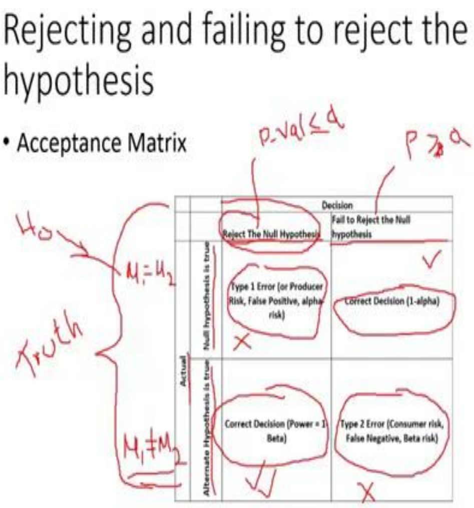
Introduction to Data Analytics
Prof. Nandan Sudarsanam and
Prof. B. Ravindran
Department of Management Studies and
Department of Computer Science and Engineering
Indian Institute of Technology, Madras

Module - 03
Lecture - 12
Inferential Statistics - Type 1 and Type 2 Errors

Hello and welcome to our lecture on Type 1 and Type 2 Errors. This is something that you might have heard and it is a fairly central and important concept with respect to hypothesis tests.

(Refer Slide Time: 00:23)



So, let us get into, what the core concept of type 1 and type 2 errors are. So, the idea is the following. Take any of the hypothesis tests that you learnt so far, in this course we provided you with a template for how the hypothesis tests are conducted. And in this template, you would notice the first step is to form a null and an alternative hypothesis and your last step is that you essentially get a p value associated with this particular test and based of the p value, you either reject the null hypothesis or you fail to reject the null hypothesis. So, this is what we have characterized in this table.

So, you actually make a null and an alternative hypothesis and you do not know which is true, that is why you are doing this test. So, your null hypothesis was that $\mu_1 = \mu_2$. So, if you knew that was true you would not be doing this test, so you do not know that is true, that is why you are doing this test. So, you might have a null hypothesis that says μ_1 = 4.8, which is a single sample case or you might have a null hypothesis, which says that $\mu_1 = \mu_2$.

In either case, you do not know which is true. But, let us say there was this, all knowing world, where you knew which was true. So, that is what marked out here in actual, you say actually what is true. So, here I am saying the null hypothesis is true and in this part, I am saying the alternative hypothesis is true. So, here I am saying μ_1 is equal to, let us say μ_2 and here the alternate hypothesis is true, $\mu_1 \neq \mu_2$. Now, this is the truth.

But, based on some data, you did a test. So, you might have done a two sample z test or you might have done a two sample t test and let say, that you did the test you actually computed the z or the t statistic. Based off of the z or t statistic, you calculated a probability or a p value and based of a p value, you took a decision and the core idea has always that p value is too small, then you reject the null hypothesis and the term too small is a subjective term and typically people use some line in the some threshold and that threshold is called α.

So, if the p value $< \alpha$, we reject the null hypothesis. Typically, the value of α tends to be something like 0.05 or 5 percent. So, based off of this, let us say you rejected the null hypothesis, that is one decision and here you fail to reject the null hypothesis, for whatever reason. So, the idea is that out here you would have rejected the null hypothesis, because your p value would have been less than your α.

So, p value is less than α, so you rejected the null hypothesis. Here the opposite was true, p was greater than or equal to α that is it. You do not have to have that equal to sign or so, just say greater than and I am not sure, you might want to say less than or equal to. So, this is what you did; now here is the problem. If the truth was that the null hypothesis was true, the null hypothesis was true which means that $\mu_1 = \mu_2$.

But, you went ahead out here in this quadrant, you went ahead and you rejected the null hypothesis. You said $\mu_1 \neq \mu_2$, but in reality $\mu_1 = \mu_2$. You did something wrong and that error is what is called is a type 1 error. It is when the null hypothesis was true, but you went ahead and you said that I am rejecting the null hypothesis and we are going to come back to quantifying that value which is called the type 1 error.

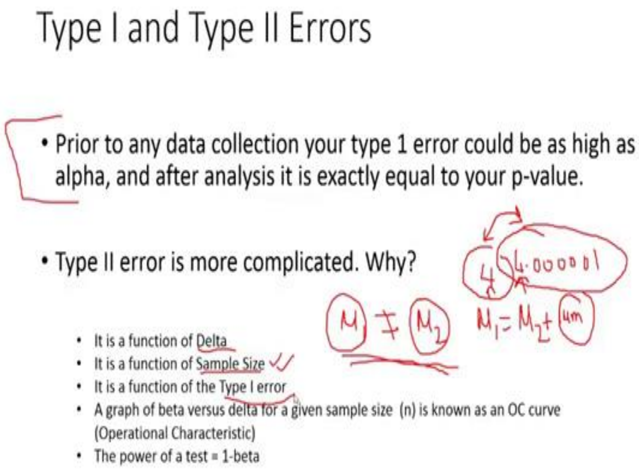
It is just noteworthy that is also called the producer risk, in sometimes the false positive or the alpha risk. We will not go in to each of those terms, for instance producer risk is seen more from a manufacturing context, and false positive is seen more from a medical context. But, a few kind of think about it, it definitely makes sense. Now, here is the case out here, where the null hypothesis was true and you fail to reject the null hypothesis.

So, that is okay, you are happy with that, reality was that the null hypothesis was true and you did not find any evidence to reject the null hypothesis. So, you did the right thing, so you are happy in this quadrant and let us come to this quadrant out here. Here, the null hypothesis in reality, because this is truth, so this is truth. So, here $\mu_1 \neq \mu_2$ and you correctly rejected the null hypothesis, this is still the null hypothesis.

Your null hypothesis is this, this $\mu_1 \neq \mu_2$ is the alternate hypothesis, but it is so happened, that the alternate hypothesis was reality that the null hypothesis was wrong and when the null hypothesis was wrong, you correctly rejected it. So, you did the right thing even here and that is a very good thing that you did. You actually detected that $\mu_1 \neq \mu_2$ and you said no, I am rejecting the null hypothesis, so that is a great thing.

Now, come to the final quadrant. Here, the truth was that $\mu_1 \neq \mu_2$, but you failed to reject the null hypothesis. You are not able to say, I am rejecting the idea that $\mu_1 = \mu_2$, I am rejecting the null hypothesis. So, you are not able to reject the null hypothesis, whereas you should have, because $\mu_1 \neq \mu_2$ and that is the truth. So, that is also not a great thing that you did and this error is called the type 2 error and again, it is call the consumer risk in the more manufacturing production settings, it is called the false negative in medical settings in couple of other settings or beta risk.

(Refer Slide Time: 07:06)



Now, in really simple words, type 1 error is the concept. So, just to be really clear, we are right here. So, the idea is that your type 1 error, even before you see any data, even before you do anything could be as high as your α and I say after analysis; it is exactly equal to your p values. So, let me explain that a little bit. So, let see you have not collected any data, we just discussed what α is. Alpha is the idea that you are going to do this hypothesis test and you are going to get some p value and you already said, if the p value is too low I am going to reject the null hypothesis. How low is too low that is the line in the stand that you draw and that value is called α.

So, if your p value is less than α, you are going to reject the null hypothesis. What is that mean when your $p < \alpha$? First of all, you are going to get a very low p value, let us say you are going to get a low p value and this $p < \alpha$. Because, you type one error first of all comes about only, when your null hypothesis is true and you go ahead and reject the null hypothesis.

So, by definition; that means, your p value must have been less than α, only then you would have rejected your null hypothesis. Now, only when you reject your null hypothesis or you even putting yourself up for the possibility of a type 1 error. So, the idea is that I am going to tell you that up front, your type 1 error is associated with this idea that you rejecting the null hypothesis and you reject the null hypothesis when your p value is less than alpha, which means your p value is really low.

Now, let us take a step back and think about, what the p value is. The p value is this idea that, if the null hypothesis is true, this is the definition that we looked at much earlier, even before we discussed type 1 and type 2 errors. P values by definition express the idea that if your null hypothesis is true, this is the probability of seeing a test statistics as extreme as this, if your null hypothesis is actually true. So, this probability is really low, then essentially what you are saying is that, you are willing to take the risk and saying, the probability of seeing this statistics is just 1 percent if the null hypothesis is true.

If the null hypothesis is true, the probability of seeing something so extreme is just so small, it is just 1 percent, then I am willing to take the risk and reject the null hypothesis. So, write there by definition, the risk that you took was that 1 percent risk, it was that p value. So, whatever your p value is, essentially if you do the entire calculation and you calculate the p value, then your p value is essentially your type 1 error, that is the risk you are taking, that because that is you actually calculated the probability that this data could actually occur with the null hypothesis being true.

So, if you are still going ahead and rejecting the null hypothesis fully knowing this probability, then that is the risk you are taking of making a wrong decision. Now of course, before you even see any data and even before you calculate the p value, given that you set yourself with an upper threshold of α, means that you could get a type 1 error as high as α. So, I hope that makes type 1 errors fairly clear.

Now, type 2 errors are more complicated, now and there is a reason for it. It is a function of variety of parameters, it is a function of something called delta, which is not the delta that we would have discussed so far and the core concept behind delta is the following. Type 2 errors, when do they occur? They occur in the situation, where your alternate hypothesis is actually true. So, $\mu_1 \neq \mu_2$, but you failed to reject the null hypothesis.

But, in order for me to quantify how likely that is, you need to tell me, how much is $\mu_1 \neq \mu_2$. So, if I said $\mu_1 \neq \mu_2$, because $\mu_1 = \mu_2 +$ {a very, very, very, very, very small numbers}. So, let us say 1 micro meter or whatever it is, whatever the metric of μ_1, μ_2 is, but it is a very small numbers. So, μ_1 let say is 4, but μ_2 =4.000001, then it is true $\mu_1 \neq \mu_2$. The alternate hypothesis is true, but the probability that I am going to reject the null hypothesis, just became very low.

My ability to discern between a difference, this difference between 4 and 4.000001 is much lower than my ability to discern between the difference of 4 and 5, all other things being equal. So, it is really a function of how different are they and that is what gets captured in delta, it is also a function of the sample size we take, which is as the sample size becomes infinite. The uncertainty around μ_1 & μ_2 becomes smaller and smaller and theoretically, even a small difference between 4 and 4.000001 could potentially be found as long this sample size is large enough.

It is a function of type 1 error, because your type 1 error tells you, how conservative or liberal you are being in rejecting the null hypothesis. So, at the more conservative you are in type 1 error, meaning that you do not want to make that type 1 error, the more likely you are to make a type 2 error and similarly, the more liberal you are with the type 1 error, meaning you are with some amount of error the better you are going to be with type 2 errors.

So, it is a little bit of given take in terms of the type 1 and type 2 error and essentially, a lot of people are very interested in understanding the relationship of beta, which is this type 2 error. So, it is a bad thing versus delta, which is something we would discussed here and for a given sample size and they will show it from many different sample size and that is known as an OC curve or an Operational Characteristic curve.

Another word that you might come across, it is known as the power of the test and that is equal to 1 - beta and it essentially a very positive thing. It is 1 - a bad thing. So, the higher the power of the test; that means, the more strength you have in being able to detect a difference between μ_1 & μ_2 and I have use the concept of μ_1 & μ_2 coming from the two sample scenarios, but all of these core concepts also applied to a single sample tests. I hope that clarified and that give you an idea of type 1 and type 2 errors.

Thank you.