

F統計量と分散説明率

F統計量と分散説明率（決定係数）の定義の確認から始めよう。

$$F = \frac{MS_{model}}{MS_{residual}} = \frac{SS_{model}/p}{SS_{residual}/(n-p-1)}$$

$$R^2 = \frac{SS_{model}}{SS_{total}} = 1 - \frac{SS_{residual}}{SS_{total}} = \frac{SS_{model}}{SS_{model} + SS_{residual}}$$

分散説明率の式を変形すると,

$$\begin{aligned} SS_{model} &= R^2(SS_{model} + SS_{residual}) \\ (1 - R^2)SS_{model} &= R^2SS_{residual} \\ SS_{model} &= \frac{R^2}{1 - R^2}SS_{residual} \end{aligned}$$

これをF統計量の定義式に代入すると,

$$F = \frac{\left(\frac{R^2}{1 - R^2}SS_{residual}\right)/p}{SS_{residual}/(n-p-1)} = \frac{R^2/p}{(1 - R^2)/(n-p-1)}$$

となる。