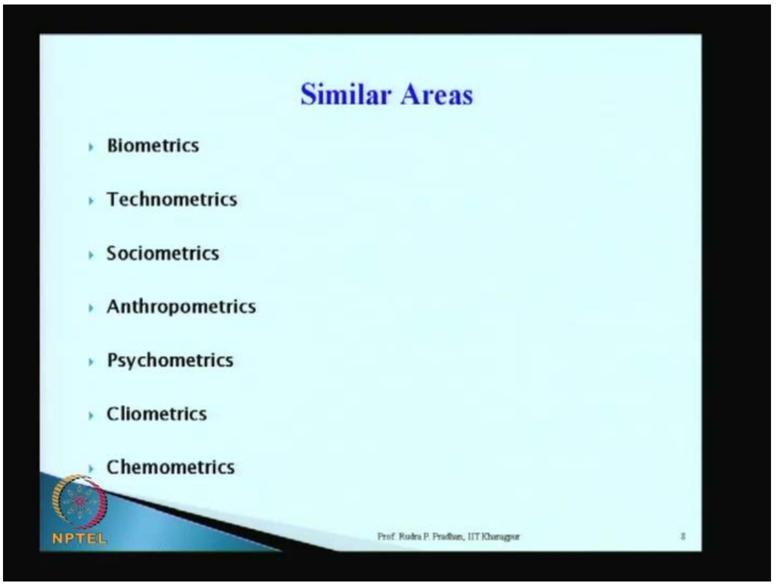
Econometric Modelling
Prof. Rudra P. Pradhan
Department of Management
Indian Institute of Technology, Kharagpur

Lecture No. # 40
Concluding Remarks

Good after noon, this is Doctor Pradhan here is welcome to NPTEL project on econometric modeling. So, this is our last lectures for this particular course econometric modeling. So, today will just summarize what we have discussed, in our couple of lectures. So, this is the summarize fact what is all about econometric modeling how we have started and how you have to contain end. So, that is our agenda. So, today's discussions, very specifically we start with you know basics of econometric modeling. So, we define what is all about econometrics and how it is fitted with modeling concepts then we have means, what is econometrics, why you know econometrics and how we have to integrate with modeling then in the same times we heavy discuss various features various applications and a various.

(Refer Slide Time: 01:38)



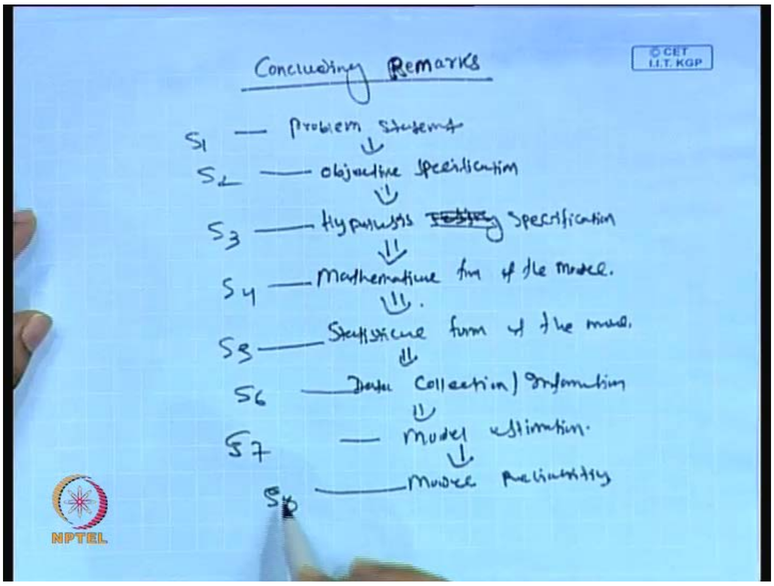
You know structures and also related area like you know Biometrics, Technometrics, Sociometrics, Anthropometrics, Cliometrics, Chemometrics etcetera. These are all related area's we have means econometric has an applications so; that means, a we can make a guess what is the importance how it is important. So, for as a feature is comes on means, that it is with respect econometric modeling. So, econometric modeling is you know fundamental it is a it is root from math's start endogenous the economics. And you know it has you know spreader our many areas as a result there are many applications, now a days without econometrics it is very difficult to handle so many problems so, what I will. So, write know I will just briefing the concepts what we have discussed, in the last lectures.

So, we started with you know various definition about econometric modeling that to you know there after we have discuss, basics of econometric modeling that two univariate statistics, vicariate statistics, multivariate statistics and also it is related concept like hypothesis testing probability etcetera. So, then you know we have category divided by into various structures through, which a we get to know the concepts of econometric modeling and you know what you can call it is a integration with business forecasting. So, today what I will do I will just give you briefing about this particular structures.

So, what we can do here is because in the beginning I started with you know the modeling strategy. So, now, we are in the and so, we have to summarize how the modeling starter begins and how it comes to end infect it has a many areas many application many branches. So, it is very it is not state forwards. So, we have to start here is and we have to end here is, if it is a particular technique then you cannot what is this starting point and what is the ending point. But here the thing is that we have in between we have discuss, so many means techniques, so many issues, so many problems, so many starters. So, it is very difficult to you know a put it in a in a summary forms. So, because every application as a different issues, different application and different starter

So, we will give you general frame work here is where's. So, how you have to start the econometric modeling. So, what are the procedurals measures, so, how we have to begin, so, then how you have to analyze and how you have to end. So, that is so, we will briefing this particular starter strategic issues then will come to the summary part.

(Refer Slide Time: 04:38)



So, we started with you know the means the modeling starter completely for econometric modeling begins with problem statement. So, we stood we start with problem statement because without having complete problems earlier got issues then econometric modeling cannot be applied properly. So, without any theory without any serious problems at it is very difficult to say, which particular econometric technique you to apply to analyze that particular issue. So, problem statement is very important that is the beginning of this econometric modeling. So, with the problems statement.

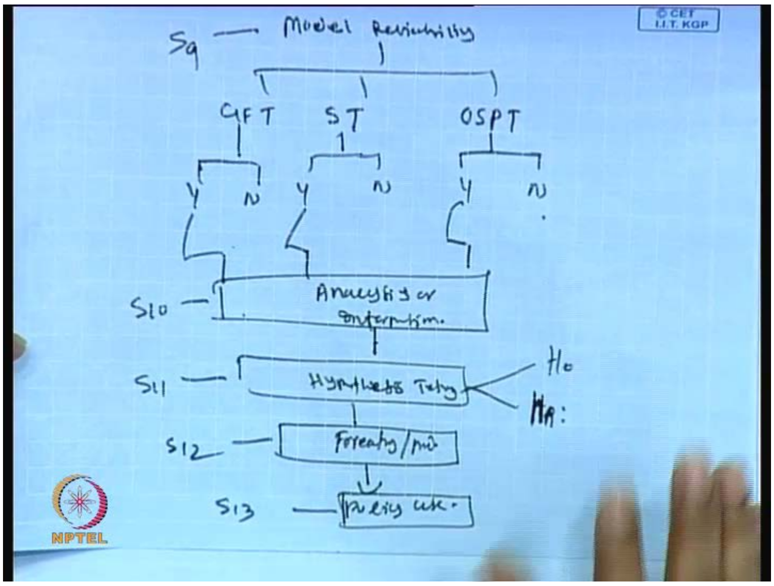
So, we have to come out with the objective specifications objective specifications we have to comma to with objective specification, then followed by hypothesis testing hypothesis testing so; that means, once you know the problem statements. So, with respect to particular problems so, we have to form objectives and corresponding to that objective you have to set the hypothesis this is agenda hypothesis is a you know just to investigate. So, hypothesis means it is a statement, which is not verified which we like to verified it is usually handle through the integration of null hypothesis and alternative hypothesis. So, which, we have discuss long back.

So, the first step is problem statement, then second step is objective specification, third step is a hypothesis a not hypothesis testing hypothesis specification. So, we will go so, testing later part hypothesis specification so for as testing is concerned. So, we have to apply the null hypothesis for one alternative hypothesis for done. So, then after getting or hypothesis specification, then will go for formulation of mathematical form of the model. So, this is we have to formulate the mathematical form of the model then so, we have to move to statistical form of the models, then it has to the followed by data collections, because you see here data collection or you can say information. You see so, once you have problem you have to specify a particular objective or many particular objective. So, it may be single, it may be multiple objectives.

So, corresponding objective have to formulate the hypothesis now, once you have hypothesis in front of view. So, according, to the problem specification and hypothesis setup. So, you have to formulate models that are mathematical form of the model. So, now, that mathematical form of the models it has to be transfer into statistical form of the model, because we need to verify whether that model is perfectly or not. So, now, because we have to apply starter with respective that model. So, statistics has to investigate whether that particular models, which we have design for that particular specific means specification of problem.

So, whether it is fruitful or not, so, we have to verify that one first in statistically then we have to apply it is in a real life situation. So, now a once you have statistical form of the model. So, then what you have to you have to go for estimation, before estimation so, you need information and you that is nothing but you need data. So, data collection and information must be theirs. So, then we have to go for model estimation, then we have to turn into model reliability. So, we have to we start with the problem statement objectives specification, hypothesis specification mathematical form of the model, statistical form of the model data collection then model estimation then model reliability. So, now, what is all about this model reliability let me here explains.

(Refer Slide Time: 08:46)



So, now model reliability has three parts one is called as a goodness fittest G F T then specification test S T then out of sample prediction test voice P T sample prediction test. So, G F T test it will give you overall fitness of the model. So, this is a with respect to parameter weight age, then this is with respect to residual aspects so; that means, it is most only diagonal D J test. So, now o for as reliability is concerned. So, we have three different tests. So, goodness fit test, specification test out of sample prediction test so; obviously, the test will be clarify you whether this model is a reliable one or not so; that means, every case, there is a yes no situations. So, every time there is a yes no situation.

Let us, assume that the every case there is you can say yes yes yes. So, now, if it is say yes yes yes, if it is yes yes yes. So, then you have to go to next part of the model it is called as an analysis. So, we have to analysis analyze this a analysis or you have to interpret the models interpretations of the models So, now with respect to interpretation we have to go for hypothesis testing hypothesis, then you can say forecasting. So, once hypothesis nuclear here we have to apply null hypothesis and alternative hypothesis our m is to reject the null hypothesis assume to be a correct statement S what we have to reject this statement. So, that a we have to we have to prove the factor.

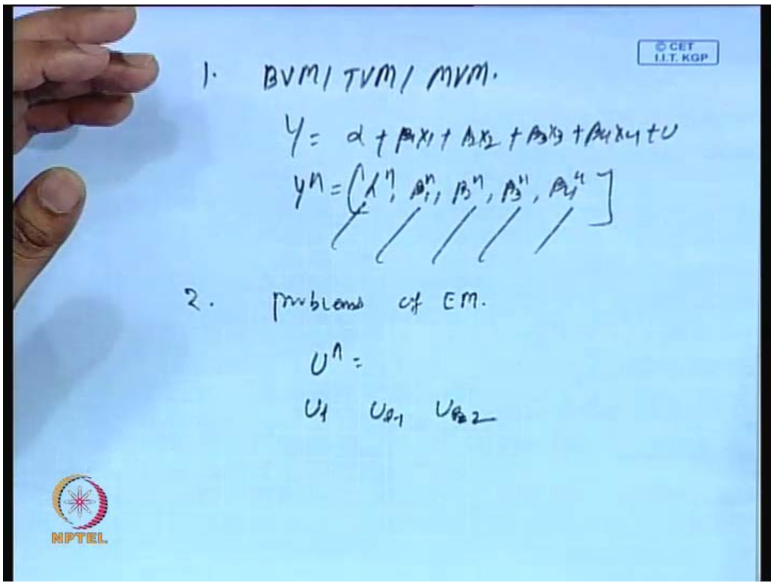
So, hypothesis testing then it has to be a forecasting or predictions then you have to finally, you have to use for policy use this is these are the steps, which we have to insert in the case of means, these are all modeling strategy. So, that means, what is our structure here is, so, if you quality step one, then this is step two, this is step three, this is step four so, this is step five, this is step six, this is step seven, this is step eight. So, then followed by this is step nine. So, now, in the step nine so, what we have done? So, we have gone through three different tests. So, if all this states are clarify that the model is reliable then we have to go through for step number ten, followed by step eleven, followed by step twelve, then followed by step thirteen.

So, these are the number of steps, which we have inserted. So, these are the following a following starter you can say following steps, which we have to follow in the case, of modeling starter, but if it is a in case no no no then; obviously, we have to go back to a again, step four. So, if it is no then we have to go back to step five step four then again, you have to read a join the mathematical form of the models. You have to read join the statistical form of the model and you will same data structures. Then you have to go for model estimation, then you have to go for model reliability, then again, you have to come to model reliability you have to go for specification test goodness fittest.

And you know out of sample prediction test then, if all this will clear, then obviously, will go for interpretation analysis then followed by hypothesis testing for then you have to test for prediction and forecasting. Then finally, that model is useful for means you available for policy use, but if it is again non then again, you still go to step four again you have to realigned model still a you get the best model through, which we can predict the fact or you can say predicate structures. Then you have to analyze this means, you have to use for policy you have to apply this model for policy use. So, that is the modeling starter, which we have started in the early face of our lectures so, with keeping in mind all this strategic rules.

So, we have to we have to enter each and every problem then accordingly, we have come to a conclusions. So, in between the interpretation analysis are very complicated and very you know interesting sometimes it is very starter. So, what we have already discuss, all this details. So, I am not going to discuss each and every component right now. So, what I have what I will do now. So, I will just briefly, highlight what are the techniques we have discuss, and we will highlight little bit a what is; that means, what is all about that techniques and final will give the conclusions.

(Refer Slide Time: 13:54)



So, the modeling started with we start with you know the regression analysis particularly, so, bivariate modeling, then trivariate modeling, then multivariate modeling. So, we started with regression analysis that is the entry for and what we discuss, regarding probability hypothesis is then introductions then related things what we have discuss, these are all in basics these a papers, but basically, the real fact is starting from the regression modelling that two bivariate modelling, trivariate modelling and multivariate modelling.

So, we have gone through details estimation process, because the basic fact is that. So, the moment you will let transfer mathematical form of the model to statistical form of the model. Then ultimately we need to estimate the parameter's and where we have to minimize the error terms every times where here a standard fact is that. So, the structure is depend one dependent variables and several one or several independent variables and something error terms, which you know, not captured in your models that will take care of. So, there are three parts all together. So, dependents structures independent structures then error structures.

So, the way will minimize the structure minimize the error term then we will be get alpha means a parameter estimated parameter's then through the estimated parameter's. So, we have to go for all this reliability test, that is goodness fittest specification test out of sample prediction basically, what we have dome's. So, we have gone through the parameters statistics that is suppose, this model will be like this y equal to say alpha plus summation you know beta 1 x 1 plus beta 2 x 2 beta 3 x 3 a something like this way beta 4 x 4. So, then what is that this is complete model. So, then we ultimately we get I hats. So, that is alpha hat beta hat's beta 1 hat beta 2 hat beta 3 hat then beta 4 hat.

So, these are the things we need to calculate. So, then so, it has to be significant all this things then followed by a followed by the overall fitness of the model is to be a you can say considerable very high that is r square adjusted r square and also means, moralist that is an our tables analysis a variance. So, we always kept were with the r square adjusted r square a residual some square explain some squares then f statistics and accordingly it is significance label. So, this is how we have procedure we have to conclude this particular component.

So, then after discussing, all this things all this things with respect to bivariate modelling trivariate modelling multivariate modelling infect in the multivariate modelling we have discuss, generalize a plus and symmetric approach so modeless effect is that. So, ultimately, we need to have estimate the parameters then we have to check the significance of the parameters and finally, we have to check the overall fitness of the models and finally, we have to go through all this you know diagnostic test. So, these are the facts we have to study under all this conditions then finally, we have discuss various problems aspects problems of you know econometric modelling.

So, that two we have discus, multicolinearity issue, autocorrelation issue, heteroscedasticity issue and so other diagnostic test so, multicolinearity issue is basically, a give you this signal of the means when we if it a model. So, generally, we start with the O l S technique to minimize there are some squares. So, where we have standard assumptions, the standard assumption ultimately turn into all this problems aspects sub regression modelling. So, multicolinearity one of problem where we like to know what is the existence of linear relationship among the regression that is independent variable. If all this regressors are independent then; obviously, model is completely and we can consider as a best fitted models otherwise, it is a we have to go round after round till you get the best fitted model. This is all about the multicolinearity issue.

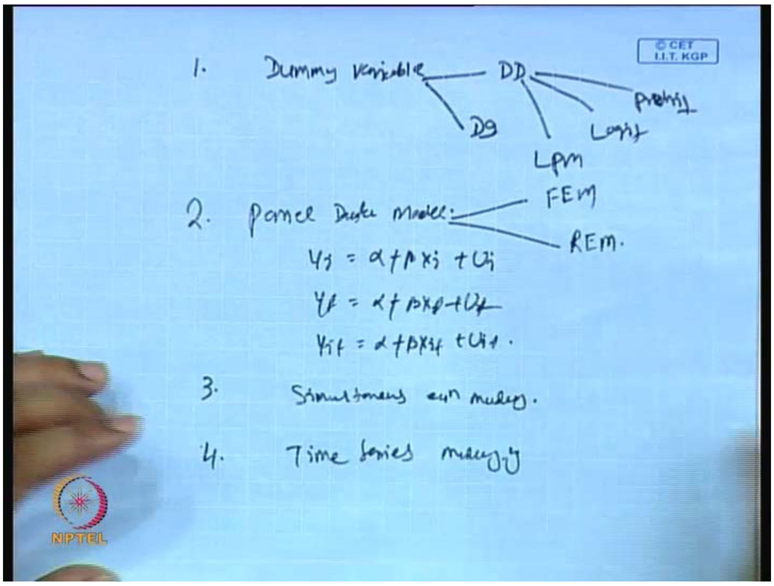
Then autocorrelation issue so, it is a problem regarding the error terms. So, means, when we fit an error terms when will get error terms, why you hat and obviously, will get U t

U t minus 1 U t minus 2 etcetera. So, we like to know like you know like the a relationship it in your among repressors. So, we may have relationship among the error terms U t. U t minus one U t minus 2 like this or U 1 U 2 U 3 etcetera. So, these term should be also independent so, but the reality is there cannot be independent. So, there is a little bit they may have interdependent. So, we like to see means our standard assumption that the correlation or covariance between U t error terms should be equal to 0, if it is not 0 then there is autocorrelation problem.

So, if the there is autocorrelation problem means, existence of autocorrelation then we have to see the range with this range we have to see whether, we can we can proceed for prediction forecasting or we have to re redesign this model. So, this is the fact of autocorrelation then finally, heteroscedasticity also problem of error terms it is related to problem related to error terms. So, in the case of autocorrelation we take care the covariance issue, but in the case of heteroscedasticity we take care the variance issues so; that means, error our standard assumption is that error variance should be equal constant for each and every unit. But if error variance or not constant then; obviously, what you have to do we have to redesign get the structures where the error variance will be somewhat or constants.

So, then that model can be constant as the best fitted model. So, autocorrelation is covariance of error terms this is variance heteroscedasticity variance of error term then a in the mean times we have discussion, so many other issues digenetic tests. So, like you know goodness of fittest of models and also specification model, like the usual P I C statistics a chi information criteria final production criteria etcetera. These are the components through which we have to come to a conclusion whether, this model is consider means be consider as the best models are not.

(Refer Slide Time: 20:40)



So, these are the facts we have discuss, in this you know entire structures then we have also discuss, very hardcore problem. We said, this problems we have discuss, severe problems like you know dummy variable modelling. So, in the case, of dummy variable modelling, so we have discuss, under to hats dummy dependent dummy independent. So, under dummy dependent we have discussed linear probability models. So, then legit models then probity models these are very interesting component a probity model. So, dummy variable means a this you know before dummy variable whatever, concept we usually discuss this here one standard assumption is that all this variables are quantitative format and readily available.

So, we have to just apply the technique and you have to get through the results and after that you have to go for all reliability testing etcetera. But some of the problems we will find the variables are categorical binary in nature. So, in that case, it is not available in directly in quantity forms so we have to apply the retain variable technique. So, once this type of problem is their it may be, a with respect to dependent variable or with respect to independent variables. So, if it is dependent respective independent variable is called as a dummy independent problem's, if it is related to dependent variables then it is called as a dummy dependent problem.

So, the techniques are medalists. So, it is manipulate interesting. So, we have discuss, with the general format then with here respectively linear probability model or binary just models legit models then finally, probity model. So, after that we have discussed, the panel data models so, because we know we have discuss, the know we have discuss, the enter cross sectional modeling. Then we have discuss, the like you know time series modeling then you know panel data settings like this y I equal to alpha plus beta x I a plus you then y t equal to alpha plus beta x t plus U I then this is a U t then I t equal to alpha plus beta x I t plus U I t. So, these are the three forms of the models which we have discuss, in the enter econometric modeling. So, few problems are purely cross sectional type few problems are purely time series type and few problems are joint.

So, that is what is we call is a panel data modeling it may be in the with respect to pool data or with respect to pure panel data settings so, in the case, of panel data setting. So, we have discuss, to model is to interesting model called as a fixed effect model and random effect model, fixed effect model with respect to time component and you know cross sectional component and a random effect model specifically, which integrates the error variance. So, this we have discussion, in the panel data is variance testing then we have discuss, the simultaneous equation modelling a through, where we have discuss we identification problem that to you know whether, the model is a means it is simultaneous equation modelling means, variables are interdependent in natures.

So, before simultaneous equation models so we have we have this system one way system. So, that is one dependent variable with this several independent variables, but you know most of the cases will find variables are variant interdependent in nature. So, there is you know once there interdependent then; obviously, the system will be there will be system and that system is called as a simultaneous equations modelling. So, with these particular structures so, what you have discussed, the model has to be first identifying that two exact identification over identifications and you can say and a you know on the identification which cannot be identify properly.

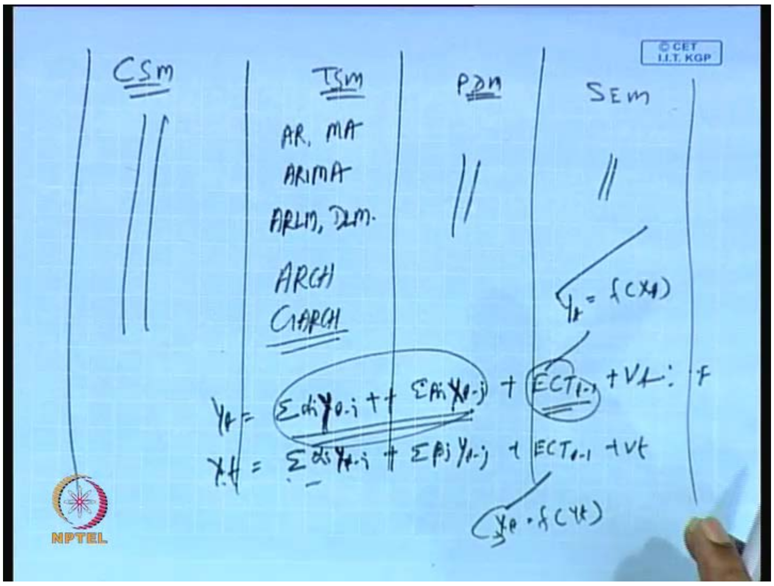
So, then under you know once it is a identify model is a identify then; obviously, will go for a estimation generally, in simultaneous equations modelling O l S is this type of application. So, we have different techniques all togethers for simultaneous equations system like you know a reduce form of the models then indirectly list square method two stage list square method three stage list square method full information maximum likelihood methods. So, many methods are there instrument method of instrumental variables. So, these are the methods through, which we can solve this simultaneously equation problems.

So, then we have discuss the structural equation modelling, where you know the it is just part of the simultaneous equation modelling, but here in the structural equation modelling there will be there should be proper structures. So, and infect in that is structure there is always to a causality. So, it is the structure of you knows we every time will like to know, what is the existence of covariance and correlation among the various items. So, how their interrelated to each other and infect basic interesting. So, because most of the problems in the real life situations are very interdependent very structures, structural equation modelling is very helpful to solve this type of problems, then we have move to time series modelling so time series modelling a time series modelling.

So, we have discuss, various aspects of time series modelling starting with basics a what is this setup of time series frame work then we have discuss, you know various models like you know auto-regressive model, distributive lag model then aroma model, auto-regressive distributive lag model, joint in the case of auto-regressive model. It is the endogenous variable as a function of lag of endogenous variables. Then in the case of distributive lag models it is endogenous variable as a function of exogenous variable and lag of exogenous variable.

So, we can also joint and we have also discussed how the distributive lag model can be transfer into auto-regressive model, because most of the schemes in that time series settings in the auto-regressive way. So, there is you can properly hand directly handle through distribution lag model. So, we can under it will also handle auto-regressive model what there is you a how you have to transfer the distributive lag model to auto-regressive model that is we, which we have discussing in the quick methods. So, then you know we have discuss volatility modeling.

(Refer Slide Time: 27:18)



So, in the case, volatility modeling so, what we have discuss with we have discuss, auto-regressive schemes then moving average schemes then aroma models like the means. Basically, what we have done means, at our structure all together is a cross sectional modeling structures, then time series modeling structures and panel data modeling structures and structural equation modeling structures these are the various structures we have discuss in enter our you know enter this class whatever, we have discuss, daring the enters you know periods. So, what this is some of the problems are pure cross sectional times some of the problems are pure time series times some of the problems are panel data setups and finally, then there is a structural equation modeling.

So, now I have already mention the structure about cross sectional modeling I have discuss also frame panel data setting structural equation modeling simultaneous structure. Then time series modeling basically, we have discuss, various schemes like you know auto-regressive models, moving average models, then arima models. Auto-regressive integrated moving average model and before that we have discuss, you know auto-regressive lag model infect auto-regressive itself is a lag model then distributive lag models. Then distributive a there are many forms of the distributive lag models like you know polynomial distributive lag model, then infinite distributive lag model, then you know a lag means auto-regressive distributive lag model there are many ways you can a discuss, this issues of you know time series modelling.

So, then we have discuss, again, the volatility modeling like you know form this auto-regressive schemes we have A R C H models we have G A R C H models and we have E G A R H, T G A R C H etcetera. But to each specifically, highlighted the issues of A R C H model auto-regressive conditional heteroscedasticity model and generalize auto-regressive conditional heteroscedasticity model may the heteroscedasticity problem we have discuss in the case of cross sectional modelling. But now, we are using this you know this particular problem in that time series setting generally, before time series modelling we our standard assumption is that are standard thinking is that autocorrelation is time series problem heteroscedasticity is cross sectional problems.

But in the real life situation or you can say in effect it is fact that correlations autocorrelation heteroscedasticity can be observe in the case of cross sectional modelling can be observe in the time series modelling. So, this is with regards to with regards to with regards to volatility modelling then we have discuss, various means specific problems of a time series modelling that two unit root unit root problem co-integration problem causality problems then vector error connection model. So, unit root test there are several test infect, but we have discuss, dickey fuller test augmented dickey fuller test then finally, we have discuss means a then co-integration that is the existence of long run equilibrium relationship will be tent to or more variables.

Infect that is the second part of the time series modelling before that unit root means stationary of the variables we must properly check then if there is the a something means, if the variables are integrated up order 1 0 2 like this and; obviously, the modeling setup will change accordingly. So, generally, we start with the stationarity issue of the time series variable that is unit root problem then we have to move to the co-integration problems, where we can discuss, the existence of long run relationship. Then finally, will we have to discuss, the causality problem because if we causality means, if you have to variables. So, x and y say then x casuist y or y casuist x before that we have a one directional causality that is one causality, but time series has a specific you know importance that there is a you know both way of causality can be derive.

So, in case of there are two variables then either I they have no causality or they there is a unidirectional causality; that means, x casuist y does not cause x or vice versa y casuist x does not cause y and there may be possibility x casuist y and y casuist x and if that is the case that is called as a bidirectional causality and if you only case of one and x casuist y does not and vice versa then it is called as a in directional causality and if x does not cause y and y does not x then; obviously, it is called as an means, there is no causality.

So, this is what we have discussed, and in the same times we have discuss, the vector error correction models, because mean the co-integration infect. In fact, co-integration we have discuss through is a procedures and you know Johnson procedures that is purely in the matrix say Eigen value and state statistic issue so in the angel methods. So, we have discuss, the linear sensitivity into variable times series variable then there error terms we have observe that stationary in natures. So, after having all this things then when will come to vector allocation models then; obviously, from the co-integrations equations where we have study the reliant existence between to variable then that error correction may be have a significant of impact on a particular variable at the current time period.

So, that is why we have a observe through vector allocation models like this you can say summation y t equal to summation alpha I x t minus I plus you can say plus summation beta I y t minus j plus error correction ECT ECT minus 1 error correction 20 minus 1 of course, there is a v t. So, similarly, x t can be retain as a summation alpha I a alpha I x t this is y. So, it should be y t minus a it should be x y t minus I this should be x t minus j. So, similarly, in this case x t alpha I x t minus I plus summation beta j then y t minus j plus error correction terms t minus one plus v t. So, this error correction term when will be integrate y as a functional y t as a function of x t that is a co-integration equation in that case x t as a function of y t that is co-integration equation.

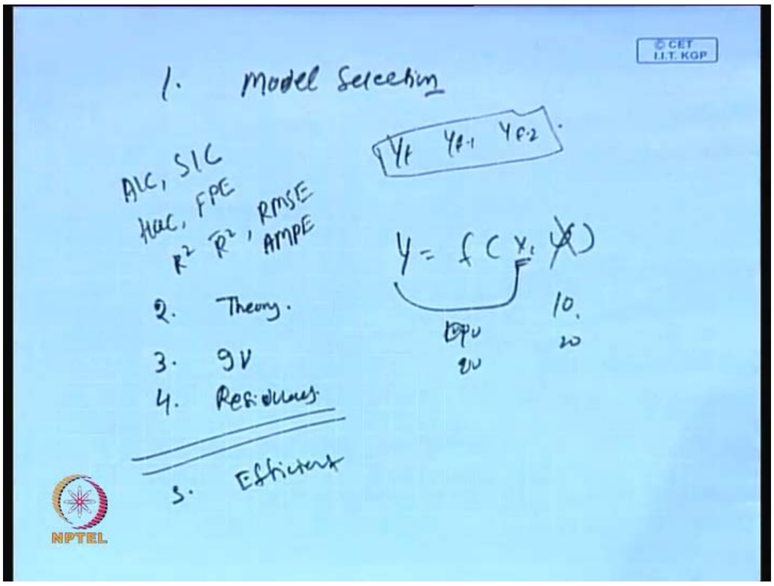
So, we like to see whether this as a impact and this as impact this is called as a long run impact these are called as short run impact this is to be investigator through f statistics so; that means, will be already discuss in our last lectures. So, now, with these particular specifications so, then we means, these are the complete structure, which we have discuss, within this particular semesters. So, what do we have to means what I to do right know. So, I will just giving you a brief frame work of modeling starter. So, what is your core idea our how you have to come to a specific structures through, which you can we can enter a particular problem very a state forwards and you can also gate the answer any accurately and you can also apply this particular model in a very authentic.

So, that is how it is very important to loan some strategic issue and strategic facts. So, here is the model a precisely highlight you know we can say various strategic rules are there it is not that you have to learn only techniques. And you know only to learn problems infect problems is always theirs objective we have derive hypothesis you can set there are several methods are there you can get to apply. But you know, which particular method you have to apply and how you have to interpret that parent testing and that is more and more you know complex game.

So, it is very important to know strategic issue how you have to enter to that problem then how you have to get out the solution it is not like that we were just entering the problem and very different to come to a particular conclusions. So, that is why, if you know very proper structure and proper setup and then; obviously, will quickly enter to that particular problem and you can get quickly answers. And that answer will be very perfect can be use persons some policy use otherwise, you just it is you know, if you has a (( )) are just going to estimate something or you can say doing something in that is not the econometric modeling.

You know, you have to apply your strategic idea you have to bring the analysis in a strategic way that is more important. So, I will be briefly highlight be important points. So, that will be you know you can consider these are also strategic rules for econometric modeling.

(Refer Slide Time: 36:48)



So, first thing is model selections, first is a model selection you see awe have all together difference shut up of models cross sectional modeling time series modeling panel data modeling structural equation modeling. So, every times means must of the cases, when you are in the huge problem area and differ in the side then; obviously, it will be problem will be modeless multivariate in natures. So, once your generally, assume that you are in multivariate problem, because bivariate is not the utility of bivariate model is very less while, the utility of multivariate model is very very high. Provided all this parameter should be significant and overall fitness of the model will be very significant then by default that multivariate model will be better chase so now, when will go for multivariate modeling.

So, you must be very carful's. So, whether it is in the cross sectional setup or you can time series setup. So, you have to learn few things and you have to clarify your cell few things, first thing is model selection for instance, if we take a cross sectional modelling. So, if it is multivariate models let us start with you can say twenty variables and thirty variables. So, there may be issues that twenty variables may not be relevant and you know infect you are using unnecessary variables and necessary variables you are excluding.

So, that is how so, models selection criteria are very important. So, it will give you a whether your specification with respect to a inclusion are exclusion variable or not this is one way and in the same kinds to put in the case, of time series modeling. So, we you know we start with the single variable called as a univariate modeling univariate time series modeling then multivariate time series modeling in whatever situation. So, we will create several a variables independent variables with respect to it is here lag for instance like y t y t minus 1 y t minus 2 like this way. So, you know in the time series modelling modeling it is very it is beauties.

So, you can create a number of variables for that, but you know, if you create a number of variables you are degree of freedom will be get a reduce so, you must be very careful. So, that is y how much lag length you have to consider that is very important; that means, in the cross sectional modeling how many variables you have to introduce here how many lag variables you have to introduce, that is in the case of time series modeling. So, in all the cases there is t how you choose this optimum size of the model. So, there are several statistics say A I C statistics S I C statistic is H U C statistics a final prediction error test fix, when we R square adjusted R square and root means square errors these are you know means percentage errors. So, absolute means errors.

So, these are all you can say a these are the methods through, which you can we can observe the model selections whether, model is perfectly or not this is first problem means first starter idea behind this modeling econometric modeling, second is a model must be constant with the theory. So, model must be consistent with theory, because theory will help in the modeling into wage first it will give you introduction means it will give you setup for objectives specification and hypothesis specification and also it will give you a at the end of the modeling. So, that is they when will go for interpretation after getting the reliability of the model if the reliability model is.

Then; obviously, you have to go for interpretation. So, that interpretation basically you may it has a strong connection with a theory what is the theory only basics of theory to interpretation the enter models. That is why a theory must you must learn or you must learn theory very well before you go for a intertype econometric modeling. Econometric modeling is just know, not statistical mathematics because it is purely a application or you can subjects. So, you need to apply math’s and start just to verify that theory or you justify that theory that is the may core agenda of econometric modeling

So, theory is the most so, then a explanatory variables should be uncorrelated. So, that means, independent variables should be totally independent whether, it is a time series modeling whether, it is cross sectional modeling a you know what ever variable’s in your system in that right side of the modeling that is you know in the case, in the form of exogenous variables or independent variables. The should be completely independent if you are having a food multivariate models and all this you know independent variables are not correlated and; obviously, that model is perfectly very fine and perfectly. Whether, it is time series setup or multiple you know cross sectional setup generally, when will enter more and more variables then there may be some interdependence among them.

So, you have to check whether, there is any interdependent, if there no such interdependent then; obviously, it is a you are your absolutely (( )), if it is there then it is a the relates to you can say multicolinearity. Basically, in this particular setup you have to face always theory for an problems first is multicolinearity problem auto-correlation problem and heteroscedasticity problem. So, these three problems are always theirs. So, that is y. So, you must be very careful. So, model must be means and you know this you know independent variable independent variable should be you can say totally independent.

So, for as a best fitness of the model is concerned, if you are getting to that situation then it is otherwise, you have to find out the possibility what range you have to consider where you know your model fitness will not be an effected. So, that is how the it is you know independent variables that is explain at the variables should be uncorrelated then for random variables residuals. That is residuals should be purely random in natures you remember in our model we have a three difference structures all together first structure is a you can say first structure is dependent classification, independent classification and error term classification.

But when we set dependent variables whether, it is a structurally equation modelling or simple cross sectional modelling time series modelling it is not issue in the case, of structural modelling instead of one equation you may have a multiple equation. Similarly, in the case, of simultaneous equation, but every cases, what you have to see there is one equation at least, if you consider equation. There is a definitely, there is a dependent structure independent structure and error structures, but you know error is just to capture the which we have not represent in the form of independent variable for instant like this y equal to function of x and U.

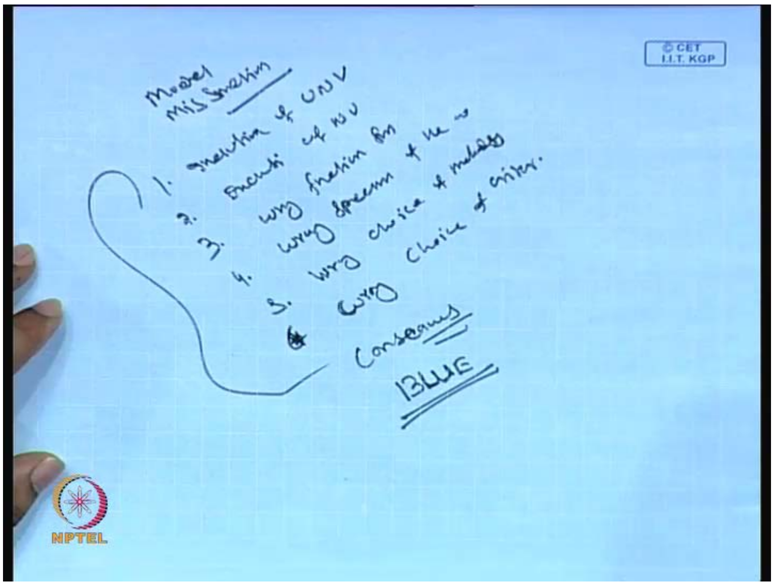
So, we if U is not their then; obviously, the x y is hundred percent x is intensive y hundred percent what it is very it is not possible every times so, as a result. So, if it is ninety then; obviously, ten will go to the error, if it is 80 20 will go to the error so; that means, there is some variable, which you cannot immediately capture. So, that variable should be considered in the form residual, but you know. So, that should be purely random in natures is you know that this variable can be measure can be quantified then issued immediately take and take into the models. So, independent structures, if not, if it is not just like the components are lock so, lock cannot be possible.

So, it is a coming in the error terms sometimes will lock can be also you can say a can be investigated in a quantity we have pi using the dummy variable technique plus that is very rare concept some of the other item also say you know earth quakes etcetera.

So, in that case, it is very difficult to credit all this things. So, it has to be all together coming under error terms. So, residual must be a must be you know random in natures. So, then finally, selected model must be efficient and significant better than other models a model must be efficient.

So, you see we once you go for modelling strategy, which we I have highlighted a long back that a there are various procedurals measures, if it is reliable one then you have to go for analysis and you know hypothesis testing then policy prediction policy use etcetera. If it is no then again, you have go it to the you know original situation, where we started with mathematical form of the model then statistical form of the model, but you know once you say that this model is best. Then; obviously, it should be very efficient than the others and a in other model can be overtake this particular models. So, then that model can be use as a use as a use as a forecasting or prediction or can be use as a for policy purpose. So, this is how model must be efficient; obviously, it will be significantly or substantially better than the other models.

(Refer Slide Time: 46:42)



So, however, if all this conditions are not specified suppose, whatever strategy given to you the these are not with the then; obviously, what are the problems we have to then; obviously, this is called as model misspecification this is called as a model misspecification. So, what is models misspecification; that means, they means what read the wage we have to misspecify the model there are certain cases, one case is a including you know unnecessarily in the systems, then second case is a excluding delivered variables in the systems.

So, inclusions of unnecessary variables unnecessary very very variables second is an exclusion of necessary variables exclusion of necessary variables unnecessary variable necessary variables. So, then wrong functional for wrong functional for then for last, but not the list wrong specification of the random term wrong rank specification or the a random term then finally, wrong math choice of methodology wrong choice of methodology no models specifications like you know criteria wrong choice of criteria this are the things where model will be specified. That means, if you are not statistically strong in a econometric modeling then; obviously, model will default will miss specified.

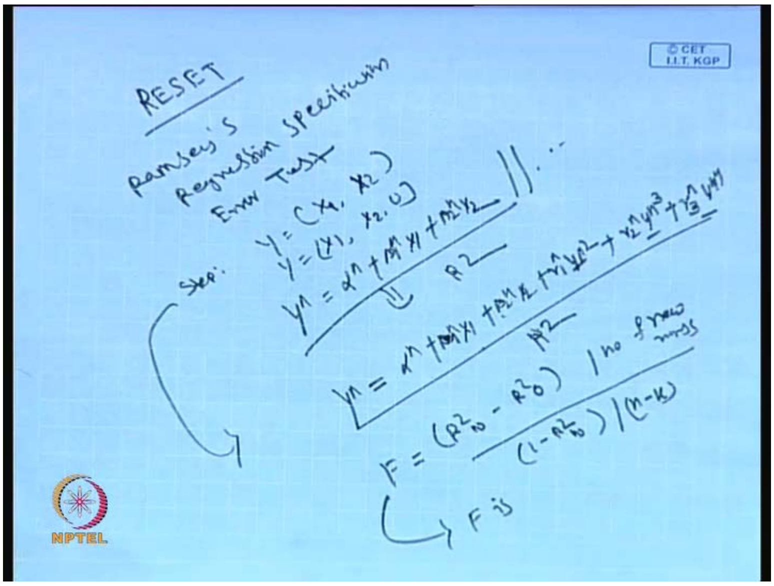
So, how do you know that this is misclassified so; obviously, so you have to go for consequences the consequences, if the model is miss specified then; obviously, the model cannot be we what we equal is a blue theorem is best linear unbiased estimators. The model cannot be a you know follow the blue theorem and it cannot be use for forecasting whether it is no cross sectional modeling or whether it is a time series modelling or whether it is a time data modelling or whether it is a simultaneous equation modelling. So, in all the cases a bit it will be it will be offered the blue theorem.

So, now these is all about the modeling strategy; that means, I have mentioned what are the strategy point we have to consider before you is a strategy point. We have to consider before you go for saying that this model is best fitted models. And because the beginning of the model you are what is the beginning of the structure you do not fit the data into that particular problem. So, that fitting in data into those particular problems it is to get beautiful models. So, that model should be consider as a best fitted model.

So, now it is not easy or it is not state forward to get the best fitted model. So, you know how you have to enter to the center to that particular problem how to get the best fitted model. So, you start with the specific you know idea then ultimately you have to you have to solve that particular problem. So, in that case as if it is not state for what in between why because there are certain strategy rules to you have to follow which I have already highlighted. So, these are the you have to take care the very carefully and I am very issue or it the endogenous of day you will get the models best fitted and that can be use for forecasting.

In any case, if your if you are not hear whether, but still you know, if all this criteria as full fill in fever of you then the model will be correctly specified. Otherwise it is miss specified still even if you get correctly specified then there is a further there is a advance test called as a RESET.

(Refer Slide Time: 50:18)



So, there is a last stage of this modelling is called as RESET this is called as a Ramsey's regression specification error test. So, what is this test? So, it helps as a altogether three for steps. So, first step is a lets we start with a models. So, y equal to function of x 1 and x 2, then what is the what is the starting procedure means, how you to enter with that particular model then; obviously, you transfer into the statistical form of the models, then this is a. So, ultimately you need to have a y hat y hat equal to alpha hat plus beta hat beta 1 hat x 1 plus beta 2 hats x 2.

So, this is estimated model so and we have gone through all this reliability measures and you all this you know you have passed all this things and finally, it is correctly specifies till it are doubtful case. So, what you have to do. So, we have to check whether it is correctly specified or not means a for advance taking. So, you get it here r square then again you estimate the models you have the estimated models by incorporating like this alpha hat plus beta 1 beta 1 hat x 1 plus beta 2 hat x 2 plus gamma 1 y hat squares plus gamma 1 hat then gamma 2 hat y 2 3 y hat to the power 3 then gamma 3 y hat gamma three hat then y hat to the power for plus error term will automatically remove then again, you estimated the strict this model and estimate the estimated model will be like that.

That means, once you estimate the model the estimated a variable will be again added to into that particular problem then you have to again, you to re-estimate and finally, you get all this estimated result and finally, you get the r square value then you construct f statistics f statistic you know r square new minus r square or then divide by number of new regressors. These are all new regressors number of new regressors divide by divide by one minus r squares new I am square new into divide by N minus k. If this f statistics is significant then model is a miss specified if it is not significant then is model is a correctly specified means this is the last option of you know model checking.

So, if it is a essential then you have got three otherwise, if the model is correctly specified all these you know reliability testing like you know goodness fittest specification test and out of sample prediction test that is nothing but the all details about this test. And obviously, model can be consider as the best fitted model, then finally, what is my conclusion is that auto before you know starting a econometric modelling. You must know theory very wells then you design the objectives very carefully, design the hypothesis very carefully design the mathematical form of the models perfectly.

Then finally, transfer this mathematical model to statistical form of the models the moment you will be statistical form of the model that is may that is you know it is beginning of your you know analysis for that is more complex model interesting most strategy etcetera. So, you have to apply very starter where then finally, will get the estimated model then ultimately go through reliability checking then reliability checking goodness fittest specification test and out of sample prediction test then finally, you have to also end of the day will go for Ramsey respected Ramsey regressions specification error test.

So, now if all these things are correctly specified then; obviously, that model can be use for prediction forecasting and also; obviously, can be use for policy use, if not then you have to continuously rotate till you get the best fitted models. So, that is why, if theory depend fundamental concept is a if theory is perfectly and you have genuine theory then; obviously, I am very sure econometric modeling will definitely have five to do something for feature or you can say for casting if the theory is not then; obviously, data will not the fitting of the data may not be a helpful to go for prediction forecasting.

So, you must be very thought about theory all this techniques and all this starter how you have to fit the data in to that particular problem and how you have to get out with the best fitted model. And then finally, how you have to use that best fitted model for forecasting and you can say policy use with this will be and this chapters. Thank you very much have a nice day bye.