Stat 6021: Hypothesis Testing in Multiple Linear Regression

In a multiple linear regression, the various tests for the regression coefficients can all be generalized by the partial F test. For the rest of this document, consider a multiple linear regression with k predictors.

1 Partial F Test

In a partial F test, we are testing whether we can drop the first r predictors from the model, where $r \leq k$. We have the following:

$$\begin{aligned} H_0 &: \beta_1 = \beta_2 = \cdots = \beta_r = 0 \\ H_a &: \text{not all of } \beta_j \text{ in } H_0 \text{ equal zero.} \end{aligned}$$

The test statistic is

$$F_0 = \frac{MS_R(\beta_1, \cdots, \beta_r | \beta_{r+1}, \cdots, \beta_k)}{MS_{Res}(\beta_1, \cdots, \beta_k)}. \tag{1}$$

Under the null hypothesis, $F_0 \sim F_{r,n-k-1}$.

2 ANOVA F Test

The ANOVA F test is just a special case of the partial F test, where we are testing whether all k predictors can be dropped, i.e., $r = k$. The hypotheses become

$$\begin{aligned} H_0 &: \beta_1 = \beta_2 = \cdots = \beta_k = 0 \\ H_a &: \text{not all of } \beta_j \text{ in } H_0 \text{ equal zero.} \end{aligned}$$

The test statistic is

$$F_0 = \frac{MS_R(\beta_1, \cdots, \beta_k)}{MS_{Res}(\beta_1, \cdots, \beta_k)}. \tag{2}$$

Notice that (2) is the same as (1) when $r = k$. Under the null hypothesis, $F_0 \sim F_{k,n-k-1}$.

3 t Test

The t test is just a special case of the partial F test, where we are testing whether a single predictor can be dropped, i.e., $r = 1$. Assuming we are looking to drop the first predictor, the hypotheses become

$$\begin{aligned} H_0 &: \beta_1 = 0 \\ H_a &: \beta_1 \neq 0 \end{aligned}$$

The test statistic is

$$F_0 = \frac{MS_R(\beta_1|\beta_2, \cdots, \beta_k)}{MS_{Res}(\beta_1, \cdots, \beta_k)}. \tag{3}$$

Notice that (3) is the same as (1) when $r = 1$. Under the null hypothesis, $F_0 \sim F_{1,n-k-1}$. It turns out that any random variable $X \sim t(n)$, $X^2 \sim F(1, n)$. Thus the test statistic from (3) gives a t statistic that is equal to $\sqrt{F_0}$ and the t statistic is compared to the t distribution with $n - k - 1$ degrees of freedom.