Applied Multivariate Statistical Modelling

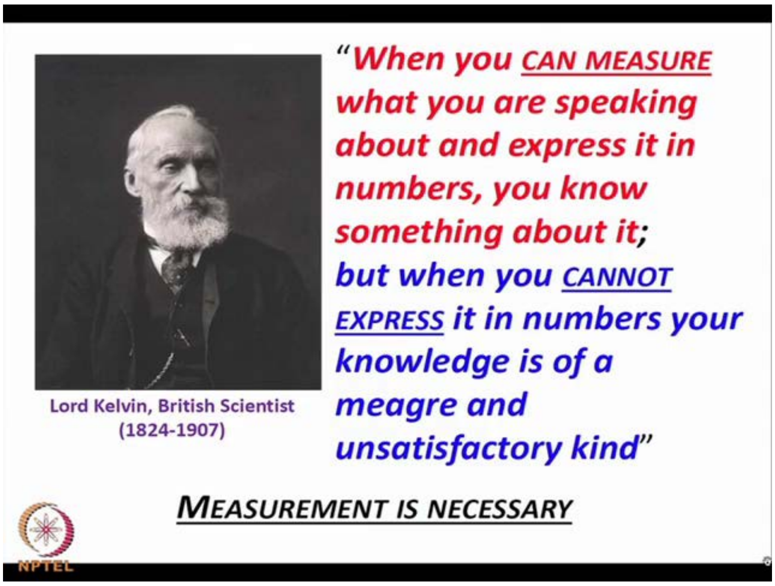
Prof. J. Maiti
Department of Industrial Engineering and Management
Indian Institute of Technology, Kharagpur

Lecture - 33
Factor Analysis

Good morning, today our discussion will be on factor analysis.

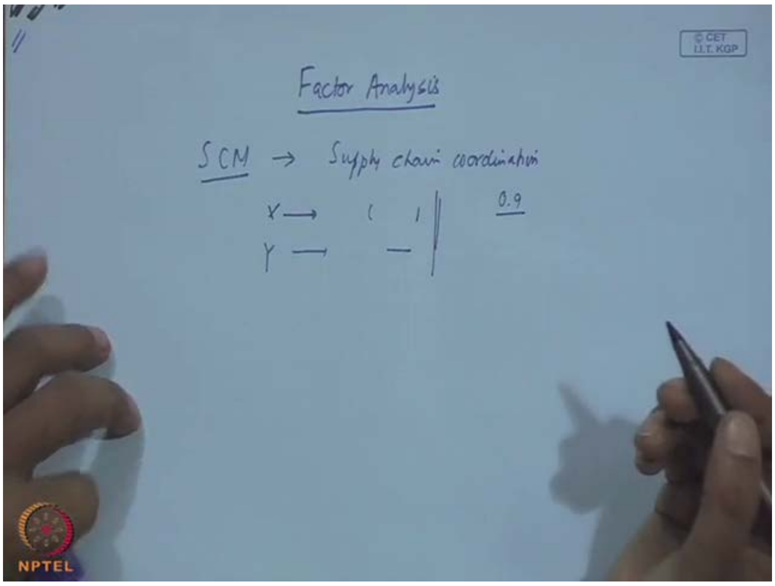
(Refer Slide Time: 00:23)

(Refer Slide Time: 00:36)



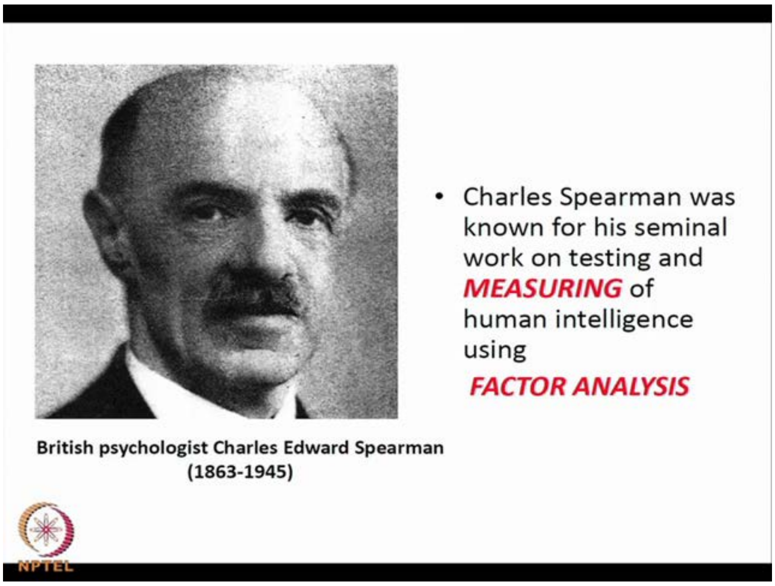
So, I will start with the famous quote of Lord Kelvin, the British scientist, what he said? When you can measure what you are speaking about and express it in numbers, you know something about it, but when you cannot express it in numbers your knowledge is of a meagre and unsatisfactory kind. Then measurement is necessary that means, what he meant to say is suppose, whenever you talk about any system of its behavior then the behavior pattern you must be able to measure. You must able to give a number to it for example.

(Refer Slide Time: 01:49)



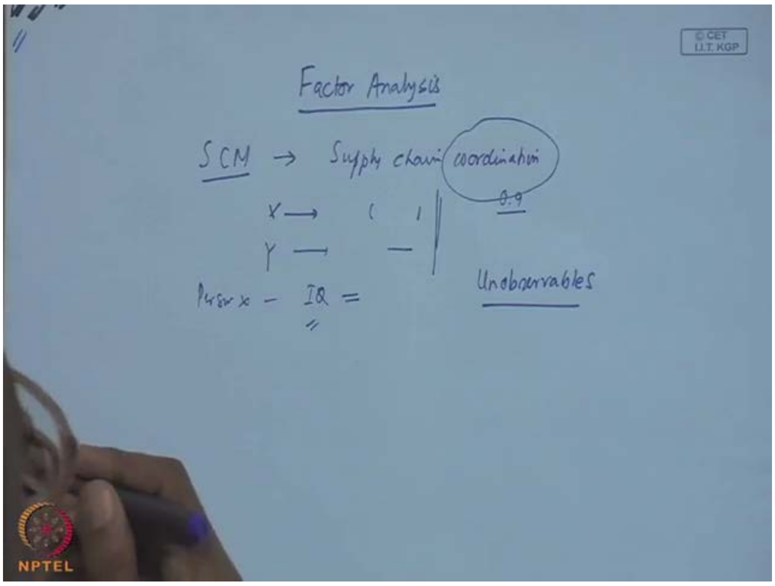
In our department there is one important area to work on, that is supply chain management and we talk much about supply chain coordination. So, we may say that X supply chain is coordination, good. Y supply chain coordination, bad. Like these when you talk up in this manner that this knowledge is not perfect. If I say that, yes suppose there is the particular food, let it be paddy supply chain in Indian context. Then you are able to tell that at the local level block level or at the state level, that the different agent how they are coordinated and the measure is 0.9. Suppose, something like this then this is a valuable knowledge that is what is important and that is said by Lord Kelvin.

(Refer Slide Time: 02:57)



Second gentleman, I think he is one of the famous statistician Charles Edward Spearman. So, Charles Spearman he measured the human intelligence that means what I mean to say.

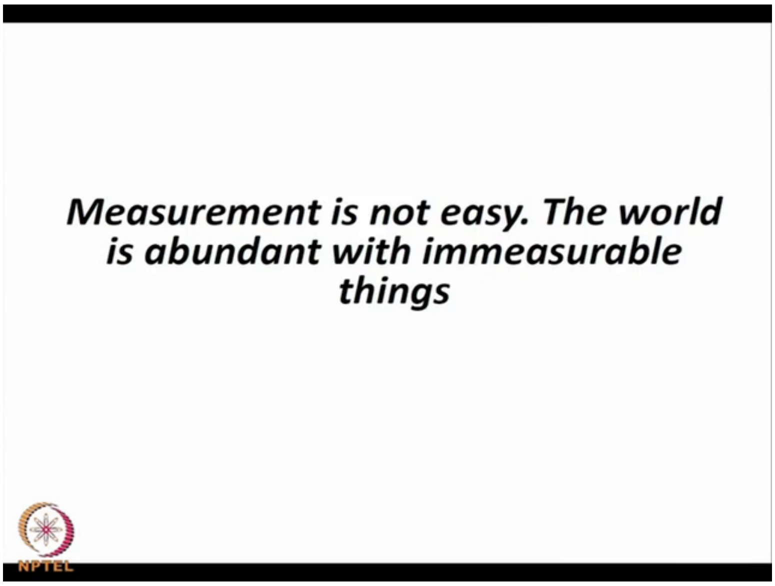
(Refer Slide Time: 03:24)



Suppose, what is in X intelligence level or IQ is somewhere low, person Y somewhere low like this. So, human intelligence you can understand it is such a concept. It is not that 5 kg of rice or 25 kg of some other product. It is not like this, you cannot just take some this type of measurement and go. It is a soft thing for example, I supplies in coordination you will say soft thing your mental measurement of that intelligence, human intelligence.

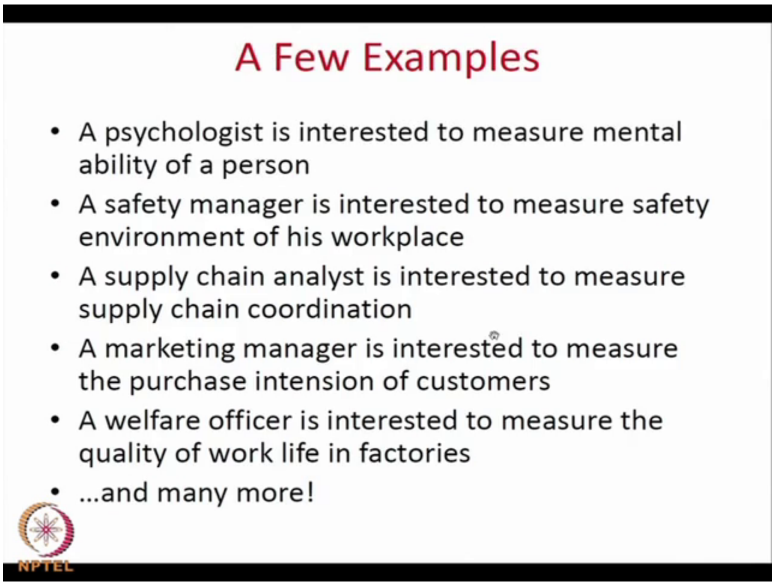
It is not visible, these all are basically unobservable things, unobservable. I can say the coordination is unobservable. You cannot directly go and measure intelligence, also unobservable. Then question comes how do you measure such unobservable? So, there are certain techniques and the factor analysis is one of such techniques and probably the one of the best techniques, that is used in measuring unobservable.

(Refer Slide Time: 04:55)



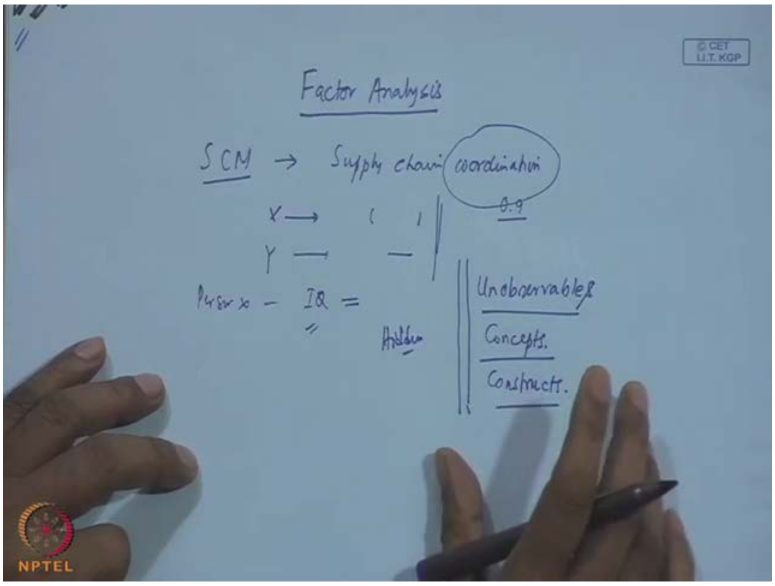
The key concept here is measurement is required, but measurement is not easy. The world is abundant with immeasurable things which cannot be measured. I will give you some example.

(Refer Slide Time: 05:13)



Here a psychologist is interested to measure the mental ability of a person. A safety manager is interested to measure safety environment of his workplace. A supply chain analyst is interested to measure supply chain coordination. A marketing manager is interested to measure the purchase intention of customers. A welfare officer is interested to measure the quality of work life in factories, and many more. You can find out hundred such items examples, where you see that it is not there is one instrument hardware based instrument. You will go and measure, it is not possible. We have a 90 percent of the cases take decisions based on this type of concepts.

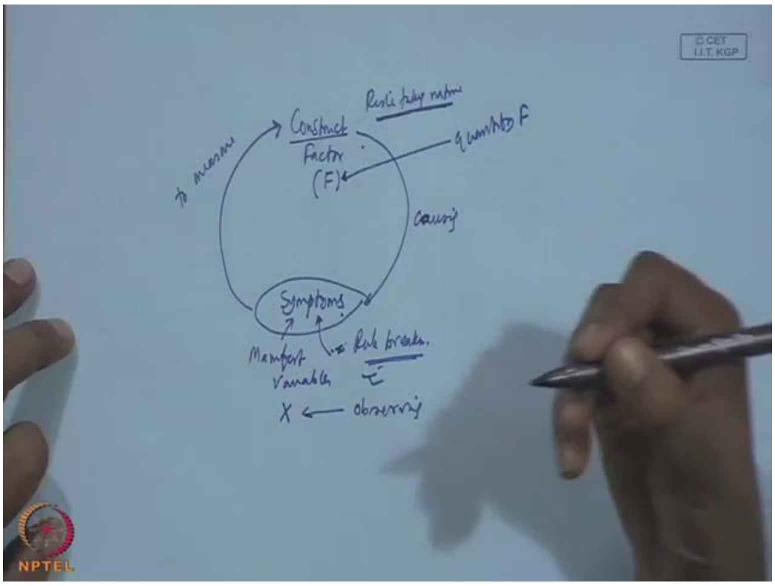
(Refer Slide Time: 06:07)



So, another word is I want to say that concepts many time, we say that constructs. So, by saying these we are basic unobservable concept construct, by saying this we are basically telling to you that there are certain things which are hidden unobservable that is not directly measureable. But those things are very important and because those things causes something to happen. For example, if there is not proper coordination in the supply chain, the supply chain performance will be poor.

Suppose, the quality of what life is not good then the human productivity safety and all other decision accompany that will be poor. So, that means although there are many things like these which are unobservable, which are concepts which constructs hidden. This cannot be directly measured, but their effects means because of the presence of these this constructs or concept of unobservable things manifested into different symptoms.

(Refer Slide Time: 07:45)

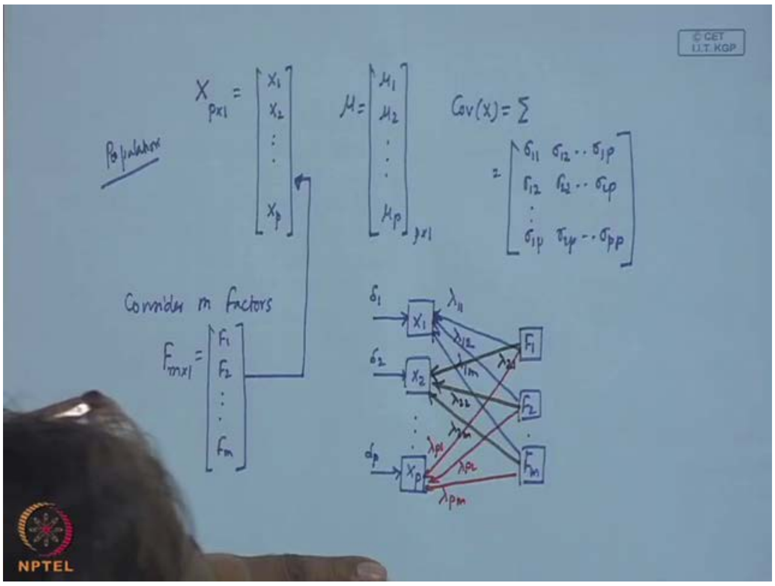


What I mean to say, I mean to say that suppose there is construct concept or we can say factor. So, these things manifested in terms of symptoms. So, because for example, someone someone is risk taking in nature risk taking in nature so what will be symptoms he will may be a rule breaker not follow the traffic rule. So, you are able to see this symptom that not following the rules may be rebellious, but it is so that risk taking that attitude.

What is his personality, that is the main reason, the cause of this rule breaking. That is why we say that that sometime hard hidden on a immeasurable constructs or factors of there reach out causing the symptoms, and these symptoms can be used to measure this factor or construct that mean factors or cofactoral construct causing the symptoms can be used to measure construct, getting me?

So, essentially that concept is like this, if I say this factor is f and there are several symptoms which we can say the manifestation or manifest variable. The symptoms are nothing, but manifestation of these hidden thing. So, this manifest variable if we say X, correct? So, that mean what is available with you are observing this X and you want to quantify F, getting me? You observing this that means with this if we can say something like this, that you are observing X which is P cross 1.

(Refer Slide Time: 10:42)



That means P number of this is a variable vector having P manifest variables, P manifest variables if differently coming from a population. Now, suppose this X is having the mean vector like this, which will also be mu 1, mu 2, mu P. This will be P cross 1 vector, this X is also having a covariance matrix which is sigma capital sigma this one is you can write sigma 1 1, sigma 1 2, sigma 1 P, sigma 1 2, sigma 2 2, sigma 2 P, like this sigma 1 P, sigma 2 P, sigma P P. So, this is what is known about the population, that mean there are P variables, P mean vector, P plus cov X vectors. The mean that mu 1 to mu P or sigma 1 1 to sigma P, these values may not be known usually, not known from the population point of view.

Now, what we have described here? We said that in this place we described that this X you are observing, because there is certain underlined causes or hidden constructs or factors. So, if this is the case, if we consider now consider m factors. So, then my F is the vector of factors whose is m cross 1. This will be your F 1, F 2, like F m and you are seeing that whatever you observe in X, this is because of the causal factor F.

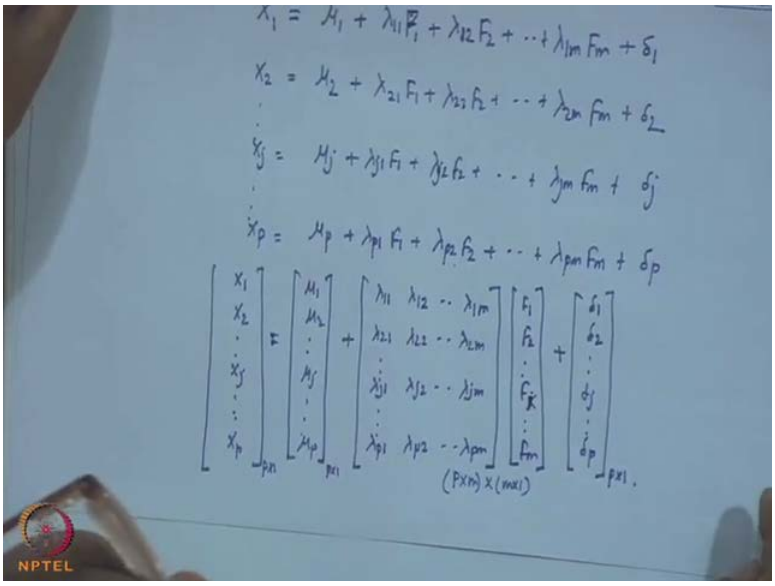
F is causing this, then we can write linear regression like equation sounds to, but first I will pictorially represent this. So, what you can write? I have X 1, this is my one object variables. My second object variable is X 2, like this my last object variable is X P and we are saying that there is factor F 1, F 2 and F m factors each are causing this X variables to occur to happen.

So, in this case I can say X 1 is caused by F 1, X 1 also caused by F 2, like this F m. What is that I am trying to say that there are m factors who is influencing this X 1, because of these factors X 1 a variable to object. So, then if I write like this F 1, X 1 is caused by F 1. I can write lambda 1 1 for X 1 1 F 1 1 then lambda 1 is caused by 2 lambda 1 is caused by m. So, what you can do? The same thing will happen to X 2.

So, what we will write? Then this one will be lambda 2 1, then lambda 2 2, lambda 2 m, then another last one is your, correct? So, lambda P 1, lambda P 2, lambda P m, when we are assuming. What we are assuming? This F causing X. So, these are the all influencing coefficient. So, factor F 1 influenced on X 1 can be measured through lambda 1 1 like this.

Now, in addition to these what we assumed at that these factors collectively cannot explain everything about X 1. So, there will be one arrow term, correct? Now, using this, but please remember here we are not talking anything about the relationship between F 1, F 2 and F m. When we are not saying anything here? We are here, we are assuming that these factors are independent orthogonal. So, we will see that sum.

(Refer Slide Time: 17:40)

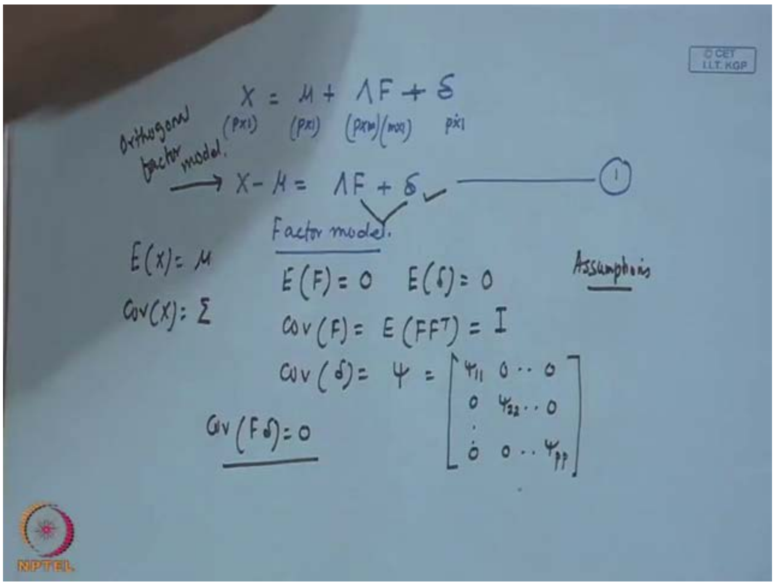


Later on with these if I want to write down certain equation, can I not write down? Regression like equation X 1 equal to mu 1 plus lambda 1 1 F 1 plus lambda 1 2 F 2 like lambda 1 m F m. See here, what you are writing? X 1 equal to lambda 1 1 F 1 lambda 1 2 F 2 lambda 1 m F m. What more is there? Delta 1 is there. So, plus delta 1. Similarly, you can write X 2 equal to mu 2 plus lambda 2 is affected by 1 F 1 lambda 2 2 F 2 like lambda 2 m F m plus delta 2.

So, in this manner you can write X j equal to mu j plus lambda j 1 F 1 lambda j 2 F 2 lambda j m F m plus delta j. And final one is X p equal to mu p plus lambda p 1 F 1 lambda p 2 F 2 lambda p m f m plus delta p, correct? So, that means in matrix form you can write X 1, X 2, X j, X p equal to mu 1, mu 2, mu j, mu p plus your lambda values up there, lambda 1 1, lambda 1 2, lambda 1 m, lambda 2 1, lambda 2 2, lambda 2 m.

So, like this lambda j 1, lambda j 2, lambda j m in the same manner lambda p 1, lambda p 2, lambda p m into your lambda f F 1 F 2, this will be your. Let it be F j, then your what will happen ultimately? F 1 to F m. So, I am not giving F j. Let it be even F k and F m what is happening here? Then you have p cross 1 equal to p cross 1 plus this is p cross m into m cross 1. And one more thing is here that is delta. So, you will write delta 1, delta 2, delta j and delta p, okay?

(Refer Slide Time: 21:50)



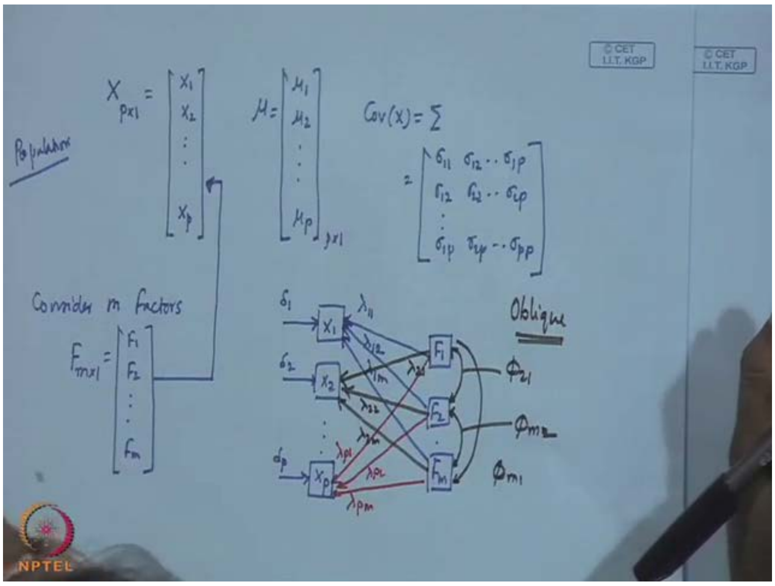
Now, I am writing this in matrix notation. X equal to mu plus lambda capital lambda F plus, I can write this is delta. You can write like this. So, what will happen actually? This one is your X, this one is your mu, this is your lambda, this one is F and this one is delta. What is the result? Your dimension p cross 1, p cross 1, p cross m, m cross 1, p cross 1. This is your desired vector.

So, you can further simplify this X minus mu lambda F plus delta. So, mean subtracted original variables is a function of the factors plus a rough. This equation if I say 1, this is the factor model. This equation is known as factor model, this equation will be further handled through several assumptions. I think all of you know that expected value of X is mu, covariance of X is sig capital sigma.

Now, what will be the E x? Our assumption is expected value of F. This one is 0, the reason is see the mean subtracted. So, expected value of F is 0 assumptions, expected value of delta that is that is 0. Now, we want to make this type of things that covariance of F is expected value of f F transfuge, because expected value of F is 0. This one this will be i and your covariance of another one, the delta that will be psi which is a diagonal matrix. psi will be p cross p and covariance of F delta, that is 0 that mean F and delta. This covariance is 0.

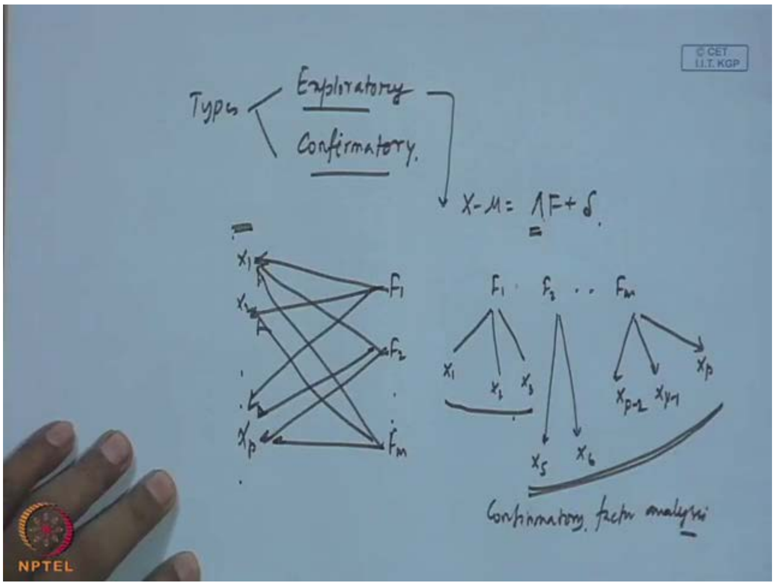
Now, this is that part of the assumptions under. If this assumptions hold, then this factor model is known as exploratory factor model orthogonal. Definitely, exploratory plus orthogonal n. So, here I think I should say little bit of as we are started with little bit of the types, we are saying orthogonal because they are independent F 1, F 2, f m.

(Refer Slide Time: 26:21)



They are independent, there will be there will be oblique factor model, you will say. No, they are not independent, they are each correlation between them, getting me? So, these correlation coefficient will be phi 2 1, this will be your phi m 2, this will be your phi m 1. So, when you allow covariance component between the factors, then your factor model will be become oblique factor model. Now, this oblique factor model will be treated separately and we will discuss orthogonal factor model. This is one, second one is important concept. Here is that we are saying the types of factor model types.

(Refer Slide Time: 27:11)



One is exploratory factor model. Another one is confirmatory factor model. Exploratory and confirmatory factor model. What is exploratory and what is confirmatory? What is the difference between these two? See, in the first case we are saying that X minus mu equal to lambda F plus delta. Here, we have taken all X and all F connected together, because of this delta matrix lambda.

This capital lambda matrix, with this capital lambda matrix we have connected both that means. What I mean to say, what I am trying to say that it is something like this. I have F 1, F 2 to f m and I have X 1, X 2 to X p. And what we are trying to say that each one of X, each one the x is dependent on F, all F. I do not know exactly which are the X values or which are the X variables that are basically coming out of F, which one which of the F s I do not know.

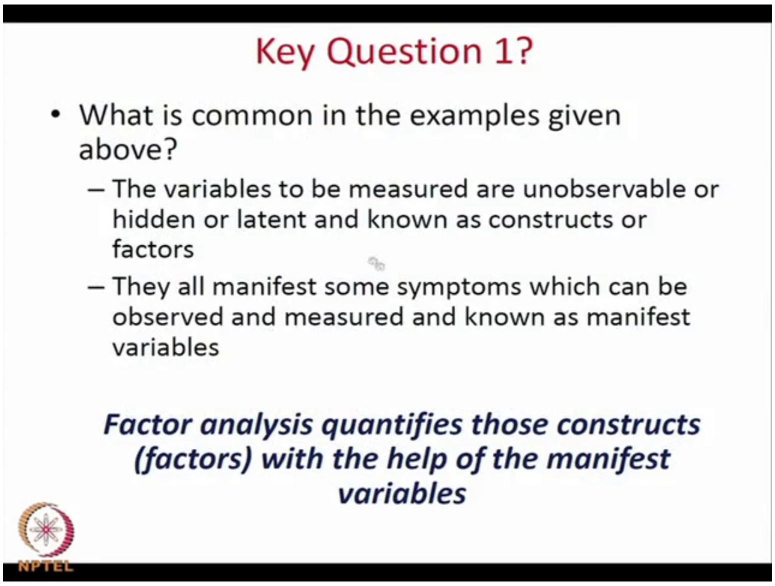
That is why what I am doing, I am allowing all the full sets of relationships here first hand and what you are saying? You are saying that this is my observed variable structure and you find out as many factors as possible. And then you say that what is this number of factors and what are those? The factors that are linked to what type of variables. Under this condition you are exploring about the possibility of hidden factors, that is why it is exploratory factor analysis.

So, here also it will be linked. This also will be linked to so everything is linked to I think this already done, getting me? Now, confirmatory factor analysis is different. What is then they are you say that F 1, F 2. Suppose, m factors are there, but you clearly know what are the X 1, X 2 maybe X 3. What are the variables x 5, x 6. So, let it be X p minus 2 X p minus 1 and X p. You know clearly, you know that three variables are because of factor 1, another 2 manifest variable because of factor 2. This variable because of factor m going in advance, because of your research work, your literature review and survey, all those things you know which of the manifest variable representing what type of which factor.

So, essentially what is happening here? So, F 1 to all other these variables the connectivity is not required, the relationship this is not required when it is coming from here. Only now this type of factor analysis is second, this factor analysis is confirmatory. Why it is confirmatory? The reason is you are saying that there are suppose, if you manifest variable, let it be 3 or 5 related to F 1 3 are some other, some manifest variables related to F 2. And you want to confirm whether you are correct or wrong. This is your hypothesis. Let it be then the first factor maybe something. Let it be the trapezin coordination.

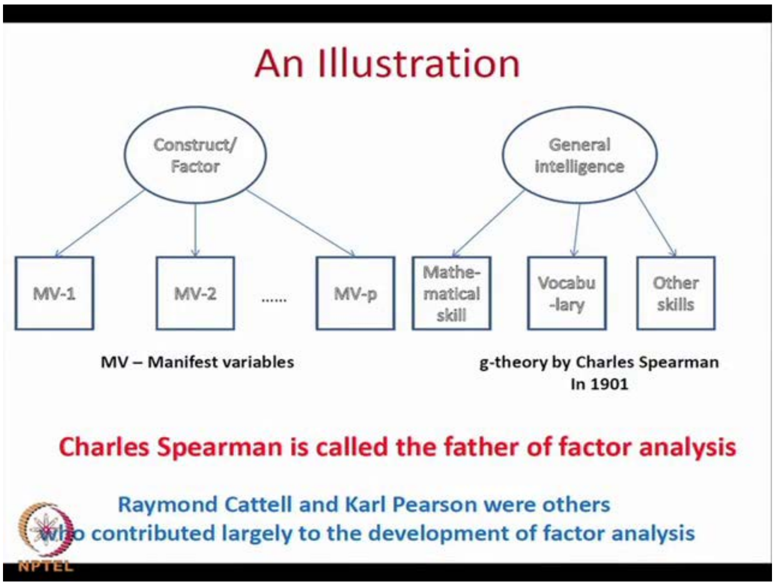
So, there are certain indicators. You are finding out and you are saying these indicators are measuring supply chain coordination, these indicators are measuring something else. Something else like this, then through statistical route you are confirming whether, what you assumed or what you hypothesized what you have proposed, they are correct or false. That is your confirmatory factor analysis. So, we will stick to exploratory factor analysis. Now, then I want to understand the purpose of exploratory factor analysis, correct? What is first question in factor analysis with respect to the examples given to you.

(Refer Slide Time: 32:35)

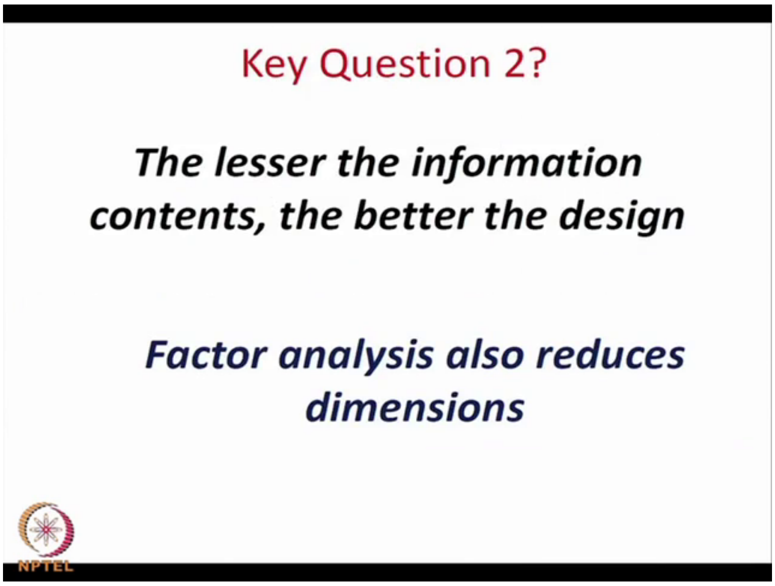


So, I think around 7, 8 examples we have discussed. My question is, what is common in the examples given above? The variables to be measured are unobservable and you have seen that hidden or latent and known as constructs or factors. We have discussed, they all manifest some symptoms like in terms of X. So, this is F, this one is X factor analysis quantifies those constructs or factors with the help of manifest variables. There is the first thing what we are doing in exploratory as well as in confirmatory.

(Refer Slide Time: 33:27)



(Refer Slide Time: 33:29)

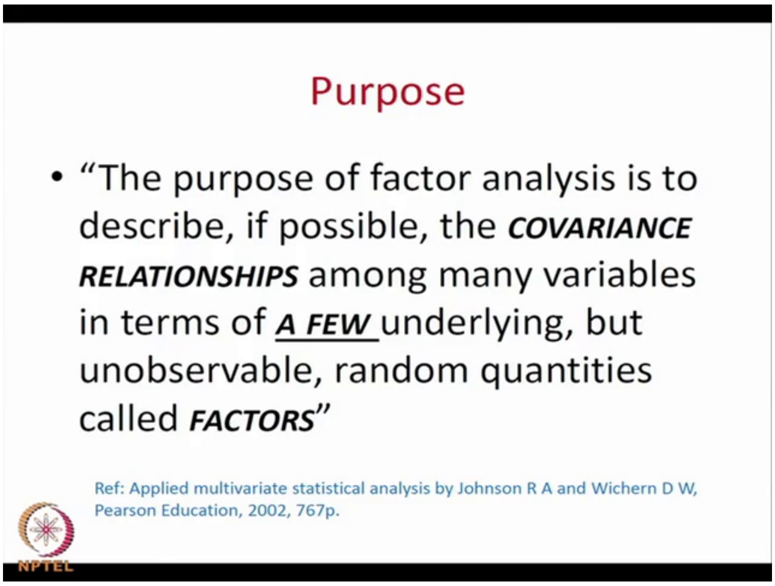


Also, you will be doing the same thing. So, this is a illustration. Now, key question 2. I have a p number of variables, that mean p information, p dimensional information based. I do not want this much of information. What I want? I want the reduced information, see if I require to know two dimensions only then I take decision. I act, I control system.

All things will be much easier if there are five different dimensions, compared to that what will happen. Suppose, you think of a point there what we used to have in this point. Suppose, there are three bottoms, one bottom to just point the lecture, one of the sentence. What another one is just from you, just call from one slide to another slide and the third one is the switch off, this type of things are there.

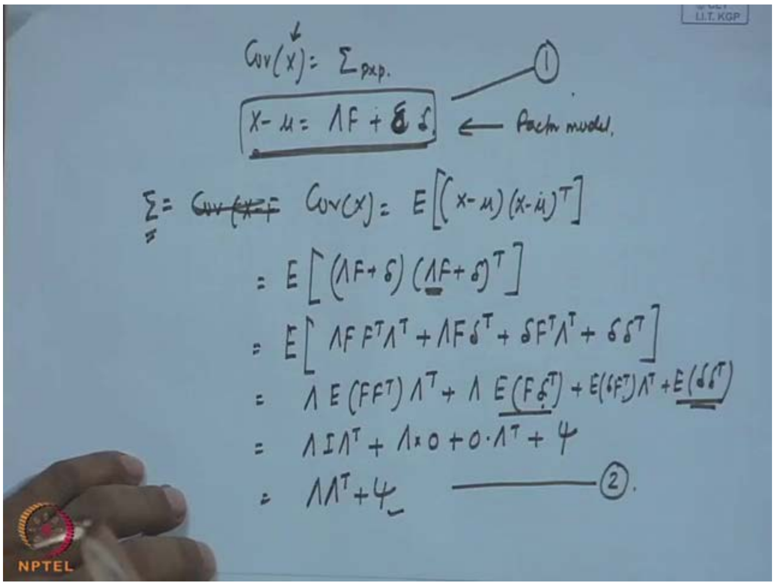
So, I must know these 3. These 3 important types information is known to me and I redo. Now, suppose they are U instead of these three, that number of information available and second thing is thing is that if you required to go for slide change, you also required to do something else before, that is the dependent things of their own selves. Then what will happen? You will be confused. The same thing will happen if us. So, that is why what we are saying the less of the information content, the better is the design, that means we want one inferent system or one recessive support system where our information will be at the reduced dimensions.

(Refer Slide Time: 35:30)



So, factor analysis also do this far from this work and what is information in factor in your data set that is the covariance matrix factor analysis uses the covariance matrix. So, the purpose is the purpose of factor analysis is to describe if possible the covariance relationships among many variables in terms of a few underlying. And but observable random quantities called factors that is the purpose. So, we will we are now in a position that we know what is construct or factor. What is what are manifest variables and what is the exploratory factor model and what are the different assumptions related to this factor model.

(Refer Slide Time: 36:27)



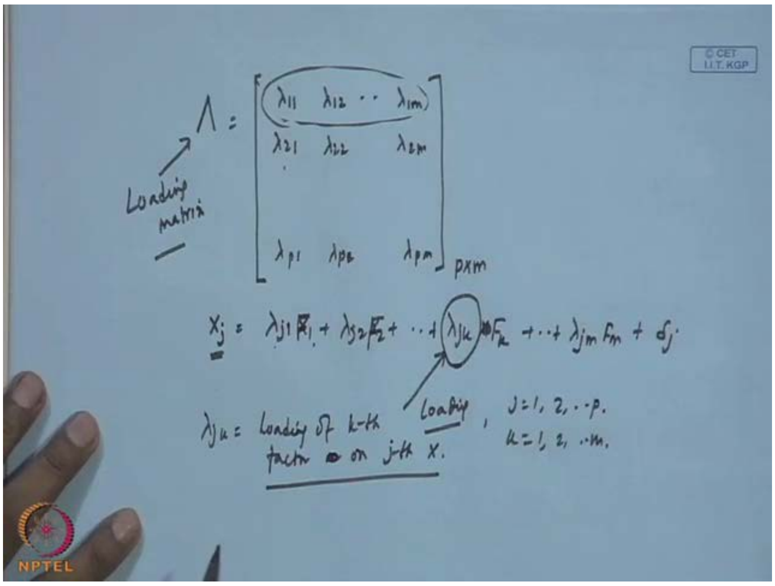
Now, the of the matter is variability explanation. We are saying that information is covariance of X which is sigma. So, you are creating a factor model which is X minus mu equal to lambda F plus delta. We are using lambda F plus delta, getting me? This is your factor model, I want to know that what is the variability. You are able to explain using this model. What whose variability of X what is that that is the covariance of X p cross p. So, if I want to understand covariance, can I not write like covariance of, sorry covariance of X equal to expected value of X minus mu X minus mu transfuge. We have seen earlier also, that mean this quantity we can write like this X minus mu.

Suppose, if I say this is my equation 1, from equation 1 I can write lambda F plus delta into lambda F plus delta transfuge X minus mu. We can do lambda X plus delta lambda m minus delta transfuge. So, I am writing like this lambda F into this one. So, this one will be F transfuge lambda transfuge plus lambda F delta transfuge plus delta F transfuge lambda transfuge plus delta deltatransfuge. Now, that capital lambda is the regression like constants, you have seen earlier. So, this constant will come.

So, lambda expected value of F transfuge, capital lambda transfuge plus lambda expected value of F delta transfuge plus expected value of delta F transfuge lambda transfuge plus expected value of delta delta transfuge. Now, you have to see the covariate this assumptions. So, I am putting the assumptions parallel here. So, what we say assumption is expected value of F ftransfuge is i. So, I will put like this lambda I, lambda transfuge plus lambda expected value of F, and delta what we have written related to F and delta this one. This is 0. So, this into 0. Similarly, this 0 into delta lambda, capital lambda transfuge plus.

What is the covariance of delta, that is expected value of delta delta transfuge this one. So, this is psi. So, then what is happening to the resultant matrix. Now, capital lambda i lambda is lambda lambda transfuge plus this, what is this? Which is your sigma value? Covariance of X is capital sigma, this is equation number 2 and this is the variability explanation point of view this equation. So, that means if you know this lambda, capital lambda and then psi, you know this, getting me?

(Refer Slide Time: 41:04)

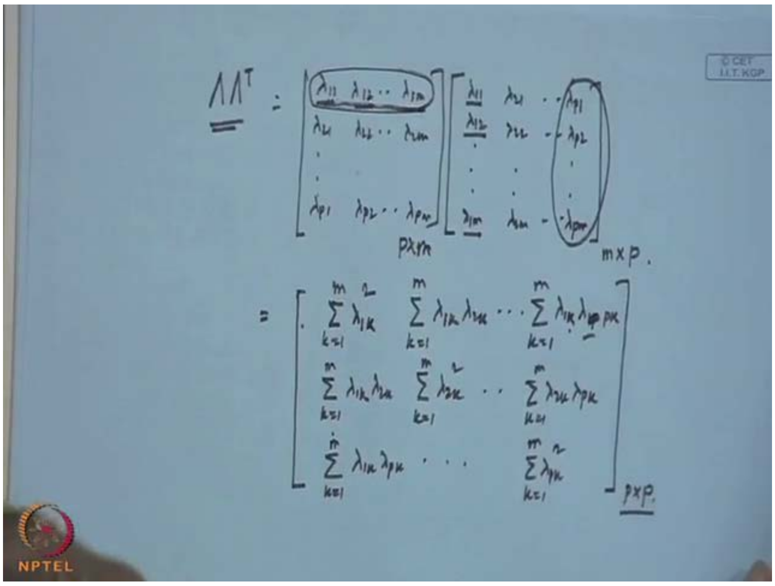


What is this capital lambda here? This one is nothing, but can you remember what we have given lambda 1 1 F 1, lambda 1 2 F 2 , like this lambda 1 m F m, lambda 2 1 F 1, lambda 2 2 F 2, lambda 2 m F m, like this. So, this lambda part I am writing now. So, lambda p 1, lambda p 2, lambda p m, this one is p cross m matrix. What is this lambda, visible?

Suppose, this one their loadings their so if you take a general formula like X j, what you have written lambda j 1 1 1 2 1 m, yes j 1 X 1 lambda j 2 X 2. Like this if I write lambda j k, your j k, sorry lambda X j u write lambda j 1 f 1 lambda j 2 F 2, lambda j k F k. So, similarly lambda your j m F m plus delta j, this is the original equation. Here, lambda j 1 F 1 like this.

So, see what is happening here. This general one is lambda j k. This is the general loading. So, it will be for j equal to 1 2 p k equal to 1 to m. Now, this one is known as loading matrix. Loading matrix mean this is really factor loading matrix and each lambda j k there are the loading of kth factor. On jth variable X variable this one is lambda j k is known as loading of kth factor, on kth factor, on jth X, correct? Now, when what is your lambda transfuge.

(Refer Slide Time: 44:00)



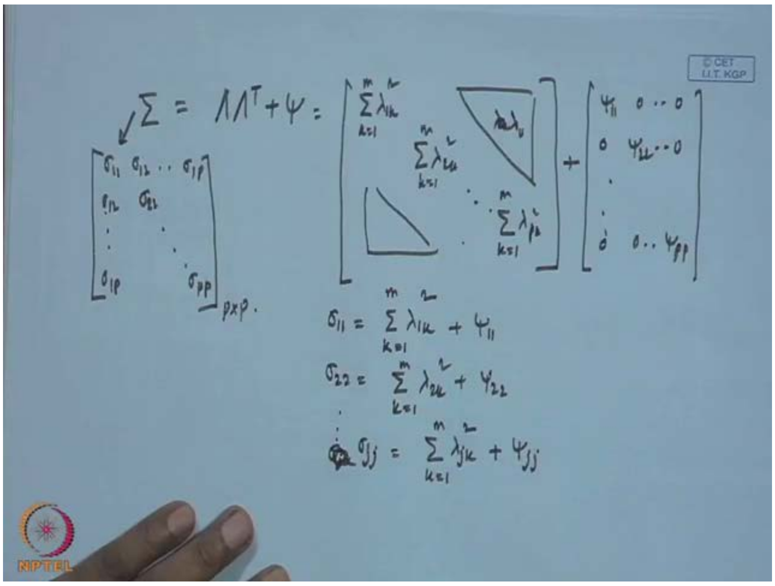
If you write down this, you have to write down lambda 1 1, lambda 1 2, lambda 1 m , 2 1, 2 2, 2 m p, 1 p 2 p m, then transfuge 1 1, 1 2, 1 m, 2 1, 2 2, 2 2, 2 m, p 1, p 2, p m. This is your p cross 1, this is now m cross p and you must get p cross p because this p cross m, p cross m, m cross p. So, if you just multiply the matrix lambda 1 1 into lambda 1 1, lambda 1 2 into lambda 1 2, lambda 1 m into lambda 1 m, what is changing here?

This one remain constant, this one that let k. I am giving k is basically 1 to m, the factor is changing k equal to 1 to m, and the square lambda 1 1 square plus lambda 1 2 square plus lambda 1 m square. Then second one will be again k equal to 1 to m, you see lambda 1 1 to 1 m and 2 1 to 2 m. So, that mean lambda 1 k, lambda 2 k, the covariance is coming here.

So, same manner k equal to 1 to m lambda 1 k and lambda. No, k is 1 to m. Now, it is 1, then 1 2 then 1 p, this is p cross p matrix, lambda 1 k, lambda p k, lambda this. This row be this column, this is coming here. So, same manner now this will be a symmetric matrix, k equal to 1 to m, lambda 1 k, lambda 2 k. This will be lambda 2 k square, k equal to 1 to m.

Same manner, lambda 2 k, lambda p k, k equal to 1 to m. In the same way it will come k equal to 1 to m, lambda 1 k, lambda p k, same manner you go it will be lambda p k square, k equal to 1 to m. This will be p cross p. So, this diagonal element is the variance, component of diagonal element is the covariance component. And this diagonal element relates each of the variables also, then if I write down further I can straightway.

(Refer Slide Time: 47:58)

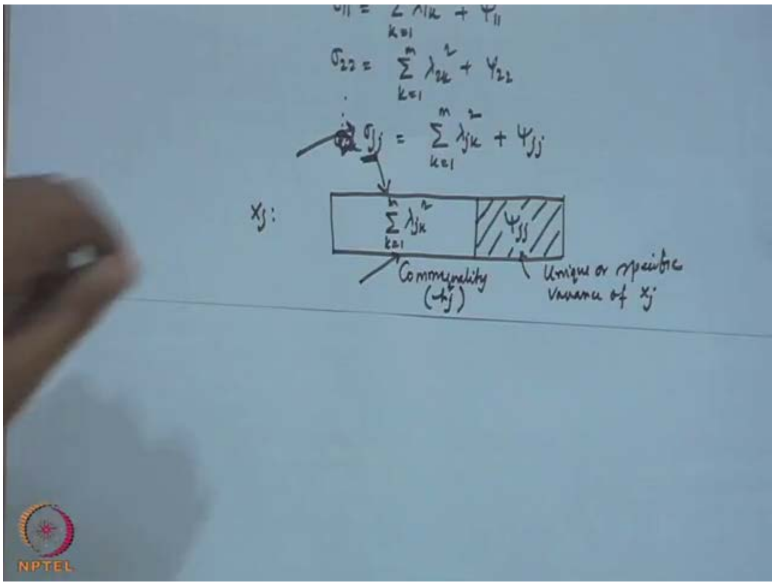


Write that lambda, lambda transfuge plus psi. If I write this is nothing, but I am not writing the off diagonal element. Now, only diagonal element I am writing 1 k square, k equal to 1 to m, lambda 2 k square, k equal to 1 to m, lambda k equal to 1 to m p k square. Then off diagonal elements, already I have written with lambda 1 k, that lambda k and something like this. Already I have written this plus our psi is psi 1 1 0 0 0 psi 2 2 0.

So, like this 0 0 psi p p and this is what is your sigma. So, if I write this sigma in the same manner, the way we have written in matrix. So, that this one will be sigma 1 1, sigma 1 2, sigma 1 p, then sigma 1 2, sigma 1 p. Now, sigma 2 2, sigma p p, the sigma is like this. So, this is p cross p, then comp if you compare this versus this diagonal element what you will get? Sigma 1 1 equal to sum total of lambda 1 k square, k equal to 1 to m, correct? Plus psi 1 1, yes or no?

So, similarly, you will be getting lambda sigma 2 2 equal to k equal to 1 to m lambda two k square plus psi 2 2. So, in the same manner if I want to understand here suppose, j j let it we will write. Now, j j here we are interested in the variance components j j. So, what will happen here, then k equal to 1 to m lambda j k square plus psi j j, then actually what is this?

(Refer Slide Time: 50:40)



See, now try to understand the variability part. Our variable is X j in this equation. Corresponding variable is X j, it has sigma j j that is the total variability. So, I can say this total is sigma j j. Now, this is divided into two parts. If I say this is the case where this person k equal to 1 to m, lambda j square and this person is phi j j. So, what I am trying to say then that sigma j j is partitioned into two halves, two parts, correct? This one is what lambda j j, you see you are extracting in factors and lambda j 1 square plus lambda j 2 square plus lambda j 3 square, like this up to lambda, where j m square.

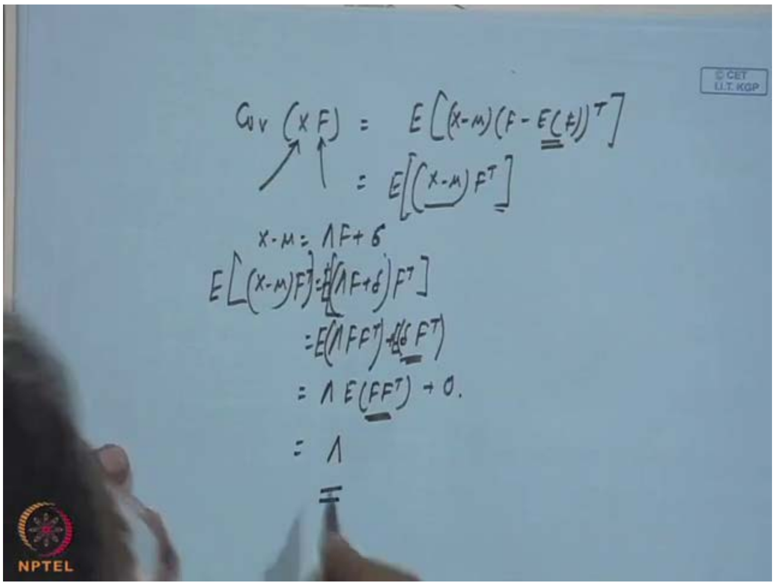
So, this is the contribution by the m common factors, this is the contribution by the m common factors and this psi j j is something which is not contributed by the common factors. That is why this one is known as communality and we denote it in terms of each j square communality and this one is unique to x j. The common factors are not able to explain this much of variability, this is unexplained, this one is unexplained one, this is explained by the common factor. This is unexplained.

So, it is we can say unique. We say unique or specific variance of X j. So, it can be like this, so unique and specific. Again some you can argue that these person which is related to X j precisely related to X j plus something which are, because you have taken factors and some other factors maybe variable correlation, some other variable you have not considered, your data collection many things. So, error also maybe included here, but in factor it is analysis, we say usually we denote like this. The variability of X j is because of the communality and because of the uniqueness of that variable, correct? So, what you want then you definitely want to maximize the communality. The password is the variability with other common factors contributing towards explaining the object variable's variability.

So, that means what is the variance part from what is the contribution for the first factor lambda 1 1 square. What is the contribution towards variability to the second factor, that is what we are saying that second factor lambda 1 that it is variable X j, then j 1 square. It is then j 2 square, j 3 square lambda that all lambda square, these are contribution from each of the factors in explaining to the variability of the X j, that factor and something you will not be able to explain.

You can explain if you suppose you extract p factors out of p x variables. Everything will come under this communality, but that is not the purpose of factor analysis. Factor analysis, one hint wants to produce the dimensions. Other end it will do in such a manner that the factors can be linked with the original variables and naming of factors can be possible, getting me?

(Refer Slide Time: 55:45)

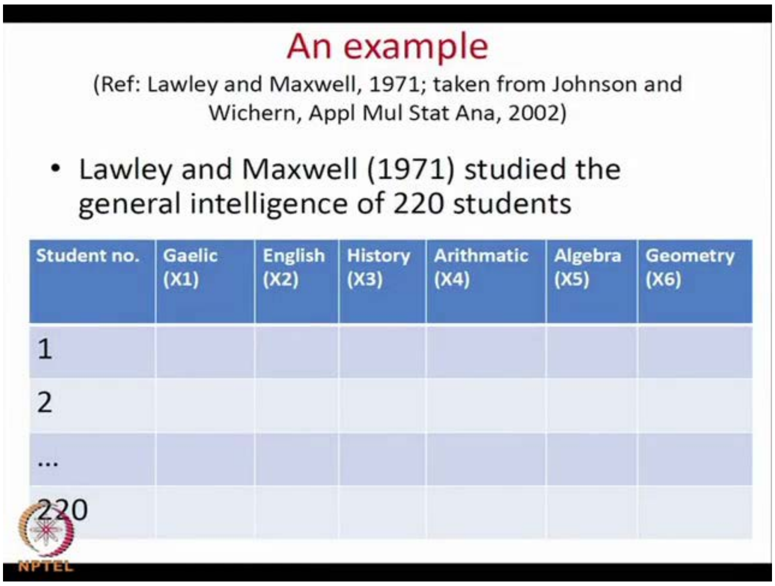


Now, with respect the another one concept, here what will be the covariance of X and F x is your original variable, F is the factor. So, your expected value of x minus mu and then F minus expected value of F transfuge, that will be the case. Expected value of F is what, 0. So, ultimately this is expected value of X minus mu F transfuge. Now, what is X minus mu X minus mu is lambda F plus delta.

So, then X minus mu F transfuge is lambda F delta F transfuge. So, lambda F, F delta plus delta F delta. Now, if I take expected value of all those things. So, ultimately this equal to expected value of this equal to expected value of this, expected value of this, it will delta will lambda will come expected value of F ftransfuge plus, this one will be 0.

Now, F ftransfuge expectable is I loading, that loading matrix is coming into coefficient variation. I think that to you I have explained the primarily, the concept conceptual model of factor analysis. This whatever we have discussed it is just the conceptual model, that what is factor analysis. What is exploratory factor analysis and how the variance of individual object variable have partitioned into communality and specific factors, and but what is required. Definitely our next topic will be estimation. Estimation of lambda. Lambdas, that lambda matrix you do not know this. This lambda matrix is not known to you, you have to estimate. There are different method of estimation.

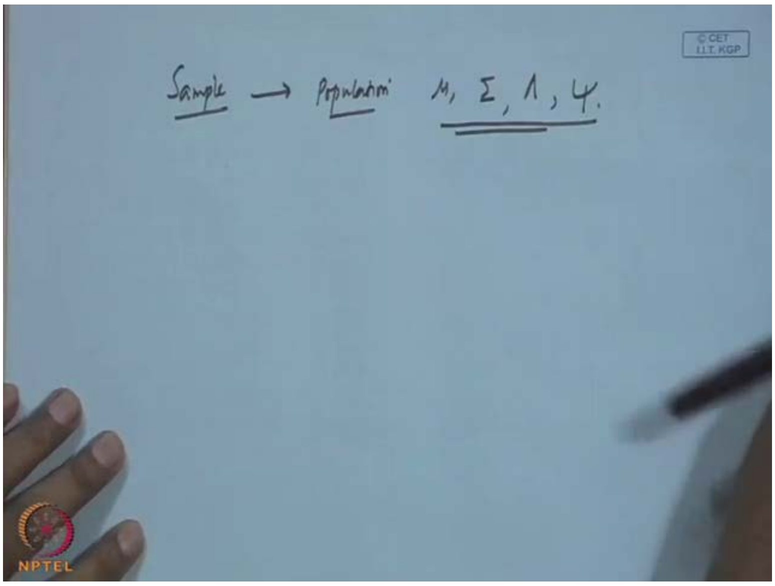
(Refer Slide Time: 58:34)



But I thought of giving you one example. Let us see one example, here this example with reference to factor analysis. This one is a study conducted by Lawley and Maxwell in 1971, the purpose is to measure the general intelligence. So, for students 220 students were considered and the variables that were taken, they are object variables. So, these are the verbal of first one gaelic, English, history. These are basically the verbal subjects and arithmetic, algebra, geometry. These are your mathematical subject. What happen you do not have to test the general intelligence of these many students.

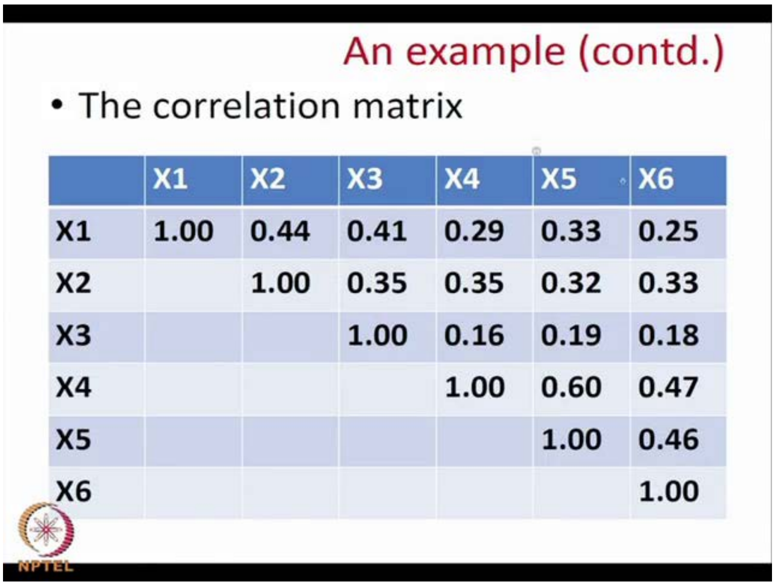
So, this six subject performance that mean performance of each student on the six subjects were what I guess is obtained, and then with these data set factor analysis was conducted. And to see that whether that the six X variables, what you are observing? They are actually six or these can be combined into several reduced factors number of factors. So, that means what is happening here that we are talking about sample.

(Refer Slide Time: 01:00:10)



But so far what I have discussed? We have discussed for population, because we have taken mu. We have taken sigma. We talk about lambda. We talk about psi, all those things related to population that will be during estimation time. We will see this sample part.

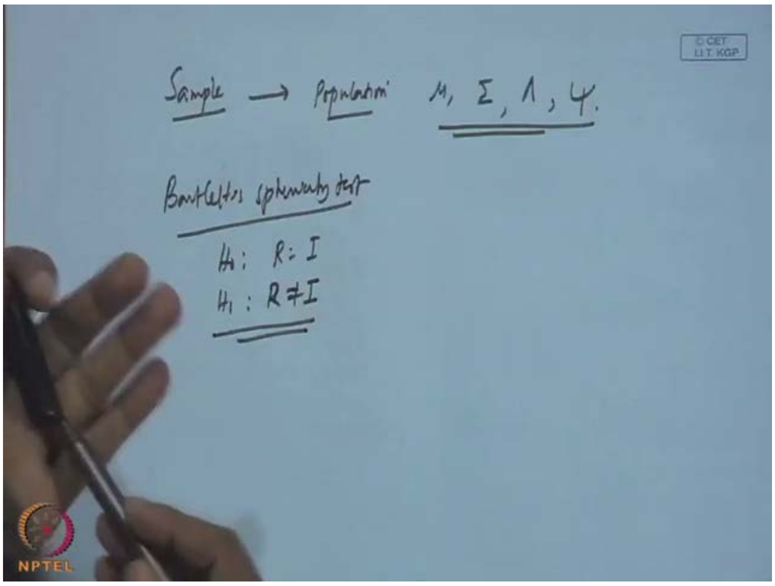
(Refer Slide Time: 01:00:43)



	X1	X2	X3	X4	X5	X6
X1	1.00	0.44	0.41	0.29	0.33	0.25
X2		1.00	0.35	0.35	0.32	0.33
X3			1.00	0.16	0.19	0.18
X4				1.00	0.60	0.47
X5					1.00	0.46
X6						1.00

But let us concentrate on this example. Now, then the immediate step toy have to see the correlation matrix. You see that if I go by the correlation coefficient between the variables, there are few variables point with correlation more than 0.3. For example, X 1, X 2, 0.44, 0.41. Like this 0.60 is the maximum. So, from correlation matrix also you can find out certain clues that whether you go for factor analysis or not, that factor ability criteria that mean the data matrix must be factorable.

(Refer Slide Time: 01:01:31)



If you have seen in principle component analysis, I say that Bartlett Sphericity test what we have assumed there? Bartlett Sphericity test we assume H 0, that R is I and H 1, R not equal to I, that mean the correlation goes distant, and combiningly you have tested this. So, here I think we definitely require Bartlett Sphericity test to see that whether we will go for factor analysis or not, because if there is no that not significant to much correlation what is the use.

But factor analysis from correlation co coefficient point of view, if there is 0.3 or more correlation coefficient, a large number of correlation coefficients having 0.3 or more. Then it is recommended, you can go for factor analysis. We will get certain benefits. So, we will stop now, and next class I will explain a parameter estimation using that loading, how to estimate the loading factor, loading matrix.

Thank you very much.