2.2: Diagnostics and Remedial Measures

Dr. Bean - Stat 5100

1. What happens if the assumptions regarding residuals are not satisfied?

- The p-values associated with our model rely on proper assumptions regarding the test statistics' "sampling distributions".

- If the residual assumptions are not satisfied, then the sampling distributions are wrong and the p-values are worthless.

2. You are running a Box-Cox transformation and the program returns no values...what happened?

Most likely problem: your response variable has negative values. Many of the candidate box-cox transformations, such as the natural log transformation, cannot be computed for negative numbers.

3. Why would we ever want to make a transformation on an X-variable?

- A transformation of the X-variable may help to restore a linear relationship between X and Y.

- More in module 3: Outlier values in X can make the estimated line fit the majority of the data poorly (called influential points).

4. Why would graphical checks (scatterplots, normal probability-plots, etc.) of assumptions be preferred to numerical checks (BF-test of constant variance, correlation test of normality, etc.) of assumptions?

- Violations of assumptions occur on a spectrum, numerical tests try to force a binary decision.

- Numerical tests are good to verify potential violations that you suspect based on graphical checks.

5. Below are two normal probability plots of residuals for a model. For which (if any) of the plots is it acceptable to assume that the residuals follow a normal distribution?
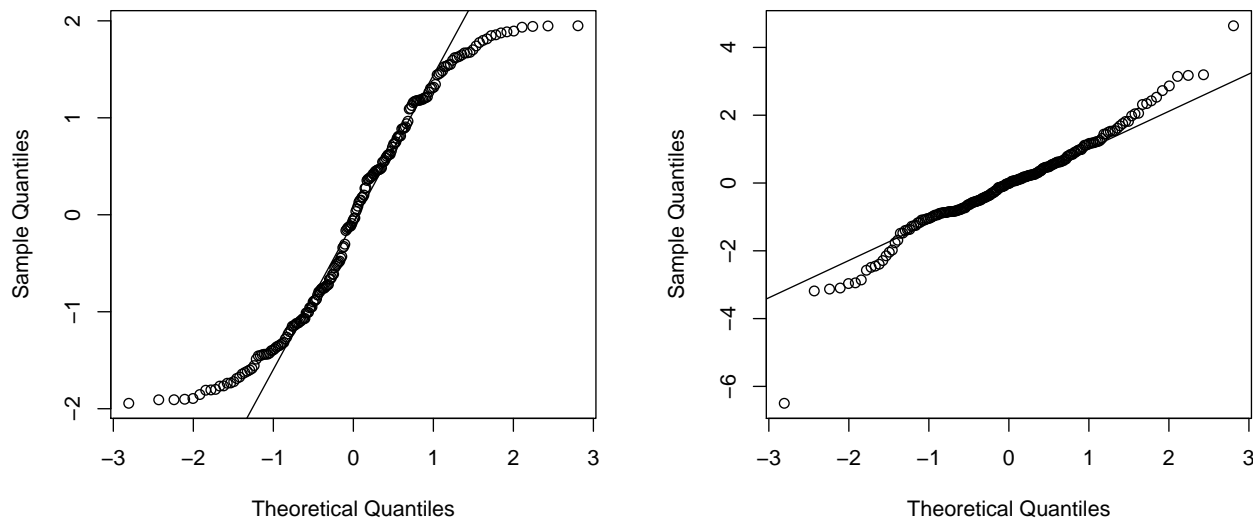
It is acceptable to assume that the left plot follows a normal distribution because the tails of this distribution are shorter than what would be expected in a normal distribution. This means that our p-values will be, at worst, slightly conservative (i.e. bigger than they might otherwise be). In general, it is better to over-estimate, rather than under-estimate p-values.

6. Based on the following output from a Box-Cox Analysis, which transformation would you recommend?
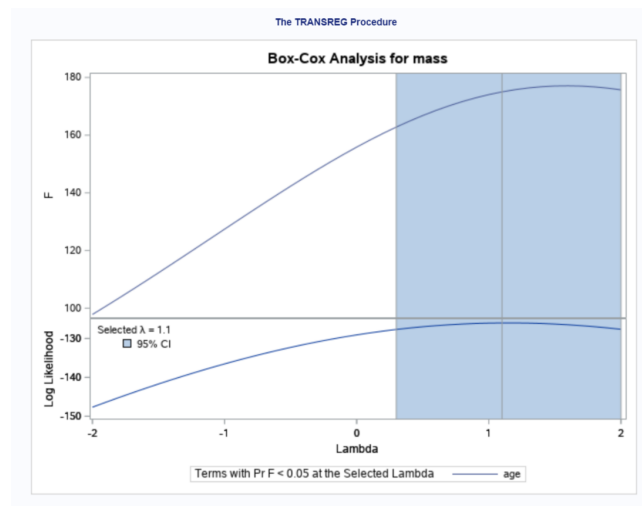
The recommended transformation is close to one. One is well within the confidence bounds of the estimated transformation. Based on this, we would recommend NO transformation.