1 Extra Sums of Squares

1.1 Definitions and Notation

- *Extra sums of squares* are used to investigate the degree to which **adding another predictor** to a model **reduces** the amount of **unexplained variation** in the responses $Y_1, Y_2, \ldots, Y_n$.
- Suppose first we have only two predictors X_1 and X_2. We have the following definitions and notation:

  ▷ Let
  $$\mathbf{SSE}(\boldsymbol{X_1}) \qquad \text{and} \qquad \mathbf{SSR}(\boldsymbol{X_1})$$
  be the error and regression sums of squares obtained from fitting the **simple linear regression model** with **just $\boldsymbol{X_1}$**.

  ▷ Let
  $$\mathbf{SSE}(\boldsymbol{X_1}, \boldsymbol{X_2}) \qquad \text{and} \qquad \mathbf{SSR}(\boldsymbol{X_1}, \boldsymbol{X_2})$$
  be the error and regression sums of squares obtained from fitting the **multiple regression** model with **both $\boldsymbol{X_1}$ and $\boldsymbol{X_2}$**.

  ▷ Define the ***extra sum of squares*** $\mathbf{SSR}(\boldsymbol{X_2}|\boldsymbol{X_1})$ to be the **reduction** in the **error sum of squares** resulting from **adding $\boldsymbol{X_2}$ to the model that already includes $\boldsymbol{X_1}$**:
  $$\text{SSR}(X_2|X_1) \;=\; \text{SSE}(X_1) - \text{SSE}(X_1, X_2) \tag{1}$$

  $\text{SSR}(X_2|X_1)$ tells us how much of the Y variation **not already explained by $\boldsymbol{X_1}$** can be **explained by $\boldsymbol{X_2}$**.

  Intuitively, $\text{SSR}(X_2|X_1)$ provides an indication of whether it's useful to add X_2 to a model that **already** includes X_1.

- This idea of an *extra sum of squares* can be extended to the case in which **more than two predictors are already in the model**. For example, the extra sum of squares
  $$\text{SSR}(X_3|X_1, X_2) \;=\; \text{SSE}(X_1, X_2) - \text{SSE}(X_1, X_2, X_3)$$
  tells us how much of the Y variation **not already explained by $\boldsymbol{X_1}$ and $\boldsymbol{X_2}$** can be **explained by $\boldsymbol{X_3}$**. It provides an indication of whether it's useful to add X_3 to a model that **already** includes X_1 and X_2.

- Likewise, an *extra sum of squares* can be used to decide if it's useful to **add more than two predictors** to a model. For example,

$$\text{SSR}(X_3, X_4|X_1, X_2) \;=\; \text{SSE}(X_1, X_2) \,-\, \text{SSE}(X_1, X_2, X_3, X_4)$$

  tells us how much of the Y variation **not already explained by $\boldsymbol{X_1}$ and $\boldsymbol{X_2}$** can be **explained by adding $\boldsymbol{X_3}$ and $\boldsymbol{X_4}$ to the model**.

- Note that another form for $\text{SSR}(X_2|X_1)$ is

$$\text{SSR}(X_2|X_1) \;=\; \text{SSR}(X_1, X_2) \,-\, \text{SSR}(X_1)\,. \tag{2}$$

  This is easy to verify using the fact that for any fitted regression model,

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

  We can generalize (2). For example,

$$\text{SSR}(X_3|X_1, X_2) \;=\; \text{SSR}(X_1, X_2, X_3) \,-\, \text{SSR}(X_1, X_2) \tag{3}$$

  and

$$\text{SSR}(X_3, X_4|X_1, X_2) \;=\; \text{SSR}(X_1, X_2, X_3, X_4) \,-\, \text{SSR}(X_1, X_2)\,. \tag{4}$$

1.2 Decomposition of Regression Sum of Squares into Extra Sums of Squares

- From (2) we have

$$\text{SSR}(X_1, X_2) \;=\; \text{SSR}(X_1) + \text{SSR}(X_2|X_1)\,. \tag{5}$$

  Recall that $\text{SSR}(X_1, X_2)$ measures the variation in Y that's **explained by $\boldsymbol{X_1}$ and $\boldsymbol{X_2}$**.

  In (5), this component of Y variation is further broken into **two parts**, variation in Y that's **explained by $\boldsymbol{X_1}$ by itself** ($\text{SSR}(X_1)$), and **excess variation** in Y (not already explained by X_1) that's **explained by $\boldsymbol{X_2}$** ($\text{SSR}(X_2|X_1)$).

- Note that by interchanging the roles of X_1 and X_2 in (2) and (5), we have

$$\text{SSR}(X_1, X_2) \;=\; \text{SSR}(X_2) + \text{SSR}(X_1|X_2)$$

- We can generalize (5) to **more than two predictors**. For example, from (3) (and then (5)),

$$\begin{aligned}\text{SSR}(X_1, X_2, X_3) &= \text{SSR}(X_1, X_2) + \text{SSR}(X_3|X_1, X_2) \\ &= \text{SSR}(X_1) + \text{SSR}(X_2|X_1) + \text{SSR}(X_3|X_1, X_2)\end{aligned}$$

  As another example, from (4),

$$\text{SSR}(X_1, X_2, X_3, X_4) = \text{SSR}(X_1, X_2) + \text{SSR}(X_3, X_4|X_1, X_2).$$

- A extra sum of squares such as

$$\text{SSR}(X_3|X_1, X_2)$$

  has **one degree of freedom** associated with it (the degrees of freedom for $\text{SSE}(X_1, X_2)$ minus the degrees of freedom for $\text{SSE}(X_1, X_2, X_3)$, or $(n-3)-(n-4)=1$). An extra sum of squares such as

$$\text{SSR}(X_3, X_4|X_1, X_2)$$

  has **two degrees of freedom** associated with it.

  In general, the **degrees of freedom** associated with an **extra sum of squares** is equal to the **number of predictors being added** to the model.

- For the case in which a **single predictor X_k** is added to a model, there are **two kinds** of **extra sums of squares**:

  ▷ The ***partial sum of squares*** (or ***Type II sum of squares***) is the extra sum of squares associated with adding X_k to a model that already includes ***all the other $p-2$* predictors**:

$$\text{SSR}(X_k|X_1, \ldots, X_{k-1}, X_{k+1}, \ldots, X_{p-1})$$

  ▷ The ***sequential sum of squares*** (or ***Type I sum of squares***) is the extra sum of squares associated with adding X_k to a model that already includes ***only the "previous" $k-1$* predictors**:

$$\text{SSR}(X_k|X_1, X_2, \ldots, X_{k-1})$$

- Note that, the **sequence** of extra sums of squares

$$\begin{aligned}\text{SSR}(X_1)\\ \text{SSR}(X_2|X_1)\\ \text{SSR}(X_3|X_1, X_2)\\ \vdots\\ \text{SSR}(X_{p-1}|X_1, X_2, \dots, X_{p-2})\end{aligned}$$

  will **depend** on the **order** in which the **predictors are specified in the model** when using **statistical software** to fit the model.

  In other words, when using **sequential sums of squares**, the contribution of a particular variable for explaining as-yet unexplained Y variation will **depend** on **which variables come before it** in the model specification.

  This is **not the case** for **partial sums of squares**. In other words, the **partial sum of squares** for a particular variable will be the **same regardless of the order** that the variables appear in the model specification.