Total Quality Management - I

Prof. Raghunandan Sengupta
Department of Industrial and Management Engineering
Indian Institute of Technology, Kanpur

Lecture - 28
Np control chart

Very good morning, good afternoon, good evening my dear friends: welcome back to this TQM one like set of lecture series under NPTEL MOOC. And I am Raghunandan Sengupta from the IME department IIT Kanpur.

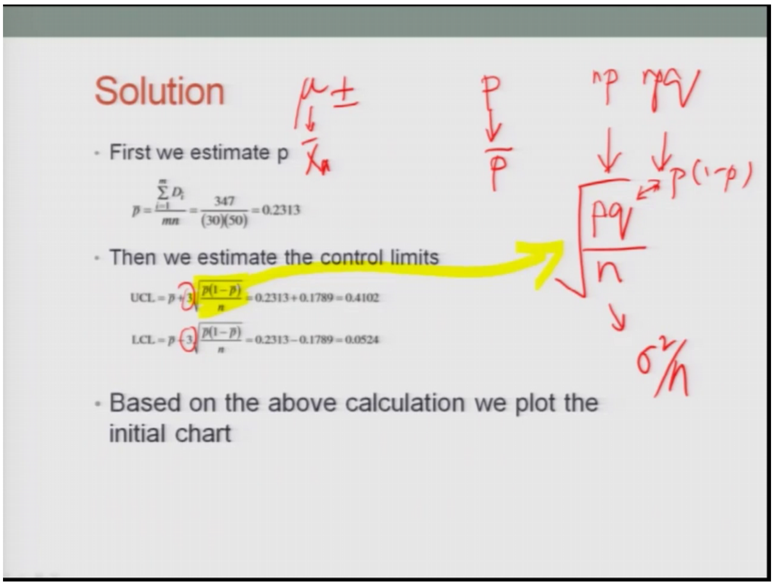
So, we are starting the 28 lecture, and as you remember we are discussing some problems related to the charts, charts being x bar charts, r charts, then which is the r charts the range charts. And then then if you do not want the range you can definitely replace that with the standard deviation charges provided standard deviation is known for the population not known. And there are different combinations you can find out and find out the estimates. Use those estimates to do the calculations and use the tables from where you can find out those coefficients of a 1 a 3 b 1 b 2 b 3 c 1 c 2 c 3 and so and so forth. And if you remember that where they are the fact part we did mention apart from the charts that how the q q plots or the quantile quantile plots could be done.

So, quantile, quantile plots are q q plots? That point came up because if you remember that when we are discussing the charts, we did mention about the chi square chart, tables, or chi square distributions. So, those come up when you are trying to find out something to do with the concept of variance for the population, provided the variance of the sample is used to find out the best estimate. And then I also did mention even though we did not solve we mentioned about the f tables or the f distribution based on the fact that you want to find out, something do with the ratios of the variance or variability.

Like as I mentioned 2 products are being produced same type in 2 factories in factory one and factory 2, in factory one the variability is very low factory 2 the variability is otherwise either is much less or much more and then you want to compare them. So, you can compare them using the ratios of the standard deviations or the populations for which you replace the population standard deviation or the variance with this standard error or the variance of the sample and do the calculations. So, there you will use the f distribution with certain degrees of freedom.

So, those loss of degrees of freedom would also come up in the discussion as you study those tables and the and the distribution.

(Refer Slide Time: 02:46)



So, first as you remember, we want to find out the proportions. Defective if you remember that we whatever the distribution is binomial Poisson, considering the central limit theorem to be true we replace that particular distribution with the normal distribution, but the fact remains that the normal distribution mean and the normal distribution standard division would be exactly to do with something with the actual population distribution based on which we are trying to draw some inference.

So if it is binomial, binomial would have a certain mean certain standard deviation. So, those will be utilized to find out the central line and the upper control limit lower control limit for the normal distribution which is now, a replace, so called replacing the concept of the binomial distribution considering central limit theorem to be true. So, first we estimate p which is the proportion. So, this is in a problem we found out the total proportion was 3, 347. And in each chunk you are taking 50 observations and such if 30 observation sets of observations you took. Obviously, the overall proportions would be 347 divided by 1500 which comes out to be 0.2313.

Then we estimate the control limits. So, then the control limits again, I am coming to the fact. Let me try to highlight it in a much better way. So, what you want to do is that this p has been replaced by p bar or p hat or whatever you say which is the average value. And if you remember for the binomial distribution, we had taken the value of say for example, the mean and the standard deviation being given n p, and n p q. Now the question comes that once you this replace this corresponding values for the population with the sample; obviously, if you remember we had divided for the normal distribution the variance with n, and basically took the square root of that and solved the problems accordingly.

So, in the same way when you basically replace that using the central limit theorem and try to basically use those formula, it basically comes out to be n p I will write the p q or technicalities p in to 1 minus p. So, the both those things are same divided by n. That the square root of that which means that basically it has something to do with sigma square by n are the square root of this is whatever you are considering.

So, now in the normal distribution what we saw? It was the mean was mu this was replaced by x bar n plus minus; that means, if you are looking from your side the central line is the central which is x double bar the upper control limit will be on this side which is much more and the control limit with the lower value. So, if you want to find out the dispersion. So, so called dispersion over and above number over and below the mean value which is x double bar. So, you have to add some coefficient is basically a fixed number depending on the overall range which you want to have.

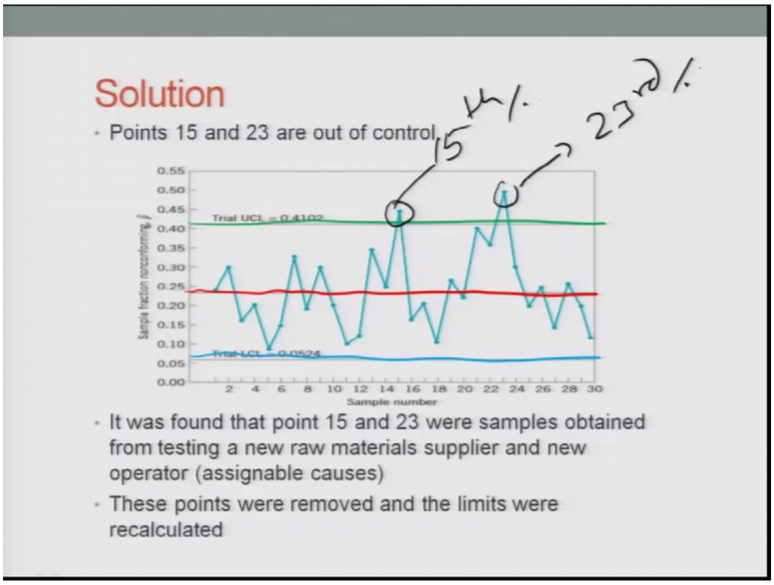
So, if it is say for example, 1 plus minus sigma, it will basically mean that overall coverage which you are doing would be about sixty 7 percent. If is plus minus 2 sigma; that means, you are going plus 2 sigma above and minus 2 sigma below it; obviously, the total coverages would be about 95 percent and if it is plus minus 3 sigma. So, the overall coverage would be above 99 percentage.

So, if you see the values of 3 which is given here again I am repeating that depends on the level of confidence which you want to have for the readings. And the values which I want to now highlight using yellow color this part, is basically exactly what we are talking about. In the case that you are taking the sample estimate in order to basically find out the, so, called control limits for the population.

So, again coming back the upper control limit becomes p gets replaced by p bar or p hat, which is the first information which is the so called central line we are aware. Next we have to find out the coefficient which is 3 in this case plus 3 and minus 3 depending on what level of confidence you want to have. Second is basically the so called standard errors square of the sample which basically comes out to be the case which is p into 1 minus p by n. So, if you want to basically find out the standard error it will be square root of that.

So, hence in one case it will be 3 multiplied by square root of p into q divide by n. And another case it will be the minus 3 into p into q divide by n. So, based on the above calculations of the central values which comes out to be upper control comes out to be 0.4102 and the lower control limit comes out to be 0.0524. You can basically plot the upper control lower control and the central line as follows.

(Refer Slide Time: 08:06)



So, the central line is this one, which I will just highlight using the red one, without trying to disturb the diagram. So, this is the central line. And I will try to draw the upper line using green one. So, this is the ucl which is happening which is 0.4102. And the lower value I will use the blue one. So, this is 0.0524. So, these are the values. Between them you plot then you can find it there would be some values which are above and below the control limits which is this one, this one and so and so forth.

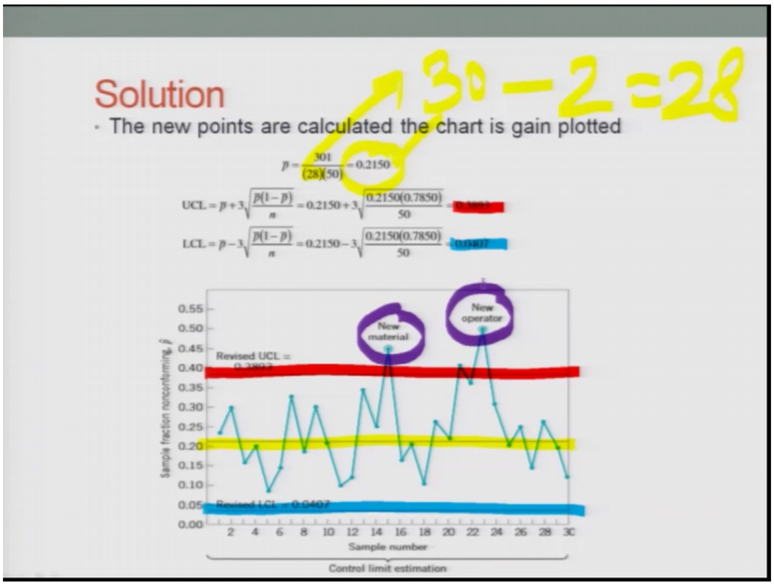
So, if you see the solution and try to solve it, this 15th one and the 23 one readings are out of control. So; obviously, there are some problems related to assignable cause unassignable cause and you have to take corrective actions. See it was found that 0.15 and 23 was samples obtained from testing a new raw materials; that means, if the raw materials have come may be the moisture content is more, may be the density is more, may be the specific gravity has changed, may be say for example, the (Refer Time: 09:29) is different. So, based on that be there may be some reasons that why those 15 under 20 third readings where basically out of control.

So, continue let us continue reading it is mentions that 0.15 and 23 were samples obtained from testing a new raw materials supplier and new operator could have been also be there. So, say for example, raw materials is fine. It could have been like a new vendor is supplying it or may be the new vendor is sending the same product, but due to some mishandling the products are getting destroyed. So, we were discussing that assignable and non assignable cause. And the reason could have been apart from the raw materials apart from the vendor say for example, thee vendor sends the materials, but the truck in which is coming has some dust particles. So, this this could have affected the raw materials, other can be the operator who is working is new or is level of confidence on working on the machine is totally different from what was done by the other operator who used to work before.

So, all this things could have happened which would have affected the 15 and the 20 third reading and based on that we can definitely say they are out of control. So, if in case if you want to be very sure that what are the reasons; obviously, you have to take corrective actions, but for the readings you will remove this points and the limits would be recalculated. So, in case the 15 and a 23rd reading are removed. So; obviously, you will removing 2 bunches, but the point was also remain that each such individual set of observations would have some sample numbers so observation numbers. So, you have to basically remove those also which means that if the 15. And say for example, 20 observations and 20 third also have 20 observations. So, forty observations would be removed and the calculations for trying to find out p bar or ucl with respect to p bar lcl with respect to this p bar which is the new p bar I am talking about would (Refer Time: 11:37) it done accordingly and the calculations done like wise.

So, in the calculation, if you remember in the last 2 last slide the denominator had the numbers m and n m was 30, n was 50. So, if I remove 2 of them; that means, it will now 30 would become 28. So, in in this case, but the samples size in each set of observation always remains 50. So; obviously, your total set of observations initially it was 1500, now it will be decreased by 100, because now it becomes 1400 and you do the calculations accordingly

(Refer Slide Time: 12:18)

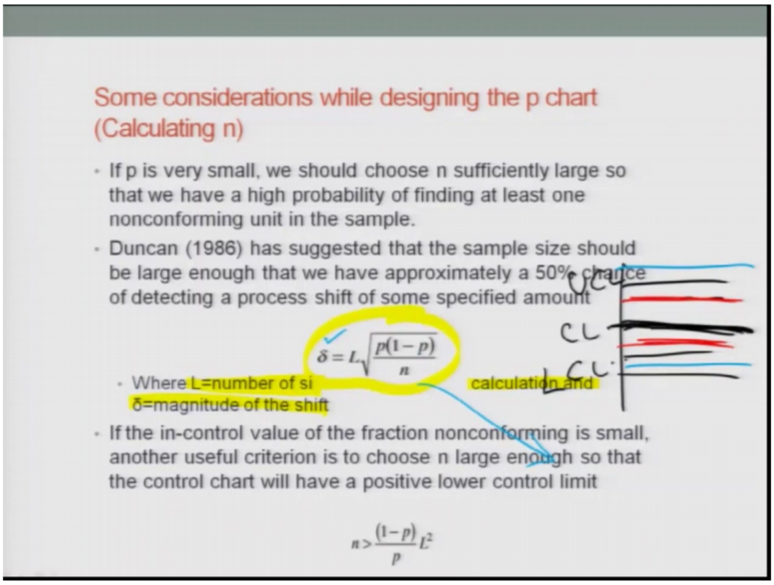


So, as shown here the new points are recalculated. So, in this equation let me highlight this becomes 28 this was initially 30. So, you talked out 2 observations. So, it will be 28 here, plus when you find out this p bar it will become 0.2150, which is this line where I am highlighting using the yellow color. When I calculate ucl which is the upper control limit again the coefficient 3 remains p bar is 0.2150. And 1 minus p bar would be accordingly 1 minus 0.2150. This n remains same yes that is important to remember, but because the chunk of observations how many such chunks you took 30.

So, you remove 2 2, but the number of observation each chunk remains 50. So, that 50 or n is not going to change based on that I (Refer Time: 13:21) is calculate the upper control limit is 0.3893, which is this line. Then the lower control limit again you find out it comes out to be let me change the color for ease of understanding, 3 comes to be this line which is 0.0407. And if you find out the calculations, it will be a very evident. So, let me use another different colors, but now say for example, violet.

So, this is the new material or new operator or whatever the reasons are you can you will be able to find out in a much better way and basically comment intelligently that were the reasons for which these aberrations occurred.

(Refer Slide Time: 14:07)



Some consideration while designing the p charts, is basically you have to find out the actual n and m based on which you will do the calculations. P is very small we should choose n sufficiently large. So, that we have a high probability of finding at least one nonconforming unit in the sample.

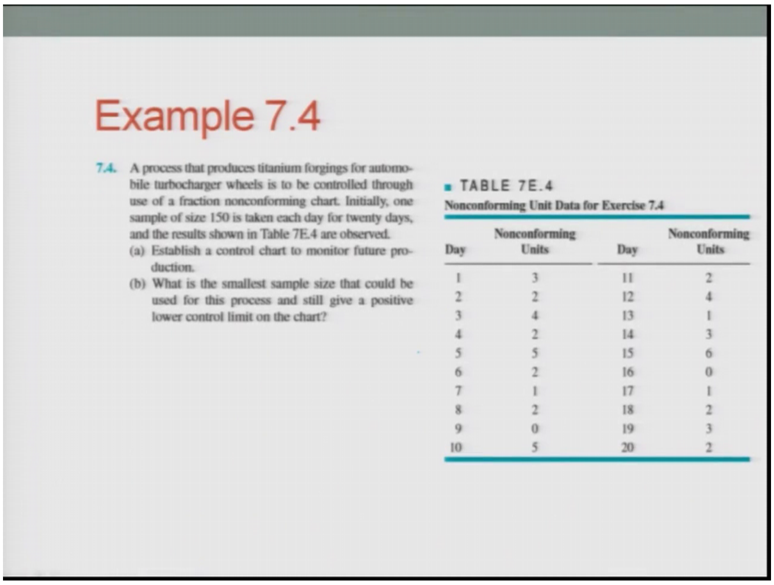
So, if the actual calculation is seeing there the proportions is very small. So; obviously, we would take a large chunk in order to find that. So, in one of the papers very famous is Duncan has suggested that the sample size should be large enough that we have approximately above 50 chance of detecting a process shift of some specified amount. So, based on that you will find it out and do the work accordingly. So, this is the delta. So, l is the basically the number of si or a calculations based on which you can do the calculations and delta is the magnitude of the shift. So, depending on the shifts; so these shifts are let me highlight. So, you have the central line this is the central line. This is the upper control line this is the lower control line. So, as the shift the central line remains the same only shift which you have let me use different colors to highlight that.

So, this movement or this movement whether moving positive or negative would positive means, they are they are moving more away from the central line and moving more down or they are coming closer would depend on the values of so, called p and 1 minus p and the corresponding values of n and based on that you can do calculations. So, if the in control value of the fraction nonconforming is small, another useful criterion which is used to choose n large enough so that the control charge will have a positive lower limit.

So, because 0 lower limit basically means that can be negative; obviously, the difference of the ranges we are trying to find out would not be basically applicable here. So, what you are trying to find is proportions of the probability of the chance or the relative frequency. Now relative frequency chance proportions probability would never may less than 0 so; obviously, they if they are 0 or little bit positive then it gives such a much better feel that why we are using the charts and we can understand through their readings how this values could be utilized.

So, if we basically put this is in the formula. So, if you square it so; obviously, it will be delta square n would go on the top which is n, and this becomes p into 1 minus p and l square comes and then you solve it to find out the value of n or a tentative value of n such that n value has to be greater than equal to that infimum value. So, here is a problem again from Montego marry.

(Refer Slide Time: 17:08)

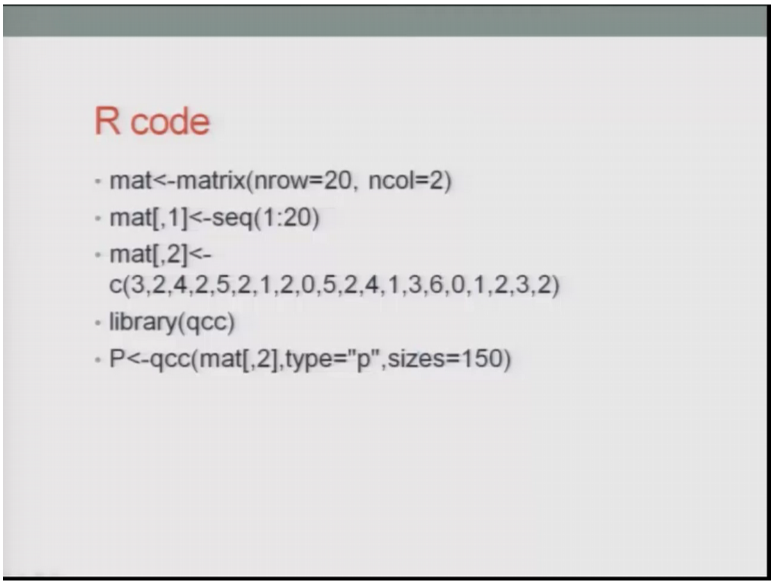


A process that produces titanium forgings for automobile turbo charger wheels is to be controlled through user fraction and conforming charts. Initially one sample of size one 50 is taken each day for 20 days. So, each you continue taking it and you continue doing it 20 days. And the results are shown in the table which is where I am pointing my palm.

So, here you have day one nonconforming units is 3. Then day 2 the nonconforming unit is 2 day 3 4 and so and so forth till the last 2 days which is nineteenth. Nonconforming number is 3 day 20 nonconforming is 2. So, now, with this information you have to establish a control charges to monitor future production processes, how it is going. And you also need to find out what is the smallest sample size that could be used for this process and still give us a positive lower control limit such that we can use those control limits judiciously in order to understand whether the process is in control or out of control. So; that means, more number of samples you take it is better is good that you will try to do the study much more thoroughly, but the flip side is there is a cost effect involved.

So, it is destructive testing; obviously, all the products are destroyed or if it testing where human beings machines man power are involved; obviously, there is a huge cost involved also. So, you have to basically be a charge accordingly such that you can and do the needful to find out the minimum value or the infimum value of n.

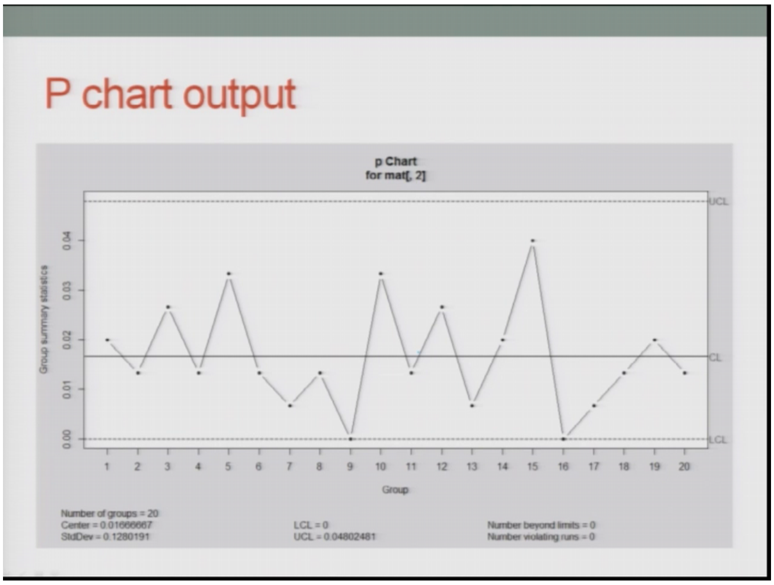
(Refer Slide Time: 18:45)



So, the R code is given. So, if I know the you basically have the matrix would be basically the number of days and the number of nonconforming.

So, based on that you put the number starting from 3 to 4, which is the 4 the third bullet point 3 2 4 till the last 2 beings 3 and 2. And you basically do the library function for the q, q plots based on that when you have the q, q plots.

(Refer Slide Time: 19:10)

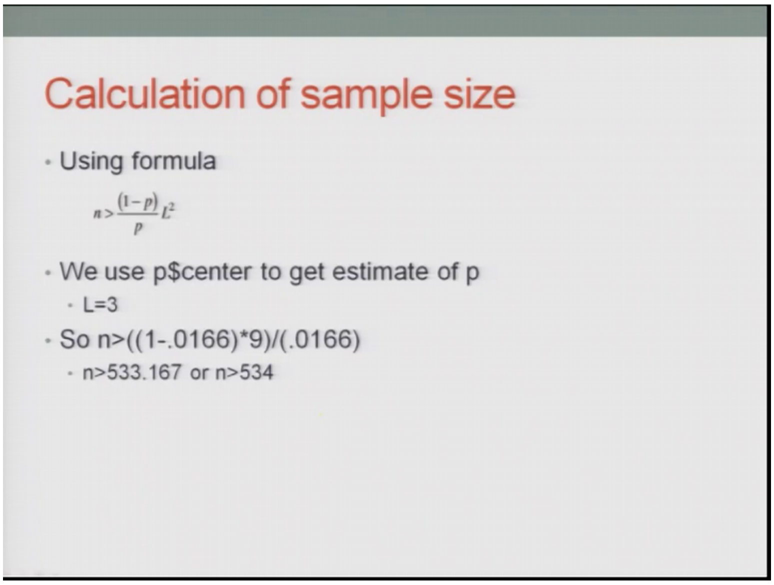


And you basically fin other tables. The central line in is given by where I am highlighting this is the central line this is upper control values are there lower control values are there and if you see the lower control values we are trying them not to be 0. So, if they are positive well and good, but; obviously, in many of cases even if you take a big sample size it may not be possible to find that for what sample size the value of the lower control limit would be greater than 0.

So, the number of groups which you are taken is 20 central line as marked by the blue pen. The value which I will try to highlight is this one is 0.0167. Standard deviation is 0.128 lower control value is 0 upper control value is 0.048, and if you want to find out if there are number of points above the central line or below the central line they are basically 0; in number because there are none of the points which are above here none of the points which are below here. Obviously, on the lower side none of the values would be lower there because they cannot be negative.

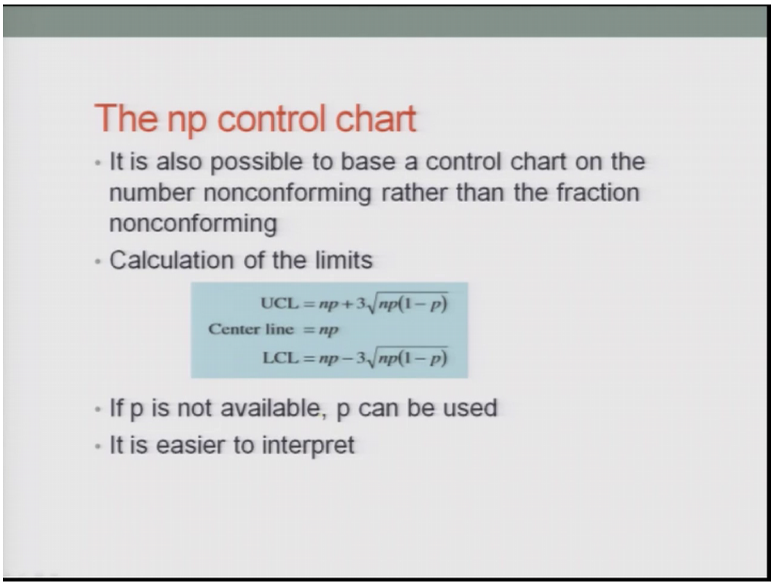
That is why when we keep repeating that the lower control limit should be any value greater than 0 it will give you much through understanding whether the control limits; obviously, upper values would be problematic, but if there are some values which are going to the other way round which is below the lower control limit. Obviously, it will it would give you some set of information how the process is doing. I am not saying they are out of control, but it will definitely give you some set of information.

(Refer Slide Time: 20:55)



Now, based on that when we do the use the formula to find out n which is the minimum sample size, we use it to find out n should be greater than equal to the maximum integer value which you can have, 1 minus p divide by p l square. So, once you put it in a calculation the value of n comes out to be about greater than equal to 533.67. So; obviously, 534 would be taken as the at the sample size for this study.

(Refer Slide Time: 21:29)



We will discuss now and pay attention to the np chapters. So, it is also possible to build a limit central line upper line and lower control line for based on the on the fact on the number of nonconforming. Till now we have done the proportions. Proportions means probability relative frequency chance, I want to find out what are the numbers. So obviously, if it your considering a very simply if p is the proportions are defective and if the total sample size is n, then the term number of defective would be calculated by p into n.

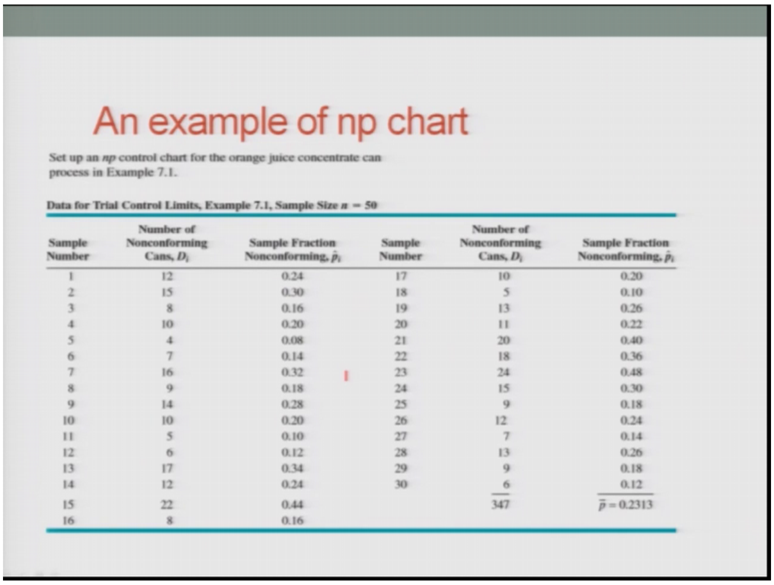
So, that would solve our problem, but that would give a some information in numbers rather than probability of chance would be much more comfortable to understand in numbers. That is why you are trying to find out the charts and the property of the charts the upper control limit lower control limit the central line based on the fact that you are using the numbers which is the sample size in the proportions are defective or non defective which ever we are looking at.

So, it is also possible to base a control charts on the number on nonconforming rather than the fraction conforming. So, here fraction is not important what is more important the numbers. So, you have the central line again comes out to be n into p actual central line. Obviously, if you do not have p you will replace with p bar so; obviously, the central line would be n into p bar. So, controlling finding out those upper control and lower control limit the formulas concept remains the same, but the formulas would change depending on what you want to find out in in the in this as in in place of the standard deviation again using the central limit theorem.

So, you have upper control as an np plus or minus for the lower control limit and the coefficient is 3. The square root value which you have initially was p into 1 minus p by n, but in this case you will basically be utilizing the fact that you want to find out the variance and the concept would be utilized for the fact when you are in interested to find out in numbers and not in the proportions. So, those values comes out to be np into 1 minus p that is square root and plus minus 3 of that would give you the upper value and the lower value.

So, if p is not available; obviously, you will use p. So, this would be easy for us to understand. So, the case is basically p hat or p bar. So, if p is not available p bar can be used and it is easier to interpret, because I said that numbers are much more easy understand then with respect to proportions of probability. Now let us consider our np charts.

(Refer Slide Time: 24:22)



An example of np chart

Set up an *np* control chart for the orange juice concentrate can process in Example 7.1.

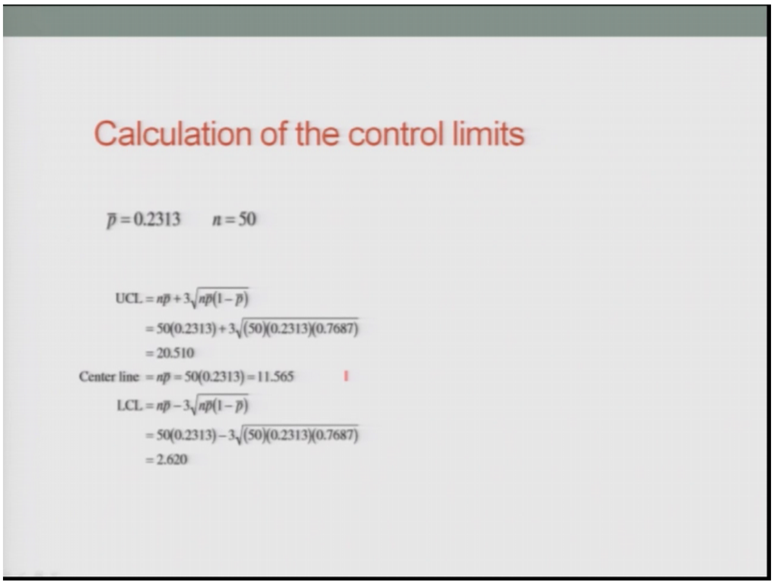
Data for Trial Control Limits, Example 7.1, Sample Size $n = 50$

Sample Number	Number of Nonconforming Cans, D_i	Sample Fraction Nonconforming, $\hat{p}_i$	Sample Number	Number of Nonconforming Cans, D_i	Sample Fraction Nonconforming, $\hat{p}_i$
1	12	0.24	17	10	0.20
2	15	0.30	18	5	0.10
3	8	0.16	19	13	0.26
4	10	0.20	20	11	0.22
5	4	0.08	21	20	0.40
6	7	0.14	22	18	0.36
7	16	0.32	23	24	0.48
8	9	0.18	24	15	0.30
9	14	0.28	25	9	0.18
10	10	0.20	26	12	0.24
11	5	0.10	27	7	0.14
12	6	0.12	28	13	0.26
13	17	0.34	29	9	0.18
14	12	0.24	30	6	0.12
15	22	0.44		347	$\bar{p} = 0.2313$
16	8	0.16			

So, what you have is the sample numbers which are given on the left most columns starting from one to 16. Number of nonconforming cans are given for sample number one it is 12 for 2 it is 15 and say for example, for the last 2 for the 15 sample number it is 20 2 and for the 16 sample number is 8.

So, sample fraction nonconforming would basically be found out from this this information sent. And they come out to be 0.24 0.30 for the first 2. And while the 15 then the 16 comes out to be 0.44 and 0.16. Now if I continue considering the sample numbers they are like 17 to 30 the number one nonconforming cans are basically are form 10 to 6 which is basically total is 347 and if you find all the sample fractions; obviously, they would come out to be 0.20 to 0.12. And here the p p bar value comes out to be 0.2313.

(Refer Slide Time: 25:27)



So, now n is 50 n is the sample size you multiply both of them find out the central values central value comes out to be as shown here. So, I just highlighting it. So, these you can find out at the central line. And the upper control and lower control would be found out by plus minus 3. And the standard division which you have in this case considering that you are using the numbers as such would be np into 1 minus p whole the whole thing is square root. Again they come out from the simple concept of central limit theorem. Then use those values to find out the upper control and the lower control this upper control comes out to be 20.510. And the lower control comes out to be 2.620.

(Refer Slide Time: 26:13)

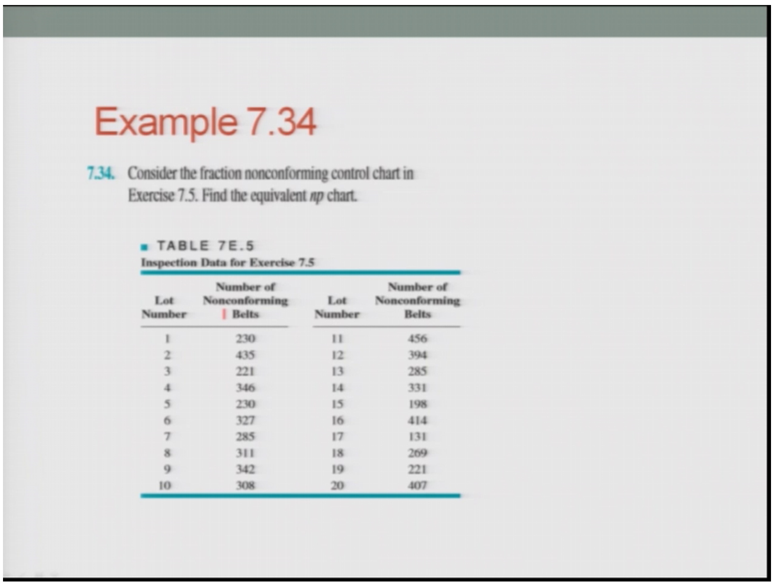
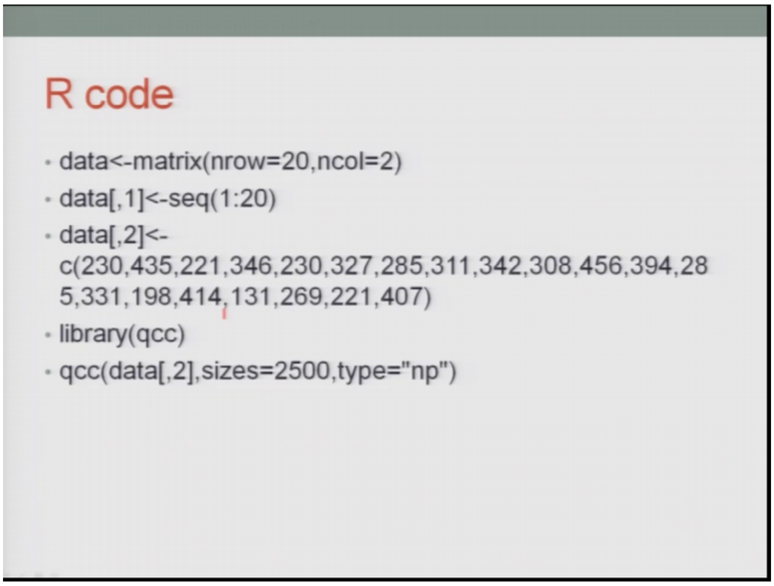


TABLE 7E.5
Inspection Data for Exercise 7.5

Lot Number	Number of Nonconforming Belts	Lot Number	Number of Nonconforming Belts
1	230	11	456
2	435	12	394
3	221	13	285
4	346	14	331
5	230	15	198
6	327	16	414
7	285	17	131
8	311	18	269
9	342	19	221
10	308	20	407

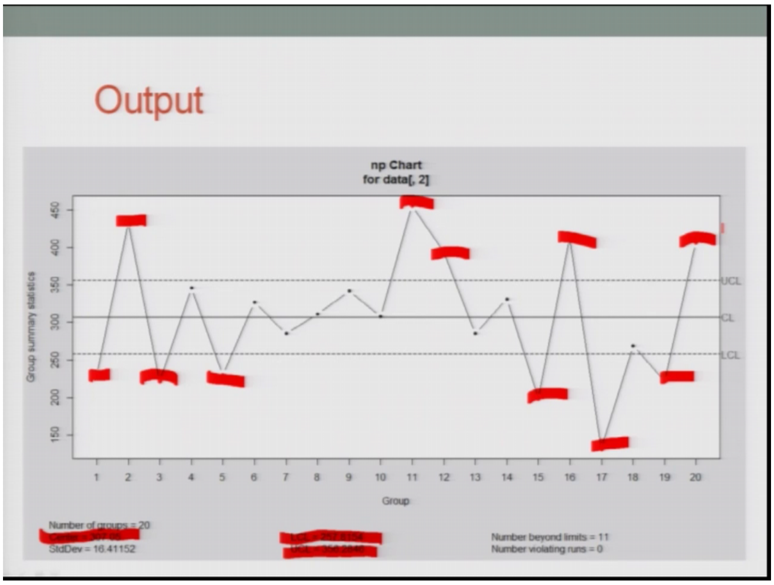
Now, let us consider the fraction nonconforming control charts and you want to find all the equivalent np charts. So, the lot number is given form one to 10 in the first column number of nonconforming belts which you are trying to test is given as 232 to 308 depending on the lot number. And again in the third column you continue the lot number from 11 to 20 and the number of nonconforming belts are basically given as 456 to 407.

(Refer Slide Time: 26:47)



So, if I put the R code. The matrix basically would be again a 20 cross 2 depending on the reading number and what is the value. So, the values are given which is 234 35 till the last 2 is basically 221 and 407. And again you basically recall the same set of functions in r and try to basically plot the graph.

(Refer Slide Time: 27:10)



So, here is the graph. The central line basically comes out to be about 307 0.05 which is highlighted here. And now you will see the lower control values would not be 0. So, you will basically find out it comes out to be as 257 0.184 and an upper control comes out to be 56.2848.

So, once you basically plot it you will find out all the values are within the control limits as given. So, number of control will limits and given and I am coming to that. So, this red light points which are marked here this one, this ones: third fourth fifth sixth 7th eighth ninth tenth eleventh. So, these are the ones which are outside the upper control and some are below the lower control and the rest of the values which are basically black dots are within a control. So obviously, we will try to find out some reason take them out from the calculations and try to basically read do your calculations accordingly say. I will come continue discussing more about TQM and the charts.

Thank you very much for your attention. Have a nice day, bye.