Probabilities

Continuous Random Variables

Mohamad GHASSANY

EFREI PARIS

Plan

Recall: **Discrete** Random Variable

Continuous Random Variable

Distribution function of continuous random variables

Function of a continuous random variable

Moments of Continuous Random Variable

$\sum$

$\int$

Recall: Discrete Random Variable

Discrete Random Variable

$X(\Omega)$ takes a finite values

- X is a ***discrete*** random variable if the set of possible values of X, $X(\Omega)$, is finite or countable.
  - The *probability distribution* defined on $X(\Omega)$ by $p_i = p(x_i) = P(X = x_i)$
  - $p(x_i) \geqslant 0$, $\sum_{i=1}^{\infty} p(x_i) = 1$, and $P(a < X \leqslant b) = \sum_{i/a<x_i \leqslant b} p(x_i)$.

$\mathbb{N}$

$X=1$

$X=2$

Recall: Discrete Random Variable

Discrete Random Variable

- X is a ***discrete*** random variable if the set of possible values of X, $X(\Omega)$, is finite or countable.
  - The *probability distribution* defined on $X(\Omega)$ by $p_i = p(x_i) = P(X = x_i)$
  - $p(x_i) \geqslant 0$, $\sum_{i=1}^{\infty} p(x_i) = 1$, and $P(a < X \leqslant b) = \sum_{i/a<x_i \leqslant b} p(x_i)$.

Distribution function of a d.r.v.

- The distribution function of X, that we note $F_X(a)$, defined for each real number a, $-\infty < a < \infty$, by $\boxed{F_X(a) = P(X \leqslant a)} = \sum_{i/x_i \leqslant a} P(X = x_i)$.

$\forall a \in \mathbb{R}$

$F_X(a) = P(X < a)$

Discrete Random Variable

- X is a ***discrete*** random variable if the set of possible values of X, $X(\Omega)$, is finite or countable.
  - The *probability distribution* defined on $X(\Omega)$ by $p_i = p(x_i) = P(X = x_i)$
  - $p(x_i) \geqslant 0$, $\sum_{i=1}^{\infty} p(x_i) = 1$, and $P(a < X \leqslant b) = \sum_{i/a<x_i\leqslant b} p(x_i)$.

Distribution function of a d.r.v.

- The distribution function of X, that we note $F_X(a)$, defined for each real number a, $-\infty < a < \infty$, by $F_X(a) = P(X \leqslant a) = \sum_{i/x_i \leqslant a} P(X = x_i)$.
  - Staircase function.
  - $F_X(a) \leqslant 1$ (it is a probability).
  - $F_X(a)$ is continuous at right.
  - $\lim\limits_{a\to-\infty} F_X(a) = 0$ et $\lim\limits_{a\to\infty} F_X(a) = 1$
  - $P(a < X \leqslant b) = F(b) - F(a)$ pour tout $a < b$

Recall: Discrete Random Variable

Discrete Random Variable

- X is a ***discrete*** random variable if the set of possible values of X, $X(\Omega)$, is finite or countable.
  - The *probability distribution* defined on $X(\Omega)$ by $p_i = p(x_i) = P(X = x_i)$
  - $p(x_i) \geqslant 0$, $\sum_{i=1}^{\infty} p(x_i) = 1$, and $P(a < X \leqslant b) = \sum_{i/a<x_i\leqslant b} p(x_i)$.

Distribution function of a d.r.v.

- The distribution function of X, that we note $F_X(a)$, defined for each real number a, $-\infty < a < \infty$, by $F_X(a) = P(X \leqslant a) = \sum_{i/x_i \leqslant a} P(X = x_i)$.
  - Staircase function.
  - $F_X(a) \leqslant 1$ (it is a probability).
  - $F_X(a)$ is continuous at right.
  - $\lim_{a\to-\infty} F_X(a) = 0$ et $\lim_{a\to\infty} F_X(a) = 1$
  - $P(a < X \leqslant b) = F(b) - F(a)$ pour tout $a < b$

Moments of d.r.v.

$\sum x_i^2 p(x_i) \leftarrow$ transfer theorem

- $E(X) = \sum_{i\in\mathbb{N}} x_i p(x_i)$
- $V(X) = E(X^2) - E^2(X)$ $= E([X - E(X)]^2)$

$E(g(x)) = \sum g(x_i) p(x_i)$

Continuous Random Variable

Continuous Random Variable - Density function and Probabilities

- Previously we have dealt with Discrete Random Variables, i.e. variables whose universe is finite or countable.
- There are however variables whose universe is infinite uncountable.
- Examples:
  - The arrival time of a train at a given station.
  - The lifetime of a transistor.

unlimited nb of possibilities

$X(\Omega) = \mathbb{R}$

$\boxed{[0,1]}$

$\cancel{\{0,1\}}$

$X(\Omega) = [0,1]$

$$f(x) = x \times \mathbb{1}_{[0,1]}(x) = \begin{cases} x & \text{if } x \in [0,1] \\ 0 & \text{if not} \end{cases}$$

[1] Not all Continuous Random Variable have a density function.

- Previously we have dealt with Discrete Random Variables, i.e. variables whose universe is finite or countable.
- There are however variables whose universe is infinite uncountable.
- Examples:
  - The arrival time of a train at a given station.
  - The lifetime of a transistor.

Definition

*X is a continuous random variable [1] **with density** if there exists a non-negative function f defined for any $x \in \mathbb{R}$ and verifying for any set B of real numbers the property*

$B \subset \mathbb{R}$

$$\boxed{P(X \in B) = \int_B f(x)\,dx}$$

two conditions:

* $f \geq 0$
* $\int_{\mathbb{R}} f(x)\,dx = 1$

$P(X(\Omega))$

The function f is called density function of the random variable X.

[1] Not all Continuous Random Variable have a density function.

- Previously we have dealt with Discrete Random Variables, i.e. variables whose universe is finite or countable.
- There are however variables whose universe is infinite uncountable.
- Examples:
  - The arrival time of a train at a given station.
  - The lifetime of a transistor.

Definition

X is a continuous random variable [1] ***with density*** *if there exists a non-negative function f defined for any $x \in \mathbb{R}$ and verifying for any set B of real numbers the property*

$$P(X \in B) = \int_B f(x)dx$$

The function f is called density function of the random variable X.

- All probability questions related to X can be treated with f.
- For example if $B = [a, b]$, we get:

$$P(X \in B) = P(a \leqslant X \leqslant b) = \int_a^b f(x)dx$$

$P(X \in [1,2])$

$= P(1 \leq X \leq 2)$

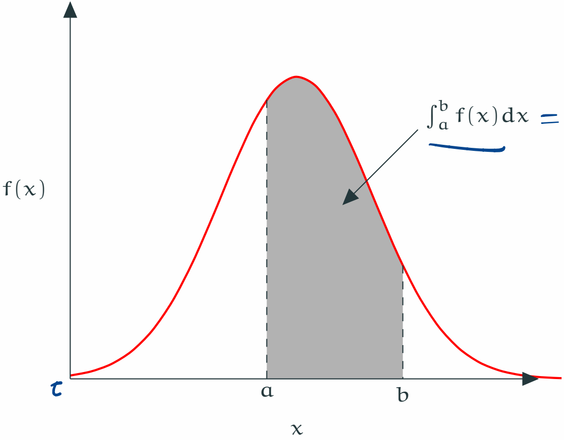
$= \int_1^2 f(x)dx$

[1] Not all Continuous Random Variable have a density function.

Graphically, $P(a \leqslant X \leqslant b)$ is the area of the surface between the x-axis, the curve corresponding to $f(x)$ and the lines $x = a$ and $x = b$.

$\int_{\mathbb{R}} f(x)\,dx = 1$

total area of the surface between f and x-axis



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

Figure 1: $P(a \leqslant X \leqslant b) =$ *area of shaded surface*

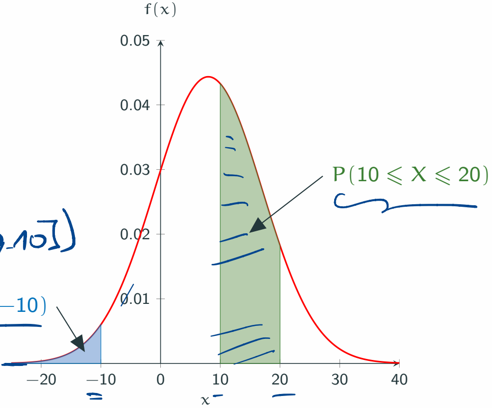


Figure 2: The colored areas corresponds to probabilities. $f(x)$ *being a probability density function.*

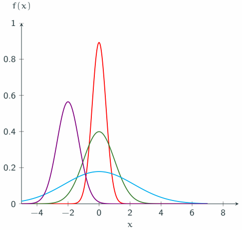
$X(\Omega) = \mathbb{R}^+$

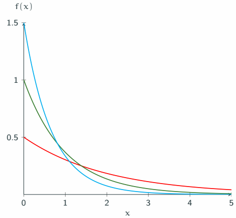
$P(X \leq 2)$

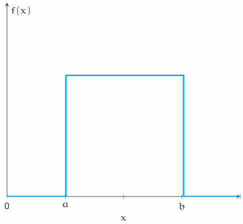
$\int_0^2 f(x)dx$

$P(X \in]-\infty, 10])$

$\int_{-\infty}^2 f(x)dx = \underbrace{\int_{-\infty}^0 f(x)dx}_{=0} + \int_0^2 f(x)dx$







Proprieties

For any continuous random variable X of density f:

$P(X=a)=0$

- $f(x) \geqslant 0 \quad \forall x \in \mathbb{R}$
- $\int_{-\infty}^{+\infty} f(x)dx = 1$
- Since $P(a \leqslant X \leqslant b) = \int_a^b f(x)dx$, if $a = b$ then $P(X = a) = \int_a^a f(x)dx = 0$
- This means that the probability of a continuous random variable taking a fixed isolated value is always zero.

$\int_{\mathbb{R}} f(x)\,dx = \int_{\infty}^{+\infty} f(x)\,dx = 1$

$P(X=a) = P(a \leq X \leq a) = \int_a^a f(x)dx = 0$

Example

Let X be the random real variable of probability density

$$f(x) = \begin{cases} \overset{\text{constant}}{k}x & \text{if } 0 \leqslant x \leqslant 5 \\ 0 & \text{if not} \end{cases}$$

$X(\Omega) = [0,5]$.

1. Calculate k.
2. Calculate: $P(1 \leqslant X \leqslant 3), P(2 \leqslant X \leqslant 4)$ and $P(X < 3)$.

1) we know that $\int_{\mathbb{R}} f(x)dx = 1$, $\int_{\mathbb{R}} f(x)dx = \int_0^5 kx\,dx = \left[k\frac{x^2}{2}\right]_0^5 = k\left(\frac{25}{2} - \frac{0}{2}\right) = \frac{25k}{2}$

so $k = 2/25$.

2) * $P(1 \leq X \leq 3) = \int_1^3 f(x)dx = \frac{2}{25}\left[\frac{x^2}{2}\right]_1^3 = \frac{2}{25}\left(\frac{8}{2}\right) = 8/25$.

* $P(2 \leq X \leq 4) = \int_2^4 f(x)dx = ---$

* $P(X < 3) = \int_{-\infty}^3 f(x)dx = \int_0^3 \frac{2}{25}x\,dx = \frac{2}{25}\left[\frac{x^2}{2}\right]_0^3 = \frac{9}{25}$.

Note that $P(X<3) = P(X \leq 3)$

Example

Let X be the random real variable of probability density

$$f(x) = \begin{cases} kx & \text{if } 0 \leqslant x \leqslant 5 \\ 0 & \text{if not} \end{cases}$$

1. Calculate k.
2. Calculate: $P(1 \leqslant X \leqslant 3), P(2 \leqslant X \leqslant 4)$ and $P(X < 3)$.

Example

Let X be a continuous random variable with density function

$$f(x) = \begin{cases} \frac{1}{6}x + k & \text{if } 0 \leqslant x \leqslant 3 \\ 0 & \text{if not} \end{cases}$$

1. Calculate k.
2. Calculate $P(1 \leqslant X \leqslant 2)$

$= \int_1^2 \left(\frac{1}{6}x + \frac{1}{12}\right) dx = \cdots$

$\int_{\mathbb{R}} f(x)dx = \int_0^3 \left(\frac{1}{6}x + k\right)dx = \left[\frac{1}{6}\frac{x^2}{2} + kx\right]_0^3$

$= \frac{3}{4} + 3k = 1$ So $k = 1/12$.

fonction de répartition

Distribution function of continuous random variables

Definition

If as for Random Variable Discrete, we define the distribution function of X *by:*

Since X is continuos

$$\begin{aligned} F_X : \mathbb{R} &\longrightarrow \mathbb{R} \\ x &\longmapsto F_X(a) = P(X \leqslant a) = P(X < a) \end{aligned}$$

then the relation between the distribution function F_X *and the probability density function* $f(x)$ *is the following:*

$$\forall \quad a \in \mathbb{R} \quad F_X(a) = P(X \leqslant a) = \int_{-\infty}^{a} f(x)\,dx$$

Definition

If as for Random Variable Discrete, we define the distribution function of X by:

$$\begin{aligned} F_X : \mathbb{R} &\longrightarrow \mathbb{R} \\ x &\longmapsto F_X(a) = P(X \leqslant a) \end{aligned}$$

then the relation between the distribution function F_X and the probability density function $f(x)$ is the following:

$$\forall \quad a \in \mathbb{R} \quad F_X(a) = P(X \leqslant a) = \int_{-\infty}^{a} f(x)dx$$

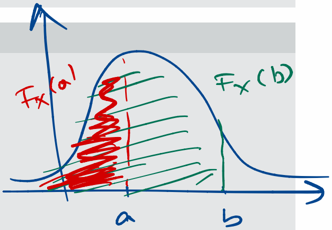
Proprieties

For a continuous random variable X:

- $F'_X(x) = \frac{d}{dx}F_X(x) = f(x).$ when F_X is derivable.
- For all real numbers $a \leqslant b$,

$$P(a < X < b) = P(a < X \leqslant b) = P(a \leqslant X < b)$$

$$= P(a \leqslant X \leqslant b) = F_X(b) - F_X(a) = \int_a^b f(x)dx$$



Distribution function of continuous random variables

The distribution function corresponds to the cumulative probabilities associated with the continuous random variable on an interval.

The distribution function corresponds to the cumulative probabilities associated with the continuous random variable on an interval.

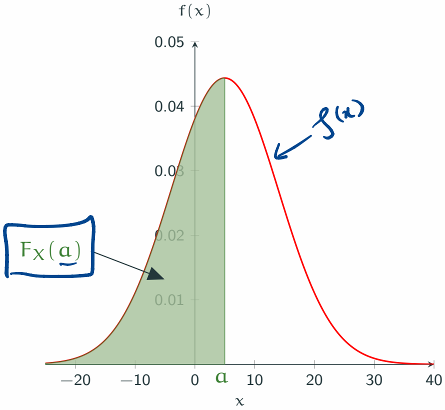
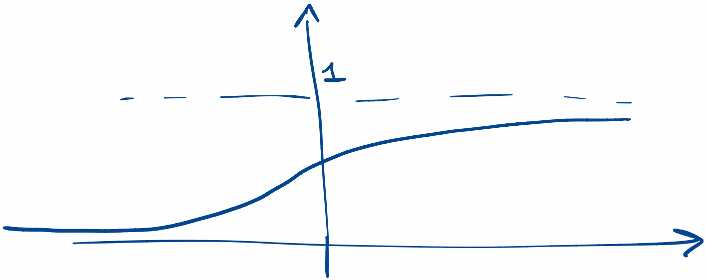


Figure 3: The area shaded in green under the curve of the density function corresponds to the probability $P(X < a) = F_X(a)$ *and is* 0.5 *because this corresponds exactly to half of the total area under the curve.*

Proprieties

The properties of the distribution function are as follows:

1. F_X is continuous on $\mathbb{R}$, derivable at any point where f is continuous.
2. F_X is increasing on $\mathbb{R}$.
3. F_X has values in $[0, 1]$.
4. $\lim_{x \to -\infty} F_X(x) = 0$ and $\lim_{x \to +\infty} F_X(x) = 1$.



Proprieties

The properties of the distribution function are as follows:

1. F_X is continuous on $\mathbb{R}$, derivable at any point where f is continuous.
2. F_X is increasing on $\mathbb{R}$.
3. F_X has values in $[0, 1]$.
4. $\lim_{x \to -\infty} F_X(x) = 0$ and $\lim_{x \to +\infty} F_X(x) = 1$.

$\forall a \in \mathbb{R}, F_X(a) = P(X \leq a)$

* if $a \leq 0$, $F_X(a) = 0$
* if $a \geq 5$, $F_X(a) = 1$
* if $a \in [0, 5]$

$F_X(a) = P(X \leq a)$

$= \int_0^a f(x)dx = \frac{2}{25}\left[\frac{x^2}{2}\right]_0^a$

$= a^2/25$

Example

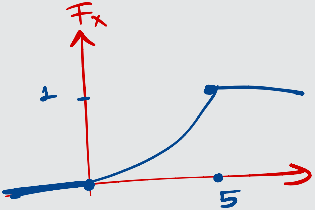
Let X and Y two random variables of density functions:

$$f_X(x) = \begin{cases} \overset{2/25}{k}x & \text{if } 0 \leqslant x \leqslant 5 \\ 0 & \text{if not} \end{cases}$$

and

$$f_Y(y) = \begin{cases} \frac{1}{6}y + k & \text{if } 0 \leqslant y \leqslant 3 \\ 0 & \text{if not} \end{cases}$$

Calculate $\boxed{F_X(a)}$ and $F_Y(a)$ for all $a \in \mathbb{R}$.



X Let $Y = g(X)$

Function of a continuous random variable

- Let X be a continuous random variable with density f_X and distribution function F_X.
- Let h be a continuous function defined on $X(\Omega)$, then $Y = h(X)$ is a random variable.

Function of a continuous random variable

- Let X be a continuous random variable with density f_X and distribution function F_X.
- Let h be a continuous function defined on $X(\Omega)$, then $Y = h(X)$ is a random variable.
- To determine the density of Y, denoted f_Y, we first compute the distribution function of Y, denoted F_Y, then we derivate it to determine f_Y.

- Let X be a continuous random variable with density f_X and distribution function F_X.
- Let h be a continuous function defined on $X(\Omega)$, then $Y = h(X)$ is a random variable.
- To determine the density of Y, denoted f_Y, we first compute the distribution function of Y, denoted F_Y, then we derivate it to determine f_Y.

Calculating the densities

Let X be a continuous random variable with density f_X and distribution function F_X. Find the density function of the following random variables:

- $Y = aX + b$
- $Z = X^2$
- $T = e^X$

$* \ \forall z \in Z(\Omega)$

$F_Z(z) = P(Z \leq z)$

$= P(X^2 \leq z)$

$= P(-\sqrt{z} \leq X \leq +\sqrt{z})$

$= \int_{-\sqrt{z}}^{+\sqrt{z}} f_X(x)dx$ or $F_X(\sqrt{z}) - F_X(-\sqrt{z})$

$X^2 \leq z$

$-\sqrt{z} \leq X \leq \sqrt{z}$

$X(\Omega)$

$Z = X^2$

$F_Z \ / f_Z$

So $f_Z(z) = (F_Z(z))' = \frac{1}{2\sqrt{z}} f_X(\sqrt{z}) + \frac{1}{2\sqrt{z}} f_X(-\sqrt{z}) = \frac{1}{2\sqrt{z}} [f_X(\sqrt{z}) + f_X(-\sqrt{z})]$

- Let X be a continuous random variable with density f_X and distribution function F_X.
- Let h be a continuous function defined on $X(\Omega)$, then $Y = h(X)$ is a random variable.
- To determine the density of Y, denoted f_Y, we first compute the distribution function of Y, denoted F_Y, then we derivate it to determine f_Y.

Calculating the densities

Let X be a continuous random variable with density f_X and distribution function F_X. Find the density function of the following random variables:

- $Y = aX + b$
- $Z = X^2$
- $T = e^X$

$* \; Z(\Omega) = [0,1]$

$\bullet$ if $z \leq 0$, $F_Z(z) = 0$; if $z \geq 1$, $F_Z(z) = 1$

$\bullet$ if $z \in [0,1]$, $F_Z(z) = P(Z \leq z)$

$= P(X^2 \leq z) = P(\cancel{-\sqrt{z}}^{\,0} \leq X \leq \sqrt{z})$ $X(\Omega)$

$= \int_0^{\sqrt{z}} 2x\,dx = [x^2]_0^{\sqrt{z}} = z.$

So $F_Z(z) = \begin{cases} 0 & \text{if } z \leq 0 \\ z & \text{if } 0 \leq z \leq 1 \\ 1 & \text{if } z \geq 1 \end{cases}$

So $f_Z(z) = 1 \times \mathbb{1}_{[0,1]}(z)$.

Example

Let X a random variable having the density function:

$X(\Omega) = [0,1]$

$$f_X(x) = 2x \times \mathbb{1}_{[0,1]}(x)$$

Determine the density function of: $Y = 3X + 1$, $Z = X^2$ and $T = e^X$.

Moments of Continuous Random Variable

$f_X(x) = 2x \times \mathbb{1}_{[0,1]}(x)$

$T = e^X$

$* \ T(\Omega) = [1, e]$

$* \ F_T(t) = \begin{cases} 0 \text{ if } t \leq 1 \\ 1 \text{ if } t \geq e \\ P(T \leq t) \text{ if } t \in [1, e] \end{cases}$

$\mathbb{1}_A(x) = \begin{cases} 1 \text{ if } x \in A \\ 0 \text{ if } x \in \bar{A} \end{cases}$

if $t \in [1, e]$,

$F_T(t) = P(T \leq t)$

$= P(e^X \leq t)$

$= P(X \leq \ln(t))$

$= \int_0^{\ln(t)} 2x \, dx = [x^2]_0^{\ln t}$

$F_T(t) = \ln^2 t.$

So $f_T(t) = \frac{dF_T(t)}{dt} = \frac{2 \ln t}{t} \times \mathbb{1}_{[1,e]}(t).$

Definition

If X is a continuous random variable of density f, we call the expected value of X, the real $E(X)$, defined by:

$$E(X) = \int_{-\infty}^{+\infty} x f(x) \, dx$$

if it exists.

Expected value

Definition

If X is a continuous random variable of density f, we call the expected value of X, the real $E(X)$, defined by:

$$E(X) = \int_{-\infty}^{+\infty} x f(x) \, dx$$

if it exists.

The properties of the expected value of a continuous random variable are the same as for a discrete random variable.

Expected value

Definition

If X is a continuous random variable of density f, we call the expected value of X, the real $E(X)$, defined by:

$$E(X) = \int_{-\infty}^{+\infty} x f(x)\,dx$$

if it exists.

The properties of the expected value of a continuous random variable are the same as for a discrete random variable.

Proprieties

Let X be a continuous random variable,

- $E(aX + b) = aE(X) + b \qquad a \geqslant 0$ and $b \in \mathbb{R}$.
- If $X \geqslant 0$ then $E(X) \geqslant 0$.
- If X and Y are two Random Variables defined on the same universe Ω then

$$E(X + Y) = E(X) + E(Y)$$

Theorem

If X is a random variable of density $f(x)$, then for any real function g we have

$$E[g(X)] = \int_{-\infty}^{+\infty} g(x) f(x) dx$$

$$E(X^2) = \sum x_i^2 P(X = x_i)$$

$$E(X^2) = \int_{\mathbb{R}} x^2 f(x) dx.$$

Theorem

If X is a random variable of density $f(x)$, then for any real function g we have

$$E[g(X)] = \int_{-\infty}^{+\infty} g(x)f(x)\,dx$$

Example

Let X a random variable of density

$$f_X(x) = \begin{cases} 2x & \text{if } 0 \leqslant x \leqslant 1 \\ 0 & \text{if not} \end{cases}$$

Calculate the expected value of $Y = 3X + 1$, $Z = X^2$ and $T = e^X$.

$\bullet\ E(X) = \int_{\mathbb{R}} x f(x)\,dx = \int_0^1 x(2x)\,dx = \int_0^1 2x^2\,dx = 2\left[\frac{x^3}{3}\right]_0^1 = 2/3.$

$\bullet\ E(Z) = E(X^2) = \int_{\mathbb{R}} x^2 f(x)\,dx = \int_0^1 x^2(2x)\,dx = 2\left[\frac{x^4}{4}\right]_0^1 = 1/2 \leftarrow$

$\bullet\ E(T) = \int_0^1 e^x f(x)\,dx = \int_0^1 2x e^x\,dx$

$2\int_0^1 x e^x\,dx$

der $\left(\begin{array}{ccc} x & \xrightarrow{+} & e^x \\ 1 & \xrightarrow{-} & e^x \\ 0 & & e^x \end{array} \right)$ int

$= 2\left[[xe^x]_0^1 - [e^x]_0^1 \right]$

$= 2(1 - e + 1) = 2(2 - e).$

Variance of a continuous random variable

The variance of a random variable $V(X)$ is a dispersion parameter which corresponds to the centered moment of order 2 of the random variable X.

Variance of a continuous random variable

The variance of a random variable $V(X)$ is a dispersion parameter which corresponds to the centered moment of order 2 of the random variable X.

Definition

If X is a random variable with expectation $E(X)$, we call the variance of X the real

$$V(X) = E\left([X - E(X)]^2\right) = E(X^2) - [E(X)]^2$$

If X is a continuous random variable, we compute $E(X^2)$ using the transfer theorem,

$$E(X^2) = \int_{-\infty}^{+\infty} x^2 f(x)\, dx$$

Variance of a continuous random variable

The variance of a random variable $V(X)$ is a dispersion parameter which corresponds to the centered moment of order 2 of the random variable X.

Definition

If X is a random variable with expectation $E(X)$, we call the variance of X the real

$$V(X) = E\left([X - E(X)]^2\right) = E(X^2) - [E(X)]^2$$

If X is a continuous random variable, we compute $E(X^2)$ using the transfer theorem,

$$E(X^2) = \int_{-\infty}^{+\infty} x^2 f(x)\,dx$$

Example

Calculate la variance of X defined in the previous example.

$f_X(x) = 2x \times \mathbb{1}_{[0,1]}(x)$

$E(X) = 2/3$, $E(X^2) = \int_0^1 x^2 f(x)\,dx = \ldots = 1/2$

$V(X) = E(X^2) - E^2(X) = \frac{1}{2} - \left(\frac{2}{3}\right)^2 = \_\_ \_\_ .$

Proprieties

If X is a random variable with a variance then:

- $V(X) \geqslant 0$, if it exists.
- $\forall\ a \in \mathbb{R}, V(aX) = a^2V(X)$
- $\forall\ (a, b) \in \mathbb{R}, V(aX + b) = a^2V(X)$
- If X and Y are two independent Random Variables, $V(X + Y) = V(X) + V(Y)$

$$V(X+Y) = V(X) + V(Y) + 2\,\mathrm{cov}(X,Y)$$

Proprieties

If X is a random variable with a variance then:

- $V(X) \geqslant 0$, if it exists.
- $\forall\ a \in \mathbb{R}, V(aX) = a^2V(X)$
- $\forall\ (a,b) \in \mathbb{R}, V(aX+b) = a^2V(X)$
- If X and Y are two independent Random Variables, $V(X+Y) = V(X) + V(Y)$

Definition

If X is a random variable with variance $V(X)$, we call the standard deviation of X the real:

$$\sigma_X = \sqrt{V(X)}$$