Hypothesis testing

Gedi Lukšys

Applied Data Science 2

October 23, 2023

Example:

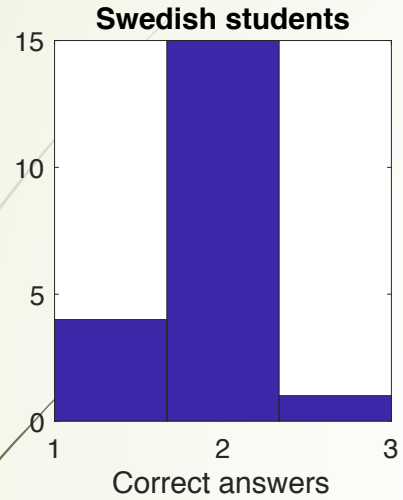
https://www.ted.com/talks/hans_rosling_shows_the_best_stats_you_ve_ever_seen

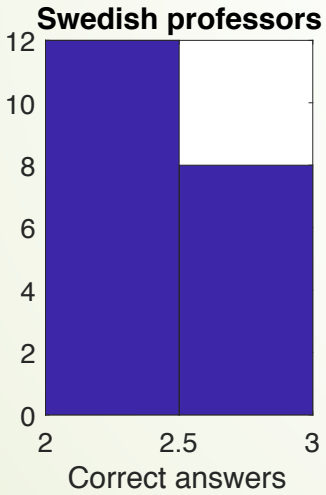
This inspired thinking shared with you by:

The Ultimate Driving Machine

Where great ideas live on.

Example:





https://www.ted.com/talks/hans_and_ola_rosling_how_not_to_be_ignorant_about_the_world

Procedure of hypothesis tests

- Formulate the Null Hypothesis, Alternative Hypothesis
- Design your experiment and collect data
- Summarise and describe your data
- Think about what you would expect if H_0 were true.
- Could your data be explained by the Null Hypothesis?
- Determine the probability of your data given H_0
- Interpret your p value and make a decision
- Be aware that hypothesis tests are not perfect

Null and alternative hypotheses

- **Null hypothesis: nothing is happening**
- In our example: Swedish students are able to distinguish child mortality between countries where it differs with ratio 2:1 **same as** chimpanzees

- **Alternative hypothesis: something is happening**
- In our example: Swedish students are able to distinguish child mortality between countries where it differs with ratio 2:1 **worse than** chimpanzees

T-test $p < 10^{-5}$ => reject null hypothesis

Statistical models: what are they?

Extract patterns from available data (*training data*)

- Can be used to predict new data (*testing data*)

Are all wrong…

…but some are useful! *(George Box)*

What are their main elements?

- **The generating process**, e.g. y = f(x)
- **Parameters** that control it, e.g.

$$f(x \mid \mu, \sigma^2) = \frac{1}{\sqrt{2\pi\sigma^2}} e^{-\frac{(x-\mu)^2}{2\sigma^2}}$$

Null and alternative hypotheses – formally speaking

- **Null hypothesis:** no difference between two parameters or a parameter and a constant
- **Alternative hypothesis:** difference between two parameters or a parameter and a constant

Types of tests

- **Two-tailed**: can be less or more (default)
- **One-tailed** (need to be justified):
  - One should be bigger than the other (*right tailed*)
  - One should be smaller than the other (*left tailed*)

Other examples: null

Null Hypothesis (H_0): "Nothing is happening"

- The effect of the medication is the same as that of a placebo.
- Lack of sleep does not affect cognitive function.
- Wildtype mice and FMRP knockout mice show the same level of synaptic activity.
- Rates of TB infection have not changed between 1920 and 2010.

Other examples: alternative

Null Hypothesis (H_0): "Nothing is happening"

- The effect of the medication is the same as that of a placebo.
- Lack of sleep does not affect cognitive function.
- Wildtype mice and FMRP knockout mice show the same level of synaptic activity.
- Rates of TB infection have not changed between 1920 and 2010.

Alternative hypothesis (H_A): "Something is happening"

- The medication has a different effect than a placebo.
- Lack of sleep affects cognitive function.
- Wildtype mice and FMRP knockout mice show different levels of synaptic activity.
- Rates of TB infection changed between 1920 and 2010.

Statistical tests: the procedure

- Use data obtained from sample(s) to make a decision whether or not to reject the null hypothesis

Swedish students sample:

2 1 2 2 1 2 1 2 2 1 2 2 1 2 3 2 2 2 2 2

Swedish professors sample:

2 3 2 3 2 3 2 2 2 2 3 2 3 2 3 2 3 2 2 3

Chimpanzees?

Presumably

50-50!...

https://www.naturalworldsafaris.com/wildlife/chimpanzees

https://media.treehugger.com/assets/images/2011/10/psychic-german-octopus-world-cup-football.jpg

Really? ☺

Example: t-tests

Swedish students sample:

2 1 2 2 1 2 1 2 2 1 2 2 1 2 3 2 2 2 2 2

Swedish professors sample:

2 3 2 3 2 3 2 2 2 2 3 2 3 2 3 2 3 2 2 3

Chimpanzees? Presumably 50% chance!...

Students vs. Professors: **$p = 7 * 10^{-4}$**

Students vs. Chimpanzees: **$p = 9 * 10^{-6}$**

Professors vs. Chimpanzees: **$p = 0.38$**

What is P value?

- **Probability** of observing a value **as or more extreme** as the one you observed **if the null hypothesis were true**.

What do we do with it?

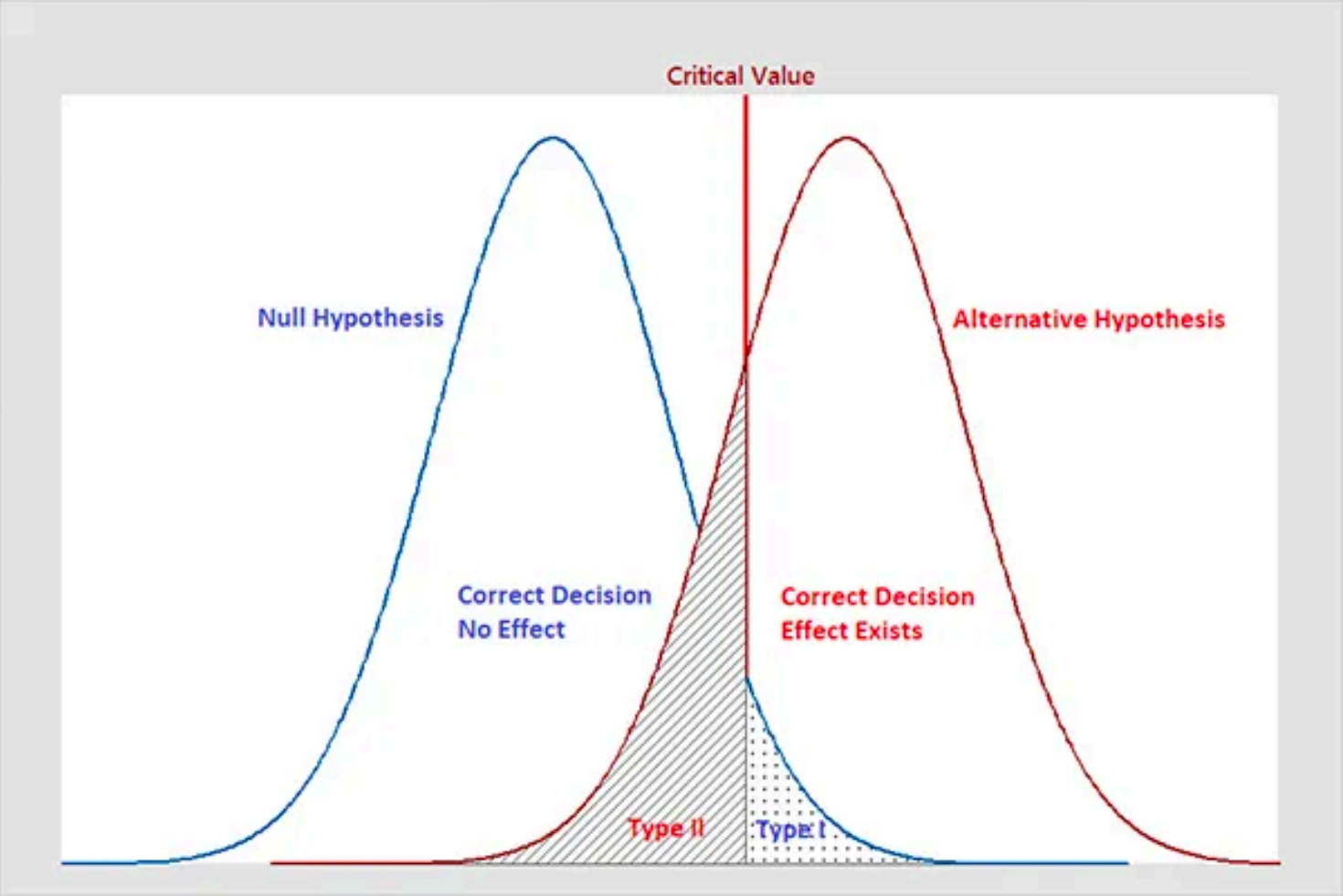
Level of significance – convention (no magic!)

- **p < 0.05: reject the null hypothesis**
- **p > 0.05: cannot reject the null hypothesis** (not the same as confirm it!)

Different outcomes

	Something is happening	**Nothing is happening**
We found something	True positive (correct)	False positive (**Type 1 error**)
We found nothing	False negative (**Type 2 error**)	True negative (correct)

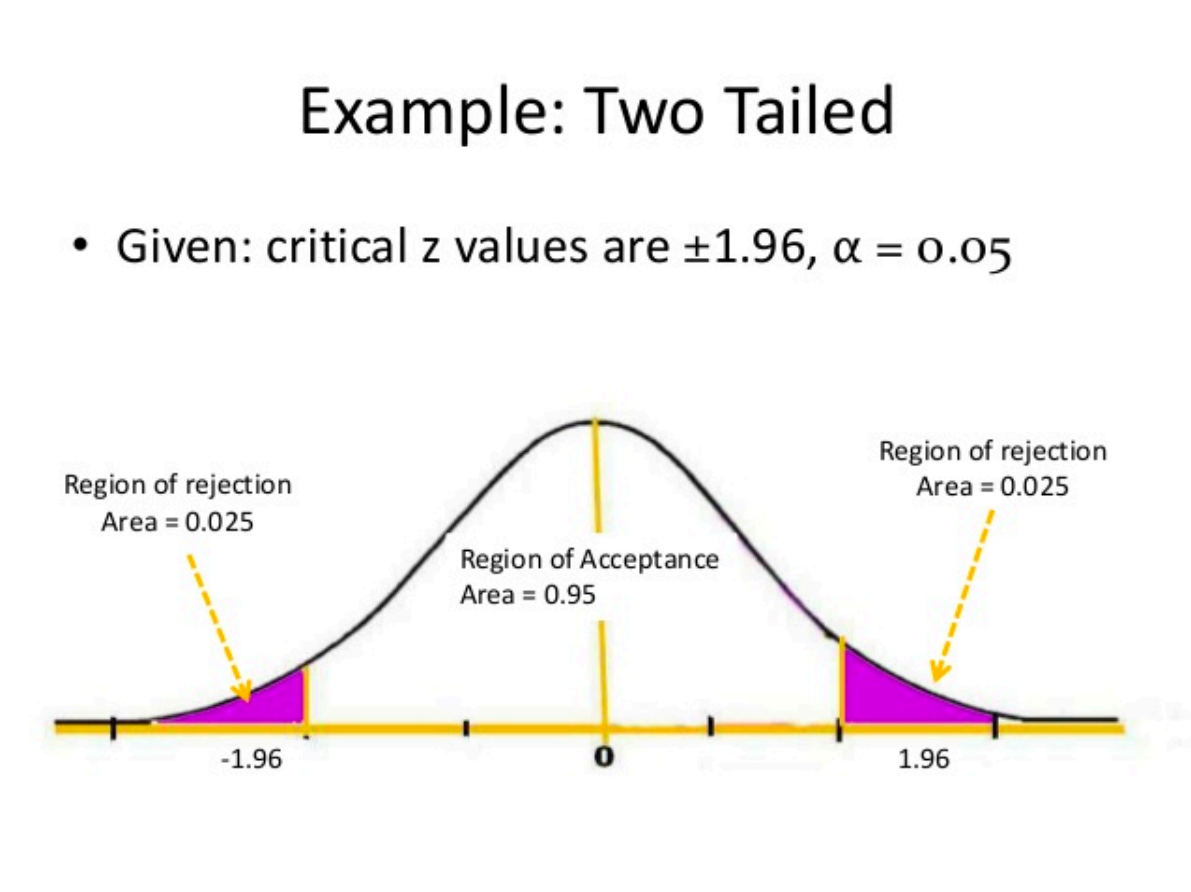
Type 1 and 2 errors



Where should the critical value be?

https://i1.wp.com/statisticsbyjim.com/wp-content/uploads/2018/07/TypesErrorHypothesisTests.png?fit=600%2C400

Critical values and critical regions



- Significance level need not always be 0.05!

https://image.slidesharecdn.com/statisticshypothesistesting-130115004420-phpapp02/95/statistics-hypothesis-testing-14-638.jpg?cb=1358210733

Hypothesis testing steps

- Formulate a hypothesis
- Choose a significance level
- Calculate the critical value
- Calculate the test statistic
- Calculate p value
- **Make a decision**

Buridan's Donkey

https://mayooshin.com/wp-content/uploads/2018/11/boridans-ass.jpg