Variabili Aleatorie Continue 1

Stefania Bartoletti

19 Aprile 2021

Indice

Variabili aleatorie continue

- Tutte le variabili aleatorie assegnano un numero ad ogni possibile esito di un esperimento in uno spazio degli esiti.

Variabili aleatorie continue

- Tutte le variabili aleatorie assegnano un numero ad ogni possibile esito di un esperimento in uno spazio degli esiti.
- Mentre le variabili aleatorie discrete assumono valori contenuti in un insieme discreto, le variabili aleatorie continue assumono valori in un insieme continuo.

Variabili aleatorie continue

- Tutte le variabili aleatorie assegnano un numero ad ogni possibile esito di un esperimento in uno spazio degli esiti.
- Mentre le variabili aleatorie discrete assumono valori contenuti in un insieme discreto, le variabili aleatorie continue assumono valori in un insieme continuo.

Variabili aleatorie continue

- Tutte le variabili aleatorie assegnano un numero ad ogni possibile esito di un esperimento in uno spazio degli esiti.
- Mentre le variabili aleatorie discrete assumono valori contenuti in un insieme discreto, le variabili aleatorie continue assumono valori in un insieme continuo.

Esempio Ammettiamo di misurare con quanti minuti di ritardo Giovanni partecipa alla sessione di Probabilità e Statistica ogni giorno. Il risultato del nostro esperimento è quindi un tempo in minuti. Siccome Giovanni arriva in ritardo di 3.4 minuti o in anticipo di 2.7 minuti, quindi con ritardo -2.7 minuti, lo spazio degli eventi consiste in tutti i numeri reali. La variabile aleatoria che descrive il risultato dell'esperimento assume valori su un intervallo continuo.

Variabili Aleatorie Continue

Una variabile aleatoria continua X assume valori all'interno di un intervallo $[a, b]$, che può essere finito o infinito, e.g. $[0, 1]$, $[-2, 1.4]$,$[-\infty, +\infty]$

Variabili Aleatorie Continue

Una variabile aleatoria continua X assume valori all'interno di un intervallo $[a, b]$, che può essere finito o infinito, e.g. $[0, 1]$, $[-2, 1.4]$,$[-\infty, +\infty]$

Definition

Una variabile aleatoria X è continua se esiste una funzione $f_{\mathsf{X}}(x)$ tale che per ogni $c \leq d$, si ha

$$\mathbb{P}\left\{c \leq \mathsf{X} \leq d\right\} = \int_c^d f_{\mathsf{X}}(x)\, dx$$

e la funzione $f_{\mathsf{X}}(x)$ si chiama funzione di densità di probabilità.

Funzione di densità di probabilità

La funzione di densità di probabilità $f_{\mathsf{X}}(x)$ di una variabile aleatoria continua X soddisfa sempre le seguenti proprietà:

- $f_{\mathsf{X}}(x) \geq 0$
- $\int_{-\infty}^{+\infty} f_{\mathsf{X}}(x) = 1$, ovvero $\mathbb{P}\{-\infty \leq \mathsf{X} \leq \infty\} = 1$

Funzione di densità di probabilità

La funzione di densità di probabilità $f_{\mathsf{X}}(x)$ di una variabile aleatoria continua X è l'analogo della funzione di massa di probabilità per una variabile aleatoria discreta. Esistono però due importantissime differenze:

Funzione di densità di probabilità

La funzione di densità di probabilità $f_{\mathsf{X}}(x)$ di una variabile aleatoria continua X è l'analogo della funzione di massa di probabilità per una variabile aleatoria discreta. Esistono però due importantissime differenze:

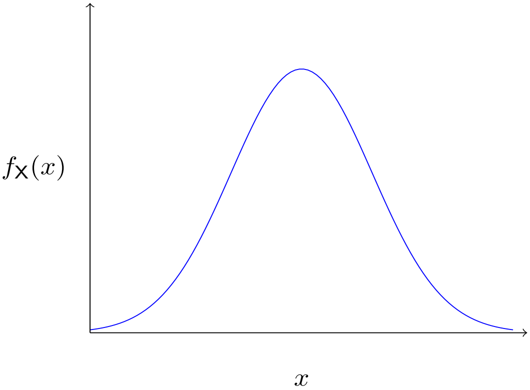
- A differenza della pmf, la pdf **non** è una probabilità, è necessario integrarla per ottenere un valore di probabilità

Funzione di densità di probabilità

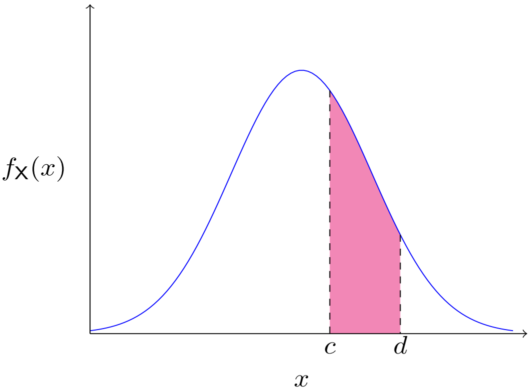
La funzione di densità di probabilità $f_X(x)$ di una variabile aleatoria continua X è l'analogo della funzione di massa di probabilità per una variabile aleatoria discreta. Esistono però due importantissime differenze:

- A differenza della pmf, la pdf **non** è una probabilità, è necessario integrarla per ottenere un valore di probabilità
- Dal momento che non è una probabilità, non vi è una restrizione sui valori di $f_X(x)$ quando X è continua, e pertanto $f_X(x)$ può assumere valori anche maggiori di 1

Interpretazione Grafica



Interpretazione Grafica



L'area totale sottesa dalla funzione di densità di massa tra $-\infty$ e $+\infty$ è pari a 1.

Esempio

Consideriamo una variabile aleatoria continua X con pdf $f_{\mathsf{X}}(x) = 3$ in $[0, 1/3]$ (ovvero $f_{\mathsf{X}}(x) = 0$ fuori dall'intervallo $[0, 1/3]$). Proviamo a disegnare la pdf e a calcolare $\mathbb{P}\{0.1 \leq \mathsf{X} \leq 0.2\}$ e $\mathbb{P}\{0.1 \leq \mathsf{X} \leq 1\}$.

Esempio

Sia X una variabile aleatoria continua nell'intervallo $[0, 1]$ con pdf $f_X(x) = Cx^2$. Qual è il valore di C?

Esempio

Sia X una variabile aleatoria continua nell'intervallo $[0, 1]$ con pdf $f_{\mathsf{X}}(x) = Cx^2$. Qual è il valore di C?
Dal momento che la probabilità su tutto l'intervallo deve essere 1,

$$\int_0^1 f_{\mathsf{X}}(x) = \int_0^1 Cx^2 = \left[\frac{Cx^3}{3}\right]_0^1 = \frac{C}{3} = 1$$

Pertanto $C = 3$. L'uso di questo tipo di costanti è molto diffuso, sono chiamate costanti di normalizzazione, in quanto normalizzano la pdf in maniera tale che la probabilità calcolata su tutto l'intervallo sia 1.

Esempio

Considerando sempre la stessa variabile aleatoria, con $f_{\mathsf{X}}(x) = 3x^2$ su $[0, 1]$, si calcoli $\mathbb{P}\{\mathsf{X} \leq 1/2\}$

$$\mathbb{P}\{\mathsf{X} \leq 1/2\} = \int_0^{1/2} 3x^2 = 1/8$$

Ma allora ...

Data una variabile aleatoria continua X:

- Quanto vale $\mathbb{P}\{a \leq \mathsf{X} \leq a\}$?

Ma allora ...

Data una variabile aleatoria continua X:

- Quanto vale $\mathbb{P}\{a \leq \mathsf{X} \leq a\}$?
- Quanto vale $\mathbb{P}\{\mathsf{X} = 0\}$?

Ma allora ...

Data una variabile aleatoria continua X:

- Quanto vale $\mathbb{P}\left\{a \leq \mathsf{X} \leq a\right\}$?
- Quanto vale $\mathbb{P}\left\{\mathsf{X} = 0\right\}$?
- Se $\mathbb{P}\left\{\mathsf{X} = a\right\} = 0$ significa che X non può valere a?

Ma allora ...

Data una variabile aleatoria continua X:

- Quanto vale $\mathbb{P}\{a \leq \mathsf{X} \leq a\}$?
- Quanto vale $\mathbb{P}\{\mathsf{X} = 0\}$?
- Se $\mathbb{P}\{\mathsf{X} = a\} = 0$ significa che X non può valere a?

Ma allora ...

Data una variabile aleatoria continua X:

- Quanto vale $\mathbb{P}\{a \leq \mathsf{X} \leq a\}$?
- Quanto vale $\mathbb{P}\{\mathsf{X} = 0\}$?
- Se $\mathbb{P}\{\mathsf{X} = a\} = 0$ significa che X non può valere a?

Pensiamo ad un esempio... quanto vale la probabilità che una persona sia alta esattamente $1.65\,\mathrm{m}$?

Calcolo

- Rispetto alle definizioni viste fino ad ora, dal punto di vista del calcolo, passare da variabili discrete a variabili continue richiede solo un passaggio da somme a integrali

Calcolo

- Rispetto alle definizioni viste fino ad ora, dal punto di vista del calcolo, passare da variabili discrete a variabili continue richiede solo un passaggio da somme a integrali
- Ovvero, partendo dall'interpretazione dell'integrale come area sottesa da una curva che rappresenta la nostra funzione integranda, o dall'approssimazione di tale area di rettangoli di dimensione dx.

Funzione di distribuzione cumulativa

La funzione di distribuzione cumulativa (cdf) di una variabile aleatoria continua X è definita esattamente come nel caso delle variabili aleatorie discrete, ovvero

$$F_{\mathsf{X}}(b) = \mathbb{P}\left\{\mathsf{X} \leq b\right\}$$

Funzione di distribuzione cumulativa

La funzione di distribuzione cumulativa (cdf) di una variabile aleatoria continua X è definita esattamente come nel caso delle variabili aleatorie discrete, ovvero

$$F_{\mathsf{X}}(b) = \mathbb{P}\left\{\mathsf{X} \leq b\right\}$$

Questa volta, sì, si tratta di una funzione di probabilità. La si può calcolare, come nel caso discreto, a partire dalla pdf, come

$$F_{\mathsf{X}}(b) = \mathbb{P}\left\{\mathsf{X} \leq b\right\} = \int_{-\infty}^{b} f_{\mathsf{X}}(x)\, dx$$

Funzione di distribuzione cumulativa

La funzione di distribuzione cumulativa (cdf) di una variabile aleatoria continua X è definita esattamente come nel caso delle variabili aleatorie discrete, ovvero

$$F_{\mathsf{X}}(b) = \mathbb{P}\left\{\mathsf{X} \le b\right\}$$

Questa volta, sì, si tratta di una funzione di probabilità. La si può calcolare, come nel caso discreto, a partire dalla pdf, come

$$F_{\mathsf{X}}(b) = \mathbb{P}\left\{\mathsf{X} \le b\right\} = \int_{-\infty}^{b} f_{\mathsf{X}}(x)\, dx$$

Nella pratica, si dice che "X ha distribuzione $F_{\mathsf{X}}(x)$" oppure "X ha cdf $F_{\mathsf{X}}(x)$"

Esempio 1

Si cerchi la cdf per la variabile aleatoria vista nell'esempio con pdf $f_{\mathsf{X}}(x) = 3$ nell'intervallo $[0, 1/3]$.

- Se $a \in [0, 1/3]$, abbiamo $F_{\mathsf{X}}(a) = \int_0^a f(x)\, dx = \int_0^a 3\, dx = 3a$

Esempio 1

Si cerchi la cdf per la variabile aleatoria vista nell'esempio con pdf $f_{\mathsf{X}}(x) = 3$ nell'intervallo $[0, 1/3]$.

- Se $a \in [0, 1/3]$, abbiamo $F_{\mathsf{X}}(a) = \int_0^a f(x)\, dx = \int_0^a 3\, dx = 3a$
- Se $a \notin [0, 1/3]$, essendo $f_{\mathsf{X}}(x) = 0$ fuori da $[0, 1/3]$, allora $F(a) = \mathbb{P}\{\mathsf{X} \leq a\} = 0$ per $a < 0$ e $F(a) = 1$ per $a > 1/3$

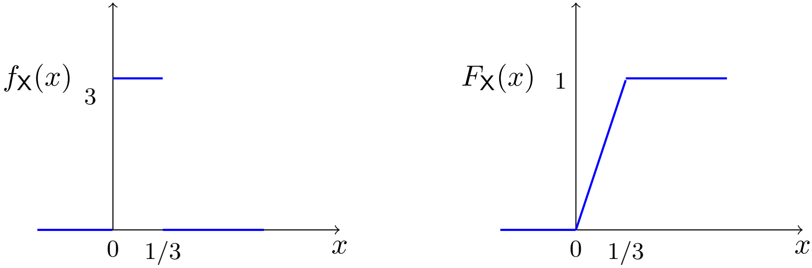
Esempio 1

Si cerchi la cdf per la variabile aleatoria vista nell'esempio con pdf $f_{\mathsf{X}}(x) = 3$ nell'intervallo $[0, 1/3]$.

- Se $a \in [0, 1/3]$, abbiamo $F_{\mathsf{X}}(a) = \int_0^a f(x)\, dx = \int_0^a 3\, dx = 3a$
- Se $a \notin [0, 1/3]$, essendo $f_{\mathsf{X}}(x) = 0$ fuori da $[0, 1/3]$, allora $F(a) = \mathbb{P}\{\mathsf{X} \leq a\} = 0$ per $a < 0$ e $F(a) = 1$ per $a > 1/3$
- Pertanto

$$F(x) = \begin{cases} 0 & \text{se } x < 0 \\ 3x & \text{se } 0 \leq x \leq 1/3 \\ 1 & \text{se } 1/3 < x \end{cases}$$

Esempio 1



Esempio 2

Si cerchi la cdf per la variabile aleatoria vista nell'esempio con pdf $f_{\mathsf{X}}(x) = 3x^2$ nell'intervallo $[0,1]$. Calcolare $\mathbb{P}\{\mathsf{X} \leq 1/2\}$

- Sappiamo che $f_{\mathsf{X}}(x) = 3x^2$ sull'intervallo $[0,1]$.

Esempio 2

Si cerchi la cdf per la variabile aleatoria vista nell'esempio con pdf $f_{\mathsf{X}}(x) = 3x^2$ nell'intervallo $[0, 1]$. Calcolare $\mathbb{P}\{\mathsf{X} \leq 1/2\}$

- Sappiamo che $f_{\mathsf{X}}(x) = 3x^2$ sull'intervallo $[0, 1]$.
- Pertanto $F_{\mathsf{X}}(a) = \int_0^a f(x)\,dx = \int_0^a 3x^2\,dx = a^3$ su $[0, 1]$

Esempio 2

Si cerchi la cdf per la variabile aleatoria vista nell'esempio con pdf $f_{\mathsf{X}}(x) = 3x^2$ nell'intervallo $[0, 1]$. Calcolare $\mathbb{P}\{\mathsf{X} \leq 1/2\}$

- Sappiamo che $f_{\mathsf{X}}(x) = 3x^2$ sull'intervallo $[0, 1]$.
- Pertanto $F_{\mathsf{X}}(a) = \int_0^a f(x)\, dx = \int_0^a 3x^2\, dx = a^3$ su $[0, 1]$
- Segue che

$$F_{\mathsf{X}}(x) = \begin{cases} 0 & \text{se } x < 0 \\ x^3 & \text{se } 0 \leq x \leq 1 \\ 1 & \text{se } 1 < x \end{cases}$$

Esempio 2

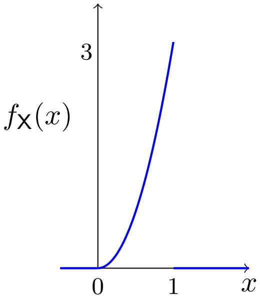
Si cerchi la cdf per la variabile aleatoria vista nell'esempio con pdf $f_X(x) = 3x^2$ nell'intervallo $[0, 1]$. Calcolare $\mathbb{P}\{X \leq 1/2\}$

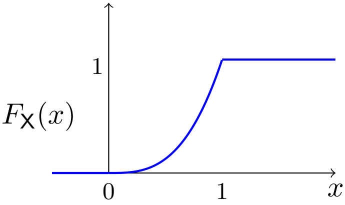
- Sappiamo che $f_X(x) = 3x^2$ sull'intervallo $[0, 1]$.
- Pertanto $F_X(a) = \int_0^a f(x)\, dx = \int_0^a 3x^2\, dx = a^3$ su $[0, 1]$
- Segue che

$$F_X(x) = \begin{cases} 0 & \text{se } x < 0 \\ x^3 & \text{se } 0 \leq x \leq 1 \\ 1 & \text{se } 1 < x \end{cases}$$

- Dunque $\mathbb{P}\{X \leq 1/2\} = F_X(1/2) = 1/8$

Esempio 2

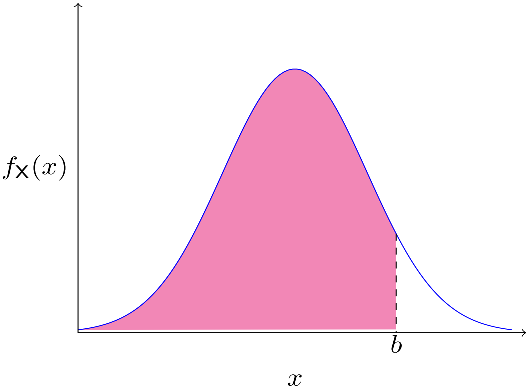




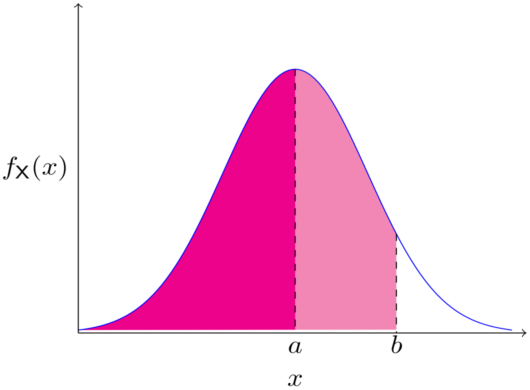
Proprietà della funzione di distribuzione cumulativa

- $0 \le F_{\mathsf{X}}(x) \le 1$
- $F_{\mathsf{X}}(x)$ è non-decrescente, ovvero se $a \le b$, allora $F_{\mathsf{X}}(a) \le F_{\mathsf{X}}(b)$
- $\lim_{x \to \infty} F_{\mathsf{X}}(x) = 1$ and $\lim_{x \to -\infty} F_{\mathsf{X}}(x) = 0$
- $\mathbb{P}\{a \le \mathsf{X} \le b\} = F(b) - F(a)$
- $F'_{\mathsf{X}}(x) = f_{\mathsf{X}}(x)$

$\mathbb{P}\{a \leq \mathsf{X} \leq b\} = F(b) - F(a)$



$$\mathbb{P}\{a \leq \mathsf{X} \leq b\} = F(b) - F(a)$$



Variabile Aleatoria Uniforme

Una variabile aleatoria è detta uniforme (o uniformemente distribuita) sull'intervallo $[0, 1]$ se la sua densità è data da

$$f_{\mathsf{X}}(x) = \begin{cases} 1 & \forall x \in [0, 1] \\ 0 & \text{altrimenti} \end{cases}$$

Variabile Aleatoria Uniforme

Una variabile aleatoria è detta uniforme (o uniformemente distribuita) sull'intervallo $[0,1]$ se la sua densità è data da

$$f_{\mathsf{X}}(x) = \begin{cases} 1 & \forall x \in [0,1] \\ 0 & \text{altrimenti} \end{cases}$$

Si osservi che la definizione rispetta le proprietà di densità, dato che $f_{\mathsf{X}}(x) \geq 0$ e $\int_0^1 f_{\mathsf{X}}(x)\, dx = 1$.

Variabile Aleatoria Uniforme

Una variabile aleatoria è detta uniforme (o uniformemente distribuita) sull'intervallo $[0, 1]$ se la sua densità è data da

$$f_X(x) = \begin{cases} 1 & \forall x \in [0, 1] \\ 0 & \text{altrimenti} \end{cases}$$

Si osservi che la definizione rispetta le proprietà di densità, dato che $f_X(x) \geq 0$ e $\int_0^1 f_X(x)\, dx = 1$.

Nel caso specifico, X assume necessariamente almeno un valore in $[0, 1]$. Inoltre, essendo $f_X(x)$ costante in $[0, 1]$, X sarà con la stessa probabilità vicino a un qualunque valore in $[0, 1]$. Per ogni $0 \leq c < d \leq 1$ abbiamo

$$\mathbb{P}\{c \leq X \leq d\} = \int_c^d f_X(x)\, dx = d - c$$

Variabile Aleatoria Uniforme

Nel caso generale, diciamo che X è una variabile aleatoria uniforme sull'intervallo $[a, b]$, e lo indichiamo con X $\sim \mathcal{U}([a, b])$, se la sua densità è data da

$$f_{\mathsf{X}}(x) = \begin{cases} \frac{1}{b-a} & \forall x \in [a, b] \\ 0 & \text{altrimenti} \end{cases}$$

Variabile Aleatoria Uniforme

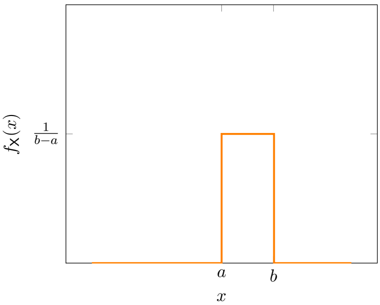
Nel caso generale, diciamo che X è una variabile aleatoria uniforme sull'intervallo $[a, b]$, e lo indichiamo con $\mathsf{X} \sim \mathcal{U}([a, b])$, se la sua densità è data da

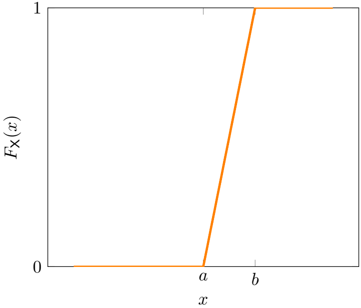
$$f_{\mathsf{X}}(x) = \begin{cases} \frac{1}{b-a} & \forall x \in [a, b] \\ 0 & \text{altrimenti} \end{cases}$$

Essendo $F_{\mathsf{X}}(c) = \int_{-\infty}^{c} f_{\mathsf{X}}(x)\, dx$, si ottiene dalla che la sua funzione di distribuzione è data da

$$F_{\mathsf{X}}(x) = \begin{cases} 0 & x < a \\ \frac{x-a}{b-a} & a \leq x \leq b \\ 1 & x > b \end{cases}$$

Grafico della pdf e cdf





Esempio 1

- Gli autobus passano a una specifica fermata a intervalli di 15 minuti a partire dalle 7; cioè alle 7, $7:15$, $7:30$, $7:45$ ecc.

Esempio 1

- Gli autobus passano a una specifica fermata a intervalli di 15 minuti a partire dalle 7; cioè alle 7, $7:15$, $7:30$, $7:45$ ecc.
- Se un passeggero arriva alla fermata in un istante uniformemente distribuito tra le 7 e le $7:30$, determinare la probabilità che egli aspetti l'autobus meno di 5 minuti oppure più di 10 minuti.

Esempio 1

Sia X il minuto dopo le 7 al quale arriva il passeggero. Dato che X è uniforme sull'intervallo $[0, 30]$, si ha che il passeggero aspetterà meno di 5 minuti se (e solo se) egli arriva tra le $7:10$ e le $7:15$ o tra le $7:25$ e le $7:30$. Pertanto la probabilità che egli aspetti l'autobus meno di 5 minuti è

$$\mathbb{P}\{10 \leq \mathsf{X} \leq 15\} + \mathbb{P}\{25 \leq \mathsf{X} \leq 30\} = \int_{10}^{15} \frac{1}{30}\,dx + \int_{25}^{30} \frac{1}{30}\,dx = 1/3$$

Esempio 1

Sia X il minuto dopo le 7 al quale arriva il passeggero. Dato che X è uniforme sull'intervallo $[0, 30]$, si ha che il passeggero aspetterà meno di 5 minuti se (e solo se) egli arriva tra le $7:10$ e le $7:15$ o tra le $7:25$ e le $7:30$. Pertanto la probabilità che egli aspetti l'autobus meno di 5 minuti è

$$\mathbb{P}\{10 \leq \mathsf{X} \leq 15\} + \mathbb{P}\{25 \leq \mathsf{X} \leq 30\} = \int_{10}^{15} \frac{1}{30}\, dx + \int_{25}^{30} \frac{1}{30}\, dx = 1/3$$

Analogamente, egli dovrà aspettare più di 10 minuti se arriva tra le 7 e le $7:05$ o tra le $7:15$ e le $7:20$, la probabilità che egli aspetti l'autobus più di 10 minuti è data da

$$\mathbb{P}\{0 \leq \mathsf{X} \leq 5\} + \mathbb{P}\{15 \leq \mathsf{X} \leq 20\} = \int_{0}^{5} \frac{1}{30}\, dx + \int_{15}^{20} \frac{1}{30}\, dx = 1/3$$

Distribuzione Gaussiana

Nel 1809, Carl Friedrich Gauss pubblicò una monografia *"Theoria motus corporum coelestium in sectionibus conicis solem ambientium"* (Teoria del moto dei corpi celesti che si muovono lungo sezioni coniche intorno al Sole) in cui presentava diverse nozioni che diventarono fondamentali per la statistica.

Distribuzione Gaussiana

Nel 1809, Carl Friedrich Gauss pubblicò una monografia *"Theoria motus corporum coelestium in sectionibus conicis solem ambientium"* (Teoria del moto dei corpi celesti che si muovono lungo sezioni coniche intorno al Sole) in cui presentava diverse nozioni che diventarono fondamentali per la statistica.

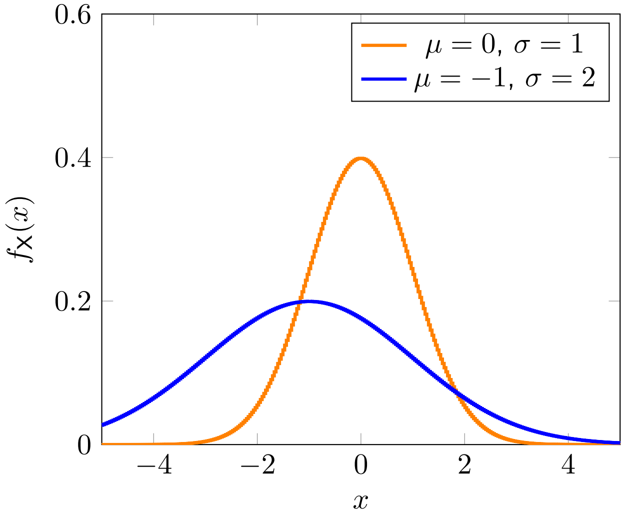
Tra queste, vi era la definizione della distribuzione normale, che pertanto si chiama anche Gaussiana e che rappresenta la più importante distribuzione tra le variabili aleatorie continue.

Distribuzione Gaussiana

Una variabile aleatoria X è detta normale (o Gaussiana) di parametri μ e σ^2, e si indica con $\mathsf{X} \sim \mathcal{N}(\mu, \sigma^2)$ se la sua densità è data da

$$f_{\mathsf{X}}(x) = \frac{1}{\sqrt{2\,\pi}\sigma} \exp - \frac{(x-\mu)^2}{2\sigma^2}$$

Distribuzione Gaussiana



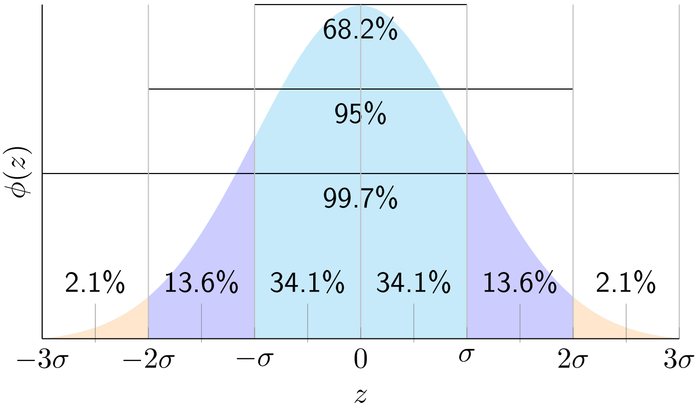
Normale Standard

Una variabile aleatoria normale $Z \sim \mathcal{N}(0,1)$ è detta normale standard. Essendo una variabile speciale, definiremo un simbolo speciale per descrivere la sua densità di probabilità

$$\phi(z) = f_Z(z) = \frac{1}{\sqrt{2\pi}} \exp \frac{-z^2}{2}$$

Regola Empirica

Si dimostra che $\mathbb{P}\{-1 \leq Z \leq 1\} \simeq 68\%$, $\mathbb{P}\{-2 \leq Z \leq 2\} \simeq 95\%$, e $\mathbb{P}\{-3 \leq Z \leq 3\} \simeq 99\%$



Esercizio

Calcolare $F_{\mathsf{Z}}(z)$...