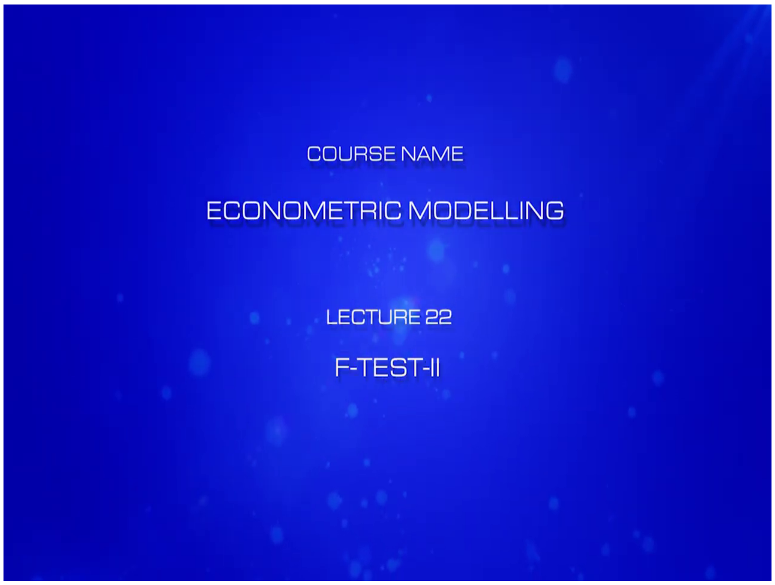
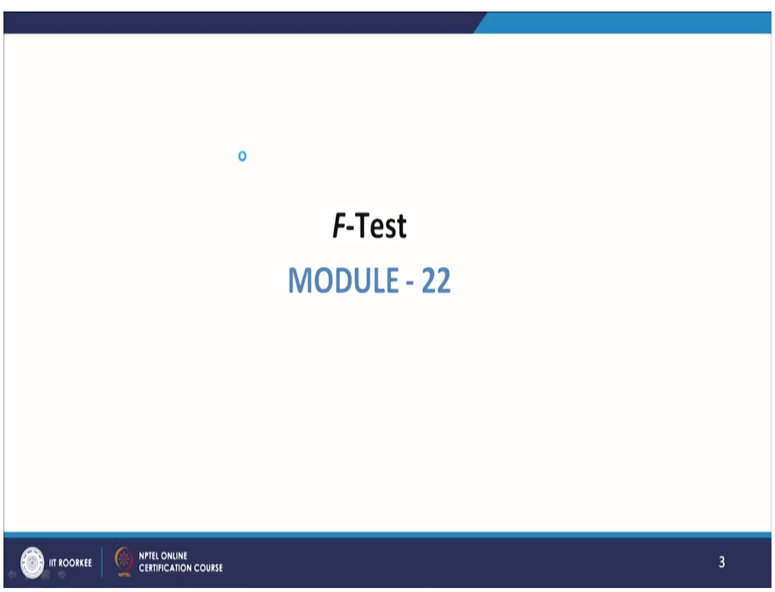
Econometric Modelling
Professor Sujata Kar
Department of Management Studies
Indian Institute of Technology, Roorkee
Lecture 22
F-Test-II

(Refer Slide Time: 00:11)



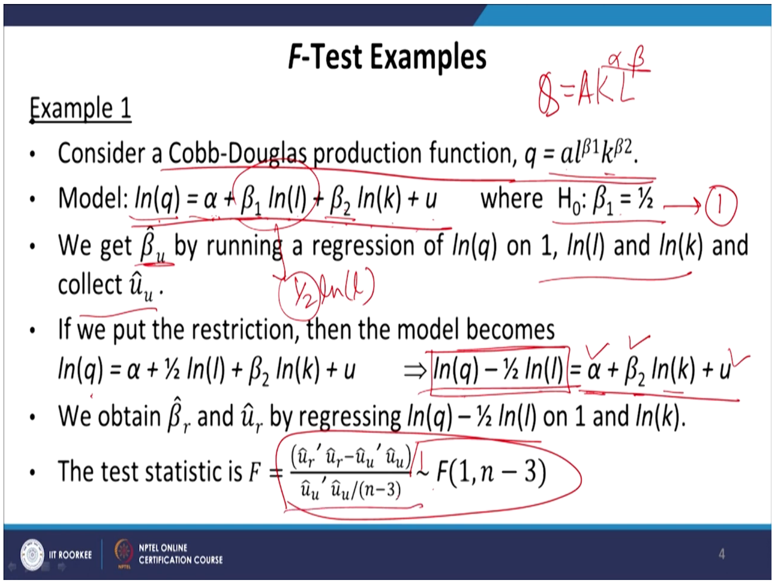
This is Module 22 where we continue with discussion on F statistic or F-test.

(Refer Slide Time: 00:35)



We further discuss some of the examples of F-test followed by applications of F-test, comparison of F-test with Wald test and t-test.

(Refer Slide Time: 00:47)



So, first, consider an example where we basically consider a Cobb-Douglass production function. If you remember, Cobb-Douglass production functions are written like this *(refer to slide time 00:47)*. In most often in economics textbooks, you would find a Cobb-Douglass production function is written as A equals K raised to the power alpha L raised to the power beta. So exactly in a similar fashion, we have written it here, where beta 1 and beta 2 are basically the share parameters.

Now, this can be converted into a linear model by taking the logarithm. So, we have a logarithm of q equals alpha plus beta 1 logarithm of l plus beta 2 logarithms of k plus u. And suppose we have a null hypothesis, which states that beta 1 is equal to half.

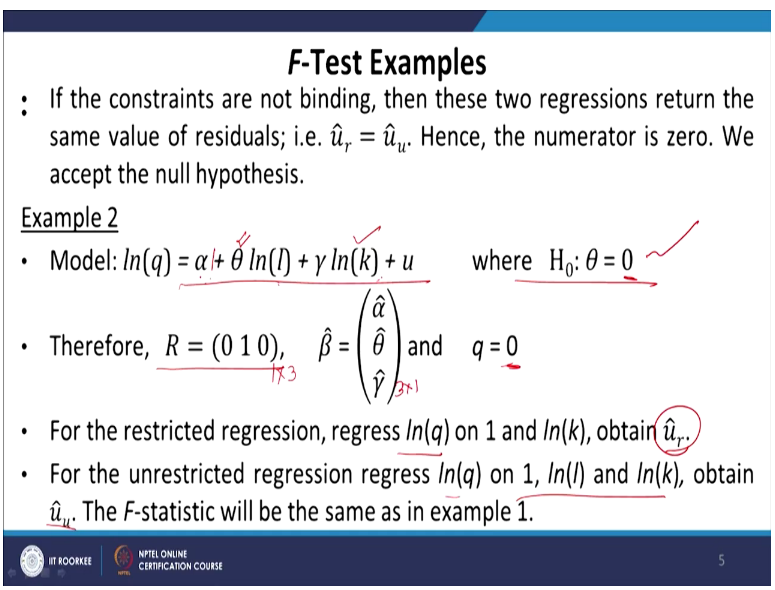
We get beta u hat, that is, the unrestricted set of estimated parameters by running a regression of logarithm of q on 1, logarithm of l, and logarithm of k. And we collect uu hat. If we put the restriction then beta 1 becomes half. So, we have alpha plus the half logarithm of l plus beta 2 logarithms of k plus u.

Now when it is the half logarithm of l then you can see that we do not have any parameter to be estimated. As a result of which we can take it to the other side. So this becomes logarithm of q minus the half logarithm of l. So, now my dependent variable is actually not the logarithm of q, but it is the logarithm of q minus the half logarithm of l. And on the hand side, we have alpha and beta 2, which are to be estimated, plus the error term *(refer to slide time 00:47)*.

So we obtain beta r hat. Now you see that beta r is a vector that consists of alpha and beta 2, while beta u hat is a vector that consists of alpha-beta 1 and beta 2, alpha-beta 1 and beta 2.

So, we obtain beta r hat and Ur hat by regressing logarithm of q minus half of logarithm of l on 1 and logarithm of k, and the test statistic F is the usual one. Now note that there is only one restriction. So, we have the numerator degrees of freedom equal to 1, and the denominator degrees of freedom is n minus 1, 2, 3, so n minus 3. So, this follows an F distribution with 1 and n minus 3 degrees of freedom.

(Refer Slide Time: 03:42)



If the constraints are not binding then these two regressions return the same value of the residuals Ur hat equals uu hat, hence the numerator is 0, and we accept the null hypothesis. But if we do not accept the null hypothesis then what happens is that we conclude that the constraints are not binding.

Now, or rather I will, I must say that if we reject the null hypothesis, then the conclusion is that the constraints are actually binding.

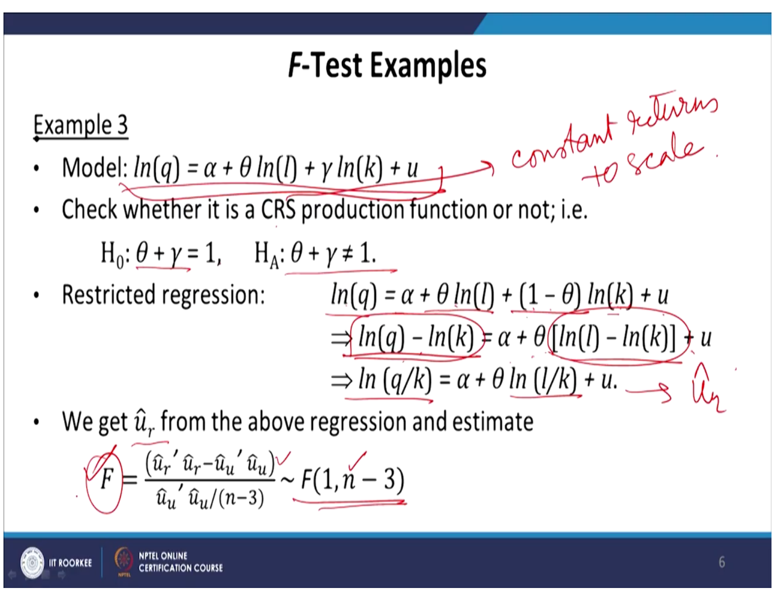
Now, consider another example *(refer to slide time 03:42)*. We continue with that Cobb-Douglas production function. Now, we have alpha plus theta logarithm of l plus gamma logarithm of k plus u. And now the null hypothesis is theta equals 0. We are putting only one restriction.

Therefore, we write it as R equals 0 1 0. You can see that there is only one restriction, so R is a 1 by 3, vector beta hat is a 3 by 1 vector, and q equals 0. For the restricted regression, we regress the logarithm of q on only 1 and the logarithm of k because theta has been put equals to 0, and that is why there is nothing to estimate.

We obtain Ur hat. For the unrestricted regression we regress logarithm of q on the usual thing, that is 1 logarithm of ln, the logarithm of k and obtain uu hat. The F statistic will be the same as that in example 1. Again, we have only one restriction and there are three basically two dependent independent variables plus one constant term so the degrees of freedom also remain the same.

F follows an F distribution or the test statistic would follow an F distribution with 1 and n minus 3 degrees of freedom.

(Refer Slide Time: 05:40)



The third example, suppose has a slightly different null hypothesis. So we again consider a Cobb-Douglas type production function where we have logarithm of q equals alpha plus theta logarithm of l plus gamma logarithm of k plus u.

Now, check whether it is a CRS production function or not. CRS stands for constant returns to scale. So, if it is a CRS production function, then theta and gamma will be equals to 1. This is implied by the fact that if it is a CRS production function then theta plus gamma will be equal to 1.

The null hypothesis is, the alternative hypothesis is theta plus gamma not equal to 1. So, the restricted regression will be the logarithm of q equals alpha plus theta logarithm of l and then gamma is replaced by 1 minus theta. So 1 minus of theta logarithm of k plus u.

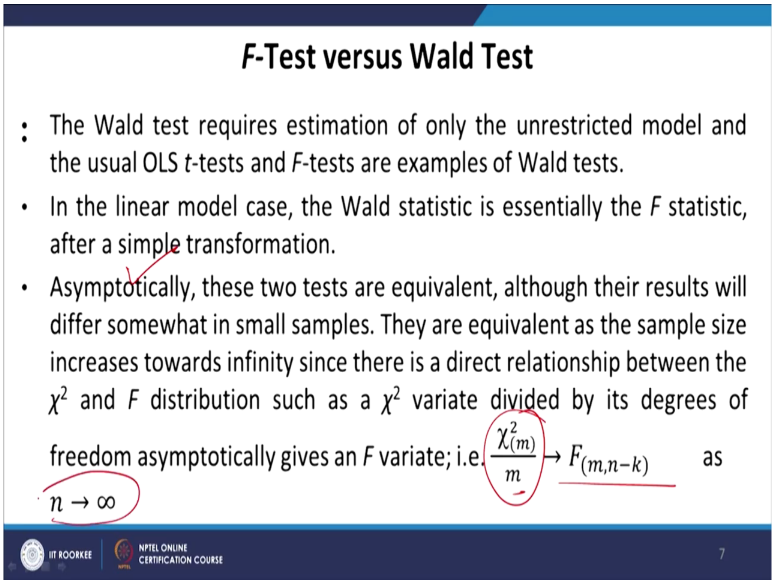
Now, by breaking this I will be having 1 plus logarithm of k, and minus theta logarithm of k. So, plus the logarithm of k is taken to the left-hand side, I have logarithm of q minus logarithm of k, then alpha plus theta logarithm of l minus logarithm of k plus u. So, I have two new variables. One is the dependent variable, which is the logarithm of q minus logarithm of k, and one new independent variable which is the logarithm of l minus logarithm of k.

So, the variables have actually changed. And then we run a regression of the logarithm of q by k equals alpha plus theta logarithm of l by k plus u. And we follow the same procedure, that from here I will be obtaining Ur hat.

So we get Ur hat from the above regression, and uu hat would come from the unrestricted regression, like the previous example. And then we calculate the F-ratio or the F statistic following this formula. This would also follow an F distribution with 1 and n minus 3 degrees of freedom.

So, once we have information on the value of n, we will be specific, more specific about what are the degrees of freedom, and then by consulting an F table, we can find out whether our calculated test statistic or calculated F-ratio is greater than or less than the critical value. Accordingly, we will either reject the null hypothesis or not reject the null hypothesis.

(Refer Slide Time: 08:24)



Now we compare the Wald test with F-test. The Wald test requires estimation of only the unrestricted model and the usual OLS t-tests and F-tests are examples of the Wald test. So this incorporation of restriction is a specific characteristic of the F-test only because Wald tests are meant for only unrestricted regression.

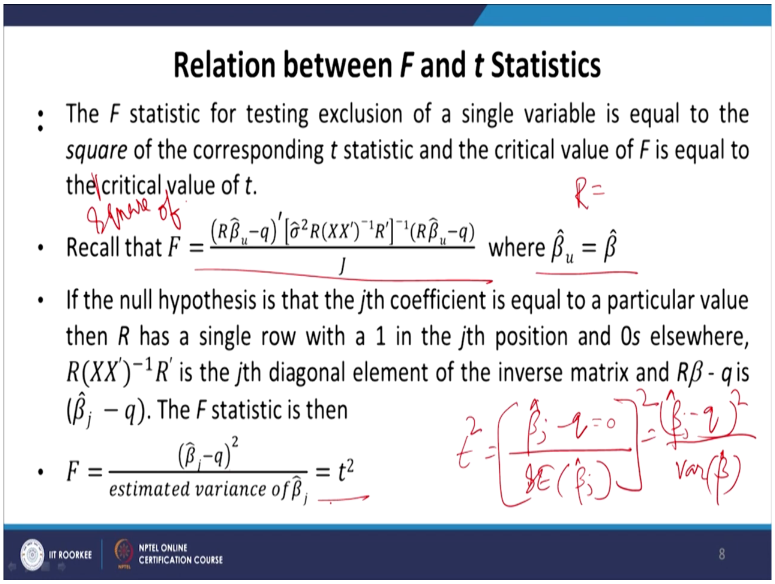
In the linear model case, the Wald statistic is essentially the F statistic after a simple transformation that we have already seen. How do we arrive at the F statistic from the Wald statistic?

Asymptotically, these two tests are equivalent although their results will differ somewhat in small samples because the reason is that Wald statistic is a large sample test, whereas while F statistic or F-test applies to both large and small samples. So, when we talk about asymptotically, that is, with a large sample both of, the tests will actually provide very similar results.

They are equivalent as the sample size increases towards infinity. Since there is a direct relationship between the chi-square and F distribution. Such as, a chi-square varies divided by its degrees of freedom asymptotically gives an F variate.

So, chi-square distribution divided by its degrees of freedom gives us an F distribution asymptotically, that is when n tends to infinity.

(Refer Slide Time: 09:56)



So, now we discuss the relationship between F and t statistics. The F statistic for testing the exclusion of a single variable is equal to the square of the corresponding t statistic, and the critical value of F is equal to basically the square of the critical value of t.

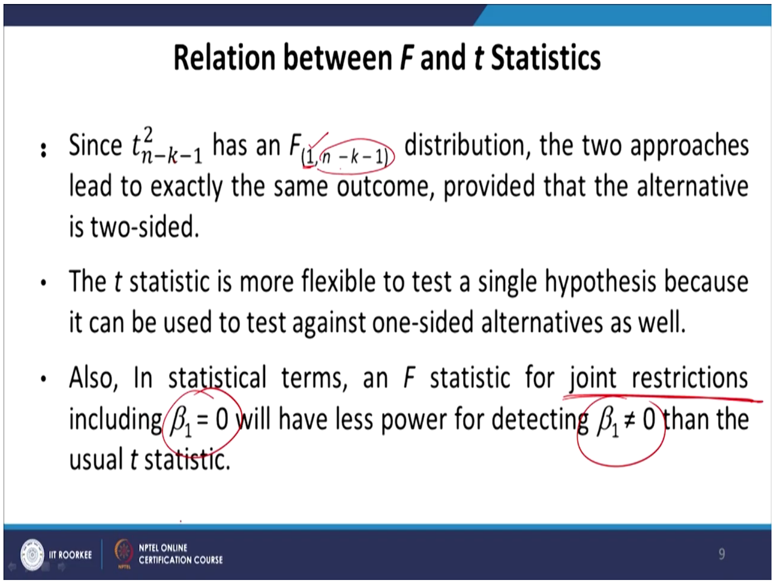
Now, why is it so? You recall that this was our expression for the F statistic where beta u hat is equaled to the beta hat. So, the unrestricted regression is simply most often the original model. So, the beta u hat is equivalent to the beta hat.

If the null hypothesis is that the jth coefficient is equal to a particular value then R has a single row. The capital R which basically considers the set of restrictions has a single row with 1 in the jth position and 0 elsewhere. R XX prime inverse R prime is the jth diagonal element of the inverse matrix, and R beta minus q will be equal to beta j minus q.

Then this F statistic is beta j hat minus q whole square divided by the estimated variance of beta j hat *(refer to slide time 09:56)*. And if you remember t statistic is beta j hat minus q, where this is the null hypothesis, and we have the standard error of beta j hat in the denominator. So, when I square it then this becomes beta j hat minus q whole square divided by variance of estimated parameter. So, we can see that this is equal to t square.

So F-ratio is actually equal to the square of the t-ratio or t statistics.

(Refer Slide Time: 11:50)

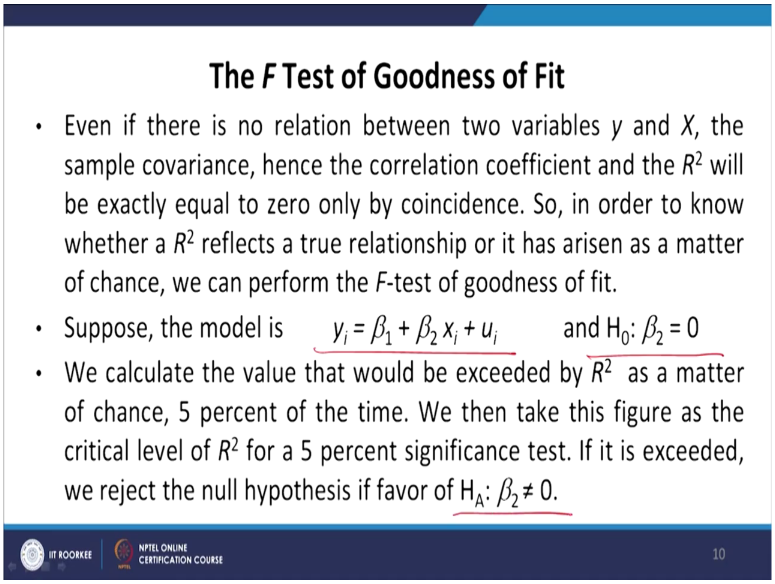


Since t n minus k minus 1 square or t square with n minus k minus 1 degree of freedom has an F 1, n minus k minus 1 distribution. Now you remember that this is, the numerator degrees of freedom are same, same as that of t's degrees of freedom. But since t statistics tests only one hypothesis at a time so there is only one restriction and F should have 1 as the numerator degrees of freedom.

So, since t square n minus k minus 1 has an F 1, n minus k minus 1 distribution, the two approaches lead to exactly the same outcome provided that the alternative is a two-sided test. The t statistic is more flexible to test a single hypothesis because it can be used to test against one-sided alternatives as well, while the F statistic is applicable only in the case of two-sided tests.

Also in statistical terms, an F statistic for joint restrictions, including beta 1 equals 0 will have less power for detecting beta 1 not equal to 0 than the usual t statistic. So, when we are going for joint restrictions or test of joint restrictions, then focusing on a single or the significance of a single parameter is more powerful if we go for t statistic compared to when we consider F statistic.

(Refer Slide Time: 13:20)



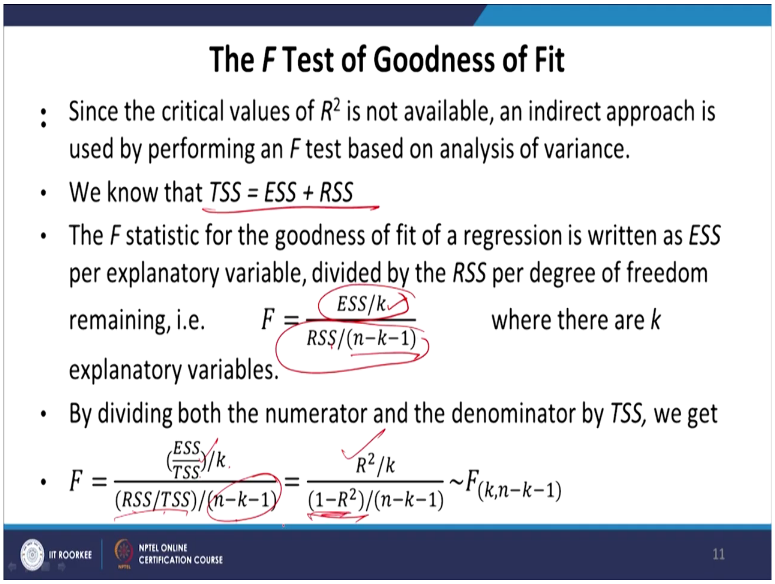
Now we talk about the F-test of the goodness of fit. Why is it so? Because even if there is no relation between two variables y and X, the sample covariance hence the correlation coefficient and the R square will be exactly equal to 0 only by coincidence, that is, by chance. So, in order to know whether an R square reflects a true relationship or not, it has arisen as a matter of chance we can perform the F-test of the goodness of fit.

- Suppose the model is $y_i = \beta_1 + \beta_2 x_i + u_i$

So, we are considering a simple regression. i stands for the ith observation where the null hypothesis is beta 2 equals 0. We calculate the value that would be exceeded by R square as a matter of chance 5 percent of the time. We then take this figure as the critical level of R square for a 5 percent significance test.

If it is exceeded, we reject the null hypothesis in favor of the alternative hypothesis, which is beta 2 not equal to 0 *(refer to slide time 13:20).*

(Refer Slide Time: 14:30)



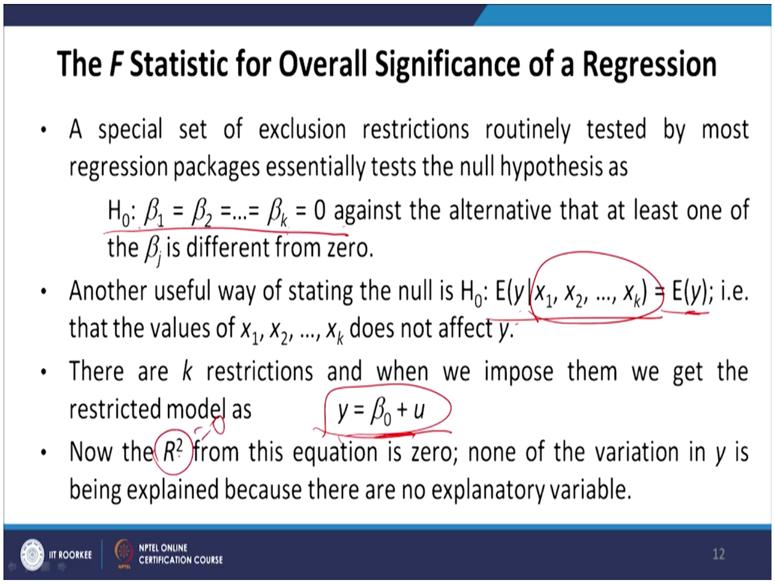
Since the critical values of R square are not available, an indirect approach is used by performing an F-test based on analysis of variance because F critical values are available with us.

We know that the total sum of squares equals the expected sum of squares plus the residual sum of squares. The F statistic for the goodness of fit of regression is written as ESS per explanatory variable divided by the RSS per degree of freedom remaining.

So, this is the number of the explanatory variables *(refer to slide time 14:30).* So ESS per explanatory variable divided by RSS per degree of freedom remaining, the remaining degrees of freedom, which is n minus k minus 1. By dividing both the numerator and the denominator by TSS, what do we get? We get ESS divided by TSS, the whole divided by k, RSS divided by TSS, the whole divided by the degrees of freedom n minus k minus 1.

So you know that R square is ESS divided by TSS. So, we have R square in the numerator, and RSS divided by TSS is actually 1 minus R square. So, this is what we have in the denominator. And it follows an F distribution with k and n minus k minus 1 degree of freedom.

(Refer Slide Time: 15:51)



Now, we talk about the F statistic for the overall significance of a regression. So, a special set of exclusion restrictions routinely tested by most regression packages essentially tests the null hypothesis as H0 equals beta 1 plus beta 2, beta 1 equals to beta 2 equals to beta k equals to 0, against the alternative that at least one of the beta j is different from 0.

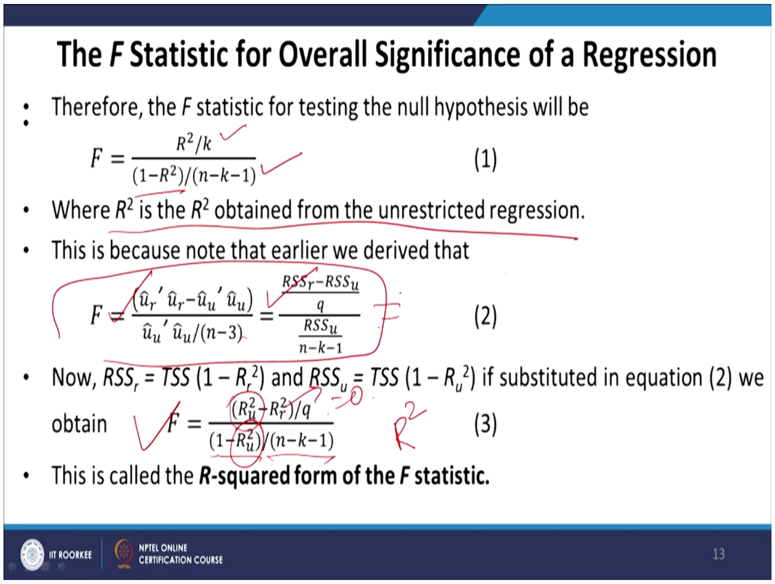
So, another useful way of stating the null is expected value of y given x1, x2 to xk is equal to the expected value of y, that is, the values of x1 to xk do not affect y altogether. Conditional upon x is as good as non-conditional upon x. So, y is independent of x, or the expected value of y is independent of x.

There are k restrictions and when we impose them we get the restricted model as y equals beta naught plus u. So, you can see that here we have k restrictions for all the explanatory variables, beta 1 to beta k. So, what I am left with for the restricted regression is actually y equals beta 0 plus u*(refer to slide time 15:51).*

We are regressing y only on the constant term under the assumption that all the explanatory variables do not contribute or none of the explanatory variables contributes to the explanation of y.

Now the R square from this regression is actually 0. None of the variations in y is being explained because there are no explanatory variables, so that is why we have R square equal to 0 from this regression.

(Refer Slide Time: 17:33)



Therefore the F statistic testing the null hypothesis will be R square divided by k divided by 1 minus R square divided by n minus k minus 1, where R square is the R square obtained from the unrestricted regression because we cannot perform or run the restricted regression without any explanatory variable altogether.

As earlier derived *(refer to slide time 17:33) a*nd on that basis, we can also derive that Ru square minus Rr square divided by q is divided by 1 minus Ru square divided by n minus k minus 1.

Now here, Rr square is actually equal to 0. So that is why we would be left with Ru square in the numerator, which is basically denoted by the R square. And again, 1 minus Ru square, which is basically 1 minus R square in the denominator. Now, this is called the R squared form of the F statistic.

So you can see that these are actually all very similar. This thing we have already seen, from there we know this one, and finally, this can further also be written as something like equation 3.

(Refer Slide Time: 18:54)

The *F* Statistic for Overall Significance of a Regression

- Because R^2 is reported in all regression, it is easy to use the this form of the *F* statistic.
- Also note that unrestricted R^2 comes first in the numerator because $R_u^2 > R_r^2$ as the unrestricted regression has more explanatory variables.
- This also ensures that the *F* statistic is always positive.
- Therefore, if $R_r^2 = 0$ in equation (3), we obtain the *F* statistic in equation (1).
- This special form of *F* statistic is valid only for joint exclusion of all independent variables.
- This is sometimes called determining the **overall significance of the regression.**

IIT ROORKEE | NPTEL ONLINE CERTIFICATION COURSE 14

The *F* Statistic for Overall Significance of a Regression

- Therefore, the *F* statistic for testing the null hypothesis will be

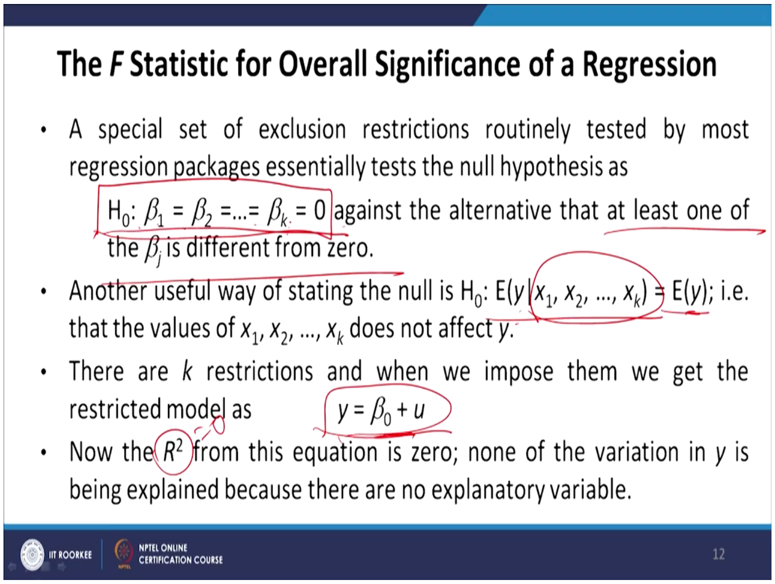
$$F = \frac{R^2/k}{(1-R^2)/(n-k-1)} \qquad (1)$$

- Where R^2 is the R^2 obtained from the unrestricted regression.
- This is because note that earlier we derived that

$$F = \frac{(\hat{u}_r{}'\hat{u}_r - \hat{u}_u{}'\hat{u}_u)}{\hat{u}_u{}'\hat{u}_u/(n-3)} = \frac{\frac{RSS_r - RSS_u}{q}}{\frac{RSS_u}{n-k-1}} \qquad (2)$$

- Now, $RSS_r = TSS\,(1 - R_r^2)$ and $RSS_u = TSS\,(1 - R_u^2)$ if substituted in equation (2) we obtain $F = \frac{(R_u^2 - R_r^2)/q}{(1-R_u^2)/(n-k-1)}$ (3)
- This is called the ***R*-squared form of the *F* statistic.**

IIT ROORKEE | NPTEL ONLINE CERTIFICATION COURSE 13



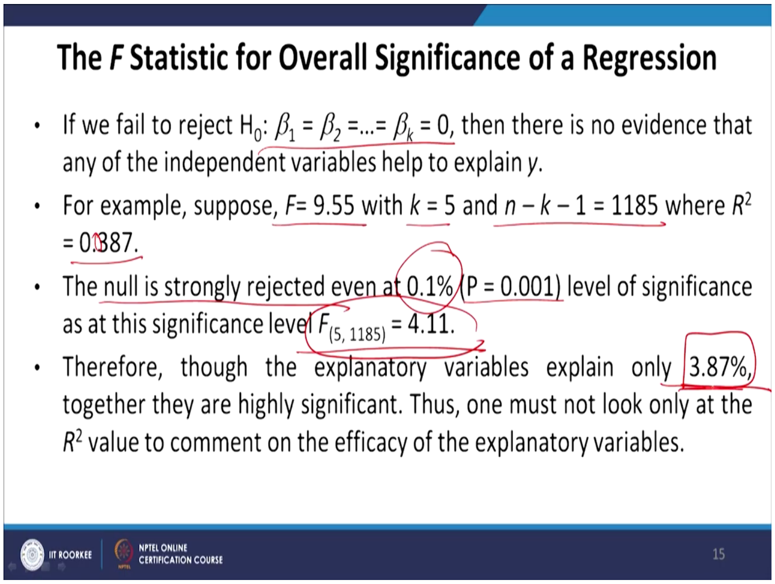
Because R square is reported in all regression, it is easy to use the form of the F statistic. Also, note that unrestricted R square comes first in the numerator because Ru square is greater than Rr square as the unrestricted regression has more explanatory variables. This also ensures that the F statistic is always positive.

Therefore if Rr square equals 0, in equation 3, we obtain the F statistic in equation 1, this is the F statistic in equation 1 *(refer to slide time 18:54)*. The special form of F statistic is valid only for joint exclusion of all independent variables. This is sometimes called determining the overall significance of a regression.

So, if we obtain an F value which is large enough then what do we do? We reject this null hypothesis. Which means that what is the alternative hypothesis? The alternative hypothesis is not rejected, which basically says that at least one of the beta j is different from 0.

So at least one of these independent variables, or since one of these parameter estimates are not 0, one of these, one of the explanatory variables actually contributing to the variations in y. So this is called the overall significance of the regression.

(Refer Slide Time: 20:26)

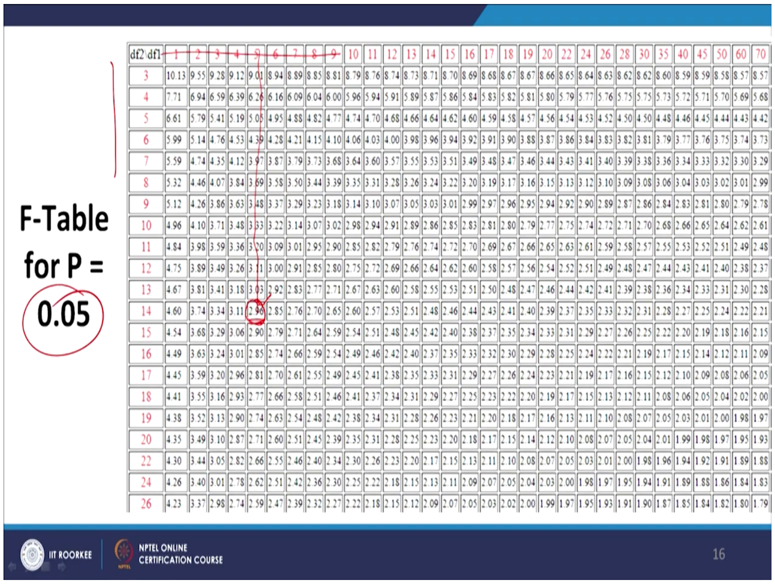


If we fail to reject the null hypothesis, which is beta 1 equals beta 2 to beta k equals 0, then there is no evidence that any of the independent variables help to explain y. For example, suppose F equals 9.55 with k equals 5, n minus k minus 1 equals 1185, and where we obtain the R square as 0.387. The null is strongly rejected even at 0.1 percent *(refer to slide time 20:26)*.

So, by considering an F table with a P value of 0.001, which implies the significance level of 0.1 percent, we are actually rejecting the null hypothesis because at this significance level the F statistic value or the F-ratio with 5 and 1185 degrees of freedom is 4.11. Therefore, though the explanatory variables explain only 3.87 percent, together they are highly significant. This should be 0.0387.

So, the model actually explains only 3.87 percent of the variations in y. Thus one must not look only at the R square value to comment on the efficacy of the explanatory variables. Even with a very small R square value, it is possible that we have some or all of the explanatory variables actually contributing to explain the variations in y, which is reflected through the F statistic despite having a very small R square number, or R square value.

(Refer Slide Time: 22:08)



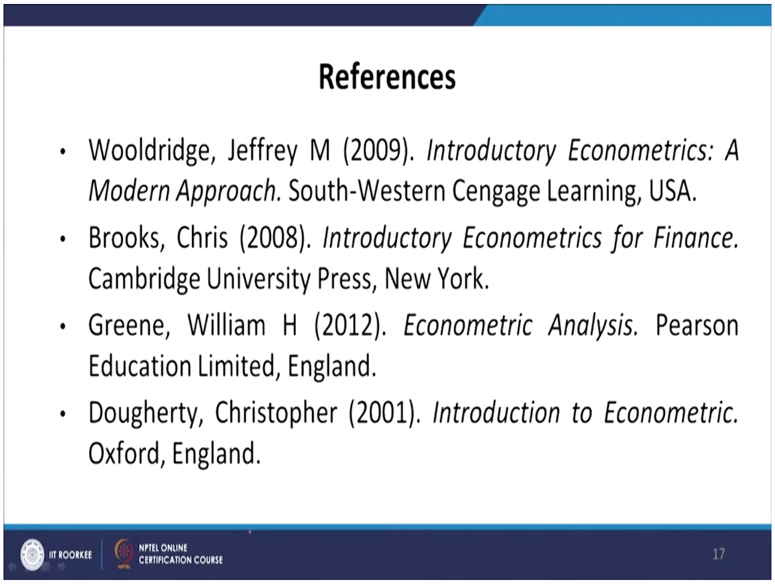
df2\df1	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	22	24	26	28	30	35	40	45	50	60	70
3	10.13	9.55	9.28	9.12	9.01	8.94	8.89	8.85	8.81	8.79	8.76	8.74	8.73	8.71	8.70	8.69	8.68	8.67	8.67	8.66	8.65	8.64	8.63	8.62	8.62	8.60	8.59	8.59	8.58	8.57	8.57
4	7.71	6.94	6.59	6.39	6.26	6.16	6.09	6.04	6.00	5.96	5.94	5.91	5.89	5.87	5.86	5.84	5.83	5.82	5.81	5.80	5.79	5.77	5.76	5.75	5.75	5.73	5.72	5.71	5.70	5.69	5.68
5	6.61	5.79	5.41	5.19	5.05	4.95	4.88	4.82	4.77	4.74	4.70	4.68	4.66	4.64	4.62	4.60	4.59	4.58	4.57	4.56	4.54	4.53	4.52	4.50	4.50	4.48	4.46	4.45	4.44	4.43	4.42
6	5.99	5.14	4.76	4.53	4.39	4.28	4.21	4.15	4.10	4.06	4.03	4.00	3.98	3.96	3.94	3.92	3.91	3.90	3.88	3.87	3.86	3.84	3.83	3.82	3.81	3.79	3.77	3.76	3.75	3.74	3.73
7	5.59	4.74	4.35	4.12	3.97	3.87	3.79	3.73	3.68	3.64	3.60	3.57	3.55	3.53	3.51	3.49	3.48	3.47	3.46	3.44	3.43	3.41	3.40	3.39	3.38	3.36	3.34	3.33	3.32	3.30	3.29
8	5.32	4.46	4.07	3.84	3.69	3.58	3.50	3.44	3.39	3.35	3.31	3.28	3.26	3.24	3.22	3.20	3.19	3.17	3.16	3.15	3.13	3.12	3.10	3.09	3.08	3.06	3.04	3.03	3.02	3.01	2.99
9	5.12	4.26	3.86	3.63	3.48	3.37	3.29	3.23	3.18	3.14	3.10	3.07	3.05	3.03	3.01	2.99	2.97	2.96	2.95	2.94	2.92	2.90	2.89	2.87	2.86	2.84	2.83	2.81	2.80	2.79	2.78
10	4.96	4.10	3.71	3.48	3.33	3.22	3.14	3.07	3.02	2.98	2.94	2.91	2.89	2.86	2.85	2.83	2.81	2.80	2.79	2.77	2.75	2.74	2.72	2.71	2.70	2.68	2.66	2.65	2.64	2.62	2.61
11	4.84	3.98	3.59	3.36	3.20	3.09	3.01	2.95	2.90	2.85	2.82	2.79	2.76	2.74	2.72	2.70	2.69	2.67	2.66	2.65	2.63	2.61	2.59	2.58	2.57	2.55	2.53	2.52	2.51	2.49	2.48
12	4.75	3.89	3.49	3.26	3.11	3.00	2.91	2.85	2.80	2.75	2.72	2.69	2.66	2.64	2.62	2.60	2.58	2.57	2.56	2.54	2.52	2.51	2.49	2.48	2.47	2.44	2.43	2.41	2.40	2.38	2.37
13	4.67	3.81	3.41	3.18	3.03	2.92	2.83	2.77	2.71	2.67	2.63	2.60	2.58	2.55	2.53	2.51	2.50	2.48	2.47	2.46	2.44	2.42	2.41	2.39	2.38	2.36	2.34	2.33	2.31	2.30	2.28
14	4.60	3.74	3.34	3.11	2.96	2.85	2.76	2.70	2.65	2.60	2.57	2.53	2.51	2.48	2.46	2.44	2.43	2.41	2.40	2.39	2.37	2.35	2.33	2.32	2.31	2.28	2.27	2.25	2.24	2.22	2.21
15	4.54	3.68	3.29	3.06	2.90	2.79	2.71	2.64	2.59	2.54	2.51	2.48	2.45	2.42	2.40	2.38	2.37	2.35	2.34	2.33	2.31	2.29	2.27	2.26	2.25	2.22	2.20	2.19	2.18	2.16	2.15
16	4.49	3.63	3.24	3.01	2.85	2.74	2.66	2.59	2.54	2.49	2.46	2.42	2.40	2.37	2.35	2.33	2.32	2.30	2.29	2.28	2.25	2.24	2.22	2.21	2.19	2.17	2.15	2.14	2.12	2.11	2.09
17	4.45	3.59	3.20	2.96	2.81	2.70	2.61	2.55	2.49	2.45	2.41	2.38	2.35	2.33	2.31	2.29	2.27	2.26	2.24	2.23	2.21	2.19	2.17	2.16	2.15	2.12	2.10	2.09	2.08	2.06	2.05
18	4.41	3.55	3.16	2.93	2.77	2.66	2.58	2.51	2.46	2.41	2.37	2.34	2.31	2.29	2.27	2.25	2.23	2.22	2.20	2.19	2.17	2.15	2.13	2.12	2.11	2.08	2.06	2.05	2.04	2.02	2.00
19	4.38	3.52	3.13	2.90	2.74	2.63	2.54	2.48	2.42	2.38	2.34	2.31	2.28	2.26	2.23	2.21	2.20	2.18	2.17	2.16	2.13	2.11	2.10	2.08	2.07	2.05	2.03	2.01	2.00	1.98	1.97
20	4.35	3.49	3.10	2.87	2.71	2.60	2.51	2.45	2.39	2.35	2.31	2.28	2.25	2.23	2.20	2.18	2.17	2.15	2.14	2.12	2.10	2.08	2.07	2.05	2.04	2.01	1.99	1.98	1.97	1.95	1.93
22	4.30	3.44	3.05	2.82	2.66	2.55	2.46	2.40	2.34	2.30	2.26	2.23	2.20	2.17	2.15	2.13	2.11	2.10	2.08	2.07	2.05	2.03	2.01	2.00	1.98	1.96	1.94	1.92	1.91	1.89	1.88
24	4.26	3.40	3.01	2.78	2.62	2.51	2.42	2.36	2.30	2.25	2.22	2.18	2.15	2.13	2.11	2.09	2.07	2.05	2.04	2.03	2.00	1.98	1.97	1.95	1.94	1.91	1.89	1.88	1.86	1.84	1.83
26	4.23	3.37	2.98	2.74	2.59	2.47	2.39	2.32	2.27	2.22	2.18	2.15	2.12	2.09	2.07	2.05	2.03	2.02	2.00	1.99	1.97	1.95	1.93	1.91	1.90	1.87	1.85	1.84	1.82	1.80	1.79

This is how an F table looks like *(refer to slide time 22:08)*. We have considered P equals to 0.05, that is, the significance level is at 5 percent. The F table would have 2 degrees of freedom. d F one, measured on, along with the columns, and dF two, measured along the rows. And suppose I have an F statistic or F-ratio with 5 and 14 degrees of freedom.

So, I come along this column *(refer to slide time 22:08)*, stop where this is 14, and 2.96 will be my estimate, my critical, F critical value, and then this is what I am going to compare with the calculated value of the F-ratio. If the calculated value is greater than this critical value, then we reject the null hypothesis. If the calculated value is lower than this critical value, then we tend to not reject the null hypothesis.

So that is broadly about F statistic. It is some of the examples, its applications, and its comparisons with two alternative tests that we have discussed earlier, the Wald test and the t-test.

(Refer Slide Time: 23:20)



And these are the references, again, that I have considered or consulted in order to come up with the discussion. Thank you.