Inference 1: Hypothesis testing (cont)

Last updated: September 18, 2017

Which test to use

- Depends on data
- S (test stat) depends on H_A (alternative)
- justifying the assumptions = CRUCIAL.

Tests we learn	num	cat	cat. vs cat.	num. vs cat.
Fisher's exact			✓	
permutation (randomization)	✓	✓	✓	✓
chi-square			✓	
z-test for proportions		✓		
t-test	✓			
two-sample t-test				✓
ANOVA				✓

Permutation test: computability $>$ precision

Fisher's exact test failed to compute on `class` vs `survived`. What to do?
Permutation test: puts computability first.
Key ideas:

- Approximate $\mathbb{P}(S \geq s | H_0)$ by simulating the distribution of S under H_0.
- Simulate S under H_0 by resampling the data without replacement.

Data example: see code.
Advantages?

Disadvantages?

Permutation test: computability $>$ precision

Fisher's exact test failed to compute on `class` vs `survived`. What to do?
Permutation test: puts computability first.
Key ideas:

- Approximate $\mathbb{P}(S \geq s | H_0)$ by simulating the distribution of S under H_0.
- Simulate S under H_0 by resampling the data without replacement.

Data example: see code.
Advantages?

- Replace math (thinking) by simulation (brute force)
- Always return an answer
- Apply to basically any situation (any data type, any H_0/H_A, any test statistic)

Disadvantages?

Permutation test: computability $>$ precision

Fisher's exact test failed to compute on `class` vs `survived`. What to do?
Permutation test: puts computability first.
Key ideas:

- Approximate $\mathbb{P}(S \geq s | H_0)$ by simulating the distribution of S under H_0.
- Simulate S under H_0 by resampling the data without replacement.

Data example: see code.
Advantages?

- Replace math (thinking) by simulation (brute force)
- Always return an answer
- Apply to basically any situation (any data type, any H_0/H_A, any test statistic)

Disadvantages?

- May not have nice theoretical properties (eg: not maximal power etc)
- If distribution is complicated, must sample for a long time
- May not be reproduceable exactly: different simulation could give slightly different p-values