Chapter 5: Confidence Interval Estimation and Hypothesis Testing

Lecture 2

1 HYPOTHESIS TESTING

1.1 Test Statistics

We first hypothesize that the true parameter takes a particular numerical value, e.g., $\beta_1 = 0$ in a regression with $K = 1$. Our task now is to "test" this hypothesis.

The *null hypothesis*

$$H_0 : \beta_1 = 0. \tag{1}$$

For this null hypothesis, three possible forms of an *alternative hypothesis*

$$H_A : \beta_1 \neq 0, \tag{2}$$

which is used for a *two-sided text*.

$$H_A : \beta_1 > 0, \tag{3}$$

which is used for a *one-sided test*.

$$H_A : \beta_1 < 0, \tag{4}$$

which is also used for a one-sided test.

Example 1: Consider the model- $HS = \beta_0 + \beta_1 GDP$, where HS denotes Housing Starts. A simple economic theory might claim that $\beta_1 > 0$, and we might be interested in testing this theory. The alternative hypothesis is then given by Equation (3).

Sometimes, the null hypothesis may not take the form that the parameter is zero. Then the alternative hypothesis changes accordingly.

Example 2: Suppose a pharmaceuticals company claims that they have discovered a drug with a higher curing rate, r (measured in percentage), than the existing drugs which cure at a rate of 40%. Then the null hypothesis is

$$H_0 : r = 40 \tag{5}$$

The alternative hypothesis is written as

$$H_1 : r > 40. \tag{6}$$

1.2 Test Statistics

A test statistic is a random variable we use to test the null hypothesis against an alternative hypothesis.

$$t = \frac{\hat{\beta}_1 - \beta_1}{SE(\hat{\beta}_1)}, \tag{7}$$

follows the t distribution with $(N-2)$ degrees of freedom. Compute the t value under $H_0 : \beta_1 = 0$:

$$t = \frac{\hat{\beta}_1}{SE(\hat{\beta}_1)}, \tag{8}$$

follows the t distribution with $(N-2)$ degrees of freedom when the null hypothesis is true. This test statistic is called the t-*statistic*, and the test based on this test statistic is called the t -*test*.

If H_0 is true, we expect a small number for $|t|$. While if $H_0$is not true, we expect a large number for $|t|$. It makes sense to choose a number t_C, and adopt the decision rule of the form

Reject H_0 if $|t| > t_C$

Accept H_0 if $|t| \leq t_C$.

Here t_C is called the *critical value*.

If the t-value falls into the *critical region* of $|t| > t_C$, then we reject the null hypothesis. If the t value falls into the *acceptance region* of $|t| \leq t_C$, then we accept the null hypothesis.

1.3 Type I and Type II Errors

Type I error: the error of rejecting a hypothesis when it is true.

Type II error: the error of accepting a false hypothesis.

Ideally, we would like to minimize both of these errors.

1.4 Choosing a Level of Significance

The *level of significance* indicates the probability of observing an estimated t-value. We first choose the level of significance. Then we choose the level of significance. Usually, it will be either 10%, 5%, and 1$. Once the level confidence is chosen, we find the critical value, t_α, from a table for the t distribution such as Table B-1.

1.5 Examples

Example (a two-tailed test):

For example, imagine that you have run OLS for the cost of another ice cream stand with the dependent variable of the cost and the independent variable of the number of the ice cream bars. Suppose that you obtained

$$\hat{Y}_i = 29.58 + \underset{(0.15)}{0.48} X_i$$

$$N = 25, R^2 = 0.79$$

Here, the standard error is in the parenthesis. We first compute the t value for the null hypothesis $H_0 : \beta_1 = 0$ from Equation (8).

$$t = \frac{0.48}{0.18} = 2.67 \tag{9}$$

We have 25 observations, hence the degrees of freedom are $(25 - 2) = 23$. For this example, we choose the level of significance to be 1%. Using Table B-1, we find that

$$P[|t| > 2.807] = 0.01. \quad (10)$$

So the critical value, t_C is 2.807.

Because the computed $|t|$ value of 2.67 is less than the listed critical t values, we accept the null hypothesis H_0 at the 0.01 level of significance.

Example (a one-tailed test): In our example of the cost of an ice cream stand, the slope coefficients expected to be positive. Hence it makes sense to test the null hypothesis of $H_0 : \beta_1 = 0$ against the alternative hypothesis of $H_A : \beta_1 > 0$; here the alternative hypothesis is one sided.

The t-testing procedure remains exactly the same as before, except that we use different critical values and only reject the null hypothesis when the t value has the correct sign.

In the example of the ice cream stand, we have 25 observations, hence the degrees of freedom are $(25 - 2) = 23$. For this example, we choose the level of significance to be 1%. Using Table B-1, we find that

$$P[t > 2.508] = 0.01. \tag{11}$$

So the critical value, t_C is 2.508. Because the computed $|t|$ value of 2.67 exceeds the listed critical t values, we reject the null hypothesis H_0 at the 0.01 level of significance by the one-tailed test.

Compare the these two examples, and note that we can reject the null hypothesis more frequently with the one-tailed test than with the two-tailed test because the critical values are smaller with the one-tailed test.

1.6 One Tailed Tests

One tailed tests are more complicated than two tailed tests. This section explains the procedure for a one tailed test for the null hypothesis, $H_0 : \beta_1 = 0$.

Step 1: Using economic theory and/or common sense, determine which alternative hypothesis makes sense $H_A : \beta_1 > 0$ or $H_A : \beta_1 < 0$. One tailed tests are used to see whether or not the data support the appropriate alternative hypothesis. It is useful to remember that

(a) We reject the null hypothesis (in favor of the alternative hypothesis) if the data support the alternative hypothesis.

(b) We accept the null hypothesis if the data do not support the alternative hypothesis.

Step 2: Examine the sign of the point estimate, $\hat{\beta}_1$.

(a) If the sign is incorrect for the alternative hypothesis, then accept the null hypothesis (even if $|\hat{\beta}_1|$ is large).

(b) If the sign is correct, then go to Step 3.

Step 3: Calculate the t-value, using the point estimate, $\hat{\beta}_1$; its standard error, $SE(\hat{\beta}_1)$; and the following formula:

$$t = \frac{\hat{\beta}_1}{SE(\hat{\beta}_1)}, \tag{12}$$

Step 4: Use Table $B - 1$ to get the critical value. Make sure to use the level of significance listed in the first row.

Step 5: If the absolute value of the t value is smaller than the critical value, then accept the null hypothesis. If not, then reject the null hypothesis.

Exercise 1: Imagine that you are selling an imported good and wish to convince your boss that you can sell more if you are allowed to lower its price. You have run a regression of Y (demand for the good) onto X (the price), and obtained the following results:

$$\hat{Y}_i = 261.091 - \underset{(0.2015)}{0.3452} \quad X_i$$
$$N = 16, R^2 = 0.838792$$

(A) Which alternative hypothesis is appropriate, $H_A : \beta_1 > 0$ or $H_A : \beta_1 < 0$?

(B) Do you reject the null hypothesis, $H_0 : \beta_1 = 0$ when you use one tailed test at the 5% significance level?

(C) Do you reject the null hypothesis, $H_0 : \beta_1 = 0$ when you use two tailed test at the 5% significance level?

(D) Imagine that the point estimate is 12.356 instead of -0.345. Do you reject null hypothesis when you use one tailed test?

Exercise 2: Imagine that you are selling an imported good and wish to convince your boss that you can sell more if you are allowed to contact richer consumers. You have run a regression of Y (demand for the good) onto X (income), and obtained the following results:

$$\begin{aligned} \hat{Y}_i = 265.093 + \quad & 0.5452 \quad X_i \\ & (0.3405) \\ N = 18, R^2 = \quad & 0.738736 \end{aligned}$$

(A) Which alternative hypothesis is appropriate, $H_A : \beta_1 > 0$ or $H_A : \beta_1 < 0$?

(B) Do you reject the null hypothesis, $H_0 : \beta_1 = 0$ when you use one tailed test at the 5% significance level?

(C) Do you reject the null hypothesis, $H_0 : \beta_1 = 0$ when you use two tailed test at the 5% significance level?

(D) Imagine that the point estimate is -32.356 instead of 0.545. Do you reject null hypothesis when you use one tailed test?