MLR Inference II: *Convergence II*

- *Convergence II: t Stats and Incremental Goodness-of-Fit*
- *… an example: bodyfat I*
- *… another example: bodyfat II*
- *… and WhatsNew*
- *Comparing MLR Models II: t stats and adjusted R2*
- *… more about t stats and adjusted R2*

Convergence II: t Stats and Incremental Goodness-of-Fit

- ***Convergence I:*** In SLR Inference, you saw the convergence of inference and assessment metrics, driven by relationship between t statistics and the R^2 measure of goodness of fit, as well as SSE/SSR:

$$t_{\hat{\beta}_1}^2 = (n-2)\frac{R^2}{1-R^2} = (n-2)\frac{SSE}{SSR}.$$

- ***Convergence II:*** We have similar results in MLR Inference, where the precision of estimation is jointly driven by the degrees of freedom (*dofs*) and now the ***marginal*** (***incremental)*** impact that each RHS variable has on R^2 or SSE's:

$$t_{\hat{\beta}_x}^2 = dofs\frac{\Delta R_x^2}{1-R^2} = dofs\frac{\Delta SSE_x}{SSR}$$

where $dofs = n-k-1$, and ΔR_x^2 (ΔSSE_x) is the increase in R^2 (SSE) when *x* is the *last* variable added to the model (R^2 and *SSR* are for the *Full Model*)

Precision in estimation is driven by:
- **the degrees of freedom, n-k-1, and**
- **incremental R-sq (SSE)**

- The SLR and MLR formulas are in fact consistent: R^2 in an SLR model is the same as ΔR_x^2 when going from no RHS variables (other than the constant term) to the SLR model.

Convergence II – An example: *bodyfat I*

	Dropping One RHS Variable (1) brozek	(2) brozek	(3) brozek	Full Model (4) brozek
wgt	0.187***	-0.136***	dropped	-0.120***
	(14.48)	(-7.08)		(-5.41)
hgt	-0.650***	dropped	-0.342***	-0.118
	(-6.29)		(-4.55)	(-1.43)
abd	dropped	0.915***	0.595***	0.880***
		(17.42)	(23.30)	(15.19)
_cons	31.16***	-41.35***	-12.12*	-32.66***
	(4.51)	(-17.14)	(-2.17)	(-5.01)
N	252	252	252	252
R-sq	0.4614	0.7187	0.6881	0.7210
mss (SSE)	6,958.1	10,837.7	10,375.8	1,0872.6
rss (SSR)	8,121.0	4,241.3	4,703.2	4,206.5

t statistics in parentheses
* p<0.05, ** p<0.01, *** p<0.001

Looking at *abd* as the *last* variable, so comparing Models (1) and (4):

$$t^2_{\hat{\beta}_{abd}} = (dofs)\frac{\Delta R^2_{abd}}{1-R^2} = 248\frac{.7210-.4614}{1-.721} = 248\frac{.2596}{1-.721} = (15.19)^2$$

$$t^2_{\hat{\beta}_{abd}} = dofs\frac{\Delta SSE_{abd}}{SSR} = 248\frac{10,872.6-6,958.1}{4,206.5} = 248\frac{3,914.5}{4,206.5} = (15.19)^2$$

And so as advertised, $t^2_{\hat{\beta}_x} = dofs\frac{\Delta R^2_x}{1-R^2} = dofs\frac{\Delta SSE_x}{SSR}$.

Convergence II – Another example: *bodyfat II*

	Dropping One RHS Variable			Full Model
	(1)	(2)	(3)	(4)
	brozek	brozek	brozek	brozek
wgt	0.187***	-0.136***	dropped	-0.120***
	(14.48)	(-7.08)		(-5.41)
hgt	-0.650***	dropped	-0.342***	-0.118
	(-6.29)		(-4.55)	(-1.43)
abd	dropped	0.915***	0.595***	0.880***
		(17.42)	(23.30)	(15.19)
_cons	31.16***	-41.35***	-12.12*	-32.66***
	(4.51)	(-17.14)	(-2.17)	(-5.01)
N	252	252	252	252
R-sq	0.4614	0.7187	0.6881	0.7210
mss (SSE)	6,958.1	10,837.7	10,375.8	1,0872.6
rss (SSR)	8,121.0	4,241.3	4,703.2	4,206.5

t statistics in parentheses
* p<0.05, ** p<0.01, *** p<0.001

- Looking at the t stats in an MLR model: the square of the t stats, $t^2_{\hat{\beta}_x}$, are directly proportional to each variable's marginal/incremental contribution to R^2 (SSE's):
  - $\frac{t^2_{\hat{\beta}_x}}{t^2_{\hat{\beta}_z}} = \frac{\Delta R^2_x}{\Delta R^2_z} = \frac{\Delta SSE_x}{\Delta SSE_z}$, for any two RHS variables *x* and *z*.
- Comparing *wgt* and *abd*:
  - Since $\Delta R^2_{abd} = .2596$ and $\Delta R^2_{wgt} = .7210 - .6881 = .0329$, we have:

$$\frac{\Delta R^2_{abd}}{\Delta R^2_{wgt}} = \frac{.2596}{.0329} = 7.88 = \frac{t^2_{\hat{\beta}_{abd}}}{t^2_{\hat{\beta}_{wgt}}} = \left(\frac{15.19}{5.41}\right)^2$$

  - And since $\Delta SSE_{abd} = 3{,}914.5$ and $\Delta SSE_{wgt} = 10{,}872.6 - 10{,}375.8 = 496.8$, we have:

$$\frac{\Delta SSE_{abd}}{\Delta SSE_{wgt}} = \frac{3{,}914.5}{496.8} = 7.88 = \frac{t^2_{\hat{\beta}_{abd}}}{t^2_{\hat{\beta}_{wgt}}} = \left(\frac{15.19}{5.41}\right)^2$$

- So variables with larger t stats have greater marginal impacts on R^2 and SSE … and *vice-versa*. ***Who saw this coming?***

Convergence II: *… and WhatsNew*$_x$

- ΔR_x^2 and ΔSSE_x can be found in the regression of *y* on *WhatsNew* about *x*, where ΔR_x^2 is the R^2 in the *WhatsNew* SLR regression, and ΔSSE_x is the SSE in that model.
- Example: Look at the previous example, and focus again on the *abd* variable: $\Delta R_{abd}^2 = .2596$ and $\Delta SSE_{abd} = 3{,}914.5$.
- And regress *brozek* on *WhatsNew* about *abd*:

```
. reg abd wgt hgt
. predict whatsnew, resid

. reg brozek whatsnew

      Source |       SS           df       MS      Number of obs   =       252
-------------+----------------------------------   F(1, 250)       =     87.65
       Model |   3914.4903         1   3914.4903   Prob > F        =    0.0000
    Residual |  11164.5263       250  44.6581053   R-squared       =    0.2596
-------------+----------------------------------   Adj R-squared   =    0.2566
       Total |  15079.0166       251  60.0757635   Root MSE        =    6.6827

------------------------------------------------------------------------------
      brozek |      Coef.   Std. Err.      t    P>|t|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
    whatsnew |    .879846   .0939765     9.36   0.000     .6947594    1.064932
       _cons |   18.93849   .4209688    44.99   0.000     18.10939    19.76759
------------------------------------------------------------------------------
```

Comparing MLR Models II: *t stats and adjusted R^2*

- Changes in adjusted R-squared ($\bar{R}^2$) are directly tied to whether or not the t stats of added variables are larger than 1 in magnitude, or not.
- $\bar{R}^2$ will always increase (decrease) when variables with t stats larger (smaller) than one in magnitude are added to the MLR model… and *vice-versa* when dropping variables from a model.
- With the addition of a RHS variable: $\bar{R}^2 \begin{bmatrix} increases \\ stays\ the\ same \\ decreases \end{bmatrix} \ when \ |t| \begin{bmatrix} > \\ = \\ < \end{bmatrix} 1$

… More about *t stats and adjusted* R^2

	(1)	(2)	(3)	(4)
	Brozek	Brozek	Brozek	Brozek
hgt	-0.650***	-0.118	-0.131	-0.138
	(-6.29)	(-1.43)	(-1.51)	(-1.55)
wgt	0.187***	-0.120***	-0.108**	-0.100*
	(14.48)	(-5.41)	(-3.18)	(-2.52)
abd		0.880***	0.883***	0.898***
		(15.19)	(15.13)	(12.62)
hip			-0.0564	-0.0723
			(-0.49)	(-0.58)
chest				-0.0348
				(-0.38)
_cons	31.16***	-32.66***	-28.64**	-25.86*
	(4.51)	(-5.01)	(-2.71)	(-2.01)
N	252	252	252	252
R-sq	0.461	0.721	0.721	0.721
adj. R-sq	0.457	0.718	0.717	0.716
rmse	5.711	4.118	4.125	4.132

t statistics in parentheses
* p<0.05, ** p<0.01, *** p<0.001

Notice that in going from Model (1) to (2), $\bar{R}^2$ increased and the added (or *last* or *incremental*) variable (*abd*) had a t stat of 15.19, well above one in magnitude. And in going from (2) to (3), and (3) to (4), $\bar{R}^2$ decreased in both cases, and the t stats of the added variables were both less than one in magnitude.

This is useful if your goal is to maximize $\bar{R}^2$. It's never a great idea to just worry about adjusted R-squared, but you wouldn't be the first analyst to do so.

onwards to Heteroskedasticity