Stat 378 – Applied Regression Analysis

Formula Sheet

- The Total Sum of Squares: Let J_n be the $n \times n$ matrix with all entries being 1.

$$SS_T = \vec{Y}'\vec{Y} - \frac{1}{n}\vec{Y}'J_n\vec{Y} = \sum_{i=1}^{n}(Y_i - \bar{Y})^2$$

- The Residual Sum of Squares:

$$SS_{Res} = \vec{Y}'(I_n - X(X'X)^{-1}X')\vec{Y} = \vec{Y}'\vec{Y} - \hat{\beta}'X'\vec{Y} = \sum_{i=1}^{n}(Y_i - \hat{Y}_i)^2$$

- The Sum of Squares for Regression:

$$SS_R = SS_T - SS_{Res} = \hat{\beta}'X'\vec{Y} - \frac{1}{n}\vec{Y}'J_n\vec{Y} = \sum_{i=1}^{n}(\hat{Y}_i - \bar{Y})^2$$

Standardized Scores

- For an individual slope:

$$t = \frac{\hat{\beta}_j - \beta_j}{\hat{\sigma}\sqrt{C_{jj}}} \sim t(n-p)$$

- For a subset of slopes:

$$F = \frac{(\hat{\beta}_1 - \vec{\beta}_1)'(X_1'X_1)(\hat{\beta}_1 - \vec{\beta}_1)}{r\ MS_{Res}} \sim F(r, n-p)$$

$$F = \frac{SS_{Res}(reduced) - SS_{Res}(full)}{r\ MS_{Res}} \sim F(r, n-p)$$

- For all slopes:

$$F = \frac{MS_R}{MS_{Res}} \sim F(k, n-p)$$

- For $E(Y|\vec{x}_0)$:

$$t = \frac{\hat{Y}_0 - E(Y|\vec{x}_0)}{\hat{\sigma}\sqrt{\vec{x}_0'(X'X)^{-1}\vec{x}_0}} \sim t(n-p)$$

- For $Y_0 = \vec{x}_0'\vec{\beta}$:

$$t = \frac{\hat{Y}_0 - Y_0)}{\hat{\sigma}\sqrt{1 + \vec{x}_0'(X'X)^{-1}\vec{x}_0}} \sim t(n-p)$$