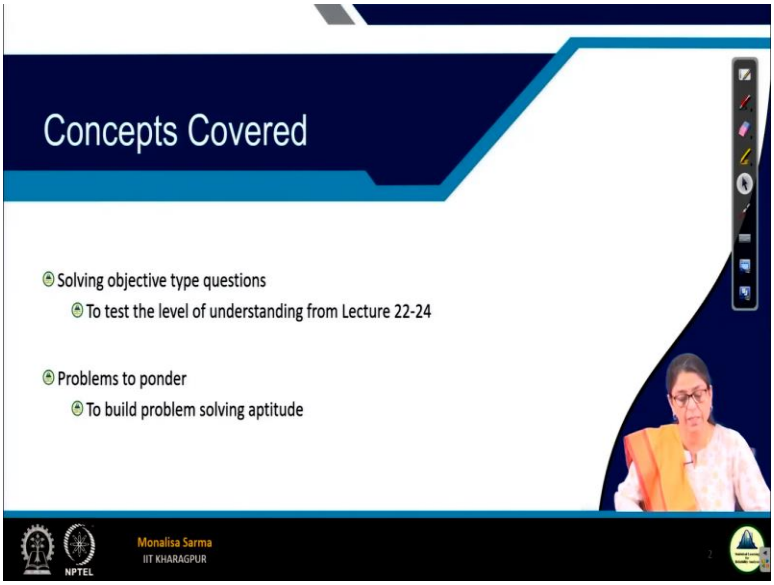
Statistical Learning for Reliability Analysis
Prof. Monalisa Sarma
Subir Chowdhury School of Quality and Reliability
Indian Institute of Technology - Kharagpur

Lecture - 25
Tutorial on Statistical Inference

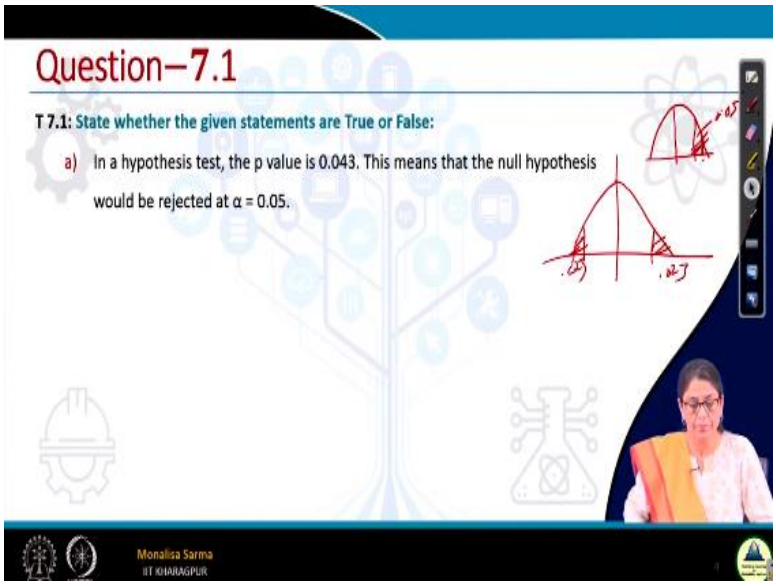
Hello guys, so in continuation of our discussion on statistical inference, today we will be doing a tutorial on whatever we have learned. It is not that statistical inference is complete, yet, there are lots of other things which we will have to learn. But, before going to those topics, till now, what we have covered, let us do a quick tutorial on that.

(Refer Slide Time: 00:47)



So, in this, we will be covering we will test the level of understanding from lecture 22 to 24 that is the last 3 lectures, and then we will also solve few problems. So, first we will see the objective type questions.

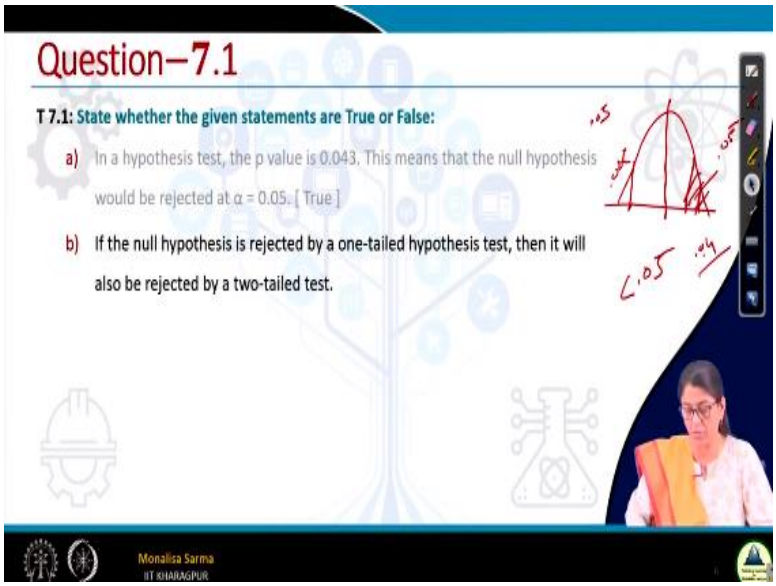
(Refer Slide Time: 01:01)



So, it is again a true false type of question let us given in a hypothesis test the p value is 0.043 that means that a null hypothesis would be rejected at $\alpha = 0.05$. So, like, what will be the answer to this question, so, if I draw the figure so, if this is my what is say sampling distribution, So, α is 0.05 means assuming both tail α is 0.05 means it will be both side it will be how much the loop 0.025, if it is both side. So, and this means that a null hypothesis will be rejected at α equal to 0.05.

This is since it is not specified whether it is 2 tails or single tail it is this written α is 0.05. So, you can just consider it as a single tail if it is a single tailed then so, it is 0.05 will be this portion is 0.05 then in that case 0.043 will fall in this region. So, in that case the null hypothesis will be rejected. However, if it would have been a 2 tailed then it would not have been rejected.

(Refer Slide Time: 02:29)

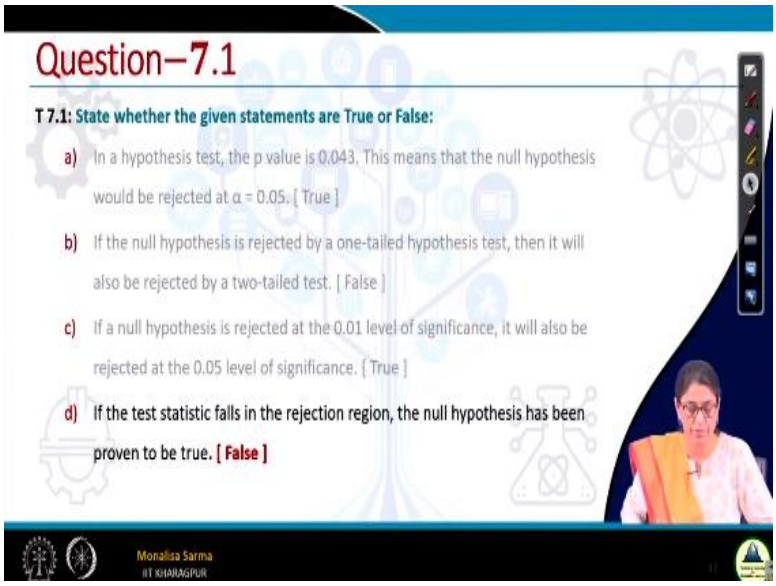


So, this is true. So, if the null hypothesis is rejected by a 1 tailed hypothesis then it will also be rejected by 2 tailed test. What it is given if the null hypothesis is rejected by a 1 tailed hypothesis test, if it is rejected by 1 tail means like just the last example what we have seen if it is rejected by 1 tail, when suppose we have taken say 0.05 is my α. So, if it is rejected by 1 tail 1 tail means my critical region is this portion is 0.05. So, if it is rejected that means it is less than 0.05. That is why it is rejected then it will also be rejected by a 2 tailed test.

So, that we cannot say it may be rejected it may not be rejected because, if it is 2 tail then what happens? 2 tail means this will be 0.025, this will be 0.025. So, that means suppose my test statistics value suppose I got 0.04. So, if it is a 1 tailed hypothesis test then 0.04 definitely it would have been rejected, is not it? But then in a 2 tailed test 0.04 will not be rejected because for in a 2 tailed test the rejection region starts only from 0.025.

So, if the null hypothesis is rejected by a 1 tailed hypothesis test then it will also be rejected by 2 tailed test no it is false, it may be rejected it may not be rejected.

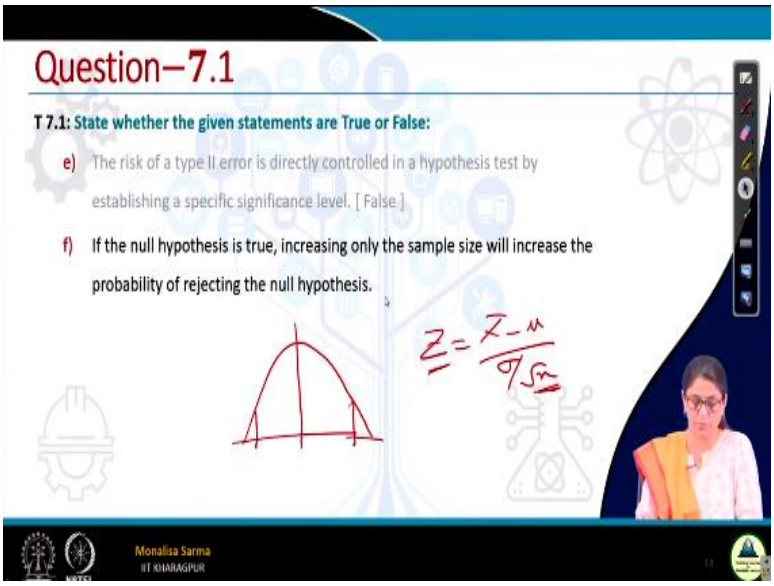
(Refer Slide Time: 04:06)



So, it will always be rejected by 2 tail test that is definitely false. So, if a null hypothesis is rejected at a 0.01 level of significance, it will also be rejected at the 0.05 level of significance. So, what it is given if it is rejected at 0.01 level of significance 0.01 it is rejected at 0.01 means my value is less than 0.01, t statistics whatever t statistics whatever it is the sample statistics value, I got is less than 0.01. That is why it got rejected.

Now, it will also be rejected at 0.05 something which is less than 0.01 will also be less than 0.05. So, if it is rejected at 0.01, level of significance it will also be rejected at the 0.05 level of significance that is true. So, if the test statistics falls in a rejection region, the null hypothesis has been proven to be true. So, if it is if the test statistic is false in a rejection region that what happens we reject the null hypothesis. So, null hypothesis program, this program proven to be true is false.

(Refer Slide Time: 05:20)



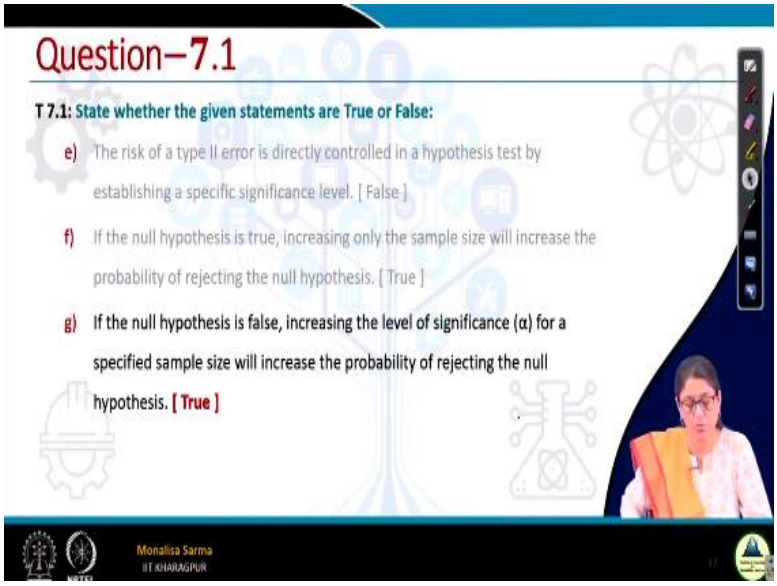
Next question the risk of a type 2 error is directly control in hypothesis test by establishing a specific significance level. Since specific significance level means so, means α by having a specific significance level α, the risks of type 2 error is not directly controlled it is indirectly controlled, because when we specify α we control we directly control the type 1 error. But that does not mean by controlling α of course, when we increase α my β decreases.

When we what to say decrease α my β increases that is of course true, but then if we control α it directly controls the type 1 error because type 2 error because we have seen when we tried to find out the operating characteristics curve, we have noticed that it is not α does not directly dictate the risk of the type 2 error risk of the type 2 error is mostly it depends on how much it is away from the true value of the population parameter to the true value of the parameter how much it is away from the hypothesis value. So, it is false.

If the null hypothesis is true increasing only the sample size will increase the probability of rejecting the null hypothesis. So, how we find out the z value if you are interested in my statistic is z, z is x bar - μ σ / √ n. So, my status the rejection region these are very much conceptual question if you can understand a concept you will be able to answer all the questions like so, rejection reason is always this means, when my z value is rejection region my z value is always more.

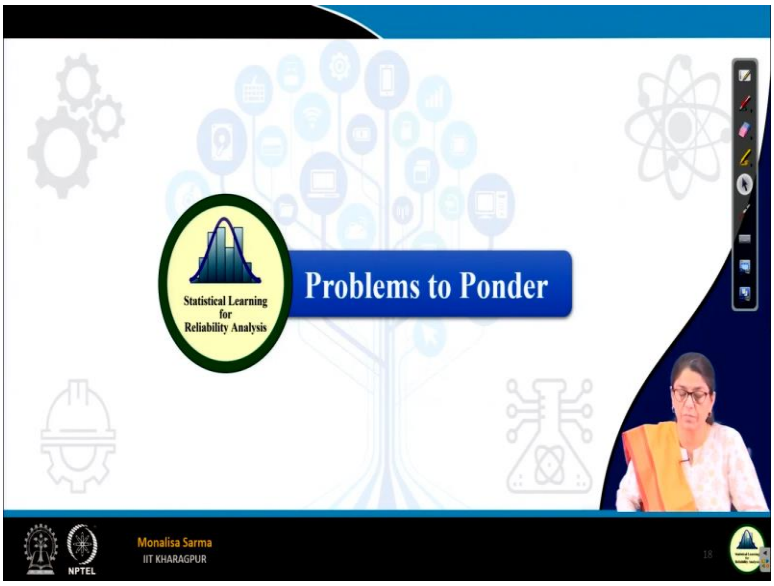
So, when my n is sample size n is increased what happened, my z value is increased. Value of z increase means same size going into the critical region becomes more and more. So, the null hypothesis is true increasing only the sample size will increase the probability of rejecting the null hypothesis is true.

(Refer Slide Time: 07:34)



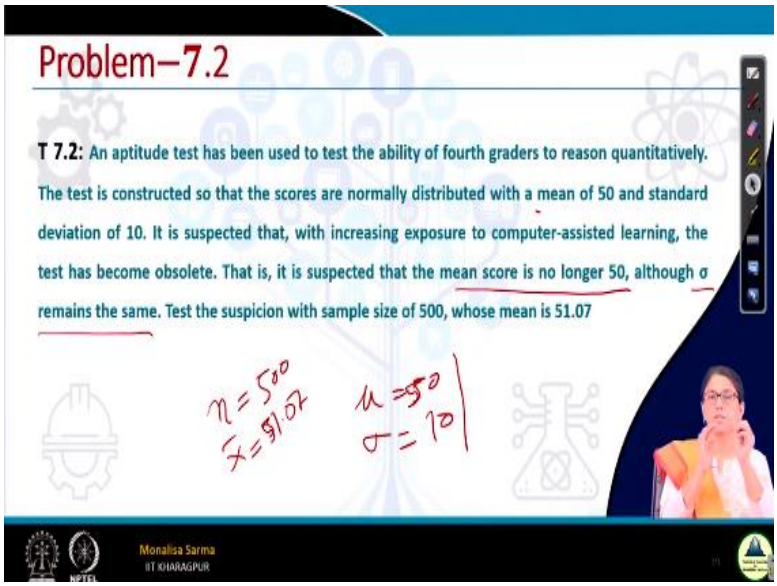
If the null hypothesis is false, increasing the level of significance for a specified sample size will increase the probability of rejecting the null hypothesis. If the null hypothesis is false given that the null hypothesis is false, but we do is that increasing the level of significance means we are increasing α, we are going to increasing α that my critical results become more. So, what happens that will increase the probability of rejecting the null hypothesis, this is true.

(Refer Slide Time: 08:03)



So, this objective question what I suggest is that immediately do not see the answer just from the question you first try to once I have already discussed now, next when you will be seeing this video again definitely it is not once you will see the video in maybe that for your exam time again, you will revise it right. So, first try to you answer yourself because, these are very much conceptual questions. If the concepts are cleared and you will be able to solve all the problems and your quiz answers for this objective will also be correct.

(Refer Slide Time: 08:34)



Next the problem we try to solve so, an aptitude test has been used to test the ability of 4th graders to reason quantitatively. The test is constructed so, that the scores are normally distributed with a mean of 50 and a standard deviation of 10. The test is constructed, so that the

test is constructed in such a way so that the test scores are normally distributed and what we got, we got a mean equal to 50 and standard deviation equal to 10 standard division is σ not σ^2.

Test is conducted in such a way so, that we get this value that means, this is what we want 50 and 10 it is suspected that with increasing exposure to computer assisted learning, nowadays there is a computer assisted learning the test has become obsolete. Whatever the tests were hardly it was there it is the suspected that the test has become obsolete. That means it has become obsolete means that means it is suspected that the mean score is no longer 50.

Although the standard division remains the same, now, a new computer assisted learning has been introduced with this the tests what was the earlier the same test it says they are suspecting that the test is no longer viable it is suspect that the mean score is no longer 50. So, it is either less than 50 or greater than 50 it is not telling anything whether you need to test for less than or it need to test for greater it is this telling that it is no longer 50.

So, although the standard deviation remains the same, it is this suspicion with a sample size of 100 whose mean is. So, we have to test the suspicion with samples we have taken a sample size of 50 and mean of the sample that is 51.07 that is the mean. Now, first thing this will have to frame the hypothesis. So, how will frame the hypothesis always I told you will discuss your hypothesis while framing the hypothesis, we have to frame the hypothesis in such a way and where type 1 error is more significant.

Because we are in a hypothesis testing procedure we specify the significance level that is α what is α? A is the type 1 error. So, while framing the hypothesis always we should frame in such a way that the type 1 error is more significant. Because we are giving more importance to type 1 error, that is why we are specifying α if we would have given more importance to be a type 2 error with a specified β.

But in hypothesis testing, we are giving more importance to A more importance to type 1 error that is why we are specifying the error probability of type 1 that is α. So, now while framing this hypothesis, we should keep this in mind. First of all, there are many things while framing the

hypothesis we should keep in this in mind first is that. Second, we understood we will have to frame it in such a way so, that type 1 error becomes more significant. And secondly, 1 important thing the type 1, the null hypothesis should always be with a equality sign.

Because we have seen if the null hypothesis is not equal, as without equal sign, if null hypothesis we have defined some parameter greater than some parameter or less than some parameter, then it is very difficult to find out α. Because in that case, we will have to find out critical region from the α what we do? From the α we find the critical region. So in that case, we will have to find the critical region for different values of the parameter.

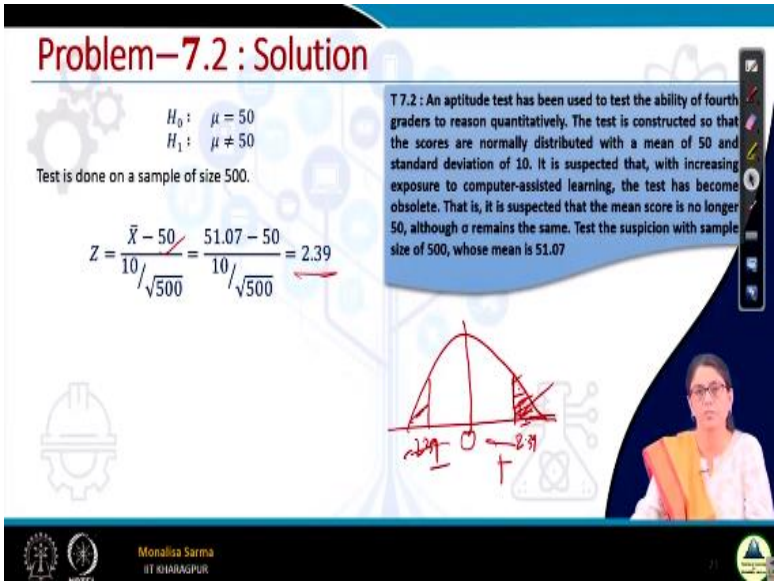
When we specify a hypothesis null hypothesis if you specify the parameter is greater than a particular value, so greater means it can take any value greater than that particular value, say x greater than x. So, we will have to take any value greater than x. So, for all says x, we will have to find out α. So, that is a very tedious process we have seen while doing it for β. So, to keep it simple, always null hypothesis is specified with an equality sign so, the second thing.

Third thing now, while specifying the significance level, it should always be specified in such a way based on the cost of the type 1 error. If the cost of the type 1 error is huge, and the cost of type 1 error is huge then we will give a very less α means we do not want to reject a true null hypothesis in rejecting a true null hypothesis should be very, very remote that means α should be very, very less the cost of the rejecting a true null hypothesis is very high, my value of α will be very, very low.

So, keeping all these things in mind, we will have to frame the hypothesis. Now, here you can see the significance level is not mentioned, but we can understand this is a first we need to understand what you will form the null hypothesis? So, here you see the null hypothesis we will be maintaining the status quo maintaining means if we have to say something, we need to either invest lots of lots of lots and lots of money, resource, time, many such things many such parameters.

So, null hypothesis always we try to maintain the status quo. So, here the status quo is that the test is no longer obsolete, that means when the test is no longer and obsolete, when you will find that μ is what is expected what is the hypothesis value μ equal to 50 that specifying that μ we are getting with the computer assisted learning also we are getting the same μ that means test is no longer what to say obsolete. So, that is my null hypothesis. So, μ = 50 is my null hypothesis.

(Refer Slide Time: 14:21)



And now next what I need to check that is the alternate hypothesis what I need to check? I need to check whether μ is not equal to 50. So, I have to check μ is not equal to 50. So, for that, I will take a sample the size of sample size is given 500 and mean of the sample is given 51.07 and I know standard deviation of the population is given 10. So, the standard deviation of the population would not have been given then I would have gone for t distribution.

Now I will be taking the standard deviation of the sample. Now, since the standard deviation of the population is given, I will consider z distribution. Now, secondly, there are 2 things first significance level is not mentioned as I told you when a significance level is not mentioned and while discussing this at all, if it is not mentioned you can do 2 things. First is you do not have to write accept reject just specify the p value, then the decision maker will decide whether to accept the null hypothesis or to reject the null hypothesis.

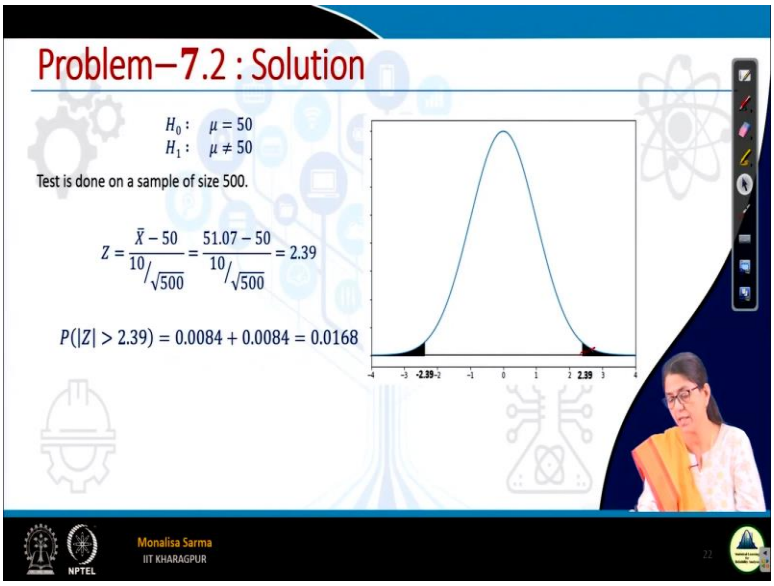
Just specify the p value that is one option. Second option is that you select your own significance level. So, that significance level usually it is considered 0.05 as a standard significance level is considered, but for this example, where if the true null hypothesis is rejected unnecessarily the whole test has to be constructed it is there because constructing a test is not an easy task while constructing a test lots of men brain has to be means you really have to work hard on it.

It may not be lots of money, but then lots and lots of effort. So, here we do not want to undo something which is working fine. So, here we may consider a very less significance level. So, now, since we will consider z distribution, so, what will be the value of the z, z we know x bar - μ σ / $\sqrt{}$ n. So, this is x bar – μ, μ = 50 so, computing we got value 2.39. So, now, we can find out the p value of this since its significant level is not mentioned. So, now see while finding out the p value, there are 2 things again you have to remember.

If it is a single tail what z value you get? For z value you get the corresponding probability of that is the p value if it is a single tail, the corresponding probability of the z value is the p value remember this z table we have seen it is not it? In z table what we get? In z table we get the z probability corresponding to different z values. Now, here z value is given. So, probability corresponding to z value maybe it is this since it is what to say standard normal distribution definitely it is zero and this side to be minus this side it will be plus this side it will be plus.

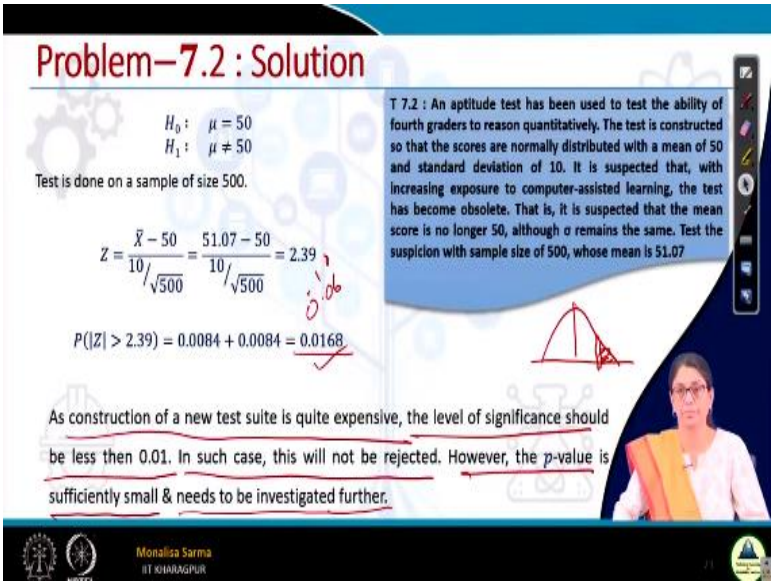
This said minus so, z we got -2.39 means some somewhere here maybe 2.39. So, whatever area is corresponding, the probability of 2.39 from the table we can get and said if it is a single tail. If it is a double tail, then what happens, z corresponding area we get on for 2.39 we will have to check for -2.39 as well. If it is a 2 tail single tail only this probability this or this, whatever it is, it is the same thing symmetric z distribution is totally symmetric. If it is double tail, we will have to find out my p value will be corresponding to the probability of 2 point + - 2.39 so, this plus this.

(Refer Slide Time: 18:05)



So, it is see 2.39, -2.39 this is the area this is the area from the z table you will see this area corresponds to 0.0084. So, this is my z value that is my p value. So, now see p value is 0.01, it is a very less actually if you consider it is very less. So, if we since constructing a new test, so, it is very expensive. So, we may consider significance level very less.

(Refer Slide Time: 18:40)



If we consider a significant levels of 0.01 in such case what happened? We have considered a significance level of 0.01 and we get the value 0.0168 it is slightly greater than point 0.168. It means if this is 0.01 my value is slightly more than this. So, since it is more than the significance level, then it is not rejected. If it is more than the significance level means it does not fall in the critical region then we do not reject a null hypothesis.

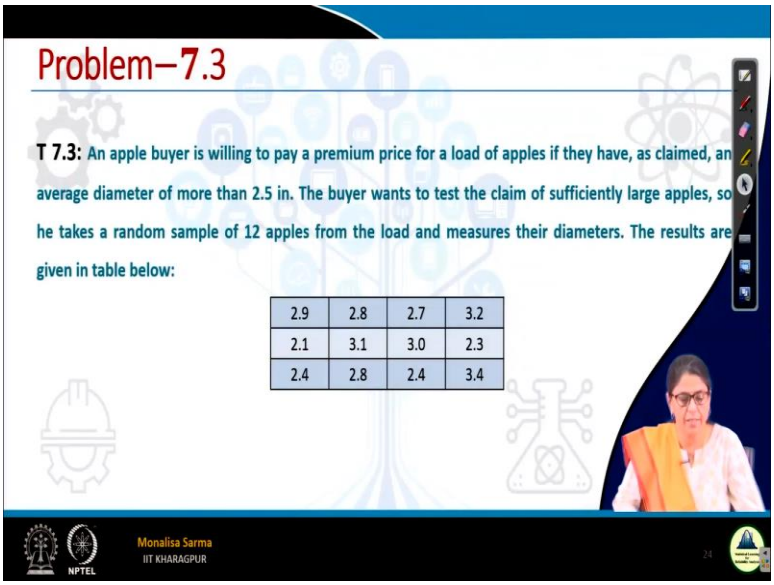
But however, as I told you remember x when we reject the null hypothesis there is only one solution one on what to say one option that means we accept an alternate hypothesis am repeating again when we reject the null hypothesis there is only one option that is we accept the alternate hypothesis. But when we cannot reject the null hypothesis when the null hypothesis is not rejected, there are 2 option either is that we accept the null hypothesis or we just consider that we could not reject the null hypothesis.

We could not reject enable to reject the null hypothesis, these are 2 different things, unable to accept the null hypothesis means we are not complacent we may still carry out the experiment again, again and again and we will try to find out whether actually it is true. Actually, my null hypothesis is true or not. If my null hypothesis is not true, the sample which I considered maybe an out layer sample, then at one time, it will really show that, that null hypothesis has to be rejected.

So, now in this case, I got a very less significance value, P value. Sorry, it is not significance P value I got a very less P value. So, I need to be it is written it needs to be investigated further, that means I will carry out the test again. I am not satisfied, because I got very less value might be that if I can carry out the test might be that it gets rejected or might be that this value, maybe quite, this value becomes more maybe that instead of 0.1, and maybe I got 0.06.

Maybe my sample was not good that is why I got 0.06 means, I am comfortable, I can accept the or maybe 0.10, so then I can accept the hypothesis, I can tell that I am accepting the null hypothesis. So, here you see properly as the construction of the new test suite is quite expensive, the level of significance should be less than 0.01, which I have already mentioned, in such case, this will not be rejected, because this is more than this. However, the P value is sufficiently small and needs to be investigated further.

(Refer Slide Time: 21:40)

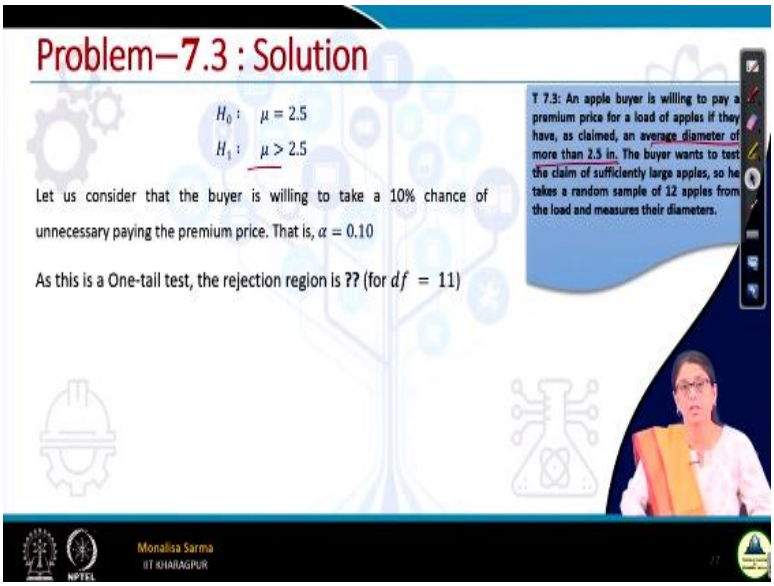


Second question see here, an apple buyer is willing to pay a premium price for a load of apples if they have as claimed an average diameter of more than 2.5. An apple buyer, someone who is buying one to buy apple lots and lots of apple he wants to he will pay a premium price means you will pay more price, if the size of the average diameter of the apple is more than 2.5. He will pay more price, the average diameter means for a big size apple, he will pay more.

The seller is telling you these apples are very big apples, I will give you bigger apples, but he is not satisfied with that. So, wants to test the claim. So, the buyer wants to test the claim of sufficiently large apples. So, he takes a random sample of 12 apples from the load and measures the diameters these are the diameters of the 12 apples which you have taken. So, now, thing is that you have to find out whether the buyer will pay the premium price or not.

Now, in this case, what will be my null hypothesis? My null hypothesis as I told you, one is maintaining status quo of course, and I should make this null hypothesis in such a way so that my type 1 error is more significant. So, here what is the null hypothesis type 1 error is most significant means my null hypothesis will be my apple is less than or equal to 2.5 that means I will tell it is my size of apple is 2.5 inch. M is 2.5 inch.

(Refer Slide Time: 23:17)



And what I want to prove that it is greater than 2.5. Now, send here also the significance level is not mentioned. Let us consider that a buyer is willing to take a 10% chance of unnecessary paying a premium price reeks of type 1 error means even if it is not large, he was fine if it even if it is not large he will take care of error that is error he will take a chance that chances 10%. 10% means α is equal to 0.10 and this question α is not mentioned.

So, based on the case the buyer maybe he can consider α is 0.10 when α is not mentioned always you can take for safety sake you can 0.05 or always specify the P value and then maybe give the decision based on 0.05. If you yourself can consider that what to say that application is very sensitive then consider 0.01 if you considered application is okay we can take care. Here say if apple is a bit small picking you will pay a bit extra or bit more.

So, then here we have to take a very high what to say significance level we do not have to take a very low significance level. Then we can take point α equal to 0.10 why we are doing this? Say here by being α is equal to 0.10 when we are increasing α my type 2 is also decreasing. Type 2 thing is also decreasing is not it? The type 2 error is also decreasing. What is type 2 error? Type 2 error means I am accepting a false null hypothesis.

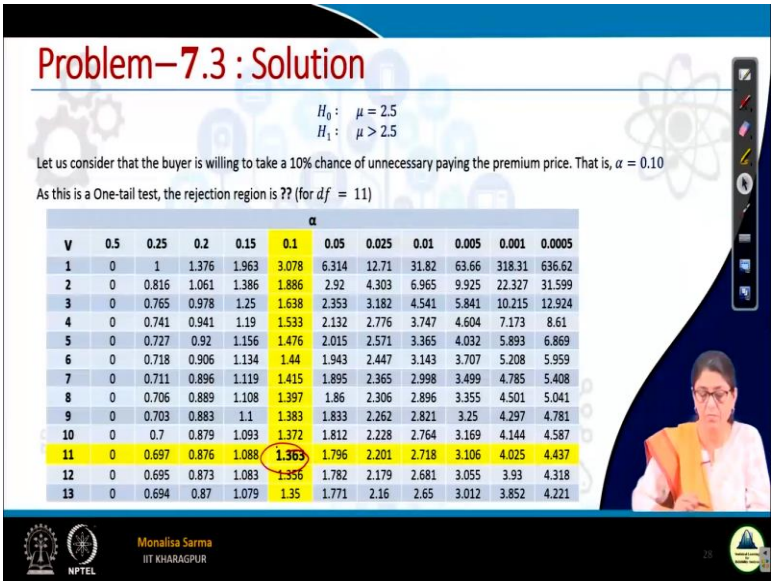
That is also I do not want a false null hypothesis means the apples are not actually big, but I am accepting it, if I make my α very small, my type 2 error will also be increased. So, in this case, I

do not want to increase the type 2 error also. So, I will keep in such a way that my α is also not very big. Anyway, you do not have to worry about this it is application specific whichever application person who wants to do this experiment and they will specify the α you just need to know the logic behind it, the science behind it.

And as well as how to do the techniques basically, as this is a 1 tailed test, because we are finding μ greater than 2.5. So, it is a 1 tail test. So, my significance level remains α 0.10 only it will not be divided by 2 and here what to say the average diameter that you see the standard deviation is not given when the standard deviation is not given it is just specified as claimed an average diameter of more than 2.5 inch.

Standard deviation is not given that means we will be considering t distribution now degrees of freedom how much degree of the freedom is 11.For degree of freedom 11 α is 0.10 what is my rejection region.

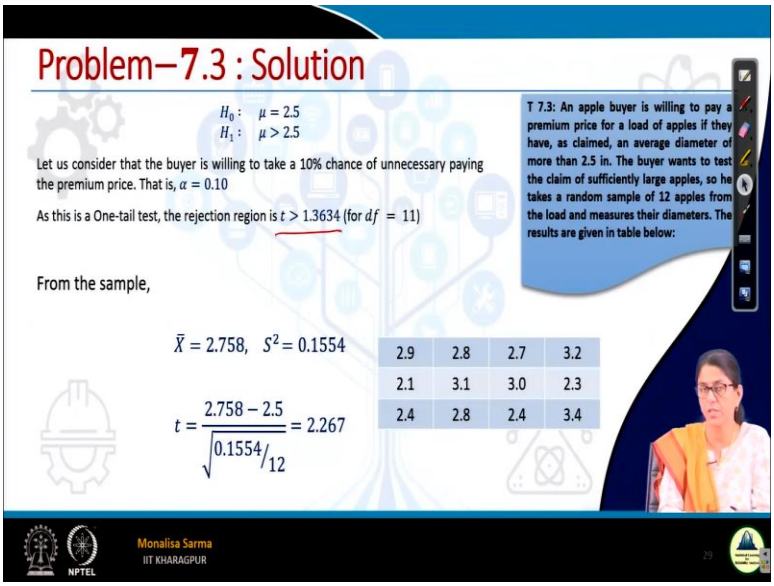
(Refer Slide Time: 26:16)



ν	0.5	0.25	0.2	0.15	0.1	0.05	0.025	0.01	0.005	0.001	0.0005
1	0	1	1.376	1.963	3.078	6.314	12.71	31.82	63.66	318.31	636.62
2	0	0.816	1.061	1.386	1.886	2.92	4.303	6.965	9.925	22.327	31.599
3	0	0.765	0.978	1.25	1.638	2.353	3.182	4.541	5.841	10.215	12.924
4	0	0.741	0.941	1.19	1.533	2.132	2.776	3.747	4.604	7.173	8.61
5	0	0.727	0.92	1.156	1.476	2.015	2.571	3.365	4.032	5.893	6.869
6	0	0.718	0.906	1.134	1.44	1.943	2.447	3.143	3.707	5.208	5.959
7	0	0.711	0.896	1.119	1.415	1.895	2.365	2.998	3.499	4.785	5.408
8	0	0.706	0.889	1.108	1.397	1.86	2.306	2.896	3.355	4.501	5.041
9	0	0.703	0.883	1.1	1.383	1.833	2.262	2.821	3.25	4.297	4.781
10	0	0.7	0.879	1.093	1.372	1.812	2.228	2.764	3.169	4.144	4.587
11	0	0.697	0.876	1.088	1.363	1.796	2.201	2.718	3.106	4.025	4.437
12	0	0.695	0.873	1.083	1.356	1.782	2.179	2.681	3.055	3.93	4.318
13	0	0.694	0.87	1.079	1.35	1.771	2.16	2.65	3.012	3.852	4.221

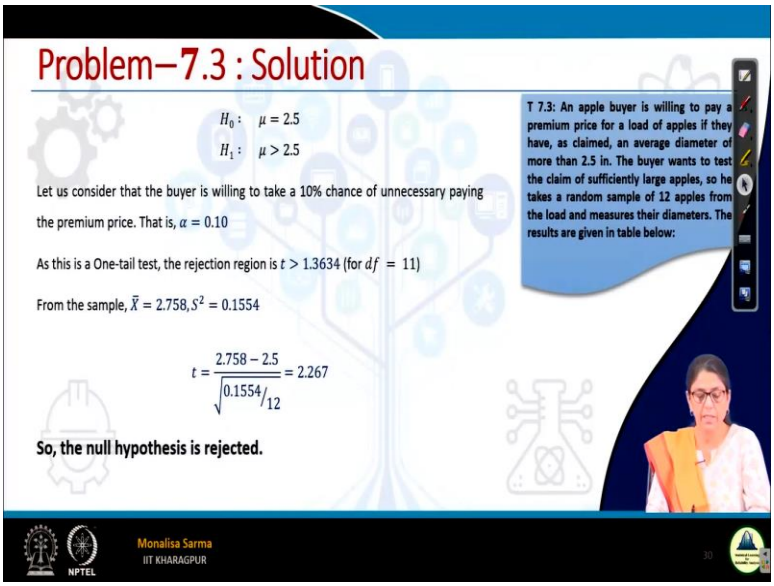
So, α 0.10, degree of the freedom is 11, 1.363 this is my rejection region. So, if from my sample, if I get value which is greater than this, then I will reject the null hypothesis that means, I can consider that means my apple or bigger size and I can pay that premium price. But if it is less than 1.363, that means the null hypothesis is I could not reject the null hypothesis. So, I will not pay the premium price.

(Refer Slide Time: 26:46)



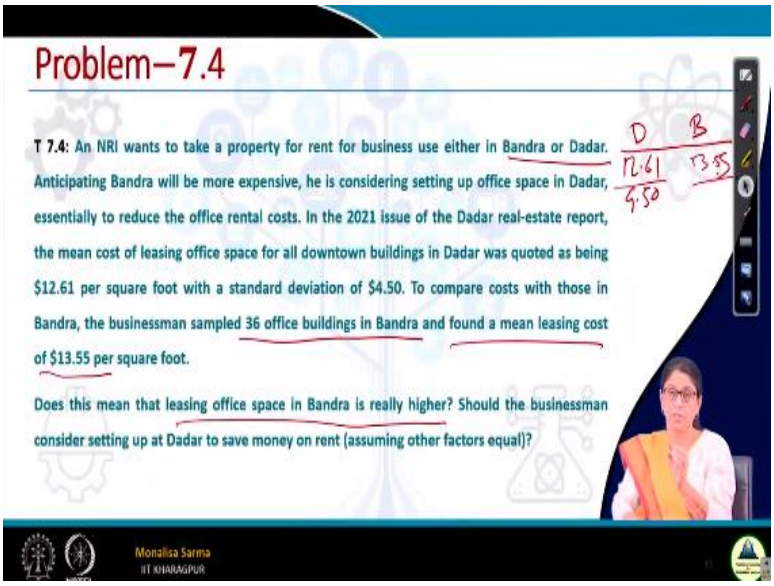
So, in this case, what happens so, the rejection region is greater than t 1.363. So, from the sample this is the sample you can calculate X bar you can calculate S^2 you can calculate by yourself S^2 is nothing but the variance you can calculate X bar is the mean. So, from the t value I got 2.267 so, 2.267 is better than this that means my null hypothesis is rejected. So, now hypothesis is rejected means alternate hypothesis accepted.

(Refer Slide Time: 27:16)



So, the buyer can pay the premium price for the apple.

(Refer Slide Time: 27:23)



So, one last question this question is a bit different in till now what we have tried to see, we have hypothesis something about the population and we are taking a sample from the sample we are trying to infer about the population whatever hypothesis value is correct or not. This is a bit different. We are just first go through the question then you will understand. See an NRI wants to take a property for rent for business use either in Bandra or Dadar, so this is Bandra, this is Dadar.

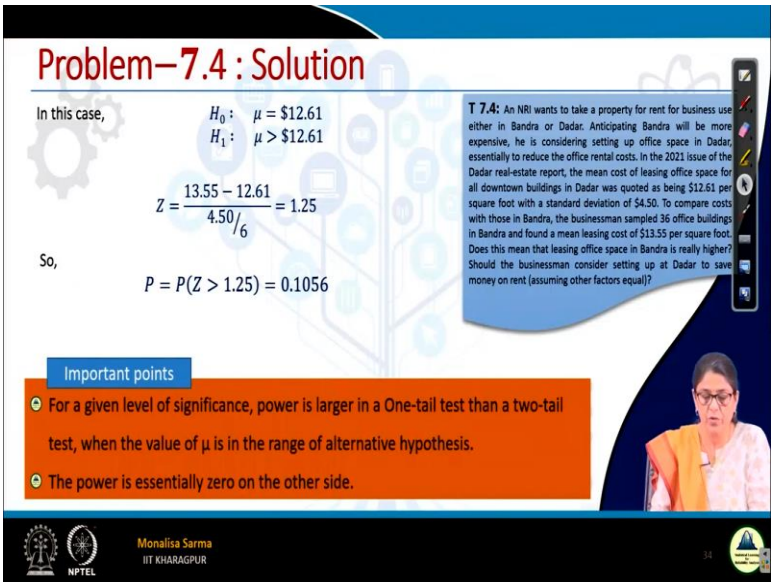
Anticipating Bandra will be more expensive he is considering setting up office space in Dadar, essentially to reduce the office rental costs so obviously he wants to reduce the office rental in the 2021 issue of Dadar real estate report, the mean cost of leasing office space for all downtown building in Dadar was quoted as being the mean costs is quoted as been 12.61 dollar with standard deviation of 4.50.

In a real estate report head is quoted for sorry, it is quoted for Dadar this is Dadar and this is Bandra. So, you want to set up in Dadar anticipating that it is less cost. So, from Dadar real estate report, he got this value of 12.61 per ^{2} foot and this is the standard division is 4.50 for Bandra, he does not have such a real estate report is not available. So, what he does to compare costs with those involved the businessman sample 36 office building in Bandra he himself sample 36 office buildings belongings he inquired basically.

He picked up different office buildings and inquired and found a mean leasing cost is 13.55 he found 13.55. See here for Dadar mean is 12.61 here it is 13.55. But in Dadar it is real estate report means it is as a whole it is given because means this is about the whole population of Dadar. And this whole Bandra has just taken some sample. Now from this sample basically he has to find out does this mean that leasing office space in Bandra is really higher.

So, from this sample, basically he has to find out whatever it is reflecting this 13.55, does that mean leasing office space in Bandra is really higher. So, that means from this sample he just he had to find out this is he got it from the sample of Bandra. So, from this he has to infer about the population of Bandra so, population of Bandra, if you find if it is very higher than he will lease space in Dadar.

(Refer Slide Time: 30:25)



So, what will be the hypothesis, definitely he wants to test because from the population, what he has for Dadar, he can assume that the same thing for Bandra. So, we can assume that μ is 12.61 and what he wants to test it is greater than 12.61 because he got 13.55 in the sample if you would have got less than 12.61 then he get tested it for less. So, he wants to test whether it is greater than this. So, now, since the standard deviation is given, he has to inquire about the mean definitely z distribution he will use. So, z is 1.25.

So, p if the significance level is not mentioned we can just specify the P value. So, P value this is again this is one tail just greater than one tail, what is the probability, corresponding to z is value is 1.25 see here we are not adding in the previous question we have added this because z was 2 tails in the single tail the probability corresponding to 1.25 is this. So, this is the P value. Now, depending on the significance level, you will take the decision or depending on what how much error he can be.

So, one important point for a given level of significance power is larger in 1 tailed test, then a 2 tailed test, when the value of μ is in the range of alternative hypotheses, why we have learned power is not it? What is power? Power is 1 - β what is power when the null hypothesis is false, we are truly we are rejecting a false null hypothesis that is power, is not it? Rejecting a false null hypothesis is power that is 1 - β.

So, given for 1 tailed test for the same level of significance, for 1 tailed test what happens α is higher because for 2 tailed tests α is becomes α / 2 significance level becomes α / 2. So, when the 1 tailed test significance of level of α remains α, so when significant level is α, that way, my β gets reduced, is not it? So, when my β gets reduced. So, what happens is 1 - β becomes more so for a given power is larger in a 1 tailed test then a 2 tailed test when the value of μ is in the range of alternative hypothesis.

When the value of μ is actually in the range of alternative hypothesis but if the value of μ is in the here see value of μ in less than it is equal to means is less than equal to when value of μ is less than that than actually the power is essentially 0. But we are not worried about that also, because we want to if it is greater than, so we need to find out the power on that. We do not want to find the power on it is less because it may be equal it may be less we are fine with that.

(Refer Slide Time: 33:12)

So that is all so these are the references as I mentioned before, thank you guys.