STAT 440Q – FORECASTING
Deriving the Least Squares Regression Equations

The basic simple linear regression model is

$$Y = \beta_0 + \beta_1 \cdot X + \varepsilon$$

For each data point, then, the error term is

$$\varepsilon = Y - \beta_0 - \beta_1 \cdot X$$

The *Sum of Squares for Error (SSE)* is the sum of the squared error terms:

$$SSE = \sum \varepsilon^2 = \sum \left(Y - \beta_0 - \beta_1 \cdot X\right)^2$$

Our least squares estimators for the intercept and slope, $\hat{\beta}_0$ and $\hat{\beta}_1$, are those values which minimize the *SSE*. To find these, we minimize the function – that is, we set the derivative equal to zero. Since there are two quantities being estimated, we need to take two (partial) derivatives and set both equal to zero:

$$\frac{\partial\, SSE}{\partial\, \beta_0} = 0 \qquad \text{and} \qquad \frac{\partial\, SSE}{\partial\, \beta_1} = 0$$

FOR THE INTERCEPT:

$$\frac{\partial SSE}{\partial \beta_0} = \sum 2 \cdot \left(Y - \beta_0 - \beta_1 \cdot X\right)^1 \cdot (-1)$$

So $$\sum 2 \cdot \left(Y - \hat{\beta}_0 - \hat{\beta}_1 \cdot X\right)^1 \cdot (-1) = 0$$

→ $$-2\sum\left(Y - \hat{\beta}_0 - \hat{\beta}_1 \cdot X\right) = 0$$

→ $$\sum\left(Y - \hat{\beta}_0 - \hat{\beta}_1 \cdot X\right) = 0$$

→ $$\sum Y - n \cdot \hat{\beta}_0 - \hat{\beta}_1 \sum X = 0$$

→ $$\sum Y - \hat{\beta}_1 \sum X = n \cdot \hat{\beta}_0$$

→ $$\frac{\sum Y}{n} - \hat{\beta}_1 \cdot \frac{\sum X}{n} = \hat{\beta}_0$$

and hence

$$\bar{Y} - \hat{\beta}_1 \cdot \bar{X} = \hat{\beta}_0$$

Note that this implies

$$\bar{Y} = \hat{\beta}_0 + \hat{\beta}_1 \cdot \bar{X}$$

In other words, once you know the slope, you find the intercept by plugging the values $\bar{X}$ and $\bar{Y}$ into the equation of the regression line.

FOR THE SLOPE:

$$\frac{\partial SSE}{\partial \beta_1} = \sum 2 \cdot \left(Y - \beta_0 - \beta_1 \cdot X\right)^1 \cdot (-X)$$

So $$\sum 2 \cdot \left(Y - \hat{\beta}_0 - \hat{\beta}_1 \cdot X\right)^1 \cdot (-X) = 0$$

→ $$-2\sum\left(X \cdot Y - \hat{\beta}_0 \cdot X - \hat{\beta}_1 \cdot X^2\right) = 0$$

→ $$\sum\left(X \cdot Y - \hat{\beta}_0 \cdot X - \hat{\beta}_1 \cdot X^2\right) = 0$$

→ $$\sum X \cdot Y - \hat{\beta}_0 \sum X - \hat{\beta}_1 \sum X^2 = 0$$

We substitute in for $\hat{\beta}_0$:

→ $$\sum X \cdot Y - \left[\frac{\sum Y}{n} - \hat{\beta}_1 \cdot \frac{\sum X}{n}\right] \cdot \left(\sum X\right) - \hat{\beta}_1 \sum X^2 = 0$$

→ $$\sum X \cdot Y - \frac{1}{n}\left(\sum X\right) \cdot \left(\sum Y\right) + \hat{\beta}_1 \cdot \frac{1}{n} \cdot \left(\sum X\right)^2 - \hat{\beta}_1 \sum X^2 = 0$$

→ $$\begin{aligned} \sum X \cdot Y - \frac{1}{n}\left(\sum X\right) \cdot \left(\sum Y\right) &= \hat{\beta}_1 \sum X^2 - \hat{\beta}_1 \cdot \frac{1}{n} \cdot \left(\sum X\right)^2 \\ &= \hat{\beta}_1 \cdot \left[\sum X^2 - \frac{1}{n} \cdot \left(\sum X\right)^2\right] \end{aligned}$$

→ $$\frac{\sum X \cdot Y - \frac{1}{n} \cdot \left(\sum X\right) \cdot \left(\sum Y\right)}{\sum X^2 - \frac{1}{n} \cdot \left(\sum X\right)^2} = \hat{\beta}_1$$

This is commonly written as

$$\frac{n \cdot \sum X \cdot Y - \left(\sum X\right) \cdot \left(\sum Y\right)}{n \cdot \sum X^2 - \left(\sum X\right)^2} = \hat{\beta}_1$$