

Anova Table for Multiple Regression

For $i = 1, 2, \dots, n$ and the model

$$Y_i = \beta_0 + \beta_1 X1_i + \beta_2 X2_i + \beta_3 X3_i + Z_i$$

we have the sum of squares identity

$$\begin{aligned} \sum_{i=1}^n y_i^2 &= \text{SSReg}(\beta_0, \beta_1, \beta_2, \beta_3) + \sum_{i=1}^n (y_i - \hat{y}_i)^2 \\ \text{TSS} &= \text{SSReg}(\beta_0) + \text{SSReg}(\beta_1, \beta_2, \beta_3 | \beta_0) + \text{RSS} \\ \text{TCSS} &= \text{SSReg}(\beta_1 | \beta_0) + \text{SSReg}(\beta_2 | \beta_0, \beta_1) + \text{SSReg}(\beta_3 | \beta_0, \beta_1, \beta_2) + \text{RSS} \end{aligned}$$

where

$$\begin{aligned} \text{SSReg}(\# | \$) &= \text{SSReg}(\$, \#) - \text{SSReg}(\$) \\ \text{TCSS} &= \sum_{i=1}^n (y_i - \bar{y})^2 = \sum_{i=1}^n y_i^2 - n\bar{y}^2 = \text{TSS} - \text{SSReg}(\beta_0) \end{aligned}$$

On computer output this is often given in the form of a table.

	Df	Sum Sq
X1	1	$\text{SSReg}(\beta_1 \beta_0) = \text{SSReg}(\beta_0, \beta_1) - \text{SSReg}(\beta_0)$
X2	1	$\text{SSReg}(\beta_2 \beta_0, \beta_1) = \text{SSReg}(\beta_0, \beta_1, \beta_2) - \text{SSReg}(\beta_0, \beta_1)$
X3	1	$\text{SSReg}(\beta_3 \beta_0, \beta_1, \beta_2) = \text{SSReg}(\beta_0, \beta_1, \beta_2, \beta_3) - \text{SSReg}(\beta_0, \beta_1, \beta_2)$
Residuals	$n - 4$	RSS

For testing $H: \beta_2 = \beta_3 = 0$ against K : at least one of the two $\neq 0$ we need to find

$$\begin{aligned} \text{SSReg}(\beta_2, \beta_3 | \beta_0, \beta_1) &= \text{SSReg}(\beta_0, \beta_1, \beta_2, \beta_3) - \text{SSReg}(\beta_0, \beta_1) \\ &= \text{SSReg}(\beta_3 | \beta_0, \beta_1, \beta_2) + \text{SSReg}(\beta_2 | \beta_0, \beta_1) \\ &= \text{SSE}_R - \text{SSE}_C \end{aligned}$$

where in the notation of the text SSE_R is RSS from the reduced model where H is assumed to be true and SSE_C is RSS from the complete or full model.