Jared Yu
Module 13 Discussion

Discuss the differences in a regression model between making the random error being multiplicative and making the random error being additive regarding how you approach estimation of the model coefficient(s), how you apply linearization for estimating the model coefficient(s), and how you obtain starting values for estimation of the model coefficient(s).

Ans:

We can look at the example of two different models. Below is model 1, with an additive error term,

$$y = \beta_1 e^{\beta_2 x} + \varepsilon$$

Below is model 2, with a multiplicative error term,

$$y = \theta_1 e^{\theta_2 x} \varepsilon$$

These two models are both nonlinear, since the derivatives w.r.t. their coefficients will depend on the other unknown coefficients.

For model 1, to linearize the model we can first take the expectation function,

$$E(y) = E\left(\beta_1 e^{\beta_2 x} + \varepsilon\right) = \beta_1 e^{\beta_2 x}$$

The above holds since the assumption thus far is that the error terms have mean zero and constant variance σ^2. Then, taking the natural log of this yields,

$$\ln E(y) = \ln \beta_1 + \beta_2 x$$

This could possibly give us some new model 1,

$$y^* = \beta_0^* + \beta_1^* x + \varepsilon$$

We need to be careful however, since in the *original nonlinear model*, the least squares implies minimizing the residual sums of squares on y, but in the *transformed model*, it is being minimized for $\ln y$ (Textbook p.398).

For model 2, applying the natural log transformation we get the following,

$$\begin{aligned} \ln E(y) &= \ln \theta_1 e^{\theta_2 x} \varepsilon \\ &= \ln \theta_1 + \theta_2 x + \ln \varepsilon \\ &= \theta_0^* + \theta_1^* x + \varepsilon^* \end{aligned}$$

Ideally, ε^* follows a normal distribution. It is worth noting here a difference, that in model 1 the natural log of the expectation required adding an error term, but in model 2 there was already the natural log of the (old) error term. The new ε^* is possibly equivalent to $\ln \varepsilon$, or there is some other step in between that the book skips out on.

In trying to find suitable starting values, it’s important to think back to the original models themselves. In this case, the example from the textbook is used. However, for example within the sciences, these could be theoretical models that have a solid foundation. It could make sense to use this domain knowledge to first estimate suitable values. Another possibility is to plot the expectation function over several values to try and get an idea of what the behavior is within different ranges. There is another technique of transforming the expectation function, to first get estimates of linear parameters and to use these as starting values.

Regarding starting values, it seems that for both the additive and multiplicative error terms, they would be rather similar, with differences only in the potential calculations if they are needed. Otherwise, they seem to be fairly similar.