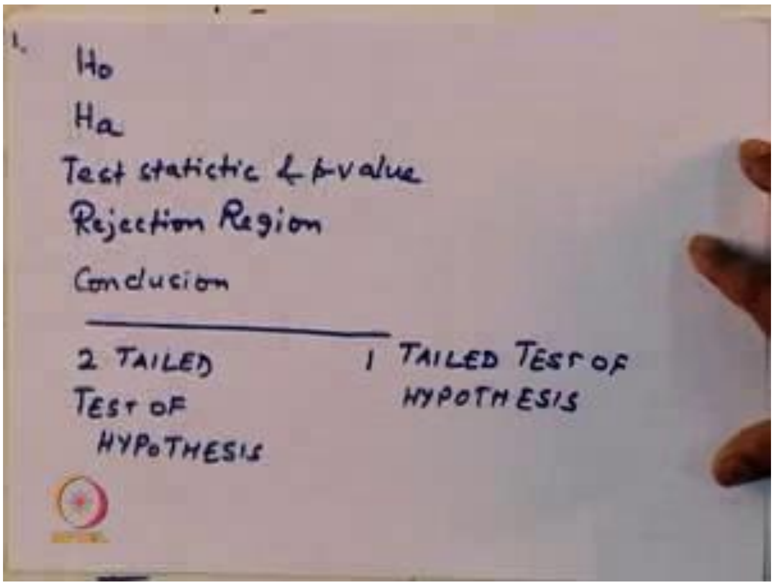
Introduction to Biostatistics
Prof. Shamik Sen
Department of Bioscience and Bioengineering
Indian Institute of Technology, Bombay

Lecture – 34
Test of Hypothesis – 4
(Type -1 and Type -2 error)

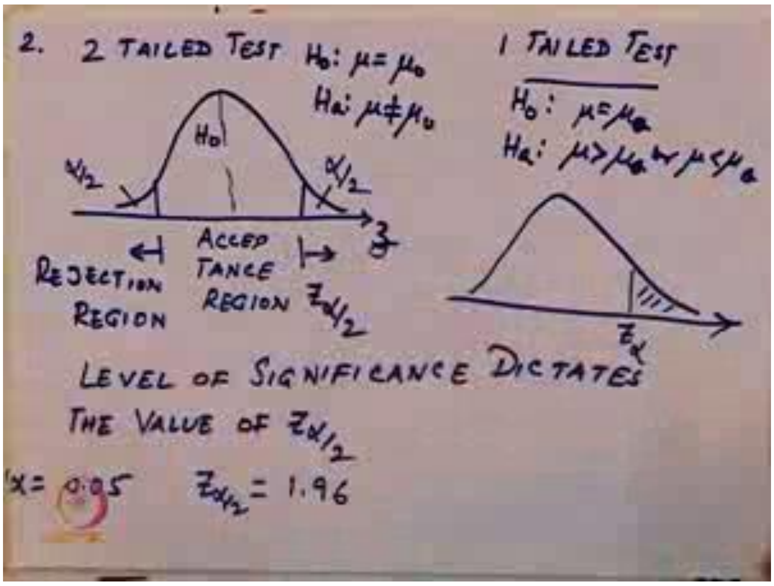
Hello and welcome to our lecture today. So, we will continue, we have our last session on a discussion with on test of hypothesis. So, just to briefly mention you have the null hypothesis, you have the alternate hypothesis, you compute a test statistics, test statistic and a p-value, you have a rejection region and you draw your conclusion. So, if you change any one of these clauses then it represents a new test of hypothesis.

(Refer Slide Time: 00:30)



And we have in last class we have discussed in great detail you can either have a 2 tailed test of hypothesis or a 1 tailed test of hypothesis.

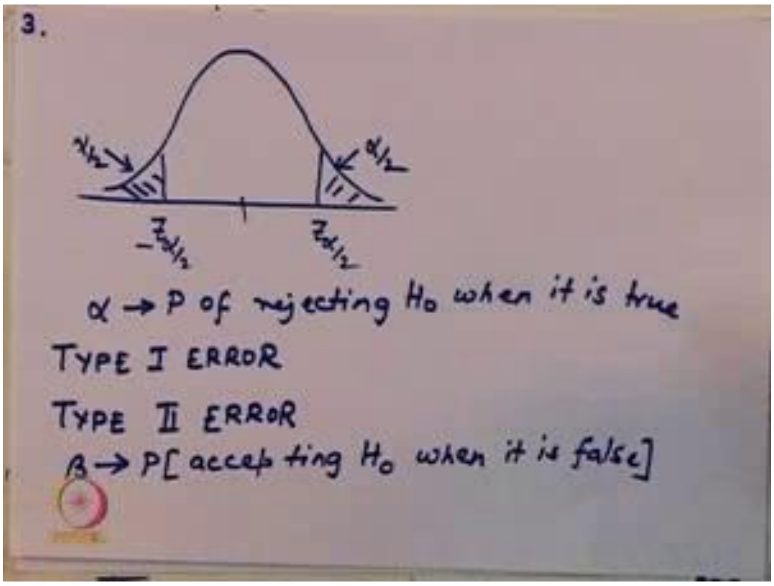
(Refer Slide Time: 01:43)



So for a 2 tailed test of hypothesis, for a 2 tailed test of hypothesis, for a 2 tailed test z value. So, this is the rejection region and here H naught is true, so this is the acceptance region. So, these values are the critical values are referred to as Z of alpha by 2. So, choice of Z of alpha x 2 is dependent on the confidence in the level of significance. So, level of significance dictates the value of Z of alpha by 2 and what we know is for 0.05 or alpha, alpha is nothing but the total area under the curve. So, both of these areas are alpha by 2 alpha is 0.5, Z of alpha by 2 is 1.96.

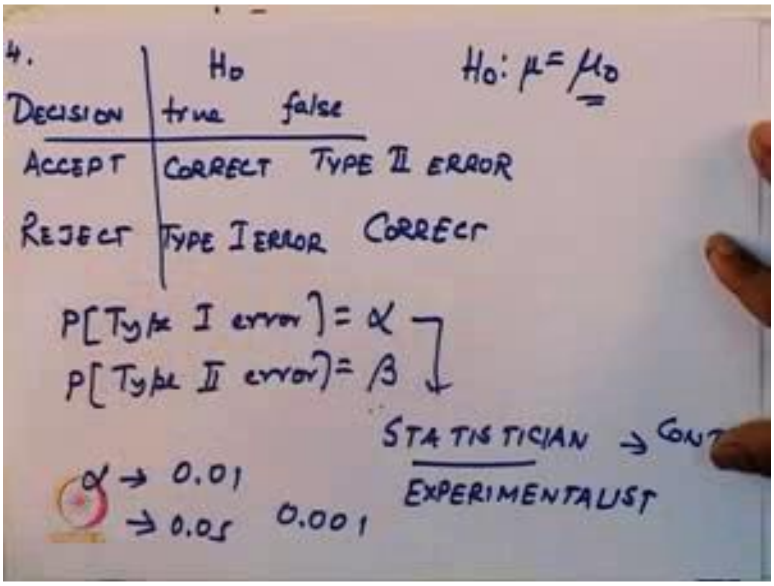
And for the 1 tailed test, so here in case of 2 tailed my, null hypothesis is mu is equal to mu naught and H a is mu not equal to mu naught. For the one tailed my H naught is mu equal to mu naught and H a is mu is greater than mu a or mu less than mu a sorry mu naught. And for the 1 tailed test you are only concerned about this area. So, you call this as Z alpha and you depending on the value of alpha you can accordingly change the value of c alpha.

(Refer Slide Time: 04:17)



So, this error that you accumulate, minus this area under this curve is alpha by 2 so the net area is alpha, so alpha represents the probability the net probability of rejecting H naught when it is true so, but alpha you have some control over. So, this is called a type I error which means that you can reject H naught when it is still true, but there is an alternate possibility of all our type II error, this. So, and beta is the probability of accepting H naught when it is false. So, you would imagine that the true effectiveness of a test, the true effectiveness of a test will depend on both of your choices of alpha and beta.

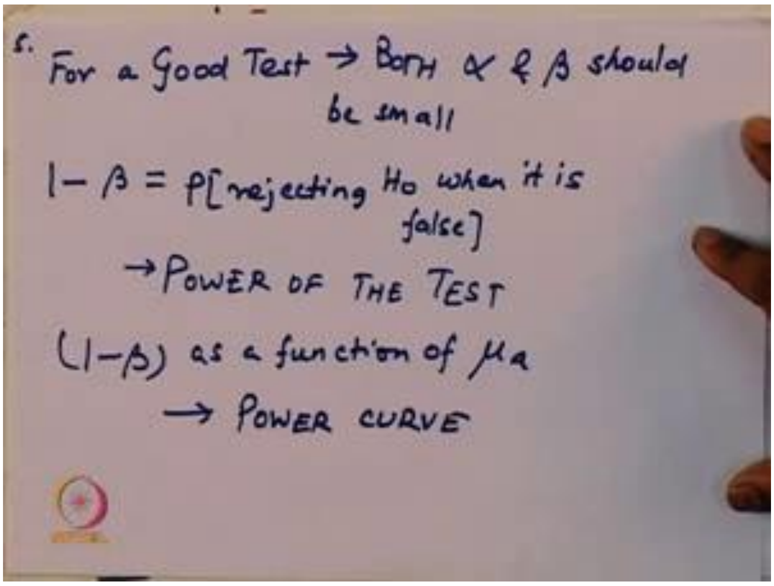
(Refer Slide Time: 05:50)



So, if I were to write down this plot again let us say you have H naught is true or H naught is false and you have a decision. If you accept that H naught is true. So, this is correct, but this is called a type II error. If you reject, if your final decision is reject this is your type I error and this is correct. So, there are 2 types of errors you can make. This is a type I error and this is your type II error.

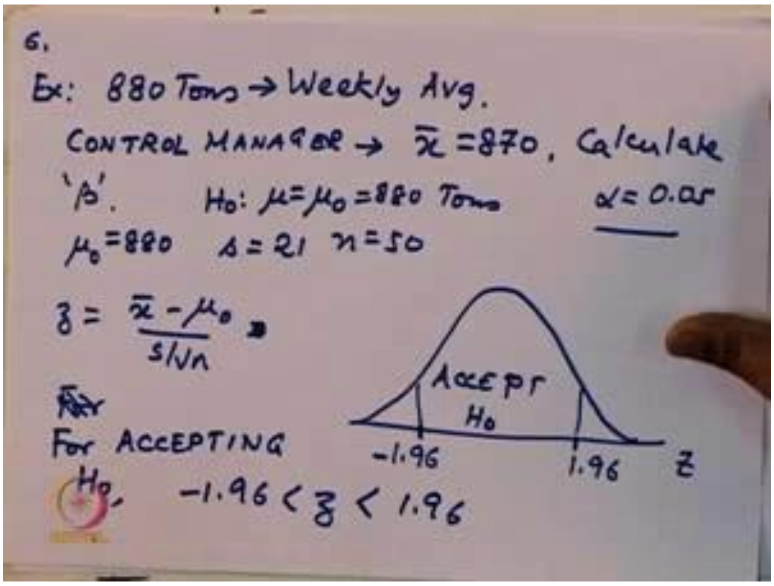
So, probability of type I error is alpha and probability of type II error is beta. Now what you see alpha this is in the steady statisticians hands or experiment lists, can control why you can choose alpha equal to either 0.01, 0.05, 0.001, so on and so forth. So, this is in your hands, but it is very difficult to choose beta why? Because beta is your type II error which is that you are accepting each term for it is false. So, your choice your H naught is mu equal to mu naught right. If your H naught is false, you have no idea of what the value of mu is. So, how can you go about setting the value of beta? So, this is very difficult to know. So, for a good test, for a good test both alpha and beta should be small.

(Refer Slide Time: 08:04)



Now I can calculate 1 minus beta which is nothing but probability of rejecting H naught. So, rejecting H naught when it is false and this is called the power of the test and you can generate 1 minus beta as a function of mu a and this curve is called a power curve.

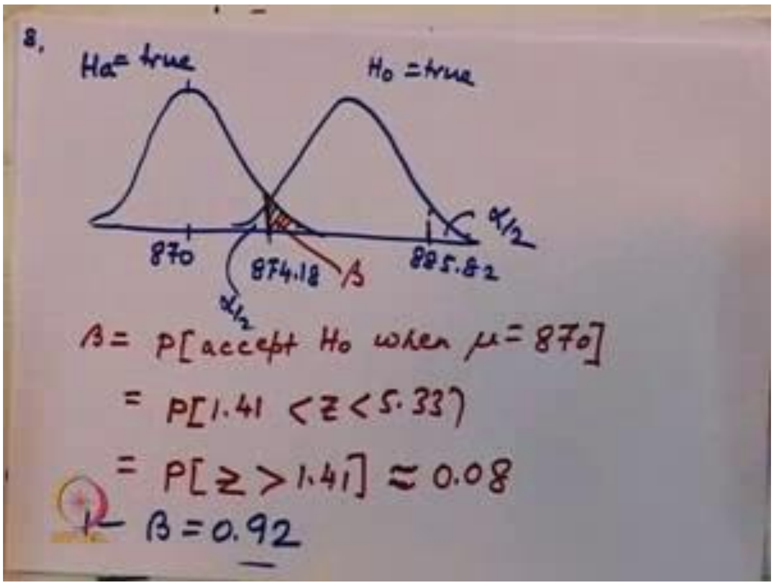
(Refer Slide Time: 09:26)



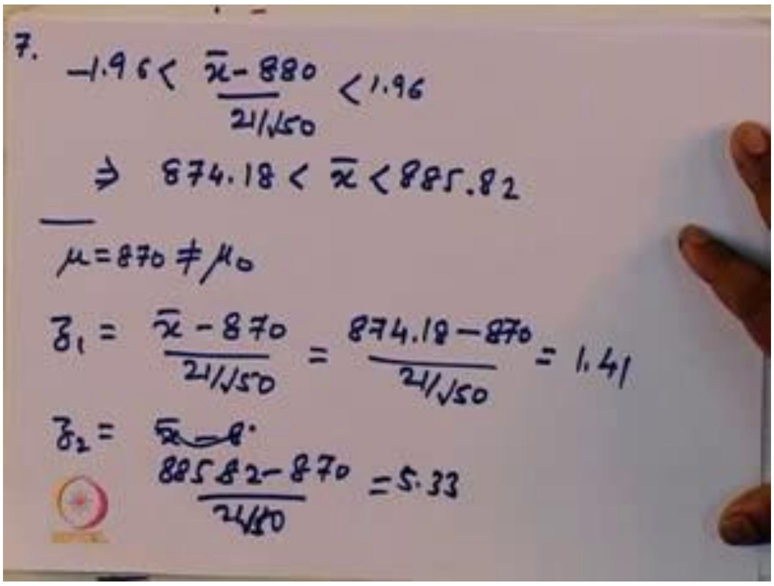
Let us see what we can say about beta or 1 minus beta. So, let us take an example. Let us go back to the example where we found was 880 tons was the weekly average, weekly average earlier known. Now for let us say, so the control manager does the statistics does the statistics and finds that x bar equal to 870. So, we want to calculate, to calculate beta if you want to calculate beta for this case.

Now, what we know, so my null hypothesis is mu equal to mu naught and this mu naught is 880 tons. So, mu is equal to 880 and from previous cases we had seen s was 21, n was 50. So, we calculate Z equal to x bar minus mu naught by s by root n. So, if you were to reject your H naught, if you were to reject your H naught. So, this Z you can calculate the value to be. So, let me draw it again. So, what is your H naught for alpha equal to 0.5? Let us say alpha equal to 0.05, for your 2 tailed test this value is 1.96.

So in order to accept your H naught your Z has to, so for accepting H naught my Z must be bound by, because this is my accept H naught area, this is a range over which I am going to accept my null hypothesis.



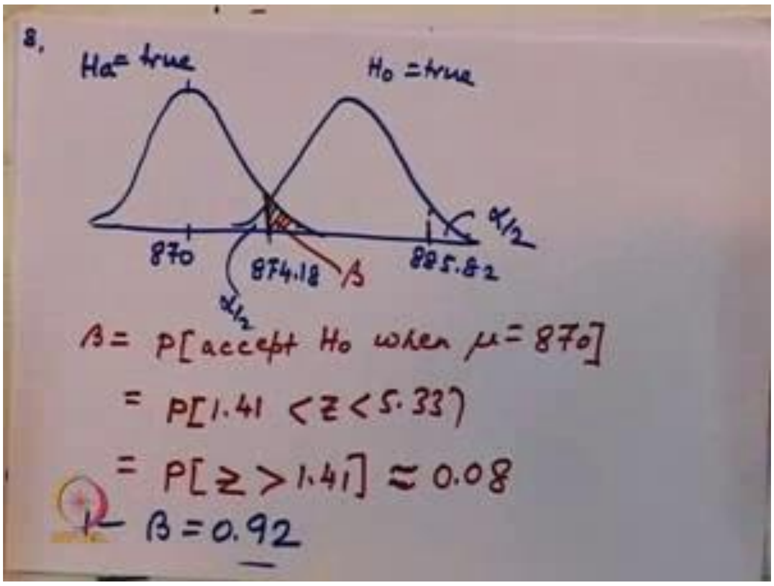
(Refer Slide Time: 12:19)



So, if you plug in the values. So, if you write down minus 1.96 less than x bar minus 880 by 21 by root 50 less than 1.96. So, if I plug in the values I will get a range of 874.18 less than x bar less than 885.82, but what I know this is if H naught was true, but what I know my mu is equal to 870 which is not equal to mu naught. So, I can calculate the test statistic z 1 and z 2 which is given by x bar minus 870 by 21 by root 50 and z 2 equal to x bar minus 870.

So using this expression I can plug in these values, I can find out corresponding to 874.18, 50 I get z 1 equal to 1.41 and z 2 is 880 5.82 minus 870 by 21 by root 50 - this gives me a value of 5.33. So, if I were to draw this curve together slide it, this is H naught equal to true, this is centered around 870, this is H a equal to true that mu is not equal to mu naught.

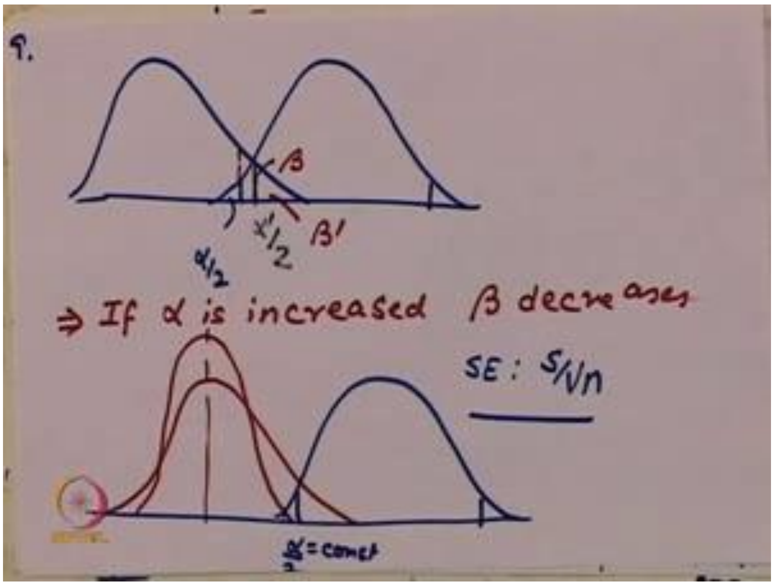
(Refer Slide Time: 14:08)



What I know is my acceptance range this I calculated at 874.18 and 885.82 this is my alpha. So, I can this is my alpha by 2 and this is a I is my alpha by 2, but I can extend this point high up and calculate. So, if I draw this line again I can calculate this total area under this part of the curve this has to be beta. Why? Because beta is probability of accept H naught when mu equal to 870, beta is equal to probability of accepting is not mu equal to 870. So, my beta is equal to probability 1.41 less than Z less than 5.33.

Now 5.33 is way to far from here, so for all practical purposes I can write quality of Z is greater than 1.41 and this probability comes out to be 0.08 roughly comes out to be 0.08. So, then I can write one minus beta is 0.92. So, this tells you the probability the probability of collected rejecting H naught when mu is true is 92 percent, probability of again probability of rejecting H naught when H a is true is 92 percent. Now one more thing you observed from this curve let me redraw this curve once more.

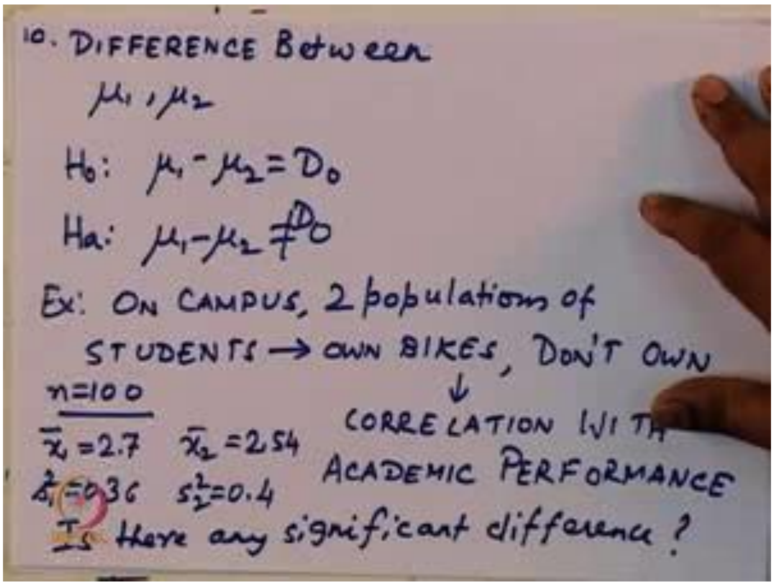
(Refer Slide Time: 16:50)



Let me draw 2 curves, let us say this is my alpha this area is alpha by 2, I extend it and this area is beta. What can I say about my choices of alpha and beta? Are alpha and beta linked by any choice? So, what you can clearly see from this curve is if I increase my alpha. So, let us say if I had chosen alpha prime equal to this let us say this is my alpha prime by 2 then corresponding to this area is going to be beta prime. So, this implies that if alpha is increased beta decreases, but if you want to prescribe, keep your alpha as constant what should be the strategy for decreasing beta right you want to decrease both alpha and beta. So, what I want to do is if this is my curve for alpha. So, imagine I want to keep alpha fixed alpha is constant, this is alpha by 2 and this is constant.

So, I can draw a curve which is this one or I can draw a curve because the spread of this is dictated. So, you know your standard deviation SE standard error is s by root of n right. So, if you increase your sample size then this will keep on decreasing that for the same value of alpha you keep on reducing the value of beta. So, this tells you how you might be able to, like for lower value of alpha you can even decrease beta by increasing the sample size.

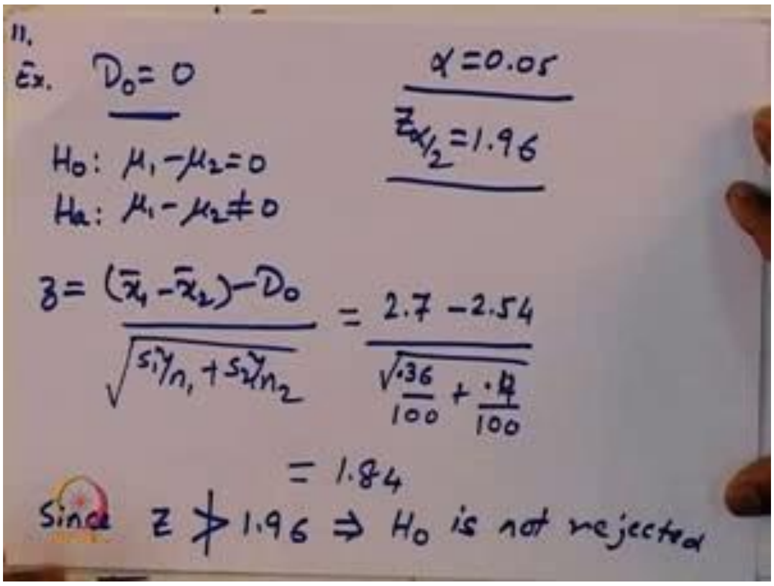
(Refer Slide Time: 19:33)



Now let us see for finding out the difference between population means mu 1 and mu 2 what should be your null hypothesis? Your null hypothesis can be something like mu 1 minus mu 2 equal to D naught, but D naught is some prescribed value that you control and your alternate hypothesis is for a 2 tailed distribution you have mu 1 minus mu 2 not equal to D naught. Let us take an example. So, imagine, you want to test. So, on campus 2 populations of students, students those who owned bicycles or bikes and those who do not own and your agenda is to see whether there is any correlation with academic performance and whether there is any correlation with academic performance .

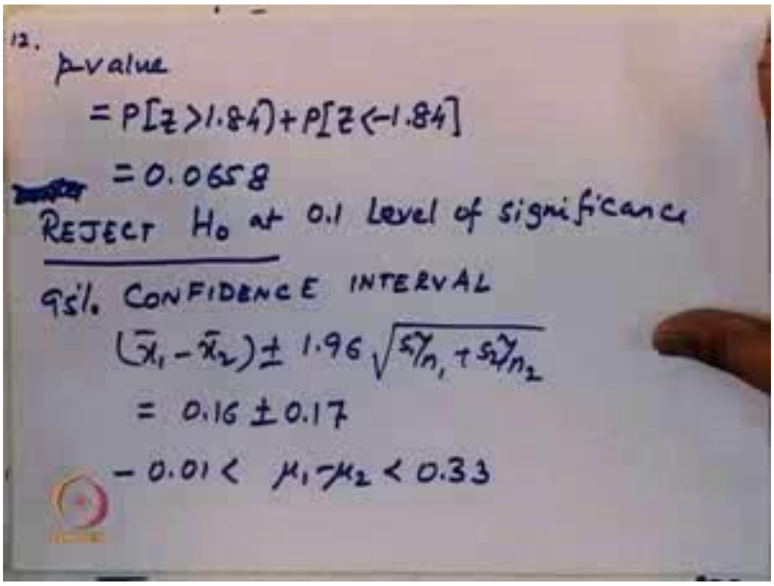
So, what you are given, let us say you have n equal to 100 for both x 1 bar is 2.7, x 2 bar is 2.54, s 1 square is 0.36, s 2 square is 0.4. So, is there any significant difference in the academic performance? So, for addressing this, for addressing this, you just want to know whether there is a difference in their academic performance or not which means that in our case this particular case. So, this example continued you want to choose D naught equal to 0, you choose D naught equal to 0. So, your H naught is mu 1 minus mu 2 equal to 0 and H a is mu 1 minus mu 2 not equal to 0.

(Refer Slide Time: 22:23)



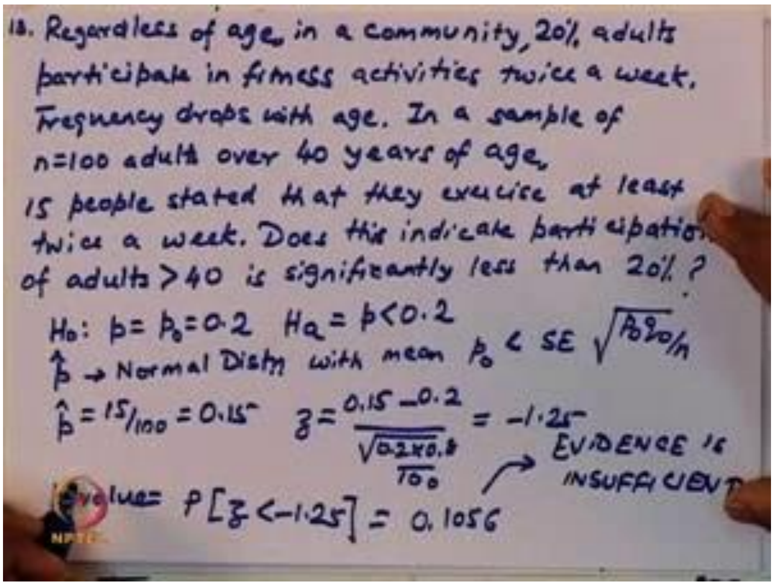
So, in the general case Z in the test statistic has to be defined as. Since for this particular case our D naught is 0 we can simply put in these values and you get a value for the test statistic as 1.84. So, how will you know whether this is significant or not? So, corresponding to alpha equal to 0.05 you are doing a 2 tailed test right a 2 tail test. So, Z of alpha by 2 is 1.96. So, since Z is not greater than 1.96, I will say H naught is not rejected and this is also cleared by computing the p-value (Refer Time: 23:59).

(Refer Slide Time: 23:59)



The p-value turns out to be probability of Z is greater than 1.84 plus probability Z is less than minus 1.84 and this you get a value of 0.0658. So, what you see is you can reject. So, this is possible to reject H naught at 0.1 level of significance, so in this case you do not reject H naught do not reject. So, we do not reject H naught we do not reach. So, we can reject it is not at 0.1 level of significance, at 0.05 level of significance we do not reject H naught. So, and you can see that if you calculate the 95 percent confidence interval, the 95 percent confidence interval comes out to be x 1 bar minus x 2 bar plus minus 1.96 times SE which is root of x 1 square by n 1 plus s 2 square by n 2 this comes out to be 0.16 plus minus 0.17. So, the bounds become minus 0.01. So, because it is possible to have a value of 0 that is no difference or 0.2 which is greater or minus 0.21, so that is why you see that you cannot reject, you cannot say that they are statistical difference between these 2 populations.

(Refer Slide Time: 26:17)

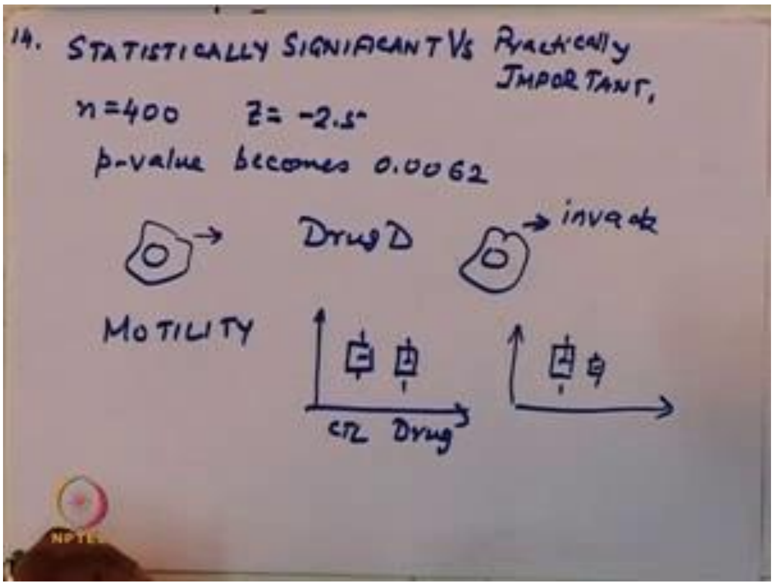


Let us do one more example. So, regardless of age in a community 20 percent adults participate in fitness activities fitness activities twice a week, and this frequency drops with age this drops with age. Now let us say in a sample of n equal to 100 adults. So, what you find adults over 40 years of age, 15 people stated that they exercise at least twice a week. So does this indicate participation of adults greater than 40 is significantly less than 20 percent?

So, what you are given in this problem, so 20 percent, you have a problem about proportions, your null hypothesis H o is p equal to p naught equal to 0.2 and your alternate hypothesis is p less than 0.2. So, your proportion, proportion p hat say know it follows normal distribution with mean p naught, p naught and standard error root of p naught q naught by n. But in your case, so in a specific case you note p hat is 15 out of 100 which is equal to 0.15. So, if you calculate the Z statistics it is 0.15 minus 0.22 minus 1.25. So, the p-value comes out the probability of Z less than minus 1.25 and this is 0.1056. So, this would tell you that your evidence is insufficient this would tell you that your evidence is insufficient.

So this also brings us to another point that how do you make a difference or how do you make a difference between statistically significant versus practically important. How do you make a difference mean statistically significant and practically important.

(Refer Slide Time: 31:14)



So, in the above case whatever this case you calculated if let us say you increase your n equal to 400 then your p-value becomes 0.0062 because your Z becomes minus 2.5. So, now, it is significant, but it may not always be practically important. For this you need to see your specific case for example, if I want to see. So, I have a cell a cancer cell which kind of is invasive and I have drug D which targets this it is ability to invade. So, if you measure it is motility. So, imagine you have a plot like this.

So, there is as you can clearly see from this plot there is very little difference between the control group and your drug group, but sometimes let us say if I were to draw in a different case, what the drug has done is the extent of variation in the population is reduced, but the mean has not changed or it has just shifted slightly to the left or up or down. So, even though this is might be statistically significant it may not help hold any practical importance. So, this has to be exercised at the on a case by case basis.

With that I thank you for your attention and I look forward to next class. In the next class we will start discussing about the T-test, the students T-test.

Thank you.