The standard confidence interval for positive parameters

Bo Lindqvist

The standard confidence interval

Let \( \hat{\theta} \) be the MLE of a (one-dimensional) parameter \( \theta \). Then general theory (beyond this course) states that, if the number of observations \( n \) is large,

\[
\frac{\hat{\theta} - \theta}{\text{SD}(\hat{\theta})} \approx N(0,1) \quad \text{(approximately)}
\]

This approximation still holds if we replace \( \text{SD}(\hat{\theta}) \) by an estimate \( \hat{\text{SD}}(\hat{\theta}) \), obtained for example by replacing the \( \theta \) appearing in the expression for \( \text{SD}(\hat{\theta}) \) by \( \hat{\theta} \).

It follows that

\[
P(-1.96 < \frac{\hat{\theta} - \theta}{\text{SD}(\hat{\theta})} < 1.96) \approx 0.95
\]

Rearranging the inequalities within \( P(\cdot) \) we get

\[
P(\hat{\theta} - 1.96 \hat{\text{SD}}(\hat{\theta}) < \theta < \hat{\theta} + 1.96 \hat{\text{SD}}(\hat{\theta})) \approx 0.95 \tag{1}
\]

which defines the standard 95\% confidence interval for \( \theta \):

\[
\hat{\theta} \pm 1.96 \hat{\text{SD}}(\hat{\theta}).
\]

(Of course we may change the 1.96 to obtain other percentages than 95\%.)

The standard confidence interval for positive parameters

This method is used by MINITAB to compute confidence intervals for any positive parameter.

We will use the following property of MLE:

- If \( \hat{\theta} \) is the MLE of a parameter \( \theta \) and \( g(x) \) is some function, then \( g(\hat{\theta}) \) is the MLE of \( g(\theta) \) (the theorem of substitution).

By the approximate normality of any MLE (see beginning of this note), and the fact that \( \ln \hat{\theta} \) must be the MLE of \( \ln \theta \) by the just mentioned property of MLE, it follows that

\[
\frac{\ln \hat{\theta} - \ln \theta}{\text{SD}(\ln \hat{\theta})} \approx N(0,1) \tag{2}
\]
But from Taylor expansion of the natural logarithm we have
\[ \ln \hat{\theta} \approx \ln \theta + \frac{1}{\theta} (\hat{\theta} - \theta), \]
and therefore,
\[ \text{Var}(\ln \hat{\theta}) \approx \frac{\text{Var}(\hat{\theta})}{\theta^2} \quad \text{and hence} \quad \text{SD}(\ln \hat{\theta}) \approx \frac{\text{SD}(\hat{\theta})}{\theta}. \]

Substituting this in equation (2) we get
\[ \frac{\ln \hat{\theta} - \ln \theta}{\text{SD}(\hat{\theta})/\theta} \approx N(0,1) \]
and by estimating the standard deviation of \( \ln \hat{\theta} \) we get further
\[ \frac{\ln \hat{\theta} - \ln \theta}{\text{SD}(\hat{\theta})/\hat{\theta}} \approx N(0,1) \]

In a similar way as in equation (1) we now get
\[ P(\ln \hat{\theta} - 1.96 \frac{\text{SD}(\hat{\theta})}{\hat{\theta}} < \ln \theta < \ln \hat{\theta} + 1.96 \frac{\text{SD}(\hat{\theta})}{\hat{\theta}}) \approx 0.95. \]

The inequalities inside the \( P(\cdot) \) are of course equivalent to the ones we get by taking the exponential functions of all terms, i.e. we have,
\[ P(e^{\ln \hat{\theta} - 1.96 \frac{\text{SD}(\hat{\theta})}{\hat{\theta}}} < e^{\ln \theta} < e^{\ln \hat{\theta} + 1.96 \frac{\text{SD}(\hat{\theta})}{\hat{\theta}}}) \approx 0.95 \]
or
\[ P(\hat{\theta}e^{-1.96 \frac{\text{SD}(\hat{\theta})}{\hat{\theta}}} < \theta < \hat{\theta}e^{1.96 \frac{\text{SD}(\hat{\theta})}{\hat{\theta}}}) \approx 0.95. \]

This defines the standard interval for positive parameters, which in short be written
\[ \hat{\theta} e^{\pm 1.96 \frac{\text{SD}(\hat{\theta})}{\hat{\theta}}}. \]