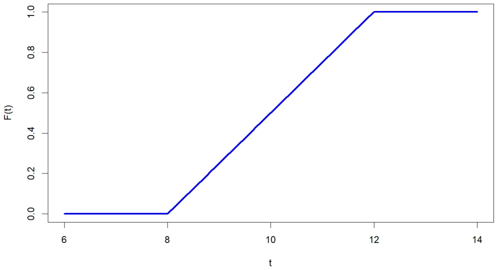
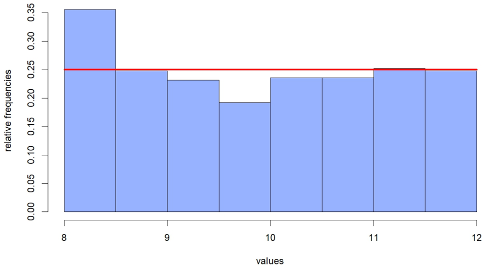
Uniform distribution (Lecture 12)



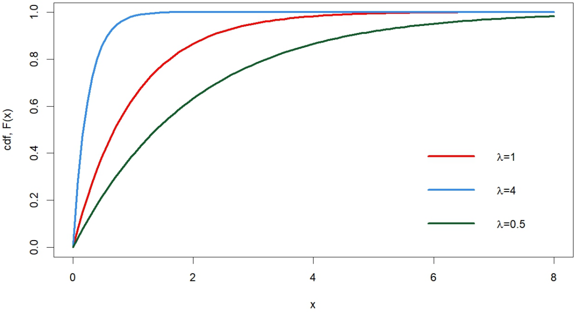
Cumulative distribution function of the uniform distribution on the interval $[8, 12]$

Uniform distribution



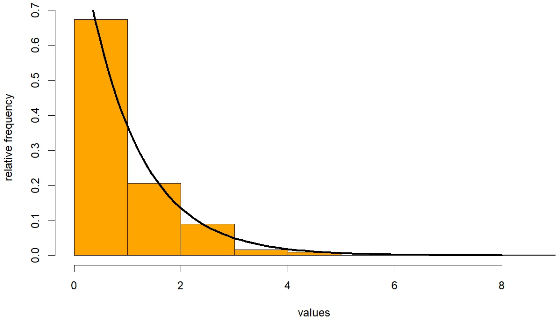
Histogram of a sample of size 500 with uniform distribution on the interval $[8, 12]$

Exponential distribution



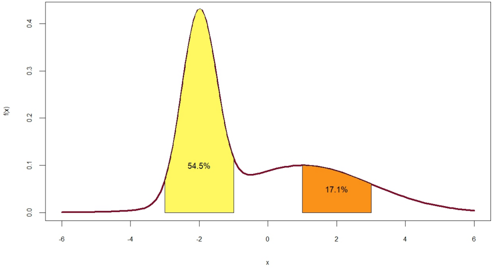
Cumulative distribution function of exponential distribution with different parameters: $\left(\lambda = \frac{1}{2}, 1, \text{ and } 4\right)$

Exponential distribution



Histogram of a sample of size 500 from exponential distribution with parameter $\lambda = 1$

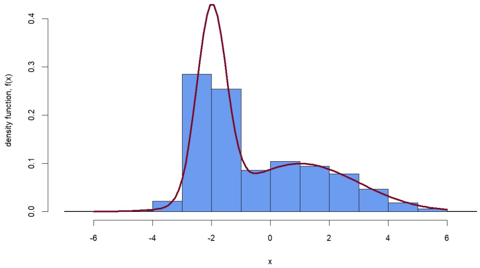
Density function



If random variable X has density function f (shown on the figure): $\mathbb{P}(-3 \leq X \leq -1) = \int_{-3}^{-1} f(x)dx = 54.5\%$;

$\mathbb{P}(1 \leq X \leq 3) = \int_{1}^{3} f(x)dx = 17.1\%$.

Density function



A density function and the histogram of a sample of size 1000 from the same distribution;

higher values of the density function $\rightarrow$ higher frequency;
sample: independent random variables such that all have density function f

Density function: definition

Random variable $X : \Omega \to \mathbb{R}$ has **density function** $f : \mathbb{R} \to \mathbb{R}$ if

$$\mathbb{P}(X \leq t) = \int_{-\infty}^{t} f(x)dx$$

holds for every real number $t \in \mathbb{R}$.

Density function: definition

Random variable $X : \Omega \to \mathbb{R}$ has **density function** $f : \mathbb{R} \to \mathbb{R}$ if

$$\mathbb{P}(X \le t) = \int_{-\infty}^{t} f(x)dx$$

holds for every real number $t \in \mathbb{R}$.

There exists random variables that do not have density function, for example, the discrete ones. If X **has density function**, then we say that X is **absolutely continuous**.

If random variable X has density function f, then for all real numbers $a < b$ we have

$$\mathbb{P}(a < X < b) = \mathbb{P}(a \le X \le b) = \int_{a}^{b} f(x)dx.$$

Properties of the density function

Let X be absolutely continuous random variable with cumulative distribution function F. (a) If f is the density function of X, then for every $t \in \mathbb{R}$ we have

$$F(t) = \mathbb{P}(X \leq t) = \int_{-\infty}^{t} f(x)\, dx.$$

(b) The function $f(t) = F'(t)$ is the density function of X (defined for t where F is differentiable).

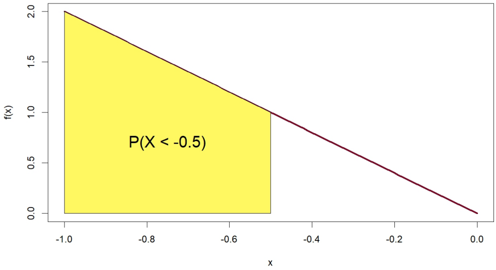
If $f : \mathbb{R} \to \mathbb{R}$ is a density function, then

(i) $f(x) \geq 0$ holds for "almost every" $x \in \mathbb{R}$ (for example, finitely many or countably many exceptions are allowed).

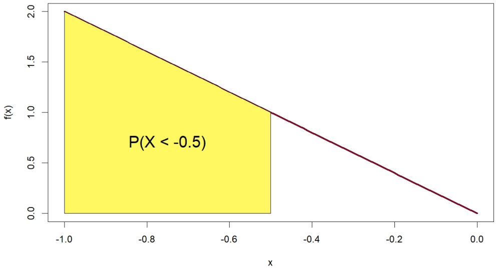
(ii) $\int_{-\infty}^{\infty} f(x)\, dx = 1$.

Other direction: if f satisfies the two properties above, then there exists a random variable whose density function is f.

Density function: example



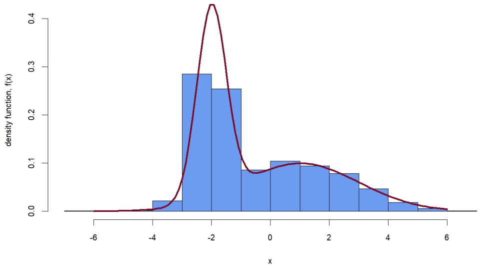
Density function: example



If X has density function $f(x) = 2|x|$, if $-1 < x < 0$, and 0 otherwise, then

$$F(-1/2) = \mathbb{P}(X \le -1/2) = \int_{-\infty}^{-1/2} f(x)\,dx = \frac{3}{4}.$$

Density function



A density function and the histogram of the corresponding sample of size 1000 – what can be the **expected value** and the **standard deviation** of a random variable with distribution function f?

Discrete and absolutely continuous case

Expected value will have a different form
in the absolutely continuous case,
because $\mathbb{P}(X = x) = 0$ holds for every x. Instead:

discrete

possible values of X: $x_1, x_2, \ldots$

$$\mathbb{E}(X) = \sum_{j=1}^{\infty} x_j \cdot \mathbb{P}(X = x_j)$$

$$\mathbb{E}(X^2) = \sum_{j=1}^{\infty} x_j^2 \cdot \mathbb{P}(X = x_j)$$

absolutely continuous

density function of X: f.

Discrete and absolutely continuous case

Expected value will have a different form
in the absolutely continuous case,
because $\mathbb{P}(X = x) = 0$ holds for every x. Instead:

discrete

possible values of X: $x_1, x_2, \ldots$

$$\mathbb{E}(X) = \sum_{j=1}^{\infty} x_j \cdot \mathbb{P}(X = x_j)$$

$$\mathbb{E}(X^2) = \sum_{j=1}^{\infty} x_j^2 \cdot \mathbb{P}(X = x_j)$$

absolutely continuous

density function of X: f.

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

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

Discrete and absolutely continuous case

Expected value will have a different form
in the absolutely continuous case,
because $\mathbb{P}(X = x) = 0$ holds for every x. Instead:

discrete

possible values of X: $x_1, x_2, \ldots$

$$\mathbb{E}(X) = \sum_{j=1}^{\infty} x_j \cdot \mathbb{P}(X = x_j)$$

$$\mathbb{E}(X^2) = \sum_{j=1}^{\infty} x_j^2 \cdot \mathbb{P}(X = x_j)$$

absolutely continuous

density function of X: f.

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

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

$$D(X) = \sqrt{\mathbb{E}\left((X - \mathbb{E}(X))^2\right)} = \sqrt{\mathbb{E}(X^2) - \mathbb{E}(X)^2}$$

Discrete and absolutely continuous case

Expected value will have a different form
in the absolutely continuous case,
because $\mathbb{P}(X = x) = 0$ holds for every x. Instead:

discrete

possible values of X: $x_1, x_2, \ldots$

$$\mathbb{E}(X) = \sum_{j=1}^{\infty} x_j \cdot \mathbb{P}(X = x_j)$$

$$\mathbb{E}(X^2) = \sum_{j=1}^{\infty} x_j^2 \cdot \mathbb{P}(X = x_j)$$

absolutely continuous

density function of X: f.

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

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

$$D(X) = \sqrt{\mathbb{E}((X - \mathbb{E}(X))^2)} = \sqrt{\mathbb{E}(X^2) - \mathbb{E}(X)^2}$$

$$\mathbb{E}(X^k) = \sum_{j=1}^{\infty} x_j^k \cdot \mathbb{P}(X = x_j)$$

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

$$\mathbb{E}(g(X)) = \sum_{j=1}^{\infty} g(x_j) \cdot \mathbb{P}(X = x_j)$$

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

Expected value and standard deviation

Let X be an absolutely continuous random variable with density function f. Then the **expected value** of X is defined as follows:

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

if this integral exists and it is finite.

Suppose that the random variable X is absolutely continuous, its density function is f, and $\mathbb{E}(X^2)$ exists, that is, the integral $\int_{-\infty}^{\infty} x^2 f(x)\, dx$ is finite. Then the **variance** of X is:

$$D^2(X) = \mathbb{E}\big((X - \mathbb{E}(X))^2\big) = \mathbb{E}(X^2) - \mathbb{E}(X)^2,$$

and its **standard deviation** is

$$D(X) = \sqrt{\mathbb{E}\big((X - \mathbb{E}(X))^2\big)} = \sqrt{\mathbb{E}(X^2) - \mathbb{E}(X)^2}.$$

The definition of standard deviation is the same as in the discrete case.

Expected value of the uniform distribution

Density function of the uniform distribution on the interval $[a, b]$:

$$f(x) = \begin{cases} \frac{1}{b-a}, & \text{if } a \leq x \leq b; \\ 0, & \text{otherwise.} \end{cases}$$

If X has uniform distribution on the interval $[a, b]$, then its expected value is:

$$\mathbb{E}(X) = \int_{-\infty}^{\infty} x \cdot f(x) dx = \int_a^b x \cdot \frac{1}{b-a} dx = \frac{1}{b-a} \left[\frac{x^2}{2} \right]_{x=a}^{b} =$$

$$= \frac{1}{b-a} \cdot \frac{b^2 - a^2}{2} = \frac{a+b}{2},$$

because the primitive function of x is $\frac{x^2}{2}$, and $b^2 - a^2 = (b-a)(b+a)$.

Standard deviation of the uniform distribution

Density function of the uniform distribution on the interval $[a, b]$:

$$f(x) = \begin{cases} \frac{1}{b-a}, & \text{if } a \leq x \leq b; \\ 0, & \text{otherwise.} \end{cases}$$

In this case, we have

$$\mathbb{E}(X^2) = \int_{-\infty}^{\infty} x^2 \cdot f(x)dx = \int_a^b x^2 \cdot \frac{1}{b-a}dx = \frac{1}{b-a}\left[\frac{s^3}{3}\right]_{s=a}^{b}$$
$$= \frac{1}{b-a} \cdot \frac{b^3 - a^3}{3} = \frac{a^2 + ab + b^2}{3},$$

as the primitive function of x^2 is $\frac{x^3}{3}$, and $b^3 - a^3 = (b-a)(a^2 + ab + b^2)$.

Standard deviation of the uniform distribution

Density function of the uniform distribution on the interval $[a, b]$:

$$f(x) = \begin{cases} \frac{1}{b-a}, & \text{if } a \le x \le b; \\ 0, & \text{otherwise.} \end{cases}$$

In this case, we have

$$\mathbb{E}(X^2) = \int_{-\infty}^{\infty} x^2 \cdot f(x)dx = \int_a^b x^2 \cdot \frac{1}{b-a} dx = \frac{1}{b-a} \left[\frac{s^3}{3} \right]_{s=a}^{b}$$
$$= \frac{1}{b-a} \cdot \frac{b^3 - a^3}{3} = \frac{a^2 + ab + b^2}{3},$$

as the primitive function of x^2 is $\frac{x^3}{3}$, and $b^3 - a^3 = (b-a)(a^2 + ab + b^2)$.

$$D^2(X) = \mathbb{E}(X^2) - \mathbb{E}(X)^2 = \frac{a^2 + ab + b^2}{3} - \left(\frac{a+b}{2} \right)^2$$
$$= \frac{a^2 + ab + b^2}{3} - \frac{a^2 + 2ab + b^2}{4} = \frac{a^2 - 2ab + b^2}{12} = \frac{(b-a)^2}{12}.$$

Uniform distribution

The random variable X has **uniform distribution** on the interval $[a, b]$, if its density function is

$$f(x) = \begin{cases} \frac{1}{b-a}, & \text{if } a \leq x \leq b; \\ 0, & \text{otherwise.} \end{cases}$$

Then

(i) the cumulative distribution function of X is as follows:

$$F(t) = \mathbb{P}(X \leq t) = \begin{cases} 0, & \text{if } t \leq a; \\ \frac{t-a}{b-a}, & \text{if } a < t < b; \\ 1, & \text{if } t \geq b. \end{cases}$$

(ii) If $a \leq c \leq d \leq b$, then

$$\mathbb{P}(c \leq X \leq d) = \int_c^d f(x)\,dx = \int_c^d \frac{1}{b-a}\,dx = \frac{d-c}{b-a}.$$

(iii) The **expected value** and **standard deviation** of X:

$$\mathbb{E}(X) = \frac{a+b}{2}; \qquad D(X) = \frac{b-a}{\sqrt{12}}.$$

Exponential distribution: definition and properties

Let $\lambda > 0$ be fixed. The random variable X has **exponential distribution** with parameter λ, if its density function is

$$f(x) = \begin{cases} \lambda e^{-\lambda x}, & \text{if } x > 0; \\ 0, & \text{otherwise.} \end{cases}$$

Then

1. the cumulative distribution function of X is the following:

$$F(t) = \mathbb{P}(X \leq t) = \int_{-\infty}^{t} f(x)\,dx = \begin{cases} 1 - e^{-\lambda t}, & \text{if } t > 0; \\ 0 & \text{otherwise.} \end{cases}$$

2. The expected value of X: $\mathbb{E}(X) = \frac{1}{\lambda}$, its standard deviation: $D(X) = \frac{1}{\lambda}$.
3. **Memoryless property.** Let s, t be positive numbers. Then we have

$$\mathbb{P}(X \geq s + t | X \geq s) = \mathbb{P}(X \geq t).$$

Normal distribution

Let m be a real number, and $\sigma > 0$. We say that the random variable Y has **normal distribution** with expected value m and variance σ^2, if its **density function** is the following:

$$f(x) = \frac{1}{\sqrt{2\pi} \cdot \sigma} \exp\left(- \frac{(x-m)^2}{2\sigma^2} \right) \qquad (x \in \mathbb{R}).$$

Notation: $Y \sim N(m, \sigma^2)$.

If $Y \sim N(m, \sigma^2)$, then $\mathbb{E}(\mathrm{Y}) = \mathrm{m}$, $\mathrm{D}(\mathrm{Y}) = \sigma$.

Standard normal distribution: normal distribution with expected value $m = 0$ and standard deviation $\sigma = 1$. **Cumulative distribution function:** Φ, density function: φ, where

$$\Phi(t) = \int_{-\infty}^{t} \varphi(x)\, dx; \qquad \varphi(x) = \frac{1}{\sqrt{2\pi}} \exp\left(- \frac{x^2}{2} \right).$$