Confidence Intervals for Coefficients

We can also find confidence intervals for the regression coefficients, \( \beta_i \) for \( i = 0, \ldots, k \), on the assumption that the residuals are independent and normally distributed.

To find a level \( C \) confidence interval, we first calculate
- \( b_i \), the least squares estimate of \( \beta_i \).
- \( SE(b_i) \), the standard error for \( b_i \), calculated from the \( x \) values, and \( s \), the estimated standard deviation of the residuals.
- \( t^* \), the value such that the area under the \( t \) distribution density curve with \( n-k-1 \) df between \(-t^*\) and \( t^* \) is \( C \).

Using these values, we compute the level \( C \) confidence interval for \( \beta_i \) as
\[
(b_i - t^*SE(b_i), \ b_i + t^*SE(b_i))
\]

Predicting the Response in a New Case When the True Parameters are Known

Suppose we find out the values of the explanatory variables for a new case, and wish to predict the response variable.

If the explanatory variables for the new case are \( x'_1, \ldots, x'_k \), and if we knew the true values of the regression coefficients \( \beta_i \), we would predict the response, \( y \), by its mean:
\[
\mu_y = \beta_0 + \beta_1 x'_1 + \cdots + \beta_p x'_k
\]

We could express the uncertainty in this prediction by the standard deviation of the residuals, whose true value is \( \sigma \).

C. I. for the Mean Response

In practice, we don’t know the \( \beta_i \), so we don’t know \( \mu_y \).

We can find a level \( C \) confidence interval for \( \mu_y \) however, as follows:
\[
(\hat{y} - t^*SE(\hat{y}), \ \hat{y} + t^*SE(\hat{y}))
\]

where \( \hat{y} \) is our estimate of the mean response at \( x' \) based on our estimates for the regression coefficients:
\[
\hat{y} = b_0 + b_1 x'_1 + \cdots + b_p x'_k
\]
and \( SE(\hat{y}) \) is our estimate of the standard deviation of \( \hat{y} \), which will depend on \( x' \), on the \( x \) values in the observed cases, and on the estimated standard deviation of the residuals, \( s \).

As before, \( t^* \) is the value such that \( C \) is the area between \(-t^*\) and \( t^* \) that lies under the \( t \) distribution density curve with \( n-k-1 \) df.

C. I. for a New Observation

We can also find a level \( C \) prediction interval for a new observation, \( y \), at \( x' \). This is not the same thing as a C. I. for the mean of this new observation — predicting a particular \( y \) is harder than guessing the mean of \( y \) for this \( x' \).

The prediction interval is centred at the same place as the C. I. for \( \mu_y \) however:
\[
\hat{y} = b_0 + b_1 x'_1 + \cdots + b_p x'_k
\]

The prediction interval is \( (\hat{y} - t^*PE, \ \hat{y} + t^*PE) \), where \( t^* \) is as before, and \( PE \) is analogous to a standard error, but isn’t the same as \( SE(\hat{y}) \).

With one explanatory variable:
\[
SE(\hat{y}) = s\sqrt{\frac{1}{n} + \frac{(x' - \bar{x})^2}{\sum(x_i - \bar{x})^2}}
\]
\[
PE = s\sqrt{1 + \frac{1}{n} + \frac{(x' - \bar{x})^2}{\sum(x_i - \bar{x})^2}}
\]
Example: Prediction from the Data on Life Expectancy in 38 Countries

Here again is the regression line of female life expectancy on the square root of televisions per person. Also shown are $\hat{y}$ plus or minus $SE(\hat{y})$ and plus or minus $PE$, for each $x'$ value:

![Graph showing regression line with confidence intervals.]

Behaviour of Predictions
- For what $x'$ are predictions most accurate?
- What happens to $SE(\hat{y})$ as $n$ gets bigger?
- What happens to $PE$ as $n$ gets bigger?