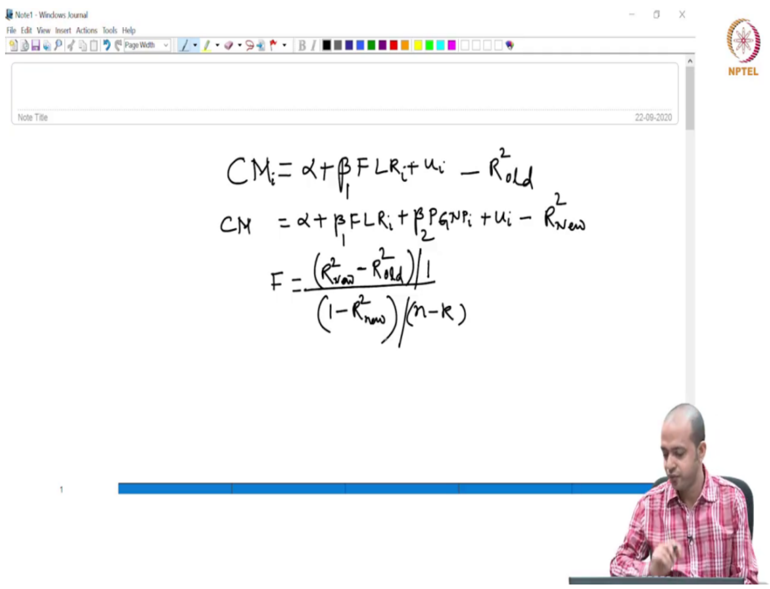
Introduction to Econometrics
Professor Sabuj Kumar Mandal
Department of Humanities and Social Sciences
Indian Institute of Technology, Madras
Lecture – 27
Multiple linear regression model and application of F statistics Part - 6

So, we were discussing about hypothesis testing in the context of multiple linear regression model. And so far we have discussed testing three types of hypothesis here. First one was the first hypothesis that we were testing was basically the marginal contribution of an explanatory variable.

That means when you know that one variable is already there in the model if I add an additional explanatory variable then what would be the additional contribution of that variable in the explanatory power of the model. And I will just quickly recap that model and for your own for your understanding.

Then I will be using the same data and I will estimate the model and I will also show you how to conduct that hypothesis testing. So, here we are interested in child mortality rate as a function of female literacy rate and PGNP as you know. So, let us say that our initial model is CM.

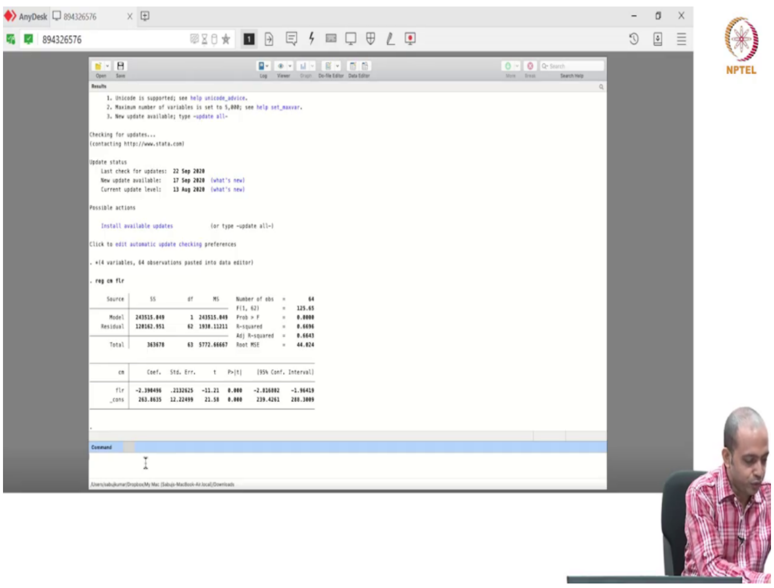
(Refer Slide Time: 01:30)



CM equals to alpha plus beta1 FLR plus ui. We call this model as old. And from estimating this model we will get let us say R square old. And then when we add the PGNP variable in the model then that will become CM equals to alpha plus beta1 FLR plus beta2 PGNP plus ui.

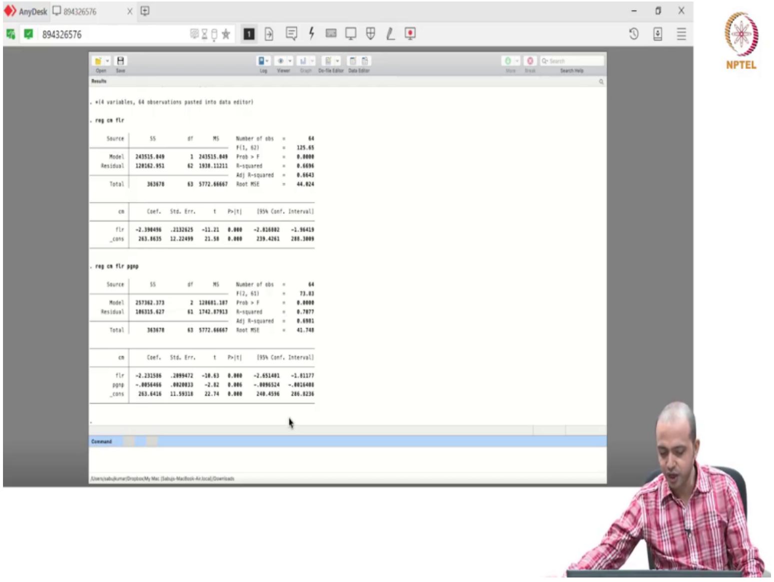
And if you estimate this model you will get another R square which is let us say R square new. And then how we have defined our F statistic? We define F statistic as R square new minus R square old divided by number of new explanatory variables or number of new explanatory variables added in the model which is 1 here. Which is 1 divided by 1 minus R square mu divided by n minus k where k is the total number of parameters to be estimated in the new model, k is the total number of parameters to be estimated from the new model. Now, we will estimate the model and will put the values of R square new, R square old, n and k in this function then we will get the calculated value of F.

(Refer Slide Time: 04:10)



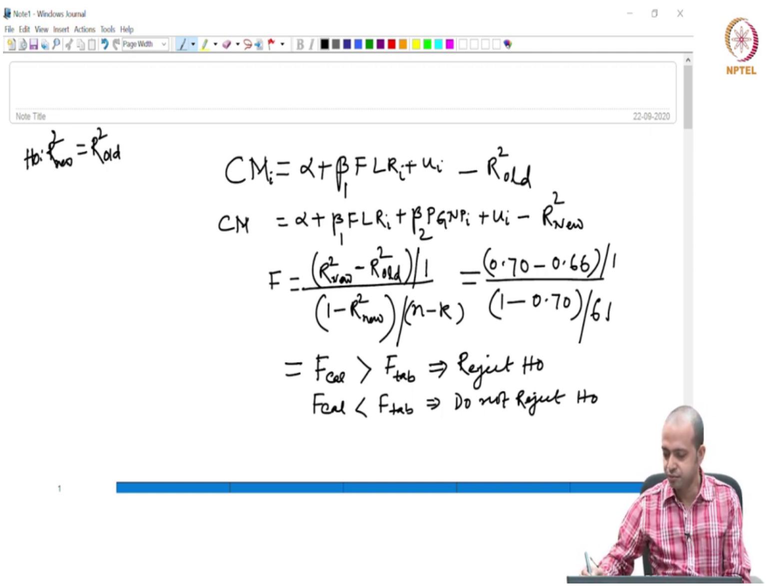
So, I will now show you how to estimate the model. So, here what you have to do just I will run the first model reg CM on FLR and if you look at the R square value is point 0.6696 0.6696. So, after this I will run the complete model.

(Refer Slide Time: 04:50)



Which is a reg CM FLR and then pgnp and this should become your complete model or the new model. And you will get the R square new as 0.70. So, what I have to do now is to use this 2 R square values and put it into the function what we have defined earlier. Alright.

(Refer Slide Time: 05:24)



So, that means you have to say that here this would become 0.70 minus 0.66 divided by 1 into 1 minus 0.70 divided by n which is actually if you look at your total number of observations is 64 and total number of parameters to be estimated from the model is 3 so 64 minus 3 is 61.

So, here I will put n minus k that means 64 minus 3 is 61. Alright. So, once you do that you will get the calculated value of F. And that calculated value of F you have to compare with the tabulated value of F which is available at the end of your textbook. For a specific level of alpha equals to let us say 1 percent, 5 percent or 10 percent.

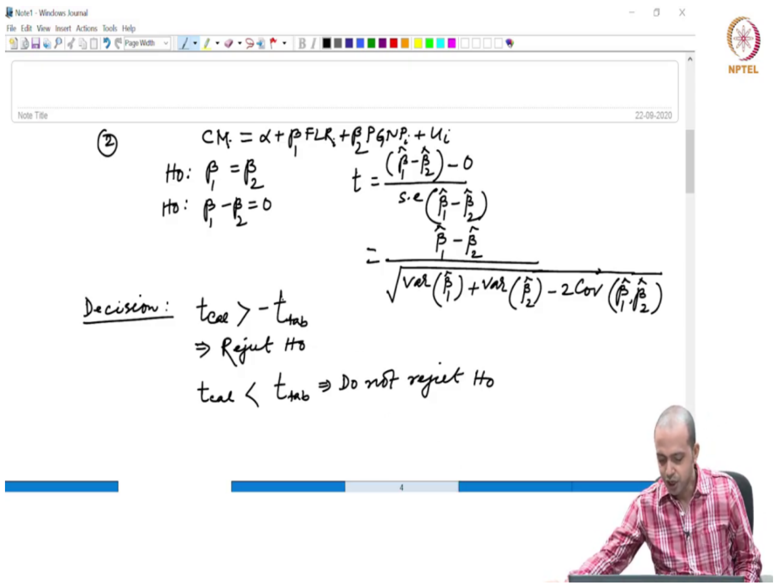
So, you can easily calculate so this would become your F. I will write F which is calculated. And then you have to compare this calculated F with the tabulated F. And if this is greater than F tabulated then what will do your calculated value is greater than the tabulated value you have to reject your null, reject your H naught.

And what was your H naught? If you remember our H naught was R square new equals to R square old. So, that means there is no significant difference between the new R square and the old R square so alternatively we can say that the new explanatory variable which is added in the model has insignificant explanatory power.

And if the F calculated is less than F tabulated then what would be your decision- do not reject the null. This should be your decision. Alright. So, you have to calculate the F value and compare this with the tabulated one. This is how you can test the hypothesis whether R square new equals to R square old.

So, that means you can justify adding the new variable PGNP in the model if and only if you can reject this null hypothesis- R square new equals to R square old. You have to reject this null hypothesis then only you can establish that yes, the new variable PGNP has significant explanatory power. So, this was the first hypothesis we tested yesterday we are talking about.

(Refer Slide Time: 09:15)



And what was your second hypothesis? Second hypothesis that we were testing in the context of a production function or here in this context also you can check it when your model wasCM equals to alpha plus beta1 FLR plus beta2 PGNP plus ui. So, what was the second hypothesis we are testing? second hypothesis is add equality between two regression coefficients. So, that means our null hypothesis here we were testing beta1 equals to beta2. That is our null hypothesis that we are going to test.

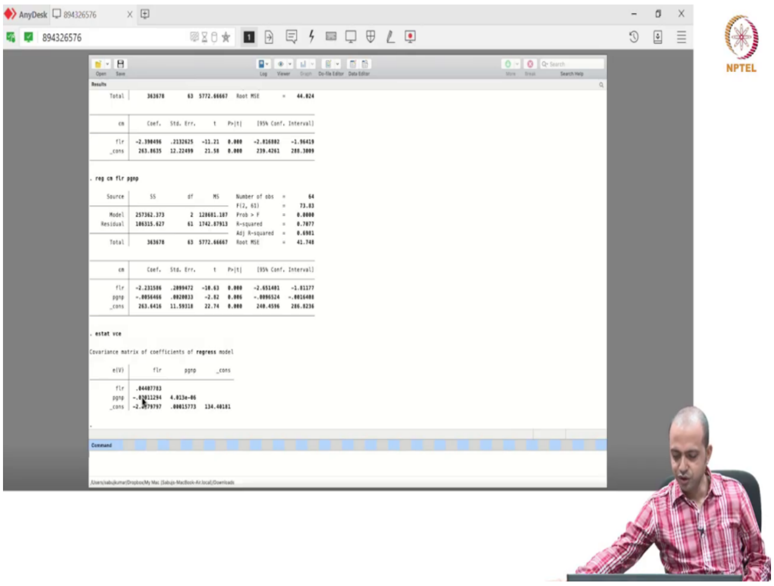
And then how we are testing? how we have constructed the test statistic? this can be alternatively written as H naught beta1 minus beta2 equals to 0. And how we have defined t? t was defined as beta1 hat minus beta2 hat minus 0 actually divided by standard error of beta1 hat minus beta2 hat.

And what is standard error of beta1 hat? that is basically square root of variance of beta1 hat plus variance of beta2 hat minus 2 covariance of beta1 and beta2. So, that means this can be written as beta1 hat minus beta2 hat divided by variance of beta1 hat plus variance of beta2 hat minus 2 into covariance of beta1 hat beta2 hat.

This is how you have to calculate your t. How will you take your decision if t calculated is greater than t tabulated? that implies that you have to reject your null. So, this will give you t calculated value and you have to check the tabulated t at a specific level of alpha either 1 percent, 5 percent or 10 percent.

Then depending on the value of t tabulated you can easily decide whether to reject your null hypothesis or not. And if t calculated is less than t tabulated then what would be your decision? do not reject your null, do not reject your H naught. So, this is the simple rule of decision rule of your hypothesis testing. Now, from this model what I will do I will now show you how to implement this particular hypothesis testing using the software.

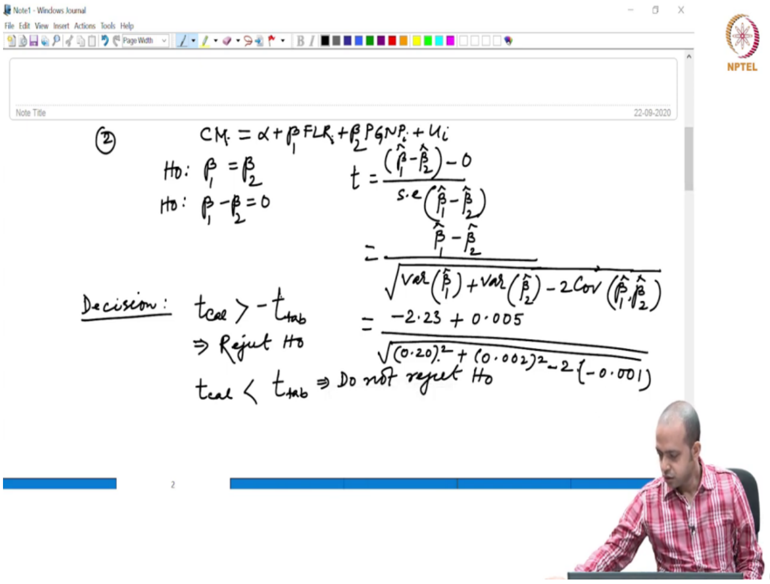
(Refer Slide Time: 13:56)



So, I have already run the complete model. So, from here see I have everything. so variance of beta1 hat, variance of beta2 hat, you have standard error so if you square it up that will give you variance of beta1 hat. That means variance of this FLR, variance of beta2 hat which is nothing but square of this. And then you have to take 2 into covariance of this and this. Now, only covariance part is not given here that you have to additionally get. And for that we will be using a specific STATA command and you have to remember the command for getting the covariance for this two variable FLR and PGNP coefficient.

And what is that command? the command is estat vce. So, immediately after running the complete model if you put this estat vce command then that will give you this type of result. This is called the variance covariance matrix and from here what you have to look at this PGNP and FLR this is the covariance. This is the covariance 0.000001 so, minus triple zero one is basically the covariance of FLR and PGNP. So, this you have to put it into the formula so that means from here I will take square root of this then square root of this plus 2 into this value 2 into this value. So, what I will do square root of

(Refer Slide Time: 15:59)

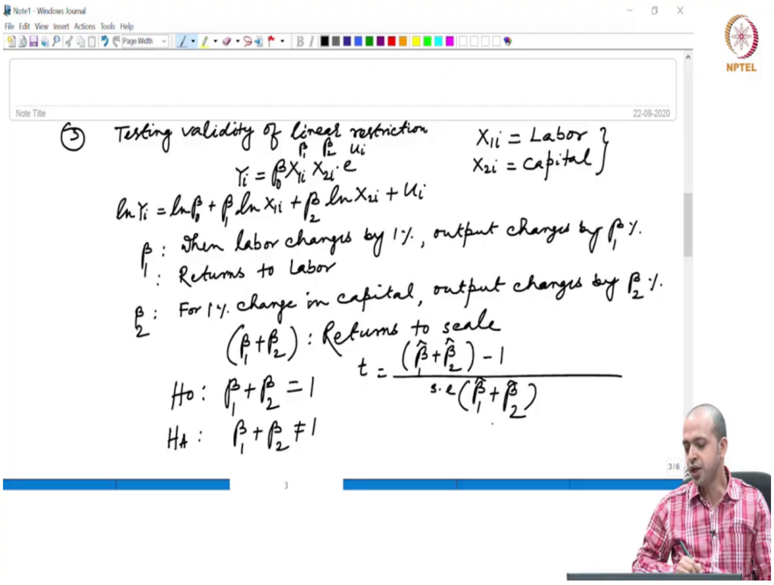


So, that means beta1 hat minus beta2 hat in the numerator and that is beta1 hat is basically -2.23. Then again minus of this so it will become +0.005 so -2.23 plus 0.005.

I think that is the value 0.005 yeah and then divided by divided by 0.20 whole square plus plus 0.002 square plus 2 into 2 into sorry this is minus actually minus 2 into 2 into 2 into what is the value you got there covariance is -0001.

So, we have 1 minus here and if you put minus 0.0001 this is how you have to calculate your t. And this calculated t should be compared to the tabulated value of t at a specific level of significance alpha equals to 1 percent, 5 percent or 10 percent. So, this is our second hypothesis. Wherein, we were testing about the equality between 2 regression coefficients. And then what is your third hypothesis?

(Refer Slide Time: 18:44)



We are testing about the third hypothesis that we are testing. That was testing the validity of linear restriction. So, please remember what we are testing- testing validity of linear restriction linear restriction linear restriction. So, this particular linear restriction we can discuss in the context of a production function that would be better it would be more meaningful.

So, what I will do I will take you to the production function that we are discussing yesterday. The production function what we are discussing was yi equals to beta0 then beta sorry beta0 it was beta0 into x1i to the power beta1 then x2i to the power beta2 into e to the power ui.

And after linearization so this is the production function and after linearization what we got log if you take log of both side log yi equals to log of beta0 plus beta1 log x1i plus beta2 log x2i plus ui. This is our linearized production function and where we said that x1i equals to labor and x2i is capital. How we have defined here x1i equals to x1i equals to labor two factors of production and x2i equals to capital capital. So, these are the two factors of production we are using in the production function. And then we also said that since this is a linearized model log log model in this log log model the interpretation of beta1 and beta2 is actually the direct elasticity measure.

So, that means for 1 percent change in labor what would be the change in your output and beta2 indicates for 1 percent change in capital what would be the change in output. So, that means I can interpret beta1 as when labor changes by 1 percent then output changes by beta 1 percent direct elasticity measure direct elasticity measure.

And we have also denoted beta1 as returns from returns to labor returns to labor returns to labor. And then this is beta 2. beta 2 how we have defined what is the what would be the interpretation that means for 1 percent change for 1 percent change for 1 percent change in capital in capital output changes by beta 2 percent.

But, please do not forget to mention on an average and keeping the impact of other factor constraint. So, that means when I am saying that interpretation of beta1 when labor changes by 1 percent output changes by beta1 percent keeping capital constant. Similarly, when capital changes by 1 percent on an average output changes by beta 2 percent keeping labor constraint. That means I am not changing labor when I am changing capital. And these two are when you keep the other factor constant, we said that that is basically returns to factor.

And when both the factors change by 1 percent then the relevant concept is returns to scale. So, that means we said that beta1 plus beta2 basically indicate returns to returns to scale. And what we are testing our claim is that this particular production function let us say exhibit either increasing returns to scale or diminishing returns to scale.

That is our claim. And if you nullify your claim then your hypothesis null hypothesis should be beta1 plus beta2 equals to 1 this is your null hypothesis. And alternative hypothesis would be what would be your alternative hypothesis would be beta1 plus beta2 not equals to 1. So, that means this consists of two things it can be either greater than 1 in that case we will say that production function exhibits increasing returns to scale. It can be less than 1 also and in that case we will say that production function exhibits diminishing returns to scale diminishing returns to scale.

And how we have formulated the test statistic here the test statistic was formulated as how you have defined our t so when this is your null hypothesis, we can say that t was defined as beta 1 hat plus beta 2 hat minus 1 divided by standard error of this standard error of beta 1 hat plus beta 2 hat.