§9. Confidence intervals

s.1. Basic definitions.

A confidence interval gives an estimated range of values which is likely to include an unknown population parameter, the estimated range being calculated from a given set of sample data.

Let be a point estimate of the population parameter calculated from a given set of sample data.

Let and

The less is the more accurate the estimation is.

Remark.

The statistical methods do not allow us to say that the estimator satisfies the inequality

We can only say about probability with which this inequality holds.

Confidence coefficient (confidence level) is a probability with which the inequality takes place, i.e.

The interval

which covers the unknown parameter with prescribed probability is called confidence interval (CI).

— estimation accuracy.

The construction of the CI:

1) point estimate calculation;

2) the choice of confidence level (0,95; 0,99; 0,995);

3) the calculation of the accuracy .

s.2. Distributions of the RV, which are often used in statistics.

Chi-squared distribution

Let RV are independent and

Then RV

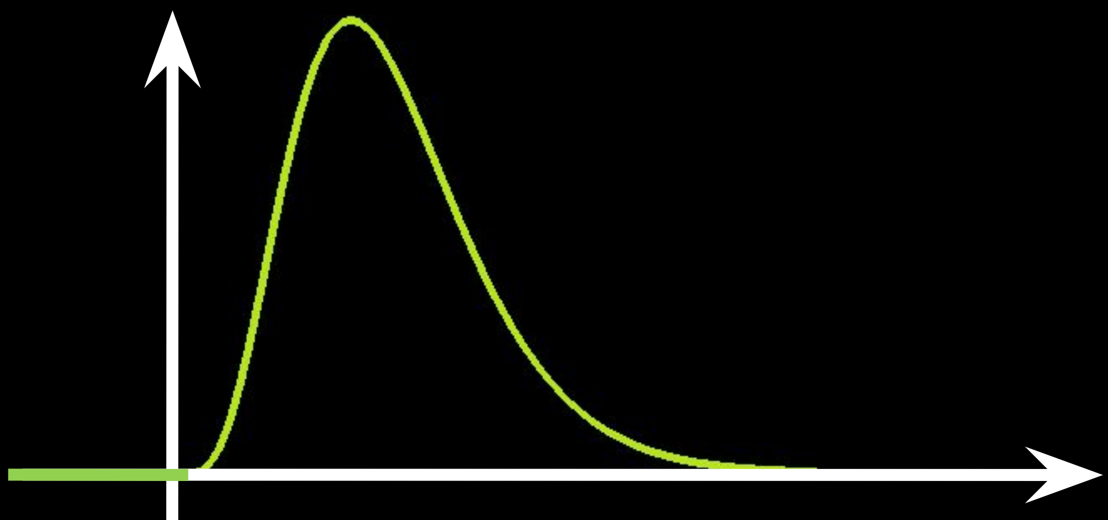
is called to be distributed according to the chi-squared distribution with k degrees of freedom.

Probability density function:

where

— Gamma function.

The plot of



The expected value:

The variance:

The quantile of the distribution, which corresponds the statistical significance , is a such value that the following inequality holds:

Remark.

The values of the quantiles can be found in special tables.

Student's t-distribution (t-distribution)

Let

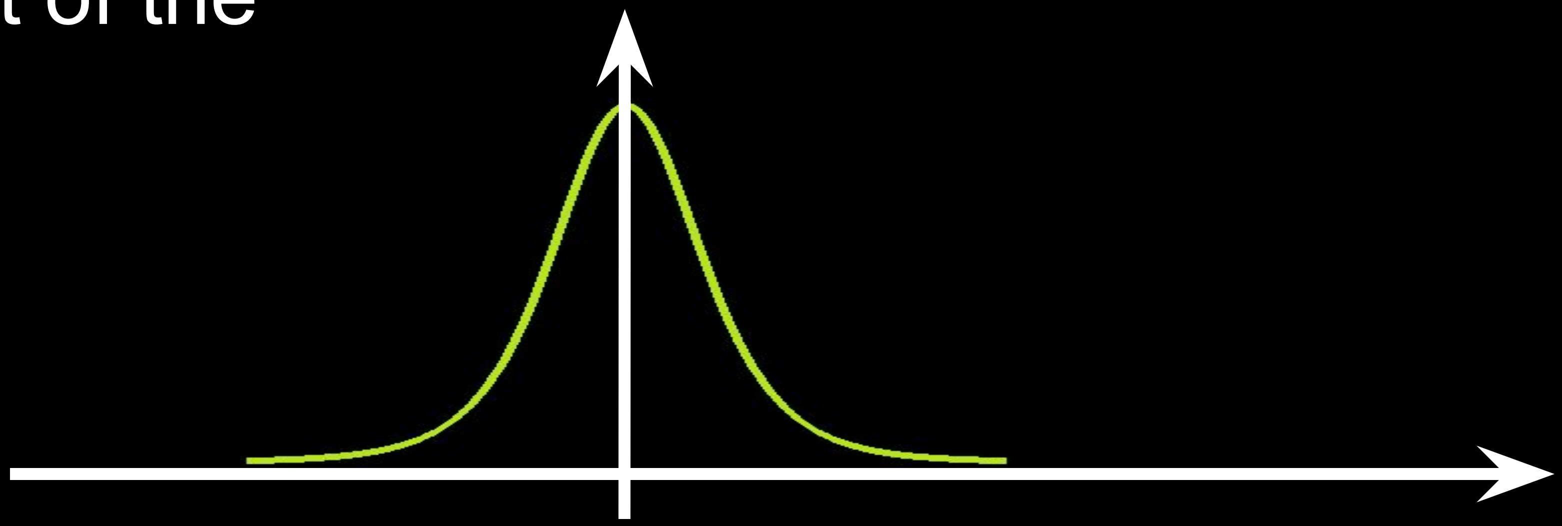
RV has the chi-squared distribution with k degrees of freedom.

Then the RV

is called to be distributed according to the t-distribution with k degrees of freedom.

Probability density function:

The plot of the



The expected value:

The variance:

The quantile of the t-distribution, which corresponds the statistical significance , is a such value that the following inequality holds:

Remark.

The values of the quantiles can be found in special tables.

s.3. Confidence Intervals for Unknown Mean and Known Standard Deviation.

Let

We know

We should find the CI for the a with confidence level .

The point estimate for the mean is

Let's find accuracy .

Let

is a sample obtained from the observations for the RV X.

The values

change from sample to sample.

Therefore, we can assume that

Besides, the sample mean is also a RV which has normal distribution, and i.e.

Since

then

Let us denote

Then

and

Therefore

i.e. with confidence level we can assert than CI

covers unknown parameter a, and the accuracy of the estimation is

Example. Let we have sample of the RV

Find 95% confidence interval for the mean.

Solution.

Confidence interval:

s.4. Confidence Intervals for Unknown Mean and Unknown Standard Deviation**.**

Let

We know

We should find the CI for the a with confidence level .

The point estimate for the mean is

Let’s find accuracy .

Let S is a standard error.

Consider the following RV

We can prove that T has t-distribution with degrees of freedom.

Let's find so that

Let us divide the both sides of the inequality in brackets on

or

Let us denote

Then

The value can be determined by the t-distribution table.

Therefore

or

i.e. with confidence level we can assert than CI

covers unknown parameter a, and the accuracy of the estimation is

Example. In the previous example find CI for the unknown mean, if standard deviation is unknown.

Solution.

Since the confidence level is

then statistical significance

The amount of the degrees of freedom

Using the special table

Accuracy

Confidence interval: