Lecture 07. Statistics and probability theory.

Important discrete distributions

1. Two point discrete distribution: $X : S \to \{0,1\}$, $\mathrm{P}(X = 0) = 1 - p$, $\mathrm{P}(X = 1) = p$, where $0 < p < 1$,

   $\mathrm{E}X = 1 \cdot p + 0 \cdot (1-p) = p$,

   $\mathrm{D}^2[X] = \mathrm{E}X^2 - (\mathrm{E}X)^2 = p - p^2 = pq$

   usually the event $\{X = 1\}$ is called "success" and the event $\{X = 0\}$ is called "failure".

2. Binomial distribution (the Bernoulli scheme): $X : S \to \{0,1,2,3,\ldots,n\}$, $\mathrm{P}(X = k) = \binom{n}{k} p^k q^{n-k}$, where $0 < p < 1, k = 0,1,2,\ldots,n$, $q = 1 - p$, notation: $b(n,k,p) = \binom{n}{k} p^k q^{n-k}$,

   $\mathrm{E}X = np$, $\mathrm{D}^2[X] = npq$

   Example. Suppose you toss a biased coin, with probability p for "head" and $q = 1-p$ for "tail" repeatedly n times. Let X denote the number of "successes ("heads") obtained. Since then any outcome ω is a sequence of "heads" and "tails" of length n, the probability distribution function is as follows

   $$p_k = \mathrm{P}(X = k) = \binom{n}{k} p^k q^{n-k},\ k = 0,1,2,3,4,\ldots,n$$

   Theorem 1. *Let $X_1, X_2, \ldots, X_n : S \to \{0,1\}$ be independent random variables having the same two-point distribution $P(X_i = 0) = 1 - p$, $P(X_i = 1) = p$, where $0 < p < 1$, $i = 1,2,\ldots,n$. Then the random variable $S_n = X_1 + X_2 + \cdots + X_n$ has the binomial distribution.*

   Example. A die is rolled repeatedly n times. Let S_n denote the number of "sixes" obtained. In our notation $S_n = X_1 + X_2 + \cdots + X_n$, where $X_i = 1$ if the "six" ocurred in the i-th roll, and $X_i = 0$ otherwise. Our probabilistic model is the Bernoulli scheme with the probability of "succes" $p = 1/6$. We find

   $$p_k = \mathrm{P}(S_n = k) = \binom{n}{k} \left(\frac{1}{6}\right)^k \left(\frac{5}{6}\right)^{n-k}, \quad k = 0,1,2,3,4,\ldots$$

3. Geometrical distribution with success probability p:

   $X : S \to \{0, 1, 2, 3, \dots, n, \dots\}$,

   $\mathrm{P}(X = k) = p(1-p)^{k-1}$, where $0 < p < 1, k = 1, 2, 3, \dots,$

   $\mathrm{E}X = 1 \cdot p(1-p)^0 + 2 \cdot p(1-p)^1 + 3 \cdot p(1-p)^2 + \dots = \frac{1}{p}$,

   $\mathrm{D}^2[X] = \mathrm{E}X^2 - (\mathrm{E}X)^2 = (1^2 \cdot p(1-p)^0 + 2^2 \cdot p(1-p)^1 + 3^2 \cdot p(1-p)^2 + \dots) - \frac{1}{p^2} = \frac{1-p}{p^2}$.

   Example. Suppose you toss a biased coin, with probability p for "head" and $q = 1 - p$ for "tail" repeatedly until a head turns up. Let X denote the number of tosses it takes until this happens, so that $\{X = k\}$ means $k - 1$ tails before the first head. Since then the favorable outcome is just the specific sequence $(\underbrace{T, T, T, \dots, T}_{k-1}, H)$, its probability is

   $$p_k = \mathrm{P}(X = k) = q^{k-1}p, \ k = 1, 2, 3, 4, \dots$$

   The random variable X is called the waiting time, for heads to fall, or more generally for a "success."

4. The Poisson distribution with a parameter $\lambda > 0$ of the random variable $X : S \to \{0, 1, 2, 3, \dots, n, \dots\}$:

   $\mathrm{P}(X = k) = e^{-\lambda}\frac{\lambda^k}{k!}$, where $0 < p < 1, \ k = 0, 1, 2, 3, \dots,$

   $\mathrm{E}X = \lambda$, $\mathrm{D}^2[X] = \lambda$

Relation between the binomial and the Poisson distributions

Theorem 2. (Poisson) *If $n \cdot p_n \to \lambda$, when $n \to \infty$ then for every natural $k \geq 0$ it holds*

$$\lim_{n\to\infty} b(n, k, p_n) = \lim_{n\to\infty} \binom{n}{k} p_n^k (1-p_n)^{n-k} = e^{-\lambda}\frac{\lambda^k}{k!}$$

in other words:

the binomial distribution		the Poisson distribution
$P(X=1)=b(n,1,p_n)$	$\to$	$P(X=1)=e^{-\lambda}\frac{\lambda^1}{1!}$
$P(X=2)=b(n,2,p_n)$	$\to$	$P(X=2)=e^{-\lambda}\frac{\lambda^2}{2!}$
$P(X=3)=b(n,3,p_n)$	$\to$	$P(X=3)=e^{-\lambda}\frac{\lambda^3}{3!}$
$\vdots$		$\vdots$
$P(X=n)=b(n,n,p_n)$	$\to$	$P(X=n)=e^{-\lambda}\frac{\lambda^n}{n!}$
0		$P(X=n+1)=e^{-\lambda}\frac{\lambda^{n+1}}{(n+1)!}$
$\vdots$		$\vdots$

2. Continuous random variable with a density

Let $f(x)$ be a function satisfying the following conditions:

(1) $f(x) \geq 0, \text{ for } x \in R$

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

The function $f(x)$ determines the probability distribution in the sample space $S = R$ as follows

$$\mathrm{P}(A) = \int_{-\infty}^{\infty} f(x)1_A(x)dx$$

where $1_A(x)$ is a characteristic function of the set A:

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

Note

The integral

$$\int_{-\infty}^{\infty} f(x)1_A(x)dx$$

is a surface area under the graph of the function $f(x)$ over the set A. Such an integral exists if and only if the set is so called the Borel set, which means that the probability $\mathrm{P}(A)$ is defined only for such sets (events).

The random variable $X : S \to R$ is called a continuous random variable with the density $f(x)$, if its distribution is gven as follows

$$\mathrm{P}(X \in A) = \int_{-\infty}^{\infty} f(x) 1_A(x) dx$$

Example. .

Let $A = [a, b]$, (a, b), $(a, b]$ or $[a, b)$, $a, b \in R$. Then $\mathrm{P}(X \in A) = \int_a^b f(x)dx$.
Let $A = \{a\} = [a, a]$, $a \in R$. Then $\mathrm{P}(X = a) = 0$.

The cumulative distribution function $F(x)$ of the continuous random variable X

Let $f(x)$ be the density function. Then the following relations hold

$$F(x) \stackrel{\text{def}}{=} \mathrm{P}(X \le x) = \int_{-\infty}^{x} f(z)dz, \quad f(x) = F'(x)$$

Example. The continuous random variable X uniformly distributed in the interval $[a, b]$. In this case the density is defined as follows

$$f(x) = \begin{cases} \frac{1}{b-a}, & \text{if } x \in [a, b] \\ 0, & \text{if } x \notin [a, b] \end{cases}$$

Then

$$\mathrm{P}(X \in [c, d]) = \int_c^d f(x)dx = \frac{d-c}{b-a}$$

for any interval $[c, d] \subseteq [a, b]$.

Example. The normal distribution with the parameters μ and $\sigma > 0$. Let

$$f(x) = \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{(x-\mu)^2}{2\sigma^2}}.$$

The function $f(x)$ is a density of the normal distribution with parameters μ and $\sigma > 0$. This distribution is denoted as $\mathcal{N}(\mu, \sigma)$. The distribution $\mathcal{N}(0, 1)$

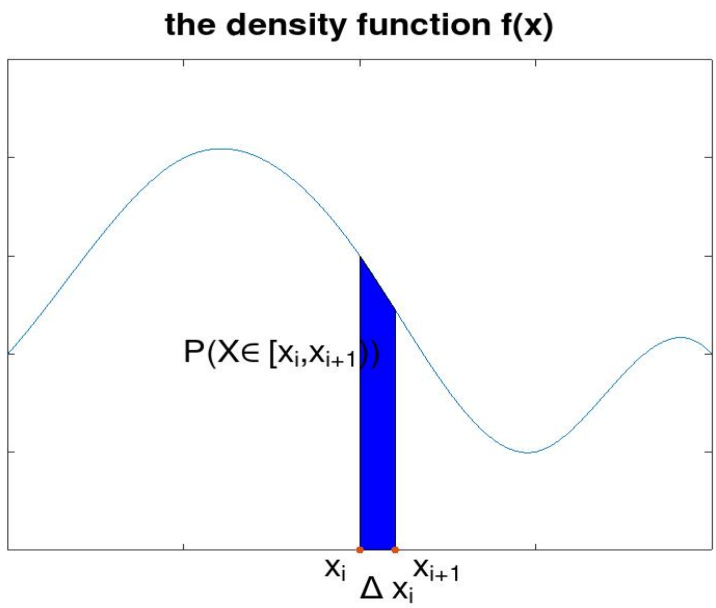
is called the standard normal distribution.

The fundamental characteristics of the continuous random variable X with the density $f(x)$:

- **Mean value (Expected value):** $\mu = \mathrm{E}X = \int\limits_{-\infty}^{\infty} xf(x)dx$ (if the integral exists)

  Motivation for the formula: Let $-\infty < x_1 < x_2 < x_3 < \cdots < x_n < \infty$

  $$\int\limits_{-\infty}^{\infty} xf(x)dx \approx \sum_{n=1}^{\infty} x_i \cdot f(x_i)\Delta x_i$$

  However

  $$p_i = f(x_i)\Delta x_i \approx \mathrm{P}(X \in [x_i, x_{i+1}]) \longrightarrow \mathrm{P}(X = x_i)$$
  $$\int\limits_{-\infty}^{\infty} xf(x)dx \approx \sum_{n=1}^{\infty} x_i \cdot p_i$$

  When $x_1 \to -\infty$, $x_n \to \infty$ and $\max \Delta x_i \to 0$, then this approximation becomes more accurate and finally we get the expected result.

Remark. Let $g : R \to R$ be a function and let $Y = g(X)$. Then

$$\mathrm{E}Y = \mathrm{E}g(X) = \int_{-\infty}^{\infty} g(x)f(x)dx.$$

- **Variance:** $\mathrm{D}^2[X] = \sigma^2 = \mathrm{E}(X-\mu)^2 = \int_{-\infty}^{\infty} (x-\mu)^2 f(x)dx$, the squared root of the variance, σ is called the standard deviation of X.

Example. Let X be a continuous variable with the normal distribution $\mathcal{N}(\mu, \sigma)$. The letters μ and σ are not used accidentially

- **Mean value (Expected value):** $\mathrm{E}X = \int_{-\infty}^{\infty} x\frac{1}{\sigma\sqrt{2\pi}}e^{-\frac{(x-\mu)^2}{2\sigma^2}}dx$, the calculation of the integral

$$\int_{-\infty}^{\infty} x\frac{1}{\sigma\sqrt{2\pi}}e^{-\frac{(x-\mu)^2}{2\sigma^2}}dx, \quad z = \frac{x-\mu}{\sigma},\ x = \sigma z + \mu,\ dx = \sigma dz$$

$$\int_{-\infty}^{\infty} x\frac{1}{\sigma\sqrt{2\pi}}e^{-\frac{(x-\mu)^2}{2\sigma^2}}dx = \int_{-\infty}^{\infty} (\sigma z + \mu)\frac{1}{\sigma\sqrt{2\pi}}e^{-\frac{z^2}{2}}\sigma dz =$$

$$\sigma\int_{-\infty}^{\infty} z\frac{1}{\sqrt{2\pi}}e^{-\frac{z^2}{2}}dz + \mu\int_{-\infty}^{\infty} \frac{1}{\sqrt{2\pi}}e^{-\frac{z^2}{2}}dz = 0 + \mu = \mu$$

Since it is known that

$$\int_{-\infty}^{\infty} z\frac{1}{\sqrt{2\pi}}e^{-\frac{z^2}{2}}dz = 0 \text{ and } \int_{-\infty}^{\infty} \frac{1}{\sqrt{2\pi}}e^{-\frac{z^2}{2}}dz = 1$$

the calculations give $\mathrm{E}X = \mu$

- **Variance:** $\mathrm{D}^2[X] = \mathrm{E}(X-\mu)^2 = \int_{-\infty}^{\infty} (x-\mu)^2\frac{1}{\sigma\sqrt{2\pi}}e^{-\frac{(x-\mu)^2}{2\sigma^2}}dx$, the calculation of the integral

$$\int_{-\infty}^{\infty} (x-\mu)^2 \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{(x-\mu)^2}{2\sigma^2}} dx, \quad z = \frac{x-\mu}{\sigma},\ x = \sigma z + \mu,\ dx = \sigma dz$$

$$\int_{-\infty}^{\infty} (x-\mu)^2 \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{(x-\mu)^2}{2\sigma^2}} dx = \int_{-\infty}^{\infty} \sigma^2 z^2 \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{z^2}{2}} \sigma dz =$$

$$\sigma^2 \int_{-\infty}^{\infty} z^2 \frac{1}{\sqrt{2\pi}} e^{-\frac{z^2}{2}} dz$$

Since it is known that

$$\int_{-\infty}^{\infty} z^2 \frac{1}{\sqrt{2\pi}} e^{-\frac{z^2}{2}} dz = 1$$

the calculations give $\mathrm{D}^2\left[X\right] = \sigma^2$

Complementions

1. $\int\limits_{-\infty}^{\infty} \frac{1}{\sqrt{2\pi}} e^{-\frac{z^2}{2}} dz = 1:$

$1.a) \int\limits_{-\infty}^{\infty}\int\limits_{-\infty}^{\infty} e^{-\frac{u^2+v^2}{2}} dudv,$

repacement: $u = r\cos\phi,\ v = r\sin\phi,\ 0 \le \phi < 2\pi,\ 0 \le r < \infty,\ dudv = rdrd\phi$

$= \int\limits_{0}^{2\pi} d\phi \int\limits_{0}^{\infty} e^{-\frac{r^2}{2}} rdr = 2\pi$

$1.b)\ 2\pi = \int\limits_{-\infty}^{\infty}\int\limits_{-\infty}^{\infty} e^{-\frac{u^2+v^2}{2}} dudv = \int\limits_{-\infty}^{\infty}\int\limits_{-\infty}^{\infty} e^{-\frac{u^2}{2}} e^{-\frac{v^2}{2}} dudv = \left(\int\limits_{-\infty}^{\infty} e^{-\frac{z^2}{2}} dz \right)^2$

hence: $\int\limits_{-\infty}^{\infty} \frac{1}{\sqrt{2\pi}} e^{-\frac{z^2}{2}} dz = 1$

2. $\int\limits_{-\infty}^{\infty} z \frac{1}{\sqrt{2\pi}} e^{-\frac{z^2}{2}} dz = 0:$

$$\int\limits_{-\infty}^{\infty} z \frac{1}{\sqrt{2\pi}} e^{-\frac{z^2}{2}} dz = \int\limits_{-\infty}^{0} z \frac{1}{\sqrt{2\pi}} e^{-\frac{z^2}{2}} dz + \int\limits_{0}^{\infty} z \frac{1}{\sqrt{2\pi}} e^{-\frac{z^2}{2}} dz =$$

$$\int\limits_{\infty}^{0} (-u) \frac{1}{\sqrt{2\pi}} e^{-\frac{u^2}{2}} d(-u) + \int\limits_{0}^{\infty} z \frac{1}{\sqrt{2\pi}} e^{-\frac{z^2}{2}} dz = -\int\limits_{0}^{\infty} u \frac{1}{\sqrt{2\pi}} e^{-\frac{u^2}{2}} du +$$

$$\int\limits_{0}^{\infty} z \frac{1}{\sqrt{2\pi}} e^{-\frac{z^2}{2}} dz = 0$$

3. $\int\limits_{-\infty}^{\infty} z^2 \frac{1}{\sqrt{2\pi}} e^{-\frac{z^2}{2}} dz = 1:$

$$\int\limits_{-\infty}^{\infty} z^2 e^{-\frac{z^2}{2}} dz = \int\limits_{-\infty}^{\infty} z \cdot z e^{-\frac{z^2}{2}} dz = \underbrace{z(-e^{-\frac{z^2}{2}})\Big|_{-\infty}^{\infty}}_{=0} + \underbrace{\int\limits_{-\infty}^{\infty} 1 \cdot e^{-\frac{z^2}{2}} dz}_{=\sqrt{2\pi}} = \sqrt{2\pi}$$

Rules for means

Rule 1. If X is a random variable and a and b are fixed numbers, then

$$\mu_{a+bX} = \mathrm{E}(a+bX) = a + b\mathrm{E}X = a + b\mu_X.$$

Rule 2. If X and Y are random variables, then

$$\mu_{X+Y} = \mathrm{E}(X+Y) = \mathrm{E}X + \mathrm{E}Y = \mu_X + \mu_Y.$$

Rules for variances and standard deviations

Rule 0. Formulas for the variance

$$\sigma^2 = \mathrm{D}^2[X] = \mathrm{E}(X - \mathrm{E}X)^2 = \mathrm{E}(X^2 - 2X \cdot \mathrm{E}X + (\mathrm{E}X)^2) = \mathrm{E}X^2 - (\mathrm{E}X)^2.$$

Rule 1. If X is a random variable and a and b are fixed numbers, then

$$\sigma^2_{a+bX} = b^2\sigma^2_X$$

Rule 2. If X and Y are independent random variables, then

$$\sigma^2_{X+Y} = \sigma^2_X + \sigma^2_Y$$
$$\sigma^2_{X-Y} = \sigma^2_X + \sigma^2_Y$$

This is the addition rule for variances of independent random variables.

Rule 3. Let ρ be the correlation coefficient of X and Y. Then

$$\rho = \frac{1}{\sigma_X\sigma_Y}\mathrm{E}(X - \mathrm{E}X)(Y - \mathrm{E}Y)$$

$$\sigma^2_{X+Y} = \sigma^2_X + \sigma^2_Y + 2\rho\sigma_X\sigma_Y$$
$$\sigma^2_{X-Y} = \sigma^2_X + \sigma^2_Y - 2\rho\sigma_X\sigma_Y$$

This is the general addition rule for variances of random variables. To find the **standard deviation**, take the square root of the variance.

Important continuous distributions

1. Uniform distribution: $X : S \to [a,b]$, the distribution density $f(x) = \frac{1}{b-a}$ for $x \in [a,b]$, $f(x) = 0$ for $x \notin [a,b]$, the cumulative distribution function

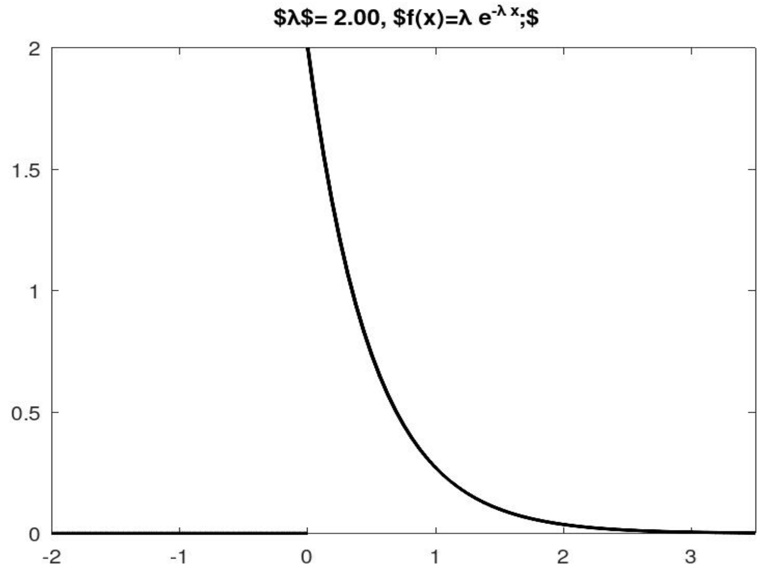
$$F(x) = \begin{cases} 0 & \text{for } x < a, \\ \frac{x-a}{b-a} & \text{for } x, \in [a,b], \\ 1 & \text{for } x > 1, \end{cases}$$

$\mathrm{E}X = \frac{a+b}{2}$, $\mathrm{D}^2[X] = \frac{(b-a)^2}{12}$.

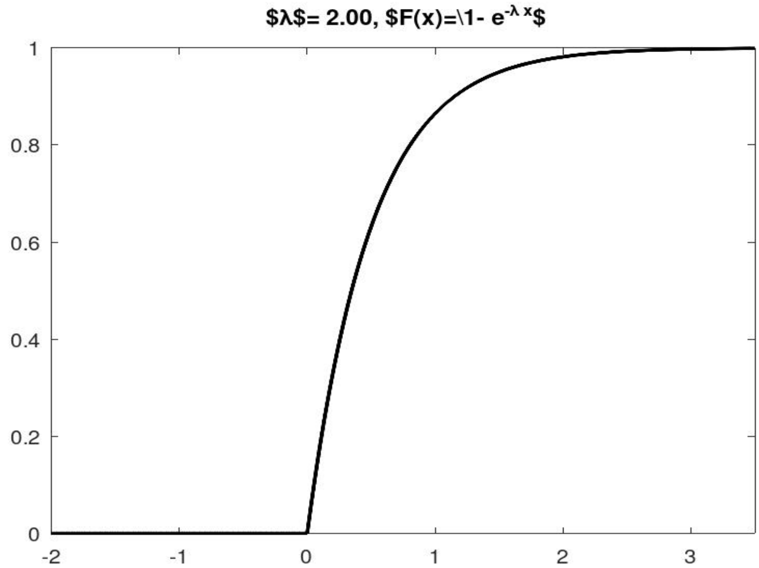
2. Exponential distribution with a parameter $\lambda > 0$: $X : S \to [0, \infty)$, the distribution density $f(x) = 0$ dla $x < 0$, $f(x) = \lambda e^{-\lambda x}$ dla $x \geq 0$, the cumulative distribution function

$$F(x) = \begin{cases} 0 & \text{dla } x < 0 \\ 1 - e^{-\lambda x} & \text{dla } x \geq 0 \end{cases}$$

$\mathrm{E}X = \frac{1}{\lambda}$, $\mathrm{D}^2[X] = \frac{1}{\lambda^2}$.



the density function



the cumulative distribution function

Random variables with the exponential distributions are useful in description of queueing problems, because of the following property of the exponential distribution

Theorem 3. *If T is a random variable with the exponential distribution, then $P(T > t + t_0 | T > t_0) = P(T > t)$*
(We've been waiting in line for at least t_0 minutes now, but that doesn't affect how long it will take)
Justification:

$$P(T > t) = \int_t^\infty \lambda e^{-\lambda z} dz = e^{-\lambda t}$$

$$\{T > t + t_0\} \subseteq \{T > t_0\} \Longrightarrow \{T > t + t_0\} \cap \{T > t_0\} = \{T > t + t_0\}$$

$$\mathrm{P}(T > t + t_0 | T > t_0) = \frac{\mathrm{P}(T > t + t_0)}{\mathrm{P}(T|t_0)} = \frac{e^{-\lambda(t+t_0)}}{e^{-t_0}} = e^{-\lambda t} = \mathrm{P}(T > t)$$

3. The Weibull distribution with a parameters $\lambda, \alpha > 0$ of the random variable $X : S \to [0, \infty)$: the distribution density $f(x) = 0$ dla $x < 0$, $f(x) = \alpha\lambda e^{-\lambda x^\alpha}$ for $x \geq 0$, the cumulative distribution function

$$F(x) = \begin{cases} 0 & \text{dla } x < 0 \\ 1 - e^{-\lambda x^\alpha} & \text{dla } x \geq 0 \end{cases}$$

4. The normal distribution $\mathcal{N}(\mu, \sigma)$, $\sigma > 0$ of the random variable $X : S \to R$, the distribution density function $f(x) = \frac{1}{\sigma\sqrt{2\pi}} e^{-(x-\mu)^2/2\sigma^2}$ dla $x \in \mathbb{R}$, the cumulative distribution function

$$F(x) = \frac{1}{\sigma\sqrt{2\pi}} \int_{-\infty}^{x} e^{-(t-\mu)^2/2\sigma^2} dt$$

$\mathrm{E}X = \mu$, $\mathrm{D}^2[X] = \sigma^2$.

The standarization of the random variable X with the normal distribution $\mathcal{N}(\mu, \sigma)$:

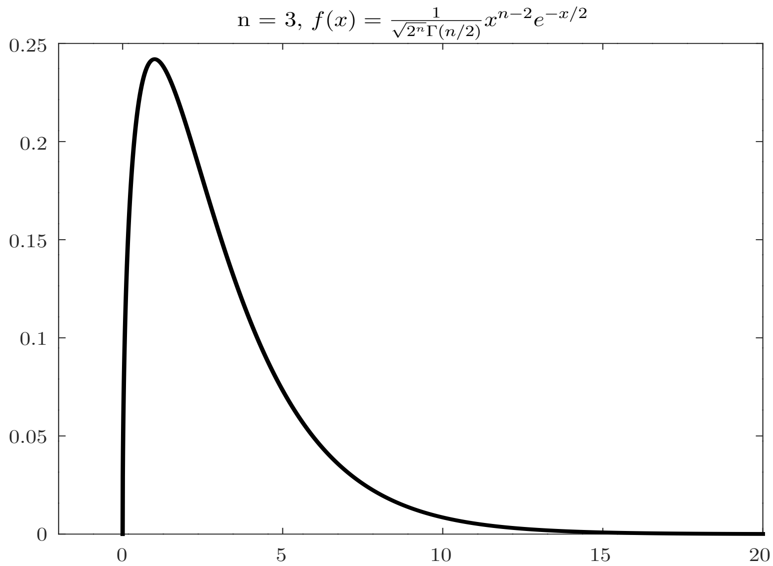
$$X \sim \mathcal{N}(\mu, \sigma) \longrightarrow \widetilde{X} = \frac{X - \mu}{\sigma} \sim \mathcal{N}(0, 1).$$

The very useful property of random variables with the normal distributions:

Theorem 4. *Let X_1, X_2,..., X_n be independent random variables with normal distribution. Then the random variable $X = X_1 + X_2 + X_3 + \cdots + X_n$ has also a normal distribution*

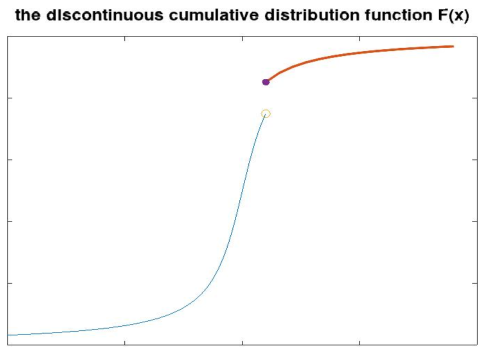
5. The probability distribution χ^2_n. This is the distribution of the random variable $X = X_1^2 + X_2^2 + \cdots + X_n^2$, where X_i, $i = 1, 2, 3, \ldots, n$ are independent random variables with the standard normal distribution $\mathcal{N}(0, 1)$, it is called the probability distribution χ^2 with n steps of freedom and it is denoted as χ^2_n. It is one of the important probability distributions in statistics. Its distribution density function is the function

$$\gamma_{n/2,1/2}(x) = \frac{1}{\sqrt{2^n}\Gamma(\frac{n}{2})} x^{\frac{n-2}{2}} e^{-\frac{x}{2}} 1_{(0,\infty)}(x).$$



Remark. There are random variables that are neither continuous and nor discrete.

Example. The function with the discontinuous graph



is a cumulative distribution function $F(x)$ of the random variable that is neither continuous and nor discrete.