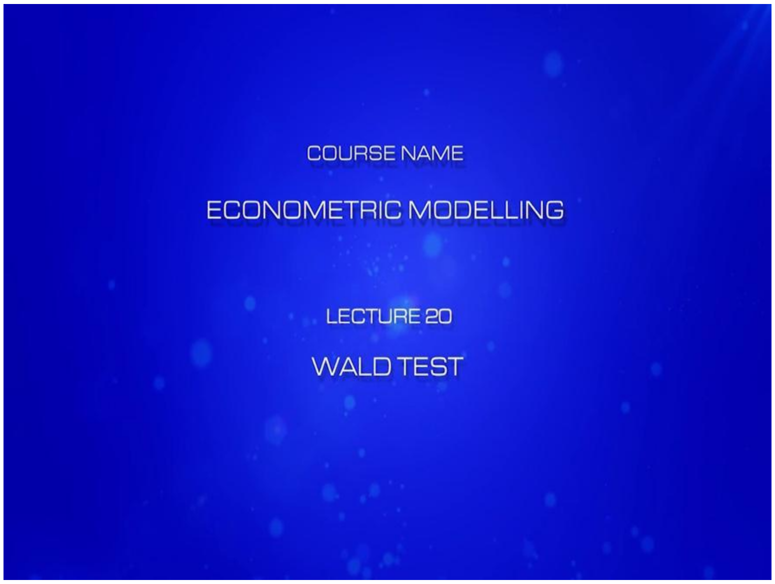
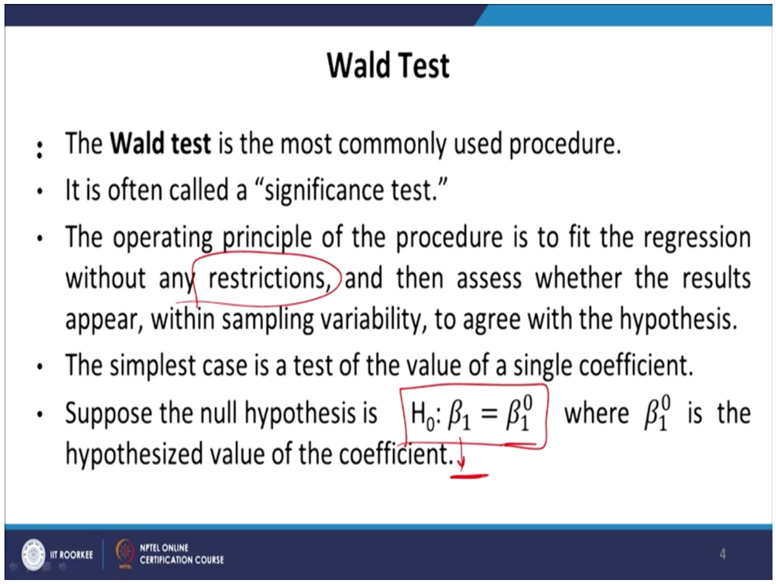
**Econometric Modelling**
**Professor Sujata Kar**
**Department of Management Studies**
**Indian Institute of Technology, Roorkee**
**Lecture 20**
**Wald Test**

(Refer Slide Time: 00:12)




And this is Econometric Modelling, Module 20. Currently, we are discussing statistical inferences. In the previous module, I have discussed the t-test and, in this module, Module 20, I am going to discuss Wald Test.

(Refer Slide Time: 00:37)




Now, the Wald test is the most commonly used procedure though most of us might not be using the term Wald test. Now, why this is so that I will be discussing, probably more in the next module when we introduce other tests.

It is also called the significance test. The operating principle of the procedures is to fit the regression without any restrictions, and then assess whether the results appear within sampling variability to agree with the hypothesis.

Now, one very important characteristic of the Wald test is that though it is the most commonly used test, but it is not designed to incorporate any restriction in any kind of regression.

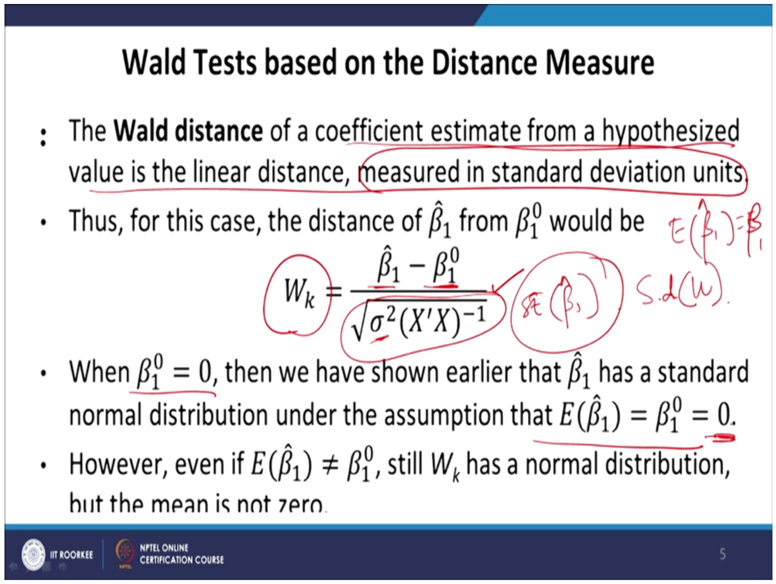
Now, when we talk about restrictions, you may recall that in Module 13, we learned that a certain amount of restrictions could be put on some of the coefficients or parameter estimates, and then how to test for that? Those kinds of analysis are not possible under the Wald test.

Wald test can perform a significance test for a complete model when we do not have any specific restrictions on any of the variables. Now, what does this mean would be even clearer when we discuss the next test that is F-test.

Now, the simplest case is the test of the value of a single coefficient. So, we begin with the Wald test in a simple linear regression format. Suppose the null hypothesis is (refer slide time: 2:17). We are testing for a single hypothesis or restriction. It can be a multiple linear regression, it can be a simple linear regression, but we are testing for a single restriction and that is why it is easy to assume that we are dealing with simple linear regression.

And the value of $\beta_1$, the population parameter is hypothesized to be equal to some value which is $\beta_1^0$.

(Refer Slide Time: 02:51)




The Wald distance of a coefficient estimate from a hypothesized value is the linear distance measured in standard deviation units. So, you see that this is actually very similar to the kind of t ratio we have discussed, but the Wald test is essentially based on a slightly different concept. So, it is based on distance measures.

Thus, for this case, the distance from (refer slide time: 3:20). So, this is the standard error of $\hat{\beta}_1$, and that is why we say that the Wald distance of a coefficient estimate from a hypothesized value is the linear distance measured in standard deviation units.

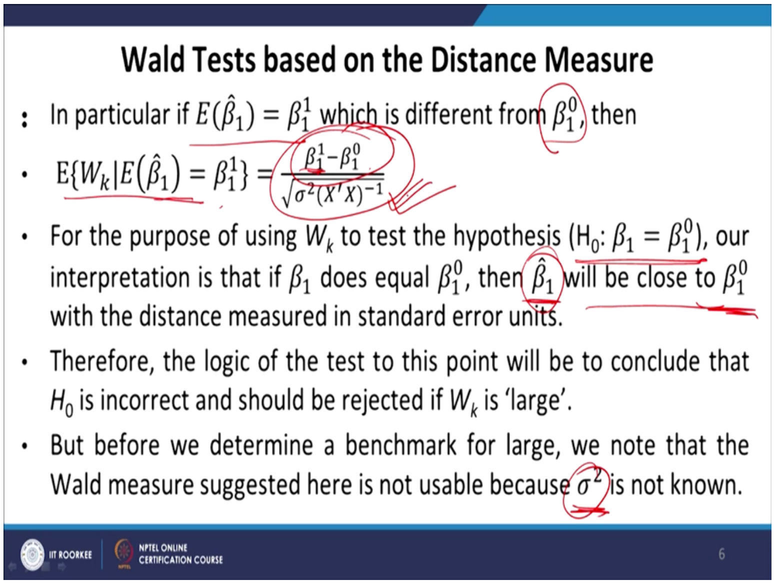
We are measuring it in terms of standard deviation or standard error of the estimated parameter. But yes, this is not the standard error of the estimated parameter because we are using $\sigma^2$ and not $\hat{\sigma}^2$, as a result of which this is actually the standard deviation of the population error term, the entire denominator.

So that is why we say that it is measured in standard deviation units. (Refer slide time: 4:23- 4:48).

Now, (refer slide time: 4:49), still $w_k$, the Wald statistic has a normal distribution, but the mean is not 0.

So, this has a standard normal distribution while if (refer slide time: 5:15), which could be equal to 0, which may not be equal to 0, then $w_k$, still has a normal distribution, but the mean is not 0.

(Refer Slide Time: 05:33)




In particular, (refer slide time: 5:35). Then expected value of $w_k$, that is the Wald statistic, (refer slide time: 5:47), which is the hypothesized value of the population parameter, divided by again, the standard deviation of the population error term.

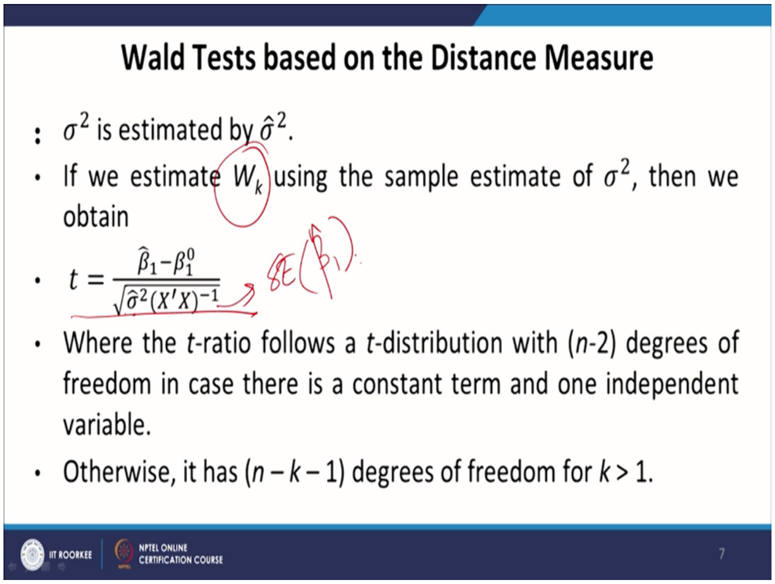
For the purpose of using $w_k$, to test the hypothesis, (refer slide time: 6:15), with the distance measured in standard units.

So, it is quite possible (refer slide time: 6:32) and the distance between them is measured in terms of the standard deviation units.

Therefore, the logic of the test to this point will be to conclude that $H_0$ is incorrect and should be rejected if $w_k$, is large. We go by similar logic that if this turns out to be a large value, then we reject the null hypothesis, and our $H_0$: $\beta_1 = \beta_1^0$.

But before we determine a benchmark for large, we note that the Wald measure suggested here is not usable because sigma square is not known. So, we are using population standard deviation here, but $\sigma^2$ is not known. So, we need to replace it with the sample standard error of the estimated parameter.

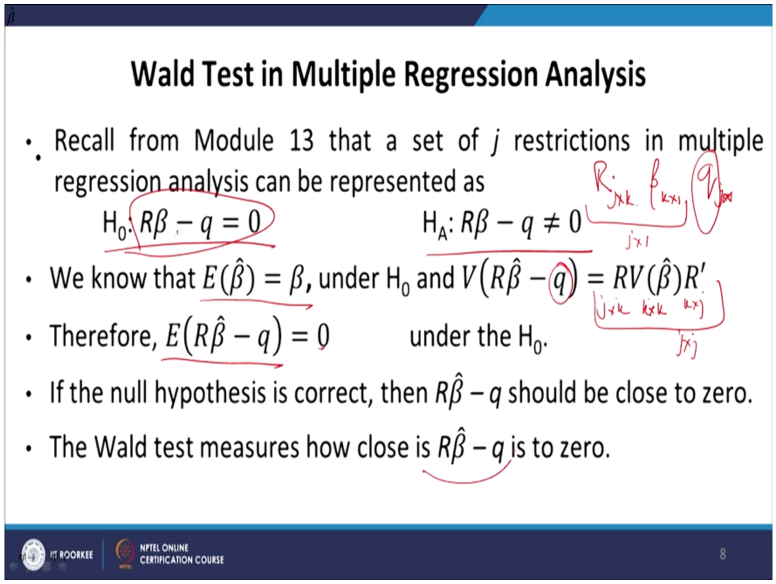
(Refer Slide Time: 07:41)




Now, $\sigma^2$ is estimated by $\hat{\sigma}^2$, which is an unbiased estimator of the population error variance. If we estimate $w_k$ using the sample estimate of $\sigma^2$, then we obtain the t-statistic, the t-ratio, (refer slide time: 8:01).

But the t-ratio follows a t-distribution with $n - 2$ degrees of freedom in case there is a constant term and one independent variable, otherwise, the t-distribution, will have a degree of freedom which is equal to $n - k - 1$, for any k greater than 1.

(Refer Slide Time: 08:45)

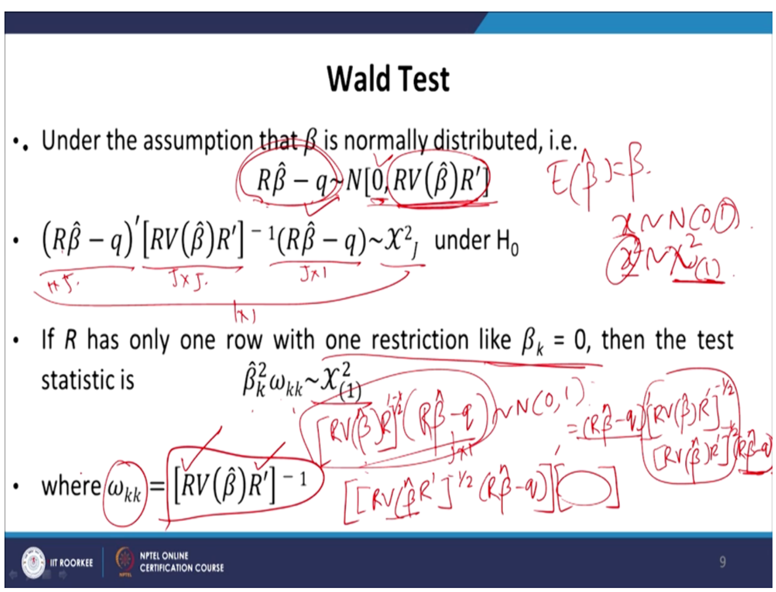


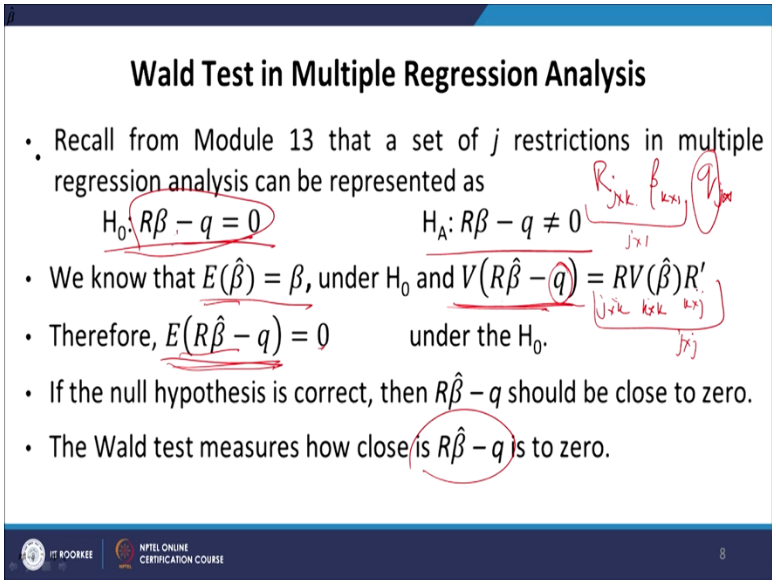

Now, I talk about the Wald test in multiple regression analysis. Now, recall from Module 13 that a set of j restrictions in multiple regression analysis can be represented as (refer slide time: 8:51).

Now, just to recall it slightly more, I tell you that (refer slide time: 9:06- 9:50).

Why is this so happening? Because q is a constant term, or it is a vector of constant numbers. (Refer slide time: 10:00- 11:15 )

(Refer Slide Time: 11:15)






Now, another assumption is that β is normally distributed, (refer slide time: 11:20- 12:45)

.

Now, if you remember, in the previous module we discussed that if $x$ follows a standard normal distribution then $x^2$ follows a chi-square distribution, with one degree of freedom, when $x$ is one variable.

Now, you note that this can be written as I write this (refer slide time: 13:07).

So, I make it a power of minus half. And this follows a standard normal distribution. Now, if I square this term, then this term will follow a chi-square distribution with $j$ degrees of freedom. Here, we have $j$ degrees of freedom because $R$ has $j$ restrictions.

Since it was a single variable, we had one degree of freedom, when there are $j$ variables, we have $j$ restrictions. And vice -versa.

Now, how do we go for a square of this term? Note that we have already shown here that this is (refer slide time: 14:15), the entire thing put under a transpose multiplied by the same thing without a transpose.

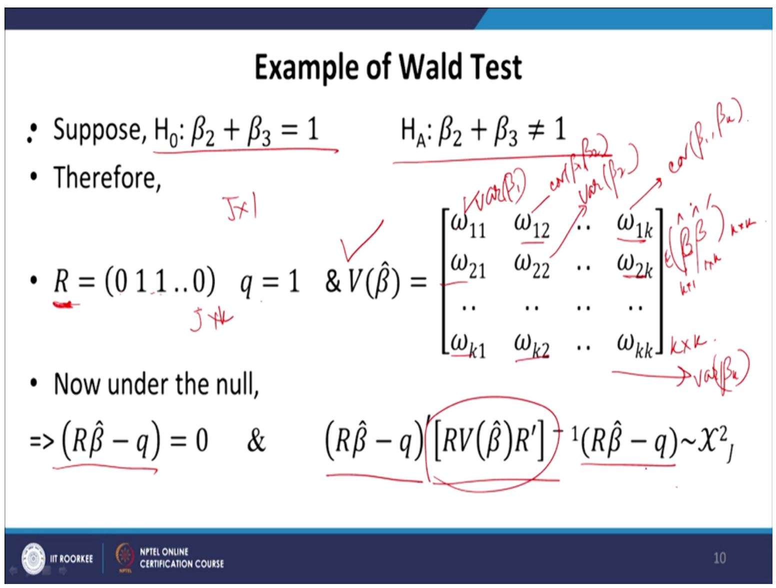
When I expand it, then (refer slide time: 14:50). Now, since these are variance co-variance matrices and these are symmetric matrices, (refer slide time:15:18- 16:50 ).

So, I repeat, that when there is only 1 variable, it follows a chi-square distribution with 1 degree of freedom.

Similarly, when there is only 1 restriction, then it follows a chi-square distribution with 1 degree of freedom. Since we have only 1 restriction, then it follows a chi-square distribution with 1 degree of freedom where $w_{kk}$ is this matrix but since we have only 1 non-zero component in this matrix and, in this matrix, so it actually turns out to be a single number, which is $w_{kk}$.

Now, I will explain it, rather in detail, in the next slide, when we go for an example of the Wald test with multiple restrictions.

(Refer Slide Time: 17:33)



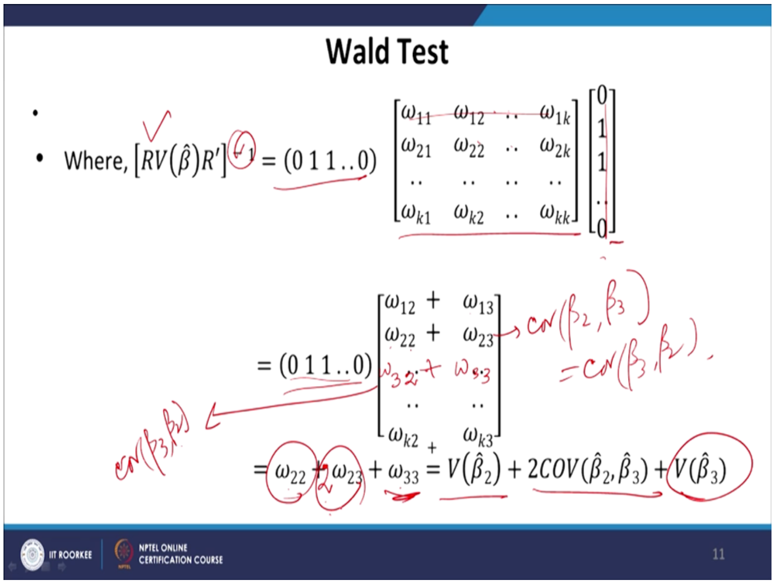

So, suppose we consider, basically here, the restriction is 1 but we have more than 1 variable. So, my null hypothesis is (refer slide time: 17:46- 18:32 ). These things in an earlier context probably were derived for $\hat{\beta}$ as well as the population error variance.

So, (refer slide time: 18:39) And what are these terms? These cross-terms are the co-variances between different parameter estimates.

So (refer slide time: 19:00- 20:09 ).

I have just mentioned them to facilitate explanation.

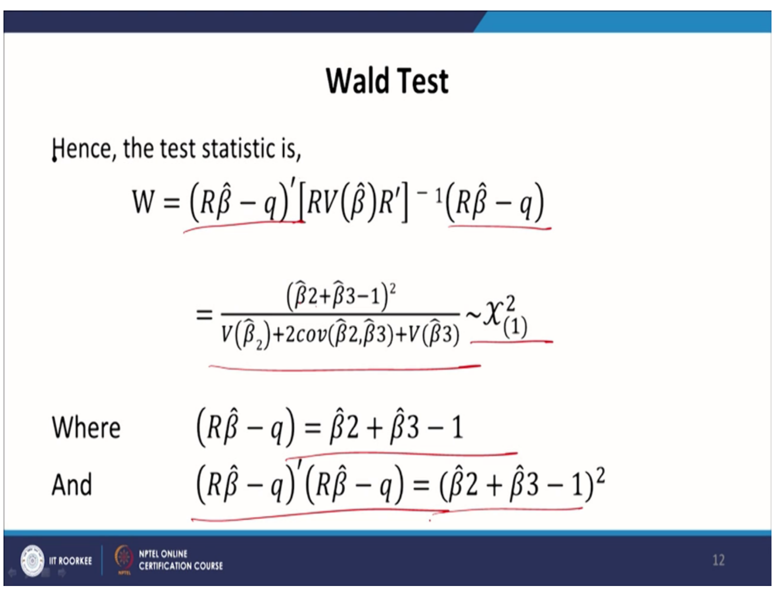
(Refer Slide Time: 20:12)

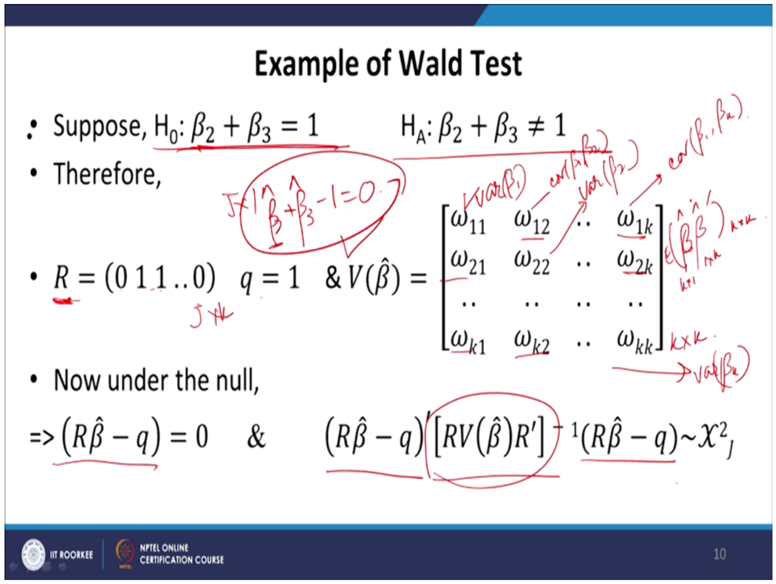


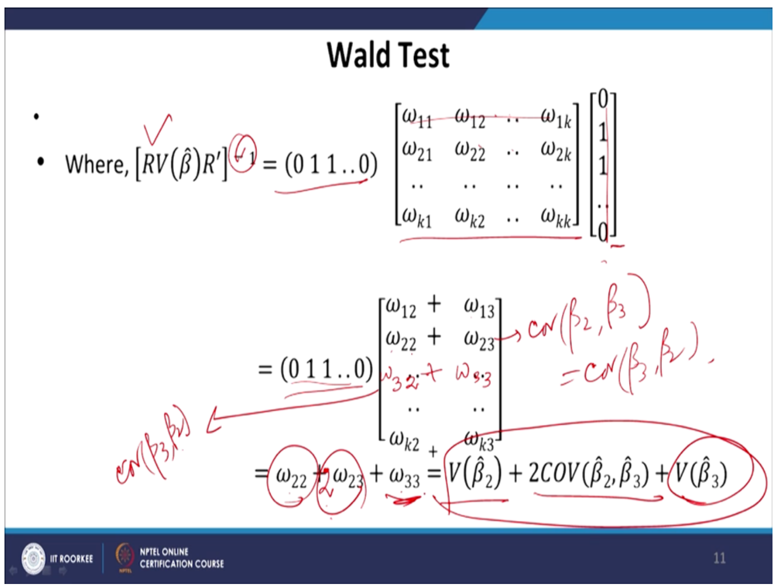

So, now we go for the derivation of the expression (refer slide time: 20:17- 22:12).

.

(Refer Slide Time: 22:12)



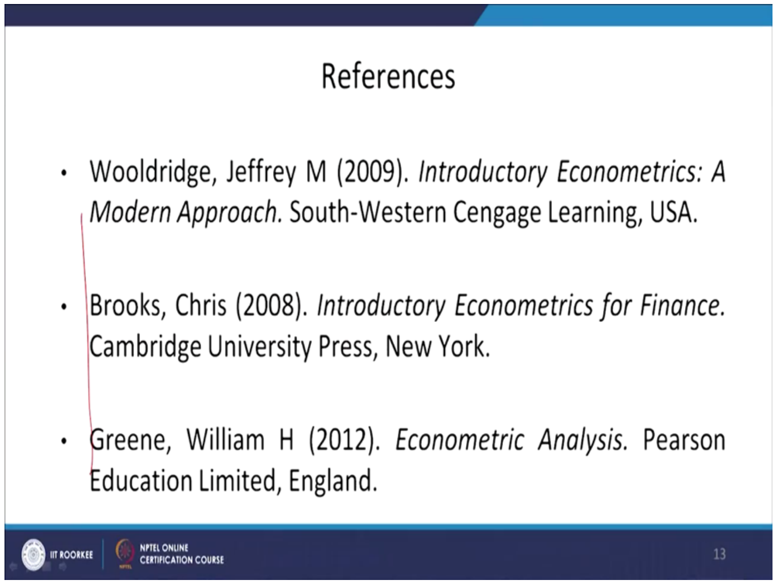



Hence, the test statistic is, (refer slide time: 22:17- 23:10).

This follows a chi-square distribution with one degree of freedom because there is only one restriction. Now, this, I have mentioned here that (refer slide time: 23:19).

(Refer Slide Time: 23:34)




So, we have discussed Wald statistics. These are the references that I have followed. But Wald statistic is a large sample and or asymptotic test. Why do we call it a large sample or asymptotic test, which I will explain probably in the next module? But one major characteristic of Wald statistic is that it does not facilitate estimation of the restricted model.

So, we cannot have restrictions on the parameter estimates. We generally have to estimate complete models. So, thank you. That is all about the Wald test.