F Test for Regression Relation

To test whether there is a regression relation between the response variable Y and the set of X variables $X_1, \ldots, X_{p-1}$, i.e., to choose between the alternatives:

$$\begin{aligned} &H_0\colon \beta_1 = \beta_2 = \cdots = \beta_{p-1} = 0 \\ &H_a\colon \text{not all } \beta_k \ (k = 1, \ldots, p-1) \text{ equal zero} \end{aligned} \tag{6.39a}$$

we use the test statistic:

$$F^* = \frac{MSR}{MSE} \tag{6.39b}$$

The decision rule to control the Type I error at α is:

$$\begin{aligned} &\text{If } F^* \leq F(1-\alpha; p-1, n-p), \text{ conclude } H_0 \\ &\text{If } F^* > F(1-\alpha; p-1, n-p), \text{ conclude } H_a \end{aligned} \tag{6.39c}$$

The existence of a regression relation by itself does not, of course, ensure that useful predictions can be made by using it.

Note that when $p - 1 = 1$, this test reduces to the F test in (2.60) for testing in simple linear regression whether or not $\beta_1 = 0$.

Coefficient of Multiple Determination

The coefficient of multiple determination, denoted by R^2, is defined as follows:

$$R^2 = \frac{SSR}{SSTO} = 1 - \frac{SSE}{SSTO} \tag{6.40}$$

It measures the proportionate reduction of total variation in Y associated with the use of the set of X variables $X_1, \ldots, X_{p-1}$. The coefficient of multiple determination R^2 reduces to the coefficient of simple determination in (2.72) for simple linear regression when $p - 1 = 1$, i.e., when one X variable is in regression model (6.19). Just as before, we have:

$$0 \leq R^2 \leq 1 \tag{6.41}$$

where R^2 assumes the value 0 when all $b_k = 0$ $(k = 1, \ldots, p-1)$, and the value 1 when all Y observations fall directly on the fitted regression surface, i.e., when $Y_i = \hat{Y}_i$ for all i.

Adding more X variables to the regression model can only increase R^2 and never reduce it, because SSE can never become larger with more X variables and $SSTO$ is always the same for a given set of responses. Since R^2 usually can be made larger by including a larger number of predictor variables, it is sometimes suggested that a modified measure be used that adjusts for the number of X variables in the model. The *adjusted coefficient of multiple determination*, denoted by R_a^2, adjusts R^2 by dividing each sum of squares by its associated degrees of freedom:

$$R_a^2 = 1 - \frac{\dfrac{SSE}{n-p}}{\dfrac{SSTO}{n-1}} = 1 - \left(\frac{n-1}{n-p}\right)\frac{SSE}{SSTO} \tag{6.42}$$

This adjusted coefficient of multiple determination may actually become smaller when another X variable is introduced into the model, because any decrease in SSE may be more than offset by the loss of a degree of freedom in the denominator $n - p$.

Comments

1. To distinguish between the coefficients of determination for simple and multiple regression, we shall from now on refer to the former as the coefficient of simple determination.

2. It can be shown that the coefficient of multiple determination R^2 can be viewed as a coefficient of simple determination between the responses Y_i and the fitted values $\hat{Y}_i$.

3. A large value of R^2 does not necessarily imply that the fitted model is a useful one. For instance, observations may have been taken at only a few levels of the predictor variables. Despite a high R^2 in this case, the fitted model may not be useful if most predictions require extrapolations outside the region of observations. Again, even though R^2 is large, MSE may still be too large for inferences to be useful when high precision is required. ■

Coefficient of Multiple Correlation

The coefficient of multiple correlation R is the positive square root of R^2:

$$R = \sqrt{R^2} \tag{6.43}$$

When there is one X variable in regression model (6.19), i.e., when $p-1 = 1$, the coefficient of multiple correlation R equals in absolute value the correlation coefficient r in (2.73) for simple correlation.

6.6 Inferences about Regression Parameters

The least squares and maximum likelihood estimators in $\mathbf{b}$ are unbiased:

$$\mathbf{E}\{\mathbf{b}\} = \boldsymbol{\beta} \tag{6.44}$$

The variance-covariance matrix $\boldsymbol{\sigma}^2\{\mathbf{b}\}$:

$$\underset{p\times p}{\boldsymbol{\sigma}^2\{\mathbf{b}\}} = \begin{bmatrix} \sigma^2\{b_0\} & \sigma\{b_0, b_1\} & \cdots & \sigma\{b_0, b_{p-1}\} \\ \sigma\{b_1, b_0\} & \sigma^2\{b_1\} & \cdots & \sigma\{b_1, b_{p-1}\} \\ \vdots & \vdots & & \vdots \\ \sigma\{b_{p-1}, b_0\} & \sigma\{b_{p-1}, b_1\} & \cdots & \sigma^2\{b_{p-1}\} \end{bmatrix} \tag{6.45}$$

is given by:

$$\underset{p\times p}{\boldsymbol{\sigma}^2\{\mathbf{b}\}} = \sigma^2(\mathbf{X}'\mathbf{X})^{-1} \tag{6.46}$$

The estimated variance-covariance matrix $\mathbf{s}^2\{\mathbf{b}\}$:

$$\underset{p\times p}{\mathbf{s}^2\{\mathbf{b}\}} = \begin{bmatrix} s^2\{b_0\} & s\{b_0, b_1\} & \cdots & s\{b_0, b_{p-1}\} \\ s\{b_1, b_0\} & s^2\{b_1\} & \cdots & s\{b_1, b_{p-1}\} \\ \vdots & \vdots & & \vdots \\ s\{b_{p-1}, b_0\} & s\{b_{p-1}, b_1\} & \cdots & s^2\{b_{p-1}\} \end{bmatrix} \tag{6.47}$$

is given by:

$$\underset{p\times p}{\mathbf{s}^2\{\mathbf{b}\}} = MSE(\mathbf{X}'\mathbf{X})^{-1} \tag{6.48}$$