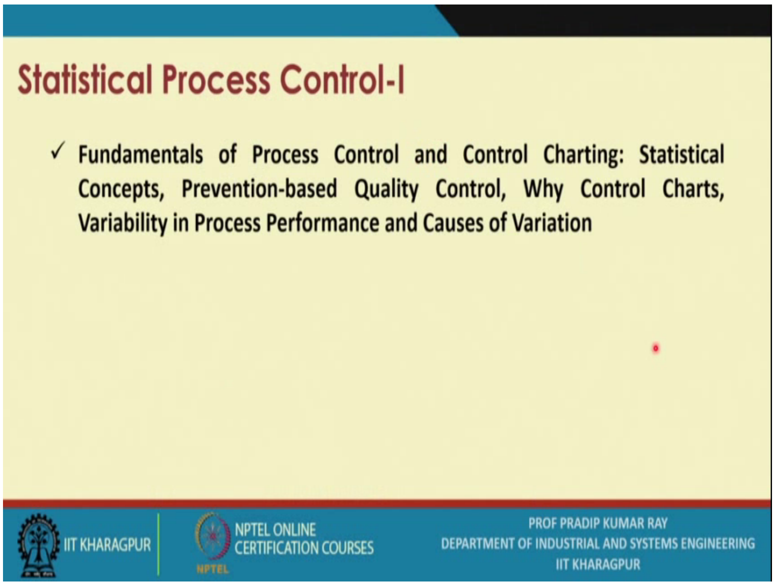
Quality Design and Control

Prof. Pradip Kumar Ray
Department of Industrial and Systems Engineering
Indian Institute of Technology, Kharagpur

Lecture – 17
Statistical Process Control-I (Contd.)

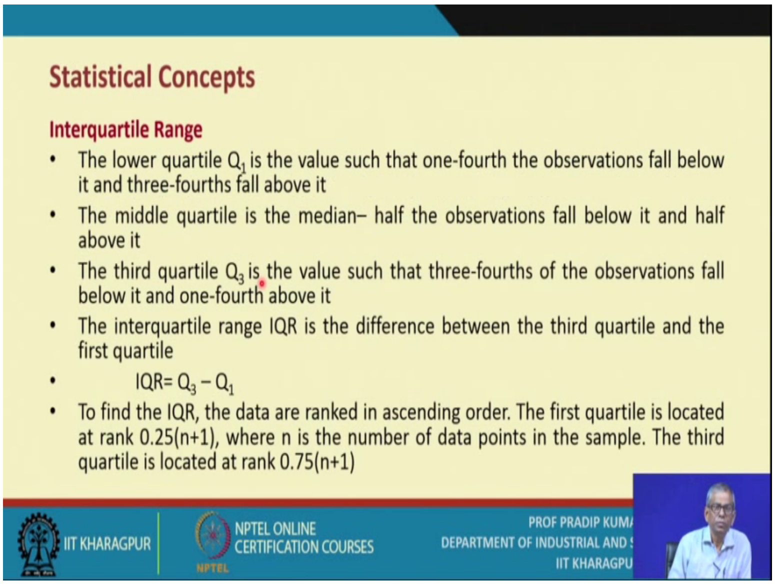
So, related to statistical process control we have several topics to discuss.

(Refer Slide Time: 00:21)



The statistical concepts knowledge in statistical concepts is a necessity and while we discuss this topic that is the fundamentals of process control and a process design and control. So, let me first talk about the statistical concepts. There are certain other concepts you must have who first have an understanding before we go on discussing the control charting and all other the tools and techniques in statistical process control.

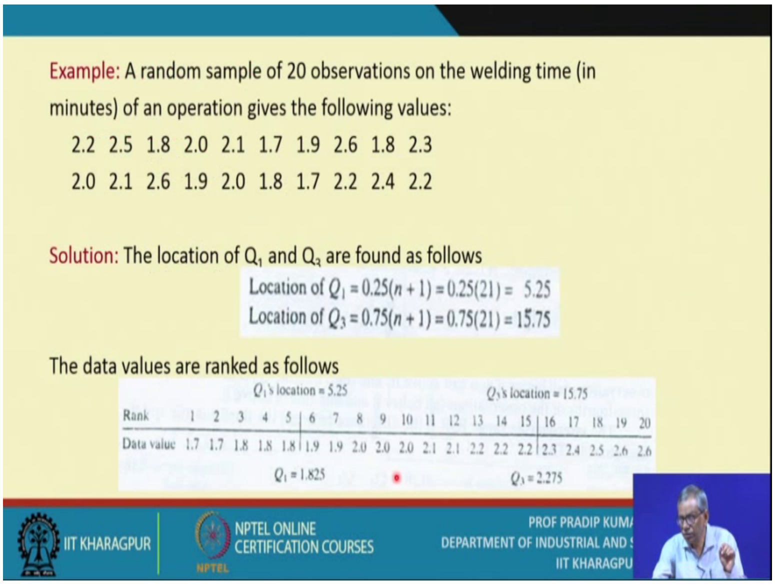
(Refer Slide Time: 01:03)



The yesterday or in the last class we mentioned about the measures of dispersion and we discussed already the first two measures that is the range and the variance of the standard deviation. The third one that is the interquartile range is also is used in many cases. So, the interquartile range is essentially we refer to the lower quartile that is a Q 1 and how do you define a Q 1 is the value such that one-fourth of the observations fall below 8 and three-fourth fall above it. So, that is the definition of Q 1. The middle quartile is the median half the observations fall below it and half above it.

Similarly, we have the third quartile that is the value and such that the three-fourth of the observations fall below it and one four above it I think it is clear to you. So, the interquartile range is given by the IQR is Q 3 minus Q 1. So, this is the traditional practice. To find the IQR what you have to do the data are to be ranked in ascending order that is the first step. The first quartile is located at the rank of 0.25 into n plus 1 where n is the number of data points in the sample say 100. So, 100 is suppose the sample size. So, the third quartile we located at rank 0.75 into n plus 1.

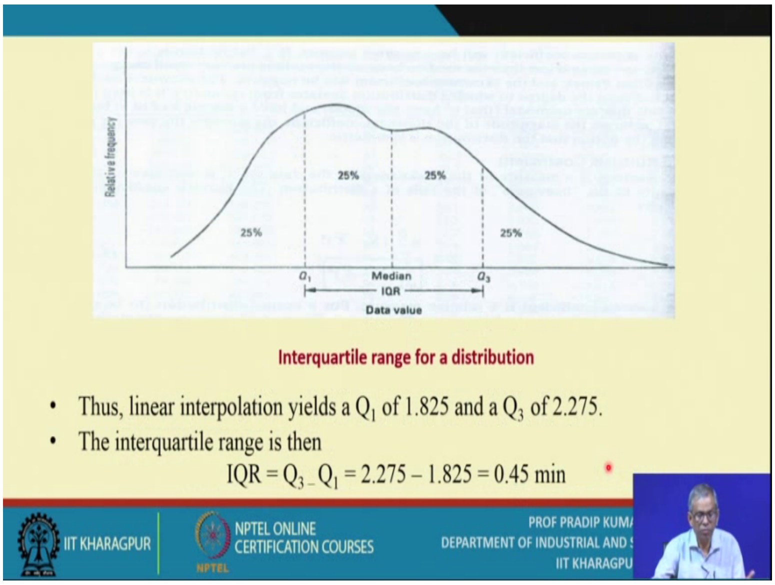
(Refer Slide Time: 02:57)



So, what is the first quartile? That means, the 0.25 n plus 1 and what is the third quartile that is 0.75 n plus 1. So, here in as example, you have trying to observations these are the values like 2.2, 2.5, 1.8 and so on. So, how do you determine the location of Q 1 and Q 3 that is straight we are applying the formula that is 0.25 into 21, 21 is n plus 1 twenty plus 1 that is 5.25 and similarly location of Q 3 is 0.75 into 21 that is 15.75. So, how do you do? That means, you erase the data in the ascending order is it; that means, the first data point is 1.7 and the last one that is 20th one that is 2.6.

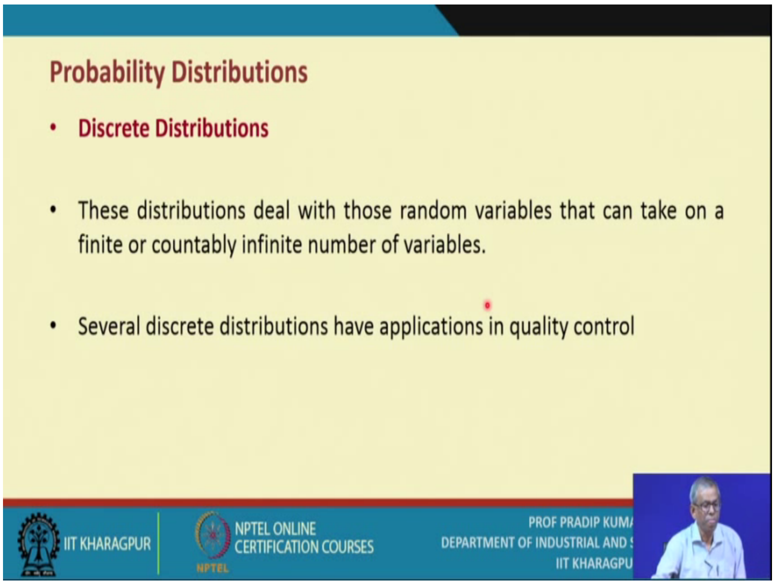
So, at this location that is 1.825 you have the value; that means, the rank is 5.25; that means, this is the rank there is a fifth rank between fifth rank and the sixth rank. So, what is the quartile one that is 1.825 it is clear. And similarly as this is the location of Q 3 is the rank is 15.75 that means, almost you know it is near to 16. So, the corresponding value you can compute that is 2.275. So, this simple rule we follow.

(Refer Slide Time: 04:26)



Obviously, you know now you can determine the interquartile range right it is a measure of dispersion. So, how do you do that? That means, Q 1 1.8 to 5 we have already measured and Q 3 is 2.275. So, interquartile range is the difference that is 0.45 millimeter. So, this is the relative frequency and this is versus the data value. So, 25 percent is of the points the data points are less than say Q 1 and 75 percent of the data points are less than Q 3, is it ok. So, this is what are you (Refer Time: 05:05) range.

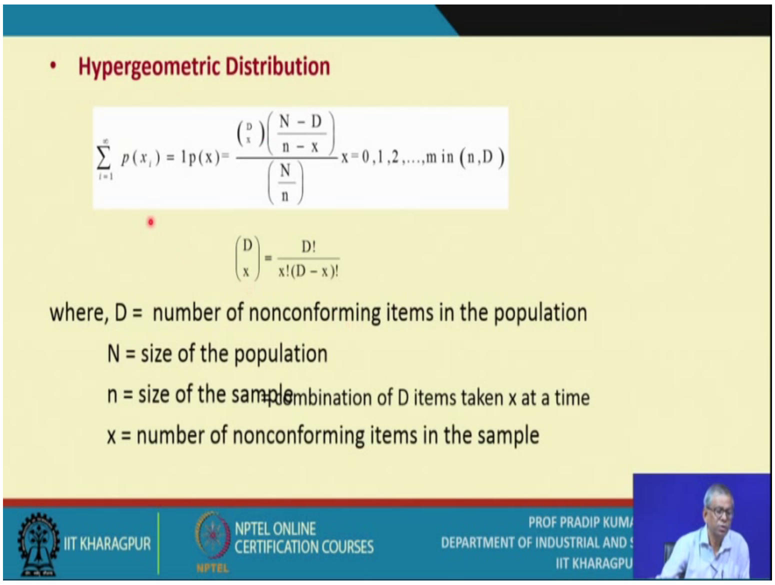
(Refer Slide Time: 05:06)



Now, we are referring to the probability distributions. You know there are you are supposed to use various kinds of tools and techniques and these tools and techniques many such tools and techniques being used in quality design and control and you know based on certain distributional assumption right. We have already mentioned that you need to deal with the random variable either discrete type or continuous type in quality control and improvement purposes.

So, what do you try to do? That means, against a particular say random variable which you are dealing with in a given case you must be higher of what kind of you know the probability distributions it may have. So, now, this just I will discuss briefly some of the important probability distributions and their characteristics which are commonly used in a quality control and design. So, this probability distributions are grouped under two categories the first one is a discrete distribution; that means, the corresponding you know that say random variable is of discrete type. So, these distributions deal with those random variables that can take on a finite or countably infinite number of variables essentially they can assume the integer values. Several discrete distributions have applications in quality control. So, I will just name them, like the first one is the hypergeometric distribution ok.

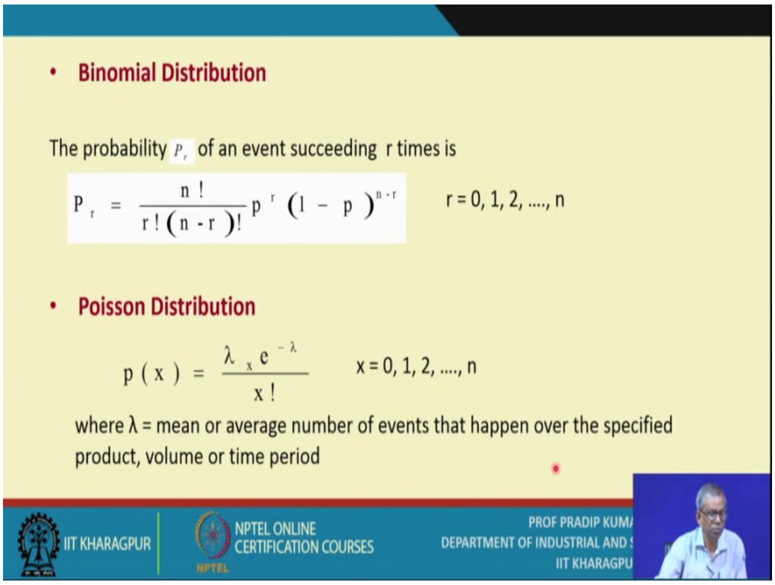
(Refer Slide Time: 06:46)



So, the hypergeometric distribution you know there is lot of the applications and particularly you know by an application related to say acceptance sampling when you deal with isolated single lot. So, the distributional assumption is hypergeometric.

So, hypermetric geometric this is the probability mass function, this is the probability mass function and you have so, this is basically you know sigma or the px is given by D c x N minus D, n minus X and this is the equations you use an X could be; that means, the number of is number like say when you deal with say the number of nonconforming units in a sample. So, it may assume several any one of these integer values and in n and D; that means, this is n not m this is n n into n comma D. So, maximum could be n and the in the X values and n is within or either it could be n or it could be D between D and n. So, this is the formula use and, the number of nonconforming items in the populations the size of the population capital n is the size of the population small n is the size of the sample combination of D items taken at taken at a time and number of nonconforming items in the sample. So, this is the hyper geometric distribution.

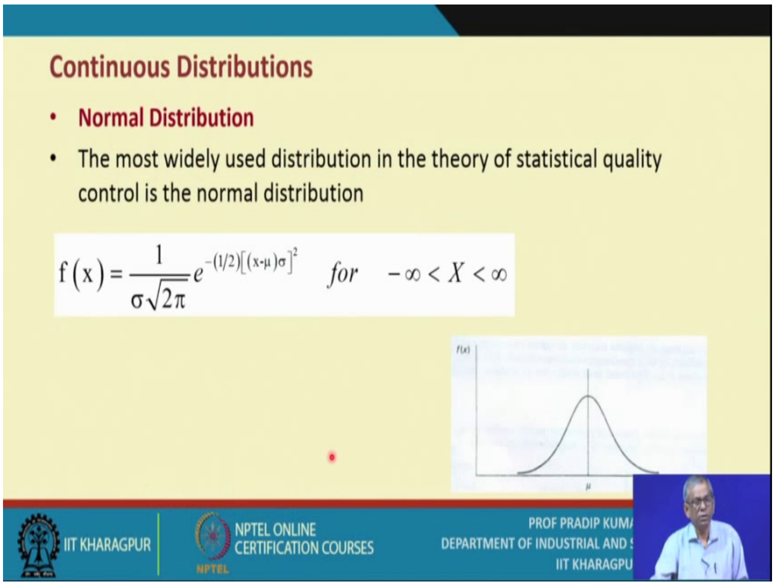
(Refer Slide Time: 08:38)



So obviously, what I am asking you to do like say for any say the distribution you come across there are there are parameters. So, what will be the expression with these parameters, particularly you know mean and standard deviation. So, that always you need to compute. Like say given you know the probability must functions for the hypergeometric distribution. So, what will be the expression for, what will be the expression for you know the mean and the standard deviation? So, that always you can compute.

Similarly we come across binomial distribution and binomial distribution this is n c x or n c r, p to the power r into 1 minus p to the power n power r. So, the value of r could be 0 1 2 up to n, n is a sample size then in many cases we assume that that the random variable is Poisson distributed. So, this is say the probability mass function of our Poisson distribution lambda to the power e to the power minus lambda, lambda to power X into factorial X this should be lambda to the power X by factorial X is it ok, where lambda is equals to the mean or average number of events that happen over the specified product volume or the time period.

(Refer Slide Time: 10:14)

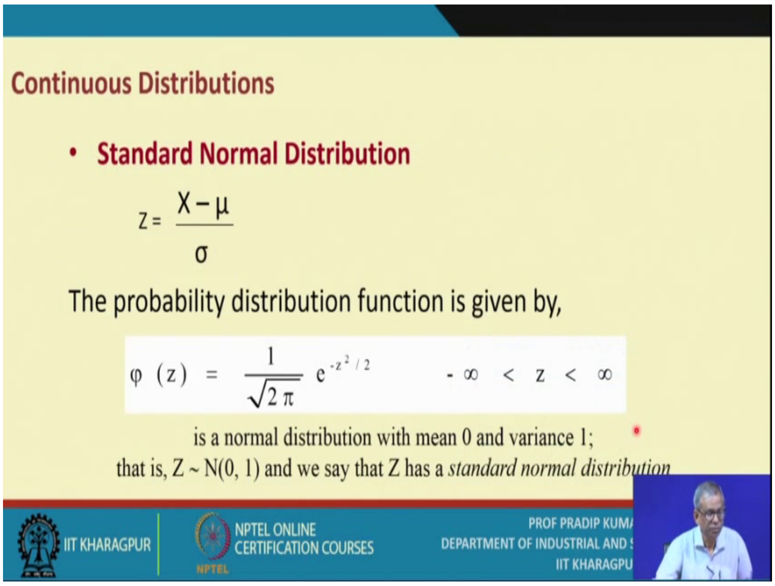


Now, these are the 3 distributions we come across frequently one is the hyper metric hypergeometric distribution, the second one is the binomial distribution and the third one is the Poisson distribution. In many cases what we do I make I make on the certain conditions, you may go for Poisson approximation to binomial distribution. So, when we refer to these cases will have these assumptions or these approximations. So, later on will take off numerical problems and where in many cases you go for Poisson approximation for binomial distributions.

As far as continuous distributions are concerned we have many, but here at this point in time I will be referring to the normal distributions. The most widely used distributions in the theory of statistical quality control is the normal distribution. So, this is the probability density functions you are all aware of. So, so this is one and there these there are two parameters we have, one is the sigma and the second one is the mu. So, these are the mu is the mean and sigma is the standard deviation.

So, this is the typical you know of the distribution f X versus X and this is a symmetric distribution and with mean mu and the standard deviation sigma, is it ok.

(Refer Slide Time: 11:46)

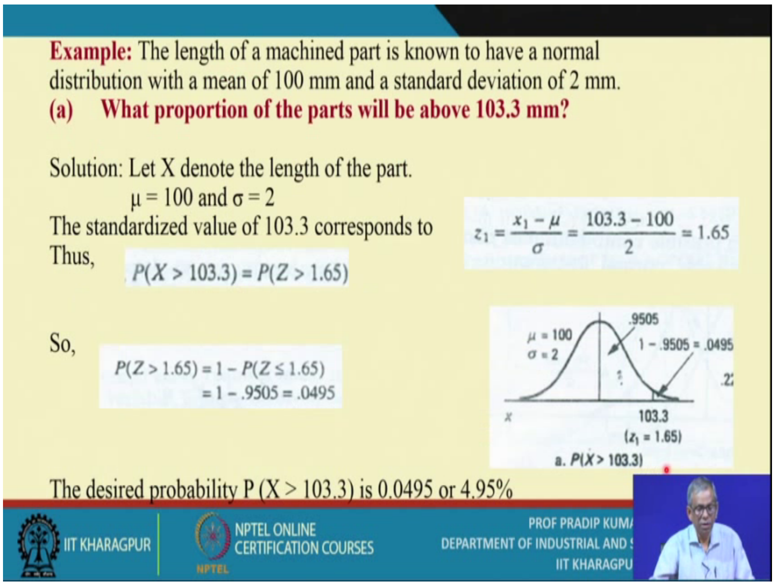


Now, what we try to do? That means, we standardize the normal distributions and we refer to the standard normal distributions where X is transformed with Z and Z is essentially you know is nothing, but the difference of value of a particular say, particular value of the random variable from the mean and these difference is expressed in the standard deviation unit.

So, the difference is X minus mu. So, mu is one of the parameters. So, X minus mu these difference expressed in the standard deviation is to this way you know the Z has been defined and so obviously, the probability distribution function, probability density function it is not distribution function density function is phi z is 1 minus e to the power minus z by 2, is it ok, where Z may live anywhere between minus infinity to plus infinity. So, here; obviously, with this transformation, this normal distribution we will have a mean of 0 and variance of 1, is it ok. So, normally this is specified as Z you know

N 0 1 and we say that Z has a standard normal distribution, is it ok. So, X is transformed into Z.

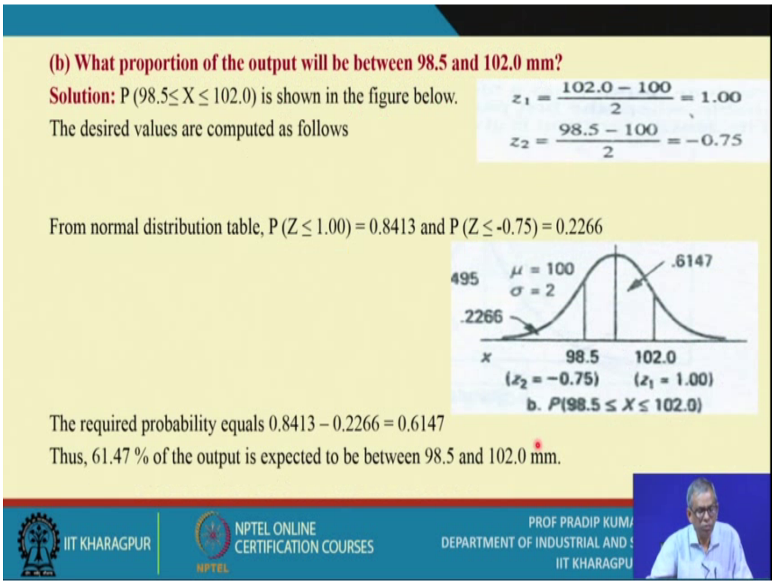
(Refer Slide Time: 13:23)



This is an example just as the simple examples I have taken the length of a machine part is known to have a normal distribution with mean of 100 millimeter and the standard deviation of 2 mm is it. So, what proportion of the parts will be above 103.3 mm?

So, what do you need to do? Let X denote the length of the part. So, mu 100 sigma is equals to 2. So, the standardized value of 103.3 corresponding to say X 1 103.3 minus 100 divided by 2 because sigma is 2 and so it is 1.65. So, what you refer to that means, area under the curve beyond 1.65. So, this is the probability the Z is greater than 1.65. So, here it is 1.65. So, if you refer to the standard normal table. So, you will find the area under the curve beyond Z equals to 1.65 is 0.0495; that means, 4.95 percent is the areas. So, this way we calculate we refer to the standard normal distribution.

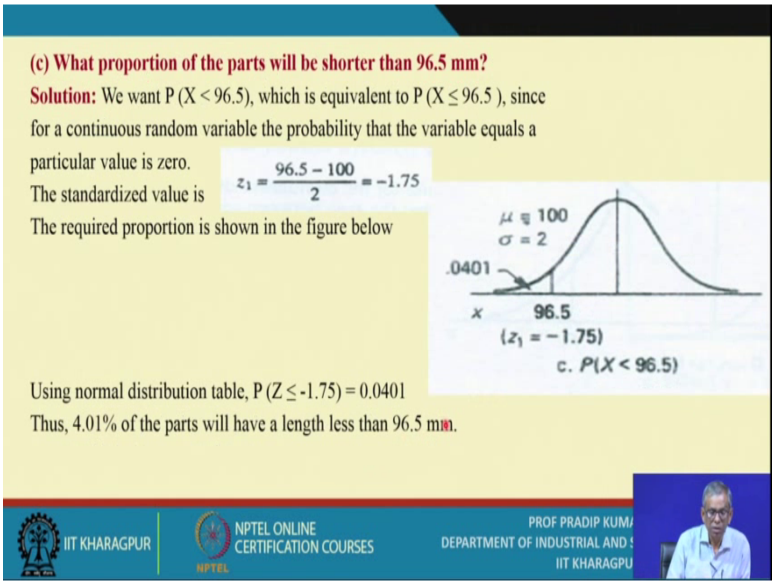
(Refer Slide Time: 14:37)



So, what proportion of the output will be between 98.5 and 102, what you try to do against 98.5? You calculate corresponding Z 1 is it when the mean and the standard deviation is known. So, this is 1.00 and similarly against 98.5 what is Z 2? That means, 98.5 minus 100 divided by 2; that means, minus 0.75.

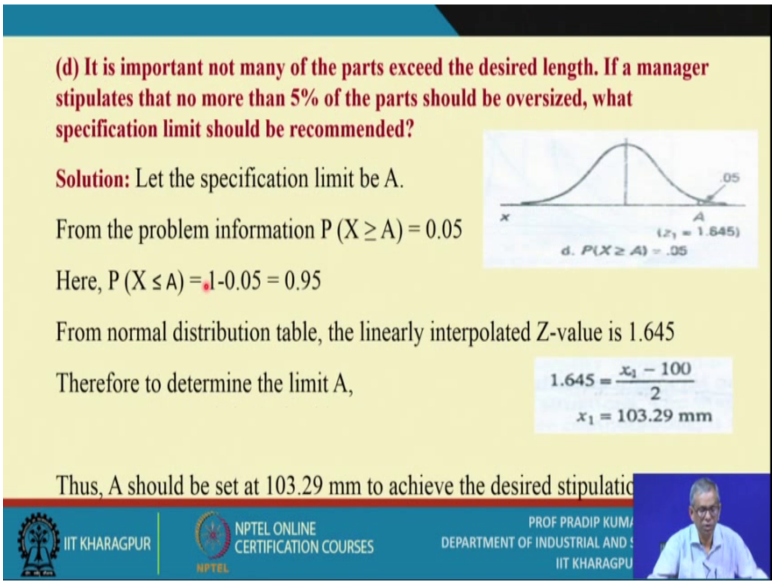
So, now what you try to do? That means, area you know area under the curve between say Z equals to 1.0 and Z equals to minus 0.75. So, minus 0.75 is somewhere here and 1.0 Z equals to 1.0 is here, is it ok. So, the area is this one is it that you have to compute. So, the required you were referred to the standard normal distribution the corresponding table and the required probability equals this 0.6147, is it ok. So, this value you get by referring to the standard normal distribution. Similarly this, this area also you get and when you subtract these 2 areas from 1. So, you get the value of 0.6147, it is clear. So, you know how to use this term like standard normal distribution table.

(Refer Slide Time: 16:05)



What proportion of the parts will be shorter than 96.5? So that means, against 96.5 again you calculate its Z value that is minus 1.75 with reference to 100 that is the mean and the standard deviation is 2. So, this area again you refer to the standard normal distribution table. So, against a value of Z of minus 1.75, area under the curve beyond this value of 2 minus infinity that is 0.0401. So, 0401 that means, 4.01 percentage of the points they will be shorter than 96.5 all the parts ok. So, these are the simple examples.

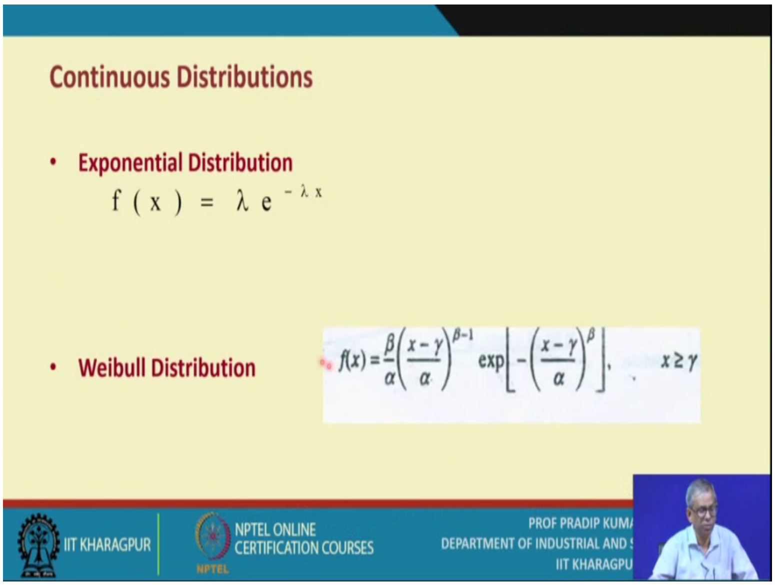
(Refer Slide Time: 16:55)



Similarly it is important not many of the parts exceeds the desired length; obviously. If a manager stipulates that no more than 5 percent of the part should be oversized what specification limit should be recommended. So, here the problem is that X must should be greater than equals to the probability is 0.05. So, this is 0.95 less than equals to A. So, corresponding is a Z value that is 1.645. So, that you get and you need to determine X 1. So, this is known, this is known, this is also given and. So, what you need to do? The what will be the corresponding value of X 1.

So, therefore, this way we say that this way you determine the value of X 1 and we say that a should be set at 103.29 to achieve the desired stipulation is it. So, this is another example.

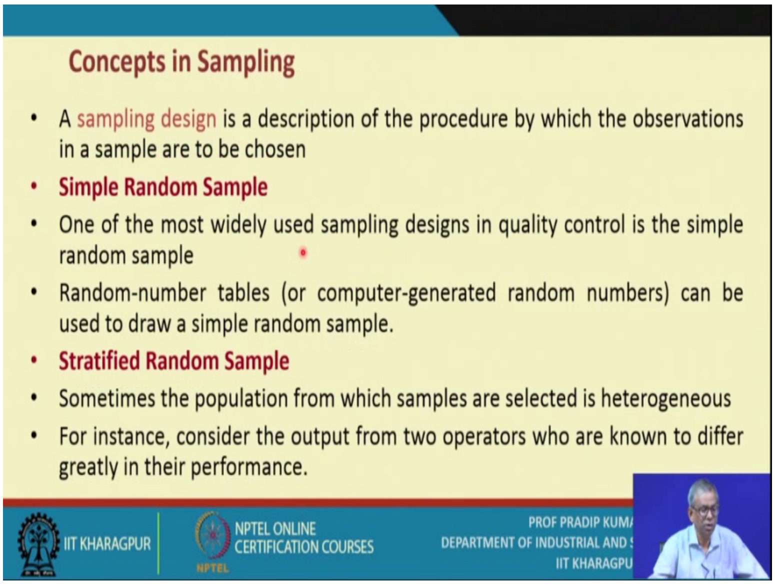
(Refer Slide Time: 17:50)



Now, under continuous distributions say exponential distribution also is used and we will distribution particularly when you try to model the reliability function of a component; and what is reliability? Reliability as you know this is one of the dimensions of the quality.

So, this is the probability of the mass function probability, probability density function for the Weibull distribution and this is exponential distribution is a particular case of Weibull distribution. So, Weibull distribution is having 3 parameters that is you know the gamma, gamma is referred to as the location parameter, beta is the shape parameter and alpha is scale parameters. So, these we come across.

(Refer Slide Time: 18:41)

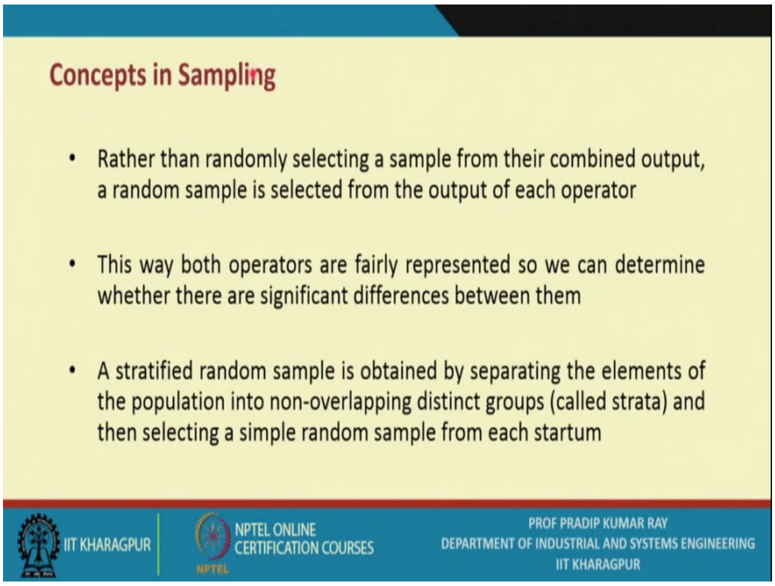


Now, we also must have an idea about the sampling because you know when you go for statistical process control through control charting. So, why do you construct the control chart what you need to do; that means, from the process of the from the population you have to create a sample; obviously, while you create a sample; that means, you must have a thorough idea about the sampling procedures.

So, just I will be referring to the concepts what is the basic concepts in sampling. So, a sampling design is the description of the procedure by which the observations in a sample are to be chosen it is very simple. So, what is a simple random sample? One of the most widely used sampling designs in quality control is the simple random sample. So, random number tables you refer to random number tables compute or computer generated random numbers can be used to draw a simple random sample. What is a stratified random sample sometimes the population from which the samples are selected is heterogeneous; that means, suppose there is in the population of for the persons. So, there could be you know the male population there could be female population is just one example.

And you have to, you know you have to collect the sample from the population. So, and it must be representative. So, what you try to do? That means, you collect one sample from say the male population and another sample from the female population. So, this is just a simple example. So, this is referred to as a stratified random sample. So, for instance consider the output from two operators who are known to differ greatly in their performance same example.

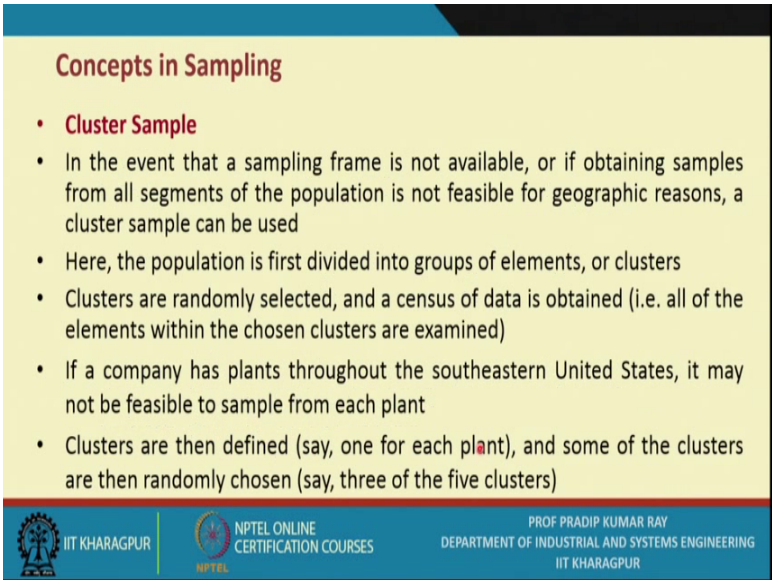
(Refer Slide Time: 20:39)



So, you have to collect sample; that means, from one operator and similarly for the second operator. So, I will elaborate it later on the stratified sampling method is widely used in the field of you know the quality control. So, rather than randomly selecting a sample from their combined output, a random sample is selected from the output for the each operator this way both operators are fairly represented.

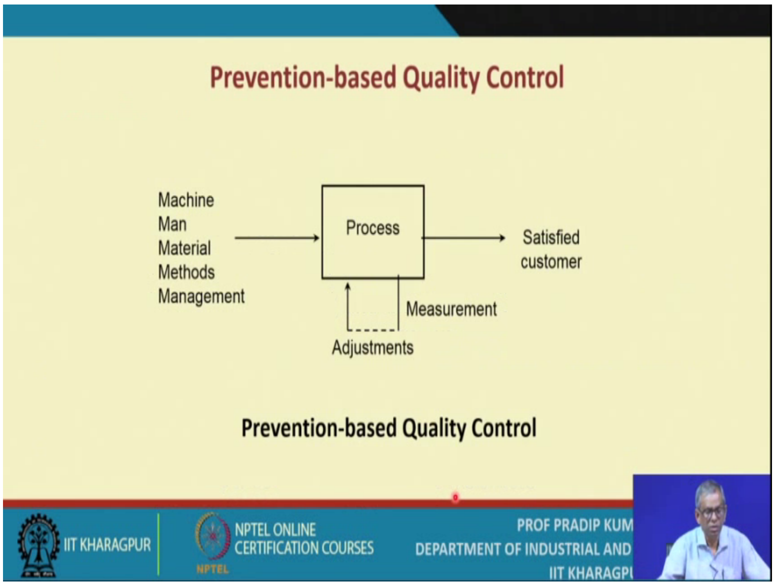
So, we can determine whether there are significant differences between them, is it ok. So, whatever the quality characteristics you are referring to. So, stratified random sample is obtained by separating the elements of the population into 2 into non overlapping distinct groups called strata is it ok, like one group is female another group is the male is it. So, the data sets for one operator and the data sets for the second operator. So, these are called different there are as the strata and then selecting a sample random sample from simple random sample from each stratum.

(Refer Slide Time: 21:56)



So, similarly we have a cluster sample the cluster is formed and many a time no you come across such situations the clusters are formed is it ok. So, what do you, many a time what you do the entire population is grouped into several clusters based on curtained the characteristics. So, if a plant has plants throughout the southern eastern United States you may not be feasible to sample from each plant is it ok. Since like say, cluster are then defined say one for each plant and some of the clusters are then randomly chosen say 3 of the 5 clusters, is it ok.

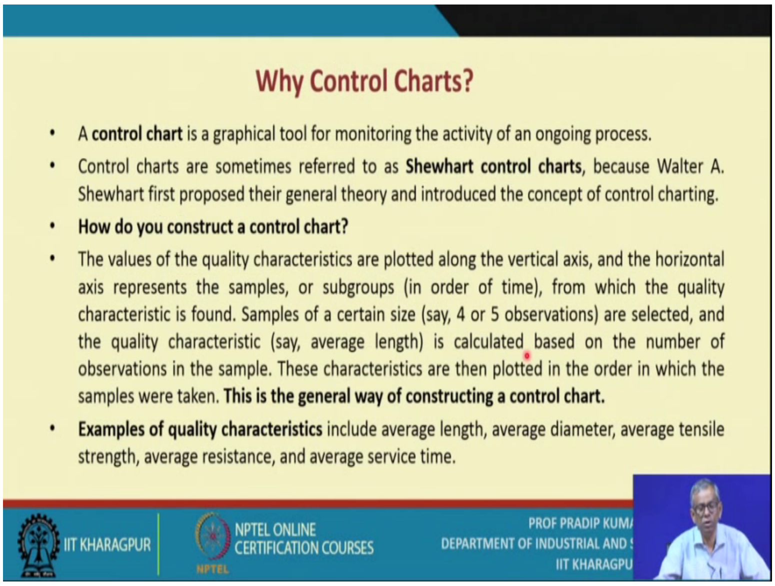
(Refer Slide Time: 22:38)



So, this is the concept of sampling. Now, before I refer to certain fundamentals you know of the statistical process control or the techniques. So, I am again referring to the prevention based quality control this diagram as we took as I have already told you that said that the first stage in the evolution of the quality system is inspection and the next one is the quality control.

So, whatever you know the statistical process control what we are talking about we assume that we are in a stage in a condition where prevention based quality control you know is a necessity and this control chart which you are going to use for statistical process control; obviously, you know the system should be the prevention based quality control. So, you know what is it; that means, you have several kinds of resources, you have the process and then as a worker you had as an operator the process what you need to do; that means, you need to go for measuring the quality characteristics and if adjustment you need to have in the measurements you must be able to do the adjustments is it. And ultimately the output should be you know the (Refer Time: 24:11) to you and so that you get a satisfied customer. So, this is in this prevention mode, prevention based quality control the control chart is to be used.

(Refer Slide Time: 24:19)



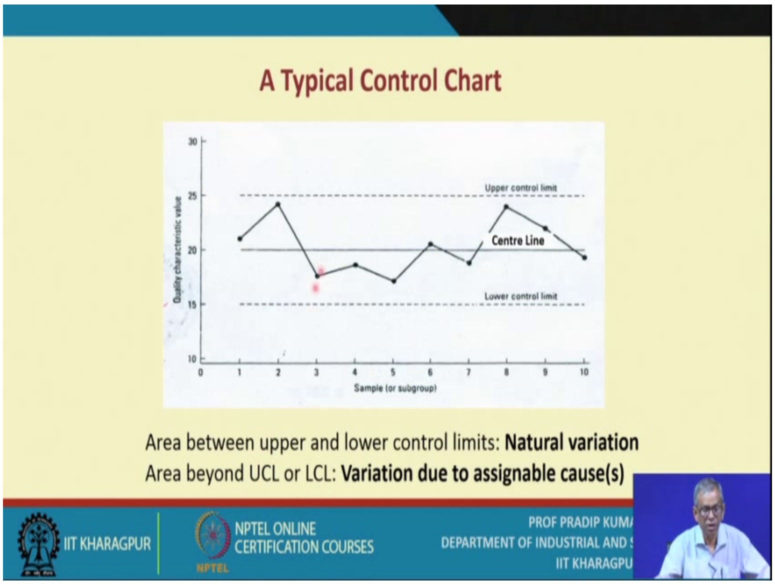
Now, I will just briefly tell you as that why control charts what is a control chart. So, what is a control chart? Control chart is a graphical tool for monitoring the activity of an ongoing process. So, it is an ongoing process; that means, through control by using control chart what you can do require definitely you can go for online real time control.

Control charts are sometimes referred to as the Shewhart control chart because Walter A Shewhart, he actually the first you know way back in 1920s he introduced the concept of control charting and the subsequently his concept of control charting was implemented in several well known organizations throughout the world.

Now, how do you construct a control chart? It is very simple, you have x axis, you have the y axis. Now, supposing for a given quality characteristic you need to you know the construct the control chart. So, what you try to do you need to collect the values of the quality characteristics from the process the values of the quality characteristics are plotted along the vertical axis; that means, the y axis and the horizontal axis represents the samples or the subgroups as I have already pointed out that the samples which you draw from the population or from the process there are certain rules there are certain norms later on will discuss this one. So, and these the samples you have to draw as the process is on; that means, the time order is to be mentioned it means that if you say it is the first sample; that means, the first sample is drawn first if you say is the second sample; that means, the second sample is drawn next like this. So, trying order is to be mentioned.

So, this you know the x axis represents the samples from which the quality characteristic is found; that means, instead of measuring or all the units in the population you just rise sample and within the sample you have you know a number of values and those values are to be collected and then you get its the average value. So, say the average represents a particular sample quality characteristic. So, these average is calculated based on the number of observations in the sample, that means the sample size and these characteristics are then plotted in order of time in which the samples are taken this is the general way of constructing the control chart. So, it is very very simple. So, examples of quality characteristics like average length, average diameter, average tensile strength, average resistance and so on and so forth.

(Refer Slide Time: 27:23)

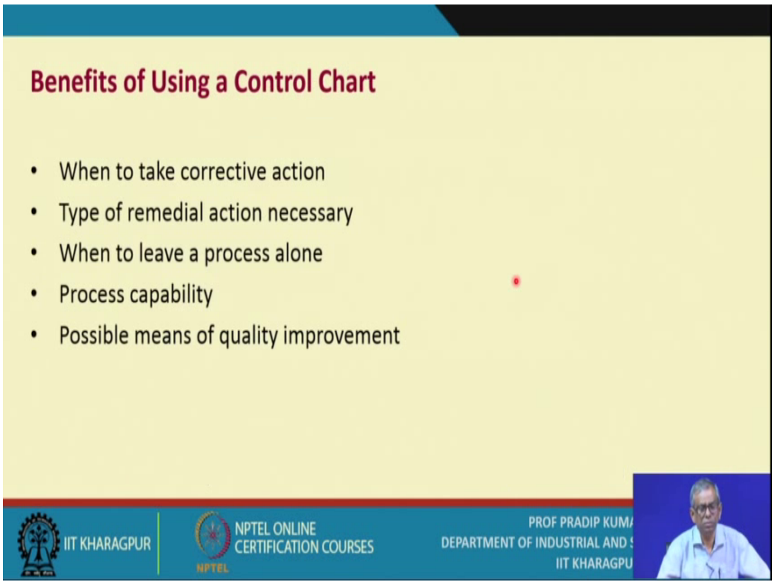


So, this is the typical control chart. So, this is actually the y axis, this is the x axis and in a typical control chart you have a central line and you have the upper control limit, you have the lower control limit is it and usually they are at equal distance from say the center line. So, on one side you have the upper control limit on the other side there is the same distance you place the lower control limits.

Area between the upper this point is to be noted. The area between the upper and the lower control limits these reflects actually the natural variation of the process whereas, the area beyond UCL; that means, this area or the area beyond LCL that means, this area this you know represents actually the variation due to assignable causes. That means, if a certain sample point falls over here you assume that there is a variation definitely, but this is within the natural variation. But if the point is somewhere here; that means, against the fourth sample; that means, you will find that there is a variation because your target is a centerline, so this much variation. Now this must variation is due to some assignable causes and something has gone wrong at that point in time while you collect data for the fourth sample.

So, this is referred to as assignable cause. And whenever you find that the point is the plotted says the fourth point is plotted say outside say he and all the points are plotted within the control zone. So, you may assume that the necessary condition for in control state has been achieved.

(Refer Slide Time: 29:18)



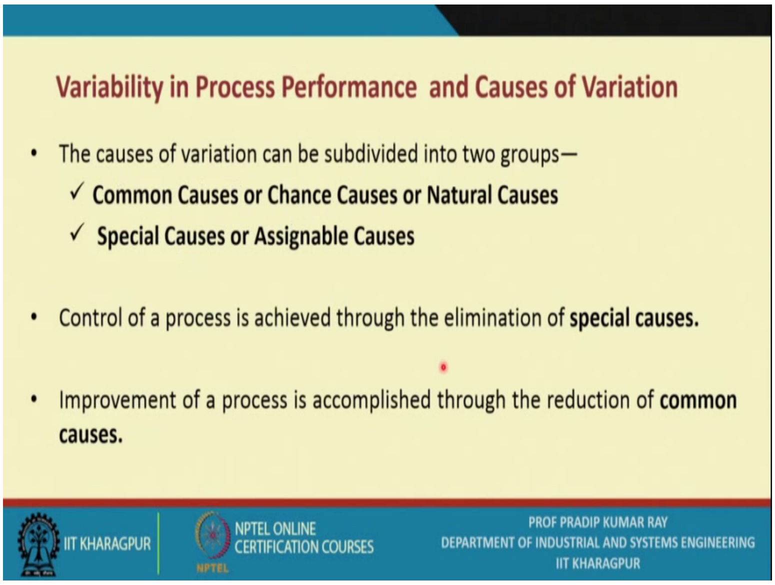
So, by using a control chart you know whether the process is in control or the process has gone out of control. So, if a point for the plots outside of the control zone we assume and we assume that the process has gone out of control; that means, out of control state. And if you find that the point plots within the control region; that means, the area specified by the upper control limit and lower control limit we assume that the process is in control or is in control state.

So, what are the benefits? When to take corrective action is it ok, as soon as you find that the process has gone out of control now you have to take some corrective actions so that the process is brought back to a state of control is it ok, from the out of control to in control state. Type of remedial action necessary; that means, when you start using a control chart against a process; that means, you are continuously using this control chart you know the behavior of the process; that means, ins and outs of the process you have a thorough knowledge and, so what kind of remedial action you need to take to bring the process back to you know the in control state that you will come to know.

When to leave a process alone; that means, if you find that that the process is a continuing in control state. So, in many cases you do not go for you know the controlling the process you leave the process alone. So, if you can create such a condition you assume that my purpose of using a control chart is fulfilled. You know will be will be discussing the process capability and when you start collect, we start collecting data through control charting you know later on these data you can use to the measure the capability of a process, is it ok. So, if you start using a control chart, so based on the information you collect, you can measure the capability of the process.

Possible means of quality improvement is it ok. So, always you go for it.

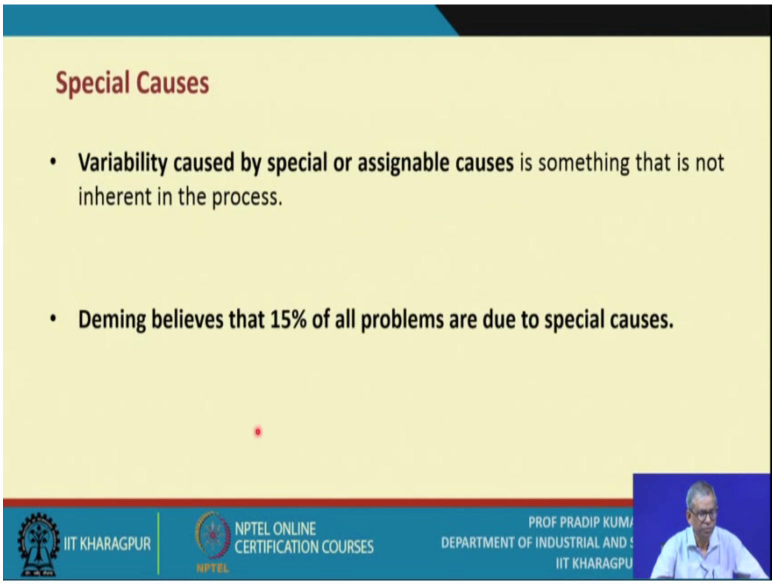
(Refer Slide Time: 31:47)



So, there are two types of the variations. So, and what we have mentioned that that the two types of causes of variation and what we have been stressing that any exercise and quality is essentially an exercise on controlling the variability. So, you must know that the causes of variation that there are two types of causes of variation, one is the common cause these also referred to as the chance cause even note down or the natural cause. Whereas, the second kind of the causes they refer to the special causes or assignable causes.

So, control of a process is achieved through the elimination of special causes. An improvement of a process is accomplished the reduction of the common causes. That means, common causes actually are a part of the system.

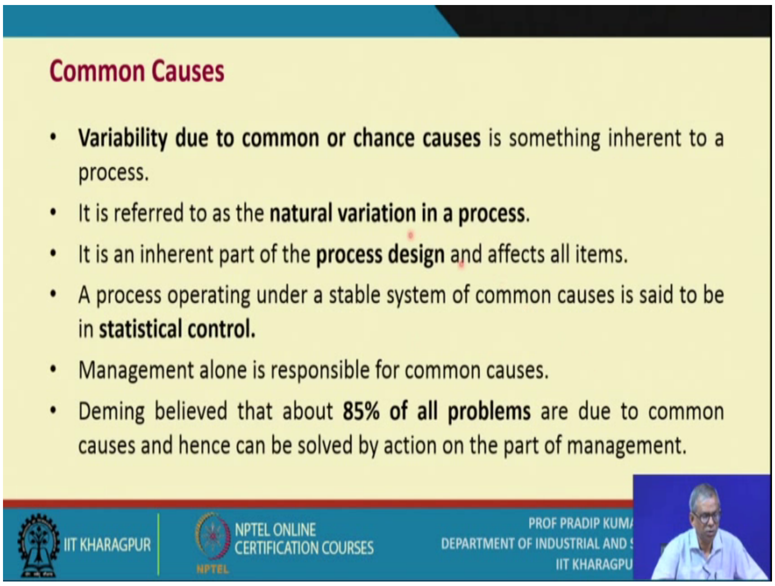
(Refer Slide Time: 32:44)



So, the variability caused by special or assignable causes is something that is not inherent in the process. That means, for the time being it has it has happened, but if you take some remedial action obviously, this cause will be will be eliminated and that that the process has become free of this particular cause.

Deming believes that fifteen percent of the all problems are due to special causes.

(Refer Slide Time: 33:11)



Now, I will just you will take 2 to 3 minutes of time to 2 more minutes to explain the common causes. The variability due to common or chance causes is something inherent to a process; that means, you are living with this. So, any systems we are talking about; that means, that the we try to reduce the variation, but there will be some variation. So, at any point in time you must know that what is this allowable variation and these variation is due to the common causes. So, this is the meaning.

And the process operating under a stable system on common causes is said to be in statistical control we will be using this term this is referred to as the statistical control or not. So, management alone is responsible for common cause that is a part of the systems as the management has created the system and Deming believed that about 85 percent of all problems are due to common causes and hence can be solved by action on the part of management. This is through his study he concludes.

So, here I conclude, but you know later on will discuss on the common causes of variation as well as the assignable causes of variation. So, if you can remove the assignable causes the process becomes in control state and if you can consider the common causes, conseque orders and you stripe you can eliminate some of the common causes we say that the process or the quality control system has improved.