Type I Error and Type II Error

Martyn Shuttleworth, Lyndsay T Wilson187.5K reads

Experimental Errors in Research

While you might not have heard of Type I error or Type II error, you're probably familiar with the terms "false positive" and "false negative."

A common medical example is a patient who takes an HIV test which promises a 99.9% accuracy rate. This means that in 0.1% of cases, or 1 in every 1000, the test gives a 'false positive,' informing a patient that they have the virus when they do not.

On the other hand, the test could also show a false negative reading, giving a person who is actually HIV positive the all-clear. This is why most medical tests require duplicate samples, to stack the odds in our favor. A 1 in 1000 chance of a false positive becomes a 1 in 1 000 000 chance of two false positives, if two tests are taken.

With any scientific process, there is no such thing as total proof or total rejection, whether of test results or of a null hypothesis. Researchers must work instead with probabilities. So even if the probabilities are lowered to 1 in 1000 000, there is *still* the chance that the results may be wrong.

How Does This Translate to Science?

Type I Error

A Type I error is often referred to as a "false positive" and is the incorrect rejection of the true null

hypothesis in favor of the alternative.

In the example above, the null hypothesis refers to the natural state of things or the absence of the tested effect or phenomenon, i.e. stating that the patient is HIV negative. The alternative hypothesis [1] states that the patient is HIV positive. Many medical tests will have the disease they are testing for as the alternative hypothesis and the lack of that disease as the null hypothesis.

A Type I error would thus occur when the patient doesn't have the virus but the test shows that they do. In other words, the test incorrectly rejects the true null hypothesis that the patient is HIV negative.

Type II Error

A Type II error is the inverse of a Type I error and is the false acceptance of a null hypothesis that is not actually true, i.e. a false negative. A Type II error would entail the test telling the patient they are free of HIV when they are not.

Considering this HIV example, which error type do you think is more acceptable? In other words, would you rather have a test that was more prone to Type I or Type II error? With HIV, it's likely that the momentary stress of a false positive is better than feeling relieved at a false negative and then failing to take steps to treat the disease. Pregnancy tests, blood tests and any diagnostic tool that has serious consequences for the health of a patient are usually *overly* sensitive for this reason – it is much better for them to err on the side of a false positive.

But in most fields of science, Type II errors are seen as less serious than Type I errors. With the Type II error, a chance to reject the null hypothesis was lost, and no conclusion is inferred from a non-rejected null. But the Type I error is more serious, because you have wrongly rejected the null hypothesis and ultimately made a claim that is not true. In science, finding a phenomenon where there is none is more egregious than failing to find a phenomenon where there is. Therefore in most research designs, effort is made to err on the side of a false negative.

Replication

This is the key reason why scientific experiments [2] must be replicable.

Even if the highest level of proof is reached, where P < 0.01 (probability [3] is less than 1%), out of every 100 experiments, there will still be one false result. To a certain extent, duplicate or triplicate samples reduce the chance of error [4], but may still mask chance if the error [5]-causing variable [6] is present in all samples.

But if other researchers, using the same equipment, replicate the experiment and find that the results are the same, the chances of 5 or 10 experiments giving false results is unbelievably small. This is how science regulates and minimizes the potential for both Type I and Type II errors.

Of course, in certain experiments and medical diagnoses, replication is not always possible, so the possibility of Type I and II errors is always a factor.

One area that is guilty of forgetting about Type I and II errors is in the legal system, where a jury is seldom told that fingerprint and DNA tests may produce false results. There have been many documented failures of justice involving such tests. Today courts will no longer accept these tests alone as proof of guilt, and require other evidence to reduce the possibility of error to acceptable levels.

Type III Errors

Some statisticians are now adopting a third type of error, Type III, which is where the null hypothesis was correctly rejected …but for the wrong reason.

In an experiment, a researcher might postulate a [hypothesis](https://explorable.com/research-hypothesis) [1] and perform research. After analyzing the results statistically, the null hypothesis is rejected.

The problem is that there may indeed be some relationship between the [variables](https://explorable.com/research-variables) [6], but it's not the one stated in the hypothesis. There is no error in rejecting the null here, but the error lies in accepting an incorrect alternative hypothesis. Hence a still unknown process may underlie the relationship, and the researchers are none the wiser.

As an example, researchers may be interested to see if there is any difference in two group means, and find that there is one. So they reject the null hypothesis but don't notice that the difference is actually in the opposite direction to what their results found. Perhaps random chance led them to collect low scores from the group that is in reality higher and high scores from the group that is in reality lower. This is a curious way of being both correct and incorrect at the same time! As you can imagine, Type III errors are rare.

Economist Howard Raiffa gives a different definition for Type III error, one that others have called Type 0: getting the correct answer to an incorrect question.

Additionally, a Type IV error has been defined as incorrectly *interpreting* a null hypothesis that has been correctly rejected. Type IV error comes down to faulty analysis, bias or fumbling with the data to arrive at incorrect conclusions.

Conclusion

Errors of all types should be taken into account by scientists when conducting research.

Whilst replication can minimize the chances of an inaccurate result, it is no substitute for clear and logical research design, and careful analysis of results.

Many scientists do not accept [quasi-experiments](https://explorable.com/quasi-experimental-design) [7], because they are difficult to replicate and analyze, and therefore have a higher risk of being affected by error.

Source URL: https://verify.explorable.com/type-i-error

Links
[1] https://explorable.com/research-hypothesis
[2] https://explorable.com/conducting-an-experiment
[3] https://explorable.com/experimental-probability
[4] https://explorable.com/random-sampling-error
[5] https://explorable.com/experimental-error
[6] https://explorable.com/research-variables
[7] https://explorable.com/quasi-experimental-design