

VARIABILI ALEATORIE CONTINUE

Se X è una variabile aleatoria continua, la probabilità che X assuma un certo valore x fissato è in generale zero, quindi non ha senso definire una distribuzione di probabilità con lo stesso procedimento seguito per una variabile aleatoria discreta. Nel caso di una variabile aleatoria continua ha senso invece calcolare la probabilità che X sia compresa fra a e b , dove a e b sono costanti, con $a \leq b$.

Esempio: Un nucleo emette un neutrone, il tempo necessario per l'emissione può assumere qualunque valore reale positivo. La probabilità che questa avvenga esattamente dopo tre secondi è nulla, ha senso invece assegnare una probabilità per una emissione tra 2,9 e 3,1 secondi.

VARIABILI ALEATORIE CONTINUE

Le variabili aleatorie continue X assumono **valori in un intervallo di numeri reali** (es. $[0, \infty)$, $[0.4, \infty)$, $(-\infty, \infty)$, $[0,1], \dots$).

Una variabile aleatoria continua è descritta da una funzione $\rho(\mathbf{x})$ detta **densità di probabilità** che assegna ad ogni numero reale x un valore

$$\rho(\mathbf{x}) \geq 0 \quad \forall \mathbf{x} \in \mathbb{R}$$

e che soddisfa la condizione

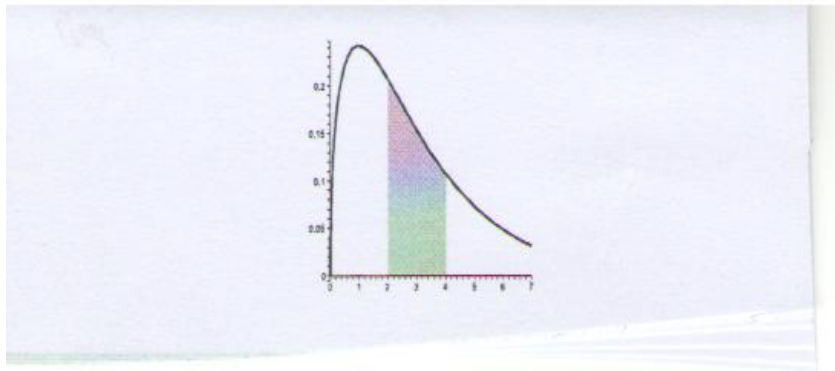
$$\int_{-\infty}^{\infty} \rho(\mathbf{x}) d\mathbf{x} = 1$$

Si definisce poi la probabilità che X sia compresa fra a e b come

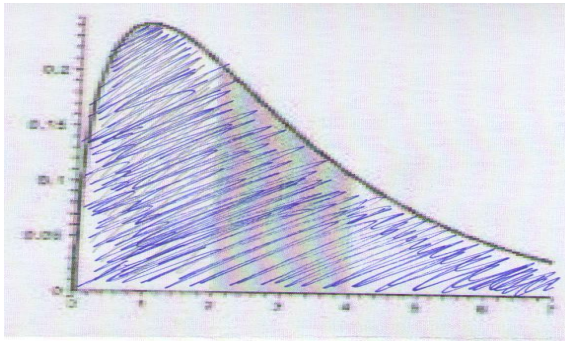
$$\mathbf{P(a \leq X \leq b)} = \int_a^b \rho(\mathbf{x}) d\mathbf{x} \geq 0$$

Si può dimostrare che questa definizione soddisfa gli assiomi della teoria della probabilità .

$P(a \leq X \leq b) = \text{area della regione delimitata dall'asse } x, \text{ il grafico di } \rho(x) \text{ e le rette } x=a \text{ e } x=b$



L'area sotto il grafico tra due punti a e b non varia se gli estremi sono inclusi od esclusi: $P(a \leq X \leq b) = P(a < X < b)$.



La condizione $\int_{-\infty}^{\infty} \rho(x) dx = 1$ significa che **l'area** della **regione** del piano **contenuta tra l'asse delle x e il grafico di ρ sia uguale ad 1** ed è coerente con il fatto che $\int_{-\infty}^{\infty} \rho(x) dx = P(-\infty < X < +\infty)$.

Esercizio Un punto viene collocato sul segmento $[0,1]$ con probabilità uniforme. Siamo interessati a conoscere il valore x dell'ordinata corrispondente alla posizione del punto. Uniformità significa che questo numero assume tutti i valori reali in $[0,1]$ con uguale densità di probabilità .

Sia α una costante e sia data la funzione:

$$\rho(x) = \alpha \quad 0 \leq x \leq 1$$

$$\rho(x) = 0 \quad \text{altrimenti}$$

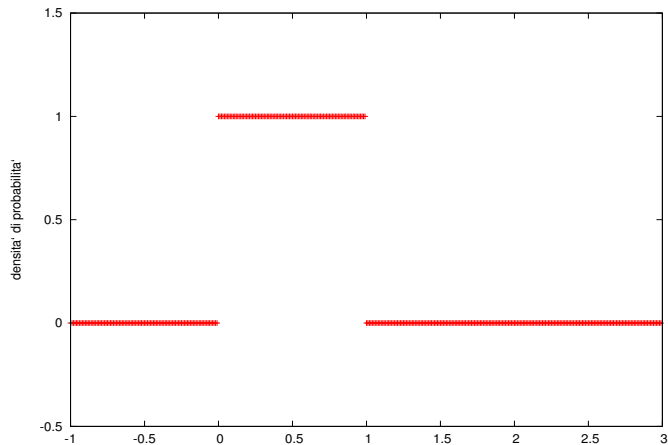
Determinare il valore di α per il quale è una densità di probabilità .

La condizione $\rho(x) \geq 0, \quad \forall x \in \mathbb{R}$ è soddisfatta per $\alpha \geq 0$.

Mentre

$$\int_{-\infty}^{\infty} \rho(x) dx = \alpha \int_0^1 dx = \alpha [x]_0^1 = \alpha$$

è uguale a 1 per $\alpha = 1$.



La densità di probabilità $\rho(x)$ è uguale a 1 se $0 \leq x \leq 1$ ed è uguale a 0 altrimenti.

Esercizio: Si consideri quindi $\alpha = 1$ e si assuma che $0 \leq a \leq b \leq 1$, si calcoli $P(a \leq X \leq b)$:

$$P(a \leq X \leq b) = \int_a^b \rho(x) dx = \int_a^b dx = [x]_a^b = b - a.$$

Si assuma ora $a \leq 0 \leq b \leq 1$, si calcoli $P(a \leq X \leq b)$:

$$P(a \leq X \leq b) = \int_a^b \rho(x) dx = \int_0^b dx = [x]_0^b = b.$$

Si assuma infine $a \leq 0 \leq 1 \leq b$, si calcoli $P(a \leq X \leq b)$:

$$P(a \leq X \leq b) = \int_a^b \rho(x) dx = \int_0^1 dx = [x]_0^1 = 1.$$

FUNZIONE DISTRIBUZIONE

Definizione. La funzione

$$F(x) = P(X \leq x), \quad x \in \mathbb{R}$$

è detta **funzione distribuzione (o di ripartizione)** della v.a. X .
Per variabili aleatorie continue si ha $P(X \leq x) = P(-\infty \leq X \leq x)$
quindi usando le nozioni già acquisite

$$F(x) = P(X \leq x) = \int_{-\infty}^x \rho(x') dx'$$

La densità di probabilità è non negativa per cui la funzione distribuzione è una funzione non decrescente della x .

FUNZIONE DISTRIBUZIONE

La funzione distribuzione è una funzione non decrescente della x e soddisfa le relazioni:

$$\lim_{x \rightarrow -\infty} F(x) = 0 \quad \lim_{x \rightarrow \infty} F(x) = \int_{-\infty}^{\infty} \rho(x) dx = 1$$

Inoltre $F(x) = P(X \leq x)$ e $P(a \leq X \leq b) = P(X \leq b) - P(X \leq a)$ implicano

$$P(a \leq X \leq b) = F(b) - F(a)$$

Esempio:

Si consideri la densità di probabilità dell'esercizio (uniformemente uguale a 1 nell'intervallo $[0, 1]$ e nulla al di fuori di esso)

Si consideri $x \leq 0$, si calcoli $F(x)$:

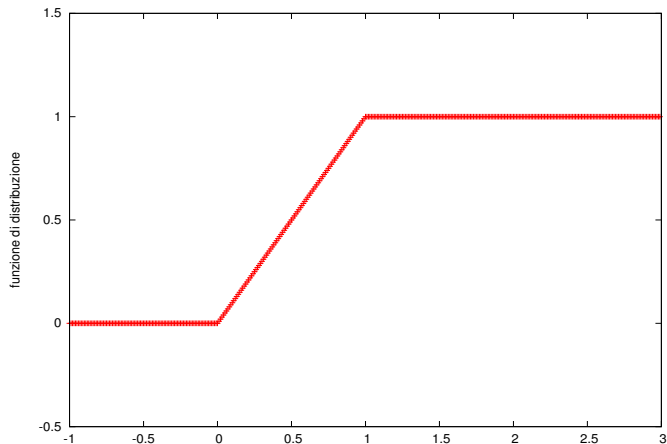
$$F(x) = \int_{-\infty}^x \rho(x') dx' = 0.$$

Si assuma ora $0 \leq x \leq 1$:

$$F(x) = \int_{-\infty}^x \rho(x') dx' = \int_0^x dx' = [x']_0^x = x.$$

Si assuma infine $1 \leq x$:

$$F(x) = \int_{-\infty}^x \rho(x') dx' = \int_0^1 dx' = [x']_0^1 = 1.$$



La funzione distribuzione $F(x)$ è una funzione non decrescente che verifica $\lim_{x \rightarrow -\infty} F(x) = 0$ e $\lim_{x \rightarrow \infty} F(x) = \int_{-\infty}^{\infty} \rho(x) dx = 1$.

VALORE ATTESO

Definizione. Il **valore atteso** (o valore medio o speranza matematica) della variabile aleatoria continua X , indicato con $\mathbb{E}(X)$ è il numero

$$\mu = \mathbb{E}(\mathbf{X}) = \int_{-\infty}^{\infty} \mathbf{x}\rho(\mathbf{x})d\mathbf{x}$$

Esempio

Si consideri di nuovo la densità di probabilità dell'esercizio (uniformemente uguale a 1 nell'intervallo $[0, 1]$ e nulla al di fuori di esso)

$$\mathbb{E}(\mathbf{X}) = \int_{-\infty}^{\infty} \rho(\mathbf{x})\mathbf{x}d\mathbf{x} = \int_0^1 \mathbf{x}d\mathbf{x} = \left[\frac{\mathbf{x}^2}{2}\right]_0^1 = \frac{1}{2}.$$

Le proprietà del valore atteso sono le stesse della media campionaria: se a e c sono costanti $E(aX + c) = aE(X) + c$.

VALORE ATTESO DI UNA FUNZIONE DI X

Sia g una funzione, allora $g(X)$ è una variabile aleatoria ed il suo valore atteso è dato da:

$$\mathbb{E}(g(\mathbf{X})) = \int_{-\infty}^{\infty} g(\mathbf{x})\rho(\mathbf{x})d\mathbf{x}$$

Esempio: si consideri di nuovo la densità di probabilità dell'esercizio (uniformemente uguale a 1 nell'intervallo $[0, 1]$ e nulla al di fuori di esso) e sia $g(x) = x^2$.

$$\mathbb{E}(g(\mathbf{X})) = \mathbb{E}(\mathbf{X}^2) = \int_{-\infty}^{\infty} \rho(\mathbf{x})g(\mathbf{x})d\mathbf{x} = \int_0^1 x^2 d\mathbf{x} = \left[\frac{x^3}{3} \right]_0^1 = \frac{1}{3}$$

VARIANZA E DEVIAZIONE STANDARD

Definizione. Sia X una variabile aleatoria con valore atteso $\mu = \mathbb{E}(X)$, la **varianza di X** denotata con $\text{var}(X)$ è

$$\text{var}(\mathbf{X}) = \mathbb{E}\left((\mathbf{X} - \mu)^2\right) = \int_{-\infty}^{\infty} (\mathbf{x} - \mu)^2 \rho(\mathbf{x}) d\mathbf{x}$$

Notate che la varianza è sempre non negativa ed è uguale a 0 solo se la variabile è certa (può assumere un solo valore).

Lo scarto quadratico medio o deviazione standard è invece:

$$\sigma = \sqrt{\text{var}(\mathbf{X})}.$$

Notare che le tutte le definizioni sono le stesse delle variabile aleatorie discrete, con le somme sostituite dagli integrali.

Notare inoltre la seguente proprietà (che è la stesse della statistica descrittiva): **$\text{var}(\mathbf{aX} + \mathbf{c}) = \mathbf{a}^2 \text{var}(\mathbf{X})$**

FORMULA UTILE

Sviluppando il quadrato della varianza $\text{var}(\mathbf{X}) = \mathbb{E}\left((\mathbf{X} - \mu)^2\right)$ si ha

$$\begin{aligned}\text{var}(\mathbf{X}) &= \mathbb{E}\left(X^2 - 2\mu X + \mu^2\right) = \mathbb{E}(X^2) - 2\mu\mathbb{E}(X) + \mu^2 \\ &= \mathbb{E}(X^2) - 2\mu^2 + \mu^2 = \mathbb{E}(\mathbf{X}^2) - \mu^2\end{aligned}$$

Dunque

$$\text{var}(\mathbf{X}) = \mathbb{E}(\mathbf{X}^2) - \mathbf{E}(\mathbf{X})^2 = \int_{-\infty}^{\infty} x^2 \rho(x) dx - \left(\int_{-\infty}^{\infty} x \rho(x) dx \right)^2$$

Una ulteriore proprietà molto importante (**già vista per le variabili discrete**) è la seguente: se X ed Y sono due variabile aleatorie allora,

- $\mathbb{E}(\mathbf{X} + \mathbf{Y}) = \mathbb{E}(\mathbf{X}) + \mathbb{E}(\mathbf{Y})$.

Inoltre se X ed Y sono due variabile aleatorie **indipendenti** allora,

- $\mathbf{var(X+Y) = var (X)+var (Y)}$.

Si tenga presente che per definizione due variabili aleatorie X ed Y si dicono **indipendenti** se gli eventi $\{X = x\}$ e $\{Y = y\}$ sono indipendenti per ogni x , valore di X ed y , valore di Y .

Esercizio: si consideri la densità di probabilità dell'esercizio (uniformemente uguale a 1 nell'intervallo $[0, 1]$ e nulla al di fuori di esso) e sia $g(x) = x^2$. Calcolare varianza e deviazione standard. Abbiamo già calcolato

$$\mu = \mathbb{E}(\mathbf{X}) = \frac{1}{2}, \quad \mathbb{E}(\mathbf{X}^2) = \frac{1}{3}.$$

La varianza, usando la "formula utile" si ottiene nel seguente modo

$$\text{var}(\mathbf{X}) = \mathbb{E}(\mathbf{X}^2) - \mu^2 = \frac{1}{3} - \left(\frac{1}{2}\right)^2 = \frac{1}{12}$$

Infine lo scarto quadratico medio o deviazione standard è :

$$\sigma = \sqrt{\text{var}(\mathbf{X})} = \frac{1}{\sqrt{12}} \simeq 0,3.$$

Nelle scienze a volte si usa la notazione $X = \mu \pm \sigma$, nel nostro caso $X = 0,5 \pm 0,3$.