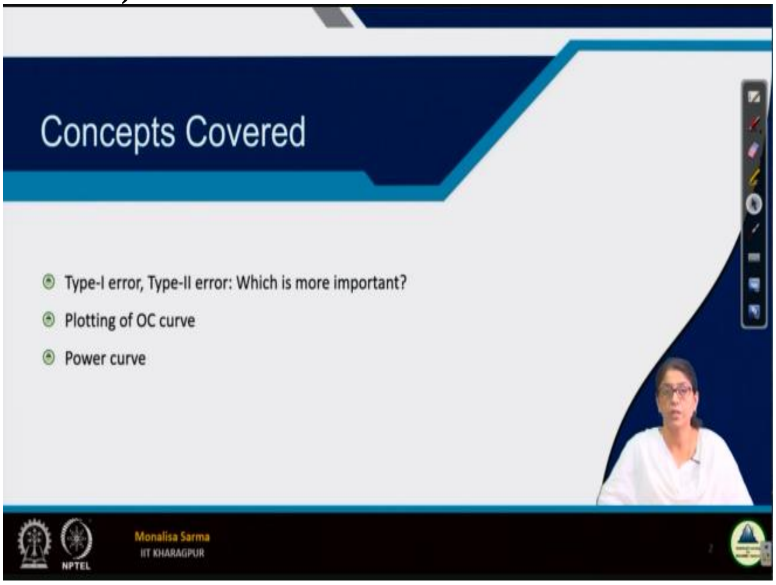
Statistical Learning for Reliability Analysis
Prof. Monalisa Sarma
Subir Chowdhury School of Quality and Reliability
Indian Institute of Technology – Kharagpur

Lecture - 24
Statistical Inference (Part 3)

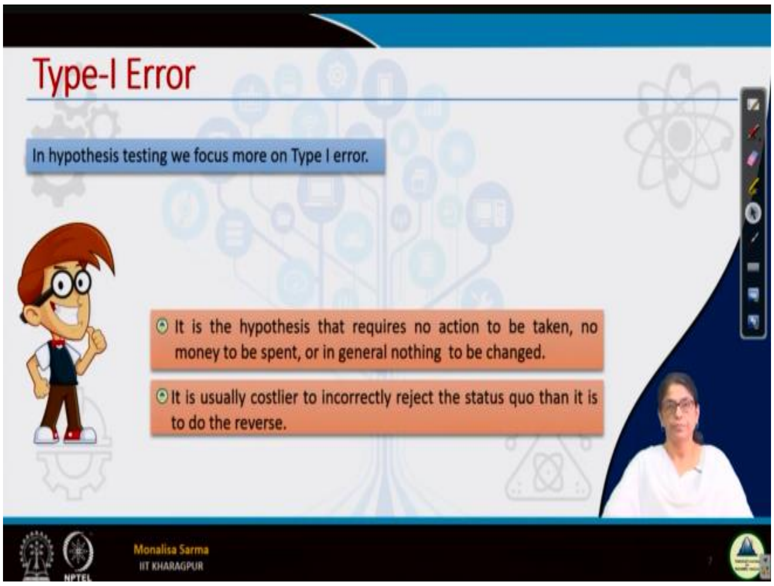
(Refer Slide Time: 00:33)



Hello, everyone so in continuation of our earlier lecture on statistical inferences. Today, we will be learning few more topics. So, in my earlier lecture, we have learned what is type 1 error, what is type 2 error. Now we will see which is more important, why we should, mainly we will see we focus more and more on type 1 error. That is why we have that significance level what I have mentioned significant level is mainly with respect to the type 1 error.

I did not mention about type 2, error, why we focus more on type 1 error, why it is considered more important? We will be discussing that and then how to plot the OC curve, what is an OC curve? How to plot that? And then we will discuss what is an power curve.

(Refer Slide Time: 01:08)



So, in hypothesis testing, we focus more on type 1 error, why? That you have noticed we focus more on type 1 error why? Because see, when the first step is, we specify a 0 h 1 and a significant level what is the significant level? Significant level is the maximum acceptable probability of rejecting a true null hypothesis, is not it? So, that is α that is what type 1 I did not mention anything about type 2 that means what type 2 is not important, we will see that. So, what is basically the null hypothesis?

Null hypothesis is the hypothesis that requires no action to be taken no money to be spent, or in general nothing to be changed. Already I have mentioned about when I talked about what is and how to form the hypothesis. So that is my null hypothesis. Null hypothesis is the maintaining the status quo like where we do not want to do any action. No, we do not have to do any more.

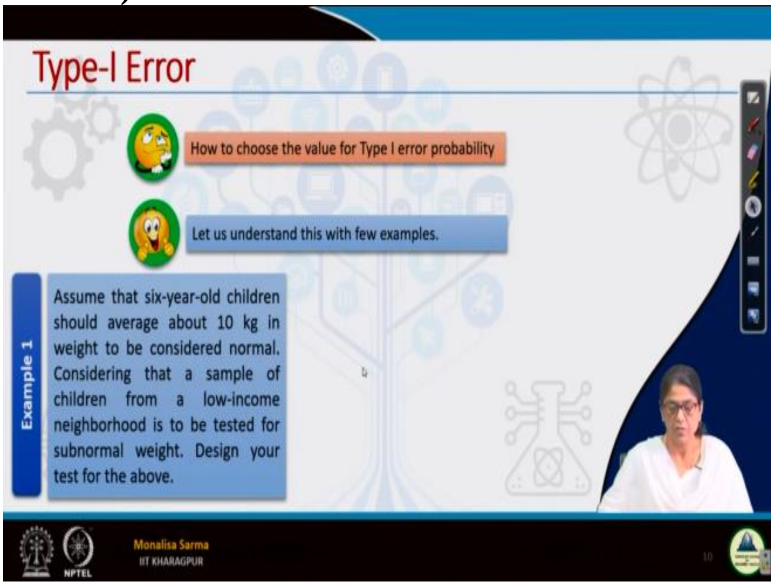
We do not have to spend any money we do not have to change anything so that is my null hypothesis. So say, like in my example, that machine that fills medicines in the 8ml bottle, so that there is a machine which fills exactly 8ml medicine in this bottle that is my null hypothesis. Remember, I have specified my null hypothesis $\mu = 8$, then what I wanted to prove? I wanted to prove μ is not equals to 8. So, see why we have focused here on type 1 error.

Now, suppose I have rejected a true null hypothesis. What happens if I have rejected a true null hypothesis suppose that the medicines that was actually producing 8ml actually filling up 8ml medicines only in the bottle, but then the sample results but I got we already saw. So when I discussed about the p value remember, if you forgot, then please, I suggest you to go back to those slides and see again, when I sometimes what have been a slight change for a slight change, we can have different results is not it?

So, maybe it is for such sample we got such a result for with my value is falling the rejection region, my value is falling in the rejection measured region means I am rejecting the null hypothesis. But what happens, but actually maybe my null hypothesis is actually true, that was a stray sample which I took and I got that value. So, what happens are I have rejected the null hypothesis.

When I rejected the null hypothesis, that means the machine is not working properly, I will have to change the whole machine maybe I have to invest crores of money for changing the machine or whatever spending lots of money, lots of time, energy, everything, which actually it is not necessary. So, it is easily costlier to incorrectly reject the status quo than it is to do the reverse. So that is the main thing why we focus on type 1 error.

(Refer Slide Time: 04:21)



So now, knowing this, so hypothesis testing, our first step is we specify the hypothesis null and alternate. And we specify the significance level, significance level is totally depends on the null hypothesis means significance level corresponds to the acceptable probability of the rejection of the true null hypothesis. So, now the question is, how do we understood why we had to focus on type 1 error.

So, now for a given application when we have to do hypothetical testing, so we will have to frame our hypothesis in such way, so that our null hypothesis is that which maintains the status quo. So, and when it is that which maintains the status quo and then our significance level should be such that we have to select the significance level such that, so, that we can even if there is some error there will not be much harm.

So, there as I told you depending on the situation the significance level is selected. So, when I also have mentioned when I talk a p value the decision maker will decide based on different factors what should be the rejection region. So, depending on the how costly it will be if I reject a true null hypothesis? How difficult it will be if I reject a true null hypothesis? Based on that level of difficulty based on that level of expenses.

I can decide my type 1 error probability that is my value of α suppose it is a very critical application. So, if I cannot manage to reject the null hypothesis for no reason if I cannot manage to reject a true null hypothesis very critical it is, because if I reject the true null unnecessarily I will spend crores of money then I will keep my type 1 error probability very, very less, maybe say 0.001.

So, however, maybe if some sort of some error I can manage, I do not need to bother much about the even if it is a type 1 error that will do not make much of a difference then I will keep my type 1 error probability a bit higher. Now, thing is that why should I keep it higher? Because already in my first lecture on statistical inferences, I have already seen which I will again talk more details on it was I have already seen when I am increasing α my β decreases that means, if I increase my type 1 error, my type 2 error decreases.

If I decrease my type 1 error, if I take my type 1 error very, very less value, my type 2 error will become very, very high, one increases the other decreases. So, for situation where the type 1 error is very, very critical, I will have a very less type 1 error probability, but at the same time my type 2 error will increase. So, since because of this relation, so, wherever the situation that I can bear a bit greater type 1 error probability then I will take a better and greater type 1 error probability so, that I get my type 2 error also reduces.

So, we will understand this with few examples. So, first consider this case assume that a 6 year old children would average about 10kg in weight and that is considered normal. So, what are the; suppose consider that a sample of children from a low income neighbourhood is to be tested for subnormal weight suppose in a slum area like this children are malnourished. So, government is trying to bring out some nutrition program it just in the first day.

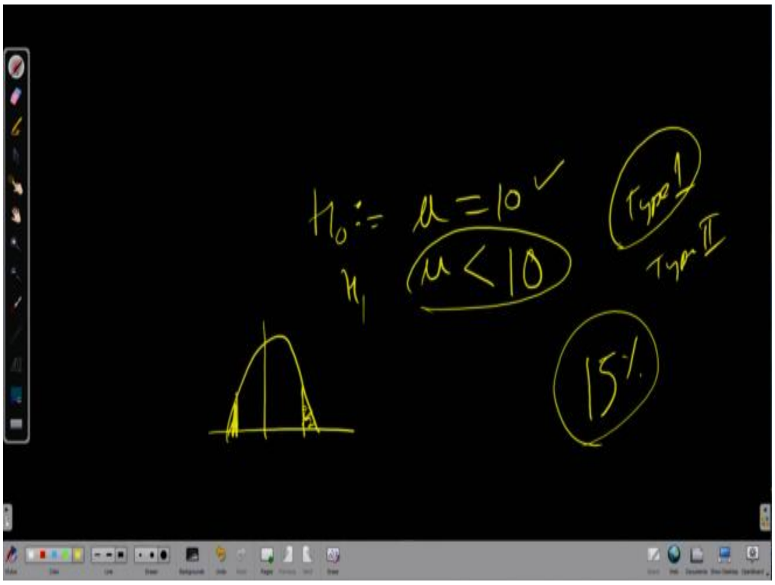
They will first find out whether the children are really malnourished if they are malnourished some program will be some meal some nourishing program government will start. If they find that the children are okay the children weight of 6 year old children on an average weight is 10kg then government need not start that program. So, now see here what will be my normal? What will be my null hypothesis?

What will be my alternate hypothesis what will be my significance level? So, here my null hypothesis definitely μ is 10kg that is I want to maintain definitely I will always want to maintain the status quo that will I will μ = 10kg but what I want to prove that μ is less than 10kg because if it is less than 10kg, then I will have to start some nutritional program for the children of that slum area. So, what I want to prove is less than 10kg.

Now here if I consider what to say μ = 10kg and I sorry let us not take this this way. For this example, what I want to prove is that μ greater than 10kg why? I want to prove μ greater than 10kg if I want to prove me greater than 10kg, then I will not have to start this nutritional program why it actually totally depending on the situation that different hypotheses are formed.

So, here maybe if my intention is if not required, I will not start a nutritional program in that case I will make the type alternate hypothesis is μ greater than 10 but see the significance here. So, that means when my null hypothesis is 10 if I reject a null hypothesis if it is true and I still reject the null hypothesis let me first write it.

(Refer Slide Time: 10:21)



So, this is that it is colon μ = 10. Then let us take that will be more easier, where μ is less than 10 this will fit more as an example this we are taking just an example to bring to the point how we select the significance level. So, here if I take μ = 10. So, if my null hypothesis is rejected, then what happens this hypothesis will be accepted H 1 will be accepted and the nutritional program will start.

So, if what happens if I do not if by mistake the null hypothesis is actually true. So, here, but null hypothesis is actually true, but I have rejected a true null hypothesis under what situation I have accepted it rejected a true null hypothesis that means, if it falls in my rejection region, null hypothesis is actually true, but the sample what I got from the sample I found that no null hypothesis and the value is falling the critical reason.

So, I will reject the null hypothesis then what happens a nutritional program will start an unnecessary expense to the government. But then you see the other way around what happens if the null hypothesis is not true actually, but I still accept it that means a type 2 error I got such a sample from the sample I found that null hypothesis is true only. But actually the null hypothesis is not true.

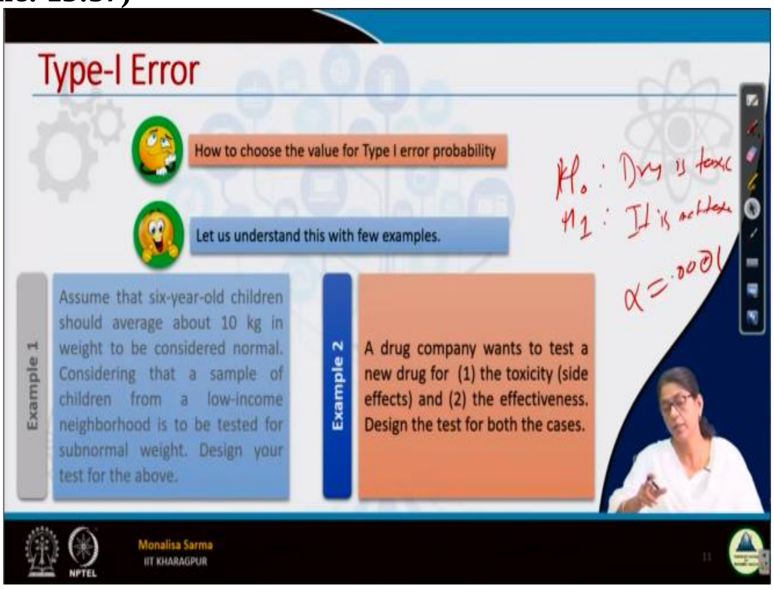
Actually the sample what I got that was a some outlier I got and I found that, the weight of the children is more equal or greater than 10. So, that means I do not have to reject the null hypothesis then what happens the nutritional program did not start now which will have more effect. So, if in the slum area if there is children even if they do not need it, if the nutrition program is started.

Government will spend some money that is all but this children definitely it will be good for the children only because as it is their poor children they will get to eat. So, what so, here if my type 1 error is not very serious, just that government has to spend some money, but if I commit type 2 error, what is type 2 error? Type 2 error means my null hypothesis is not true, but still I accept it, then it may be dangerous, the children are malnourished, still the government did not start the program.

So, it will have a dangerous effect on the children. So, here my type 2 error is very important understood the point. So, in such case hypothesis testing we always specify the type one error acceptable type one error. So, this is the case where type 1 error is not very critical. So, what I will do, I will specify a greater type 1 error so that I get a lesser type 2. So, maybe it is the in standard type 1 error is usually 5% or 1%, maybe I take some type 1 error of say 15%.

Poor children even if their weight is fine if the null hypothesis if and by mistake also we have rejected the null hypothesis, let the nutritional program start. So, I am ready to accept it 15% error probability understood so, now, this is 1 example.

(Refer Slide Time: 13:57)



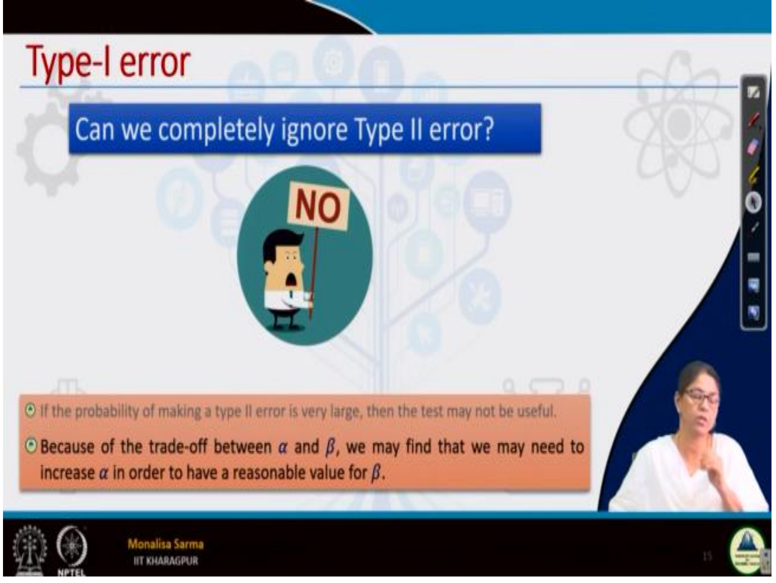
The next say the next example, a drug company wants to test a new drug for the toxicity it wants to test the new drug for toxicity see here. So, here what will be the null hypothesis will be a drug is toxic alternate hypothesis, because what I want to prove that it is not toxic. Here again, the as I told you, this formation of this hypothesis is totally depends on who is trying to do here the drug company itself is trying to do.

So, drug companies what to say its objective will be trying to prove that it is not toxic. But if some third party a third party is trying to prove third party will try to prove that it is toxic. So, then accordingly my hypothesis will be different. It totally depends on who forms the hypothesis and which is more important while forming that hypothesis. Alternate hypothesis is something which we always want to test that is the alternate hypothesis. So, my H 1 is it is not toxic.

So, now here the drug is toxic. So, here suppose here what happens if I by mistake if I reject the null hypothesis then what happened given the drug is toxic I have reject the null hypothesis then what happens the drug is then it is proved that drug is not toxic see the effect it will have the drug is toxic and you got your specifying the drug is not toxic. So, it will have a very, very dangerous effect.

So, in this case my type 1 error is very, very significant in such case maybe I will select my type 1 error will be say 0.0001 let β the more we will see even if the β is more it will not have much effect we will see what but here let β be more but we cannot manage a higher type 1 error because a higher type 1 error means it will result in a toxic drug turned as non toxic.

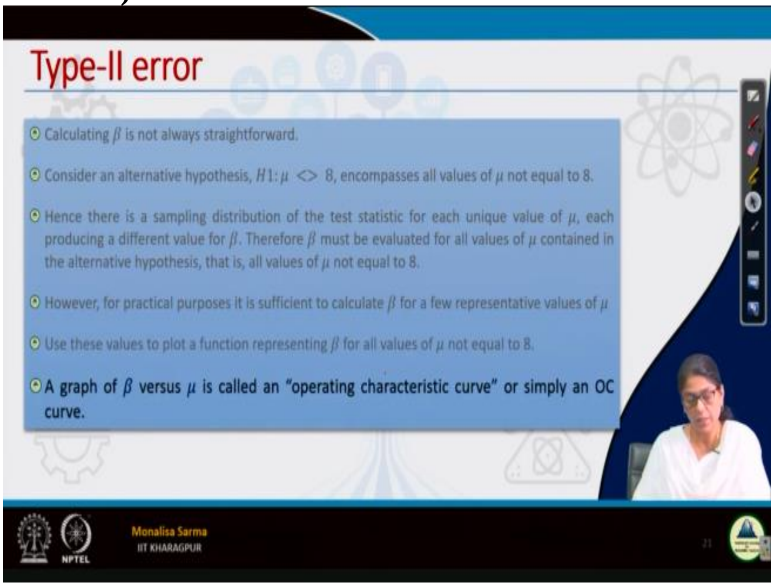
(Refer Slide Time: 16:15)



So, now, the question is can we completely ignore type 2 error? Now, in the second question see, if we have selected a very, very, very small type 1 error that means our type 2 error will become very big. So, that means is it that we can completely ignore type 2 error? No, we cannot ignore, but there is a twist to it, we will see how that is what if the probability of making a type 2 error is very less than this may not be useful also.

So, we will see, because the trade off between α and β, we may find that we may need to increase α in order to have a reasonable values of β. So, that totally depends on the application what α will keep, which is more important type 1 or type 2.

(Refer Slide Time: 16:54)



Now, calculating β remember in the first lecture, I told 2 things that we will be discussing one is how to calculate β though for the toy example that is box of chocolate I calculated β but this is a toy example that is not applicable for any that is just to understand the concept, but calculating β is not a very easy job it is not straightforward. That is so that is the reason I did not show, how to calculate β over there.

Now, you have enough knowledge. So, now we can see how to calculate β. Secondly, I also mentioned one thing and that lecture itself the null hypothesis always should be with equal sign that why I will come here thirdly, why we were to say focus more on type 1 error that we have seen here. Now, how to calculate β and why null hypothesis we need an equal sign that we will see here.

So, see any sort of hypothesis testing my backbone is the; my main CPU I should say my CPU is the sampling distribution everything I put data I give to the simple CPU and the CPU gives me the results, is not it? So, every data I give to the sampling distribution from the sampling distribution, whatever result I get based on the result only I give the decision. So, for sampling, when I have the sampling distribution, what happened my sampling distribution means that is one distribution which is normally distributed.

And which has a mean and its mean is what its mean is equals to the population mean and it has is variance is population mean divided by population variance divided by sample size that is the variance, forget about the variance. Now, the sampling distribution, it mean is equals to the population mean. So, if my hypothesis if always I try to find out the data based on my null hypothesis what we have a hypothesis that is $\mu = 8$.

So, when I find out the sampling distribution in my sampling distribution, so, my mean is 8. So, if I do not have an equality sign, then I cannot have 1 sampling distribution, I will have to have many distribution with differs each values of μ if I write μ not equals to suppose that medicine example that is mean is equals to 8 mean not equals 8, it means not equals to 8 means I will have to have distribution for all different all possible values of 8 is not it? All possible values of 8 means 7.9.

For 7.9 that means, I will have to have a distribution μ 7.9 and variance whatever it is, then whatever suppose 8.1, I will have to have a value for distribution for 8.1. So that way, I will have to find out the α values for all different values of μ. That was the reason why to simplify the process. My null hypothesis always need to have an equal sign so that we can have 1 sampling distribution which mean is equals to the population mean.

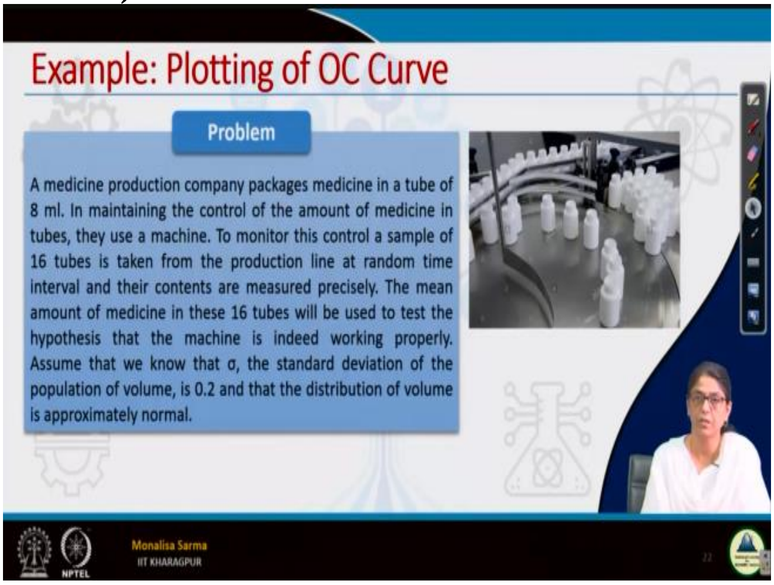
Now, when we are trying to find out the type 2 error for finding out the type 2 parameters, we have to find out the rejection regions or basically type 2 error means, we are not rejecting a true null hypothesis means, it is false but it is falling in the acceptance region. So, definitely we need to have a sampling distribution. So, when we have a sampling distribution we need to have the mean of the distribution.

So, when the value is not equals to 8 what will be the value? So, that means, we will have different distribution for different values of μ that is why calculating β is not always straightforward that was a toy example that box of chocolates so, we could do it very easily. So, here so, hence the sampling distribution hence, there is a sampling distribution of that for each unique value of μ for each unique value of μ there will be 1 sampling distribution each producing a different value of β.

Therefore, β must be evaluated for all values of μ containing the alternate hypothesis that is all values of μ not equals to 8 we will have to take μ not equal to 8 there can be any value in final values. So, for all values of μ not equals to 8 we will have to find out β. So, however, for practical purpose we will see it is sufficient to calculate β for a few represented values of μ if we take some few representative values of μ.

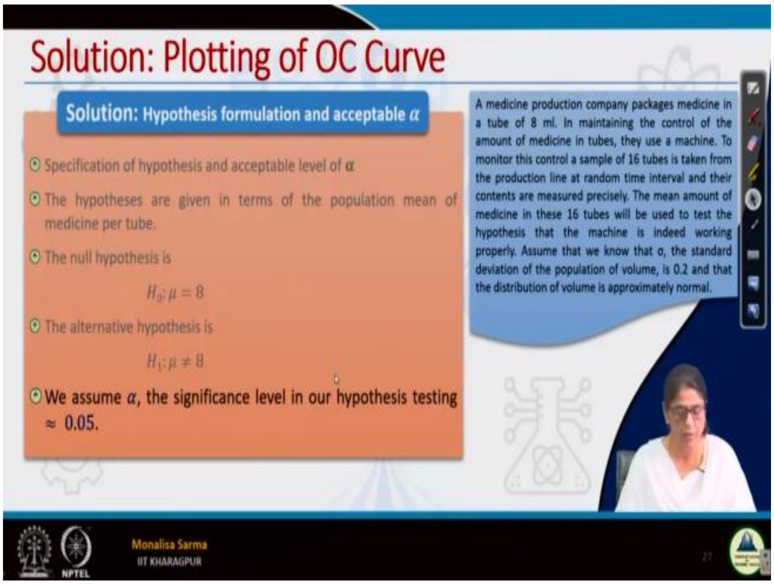
Then it is sufficient to find out the value of β then this represent the values of β for this values of μ different values of μ what β we get, we can plot this and what we call this this plotting is nothing but it is called operating characteristic curve. So, we will see what is an operating characteristic curve.

(Refer Slide Time: 21:48)



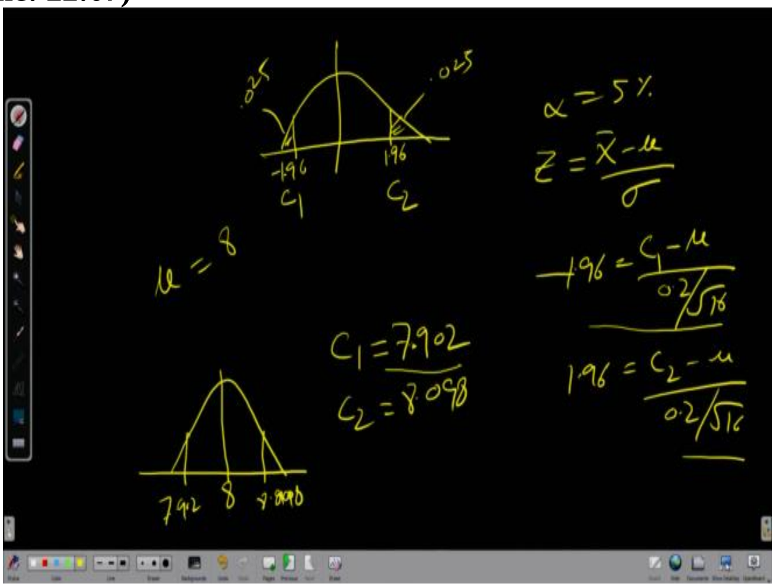
So, we have again taken the same example to plot the operating characteristic curve.

(Refer Slide Time: 21:55)



So, here you see what was our hypothesis, hypothesis is μ = 8 alternate hypothesis is μ not equals to 8 significance level we have considered 5% significance level only.

(Refer Slide Time: 22:07)



So, here when I have found out the significance level, the see very carefully z is equals sorry significance level α = 5%, when α = 5% means, what if this is my what to say distribution that means, this bellow this area this area is 0.525 and this area is 0.025. Then I found out what is the z value corresponding to this z value corresponding to this. So, I think you remember it was 0.96 and - 1.96 this is in terms of z, is not it?

But I need the value in terms of to find out the rejection region I may also need to value in terms of the actual parameter value random variable what does my random variable my random variable was x based on that I found out x bar, is not it? So, this z I got value converting from x to z how are we what is the z value that is equals to x bar - μ / σ is not it? This is how we calculate the z value.

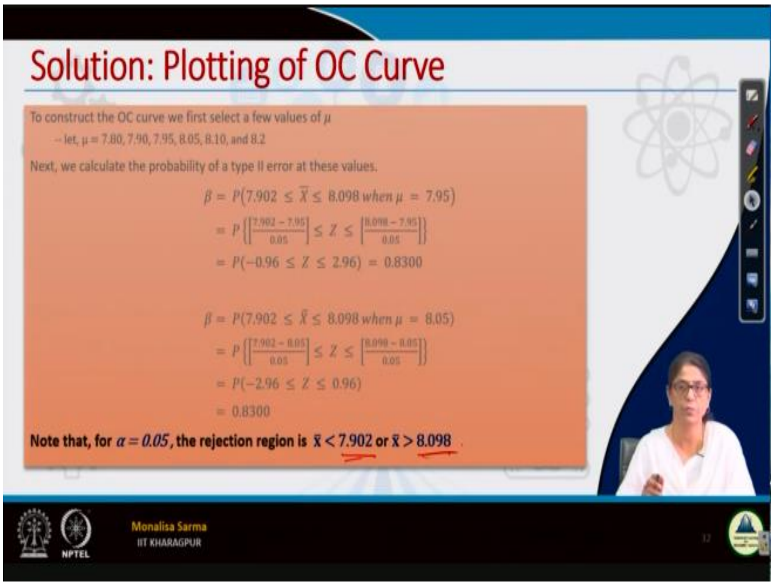
Now, this let me tell this as this counterpart in the as a random variable. So, this is c 1 this is c 2 then what is the z value - 1.96 = c 1 - μ is 8 and what was σ? Σ was in this example is I think 0.2 divided by 16. We have taken a sample of size 16 and standard deviation is 0.2. So, similarly again 1.96 = c 2 - μ 0.2 / √ 16. So, we calculate the 3 we are from here we calculate the value of c one from here we calculate the value of c 2.

So, c 1 we get if I can simplify it c 1 will get I think 7.902 I remember correctly and c 2 will we will be getting some 8.0 90 so, my rejection region is μ = 8, so, my rejection region will start from 7.902 to 8.908. That means, here I will draw I am drawing the graph again. So, if this is the thing, this is my rejection region 7.902 and this is 8.098. So, this is my rejection region here this is 8 that means acceptance region μ = 8.

Remember why what is this acceptance is I am repeating it again μ = 8. So, μ = 8 means the under what situation I will accept it μ is 8 but still a bit of variance I can consider that so, that was my logic remember bit of variance I can consider. So, how much variance I can consider maximum and this side it is 7.9 and this side it is 8.1 maximum this much variants I can consider, is not it?

That is what so, basically that is I have just told 7.9 or 8.1, but when I specify the significance level from the significance level I could find out what is my this critical region or start of the rejection region.

(Refer Slide Time: 25:26)



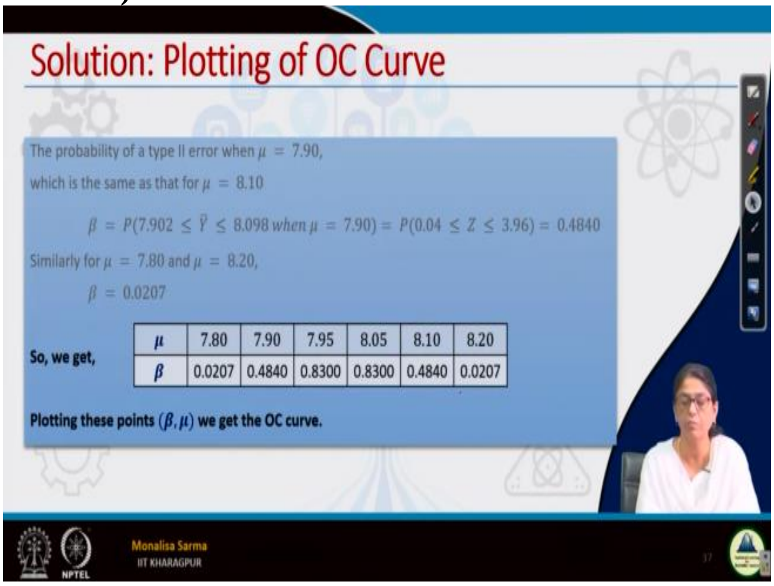
So, now here so, now, we will have to find out the value of β for the different values of μ, because μ is not equal 8. So, different value we have taken some representative value. So, what we have taken we have taken see here 7.80. And this side we have taken 8.20 we have taken equal what to say symmetric value both the side. So, we have taken 7.90 have taken 8.1, 7.95, 8.05.

So, now, we have to calculate the type 2probability, type 2 probability means, actually, my null hypothesis is false, but I am accepting it that means it should move on when a null hypothesis is false, but my value is falling in the acceptance region that means μ value is falling between 7.9 to 8.1, but actually it is false. So, that is my type 2 error. So, β is a probability that my X bar value falls within this range when μ is actually my mean with 7.95 μ is not a 8.

But when my μ is 7.995 my X bar is falling within this range that is my β. So, what is this corresponding to this X bar value, corresponding to this X bar value what is that value I can find out so, this is the Z value. So, what is the probability of β is 0.8300 from the Z table I can find out and I can remember how do we find out this and again repeating it we got it from here this portion then this portion.

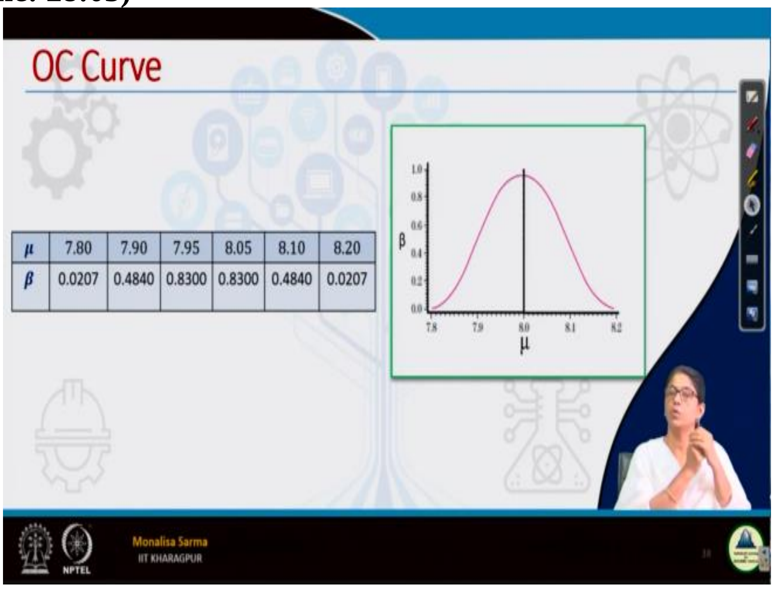
So, then this portion minus this portion will give me this portion. So, this we found out for musical 7.95 my β value is 0.8300. Similarly, I found out for μ = 8.05 what is my β value, I got the same thing, same β value, it is symmetric. So, considering this is a 7.902 and 8.098 is the rejection region that is what I have just now shown why this is the reduction in corresponding to minus 1.96 and + 1.9645% significance level we got this as the x value.

(Refer Slide Time: 27:35)



So, now, we will consider for μ = 7.10 and 8.10 we will see that if we calculate we will get this value 0.4840. Similarly, if we calculate for 7.80 and 8.20 we will get this this β value. So, for this represented values of μ different values of μ this is the these are the values of β we got. Now, we can plot this value.

(Refer Slide Time: 28:03)



So, if we plot this value, this is the this curve is called operating characteristic curve operating characteristic curve means, where we are trying to plot β for different μ what is β for different μ, where μ is not illustrated, you see here the beauty of this curve you see, see when my μ actual μ that means, population mean when my mean is very close to the hypothesized value, then what happens then my type 2 error is very high.

It in fact β in fact, a process 1 - α, whatever α, we specified β process 1 - α, when my actual value approaches goes very near to the hypothesized value. So, see what happens when my actual value is very near to the hypothesized value, then my type 2 error is very high meaning what means I am accepting a false null hypothesis. But you see when my value is not very different, but very near to the hypothesis value.

If I am even if I am accepting a false null hypothesis, then also it will not have a very serious effect that is one thing. Second point is that already I told you that when my null hypothesis is not rejected, that does not mean that I am accepting the alternate hypothesis that does not mean that when a null hypothesis is not rejected, does not mean that I am accepting the null hypothesis. It may mean 2 things.

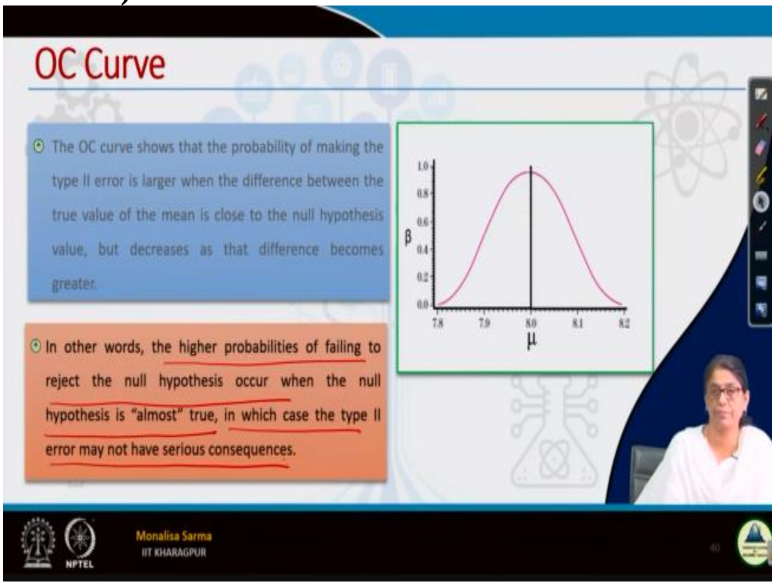
It may mean I am accepting the null hypothesis. It may mean I am I am not able to reject the null hypothesis. So, if I am accepting the null hypothesis, that means I am complicit I am okay fine, my null hypothesis I have accepted, but I am not being able to reject the null hypothesis means, again, I will do in future again, I will do the experiment and again, I will see whether really, my machine is really working fine or not.

I will again check put on my check whether my null hypothesis is accepted or rejected, I will maybe I will continue this experiment for 2 times 3 times 4 times. And if it is really, if my population value is actually a really away from 8, the time will come when I will see that actually it is it is falling in a critical region and a null hypothesis will be rejected. So that is what when my even if I have a very high type 2 probability.

Since my value is very close to μ, it will not have a serious effect that is the first point. Second point on if the μ is really different after the 1, 2 experiment, it will definitely come to the conclusion that null hypothesis actually rejected when will when I will take one sample again after sometimes I will take another sample for some sample maybe bias all sample will not be bias.

So, I may gradually see that my null hypothesis is rejected that is the first point now, when my actual value comes very much away from the null hypothesis, then type 2 error is very less means that times accepting a false null hypothesis is very less and that is good that is required, is not it? That means when I require my type 2 error to be very less, I am getting that when it is very much away from the hypothesis below my type 2 error is very less.

(Refer Slide Time: 31:46)



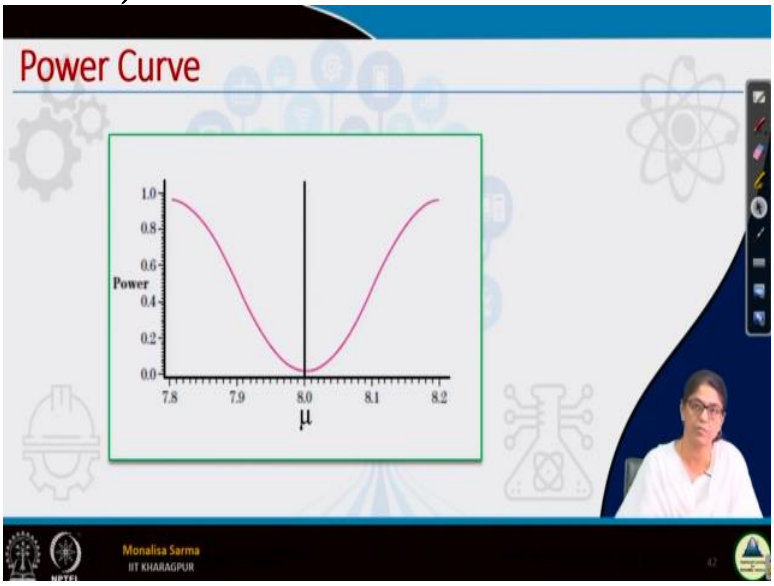
So, the OC curve shows that the probability of making the type 2 error is larger when the difference between the true value of the mean is close to the null hypothesis value, but decreases as the difference becomes greater. In other words, the higher probabilities of failing to reject the null hypothesis occur when the null hypothesis is almost true, in which case the type 2 error may not have serious consequences.

(Refer Slide Time: 32:18)

Now there is one more concept that is the last concept in this lecture, we shall be discussing that is power. Power is nothing but it is β is type 2 error, power is nothing but 1 - β. So, basically as a practical method we are easily more interested in the probability of not making a type 2 error, practically if we see we will be more interested in making not making type 2 error. That is the probability of correctly rejecting the null hypothesis when it is false.

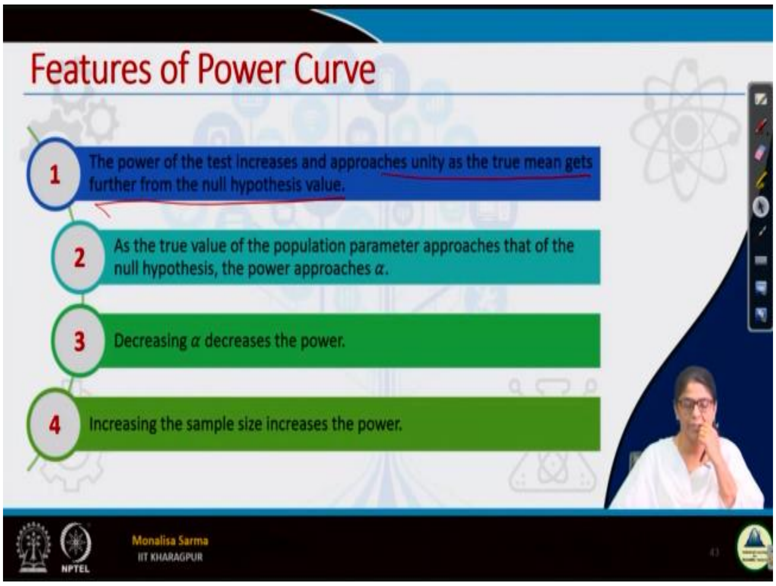
Not making the type 2 error, not making the error means when the null hypothesis is false we are correctly rejecting it, instead of accepting it, that is my power. The power of the test is the probability of correctly rejecting the null hypothesis when it is false, we define as a power of a test. So, power of test is nothing but 1 - β, β is accepting a false hypothesis so 1 - β is rejecting a false hypothesis. The graph of power versus all values of μ is called a power curve. This whatever OC curve we got is just a power curve is just an opposite of that.

(Refer Slide Time: 33:22)



So, this is the power curve OC curve is β versus μ and what is power curve? 1 - β versus μ. So, my power is very less when my actual value is very closer to the null hypothesis value. When my actual value goes far away from my null hypothesis value my power is very high that is desirable.

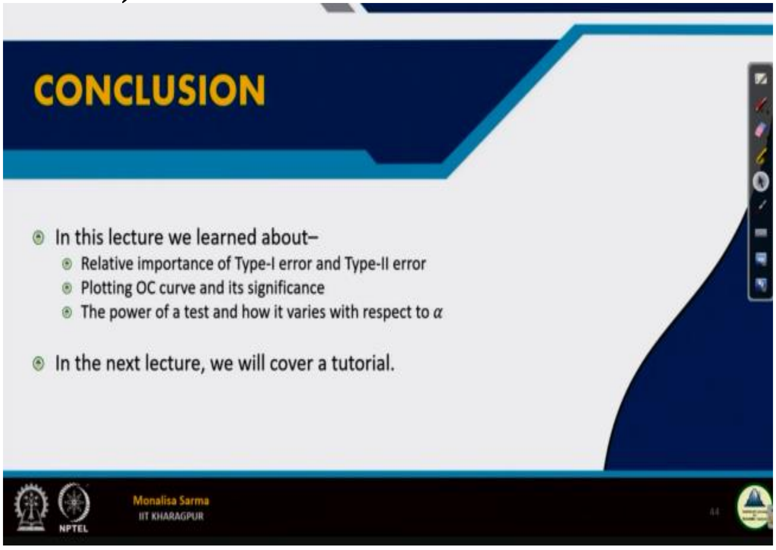
(Refer Slide Time: 33:51)



So, the power of the test increases and approaches unity as the true mean gets further from the null hypothesis value. As the true value of the population parameter approaches that of the null hypothesis, the power approaches α. When the true value of the population it approaches near the null hypothesis value my power becomes very less, it approaches α that is it approaches to the value of type 1 error.

Decreasing α decreases the power why? When it decreases α β increases when β increases 1 - β decreases is not it? So, decreasing α decreases the power. However, increasing the sample size increases the power.

(Refer Slide Time: 34:33)



So, to conclude this in this lecture today we have learnt couple of concepts we have learned relative importance of type 1 error type 2, error how we plot OC curve and what is its significance; the power of a test and how it varies with respect to α. In the next lecture we will cover a tutorial encompassing all 3 lectures what we have done in statistical inference. We will move to statistical inference we shall be discussing.

Before you forget all the things what we have discussed let us see the tutorial first and then we will be going to do lectures other lectures other topics on statistical inference.

(Refer Slide Time: 35:07)

So, these are my references and thank you guys.