Econometric Modelling

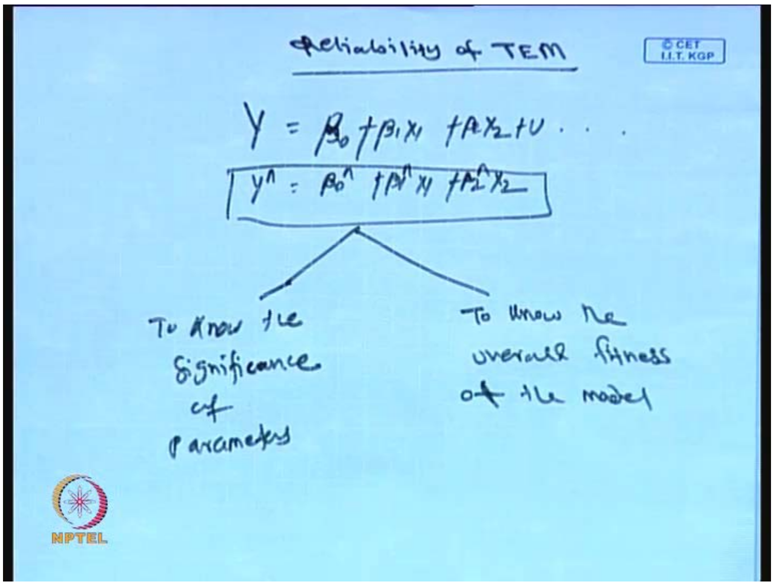
Prof. Rudra P. Pradhan
Department of Management
Indian Institute of Technology, Kharagpur

Lecture No. # 15

Reliability of Trivariate Econometric Modelling

Good afternoon, this is doctor Pradhan here. Welcome to NPTEL project on econometric modelling. Today, we will discuss the reliability of trivariate econometric modelling. In the last lectures, we have discussed the entire structure of trivariate econometric modelling, where the system consist of three variables Y, X 1 and X 2; that means, one dependent variable and 2 independent variables.

(Refer Slide Time: 01:03)

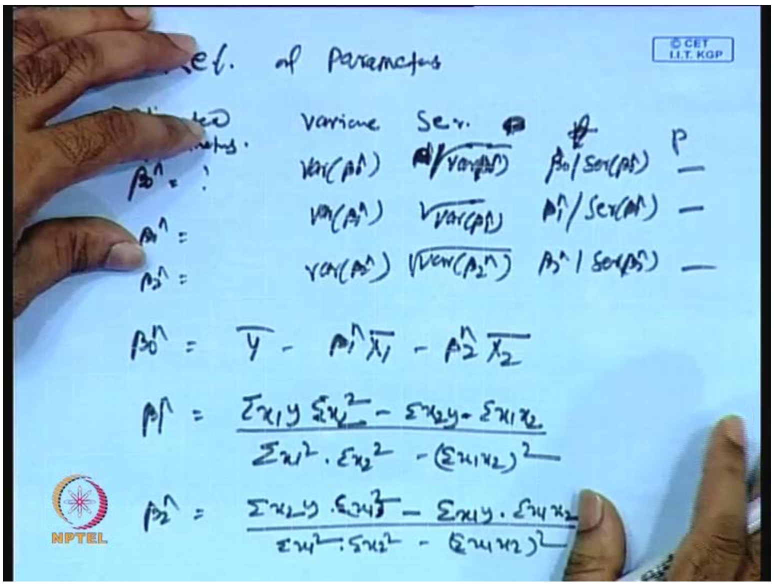


The way we have last class, we have received the estimated model just we highlight here. We start with Y equal to beta 0 plus beta 1 X 1 plus beta 2 X 2 and plus error terms, so through which we have received the estimated models Y hat equal to beta 0 hats plus beta 1 hat X 1 plus beta 2 hat X 2.

Now the moment we will get the estimated model, the next step is to go for the reliability testing, so far as reliability is concerned like bivariate econometric modelling we have to go for again two different test structures that is reliability with respect to parameters and reliability with respect to overall fitness of the models. Now the moment we will get the estimated models, so the testing or you can say reliability is all about we can say two parts, so first is to know the significance of significance of to know the significance of parameters and to know the overall fitness of the model to know the overall fitness of the model overall fitness of the model. So, this is how this is how we have we have to go for this assignment.

Now, let us we start with the estimated parameter side first then we will go for overall fitness of the models later. So, then what we have to do, so far as overall fitness of the model is concerned, so we have to so far as reliability of parameters are concerned, so we have to receive all such information, so that we can get to know whether this X 1 is significant one or X 2 is significant one, so that means we have to we have to take care of this beta 1 hat and beta 2 hat.

(Refer Slide Time: 03:19)



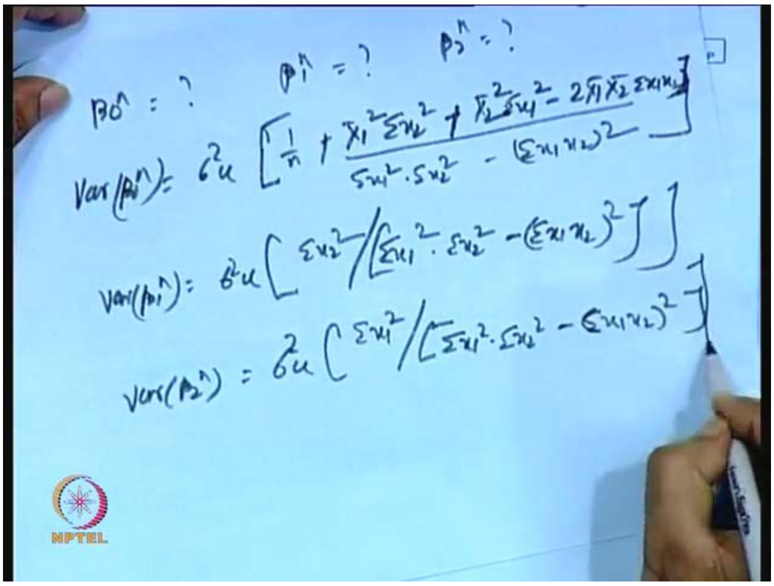
So, let me highlight how is all about this particular structure. So that means, in the first case it is reliability of parameters reliability of parameters. We we have to report the results like this, the estimated parameters estimated parameters, then estimated parameters then variance variance, then standard error then t statistics and corresponding probability, so estimated parameters we have to receive beta 0 hat, then beta 1 hat, then beta 2 hat corresponding variance of beta 0 hat, variance of beta 1 hat, variance of beta 2 hat.

This is square root of variance of beta 0 hat, then square root of variance of beta 1 hat, then variance of variance of beta 2 hats, so we will call it t here then we will call it p here. Now t is nothing but beta 0 hat by standard error of beta 0 hat, then beta 1 hat by standard error of beta 1 hat then beta 2 hat by standard error of beta 2 hat, so then corresponding probability we have to check it.

What is this actually beta 0 hat? So beta 0 hat basically beta 0 hat is equal to Y bar minus beta 1 hat beta 1 hat X 1 bar minus beta 2 hat X 2 bar; so that means, we first get to know what is beta 1 hat, then what is beta 2 hat, so once we have beta 1 hat and beta 2 hat then we can able to calculate beta 0 hat. So, the initial step is to know what is beta 1 hat and beta 2 hat then everything will be coming automatically. Now, beta 1 hat is equal beta 1 hat is equal to summation x 1 y, summation x 2 squares x 2 squares minus summation x 2 y into summation x 1 x 2 divide by summation x 1 square into summation x 2 square minus summation x 1 x 2 whole square.

Similarly, beta 2 hat is equal to summation x 2 y into summation x 1 x 1 summation x 1 square, then minus summation x 1 y into summation x 1 x 2, so divide by summation x 1 square into summation x 2 square minus summation x 1 x 2 whole square, so this is how the estimated parameters has to be obtained. Now after getting all these beta 1 and beta 2 hat, so if we will put to here beta 2 hat here and beta 1 hat here, then you get to know the beta 0 hat. Now we like to know how is the structure of variance of beta 0 hat and variance of beta 1 hat and variance of beta 2 hat.

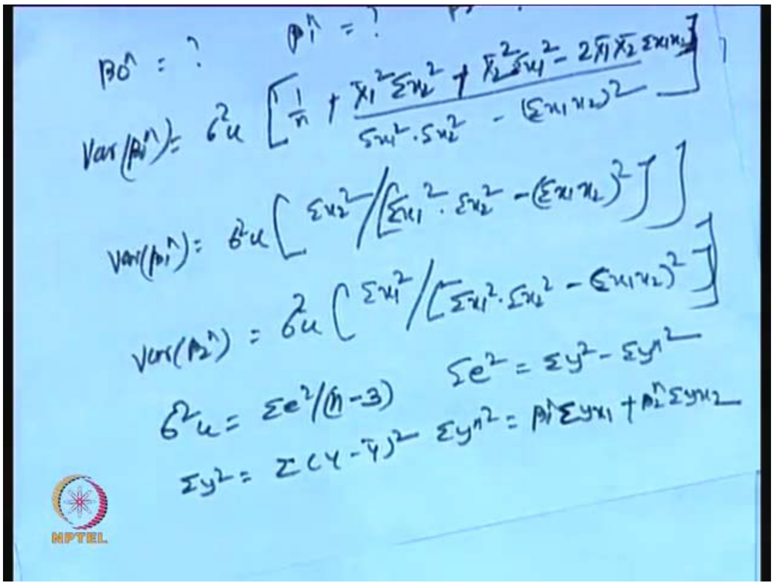
(Refer Slide Time: 06:59)



So, after getting beta 0 hat, beta 1 hat then beta 2 hat, so next step is to know what is the variance and once you get the variance then you can able to calculate the standard error and followed by t statistics then you have to go for the level of significance, so the next step is just to calculate the variance of beta 1 hat, variance of beta 2 hat and also variance of beta 0 hat. That means here, variance of beta 0 hat variance of beta 0 hat equal to sigma square u 1 by n plus X X 1 bar square summation x 2 square plus X 2 X 2 bar square into summation x 1 square minus 2 X 1 bar X 2 bar summation x 1 x 2 summation x 1 x 2 like this summation x 1 and x 2 divided by summation x 1 square into summation x 2 square minus summation x 1 x 2 whole square, this is what the beta 0 hat variance. Similarly, variance of beta 1 hat is equal to sigma square u sigma square u summation x 2 square divide by summation x 1 square into summation x 2 square minus minus summation x 1 x 2 whole squares, so this is variance of beta 1 hat.

Similarly, variance of beta 2 hat is equal to sigma square u summation x 1 square divided by divided by summation x x 1 square into summation x 2 square minus summation x 1 x 2 whole square whole square whole square. Now we have beta 0 hat, estimated beta estimated of beta 0 hat, beta 1 hat and beta 2 hat, then followed by variance of beta 0 hat, variance of beta 1 hat and variance of beta 2 hat.

(Refer Slide Time: 09:38)

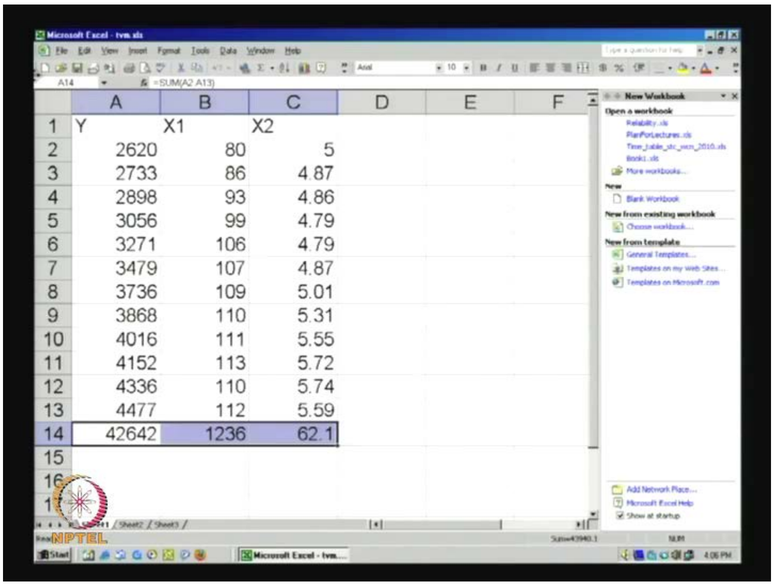


So, now to get the exact value of variance of beta 0 hat, beta 1 hat and beta 2 hat, so we like to know the value of sigma sigma square u first. So, the moment you will get the sigma square u then everything can be possible here. So, what is sigma square u? So sigma square u is equal to summation e square by n minus k; obviously, in this trivariate it will equal to n minus 3, so now sigma square equal to summation e square by n minus 3. Now, we like to know what is summation e square? Summation e square is error sum square, which is nothing but summation y square minus summation y hat square. Summation summation y square equal summation y square equal to summation y minus y bar whole squares, but summation y hat square is equal to beta 1 hat summation summation y x 1 plus beta 2 hat summation y x 2. So, this is how summation y hat square we have to receive.

Now now everything everything is with respect to this calculation, so now we have summation x 1 square, summation x 2 square, summation x 1 x 2, sigma square u, summation e square, then summation y square, summation y hat square. Now, we have to just integrate properly to get to know whether the parameters are statistically significant or not. What we have to do? So, the moment you will get this a variance of beta 0 hat, beta 1 hat and beta 2 hat, then you can just take a square root you will get standard error of beta 2 hat, standard error of beta 1 hat, standard error of beta 2 hat then followed by t equal to beta 0 hat by standard error of beta 0 hat, standard error of beta 1 hat by standard error of beta 1 hat, then beta 2 hat by standard error of standard error of beta 2 hats, so this is how you have to proceed.

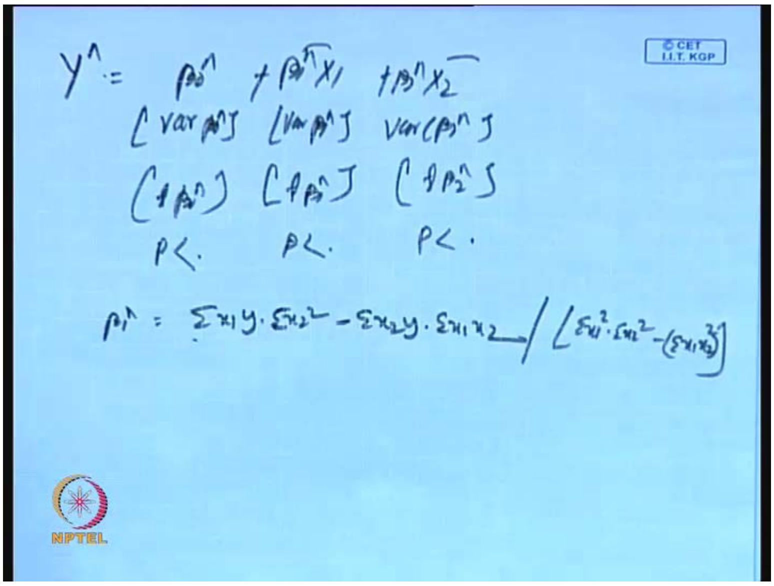
Now what is our agenda here? So, our agenda is our agenda is to get to know what this particular component and whether this particular component is statistically significant. These are the formulas or structures through which we have to examine the level of significance or test the reliability of estimated parameters for the trivariate econometric modelling. To justify the entire themes we have to we have to take numerical problems then we can able to interpret everything, because these are all theoretical way we are just bringing the structures, but actual thing will be possible when you have the data in your hand.

(Refer Slide Time: 12:09)



So, in this particular in this particular setup we have taken a problem here consists of 9 observations and there are 3 variables here Y X 1 and X 2. So, by by the time being we should not we we should not discuss anything about the name of Y name of X 1 and name of X 2. In fact, when we will go for interpretation then that times the variable naming is very much important, because by the way of variables information we can interpret the particular things. It is after reliability test and when we will go for policy implication or you can say detail discussion about these estimated models. In the time being, so keeping in mind these three variables, we like to know how is the shape and structure of reliability of estimated parameters.

(Refer Slide Time: 13:09)

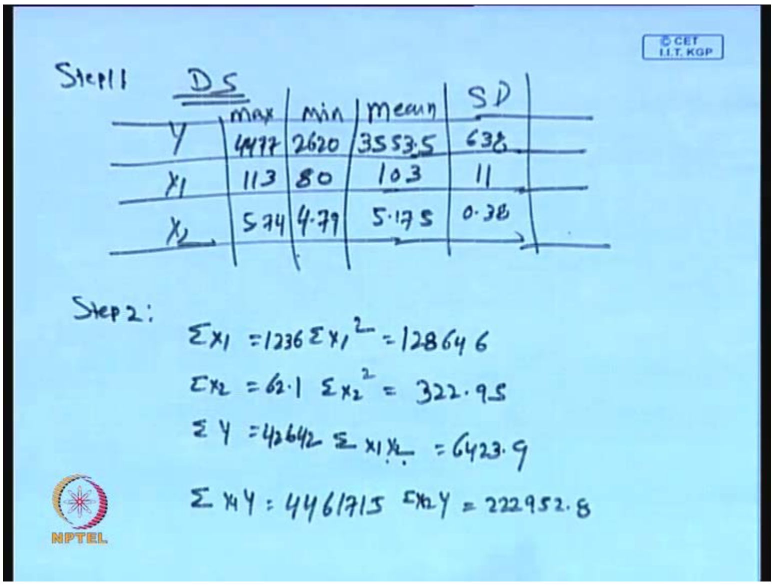


Now, we need to know first, what is estimated model? So, the moment we will get the estimated model then you have to follow the sequence. So, what is the setup here? So, all together we have to set like this Y hat equal to beta 0 hat plus beta 1 hat X 1 plus beta 2 hat X 2, then obviously this is variance of variance of beta 0 hat, then variance of beta 1 hat, then variance of beta 2 hat, so followed by t of beta 0 hat so, I am keeping the standard error component, so t beta 1 hat, then then t beta 2 hat then followed by probability level of significance probability level of significance probability level of significance.

So, now what is Y hat here? Now now we need to first calculate beta 1 hat beta 2 hat. So, beta 1 hat is equal to what is beta 1 hat here the summation x 1 y into summation x 2 square minus summation x 2 y into summation x 1 x 2 divided by divided by summation x 1 square into summation x 2 square minus summation x 1 x 2 x 2 whole square. This is how the entire structure is is all about the beta 1 hat.

Now, in this particular information we have to calculate all these things, but you remember one thing here this particular items are in deviation format. So, what we have to do for this time being, so we have to we have to go for in a particular structure. So, what is this particular structure? So here the theme is for this particular problem, let me first we will highlight all these detail steps through which we will get the estimated model.

(Refer Slide Time: 14:56)

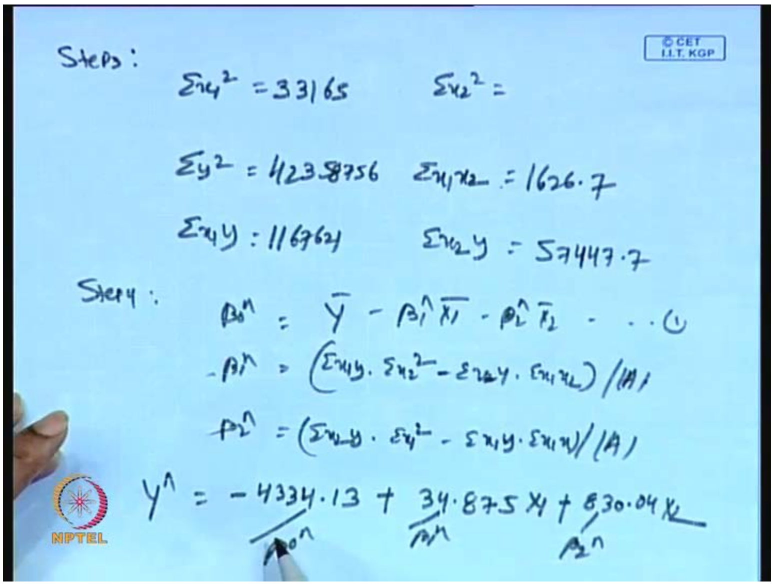


So, now this is this step 1 step 1 as usual we like to we like to present the descriptive statistics. So, this descriptive statistics will give you the briefing about these particular structures. So, this step 1 process is to know the descriptive statistics. So, descriptive statistics means the structure will be Y Y, X 1 then X 2, so we need to we need to know the maximum value, minimum value, then mean, then standard deviations, then some skewness kurtosis also requires. In the mean time, we are summarizing with respect to maximize maximum value, minimum value, mean and standard deviation, because these are all essential for this to check the reliability of the estimated model. So, now for this particular problem, so the maximum of Y is equal to 4477, then this is 113 and this is 5.74, then minimum is 2620, then this is 80, then 4.79. So, mean is coming 3553.5, then it is 103, then this is 5.175, corresponding to standard deviation is 638, then this is 11 and this is 0.38 0.38. So, this is how the entire descriptive statistic for this particular problem.

Now what we do? So next step is next step is you have to give again the summery of the entire variables, so what is the summery of entire variables? So now, for these particular problems, we need to have summation x 1, summation x 2, then summation summation y, then summation x 1 square, then summation x 2 square, then summation summation x 1 x 1 x 1 x 2, then then summation x 1 y, then summation x 2 y x 2 y. This much of information, we need to know. What is summation x 1 summation x 1 here? Summation X 1 is summation X 1 is 1236, so 1236 then summation x 2 is 62.1 62.1, then summation y is 42642. So, this is how the entire structure and in fact, I have not added other columns here. I am just giving the summery of this particular result. So, sum sum x 1 sum x 1 square is equal to 128646 then sum x 2 square is equal to 322.95, summation x 1 and x 2 is equal to 6423.9, then summation x 1 y is equal to 4461715, then summation x 2 y is equal to 222952.8.

So, this is the this is the entire summary of this particular problem that is with respect to Y, X 1 and X 2. So, what we have to do first, so you need to quickly calculate summation y, summation x 1, summation x 2, summation y square, summation x 1 square, summation x 2 square, summation x 1 x 2, summation x 1 y, summation x 2 y, then we have to move to second step. So, these are the entire information we need to we need to proceed for the significance of the parameters.

(Refer Slide Time: 19:21)



So, then we have to move to the step 3, what is this step 3 process? On the step 3 process is to transfer these all these components what we have received from the step 2 that in a deviation format. So that means, what do we need? we need summation x 1 square, then summation x 2 square, then summation y square, then summation x 1 x 2, then summation x 1 y, then summation x 2 y. I think this much this much is enough, three such deviation format is required. So, here x 1 small x 1 is deviation format that is nothing but X 1 minus X 1 bar, so that is in original figures.

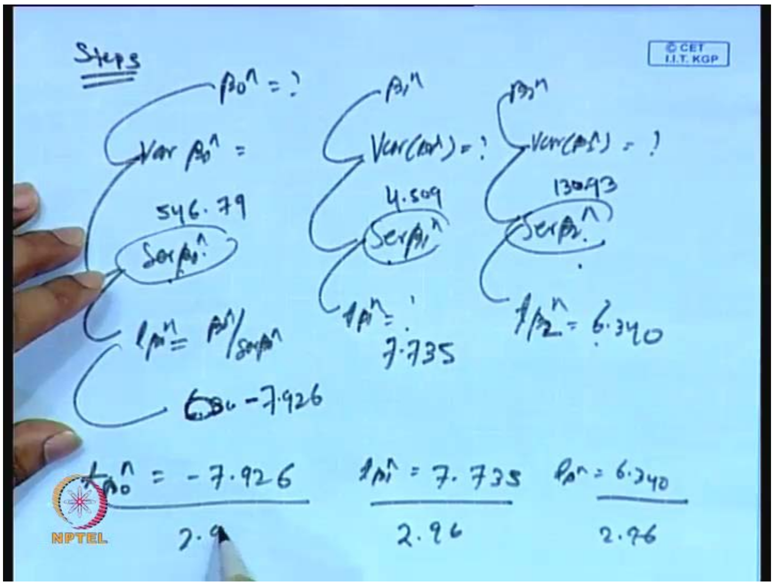
So, now similarly, we have to calculate summation x 2 square, summation y square, summation x 1 x 2, summation x 1 x 2 means sum sum of x 1 minus x 1 bar into x 2 into minus x 2 bar. Similarly, summation x 1 y means summation x 1 minus x 1 bar into y minus y bar, so this is all about in a deviation format. So, we directly transfer entire item into deviation format, so that we can quickly we can quickly have the t calculated t calculated value.

Now, we like to know what is summation x 1 square? Summation x 1 square equal to 33165, then summation y square equal to 42358756, then summation x 1 y equal to 1167621, then summation x 1 and x 2 equal to 1626.7 and summation x 2 y equal to 57447.7. So, these are the items we have to receive we have to receive in a deviation format. So, the moment we will get in a deviation format then your step is very smooth now, so we can proceed very quickly.

So, what what you do now. So, we need to know what is in the step 4 step 4 we like to know the beta 0 hat value, then beta 1 hat value, then beta 2 hat value. So, what is beta 1 hat here? So, beta 1 hat beta 1 hat is equal to here summation summation x 1 y into summation x 2 square minus summation x 2 y into summation x 1 x 2, so divided by that summation x 1 square into summation x 2 square minus summation x 1 x 2 whole square, so that is called as a determinant A. Let us assume that this is determinant A. So similarly, this is nothing but summation x 2 y into summation x 1 square minus summation x 1 y into summation x 1 x 2 divided by mode A divided by mode A.

So, now if once you get this particular value then you put in beta 0 hat. By the calculations what we have receives? So, we have receives we have receives the items let say, we can directly move to step 4, so we have to write like this Y hat equal to beta 0 hat, which is nothing but minus 4334.13, then plus 34.875, then this is obviously X 1 plus 830.04 X 2. That means, this is beta 0 hat, this is beta 1 hat and this is beta 2 hat. So, I have here omitted number of steps, because because of lack of time the way because we have already reported here the all these figures here, so you just put all these figures here then you calculate. If you apply in scientific calculator or excel sheet you can directly get this, but this is how you have to proceed and ultimately you will get this particular figure. So, we will get later, first we will calculate this one, then you have to calculate this one, then put this and this in equation 1 that is what is known as Y bar minus beta 1 hat X 1 bar minus beta 2 hat X 2 bar. If it is equation 1, then you put this one then you will get beta 0 hat. So that means; the estimated value of beta 1 beta 0 hat is 4334.13 then, beta 1 hat is 34.875 and beta 2 hat is 830.04.

(Refer Slide Time: 24:48)



So now we like to know, what is the variance of these particular structures? So what is the variance of these particular structures? So now what we have to do? So, we have to move to step 5 you have to move step 5. So, now once you have once you have beta 0 hat, beta 1 hat and beta 2 hat, so then we have to calculate variance of beta 0 hat, so variance of beta 0 hat is equal to you have to find out and variance of beta 1 hat you have to also find out and variance of beta beta 2 had you have to find out.

So, now I am not writing directly. So, ultimately the way we will move to variance of beta 0 hat or from beta 1 hat to variance of beta 1 hat and variance beta 2 hat to variance of beta 2 hat then of course we have to move again for standard error of standard error of beta 0 hat, then we have to again move for standard error of beta 1 hat, then again we have to move for standard error of beta 2 hat, so this is how you have to proceed. Ultimately, again we have to move for t of beta 0 hat, then again we have to move for t of beta 1 hat again we have to move for t beta 2 hat. Of course, we need to know this and this, this and this then t beta 0 hat equal to beta 0 hat by standard error of beta 0 hat, so similarly t beta 1 hat and similarly t beta 2 hat.

So, now we like to know, what is this value, what is this value and what is this value? So, this particular value if we will calculate by putting all these values in that particular equation, then a standard beta 0 hat we have to receive here 546.79. By the way I am omitting few steps here, because if you will go each steps one by one then it will take lots of time. So, you just remember this particular formula and I have already calculated the deviation values, so you put all these values in that particular equation then you will able to get variance of beta 0 hat, variance of beta 1 hat and variance of beta 2 hat. Now I am not directly calculating all these figures. So, here I am just reporting, so that we can quickly proceed for the significance level, because ultimately our aim is to know whether this particular item is statistically significant or not. But we need to have calculated value first and we need to have a tabulated value. So, calculate value you have to calculate by this process, so the basic idea is just to know briefly descriptive natures and actually this descriptive nature in fact in the initial it will not help it will help in the later stage when we will go for interpretation. So by the way, as far the step wise process, so we have to little bit know about the descriptive, because it it can give you the variation between all sample points.

So, then we have to go for step 1, so where you will represent the summary of all these variables, then you transpose this summery of variables into deviation format in step 3, then now in the step 4 you just quickly calculate the parameters beta 0 hat, beta 1 hat and beta 2 hat. Then in the next step step 5 we have to calculate the variance of beta 0 hat, beta 1 hat and beta 2 hat. So, the formulas I have already presented here. So, now what I have to do? I am just already I have already figure out this particular problem, so all these information, which readily available here it can be transferred into the deviation format. Just you add three columns here Y X 1, Y X 2 then X 1 and X 2 X 1 square X 2 square Y square. So, these 6 columns once you add then obviously you make it sum total then the entire summary of these particular variables in capital format you will receive immediately. So, by the way in mathematical formula we have to transfer into the deviational format and through the deviation format then everything will be in a sequence.

So, now by this process, we have received a standard error of beta 0 hat equal to 546.79 standard error of beta 1 hat equal to 4.509 and standard error of beta 2 hat is nothing but, 130.93. By the way even if you can also skip this standard error of beta 0 hat, standard error of beta 1 hat and standard error of beta 2 hat. You can directly get the t beta 0 hat, t beta 1 hat, t beta 2 hat. But before that once you have standard error then you have to then you have to integrate with the hypothesis that is null hypothesis and alternative hypothesis. Null hypothesis is that beta 0 hat equal to 0 beta 1 hat equal to 0 beta 2 hat equal to 0; so that means, we start with that the coefficients are not at all significance, they there not contributing anything. Then ultimately we have to set the alternative hypothesis corresponding to beta 0 hat, beta 1 hat and beta 2 hat. So, here in the case of alternative hypothesis, beta 0 hat not equal to 0 and beta 1 hat not equal to 0 then for beta 2 hat alternative hypothesis is beta 2 not equal to 0. So that means, in the null hypothesis we assume that beta 0 hat equal to 0 beta 1 hat equal to 0 and beta 2 hat equal to 0 then, alternative hypothesis we will set beta 1 hat not equal to 0 beta 2 hat not equal to 0 and beta 3 hat not equal to 0.

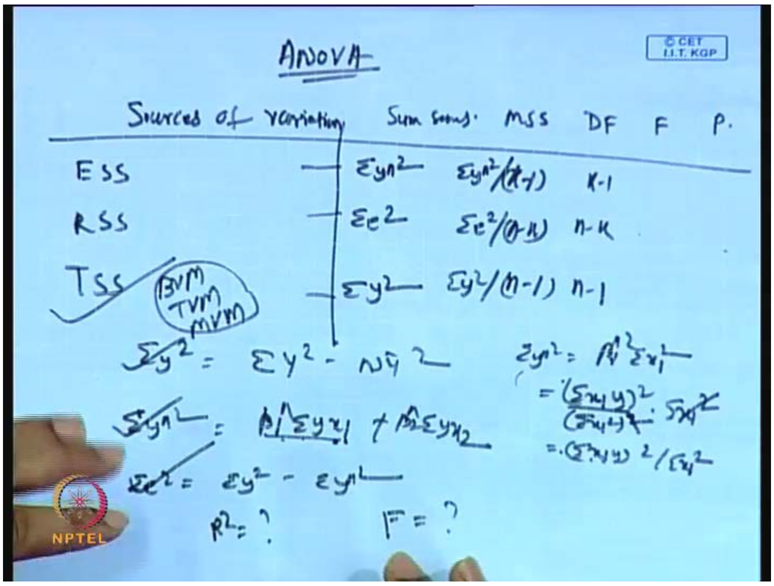
The moment we will put not equal to 0 alternative hypothesis that means it will give signal in two ways; that means, we can go for either 1 tailed test or we can go for 2 tailed test. So, that is the specialty how you are your setting the null hypothesis and alternative hypothesis. By the help of null hypothesis setting and the standard error of all these parameters then we can able to calculate the t statistics. So, now if we will apply by this particular process, then t of beta 0 hat you will get here 6.340 t of beta 0 hat is equal to it is coming negative that is that is minus 7.926, then t of beta 1 is coming here 7.735 and t of beta 2 hat is coming coming here 6.340. So that means, we will just bring it here, so 2 beta 0 t beta 0 hat equal to minus 7.9226, then t of beta 1 hat is equal to 7.735 and t of beta 2 hat equal to 6.340. So, this is also very high this is also very high and this is all very high.

So, now what you have to do. In fact, the high may be create some distort, high does not mean that the variable will be statistical significant, because this significance of this particular item depends upon the tabulated value and tabulated value also depends upon the level of sample size. So, now sample size and the involvement of variables are the two indicators through which we can get the significance level. So now, very quickly we can assess the tabulated value from the tables, so that is with respect to these degrees of freedom, so where n equal to 9 and k represents 3. So, now corresponding to this n and k and 1 tailed and 2 tailed test at 1 percent level and 5 percent level, 10 percent level, so you start with 1 percent first, if it is ok, then you proceed, if it is not then you have to go for 5 percent, if not then you have to go for 10 percent.

So, let us assume that at 1 percent level, so it is actually say 2.96 and this is called 2.96 and this is called 2.96. That means, in this particular problem X 1 is statistically significant X 2 is statistically significant. That means, the beta 1 coefficient is statistically significant, beta 2 coefficient is statistically significant; that means both X 1 and X 2 are significant factors through which y can be changed or you can say y significantly influence by X 1 impart and X 2 impart, this is how the model can be interpreted. So that means; in these particular models X 1 X 2 are highly significant factors through which Y will be changed in this particular setup.

So, now this is one part of the problem so far as a reliability of the parameters is concerned. So similarly, we quickly have the structure of the ANOVA. So, that is analysis of variance that will bring the overall fitness of this particular system. So, the way we have already proceeded the calculation of beta 0 hat to t of beta 1 beta 0 hat and beta 1 hat to t of beta 1 hat and beta 2 hat to t of beta 2 hat then we have followed the statistical procedures through which we get to know the statistical significance of the parameter. Similarly, ANOVA is also in that particular track. In fact, we have discussed details the ANOVA analysis of variance in the bivariate setup. So, now same thing it can be represented here just an addition of another variable.

(Refer Slide Time: 34:30)



Let me explain, what is this particular structure of ANOVA? So, now second part of this particular reliability is ANOVA; so ANOVA means analysis of variance, which we have already discussed. So, ANOVA structure is like this, so we like to know the sources of variations sources of variations. In fact in fact, the structure of ANOVA is, in fact structure of the estimated parameters also more or less same in comparison with bivariate setup. In the bivariate setup, there is only 2 variables Y X 1 say and in the case of trivariate it is Y X 1 and X 2 and again for multivariate it will depends upon how many independent variable you are integrating in the system. So, after this lecture we can quickly proceed to multivariate analysis and we can discuss detail about that issue.

In the meantime, the way we have finalized the bivariate setup the same thing we have here also in the trivariate setup. So, more or less in the case of ANOVA, a table is very unique. There is no such change, but so far as the value of that particular component concerned component like error error sum square, residual sum square there is change, because this formula is slightly different from this particular bivariate setup, why because if there is involvement of hard variable in this particular system.

So obviously, first of all we bring this ANOVA tables then we quickly highlight what is this setup here in the trivariate setup. So now, this is sources of variations. So, then next is sum squares, then mean sum squares, so then degrees of freedom, then F statistic and probability level of significance P. So, sources of variations ESS, then RSS, then TSS, which we have already which we have already received. So, what is this sum square here? Sum square for ESS it is summation y hat square, this is summation e square and this is summation y square. So, corresponding to this one, so your mean sum square equal to summation y hat square divided by k minus 1 then this is summation e square by n minus k and this is summation y square by n minus 1, k k cancel it is the sum of these two.

So, now this is k minus 1 degrees of freedom, this is n minus k degrees of freedom, this is n minus 1 degrees of freedom. So, we come to the F statistic subsequently. So, you see here the entire structure is more or less same like bivariate issue. So, only only difference here is that the k value only. In the case of bivariate setup k equal to 2 obviously n according to your problem setup, but. But mathematically in the case of bivariate k equal to 2, but n depending upon your sample size. In the case of trivariate, this k becomes 3 that is the only difference in this particular structure. Of course, there is there is change in between the inside picture of this particular sum squares.

So, what is this difference? Let us see here now, as usual summation y square will remain same, summation y square equal to summation y square minus N Y bar square, so, this is how we have to receive summation y minus y by r square this is in deviation format. You remember one thing, when we move from bivariate to trivariate or trivariate to multivariate so that means, when we move from k equal to 2 to k equal to 3 k equal to 4 k equal to 5 k equal to n, so in all the cases whatever size of your model, so your summation y square is always same. It will not change a terms only change is on the right hand side. So, that is how the summation y square that is in the left side, which is considered as a dependent variable is remained fixed, only change is with respect to independent variables that is what the trivariate and multivariate is all about.

So, now summation y square summation y square case is almost all same for bivariate, trivariate and multivariate, but there is a change with respect to ESS with respect to RSS only. So that means, in the case of TSS, so it is it is same for bivariate models, trivariate model and multivariate model. So, it is not at all different, s, it is more or less more or less same.

So, now we will see what is ESS here? So, ESS is nothing but summation y hat square so summation y hat square. So summation y hat square can be written like this summation y hat square is ESSENTIALLY, so which is nothing but written like beta 1 hat summation y x 1 plus beta 2 hat summation y x 2. So, in the case of bivariate models, we write summation y hat square equal to beta beta 1 hat square summation x square. So, beta 1 actually beta 1 hat square summation x square, so beta (( )) square equal to summation x 1 y, summation x square so, here also same thing.

So, what we have to do, we have put it in a different particular format. So, beta 1 hat square summation x square means summation if you put instead of x equation put x 1 then it is nothing but summation x 1 y divided by summation x 1 square. So, it is summation x 1 square so, summation x 1 square this is also square because beta 1 square here, so this is summation x 1 square, so, summation x 1 square, summation x 1 square cancels. So, it is nothing but summation x 1 y whole square by summation you can say x 1 square, this is what is called as a summation y hat square.

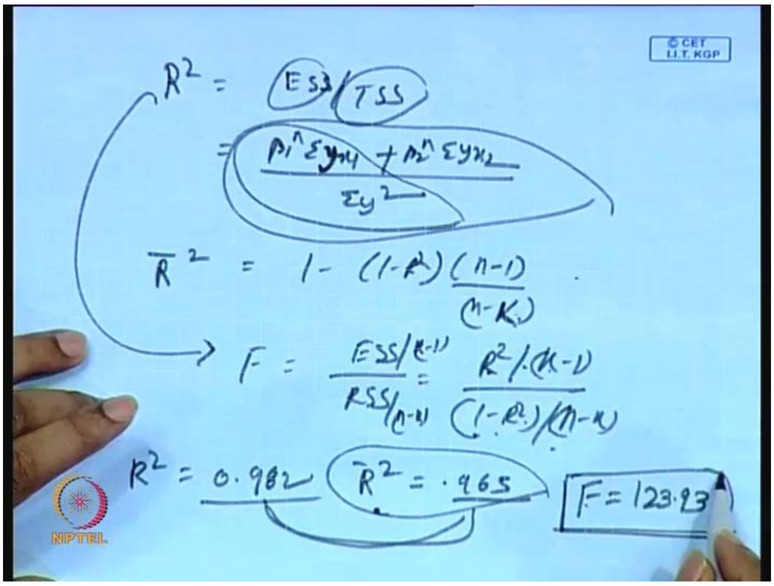
So now similarly, it can be written like beta 1 hat summation y x 1. So, the way we will redesign, so it will be only this particular format. Now, since there is another involvement variable involvement, so there is another term here involve in the numerator side. So, in the below the summation y square is there it remains constant that is total sum square. So, this is y hat square so obviously, summation e square equal to summation y y square minus summation y hat square. So what is the procedure here? The procedure is we first have summation, obviously, summation y square in the very beginning we have, so we have to calculate summation y hat square, because beta 1 hat is already with us beta 2 hat is already with us, summation y x 1 we have already calculated summation y y x 2 also we have already calculated.

So, now we have to just process it then ultimately we will get summation y hat square. So, the moment you will get summation y hat square and you have already summation y square, so you just make subtractions then you will get the summation e square. So, once you have this one this one and this one, so then you can able to get the mean sum squares with respect to error sum square, residual sum square and total sum square provided, it has to be integrated with the proper degrees of freedom.

Now ultimately in this ANOVA, we have two different objectives all together. First objective is to know what R square value is and second objective is whether this R square value is statistically significant, for that we need to have F value; so that means, ANOVA is mean to bring the overall fitness of the model, which is judge by R square that is coefficient of determination and on the other side F, which can justify the significance of this overall fitness of the model that is R square.

In fact, here the adjusted R square, which we have discussed in the bivariate has lots of meaningful interpretation in the case of trivariate, because adjusted R square means R square is adjusted with respect to degrees of freedom. In the case of bivariate, the degrees of freedom cannot change let say n minus 2 its same, but the degrees of freedom can change here, because we have introduced third variable, that means, k is 3 here. So obviously, adjusted R square will get affected, means in frankly in reality adjusted R square has a meaningful interpretation, if we will integrate both n and k, so that means, it depends it adjusted with respected to sample size as well as the involvement of variables in the systems, so n represents sample size and k represents the involvement variables in the system. So, if we will integrate together then adjusted R square has to be calculated.

(Refer Slide Time: 43:55)



Let me highlight what is all these issues. So, the issue is here, so R square we have to calculate R square that is coefficient of determination. Coefficient of determination means the physical interpretation is that the percentage influence of independent variable to dependent variable. So, percentage influence means it is the ratio between explained sum square by total sum squares. So now it is the ratio between explained sum square by total sum square, so means if we will go by mathematical formulation then it will be coming here beta 1 hat summation beta 1 hat summation y x 1 plus beta 2 hat summation y x 2, so divide by TSS that is summation y square. This is how you have to receive this R square component.

So, corresponding to R square we will get adjusted R square, adjusted R square is nothing but 1 minus 1 minus R square into n minus 1 by n minus k. So that means, this particular structure is also in the case of bivariate setup, but bivariate setup this k is equal to 2 here, we need k equal to 3, so accordingly we have to get the adjusted R square. The moment the moment the moment will add one after another variables in the system then your R square value will be increased frequently. The reason is that, because it is the ratio between explained sum square by total sum square, but total sum square in every case is remain same whether it is bivariate or trivariate or multivariate. But, in the upper side it is adding frequently, because when we will go for bivariate setup then the item is this much only.

So, when we will go for trivariate then it is this much only means all together. So now, when we will go for addition of fourth variable then it will add another item here beta 3 hat summation y x 3. So, if we will add another variable then beta 4 hat summation y x 4. So, like this you have to proceed subsequently. So that means, if we will follow that particular path every time every time the numerator numerator value will start increasing, but in the same time y square summation y square that is in the below side it remains constant whether it is bivariate or whether it is trivariate or whether it is a multivariate case. So, in every case lower portion is always constant. So, once you add one after another variable, so R square will be substantially high and high. So, that is how when we will go for trivariate modelling or you can say multivariate modelling it is better to have adjusted R square or better to use adjusted R square rather you can say R square and when you will go for purely multivariate models that time adjusted R square is the much reliable component to justify the overall fitness of the model, because it adjusts with the degrees of freedom that is with respect to number of variables involvement and number of sample size involvement.

Now now corresponding to R square, this is how we have to bring the overall percentage influence of independent variable to dependent variable, so now it has to be tested. So, how do we test for that? So, testing follows F statistics, F is the ratio between ESS by RSS ESS by RSS of course it is nothing but R square by k minus 1 divide by 1 minus R square into n minus k, this is how this is k minus 1, this is n minus k. Just we have borrowed from this particular this particular setup here. So, this is ESS and this is RSS, so the ratio between this particular two this particular two will give you the F statistic this will give you the F statistic.

So, now once you have F statistic then you have to proceed here for this significance level. So, now now as usual we have discussed in the t case, f case is also same, so with respect to you can say variables involvement and sample size we have to see the F tables and get the tabulated value. So, this is calculated value, so accordingly we have to get the tabulated value then we have to make a comparison. By the way when we will process all these information, because we have already received the ESS value we have already received the TSS value that means we get we can able to now get the R square value. So, R square by the process is equal to 0.982, in fact I am omitting at this step, so I am directly writing the results, means if you will use all these information in this particular formula then ultimately we will get R square equal to 0.982. So, similarly adjusted R square we will receive here is 0.965. So that means, there is a drastic difference between R square and adjusted R square in fact between the two this is more reliable than this one, because it is not adjust with degrees of freedom rather it is adjust with degrees of freedom that is why adjust R square is 0.965.

So, suppose you are you are very much interested for judging the significance of the overall fitness of the model, so again in this case you can also apply F statistic and if you apply adjusted R square then F statistic corresponding it is adjusted R square by k minus 1 then provided 1 minus R square by n minus k. So, this is how it can be adjusted with degrees of freedom. So, now by the process we have to receive the F value, so now once you will use this R square value then f can be calculated as 123.993, so this is how the calculated value of F, so this is calculated value F calculated value F and we have already received the tabulated value.

So, now we have to make a comparison, so accordingly we have to give a conclusion. But if you remember one thing there is always thumb rules whether it is significance of the parameter or significance of the overall fitness of the model. If the t value it is coming high then (( )) checking the tabulated value; generally those who are very experts, great statistician or econometrician; so they can guess whether this particular variable can significant or not. So, here in this particular (( )) I am saying, because F is coming very high 123.93 with respect to this sample size. So that means, you close just close your eye and you can say that it is highly significant it is highly significant at 1 percent levels, so that means that is the best so far as significance is concerned.

In fact, in the bivariate models what we have received both the both the parameters beta 0 hat and beta 1 hat significant and this case also same things we have received means both R square is significant and beta 0 hat, beta 1 hat and beta 2 hat significant. In fact, over the long run when we will go for multivariate then supporting component that is beta 0 hat has very less impact. We like to always look what is the supporting factor to all variables that is with respect to beta 1 beta 2 here.

Now here the major findings is that, so beta 1 hat is highly significant beta 2 hat is highly significant and in the same times R square is also highly significant, so that means the conclusion is that the model is reliable one, so it can be used for forecasting or policy use, but if by any chance if beta 1 hat, beta 2 hat are highly significant and other case R square is not highly significant or vice versa like beta 1 hat, beta 2 hat is not significant and R square is highly significant then the model reliability has a weakness. It is not like that you have to look only one part of the problem. So, both the part has to be taken care together. So, then it will give you a bright picture of this modelling whether it is a bivariate setup or multivariate setup every time we have to go through particular process that is to check the reliability of the parameters and to check the overall fitness of the models. So, if both are going in a right direction then your model can be best fitted and can be used for forecasting and policy use.

With this, we have we have finished this particular trivariate econometric modelling. In the next class, we will discuss the multivariate modelling, that is more interesting and more classics, we will discuss in detail in the in the next class. Thank you very much, have a nice day.