Lecture 3 – F tests and other tests

3.1 Introduction to the $F$ test
The $F$ test is one of the most frequently quoted statistics as a measure of explanatory power. It can be used to determine the overall significance of a regression. Formally this involves testing the null hypothesis:

$$H_0 : \beta_1 = \beta_2 = \ldots = \beta_k = 0$$

So the assumption is that constant term is the only non-zero coefficient, and it will equal the mean value of $y$ (the dependent variable) i.e. $y = \bar{y} = \beta_0$. The alternative hypothesis is:

$$H_1 : \text{not all the } \beta_s = 0$$

If the null hypothesis is true, that is if all the true parameters are zero, there is no linear relationship between $y$ and the independent variables. The derivation of this test comes from analysis of variance theory. We shall not be going into that theory, merely presenting the results. Essentially, to conduct the $F$ test you first calculate:

$$F^* = \frac{\sum (\hat{y} - \bar{y})^2 / (k-1)}{\sum e^2 / (n-k)}$$

**Figure 1.1**

Then this value is compared with the theoretical value of $F$ obtained from the tables. This has two degrees of freedom: $v_1 = k-1$ and $v_2 = n-k$. If $F^* < F$ we accept the null hypothesis, that is we accept that the overall regression is not
significant. e are the residual from the regression, k the number of parameters, n the number of observations, \( \hat{y} \) the predicted value of y from the regression.

3.2 Testing the improvement of fit obtained from additional explanatory variables

This can be done by looking at the t statistics, but here we present an alternative way. Suppose we have the model:

\[
y = f(x_1, x_2, \ldots, x_m, x_{m+1}, \ldots, x_k)
\]

**Figure 2.1**

We first regress y on the m variables \( x_1, x_2, \ldots, x_m \) obtaining:

\[
\hat{y} = \hat{\beta}_0 + \hat{\beta}_1 x_1 + \ldots + \hat{\beta}_m x_m
\]

**Figure 2.2**

We then regress y on all the variables:

\[
\hat{y} = \hat{\beta}_0 + \hat{\beta}_1 x_1 + \ldots + \hat{\beta}_m x_m + \hat{\beta}_{m+1} x_{m+1} + \ldots + \hat{\beta}_k x_k
\]

**Figure 2.3**

The relevant F* ratio is:

\[
F^* = \frac{\sum (\hat{y} - \bar{y})^2 - \sum (\hat{y} - \bar{y} - \hat{\beta}_0)^2}{k - m} / \frac{\sum e^2}{n - k}
\]

**Figure 2.4**

This is, again, compared with the corresponding F statistic, with \( \nu_1 = k - m \) and \( \nu_2 = n - k \) degrees of freedom. The residual are divided by n-k this indicates that they relate from the regression in 2.3, i.e. the full regression.

3.3 Chow forecast test

The Chow forecast test is used to test for equality between coefficients obtained from different samples. Suppose we have two samples on the variables x and y, the one containing \( n_1 \) observations and the other \( n_2 \) observations. We thus obtain two estimates of the same relationship for two different periods (in cross-section they may be broken up into sub-samples through other reasoning):
\[ \hat{y}_1 = \hat{b}_0 + \hat{b}_1 x_1 \]
\[ \hat{y}_2 = \hat{b}_0 + \hat{b}_1 x_2 \]

Figure 3.1

We can also 'pool' the two sets of observations to get a third set of estimates:
\[ \hat{y}_p = \hat{a}_0 + \hat{a}_1 x_p \]

Figure 3.2

The 'unexplained variance' is denoted as before:
\[ \sum e_p^2 = \sum (y_p - \bar{y}_p)^2 - \sum (\hat{y}_p - \bar{y}_p)^2 \]

Figure 3.3

I do think the above is almost superfluous it is simply defining an error term. We can then form the ratio:
\[ F^* = \frac{\left[ \sum e_p^2 - (\sum e_1^2 + \sum e_2^2) \right] / k}{\left( \sum e_1^2 + \sum e_2^2 \right) / (n_1 + n_2 - 2k)} \]

Figure 3.4

We then test this \( F \) statistic in the normal way by comparing it with one obtained from the table, at the appropriate level of confidence, and with \( v_1 = k \) and \( v_2 = n_1 + n_2 - 2k \) degrees of freedom (note that degrees of freedom are such that \( v_1 = k \) is the divisor in the numerator and \( v_2 = n_1 + n_2 - 2k \) is the divisor in the denominator – this is always the case in the \( F \) test).

3.4 Testing of a restriction imposed on the relationship between two or more parameters of a function

To demonstrate this we take the Cobb Douglas production function:

\[ Q = \beta_0 L^{\beta_1} K^{\beta_2} \]
\[ \ln Q = \ln \beta_0 + \beta_1 \ln L + \beta_2 \ln K \]

Figure 4.1

We might wish to test whether we have constant returns to scale, i.e. whether \( \beta_1 + \beta_2 = 1 \). More generally we may be faced with the regression:
\[ y = \beta_0 + \beta_1 x_1 + \ldots + \beta_k x_k \]

**Figure 4.2**

Assume that we believe that \( \beta_1 = 1 \) and \( \beta_2 = \beta_3 \). We want to test the null hypothesis:

\[ H_0 : \text{the two restrictions are true} \]

Against:

\[ H_1 : \text{not all restrictions are true} \]

We first apply OLS to the original **unrestricted** function and obtain the sum of squared residuals \( \sum e^2 \) with \( n-k \) degrees of freedom. We then incorporate the restrictions into the model and obtain the **restricted form**.

\[ (y - x) = \beta_0 + \beta_2 (x_2 + x_3) + \beta_4 x_4 + \ldots + \beta_k x_k + \text{error} \]

**Figure 4.3**

Note the above is derivable from the restrictions. We are now estimating \( k-2 \) coefficients and therefore have \( n-(k-2) \) degrees of freedom for the 'restricted' sum of squared residuals \( \sum e^2_R \). Therefore, \( \sum e^2_R - \sum e^2 \) has \( n - k + 2 - n + k = 2 \) degrees of freedom. We can then calculate the \( F \) statistic:

\[ F* = \frac{\left( \sum e^2_R - \sum e^2 \right)/2}{\sum e^2/(n-K)} \]

**Figure 4.4**

This is then tested in the standard way.

**3.5 An intuitive view of the F test**

We will place this within the context of section 3.2, testing the improvement of fit obtained from additional explanatory variable. We had the \( F \) statistic given by:

\[ F* = \frac{\left( \sum (\hat{y} - \bar{y})^2 - \sum (\hat{y} - \bar{y})^2 \right)/k - m}{\sum e^2/(n-k)} \]

**Figure 2.4**
The denominator represents the unexplained variance divided by the number of
degrees of freedom, i.e. the average unexplained variance per degree of
freedom (to vary). The numerator represents the improvement in fit (or reduction
in the unexplained variance) divided by the increase in the number of explanatory
variables (k-m).

Now, k-m also represents the reduction in the number of degrees of freedom that
results from the addition of these explanatory variables.

If the additional explanatory variables are not valid determinants of the
independent variable then we would still expect some reduction in the
unexplained variance, simply because they have used up k-m degrees of
freedom. In fact as each degree of freedom gives rise to an average unexplained
sum of squares of \( \sum e^2/n - k \), then even if the additional variables have no
explanatory power we would expect the reduction of k-m degrees of freedom to
reduce the unexplained sum of squares by \( \left[ \sum e^2/(n - k) \right] k - m \). For example, suppose:

\[
\begin{align*}
\sum e^2 &= 1000 \\
  n &= 105 \\
  k &= 5 \\
  m &= 3 \\
\sum (\hat{y} - \bar{y})^2 &= 2000 \\
\sum (\hat{y} - \bar{y})^2 &= 1500
\end{align*}
\]

then:

\[
\frac{\sum e^2}{n - k} = \frac{1000}{100} = 10
\]

**Figure 5.1**

The average unexplained sum of squares per degree of freedom is 10 (*this is the
denominator in 2.4*). Now the introduction of 2 additional variables reduces the
degrees of freedom by 2. If these additional variables are not valid we would
expect them to cause a reduction in the total unexplained variance of 20 (10 for
each degree of freedom). In actual fact they cause a drop of 2000-1500=500,
which is a reduction of 250 per degree of freedom (or per additional variable; *this
is the numerator in 2.4*). This is considerably more than 10, and we can safely
conclude that they are actually significant determinants of the dependent variable
\( y \). Using the equation in Figure 2.4 we get a ratio of 250/10=25. If the two
variables were not significant we would expect a ratio of 1. 25 is obviously much
greater than 1, and we therefore accept the two additional variables as
significant.

Figure 5.2

25 lies well to the right of the critical value.

3.6 Ramsey RESET test
Run a regression of the form:

\[ y = X\beta + u \]

Figure 6.1

Which gives \( y^p = X\beta \). Then calculate \( y^{p^2} = y^p \times y^p \). Do a secondary regression:

\[ \hat{u} = a + by^{p^2} \]

Figure 6.2

i.e. regress the estimated residuals on \( y^{p^2} \). This is a test for correct functional
specification of the model. A more general format would be to estimate:

\[ y = X\beta + u \]

Figure 6.3

Then calculate \( y_i^{p^2} \) and estimate:

\[ y_i = X\beta + \alpha_i (y_i^{p^2})^2 + \theta_i (y_i^{p^3})^3 + \delta_i (y_i^{p^4})^4 + e_i \]

Figure 6.4
that is taking higher powers. Follow this by doing an $F$ test to see if $6.4$ is significantly better than $6.3$. It is distributed as $F(3, n - j - 4)$. If significant then reject the null hypothesis that $\alpha_i = \theta_i = \delta_i = 0$.

There is some disagreement as to what exactly this test does. For a linear model which is properly specified in functional form, nonlinear transforms of the fitted values should not be useful in predicting the dependent variable. Thus it seems a test of correct functional form. However, STATA says “Ramsey regression specification-error test for omitted variables” and the literature on the shadow economy e.g. has used it in this way. It can be used for either cross section or time series.

### 3.7 ARCH test

This is a test for AutoRegressive Conditional Heteroscedasticity done by regressing:

$$e^2_i = c_0 + \sum_{i=1}^{G} \gamma_i e^2_{i-i}$$

**Figure 7.1**

We are testing for significance of $\gamma_i \forall i$. This can be done using a Lagrange multiplier (LM) or an $F$ test. $e_i$ is the residual from a regression, not just an OLS regression. Indeed all these tests tend to be quite general with respect to technique. Clearly as time is involved in 7.1. It is a time series test only and cannot be used for cross section analysis.

### 3.8 Comfac test

The PC GIVE Manual explains this thus:


COMFAC tests for the legitimacy of common-factor restrictions of the form:

$$\alpha( L) b( L) y = \alpha( L) \sum_{i=1}^{r} b_{i}( L) x_{i} + u_{i}$$

(eq:18.10)

where $\alpha( L)$ is of order $r$ and denotes polynomials of the original order minus $r$. The degrees of freedom for the Wald tests for COMFAC are equal to the number of restrictions imposed by $\alpha( L)$. Failure to reject common-factor restrictions does not entail that such restrictions must be imposed. For a discussion of the theory of COMFAC, see Hendry and Mizon (1978) for some finite-sample Monte Carlo evidence see Mizon and Hendry (1980). COMFAC is not available for RALS.
L is a lagged operator, for example: \( a(L)y_t = a_1y_{t-1} + a_2y_{t-2} + a_3y_{t-3} + \cdots \)

Eq. 18.10 is thus a lag structure being imposed on the data.

As a specific example, Suppose we have a regression:

\[
y_t = \beta_0 + \beta_1 x_{it} + u_t
\]

**Figure 8.1**

Where it is assumed that \( u_t = \rho u_{t-1} + \varepsilon_t \). Now lag 8.1, and premultiply by \( \rho \):

\[
\rho y_t - 1 = \rho \beta_0 + \rho \beta_1 x_{it-1} + \rho u_{t-1}
\]

Subtract from 8.1:

\[
y_t - \rho y_{t-1} = \beta_0 (1 - \rho) + \beta_1 x_{it} - \beta_1 \rho x_{it-1} + \varepsilon_t
\]

**Figure 8.2**

Therefore this follows on to:

\[
y_t = \beta_0 (1 - \rho) + \rho y_{t-1} + \beta_1 x_{it} - \beta_1 \rho x_{it-1} + \varepsilon_t
\]

**Figure 8.3**

Do the expression in Figure 8.1 using ARI or some other first order autoregressive technique, and then estimate the expression in Figure 8.3 using OLS (can do as residuals (\( \varepsilon_t \)) should be white noise). If there is no significant improvement when using LM test or \( F \) test of 8.3 over 8.1 then accept the expression in Figure 8.1.

### 3.9 Jarque-Bera test for normality

Based on a paper from 1980, the key formula for this is:

\[
\chi^2 = \frac{T - k}{6} \left( SK^2 + \frac{1}{4} EK^2 \right)
\]

**Figure 9.1**
This gives the BJ stat. \( k=0 \) when \( x_i \) is an observed data series, and \( k \) is the number of regressors when \( x_i \) is the residual of the regression. \( SK \) is a measure of skewness:

\[
SK = \frac{1}{T-1} \sum_{i=1}^{T} (x_i - \bar{x})^3 \\
\left( \frac{1}{T-1} \sum_{i=1}^{T} (x_i - \bar{x})^2 \right)^{1.5}
\]

*Figure 9.2*

\( EK \) is a measure of excessive kurtosis:

\[
EK = \left[ \frac{1}{T-1} \sum_{i=1}^{T} (x_i - \bar{x})^4 \right] - 3
\]

*Figure 9.3*

\( T \) is the number of observations. The statistic is approximately distributed as \( \chi^2 \) with two degrees of freedom. Reject the null hypothesis of normality if BJ exceeds the critical \( \chi^2 \) value. The test is most frequently applied to the residuals from a regression.

*From Wikipedia:* The chi-square approximation, however, is overly sensitive (lacking specificity) for small samples, rejecting the null hypothesis often when it is in fact true. Furthermore, the distribution of p-values departs from a uniform distribution and becomes a right-skewed uni-modal distribution, especially for small p-values. This leads to a large Type I error rate. The table below shows some p-values approximated by a chi-square distribution that differ from their true alpha levels for very small samples.

### 3.10 Hendry forecast test

Calculate \( y_t = X\beta + u \) where \( E(u^2) = \sigma_u^2 \) and is estimated as \( \sum u_t^2 / (n-k) \). Use this equation to forecast into the future and let \( f_t \) be the forecast error. Calculate \( \sum f_t^2 / \sigma_u^2 \) and this is distributed as \( \chi^2(N)/N \) where \( N \) is the number of forecast periods.
3.11 Cointegration – Augmented Dickey-Fuller

This is done in more detail later. In the Augmented Dickey-Fuller test you do:

$$\Delta y_i = \sum_{i=1}^{L} x_i \Delta y_{i-1} + \beta y_{i-1}$$

Figure 11.1

This is testing whether $\beta$ is significantly different from zero, and is negative. Approximate critical value is -2.9 at 5%.

Useful references: