restricted vs. full F tests

In many situations, we want to compare the fit of two linear models

$$E(Y) = \beta_0 + \beta_1 X_1 + \beta_2 X_2 + \beta_3 X_3 + \dots + \beta_q X_q + \beta_{q+1} X_{q+1} + \dots + \beta_p X_p$$

$$E(Y) = \beta_0 + \beta_1 X_1 + \beta_2 X_2 + \beta_3 X_3 + \dots + \beta_q X_q$$

where the X variables in the restricted model are a subset of the X variables in the full model.

If X_{q+1} to X_p are not included in the model, their coefficients are being set to 0.

The restricted model estimates β_1 to β_q under the hypothesis that ***all*** of β_{q+1} to β_p are equal to 0 in the population.

The restricted vs. full F test uses the SSE from each model and the corresponding dfe to construct a formal hypothesis test.

The SSE from the full model will always be less than or equal to the SSE from the restricted model.

The test answers the question: does the reduction in SSE provide enough evidence that the full model describes the population more accurately than the restricted model?

The F statistic is calculated as:

$$(SSE_{rest} - SSE_{full}) / (dfe_{rest} - dfe_{full}) \quad \big/ \quad SSE_{full} / dfe_{full}$$

(The denominator is equal to the MSE for the full model.)

The degrees of freedom for this F ratio are $dfe_{rest} - dfe_{full}$ and dfe_{full}
For the test set up here, the difference $dfe_{rest} - dfe_{full}$ is equal to $p - q$,
the number of variables deleted from the restricted model.

The hypotheses may be stated in terms of the coefficients

H_0 : $\beta_{q+1} = 0$ AND $\beta_{q+2} = 0$ AND ... AND $\beta_p = 0$

Ha : not all of $\beta_{q+1}, \beta_{q+2}, \dots, \beta_p$ are equal to 0

The alternative hypothesis says ***at least one*** of the coefficients $\beta_{q+1}, \beta_{q+2}, \dots, \beta_p$ is not 0.
The alternative hypothesis ***does not say*** that all of $\beta_{q+1}, \beta_{q+2}, \dots, \beta_p$ are not equal to 0!!

To use an F test, the models must be "nested" –
the variables in the restricted model must all be in the full model.

Notes:

1. The use of the F distribution to obtain P values (or critical values for some specified α level) requires the assumption that the ε are normally distributed.

2. The restricted vs. full F test is a relative comparison of model fit, not an absolute one.

It is useful to examine the root MSE and the R^2 from the two models to help determine the quality of the fit.

3. The full and restricted models are also relative; one model may be the full model in one test and the restricted model in another test!

Implementation in Stata:

Be sure that you can calculate the F statistic by fitting the full and restricted models, getting the SSEs and the corresponding dfe values, and using the formula above.

The Ftail function in stata can be used to find P values

If the computed F statistic is 4.83 with 2 and 45 df,

`display Ftail(2, 45, 4.83)` will give the P value.

Be careful not to space between Ftail and the opening parenthesis.

To carry out the test in Stata

Fit the full model with the regress command

Use the test command followed by the names of the variables deleted from the restricted model.

example:

```
regr poverty metro hsgrad snglpar

test metro snglpar
```

will give the F ratio to test the null hypothesis that the coefficients on metro and snglpar are both equal to 0, comparing the model with 3 predictors to the model that just includes hsgrad.

examples with the Agresti and Finlay crime data

```
. regr poverty metro white hsgrad snglpar

      Source |       SS           df       MS      Number of obs   =        51
-------------+----------------------------------   F(4, 46)        =     34.22
       Model |  786.486789         4  196.621697   Prob > F        =    0.0000
    Residual |  264.276748        46  5.74514669   R-squared       =    0.7485
-------------+----------------------------------   Adj R-squared   =    0.7266
       Total |  1050.76354        50  21.0152707   Root MSE        =    2.3969

------------------------------------------------------------------------------
     poverty |      Coef.   Std. Err.      t    P>|t|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
       metro |  -.0324866    .016552    -1.96   0.056    -.0658041    .0008309
       white |   .0460236   .0363422     1.27   0.212    -.0271294    .1191767
      hsgrad |  -.5520469    .064914    -8.50   0.000    -.6827121   -.4213817
     snglpar |   1.141768   .2121786     5.38   0.000     .7146742    1.568861
       _cons |   41.72504   6.427176     6.49   0.000      28.7878    54.66227
------------------------------------------------------------------------------
```

Can we take out both metro and white from the model? Both have P > .05 ?

(We already know the answer is No, from our model using metro, hsgrad, and snglpar)

```
. regr poverty  hsgrad snglpar

      Source |       SS           df       MS      Number of obs   =        51
-------------+----------------------------------   F(2, 48)        =     58.55
       Model |  745.265099         2   372.63255   Prob > F        =    0.0000
    Residual |  305.498437        48  6.36455077   R-squared       =    0.7093
-------------+----------------------------------   Adj R-squared   =    0.6971
       Total |  1050.76354        50  21.0152707   Root MSE        =    2.5228

------------------------------------------------------------------------------
     poverty |      Coef.   Std. Err.      t    P>|t|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
      hsgrad |  -.5369575   .0653997    -8.21   0.000    -.6684525   -.4054625
     snglpar |    .874346   .1723884     5.07   0.000     .5277359    1.220956
       _cons |   45.28522   5.750212     7.88   0.000     33.72365     56.8468
------------------------------------------------------------------------------
```

SS resid restricted	305.498	df resid = 48
SS resid full model	264.277	df resid = 46

F statistic = [(305.498 − 264.277) / (48 − 46)] / [264.277 / 46] = 3.5875

Stata: `. display Ftail(2,46,3.5875)`   .03565806

The restricted model never fits better, but sometimes it fits ***as well as*** the full model.

Implementing in Stata:

```
. regr poverty metro white hsgrad snglpar

      Source |       SS           df       MS      Number of obs   =        51
-------------+----------------------------------   F(4, 46)        =     34.22
       Model |  786.486789         4  196.621697   Prob > F        =    0.0000
    Residual |  264.276748        46  5.74514669   R-squared       =    0.7485
-------------+----------------------------------   Adj R-squared   =    0.7266
       Total |  1050.76354        50  21.0152707   Root MSE        =    2.3969

------------------------------------------------------------------------------
     poverty |      Coef.   Std. Err.      t    P>|t|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
       metro |  -.0324866    .016552    -1.96   0.056    -.0658041    .0008309
       white |   .0460236   .0363422     1.27   0.212    -.0271294    .1191767
      hsgrad |  -.5520469    .064914    -8.50   0.000    -.6827121   -.4213817
     snglpar |   1.141768   .2121786     5.38   0.000     .7146742    1.568861
       _cons |   41.72504   6.427176     6.49   0.000      28.7878    54.66227
------------------------------------------------------------------------------

. test metro white

 ( 1)  metro = 0
 ( 2)  white = 0

       F(  2,    46) =    3.59
            Prob > F =    0.0357
```

If the models differ by a single variable, the restricted vs. full F test gives exactly the same answer as the t test on that variable in the full model. $\{\text{t statistic}\}^2 = \text{F statistic}$

```
. regr poverty metro white hsgrad snglpar

      Source |       SS           df       MS      Number of obs   =        51
-------------+----------------------------------   F(4, 46)        =     34.22
       Model |  786.486789         4  196.621697   Prob > F        =    0.0000
    Residual |  264.276748        46  5.74514669   R-squared       =    0.7485
-------------+----------------------------------   Adj R-squared   =    0.7266
       Total |  1050.76354        50  21.0152707   Root MSE        =    2.3969

------------------------------------------------------------------------------
     poverty |      Coef.   Std. Err.      t    P>|t|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
       metro |  -.0324866    .016552    -1.96   0.056    -.0658041    .0008309
       white |   .0460236   .0363422     1.27   0.212    -.0271294    .1191767
      hsgrad |  -.5520469    .064914    -8.50   0.000    -.6827121   -.4213817
     snglpar |   1.141768   .2121786     5.38   0.000     .7146742    1.568861
       _cons |   41.72504   6.427176     6.49   0.000      28.7878    54.66227
------------------------------------------------------------------------------

. test white

 ( 1)  white = 0
       F(  1,    46) =    1.60
       Prob > F =    0.2117
```

$(.046/.03634)^2 = 1.60$

two-sided P values the same

the not-so-useful overall F test

null hypothesis model: all of the Xs in the model have population slopes = 0

none of the Xs are useful in explaining variability in Y

model E(Y) = constant with constant estimate = $\bar{y}$

alt hypothesis model: the regression does have some useful information
for explaining variability in Y values

some of the Xs, but not necessarily all,
have population slopes that are not 0

model E(Y) = $\beta_0 + \beta_1 X_1 + \beta_2 X_2 + \beta_3 X_3 + \ldots + \beta_p X_p$
describes the population better

Every Statistical Software system does this test for every regression fit.

Under the null hypothesis, the Sum of squares residual is simply $\sum_{i=1}^{n} (y_i - \bar{y})^2$ = SS total

The F statistic becomes

{ (SStotal – SSresid) / ([n – 1] – [n – (p + 1)]) } / { SSresid / [n – (p + 1)] }

{ SS model / p } / MS resid or MS model / MS resid

with df p and [n – (p + 1)]

model for poverty with metro, white, hsgrad and snglpar

MS model = 786.487 / 4 = 196.62

Ms resid = 264.277/ 46 = 5.745

F = 34.22 4, 46 df P < .00005

Our model is better than model with no Xs, even though we have P = .212 on white.

How can we compare “non-nested” models?

Compare adjusted R^2 values.

```
. regr poverty metro hsgrad

      Source |       SS           df       MS      Number of obs   =        51
-------------+----------------------------------   F(2, 48)        =     30.23
       Model |  585.775516         2  292.887758   Prob > F        =    0.0000
    Residual |  464.988021        48  9.68725043   R-squared       =    0.5575
-------------+----------------------------------   Adj R-squared   =    0.5390
       Total |  1050.76354        50  21.0152707   Root MSE        =    3.1124

------------------------------------------------------------------------------
     poverty |      Coef.   Std. Err.      t    P>|t|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
       metro |  -.0132573   .0200467    -0.66   0.512    -.0535639    .0270493
      hsgrad |  -.6100675   .0787127    -7.75   0.000      -.76833    -.451805
       _cons |   61.65373   6.170625     9.99   0.000     49.24686    74.06061
------------------------------------------------------------------------------

. regr poverty metro snglpar

      Source |       SS           df       MS      Number of obs   =        51
-------------+----------------------------------   F(2, 48)        =     12.65
       Model |  362.693435         2  181.346718   Prob > F        =    0.0000
    Residual |  688.070101        48  14.3347938   R-squared       =    0.3452
-------------+----------------------------------   Adj R-squared   =    0.3179
       Total |  1050.76354        50  21.0152707   Root MSE        =    3.7861

------------------------------------------------------------------------------
     poverty |      Coef.   Std. Err.      t    P>|t|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
       metro |  -.0454657   .0252529    -1.80   0.078      -.09624    .0053086
     snglpar |   1.307678   .2613634     5.00   0.000     .7821721    1.833185
       _cons |   2.512668   3.053136     0.82   0.415    -3.626074    8.651409
------------------------------------------------------------------------------
```