Econometric Modelling

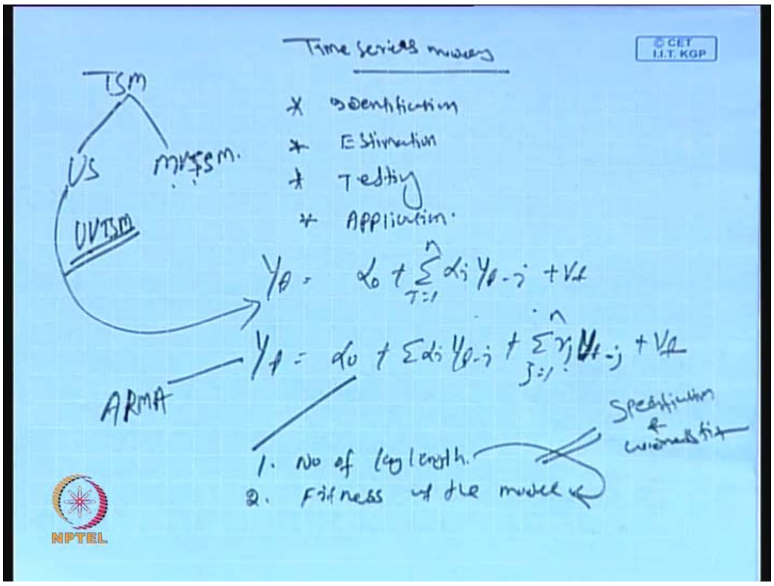
Prof. Rudra P. Pradhan
Department of Management
Indian Institute of Technology, Kharagpur

Lecture No. # 37
Time Series Modelling (Contd.)

Good evening, this is doctor Pradhan here, welcome to NPTEL project on econometric modelling. So today, we will continue the same time series modelling. So in the last lectures, we briefly highlighted about the a time series setup, a you are it is a problem related to time series, how it can be designed with respect to various models like you know, univariate time series models, multivariate time series models, then under multivariate time series models, so we have discussed autoregressive distributed lag model and autoregressive lag models.

So, in the case of univariate time series modeling, we have discussed various issues like a autoregressive schemes moving, average schemes and ARMA schemes, autoregressive moving average schemes. And there is a structure means, just introduce the concept of ARMA autoregressive moving average with respect to orders. If order will change, then obviously the setup will be change accordingly. So with this basic introductions means, with the time series setup, we have to go little bit something more about the time series modeling. So now, first of all what is the basic objective, means I will first highlight, what is the basic objective of time series modeling? What are the problems we will face? And what are the special features we will receive? So how you to proceeds or how we will get out with a best model, and that can be used for forecasting.

(Refer Slide Time: 02:05)



So for as times series modeling is concerned, so there are three important issues. Identifications, then second is estimations, and third is the testing, and of course another model is called as an applications. So that means, there are four important objectives of this time series setup. So you have to identify the models, because it is a lag issue is very important here. So how many variables in the systems, and how much lag length is means how much lag length you have to introduce in the system. So numbers of variables, number of lags you have to introduce in the system, these two important information you have to be very careful. So that is how identification is the most important tricks of the time series modeling.

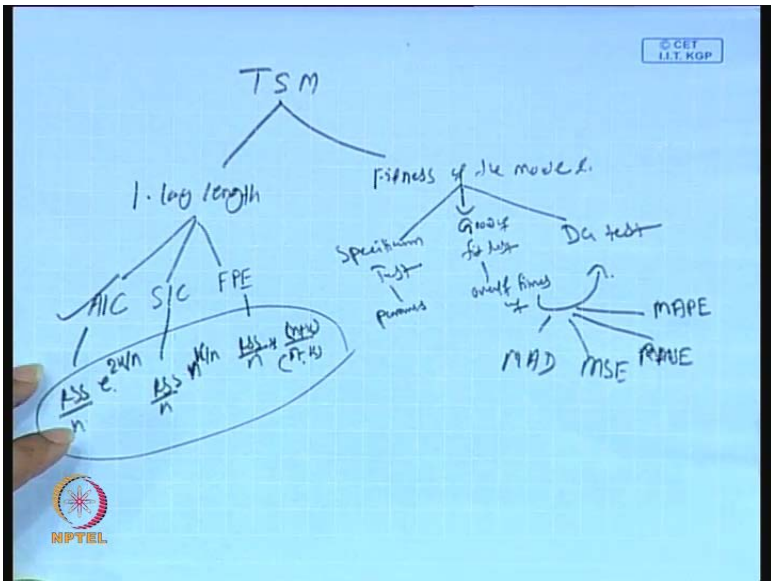
So identification, then estimation the moment you will identify the model, then you to go for estimation, so that is, to get the estimated value usually parameters and all these you know anova statistics that is analysis of variance. Then testing means, as usual, so when you will go for any econometric modeling, you have estimated outputs, then you have to go for testing with respect to reliability of the models, because ultimately we need a model which should be best fitted and we should be considered as the, can we use for forecasting or policy use. So that is why estimation estimated results we need. Then finally, we have to go for testing with respect to specification test, then goodness fit test, then also there is a term called as DG test. I will highlight DG test here, because we have now discussed details in the DG test in the earlier in our earlier lectures.

So here I will typically highlight, because DG test is very important in the case of time series setting. And finally, we will come out with the applications. So now, let me first briefly highlight the little bit structure about the time series setup. So the brief starting point of time series is like this, Y t equal to alpha plus summation, I will call it alpha 0 plus alpha i Y t minus i i equal to 1 to n plus b t, so this is or U t, this is a simple time series modeling. So means, basically time series modeling is in a univariate setup and its multivariate setup univariate setups then it is a multivariate setups. So this is univariate statistics, univariate time series modeling, so it is called as a univariate time series modeling, UVTCM. Univariate time series modeling multivariate multivariate (( )) means multivariate time series modeling MVTS, MUVTSM time series modeling with respect to univariate and with respect to multivariate.

Now, the moment you will have this type of model. So, let me put it in other way, because it is not only the lag length is important. Again lag length with respect to its error term is also important, so that means, if in a broader sense the model will be alpha 0 plus summation alpha i Y t minus i plus summation gamma j Y t minus U t minus j U t minus U t minus j j equal to 1 to n plus say b t, so this is the complete models, this particular setup is called as a ARMA setups, autoregressive moving average, this is called as a autoregressive moving average models autoregressive moving average models. So now you see here is, there are two problems here. First problem, there are two problems first problem is a number of lag length. Then second problem, the fitness of the models fitness of the models. We need to highlight two things separately. So, what is you know, number of lag length, we have to choose in this particular schemes and fitness of the model means reliability of the models. Of course, we in between number of lag length and fitness of the models, in between there is a specification test and goodness fit test goodness fit test goodness fit test. So specification test, goodness fit test must be there in addition to that we will apply DG test.

Now, here is time series modeling basically looks for good identifications, means model must be perfectly with respect to number of variables, with respect to its lag and with respect to its structures, this is how the identification is all about. Then, we have to go for estimation. To get the estimated values of the parameters and all and its related statistics like RSS, ASS and also some of the statistic like you know, SSIC, etcetera; we will highlight all these things. Then you have to go for testing with respect to all these estimated results. We have to go for proper testing to get whether the parameters are okay, whether the overall fitness of the model okay, then finally, you have to use application means you whether you it is use for forecasting or it can be used for policy making. So, that is how we have to be very careful. So now, two things are important number of lag length and fitness of the model. What is the important issue here, in the case of time series modeling?

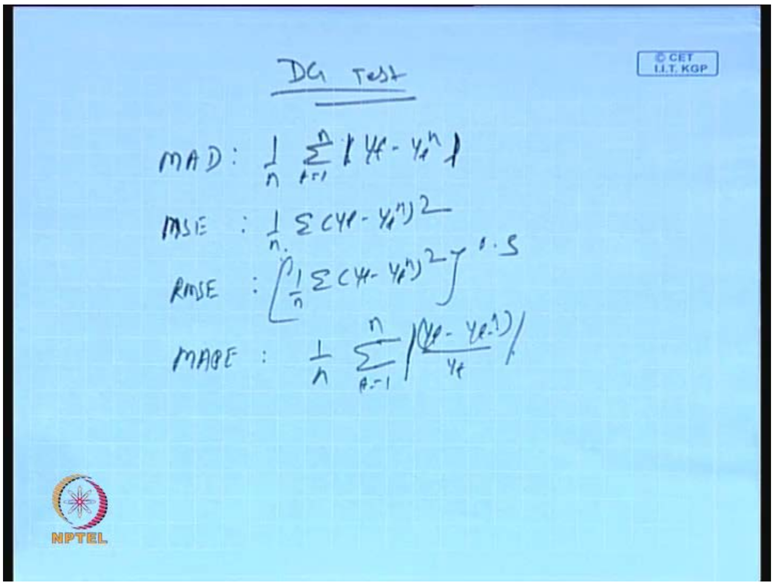
(Refer Slide Time: 08:00)



So, two aspects are very important, one is called as a lag length test, lag length issue and another is called as a fitness of the model. So for lag length test, we have discussed under AIC statistic, SIC statistics, final prediction errors. In the fitness of the models, we use generally in fact three different structures. One is called as a specification test specification test specification test, then goodness fit test goodness of fit test, and then this is DG test. So specification test with respect to parameters, this is overall fitness of the model and this is other way also can be designed through DG test.

So AIC, we have already discussed. The very briefly error AIC is nothing but RSS in into e to the power 2 k by n, then this is SISC (( )) information criteria is RSS k n to the power k by n. And finally, final prediction error RSS by n this is in fact a RSS by n e to the power 2 k by n RSS by n to the power k by n RSS by n into n plus k by n minus k. So this is how, we have already discussed; this is how this choice of lag length has to be decided, minimum the error then lag length has to consider. In the case of DG test, we have number of test like you know MAD, MSE mean square, root mean square error. Then you can say mean absolute deviation percentage error mean absolute deviation percentage error. So that means, mean absolute deviations, mean square errors root mean square errors, mean absolute percentage errors. Let me highlight what is all this things that means DG test is very important here. So in the case of DG test, we like to have mad first is mad mean absolute deviations.

(Refer Slide Time: 10:31)

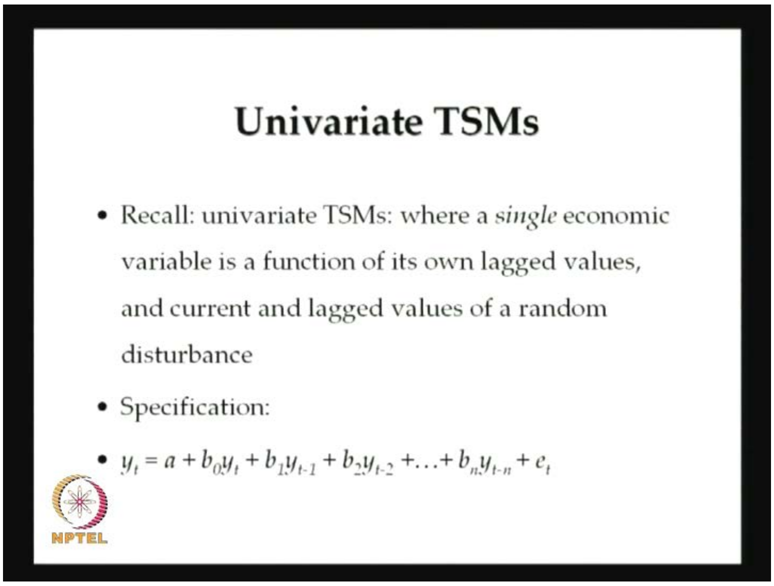


So this is nothing but summation, Y t minus Y t bars, in fact modulus has to be considers 1 by n t equal to 1 to n, so this is minimum absolute deviations. Then mean square errors mean square errors mean square error is nothing but summation Y t minus Y t bar whole square 1 by n, so mean square error. Then root mean square error root mean square errors, summation Y t minus Y t bar whole squares 1 by n then (( )) 0.5 square root then mean absolute percentage error mean absolute percentage error 1 by n summation Y t minus Y t hats divide by Y t.

So this is t equal to 1 to n. This is how the DG test has to be tested. So that means, it is altogether every times there is a residuals. So Y t minus Y t bar means, it is an error term only. So if we will take squares means, sum square etcetera, then we will get residuals value only. So that means, residual is most important factor which can be discussed in the case of time series modeling. So, it is not only variable is important, in fact cross sectionals getting a error component is very important, but in these case, it is more important like you know direct independent variables.

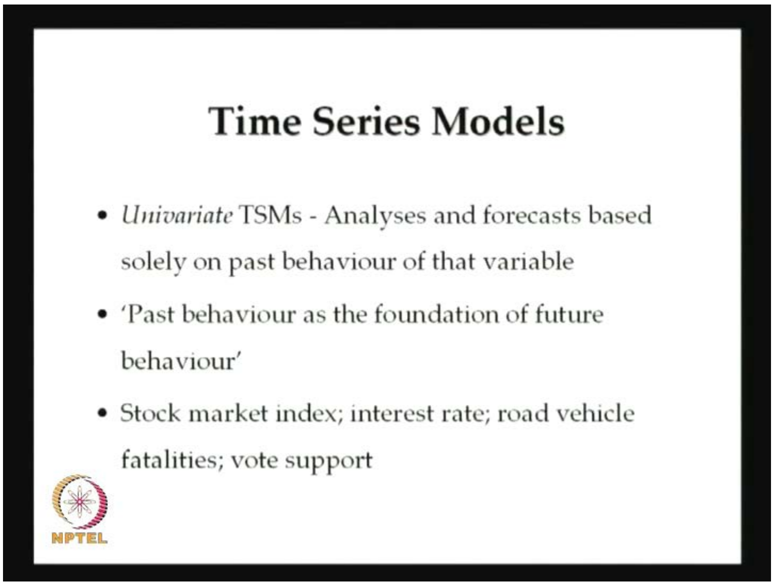
So here, even if one independent variable you are creating several independent variables and in the same times error terms will be you know, error term has to be introduced and that has to be tested very typically, so that the model can be considered as the best fitted model. Now, here once we will go for you know various aspects of time series modeling, then obviously there are two specific problems you will be generally face. One specific problem is, obviously by default there will be serious multicolinearity issue. And there is a serious you can say, autocorrelation issue. So that is why you must be very careful before you handling the time series modeling. So now, before I move to you know there is a concept of volatility modeling, let me highlight the concept of ARCH and GARCH models. Then means ARIMA models we will specifically highlight in a better way. Then we will move to this volatility modeling like you know, ARCH, GARCH, etcetera. So let me first highlight the structure of the particular structure here is what we have discuss till now, I will just summarize, then we will come down to this (( )).

(Refer Slide Time: 13:28)



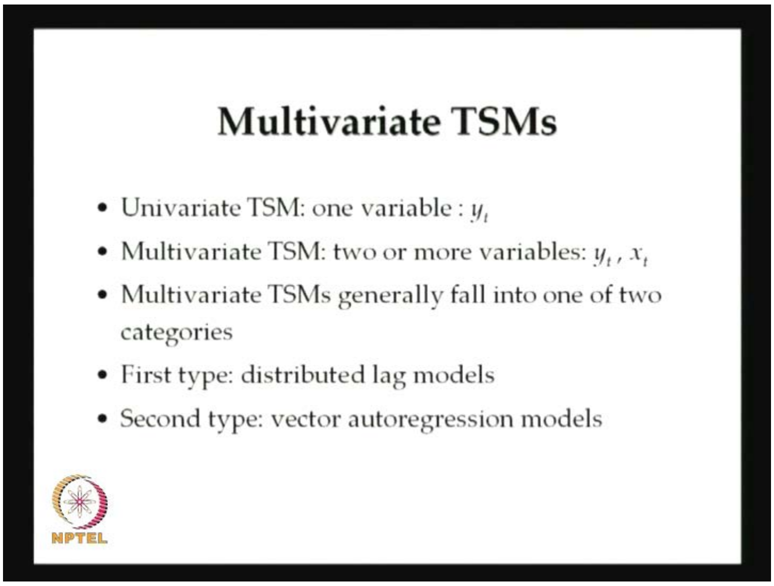
So this is univariate time series modeling, so that to y t as a function of y t y t minus 1 y t minus 2 y t minus n. So here the specific object is to estimate the parameters like a beta 0 beta one, a b 0, b 1, b 2, up to b n. Then of course, we need to have error terms e t.

(Refer Slide Time: 14:01)



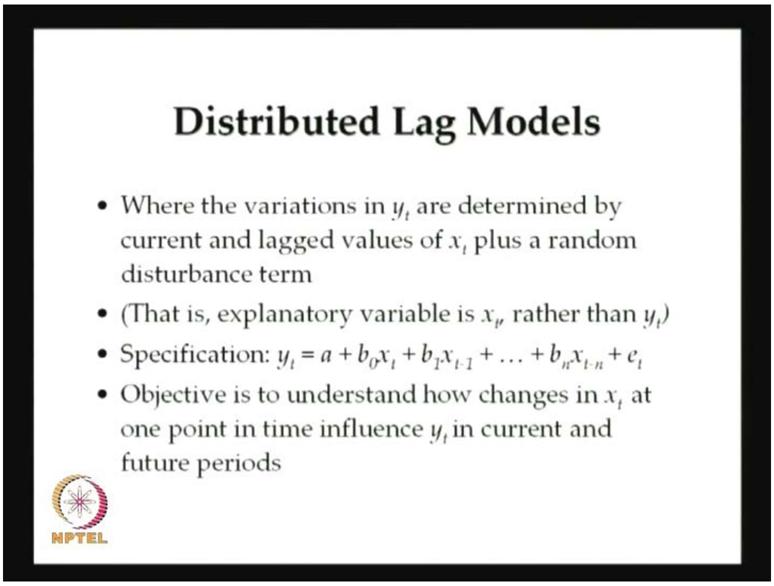
The moment you will get the error terms e t, then obviously you will go for you know, moving or schemes again. So till now, this is you know only autoregressive scheme. Then obviously, this is univariate setup, so which we already discuss.

(Refer Slide Time: 14:09)



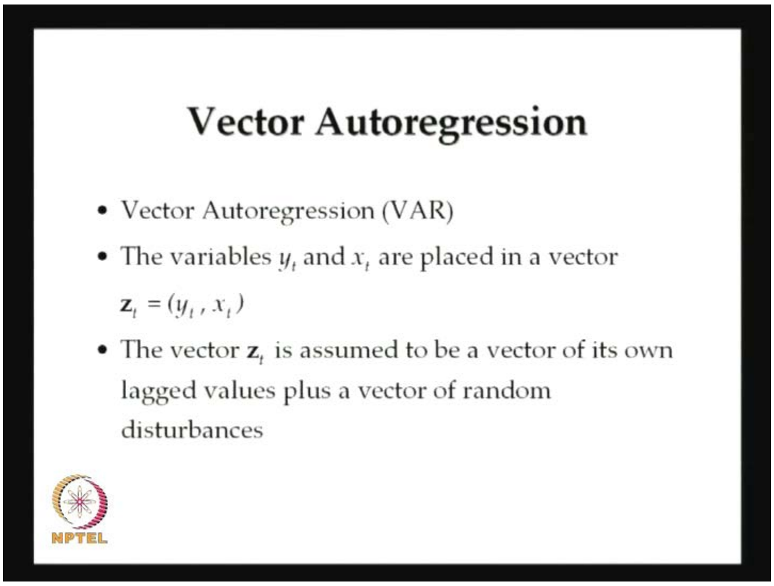
First behavior of that variables obviously, it is a non doubt about it. One variable y t, then multivariate models with respect to 2 n mode variables, let us say y t and x t. Multivariate TSM, and generally fall in to one of the two categories. First type is distributed lag models, second is vector autoregressive model, which we have already highlight in details.

(Refer Slide Time: 14:28)



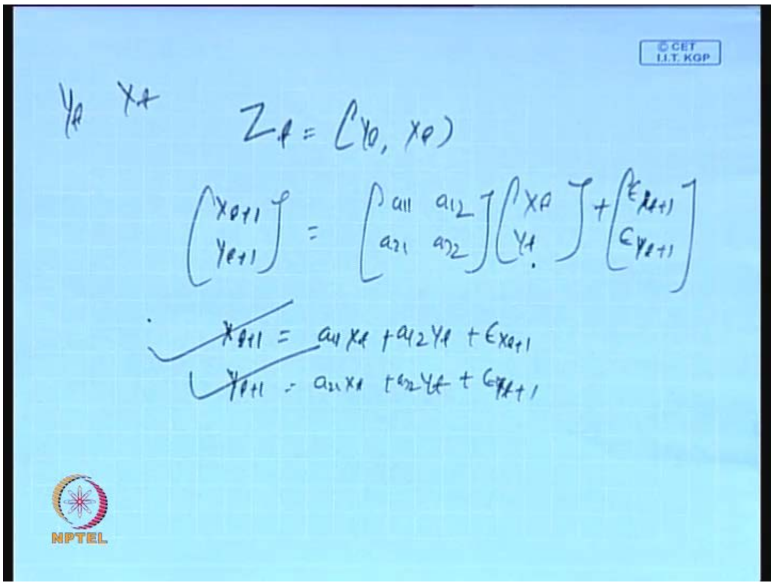
Now in the case of distributed lag models, I have already mentioned, so it is the endogenous variable as a function of exogenous variables. And it is lag, so obviously y t has a function of a beta 0 x t, beta 1 x t minus 1, beta t x t minus 2, like this beta 2 x 2 minus 2 then beta n x t minus n. So objective to understand, how changes in x t at one point of time influence, y t in current and its future periods. That is very important fact; we have to observe in the case of distributive lag models.

(Refer Slide Time: 15:04)



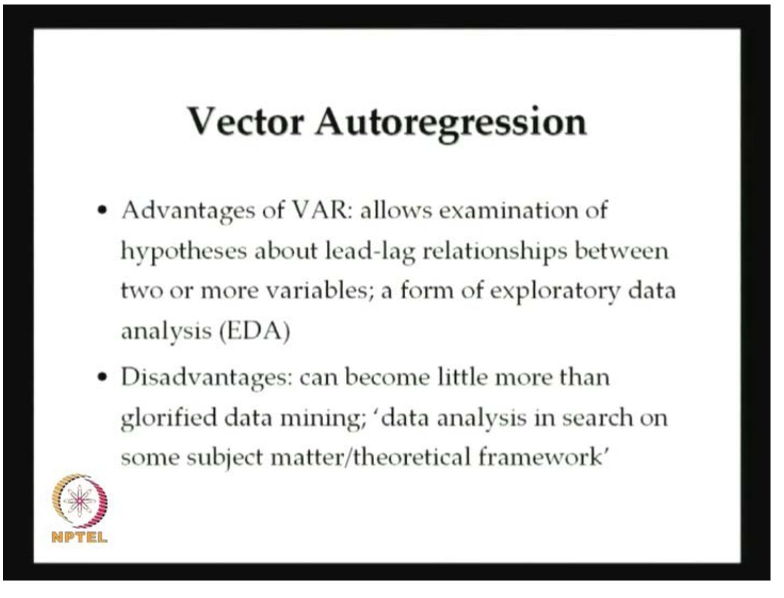
Then this is you know like similarly, you know like distributive lag models there is a autoregressive lag models where endogenous variable y t as a function of exogenous variables x t and its lag x t minus 1 x t minus n and its lag of endogenous variable like y t minus 1, y t minus 2, y t minus 3, up to y t minus n, so we will come down to that particular model again here. There is a VAR scheme here. VAR basically, you know z t equal to y t minus x t, so that means the VAR model will be little like this way.

(Refer Slide Time: 15:38)



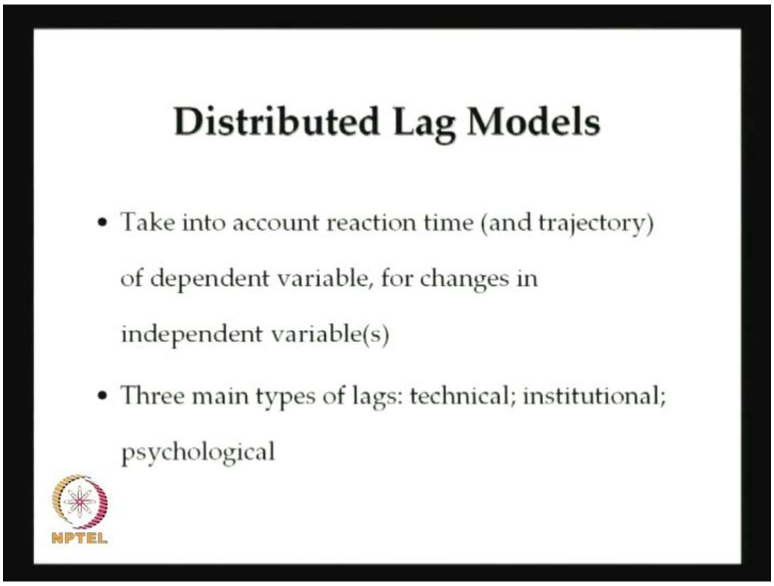
In the framework of y t and x t, so that means in the VAR setup, you must have two variables y t and x t. So this is generally written in various different ways, but may highlight, how we I will start a VAR model. Let us start with x t plus 1 then y t plus 1 then which is in the matrix format which is equal to let us say a 11, a 12, a 21, a 22 into x t y t plus summation epsilon t x t plus 1, then epsilon y t plus 1. So let us say it is time period t plus 1 is a current time periods. So now, x is a first values. If I will put y t, then obviously first value is y t minus 1. If I will put current t plus 1, then it is first value obviously a t. If I will put it in explicit format, then the equation can be written like this way x t plus 1 is equal to 11 x t plus a 12 y t plus epsilon x t plus 1. And then y t plus 1 is equal to a 21 x t then a 22 y t plus epsilon x y t plus 1, this is how two different models can be represent, this particular structure is called as a VAR model. So the vector is assumed to be a vector of its own lag values plus a vector random disturbance that is what the VAR scheme all about.

(Refer Slide Time: 17:19)



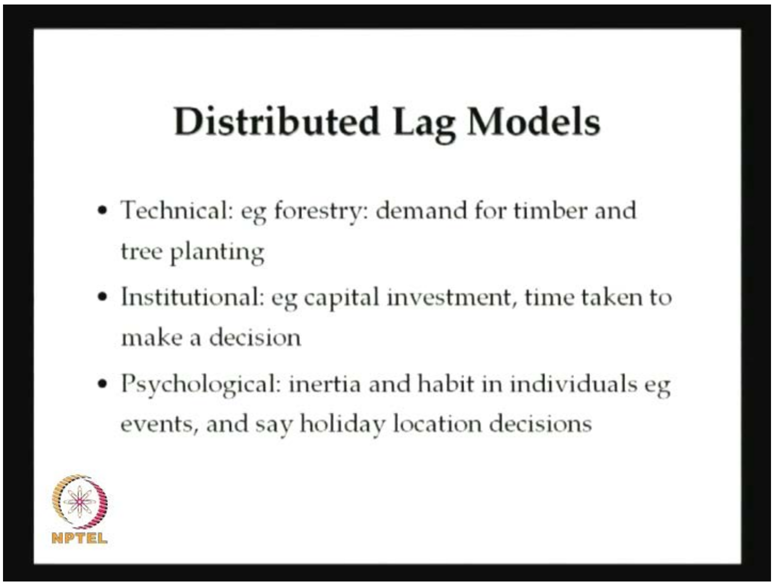
Advantage of VAR you know, specifically allows examination of hypothesis about lead lag relationship between two or more variables from exploratory data analysis EDA. Disadvantage is that can become little more than glorified data mining, data analysis in search on some subject matter theoretical framework. So it is a very much, you know means the proper structure is, how you will be go for the sampling, etcetera. So your sample size should be absolutely very high.

(Refer Slide Time: 17:53)



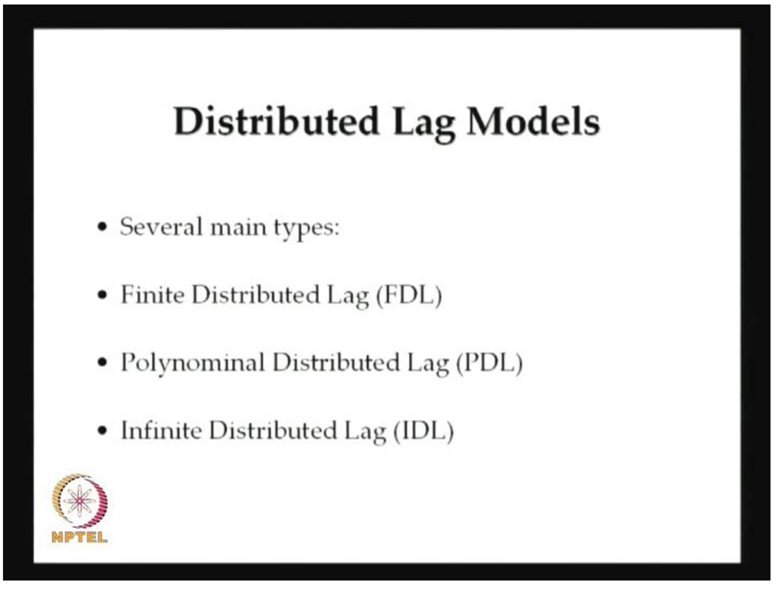
And this is the different structure of distributed lag modelings. Why there is distributed lag model? There are three different a specific regions, technical aspect institutional aspect and psychological aspect.

(Refer Slide Time: 17:56)



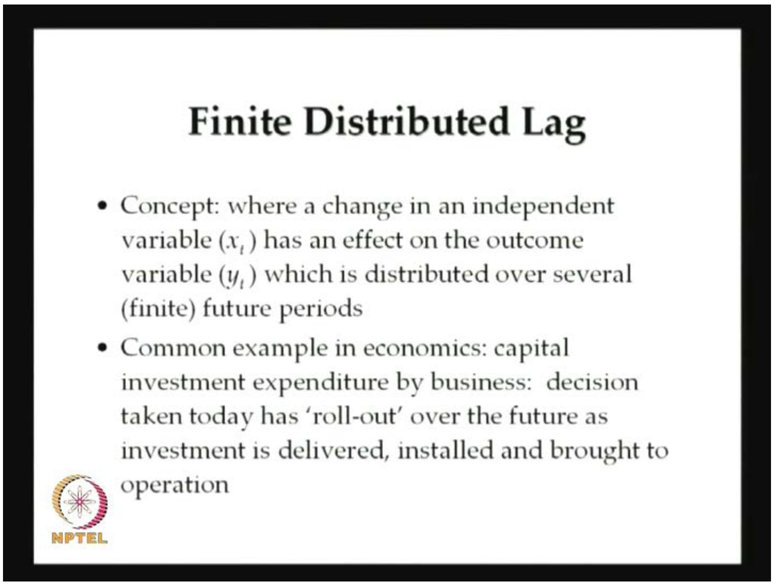
So I am not going in detail all these (( )), because, generally I have already mentioned in the time series setup. So we must be very careful, so the introduction of lag, because here one of the standard assumption is that the current value of a current variables depends upon its first end. If the first means, it is we are connected with its first end. So the thing is that or the issue is that so how many first observation it depends? So that is you know how many first periods time periods we need to incorporate, in that particular time periods, means current time periods that is very important. That means, if will you take a x t is a variable, then how quickly it has integrated with its first values like you know, whether it is x t minus 1 or x t minus 2 or x t minus k. Sometimes there are certain models, it start with integrating future values like x t plus 1, x t plus 2, x plus 2 up to x t plus k, so that means, it is x t is current then it is first integrations and it is in a future integration. Then you have to having the collective information, then we will go for further future forecasting.

(Refer Slide Time: 19:33)



So that is very important in the issue of time series modeling. As I have already mentioned, there are three different structures you will find in the distributive lag schemes. One is called as a finite distributive lag schemes, then polynomial distributive lag schemes, then infinite distributive lag schemes.

(Refer Slide Time: 19:41)

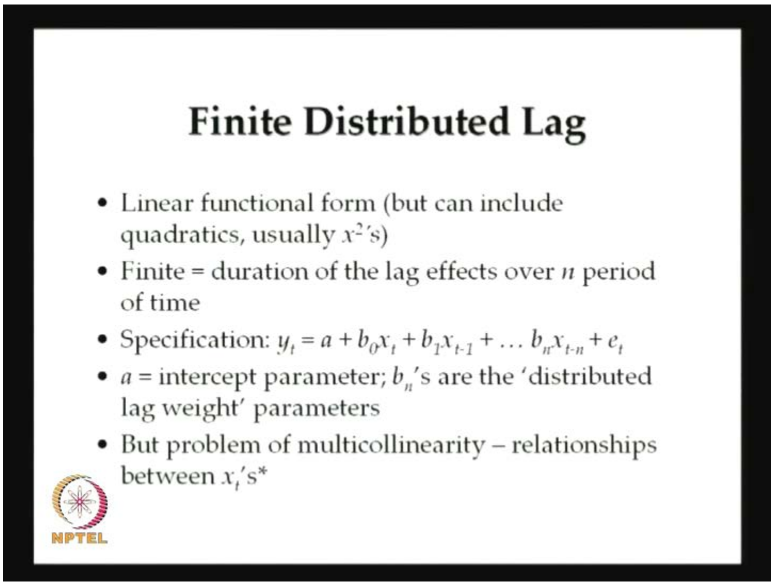


These are all we discuss, because the moments you will go for distributive lag models, then obviously you must have two different variables together. Because, altogether it is a multivariate time series modeling scheme. In the multivariate time series modeling scheme, at least you start with two variables, otherwise it cannot be possible. Then, if not two variables, then obviously we will come down to this univariate time series setup. Now in this particular setup, if there are two variables, then it is a time series multivariate time series modeling, that to distributive scheme. Under distributive scheme, we have different game different models altogether again. So, one of such model is equal to is finite distributive lag schemes.

So that means, once you have a variables y t as a function of x t x t minus 1 x t minus 2 up to x t minus k, so this is called as a finite lag scheme, that means, end we know. Sometimes, end we may not know that is called as infinite lag scheme, so that k may be finite k may means, k less than infinite k may be equal to infinite. So equal to infinite means, you are not sure about how much lag length will be considered. Let us say, I have 10000 sample points. We will get end number of you know, number of lag variables in the systems, but provided it should be optimum every time. So your moral improvement will be going on when we will add one after another variable in the systems.

It is just like you know, as I have mentioned it just like a stepwise regressions. You have to enter one after another lag variable in the systems. Then every times you have to check the overall fitness of the models. That means the parameters which is already theirs so that should not be affected and overall fitness of the model that should not be affected. But the moment you will be add another variable, then keeping all these variables specifications, entry variable new entry variable should be significant and in the same times the overalls specification and moral fitness should be same conditions or you can say it should be improve in fact. If it is same, it is if it is declined then you have to stop theirs.

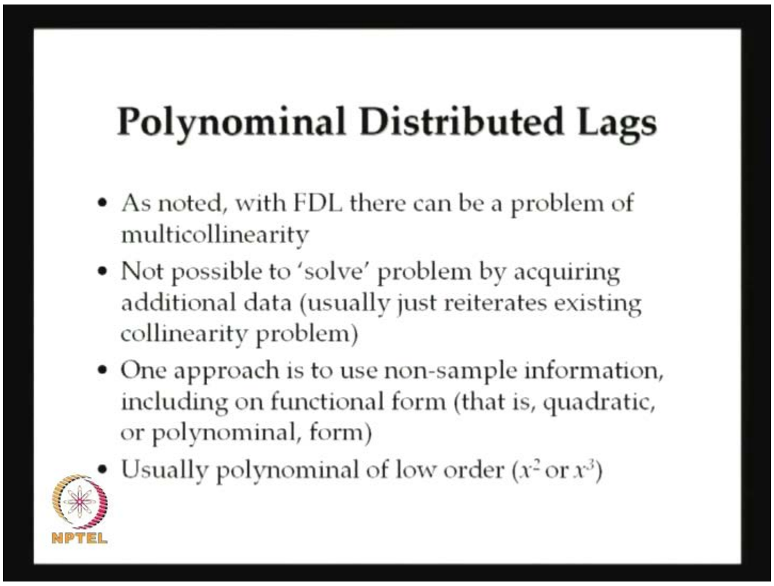
(Refer Slide Time: 22:04)



So that is how you have to go for you have to go for introducing lag in that particular system. This is finite lag schemes; the model of the finite lag scheme is like this way that means its limit is already there. As I have already mentions, so there are two major problems we will face here in this lag models. One is called as a multicolinearity problems, another is autocorrelation problem. So mostly, that means, in when we will go for time series modeling and its introduction of lag, then these two problems are obvious multicollinearity problem and autocorrelation problem is very obvious.

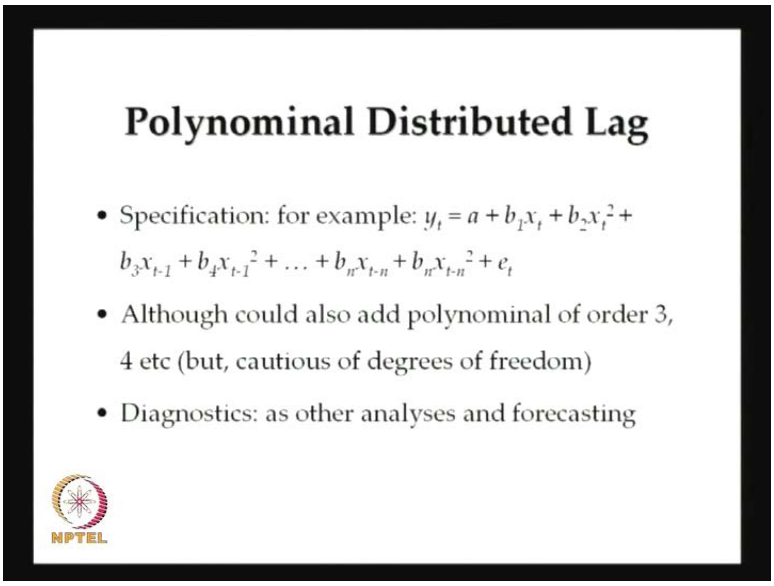
So obviously, you have to it is mandatory to check the multicolinearity issue and it is mandatory to check the autocorrelation issue. So if the autocorrelation is serious problem, then you have to modify. Because there are, in fact I have already mentioned you know these particular structures, which we have a started with the initial beginning setup the detections. Detection means you have to detect whether the model is in accurate form. That means, the variables which you are in a huge in that process may not be appropriate and for that you need to have a transformation, appropriate transformations or appropriate setup, appropriate mathematical form, So that you will get the model in a better (( )).

(Refer Slide Time: 23:19)



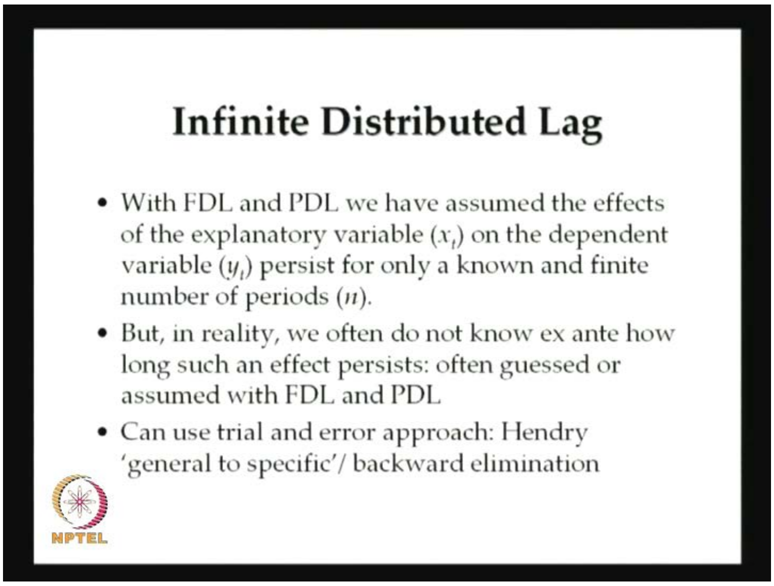
So this is you know mod specification about finite lag distributed model. Then, in fact come down to polynomial distributed lag models. In fact polynomial means, it is the non-linear formats.

(Refer Slide Time: 23:26)



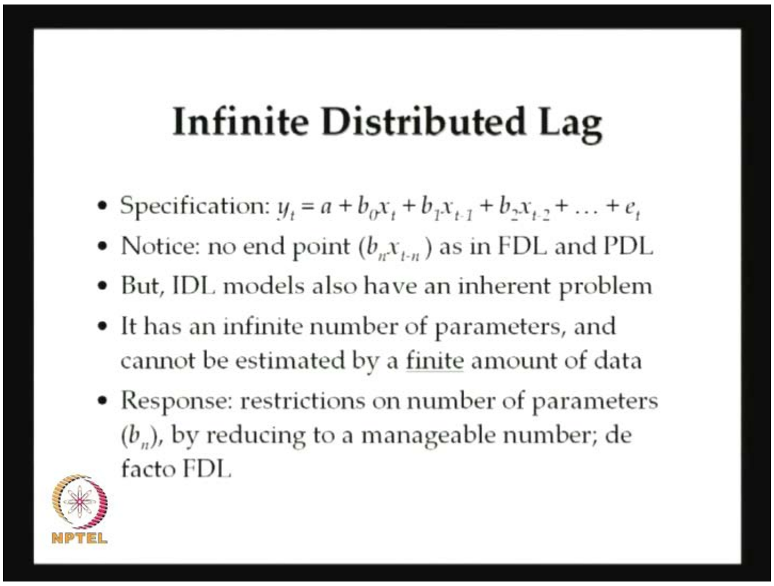
Then both non-linear not with respect to coefficients, but with respect to the variables only, which we have already discussed in the case of cross sectional modeling. So variable, you can say particularly we have discussed this issue in the case of this particular qualitative response regression modelling that to you know probability models. So, that means logic and profit model case, we have to discuss all these details. The thing is that here, polynomial distributed lag, here parameters are specifically constant. But in the variables, we are putting in a in a different sets like you know beta 1 x t beta 2 x t squares beta 3 x t minus 1 beta 4 x t minus 1 square like this way.

(Refer Slide Time: 24:24)



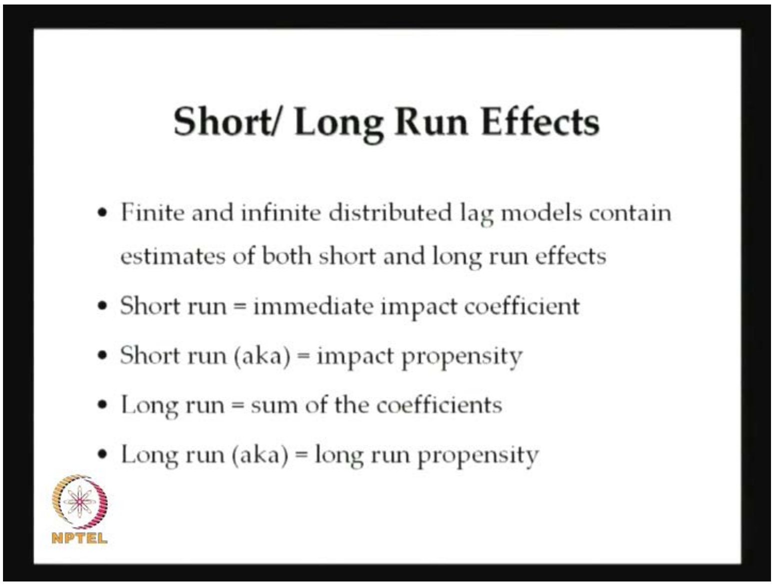
So this is in this different schemes altogether, so you have to be very careful about his one. Then infinite lag distributed lag models, so here is, here the limit is not you know, fixed here. That means we do not know what is n here.

(Refer Slide Time: 24:28)



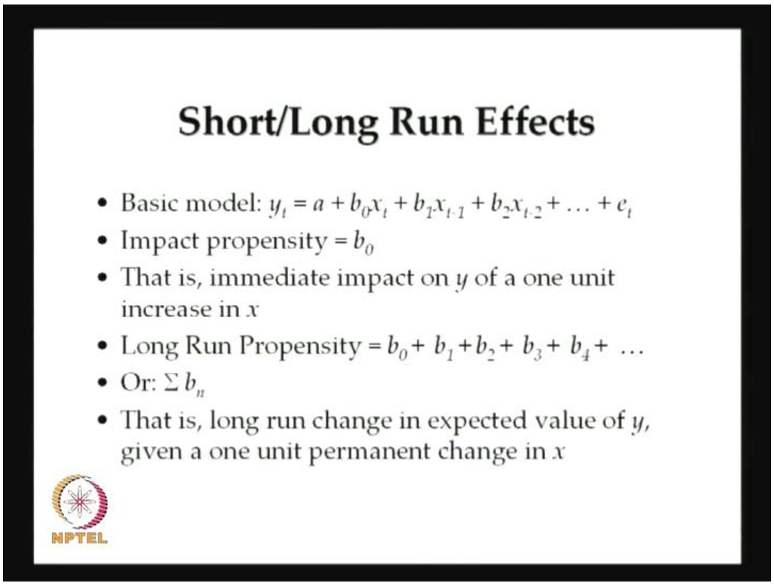
So n may be means, we are considering that it is an infinite in natures. But it is very difficult to handle this type of problem. So this is short run effect and long run effect. Some so basically, you see, when there is a time series issue, then obviously there are two specific objectives. What is this short run impact? And what is the long run impact?

(Refer Slide Time: 24:44)



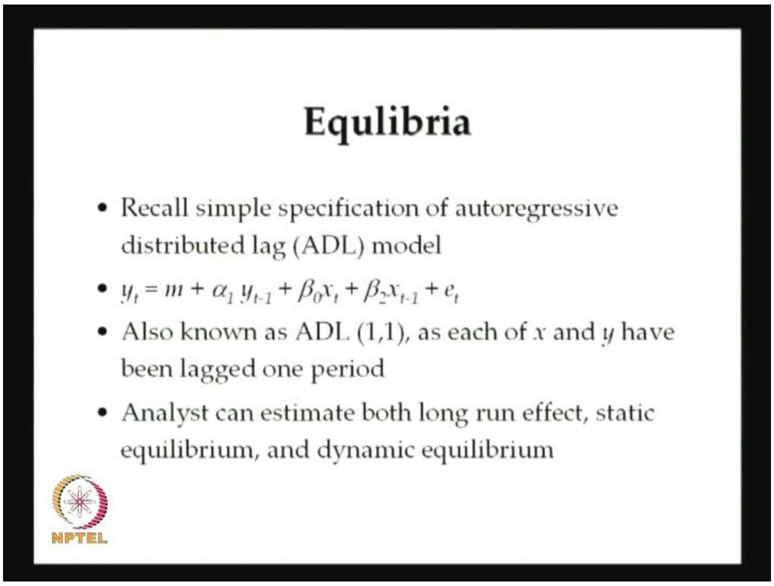
So, it the structure is completely different with respect to short run impact and with respect to long run impact. The more long run behavior the model is modeling structure is completely different and short run the modeling structure is completely different. Sometimes you know, there is a model called as a vector error correction model which it integrates a short run impact along with long run impact. So, we will discuss in details.

(Refer Slide Time: 25:26)



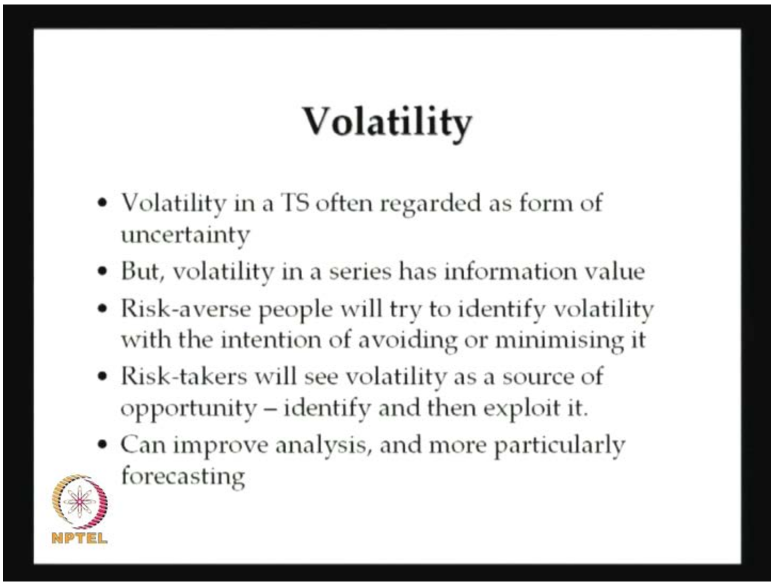
Let us, this is how you know short run impact, long run impact here. We specifically in this particular structures, we can study with respect to the parameter value. So the impact propensity that is short run impact which is called as a beta 0. Then long run propensity to you knows sum of the coefficients will equal as a long run impact of this particular systems.

(Refer Slide Time: 25:53)



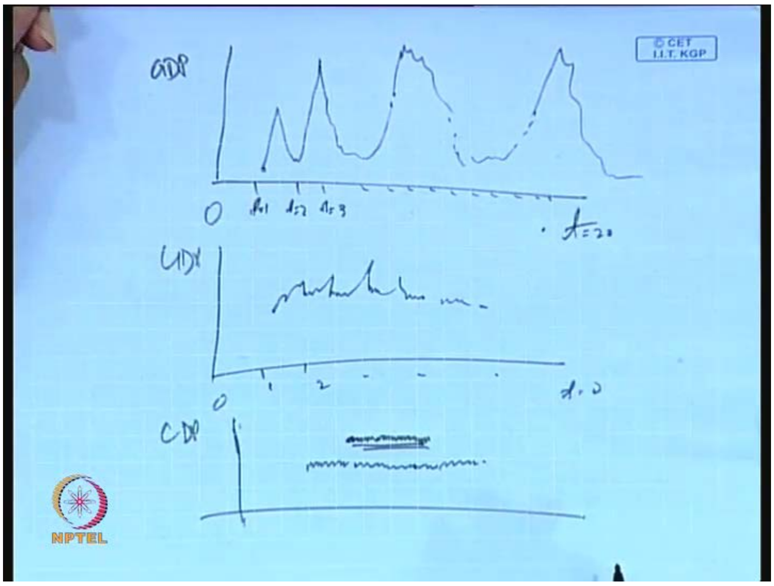
Equlibria that is this is another autoregressive distributed lag schemes. Now, you come down to volatility. So this is serious structures which we want to discuss now. You know volatility is one of the interesting problems which usually we observe in various issue.

(Refer Slide Time: 25:59)



By the time series issue modeling is concerned, volatility is must like this. So what is time series modeling altogether, so I have a time frame one side and another side the behavior of the variable. So the behavior variables will be like this.

(Refer Slide Time: 26:27)



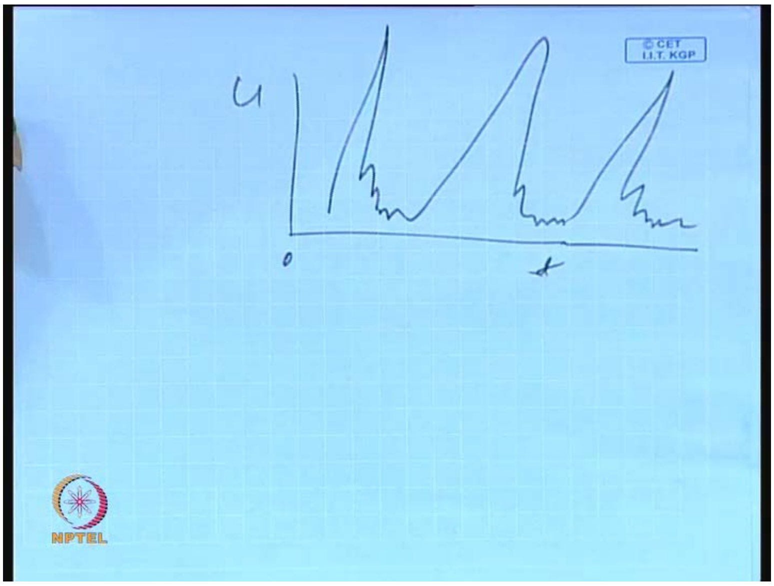
Let us say I will just plot a sum of the variable like this with respect to time. Let us say I will take a GDP structure GDP, so t equal to 1, t equal to 2, t equal to 3, so up to t equal to say 20, so 20 periods I have taken. But 20 periods is not reliable sample for time series modeling. You might at least have say 200 issue. Anyway, so for you know classroom just you for clarification I am taking 20, so that there may be lots of simplicity here. What I will do? I like to know the volatility issue. What is all about the volatility, that I have to highlight here in this particular setup. So, what I will do? I will just plot the variables. Let us the plot variables that come like this way. This is a difference of structural variables you will find in this setup. This is how one of the structures, find you will find the trained like this way.

I will put another structures here, is equal to t to 20, 1, 2, 3, up to like this way continuous. Then, I will get a structure like this. This is another type of structures, lets I will take another schemes. So, this is another scheme, here GDP this sides and GDP this sides, I will just go through like this. This is how the scheme will be altogether. So in this out of all this three models, you see I am just giving a little bit you know, in a decreasing trend. So this particular models, if will you make a look here, then it is very volatile in nature, this is less volatile, this is you know, very very minimum volatility.

In fact, the thing is that what is first thing is what is volatility? Why is it important? And whom it is important? These three questions is a means, these three questions are very much important, before we going to analyze the volatility modeling. So first of all, you know take a case of investment decisions. When we will go for any investment, particular this volatility modeling is a most of the cases, we handle in the case of times financial time series, financial modeling, etcetera. So when we will go for investment decisions, then obviously there are two items you have to observe. First is how much return you will have in this investment and how much risk you to bear in the investment. So that means, one side it will return and one side it will risk. So there is obvious game between risk and returns. So the structure is that higher the risk and higher the returns and lower the risk lower the returns, but optimal a person who means an investor, who want always you know with a high return with a minimum risk. So, high return with minimum risk, it is very very difficult to get sometimes.

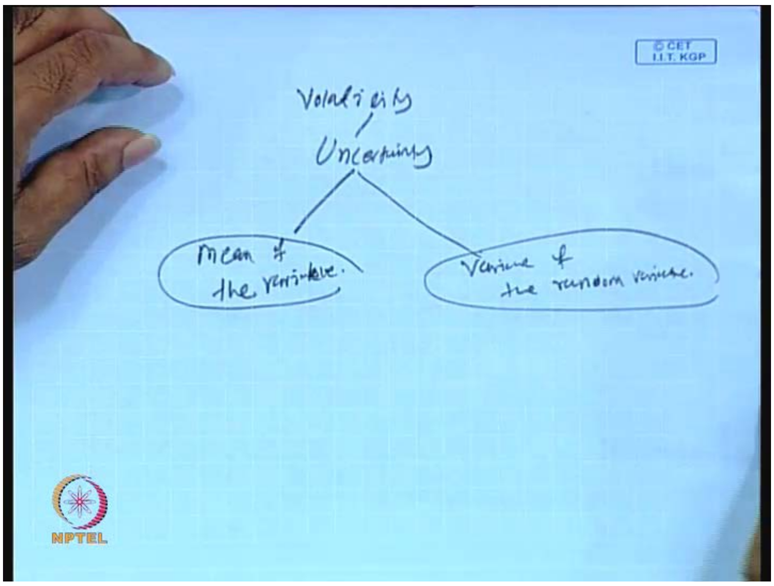
So what you have to do? You will go for maximum return with optimum risk, so that is the structure you have to set. So, there are different models altogethers. And one of the ways we will observe this is nothing but this volatility modeling, so you have to see how much volatility is there. If there is less volatility, then obviously it risk is very less and obviously, in the same times you have to check, how much return is the in this particular investment. And if the volatility is very high, then obviously, it is very risk that means, it may give you high returns or it may not give you high returns. So, there is you know totality is uncertainty issue. So that means, there are two different models we will find. One is called as it is a certainty model with certainty and model with uncertainty. Model with certainty means, it is a stable models, stability situation and uncertain means, it is not instability. In fact, it is not a stable, it is very in the instability case, it is very difficult to predict something or it is very difficult to forecast something. But in the case of stable, you stability you can forecast something very accurately, so it is not a problem. But you know, in the case of volatility it is more interesting, so the system should not be too stable and it should not be too instable. So, it should be in between these two. So these are the two extreme, if it is too stable. For instance, I will put like this way I will put like this way only, so that means there is no such change at all’s there is no such change at all’s. Then in that case, it is very difficult to go for investment. If there is no such volatility, nobody will take a risk. If you do not take risk, then obviously return cannot be obtained easily.

(Refer Slide Time: 31:52)



What you have to do? We need a very serious issue. For instance, this is not a problem, but it is more interesting. This is somewhat problem, but it is also interesting. But I will give you another type of variations you may face, this is 0 t and this is G, let us this type of structure; it is very volatile in nature. In that case, it is very difficult to handle or predict in this particular angles. So in that case, either you can handle with volatility modelling or what you have to do, you have to go for different transformation technique. If will you go for different transformation technique, then it will give you beautiful results and you know forecasting issue.

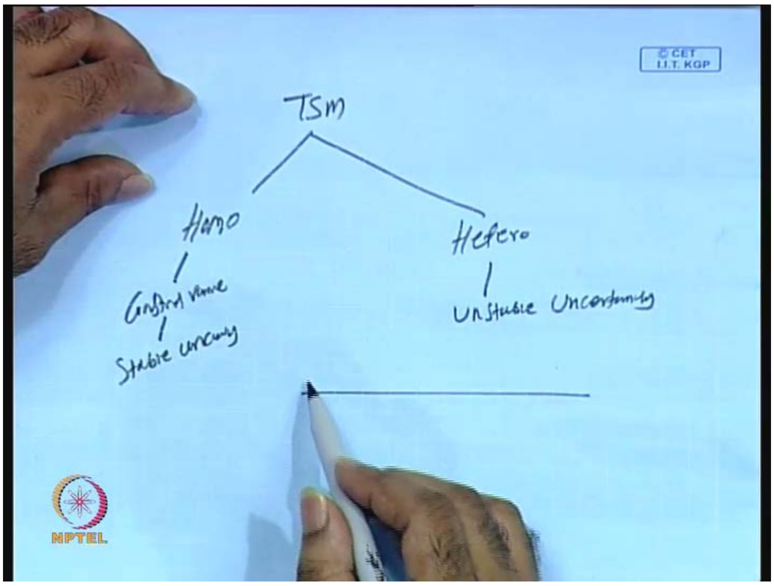
(Refer Slide Time: 32:29)



So let me highlight, what is exactly this volatility? Volatility basically means volatility basically is uncertainty. You know uncertainty, there are two forms. So, mean of the random variance mean of the random variance and variance of the variance of the random variables mean of the mean of the variables and variance of the random variables mean of the random variables, this is variable in fact mean of the random variable and variance of the random variables. So we have to see what is the mean? And what is the variance? There is a concept called as a stationary issue, we will discuss the stationerity issue in the later stage. Stationerity means it will give you the stability issue. And once means, when you will go for modelling, then one of the interesting requirement is that variables should be stationery in natures. So what is mean by stationery? A variable will be called as stationery, if its mean and variance are almost all constant over the time frame.

If it is you know not constant, then there is a question of volatility and that variables cannot be used for, you can say or it cannot be directly used for any forecasting and policy use. You have to first bring here to stationerity, to what level it is stationery, then that level has to be used for you can say preparing a good models and that to be a that can be considered as the best model and can be used for forecasting and policy use.

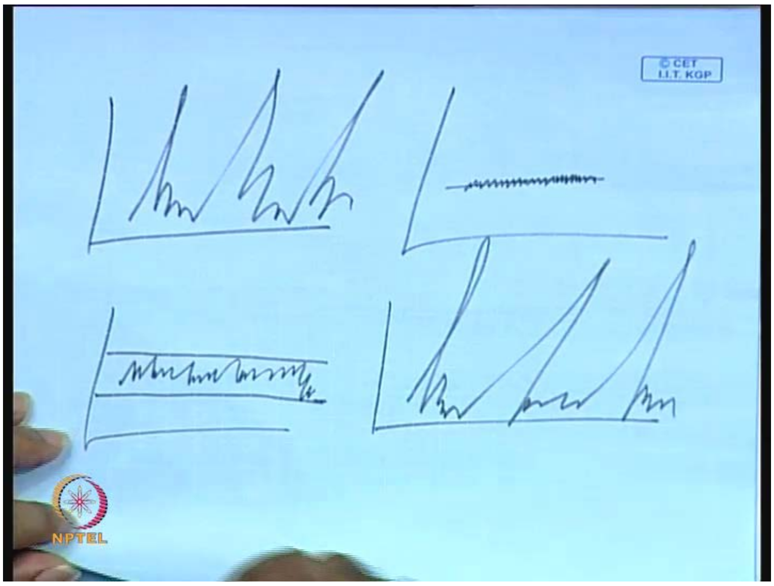
(Refer Slide Time: 34:20)



Anyway, so far the time series modelling is concerned, so far as time series modelling is concerned there is two issues. One is called as a homoscedasticity and another is called as a heteroscedasticity, which we have discussed. There is chapter, we have already discussed in the heteroscedasticity issue. So heteroscedasticity means, the error variance may not be constant. Over the means obviously error variance, if error variance is constant. Then it is homoscedasticity issue, if error variance is not constant then it is called as a heteroscedasticity issue. I have already mentioned why there is a heteroscedasticity always in the problem. Like you know, there is concept of error learning, then you know outliers problems, different mathematical form of the models or you can say wrong variable enter to the models or relevant variables not incorporate in the models. So these are the things, why there is system will not give you homogenous rather than heterogeneous. If it is you know, you getting heterogeneous issue then obviously it are very problem.

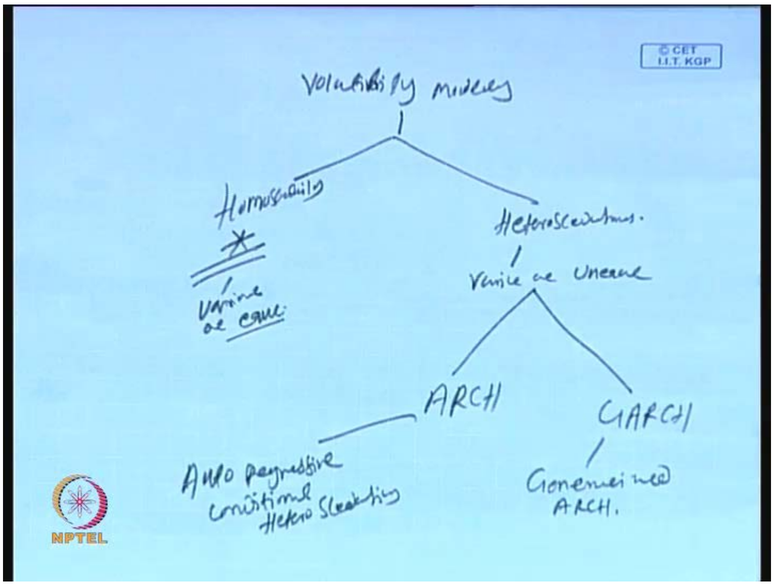
So in the case of homoscedasticity, it is constant variance it is constant variance over the time or over the samples, then you know it will give you stable uncertainty it will give you stable uncertainty. But heteroscedasticity, it will give you unstable uncertainty unstable uncertainty. In fact, certainty term is inside, in fact it is a risk term. Sometimes you know, I am using the term stable uncertainty and unstable uncertainty, for instance like this.

(Refer Slide Time: 35:59)



As I have already mentioned, as I have mentioned already, this is two extreme and another is a too much, you know too much rigid, this is another setup. Now, what is stable uncertainty means, this is not stable uncertainty, there is no question of uncertainty here. So let me, I will put like this way, you know unstable uncertainty unstable uncertainty. Stable uncertainty means, I will put like this way. That means the uncertainty can be under a certain range only uncertainty can be under certain range. So that is how it is called as a stable uncertainty. But unstable uncertainty means, there is a no such structure like this, it will become like this way. So you are not very sure, it is has huge ups and huge downs. So in that case, it is very difficult to predict altogether. But if this is the stable certainty, then it is very easy to handle this problem. Now, there are two different you can say, problems you have to handle through you know volatility modelling.

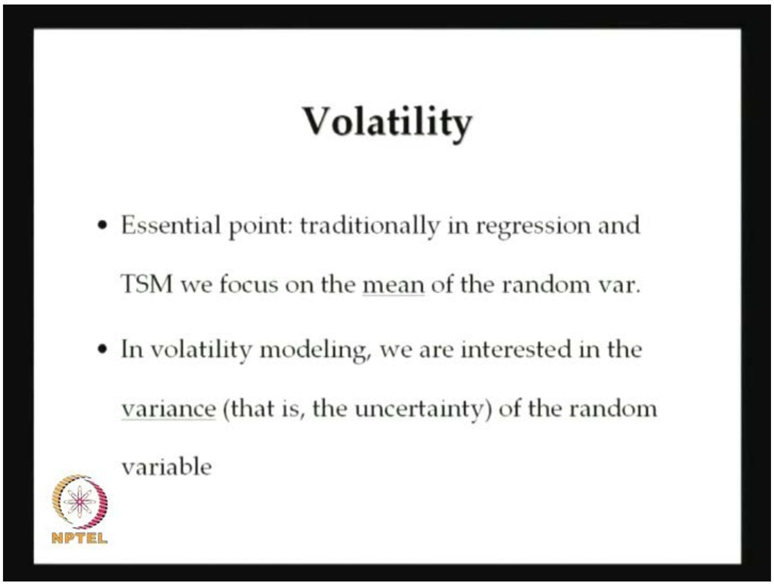
(Refer Slide Time: 37:18)



So volatility modelling, basically volatility modelling volatility modelling, basically deals with two aspects, homoscedasticity aspect, homoscedasticity and this is not our job right now, so there is another is called as a heteroscedasticity. So that means, variance are unequal, here variance are equal. Variances are equals means variance are equal, if variance are equal, then obviously your game is over. So that means it is already in the positive side nothing to do extra, so just you have to continue. But if you know variances are unequal, then it will give you additional problem. So you have to solve, again that additional problem before you use that model for forecasting.

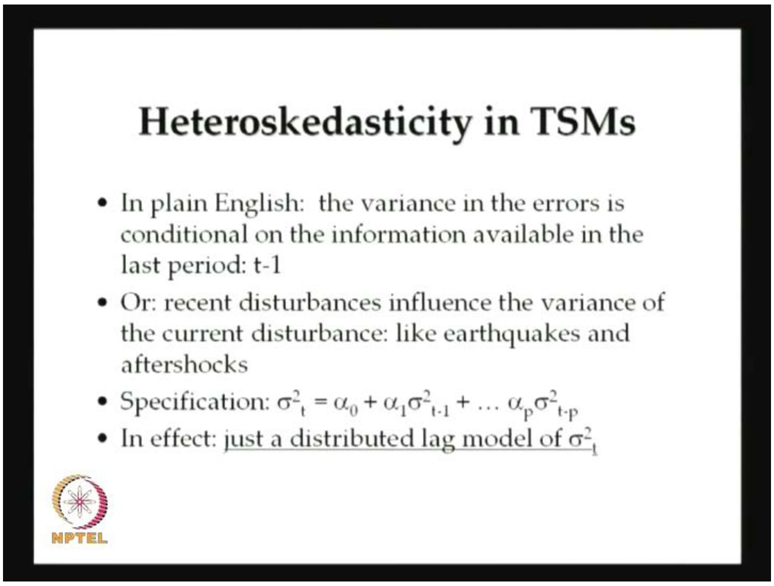
So now, once variances are equal, then we have different models. There are two, there are various forms of the models are there with respect to heteroscedasticity issue. heteroscedasticity issue means, error variance are unequal over the different time frame, because of various reasons, as I have already pointed out, outliers and you know inclusion of the unnecessary variable, exclusion of necessary variables, different mathematical form of the model, imperfection of error term, etcetera. So many ways it may be they are. So basically, error variances are not equal over the time frame. So you have two different models, one is called as ARCH model another is called as GARCH models. This is called as an autoregressive conditional heteroscedasticity autoregressive autoregressive conditional heteroscedasticity and this is called as a generalized autoregressive conditional heteroscedasticity. I will come down to here is right now, now volatility is a time series. Times series often regarded as a form of uncertainty that I have already mentioned.

(Refer Slide Time: 39:53)



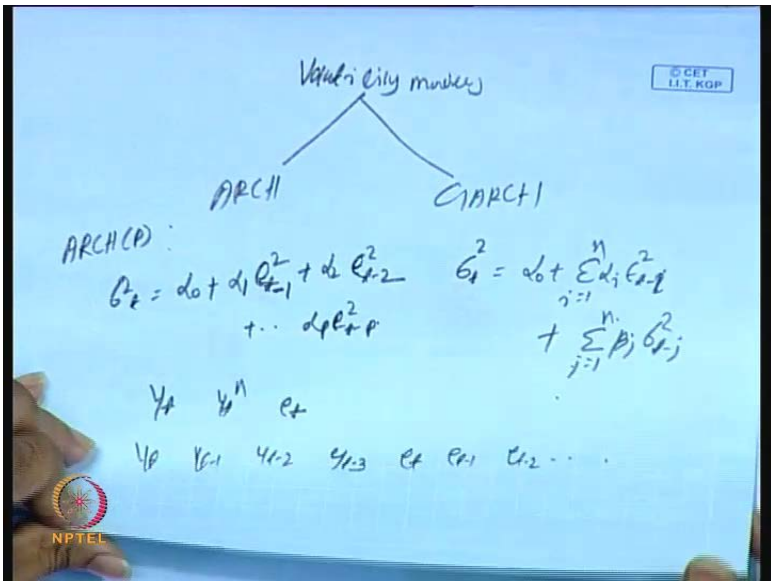
Generally, risk taking peoples will be very means they need all these issues before you know, explaining their status or investment plan. So they need volatility issue more specifically.

(Refer Slide Time: 39:59)



These are the basic features of volatility. I am just skipping all these things then I will come down to particular idea behind this volatility. This is the general specifications. So far as a heteroscedasticity problem is concerned, we usually handle two different models one is called as ARCH model and GARCH models.

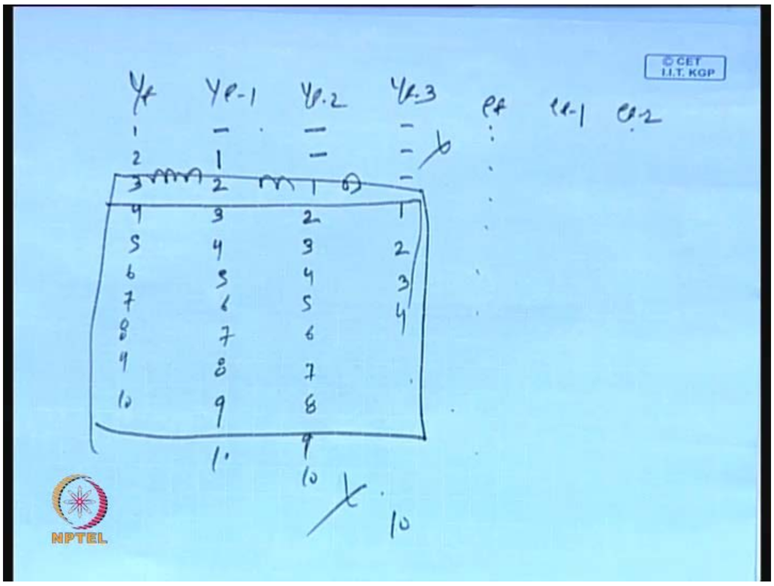
(Refer Slide Time: 40:16)



So what is an ARCH model? You see here is. So volatility modelling, basically deals with two aspects ARCH models and GARCH models. ARCH models, ARCH of order p, I will call it here is sigma square t is equal to alpha 0 plus alpha 1 epsilon square t minus 1 alpha 2 epsilon square t minus 2, it will continue alpha p epsilon square t minus p. So this is general form of the, you know autoregressive conditional heteroscedasticity. And in that case GARCH models, sigma square t is equal to alpha zero plus summation alpha i epsilon square t minus i i equal to 1 to n plus beta j sigma square t minus j j equal to 1 to n, this is two different models altogether. So what is the procedural measures here how you have to go for that one's. What you have to do? You start with two variables first, let us say y t as a function of y t minus 1, y t minus 2, like this way.

So then, you get the estimated coefficients. With the help of estimated coefficient, you have the estimated model and y minus y hat that is estimated model will get the error term. Now the moment, you will get like this. So you will find like this, y t then y t hats so then, e t then this is epsilon t, its better you put e square t minus 1, e square t minus 2, e square t minus p, so this is better error terms. It is there, the way we are creating y t, y t minus 1, y t minus 2, y t minus 3, like this way. Then we will create e t, then e t minus 1, e t minus 2, like this way. You know how the setup will be coming; generally the setup will be like this. This you know, y t y t is like this then you know y t minus 1 like this way.

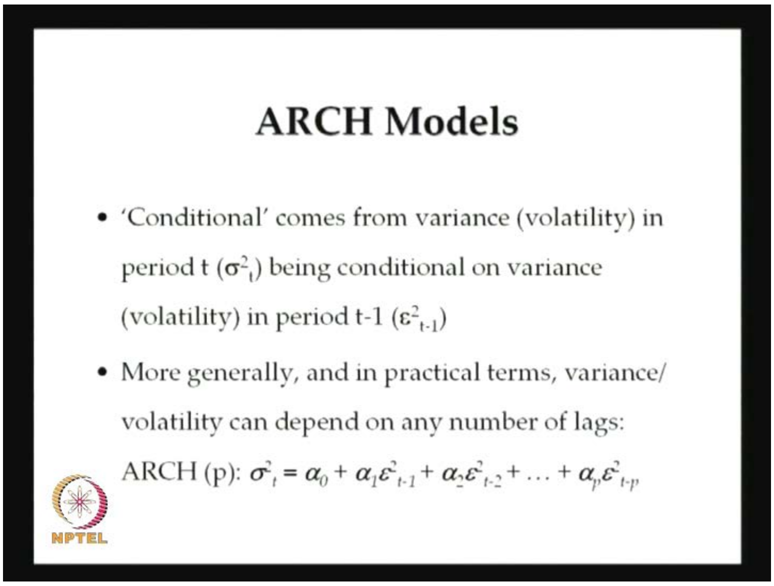
(Refer Slide Time: 42:31)



So this is 1, 2, 3, 4, 5, 6, 7, 8, 9, and 10. Then y t minus 1 will start from here, so 1, 2, 3, 4, 5, 6, 7, 8, 9. Then, y t minus 2, so the scheme will be like this, then it will start here is 1, 2, 3, 4, 5, 6, 7, 8 then 9, 10, so always there will be ten here. Then, y t minus 3 so 1, 2, 3, then it will start 1, 2, 3, 4, 5, it will be again ten here. So this is how, the sequence will grow. So that means you see here. One of the conditions of the modelling is that yours sample observation for all the variables should be uniform. So as a result, first, if will you like to considers, you know up to y t 3 minus 1, then you have to take these much of clusters you have to take these much of cluster, that means, this is rejected and this is rejected. So that means altogether, you are losing here three different samples; we are losing three different samples. This is three means; you will be start with actually four you will start with actually four. This is not correct, so we are losing three samples mean, so 1, 2, 3. This is the proper sequence, so three samples you are losing.

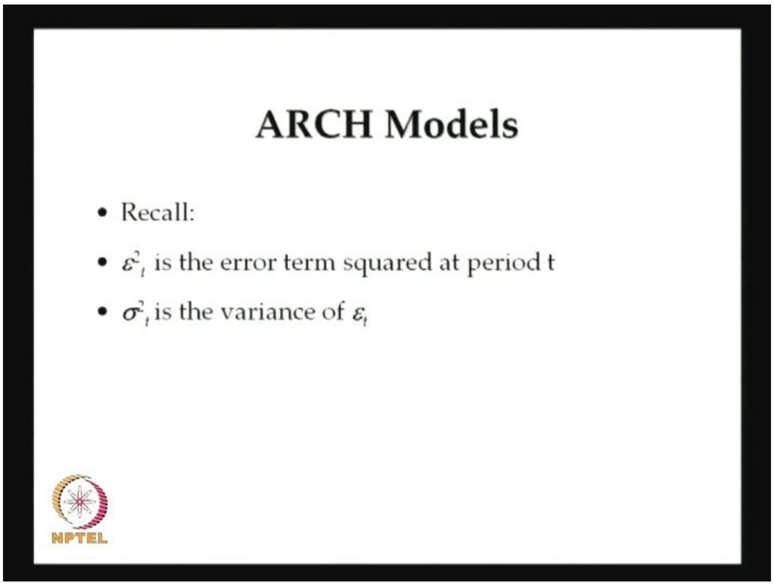
So now, if will we add further, then you will lose another sample; as a result, it will be give you degrees of freedom problem, so you must be very careful how you have to handles. And similarly, if will you go for error term again, so e t, it will be first you calculate e t with respect to y t, it can be calculated. Then you will go to e t minus 1, and then you will go to e t minus 2, so like y t minus 1, y t minus 2, y 3 minus 1 so similarly, we will go to get e t minus 1, e t minus 2, etcetera. So now, what is the model importance is that, you have to find out the variance of error terms; sigma square t, which is function of its error component, square of the error component. If means, here the objective is all this coefficients should be all this coefficients should be statistically significant. So that means they are different from 0. If they are not different from the 0, then that that means there is no volatility issue.

(Refer Slide Time: 45:03)



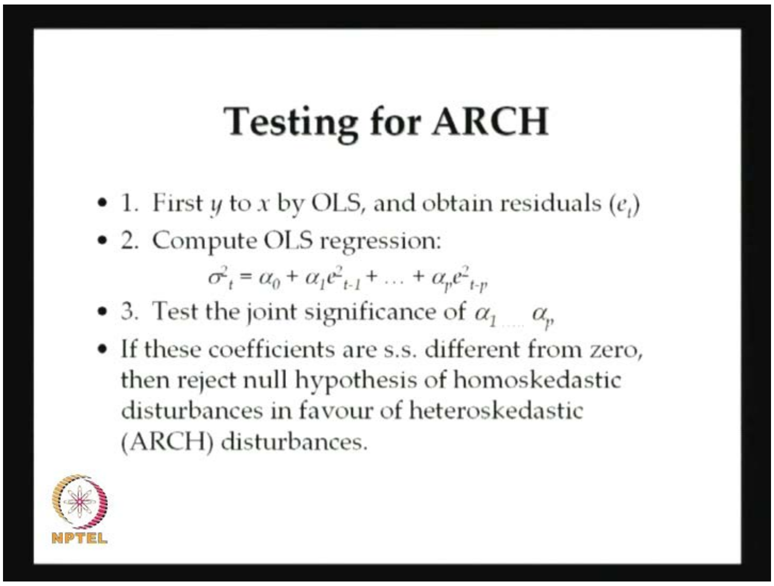
This is general generalized format of you know, volatility model that too arch models. More generalized and more practical way, we have to establish. We must get equal to alpha 0 alpha square e square epsilon square t minus 1 alpha 2 epsilon square t minus 2, then alpha p epsilon square t minus p, so this is the generalized formula.

(Refer Slide Time: 45:26)



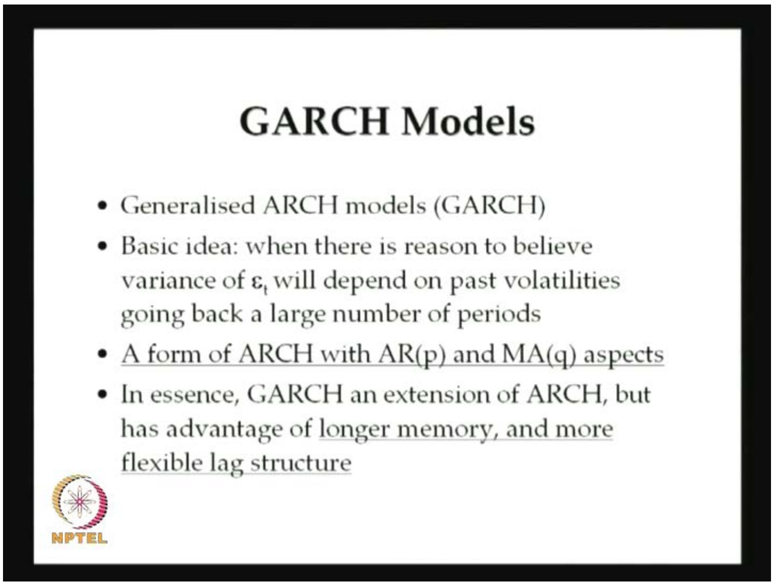
So what we will do? You have to find out the, means here epsilon square t means its error terms here at time period t. Similarly, epsilon square t minus 1 means the error term square at time period t minus 1. So similarly, sigma square t is the variance of error terms that is epsilon t. So we have to regress the variance with the square of the error terms, then we have to observe the volatility issue. If the variables are totally significant, then obviously there is an ARCH effect. If not then you know there is no such ARCH effect. So accordingly, you have to be very careful about that issue.

(Refer Slide Time: 46:03)



So that means ultimately, we like to establish whether there is a volatility modelling, means in that particular problem or in that particular for what we have discussed in the investment decision issue. So we like to know, what is the level of volatility in that particular problem or particular investments plan. If their issue is volatility, then accordingly the investor will take a decision. If it is low volatility, he will take, of course, accordingly to as per his need and choice. So that is why observing volatility is very important and is very, you know typical issue in the time series modelling. It will give you long term trend and lot of setups how you have to build the entire setups.

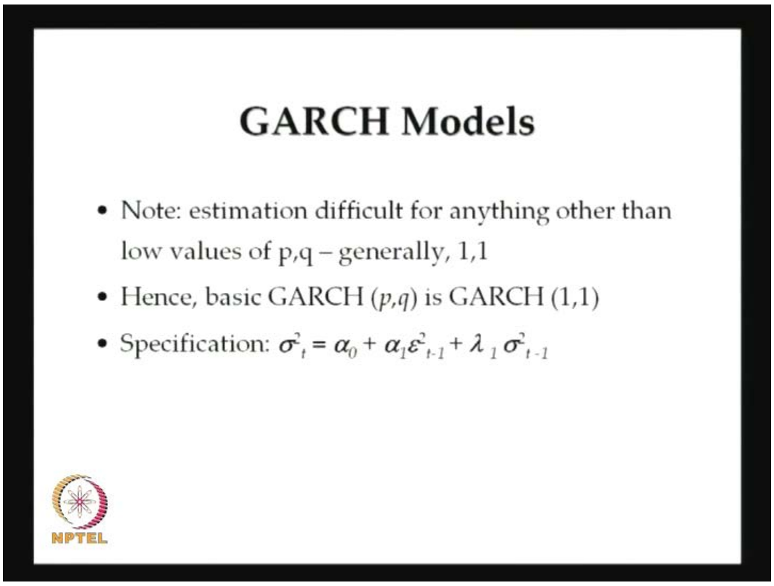
(Refer Slide Time: 46:52)



This is testing of ARCH models, so how you have to go for all this testing anyway. Then it is; now it is come for advanced portion of GARCH model. Sometimes you know, GARCH model is generalized more advanced versions or broader concept than the ARCH model autoregressive conditional heteroscedasticity model. In the case of GARCH model, you mean here the objective is you have to go for ARCH effect and GARCH effect. Sometimes, what happens? ARCH effect may be theirs, GARCH effect may not be theirs. But once ARCH effects there, then you have to go for GARCH effect; if ARCH effect is not there, then no point to go for ARCH effect.

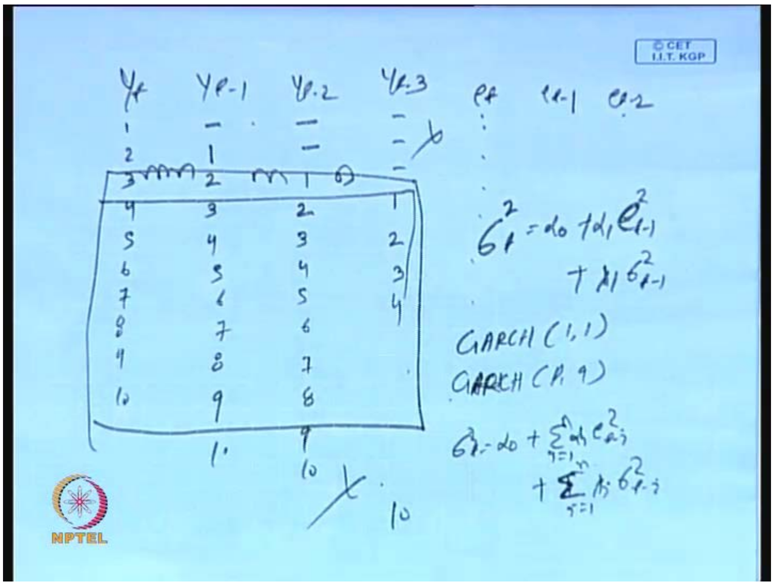
So GARCH effect is more generalized concept, if there is GARCH effect then it is a more general way you have to explain the volatility issue. It is you know, just like you know, we have discuss ARMA models where we will start with the autoregressive schemes, then moving average scheme, then you know autoregressive with moving average scheme. So here also same things, we have ARCH model and GARCH model. In the case of ARCH model, it is like you know moving average schemes. Then in the case of GARCH, it is the combination of autoregressive schemes as well as moving average schemes that is why we can explains here is. So the general framework of GARCH model is like this way.

(Refer Slide Time: 48:03)



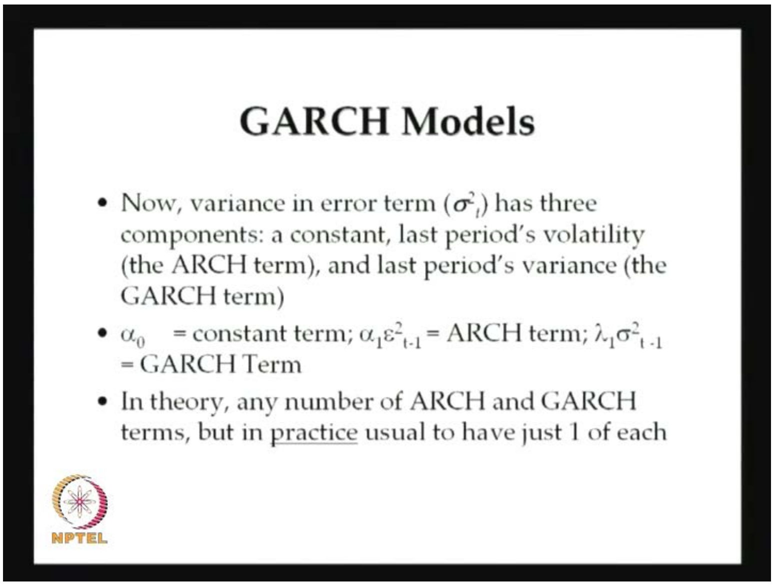
So it is a sigma square t error variance is depends upon the square of error terms the time period t minus 1 and the time you know error variance with respect to time period t minus 1. So that means, we like to know the first we start with a regressing y 2 with a y t minus 1 or x t minus 1 something something. So we have to find out the error component then you have y t with respect to its lag and e t with respect to lag. Then we like to know, we have to find out the error variance, then we have to regress error variance with a square of the error terms, this is one site of the game. In another site of the game, what you have to do? We have to find out the error variance with respect to square of the error terms and the lag of error variance. That means, you have to see the see the effect of the error terms and the effect of the error variance. If your objective is like that way, this particular structure is called as a GARCH models. Generalized, if we it is ARCH of one is to one. So, one is to one means, we are taking the error square of error term with respect to lag one and square of error variance with respect to one.

(Refer Slide Time: 49:33)



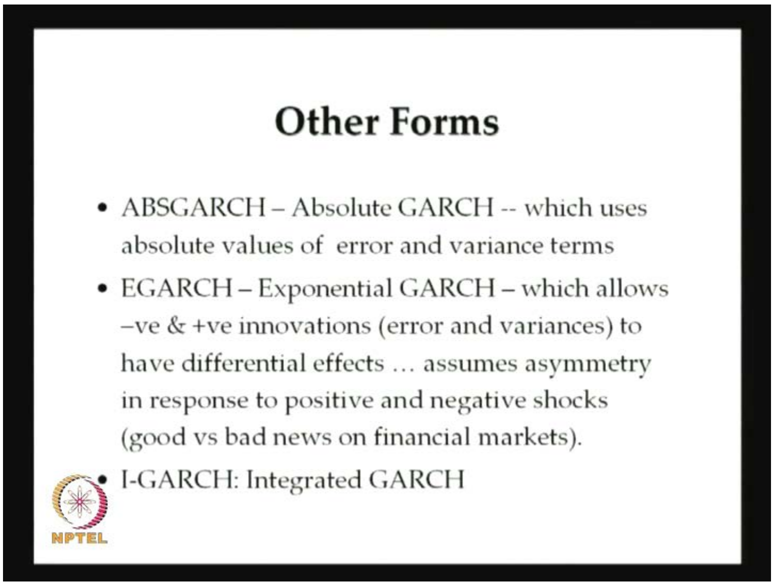
So it can be extended up to p k terms like this way. So general trend of the structure is a sigma square t alpha 0 plus alpha 1 e square t minus 1 plus gamma 1 sigma square t minus 1, this is GARCH of this is GARCH of one is to ones. So similarly, if will we go for GARCH of p q, GARCH of GAR GARCH of GARCH of p q, then obviously the model will be like this way, so sigma square t equal to alpha 0 plus summation alpha i e square t minus i i equal to 1 to n plus gamma summation plus summation gamma i sigma square t minus i i equal to 1 to n.

(Refer Slide time: 50:36)



So this is how the GARCH effect can be fitted. Now, here the objective is you have to find out the constant term. Then arch effect, that is alpha 1error square t minus 1 and you know GARCH term, lambda 1 sigma square t minus 1. So that is why, alpha 1 and sigma 1 is more important, so that should be that is that should be tested. That is our hypothesis that alpha 1 and lambda 1 should not be equal to 0. If it is equal to 0, then obviously it is very difficult to interpret the volatility issue.

(Refer Slide Time: 51:21)



So this is how means, the typical issue here mean null hypothesis to test alpha 1 equal to 0 against alpha 1 not equal to 0; similarly, lambda 1 equal to 0 against lambda 1 not equal to 0. Of course, we had discussed ARCH model and GARCH model, but there is a various difficult for you know, means there is a additional structure of the volatility modelling like EGARCH, GGARCH, IGARCH, etcetera. And it is more complicated, more systematic in fact and because of time constraints, it is not possible to highlight all this details, because it is more advanced and purely it is a research oriented problem. So that is why we are not means time will not permit to go enter deep in to this particular problems. What is my main idea is to highlight the detail about the volatility modelling. So basically, the starting point of ARCH model and more generalized format is GARCH models. Then you know, with respect to GARCH there is lot of different clusters like EGARCH, GGARCH, IGARCH, and etcetera.

This is you know, some partitions with different condition and constraints, but basic thing is the generalized autoregressive conditional heteroscedasticity models, and autoregressive conditional heteroscedasticity model. So our main concern is to see, what is ARCH effect and GARCH effect? So, with respect to different problem and different scenario, different situation, we will go for different step of the GARCH model. So this is not in our aim means, this is not our aim to discuss all these details. So here, we are just highlighting, what is the volatility issue of the time series modelling. So with this, we will conclude this particular session. Thank you very much, have a nice day.