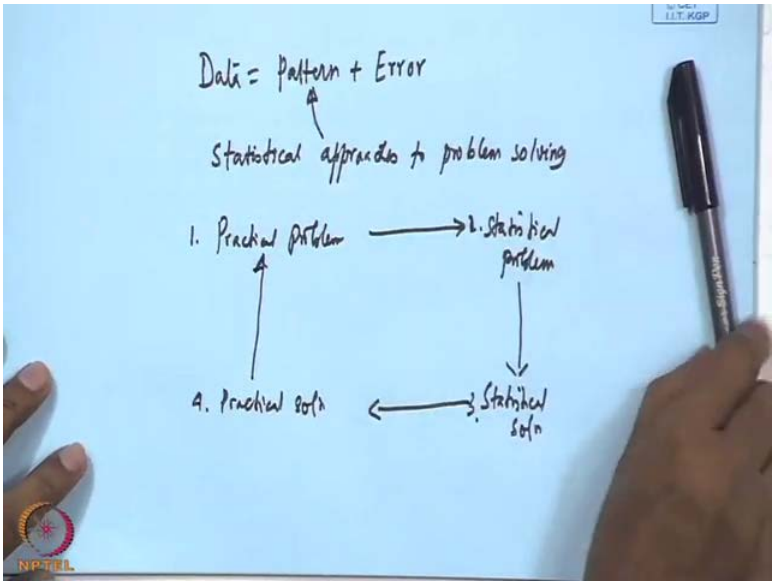
Applied Multivariate Statistical Modeling
Prof. J. Maiti
Department of Industrial Engineering and Management
Indian Institute of Technology, Kharagpur

Lecture - 2
Introduction to Multivariate Statistical Modeling
(Contd.)

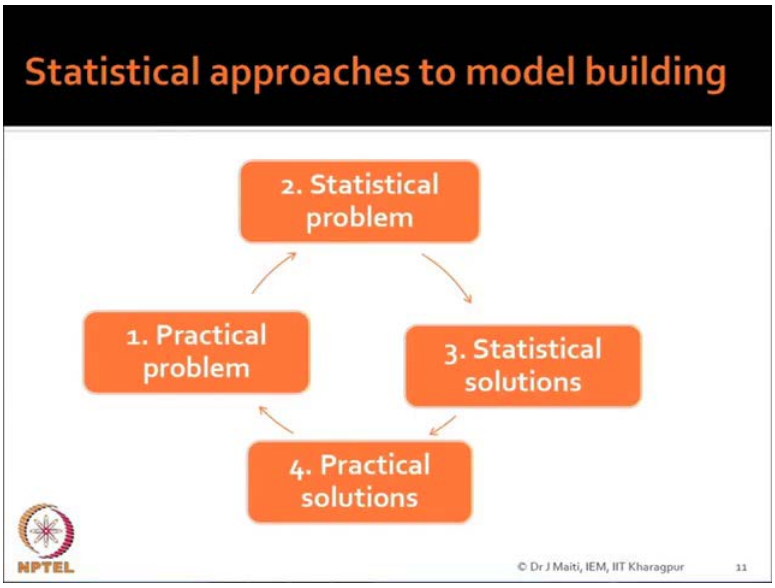
(Refer Slide Time: 00:28)



So, good morning, last class what we have discussed that your data equal to pattern plus error pattern plus error. So, how do we extract pattern that was the issue, so with respect to this we will discuss. Now, what are the statistical approaches, so statistical approaches to problem solving. If you see the slide, you see that here we will start with practical problem then this practical problem will be converted into statistical problem. Statistical problem will help us to generate statistical solution, and this statistical solution.

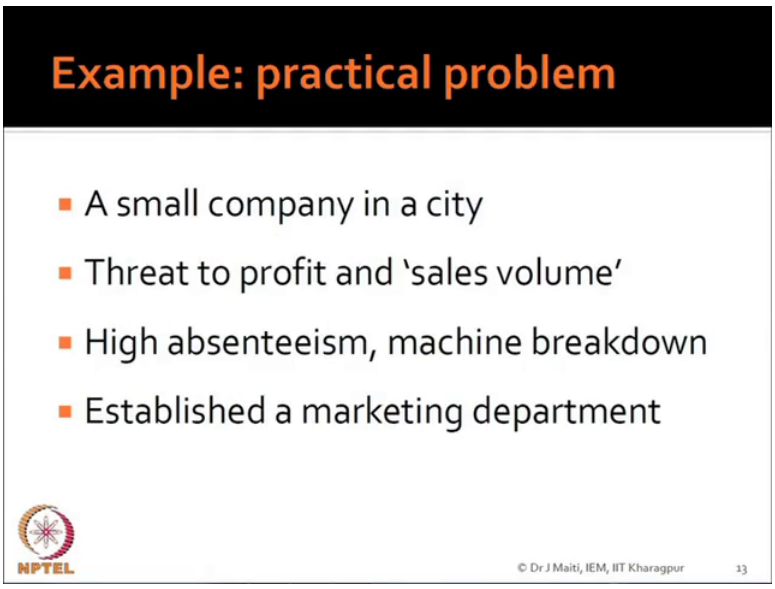
So, if I say that first one is practical problem, second is statistical problem, third one is statistical solution, then this statistical solution must be practical enough. So, that is practical solution, now whatever solution you get statistical solution is converted to practical solution. It should be checked whether the practical problem is truly solved or not, that is why the cycle.

(Refer Slide Time: 02:39)



If you see the slide, you see that it is basically a cyclic one your practical problem and that is the crux of the matter. If you find out the practical problem, nicely identify and define the problem nicely, because rest of the things are following it. So, then the statistical problem then your statistical solution, and finally the practical solution. Many a times, we falter in effectively defining the practical problem, getting me.

(Refer Slide Time: 03:24)

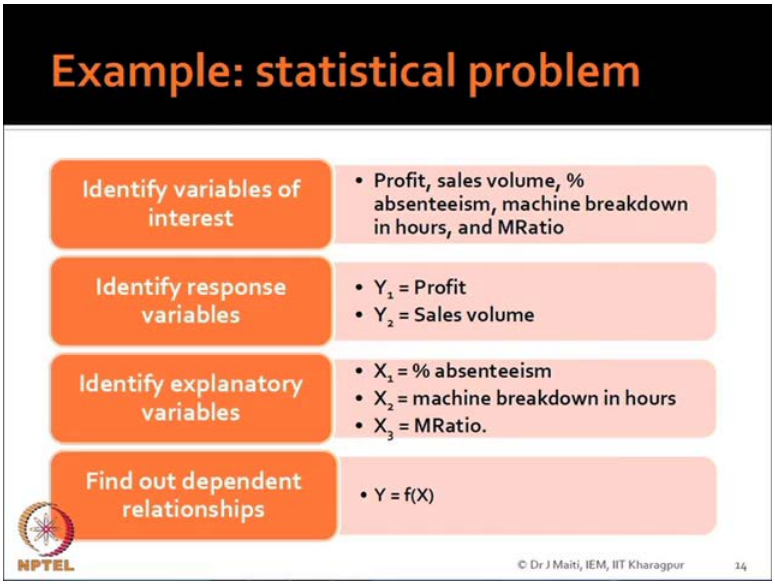


So, what is practical problem how do you understand a practical problem. In last class, that a small company in a city doing business primarily local at the local level and the

company is monitoring its profit and sales volume. It is observed that absenteeism is quite high and there are machines breakdowns also and in order to improve the business performance marketing department is also established may be recently. Then they want to check that how the marketing department is performing. Now, here what is your practical problem if I go back to this slide again, what is the practical problem.

The practical problem is, as I told you that there is threat to profit and sales volume, high absenteeism and occasional to frequent breakdowns. May be the poor performance of the marketing department may not be, but I am assuming that performance of marketing department is also not. That means threat to how to overcome threat to profit and sales volume if high absenteeism is causing lower profit. What is the extent of that effect, then it is better for the management to take actions.

(Refer Slide Time: 05:27)



Now, we will convert it into a statistical problem, if you really want to built statistical problem from practical problem these are the few items or steps which you must follow. First one is identify variables of interest, in these example profit sales volume percentage absenteeism machine breakdown in hours and M ratio. These are the variables what we have identified, now identify response variables out of those variables what are the response variables.

Now, here response by response variable what I mean to say by response variable we want to mean to say that this variables are affected by presence of other variables in the

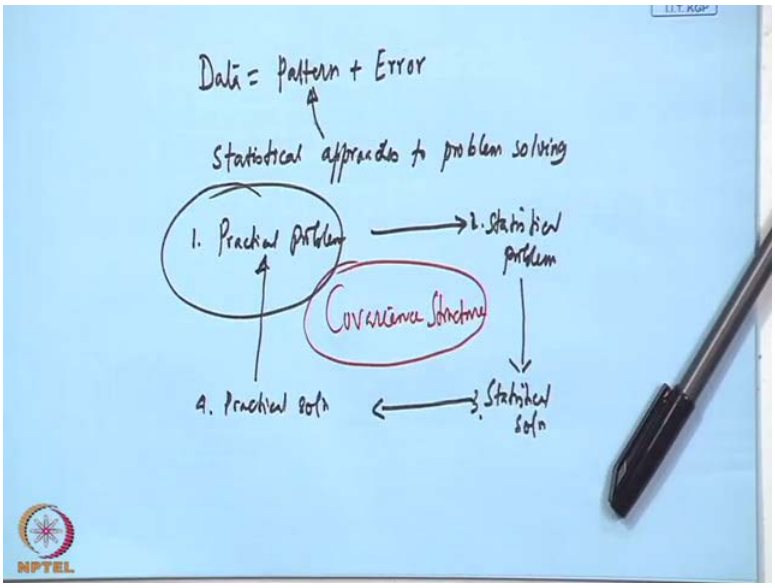
system, getting me. In other words, what we can say dependent variables, so Y 1 and Y 2 are response variable they are dependent variable, why because the absenteeism machine breakdown

M ratio can influence profit and sales volume, so immediately the next objective is identify the exploratory variables and find out the dependence relationship, getting me. For this particular problem, although it is little bit difficult at this moment that why suddenly we are going for this dependence modeling what are the different kinds of dependence modeling techniques are there.

But, you must be, what I mean to say, you should not be governed by the techniques that is one of the threats. Many a times, what happens you know many multivariate techniques and you think that your problem will fit to that technique please, do not do this. That is why I have written here in finding the relationship Y is function of X we have not said it is a linear relationship or non linear relationship or it is a regression. It is something some other way of doing things nothing mentioned, so you must be driven by the problem at hand.

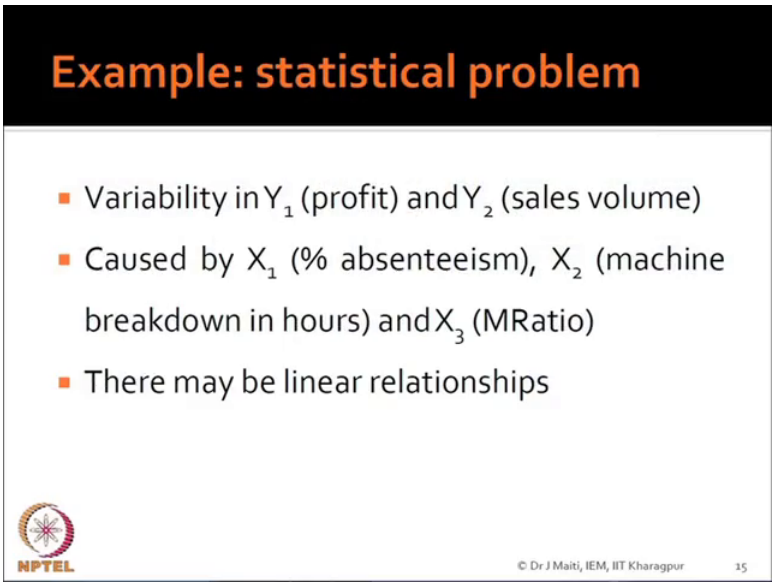
So, our problem at hand is threat to profit and threat to your sales volume and so we got the response variable where we want to concentrate. Accordingly, the other explanatory variables means the variables which explain why there is variability in profit and variability in sales volume. Function of Y is f X in statistics statistical modeling the variability, variability is the crux of the method. Please understand, you must understand the variability structure in case of one variable there is variance. In case of several variables, there is variance and covariance in totality we say that is covariance structure.

(Refer Slide Time: 08:56)



So, that covariance structure, covariance structure this will give you the pattern, the pattern you want to extract this will be given by the covariance structure.

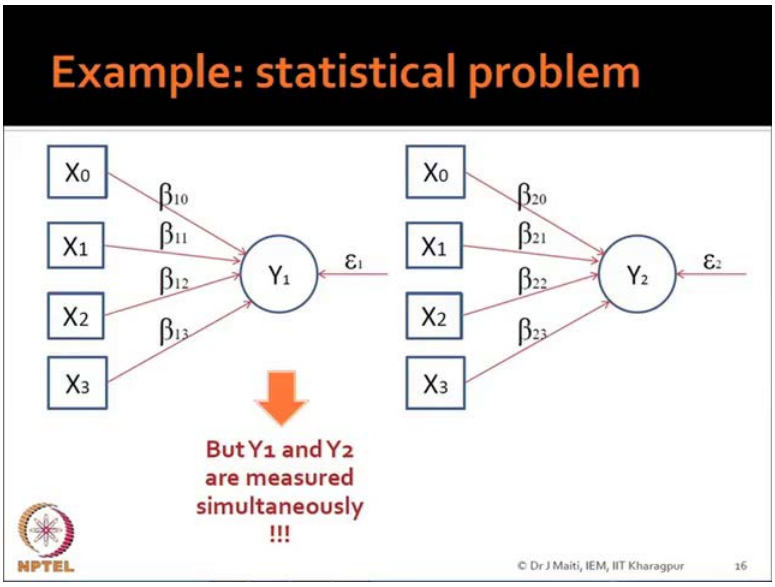
(Refer Slide Time: 09:18)



So, let us see the statistical problem, here the statistical problem is that the variability in profit and sales volume. By variability, we are saying that definitely there is high variability caused by X 1 which is absenteeism X 2 machine breakdown in hours and X 3 M ratio there may be linear relationships. Now, you are coming to statistics domain, you

have to assume something there may be linear relationship any problem up to this. Do you have any query and query from your side, it is this side, no problem, yes.

(Refer Slide Time: 10:16)



Then I want to show, we say that fine there is a your relationship statistical way you can examine that relationship. And we also assume that their relationship is linear. Now, let us see, then pictorially if you see this see this particular figure, here what we have said that Y 1 is affected by X 1, X 2 and x 3 what is Y 1, Y 1 we have said Y 1 is profit correct. So, Y 1 is profit, X 1 is percentage absenteeism, X 2 is breakdown and X 3 is machine hours, so Y 1 is profit, X 1, X 2, X 3 absenteeism and your breakdown hours and M ratio.

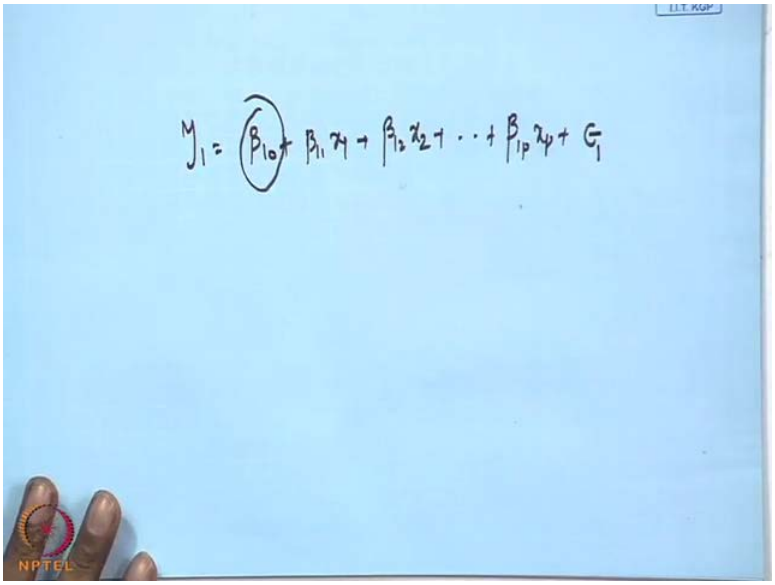
Now, this is the linear relationship between Y 1 and X 1, X 2 and X 3, you may be wondering that what is X 0 then that X 0 what is constant value which will be later on in multiple regression. You will be knowing that X 0 will be given a value of one, for all observation one value will be given and beta 1 0 will be the intercept means, that irrespective of the explanatory variable considered here. But, there will be still some value for Y and that will be determined by beta 1 0, let it be slowly you do it later on, we will see.

But, the sole purpose of this particular figure is that what I mean to say that Y 1 if you want a linear relationship, you can pictorially represent like this. Similarly, for sales volume also you can represent like this in the pictorial representation. Please keep in

mind two things, first of all the arrow head is see that it is suppose X 1 versus Y 1. If you consider that the arrow basically starts with X 1 and ends with Y 1, and arrow head is at the Y 1 level. So, that it simply indicates that Y 1 is a dependent or affected variable and X 1 is the causal variable or explanatory variable.

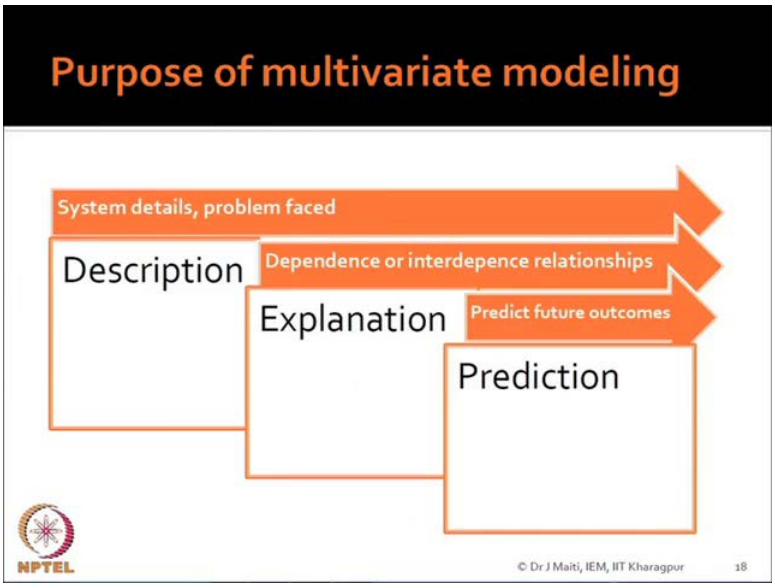
Another issue is the epsilon 1 that is basically some error component as I told you that two parts, one is pattern and error. Apart from this arrow, the rest of the things are basically pattern beta 1 0, beta 1 1, beta 1 2, beta 1 3, all will reflect the pattern, similarly for Y 2. But, for this particular company, your Y 1 and Y 2, they are sales and volume and profit and for the same one they are simultaneously occurring. So, if you go by two different models, linear models will it suffice that is why what I have written Y 1 and Y 2 are measured, simultaneously they are co varying also like beta 1, p X p epsilon 1.

(Refer Slide Time: 14:33)



So, what is this, now what I am saying further that Y 1, Y 2 you are measuring, simultaneously do you want to keep the structure intact while estimating, while modeling, what is the structure. Here, both things you are measuring simultaneously, you see the next slide here, well next slide left one, what you are doing. Here, you are basically saying, no both are occurring simultaneously and they are multivariate observations only. So, what I want to want to keep them in one model, I will not go for two linear models only in one linear model I will do this.

(Refer Slide Time: 15:46)



But, there is another problem, what is other problem other. Problem is Y 1 is basically and Y 2 3 may be relationship between the two, it may so happen that Y 1 may affect Y 2. That means Y 1 is not only a dependent variable or response variable, it also becomes a causal variable or explanatory variable, it all depends on the situation. Suppose, any of the situation this is not the case no problem, but many a times what will happen you will find out that this type of structure is there. So, when such type of a structure is present in the practical problem, how can you ignore it, you cannot ignore it.

That means you have to keep the practical that behavior real behavior in fact and then find out the model. Not the other way round, find out the model with the data and accordingly you say that is the behavior of the system studied, it is not like this. So, actually what happened there are few models, by saying this few models are I say discussed pictorially one is multiple regression, one is multivariate regression, then path model. Later on, we will see that what are all those things, now what is the purpose effectively, basically from statistical sense I said that what is the purposes of multivariate modeling at the beginning?

But, from statistical sense, what are the purpose you see what first purpose is description any model you built is definitely describe what is the problem at hand. What are the purpose of that study and what are the different variables involved in this particular problem and the how these variables are measured. How these variables, I can say are

stored or kept from what data source it is found source, whether it is a primary data mane you have just gone and collected or taken from some other source. So, all those things are coming under description, then these description part should be done ritually. Infact, everything should be done ritually descriptive, you should not falter because this is the problem definition part then explanation, what is explanation, the relationships between the variables amongst the variables, what is the relationship that is coming under explanation then prediction.

So, whenever we talk about any model, we talk about description of the model, explanation of the relationships of the variables, building which is basically used to build the model and prediction. Now, many models are not able to predict, so that means in a model, any model if the first two portion is missing then that is not a model. You will just develop a structure like this, a schematic diagram like this and describe something and you will say that is my model. That is the description part, that is not the model, you have to explain the relationship, that means in this particular diagram, this is the description plus relation, where is the relation.

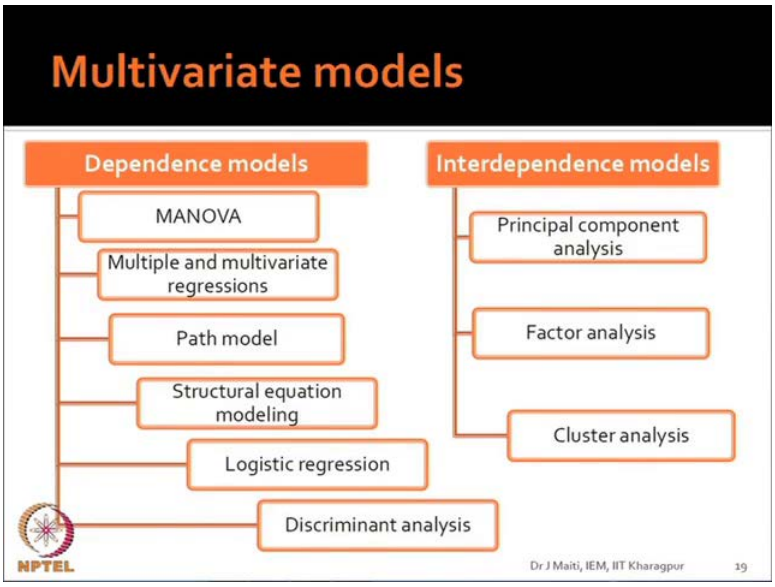
Relation is gamma 1, gamma 2, gamma 1 3, gamma 2 1, gamma 2 2, these are the relations and without this gamma and beta values, this is only a description. If you use certain model which will give you all those beta gamma estimates, and then explanation is completed. Now, using this side, the left hand side plus the estimates of beta gamma and errors if you predict some values for Y 1 or for Y 2, that is what is prediction any model.

I understand the problem, this is the first class, basically second lecture here, now this what i am trying to give you a pictorial view here. Now, how this beta 1 gamma 1 or beta 1 gamma 1 on all those things mean what is the issue and how it will be estimated, all those things will come slowly in the subsequent lecture. You will be knowing the estimation process where this is the only diagram fine I understand, now your question is this one, this side actually in this case this is a representation of multivariate regression.

Here, we consider all these variables are purely independent and in this model case, this is a pictorial representation of path model. Here, we consider that the independent variable can vary, co vary, getting me, but in reality you will not find out that all variables are independent, truly independent. So, this is better representation, slowly just

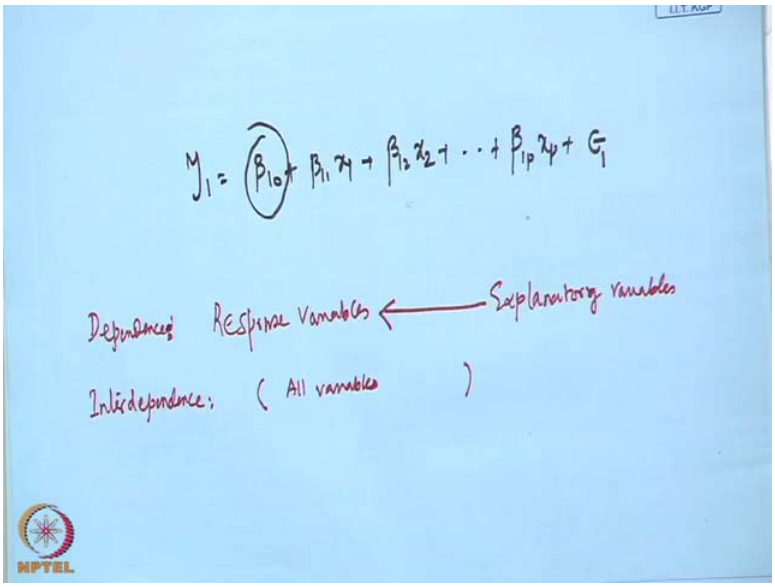
if multicollinearity is a problem, then this is a case, in few later subsequent lectures it is there.

(Refer Slide Time: 22:55)



Now, what are the different multivariate models that will be discussed in this subject. So, I have given you some idea of that what is multivariate analysis, what is multivariate statistical modeling, what do you mean by multivariate and also different data types, how practical problem can be converted to statistical problem. In statistical domain, what are the problems then there we will now see that there are different types of multivariate models and for primarily these models will be discussed in the subject. So, multivariate models are categorized into two broad groups, one is dependence models and the other one is interdependence model. By dependence model, we say there are two sets of variables.

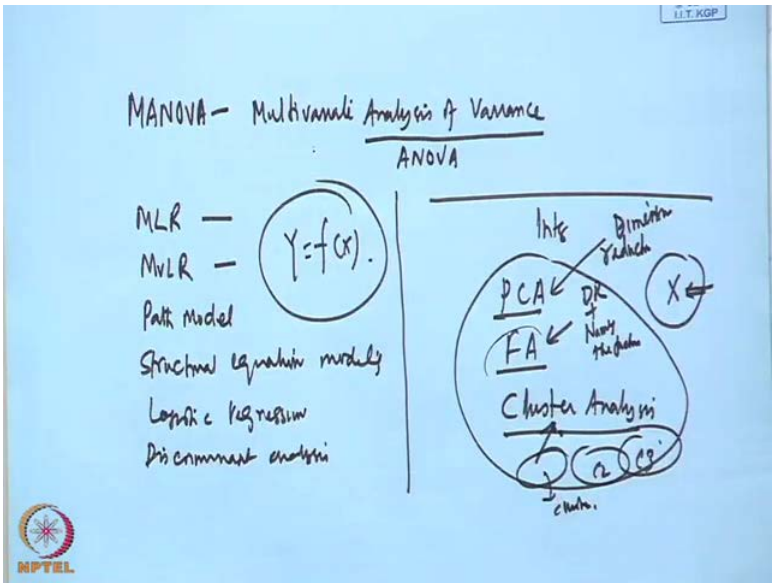
(Refer Slide Time: 23:59)



One is response side or response variables another is explanatory variables, getting me. There are two sides, explanatory variables are used to explain the response variables, so that is what is dependence then that is the dependence structure and interdependence is interdependence. If you see that is interdependence model, there is no response side all, variables are coming under one bracket, all variables under one bracket.

What do you mean to say there is relationship definitely, but within this the same set of variables, there is no categorization, that ne is dependent variable or response variable other is explanatory variables. So, based on this concept, so there are dependence model, there are interdependence model, so in dependence model MANOVA is there which is multivariate analysis of variance.

(Refer Slide Time: 25:29)



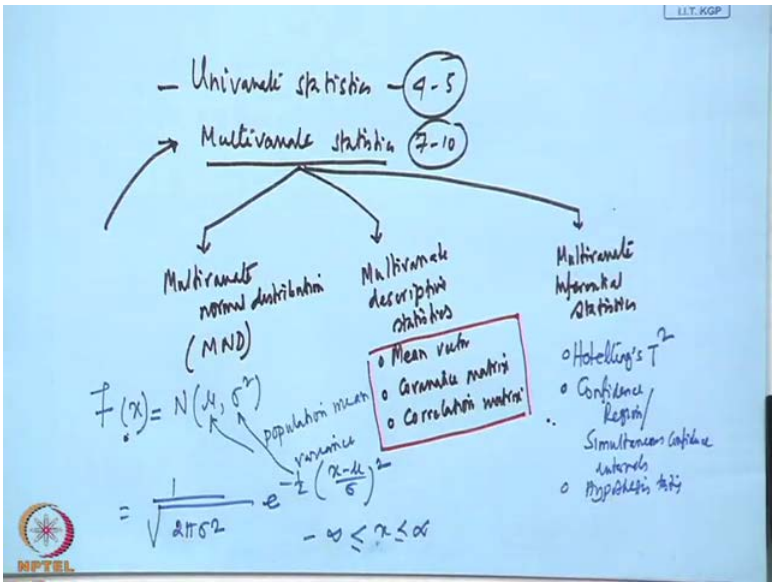
Multivariate analysis of variance, I think all of you what is analysis of variance that is ANOVA and it is presumed that the ANOVA is known, then only MANOVA can be understood. Then multiple linear regression, then multivariate linear regression you have path model structural equation, modeling is another component in addition logistic regression discriminant analysis, there are many more dependence model. But, in these sets of models what is the basic feature is that we have two sets of variables, one can be affected by other set of variables. We can build model like Y equal to function of X, this type of relationship model we can build other sets other group that is what I say the interdependence model.

So, under interdependence model case they are principal component analysis, factor analysis, cluster analysis. Here, we will not segregate the variables into dependent or independent side, we take everything as just one set of variables. Then we want to see the covariance structure of these of this X variables, there are many X variables. Based on this structure, our primary objective in principal component analysis is reduction of dimension, so dimension reduction factor analysis not only dimension reduction plus naming the factors. So, dimensional reduction plus naming the factor in cluster analysis, we want to group the individuals not the variable factor analysis and P C A.

We basically group the variables, so then we reduce the dimension and then in cluster analysis we will not group the variables. We will group the individuals or items or

objects on which data is collected and then we make several clusters, cluster 1, cluster 2, cluster 3. So, like this several clusters will be formed before that because we will try to cover so many models it may not before possible to cover all. But, substantial number of models will be covered here, but before that what the basics of multivariate statistics is very, very important that will that will be covered.

(Refer Slide Time: 29:05)



So, as far as the sequence of the lecture is concerned next lecture we will be covering your univariate statistics a few lectures on univariate statistics followed by multivariate statistics. So, there may be few hours, 4 hours or 5 hours, let it be 4 to 5 hours of this univariate case then multivariate statistics may be around 7 to 10 hours will be on multivariate statistics. Understanding of multivariate statistics is very, very important for all of us, otherwise you will not be able to explain the multivariate models. Now, what are those models, where from these models are coming, what is the utility of these models, we will say many things and you will not be able to explain.

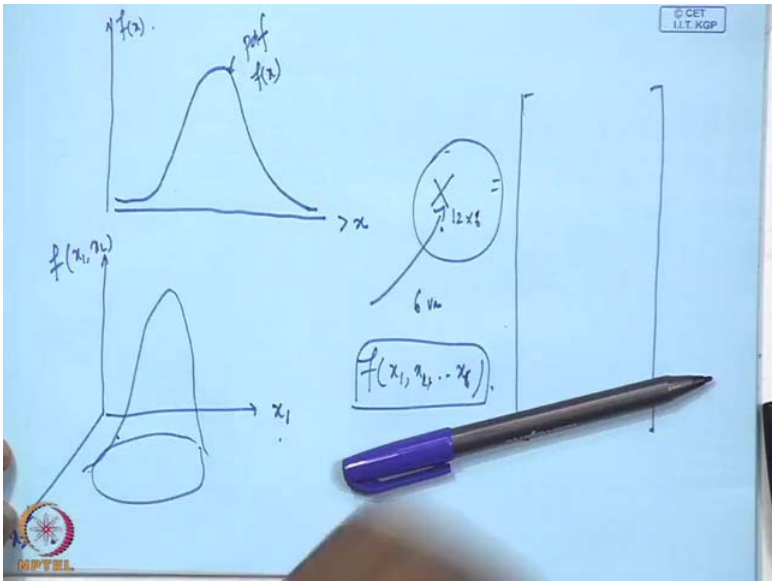
So, the backbone for all these multivariate models that I have shown you just few minutes back is multivariate statistics. Under multivariate statistics, there will be multivariate normal distribution, multivariate normal distribution which we will be denoting like MND and followed by that will be multivariate descriptive statistics. Descriptive statistics which covers primarily mean vector, covariance matrix correlation matrix, so under multivariate descriptive statistics we will be concentrating on this. Then

under multivariate statistics, the next issue is multivariate inferential statistics, so under multivariate inferential statistics, what are things we will be covering.

First is hotelling's T square, getting me hotelling's T square then that is confidence region and simultaneous confidence interval, confidence intervals then your hypothesis testing. So, whatever you have learnt in univariate statistics, that univariate inferential statistics that counterpart in multivariate domain that will be discussed upto hypothesis testing. That in case of one population, in case of two population all those things will be discussed and you will see later on. As I told you that multivariate, your multivariate normal distribution is the crux of the matter, so univariate normal distribution, all of you know and that we will denote it like this N mu and sigma square, where mu is the population mean.

That population is characterized by one variable that is x and sigma square is the component variance of the population with respect to x. This one by square root of 2 pi of sigma square e to the power minus half x minus mu by sigma square, where your this x value lies minus infinity to plus infinity it is visible. So, now what will be the multivariate counterpart of this.

(Refer Slide Time: 33:53)



So, normal distribution if you draw you will be finding like this, this is your univariate normal distribution which is known as probability density function or other way we say F x. This direction x axis will be y variable and your y axis will be your F x, so you want

you have to understand its multivariate counterpart that will be very difficult one, it is not that easy. Suppose there are two variables x 1 and x 2, and then there will be joint distribution x 1 and x 2, so what is the distribution. Something like this you will be getting a two dimensional picture when we talk about only two variables, when we talk about more than two variable that is not possible to draw, but you have think that concept.

So, that abstraction level of thinking is required for understanding multivariate normal distribution. As I told you that multivariate normal distribution is very, very important because most all the models in multivariate statistical modeling subject follows multivariate normal distribution, because if the data follows multivariate normal distribution what will happen.

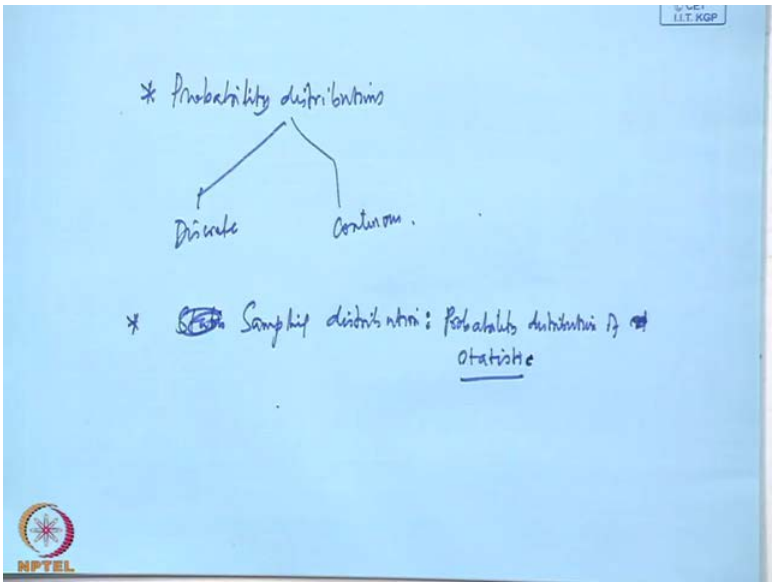
Most of the test can be possible, you can measure the goodness of fit with appropriate statistical distribution, how the statistics what will be used in those models. So, you will know all those things mean vector covariance, matrix correlation matrix, it is like this. Now, with respect to this small company data I think I have given you that the data is 12 into 6 how many 12 into 6 data was given. So, you have seen the matrix, you have also seen this matrix, now if I ask you what will be the distribution of this distribution as there are 6 variables, getting me.

So, it will be a 6 dimensional issue and your joint distribution will be F x 1, x 2 to x 6, so how to get this distribution, so you cannot work now in scalar domain like the univariate case you have to go to the matrix domain. As a result, you will be see that I have written under multivariate distribution that is few things, one is your mean vector, vector is coming into consideration. Second one is covariance matrix, third one is correlation matrix, so we will be discussing this after univariate statistics. Once these will multivariate normal distribution and descriptive statistics is covered we do for multivariate inferential statistics which we will basically discuss based on hotelling's T square.

We will be using extensively the hotelling T square distribution, so apart from this I request all of you to do one more thing, that you please go through probability distributions little bit of this because this knowledge is very, very important. Both discrete and continuous as well as continuous the probability distribution will not be

taught, I will definitely bring some probability distribution. But, I cannot explain in each of the distribution, so you have to go through this.

(Refer Slide Time: 38:24)



Another issue in this subject is that statistical distribution no not statistical, sorry sampling distribution it is also probability distribution. But, it is related to a that is probability distribution of statistic probability distribution of suppose one probability distribution of statistic statistic statistic or there will be statistics. But actually statistics by statistics, we are basically talking about the random variable I will explain in a univariate as well as multivariable by describing a statistics case, how the statistics, how it is a random variable all those things will be discussed to you.

(Refer Slide Time: 39:35)

Now, let us come back to this again introduction part that I told you that multivariate analysis is purely for this particular subject, here it is for the practice for the people who will be using it for solving the real life problems. Now, real life problems what we mean, here some illustrative examples quality control, any idea about quality control for example, you think that suppose in any Kharagpur, there is Tata bearings, there is Tata metallics. Now, Tata bearings will produce the ball bearings roller, bearings this component they produce, so this component has certain quality features.

For example, the dimensions, for example the strength, now customers require the bearings the product with certain quality features. These quality features basically converted into specifications if you produce beyond specifications means not within the specification provided by the customer, what will happen that product will be rejected by the customer. But, if you want to find out why you are producing rejects then you will find out that there is a problem with the process manufacturing process.

Now, manufacturing process variables will lead will affect the product quality, so it is the product quality variables will be the dependent side and the process variables will be the independent side. You can model that product quality a viz process variables, enormous examples of quality control issues are there in the literature using statistical techniques. Second example, basically say one hotel room service case service example, suppose you see that rack storage that hotel such a big hotel I am basically talking about

this hotel, they are basically providing service to many customers. So, they also require some sort of service quality management otherwise impossible, they will not yield to the profit they are interested in.

So, there also lot of data is generated and you have to use those data and develop model and solve the purpose for which the model is developed. Basically, the hotel room service can be much better test of medicine, another issue test of medicine means every time every month you will see that there is a different medicine coming for different diseases. Means for every single disease, there may be 4, 5 medicines are there, so my question is how do I know that medicine works. Medicine a works for some groups, medicine b may work for some other groups even though the same disease or some medicine are performing better that the other medicine for the same disease.

So, that type of testing is also possible using if you collect appropriate data that is possible, now vendor selection is a very tricky problem very, very essential. Critical problem in almost all small medium large enterprises, because your vendor the supplier's quality is very, very important. If you get poor items, poor raw materials supplied by your vendor ultimately your product quality will suffer, so how to go about it, what modeling can be possible there, so that you will select a base vendor. There were different ways of selection, may be someone may go for not statistical route some other route, but multivariate statistics also helps you in selecting vendor safety management.

So, as I told ((Refer Time: 44:09)) few minutes back there are many variables affecting the people's safety, I want to know that what are the variables affecting more. What are the variables affecting less so that I can take appropriate actions which will improve the overall safety standard of the plant of the shop floor of the work everywhere. Now, lean production you see this lean production lean production means what is there in the suppose inventory control is a big issue you are producing 10 items.

But, from raw material side you are keeping a huge raw materials item in inventory it is a loss to the company, so what you want you want to minimize the inventory. But, you cannot minimize inventory unless you have some steps taken some buffer, or something else is there through which you can satisfy the customer at the same time based on your product requirement. Now, lean production says that you minimize everything in such a manner that you will satisfy the customer. But, as well as from inventory point of view,

there are also at the minimum level raw material inventory work, in process inventory everything will be at the minimum levels.

There are very good models in or present research area, but statistics can also be useful there, now damage during transportation you see that damage during transportation, it is a logistic problem, I will show you some. I think one tutorial I will give you later on, that how to model that damage part with respect to that what mode of transport you are using, what is the distance that you are transporting, what type of packaging systems you are using. So, all those things ultimately lead to damage and it is a huge cost to the company because you are basically supplying that products to the customer effectiveness of training there is.

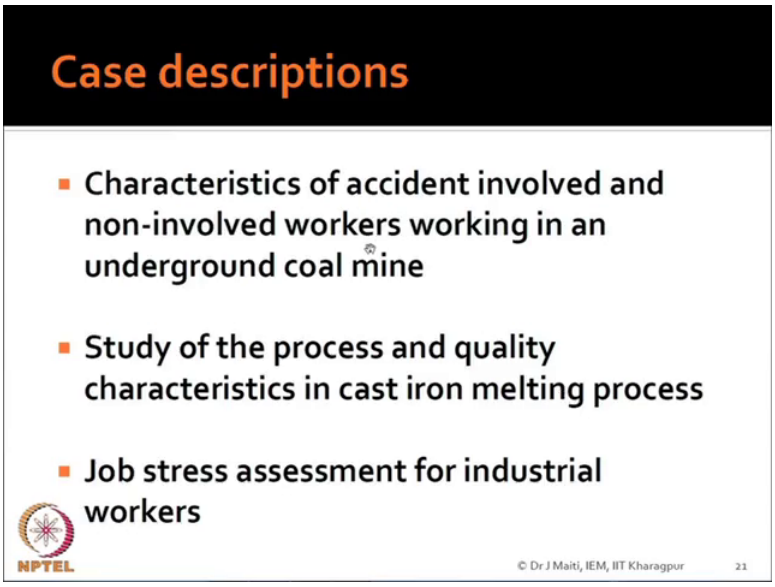
Now, that whether ppt based training is a teaching is based or class black board or chalk and talk there is question like that in the training. In the industry level that whether it is basically a class room training or on shop floor training, what will be there, what will be the teacher's combination, the faculty's combination so many issues are there. So, if you collect data effectively, you will be able to use multivariate statistical models to find out that which mode of training or which type of training or which type of training is better. Marketing is a area where safety, sorry your statistics is used very much for example if you go through any marketing journal you will be finding out that full of statistics.

Many models comes from that side, also marketing research that is the area where you will be using statistics by marketing performance. Basically, we have conducted one study a to understand the purchase intention of customers can be use structural equation modeling there getting structural equation model in there. So, me illustrative examples means some other illustrative examples can be framed, but you are the person who knows your system and please find out some case.

Some example from your work place your domain of expertise, so that whatever I will teach here you will be able to translate to your own system. In an administrative science management, science, social science everywhere this multivariate statistics is used and you have to find out all those things and I will be giving one after another. You will just based on one of this explanation and the things you will be learning, here you try to find out analogy to your system and develop accordingly. Then learning will be complete,

otherwise learning will become incomplete you have to very, very careful for this purpose.

(Refer Slide Time: 48:52)



I will show you later on some cases, so that you understand that the totality as I told you that starting from practical problem to practical solution. That mean practical problem statistical problem, statistical solution practical solution, this total cycle will be completed. Using these cases, first case is characteristics of accidents involved and non involved workers working in a underground coal mine I told you that in safety management. Just I told you that one that in safety management one of the issue is one of the theories were there was there which is now not that much popular.

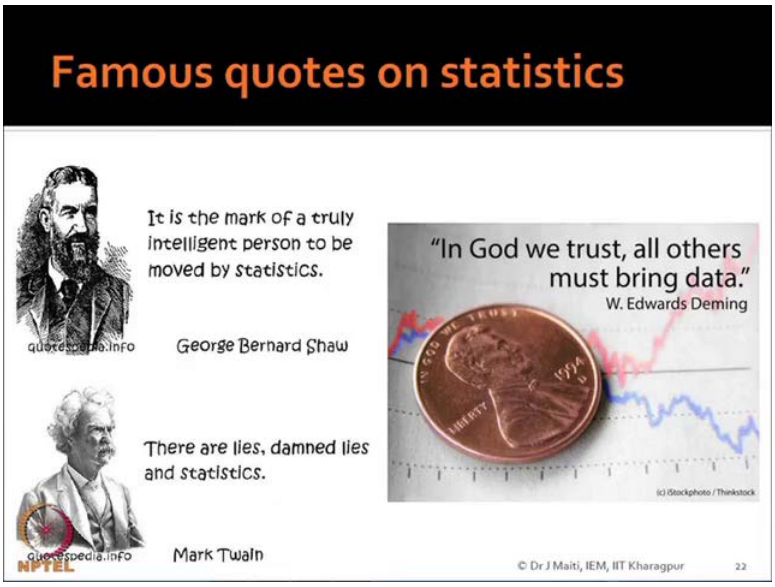
But, this theory still working in that sense that some people are inherently accidental irrespective of the situation they meet an accident. But, some people are able to avoid accident, so that is the issue then we started thinking that whether that people can be categorized as accident prone or not accident prone. So, there are twenty two variables we have collected, and I will show you this case wherever it is required based on the need of the model that will be just right, then study of process quality process variables as well as quality characteristics. In a cast iron melting process, that is also a real case study like earlier one and that where we have that two sets of variables we considered.

We found out the relationship between the two sets and based on this relationship how effectively the quality of melting process can be improved. Third one is job stress

assessment for industrial workers, so there are from officials to that shop floor may be your worker like mechanic or the someone who is doing welding. So, different job profiles are there, their responsibilities and roles vary and accordingly what happened they may suffer from different level of job stress.

So, I will show you in that how the job profile as well as their demographic variables will affect the job stress and you will see that what way it will be done. The similar situation you will face in your work also, then effect you will be able to just straight way translate this to your work. It is possible getting me any question because this is a purely qualitative one lecture. And but next class onwards it will be mathematical and then if there are some, I think all of you know this.

(Refer Slide Time: 52:45)

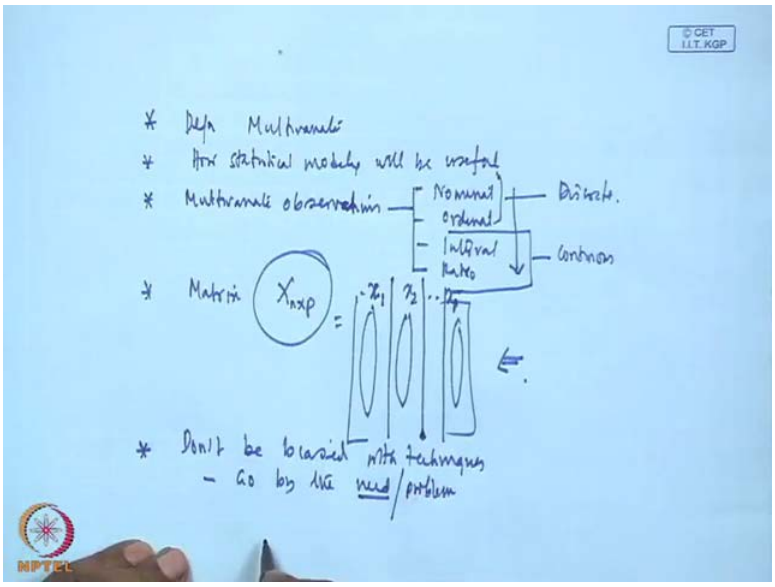


But, there are some famous quotes for statistics in today's, now this lecture with these quotes you know George Bernard Shaw what he has basically said. You see it is the mark of a truly intelligent person to be moved by statistics, now you know Mark TWAIN also, but he says there are lies damned lies and statistics. The two different philosophy Mark Twain is saying that do not believe in statistics because statistics is the maximum lies and that level.

But, we believe in Edward Deming's because ultimately Edward Deming is the quality guru that sense he used statistics in the quality domain, quality management domain. He showed that the statistics can do wonderful in improving quality in terms of totality

quality management what he said in God with trust all other must bring data, if there is no data I will not believe. So, what I we can then summarize what you have learnt today.

(Refer Slide Time: 54:25)



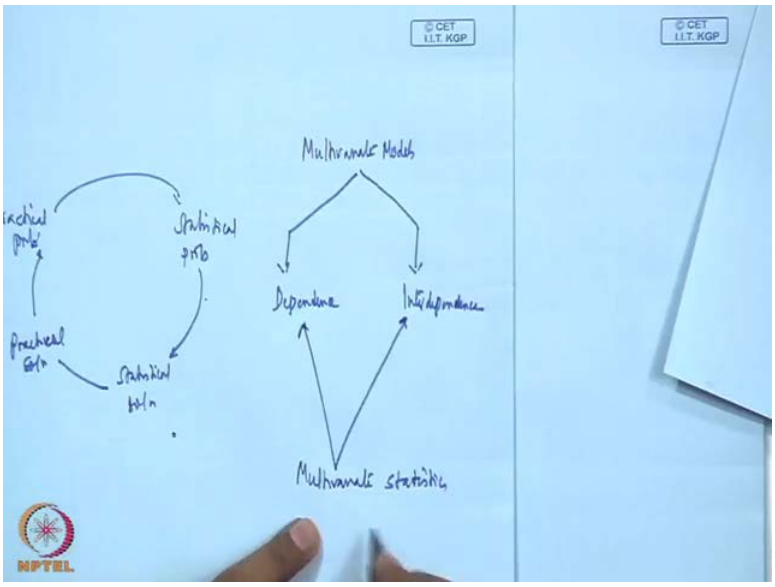
So, today ultimately we started with first is the definition of the word multivariate then we said that how statistical modeling will be useful. We also said that when you talk about multivariate that multivariate observations needs to be carefully measured you require careful measurement of multivariate observations. You must be thorough about the data types like nominal data, your ordinal data, your interval data, your ratio data, if possible you go by this order means if possible collect ratio data.

We have seen also that ratio data and interval data are basically common need to this continuous and first two come under discrete. But, by discrete data we also mean we understand to count data that we will keep in mind then I show you basically that multivariate analysis. Means you have to work in the domain of matrix because X n cross p that is what is our multivariate observation and this one can be written like this. There may be several your vectors like x 1 like x 2 like x p where x 1 denotes the first column, all observations x 2 denotes the second column observation these are the vector variable.

Vector 1 variable, vector 2 variable, vector 3 that is very, very important for all of us, then another thing what we said that do not rely on technique do not be biased with

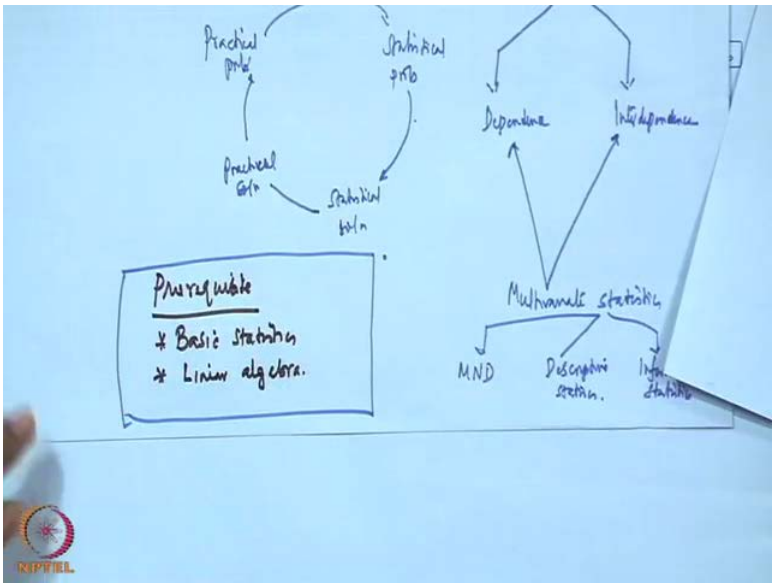
techniques. You go by problem go by the need means what I mean to say go by the problem and you carefully define these and that definition.

(Refer Slide Time: 57:20)



So, that means you are basically going from practical problem to statistical problem, statistical problem to statistical solution, statistical solution to practical solution. Then practical solution to this and what you have understood, here also that you have understood that there are in multivariate models there are two sets of modeling techniques. One is dependence and the other one is interdependence correct, but please keep in mind that these two will not be fully understood if you do not understand multivariate statistics. Mainly the descriptive statistic part, three things you have to understand, here these three things are multivariate normal distribution, descriptive statistics and inferential statistics.

(Refer Slide Time: 58:50)



Last, but that is also very, very important one is that your prerequisite is basic statistics and I think linear algebra that will be better.

So, thank you very much, we will meet tomorrow again.