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

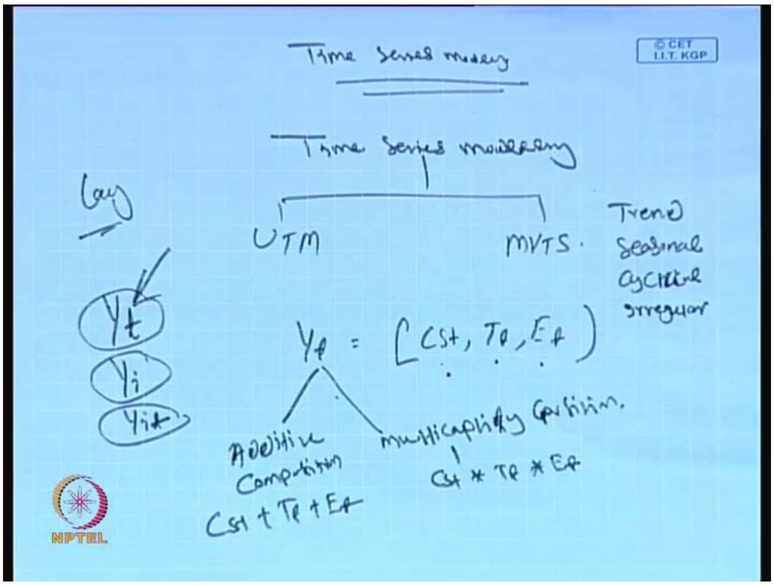
Lecture No. # 36
Time Series Modelling

Good afternoon, this is Doctor Pradhan here. Welcome to NPTEL project on Econometric Modelling. So today, we will start a new chapter called time series modelling. So, in the last couple of lectures, we have discussed the entire structure with respect to cross sectional modelling. Sometimes with Cross Sectional, you need sometimes with time series, you need sometimes with the integration of both time series and cross sectional; that is what we call it is as panel data settings. So today, our discussion is completely, somewhat other side of the pictures, where we will directly handle with the time series modelling. Basically, it will talk about the time series modeling, the structure is completely or somewhat, it is different from the discussion which we have already done.

So, it is partly you know it is something you know very interesting and also something very complex type of problems. So, we will discuss various issues under this particular topics like (( )) modelling and then you need to (( )) co-integration, bar modelling etcetera. With specifically, with respect to time limits, we have to restrict our discussion with certain components. In fact, time series modelling itself is an another subject. So, it is not possible to cover each and every components of the time series modelling, but what is essential points as per the needs, we have to discuss that respects on the…

So the anyway, the time series modeling, what is all about the time series components.

(Refer Slide Time: 02:04)



Time series modelling basically, deals with the two types of problems. One is called as a univariate time series modelling and another is called as a multivariate time series modeling. For instance, when will go for any typical modeling then obviously, the structure is that, we must have a dependent variable and we must have a independent variable. Sometimes, independent variables, there are two different games altogether. In one side, there is one dependent variable with one or multiple independent variables. In another structures, there are series of independent variables and series of dependent variables.

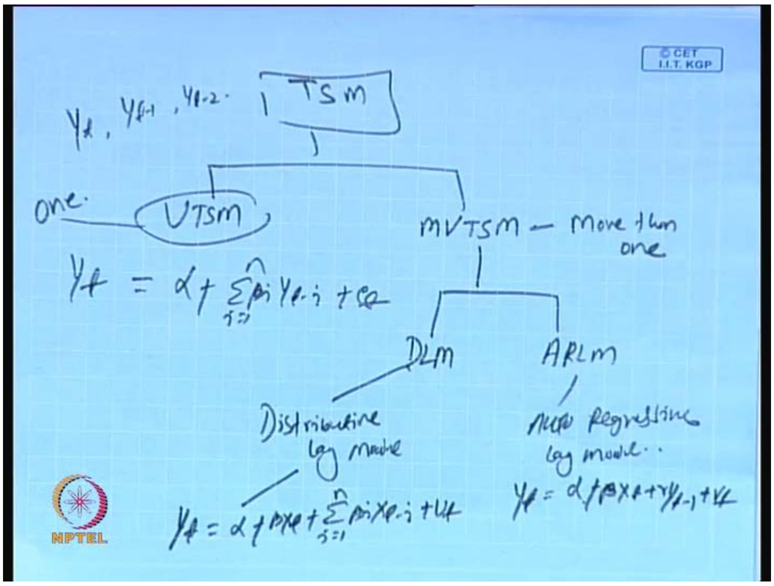
So, in that case, we have discuss the problems like structural equation modelling. Here also, the same things we will discuss, but with respect to time effects only. What is the speciality of time series here? that is very important in this particular lectures. Time series basically, deals with variable with its lag. That means, lag is the most important issue here in the case of time series modelling. Now, what is this lag issue? Basically, before we go to time series modelling in detail, I like to highlight a few things here.

Why there is a time series? Because, time series, let us a variable say Y then if I will put like this t then it is called as a time series component. Same thing, I can put it as Y i, I can put it as Y i t. This is panel data setting, this is cross sectional setting, this is time setting. Now, Our discussion today is, with respect to Y t settings. What is all about the this y t? The specialty of Y t is, it is a function of Cs t, T t and E t. It is nothing, but it is basically divided into two parts; it is additive compositions and this is multiplicative compositions. This is additive compositions, that means, the structure is a like this CS t plus T t plus E t here is and in the multiple sides, CS t into T t into E t. Infact, We can take it is also logger than then you can transfer into different formats.

So, the structure is, what is CS t here is, there means, it is basically, time series a depends upon so many factors. One is trend factors, seasonal factors, irregular components and infact, trend components, seasonal components and irregular components. These are the three which we have discuss here. CS t transfer seasonal trends, this is trend, cyclic trend and you can say irregular trend and you can say sometimes it may be cyclical. So, trend basically it is trend effect then seasonal effect then cyclical effect then it is irregular effect. So, there are four different compositions through time series can be a decompose with respect to additive and with respect to multiplication. That means, why there is a lag?

There are various regions for that. One of means basically we divided all this composition into four different aspects; one is called as a trend components, seasonal component, cyclical component and irregular component. What will discuss today; this is not the class to discuss the basic features of time series. We start with little bit advance label so that we will highlight few important or interesting problems. Now, if will we go, pure time series component with respect to econometric point of view then obviously, we start with like this way.

(Refer Slide Time: 06:44)



Time series modeling, basically, divided into two parts as I have already highlighted, univariate time series modelling then multivariate time series modeling. We start with, let us say variables Y t. I have already mentioned the moment the specialty of time series modelling is that it is lag introduction.

If there is a one variable, we can create a multiple models or you can create multivariate models with respect to its lag length. If we put adding one after another lag, then obviously, new variables will be created automatically. But instantce here is, I will write Y t equal to function of alpha plus summation beta i Y t minus i equal to 1 to n plus E t. This is one of the structure of time series modeling that too is a univariate time series modeling. That means, the origin is the Y t, we are creating Y t minus 1, Y t minus 2. Then, we are integrating Y 2 with respect to Y t minus 1, Y t minus 2 upto Y t minus k depending upon the structure and setup. Very important thing is the, how much lag length you to introduce in this particular systems. The most interesting or you can say complex process of time series modelling is that it is lag length.

The choice of lag length is very important factors. There are two basic reasons for that; one region is that, it is not continuous process like you have to create a one after another variables by introducing lag. The moment you will introduce one after another lag then; obviously, one problem immediate problem will face is that the degrees of freedom. So, you will lose degrees of freedom continuously. Now, to avoid degrees of freedom, what you have to do? You have to keep lag length it the optimum label, there are certain techniques through which we can choose the lag length. We will discuss in details. Let me first highlight the structure of time series modelling then we will go in details about the choice of lag length or you can say a model setting etcetera.

Especially, there are two structures here is; univariate time series modelling structures then multivariate time series structures. In the case of univariate time series structures, the structure is Y t as a function of Y t minus 1, Y t minus 2 upto Y t minus k. Now, in the case of multivariate time series modeling, it can be divided into, means here, the specialty of this particular is, it is one variable setup.

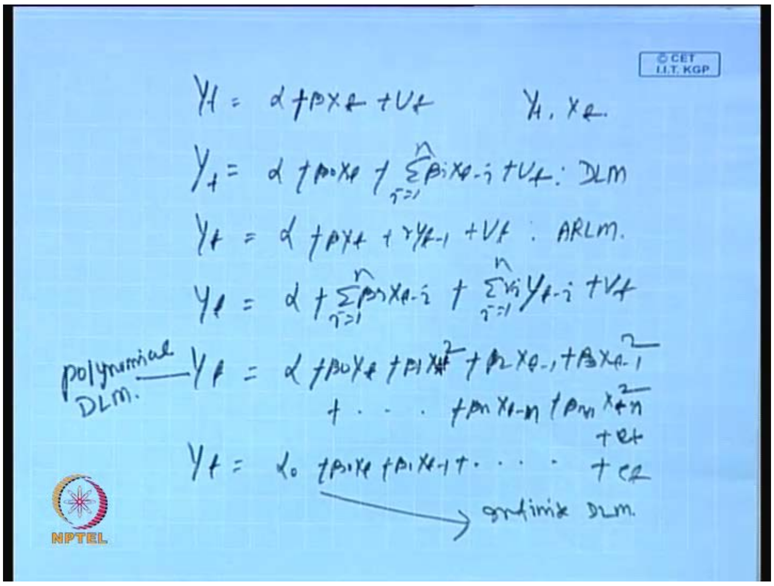
This is more than one variables means the variable is one here, but we are creating additional variables with respect to its lag only. That means, the game is only with respect to single variable. So, now, we are creating with respect to its past observations or past trends then obviously, we are creating additional variable, but in the case of multivariate time series modeling, we at least have two different variables. In that case, if it is more than one time series variable then the model can be further divided into two types; one is called as distributive lag models, another is called as auto regressive lag models.

So, DLM, it is called as distributive lag models then this is called as a auto regressive lag models. what is distributive lag models? It is the structures like, while I put like this; Y t equal to alpha plus beta X t plus summation beta i X t minus i i equal to 1 to n plus U t. This is form of distributed lag models. In the auto regressive lag models, we can write like this way; Y t equal to alpha plus beta X t plus gamma Y t minus 1 plus V t. So, I can right like this way. That means, you see here is, the clear cut difference, with respect to auto regressive and distributive model is that, in the case of distributive lag model, it is the endogenous variable as a function of exogenous variable, but in the case of auto regressive model, it is endogenous variable as a function of both endogenous variable and exogenous variable.

Altogether, time series modelling can be divided into 3 groups; one is called as a univariate time series modeling, another is called multivariate time series modelling that to; auto regressive scheme and another is called as a distributive lag scheme.

We like to know details about what is the auto regressive kind and what is the distributive lag scheme. Let me first highlight various forms of time series modelling that we usually handle in the econometric modelling.

(Refer Slide Time: 12:06)



You see here, generally, we start with like this way, Y t equal to alpha plus beta X t let us say. I am starting with multivariate frame work only, Y t equal to alpha plus beta X t plus U t. This is a simple time series modelling with respect to two variables Y t and X t.

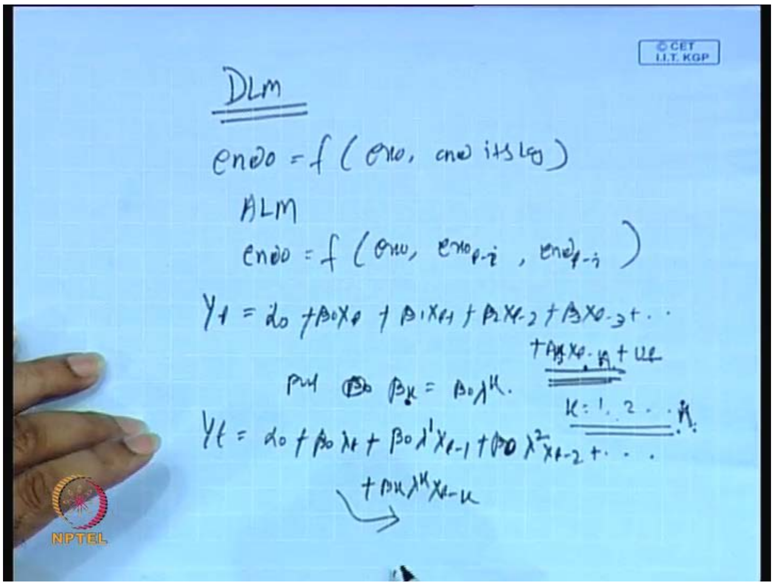
Within this particular setup, I will create an another model here is, Y t equal to alpha plus beta 0 X t plus summation beta i X t minus i i equal to 1 to n plus U t. This is another model, it is called as a distributive lag model. I can write another model, Y t equal to alpha plus beta X t plus gamma Y t minus 1 plus V t. This is another model. This is called as a distributive lag model and this is called as auto regressive lag model that what we have already discussed. Then, another form of the models will present here is, Y t equal to alpha plus summation beta i X t minus i i equal to 1 to n plus summation gamma i Y t minus i i equal to 1 to n plus V t. This is another form of the models. This is lag model, with various lag effect. Then another model, I will put it here, Y t equal to alpha plus beta 0 X t plus beta 1 X t squares then beta 2 X t minus 1 plus beta 3 X t minus 1 squares plus continue beta n X t minus plus beta n plus 1 X t minus n squares plus e t, error terms. So, this particular structure is called as a polynomial distributive lag model.

I will write another model, Y t equal to alpha plus beta 0 X t plus beta 1 X t minus 1 plus continue plus e t. So, this particular scheme is called as a infinite distributive lag models. Altogether, we have various forms of the models, starting with the univariate frame work where Y t as a function of its lag length means you have to create, the moment we will use a lag length then obviously, will create additional variables like Y t minus 1, Y t minus 2, upto Y t minus k. But you cannot create a infinite number of variables in a particular system. Because, if the sample size is limited then obviously, every creation of lag additional variable will lead to loose one degrees of freedom. That means, we have to loose every times one degree of freedom. You must be very careful before handling the or before you proceeding to the time series modeling. These are the schemes under, you can say, various forms of time series modelling.

There may be, generally, if you will go for multivariate schemes, you will find there are two different methods altogether. In one case, it is called distributive lag schemes and another case, it is called as a auto regressive lag scheme.

Generally, between the two, auto regressive lag scheme is much important than the distributive lag schemes because here is we are handling the lag length of both the variables together. In the case of distributive lag model, we are just incorporating dependent variables say Y t as a function of X t and its lag length. That means, what is the exact difference here is? In the case of distributive lag model, it is the endogenous variable as a function of exogenous variable and its error term, but in the case of auto regressive lag model, it is the endogenous variable as a function endogenous lag and exogenous variable and exogenous lag.

(Refer Slide Time: 16:43)



That means, it is like this, in the case of distributive lag models, it is the endogenous variables as a function of exogenous variable and its lag. This is the structure of distributive lag models. In the case of auto regressive lag models, it is the endogenous variables as a function of exogenous variables and exogenous lag t minus i then endogenous lag t minus i. This is how the structure is all about in the case of distributive lag models and auto regressive lag models. I hope you understand the concept of endogenous and exogenous. Endogenous is nothing but dependent concept and exogenous is nothing but independent variables.

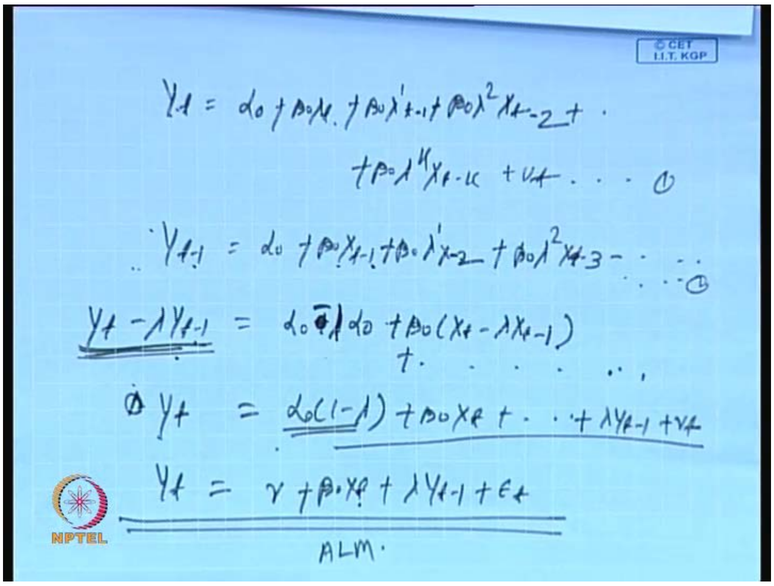
The thing is that, in the time series modeling, it is very difficult to say that it is a variable which is totally independent. In the times series setting, a variable which must have a integration with the previous observation. So, that is why, a variable in the case of time series modeling, has a less impact that it is totally independent to others. There are various problems you have to find out and as a result there are lots of interdependents exist in the case of time series setting. As a result, the variables which you are considering in a particular system, it most have some lag scheme.

That lag scheme we have to highlight and we have to play the game with respect to various lags. Infact, what is the important agenda is here there? Most probably, auto regressive lag models is better choice than the distributive lag model because, here, there is involvement of lag with respect to both the variables; endogenous variable, exogenous variable. In the case of distributive lag model, it is only endogenous variable as a function of exogenous variable and its lag length. But the liking point is that, there is no such endo lag here. But, if the system has a endogenous variable in a system where endogenous variable as a function of lag endo and lag exo including original exogenous variables then obviously, the system will be more appropriate and more authentic. That is why, auto regressive lag model is better choice than distributive lag models, but distributive lag model can be transfer into auto regressive lag model.

Let me highlight, how we can transfer all these things. We will start with a model like this, let us say, Y t equal to alpha 0 plus beta 0 X t plus beta 1 X t minus one plus beta 2 X t minus 2 plus beta 3 X t minus 3 plus continue beta k X t minus k plus U t. So, now, what I will do, I will put beat 0 or I will put like this way, beta k is equal to beta 0 lambda to the power k. So, beta k means, k here stands for 1, 2 upto infact we can this beta k is 1 to upto k or you can take it n. So, then you will put it here n then obviously, it is t minus n. Let us assume that since it is a k number of series, we are putting beta k equal to beta 0 lambda to the power k. As a result, this series equal to continue like this way. I can transfer this particular structures, Y t equal to alpha 0 plus beta 0 X t plus beta 1 I can write it here is, then beta 0 lambda to the power, if I will put beta 1, then it is beta beta 0 lambda to the power 1 then X t minus 1.

Then, if I will put beta k equal to 2, then beta 2, sorry beta 0 lambda to the power 2, then X t minus 2 plus it will continue, then finally, beta k lambda to the power k X t minus k, this particular component, beta k X t minus k. There is no point to again introduce n, since it is k number of variables or k number of lag length. So, obviously, will take this schemes in that fissions. Automatically, the series will transfer to by this particular structure.

(Refer Slide Time: 21:36)



So, what I will do here is, ultimately, we have Y t equal to alpha 0 plus beta 0 lambda t plus beta 0 lambda 1 X t minus 1 plus beta 0 lambda squares x t minus 2, then it will continue then beta 0 lambda to the power k X t minus k plus U t.

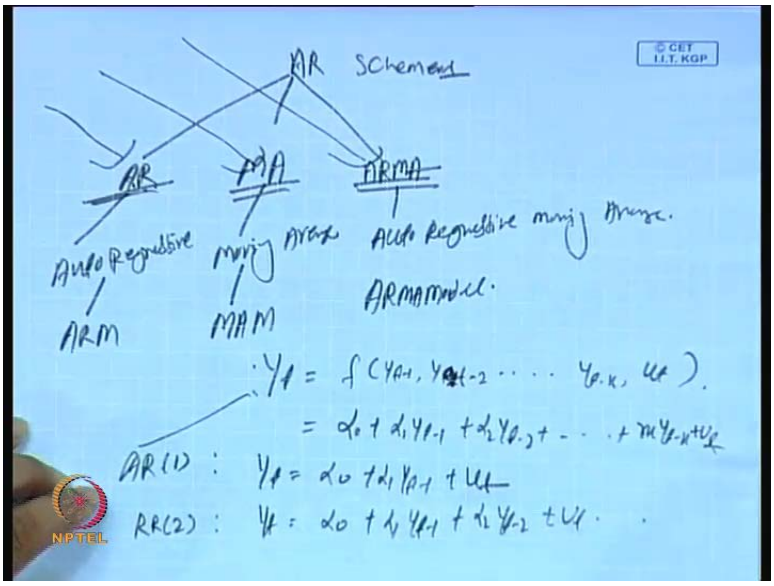
So, what I will do now, I will create another variables here, let us say, Y t minus 1 then alpha 0 plus beta 0 then X t minus one then beta 0 lambda to the power 1 X t minus 2 plus beta 0 lambda to the power 2 X t minus 3, it will continue like this way. Then what I will do, I will subtract it, Y t minus lambda Y t minus 1. Let me highlight like this way, I will create another schemes here is, what i have done is, I multiply lambda in both the sides then I subtract from this equation to this equatio. So, this is equation number 1, this is equation number 2. Now, I subtract the equation number 1 minus lambda upon equation number 2. Then, I will get Y t minus lambda Y t minus 1, then in the sametimes, I will get it here is, lambda 0 , lambda 0 cancels, it will be get, lambda 0 plus lambda, sorry minus lambda alpha 0 then beta 0 X t minus lambda X t minus 1 then this will continue like this way. So, I will call it this one, Y t. Then, this is alpha 0 into 1 minus lambda then I will call it beta 0 X t plus it will continue plus lambda Y t minus 1 plus V t. Here is, what I have done is, Y t then minus lambda Y t minus 1. So, I got this equation in the left side. In the right side, this is the structures. If I will simplify, I will get like this way. I have not generalizing this concept.

Now, what I have done, I am putting Y t equal to here, then lambda Y t minus, I am taking in the right side. So, finally, I will call it is, let us say, gamma then beta 0 x t then, other item will be remain constant, if I will put it in sequence. Then obviously, it will come down to this a particular structure, beta 0 x t plus lambda Y t minus one plus epsilon t. So, this is the form of auto regression lag models. So, now, this particular structure is called as a time series modelling with respect to distributive lags schemes to auto regressive lag schemes.

What we conclude here is, that auto regressive lag model is a more important than the distributive lag models because, there is involvement of lags with respect to exogenous component and endogenous component.

We like to understand, what is auto regressive schemes and how it is integrated with the system?

(Refer Slide Time: 26:23)



So, we will start with auto-regressive schemes. What is the auto-regressive scheme here is, auto-regressive scheme, we start with, there are various forms of the auto-regressive schemes. Initially, it is the simple auto regressive models, then moving average models then its integration ARMA models. So, auto regressive, this is stands for auto regressive, this is stands for moving average, this is stands for auto regressive moving average. So, now, we start with the, this is one scheme, this is another scheme and this is another scheme; that means, the modeling can be applied to this one, modeling can be applied to this one, modelling can be applied to this one. If I will apply modeling to auto regressive scheme, then it will called as auto-regressive models. If I will apply modelling to this one, then it is called as moving average models. If I will apply modelling to this one, then it is called as ARMA models. Let me start with, what is this ARMA schemes, what is the ARMA, since here is that, let us say, Y t is a function of Y t minus one, Y t minus 2, it will continue, Y t minus k plus something error terms u t. This is the simple model, I derived. In the time series modeling, there are two different structures altogether, it is called as univariate structure and multivariate structure. In the case of multivariate structure, there should be at least two variables with respect to its lag, means, with respect to two variables; say Y t and X t, then X t with respect to its lag and Y t with respect its lag. That means, finally, how many number of variables in the systems, you do not know. So, it is the technique will decide, what is the total number of lag involve in this particular structures and what should be the size of the models altogether.

The thing is that, since there are two schemes univariate time series modelling and multivariate time series modeling, in the univariate structures, Y t as a function of Y t minus 1, Y t minus 2 upto Y t minus k. So, what we will do here is, in this particular univariate time series setting, we will like to connect with moving average concept and its integration with auto regressive concept. So, the scheme, Y t as a function of Y t minus 1, Y t minus 2 upto and error terms, that scheme is a called as auto regressive scheme, that is, very very important structures.

So, Y t as a function of its lag, that means, it is a endogenous as a function of endo lags. This particular structure is called as auto regressive scheme, that is the special features of time series modelling. That is how, auto-regressive model is more important than the distributive lag models. Because, we are starting, in the time series setting, then obviously, a particular variable always depends upon its past variable. For instance, let us take a case of a stock price, then stoke price today depends upon the yesterday stock price value. Take case of GDP, GDP for the today, say for this year, GDP t depends upon GDP t minus 1. That means, in the time series setting, one of the standard assumption is that, the variables in the present contest depends upon its past contest; that means, GDP at a time period setting, that is called as a current time period, then it depends upon its past time periods like t minus 1, t minus 2, t minus 3 like this way.

So, theoretically, if you say, suppose I will like to invest today, the impact will not get immediately. For instance, if suppose, I have some amount of money and I like to invest in educational sectors. So, my impact, my objective is to know, if I invest this much of money, then what is the impact of education. So, today, if you have to spend huge money, that is called as educational spending. Then obviously, the impact cannot be immediate. The impact will be in the long runs and that too next year's only, that means, suppose I like to study the impact of the this spending with respect to its effect. Effect means, whatever may be the concept or means, you say, literacy label or something, something. In that case, we will say output after 5 years, then 5 years which we have to connect to with the present this investment, that means, it is the 5 years lag we have to introduce.

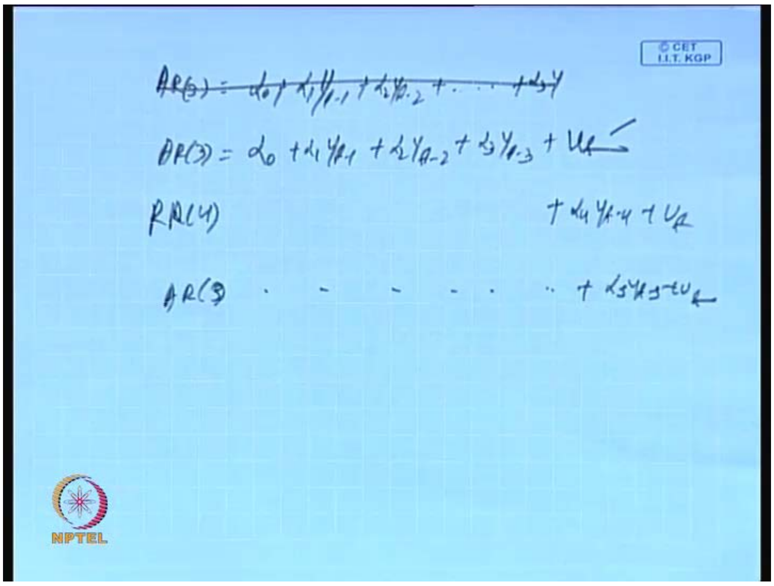
If that is the Y t and obviously, Y t depends upon its Y t minus 5. This is how; the last scheme is integrated in the system. It should be continued, it should be very systematic, should be very visible in natures. For instance, you cannot directly jump Y t to Y t minus 5 immediately. You go in sequence, then you test it, after that you can come to a conclusion, how much reliant you can incorporate to get the particular models or occurs the objective is moralist. So, we have to find out the optimal models.

That model should be best fit and can be use for forecasting year policy that objective always there see whatever component will discuss, means that we have already discussed and what we will discuss in feature also. Every time objective is to fix a model should be very reliable, very feasible, very accurate and it should be with forecasting and policy use. So, keeping in mind, we are moving here and there, send putting difference structures saves and setup so that the model can be consider as the best fit and can be use as for forecasting and policy use.

Now, our system is here to discuss the auto regressive schemes. So, auto-regressive scheme is basically three different formats; auto regressive itself it can be model, then moving average concepts, then you can say ARMA concept. We have to see how auto regressive and moving average and ARMA can be finally concluded, means created. Let me start with the variables Y t the when I connect Y t with respect to its lag length, this particular structure is called as auto regressive scheme.

So, now this particular structures I will call as a auto regressive scheme. For instance, if I will put it in a explicitly format, the picture will be coming like this way, alpha 0 alpha 1 Y t minus once plus alpha 2 Y t minus 2 plus continue alpha k Y t minus k plus U t. Now, you see, auto regressive scheme for 1, if I will have auto regressive scheme of 1, then I will write, equal to simply alpha 0 plus alpha 1 Y t minus 1 plus U t. If I will say, auto regressive upto, then you have to, this 2 1 to represent the lag length schemes. So, Y t as a function of alpha 0 alpha 1 Y t minus 1 plus alpha 2 Y t minus 2 plus U t. This is how the structure will be continued..

(Refer Slide Time: 34:32)

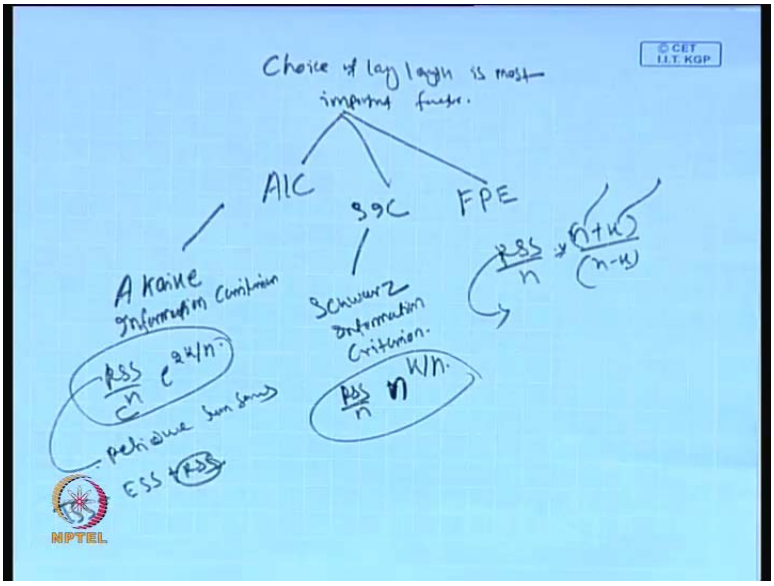


I will put another structure, auto regressive 3 is equal to alpha 0 plus alpha 1 Y t minus 1 plus alpha 2 Y t minus 2 plus alpha 3 Y t minus 3 plus U t. This is another auto regressive scheme of order 3. If I will put here auto regressive 4, then again, you have to add further, alpha 4 Y t minus 4 plus U t. So, if I will put auto regressive 5 then; obviously, it will be continued plus alpha 5 Y t minus 5 plus U t. This is how, it will be continued. The thing is that, mathematically if will write, then it is very interesting and you are just adding one after another and you just know the structure and you can continuously follow this set up. But in reality, you cannot create unlimited variables within a particular setup. So, there is a systematic approach, how much lag length you have to consider.

It is not you to decide, it is the system will decide, what should be lag length, is it auto regressive model should be best with respect to auto regressive scheme 1, auto-regressive scheme 2, auto regressive scheme 3. Mathematically, when will go for proceeding like this way, of course, what you have to do, it is just like a stepwise process. Infact, we have a system called as stepwise regression. So, that is specially use for solving multicollative problem, but here the scheme is auto regressive scheme is just like stepwise scheme. So, what you have to do here, one of the most important agenda is, how much lag length you have to incorparating the system. If to the current year outcome of particular variable depends upon its how much lag length.

So, then it is what should be fast observations, that is how it is a means you differ predicting the Y t with respective feature. We have to know, how it is continuous? Let us with the past information, that is our entire objectives; it may be 1, it may be 2, it may be k. So, it is not you to decide, means it is the structure or technique will decide the same.

(Refer Slide Time: 37:21)



Let me, how we will find out this structure, choice of lag length is most important factor. It is very, very important issue. Why means, there are two specific region; first region is a the moment will introduce one after another lag then; obviously, you will get every time will get one after another parameters. That is to be estimated in the same times when will add one after another variable you will like to loose one degrees of freedom. So that means the moment will go for one extra lag then obviously one variable has to be increased in other side, degree of freedom will be slightly sorts, that means, one additional parameter with one degrees of freedom, that is your equation. If will another variable then; obviously, two more parameters will be added, then two degrees of freedom you are going to loose. So, this is how because of this particular region, it is very, very important to understand the optimum lag length. Infact, in the time series modeling, one of the condition is that, frank condition is that, your number of sample observations should be exclusively very high. If it is not very high then it is very difficult to choose the lag length or to fix how much variable you have to create within a particular system.

Now, what is our conclusion is that, choice of lag length is most important effect to decide the optimum models time series model. There are many techniques are available. I will give briefly, you have 3 different important criteria, through which you can choose this lag length; one is called as a AIC statistics, then SIC statistics, then FPE final prediction error. So, AIC represents A kaihe information criteria. So, the statistic is, it is nothing, but RSS by n e to the power 2k by n. This is the formula, RSS represents residual sum square. So, we know total some square equal to explain sum square plus residual sum square.

We need some square, infact, in the case of time series modeling, the residual factor is very important. Because, it is very important to handle various issues, various structures, various setup. One of the agenda to decide the best fit of the model is to see the structure of RSS. So, now, through A kaihe information criteria, you have to find out the best fitness, means optimum lag length and goodness of fit of the model.

So, another statistic called as Schwarz information criteria. So, this is nothing, but RSS by n n to the power k by n. Here, RSS represents residual sum square, n equal to total number of observations. k represent number of variables in the systems. For instance, if it is lag length 1 then; obviously, there are two variables in the system. If lag length 2, then there are three variables in the system; Y t, Y t minus 1, Y t minus 2. Similarly, if it is lag length 3 then; obviously, there are four variables in the system. This is how, you have to find out the number of the k value. So, SIC will be calculated like this way, then AIC will be calculated like this way then final prediction error is nothing, but RSS by n into n plus k by n minus k. So, this is called as final prediction error. So, now, you see here n already defined, k already defined, RSS already defined. Like this, there are several other criteria also to decide the optimum lag length and best fitness of the model. In all the cases, you will find residual sum square is the criteria through which you have to decide the structure.

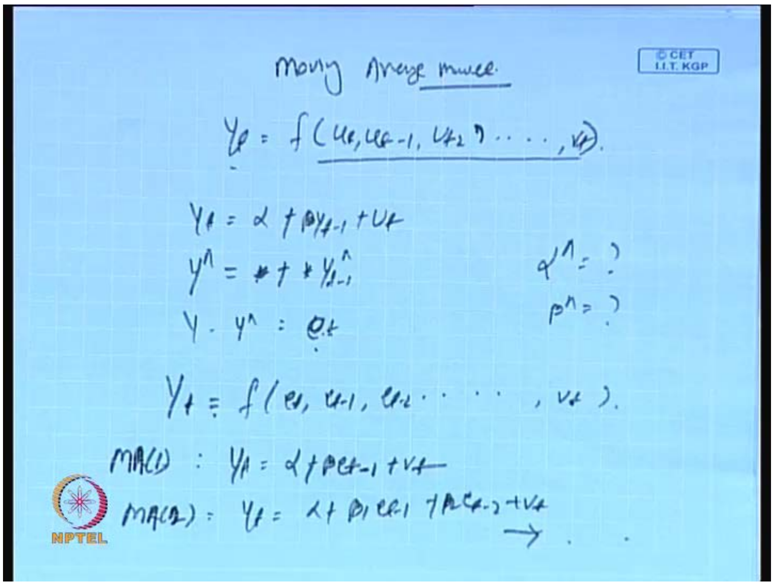
Three important items are very important in this particular time series modeling. One is RSS that is residual sum squares, then n represents total number of observations and k represents total number of variables in the system. Say lag length; if the lag length is one, there are two variables. If lag length is two, then there are three variables in the system. So, once you add one after another lag, then you will get additional variables one after another, but in the sametime, you are going to loose lots of degrees of freedom. So, that is why you must be very careful.

By this criteria, how to decide, which particular lag length is important means we have to fix in the systems . In a particular system, if the all the statistics are very low and it is continuously declaim then; obviously, you have to choose the lag length continuously otherwise if the value of these statistics are very high then you have to stop the lag length. If will make a comparative analysis, two different time periods of two different models, where AIC statistic is less that is to be considered. Over the time periods, it is one setup Y t, Y t minus 1; another setup Y t, Y t minus 1, then Y t minus 2; another setup Y t, Y t minus 1, Y t minus 2, Y t minus 3. So, every case you have to find out AIC statistic or SIC statistic or FPE statistics final prediction error statistic. Then you check it,where it is low. So, that particular model has to be considered as the best model.

So, once it is considered as the best model then further additional variable cannot be introduce in the systems. So, you have to continuously check, one AIC statistic declining; obviously, you have to add one after another variable in the system. The moment it will be stagnant or start increasing then; obviously, you have to stop there and you have to finally, say that this is the optimal lag. You have to consider for the best fitness of the models So, this is the scheme called as auto regressive schemes.

Now, coming to moving average schemes.

(Refer Slide Time: 44:30)



So, there is a another called as moving average models. Moving average model is the systems it is called as Y t as a function of U t, U t minus 1, U t minus 2 then it will continue, another variable say V t, error term.

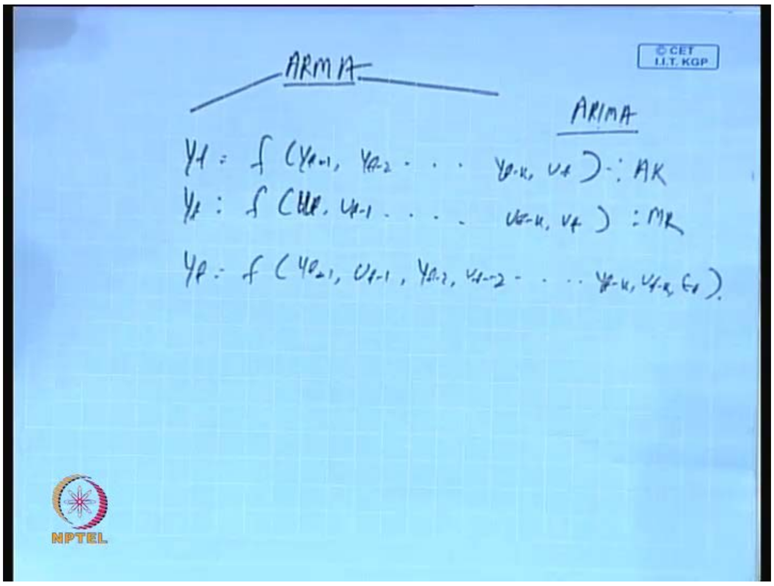
So, U t as a functional error term and its leg length. So, this is more interesting infect what we know earlier. So, we start with like this, let us say, Y t equal to alpha plus beta Y t minus 1. So, let us say this we start with model like this U t alright. So, then ultimately we need to find out estimated model Y head. For that, we need to have alpha head and we need to have a beta head. So, once will get it, this one then accordingly, we have to fit this models Y t minus 1 head then; obviously, error term will be removed so; that means, Y minus Y head indicate the error terms e t or u t, you have v t then; obviously, the modelling will be start with respect to e t again. So, Y t is a function of e t minus one, e t minus 2, it will be continued with error another variable scheme. Like auto regressive schemes, we can prepare for moving average scheme also. For moving average, this is not a moving average, this is moving average 1, then this system will be Y t equal to alpha plus beta e t minus 1, this is v t.

So, now moving average upon 2, then it is nothing, but Y t equal to alpha plus beta 1 e t minus 1 plus beta 2 e t minus 2 plus v t. So, this is how, it will be continuously increased. In one setup, you get the auto regressive scheme. in another setup you get the moving average scheme. In the case of auto-regressive model, it is the endogenous variable as a function of endogenous lag of endogenous variable. So, if there are three lags then; obviously, Y t minus 1, Y t minus 2, Y t minus 3. If it is four lag; Y t minus 1, Y t minus 2, Y t minus 3, Y t minus four. So, like this you will continuously increase the lag length. So, the moment you will increase the lag length then; obviously, the model complexity will start increasing. So, your work load will be very high because have to estimate the parameters old together

And also you have to check the reliability of the model or overall fitness of the model. The structure is almost all simple econometric modelling which we have already discussed in the cross sectional setting and penal data setting. So, here is, the setting is, you start with one variable or two variable or five variables then what is the more additional part or interesting part is that, you have to introduce the lag component. If there are suppose two variables Y t and X t, then you have to kaihe Y t minus 1, X t minus 1, Y t minus 2, X t minus 2 continues form this particular structures. There are many cases like this.

So, you have to find out most important one which will be considered as the best and can be use for forecasting or policy use. It is altogether, it is decision making process. Out of many possibilities, you have to find out the particular case which is more important for you, that is how you have to choose this particular setup.

(Refer Slide Time: 48:45)

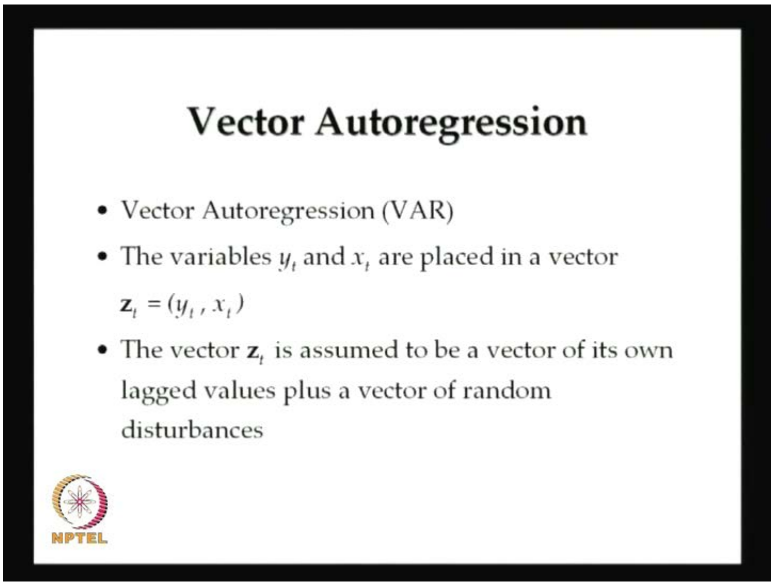


Now, another scheme is called as ARMA. We know, what is auto regressive, what is moving average. It is called as auto regressive moving average, sometimes it is called as ARIMA, autoregressive integrated moving average methods. So, we will know what is integrated that is order concept, integrated of order what so; that means, first let me highlight, ARMA then will go for integrations order of integrations

So, we know Y t as a function of Y t minus 1, Y t minus 2 upto Y t minus k then U t. So, then this is moving auto regressive schemes. Then, moving average scheme is a function of Y t u t minus 1, it is u t minus k, you can say v t. So, now, what will do this is auto-regressive scheme in this is moving average scheme. So, now, what will do will integrate together

So, now Y t is a function of Y t minus 1, U t minus 1, Y t minus 2, U t minus 2 it will continue Y t minus k, U t minus k then error terms say epsilon t. There are three different structure of models you will find here is; one is called as auto regressive schemes, another is called as moving average schemes, then finally, you have to add together then that is structure is called as autoregressive moving average scheme. In the case of auto-regressive moving average scheme, it is the endogenous variables as a function of endogenous lag and error lag. How do you get all this concepts, now you see here is, I will highlight here is, this is let me find list first. take it this case, this is how univariate time series, multivariate time series.

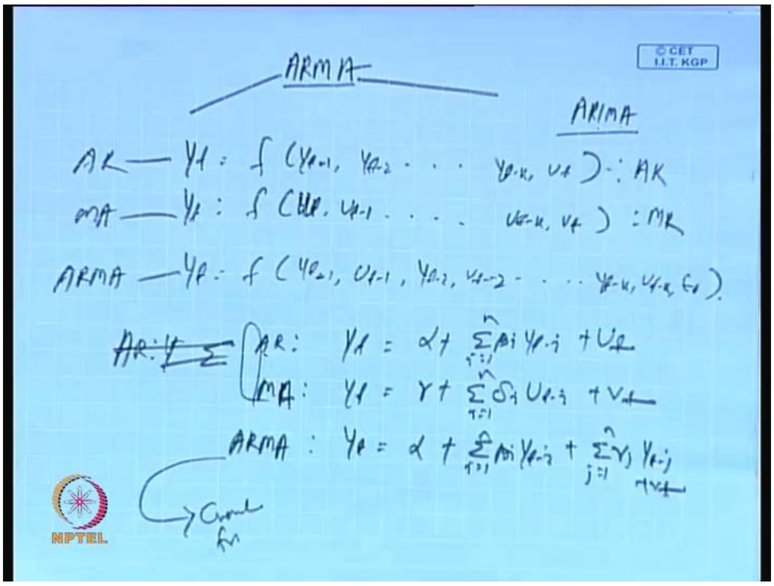
(Refer Slide Time: 51:05)



So, this is a multivariate auto regressive scheme. Here, there are two variables in the systems x t and y t. So, now, we are not integrating x t, y t here together. So, what we are doing, we are taking one variables then we are integrating with auto regressive scheme, moving average scheme and ARIMA schemes

So, now, I have already highlighted with this setup of auto-regressive scheme, what is the setup of moving average scheme, then what is structure of the integration of auto-regressive and moving average scheme. Let me highlight one thing here is model explicitly.

(Refer Slide Time: 51:45)



So, this is auto-regressive, this moving average. So, this is ARMA. What will do in the case of auto-regressive scheme, I will write y t sorry in the case of auto-regressive schemes. I will write y t is equal to alpha plus summation beta i y t minus i equal to 1 to n plus u t, this is auto-regressive schemes and moving average scheme is nothing, but y t equal to gamma plus summation and delta i u t minus i i equal to 1 to n plus v t. So, this is auto-regressive and moving average this is auto-regressive and this is moving average then I will write auto-regressive moving average auto-regressive moving average which is nothing, but y t equal to function of let alpha plus beta i beta i y t minus i i equal to 1 to n plus summation gamma j y t minus j j equal to 1 to n plus v t

So, this particular structure is called as a ARMA models or see this is ARMA in general format this is ARMA in general format this auto-regressive and moving average in a general frame works; that means, altogether there are 3 different schemes you will find in this particular setup one is called as a auto-regressive setup moving average setup and ARMA setups. Altogether what we have discuss today is that the basic of time series modeling, that too you know time series modelling is divided into two parts; one is called as univariate time series schemes and another is called as multivariate time series scheme. In the case of univariate time series scheme, it is the endogenous variable as a function of its a lag means endogenous variable as well as a function of a another endogenous variable it is sorry same endogenous variable its lag length. For instance, y t as a function of y t minus 1 y t minus 2 like y t minus k, but in the case of multivariate time series modelling. So, there are at least two variables then; obviously,. So, we will in onre case it is called as a auto-regressive lag models another case it is called as distributive lag models in the case of auto-regressive lag model it is a endogenous variable as a function of a exogenous variable and it is lag and a endogenous lag and in the case of distributive lag model it is the endogenous variable as a function of exogenous variable and exogenous lag. So, that is the complete a means clear cut difference between univariate time series modelling multivariate time series modelling that to distributive lag scheme and you can say auto-regressive lag scheme

So, now what we have discussion another scheme is called as a ARMA scheme. So, that to with respect to univariate setup only. In the case of univariate setup. what we have to done, we three difference modeling, that is auto-regressive models, moving average models and auto-regressive moving average models. So, the scheme of auto-regressive is that it is the endogenous variable as the function of endogenous variable and it is lag and moving average scheme. It is the endogenous variable as a function of error term and its error lag. For instance, in that particular case. So, first you see start with the regressive y t with the y t minus 1 if provided the system will be say y t as a function of y t minus 1 only. So, then you find out the error term then again once we have a error term create a error lag y t, e t minus 1, e t minus 2 like this way. So, then you have to connect y t with the with it e t minus 1, e t minus 2, e t minus 3 like this way. So, it is this say to way process in that process it is called as moving average scheme then finally, it is called as a ARMA, it is the integration of average auto-regressive scheme and moving average scheme; that means, it is a endogenous variable as a function of endogenous lag and error terms and error lag

So, this particular structure is x called as A R ARMA structures infect ARMA has a different order of integrations like you one to the moment we have put out auto-regressive 1 auto-regressive 2 moving average 1moving average 2 like this. So, similarly called ARMA 1, ARMA 2, ARMA 3 like this way. So, 1, 2, 3 is called as a degree of order of the integration.

So, we will discuss in detail in the next class. So, with this, we conclude this particular session.

Thank you very much. Have a nice day.