3 Una famiglia di funzioni: interpretazione probabilistica

3.1 Funzione di densità

La funzione

$$f(x) = \frac{1}{\sqrt{2\pi}\sigma} e^{-\frac{(x-\mu)^2}{2\sigma^2}}$$

è una funzione di densità per ogni $\mu, \sigma \in \mathbb{R}$*, con* $\sigma > 0$.

Infatti vale:

- $f(x) \geq 0 \ \forall x \in \mathbb{R}$
  dato che la funzione esponenziale assume solo valori positivi e σ è positivo per ipotesi[a]
- $\displaystyle\int_{-\infty}^{+\infty} f(x)\,\mathrm{dx} = 1$
  Si può dimostrare[b] che $\int_{-\infty}^{+\infty} e^{-\frac{(x-\mu)^2}{2\sigma^2}}\,\mathrm{dx} = \sqrt{2\pi}\sigma$. Pertanto il fattore $\frac{1}{\sqrt{2\pi}\sigma}$ che compare nella funzione f serve proprio affinché l'integrale di f su $\mathbb{R}$ sia 1

Definizione. La funzione f si dice funzione di **densità normale** (con parametri μ e σ^2) e la variabile aleatoria (spesso abbreviata in v.a.) ad essa relativa si dice **variabile aleatoria normale**.

[a] Anzi $f(x) > 0 \ \forall x \in \mathbb{R}$.
[b] Con strumenti di analisi che vanno oltre le competenze liceali.

Osservazione. La probabilità che X sia compresa tra due valori $a, b \in \mathbb{R}$, come per ogni v.a. continua, si può esprimere nella forma:

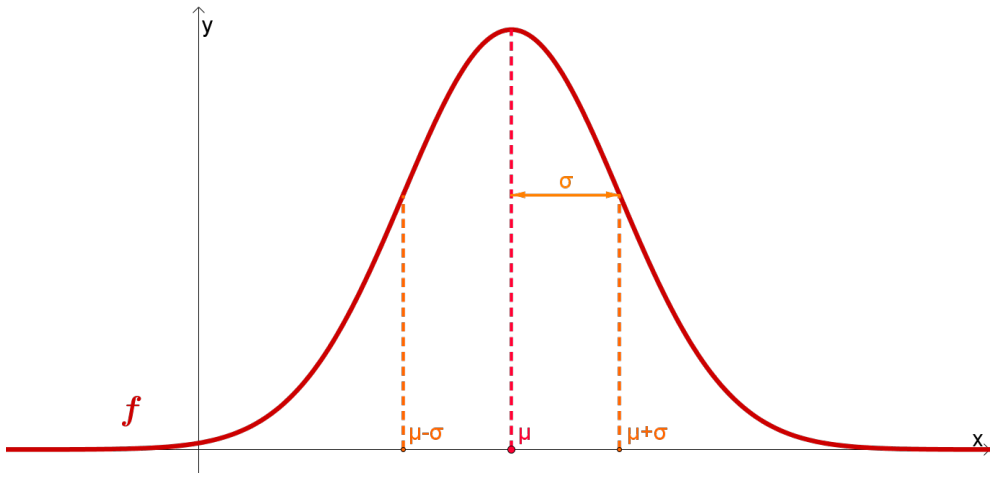
$$P(a \leq X \leq b) = \int_a^b f(x)\,\mathrm{d}x$$

dove f è la funzione di densità normale.
C'è però un problema: tale integrale *non* è risolvibile elementarmente, cioè la primitiva di f non si può esprimere mediante una formula che contenga le funzioni elementari e le operazioni usuali. Vedremo nel seguito come risolvere tale questione.

3.2 Significato probabilistico dei parametri

Per comprendere il significato probabilistico dei parametri μ e σ, iniziamo con l'esaminare l'espressione analitica e il grafico della densità:

$$f(x) = \frac{1}{\sqrt{2\pi}\sigma} e^{-\frac{(x-\mu)^2}{2\sigma^2}} \quad dove\ \sigma > 0$$



Da essi deduciamo il significato geometrico dei due parametri, come visto nel paragrafo precedente (2.3).

Ora però compiamo un ulteriore passo: utilizziamo il fatto che f è una densità di *probabilità*. Queste due considerazioni[15] insieme, cosa permettono di affermare sui parametri?

Intuitivamente possiamo dire che:

- μ rappresenta il "centro" della distribuzione
- σ^2 è una misura della "larghezza" della distribuzione, rispetto al "centro"

Ci *aspettiamo* dunque che:

- μ sia uguale o legato al valore atteso della v.a. normale
- σ^2 sia uguale o legato alla varianza della v.a. normale

Precisamente, mediante il calcolo integrale si può dimostrare il seguente risultato:

Proposizione. Sia X una v.a. normale, cioè una v.a. continua che ha densità

$$f(x) = \frac{1}{\sqrt{2\pi}\sigma} e^{-\frac{(x-\mu)^2}{2\sigma^2}} \quad dove\ \sigma > 0$$

Allora la media di X è μ e la varianza di X è σ^2.

Proveremo questo risultato nel caso della v.a. normale di parametri $\mu = 0$ e $\sigma^2 = 1$. Tale variabile si indica spesso con Z e si dice v.a. **normale standard**.

[15] Ossia il significato geometrico dei parametri e il fatto che f è una densità di probabilità.

3.3 Media e varianza della distribuzione normale standard -attività-

Consideriamo la variabile aleatoria normale standard, cioè la v.a. continua che ha densità

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

Mostra che la media di tale variabile è 0 *e la varianza è* 1.

Suggerimenti

a) Ricorda che, per definizione di media, la media della v.a. normale standard è

$$\int_{-\infty}^{+\infty} x\, f(x)\, \mathrm{d}x$$

Per determinare il valore di questo integrale non serve scrivere esplicitamente una primitiva della funzione $xf(x)$: basta osservare che essa è dispari.

b) Secondo la definizione di varianza, quella della v.a. normale standard risulta uguale a

$$\int_{-\infty}^{+\infty} x^2 f(x)\, \mathrm{d}x$$

Per determinare il valore di questo integrale, osserva che la funzione $x^2\, e^{-x^2/2}$ si può scrivere nella forma $x\,(xe^{-x^2/2})$ e dunque si può integrare per parti.

Risoluzione

a) Per trovare la media della v.a. normale standard calcoliamo dunque il seguente integrale:

$$\frac{1}{\sqrt{2\pi}} \int_{-\infty}^{+\infty} x\, e^{-x^2/2}\, \mathrm{d}x$$

Osserviamo che la funzione $g(x) = xe^{-x^2/2}$ è dispari, ossia vale $g(-x) = -g(x)\ \forall x \in \mathbb{R}$. Ciò significa che il grafico della funzione g è simmetrico rispetto l'origine degli assi. Tale proprietà ci permette di dedurre direttamente il valore dell'integrale in esame.

Per comprenderlo vediamo una situazione più semplice e consideriamo una funzione h dispari, come quella rappresentata in figura 8. Per ragioni di simmetria, l'area del sottoinsieme del piano delimitato dal grafico di h e dall'asse x nell'intervallo $[-t; 0]$ è uguale all'area del sottoinsieme del piano delimitato dal grafico di h e dall'asse x nell'intervallo $[0; t]$, per qualsiasi reale $t > 0$. Ma la funzione h assume segni opposti in questi due intervalli, per cui

$$\int_{-t}^{0} h(x)\, \mathrm{d}x = -\int_{0}^{t} h(x)\, \mathrm{d}x$$

Pertanto vale anche

$$\int_{-\infty}^{0} h(x)\, \mathrm{d}x = -\int_{0}^{+\infty} h(x)\, \mathrm{d}x$$

e quindi

$$\int_{-\infty}^{+\infty} h(x)\, \mathrm{d}x = 0$$

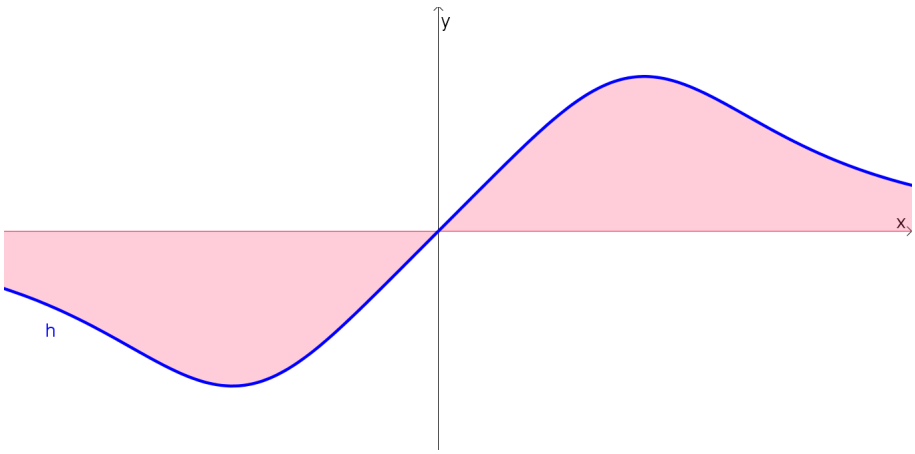


Figura 8: Rappresentazione del sottoinsieme del piano delimitato dall'asse x e dal grafico della funzione $h(x) = \dfrac{x}{1+x^4}$.

Più in generale, ragionando come nell'esempio, si ottiene che l'integrale su $\mathbb{R}$ di una funzione (integrabile) dispari è 0. Per cui anche nel nostro caso

$$\frac{1}{\sqrt{2\pi}} \int_{-\infty}^{+\infty} x\, e^{-x^2/2}\, \mathrm{d}x = 0$$

Ossia la media della v.a. normale standard è 0.

b) Vediamo ora il calcolo della varianza della v.a. normale standard. Nel nostro caso essa è uguale a

$$\frac{1}{\sqrt{2\pi}} \int_{-\infty}^{+\infty} x^2 \, e^{-x^2/2} \, \mathrm{d}x$$

Integriamo per parti:

$$\begin{aligned}
\frac{1}{\sqrt{2\pi}} \int_{-\infty}^{+\infty} x \left(x e^{-x^2/2} \right) \mathrm{d}x &= \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{+\infty} x \left(-e^{-x^2/2} \right)' \mathrm{d}x \\
&= -\frac{1}{\sqrt{2\pi}} \left[x e^{-x^2/2} \right]_{-\infty}^{+\infty} + \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{+\infty} e^{-x^2/2} \, \mathrm{d}x \\
&= 0 + 1 = 1
\end{aligned}$$

Il primo addendo è 0: essenzialmente ciò è dovuto al fatto che[16] $\lim_{x \to +\infty} (x e^{-x^2/2}) = 0$. Per determinare il secondo addendo, invece, utilizziamo il fatto che f è una densità (si veda paragrafo 3.1) e dunque vale

$$\int_{-\infty}^{+\infty} f(x) \, \mathrm{d}x = 1$$

Concludiamo così che la varianza della v.a. normale standard è 1.

[16] Precisamente:

$$\left[x e^{-x^2/2} \right]_{-\infty}^{+\infty} = \lim_{k \to -\infty} \left[x e^{-x^2/2} \right]_{k}^{0} + \lim_{h \to +\infty} \left[x e^{-x^2/2} \right]_{0}^{h} = \lim_{k \to -\infty} (-k e^{-k^2/2}) + \lim_{h \to +\infty} (h e^{-h^2/2}) = 0 + 0$$

dove con la prima uguaglianza abbiamo esplicitato il significato della prima scrittura.