Lecture Summary: Feb. 24, 2023

- F-tests

Examples: Y, x_1, x_2, x_3

1. $\mathrm{H}_0 : \beta_3 = 0$ vs $\mathrm{H}_1 : \beta_3 \neq 0$

Full model: $Y = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \beta_3 x_3 + \epsilon$.

SSE(F)=SSE(x_1, x_2, x_3).

Reduced model (under H_0): $Y = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \epsilon$.

SSE(R)=SSE(x_1, x_2).

$$\begin{aligned} F &= \frac{\{\mathrm{SSE(R)} - \mathrm{SSE(F)}\}/\{\mathrm{df_R} - \mathrm{df_F}\}}{\mathrm{SSE(F)}/\mathrm{df_F}} \\ &= \frac{\{\mathrm{SSE}(x_1, x_2) - \mathrm{SSE}(x_1, x_2, x_3)\}/1}{\mathrm{SSE}(x_1, x_2, x_3)/(n-4)} \\ &= \frac{\mathrm{SSR}(x_3|x_1, x_2)/1}{\mathrm{MSE}(x_1, x_2, x_3)} \\ &= \frac{\mathrm{MSR}(x_3|x_1, x_2)}{\mathrm{MSE}(x_1, x_2, x_3)}, \end{aligned} \tag{1}$$

where $\mathrm{df_R} = n - 3$ and $\mathrm{df_F} = n - 4$. Under H_0, $F \sim F_{1,n-4}$.

2. $\mathrm{H}_0 : \beta_2 = \beta_3 = 0$ vs H_1 : Not H_0

Full model: Same as above.

Reduced model (under H_0): $Y = \beta_0 + \beta_1 x_1 + \epsilon$.

SSE(R)=SSE(x_1).

The F-statistic has the same expression as the first line in (1) but the definition of SSE(R) is different, and now $\mathrm{df_R} = n - 2$ [SSE(F) and $\mathrm{df_F}$ remain unchanged]. Under the new H_0, $F \sim F_{2,n-4}$.

- Coefficient of partial determination

$R^2_{Y,x_2|x_1}$, or $R^2_{Y,2|1}$, is defined as

$$\frac{\mathrm{SSR}(x_2|x_1)}{\mathrm{SSE}(x_1)} = 1 - \frac{\mathrm{SSE}(x_1, x_2)}{\mathrm{SSE}(x_1)}.$$

It measures the proportionate reduction in the variation of Y due to adding x_2, given that x_1 is already in the model.