SEARCH   Search ›

Home » Support

Frequently Asked Questions

SEARCH FOR FAQ

Search

The distinction between confidence intervals, prediction intervals and tolerance intervals.

FAQ# 1506   Last Modified 1-July-2009

 When you fit a parameter to a model, the accuracy or precision can be expressed as a confidence interval, a prediction interval or a tolerance interval. The three are quite distinct. The discussion below explains the three different intervals for the simple case of fitting a mean to a sample of data (assuming sampling from a Gaussian distribution). The same ideas can be applied to intervals for any best-fit parameter determined by regression.

Confidence intervals tell you about how well you have determined the mean. Assume that the data really are randomly sampled from a Gaussian distribution. If you do this many times, and calculate a confidence interval of the mean from each sample, you'd expect about 95 % of those intervals to include the true value of the population mean. The key point is that the confidence interval tells you about the likely location of the true population parameter.

Prediction intervals tell you where you can expect to see the next data point sampled. Assume that the data really are randomly sampled from a Gaussian distribution. Collect a sample of data and calculate a prediction interval. Then sample one more value from the population. If you do this many times, you'd expect that next value to lie within that prediction interval in 95% of the samples.The key point is that the prediction interval tells you about the distribution of values, not the uncertainty in determining the population mean.

Prediction intervals must account for both the uncertainty in knowing the value of the population mean, plus data scatter. So a prediction interval is always wider than a confidence interval.

Before moving on to tolerance intervals, let's define that word 'expect' used in defining a prediction interval. It means there is a 50% chance that you'd see the value within the interval in more than 95% of the samples, and a 50% chance that you'd see the value within the interval in less than 95% of the samples. Imagine doing lots of simulations, so you know the true value and thus know if it is in the prediction interval or not. You can then tabulate what fraction of the time the value is enclosed by the interval. Repeat with many sets of simulations. On average the value will be 95%, but it might be 93% or 98%. Half the time it will be less than 95% and half the time it will be more than 95%.

What if you want to be 95% sure that the interval contains 95% of the values? Or 90% sure that the interval contains 99% of the values? Those latter questions are answered by a **tolerance interval**. To compute, or understand, a tolerance interval you have to specify two different percentages. One expresses how sure you want to be, and the other expresses what fraction of the values the interval will contain. If you set the first value (how sure) to 50%, then a tolerance interval is the same as a prediction interval. If you set it to a higher value (say 90% or 99%) then the tolerance interval is wider.
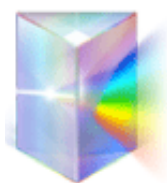
Need to learn Prism 6?

These guided examples of common analyses will get you off to a great start! **CLICK HERE > (/scientific-software/prism/?tab=5)**

(/data-analysis-resource-center/)

Questions about curve fitting or statistics?
More >

On-site training
LEARN MORE > (/support/on-site-training)