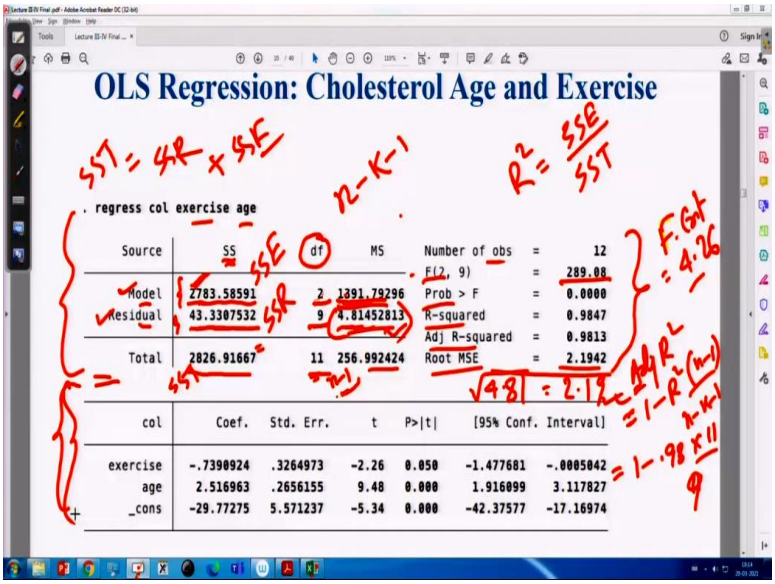
**Applied Econometrics**
**Prof. Tutan Ahmed**
**Vinod Gupta School of Management**
**Indian Institute of Technology - Kharagpur**

**Module - 6**
**Lecture - 50**
**Regression Table**

Hello and welcome back to the lecture on Applied Econometrics. So, we have been talking about multiple regression. And in the previous lecture, we have seen something called Simpson's paradox, and we run a regression where we had our independent variable age and independent variable exercise, whereas our dependent variable was the cholesterol level. Now, we have done that, and we kind of explained that table also.

But we have seen, when we are actually doing that regression table, we have seen, there are many elements in the regression table. And we decided that we will actually talk about these different elements in the regression table in the next lecture. So, this is what we are going to do in this lecture. We will basically talk about the different elements in the regression table.

**(Refer Slide Time: 01:07)**




Now, let us take a look at the regression table. We see that we have this multiple regression where we have actually other 2 components in the regression table. So, one component, we have here, this part, which is essentially an ANOVA, and we will explain that. And the second part, which is more important to us, because we will talk about the coefficients here, is essentially the hypothesis testing for the coefficients.

So, let us first start with the ANOVA part; so, the first component of the regression table. So, let us see what we see here. So, we see that this SS; SS is sum of square. And this is model; this is residual; and there is a number here; there is a number here. So, what these things are actually? Now, if you go back to the concept of regression and where we actually drew the regression line, you remember that we had this sum of square explained by the model.

So, you have that Y bar and Y hat; so, that difference was something that is explained by the model. And that part is essentially, is something that kind of tells you how well is your model ease; I mean, if it is able to explain lot of these variations, so, then your model is good; but if it is not, then perhaps, we will have to rethink; whereas, the residual is essentially the difference between the actual Y i and the Y hat.

So, what my model is not able to explain, the real data, that part, which part is not explained by the model, is the residual, residual sum of square. So, these 2 components are essentially talking about that. So, this part is sum of square which is explained by the model. And this part is sum of square which is not explained by the model, which is residual. So, we can also say, sometimes we have used this notation sum of square explained, and this is sum of square residual.

So, this is essentially whatever we have actually seen when we drew that line and we divided this error terms into different components, essentially, that is what exactly represented here. Now, that is the first part. Second part is the degrees of freedom. And we explained in detail the concept of degrees of freedom and how this concept actually varies, or actually how we can apply this idea of degrees of freedom in different contexts.

So, when it comes to model, you have to be very carefully here; when it comes to model, we are actually taking the number of explanatory variable as the degrees of freedom. So, I have my 2 explanatory variable here, exercise, age; and my degrees of freedom is equal to 2. When I am talking about the residual, this is actually coming from the number of observation n - k is a number of explanatory variable -1.

And this one, for the total, for the entire, if you just do not consider all the model residual, but if you just consider the data set in one go, so, it will be like n - 1. Now, the point is, why these degrees of freedoms are different, and why in different contexts, we are using different

expression for degrees of freedom? I have explained that. When we talked about the concept of degrees of freedom, we explained how it actually varies from context to context.

I will briefly touch on that. We will do a little bit of recap on that. So, for the model, we have n - 1. And the reason here is that; why I have to have this 1 constant? When we talked about degrees of freedom, we said that my mean is constant, let us say, or my variance is constant, whichever property I want to keep constant, keeping that constant, I am actually trying to vary the different observations. Now, why I am doing that?

The reason I am doing that is, it is a estimation problem, right? So, it is something that I am trying to go from sample to population. So, and we know when we talked about the properties of estimator; so, we know that the best estimator for a sample mean or a population mean is the expected value of the sample mean. So, basically, population mean is equal to sample mean.

So, when we actually assume that, because you are assuming it is an estimator, which is essentially, it is an unbiased estimator, so, if the estimator has to be unbiased, I have to have my expected value of sample mean has to be equal to the population mean. Now, the moment I fix it, so, to ensure that, that property is satisfied, I have to keep one, this parameter constant, and which is your mean.

And the moment I keep the mean constant, I will have only n - 1 degrees of freedom. And model, we have explained that it will have number of explanatory variable, because that is the number of way, number of different dimensions you can explain the variability of your model. And residual, we explained in that plane, how depending on the number of observation and number of explanatory variable, you have your degrees of freedom.

So, the concept remains same, but the expressions are different because they are explaining different things. Now, when I want to get the equivalent variance; so, when we get variance, we normally divide by n, because we try to see the variation par observation, right? So, we actually take the sum of square of the errors, and then we divide by n, because we want to see the variation par unit of observation.

Now, here; so, in that case, I do not have any constants, so, I am able to divide it by n; but here, I have some constant, right? There is a limit to what extent I can vary things. Now, for model, I can vary, like, I can have 2 variables. So, I can actually; so, k explanatory variables, I can actually vary in k ways. That is why I can divide it by k. So, it is equivalent to variance, but you divide it by k.

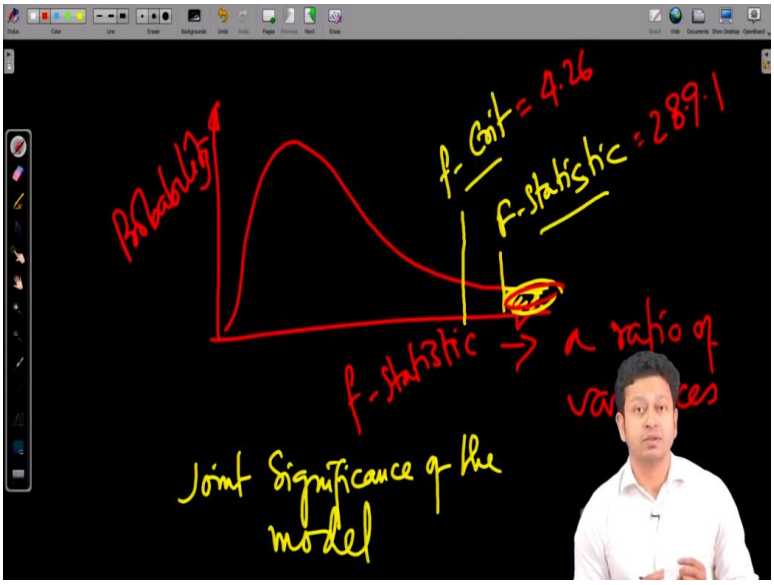
Whereas, for residual, you have n - k - 1 way you can vary the residual, not more than that, because we are, you remember the plane where you have this constant, so, you cannot go beyond that. And that is why you have to divide it by n - k - 1. When you do that, you got this mean values, mean squares. And now, these are the, essentially, you divide this 2783.58 by 2 here. And here, you divide 43.33 by 9 here. And here you get 1391.79. And here you get 4.81.

So, that is what you get. And here also, you get this; you essentially add these two things up, these two sum of square. And we know that SST is essentially equal to SSR plus SSE. And this is nothing but my SST. And this is my, as I said, SSE; and this is my SSR. And of course, you divide this one also with the degrees of freedom. Now, we have got the first part of the table.

Now, we have to understand the second part of the table, how we have got all these different estimates of F? What we mean by this? What we mean by this one? What we mean by R-squared? What we you mean by adjust R-squared? What we mean by root MSE? The number of observation is of course given. We have seen in the previous lecture that we are only dealing with 12 observations.

Now, let us try to understand each of these different terms. Let us start with the first one, F. So, F-distribution, if you know, so, F-distribution is a probability distribution where you plot, basically, where you see the probability of F-statistic to take different values. So, how do we really draw it? So, we draw it.

**(Refer Slide Time: 09:32)**

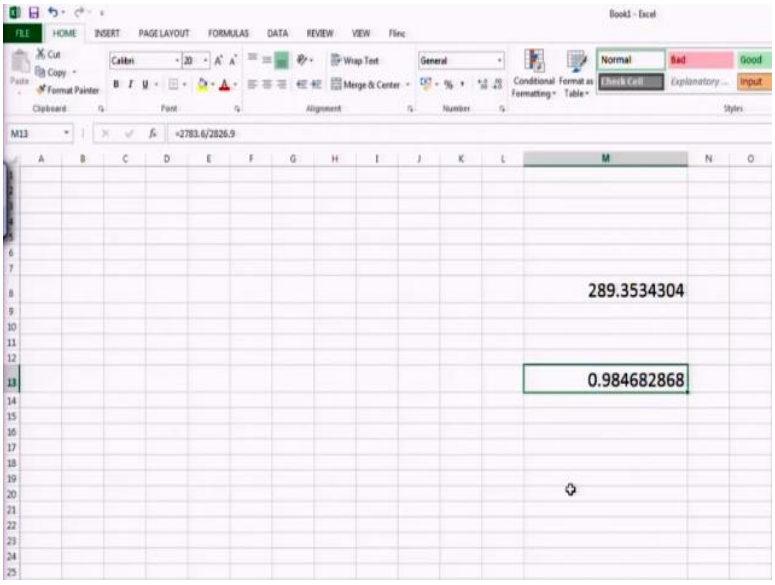



So, usually, we have, the F-distribution looks like this. So, we have this axis. Here, let us say we have probability; and here we have F-statistic value; and we have something like this. And the reason we have this kind of distribution is that, F-statistic is nothing but a ratio of variances. So, in numerator, you have a variance term, in denominator, you have a variance term.

When you are actually getting the ratio of variances; now, variance, we know that there were positive, right? So, we can never have a F-statistic less than 0. It is always positive, because you are taking variances. Now, what variances you take? So, F-distribution, that could be like, in different context, you can take different ratio variances. So, in this case, in this particular, when you are talking about regression, what we do here is, we essentially take the variance for essentially the model and residual, like when they are divided by the respective degrees of freedom.

So, essentially, here, I will take the variance of these two terms. I will take these two terms and I will try to see the ratio of the variances. So, if I do that; so, I have 1391.79. So, let me actually use the Excel sheet to actually do this. So, let us do this.

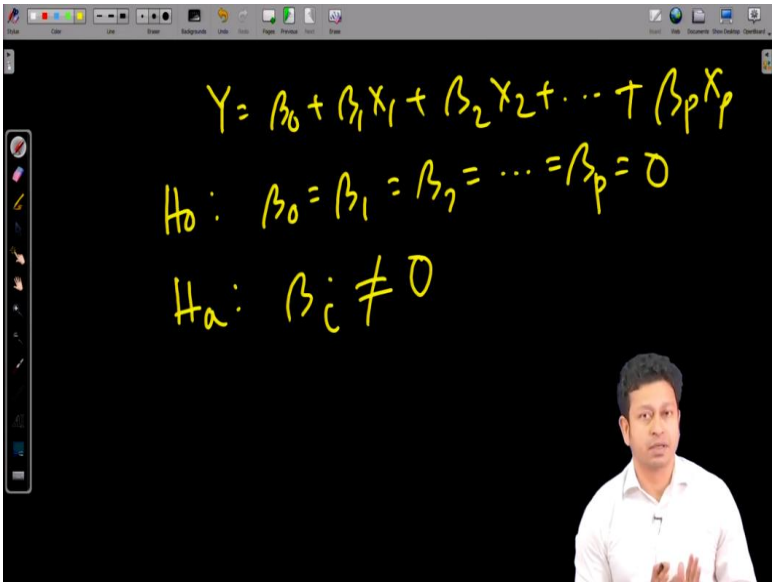
**(Refer Slide Time: 11:22)**




So, I am going to take a variance where I have my numerator equal to 1391.79 by 4.81. And what we will get is 289.35. And if we see here, essentially, it is talking about the same thing. So, you will see that this is essentially, this one, 289.08. So, that is essentially your F-statistic. So, F-statistic is essentially the ratio of these 2 variances. So, the model by the residual. So, essentially, it talks about, to what extent my model is able to explain the variation now.

How do we really interpret F-statistic? We interpret F-statistic, or the way we see the F-distribution, we say that, what we actually do here is that, we try to see if the; like, F-statistics actually measures the joint significance; let me actually write down; let me actually write down, joint significance of the model. That is what F-statistic does. So, now, how do I actually express that?

**(Refer Slide Time: 12:52)**



So, by joint significance of the model; so, let us say I have a model where my, say Y is equal to, say beta 0 plus beta 1 X 1 plus beta 2 X 2, beta p X p. And if that is the model, so, my joint significance of the model would say that H naught would mean that my beta 0 is equal to 0, my beta 1 is equal to 0, my beta 2 is equal to 0, or all my betas are equal to 0, so that my model does not have any explanatory power.

So, I can write it down like beta p is equal to 0. So, none of them have any explanatory power. But the moment you have any of the coefficients with any explanatory power, your model is actually having some explanatory power. So, then you will reject the null hypothesis that your model does not have an explanatory power. So, essentially, your alternative hypothesis is, any of the beta is not equal to 0. So, that is alternative hypothesis.

So, the moment you have any beta, any of this coefficient coming out to be significant, your model has some explanatory power. And the moment your model has some explanatory power, you reject the null hypothesis. And we reject it using your F-statistic. Now, how you do that? So, essentially, if you again go back to this; so, what you do is, so, there is a critical value like any other distribution.

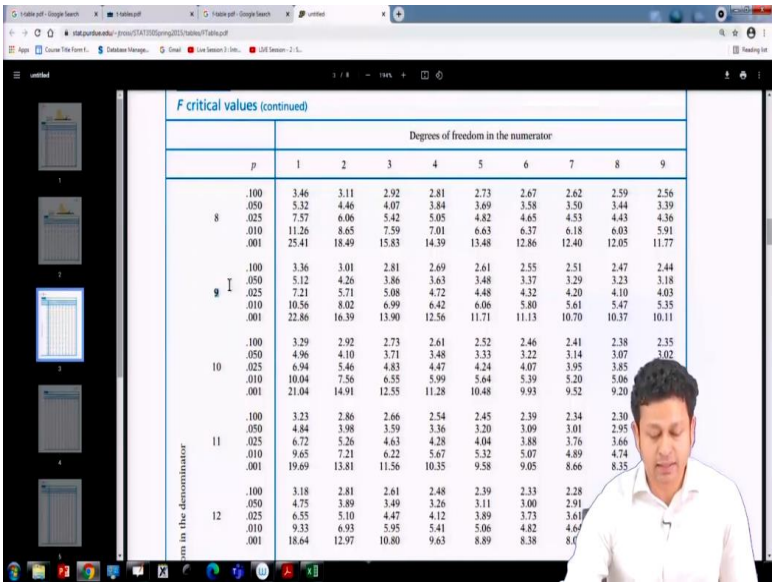
So, your critical value depends on the degrees of freedom and the confidence interval. So, if your F-value, so, whatever F-statistic you have got, if that goes beyond this F-critical, so, then what we say is that, we calculate the P-value. P-value is the area which is right to the F-statistic. It is very slow and we actually reject the null hypothesis because of the fact that; the F-statistic is explaining that your alternative hypothesis is actually falling far away from the null hypothesis.

And so much so that it is going beyond the F-critical, and you are actually rejecting the null hypothesis. So, this is what we will try to see here, whether your calculated, the F-statistic that you observe is actually bigger or smaller than the F-critical. So, we have seen that, here my F-statistic for 2 and 9 degrees of freedom, we have 289. So, that we have observed. So, now, I have to actually take a decision on whether to actually say my model is significant or not.

So, I am talking about the joint significance; I am talking about all the variables together. Now, for that, what I will see is, I will actually go to the P-value, like just what we have

explained, the P-value, we will try to get this P-value. And to get that, what you have to consult is a F-table. So, let us actually see how F-table looks like and how do we actually calculate this.

**(Refer Slide Time: 16:01)**




So, the moment we want to see a F-table, we have to keep in mind that there are 2 different degrees of freedom here. So, for our case, it is 2 and 9; it is given here, it is 2 and 9. Now, for 2 and 9 degrees of freedom, and 2 is the degrees of freedom in the numerator, and 9 is the degrees of freedom in the denominator. So, here it is very clearly given, the degrees of freedom in the numerator and denominator.

So, the degrees of freedom in the numerator is 2 and my degrees of freedom in the denominator is 9. So, we can actually see if we scroll down what is the corresponding F-statistic that we are getting. So, here we have 2 and here we have 9. Now, here, F-critical values they have given; so, let us say it is 0.05, which is like 95%; for 2 and 9, we have something like 4.26. So, my F-critical is 4.26.

Now, if you now go back to our F-table, what we will see is, it is 4.26 and this 289 point something. Now, so, it is very clear that my F-statistic is falling far ahead of my F-critical and then my model; if my F-statistic is far ahead of my F-critical, then I would say my model is significant, because my model has some explanatory power, because the P-value is really low; and it is not really able to, we cannot really attribute to the variations in the model to all randomness.

So, that is what we have to keep in mind; so, how we actually take the decision. So, that is about the F-statistic part. Now, let us look at the other components of the table. Now, we come to R-squared. Now, R-squared, we have already explained. R-squared is rather simple thing here, and it is just a ratio. So, R-squared is nothing but the ratio of sum of square explained by sum of square total, because we know R-squared is kind of talking about how powerful our model is.

So, if my model is able to explain all the errors, all the variations, so, then my R-squared is actually high. So, here, what you do is, you actually take this ratio of these two. So, let us actually do that. So, 2783.6 by 2826.9. And that is my, it is 0.984. And that is precisely the R-squared, 0.984. Now, so, R-squared is straightforward. And we also explained the concept of adjusted R-squared, and that is something we explained why it is important, because you need to, if you will keep on increasing the number of explanatory variable, there has to be some means for actually penalising the model, because, in general, increasing the number of explanatory variable will actually explain the explanatory power of the model.

But it may so happen that your explanatory variable is actually redundant, it does not have really any significance. So, that is why we need to create a penalty for that. And the formula we use for adjusted R-squared, the formula was that; squared is equal to 1 minus R-squared into n - 1 by n - k - 1. So, what happens here is, the moment you keep on increasing your k, what happens is that, this whole adjusted R-squared value actually decreases.

So, if we plot, if we put these numbers here, you will actually get the adjusted R-squared. So, we can do it or you can do it on your own. Let me actually try to do that. So, 1 - 0.98 into n - 1 is 11 by you will have n - k - 1, which is 12 - 2 - 1, which is 9. So, you will get a value for adjusted R-squared. I will not do it. So, the root mean square error is that; the last term is the root mean square error.

And this is equal to; this is essentially nothing but the square root of the error due to the residuals. So, this part. So, if you just do that, square root of the residuals, so, you will find this value. So, if you take square root of 4.81, you will get something like 2.19. So, you get the root mean squared. So, it is important to have root mean squared error, because, what you made here is that, you try to get an idea about the residual variation, par unit observation.

So, you are basically, this is sort of a variance term and this corresponding, like, you can think like it is standard deviation term for each of these observations. So, this is another measure. So, of course, you will want your root mean square error to be as minimum as possible, because, like, that is unexplained part of the error. So, that is basically the first part of the regression table. So, with this, I conclude this lecture. And in the next lecture, I am going to talk about the second part of the regression table, which is the hypothesis testing for the coefficients. Thank you.