_{x \to +\infty} F(x) = 1$;
4. $\lim_{x \to x_0^+} F(x) = F(x_0)$, ovvero $F(x)$ è continua da destra.

Le dimostrazioni di **1.** e **2.** sono identiche a quelle relative ad una v.a. discreta. Le dimostrazioni di **3.** e **4.** sono un po' più complicate e rimandiamo al libro di testo, per approfondimenti.

Definizione 2 Una v.a. continua X si dice *assolutamente continua* se la sua funzione di distribuzione (f.d.d) $F(x)$ è continua, ed esiste una funzione $f(x) \geq 0$, integrabile in $\mathbb{R}$, tale che:

$$(i) \qquad \int_{-\infty}^{+\infty} f(x) dx = 1$$

(ii) $$F(x) = \int_{-\infty}^{x} f(t)dt.$$

La funzione $f(x)$ si chiama la *densità* (continua) della v.a. X.
Da (ii) si ottiene

$$P(a \leq X \leq b) = F(b) - F(a) = \int_a^b f(t)dt,$$

e, in generale, se $A \subset \mathbb{R}$:

$$P(X \in A) = \int_A f(t)dt.$$

Inoltre, se $f(x)$ è continua, per il teorema di Torricelli, si ha che la f.d.d. $F(x)$ è derivabile e risulta $F'(x) = f(x)$, ovvero $f(x) = \frac{d}{dx}P(X \leq x)$.

Se X è una v.a. assolutamente continua, allora

$$P(a < X < b) = P(a \leq X < b) = P(a < X \leq b) = P(a \leq X \leq b),$$

ovvero, per ogni valore x, risulta $P(X = x) = 0$. Ciò segue dal fatto che $P(X \in A) = \int_A f(t)dt$, e questo, come ben si sa, non cambia se aggiungiamo ad A un punto, o un'infinità mumerabile di punti.
Il significato della densità $f(x)$ è fornito dalla seguente considerazione. Abbiamo:

$$P(x < X \leq x + \Delta x) = F(x + \Delta x) - F(x) = \int_x^{x+\Delta x} f(t)dt$$

che, per il teorema della media integrale, essendo $f(x)$ continua, vale $f(\bar{x})$, dove $\bar{x} \in (x, x + \Delta x)$ è un conveniente punto. Quindi, si ha:

$$P(x \leq X \leq x + \Delta x) \approx \Delta x f(x), \ \Delta x \approx 0.$$

Nella tabella qui sotto, mettiamo a confronto le quantità essenziali, riguardanti v.a. discrete e continue.

X v.a. discreta	X v.a. assolutamente continua
$p(x) = P(X = x)$ densità discreta	$f(x) \geq 0$ densità (continua)
$\sum_i p_i = 1$	$\int_{\mathbb{R}} f(x)dx = 1$
$F(x) = P(X \leq x) = \sum_{i:x_i \leq x} p(x_i)$	$F(x) = \int_{-\infty}^{x} f(t)dt$
$p(x_k) = P(X = x_k) = F(x_k) - F(x_k^-)$	$f(x) = F'(x)$
$P(X \in A) = \sum_{i:x_i \in A} p(x_i)$	$P(X \in A) = \int_A f(x)dx$

Esempio 1. (v.a. uniforme nell'intervallo (a, b))
Ha densità

$$f(x) = \begin{cases} \frac{1}{b-a} & se \ x \in (a, b) \\ 0 & altrimenti. \end{cases}$$

La f.d.d. è

$$F(x) = \int_{-\infty}^{x} f(t)dt = \begin{cases} 0 & se \ x < a \\ \frac{x-a}{b-a} & se \ a \leq x \leq b \\ 1 & se \ x > b. \end{cases}$$

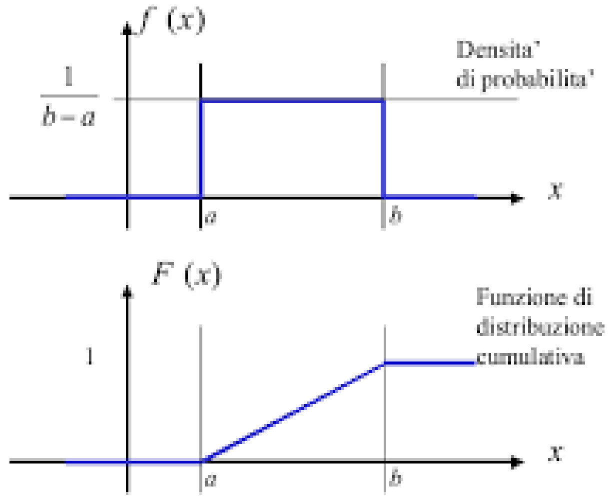

Figure 2: densità e funzione di distribuzione di $X \sim Uni(a, b)$

In particolare, $X \sim Uni(0, 1)$ ha densità

$$f(x) = \begin{cases} 1 & se\ x \in (0, 1) \\ 0 & altrimenti. \end{cases}$$

e

$$F(x) = \int_{-\infty}^{x} f(t)dt = \begin{cases} 0 & se\ x < 0 \\ x & se\ 0 \leq x \leq 1 \\ 1 & se\ x > 1. \end{cases}$$

Esempio 2. (v.a. esponenziale)
Per $\lambda > 0$, sia X una v.a. con densità

$$f(x) = \begin{cases} \lambda e^{-\lambda x} & se\ x > 0 \\ 0 & altrimenti. \end{cases}$$

X si chiama v.a. esponenziale di parametro λ. La f.d.d. è

$$F(x) = \int_{-\infty}^{x} \lambda e^{-\lambda t} dt = 1 - e^{-\lambda x}.$$

Per gli Esempi 2 e 3 è immediato verificare che $\int_{\mathbb{R}} f(x)dx = 1$.

* * * * *

Proprietà di mancanza di memoria di una v.a. esponenziale

Una variabile aleatoria continua $X \geq 0$ gode della proprietà di mancanza di memoria se per ogni $t, u > 0$:

$$P(X > t + u | X > t) = P(X > u) \qquad (A3)$$

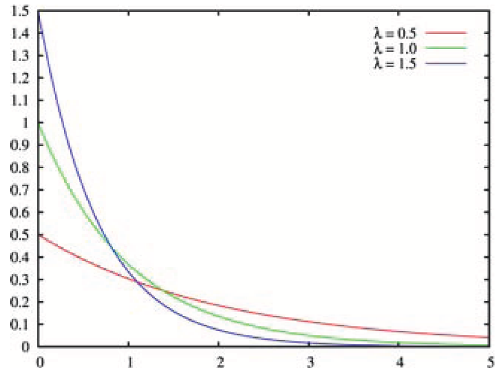

Figure 3: densità di X esponenziale di parametro λ

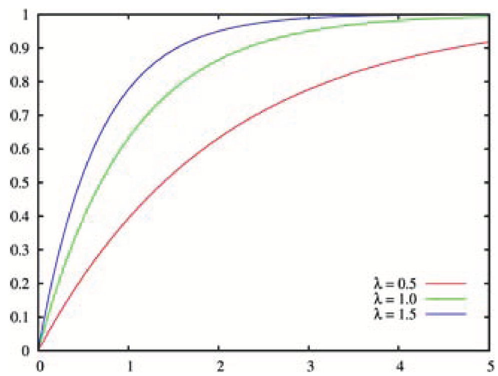

Figure 4: funzione di distribuzione di X esponenziale di parametro λ

ovvero

$$\frac{P(X > t + u)}{P(X > t)} = P(X > u), \tag{A4}$$

visto che il I membro di (A3) è uguale a:

$$\frac{P((X > t + u) \cap (X > t))}{P(X > t)} = \frac{P(X > t + u)}{P(X > t)}.$$

Se denotiamo con $F(t) = P(X \leq t)$ la funzione di **distribuzione** di X e con $S(t) = 1 - F(t) = P(X > t)$ la funzione di **sopravvivenza**, per la (A4) la proprietà di mancanza di memoria diventa:

$$\frac{S(t + u)}{S(t)} = S(u), \ \forall t, u > 0 \tag{A5}$$

1. Analogamente alla distribuzione geometrica modificata, se X ha distribuzione esponenziale di parametro $\lambda > 0$, allora X gode della proprietà di mancanza di memoria, ovvero vale la (A4). Infatti, si ha $F(t) = 1 - e^{-\lambda t}$, $S(t) = e^{-\lambda t}$ ed il primo membro della (A4) diventa

$$\frac{e^{-\lambda(t+u)}}{e^{-\lambda t}} = e^{-\lambda u} \equiv S(u).$$

Se X è per esempio un tempo di attesa, la (A4) ed equivalentemente la (A3), si interpretano dicendo che la probabilità condizionale che il tempo di attesa X sia maggiore di $t + u$, dato che esso è maggiore di t, non dipende da quanto si è già atteso, ossia da t.

La v.a. geometrica modificata e quella esponenziale sono, rispettivamente tra quelle discrete e quelle assolutamente continue, le **uniche** v.a. che godono della proprietà di mancanza di memoria. Infatti, vale la seguente:

Proposizione. *Sia $X \geq 0$ una variabile aleatoria continua, con densità continua f in $(0, +\infty)$ e tale che esista $\lim_{t\to 0^+} f(t) = f(0^+)$. Allora, se X gode della proprietà di mancanza di memoria (A5), X ha distribuzione esponenziale.*

Dim. Risolvere la (A5) equivale a trovare una funzione non crescente $S : [0, +\infty) \longrightarrow [0, 1]$, derivabile in $(0, +\infty)$ e provvista di derivata destra in 0, tale che

$$\begin{cases} S(t+u) = S(t)S(u) \\ S(0) = 1, \ S(+\infty) = 0 \end{cases} \tag{A6}$$

(S è la funzione di sopravvivenza, cioé $S(t) = 1 - F(t) = P(X > t)$).
Si osservi che per una v.a. continua, con densità f continua in $(0, +\infty)$, la derivabilità di F per $t > 0$ e quindi di S, è garantita dal Teorema fondamentale del Calcolo integrale, avendosi $F'(t) = f(t)$.
Si ha ora, per la (A6):

$$\frac{S(t+u) - S(t)}{u} = \frac{S(t)S(u) - S(t)}{u} = S(t)\frac{S(u) - 1}{u}$$

Passando al limite per $u \to 0^+$, si ottiene, per $t > 0$:

$$S'_+(t) \equiv S'(t) = S(t)S'_+(0) \tag{A7}$$

ovvero

$$\frac{S'(t)}{S(t)} = S'_+(0) = -\lambda \leq 0$$

e ancora $(\ln S(t))' = -\lambda$, da cui $\ln S(t) = -\lambda t + c$, $S(t) = e^c \cdot e^{-\lambda t}$. Infine, utilizzando il fatto che S(0)=1, si ottiene $S(t) = e^{-\lambda t}$, ovvero $F(t) = 1 - e^{-\lambda t}$, e quindi X ha distribuzione esponenziale di parametro $\lambda > 0$; si esclude infatti che sia $\lambda = 0$, poiché in tal caso sarebbe $S(t) \equiv 1$, che non è la funzione di sopravvivenza di alcuna v.a. X (S deve essere non crescente, $S(0) = 1$, $S(+\infty) = 0$).

Si osservi per inciso che, se non si assumesse la derivabilità di S per $t > 0$, ma solo l'esistenza della derivata destra di S in 0, comunque la (A7) implicherebbe l'esistenza della derivata destra di S in $t > 0$; si avrebbe infatti $S'_+(t) = -S(t)f(0^+)$, essendo $S'_+(0) = -f(0^+)$.

* * * * *

Esempio 3. (v.a. normale o Gaussiana, $X \sim \mathcal{N}(m, \sigma^2)$)
Per $m \in \mathbb{R}$ e $\sigma > 0$, si chiama v.a. normale, o Gaussiana, di parametri m e σ una v.a. con densità

$$f(x) = \frac{1}{\sigma\sqrt{2\pi}} e^{-(x-m)^2/2\sigma^2}, \ x \in \mathbb{R}.$$

Verificheremo in seguito che $\int_{\mathbb{R}} f(x)dx = 1$.

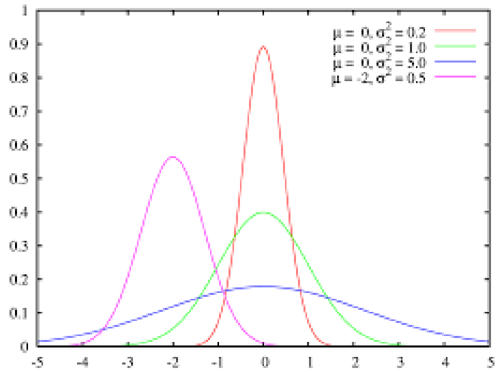

Figure 5: densità di $X \sim \mathcal{N}(\mu, \sigma)$

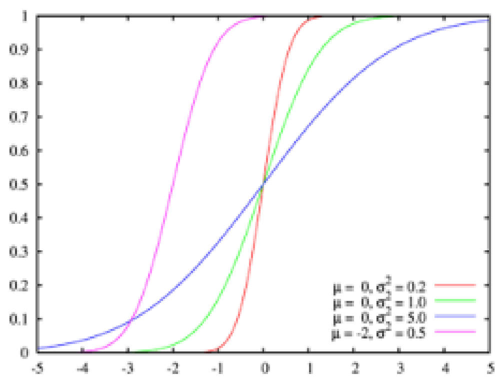

Figure 6: funzione di distribuzione di di $X \sim \mathcal{N}(\mu, \sigma)$

Importante è il caso in cui $m = 0$ e $\sigma = 1$; in tal caso si parla di v.a. normale standard $X \sim \mathcal{N}(0, 1)$, e la densità è:

$$f(x) = \frac{1}{\sqrt{2\pi}} e^{-x^2/2}, \ x \in \mathbb{R}.$$

Il grafico della densità di una v.a. $X \sim \mathcal{N}(m, \sigma^2)$ ha la caratteristica forma a *campana* incentrata nel punto m e con vertice nel punto di ascissa m e ordinata $\frac{1}{\sigma\sqrt{2\pi}}$; al diminuire di σ la campana diviene sempre più stretta e il punto di max si alza (viceversa, al crescere di σ la campana diventa sempre più slargata e il punto di max si abbassa). Per $x \to \pm\infty$, $f(x)$ tende a zero, più che esponenzialmente.

Esempio 4. (v.a. con densità Gamma, $X \sim \ Gamma(\alpha, \lambda)$)
Per $\alpha, \ \lambda > 0$ è la v.a. con densità

$$f(x) = \begin{cases} \frac{\lambda^\alpha}{\Gamma(\alpha)} x^{\alpha-1} e^{-\lambda x} & \textit{se } x > 0 \\ 0 & \textit{altrimenti}. \end{cases}$$

Qui

$$\Gamma(\alpha) = \int_0^{+\infty} x^{\alpha-1} e^{-x} dx$$

è la funzione Gamma di Eulero. Risulta $\Gamma(1) = \int_0^{+\infty} e^{-x} dx = 1$; inoltre, integrando per

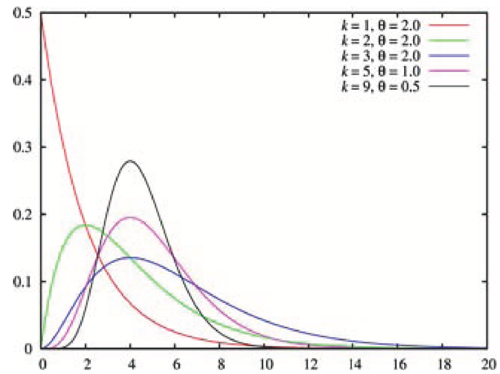

Figure 7: densità di $X \sim Gamma(k, \theta)$

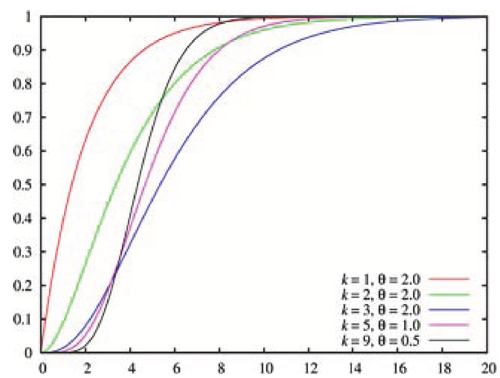

Figure 8: funzione di distribuzione di $X \sim Gamma(k, \theta)$

parti:

$$\Gamma(\alpha+1) = \int_0^{+\infty} x^\alpha e^{-x} dx = \left[-x^\alpha e^{-x}\right]_0^{+\infty} + \alpha \int_0^{+\infty} x^{\alpha-1} e^{-x} dx = \alpha\Gamma(\alpha).$$

Da ciò si ricava $\Gamma(n+1) = n!$, ovvero la funzione Γ interpola il fattoriale.
Come si vede dalla definizione, se $\alpha = 1$, allora $X \sim Gamma(1, \lambda)$ non è altro che una v.a esponenziale di parametro λ.

Verifichiamo che l'integrale di $f(x)$ vale 1. Infatti:

$$\int_{-\infty}^{+\infty} f(x)dx = \frac{\lambda^\alpha}{\Gamma(\alpha)} \int_0^{+\infty} x^{\alpha-1} e^{-\lambda x} dx;$$

col cambio di variabile $t = \lambda x$, otteniamo

$$\frac{\lambda^\alpha}{\Gamma(\alpha)} \int_0^{+\infty} \left(\frac{t}{\lambda}\right)^{\alpha-1} \frac{e^{-t}}{\lambda} \, dt$$

$$= \frac{\lambda^\alpha}{\Gamma(\alpha)} \frac{1}{\lambda^\alpha} \int_0^{+\infty} t^{\alpha-1} e^{-t} dt;$$

Ma, l'ultimo integrale scritto vale $\Gamma(\alpha)$, quindi si ottiene infine

$$\int_{-\infty}^{+\infty} f(x)dx = \frac{\lambda^\alpha}{\Gamma(\alpha)} \frac{1}{\lambda^\alpha} \Gamma(\alpha) = 1.$$

Non sono disponibili formule semplici per la f.d.d. di una v.a. $X \sim Gamma(\alpha, \lambda)$, a meno che α non sia un numero intero. In tal caso, cioè se $\alpha = m$ è intero, si può dimostrare che la f.d.d. di $X \sim Gamma(m, \lambda)$ è:

$$F(x) = P(X \leq x) = 1 - e^{-\lambda x} \sum_{i=0}^{m-1} \frac{(\lambda x)^i}{i!}.$$

Le leggi Gamma, quando il parametro $\alpha = m$ è intero, vengono anche dette leggi di Erlang di parametri m e λ.

1.1 Calcoli con densità

1. Sia X v.a. con densità $f_X(x)$, e $a, b, \in \mathbb{R}$, $a \neq 0$; qual è la densità della v.a.

$$Y = aX + b?$$

(i) $(a > 0)$ la f.d.d. di Y è:

$$F_Y(y) = P(Y \leq y) = P(aX + b \leq y) = P(X \leq \frac{y-b}{a}) = F_X(\frac{y-b}{a});$$

derivando:

$$f_Y(y) = F'_Y(y) = \frac{1}{a} f_X(\frac{y-b}{a}).$$

(ii) $(a < 0)$ la f.d.d. di Y è:

$$F_Y(y) = P(Y \leq y) = P(aX + b \leq y) = P(X \geq \frac{y-b}{a}) = 1 - F_X(\frac{y-b}{a});$$

derivando:

$$f_Y(y) = F'_Y(y) = -\frac{1}{a} f_X(\frac{y-b}{a}).$$

Pertanto, mettendo insieme i due casi:

$$f_Y(y) = \frac{1}{|a|} f_X(\frac{y-b}{a}). \tag{1.2}$$

Esempio Se $X \sim \mathcal{N}(0,1)$, calcolare la densità di $Y = \sigma X + \mu$, $\sigma > 0$.
Si ha;

$$f_X(x) = \frac{1}{\sqrt{2\pi}} e^{-x^2/2}, \ x \in \mathbb{R};$$

allora, usando (1.2) con $a = \sigma$ e $b = \mu$, si ottiene

$$f_Y(y) = \frac{1}{\sigma} f_X(\frac{y-\mu}{\sigma}) = \frac{1}{\sigma\sqrt{2\pi}} e^{-(y-\mu)^2/2\sigma^2}, \ y \in \mathbb{R},$$

ovvero:

$$X \sim \mathcal{N}(0,1) \ \Rightarrow \ Y = \sigma X + \mu \sim \mathcal{N}(\mu, \sigma^2),$$

viceversa:

$$Y \sim \mathcal{N}(\mu, \sigma^2) \;\Rightarrow\; X = \frac{Y-\mu}{\sigma} \sim \mathcal{N}(0,1).$$

Ciò implica che la densità Normale standard è importante, in quanto qualunque densità Normale si scrive in funzione di quella standard. Per quanto riguarda la f.f.d., se $Y \sim \mathcal{N}(\mu, \sigma^2)$, allora $Y = \sigma X + \mu$, dove $X \sim \mathcal{N}(0,1)$; quindi, se $\Phi(x) = P(X \le x)$ denota la f.d.d. di una v.a. Normale standard X, si ha:

$$F_Y(y) = P(Y \le y) = P(\sigma X + \mu \le y) = \Phi\left(\frac{y-\mu}{\sigma}\right).$$

I valori della f.d.d. della Normale standard, $\Phi(x)$, sono stati calcolati mumericamente e tabulati, essendo impossibile calcolare esplicitamente l'integrale

$$\Phi(x) = \int_{-\infty}^{x} \frac{1}{\sqrt{2\pi}} e^{-t^2/2} dt;$$

il grafico di $\Phi(x)$ ha la caratteristica forma ad "S", con $\lim_{x\to-\infty} \Phi(x) = 0,\ \lim_{x\to+\infty} \Phi(x) = 1$ e $\Phi(0) = 1/2$.

2. Sia X v.a. con densità $f_X(x)$; qual è la densità della v.a.

$$Y = X^2?$$

La v.a. Y assume valori non negativi; quindi, se $y \ge 0$:

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

Derivando, si ottiene la densità di Y :

$$f_Y(y) = F_Y'(y) = \frac{1}{2\sqrt{y}}(f_X(\sqrt{y}) + f_X(-\sqrt{y}),\ y \ge 0.$$

Per esempio, se $X \sim \mathcal{N}(0,1)$, la v.a. $Y = X^2$ ha densità:

$$f_Y(y) = \frac{1}{\sqrt{y}}\phi(\sqrt{y}) = \frac{1}{y\sqrt{2\pi}} e^{-y/2},\ y \ge 0,$$

poiché la densità normale standard

$$\phi(x) = \frac{1}{\sqrt{2\pi}} e^{-x^2/2}$$

è una funzione pari, ovvero $\phi(-x) = \phi(x)$.

3. Sia X v.a. con densità $f_X(x)$ e $g : \mathbb{R} \longrightarrow \mathbb{R}$ una funzione derivabile, strettamente monotona (crescente o decrescente); qual è la densità della v.a. $Y = g(X)$?
(i) $g(x)$ crescente; si ha:

$$F_Y(y) = P(Y \le y) = P(g(X) \le y) = P(X \le g^{-1}(y)) = F_X(g^{-1}(y)).$$

Derivando rispetto a y, si ottiene:

$$f_Y(y) = F'_Y(y) = \frac{d}{dy}\left[F_X(g^{-1}(y))\right] = F'_X(g^{-1}(y))(g^{-1})'(y) = f_X(g^{-1}(y))\frac{1}{g'(g^{-1}(y))}.$$

(ii) $g(x)$ decrescente; si ha:

$$F_Y(y) = P(Y \leq y) = P(g(X) \leq y) = P(X \geq g^{-1}(y)) = 1 - F_X(g^{-1}(y)).$$

Derivando rispetto a y, si ottiene:

$$f_Y(y) = F'_Y(y) = \frac{d}{dy}\left[1 - F_X(g^{-1}(y))\right] = -F'_X(g^{-1}(y))(g^{-1})'(y) = -f_X(g^{-1}(y))\frac{1}{g'(g^{-1}(y))}.$$

Pertanto, mettendo insieme i due casi ($g(x)$ strettamente monotona, cioè invertibile):

$$f_Y(y) = f_X(g^{-1}(y))\frac{1}{|g'(g^{-1}(y))|}. \tag{1.3}$$

4. Siano X e Y v.a. indipendenti, con densità $f_X(x)$ e $f_Y(y)$ rispettivamente; allora la densità di $Z = X + Y$ è:

$$f_Z(z) = \int_{-\infty}^{+\infty} f_X(x) f_Y(z - x)dx.$$

La dimostrazione di questa formula la vedremo più oltre; osserviamo che essa è analoga a quella che fornisce la densità discreta della somma di due v.a. discrete indipendenti; naturalmente la sommatoria è qui sostituita da un integrale.

1.2 Leggi Normali

$X \sim \mathcal{N}(0, 1)$ (Normale standard) ha densità

$$f_X(x) = \phi(x) = \frac{1}{\sqrt{2\pi}}e^{-x^2/2}, \; x \in \mathbb{R}.$$

$Y = \sigma X + \mu \sim \mathcal{N}(\mu, \sigma^2)$ (Normale di parametri μ e σ^2, con $\sigma > 0$) ha densità

$$f_Y(y) = \frac{1}{\sigma} f_X((x - \mu)/\sigma) = \frac{1}{\sigma}\phi((x - \mu)/\sigma) = \frac{1}{\sigma\sqrt{2\pi}}e^{-(y-\mu)^2/2\sigma^2}, \; y \in \mathbb{R}.$$

$$P(X \leq x) = \Phi(x) = \int_{-\infty}^{x} \phi(t)dt.$$

Osservazione Se $X \sim \mathcal{N}(0, 1)$, anche $-X \sim \mathcal{N}(0, 1)$.
Infatti,

$$P(-X \leq t) = P(X \geq -t) = 1 - P(X \leq -t) = 1 - \Phi(-t),$$

ma, essendo $\Phi(-t) = 1 - \Phi(t)$, si ottiene:

$$P(-X \leq t) = \Phi(t) = P(X \leq t),$$

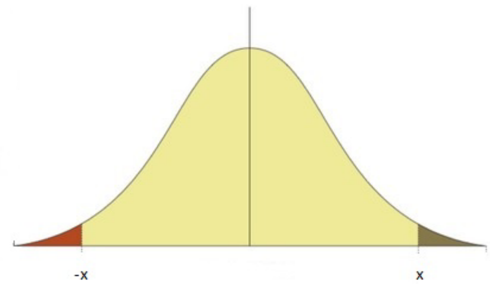

Figure 9: Spiegazione grafica del perché $\Phi(-x) = 1 - \Phi(x)$

ovvero X e $-X$ hanno la stessa distribuzione Normale standard (e quindi la stessa densità). Si osservi che, dire che due v.a. hanno la stessa distribuzione non vuol dire che esse siano uguali. Per es., se $X, Y \sim \mathcal{N}(0,1)$, esse hanno uguale distribuzione, ma, posto $Z = X - Y$, si ha $P(X = Y) = P(Z = 0) = 0$, essendo Z una v.a. assolutamente continua.
Il fatto che $\Phi(-t) = 1 - \Phi(t)$ si può dimostrare in tanti modi; per esempio, in maniera geometrica, osservando che $\Phi(-t)$ rappresenta l'area della regione di piano situata sotto il grafico della campana di Gauss, avente l'ascissa $\leq t$, mentre $1 - \Phi(t)$ rappresenta l'area della regione di piano situata sotto il grafico della campana, avente l'ascissa $\geq t$, e le due aree sono uguali, essendo la curva a campana simmetrica rispetto a 0 (vedi Fig. 9).

Quantili di una v.a. $X \sim \mathcal{N}(0,1)$
Se X è una v.a. e $\alpha \in (0,1)$, si chiama quantile di X di ordine α la quantità ϕ_α tale che

$$P(X \leq \phi_\alpha) = \alpha.$$

Se X è normale standard, siccome $-X \sim X$, si ha:

$$P(X \leq -\phi_\alpha) = P(-X \leq -\phi_\alpha) = P(X \geq \phi_\alpha) = 1 - P(X \leq \phi_\alpha) = 1 - \alpha.$$

Pertanto, abbiamo ottenuto, per $X \sim \mathcal{N}(0,1)$, che:

$$\phi_{1-\alpha} = -\phi_\alpha.$$

Inoltre, se $X \sim \mathcal{N}(0,1)$, allora $Y = \sigma X + \mu \sim \mathcal{N}(\mu, \sigma^2)$; allora

$$P(Y \leq y) = P(\sigma X + \mu \leq y) = P(X \leq (y - \mu)/\sigma) = \Phi((y - \mu)/\sigma).$$

Per i quantili, q_α, di una v.a. $Y \sim \mathcal{N}(\mu, \sigma^2)$, si ha:

$$\alpha = P(Y \leq q_\alpha) = P(\sigma X + \mu \leq q_\alpha) = P(X \leq (q_\alpha - \mu)/\sigma);$$

abbiamo trovato allora che $P(X \leq (q_\alpha - \mu)/\sigma) = \alpha$, e quindi per definizione di quantile di X, deve essere $(q_\alpha - \mu)/\sigma = \phi_\alpha$, ovvero

$$q_\alpha = \sigma\phi_\alpha + \mu.$$

Theorem 1.1 *(Somma di v.a. indipendenti con legge Normale) Siano X e Y v.a. indipendenti, con $X \sim \mathcal{N}(\mu, \sigma^2)$, $Y \sim \mathcal{N}(\lambda, \tau^2)$; allora*

$$Z = X + Y \sim \mathcal{N}(\mu + \lambda, \sigma^2 + \tau^2).$$

Rimandiamo più oltre la dimostrazione di questo importante teorema. Ci limitiamo ad osservare che esso vale anche per più di due v.a. indipendenti con leggi Normali, ovvero la somma di n v.a. indipendenti, di leggi $\mathcal{N}(\mu_i, \sigma_i^2),\ i = 1, \dots, n$ ha legge $\mathcal{N}(\mu_1 + \cdots + \mu_n, \sigma_1^2 + \cdots + \sigma_n^2)$.

1.3 Ancora sulle Leggi Gamma

Ricordiamo che per $\alpha,\ \lambda > 0$ una v.a. X ha distribuzione Gamma (α, λ) se la sua densità è:

$$f(x) = \begin{cases} \frac{\lambda^\alpha}{\Gamma(\alpha)} x^{\alpha-1} e^{-\lambda x} & se\ x > 0 \\ 0 & altrimenti. \end{cases}$$

Vale il seguente risultato:

Proposition 1.2 *Se $X \sim \mathcal{N}(0, \sigma^2)$, allora $Y = X^2 \sim Gamma(1/2, 1/(2\sigma^2))$.*

Infatti, per la formula già vista per la densità del quadrato di una v.a., se $y > 0$ si ha:

$$f_Y(y) = \frac{1}{2\sqrt{y}}[f_X(\sqrt{y}) + f_X(-\sqrt{y})] = \frac{1}{2\sqrt{y}} \frac{1}{\sigma\sqrt{2\pi}} \left(e^{-y/2\sigma^2} + e^{-y/2\sigma^2}\right)$$

$$= \frac{1}{\sigma\sqrt{2\pi}} y^{1/2-1} e^{-y/2\sigma^2},\ y > 0;$$

questa (confrontando con la definizione della densità Gamma) è la densità di una v.a. $Gamma(1/2, 1/(2\sigma^2))$, purché risulti:

$$\frac{1}{\sigma\sqrt{2\pi}} = \sqrt{\frac{1}{2\sigma^2}} \Big/ \Gamma(1/2),$$

il che implica l'utile uguaglianza $\Gamma(1/2) = \sqrt{\pi}$.

Theorem 1.3 *(Somma di v.a. indipendenti con legge Gamma) Siano X e Y v.a. indipendenti, con $X \sim Gamma(\alpha, \lambda)$, $Y \sim Gamma(\beta, \lambda)$; allora*

$$Z = X + Y \sim Gamma(\alpha + \beta, \lambda).$$

Rimandiamo più oltre la dimostrazione di questo importante teorema. Ci limitiamo ad osservare che esso vale anche per più di due v.a. indipendenti con leggi Gamma, ovvero la somma di n v.a. indipendenti, di leggi $Gamma(\alpha_i, \lambda),\ i = 1, \dots, n$ ha legge $Gamma(\alpha_1 + \cdots + \alpha_n, \lambda)$. Utilizzando questo teorema, si può ottenere:

Se $X_1, X_2, \dots, X_n$ sono v.a. indipendenti con distribuzione $\mathcal{N}(0, \sigma^2)$, allora:

$$Y = X_1^2 + X_2^2 + \cdots + X_n^2 \sim Gamma(n/2, \sigma^2/2)$$

Infatti, per la Proposizione 1.2, $X_i^2 \sim Gamma(1/2, \sigma^2/2)$ da cui, per il teorema della somma di v.a. Gamma, segue subito il risultato.
Una v.a. che ha densità $Gamma(n/2, 1/2)$ si dice che ha distribuzione del Chi quadrato con n gradi di libertà, e si indica con $\chi^2(n)$.