STAT6038 week 6 lecture 16

Rui Qiu

2017-03-29

Transformations on SLRs

In fitting regression models, we can experiment with changes of scale (transformation) for the variables involved and then assume the regression model holds on the transformed, rather than the untransformed scale.

The transformations we typically use are ones that are typically monotonic, increasing and are considered to be linearizing, and/or variance-stablizing.

A mid-range example of a variance-stablizing transformaion is the log() transformation.

A log-log transformation of both X and Y often works in this situation. (see the "brains" example in tomorrow's lecture)

Note: the default for log() in R is base e, i.e. natural logarithms to the base e, mathematically, as ln(), not common logarithms $\log_{10}$ or binary logarithm $\log_2$.

In R, we have log(..., base=10) or log10(), log(..., base=2) functions respectively.

Transformations continued

A log transformation can also be "linearizing" in that it makes a multiplicative non-linear relationship on the untransformed scale into a linear relationship on the transformed scale.

$$Y = cX^d$$

becomes

$$\ln(Y) = \ln(cX^d)$$
$$= \ln(c) + \ln(X^d)$$
$$= \ln(c) + d \cdot \ln(X)$$

So, let $Y^* = \ln(Y), X^* = \ln(X), \beta_0 = \ln(c), \beta_1 = d$.

$$Y^* = \beta_0 + \beta_1 \cdot X^*$$

This model is **linear.** Linear in the coefficients $\beta_0, \beta_1, \beta_2, \beta_3 \dots, \beta_k$ not necessarily linear in the original variables X, Y.

Finally, log() is just a mid-range example of transformations such as sqrt(), negative inverse $(-\frac{1}{x})$, that may help.