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

Lecture – 23
Variable Sample Size in X bar and R chart

Welcome back my dear friends; a very good morning, good afternoon, good evening to all of you; I am Raghunandan Sengupta from IME Department; IIT, Kanpur and this is the total quality management; one which is under the NPTEL MOOC platform and today would be the twenty third lecture. And if you remember; the last twenty second; we started about the concepts of control charts and it will mention that we will consider the movement or the variations of central tendency and try to study how control charts can be used to study the mean and its deviations.

Obviously the central tendencies are; few others are the median and the mode but we would not consider that. Then we also considered that variability or dispersion and rather than considering is the standard deviation of the variances, we consider the range; we choose the differences between the maximum and the minimum. One thing should be remembered which I did mention time and again; the graphs which were shown, whether with respect to the normal distribution; considering that central limit theorem is true which an overwhelming fact, we can use that to find out the central tendencies and the variability is in the concepts for the range. And if you also remember that; I did mention time and again that we will pick up samples of size n and we will do it m number of times.

So, if there are 20 observations in each group and we consider such groups 30 number of times, so it will be 20 into 30; such total set of observations. So, we will try to basically find out the best predictor, best estimator or best proxy of the population mean; coming from the sample mean, such that you will be taking the averages of the sample average so; that means, the averages of the average that is why, if you have seen it was x double bar.

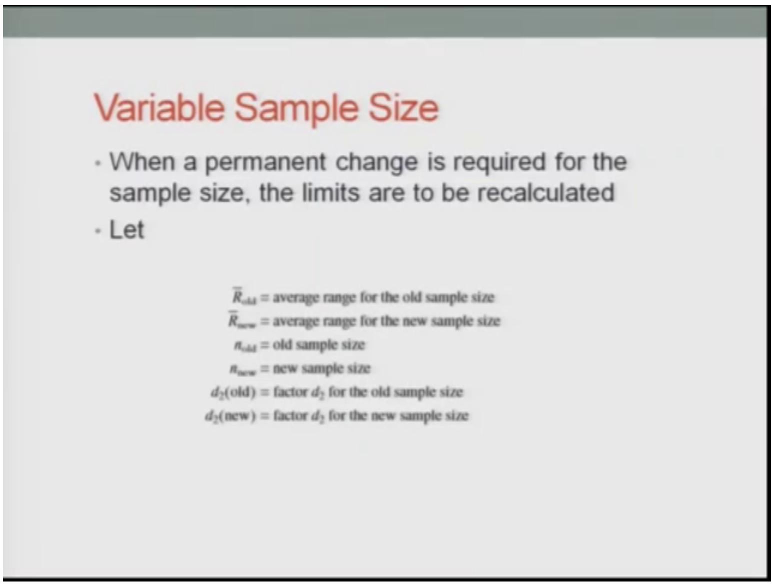
So, it that was symbol was used that was the central line and the variability; I am not talking about the range, I am not talking about the standard division; the variability or the dispersion of the mean value plus minus over the central line, below the central line would be decided by what amount of dispersion which you like for a mean values itself.

So, if you remember we considered the concept that x double bar plus minus some values would be given. So, those values would be basically dependent on the concepts of the standard deviations and some values of a or b. So, those who are they would depend on the sample size.

Later on we considered the concept of variability as dispersion; which is actually we want to stud study the standard deviation, but we consider the ranges. Range was basically the difference; for any sample that sample means the number 20; which you have taken. So, we have found out the difference between the maximum, the minimum for any sample and then averaged it out for all the 30 such set of observations we took and found out R bar and correspondingly we also found out; what was the deviation for that range plus minus and based on that we drew the R charts also.

So, continuing that in our 23rd lecture; so, when variable sample size would be considered.

(Refer Slide Time: 03:57)



So, when a permanent change is required for a sample size. So, may be that the production is decreased or maybe you want to take more sample sizes to understand how the overall distribution characteristics are changing, because if you remember this point was haft timed again and during the initial part of this TQM 1 course also; when we started the concept of the control charts.

The fact remains that they may be some observations which you want to take and they would some white noises or noises. So, some would be controllable some would be uncontrollable; uncontrolled means they are random; obviously, they would be there that is why the variability is coming.

But if there is a change in the variability due to problems happening; say for example, due to humidity, due to temperature change, due to machinability problems, due to the error from the point of view of the work man or due to raw materials from being a problem or being the coolant which is being used to cool as the sense the machine works there is a problem. Or say for example, the different environment factors which are there; if there are fluctuating which can be control, if they are fluctuating; obviously, it will have an effect on the readings and that would basically be coming out from the readings from the x bar charts and the R charts.

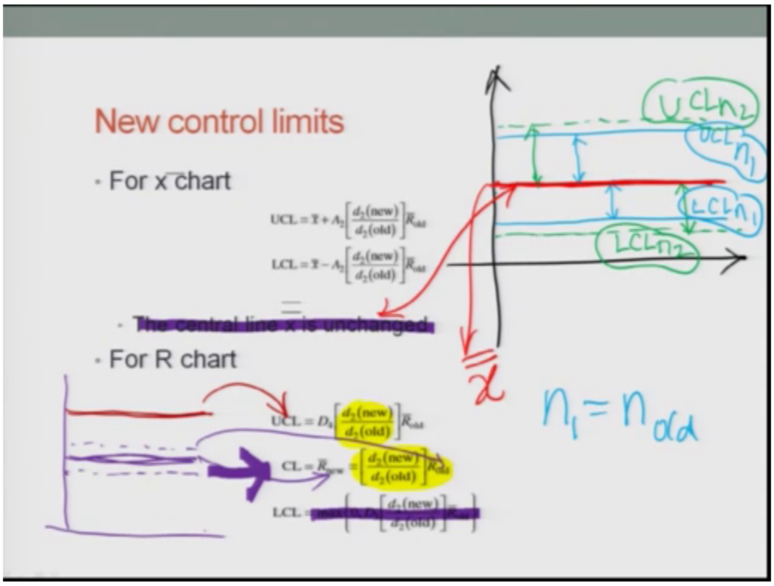
So, when again continuing what is written there on the slides when a permanent change is required for the sample size; the limits are to be recalculated. So, say for example, we know that for the R charts; there is a old average R bar suffix old which means the average range for the old sample size; whatever the sample size it consider that is n 1.

Now, do not confuse this n 1 with the n 1 and n 2 we have considered for the single sampling, double sampling and the sequential sampling plans. So, consider this is n 1 and R bar new is the average range for the new sample size which we consider as n 2. So, n 2 and n 1 are distinct and here it is basically mentioned as for our R is of understanding is n suffix old and n suffix new; which is the old sample size and the new sample size respectively.

So; obviously, when you find out based on the tables which we saw; that means, on the left column; you had the sample sizes, what was n and corresponding to that you had different values of a, d, b and so, on and so forth based on which you can calculate the ranges for the x bar charts, ranges for the R bar charts so on and so, forth. So, d 2 old would be the factor which you will be utilizing to multiply to find out the so called ranges for the old sample which is m n old and d 2 new would be the factors for d 2;

which would be used for the case when the new sample size which is n new which is being utilized.

(Refer Slide Time: 07:05)



So, in this case when you are using the x bar charts. So, that will be the upper control limit UCL and LCL are the upper control limit and the lower control limit. So, in that case we need to find out the new UCL and the LCL; it will again be x bar remains the same in both the cases; in one case as you know it will be plus A 2 and another case it will be minus A 2. So, the ratios of the d 2 and d old would be taken to multiply it by R bar; in order to find out basically the new control limits.

Now, one important things to be remembered that x double bar which is the central line let me come back to it; let me use the black color; it will be easy for me to specify. So, this is the X axis, this is the Y axis; Y axis this the X axis and you have the central line as given. So, this is actually what we mean by x bar bar and when I calculate the limit. So, this is the upper control limit; in the first case lower control limit in the second case.

Then as the sample size changes; that means, due to the fact you will have d 2 new and d 2 old; the values can be like it has gone up this is gone down. So, this would be the range in first case; upper range, this would be the range for the second case for the new sample size. And the corresponding values, if you remember this would be the case for the upper control for the old sample and this is the lower control for the new sample. So, we will say this would be the upper control limit for n 1; so, n 1 basically I mean n old.

Just a nomenclature do not we confuse too much about this; this is the lower control limit n 2, when I come back to this; this would be the lower control limit; sorry my mistake, I should change it. I should change it this becomes; so, when I come to the old one; this will be the upper control limit for n 2. So, this would go along with this and this would go along with this. But the important factor remains; the central line which you have is this one remains same as mentioned here.

So; obviously depending; so, another fact you may be thinking why I have drawn the green line which is the UCL; n 2 and LCL; n 2 above UCL n 1 and LCL and below LCL n 1 respectively. So, that would depend on the ratio; so, you say for example, if d 2; new by d 2 old R is greater than 1 or less than 1. Obviously, the limits for UCL and LCL would be accordingly done. So, depending on how many more numbers samples you have taken or you have reduced the sampling that would basically decide on the facts of what are the control limits.

Now, when I am coming to the R charts; so, again in the R charts; obviously, the values of the coefficient of the factors which we used to multiply depending on the sample size would also now depend on n old and n new; which is basically n 1 and n 2 as I am using the example. So, if I highlight it now; I highlight using the highlighter to basically make sense. So, the diagram which I had drawn here is specific to the x bar charts.

Now, I will try to highlight it using the yellow color. So, this would be the upper control limit based on the fact that what is the ratio here. So, this ratio is it would exactly be not the value would why would be same; but it would be the same concept as you have used for the x bar charts. So, in this case when you are using d 2 new and d 2 old; here also you are using d 2 new and d 2 old. But the values why I am saying they are not exactly the same would depend on the sample sizes; which are changing.

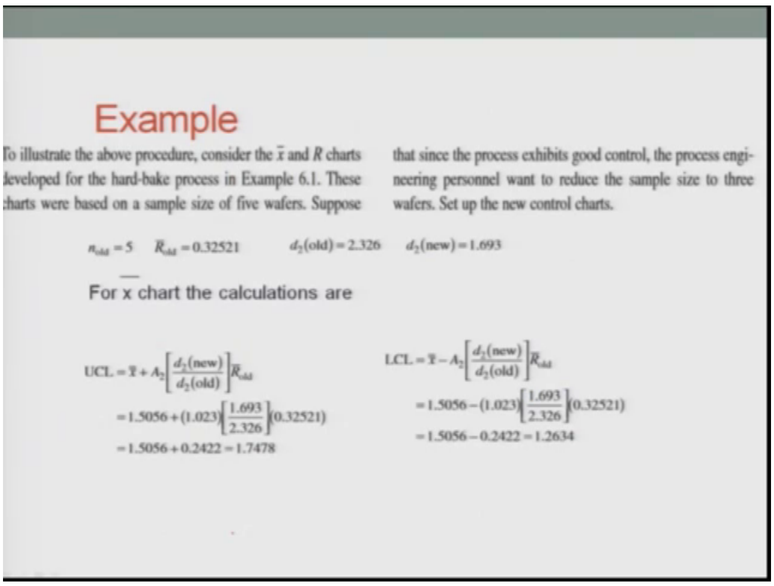
So, once I have the central limits. So, R new would be based on the fact of what is the factors which is being multiplied by R old. And similarly the lower control limit would basically be; lower come to where you are going. So, it can be maximum of 0 or some positive value; it cannot be negative because range cannot be negative because I am trying to find out the difference between the maximum and the minimum.

One very important thing to note down is this; so, I had use a highlighter of a different color. Here; it mentions for the x bar charts that x double bar does not change. But if we note down here where I am trying to basically highlight using the violet color with an arrow; it says the central line for R bar new and the R bar old will be different depending on the sample size difference; that is the ratios. So, that is why I said; the values may not be equal and; obviously, the upper control would be the values based on the ratios of d 2 new and d 2 old; similarly for the LCL.

Only one fact remember here; it will be the maximum value of 0 and the value which are getting based on the new calculation. Because again I am mentioning the lower control limit cannot be negative. So, in this case if I draw let me use diagram accordingly. So, if I draw here this value which is there; which was basically R new would; obviously, be different either above or below R old. So, this becomes R old, this because R new and the upper control limit would be some value which I will try to utilize a different color amongst this and try to use the brown.

So, this is brown not dark brown or not red. So, this would basically go to the upper control limit and the lower control limit; I am again using the same color. So, lower control limit cannot be 0. So, this is the 0 line; so it cannot be negative that is why it says it is the maximum 0 whatever value you have.

(Refer Slide Time: 14:22)



So, to illustrate these example but the above procedure consider the x bar and R charts developed for the heartbreak process which was there in example 6.1. So, again I am please; please requesting all of you, have a look at this book; this is a fantastic book

Montgomery, it may be little bit in depth but it will give you a huge amount of information based on which how you can use the control charts for the statistical process; control and understanding total quality management.

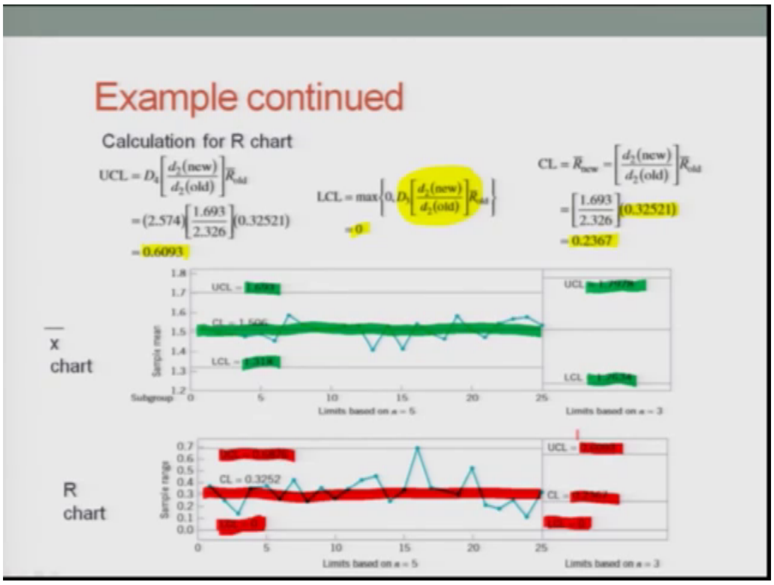
So, those charts or base; so these chats or base initially for the sample size of five wafers. So, if you remember you were taking five wafers and some number m was given. So, n old was five based on that we find out R bar old; that means, for the five was about 0.32 and the d value for the wafer sample size being 5; it was about 2.326. So, I will just mention the second place of decimals; so, that should be clear to you.

Now, when we suppose that since the process exhibits goods good control the process engineer wants to reduce the sample size and want to now take three wafers now n is 3. So, number of m can be same so; obviously, as n changes then d 2 becomes 1.69 about approximately 1.69. So, you have to basically calculate; so, once you do the calculations based on this. So, first let us consider the x bar charts; so, x bar charts, x bar or double bar; obviously, remains the same. So, let me highlight using the color yellow.

So, this value remains as 1.5056 as I am highlighting; but when I come to the value calculation of the ratios of the plus and minus limit. So, they would basically be a ratio of new toad; which comes out to be less than 1; because it is 1.69 divided by 2.36. So; obviously, the values which you calculate accordingly now is 1.74 and 1.26.

So, when you are considering the change of the values of n old to n new; which is from n 1 to n 2, you will have to basically recalculate the upper control and the lower control. Basically, find out that how the points are basically now fluctuating within the changed LCL and the UCL.

(Refer Slide Time: 16:59)



So, now I get let us consider. So, now we in the next step; obviously, you have to calculate the values of R bar values; R bar will change upper control would change. So, again when we consider the upper control and lower control for the R charts. Again I am highlighting using the yellow color; so, the upper control limit becomes about 0.6093 and the lower control limits basically because this is maximum is 0 and this value.

So, this value if you note down which is I am highlighting here; obviously, if it is maximum between 0 and that value; it has to be negative that is why it is a 0 and the central line you have to recalculate. So, R new basically with respect to the old; so, R new becomes 0.237, 2367 on the R old is 0.3251.

So, when I basically convert into the tables of the chart. So, let us go and see it; so, I use the green highlighter for the x bar charts and the red highlighter for R bar charts. So, for the green one I am use using; so, where I am hovering the pen. So, this is the central line which is x double bar which remains the same. And the upper control limit values are at calculate R 1.693, 1.318; but if you consider the values in the old sample; so, those 1.79, 1.26.

Now, if you see they have reduced; either can they expand at reduce; obviously, you can find out that as they were within limits. So, shrinking and making it more type would not effect because; that means, you are spending less amount of money by taking three observations in same sample, still you are able to meet the requirements. Technically means that your total cost was sampling has decreased.

Now, when I go to the R charts; as I said I will be using the red highlighter. So, here this central line; it is now the values of the central line is 0.236, where I am hovering because this yellow color I am not going to change it. So, 0.2367 is basically coming from here and the actual central value initially was here; when I am again highlighting the pointed what was 3.32521; which is here and the actual values which I calculate using the upper control and the lower control values; was 0.6093 and 0 and in the initial cases these values were 0.68 and 0. So, this 0 value remains the same; obviously, the upper control values changes.

Now, what you are do doing in this cases is 2 things; number one we know that the number of observations we are taking for each sample has to be reduced; because initially we saw and I did not mention it. But we have seen the slides for the case when he had drawn the x bar charts for the case when you are taking five observations as in each set; the overall width of the upper control limit or the lower control limit were quite wide. In sense that what the calculation gave; but the values or the dots or the each sample observation which we took or within a very small range, so; obviously, decreasing the upper control limit and the lower control limit would be beneficial because as I said cost reduction is there.

Now, if you want to basically compensate that with trying to find out whether these variability in the each sample. Now, I did not mention this point very explicitly but one thing should be understood; that when you are trying to draw the x bar charts and the R bar charts. So, the x double bar is the average sample of the sample which is the averages average; you are taking. So, you had a sample of 5; you are collaborating everything and trying to find out the average of the whole group.

And in the R bar charts very interestingly it gives you the specification for each sample. So, technically the R or the range you are trying to find out the variability of the dispersion for each sample and trying to find out what is the average and what is the dispersion. But when you go into basically trying to understand the x bar charts, here what you do is that you try to find out the variability or the difference in the averages of each sample with respect to the averages of the averages technically which should be actually the populations.

So, what I am trying hint is that the value again I will not highlight; but I am trying to hover my pen. This value of the central limit value which you have; which is 1.506 or 1.779; technically in the long run if you take more and more observations; you should actually go to the new value. But here we are not interested basically to actually attain that efficiency because the cost would be very high; that is why we had taken only three observations per set.

So, these two charts which are there in front of you; with respect to the charts which we are drawn considering five wafers in each sample, would give you a feel that whether than sample sizes should be decreased.

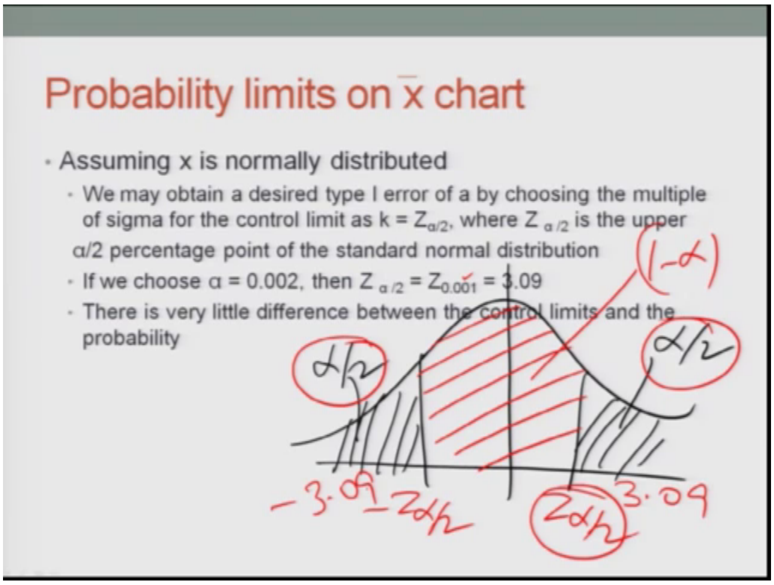
Another important point; when you had drawn the values of the R charts. So, the variability would give you some points which are coming out of the range. So; obviously, you have to take some decisions that variability for any sample observations we are taken whether they have exceeded limit or the ranges are much more than what is they expected.

So, that may be due to some assignable causes or due to some unassignable causes. So, for that you have to basically do an in depth analysis and try to analyze or try to study where the problem actually lies and if at all there is a problem; because the variability would definitely be there. Like when you are tossing a coin; they would be head and a tail, when you are rolling a die they would be number 1 or 2 or 3 or 4 or 5 or 6.

So, you have two take a decision accordingly where the viability is actually happening due to the random events; which are not under your control or whether the variabilities happening due to some events which are controllable but which are going slowly out of hand. Like as I mentioned; the coolant being not used properly, the humidity is high, the temperature is high; machine ability is not proper; poor workmanships and all these things.

So, now you have to basically understand something about the probability limits on x bar chart.

(Refer Slide Time: 24:02)



So, control limits in the charts are often expressed in terms of multiples of standard deviations. So, now this will become a little bit more interesting; if you remember I will speak a little bit more here without going to the slides immediately and I will basically sketch some diagrams accordingly.

Now, if you consider the dumbbell shaped curve which is the inverted bell of for the normal destruction and if we remember in one of the slides in the last class and previously also; we considered there is a mean value which is a mu and there is some standard deviation. Now, and these normal distributions are basically coming whether the distribution is normal and even if the distributions are not normal, we can consider the central limit theorem to be true and hence we use this.

Now when we convert into a standard normal deviate; so, I am saying would be true for the case when we convert into a standard normal deviate or whether you keep the mu value and the standard deviation as it is. Now, if you go plus 1 sigma and minus 1 sigma; from your side if you looking plus would be on my left which is your right and minus sigma would be on my right which is your left.

So, if you consider that the total area coverages is as I mentioned again; if it is about 67 percent; I am using approximated value to be 67 point something. Then if you basically make the divisions plus 2 sigma and minus 2 sigma, it will be about a 95 percent and if you basically make it plus 3 sigma and minus 3 sigma; the values would be about 99 percent.

Now, this is where the important come. So, depending on how good or bad your limits are; you will basically try to multiply those factors; if you remember the factors which you had discussed like a, b, d and all these things are depended on the values on sample size which is n. So, that will also give you an information that what excuse me what is overall breadth off of the values of the upper control limit and the lower control limit you are trying to keep.

And that will also give you how good or bad or a how efficient and robust your overall statistical process control techniques are which would definitely becoming out from your process control charts. So, continuing this; we are also able to define or possible to define the control limits by specifying the type of errors. Now, I will pause here if you remember when you had done the OC curves; that alpha, beta type 1 error, type 2 error, the level of consumer risk producer risk; it basically means that there are some good items which are being rejected and there are some bad items; which are being accepted.

So; obviously, it is basically a type of risk which would be positive; if you consider both the risk, it will be positive for one case for the producer and negative for the producer also. But when I basically a look at the picture from the consumer sides then the producer risk; positive sense would become my negative sense risk and the negative risk for the producer would become a positive risk for me.

So, the level of risk would basically be decided accordingly. So, it is also; as I had mentioned is also able to define the alpha and beta values. So, these are called the probability limits of the control charts; which will basically continue and study as we proceed with the course.

So, now considering x is normally distributed. So, we may obtain a desired level or type 1 error or by choosing the multiple of sigma's of the control value. So, now if we consider Z as alpha suffix alpha by 2; it means that that the value the dispersions which we have; let me basic it highlight and use the black color. So, this is the distribution which you had; these are the values which I marked. So, this is basically alpha by 2, this is also alpha by 2 and overall area which I have. So, let me mark with the red pen; this area I have is basically 1 minus alpha.

Hence the values which I have on the left on the right which is Z alpha by 2 and plus Z alpha by 2 would basically denote my limits. So, reading continue reading; that I will come to that later on also in more details. So, we may obtain a desired level of type 1 error by choosing the multiple of sigma for the control limit.

So, which is Z alpha by 2; so, that the coefficients where Z alpha by 2 is the upper alpha by 2 percentage; which is here. And similarly we will basically have Z minus Z alpha by 2 on the left hand side. So, if we choose alpha by as about 0.002; so, this would be 0.002 by 2 and similarly this 4 value would be. So, we get the values of Z; 0.001; which is given here as about 3.09 here as minus 3.09 here and we do the calculations accordingly. So, with this I will end this lecture; I am not abrupt in ending. I will continue in this discussions in the subsequent lecture; this is the twenty forth and the twenty fifth one.

Thank you very much for your kind attention; have a nice day, bye.