Hypothesis Testing

DSC-ATL

Motivation

- https://amstat.tandfonline.com/doi/pdf/10.1080/00031305.2016.1154108?needAccess=true

Hypothesis Testing

The Steps:

1. Define your hypotheses (null, alternative)
2. Specify your null distribution
3. Perform an experiment
4. Calculate the p-value of what you observed
5. Reject or fail to reject the null hypothesis
6. State conclusion

The Null and Alternative Hypothesis

- **Null hypothesis**. The null hypothesis, denoted by H0, is usually the hypothesis that sample observations result purely from chance(**assumed to be true**).

- **Alternative hypothesis**. The alternative hypothesis, denoted by H1 or Ha, is the hypothesis that sample observations are influenced by some non-random cause(**claim to be proved**).

One-Tail vs. Two-Tail Hypothesis

One-Tail Test - when you want to know if a parameter from the treatment group is greater than (or less than) a corresponding parameter from the control group.

Two-Tail Test - when you want to know if a parameter from the treatment group is different from a corresponding parameter from the control group.

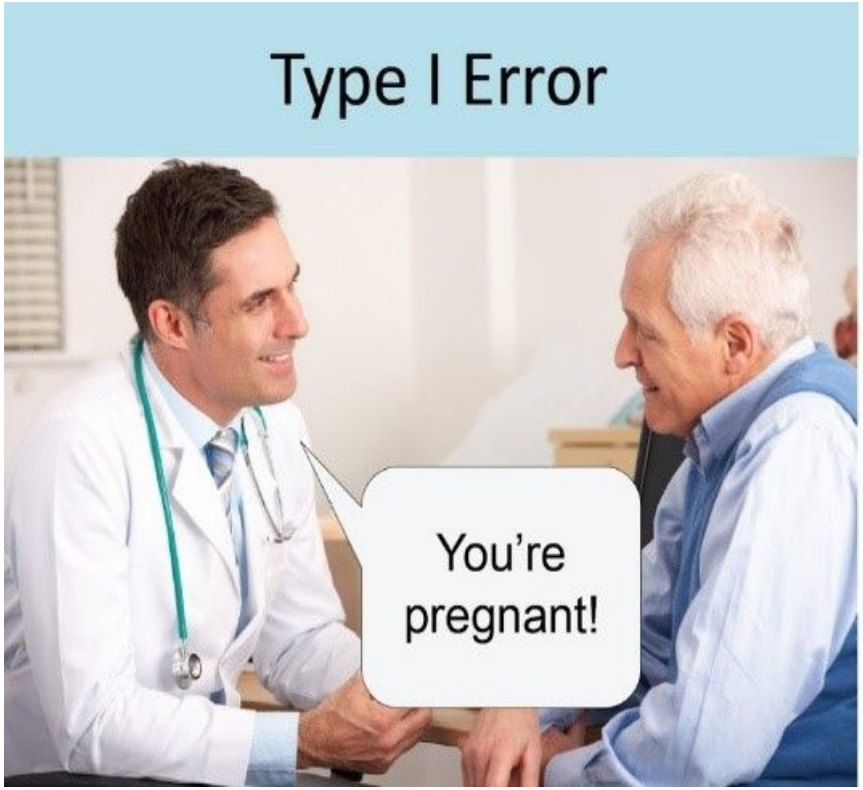
Errors

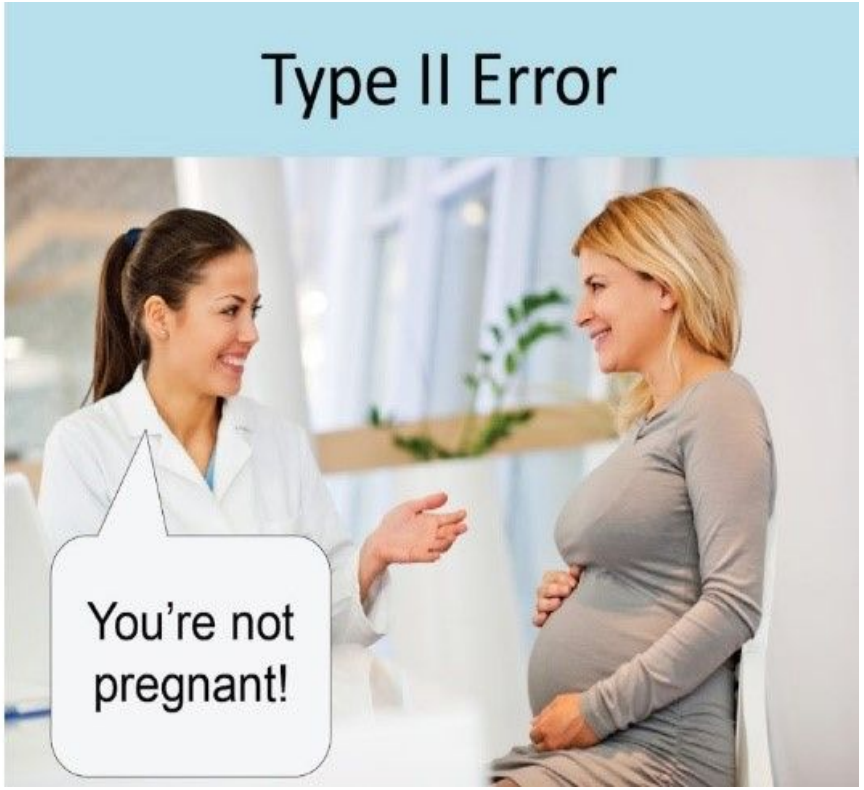
Type-I Error (also known as “α” and “False Positive”): Rejecting the null when the effect isn’t real.

- Incorrectly reject a true null hypothesis

Type-II Error (also known as “β “ and “false negative”): Failing to reject the null when the effect is real.

- Failing to reject a false null hypothesis.





- 'man cannot be pregnant' is the null hypothesis and 'man is pregnant' is the alternative hypothesis.
- 'woman is not pregnant' is the null hypothesis and 'woman is pregnant' is the alternative hypothesis.

Example: Type I and Type II

Your Statistical Decision	**True state of null hypothesis**	
	H_0 True (example: the drug doesn't work)	H_0 False (example: the drug works)
Reject H_0 (ex: you conclude that the drug works)	*Type I error (α)*	*Correct*
Do not reject H_0 (ex: you conclude that there is insufficient evidence that the drug works)	*Correct*	*Type II Error (β)*

Level of Significance

α : The marginal threshold at which you're okay with rejecting the null hypothesis (**Upper bound on Type 1 Error**).

P-Value

- Rejecting H_0 at a specified α does not full convey the information in the data.
- More useful to report the smallest α--level at which the observed result is significant.
- This smallest α--level is called the **P-Value** (observed level of significance).
- Smaller the P-Value the more significant the test result.

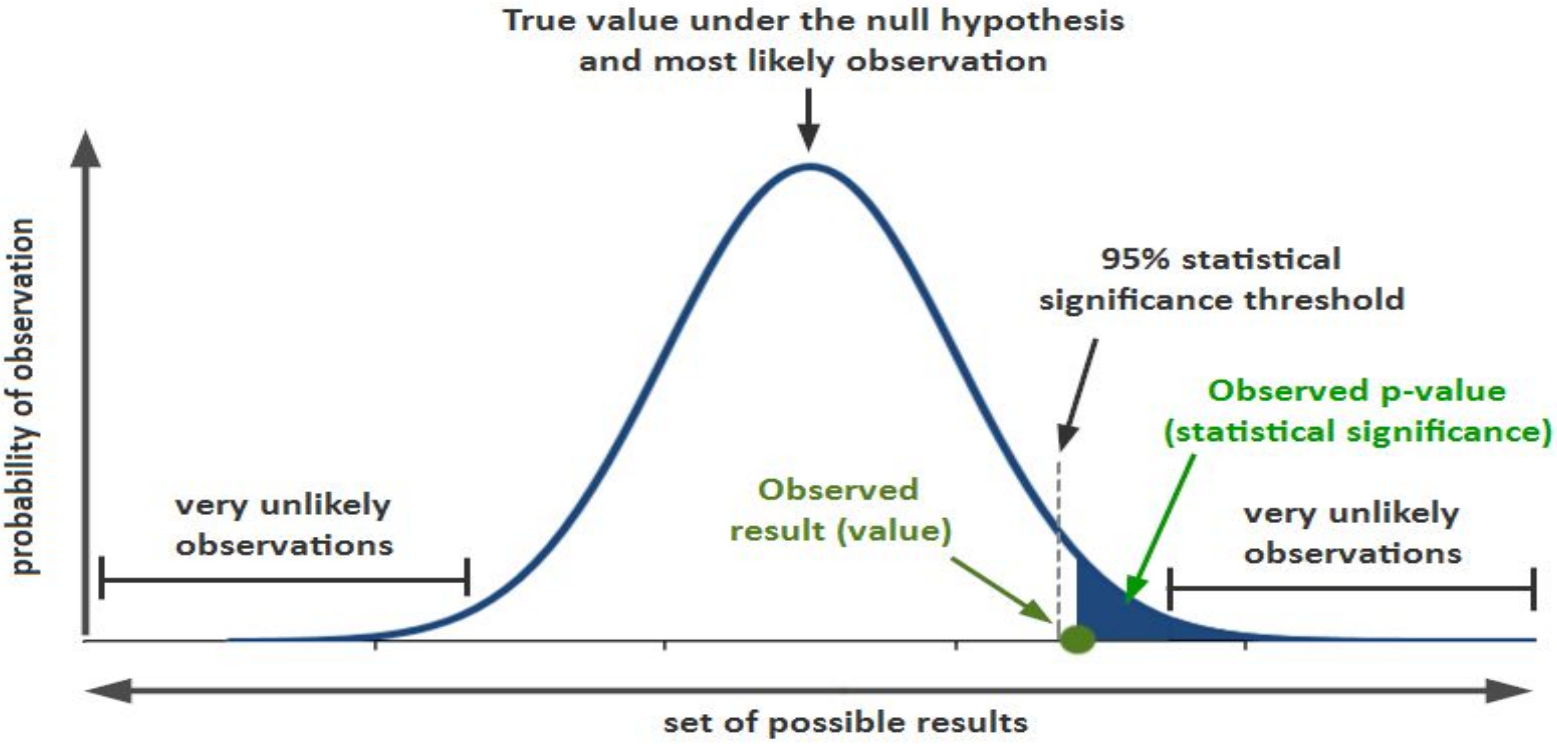
Decision Rule

$$P - Value < \alpha \quad reject \quad H_0$$

Alternative Definition P-Value

- It is the probability under H_0 of obtaining a test statistic at least as "extreme" as the observed value.
- The statement above can be interpreted as a small P-Value indicates that the observed result is rare under the assumptions of H_0

Probability & Statistical Significance Explained



Careful on Interpretations

- Results that are not statistically significant should not be interpreted as "evidence of no effect," but as “no evidence of effect”

T-Test Background

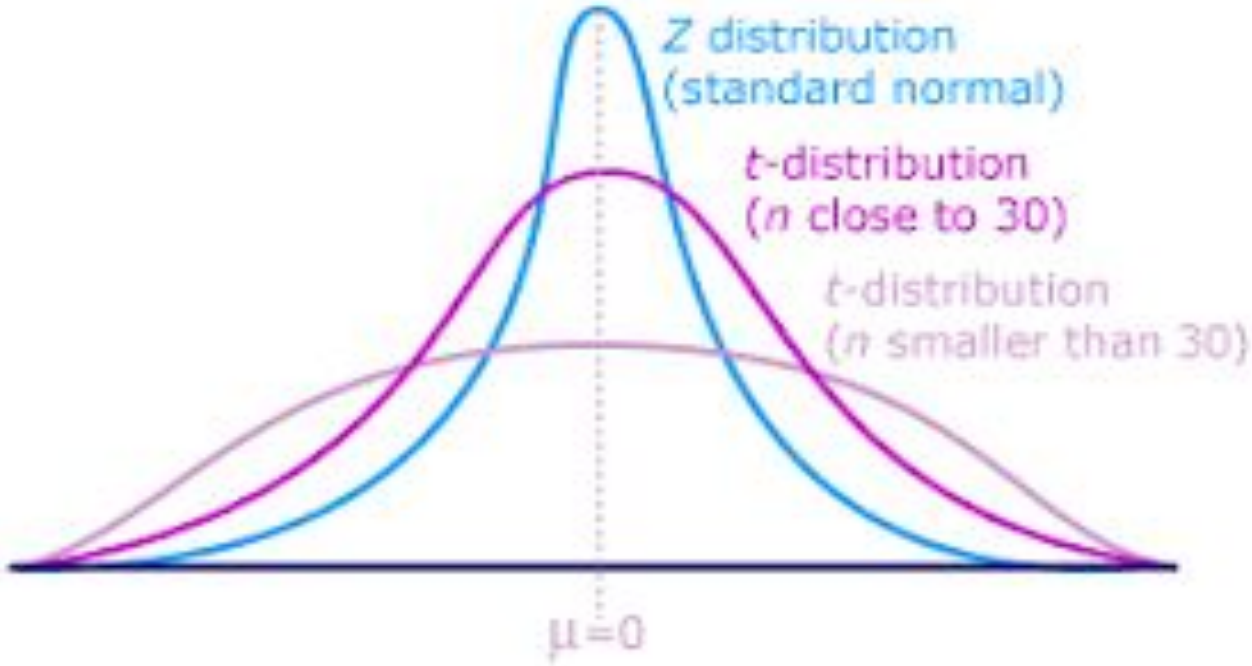
- The **t-test** is used to test hypotheses about means **when the population variance is unknown** and the sample size is small say less than 100.

  Closely related to z, the unit normal.

- Single sample, independent samples, and dependent(paired) samples.

T-Dist Again

- The *t* distribution is the short, fat relative of the normal. The shape of *t* depends on its *df.* As *N* becomes infinitely large, *t* becomes normal.


Z distribution
(standard normal)
t-distribution
(n close to 30)
t-distribution
(n smaller than 30)
μ=0

What Type of T-test is it?

- Single sample t – we have only 1 group; want to test against a hypothetical mean.
- Independent samples t – we have 2 means, 2 groups; no relation between groups, e.g., people randomly assigned to a single group.
- Dependent t – we have two means. Either same people in both groups, or people are related, e.g., husband-wife, hospital patient and visitor.

Tails

One-Tail Test (left tail)	Two-Tail Test	One-Tail Test (right tail)
$H_0 : \mu = \mu_0$	$H_0 : \mu = \mu_0$	$H_0 : \mu = \mu_0$
$H_1 : \mu < \mu_0$	$H_1 : \mu \neq \mu_0$	$H_1 : \mu > \mu_0$
		

Tails and P-Values Two Sample

A **two-tailed P-Value** answers this question:

> Assuming the **null hypothesis** is true, what is the chance that a randomly selected sample would have a mean as far or further than you observed in this experiment with either group having the larger mean?

To interpret a **one-tail P-Value**, you must predict which group will have the larger mean before collecting any data. The **one-tail P-Value** answers this question:

> Assuming the **null hypothesis** is true, what is the chance that a randomly selected samples would have a mean as far apart as (or further) than observed in this experiment with the specified group having the larger mean?

Note

- The one-tail P-Value is half the two-tail P-Value.

- The two-tail P-Value is twice the one-tail P value.

Example

Effect Size