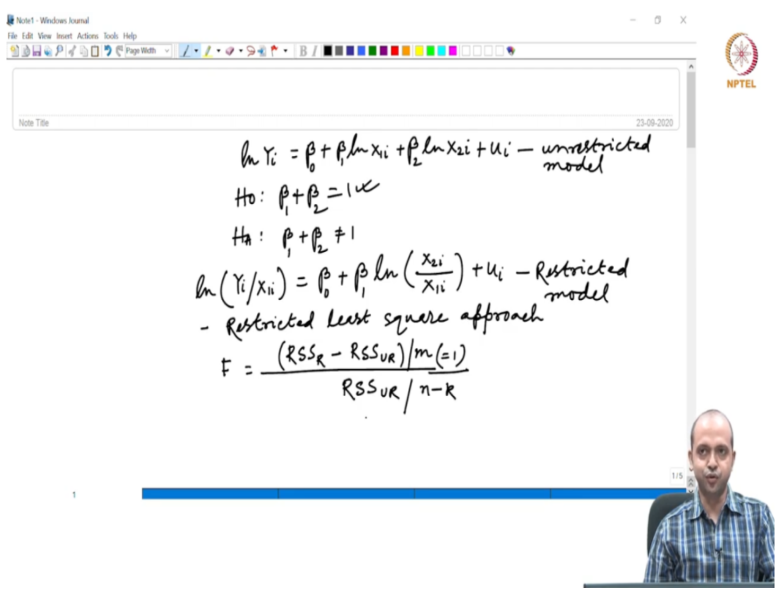
Introduction to Econometrics
Professor Sabuj Kumar Mandal
Department of Humanities and Social Sciences
Indian Institute of Technology, Madras
Lecture 29
Structural break analysis using Chow test Part - 2

Good afternoon so we will continue again the hypothesis testing part that we were discussing yesterday in the context of multiple linear regression model and the last hypothesis, if you recall what we were discussing was testing the validity of linear restriction using our traditional F statistic. So, I will just quickly recap what we were discussing in the context of a production function. So, I will quickly summarize what we were discussing yesterday. So, I will just quickly write.

(Refer Slide Time: 1:09)



So, our model was this was log Yi equals to beta 0 plus beta 1 lnX1i plus beta 2 lnX2i plus ui. This is the model where X1i is labor and X2i is for capital and the particular test what we were doing was our null hypothesis was H naught beta 1 plus beta 2 equals to 1. This is the null hypothesis that we were testing and what is the implication of this particular null hypothesis? Beta 1 plus beta 2 equals to 1 indicates this particular production function exhibits constant returns to scale.

So, this is what we are testing and our alternative hypothesis was beta 1 plus beta 2 not equal to 1. This is our alternative hypothesis. And what is the approach that we are discussing? We are discussing about the restricted least square approach that means we will impose this restriction in this production function and we will derive the restricted form of this production function and then we will estimate the restricted version of this production function using the same OLS method. That is why this method of applying OLS into the restricted version of the production function is called a restricted least square approach.

So, if you impose the restriction, beta 1 plus beta 2 equals to 1 in this production function, you will get beta 1 equals to 1 minus beta 2. 1 minus beta 2 or beta 2 equals to 1 minus beta 1 as we discussed yesterday, we will get two versions of the restricted model from this original production function but for economies the more appealing version is this where log of output by labor which is X1i equals to beta 0 plus, let us say beta 1 log of X2i divided by X1i capital labor ratio plus Ui. So, this is the restricted version of the original production function and we are applying OLS in this equation.

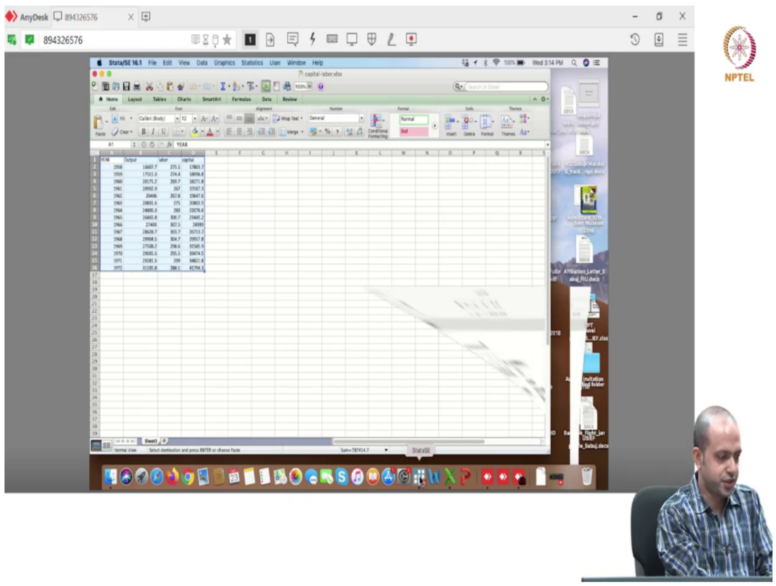
This is the restricted model and applying OLS into the restricted model is called restricted least square approach. So, what is the approach that we are following this is called estimating this restricted form of the production function using OLS method is called restricted least square approach. Restricted least square approach and the F statistic what we are constructing was this is the unrestricted version. This is the restricted model so this is called unrestricted model.

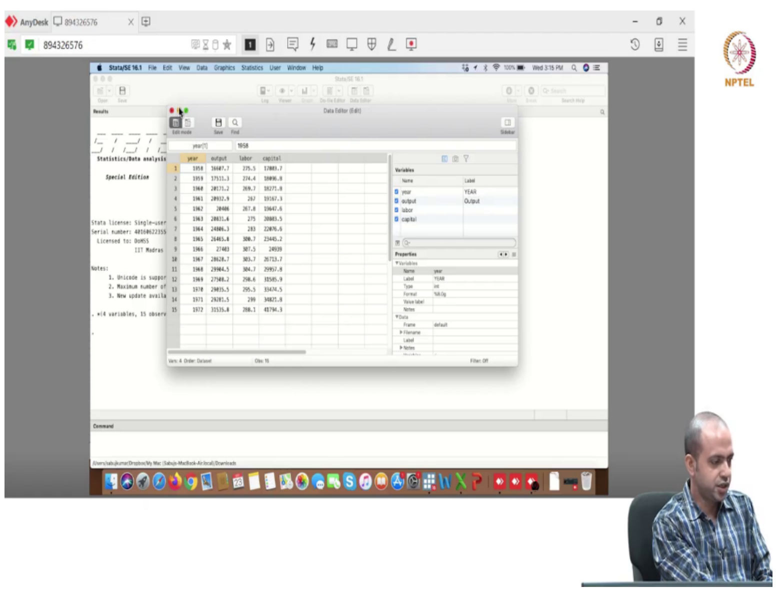
This is the restricted version, so we will estimate both unrestricted and restricted model and then we will collect the RSS, Residual Sum of Square from both the model. Then our F statistic will become RSS. RSS restricted minus RSS unrestricted divided by m, what is m here? m is the number of linear restrictions and what is the number of linear restrictions? This is the restriction we are going to impose and that is the reason m equals to 1 here.

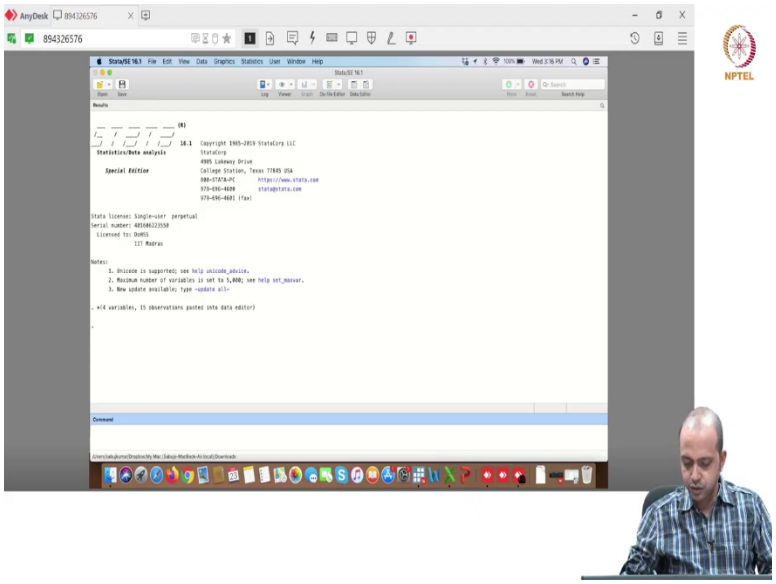
m equals to 1 and in the denominator there would be RSS unrestricted divided by n minus k where k is the total number of parameters to be estimated from the unrestricted model. That is the formula. So in yesterday's lecture I made a small mistake here in the denominator I made 1 minus RSS R. So instead of 1 minus RSS R, that is actually RSS UR. So, this is the corrected version of the F statistic.

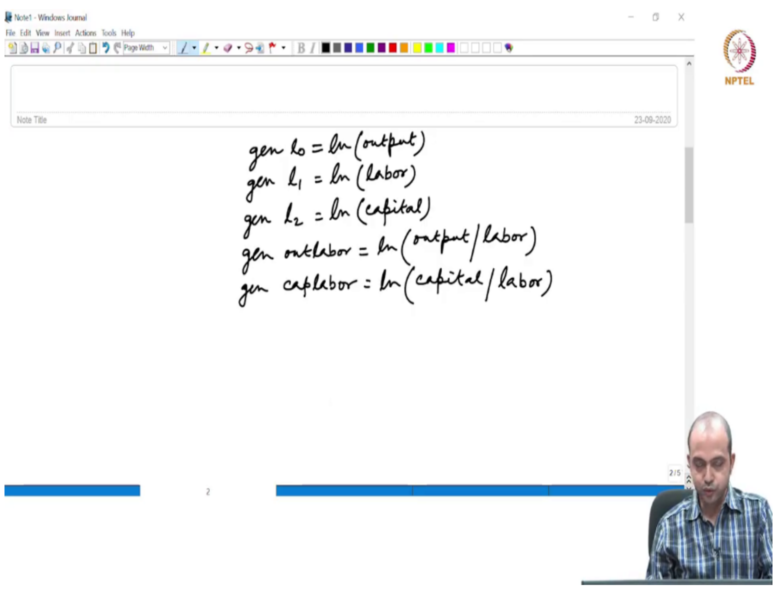
So, please make the correction noted. Numerator should be RSS R minus RSS UR divided by m and this is RSS UR, the denominator is RSS UR, RSS from the unrestricted model divided by n minus k, where k is the total number of parameters in the unrestricted model. We have this formula and now I will quickly go into the software part. I will be again using the statistical software and the same data set to estimate the model RSS R, RSS UR and then we will see what is the value of that, right.

(Refer Slide Time: 9:18)







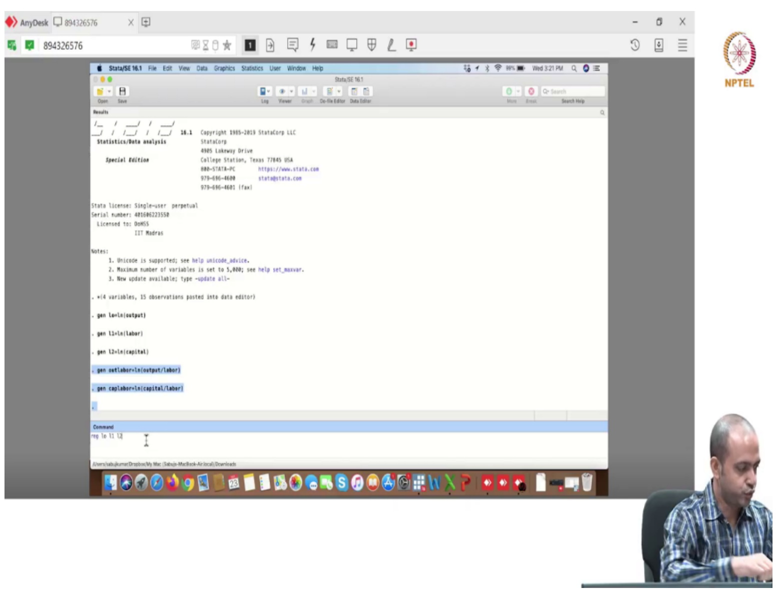


So, we will now go to the statistic. So, here I will first go to the data part. Yeah, so what I will do, I will open the same data set showing the production function, same data set I will be using. This is my data set where I have renamed the variable I have made output labor and capital, these are the three variables and I have data from 1958 to 1972. This is a time series data I have and then I will copy this and I will now put this into statistical software stata.

I will go to data editor and then I will put it paste the data and once you try to paste it then stata will ask you whether the first row of your data set should be considered as variable name, you have to select that part. So, now the variable is taken and you can work on that. So, now as I said yesterday, that our variables are in level form all the variables output, labor and capital everything is level form but we are trying to estimate a model which is in log form. So, that means left hand side dependent variable, right hand side all the independent variables should be converted into log form and there is a specific command that I am going to use for converting the variable from level form to the log form and what is the command that I am going to use? The command is I will first write the command here, the command is gen and you have to put a name. So, when you are converting the output variable, you can put any name let us say I am giving output as, let us say L0.

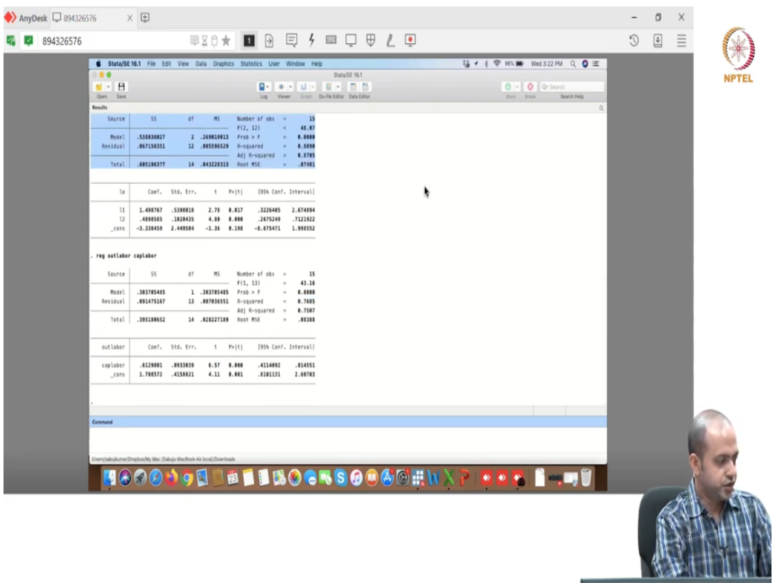
This is my variable name and how I am defining this variable? This is Ln then in the bracket output. Similarly, L generate gen L1, this is the variable name equals to Ln labor and then gen L2 equals to Ln capital. This is the stata command to convert the variable from log form to the level form to log form. Similarly I have to, for the restricted version, I need to generate two additional variable that is gen, let us say, output divided by labor which I am giving name as out labor equals to log of output divided by labor and gen cap labor equals to log of capital divided by labor. So, using this gen command you can easily convert the variables from level form to its log form. So, now I will go back to stata once again and I will estimate the model.

(Refer Slide Time: 13:45)



So, what is the command that I am using? Gen you have to say gen L0 which is equals to Ln output. So, I have generated variable output into log form. Then gen L1 equals to log labor. Then gen L2 which is equal to log of capital. Then gen output per unit of labor which I have given name as out labor, equals to Ln, in the bracket output per labor. And then gen cap labor, capital per unit of labor equals to Ln capital divided by labor. So, this is how I have named the variable. Now, you can run the unrestricted as well as restricted version of the model. And what is the unrestricted version? It is simply a rig then your dependent variable is now L0 equals to this is actually not 0 this is OLO, then your independent variable is L1 and L2.

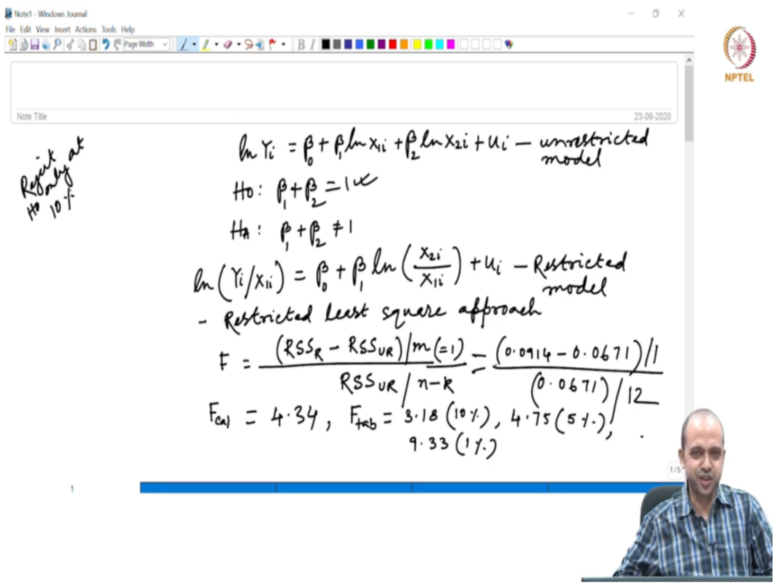
(Refer Slide Time: 16:15)

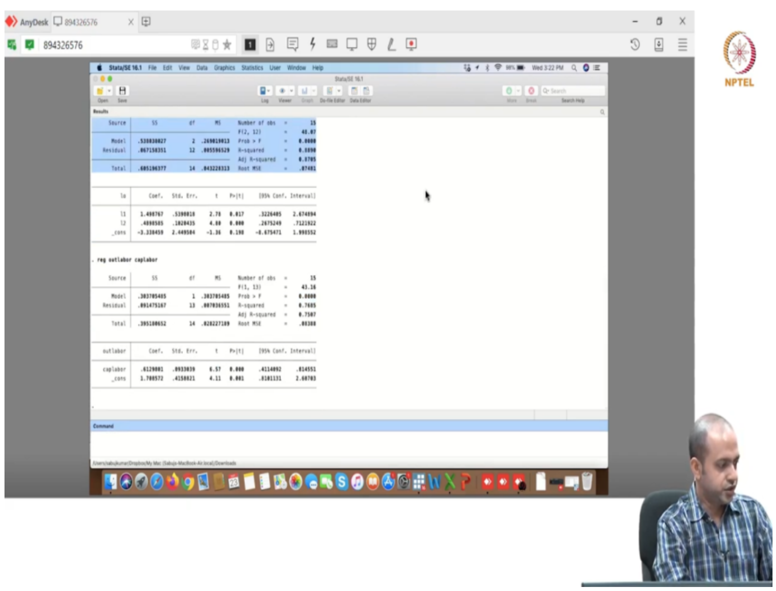


So, this is your model okay. This is the model you have estimated, this is the unrestricted version and from unrestricted version what you have to look at is the RSS. We have to use the RSS from the unrestricted model. So, if you look at the residual indicating RSS which is 0.0671. Now, you run the restricted version reg, what is the name I have given if you look at out labor output part labor out labor.

This is the dependent variable and independent variable is cap labor. Cap labor, this is the restricted version and from the restricted version you can look at the RSS is 0.0914. So, these are the two things you have to keep in mind and total number of observations is 15. So, that is why when you are constructing the F statistic, if you look at in the numerator you have RSS R minus RSS UR divided by m, m equals to 1 here and the denominator is RSS UR divided by n minus k where K equals to total number of parameters in the unrestricted model which is 3 capital labor and 1 intercept. So, 15 minus 3 equals to 12. So that you have to keep in mind. So, what I will do now, I will use this formula to construct this test statistic and how I will do that? So, I will now go to the formula that we were using.

(Refer Slide Time: 18:52)





So, in this formula look at, so here what I will do, I will now use RSS R. What is RSS R? RSS R by RSS UR, so that would become RSS R is what is the value we got, RSS R, the restricted version is 0.0914 divided by 1 because m equals to 1 and what is RSS UR? RSS sorry RSS UR is 0.0671. 0.0 sorry, I made a mistake here.

So, 0.0914 minus 0.0671 divided by 1 and numerator would be unrestricted, which is 0. What is the unrestricted RSS? 0.0671 and this should be divided by n minus k, where n is 15 and k is 3. So this is 12 and if you compute this, then you will get a value which is 4.34. So, this is your F which is calculated. So, calculated F is 4.34 and now what you have to do?

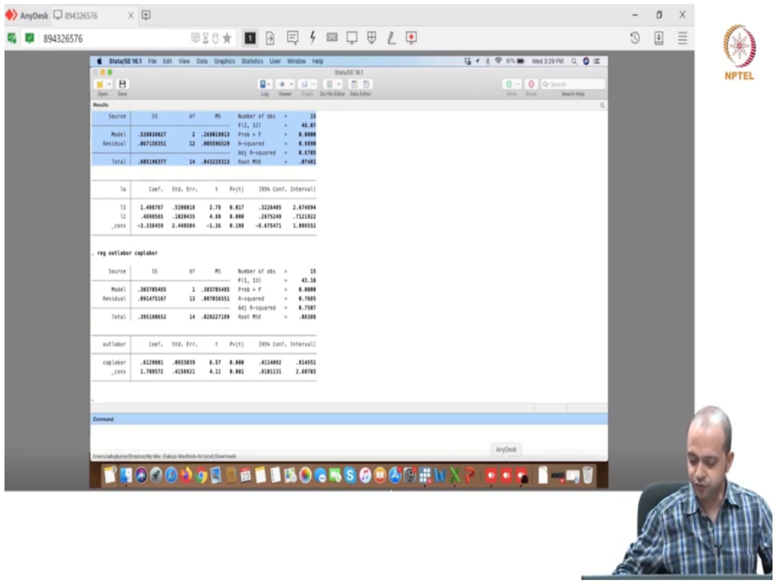
You have to look at the value of F from the F table at one percent, five percent and ten percent level of significance and if you look at then you will see that F tabulated, F tab equals to and what would be the degrees of freedom? Please keep in mind, the degrees of freedom would be 1and 12. So in the F table if you specify N1 equals to 1 and N2 equals to 12, then you will get the value for a specific level of significance. So, the value of F is 3.18 at 10 percent, then you have 4.75 at 5 percent level of significance and 9.33 at 1 percent level of significance.

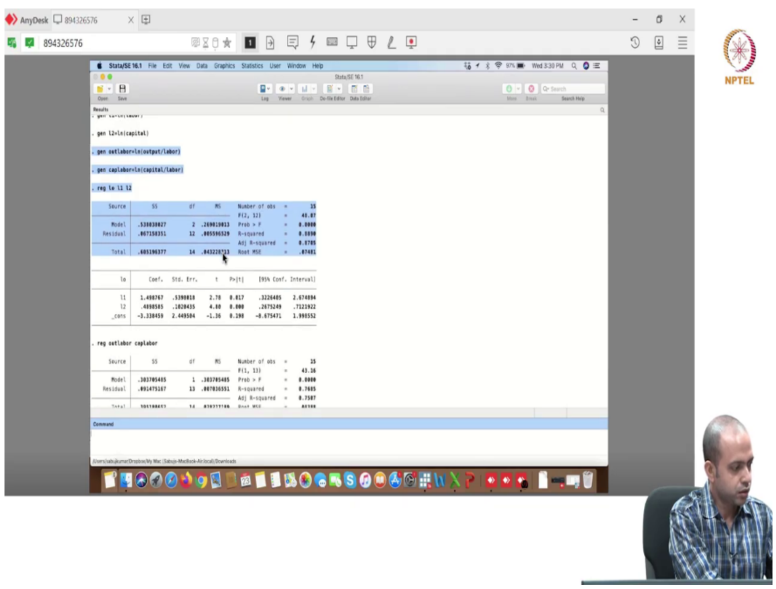
So, now if you compare the calculated value of F with that of the tabulated value of F, you can easily understand that calculated F is greater than the tabulated F only at 10 percent level of significance. So, that means if you consider alpha equals to 10 percent then only your calculated value is greater than the tabulated one. But for 5 percent and 1 percent, it is always lower. So, that means what would be your conclusion?

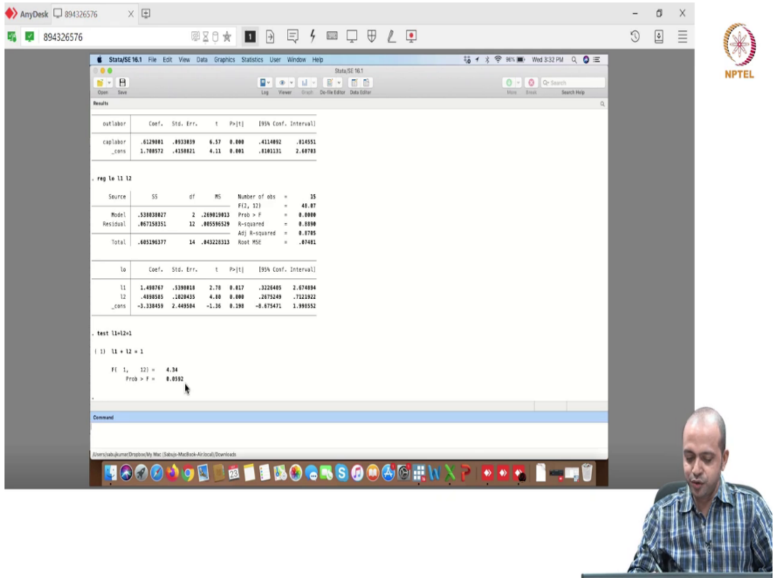
You can reject your null hypothesis. So, the conclusion is reject your null, reject H naught only at 10 percent. But as we have discussed earlier, generally econometrician, they generally do not advise to reject your null at 10 percent level of significance. It is generally recommended that we should take our decision at 1 percent and 5 percent level of significance.

So, we can say that at 1 percent and 5 percent level of significance. What is my null? Beta1 plus beta 2 equals to 1, so that means we come to a conclusion that this production function exhibits constant returns to scale. That is how you have to construct the F statistic and conduct the test manually.

(Refer Slide Time: 24:32)





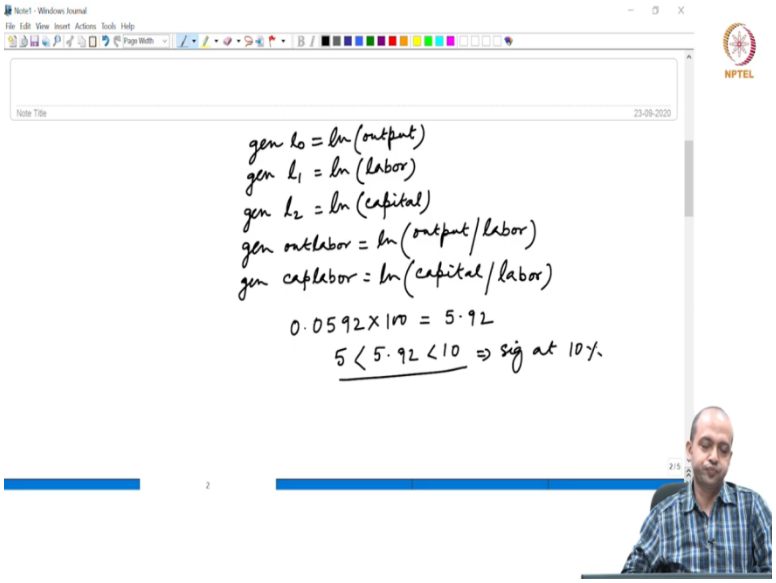


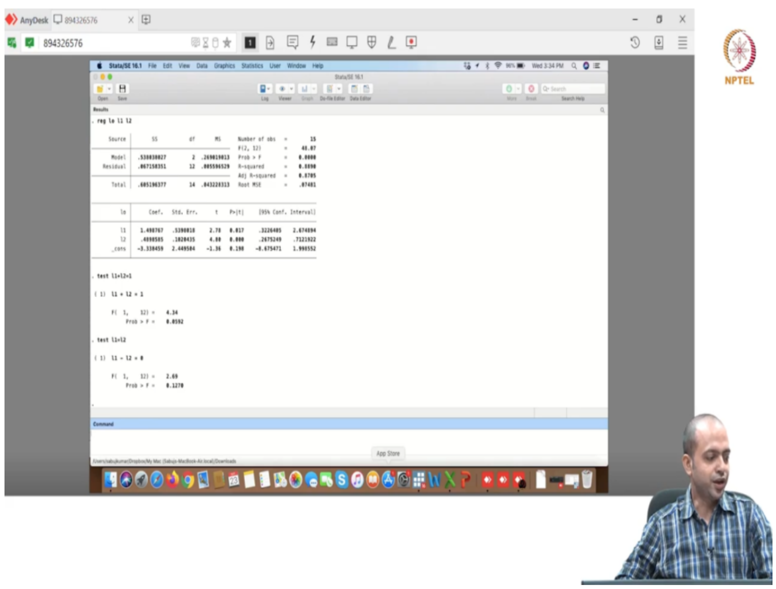
But in stata, there is one inbuilt option, you do not have to calculate all these things, in stata there is an inbuilt option for conducting this particular hypothesis testing and for that you have to use one specific command. I will just copy the command once again, so just run the model once again. Reg LO L1 L2 this is your model and immediately after running your complete model, if you specify the command which is test. This is a test command, test what you are testing?

L1 plus L2 equals to 1 that is what you are going to test okay. L1 plus L2 equals to 1 and if you put enter, then you will come up with this particular test. Now, look at the calculated value of F which we calculated manually and the calculated value of F what stata is supplying exactly matching. So, this is also 4.34. So the stata is giving you the ready-made result for the test statistic F, if you use this test command.

Test command is a very very powerful command. So, please remember this command whenever you are applying F test. Please remember in stata, you have to always use only the test command. And the advantage of conducting this test using stata is that you do not have to compare the tabulated value of F with the calculated value of F rather, F is also supplying is this value. Probability greater than F which basically says how many type 1 error you are committing while rejecting your null hypothesis. And from here look at the value, the value is 0.0592. And what is the procedure? If you remember we have to multiply this value with 100 and if you multiply, then this would become 5.92. Since 5.92 is greater than 5 but less than 10.

(Refer Slide Time: 27:21)





So you multiply 0.0592 by 100, then you will get 5.92. Now this 5.92 is less than 10 but greater than 5. So, that is why this implies significant at 10 percent only. So that means you can reject your null only at 10 percent. This is how you can conduct the important test of linear restriction using F statistic in stata. Now, suppose you want to test whether the returns from labor and returns from capital they are same or different.

They are same or different that also can be conducted using the same test command and in that case, the test command would be again T-E-S-T whether L1 equals to L2 or not, this is the command. L1 equals to L2 and if you put enter then look at what is happening? It is not the value of F is 2.69 and P greater than F is 0.1270. If you multiply that by 100 it is 12.70, so I cannot reject my null hypothesis.

I cannot reject my null hypothesis and what is my null hypothesis? Returns from labor and returns from capital they are almost same. There is no significant difference in the returns to capital and returns to labor. That is what you can say. So, in this way you can apply the F statistic to test the equality between two regression coefficients, the equality between two regression coefficients. Now, after this we will discuss one more application of this F statistic, one more hypothesis testing that we are going to conduct using the same F statistic.