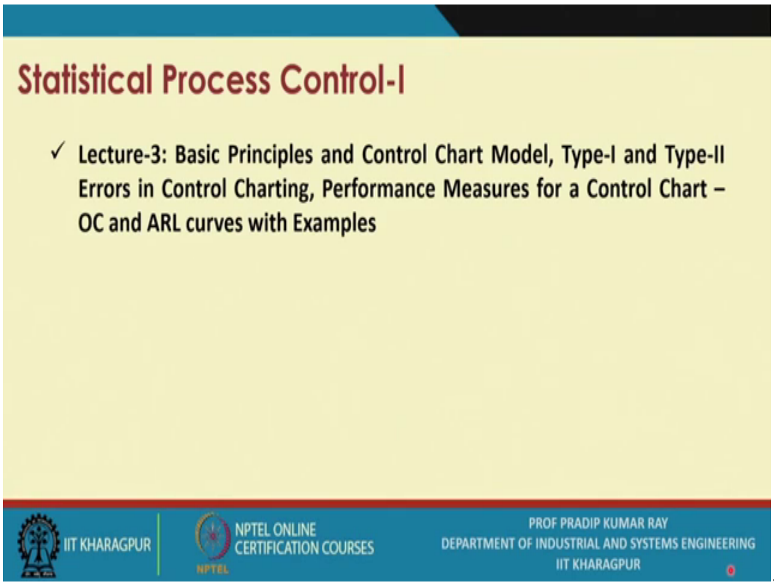
Quality Design and Control
Prof. Pradip Kumar Ray
Department of Industrial and Systems Engineering
Indian Institute of Technology, Kharagpur

Lecture - 18
Statistical Process Control-I (Contd.)

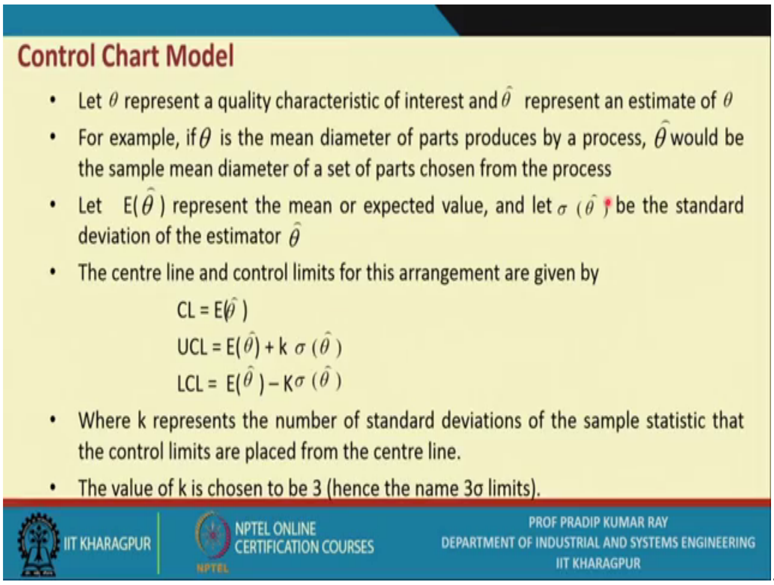
In this session on Statistical Process Control.

(Refer Slide Time: 00:19)



I will be discussing few other aspects related to control charting that basic principles under a control chart and what is a control chart model which is generalized control chart model. The type 1 and type 2 errors in control charting this is an important concept you must be aware of performance measures for a control chart and particularly 2 kinds of the control chart or say performance measures we will be using and we will be referring to one is operating characteristic curves how to draw an operating characteristics curve for a control chart and how to draw that average run length curves, these are the 2 performance measures we are going to discuss.

(Refer Slide Time: 01:20)



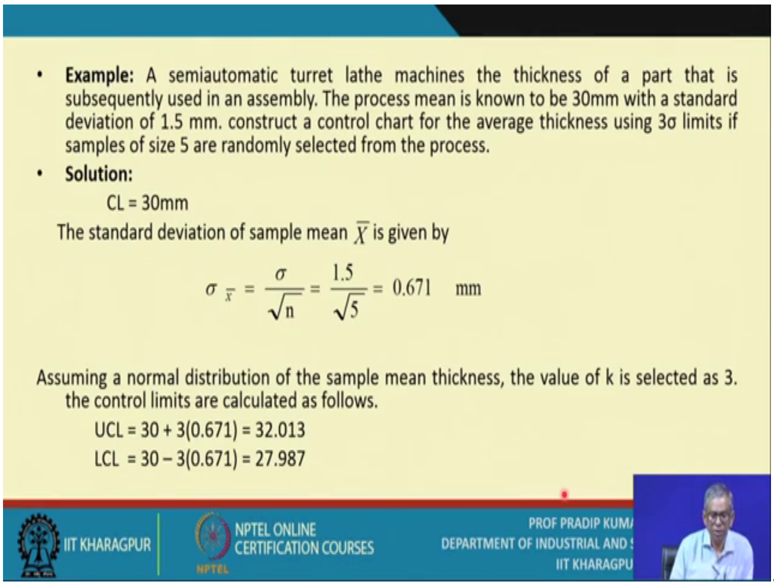
Now, let us the talk about a typical control chart model. So, let me just go through certain points like let say theta represent a quality characteristics of interest say length or diameter of a of a part and theta hat represents an estimate of theta is it; that means, you have to collect data you have to measured the quality characteristics. For example, if theta is the mean diameter of parts produced by a process theta hat would be the sample mean diameter of a set of parts chosen from the process is it so; obviously, you know the true mean is never known; that means, the population mean may not be known.

So, what you try to do; that means, from the population you collect the sample of and the sample becomes a representative of the population and then you measure individual values individual you know you get the individual values or individual units related to the quality characteristics and you have the sample values. So, from the sample values you calculate the mean. So, how many such samples you will consider you will be considered several of samples is it because if you remember when you refer to the control chart you say that I will be the plotting the sample values over the samples and these samples are drawn you know maintaining the time order is it. So, what is the expected value of theta hat, it represents the mean or expected value and against theta hat.

You have the several sample values so; obviously, from several samples values you can calculate it is standard deviation. So, the notation is sigma theta hat standard deviation of the estimator theta hat. So, how do you what will be the expression for central line that is the expected value of theta hat upper control limit is expected value plus k times this one sigma of theta hat and similarly the lower control limit is expected value of theta hat minus small k times sigma theta hat. So, we are small k represents the number of standard deviations of the sample statistic that the control limits are placed from the centre line is it, usually unless otherwise stated is it.

Usually the value of a k is assumed to be 3 and that is why many a time we refer to the control limits as plus minus, 3 sigma limits is it.

(Refer Slide Time: 04:32)

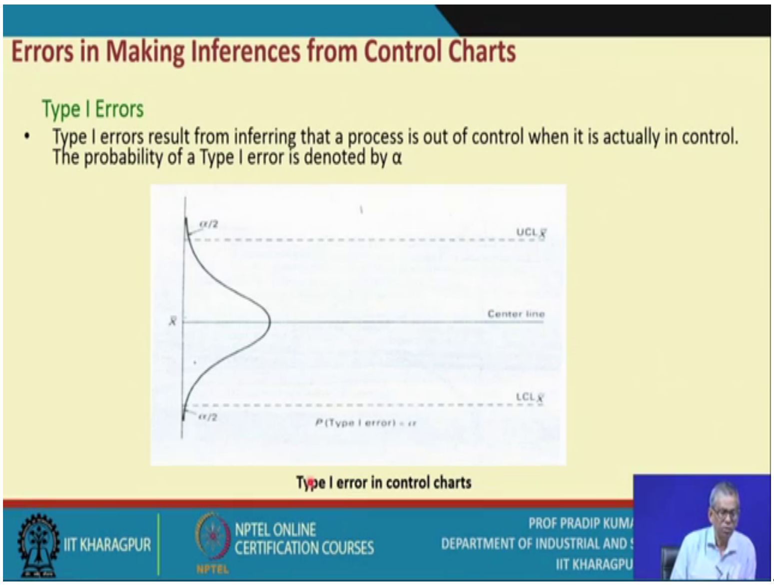


So, here is an example a semi automatic turret lathe machines the thickness of a part that he subsequently used in an assembly. So, thickness is your basic quality characteristic of interest the process means is known to be 30 mm these value is given to you with a standard deviation of 1.5 mm; that means, this is also the process standard deviation or the population standard deviation construct a control chart for the average thickness using 3 sigma limits if samples of size 5 are randomly selected from the process.

So, this is your problem, how do you get the solution control limit is known as 30 mm this is the centre line is 30 mm the standard deviation of the sample mean as per the central limit theorem here sigma X bar equals to sigma upon root n is it. So, this is 1.5 divided by root over 5 that is 0.671 mm, what will be the upper control limit that is 30 plus 3 times 0.671 that is 32.013 and the lower control limit is 30 minus 3 times 0.671

that is 27.987. So, I think there will there is a no problem in getting these ideas and in understanding these ideas. So, it is the most simple problem and we are assuming that control limits are at plus minus 3 sigma limits, UCL is known LCL is known.

(Refer Slide Time: 06:20)



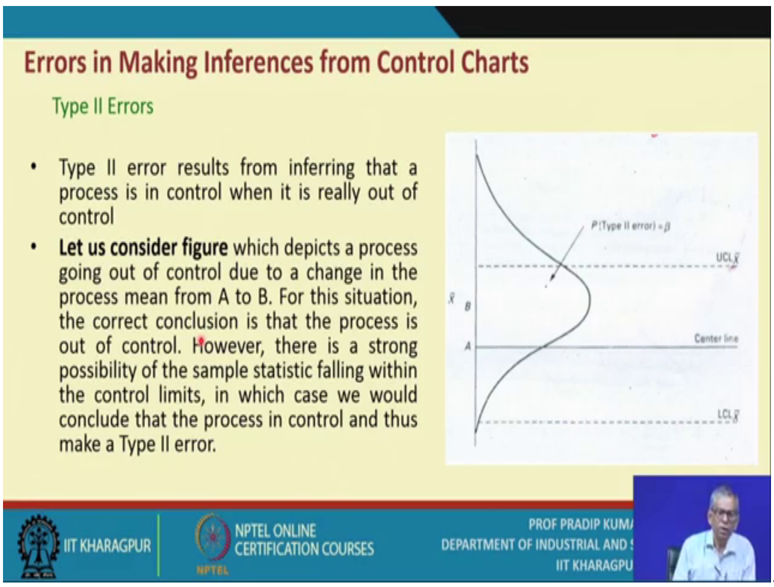
So, now you know you are trying to control the process and the desirable condition that you ask for in a process that is in control condition or in control state and if the process remains in control state; that means, that process performance is guaranteed and the process becomes sustainable now how to monitor the state of the state of the process; obviously, you know you are using one or more tools one such tool that is commonly used.

That is the control chart now whenever you use certain kinds of tools of the control charts while you make a inferences about the state of the process there will always there is a chance that you will be make some mirror and this is this is known implicitly we would like say any tool techniques you use ; obviously, in the unity of the cases you expect that there will not be error, but the error will be there and it is your duty while you design a control chart related to a process you design the control chart in such a way that this value of values of these errors should come down or should be minimum now in respect of control charts the 2 types of error you may commit while you try to make inferences about the state of the process this 2 errors or type 1 error and type 2 errors.

So, what is a type 1 error now the type 1 error means; that means, the control chart source that the process has gone out of control and it is assumed, but when you do actually verify your process you may find that the process remains in control, this is type 1 error; that means, we are inferring that the process is out of control by looking at the control chart formation but actually it is in control state, the probability of a type 1 error is denoted by alpha. So, at any point in time while you use a control chart you must be in a position to calculate this calculate this type 1 error here for example, you have the upper control limit and you have the lower control limit for say kind of chart that is X bar chart that is why it is written in the subscript. So, that is UCL X bar this is a central line and you have LCL X bar. So, now, what do you have you have plotted several values and these values when the process is in control actually, it shows you know a normal distribution these are the values.

Now out of all these values you have plotted is it, this is the summary what do you find that the certain values are beyond upper control limit whereas, certain other values are beyond you know the lower control limit and if it is a normally distributed then; obviously, there is the area under the curve beyond UCLl and LCL this is alpha by 2 plus alpha by 2; that means, alpha is it. So, this is a symmetric distribution that is why these 2 areas must be are same, this way you know even if the process is in control you will find that you cannot avoid these 2 areas these are shown beyond upper control limit as well as the lower control limit is it listen.

(Refer Slide Time: 10:30)



.

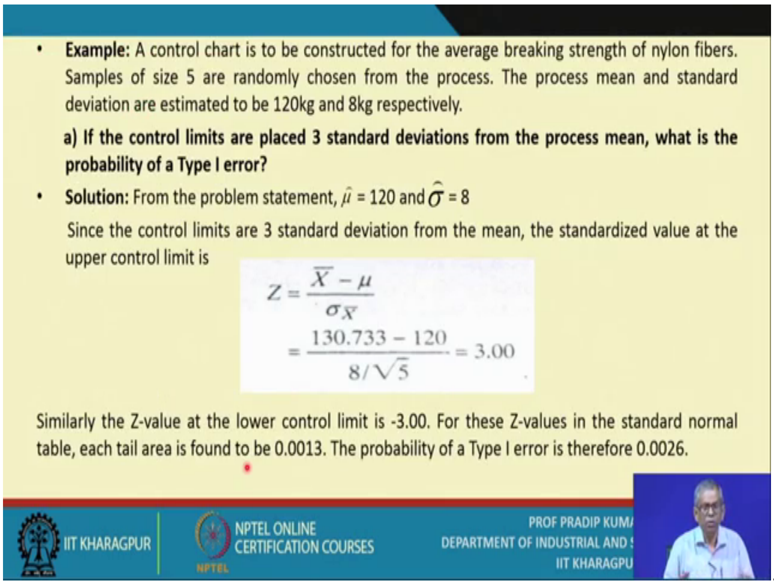
This is the kind of error and sometimes this type 1 error is also referred to as you know the false warning is it; that means, unnecessarily you know you stop the process you try to find out that is there any assignable cause or not and you do not find any, you say the process remains in control.

But the second type of error it is it becomes very very critical in most cases and this is type 2 error. So, what is type 2 error, type 2 error is; that means, as per the control chart information is it you assume that the process is in control, but actually it has gone out of control is it so; obviously, you know it is if you assume so; that means, we were unaware of you know assignable causes impacting the performance of the process and that any point in time the process may stop functioning.

So, the type 1 error results from inferring that the process is in control when it is really out of control. So, let us just look at this particular figure this figure depicts a process going out of control due to a change in the process mean from A to B. So, when you look at a particular control chart actually the central line is indicative of that the process is under control and should be under control you know a along this centre line is it. So, now what has happened that the process was in control around A; that means, the values will be plotted like this what is happened that due to some reasons that there is a shift in the value of A and is a permanent shift suppose; that means, now the process is working definitely, but the value of A has now become value of B, this is an upward the shift is it.

So, now the entire processes shifted around this, but you still are using the same control chart and what do you find that this process actually which is gone out of control is it, but this is the area; that means, upper control limit and the upper control limit this entire area; that means, this is the probability it shows that this is the probability that you assume that the process is in control. So, process mean has shifted from A to B for this situation the correct conclusion is that the process is out of control, but that you cannot do; obviously, you have not change your control it is your control chart with respect to B. There is a strong possibility of the sample statistics falling within the control limits; that means, this is the area; that means, many such points is it many degree of the points it still fall in this particular zone. So, this, we would conclude that the process is in control and thus make a type 2 error, there is a high probability.

(Refer Slide Time: 14:02)



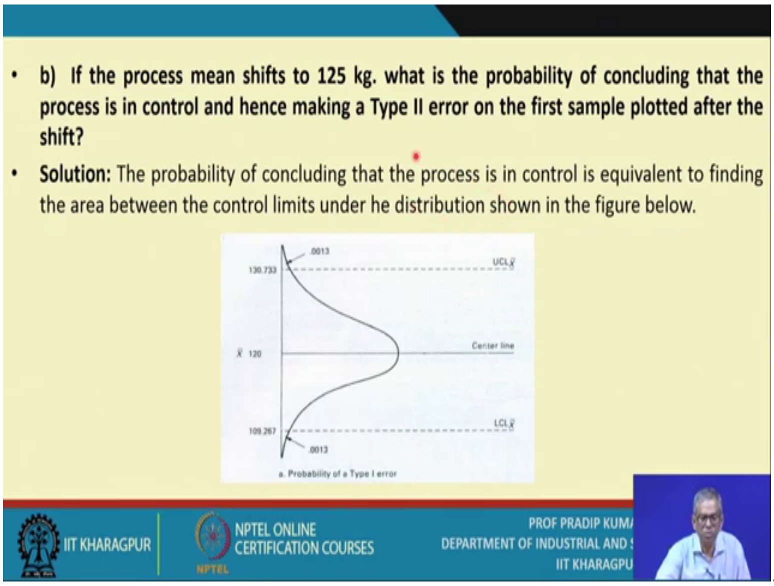
So, your system should be such that you must be able to detect the type 2 error as quickly as possible it is unavoidable, but make sure that you are prompt enough and you are sensitive enough and you have mechanism with you with which you will be able to detect the occurrence of type 2 error and you will be in a position to say you know the bring back the process from an out of control state to an in control state. So, here is an example I will just quickly I will go through this example a control chart is to be constructed for the average baking strength of nylon fibers samples of size 5 are randomly chosen from the process.

This is you have to do the process mean and the standard deviation are estimated to be 120 kg and 8 kg respectively right, these are known. So, these 2 values are given, so we are assuming that the control limits another do exist at plus minus, 3 sigma limits from the mean, what is the probability of a type 1 error. So, what do you do; that means, you have you know what is the upper control limit, we have already determine this is the upper control limit and this is the say the process mean and this is basically the sample standard deviation and the sample standard deviation is given by you know sigma X bar is equals to sigma upon root n as per central limit theorem.

So, what is central limit theorem, central limit theorem says that whatever may be the distribution of X, the distribution of any statistics statistic of X a X bar is it is approximately normal when the sample size is very large is it. So, we are assuming at this stage that the central limit theorem you know central limit theorem is applicable and the normality assumption holds. So, the Z values at the lower control limit. So, this is at the upper control limit that is 130.733. So, we have computed the upper control limit and corresponding Z value we have computed.

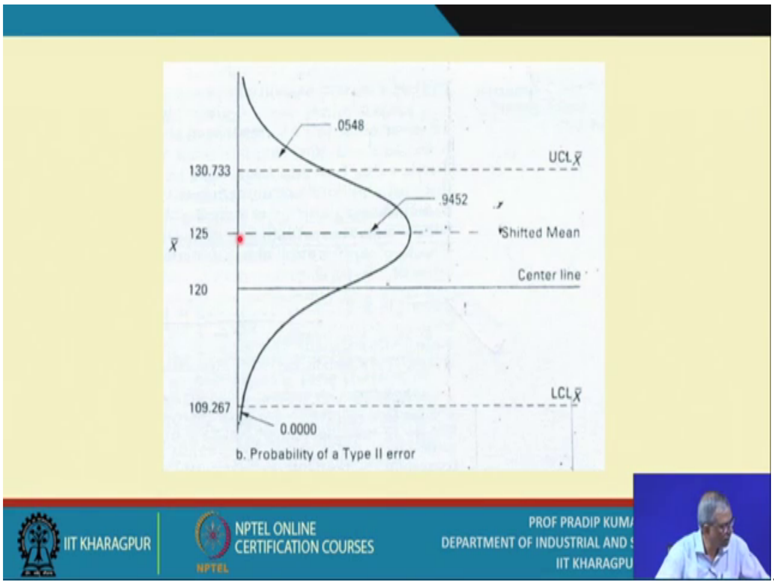
Is it; that means, difference from the mean that is 120 and expressed in the standard deviation units that is 3 similarly if you compute the Z value at the lower control limit you get a value of minus 3. So, these Z values in the standard normal table each tail area is found to be 0.0013 is it, the probability of a type 1 error is therefore 0.0026 clear. So, you should remember these values; that means, in many a time you come across a situation where control limits are at plus minus 3 sigma limits and you assume that the normality assumption is valid, hence you should remember this one. That is the probability of making type 1 error that is 0.0026.

(Refer Slide Time: 17:23)



Now, if the process mean shifts to 125 kg how do you compute the type 2 error it is clear; that means, previously it was 120 now it has become say 125 is it. So, this is you know upper control limit we have already determined 130.733 lower control limit is 109.267 is right. So, this is the area alpha by 2 that is 0.0013 the process is in control around 120 and this is on the other side 0.0013. So, if you add these 2 alpha by 2 plus alpha by 2, it becomes 0.0026.

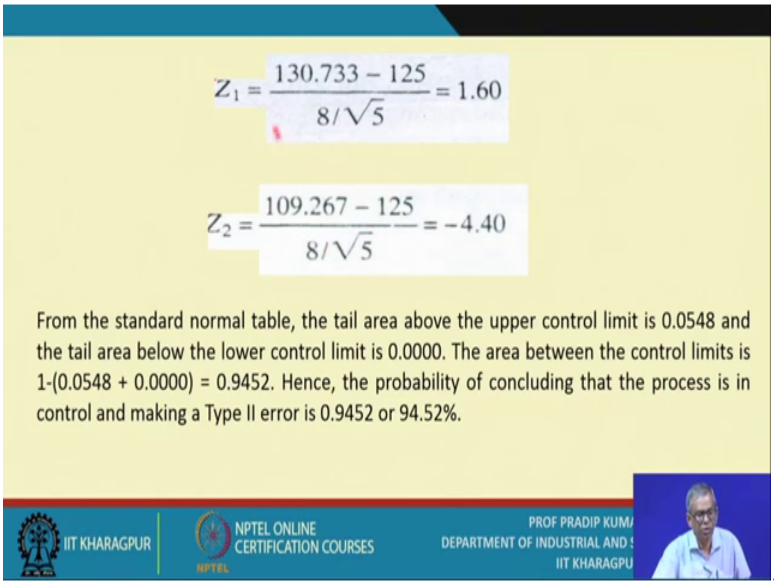
(Refer Slide Time: 18:10)



So, how to compute the probability after of making type 2 error, now, initially it was 120 the process you know the central the value the mean of the process is 120 now it has shifted 125 is it so, but the old you know the control chart is still using and. So, what do you find that this process has shifted around 125 is it. So, what do you find that you need to compute this area there is a probability of making is it type no this is the probability of making say a right conclusions actually, this area is you know the shown as out of control is it.

Similarly, this area is also shown as out of control whereas, the area between this and this within this there in a steel shown within the say upper control limit and lower control limit. So, this area you must know; how do you do this. So, it is very simple; that means, area; that means Z value around say Z value at 130.733 when the mu is 125. So, that you calculate and we are assuming the sigma remains same and you calculate the Z value at 109.267 and the mean is shifted to 125 is it.

(Refer Slide Time: 19:59)

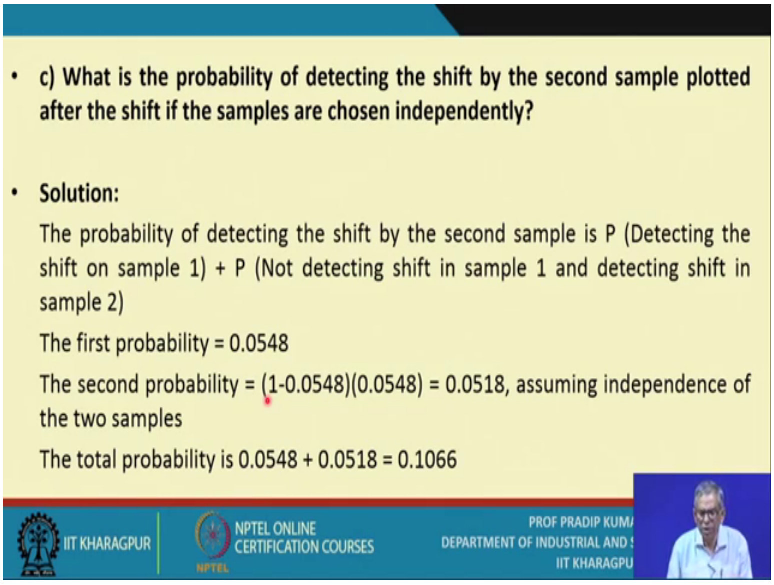


So, this calculation you do; that means, the Z 1 at 130.733 the change to mean value is 125 whereas, we are assuming that the standard deviation does not change, but just make a note that usually you know the both the values you know may change and we may assume that the sigma and the mu they are in many a time they are related is it whereas, right now we are assuming that the sigma is not related to mu is it. So, this is just a fast approximation, this value is against 130.733 now it has become 1.60 and z value at 109.267 is it that is the lower control limit at you know when mu is changed to 125 this has become minus 4.40.

So, from this now you refer to the standard normal table what do you find that the tail area above the upper control limit is 0.0548 and the tail area below the lower control limit that is well things say this is almost 0 ; that means, up to 4 decimal place it is 0. So, the area between the control limits is one minus this plus this; that means, 0.9452 that means, what is your type 2 error probability that is 0.9452 that means, a 94.52 percent of the cases you will be unable to detect the shift.

The shift from 120 to 125 ; that means, though the process has gone out of control a 94.5 percent of the cases you will assume that the process remains in controlled, you will be making type 2 error.

(Refer Slide Time: 21:50)

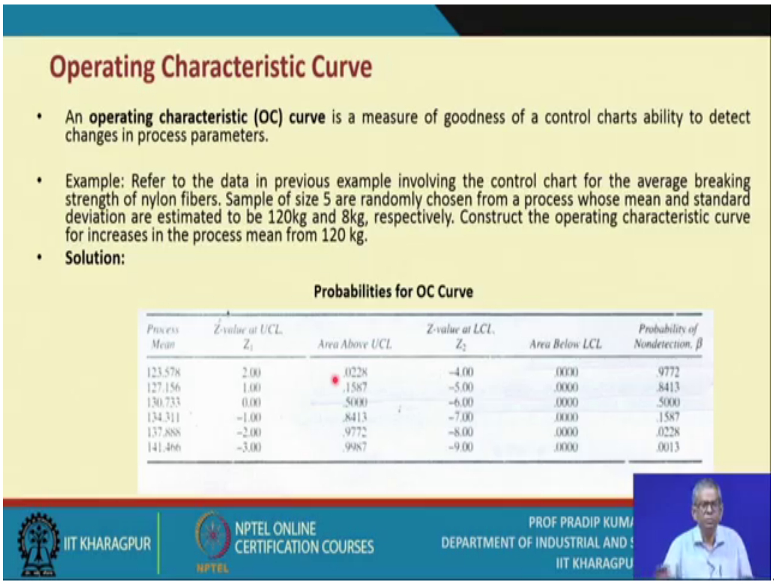


What is the probability of detecting the shift by the second sample plotted after the shift if the samples are chosen independently, what we are assuming at this stage, when you are going to plot all these sample points we are assuming just the point to be noted that is all these sample points are independent with one another; that means, no way one single you know the sample point influences the other the sample points of the sample values.

So, how do you get an answer to this question the probability of detecting the shift by the second sample is the probability that detecting the shift on the sample 1 that is the first event, second event is you are probability that you are not detecting the shift in the sample 1, but you are detecting the shift in sample 2 is; obviously, the first probability is 0.0548, the second probability is 1 minus 0.0548 into 0.0548 that is 0.0518 assuming independence of the 2 samples.

So, this is to be noted right means we are assuming that all the sample points are independent with to another. So, the total probability is this one that is 0.1066 that means 10.66 percent is. So, this is a typical problem we come across and instead of having you know say here by the second sample. So, you may compute the, what is the probability that this is detected by say the third sample or by fourth sample by nth sample is it.

(Refer Slide Time: 23:42)



So, now we will be just a referring to the 2 types of performance measures which are usually used against control chart; that means you are using a control chart and; obviously, there is a chance that you will be making type 2 error now at a certain point in time while we use the control chart you must be able to assess your performance, whether your control charts performances is acceptable to you or not. So, for that 2 kinds of you know the performance measures we recommend the first one is the operating characteristics cause that you have to draw and the second one is average run length and average run lengths curve is it. So, I am going to discuss first the operating characteristic curve what is an operating characteristic curve.

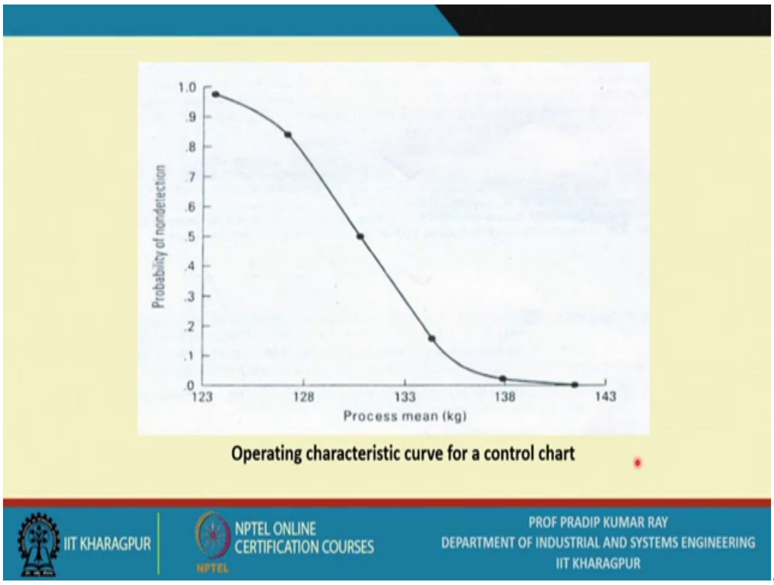
It is a measure of goodness of a control chart control charts ability to detect changes in the process parameters. So, essentially it is a plot of beta versus the you know the process parameter values and you know that parameter depends on which kind of control chart is you use . So, if you use a P control chart later on we will come to know there are different types of control charts you use under you know the different situations, the different conditions we will explain in detail.

So, against a particular control chart you must know what is the process parameter and you also must know what are the possible process parameter values and against each process parameter value you select you must be able to determine say the probability of making type 2 error like in the previous examples we say that the initially the process mean was 120 so; that means, that is a particular value of the process parameter now it changes to 125. So, when it changes to 125 we have computed, what is the probability of making type 2 error. So, this is just one computation it may this process parameter may assume other values like say 130,140,135 and so on against all such possible values which a process may have due to various reasons what is the corresponding probability of making type 2 error.

So, this is just an example we are giving, this way we get the different values of beta and beta values are plotted against several values of theta. So, like say here the process mean there are so many you know the process main values we have assumed and against a particular process mean you know we can compute the Z value and average the area above UCL area below LCL probability of non detection of the shift in the process parameter value that it that way we define the probability of making type 2 error. So, this is your beta value.

Similarly, suppose it is 127.56, corresponding value at that is Z 1 is 1 against that particular control chart which you use is it. So, for which the upper control limit and the lower control limit is in turn. So, Z value at upper control limit that is 1, area above upper control limits 0.1587 and Z value at LCL is minus 5. So, this is 0, the area you know you between upper control limit and lower control limit will be 0.8413. So, this way for several values of process mean I considered you calculate the corresponding values of beta and next you plot them.

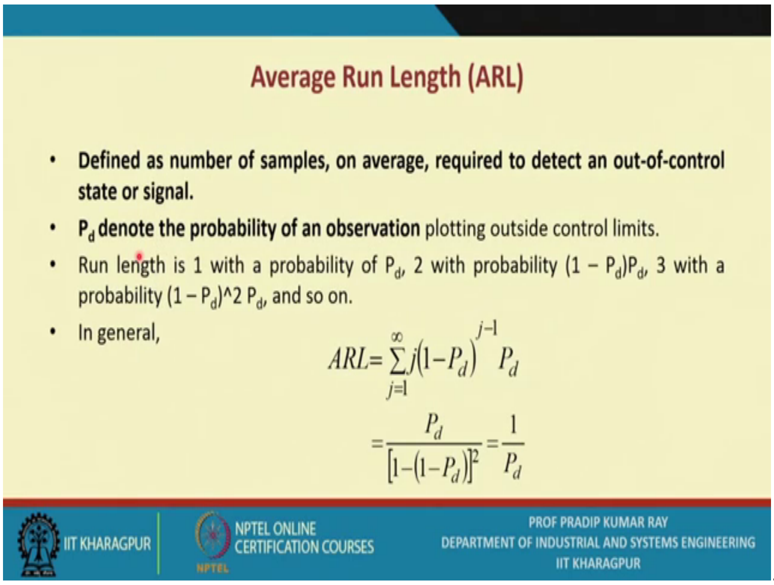
(Refer Slide Time: 27:53)



That means, here we are assuming that say these are the possible values like say 123,128,133,138 these are the possible values; that means, these values are technically feasible for the process parameters and suppose the control chart is drawn over there at 120 is it, that is the base. So, what we are assuming that is you know the positive shift is it there could be negative shift also, but we are just you know by showing that how to the draw the operating characteristics curves. So, just we are you know the positive deviation we have considered possible deviations.

So, we look at this figure what you will find that as the shift magnitude changes is it either on the positive side or on the negative side what we will find that the probability of non detection decreases, but if the ship magnitude is very small is it. So, this is one of the characteristics of a typical that is what control chart; that means, when the shift is small the probability of it is detection is not that high, but ; obviously, this if the shift magnitude is substantial the probability of point detection comes down.

(Refer Slide Time: 29:24)

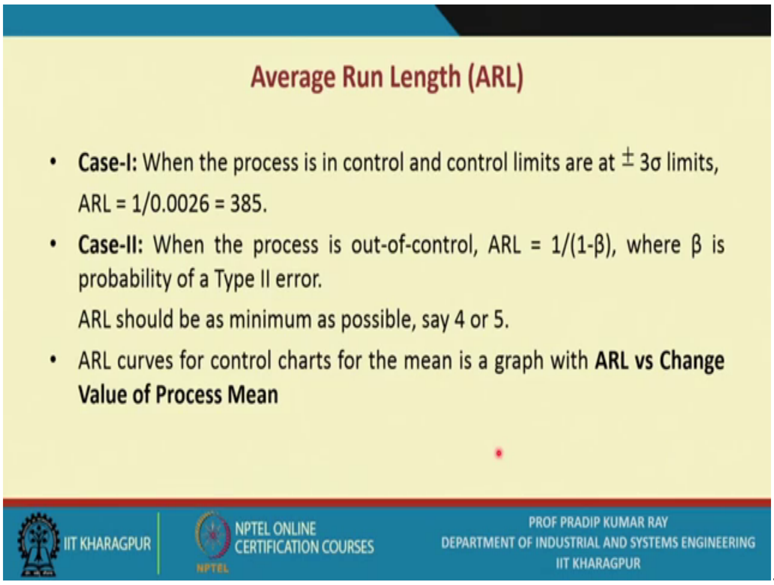


So, this is a typical one the next one is you know there will important say the performance measures many a time we recommend that is the Average Run Length. Now what is the average run length average the run length is defined as the number of samples on an average that you require to detect an out of control state or signal I repeat; that means, average run length is, you start using a control chart at any point in time the process may go out of control. So, and you are basing your decision on the sample values and you are referring you are comparing the sample values with the upper control limit or the lower control limit.

So, how many you know if the process goes out of control, how many were in the sample points you are required to plot before you will be able to detect this condition out of control condition. So, supposing Pd denotes the probability of an observation plotting outside the control limits that is the probability, the run length is one with a probability of Pd is it; that means, you plot this point what is the probability that this point will be plotted beyond the upper control limit or beyond lower control limit that is the Pd, so it will be one if it is two; that means, the first point when you plot is it is not detected as an out of control point , but the second point definitely and that is why the run length is 2 it could be three; that means, the first to sample the points is it, they are not shown as out of control point, but the third one is shown as are out of control points. So, corresponding probability is this.

So, average or the expected value of run length how do you calculate; that means, the general formulation is sigma Z equals to 1 to infinity Z into 1 minus Pd to the power Z minus 1 into P d is it and you know the convergent infinite series so; obviously, you know the ARL ultimately becomes one upon P d, this is the most simple formula.

(Refer Slide Time: 31:43)



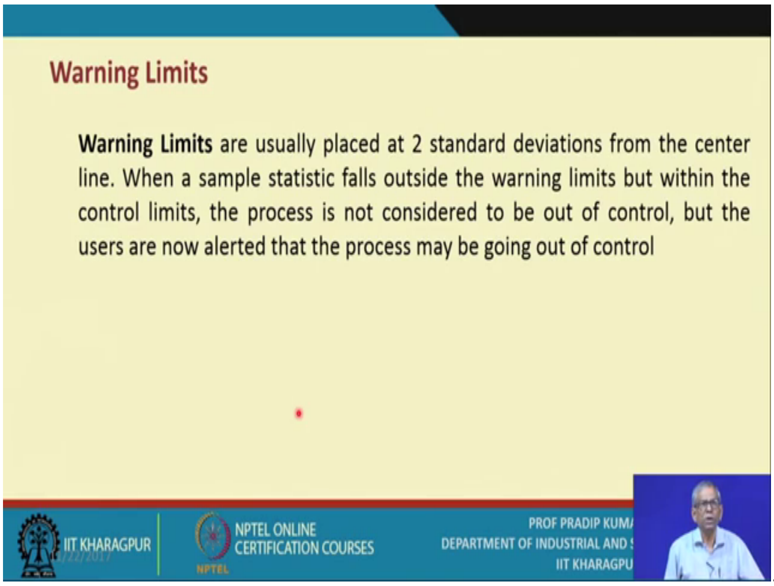
Now you always come across 2 situations as far as you know the process condition is concerned the first condition is the case 1 when the process is in control and control limits are at plus minus 3 sigma limits this is a very common occurrence. In fact, the process is in control. So, what is the value of ARL is; obviously, 1 upon.

Student: (Refer Time: 32:10).

1 upon alpha, alpha is 0.0026 so; that means, it is 385 that means, on an average every after 385 the sample points were going to plot those 385th point may be shown as an out of control point. So, it is clear the next case is when the process is out of control is it there could be many reasons and you know there are many reasons you will come to know later on. So, what will be the expressions for ARL, ARL is 1 upon 1 minus beta. So, what is beta, beta is a probability of detection of the shift and what is 1 minus beta; that means, the probability of 1 minus beta is a probability of detection and beta is probability of non detection and the beta is the probability of type 2 error. So, ARL should be as minimum as possible say 4 to 4 or 5, when the process goes out of control; that means, you must know that; what is that probability of making type 2 error; that

means, probability point detection? So, ARL curves for control charts for the mean is a graph with ARL versus change value of the process mean.

(Refer Slide Time: 33:30)



Now, here what do we conclude; that means, sometimes what do you do you give the warning limits is it and these are all actually you know the placed at plus minus 2 sigma from the centre line. So, it; that means, in order to have an online real time control you must place the warning limits also is it.

(Refer Slide Time: 33:50)

Now the basic issue is that the process may be in control or process may go out of control. So, if the process is in control is it, you try to know that what are the reasons that why you know, what is specifically related to a process, you must be aware of the process settings and what extent you relate the process condition or the state of the process with the process settings this is first, part the second part is that at any point in time the process may go out of control.

So, immediately you know the first thing you have to do; that means, how quickly you are able to you know the detect this condition and if you are able to detect the conditions you must be able to identify the assignable causes and then you have to take some you know the corrective measures at they at the first stage and later on you have to take some preventive measures. So, whenever someone starts using the control chart process we are assuming that he believes in the concept of prevention based quality control, other aspects we will discuss in the subsequent lecture classes.