

Lab 9 - Inference

Spring 2017

1 Introduction

The goal of this handout is to provide you with an introduction to inference i.e. testing different hypotheses using Stata. The implementation of a variety of very complicated tests is easily executed in Stata. Once a specification is estimated, a simple post-estimation command will provide you with a p-value for the corresponding test. Given the simplicity of testing in Stata we will use the opportunity to provide you with a quick introduction to hypothesis testing as well as present some examples.

2 Inference: the basics

Econometric analysis usually starts with a specification of a model. After allowing for several fundamental assumptions about the population from which our dataset was sampled we can proceed to estimate the specified model. Analyzing the estimation results is the ultimate goal of econometric research. We specify and estimate models in order to investigate a hypothesis about the effect of a factor or a combination of factor on a certain process. This last stage of the analysis where we test our hypothesis is called the inference stage and it has the following structure:

1. Hypothesis specification.
2. Calculation of a test statistic based on the estimation.
3. Calculation of the percentile of the test statistic.
4. Comparison of the percentile of the test statistic to the level of significance we are willing to tolerate.

2.1 Hypothesis Specification

- We will never be able to observe the true population parameters even after we estimate our regression. What we can do is to quantify how likely different values of the population parameters are.
- *Examples:* (1) we can quantify how likely it is (based on our estimated coefficients) that one of the true population parameters is zero: $\beta_1 = 0$. This is the hypothesis we specify when we do inference on the statistical significance of a parameter. (2) we can quantify the likelihood

that one of the true population parameters takes a specific values: $\beta_2 = 1, 100, -3$. (3) we can quantify the likelihood that one of the true parameters is equal to another one: $\beta_3 = \beta_4$. (4) the first parameter is half the size of the second parameter parameter: $\beta_1 = \frac{1}{2} \times \beta_2$. (5) the first parameter is strictly positive: $\beta_1 \leq 0$.

- Note that the examples above involve β 's not $\hat{\beta}$'s. We are hypothesizing about the population parameters not about the actual parameters we have calculated from our current sample - we already know these estimates.
- Note that the β 's in our regression represent partial effects (we hold all other factors fixed). As a result our hypotheses should be interpret as specifying the value of the effect after we have controlled for all other factors.
- Sometimes people refer to these hypotheses about the population parameters as restrictions on the specification of the model. They are called this way because they restrict the true value of the population parameter and require that our sample abides to them.
- All of the above examples involve one particular parameter (or in the example (3) compare one parameter to another). We refer to these as a single-restriction hypotheses.
- Alternatively we can have multiple-restriction hypotheses. These involve simultaneous restrictions about more than one parameters in the population.
- *Examples:* (1) the first three parameters in the model we have specified about the population are zero: $\beta_1 = 0$ AND $\beta_2 = 0$ AND $\beta_3 = 0$. (2) the first parameter is equal to the second and the third is equal to zero: $\beta_1 = \beta_2$ AND $\beta_3 = 0$.
- All the hypotheses we have given as examples are referred to as the "Null Hypothesis" or H_0 .
- By construction our hypothesis has an alternative. In econometrics (and in logic) the alternative of our hypothesis is everything else. For example the alternative of $\beta_2 = 1$ is $\beta_2 \neq 1$. The alternative of $\beta_1 = 0$ AND $\beta_2 = 0$ AND $\beta_3 = 0$ is that one parameter is non-zero or that two parameters are non-zero or that all of them are non-zero. The alternative is referred to as the "Alternative Hypothesis" or H_1 .
- You have learned that one of the fundamental regression assumptions is that the model we specify is the true model. When we conduct inference we also assume that H_0 is true about the population.
- The idea behind inference is to allow for H_1 in your specification even though H_0 is really true, to estimate your model still assuming that H_0 is true, and to quantify the likelihood that H_0 is true in your inference stage. If H_0 is indeed true the likelihood should be "high." If the likelihood is not "high" you will reject H_0 . Otherwise you will fail to reject H_0 .

2.2 Test statistic for your H_0 , the percentile of the test statistic, and test outcome

When quantifying the likelihood of the restriction/s in our hypotheses we differentiate between the cases of one and the cases of multiple restrictions. It may seem that we can analyze multiple restrictions just by looking at a group of single restrictions. For concreteness consider the multiple restriction: $H_0^{joint} : \beta_1 = 0$ AND $\beta_2 = 0$. It is tempting to look at the two individual hypotheses

$H_0^1 : \beta_1 = 0$ and $H_0^2 : \beta_2 = 0$ and to conclude that unless we reject both hypotheses we cannot reject the joint hypotheses. This is not necessarily true. It is important to emphasize this point because it underlines the motivation for conducting multiple-restrictions tests. If the two independent variables that we are testing are highly correlated one of them will not be statistically significant. Yet both of them may be important in explaining the dependent variable.

2.2.1 Single Restriction

- The test statistic we use for the case of a single restriction is called the t ratio. The t ratio is calculated for a given parameter: $t_{\hat{\beta}}$. It is equal to:

$$t_{\hat{\beta}} = \frac{\hat{\beta} - \beta_{H_0}}{\sqrt{\widehat{\text{Var}}(\hat{\beta})}} \sim t_{n-1-k} \quad (1)$$

- Replace β_{H_0} with whatever you hypothesize β is according to H_0 . Remember that we assume that H_0 is true in the population when we estimate our regression. You will be able to find $\hat{\beta}$ and $\widehat{\text{Var}}(\hat{\beta})$ from the regression output.
- Under the classical regression assumptions the t ratio is distributed according to the t distribution with $n - 1 - k$ degrees of freedom where k is the number of parameters outside of the intercept that we estimate.
- Once we have calculated $t_{\hat{\beta}}$ we can use our favorite statistical software to tell us what is the percentile of the estimate according to the distribution that it follows. In Stata you will be able to see both the critical value and the percentile in most of the output results (more on this later).
- *Example:* (1) If $H_0 : \beta = 0$ then use $\beta_{H_0} = 0$. If $\hat{\beta} = .396$ and $\widehat{\text{Var}}(\hat{\beta}) = .04$ and $n - 1 - k = 120$ then the t ratio is in the 95th percentile. This means that if H_0 is indeed true (as we have been assuming all along) and we have 100 samples of the population then roughly 95 of these samples would lead to a t ratio below what we calculated. If we reject H_0 when the t ratio is in the 95th percentile this means that we are willing to tolerate making a mistake in more than five of the 100 samples of the population. If we assume a 10% significance level we will reject H_0 when the t ratio is in the 95th percentile and will falsely reject H_0 even if it is true in about 10 samples. This is because we can still have a dataset that produces a t ratio in the 95th when the true parameter is 0 (albeit very rarely). Note that if $\hat{\beta} = -.396$ and $\widehat{\text{Var}}(\hat{\beta}) = .4$ and $n - 1 - k = 120$ then the t ratio is in the 5th percentile. A t ratio in the 5th occurs only 5% of the time and has a small probability of occurring if H_0 were indeed 0.
- *Example:* (2) If $H_0 : \beta = 3$ then use $\beta_{H_0} = 3$. If $\hat{\beta} = 3.396$ and $\widehat{\text{Var}}(\hat{\beta}) = .04$ and $n - 1 - k = 120$ then the t ratio is in the 95th percentile. Same result as above follows.
- *Example:* (3) If $H_0 : \beta \leq 0$ then use $\beta_{H_0} = 0$. If $\hat{\beta} = .396$ and $\widehat{\text{Var}}(\hat{\beta}) = .04$ and $n - 1 - k = 120$ then the t ratio is in the 95th percentile. If we can tolerate more than 5% error in our conclusion (from the perspective of repeated sampling) we will reject t ratios in the 95th. If we can tolerate a 10% chance of error we will reject H_0 for t ratios above the 90th percentile. VERY IMPORTANT: if the t ratio is in ANY percentile below the 90th we fail to reject H_0 if we set a 10% confidence level. This is true because we have a one-sided test. We only care if the true beta is above zero so we ignore t ratios in the percentiles below the 10th or the 5th as they indicate that the true beta could be below zero and that does not reject our H_0 .

2.2.2 Multiple Restrictions

- Here we hypothesize that at least two of the population parameters take particular values. This is the H_0 that we believe is the truth. As discussed above, during inference we allow for H_1 and proceed to estimate the model under H_1 while believing that H_0 is true.
- For concreteness, take $H_0 : \beta_1 = 0$ AND $\beta_2 = 0$ AND $\beta_3 = 0$. The alternative, H_1 , is that at least one (meaning that one, two, or all) are different from zero.
- Implementing a simple test for multiple restrictions called the F test we follow three steps:
 1. We estimate our model allowing for H_1 . This means that we allow all three betas to take any value. We call this out unrestricted model. We write down $A_{ur} = SSR_{ur}/(n-k-1)$ for this estimation. A_{ur} is the sample variance of the estimated residual from this regression.
 2. We estimate our model allowing for H_0 . This means that we allow all three betas to be exactly zero (we drop them from the estimation). We call this the restricted model. We write down the increment in SSR going from H_1 to H_0 : $SSR_r - SSR_{ur}$. The increment is always non-negative since applying a restriction to our model will require to a bigger proportion of the variation in dependent variable to be explained by the residual. This is the reason why the F distribution is non-negative. Write down $B_r = (SSR_r - SSR_{ur})/q$ where q is the number of conditions in H_0 .
 3. The test statistic is:

$$F = B_r/A_{ur} \sim F_{q,n-k-1} \quad (2)$$

- Once we find F we can find its percentile according to $F_{q,n-k-1}$. Note that this statistic is always non-negative. If H_0 were indeed true in the population the increment in SSR per condition B_r relative to A_{ur} will NOT be high and F will be in a LOW percentile. Only when F is in a high percentile e.g. the 95th can we reject H_0 . The reason for this is that SSR is a rough measure of how well our model explains the dependent variable. If the increment in SSR going from H_1 to H_0 is too "big" this means that imposing H_0 leads to a deterioration in the fit of our model captured by the larger variance of the residuals in the restricted model.
- *Example:* We test whether three different measures of intelligence matter for the wage one receives: $H_0 : \beta_{SAT} = 0$ AND $\beta_{GPA} = 0$ AND $\beta_{IQ} = 0$. All three will be correlated so we cannot look at individual t ratios. We estimate under H_1 and find that $SSR_{ur} = 190$ with $n - 1 - k = 380 \implies A_{ur} = .5$. We estimate again under H_0 by dropping all intelligence variables. We find that $SSR_r = 196$ so the SSR increment is 6. The increment per condition in H_0 is $B_r = 6/3 = 2$. $F = 2/.5 = 4$. 4 is in the 99th percentile of $F_{3,380}$ so it looks like H_0 is pretty unlikely.

3 Inference in Stata

In this section we apply the basics we learned in the previous section. We will use the dataset `wine` to conduct inference.

3.1 Single Restrictions

- We begin by specifying the model:

$$deaths_i = \alpha + \beta_1 \times alcohol_i + \beta_2 \times heart_i + \beta_3 \times liver_i + u_i \quad (3)$$

- We hypothesize that after controlling for other health factors average alcohol consumption does not impact average deaths in a country.

- **3.1.1** $H_0 : \beta_1 = 0$ and $H_1 : \beta_1 \neq 0$

- We estimate the above model under H_1 but assume that H_0 is indeed true. Remember that under the classical regression assumptions we can include insignificant factors without a major violation of the properties of the estimated coefficients.

- Under H_1 we find

deaths	Coef.	Std. Err.	t	P> t	[95% Conf. Interval]
alcohol	-5.619184	8.716553	-0.64	0.528	-24.0095 12.77113
heart	1.259491	.2504379	5.03	0.000	.7311134 1.787869
liver	3.94427	1.700801	2.32	0.033	.3558938 7.532646
_cons	532.1875	69.85662	7.62	0.000	384.8029 679.5721

- To test H_0 we can either use the t ratio or compute the F . We can read the t ratio from the regression output. It is in the 26th percentile so we fail to reject H_0 .
- How did we know that the t ratio is in the 26th percentile? If you look carefully at the output columns you can see that the p-value is for P>|t| and it gives us the probability that the t ratio is below -.64 and above .64. Dividing the probability value by 2 will tell us what is the probability that the t ratio is below -.64 i.e. the percentile of -.64.
- To compute the F for this test use `test alcohol`. `alcohol` refers to the population coefficient on the variable alcohol as specified in the H_0 . We could alternatively write the command `test alcohol=0`. It is equivalent to the original command. This give you:

```
( 1) alcohol = 0
      F( 1, 17) = 0.42
      Prob > F = 0.5278
```

- You should note a couple of things: 1) the output explicitly shows that there is only one restriction being tested: `alcohol = 0`. 2) Since the F test cannot distinguish between a rejection of H_0 when the estimated coefficient is very negative or very positive we get that the percentile of the F is equal to the combined probability of the rejection of H_0 in either of the cases. The combined probability is equal to the size of the percentiles below $-t_{\hat{\beta}}$ or above $t_{\hat{\beta}}$ if we were using a single-restriction test.
- We fail to reject H_0 with the F test as well.

- **3.1.2** $H_0 : \beta_2 = 1$ and $H_1 : \beta_2 \neq 1$

- We hypothesize that in the population β_2 is actually 1. We estimate under H_1 and should get the output as above. To test the H_0 we have two options:

- t ratio: $t_{\hat{\beta}_2} = \frac{1.26-1}{.25} = 1.04$ (check the formula in the previous section if this is confusing). To find the percentile of $t_{\hat{\beta}_2} = 1.04$ use `ttail(n-1-k, 1.04)` where $n - 1 - k$ in this regression is 17. Executing `ttail(17,1.04)` gives you .157. This is the probability that the t ratio is bigger than 1.04. The t ratio is thus in the $(100 - 16)^{th}$ percentile. We fail to reject H_0 .

- F test: Execute `test heart=1` to get:

```
( 1) heart = 1
      F( 1, 17) = 1.07
      Prob > F = 0.3146
```

we have that F is in the $(100 - 31)^{th}$ percentile so we fail to reject H_0 . Note that the F will not distinguish between rejections of H_0 for very high or very low t ratios - it combines the probability of rejection.

- **3.1.3** $H_0 : \beta_2 = \beta_3$ and $H_1 : \beta_2 \neq \beta_3$

- This hypothesis is more interesting as it says that after controlling for the average consumption of alcohol and average heart disease the effect of average liver disease on average deaths is the same as the effect of average heart disease on average deaths after controlling for alcohol and liver.
- In economics lingo we are testing whether heart and liver disease are perfect substitutes in their effect on average deaths.
- Note that this is a single restriction hypothesis. We can test it using the F test: `test heart=liver` to get

```
( 1) heart - liver = 0
      F( 1, 17) = 2.63
      Prob > F = 0.1235
```

The F is in the $(100 - 13)^{th}$ percentile so we fail to reject H_0 . There is some probability that liver and heart disease are perfect substitutes with respect to average deaths.

- A more interesting way to test this hypothesis is to use our economics background and to reason the following way: Let us believe that H_0 is indeed true and that heart and liver are perfect substitutes. This means that what impacts deaths is the total of heart and liver. We therefore generate a new variable equal to heart plus liver. Call it `hl`.
- Under H_0 we believe the model is

$$deaths_i = \alpha + \beta_1 \times alcohol_i + \beta \times hl_i + u_i \quad (4)$$

- To test this we estimate allowing H_1 : if both liver and heart are important separately then not just the sum will be important but also the quantity of either one:

$$deaths_i = \alpha + \beta_1 \times alcohol_i + \gamma \times liver_i + \beta \times hl_i + u_i \quad (5)$$

To understand this we need to remember the correct interpretation of γ : after controlling for the sum of liver and heart as well as alcohol, liver impacts deaths by γ . If γ is not significant then liver and heart are perfect substitutes.

- Estimating this regression we get

deaths	Coef.	Std. Err.	t	P> t	[95% Conf. Interval]
alcohol	-5.619184	8.716553	-0.64	0.528	-24.0095 12.77113
liver	2.684779	1.656735	1.62	0.124	-.8106259 6.180183
hl	1.259491	.2504379	5.03	0.000	.7311134 1.787869
_cons	532.1875	69.85662	7.62	0.000	384.803 679.5721

You can see that $t_{\hat{\gamma}} = 1.62$ is in the $(100 - 6)^{th}$ percentile in agreement with the F test above. Since we fail to reject $H_0 : \gamma = 0$ we have evidence that heart and liver are perfect substitutes.

3.2 Multiple Restrictions

- Note that all three variables in our original specification are highly correlated:

	alcohol	heart	liver
alcohol	1.0000		
heart	-0.6613	1.0000	
liver	0.7372	-0.6129	1.0000

In particular alcohol is highly correlated with liver. A standard result about multicollinearity states that the coefficient of an independent variable that is highly correlated with the other independent variables in the model will be imprecisely estimated. We already noticed this when we tested $H_0 : \beta_{alcohol} = 0$.

- In a situation like this we may not want to just exclude alcohol when the single-restriction test failed to reject H_0 exactly because of the high correlation with the rest of the independent variables.
- What we do is to test whether the GROUP of highly correlated independent variables are statistically significant.
- A group test requires the framework of a multiple-restrictions test.

- **3.2.1** $H_0 : \beta_1 = 0$ **AND** $\beta_2 = 0$. $H_1 : \text{All the other cases}$

- This test can be implemented the following way:

```
test (alcohol=0) (liver=0)
( 1) alcohol = 0
( 2) liver = 0

F( 2, 17) = 3.00
Prob > F = 0.0766
```

- Note that the syntax is similar to the single-restriction test. The only difference is that when we have more than one restriction we have to contain each restriction in parentheses.
- As you can see the F is in $(100 - 8)^{th}$ percentile so we reject H_0 at 10%. This implies that liver and alcohol are jointly significant even though alcohol is not significant by itself.

- **3.2.2** $H_0 : \beta_1 = 0$ **AND** $\beta_2 = \beta_3$. $H_1 : \text{All the other cases}$

- This hypothesis extends our analysis in the single-restriction case. Since we have noted that alcohol is highly correlated with liver we may want to include the restriction that alcohol is not significant while liver and heart and perfect substitutes with respect to deaths.
- The test is implemented the following way:

```
test (alcohol=0) (liver=heart)
( 1) alcohol = 0
( 2) - heart + liver = 0

F( 2, 17) = 1.44
Prob > F = 0.2654
```

- As you can see the F is in $(100 - 27)^{th}$ percentile so we fail to reject H_0 . The sample does not give us enough evidence that alcohol is important in determining average deaths once we have controlled for average heart and liver and restricted them to have the same impact on deaths.