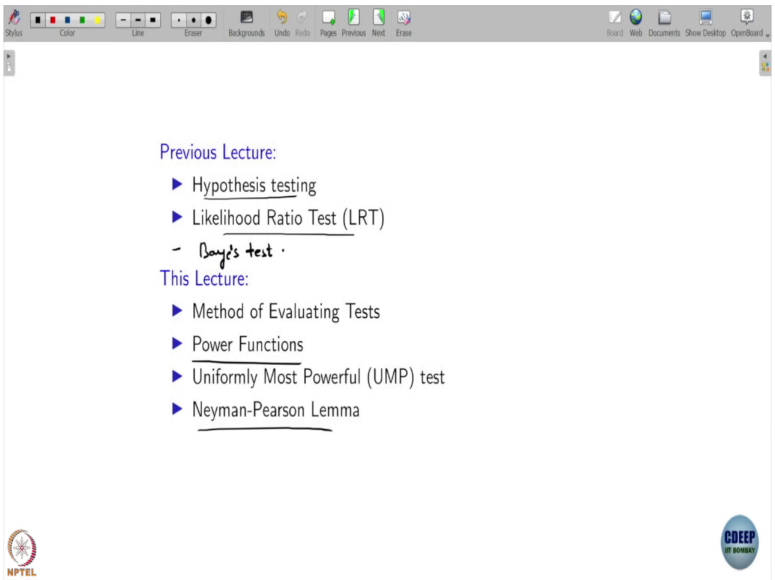
Engineering Statistics
Professor Manjesh Hanawal
Department of Industrial Engineering and Operations Research
Indian Institute of Technology, Bombay
Lecture 43
Type I and II errors, Power Functions

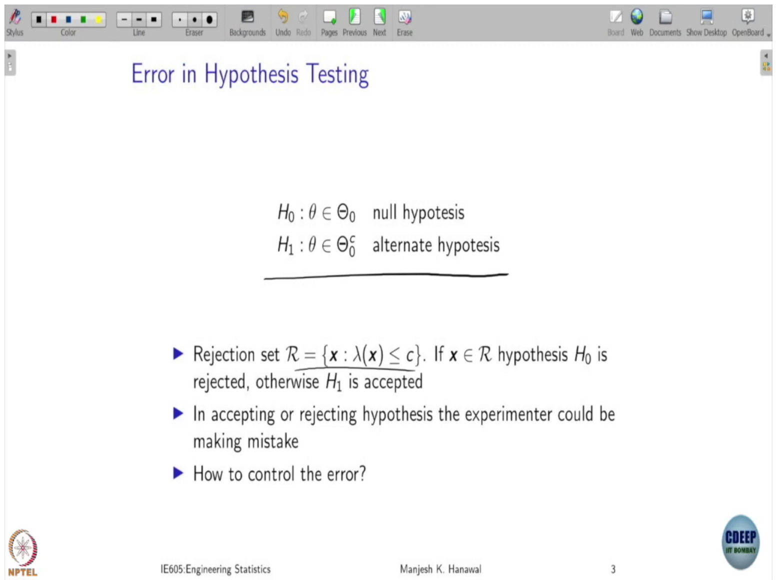
(Refer Slide Time: 00:22)

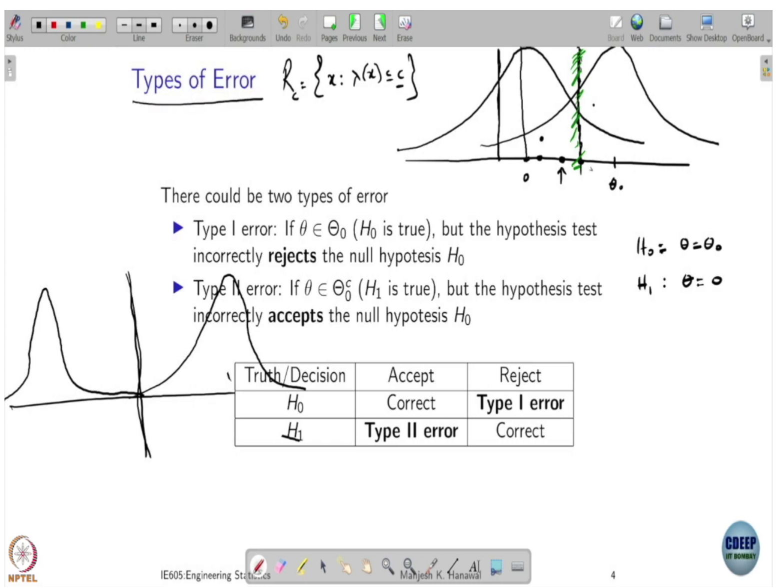


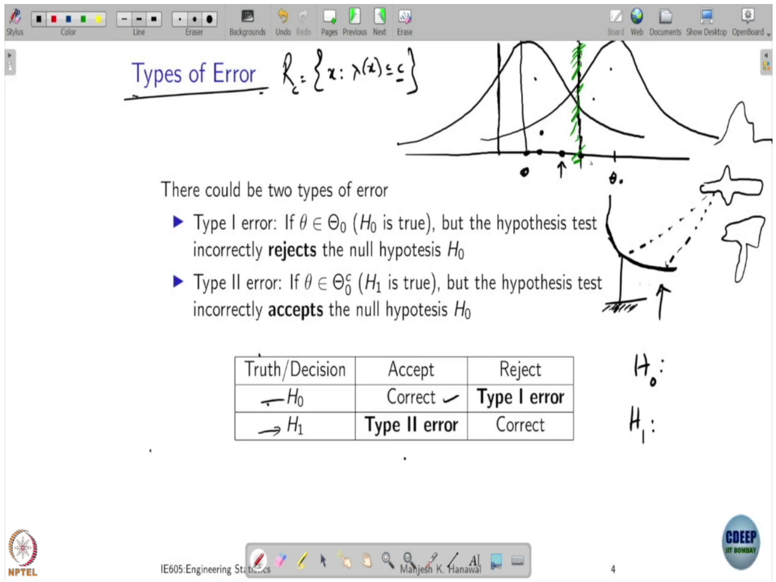
So, we just started discussing about this hypothesis testing, we covered likelihood function, likelihood ratio test, we actually also covered Bayes tests. And now, again what is a way to evaluate which test is good? We have a method of based on LRT, we have a based on, method based on Bayes, if you have multiple options given to you, which one to take? So, in the case of estimators we had an criteria, we had say that given multiple unbiased estimator, which one you would have preferred? We said among the ones which gives me the smaller variance is better, if all the estimators have the same, all are unbiased that when their bias is 0, then the only criteria you will look is something which has lower variance. And Cramer-Rao bond, what is the smallest variance one can anticipate.

Now similarly, there should be method, if you are given multiple tests like this, to accept or reject a hypothesis which one is better? And how to evaluate them? So, that is where we are going to develop some method there is something called power functions, and based on that we will go into define some uniform most powerful test. And some statement on Neyman-Pearson lemma, we make, and I do not know we will may not have much time to discuss this Neyman-Pearson lemma, but we will just state it.

Refer Slide Time: 02:16)







So, we said, when you are given a hypothesis like this, and we have a rejected region. In accepting or rejecting the hypothesis, can there be errors? There can be errors, and then the question is how to control those errors? Let me consider these two, I am going to let us say this is a 0 here, I have one Gaussian distribution with the mean theta here, and another Gaussian distribution with mean 0. Now, I have only two hypothesis null hypothesis is theta is equals to theta 0, and the H1 is theta equals to 0, so there are. My samples can come only possibly from two hypothesis, either with mean theta 0, or mu, I do not know which is.

Now, let us focus on our likelihood ratio test. My rejection region is what? All x such that lambda of x is less than or equals to c. Depending on the c value, let us say we will end up with this decision boundary, these are boundary, what is going to say that all the samples that are above this are going to be accepted as belonging to null hypothesis, and all the samples that are going to fall below this point here. So let me make it a little bit bold, will be rejected, to belong to alternate hypothesis.

But if you draw a sample from this Gaussian distribution with parameter theta not, it may end up here, you may get a sample like here, you may get a sample here. Because Gaussian is covering the entire region, when you the sample actually could drawn can be here, in that case will it be, if this sample going to be get accepted as a null hypothesis or alternate hypothesis. It is going to be accepted alternate hypothesis, and not as a null hypothesis. So, there is a error right now, because a sample actually came from null hypothesis, but you are rejecting it.

But you may improve, and say that why this boundary here, I am going to shift this boundary here, I will get rid from this guy here, and I will shift this guy here. So, in that way you are trying to cover here, but what you are doing on the other hand, a distributions from this may generate a sample here. But this sample generated from this distribution will it be accepted as null hypothesis? It will be accepted as null hypothesis. So, where to put this then additional boundary? There is a dilemma here.

So, suppose let us this this hypothesis are very well separated, let us take this case, like one hypothesis is here, almost getting killed, and another one is here, like here. So, their overlap is very small. Like, I mean, the probability that for this distribution gain here, this is almost like a small probability. And whereas, this guy encroaching, or maybe let us even make it simpler, very well separated like let us say this guy, very high and almost 0, that is hitting 0, it will never hit 0, but let us say it is 0, and the another one is like almost like this. They are very well separated, maybe you can put addition boundary here, then you kind of they are well separated, so you do not care about this.

But it may happen that you may end up with this situation, they are very close, then the question is how to put the boundary so that you do not get confused one guy for the other? Because they are overlapping here? Whatever this guy is generating that is actually overlapping what the space of the other guy also, and because of that, there is a possibility of making mistake here. And that is where to study this we have introduced something called type of errors. So, to give you another examples, so let us do like we this. Another example, is very useful is in this signal processing, hypothesis testing is pretty much used in signal processing and all.

So, how many of you know radars? How radars work? None of you know radar? You know, how aircraft flies? Or how the communication happens? Like who, how the aircrafts get all these communications? You know that big, big antennas? Put. All those, fine. We do not need to get into that. But, let us understand. A very stripped down version of this. So, the radar basically works on the principle that you send signal, and if the object is there, the signal gets reflected back, and if the object is not there, the signal will not come back to you. And now depending on the time taken for your signal to come back, when there is an object, you can identify how far the object is from you, is that clear?

So, let us say this is your, let us say I am just putting antenna like this, or make some big antenna like this, and let us say there is a some aircraft like. Now, you can send some signals to this guy, and it will get reflected, and come backs to you. And the time taken for you know, we know the velocity of light, we know once I send a signal, how much time if it is going to come back after certain time, I know how much distance it traveled. So, we can use this to see that how far my aircraft is away. And we can also use it not only to detect, and also to track it, like how fast it is approaching me and all. So, this is how all this warplanes work, and you detect and try to engage them, if it is a war kind of scenario and all.

Now, suppose now, I want to use, if I want to use in such scenario fine, this is a basic concept I want to use it the property of light traveling its reflection to compute whether the there is a some object there and all. But I am not in an ideal scenario, there could be some objects, when this is there could be some tree here, or maybe some mountains here, or maybe some installations are there and all. They will also reflect my signal, and I should not confuse those signal with that reflected by this aircraft. And in addition to that, there are so many other stochastic behavior in a what we call it as a wireless channel, it depends on your weather, it depends on your humidity, so many other things which we will not get into.

But suppose let us say this aircraft is not there, you are testing your equipment. And here you know like the, only this aircraft is not there, but you know, there is a tree, there is a mountain, and some other installation are there. Like you know, like when you send your signal what happens to that when it returns and all, you characterise there. Let us say that is 1, I will actually call it an alternate hypothesis, that hypothesis, like that hypothesis when the aircraft is not there, what is the observations you are going to make? And when there is an aircraft, actual aircraft is there of course, the scenario is different. And though you are going to call it as a null hypothesis, what happens?

And it may happen that with some chances, your system, even when the aircraft is present, it may say that it is not present, and it may otherwise also happen that the aircraft is actually not there, but your system may say that aircraft is there, this is because this overlapping regions. So, when you are, see when the aircraft is there, maybe this is the guy who is going to be reflecting most of my power standing back. So, it may be having a higher mean of, and is receiving my signal with higher mean. When this aircraft is not there only the small guys like trees and mountains they will be reflecting, so my reflections are going to be weaker, it is like just like noise.

But it may happen that the noise is so powerful at some instances that you may confuse that with a aircraft being present, similarly. So, now that what we want to study now, we have hypothesis let us say simply hypothesis, the aircraft I, when you see all these aircrafts, and navy ships, and all, you will see big, big radar screens on them, or big, big screens on them. They will keep on showing some object is present or not. Suppose on, my radar screen start showing that an object is present.

Now, that does not mean that I should start firing my missiles, or start shooting, I need to ascertain that the object is indeed present, and that is where I need to see that, if it is saying present, it is really present, or with what percent it is saying that is present, that is where this types of error come into picture, and we are going to put it, and we will see two types of error basically.

Suppose, so, you have two hypothesis, and underlying data is actually generated as per your null hypothesis, and you accept it to be, and you accept it as null hypothesis, then you are correct. But underlying this, underlying samples are coming from your null hypothesis, but you reject it, then it is going to be type I error. And similarly, underlined data may be generated according to alternate hypothesis, but you accept it to be null hypothesis, then it is called a type II error. And similarly, if you it is an alternate hypothesis, but you kind of reject the null hypothesis then that is fine, that is in this category.