DATA 8
Spring 2022

Lecture 21

Examples

Announcements

More on Hypothesis Tests

Discussion Question

- Manufacturers of Super Soda run a taste test
- 91 out of 200 tasters prefer Super Soda over its rival

Question: Do fewer people prefer Super Soda than its rival, or is this just chance?

Null hypothesis: The same proportion of people prefer Super as Rival

Alternative hypothesis: A smaller proportion of people prefer Super

Test statistic: Number of people (out of 200) who prefer Super

p-value: Start at the observed statistic and look which way? Left

(Demo)

Hypothesis Test Concerns

The outcome of a hypothesis test can be affected by:

- The hypotheses you investigate:
  How do you define your null distribution?
- The test statistic you choose:
  How do you measure a difference between samples?
- The empirical distribution of the statistic under the null:
  How many times do you simulate under the null distribution?
- The data you collected:
  Did you happen to collect a sample that is similar to the population?
- The truth:
  If the alternative hypothesis is true, how extreme is the difference?

Hypothesis Test Effects

Number of simulations: Make it as large as possible so that the empirical distribution of the test statistic under the null distribution is good. No new data needs to be collected.

Number of observations: A larger sample will lead you to reject the null more reliably if the alternative is in fact true.

Difference from the null: If the null hypothesis is false, but the truth is similar to the null hypothesis, then even a large sample may not provide enough evidence to reject the null.