Probabilità e Processi Stocastici (455AA)

Lezione 10

Dario Trevisan – *https://web.dm.unipi.it/trevisan*

22/10/2025

Variabili aleatorie gaussiane

Cosa vedremo:

- la definizione e le principali proprietà (sia nel caso reale che nel caso vettoriale)

Cosa vedremo:

- la definizione e le principali proprietà (sia nel caso reale che nel caso vettoriale)
- come stimare i parametri sulla base di osservazioni indipendenti (un campione) tutte con gli stessi parametri

Cosa vedremo:

- la definizione e le principali proprietà (sia nel caso reale che nel caso vettoriale)
- come stimare i parametri sulla base di osservazioni indipendenti (un campione) tutte con gli stessi parametri
- l'analisi delle componenti principali (PCA) e una giustificazione tramite opportune variabili gaussiane

Cosa vedremo:

- la definizione e le principali proprietà (sia nel caso reale che nel caso vettoriale)
- come stimare i parametri sulla base di osservazioni indipendenti (un campione) tutte con gli stessi parametri
- l'analisi delle componenti principali (PCA) e una giustificazione tramite opportune variabili gaussiane
- qualche rudimento della regressione, in particolare il metodo dei minimi quadrati, giustificato tramite opportune ipotesi di gaussianità

Cosa vedremo:

- la definizione e le principali proprietà (sia nel caso reale che nel caso vettoriale)
- come stimare i parametri sulla base di osservazioni indipendenti (un campione) tutte con gli stessi parametri
- l'analisi delle componenti principali (PCA) e una giustificazione tramite opportune variabili gaussiane
- qualche rudimento della regressione, in particolare il metodo dei minimi quadrati, giustificato tramite opportune ipotesi di gaussianità
- un cenno ai metodi principali per giustificare l'ipotesi di gaussianità

Cosa vedremo:

- la definizione e le principali proprietà (sia nel caso reale che nel caso vettoriale)
- come stimare i parametri sulla base di osservazioni indipendenti (un campione) tutte con gli stessi parametri
- l'analisi delle componenti principali (PCA) e una giustificazione tramite opportune variabili gaussiane
- qualche rudimento della regressione, in particolare il metodo dei minimi quadrati, giustificato tramite opportune ipotesi di gaussianità
- un cenno ai metodi principali per giustificare l'ipotesi di gaussianità
- un cenno al metodo di Laplace per approssimare densità generali con opportune gaussiane.

Il caso reale: definizione veloce

Si dice che una variabile aleatoria $X \in \mathbb{R}$ ha densità continua gaussiana se vale

$$p(X = x) \propto \exp\left(ax^2 + bx\right), \quad \text{per ogni } x \in \mathbb{R},$$

per degli opportuni parametri a, $b \in \mathbb{R}$.

- la densità è l'esponenziale di un polinomio di secondo grado dei possibili valori $x \in \mathbb{R}$.

Il caso reale: definizione veloce

Si dice che una variabile aleatoria $X \in \mathbb{R}$ ha densità continua gaussiana se vale

$$p(X = x) \propto \exp\left(ax^2 + bx\right), \quad \text{per ogni } x \in \mathbb{R},$$

per degli opportuni parametri a, $b \in \mathbb{R}$.

- la densità è l'esponenziale di un polinomio di secondo grado dei possibili valori $x \in \mathbb{R}$.
- dovendo essere $\int_{-\infty}^{\infty} p(X = x)dx < \infty$, allora $a \in \mathbb{R}$ necessariamente deve essere $a < 0$

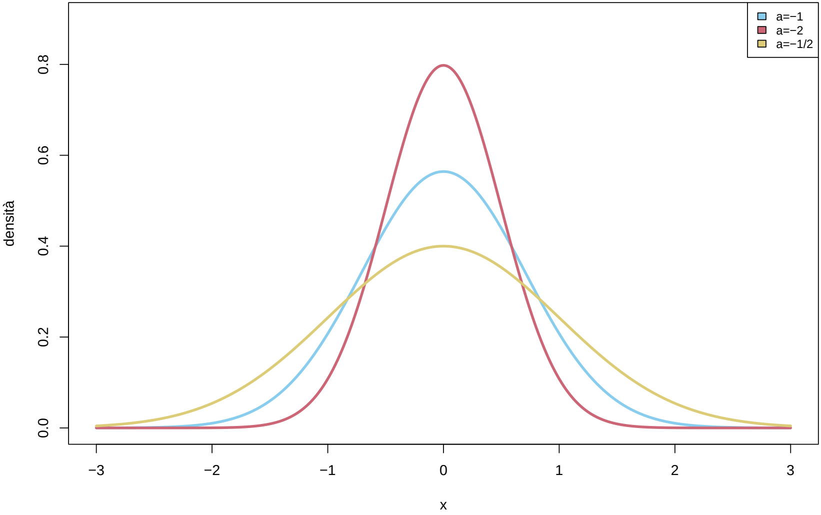

Figure 1: densità gaussiana al variare del parametro $a < 0$, $b = 0$

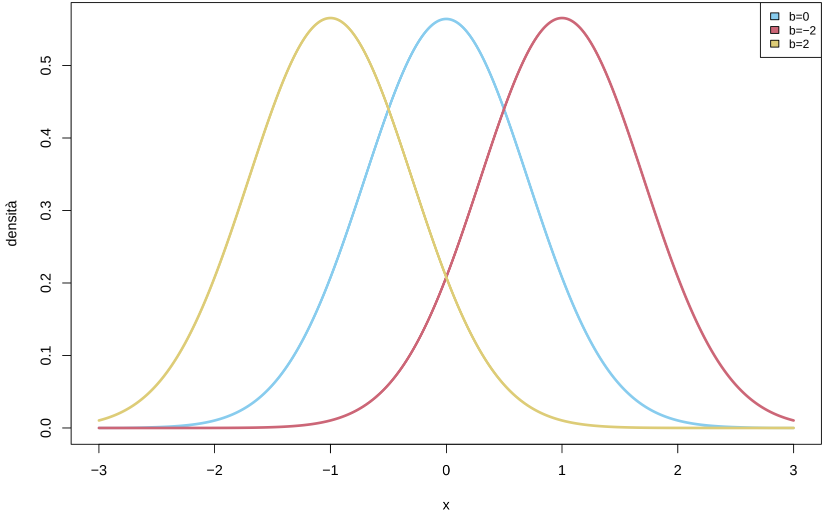

Figure 2: densità gaussiana al variare del parametro b, $a = 1$

Intepretazione dei parametri

Sia X una variabile con densità gaussiana

$$p(X = x) \propto \exp\left(ax^2 + bx\right).$$

Allora vale

$$a = -\frac{1}{2\sigma_X^2}, \quad b = \frac{\mathbb{E}[X]}{\sigma_X^2},$$

ossia

$$\mathrm{Var}(X) = \sigma_X^2 = -\frac{1}{2a} \quad \mathbb{E}[X] = -\frac{b}{2a}.$$

Dimostrazione

Definizione usuale

Si dice che $X \in \mathbb{R}$ ha densità continua gaussiana di valor medio $m \in \mathbb{R}$ e varianza $\sigma^2 > 0$, e si scrive brevemente $\mathcal{N}(m, \sigma^2)$, se

$$p(X = x) \propto \exp\left(-\frac{1}{2}\frac{(x-m)^2}{\sigma^2}\right).$$

- Più esplicitamente, si può mostrare che vale l'identità

$$p(X = x) = \exp\left(-\frac{1}{2}\frac{(x-m)^2}{\sigma^2}\right)\frac{1}{\sqrt{2\pi\sigma^2}}.$$

Definizione usuale

Si dice che $X \in \mathbb{R}$ ha densità continua gaussiana di valor medio $m \in \mathbb{R}$ e varianza $\sigma^2 > 0$, e si scrive brevemente $\mathcal{N}(m, \sigma^2)$, se

$$p(X = x) \propto \exp\left(-\frac{1}{2}\frac{(x-m)^2}{\sigma^2}\right).$$

- Più esplicitamente, si può mostrare che vale l'identità

$$p(X = x) = \exp\left(-\frac{1}{2}\frac{(x-m)^2}{\sigma^2}\right)\frac{1}{\sqrt{2\pi\sigma^2}}.$$

- La costante $1/\sqrt{2\pi}$ è interessante da calcolare analiticamente, ma non troppo utile nelle applicazioni.

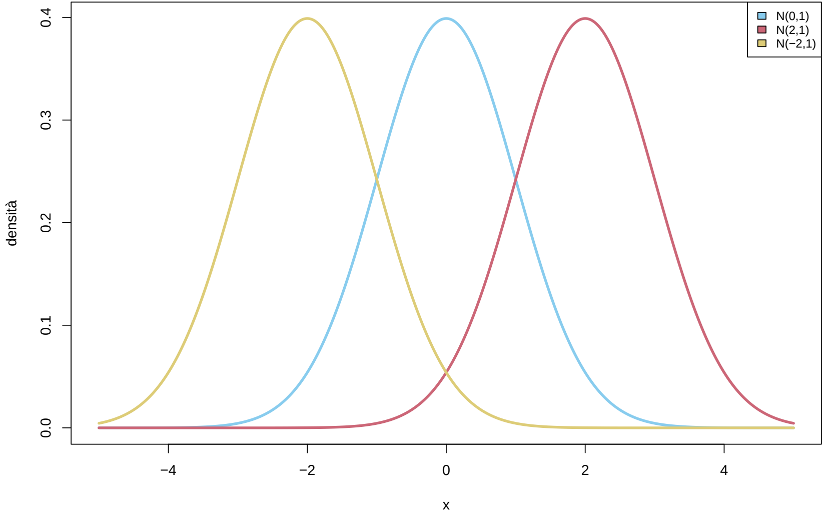

Figure 3: densità gaussiana al variare del parametro m (con $\sigma = 1$ costante

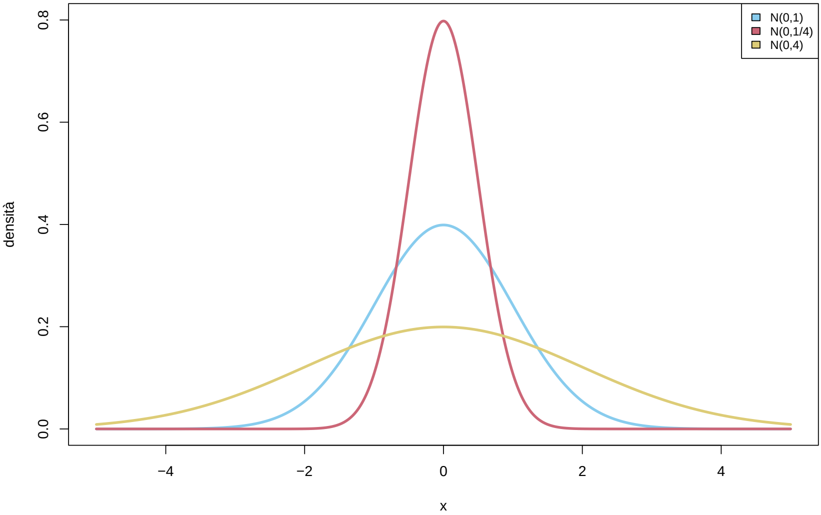

Figure 4: densità gaussiana al variare del parametro σ (con $m = 0$ costante

Proprietà di massima entropia

Al variare di tutte le possibili densità continue per una variabile X, $p(X = x)$, con $x \in \mathbb{R}$, tali che il valor medio e la varianza di X siano fissati

$$\mathbb{E}\left[X\right] = \int_{-\infty}^{\infty} xP(X = x)dx = m, \quad \mathrm{Var}\left(X\right) = \int_{-\infty}^{\infty} (x - m)^2 P(X = x)dx =$$

la densità gaussiana $\mathcal{N}(m, \sigma^2)$ è quella di **massima entropia**.

- Pertanto, seguento principio di massima entropia, avendo a disposizione come informazione su una variabile aleatoria (reale) solamente il suo valor medio m e la varianza σ^2, sia imporrà che sia una densità gaussiana $\mathcal{N}(m, \sigma^2)$.

Trasformazione affine

Sia X una variabile con densità continua $\mathcal{N}(m, \sigma^2)$ e siano $\lambda \neq 0$, $c \in \mathbb{R}$. Allora la variabile $Y = \lambda X + c$ ha densità continua gaussiana, di parametri $\mathcal{N}(\lambda m + c, \lambda^2 \sigma^2)$.

- se X ha densità gaussiana $\mathcal{N}(m, \sigma^2)$, la sua standardizzata

  $$\frac{X - m}{\sigma} \quad \text{ha densità continua } \mathcal{N}(0, 1),$$

  pertanto detta anche densità **gaussiana standard**, che ha densità

  $$\exp\left(-\frac{1}{2}x^2\right)\frac{1}{\sqrt{2\pi}} \quad \text{per } x \in \mathbb{R}.$$

Trasformazione affine

Sia X una variabile con densità continua $\mathcal{N}(m, \sigma^2)$ e siano $\lambda \neq 0$, $c \in \mathbb{R}$. Allora la variabile $Y = \lambda X + c$ ha densità continua gaussiana, di parametri $\mathcal{N}(\lambda m + c, \lambda^2 \sigma^2)$.

- se X ha densità gaussiana $\mathcal{N}(m, \sigma^2)$, la sua standardizzata
$$\frac{X - m}{\sigma} \quad \text{ha densità continua } \mathcal{N}(0, 1),$$
pertanto detta anche densità **gaussiana standard**, che ha densità
$$\exp\left(-\frac{1}{2}x^2\right)\frac{1}{\sqrt{2\pi}} \quad \text{per } x \in \mathbb{R}.$$
- Se $\lambda = 0$, la variabile $\lambda X + c = c$ è costante. Per uniformare le notazioni, si conviene di considerare anche le variabili costanti come caso *degenere* di una densità gaussiana.

Dimostrazione

CDF

La funzione di ripartizione gaussiana (anche nel caso standard) non è esprimibile in termini di funzioni elementari.

- Il comando R per ottenerne i valori è `pnorm()`.

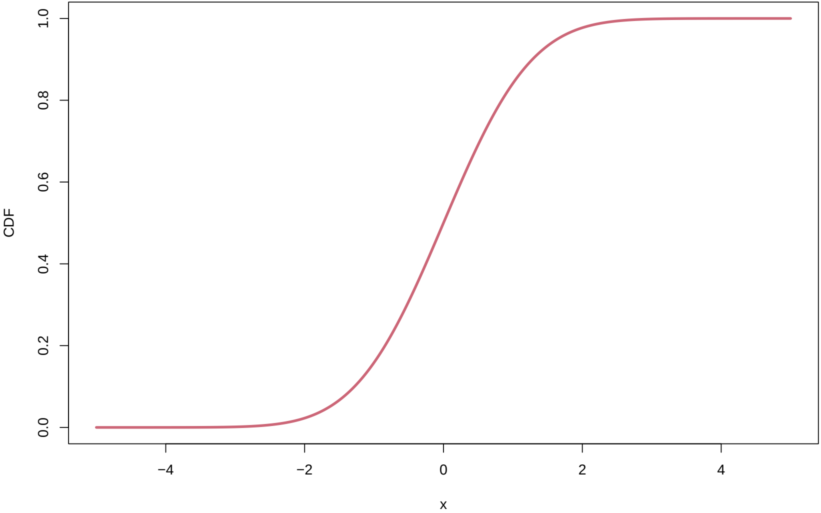

Figure 5: CDF di una variabile gaussiana standard

MGF e funzione caratteristica

Sia X una variabile con densità continua $\mathcal{N}(m, \sigma^2)$. Allora

$$\mathrm{MGF}_X(t) = \exp\left(mt + \frac{\sigma^2}{2}t^2\right),$$

e

$$\varphi_X(\xi) = \exp\left(im\xi - \frac{\sigma^2}{2}\xi^2\right).$$

Dimostrazione

Problema 1

L'orario d'arrivo a lezione degli studenti di ingegneria robotica segue approssimativamente una distribuzione gaussiana di media 8:25 e deviazione standard 5 minuti. Preso uno studente a caso,

1. calcolare la probabilità che arrivi dopo l'inizio delle lezioni (8:30);
2. calcolare il ritardo medio (in minuti).

Esprimere eventualmente i risultati come opportuni integrali o indicare un comando R per il calcolo numerico.

Problema 2

L'altezza degli studenti (maschi) del corso di ingegneria è rappresentata da una distribuzione gaussiana di media 175 cm e deviazione standard 10 cm. L'altezza delle studentesse (femmine) è pure una gaussiana di media 160 cm con deviazione standard 10 cm. Preso uno/a studente a caso, si osserva che è alto/a 165 cm. Dire se è più probabile che sia maschio o femmina,

1. senza conoscere la percentuale di studenti maschi e femmina nel corso;
2. sapendo anche che i maschi rappresentano il 70% degli studenti di ingegneria e il 30% è femmina (solo per semplicità di calcolo escludiamo le persone non identificate in uno dei due generi).