# Lecture Summary: Feb. 22, 2023

• ANOVA: Define SSTO $= \sum_{i=1}^n (Y_i - \bar{Y})^2$, SSR $= \sum_{i=1}^n (\hat{Y}_i - \bar{Y})^2$, SSE $= \sum_{i=1}^n (Y_i - \hat{Y}_i)^2$, where $\hat{Y}_i = \hat{\beta}_0 + \hat{\beta}_1 x_{i,1} + \cdots + \hat{\beta}_{p-1} x_{i,p-1}$. Then, we have

$$\text{SSTO} = \text{SSR} + \text{SSE}.$$

ANOVA table:

| Source | SS | d.f. | MS | F |
|---|---|---|---|---|
| Regression | SSR | $p-1$ | MSR$=$ SSR$/(p-1)$ | $F =$ MSR/MSE |
| Error | SSE | $n-p$ | MSE$=$ SSE$/(n-p)$ | |
| Total | SSTO | $n-1$ | | |

Under $\text{H}_0 : \beta_1 = \cdots = \beta_{p-1} = 0$, the $F$ in the ANOVA table has an $F_{p-1,n-p}$ distribution.

The definition of $R^2$, and its interpretation, are the same as in simple linear regression.

• Extra SS

Example: $Y$, $x_1$, $x_2$, $x_3$

Q: Is $x_2$ important given that $x_1$ is already in the model?

$$\begin{aligned} \text{SSR}(x_2|x_1) &= \text{SSR}(x_1, x_2) - \text{SSR}(x_1) \\ &= \text{SSE}(x_1) - \text{SSE}(x_1, x_2). \end{aligned}$$

Similarly, we have

$$\begin{aligned} \text{SSR}(x_3|x_1, x_2) &= \text{SSR}(x_1, x_2, x_3) - \text{SSR}(x_1, x_2) \\ &= \text{SSE}(x_1, x_2) - \text{SSE}(x_1, x_2, x_3); \\ \text{SSR}(x_2, x_3|x_1) &= \text{SSR}(x_1, x_2, x_3) - \text{SSR}(x_1) \\ &= \text{SSE}(x_1) - \text{SSE}(x_1, x_2, x_3). \end{aligned}$$

• D.f. for extra SS

SSR($x_2|x_1$) has 1 d.f.

Similarly, SSR($x_1|x_2$), SSR($x_3|x_1, x_2$) both have 1 d.f., but SSR($x_2, x_3|x_1$) has 2 d.f.

- ANOVA with extra SS

$$\mathrm{SSR}(x_1, x_2, x_3) = \mathrm{SSR}(x_1) + \mathrm{SSR}(x_2|x_1) + \mathrm{SSR}(x_3|x_1, x_2).$$

| Source | SS | d.f. | MS |
|---|---|---|---|
| Regression | SSR($x_1, x_2, x_3$) | 3 | MSR($x_1, x_2, x_3$) =SSR($x_1, x_2, x_3$)/3 |
| $x_1$ | SSR($x_1$) | 1 | MSR($x_1$) =SSR($x_1$)/1 |
| $x_2\|x_1$ | SSR($x_2\|x_1$) | 1 | MSR($x_2\|x_1$) =SSR($x_2\|x_1$)/1 |
| $x_3\|x_1, x_2$ | SSR($x_3\|x_1, x_2$) | 1 | MSR($x_3\|x_1, x_2$) =SSR($x_3\|x_1, x_2$)/1 |
| Error | SSE($x_1, x_2, x_3$) | $n-4$ | MSE($x_1, x_2, x_3$) =SSE($x_1, x_2, x_3$)/($n-4$) |
| Total | SSTO | $n-1$ | |