

VARIANCE AND STANDARD DEVIATION OF DISCRETE VARIABLES

Definizione. Given a random variable X with expected value $\mu = \mathbb{E}(\mathbf{X})$, the variance of $\mathbf{X}$ is defined as

$$\text{var}(\mathbf{X}) = \mathbb{E}\left((\mathbf{X} - \mu)^2\right) = \sum_{i=1}^n (\mathbf{x}_i - \mu)^2 \mathbf{p}(\mathbf{x}_i)$$

The standard deviation is simply the square root of the variance:

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

Note that all definitions are the same as for descriptive statistics, with relative frequencies replaced by probabilities. For variance, this implies

$$\text{var}(\mathbf{aX} + \mathbf{c}) = \mathbf{a}^2 \text{var}(\mathbf{X}).$$

Also note that, similarly to the sample variance case, we can write

$$\text{var}(\mathbf{X}) = \mathbb{E}\left((\mathbf{X} - \mu)^2\right) = \sum_{i=1}^n (\mathbf{x}_i - \mu)^2 \mathbf{p}(\mathbf{x}_i) =$$

$$\sum_{i=1}^n \mathbf{x}_i^2 \mathbf{p}(\mathbf{x}_i) + \mu^2 \sum_i \mathbf{p}(\mathbf{x}_i) - 2\mu \sum_{i=1}^n \mathbf{x}_i \mathbf{p}(\mathbf{x}_i) = \mathbb{E}(\mathbf{X}^2) - \mu^2$$

. that is:

$$\text{var}(\mathbf{X}) = \mathbb{E}(\mathbf{X}^2) - [\mathbb{E}(\mathbf{X})]^2$$

.

Exercise 1. Assume that $\mathbf{X}$ is a random variable which only takes the **values 1 and -1** with equal probability $1/2$. Also assume $\mathbf{Y} = 3\mathbf{X}$. Compute expected value and variance of both.

$$\mathbb{E}(\mathbf{X}) = 1 \cdot \frac{1}{2} - 1 \cdot \frac{1}{2} = 0$$

$$\text{var}(\mathbf{X}) = \mathbb{E}(X^2) - [\mathbb{E}(X)]^2 = \mathbb{E}(X^2) = 1 \cdot \frac{1}{2} + (-1)^2 \cdot \frac{1}{2} = 1$$

The random variable $\mathbf{Y}$ only may assume the two **values 3 and -3 with equal probability**. We easily find its expected value and its variance as for $\mathbf{X}$.

Nevertheless, we can directly use $\mathbb{E}(\mathbf{Y}) = 3 \cdot \mathbb{E}(\mathbf{X}) = 0$ and $\text{var}(\mathbf{Y}) = 3^2 \cdot \text{var}(\mathbf{X}) = 9$.

A very important property of variance states that if X and Y are two **independent** random variables, then

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

Definition. Two random variables X e Y are defined **independent** if the events $\{X = x\}$ e $\{Y = y\}$ are independent for any value x of X and for any value y of Y .

If two variables are independent, then $\mathbb{E}(XY) = \mathbb{E}(X) \cdot \mathbb{E}(Y)$, which is sufficient to demonstrate the above linear property of the variance.

Notice that for the expected value, independence is not a necessary condition for the analogous linear property.

Exercise 2.

Roll a die and calculate the mean and variance of the resulting value X .

$$\mathbb{E}(X) = \frac{1}{6}(1 + 2 + 3 + 4 + 5 + 6) = 3,5$$

$$\text{var}(X) = \mathbb{E}(X^2) - [\mathbb{E}(X)]^2 = \frac{1}{6}(1+4+9+16+25+36) - (3,5)^2 = 2,916$$

Exercise 3, method 1.

Roll two dice and calculate the mean and variance of the resulting value Z .

The rolls are independent and the result Z is the sum of X e Y which are the results of the two dice.

$$\mathbb{E}(Z) = \mathbb{E}(X) + \mathbb{E}(Y) = 3,5 + 3,5 = 7$$

$$\text{var}(Z) = \text{var}(X) + \text{var}(Y) = 2,916 + 2,916 = 5,83$$

Exercise 3, method 2.

Roll two dice and calculate the mean and variance of the resulting value. Z .

The variable Z assumes the values

$$2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12$$

with probabilities

$$1/36, 2/36, 3/36, 4/36, 5/36, 6/36, 5/36, 4/36, 3/36, 2/36, 1/36.$$

Then:

$$\mathbb{E}(Z) = \frac{1}{36}2 + \frac{2}{36}3 + \dots + \frac{1}{36}12 = 7$$

$$\text{var}(Z) = \mathbb{E}(Z^2) - [\mathbb{E}(Z)]^2 = \frac{1}{36}4 + \frac{2}{36}9 + \dots + \frac{1}{36}144 - 7^2 = 5,83$$

BINOMIAL DISTRIBUTION

Let us consider again the case of dichotomous choices, the single individual k may answer yes ($X_k = 1$) with probability μ or no ($X_k = 0$) with probability $1 - \mu$, which implies $\mathbb{E}(X_k) = \mu$ for all k . We are interested in the variable $X = \sum_{k=1}^n X_k$. Since X_k may only take the two values 0 and 1, one has

$$X_k^2 = X_k \quad \rightarrow \quad \mathbb{E}(X_k^2) = \mathbb{E}(X_k) = \mu$$

which implies

$$\text{var}(X_k) = \mathbb{E}(X_k^2) - (\mathbb{E}(X_k))^2 = \mu - \mu^2 = \mu(1 - \mu).$$

Finally, given the independence of the rolls:

$$\text{var}(X) = \sum_{k=1}^n \text{var}(X_k) = n\mu(1 - \mu).$$

DIRECT COMPUTATION OF THE VARIANCE

The expected value of X^2 can be easily computed. Given that

$$\begin{aligned}\sum_{x=0}^n x^2 \binom{n}{x} a^x b^{n-x} &= a \frac{\partial}{\partial a} \left(a \frac{\partial}{\partial a} \sum_{x=0}^n \binom{n}{x} a^x b^{n-x} \right) = \\ &= a \frac{\partial}{\partial a} \left(a \frac{\partial}{\partial a} (a+b)^n \right) = na(a+b)^{n-1} + n(n-1)a^2(a+b)^{n-2}\end{aligned}$$

by choosing $a = \mu$ and $b = 1 - \mu$ one obtains the expected value

$$\mathbb{E}(X^2) = n\mu + n(n-1)\mu^2$$

$$\text{var}(X) = \mathbb{E}(X^2) - (\mathbb{E}(X))^2 = n\mu(1 - \mu),$$

where we have used the $\mathbb{E}(X) = n\mu$ which we previously obtained.

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

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

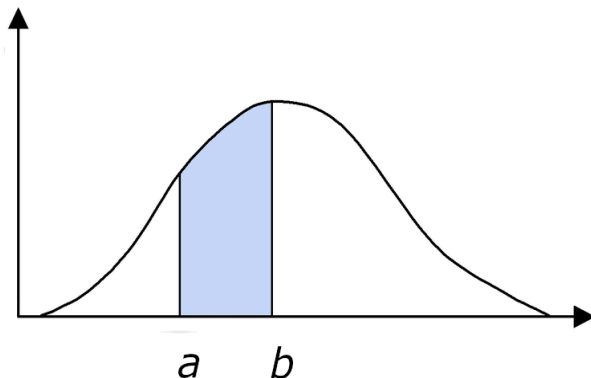
e che soddisfa la condizione

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

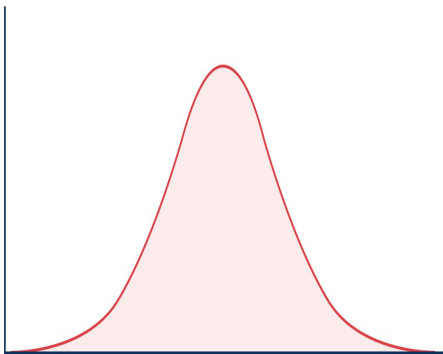
Si definisce poi la probabilità che X sia compresa fra a e b come

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

$P(a \leq X \leq b)$ = area della regione delimitata dall'asse x , il grafico di $\rho(x)$ e le rette $x=a$ 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 $\int_{-\infty}^{\infty} \rho(\mathbf{x}) d\mathbf{x} = 1$ 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 $\int_{-\infty}^{\infty} \rho(\mathbf{x}) d\mathbf{x} = \mathbf{P}(-\infty < \mathbf{X} < +\infty)$.

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:

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

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

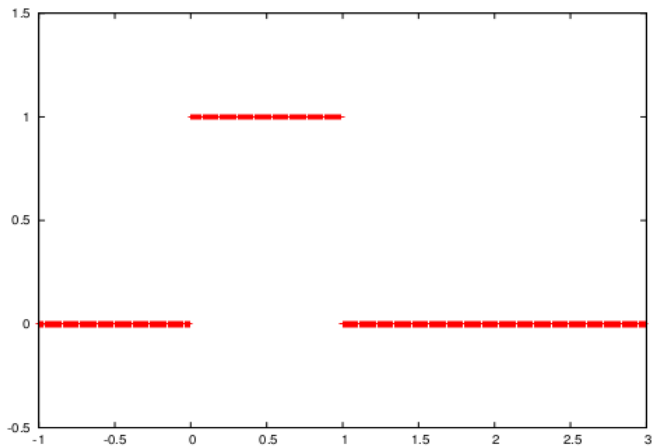
Determinare il valore di α per il quale è una densità di probabilità .

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 $P(a \leq X \leq b)$:

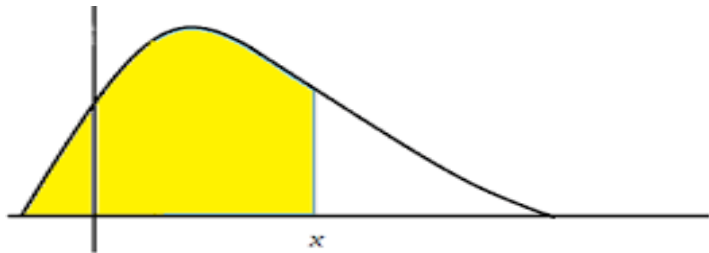
$$P(a \leq X \leq b) = \int_a^b \rho(x) dx = \int_0^b dx = [x]_0^b = b.$$

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

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

Definition. The **distribution function** of the random variable X is defined as

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



The probability density is non-negative, so the distribution function (yellow area in the figure) is a non-decreasing function of x .

One also has

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

Which is also coherent with the fact that $F(x)$ is a non decreasing function of x , in fact, $P(a \leq X \leq b)$ is a non negative probability, which implies $F(b) \geq F(a)$.

Finally, it is easy to find

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

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 $F(x)$:

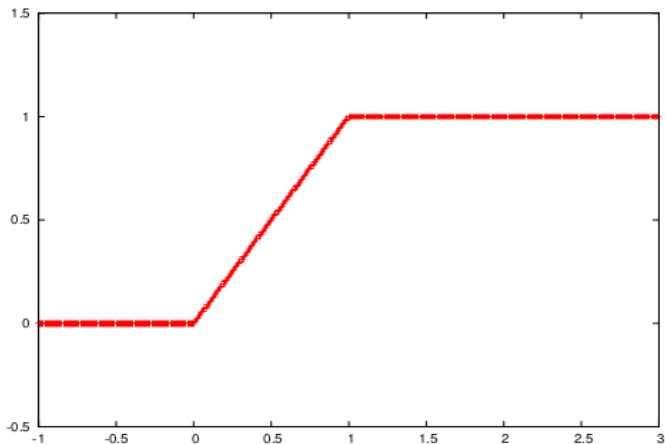
$$F(x) = \int_{-\infty}^x \rho(x') dx' = 0.$$

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

Esercise:

Consider now a probability density which vanishes when $x < 0$ and equals e^{-x} when $x \geq 0$ and compute $F(x)$.

Assume $x \leq 0$:

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

Assume $x \geq 0$:

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

As expected

$$F(-\infty) = 0, \quad F(\infty) = 1$$