Econometric Modelling

Prof. Rudra P. Pradhan

Department of Management

Indian Institute of Technology, Kharagpur

Module No. # 01

Lecture No. # 34

Structural Equation Modelling

Good afternoon, this is doctor Pradhan here. Welcome to NPTEL project on Econometric Modelling. So, today we will start a new component, Structural Equation Modelling. So, in the last lectures, we have highlighted the details about simultaneous equation modelling, so we have discussed various issues regarding a single equation modelling and you can in the same times, there is a simultaneous equation modeling, so in the case of single equation modelling, there is the one way causality, but in the case of simultaneous equation modelling, there is the two way causality because, there is, there are interdependence of variable, interdependence among the various variables.

So, single equation modelling is very easy to understand, easy to calculate or easy to estimate whereas, a simultaneous equation modelling is very complex, very confusing or you can say very time taking component, but it is more interesting because it is, because, most of the problems real world problems are very interdependent in natures. As a result, it has to be discussed with a with an angle of simultaneous equation modelling, so without having sufficient knowledge and simultaneous equation modelling, it is very difficult to understand the real world biggest problems.

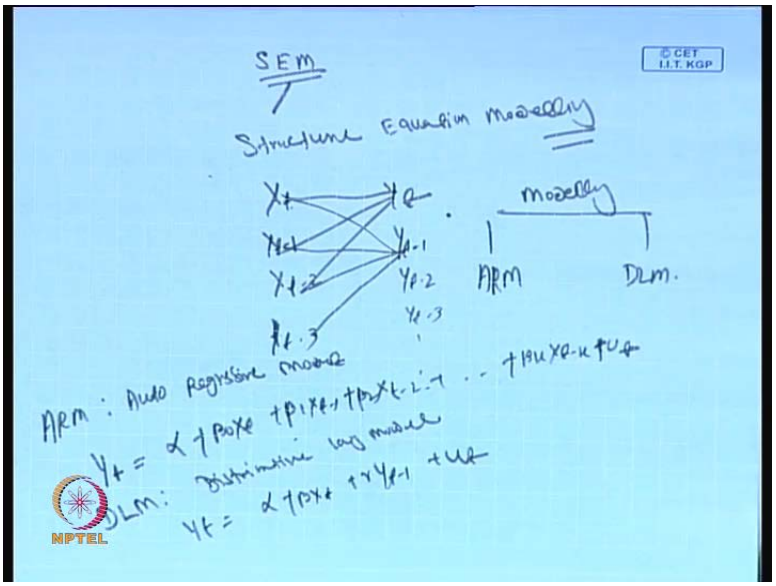
So, means in brief, we have made a difference between single equation modelling and simultaneous equation modelling, so with the basic information about the simultaneous equation modelling, we have to move to structural equation modelling, because a simultaneous equation modelling, will give you route to the structural equation modelling. In the structural equation modelling, it is also joint dependence of the variables, means existence of joint dependence among the variables, but here something and there is more interesting than the simultaneous equation modelling.

So, before you proceed to structural equation modelling, I briefly highlight some of the issues and concepts under simultaneous equation modelling, then we will converse to structural equation modelling. So, what we have discussed in the simultaneous equation modelling, is that means, the objective is to describe the joint dependence among the variables, you cannot estimate the parameters of particular single equation without having information about information about the parameters of other equation in the model.

So, now the moment you will go for simultaneous equation modelling, if basically, means, if we will move from single equation modelling to simultaneous equation modelling, then obviously, there are two (()) we have to, we can say overcomes, so that these are you can say, means, one is multicolinearity problem and another is a autocorrelation problem.

So, in the case of when we will apply single means, any problems using single equation modelling, may have a multicollinearity, may have autocorrelation, because some of the things which are interdependent, we are not incorporating, but the moment when you go for simultaneous equation modelling, so we are integrating the system in an accurate way, in a feasible way, so that the problem of multicolinearity and autocorrelation can be removed very easily. So now, you know, once you proceed to the, you can say, structural equation modelling, we will know more something more about the interdependence technique, it is one of the multivariate, you know, data analysis.

(Refer Slide Time: 03:59)



So, structural, S E M, the popular name is called as an S E M, this is nothing but, structural equation modelling, Structural Equation Modelling, so before I go to this structural equation modelling completely, so, I will slightly highlight some of the issues which we had not means, we have not covered details in the case of simultaneous equation modelling.

So, when we will go for simultaneous equation modeling, then obviously, the system will be very interdependent, that means, it is a problem where there are there are groups of dependent variables and groups of independent variables and it is very difficult to describe sometimes, who are the independent and who are the dependent, but we have to categorically classify, that is how identification problem is very important.

So, once you will identify properly, then obviously, it is not a problem for estimation, but the problem if I once if you identify, then obviously, the estimation will be automatically come means, accurate estimation will accurate perfectly come into the picture, so that is how we have to be very careful about this a system about the simultaneous equation system. So now, so, in the simultaneous equation system, so we have we have two groups of variables, one is called as endogenous cluster and another is called as a exogenous clusters.

So, there are many ways we can express this particular problems, in fact, before this simultaneous equation modeling, we have we have highlighted various issue regarding

you know, simple modelling, cross, that is cross sectional modelling, time series modelling, panel data modelling, so then, in the then, we are discussing here, the simultaneous equation modelling. So now, in this particular setup, so it is not, in the simultaneous equation modelling, in fact, the time integration is very high, for instance with a particular variable, we will create several dependent, independent variables with respect to lag and this is one type of, you know, creating simultaneous equation modelling and if it is a, if once you have a time lag then, obviously, multicolinearity problem and autocorrelation problem is (( )).

So, that is, how to avoid all such problems, so we are using simultaneous equation modelling, but, structural equation modelling is more advanced, more you know, more accurate. In fact, it may not, it can be applied for these types of problems and also it is very useful for cross sectional modelling where, there are groups of independent variables and groups of dependent variables, without any integration of time lag.

So, now if we you have the time lag concept then obviously, with 1, for instance if we have X t and only Y t, so, this t represents time lag and Y t, here also Y variable with respect to time lag. So now, we can create several independent variables, this is independent variables and this is dependent variable in the time frame, in the in the time being, let us say, so, then, we will create several independent variables X t minus 2, X t minus 3, X t minus 3, like this, it will continue. Similarly Y case, Y t minus 1, Y t minus 2, Y t minus 3 like this we can continue, so there are series of X, we will create here, there are series of Y, you can create here.

So, then there is, maybe you know, sometimes Y t depends upon X t, Y t depends upon X t minus 1, Y t depends upon X t minus 2, similarly you know, Y t minus 1 depends upon this, Y t minus depends Y 2 minus depends upon this one. So, Y 2 minus 1 depends upon X 3 minus t, this is how there, all together integrated to each other. So now, no before jump to structural equation modeling, I briefly find out, means highlight the difference about these particular structures.

So basically, when we will go for time series modelling and its lag, so far as a simultaneous equation modelling systems are concerned, then obviously, there are two ways you have to create the simultaneous equation systems. So, one is called as an auto regression scheme, another is called as a distributive log schemes.

So, modelling basically, the time series modelling basically, divided into two parts - it is called as an auto regressive, auto regressive models and this is called as a distributive log models. ARM stands for Auto Regressive Models, Auto Regressive Models then, what are you what is your auto regressive model auto regressive models? We can write like this way, Y t equal to Y t equal to alpha plus beta 0, X t plus beta 1, X t minus 1 plus beta 2 x t minus 2, it will continue, plus beta k x t minus k plus u t.

This is one type of auto regressive models, so then, there is another model called as a distributive log model, so, distributive log model. In fact, we have little bit highlighted this issue, in the case of, dynamic panel data modelling, so here is Y t equal to alpha plus beta X t plus gamma Y t minus ones plus u t, this is a simple model of distributive log models.

You see here is, so now, in the case of auto regressive models, Y t depends upon you know exogenous variables and (()) log variables, but in the case of distributive log models, so we are taking dependent variable, that is endogenous variable as a function of both endogenous variable and exogenous variable, which is slightly different to auto regressive model. In the case of auto regressive models, it is the endogenous variable, as a function of exogenous variable and its log length.

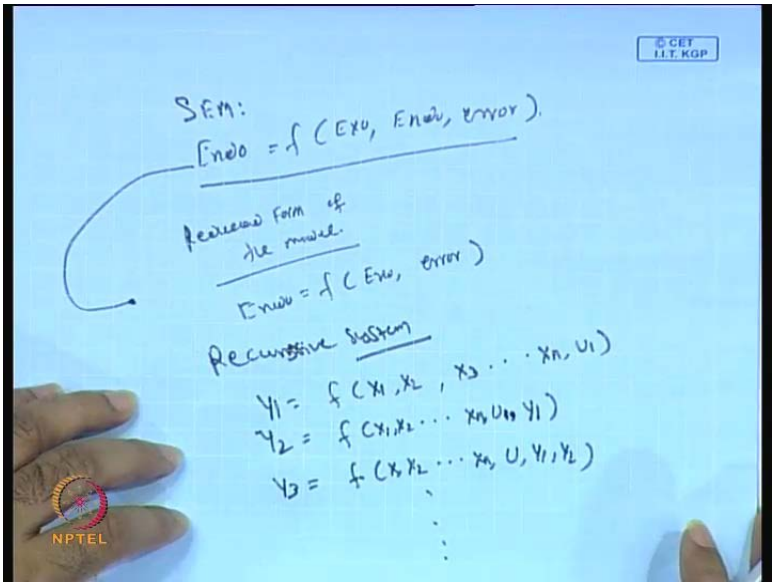
So, the moment you will enter the log, then obviously, the additional variable, will created and that too exogenous variable only, but in the case of distributive log models, it is the dependent variable and other side it is an exogenous variables and also endogenous variables and their logs, so this is the classic difference between the auto regressive model and distributive log model. That means, we start with Y t and X t then, we will create various difference, various types of models, so far as a simultaneous equation modeling system is concerned, in fact, these are all single equation systems, but we can go simultaneously with respect to, you can say, simultaneous equation systems.

So this is how, we have to describe this structure of, you can say modelling that too distributive log models and you know auto regressive log models. So now, these two are not so much important for structural equation modelling. In the structural equation modelling, so we are not concerned about the Y t minus 1 X t minus 1 etcetera. Our concern is so, what are the groups of endogenous variable and what are the groups of

exogenous variable. That is very important, so we have to highlight all this issue in this way.

So, first of all what is structural equation modelling, so in the moment you will move in to the simultaneous equation to the structural equation modelling, then obviously, this particular definition is very important.

(Refer Slide Time: 11:43)



Structural equation modeling, that means, what is structural equation modelling? Structural equation modelling, in general format is that it is the endogenous variables, endogenous variables, as a function of exogenous variables then, endogenous variables then, you say error terms, error terms.

So, this is how structural equation modeling, it is the structure of various means, various variables with respect to endogenous clusters and you know exogenous clusters. So, this is the form of structure structural equation modeling. So, similarly in the case of, there is a response to structural equation modeling, there is a concept called as a reduce form of the models. So, in the case of reduce form of the models, there is called as a reduced form of the models, reduced form of the models.

So, in the reduced form of the models, so here, endogenous variable, endogenous variable as a function of only exogenous variables and its error terms and its error term, so that means, we start with the having a models called as a structural equation

modelling, so either you have to bring proper structure and you have to find out the solutions or else you have to transfer the structural equation modelling to be reduced form of the model and then, we can directly apply the single equation model methods to estimate the parameters.

So, now in the structure, what is the difference between this a structural equation modelling and single equation modelling? In the case of structural equation modeling, there are, obviously, the model consist of at least 2 numbers of equations. Then, there is a complete integration, so first equation depends upon second equation or the second equation depends upon the first equation. So now, we have to prepare in such a way or you have, we have to create a model in such a way, so the proper they, both the integration, both the equation, will give you unique estimate, will provide you, unique estimated results.

But, if you are not smart enough to handle the structural equation modelling directly, so what you have to do, you have to go for a reduced form of the models, where endogenous variable as a function of your exogenous variable and error term. So now, this is, the moment you transfer the endogenous variable as a function of endogenous and exogenous error term to endogenous as a function of exogenous and error term, only then in that case, so, it is just like an application of senior equation modelling again; so, it is not a problem. So, only problem is how to transfer the structural equation modelling, to a say, you know, various form of the equation, that is single equation modelling.

So, that is very interesting, so the moment you are the moment you are, successful then, obviously, it is not a problem, it is as usual single equation modeling, because, the estimation process of single equation modelling is very easy to be, in compared to structural equation modelling and simultaneous equation modelling. So now, so either you can directly handle the structural equation modelling problem, so either with the help of simultaneous technique or you can say structural equation technique, so now, if not then you have to transfer into various forms of the techniques.

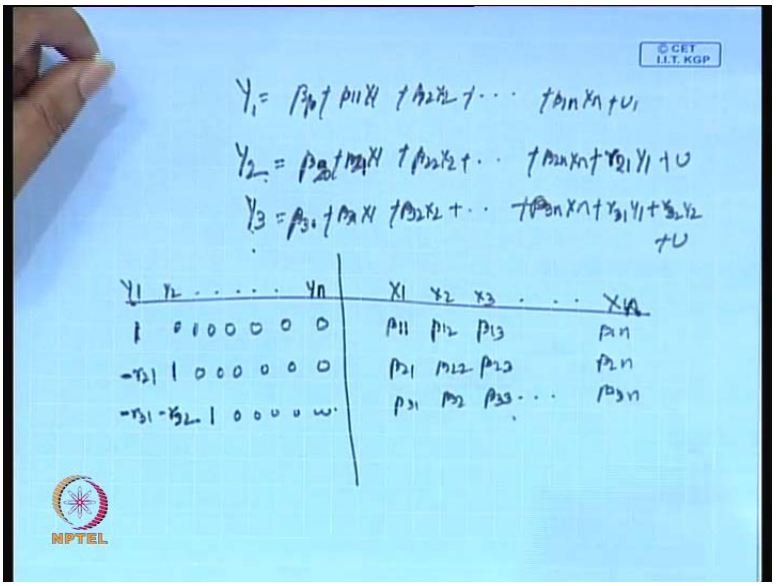
So, in the case for reduced form of the technique, so, you it is the system are a single in the format of single equation, just you have to estimate the parameters by the use of any techniques like O L S, G L S, W L S or maximum likelihood estimators. So now, in response to then means, in response to structural equation modeling and reduced form of

the model, so we have a system called as recursive systems. So, there is a concept called as a recursive systems, a recursive, a recursive, a recursive systems. Recursive system is that, it is the recursive system, it is, you know it is a chain of equations.

So, it is just like a structural equation modelling, but the first equation start with a endogenous variable as a function of exogenous variable only then, second equation will be a and second endogenous variable as a function of exogenous variable and first endogenous variable, then third equation will be third equation, third endogenous variable as a function of exogenous variable and first two endogenous variable, this is how you have to proceeds, as long as your variable in your hands.

So like this, you start with like this, so Y 1 equal to function of x 1, x 2, x 3 up to, you can say, x n then u 1. So, this is first equation where, endogenous variable as a function of all exogenous variable and error terms. So, in the second equation, Y 2 equal to function of x 1 x 2 up to x n u 1, so put u 1 error term, then, first endogenous variables Y 1. In the second case Y 3 equal to function of x 1, x 2, x n, then u, then y 1 then y 2. So, this is how it has to, it is to be, you can say, it is to be, increase subsequently.

(Refer Slide Time: 17:09)



So, if you will put it in a explicitly format then, the picture will be coming like this way, the first expression will start with like this - beta 0 plus beta 1 1 X 1 plus beta 1 2 X 2 beta 1 2 X 2 plus continue, beta 1 1 n X n plus u 1. Similarly, Y 2 this is Y 1, so, Y 2 equal to Y 2 equal to Y 2 equal to, beta 0 plus beta 1 1 X 1 plus sorry this is, 2 1 because,

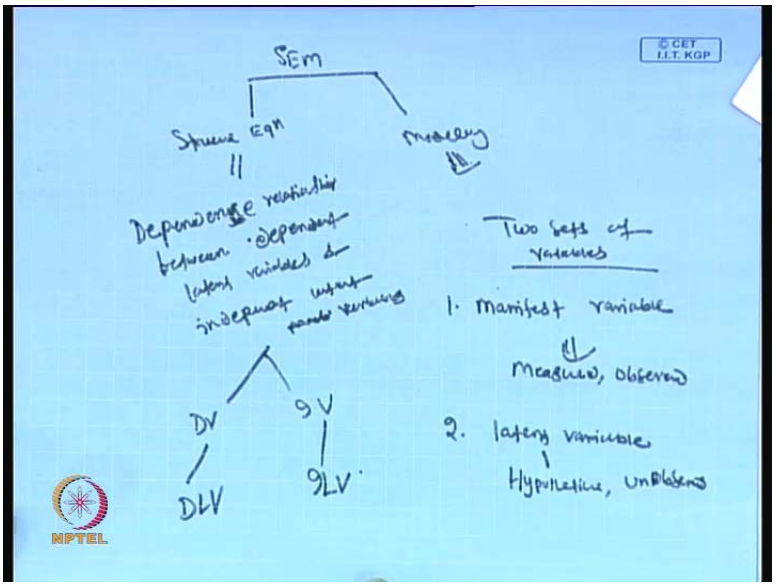
it starts from 2, so, beta beta is 2 beta 2 1. So, it is better you put here beta 1 0, this is, beta 2 0, beta 2 0 then, beta 2 1 X 1. So, beta 2 2 X 2 plus continue, beta 2 n X n plus gamma, gamma 2 1 Y 1, gamma 2 1 Y 1.

So, then similarly it will continue, so, we will put to another equation. Let us say, Y 3 equal to beta 3 0, beta 3 1 X 1 plus beta 3 2 X 2 continue, plus beta 3 n X n plus gamma 3 1 Y 1 then, plus gamma 3 2 Y 2 plus error term plus error terms. So, this is how, it is and it will proceed like this way. So, that means, if will you put it in a matrix format, this will come, like this way, so this side, X 1 X 2 X 3 up to say X k, so this side Y 1 Y 2 it will come like you know Y n and this is X n, X n, so this is beta 1 1, so I am just reporting the coefficient of variables only.

Beta 1 1, so beta 1 2, beta 1 3, then beta 1 n, similarly beta 2 1, beta 2 2, beta 2 3, beta 2 n, so beta 3 1, beta 3 2, beta 3 3, beta 3 n. Similarly, for first equation these are all 0, here, except first equation 1, in the second equation, Y 2 becomes 1 and Y 1 becomes minus gamma 2 1, gamma 2 1, so others are remaining 0. So, in the third case it is gamma 3 1, minus gamma 3 2 then, this is 1, then these are all 0, so this is how it is proceed.

So, this is another format of you know, simultaneous equation modelling. So, it is called as a recursive system. So, you, so, we have knowledge on single equation modellings, simultaneous equation modellings, structural equation modelling and recursive systems. So, with this basic setup, so we have to move to the concept called as a structural equation modelling.

(Refer Slide Time: 20:37)



So, first of all, what is structural equation modeling? So, what is structural equation modeling? Structural equation modelling has a two different setup altogether, one is called as a structural equation another is called as a modelling.

So, what is modelling? So, modelling is the process of presenting the dual valve problem as a set of mathematical equation, it is a process to transfer the involved regions problem in the form of a number of equation, that is what, we call as a modelling. So, in the first lectures we have highlight details about this modelling scenario. In the second case structural equation there, is a dependent relationship between in the structural equations, it is the dependency, dependency relation, dependence it is better to put dependence relationship between, relationship between, relationship between, dependent latent variables, dependent, dependent latent variables and independent latent variables, independent latent variables, latent variables, latent variables, that is right, independent latent variables.

So, now, in the case of structural equation modelling, so we have two different, we have two different setup, first is modelling and second is the structural equation. So, modelling means, it is the process of, it is a transformation rule to transfer the real dual business problem, into mathematical form of the equation. In the second case, it is a structural equation, structural equation, it is the dependence relationship between

dependent latent variables and independent latent variables, so that means, it is the question of dependent clusters and independent clusters.

So, I will highlight what is all this dependent cluster and independent clusters? That means, we know, all together dependent variables and independent variables, but here we are discussing about dependent latent variables and independent latent variable, so that means, it is something more than the structural equation modelling. So, now, in the structural equation modeling, there are several dependent variables, several independent variables. We like to know how they are integrated to each other, interdependence to each other, but here, structure is very important, so, that is why it is called as a structural equation modellings.

So the structure, so the way you will decide the structure, so, in that, in that case, so, we have to use the structure called as a dependent latent variable and independent latent variable. So that means, so, we have to bring a structure where, we can observe the dependence, interdependence among the various variables, with respect to endogenous cluster and exogenous variable, exogenous clusters.

So, that means, in this particular structural equation modelling, there are 2 sets of variables, so one, there are 2 sets of variable, so 2 sets of, there are 2 sets of variables, so like, you know simultaneous equation you call as endogenous variable and exogenous variable, here we will call as a manifest variables and you know latent variables. So, here, we will call this manifest variables, manifest variables and another is called as a manifest variables and another is called as a latent variables, latent variables. Manifest variables, otherwise, called as a measure, it is measured variables observed, can be observed. So, this latent variable is basically hypothetical and unobserved. It is hypothetical and unobserved.

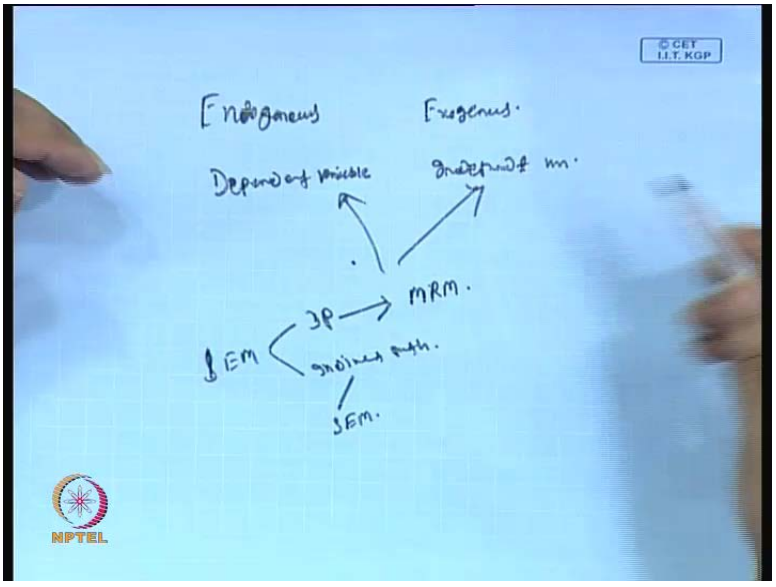
So, in the, there are two variables - one is called as a manifest variables and another is called as a latent variables. Manifest variable is nothing but, measured variable or observed variable. Latent variable is nothing but, hypothetical variables and unobserved variables, that means, data usually available for manifest variables, then ultimately, we have to create a latent variable with respect to this manifest variables, just like you know, in the case of time series modelling, we have Y t, relationship of 2 variables Y t and X t.

So, data will be available for Y t and data will be available for X t, but we will create Y t minus 1, Y t minus 2, Y t minus 3, like you know Y t minus case. Similarly, you can create X t minus 1, X t minus 2 up to X t minus k. So, within a particular, we are creating several variables, this is one type of structure. But, here the structure is that, so we have, we have, two groups of variables, so the in that times, this case it is called, as a log variables.

So here, so we are using manifest variables and you can say latent variable. Manifest variables, it is a very much observed data files are there, means data are readily available, but in the case of latent variable, it is hypothetical and unobserved, so it is not available, so we have to construct the data file with the available data files, existing data file, that means, so we have to create a something, that is, it is a we have to use the processed processed data in a different way, that is, the you know beauty of the structural equation modelling.

So, we have to integrate means, the original variables with the latent variables so that the latent, the integration of latent variable and original variable will give you the proper structure or you can say beautiful modelling setup, so that is how we have to, we have to discuss these concept of structural equation modelling. Altogether, this picture will be coming like this way.

(Refer Slide Time: 26:48)



So, endogenous with endogenous variables, so basically it is the game between endogenous variables and exogenous variable.

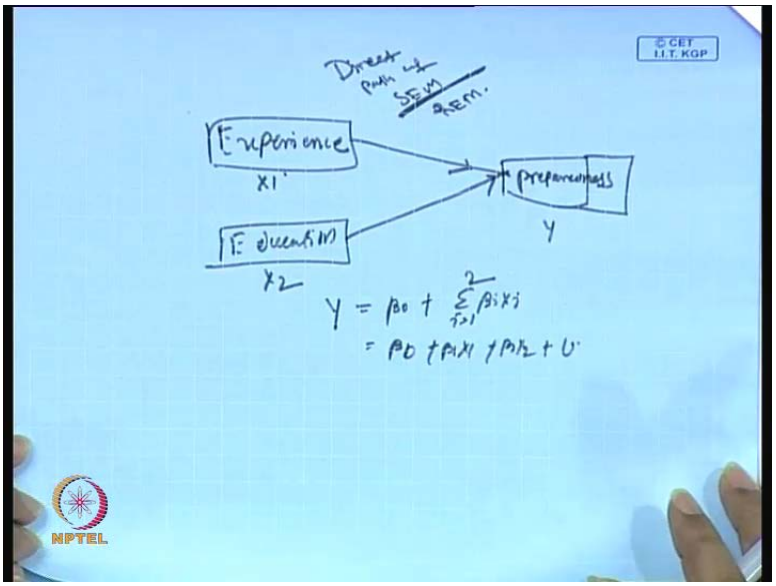
So, that means, it is the dependent variables, it is the game between dependent dependent variable and independent variables, it is the game between dependent variable and independent variables, where the variables in the models, so that means, so, so we have to give a signal, so where is the dependent variable and where is the independent variable? For instance, I will give you a proper structure here is, so the structure is here is, I will take you know a simple model here.

So, far as a structural equation modelling is concerned, there are two ways you can say, there will be a direct path and there will be indirect path, indirect path, so, so indirect path is basically the form of structure, appear structural equation modelling, direct path is purely a multivariate, multivariate regression modelling, it is just like your multivariate regression modelling.

Before I go to highlight details about structural equation modelling, its origin and its structure, its methodology and its application, so I briefly highlight how this multi means, multivariate, it means the origin of this you know simultaneous equation modelling and structural equation modelling is from the multiple regression only, so it is the, we are handling multiple regression in a different step altogether, so that is in the form of simultaneous equation modelling and structural equation modelling.

So, now, I will take one simple example, you know, how we can analyze that particular problem, in the case of, you can say multivariate, multivariate regression modelling and how we will put altogether in the case of, structural equation modelling. Let us see, here is or in between, how simultaneous equation modelling can be, you can say, discussed.

(Refer Slide Time: 29:11)

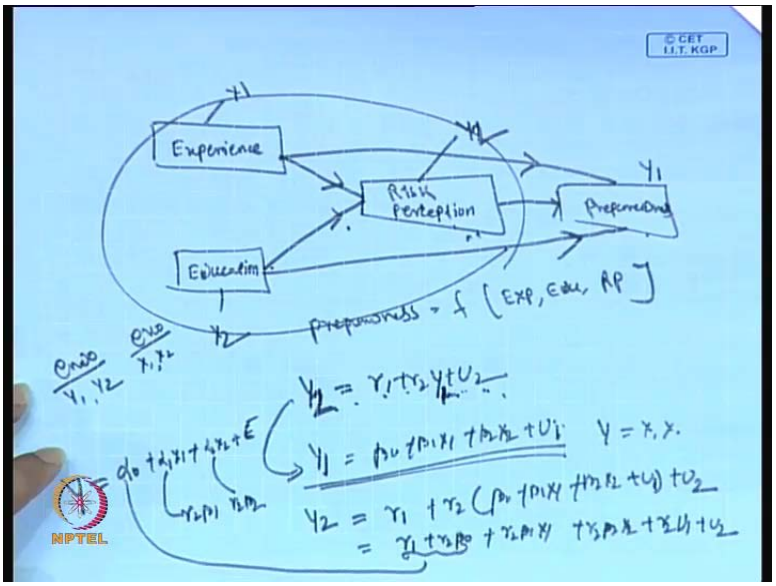


So, the proper setup is like this, so, the proper setup is like this, let us say, so, I will take two variables. Here is, you can say, let us take case of experience is a one variables, then another variable, I will take it educations. Then these two are, I will, I will assume that these two are independent variables. Then, you know, here, I will put it, a component called as a preparedness, preparedness, there is another variables. So, now, this, the, you know simple direct path is like this. So, this indicates; that means, this I will call it x 1, this I will call it x 2, this I will call it y.

So, now when when I will go for, when I will go for, you can say, simple direct path then obviously, I would apply direct multivariate regression analysis, that too multivariate multiple regression analysis, so that times I will, I will mention like this way, beta 0 plus summation beta x I. I equal to 1 to 2. So, in other words, we can write like this beta 0 equal to beta 1 x 1 plus beta 2 x 2 plus you can say error terms, alright but in the case of this is simple structure, this is direct path, direct path of, direct path of more structural equation modelling, direct path of structural modelling is in something or nothing, but regression regression equation modeling, it is something regression modelling only.

So now, so how is it how is it different to structural equation modelling? I will, let us say, I will, I will put here another diagram here, then we will, I will highlight the structural equation modelling.

(Refer Slide Time: 30:57)



So, let us say, same problems I will have highlight here, so this is an experience, this is experience and this is educations, this is experience and this is education, then this is, this perceptions. I will take another component called as a risk perception, then preparedness. You see here, generally when we have a particular problem, we start with because, ultimately this particular technique is also causality technique.

So, here the only difference is, here we are having, you know, bidirectional causality or two way causality or in the case, on the contrary, in the case of single equation modelling, we always look for one way causality, but here in the case of particular structural equation modelling, structure is very important, then in the same times, we look for, you know, two way causality. For instance, you know, we, in this particular case we highlighted the model or problem where, preparedness as a function of experience and educations, preparedness as a function of experience and educations.

So, our aim is here means, very beginning our aim is, how means, objective is, how experience and education are related to preparedness, whether we will put more experience to get preparedness or we will put, you know, more education to prepare more preparedness, so this is how the multiple regression can be set. But you know, this in between if you will add any variables, or you can say, subtract another variable and the system will be completely changed, but you know, when there is a talking of structural equation modelling, so within a particular system.

So, if will you add something, then, it will be create a structures and then, if you will form a model so that form, that model is called as a structural equation modelling, so that means within a particular system, so you have to bring particular factorial variables, so either by a natural process or you have to create, like in the form of latent variables. So now, within that particular structure, it will you integrate all together, then, it is called as a structural equation modelling. Let us see, how is the loop of structural equation modelling?

So, this is how, this is how, the structures, so now, we have already mentioned, so experience and the risks, has a, experience has a impact on risk perceptions, education has a impact on risk perceptions, then you get preparedness, so preparedness we have, already highlighted, there is integrations. So, similarly, you have to highlight here, with a concept preparedness, experience has a impact on preparedness, then finally, risk perception has a impact on preparedness that means, you see here, this picture is more slightly complicated, so you see, so we can directly write like this way.

So, for instance, preparedness is equal to function of experience, educations and risk perceptions, you can write like this way, then, you can analyze, but in that case, in that case, we cannot get a particular structure or we cannot highlight the hidden causality in between these two problems, so, that means, so our, if will we apply you know, particular problem with like to means, with the application of regression analysis, then, obviously, experience, education and risk perception should be completely independent.

So, that means, there should not be any such linear relationship among them, but here, so, we find risk perception has a function of experience and education, so, that means, there is a some kind of problem in between. So, we have to first sort out that particular problem then, you have to go for, you can say, its solution. For instance, if I will highlight this one, if I will call it, experience as X 1 and you know education is X 2 then, risk perception is you can say Y 1, so, then, this is Y 1.

So, we will call it, this is Y 1 and this is Y 2, so that means, here and the, and the endogenous cluster is a y 1, y 2, so, exogenous cluster will be x 1 and x 2, so only the thing is that, only a classifying endogenous variable and classifying exogenous variable will not solve your problem or will not give you any kind of solutions for this, you know interdependence, multivariate interdependence technique, so here what is the important

thing is that you have to find out the structure, hidden structures lies between this particular problem set, means whatever problems we are discussing.

So, now here, so we like to know, how endogenous means, there are two specific you know ideas here, one specific idea or objective is to classify who are the endogenous variable and who are the exogenous variable; the second specific objective is that, we have to find out to whether there is any structure in between these variable; if there is any structure in between these variables then you have to directly apply the structural equation modelling, if there is no such structure then you have to go for simple multivariate regression only, if there is no structure.

Then obviously, we will assume that these variables are completely independent to each others, so we should not go to complex problem, so it is better to serve this particular program in a simple manner by using multiple regression analysis only. But within the, you know, multivariate, the way you are solving that particular problem, so if there is any structure in between, then obviously, it will be better or it is suggested to use structural equation modelling, rather than the, you can say, simple multivariate regression modelling.

Because if you do not understand the structure then obviously, the model accuracy will not be perfectly, either it will affect the goodness fit of the model, or you can say it may affect, affect the significance of that particular variable, but altogether the blue property will not be obtained. Because, we are picking up, it is obvious, of course, there is a big model which, has a close integration and we are just picking one equation and estimating the parameters without having sufficient knowledge. That is why, a very beginning, I have mentioned, before you handling any statistical modelling or before you apply any statistical modelling to a particular problem, you must have a sound knowledge on theory.

So theoretically you should be very very sound, where is the root of the problem, the identification of problems, the factors effecting that particular problem means, you have to identify what are the total number of factors and who are the dependent and who are the independent, whether there is any extra variables which can inserted and what whether the variables which represents are all relevant variables? These are the issues

you have to, you have to take care before you enter to this particular structure of particular setup of modelling.

So now, generally what people used to do, so they do not know all these things, simultaneous structures or structural equation modelling etcetera. So, they simply, you know, they, if you give them problems, so if you just give a clue that, these are all independent variables and these are all dependent variables, sorry, only one dependent variables and these are all independent variables, so then obviously, so they will just simply apply the multiple regression, just they enter the data in the particular software click click they will get the results.

So, these are not the factors. So here, even if the, here the problem is that, we are very much keen, what is the hidden agenda in that particular problem? It is not only, obviously, the system will be very efficient, very feasible, very practical if there are more number of variables in the systems but, more number of variables in the system means, in the one hand, it will be give you more feasibility, more practical, means, more practical things, but in the same times it will add lots of complexity.

For instance, if you will add more variables then, obviously, there may be question of a, means, there is, there may be problem of simultaneous equation modelling or there may be structural equation modelling, means, that means, there may be some interdependence among the variables because all, the moment, you will use all these means more number of variables in a particular setup then obviously, all variables cannot be exclusively independent.

So obviously, there should be some kind of integration. So, if some kind of integration there and you directly apply this regression modelling then obviously there is a standard problem with multicolinearity and autocorrelation. But, if will we go indirectly, exactly knowing the, means, once equation and how this equations are, means, what are these variables and how these variables are whether these variables are influence with other variables or they are totally independent, if they are totally independent then, absolutely fine, if they are not means, totally means, they are not totally independent then, obviously, there is a serious problem.

So, for that you have to find out a particular solution for that. So now, for this particular problem, so what I will do, so I will go first like, you know, what is we have already

highlighted y 1, y 1 equal to, y 1 equal to gamma 1 plus gamma 2 y plus u 2, gamma 1 plus y 2. But, we have already highlighted this is gamma 2 y 2, but we have highlighted y 1 equal to y 1 equal to beta 0 plus beta 1 x 1 plus beta 2 x 2 plus u. So this is which, we have highlighted this particular setup only, sorry, means in this particular setup, y equal to x 1 and x 2 so, this is how we have cited here.

This is game between, this is if I call y 1 then, this is nothing but, y 1 as a function of x 1 and x 2 only, so now, it is the function of x 1 and x 2, so now, if we integrate this one here then obviously, this first equation is written like this way, so in the second case we are using the risk perceptions, so how risk perceptions are integrated, that means, here is risk perception, has a integration with y 1 so obviously, one of the standard equation you will find out here is, y 1 is equal to gamma 1 plus gamma 2 y 2 so, plus u 2.

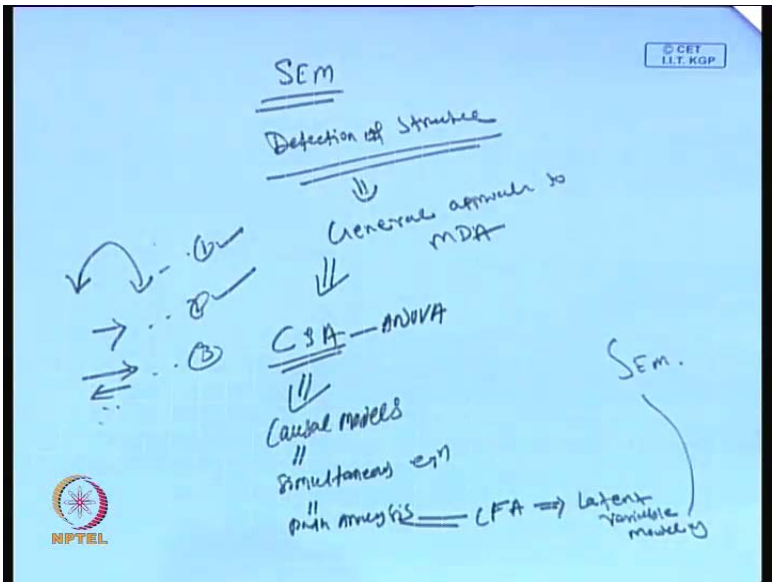
So, then what we have to do, to solve this particular problem we have to summarize this one. So, this is, this is, obviously y 2, this is obviously y 2 and this is obviously, y 1 this is obviously, y 1. So, if it is y 1 here, is alright. So now, so what you have to do, solve y 2, so we have to put gamma 1 plus gamma 2 into beta 0 plus beta 1 x 1 plus beta 2 x 2 plus u plus, let us say this is this is u 1 this is u 2, so this is u 1, this is u 2. So now, you simplify this one, so this is gamma 1 plus gamma 2 beta 0 plus gamma 2 beta 1 x 1 plus gamma 2 beta 2 x 2 plus gamma 2 u 1 plus u 2.

So now, what I will do, what I will say, what I will say means, the moment I will summarize then, I can write like this way, simply, you know, alpha 0 plus alpha 1 x 1 plus alpha 2 x 2 plus simply say, error component, it is done . So now, see here is, alpha 0 is nothing but this particular item, similarly alpha 1 is nothing, but, gamma 2 beta 1 then alpha 2 is nothing, but gamma 2 beta 2, so then accordingly, we have to find out a solution. So, this is now just like a single equation modelling, we can solve this particular problem.

But this is very much, if you understand inside the problem, what is the structure you know, hidden in this particular problem. Until and unless, you know the structure, you cannot sort out that particular problem or you can say examine that particular problem. With this basic setup, so we have to we have to proceed to this particular simultaneous, what simultaneous equation modeling, that means, our main agenda is what is this

simultaneous, sorry, structural equation modelling and how this, what is the route and how you have to solve this particular problems?

(Refer Slide Time: 45:01)



So, now structural equation modelling, basically structural equation modelling, is the, basic objective, is to identification of structure with set of observed variables. So, the basic idea about structural equation modelling is that, so you have to find out the proper structure within the observed variables that means, existing setup, so we have to find out to which structure (()) the structure, is very important in the case of structural equation modeling, so detection of structures, so detection of structures, detection of structures, is one of the prime objectives of structural equation modelling.

So that means, but ultimately, it is generally a approach to multivariate, it is a approach to multivariate data analysis so that means, so we can write, here is general approach to, general approach to multivariate data analysis, so that means, you see here is, so structural equation modeling has two important features, so first thing so we have discussed various aspects of econometric modelling, so then finally, this is almost of one part of the, you know, problem called as a regression modelling, then, we will move to the time series modelling.

In the end part of the regression modelling, is the structural equation modeling, so here we have to means, we are at the extreme, complex environment where we have two specific objectives - first objective is, a means, particularly, for the structural equation

modelling, it is the generalization approach of multivariate data analysis. Though we have discussed multivariate data analysis in the case of, with the help of multiple correlation, multiple regressions and also simultaneous equation modeling, so similar, in fact, simultaneous equation modelling, structural equation modelling, there are very close connections, more or less same, there is a slight difference is there, so we will highlight what is the slight difference, but here, in this structural equation modeling, so two important feature is, it is generalize, generalize, generalize approach to multivariate data analysis and most important thing, most important feature is, to find out a particular structure which usually in hidden, so you have to explore it properly, how they are integrated usually.

And another thing is that, it is very difficult to find out the structure, without having theoretical knowledge and sufficient knowledge about the problem. Generally the theory and problem surrounding, will give you, to prepare a particular structures. Mathematics, mathematically, we can bring the equation, we can set the equation, but it is very difficult to set in a structure, so because, various variables how they are integrated, we cannot hypothetically just represent, we cannot hypothetically, can write something, there is a something available in our hand then with the availability we have to prepare the models, so that is the beauty of the structural equation modelling.

So, there are various form, various ways, we can represent the structural equation modeling, it is called as a covariance analysis, covariance, it is covariance structure analysis, CSA, it is otherwise called as a CSA, covariance structure analysis, sometimes it is called as a causal models, causal models, sometimes it is, sometimes it is called as a causal model and this is otherwise called as a an Annova, analysis of variance, so then, simultaneous equations, it is the question of simultaneous equation, simultaneous equation.

Then, there is a concept of path analysis, there is a cause concept of path analysis, then, then, there is a integration to confirmatory factor analysis, then latent, then it leads to latent latent variable modelling, latent variable modelling, latent variable modelling, so, this is, that latent variable modelling is nothing but, the question of structural equation modelling, latent variable modelling is nothing but structural equation modelling, that means, the structure is like this, so you have dependent classification, that is a

endogenous variable cluster, then you have also exogenous clusters, that is independent variable.

So, within the endogenous variable setup, so we have to find out the latent variables, so that means, within the endogenous, how the factor can be consider, so factor can be created, that means, it is basically integration of path model and factor models. So, now, you have to understand what is what is factor model? Factor factor analysis is the interdependence technique, through which, we have to transfer the original variable to a smaller set of variables which is a linear combination of original variable.

So, that is how the structure is all about, that means, we have to, we have to create a latent variable, which, has a particular structure or integration with the original variables, so that is how, it is called as a structural equation modeling. So there are many things, we have to integrate here, is with respect, with regards, to this structural equation modelling, but originally originally, we look for two different tools here, is two different objectives, so one is correlations and another is causality; so that means, the structure will be like this way, so it, the indication will be like this way, one way causality or two way causality, so these are, this is one, this is second, this is third, so these three are common common, agenda for the structural equation modeling.

So that means, so, when there are a large number of variables in the systems and that is in that particular system, if we have any structure then obviously, how you have to bring that particular structures, so either the structure is a one way causality or two way causality or there is a question of correlations, covariance.

Two way causality means, it is a reverse direction is not allowed in fact here, but we are calling it is a basically, it is game between these two only one and two. So, it is just, we are showing that, there is a question of bidirectional causality because, sometimes you know, same variable will be effecting indirectly through other variables. So, that is how, that means, here we have to, means, we discussed already, correlation, covariance technique which is the, which can express the relationship between two variables, two independent variables, but in the case of causality we have to know who causes what.

So, now once you find out the structure then obviously, we get to know there are a some things which we have causal relationship and something which observes simple relationship, so that is the correlations, so structural equation modelling, the basic

framework or the basic idea is to find out, the correlation issue and the regression issue. So, that is the, that is the basic or measured agenda for the structural equation modelling. But, it is, it is not so simple. Correlation is a very simple technique and regression is also somewhat you can say in fact it may be little bit complex, but it is so much easy like you know with respect to structural equation modelling.

So correlation is easy to understand, easy to calculate, regression is easy to understand, easy to calculate, so like that you should have some knowledge on this particular statistical theory, then then, we have to enter to the structural equation modeling. In the case of structural equation modeling, the most important thing is, you have the systems, then, you have to find the, examine the structure or detect the structure, within their detection structures, we have to see where we can, where we have, correlation and where we have a regression.

So, the issue of correlation and regressions with the particular setup is called as a structural equation modeling, so, it is very interesting, so, because, it has a lots of application, in fact, the origin started from the psychology, then, you can say, psychology, then economics, then you can say biology, so it has a many application in many fields in various ways. So basically, the structural equation modelling is the integration of path model and factor model, so we briefly first highlight little bit about the factor model, then we will briefly highlight the path model then, we will integrate these two to get the structural equation modelling.

So, we will discuss details about this setup in the next class. With this we can conclude this particular session. Have a nice day. Thank you very much.