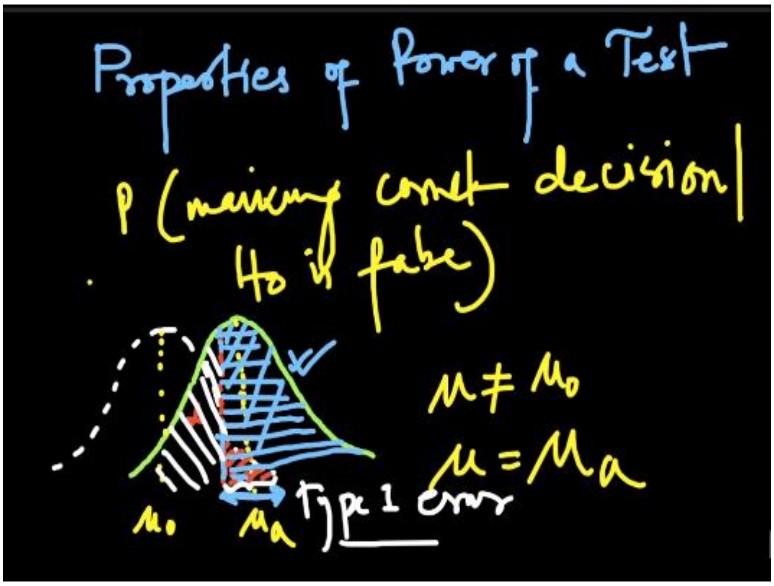
Applied Econometrics
Prof. Tutan Ahmed
Vinod Gupta School of Management
Indian Institute of Technology-Kharagpur

Lecture - 39
Properties of Power of a Test

Hello and welcome back to the lecture on Applied Econometrics.

(Refer Slide Time: 03:00)



In this lecture we are going to talk about power of a test in detail along with the properties. And power of a test is a very important concept when it comes to hypothesis testing and basically designing some research, okay. Now what we understood by power, we already have explained what we mean by power. I am just writing it again.

So essentially, it is probability of making correct decision, making correct decision when H naught is false, right where my null hypothesis is false, when H naught is false. Now how do I actually represent it diagrammatic, okay? So let me just try to do that. Now when I say H naught is false. So let us say we have two distribution. One is the distribution for the null hypothesis. And then we will have a true distribution, right?

The alternative hypothesis is correct here. So let us say I draw it with a solid line. This is my other distribution, the alternative hypothesis. And here I had my, let me use a different color again. Here I had my population mean as proposed by the H naught and here I had the population mean as per the alternative hypothesis H a. So I would write it as let us say mu a okay. All right.

Now in this case, the true population parameter mu is not equal to mu 0. Whereas, it is equal to mu a, right. And with this, this is you know where we actually in this scenario, when we actually make the correct decision, we call it the power, right. Now if we have let us say, if we in this distribution in the null hypothesis of distribution, let us say I have the level of significance alpha let us say, is listed here, whatever that alpha value is.

Now if I, when my H naught is false, when my H naught is false, so this part of the distribution, this part is saying something correct, right? Because it is saying that the H naught is false. But unfortunately this is conditional on the wrong distribution, like the distribution corresponding to the null hypothesis. This is conditional on the wrong distribution.

So we want it to be conditional on the true distribution, and if we sort of draw this line towards it, we draw this line further, what do you get is a different color, what I get here. Now I am basically conditioning my state sorry my statement on the true distribution, right? Now I am saying that H naught is false and I made a correct decision, right? My H naught is false.

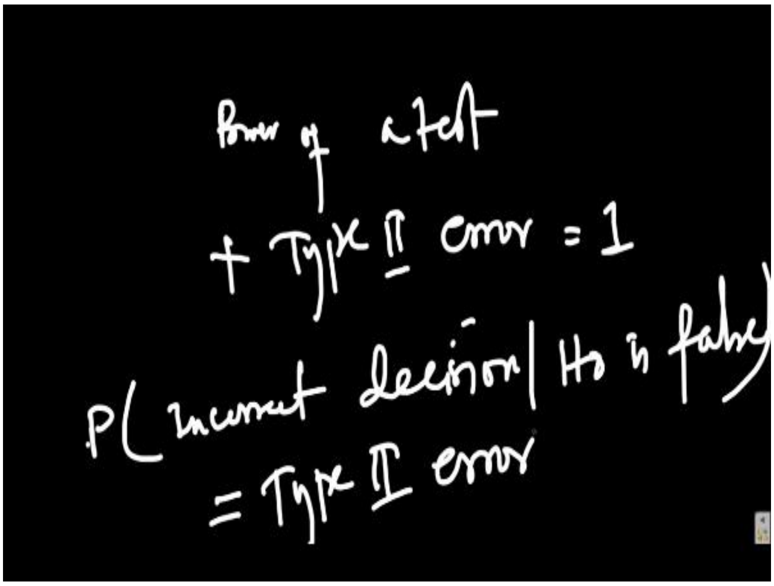
So it has to be, when H naught is false, it has to be for the first distribution, it has to be here. Now when it has to be here, if I just correspond that for the second distribution, then I will get this area, okay. So this is something we have to understand. So this as it is a sliding bar here, okay? And I will explain the importance of the sliding bar. And this one is basically the power of the test, okay.

This one is basically the power of the test. All right. So it is conditional on the true distribution. And it is also stating that H naught is false, right? And it is also for the first distribution, this part is also taken care of. So it is actually falling in this region where it has to fall, right? Okay, so now we got it. Now we have to understand some relationship.

We did explain the relationship before but just for a recap purpose we have, now this part, this part is my type I error, we know that this part is my type I error, right? If my actually null hypothesis is true, and I actually make the mistake, then that will constitute type I error. Okay, so we got type 1 error and we got what is the power. Now look at it.

This for this for the true distribution we have, of course we have seen the blue part, which is the power. And then there is the other empty part and I am going to fill it up with a white color. And this looks like this. So this is basically the type II error, right? So this part is type II error and we know that power of a test, let me use a different page.

(Refer Slide Time: 06:14)



We know that power of a test plus type II error would lead to a probability is equal to 1, okay. Now what is the type II error again? Like we have constructed this, but what is the type II error? So essentially type II error says that we make an incorrect decision, probability of incorrect decision when H naught is false, so when H a is true okay, incorrect decision when let us say H naught is false, okay.

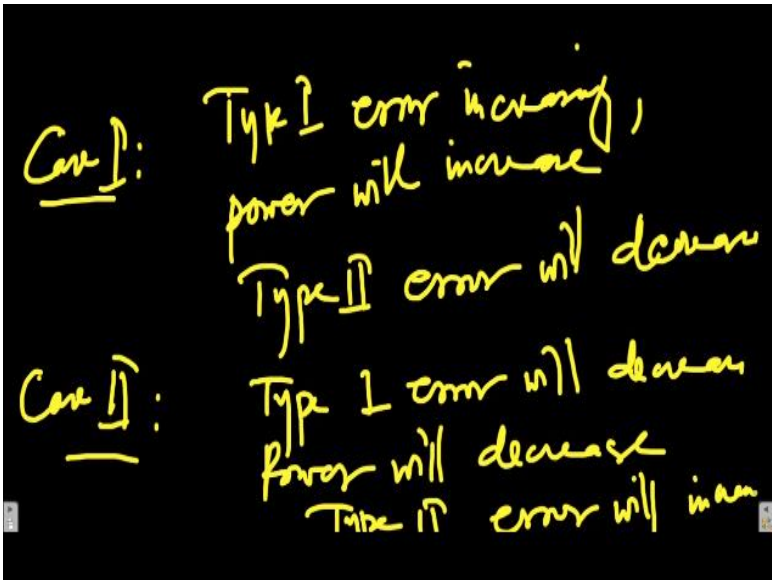
So when H naught is false, so we have basically my observations are conditioned on the second distribution, the green distribution. But then I am making an incorrect, this is type II error, all right. So type II error. So we have seen in this lecture that we have seen power of a test, we have seen type I error, we have seen type II error.

Now we need to sort of, you know in our mental math, we have to kind of see how these three things look, go together. So here, what is most important is that this line here, this line that we draw, okay? The sliding bar, let us say, the moment sorry, the moment the bar, the moment, the moment the bar, somehow I am not able to align this line, but I do not think we need it.

So the moment we, the moment we actually push the bar to this side, so what is happening here? So the moment we push the bar to the side, I am actually increasing the type I error right, the type I error is increased, because I pushed the bar in this side. So what happened because of that? The power also got increased interestingly.

The power got increased, the type I error got increased, and of course type II error got decreased. On the other hand, if I move this slide, this sliding bar to this side what is happening is that the power actually is declining, decreasing, right? And the type II error is increasing, but at the same time type I error is decreasing.

(Refer Slide Time: 08:54)

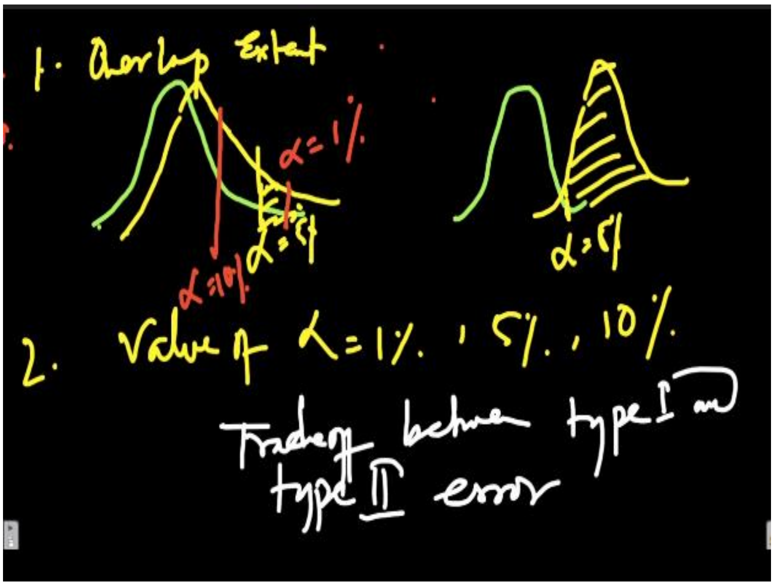


So if I write it down, so what the relationship would look like, so let us say type I error, if it is increasing, so essentially, it means I am sliding the bar on the left. So that your mean, so basically let me write down case I. So I have a case where I can have, I increase the type I error by sliding the bar on the left. And then it will also correspond with the power will increase. And type II error will decrease.

And just the reverse for case II. If I slide the bar on the right side, case II, if I slide the bar on the right side what will happen? Type I error will decrease, power will decrease and type II error, what will happen to type II error? Type II error will increase, right? Type II error will increase, okay. So that is basically how these three concepts are related.

Now let us now actually take a dip into the properties of the power, okay. Let us say, we sort of do a mental exercise here. And let us say I have, again, two distributions, okay.

(Refer Slide Time: 10:44)



So let us say I have a distribution like this, I am not going to be very accurate. And then I have another distribution, let us say it is pretty much overlapping, alright. And then let us say it is not a normal distribution, but just let us assume that this is a normal distribution that I have here. And another distribution I have here.

So the difference that we have for the first case, and second case is that in the first case, the distributions are pretty much overlapping. Whereas in the second case, they are not overlapping, right? So what will happen here? So let us just try to see what happens. Let us say, if I have for this case, if I actually have some alpha, whatever that alpha is, let us say 5%, or whatever.

So essentially, I am ending up getting a very low power, right? I mean, you are getting a very low power here, very low power. Whereas in this case, let us say the distribute is alpha is here, whatever this 5%, alpha is 5%. But here because they are not overlapping, what is happening here I am getting a very high power, right? So ideally, I will want these distributions to be non-overlapping, as non-overlapping as possible.

Because only then it is possible to a level of significance alpha. Whereas for non-overlapping distributions will have a high sort of power right, vis-à-vis the overlapping distribution, okay. Now what will happen? Now let us, this is the case I, let us say. I will basically talk about these different cases. There is a overlap okay, extend of overlap, overlap extend. In the second case, I will write down, let us say, the value of alpha.

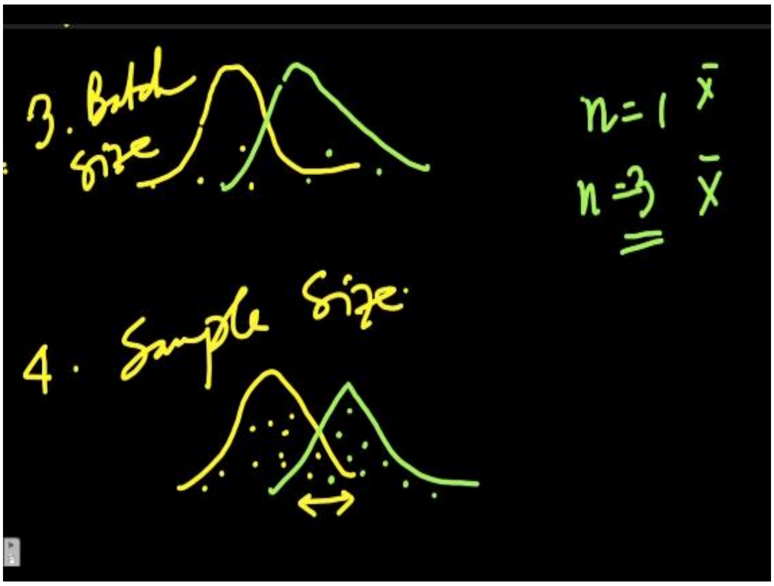
Value of alpha, let us say it could be 1%, it could be 5%, or it could be 10%. And I will actually ask you to think. Now by now we should be able to think if I have my alpha is equal to 10%, what will happen? What will happen if my alpha is equal to 10%? Well, so if my alpha is equal to 10%, basically what is going to happen is that the sliding bar is actually going to move towards the left, right? Alpha is 10%.

So what is going to happen? I am going to get, I am going to get a bigger power, right? I am going to get a bigger power. My alpha on the other hand is 1%. Let me write down in the top. What is going to happen? Well, since what I have done, essentially I have you know sort of pushed the sliding bar towards right and that means I have a very low type I error, high type II error and a very lower power, right?

And this is the two case; one here, this one here, as well as case two here. Case two, okay. So in both cases, it will apply, right. Alright, now the other point, let us say, so okay. So here one you know simple question that often we actually like to ask is that the trade off, the tradeoff of between type I and type II error, okay. And by this, I mean with this explanation, it should be clear how this tradeoff is happening, okay. All right.

Now let us talk about another case of sample size. And I think you should be able to guess this, because we always know more sample size is better. But let me explain that with a diagram how it is really happening.

(Refer Slide Time: 14:26)



Let us say I have this two distribution again. Here is the two distribution. And let us say I have the best size or in I mean, I draw like in every draw let us say in one case I draw 1. And in the other case, I draw n equal to 3, okay. Now what will happen if n equal to 1 and n equal to 3. Now if I draw n equal to 1, so what I will have? So I will have one observation here, let us say.

Then the other observation is going to be here. And the other observation let us say is going to be here. So variance is going to be pretty high, right? But whereas if I have n is equal to 3, so what will happen? I am ending up, let us say drawing three observations from here. So essentially, they are cancelling out, they are cancelling out their positive and negative biases.

And on that same vein, when I draw observations from here, they are cancelling out their positive and negative biases. So what is happening here, I am actually getting a low variance case. And I will be able, so I will be able to sort of get the resultant sample mean, the x bar values. I will be able to get the resultant sample mean, actually going to be close to the population mean.

So when you do that, for both the distributions, we have the resultant observations we are getting, they are actually pretty close to the corresponding population mean. And because of that, we are actually going to get a good result. All right. So it is like, let us say this is the case of batch size okay, batch size. So a good batch size is always better, right? Now let us talk about the sample size, okay.

So I basically I draw you know basically increase the number of samples. I am not talking about the batch size, whatever is the batch size. I am talking about the sample size. And here, what I will get, let us say, I will again explain this. Very poor normal distribution, but let us proceed with this. What is happening here is that when we have this kind of distribution, so we have let us say, we get some observation, okay.

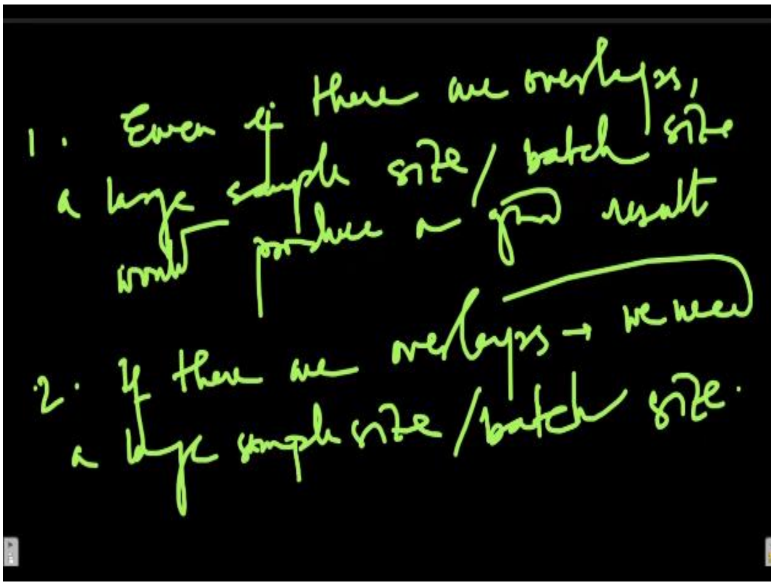
If my sample size is low, so let us say I have some observation of this distribution here and some observation of this distribution here, okay. And it is difficult for us to actually distinguish. It can actually you know the yellow dots would be here and the blue dots would be here and green dots would be here on the extremes. But that is not necessarily that is going to happen.

So when we have small sample size, we have more sort of uncertainty involved, okay. But whereas, if we have large samples, large sample size, so what will happen is, we will have observation from all these different points. And similarly, we will have observation from all these different points. And when we do the, we sort of estimate their statistic, so we are going to more likely to get the true statistic okay.

Then essentially, why is it important? Now essentially, this batch size and sample size, they are you know talking about the overlap, right? So if I have a sort of big sample size or I have a large batch size, so then even if there is overlap, I can actually take that you know handle that with a large batch size or with a large sample size. But if there is overlap and my batch size or sample size is small, then there is a problem, right?

And then the, you know sort of estimation of power would be a problem, okay. So we will see that, more about it in detail.

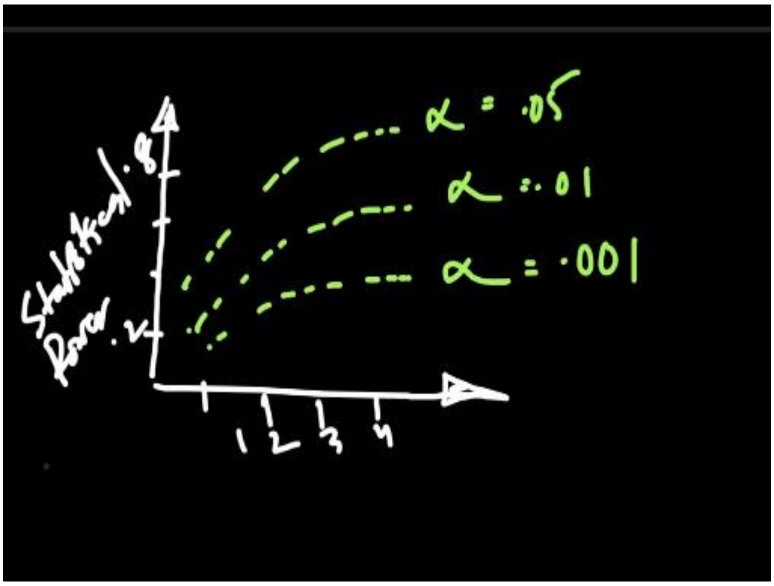
(Refer Slide Time: 18:24)



So essentially, we have to even if there are overlaps, a large sample size or batch size good result, okay. It will produce a good result and other inference could be if there are overlaps, so basically then that means if there are overlaps, if there are overlaps, so it automatically means, it means we need a large sample size, batch size, okay. If we do not, then the results will not be good.

The results will not be good. All right. Okay, now let me conclude this lecture with a couple of more than and let me draw them here.

(Refer Slide Time: 19:53)



So let us actually try to plot, let us say statistical power and we will just see how we can actually conceive these ideas. So this is my statistical power, okay. And this is let us say 0.2, 0.4, 0.6 and let us say 0.8, okay. And then I have my sample size or whatever, you can take by size it does not matter. Let us say I am taking batch size 1, 2, 3, 4. Now if I plot the level of confidence or level of significance alpha, so what I get is, I will get something like this.

Alpha is equal to 0.05. Alpha, this is going to give me alpha is equal to 0.01. And let us say another line will be if I want to draw alpha is equal to 0.001. Now how we can really make sense of this? So it shows that as we increase the sample size, as we increase the sample size for a given or batch size for that matter, for a given alpha, we are actually increasing the statistical power, right?

So now if we, so that makes sense, right. And if we actually decrease the value of alpha, so what is happening, we have already seen that. By decreasing the value of alpha, we are actually, you know sort of sliding the bar towards the right, sliding the bar towards the right. And if we slide the bar towards the right, naturally I will have less power, okay. So that is basically the idea, okay.

So with this, I will end the lecture here. And we will again, there is something very important, that we will follow up from here, and that is minimum detectable effect. And minimum detectable effect is something that is really crucial when it comes to any sort of experiment design that we do with randomized control trial, particularly, it is really important when you have to basically understand.

We will talk about minimum detectable effect not now but later when we talk about all these in the third module of experiment design. We will talk about causality, then we will, you know sort of get back to the idea of minimum detectable effect, and we will explain that in details. With that we end the lecture here. Thank you.