

# EEP/IAS 118 - Introductory Applied Econometrics, Lecture 7

Gregory Lane

July 2017

# This Lecture

## Topics

- Review
- P-values
- Linear combination tests
- Testing multiple parameters

## Assignments

- Problem Set 3 Due Wednesday
- Quiz 3 Tomorrow
- New Daily Assignments Posted

## Midterm

- This is the last lecture that will be included on the Midterm (one week from today) - bring a calculator!

## Quiz 2

Let's review Quiz 2:

- **Question 1:** We have downward bias. Why?
  - $\beta_{age} > 0$
  - $\rho_{age,educ} < 0$
  - $\Rightarrow Bias = \beta_{age} * \rho_{age,educ}$  is negative
- **Question 2:** Having a random sample will not fix any problem stemming from omitted variable bias.
- **Question 3:**

$$\left[ 10 - 2.262 * \frac{2}{\sqrt{10}} , 10 + 2.262 * \frac{2}{\sqrt{10}} \right]$$

## Problem Set 2

Let's review issues from PS2:

- Remember to specify "holding all else constant", when interpreting  $\hat{\beta}$  in a regression with multiple variables
- Be clear about your reasoning when inferring correlations between variables in OVB questions
- **Exercise 2, Question 5:** What should we infer from the fact that the coefficient on *pencil* was a similar magnitude as the coefficient on *computer*?
  - Since it is silly to think people are paid a higher wage for being able to use a pencil, this implies that there is a common OVB problem in these regressions. We therefore should question the coefficient on *computer*, as the omitted variables that biased our estimate for the pencil regression are likely to also bias our estimate for computers

## Specify Null and Alternative

We observe the average consumption on temptation goods (cigarettes, alcohol, and tobacco) for clients and non-clients of a microfinance institution (MFI).

- Want to test that the average consumption of temptation good is the same between non-MFI clients and MFI clients. Call  $D$  the difference in mean consumption of the two groups.

*Specify the null and alternative hypothesis*

- Want to test that the average consumption of temptation good is higher among non-MFI clients. Call  $D$  the difference in mean consumption of the two groups.

*Specify the null and alternative hypothesis*

## Specify Null and Alternative

- Want to test that the average consumption of temptation good is the same between non-MFI clients and MFI clients. Call  $D$  the difference in mean consumption of the two groups.

*Specify the null and alternative hypothesis:*

$$H_0 : D = 0, \quad H_1 : D \neq 0$$

- Want to test that the average consumption of temptation good is higher among non-MFI clients. Call  $D$  the difference in mean consumption of the two groups.

*Specify the null and alternative hypothesis:*

$$H_0 : D = 0, \quad H_1 : D > 0$$

## Specify Null and Alternative

According to the 2010 U.S. Census (i.e., US population), the average household income was \$100,000. Suppose in the year 2015 you draw a random sample of two thousand U.S. households that has a sample mean of \$105,000 with a sample standard deviation of \$6,000.

We want to conduct a hypothesis test to evaluate whether average incomes in 2015 are larger than those in 2010. Write down the null and alternative hypotheses

## Specify Null and Alternative

We want to conduct a hypothesis test to evaluate whether average incomes in 2015 are larger than those in 2010. Write down the null and alternative hypotheses

$$H_0 : \mu_{2015} = 100,000$$

$$H_1 : \mu_{2015} > 100,000$$

What were our clues?

- In description of 2010 data "i.e. US population"  $\rightarrow$  100,000 is the true population parameter, not an estimator
- There was no standard deviation given for the 2010 data, if there is a difference in means test you will need to calculate  $s^2$  for *both* samples



## Aside: Relationship Between CI and Hypo. Tests

We have talked about these procedures separately, but they are related

- The basic formula for a CI is:

$$[\hat{\theta} - c \cdot se(\hat{\theta}), \hat{\theta} + c \cdot se(\hat{\theta})]$$

- The basic formula for a test stat which we compare to  $c$  is:

$$\frac{\hat{\theta} - \theta_{H_0}}{se(\hat{\theta})}$$

- If  $\theta_{H_0}$  falls outside of the the CI for a given size (e.g. 90%, 95%, 99%), then we would also reject at the same significance level that the null hypothesis is true

# P-Values

Selecting the significance level ( $\alpha$ ) can be arbitrary

- Two reasonable people could choose different  $\alpha$  and reach different conclusions
- Is a t-stat of 1.94 really that different from a t-stat of 1.96?
- Binary reject / fail-to-reject obscures the continuous nature of the random draw

# P-Values

Instead we can examine the p-value of a test for an objective idea of the robustness of the result. What is a p-value?

Here are 2 definitions of p-values that say the same thing using different words:

- The p-value is the smallest significance level at which the null hypothesis would be rejected.
- The p-value is the probability of obtaining a value of the test statistic as extreme or more extreme than the one actually obtained from the sample under the null (i.e if the null is true).

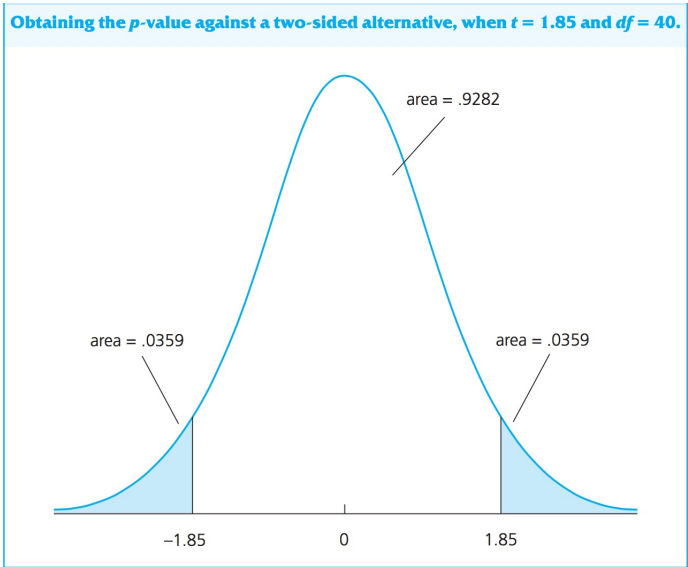
# P-Values

- P-values close to zero constitute strong evidence against the null  $H_0$
- Large p-values close to one constitute weak evidence against the null  $H_0$ .

**Example:** Suppose we calculate a test statistic of  $t=1.85$ , with 40 degrees of freedom (two sided). We can find the p-value

$$pvalue = P(|T| > 1.85 | H_0) = 2P(T > 1.85) = 2(0.0359) = 0.0718$$

P-value of 0.0718:



## p-values

In this example, if the null were true (e.g.  $\beta = 0$  or  $\mu = 10$ , etc.), then:

- We would observe a t-stat as large as 1.85 about 7.2% of the time
- We would *fail* to reject the null at 5% significance level
- We would reject the null at 10% significance level

p-values are reported as the fourth column in the Stata output

## p-values in Stata

Source	SS	df	MS	Number of obs	=	526
-----+-----				F(3, 522)	=	94.75
Model	52.2939096	3	17.4313032	Prob > F	=	0.0000
Residual	96.0358418	522	.183976708	R-squared	=	-----
-----+-----				Adj R-squared	=	0.3488
Total	148.329751	525	.28253286	Root MSE	=	.42893
-----						
lwage	Coef.	Std. Err.	t	P> t	[95% Conf. Interval]	
educ	.0912897	.0071232	-----	0.000	-----	-----
exper	.0094139	.0014493	6.50	0.000	.0065667	.012261
female	-.3435967	.0376668	-9.12	0.000	-.4175939	-.2695996
_cons	.4808357	.1050163	4.58	0.000	.2745292	.6871421

- The p-value stata shows you is the p-value for the null hypothesis that  $\hat{\beta} = 0$ . This is not necessarily the one you want
- We now know almost everything the regression output produces

# Sign, Size, and Significance

We can now address the last step in our interpretation of  $\hat{\beta}$  from regression output:

- In addition to sign and size, we need to comment at what level the coefficient is statistically significant
- Statistically significant means: *we reject the null hypothesis that  $\beta$  is equal to zero*
- We can determine this by looking at the p-value (or t-stat) for the coefficient in the regression output
- **Language:** " $\hat{\beta}$  is statistically significant at the 10% level, but not at the 5% level (p-value of 0.07)"
  - Stick to the standard significance levels (10%, 5%, 1%). If the estimator is not significant at 10% we say it is statistically insignificant



# Linear Combination Test

Up until now, we have restricted ourselves to testing whether  $\beta$  is equal to a constant (most often zero)

However, that is not always what we care about. Often we want to know the relative size of two  $\beta$ . E.g. does

$$\beta_2 = \beta_3$$

Or maybe we want to know if two effects cancel each other out. E.g. running (good for you) in a badly polluted area (bad for you). In this case maybe we want to know if

$$\beta_{\text{running}} + \beta_{\text{pollution}} = 0$$

We call these “linear combination” tests

# Linear Combination Test

How do we go about this process?

Theoretically, this is similar to process of testing the equality of two estimated means:

- 1 Define hypothesis:  $H_0 : \beta_1 = \beta_2$     $H_1 : \beta_1 \neq \beta_2$
- 2 Write the test statistic:

$$\frac{\hat{\beta}_1 - \hat{\beta}_2}{se(\hat{\beta}_1 - \hat{\beta}_2)}$$

- 3 Stop because we don't know how to calculate  $se(\hat{\beta}_1 - \hat{\beta}_2)$

We *can* calculate this value, but it is complex. There are two easier ways to do this in Stata.

# Linear Combination Test

First option:

- 1 Change variables in the regression so that the output tests equality directly

Lets say we have this model

$$lwage = \beta_0 + \beta_1jc + \beta_2univ + \beta_3exp + u$$

We want to test whether the returns to a year of junior college are the *same* as a year of four year college

- 1 Define our hypothesis

$$H_0 : \beta_1 = \beta_2 \text{ or } \beta_1 - \beta_2 = 0$$

$$H_1 : \beta_1 - \beta_2 \neq 0$$

# Linear Combination Test

Now, we need to find an estimate of  $\beta_1 - \beta_2$  and the standard error of this object. We accomplish this through changing the variables in our regression:

Define  $\theta_1 = \beta_1 - \beta_2$

$$\begin{aligned}lwage &= \beta_0 + \beta_1jc + \beta_2univ + \beta_3exp + u \\ &= \beta_0 + (\theta_1 + \beta_2)jc + \beta_2univ + \beta_3exper + u \\ &= \beta_0 + \theta_1jc + \beta_2(jc + univ) + \beta_3exper + u \\ &= \beta_0 + \theta_1jc + \beta_2(totcoll) + \beta_3exper + u\end{aligned}$$

So we can test if  $\theta$  is different from zero directly. **NOTE:** we could have also replaced  $jc$  with  $totcoll$  and we would also get the same test on  $\theta$

# Linear Combination Test

Option 2:

After regression output, write command

$$\text{test } Var1 == Var2$$

or

$$\text{test } 2 * Var1 == Var2$$

The output will give you a test statistic and a p-value for the test you specify. This is useful to know for your problem sets, but doesn't help you understand what is going on.

# F-Test

Suppose we want to test whether two parameters are *jointly* different from zero. Conceptually this is different from a linear combination test because we may want to know if two  $\beta$  are jointly significant even if they cancel each other out

- E.g. if  $\beta_2 = -2$  and  $\beta_3 = 2$  both significant, a linear combination test of  $\beta_2 + \beta_3$  would not reject the null, but an F-test of joint significance would reject the null

To do this we use an F-test. It's easiest to see how to do this from an example from a development data set on women's autonomy.

# F-test

- Have data on women's autonomy, their current age, the age they got married, husband education, and urban or rural.
- Want to test whether age in general (*crage*, *mrage*) matters for autonomy

Steps:

- 1 Define hypotheses:

$$H_0 : \beta_{crage} = 0 \quad \& \quad \beta_{mrage} = 0$$

$$H_1 : \beta_{crage} \neq 0 \text{ or } \beta_{mrage} \neq 0 \text{ or both}$$

Note the form of the null and alternative hypothesis!

# F-Test

- 2 Write down the two models the null hypothesis implies:

*Unrestricted model:*

$$\begin{aligned} \textit{autonomy} = & \beta_0 + \beta_1 \textit{mrage} + \beta_2 \textit{crage} + \beta_3 \textit{husbedu} \\ & + \beta_4 \textit{urban} + u \end{aligned}$$

*Restricted model:*

$$\textit{autonomy} = \beta_0 + \beta_3 \textit{husbedu} + \beta_4 \textit{urban} + u$$

- We call the regression with the variables we are testing the **unrestricted model**.
- The regression without these variables is the **restricted model**
- Estimate both these models separately



# F-Test

- **Unrestricted model:** is the model where we allow  $\beta_{crage}$  and  $\beta_{mrage}$  to take on any value that best fits the data
- **Restricted model:** is the model where we force  $\beta_{crage}$  and  $\beta_{mrage}$  to be equal to zero
- For the F-test, we estimate both these models separately
  - We use the regression output from both models to calculate our test stat

# F-Test

- 3 Write our F-stat from the two regression outputs. **Note:** there are two equivalent ways to calculate the test statistic:

$$F = \frac{(SSR_R - SSR_{UR}) / q}{SSR_{UR} / (n - k_{UR} - 1)}$$

$$F = \frac{(R_{UR}^2 - R_R^2) / q}{(1 - R_{UR}^2) / (n - k_{UR} - 1)}$$

Where

- $q$  is the number of restrictions
- $k$  is the number of variables in the unrestricted model
- F-stat is distributed  $F_{q, n-k-1}$  - i.e. there are **two** degrees of freedom to track, the numerator ( $q$ ) and the denominator ( $n - k - 1$ )

# F-test

- 4 Compare the F-stat to the correct critical value found in the F-table. You will need to keep track of **both** numerator degrees of freedom ( $q$ ) and denominator degrees of freedom ( $n - k_{UR} - 1$ )
- 5 Interpret: We reject / don't reject the null hypothesis that both  $\beta_1$  and  $\beta_2$  are equal to zero

# F-Test: Example

```
reg autonomy marr_age curr_age husb_educ urban
```

Source	SS	df	MS
Model	949.690268	4	237.422567
Residual	15809.3097	971	16.2814724
Total	16759	975	17.1887179

```
Number of obs = 976  
F( 4, 971) = 14.58  
Prob > F = 0.0000  
R-squared = 0.0567  
Adj R-squared = 0.0528  
Root MSE = 4.035
```

```
reg autonomy husb_educ urban
```

Source	SS	df	MS
Model	486.726699	2	243.363349
Residual	16317.306	975	16.7356985
Total	16804.0327	977	17.1996241

```
Number of obs = 976  
F( 2, 975) = 14.54  
Prob > F = 0.0000  
R-squared = 0.0290  
Adj R-squared = 0.0270  
Root MSE = 4.0909
```

Fill in the blanks:

$$F = \frac{(SSR_R - SSR_{UR}) / q}{SSR_{UR} / (n - k_{UR} - 1)}$$

# F-Test

Fill in the formula:

$$F = \frac{(SSR_R - SSR_{UR}) / q}{SSR_{UR} / (n - k_{UR} - 1)}$$

- $SSR_R = 16317.3$
- $SSR_{UR} = 15809.3$
- $R_R^2 = 0.029$
- $R_{UR}^2 = 0.057$
- $n = 976$
- $k_{UR} = 4$
- $q = 2$

$$F = \frac{(16317.306 - 15809.3097) / 2}{15809.3097 / (976 - 4 - 1)} = 15.6_{\{2,971\}}$$

# F-Table

818

Appendix G Statistical Tables

**TABLE G.3a**  
10% Critical Values of the F Distribution

		Numerator Degrees of Freedom									
		1	2	3	4	5	6	7	8	9	10
D e n o m i n a t o r	10	3.29	2.92	2.73	2.61	2.52	2.46	2.41	2.38	2.35	2.32
	11	3.23	2.86	2.66	2.54	2.45	2.39	2.34	2.30	2.27	2.25
	12	3.18	2.81	2.61	2.48	2.39	2.33	2.28	2.24	2.21	2.19
	13	3.14	2.76	2.56	2.43	2.35	2.28	2.23	2.20	2.16	2.14
	14	3.10	2.73	2.52	2.39	2.31	2.24	2.19	2.15	2.12	2.10
D e g r e e s	15	3.07	2.70	2.49	2.36	2.27	2.21	2.16	2.12	2.09	2.06
	16	3.05	2.67	2.46	2.33	2.24	2.18	2.13	2.09	2.06	2.03
	17	3.03	2.64	2.44	2.31	2.22	2.15	2.10	2.06	2.03	2.00
	18	3.01	2.62	2.42	2.29	2.20	2.13	2.08	2.04	2.00	1.98
	19	2.99	2.61	2.40	2.27	2.18	2.11	2.06	2.02	1.98	1.96
o f	20	2.97	2.59	2.38	2.25	2.16	2.09	2.04	2.00	1.96	1.94
	21	2.96	2.57	2.36	2.23	2.14	2.08	2.02	1.98	1.95	1.92
	22	2.95	2.56	2.35	2.22	2.13	2.06	2.01	1.97	1.93	1.90
	23	2.94	2.55	2.34	2.21	2.11	2.05	1.99	1.95	1.92	1.89
	24	2.93	2.54	2.33	2.19	2.10	2.04	1.98	1.94	1.91	1.88
F r e e d o m	25	2.92	2.53	2.32	2.18	2.09	2.02	1.97	1.93	1.89	1.87
	26	2.91	2.52	2.31	2.17	2.08	2.01	1.96	1.92	1.88	1.86
	27	2.90	2.51	2.30	2.17	2.07	2.00	1.95	1.91	1.87	1.85
	28	2.89	2.50	2.29	2.16	2.06	2.00	1.94	1.90	1.87	1.84
	29	2.89	2.50	2.28	2.15	2.06	1.99	1.93	1.89	1.86	1.83
	30	2.88	2.49	2.28	2.14	2.05	1.98	1.93	1.88	1.85	1.82
	40	2.84	2.44	2.23	2.09	2.00	1.93	1.87	1.83	1.79	1.76
	60	2.79	2.39	2.18	2.04	1.95	1.87	1.82	1.77	1.74	1.71
	90	2.76	2.36	2.15	2.01	1.91	1.84	1.78	1.74	1.70	1.67
	120	2.75	2.35	2.13	1.99	1.90	1.82	1.77	1.72	1.68	1.65
	∞	2.71	2.30	2.08	1.94	1.85	1.77	1.72	1.67	1.63	1.60

Example: The 10% critical value for numerator  $df = 2$  and denominator  $df = 40$  is 2.44.  
Source: This table was generated using the Stata® function invfprob.

819

Appendix G Statistical Tables

**TABLE G.3b**  
5% Critical Values of the F Distribution

		Numerator Degrees of Freedom									
		1	2	3	4	5	6	7	8	9	10
D e n o m i n a t o r	10	4.96	4.10	3.71	3.48	3.33	3.22	3.14	3.07	3.02	2.98
	11	4.84	3.98	3.59	3.36	3.20	3.09	3.01	2.95	2.90	2.85
	12	4.75	3.89	3.49	3.26	3.11	3.00	2.91	2.85	2.80	2.75
	13	4.67	3.81	3.41	3.18	3.03	2.92	2.83	2.77	2.71	2.67
	14	4.60	3.74	3.34	3.11	2.96	2.85	2.76	2.70	2.65	2.60
D e g r e e s	15	4.54	3.68	3.29	3.06	2.90	2.79	2.71	2.64	2.59	2.54
	16	4.49	3.63	3.24	3.01	2.85	2.74	2.66	2.59	2.54	2.49
	17	4.45	3.59	3.20	2.96	2.81	2.70	2.61	2.55	2.49	2.45
	18	4.41	3.55	3.16	2.93	2.77	2.66	2.58	2.51	2.46	2.41
	19	4.38	3.52	3.13	2.90	2.74	2.63	2.54	2.48	2.42	2.38
o f	20	4.35	3.49	3.10	2.87	2.71	2.60	2.51	2.45	2.39	2.35
	21	4.32	3.47	3.07	2.84	2.68	2.57	2.49	2.42	2.37	2.32
	22	4.30	3.44	3.05	2.82	2.66	2.55	2.46	2.40	2.34	2.30
	23	4.28	3.42	3.03	2.80	2.64	2.53	2.44	2.37	2.32	2.27
	24	4.26	3.40	3.01	2.78	2.62	2.51	2.42	2.36	2.30	2.25
F r e e d o m	25	4.24	3.39	2.99	2.76	2.60	2.49	2.40	2.34	2.28	2.24
	26	4.23	3.37	2.98	2.74	2.59	2.47	2.39	2.32	2.27	2.22
	27	4.21	3.35	2.96	2.73	2.57	2.46	2.37	2.31	2.25	2.20
	28	4.20	3.34	2.95	2.71	2.56	2.45	2.36	2.29	2.24	2.19
	29	4.18	3.33	2.93	2.70	2.55	2.43	2.35	2.28	2.22	2.18
	30	4.17	3.32	2.92	2.69	2.53	2.42	2.33	2.27	2.21	2.16
	40	4.08	3.23	2.84	2.61	2.45	2.34	2.25	2.18	2.12	2.08
	60	4.00	3.15	2.76	2.53	2.37	2.25	2.17	2.10	2.04	1.99
	90	3.95	3.10	2.71	2.47	2.32	2.20	2.11	2.04	1.99	1.94
	120	3.92	3.07	2.68	2.45	2.29	2.17	2.09	2.02	1.96	1.91
	∞	3.84	3.00	2.60	2.37	2.21	2.10	2.01	1.94	1.88	1.83

Example: The 5% critical value for numerator  $df = 4$  and large denominator  $df (= \infty)$  is 2.37.  
Source: This table was generated using the Stata® function invfprob.

# Hypothesis Testing & CI Overview

## Standard Deviation vs. Standard Error:

Test Type	Variance of $X$ or $u$	StDev of $X$ or $u$	Var. of Estimator	StDev Estimator
$\mu$ , known $\sigma^2$	$\sigma_X^2 = E[(X_i - E[X_i])^2]$	$\sqrt{\sigma_X^2}$	$\frac{\sigma_X^2}{n}$	$\sqrt{\frac{\sigma_X^2}{n}}$
Unknown $\sigma^2$	$s_X^2 = \sum_i (X_i - \bar{X})^2$	$\sqrt{s_X^2}$	$\frac{s_X^2}{n}$	$\sqrt{\frac{s_X^2}{n}}$
Proportion ( $p$ )	$Var(X) = p(1-p)$	$\sqrt{p(1-p)}$	$\frac{p(1-p)}{n}$	$\sqrt{\frac{p(1-p)}{n}}$
$\beta$ , known $\sigma^2$	$\sigma^2$	$\sigma$	$\frac{\sigma^2}{SST_{X_j}(1-R_j^2)}$	$\sqrt{\frac{\sigma^2}{SST_{X_j}(1-R_j^2)}}$
Unknown $\sigma^2$	$\left( \hat{\sigma}^2 = \frac{\sum \hat{u}_i^2}{n-k-1} \right)$	$(\sqrt{\hat{\sigma}^2})$	$\frac{\hat{\sigma}^2}{SST_{X_j}(1-R_j^2)}$	$\sqrt{\frac{\hat{\sigma}^2}{SST_{X_j}(1-R_j^2)}}$

# T-Stats and Distributions Overview

Fill in the blanks!

Test Type	Test Statistic	Distribution
Population mean e.g. $H_0 : \mu = \mu_0$		
Difference in population means e.g. $H_0 : \mu_1 - \mu_2 = \mu_0$		
Population proportion e.g. $H_0 : p = p_0$		
Difference in population proportions e.g. $H_0 : p_1 - p_2 = p_0$	$z = \frac{(\hat{p}_1 - \hat{p}_2) - p_0}{\sqrt{\frac{\hat{p}(1-\hat{p})}{n_1} + \frac{\hat{p}(1-\hat{p})}{n_2}}}$	$z \sim N(0, 1)$
True regression parameter ( $k$ other vars) e.g. $H_0 : \beta = \beta_0$		
Multiple restrictions in regression ( $q$ restrictions, $k$ total variables in UR model)	$F = \frac{(R_{UR}^2 - R_R^2)/q}{(1 - R_{UR}^2)/(n - k - 1)}$	$F \sim F_{q, n-k-1}$



# T-Stats and Distributions Overview

Test Type	Test Statistic	Distribution
Population mean e.g. $H_0 : \mu = \mu_0$	$t = \frac{\bar{x} - \mu_0}{\sqrt{\frac{s^2}{n}}}$	$t \sim t_{n-1}$
Difference in population means e.g. $H_0 : \mu_1 - \mu_2 = \mu_0$	$t = \frac{(\bar{x}_1 - \bar{x}_2) - \mu_0}{\sqrt{\frac{s_1^2}{n_1} + \frac{s_2^2}{n_2}}}$	$t \sim t_{n_1 + n_2 - 2}$
Population proportion e.g. $H_0 : p = p_0$	$z = \frac{\hat{p} - p_0}{\sqrt{\frac{p_0(1-p_0)}{n}}}$	$z \sim N(0, 1)$
Difference in population proportions e.g. $H_0 : p_1 - p_2 = p_0$	$z = \frac{(\hat{p}_1 - \hat{p}_2) - p_0}{\sqrt{\frac{\hat{p}(1-\hat{p})}{n_1} + \frac{\hat{p}(1-\hat{p})}{n_2}}}$	$z \sim N(0, 1)$
True regression parameter ( $k$ other vars) e.g. $H_0 : \beta = \beta_0$	$t = \frac{\hat{\beta} - \beta_0}{SE(\hat{\beta})}$	$t \sim t_{n-k-1}$
Multiple restrictions in regression ( $q$ restrictions, $k$ total variables in UR model)	$F = \frac{(R_{UR}^2 - R_R^2)/q}{(1 - R_{UR}^2)/(n-k-1)}$	$F \sim F_{q, n-k-1}$

# Practice

Fill in the blanks:

```
. reg lwage educ exper female nonwhite
```

Source	SS	df	MS	Number of obs =	2000
Model	182.711923	4	45.6779807	F( 4, 1995) =	186.03
Residual	489.864945	1995	.245546338	Prob > F =	0.0000
				R-squared =	.....
Total	672.576867	1999	.336456662	Adj R-squared =	0.2702
				Root MSE =	.49553

lwage	Coef.	Std. Err.	t	P> t	[95% Conf. Interval]	
educ	.1166997	.0053153	.....	0.000	.1062756	.1271237
exper	.0108872	.0008691	12.53	0.000	.0091827	.0125917
female	-.2533177	.0222198	-11.40	0.000	.....	.....
nonwhite	-.0374311	.....	-1.20	.....	-.0985117	.0236495
_cons	1.061903	.0759003	13.99	0.000	.9130514	1.210756