PSC 400 SYRACUSE UNIVERSITY

DATA ANALYTICS FOR POLITICAL SCIENCE

QUANTIFYING UNCERTAINTY

CI SAMPLE MEAN

95% CONFIDENCE INTERVAL FOR THE SAMPLE MEAN

$$\overline{Y} - 1.96 \times \sqrt{\frac{var(Y)}{n}}, \quad \overline{Y} + 1.96 \times \sqrt{\frac{var(Y)}{n}}$$

CI DIFFERENCE IN MEANS

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

LOWER LIMIT:

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

UPPER LIMIT:

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

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

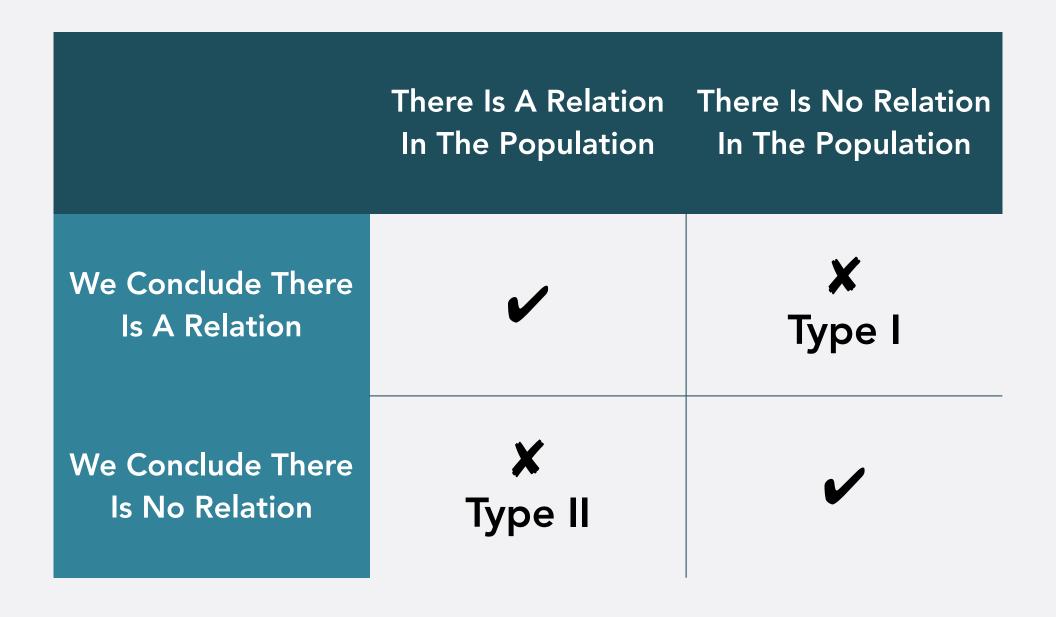
NULL HYPOTHESIS

- In the population, there is *no relationship* between dependent and independent variable
 - H₀

ALTERNATIVE HYPOTHESIS

- There is a relationship between the independent and dependent variable in the population
 - H_a or H₁

ERRORS



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 H₀

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 H₀

DECISION

- It's really bad if we conclude there is a relationship when in reality there is not
- Type I error: falsely rejecting H₀
- 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