PAD 705 Handout: Hypothesis Testing on Multiple Parameters

In many cases we may wish to know whether two or more variables are *jointly significant* in a regression. T-tests are only useful in telling us whether a single variable is, by itself, statistically significant. It may be that two or more variables have statistically insignificant t scores but are jointly significant. This may occur if two or more variables are collinear with one another (i.e., have relatively high levels of correlation with one another). We may also wish to see if a group of unrelated variables with small t scores are jointly significant – if they aren't, we could drop the entire set, thus gaining power on our included variables. Finally, joint significance tests let us tell whether variables that measure the same information are all insignificant – for instance, we can only be sure "age" is insignificant in a regression where we used a quadratic form if we test that both "age" and "age2" are jointly insignificant.

In many large datasets, it usually isn't terribly important to reduce the number of variables included in the regression – there are usually degrees of freedom to spare. Also, removing statistically insignificant variables may introduce bias to our coefficient estimates, so joint significance tests should never be used as the sole criteria by which to remove variables with small t scores.

The manual way to calculate joint significance is to run an "unrestricted regression" – one which includes all the variables of interest – and then run a "restricted" regression – one where variables with small t scores are dropped. The standard we use to determine whether the variables are jointly significant is whether the increase in "unexplained" variation in the model (Error Sum of Squares or "ESS" – which Stata unhelpfully calls the "Residual Sum of Squares") grows "too much" – as measured by the F distribution. Remember, we can increase the Regression Sum of Square (RSS) (the part of the variation explained by our regression model – which Stata unhelpfully calls "Model Sum of Square") by adding additional variables – but some of those added variables may not increase RSS enough to significantly improve our model's power. In those cases, we may wish to drop the set of low t score variables.

Set-up

```
. use "H:\Rockefeller Courses\PAD705\Problem Set Data\cps83.dta", clear
. gen exper2=exper^2
. gen hwage= wklywage/ wklyhrs
. gen lhwage=log( hwage)
. gen fem=(sex==2)
. gen fexper=fem*exper2
. gen fexper2=fem*exper2
```

The "Unrestricted" Regression -

```
. reg lhwage exper exper2 yrseduc fem fexper fexper2
```

Source	SS	df	MS
Model	115.86953	6	19.3115884
Residual	**194.215486**	993	.195584578
Total	310.085016	999	.310395411

Number of obs	=	1000
F(6, 993)	=	98.74
Prob > F	=	0.0000
R-squared	=	**0.3737**
Adj R-squared	=	0.3699
Root MSE	=	.44225

lhwage	Coef.	Std. Err.	t	P>|t|	[95% Conf.	Interval]
exper	.0555183	.0042959	12.92	0.000	.0470882	.0639483
exper2	-.0008826	.0000882	-10.01	0.000	-.0010557	-.0007095
yrseduc	.0773751	.0053197	14.54	0.000	.0669358	.0878143
fem	-.0329736	.0617353	-0.53	0.593	-.1541203	.0881731
fexper	-.0263973	.0065542	-4.03	0.000	-.0392589	-.0135357
fexper2	.0003684	.0001361	2.71	0.007	.0001012	.0006355
_cons	.5645578	.0801279	7.05	0.000	.4073183	.7217972

The "Restricted" Regression – testing whether fexper and fexper2 are jointly significant

```
. reg lhwage exper exper2 yrseduc fem
```

Source	SS	df	MS
Model	109.934536	4	27.4836341
Residual	**200.150479**	995	.201156261
Total	310.085016	999	.310395411

Number of obs	=	1000
F(4, 995)	=	136.63
Prob > F	=	0.0000
R-squared	=	**0.3545**
Adj R-squared	=	0.3519
Root MSE	=	.4485

lhwage	Coef.	Std. Err.	t	P>|t|	[95% Conf.	Interval]
exper	.0435066	.0032881	13.23	0.000	.0370541	.0499591
exper2	-.0007107	.0000687	-10.34	0.000	-.0008456	-.0005758
yrseduc	.0789745	.0053867	14.66	0.000	.0684038	.0895452
fem	-.3188769	.0285882	-11.15	0.000	-.3749769	-.2627768
_cons	.6709668	.0785873	8.54	0.000	.5167509	.8251828

For the regression equation:

$$\text{lhwage} = \beta_o + \beta_1 \text{exper} + \beta_2 \text{exper2} + \beta_3 \text{yrseduc} + \beta_4 \text{fem} + \beta_5 \text{fexper} + \beta_6 \text{fexper2} + \varepsilon$$

the null hypothesis is:

$$H_O: \beta_5 = \beta_6 = 0$$

Compute the F-statistic:

$$F_{q,\,N-k} = \frac{(RSS_R - RSS_{UR})/q}{RSS_{UR}/(N-k)}$$

Where:

RSS_{UR} = Residual sum of squares, unrestricted
RSS_R = Residual sum of squares, restricted
q = number of restrictions (here, the number of variables set equal to zero)
N = population size
k = number of variables in the regression, including the constant

Here, we get:

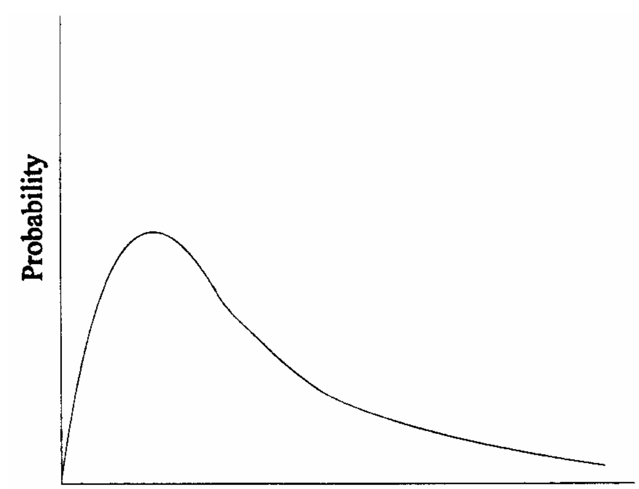
$$F = \frac{(200.150479 - 194.215) / 2}{194.215 / (1000 - 7)} = 15.1725$$

To test hypotheses regarding more than one regression parameter, we use the F distribution. This distribution is appropriate in this case, because the ratio of two Chi-Square distributed variables is distributed as F with the two degrees of freedom corresponding to the degrees of freedom in the numerator and denominator of the ratio, respectively. The test statistic we compute is the ratio of sums of squares, and is therefore the ratio of two Chi-Square variables.

The F-distribution takes only positive values. In the case of testing restrictions in a regression equation, clearly the computed F-statistic will always be positive since the error sum of squares cannot be made smaller by restricting the regression. The F-test is testing whether the error sum of squares is made “significantly” larger by imposing the restriction or whether the regression equation fits roughly equally well with or without the restrictions (in which case we do not reject the null hypothesis that the restrictions belong).

To perform the hypothesis test, compare the computed value of F to the critical value $F_{q,N-k}$ for a particular significance level (5% for example), where q is the number of restrictions imposed on the regression equation (here, q = 2) and N-k is the number of observations minus the number of variables in the regression (including the constant – here, k = 7). If the computed value exceeds the critical value, we can reject the null hypothesis.

Graphically, if the computed value exceeds the critical value, we are in the tail of the distribution:



The cutoff for 5% significance for F with a numerator degrees of freedom of 2 and a denominator degrees of freedom that is greater than 120 (see page 606 in Pyndyck and Rubinfeld) is 3.00. Thus there is no chance that these coefficients are actually equal to zero.

There is also a second way to calculate joint significance which relies on the *unadjusted* R^2. The general formula for this calculation is:

$$F_{q,\, N-k} = \frac{(R^2_{UR} - R^2_R) / q}{(1 - R^2_{UR}) / (N-k)}$$

Where:

R^2_{UR} = unadjusted R^2 of unrestricted regression
R^2_R = unadjusted R^2 of restricted regression
N, k, and q as defined above

In this case we get:

$$F = \frac{(0.3737 - 0.3545) / 2}{(1 - 0.3545) / (1000 - 7)} = 14.7681$$

Notice that the results, while similar, are not exactly the same because the R^2 statistics are rounded to only four decimal places. What this suggests is that when the appropriate measures of sums of squares are available you should always use the sum of squares formula instead of the R^2 formula.

Stata, of course, will run a joint significance test for you by invoking the `test` command after you run the *unrestricted regression.* Stata can execute several types of tests. First, let's test to see if both `fexper` and `fexper2` are equal to zero:

```
. test  fexper fexper2

 ( 1)  fexper = 0.0
 ( 2)  fexper2 = 0.0

       F(  2,   993) =   15.17
            Prob > F =    0.0000
```

We might also wish to test if two variables are equal to one another. The squared term tells us the rate at which the change in wage accelerates or decelerates, so it might be of interest to know whether acceleration/deceleration is equal for men and women. In this case, the hypothesis is:

$$H_O: \beta_2 = \beta_6$$

```
. test exper2 = fexper2

 ( 1)  exper2 - fexper2 = 0

       F(  1,   993) =   37.58
            Prob > F =    0.0000
```

In this case, the rate of deceleration for men and women is not equal – it is faster for women than for men. Notice also that the denominator degree of freedom is only 1. This is because we are not "restricting" the value of two coefficients – only one. When we ask whether two coefficients are really equal to zero, we are "restricting" the value of both β_5 and β_6 to be zero. Thus the number of restrictions, q, is equal to 2. When we ask whether $\beta_2 = \beta_6$, we are still allowing either β_2 or β_6 to be whatever they were originally estimated to be. So if we let β_2 be whatever we estimated to be, β_6 is "restricted" to being equal to the value established for β_2. For this reason q = 1.

We could ask another question: do the two parameters add up to some value. For instance, we might want to know whether $\beta_2 + \beta_6 = -0.03$. Here is the null hypothesis:

$$H_O: \beta_2 + \beta_6 = -0.03$$

```
. test exper2 + fexper2 = -0.03

 ( 1)  exper2 + fexper2 = -.03

       F(  1,   993) =79133.39
            Prob > F =    0.0000
```

Notice that, once again, the numerator degree of freedom (i.e., q) is equal to 1. As in the previous example, the value for β_2 or β_6 may float freely. However, once a value for β_2 is established, the hypothesized relationship establishes what the value of β_6 must be. For instance, if $\beta_2 = -0.20$, then β_6 must be equal to -0.10 according to the null hypothesis. Since only one restriction is created the numerator degree of freedom is just 1.

In this case, asking whether $\beta_2 + \beta_6 = -0.03$ is rather non-sensical. In other situations that is not the case. For instance, when evaluating Cobb-Douglas production functions, we often wish to know whether the coefficients on labor and capital add up to one, since that tells us whether this particular production technology generates increasing, decreasing, or constant returns to scale.

All of the joint significance tests done using the `test` command can be done manually as well. The trick is to substitute the hypothesized relationship into the original regression equation. Let's work this out for the last example. Again, the hypothesis is:

$$H_O: \beta_2 + \beta_6 = -0.03$$

And the regression equation is:

$$lhwage=\beta o + \beta_l exper + \beta_2 exper2 + \beta_3 yrseduc + \beta_4 fem + \beta_5 fexper + \beta_6 fexper2 + \varepsilon$$

We can solve the hypothesis in terms of one of the coefficients – I'll solve for β_6 in this case:

$$\beta_6 = -0.03 - \beta_2$$

We now need to substitute into the originally regression equation:

$$lhwage=\beta o + \beta_l exper + \beta_2 exper2 + \beta_3 yrseduc + \beta_4 fem + \beta_5 fexper + (-0.03 - \beta_2)fexper2 + \varepsilon$$

We need to collect terms to make this more amenable to regression, so first multiple (-0.03 - β_2) into fexper2:

$$lhwage=\beta o + \beta_l exper + \beta_2 exper2 + \beta_3 yrseduc + \beta_4 fem + \beta_5 fexper - 0.03fexper2 - \beta_2 fexper2 + \varepsilon$$

The expression (- 0.03fexper2) is just a constant – it is the variable fexper2 multiplied by -0.03. Since this expression has no coefficient in front of it, it needs to be subtracted from the dependent variable. We also need to collect the two terms that are related to β_2. So now the regression equation is:

$$lhwage - 0.03fexper2=\beta o + \beta_l exper + \beta_2(exper2 – fexper2) + \beta_3 yrseduc + \beta_4 fem + \beta_5 fexper + \varepsilon$$

This equation looks a little funny, but we can transform the variables so that it may be regressed – and in fact, Stata is doing the transformations on the fly when you use the `test` command. To do this manually, we have to generate two new variables: first, we need to create a new variable equal to (lhwage - 0.03fexper2); we'll call it `newDV`. `newDV` is now the dependent variable in the "restricted" regression. Next, we need to create a second new variable that is equal to (exper2 – fexper2); we'll call this one `newbeta2`. Let's look at how to do this in Stata:

```
. gen newDV =  lhwage - .03*fexper2

. gen newbeta2 =  exper2- fexper2
```

Unrestricted Regression

```
. reg  lhwage exper exper2 yrseduc fem fexper fexper2

      Source |       SS       df       MS              Number of obs =    1000
-------------+------------------------------           F(  6,   993) =   98.74
       Model |  115.86953     6  19.3115884            Prob > F      =  0.0000
    Residual |  194.215486   993  .195584578           R-squared     =  0.3737
-------------+------------------------------           Adj R-squared =  0.3699
       Total |  310.085016   999  .310395411           Root MSE      =  .44225

------------------------------------------------------------------------------
      lhwage |      Coef.   Std. Err.      t    P>|t|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
       exper |   .0555183   .0042959    12.92   0.000      .0470882    .0639483
      exper2 |  -.0008826   .0000882   -10.01   0.000     -.0010557   -.0007095
     yrseduc |   .0773751   .0053197    14.54   0.000      .0669358    .0878143
         fem |  -.0329736   .0617353    -0.53   0.593     -.1541203    .0881731
      fexper |  -.0263973   .0065542    -4.03   0.000     -.0392589   -.0135357
     fexper2 |   .0003684   .0001361     2.71   0.007      .0001012    .0006355
       _cons |   .5645578   .0801279     7.05   0.000      .4073183    .7217972
------------------------------------------------------------------------------
```

Restricted Regression

```
. reg   newDV exper  newbeta2 yrseduc fem fexper

      Source |       SS       df       MS              Number of obs =    1000
-------------+------------------------------           F(  5,   994) = 2712.51
       Model |  228817.983     5  45763.5967           Prob > F      =  0.0000
    Residual |  16770.0623   994  16.8712901           R-squared     =  0.9317
-------------+------------------------------           Adj R-squared =  0.9314
       Total |  245588.046   999   245.83388           Root MSE      =  4.1075

------------------------------------------------------------------------------
       newDV |      Coef.   Std. Err.      t    P>|t|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
       exper |   .0468307   .0398979     1.17   0.241    -.0314631    .1251246
    newbeta2 |  -.0005619   .0008192    -0.69   0.493    -.0021694    .0010456
     yrseduc |   .2054479   .0492387     4.17   0.000     .1088242    .3020717
         fem |   8.802797    .499299    17.63   0.000     7.822996    9.782598
      fexper |   -1.38974   .0425876   -32.63   0.000    -1.473311   -1.306168
       _cons |  -1.096138   .7423132    -1.48   0.140    -2.552819    .3605427
------------------------------------------------------------------------------
```

$$F = \frac{(16770.0623 - 194.215486) / 1}{194.215486 / (1000 - 7)} = 84750.284$$

Rounding errors explain the difference between this statistic and the F statistic reported by Stata. In any event, this procedure provides a way to manually test certain joint significance hypotheses. It can be replicated for any hypothesized relationship between two or more variables in a data set.

Note: It is fair game on the exam for me to ask you to calculate the F statistic from a restricted and unrestricted regression output. Do not become overly reliant on Stata.