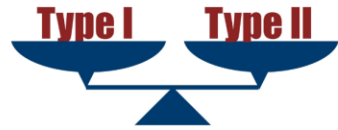


Type I and type II errors

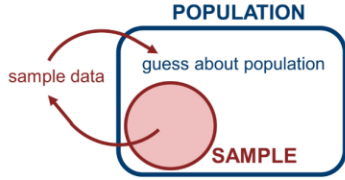


btw, errors aren't mistakes



Doing statistics

When we do statistics, we typically use a sample to make an educated guess about the population.



But sampling error (i.e., noise) means that our samples are sometimes misleading.

This is statistical **error** due to randomness,

This is not a **mistake** on our part - unless we don't realize it might be happening ... that's a mistake.

Doing statistical tests

When we do statistical tests, we use a null and alternative hypothesis.

Reality: either H_0 or H_A is true.

We then do stats and get a *decision* or *conclusion*.

Decision: either H_0 or H_A is true.

We can avoid mistakes (like badly designed hypotheses), but sampling error *may* cause the wrong decision.

H_0 : the green one is better
 H_A : the red one is better.



Doing statistical tests

We *cannot* eliminate both types of error.

We *can* shift our requirements for the decision.

But this is a tradeoff ... reducing risk of Type I increases the risk of Type II and vice versa.

Our decision
Accept H_0 Reject H_0

H_0 true	Correct	Type I error
H_0 false	Type II error	Correct

Criminal justice system analogy

This logic and these types of risk are common sense.

H_0 : person is innocent
 H_A : person is guilty

Type I error
Convicting an innocent person

Type II error
Letting a guilty person go unpunished.

Our decision
Accept H_0 Reject H_0

H_0 true	Correct	Type I error
H_0 false	Type II error	Correct



Doing statistical tests

Reality: either H_0 or H_A is true.
Decision: either H_0 or H_A is true.

Type I error
Rejecting a true H_0

Type II error
Accepting a false H_0

Our decision
Accept H_0 Reject H_0

H_0 true	Correct	Type I error
H_0 false	Type II error	Correct

Justice system analogy

Type I error: Convicting an innocent person

Type II error: Letting a guilty person go unpunished.

Ideal tradeoff varies based on the perceived consequences of an error.

Criminal cases: "Beyond a reasonable doubt"
Hard to reject H_0 (i.e., p value extremely small)
Civil cases: "According to a preponderance of the evidence"
Easier to reject H_0 (i.e., p value somewhat small)

Conservatives vs **liberals**

	Accept H_0	Reject H_0
H_0 true	Correct	Type I error
H_0 false	Type II error	Correct



Error is inevitable, but when you know your enemy ...

Type I error: also called "**false positives**" sometimes.
Type II error: also called "**false negatives**" sometimes.

Our decision
Accept H_0 Reject H_0

H_0 true	Correct	Type I error
H_0 false	Type II error	Correct

Our test result
Negative Positive

Negative	Correct	False positive
Positive	False negative	Correct

How to interpret test results

The wrong way: The small p value indicates that the null hypothesis can be rejected ~or~ a large p value indicates that the null hypothesis should be accepted and that's THE ANSWER.

The right way: The small p value indicates that the null hypothesis is probably wrong, but we may be making a **type I error** ~or~ the large p value indicates that the null hypothesis is probably true, but we may be making a **type II error**.

STATS IS NOT PROOF