Econometric Modelling

Prof. Rudra P. Pradhan
Vinod Gupta School of Management
Indian Institute of Technology, Kharagpur

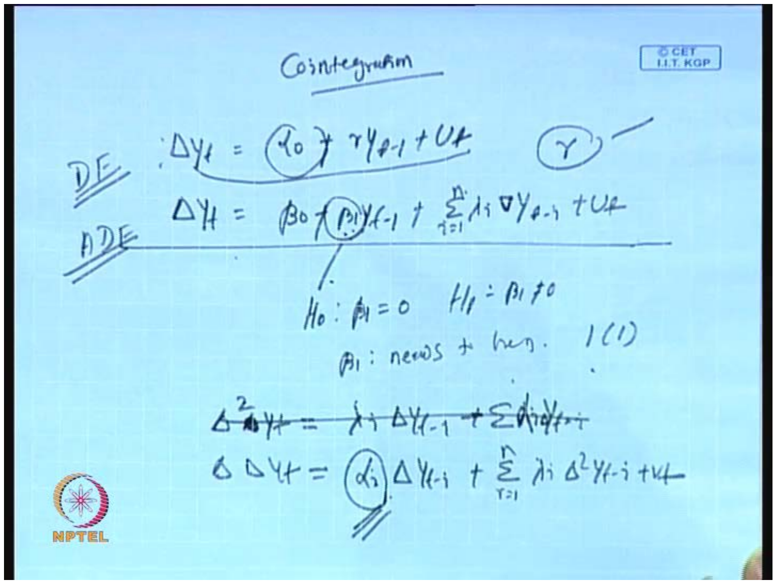
Lecture No. # 39
Cointegration

Good evening, this is doctor Pradhan here. Welcome to NPTEL project on econometric modeling. So, today we will discuss cointegration technique, it is one of the interesting techniques under time series modeling. So, in the last lectures we have highlighted the basic framework of unit root problem because without knowing unit root or you know, stationary issue, you cannot start cointegration.

So, unit root, you, you must know the nature of the variable first, then accordingly you have to apply the cointegration. So, so first I will briefly highlight what is exactly unit root problem? Then, we will come down to cointegrations; then also we will discuss the, in same times, the causality issue. So, now, what is first of all the unit root problem? So, unit root problem objective is to know the order of integrations, where the variable will be stationary in nature.; Is it order of zero, is it order of one or is it order of two and so on, so that, the variables can be, get to know whether it is stationary in nature or non-stationary in nature. So, depending upon the order of integration, where the variables are stationary in nature, we can apply the cointegration and causality technique.

So, without knowing the unit root problems or the stationary issue it is very, very difficult or it is not possible to enter to the cointegration and causality test. So, as I have mentioned in the last lectures, there are many techniques available to check the stationary problem, so that unit root problem. So, we start with Dickey-Fuller test, then we, we discuss with augmented Dickey-Fuller test, KPSS, Phillips-Perron test, NG test and so many things are there. So, basically, in the last lecture we have highlighted the concept of Dickey-Fuller test and augmented Dickey-Fuller test. So, I just briefly bring that, these two two techniques little bit, then we will enter to the cointegration techniques.

(Refer Slide Time: 02:55)



So, Dickey-Fuller test, the standard form of Dickey-Fuller test is delta Y t equal to alpha 0 plus gamma Y t minus 1 plus U t. So, here alpha is need to be significant. If alpha is significant, then the variable is stationary, attend stationary or it does not attend the stationary. So, this is standard Dickey-Fuller test. So, corresponding to Dickey-Fuller test, the augmented Dickey-Fuller test statistic will be written like this, delta Y t equal to beta 0 plus beta 1 Y t minus 1 plus summation, you can say, lambda I delta, delta, delta Y t minus I plus U t, I equal to 1 to n. So, this is, this is the standard tricks of, this is ADF test equation, this is Dickey-Fuller test equation.

So, by the way, there are three forms of mathematical equation are available with Dickey-Fuller test and augmented Dickey-Fuller test. So, you know, we start with the delta Y t equal to lambda Y t minus 1 plus U t. Then, this is the starting point means if we will remove this one, then the starting point is this, this and this. So, then, this is the starting point and with respect to this particular model, so we can add drift component, we can add drain component and we can add both the components.

So, I am not highlighting all these details here, because we have already discussed all these issues in the, in the last lectures. So, here, if we will get the, means, in this particular equation, augmented Dickey-Fuller equation, because it is more advanced than Dickey-Fuller, a simple Dickey-Fuller test. So, beta 1 is the important, you know, parameter. By the way, beta 1 should be statistical significant to know the stationary levels or order of integration through which you can say, the variable is stationary in nature.

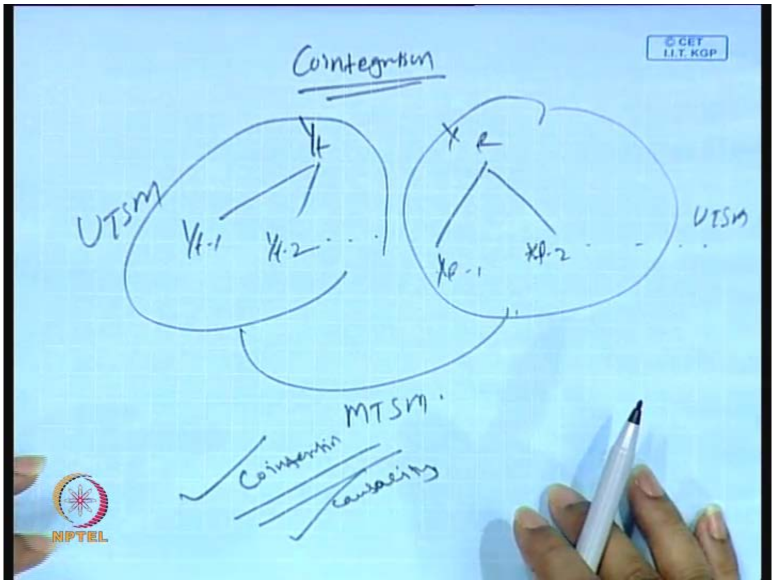
Suppose, suppose this is integrated, this is a beta, here alpha 0, so means, null hypothesis is, that beta 1 should be equal to 0 against beta, beta 1, sorry, beta 1 not equal to 0. So, that means, beta 1 needs to be significant; beta 1 needs to be significant. If it is significant, then we will call it is, it is integrated of order 1; here is integrated or order 1.

Suppose it is, it is not stationary at this level, then you go for second difference, it is called as del square, del, del square Y t. It means, we can write like this way, del, del Y t or delta Y square delta square Y t, which is equal to lambda I delta Y t minus I plus summation, summation, lambda I, it is lambda is already there here, so then you put it here, this is alpha I Y t minus I. Of course, delta is, there is, once again I am writing here, delta Y t equal to alpha I delta Y t minus I plus, you know, lambda I delta square Y t minus i equal to 1 to n plus v t.

Now, again this is very important here is lambda. In fact, this is, this is a, this particular item should be statistically significant; this should be statistically significant. So, depending upon the significance levels, so we can check whether the variable is stationary in nature or non-stationary in nature. So that means, first you start with the simple augmented Dickey-Fuller test, then you see whether the variable is stationary, attends stationary or not?

So, now, your objective is to know at what level the variable will be stationary, is it at the level data, is it at the first difference, is it at the second difference, is it at the kth difference? If it is, if it is stationary at the level data, then it is integrated order 0. If it is attends stationary at the first difference level, then you call it is a, its order, order of, order of integration is 1. Similarly, if it is, you know, attends stationary at the kth, kth difference levels, then you can say, that it is integrated of order k. So, that is, that is you know, a requirement, one of the most important requirements for cointegration technique. Until and unless you know the stationary issue, then you cannot move in to the cointegration techniques.

(Refer Slide Time: 08:17)



So, now, you come down to the co integration structures. So, what is called together the cointegration? You see, the basic definition for integration is that, so we like to know, what is the long run association between variables? So, long run association between variables, you remember one thing very carefully, in the time series modeling your sample size should be, should be very, very, very high, otherwise it is very difficult to handle time series models where the lag is in the, you know, with you.

So, the, once you are using lag, then obviously, the sample size should be substantially high. The moment sample size is substantially high, then it is the question of long run analysis. So, when there is long run analysis and when you are handling more variables at a time, then you know what we have discussed in the case of simultaneous equation modeling or structural equation modeling. So, in a particular system, once number of variables is very high or large in numbers, then obviously, there may be question of interdependence. So, the question, existence of interdependence in the time series setting is called as cointegration.

So, we like to know what is the long run association between these two variables is there. Are they convergent or are they divergent? So, that means the cointegration will become in to pictures. So, cointegration represents the existence of long run equilibrium relationship between two or more variables.

So, let us we start with two variables at a time. That means, cointegration technique is a multivariate technique, so with one variable cointegration cannot be possible. So, we start with two, two at least two two variables, like you know, correlation or regressions. So, cointegration also starts with the, you can say two two variables together. So, it is not like that, like you know, in the time series modeling. So, we have, Y t, Y t minus 1 Y t minus 2. So, with one, that is univariate time series modeling with single variables, you will create additional variables.

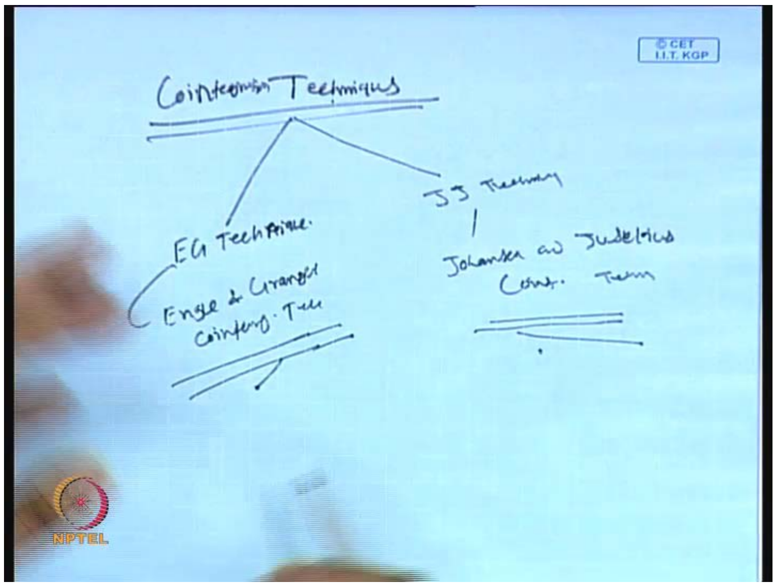
So, that is not the structure of cointegration. In the cointegration textures there should be two different variables, let us say, Y t and X t, like this; so, Y t and X t. So, we need to create here is Y t minus 1 Y t minus 2, that is one structure continued. Then, it is X t minus 1 X t minus 2, X t minus 2, it will continue, so this is one type of situation. If this game boundary is this side then it is called as a univariate time series modeling and this is, this is called as a univariate time series modeling. But when we will integrate these two, this is called as a multivariate time series modeling. So, now, there are lots of techniques under multivariate time series modeling. So, one is called as cointegrations, so another is called as a causality, that, that causality, causality.

So, our main, today's discussion is with respect to this, cointegrations in causality. So, what is, what is all about the cointegration, how to detect it and what, how, why it is there and what is its relevance? Then, we will come down to causality. So, cointegration is the middle way of this time series setting. So, the initial setup is unit root, then end setup is causality, in between cointegration will play a middle role. So, it is just like a middle order batsman.

So, now we have to see how cointegration will give you signal for the causality. So, it will give you, it will give you, means, it will indicate, that there is a long run association between these two variables. That means, if there is cointegration, of course by definition we can say, that there is an association between these two variables, that too in the long run and if it is, it would give you only the long run association between these two variables. But it will not give you the indication of direction of causality because time series modeling's one of the major objective is to detect the direction of causality, which is a special feature of time series modeling, which is not available in the case of cross-sectional setting where the causality is one way here we have two way causality there is possibility of two way causality.

So, that is why, because it is all about time series setting, so we first know what is the cointegration (( )). So, cointegration means, the existence of long run, long run equilibrium relationship between two or more time series variables. So, that is very important time series variables.

(Refer Slide Time: 13:04)

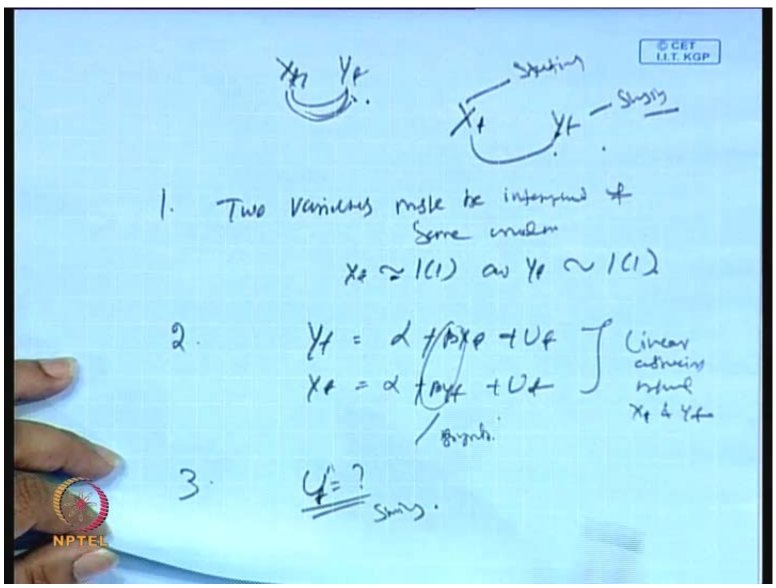


So, now, like unit root, like unit root techniques are many techniques are there, cointegration techniques cointegration, cointegration techniques like, like you know, like unit root cointegration techniques are also multiple in nature, but there are two standard forms of cointegration technique, one is called as EG test, EG technique, technique, and another is called as JJ technique. So, this is called as Engle-Granger cointegration technique and this is called as Johnson and Juselius, Juselius cointegration technique.

This is much simpler, easy to understand, easy to calculate; this is too complex, difficult to calculate and it is too much time taking. It is very easy to handle, this cannot be easy to handle, with, means, easy to handle means, it is, manually it is very easy to, it is not so easy to handle when the setup is very complicated, but it is more advanced, more accurate, more reliable and you know, more useful, in fact. So, we start with the process of easy technique, then, then you will converge to Johnson technique.

So, Johnson technique, generally when you will use software, it is better to use Johnson technique rather than EG test. And the problem, if you know, just like you know, in the optimization problem we have graphic techniques and (( )) every techniques, so here also same thing. So, if we, the variables problem setup is with respect to two variables only, then simply we can apply the Engle-Granger test. In fact, it can be also Johnson-Granger test. In fact, you can also apply Johnson-Juselius (( )) test. But, Engle and Granger test is very much useful for the two variable setting.

(Refer Slide Time: 15:30)



So, how, what is that test procedure here is, the one of the condition for this set test is, that there must be two variables, X t and Y t. This, you know, this Engle and Grangers defined the possibility of, you know, convergence between these two. So, that convergence is called as a cointegration. They, they are very, they are, you know, they got Nobel prize for this particular techniques and they are very, they are very intelligent and basically, they are statistician, they have developed a very beautiful setting in the time series problems to know the existence of long run equilibrium relationship between two variables.

So, basically it is three standard steps. First step is, let us say, there are two variables X t and Y t. So, now, if we have X t and Y t, so then obviously, so, you like to know X, is there any relationship between X and Y t. So, that is our target. Obviously, you have to first find out the stationary issue here is, you also find out the stationary issue here is. So, this is our first job.

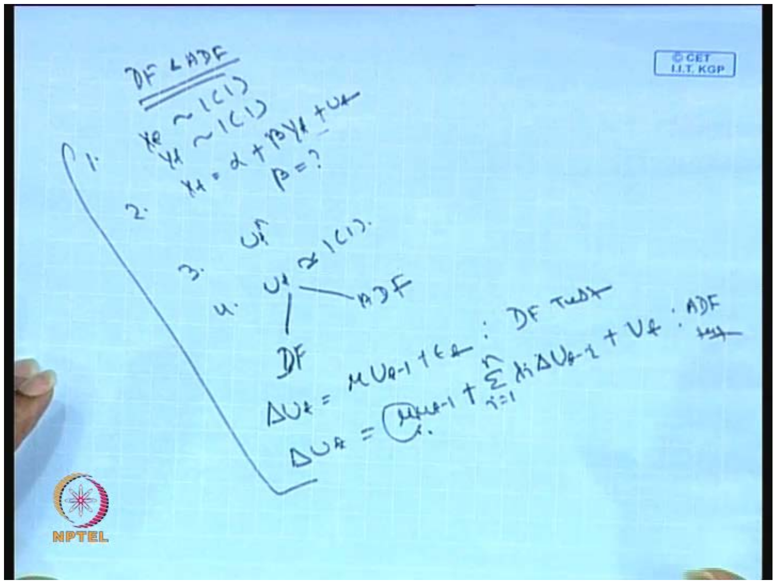
So, once you have the stationary issue, then come down to existence of cointegration. It will give you; once you know the order of integration, it will give you signal how you have to handle the cointegration technique. So, that means, it will give you a signal, that there may be possibility of long run equilibrium, so, but it will not give you positive signal, that there is a long run, long run relationship. So, for that you have to apply the cointegration techniques. So, this is a test, basically, give you, means, it, it, it consists of three steps altogether.

So, first step is two variables must be integrated of same orders. So, two variables, two variables must be integrated of, integrated of same order, preferably, preferably like this, X t X t integrated of order one and Y t integrated of order one. So, this is, this is the first condition of the cointegration technique, that too EG, under EG test. So, two variables must be integrated of same orders. So, let us assume, that the, that means, they are stationary at the, at the first difference levels, at the first difference level. So, now, X t is, X t is integrated of order one; Y t is integrated of order one. So, let us assume, that they are integrated of order one only.

So, then, you can come to next step. Only if they are not, if there is mismatch in the order of integration, then EG technique will not, EG help you to detect the cointegration means, existence of long run equilibrium relationship. So, what you have to do in the, once it is satisfied, then you come down to second step. Second step is to establish the existence, to establish the linear relationship between the two variables. That means, so Y t equal to alpha plus beta X t plus U t, then, or else X t equal to alpha plus beta Y t plus U t. So, there must be linear association, linear association between, between X t and Y t. So, the, so, we, we, we, task there is a linear relationship between these two and beta needs to be significant; beta needs to be significant. So, first condition is, that variables must be integrated of order one and second condition is, that, so there must be linear relationship between these two.

Then, third condition, once you have linear relationship, then you will find the error terms; you will find the estimated error terms. So, now, we have to move to the third condition, that error term should be stationary in nature. So, the error obtained error term must be further stationary in nature. So, obtain, obtain error term, must be stationary in nature and for that, so you again apply DF statistics or ADF statistics.

(Refer Slide Time: 19:42)

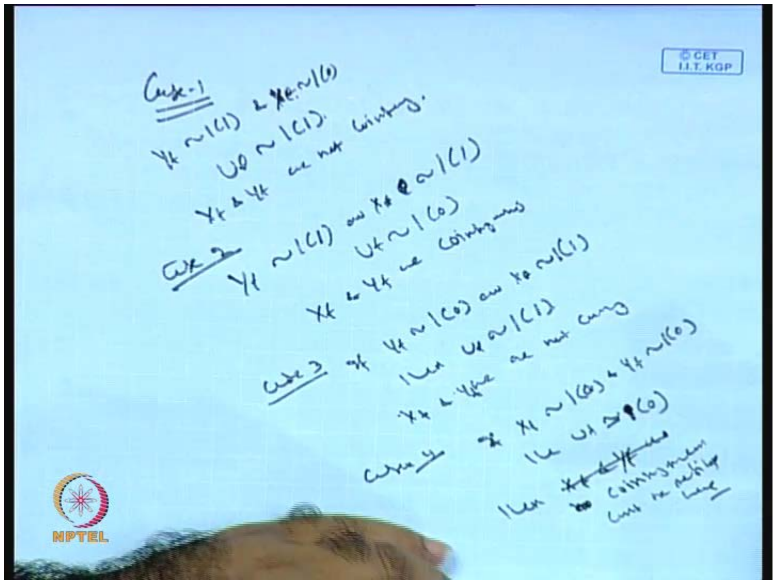


So, that means, you see, there are three condition altogether. First condition, X t, X t follows 1, 1; Y t follows 1, 1. Second, X t is a function of Y t, there is a linear relationship between X and Y, U t. Let us assume, that the linear relationship is here, like this beta should be significant and third condition, you had the U t hat. Then, fourth, U t hat should be stationary in nature, it should be integrated of order 1 or 0; it should be stationary in nature. So, for that you have to apply Dickey-Fuller test or augmented Dickey-Fuller test.

So, once you apply Dickey-Fuller test, then the statistic will be like this, delta Y t equal to, you can say, mu U t minus 1 plus epsilon t or this is DF, DF test, Dickey-Fuller test and, or you can apply delta Y t equal to mu U t minus 1 plus summation lambda i delta U t minus I, I equal to 1 to n plus, you can say, V t. This is ADF test, this is ADF test; this is ADF test.

So, now, of course, this has to be statistically significant, this has to be statistically significant, so that it will be stationary in nature. So, that means, these are all, you can say, procedural measures for checking the existence of cointegration between X t and Y t, provided X t and Y t are, are integrated of same order, that too here, 1, 1. So, that means, they are integrated of order, order one only. So, now, we will see if there are two variables and we are assuming that both the variables should be integrated of order one, if not, then what are the problems.

(Refer Slide Time: 22:17)



So, it will mean, if there are two variables, so as far as order of integration is concerned, then we may have four different situations. So, we may not, may, there is always four different possibilities; there are four different possibilities.

See, case one, the case one is, if you see Y t is integrated of order one and X t, X t is integrated of order 0, in that case, in that case U t will be integrated of order one; integrated of order one. And here, X t and Y t are not cointegrated, are not cointegrated. This is case one.

Case two, if you know Y t integrated of order 1 and X t integrated of, X t integrated of order 1 then, U t will be integrated of order 0. So, in that case, X t and Y t are cointegrated, are cointegrated.

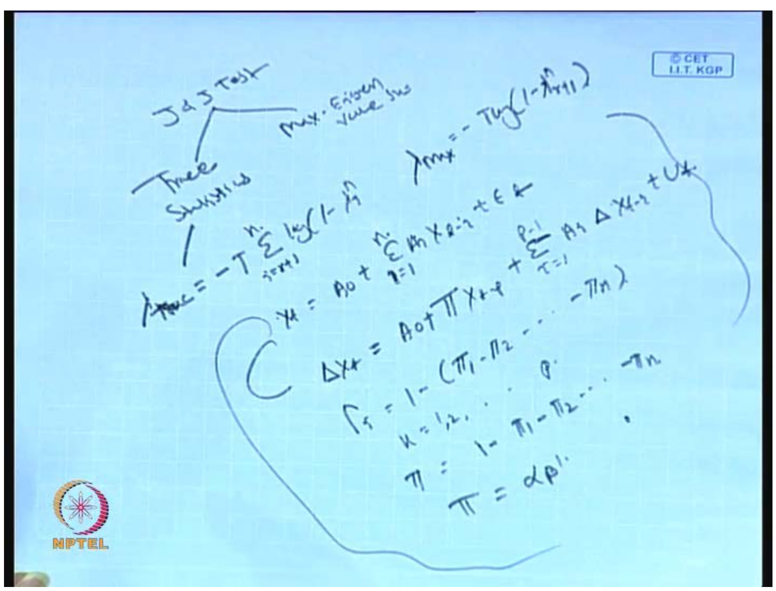
Case 3 case 3, if Y t integrated of 0, order 0 and X t integrated of order 1, 1, then U t will be integrated of order 1; integrated of order 1. In that case, X t and Y t are not cointegrated.

Case 4, case 4, if X t integrated of order 0 and Y t integrated of order 0, then U t must be integrated of order, integrated of order, order, order 0, sorry, it is integrated of order 0; of course, integrated of order 0. Then, X t and Y t, X t and Y t, X t and Y t are, are means, cointegration here. That means, here there is no such cointegrations, there is a cointegration cannot be possible. It is better to write cointegration, cointegration cannot be possible here, this is all about, this is all about the EG test.

So, Engle-Ganger test follows four different steps. So, that means, if there must be two variables in the systems, at least two variables in the systems and both must have same order of integrations and there must be linear relationship between the two, then obtained error term must be stationary in nature, that too by either applying ADF test and ADP test. There is also, you know, item called as a CRDW, cointegrating regression Durbin Watson statistics.

So, you know, we have discussed autocorrelation problem. So, same thing is here. So, we once, the order of integration are same, then we will go for cointegrating equations. So that means, they, they must have linear relationship. So, you know, in that case when we will estimate that cointegration equation, you will find there will be Durbin-Watson statistic. Again you check the, like autocorrelation, whether that particular, statistic, statistically significant or not. If it is, so then there is a cointegration, so otherwise it is not a cointegration. So, this is how you have to go for this, you know, checking the cointegration test.

(Refer Slide Time: 26:11)



Another advanced test is called as Johnson and Juselius, Juselius test, JJ test. Basically, this test deals with two statistics, one is called as trace statistics, trace statistics, statistics, another is called as maximum Eigen, Eigen value, Eigen value statistics. This is, basically, denoted as lambda max, sorry, lambda trace, which is equal to minus T summations log 1 minus lambda I hat i equal to r plus 1 to n.

Similarly, maximum Eigen value, we will write (( )) max equal to minus T log 1 minus lambda hat r plus 1, which is followed by the two equations, X t equal to A 0 plus summation a I X t minus I plus epsilon t I equal to 1 to n. This can be rewritten as delta X t is equal to A 0 A 0 plus, you know, product X t minus p plus summation I equal to 1 to p minus 1 p minus one A i, A i and delta X t minus i plus U t. So, while, while, while 1 minus lambda 1 pi 1 minus pi 2 minus pi n, K equal to 1 to up to p and pi equal to 1 minus pi 1 minus pi 2 minus pi n. So, obviously, this is nothing, but alpha plus beta coefficient; alpha plus beta coefficient. So, this is your, these are all the derivations.

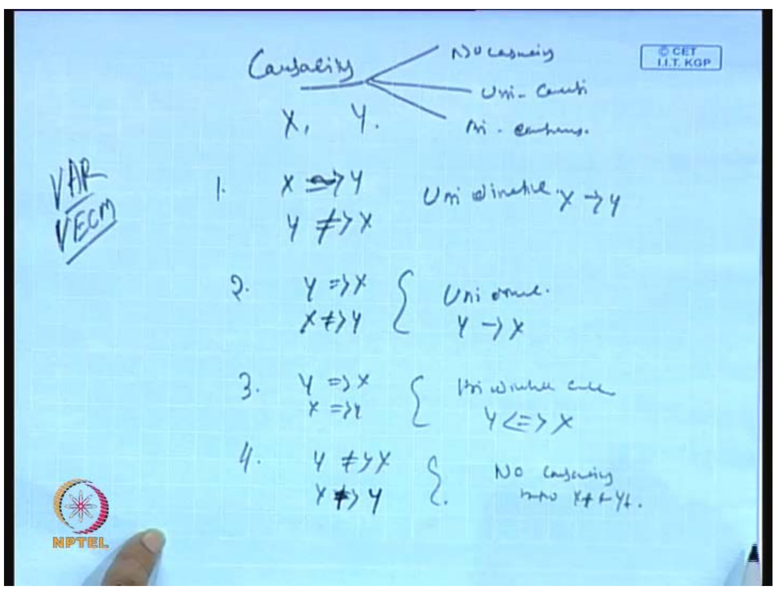
So, I am not going to derive all these details here because it will take lots of time. So, it is not possible to derive all these details with limited lectures. So, what I will suggest you, so for Johnson and Juselius test, so you have to just pick up two (( )) maximum Eigen values. So, we have a tabulated value, means, we have a critical value. So, with the basis of critical value you have to compare the calculated value of trace statistics, lambda trace and lambda max, maximum Eigen value statistics. Then, we like to know, it is actually matrix format. So, with the help of its Eigen value, Eigen matrix, so it will give you signal how many, how many cointegrating vectors is there, whether they are, there is any cointegration? That means, if the statistic is significant it will give you hint that there is cointegration.

So, now, how many cointegrations are there depending upon the number of variables involvement and after? Suppose, there are two variables involvement and then obviously, there is only one cointegration equations, two cointegration X t Y t Y t X t. Similarly, if there are multiple numbers, so you have to find out how many possibilities are cointegrating variables and accordingly, you have to justify what are the number of cointegrations? So, that will give you signal of you know, long run equilibrium relationship between two or more variables. So, that is nothing, but you can say cointegration technique. So, now, cointegration will give you the signal, whether, whether there is existence of long run equilibrium relationship between two, two or more variables.

So, now, once it will give you detection whether the variables are, you know, exist, whether there is existence of long run equilibrium relationship between the two variables or more. So, it will not give you the signal of the direction of causality, that means, whether all these variables are influencing single variables, so all these variables are interdependent to each other. So, that means, you see, I, in the very beginning I have mentioned, so these three statistics are very integrated to each other, that means, unit root problem, cointegration problem and causality problem.

So, unit root test will give you signal to go for cointegration, cointegration will give you signals in which green signal to the causality issue. Ultimately, we like to know, what is the direction of causality of this particular problem.

(Refer Slide Time: 31:05)



So, now, so, what we will do here is, so we will discuss the causality problems. So, now, the next problem we highlight is the causality problem. So, now, causality means, it will mean long run existence of cointegration will automatically give you a signal, that there is causality. But that that is not, that is essential or necessary condition, but it is not sufficient condition. For sufficient condition you have to go for causality test and this is how Granger is very famous. So, C. W. Granger has developed a test, according to his, according to his name it is known as a GC Test, Granger Causality test.

So, we have to go for Granger causality test to check whether there is a direction of causality. If it is so, whether it is in one, unidirectional causality or bidirectional causality, that means, there are, of, of course, when we will move to cointegration, then the system will be itself multivariate, it is not univariate, like you know in the case of unit, unit root test. So, the moment you are in the cointegration, then the system will be multivariate, then obviously, there is possibility of causality.

Then, when there is possibility of causality there are three different, you know, outcomes you may get. So, there may be no causality between the, between the two or more variables and there may be unidirectional causality and there may be bidirectional causality. So, let me highlight what is all about this causality issue. For instance, let us start with X and Y. So, that means, if X causes Y, X causes Y and Y does not cause X, X causes Y, Y does not causes X, so this is called as a unidirectional causality, unidirectional causality.

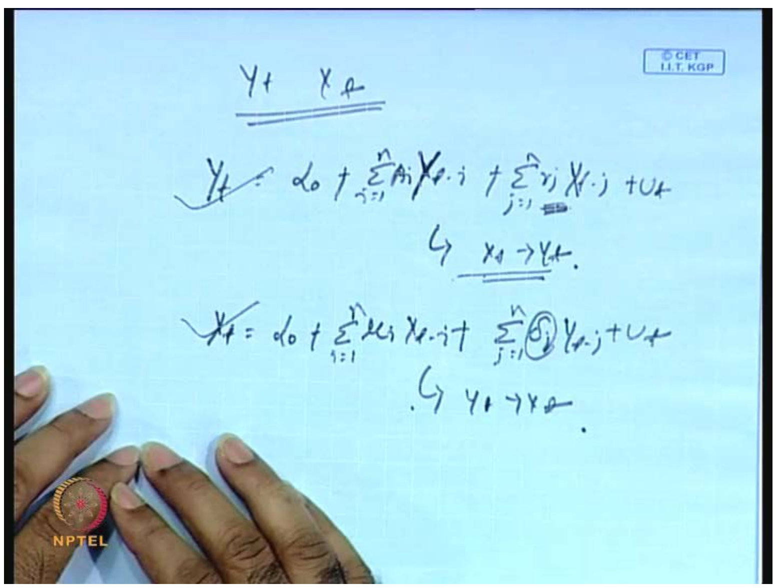
That means, causality is three forms: no causality, no causality, unidirectional causality, then bidirectional causality, bidirectional causality. So, now with two variables X and Y, we start with very simple models, then we will go for, you can extend to multivariate models. So, like you know, there is a concept called as a vector autoregressive model and vector error correction models. So, once you will attack with these systems, then we can be able to know, what is vector error correction model and model vector autoregressive model.

So, second, second case is Y influence X, and X, Y influence X and X does not influence Y, this is also called unidirectional causality; this is called as a unidirectional causality from Y to X. So, this is unidirectional causality from X to Y, but reverse causality is not there. So, third case is, this is third case is Y causes X and X causes Y, so this is called as bidirectional causality, bidirectional causality. So, that means, Y and X, both causes each other. Then, fourth, fourth possibility is, Y does not cause X and X does not, does not cause Y, X does not cause Y. So, this is, if this is the case, so then this is called as a interpretation, as a no causality, no causality between Y t and X t, X t or Y t.

So, that means, for a particular systems having two variables together, so there are three different steps, in fact, four different steps. No causality, that means, if X does not cause Y, Y does not cause X. Unidirectional causality means, X causes Y, then X, mean, in the other side Y does not influence X. Then, other unidirectional causality is Y causes X, but reverse is not true. Then, final is, X causes Y and Y causes X, so this is called as bidirectional causality.

So, we, we have to, means, cointegration will give you, that there is a possibility of causality, but it will not give you the integration of direction of causality. That means, whether there is really any causality or if it is so, whether it is unidirectional problem or it is a bidirectional problem that is our main agenda of this causality, Granger Causality test.

(Refer Slide Time: 35:48)



So, now, how you have to detect all these things, there is a technical procedure how you have to do all these things. So, let us start with Y t and X t. So, how do you check the Granger causality test? So, I will, I will directly prefer a model here. So, to check the Granger causality test, so we, we start with Y t equal to alpha 0 plus summation beta I X t minus i equal to 1 to n plus summation gamma J Y t minus J j equal to 1 to n plus U t. You know, this is one model, which we have to use, this is called as autoregressive model, as I have mentioned earlier.

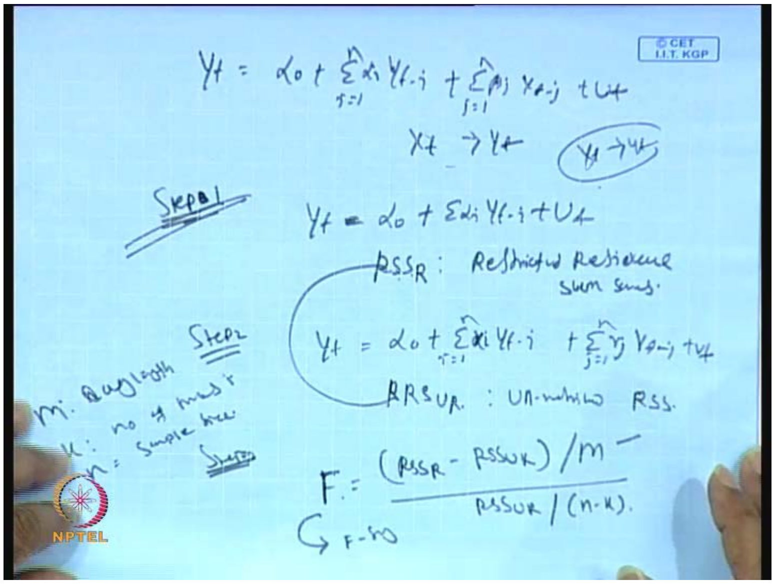
So, the divisions of models may be distributing lag model, may be autoregressive lag model, but the application of autoregressive model is much higher than the distributive lag model. In fact, there is a way how to transfer the distributive lag model to autoregressive lag model. So, with this particular setup Y t and X t, so we need to know whether Y causes X or X causes Y.

So, now, in this particular equation the moment I will write like this way, so it will give you signal. So, means, give you signal or our aim is to know whether it is X t will influence Y t. So, it is better we start with here Y t, that means, you start here, Y t minus I and it is X t minus J. So, that means, this particular equation give you signal whether X t influence Y t only. Similarly, we should know, whether Y t influence X, X, X t. So, that means, for that we need to have another equations. So, alpha 0 plus summation, you know, mu, mu I, mu i X t minus i, i equal to 1 to n plus summation delta j, delta j, Y t minus j, j equal to 1 to n plus U t. So, then this particular equation will give you signal, whether Y t influence X t.

So, now, once this is the case, then our null hypothesis is, that these are, these are, these coefficients should at least, one of the coefficients should not be equal to 0, so that means, this parameter should be statistically significant. When this is your task, then obviously, this delta J parameter, at least one of the delta J parameter, is, should not be equal to 0. This is our standard null hypothesis.

So, now, here we, we are now targeting to check whether X t influence Y t and here we are targeting to check whether Y t influence X t. So, now with two, two variables in the system, so we have to set two equations. So, similarly, there are three variables in the system; we will put three different variables, so three different equations. If it is four variables in the system, then we will apply four different equations. So, we like to know every time what are the other variables, whether causes other way, first variable or second variable, third variable or fourth variable. But when the system (( )) more, more and more number of variables, it is very difficult to handle manually. In that case, you have to apply the (( )) model vector, autoregressive model, alright. So, now, let us assume that with a simple problem Y t and X t means, the system should involve only two variables, Y t and X t. So, we like to know how is the procedure to check the causality issue here is, alright.

(Refer Slide Time: 39:27)



So, now, the standard procedure, let us, we start with Y t equal to alpha 0 plus summation alpha i Y t minus i, i equal to 1 to n plus summation beta j, beta j X t minus X t minus j, j equal to 1 to n plus U t. So, let us say this is our equation. So, that means, here our agenda is to know whether X t influence Y t. So, altogether there are three steps here. So, like cointegration, so is a cointegration test? So, it has three different steps.

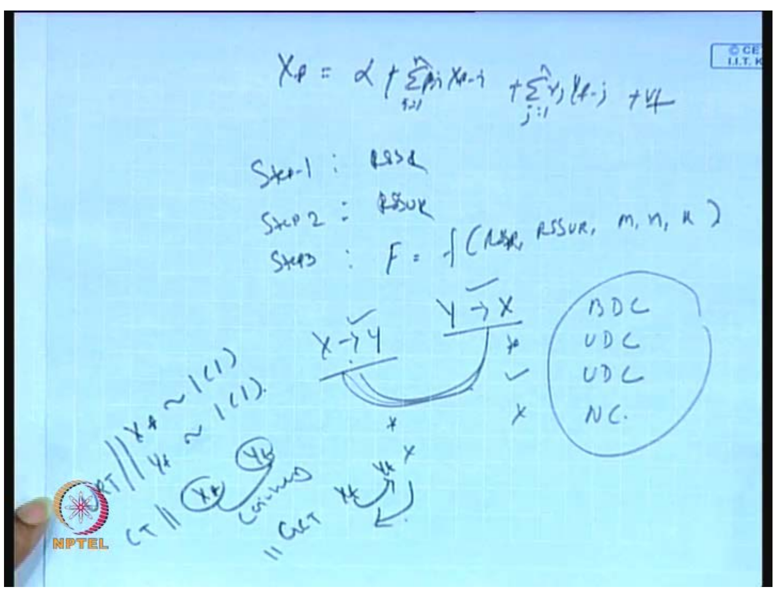
So, first test is, step one, (( )), first step is, there are three steps, first step, step one, step one is to have an equation like this, Y t equal alpha 0 plus summation alpha i Y t minus i, purely autoregressive model, Y t minus i, then plus U t. Then, you regress all these models, means, all the, all this regress, these models, then you have residual sum, obtain the residual sum square, RSS, that we will call it RSS, restricted residual sum squares. This is called as a restricted, restricted residual, residual, sum square, residual sum square, restricted residual, residual, restricted residual sum squares, RSS, RSS.

So, similarly, (( )) to step two. So, you fit the equation like this, Y t equal to alpha 0 plus summation alpha, alpha i Y t minus i, i equal to 1 to n plus summation gamma j, gamma j X t minus X t minus j, j equal to 1 to n plus v t. Then, you again have, again have a residual sum square, that we will call as an unrestricted, this is called as an unrestricted, unrestricted, unrestricted RSS residual sum squares. So, step one is to have residual sum square, restricted residual sum square. So, that is nothing, but to regress Y t with the Y t minus i only. So, then, obtain the residual sum square and call it restricted residual sum square.

Then, second, regress the full model, then you have the residual sum square. In that residual sum square you have, your name is unrestricted residual sum squares. So, now, in the step three, so you integrate this, this and this. So, how you have to integrate, for that you have to apply F statistics. So, F statistic is nothing, but RSS R minus RSS UR divide by M or divide by RSS UR divided by n minus k, M represents, here M represents lag length, lag length, lag length, k represents number of parameters in the systems and n represents sample size, n represents sample size.

So, now, you get the F statistics here. So, now, if F is significant, that means, this is calculated value of F and you have a tabulated value F, provided with this degree of freedom. So, once you, you find, that f is statically significant, then you will conclude, here X t causes, X t causes Y t, X t causes Y t. Again, for, you know, again for, you know, reverse causality, you feed the equation similar way.

(Refer Slide Time: 43:42)



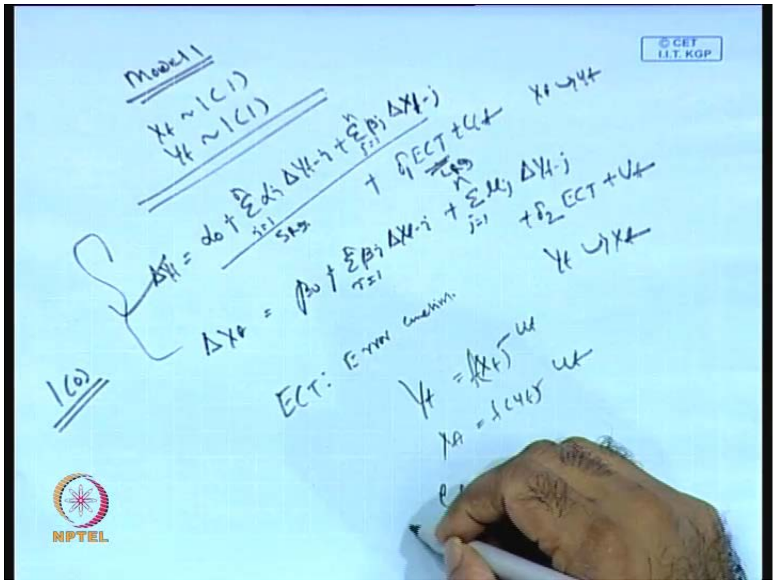
You know, in this way in this, you know, for reverse causality you feed equation X t equal to alpha plus summation beta A i X t minus i equal to 1 to n plus summation gamma J Y t minus J j equal to 1 to n plus v t.

So, again you have to follow the step one, step two, step three; step three, step two, step one, that will get RSS R, you will get RSS UR and you will get F statistic, which is function of RSS R, RSS UR divide followed by lag length sample size n degrees of freedom. So, this is how you have to calculate the F statistics. So, now, if by any chance it is significant, then you will call as X, influence X. So, now, in the first case we have checked X influence Y and here, we are checking Y influence X. Now, you compare these.

So, if it is true, if it is true, then it is called as bidirectional causality. If this is true, this is not true, then it is called as a unidirectional causality. This is true, this is not true, then this is called as again, unidirectional causality. This is not true, this is not true, then this is called as no causality. So, these are the four outcomes we have to find in the case of, in the case of, in the case of, you know, Granger causality. Here, one of the standard tricks is, that because before, see this is simple Granger causality test. So, here one of the standard assumption is, that X t, X t should be integrated of order 1, Y t should be integrated of order 1; X t integrated of order 1, Y t integrated of order 1. So, this is how we have checked through unit root test, unit root test; this is how we have checked through unit root test.

And similarly, next step, we know what is the X t and Y t relationship cointegrations. So, then we like to know, in the third this is cointegration technique. Then, this is, you know, in the third we have a Granger causality test, that too, you know, whether X t Y t are causing each other, both way causing each other or no causality each other. This is how you have to proceed here for checking the Granger causality test.

(Refer Slide Time: 46:25)



So, now, Granger causality has some advanced versions also. So, it has some advanced version. So, model 2, this is what we have discussed or else, let us say, model 1, model 1. What is the condition? X t is followed by 1, 1 order of integration 1 and Y t is the order of integration 1. So, then you apply what we have already discussed, but there may be other function model. So, I can write like this way. So, delta Y t is equal to alpha 0 plus summation alpha I delta Y t minus i, i equal to 1 to n plus summation beta j delta X t minus j, t, t minus j, j equal to 1 to n plus delta ECT, delta ECT plus U t.

This, this can be one model and another model, I will call it delta X t, which is nothing, but let us say beta 0 plus summation beta I delta X t minus i, i equal to 1 to n plus summation mu j delta Y t minus j, j equal to 1 to n plus, this I will call it delta 1, this I will call it delta 2 ECT, ECT plus v t. So, this is another form of the model. So, the, this particular structure will give you a signal whether X t causes Y t. This will give you a signal, whether Y t causes X t and this will give you long run impact, long run impact and these are all give you certain impact, certain impact.

So, that means, once, once you check the unit root component, then obviously it will give you the order of integration. So, by the help of order of integration you get to know what type of models further you have to use for cointegrations and causality. So, once you have order of integration, it will give you signal or how to check the cointegrations and how to check the causality.

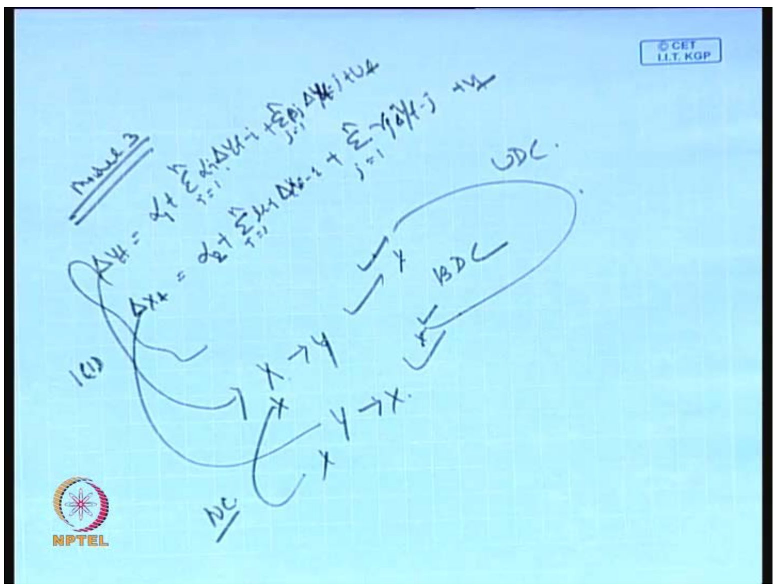
So, the moment you will get the variables means, you get the information about the stationary issue. Then, obviously, you have to enter the cointegration because it has already given the signal, that there may be long run equilibrium relationship and for that you have to apply a cointegration technique. So, once you apply the cointegration technique, then obviously, if it is there, then obviously, it will give you the signal to the Granger causality test means, existence of causality, so for that we need to apply the Granger causality test. Then, once you apply the Granger causality test, it will give you signal, whether, whether there is really causality or not, if it is causality, then whether it is unidirectional causality or bidirectional causality or no causality at all? So, this way you have to, you know, calculate the entire issue or you want to…; you can explore how to check the causality issue.

The thing is, here is, you remember here is, this particular ECT here, ECT stands for error correction terms, error correction terms here is, here is the standard format, is the variables are cointegrated, but we are using 1 0, 1 0 data. So, that means, it is integrated of order 0, integrated of order 0, so that, that data you have to use to justify the causality. That means, how to get the ECT, other items are always available. So, you have to just apply the first difference and you, you can get through, but in the case of ECT it is not available directly. So, you have to get through it.

For instance, how do you get an ECT, you know, because once the variables are integrated stationary, then you have to check the cointegrations. So, for that you need to, you need to know, whether there is any linear relationship between the two, that is, you know, Y t, Y t as a function of X t, function of X t, then or X t as a function of Y t. So, obviously, there is error term U t and there is a error term U t. So, the moment you will get the error term, so that error terms, once you get the error term, that error term will be considered as an error correction term. Again, it has to be inserted in a particular, in a particular equation. Then, you have to check the impact of ECT on this particular variable at the first difference levels. So, if it is, so then obviously, so you can get to know whether means, here in this particular format.

So, you have two different objectives, first objective is the direction of causality and second objective is the certain impact and long run impact. So, some of the, it is very essential because it is a time series model altogether. So, we have a certain interval and we have the long run intervals. So, now according to these setups, so we can check the structures and you can get to know the results. So, this is, this is, this particular structure we can call it a model two. So, there is another model, another model.

(Refer Slide Time: 52:02)



So, we can write like this way model 3, model 3. So, I can simply write delta Y t equal to alpha plus summation alpha i delta Y t minus i, i equal to 1 to n plus summation beta j delta Y t minus j, j equal to 1 to n plus U t.

So, this is one equation and another equation, I will call it delta X t equal to alpha plus. This is, I will call it alpha 1, this is I call it alpha 2, then I will call it say, mu I delta X t minus I, I equal to 1 to n plus summation, you can say, alpha, beta, gamma, let us put gamma here is, then, then this is Y t minus j. So, this should be X t minus j, this is X t minus j and this is X t minus i. Then, gamma j Y t minus j Y t minus j. Of course, delta is here, delta is here, then j equal to 1 to n plus v t. This is another model, this is another model where we are using one, one data, here is one, one data.

So, now if this is, if in this particular model our aim is to check whether it is, whether it is X causes Y, it is causes, X causes Y and this particular, yes this, this will give you the signal X causes Y and this will give you signal Y causes X, Y causes X. So, that means, if both are true, then it is called as a bidirectional causality.

If only one is true, others are not there, so in that case it is called as unidirectional causality. Either this is correct and this is not correct, or this is not correct, this is correct. This is called as a unidirectional causality. If both are rejected, then it is called as a no direction, no causality, no causality. Both are rejected, then no, it is question of no causality. So, this is, this is altogether the structure of causality tests.

So, what we have learnt in the couple of lectures. So, major issue of time series, major problems in the time series modeling is the unit root problem, cointegration problem and causality problem. So, you need to give a signal to know the stationary or order of integration of particular variables and that too. That is very important and very much useful, particular in the financial time series.

Then, with the help of unit root test, then we have to move to the cointegration, cointegration technique will give you signal about the existence of long run or yes, long run relationship between two or more variables. So, once you get to know the existence of long run relationship, then obviously, you can able to proceed to check the causality issue, direction of causality because unit root will give you green signal to the cointegration, but it will not give you integration, whether there is a long run association there, as a result you have to go for again, cointegration technique like EG test and JJ test, Johnson test.

So, if, if you get through the cointegration results, that means, if there is a possibility of long run association, then it will give you signal, that there is a causality, but it will not give you the signal whether it is unidirectional causality or bidirectional causality or no causality at all. So, in that case you have to go for another causality test with a different model setup with the help of different structure, different setup. You have to pick up a particular form of the model and end of the day you have to check, whether there is any kind of causality, if there is causality, then whether it is unidirectional causality or bidirectional causality. So, with this we can conclude these particular sessions. Have a nice day, thank you very much.