INDIAN INSTITUTE
OF
TECHNOLOGY
KHARAGPUR

NPTEL
National Programme
on
Technology Enhanced Learning

Applied Multivariate Statistical Modeling

Prof. J. Maiti

Department of Industrial Engineering and Management
IIT Kharagpur

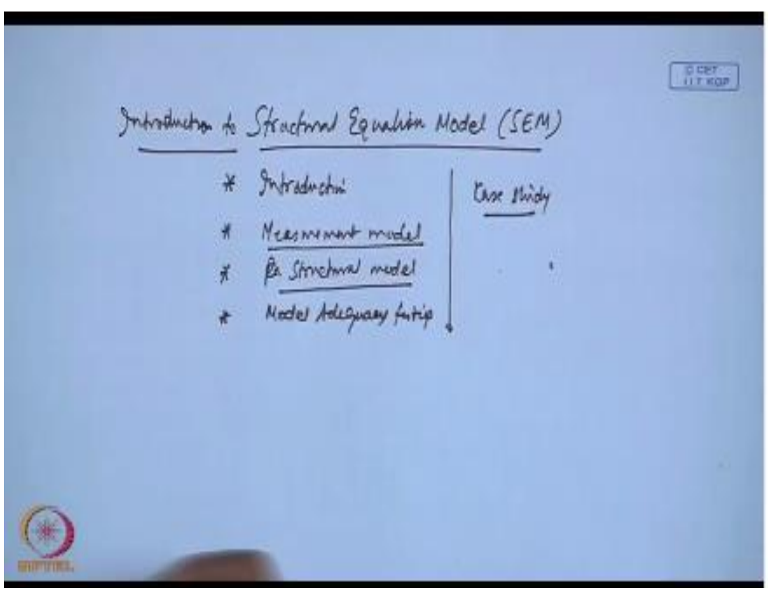
Lecture – 38

<u>Topic</u>

Introduction to Structural Equation Modeling

Good evening, today I will discuss structural equation modeling.

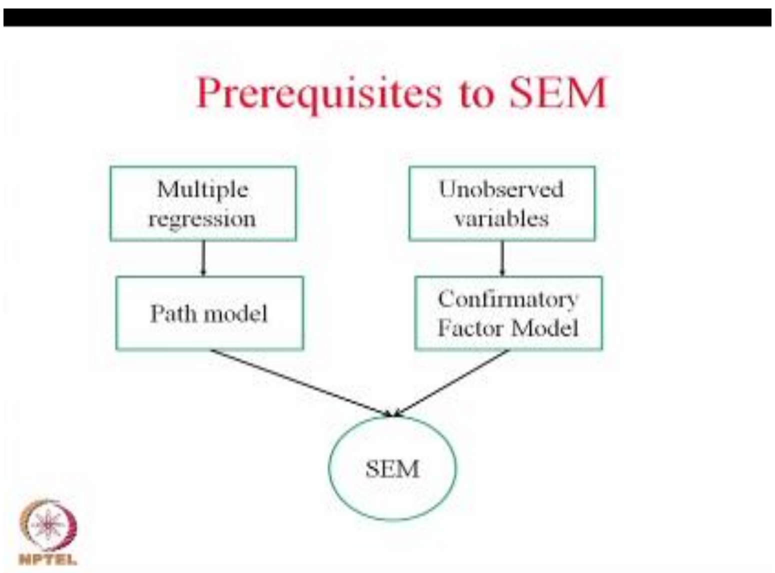
(Refer Slide Time: 00:24)



Also known as SEM. It is a little bit complex, advanced and it requires knowledge of some other tools and techniques what already covered in this course and I have partitioned this total lecture into three primary headings one is today's introduction. So today I will show you that what is structural equation modeling and how it is applicable to real life problem solving. Then also we will discuss that what are the components of structural equation modeling.

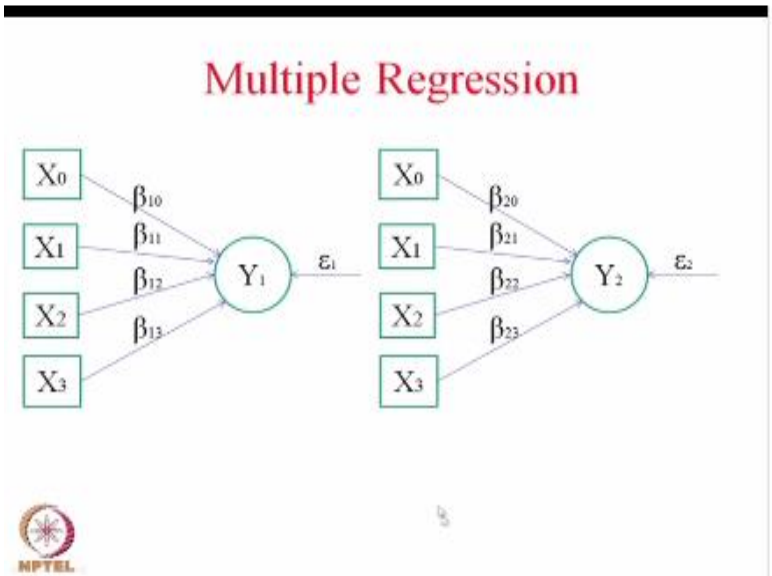
For example it consists of measurement model, it also consists of path model or structural model. So we will elaborately discuss this measurement model in next class then we go for path model discussions and then there will be the model adequacy testing, adequacy testing and obviously I will discuss case study one case study all through, okay. Today's lecture is introduction to structural equation modeling, let us see that what are the prerequisites to structural equation modeling.

(Refer Slide Time: 02:39)



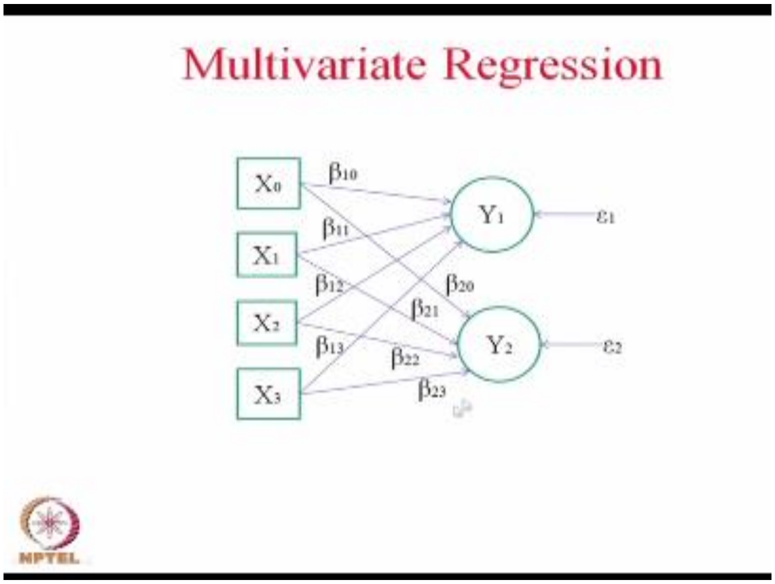
You must know multiple regression, you must know path model although I have not described path model till date but within the system we will be describing path model. You must know confirmatory factor model whose pre requisites is exploratory factor model so that unobserved variables can be measured through exploratory factor model. Now what is the difference between confirmatory and exploratory factor model that we have already discussed. So that means if you know confirmatory and exploratory factor model you know path model which is basically the structural model then you will be able to apprehend structural equation modeling.

(Refer Slide Time: 03:35)



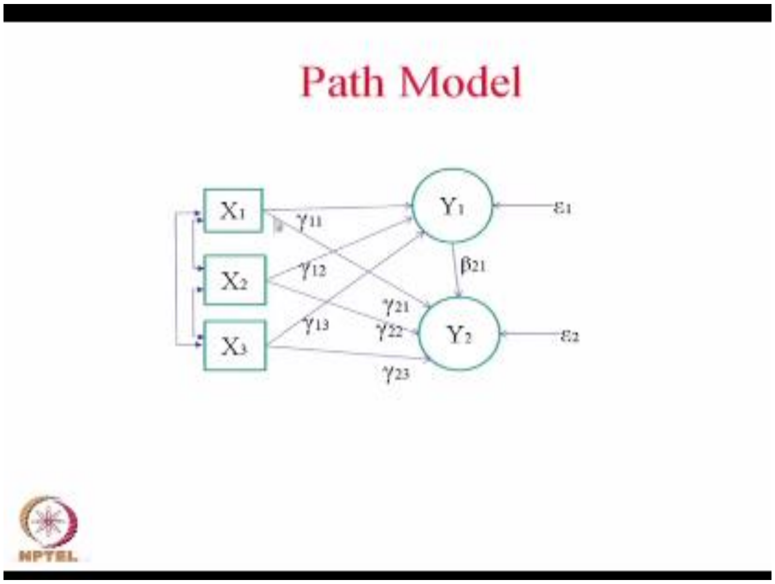
Let me recapitulate little bit of the prerequisites, I know that all of you have seen this type of pictorial representation for multiple regression, where Y1 here Y1 is affected by several variables as well as Y2 is also affected by the same sets of variables X1, X2, X3 and when we have more than one dependent variable we have seen that.

(Refer Slide Time: 04:09)



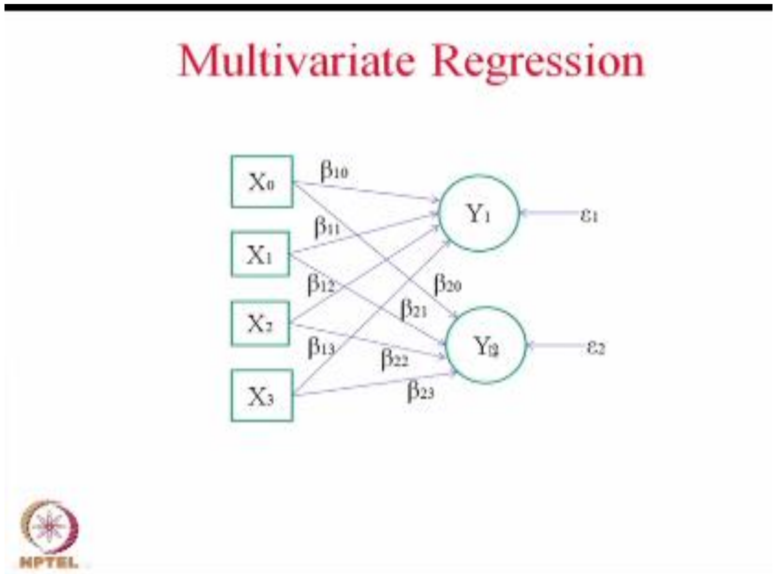
That multivariate regression is preferable. Now the question comes if these variables are also correlated and the independent variables are also correlated or if there is even the causal relationship between that Y1 and Y2 so what will happen? So ultimately in order to incorporate the covariance relationship.

(Refer Slide Time: 04:38)



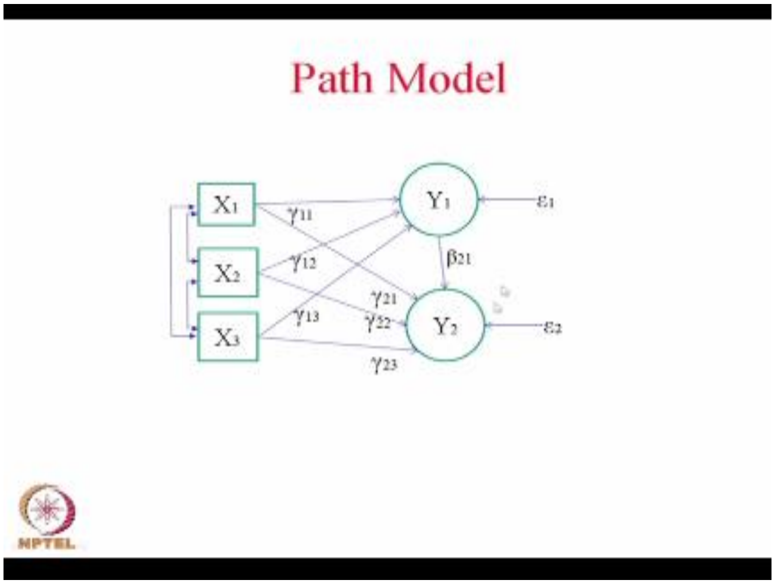
Between the X variables, so you create something like this where X1,X2 and X3 these three are the independent variable Y1, Y2 are dependent variables but the difference of this structure with multivariate regression structure is that.

(Refer Slide Time: 05: 00)



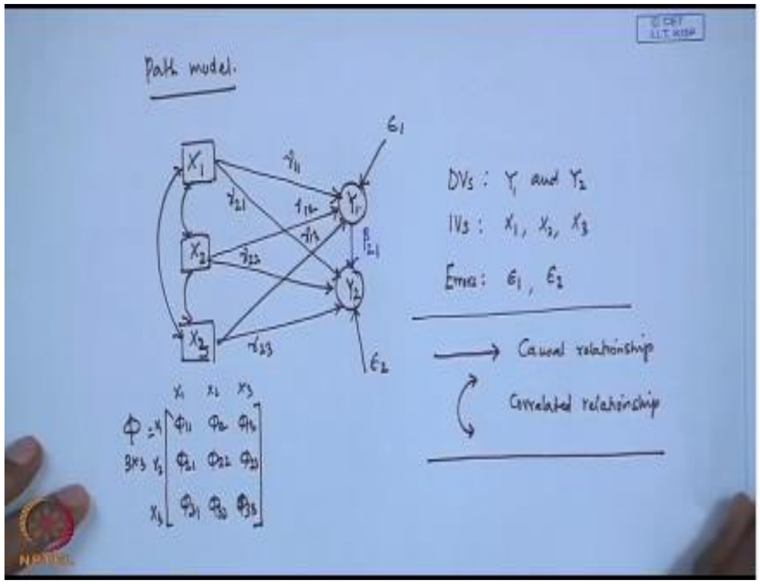
In multivariate regression you are assuming that X1, X2, X3 truly independent they are not correlated and there is no causal relationship between Y1, Y2 also.

(Refer Slide Time: 05:12)



But in path model what you are making, you are considering that X1, X2 as well as X2, X3 or X1 X3 they can go high and in addition what we are seeing that Y2 can be affected by Y1 also in addition to the effect of X1 to X2 to X3, okay. Now let us see equation for this path model. So we are deriving a equation for from path model.

(Refer Slide Time: 05:51)



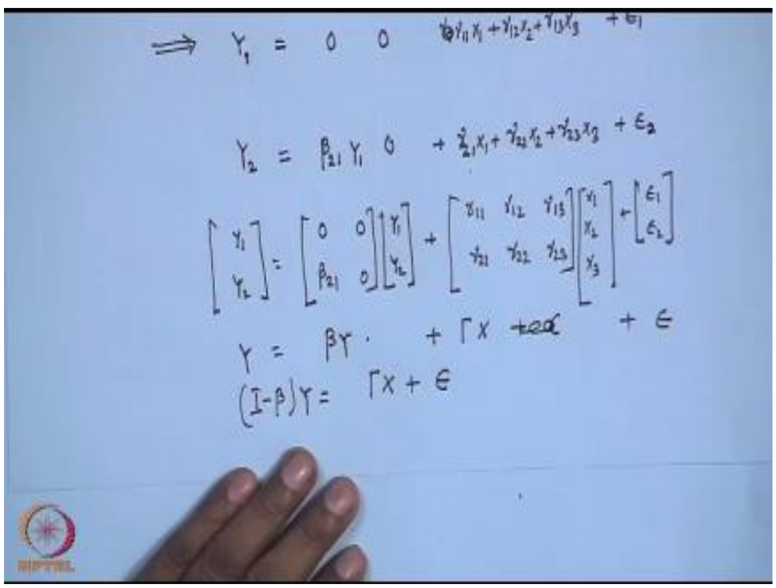
Given like this X1 is a independent variable in this model but correlated weight other independent variables. So that is basically this correlation can be captured through a matrix which is ϕ, If we can say this 3 ϕ is a 3 cross matrix we can write like this, so X1, X2, X3, X1, X2, X3, it is ϕ_{11} then we can write ϕ_{21} we can write ϕ_{22} and we can write ϕ_{31}, ϕ_{32} and ϕ_{33}. Definitely this is a symmetric matrix and this will be ϕ_{12} which is basically $\phi_{21} = \phi_{21}$. Then $\phi_{= 13} = \phi_{31}$ and $\phi_{23} = \phi_{32}$.

And in case of multivariate and multiple regression we have assumed that this is structure of diagonal elements are 0, okay. Now we have considered here two independent variables Y1 and Y2 and we are assuming that X1, X2 and X3 they are affecting Y1 with coefficient like this γ1 is affected by 1 γ1 is affected by2, γ1 is affected by 3. Similarly from independent side point of view.

This γ2 is affected by 1 that means Y2 affected by Y1, Y2 affected by X2 and Y2 affected byX3.Let me repeat here Y2 affected by X1 that is coefficient is γ21, Y2 affected by X2 coefficient is γ22, Y2 affected by X, coefficient γ23. As these are regression lines like equations so this is having an error term. Now this also will have an error term, in addition here we are assuming that this Y2 is also affected by Y1. If this is the case you have to give a symbol like ß, Y22 and Y11 and this ß 21. So if you want to write an equation for this, then there are how many variables are there, How many variables are there, there are different types. One set is DVs that is Y1 and Y2, another set is IVs that is X1, X2, and X3, another set is errors that are ε1 and ε2. Then what types of relationships are available here?

One relationship is like this, if you see that X1 to Y1 this single arrow this is known as causal relationship. Another type of relationship with this giving here like this card bidirectional this is known as correlated relationship, okay. So these two types of relationships are available in this these two types of relationships are estimated in path model, okay.

(Refer Slide Time: 10:47)

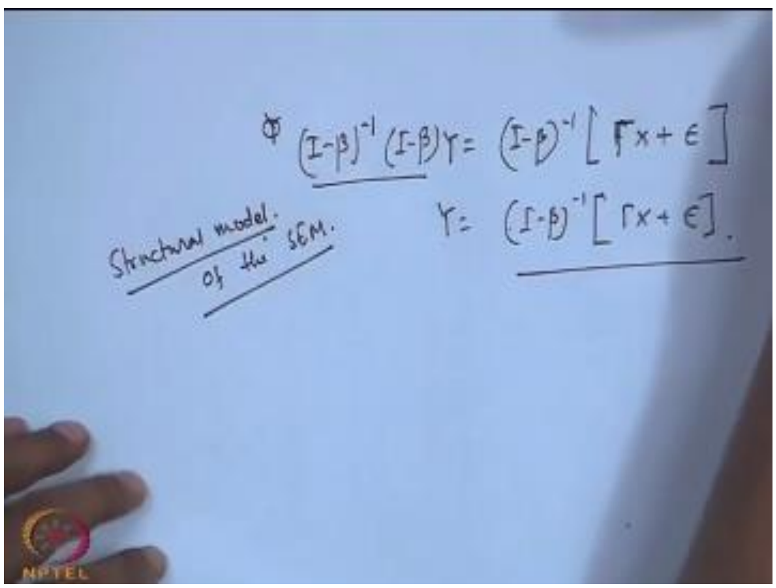


Now, if you want to write down the equations for these now straightway we can write one is DVs I can write Y1 and definitely Y2 is there this side also you can write DVs you can write IVs you can write IVs. How many DVs are there? Two DVs, Y1 and Y2, X1, X2, X3, okay then errors are definitely this. So the equations will be like this Y1 equal to that Y1 is not dependent on that Y1 again it is not like this.

Similarly, if I see that Y1 is dependent on X1, X2 and X3 and this arrow term, not Y2 so nothing will be written here. It will be straightway γ11 then X1 then plus γ12 X2 + γ13 X3 + ε. This is the regression equation forY1. Similarly for Y2 if we see that Y2 is also affected by Y1 so we can write here that ß2 is affected by 1, Y1 + γ21 X1 γ22 X2 γ23 X3 + ε2. So in matrix form you can write Y1 = Y2.

Then this side you see here it is nothing means 0, 0, 0 so ß21, 0 definitely again you are writing Y1 and Y2 which is the difference from regression model and then here you have γ11, γ12, γ13, γ21, γ22, γ23 into X1, X2, X3 plus you are getting ε1 ε2. This is your path model or Y = ßY + γX + ε, ßY + γX + ε. If you do little more manipulation then this will be i- ßY = γX + ε, okay.

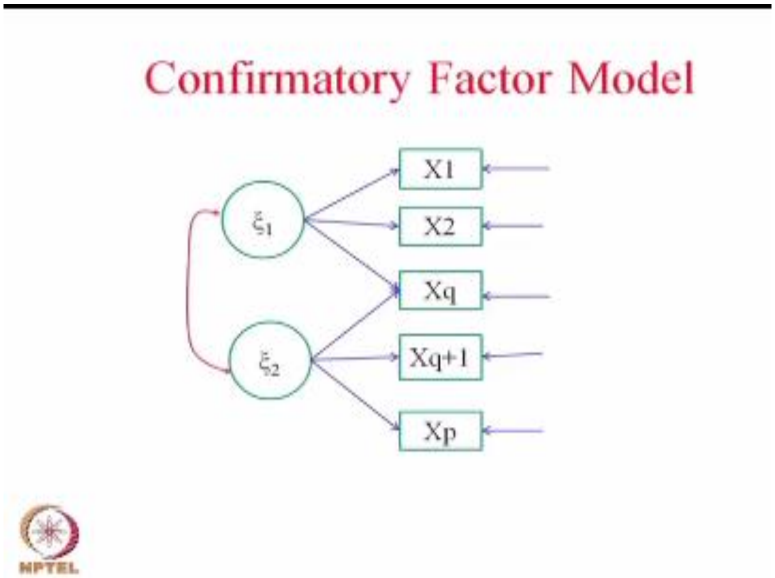
(Refer Slide Time: 14:14)



Then further you may be interested to write this that Y if I multiply Y – ß$^{-1}$ where with I - ß Y, then I- ß$^{-1}$ you are getting that γX + ε and this will be I. So Y = I – ß$^{-1}$ F(X) + ε. The details of this how to estimate this parameter ß, ϕ and all those things, ε ϕ all those things will be discussed under structural equation model, structural model under structural model of the structural equation modeling, okay.

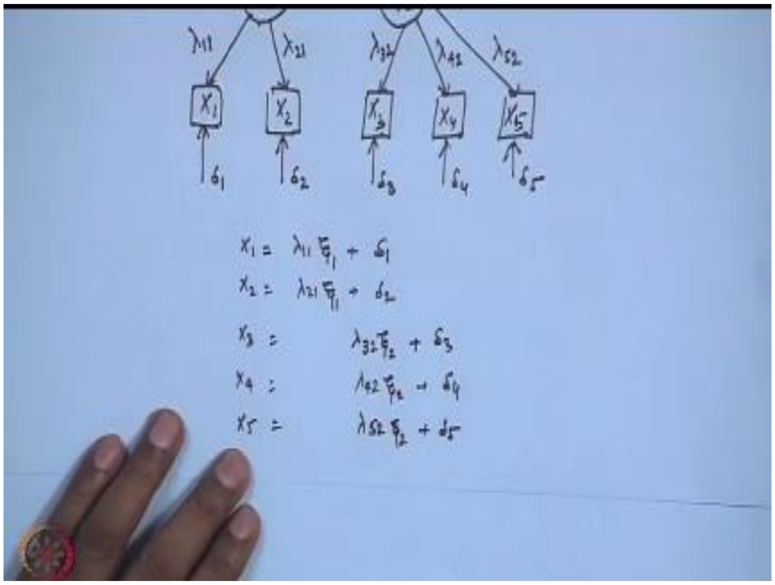
So let us assume that we know this for the time being that this correlation is possible to understand.

(Refer Slide Time: 15:22)



Then you come to confirmatory factor model. I think you can collect yesterday's lecture will particularly I think last two lectures were on cluster analysis before that.

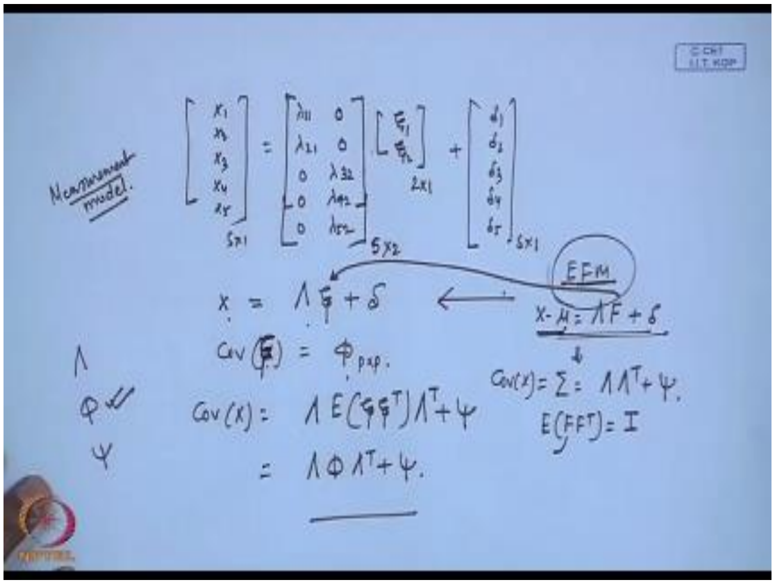
(Refer Slide Time: 15:37)



Then if I write like this what you are doing, then here in this case let us concentrate on only two factors and let they are related with X1, first one is related with X1, X2 maybe second one is related with X3, X4 and X5. There is definitely covariance between the two that is ϕ 21 and each of the observed variable X1 are caused by ϕ1, these are caused by ϕ2, ϕ1 and ϕ 2. Now if I write this one is λ.

Then for X1 I am writing 1 for ϕ 1. I am writing 11X1 is affected by ϕ 1, then λ 2 X2affected by ϕ 2, then λ 32, then λ42, then λ 52. So this one will be δ 1, this will be δ 2 and this will be δ3 this will be δ4 this will be δ5 and this is what you have seen in confirmatory factor analysis we have written this one in this manner that X1= λ11 ϕ 1 + δ1,X2 = λ21 ϕ1 + δ 2, X3 = λ32 ϕ 2 + δ 3 and X4 = λ 42 ϕ 2 + δ 4, X5 = λ 52 ϕ 2 + δ5.

(Refer Slide Time: 18:07)



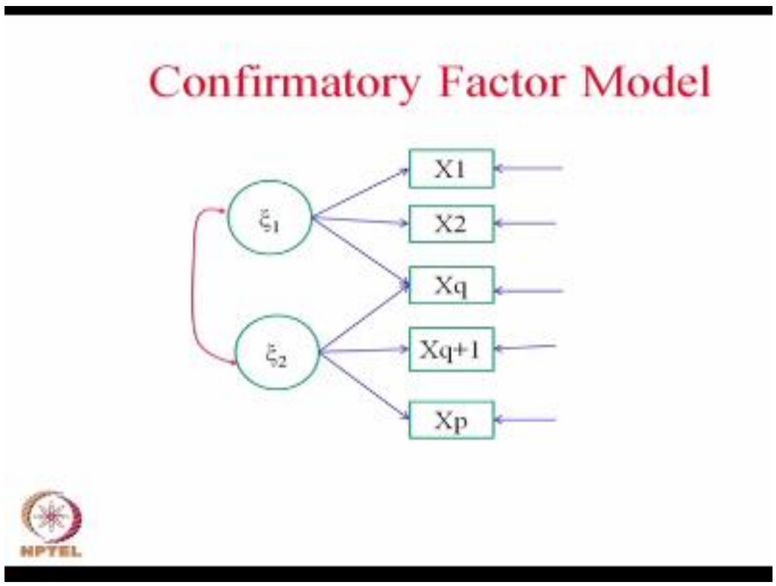
So the resultant matrix you have written like this. So X1, X2, X3, X4, X5 this is equal to your ϕ 1, ϕ2 is there, so we write down λ11, 0, λ 21, 0 then 0, λ 32, 0 λ42, 0, λ 52. So this is 5 x 1. This is your 5 x 2 and then ϕ 1 and ϕ 2 this is 2 x 1 and plus δ1, δ2, δ3, δ4, δ5 this is 5 x 1.The resultant you can write it ax = this is your λ this one is ϕ plus we can write this as δ, okay and if you recollect called exploratory factor model.

What we have written there is X - μ = λF + δ. It is coming similarly only that -μ is not there but here also we assume that this subtraction is made. So actually what happened is that you are getting the same factor model there is one difference in this case the difference is in this exploratory factor model, what you found out, when you found out that covariance of X which is σ that one is $\lambda \lambda^{-1} + \phi$.

Because we assume that expected value of $F F^{-1}$ this is I that covariance between the factors. In this case we are assuming that covariance of X is ϕ depending on, so there are p variables it will be p + p. So as a result the covariance here we are assuming covariance of ϕ I am extremely sorry covariance of ϕ this one correspondence to this corresponds to this covariance of ϕ is this. Then covariance of this will be λ expected value of $\phi \phi^{-1} \lambda + \phi$, okay.

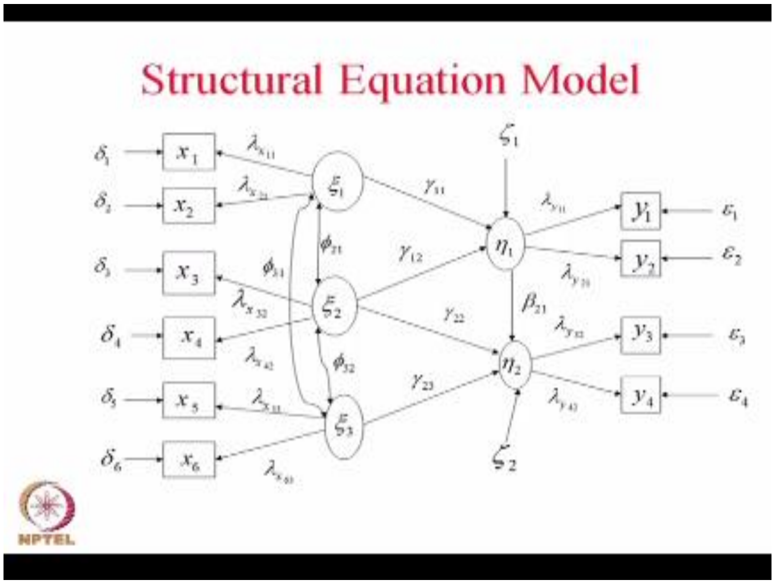
So for the time being we stick to this only. Now question is how to estimate this λ ϕ and ϕ and all those things related to exploratory factor analysis we have seen that how this can be estimated but in exploratory initial direction this ϕ matrix was not there. Here ϕ matrix is there also we will be discussing this under measurement model.

(Refer Slide Time: 21:53)



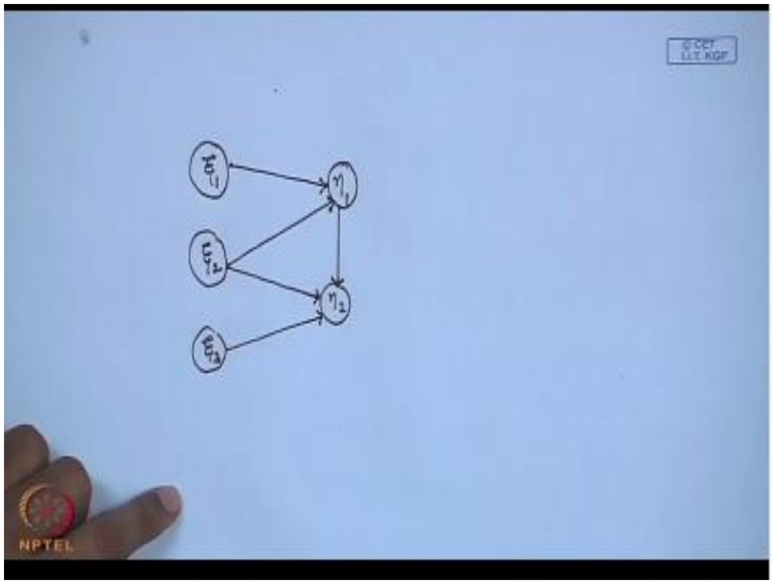
So these two are very essential for structural equation modeling.

(Refer Slide Time: 21:55)



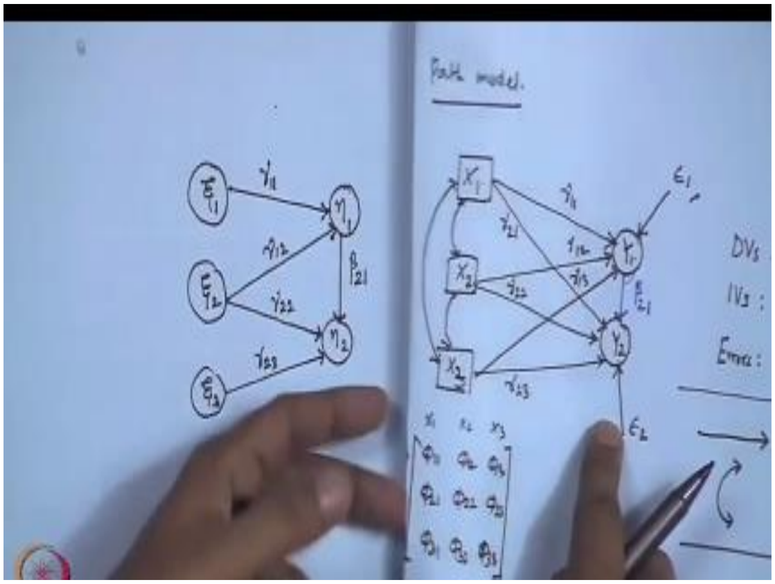
Now see one big picture where we are doing like this. First let us think of that there is one variable η1another variable η2.

(Refer Slide Time: 22:15)



And these two variables are affected by some variables called ϕ1, ϕ2, and ϕ 3, okay. Let the link is like this η1 is affected by ϕ1 as well as ϕ2 and η2 is affected by ϕ2as well as ϕ3 and η2 is also affected by η1. So under this condition so I can if you see the structure path model equation, so in path model equation let me find out this path model.

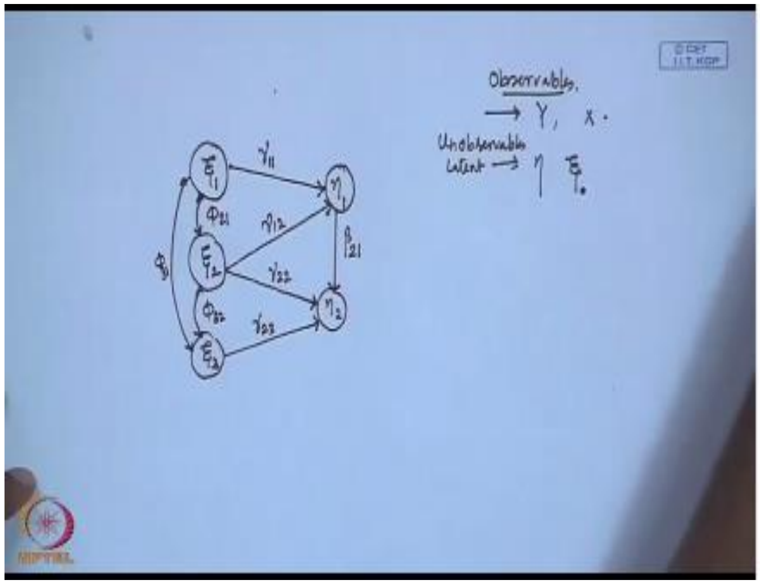
(Refer Slide Time: 23:14)



Yes this is our path model equation. So if I keep side-by-side this to side-by-side, so what is happening here is you see we are writing η1 in place of Y1, η2 in place of Y2, ϕ1 in place of X1, ϕ2 in place of X2, ϕ3 in place of X3 and here everything is linked with all the X variables are linked by Y. Here we are restricting that there were that η1 is affected by ϕ1 and ϕ2 and η 2 is affected by these two.

So there has been this structure is similar to the structure except the some of the links are omitted, okay. So I can write this that γ1 η1 is affected by ϕ1 means γ11. So this one is γ12 then this is γ22, this is γ23 and this one is ß2 affected by 1. The coefficient as usual we are writing here, okay.

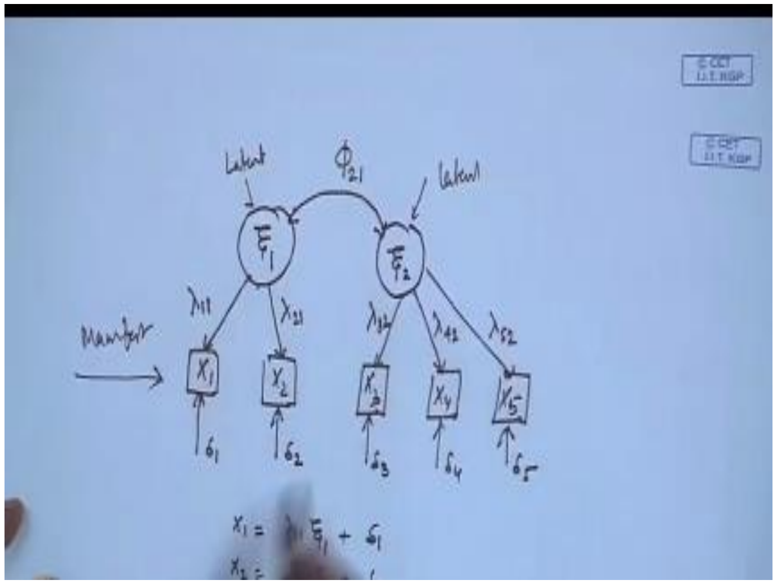
(Refer Slide Time: 24:15)



Now the difference between this path model what I have shown I have shown in path model in terms of Y and X, when we talk about in terms of Y and X we say they are observable, okay. Now when we are saying in terms η and ϕ they are latent and unobservable or latent that is the difference. So if this is the case then what you require to do? You require measuring these unobservable or latent things, okay.

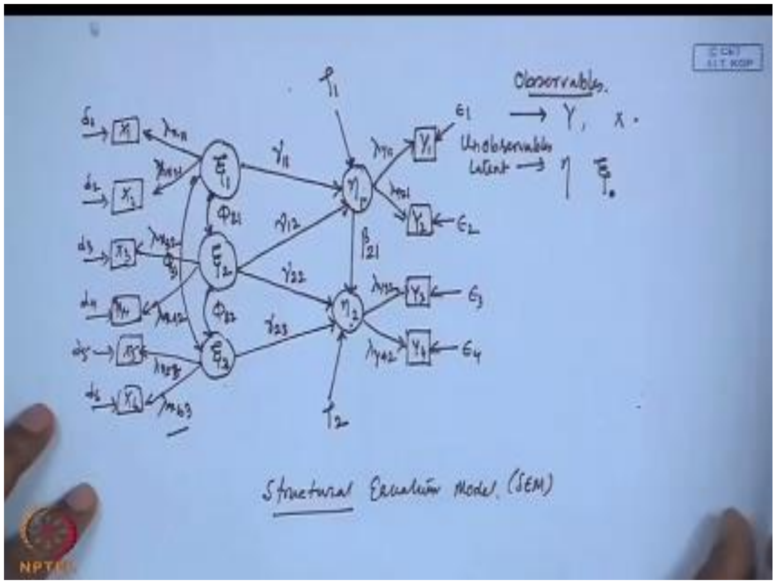
We will see this but before that some more things that these are cohering, let me write. This is nothing but $\phi21$, this is my $\phi32$ and this is my $\phi31$. These structures are there, okay. So as I told you that this η and your ϕ they are immeasurable, unobservable, latent we require measuring this. How do you measure? We have seen earlier that our confirmatory factor analysis or exploratory factor analysis will help us to measure. What it is?

(Refer Slide Time: 25:50)



These are latent this is latent variable, this is also latent and these are manifest variables. So these manifest variables are used to measure these latent variables.

(Refer Slide Time: 26:07)



Now we will do the same thing here. We may create here like this, let we denote it is as Y1 this one isY2 here also Y3 and Y4 and definitely there will be error, that error will be suppose $\varepsilon1$ this error is $\varepsilon2$, this error is $\varepsilon3$ and this error is $\varepsilon4$,error head. Please give it properly. Then what is the difference between this structure and this structure. If I place these to linking to three indicators linking to two indicators, variables here $\eta1$ is linked with two indicator Y variables. H 2 is linked with two Y variables.

What we are saying these two are latent, these are the manifest variables. This $\eta1$ and $\eta2$ can be measured using this. This is nothing but a confirmatory factor analysis this is what confirmed confirmatory factor model is. So, you can give a notation here. So, we have given λ I think, let us use one another subscript Y and this is 11 then this one will λ Y2 is affected by 1 then this will be λ Y3 is affected by 2.

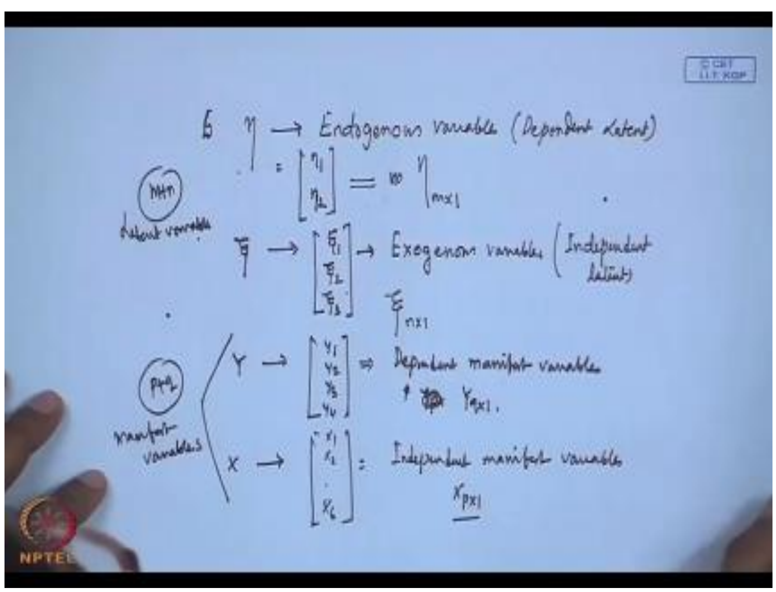
This is λY4 is affected by 2. So we also require to measure this. So let us assume that this can be measured by indicated manifest variable X1 then your X2 then this is your X3, this one is X4, then your X5, X6 and this side $\delta1$, this is your $\delta2$, this is $\delta3$, $\delta4$, $\delta5$ and $\delta6$.Now if you look into this side what is there? Is there any difference between this, between this and this here $\phi1$, $\phi2$?

You have taken three latent variables and these are the indicator five and indicators two are linked with $\phi_1$and three linked with ϕ_2. Here there are sixteen indicators two each are linked withϕ_1 and ϕ_2. So, that means this is also a factor analysis part confirmatory part. So then this one we can write like this λ we are taking as X11 this is your $\lambda X21$, this is your λ X32, $\lambda X42$, λ X53, λ X63, okay. What is missing here?

These are from confirmatory factor point abilities, these are the dependent and these are the causes here these depend on these causes. So arrow head arrow terms are given, arrow terms are given but here from the path model point of view these are the dependent side and this is independent. So you have to give some η, okay. This is what is known as in path diagram this is known as structural equation model.

Why it is structural equation model because the latent variables they also have certain structure some structural relations which they are. For example $\phi_1$is affecting η_1 but η_1 is affecting η_2. So there is a structural path from $\phi_1$to ϕ_7 even though here ϕ_1 is not directly affecting η_2 but ϕ_1 can have effect through η_1 to η_2. So that means that the relationship structure is correctly preserved. So that is why this name is structural equation modeling is given here, okay.

(Refer Slide Time: 31:37)

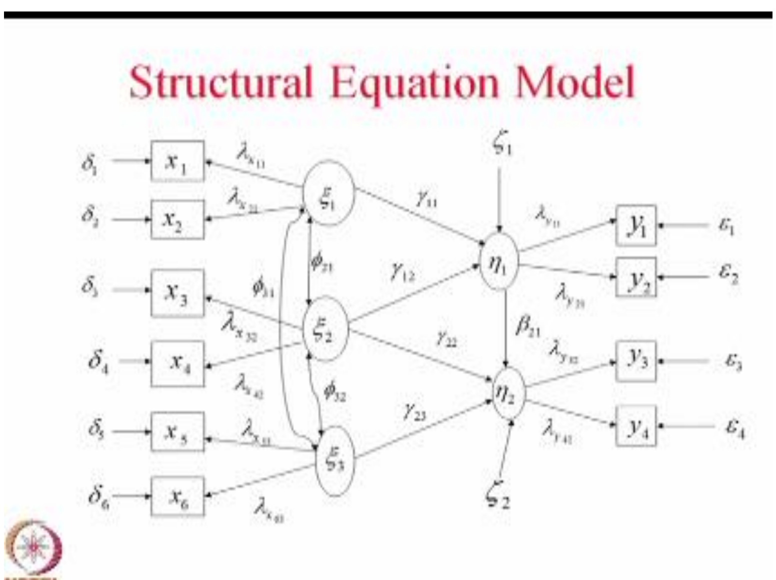


Now how to have the estimation of this? This is a big issue and we will be describing later on that all those things but here we will want to give you some names in the structure. I want to give you some names of the variables, in structural equation modeling we will use the term like these η and all η will be termed as endogenous variable. Endogenous, this is basically dependent latent, in the example given this one each this η this is basically η1 and η2.

Now there is another variable which is ϕ which is in this example this is ϕ1, ϕ2, ϕ3,which is known as exogenous variable, exogenous variables. This is independent latent, okay. So in general, there may be your η can be n x 1,ϕ can be n x 1 those many things will be there or any other means you can use something else.

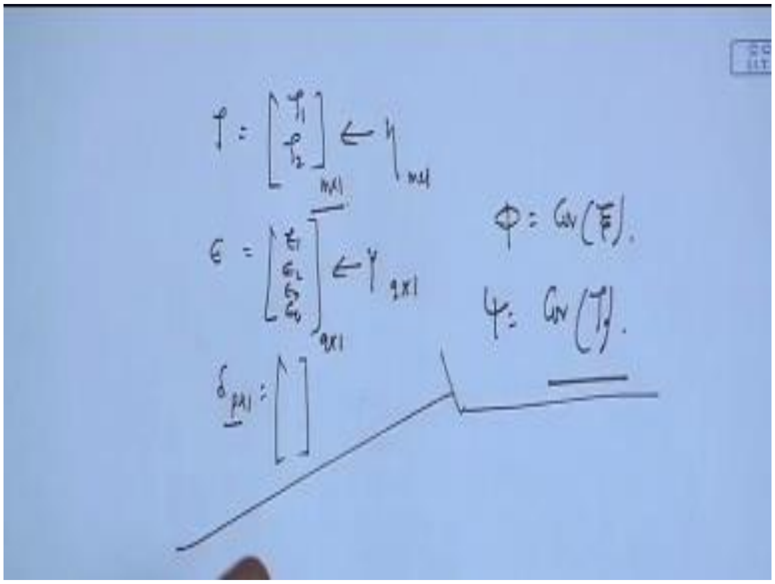
Then we have Y which is in our case in this example Y1, Y2, Y3 and Y4. So this is known as dependent manifest variable. It can be p, so y can be q x 1. Similarly you have X here we have X1, X2 to X6. These are independent manifest variables it can be X that is p x 1. That means in totality you will be having p + q manifest variables and m + n latent variables. In addition you have other variables like ϕ.

(Refer Slide Time: 34:35)



Other variable in this case for example is zη.

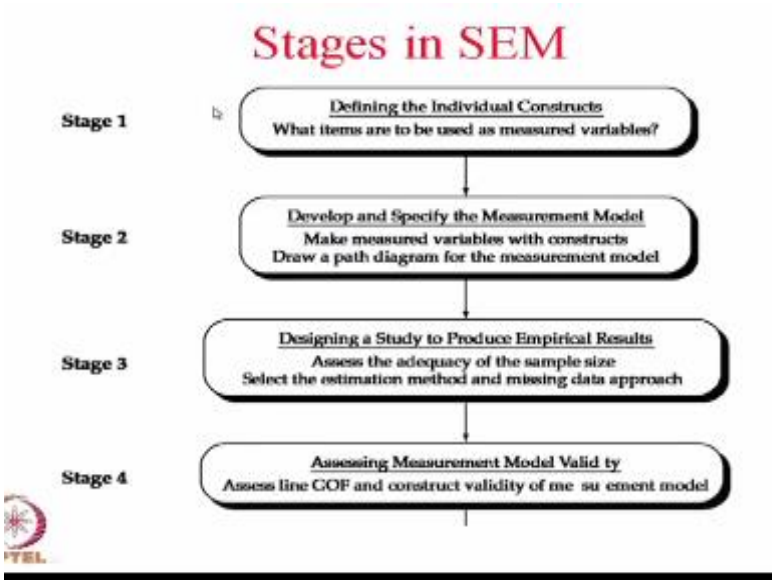
(Refer Slide Time: 34:39)



Zη is nothing but the error terms. In this case, it is zη 1, zη 2. It is related to η. So depending on the η, if η is m x n then this also be m x 1, m x 1. Then another one is your ε that will be related to ε1 to ε 2 to ε 4 in this case, this is related to your Y. It all depends on what is the value of Y. So this will be same. Similarly related to X, there is δ. The δ is your p x 1 variable vector these three.

Now you have another parameter called ϕ which is basically covariance of this you will also have you will be getting ϕ as your covariance of your zη, okay. So there are a huge every large number of what I can say parameters to be estimated. Let us assume that yes everything is possible and we are going for its stages.

(Refer Slide Time: 36:10)

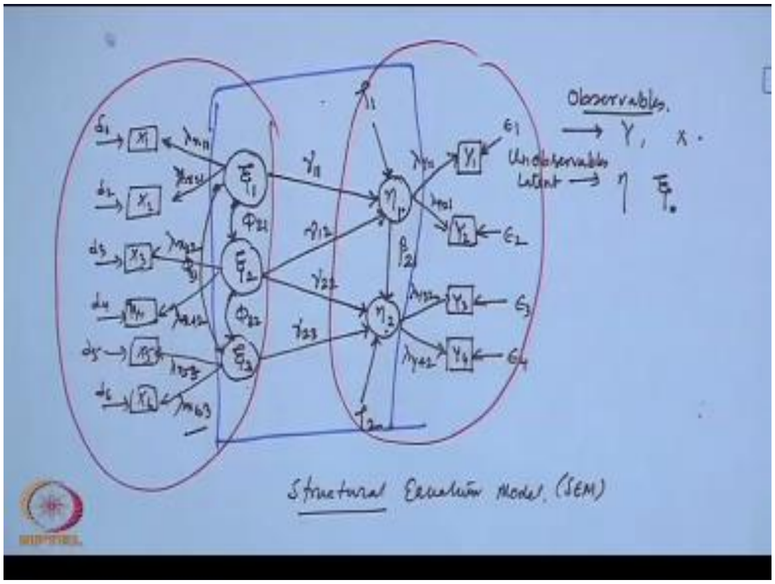


What is stage one is? It is defining the individual constructs what items are to be used as measured variables. Then stage two is develop and specify the measurement model what is the confirmatory factor model, make measured variables construct, draw a path diagram for the measured variables. Then stage three is designing a study to produce empirical results. All multivariate modeling as such any model building case the design study design is very important.

Here what we are saying assess the adequacy of sample size select the estimation method and missing data approach. This is true for multiple regression, multivariable regression and path model factor analysis, compound principle, compound analysis, structural model, cluster model, everywhere it is applicable. Then assessing measurement model validity means these few first four stages.

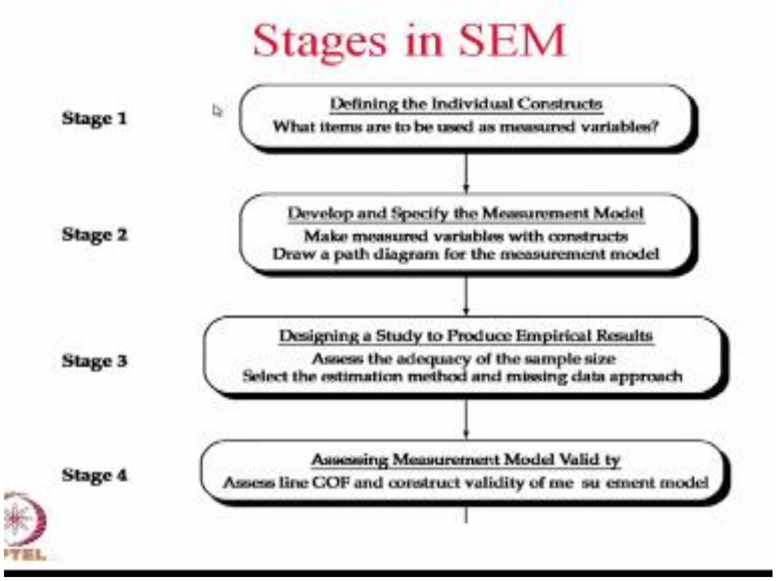
We are talking about the measurement model. Then what is this measurement model? Measurement model is this part.

(Refer Slide Time: 37:30)



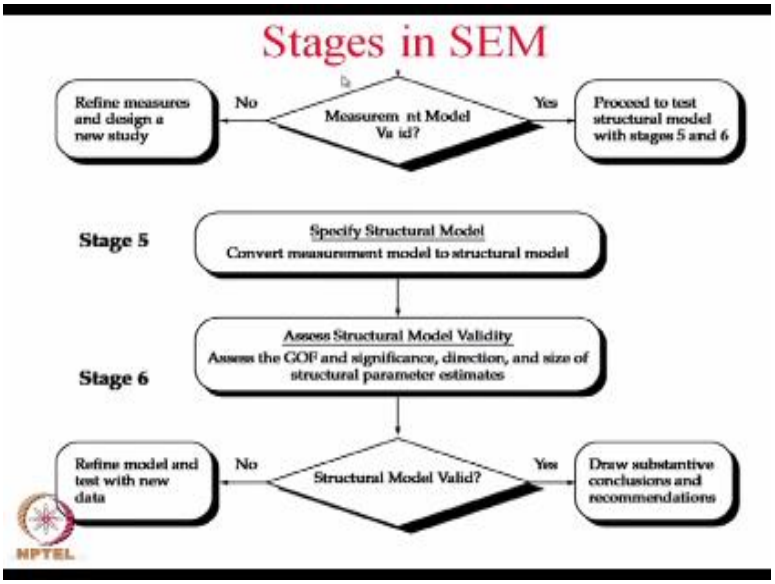
This factor analysis part this is your measurement effect. Then what is your structural side? Structural side is this one. So first what we have it is said that in the stages and find out the manifest variables.

(Refer Slide Time: 37:57)



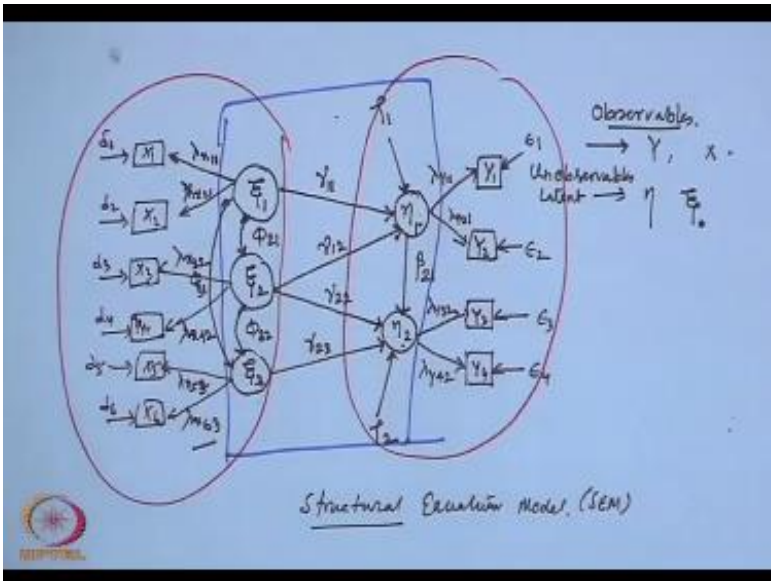
Then develop the constructs and also measure this the adequacy of this model.

(Refer Slide Time: 38:07)



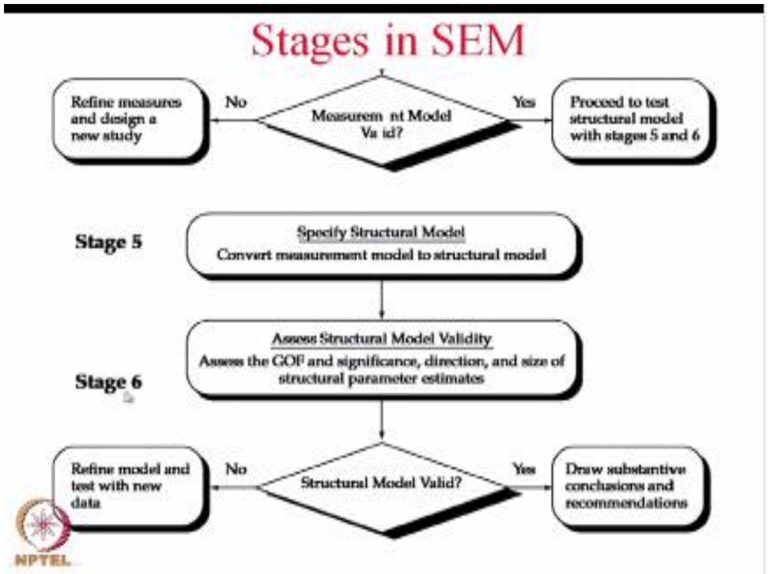
Then you have to see if the measurement model is valid then you have to go through five to six. If the measurement model is invalid then you have to redefine the measures and other things. Now specify then your structural model is coming. Specify the structural model convert measurement model to structural model. I think you have seen here ultimately.

(Refer Slide Time: 38:35)



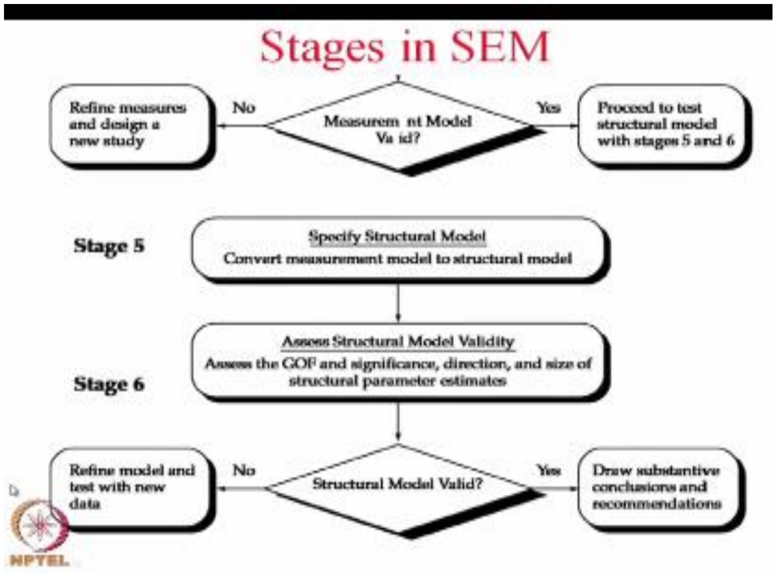
I will repeat this again.

(Refer Slide Time: 38:38)



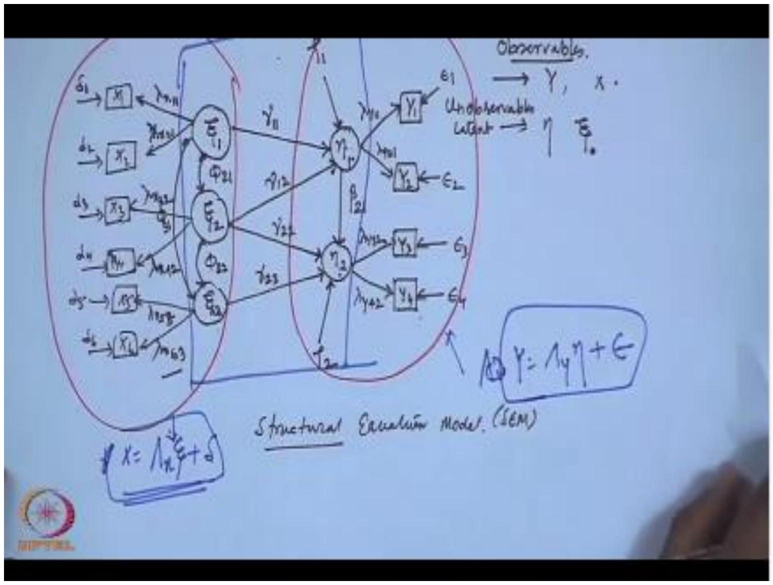
Then stage six is assessing structural model of validity then if structural model is valid then draw substantive conclusion if it is not valid, refine the model and test with a new data, okay. In this case what is it actually given?

(Refer Slide Time: 38:56)



It is given in a two stage model building. One is first you develop the measurement model. First you develop the measurement model.

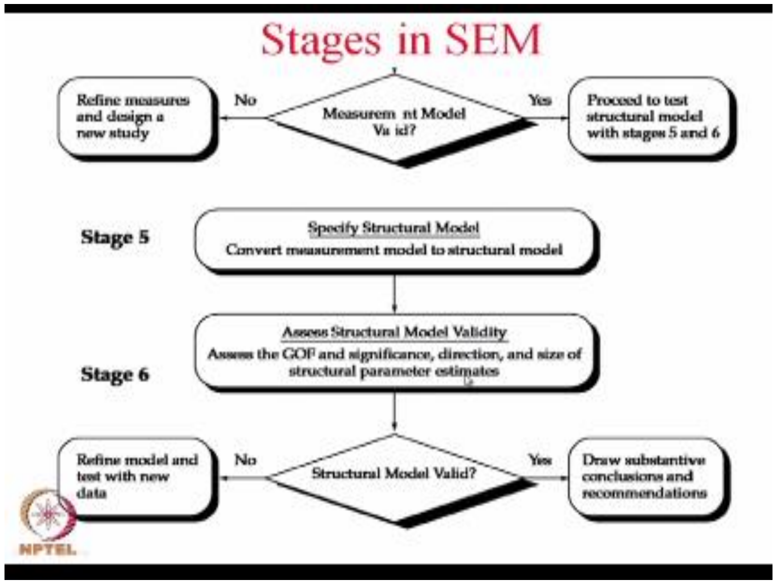
(Refer Slide Time: 39:07)



Irrespective of your in this case forget about the structural part. What you have? You have these many X variables, these many Y variables. Irrespective of η and ϕ, you first go for it on measurement model where you will basically factorize confirmatory you factorize this η 1, η2, ϕ1, ϕ2, ϕ3 from Y1 to Y4 this is the first.

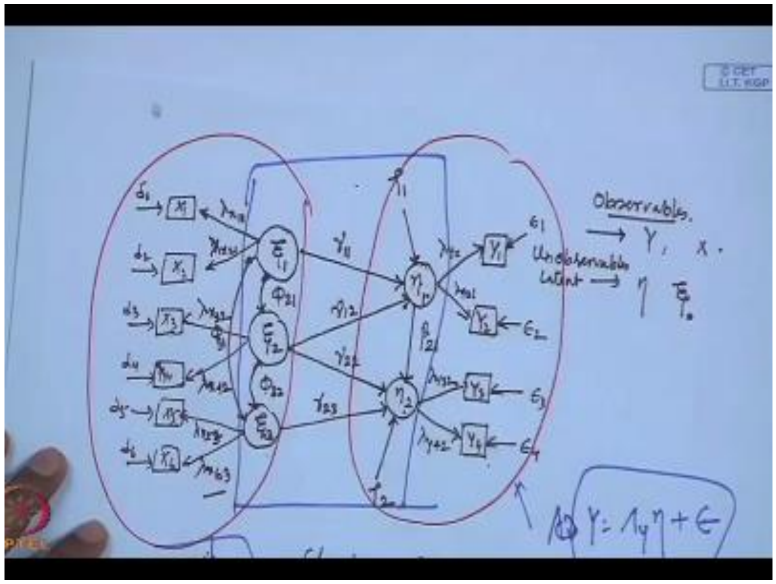
And once you are happy with this, so that means that if I just follow this one this side, it is X is λ x ϕ plus δ. This is my measurement model for this site. This side it will be Y = λy x η +ε. This is my contrary path. But the process of model building stages is that that you go for measurement model. First consider all these and then do a factor confirmatory factor analysis. Find out the correlation or covariance structure of all the latent variables.

(Refer Slide Time: 40:39)



See here. Then once you have all your latent variables, everything is crystal clear and then you go for structural model the structural model is the building between this path.

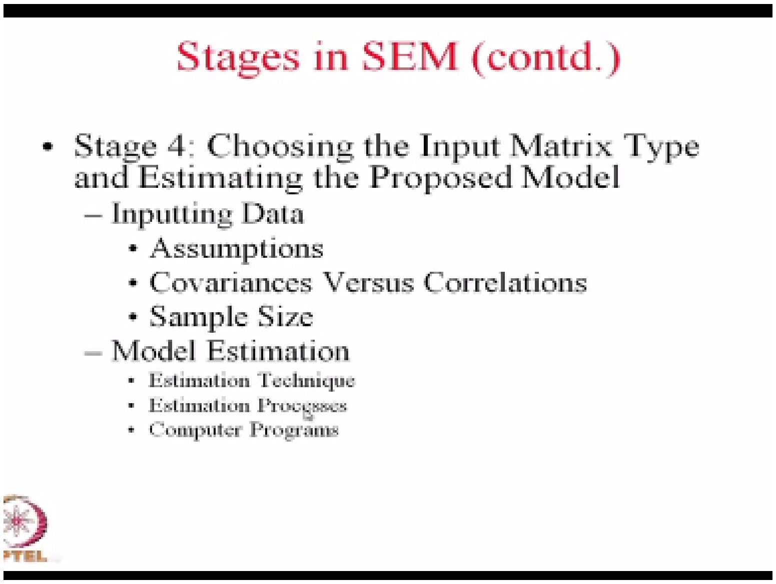
(Refer Slide Time: 40:49)



This one you will do and you will estimate and all those things. So ultimately you will first prepare measurement model then from the measurement model you get the covariance or correlation matrix for the latent variables using this correlation or covariance matrix you will estimate the regression like coefficients regression coefficients for this structural model this is one.

Second one is you may be interested to do everything at a time. That means you will feed the model when both structural and measurement part keeping intact at one go. But because of so many parameters to be estimated and so many computations, permutations, combinations what will happen ultimately, it is highly likely that you may get several of offending estimates and how to delete offending estimates it is an issue also in structural equation modeling.

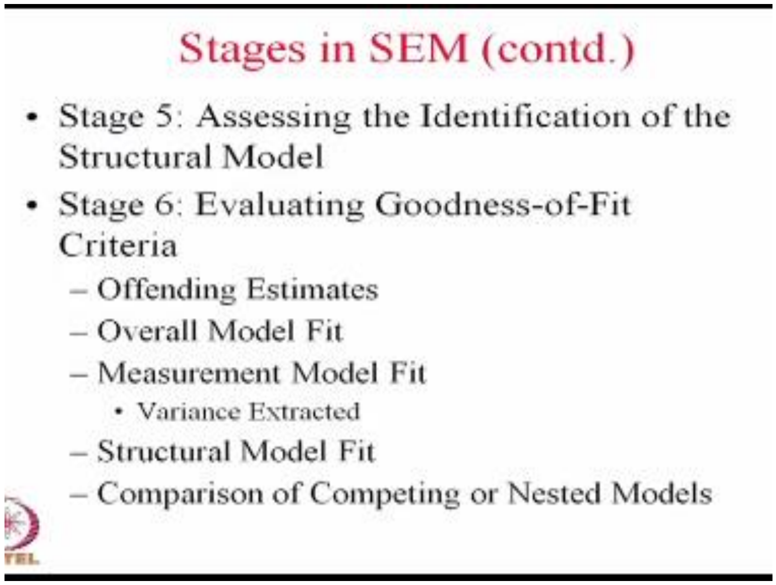
(Refer Slide Time: 42:09)



So now what I will do? This is the structure I have shown you already I have shown you these are the stages that what we have discussed. These equations also you have seen X and Y these equations you have seen then the stage four is that what matrix you are going to use you have to choose one input type of matrix that is correlation or covariance matrix. There are different assumptions.

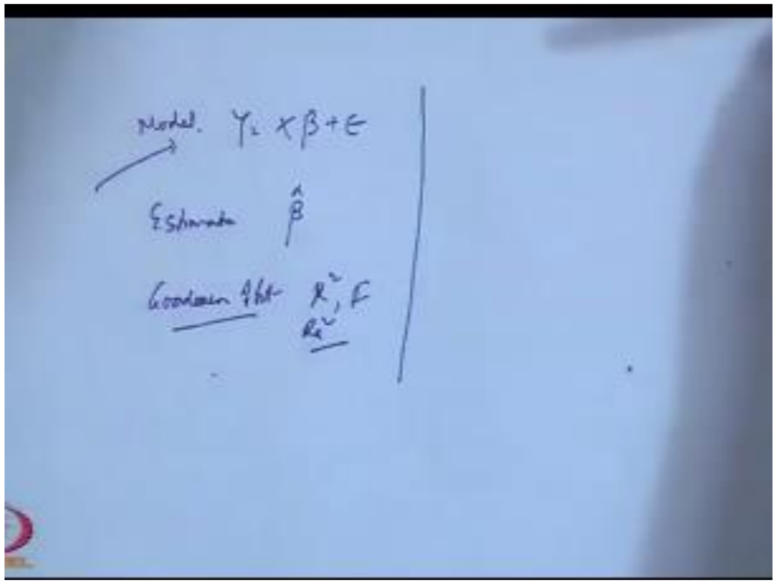
Then there is your estimation technique to be used process, computing programs these are all the stages one by one you have to do.

(Refer Slide Time: 42:30)



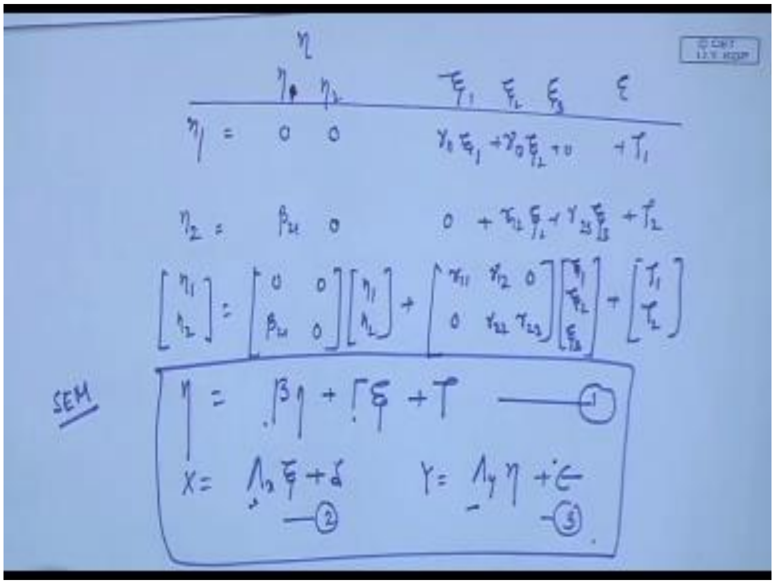
Then suppose this is like your regression model also.

(Refer Slide Time: 42:41)



You have seen that regression equation is y = xß + ε this is my model. Then what you have done you have gone for estimation that is your ß^ then you have done? You have gone for goodness of fit, model adequacy goodness of fit or model adequacy that is R^2 FRe^2 all those things that we have used. So here also if I see that this is my structural equation modeling, then are you not getting a structural equation like this?

(Refer Slide Time: 43:16)

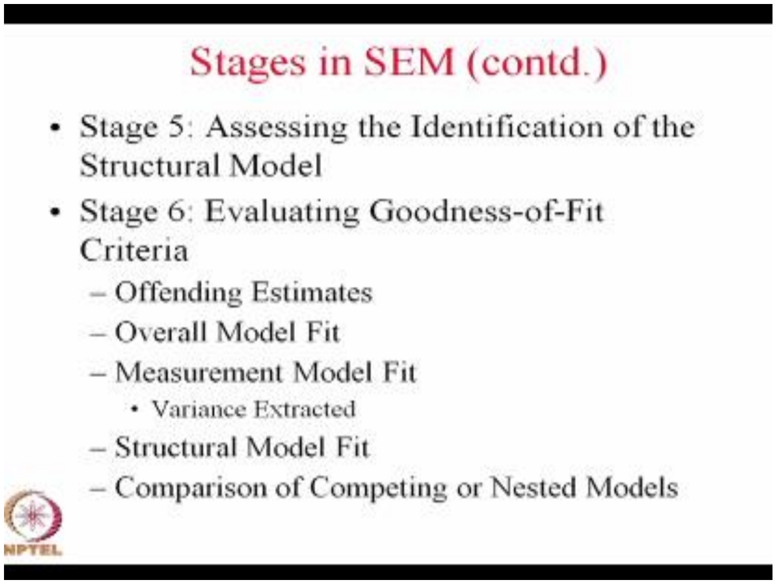


The equation will be your first one that η1 η 2 in this case and definitely here will be η 1 and ϕ, sorry η and ϕ and error that path structure under this η 1 and η 2, if I write η here then this is ϕ1, ϕ 2, ϕ 3 and all those things so you have to write the equations. If I write for this particular model the equations, η1 is not affected by any of the data 0, 0.So, γ11 ϕ 1 + γ 12 ϕ 0 + zη 1 and this one is ß 21 and this is zero.

Then this one is also 0 + γ22ϕ2 + γ23 ϕ3 zη2.Then what is your matrix? It is η1 η2 = to 0, 0 ßη 21 0 η1 η2 plus, you are getting γ11 γ12 0, 0 γ22 γ23 x ϕ1, ϕ2, ϕ3 + zη1 zη2 that is my structural part. So η = ß η + γϕ + zη this is one equation. Another equation is X= λ x ϕ + δ and Y = λ y η + ε. This is equation number two equation number three. So your model this is my structural SEM model, okay.

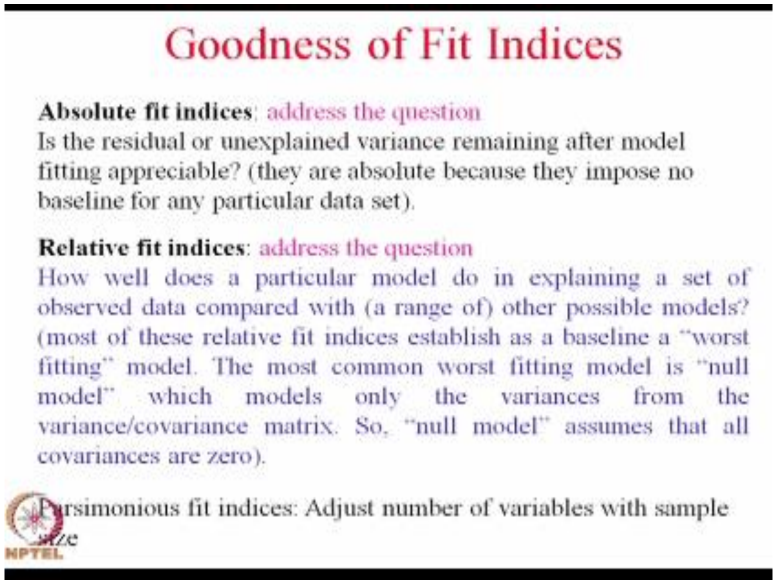
And you require to estimate ß γ then γx, λx, λ y covariance of this covariance between these so many parameters you have to estimate. Second is parameter estimation. So, we are not discussing here the details of parameter estimation and we will discuss later on separately.

(Refer Slide Time: 46:14)



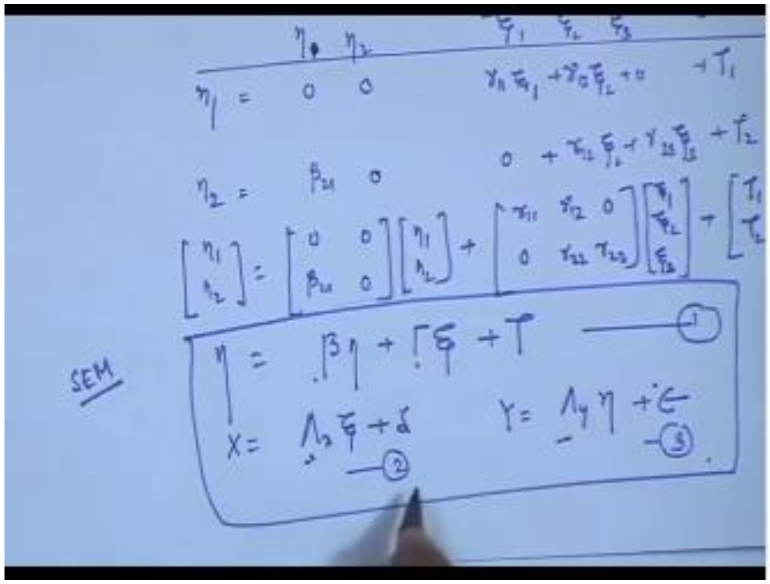
So once you have the model you require to know.

(Refer Slide Time: 46:22)



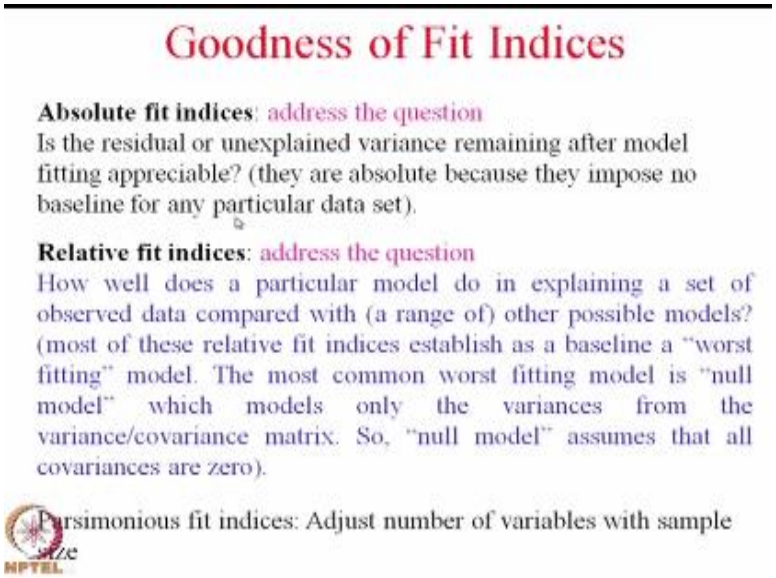
That what will be the goodness of fit measures. So my model is this.

(Refer Slide Time: 46:27)



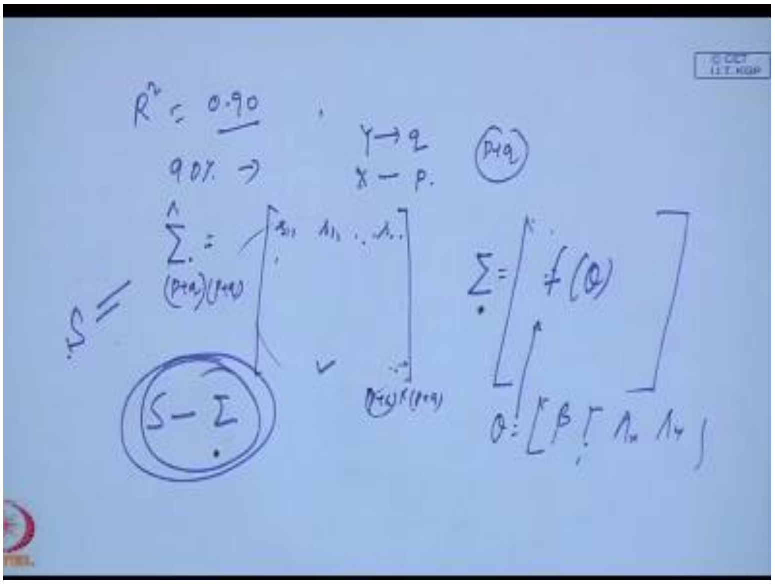
And I know that somewhere I will estimate all those things. Then I have to go for fit measures.

(Refer Slide Time: 46:39)



There are three types of fit measures used in structural equation modeling, one is absolute fit indices, relative fit indices and parsimonious fit indices. In absolute fit indices the question is the residual or unexplained variance remaining after model fitting appreciable? They are absolute because they impose no baseline for any particular data similar to R^2.

(Refer Slide Time: 47:12)

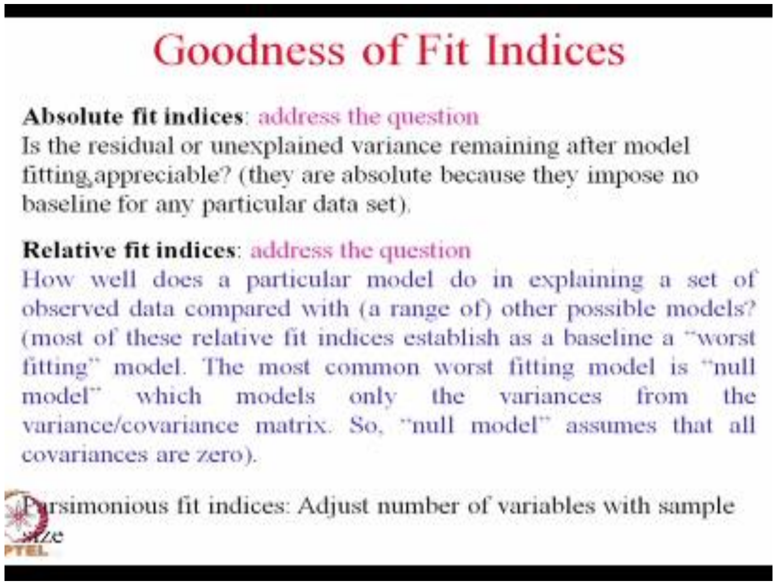


R^2 if it is 0.90, it says 90% variance of Y is explained by this model. Here also we are basically talking about what will happen to this. Actually although we will be describing in detail later on, the estimation what is observed with you is this, this is basically if I how many Y variables Y is q observed variables and Xp observed variable you have total p + q observed variables. So you have a covariance matrix of p + q or p + q. So this is the observed value and this is nothing but this is what is observed.

Now using these equations, this η X and Y this equation you will also be able to estimate or other way I can say fit this one you can calculate this. Here if I use S then it is s11, s12 like s that p, q …… p + q. So similar values we will be getting and here instead of now I am writing instead of writing s here we are writing the population sigma then that will be in terms of this ß γλ X λ Y.

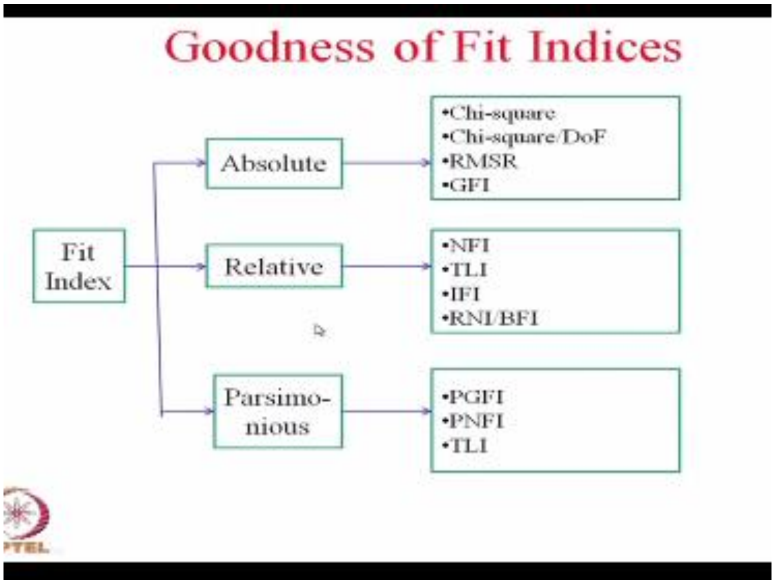
So there will be function of the sorry, so what I mean to say is we can find out the that η in the nut shell I am talking about that ß, γλ X, λ Y like this it is possible to frame this then this s and these are compared there are several ways to compare this but and then here p + q x p + q, this is the end. So how many unique elements are there those will be compared, when we initialize some value here some value here, those will compared. Then we finally get the value. Then based on this difference when we need something like this, this is total.

(Refer Slide Time: 49:43)

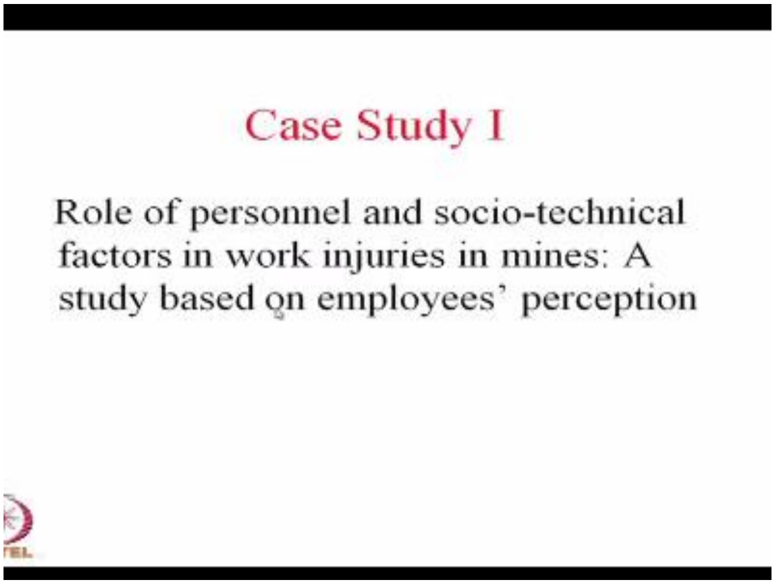


Relative feed index that is the question that how well does a particular model do in explaining a set of observed data compared with other possible models. You can go for several models and then compare which one is your best. The reference one is null model where we assume that there is no covariance value, relationships in the observed variables. Parsimonious fit index is similar to your adjusted R^2 where it is the number of parameters estimated are adjusted.

(Refer Slide Time: 50:21)

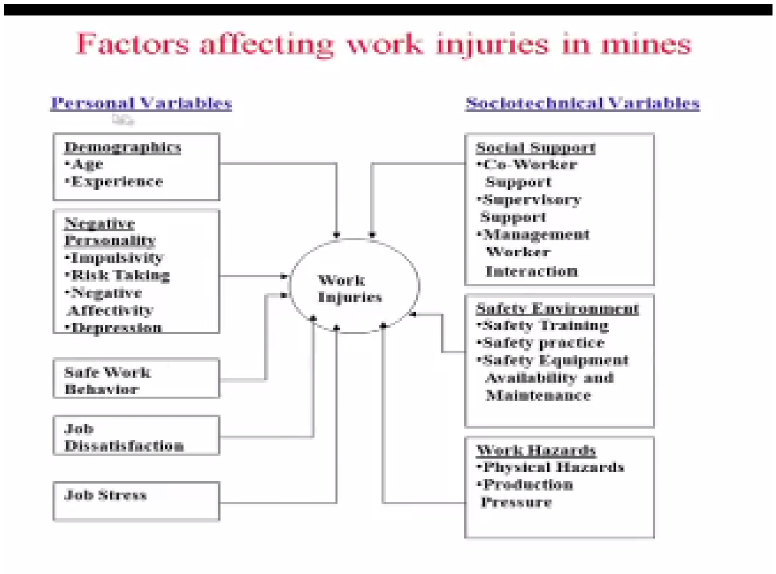


(Refer Slide Time: 50:26)



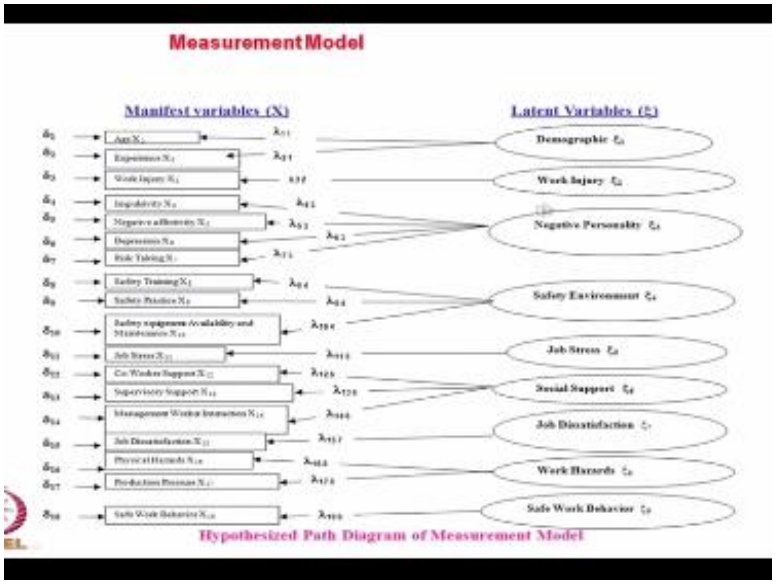
There are different fit indices I will show you one case here by five minutes of time. This is role of personnel and social technical factors in work injuries in mines a study based on employees' perception this is published in organic general in 2008.

(Refer Slide Time: 50:42)



So we have captured several variables under personal and social technical variables. Then what we have done?

(Refer Slide Time: 50:53)



We have first gone for measurement model all the manifest variables are used simultaneously and then the latent variables are first latent variables are identified and link to them was done.

(Refer Slide Time: 51:23)

Correlations amongst constructs

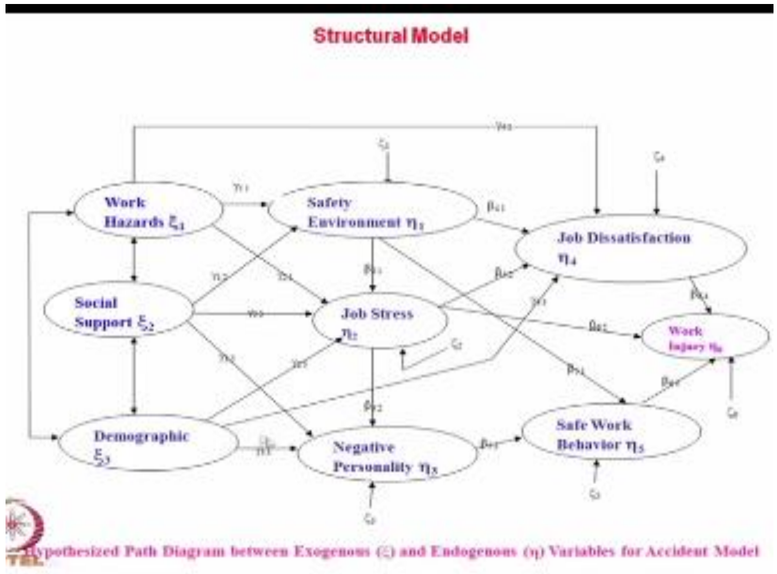
Structural Correlations among Latent Variables Presented in the Measurement Model for Injury/Accident Causation

Demographic	1.00								
Work injury	0.29*	1.00							
Negative personality	-0.10*	0.41*	1.00						
Safety environment	0.04	-0.42*	-0.94*	1.00					
Job stress	-0.09	0.17*	0.86*	-0.73*	1.00				
Social support	0.06	-0.30*	-0.91*	0.83*	-0.75*	1.00			
Job dissatisfaction	0.01	0.31*	0.65*	-0.75*	0.62*	-0.70*	1.00		
Work hazards	0.17*	0.30*	0.67*	-0.77*	0.63*	-0.78*	0.73*	1.00	
Safe work behavior	0.04	-0.22*	-0.51*	0.49*	-0.26*	0.48*	-0.29*	-0.26*	1.00

* indicates 0.01 probability level of significance

Then confirmatory factor was done, the output of the confirmatory model, this is the correlation matrix, this is what the output is? Once we get this output this is input to our structural model.

(Refer Slide Time: 51:30)



This is the first phase where you have to identify the manifest variable and the latent constructs. Then you do confirmatory factor analysis find out the correlation or covariance structure of the latent variables because in SEM your input is either correlation covariance structure. So for the measurement model you give the covariance structure they will give you the latent correlation or covariance structure.

Now for structural model again use the latent correlation or covariance structure whatever you are getting from the latent variables and then you feed like this. So you will be getting picture like this.

(Refer Slide Time: 52:18)

Design of questionnaire for the identified factors

Identified factors	No. of questions	Sample question
Risk taking	11	Are you ready to take risk in order to increase your income?
Negative affectivity	15	Do you become disappointed easily when your thinking does not match with other?
Job Dissatisfaction	13	Do you think your life has become burden to you because of this job?
Impulsivity	12	Do you get excited when any new idea comes in your mind?
Depression	5	Have you got depressed when you are unsuccessful in your work?
Job stress	12	Do you face any problem to complete the excess amount of work hurrily?
Safety training	8	Do you think that the safety training facilities provided to you are adequate?
Safety practice	27	Do supervisor always check that miners wear the special shoes given by the company before going to the underground?

Then definitely these are the some issues what we may be discussing later on if time permits in other classes.

(Refer Slide Time: 52:31)

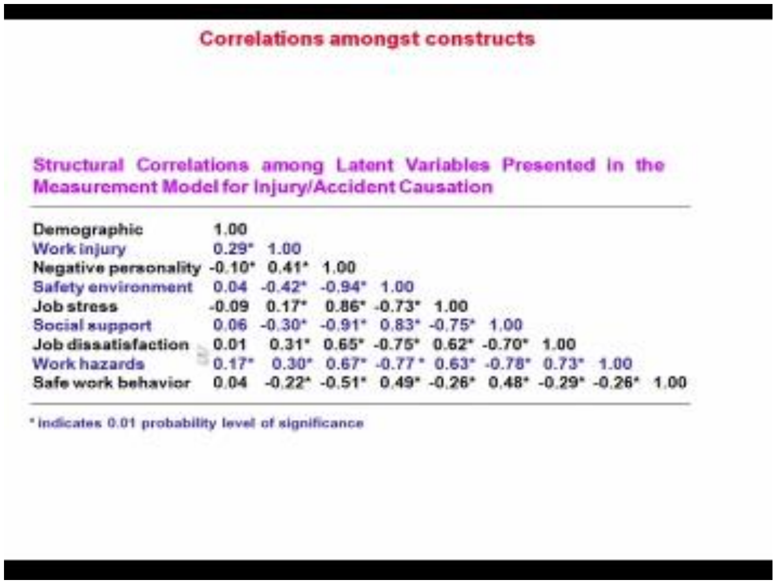
Measurement Model Fit Indices

Goodness-of-Fit Indices for the Measurement Model of Causal Accident Model

Parameter	Values
Chi-square with 99 degree of freedom	257.24
Root Mean Square Residual (RMR)	0.06
Goodness of Fit Index (GFI)	0.98
Normed Fit Index (NFI)	0.97
Comparative Fit Index (CFI)	0.99
Incremental Fit Index (CFI)	0.99

Ultimately this is what is the observed correlation this is the factor and measurement model. This is the fit of the measurement model that I told you that I will be discussing fully in the next class.

(Refer Slide Time: 52:38)



Correlations amongst constructs

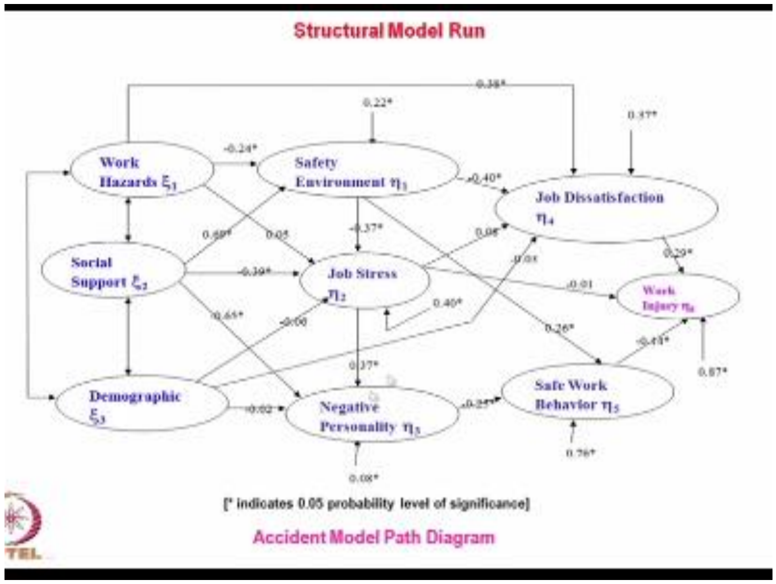
Structural Correlations among Latent Variables Presented in the Measurement Model for Injury/Accident Causation

Demographic	1.00								
Work injury	0.29*	1.00							
Negative personality	-0.10*	0.41*	1.00						
Safety environment	0.04	-0.42*	-0.94*	1.00					
Job stress	-0.09	0.17*	0.86*	-0.73*	1.00				
Social support	0.06	-0.30*	-0.91*	0.83*	-0.75*	1.00			
Job dissatisfaction	0.01	0.31*	0.65*	-0.75*	0.62*	-0.70*	1.00		
Work hazards	0.17*	0.30*	0.67*	-0.77 *	0.63*	-0.78*	0.73*	1.00	
Safe work behavior	0.04	-0.22*	-0.51*	0.49*	-0.26*	0.48*	-0.29*	-0.26*	1.00

* indicates 0.01 probability level of significance

Then this is the correlation among the constructs.

(Refer Slide Time: 52:41)



And then this is your structural part and these are the structural regression coefficients.

(Refer Slide Time: 52:47)

Structural Model Fit Indices

Goodness-of-Fit Indices for the Structural Model of Accident Path Model

Parameter	Values
Chi-square with 15 degree of freedom	212.23
Root Mean Square Residual (RMR)	0.06
Goodness of Fit Index (GFI)	0.87
Normed Fit Index (NFI)	0.88
Comparative Fit Index (CFI)	0.88
Incremental Fit Index (CFI)	0.88

Then you see whether the structural model is fit or not. If your structural model is fit then go for interpreting the things. So here ultimately there are several structural that fit measures like chi-square with 15degrees of freedom this should be as small as possible, root mean square is 0.06 this should be less than 0.05, less than or equal to 0.05 the goodness of fit index is 0.87,0.9 or more is better means adequate non fit index but this is almost 0.9. We can say the structural model is fitting the data.

(Refer Slide Time: 53:52)

Measurement Model Fit Indices

Goodness-of-Fit Indices for the Measurement Model of Causal Accident Model

Parameter	Values
Chi-square with 99 degree of freedom	257.24
Root Mean Square Residual (RMR)	0.06
Goodness of Fit Index (GFI)	0.98
Normed Fit Index (NFI)	0.97
Comparative Fit Index (CFI)	0.99
Incremental Fit Index (CFI)	0.99

Similarly we have also seen that the measurement model also fits the data. So that means our model or my model fits very much to the data then using this structural equation modeling, you will be also able to find out the total effect that is direct effect plus indirect effect and also you will be able to based on the effect you will be able to rank the variables.

(Refer Slide Time: 53:58)

Contributions of constructs

Total Effect of the Significant Variables on Work Injury

Variables	Direct	Indirect	Total	Rank Order
Work hazards	----	0.15*	0.15*	3
Social support	----	-0.14*	-0.14*	4
Safety environment	----	-0.16*	-0.16*	2
Job dissatisfaction	0.29*	----	0.29*	1
Safe work behavior	-0.14*	----	-0.14*	4

* indicates 0.05 probability level of significance

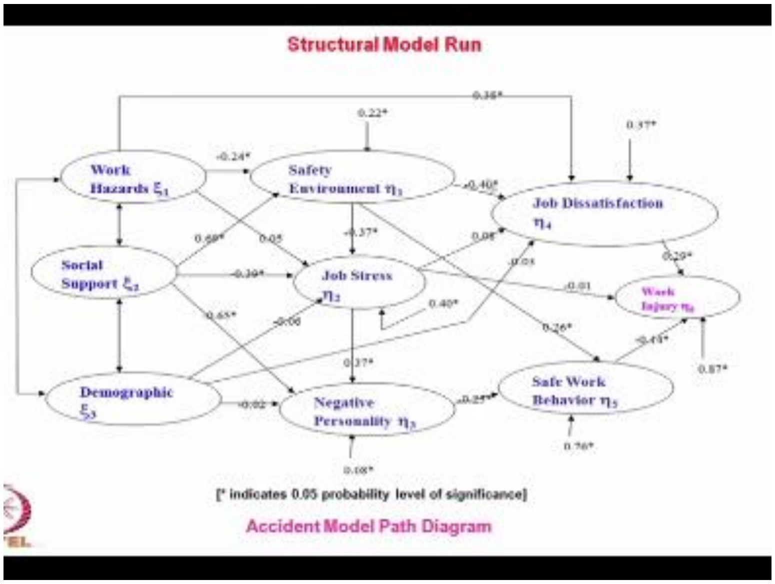
Total Effect of the Significant Variables on Safe Work Behavior

Variables	Direct	Indirect	Total	Rank Order
Work hazards	----	-0.07*	-0.07*	5
Social support	----	0.40*	0.40*	1
Safety environment	0.26*	0.03	0.29*	2
Job stress	----	-0.09*	-0.09*	4
Negative personality	-0.25*	----	-0.25*	3

* indicates 0.05 probability level of significance

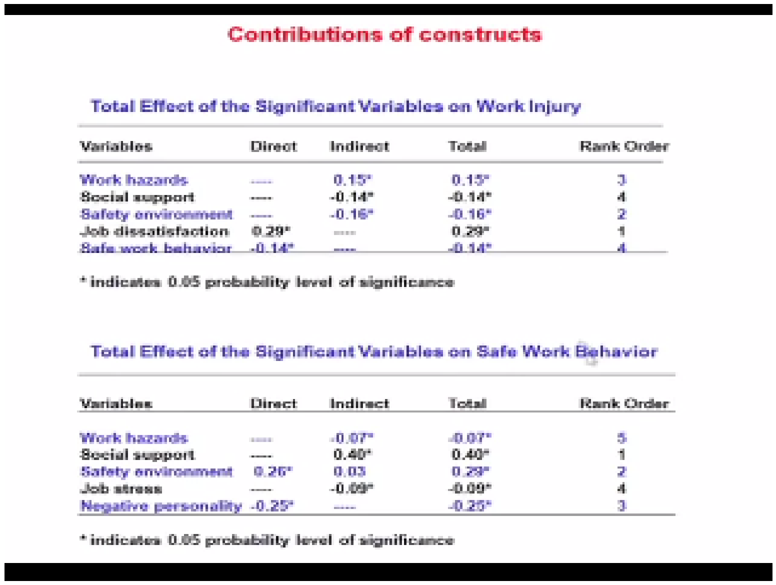
Now what do we see that we found that work hazard visa work injury work hazard has no direct effects in the model we have failed but it is indirect effect. Total effect is this social support this indirect affect this safety environment indirect effect this and total effect this job dissatisfaction direct effect and safety behavior this is the case. So ultimately based on this total of effect we have given the rank one, two, three, four and five, like this. Similarly as there are many variables, many dependent variables or dependent constructs you can find out here.

(Refer Slide Time: 54:42)



See these are the exogenous but all are endogenous. So for all any of the endogenous level variable, you can find out that what are the direct and indirect effects. This will definitely help you in making better policy, okay.

(Refer Slide Time: 55:02)



Contributions of constructs

Total Effect of the Significant Variables on Work Injury

Variables	Direct	Indirect	Total	Rank Order
Work hazards	----	0.15*	0.15*	3
Social support	----	-0.14*	-0.14*	4
Safety environment	----	-0.16*	-0.16*	2
Job dissatisfaction	0.29*	----	0.29*	1
Safe work behavior	-0.14*	----	-0.14*	4

* indicates 0.05 probability level of significance

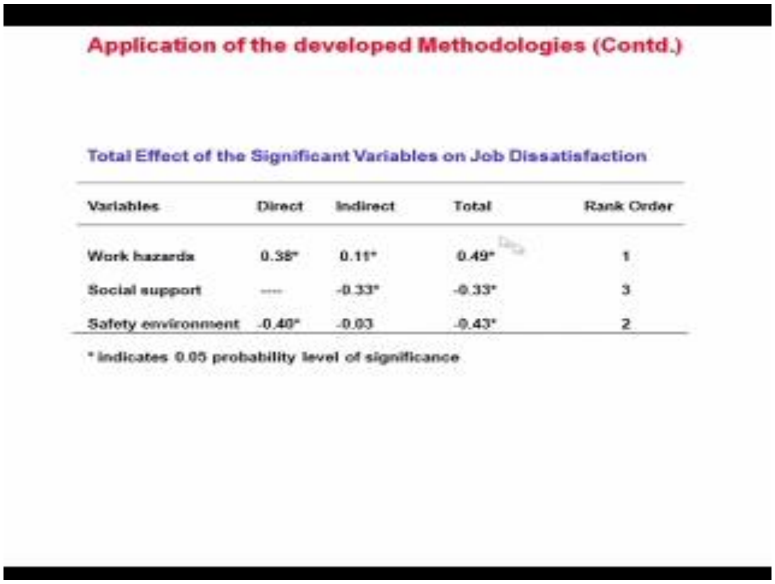
Total Effect of the Significant Variables on Safe Work Behavior

Variables	Direct	Indirect	Total	Rank Order
Work hazards	----	-0.07*	-0.07*	5
Social support	----	0.40*	0.40*	1
Safety environment	0.26*	0.03	0.29*	2
Job stress	----	-0.09*	-0.09*	4
Negative personality	-0.25*	----	-0.25*	3

* indicates 0.05 probability level of significance

Total effect, so we calculated for this case work behavior for work injury for job dissatisfaction, what are the variables.

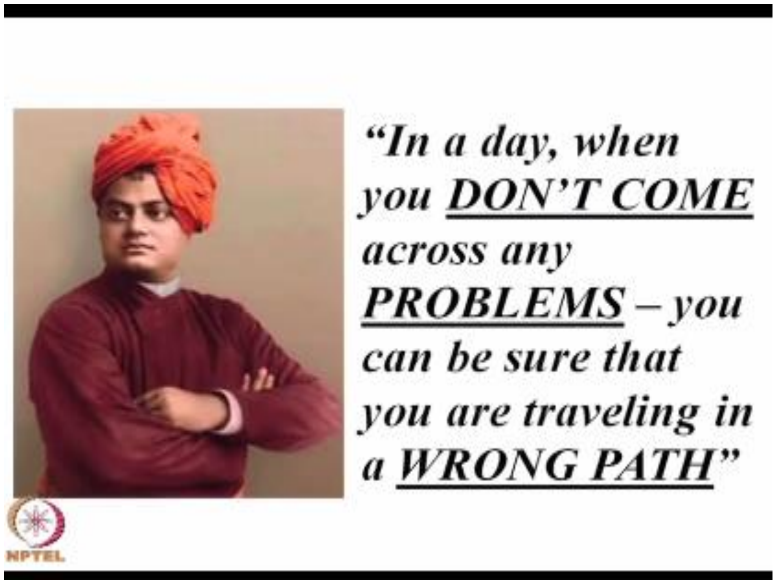
(Refer Slide Time: 55:07)



Variables	Direct	Indirect	Total	Rank Order
Work hazards	0.38*	0.11*	0.49*	1
Social support	----	-0.33*	-0.33*	3
Safety environment	-0.40*	-0.03	-0.43*	2

And what is the rank order.

(Refer Slide Time: 55:12)



So thank you very much. Next class we will discuss confirmatory factor model with respect to structural equation modeling followed by my structural model which is basically a path model. Thank you very much.