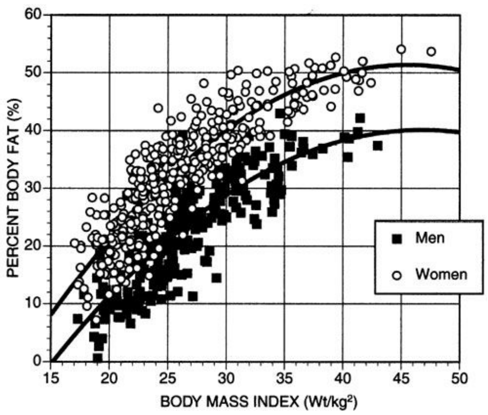
Playing with Bodyfat: F tests in action

About that BMI metric: $BMI = \frac{wgt\_kg}{hgt\_m^2}$

If the BMI is a useful metric, then it should do a good job explaining body fat levels:



$$Brozek = \alpha BMI^{\beta} = \alpha \left[\frac{wgt\_kg}{hgt\_m^2}\right]^{\beta} \quad \ldots \text{ or}$$

$$\ln(Brozek) = \ln(\alpha) + \beta \ln(BMI)$$
$$= \ln(\alpha) + \beta \ln(wgt\_kg) - 2\beta \ln(hgt\_m)$$

So if we regress Brozek on lnhgt and lnwgt, we should find the lnhgt coeff will be twice the magnitude of the lnwgt coeff and opposite in sign.

So let's try that.

```
. bcuse bodyfat

. gen lnBrozek=ln(Brozek)
(1 missing value generated)

. gen lnhgt=ln(hgt_m)
. gen lnwgt=ln(wgt_kg)

. reg lnBrozek lnhgt lnwgt

      Source |       SS       df       MS              Number of obs =     251
-------------+------------------------------           F(  2,   248) =   79.37
       Model |  22.6698706     2  11.3349353           Prob > F      =  0.0000
    Residual |  35.4177456   248   .14281349           R-squared     =  0.3903
-------------+------------------------------           Adj R-squared =  0.3854
       Total |  58.0876161   250  .232350465           Root MSE      =  .37791

------------------------------------------------------------------------------
    lnBrozek |      Coef.   Std. Err.      t    P>|t|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
       lnhgt |  -1.512178   .3718011    -4.07   0.000    -2.244469   -.7798881
       lnwgt |   1.978279   .1576453    12.55   0.000     1.667784    2.288773
       _cons |  -4.958816   .6739082    -7.36   0.000    -6.286129   -3.631503
------------------------------------------------------------------------------
```

Oops… while they do have opposite signs, the lnhgt coefficient is definitely not twice the size of the lnwgt coefficient… in fact, the lnhgt coefficient is actually smaller in magnitude!

So much for the BMI! … or maybe not!

```
. bcuse bodyfat
```

Looking for data issues I… Compare reported Brozek with calculation (Brozek2)

```
. list Case Brozek Brozek2 if abs( Brozek- Brozek2)>1

     +---------------------------+
     | Case   Brozek     Brozek2 |
     |---------------------------|
 33. |   33     13.4    12.14574 |
 48. |   48      6.4    14.30445 |
 76. |   76     18.3    14.26428 |
 96. |   96     17.3    1.594741 |
182. |  182        0   -2.079881 |
     +---------------------------+
```

Looking for data issues II… compare reported heights and weights

```
. scatter hgt wgt
.
. list if hgt < 40

-----------------------------------------------------------------------------------+
Case | Brozek | Siri | density | age | wgt |  hgt |  BMI | fatfre~t | neck | chest |   abd |
  42 |   31.7 | 32.9 |   1.025 |  44 | 205 | 29.5 | 29.9 |    140.1 | 36.6 |   106 | 104.3 |
--------------+------+---------------------------------------------------------------------|
  hip | thigh | knee | ankle | biceps | forearm  | wrist  |   wgt_kg  | hgt_m  |      BMI2 |
115.5 |  70.6 | 42.5 |  23.7 |   33.6 |    28.7  |  17.4  | 92.98636  | .7493  | 165.6181  |
-------------------------------------------------------------------------------------------|
                Brozek2                  |                     Siri2                        |
               31.65366                  |                  32.92683                        |
-------------------------------------------------------------------------------------------+
```

Yes, that's right…. 29.5 inches tall and weighs 205 pounds! *No doubt!*

What happens when we drop those curious observations?

```
. drop if Case == 33 | Case == 48 | Case == 76 | Case == 96 | Case == 182 |
Case == 42
(6 observations deleted)

. reg lnBrozek lnhgt lnwgt

      Source |       SS       df       MS              Number of obs =     246
-------------+------------------------------           F(  2,   243) =  100.04
       Model |  25.5898191     2  12.7949096           Prob > F      =  0.0000
    Residual |  31.0781474   243  .127893611           R-squared     =  0.4516
-------------+------------------------------           Adj R-squared =  0.4471
       Total |  56.6679666   245  .231297823           Root MSE      =  .35762

------------------------------------------------------------------------------
    lnBrozek |      Coef.   Std. Err.      t    P>|t|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
       lnhgt |  -4.839145   .7198636    -6.72   0.000    -6.257114   -3.421176
       lnwgt |   2.404657   .1701085    14.14   0.000     2.069582    2.739733
       _cons |  -4.895534   .6447506    -7.59   0.000    -6.165547   -3.625521
------------------------------------------------------------------------------
```

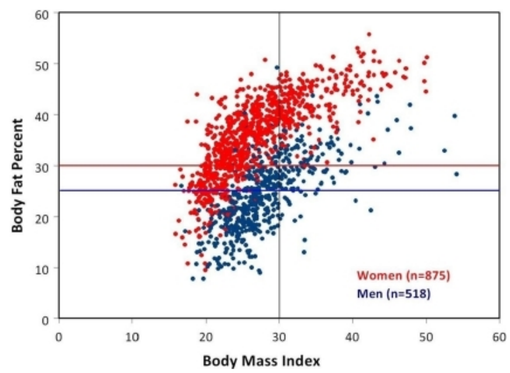
Bingo! 2-1 ratio! … we can use the F test to test for the BMI restriction.

Assume that the restriction in true (the Null hypothesis is that the magnitude ratio is 2-1).

```
. test lnhgt = - 2*lnwgt

 ( 1)  lnhgt + 2*lnwgt = 0

       F(  1,    243) =    0.00
            Prob > F =    0.9617
```



We fail to reject the restriction, since the Prob > F is so large … we typically reject when that probability is less than .1, .05, .01 etc. This probability is similar to the p-value… it's the probability that we will falsely reject the Null hypothesis.

So let's impose the restriction and see what we get:

```
. gen lnBMI2=ln(BMI2)
. reg lnBrozek lnBMI2

      Source |       SS       df       MS              Number of obs =     246
-------------+------------------------------           F(  1,    244) =  200.91
       Model |  25.5895283     1  25.5895283           Prob > F      =  0.0000
    Residual |  31.0784383   244  .127370649           R-squared     =  0.4516
-------------+------------------------------           Adj R-squared =  0.4493
       Total |  56.6679666   245  .231297823           Root MSE      =  .35689

------------------------------------------------------------------------------
    lnBrozek |      Coef.   Std. Err.      t    P>|t|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
      lnBMI2 |   2.404336   .1696284    14.17   0.000     2.070213    2.738459
       _cons |  -4.911768   .5479379    -8.96   0.000     -5.99106   -3.832476
------------------------------------------------------------------------------
```

Reported R-squared didn't change, but SSRs increased (barely) from 31.078147 to 31.078438. (Recall that SSRs must increase (or more precisely, cannot decrease) if you impose a restriction.)

The F-test basically looks at how much the SSRs increase when you impose a restriction (or restrictions):

- If SSRs increase by a small amount, then little explanatory power is lost when the restriction is imposed, and the F test fails to reject the Null Hypothesis that the restriction is OK.
- If the SSRs increase by a lot, then the restriction is leading to a much poorer model (as measured by R squared) and the restriction is rejected.

Here's the calculation of the F Statistic for the F Test:

$$F = \frac{(SSR_R - SSR_{UR})/SSR_{UR}}{q/(n-k-1)} = \frac{(31.078438 - 31.078147)/31.078147}{1/243} = 0.00227$$

F tests the Easy Way:

Incorporate the restrictions in the model and test for differences

```
. bcuse bodyfat

. gen lnbfat=ln(Brozek)
. gen lnwgt=ln(wgt)
. gen lnhgt=ln(hgt)
.
. reg lnbfat lnwgt lnhgt

      Source |       SS       df       MS              Number of obs =     251
-------------+------------------------------           F(  2,   248) =   79.37
       Model |  22.6698663     2  11.3349331           Prob > F      =  0.0000
    Residual |  35.4177498   248  .142813507           R-squared     =  0.3903
-------------+------------------------------           Adj R-squared =  0.3854
       Total |  58.0876161   250  .232350465           Root MSE      =  .37791

------------------------------------------------------------------------------
      lnbfat |      Coef.   Std. Err.      t    P>|t|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
       lnwgt |   1.978279   .1576454    12.55   0.000     1.667784    2.288773
       lnhgt |  -1.512178   .3718011    -4.07   0.000    -2.244468   -.7798875
       _cons |  -.9685202   1.597071    -0.61   0.545    -4.114073    2.177033
------------------------------------------------------------------------------

. test lnhgt = -2*lnwgt

 ( 1)  2*lnwgt + lnhgt = 0

       F(  1,   248) =   32.82
            Prob > F =    0.0000

. di sqrt(32.82)
5.7288742
.
. gen lnbmi = lnwgt - 2*lnhgt
```

lnbmi incorporates the constraint; the coefficient on lnhgt in the following regression will pick up differences… and we can use the t- test to test the null hypothesis that the differences are zero.

```
. reg lnbfat lnbmi lnhgt

      Source |       SS       df       MS              Number of obs =     251
-------------+------------------------------           F(  2,   248) =   79.37
       Model |  22.6698663     2  11.3349331           Prob > F      =  0.0000
    Residual |  35.4177498   248  .142813507           R-squared     =  0.3903
-------------+------------------------------           Adj R-squared =  0.3854
       Total |  58.0876161   250  .232350465           Root MSE      =  .37791

------------------------------------------------------------------------------
      lnbfat |      Coef.   Std. Err.      t    P>|t|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
       lnbmi |   1.978279   .1576454    12.55   0.000     1.667784    2.288773
       lnhgt |    2.44438   .4266716     5.73   0.000     1.604018    3.284742
       _cons |  -.9685202   1.597071    -0.61   0.545    -4.114073    2.177033
------------------------------------------------------------------------------
```

Note that everything except for the lnhgt coeff and Std. Err. Is the same in the two regressions… and the t-stat is the sqrt of the F stat above. So the two tests are basically the same.

Now, let lnwgt pick up differences… and we can use the t- test to test the null hypothesis that the differences are zero.

```
. reg lnbfat lnbmi lnwgt

      Source |       SS       df       MS              Number of obs =     251
-------------+------------------------------           F(  2,   248) =   79.37
       Model |  22.6698663     2  11.3349331           Prob > F      =  0.0000
    Residual |  35.4177498   248  .142813507           R-squared     =  0.3903
-------------+------------------------------           Adj R-squared =  0.3854
       Total |  58.0876161   250  .232350465           Root MSE      =  .37791

------------------------------------------------------------------------------
      lnbfat |      Coef.   Std. Err.      t    P>|t|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
       lnbmi |    .756089   .1859006     4.07   0.000     .3899438    1.122234
       lnwgt |    1.22219   .2133358     5.73   0.000      .802009    1.642371
       _cons |  -.9685202   1.597071    -0.61   0.545    -4.114073    2.177033
------------------------------------------------------------------------------.
```

Surprise! Same test and virtually the same results (only coeffs and Std Errs differ!

```
. esttab , r2 ar2 scalar(df_r F rmse) plain

                     est1         est2         est3
                      b/t          b/t          b/t
lnwgt            1.978279                   1.22219
                 12.54892                  5.728949

lnhgt           -1.512178      2.44438
                -4.067169     5.728949

lnbmi                         1.978279      .756089
                              12.54892     4.067169

_cons           -.9685202    -.9685202    -.9685202
                -.6064351    -.6064351    -.6064351

N                     251          251          251
R-sq             .3902702     .3902702     .3902702
adj. R-sq         .385353      .385353      .385353
df_r                  248          248          248
F                79.36878     79.36878     79.36878
rmse             .3779067     .3779067     .3779067
```

Essentially the same test stats: The F-stat is the square of the t-stats!

And the SRF's from the three models are identical!

$$SRF_1 : \hat{y} = -.97 + 1.98\, lnwgt - 1.51\, lnhgt$$

$$SRF_2 : \hat{y} = -.97 + 1.98\, lnbmi + 2.44\, lnhgt = -.97 + 1.98\,(lnwgt - 2\, lnhgt) + 2.44\, lnhgt$$

$$= -.97 + 1.98\, lnwgt - 1.51\, lnhgt$$

$$SRF_3 : \hat{y} = -.97 + .76\, lnbmi + 1.22\, lnwgt = -.97 + .76\,(lnwgt - 2\, lnhgt) + 1.22\, lnwgt$$

$$= -.97 + 1.98\, lnwgt - 1.51\, lnhgt$$

Now drop those six problematic observations:

```
.
. drop if Case == 33 | Case == 48 | Case == 76 | Case == 96 | Case == 182 |
Case == 42
. reg lnbfat lnwgt lnhgt

      Source |       SS       df       MS              Number of obs =     246
-------------+------------------------------           F(  2,   243) =  100.04
       Model |  25.5898178     2  12.7949089           Prob > F      =  0.0000
    Residual |  31.0781487   243  .127893616           R-squared     =  0.4516
-------------+------------------------------           Adj R-squared =  0.4471
       Total |  56.6679666   245  .231297823           Root MSE      =  .35762

------------------------------------------------------------------------------
      lnbfat |      Coef.   Std. Err.      t    P>|t|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
       lnwgt |   2.404658   .1701086    14.14   0.000     2.069582    2.739733
       lnhgt |  -4.839147   .7198639    -6.72   0.000    -6.257116   -3.421178
       _cons |   10.97766   2.722847     4.03   0.000     5.614268    16.34106
------------------------------------------------------------------------------

. test lnhgt = -2*lnwgt

( 1)  2*lnwgt + lnhgt = 0

       F(  1,   243) =    0.00
            Prob > F =    0.9617
.
. qui: reg lnbfat lnbmi lnhgt
. qui: reg lnbfat lnbmi lnwgt
```

	(1) lnbfat	(2) lnbfat	(3) lnbfat
lnwgt	2.405***		-0.0149
	(0.000)		**(0.962)**
lnhgt	-4.839***	-0.0298	
	(0.000)	**(0.962)**	
lnbmi		2.405***	2.420***
		(0.000)	(0.000)
_cons	10.98***	10.98***	10.98***
	(0.000)	(0.000)	(0.000)
N	246	246	246
R-sq	0.452	0.452	0.452
adj. R-sq	0.447	0.447	0.447
df_r	243	243	243
F	100.0	100.0	100.0
rmse	0.358	0.358	0.358

Same p-values for the F and t tests! (They were all .0000 above, so we didn't know for sure… but the p-values will all agree.

So Run the F test… or incorporate the constraint and allow for differences… and use the t test to test the Null Hypothesis of no difference.

F tests continued: *More about bathwater and babies…*

```
. bcuse sovdebt02

. gen badbad = corrupt
. replace badbad = 6.9 in 16

. reg NSRate corrupt badbad eurozone gdp inflation deficit_gdp debt_gdp

      Source |       SS       df       MS              Number of obs =     108
-------------+------------------------------           F(  7,   100) =   58.33
       Model |  269.066594     7  38.4380849           Prob > F      =  0.0000
    Residual |  65.8966464   100  .658966464           R-squared     =  0.8033
-------------+------------------------------           Adj R-squared =  0.7895
       Total |  334.963241   107  3.13049758           Root MSE      =  .81177

------------------------------------------------------------------------------
      NSRate |      Coef.   Std. Err.      t    P>|t|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
     corrupt |  -10.78155   8.422346    -1.28   0.203    -27.49125    5.928139
      badbad |   11.39931     8.4228     1.35   0.179     -5.31129     28.1099
    eurozone |   .6960289   .2478395     2.81   0.006     .2043223    1.187735
         gdp |    .000254   .0000513     4.95   0.000     .0001522    .0003558
   inflation |  -.0150957   .0210322    -0.72   0.475    -.0568231    .0266317
 deficit_gdp |   -.057536   .0155416    -3.70   0.000    -.0883701   -.0267018
    debt_gdp |  -.0109844   .0027306    -4.02   0.000    -.0164018    -.005567
       _cons |    3.65045   .2644217    13.81   0.000     3.125845    4.175055
------------------------------------------------------------------------------
```

Neither corrupt nor badbad is statistically significant at the 15% level, let alone 5%:

```
. test corrupt
 ( 1)  corrupt = 0

       F(  1,   100) =    1.64
            Prob > F =    0.2035

. di sqrt(1.64)
1.2806248

. test badbad
 ( 1)  badbad = 0

       F(  1,   100) =    1.83
            Prob > F =    0.1790

. di sqrt(1.83)
1.3527749
```

… but don't toss them both out!

```
. test corrupt badbad

 ( 1)  corrupt = 0
 ( 2)  badbad = 0

       F(  2,   100) =  106.58
            Prob > F =    0.0000
```

F Statistics, Adjusted R^2 and Mean Squared Error

F statistics, SSRs and dofs (degrees of freedom):

Under MLR.1-MLR.6, the following F statistic will have a $F(q, n-k-1)$ distribution:

$F \equiv \dfrac{\left(SSR_R - SSR_{UR}\right)/q}{SSR_{UR}/(n-k-1)}$, where q is the number of restrictions, and $n-k-1$ is the number of degrees of freedom in the unrestricted model.

Another interpretation of the F statistic: $F \equiv \dfrac{\left(SSR_R - SSR_{UR}\right)/SSR_{UR}}{q/(n-k-1)} = \dfrac{\%\Delta SSR}{\%\Delta dofs}$

so the F statistic is an elasticity, and tells you the %change in SSRs for a given %change in degrees of freedom. If $\%\Delta SSR \gg \%\Delta dofs$, then this elasticity is large, the F statistic is large, the associated probability level is small, and the restrictions are rejected. If the elasticity is small, then the restrictions had a much smaller impacts on SSRs , the associated F probability is large, and the restrictions are not rejected.

Adjusted R^2: $\bar{R}^2 = 1-(1-R^2)\left[\dfrac{n-1}{n-k-1}\right] = 1-\dfrac{SSR}{SST}\dfrac{(n-1)}{(n-k-1)}$. Since $\dfrac{(n-1)}{(n-k-1)} > 1$, $\bar{R}^2 < R^2$, but as n gets very large or k approaches zero, $\bar{R}^2 \to R^2$. Since $R^2 = 1-\dfrac{SSR}{SST}$, $\bar{R}^2 = 1-\dfrac{SSR/(n-k-1)}{SST/(n-1)} = 1-\dfrac{MSE}{S_{YY}}$. Thus $\bar{R}^2$ increases iff MSE decreases.

Consider two models: UR – unrestricted and R- restricted (q = #restrictions). Then the change in adjusted R^2 in going from one model to the other is defined by:

$$\bar{R}^2_{UR} - \bar{R}^2_R = \frac{1}{S_{YY}}\left(MSE_R - MSE_{UR}\right) = \frac{1}{S_{YY}}\left(\frac{SSR_R}{(n-k-1)+q} - \frac{SSR_{UR}}{n-k-1}\right).$$

But this is:

$$= \frac{qSSR_{UR}}{S_{YY}((n-k-1)+q)(n-k-1)}\left(\frac{(n-k-1)\left[SSR_R - SSR_{UR}\right]}{qSSR_{UR}} - 1\right)$$

$$= \frac{qSSR_{UR}}{S_{YY}((n-k-1)+q)(n-k-1)}\left[F-1\right] = \frac{q}{(n-k-1)+q}\frac{MSE_{UR}}{S_{YY}}\left[F-1\right]$$

$$= \frac{q\left[1-\bar{R}^2_{UR}\right]}{(n-k-1)+q}\left[F-1\right]$$

Interpretation: The sign of this expression will depend on whether the F statistic is greater or less than 1: adjusted R^2 increases iff the F statistic exceeds 1.

Recall that for a single restriction, the F statistic is the square of the t stat, and so adjusted R^2 increases iff the magnitude of the t stat of the added RHS variable exceeds 1.