Assignment 6 Q56 (june 2018)

Abhishek Ajit Sabnis

10 March 2022

56) To test the equality of effects of 10 schools against all alternatives, we take a random sample of 5 students from each school and note their marks in a common examination. "Between sum of squares" and "total sum of squares" are found to be 180 and 500 respectively. What is the p-value for the standard F-test?

Ans

Let X be n×m matrix where n is number of students from each school and m is total number of schools.

Let $\overline{X}$ be m × 1 column vector of mean marks of students of a particular school.

Let $\overline{\overline{X}}$ be mean marks of all students.

$$\Rightarrow \overline{X} = (X^T_{m\times n} 1_{n\times 1})/n \tag{1}$$

$$\begin{aligned} \Rightarrow \overline{\overline{X}} &= (1^T_{1\times m}\overline{X}_{m\times 1})/m \\ &= (1^T_{1\times m}X^T_{m\times n}1_{n\times 1})/mn \end{aligned} \tag{2}$$

The **Total sum of squares (SST)** is defined as -

$$\begin{aligned} SST &= Tr(X^TX) + mn(\overline{\overline{X}})^2 - 2\overline{\overline{X}}[1^T_{1\times n}X_{n\times m}1_{m\times 1}] \\ &= Tr(X^TX) - \frac{(1^TX^T1)(1^TX^T1)}{mn} \end{aligned} \tag{3}$$

The **Within sum of squares (SSW)** is defined as -

$$\begin{aligned} SSW &= Tr(X^TX) + n\overline{X}^T\overline{X} - 2[1^T_{1\times n}X_{n\times m}\overline{X}_{m\times 1}] \\ &= Tr(X^TX) - \frac{1^TXX^T1}{n} \end{aligned} \tag{4}$$

The **Between sum of squares (SSB)** is defined as -

$$\begin{aligned} SSB &= n[\overline{X}^T\overline{X} + m(\overline{\overline{X}})^2 - 2\overline{\overline{X}}(1_{1\times m}^T\overline{X}_{m\times 1})] \\ &= \frac{1^TXX^T1}{n} - \frac{(1^TX^T1)(1^TX^T1)}{mn} \end{aligned} \tag{5}$$

Add equations 4 and 5-

$$\Rightarrow SSB + SSW = SST \tag{6}$$

Now, Let the null Hypothesis be H_O = All the schools are same (true mean of marks are same). We conduct F-test to verify this.

$$\begin{aligned} F &= \frac{Var_B}{Var_W} \\ &= \frac{SSB/dof_b}{SSW/dof_w} \\ &= \frac{SSB/(m-1)}{SSW/(m(n-1))} \end{aligned} \tag{7}$$

Given SSB = 180 and SST = 500. From equation (6), SSB = 320. From question, m=10 and n=5

$$\Rightarrow F = 2.5 \tag{8}$$

Therefore, p-value is -

$$P[F_{9,40} \geq 2.5] \tag{9}$$