

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

Lecture - 17
Hypothesis Testing

Hello and welcome to session 17 of our course on Quality Control and Improvement with MINITAB. I am Professor Indrajit Mukherjee, Shailesh J. Mehta School of Management ok. So, last time we are discussing about confidence interval and that gives a lead to another important concept which is extensively used in quality and improvements control and improvements which is known as Hypothesis Testing.

So, I will give a brief idea of hypothesis testing over here. And then we will see what type of test is generally conducted which will help us in assessing the qualities like in many situations that we encounter in real-life problems like that ok. So let me start with hypothesis testing speaking on hypothesis testing.

(Refer Slide Time: 00:59)

Quality Control and Improvement using MINITAB

Hypothesis Testing

- Procedure of **making inference about population**.
- Allows us to determine whether enough statistical evidence exists to **conclude a belief** (i.e. hypothesis) **about a parameter** is supported by the data.

```
graph LR; Sample[Sample] -- Inference --> Population[Population];
```



Prof. Indrajit Mukherjee, SJMSOM, IIT Bombay



So, and we have discussed about population and samples. Population means if it is in manufacturing I told it is infinite population and from there whatever is the CTQ interest and what we do is that we take samples and measure the CTQs and try to estimate the parameters of the population. Let us say mean I am interested in because location, accuracy and precision both are important whether it is mean or variance like that.

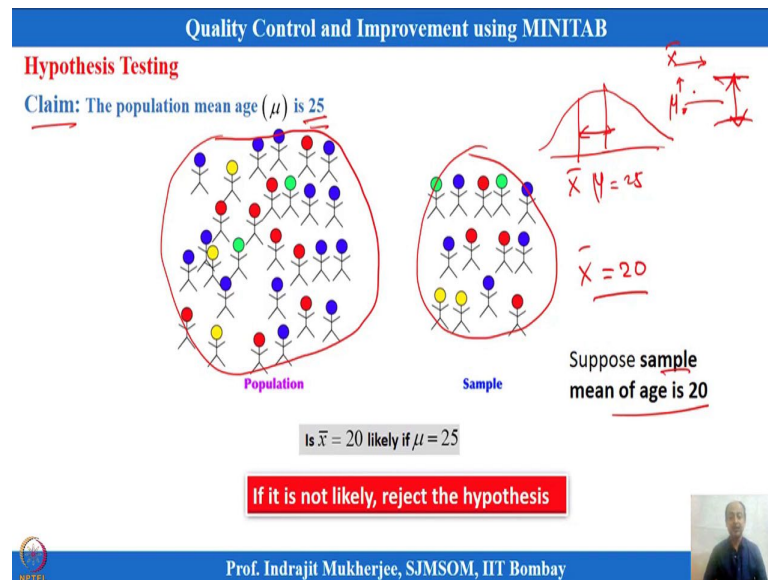
So, data has to be collected from the population. So, what we do is that? We take certain amount of samples which is adequate like that and then what we do is that we try to assess based on the sample observations. So, we calculate the mean and we try to say that this is the confidence interval of the population parameter or population mean like that ok or variance of the other.

So, there will be bound; there will be a boundary that we can determine like upper bound and lower bound like that ok. For mean and also for variance like that we have seen in confidence interval. And in confidence interval what is important to observe is that the statistician has provided a measure that we can be wrong, our confidence interval can be wrong and that is the level of significance that is the level of significance.

So, which is taken generally as 5 percent α level of significance that we have seen last time. So, with the given level of α we can determine the confidence intervals. So, if you want to minimize the α in that case or minimize the error in your confidence interval. So, what is required is that either I increase the sample size like that or what we can do is that? We can just enhance the confidence band; enhance the confidence band or increase the confidence band like that.

So, as α decreases in that case what will happen is that my confidence band will increase like that. So, this extending and reducing this depends on the α values like that depends on the α value which will define the boundary conditions, which will define the upper bound and lower bound or the confidence interval which will define the confidence interval.

(Refer Slide Time: 03:13)



So, based on this concept itself this was extended in hypothesis testing which is nothing, but a statement or claims that is proved or disproved like that. Basically we are interested to generally assess some of the parameters and what is the value of the parameters, is it close to certain values is it different from certain values like that ok.

This is a simple example where you can see that population mean is given as 25, but I have got a value from the sample observation I have got a value of 20 over here. So this is the sample observation and this is the population and it is expected that the population from experience or something they expect that it should be close to 25 like that.

But, I have taken a sample observation over here and the X-bar observation that we are getting over here is around 20 ok. So, is 20 close to 25, can we say that although it is coming out to be 20, but we expect that if this is natural. So, when the population mean is at 25, this value is quite natural like that. So, this is a let us say normal distribution assumption and mu is said to be 25 and I have got some value which is 20 over here X-bar is I am getting a value of 20.

So is it very close to the mean value over here that is the thing that we want to check over here ok. So, and I will do it once like that I cannot take repeated samples like that privilege I do not have. So, with one sample observation I am getting a mean value of this. So, either I have to prove or disprove the claim that is made over here.

So, this is the statement that is I have to prove or disprove over here whether it is close to 25 or it is different from 25, that means, it is very far away from 25 like that ok. For these types of claims or CTQ whatever you can think of so any CTQ values and we expect some values and we are getting some values. So, is the center or has the center really shifted like that. So, in process control chart what we have seen is that.

So, we define some control limit line so this is based on some μ observation over here, that means, population μ observation is somewhere over here. But, I am getting some values mean observation over here, let us say average values over here. So, this control chart should give me a signal whenever is the whenever the mean is shifting like that. So, that, type of scenario also exists in control chart like that.

So I am proving and disproving like that so, it is in control and out of control like that. So, that is also done in control chart techniques like that ok. So, here I am taking a simple example over here I have some data observation or population mean age over here target population whatever it is.

So, in this case my expectation is it is near to 25, but I am getting 20 is 20 close to 25 that I want to prove or disprove like that ok. So, how do I do that by hypothesis testing? So, statistician has given us some options over here whether I am close to 25 or I am not, I am different from 25 basically.

(Refer Slide Time: 06:10)

Quality Control and Improvement using MINITAB

Null and Alternate Hypothesis

A non-statistical example

A criminal trial Example

- In a trial a judge must decide between two hypotheses.

The ~~null hypothesis~~ may be

H_0 : The defendant is innocent

The ~~alternative hypothesis~~ can be

H_1 : The defendant is guilty

The judge does not know which hypothesis to accept or reject. He must make a decision on the basis of evidence produced in the court.



Prof. Indrajit Mukherjee, SJMSOM, IIT Bombay

So, that can be done in hypothesis testing and. So, when we write the hypothesis testing what we do is that we say something is null statement or void statement over here which is known as H_0 and alternate statement over here which is known as H_1 like that.

So, whenever a person is taken to court in that case judge what he has to decide is that either the person is an innocent person or the person is guilty basically. So, there are two hypotheses over here. So, initially what judges assume that that person is innocent like that so, that is the null hypothesis or void statement like that and the alternate statement will be the person is guilty basically.

So, that will be the alternate statement like that ok. So, this is the way we define statements like that. So, this is a nonstatistical example we can think of. So, whenever somebody is in court so, either he is an innocent person or is a guilty person. So, either I go for one of the claims over here and or I disprove this one that the person is guilty like that.

So, based on the evidence that is provided by the police and other from the society what we gather like that we either prove or disprove the statement. So, over here statement is that person is innocent and the alternate statement is person is guilty like that.

(Refer Slide Time: 07:27)

Quality Control and Improvement using MINITAB

Null and Alternate Hypothesis

A statistical example

- Suppose that we are interested in the **mean burning rate (CTQ)** of the solid propellant.
- Now, **burning rate is a random variable** that can be described by a probability distribution.
- Suppose that our interest focuses on the **mean burning rate** (a parameter of this distribution).
- Specifically, we are interested in deciding whether or not the **mean burning rate is 50 centimeters per second**. We may express this formally as

$$\rightarrow H_0 : \mu = 50 \text{ cm/s}$$

$$\rightarrow H_1 : \mu \neq 50 \text{ cm/s}$$

The Null Hypothesis

The Alternate Hypothesis



Prof. Indrajit Mukherjee, SJMSOM, IIT Bombay

Source: Montgomery, D. C. (2005). *Applied statistical probability for engineers*. John Wiley & Sons



Similarly, in manufacturing example let us say and this is taken from Montgomery the mean burning strength I am interested in that is the CTQ, I am interested in over here and

we expect it should be near to 50 centimeters per second like that and we want to prove or disprove the statement like that. So, one is null statement over here one is alternate statement.

So, H_0 versus H_1 over here that is a null and alternate hypothesis that we try to check generally ok. So, we will have some evidence so, in that case what we will do is that we will collect some sample observation and based on that mean burning that we get over here $\bar{X}$ information and standard deviation information that we get based on that what we will do is that either we go for this statement over here μ equals to 50 or μ_0 naught equals to 50.

Now, my sample observation cannot be exactly equals to 50. Now, then the idea of confidence interval is used like that the similar idea is used over here to assess and I will give a upper bound and lower bound. If this statement is true and then I will see that if $\bar{X}$ falls within this or not does not fall within this. So, with the help of a state statistic what we will do is that? We will either go for this statement or go against this statement.

So, I am trying to prove whether it is equals to 50 or whether it is not equals to 50, this is null and this is alternate, generally null is written like that. So, sometimes we will see that greater than equals to on one side and less than like that. So null can be greater than equals to or less than equals to and the opposite will be greater than or less than conditions like that. So, this is also possible the way in many of the books you will see ok.

So, what I am trying to say over here is that one will be null one will be alternate hypothesis. So, I want to prove or disprove the statement like that ok, whether improvement has happened or improvement has not happened like that in any quality measures that we are doing. So, that is also an important issue when we say that we have done improvements whether we have really done or it is not statistically significant basically. So, that we prove or disprove like that.

(Refer Slide Time: 09:25)

Quality Control and Improvement using MINITAB

One-Sided and Two-Sided Hypothesis

Two-Sided

$$\begin{aligned} H_0 : \mu &= \mu_0 \\ H_1 : \mu &\neq \mu_0 \end{aligned}$$

One-Sided

$$\begin{aligned} H_0 : \mu &= \mu_0 \\ H_1 : \mu &< \mu_0 \end{aligned} \quad \text{or} \quad \begin{aligned} H_0 : \mu &= \mu_0 \\ H_1 : \mu &> \mu_0 \end{aligned}$$

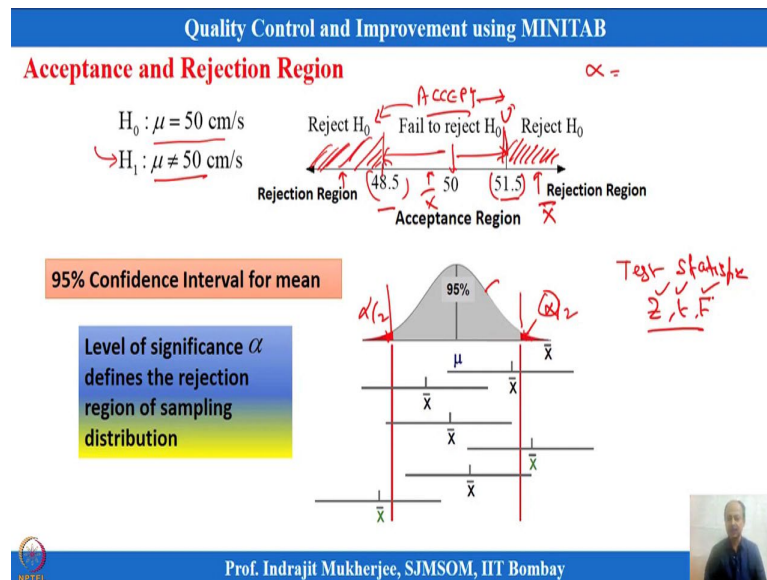
 Prof. Indrajit Mukherjee, SJMSOM, IIT Bombay 

So, for that we use hypothesis testing concept like that. So, hypothesis testing can be two sided hypothesis testing that means, what does it say? That not equals to condition when this is the condition when we say that it can go on either side of μ_0 like that. So, either it is equals to or not equals to. So, it can be greater than it can be less than conditions like that.

Hypothesis testing can also be one-sided over here, that means, this is one-sided over here it is less than equals to or it is less than or it is greater than like that. So, it is strict condition that we are imposing over here either less than or greater than like that. This is one-sided statement what we say one-sided statement that we make ok.

So, and if this is the less than condition then in that case this should be greater than equals to condition basically that this should be the other condition. So, I can write equals to or equivalent so greater than equals to ok. If it is a continuous distribution in that case it does not make much difference. So whether it is equals to greater than equals to condition like that ok.

(Refer Slide Time: 10:26)



So, in this case we have a concept of hypothesis testing. So, what statistician has given us that if the mean burning rate lies over here then I can define a boundary condition over here. I can define that this is the upper condition and this is the lower condition and this is the acceptance zone we can say that this is accept zone and this will be the rejection zone this will be the rejection zone over here.

So, what statistician says is that I can define this one, now with a level of significance α , that means, with a confidence level of α over here. That means, I can be wrong; I can be wrong, but I am defining a zone over here and if my value of $\bar{X}$ fall within this whatever sample statistic I have calculated. If it falls within this zone what we can say is that it is not different from 50 basically it is not different from 50.

But if some observation coming out over here which is very far away over here and that is not lying within the acceptance zone over here. And that is defined by this to this zone over here, then in that case we will say this is unnatural observation this is unexpected. So, in that case we expect that the population mean is not 50 at least.

So if you are seeing extreme observation which is away from this one in that case what will happen is that we can say that these observations are unusual. So, in that case mean this statement over here μ equals to 50 does not hold true. So, I will go by this statement over here and I have evidence to go for this μ not equals to 50 conditions like that.

Because and I but, however, hypothesis testing always whatever decision I am taking so, in hypothesis testing what we are doing is a we are making inference basically over here. So, hypothesis testing is all about inference. So earlier what we are doing is that we are not taking any conclusive decisions in the earlier cases.

But now we want to accept or reject some statements like that whether improvement has happened not happen based on statistical evidence we want to conclude basically ok. So, and that is done in quality everywhere we have to decide. So, control chart also we have to decide stable unstable, stable unstable scenarios like that for that there also we are doing hypothesis testing basically.

Unknowingly, we have not mentioned we I have not mentioned that one, but you can also think about that as a hypothesis testing what we are doing with the control limit lines the UCL and LCL like that ok. So, over here, so how they are defining this upper limit and lower limits. So, that is based on confidence interval that α level of significance that you have considered earlier also.

So, $\alpha/2$ on this side and $\alpha/2$ both sides test not equals to condition. So, in this case rejection region will be on one side it can be also on the other side over here. So, in this case with $\alpha/2$ and based on the distribution assumptions over here what we can do is that? We can draw the demarcation line like this. Whatever is drawn 48.5 this value can be gathered based on α level of significance.

And what is the distribution we have to, we can assume over here what distribution it is falling based on that we can define the upper limits and we can define the lower limits for acceptance zone and the rejection zone is beyond that basically rejection zone is beyond that like that ok.

And for doing hypothesis testing, so they will use a test statistic basically they will use a test statistics like that test statistic. So, this test statistic can be Z state test statistics t test statistics like that so, F statistics so many options are there over here.

So, generally Z distribution t distribution F distribution these are the things that is this type of test statistics are generally used to make a conclusion on mean or variance like that ok. So, what is the formulation of that, so I have to define the reason and based on

that if it goes outside that if any value of the CTQ goes outside that we will accept or reject the statement like that ok.

(Refer Slide Time: 14:12)

Quality Control and Improvement using MINITAB

Type-I and Type-II Error

- **Type-I Error:** Reject Null hypothesis, when it is true.
- **Type-II Error:** Fail to reject Null hypothesis, when it is false.

Decision	If H_0 is True	If H_0 is False
Fail to reject H_0	No error	Type II error (β)
Reject H_0	Type I error (α)	No error

$\alpha = P(\text{type I error}) = P(\text{reject } H_0 \text{ when } H_0 \text{ is true})$

So, and whenever we are doing this that I told that there will be an error associated with this is known as Type-I error which is easy to see also, that means, if I am; if I am defining the zone over here what is possible over here is that although the mean is at μ_0 let us say 50 or something like that over here ok.

So, I have defined the zone over here. So, and the values $\bar{x}$ that we are getting can fall anywhere over here anywhere over here, but my value has fallen over here. So, I have rejected the null hypothesis and gone by the alternate hypothesis like that. But you can also think of think about if it is a normal distribution some values will always fall outside the boundary conditions that you have specified with a given level of α over here.

So, with a given level of α , $(1 - \alpha/2)$ on this side or $\alpha/2$ on this side region so, in that case area over here. So, in this case $\alpha/2$. So, in this case what can happen is that? Some values naturally can go outside also; some values can naturally go outside also.

So, that type of error when I make when μ equals to μ_0 , but I reject this hypothesis when I reject this hypothesis. Although this $\bar{x}$ what value I have got may be natural over here because some values will always go outside this because normal distribution does

not touch the axis over here. So, in this case some error I am committing over here this is known as Type-I error. So, α is also known as Type-I error over here.

So, and Type-I error talks about rejecting the null when it is true basically although, the μ is at μ_0 , but I am rejecting this hypothesis over here. I am rejecting the null over here, that means, although I have got a value which is falling outside the acceptance zone over here, but that can be completely natural. So, that type of miss judgment can happen 5 percent of the time which I have taken α value of 5 percent over here.

A level of significance is considered which is my acceptance zone, which is my rejection zone that will be defined by what is the value of α I have assumed in my experimentation basically ok. So, 5 percent what we are assuming over here most of the time you will find people are taking 95 percent as the, this acceptance zone for defining the acceptance zone and 5 percent is the rejection zone basically when it is defined.

So, it can be one sided or it can be both sided test by whenever I am doing one sided and both sided. So, when both sided test we are doing so, area can be on this side rejection region can be on both side if I am doing one sided test on the greater than this one greater than condition. So, it can be higher than this side.

And if it is less than condition so, μ is greater than when I am trying to prove some values 50 over here. So, this is the region, so this will be 5 percent over here. And when I am doing two sided test over here. This will be 2.5 percent on this side this can be 2.5 percent on this side like that so, one on this side one on this side.

So, in this case this will be α by 2 on this side and this will be α by 2 on this side this is complete α on one side. So, this is completely because other side I am not interested. So, all the weightage I will put on one side like that ok. So all the this Type-I error should be on one side. So, that, that is a that is the idea over here.

So, in this case Type-I error is something that we reject the null hypothesis when it is true this is very serious we reject that. So, when we are talking about let us say we are making some drugs and in that case now, and void statement is it does not have any effect an alternate statement is it has an effect it has an effect. So, in case I reject the null when it is, when it is true, that means, the drug is not effective, but I am saying it is effective basically.

So, that, that is a very serious condition like that. So, Type-I error is very serious like that. So, in this case if I am committing such kind of error and I want to minimize this one how can I minimize this one?

Increase the sample size increase the sample size that is mathematically we can show that if you increase the sample size in that case what will happen is that your judgment will be fine because you see apple tree more apples you take in that case judgment will always be quite close to the population average like that.

So, one easiest way is to how you can reduce the Type-I error over here I in hypothesis testing you increase the sample size, but you cannot increase so much that it is it should be cost effective also. So, there is optimal sample size calculation also when we are doing hypothesis testing when we are doing hypothesis testing like that.

So, if I can demark it like that there are two kinds of error that can happen in hypothesis testing. So, you can see some other lectures on hypothesis testing which is easily available in any or any web course where you can NPTEL web course.

So, where we have already given about details of hypothesis testing like that or any lectures which is which you can see from YouTube's or some other where detail explanation of Type-I and Type-II error is given I have given a brief idea that type one error is more serious.

And it talks about rejecting the null when it is true basically void condition I am rejecting although it is true, that means, there is no improvement, but I am saying that there is improvement like that. So, that is more serious in my opinion ok. So, then there is a Type-II error also, I failed to reject, that means, I accept the null over here.

So, although the drug is effective in that case I say no the drug is not effective like that. So, that is the Type-II error that we can commit over here, when the hypothesis is false we have to go for the alternate, but we have not gone because the data does not show like that and we have made a miss judgment over here, we have also taken a miss or decision is incorrect like that. So, that can also go wrong ok.

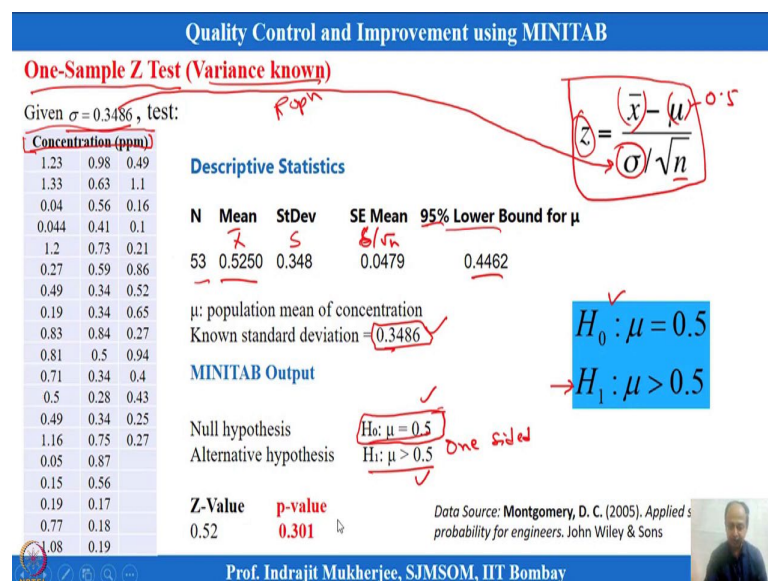
So, I fail to reject the null when it is when it is false. So, that is known as beta error over here one is α error one is beta error over here, α error is what we should be concerned

about ok while doing hypothesis testing like that. To reduce the beta error or Type-II error over here, again the same and Type-II error is not easy to calculate because where the mean has shifted that information is required.

So, what I will suggest is that we define the α and based on that we do the test. We want to minimize this one if you have to minimize this one and you increase the sample size that is the simplest way. So, more and more sample size α and beta both can be reduced like that ok.

So, that is the what we do in hypothesis testing basically ok one is Type-I error please remember one is Type-I error and Type-II error. So, α based on α we make a judgment basically when we and we increase the sample size to reduce the beta error over there ok. So, and that is the one way we can go about it ok. So, hypothesis testing is all about testing some statement whether it is, so one will be null statement one will be alternate statements like that ok.

(Refer Slide Time: 20:51)



So, first important test that we can think of in hypothesis testing is one sample Z test over here, ok one sample Z test. So, one example is taken where sigma is known over here sigma is known and some consideration is information last time also we have seen. So, for defining the confidence interval what we have taken is that some sample observation was taken over here and this example is taken from Montgomery again ok.

So, this data information gives me a descriptive statistics what we have seen is that mean value of 0.524 5 to 5. This sample observation what observations we have based on that $\bar{x}$ was calculated and S standard deviation was calculated like that and standard error is nothing, $\sigma/\sqrt{n}$ over here.

So and bound condition can also be defined. So, this is also defined lower bound like that and this is the 95 percent confidence interval that we have considered over here. And now we want to test, so known standard deviation it is saying that there is a known standard deviation of this value over here now I want to test.

So, whenever variance is known there can be scenarios where variance is known population variance basically population variance is known and population variance is unknown. So, in this case population variance is given over here and I want to test the null hypothesis mu equals to this one or it is greater than. So, one-sided test one-sided test I want to do basically over here ok.

So, and in this case I want to test that whether the mean is equals to 50 or mean is differing from 50 like that. Because I am interested in characteristics that means, the outcome over here, concentration should be close to 0.5, that may be the target what is given by the ok customer like that it should be close to target. And I want to show that statistically to the customer like that. So, and this data set, so one test statistics is used test statistics is used over here.

So, this is the test statistics that is used to confirm whether the whether I will go for this statement or I will go for this statement like that. So X-bar can be calculated over here and hypothetical mean is over here as 0.5 and this sigma is also given which is also mentioned over here. So, this information will go over here.

(Refer Slide Time: 23:02)

Quality Control and Improvement using MINITAB

P-value :

P-value : '**Smallest**' level of significance that would lead to rejection of Null Hypothesis. (or '**observed**' level of significance)

- Compare the p-value with the selected value of α
- If the **p-value is less than or equal** to the level of significance, we judge the p-value to be small enough to reject the null hypothesis.
- If the **p-value is greater than level of significance**, we do not reject the null hypothesis.

 Prof. Indrajit Mukherjee, SJMSOM, IIT Bombay 

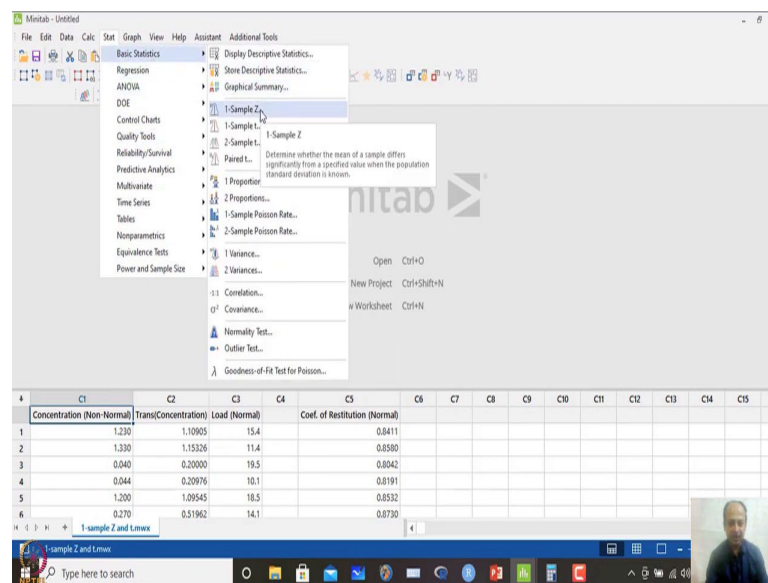
And n information is also given because number of observations also is known to me number of observations is also known to me over here. And based on that I have to make a conclusion based on the Z value I have to make a conclusion over here, whether I will go for this statement or whether this statement is true like that ok.

How do we do it in MINITAB? So, I will show you that one and beforehand I will also like to say that this Z test is used assuming that the variable for normal distribution although this is robust because I am making a Z conversion over here from the mean value. So, we expect that the test statistics will work over here and in this case I am making a Z conversion over here. So, $\bar{x}$ is converted into Z and I am using a standard normal distribution to get that one.

But, underlying assumption is that this is not highly skewed this data is not highly skewed and it follows normal distribution then we are assuming that that to be true when we are applying Z test over here ok.

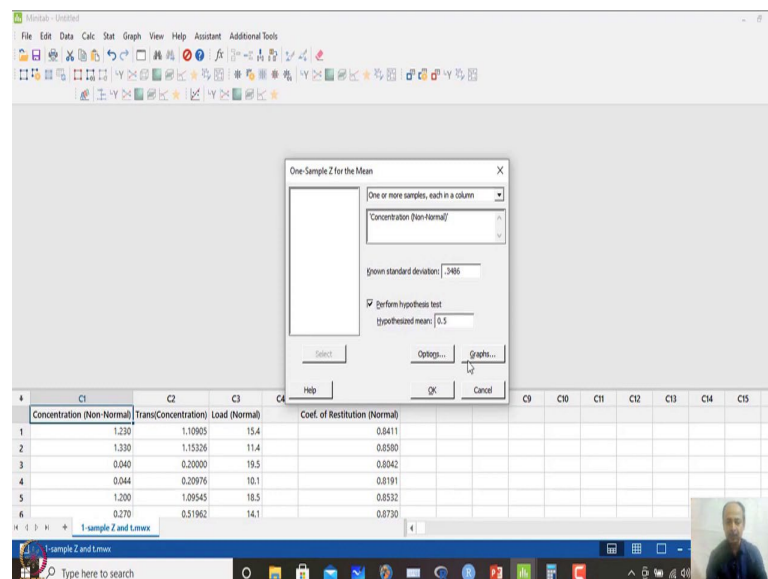
Let us try to see assuming that this is true and we do the one sample Z test over here, to see whether it is $\mu = 0.5$ or $\mu_0 = 0.5$ like that. So, then for that doing the test what is required is that I need to see some other important aspects like what I mentioned is P-value that that we have to see take a note of that.

(Refer Slide Time: 24:21)



So, now this is the data set that we are having. So, concentration is the data set that I am using over here. So, what I will do is that? I will go to statistics and then what I will do is that basic statistics and then I will go to one sample Z. I will go to one sample assuming that it follows normal distribution over here.

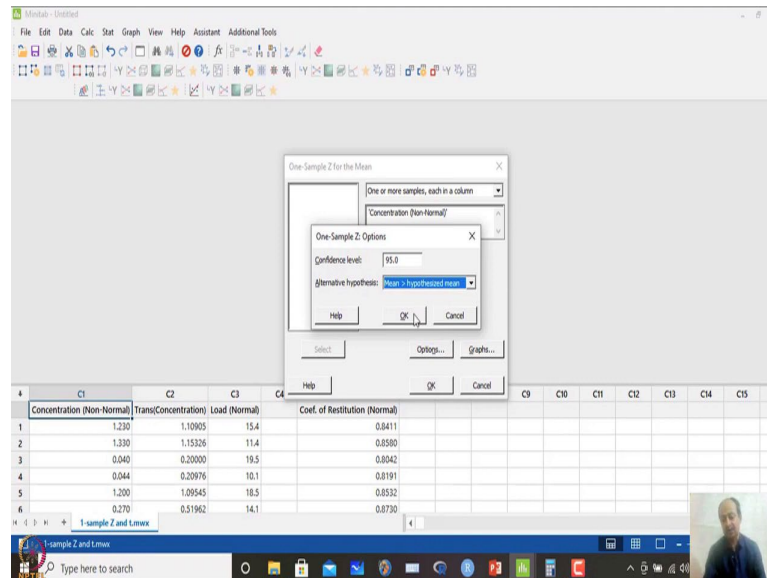
(Refer Slide Time: 24:36)



So, in this case I will take concentration over here and known standard deviation is given over here as, if I go back to the 0.3486 that is the standard deviation that is given over here and I have mentioned that one over here. So, I have just typed this one. So, perform

hypothesis testing this I have clicked over here and hypothesis that is that we want to test is whether it is 0.5 or whether it is different from 0.5 like that.

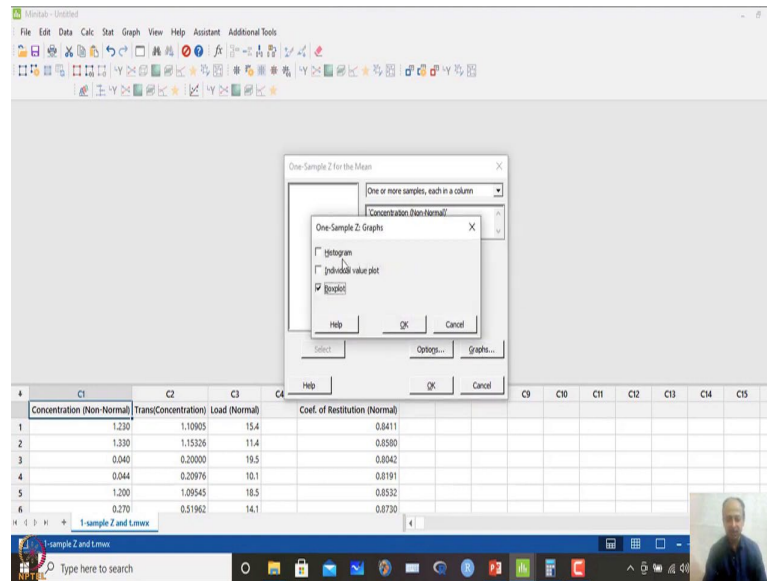
(Refer Slide Time: 25:12)



So, I have just ticked this one. So, options may not be there, so you have to click this one and then right over here 0.5, 0.5 like that. So, when you do that then you go to the options and the confidence level to define the confidence interval this is the confidence interval 95 percent.

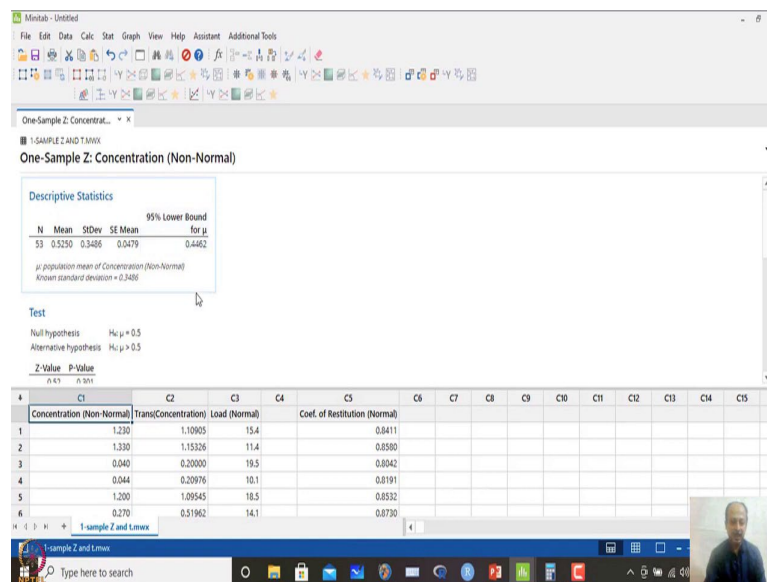
So, level of confidence, confidence level is 95 percent means 5 percent is the our that is level of significance that is the level of significance over here ok. So, I want to test whether mean is greater than condition. So, mean not equals to is given over here less than is given over here mean is greater than. So, I want to test mean is greater than one-sided test I want to do and with a α level of significance considering over here as 5 percent like that I click ok.

(Refer Slide Time: 25:47)

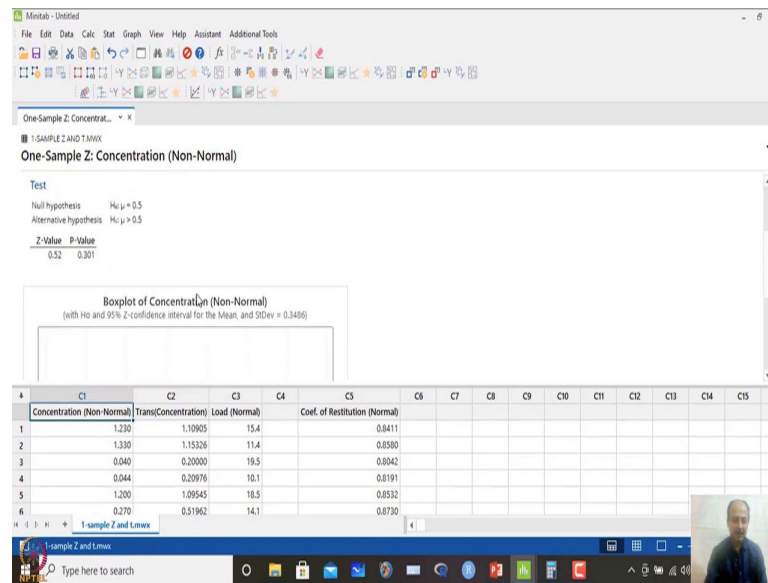


And graph what you can do is that you can see box plot of this data like that and histogram is also possible. So, over here and I click ok.

(Refer Slide Time: 25:56)



(Refer Slide Time: 25:57)



And when I do that what I see is that this test was performed over here null and the corresponding Z value that is the Z statistics was calculated. So, this Z value what you see this Z value was calculated over here. And that is reported by MINITAB, in this analysis that is reported as 0.52, but whether to accept null or to reject null we will make a decision based on this P-value and let us remember the P-value is 0.301 over here ok.

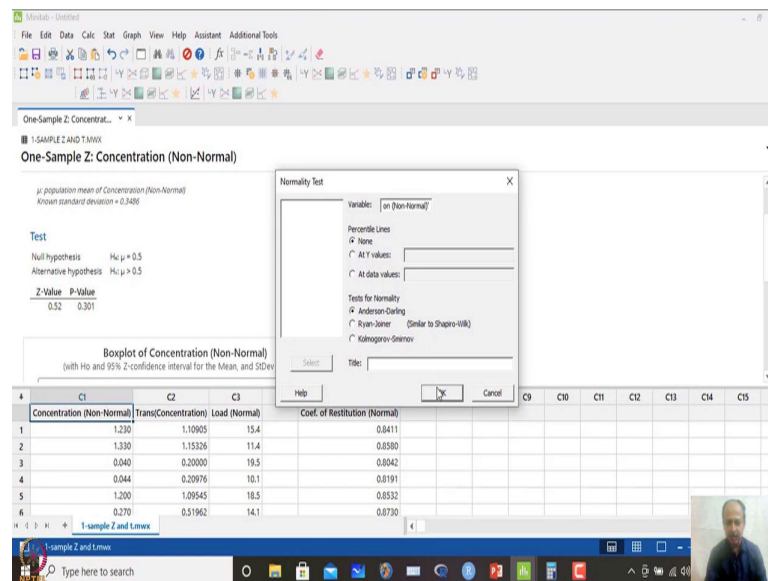
So, if the P-value is more than 0.05 you can note it down if the P-value is more than 0.05, in that case we cannot reject the null hypothesis ok. So, only if P-value is less than 0.05, then we will because 5 percent is the level of significance we are considering if the P-value is less than 0.05. Then we will reject the null hypothesis and we will go by the alternate hypothesis over here.

As the P-value is more than 0.05, so we cannot reject that one, that means, concentration is not different from 0.5 basically over here. So, in this case we will accept the null hypothesis and we will say we do not have evidence to prove that μ is very different from 0.5, statistically we cannot prove that. So, we will go by the statement null hypothesis over here.

So, this is one thing that we have to note, but also we have to note that before we do this test Z test what is required is that this data should be normally distributed like that. Now, while this data set is taken from example is taken from Montgomery over here, but we can also test whether it is normal or not.

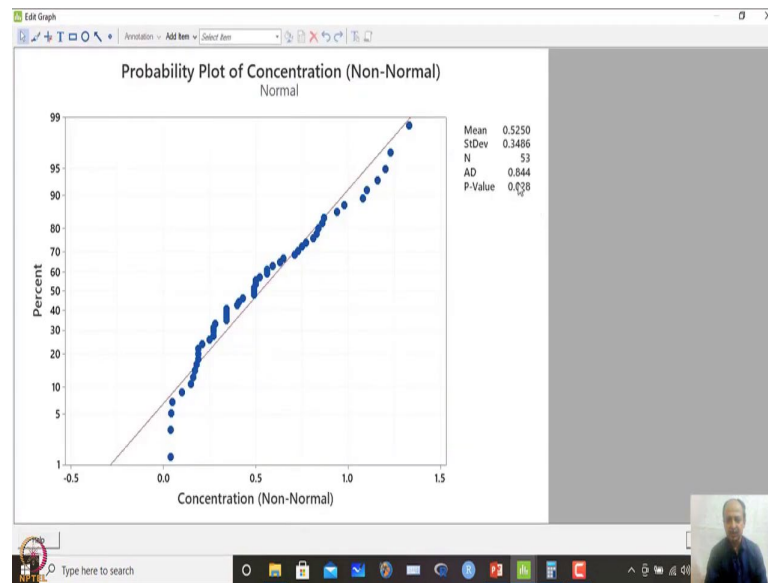
So, in this case what we will do is that just to cross check, but I am saying that this Z test is quite robust ok. Even if some small deviation happens that will not affect the outcome of the test, but we have alternatives also over here to do something more over here. So, what we can do is that? We can go to stat and what we can do is that in control chart I told there is a transformation that is possible.

(Refer Slide Time: 27:52)



But, first let us try to see whether it is normal or nonnormal, basic statistics I will go to normality test and in normality test what we will do is that we will take concentration over here and I will do Anderson-Darling test.

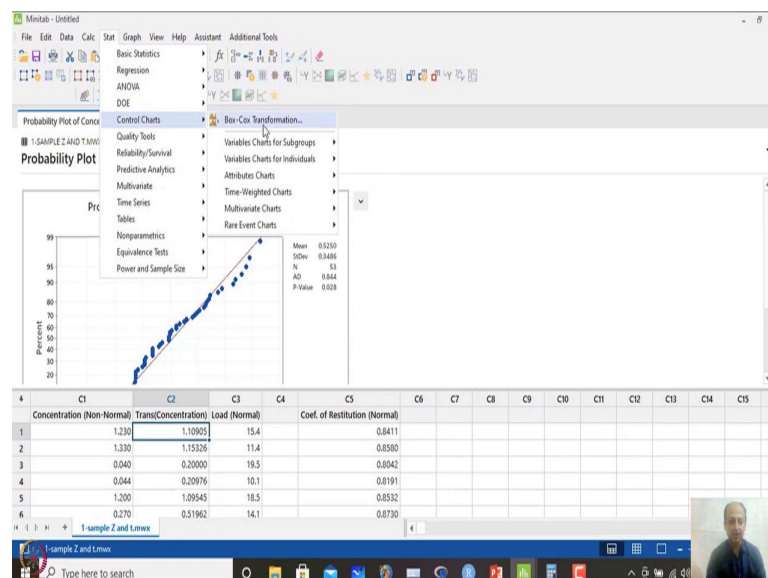
(Refer Slide Time: 27:56)



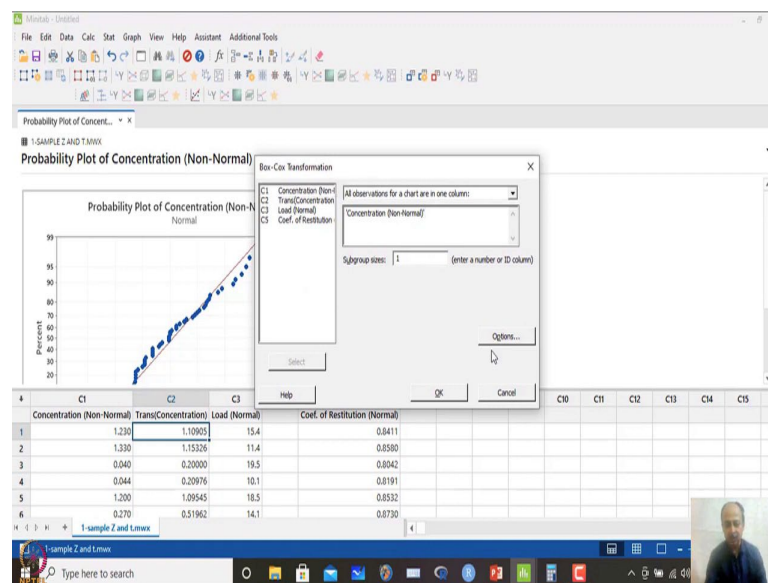
So, when I do Anderson-Darling test, what happens is that what we see is that P-value is less than 0.05 over here, when P-value is less than 0.05 this data is nonnormal, when this data is non normal it will deviate from this line central line that you see straight line that you see. So, deviation has happened.

So, it is non-normal scenarios that we are facing with this data set over here. So, if it is not normal if it is a non-normal in that case what we can do is that we can transform the data.

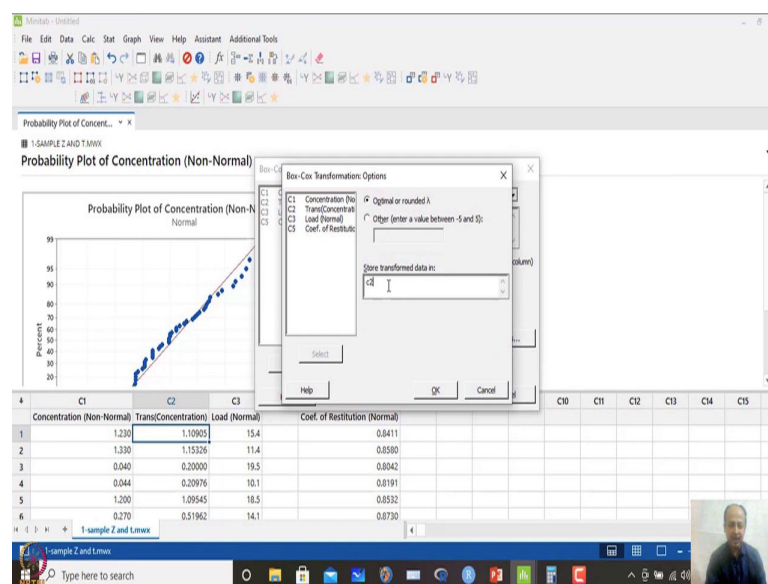
(Refer Slide Time: 28:30)



(Refer Slide Time: 28:32)

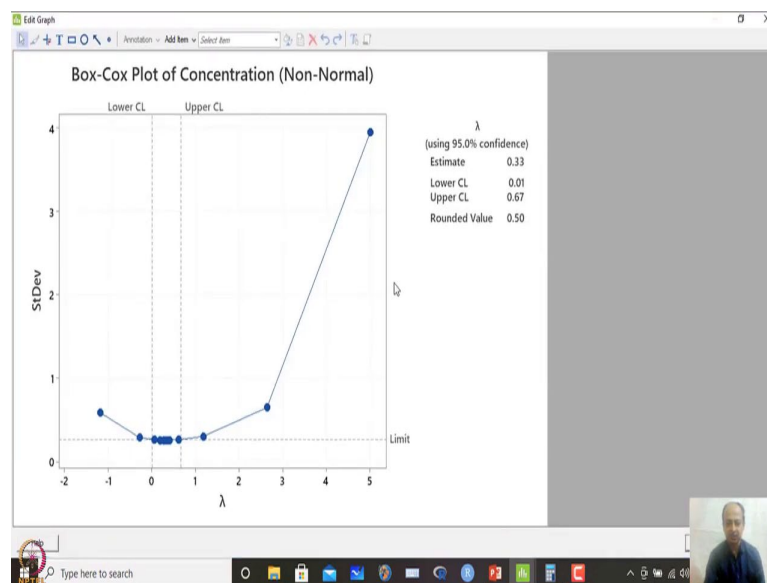


(Refer Slide Time: 28:38)



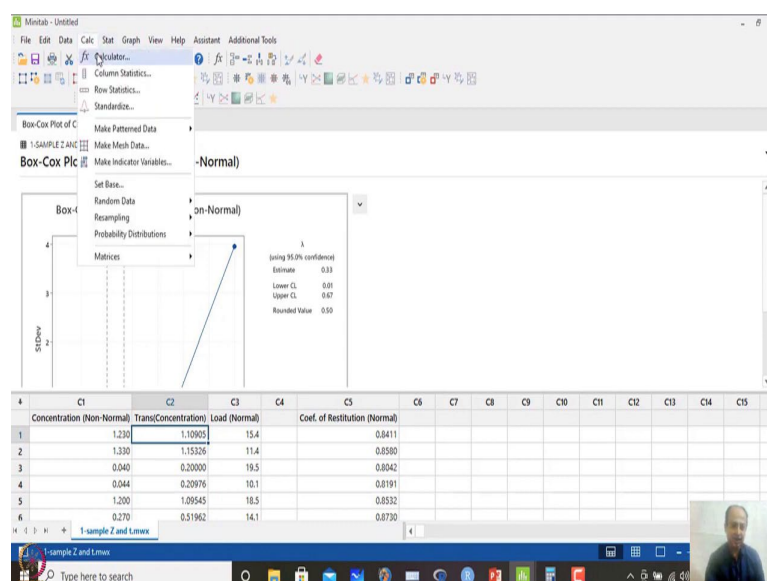
So, what was done is that in this case, so I can transform the data how do I transform the data I go to control chart I go for box-cox transformation over here. So, what I did is that box-cox transformation on this data set, so I highlighted the data subgroup size is 1 over here.

(Refer Slide Time: 28:46)

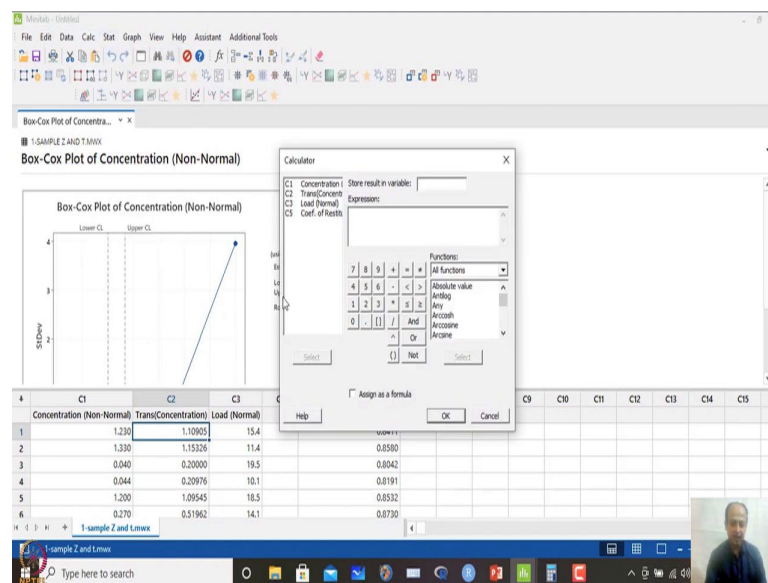


So, in this case options where I should save I will save in c 2 like that. So, I have done c 2 over here and I click ok and when I do that ok what happens is that this graph comes over here. So, this is the box-cox transformation that is recommended over here and the recommended value that you can see is 0.5, when recommended y to the power lambda is 0.5 that represents square root transformation basically that represents square root transformation.

(Refer Slide Time: 29:09)



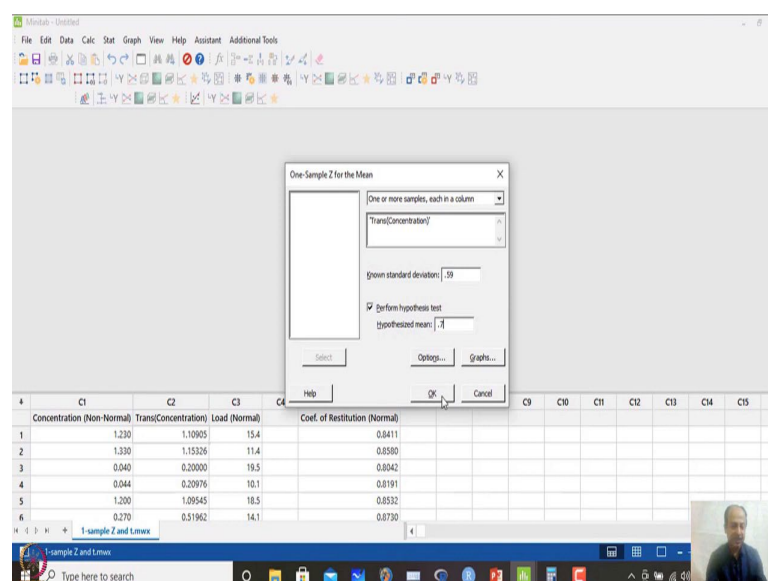
(Refer Slide Time: 29:10)



So, what was done is that one square root transformation was done over here, this can also be done by using other options over here, that means, calculate and calculator we can use and the columns where I want to save. So, this can be done in calculators over here, so that is also possible.

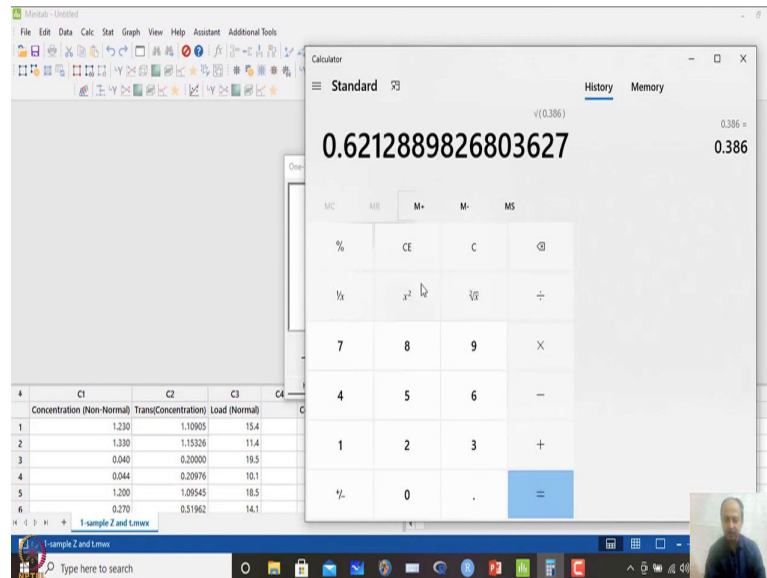
But, this is the square root transform value over here. So, this is the square root transform value over here. So, then what you can do this is a transform value. So, then what we can do is that we can do the same Z statistics 1 sample on the transform data basically.

(Refer Slide Time: 29:31)



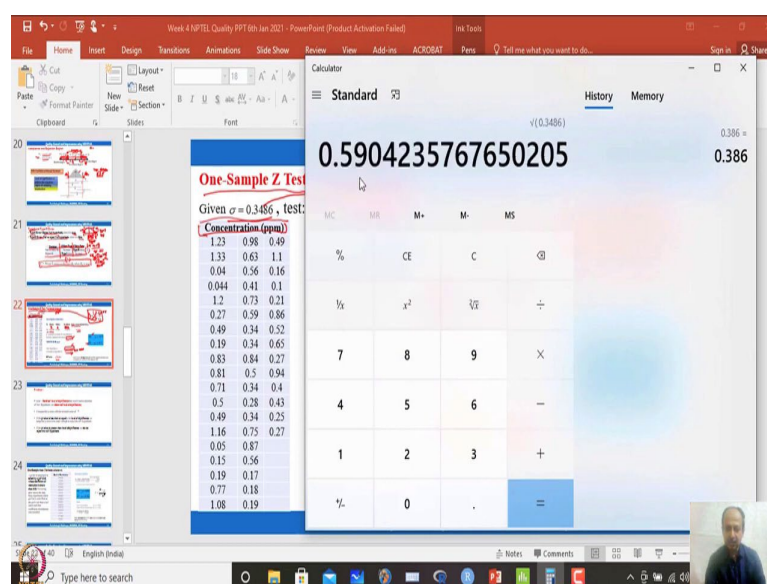
So, I will do 1 sample Z. So, in this case what I will do is that transform data I will consider over here and instead of 0.34, I will take square root transformation over here.

(Refer Slide Time: 29:41)



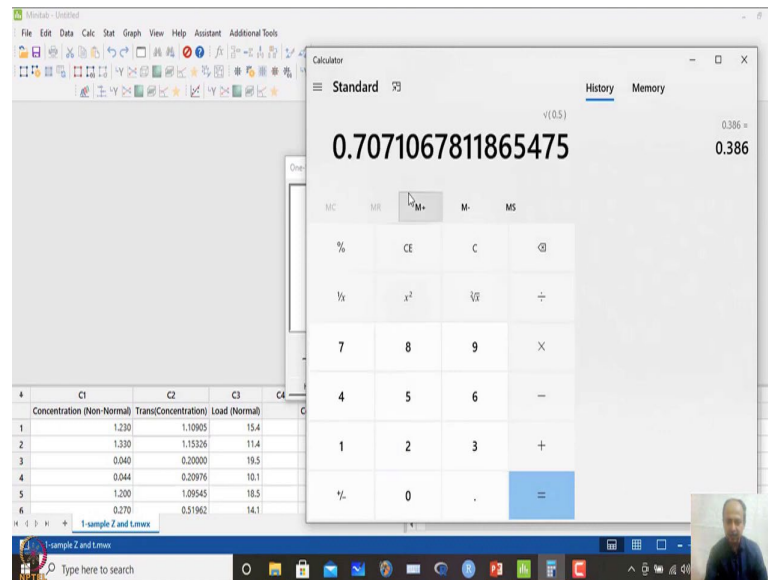
So, if I go to calculators over here 0.3486 or something that the value that we have considered over here. So, this can be I can take square root transformation over here. So, let me just cross check what was the value 3486. So, 3486 sorry this needs to be corrected over here.

(Refer Slide Time: 29:59)



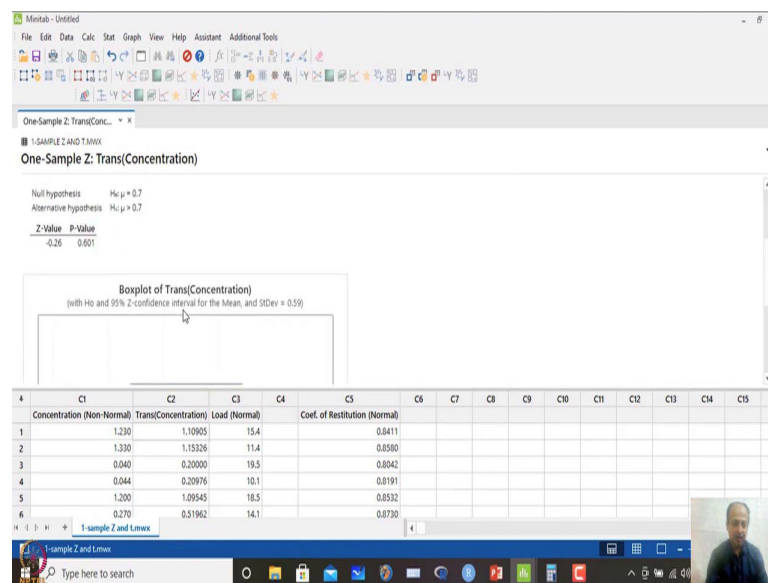
So, 0.3486 over here and I take a square root value of this. So, 0.59 approximately equals to 0.59 will be the standard deviation and let us do 0.59. So, what is required is that? I will type over here as 0.59 that is the square root over here and perform hypothesis 0.5 has to be also converted.

(Refer Slide Time: 30:22)



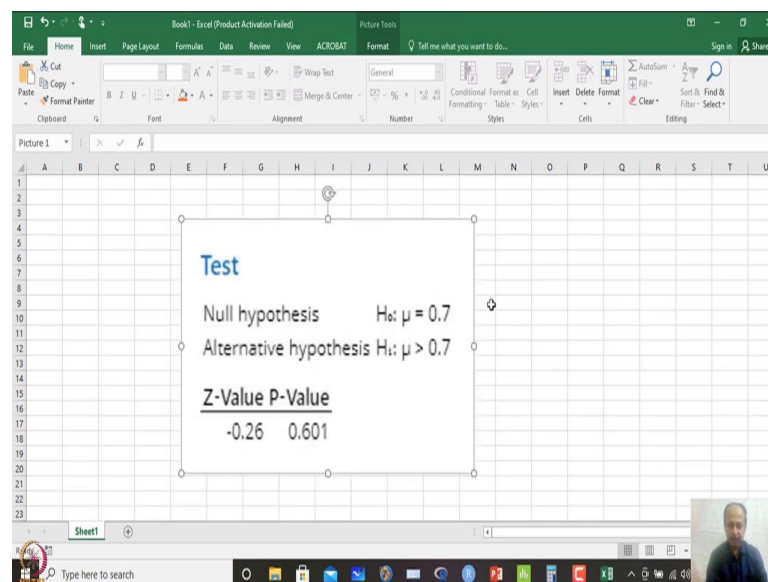
So, what is the corresponding value of 0.5 square root of that 0.5 if I take and then take a square root of this and the value is 0.707. So, what I will do is that point I want to test 0.7 something like that, so this is the converted values. So, I am doing the same test only I have done the conversion of the CTQs basically over here.

(Refer Slide Time: 30:42)



So, when you do the options over here and do the testing over here the analysis P-value is again you see the P-value when I am let me just copy this one. So, if I copy this as a picture over here and I paste it in let us say excel.

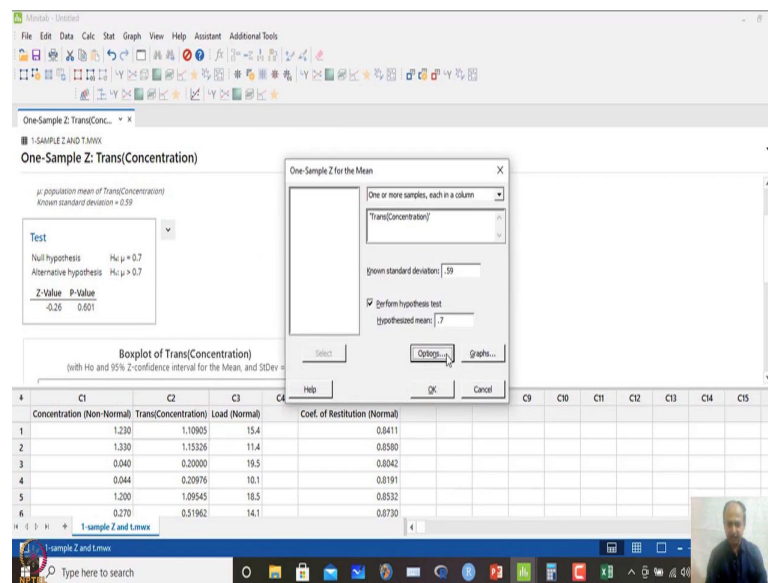
(Refer Slide Time: 31:19)



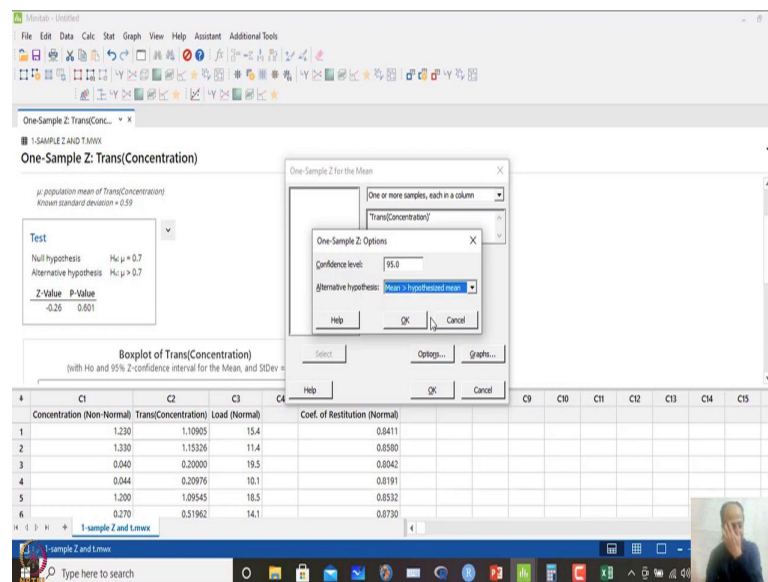
So, we can do that and we can enlarge the image like that. So, if I am doing that let us try to see what happens. So, this may be quite small window what we are seeing. So, what I will do is that I will just show you what is the results that we are getting over here ok. So, I will open a blank sheet over here and I will copy paste this one.

So, and I will enlarge this one. So, you can see that Z statistics is coming out to be this value and P-value is coming out to be 0.601. So, and we are doing one-sided test over here. So, one-sided test when we are doing one-sided test what we have given is that in this hypothesis testing I have mentioned that options it should be greater than condition.

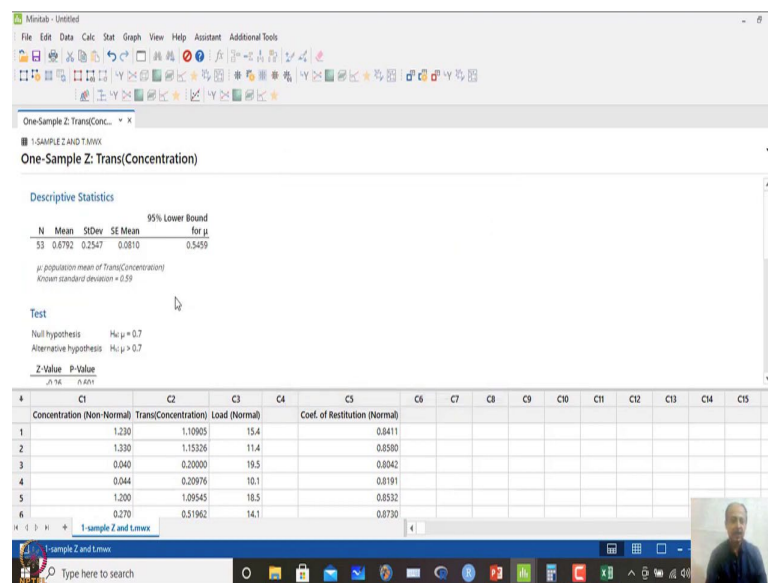
(Refer Slide Time: 31:41)



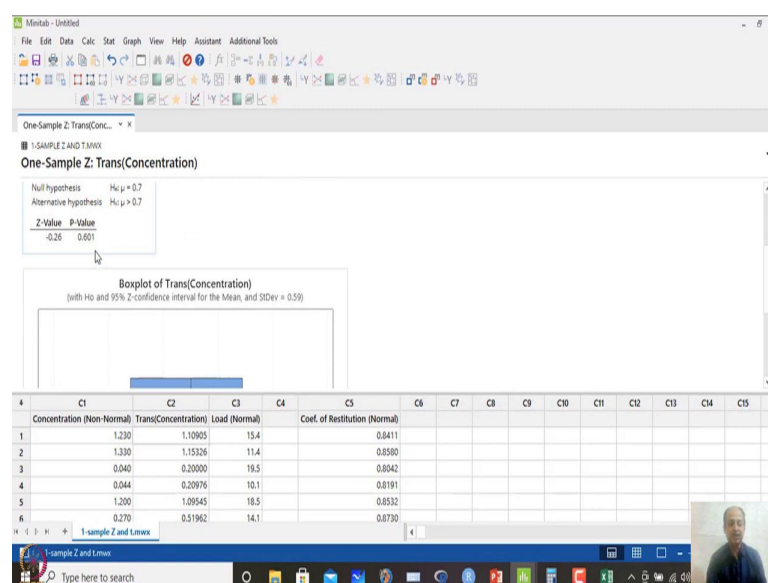
(Refer Slide Time: 31:42)



(Refer Slide Time: 31:49)



(Refer Slide Time: 31:50)



So, this you have to remember what you are doing one-sided test or both sided test like that. So, if you click ok you will get the results. So, even after transformation what is happening is that? I am going by the same decision. So, I have not transformed and the decision was go for H_0 , and if I am transforming also I am going for H_0 like that ok, if transformation works if transformation works.

So, in case transformation does not work then we will see what is to be done in our next session like that if transformation does not work what is the options that we have that we

will see in the next sessions like that ok. So, thank you for listening we will continue from here and we will see some more examples of hypothesis testing ok.

Thank you.