

PSC 400

SYRACUSE UNIVERSITY

DATA ANALYTICS FOR POLITICAL SCIENCE

QUANTIFYING UNCERTAINTY

CI SAMPLE MEAN

95% CONFIDENCE INTERVAL
FOR THE SAMPLE MEAN

$$\left[\bar{Y} - 1.96 \times \sqrt{\frac{\text{var}(Y)}{n}}, \quad \bar{Y} + 1.96 \times \sqrt{\frac{\text{var}(Y)}{n}} \right]$$

CI DIFFERENCE IN MEANS

95% CONFIDENCE INTERVAL FOR THE DIFFERENCE-IN-MEANS ESTIMATOR

LOWER LIMIT:

$$\bar{Y}_{\text{treatment group}} - \bar{Y}_{\text{control group}} - 1.96 \times \sqrt{\frac{\text{var}(Y_{\text{treatment}})}{n_{\text{treatment group}}} + \frac{\text{var}(Y_{\text{control}})}{n_{\text{control group}}}}$$

UPPER LIMIT:

$$\bar{Y}_{\text{treatment group}} - \bar{Y}_{\text{control group}} + 1.96 \times \sqrt{\frac{\text{var}(Y_{\text{treatment}})}{n_{\text{treatment group}}} + \frac{\text{var}(Y_{\text{control}})}{n_{\text{control group}}}}$$

EXERCISE

- **UA_survey.csv**
- **Compute difference-in-means for pro-Russian vote between those with and without access to Russian TV**
- **Compute the 95% confidence interval of that difference**

POPULATION VS. SAMPLE, AGAIN

- Want to know: does Russian TV have effect on pro-Russian votes in the *population*?
- We only have data from a *random sample*
- Idea: Use relation between two variables in *sample* to make inference about relation between two variables in *population*
 - Of course, means we can make mistakes

NULL HYPOTHESIS

- In the population, there is *no relationship* between dependent and independent variable
 - H_0

ALTERNATIVE HYPOTHESIS

- There *is* a relationship between the independent and dependent variable in the population
 - H_a or H_1

ERRORS

	There Is A Relation In The Population	There Is No Relation In The Population
We Conclude There Is A Relation	✓	✗ Type I
We Conclude There Is No Relation	✗ Type II	✓

TYPE I ERROR

- We conclude there is a relationship between X and Y when in reality there is not
 - "Type I error"
 - We falsely reject H_0

TYPE II ERROR

- We conclude there is no relationship between X and Y when in reality there is
 - "Type II error"
 - We falsely do not reject H_0

DECISION

- It's really bad if we conclude there is a relationship when in reality there is not
- Type I error: falsely rejecting H_0
- We only want to reject H_0 based on our sample if chance of committing Type I error is relatively small
 - Typically: 5% or less