LM02 Evaluating Regression Model Fit and Interpreting Model Results

Goodness of fit

R^2 and adjusted R^2

R^2 measures the percentage of variation in Y that is explained by the independent variables. In multiple regression, R^2 increases as we add new independent variables, even if the amount of variation explained by them is not statistically significant. Hence, the adjusted $\overline{R}^2$ is used because it does not automatically increase as independent variables are added to the model.

AIC and BIC

When evaluating a collection of models that explain the same dependent variable, we cannot rely on the adjusted R^2 alone. Two commonly used statistics for this purpose are Akaike's information criterion(AIC) and Schwarz's Bayesian information criterion (BIC). These stats are often provided as part of the regression software output.

Lower values of both measures are better.

When do we prefer one measure over the other?

- AIC is preferred if the model is used for prediction purposes
- BIC is preferred when the best goodness of fit is the goal

Testing joint hypothesis for coefficients

Hypothesis tests of a single coefficient

The hypothesis tests of a single coefficient in a multiple regression are identical to those in a simple regression.

If we are testing simply whether a variable is significant in explaining the dependent variable's variation, the hypotheses are:

$H_0: b_j = 0$

$H_a: b_j \neq 0$

The statistical software produces the t-statistics and P-values for a test of the slope coefficient against zero for each independent variable in the model. t-stats greater than the critical value indicate that the variable is significant.

Joint F-test

The joint F-test is used to jointly test a subset of variables in a multiple regression, where the "restricted" model is based on a narrower set of independent variables nested in the broader "unrestricted" model. The null hypothesis is that the slope coefficients of all independent variables outside the restricted model are zero.

It is calculated as:

$$F = \frac{(\text{Sum of squares error restricted model} - \text{Sum of squares error unrestricted})/q}{\text{Sum of squares error unrestricted model}/(n-k-1)}$$

where: q is the number of restrictions, i.e. the number of variables omitted

If the calculated F-stat exceeds the critical F-value we can conclude that at least one of the omitted variables is statically significant.

General linear F-test

The general linear F-test is an extension of the joint F-test, where we test the significance of the whole regression equation. The null hypothesis is that the slope coefficients on all independent variables are 0, against the alternative that at least one coefficient is different from 0.

The F-stat is calculated as:

$$F = \frac{\text{mean regression sum of squares}}{\text{mean squared error}} = \frac{MSR}{MSE}$$

If the calculated F-stat exceeds the critical F-value we can conclude that at least one of the slope coefficients is different from 0.

Forecasting using multiple regression

To forecast the value of a dependent variable using a multiple linear regression model, follow these three steps:

1. Obtain estimates $\widehat{b_0}, \widehat{b_1}, \widehat{b_2}, \ldots, \widehat{b_k}$ of the regression parameters $b_0, b_1, b_2, \ldots b_k$. The ˆ symbol indicates that the values are estimated.
2. Determine the assumed values of the independent variables $\widehat{X_{1\imath}}, \widehat{X_{2\imath}}, \ldots \ldots, \widehat{X_{k\imath}}$
3. Compute the predicted value of the dependent variable, $\widehat{Y\imath}$ using the following equation: $\hat{Y}_i = \widehat{b_0} + \widehat{b_1}\hat{X}_{1i} + \hat{b}_2\hat{X}_{2i} + \cdots \ldots + \widehat{b_k}\hat{X}_{ki}$

The level of uncertainty around the forecast of the dependent variable is called the standard error of forecast. This forecast error depends on how well the independent variables (X_1, X_2, X_3…) were forecasted i.e. the sampling error, as well as the model error.

The larger the sampling error, the larger is the standard error of the forecast of Y and the wider is the confidence interval.