

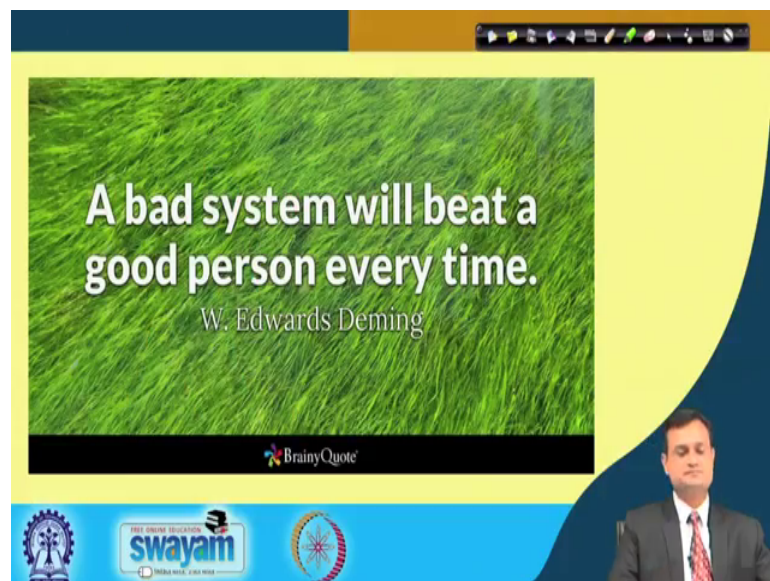
Six Sigma
Prof. Jitesh J Thakkar
Department of Industrial and Systems Engineering
Indian Institute of Technology, Kharagpur

Lecture - 27
Non - Normal Process Capability Analysis

Hello friends, once again I welcome you to our ongoing journey on Six Sigma and today we are discussing lecture 27 that is Non-Normal Process Capability Analysis. We have devoted three lectures process capability analysis for normal and we had seen the first lecture; the concepts related to process capability what is its need, what does it tell and what are the various cases three cases when my process is meeting, barely meeting, the specifications, my process is barely capable rather my process is not capable and my process is adequately capable.

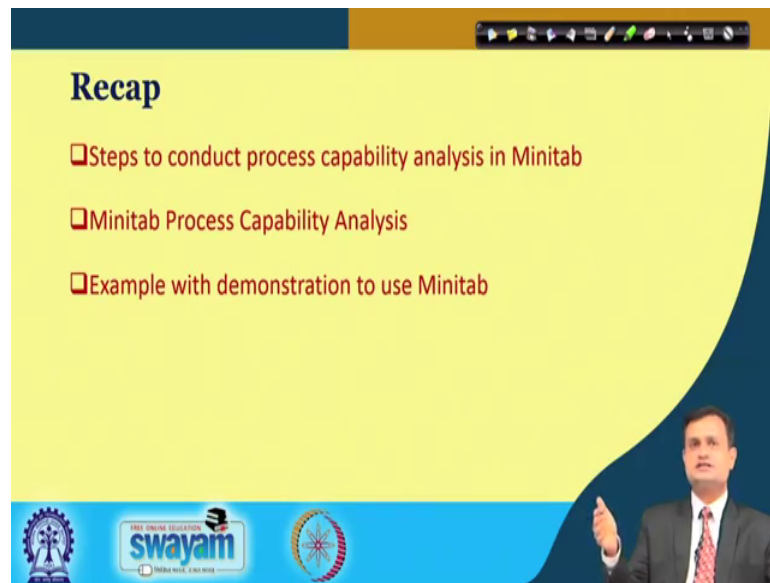
And then we have seen the various indices and followed by that we have seen the application of mini tab for conducting say process capability analysis for normal data. This lecture 27 will help you to understand: what is non normal process capability, when it should be conducted, how it should be carried out and how we can use the mini tab to easily conduct this non-normal process capability analysis.

(Refer Slide Time: 01:42)



So, once again I would like to remind a bad system will beat a good person every time. So, correct the system first.

(Refer Slide Time: 01:53)



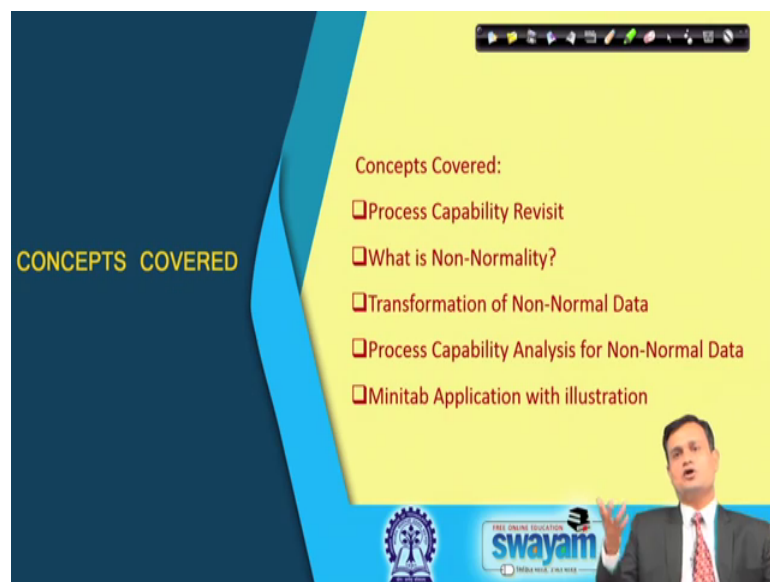
The slide is titled "Recap" in a bold, dark blue font. It lists three bullet points in red text: "Steps to conduct process capability analysis in Minitab", "Minitab Process Capability Analysis", and "Example with demonstration to use Minitab". The slide has a yellow background with a dark blue curved border on the right. At the bottom, there is a blue banner with logos for "swayam" and "INDIA WISE, SKILL WISE". A small inset image of a man in a suit is visible in the bottom right corner.

Recap

- ❑ Steps to conduct process capability analysis in Minitab
- ❑ Minitab Process Capability Analysis
- ❑ Example with demonstration to use Minitab

So, we have seen this that what is the use of mini tab, how it can be conducted. There is six pack process capability analysis and there is simple capability analysis and what are the outputs it can generate and how to interpret these outputs.

(Refer Slide Time: 02:10)



The slide is titled "CONCEPTS COVERED" in a bold, yellow font. It lists five bullet points in red text: "Process Capability Revisit", "What is Non-Normality?", "Transformation of Non-Normal Data", "Process Capability Analysis for Non-Normal Data", and "Minitab Application with illustration". The slide has a yellow background with a dark blue curved border on the left. At the bottom, there is a blue banner with logos for "swayam" and "INDIA WISE, SKILL WISE". A small inset image of a man in a suit is visible in the bottom right corner.

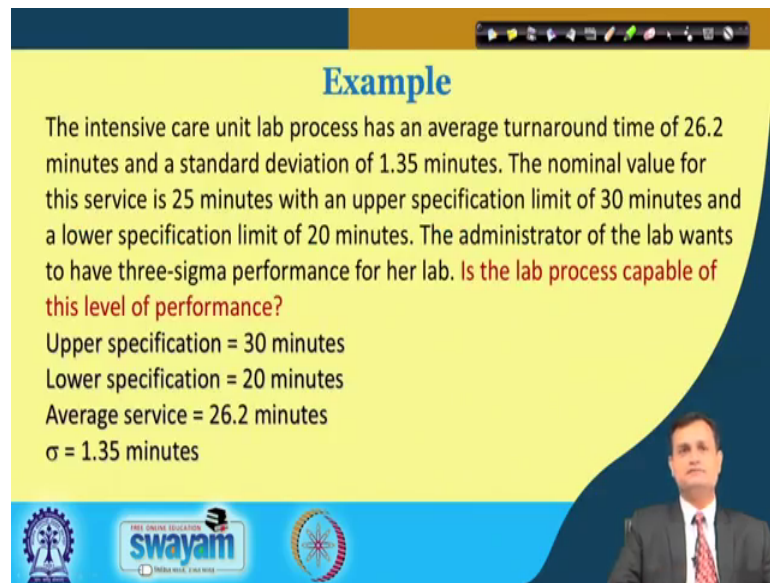
CONCEPTS COVERED

Concepts Covered:

- ❑ Process Capability Revisit
- ❑ What is Non-Normality?
- ❑ Transformation of Non-Normal Data
- ❑ Process Capability Analysis for Non-Normal Data
- ❑ Minitab Application with illustration

This particular lecture 27 will focus on process capability revisit just a small example. What is non-normality? Transformations of non normal data; what are the ways you can conduct the process capability analysis and mini tab application with illustration.

(Refer Slide Time: 02:36)



Example

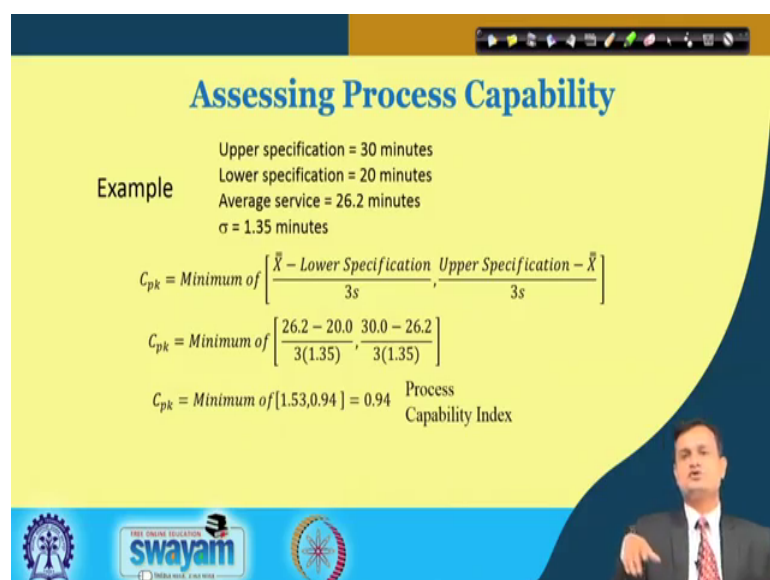
The intensive care unit lab process has an average turnaround time of 26.2 minutes and a standard deviation of 1.35 minutes. The nominal value for this service is 25 minutes with an upper specification limit of 30 minutes and a lower specification limit of 20 minutes. The administrator of the lab wants to have three-sigma performance for her lab. **Is the lab process capable of this level of performance?**

Upper specification = 30 minutes
Lower specification = 20 minutes
Average service = 26.2 minutes
 $\sigma = 1.35$ minutes

The slide features a yellow background with a blue header and footer. The footer includes logos for 'swayam' and other educational institutions. A presenter in a suit is visible in the bottom right corner.

So, if we just revisit our concept with an example, let us say there is an intensive care unit lab process has an average turnaround time of 26.2 minute in standard deviation 1.35 minutes. The nominal value for this service is 25 minutes with an upper specification 30 minutes and lower specification 25 minutes and sorry 20 minutes, the administrator of the lab wants to have three sigma performance for her lab and is the lab process capable of this level of performance that is the issue that I want to investigate. Now your upper specification limit USL 30 LSL 20, average service 26.2 and sigma 1.35 minutes.

(Refer Slide Time: 03:26)



Assessing Process Capability

Example

Upper specification = 30 minutes
Lower specification = 20 minutes
Average service = 26.2 minutes
 $\sigma = 1.35$ minutes

$$C_{pk} = \text{Minimum of } \left[\frac{\bar{X} - \text{Lower Specification}}{3s}, \frac{\text{Upper Specification} - \bar{X}}{3s} \right]$$
$$C_{pk} = \text{Minimum of } \left[\frac{26.2 - 20.0}{3(1.35)}, \frac{30.0 - 26.2}{3(1.35)} \right]$$
$$C_{pk} = \text{Minimum of } [1.53, 0.94] = 0.94$$

Process Capability Index

The slide features a yellow background with a blue header and footer. The footer includes logos for 'swayam' and other educational institutions. A presenter in a suit is visible in the bottom right corner.

So, I will just plug in the values into my expression for C_{pk} and it basically tags the minimum of \bar{X} minus lower specification divided by $3s$ and upper specification minus \bar{X} divided by $3s$. So, computing this I get C_{pk} minimum of 1.53 and 0.94. So, I will check 0.94.

(Refer Slide Time: 03:58)

Assessing Process Capability

Example $C_p = \frac{\text{Upper Specification} - \text{Lower Specification}}{6\sigma}$

$C_p = \frac{30 - 20}{6(1.35)} = 1.23$ Process Capability Ratio

Does not meet 3σ ($1.00 C_{pk}$) target due to a shift in mean (Note variability is ok since C_p is over 1.0)

Before Process Modification

Upper specification = 30.0 minutes Lower specification = 20.0 minutes

Average service = 26.2 minutes

$\sigma = 1.35$ minutes $C_{pk} = 0.94$ $C_p = 1.23$

The slide includes logos for Swamyam and other educational institutions at the bottom.

So, this is my process capability index. I would also like to find the C_p which basically measures the spread and variability of the process. So, USL minus LSL divided by 6 sigma and I get 1.23. So, here I can say that does not meet 3 sigma which is typically 1 C_{pk} target due to shift in mean. And if we note here, then variability part is C_p is ok it is greater than 1, but my C_{pk} has a problem and typically my C_{pk} value is my C_{pk} value here is 0.94.

So, little bit less we have the standard that, it should be around 1.2, 1.3 or at least 1 and in this case my variability is C_p is ok, but there is a shift in mean and hence my C_{pk} has some problem.

(Refer Slide Time: 05:19)

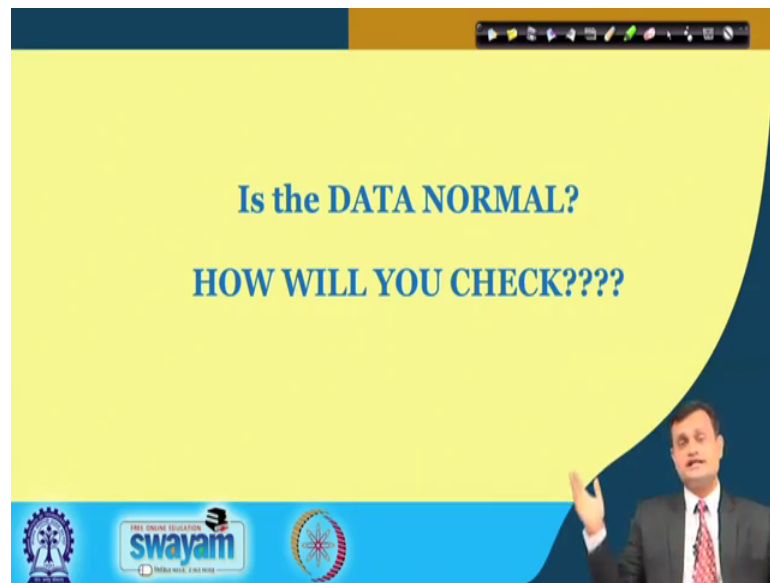
The slide is titled "Non-normality in the data" in blue text on a yellow background. It contains two bullet points: "▪ Cpk is an index much misused: It is only applicable to normally distributed variable data that is measured in the short term." and "▪ Capability studies for non-normal data (Weibull) do not give Cpk values". In the bottom right corner, there is a small video feed of a man in a suit and tie. The bottom of the slide features logos for "swayam" and other educational institutions.

- Cpk is an index much misused: It is only applicable to normally distributed variable data that is measured in the short term.
- Capability studies for non-normal data (Weibull) do not give Cpk values

So, you have the upper specification 30 lower specification 20, average service 26.2 sigma is 1.35, Cpk 0.94, C p is 1.23 before any modification now, you need to take the corrective action short term or long term. As I mentioned short term you can just shift the mean and try to avoid producing the scrap that will give you the cost saving, but this is just the short term temporary measure. For long term you try to replace the part, set the process, develop the skill of the operator and see that your process can remain at a particular targeted value for a long operation period of time. Now let us talk about the non-normality in the data which is the focus of this particular lecture.

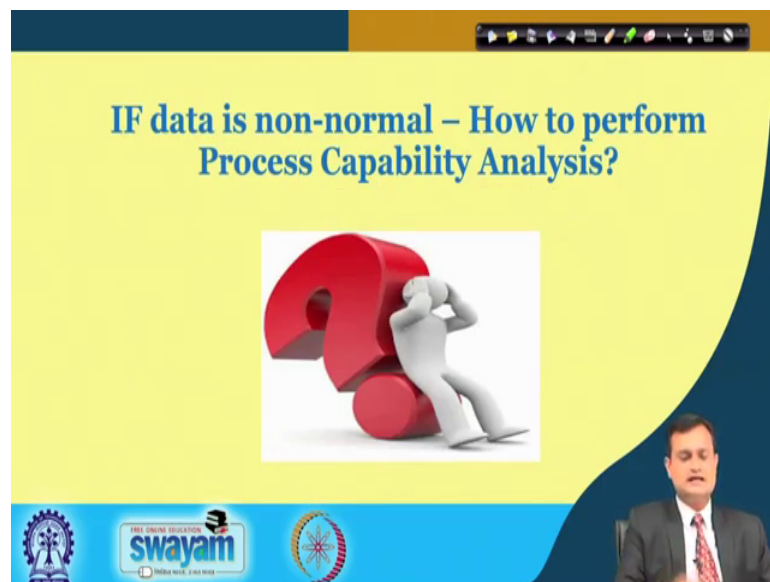
So, Cpk is an index which is many a times misused. Remember that this index is only applicable to normal distributed, where normally distributed variable data and this is typically measured for short term. Capabilities studies for non-normal data, let us say Weibull distribution do not give Cpk values and if you will just try to fit let us say this nonnormal data and try to figure out or find out the values of C p and Cpk then it would be misleading misleading. So, my assumption about the normality of the data must be checked first and then only the question of process capability analysis comes. If the data is non-normal you must conduct the non-normal process capability analysis, if it is normal then you go for normal process capability analysis.

(Refer Slide Time: 07: 06)



So now, the question is that, if I have the data non-normal or if I have to check rather my data is normal or not normal then how will I check?

(Refer Slide Time: 07:20)



So, there is a procedure and that can help us to check whether my data is normal or not normal, non-normal and what to do. So, if data is non-normal how to perform process capability analysis that is the second question. So, you will just struck your head and think that now my data is not-normal what to do.

(Refer Slide Time: 07:44)

IF data is non-normal – How to perform Process Capability Analysis?

We have two options:

Option 1: Find the best fit distribution (data—exponential, chi-square, Weibull etc.) and perform directly the process capability analysis in Minitab under “Capability Analysis (Non Normal).”

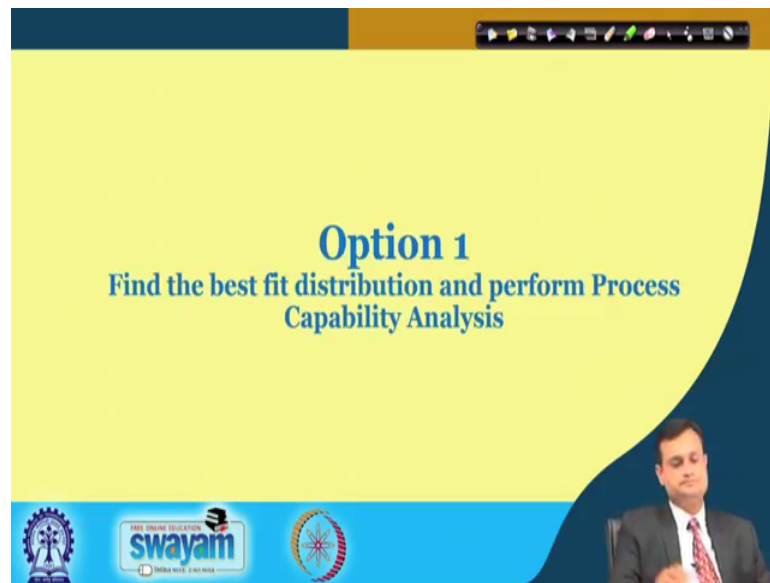
Option 2: Transform the NON-NORMAL data into NORMAL using transformation and perform the process capability analysis for normal data.

Logos at the bottom: Swamyam, Free Online Education, and a circular logo.

So, if data is not normal, it does not follow typically the normal pattern when you look at the histogram or normality plot, then how to perform is a question. So, here I would recommend that there are two options possible; option number 1 find the best fit distribution let us say your data fits to exponential, chi square, Weibull and perform directly the process capability analysis in mini tab under capability analysis non normal. So, this provision is that, but for that you need to figure out that what is the best fit distribution for your data.

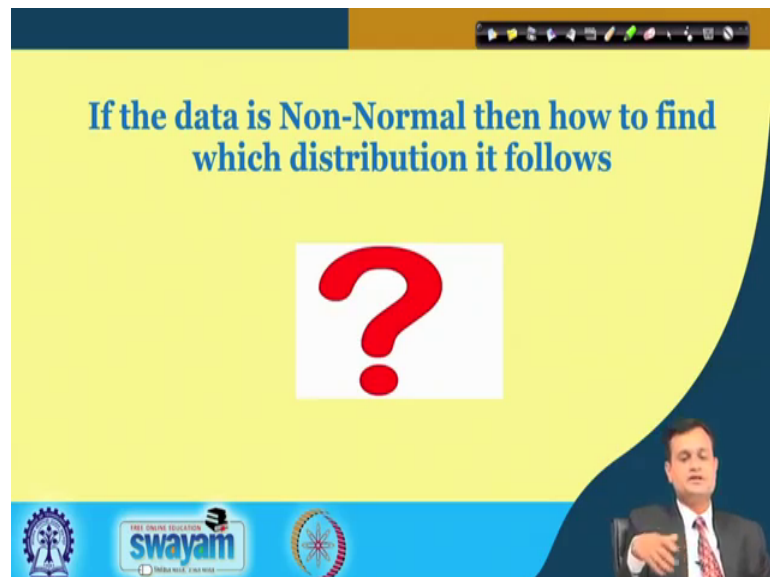
If you do not want to do this, the option 2 is to transform your data which is not normal into normal using some kind of transformation, mathematical transformation and then perform the process capability analysis with the simple normal data assumption. And, you can get the same result both way either option 1 option 2.

(Refer Slide Time: 08:58)



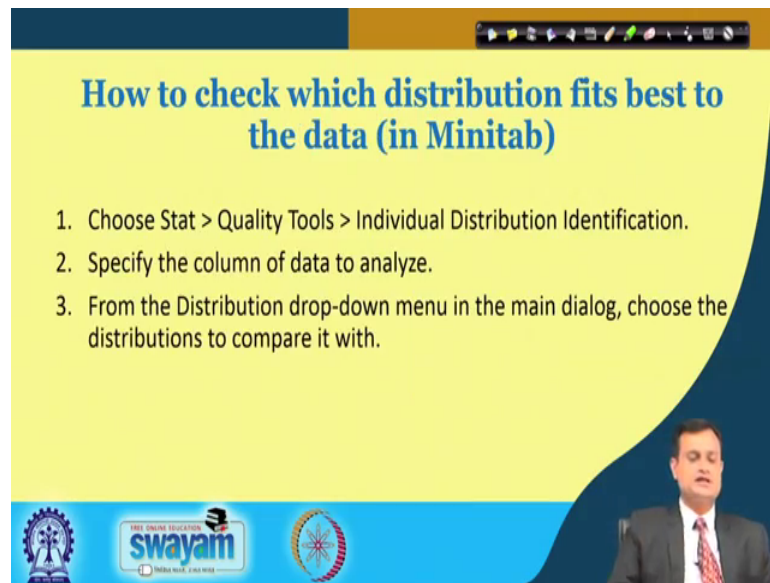
So, let us try to see that how to execute different options. So, option 1 find the best fit distribution and perform the process capability analysis. So, I am interested in finding the best fit distribution; it may be exponential, Weibull, triangle whatever I want to see that my data fits best fit to which particular distribution.

(Refer Slide Time: 09:24)



So, if data is non-normal then how to find the distribution.

(Refer Slide Time: 09:30)



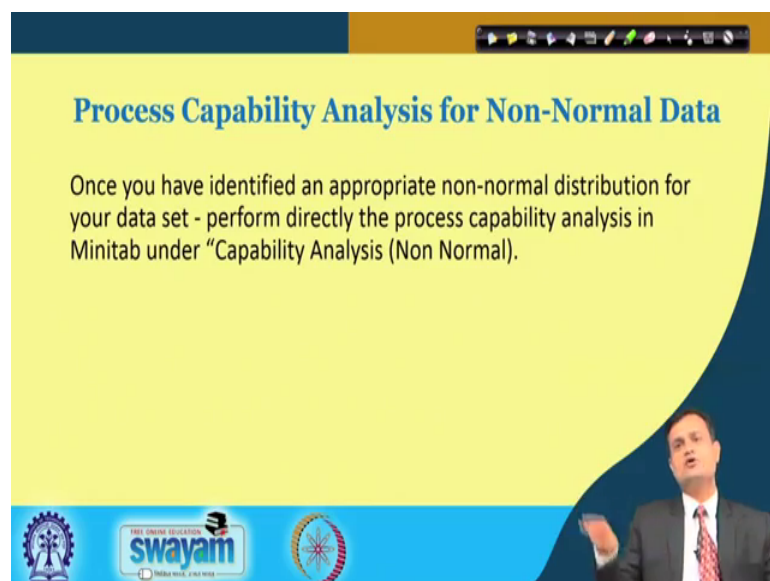
How to check which distribution fits best to the data (in Minitab)

1. Choose Stat > Quality Tools > Individual Distribution Identification.
2. Specify the column of data to analyze.
3. From the Distribution drop-down menu in the main dialog, choose the distributions to compare it with.

The slide features a yellow background with a blue header and footer. The footer contains logos for 'swayam' and 'INDIA RISE, EDUCATION RISE' along with a circular emblem. A presenter in a suit is visible in the bottom right corner.

So, you can use mini tab very easily to do this exercise and choose the stat quality tools, individual distribution identification, specify the column of data to analyze and from the distribution drop down menu in the main dialog you can choose the various distributions to compare with. And, then your mini tab will give you the best fit for various distribution comparing and you can find out that which distribution your data basically is resembling.

(Refer Slide Time: 10:06)



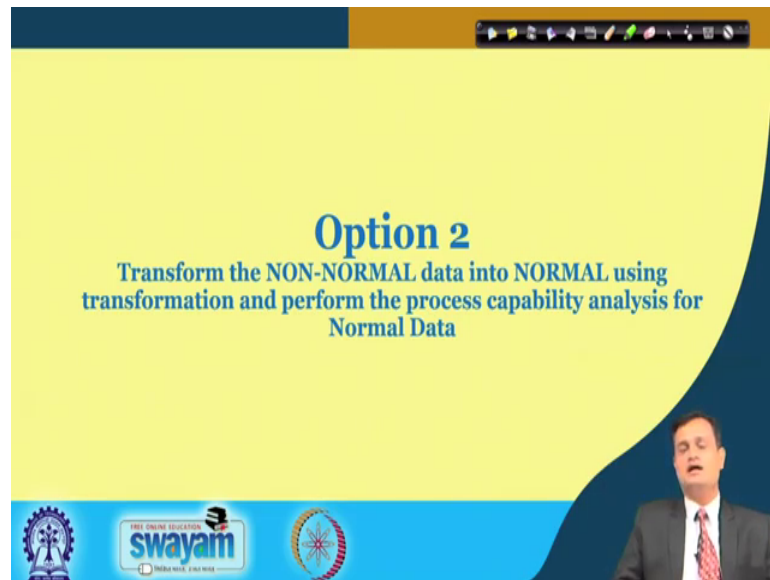
Process Capability Analysis for Non-Normal Data

Once you have identified an appropriate non-normal distribution for your data set - perform directly the process capability analysis in Minitab under "Capability Analysis (Non Normal)".

The slide features a yellow background with a blue header and footer. The footer contains logos for 'swayam' and 'INDIA RISE, EDUCATION RISE' along with a circular emblem. A presenter in a suit is visible in the bottom right corner.

So, process capability for non-normal data once you have identified figured out your non-normal distribution, best fit distribution then directly go to the option capability analysis non-normal perform it and you will get the results.

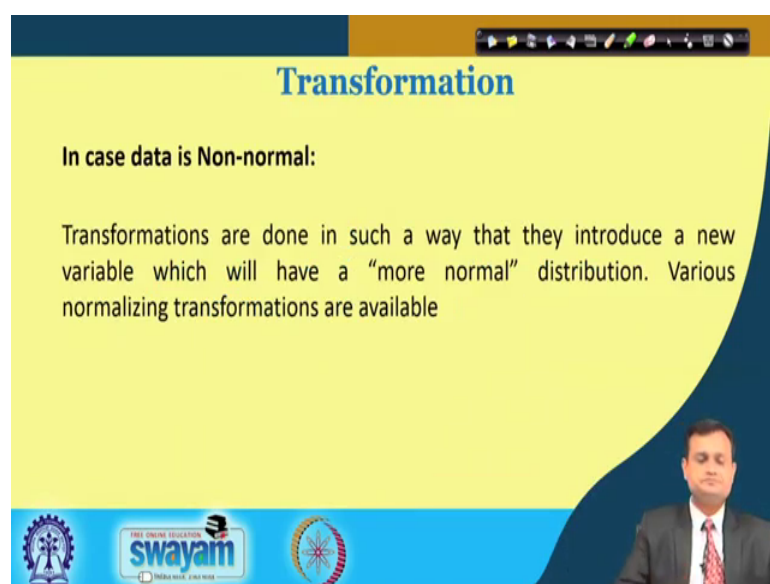
(Refer Slide Time: 10:26)



The slide has a yellow background with a dark blue curved border on the right. At the top, there is a presentation toolbar. The main text is centered and reads: **Option 2**
Transform the NON-NORMAL data into NORMAL using transformation and perform the process capability analysis for Normal Data. At the bottom left, there are logos for 'swayam' and other educational institutions. A small video inset of a man in a suit is in the bottom right corner.

So, option 2 is basically to transform. Now, when I cannot do the regular capability analysis which is made for normal data and if my data is non-normal I have to transform the data this is option 2.

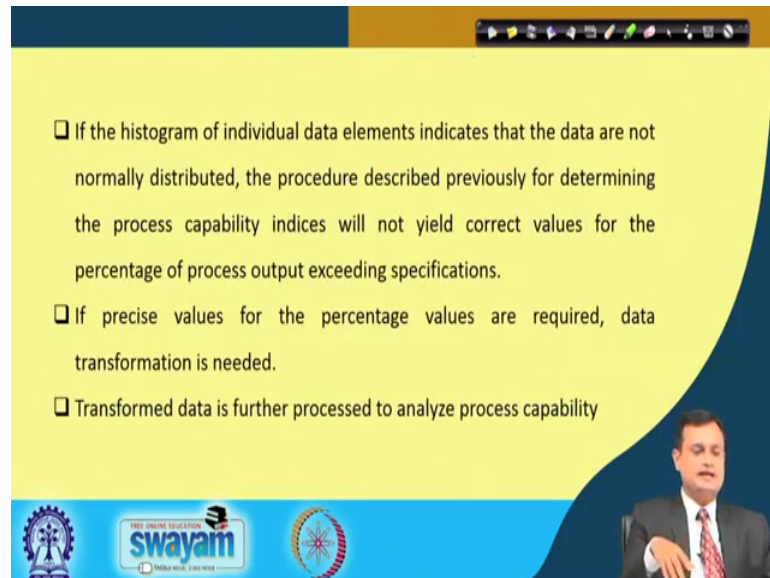
(Refer Slide Time: 10:45)



The slide has a yellow background with a dark blue curved border on the right. At the top, there is a presentation toolbar. The main text is centered and reads: **Transformation**
In case data is Non-normal:
Transformations are done in such a way that they introduce a new variable which will have a "more normal" distribution. Various normalizing transformations are available. At the bottom left, there are logos for 'swayam' and other educational institutions. A small video inset of a man in a suit is in the bottom right corner.

So, you have the non-normal data and try to make it more normal by transforming it through some mathematical function.

(Refer Slide Time: 10:56)



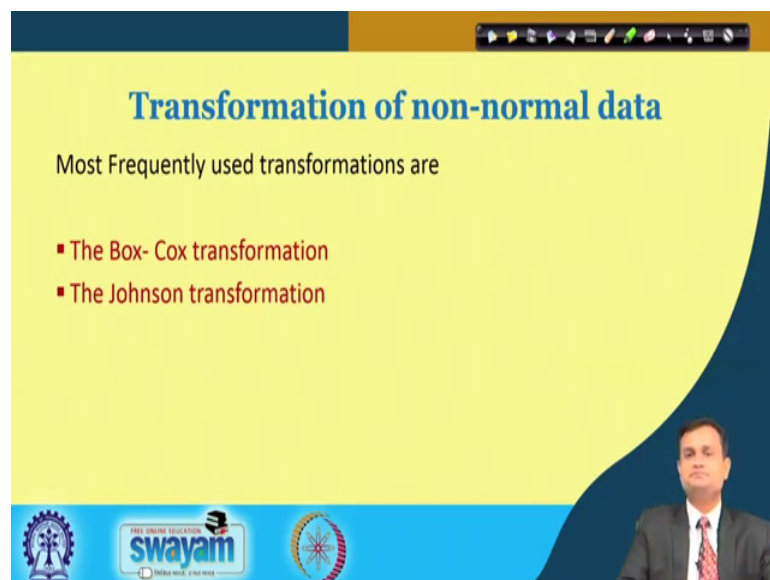
This slide contains three bullet points explaining when data transformation is necessary for process capability analysis. The background is yellow with a blue wavy border on the right. A presenter is visible in the bottom right corner.

- ❑ If the histogram of individual data elements indicates that the data are not normally distributed, the procedure described previously for determining the process capability indices will not yield correct values for the percentage of process output exceeding specifications.
- ❑ If precise values for the percentage values are required, data transformation is needed.
- ❑ Transformed data is further processed to analyze process capability

Logos for IIT Bombay and SWAYAM are at the bottom.

So, you can first check through histogram or your normal probability plot that whether your data is normal or not, and if you feel confident that yes my histogram as well as normal probability plot indicates that my data is not normal, then you transform the data.

(Refer Slide Time: 11:18)



This slide lists the most frequently used transformations for non-normal data. The title 'Transformation of non-normal data' is in blue. The background is yellow with a blue wavy border on the right. A presenter is visible in the bottom right corner.

Transformation of non-normal data

Most Frequently used transformations are

- The Box- Cox transformation
- The Johnson transformation

Logos for IIT Bombay and SWAYAM are at the bottom.

So, you have two options to transform the data two transformations rather one is Box Cox transformation other is the Johnson's transformation, it is extremely difficult to do it

manually I will give you the mathematical expression for your understanding. But, we take the help of software which has inbuilt function mathematical function and that helps us to transform the data very easily.

(Refer Slide Time: 11:48)

Box-Cox Power Transformation

The original form of the Box-Cox transformation, as appeared in their 1964 paper, takes the following form:

$$y_i^\lambda = \frac{y_i^\lambda - 1}{\lambda} \quad \text{for } \lambda \neq 0,$$

and $y_i^\lambda = \ln(y_i)$ for $\lambda = 0$

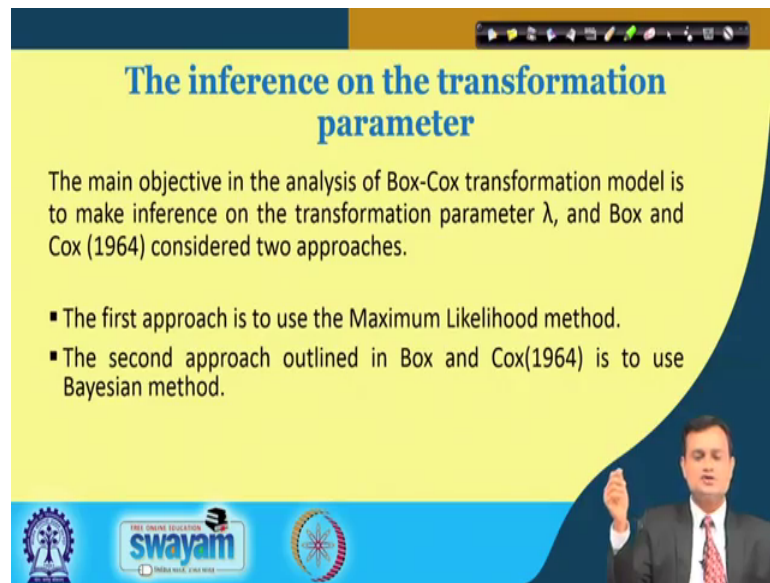
Here y is the observed data point and λ is a random variable. Software packages are used calculate λ .

The slide features a yellow background with a blue header and footer. The footer includes logos for 'swayam' and 'INDIAN INSTITUTE OF TECHNOLOGY KANPUR'. A small video feed of a presenter is visible in the bottom right corner.

So, Box Cox transformation, the critical element here is the lambda which raises the power of the data to that exchange where your data becomes more normal. So, just see the expression that y_i rise to lambda is equal to y_i rise to lambda minus 1 divided by lambda for lambda not equal to 0.

And if you have lambda equal to 0 then y_i rise to lambda is equal to $\log y_i$. So, lambda typically is a random variable and software packages are typically used to calculate the lambda and y is basically your absorbed data point, typically your data which is non normal and this lambda will help you to transform your non-normal y_i data into more normal.

(Refer Slide Time: 12:41)



The inference on the transformation parameter

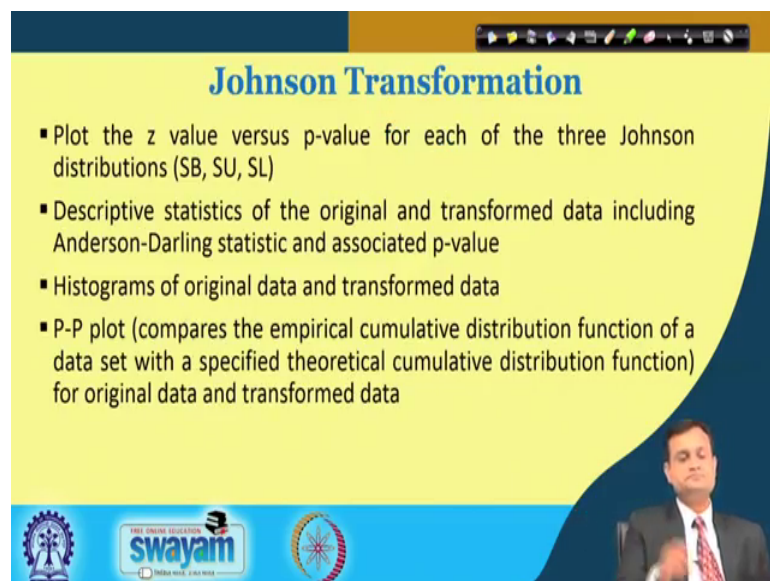
The main objective in the analysis of Box-Cox transformation model is to make inference on the transformation parameter λ , and Box and Cox (1964) considered two approaches.

- The first approach is to use the Maximum Likelihood method.
- The second approach outlined in Box and Cox(1964) is to use Bayesian method.

The slide features a yellow background with a blue header and footer. The footer contains logos for 'swayam' and other educational institutions. A presenter in a suit is visible in the bottom right corner.

So, you have certain influence on the transformers and parameter. So, the main object you have Box Cox transformation model is to make inference on the transformation parameter the critical parameter lambda, and Box Cox in 1964 this particular method was proposed he advocated two approaches. Number 1 is to use the maximum likelihood method you may have to read a little bit on this. And, the second approach is to use the Bayesian approach and considering the previous data posterior distribution and you can little bit read on these to just get familiarized with this terminology.

(Refer Slide Time: 13:27)



Johnson Transformation

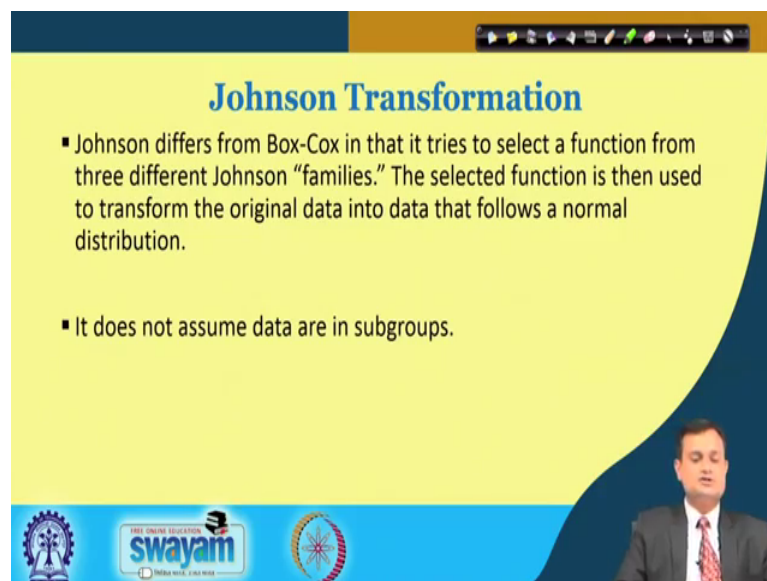
- Plot the z value versus p-value for each of the three Johnson distributions (SB, SU, SL)
- Descriptive statistics of the original and transformed data including Anderson-Darling statistic and associated p-value
- Histograms of original data and transformed data
- P-P plot (compares the empirical cumulative distribution function of a data set with a specified theoretical cumulative distribution function) for original data and transformed data

The slide features a yellow background with a blue header and footer. The footer contains logos for 'swayam' and other educational institutions. A presenter in a suit is visible in the bottom right corner.

So, you have another option that is Johnson's transformation and the Johnson's transformation typically plot the z value to your normal critical value versus p value that is the probability value for each of the three Johnson distributions. So, he has given a typical family or three distributions SB SU and SL. So, descriptive statistics of original and transform data including Anderson darling statistics which basically checks the normality of the data and the associated p value.




So, histogram of the original data and the transformed data can also be compared for visual purposes and you can have the PP plot which typically say compares your empirical distribution function and theoretical cumulative distribution function. So, empirical distribution is based on the sample, how your events are occurring and the theoretical distribution is based on the large data set available. So, we try to compare these two and see that to what extent my transform his son is acceptable.


(Refer Slide Time: 14:47)



Johnson Transformation

- Johnson differs from Box-Cox in that it tries to select a function from three different Johnson “families.” The selected function is then used to transform the original data into data that follows a normal distribution.
- It does not assume data are in subgroups.



So, you have Johnson's transformation typically differs from Box Cox and the difference is that it tries to select the function from three different Johnson families. SB SU SL and the selected function is then used to transform the original data in case of Box Cox that transformation component was lambda, here you have three distribution families and typically it does not assume the data in subgroups.

(Refer Slide Time: 15:23)

Johnson Transformation

Determining the optimal transformation

Johnson Transformation uses the following algorithm to determine an optimal transformation from the three families of distribution: SB, SL, and SU, where B, L, and U refer to the variable being bounded, lognormal, and unbounded

Bounded - something is of finite size, and that this is the case if it is smaller than some other object that has a finite size. (Otherwise it is **unbounded**.)

The slide features a yellow background with a blue wavy border on the right. At the bottom, there are logos for 'swayam' and other educational institutions, along with a small video inset of a man in a suit.

So, you have typically this Johnson transformers families, which is SB, B stands for bounded L stands for log normal and U stands for unbounded. So, bounded is something of finite size and this is the case if it is smaller than other object, that has a finite size otherwise it is unbounded.

(Refer Slide Time: 15:49)

Johnson Transformation functions

Johnson family	Transformation function
SB	$\gamma + \eta \ln[(x - \xi) / (\lambda + \xi - x)]$
SL	$\gamma + \eta \ln(x - \xi)$
SU	$\gamma + \eta \sinh^{-1}[x - \xi / \lambda]$ Where, $\sinh^{-1} x = \ln[x + \sqrt{1 + x^2}]$

The slide features a yellow background with a blue wavy border on the right. At the bottom, there are logos for 'swayam' and other educational institutions, along with a small video inset of a man in a suit.

So, you can see the expressive mathematical expression and as I mentioned that it is difficult to do it manually, it is always better to use the software. But obviously, we should appreciate the logic behind the software has basically created a platform and

using it. So, you have SB, which is typically express in terms of a transformers and function $\lambda + \eta \log x - \eta x$ and same way your SL and SU. So, I am not going into the details of these mathematical expressions, but I would like to make a mention that depending upon your nature of data it will select one of these mathematical function and transform the data to the more normal data.

(Refer Slide Time: 16:45)



The slide features a yellow background with a dark blue header and footer. The title "Non-normal Process Capability in Minitab" is centered in blue. Below it, the word "Sources:" is followed by two bulleted links. The footer contains three logos: a circular emblem on the left, the "swayam" logo in the center, and a circular emblem on the right. A small inset image of a man in a suit is visible in the bottom right corner.

Non-normal Process Capability in Minitab

Sources:

- a) <https://blog.minitab.com/blog/statistics-in-the-field/5-simple-steps-to-conduct-capability-analysis-with-non-normal-data>
- b) <https://sixsigmadsi.com/non-normal-capability-analysis/>

swayam

So, let us try to see how we can conduct this analysis that is non normal process capability in mini tab, I have highlighted here two sources from which you can learn this application into greater detail I also use them for explaining this application to you; so, illustrative example 1.

(Refer Slide Time: 17:09)

Steps involved in Minitab for Non-normal data

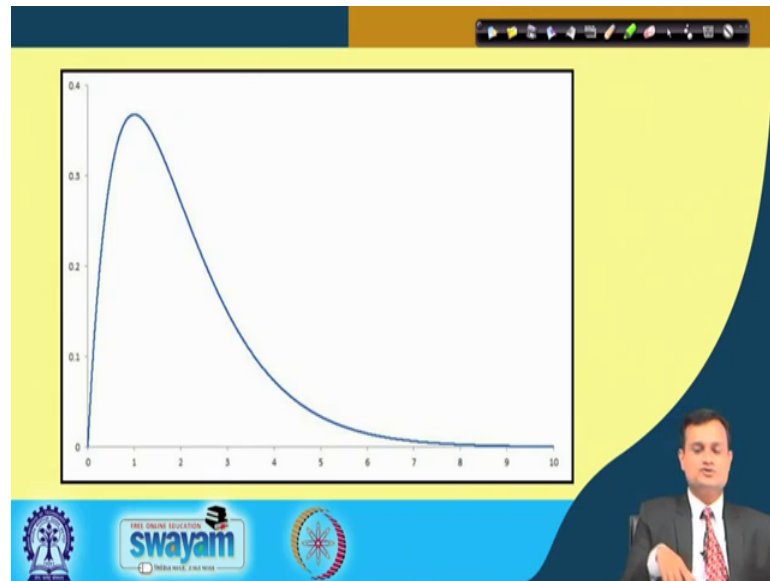
- ❑ Consider a Lean Six Sigma project to reduce the lead time required to install an information technology solution at a customer site. It should take no more than 30 days—working 10 hours per day Monday–Friday—to complete, test and certify the installation. Following the standard process, the target lead time should be around 24 days.
- ❑ Twenty-four days may be the target, but we know customer satisfaction increases as we complete the installation faster. We need to understand our baseline capability to meet that demand, so we can perform a capability analysis.
- ❑ We know our data should fit a non-normal (positively skewed) distribution.

The slide features a yellow background with a blue header and footer. The footer includes logos for 'swayam' and 'INDIA RISE, EDUCATION RISE' along with a small video inset of a man in a suit.

Let us say I have the minute app software and typically I need to follow the steps like this. Consider typically a lean six sigma project to reduce the lead time required to install typically an information technology solution at a customer site and it should take no more than 30 days. So, working is for 10 hours per day Monday Friday to complete Monday to Friday and test and certify the installation.

So, following the standard process the target lead times should be 24 days and I need to conduct the non normal process capability to see that to what extent I am capable enough to meet this target. So, 24 days maybe the target, but we know that customer satisfaction increases, if we complete the installation faster and we basically need to understand or figure out that our baseline capability to meet that demand what is it. So, we can perform a capability analysis.

(Refer Slide Time: 18:22)



Now, just see this and you would find that your distribution is positively skewed and it does not out the normal distribution pattern of the data and I have to conduct the non-normal process capability analysis.

(Refer Slide Time: 18:41)

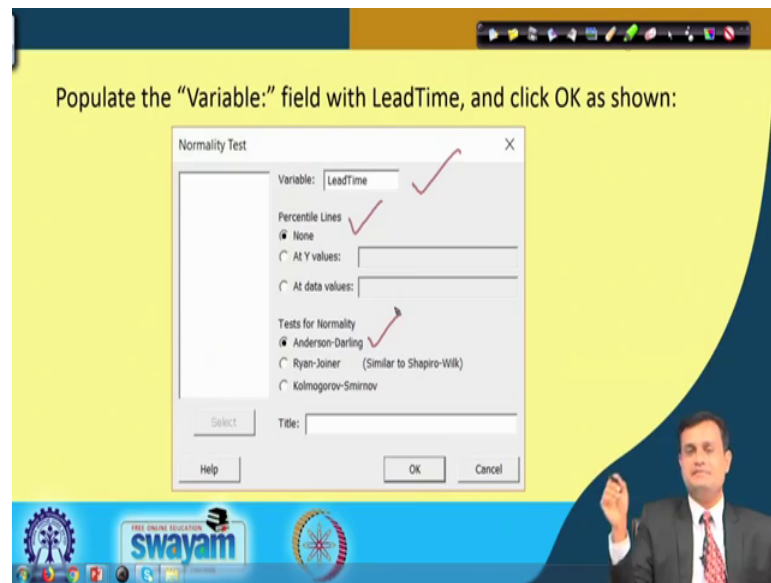
Steps involved...

- **Step 1: Collect data:** First we must gather data from the process. In this scenario, we are collecting sample data. We pull 100 samples that cover the full range of variation that occurs in the process. In this case the full range of variation comes from three installation teams. We will take at least 30 data points from each team.
- **Step 2: Identify the Shape of the Distribution:** We know that the data should fit a non-normal distribution. As Lean Six Sigma practitioners, we must prove our assumption with data. In this case, we can conduct a normality test to prove non-normality. We are using Minitab as the statistical analysis tool, and our data are available in this worksheet.
- From the menu, select "Normality Test" found under "Stat -> Basic Statistics -> Normality Test ..."

The slide is titled "Steps involved..." and contains three bullet points. The first bullet point describes Step 1: Collect data. The second bullet point describes Step 2: Identify the Shape of the Distribution. The third bullet point describes the menu path to select the Normality Test. The slide is displayed on a yellow background with a blue border. At the bottom, there are logos for 'swayam' and 'INDIA WIDE, TIME WIDE'.

So, steps are very simple to follow, you collect the data identify the shape of the distribution. So, for that step 2 you go to stat basic statistics normality test.

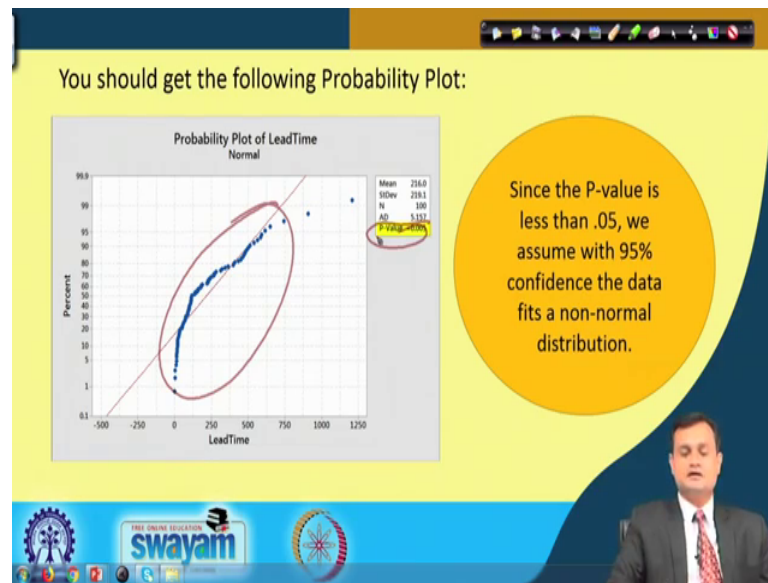
(Refer Slide Time: 18:54)



So, I am showing you the window also so, that you get clear on idea about executing the non normal process capability in mini tab. So, you have this particular window where you have percentile lines. So, I am saying none you are the variable of interest, I am putting lead time test for normality I am selecting Anderson Darling. So, I would like to just mention here when we would be discussing the various test in the subsequent lectures statistics is the tool by which inferential statistics by which I can check my hypothesis, I can check my claim at certain level of significance and say that my assumption or my claim is true or it is false.

So, when I want to do this, with certain kind of assumption then there are various trades available for this purpose I would use the Anderson darling test which is widely used to check the normality. So, in lecture one I clearly mentioned that to learn this course you need to have as a prerequisite some basic knowledge of the statistics anyhow we are discussing the important concepts, but it is always desirable that you have some knowledge about the basics of the statistics including distribution and probability theory and other. So, that part you should keep in mind.

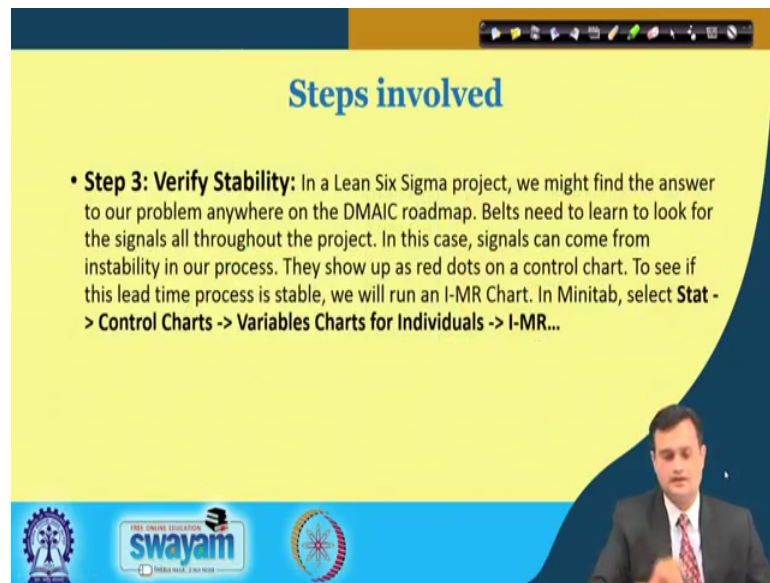
(Refer Slide Time: 20:45)



Now, let us see that what is the next step in my analysis using mini tab. So, I can see here very clearly that my normal probability plot shows that my data is not fitting the straight line, then my assumption about normality of the data is not valid, it means my data is non normal. You can also check the P value we have not discussed the concept of P value so, far, but you can just see that your P value is point 0.05, which is quite less than the level of significance you might have assumed 0.05.

And, hence it is in the rejection region we will see this in the subsequent lecture how to test the hypothesis and again this inferential measure indicates that my assumption about normality is not true and my data is non normal.

(Refer Slide Time: 21:56)



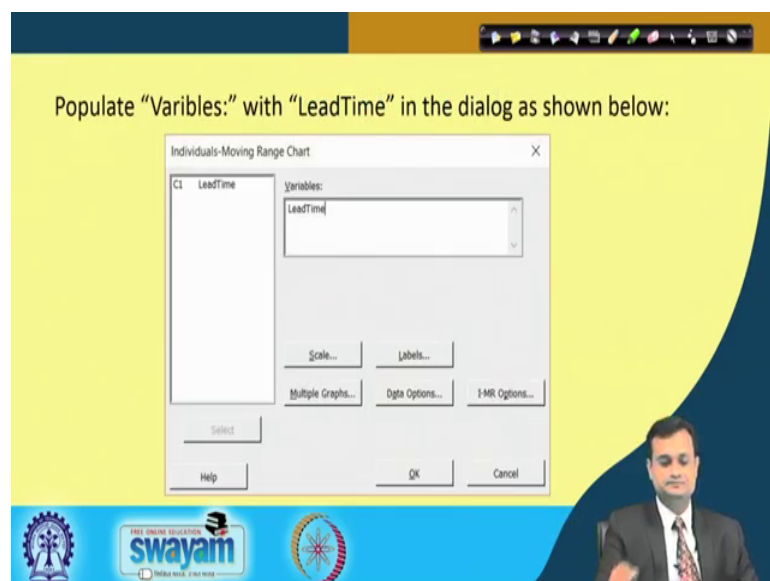
Steps involved

- **Step 3: Verify Stability:** In a Lean Six Sigma project, we might find the answer to our problem anywhere on the DMAIC roadmap. Belts need to learn to look for the signals all throughout the project. In this case, signals can come from instability in our process. They show up as red dots on a control chart. To see if this lead time process is stable, we will run an I-MR Chart. In Minitab, select **Stat -> Control Charts -> Variables Charts for Individuals -> I-MR...**

The slide features a yellow background with a blue header and footer. The footer includes logos for Swamyam and other educational institutions. A small video inset of a man in a suit is visible in the bottom right corner.

Now, step 3 verify the stability. So, typically in lean six sigma project which we are trying to study for installation of the software, I have a data set, but I would like to check the stability through individual moving average chart and this will tell me that over a period of time my processes in the statistical control or not. If not then there is no point in conducting the process capability analysis whether normal or not normal for the collected data.

(Refer Slide Time: 22:32)



Populate "Variables:" with "LeadTime" in the dialog as shown below:

The slide displays a screenshot of the Minitab 'Individuals-Moving Range Chart' dialog box. The 'Variables:' field is populated with 'LeadTime'. The dialog box includes buttons for 'Scale...', 'Labels...', 'Multiple Graphs...', 'Data Options...', 'I-MR Options...', 'Select', 'Help', 'OK', and 'Cancel'. The slide has a yellow background with a blue header and footer. The footer includes logos for Swamyam and other educational institutions. A small video inset of a man in a suit is visible in the bottom right corner.

So, I can do this very easily in the mini tab.

(Refer Slide Time: 22:37)



I am putting the variable each time, I can generate this chart already I was shown in the previous slide and I can see that yes there are some of the points which are falling outside and that is a stability issue. So, either I can discard this point and do it again or I can take certain corrective actions to see that my process becomes statistically controlled and stable so, that whatever data I collect that are adequate and good enough to continue conduct the process capability analysis.

(Refer Slide Time: 23:18)

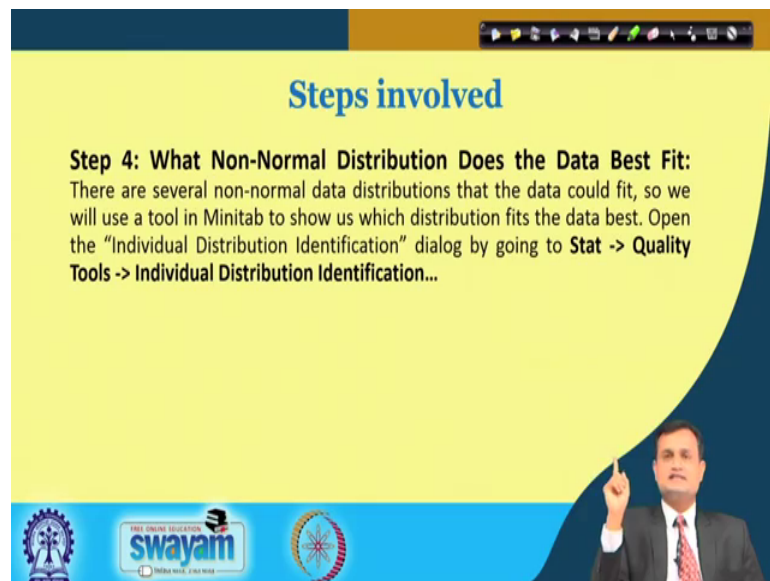
Observation from I-MR Chart

- These data points indicate abnormal variation, and their cause should be investigated. These signals could offer great insight into the problem you are trying to solve.
- Once identified and resolved the causes of these points, you can take additional data or remove the points from the data set.
- **In this scenario, we will leave the two points in the data set.**

swayam

Now I have the observation that these data points which are indicated outside the upper control limit, they indicate the abnormal variation and once identified resolve the cause of this points as I said take the corrective action. And, you can take the additional data or remove the data points from the data set and then you can conduct the process capability analysis for the remaining data.

(Refer Slide Time: 23:47)



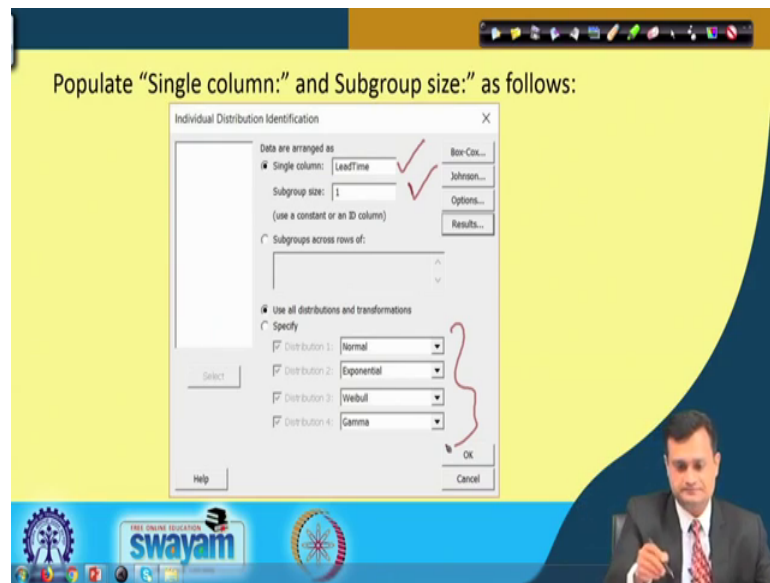
Steps involved

Step 4: What Non-Normal Distribution Does the Data Best Fit:
There are several non-normal data distributions that the data could fit, so we will use a tool in Minitab to show us which distribution fits the data best. Open the "Individual Distribution Identification" dialog by going to **Stat -> Quality Tools -> Individual Distribution Identification...**

The slide features a yellow background with a blue header and footer. The footer contains logos for 'swayam' and 'INDIAN INSTITUTE OF TECHNOLOGY KANPUR'. A small inset image of a man in a suit pointing upwards is located in the bottom right corner of the slide.

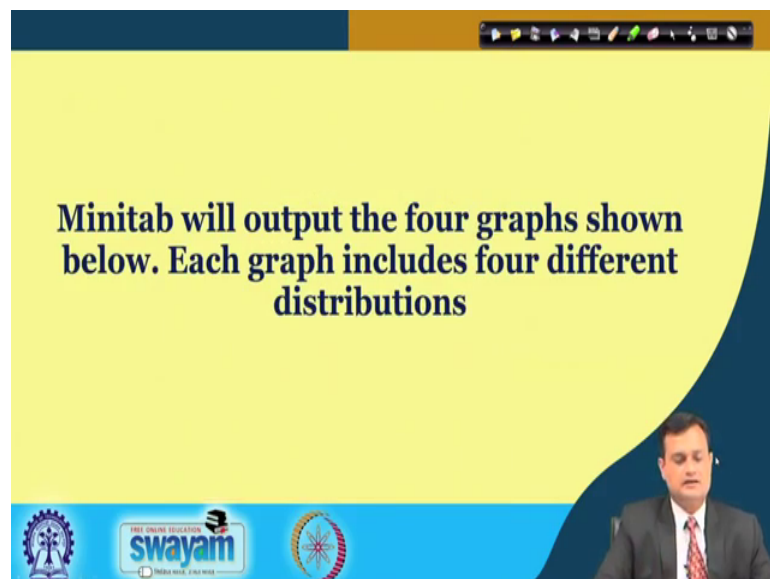
Step 4: what non-normal distribution does the data based fit? So, as I mentioned you have two options available option number 1 you can find out what is the best fit distribution and directly use the option for non-normal process capability in mini tab and conduct the analysis or else you can transform the data to normal and then you use the normal say process capability analysis in mini tab. So, both in a way it is ok.

(Refer Slide Time: 24:19)



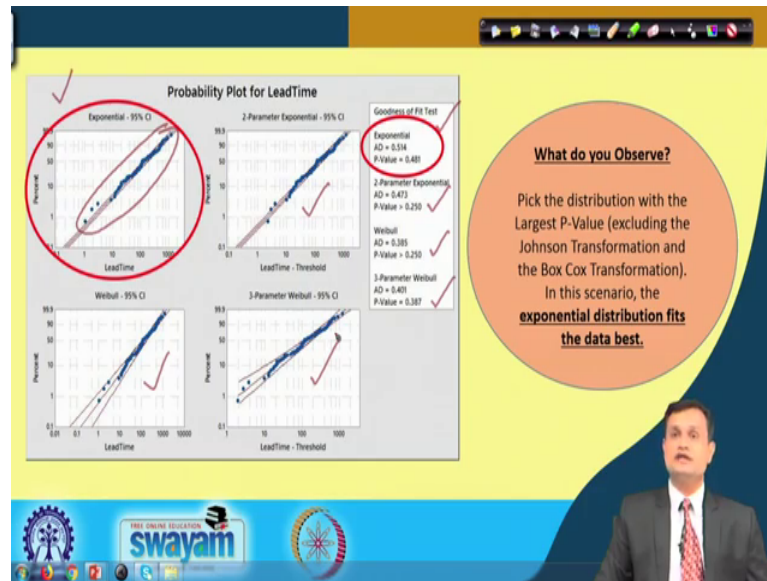
So, here I am just trying to see that what could be the best fit distribution and I am trying to do it with the option 1. So, just see this that my variable of interest is lead time subgroup size I have considered one and I am trying to select the various distributions normal exponential Weibull gamma and my say output will help me to understand what is the best fit distribution.

(Refer Slide Time: 24:57)



. So, here mini tab will give the output for the four graphs typically shown in this slide.

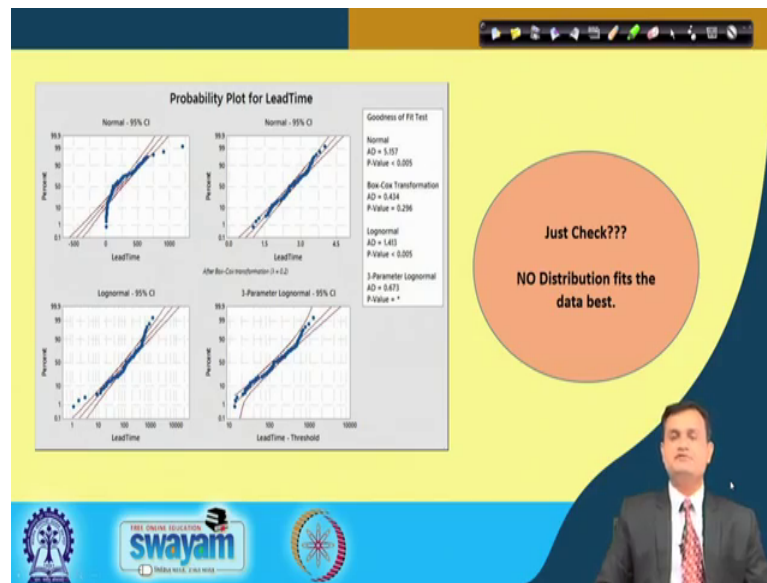
(Refer Slide Time: 25:03)



And what you can see here that, exponential non my probability plot exponential the data best fits. Also you can verify this here that Anderson Darling value is 0. 5 and 4 which is quite acceptable we have not discussed this test in detail. So, leave it, but you can at least see for the P value 0. 481. So, this is greater than the alpha which is 0. 05 in this case and it is acceptable.

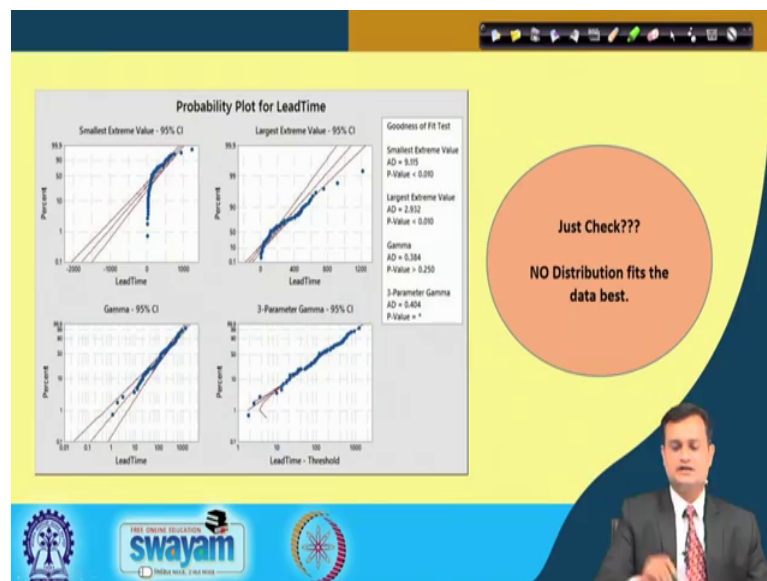
So, visually also it is acceptable as well as statistically also it is acceptable and in all other cases if you see then P value as well as Anderson-Darling value they do not conform, and if I look at the nor probability plot again it is not conforming that my data is fitting to that particular distribution. So, here are my data best fits to the exponential distribution and fine I can conduct the non-normal process capability choosing the exponential distribution.

(Refer Slide Time: 26:23)



So, you can also compare it with the other distribution and just check it the best fit.

(Refer Slide Time: 26:29)

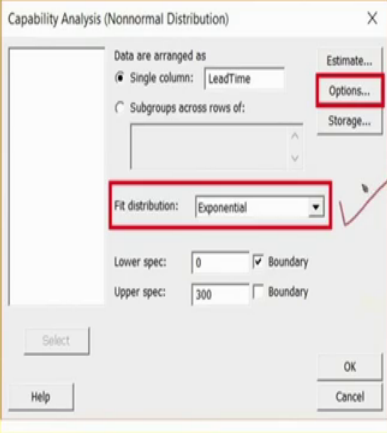


Just check for other distribution. So, no way it is meeting just visually you can see the probability plot and you can figure out that it is not fitting that particular distribution.

(Refer Slide Time: 26:40)

Steps involved

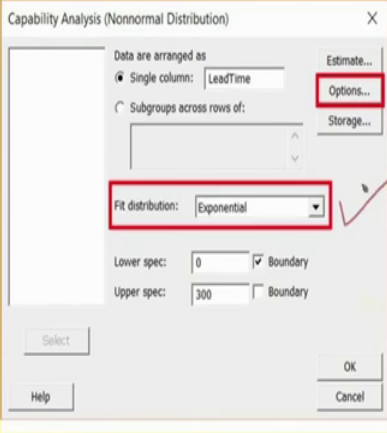
- **Step 5: What Is the Process Capability:** Now that we know the distribution that best fits these data, we can perform the non-normal capability analysis. In Minitab, select **Stat -> Quality Tools -> Capability Analysis -> Nonnormal...**
- Populate the "Capability Analysis (Nonnormal Distribution)" dialogue box as seen below. Make sure to select "Exponential" next to Fit distribution. Then, Click on "Options".



The slide shows a yellow background with a blue header and footer. The header contains the title "Steps involved". The footer contains logos for "swayam" and "INDIA WIDE, FREE WIDE". A small video inset in the bottom right corner shows a man in a suit and tie.

So, now step 5 what is the process capability? So, you use the options stat quality stat tools, capability analysis, non-normal I am exercising option one because it is my non-normal process capability, I have figured out what is the best fit distribution.

(Refer Slide Time: 27:00)

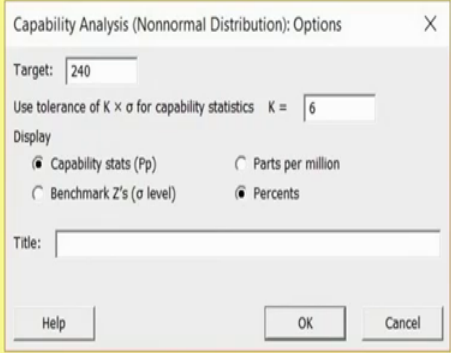


The slide shows a yellow background with a blue header and footer. The header contains the title "Steps involved". The footer contains logos for "swayam" and "INDIA WIDE, FREE WIDE". A small video inset in the bottom right corner shows a man in a suit and tie.

So, you have chosen here this exponential distribution which you found in the previous analysis.

(Refer Slide Time: 27:18)

Fill in the "Capability Analysis (Non Normal Distribution): Options" dialogue box with the following:

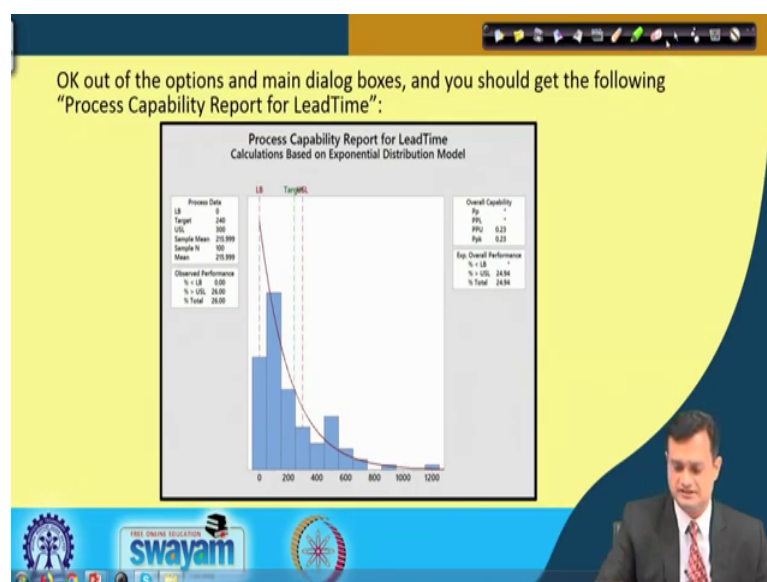


We chose "Percents" over "Parts Per Million" because in this scenario it would take years to produce over one million outputs (or data for each installation time).

swayam

And now with this you conduct basically the process capability analysis. So, you put the target value, you put the k value, that is your tolerance came to sigma. So, six sigma let us say capability stats Pp percent and you try to find out the various values.

(Refer Slide Time: 27:39)

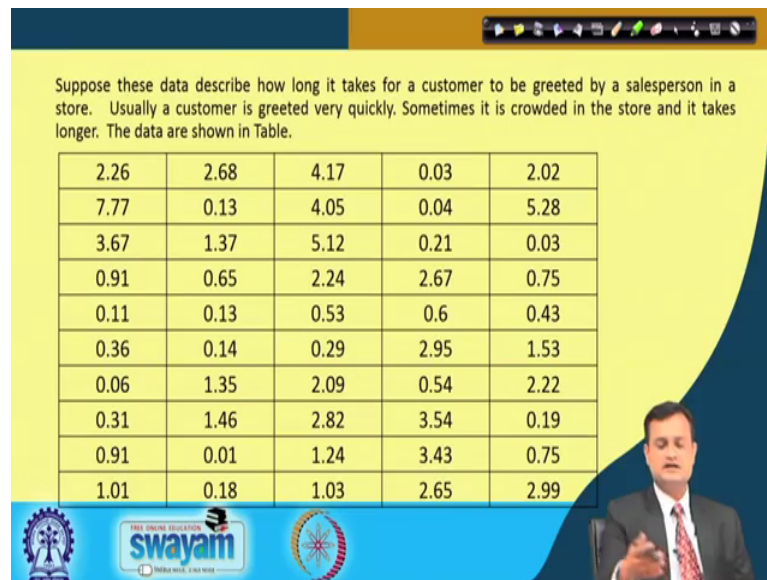


So, you will see that PPU is point 2 3 pp k is point put 23. So, my process is not capable to meet the required specifications or my target of completing a project and there is a need to improve upon my implementation plan I included another example you can quickly go through this, but the same procedure is applied.

(Refer Slide Time: 28:09)

Suppose these data describe how long it takes for a customer to be greeted by a salesperson in a store. Usually a customer is greeted very quickly. Sometimes it is crowded in the store and it takes longer. The data are shown in Table.

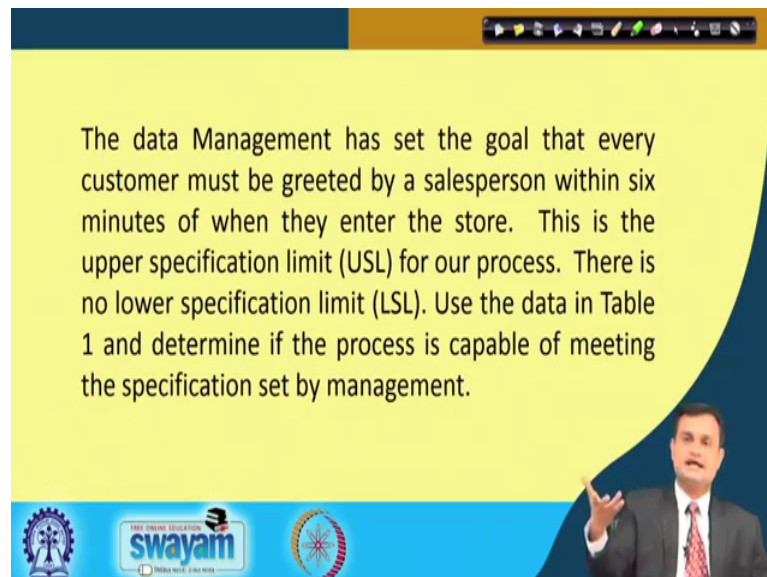
2.26	2.68	4.17	0.03	2.02
7.77	0.13	4.05	0.04	5.28
3.67	1.37	5.12	0.21	0.03
0.91	0.65	2.24	2.67	0.75
0.11	0.13	0.53	0.6	0.43
0.36	0.14	0.29	2.95	1.53
0.06	1.35	2.09	0.54	2.22
0.31	1.46	2.82	3.54	0.19
0.91	0.01	1.24	3.43	0.75
1.01	0.18	1.03	2.65	2.99



I have the data set describing the customer to be agree they are greeted by a sales person and how long this process takes. So, usually customer is greeted very quickly, and if you keep customer waiting and then you greet then it will lose the value of greeting and welcoming.

(Refer Slide Time: 28:33)

The data Management has set the goal that every customer must be greeted by a salesperson within six minutes of when they enter the store. This is the upper specification limit (USL) for our process. There is no lower specification limit (LSL). Use the data in Table 1 and determine if the process is capable of meeting the specification set by management.

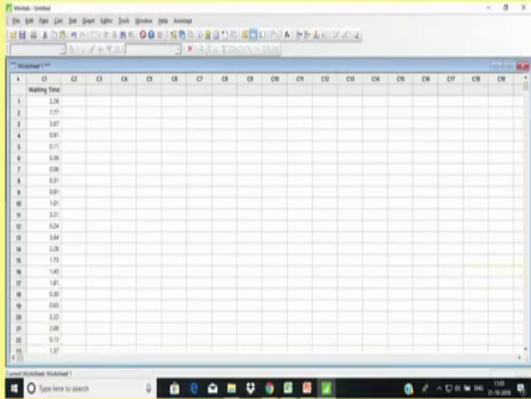


So, this is the data I have and this data said you can use in fact, to conduct the analysis as a hands on exercise I have given here the USL this is the upper specification limit for our process and within six minute it should be done and there is no lower specifications as

early as possible, the moment customer comes you greet welcome him and let us say you have the data given in your table one.

(Refer Slide Time: 29:03)

Enter Data in Minitab Work Sheet as shown



The screenshot shows a Minitab worksheet with a column named 'Waiting Time' containing 25 data points. The data values are: 2.26, 1.77, 1.67, 0.87, 0.77, 0.36, 0.86, 0.87, 0.87, 1.07, 3.27, 0.34, 0.64, 2.26, 1.76, 1.45, 1.87, 0.87, 0.87, 2.07, 2.08, 0.77, 0.77, 1.77.

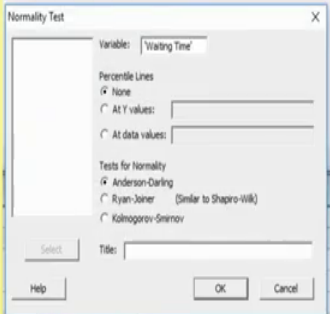
swayam

So, open these window enter the data.

(Refer Slide Time: 29:06)

Normality Test

- Stat->Basic Statistics->Normality test
- Populate the "Variable ." field with waiting time and click OK as shown.

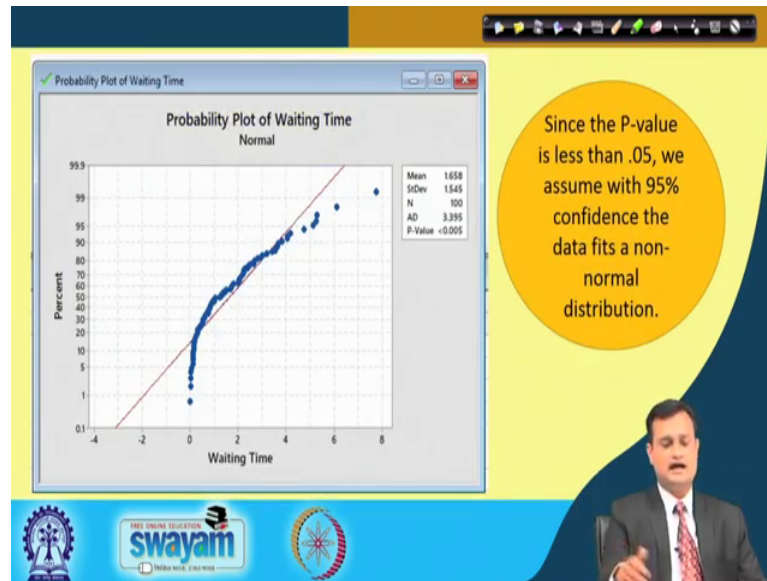


The screenshot shows the 'Normality Test' dialog box in Minitab. The 'Variable' field is set to 'Waiting Time'. Under 'Percentile Lines', 'None' is selected. Under 'Tests for Normality', 'Anderson-Darling' is selected. The 'OK' button is highlighted.

swayam

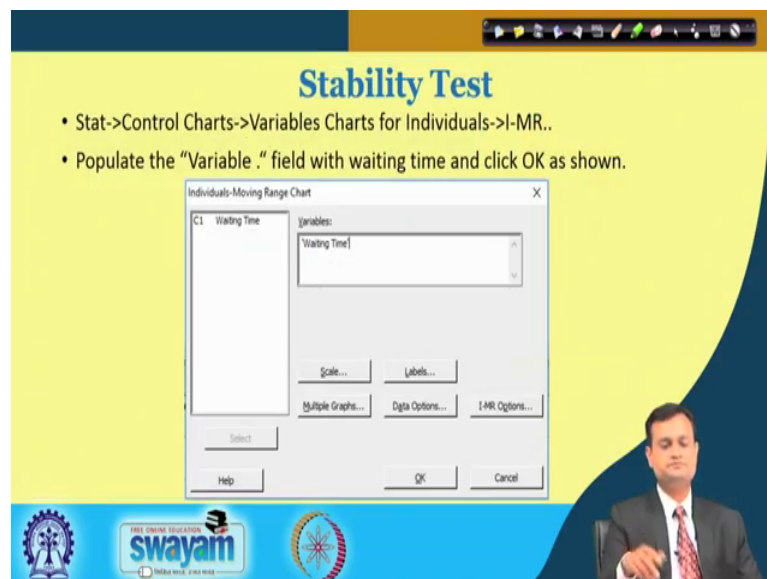
Choose this normality test first let me check whether my data is normal or not normal.

(Refer Slide Time: 29:14)



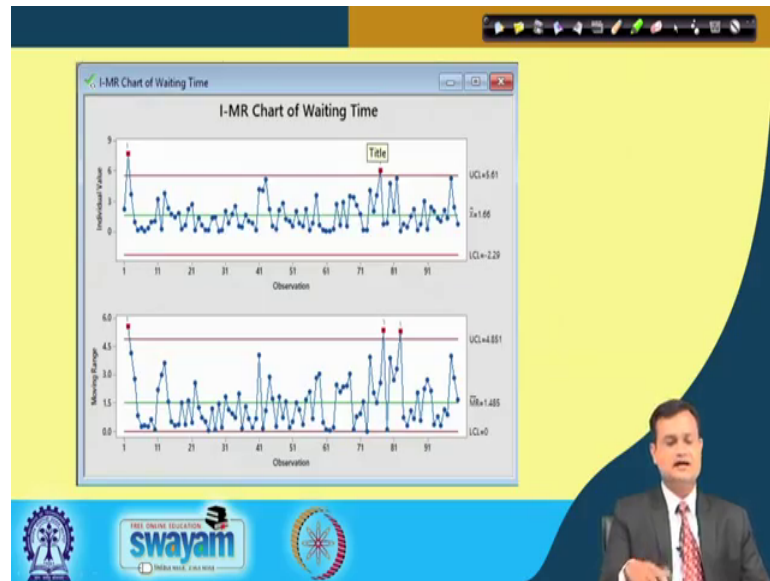
I will get this probability plot clearly I can see it is not normal.

(Refer Slide Time: 29:20)



Then I can check the stability.

(Refer Slide Time: 29:23)



I will see that couple of points are falling outside indicated in red. So, take the corrective action remove this point consider the other points.

(Refer Slide Time: 29:33)

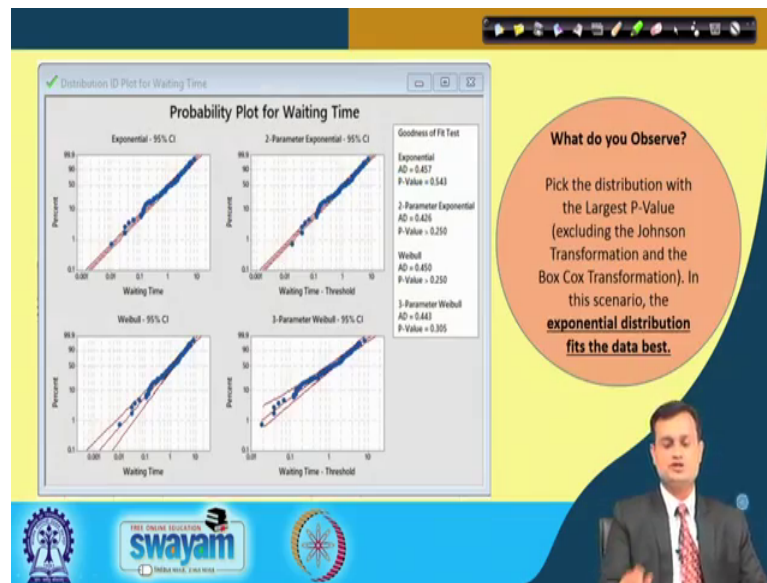
Individual Distribution Identification

Stat->Quality Tools->Individual Distribution Identification...

The slide has a yellow background with a blue curved border on the right. At the bottom, there are logos for 'swayam' and 'INDIA WISE, 2020 WISE'.

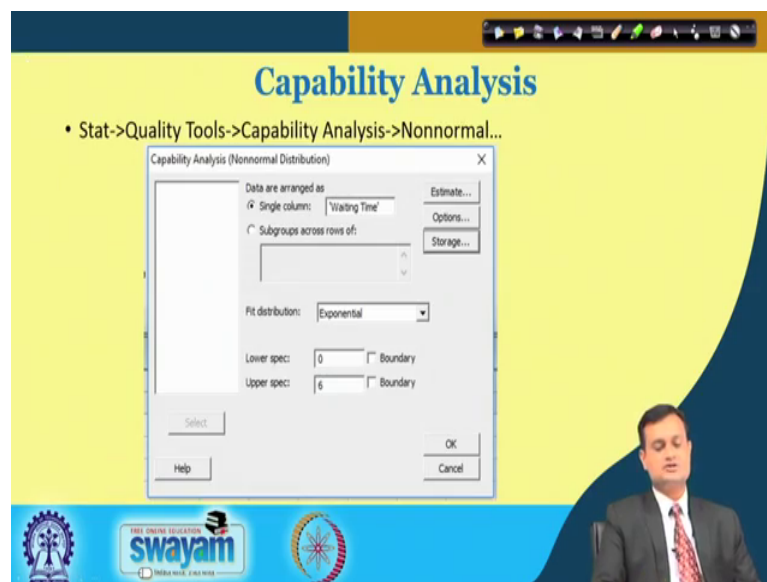
Go to stat quality tool individual distribution.

(Refer Slide Time: 29:37)



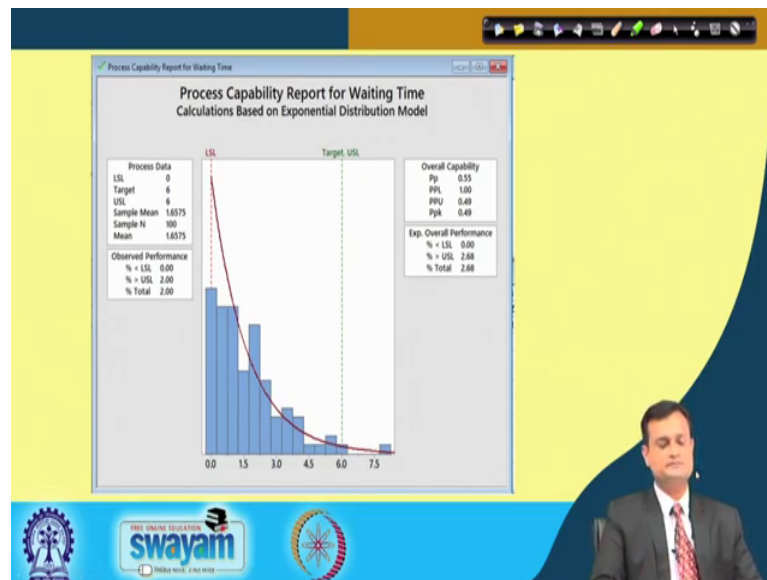
Conduct the analysis again I can see that my best fit distribution here is exponential. So, I would consider this as the best for distribution for conducting my process capability analysis.

(Refer Slide Time: 29:56)



So, then you put the consider the exponential define the variable in column upper specification, that is lower specifications zero immediately you can read welcome the customer.

(Refer Slide Time: 30:08)



You will get the analysis with says Pp is lower than 1 not acceptable PPL is just one both PPU when people are less than 1. So, the sales people or the people at the desk they are not very well trained or the rush of customer is very high I need to really correct my process of greeting welcoming the customer so, that, that can give customer a better satisfaction.

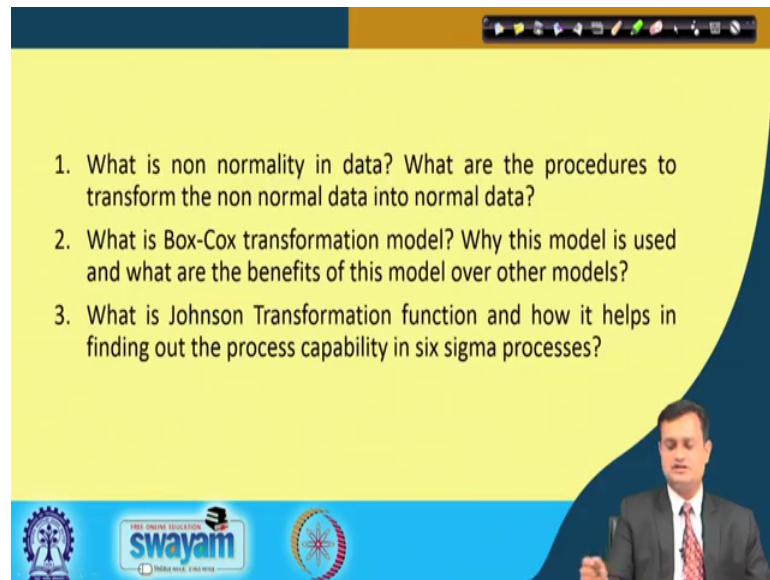
(Refer Slide Time: 30:51)



So, this is the capability analysis six pack that will also help you to understand the capability of your process, but how could we give use six different say display in one

window and r chart I chart stability of the process then observation plotted over a period of time there is there a trained exist and then normality through histogram as well as best fit distribution and then process capability indices.

(Refer Slide Time: 31:25)



1. What is non normality in data? What are the procedures to transform the non normal data into normal data?
2. What is Box-Cox transformation model? Why this model is used and what are the benefits of this model over other models?
3. What is Johnson Transformation function and how it helps in finding out the process capability in six sigma processes?

The slide is part of a presentation. At the bottom left, there are logos for IIT Bombay and the swayam platform. At the bottom right, there is a small video inset showing a man in a suit and tie, likely the presenter. The slide has a yellow background with a dark blue curved border on the right side.

Let me just end the session with a small think it. What is non-normal process capability and what is the non-normality in the data and what are the procedures you would like to follow to convert your normal data into the normal data. What is Box Cox transformation, why this model is used and what is Johnson transformation function, how it helps in finding out the process capability in six sigma processes.

(Refer Slide Time: 31:53)



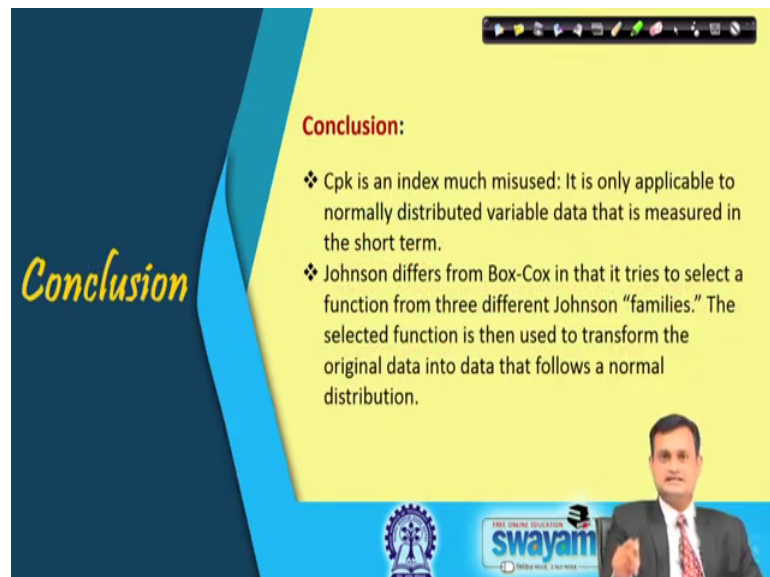
References:

- ❑ T. M. Kubiak, Donald W. Benbow, The Certified Six Sigma Black Belt Handbook, Pearson Publication.
- ❑ Forrest W. Breyfogle III, Implementing Six Sigma, John Wiley & Sons, INC.
- ❑ Mitra, Amitava. Fundamentals of Quality Control and Improvement, Wiley India Pvt Ltd.
- ❑ Montgomery, D.C. Statistical Quality Control: A modern introduction, Wiley.

The slide features a dark blue background on the left with the word 'References' in a yellow script font. The right side is yellow. A presenter in a suit is visible in the bottom right corner. The Swayam logo and 'FREE ONLINE EDUCATION' text are at the bottom.

You can use this references which has given the details of non-normal process capability, but I would insist that you use the same data or some other data and have hands on exercise in mini tab. So, that you feel confident in conducting as well as interpreting the results of non-normal process capability analysis.

(Refer Slide Time: 32:18)



Conclusion:

- ❖ Cpk is an index much misused: It is only applicable to normally distributed variable data that is measured in the short term.
- ❖ Johnson differs from Box-Cox in that it tries to select a function from three different Johnson "families." The selected function is then used to transform the original data into data that follows a normal distribution.

The slide features a dark blue background on the left with the word 'Conclusion' in a yellow script font. The right side is yellow. A presenter in a suit is visible in the bottom right corner. The Swayam logo and 'FREE ONLINE EDUCATION' text are at the bottom.

So, Cpk is an index, misused only use it when the data is normal otherwise convert your data non-normal into normal or choose the best fit distribution automatically your

software will take care. And, it will convert it into normal and give you the best result appropriate result of the process capability.

Thank you very much, keep revising be with me enjoy.