DATA 8

Fall 2016

Lecture 36, November 21

A/B Testing

Slides created by Ani Adhikari and John DeNero

Announcements

- Project 3 checkpoint Tuesday 11/22 (tomorrow), final deadline Tuesday 11/29
- Current homework:
  - Early submission: Wed 11/23 (usual schedule)
  - “Regular” submission: Monday 11/28 after the break
- GSI/Tutor office hours today 2-5 Etcheverry 3106, tomorrow 2-5 458 Evans and B6 Evans

A/B Testing

- Two random samples:
  - Sample A
  - Sample B

- Question: Are they drawn from the same underlying distribution?

- Answer by **A/B testing**

The Hypotheses

- **Null:**
  - The two samples are drawn from the same underlying population distribution; they look like random draws from the same set.
- **Alternative:**
  - The samples are drawn from different distributions; they don’t look like random draws from the same set.

Permutation Test

- **Null:** The two samples are drawn randomly from the same underlying distribution.
- If the null is true, all rearrangements of the variable values among the two samples are equally likely. So:
  - compute the observed test statistic
  - then shuffle the attribute values and recompute the statistic; **repeat**; compare with the observed statistic

The Test Statistic

- If the samples are categorical, then for this null and alternative hypothesis a natural test statistic is the total variation distance. It measures the difference between the distributions in the two samples.

(Demo)

- If the samples are numerical, often a simpler statistic is just fine, such as the absolute difference between the two sample means.

(Demo)

How Big is the Difference?

If you think that the two underlying population means might be different, you'll want to know how different they are.

- So instead of just running a "same/different" test, don't make any hypotheses. Just estimate the difference between the two population means.
- You can do this by bootstrapping the sample and constructing a confidence interval for the parameter: "difference between the population means". (Demo)