ANOVA II. Sequential SS and Partial SS

Miaoyan Wang

Department of Statistics
UW Madison

Extra Sums of Squares

- Basic ideas: An **extra sum of squares** measures the marginal reduction (or increase) in the SSE (or SSR) when one or several explanatory variables are added to the regression model, given other explanatory variables are already in the model.
- Extra sums of squares are useful for constructing tests about subsets of regression coefficients.
- Recall the general linear test approach.

General Linear Test Approach

- Consider the **full model** (or, **unrestricted model**)

$$Y = \beta_0 + \beta_1 X + \varepsilon, \quad \varepsilon \sim \text{iid } N(0, \sigma^2)$$

  and obtain SSE(F).
- Consider the **reduced model** (or, **restricted model**) under the $H_0 : \beta_1 = 0$

$$Y = \beta_0 + \varepsilon, \quad \varepsilon \sim \text{iid } N(0, \sigma^2)$$

  and obtain SSE(R).
  Note that we always have SSE(F) $\leq$ SSE(R).

Example 1

- Response variable Y and 2 explanatory variables X_1, X_2.
- The full model is

$$Y = \beta_0 + \beta_1 X_1 + \beta_2 X_2 + \varepsilon$$

  and denote its sum of squared error by $\text{SSE}(X_1, X_2)$.
- To test $H_0 : \beta_2 = 0$, what is the reduced model?

$$Y = \beta_0 + \beta_1 X_1 + \varepsilon,$$

  and denote its sum of squared error by $\text{SSE}(X_1)$
- Compare $\text{SSE}(X_1)$ and $\text{SSE}(X_1, X_2)$.

$$\text{SSE}(X_1) \geq \text{SSE}(X_1, X_2)$$

  Adding another explanatory variable will *never* increase SSE.

Example 1

- Define

$$\begin{aligned}\mathrm{SSR}(X_2|X_1) &= \mathrm{SSE}(X_1) - \mathrm{SSE}(X_1, X_2)\\ &= \mathrm{SSR}(X_1, X_2) - \mathrm{SSR}(X_1)\end{aligned}$$

- Interpretation:
  $\mathrm{SSR}(X_2|X_1)$ measures the decrease in the SSE when X_2 is added to the regression model, given X_1 is already in the model.

Partial F Test: Example 1

- The test statistic for $H_0 : \beta_2 = 0$ is

$$
\begin{aligned}
F^* &= \frac{\frac{\mathrm{SSE}(X_1) - \mathrm{SSE}(X_1, X_2)}{(n-2)-(n-3)}}{\frac{\mathrm{SSE}(X_1, X_2)}{n-3}} \\
&= \frac{\frac{\mathrm{SSR}(X_2 | X_1)}{1}}{\frac{\mathrm{SSE}(X_1, X_2)}{(n-3)}}
\end{aligned}
$$

- Under the H_0,

$$F^* \sim F_{1,n-3}.$$

- The decision rule is to reject H_0 if $f^* > f_{1,n-3,\alpha}$.
- Relation to a T-test for β_2 in the full model? as $(t^*)^2 = f^*$.

Example 2

- Response variable Y and 3 explanatory variables X_1, X_2, X_3.
- The full model is

$$Y = \beta_0 + \beta_1 X_1 + \beta_2 X_2 + \beta_3 X_3 + \varepsilon$$

  and denote the sum of squared error by $\mathrm{SSE}(X_1, X_2, X_3)$.
- To test $H_0 : \beta_1 = \beta_3 = 0$, what is the reduced model?

$$Y = \beta_0 + \beta_2 X_2 + \varepsilon.$$

  Denote the sum of squared error by $\mathrm{SSE}(X_2)$.
- Compare $\mathrm{SSE}(X_1, X_2, X_3)$ with $\mathrm{SSE}(X_2)$.

Example 2

- The extra sum of squares $SSR(X_1, X_3|X_2)$ is defined as

$$SSR(X_1, X_3|X_2) = SSE(X_2) - SSE(X_1, X_2, X_3)$$

- Interpretation:
  $SSR(X_1, X_3|X_2)$ measures the decrease in the SSE when X_1 and X_3 are added to the regression model, given X_2 is already in the model.
- Equivalently,

$$SSR(X_1, X_3|X_2) = SSR(X_1, X_2, X_3) - SSR(X_2)$$

- Interpretation:
  Equivalently, $SSR(X_1, X_3|X_2)$ measures the increase in the SSR when X_1 and X_3 are added to the regression model, given X_2 is already in the model.

Partial F Test: Example 2

- The test statistic for $H_0 : \beta_1 = \beta_3 = 0$ is

$$
\begin{aligned}
F^* &= \frac{\frac{\text{SSE}(X_2) - \text{SSE}(X_1, X_2, X_3)}{(n-2)-(n-4)}}{\frac{\text{SSE}(X_1, X_2, X_3)}{n-4}} \\
&= \frac{\frac{\text{SSR}(X_1, X_3 | X_2)}{2}}{\frac{\text{SSE}(X_1, X_2, X_3)}{(n-4)}}
\end{aligned}
$$

- Under the H_0, $F^* \sim F_{2,n-4}$.
- The decision rule is to reject H_0 if $f^* > f_{2,n-4,\alpha}$.

Decomposition of SSR into Extra Sums of Squares

- Begin with

$$\text{SSTO} = \text{SSR}(X_1) + \text{SSE}(X_1).$$

- Since $\text{SSE}(X_1) = \text{SSR}(X_2|X_1) + \text{SSE}(X_1, X_2)$, we have

$$\text{SSTO} = \underbrace{\text{SSR}(X_1) + \text{SSR}(X_2|X_1)}_{\text{explained by regression SSR}(X_1, X_2)} + \underbrace{\text{SSE}(X_1, X_2)}_{\text{explained by error}}.$$

Sequential SS in ANOVA Table

For $X_1, \ldots, X_{p-1}$ in general, we may summarize the decomposition of SSR into extra sums of squares in an ANOVA table:

Source	SS	df
Regression	$\text{SSR}(X_1, X_2, \ldots, X_{p-1})$	$p-1$
X_1	$\text{SSR}(X_1)$	1
X_2	$\text{SSR}(X_2 \mid X_1)$	1
$\cdots$	$\cdots$	
X_{p-1}	$\text{SSR}(X_{p-1} \mid X_1, \ldots, X_{p-2})$	1
Error	$\text{SSE}(X_1, X_2, \ldots, X_{p-1})$	$n-p$
Total	SSTO	$n-1$

Order of Fitting

- The order of the explanatory variables is arbitrary. For example,

$$\begin{aligned} \text{SSTO} &= \text{SSR}(X_1) + \text{SSR}(X_2|X_1) + \text{SSE}(X_1, X_2) \\ \text{SSTO} &= \text{SSR}(X_2) + \text{SSR}(X_1|X_2) + \text{SSE}(X_1, X_2). \end{aligned}$$

- Generally, decomposition depends on order of explanatory variables.
- The number of possible orderings becomes large as the number of explanatory variables increases.
- The extra sums of squares in the ANOVA table above are called **sequential SS**.
- When is sequential SS useful?
  when there is a pre-determined order for selecting explanatory variables (e.g. main effect, interaction effect).

Partial SS in ANOVA Table

For $X_1, \ldots, X_{p-1}$ in general, we may summarize the decomposition of SSR into **partial sums of squares** in an ANOVA table:

Source	SS	df
Regression	$SSR(X_1, X_2, \ldots, X_{p-1})$	$p-1$
X_1	$SSR(X_1 \vert X_2, X_3, \ldots, X_{p-1})$	1
X_2	$SSR(X_2 \vert X_1, X_3, \ldots, X_{p-1})$	1
$\cdots$	$\cdots$	
X_{p-1}	$SSR(X_{p-1} \vert X_1, X_2, \ldots, X_{p-2})$	1
Error	$SSE(X_1, X_2, \ldots, X_{p-1})$	$n-p$
Total	SSTO	$n-1$

- The results are independent of the order of the explanatory variables.
- The partial sums of squares do not add up to anything meaningful.

Coefficient of Partial Determination

- Coefficient of partial determination: measures the marginal contribution of one explanatory variable when all others are already included in the regression model.
- For example, with 3 explanatory variables, the coefficients of partial determination are

$$
\begin{aligned}
R^2_{Y1|23} &= \frac{\mathrm{SSR}(X_1|X_2, X_3)}{\mathrm{SSE}(X_2, X_3)} \\
R^2_{Y2|13} &= \frac{\mathrm{SSR}(X_2|X_1, X_3)}{\mathrm{SSE}(X_1, X_3)} \\
R^2_{Y3|12} &= \frac{\mathrm{SSR}(X_3|X_1, X_2)}{\mathrm{SSE}(X_1, X_2)}
\end{aligned}
$$

- Coefficient of partial correlation: square root of a coefficient of partial determination with the same sign as the corresponding fitted regression coefficient.