Data Science for Everyone

Week 10: Permutation testing

Angela Lai

New York University

Outline

- Logistics
- Demo with notebook
- Questions?

Logistics

- Lab 7 out, due at 8 p.m. ET on April 15
- Homework 3/4 will be out tonight, due at 8 p.m. ET on April 27

Logistics

- If you filled out the Google Form for feedback on a data set, you should have gotten a response last night
- Note this project requirement: ”You may select any topic and use any dataset that you like as long as it’s publicly available and it contains **two continuous variables whose association you are interested in examining.**”
- The focus of the project analysis is simple linear regression

Concept Review

Researchers are conducting a test of hypotheses using 3% as the cutoff for the P-value.

If the null hypothesis is true, the chance that the test reaches the correct conclusion is

(a) 100%

(b) 0%

(c) 2%

(d) 3%

(e) 97%

(f) 50%

Concept Review

Researchers are conducting a test of hypotheses using 3% as the cutoff for the P-value.

If the null hypothesis is true, the chance that the test reaches the correct conclusion is

(a) 100%

(b) 0%

(c) 2%

(d) 3%

(e) **97%**

(f) 50%

Concept Review

The P-value of the test comes out to be 2%. The conclusion of the test is that

(a) the data support the null hypothesis more than they support the alternative, because there is a 98%chance that the null hypothesis is true.

(b) the data support the alternative hypothesis more than they support the null, because if the null were true then something unlikely has occurred

(c) the data support the alternative more than they support the null, because there is only a 2% chance that the null hypothesis is true

Concept Review

The P-value of the test comes out to be 2%. The conclusion of the test is that

(a) the data support the null hypothesis more than they support the alternative, because there is a 98%chance that the null hypothesis is true.

(b) **the data support the alternative hypothesis more than they support the null, because if the null were true then something unlikely has occurred**

(c) the data support the alternative more than they support the null, because there is only a 2% chance that the null hypothesis is true

Demo Time

Time for a demo using Will's notebook!

Hypothesis Testing

Reminder:

	Null is true	Null is false
Fail to reject null	Correct inference	Type II Error
Reject null	Type I Error	Correct inference

Questions?

Any questions?