Total Quality Management - I
Prof. Raghunandan Sengupta
Department of Industrial and Management Engineering
Indian Institute of Technology, Kanpur

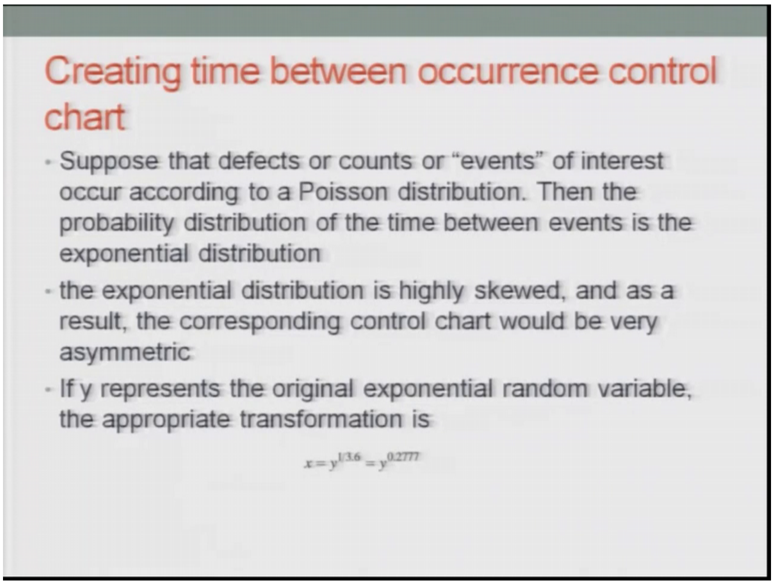
Lecture - 32
Introduction to Process Capability Analysis

Welcome back my dear friends; very good morning, good afternoon, good evening to all of you. And this is the TQM 1 lecture on the NPTEL MOOC series. And I am Raghunandan Sengupta from IME department IIT Kanpur.

So, if you remember we were discussing on the different type of charts. The initially like if I go back about 5 6 lectures back. So, we have considered the x bar charts R charts which is the range then if range was not available, you wanted to go into the standard deviation then we considered the conformity p c charts, u charts, g charts, h charts, but the concept remains the same that you need to basically have the central value central line the upper control, and the lower control basically depend on what is the standard deviation, the standard deviation would basically be used in depending on what is the overall coverage probability of the level of confidence which you have.

So, if you want to basically cover 99 percent, you will use plus minus 3 sigma. If you basically want to cover 95 would plus minus 2 2 sigma and so on and so forth. And also be considered that if the sample size was wearing we would basically divide by a changing value of n is n I n 1 n 2 n 3. In case if it was not possible we will use the average value which is n bar and do our calculations. So, continuing the same discussion; so creating time between occurrences and the control charts.

(Refer Slide Time: 01:46)



So, we did mention time and again yes consider the set of observations at 9 o clock 10 o clock 11 o clock. So, say for example, one hour interval. So, there is a one hour was just an arbitrarily example.

So, suppose that defects or counts or events of interest occur according to a poisson distribution. Then the probability distribution of the time of arrival would be exponential. So, here I should basically mention something. So, generally considered this there is a tale there is a tailor machine in the bank or a machine in the shop floor and jobs are coming. Jobs are coming at a certain interval and a certain rate and they are being processed and then basically taken out of the system; now considered u as the person who is measuring the number of people who are coming per unit time.

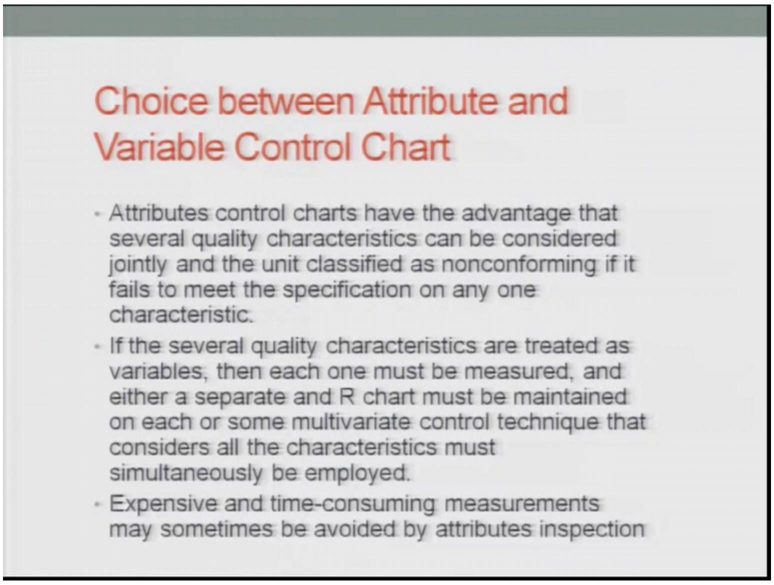
So, what that per unit time would basically depend on the type of example which you are doing. And another person it was also basically trying to find out that what was the rate of processing for each unit which was arriving. So, technically if the per unit time and if the time intervals between those per unit time was IIT's which was independently and identically distributed and if I were able to make that time interval as small as possible. So, technically in the number of items which are coming between the time interval is discrete which is poisson process. And the time to take taken to process is individual is a continuous distribution it will basically be in exponential distribution.

So, that is why the relationship of exponential and Poisson distribution is mentioned in this slide. So, again I will repeated suppose that defects or counts or events of interest occur according to Poisson distribution. Then the probability distribution of the time between events is the exponential distribution. The exponential distribution is highly skewed and as a result the corresponding control charts would be very assent asymmetric. So, if y represents the original exponential random variable then and then the appropriate transformation which will try to utilize would be basically ratio 1 by 3 point 6, but obviously, if the sample size increases very large then; obviously, you can you can obviously, with some obvious they would be error, but it would be quite nice that if we use the central limit theorem, have a certain mean have a certain standard deviation.

The sample standard deviation and the mean which you will be utilizing for the normal distribution would basically be coming from the actual population distribution from where you picking up the observations and doing the work. Now we will basically try to understand the choice is that when we will use the attribute charts and the concept when we will use the variable charts and the concept is the variable control charts. So, attributes control charts have the advantages that several quality characteristics can be considered jointly and the unit classify classified as nonconforming, if it fails to meet the specification on any one of the characteristics.

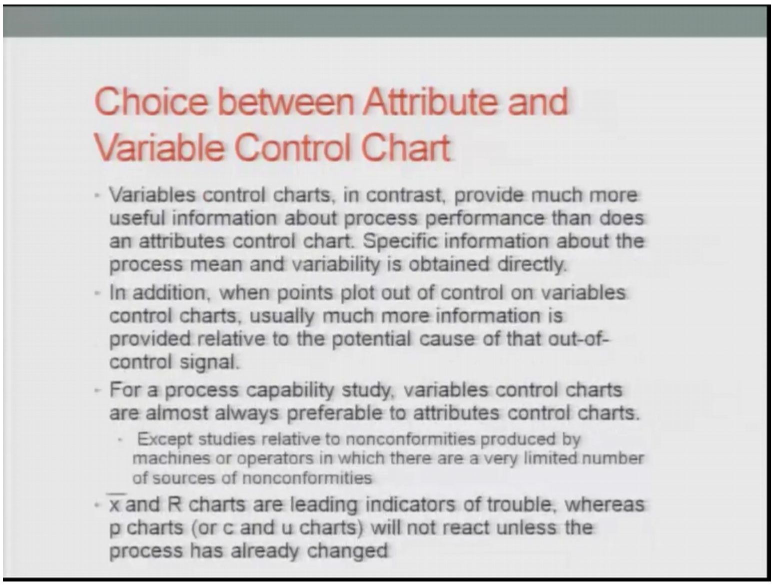
So, if the several quality characteristics are treated as variables then each one must be measured and basically you should have to take corrective actions and pass the decisions accordingly. And either a separate and R charts must be mentioned on each of this variables or a multivariate distribution chart should be drawn in order to understand the variations in a must depth.

(Refer Slide Time: 05:14)



So, they are basically expensive and time consuming measurements may sometimes be avoided by, but the attribute inspections themselves.

(Refer Slide Time: 05:28)



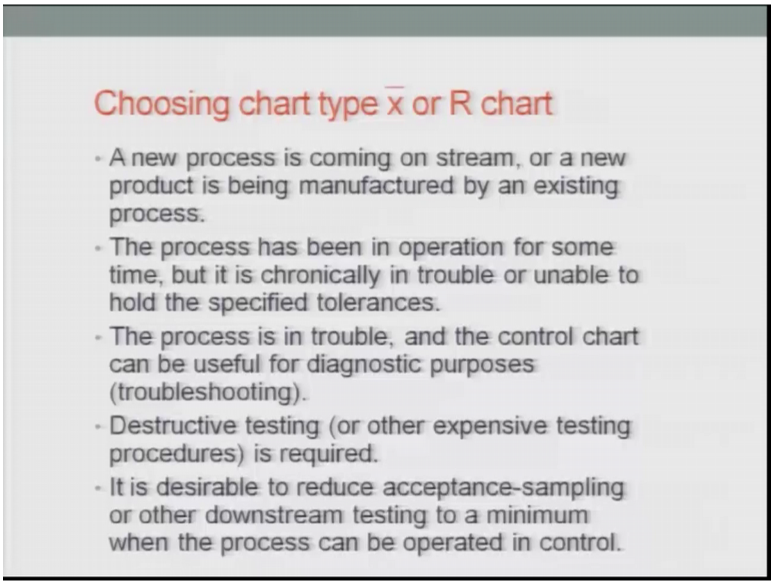
Variable control charts in contrast provide much more useful informations about process performance. So, because the attributes are give you some quality. So, good bad yellow green whatever it is, but the at the characteristics of the variables would basically give you some information based on which you can find out it can be either the width length height of the numbers whatever it is. So, quantify (Refer Time: 05:54) which you can measure.

So, continue may reading along with me variables control charts in contrast provide much more useful information about the process performance than does an attributes control charts. Specific information about the process mean and variability is obtained directly. In addition when points plot out of control on variable control charts usually much more information is provided relative to the potential cause of that out of control signal. So, we are able to basically pass judgements for the variable control charts much more authentically and much more logically than for the attribute control charts.

For a process capability study variable control charts are almost always preferred preferable to attributes of the control charts. So, except studies related to the nonconformities produced by the machines they good bad or several colour change is there black or white whatever it is. Produced by the machines or operators in which there are very limited number of source of information for the nonconformities charts and at values.

So, x bar and R charts are leading indicators of the trouble whereas, p or c or u are basically they are the tables which are basically used for the for the attributes. So, so as it reads they would not react unless the process has already been changed. So, they would basically continue to use the p u and c charts for the attributes and basically x bar and R charts and also sigma charts for the variable charts.

(Refer Slide Time: 07:30)

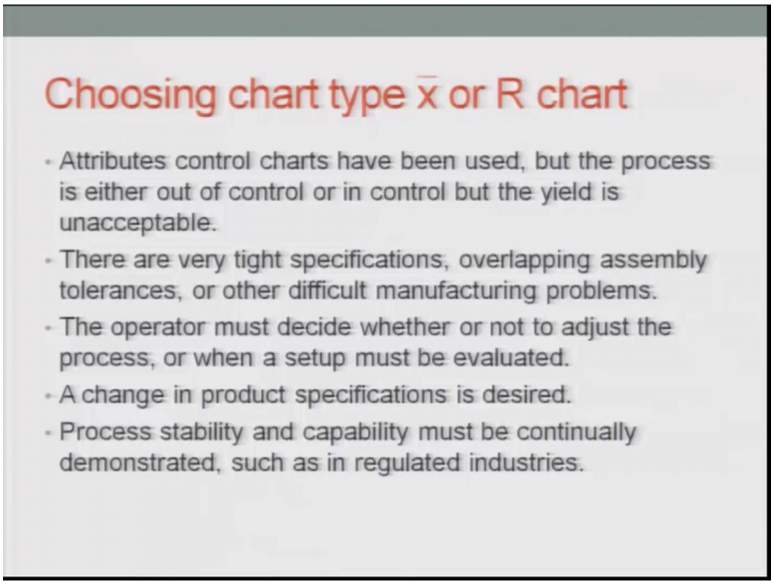


Choosing between x bar and R charts a new process is coming on stream or a new product is being manufactured by an existing process. So, in that case what you will do the process has been in operation for some time, but it is chronically in trouble or unable to hold the specific allowable tolerance so obviously, you have to take some decisions accordingly which means there is a huge amount of variability if there is a variability you will be you will definitely should use the r on the range charts. The processes in trouble the control charts can be used for diagnostic purposes or troubleshooting and then you take corrective actions accordingly.

So, destructive testing or other expensive testing procedures is required so obviously, it would mean that the total cost is quite height. So, you have to be careful and basically do the testing destructive testing for those products, where assure you will get some informations which is useful in order to pass some logical and very, very intelligent statement about the process capability of that hold process.

So, it is desirable to reduce an reduce acceptance sampling or other downstream testing to a minimum and the process can be operated in control.

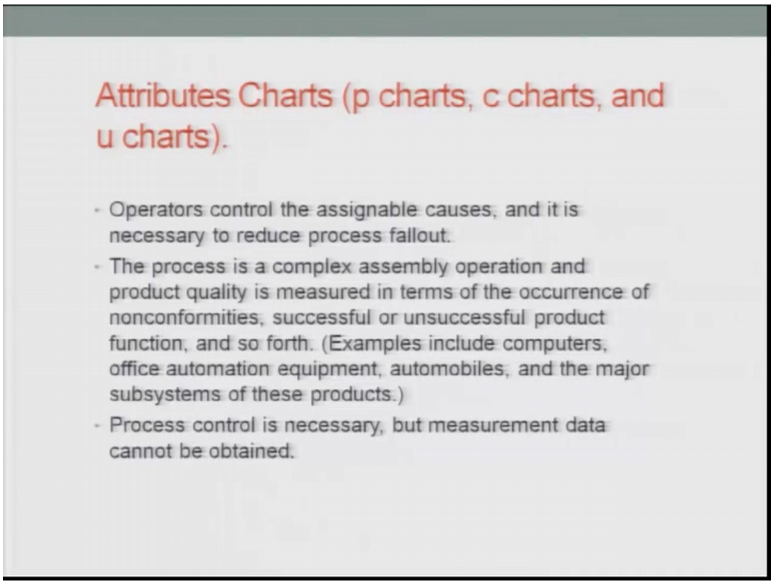
(Refer Slide Time: 08:36)



So, continue continuing the discussion of choosing the chart, chart type whether the x bar or the R charts. So, attributes control charts have been used, but the process is either out of control or in control, but the yield is unacceptable based on that whatever the information which you have. So, they are very tight specifications overlapping assembly tolerances or other difficult manufacturing problems. So, the operator must decide whether or not to adjust process or when a setup must be evaluated and then corrective actions should be taken as necessary.

Or a change in product specification is desired and then you will basically want to implement those specifications for the production process as required. Process stability and capability must be continually demonstrated such as in regulated industry and you will basically take back the information from the industry and basically make your studies of the of the charts whether they are related to quality control or whether related to some variables being studied whether they are related to some attributes being studied; obviously, we will take corrective actions to find out whether problem is, but before that the charts have to be drawn.

(Refer Slide Time: 09:46)

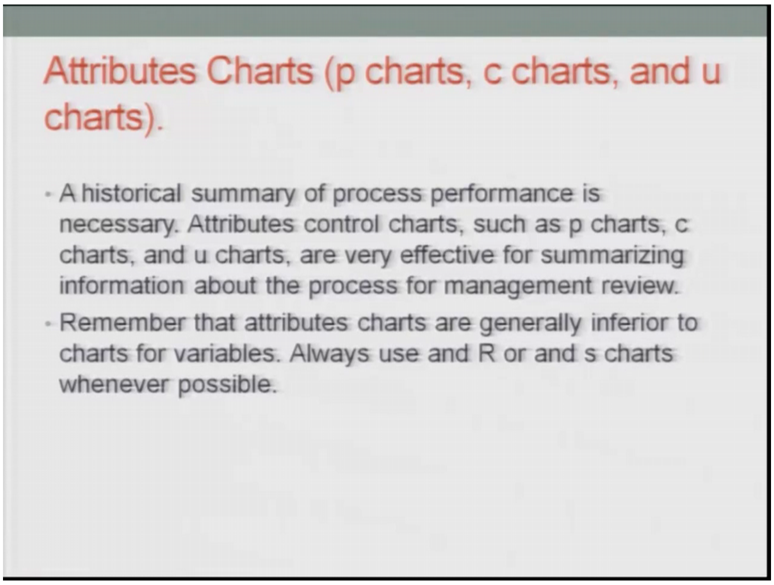


Coming back to the attribute charts we mentioned that for the variable charts, we use the x bar and the r and sometimes the sigma charts which is the basically the standard deviations. Now in the attribute charts you have the p charts the c charts and the u charts. So, operators control the assignable causes and it is necessary to reduce process fallout. So, if there are in variations the process is a complex assembly operations and product quality is measured in terms of the occurrence of nonconformities and so on and so forth.

So, examples include computers, office automations, equipments, automobiles and the major sub subsystems of these products which are being basically considered. Process control is necessary, but measurement data cannot be obtained. So, in this case if they are not obtains. Obviously, you will basically have the attributes, attributes you will give you basically some information whether the process is as per the normal outside.

So, if you are not able to get exact data of the variables you will try to utilize the attribute characteristics in order to study the overall process.

(Refer Slide Time: 10:50)

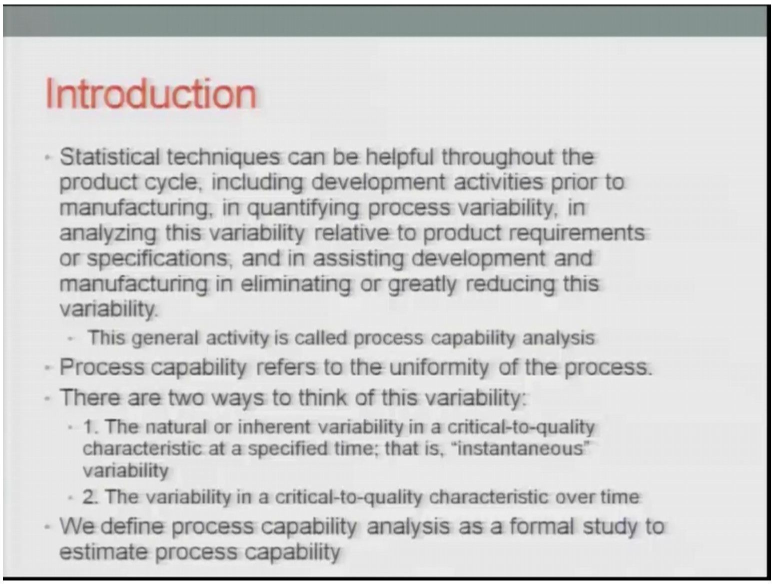


Continuing with the attribute charts; so p c and the u and later on we saw that we had the g and the h charts also. So, historical summary of the process performance is necessary. So, attribute control charts such as the p charts c chart u charts are very effective for summarizing information about the process for management review and then the management basically can study and pass concrete judgement about the whole process how it is going.

Remember that attributes charts are generally inferior to charts for variables which we did not mentioned. So, always use the r or the s charts whenever a possible. So, if you remember the r was the range and s was the standard error. So, they would be utilized because they would give you some more extra set of information about the variability or the process is occurring. Because say for example, the mean value is fixed, but if the variability is changing quite dramatically is; obviously, there are some problem the overall process. So, that would basically be taken into consideration when you want to compare between where to use x bar charts and where to use r and sigma and the s charts.

So obviously, if you remember I did mention frettingly though that the r and the sigma the s charts are the best way of trying to get the information, but; obviously, you will only do that in case if getting such information was not possible when you are trying to draw the charts using the attributes.

(Refer Slide Time: 12:18)

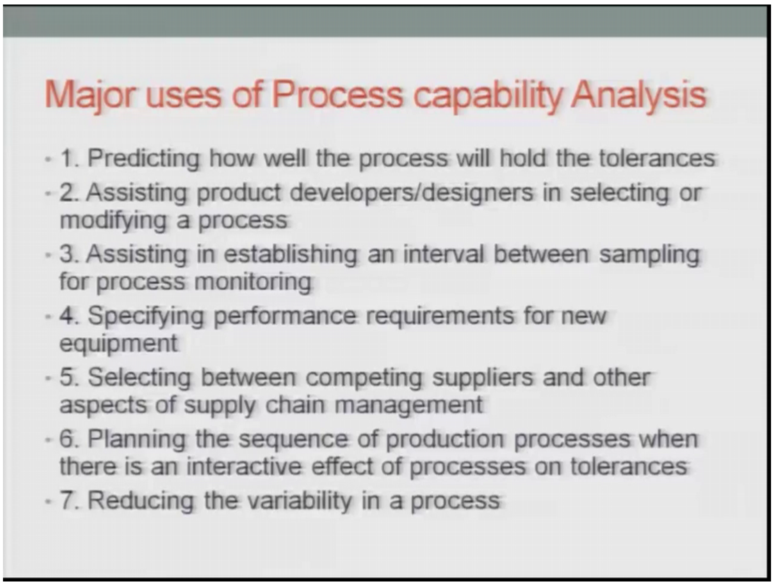


So now we will basically continue some discussion over the statistical techniques in further details. So, in statistical techniques can be useful throughout the product cycle including development activities prior to the manufacturing process in quantifying process variability, in analyzing the analyzing this variability related to the product requirements; or specifications and in assisting development and manufacturing and in eliminating or greatly reducing the variability.

So, these general activities call the process capability analysis and one the process had started. So, it will be monitor so obviously, the information which is coming out whether there the process under control out of control we do not know. So but obviously, the information is coming we are trying to monitor it or the team is trying to monitor it. So, the process capability refers to the uniformity of the process. So, there are 2 ways to think of the variability so obviously, in any process they would be variability some would be external noise or which you cannot control some would be internal one which you should definitely control.

The natural or inherent variability is a critical to quality characteristics at a specified time. So, that is instantaneous variability and that can be utilized according to what the person who is doing the study can pass corrective judgement about whether to use the variable charts or the attribute charts and take corrective actions accordingly. The variability is a critical to on the critical to quality characteristics over time we define process capability analysis as a formal study to estimate the process capability. Major use of process capability analysis are predicting how well the process will hold the tolerance assisting product development designer in selecting or modifying a process.

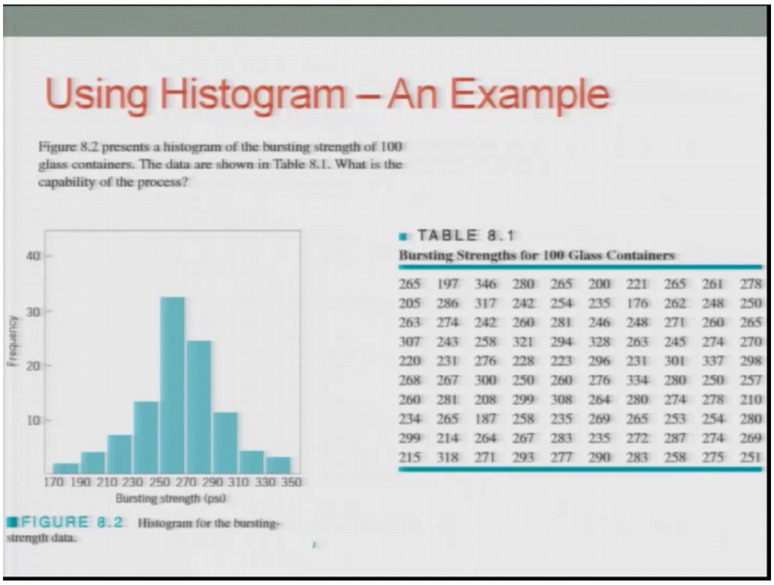
(Refer Slide Time: 14:01)



Assisting in establishing an interval between sampling for process monitoring there should also be considered. Then specifying the performance requirement for new equipment should be mentioned depending on new comments are coming or not. So, we should also look at the other on the major uses a process capabilities would be selecting between competing suppliers another aspects of the supply chain management. Planning the sequence of product processes when there is an interactive interactive effort of processes and tolerances, and you basically try to reduce the variability in the overall process.

So, this you will be basically tempted to use the process capability analysis.

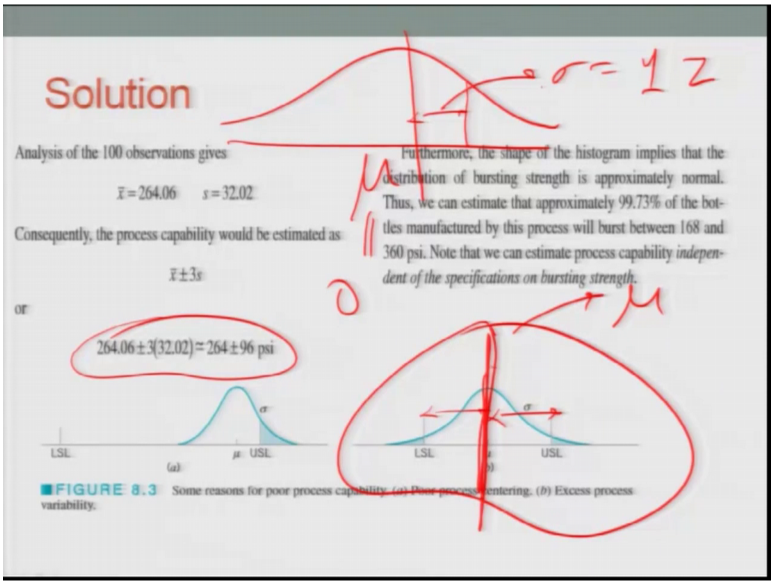
(Refer Slide Time: 14:53)



So, here is a diagram. So, figure is 8.2 represents the histogram of the bursting strength of 100 glass containers. So, the glass containers are this underway industructative testing and you find out the bursting strength for 100 of them. So, the level of bursting strength is basically giving along the x axis mm and the frequency or the relative frequency or the probability that basically technically mention along though y axis.

So, the table gives you the bursting strength numbers. So, you basically divide in 2 intervals find out how many intervals are there and the finite you do better would be a basically going move towards the normal distribution, but; obviously, it will depend on the what is the whole set of data which you have in this caseyou basically 100 such infra set of informations are there.

(Refer Slide Time: 15:51)

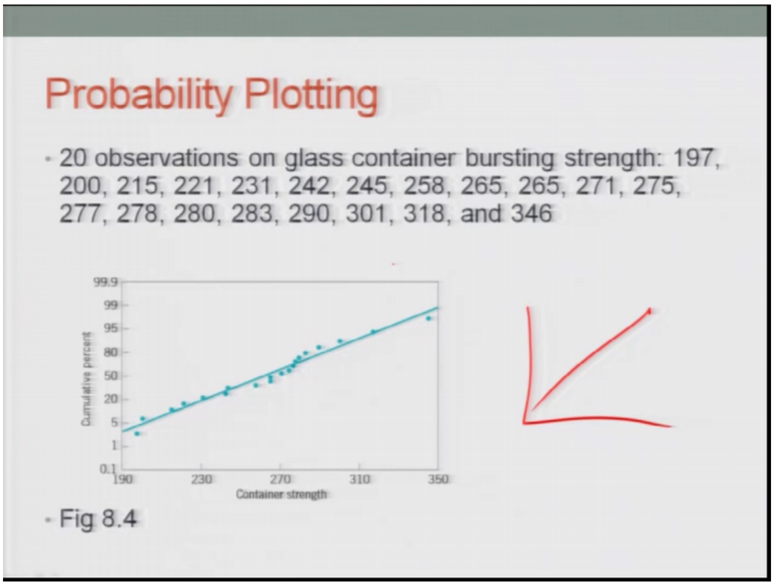


So, administrative 100 observation gives you the bar if the central value which is x bar is 264.06 s which is the standard deviation comes up to 32.02 consequently the process capability would be given an estimated at x bar plus minus the value of level of confidence which you want to place multiplied by the standard error. So, it would be basically x bar plus minus 3, 3 would basically the overall coverage is 66, about sorry 99 multiplied by 32.02 which 32.02 it basically the standard error of the whole process.

So, if I basically find out the intervals. So, as I was mentioning this is to 264.06 is the average value plus minus 3 is 32.02 and you found find out the psi which is the unit of the pressure basically comes out to 264 point plus minus 96 psi. So, furthermore the shape of the histogram in implies the distribution of bursting strength is approximately normal. So, this is what I am saying is that as you keep increasing the sample size they would slowly tend to 2 become the normal distribution. Thus we can estimate the approximate 99.73 of the bottles manufactured by the process, will burst between 168 and 360 psi. So, no not know that we can estimate process capability inde independent the specifications on a bursting strength.

So, based on that you try to find out that what is the level average value and what is the dispersion which is there below and above the mean value. So, in this second diagram which you see, the central line is given as mu. And the dispersion on to the right is given by the USL which is the upper control, then you have basically the lower control and depending on how the process has been done we can basically find out the UCL and the LCL. Now one thing should be remembered the sigma value which we have pertains to the standard deviation of the process. So, this I have been mentioning time and again and the level of confidence will come from the so called value of z. So, it is 3 or 2 or one would depend on the level of confidence. So, let me draw the standard normal.

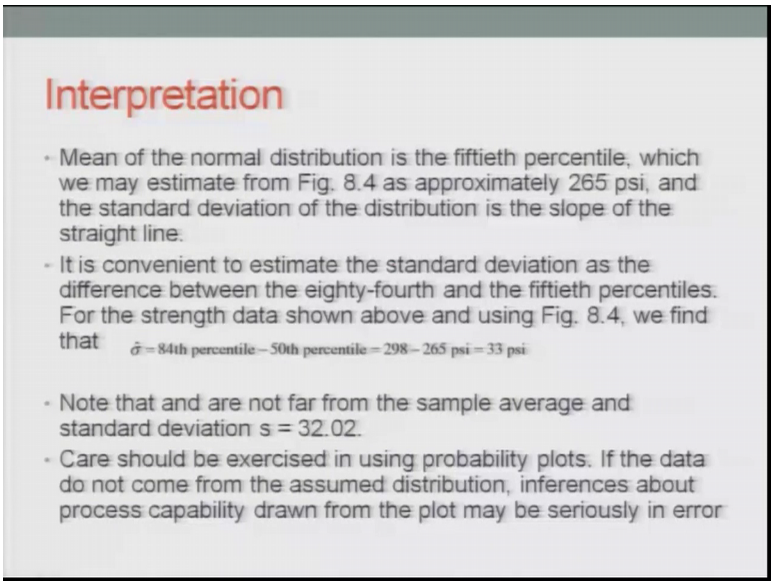
(Refer Slide Time: 19:55)



So, this is mu, technically this would be 0 for the for the z distribution. And this value which we have is sigma this technically would be 1 for the z distribution. And if you move, move 3 such units on to the right and 3 such units on to the left. So, that would basically cover about 99 percent of the overall coverage. I am repeating it please do not mind because the more you hear it would be much become much clear that how can use the concept of mean the standard deviation to draw the limits concerning the distribution is normal using the concept of central limit theorem based on the fact the mean and the standard division all this of the normal distribution which you are using basically coming from the actual original population distribution.

So, if it is a poisson we will use the poisson distributions mean and the variance, if it is a normal we will use the mean and the variance on the normal distribution and basically sorry binomial one we will use the binomial distributions mean and variance and use that in the central limit theorem. So, once you plot the observations for the bursting strength. So, these basically comes out and if you remember the qq plots which I mentioned the 45 degrees line they are not actually 45 there are some aberrations so obviously, is a not a normal distribution. So, the values are given as one 97 for the 20 observations 200, 250 till the last 3 one which is 300: and 1 300 and 18 and 346.

(Refer Slide Time: 20:22)

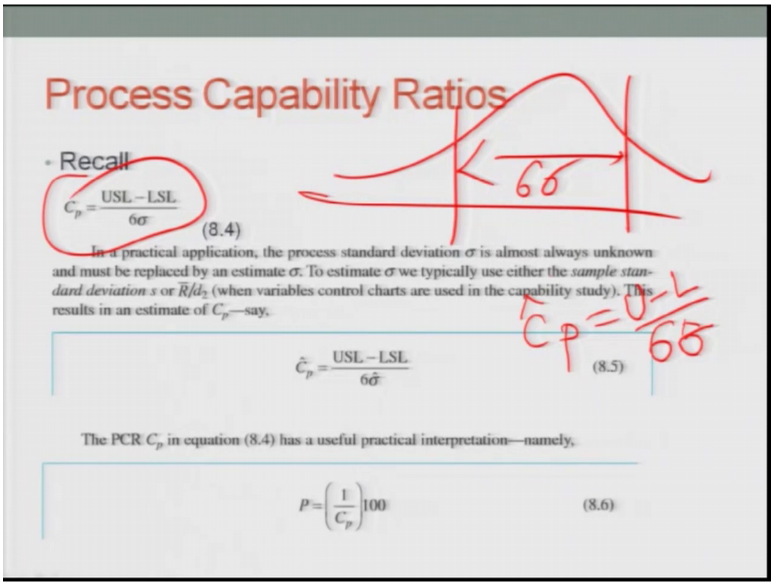


So, you want to basically interpret the results. So, the interpretation of the results would be mean of the normal distribution is the 50th percentile which will estimate from figure 8.4 as just mentioned. So, basically it will be 265 psi. So, pounds per square inch the pressure which is there. So, that would basically depend on what is the average value and plus minus of that would give you the overall so called dispersion based on which you can see yes within this range the bottle should not burst. And the standard deviation of the distribution is the is the slope of the straight line. Now technically it means that when you are trying to basically have more and observations are less and less observations, the values of the distribution would basically be either very shrunk or very large.

So, this I am going to basically mention it very say the simply. Now consider the distribution is normal where I am drawing the x is normal with certain mean and certain variance, sigma square when you find out with the best estimate of mu would basically x bar and with a certain mean which is again mu and sigma square by n is basically the variance. Because that would basically depend that for each set of variance the variability between them will slowly started decreasing. So, if you take the value of n as large the variance basically would now happen like this I will try to draw it. So, say for example, for some n value it is this if n basically increases it will because it become a more picked one and it would basically all packed up together. So, let me delete it. So, the second point says it is convenient to estimate the standard deviation as the difference between the 84th and the 50th percentile for the strength of the data shown and the value comes out to be about 33 psi.

Note there are not far from the sample average and the standard deviation as we know is about 33 point 32.02. So, care should be exercised in using the probability plots if the data do not come out from the assumed distribution. Inferences about the process capability drawn from the plot may be seriously in an error because the actual information based on which we are trying to draw the standard normal deviate or the distribution for the process control, they would be some conceptual error in what we are trying to consider. So, you want to find out the process capability ratios.

(Refer Slide Time: 23:11)

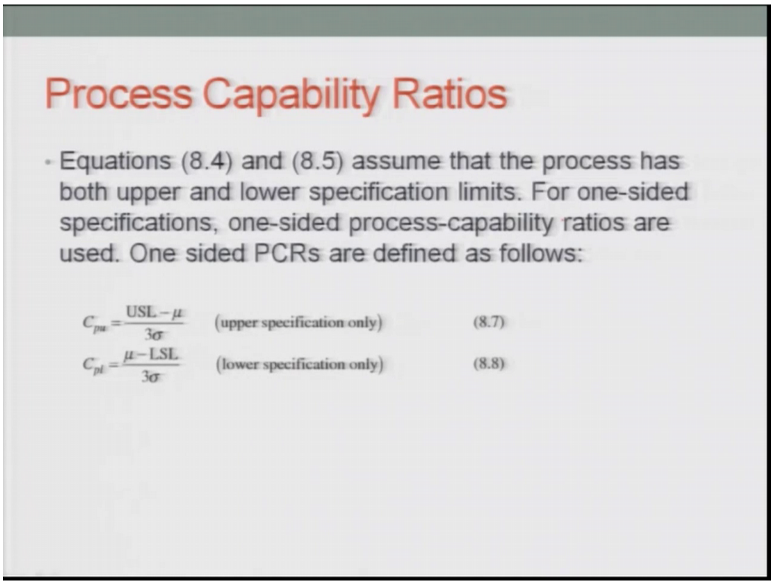


So the recall; now, you want to find out that what is the overall coverage. So, if you see this formula this would immediately make sense that what I have been repeating time and again. So, what we have from your side if you are looking the upper control limit is on my left hand side the lower control limit on the right hand side on my right. So, the difference would basically be if you are g1 plus 3 sigma on to the right and plus 3 sigma on to the left so obviously, the overall coverage would be 6 sigma. So, if you want to find out that what is the coverage so obviously, it will be upper control limit minus the lower control limit divide by 6, 6 sigma. 6 sigma basically been the standard deviation that would give you the process capability ratios. So, this is what is happening.

So, if I have the upper control limit lower control limit difference between that divided by 6 sigma would give me the ratios. So, in practical application the process standard deviation sigma is almost always unknown. So, this difficult to know, and must be replaced by an estimator of sigma in the other initial case we found out that the best so called way of trying to find out the dispersion about the range to estimate sigma we typically use either the sample standard deviation. S or the concept we based on which we are trying to get the standard and the standard error of the sample using the concept of the range. So, that was basically r bar divide by d 2 would give you some concept of what the standard error of the sample was. So, continuing reading it means that when variable control charts are used in a capability study we will be tempted to use or we will be required to use the value of r bar by d 2 where d 2 was constant which depends on the sample size.

This results in an estimate of c which is the ratio of c p which is the capability ratio. And if we use the upper control limit and the lower control limit. So, initially what we did the ratio which was finding out was based on the fact of u, u I am trying to use as a upper control and l as the lower control. So, this was basically 6 sigma. Now if we replace a sigma by sigma hat, so obviously, the c value would also with an estimate. So, hence it would be replaced by c hat. So, the process capability ratios have an useful practical implications namely they give me the overall the value of p based on which I can do the calculations accordingly.

(Refer Slide Time: 25:45)



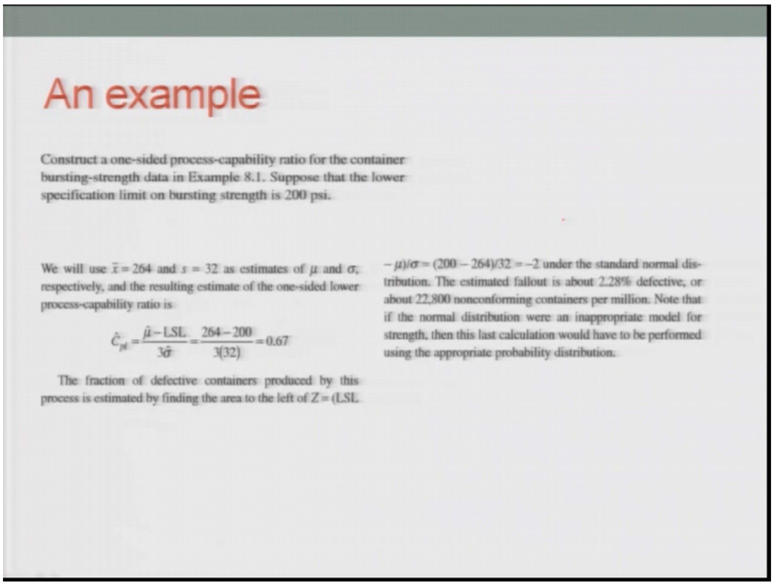
Continuing the continuing the process capability ratios. So, these equations assume that the process has both upper and lower specification limits for one sided specifications one sided process capability ratios can be utilized. So, in this case it will be basically the upper control limit minus the mean value or the mean value minus the lower control limit. So obviously, in that case it would not be divided by 6 sigma now because 6 sigma was the overall coverages.

So, it now will basically be divided by 3 sigma if it is upper control minus mu divided by 3 sigma you want to find out the process capability ratios or PCR and in the other case if you are going on from your side you are going to the left. So, it will basically mean minus the lower value divided by 3 sigma. Now one should remember this 6 and 3 which we were using time and again would change basically to 4 and 2 depending on the overall level of confidence which we have for the problem if you want to cover an area of 95 percent.

In the case we are basically certain that we will cover about 67 percent of the area then in that case the process capability ratios if the whole coverage was done then the division would basically be 2 sigma. In the case say for example, we are taking only one side. So, it would basically be divided by, but the ratio would that now become basically the numerator would be the upper control minus the mean value divided by sigma and the on the right hand side from for my case if you are looking from your side is the left hand side for the normal distribution. It would basically be the mean value while you minus the lower control limit divided by either 2 sigma or sigma depending on how you are basically frame the problem.

So, you are basically. So, this will be one sided process capability. So, you are required to find out the process capability for one sided for the continuous bursting example.

(Refer Slide Time: 27:42)



So, suppose the lower specific limit of the bursting strength is given by 200 pound per square inch. So, you will use the x bar as 264 s as we already known is 32 and based on that when we find out the values of the process capability ratio, for a level of one sided level of process capability ratios depending on the overall coverage as 6 sigma then it gets divided by 3 sigma. And then sigma is replaced by sigma hat because this is not known the value comes out to be 0.67. The fraction on the defective containing produced by this process is estimated by finding the area to the left of z.

So, this is what I am saying, saying when we find out x based on that x when we convert into z we always use this formula which is standard normal deviate; so based on that if we find out minus 2. So, under this standard normal distribution the estimate fall is about 2.28, defective or about 22800 nonconforming containers per million. So, based on that you can find out that in a in a one million or 10 million or one thousand or 10 thousand or one lakh, then what would be the number of such defectives.

So, note that if the normal distribution were and a approximation model for the strength then the last calculation would have to be performed using the appropriate probabilities distribution. And we will basically do the calculations accordingly and let me mention one time one thing again. The mean and the standard deviation have to be calculated depending on what is the actual probability distribution for the population based on that we again I am repeating please forgive me for that and it will basically be the central limit theorem coming into action then you have to find out what is level of confidence based on that we replace them and do the calculations accordingly.

With this I will end the 30 second lecture and continue the rest 8s such that the overall coverage of the course would be done within the time frame of forty lectures. Have a nice day.

Thank you very much.