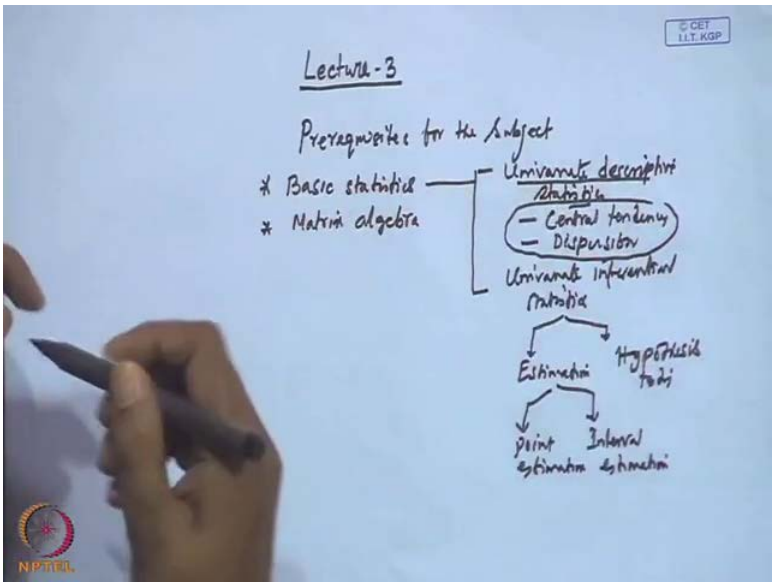
Applied Multivariate Statistical Modeling
Prof. J. Maiti
Department of Industrial Engineering and Management
Indian Institute of Technology, Kharagpur

Lecture - 3
Univariate Descriptive Statistics

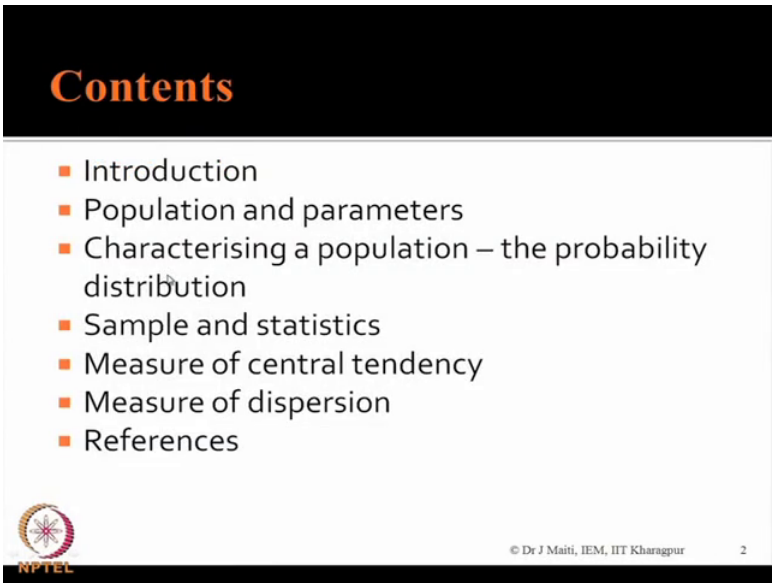
(Refer Slide Time: 00:39)



Good afternoon. Last class we have described the multivariate statistical modeling from the purpose as well as different modeling techniques point of view, and we ended that lecture with prerequisites for the course, prerequisites for this course or subject. And what we have described there that basic statistics is one of the prerequisites and I told u that you also require to know matrix algebra a bit. Now, under basic statistics univariate statistics, the univariate descriptive statistics and univariate inferential statistics are important.

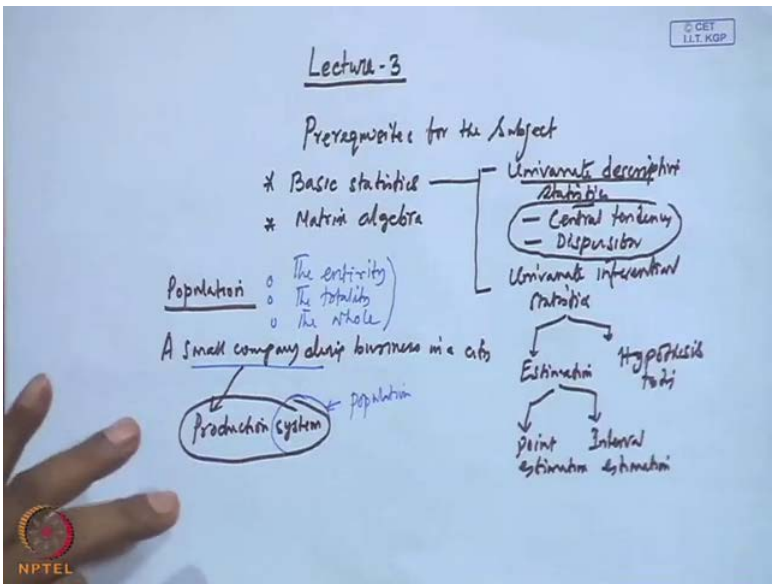
So, again under univariate descriptive statistics usually the central tendency and dispersions, these two issues are described under descriptive statistics. Under inferential statistics estimation and your hypothesis testing, under estimation there will be point estimation and interval estimation. Today, we will in this lecture we will describe this one univariate descriptive statistics.

(Refer Slide Time: 03:08)



You see the content of today's this lecture, we will start with population and parameters then we will describe probability distribution. Particularly, the normal probability distributions then we discuss sample and statistics followed by measure of central tendency, measure of dispersion and followed by references. Now, do you have any idea about population?

(Refer Slide Time: 03:42)

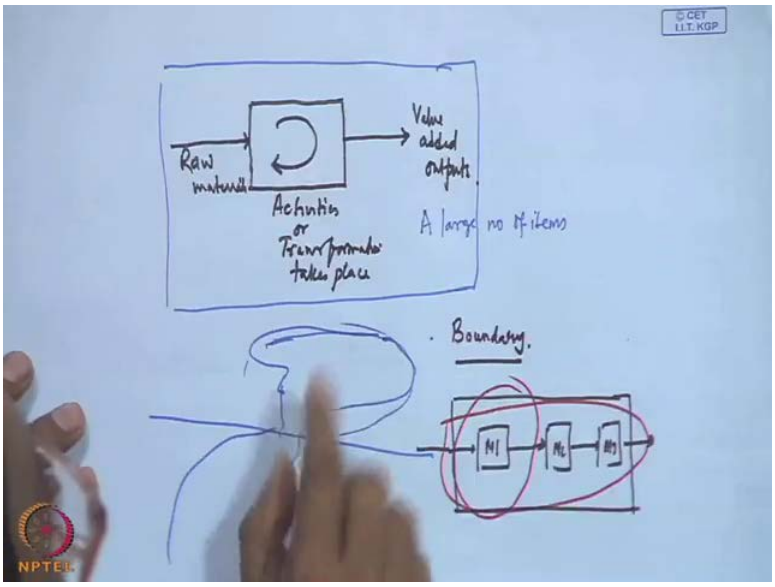


What do you mean by population, general sense we say that the population of West Bengal population of India, but in statistics this population has much broader sense. For

example, in last class we have described one example that a small company doing business in a city. So, the company has a production system, can it be a population if you define population form statistics point of view population if the entirety totality or the whole population. The entirety or the totality or we can the whole when we talk about the population of West Bengal that means, each and every resident legal residents of West Bengal is considered from the production system point of view.

The system, this word also represents population the way we understand application of statistics, so the system is also synonymous for us. It is also population, because system can be characterized by different variables for example, for this company there are profit sales volume absenteeism. So, may other variables are we have discussed, so these are basically which characterize the population or the system another, word could be for us that is a process.

(Refer Slide Time: 06:19)



A process also we can think in this line also the process can be from our purpose point of view, a process is something where transformation or activities taken place activities or transformation takes place. For example, you give inputs as a raw materials and the process production process it converts into value added output. Now, if we consider the total life cycle of this process, then it will produce a large number of items, so large number of items will be produced. All items collectively is the entirety the totality or the whole, so that things with respect to the items produced by this production process.

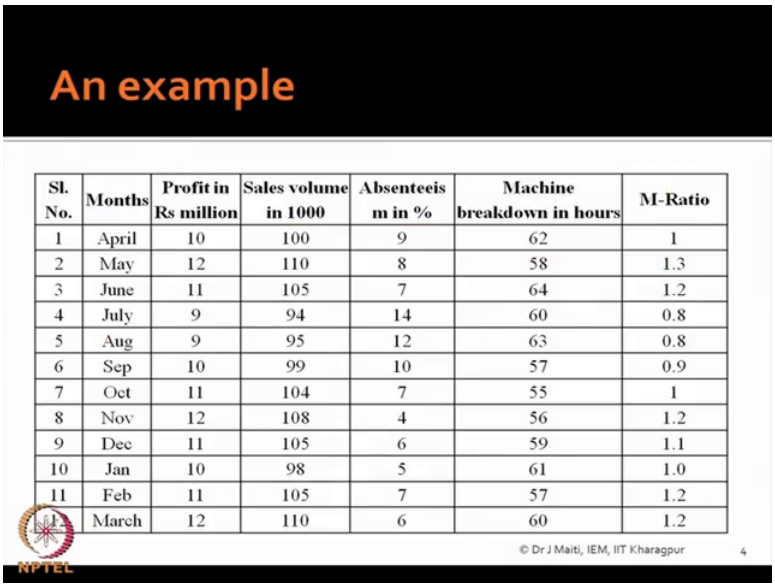
We can define population from, if you go to the service sector for example, the health care system or the banking system there also you can define population. So, essentially if you want to define population and you required to keep in your mind two things that when I talk about the population of West Bengal. Suppose this type, this figure let it be the portion, now the hilly regions population, at the hilly region that is different than the population in the west of West Bengal or south of West Bengal. Now, for a particular purpose you may be interested to understand what the educational status of the people is, if the hilly people of West Bengal.

Then what is happening is you are making a boundary, creating a boundary for the system, so this boundary is the hilly region. So, in that case your population is this hilly region only, now if you think from the voting point of view, suppose the election time. So, this and all the legal that voters they go for voting in that case all the voters of the total West Bengal, they are the population. So, in that sense what is happening that means if you really want to define population, the boundary is very important, getting me.

Boundary in the sense, if you go you come back again the manufacturing scenario, in manufacturing scenario you will find out that the total production system may be composed of several half system that. For example, this may be machine 1, machine 2, and machine 3 and they are doing different operations, raw material coming here and transforms to machine M 1. Then going to machine M 2, some or the other activities is going on, now if you are interested to infer something about machine 1. Suppose you want to infer something about machine 1, then your population is this if you think that are some common characteristics applicable to all the machines.

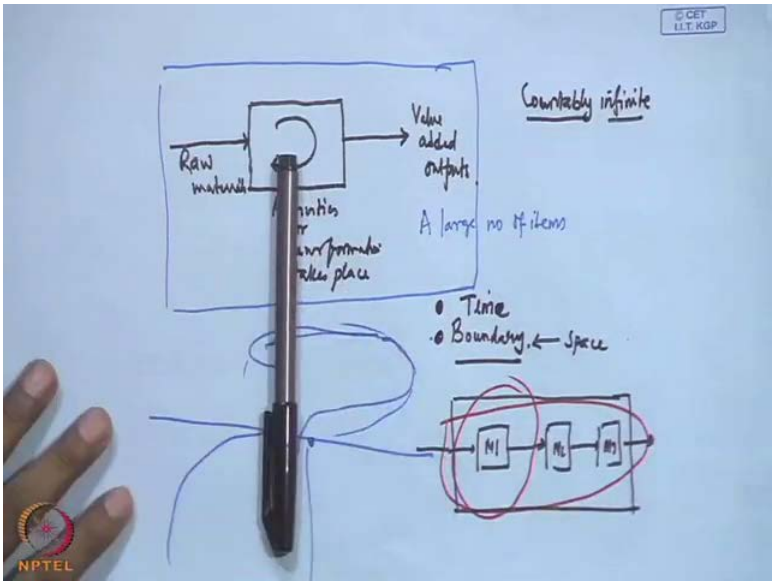
Then you may be interested to see the totality including all the machines, then your population will consider all the machines, this is very important. Unless we understand population, there is no use of statistics because statistics is used to infer about the population, inference related to many things. During inferential statistics, we will be telling you what are the different inferences possible, but for the time being you please understand that when we talk about population, we talk about a system or a case for. Why we require to study the process or the population, because we want to understand the behavior of the process or the system or in terms of the population you want to study the behavior.

(Refer Slide Time: 11:50)



Now, if you see the size of population what will happen? Population can be finite, can be infinite when I am talking about, suppose the production of a process for 1 year, number of items produced per year. If that is my population then it is a finite population, so time is another aspect which also defines, used to define the population.

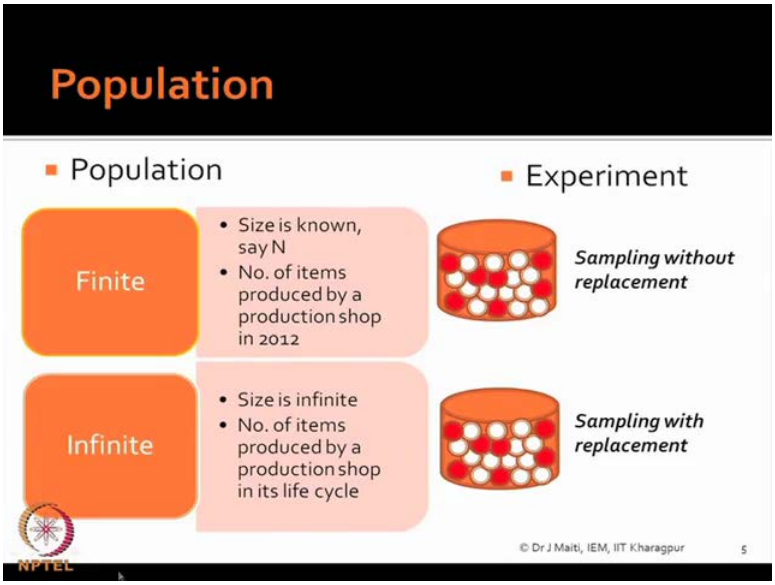
(Refer Slide Time: 12:30)



So, one is the boundary another one is the time, so in two this is basically boundary in the sense, space boundary and the time boundary. So, if you go for the entire lifecycle of a process then what will happen can you count that what are the number of outputs it is

very, very difficult. So, if we talk about the entire lifecycle, total time of the life of the process what will happen the number of items produced will be countable infinite, whether countable infinite or infinite, we will basically define in statistics in two senses.

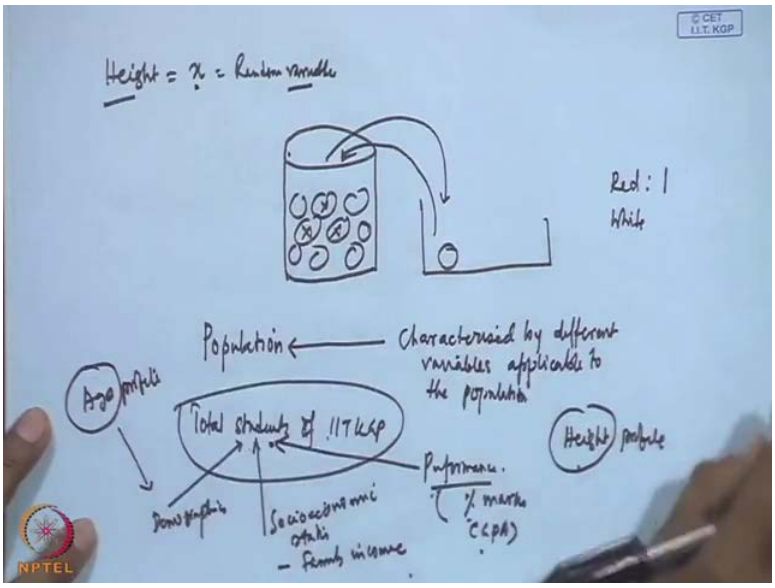
(Refer Slide Time: 13:20)



One is that your population will be finite population or infinite population and finite population will means the size is known that is N. For example, number of items produced by a production shop in 2012, infinite population size is infinite number of items produced that is on the lifecycle of the process that is countable infinite. If you need further explanation as I told you in the last class, that random experiment is the issue in statistics deals with random variables.

Random variables comes from random experiment, we generate random variable based on the experiments conducted. So, if we do one experiment like this, you see this figure inside this if I say this is basically all and inside this there are red and white balls. Now, you pick up one ball, next one ball like this one after another without replacement what will happen after sometime there will be no ball to pick up experiment will end, this is finite population. Now, in other cases see that what we do in the second experiment you pick up again replace, so what will happen in that case.

(Refer Slide Time: 15:02)



In that case, the number of ball will never exhausted, there are so many balls red white what you are doing you are picking up and finding out whether it is a red or white. So, either red or white, so you are counting that red then again you are replacing this, similar manner you are continuing this experiment, the size of the population what will happen. The number of balls will remain as it is from experimental point of view it will be different, so this is what is infinite population?

In statistics most of the issues what will be discussed later on we consider infinite population, so in reality they are it may not be 100 percent true that all populations are infinite. But, countable infinite populations are many and for our practical purposes, if we consider this infinite population there is no problem. Population behavior if you measure, you require to know that what are the variables, that is governing the population in sense characterize the population. So, population is characterized by different variables applicable to the population.

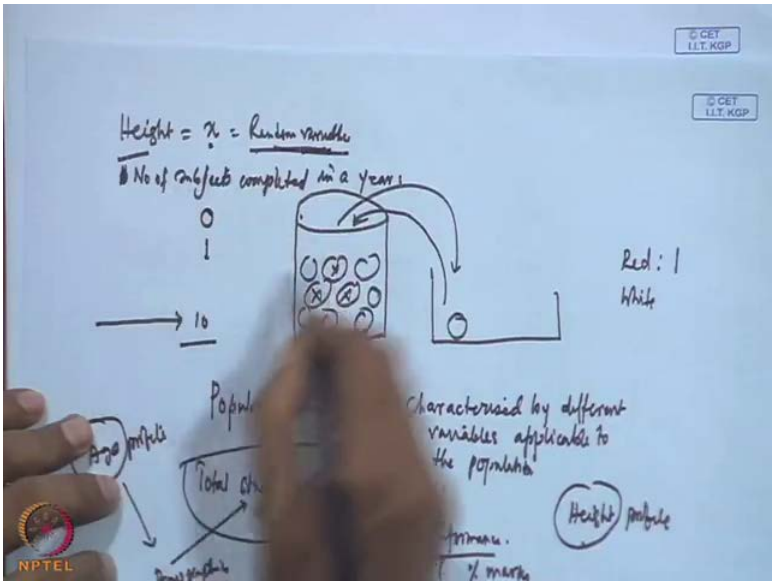
For example, if we consider the total students of IIT Kharagpur, all students of IIT Kharagpur this is my population all students of IIT Kharagpur and the Kharagpur students they come from different demographic. Their demographics differ their socio economic family, socio economic status differ their performance in the graduation that mean in IIT Kharagpur exams that also differ. So, for performance you may be interested

to see that what is percentage of marks of tenth or CGPA your cumulative grade point average or somewhere related to demography.

You may be interested to see that what is the age, profile age sometimes we may be interested to know their height profile you see age, sorry height, age, percentage of marks CGPA. Under socio economic status, family income, all are basically coming under these are all variables which characterize the students of IIT Kharagpur. So, if you want to understand population, not only the space and time boundary we also require to understand what are the variables that governs the population, that is what we see basically.

If you consider any of the variables let height, I am writing height is the students and this I am denoting as x which is a random variable, let it be. Here, we are saying it is random because if we just pick up one student you do not know what is his height whatever you measure you find out from height that is it. So, it is x is, so I want to characterize the students in terms of their height or you may be interested to characterize the students in terms of their number of subjects completed in a year. We will find out that there are many back lock cases, many students could not complete.

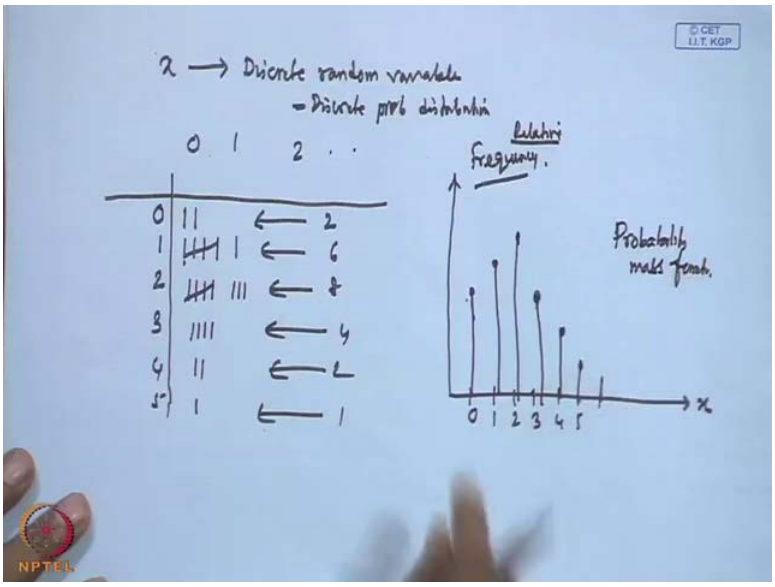
(Refer Slide Time: 20:23)



So, in that sense it may so happen that if we consider that there are 10 subjects to be completed, it may so happen that you will find some students subjects completed, some students 1 or may be like this up to 10, although it will be heavily biased towards 10.

But, this is possible should and depending upon what type of random variable you have considered and accordingly you require to use certain probability distribution.

(Refer Slide Time: 20:55)

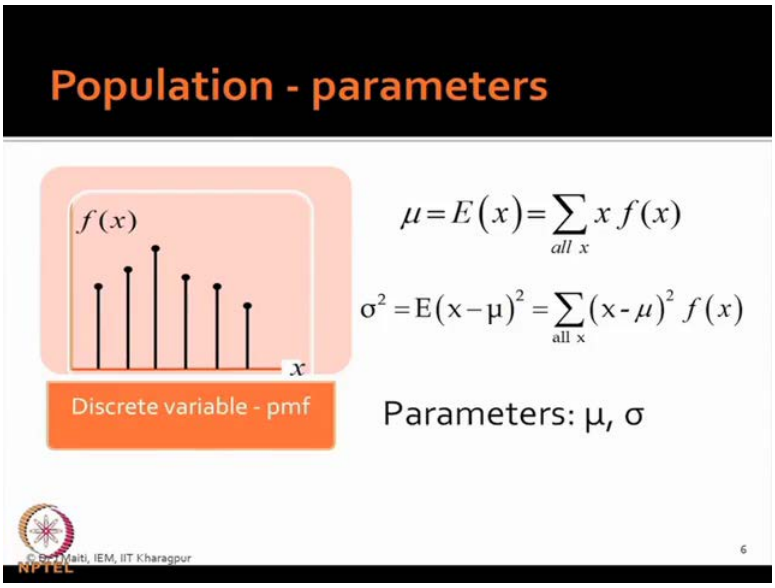


Last class I told you that if the variable is discrete suppose x is a discrete variable, discrete random variable then you have to use discrete probability distribution we discussed last class. But, we have not said what are those probability distribution later on we will see, but what you can see very easily that suppose x is discrete variable it can take values 0, 1, 2, 3 like this.

Then if you make a tally chart, tally chart in the sense frequency 0, 1, 2, 3, 4, 5 like this suppose when you are getting 0 counts you are putting one like this. Then again suppose 0 count then similarly like this what this is the tally count what happened, what is the occurrence of 0 2 times, this 1 6 time, this 1 8 times, this 1 4 times, this 1 2 times, this 1 1 time. So, by categorization what do you mean, here we mean that I have my discrete random variable which can take different values suppose 0, 1, 2, 3, 4, 5 and it appears for different times, I think all of you know.

This is nothing but the frequency diagram and this frequency diagram if I know the total number and if you divide each of the frequencies by their total then you will be getting relative frequency. That relative frequency will give you that empirical probability distribution and this distribution is known as probability mass function.

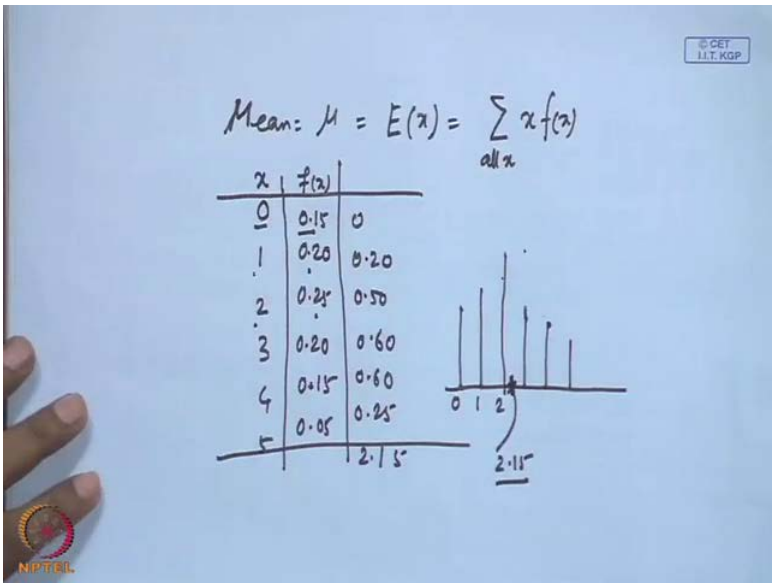
(Refer Slide Time: 23:36)



What we have said this p m f, here is see that this discrete variables when you get this type of plot, you basically developing probability mass function. But, I told you that we will be considering infinite population, so infinite population means that totality is not known. Second thing is that our variable is random, what will happen next minute what value it will assume, we do not know.

So, anywhere in the population domain you cannot get that value and immediately do when you are in, then population domain yes we will get the values when we go for the sampling. But, at least before sampling we do not have all those values, so what you can do for a particular variable which concerned, you can expect something what is this expectation. Suppose, we want to know what is the average height of IIT students this is nothing but the expected value of x, so that expected value of x or the variable of interest this is known as mean.

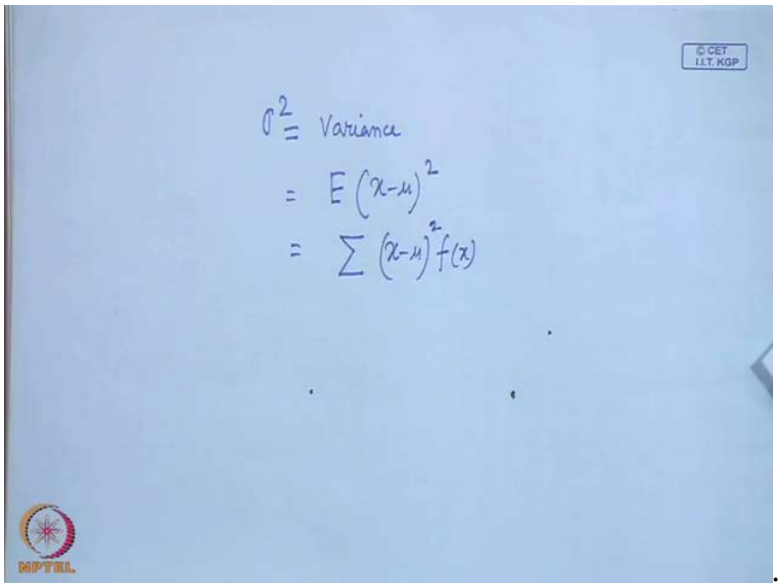
(Refer Slide Time: 25:00)



That is mean, mean stands for mean, mean a expected value of x when your variable x is discrete variable, so you will get like this x f x for this is for all I, sorry all x. Whatever may be the your number for all x if you see this example, here if you see this example, so we are saying here that x can take this value. This value like this there are 1, 2, 3, 4, 5, 6 values, so if I assume that these values are nothing but 0, 1, 2, 3, 4, 5 and so these are all x values. If I assume that they are then sitting, here as discrete also the probability values and their probability values is like this. Suppose the first one is 0.15, second one is zero 0.20, third one is 0.25, fourth one again you can write 0.20, fifth one suppose 0.15 then what is left that will be 40, 60, 95, so 0.05, so then what is your expected value.

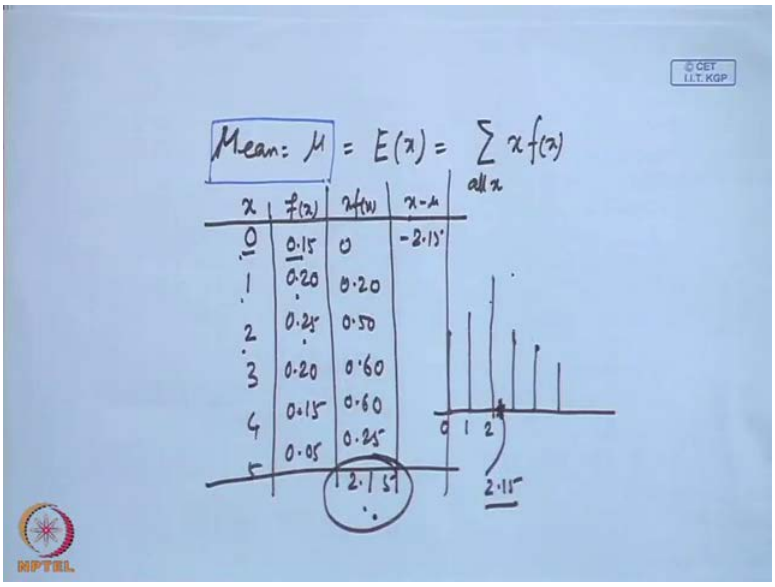
Here, x into f x you have to find out 0 into 0.15 is 0, 1 into 0.2 is 0.2, 2 into 0.25 is 0.50, 3 into 0.2 is 0.60, so like this again 0.60 and this will be 0.25. If you add what you will get, you add 5 6 plus 2 8, 14, 19, 21 so 2.15 so that means that your expected value is if this is 0 this is 1 and this is 2 somewhere here. So, if I draw here I can say that suppose this is my 0 value, this is 1 values, this 1 2 values, and 1 this one is 3 values, this is your 4 and then 5. So, this is 0, 1, 2; somewhere this your value is 2.15 this is what is expectation, but another measure here it is there.

(Refer Slide Time: 28:12)



So another one is sigma square, so we have said here your mu that is mean we have just said as well as, let there is another measure which is sigma square that is the variance what is variance is expected value as x minus mu whole square. So, for this case your discrete case you will write x minus mu whole square f x.

(Refer Slide Time: 28:52)



You have computed here two things, what are those things that you computed x, then your this one with the x f x you know the mu value, mean value you know. Now, you can create x minus mu that means 0 minus 2.15, that is minus 2.15 like this you can

calculate and again you square it multiply it then add it. So, you will be getting the sigma square that is variance part then what we have assumed, here we have assumed that x can take these five values only and this is the probability mass function. So, what will be the sum total of these probability values then what is mu and sigma or sigma square?

Student: mu is long run mean

Correct.

Student: Sigma is...

Long run standard deviation that means you are saying that mean and mean and standard deviation will vary for a population for a particular characteristics.

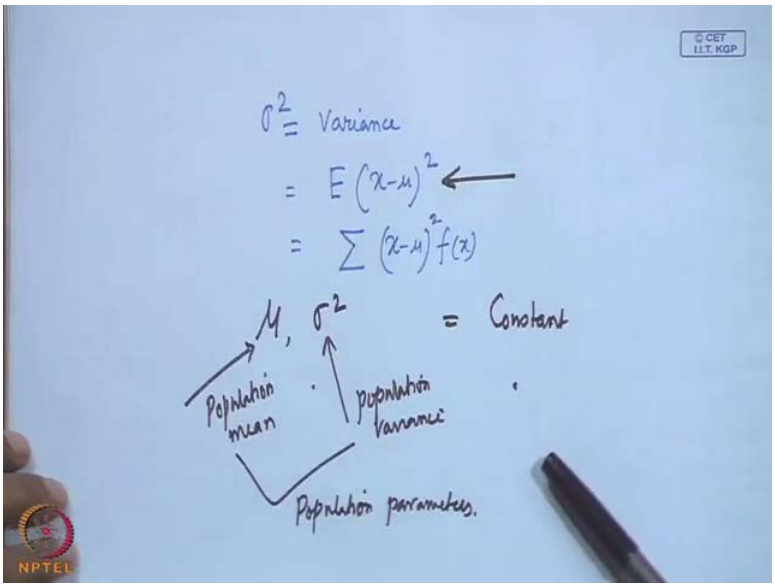
Student: It should not be for a large population.

No, even for small population.

Student: It should not vary.

It should not vary it is a constant, when we talk about a parameter.

(Refer Slide Time: 30:38)



So, actually these are here mu and sigma square in statistically we say population mean and population variance they are constant then another issue will be there they are not

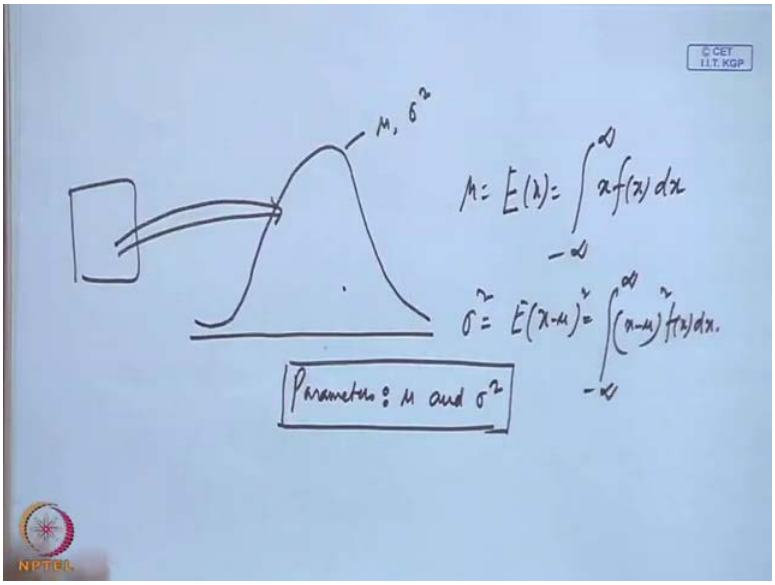
known also. Here, we have assumed a very finite population a very small population and we have calculated something like this population size infinite you will not get all values of x you will never get.

If the size is infinite, that means you cannot measure this or say compute this, but probably what you can do you can expect something that is why the expectation term is used, here expectation is used here. So, if I say population parameter, now you can understand that these two are population parameter it is by saying these two are population parameter.

Please do not consider them there is no other population parameter, these two are some of the population parameters many of the population parameters these two man and standard deviation. Standard deviation or variance they are population parameters, why we go for population parameter, because the lecture is today's topic is very simple topic calculation point.

Understanding point of view, we must understand why we require population parameter, we require population parameter because if you know these two parameter and you also know that your x is random variable and that can follow certain probability distribution. If you know that distribution and also if you know the parameters, what happens you do not require to go for that particular process or system, for further study for this particular variable is concerned.

(Refer Slide Time: 33:04)



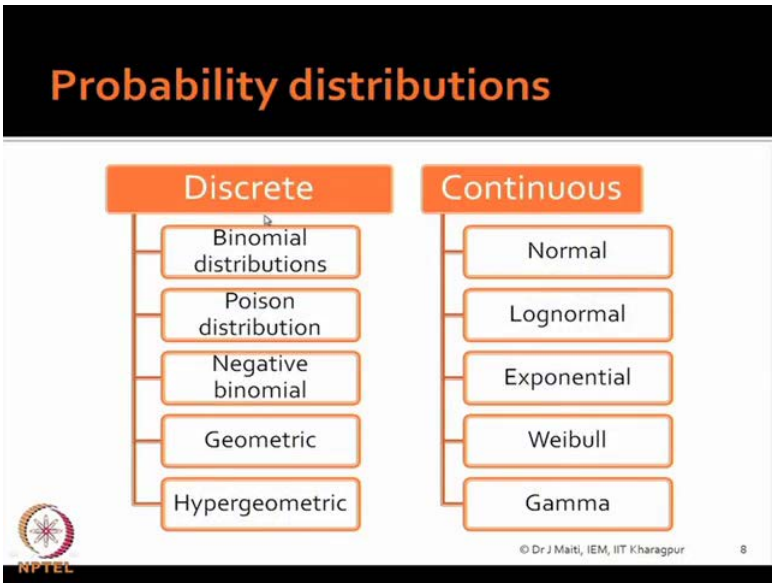
If I say that the absenteeism in the shop floor for the production shop considered it follows something like normal distribution and we know that it is mean. If mu and sigma square is the variance component that means I am in a position to know this, so if I know the distribution what is actually happening. Here, that real production shop from worker performance point of view, the absenteeism is converted to a mathematical equation, a statistical equation. That is the advantage that means if I know truly I know that what is the probability distribution with respect to a variable and what are the population parameters for that variable I have the distribution at my hand.

So, I do not require to go further so long the process will not change by process will not change what I mean to say that suppose it is a machine works overtime machine condition deteriorates that means today a new machine it is performance. Now, after 10 years the machine will not perform same at the same level that means what happens the characteristics changes. So, long the characteristics not changing even the distribution is itself enough for you, now if your variable will not discrete your variable is continuous. You see what is this one left hand side this is p m f or p d f, p d f this is probability density function, now why in continuous case we say probability density function.

Whereas, in the discrete case we say probability mass function you think this one, so here also if we know that this particular population it has mean and variance component. When it is in the continuous level you have to use these two equations for expectation, so basically integration will come into picture. This is integration minus infinite to plus infinite depending on the range for which the variable is defined then f x d x and your sigma square is nothing but again x minus mu the whole square.

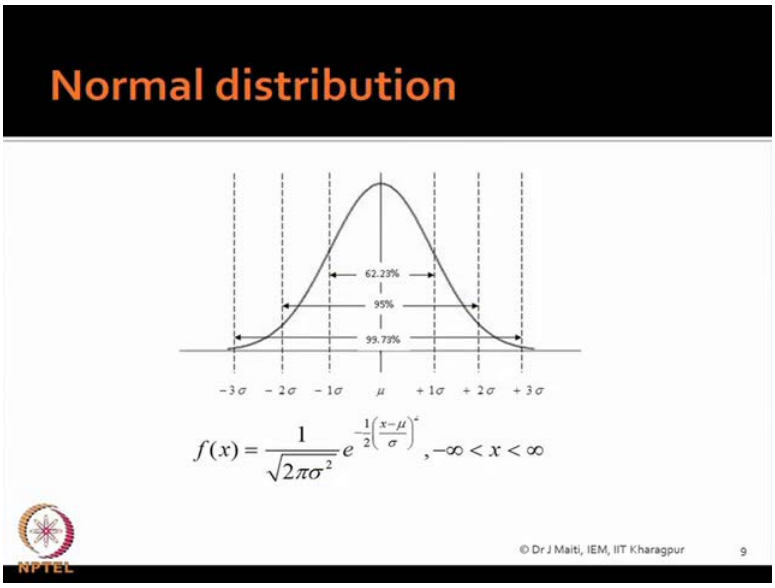
This is infinite to infinite that x minus mu square f x d x, so I am saying the parameter mu and sigma square, here I hope that you understand. Now, what is population and population is characterized by probability distribution if the random variable has a probability distribution. If you know that for that variable the parameters of the distribution, you have characterized the population that is what is known as characterization of population in terms of probability distribution now there are many probability distributions.

(Refer Slide Time: 36:57)



You see here we have we can see, here that under two heads discrete distribution and continuous distribution. Under discrete distribution, binomial distribution, poison distribution, negative binomial, geometric, hypergeometric many more the series. Similarly, continuous normal lognormal exponential Weibull, gamma, so many distribution they are probability distribution we will not discuss all the distributions. We will discuss only normal distribution, here because in multivariate statistical modeling normality assumption this normality assumption is very valid vital one. Many of the models assume normality of the data definitely at the multivariate level that will be multivariate normality, so we will discuss only normal distribution other distribution.

(Refer Slide Time: 37:55)



You can follow Johnson book is there, you can follow this and I am sure all of you are familiar with this distribution this is what is normal distribution. It looks like this and this mu is the center point, here and here that it is basically symmetric. So, maximum number of observations also you will find along this level and it gradually both sides it will gradually reduce. Finally, after 3 sigma level it will almost negligible to 0 level like this, now how to read this normal distribution you see that within 1 sigma plus minus 1 sigma this is very important minus 1 sigma to plus 1 sigma 62.23 observations fall within this.

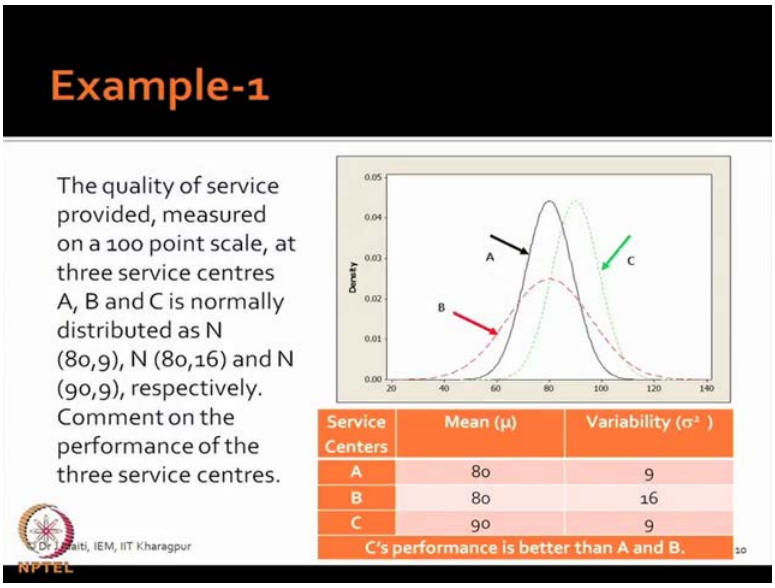
Then within plus 2 sigma level it will be little more than 95 percent, but not 96 95 point something and if you consider plus minus 3 sigma level then your 99.73 percent observation will fall under this category this zone. Now, this is important because suppose you think you are producing something and your variable of interest follows normal distribution, now from data's you will get like this distribution is like this what will happen. That the spread of this particular variable values within plus minus 3 sigma level 99.73 percent of the items produced will fall under within this, now what will happen if the customer will not be interested at this wide range.

(Refer Slide Time: 40:04)

For example, suppose this is my quality characteristics x and this is the lower specification limit this I am giving the physical interpretation. Suppose this is your upper specification limit then what will happen ultimately, suppose this is the mean one this follows normal distribution and your distribution may be like this. So, this may be let it be minus 2 sigma plus 2 sigma, so what will happen ultimately that 5 percent almost 5 percent of your production is rejected product because people that customer will not accept it.

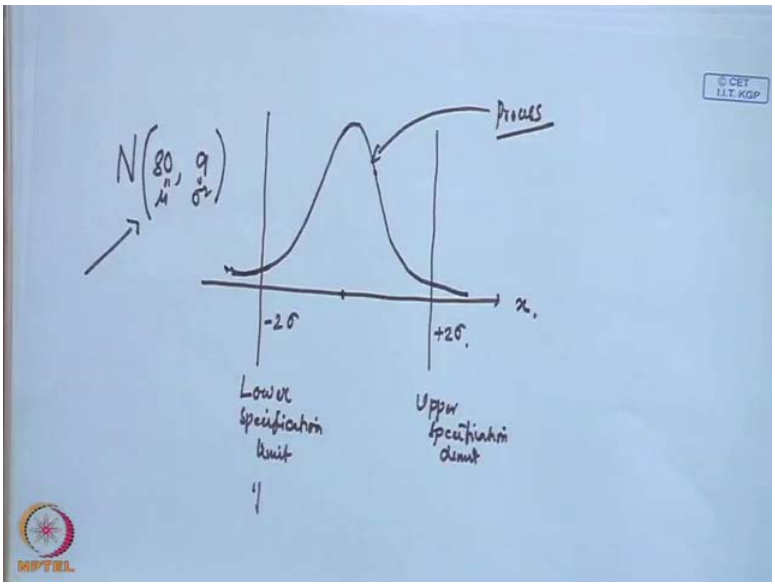
So, when I talk about or say that characterization of the process that means with respect to this distribution, this is your process exactly that this is the shop representation of the process you are not going to the shop floor. But, this is the case your customer region is here and you are producing at this level, 5 percentage of your production is not accepted by the customer you want to improve it you are getting me, you want to improve it how you will do it.

(Refer Slide Time: 41:39)



Can you explain this figure, this figure you see this is again basically as I told you characterization of a process through probability distribution this is another example. The quality of service provided measured on a 100 point scale at three service centers A, B and C is normally distributed as N 80, 9; 80 stands for the mean value, 9 stands for the variance, because our general notation what we will be following is normally distribution.

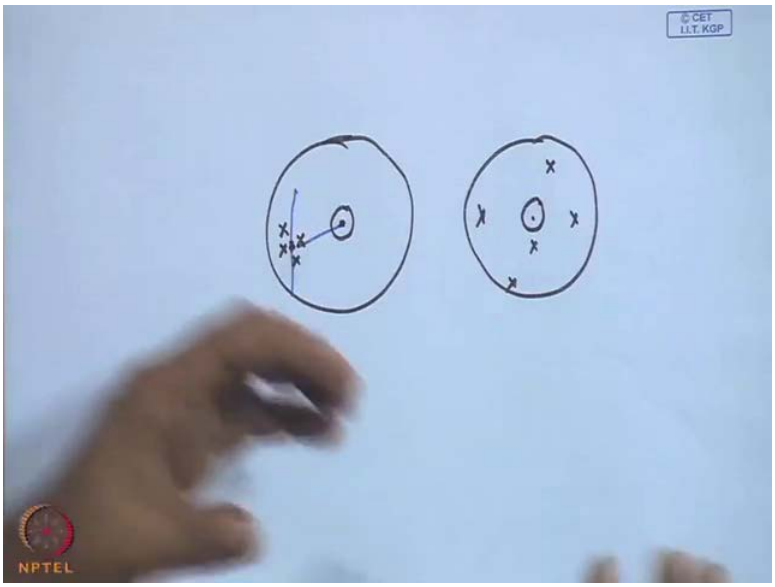
(Refer Slide Time: 42:13)



N stands for normally distributed mu, first is mu that is sigma square basically, so mu is this and sigma square is this, this is the general notation we will be following all through. Then your second process, suppose B 80, 16 and third one is 90, 9 and I have plotted the probability distribution for all the three processes A, B and C and please remember that your variable of interest of quality of service provided. If I ask you which process is better you will say C yes or no, yes why mean yes you are right mean is 90 and is quality of service provided on a 100 point scale.

You are measuring the higher the value better the process, but parallely you see the variability is 9. So, both from mean point of view it is at a higher level and variability point of view it is at the lowest level when you compare the three processes. Now, I ask you from compare A and B which one is better A, because the variability is low mean at the same level, so what is the physical interpretation of this. Then physical what happens it is the variability, the most difficult parameter very difficult to control variability I am giving you another important good example here.

(Refer Slide Time: 44:22)

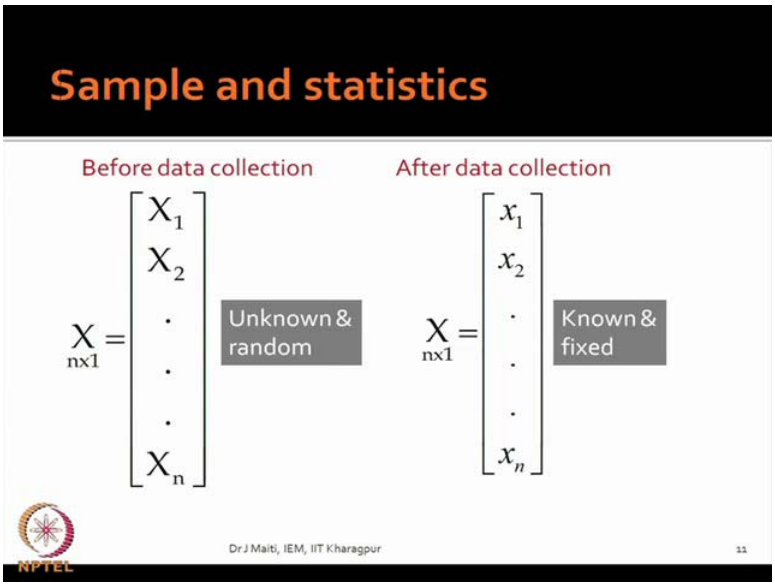


Suppose, you will think that you all know that archery, suppose this is the bulls eye that one our gold medal winner, what is his name that bullet, anyhow this is the target. Now, someone all shoots are here and someone shooting is like this is the bull's eye, you are the trainer two shooters A and B who by training who will be improved first. First one

the precision is first one, the precision level is higher than the second one what will happen you can shift this is mean value to this.

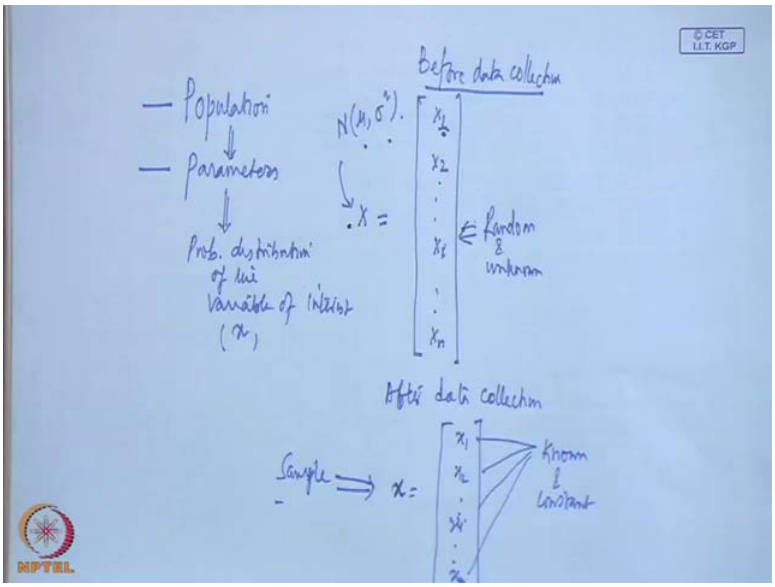
But, here this is a variable one you variable when something is variable ever for the student's point of view, some students are very erratic variable very, very difficult. But, some student may be because of some reason very regular, but suddenly or basically they are coming late by some few minutes. It is always we can, but they are still coming you can motivate them, so this is the physical meaning of the distribution.

(Refer Slide Time: 45:58)



Now, we will come to the next important concept is called sample and statistics.

(Refer Slide Time: 46:08)



So, we have seen so far population and parameters and please remember the random variable is there everywhere. So, population and parameters when we talk about population we definitely talk about parameters, we talk about particularly this next the probability distribution also. The probability distribution of the variables of interest, variable of interest that is x variable of interest x, now see you are planning to collect data what you have thought of that, I know my variable.

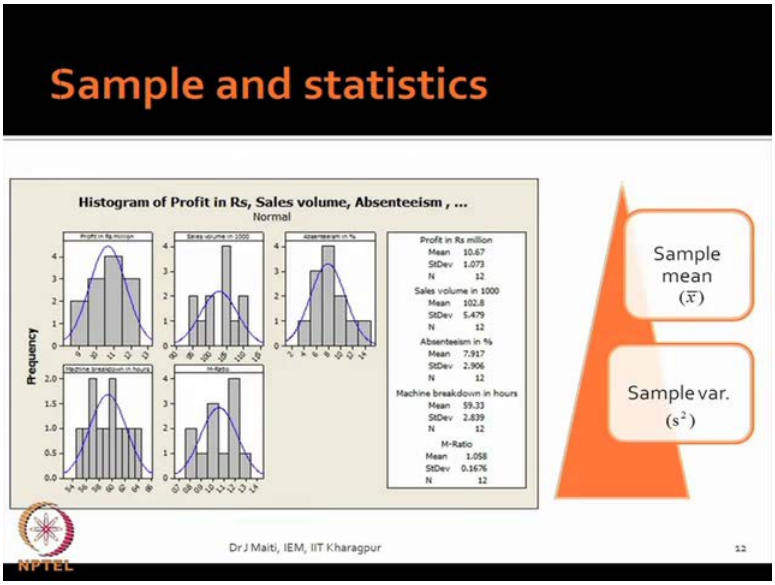
That variable is x, I want to collect data, how many data you want to collect you want to collect N data points what can you say about each of the observation what can you expect, getting me, so if x is normally distributed. Suppose x is normally distributed with mu and sigma square that means x 1 also normally distributed with mean and sigma square, x 2 also normally distributes with mu and sigma square because they are coming from the same population. There is no guarantee that when you if you go, you will observe some value of x somebody else will go he will observe some different value of x even though the observation is the first observation for you also it is first for him also it is.

First keep in mind this one this is very important you have not collected data that is before data collection you are planning that you will be collecting data N data. So, x 1 to 1 first you observe x 1 like x 2 like x n, now what is the issue here this all these values are random as if they are basically random and unknown. Now, you thought that you

have collected data, after data collection what will happen, so you will collect data I am denoting in terms of small x. Let it be this small x, so x 1, x 2, x i, x n what are these values known values realized, every value is realized known and constant this is in the population domain.

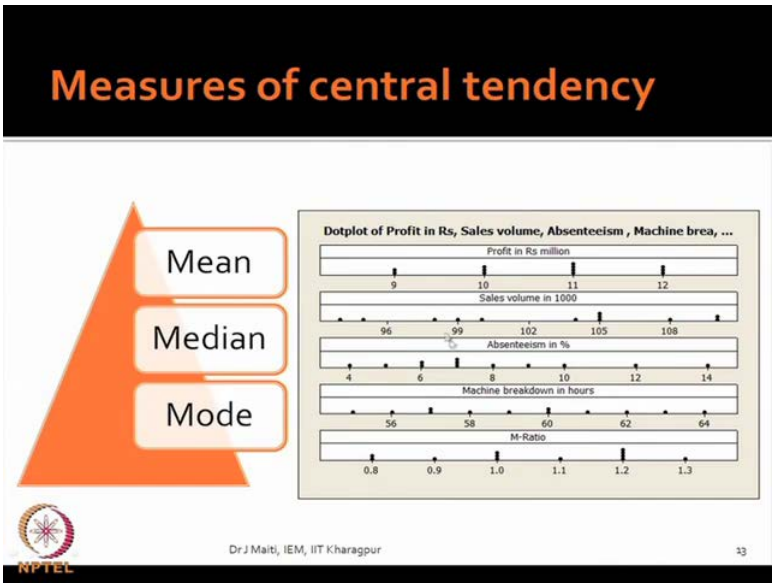
This is now sample of size N you see this slide, here before data collection you have planned to collect N data points and these are unknown and random. When you collect data after collection they are already known, fixed values that is the big difference and another important concept you keep in mind that all the observations each of the observation will follow the same probability distribution. We meaning that it is normal distribution with mu and sigma square as mean and variance x 1 has also normal distribution with mean mu sigma square as variance. Once data you have collected forget about all distribution there is fixed value, no randomness in the data it is already collected.

(Refer Slide Time: 50:09)



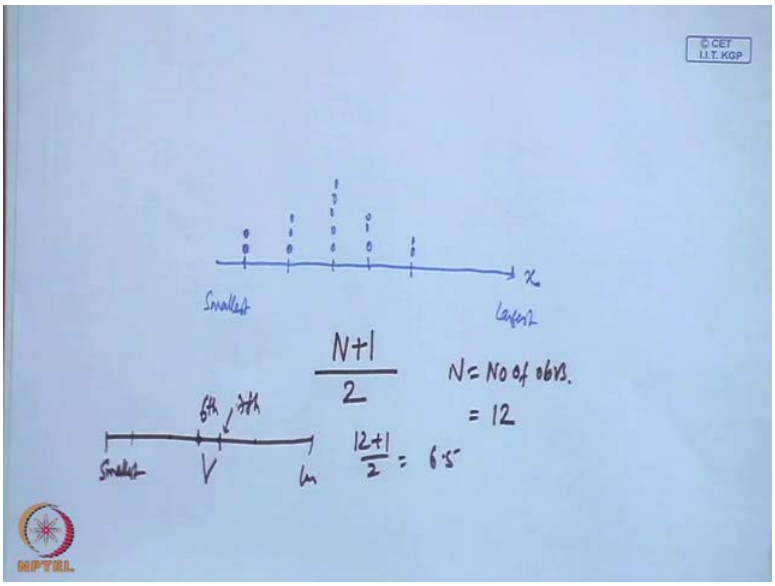
This one can example last example that is small company I can we can give a name to the company, I have given city cam. Later on I will use city cam, the company name is city cam, so these are the profit sales volume absenteeism all those things what is this. Basically, we are characterizing the that city cam process in totality in terms of these variables and your sample mean and sample variance, these are the statistics with respect to population mean and population variance, correct.

(Refer Slide Time: 51:02)



Do you know what is this is dot plot, dot plot is something like this dot plot is something like this.

(Refer Slide Time: 51:17)



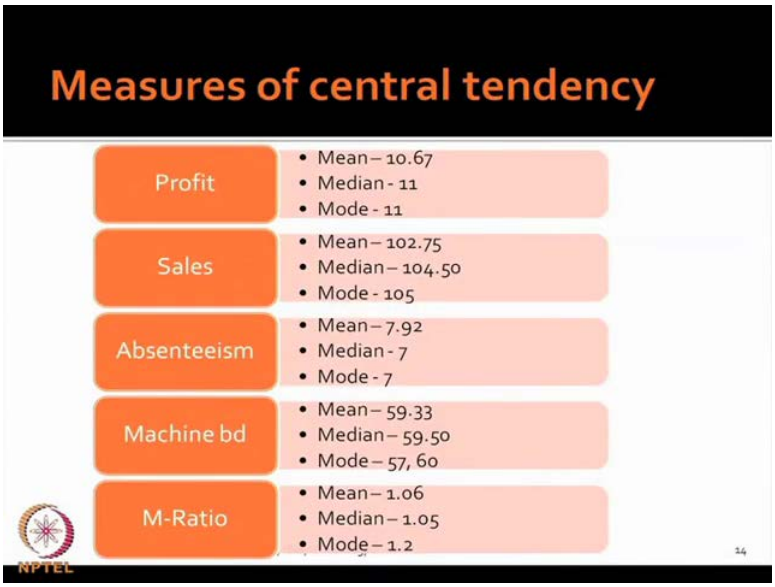
Your variable is x you arrange from the smallest to the largest, now suppose this is one value, this is second value, and this is third value like different values are there. Suppose this value, there is two you have two observations there let it be three observations, here let it be five observations, this type of plotting you are doing here. So, again this suppose

this again two values this is, let it be two values like this is known as dot plot, dot plot again it is similar to histogram plot.

Now, here you are able to count the number of observations against each of the values of the x, so it will help you to find out the mode suppose if I say for this example profit in rupees million you say 9 million rupees case. It is two observations for 10 it is 3, for 11 it is 4, for 12 it is 3 that means the mode of the data points for profit is 11 mean you have already seen median is the middle value. How do you compute median, for computation of median you find out the position N plus 1 by 2 where N is the number of observations.

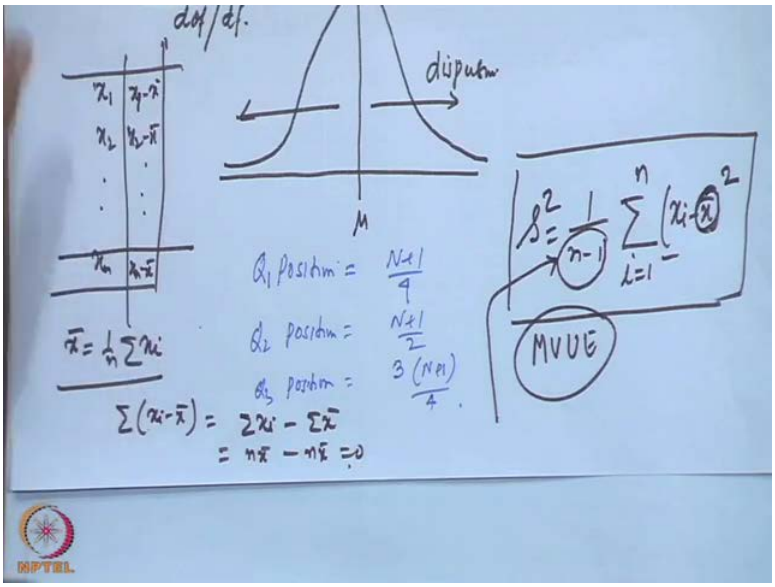
In this particular example, there N equal to 12 because 12 months data, so 12 plus 1 by 2 that means 6.5. So, 6.5 means when you arrange your data from smallest to the largest you just find out the position sixth and seventh position. Suppose this is a sixth position, this is your seventh position you take the average of these two value sixth position value and seventh position value, so and what is the mode is the value of x which there are maximum occurrences.

(Refer Slide Time: 53:30)



This is the calculation for that data and by using excel sheet, you can very easily calculate this thing, now measure of dispersion measure of dispersion is this.

(Refer Slide Time: 53:45)



What we have seen that if the data follows suppose normal distribution, then this one is mu and this side how much it is going. This side that is the dispersion there are several ways to measure dispersion, one is range that is minima maximum minus minimum value another. One is the inter quartile range which is the third quartile range minus first quartile, first quartile is basically N plus 1 by 4 and your quartile 1, Q 1 position is N plus 1 by 4 then Q 3, Q 2 position is the median which is N plus 1 by 2. Q 3 position is that is third quartile which is 3 into N plus 1 by 4, so all those position values you have to find out and then appropriately you have to manipulate the data.

So, if there are two values where N is coming in the middle, then you take the average. If it is coming, not middle may be right hand more than the middle seventh 0.75 position, suppose 3.75 position, so accordingly 0.75 that weight age to be given for that data. These are all very simple things you will be able to find out, these are little equally important and you require to know also these things. You know the variance also yes or no, how to compute variance statistical sense s square is 1 by n minus 1 sum total of i equal to 1 to n then x i minus x bar square.

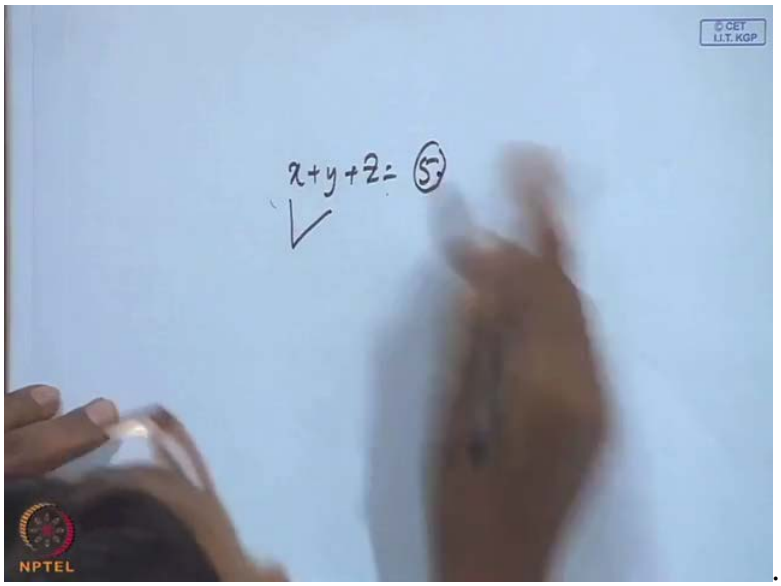
This is the variability measure why this n minus 1 minimum variance unbiased estimator any other explanation, now later on we will be discussing very much very frequently the degrees of freedom, getting me. Degrees of freedom we will be using d o f or d f, now

see in this case this n minus 1 is coming because of degrees of freedom because you have n data points x 1 to x n.

When you are computing this variance, you require what you require, you require x bar to be computed, so as x bar is computed with this formulation. So, what has happened ultimately, here when you find out that when x i minus x bar that is x 1 minus x bar, x 2 minus x bar like this for the last one you do not require to compute it is automatically computed.

So, I will write here suppose x n minus x bar what I mean that suppose if I write like this sum of x i minus x bar what will be the value x I bar. So, now what is mean value, so that means summation of x i minus summation of x bar this is n x bar minus n x bar. So, this is 0, so what will happen ultimately you are not getting n x minus x minus x bar value one value is 1, very simple other way if you say.

(Refer Slide Time: 58:21)



Suppose I have given you one equation x plus y plus z equal to 5, what is the degree of freedom. Here, you see if you change x and y, z cannot be changed further it is fixed even though the three values are there. You have two degrees of freedom because I have made it is made at 5 and in this case also the same thing is happening that is the sum of all this will be 0. So, you require how many data points you have in s minus 1 square equation n minus 1 that is the another explanation of why n minus one will be divided by 1 while computing this.

(Refer Slide Time: 59:03)

Now, I will finish this lecture, see we have we have told you that normal distribution is very important one and later on for this subject multivariate normal distribution. So, we must know that who is the father of normal distribution and you see that Abraham De Moivre, French born English mathematician. He basically has given the general form of this normal distribution, what form you will see that one by root over 2 pi sigma square e to the power minus half x minus mu by sigma to the power square.

Now, he is considered the father of normal distribution, but only equation will not become sufficient later what happened that Gauss, he is another famous mathematician and statistician what has he has given the properties. So, all statistical properties of normal distribution is identified tested by Carl Friedrich Gauss, he is specifically that German mathematician, they are not mathematician.

If you see that thing this is very interesting, it is not knowledge, but it is act of learning not possession, but the act of getting there which grants the greatest enjoyment. So, suppose you are doing PHD, so long you are not getting PHD you are thinking once I get PHD, I will be very happy, but it is not true once you get within 2, 3 days you will find out that you are the same person. But, learning going there, the act of going there that is what is very, very important and this famous people they have quoted and we must obey to their all suggestions.

Thank you very much, next class I will tell you sampling distribution.