Statistical Learning for Reliability Analysis
Prof. Monalisa Sarma
Subir Chowdhury School of Quality and Reliability
Indian Institute of Technology - Kharagpur

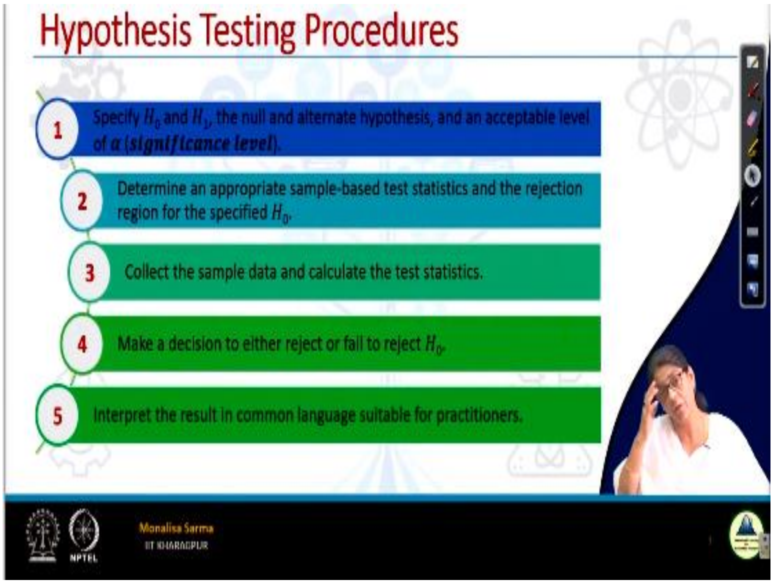
Lecture - 23
Statistical Inference (Part 2)

Welcome back guys so, today in continuation of our earlier lecture on statistical inference, today is the second lecture on statistical inference.

(Refer Slide Time: 00:35)

So, in this lecture, we will learn about hypothesis testing procedure. And we will see some test cases I mean some case studies for hypothesis testing and we will also see what is p value in the context of hypothesis testing.

(Refer Slide Time: 00:48)



So, like in the last class, if you remember, we have discussed what do we mean by statistical inference. And what are the different types of error that can happen in a statistical inference how to form the hypothesis, what does rejection region mean rejection region and critical region what are the same thing actually, what does this mean those we have discussed in my last lecture. So, today will directly come to the hypothesis testing procedure.

So, there are many steps in hypothesis testing procedure many means, there are total 5 steps. So, what is the first step? First step is that we have to specify H 0 and H 1 what is H 0 and H 1 I have already mentioned in my last class, where we specify H 0 and H 1 the null and alternate hypothesis and an acceptable level of α what is α? Remember, α is called the probability of type 1 error. What is type 1 error?

Type 1 error is that we are rejecting a true null hypothesis when a null hypothesis is true and we are rejecting it that is the called type 1 error. So, α is the probability of that and this α is also called significance level. So, since this α is called significance level in hypothesis testing procedure, we use the significance level. So, this hypothesis testing is also called significance testing.

So now, then the after specifying H 0, H 1 null and alternate hypothesis then what we need to do, we need to find out an appropriate sample based test statistics. We will have to find from the sample will have to calculate the test statistic and the rejection region for the specified H 0. After calculating the test statistics in a second step, what we will do? We will find out what is an appropriate test statistics.

Like when I talk about appropriate test statistics means, if I want to do sampling distribute if I were to infer about the population mean, then my test statistics maybe Z and my test statistic maybe T value also depending on the whether, I know the standard deviation of the population or not, then again if I want to infer something about the population variance, then my sample test statistics will be chi square value. If I want to compare 2 variants, then my test statistics will be F value.

So, accordingly based on the problem, we will find out by sample based test statistics and the rejection region. How do I find out a rejection region? Rejection region is I find the rejection region based on the value of α that is the significance level that we have already seen in the last class. Then, we will collect the sample data calculate the test statistics. Next step after calculating the test statistics.

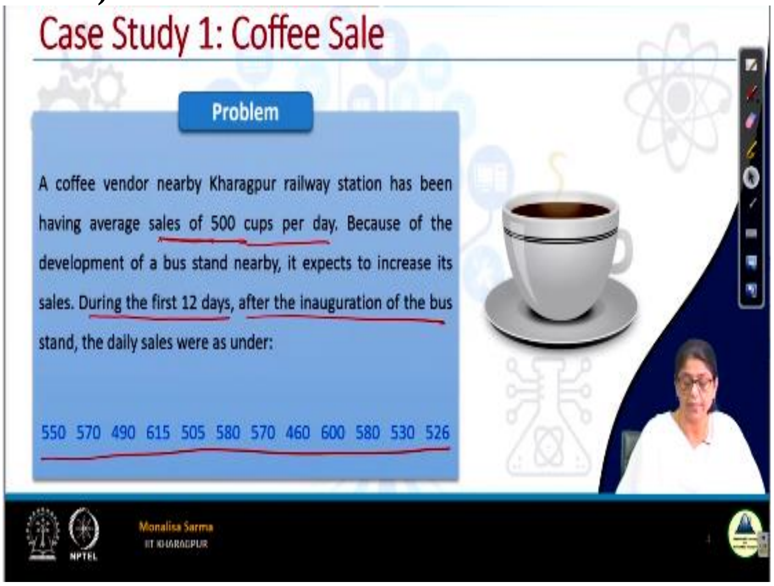
Since we know what is the rejection region then based on a test statistic we will be able to say whether it falls in a rejection region or it does not falls in the rejection region. If it falls so, based on that we will make a decision to either reject or fail to reject H 0. Then finally, the last step is interpret the result in common language suitable for practitioners. After once I have reject the null hypothesis or we do not reject a null hypothesis.

Then after that, we will interpret the result in common language that is depending on the application. Now, from point number 4 when we make a decision to either reject or fail to reject, here I want to bring to your notice 1 important point is that one thing. If my null hypothesis is rejected, then definitely null hypothesis is rejected means the alternate hypothesis is accepted there is no other way out.

But if my null hypothesis is not rejected, that does not only mean that I accept my null hypothesis, it may mean I accept my null hypothesis, or I can also say failure to reject the null hypothesis unable to reject the null hypothesis. Both are accepting the null hypothesis or unable to reject the null hypothesis both are not different and both are not same. We will see in some examples how in what way it is different you may think both are the same thing.

We are not rejecting the null hypothesis means we are accepting no that is not both are 2 different perspectives. We will see with examples.

(Refer Slide Time: 05:08)

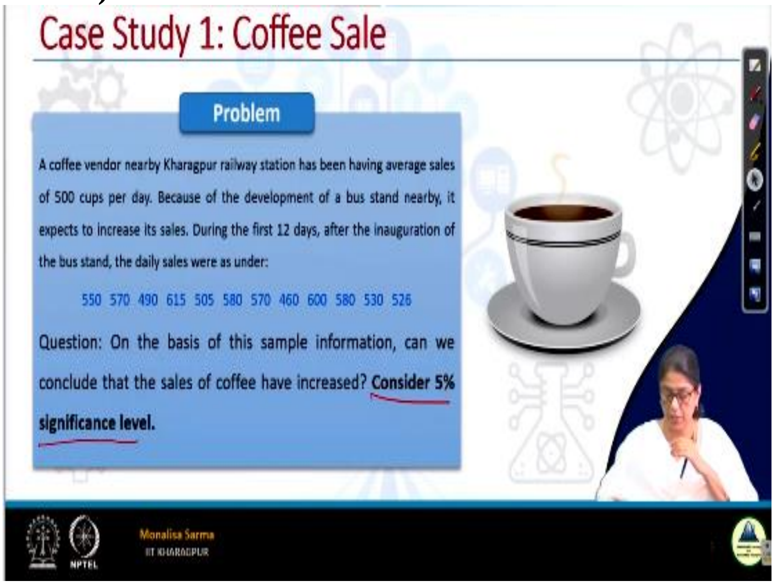


So, now this hypothesis testing procedure, we will see with some very simple example. First is a very simple example see, what it is given a coffee vendor nearby Kharagpur railway station since I am from Kharagpur I have mentioned it is Kharagpur railway station it can be any railway station. So, having an average sale of 500 cups per day this pen is really irritant do you know. So, as an average sale of 500 cups per day that is a hypothesized value.

Now, we will be using the hypothesized value. Because it has been this coffee shop has been running for ages it is not possible for a person to daily find out the average and come up with a service based on an as we know hypothesis, maybe an educated guess or based on some solid evidence. So, it can be hypothesized that sale is around 500 cups per day that means mean value is 500.

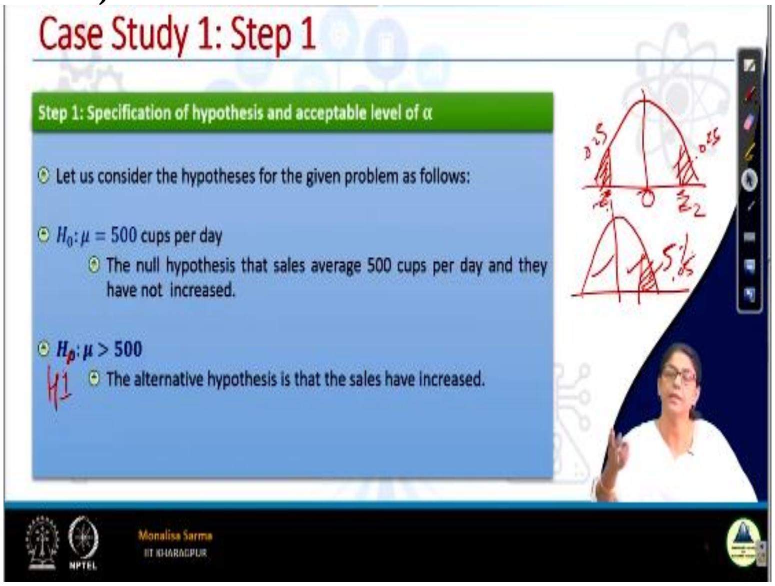
I can say mean of the population that is μ is 500 because of the development of a bus stand nearby it expects to increase it sells, obviously, bus stand is there so, many people will come. So, during the first 12 days after the inauguration of the bus then that daily sales were as under so, after the inauguration of bus stand, I noticed that sales amount of sales per day for 12 days and this is the data.

(Refer Slide Time: 06:50)



So, on the basis of this sample information, can we conclude that the sales of coffee have increased? We have where it has once and here we can consider 5% significance level. That means my significance level α is 0.05. So, now here null hypothesis is maintaining the status quo. Alternative hypothesis is what we want to test. Null hypothesis is with maintaining the status quo that means μ = 50. That will make H 0 alternate hypothesis what we want to test we want to test whether the coffee sales has increased more than 500 earlier mean was 500.

(Refer Slide Time: 07:27)



So, specification of first step specification of hypothesis and acceptable level of α, so, this is my first step, μ = 500 that is H 0 and this is H 1, sorry this is not H 0 this is H 1 alternate hypothesis H 1. Alternate hypothesis μ is greater than 500. The alternate hypothesis is that sales have increased means I want alternate hypothesis something what I want to test. So, and significant level of acceptable level of α is 5%. Why it is called an acceptable level of α? Significance level α what does that mean actually?

In a hypothesis test α is the probability which is the maximum probability which is acceptable maximum probability of rejecting a true null hypothesis maximum accepting probability of rejecting a true null hypothesis. That means, α 1 we call it a significance level that is like we can this much we can accept this much significance level this much type 1 error we can accept.

The type 1 definitely in all hypothesis as I mentioned before yesterday only in hypothesis testing is making an error is an inescapable part there may not be error, but then we cannot say that there will never be error, it is an inescapable part of null hypothesis testing. So, when it is an inescapable part let us accept the fact. So, what we accept and we accept give specifying α. we will accept the maximum acceptable level of rejecting a true null hypothesis this is α. I will maximum I will accept 5% that does that is the meaning of significance level.

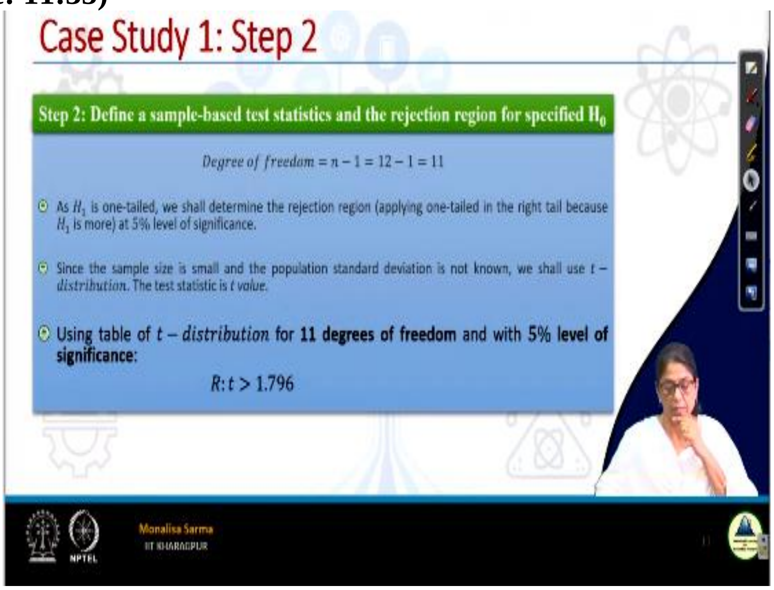
So, here it is 5% now, again one more thing to say here, when as α is 5%. Now, I yesterday we had discussed what is 2 tailed tests and what is single tailed test. So, if it is a 2 tailed test, then rejection region will be in the both side rejection reason you say or critical reason you say it will be in the both side. If it is a 2 tailed test. So, when I specify 5% that means this 2 is put together is 5% that means, what is this? This is 0.025 this is 0.025.

When it is 2 tail and when it is single tail it will be one way either this way or this way, whatever way it is depending now, it is I want to prove that μ greater than 500. So, definitely my rejection region will be this side. So, it is 5% so, 5% is total this area will be 5%, that is 0.05. Because this side we are not considering. So, why we need α based on this α we will find out the corresponding Z value. What is the Z value cost?

Because Z value give me this value. But is this is the axis is the y axis what is the value corresponding to this point that is the Z value that is the starting of my critical region try to remember all this, this is just starting on my critical region or rejection region. So, this I am finding it in terms of Z distribution. So, this is 0 this side it will be minus this side it will be plus. So, once I found out my critical region Z let me tell this Z 1, let me tell this Z 2, this are my 2 critical region.

So, this is when I convert it to Z distribution, but however, in a given problem it is not in Z distribution it will I will be getting a sample and say any random variable say x. So, corresponding to this Z, what is my rejection region that also I can again find out there is nothing corresponding to Z 1 value what is the x bar value? Correspondingly what is Z 2 value what is the x bar value? That is my starting of the critical region we will see in some examples. So, now, given the; acceptance of 5% level of significance.

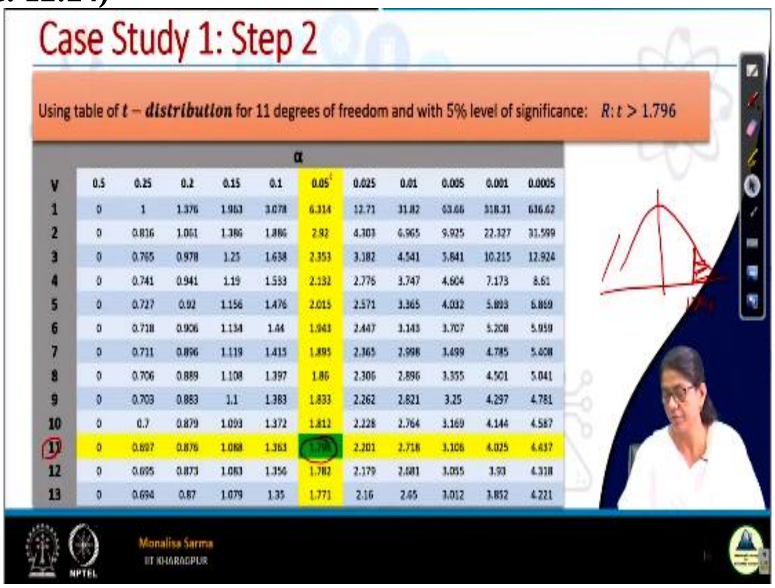
(Refer Slide Time: 11:35)



So, now, step 2 is, define a sample based test statistics and the rejection region. So, what will be the test statistics here see here what it is given? The mean of the population is given but the standard division where the population is not known, when the standard deviation of the population is not known, and we have to infer about the population mean then definitely we will be using on my test statistics will be t value that means, we will be using t distribution.

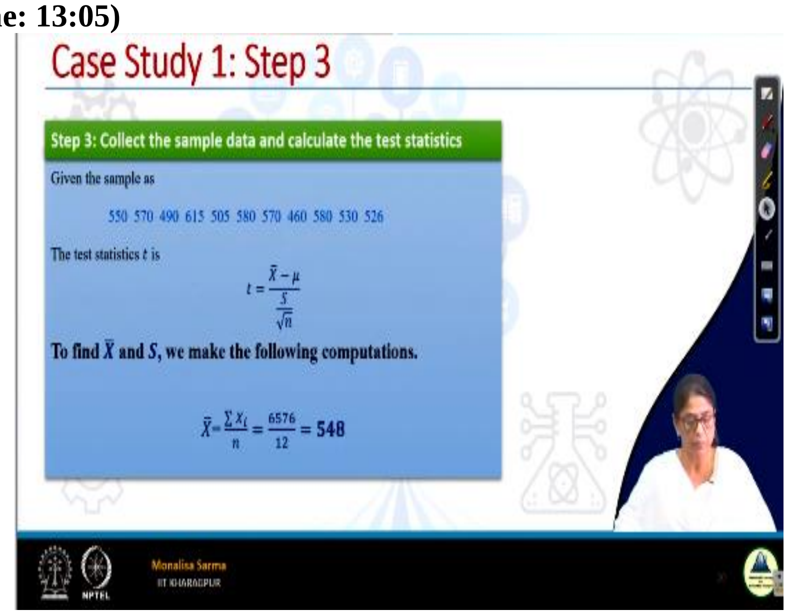
So, with the degrees of in t distribution I have 1 parameter that is the degrees of freedom. Degrees of freedom will sit and minus 1 sample size minus 1. So, it is 11 degrees of freedom and at 5% level of significance I need to find out what is the t value.

(Refer Slide Time: 12:24)



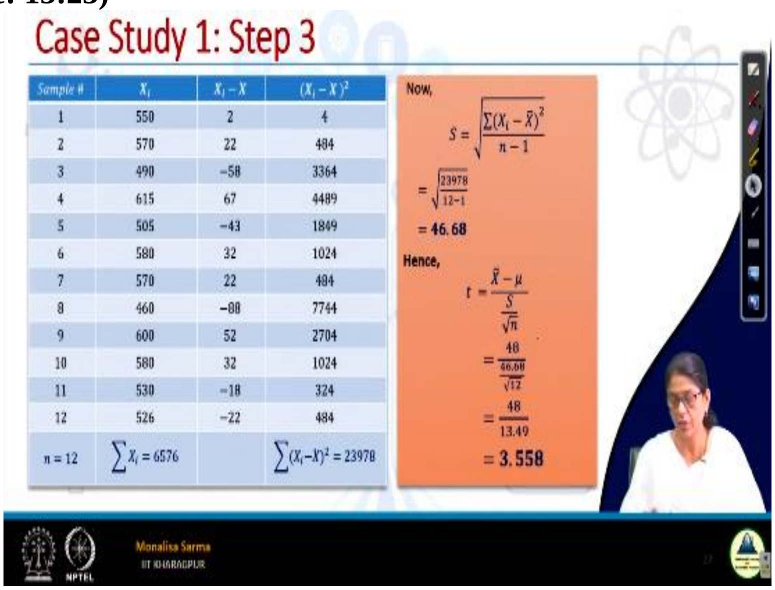
So, this if you see the table t table ere, see the t, this is 11 and this is corresponding to because 5% 0.05 it is 1 tail. So, I will take total 5% in one side only. So, it is 11 degrees of freedom 11 point by 5% significance level. So, this is my rejection region 1.796 that means, t distribution is 1.796 so, my rejection rates are 1.796. This is my rejection region t distribution fatter tails, so, I have drawn it this way. So, my rejection region is somewhere here. If my test statistics value falls in this region, then I reject the hypothesis.

(Refer Slide Time: 13:05)



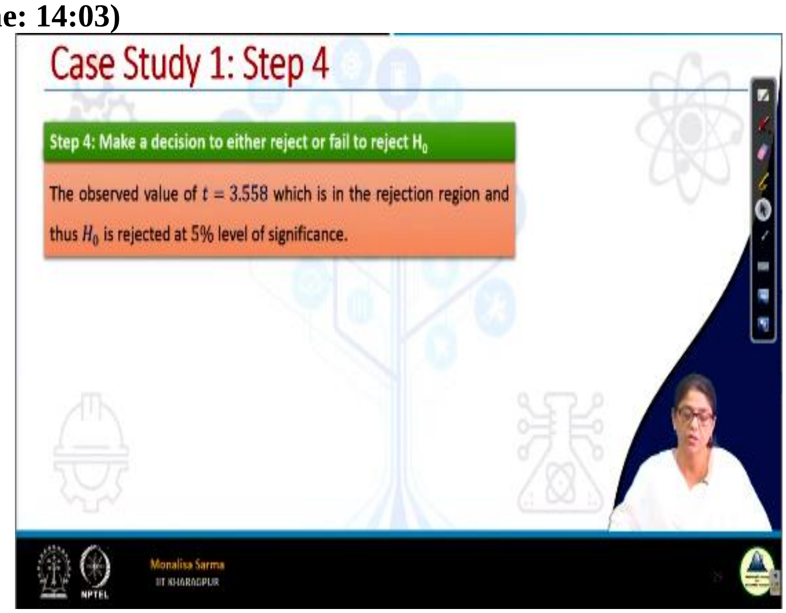
So, given the sample is given from the sample I need to find out this value x bar value I need to find out the standard deviation that is S. So, given from a set of data how to calculate x bar? That is the mean how to calculate the standard deviation that you know it I am just skipping it.

(Refer Slide Time: 13:23)



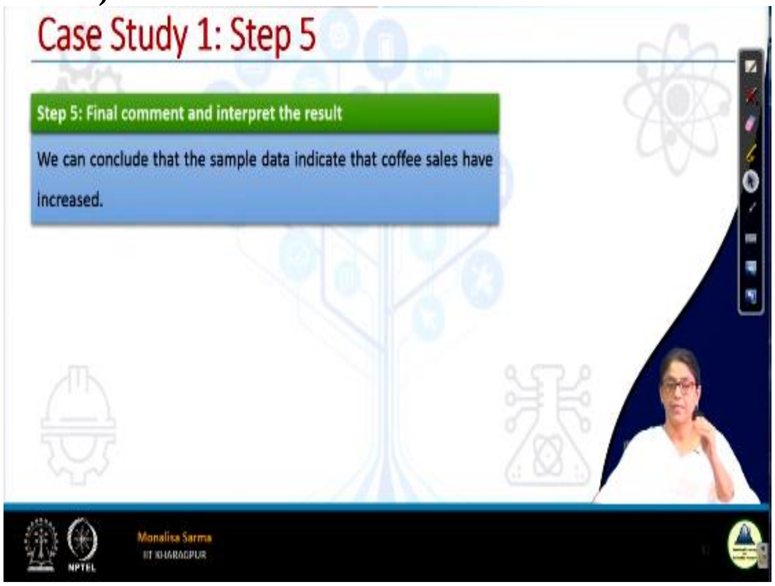
So, I got the mean value I got standard deviation value then I put it into a t value and then I got my t is 3.558. Where is my what to say my what was my rejection region starting up rejection region my value was remember 1.796 my rejection region value was 1.7 this value is 1.796 and this value t value I got something 3 point something so, definitely it is in this area is not it? 3 point something would be maybe somewhere maybe. So, that means it lies in the rejection region.

(Refer Slide Time: 14:03)



So, the observed value of t = 3.5, which is in the rejection region and H 0 that is null hypothesis is rejected at less than 5% level of significance.

(Refer Slide Time: 14:16)



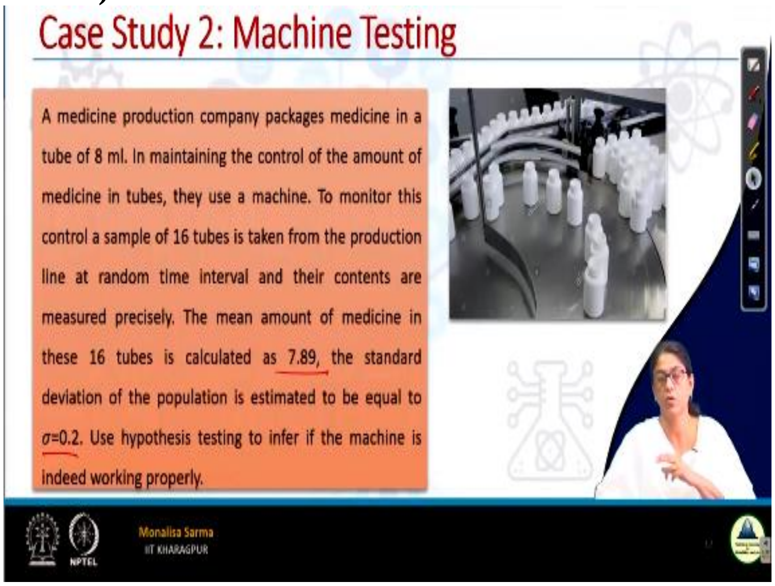
Now, the final comment based on the application we can give the form it is come in simple language, we can conclude that the sample data indicate that a cup of coffee sales have increased because we are rejecting the null hypothesis means we are accepting the alternate hypothesis. What is the alternate hypothesis μ is greater than 500. So, that means we are accepting the fact that coffee sales have increased.

Now here I also want to bring to your notice is that some we call a result statistically significant when we call a result as a statistically significant. Please, this is very important. Please try to remember when a statistically significant is basically when we reject a null hypothesis how we reject the null hypothesis because the value that we get from the sample it is very much it is significantly different from the statistically different from the null hypothesis value.

That is why only we are rejecting the null hypothesis is not it? That why we are rejecting the null hypothesis? Because in a null hypothesis we have assumed a certain value of the mean assuming the value of the mean from the sample whatever data we got, if that data correlates with this mean, then this is that means the mean is correct, is not it? But when we reject when we get such a value, that it falls probability of which is very, very less that we specified whatever probability is acceptable to us and accordingly we specify the rejection region.

So, if the value sample test statistics values fall in the rejection region, that means, what sample value is statistically significant compared to the null hypothesis value. So, such results are called statistically significant, that means, when we say a result is statistically significant that means, definitely we are rejecting the null hypothesis. So, statistically significant is important term please try to understand this.

(Refer Slide Time: 16:14)



So, we will see one more example. So, say this example thing we have written I have explained it while explaining the hypothesis testing, how to form the; what to say hypothesis. So, here however, there are some other information are given compared to the previous one. Previously, and I am just quickly going because we have already discussed this medicine production company packages medicine in a tube of 8 ml that is my mean is 8 ml in maintaining the control of the amount of medicine in tubes, they use a machine.

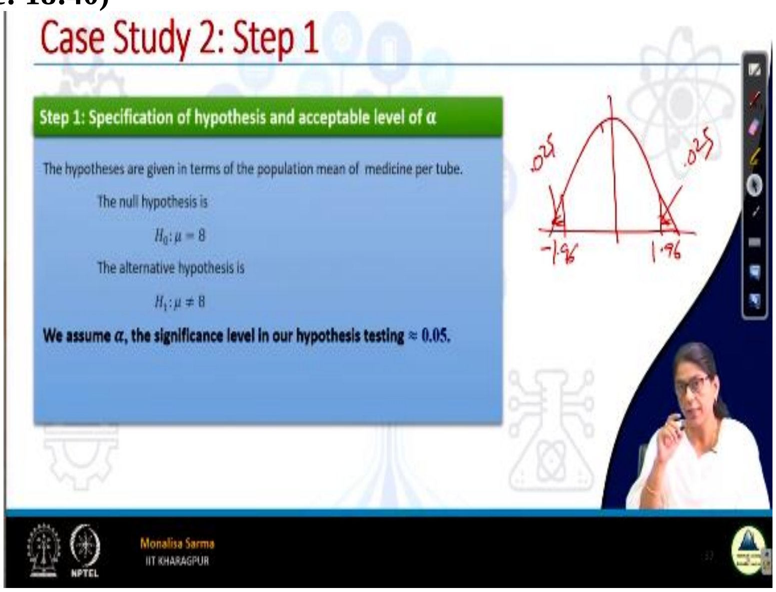
To monitor this control a sample of 16 tubes is taken from the production line at random time interval and the contents are measured precisely. We have already done that that was we have already formulated the hypothesis. What is now what is given the mean amount of medicine of the 16 tubes is calculated as we have taken a sample from the sample the mean amount sample of 16, mean amount is calculated 7.89.

And the standard deviation a population not the sample, but the population is estimated to be 0.02 use hypothesis testing to infer it a machine is indeed working properly. So, see here there are 2 things so, in some problem the significance level may not be mentioned. So, there are 2 ways one is it a significance level is not mentioned then we will report a P well what is p value I will come later.

Now, another test for another way, if the significance level is not mentioned, in general 5% significance level is considered. If it is not mentioned, or if it is if you are asked to get a P value, then it is a different thing then you do not have to consider a significant level. But if you are not asked to give the p value, what is p value we will later again. So, then you will have to assume is 5% significance level, here the significance level is not given.

So, let us take the significance level of 5%. Now here how to formulate what to say hypothesis first is definitely $\mu = 8$ ml status quo that means we know that mean is 8 ml, we are hypothesizing that it is 8 ml. And what we want to test that it is not equal 8 ml not neither less or greater, we are not interested in less or greater, we have what we want to test that it is not equal to 8 ml. So, that is my alternate hypothesis, alternate hypothesis is $\mu \neq 8$.

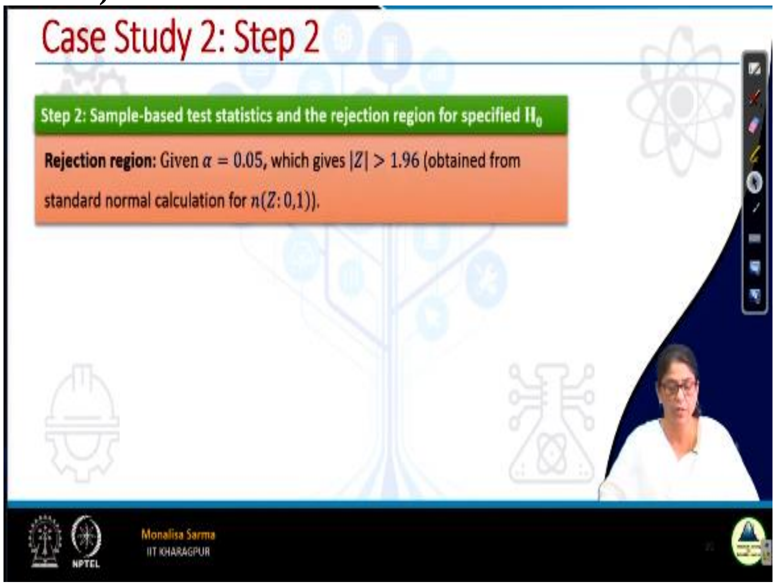
(Refer Slide Time: 18:40)



So, this is my that is how I found a hypothesis $\mu = 8$ H 1 $\mu \neq 8$, then some significant level 5%. We have assumed, so, since it is 5%, it is 2 tailed means one side it will be 0.025, aside will 0.025 and as I said it will be 0.025. So, my when I talk of my rejection region. So, 5% means say this is 2 things 2 sides this and this both put together to 5 person so that means this will be 0.025, this will be 0.025 and then I will have from the table I will be able to find out the Z value corresponding to this 0.025.

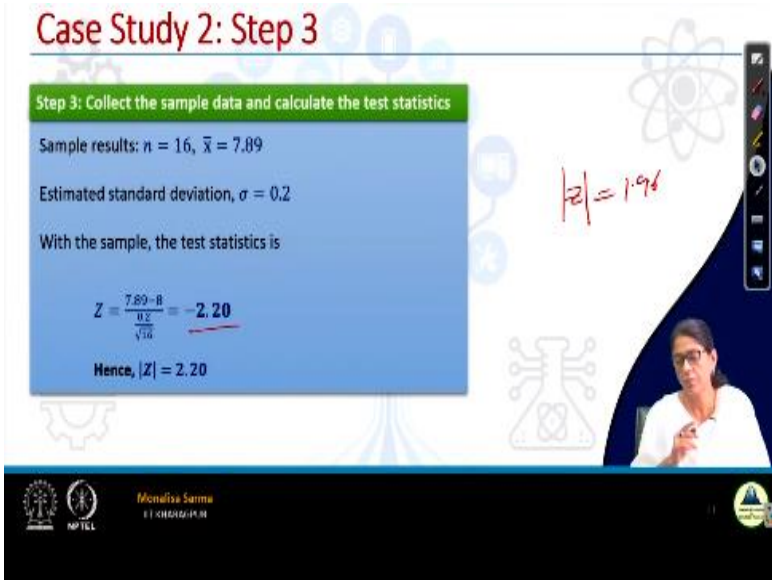
So, Z value corresponding to this is I can remember actually it is minus 1.96. And this is plus 1.96. So, for the thing to be added to and null hypothesis to be accepted my values should be within minus 1.96 to plus 1.96. You can see the table.

(Refer Slide Time: 19:49)



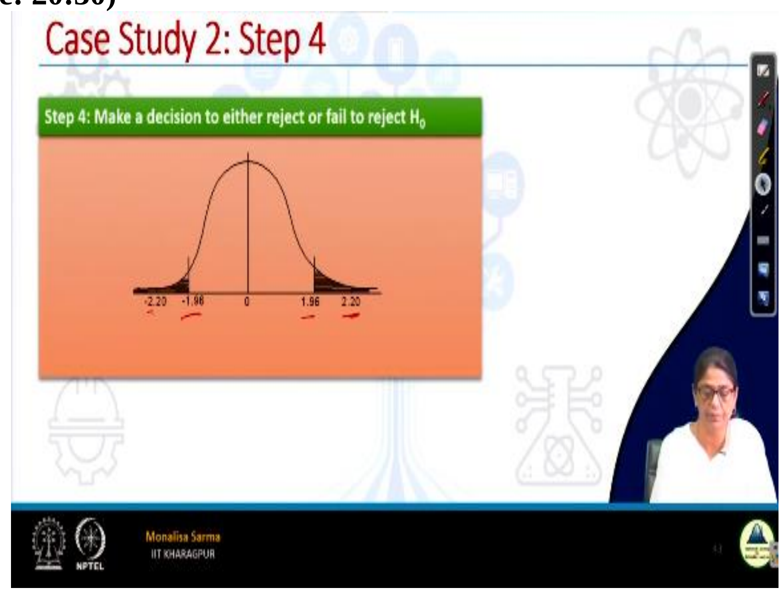
So, which sample based test statistics we will be using here. See here what it is given population mean is given population standard deviation is given, we have to infer about the population mean then definitely we will be using their distribution.

(Refer Slide Time: 20:03)



And collect from the data we will come to find out the Z value. So, that will what we got minus 2.20. So, minus 2.20 means, it is greater than the rejection region is not it? What is my rejection region? Rejection region is Z = 1.96 from minus 1.96 to plus 9.16 and my calculated Z value is greater than that and so, it is in the rejection region.

(Refer Slide Time: 20:30)

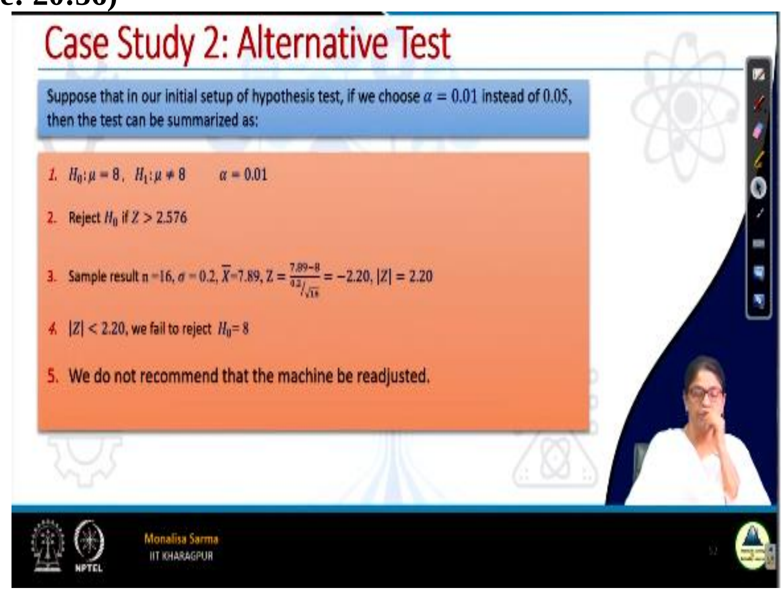


So, you see this figure this is my rejection region and I got my value here in this region it may be positive or negative whatever it is I got in this region so, we reject H 0.

(Refer Slide Time: 20:45)

So, we conclude that μ is not equal to 8 and recommend that a machine be adjusted that means, there is some problem in the machine and we recommend that that machine is adjusted.

(Refer Slide Time: 20:56)



Now, why I have taken this example actually, to prove a point, I basically wanted to come gradually to p values that you have to this example again though we have discussed this example in my first lecture, see here specifying significance level for the same problem, same problem everything data remains everything same. Now, let me take a significance level of $\alpha = 0.01$ instead of 0.05. So, if I take a significance level of $\alpha = 0.01$.

So, what happens, my reject H 0 my Z value will be 2.576, corresponding to 0.01. So, what it will be my both sides it will be 0.005. This side is 0.005, this side it will be 0.005. If you see the

Z table, you will get corresponding value to that is 2.576. So, from the test statistics what value I got from the test statistics I got value 2.2. So, now, that means, it is not rejected failed I failed to reject the status the data statistics the sample data failed to reject the null hypothesis I got value 2.0 and this is 2.57 says it is less than that.

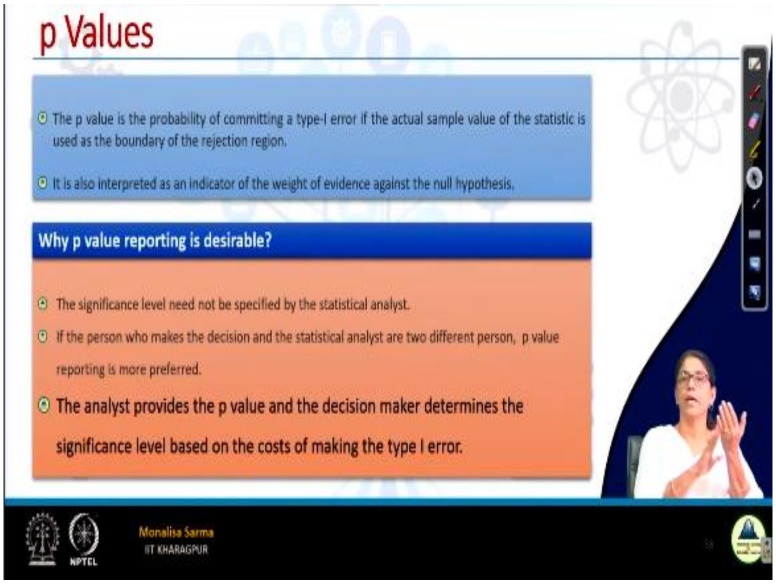
So, similarly the same problem suppose, as I previously when I was discussing sampling distribution I told them if you take different sample data statistics from the sample will different is a very slight chance that statistics of many 2 samples are same. So, I have taken a sample from the sample I got X bar is 7.89. In this example, I have that is what I have assumed this from X bar I got 7.89 suppose I took a different sample and from there suppose in the X bar I got X bar = 7.91.

Let me happen I took a sample 1 sample from which I got 1.89. Another sample I got from which I got X bar = 7.91. Now, for the 7.91 only, if I take 5% significance level only no need of taking 1%. If I take 5% significance level only still what will happen my I will fail to reject H 0 null hypothesis. See such a slight change from 7.89 and I got 7.91, for 7.89 I had to reject the hypothesis for 7.9. I could not reject the hypothesis.

If I change the significance level 5% significance level I have rejected the hypothesis for 1% significance level we fail to reject the hypothesis then what about if I take 2% significance level if I take 3% significance level. So, that is why usually this sort of problem is not carried out manually the sort of problem is the problem is done automatically in a computer. So, in the computer when we feed the algorithm basically for this particular this is the critical reason.

Then from what sample whatever data it gets, and what the computer checks this data is less than this then do not reject if it is greater than this reject it is just reject or no reject there is no other option here. But see for slight changes of data we get different results. So, that is why there is one more approach instead of directly telling it is reject or not reject one more approach of mentioning the statistical results the same significance test results or hypothesis test results is by reporting the p value.

(Refer Slide Time: 24:31)



So, instead of directly telling that to reject the hypothesis or do not reject the hypothesis, we do not what we do is that we report the p value when first of all when this p value is used mainly suppose the person who takes the decision is someone else and the person who does the statistical test is someone else. So, suppose I am doing the statistical test. So, I got all the data and I will just report the data to the decision maker.

The decision maker will see the data and will try to find out whether, we need to based on this data whether I need to reject the hypothesis or I need to accept it that is totally based on the decision maker because he has total knowledge about the system, he will be able to take proper decision whether to reject or not to reject based on the, how the data has come. So, as a statistician, I am just doing all the data, I am just collecting the data calculating all the values and giving you so now in that case.

So, when I am not directly specifying, reject or not reject what I am doing is I am specifying the p value. So, what is p value? p value is that basically if u see the definition p value is the probability of committing a type 1 error is the actual sample value of the statistic is used as the boundary of the rejection region, it is the definition of p value. Let us not go into those definitions just simple p value basically what we from the statistics whatever value you got like here this is the from the statistics you got the value 2.20.

From the sample this is a sample statistic these statistics we got from the sample. So, what we will do is that from the table, we will find out what is the probability of this occurrence that is nothing but the p value of this set of data, what is the probability of this 2.20 that is nothing but the p value. So, in fact, now, here this is the p value is the probability of committing a type 1 error in the actual sample value of a statistic is used as a boundary region or the rejection region.

Why is it telling that why is defining p value in such a way when suppose, in significance level is not mentioned. So, whatever p value I get based on that if I reject the null hypothesis that means, that for the p value corresponding Z value corresponding for that p value that is my starting region that is my rejection region starting of the rejection region. So, you do not have to remember this definition just if you understand what is p value, p value is the probability of getting that t statistic value.

So, it is also interpret it as an indicator of weight of evidence against the null hypothesis why it is indicator of weight of evidence? Because we have assumed something null hypothesis is what we have assumed what we have hypothesized and from the sample whatever we got now, as you mean this what we got what is the probability of weight that is the way it is, is not it? It is the indicator of the weight of evidence against the null hypothesis.

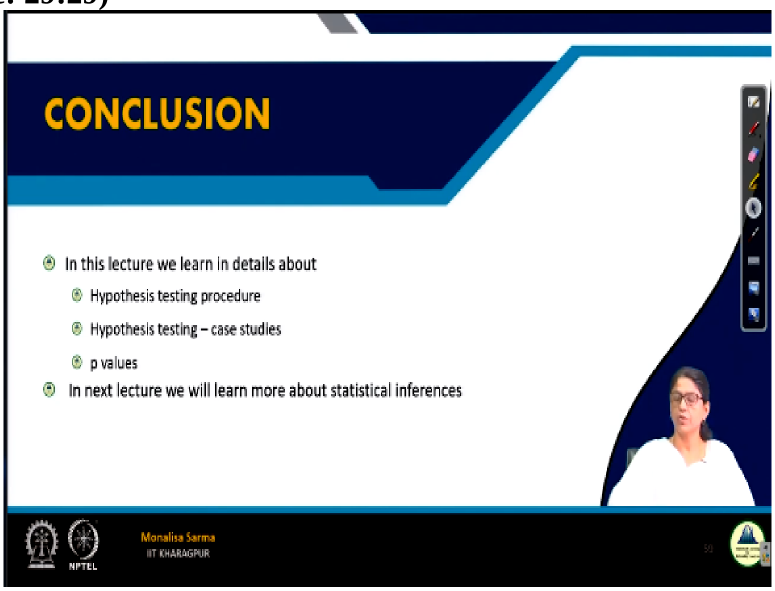
So, why p value reporting is desirable? The significance level need not be specified by the statistical analyst. It the person who makes the decision and the statistical analysts are 2 different person p value reporting is more preferred. The analysts provide the p value and the decision maker determines the significance level based on the cost of making the type 1 error. If the cost type 1 error already I have specified I have mentioned some what is this type significance level?

That is the type 1 error significance levels is the maximum acceptable significance level or maximum probability of type 1 error that will be accepted maximum acceptable type 1 error acceptable means what? Acceptable probability of rejecting a true null hypothesis. So, when so, it is the analyst a decision maker will decide how much acceptable well how much acceptable error I can take 1% error 2% error, it totally depends upon the application if the application is a very critical application, then my error what do I mean by critical applications?

How do I select a null hypothesis? How do I select a significance level? I will be discussing in my next lecture. So, basically, so, there are different factors based on the application. So, it is so, the decision maker will based on these different factors will take the decision whether this is the p value of the statistical analysis given me this is the p value. Now, based on I know what sort of what is what application is this?

What is the cause of the cost of an error? So, based on that I can decide whether I will accept the null hypothesis or I will not. Whether I will reject the null hypothesis or I will not reject a null hypothesis.

(Refer Slide Time: 29:29)



So, in this lecture we learn in detail about the hypothesis testing procedure. We have seen few case studies we have also understood the concept of p values. And in the next lecture, we will learn some more about statistical inferences.

(Refer Slide Time: 29:42)

So, these are the references and thank you guys.