Applied Statistics and Econometrics

Professor Deep Mukherjee
Department of Economic Sciences
Indian Institute of Technology, Kanpur

Lecture No. 33

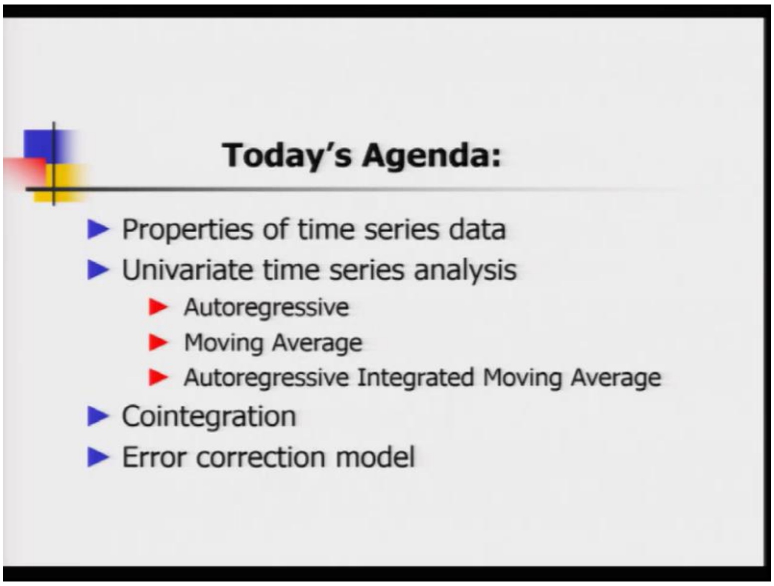
Time Series Regression with Non-Stationary Data

Hello friends welcome back to the lecture series on Applied Statistics and Econometrics. So, today we are going to continue our discussion on time series econometrics. So, broadly speaking there are 3 major time series schools and one of them we have already dealt with in the previous lecture. So, that school believes that if you have stationary data, then you can conduct regression analysis and then you can have different types of models like autoregressive distributed lag models and all.

And there are two other schools and we are going to talk about them in today's lecture. One school models time series by utilizing only one variable. So, that is called univariate time series analysis and this school was mostly proposed by Box and Jenkins back in 1970s and then there is another school which is from United Kingdom and there are some major names who are acclaimed time series econometricians in this school like Clive Granger and Robert Engel.

And they say that well you can still do econometric modelling in terms of regression analysis whenever you have stationary and non-stationary data. So, in today's lecture, we are going to talk about the second school and the third school. The second school being the univariate time series analysis school and the third school is basically a more structural school.

(Refer Slide Time: 1:56)



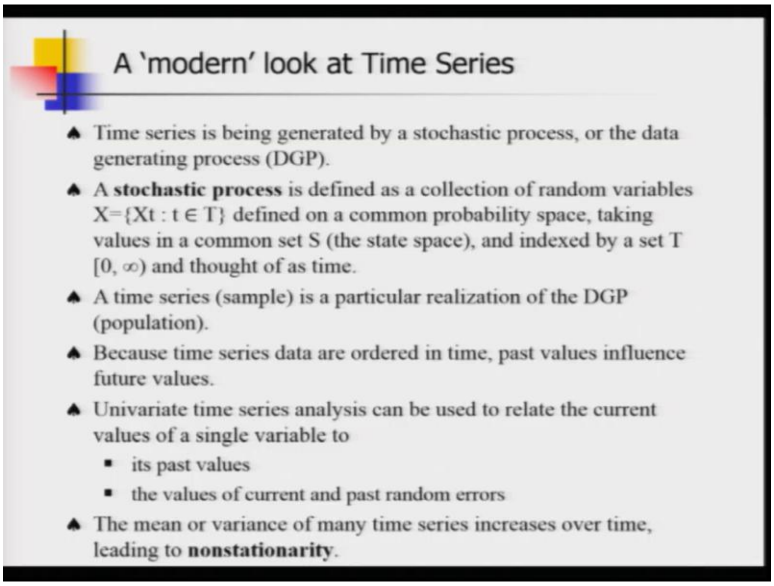
So, before we start talking about the theoretical models, let us have a look at today's agenda items and first we are going to briefly talk about properties of time series data and it is important to remind you again the basic features of time series data although we have discussed little bit of that in the previous lecture because time series is very special.

And the second topic of today's lecture is univariate time series analysis and there I am going to discuss about 3 types of models and they are autoregressive, moving average and autoregressive integrated moving average and the last one has an acronym which is more popular than the full name and that is AIRMA.

Now the third topic of today's lecture is going to be based on the work by Nobel Lodiate econometrician Clive Granger and the topic is cointegration and error correction model is also another area where Clive Granger contributed. And this is the last topic of today's lecture.

So, now we are going to remind you about some special properties of time series data. Some of them will be repetition of what we have discussed in the previous lecture but I am going to talk about some new concepts and maybe I will give you a different perspective on the same old concept over which we have already had a discussion.

(Refer Slide Time: 3:34)



So, time series can be seen as a stochastic process. And here the data is being generated by a stochastic process. So, the data generating process is purely stochastic. Now, what is a stochastic process? A stochastic process is defined as a collection of random variables X. So, here you see I am defining that set X here, Xt actually is the observation of time series variable at a time point small t and then small t actually belongs to a set of time series periods and that is denoted by the exhaustive set is denoted by capital T. And this collection of variables are defined on a common probability space taking values in a common set S the state space and indexed by a set capital T.

So, capital T we have already defined here, so this capital T is basically any number between 0 and infinity. And this is basically the nature of the time series data. So, a time series sample is basically a particular realization of the data generating process which is basically the population. And because time series data are all ordered in time, past values influence the future values. So, now let us talk about the first 3 bullet points in the slide that I have shown you.

So, here, you think about a time series data, it can be an annual data on some economic variable or it could be even daily data on some variable, it does not matter, but the point is that when you are observing a particular value of a time series variable for a particular time period, that is basically a random drop from a population.

And what value is to be chosen it is not in your control of course, because even not drawing the random sample, like in the case of cross sectional data, you have this flexibility that you can actually design a sample survey and then you can have some strategy to pick random samples from the population.

But when you are dealing with time series data, it is very special in the sense that although the observed value at a time period t is a random drop from the probability distribution that pertains to population, but you are not drawing that random sample. So, it is basically Mother Nature or God is pulling one number from that distribution.

So, basically this distribution is there and all the realised values of time series variables are the realizations from that distribution. So, basically behind this numbers that you observe in a time series dataset, there is a stochastic process and that is what I have talked about in the first 3 bullet points.

Now, as the time series data are all ordered in time, past values influence the future values. And this is a very important point, this fourth bullet point that I just have shown you on the slide because that leads to the second school of thought which is basically the univariate time series analysis. So, here the proponents of this school say that as the data is ordered in time, and there are strong relationships between the variable values in the past and the present period, we do not have to have any exogenous variable from outside to explain the variation in the variable in question.

We do not have to have any explanatory variable in the regression equation to predict the future values of the variable in question. We can make use of the past values of the same variable to model the variable in question if that is not sufficient, we can actually make use of the stochastic random error term also to help us in predicting the future values or modelling the variable in question. So, this is basically in a nutshell, the second school talks about. So, this is the univariate time series data analysis school.

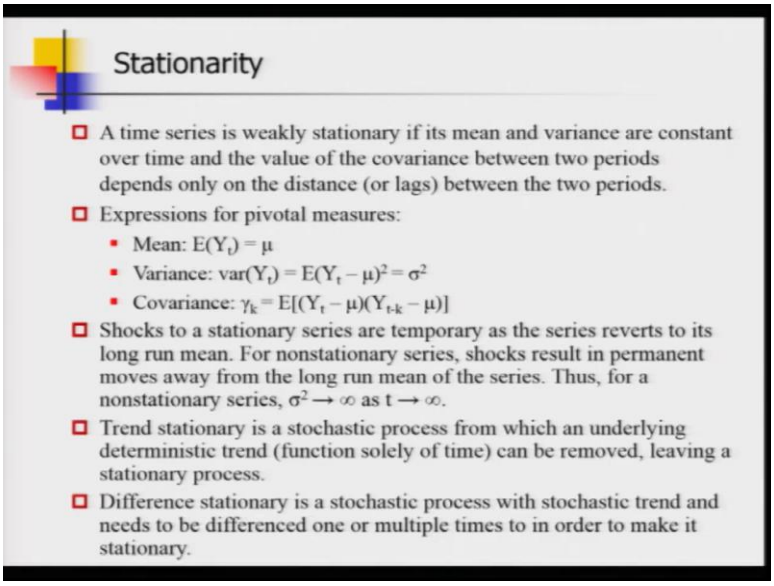
Now I am going to end this slide by talking about another property or feature of time series data and this is going to be the pivotal point of today’s lecture. So, the last bullet point says that mean or variance of many time series increases over time, leading to non-stationarity. So, let us have some discussion over this last bullet point because I already said that this is basically the pivotal term for today’s lecture. So, the entire lecture is based on an assumption that your time series data is non-stationary.

So, in the previous lecture we ended by defining what do we mean by stationarity and violation of those conditions, those 3 conditions that I have talked about constant mean, constant variance and covariance depending on only the distance between two time periods or the lag if you use the formal jargon from time series econometrics. If any one of these conditions are violated, then basically you have non-stationary time series data.

So, all the regression models that we have actually conducted in the previous lecture they assumed stationarity of the data. There could be other problems like autocorrelation and all, but we have assumed that stationarity is there. But, as I also told you that there are many macroeconomic and other economic time series variables which are not stationary, actually if you plot the data over a large period of time periods, then you will find that not only there is a trend but the variance is also changing with respect to time. So, may be the variance is increasing as the time increases, we do not know a priori.

So, here in this particular lecture, I am going to teach you some very basic constant and models that deal with non-stationary data because if your data set is non-stationary, then whatever we have learnt in the last lecture you cannot apply.

(Refer Slide Time: 10:37)



So, in this slide let us talk about stationarity once more, it is basically brief recap of what I have already talked about in the last lecture, but if you a forgotten no need to worry. So, for better continuity I am repeating some of the stuff from the last lecture. So, a time-series is called weakly stationary or simply stationary if its mean and various are constant over time and the value of the covariance between two time periods depends only on the distance between those to time. Periods.

And I have already show the conditions in the last lecture, now if you want to express those conditions positionality in mathematical or statistical terms so, here I have the expressions for you in the second bullet. Now why stationarity is so important? So, of course, there are some statistical reasons, but if you do not want get in to the complex statistical analysis and theorems and proof, can you say in simple language why we are so crazy about no stationarity an all? What is the problems if you have non stationary data? What will be a big problem if I do not have stationarity if my mean and various of the time series variable is increasing with time?

So, basically the problem lies here, if there is a shock, then you will be way far from the normal trend line or normal pattern or normal over all long run trend in the data. But if there is non-stationarity, then actually the chances are very less that you are going to come back to that original equilibrium and you will follow the previous trend. So, you will move actually further away from the trend that you were following when the shock happened. So, basically there is a departure from the initial trend as you note time increases.

So, that is why non-stationarity actually very bad thing because it explodes the time series data. So, the point that I just has spoken about can be nicely captured by one single statistical statements which says that sigma square will move towards infinity as T tends to infinity. Now we are going to differentiate between two stationary concepts. There are two types of stationarities that we come across in literature and one is called trend stationary and the other one is called difference stationary.

We, before we look into the formal definition, textbook definitions for these two stationarities. Let us have a brief discussion using simple terms. So, if you remember when we did even classical time series analysis I said that there are some components of a time series variable namely trend, seasonal, cyclical and the random component.

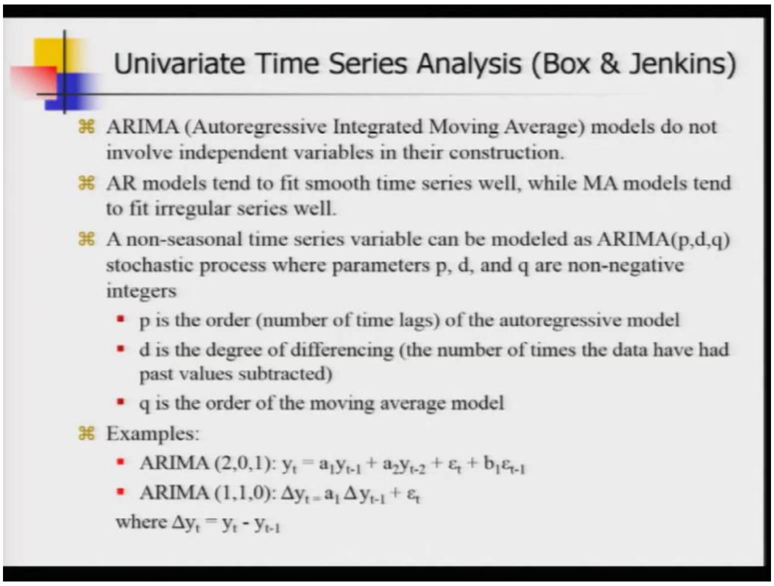
Now, this stationarity is all about this trend component. Now, trend could be of two types, one of course, we have already handled that and that is like a deterministic trend and how do you get that deterministic trend component? By fitting some regression line through the data points.

So, we have seen that there could be linear trend, there could be quadratic trend, but whatever trend is there, the crucks is the it is a deterministic thing. So, it is not going to change with time. Whereas, if trend also changes with time, then trend becomes stochastic and then you have whole lot of problem. So, this distinction whether your trend is deterministic or your trend is stochastic, that leads to this concepts trend stationarity and difference stationarity.

So, now let us have a look at the two definitions. Trend stationary is a stochastic process from which an underlying deterministic trend, solely function of time can be removed and that will leave a pure stationary process and this pure stationary process is sometimes called White noise in textbooks. And in contrast, difference stationary is a stochastic process only but with a stochastic trend and it needs to be, the data needs to be differenced one or many times to make it stationary.

So, I have already spoken about how to take difference in the previous lecture, we are again going to come back to this issue of differencing in this lecture also, but the crucks of difference stationary is that if you take 1 or 2 rounds of difference, then that stochastic trend disappears from your data and you are left with only the white noise.

(Refer Slide Time: 16:10)



So, now in this slide I am going to talk about the univariate time series analysis which is basically the second school of time series and this was mostly established by Box and Jenkins in 1970s. So, Box and Jenkins proposed these terms ARIMA which became very popular later on and in this lecture we are going to devote some time to understand what these terms are and if I follow Box and Jenkins, then how I can model non-stationary time series variable.

So, first I will start with the full form of the acronym ARIMA, and the full form is Autoregressive Integrated Moving Average. So, in the last lecture only we have spoken about autoregressive and moving average. Integrated is the new term and of course, we are going to see what integrated means in next couple of slides.

So, this ARIMAs types of models, then do not involve any independent explanatory variable in their construction because they believe that the past values of the variable in question are good enough of course, together with the error structure it is good enough to predict the future values of the variable in question.

So, in ARIMA, there are 2 major models and one is autoregressive and one is moving average. So, the autoregressive models tend to fit smooth time series well, while the moving average models tend to fit the irregular series well. So, you have to first plot the data against time and then you have to see what kind of fluctuations you are observing, you see too many and too much fluctuation, then basically the data is telling you that moving average model is a better choice and if you see more or less smooth pattern, less fluctuation, then AR could be the better option.

But you just cannot choose a model by eyeballing like this. So, there has to be some statistical criteria by following which you are going to choose your time series model. So, we are going to discuss how you can make use of statistical knowledge to pick a good model for your non-stationary data. In fact, in many cases we see that a combination of AR and MA models are actually sort by the data. So, you can actually have both autoregressive and moving average components in the ARIMA model that Box and Jenkins proposed.

At this movement, let me make an assumption because if I do not make an assumption, then you will be completely mislead. So, I told you that when we studied the classical time series analysis, then there are some components in the time series variable and seasonality is one of them and we have discussed how to deduct the seasonal component from the data. Now what I am saying here, if you want to apply the models that I am going to show you today, you have to make sure that your variable does not have a seasonality component.

So, what I am trying to say that the kind of analysis I am going to show you here, it is better if you apply on annual time series data where you will not find seasonal component, but if you actually have a seasonal data like if you have monthly observations or quarterly observations, then actually this ARIMA model that I am going to show you, it is going to be even more complicated. So, we are not going to cover that in today's lecture, but there is a way out.

Although I am not showing you how to handle seasonality in this ARIMA framework because we do not have enough time, but still you can apply it judiciously, you have to make adjustment. What adjustment? I have already taught you have to take out the seasonal component from the data.

And basically you can adjust the data for season. So, you can actually apply this tools that I am going to discuss you with now on a seasonally adjusted data so that the seasonal component that is embedded in the time series variable, it comes to 0 or if it is not 0, then the role of seasonality component drops down significantly so that you, it is negligible.

So, the crucks is that even if you have monthly or quarterly data, you first deseasonalize the data or you get seasonally adjusted time series data and then you apply these tools that I am going to show you next. So, a non-seasonal time series variable can be modelled as a ARIMA p d q stochastic process where parameters p, d and q have their own meanings, they are nom-negative integers, all of them.

And now let us look at them individually what they are. So, p is the order number of time lags of the autoregressive model, d is the degree of differencing. So, number of times the data had past values subtracted and then the last one is q which is the order of moving average model.

Now I am showing you 2 very simple examples of ARIMA models and I am going to change the values of the parameters p, d, q so that you can see different models will be generated by assuming different values of p, d and q. So, I am first going to show you an example of ARIMA 2, 0, 1.

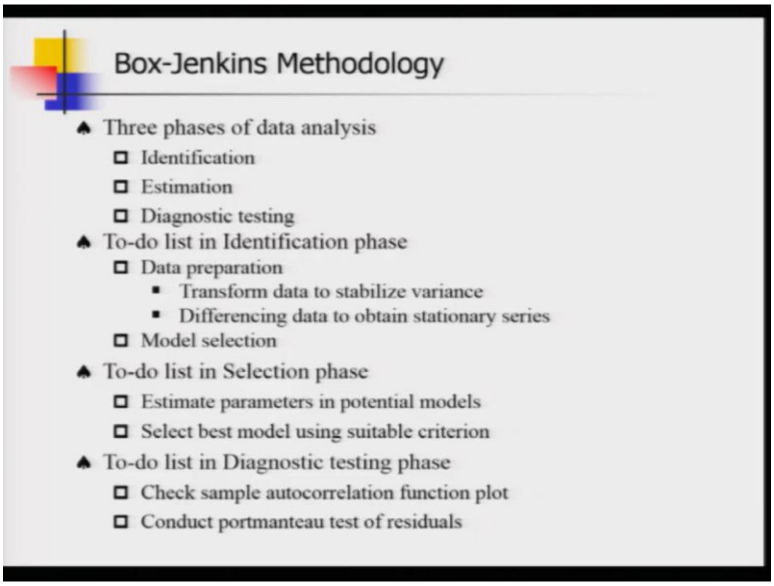
So, here p is equal to 2, d is equal to 0 and q is equal to 1. So, p is the order of autoregressive model and that is 2. So, that means that if y is the variable in question that we are trying to model, then we have to take two past values of the variable and they are yt minus 1 and yt minus 2. We are using 2 lags because p value is equal to 2.

And then 0 means that we do not have to difference it so there is no difference term here and the q term is 1, so you have to have a moving average model where you also use 1 period past value of the stochastic random error term. So, you have epsilon plus b1 times epsilon t minus 1 in the model.

So, now you see that I am using the difference notation here. Why? Because here the d is taking value 1 so that means that you have to make the first order difference. If you have forgotten how to take difference, I am showing you here at the bottom of the slide how you can take difference and then there is delta operator.

So, you are now using this delta operator to write your equation and the equation is simply delta yt equals to a1 times delta yt minus 1 plus epsilon t minus 1. So, if you expand this, you will see that I am making use of 1 time period lag of the y variable that is in question.

(Refer Slide Time: 24:02)



So, I am to now talk about the 3 phases of data analysis using Box-Jenkins Methodology. So, these 3 phases are identification, estimation and diagnostic testing. Now, here I am presenting a to-do list for each of these phases. So, first we are going to look at the to-do list for identification phase and here we have 2 major analysis to conduct, one is data preparation and the other one is model selection. Now, what is data preparation? So, I told you in the previous lecture that if you suspect that there is non-stationarity in the data, then you can take care of that by doing some data manipulation.

And first you can take logarithm transformation and then you can take first difference and most likely you are going to see a stationary time series after that. But whether you have actually got a stationary time series or not, that of course, is subject to hypothesis testing or statistical analysis. But this is the zeroth step you can say before you actually start the Box-Jenkins time series analysis.

Then comes the model selection. Now, when you have a stationary time series data after taking rounds of differencing and maybe taking log at the very beginning, then what to do? So, you have to now fit a ARIMA p d q model. Now to fit a model, you have to first understand how many parameters will be there in your model because here it is very different from what we have done in the past in this course. So, in the past when we were dealing with regression, we used to have some explanatory variables and they were part of the regression and it varies like if you have data on 7 variables, you take 7 variables to explain the variation in y.

But here in time series, this is univariate time series case, Box-Jenkins methodology, there they are not going to make use of any X values that is outside the system, they are going to make use of the past values of the same variable and the past values of the errors. So, that is basically their philosophy. Now, this p, d, q are 3 different parameters of ARIMA model and they are very crucial. What are the saying?

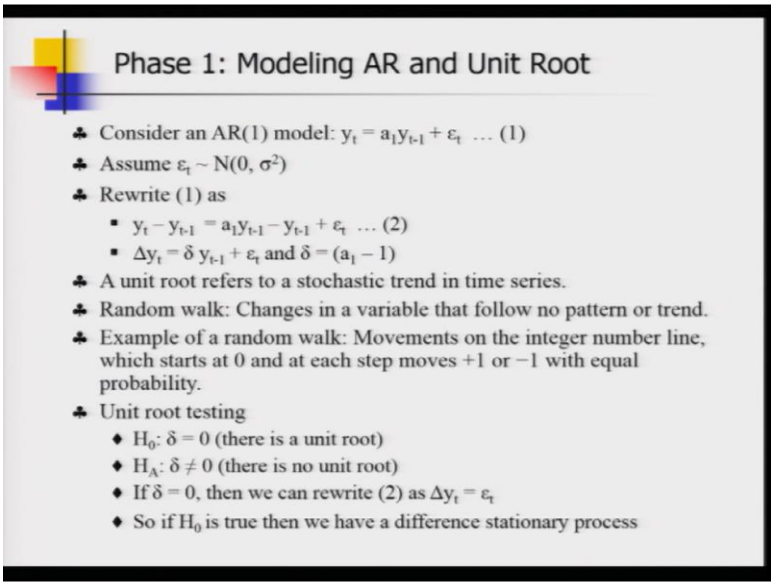
They are saying how many past values of the variable y you are going to choose. So, that is basically the selection for p. How many past values of the stochastic disturbance term you are going to keep in your model, that is basically answering the q parameter question. What will be the value of q?

And d of course, we know after how many differencing we land up what is called a stationary process. So, you know the value of d because that is how you actually your stationary time series data. Now you have to identify what are the possible values for p and q. And not only that, you also have to decide whether you are going to apply only AR component of the ARIMA model or you are going to apply MA component of the ARIMA model or you are going to have a mixture of both AR and MA components.

So, in the second part or second phase, Box-Jenkins suggest that once you identify your model, identify means that you have figured out the values for p and q and the nature of the model is known, now you go for model estimation. There is another step in the second phase and after you run alternative models, you have to choose the best suitable model by following some statistical criterion.

And in the third and last phase, you have to conduct some residual analysis just to figure out whether the residual generated from your ARIMA model is a pure white noise or not. So, now in the next lecture we are going to talk in detail about phase 1.

(Refer Slide Time: 28:39)



So, we start with a simple AR(1) model and it is given in equation 1 yt equals to a1 tis delta yt minus 1 plus epsilon t and now we are going to assume normality for the stochastic disturbance term so that we can say our random error is white noise and now you rewrite equation 1 as you are seeing here.

So, basically I have now used the delta operator to rewrite the expression and now you see I have a revised equation or rewritten equation which says delta yt equals to this another delta sign that is a different coefficient, it is not that triangle shaped delta. And this delta is a kind of a coefficient associated, it is not an operator like the previous triangle shaped delta. So, this delta time yt minus 1 plus epsilon t is my model.

Now, note that this delta is basically a1 minus 1. Now, we say that there is a unit root problem in non-stationary time series data and a unit root refers to stochastic trend in the time series. Now what is unit root? So, here if delta takes value 0, it implies that if a1 takes value equals to 1, then we have this unit root problem. Why it is called unit root, because note that if I concentrate on this equation 1 again now a1 is basically root from the characteristic equation corresponding the variable yt and that is why this a1 if it takes value equals to 1, it is called unit root.

Now we are going to talk about another term which you will come across often in time series literature and that is related with this unit root concept and the term is known as random walk. What is a random walk? It is a stochastic process which does not show any trend over a long time period. So, an example of a random walk could be movements on the integer number line which starts at say 0 or some positive number 1 or 2 and at each step moves plus 1 or minus 1 with equal probability.

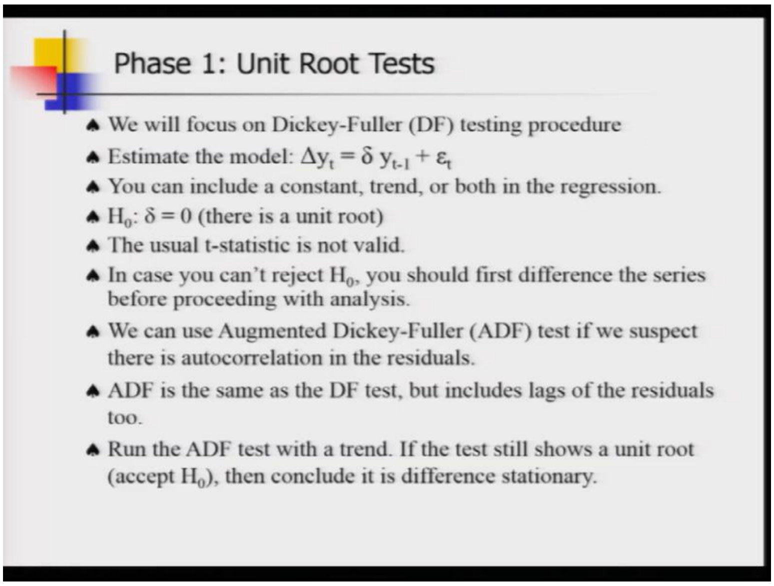
So, you see if you start at 0, then there will be a constant fluctuation of plus 1 minus 1, plus 1 minus 1 around 0. So, it is not showing any patterns. So, this is a simple example of a random walk. Now why random walk and this unit root unit are related because if unit root actually exists, then what will be the revised form of this equation 1 let us our look at the slide again. So, now you concentrate our equation 1 here and you see if there is unit root, then a1 will be equal to 1 so, we will have yt equals to yt minus 1 plus epsilon t. So, this is the exact situation that I know I was talking about through that example of random walk.

Now we are going to talk about unit root testing because we have to see whether our data has unit root or not. If it has unit root, then that means that there is a stochastic trend and we have to get rid of that stochastic trend and before we take the necessary actions, first we have to know whether the level values that we have actually observed in real life, whether there is some unit root or not and there are several unit roots available in time-series econometrics literature, we do not have time to go through all of them.

Here we are going to talk about 2 popular once and they are called Dickey Fuller test and augmented Dickey Fuller test. So, here in a nutshell, a unit root actually is talking about H naught is delta is equal to 0. So, there is a unit root and of course you have to now set an alternative hypothesis which says that data is not equal to 0. So, the interpretation is there is no unit root. So, if data is equal to 0, then actually we can rewrite the equation 2 as delta yt equals to epsilon t.

So, what do I means by this expression here? So, that means that if you take first difference, if your data has unit root, then actually after first differencing you are left with pure stochastic white noise. So, that is a very good news for you. So, if the null hypothesis indeed true, then actually we have a difference stationary process. So, if I now take a difference, then I will get a white noise which is stationary. Now we are going to talk about difference steps that are associated with Dickey Fuller test and augmented Dickey Fuller test.

(Refer Slide Time: 34:07)



So, first you estimate the model delta yt is equal to delta yt minus 1 plus epsilon t and then it is your choice you can include the constant, a trend here of course, I mean deterministic trend line say some constant time variable or both in the regression and I have already spoken about the null hypothesis and you then write your write statistics but it is not going to be the usual test statistic that we have seen in the case of cross sectional data based regressions.

Now, here is the problem, most of the time series tests you are going to see that they are very non-traditional. So, they do not follow traditional distributions. So, I am talking about the test statistics, once you set your null and alternative hypothesis you have to now write down your test statistic. Now the test statistic in the previous cases that we have come across follows nice trendily distributions like f, chi square, etcetera.

But here they are not going to follow t distribution or f distribution or z distribution. So, here they are actually going to follow very special distributions and the proponents of this test Dickey and Fuller they have come up with the critical values for the Dickey Fuller test statistic, but of course, in any econometrics textbook you are not going to get the full length table for Dickey Fuller critical values. So, softwares these day interpolate numbers and then show you on screen that what could be the interpolated critical values for Dickey Fuller test statistic.

But you do not have to worry about all this, you can follow the traditional approach, not traditional approach because here you are not going to match the critical value and the tabulated value, you can follow the p value approach and if the p value is less than the set level of significance say 0.05, you can always reject the null hypothesis. Now if you cannot reject the null hypothesis, then you have unit root in your model. So, you should now first difference the series before proceeding with the analysis.

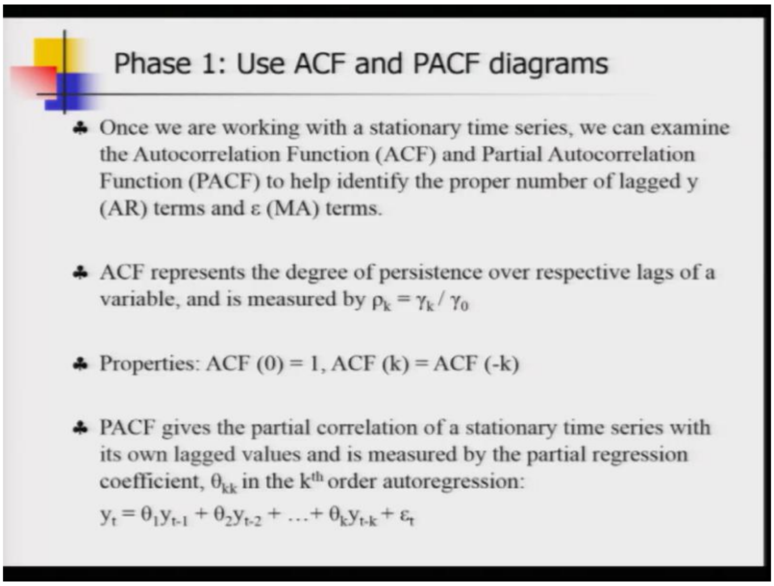
You can also use augmented Dickey Fuller test if you suspect that there is autocorrelation in the residuals. So, of course, when we conducted the Dickey Fuller test, we are not assuming that there is a problem of autocorrelation, but in most time series you can suspect that there could be some degree of autocorrelation problem. So, it is better to actually apply augmented Dickey Fuller because the test is readily available and you do not have to compute anything, software will compute for you, you just have to take a decision.

So, ADF is also similar like the DF test, but it only includes some extra variables and these are the lags of the residuals as explanatory variables in the regression. Now run the ADF test with a deterministic trend if possible if you want, then if the test shows unit root, then you conclude that it is difference stationary. Now suppose that you have come to know that there is non-stationarity problem, unit root exists, so you take difference and you get a stationary dataset now.

So, when you finally have a time series variable which is stationary, now the next step is to choose for the values of p and q because d is already known to you because you have taken differences, so your data maybe stationary after taking first difference or you may have to take difference once more. So, after second difference it will be stationary, but if you have stationary data, then you already know the value of d, you have to know choose for the values for p and q.

Now there are graphical methods to help you to choose the values for p and q. You do not have to arbitrarily choose p and q values and here we are going to introduce to new concepts and they are called autocorrelation function and partial autocorrelation function.

(Refer Slide Time: 38:37)



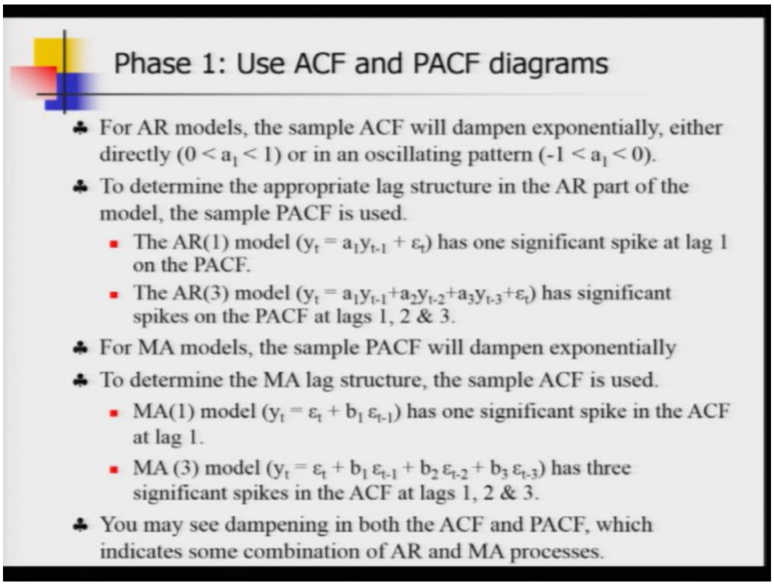
So, we examine autocorrelation function and partial autocorrelation function to identify the proper number of lagged y terms and lagged epsilon terms for the AR model and MA model respectively. So, what is autocorrelation function? An ACF represents the degree of persistence over respective lags of a variable and it is measured by rho k equals to gamma k divided by gamma naught. Now rho k I have already introduced in the previous lecture, so this is known to you and gamma k is basically the k lag covariance between two time series variables and that gamma naught is basically the variance term.

So, here I do not have enough time to talk about the properties of autocorrelation functions, but I am just listing down two very important autocorrelation function properties. So, you see that at zeroth lag, then ACF will take value 1 and ACF at kth lag and ACF at minus kth lag they have same values. Now there is another concept called partial autocorrelation function, so a PACF gives the partial correlation of stationary time series with its own lagged values and it is measured by the partial regression coefficient.

So, now let us have a look at these autoregressive problem where yt is being regressed on its previous values and we are going in the past by k time period so we are now using the past values like yt minus 1 yt minus 2 yt minus k. So, now you this theta 1 theta 2 theta k coefficients are talking about the partial correlation between the current value of the variable yt and the past values of the same variable. So, theta 1 actually talks about the correlation between yt and one period past value of the same variable y.

So, now you can run this kth order autoregression model and see how many of these thetas are significant or not. So, now this choice of p and q is very lengthy discussion to cut it short, I am going to adopt a cookbook approach and I am going to show you the basic results, some fundamental rules that you must remember when you are looking at a PACF, ACF diagram and choosing the values for p and q.

(Refer Slide Time: 41:20)



So, for autoregressive models, the sample ACF will dampen exponentially either directly for the positive values of a1 or in an oscillating manner. So, if you follow that there is a exponential damping from the ACF graph for the stationary time series, then you can say that probably an AR model fits the data better. And then you have to determine the appropriate lag structure in the AR part of the model. And for that you have to use the sample PACF. Why sample PACF and why sample ACF?

Actually the first step also when you are looking at the ACF, it has to be based on the sample because of course, you do not have the value for population rho, so you have to get it estimated by sample values. So, basically what you have to do? You have to get the residuals from the OLS regression and then you have to lag them and then you have to compute the Pearson correlation coefficient between these lagged values and the current values of the residuals and that is the way you will get the different values of autocorrelation coefficient at different lags. So, this is called sample autocorrelation.

So, similarly you can also run the regression based on the fitted residuals and you will have sample partial autocorrelation function. So, I am going to talk about 2 examples here, how you can choose the orders for AR. So, in the first sub bullet point under second bullet point I am showing the case for AR(1) model. So, AR(1) model is written here, so if the data actually follows AR(1) model, then there will be one significant spike at lag 1 of the sample

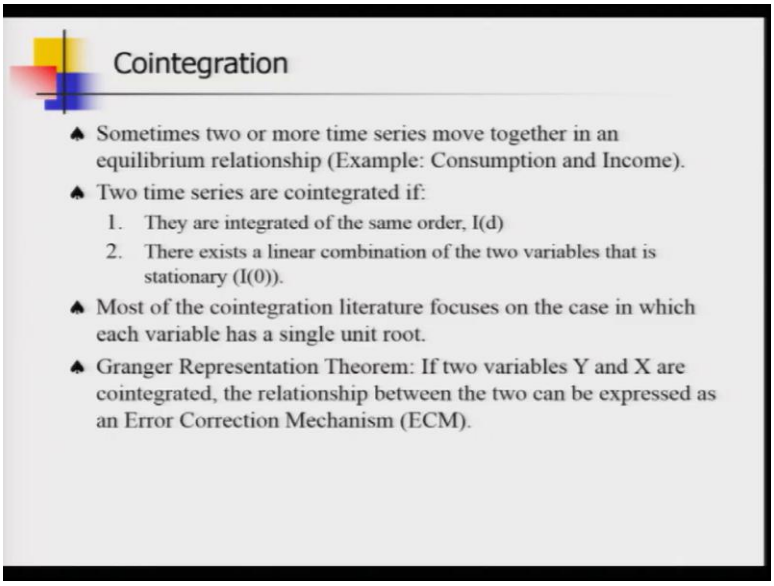
PACF. And if there are 3 such spikes on the sample PACF diagram, then actually you shall pick the value p equal to 3 and you will frame and AR(3) model.

Now, for moving average model, the story is somewhat different. So, here whether the moving average model works better for the data or not, for that we have to sample PACF diagram and if we say that there is an exponential decay in that sample PACF diagram, then actually we can tell us that most likely our data is going to follow moving average process. And now how to choose the value of q to fit a moving average model? For that now you have to look at the sample ACF diagram.

So, if you see that there is only one significant spike in the sample ACF diagram at lag 1, then you should model it like MA 1 model and if you see that there are 3 significant spikes in the sample ACF diagram at lags 1, 2, 3, then you write an MA 3 model and I am showing you how to write the MA 3 model here. Now, if you see that there is this damping exponential damping behavior in both the ACF and PACF diagrams, then you can actually think that well maybe a mixture of AR and MA models would work best in this case and you can then have a proper ARIMA model.

Now, in the second step, you go to software and ask it to estimate the model for you, it is complicated, I am not going to get into details of that and once the model is estimated, you can generate the residuals and in the third step of the Box-Jenkins methodology, now you should also check for the white noisness of the generated residuals so that you can be sure that you do not have any non-stationarity issue left in the analysis. And for that there are many tests available, but we do not have time, so we are not going there, but remember that after you get the residuals from ARIMA, you should check for the white noisness of the generated residuals.

(Refer Slide Time: 45:33)



Now, in the last 10 minutes of the lecture, I am going to talk about the other schooling time series analysis which is mostly based in United Kingdom and led by famous time series econometricians like Clive Granger who was also Nobel Laureate for his fundamental and tremendous contribution in the field of time series analysis. So, here we are going to talk about 2 major concepts very briefly and they are called cointegration and error correction model.

Now, before we go to the slides and discuss the bookish definition and all, let us talk about what is the philosophy behind this error correction model and cointegration. So, Clive Granger opined that there are many economic variables for which you can think of a long run equilibrium. So, these 2 or more variables could be like linked over long run and there could be an equilibrium relation between or among the variables.

So, one example could be the consumption and the income data. Of course, it is not at all a bad idea to assume that there is a long run equilibrium between households income and households consumption expenditure and if you have these kind of long run equilibrium in the data, then how do you model it.

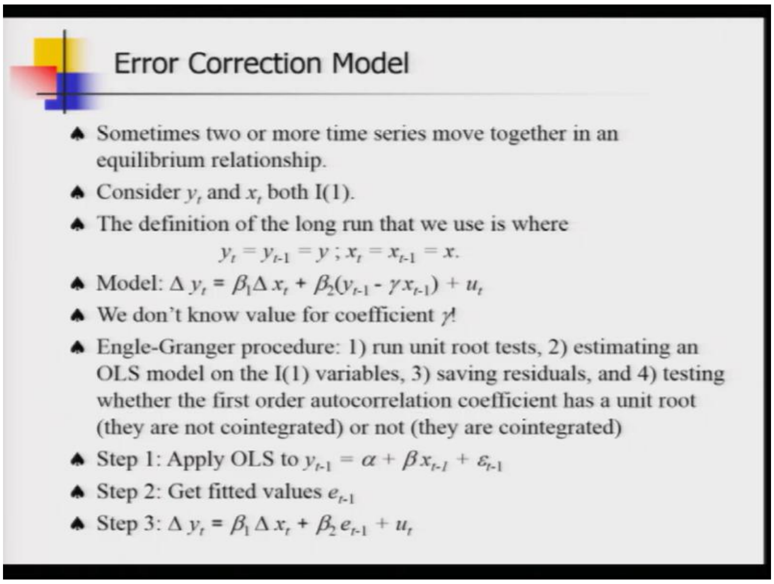
So, then Granger and others proposed this error correction model. So, they are of the opinion that if there is a long term relationship between two variable say Y and X, if there is a short term fluctuation by some shock, then there is a tendency that within the time series system, there is a dynamics which will actually now help that way of point the come back to the natural trend line if there is any. So, that correction is called error correction because as you have moved away from the long run equilibrium relation between two variables due to some shock, not that is an error and there will be some correction corrective nature taken over time and you will again come back or come home to the long run relationship between the variables. So, that is basically known as the error correction model.

So, now Granger defined two time series are cointegrated if they are integrated of the same order. So, the d value will be the same for both variables and there exists a linear combination of these two variables that is stationary. So, although you are working with non-stationary variables X and Y and they are integrated of order d, but if you take a linear combination of these two variables X and Y, then the resulting variable will be a stationary variable. It implies you do not have to take difference to make it stationary. Or in other words, there is no stochastic trend.

Now most of the cointegration literature focuses on the case in which each variable has a single unit root and that is what we are going to follow. So, now we are going to briefly mention a very important result knows as Granger Representation Theorem. It says that if two variables Y and X are cointegrated, then the relationship between the two can be expressed as an error correction mechanism.

We have only 5 or 6 minutes left in today's lecture, so I am not going to get into deeper details of error correction model, but I am just going to show you some steps that you must follow if you are interested to conduct an error correction model analysis.

(Refer Slide Time: 49:32)



So, suppose you have 2 variables Y and X and you have time series data yt and xt and both are say I1 processes, so if you take first difference of both variables, then you will have stationary data, but the problem is that if you take difference, then this long term relationship, that long run equilibrium will disappear from the picture all together. How, that we are going to discuss the next.

So, the definition of the long run that we use in error correction model idea is represented by this two conditions, yt equals to yt minus 1 equals to some common y and xt equals to xt minus 1 equals to some common x. So, you see that if you now take the difference, so xt xt minus 1, so then there will be 0. So, this long term relationship or this common equilibrium is gone. So, basically this error correction model is proposed that I am showing here in the slide. So, it says delta yt is equal to beta 1 times delta xt plus beta 2 times some error correction mechanism that is yt minus 1 minus gamma times xt minus 1 plus ut.

So, this thing in within bracket that is called the error correction component. But the problem is that we do not know the value of gamma, it is an unknown parameter, so what to do? How can I run a regression model like this where I have this gamma parameter in this error correction component?

So, Engle and Granger, they have suggested a 4 step procedure and in step 1 you have to run the unit root test. In the second step you have to estimate this regression equation on I1 variables via OLS method. So, I1 variable means that you are running OLS on the level values of the variables.

And in step 3 you have to save the residuals and in step 4, you have to test whether the first order correlation has a unit root or not. Now you note down the steps again. So, in step 1 apply OLS to the model yt minus 1 equals to alpha plus beta xt minus 1 plus epsilon t minus 1. Then basically from this OLS you can generate the estimated coefficient values for alpha and beta and once you have that you can generated the fitted values for the random term also which is epsilon t minus 1 hat or et minus 1 hat.

So, in step 2, now you get those fitted values et minus 1 and then in step 3, you basically replace this yt minus 1 minus gamma times xt minus 1 component by the fitted lagged residual value which is et minus 1 that you obtain from step 1 and network you have a redefined regression problem where you regress delta yt on beta 1 times delta xt plus beta 2 times et minus 1 plus ut.

Now, you note that as et minus 1 is observed, then that problem of unknown gamma has disappeared. So, you can actually apply OSL on this redefined regression model and estimate an error correction model. So, this is it on time series econometric. I know that probably I rushed through a couple of concepts, but time series econometrics is an ocean, really I mean it.

And it is very difficult to give you a very good idea about an ocean in 50-55 minutes time. I tried to give you an idea about some basic concepts and the popular terms and if you are interested, feel free to consult other textbooks in econometrics where you will find detailed analysis on these issues. So, this is it for the time being. Thank you, bye.