Hypothesis Tests and Confidence Intervals in Multiple Regressions

SW Chapter 7

Overview of tests

Extended example: Angrist, Lang, and Oreopoulos (2009)

- Testing one restriction, multiple coefficients
- Testing multiple restrictions
- Testing overall significance of a regression

Learning objectives

- Construct and interpret tests of joint hypotheses
- Construct and test hypothesis test involving one restriction and multiple coefficients

Overview of tests

Three types of tests

1. Hypothesis tests with one restriction, one coefficient
   - Example: $H_0 : \beta_j = \beta_{j,0}$ vs. $H_a : \beta_j \neq \beta_{j,0}$
2. Hypothesis tests with one restriction, multiple coefficients
   - General: $H_0 : \beta_j = \beta_m$
   - Example: $H_0 : \beta_1 = 0$
3. Hypothesis tests involving a multiple tests at once (joint hypothesis tests)
   - General: $H_0 : \beta_j = \beta_{j,0}, \beta_m = \beta_{m,0}, \ldots$
   - Example: $H_0 : \beta_1 = \beta_2 = \beta_3 = 0$
   - Special case - test of *all* regressors

1. One restriction, one coefficient, $\beta_j = \beta_{j,0}$

1. Select your significance level ($\alpha = 0.01$)
2. State your null hypothesis:

$$H_0 : \beta_j = \beta_{j,0}$$
$$H_a : \beta_j \neq \beta_{j,0}$$

3. Compute the t-statistic:

$$t = \frac{\widehat{\beta}_j - \beta_{j,0}}{SE(\widehat{\beta}_j)}$$

4. Compare the t-statistic to your critical value (**2.58**) and reject null if $|t| > c$
5. (Optional) Construct your confidence interval:

$$(\widehat{\beta}_j - \mathbf{2.58} SE(\widehat{\beta}_j), \widehat{\beta}_j + \mathbf{2.58} SE(\widehat{\beta}_j))$$

2. One restriction, two coefficients, $\beta_j = \beta_m$

1. Select your significance level ($\alpha = 0.01$)
2. State your null hypothesis: $H_0 : \beta_1 = \beta_2$ vs $H_a : \beta_1 \neq \beta_2$
3. Transform your regression:

$$\begin{aligned}
y_i &= \beta_0 + \beta_1 x_{1i} + \beta_2 x_{2i} + u_i \\
y_i &= \beta_0 + \beta_1 x_{1i} + \beta_2 x_{1i} - \beta_2 x_{1i} + \beta_2 x_{2i} + u_i \\
y_i &= \beta_0 + (\beta_1 - \beta_2) x_{1i} + \beta_2 (x_{1i} + x_{2i}) + u_i \\
y_i &= \beta_0 + \gamma_1 x_{1i} + \beta_2 w_i + u_i
\end{aligned}$$

   and instead test $H_0 : \gamma_1 = 0$ vs. $H_1 : \gamma_1 \neq 0$
4. Repeat remaining steps

3. Multiple restriction null restrictions, under homoskedasticity

There are lots of variants, but this is the one we will compute by hand

1. Select your significance level ($\alpha = 0.01$)
2. State your null hypothesis: $H_0 : \beta_1 = \beta_2 = \beta_3$ vs any part not true
3. Estimate model with variables you are testing (unrestricted) and without (restricted)
4. Calculate F-statistic

$$F = \frac{(SSR_r - SSR_{ur})/q}{SSR_{ur}/(n - k_{ur} - 1)} = \frac{(R^2_{ur} - R^2_r)/q}{(1 - R^2_{ur})/(n - k_{ur} - 1)} \sim F_{q,n-k_{ur}-1}$$

5. Compare the F-statistic to your critical value from a $F_{q,n-k_{ur}-1}$ critical value and reject null if $F > c$ (usually, this will be from $F_{q,\infty}$ distribution)

Extended example: Angrist, Lang, and Oreopoulos (2009)

Angrist, Lang, and Oreopoulos (2009)

Incentives And Services For College Achievement: Evidence From A Randomized Trial

By JOSHUA ANGRIST, DANIEL LANG, AND PHILIP OREOPOULOS*

This paper reports on an experimental evaluation of strategies designed to improve academic performance among college freshmen. One treatment group was offered academic support services. Another was offered financial incentives for good grades. A third group combined both interventions. Service use was highest for women and for subjects in the combined group. The combined treatment also raised the grades and improved the academic standing of women. These differentials persisted through the end of second year, though incentives were given in the first year only. This suggests study skills among some treated women increased. In contrast, the program had no effect on men. (*JEL* I21, I28)

We're going to think about predictors of year 1 college GPA.

A satellite campus of a “large, Canadian university”

```
. regress GPA_year1  female mt_french mt_other hs_q2 hs_q3

      Source |       SS           df       MS      Number of obs   =     1,374
-------------+----------------------------------   F(5, 1368)      =     57.06
       Model |  186.639466         5  37.3278932   Prob > F        =    0.0000
    Residual |  894.929455     1,368  .654188198   R-squared       =    0.1726
-------------+----------------------------------   Adj R-squared   =    0.1695
       Total |  1081.56892     1,373  .787741385   Root MSE        =    .80882

------------------------------------------------------------------------------
   GPA_year1 |      Coef.   Std. Err.      t    P>|t|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
      female |  -.1296654   .0442666    -2.93   0.003    -.2165032   -.0428276
   mt_french |  -.4658419   .3079241    -1.51   0.131    -1.069896    .1382127
    mt_other |  -.0755215     .04796    -1.57   0.116    -.1696047    .0185616
       hs_q2 |   .3743258   .0545463     6.86   0.000     .2673224    .4813292
       hs_q3 |   .8857912   .0532672    16.63   0.000      .781297    .9902854
       _cons |   1.484259   .0467597    31.74   0.000     1.392531    1.575987
------------------------------------------------------------------------------
```

with controls for gender, mother tongue, and HS quartile (no top quartile!)

Avoid dummy variable trap!

```
. regress GPA_year1  female mt_english mt_french hs_q2 hs_q3

      Source |       SS           df       MS      Number of obs   =     1,374
-------------+----------------------------------   F(5, 1368)      =     57.06
       Model |  186.639466         5  37.3278932   Prob > F        =    0.0000
    Residual |  894.929455     1,368  .654188198   R-squared       =    0.1726
-------------+----------------------------------   Adj R-squared   =    0.1695
       Total |  1081.56892     1,373  .787741385   Root MSE        =    .80882

------------------------------------------------------------------------------
   GPA_year1 |      Coef.   Std. Err.      t    P>|t|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
      female |  -.1296654   .0442666    -2.93   0.003    -.2165032   -.0428276
  mt_english |   .0755215     .04796     1.57   0.116    -.0185616    .1696047
   mt_french |  -.3903204   .3093578    -1.26   0.207    -.9971875    .2165467
       hs_q2 |   .3743258   .0545463     6.86   0.000     .2673224    .4813292
       hs_q3 |   .8857912   .0532672    16.63   0.000      .781297    .9902854
       _cons |   1.408737   .0575791    24.47   0.000     1.295785     1.52169
------------------------------------------------------------------------------
```

Choice of excluded group only affects interpretation

English mother tongue excluded:

$$\widehat{GPAyear1} = 1.484 - 0.130 female - 0.466 mt_{french} - 0.076 mt_{other} + 0.374 hs_{q2} + 0.886 hs_{q3}$$

"Other" mother tongue excluded:

$$\widehat{GPAyear1} = 1.409 - 0.130 female - 0.076 mt_{english} - 0.390 mt_{french} + 0.374 hs_{q2} + 0.886 hs_{q3}$$

Choice of excluded group only affects interpretation

For simplicity, assume $female = 0$, $hs_{q2} = 0$, and $hs_{q3} = 0$

English mother tongue excluded:

$$\widehat{GPAyear1} = 1.484 - 0.130 female - 0.466 mt_{french} - 0.0755 mt_{other} + 0.374 hs_{q2} + 0.886 hs_{q3}$$

$$E[GPAyear1|mt_{english} = 1] = 1.484$$
$$E[GPAyear1|mt_{french} = 1] = 1.484 - 0.466 = 1.018$$
$$E[GPAyear1|mt_{other} = 1] = 1.4843 - 0.0755 = 1.409$$

Choice of excluded group only affects interpretation

For simplicity, assume $female = 0$, $hs_{q2} = 0$, and $hs_{q3} = 0$

"Other" mother tongue excluded:

$$\widehat{GPAyear1} = 1.409 - 0.130 female + 0.0755 mt_{english} - 0.390 mt_{french} + 0.374 hs_{q2} + 0.886 hs_{q3}$$

$$E[GPAyear1 | mt_{english} = 1] = 1.4087 + 0.0755 = 1.484$$

$$E[GPAyear1 | mt_{french} = 1] = 1.4087 - 0.3903 = 1.018$$

$$E[GPAyear1 | mt_{other} = 1] = 1.4087 = 1.409$$

Questions we could want to answer

```
. regress GPA_year1  female mt_french mt_other hs_q2 hs_q3

      Source |       SS           df       MS      Number of obs   =     1,374
-------------+----------------------------------   F(5, 1368)      =     57.06
       Model |  186.639466         5  37.3278932   Prob > F        =    0.0000
    Residual |  894.929455     1,368  .654188198   R-squared       =    0.1726
-------------+----------------------------------   Adj R-squared   =    0.1695
       Total |  1081.56892     1,373  .787741385   Root MSE        =    .80882

------------------------------------------------------------------------------
   GPA_year1 |      Coef.   Std. Err.      t    P>|t|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
      female |  -.1296654   .0442666    -2.93   0.003    -.2165032   -.0428276
   mt_french |  -.4658419   .3079241    -1.51   0.131    -1.069896    .1382127
    mt_other |  -.0755215     .04796    -1.57   0.116    -.1696047    .0185616
       hs_q2 |   .3743258   .0545463     6.86   0.000     .2673224    .4813292
       hs_q3 |   .8857912   .0532672    16.63   0.000      .781297    .9902854
       _cons |   1.484259   .0467597    31.74   0.000     1.392531    1.575987
------------------------------------------------------------------------------
```

- Do French speakers have same GPA as English speakers?
- Do non-English/French speakers have same GPA as English speakers?
- Do French speakers have same GPA as non-French/English speakers?
- Are there differences in GPA by mother tongue?

One restriction, multiple coefficients

Do French speakers have same GPA as non-French/English speakers?

Population model:

$$GPAyear1_i = \beta_0 + \beta_1 mt_{french} + \beta_2 mt_{other} + \beta_3 hs_{q2} + \beta_4 hs_{q3} + u$$

Test whether $H_0 : \beta_1 = \beta_2$ vs $H_a : \beta_1 \neq \beta_2$

Equivalent to asking whether $\beta_1 - \beta_2 = 0$

One restriction, multiple coefficients

Do French speakers have same GPA as non-French/English speakers?

```
. test mt_french = mt_other

 ( 1)  mt_french - mt_other = 0

       F(  1,  1368) =    1.59
            Prob > F =    0.2073
```

Multiple restrictions (joint test)

Are there differences in GPA by mother tongue?

Population model:

$$GPAyear1_i = \beta_0 + \beta_1 mt_{french} + \beta_2 mt_{other} + \beta_3 hs_{q2} + \beta_4 hs_{q3} + u$$

Test $H_0 : \beta_1 = \beta_2 = 0$ vs. H_a any one equality is not true (either $\beta_1 \neq 0$ or $\beta_2 \neq 0$)

Equivalent to asking whether the coefficients are jointly significant

Testing multiple restrictions: key terms

- **Exclusion restriction** Test of whether certain covariates can be excluded from population model
- **Restricted model**: The model with *fewer* covariates
  - “Restricted“ because we are imposing that coefficients of missing covariates are zero
- **Unrestricted model** The model with *more* covariates
  - “Unrestricted“ because coefficients on those covariates can be zero or anything else

Estimation of the unrestricted model

Idea: How would the model fit be if these variables were dropped from the regression? We can consider how much variation is unexplained using *SSR*

```
. regress GPA_year1  female mt_french mt_other hs_q2 hs_q3

      Source |       SS           df       MS      Number of obs   =     1,374
-------------+----------------------------------   F(5, 1368)      =     57.06
       Model |  186.639466         5  37.3278932   Prob > F        =    0.0000
    Residual |  894.929455     1,368  .654188198   R-squared       =    0.1726
-------------+----------------------------------   Adj R-squared   =    0.1695
       Total |  1081.56892     1,373  .787741385   Root MSE        =    .80882

------------------------------------------------------------------------------
   GPA_year1 |      Coef.   Std. Err.      t    P>|t|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
      female |  -.1296654   .0442666    -2.93   0.003    -.2165032   -.0428276
   mt_french |  -.4658419   .3079241    -1.51   0.131    -1.069896    .1382127
    mt_other |  -.0755215     .04796    -1.57   0.116    -.1696047    .0185616
       hs_q2 |   .3743258   .0545463     6.86   0.000     .2673224    .4813292
       hs_q3 |   .8857912   .0532672    16.63   0.000      .781297    .9902854
       _cons |   1.484259   .0467597    31.74   0.000     1.392531    1.575987
------------------------------------------------------------------------------
```

Estimation of the restricted model

The sum of squared residuals *necessarily* increases, but is that increase statistically significant?

```
. regress GPA_year1  female hs_q2 hs_q3

      Source |       SS           df       MS      Number of obs   =     1,374
-------------+----------------------------------   F(3, 1370)      =     93.41
       Model |  183.662596         3  61.2208653   Prob > F        =    0.0000
    Residual |  897.906326     1,370  .655406077   R-squared       =    0.1698
-------------+----------------------------------   Adj R-squared   =    0.1680
       Total |  1081.56892     1,373  .787741385   Root MSE        =    .80957

------------------------------------------------------------------------------
   GPA_year1 |      Coef.   Std. Err.      t    P>|t|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
      female |  -.1333114   .0442706    -3.01   0.003    -.2201568   -.0464659
       hs_q2 |   .3686178   .0545086     6.76   0.000     .2616884    .4755472
       hs_q3 |   .8780338   .0531684    16.51   0.000     .7737334    .9823341
       _cons |   1.466264    .045171    32.46   0.000     1.377652    1.554875
------------------------------------------------------------------------------
```

Test of joint significance, under homoskedasticity

$$F = \frac{(SSR_r - SSR_{ur})/q}{SSR_{ur}/(n - k_{ur} - 1)} \sim F_{q,n-k-1}$$

- q is the number of restrictions you're examining.
- $n - k_{ur} - 1$ is degrees of freedom in unrestricted model

The relative increase of SSR going from restricted to unrestricted follows a F-distribution if H_0 is true.

Rejection rule: The F-distribution

- A F-distributed variable takes on positive values, reflects fact that SS can only increase as we add more variables (move from restricted to unrestricted)
- Choose critical value so that the null hypothesis is falsely rejected in α% of cases (type I error)
- For now, let $\alpha = 0.1$

Test of joint significance

$$F = \frac{(897.91 - 894.93)/2}{894.93/(1374 - 5 - 1)} = 2.28$$

$$F \sim F_{2,1368} \rightarrow c_{0.1} = 2.30$$

$$P(F > 2.28) = 0.1032$$

We cannot reject null at conventional levels.

The relative increase of SSR going from restricted to unrestricted follows a F-distribution if H_0 is true.

In Stata...

```
. regress GPA_year1  female mt_french mt_other hs_q2 hs_q3

      Source |       SS           df       MS      Number of obs   =     1,374
-------------+----------------------------------   F(5, 1368)      =     57.06
       Model |  186.639466         5  37.3278932   Prob > F        =    0.0000
    Residual |  894.929455     1,368  .654188198   R-squared       =    0.1726
-------------+----------------------------------   Adj R-squared   =    0.1695
       Total |  1081.56892     1,373  .787741385   Root MSE        =    .80882

------------------------------------------------------------------------------
   GPA_year1 |      Coef.   Std. Err.      t    P>|t|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
      female |  -.1296654   .0442666    -2.93   0.003    -.2165032   -.0428276
   mt_french |  -.4658419   .3079241    -1.51   0.131    -1.069896    .1382127
    mt_other |  -.0755215     .04796    -1.57   0.116    -.1696047    .0185616
       hs_q2 |   .3743258   .0545463     6.86   0.000     .2673224    .4813292
       hs_q3 |   .8857912   .0532672    16.63   0.000      .781297    .9902854
       _cons |   1.484259   .0467597    31.74   0.000     1.392531    1.575987
------------------------------------------------------------------------------
```

```
. test mt_french = mt_other=0

 ( 1)  mt_french - mt_other = 0
 ( 2)  mt_french = 0

       F(  2,  1368) =    2.28
            Prob > F =    0.1032

. testparm mt_french mt_other
```

Also a good idea to adjust for heteroskedasticity!

```
. regress GPA_year1  female mt_french mt_other hs_q2 hs_q3,robust
```

Linear regression

Number of obs	=	1,374
F(5, 1368)	=	57.50
Prob > F	=	0.0000
R-squared	=	0.1726
Root MSE	=	.80882

GPA_year1	Coef.	Robust Std. Err.	t	P>|t|	[95% Conf.	Interval]
female	-.1296654	.0445133	-2.91	0.004	-.2169872	-.0423437
mt_french	-.4658419	.2451495	-1.90	0.058	-.9467516	.0150677
mt_other	-.0755215	.0484309	-1.56	0.119	-.1705283	.0194853
hs_q2	.3743258	.0535636	6.99	0.000	.2692501	.4794015
hs_q3	.8857912	.0528873	16.75	0.000	.7820421	.9895402
_cons	1.484259	.045705	32.47	0.000	1.3946	1.573918

Also a good idea to adjust for heteroskedasticity!

```
. test mt_french = mt_other=0

 ( 1)  mt_french - mt_other = 0
 ( 2)  mt_french = 0

       F(  2,  1368) =    2.86
            Prob > F =    0.0573

. testparm mt_french mt_other

 ( 1)  mt_french = 0
 ( 2)  mt_other = 0

       F(  2,  1368) =    2.86
            Prob > F =    0.0573
```

Test of overall significance of a regression

- Can we exclude all the regressors?
- General population model (unrestricted) :

$$y_i = \beta_0 + \beta_1 x_{i1} + \beta_2 xi2 + .. + \beta_k x_{ik} + u_i$$

- Test $H_0 : \beta_1 = \beta_2 = ...\beta_k = 0$
- Restricted model:

$$y_i = \beta_0 + u_i$$

$$F = \frac{(SSR_r - SSR_{ur})/q}{SSR_{ur}/(n-k-1)} \sim F_{q,n-k-1} = \frac{R^2/k}{(1-R^2)/(n-k-1) \sim F_{k,n-k-1}}$$

The test of overall significance is reported in most regression packages. Usually, null hypothesis is overwhelmingly rejected.

Test of overall significance of a regression

```
. regress GPA_year1  female mt_french mt_other hs_q2 hs_q3
```

Source	SS	df	MS
Model	**186.639466**	**5**	**37.3278932**
Residual	**894.929455**	**1,368**	**.654188198**
Total	**1081.56892**	**1,373**	**.787741385**

Number of obs	=	**1,374**
F(5, 1368)	=	**57.06**
Prob > F	=	**0.0000**
R-squared	=	**0.1726**
Adj R-squared	=	**0.1695**
Root MSE	=	**.80882**

GPA_year1	Coef.	Std. Err.	t	P>|t|	[95% Conf.	Interval]
female	**-.1296654**	**.0442666**	**-2.93**	**0.003**	**-.2165032**	**-.0428276**
mt_french	**-.4658419**	**.3079241**	**-1.51**	**0.131**	**-1.069896**	**.1382127**
mt_other	**-.0755215**	**.04796**	**-1.57**	**0.116**	**-.1696047**	**.0185616**
hs_q2	**.3743258**	**.0545463**	**6.86**	**0.000**	**.2673224**	**.4813292**
hs_q3	**.8857912**	**.0532672**	**16.63**	**0.000**	**.781297**	**.9902854**
_cons	**1.484259**	**.0467597**	**31.74**	**0.000**	**1.392531**	**1.575987**

Test of overall significance of a regression

```
. testparm *

 ( 1)  female = 0
 ( 2)  mt_french = 0
 ( 3)  mt_other = 0
 ( 4)  hs_q2 = 0
 ( 5)  hs_q3 = 0

       F(  5,  1368) =   57.06
            Prob > F =    0.0000
```

When to use t-tests? F-tests?

- T-test: test of just one regressor, one-sided or two-sided
- F-test: test of just one regressor or multiple regressors, two-sided only

When working with just one regressor, $F = t^2$, so results are functionally equivalent!

Conclusion

Overview of tests

Extended example: Angrist, Lang, and Oreopoulos (2009)

- Testing one restriction, multiple coefficients
- Testing multiple restrictions
- Testing overall significance of a regression