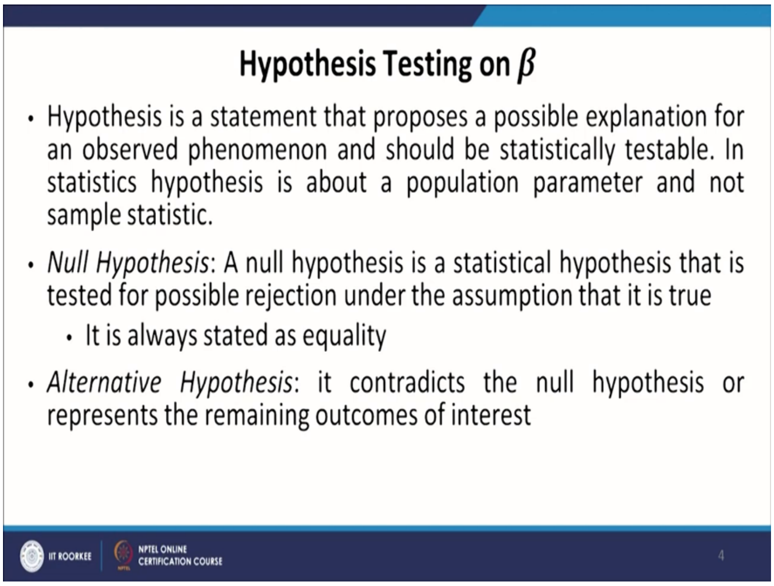
Econometric Modelling
Professor Sujata Kar
Department of Management Studies
Indian Institute of Technology Roorkee
Lecture 10
Hypothesis Testing

Hello and this is the tenth module of the course on Econometric Modelling. This is the last module of the second part which is an overview of a classical linear regression model. So here we are going to discuss hypothesis testing.

(Refer Slide Time: 00:42)

Hypothesis testing is basically a, we are not going to get into any testing procedure right away. As you can see that the testing procedures are probably discussed at length in part 4 where we are getting into statistical inferences in a big way. So, I am here right now just going to introduce you to the concept of hypothesis testing that is what is hypothesis, how the testing is generally done without talking about any specific distribution or testing procedure.

(Refer Slide Time: 01:24)

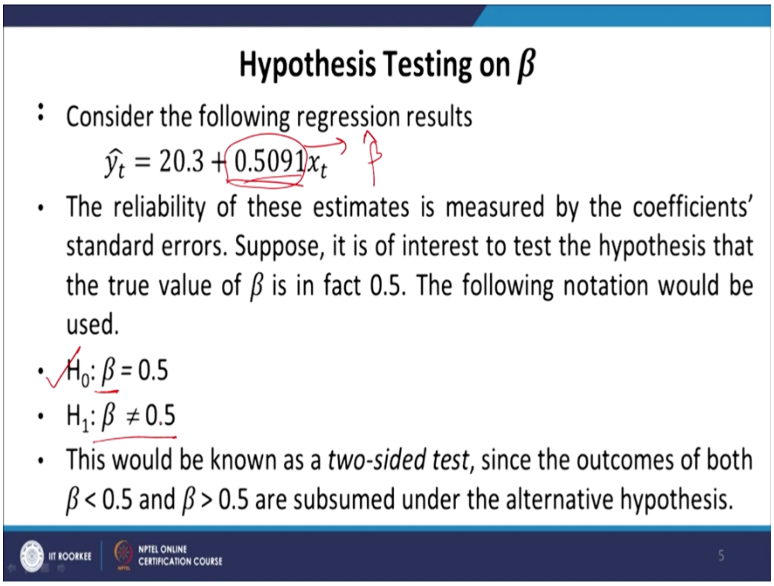


So, we begin with hypothesis testing on β. It can be α *or* β. Most often we talk about β only. A hypothesis is a statement that proposes a possible explanation for an observed phenomenon and should be statistically testable. In statistics, the hypothesis is about a population parameter and never for sample statistic. So, whenever we form the hypothesis, they are with respect to the population parameter α *or* β.

We have two types of hypotheses when we go for the testing of hypotheses. One is the null hypothesis. A null hypothesis is a statistical hypothesis that is tested for possible rejection under the assumption that it is true. So, since null implies void so most often the way null hypotheses are formed, should ideally be rejected. It is always stated as equality.

And we have the second type of hypothesis while testing for the hypothesis. And that is the alternative hypothesis. It contradicts the null hypothesis or represents the remaining outcomes of interest. So null hypothesis is the void hypothesis which we generally expect to reject; while the alternative hypothesis is probably the alternative that we are expecting to come out.

(Refer Slide Time: 02:54)

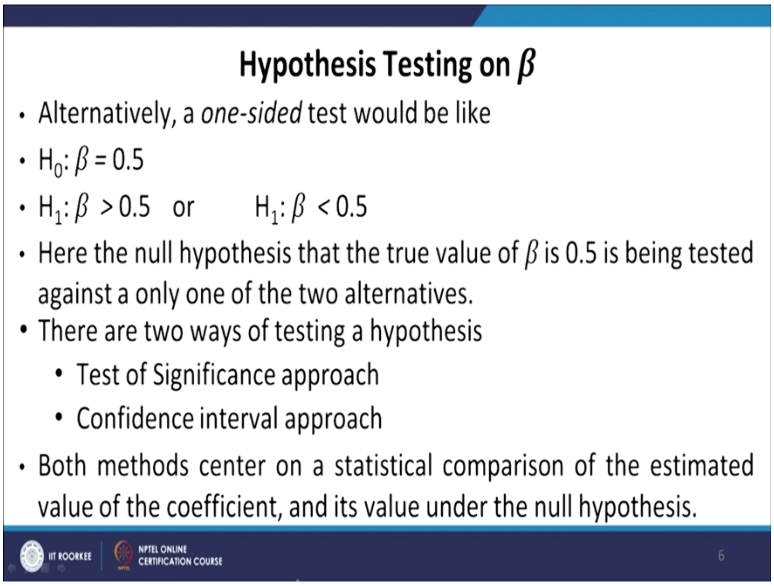


Consider the following regression results, that is, (refer slide time 2:59). So, so far this is basically the estimated part of the observations on y that we have discussed. Now we did not talk about how good this estimated value of 0.5091 is. So, this is a time when we actually just introducing you that how we can test, whether how good this estimated parameter is. The reliability of these estimates is measured by the coefficients' standard errors.

Suppose it is of interest to test the hypothesis that the true value of beta is in fact 0.5. The following notation would be used. This is how we write a null hypothesis (refer slide time 3:55). So, though we are trying to test whether we are working with the estimated parameter that is 0.5091 which is actually denoted by $\hat{\beta}$, the null hypothesis is for the population parameter and not $\hat{\beta}$, its estimated counterpart.

And H1 is the alternative hypothesis. And this is beta not equal to 0.5. This would be known as a two-sided test since the outcomes of both beta less than 0.5 and beta greater than 0.5 are subsumed under the alternative hypothesis. So not equal to 0.5 simply implies that it could be greater than 0.5 or less than 0.5. Both the possibilities are kept. And that is why it is called the two-sided hypothesis or two-sided test.

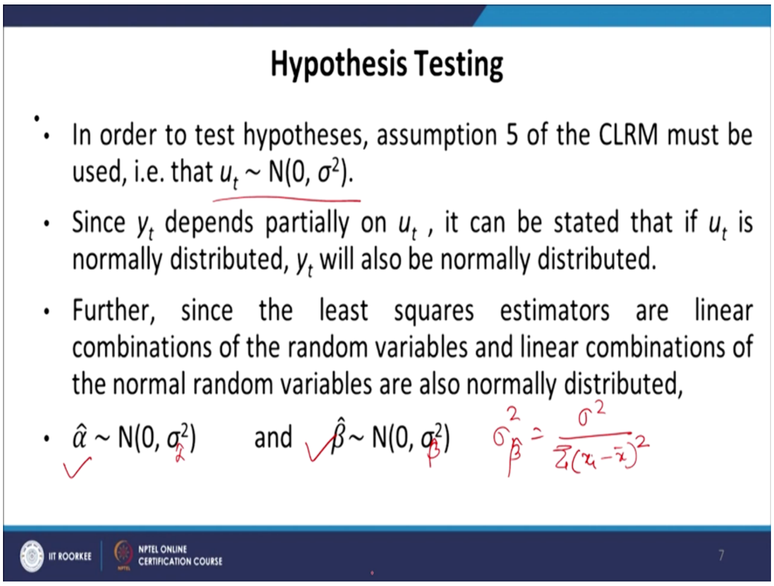
(Refer Slide Time: 04:54)



Alternatively, a one-sided test would be like (refer slide time 5:00). So null hypothesis is always in terms of equality, while the alternative hypothesis is either beta equals beta greater than 0.5 or beta less than 0.5. Here the null hypothesis that the true value of beta is 0.5 is being tested against only one of the two alternatives.

There are two ways of testing a hypothesis that is two alternative approaches. They lead to the same results; test of significance approach and confidence interval approach. Both methods center on a statistical comparison of the estimated value of the coefficient and its value under the null hypothesis. So, it compares the estimated value of the coefficient with a statistical value that is given by the distribution or whatever testing procedures we follow.

(Refer Slide Time: 06:00)



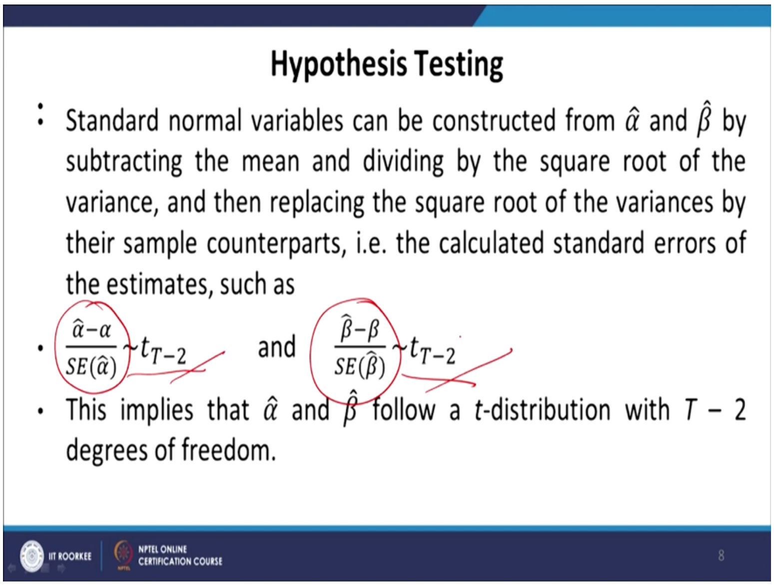
In order to test the hypothesis, assumption 5 of the CLRM that is a Classical linear regression model must be used, that is u_t follows a normal distribution with zero mean and constant variance equals to sigma square. Now, this is actually not something we are going to readily use. But to begin with, we can follow this assumption. Since y_t depends partially on u_t, it can be stated that if u_t is normally distributed y_t will also be normally distributed. And if you remember $\hat{\alpha}$ *and* $\hat{\beta}$, or α, β are population parameters. $\hat{\alpha}$ *and* $\hat{\beta}$ are estimated. They can vary from sample to sample.

But from one sample or for one particular sample x_t is a variable. We can consider it to be non-random once the sample has been collected. So u_t follows a normal distribution. And that is how y_t would follow a normal distribution. Further, since the least square estimators are linear combinations of the random variables and linear combinations of the normal random variables are also normally distributed, we have $\hat{\alpha}$ *and* $\hat{\beta}$ also following a normal distribution with mean and variance 0 and sigma square.

This is actually (refer slide time 7:35) because we know that the variance of an alpha hat and beta hat is not exactly sigma square. We can denote them by sigma square alpha hat and sigma square beta hat; while the (refer slide time 7:50). And similarly, we will have an expression for sigma square alpha hat as well.

So we, for the time being, formally assumed that $\hat{\alpha}$ *and* $\hat{\beta}$ also follow a normal distribution. And that is because we have a large number of or we can have a large number of $\hat{\alpha}$ *and* $\hat{\beta}$ when we go for repeated samples from the same population.

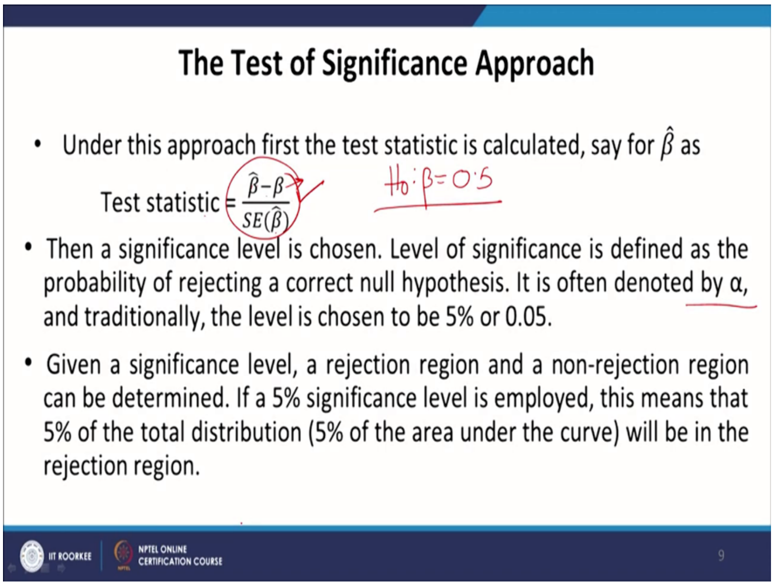
(Refer Slide Time: 08:21)



Now standard normal variables can be constructed from $\hat{\alpha}$ *and* $\hat{\beta}$ by subtracting the mean and dividing by the square root of the variance and then replacing the square root of the variances with the sample counterparts that is the calculated standard errors of the estimates. So, what we obtain here is this; that is, (refer slide time 8:44).

Now, these two follow t-distribution with T -2 degrees of freedom. Why do they follow t-distribution? That we would discuss at length while discussing the t-test in part 4 while we deal with statistical inference at length. As such for the time being, we are also not going to use these assumptions or these expressions much. Just to begin with I am introducing this concept. So this implies that $\hat{\alpha}$ *and* $\hat{\beta}$ follow a t-distribution with T -2 degrees of freedom.

(Refer Slide Time: 09:26)

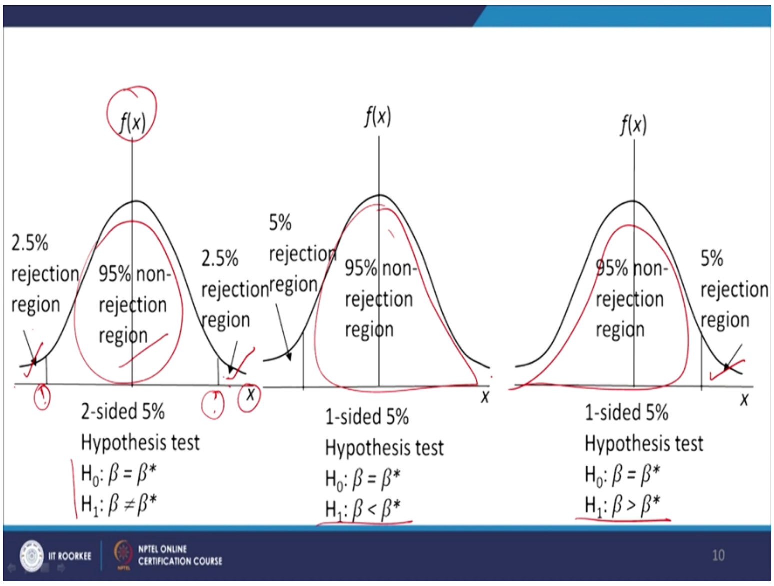


So, under this approach first, the test statistic is calculated. Say for (refer slide time 9:36). Now you would say that we did not know the value of beta. So, for that, we have the null hypothesis. So, the null hypothesis is always about the beta. So, if the null hypothesis is beta is equaled to 0 then my test statistic would be (refer slide time 9:55).

If the null hypothesis is (refer slide time 10:03). So, we can always construct a test statistic given the null hypothesis or under the null hypothesis. Then a significance level is chosen. The level of significance is defined as the probability of rejecting a correct null hypothesis. It is often denoted by α and not to be confused with the constant term alpha in a regression model.

So, significance or level of significance is also denoted by alpha. And traditionally the level is chosen to be 5 percent or 0.05 percent. But one can go for other significance levels also. The other two most commonly used possibilities are 1 percent and 10 percent. Now given a significance level, a rejection region and a non-rejection region can be determined. If a 5 percent significance level is employed this means that 5 percent of the total distribution, that is 5 percent of the area under the distribution curve will be in the rejection region. And if my test statistic falls in that region, we will be rejecting the null hypothesis, otherwise, we will not.

(Refer Slide Time: 11:43)



So, this is actually explained using a diagram. So first we explain the two-sided hypothesis. So, we are measuring x on the horizontal axis and the frequency on the vertical axis, the way a distribution curve is generally drawn. And it may follow a normal distribution. It may follow t-distribution. I am generally going for a nice bell-shaped curve for the time being.

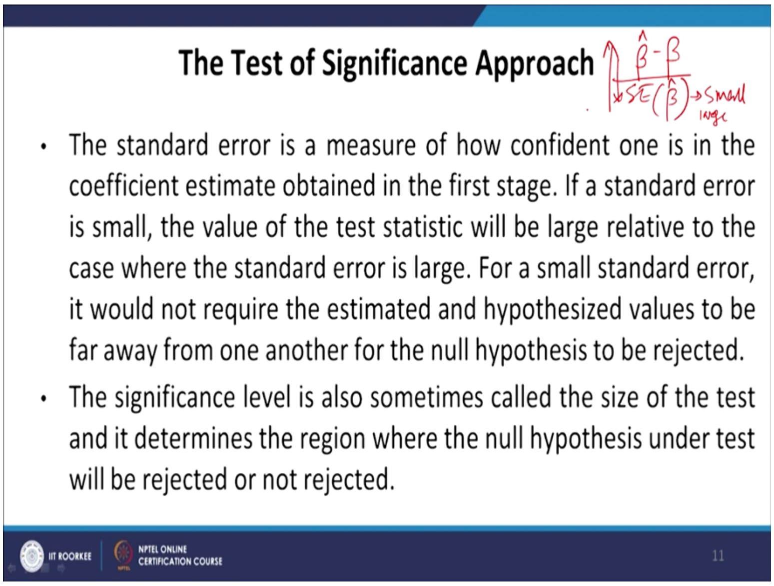
So, if it is a two-sided test at a 5 percent level of significance then what we do is that we divide 5 percent for both sides. So, we have 2.5 percent on each side. So, 2.5 percent rejection region and this is another 2.5 percent rejection region and in the middle, we have 95 percent non-rejection region. And I calculate the test statistic based on the null hypothesis or under the null hypothesis. If my test statistic falls in either of these regions then I will reject the null hypothesis, otherwise, I will not reject the null hypothesis. So, we call it the non-rejection region.

Now how do I understand whether it falls here? Of course, I will be having a critical value corresponding to these points, and then if my calculated test statistic is larger than either of these two critical values; that is smaller than this critical value and larger than this critical value, that means my test statistic is falling either in this region or in this region.

Now we go for one-sided 5 percent hypothesis testing, that is, again the significance level is 5 percent. But this is a one-sided test where the alternative hypothesis is beta is less than a beta star. So, we have here only 5 percent rejection region that is on the left, and 95 percent rejection region consists of this entire region.

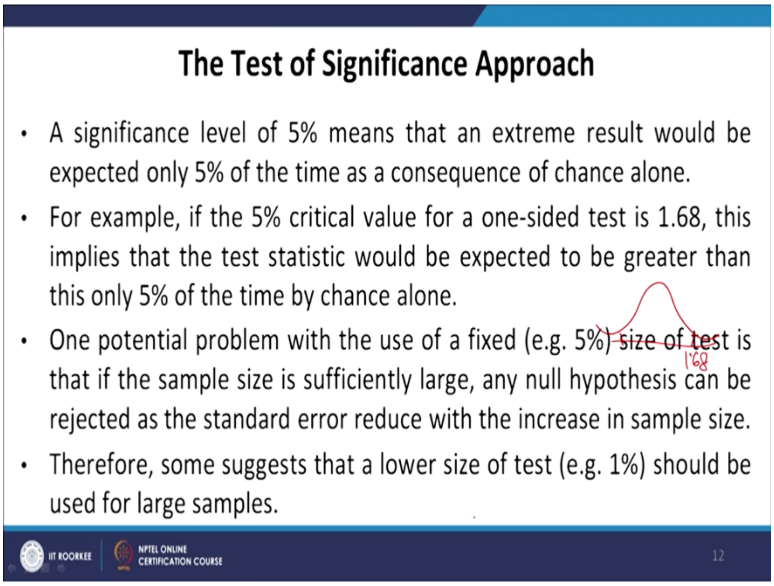
Similarly, we again go for another one-sided 5 percent significance level hypothesis test where the alternative hypothesis is beta is greater than a beta star. So of course, we will be considering the right-hand tail; and this is my 5 percent rejection region and the rest is actually 95 percent non-rejection region. So graphically this is what we try to do.

(Refer Slide Time: 14:21)



So, the standard error is a measure of how confident one is in the coefficient estimate obtained in the first stage. If a standard error is small the value of the test statistic will be large relative to the case where the standard error is large. You remember while we calculated test statistics, we have (refer slide time 14:44). So that is what is being mentioned here. That for a small standard error it would not require the estimated and hypothesized values to be far away from one another for the null hypothesis to be rejected. The significance level is also sometimes called the size of the test and it determines the region where the null hypothesis under test will be rejected or not rejected.

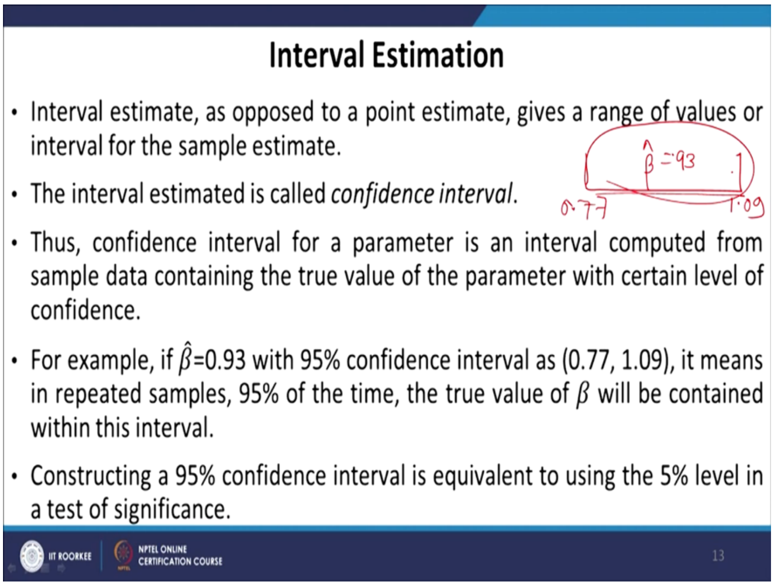
(Refer Slide Time: 15:33)



A significance level of 5 percent means that an extreme result would be expected only 5 percent of the time as a consequence of chance alone. For example, if the 5 percent critical value for a one-sided test is 1.68 percent. So, I mentioned that in this bell-shaped curve there will be critical values. So, suppose that critical value is 1.68 here. This implies that the test statistic would be expected to be greater than this value, that is, 1.68, only 5 percent of the time by chance alone.

One potential problem with the use of a fixed, for example, 5 percent size of the test is that if the sample size is sufficiently large, and null hypothesis can be rejected, as the standard error reduces with the increase in sample size. Therefore, some suggest that a lower size of the test, for example, 1 percent should be used for large samples. So, for large samples, we would recommend going for a lower sample size for the reasons explained here.

(Refer Slide Time: 16:45)

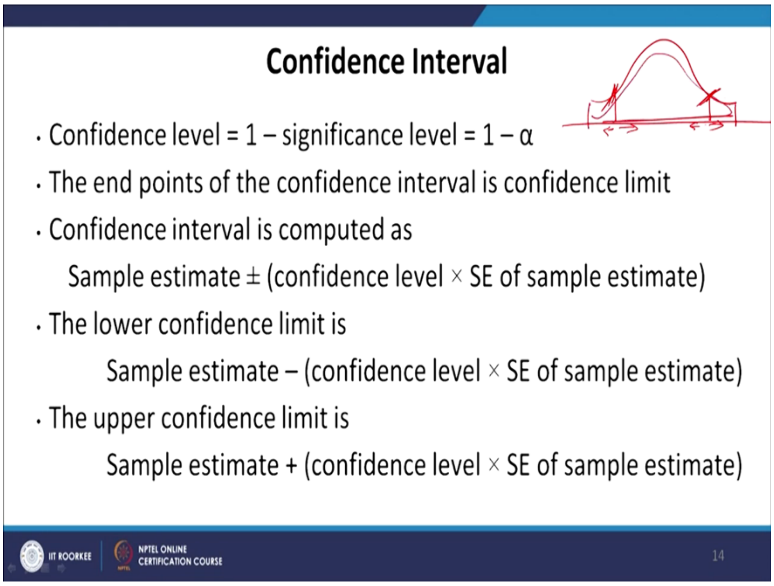


Now coming to the interval estimation. So, so far, we were talking about the significance level approach. Now we are getting into the confidence interval approach. An Interval estimate as opposed to a point estimate gives a range of values or intervals for the sample estimate. The interval estimated is called the confidence interval. Thus, the confidence interval for a parameter is an interval computed from sample data containing the true value of the parameter with a certain level of confidence.

For example, (refer slide time 17:26). Then it means in repeated samples 95 percent of the time the true value of β will be contained within this interval. So, it is actually the long, it is between this interval, this is actually close to the true value.

Constructing a 95 percent confidence interval is equivalent to using the 5 percent level in the test of significance. So, they are basically the same concept. 5 percent significance level tells us the possibility of rejecting a correct null hypothesis while 95 percent confidence interval tells us that we can say with 95 percent confidence that our estimated or the true parameter will lie within this interval.

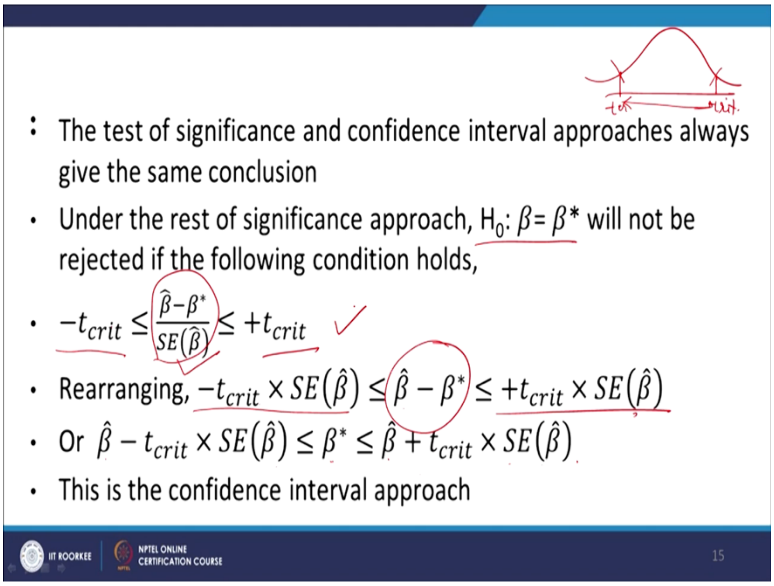
(Refer Slide Time: 18:48)



So, the confidence interval is actually equal to 1 minus significance level or 1 minus α. The endpoints of the confidence interval are the confidence limit. So, when we are actually drawing the lines for critical values; or this is basically the beginning of the significance level and also the beginning of the confidence interval. So, confidence interval this way; significance levels this way. And these are called confidence limits. This is a lower confidence limit. This is the upper confidence limit.

Confidence interval is computed as sample estimate plus-minus confidence level, that is say, 95 percent, multiplied by the standard error of sample estimates. So, the lower confidence limit, that is this, is sample estimate minus confidence level multiplied by the standard error of sample estimate and the upper confidence limit is sample estimate plus confidence level multiplied by the standard error of sample estimate. So, this is the upper confidence limit. If I reduce my significance level to 1 percent, then my confidence interval would be 99 percent. So, as I reduce the significance level my confidence interval grows.

(Refer Slide Time: 20:08)



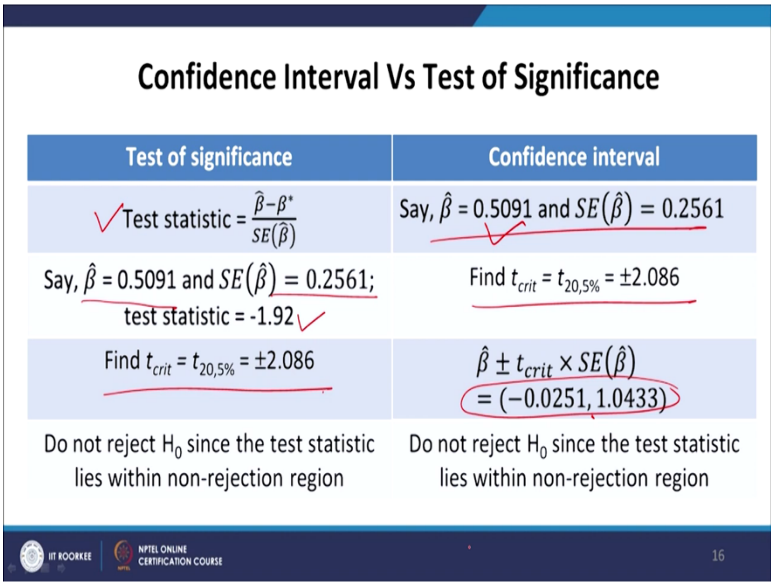
The test of significance and confidence interval approaches always give the same conclusion. Under the test of significance approach (refer slide time 20:19) will not be rejected if the following condition holds. So, we are trying to draw an analogy between the significance approach and the confidence interval approach.

So, we write that, we generally write that t critical, as I am trying to explain to you that I have my t critical here, if it follows the T statistic then I have t critical value here and t critical value here. Note that, when I am drawing the t critical values, they are actually equivalent to lower confidence limit and upper confidence limit.

So, what we say under, level of significance approach is that if my calculated test statistic which is (refer slide time 21:11), if it lies within this region then I am not going to reject the null hypothesis. So minus t critical greater than equal to the test statistic less than equal to plus t critical. So, for this the null hypothesis, beta equals beta star will not be rejected.

Now if I rearrange terms then I have (refer slide time 21:44 – 22:15). This is the confidence interval approach. So, we can deduce the level of significance approach or from the level of significance approach the outcome of the confidence interval approach. So that is why we say that they both give the same results.

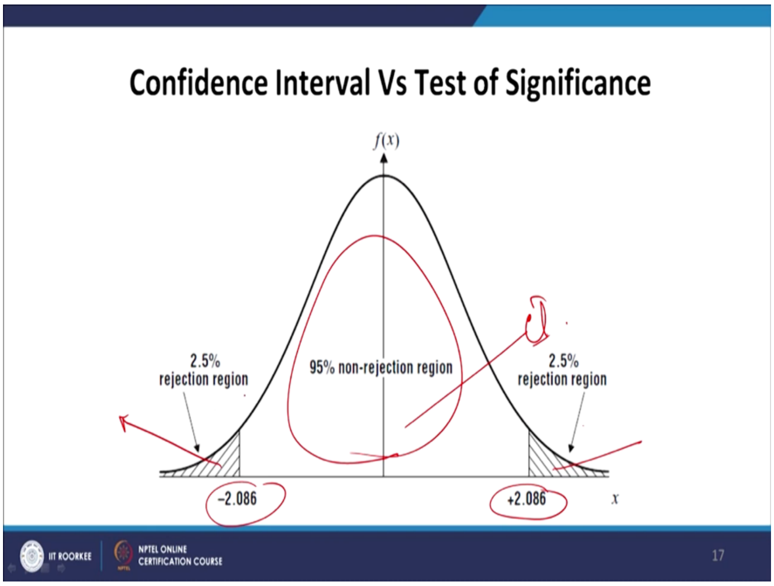
(Refer Slide Time: 22:30)



So, this is basically a table that compares the test of significance and confidence interval approaches. So, this is how we first calculate the test statistic, (refer slide time 22:45). So, we do not reject the null hypothesis since the test statistic lies within the non-rejection region.

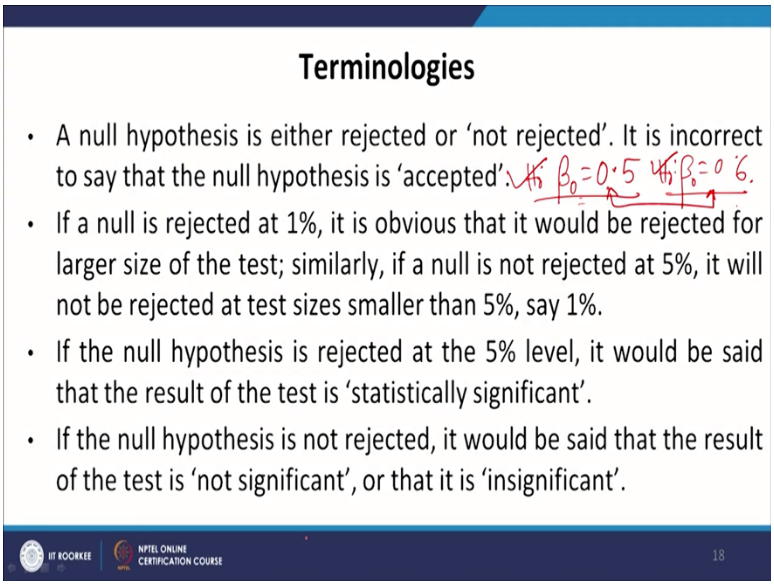
Now for the confidence interval, we again continue with the same example. We find t critical which are given here. And then we calculate the confidence interval as -0.0251 to 1.0433. And since my $\hat{\beta}$ hat lies within this interval, so we do not reject the null hypothesis since the test statistic lies within the non-rejection region. It basically leads to the same outcome.

(Refer Slide Time: 23:41)



Graphically what we show, is a 95 percent non-rejection region. This diagram I have almost drawn on every slide. And this is my confidence interval. And this is the rejection region of 2.5 percent here and 2.5 percent here. I have also mentioned here the critical values; now which implies that these two add up to their significance level. So again, they give us the same results, confidence interval versus test of significance.

(Refer Slide Time: 24:17)



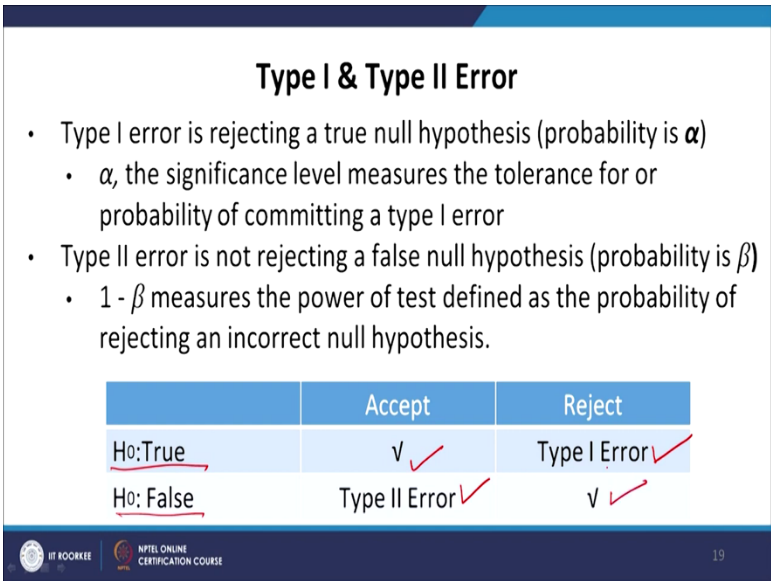
Now I talk about certain terminologies. A null hypothesis is either rejected or not rejected. It is incorrect to say that the null hypothesis is accepted. So, so far you have also seen that it while drawing that bell-shaped curve always mentioned non-rejection region. I never have mentioned the acceptance region. So null hypothesis is accepted is incorrect.

When we do not reject the null hypothesis, it actually does not mean that we accept the null hypothesis. This is also possible that some alternative null hypothesis can also not be rejected. So, between two alternative null hypotheses which one we are going to accept? We basically do not reject it. For example, my null hypothesis, one is 0.5. And the other null hypothesis say is 0.6.

Now if suppose I do not reject any one of them or given the test statistic, I fail to reject any one of them; but both cannot be accepted simultaneously. So that is why we say that we have not rejected these. We do not become very affirmative about the fact that we have accepted them, because we are not sure between these two which one can be accepted. And it is quite possible that both of them are not rejected. If a null hypothesis is rejected at 1 percent, it is obvious that it would be rejected for a larger size of the test.

Similarly, if a null is not rejected at 5 percent, it will not be rejected at-test sizes smaller than 5 percent, say 1 percent. If the null hypothesis is rejected at a 5 percent level it would be said that the result of the test is statistically significant. If the null hypothesis is not rejected it would be said that result of the test is not significant, or that is insignificant. So, these are the terminologies that we use. If we do not reject a null hypothesis, we say that the result of the test is insignificant or not significant and vice versa.

(Refer Slide Time: 26:44)

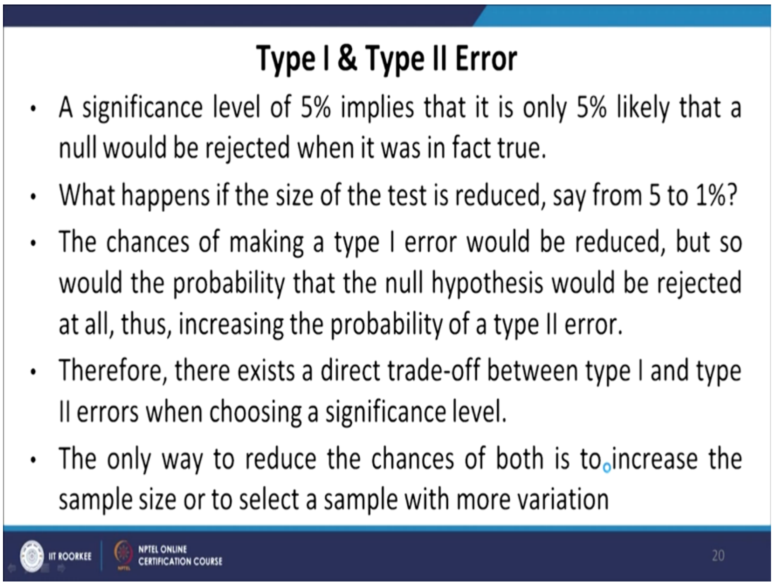


Now we talk about Type 1 and Type 2 errors. Type 1 error is rejecting a true null hypothesis, probability is α. So, the Type 1 error is the probability of rejecting a correct null hypothesis. α , the significance level measures the tolerance for or probability of committing a Type 1 error. So, the moment we specify the significance level at 5 percent this means that the probability of rejecting a correct null hypothesis is only 5 percent which is considered to be pretty low and generally created by or resulted because of certain randomness in the results obtained or data generated.

Type 2 error is not rejecting a false null hypothesis, probability is β. That is Type 2 error measures the probability of not rejecting a null hypothesis that is incorrect or false. So, $1 - \beta$ measures the power of the test defined as the probability of rejecting an incorrect null hypothesis. $1 - \beta$ is the probability of rejecting an incorrect null hypothesis. β is the probability of accepting a false null hypothesis.

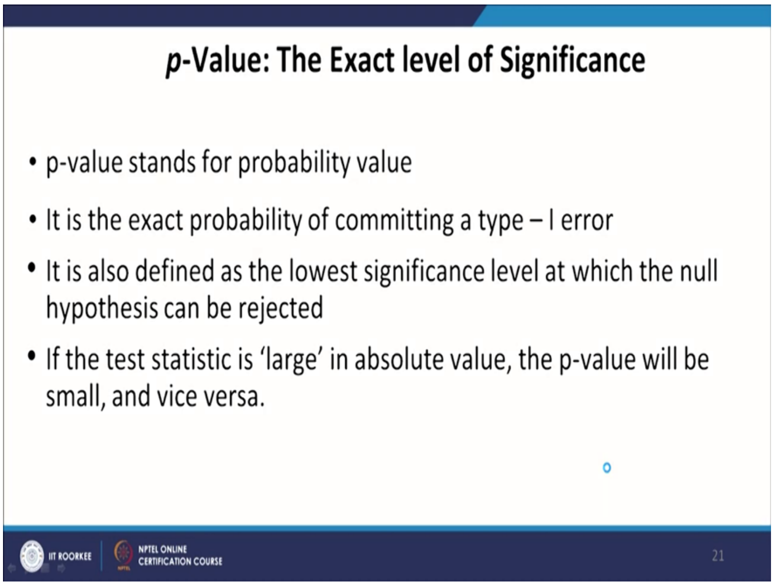
So, this is how we prepare a table. This is a situation where the null hypothesis is true. We are accepting it, not making any mistakes. We are rejecting a correct null hypothesis; we are making a Type 1 error. The null hypothesis is false. So, when we accept a false null hypothesis, we make a Type 2 error. If we reject a false, null hypothesis then there is nothing wrong with it.

(Refer Slide Time: 28:33)



A significance level of 5 percent implies that it is only 5 percent likely that a null would be rejected when it was in fact true. What happens if the size of the test is reduced, say from 5 percent to 1 percent? The chances of making a Type 1 error would be reduced but so would the probability that the null hypothesis would be rejected at all, thus increasing the probability of Type 2 error. So, there is always a tradeoff between Type 1 and Type 2 errors while choosing a significance level. The only way to reduce the chances of both is to increase the sample size or to select the sample with more variation.

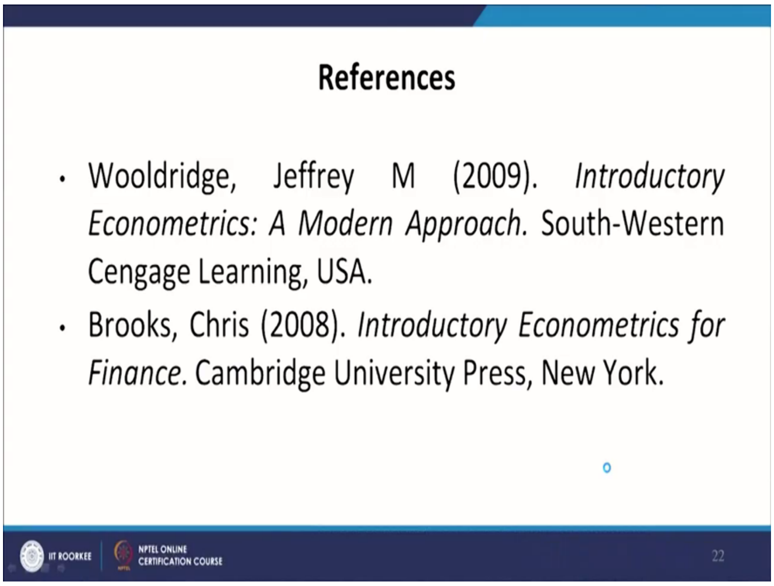
(Refer Slide Time: 29:21)



Now we talk about the final thing that is the p-value or the exact level of significance. p-value stands for probability value. It is the exact probability of committing a Type 1 error, that is, the probability of rejecting a correct null hypothesis. It is also defined as the lowest significance level at which the null hypothesis can be rejected. If the test statistic is large in absolute value, the p-value will be small and vice versa.

So, p-values are most often reported by the software, whatever software you use, you would see that a p-value is reported. So, on the basis of the p-value reported we can determine whether my null is rejected or not rejected, at what significance level 5 percent, 1 percent, 10 percent, and so on. So, p-values are basically the exact probability of committing a Type 1 error. A very small p-value implies the probability of committing a Type 1 error is absolutely or almost 0 and this happens when the sample size is very large.

(Refer Slide Time: 30:34)



So that brings us to the end of module 10 on hypothesis testing. These are the books that I have followed for discussion on these topics. Thank you.