1 The F-Statistic

1.1 Joint Significance

Suppose we regress y on three different regressors, w, x and z, and both slope coefficients for x and z have high enough p-values that we conclude each one is statistically insignificant. It is still possible, however, that they may be *jointly* significant, even if they are individually insignificant. In other words, we want to test simultaneously that

$$H_0 : \beta_x = \beta_z = 0,$$
$$H_A : \text{at least one of } \beta_x, \beta_z \neq 0.$$

Think of H_0 as being a *restriction* placed on β_x and β_z that we want to test. If the restriction leads to goofy results, then we conclude that the restriction is a bad one and we instead accept the alternative, namely, that β_x and β_z are jointly significant.

The test proceeds as follows. The first thing to do is take the model where β_x and β_z are unrestricted (that is, a regression where x and z are included and thus their coefficients are estimated) and find its sum of squared residuals, call it RSS_{ur}. Then make the restrictions (by not even including them in the regression, which implicitly sets them equal to zero) and find that model's sum of squared residuals, call it RSS_r.

If β_x and β_z are jointly insignificant, i.e. if H_0 is true, then you would expect the difference between the two RSS terms to be small since the RSS represents unexplained variation in y. In other words, if β_x and β_z are jointly insignificant, then we shouldn't expect much difference in how well the model explain things (or how badly the model explains things, since we're using RSS) whether they're both simultaneously included or not. In other-other words, we're testing whether or not the regression is significantly (in the statistical sense) less bad when we include regressors x and z.

We just need to formalize what we mean by a "small" difference between the two. This is given by the F-statistic,

$$F \equiv \frac{(\text{RSS}_r - \text{RSS}_{ur})/(k-g)}{\text{RSS}_{ur}/(n-k)}, \tag{1}$$

which is drawn from the $F_{k-g,n-k}$ distribution, where

- n is the number of observations;
- k is the number of parameters being estimated in the unrestricted model, in this case

$k = 4$ because we estimate the intercept plus slope coefficients for w, x, and z;

- g is the number of parameters being estimated in the restricted model, in this case $g = 2$ because the restricted models omits x and z and hence only estimates the intercept and slope coefficient for w;
- $F_{k-g,n-k}$ is the F-distribution with $k - g$ parameters included in the restriction and $n - k$ is the unrestricted degrees of freedom.

Sometimes we write $q = k - g$ for the number of restrictions instead. In fact, I prefer that and it's very common and I will use it instead. Then we have

$$F \equiv \frac{(\text{RSS}_r - \text{RSS}_{ur})/(q)}{\text{RSS}_{ur}/(n-k)}, \tag{2}$$

which is drawn from the $F_{q,n-k}$ distribution.

1.2 Overall Significance

At the extreme end, we can test whether *all* regressors are jointly significant by comparing it to a regression with *no* regressors. There are k estimates, $k - 1$ of them are from regressors, therefore we are making $q = k - 1$ restrictions. Since we are imposing that all regressor coefficients are zero, it means the restricted model has zero explanatory power and therefore $\text{RSS}_r = \text{TSS}$. Using these two pieces of information, we can write the F-statistic as

$$F \equiv \frac{R^2/(k-1)}{(1-R^2)/(n-k)}, \tag{3}$$

where R^2 is from the unrestricted regression.

1.3 Individual Significance

At the other extreme end, we can use the F-test to test just one restriction, for example $H_0 : \beta_x = 0$ against $H_A : \beta_x \neq 0$. This looks like a simple, ordinary hypothesis test, and indeed, the F-statistic in this case will be the t-statistic squared.

1.4 Inference

The unrestricted model can never explain less than the restricted model, and it follows that $\text{RSS}_{ur} \leq \text{RSS}_r$. This means that the F-statistic can never be negative; and furthermore it means that we only reject the null hypothesis when the F-statistic is too large in the positive direction. Therefore inference only looks at the right-tail, so there is no need to chop the significance in half when finding critical values or multiply by 2 when finding the p-value. Also note that ***this inference is only valid with homoskedasticity!!!!11*** Could be a multiple choice question.

2 Examples

I now provide some worked-out examples from Colin Cameron's old exams.

2.1 Final 2016, Problem 3c

The question asks us to test whether the variables *radio*, *newspaper*, *tvbynews*, *region*1, and *region*2 are jointly significant. (Thus we already know that $q = 5$). That is, compared to the simple regression $sales = \beta_1 + \beta_2 tv + u$, we want to see if the regression

$$sales = \beta_1 + \beta_2 tv + \beta_3 radio + \beta_4 newspaper + \beta_5 tvbynews + \beta_6 region1 + \beta_7 region2 + u$$

is worth a damn. The test is

$$H_0 : \beta_3 = \beta_4 = \beta_5 = \beta_6 = \beta_7 = 0,$$
$$H_A : \text{at least one of } \beta_3, \beta_4, \beta_5, \beta_6, \beta_7 \neq 0.$$

Unrestricted Model. The unrestricted model is the regression

$$sales = \beta_1 + \beta_2 tv + \beta_3 radio + \beta_4 newspaper + \beta_5 tvbynews + \beta_6 region1 + \beta_7 region2 + u.$$

Thus the unrestricted model estimates $k = 7$ parameters. The Stata output on the exam indicates that $\text{RSS}_{ur} = 518350292$.

Restricted Model. The restricted model, which omits the variables in question, is

$$sales = \beta_1 + \beta_2 tv + u,$$

and thus the restricted model estimates $g = 2$ parameters. The Stata output on the exam indicates that $\text{RSS}_r = 2102500000$.

***F*-Statistic.** There are $n = 200$ observations. Hence the F-statistic is given by

$$F = \frac{(2102500000 - 518350292)/5}{518350292/(200-7)} \approx 117.97.$$

We are told that the test has a critical value of 2.261, and hence we reject the null (that the additional regressors are jointly insignificant) because $117.97 > 2.261$. The conclusion, then, is that the additional regressors *are* jointly significant.

2.2 Final 2016, Problems 6c and 6d

Part C, Method 1. By default, Stata does a test of overall significance, the F-statistic for which is reported in Stata as `F(2, 18) = (C)`. The number 2 comes from the fact that we are making $q = 2$ restrictions in the test for overall significance (one restriction for β_x and one for β_z), and the number 18 comes from the fact that the unrestricted model makes k estimates and therefore has $n - k = 21 - 3 = 18$ degrees of freedom.

Our task is to calculate this F-statistic. The formula for an overall significance test is

$$F \equiv \frac{R^2/(k-1)}{(1-R^2)/(n-k)}.$$

We have $k = 3$ and $n = 21$. R^2 is the explained variation of y around its mean, i.e.

$$R^2 = \frac{540}{720} = 0.75 \quad \Longrightarrow \quad F = \frac{0.75/(3-1)}{(1-0.75)/(21-3)} = 27.$$

Part C, Method 2. Alternatively, consider the restricted model to be the one with no regressors, that is, where $H_0 : \beta_x = \beta_z = 0$. In this case, there is no explained sum of squares since there are no regressors doing any explaining! Thus $\text{TSS} = \text{RSS}_r = 720$ and $q = 2$ because we are testing the two slope coefficients.

For the unrestricted case, i.e. the one where all of the regressors are used and estimated, we have $\text{TSS} = \text{ESS}_{ur} + \text{RSS}_{ur}$ implies that $\text{RSS}_{ur} = 720 - 540 = 180$, and $k = 3$. Hence the F-statistic is

$$F = \frac{(\text{RSS}_r - \text{RSS}_{ur})/(q)}{\text{RSS}_{ur}/(n-k)} = \frac{(720-180)/2}{180/(21-3)} = 27.$$

Part D. This is just an ordinary two-sided test, so we can calculate the t-statistic

$$t = \frac{2-0}{1} = 2 \quad \Longrightarrow \quad F = 2^2 = 4.$$