Econometric Modelling

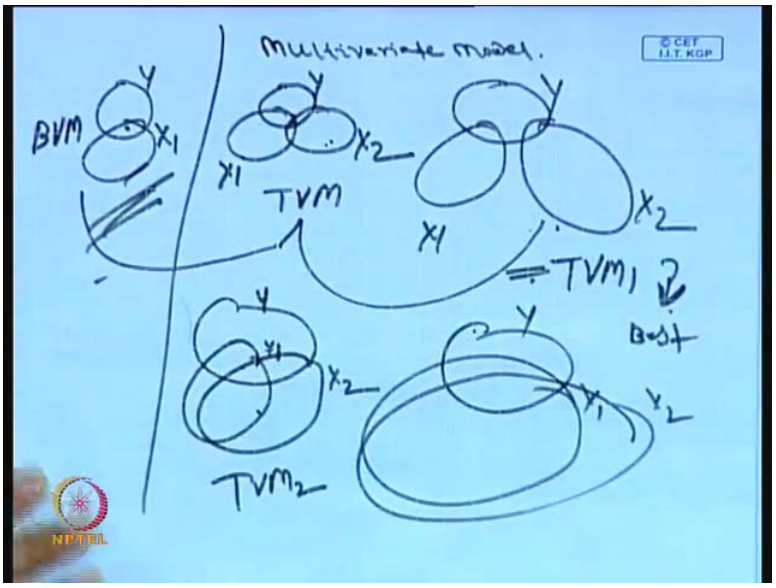
Prof. Rudra P. Pradhan
Department of Management
Indian Institute of Technology, Kharagpur

Lecture No. # 16
Multivariate Econometric Modelling

Good evening this is Doctor Pradhan here. Welcome to NPTEL project on econometric modelling. So, today we will discuss the component multivariate econometric modelling. So, in last couple of lectures, we have discussed the entire structure of econometric modelling that to bivariate analysis and trivariate analysis. In fact, trivariate itself is entry point to multivariate. But for simplicity or you can say understanding point of view, so we have classified in to three different setups, one is bivariate set-up, trivariate setup and multivariate setups. So, we are just entering one after another complex problems.

But, the starting point of econometric modelling is with respect to two variables. If, the system consists of two variables then we can proceed for analysis or you can go for fitting econometric models. So, the starting point of econometric modelling involves two variables, one is dependent and another is independent. Now, the process will be very complicated and means in fact, more complex if you proceed or if you enter one after or it will add one after another variable in this particular systems.

(Refer Slide Time: 01:44)



So, like this, so the starting point of the case is like, one is Y, another is X 1 here. So, we like to know what is the influence of Y and X 1? So, this particular structure is called as a bivariate econometric modelling. In the case of trivariate, we have one dependent variable Y and we have two independent variables X 1 and X 2. So, like Y here and we have X 1 and we have X 2. So, we like to know how X 1 and X 2 has integrated in the case of Y. So, this is case of trivariate modeling.

So, here and here there are two independent variables. And we like to know, what is the influence of 2 independent variables to dependent variables? You see here when we will move from bivariate to trivariate then, in fact the number of variable size is very high. In the same time the complexity of this modelling is also substantially increased or you can say substantially very high. For instance, here very interesting point is that we can able to track what is the exact relationship between X 1 and X 2.

But so far as econometric modelling is concerned means, best fitted model is concerned. So, particularly in the case of trivariate model and onwards like if you will proceed to multivariate then in that case the problem is that the relationship between these independent variables what we usually call as a multicollinearity problem. So, here there is no such problem in the case of bivariate model the problem. The structure itself is a very simple one where the complexity is almost all very negligible or at the low level.

But the complexity or the game is more interesting when we will move to trivariate or multivariate.

I will I will little highlight here. So, what is the basic problems or setup when we will move from bivariate to trivariate?. So, the joining starts from bivariate then ultimately you have to move to trivariate. So, then I will just before moving to the trivariate econometric means multivariate econometric modelling. So, I will little bit highlight here the basic setup of this particular problem.

So, one of the standard problem here is the multicollinearity issues. So, what is all about this multicollinearity issue? Multicollinearity issue means, the existence of linear relationship among the regressors what we call it a independent variables independent variables in this particular setup.

So, now in this particular structure this model accuracy model accuracy is not ok or means the model cannot be best fitted, it may be close to best fitted, but I will give you another example here. So, that you can get to know how is the model fitted?

So, let us say I I will have a structure like this, this is Y and this is X 1. So, I will create another another variable here X 2 this particulars this is one problem and this is another problem. But if you will compare these two this is trivariate model and this is another trivariate model, but now between these two, this is considered as the best models.

Because, there may be some other problems but by external outlook you can make a guess that this problem is probably it will be best fitted. The reason is that because, the additional problem which will receive from the trivariate model is the multicollinearity issue, which is not clearly visible here. So, the moment if you will put in a set form. So, since there is no intersection between X 1 and X 2 obviously, we are making a guess.

It is not a question of guess, the moment if the sample information say is a no close means no connection then, obviously by default itself they are totally independent. But this will not be completely independent to Y. If there is completely independent to Y then there is no point to put it any econometric models.

So, X 1 has a influence on Y, X 2 must have a influence on Y, but in the same time X 1 should not have any connections. So, if there is any connection then obviously, the

problem will be more interesting or you can say more complexity. So for instance, if by any chance if there is such relationship between X 1 and X 2 then, ultimately ultimately the complexity will start increasing.

For instance, I will I will I will take another issue here. So, this is Y this is X 1 this is Y this is X 1 and this is X 2. So, this is another trivariate model here. So, now you see here. This is one different problem, this is another problem, this is all together another problem. So, now if we compare all this forgets about this bivariate analysis now. Because, we are already moved to trivariate because, multivariate has a little bit connection to the trivariate and more or less trivariate multivariate are in the same groups. So, it is one structure.

So, now in this particular setup, if this is the case then by look you can say that it is a best fitted model and that model can be considered as the best fitted. Obviously, we have to go for estimation process, then reliability check etcetera etcetera. But in the mean times if we will find this setup is like the case 1 and case 3 then, it is little bit negative side of the reliability or negative side of the econometric models.

And again if we will compare 1 and 3 then obviously these three is more complicated, And, this this problem is more dangerous because here the association between X 1 and X 2 is a too much high. So, as a result by default means before going for estimations if you will go for cross checking of all these data and information then obviously, you have to find out solution for that. So, that before facing problem you have to clear it. Otherwise, if you go with this type of structures then I am very sure, one point of time your reliability will get affected.

So, what you have to do? So, before you going for estimations then obviously, you have to clarify all these details and process so, that it cannot be problem at all. So, sometimes altogether it is not possible. Because, if the sample observations are not straight forward and too much high then, that time it is very problematic.

So, anyway by any chance if you can make a guess then obviously, you have to find out its solution before you entering to the estimation and reliability checking. Otherwise, it is altogether step by step process and every step, for every stage you have to be very careful about all these problems or complexity. So, that when when we will go for estimation and reliability you may not face this type of problem further.

Because, the moment you entering to the estimation and reliability checking. So, you have already spent lots of time, effort etcetera. So, though it is a continuous process, still it will create unnecessary headache for you to go for this checking etcetera. So, it is better to find out its solutions.

So similarly, I will take another case here. So, this is Y this is this is X 1 and this is X 1, this is in fact this is in fact, more dangerous because, here the degree of association between these two variable is we can say substantially very high. So, that is why you have to be very careful about it. If in case you are facing problems like this means, no this is how we are making a guess with respect to information.

Because, we have not we have not entered anything about estimation, but plotting all these points with respect to X 1 and X 2. So, we are checking how they are? Whether there is an association? You see if you have Y information and X information let us say, some Y informations are there with respect to time or with respect to processional unity then you may just apply this said theory and you find out the intersections.

If Y X 1 intersection cannot be 0 then, that will be fitted. If, Y X 2 intersection is not 0 then it is fitted. But the essential requirement to fit that particular model Y with respect to X 1 and X 2 is that. So, the intersection between X 1 and X 2 should be 0, there should not be equal to any common element. If there is such common element then, there is a problem of multicollinearity and that is one of the obstacles of econometric modelling.

So, you have to bring a particular system, where Y and X 1 has a connection, Y X 2 has a connection, but in the same times X 1 and X 2 has no connection. But but in the real world problems it is very difficult to find out a situation where Y X 1 has a linkage, Y X 2 has a linkage and in the same times X 1 and X 2 has no linkage.

So, that is particularly for trivariate econometric modelling. And the complexity is more interesting when we will go for multivariate modelling. So, now in the case of trivariate it is very difficult because most of the variables are very very interdependent to each other. So, that dependency depends upon the degree. So, if the degree is very low then obviously, obviously the model accuracy will be very high. But the degree of collinearity is drastically high then the model accuracy will be very low. It is inversely related to each other. So, that means model accuracy depends upon the degree of linearity among the regressors.

So, if the degree of linearity among the regressor is very high then obviously, model accuracy will be very low. But if model accuracy is very high by default we are guessing or we can observe that there is a substantial chance that the degree of association between X 1 and X 2 will be substantially low. This is the basic starting point of this trivariate modelling.

So, my intention is here that so, I I I like to highlight when we will move from basic to complex; that means, in this particular structures if you will move from bivariate to trivariate and trivariate to multivariate what type of problems you usually face and what sort of advantage and disadvantage? Obviously, so far as overall model fitness is concerned, then the multivariate model is always very strong enough or you can say very beautiful to justify something.

But, in the in the same times there are some additional problems, some complexity are there that you have to take care very properly. So, if you could not take that type of problems then, obviously it will affect the model accuracy or you can say model reliability. So, you must be very careful about all these problems. So, with this basic in fact, there are certain other problems also which is not typically visible in the case of trivariate models, but it can be visible in the case of multivariate model.

So, we will see how this particular problem can be a more complicated, more complex once we move to the multivariate process or multivariate modeling. But you remember one thing if you are very comfortable with the bivariate structure then you also feel enjoyment in in the trivariate structure and multivariate structure, because the structure is more or less same. Only the only the additional part is just we are adding one after another variable to the system.

So, far as a calculation or formulation is concerned it is little bit complicated. And, for different problems or you can say different degree of involvement we have to use mathematics accordingly. For instance, in the multivariate models, means once you have a bivariate setup then it is very easy to handle the situation. If trivariate it is still easy to handle the particular situation. But if it is a multivariate consist of you can say 10 variables or 20 variables so, it is very difficult to handle together

And, in that particular contest by simple mathematics it cannot be possible to solve this particular problem. Particularly, if it is in the classroom class work, but still it can be able to manage provided you must have a strong mathematical knowledge.

For instance, means particularly mathematical to matrix is very useful for multivariate technique until and unless you have complete information of matrix knowledge, then it is very difficult to solve multivariate problems in particular classrooms. Yes of course, when you will go for research oriented problems or project oriented work, then that times in fact, we we have to handle in a proper statistical software.

So, that time just you have to know the themes and know the structures, know the setup and you can say interpretation means you must have complete knowledge about this modelling. Then you have to operate properly that software(s) that is all. But this theme is just you have to interpret properly with respect to the obtained result and it is a interpretation.

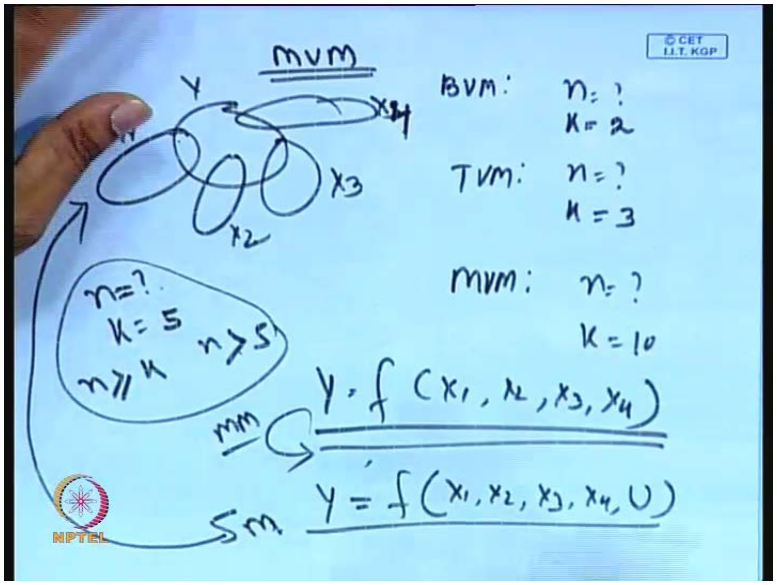
But, when you are in the class you have to be very careful and you should be little bit smart enough how you have to handle the situation with quick time without having the use of statistical software. Of course, some of our lectures we allow excel sheet to handle the situations, but excel sheet is just like a calculator it will help in middle levels only. It will not give you the final result.

The moment you will go for software use, then it will give you directly it will give you the final results. So, the intermediate process it is very difficult to have through software(s). Of course, if you use the programming then if you will go step wise then obviously, you can have this information. But this is more or less complicated ultimately we have to justify something else.

So, our point is not to make beautiness in programming or make beautiness in software(s) or make beautiness in a mathematical formulation etcetera. Our structure is to fit the estimated models and that should be considered as the best. That is how we are doing all sync. These are the helpful tools through which we can we can help the results very quickly so, that we can manage the themes in a less time.

So, however this means that these are the basic fundamental issue or basic information before you proceed to multivariate analysis. So now, first thing is what is all about multivariate analysis? Let me highlight here.

(Refer Slide Time: 16:57)



So, multivariate analysis, multivariate analysis multivariate modelling basically, here the system is like this. I am just graphically highlighting the issue. So, here Y is considered as a dependent variables and we have X 1 and we means I am now moving at a at a good structural level so, that the complexity automatically removed. Let us let us first try in a simple structured way and then we will go for complexity.

So, what is simple structured way? The simple structure way is that here game boundary k is more than 3 because we are altogether multivariate. That means, Y is 1 obviously, means in the other side(s) k must be greater than 2. So, in the case of bivariate model what we receives forget about n, but k is important here, k is equal to 2 here. So, in the case of trivariate models k has lots of integration with n because, the reason is that your model accuracy or model reliability depends upon your sample size and number of involvement.

For instance, suppose you have a model consists of 20 variables. And, you have another model with 2 variables. That means, it is a bivariate and it is a multivariate, but here k equal to 2 and here k equal to 20 sorry 19 means in addition to that dependent variables. Anyway if you add together means k is considered as 2 then obviously, k is considered

as 20 here. So, with 20 sample observation you you can go for estimation. But if with 20 samples and 20 variables you cannot go for this estimation and when will you add one after another variable then obviously, you make or you must be in a position to increase sample size accordingly.

So, the moment you will add one after another variable the same sample cannot be used accordingly. So, for instance if it is 20 samples and you are going bivariate to trivariate it is not a problem or if you will add another variable it is not a problem. But if you will add 15 more variables at a stretch then obviously, it will be it will be very complex and by default you will get model inaccuracy.

So, you will not get a best fitted model. So, that is why the n and k has lots of integration. Because, it is one of the check point through which, you get to know whether the estimated model will be a most reliable and it can be best fitted and it can be used for forecasting.

So, these are all different cross checking methods through which you can make a judgment before you go for estimation and reliability checking. If all these cross checking is something negative or you can say other way around then obviously, you have to you have to clear or you have to first erase all these problems before you go to the estimation and reliability checking.

So, what we have observed here first observation is the relationship between independent variables is more complex when you will go to multivariate model and in the same times the relationship between n and k has to be taken care properly. These 2 things are very important here in this case of multivariate setting.

So, now in the case of trivariate here k equal to 3. So, when we will go for multivariate model. In fact, multivariate models with respect to how many variable that is very important we can represent here. So, let us say k equal to say 10. So that means, this system is consists of multivariate where number of variables equal to 10 only. So, 10 means, one dependent and other 9 are independent variables. So, this particular problem is more interesting more, structured so far as multivariate is concerned.

So, let us assume that very simplicity so, what you can do? We will start with say 4 variables together. So, that 4 variables means it can be considered as also multivariate.

So, let us see if we will go for 4 variables it is very very simple structure then, you will assume that whatever independence variables in our system we have. So, these should not have any correlation. Means, they should not be related each other again. So that means, we have another variable say X 2 and we have another variable say X 3 and we have another variable say X 5.

So, right now we need not require what is the n structure of course, it should be obviously substantially greater than k but here k is how much? X 1, X 2, X 3 let us say this is X 4. So, altogether there are 5 variables here. That means, the essential condition is that n should be substantially greater than k. So that means, your n should be obviously greater than at least greater than even if it is equal to 5 the result cannot be, the model cannot be best fitted or the accuracy itself.

Sometimes, if you will put in software it may not operate properly, it will give you message that insufficient information. So, this is how you have to be very careful. So, in the mean time what you have to do? You have to take care of this particular problem this is very serious issue before you going to multivariate models. So, now, the way I have here structured the particular problem is that. So, Y is here so that means, Y is function of X 1, X 2, X 3 and X 4. So, this is the mathematical formulation of this particular problem.

Because, this econometric modelling has lots of beauty. So, one beauty is that you must have very theoretical knowledge and that theory must be very perfect and it should be very interesting or say it must be very challenging. Then, through the theory you have to build a mathematical model. So, that mathematical model has also beauty. Because, there are many ways you can formulate the mathematical models.

Then, once you will get a proper mathematical model then that can be transfer in to proper statistical model. Then, our starting point will be means that is the starting point for our econometric modeling. That means before we start any econometric process or econometric modelling. So, you have to do lots of homework with respect to theory, exact problem, objectives and of course, there is also hypothesis which you have fix before going for this model fitting.

This is altogether in resource you have to do all these assignments. If it is a small classroom problem that is different task. But the the actual issue in econometric model is

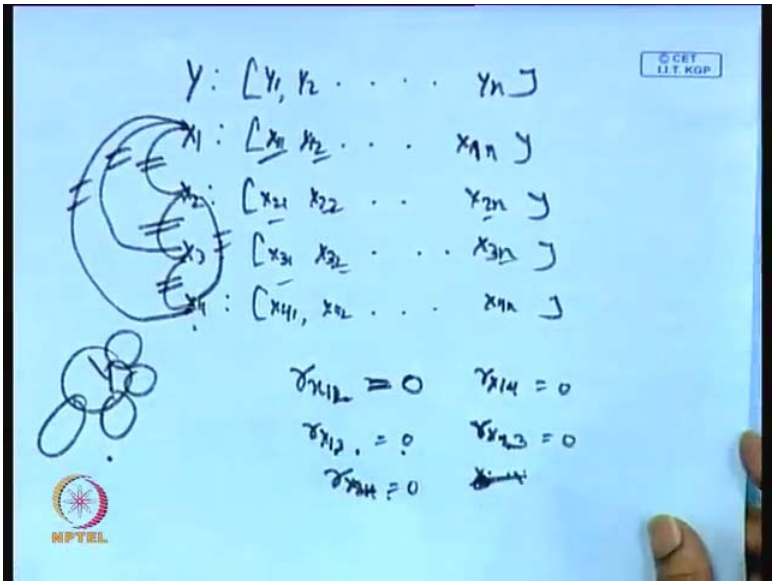
that it should be substantially attached with the particular theory and problems, objectives and hypothesis.

So, these are all your home assignments you have to do before this fitting of econometric modelling. The moment you have these, you have to do all cross checking first then, you have to proceed further. So, in this particular setup, what we have observed? So, this is the mathematical formulation of the model. But in this particular context if we transfer this mathematical form of the model to statistical form of the model then the then the formulation will be like this X 1, X 2, X 3, X 4 then U. So, this is the this is the statistical form of the model. This is statistical model and this is mathematical model.

So, of course, this mathematical model this is in implicit format can be represented in many ways, this is also implicit form can be represented in many ways. For instance, they may be linear related, there may be nonlinearly related. Again in nonlinearity there are various steps like it may be logarithmic in nature, it may be exponential in nature like this it may be some kind of powers functions. So, there are certain rules and regulation through which you have to observe all these details.

So now, once you have a statistical form of the model. Now, if you integrate with the statistical models typically these econometric issues then obviously, if plotting of information is coming like this then obviously, that means the structure is like this.

(Refer Slide Time: 26:22)



The structure is in this particular case. We have Y consists of Y 1, Y 2 then you can say Y n. So, this is how the structure is all about. So, these are all sample information(s). Similarly, we have we have 4 different independent variables. So, we will take X 1, X 1 is nothing but X 1 1, X 1 2 then it is nothing but X 1 n X 1 n.

Similarly, for X 2 the structure is X 2 1, X 2 2, X 2 n. Then, X 3 consists of X 3 1, X 3 2 then X 3 n. Then, we have to move to X 4 X 4 1, X 4 2 then X 4 n. So, for k number of variables, it will be extend to k. For this particular formulation, we have 4 independent variables and 1 dependent variable. So, you can set the sample information like this way. In fact, these are all theoretical setup these are all theoretical setup there may be some quantity figure in all these details.

So, what you have to do? When this model will be like this Y and X 1, X 2, X 3, X 4 like this. So, there is no common element in between. That means, there is no connection between this and this, there is no connection between this and this and there is no connection between this and this. Means these are all no correlation again the complexity start with X 1 and X 3, and the complexity will start with X 1 and X 4 and then, X 2 to X 4 again.

So, these are the possible ways you have to find out the linear relationship among the regressors. And, these are all essential before you are going for before you are going for estimation of this particular multivariate econometric model.

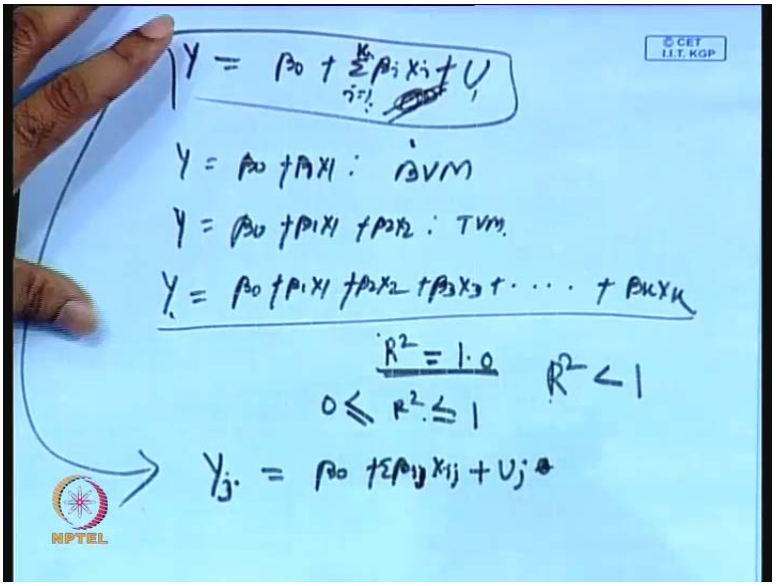
So, by graph we are getting that there is no there is no such intersections. Obviously, there is no correlation between all these items. So, we will highlight how it can be structured all together. For instance, for X X 1 X 2, what is this collinearity? we can we can get the correlation coefficient X 1 2. So obviously, if your structure is like this, if you will correlate X 1 and X 2 then, correlation X 1 and X 2 obviously it will be equal to 0 because there is no common points here.

Similarly, r X 1 3 is equal to 0 then, r X 1 4 is equal to 0 then, r X 2 3 is also 0, then r x 2 4 is also equal to 0. These are the these are the possible ways you have to justify X 1 2 3 X 1 3 is left here. X 1 3 this is X 1, X 2 this is X 1, X 4 this is X 1, X 3 then X 1, X 4 then X 1, X 3 this is X 1 2 3. So, there is X 1 3 is already here. So, X 1, X 2, X 3 then X 2 X 4. So, these are the possible case 1, 2, 3, 4, 5 then another is X 2 4 so, X 2 4 is equal to 0.

So, these are the possible ways you have to establish the relationships. So, these are all equal to 0. So now, in this particular structure if will we move for econometric modelling and its estimation and reliability checking most probably you will get the best fitted models. So, now we will proceed to discuss the details about this econometric modelling that too multivariate setup. So, in addition to these particular problems, we will face very interesting very complex issues also. Let us highlight what are these complex issues.

So, how would we start with this, particular multivariate models. And, how can apply the beautiness so, that the model can be very accurate or very perfect to go for the mathematical derivation or you can say mathematical estimations. Let us proceed with that particular setup. So, let us assume that this system consists of k variables.

(Refer Slide Time: 31:32)



So, if the system consists of k variables including the dependent variables. So that means, k minus 1 independent variables in the particular system. So, k minus independent variable and k 1 dependent variable. We started with this concept degrees of freedom n minus k. So, that k is in fact, it is unknown, but it may be k equal to 10, it may be 20, it may be 30 like this way there are many ways it can be represented.

But, let us we start with this k setup then will we summarize it. So obviously, I will write like this Y equal to beta 0 plus summation beta i X i, i equal to 1 to k. So, this is the mathematical starting point of mathematical starting point of multivariate models. So, if we will make a look at this particular model.

Then, in the case of trivariate setup, we put i equal to 1 to 2 and in the case of bivariate model, i equal to 1 only. Because, it is nothing but beta 0 plus beta 1 X 1 like this Y equal to beta 0 plus beta 1 X 1. So, that is bivariate setup, bivariate model. So, when i equal to 1 to 2 then obviously, the structure will be Y equal to beta 0 plus beta 1 X 1 plus beta 2 X 2 that is called as a trivariate model. Which we have discussed just now.

So, the complexity will start when we will go for i equal to 1 to k. So, that means the system will be like this beta 0 plus beta 1 X 1, plus beta 2 X 2, plus beta 3 X 3, plus continue beta k and X k. So, this is the mathematical generalization of multivariate models. So, this is altogether mathematical form of the model here the model model speaks that your k number of independent variables has an integration with Y. So, when we will fit like this way, there are lots of chances that there are other variables which can influence this also Y.

Because, the moment we are putting Y equal to beta 0 plus beta 1 X 1 up to beta k X k then obviously, we are assuming that only k number of variables which can influence Y. If we will not put anything extra and we are putting this statement that means, this particular left hand side, that means X 1 to X k has a 100 percent influence on Y. But still if you are saying 100 percent information statistician does not believe or economist does not believe it.

So, they assume that there is a possibility of some factor which can also influence the Y in this particular setup. So, as a result it has to be it has to be investigated or it has to be verified. So, the moment the question of investigation or verification is there. So obviously, your first target is to get the estimated parameters. So, then accordingly you will get the estimated model. So, the moment you will get the estimated model so, you have to go for reliability that too significance of parameter and significance of overall fitness of the model.

So, now the overall fitness of the model will give you signal whether the starting of this mathematical form of the model and later transfer into statistical form of the model is perfectly fitted, 100 percent fitted like this. So, the justification that means, if you will say that these are the variables which are only influencing Y. Then in that particular case, we we know the overall fitness depends upon R square.

When all these variables has 100 percent influence on Y and obviously, R square should be exactly equal to 1.0. Because, R square is a coefficient of determination which range is always 0 less than R square less than 1. So, if by any chance all these variables are not sufficient to influence or not having 100 percent influence on Y, then obviously, the R square cannot be equal to 0.

So that means, most of the cases R square will be less than 1. But you remember, R square cannot be negative R square cannot be negative because it is a square. In fact, this R square and which we have discussed long back r square hope that to this particular R square, small r is correlation coefficient which can be negative or positive, but if you will go for square of the correlation coefficient then obviously, it will be always positive. But when you will calculate simple correlation like the way we have described this r X 1 and X 4 like this. These are all simple correlation coefficient that can be negative, that can be positive, but that is not the issue.

So, in that particular case there is a possibility negative and positive, but when we will go for r square, it cannot be negative. So, it will be either 0 or 1 or in between 0 and 1. So, the model accuracy or means model reliability is very high if the R square value is close to 1. If the model accuracy and model reliability is very low if the r square value is close to 0 which we have little highlighted in the case of bivariate analysis and In fact, in the case of trivariate analysis.

So, the major problem here is that. So, when we are putting k variables in the systems then obviously, there are k number of parameters. So, when R square is close to 1 then the essential requirement is that whatever parameters involve in that particular models it should be it should be highly statistically significant. That means, all these individual parameters has to be statistically significant and at the higher level that too at least at 1 percent level.

So now, if R square 1 and the other situations are not ok then obviously, the model accuracy is something wrong or model accuracy is somewhat low. So, you have to find out the situation. Otherwise, otherwise R square may not be exactly equal to 1. So, if R square equal to 1 then obviously, there is lots of danger signals.

So, because when R square is there and all other variables are not statistically significant then, definitely it is an clue or the model itself will give you clue that there is some kind

of problem. Either in estimations or in model formulations or the inclusion of variables or something something, mainly with respect to sometimes the sample size also the problem integrations so many things will be coming in to the pictures

So, but once you have the estimated model, everything can be discussed or you can make a guess. But until and unless you go for this estimation or if you do not have a estimated model then obviously, it is very difficult to say something or you cannot make a comment.

For instance, for econometric model case the game is more interesting with respect to multivariate models only. The moment you will have a multivariate model then, there are various additional problem, various interesting problems that to you know dynastic check etcetera. In fact, multivariate is the starting point, one of the one of the stage through which we have to discuss so, many aspects so, many different problems on the econometric modelling. So, we have lots of interesting games under this econometric model.

So, like autocorrelation problem, hetero correlation problem, serial correlation problems then domain variable modelling. In fact, domain can be also applied in the bivariate structure, then structural equation modelling and panel data modeling. So, what I want to say that all these components and structures are more interesting once you handle or once we discuss this with a multivariate setup.

But, if the setup is very limited or in a bivariate then obviously, these discussions have a less accuracy and less reliable. Even in the case of bivariate structures, if you will get models where all these parameters are highly statistically significant and R square is a highly statistically significant. Even R square equal to 1 and all these statistical, all these parameters are or all the variables have 1 percent level of significance still there is a lots of confusion.

So, it is also in fact, more dangerous. If for a bivariate setup if you are getting R square 1 and all variables are highly significant by default it will give you wrong signals. And, that model may not be fitted for this may not be considered for policy use or you can say forecasting. But on the other side if you have a multivariate models where you will get a few variables statistical significant and R square is less than 1 that model can be considered as the best and can be used for forecasting and police use.

But, the reverse case where R square is 1 and all parameters are statistical significant then your game boundary is just limited to 2 variables Y and X. That means more and more you can say wrong way to go for forecasting and policy use. So, one of the interesting criteria or interesting interesting trick of forecasting and policy use is that you have to fit or you have to use this estimated models, Provided, it should be in a multivariate setup or multivariate setting.

Because, once you prepare to fit a model Y with respect to certain variables, then Y must have some name. Because, without a theory or without proper naming then, econometric itself is meaningless. Because, it is not just statistics or mathematics, it is the application of statistic and mathematics to some kind of theory and problems. So, the root is something different, the root is you have to find out the exact statistical problems. So, that through the problems we have to use mathematic, statistic to bring a particular shape. So, that is how it is very interesting.

So, you have to note that very carefully that econometric is not only for mathematician statistic. It is the application of mathematics statistics where, you must have a sufficient evidence about the theory that means, that problem which we identify or which you like to investigate it should be very accurate and very perfect, very authentic. So that, so that you will enjoy this game and it will be more practical to use econometrics. Otherwise, without sufficient evidence, sufficient theory and you just bringing information and just processing it by mathematically and statistically there is no utility at all.

Econometrics is altogether using some different problems and mathematics statistics. So, we discuss that particular problem or represent that particular problem in a more accurate way or more authentic way. So, that is how econometrics has a lot of uniqueness. So, you you must be very careful about that and you must remember all these things. It is not just like simple mathematics and simple statistic or any simple subjects. But the beautiness of this particular structure in setup is that the problem can be from various angles, various areas. But we have analyze here with the application of statistics and mathematics. That is the more interesting about this econometric modeling.

So, for instance some of the mathematical papers it is very bounded and the application is very limited. But the specialty of this particular econometric modelling is its features is very much infinite in natures, you cannot imagine you cannot say that where it cannot be

applied. It will be applied in each and every field wherever you like when there is a question of some argument or verification or justification that times you close your eye and just apply the econometrics. Provided, you must have sound logic, sound theory means proper theory and proper problem setup. Then, you can apply accordingly you will get this process very carefully.

So, this is how this is how you have to get the multivariate econometric modelling. So now, the starting point means the briefing of this multivariate econometric modelling is like this. Where i equal to 1 to k and the problem setup is like this and if you will represent in econometric form econometric models then there should be error terms. Otherwise, if we will just simply write beta 0 plus summation beta X I then it is just mathematical equation or that means, it is a mathematical model.

So, but it will add u then obviously, it is represented as a econometric model. But in the bivariate setup and trivariate setups we mention very clearly. So, this particular model can be represented in 3 different forms with respect to your data setup like a cross sectional samples and time series samples and panel data samples. So obviously, I have not mentioned here the sampling structure. So, there it is a cross samples or time samples or panel data samples.

Accordingly we have to use here. So, sampling information for instance instead of Y have to put i, but here I have already used. So that means, you can use j also. I can write like this for cross section modelling say Y j for this particular I am just transferring it to a cross sectional formulation.

Here in this if I will write like this way then there is no clear cut identification. So, whether it is a cross sectional modelling or whether it is a time series modelling or whether it is a panel data modelling. So, this in fact, represents the implicitly format of the model. This itself has a lot of beautiness because something you are keeping in a hidden. So, people has to explore and the moment you will put like this way then, by look you cannot make a guess what type of problem is this? Or what is the estimation process?

So, to know this one they will read lot or they will go through your theory etcetera. So, that is why in research problems they start with like this. But instead they will mention all details. So, that anybody can go through it properly. But for the class point of view or

for beginners you must be very careful about this particular formulation. And, now for cross sectional modelling the structure will be like this Y j equal to beta 0 plus summation beta i j and X i j. No, in fact, beta i only this is X i j because the sample is with respect to variable then u j. So, this is how the formulation all about.

So similarly, when we will we go for time series modelling then j will be transferred to t. So, when you will go for panel data setting then j will transfer to j t similarly, x i t. So, this is how you have to you have to move from different sample problem that means, cross sectional to time series and time series to panel data.

In fact, the panel data itself has a lot of beautiness. For instance it will increase the sample size, but we have not discuss anything the moment you will go for panel data modelling then, it is the integration of cross sectional observation and time series observation.

So that means, i is not constant here. So, i itself will vary and t is not a constant, t will vary. For instance, there are 5 different form analysis and 5 different time periods. So, 5 into 5 different structures. So that means, it is all altogether keeping t constant where t equal to 5, i can change, j can change with respect to 5 different industry setup. Then accordingly, if you keep this constant then t can change with respect to different time frame.

So, that type of problem is a really panel data analysis. So, if you go to set this type of structure then altogether it is a different ball game. So, means you will find there is a random effect or you can say 3 d effect etcetera. These are all the details means, we will discuss when we will go for this panel data modelling we have a separate lecture for that.

But, in the mean time you just make a look or you must understand the concept multivariate modelling how you have to deal the issues? How you have to deal the structures? Then you have to go for this estimation process then, how you have to go for the reliability testing etcetera.

There is a lot of interesting games you have to find out lots of enjoyment you will feel when we will handle the multivariate problems. You will find there are lots of additional techniques, additional tools we have to apply when we are in the multivariate setup. But which is not practically feasible in the case of bivariate setup and trivariate setup. With

this basic framework so, we will stop here, because time will not permit to discuss the estimation process of this multivariate modelling. So, in the next class we will discuss the detailed setup of multivariate modelling. So, with this we will conclude this particular class thank you very much, have a good day.