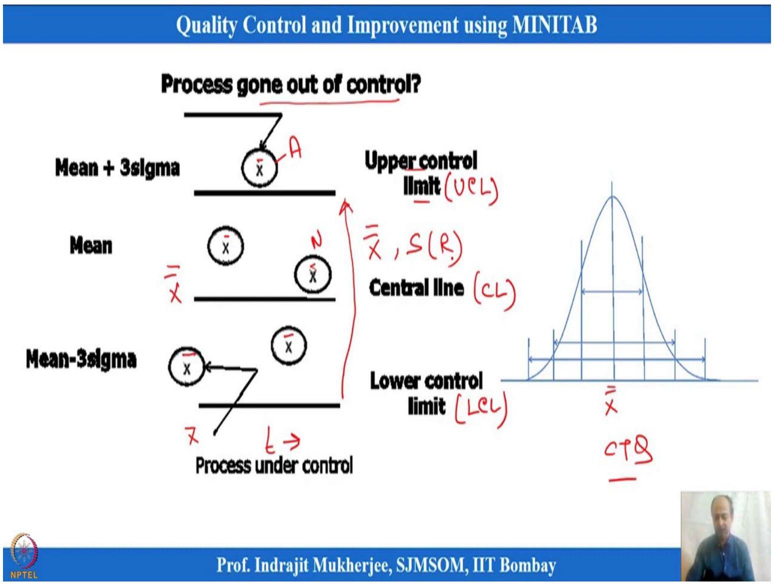
Quality Control and Improvement with MINITAB
Prof. Indrajit Mukherjee
Shailesh J. Mehta School of Management
Indian Institute of Technology, Bombay

Lecture - 09
X-bar and R Chart

Hello and welcome to session 9 of our lecture on Quality Control and Improvement using MINITAB. So, I am Professor Indrajit Mukherjee from Shailesh J Mehta School of Management. So, last time what we were discussing about control charts which is used to differentiate between normal and abnormal scenarios with respect to time.

(Refer Slide Time: 00:41)



So, what we are discussing last time we will start from there, and this was the basic details that we are discussing over here. If it is a normal distribution, so in that case what is expected that 99.73 of the observation will should fall within plus or minus 3 standard deviation. So, if we can draw a two demarcation line one is known as upper control limit line over here which is also in short form you can write as UCL, and lower control limit line as LCL.

So, if you can define this line with a central line over here which is known as CL over here, and that will help us to differentiate between normal and abnormal scenarios. So, this is a normal scenario, this is an abnormal scenario what we discussed last time, ok. Abnormal scenario means assignable cost, due to certain assignable cause or cause that is that leads to a abrupt change in the process mean over here because central line is basically mean over here.

So, overall process and this is the mean over here what you see over here. So, suddenly there is something has gone wrong and the average has come out which is this observation over here. This will be some average which is plotted at every time point and at different time points we are taking information that also we discussed last time.

And we are plotting average because average tends to be much smoother and it follows normal distribution from the central limit theorem which we have told last time, ok. So, if we can and this is the CTQ; we are monitoring the CTQ basically, this X bar is for a specific CTQ let us assume, ok.

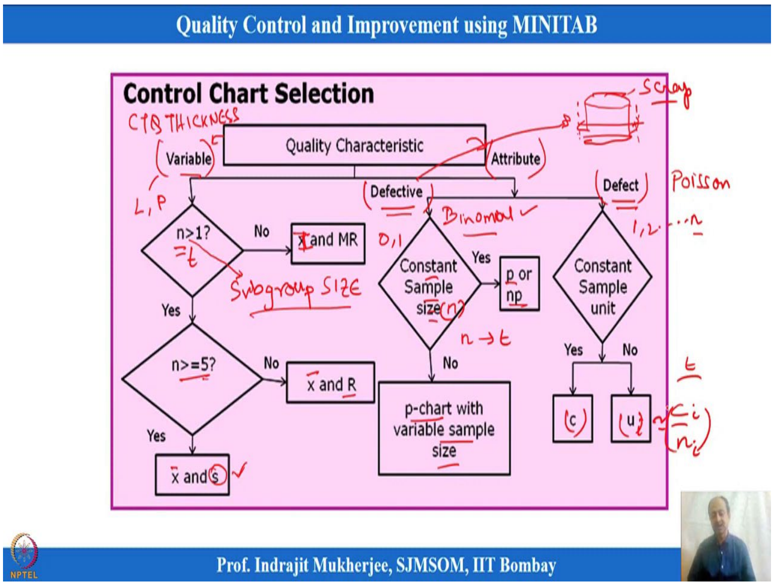
So, if we can define the upper limit line and lower limit line and a central line, so in that case it becomes easier for us to understand when the process has gone out of control, ok. So, this is, if the process if something goes outside the limit line we say process has gone out of control basically, ok. So, then what we do is that we stop the process and take some precautionary actions and so, that this does not recur this type of incidence.

And if it is a mass manufacturing we may not be able to stop, but later on we can diagnose and see that which is the assignable cause which we can block in next operation or next shift maybe it will not recur like that. So, it uses a signal kind of, signal alarm kind of scenarios and it is a proactive actions that we generally try to take, so that it does not recur basically, ok.

So, this is the concept of control chart. So, we plot with respect to time. At different time points what we will do is that we will monitor the average values over here and for that we will define a limit lines over here based on process average, overall process average. And based on the information of standard deviation of the process or range of the process maybe, we will be try to differentiate where to use which one over here.

So, mostly people prefers to use R or S like that for expressing the variability and so, those things we will discuss how to implement that in MINITAB. So, assuming that one CTQ we are monitoring and we want to use control chart for that, and try to figure out when there is a abnormal scenario, when there is a normal scenario like that, ok.

(Refer Slide Time: 03:54)



So, there are different types of control charts suggested by researchers like that, ok. So, one is known as variable control chart what you see on the left hand side, and one is known as attribute control chart over here. Variable control chart like thickness what we have mentioned, so that can be considered as variable for that CTQ we can use variable control chart.

Because a variable control chart where we are interested in location and also in precision both the things are important for me. Because variables can be measured and in that case continuous variable mostly and in that case or ratio scale variables we can think of. So, that is what we can monitor. And for that mean accuracy and precision both is important. So, in that case both can be measured and can be monitored like that.

An attribute like defects what you see over here and defectives. So, difference between defects and defectives we need to understand, when to use defect control chart, when to use defective control chart. So, in a product what happens is that there can be innumerable problems or defect types like that, some can be rectified, some may not be rectified like that, ok. So, when we can rectify all then we have a rework, we have reworked all the defects and we have a good products like that, ok.

But sometimes what happens is that one of the defects has gone beyond my control. So, if there is a bore dimensions which has to be within certain specification over here and I have manufactured something which is having a dimension more than that one; that means, I have basically the bore dimension has increased over here due to processing or some other mistakes like that.

But this cannot be rectified. You see these kind of defects cannot be rectified, so that means, this bore is of no use for the next operations like that, ok. So, this goes as a scrap basically, this goes as a scrap over here. So, that whenever it goes as a scrap and we cannot do anything about that, so then we define that as a defective items like that, then we define that as a defective item over here. So, whenever.

So, to create, so to express a defective items we need at least one defects which cannot be rectified basically, ok. So, attribute control chart talks about defects and defectives. This is a primary control chart which is implemented in processes like assembly processes, or any other manufacturing processes. First our identification whether defects are going abnormal, defectives are going abnormal and then if it is so then can we link it with some CTQs and monitor that one.

So, it has to be converted because defects has no meaning. Defects because of what? Some CTQs is like that. So, we have to convert the defects into some CTQ and monitor that CTQ, so that we take more proactive actions on the CTQs like that when that is going out of controls like that, because that will create defects, that will create defects.

So, this is a primary level of data analysis what we do by control charting, but if you want to go depth in depth like that then we go for variable control charts like that because defects are linked with some CTQs like that, ok. So, if you can link defects with CTQs, multiple CTQs can be and individual CTQs can be monitored or together also there are control charts to be monitored which is known as multivariate control chart, ok.

So, the basic difference between defects and defective is that defectives cannot be rectified, defects can be rectified, but minimum one defect is required to define a defective items like that. So, there are different types of control chart to monitor that one. Now, in variable control chart we have to understand that at a given time point t how many observations we are taking over here. How many observations we can take for calculating the average of the process like that, average of the process.

So, we may manufacture let us say baby foods and in that case 4, we can take let us say 50 packets are coming out of the process at a given time point and we will take only 4 out of that. So, that 4 number which we are taking at a given time point to calculate average weights of the packets like that, so in that case that is known as subgroup size, that is known as subgroup size in quality control, subgroup size, ok.

So, this subgroup size is an important aspects. So, we need to ensure some subgroup size over here based on the process details. So, we have to define the subgroup size. So, generally subgroup size is taken as 5 or more like that, 4 to 5 or more than that also can be taken depending on how much we can; if it is destructive kind of testing in that case it becomes difficult to take more samples like that.

But 4 or 5 is sufficient enough to draw the primary control chart like what is known as X bar R chart or X bar S chart. So, whenever the sample size some guideline is given this is not absolute guideline what we are using over here. So, maybe if the sample size is more than 5 we can go for X bar and standard deviation s chart. So, this is known as X bar and s chart, and this is known as X bar R chart. So, if it is less than or equals to 5 maybe we can restrict to X bar R chart like that.

So, if it is a variable and more than 1 subgroup size is there, so n is greater than 1 what is written over here n is known as the subgroup size. And if n equals to 1 sometimes in chemical process what happens is that we do not need more than 1 samples, because the variation will be very less sample to sample at a given time point may be viscosity or something like that we are talking over here.

In that case, there is little variation; even if I take more than 1 samples readings will be more or less it is same. So, in that case, it is unnecessary to take more number of samples like that. So, if that is the scenario, what we do is that we plot individual moving range like that we can write I MR type of chart which can be written as I MR over here.

Or if it is more than one samples can be gathered and we know that there will be sample to sample variation, so in that case whether, we have we can take about 5 samples like that; if it is within 5 samples what we will do is that we will use X bar R chart like that and in case it is more than 5 also we can afford. So, in that case more appropriate chart will be X bar S chart like that.

So, we will go variable control chart we will do first and then we will see attribute control chart, how to do it in MINITAB like that, ok. So, variable control chart within that also we are restricting over here to for larger shift the control chart that I use to detect large shifts, and in that case what Shewart what has recommended one is X bar R chart, X bar s chart. And if it is n equals to 1 then how to use individual moving range chart which is known as I MR chart like that. That we will discuss first then we will go to attribute chart.

Within attribute charts also there are bifurcations over here. What you can see is that if it is defective chart we have constant sample size. The number of samples that is subgroups that is taken n whether it is constant or whether it is varying like that. So, in case it is constant in that case we can use p or np charts like that.

We can also use np charts over here or p charts, p and np chart is more or less same, only for the operators benefit of the operators what we do is that we can multiply it with n and that gives the operator some sense that defectives. If we write 0.00 something defectives like that we will see how calculation is done. So, np chart is only to facilitate the operator to monitor like that.

They understand 2 number of defects, 3 number of defects, 2.5 defects that is understood. But if we write defects in decimal place or something like that, defectives in decimal place that is not well understood by the operators. So, in that case what happens is that somebody prefers to use np chart instead of p charts like that, ok.

But there is also possibility that we can vary the sample size when n is not constant at any given time point. So, n keeps on changing with respect to time like that. So, if we keep on changing the n in that case we have a variable control chart p chart with variable sample sizes like that. That we will also demonstrate in that; if that is the scenario how the control chart needs to be used and how it is monitored like that, ok.

Similarly, in case of defective, so defective is like how many assemblies you are monitoring the assembly operations like that and 20 engines are manufactured and in that case you take some sample and or 100 engines are manufactured out of that you are taking let us say 5 of them and you are checking that whether it is, whether it is defective or no defect 0, 1 condition basically, whether it is working or not working like that, that is the condition, ok.

Defective is 0, 1 condition, so either it is working or not working type of scenario. Here we can have 1 defects, 2 defects like this, there can be n number of defects like that. But whenever I am saying defectives this is 0, 1 and the primary distribution that it follows we are assuming is binomial. So, this is binomial distribution that underlying distribution is binomial for this and this is a Poisson distribution which is considered over here when we are talking about defects like that, ok.

So, this way; so, defects when we are talking about defects the control limit lines will be calculated and that will be based on binomial distribution and for that different types of control chart n and p will be demonstrated and for if it is defect types also n can vary over here.

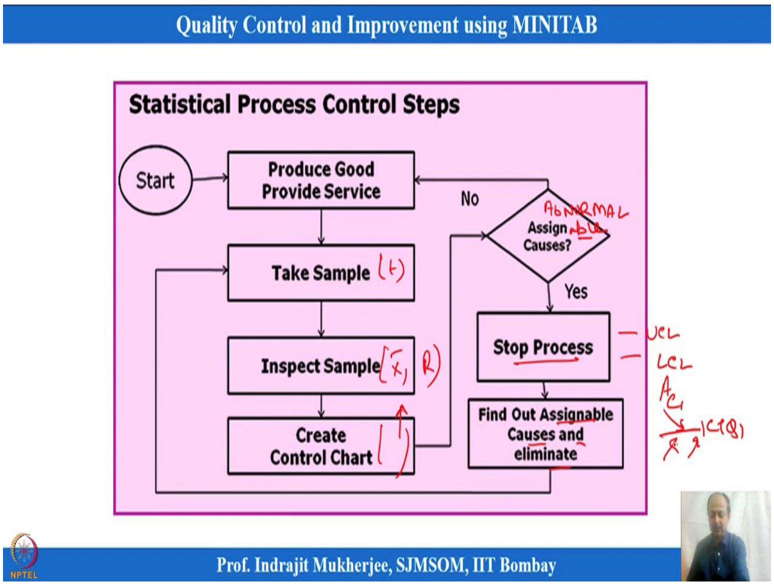
So, if n is constant in that case we have a c type of control chart and in case we have a we have varied n over here, so at a given time point there are number of samples that was inspected is different and number of defects is monitored in that case we can use u charts. So, in software industry you will find that people are talking about lines of code, so within that lines of codes how many defects there is or error that has happened basically.

So, that can be monitored because lines of codes are different at given different time points like that in a projects it can vary, so if it varies in that case u is taken as. So, this can be calculated at a defects by a number of observations. So, n_i over here and these are the c_i number of defects at a given time point t let us say. And if you take the ratio of this you will get u conditions over here and this is the u variable over here and this u will be plotted like that in the control chart instead of c charts over here, ok.

So, our basic assumption is that quality characteristics can be of different types, but if we can go to the highest level of precision then we have a variable control chart for that and if you are going for defects or defective attribute types of scenarios which is quality of data is not so high as compared to when we are going to variable control charts like that.

So, here only defects and defective that level of differentiation is only possible. So, lowest level is defectives we can assume the data quality is lowest level is defective, then defects maybe, and then variable controls are which is the highest level data we what we can gather and based on that we take a decision like that, ok. So, assuming that CTQ thickness we can measure and in that case how to implement control chart that we will try to see, ok.

(Refer Slide Time: 15:03)



So, what do you do in statistical process control basically? We do we take some sample at a given time point that I told already that at any given time point t. We will go to the process and we will try to take number of samples n equals to 4, 5 or something like that more than 5 also.

So, we take a sample, then inspect the samples and calculate the average of the samples maybe, and range of the samples may be of the 4 or 5 samples that I have taken, and then we have a control chart first we create try to create the control chart, and if it is already created then we just plot the values of what we have taken over here. So, control chart will have some limit lines upper limit lower limit that we have discussed earlier.

And then we see whether there is abnormal or normal condition like that, is it any abnormal condition is there and in that case if it is yes stop the process like that. So, this will be assignable, so assignable cause. So, in case abnormal conditions are not there we will not stop the process and if assignable cause is there or point is going outside the limit lines that is defined LCL like that.

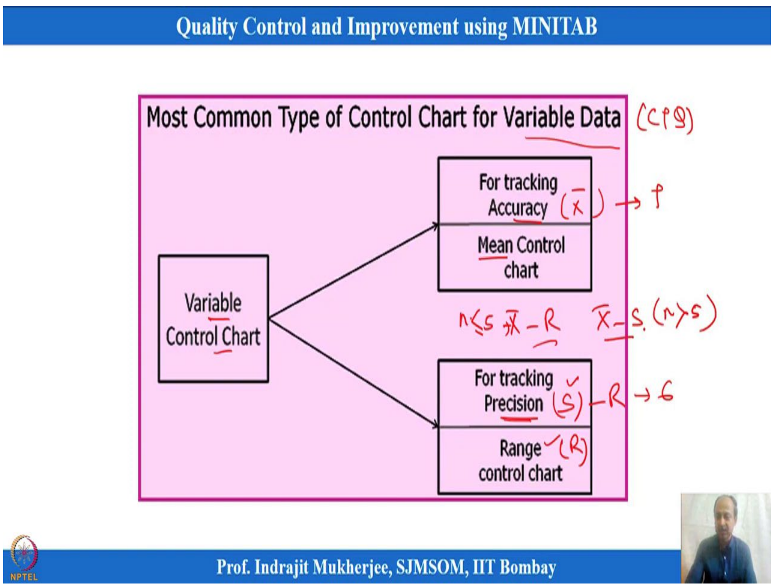
So, in that case what will happen is that we will say that some assignable cause is there and we will try to eliminate that cause. So, we will have a fishbone diagram for this. So, this is the process CTQ which is going wrong over here and this is the cause which is basically primarily influencing and for that we are getting assignable cause. How to deal with that, so that it does not recur? So, this may be assignable cause 1. So, in that way we have to; we have to think like that, ok.

So, this is the overall process. So, we try to monitor whenever product is coming out of that we do not inspect everything out of that. So, we will take some 5 samples out of that and then find out what is the average value of whatever characteristics we are monitoring over here and then plot it into control chart which is already created. Let us say and then figure out whether it is in control or out of control scenarios.

So, if it is in control, everything is fine we do not have to change the settings or adjustment is not required in case something is plotting outside figure out what is going wrong over here or adjust the process like that, either I will block the cause or in presence of cause I will change the settings, so that our again the process target values are satisfied. So, X bar is close to the target.

So, that way we have to think using the control chart. So, this is a signal alarm based type of approach that is used and proactive approach that we can use over here. So, and this control limit lines will be calculated that we will show in next subsequent slides like that, ok.

(Refer Slide Time: 17:40)

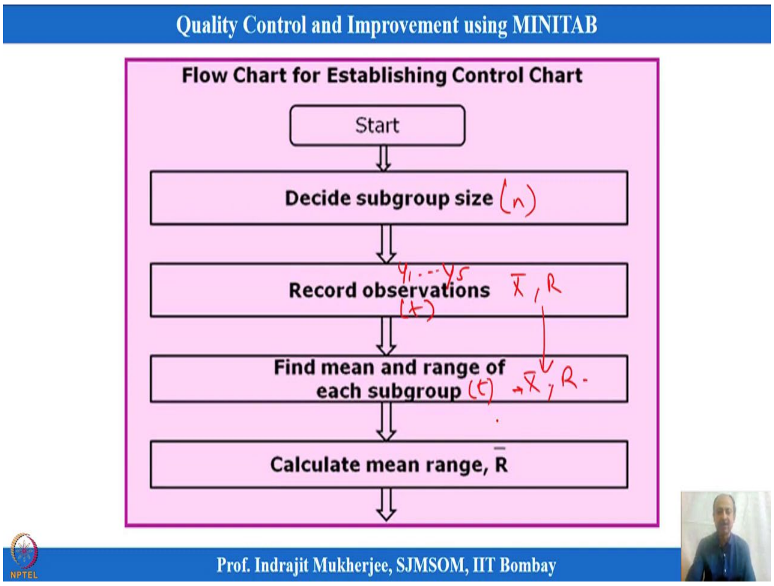


So, variable control chart, variable control chart monitors two thing, one it monitors mean and also it monitors precision. So, one is accuracy it monitors one is precision it monitors. Accuracy means how the mean is moving. So, mean value is important because that has to be monitored with respect to target and here what we are monitoring is standard deviation or range over here, so which talks about variability sigma; that means, how much variation in the process.

So, maybe S type of control chart can be used or range type of control chart or R chart can be used over here. So, combination of this X bar chart with R or X bar chart with S is generally used to monitor the variable type of CTQs, variable type of CTQs whenever I have a CTQ which is variable type in that case I will use this X bar R chart to monitor that one or X bar S chart over here.

General recommendation is if $n \le 5$, we can use this one and if in case $n > 5$ and in that case we can use X bar S chart. More or less the concept or the formulation over here is more or less same. So, in that case we do not have to; some variable, some constant will be changed otherwise the process remains same process remains same, ok.

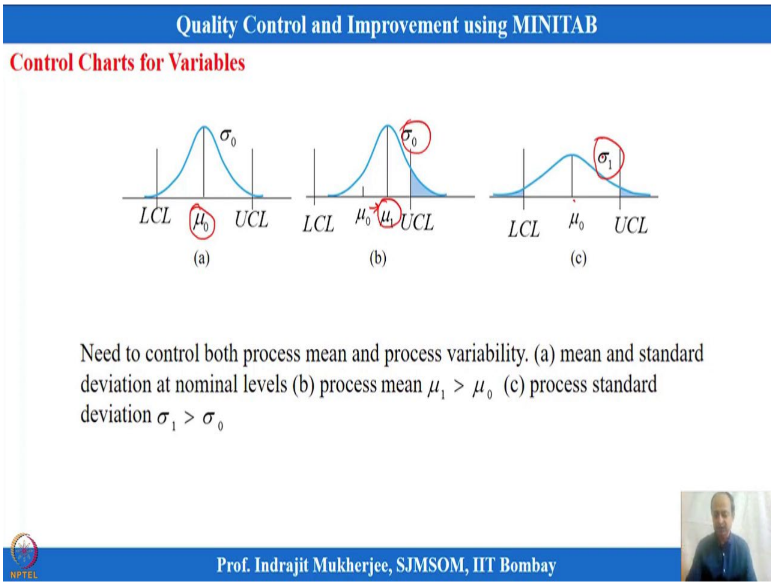
(Refer Slide Time: 19:04)



So, then what we will do is that we will try to see. So, what we do is that we try to first decide on the subgroup size that is n. So, what we do is that, we decide on the subgroup size n over here. And based on the subgroup size n then what we do is that we record the observation what is the X bar and what is the range for at a given time point t let us say, at a given time point t. And with the all information, so here also we calculate X bar and R like that.

So, here record observation, maybe I have 5 observation over here, y_1 to y_5 like that and then after recording the observation I calculate the mean of the observation and range of the observation at a given time point t, ok. So, then we calculate. So, this is the same thing what we are seeing over here. So, more or less then we calculate the average and range and that will be plotted basically, that will be plotted like that, ok.

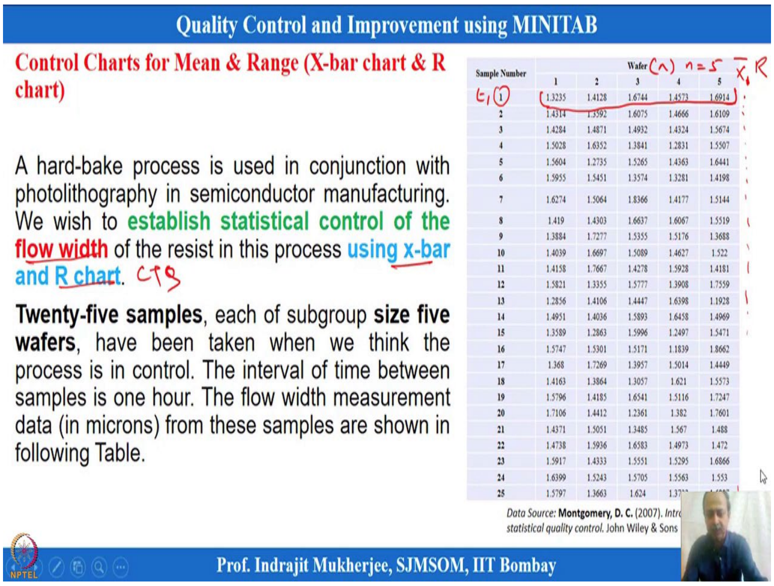
(Refer Slide Time: 20:02)



So, what we can see out of this is that whenever the mean, let us say this is the mean over here, and mean can shift over here mean can shift; that means, location can shift over here and mean may be same and then standard deviation may change over here. So, σ_0 is the standard deviation variation can also change like that. So, both will impact the control chart.

Whenever there is a variation problem it will be indicated, whenever there is, we have a mean shifting problem there also we will have a situation like that. So, that is also we will come across when we are using the control chart like that, ok.

(Refer Slide Time: 20:38)



So, here we will take a specific example to illustrate this one and this is taken from Montgomery's example, statistical quality control book. And in this case, it is saying that hard bake process is used in conjugation with the photolithography in semiconductor manufacturing and we want to establish control charts, statistical control and try to see and for that some data was gathered like that and flow width is a parameter which was monitored over here. So, this is the CTQ which was monitored.

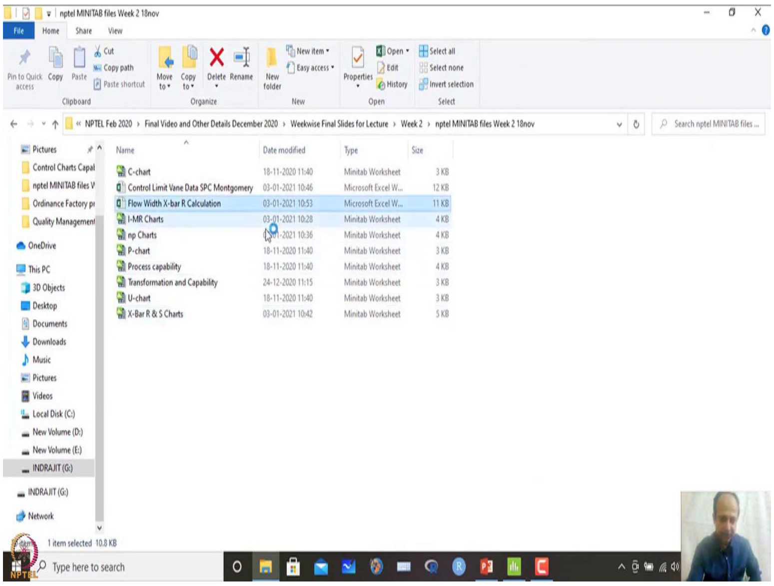
And we want to use X bar R chart for this process like that. So, you will see at a given time point t_1. Let us say first sample observation was taken and given number of subgroups over here $n = 5$. So 5 observation was measured at a given time point t_1 over here, ok. And at this instance what we can do is that we can calculate the average over here which is x over here.

So, what is the average of this? So, all this 5 observation that way we can calculate average and we can also calculate range maximum minus minimum maximum of this observation and minus minimum of this observation. So, this can be easily done when we are using excels like that, ok. So, let me take this example and show you to excel what it does. So, then we have 25 number of averages we will get, we will get 25 average over here and we will have 25 ranges over here.

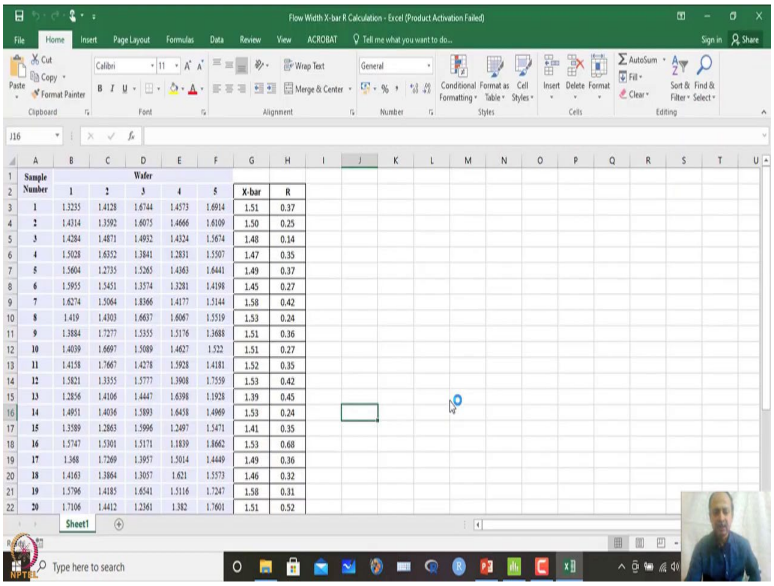
So, then what we can do is that $\overline{\overline{X}}$ can be the average of all $\overline{X}$ over here and divided by 25. So, this will give me a overall average or grand average over here. Also range average we can calculate, so we can calculate we have a R, whatever R values we can also calculate range average. So, we will have 25 range and average of that range can also be calculated like that, ok.

And then we will use the control chart to monitor this one. So, let us go to the, what we will do is that we will see that this data is already available with us and we will just demonstrate that one, ok.

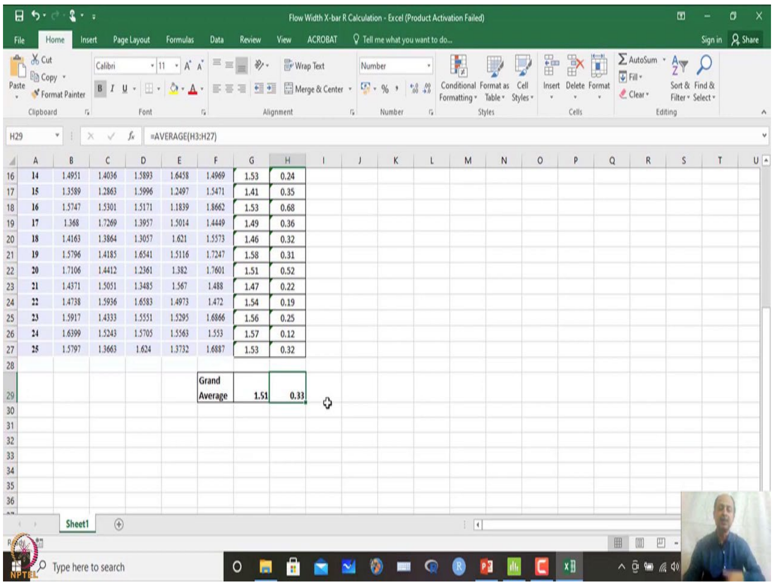
(Refer Slide Time: 22:52)



(Refer Slide Time: 22:58)

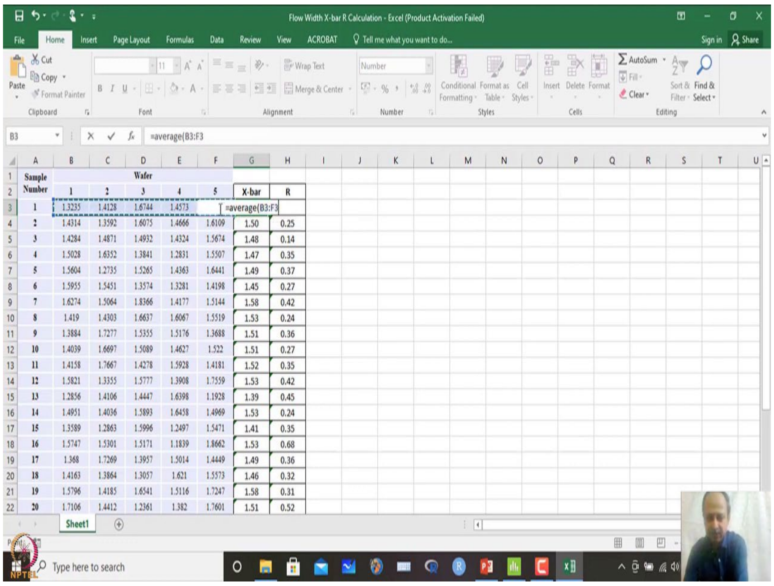


(Refer Slide Time: 23:04)



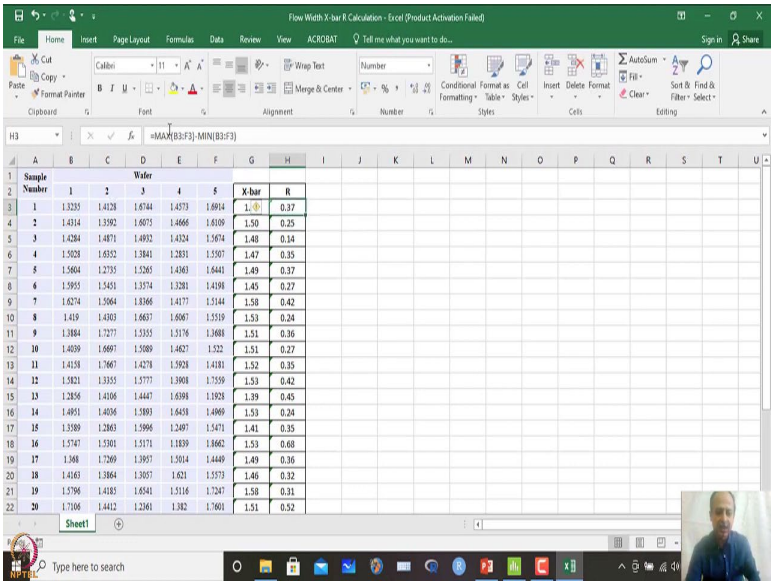
So, flow width X bar R chart. So, this is the data set that we are having. What I mentioned is that with at different time points what is done is that this is the data set, the same data set that we have shown from the examples over here, 25 data sets.

(Refer Slide Time: 23:10)



And then $\overline{X}$ was calculated which can be calculated as we can just write $\overline{X}$ is the average observation of this. So, we can write what average. So, I can highlight this one, 5 averages over here.

(Refer Slide Time: 23:21)

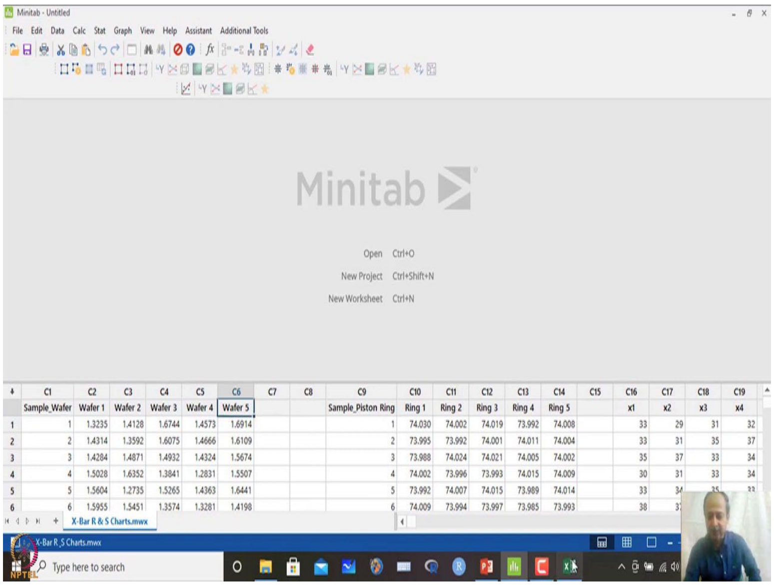


And then range can also be calculated as maximum, what you can see on the top maximum minus minimum of this observation. So, maximum of this observation, so B3 to F3. So, this is B3 to F3 and minus minimum of this. So, range can also be calculated for a specific row over here. So, similarly, we can calculate all average and all range and grand average can also be calculated which is the average of all this observation that you see over here, ok.

Similarly, range average can also be calculated. So, these two values are required to develop the control chart limit lines like that. So, $\overline{\overline{X}}$ is 1.51 over here and $\overline{R}$ is 0.33 over here. So, this will help me. So, I have 25 observations with subgroup size of 5 over here and then I can calculate individual values. So, these are the individual X bar R values.

This will be plotted in the control chart and I have a grand average, and average of the range and which will help me to define the control limit lines that is which is normal, which will differentiate between what is normal and what is abnormal basically, ok. So, this will be required when we.

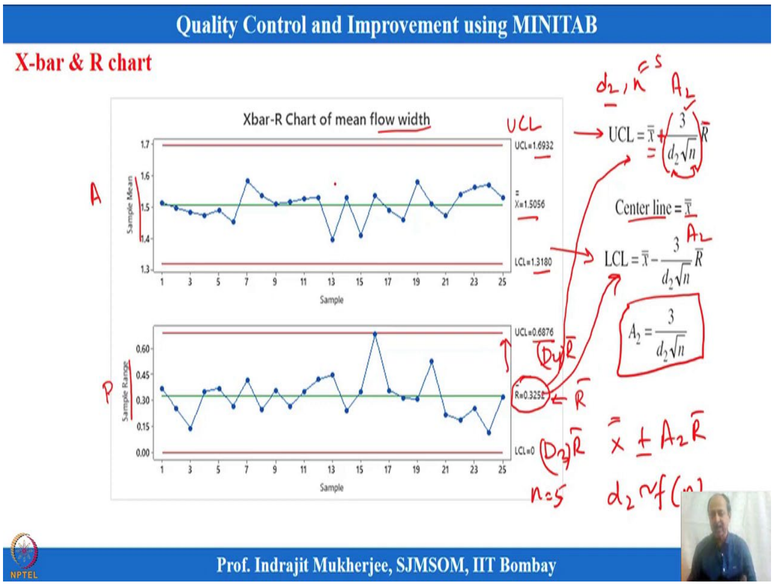
(Refer Slide Time: 24:34)



So, let us take this data set into a MINITAB. So, we will just see and it is already there in the MINITAB. So, this is the observation wafer 1 to wafer 5 I think. So, this is the same observation wafer 1 to 5. So, 1.3235 and this is 1.3235. So, that you can see over here, ok. So, and this is the new sheet which is named over here X bar S chart.

So, you can also create from the excel and you can just type in the sample observation over here. So, whenever you have typed in the data then we can create the control charts.

(Refer Slide Time: 25:07)



And what is the formula that is used to create the control chart over here? So, what you see formulation over here is that there is a MINITAB we will use a calculation method to get the upper control limit line over here and this is X double bar 1.50, and this is same as what we have calculated 1 point grand average what you what we have calculated over here. So, maybe we can just cross check in excel.

So, grand average is 1.51 and range average is 0.33. So, in this case also you will find that the MINITAB calculates 1.5056 and that is close and also range is 0.3233 or that is also close. So, MINITAB does the same calculation; so, from the same data set. So, what we have done in excel, so $\overline{\overline{X}}$ and $\overline{R}$ is calculated in same way. Only we have to understand how this UCL and LCL is calculated over here, and UCL has a specific formula.

So, UCL formulation over here you see this is $\overline{\overline{X}}$ information, so that is the grand average is taken over here and there is this is multiplied plus sign is over here. And this is a constant factor that is multiplied with the $\overline{R}$ or R average range average what we have calculated already over here, which is given over here.

So, this is used in the formulation to calculate UCL and this is also used to calculate LCL over here. So, one is plus A factor and this can be written as this is the, I can replace this one with A_2. So, I can write this as A_2 and this is also A_2 like that. So, formulation changes like this $\overline{\overline{X}} \pm A_2\overline{R}$. So, we can just write that one replacing this one, ok.

So, what is this A_2 over here? There is a constant which you see as d_2 over here observation d_2 and n is the subgroup size like that, n = 5 for our case over here, this example flow width because 5 observation was taken over here. So, n equals to 5 and this d_2 is a function of n over here.

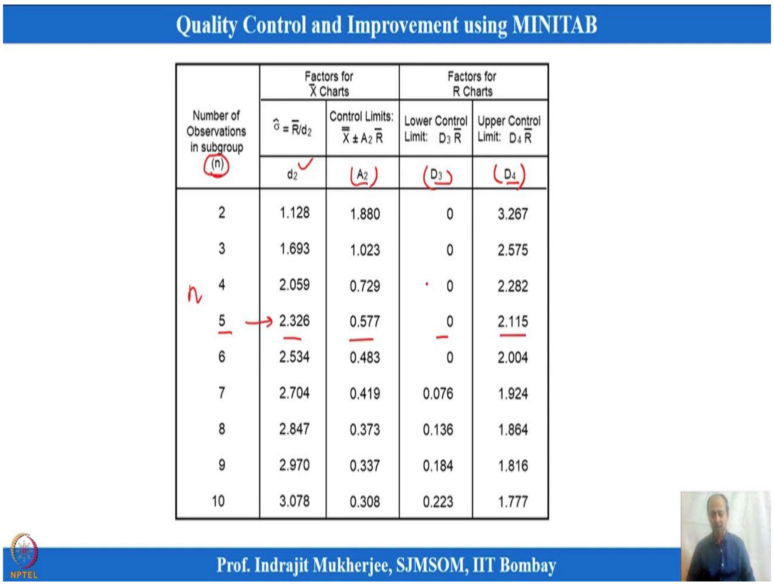
So, depending on the subgroup size this d_2 value will change like that, ok. 3 is a constant, so 3 does not change over here, but these two has a relationship with n and d_2 are related with n, so this constant will change, ok. So, this is for X bar chart over here, this is for X bar chart, so I will get a upper control limit line using this formula that is given over here.

I will get a lower control limit line using the formula over here because every information is with me, n is with me, d_2 value is with me which is we can get from table and then $\overline{R}$ is there $\overline{\overline{X}}$ is there. So, I have defined UCL as 1.69 and LCL comes out to be 1.31. Similarly, for range chart over here and central line will be $\overline{\overline{X}}$ in this case.

So, this is $\overline{\overline{X}}$ what you see. Similarly, range average will be what we have calculated will be the central line, and the upper control limit line is calculated based on certain formulas that is $D_4\overline{R}$ and lower control limit line is $D_3\overline{R}$ like that. Now, D_4, D_3 is again a constant which depends on what subgroup size I have taken over here which is equals to 5.

So, this will define the D_4, D_3 value which is a function of basically n subgroup size like that. So, that formulation is also we can see. So, this is central line will be R bar over here and one side upper control limit line with the $D_4\overline{R}$ and $D_3\overline{R}$. So, this is monitoring sample range what you see and this is monitoring mean over here. So, one in monitoring accuracy, one is monitoring precision over here.

(Refer Slide Time: 28:47)

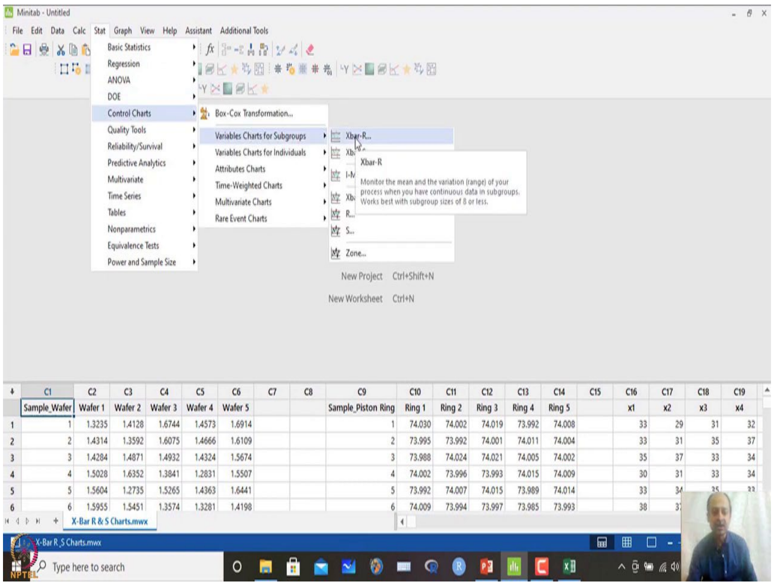


Number of Observations in subgroup (n)	Factors for $\bar{X}$ Charts		Factors for R Charts	
	$\hat{\sigma} = \bar{R}/d_2$	Control Limits: $\bar{\bar{X}} \pm A_2\bar{R}$	Lower Control Limit: $D_3\bar{R}$	Upper Control Limit: $D_4\bar{R}$
	d_2	A_2	D_3	D_4
2	1.128	1.880	0	3.267
3	1.693	1.023	0	2.575
4	2.059	0.729	0	2.282
5	2.326	0.577	0	2.115
6	2.534	0.483	0	2.004
7	2.704	0.419	0.076	1.924
8	2.847	0.373	0.136	1.864
9	2.970	0.337	0.184	1.816
10	3.078	0.308	0.223	1.777

So, and this is now this calculation can be done by hand also. Because all the charts are available, and this is the; this is the chart where you will get the values of d_2 for a given n over here, so the n is the subgroup size and for a given value of n equals to 5 and I can get what is the value of d_2. I can also value of A_2 and I can get what is the value of D_3 and see what has recommended this value, so what is the value of D_4.

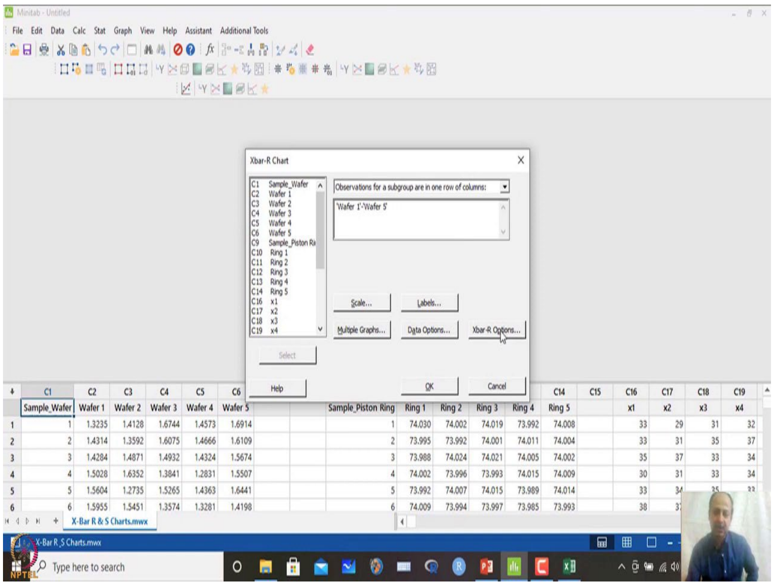
So, using these values two corresponding n changing n over here I will go to that particular specific row and I can get all the values. So, I want A_2 values, I want D_3 and D_4 values, so that will define the control limits for X accuracy and also to monitor precision like that. So, a MINITAB does it automatically for you. So, in that case, it is easy for you to get that. So, how does MINITAB does it? So, this data set is taken into MINITAB and this is very easy.

(Refer Slide Time: 29:41)



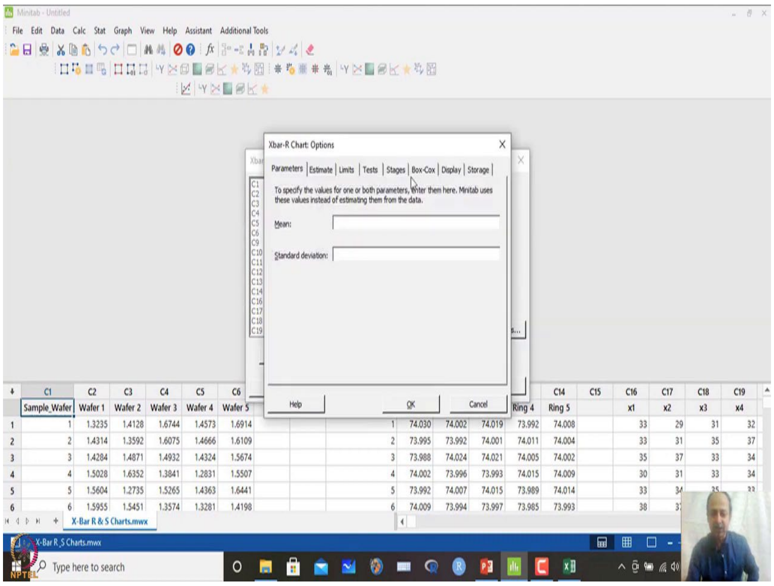
So, when we have the data in MINITAB, so this is the same data set. And what you do is that stat, control chart, variable chart, X bar R chart, this is the option that you have to go. Because I am monitoring 5, less than equals to 5, so I am using X bar R chart over here.

(Refer Slide Time: 29:54)

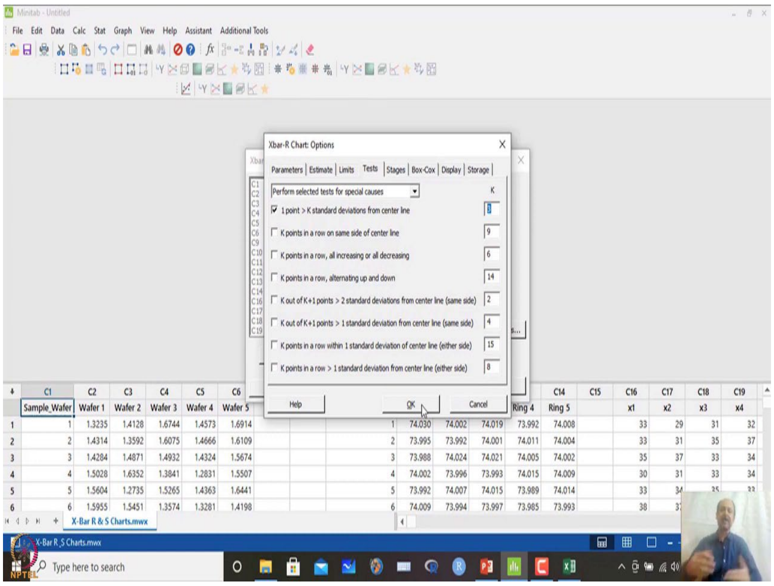


So, I will use X bar R chart then it will ask is it in same column, then I mention, no, it is in different columns like that. So, in that case, so I will mention that wafer 1 to wafer 5 is the data set you take and I select this one.

(Refer Slide Time: 30:13)



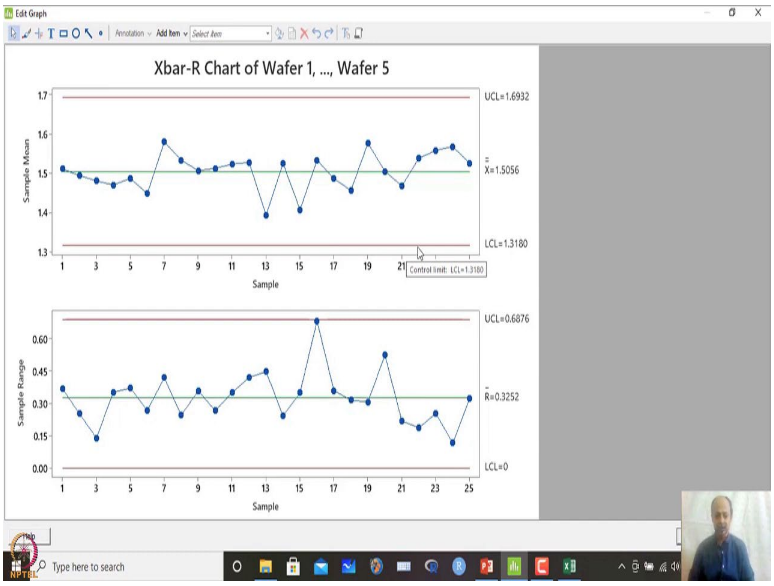
(Refer Slide Time: 30:15)



And $\overline{X}$ any point going outside the limit line we will and this is by default which is given over here as test. So, any point going outside, so I have clicked only one options over here we will discuss about this in next sessions. So, this is if it is going beyond 3σ limit we will take corrective action. So, that will be as will be an assignable cause.

So, I have taken one single condition that if point goes outside the 3σ limit line, this is the abnormal scenario. So, you monitor that one only. So, I click OK over here, and I click OK, what you will see is that you have this control chart like that, ok.

(Refer Slide Time: 30:44)



So, whenever you do that, so X bar R control chart, the same information. So, you see all the points in mean is falling within the control limit line. So, $\overline{\overline{X}}$ is same $\overline{R}$, I have the control limit line upper and lower control limit line. Lower control limit line is 0 because D_3 is 0. So, in that case, this is a 0 line or range can be equals to 0, it cannot be negative, but it can be 0.

So, there is no variation. So, that is the lower control limit line. So, and this is the upper control limit line. So, all the points are within the control limit lines. So, in this case there is no red points that is going outside this limit line what you see. So, everything is going fine over here. So, there is no out of control or abnormal scenarios over here.

So, we will stop over here and we will continue for discussion on this topic from here. Again, we will take some more examples to illustrate and how it is happening. So, abnormal scenarios also we will try to see. So, let us stop over here and we will continue in the next session.

Thank you for listening, ok. See you in next session yeah.