Prediction Intervals

Remember that Michelson was trying to estimate the speed of light using measurements taken during his experiments. Michelson was trying to estimate the true, unknown mean. However, you're not always interested in estimating the mean. Consider this scenario: Suppose that, after Michelson gathered his first 20 measurements, he wanted to predict his next measurement. Estimating the mean of all measurements is a little different from estimating the next measurement.

You use a confidence interval to estimate a population parameter. If you want to estimate the next value, or the next set of values, you would construct a prediction interval.

Here's a simple analogy. Many nations provide estimates of the average fuel efficiency and carbon dioxide emissions for each model of automobile sold, and require cars to have labels displaying the estimates.

These labels generally have a disclaimer, such as "Your actual consumption and emissions will vary." Your fuel economy will probably deviate from the population mean, depending on a variety of factors. In short, there is more uncertainty involved in estimating an individual observation than in estimating the mean of all observations.

A prediction interval is centered at the point estimate, exactly like a confidence interval. However, the upper and lower bounds are pushed farther from that point estimate to capture the additional uncertainty.

Let's consider a new scenario involving the diameter of parts, measured in millimeters. The parts are selected in rational subgroups, with five parts per subgroup. Data are collected on 20 subgroups.

You use an X-Bar and R chart, and determine that the process is stable. This tells you that the process centering and spread are predictable.

Here is the histogram and summary statistics for the data. Based on these data, you want to estimate the range of possible values for the next part produced, or for the next subgroup of parts produced.

When you construct a prediction interval, you need to specify both the confidence level and the number of future samples to be included in the interval.

For example, you might want to construct a prediction interval for the next observation. Or for the next five observations.

How do you interpret these intervals?

You are 95% confident that the diameter for the next part will fall between 16.02 and 16.26 millimeters. Likewise, you are 95% confident that the diameter measurements for the next five parts will fall between 15.99 and 16.30 millimeters.

If your parts do not fall within this interval, you have a good indication that the process might have changed.

This video does not include the details of how to compute a prediction interval because the software does this for us. You learn how to construct prediction intervals using JMP in an upcoming video.

One final note before we move on. When you construct a prediction interval, there is an assumption that your data are approximately normally distributed. However, the discussion of prediction intervals for non-normal data is beyond the scope of this course.

For more details about constructing and using prediction intervals, see the Read About It for this module.

Copyright © 2020 SAS Institute Inc., Cary, NC, USA. All rights reserved.

Close