1. Introduction

Welcome to the series of E-learning modules on Types of Errors. In this module we are going cover the concept of statistical errors, Type one and Type two errors and understand the consequences and effect of sample size on errors.

By the end of this session, you will be able to:

- Understand the concept of statistical errors
- Describe the role of Type I and Type two errors
- Understand the theory behind the statistical errors
- Understand the consequences and effect of sample size on errors

It can be said, in sum, that statistics enable scientists to test substantive hypotheses indirectly by enabling them to test statistical hypotheses directly. They test the “truth” of substantive hypotheses by subjecting null hypotheses to statistical tests on the basis of probabilistic reasoning.

Statistical errors:
In statistical test theory, the notion of statistical error is an integral part of hypothesis testing. The test requires an unambiguous statement of a null hypothesis, which usually corresponds to a default "state of nature".
For example, “this person is healthy", "this accused is not guilty" or "this product is not broken".

An alternative hypothesis is the negation of null hypothesis, for example, "this person is not healthy", "this accused is guilty" or "this product is broken".
The result of the test may be negative, relative to null hypothesis for example, not healthy, guilty, broken or positive like, healthy, not guilty, not broken.
If the result of the test corresponds with reality, then a correct decision has been made.

However, if the result of the test does not correspond with reality, then an error has occurred. Due to the statistical nature of a test, the result is never, except in very rare cases, free of error. Two types of error are distinguished: type one error and type two errors.

Type one error is the incorrect rejection of a true null hypothesis. A Type two error is the failure to reject a false null hypothesis.
Usually, a type one error leads one to conclude that a thing or relationship exists when really it doesn't.
For example, a patient has a disease being tested for when in reality the patient does not have the disease, or that a medical treatment cures a disease when it really does not.

Examples of type two errors would be a blood test failing to detect the disease it was designed to detect, in a patient who really has the disease;
or a clinical trial of a medical treatment failing to show that the treatment works when really it does.Type one and Type two errors are also called errors of the first kind and errors of the second kind.

2. Types of Decisions

Types of Decisions: Correct and Incorrect
In the typical hypothesis testing situation, we come to the final decision where we either retain or reject the null hypothesis.
Thus, a decision is made in the study. But, juxtaposed to this decision is the real truth.
The null hypothesis is either true or false. What this yields is a two by two table that indicates the various combinations of decisions you make compared to the actual truth.

Consider the following table.

Figure 1

	Null Hypothesis (H_0) is True	Null Hypothesis (H_0) is False
Retain Null Hypothesis	Correct Decision	Wrong Decision
Reject Null Hypothesis	Wrong Decision	Correct Decision

In the first column, notice that we have the two possible decisions that we can make. That is, either retain the null hypothesis or reject the null hypothesis.
In the first row, we have the two possible states of truth. That is, either the null hypothesis is true or it is false.

This sets up four possible different combinations of decision and truth. For purposes of discussion, assume for a moment that the null hypothesis is that, the population mean I Q score is one hundred.
If the population mean is really one hundred and we retain the null hypothesis, then we are in the upper left cell and have made a correct decision. To retain the null hypothesis when the null hypothesis is true is the right thing to do.
On the other hand, if the true mean is one hundred and you reject the null hypothesis, then you have made the wrong decision and you are in the lower left cell. You have made a wrong decision what is called a Type one or Alpha error.
A type one or alpha error is when you reject the null hypothesis when the null hypothesis is true.

If the null hypothesis is false and you retain the null hypothesis that is, the upper right cell, then again you have made a wrong decision. You have retained the null hypothesis when the null hypothesis is not correct. Since this is the opposite of a Type one error, this is called a Type two or Beta error.
A type two or beta error is when you retain the null hypothesis when the null hypothesis is not true.
Finally, if you reject the null hypothesis when the null hypothesis is false, then you have made

a correct decision. If the true population mean I Q score is not one hundred and you reject the null hypothesis of one hundred, then you are saying that you do not think that the mean is one hundred and you would be doing the right thing by rejecting the null hypothesis.

Let's study the above diagram in another context, a jury trial. The null hypothesis in a trial, when we give the benefit of doubt to the person on trial, is that the person is not guilty and did not commit the crime.

Look at the following diagram:

Figure 2

	H_0 is true Innocent	H_0 is false Guilty
Retain H_0 (Acquit)	Correct Decision (Let an innocent person go free)	Incorrect decision (Let a guilty person go free)
Retain H_0 (Convict)	Incorrect decision (Convict an innocent person)	Correct Decision (convict a guilty person)

If the person is really not guilty and the jury votes "not guilty" and acquits the person, then they have made the correct decision. The jury has set free an innocent person.
If the person is actually not guilty but the jury convicts the person anyway, then the jury has made a TYPE ONE or alpha error. Convicting an innocent person is an error of the "first kind".
On the other hand, if the person is really guilty and the jury votes to acquit the person, then they again have made a mistake and this type of mistake in decision is called a TYPE TWO or Beta error. This is an error of the "second kind".
Finally, if the jury votes to convict the person and the person really did commit the crime, then the jury has made the correct decision.

In the literature, the decision of rejecting a false null hypothesis is called POWER.
Although you may not have thought about it this way, the main purpose of a trial is to convict guilty persons. Certainly, we want the jury to arrive at the correct decision and if the person is not guilty, then we hope the jury will acquit the person. However, the defence side of the case that is, to defend the person on trial, is not the side to bring the case to trial. It is the prosecution side; they are the ones to instigate the entire trial process. In this context therefore, the hopes of the trial system, which in addition to finding the truth, is to convict persons of the crime they were charged with especially if they really committed the crime.

Thus, what we see here is the possibility of 4 different consequences of our decisions:
Two good and correct decisions and also two bad and incorrect decisions ones. Of course, whether it is in a jury trial situation or an experiment where we retain or reject the null hypothesis, our goal is to maximize the chances of making a correct decision and minimize the chances of making an incorrect decision.

To summarize, we will do one the following.
Make a right decision by:

- Retaining the null hypothesis when the null hypothesis is true, or
- Rejecting the null hypothesis when the null hypothesis is false

Make an incorrect decision by:

- Retaining the null hypothesis when the null hypothesis is false (TYPE TWO or beta error)
- Rejecting the null hypothesis when the null hypothesis is true (TYPE 1 or alpha error)

Generally speaking, researchers are not willing to reject the null hypothesis if the chance of rejecting wrongly is greater than about five percent.
Remember, we usually want to reject the null hypothesis but we only want to do this if the chance that we do it incorrectly is rather low.

Unfortunately, setting some criterion about how much Type one or alpha error rate we are willing to tolerate does little to regulate or establish what level of Type two error or beta we are willing to tolerate.
The literature seems to be fixated only on the rate of Type one errors without giving much attention to Type two error.

Null hypothesis is rejected with the awareness that an error might have been made, but the chances of that happening are **less than five percent.**
The conclusion of rejecting null hypothesis on an average is correct more than ninety five percent of the time.

3. Effect of Sample Size

Effect of sample size
The size of the sample is related to both types of errors.
With a fixed value of type one error and a fixed sample size n, the value of type two error is predetermined.

If type two error is too large, it can be reduced by either raising the level of type one error for fixed n, or by increasing n for a fixed level of type one error.
Although type two error is seldom determined in an experiment, researchers can be assured that it is reasonably small by collecting a large sample.

The goal of the test is to determine if the null hypothesis can be rejected.
A statistical test can either reject, (that is, prove false) or fail to reject, (that is, fail to prove false), a null hypothesis, but never prove it true that is, failing to reject a null hypothesis does not prove it true.

Tabularized relations between correct or incorrectness of the null hypothesis and outcomes of the test are given below.

Figure 3

	Null Hypothesis (H_0) Is True	Null Hypothesis (H_0) Is False
Reject Null Hypothesis	Type I Error False Positive	Correct Outcome True Negative
Fail To Reject Null Hypothesis	Correct Outcome True Positive	Type II Error False Negative

False negative error:
A false negative error is where a test result indicates that a condition failed, while it actually was successful.
A common example is a guilty prisoner freed from jail. The condition: "*Is the prisoner guilty?*" actually had a positive result (yes, he is guilty). But the test failed to realize this, and wrongly decided the prisoner was not guilty.
A false negative error is a type two error occurring in test steps where a single condition is checked for and the result can either be positive or negative.

False positive error:
A false positive error is where a test result indicates that a condition turned out successful, while it actually was a failure.
As it is conjectured that adding fluoride to toothpaste protects against cavities, the null

hypothesis of no effect is tested.
When the null hypothesis is true (that is, there is indeed no effect), but the data give rise to rejection of this hypothesis, falsely suggesting that adding fluoride is effective against cavities, a type one error has occurred.
A type two error occurs when the null hypothesis is false (i.e., adding fluoride is actually effective against cavities), but the data are such that the null hypothesis cannot be rejected, failing to prove the existing effect.

4. Theory and Consequences

Theory
From the Bayesian point of view, a type one error is one that looks at information that should not substantially change one's prior estimate of probability, but does.
A type two error is one that looks at information which should change one's estimate, but does not. (Though the null hypothesis is not quite the same thing as one's prior estimate, it is, rather, one's *pro forma* prior estimate.)

Hypothesis testing is the art of testing whether a variation between two sample distributions can be explained by chance or not.
In many practical applications, type one errors are more delicate than type two errors. In these cases, care is usually focused on minimizing the occurrence of this statistical error.

Suppose, the probability for a type one error is one percent, then there is a one percent chance that the observed variation is not true; this is called the level of significance.
While one percent might be an acceptable level of significance for one application, a different application can require a very different level.

For example, the standard goal of six sigma is to achieve precision to four point five standard deviations above or below the mean. This means that only three point four parts per million are allowed to be deficient in a normally distributed process.

Consequences
Both types of errors are problems for individuals, corporations, and data analysis.
A false positive (with null hypothesis of health) in medicine causes unnecessary worry or treatment, while a false negative gives the patient the dangerous illusion of good health and the patient might not get an available treatment.

A false positive in manufacturing quality control (with a null hypothesis of a product being well made) discards a product that is actually well made, while a false negative stamps a broken product as operational.
A false positive (with null hypothesis of no effect) in scientific research suggest an effect that is not actually there, while a false negative fails to detect an effect that is there.

Minimizing errors of decision is not a simple issue.
For any given sample size the effort to reduce one type of error generally results in increasing the other type of error.
The only way to minimize both types of error, without just improving the test, is to increase the sample size, and this may not be feasible.

5. Causes for Errors

Causes for Errors in testing:
Just because the sample mean (following treatment) is different from the original population mean does not necessarily indicate that the treatment has caused a change.
You should recall that there usually is some discrepancy between a sample mean and the population mean simply as a result of sampling error.

Since the hypothesis test relies on sample data, and because sample data are not completely reliable, there is always the risk that misleading data will cause the hypothesis test to reach a wrong conclusion.

Causes for Type one errors

- A Type one error occurs when the sample data appear to show a treatment effect when, in fact, there is none
- In this case the researcher will reject the null hypothesis and falsely conclude that the treatment has an effect
- Type one errors are caused by unusual, unrepresentative samples. Just by chance the researcher selects an extreme sample with the result that the sample falls in the critical region even though the treatment has no effect
- The hypothesis test is structured so that Type one errors are very unlikely; specifically, the probability of a Type one error is equal to the alpha level

Causes for Type two errors

- A Type two error occurs when the sample does not appear to have been affected by the treatment when, in fact, the treatment does have an effect
- In this case, the researcher will fail to reject the null hypothesis and falsely conclude that the treatment does not have an effect
- Type two errors are commonly the result of a very small treatment effect. Although the treatment does have an effect, it is not large enough to show up in the research study.

Hence we need to briefly examine the consequences of the decisions we make when testing hypotheses.
Since we are using sample data to make inferences about the population parameters, and in this context, we retain or reject some null hypothesis, we can never be sure that the decision we make is the correct one.
Since we only have sample data and not the population data, our inferences are only our best estimates of the truth.

Summary
Just like a judge's conclusion, an investigator's conclusion may be wrong. Sometimes, by chance alone, a sample is not representative of the population. Thus the results in the sample do not reflect reality in the population, and the random error leads to an erroneous inference.

A type one error (that is, false-positive) (or error of first kind) occurs if an investigator rejects a null hypothesis that is actually true in the population.
A type two error (that is, false-negative) (or Error of second kind) occurs if the investigator

fails to reject a null hypothesis that is actually false in the population.

Although type one and type two errors can never be avoided entirely, the investigator can reduce their likelihood by increasing the sample size.
Larger the sample, lesser is the likelihood that it will differ substantially from the population.

False-positive and false-negative results can also occur because of bias. (Bias of the observer, instrument, recall, etc.).
Errors due to bias, however, are not referred to as type one and type two errors.
Such errors are troublesome, since they may be difficult to detect and cannot usually be quantified.

Here's a summary of our learning in this session where we have:

- Understood the concept of Statistical Errors
- Understood the role of type one errors in statistical tests
- Understood the role of type two errors in statistical tests
- Explained the theory behind the errors
- Understood the reasons for occurrence of two types of errors
- Explained the effect of sample size on the errors
- Explained practical illustrations for the occurrence of Type one and Type two errors