Suppose that we have calculated an average of a number of values from a process. We want to estimate the true average μ of the process using the calculated result which is based on data containing randomness. Thus our estimation will not be μ exactly.

If we let the (green) dot in every diagram be our calculated average, we find that there are many values of μ that might give the average that we did get from our data.

The eight distributions symbolize only few of all possible distributions. The two (red) distributions (upper left and lower right) can however be regarded as some kind of extreme possibilities.

If the green dot lies in right '2.5%'-tail of the left distribution and in the left '2.5%'-tail of the right distribution respectively, we state that the difference between the two 'μ'-values of the distributions forms a '95%' *confidence interval* for the theoretical average (μ) of the process.

Such an interval covers the true value of the parameter being studied with 95% probability. If we want a more secure interval, e.g. 99%, we must let the green dot enter the '0.5%'-tails of the distributions. In other words we must pull the left distribution further to the left and the right distribution further to the right. In this way we get an interval that with a higher probability cover the true value of the process; although to the price of a longer interval.

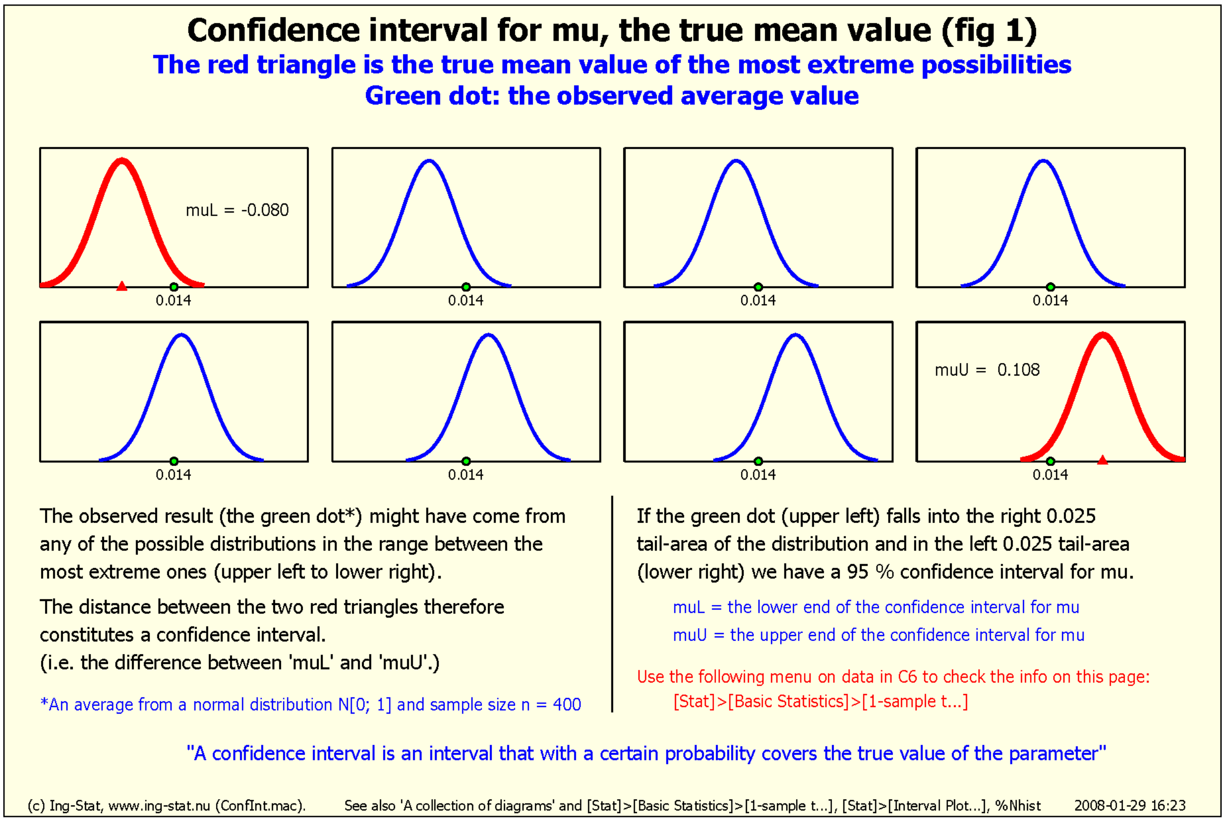
Calculating and reporting a confidence interval is an honest way to exhibit the available data and shows what conclusions can be (or can not be) drawn.

Macros and menus
%Nhist, %Boot1, %Boot2

[Stat]>[Basic Statistics]>[1-Sample t…]

Data on the printout

There is no data on the printout



%ConfInt • Confidence interval for mu, the true mean value (fig 2)

A *confidence interval* is an interval which with a stated probability covers the true value of the parameter being studied, in this case μ, the theoretical mean of the process. (This parameter is also called the expected value of the process.)

The length of the interval depends on the standard deviation s, number of data values n and the wanted confidence level given by the constant t according to the following expression:

$$\bar{x} \pm t \cdot \frac{s}{\sqrt{n}}$$

The diagram shows a number (2 - 100) of simulated confidence interval based on indata from the start of the routine. By using different, more or less extreme input values (e.g. $n = 2$), it is possible to see some of the features of the confidence interval. We expect, when calculating an e.g. 95% confidence interval, that 5% of the intervals miss the true value of the parameter. Of course, we never know whether an interval, calculated from real data, covers the true value.

It is also possible, using more or less complicated mathematics, to calculate intervals for other parameters such as the median, the standard deviation, etc.

Data on the printout

Expected value	=	Theoretical average (μ)
Standard dev.	=	Theoretical standard deviation (σ)
Sample size (n)	=	Sample size (n)
No. of intervals (m)	=	Number of simulated intervals
Confidence level	=	Confidence level

Macros and menus
%Nhist, %Boot1, %Boot2

[Stat]>[Basic Statistics]>[1-Sample t…]