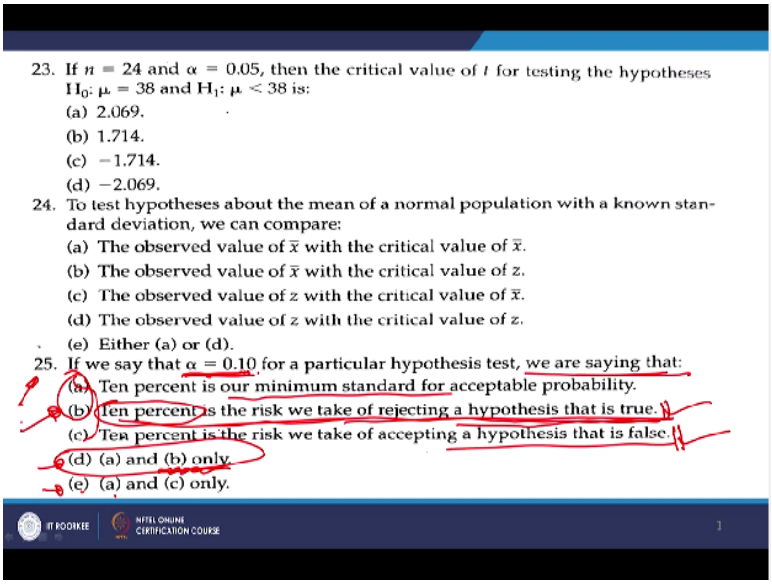
Business Statistics
Prof. M. K. Barua
Department of Management Studies
Indian Institute of Technology - Roorkee

Lecture – 36
Hypothesis Testing: One Sample Test

Hello friends. I welcome you all in this session. As you are aware in previous session, we were discussing about lot of questions related to hypothesis testing and we have worked out couple of statements and we have seen some of them were true, some of them were false. Let us continue the same exercise. We will look at question number 25 once again.

(Refer Slide Time: 00:50)



So if we say that alpha=0.10 for a particular hypothesis which we are testing, what does it mean? What are we saying actually? Does it mean that the 10% is our minimum standard for acceptable probability? Does it mean that 10% is the risk we take of rejecting a hypothesis that is true? 10% is the risk we take of accepting a hypothesis that is false. a and b and a and c. So a and c which of these options are correct?

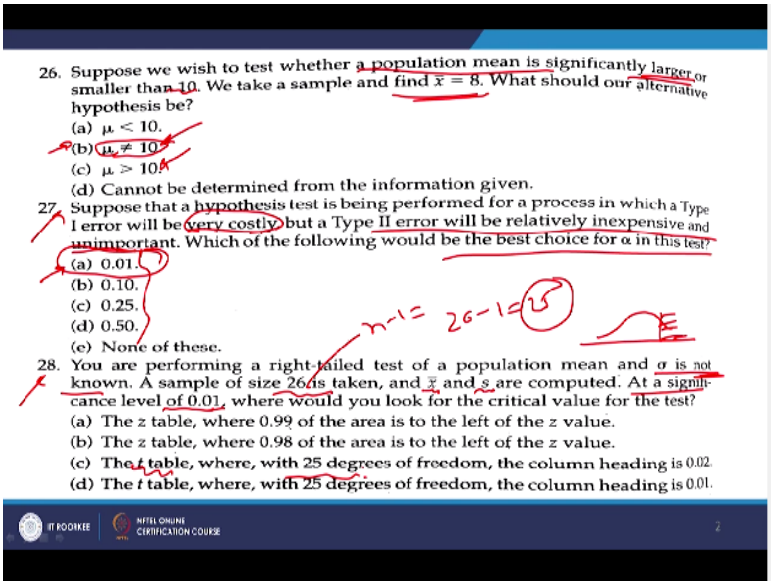
Just think over on this question for some time. So alpha is given, right. So what does alpha mean? Does it mean 10% is the risk we take of accepting a hypothesis that is false or hypothesis that is true. In fact, if you look at b and c option, in fact, though. In fact, these 2, only one of them can be correct. Is not it? You cannot have both these options correct. So when you say a

and c, it is possible that if c option is correct then you might have a situation where a is also correct.

Then this would be the answer and if let us say, b is correct, then we do not know whether a is correct or not. So if a and b, both are correct, then this would be the answer. So at this point in time, what is the correct answer for this particular question? So if you look at this question carefully, alpha is nothing but type I error. It is the probability of rejecting a hypothesis when it is true, right.

So this b part is true, right and when we say b part is true, it means this is true, means in d if you take, let us say, these b is correct, then b is there in d part, right. So b is true. 10% is our minimum standard for acceptable probability. So a is also true. So the right answer is d for this particular question, right. There is no other answer except d, right.

(Refer Slide Time: 03:15)



Let us look at one more slide on these questions. Suppose you wish to test whether a population mean is significantly larger or smaller than 10. Our population mean is significantly larger or smaller than 10. We take a sample and find sample mean is 8. What should our alternative hypothesis be? So can you think of alternative hypothesis. Is significantly larger or smaller than 10.

So the alternative hypothesis would be what? Let us say if this is smaller than, right, then alternative hypothesis would be this and if it is, let us say, larger, then again alternative hypothesis is this. So b is the correct answer for this particular question, right. Suppose that a hypothesis test is being performed for a process in which a type I error will be very costly, but type II error will be relatively inexpensive and unimportant.

Which of the following would be the best choice for alpha in this case? 0.01, 0.10, 0.25, 0.50, none of these. So what is type I error? Probability of rejecting a null hypothesis when it is true and we know that if we commit type I error, it is very costly. But if type II error, is relatively inexpensive and unimportant. So what would be the choice of alpha in this case? This can you tell me what would be the right answer out of a, b, c, d, e.

In fact, as in one of the previous examples here, what we have said? Alpha is equal to 10, right, 10%, right or is 0.10. So what we have said that? 10% is the risk we take of rejecting a hypothesis that is true, right. So 10% is the risk we take of rejecting a hypothesis that is true. So would you like this percentage to be more and more or less and less. You want this percent to be less and less. Because you should not reject the null hypothesis when it is true.

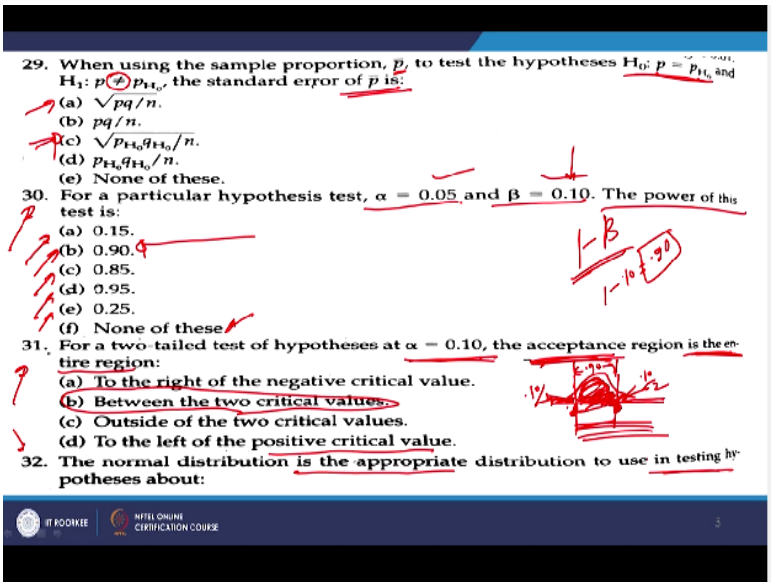
So alpha has to be as low as possible out of these 4. So the answer would be 0.01, right. So the probability of rejecting null hypothesis when it is true is just 1% here in this case. Is not it? Let us look at next question. You are performing a right-tailed test of a population mean and standard deviation is not known. Is not known, right. This is you are performing a right-tailed test, right.

A sample of size 26 is taken and x bar sample mean and sample standard deviations are computed, okay. At a significance level of this, 0.01, where would you look at for t value in the t table? So in fact, in different books, you will have different ways of representing t values. So whatever we have been following so far, you just look at the t table. At what degrees of freedom you should look at? At n-1, right.

So this is 26-1, 25 degrees of freedom, right. So we should look at the t value at 25 degrees of

freedom, right. And whatever is that value, that would be the right answer, okay.

(Refer Slide Time: 07:46)



Let us look at the next question. When using the sample proportion, p bar, to test hypothesis where null hypothesis P = P_{H0} and alternative hypothesis is not equal to. The standard error of sample proportion is what? Is it under root of pq/n. Third option is under root of pH0qH0/n, or none of these. So in fact this is quite a simple question. You need to just remember formula for this, right.

So c is the correct answer in this case, okay. Not much to think in this particular example. Let us look at the next question. For a particular hypothesis testing, alpha is this, beta is this. The power of this test is 0.15, 0.90, 0.85, 0.95, 0.25, none of these. For a particular hypothesis test, alpha this, beta this, what is the power of test. So what is power of test first of all. Power of test is the probability of rejecting a null hypothesis when it is false, right.

So this is represented as 1-beta. Is not it? So beta is this, 0.10. So 1-0.10 is 0.90 is the power of test. So the correct answer is b, right. Let us look at the next question. For a two-tailed test of hypothesis at alpha=0.10, right. So it is a two-tailed test, alpha is 0.10. The acceptance region is the entire region. To the right of the negative critical value. In fact, let us first draw how this distribution would look like, right for this question.

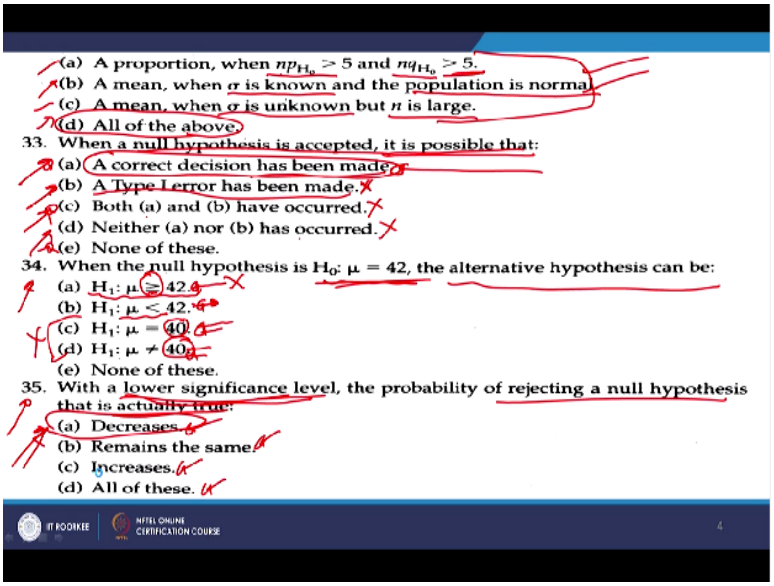
So it is a two-tailed test, right. So you have two rejection regions. So this rejection region and other rejection region. So you have got these 2 rejection regions and this alpha is 0.10. So this alpha/2, right. So it is 0.10/2 here and 0.10/2 here, right. And this is 0.90. So the acceptance region is the entire region to the right of the negative critical value, this negative critical value, this, sorry this is positive critical value.

This is negative critical value. So to the right of the negative critical value, so a part is, is this the rejection region? Or between 2 critical values, this one, between these 2, right. This negative critical value, positive critical value and this region. Outside the 2 critical values, this side or this side, then to the left of the positive critical value. This is positive critical value and left is this, right.

So what will be the correct answer? The acceptance region is the entire region. What it would be? It is very simple. It is this region, right. It is between 2 critical values. In fact, we have said initially that when you said two-tailed test, it means there are 2 rejection regions, right. And the region which is not rejection region is the acceptance region, okay. So this is how you can solve question number 31. Let us look at next question.

The normal distribution is the appropriate distribution to use in testing hypothesis about, what? So normal distribution is appropriate distribution to use in testing hypothesis about a proportion, when $np_{H0}>5$ and $nq_{H0}>5$.

(Refer Slide Time: 12:20)



Second option is a mean, when standard deviation is unknown and population is normal. So under which of these situations normal distribution is appropriate, that is the question, right. A mean, when standard deviation is unknown but n is large or all of the above. So you have to find out the answer to this question and the question is when do we apply normal distribution, right. For application of normal distribution, you need to have all these conditions.

So all these are correct. So answer is d, okay. Let us look at next question. When a null hypothesis is accepted, it is possible that a correct decision has been made, a type I error has been made. What is type I error? The probability of rejecting a null hypothesis when it is true, right. So here you are not rejecting. A null hypothesis is accepted then it is possible that a correct decision has been made or type I error has been made or both a and c have occurred.

It means correct decision has been made and type I error has been made. This is the c part, right. d part, neither a nor b has occurred. None of these. So what would be the right answer. When a null hypothesis is accepted, it is possible that then there is only one situation possible which is a. This cannot happen because you have not rejected null hypothesis. Now when b is not the answer, c cannot be answer because b is there.

d cannot be answer because b is there. Of course, so the only possible answer left is a correct decision has been made, right. So answer a is correct or option a is correct in this case. Let us

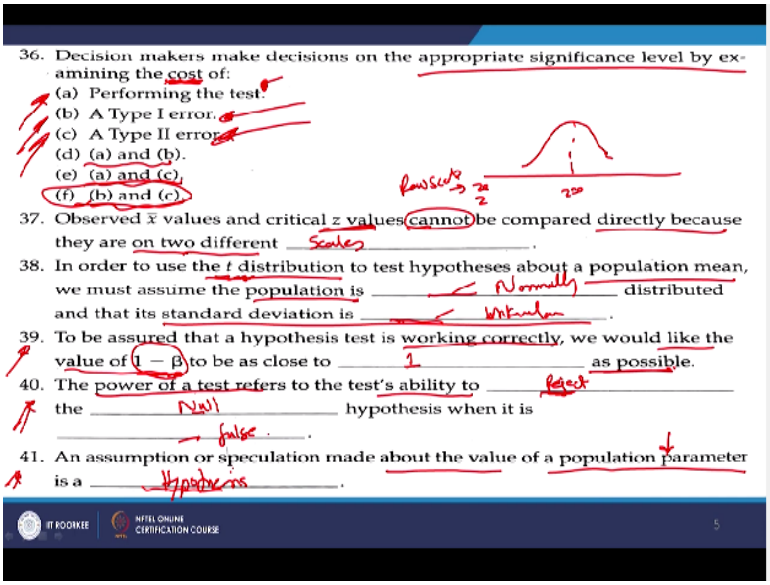
look at next question, 34. When the null hypothesis is this, mean=42, the alternative hypothesis can be, this is a very good question. So it will help you in framing null and alternative hypothesis. So when the null hypothesis is this, mean=42, the alternative hypothesis, mean=42 or greater than 42; is less than 42, the second option.

Third option is, is equal to 40, is not equal to 40. So if you look at these options very carefully, how you can have 40 and 40 in alternative hypothesis? So these 2 options are ruled out. Is not it? Okay. Now since you are saying that mean is equal to 42 and here, one of the options is greater than or equal to 42. Since equal to is already here in this question, so this cannot be option. So only option left is this.

So for 34, b is the right answer. Mean is less than 42, okay. Now let us look at the next question. With a lower significance level, the probability of rejecting a null hypothesis that is actually true decreases, remains the same, increases, all of these. With a lower significance level, the probability of rejecting null hypothesis that is actually true is decreases, right. So a is the right answer.

In fact, when you increase the significance level, let us say you test a hypothesis at 50% significance level. Then there is a probability of rejecting null hypothesis when it is true, right. So at lower significance level, this probability decreases, right. At high significance level, this probability increases, right. So answer to thirty fifth number question is a.

(Refer Slide Time: 17:54)



Let us look at next question. Decision makers make decision on the appropriate significance level by examining the cost of what? Performing the test, type I error, type II error, a and b both, a and c, b and c. Just think over it for a while and try to answer this question. Decision makers make decision on the appropriate significance level by examining the cost of, in fact, this will give you an idea.

Cost of what? So when businessmen are decision makers incur cost, they incur cost not much while performing the test but when there is type I error committed and type II error committed. So they should minimize these errors, okay. So answer to this question would be b and c. Is there any option? Yes, b and c is the option. So f is the right answer to this question. Let us look at next question.

Observed mean value and critical z values cannot be compared directly because they are on, since they are on 2 different what? Observed x values and critical z values cannot be compared directly. Generally, what do we compare? We compare the calculated z value with the critical value. Is not it? So in fact, you have a two types of scales. So you can have one x scalar, it is known as a raw scale and the other one is z scale, right or z scale, z=0, right.

Mean is 0 and standard deviation is 1. Is not it? This is what we have seen. So they cannot be compared because they are on two different, what? Scales, right. Let us look at the next question.

In order to use the t distribution to test hypothesis about population mean, we must assume the population is, what distributed. This is dash distributed and its standard deviation is what? What would be the, how would you fill up, fill in these 2 blanks.

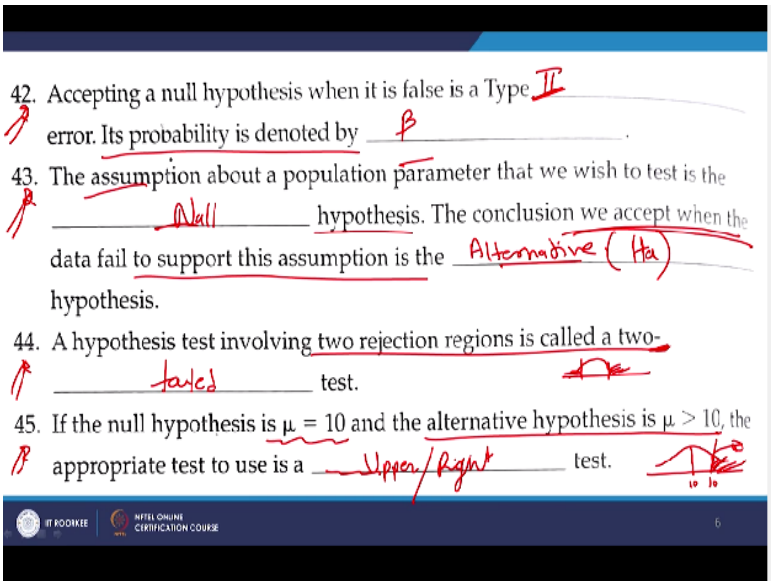
In order to use t distribution to test hypothesis about population mean, we must assume the population is, is what? Normally distributed or some non-normal distribution or what. And the standard deviation is what? This is normally distributed and standard deviation is unknown, right. So it is unknown. So for question number 38, this is as I said normally distributed and this is unknown, right.

Standard deviation is unknown. But there is one more condition. The condition is the sample size has to be less than 30, right. 39, to be assured that a hypothesis test is working correctly, we would like the value of 1-beta, what is this called? This is called power of test, right. Probability of rejecting a null hypothesis when it is false, right. So you should, what should be the ideal condition in hypothesis testing?

You should always reject the null hypothesis which is false, right. So that should be the goal which is known as power of test, 1-beta to be close to 1, right, as possible, okay. Let us look at next question, question number 40. The power of a test refers to the test's ability to dash the dash hypothesis when it is dash. Just think over it. It is not a simple question. The power of test, in fact, we just talked about power of test refers to the ability to reject the null hypothesis when it is true or false?

Tell me? The power of test refers to the test's ability to reject the null hypothesis when it is false. Is not it? So this is how you should fill up the blanks, right. Let us look at question number 41. An assumption or speculation made about the value of a population parameter is what? It is null hypothesis, alternative hypothesis, or just hypothesis. What would be the answer? It is just hypothesis, right. So let us look at next question.

(Refer Slide Time: 24:38)



Accepting a null hypothesis when it is false is a type error. Its probability is denoted by what? So we were told a lot about type I and type II error. I hope that you would be able to answer this particular question. Accepting a null hypothesis when it is false is, is type I or type II error? What is type 1 error? Type I error is rejecting a null hypothesis when it is true. But here, we are accepting a null hypothesis when it is false, is a type II error, right.

It is not type I error. Its probability is denoted by what? Alpha or beta? Alpha is type I error. So this is beta, right. Let us look at next question. The assumption about a population parameter that we wish to test is dash hypothesis. The conclusion we accept when data fail to support this assumption is dash hypothesis. Just look at carefully. The assumption about a population parameter that we wish to test is, is what hypothesis?

Is null or alternative? It is null hypothesis, right. The assumption about population parameter. The conclusion we accept when data fail to support this assumption is the alternative hypothesis. Is not it? We also represent it by H1 or Ha, is not it? Let us look at the next question. A hypothesis test involving two rejection regions is called a two test. What? A two, what? A two alpha test, two beta test, or what?

It is called tailed test, right. So there will be two rejection regions, right. This is one rejection region and this is second rejection region, right. Let us look at the last one. If the null hypothesis

is, mean=10 and the alternative hypothesis is mu>10, the appropriate test to use is what? So let me draw this distribution. So you have got mean over here, 10 and this is your null hypothesis. This is your more than 10, right.

So this is your rejection region. So this is a case of what type of tail. This is one-tail test. So you can write option over here is, let us say, the appropriate test to use is one-tailed test. That can be one of the answers or the more appropriate would be an upper tail test, right. An upper tailed test or right-tailed test. Is not it? So with this, let me wind up today's session. Let me summarize what we have done in today's class.

In today's session, we have looked at couple of fundamentals of hypothesis. Because by reading theory and by solving questions, you do understand concepts but by looking at all these 45 questions where in, there were questions on almost all the points related to hypothesis testing. So we have worked out these 45 questions and I hope that these questions would have made your fundamentals did clearer. So with this let me finish today's session.

Next session we will work out some questions on either two-tailed test using 2 samples or let us say one-tailed test using 2 samples, right. So let us see what exactly we would be doing in next class. Thank you very much.