

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

1 Interpreting Regression Coefficients: Log Transformation on Predictor and Response

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 and the response variable. The least-squares regression equation becomes

$$\log \hat{y} = \hat{\beta}_0 + \hat{\beta}_1 \log x \quad (1)$$

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

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

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

$$\begin{aligned} \log \hat{y}_{new} - \log \hat{y} &= \hat{\beta}_1 \log(1.1) \\ \implies \log\left(\frac{\hat{y}_{new}}{\hat{y}}\right) &= \log(1.1)^{\hat{\beta}_1} \\ \implies \frac{\hat{y}_{new}}{\hat{y}} &= (1.1)^{\hat{\beta}_1} \end{aligned} \quad (3)$$

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

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

Let us look at a numerical example. Suppose $\hat{\beta}_1 = 0.1$:

- For a 1% increase in the predictor, the predicted response is multiplied by a factor of $(1.01)^{0.1} = 1.000996$.
- Using the alternative interpretation: for a 1% increase in the predictor, the predicted response increases by approximately 0.1 percent.