Testing Linear Hypotheses

The data set lin.dta contains hourly pay data for 4029 individuals alongside some personal and firm characteristics. A regression of log hourly pay on some covariates gives

```
. reg lhw   female tenure edage numkids single

      Source |       SS       df       MS              Number of obs =    4029
-------------+------------------------------           F(  5,  4023) =  143.22
       Model |   210.63651     5   42.127302           Prob > F      =  0.0000
    Residual |  1183.31224  4023  .294136773           R-squared     =  0.1511
-------------+------------------------------           Adj R-squared =  0.1501
       Total |  1393.94875  4028  .346064733           Root MSE      =  .54234
------------------------------------------------------------------------------
         lhw |      Coef.   Std. Err.      t    P>|t|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
      female |  -.2595874   .0172446   -15.05   0.000    -.2933962   -.2257785
      tenure |    .015719   .0011159    14.09   0.000     .0135312    .0179067
       edage |  -.0035987   .0005904    -6.10   0.000    -.0047562   -.0024411
     numkids |  -.0035315   .0083279    -0.42   0.672    -.0198588    .0127958
      single |  -.1362066   .0212375    -6.41   0.000    -.1778439   -.0945693
       _cons |   2.140024   .0220509    97.05   0.000     2.096792    2.183256
------------------------------------------------------------------------------
```

A test of a linear restriction that the coefficient on a single variable (number of children) equals zero is done by the following command

```
. test numkids

 ( 1)  numkids = 0

       F(  1,  4023) =    0.18
            Prob > F =    0.6715
```

Since estimated F value is **below** the critical value ($F[1, \infty)^{5\%level} = 3.84$), can't reject null that true coefficient is zero

Note that this is equivalent to the t value on the numkids variable (square of t is F), the usual way of testing the significance of a single variable

A test of the hypothesis that the effect of **all** the covariates is zero is given by

```
. test   female tenure edage numkids single

 ( 1)  female = 0
 ( 2)  tenure = 0
 ( 3)  edage = 0
 ( 4)  numkids = 0
 ( 5)  single = 0

       F(  5,  4023) =  143.22
            Prob > F =    0.0000
```

Since estimated F value is **above** the critical value ($F[5, \infty)^{5\%level} = 2.21$), reject null that true effects of all explanatory variables are zero

Note that this is equivalent to

$$F = \frac{ESS/k-1}{TSS/N-k} = \frac{R^2/k-1}{(1-R^2)/N-k} \sim F[k-1, N-k]$$

which is given in the upper right hand corner of the stata output

A test of the equality of coefficients is given by

```
. test   edage=numkids

 ( 1)  edage - numkids = 0

       F(  1,  4023) =    0.00
            Prob > F =    0.9937
```

Should recognise that this is given by

$$(\beta_i - \beta_j)^2 / \text{Var}(\hat{\beta}_i - \hat{\beta}_j) \sim F(1, N-k)$$

and the appropriate values for this test can be found in the matrix of variance and covariances of the estimated coeffiecients ($\text{Var}(\hat{\beta}) = s^2(X'X)^{-1}$) along with the estimated s^2 value in the regression output above (.294)

```
. matrix list e(V)

symmetric e(V)[6,6]
             female      tenure       edage     numkids      single       _cons
 female   .00029737
 tenure   2.339e-06   1.245e-06
  edage  -1.625e-07   3.737e-08   3.486e-07
numkids   1.018e-06   1.304e-06  -8.406e-07   .00006935
 single   .00002647   6.275e-06  -4.141e-06   .00004839   .00045103
  _cons  -.00017637  -.00001437  -5.396e-06  -.00006734  -.00015142   .00048624
```

Tests for the subset of coefficients having zero effect in the model are given by

```
. test   edage numkids

 ( 1)  edage = 0
 ( 2)  numkids = 0

       F(  2,  4023) =   19.68
            Prob > F =    0.0000

. test  female tenure

 ( 1)  female = 0
 ( 2)  tenure = 0

       F(  2,  4023) =  241.87
            Prob > F =    0.0000
```

This is calculated on the basis of $F = \frac{(RSS_{restricted} - RSS_{unrestricted})/k2}{RSS_{unrestricted}/N-k} \sim f[k2, N-k]$