Stat 6021: Interpretation of Regression Coefficients with Log Transformation on Predictor

1 Interpreting Regression Coefficients: Log Transformation on Predictor

One of the reasons a log transformation is a popular transformation is that regression coefficients are still fairly easy to interpret. Consider a log transformation applied to the predictor. The least-squares regression equation becomes

$$\hat{y} = \hat{\beta}_0 + \hat{\beta}_1 \log x \tag{1}$$

When the predictor variable increases by 10%, (1) becomes

$$\hat{y}_{new} = \hat{\beta}_0 + \hat{\beta}_1 \log(1.1x) \tag{2}$$

Consider the difference between $\hat{y}_{new}$ and $\hat{y}$ using (1) and (2), i.e.,

$$\hat{y}_{new} - \hat{y} = \hat{\beta}_1 \log(1.1) \tag{3}$$

From (3), we see for a 10% increase in the predictor, the predicted response variable increases by $\hat{\beta}_1 \log(1.1)$.

1. In general, for an a% increase in the predictor, the predicted response increases by $\hat{\beta}_1 \log(1 + \frac{a}{100})$.
2. If a is small, then $\log(1 + \frac{a}{100}) \approx \frac{a}{100}$ (Taylor series). So an alternative interpretation is: for a 1% increase in the predictor, the predicted response increases by approximately $\frac{\hat{\beta}_1}{100}$.

Let us look at a numerical example. Suppose $\hat{\beta}_1 = 10$:

- For a 1% increase in the predictor, the predicted response increases by $10 \times \log(1 + \frac{1}{100}) = 0.0995$.
- Using the alternative interpretation: For a 1% increase in the predictor, the predicted response increases by approximately $\frac{10}{100} = 0.1$