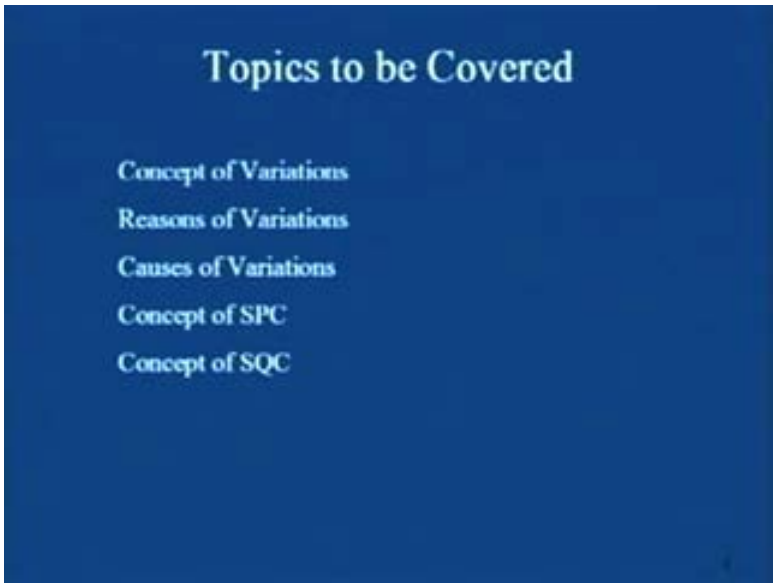
Industrial Engineering
Prof. Pradeep Kumar
Department of Mechanical and Industrial Engineering
Indian Institute of Technology, Roorkee

Module - 02
Lecture - 02
Statistical Process Control Part-2

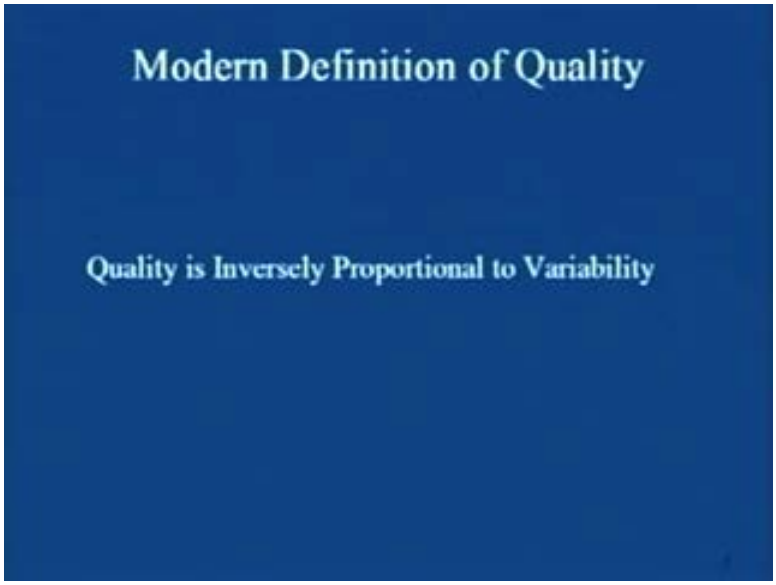
Dear viewers, I welcome you and today we will cover the part 2 of the Statistical Process Control, the part 1 of the statistical process control we have all ready covered. And in the part 1 of this statistical process control, we have covered the importance of the quality, we also had a discussion regarding the historical prospective of the quality along with the various definitions, starting from the traditional to the modern definition of the quality.

(Refer Slide Time: 01:11)



In this part that is the part 2 of the statistical process control we will cover the following topics, that is the concept of variations, reasons of variations, causes of variations and the concepts of SPC; that is the Statistical Process Control and the concept of SQC Statistical Quality Control. In the part 1, I had finished my lecture on this particular definition that is the modern definition of the quality.

(Refer Slide Time: 01:36)



And we mentioned the quality is inversely proportional to the variability, the entire lecture I will start to build up a theme regarding the variability to understand all the concepts of the variability.

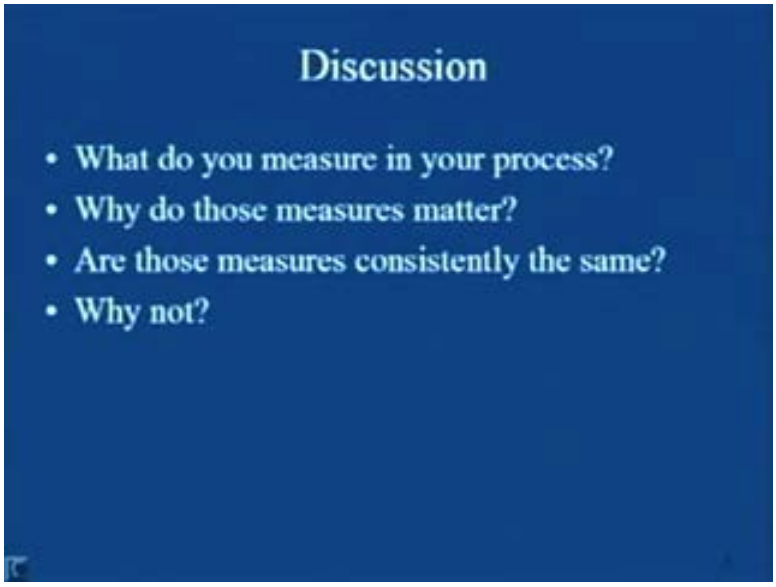
(Refer Slide Time: 02:07)

Now, just an example to build up the concepts of variability, I have written some points that have you ever shot a rifle or played darts, in both the games we have the target and we played dart on this target. These are the just two persons Ram and the Arjun, they basically played the darts on this two targets and the output of the Ram and the output of the Arjun has been shown by the black bullets.

These are there black bullets that is output of the Ram and these black bullets basically the output of the Arjun, now question arises who is the better shot. So, when we think that who is the better shot, first we need to understand what we are going to measure, if we see the Arjun's points that all the bullets or in all the darts right here, they are just close enough to each other and approximately coming to the same point.

And as far as Ram is concerned it is distributed and touching one circle and another circle and the third circle. Now, the question is that whether we are going to take the decision with the respect to the consistency or with the respective to the total evaluation, so the question again here is the, who is the better shot?

(Refer Slide Time: 03:35)



So, before we decide that who is the better shot, let see that what we should do in our processes, so whenever we do has any measurement of the any characteristics of the process we should know that what we are going to measure. Suppose, there is a some shaft that you a turning on a lathe, now the question is whether you are measuring the diameter of the shaft or you are measuring the eccentricity of the shaft, or you are measuring the surface roughness of the shaft.

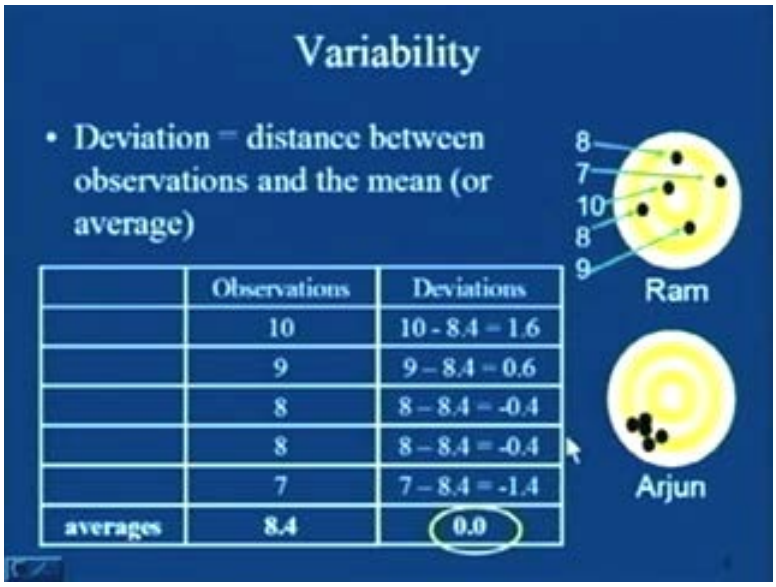
So, first we need to be very clear that what we are going to measure in our processes, second the question is why do these measure matter, means what is the importance of all these measurement. The third question is are those measures consistently the same, means when we are saying the measures consistently the same, this means whether the equipment or the measurement tools we are using the same are the equipments. Or the

tools are calibrated or the method that is being use to measure is same, or the operator is using the same method consistently first time and the every time.

Then, when we say that are those measures consistently the same, if yes good, if not then we should no why not that why these measurements are not coming the same, so before we decide as I have given the example that Ram and Arjun they have played the dart game. So, before we say that who is the better shot we should know what we are going to measure, so depending upon the measurement that we are doing, we should take the decision.

Let see because in that dart game the two possible measurements was there, one was that the count, total count, the total number, second is the consistency. So, when we say consistency means we are talking about the variability and when we are talking about variability we need to define what variability is…

(Refer Slide Time: 05:55)



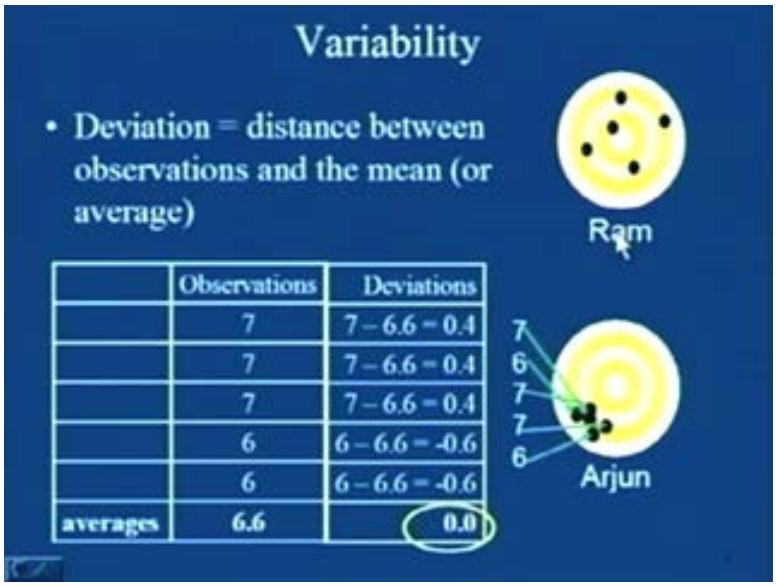
Suppose, we define variability as the deviation, deviation means distance between observations and the mean value or the average value. This is the output of the Ram as I have shown you, in addition to the output in this slide you will see that I have given the numbers, that what was the different output, the different numbers are there 8, 7, 10, 8, 9 depending upon where the shot is there.

So, if you see in this table that there are 5 observation and the average of these 5 observation is 8.4 and if we define the variability as the deviation, this means the observation minus the average value, observation minus the average value, observation

minus the average value, observation minus the average value, observation minus the average value. And if you see the first and second these two observations, the observations are higher than the average value and in the last three observations it is less than that is why the negative sign is here.

So, this figure that represents the sum of the deviations of all the 5 of observations, so this is equal to 0.0, so when we say the variability of Ram in terms of the deviation from the mean that is 0.0. Let us see what happens for the deviation with the respective to the Arjun.

(Refer Slide Time: 07:32)



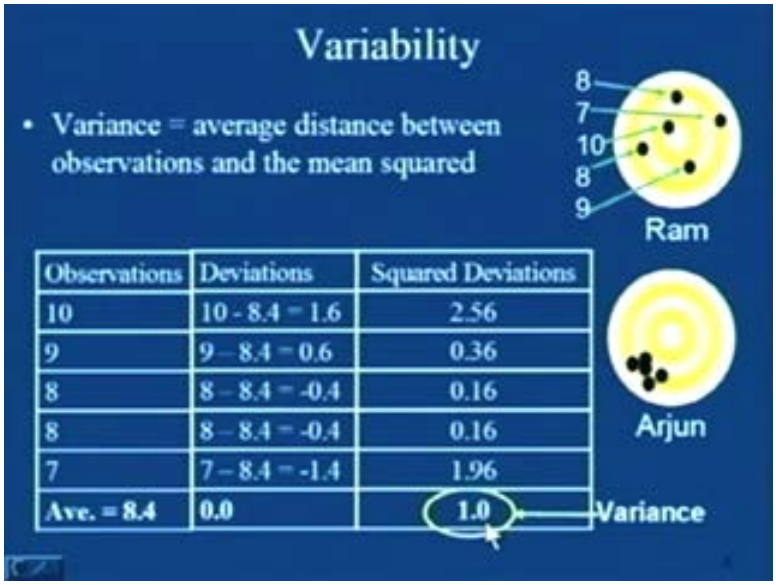
	Observations	Deviations
	7	7 – 6.6 = 0.4
	7	7 – 6.6 = 0.4
	7	7 – 6.6 = 0.4
	6	6 – 6.6 = -0.6
	6	6 – 6.6 = -0.6
averages	6.6	0.0

The Arjun same, so he has the output at 5 I have output here of 5 observations here and all these observations are just at approximately near the one or the second circle, but none of the observation is at the centre or then next circle. So, depending upon the his output the marking scheme is different, because if he marks just in the inner circle the marks for his 10, so his marks are less.

So, let us see the deviation of the Arjun with the respect to the observation, so these are the 5 of observation this gives the average value that is 6.6 and the next column shows that observation minus the mean value. So, all these 5 of observation we have calculated the deviation that is observation minus the mean value and the last two observations if you see, the observations are less than the average value and that is why the negative sign is here.

And the total observation, the sum of the deviation means this sum of that deviation is algebraic sum of the deviation and that is coming 0.0, so what is the difference, Ram has the different observations, Arjun has the different observations and the deviation is coming same. So, this shows that deviation is not the right way to distinguish between the output of the Ram and Arjun, let us see we define the variability in another way, now we define the variability as a variance.

(Refer Slide Time: 09:13)

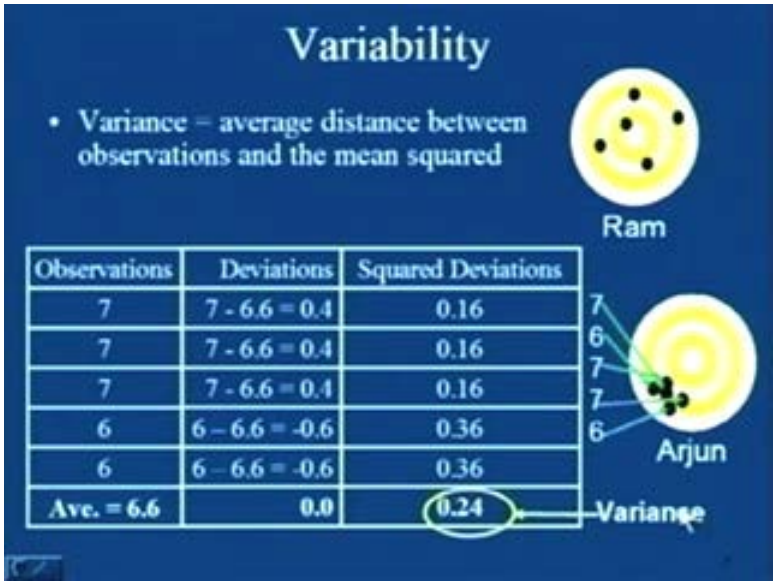


Observations	Deviations	Squared Deviations
10	10 - 8.4 = 1.6	2.56
9	9 – 8.4 = 0.6	0.36
8	8 – 8.4 = -0.4	0.16
8	8 – 8.4 = -0.4	0.16
7	7 – 8.4 = -1.4	1.96
Ave. = 8.4	0.0	1.0

What is the variance is, everybody should must be knowing, that it is average distance between observations and the mean squared; again we will calculate the variance with the respect to this definition, these are the same observations of the Ram and Arjun. The observations have been written here 10, 8, 9, 9 and 7 the average we have all ready calculated it is coming as 8.4, these are the deviations 10 minus 8.4 means observation minus the mean.

So, all the deviations from the mean values have been calculated and since we are calculating variance, this is the average distance between observation and the mean squared. So, these are the deviations from observation and the mean and in the next column we will square all these deviations that is 1.6 into 1.6 2.56, 0.6 into 0.6 0.36. Means all the standard deviations that squared deviations we are calculating, that is the square of these and this figure tells us the summation of these things. So, this 1.0 represents the variance of the output of the Ram with respect to the dart game, so let us see for the Arjun.

(Refer Slide Time: 10:37)

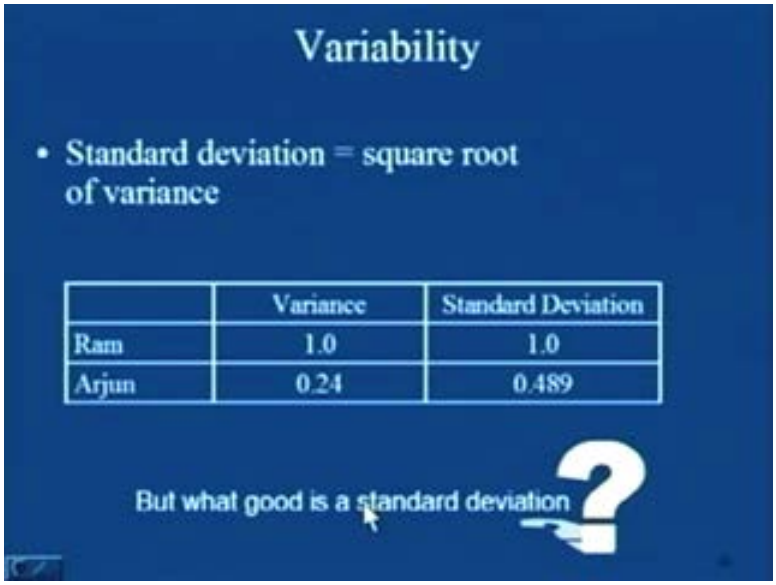


Observations	Deviations	Squared Deviations
7	7 - 6.6 = 0.4	0.16
7	7 - 6.6 = 0.4	0.16
7	7 - 6.6 = 0.4	0.16
6	6 – 6.6 = -0.6	0.36
6	6 – 6.6 = -0.6	0.36
Ave. = 6.6	0.0	0.24

These are the observation 7, 6, 7, 7 and 6 for the Arjun and observations have been listed here, the average value is 6.6 that we have all ready calculated. The deviations, that deviation means that we have all ready defined observation minus the mean, so these are the at deviations of all the observations. And the next column shows the square of these deviations, the square of all these deviations have been given and the total comes that is 0.24 this is a variance.

Now, if you remember the variance of Ram was 1.0 and the variance of Arjun is coming as 0.24, so if we define the variability with respect to the variance we can say that variance of Arjun is much less than the variance of Ram. Whereas, the deviation of Arjun as well as the Ram was same 0.0, so if we define that the our decision we have to take based upon the variability and we define the variability with respective to the variance. And we say who is the better shot, we can say that Arjun is the better shot with respect to the consistency or variance of the output of the results, let us define the variability again.

(Refer Slide Time: 12:08)



We are defining the variability with respect to standard deviation and standard deviation is what, it is square root of the variance, so the variance of Ram was 1, Arjun was 0.24 and the standard deviation is 1 and 0.489. So, even with the respect to the standard deviation also that Arjun is showing the sigma, standard deviation is represented by the symbol sigma, that sigma value of Arjun is much less than the sigma value of the Ram.

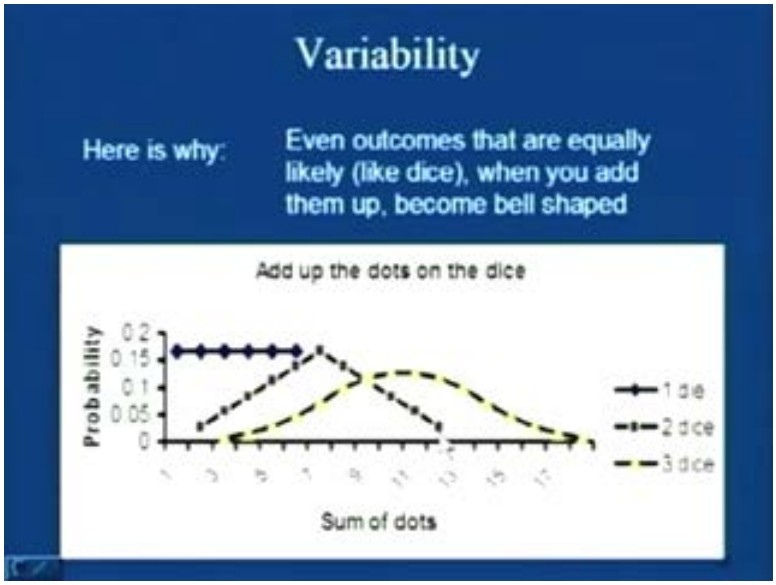
So, this shows that Arjun is more consistent in the dart game, although the total number of points maybe less, but the consistency wise he is the better one. Now, the question arises that, but good is the standard deviation, because we are talking about standard deviation, so how come this standard deviation is good enough to represent the output of a person or output of a process. Let see what good a standard deviation in the next slides.

(Refer Slide Time: 13:19)

This is a normal distribution curve and the most of the observation that they follow the bell shape, means the word tends to be bell shaped. If you are just randomly and unbiased taking the observation the observations are will follow the bell shape curved or the normal distribution curve. Means most outcomes occur in the middle, in this nominal distribution that most of the outcomes that are in the middle, fewer outcomes comes in the tails, something in upper side, something in lower side means some outcomes come in this area.

There are some very rare outcomes are there, it says even very rare outcomes are possible probability is greater than 0, means it goes beyond that middle point in positive side as well as in the negative side.

(Refer Slide Time: 14:16)

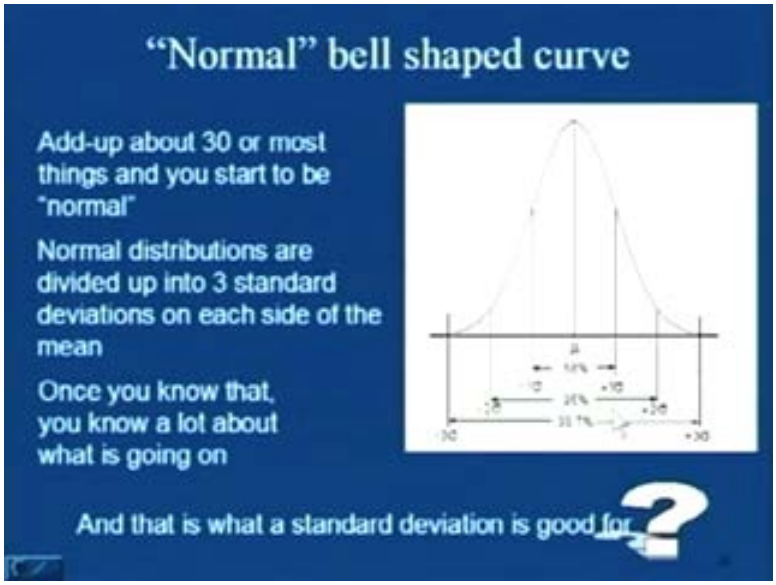


When I said that most of the outcomes tends to be bell shaped, this means even outcomes that are equally likely, like the dice when you add them up become bell shaped, because a dice has a 6 face, 1, 2, 3, 4, 5, 6. And if you just throw the dice that what is the probability of coming a particular face is we can calculate it is 1 by 6, so if I say that the probability of number 1 coming is 1 by 6 the equal probabilities for the number 2 that is 1 by 6 equal probabilities for the number 3.

And if we plot these probability it will because every dice is having a single same probability the curve will be a straight line. But, if you throw two dice then the first one will have you 1 by 36, another will have 2 by 36, so like that the total probability will keep on increasing and that if you plot all this points, all the probabilities of throwing two dice. That if I throw the two dice what is the probability of coming the sum equal to 2, what is the probability of sum coming equal to 3.

So, if we plot all these probabilities and same case is true when we are throwing three dice simultaneously, then we can find out what is the probability of starting from the 3 to 18 that if you plot all these points, these point basically follows the, approximately follows the normal distribution.

(Refer Slide Time: 16:00)



And this normal distribution graph we have all ready seen and I mentioned that most of the outcomes that come in the range, in the middle area that is approximately 68 percent of the area is covered in the middle area and this is nothing but, this is a mean value means plus 1 sigma minus 1 sigma. Means the total area if we say numerical area 2 sigma area basically, plus 1 sigma and minus 1 sigma that covers basically around 68 percent.

If we go 1 sigma further here and 1 sigma further here this total area basically of this curve is 95 percent mu plus 2 sigma and mu minus 2 sigma. And if we go further one more sigma that is mu plus 3 sigma and mu minus 3 sigma this covers 99.7 percent of the total area of the output. And it says add up about 30 or most things and you start to be normal, just add any things if they are random and independent and you will find the things start to be normal.

Normal distributions are divided up to 3 standard deviations on each side of the mean that I mentioned that this is 1 sigma, further 1 sigma and further 1 sigma, so on each side means this is mean plus 3 sigma mean minus 3 sigma. Once we know that, that this is in most of the outcomes that are random and independent the follow the normal distribution and we also know that we can divide up the normal distribution into three part plus 1 plus 2 and plus 3 sigma 1 on one side.

And similarly minus 1 minus 2 minus 3 sigma on other side, that the entire area that means, 99.7 percent of the observations cover in mu plus minus 3 sigma, and once we

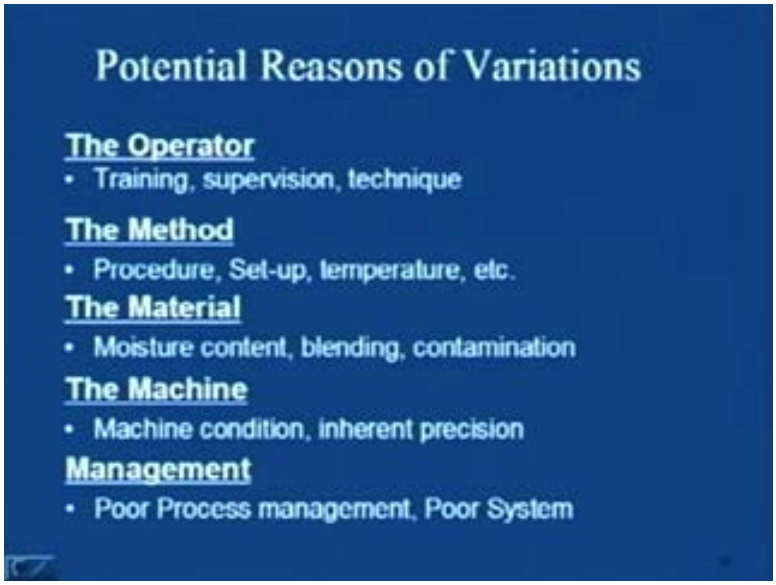
know that we know a lot about what is going on. And that the answer of our question that we asked what good a standard deviation is and we say that this is what a standard deviation is good for means… Once we have the understanding of sigma, we can find out where the observation lies, whether it is in the middle area plus 1 sigma, 2 sigma, 3 sigma, 3 sigma or beyond if it is beyond it means it is out of the control.

So, this is how we are basically understanding the concept of sigma through the variability. Now, question comes that if you remember that in the case of the Ram, that there are different output starting from the middle observation to the second, observation or to the first circle to that second circle or to the last circle, then certainly there is some reason of the variation.

And earlier as I mentioned that when we are going for the measurement of a process, we need to check our instrument, that our instrument should be calibrated and it should provide the observation or the measurement consistently. Also the method should be standardize whatever the method we us to measure the processes that method should be used consistently.

And the operator should be skilled enough, so that whatever is measuring, the way he is measuring, the procedure he is following that he should use the procedure first time right and every time right. So, but still I mention in my first part, that there are inherent causes of variation and let us find out what are the reasons, potential reasons of the variation.

(Refer Slide Time: 20:01)



The very first reason of the variation is the operator himself, that sometime there is a not proper training to the operator, sometime not proper training for the supervision and sometime the operator is not using the right technique. Sometime he use the one method, sometime he use the another method, so that basically is that he is not aware of the right technique to measure.

There is another reason that the method, that procedures are not standardized there is set up change that temperature and maybe some other conditions, ambience condition that changes from time to time. Say for example, if you take the measurement in any foundry shop regarding the temperature of the molten material in the morning and in the evening that because of the temperature range, that because in the morning the temperature in the winter season varies from morning to evening.

So, that the effect of that ambient temperature comes, so this may become the reason of the variation, if you have measurement in the morning, if you have measurement in the night there is a variation in that thing. The variation may possible be with respect to the material, the moisture content, blending, contamination, sometime what happens when in a manufacturing concern or in an manufacturing organization, sometime the material is coming from one vendor, sometime the material coming from another vendor.

If there is a possibility that whatever the material is coming from one vendor the constituent, the characteristics the properties may slightly vary from vendor to vendor, so this may become the reasons of the variation. The machine conditions, inherent precision sometime because of the usage the wear and tear happens in the machine, and because of the wear and tear in the machine there maybe some chatter, some vibrations that occurs in the machine.

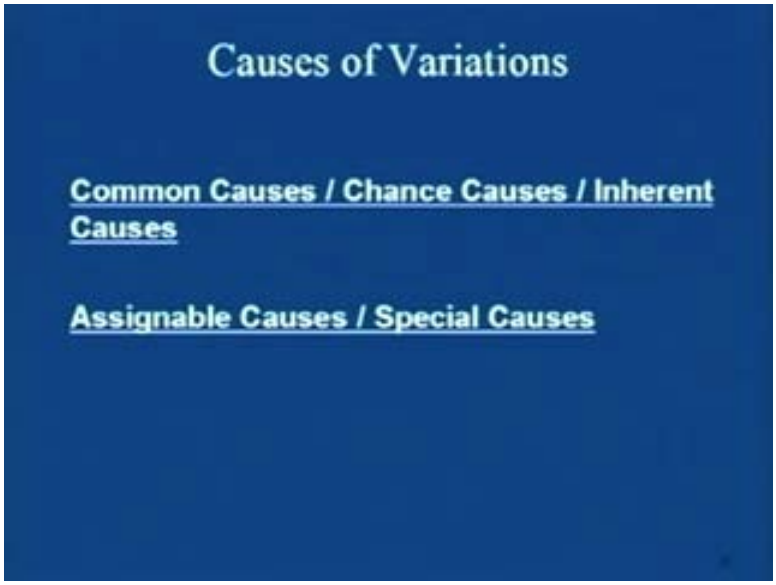
Finally the management, poor process management and poor system, sometime what happens if we are having a product or a product design or product characteristics and we have to develop a tolerance of plus minus 0.001 means up to the three place after the decimal. But, if the machines that the manufacturing section is having, the machines are not capable to produce that tolerances, then every time whatever we efforts we put in we will not get the right product.

So, then this type of cause is basically that occurs, because of the process management, you need to basically involve the management and the management has to look after that if the machines are not there, whatever effort and whatever the pressure we put on we

will not be in a position to produce the right product. Sometime it happens as a part of the process management that we are not treating the next process in the line as an internal customer.

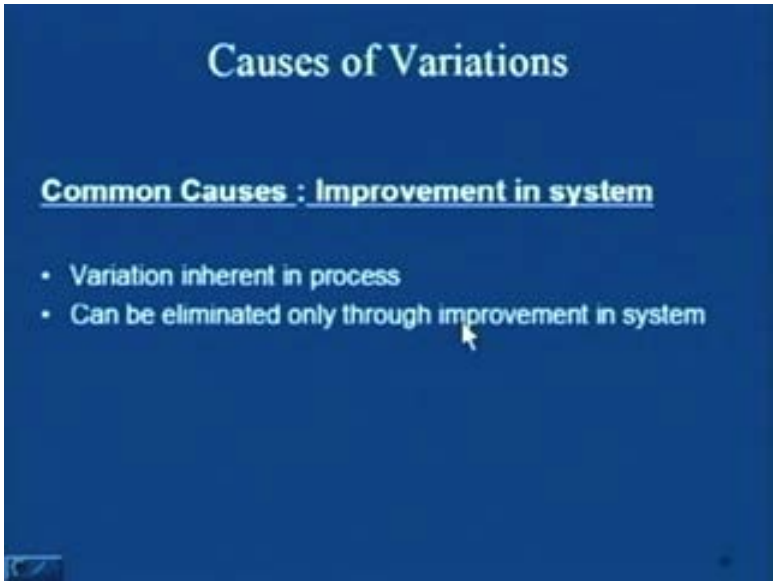
If something wrong happens in the first operation that mistake basically keeps on adding up in the further operations, so every time we should consider the next operation in the line as a customer. So, that whatever we deliver, means whatever we deliver within our organization that delivery should be 100 percent right, so that the things or the things may not get further worsen up. So, these are the potential reasons of variations I am just again highlighting the points, the reasons maybe for operator, the reason of variation maybe the method. Reason of variation maybe the material with reason of variation maybe the machine and the reasons of variation maybe process, poor process management or the management.

(Refer Slide Time: 24:13)



The causes of variations, there are two causes of variation, one is the common causes also known as chance causes or inherent causes, the another second one is the assignable causes, this causes is also known as the special causes.

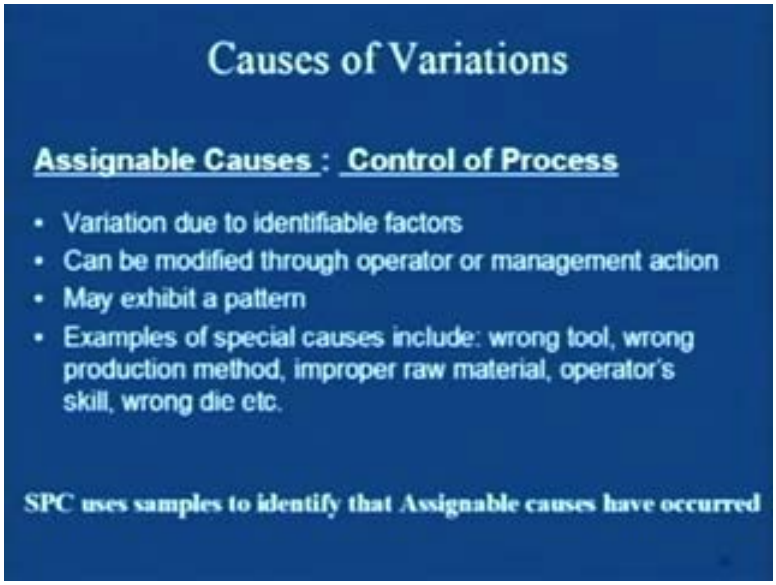
(Refer Slide Time: 24:30)



Let see the first type of cause that is the common causes also known as I mention known as the inherent causes, that basically common causes, when we remove such causes be improve the system. So, whenever we have to bring the improvement in the system, we need to control the common causes, what are the common causes, this is the cause of the variation inherent in the process that is why I mentioned, that this common causes are also called inherent causes.

These causes can be eliminated only through improvement in the system, as I have all ready given the example that if you have the machine and that machine cannot produce the tolerances required by the designer. Then there is no way out you have to have the better machines that can cater or that can take care of those tolerances that the designer is looking forward.

(Refer Slide Time: 25:31)



The second type of cause causes is the assignable causes and this assignable causes if we remove we can control the process, the common causes, if we remove the common causes this the improvement in the system occurs. And if you remove the assignable causes or the special causes, the control of the causes occurs or we can say we can have the stable process these happens, because of the variation due to the identifiable factor.

There is one factor I mentioned that if you go for a casting in the morning and in the afternoon or in the evening in the winter season, then the temperature difference has the effect. So, whenever we see that if we combine the items of morning session and evening session it is clearly visible that these are the variation and we can find out what is the special reason for this. And if we can remove that special reason the process becomes stable or process becomes in control, it can be modified through operator or management action.

We can identify, we can look back if you go to the history we can find out what happened to that particular product, sometime you are using the wrong die that maybe one special reason. Sometime when you are setting the product on a machine the possibility that tool may fell down or the work piece may fell down, so that are the special reasons. So, whenever we find the process is not coming in control or the measurements and not coming in control there is some special reason.

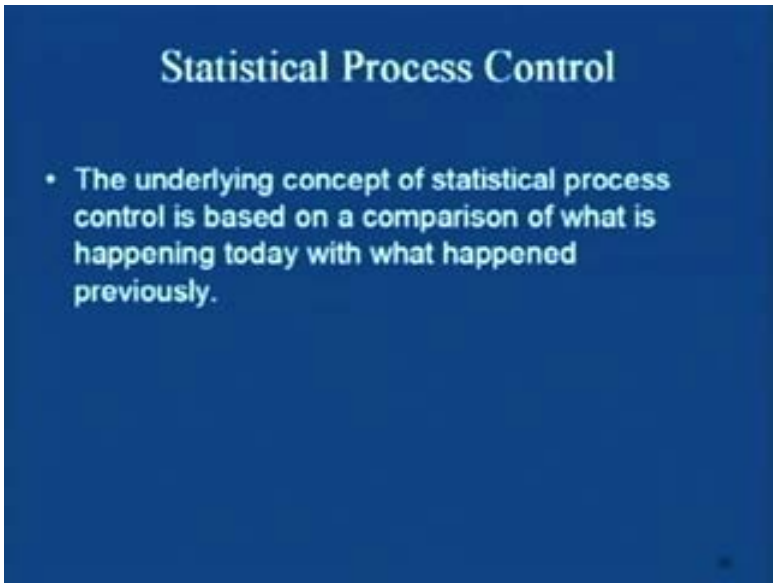
And if we remove the that special reason, means we are controlling the process or controlling the process in such a way that it becomes a stable process or a process in the

state of statistical control, it may exhibit a pattern. Sometime if you are using a tool to turn a particular component on a lathe machine and when the tool gets blunt, the next part will have the higher dimension.

Once the tool is wearing up from one part to the another part, so because of this what happens you are the trend that you are getting it is an increasing trend, every next dimension if we measure is higher than the previous one. So, this whenever you find such reason you say that there is some special reason, that is why it is happening and we need to focus our self to remove such causes. Example, as I mentioned earlier also, example of special causes include wrong tool, wrong production method, improper raw material, operator skills wrong die, etcetera.

We can there is a possibility as one example earlier mentioned that you are using the material, the material is coming from two different sources and if you combine the data with these two different sources, then certainly the variation is bound to come. So, these are the possible reasons and but important part in the assignable reason is that these are the special one and if we remove this causes, we can make our process stable in nature. And statistical process control use this samples to identify that assignable causes have occurred that all ready I have mentioned.

(Refer Slide Time: 28:51)



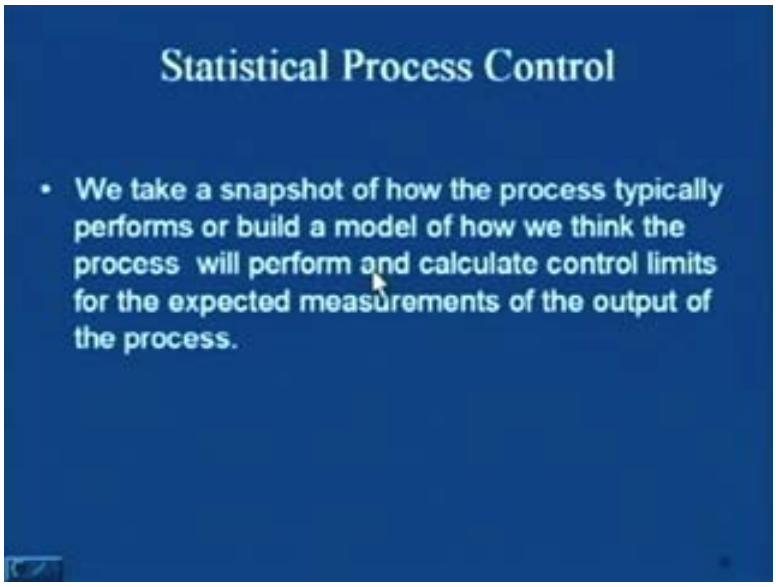
What the statistical process control is, the underlying concept of the statistical process control is based on a comparison of what is happening today with what happened previously. Means we compare the output of the processes with respect to the output that

when the process was in a state of statistical control. So, what we try to do we initially we take, we go for the manufacturing or we go for the measurement of a processes and try to find out that whether the process is in control or not.

And through the removal of the special reason or through the removal of assignable reason we bring back the process in a state of statistical control. Once it is in a state of statistical control, we fix up the upper and lower limits and then we keep manufacturing the things and providing the services in the operations. And later on when the next products that we are manufacturing, we compare the measurement of the next product with respect to the mean value or the upper and lower limits of the process.

So, this is basically the basic concept of the SPC is that, check the what is happening today with what happened previously and we go for the comparison. If what is happening today is exactly or minutely different from the earlier, means we are still in a state of control.

(Refer Slide Time: 30:23)

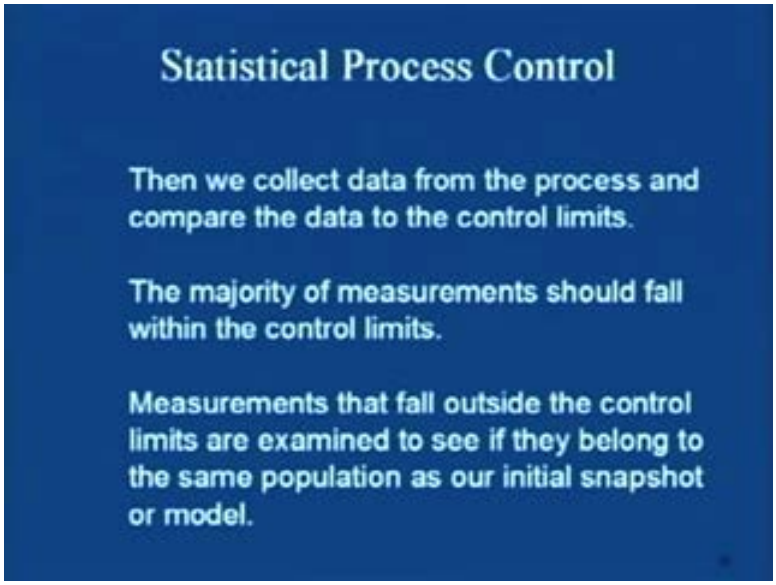


We take a snapshot of how the process typically performs or we build a model of how we think the process will perform and calculate control limits for the expected measurement of the output of the process. This all ready I have mentioned that we from the earlier observations or earlier calculations or a earlier model that we build, we basically calculate the limits, control limits.

So, once the control limits or the limits have been stabilized, then we leave the process alone, because now we understand that process is in a state of statistical control or a

stable process. Then, whatever the next production we do we keep on checking whether the next observation or the next characteristic is coming within the control limits or not.

(Refer Slide Time: 31:09)



So, after taking the snapshot of the model or understanding the process and deciding about the control limits, we collect data from the process and compare the data to the control limits. The majority of the measurement should fall within the control limits, measurements that fall outside the control limits are examined to see if they belong to the same population as our initial snapshot or the model.

(Refer Slide Time: 31:40)

The next concept is the statistical quality control, this is entirely different then the SPC here in statistical quality control, the purpose or the basic thing is to ensure in a cost efficient manner that the product shipped to customer meets their specification. Means this is more on the inspection, but we have done earlier.

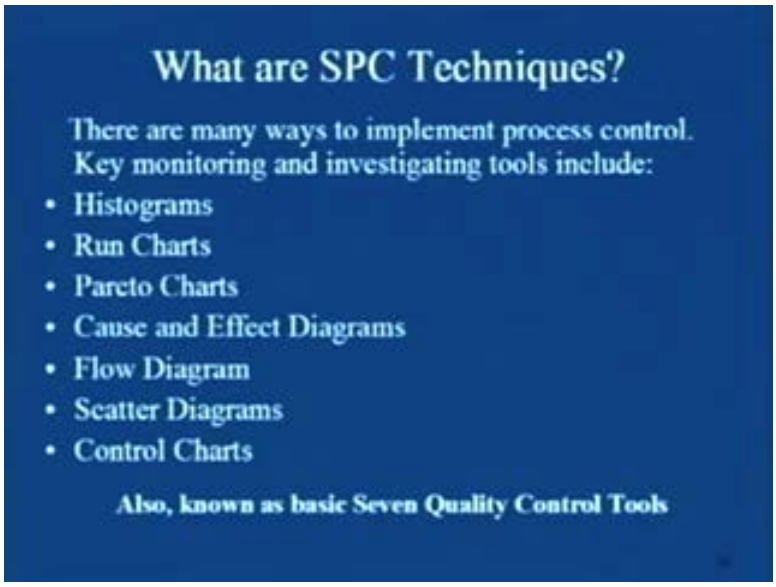
(Refer Slide Time: 32:04)

Inspecting every product is costly and ineffective, but consequences of shipping none confirming product can be significant in terms of customer dissatisfaction. If you are sending the right product to the customer, customer will come back to you, if the product is not right, not meeting the specification, not meeting the performance then your product will come back. So, if we are not producing the right products, not producing the delivering the right services means you are losing the goodwill, this already I mentioned, earlier when we defined the quality and the concept of the quality.

(Refer Slide Time: 32:47)

So, the statistical quality control it is the process that ensures or it is the process of inspecting enough product from the given lots to ensure a specified quality level. So, this SQC basically ensures that the sampling process of that whatever the outgoing quality is that outgoing quality should be sufficient enough that the production is consistent or the finish product production is consistent. And the right product is being delivered to the customer and customer is satisfied with the product.

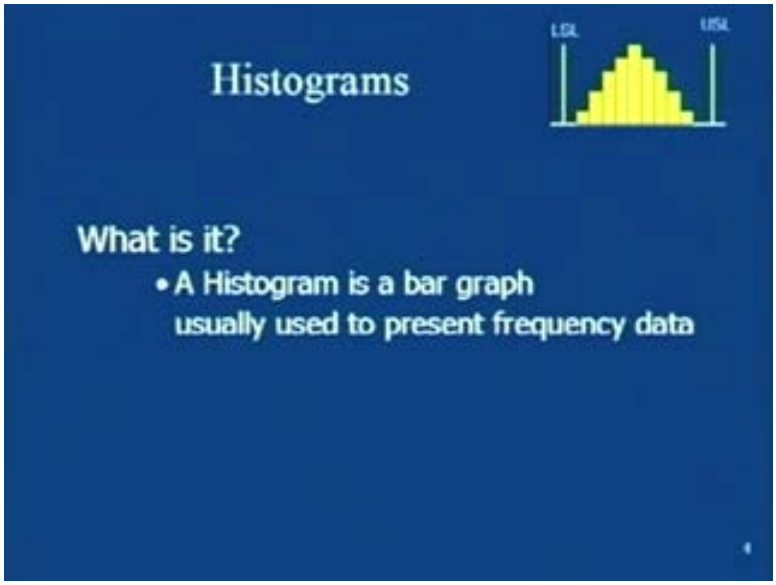
(Refer Slide Time: 33:33)



The tools and the techniques of the statistical process control are, there are many ways to implement the process control, but the key monitoring and investing tools include the

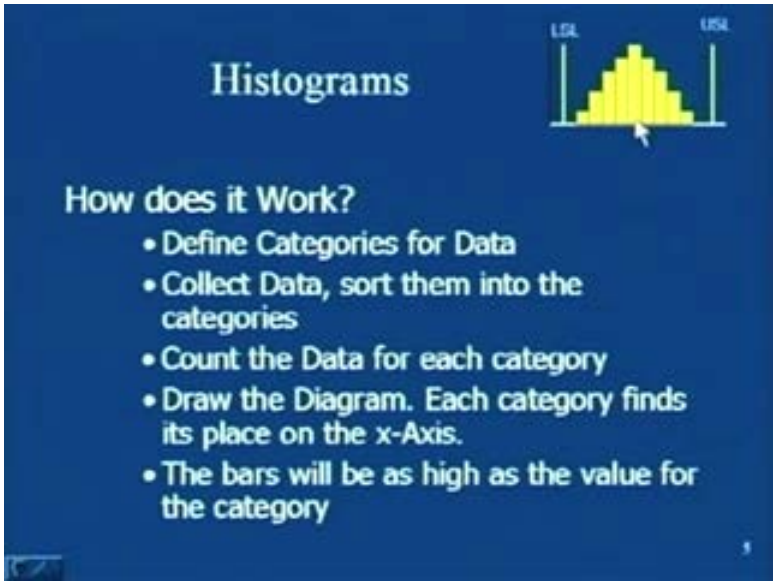
histograms, run charts, pareto charts, cause and effect diagrams, flow diagram, scatter diagram and the control charts. These 7 tools are also known as the 7 quality control tools sometime it is are they are also known as the 7 basic quality control tools, let us discuss these tools in detail.

(Refer Slide Time: 34:21)



The first tool in this category is the histogram, what is it a histogram is a bar graph usually used to present the frequency data.

(Refer Slide Time: 34:29)

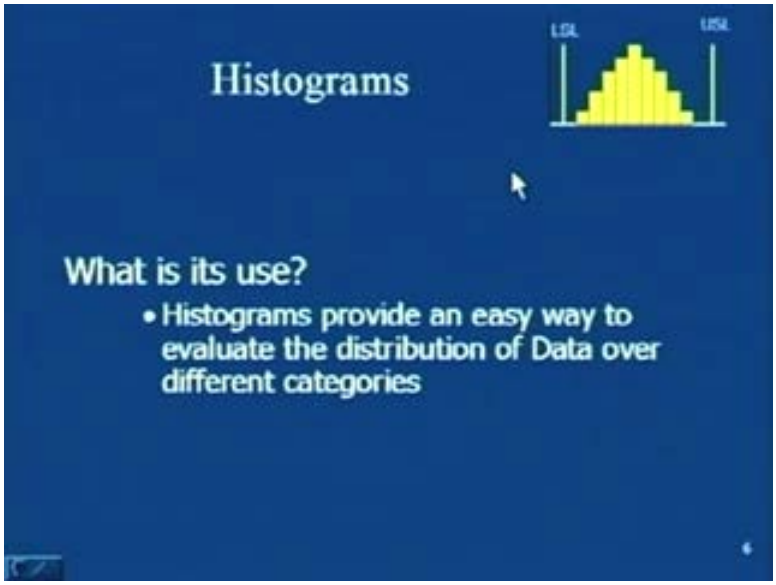


How does it work, what we do we define categories for the data, if we have lot of data available then we basically define the categories. Say for example, if the data ranges

from 10 to 1000, then we define the categories that the data from 10 to 19, 20 to 29, 30 to 39, so first define the categories, then collect data sort them into the categories. So, once the data has been sorted down, then we calculate count the data for each category that is the frequency, that how many numbers, how many data's are coming in each category.

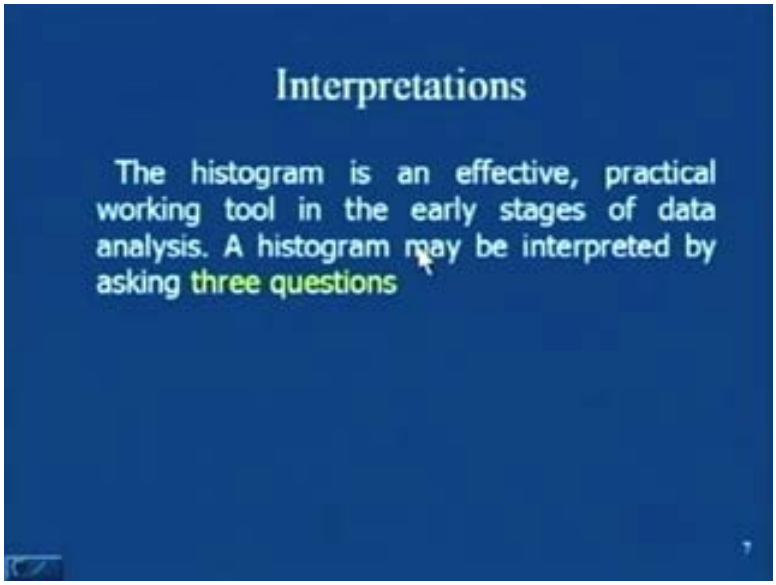
Then we draw the diagram each category finds in place on the x axis, the bars will be as high as the value for the category and if you see that histogram is shown here, that these are the different bars and this is the first category, second, third, fourth, fifth the categories are mentioned. And the height of the bar basically represents the value of that category, means how many observations are falling in one category, so if you see this that it is very clear, that maximum number of observations are falling in this category.

(Refer Slide Time: 35:49)



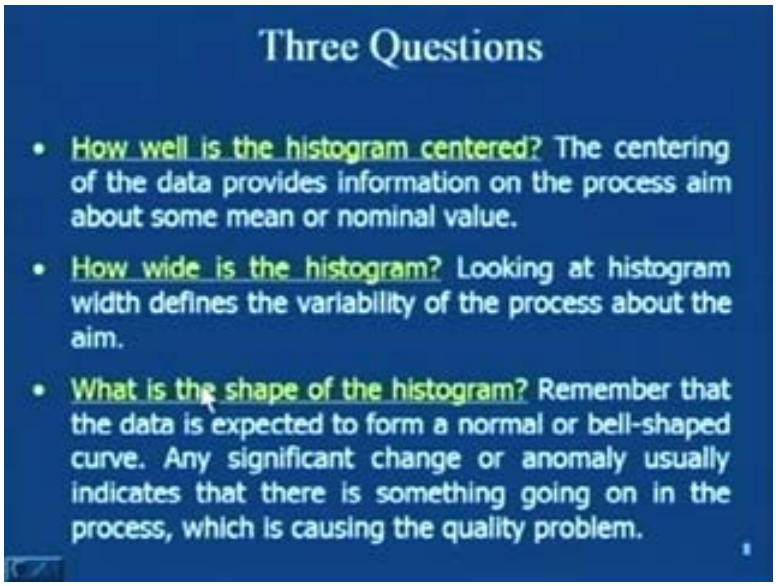
What is it is use, histogram provides an easy way to evaluate the distribution of data over different categories.

(Refer Slide Time: 36:00)



How to go for the interpretations of the histogram, the histogram is an effective, practical working tool in the early stages of the data analysis. Whenever we need to in go for the interpretation of the histogram, we should ask three questions.

(Refer Slide Time: 36:21)

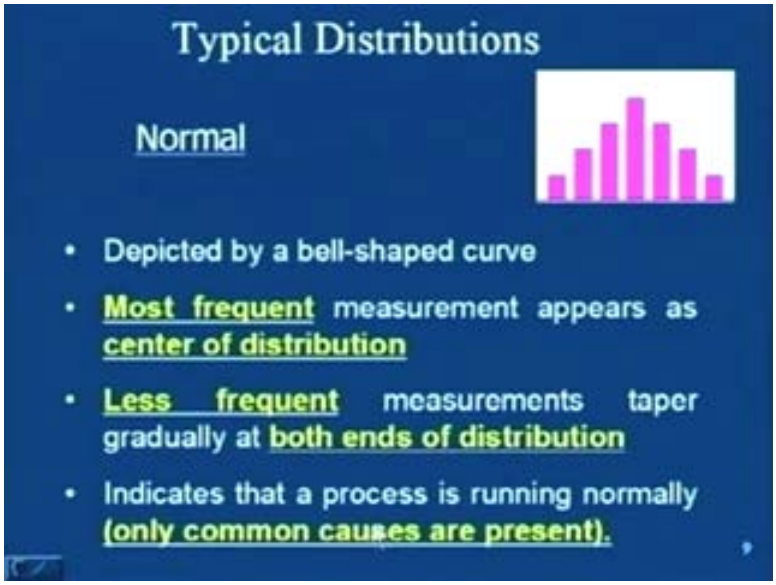


The first question, how well is the histogram centered, the centering of the data provides information on the process aim about some mean or nominal value. Means this histogram basically provides where our the mean or the target or the average value is lies. Second question we need to ask how wide is the histogram, this wideness of the histogram basically tells you about the variability, looking at histogram width defines the

variability of the process about the aim; aim is basically a target value or the mean value or the normal value.

The third question is what is the shape of the histogram, as earlier I mentioned that most of the observation follows the normal distribution, if the observations has been taken independently and in unbiased way. Remember that, the data is expected to form a normal or bell shaped curve, any significant change or anomaly usually indicates that there is something going on in the process, which is causing the quality problem. If anyhow we do not get the normal bell shaped curve, some different shape this is a clear indication that something is going wrong with the process or with the services that we are providing.

(Refer Slide Time: 37:49)

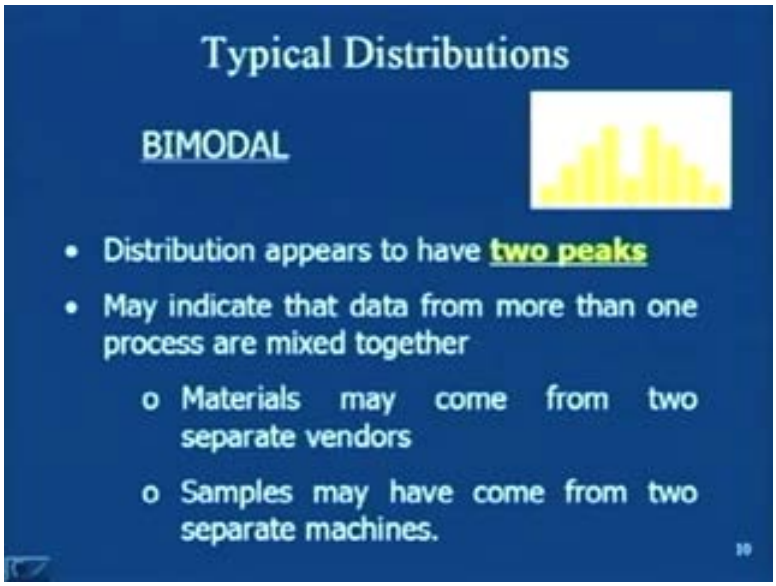


Some typical distribution, this is the normal distribution that we have talked about a lot earlier also, that it is depicted by a bell shaped curve means there is only one peak, there are two at the centre. And then observations are gradually decreasing on both the side of the mean, most frequent measurements appears at the centre of the distribution in the bell shaped curve.

Less frequent measurements taper gradually at both the ends of the distribution, that we have talked about that on both the sides these are the less frequent measurements and they are basically at both the ends of the distribution. Indicates that a process is running normally, means when you get the normal distribution of your data that clearly indicates that the process is running normally, means there is no special reason, no special causes

of variation or no assignable causes there only common causes are present. Means whatever the variations are there, inherent variations are in there only those variations are present in the data, the another distribution is the bimodal.

(Refer Slide Time: 39:01)



If you see this figure that it shows very clearly that there are two peaks, means when instead of one peak we get the two peaks that type of distribution is called the bimodal and distribution appears to have two peaks, it may indicate that data from more than one process are mixed together. So, this happens because of the many reasons, one reason is mentioned here, materials make come from two separate vendors.

And when we collect the data from the processes and in the history of the processes the batches are coming from the purchasing section, the material is coming from the purchasing section or the from the inventory. There is a possibility that one lot of the material is coming that is basically belongs to one supplier and the another lot belongs to the another supplier.

If we are combining the data of the material that is coming from two different supplier, there is a great possibility that we will get that this type of distribution that is having two peaks. There is another possibility that the samples may have come from two separate machines, means when we are collecting the data from two separate machines and we are clubbing the values or the observations of these data.

There is possibility also that, sometime when we collect the data we never bother about when this material has been or when this observations have been taken. Sometime we

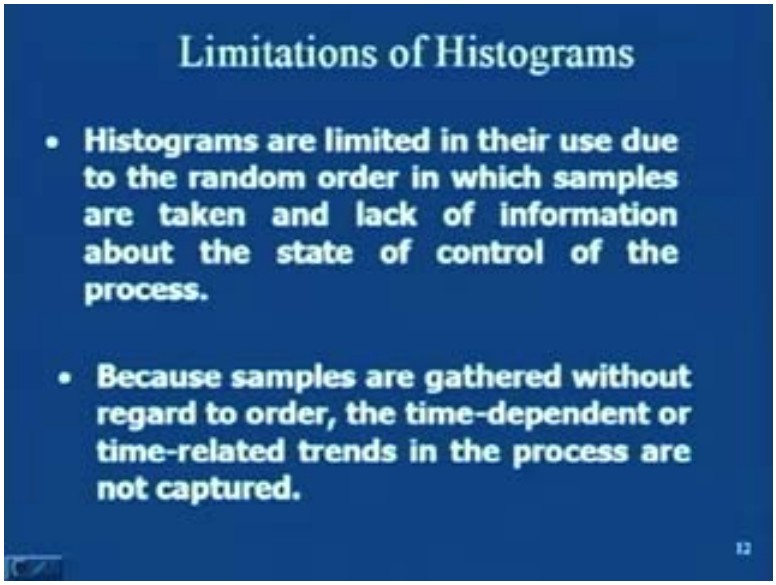
club the data of the two different operators, sometime we club the data of two different operator working in different machines, sometime we collect the data when the different operators are working in different shifts.

So, the question point here is anytime we collect the data with care should be taken into account, that the data should be from the single source, same source and from the same operator, same machine. And all care should be taken under account that the same measurement method in all the data, same tools, same material and all the properties and characteristic should be same, only then the data will behave like a normal distribution.

(Refer Slide Time: 41:23)

The third distribution is the cliff like and if you see here in the cliff like distribution that data is decreasing and this is a clear indication that once the data has been collected, some sorting has been done. The data has been arranged in the descending order, so that appears to end sharply or abruptly at one end, this indicates possible sorting or inspection of non-conforming parts.

(Refer Slide Time: 41:55)

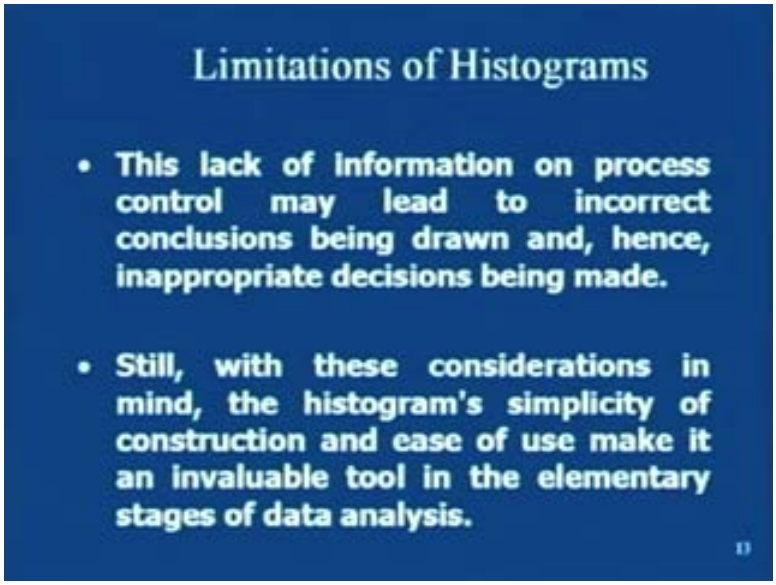


What are the limitations to histogram is a very simple tool we use in our daily life, but still because simply just saying the histogram, we know only the centering of the process we know the spread of the process by just understanding the width of the histogram. And only information we can get that the process is behaving normally and with that particular average value or the mean value as well particular width or particular value of the variability.

The limitation of the histogram that histograms are limited in their use due to the random order in which samples are taken and lack of information about the state of control of the process. So, once we have the data and then when we convert into the categories we count the frequency, but later on once the histogram is ready, then if we want to know something that what happened to that particular observation that observation is lost.

Only the information you are getting it is the data has been arranged in 5 for 10 categories and the every category has this value, so that data information is lost this is one of the limitation. The another limitations is that because samples are gathered without regard to order, the time dependent or time related trends in the process are not captured, means when the first data has been taken and when the next data has taken. So, this basically the time dependent activities or time dependent trends in the process is not captured and these are the limitation of the histogram.

(Refer Slide Time: 43:41)

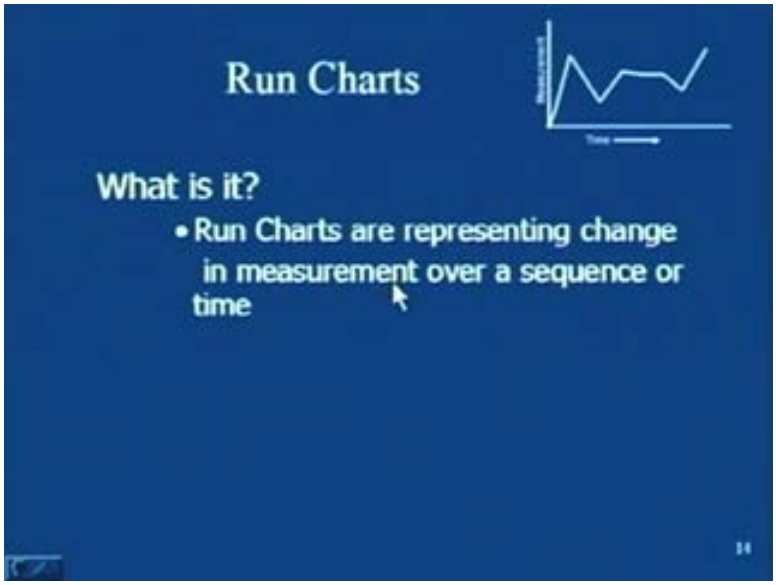


This lack of information on process control may lead to incorrect conclusions being drawn and hence, inappropriate decisions being made. Because, once we do not know when the data has been collected we cannot go back to find out from the history records about the product or about the processes, what happened to the process, why the process is not be having normally. But, still with these consideration in mind the histogram simplicity of construction and ease of use make it and invaluable tool in the elementary stages of the data analysis.

So, this although it has the limitation, but the histogram is very simple tool once we do not know anything, once we have collected the data the very first point is that, if you do not know any analysis first try to develop the histogram. Once the data is converted into the histogram it gives you some information about the process, that how the process entered whether the process is behaving normally or not.

What is the extent of variability or the deviations among the data is there from the mean value, so that information we can get from the histogram. The limitation that I mentioned earlier in the histogram that is basically the time dependent activities are not recorded, we cannot obtain the information with the respect to the time regarding the data that we are collecting, this limitation is over come in these next chart that is the run chart.

(Refer Slide Time: 45:19)



And run chart is very common, if you see the figure here, that on the one axis we have the time on the another axis we have the measurement. So, anytime we get the data we just note down the time and the data, next we collect note down the time and collect the measurement value. Like that we keep on just noting the time and at that time what is the measurement value and we just join the data with this straight line and the run chart is ready to use.

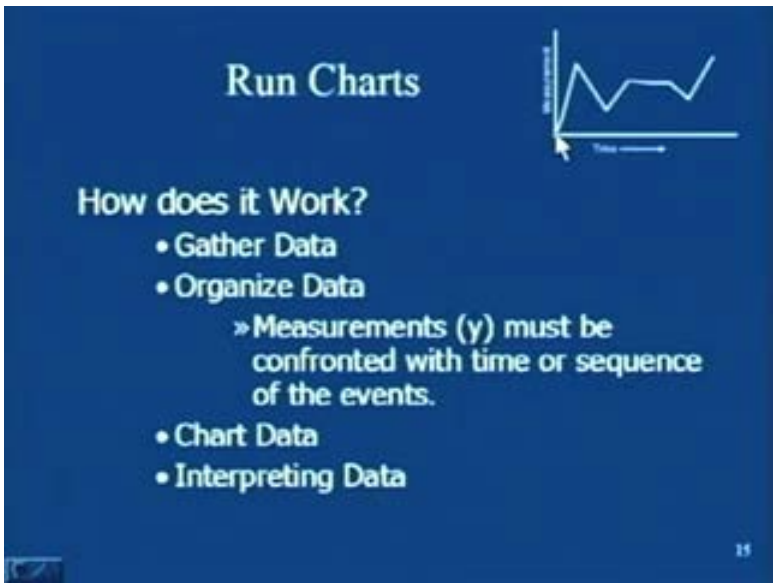
And if you see this data, that this is the run chart and these data point basically is being joint from the previous data with a line and we know all the values, if you take the average of all this value you can draw a horizontal line here. So, once we draw a horizontal line here that basically tells us the average of all the data, so average value is also be this obtained in the run chart.

And the minimum value and the maximum value we also know and once we know the minimum value and the maximum value of all the observation, we can find out what is the range, what is the variability with respect to the range of the process. So, whatever the information we are getting in the histogram that information we can also get in the run chart, because we can get the mean in the histogram.

The mean value we can get from the average of all these data and the minimum and maximum value we are getting from here. So, the variability or the width of the data we can obtain, but the additional benefit is that we are also getting the time dependent

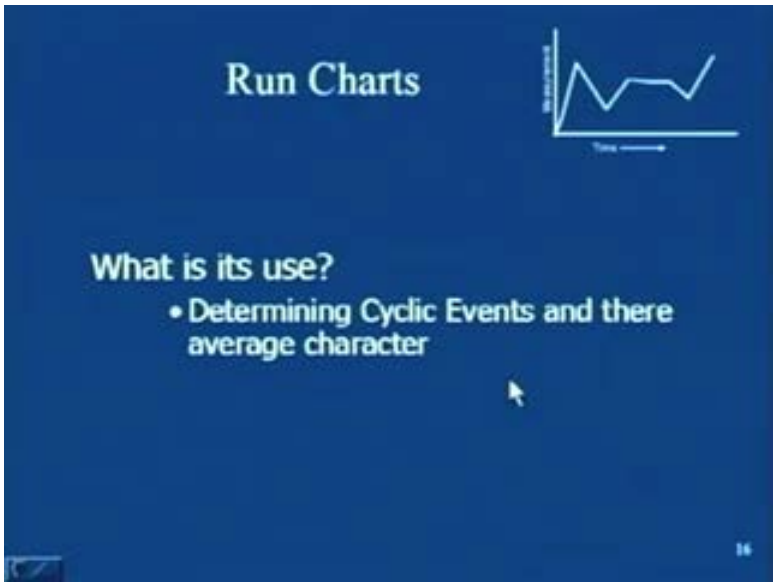
activities at what time the observation was there. Now, coming to this what is it, run charts are representing change in measurement over a sequence of time.

(Refer Slide Time: 47:19)



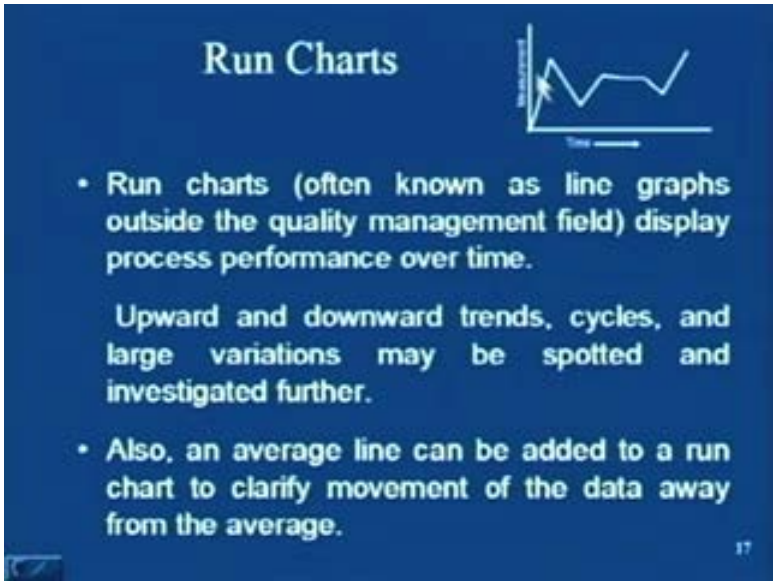
How does it work, it is very simple I earlier mentioned first gather the data organize the data, measurement y that is on the x axis, y axis must be confronted with time or sequence of the event. It is not always necessary that x axis should be time, means the first observation, second observation, third observation, fourth, so we can write down the order of the observations, then chart the data and go for the interpretation of the data.

(Refer Slide Time: 47:51)



What is it use, apart from the information that we get from the histogram we also get the ordered information, so the use is determining cyclic events and there average character.

(Refer Slide Time: 48:09)



Run charts often known as line graph outside the quality management field, it displace the process performance over time, upward and downward trends, cycles and large variations maybe spotted and investigated further. So, this is the additional benefit apart from the order of the observation that we can also, these run chart also reveals the trends, trends maybe the increasing, trend decreasing, trend increase in the dimension from one observation to the next.

Maybe some repetitive cycles are there that after every 10 observation, thus that observations values keeps on repeating and maybe some large variations, because of the ambient conditions that can be spotted by investigating and we can investigate further why this is happening. Also that I all ready mentioned an average line can be added to a run chart to clarify movement of the data away from the average, so if you know the average value we can draw a horizontal line, you can find out that how the data is deviating from the mean value.

(Refer Slide Time: 49:21)

Two ways to misinterpret the run chart, what happens that we conclude that some trend or cycle exist, when in fact we are just seeing the normal process variation and every process will show some variation this type of misinterpretation in the quality control field we say the type one error. That when even thing is behaving normally, means whatever the data we are collecting it follows the normal distribution.

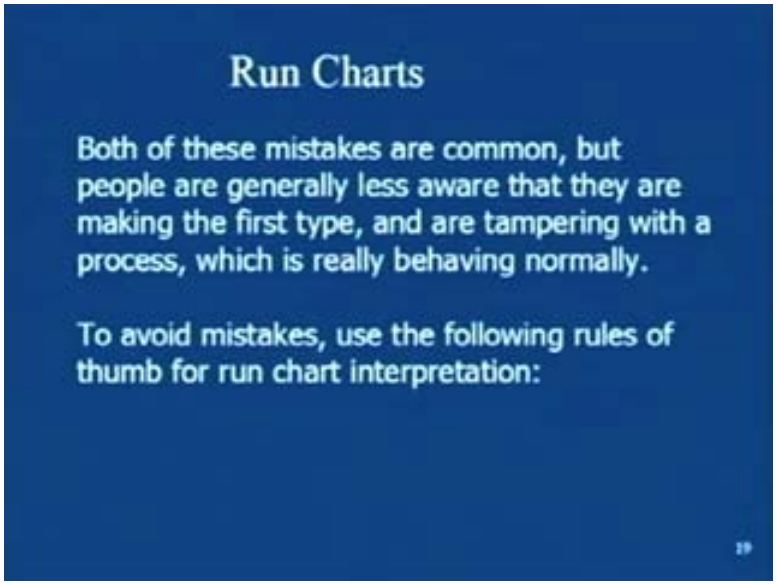
And by chance randomly if you just collect take the one observation, it falls beyond the three sigma control limit, then we say that process is not behaving although the process is behaving right. And sometime just collecting the seminal rate of 10 observations, you see that there is a some trend or some cycle exist actually there is no cycle or no trend is existing. So, this is very common misinterpretation of the run chart that we conclude something about that is something wrong is going on, although there is nothing wrong in the process.

The other error is sometime we do not recognize a trend or cycle although it exists, means process is not behaving right, there are cycles trend and in problems are there in the process, but we are not able to recognize from the run chart. Means sometime what happens, we have very small little information or few observations there and if you plot 5, 4, 6 observations, those observations are not sufficient enough to reveal whether there is cycle or trend exist.

So, this is the another observation and we normally says it is type two error, then there is a wrong in the processes, process has gone wrong, but we still we are concluding that

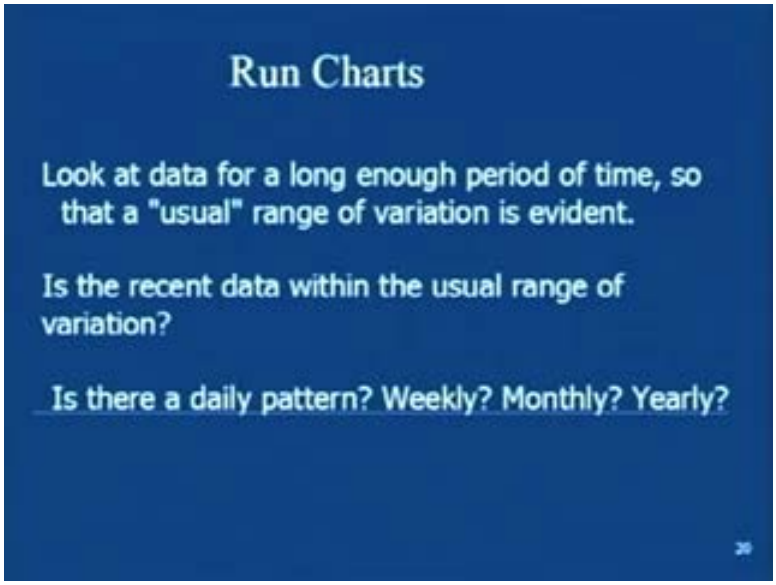
process is working fine that is the type two process. So, again I am repeating type one error basically is that when processes is in control everything is there right, but we are concluding that something wrong is there. And another type two is that when the things have gone wrong, but still we are concluding that process is behaving correct.

(Refer Slide Time: 51:36)



So, both of these mistakes are common, but people are generally less aware that they are making the first type, and are tampering with a process which is really behaving normally. So, this is very common mistake and possibilities are there, when we just decide or conclude that something is wrong although the process behaving normally, we try to correct the process, correct the correct process basically. To avoid such mistakes, use the following rules of thumb for run chart interpretation.

(Refer Slide Time: 52:14)



The first rule of thumb is that look at the data for a long enough period of time, so that a usual range of variation is evident, never conclude just by collecting very few data and taking the wrong decision. So, first let the process go and take the sufficient information or sufficient observation, then conclude it, is the recent data within the usual range of variation, also have an eye that whatever the data you are obtaining that is coming in the usual range of the variation.

Is there is a daily pattern, weekly pattern, monthly pattern or the yearly pattern, so these are the three rules of the thumb rules that you need to ask yourself and you need to check yourself. At look at the data for a long enough for the time, check is the recent data within the usual range of variation or the third point is there daily pattern, weekly pattern, monthly pattern or yearly pattern we are looking around.

Here we are finishing the part two of statistical process control right at this point and we will continue with the rest of the static 7 basic quality control tools in the third part of the statistical process control. Today in this part 2 of the statistical process control we have learnt about the concept of the variation, we also had talked about the potential reasons of the variation, causes of the variation.

And we have also got the conceptual background of SPC that is the statistical process control, the conception background of statistical quality control SQC. And the 7 basic tool and out of the 7 basic tool we discussed in detail about the histogram and as well as the run charts.

Thank you viewers.