Simple Linear Regression Beta Coefficient Derivations

BIOS 6611

CU Anschutz

Week 8

1. **Review of Least Squares**

2. **Deriving $\hat{\beta}_0$ and $\hat{\beta}_1$**

Review of Least Squares

Method of Least Squares/Least Squares Regression

$S = \sum_{i=1}^{n} e_i^2$ is called the **method of least squares** or **least squares regression** because it minimizes the sum of squares due to error (SS_{Error}):

$$SS_{Error} = SSE = \sum_{i=1}^{n} e_i^2 = \sum_{i=1}^{n} (Y_i - \hat{Y}_i)^2 = \sum_{i=1}^{n} \left(Y_i - (\hat{\beta}_0 + \hat{\beta}_1 X_i) \right)^2$$

The SSE is also known as the **sums of squares error** or **residual sum of squares**.

Mathematical Approach to Least Squares

Mathematically stated, this approach identifies estimates for β_0 and β_1, $\hat{\beta}_0$ and $\hat{\beta}_1$, such that for any other possible estimators, $\hat{\beta}_0^*$ and $\hat{\beta}_1^*$, it must be true that:

$$SSE = \sum_{i=1}^{n} e_i^2 = \sum_{i=1}^{n} \left(Y_i - (\hat{\beta}_0 + \hat{\beta}_1 X_i) \right)^2 < \sum_{i=1}^{n} \left(Y_i - (\hat{\beta}_0^* + \hat{\beta}_1^* X_i) \right)^2$$

How can we arrive at these optimal estimates for $\hat{\beta}_0$ and $\hat{\beta}_1$?

One approach is to treat our SS_{Error} as a *loss function* and minimize it over all choices for β_0 and β_1. To obtain the minimum (or maximum) of a function we find values such that the first (partial) derivatives are equal to 0. We will derive these in a separate slide set.

Deriving $\hat{\beta}_0$ and $\hat{\beta}_1$

Formulas for $\hat{\beta}_0$ and $\hat{\beta}_1$

All simple linear regression parameters can be estimates from 5 summary statistics:

1. n
2. $\sum X_i$ (can determine $\bar{X}$ by dividing by n)
3. $\sum Y_i$ (can determine $\bar{Y}$ by dividing by n)
4. $\sum(X_i^2)$
5. $\sum(X_i Y_i)$

With these 5 statistics we have

$$\hat{\beta}_1 = \frac{\sum_{i=1}^{n}(X_i - \bar{X})(Y_i - \bar{Y})}{\sum_{i=1}^{n}(X_i - \bar{X})^2} = \frac{S_{XY}}{S_{XX}}$$

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

Deriving $\hat{\beta}_0$

First we will solve for $\hat{\beta}_0$:

$$\frac{\partial}{\partial \beta_0} SS_{Error} = \frac{\partial}{\partial \beta_0} \left(\sum_{i=1}^{n} (Y_i - \beta_0 - \beta_1 X_i)^2 \right)$$

Let's cut to our "whiteboard" to work through the math...

Deriving $\hat{\beta}_1$

Let's now solve for $\hat{\beta}_1$ (which follows similar steps):

$$\frac{\partial}{\partial \beta_0} SS_{Error} = \frac{\partial}{\partial \beta_1} \left(\sum_{i=1}^{n} (Y_i - \beta_0 - \beta_1 X_i)^2 \right)$$

Let's cut to our "whiteboard" to work through the math...