F-test for linear regression

Consider an linear regression model (unrestricted model) is

$$y = w_0 + w_1 x_1 + \cdots + w_n x_n + e \tag{1}$$

where w_0 and $w_1, \ldots, w_n$ are parameters need to be estimated and $x_1, \ldots, x_n$ are input variables.

Denote $e \sim \mathrm{N}\left(0, \sigma^2\right)$, then the sum of squares of residuals (SSR)

$$SSR = \sum_{i=1}^{N} e(i)^2 = \sum_{i=1}^{N} \left[y(i) - \hat{y}(i) \right]^2 \tag{2}$$

satisfies a chi-square distribution:

$$SSR \sim \sigma^2 \chi^2 (N - n - 1) \tag{3}$$

Consider the hypothesis

$$\begin{aligned} &H_0 : w_{n_\mathrm{r}+1} = \cdots = w_n = 0 \\ &H_\mathrm{A} : w_{n_\mathrm{r}+1}, \ldots, w_n \neq 0 \end{aligned}$$

The restricted linear regression model is

$$y = w_0' + w_1' x_1 + \cdots + w_{n_\mathrm{r}}' x_{n_\mathrm{r}} + e \tag{4}$$

where w_0' and $w_1', \ldots, w_{n_\mathrm{r}}'$ are model parameters and $x_1, \ldots, x_{n_\mathrm{r}}$ are input variables.

The variance of e for the restricted model is the same as that of the unrestricted model assuming that H_0 is true.

Then

$$\begin{aligned} SSR_\mathrm{r} - SSR_\mathrm{ur} &\sim \sigma^2 \chi^2 (n - n_\mathrm{r}) \\ SSR_\mathrm{ur} &\sim \sigma^2 \chi^2 (N - n - 1) \end{aligned} \tag{5}$$

thus

$$F = \frac{(SSR_\mathrm{r} - SSR_\mathrm{ur}) / (n - n_\mathrm{r})}{SSR_\mathrm{ur} / (N - n - 1)} \sim \mathrm{f}\,(n - n_\mathrm{r}, N - n - 1) \tag{6}$$