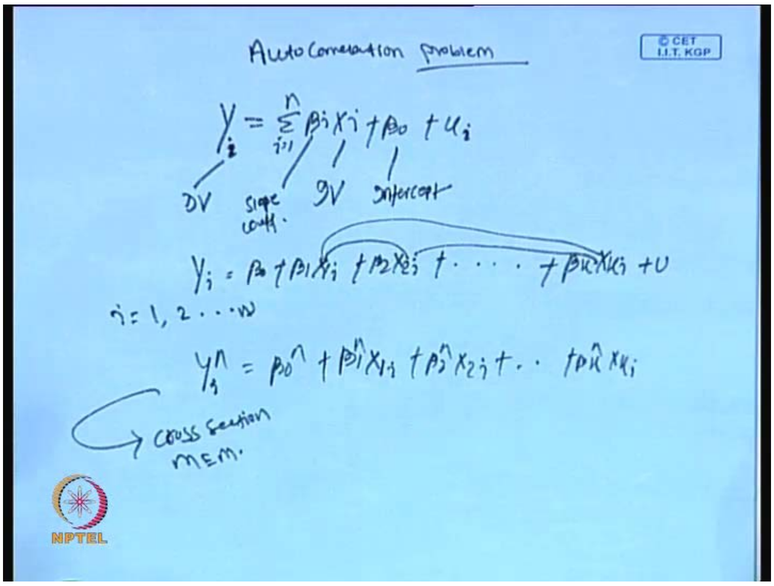
Econometric Modelling

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

Lecture No. # 22
Autocorrelation Problem

Good evening, this is Doctor Pradhan here, welcome to NPTEL project on econometric modelling. Today, we will discuss the problem called as a autocorrelation. First thing is what is autocorrelation? In the last lecture, we have discussed the concept called as a multicollinearity issue. So multicollinearity is a it is a problem where there is existence of linear relationship among the regressors. But here we have to discuss in same way but the structure is little bit different.

(Refer Slide Time: 01:05)



So, let me highlight, what is exactly the issue. So, for a particular model, say you know regression model with a multivariate framework. So, this is how the structure is all about, then Y is dependent variables then X is independent variables intercept concept, intercept component, then this is slope coefficient slope coefficient. So, now put it in another particular format, the structure can be written like this Y i equal to beta 0 plus beta 1 X 1 i plus beta 2 X 2 i plus continue plus beta k X k i. So beta 0 plus beta 1 X 1 i beta 2 X 2 i plus beta k X i plus U. So now you see here; here the model structure starts with you know i subscript here, so i represents here sample size; i equal to 1 2 up to N alright. So, N number of observations are there. So, we have k number of independent variables. Now when we discuss the concept of multicollinearity then we have to see, what is the relationship between, Y X 1 X 2 or Y X 1 X k or X 2 X 1 X k this is how we have to establish the relationship. In the similar, we have to discuss a problem called as a multicollinearity issue. So, what is all about this multicollinearity? So, multicollinearity is the means you know we have to discuss the autocorrelation issue means it is slightly different from the multicollinearity issue. So, now the multicollinearity problem is the degree of association between various associations between various independent variables. But you know in the case of autocorrelation it is the degree of association between something called as a error terms, because you see when we have a regression model so, there are three components all together: dependent variable component, independent variable component and error component.

So, some of the interesting techniques are with respect to Y and X some of the interesting techniques are with respect to X and Y and some of the interesting components are with u itself. So, now in this particular autocorrelation problem we have to discuss the issue of the u component only. So, let me first highlight the starting point of this econometric modelling and how we have to integrate with this autocorrelation problem.

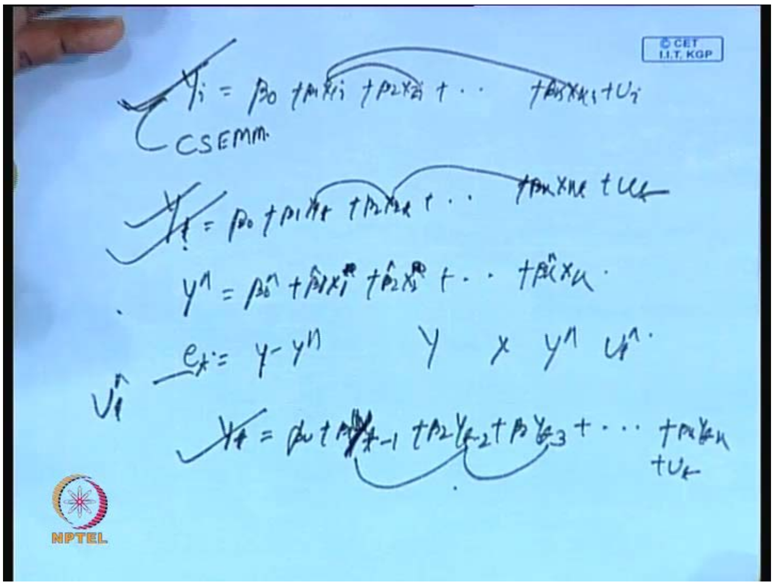
So, now the moment you will get the regression models, multivariate regression models then obviously the next step you have to go for you can say its estimations. Now, we have to see, what is the estimation process here? So, estimation process is equal to Y hat equal to beta 0 hat plus beta 1 hat plus beta 1 hat beta 1 hat X 1 X 1 i plus beta 2 hat X 2 i plus continue beta k hat X k i.

So, obviously u will be automatically removed in this structure. Now the moment you will get this, then obviously this particular structure again with respect to Y i equal to 1 to n. This is otherwise called as a cross sectional modeling. Cross sectional cross sectional multivariate econometric modeling. So, in the last class we have discussed these particular models and after that we have to see whether there is multicollinearity problem, because it is standard problem associated with the regression. Then ultimately

ultimately we have to see, whether there is any association between the independence variables.

So, now in the same lines we have to discuss this problem called as autocorrelation. You remember one thing here. Autocorrelation is a time series problem. Multicollinearity it can be you can say cross sectional problem or it can be you can say time series problem. But autocorrelation most of the cases or most of the situations it is time series issues only. So, we have to set here the time series component first then we have to integrate with the autocorrelation problem.

(Refer Slide Time: 06:17)



Let see here, so this is the cross sectional modeling. So, we have to we have to just integrate with the time series modelling. So, this is how we have to write. So Y i equal to beta 0 plus beta 1 X 1 i plus beta 2 X 2 i plus continue beta k X k i so plus U i. So this is called as a cross sectional econometric modelling econometric modeling.

So, in the line same lines we have to write here Y t equal to beta 0 plus beta 1 X 1 t plus beta 2 X 2 t plus beta k X k t plus U t. So, this is how when I will write the subscript t, then it is one way of putting in the time series issue. So, that means it is an indication that we are handling the time series problem.

So, now the major issue is here; whether the subscript is i and whether the subscript is t. If it is t then it is called as a time series issue and if it is i then it is a cross sectional issue.

So, now you know both the models are looking like same but the here there is the moment you will introduce the time series issue, then one of the major problem on the time series issue is autocorrelation problem.

So, now we have to see because in fact it is not only way to put in only t, so there is other way you can put like you know it may be t minus 1 t minus 2 t minus 3 like this. So that means, there is question of log introduction in the system. So, this is the simplest way of representing the time series model but there is lots of complexity in the time series time series itself. So, we have to discuss little bit later but in the mean time you see here. The term multicollinearity and term autocorrelations are more or less same dimension. Here, in the multicollinearity we are observing degree of association between you can say regressors. In the case of autocorrelation we are discussing the degree of association among the error terms. So that means, what is all this particular issue?

But you remember whether you have this particular format or this particular format then ultimately your target is Y hat you have to get the Y hat. Y hat equal to beta 0 hat plus beta 1 beta 1 X 1 hat plus beta 2 X 2 hat plus continue beta k beta k sorry beta 1 hat beta this is X 1 this is X 2 then beta k hat X k then obviously there is no error terms.

So, now you see here the moment you will get the estimated model, then error term is automatically removed. But you know you can get the error terms so error term is nothing but let say even if U t or you can call it e t no problem at all. So, e t equal to Y minus Y hat. So, ultimately we can get the you know, error component. So that means initially we start with Y and X then ultimately you will get Y hat and you know U t hat.

Let us say U t hat; this is error terms instead of e t we can put it U t hat error term, so this is error components. So, there are all together four variables. But our structure with respect to autocorrelation is that so we have to handle with the U t component how U t is very active in this econometric model that is the agenda we have to discuss in the autocorrelation issue.

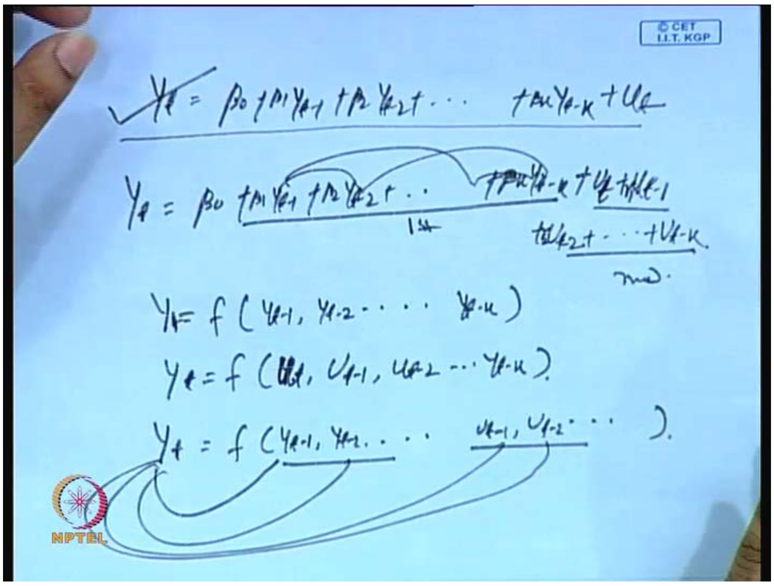
Let me briefly highlight here, so now we just make a issue here. So, I can write a model like this instead of writing like this. I can write Y t equal to beta 0 plus beta 1 X beta 1 same I can Y t I can put Y t also; beta 1 Y t minus 1, plus beta 2 Y t minus 2, plus beta 3 Y t minus 3, plus something something, plus beta k Y t minus k beta k t minus k plus U t.

So, this is how I can write a model so, this is multivariate model and this is also multivariate models, this is also multivariate.

This is purely cross sectional modelling, this is time series modelling but it is not pure time series modelling this is again time series modelling and it is in fact called as pure time series modelling. Because, time series modeling one of the interesting features is the log introductions, the moment you will introduce the log then obviously problem itself will be very interesting.

So, now you see here is the way we have discussed here in the case of cross sectional modelling or you can say simple time series modelling; we are much concerned about this association. These are called as a multicollinearity issue. Even if here also, the series is Y t minus 1, Y t minus 2, Y t minus 3 is also multicollinearity issue. But you know within the particular setups I can write here you see here I will put it in different dimension.

(Refer Slide Time: 11:13)



So, now I will write here let say Y t equal to beta 0 plus beta 1 beta 1 Y t minus 1 plus beta 2 Y t minus 2 plus beta k Y t minus k plus U t. So, this is one type of time series model with log involvement. So, now the way you are involving a log with a particular variable, then obviously the log can be applied also in the error terms.

So, this is the log model where error term is simply represented as U t. But it cannot be simply represented as U t, once you introduce log with respect to Y t. So, obviously I can write the model like this Y t equal to beta 0 plus beta 1 Y t minus 1 plus beta 2 Y t minus 2 plus continue beta k Y t minus k plus U 1 we can call it U t. Then U t minus 1, plus U t minus 2 plus it will continue plus U t minus say k.

So, now this is one division and this is another division. So this is first division and this is called as a second division. Then first division is means Y t is a function of Y t minus 1 Y t minus 2 so, Y t minus k this is 1 division. And another division is Y t equal to function of U t U t U t minus 1 U t minus 2 continue like U t minus k. So, now if you will integrate then we will get Y t equal to function of Y t minus 1, Y t minus 2 so continue like this U t minus 1 t minus 2 continue so that means this is 1 series and this is another series.

So, now the game is very interesting here. Now, our target is to regress these with this, this with these, this with these this with again these. So this is how we have to we have to fit the models. So, that means the moment if you will put it in explicit form, then obviously there should be some supporting component here. So, that means let say it is a gamma 1 Y t minus 1 gamma 2 Y t minus 2 gamma k Y t minus k.

So, that means in this particular size beta 1, beta 2, beta 3 up to beta k are the supporting component to each and each and every variable. And in the case of error terms alpha 1, alpha 2 up to alpha t alpha t minus 2 alpha 1, alpha 2 up to alpha k is also another series of series of variables which can you know interact with error terms.

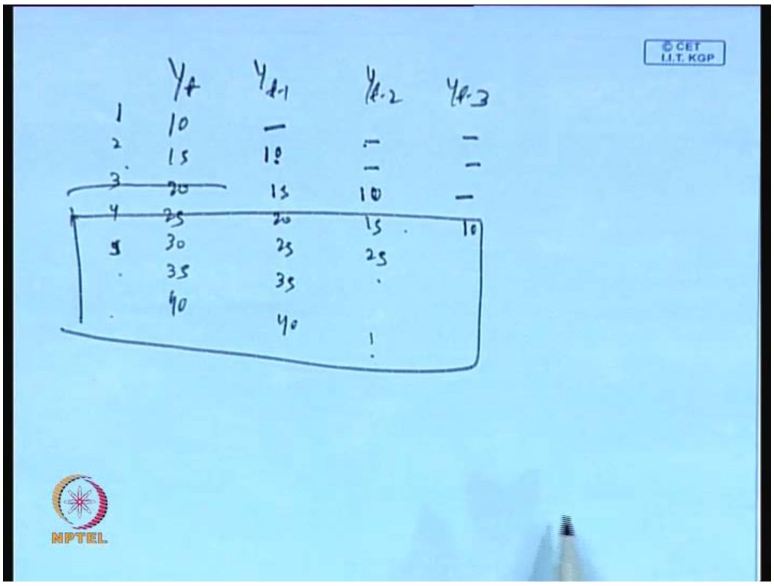
So, that means here there are two series, one with respect to direct variables and another with respect to error terms. Now, we like to know what is the association between these variables and what is this association between these variables this particular core structure is called as a multicollinearity problem and this particular structure is called as a autocorrelation problems.

So, that means if you will closely make a look in this particular models then you will find the structure of multicollinearity and the structure of autocorrelation is in a similar fashion. The only difference is it is with respect to direct variables and this is with respect to in autocorrelation, it is with respect to error components, because but you remember error is not initially with you. So, initially you start with a particular variables say Y or X.

So, means you can start with Y X both together that is you can say multivariate time series modeling, but when we will go for univariate time series modeling. So, within the univariate, within the particular variables, we can create also multi multivariate model.

So, that means we can we can integrate Y U t with U t minus 1 or Y t with Y t minus 1 like this so, this is how the complexity will start. So, that means you add one after another log then obviously you will get a complicated model and that is called as a multivariate model. But you remember in the time series that is very interesting feature but in the same times the moment you will introduce one after another log, then you know model will tends to multivariate. But in the same time there is lots of problems will be in front of you. First thing is if we will introduce log one after another then obviously you are going to lose the sample size, because in the very beginning I have mentioned when we will fit a model let say the model is Y t Y t minus 1 Y t minus 2 then obviously you need a consistent sample size. So, that means for Y t so there should be ten numbers Y t minus 1 there should be ten and Y t minus 2 it should be you can say ten so like this.

(Refer Slide Time: 16:28)



I will give you little bit indication, how is all exactly structure. So, let say Y t here so Y t Y t here the sample is 1, 2, 3, 4, 5 like this. So, now this is Y t, now this Y t minus 1 so for each sample observation then obviously I will say number 10, 15, 20, 25, 30, 35, 40 like this so, this is Y t. So, now i will create Y t minus 1, Y t minus 2, Y t minus 3. Let us say 3 variables extra variables I have to create within the system.

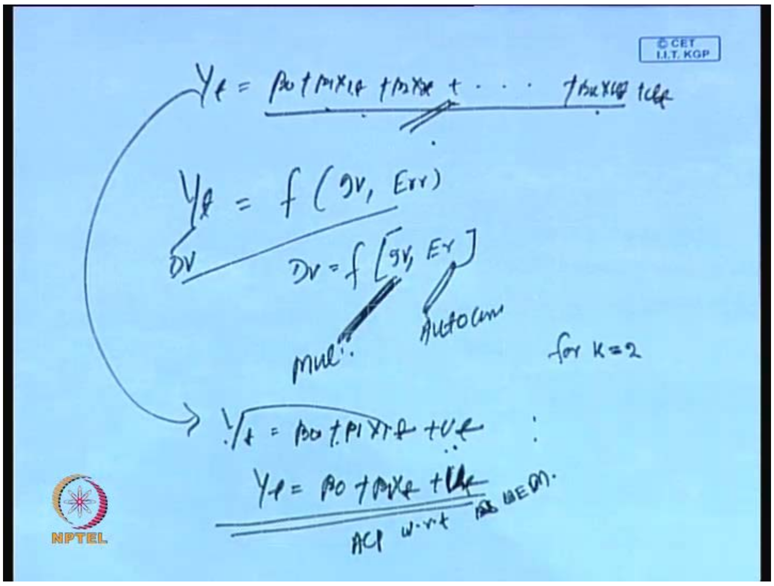
So, Y t minus 1 so for this there is no such Y t minus 1 so that means, for second then the series will start from here. This is for this one this will be Y t minus 1. So, this will be 10, then this will be 15, this will be 20, this will be 25, this will be 35, then this will be 40 so this is how it will look.

For Y t minus 1 t minus 2 so 10 will come here only. So, this is will be come here 10, 15, 25 so it will be coming here; so Y t minus 3 so it will come here 10. So that means here we are getting means losing one sample point here we are losing two sample point here we are losing three sample points. So, now ultimately when you will fix up a model then you have to take uniform sample so it is not uniform sample rather this particular structure has a uniform sample. So, that means so once you had one after another so you are going to lose one sample size.

So, obviously you must be very careful about that in fact there is standard tricks or techniques to decide what should be your log length. But one of the conditions of time series modelling is that your sample size should be exclusively very high. In the cross sectional modeling, with little sample you can do some work but you know in the time series modeling.

So, if you will start with creating multivariate models within a particular variable, then obviously your sample size is that means the sample size should be exclusively very high. If it is very low level, then obviously you cannot fit time series modelling even you fit may be it may not be the consistence result.

(Refer Slide Time: 18:58)



So, this is how the entire structure means beginning of this particular term autocorrelations. Now, I will give you little bit hint what is all about this autocorrelation problem. Now, we have written here Y t equal to beta 0 plus beta 1 X 1 t plus beta 2 X 2 t plus continue beta k X k t plus U t. In the mean time, we are not introducing the log so log model is usually called as a volatility modeling. So, we will discuss some time later, because we have a specific lecture for that. So, in the mean time I will take a simple model and you know use the subscript t subscript t means time series so the data will be available with respect to time whether means instead of any cross sectional units.

So, now once you have this type of model then, our agenda is here to discuss the multicollinearity issue. Sorry, autocorrelation issues. Now, when there is autocorrelation issue, then obviously we will not bother about this particular independent setup. So that means, this particular setup will remain handicapped. Of course, in the in between we can use this one for this autocorrelation checking or something else something else. But if you directly involve with this particular structure then it is called as a multicollinearity.

So, that means if you if you are playing game with only X regressors X 1 to X 1 to 2 X 2 to t and X k t then obviously this particular structure is multicollinearity issue. But our game will be very interesting if we will integrate Y 2 with the error terms. So, because the error term in fact is the much means much influential component with respect to this particular model.

So, now you see here. So we have two different setups here, so first setup is that we have discussed this particular multicollinearity starting point is that you should have a multivariate models, where number of regressors must be substantially or somewhat very high. So, that means number of regression at least you should have 2 if it is 2 regressor, then obviously there is question of multicollinearity detection or you can say multicollinearity means you have to play with lots of multicollinearity problems.

But if you increase one after another X then obviously the game will be more interesting. But you know if this is the way we are discussing with multicollinearity the by putting 2 at a time 3 at a time 4 at a time then obviously the complexity of multicollinearity will start increasing. So similarly, in the case of U t so you start with U t then you create U t minus 1 you create U t minus 2 and again U t up to U t minus k. Then within that particular system you are creating additional systems, so that additional systems will be perfectly okay or perfectly consistent, if not then there will be autocorrelation problem.

Now, this is the general framework of time series modeling. But before you means before you proceed to this particular autocorrelation problem. So, I like to highlight two things here; first thing is we have discussed multicollinearity in the similar line of autocorrelation because technically means inside store is more or less same more or less same means.

Why it is one hand and is a function of you know independent variable and error term and this side is dependent variable this is dependent variable so that means if you will integrate so it is a dependent variable as a function of independent variable and error terms. So, now if within the independent variable if there is any such modeling, then it is called as a multicollinearity issue and within the error term if such modelling then it is called as a autocorrelation problem. This is called as a autocorrelation problem and this is called as a multicollinearity problem.

But, multicollinearity issue we have already discussed. Now, we will see what is this autocorrelation issue; from this particular structure, we must be very careful second thing second thing is that autocorrelation can be bivariate problem can be multivariate problem. Means, if you have a model say Y t is simply function of X t only then there may be chances of multi autocorrelation but there is no chance of multicollinearity.

So, that means what we can conclude? Multicollinearity is always you can assume that it is a multivariate problem but in the case of there is no question of in fact assumptions by default multicollinearity is always multivariate problem.

But, in the case of autocorrelation, it can be bivariate, it can be multivariate. Now we have to see in fact if it is bivariate the system is very simple one. But if it is multivariate model then system will be little bit complex one. In fact it is not too much complex because the entire setup will be same way. Ultimately, you have Y t has a function of several variables and U t then, you have to integrate properly to get the U t component only. Once U t component you will receive then the game plan will be completely different. But we have to do lots of interesting games with respect to only error term keeping you know Y t X t remain constant.

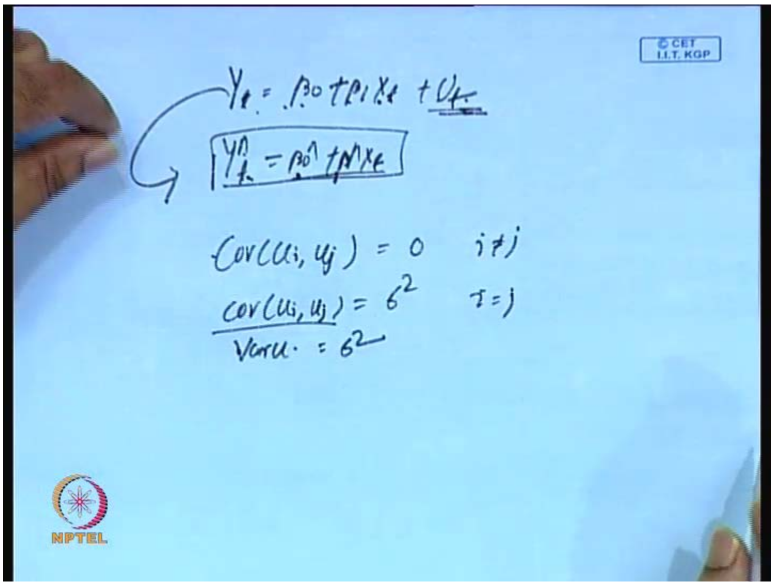
But, ultimately you have to first use Y t and X t to get the U t. So, once you will have the U t then the game will be again more interesting. So, that means here the main agenda is to find out the error term. So, the moment you will get the error term then obviously there will be issue of autocorrelation. So, now since we have already mentioned autocorrelation can be bivariate problem can be you can say multivariate problem. So, it is better it is better we will discuss this particular problem with respect to bivariate setup, because it will give you little of little bit simplicity.

So, let say we start with so we will reduce our model to Y t equal to beat 0 plus beta 1 X 1 t. So, let us assume this is plus U t so for you can say for k equal to 2 k is equal to number of variables in the system, k equal to number of variables in the system t represents number of sample point certain time periods.

So, k equal to 2 means here Y t and X t these are the 2 variables in the system. So, we are introducing error term then obviously all together the system has 3 components. Now, if we if we are reducing this model to this model, then we can simply write Y t equal to beta 0 beta 0 plus beta 1 X t plus U t because since it is only 1 X then there is no point of introducing X 1 t. If there is multiple X then of course, you have to write X 1 t, X 2 t, X 3 t because there is difference among them.

But, once you have 1 variable then its better you have to put Y t equal to beta 0 plus beta in fact this is beta 1 so beta 1 X t plus U t. Now, the autocorrelation starting point is this one. So, we will discuss this autocorrelation with respect to autocorrelation problem with respect to with respect to bivariate setup bivariate models bivariate econometric modeling.

(Refer Slide Time: 26:42)



So, then it can be generalized in the case of you can say multivariate modeling. So in the bivariate setup how is this step of autocorrelation? Let us start with that simple equation. So, Y t equal to beta 0 plus beta 1 X t plus U t so that no point to explain it once again. Because Y t is as usual dependent variable, X t is independent variable and U t is the error terms with respect to time.

So, now what is our main agenda? Main agenda is to have the estimators. So, that means we like to have Y t hat Y t hat is equal to beta 0 hat plus beta 1 hat X t beta 1 hat X t so error term will be removed automatically. Now, before going to this particular estimation models; so I like to highlight something about the U t. But you remember one thing so when there is multivariate models like this, multicollinearity models so this this will be it is better to put this one.

So, if you put like this model then obviously, you have Y hat equal to beta 0 hat beta 1 hat X 1 plus beta 2 hat X 2 plus beta k hat X k. Now once you have these type of models so what you have to do? So, we immediately have the U t components so once you have U t component then obviously you have to proceed for its modelling. So, once you have U t, so you will expand also with respect to with means the way we expand with respect to Y t and X t so, let us see here. Now the basic starting point of this you know model is

Y t equal to beta 0 hat beta 1 beta 1 X t plus u t so obviously the estimated model will be Y hat t beta 0 t beta 0 hat plus beta 1 hat X t.

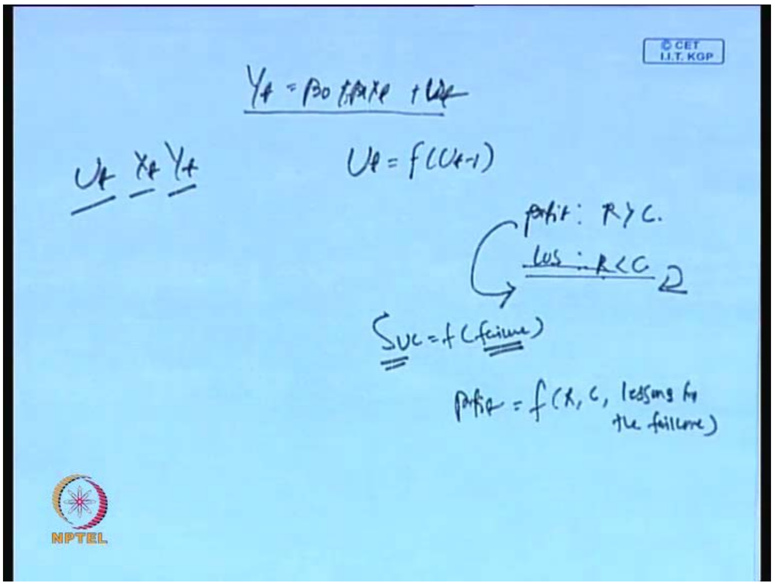
But you remember to get equation 2 from equation 1 we usually apply o 1 s technique but remember o l s technique is biased on certain assumptions. And one of such assumption is a covariance of covariance of U i U j is equal to 0 for i not equal to j and covariance of U i U j equal to sigma square provided i equal to j. This is covariance and obviously by default this if i equal to j then it is called as a variance of U variance of U equal to sigma square.

But when we will go for time series modelling pure time series modelling then it may be covariance between U t, U t minus 1 or U t, U t minus 2 like this. So, this is how the problem of autocorrelation can be analyzed. So, that means the standard assumption of o l s technique is that, there should not be any linear relationship among the error terms. If there is linear relationship among the error terms then obviously that will be lead to autocorrelation.

So, now we come to know what is autocorrelation? Autocorrelation is the problem of having linear relationship among the error terms. Like multicollinearity what is the equation of multicollinearity? Multicollinearity is having the linear relationship among the regressors. Now here we are discussing what is the linear relationship means not what is there any linear relationship among the error terms, if it is so then you know the model by itself cannot be considered as the best, so until unless you solve this autocorrelation.

But you know like multicollinearity it cannot be removed completely autocorrelation also cannot be remove completely there is some range, if autocorrelation will live on that particular range then the model can be considered as the best model. Still some component has to be satisfied. But in the mean time, if it is going beyond that range, then you have to redesign or reformulate the model or estimate the model till you get the best fitted model alright. So, this is how the starting point of autocorrelations. Now I will little bit highlight the way we will design the autocorrelation problem.

(Refer Slide Time: 30:58)



The structure is that you see here Y t equal to beta 0 plus beat 1 X t plus U t since we are here you know here all together we have Y t U t then we have X t then we have U t. So, there are three variable all together so u t of course, we will receive after having the estimated model the difference between you know Y t minus Y hat will give you the U t hats that is error component.

So, now once you have this particular models but we are not going to introduce log directly here with U Y t or you do not like to introduce X log with respect to X t so we are very much interested here to introduce log with respect to U t. So, if we will say Y t equal to beta 0 beta 1 X t plus U t so then, derivation definitely there may be some functional form of U t where U t equal to function of U t minus 1. So, that means you see sometimes what happens when you are doing some process or work then there will be some mistake so the mistake if we will call it is a errors. So, that means you know you are doing the activity continuously with respect to time. Now let us say this is one time period, this is another time period, this is another time period. So, now here when you are in the current time period then, you have to see how much mistakes you are doing.

So, now in the second time periods then again you have to see how much error term error you are committing so that means there is enough chance that here presently means here present level of error term error may be depending upon the first error terms for instance, you take it theoretically. In theoretically, just like you know it is called as a failure is the pillar of success. So, you know when you will say this statement failure is the pillar of the success the term failure itself is the first indications. So, if some things failed means it is in fact it is gone. So, that means once failed means it is log component which is otherwise in statistics it is called as a error component.

So, now so failure is the pillar of success that means we are what is mean of failure and success that means some things between you know plus minus for instance you see let us take case of profit and loss you can say profit and loss just like you know win and loss we will take example of profit and loss. So, how will you calculate profit and loss? Now, when the revenue is greater than cost and when revenue less than cost so now loss is say revenue greater than less than cost so, that we can saw the we can assume that in this you know failures this is failure and this is the success.

So, that means we are assuming that profit is a function of you know failure means your success is function of failure so if we will say failure is the pillar of success that means success is a function of failures. So this is the success of failure, that means this is first event and this is current event so similarly, so the success may be at the highest level at the lowest level like this so obviously we are assuming that so when we will do the war continuously one after another one after another then, obviously committing of error will be going down so this is our observation or it may be I think sometimes it is by default you will get some kind of less errors.

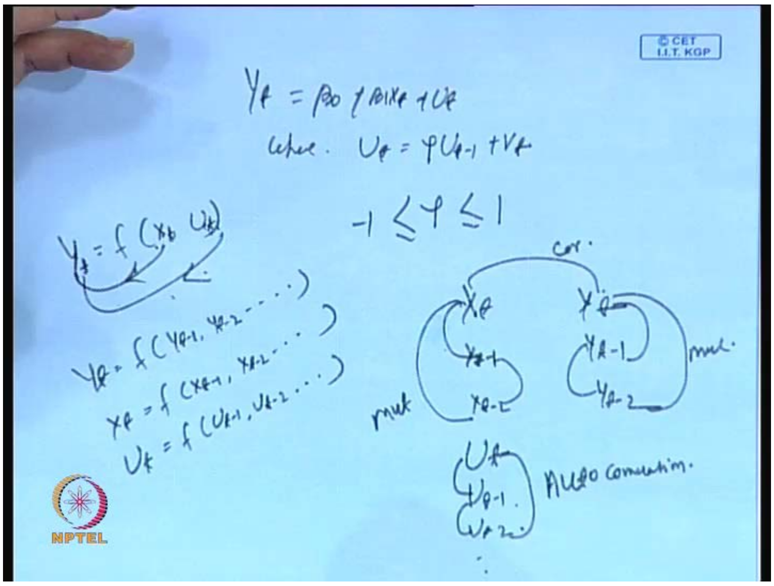
Because, once you practice one after another times continuously then, obviously the error will be you know will be tends to 0 ultimately it cannot be exactly equal to 0 but it will be close to 0 but initially the error may be very high suppose you are a beginning you have 0 knowledge of statistics so I ask you to enter few data(s). So, then obviously you make lots of mistakes but you know I will again ask you to type the data another day again I will ask the data to type the same data any another day then obviously I will check how much error you are committing mistake in the first and second and third and fourth. If you will take hundred samples like this way then in the most of the cases the error will be in a decreasing trend. So, that means it is the learning or it is the failure which you can say make you success or you can say error tends to 0.

So, that means theoretically there is linkage every times so, that means if something is present then there is some connection to failure. Now, the moment you will say U t this is the failure or error terms so that error term is because of the first failure only. So, first you make some mistake that is why there is error term for instance why there is failure or why there is loss? Because revenue is not greater than to cost here so that is why loss now we will turn to profit now, when we will turn to profit say then you know profit has a function of not only revenue and cost it will be it will be another function cause may quench loss factors failure.

So, that means I will call it in the profit case it is a profit is a function of revenue and cost and in addition to that I can say that the lessons you learn from the failures, lessons from the failure. Lessons from the failures may be another variables which can turn you to go for profit levels or which you which you make or you can say which you make an attempt you can say to go to the success levels. So, that is how the you know autocorrelation coming into the picture so that means when you will p t m models Y t equal to beta 0 beta 1 X t plus U t then obviously our autocorrelation problem means to discuss autocorrelation problems we hypnotically assume that this this error terms U t is not at all independent variable.

So, there is some connection which you see first observation or first sample points if that is the case, then we have to check it. So that means, our idea is that whether this error term is completely independent or you know completely dependent that means if completely independent then there is no other errors. So, that means we will create artificial error then we will find or we will justify that there is no significant association between the two. If there is significant association between the two, then there is a problem. If there is no such significant association, then obviously there is no such problems. So, when the problem will be always there, when there is significance association between all these error terms. So, that is how the structure is all about. So how do we write for that?

(Refer Slide Time: 38:17)



So that is in the simple way we will write like this Y t equal to beta 0 plus beta 1 X t plus U t so where, U t is function of you can say rho U t minus 1 plus V t so I will I will write like this way. So, that means we start with linear models so Y t equal to beta 0 beta 1 X t plus U t now Y t X t we are assuming that Y t is not depend on other variable and X t does not depends upon other variable, that means when we will set Y as a function of X and U so then obviously, we are assuming that Y X U are completely independent so that means only three variables which we can integrate properly.

So, that means we like to know what is influence of X and what is influence of Y on U on Y so that is our objective what we have discussed earlier. Now, when we will think about log modelling particularly then obviously so I will I will put it here Y t X t then U t so for simplicity. So, now when we will put Y t X t U t so there is problem is not only to find out whether there an association between Y t X t, Y t U t or Y X t U t so that is the case which we have. We have to discuss before coming to the autocorrelation that is the case. So, now in the that means, if that is the case then there is no such autocorrelation problem because we are assuming that they are completely independent but let us assume that they are not completely independent that means Y t may depend upon whether its first item so that means if the problem will be more complicated if Y t is function of Y t minus 1 Y t minus 2 or continue like this and X t as a function of X t minus 1 X t minus 2 and continue and U t is function of U t minus 1 U t minus 2 and continue.

So, that means so we start with Y t as a function of X t and U t but in between we find Y t is also function of Y t minus 1, Y t minus 2, Y t minus k. X t as a function of X t minus 1, X t minus 2 up to X t minus k similarly, U t can be function of U t minus 1, U t minus 2 U t minus k. For simplicity, we can write simply U t is function of U t minus 1. So obviously, U t minus 1 as a function of U t minus 2 similarly, U t minus 2 is a function of U t minus 3 so that means if you will integrate all these totals then U t has a function of U t minus 1 U t minus 2 U t minus 3 U t minus 4 like this.

So now, so that is the case where we have to look means we have to very carefully consider. So, that means here Y can be expanded with respect to means if you apply log then X can be expanded if you introduce log and U t can expand if we again introduce log. So that means but here our discussion is to talk about the autocorrelation now when there is question of autocorrelation that time, we have no serious business about Y log and X log but we have very serious business with respect to U log that is error terms. So, that means U t as a function of U t is a function of U t minus 1 but generally U t lie between minus 1 to plus 1 so it is minus 1 to plus 1 so like we have discussed correlation coefficient; correlation coefficient is a degree of association between two different variables say X and Y two different variables X is one variable Y is another variable.

But, there is some correlation coefficient between X t into X t minus 1 like this, when there is X t and when we have Y t then when we will relate then that term is called as a correlation. So, now with X t we can have X t minus 1, we can have Y t minus 1, we have X t minus 2, we have Y t minus 2. So, now there may be correlation between these two then there may be correlation between these two this may be correlation this may be correlation this may be correlation this may be correlation, this is how the entire structure is all about.

But, you know when we discuss these with these then it is econometrically in a positive side. So, that means if there is any association between two different variable there is such meaningful interpretation with respect to econometric modelling, it is as such a positive instrument but now if we will correlate X t with X t minus 1 ,X t with X t minus 2 it is negative issue with respect to econometric modelling because this particular structure is called as not correlation it is called as a collinearity or otherwise called as a multicollinearity, this particular structure is called as a multicollinearity.

So, that means you distinguish here. So, how things are very integrated? So we discuss correlation which is nothing but the degree of association between two variables not two variables two different variables. So, now there may be chance of degree of association between same variables X t X t minus 1 you know Y t Y t minus 1 or Y t minus 1 Y t minus 2 ,X t minus 1 ,X t minus 2 this is one type of problem. If your objective is to track this type of problem so, this type of problem then it is called as a multicollinearity problem.

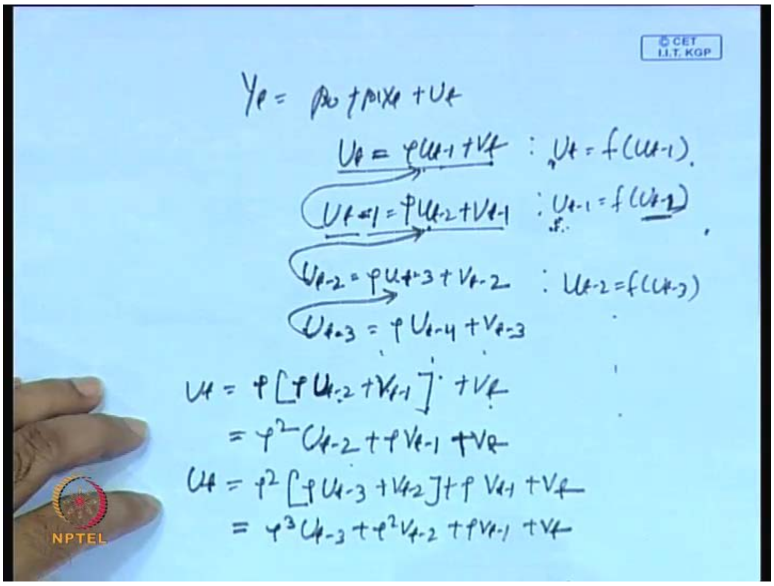
So, now within the X t Y t and if we will apply estimated equation then you will get: U t now U t has a U t minus 1, U t minus 2 like this, now you like to correlate like this you like to correlate like this, if you like to correlate like this then if this correlation found something value. So that means, it is a degree of association if the degree is somewhat positive or negative then there is a problem. So, that means if the degree is 0 then there is no such correlation that means error term is a single variable that is U t only so there is no such other that means, the component failure is the pillar of success we will be not an issue here.

But, autocorrelation is just like this statement of failure is the pillar of success that means every time the present will you know, depend upon past, failure means here we indicate that failure is because of error component only. This failure means fail means that is errors that is how your failings. So, that is how the term we have to integrate with failure and success here we have to integrate the relationship between U t Y t minus 1, U t minus 2 this particular structure is called as a autocorrelation this particular structure is called as a autocorrelation problem.

So, now what is this autocorrelation issue? So, autocorrelation means it is the degree of association between not two different variables not same variable with its log rather it is same between two error terms with respect to log that is what the autocorrelation all about. Autocorrelation means, it is the existence of linear relationship among the error terms and that too with respect to log issue only so that is what the autocorrelation all about.

So, now once you get to know this autocorrelation then we have to discuss so many issues with autocorrelation. Let me first highlight the entire structure of autocorrelation. How these error terms are well connected, and how it can be very potential and very influential so far as a econometric modelling is concerned that too best fitness is concerned.

(Refer Slide Time: 46:44)



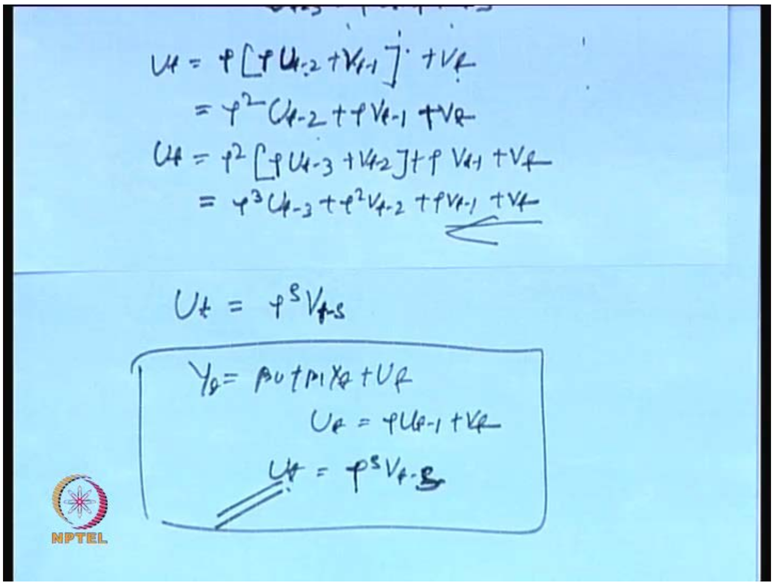
So, now you see here Y t equal to beta 0 plus beta 1 X t plus U t so then what we have mentioned U t equal to rho U t minus 1 plus V t. So, this is how we have to see now that means if we will put it like this, then we can call it U t is function of U t minus 1 so similarly, there may be chance that U t minus 1 as a function of U t minus 2 because same way the failure is the pillar of success.

So, when you are saying that failure is the pillar of success. Then again, when you are here then you are assuming that this is one type of success it may be with error but it may be at the lower level this error this error will be different. So, this may be less error so if this is the less error then obviously, so then it is you know the error is again in the past so, that less will be because of this is this knowledge only. So, that means every time there is well connected so that means U t minus 1 is function of U t minus 2 similarly, U t minus 2 as a function of U t minus 3 and it will continue like this.

So, now if U t as a function of U t minus 1 then the explicitly format is like this. So, let us assume that there linearly related to each other because we are right now, we are discussing the linear association between or among the error terms. So, that is why we have establish a linear relationship so that means Y t Y t influenced by Y t minus 1 so it is the Y t minus 1 which influence U t so that means it is supported by another term called as error component :V t. So that means this error term may not be exactly this U t may not be exactly depends upon its past it may be because of some other factor also.

So, that has to be taken care by another error term called as a V t but V t is not the groups of U t, U t U t minus 1, U t minus 2 like U t minus k it is one group. So, that means all these items will be derived from U t but V t may not directly derive from U t it is committed error which can be also know influence on the present U t.

(Refer Slide Time: 49:08)



So now similarly, we will write here U t equal to U t minus 1, U t minus 1 is equal to rho U t minus 2 plus V t V t minus 1 so this can be written. So similarly, U t minus 2 it can be written as rho U t minus U t minus 3 plus V t minus 2. So similarly, we can write U t minus 3 is equal to rho U t minus U t minus 4 plus V t minus 3. So like this it will continue. Now, ultimately you see here, so U t is function of rho U t minus 1 plus b t but rho U t minus 1 again is a function of U t minus 2 so that means U t minus 1 can be transferred way then again U t minus can be put it here, then again U t 3 minus can be put it here. So, then if we will put all these items sequence then you will get another structure or another interesting models.

So, let see what is that interesting model? Now, what you have to do? So, you put U t U t equal to rho U t minus 1 so what is U t minus 1? U t minus 1 equal to rho this is rho already so rho U t minus 2 plus V t minus 1 V t minus 1 then, it is you know this U t minus 1 I have introduced here plus V t.

So now, you simplify this one if we will simplify then this is rho square then this is U t U t minus 2, this is V t minus 1. So rho square U t minus 2 plus rho V t minus 1 rho V t minus 1, V t minus 1 plus V t. So this is how the structure all about. So, now similarly, what you have to do you will put U t minus 2 in this particular rho function then rho square into U t minus 2 what is U t minus 3 then this is rho U t minus 3 plus V t minus 2, plus rho V t minus 1 plus V t. So, this is how the then if we will simplify you will get rho to the power 3 U t minus 3 plus rho square V t minus 2 plus rho V t minus 1 plus V t so this is how it will expand.

So, now you see here so the way we will expand this way then obviously similarly, if we will put here in generalize then it will be ultimately come into rho to the power s then v t minus s. So that means ultimately, ultimately you will come to a point where there is no U terms so ultimately the entire U t will depends upon another series of error terms which called as a V t so that means it is V t V t minus 1, V t 2 minus 2, V t 3 minus 3 it will be influential factor for U t. So, this is how we already observed that means so our model is here Y t is equal to beta 0 beta 1 X t plus U t and U t is equal to rho U t minus 1 plus V t and similarly, ultimately if we will simplify U t equal to rho to the power s V t minus s rho to the power s V t minus t minus s so this is how the multicollinearity issue.

So, now we have to see the mean of error term should be exactly equal to 0 then you know variance of error term should be some constant component. So ultimately, this is how the autocorrelation is connected to this econometric modelling. So now, how it is active? And what are the causes? And how you have to detect? And what are its solutions? What is its feasibility? All these details we have to discuss in the next lecture. So, for this time being we have to stop here, thank you very much. Have a nice day.