Topic 3 – Very Basics of Hypothesis Testing

- Let's say I put forward a **null hypothesis** $H_0 : \beta = \beta_0$ against the **alternative hypothesis** $H_1 : \beta \neq \beta_0$ for some β_0, how would I know whether I am right or wrong? This is when you have to do hypothesis testing.
- Notice: Your estimator $\hat{\beta}$ never goes in the hypothesis! You goal is to use $\hat{\beta}$ and decide whether you agree with $H_0 : \beta = \beta_0$.
- When our test *incorrectly rejects a null that is in fact true*, we make a **Type I error**. The probability of making this error is called the **size**, **significance level**, **level**, or α. When you hear someone says "I reject the null at 5% significance level", it means someone has 5% chance of rejecting the null wrongly. Remember, there is nothing magical about 5%. It is just a convention.
- Our level gives a **rejection rule** by specifying a **rejection region**: if our test statistic falls into the region, we reject the null hypothesis.
- The **power** of a test is the probability that it *rejects a wrong null*.
- If the null is false and you do not reject it, you make a **Type II error**. The probability of making this error is 1 minus the power of the test.
- Given a level, we want to maximize the power. Notice that the power is a function of the alternative hypothesis (remember that u-shaped function in class).
- The **p-value** is the probability of seeing the test statistic we have if the null hypothesis is true. So small a **p-value** implies strong evidence against the null. Again, we have to decide what "strong" means by choosing the level, which is usually 5%.

▪ Now we talk about how to perform *t*- and *F*-tests in practice, based on some EViews output.

▪ We do the ***t*-test** use the now familiar CPS example:

$$\ln wage_i = \beta_0 + \beta_1 age_i + \beta_2 ed_i + \beta_3 ed_i^2 + u_i$$

Dependent Variable: LNWAGE
Method: Least Squares
Date: 01/30/07 Time: 17:20
Sample (adjusted): 5 134375
Included observations: 97742 after adjustments

Variable	Coefficient	Std. Error	t-Statistic	Prob.
C	0.976252	0.032054	30.45601	0.0000
AGE	0.023232	0.000229	101.6009	0.0000
ED	-0.039157	0.004632	-8.452841	0.0000
ED^2	0.006082	0.000178	34.21813	0.0000

R-squared	0.234489	Mean dependent var	2.518965
Adjusted R-squared	0.234466	S.D. dependent var	1.001028
S.E. of regression	0.875847	Akaike info criterion	2.572791
Sum squared resid	74975.67	Schwarz criterion	2.573180
Log likelihood	-125730.9	F-statistic	9979.609
Durbin-Watson stat	1.831277	Prob(F-statistic)	0.000000

▪ What do the *t*-stats in the result window mean? They are the *t*-stats for the null that the coefficients are **zero**. For example, the *t*-stat for the coefficient on age is calculated as:

$$t = \frac{\hat{\beta}_1 - 0}{\sqrt{Var\left(\hat{\beta}_1\right)}} = \frac{0.023232 - 0}{0.000229} = 101.6009.$$ Since it is larger than 1.96, we reject the null at 5% level (actually at all levels, as the *t*-stat is so big).

- How about we want to test the hypothesis that $\beta_1 = 0.02$? This is simply:

$$t = \frac{\hat{\beta}_1 - 0}{\sqrt{Var\left(\hat{\beta}_1\right)}} = \frac{0.023232 - 0.2}{0.000229} = 14.11354$$. Again it is larger than 1.96, and we reject the null at 5%.

- Not giving up, we test the hypothesis that $\beta_1 = 0.023$. So we calculate

$$t = \frac{\hat{\beta}_1 - 0}{\sqrt{Var\left(\hat{\beta}_1\right)}} = \frac{0.023232 - 0.23}{0.000229} = 1.0131$$. Well, it is smaller than 1.96, so we do NOT reject the null at 5% level. You can check the *t* table that we do not reject it also at 10% level, the critical value of which is 1.282.

- Again, a large number for the *t*-stat means your null hypothesis is "way off", and you choose the level (usually 5%) to define what you mean by "way off".

- Now go to a more complicated hypothesis: $\beta_1 + \beta_2 = 0$, or the coefficients for age and education sums up to zero. There is not much meaning for this hypothesis, though. We again calculate the *t*-stat: $t = \frac{\left(\hat{\beta}_1 + \hat{\beta}_2\right) - 0}{\sqrt{Var\left(\hat{\beta}_1 + \hat{\beta}_2\right)}} = \frac{\left(\hat{\beta}_1 + \hat{\beta}_2\right) - 0}{\sqrt{Var\left(\hat{\beta}_1\right) + Var\left(\hat{\beta}_2\right) + 2Cov\left(\hat{\beta}_1, \hat{\beta}_2\right)}}$.

But wait, the result window above is not enough for us to do that. Let's say you are given the covariance-variance matrix (i.e. $Var\left(\hat{\boldsymbol{\beta}}\right)$, which you can get by **View/Covariance Matrix**):

	C	AGE	ED	ED^2
C	0.001027	-2.17E-06	-0.000140	5.00E-06
AGE	-2.17E-06	5.23E-08	6.88E-08	-4.17E-09
ED	-0.000140	6.88E-08	2.15E-05	-8.03E-07
ED^2	5.00E-06	-4.17E-09	-8.03E-07	3.16E-08

- We can plug in the numbers as follows:

$$t = \frac{0.023232 - 0.039157 - 0}{\sqrt{0.0000000523 + 0.0000215 + 2 * 0.0000000688}} = -3.41941$$, and we reject the null at 5% level.

- So remember to get the denominator right when more than one coefficient are involved, otherwise your inference will be wrong.
- Next we do ***F*-test**. We want to test whether education does matter at all ($\beta_3 = 0$ and $\beta_4 = 0$). To do this test having the above output is not enough.
- Notice that you use the *F*-test whenever you have more than one equal sign in your hypothesis.
- Now you are given another regression, the regression in which the coefficients are "restricted" to be zero (you restrict them by dropping them out for this case):

Dependent Variable: LNWAGE
Method: Least Squares
Date: 01/30/07 Time: 20:49
Sample (adjusted): 5 134375
Included observations: 97742 after adjustments

Variable	Coefficient	Std. Error	t-Statistic	Prob.
C	1.420047	0.009822	144.5765	0.0000
AGE	0.028215	0.000240	117.4856	0.0000

R-squared	0.123745	Mean dependent var	2.518965
Adjusted R-squared	0.123736	S.D. dependent var	1.001028
S.E. of regression	0.937052	Akaike info criterion	2.707864
Sum squared resid	85822.18	Schwarz criterion	2.708058
Log likelihood	-132334.0	F-statistic	13802.86
Durbin-Watson stat	1.723440	Prob(F-statistic)	0.000000

- Now we can do *F*-test. Remember the formula

$$F_{(q,n-k)} = \frac{\left(SSR_{restricted} - SSR_{unrestricted}\right)/\text{\# of restrictions}}{SSR_{unrestricted}\ /(\text{Sample size} - \text{\# of parameters in the unrestricted regression})}$$

$$= \frac{\left(SSR_{restricted} - SSR_{unrestricted}\right)/q}{SSR_{unrestricted}\ /(n-k)}$$

- For our case, q=2 (we restrict two parameters), k=4 and n=97742. So by looking up the sum of squared residuals from the two regressions, we can calculate:

F=((85822.18-74975.67)/2) / (74975.67/(97742-4))=7069.735. Well, since the critical value at 5% is 3, our null hypothesis is rejected ruthlessly.

- Finally we look at a special case of the *F*-test, the **Chow test**.
- Example 1 (*Structural break over time*): We have a certain regression run over the period 1970 to 2007, but we suspect that the parameters during 1970 to 1985 are different from those during 1986 to 2007. How do we test that?
- Example 2 (*Structural break between groups*): You have a wage equation to run for a certain sample, but you suspect that the equation for men is different from that for women (e.g. the return to education is different). How do we test that?
- Let's try with our example. First I run a regression for women:

Dependent Variable: LNWAGE
Method: Least Squares
Date: 01/30/07 Time: 21:19
Sample: 1 134375 IF FE=1
Included observations: 47380

Variable	Coefficient	Std. Error	t-Statistic	Prob.
C	0.615291	0.053754	11.44640	0.0000
AGE	0.020708	0.000324	63.85072	0.0000
ED	-0.006898	0.007725	-0.892948	0.3719
ED^2	0.005266	0.000291	18.11213	0.0000

R-squared	0.216129	Mean dependent var	2.354468
Adjusted R-squared	0.216079	S.D. dependent var	0.985928
S.E. of regression	0.872933	Akaike info criterion	2.566170
Sum squared resid	36101.12	Schwarz criterion	2.566910
Log likelihood	-60788.56	F-statistic	4354.172
Durbin-Watson stat	1.910624	Prob(F-statistic)	0.000000

- Next we run the same thing for men:

Dependent Variable: LNWAGE
Method: Least Squares
Date: 01/30/07 Time: 21:20
Sample: 1 134375 IF FE=0
Included observations: 50362

Variable	Coefficient	Std. Error	t-Statistic	Prob.
C	0.998054	0.038658	25.81772	0.0000
AGE	0.025816	0.000310	83.19948	0.0000
ED	-0.026627	0.005601	-4.754321	0.0000
ED^2	0.005422	0.000218	24.87212	0.0000

R-squared	0.273532	Mean dependent var	2.673722
Adjusted R-squared	0.273489	S.D. dependent var	0.990401
S.E. of regression	0.844175	Akaike info criterion	2.499165
Sum squared resid	35886.67	Schwarz criterion	2.499866
Log likelihood	-62927.47	F-statistic	6320.326
Durbin-Watson stat	1.920528	Prob(F-statistic)	0.000000

- The formula for the Chow test is something that you should remember:

$$F_{(k,n-2k)} = \frac{\left(SSR_{Pooled} - \left(SSR_{Group1} + SSR_{Group2}\right)\right)/k}{\left(SSR_{Group1} + SSR_{Group2}\right)/\left(n-2k\right)}$$

- In our case it is *F*=((74975.67-36101.12-35886.67)/4) / ((36101.12+35886.67)/(97742-8))=1014.122, and the critical value is 2.37. So the null is again rejected.