Tutorial 3

Research Methods for Political Science - PO3110

Andrea Salvi

7 & 8 October 2019

Trinity College Dublin,
https://andrsalvi.github.io/research-methods/

Table of contents

Research Paper

- Introduction: Explanatory research question and motivation

- Short discussion of the literature and outline of theoretical argument

- Research Design, including type of data; operationalisation (valid and reliable!); data-sets; clear and testable hypotheses

- Analyze your data! What can we conclude? Are the hypotheses confirmed?

- Please indicate how each of you has contributed

Key Concepts from the Lecture

Confidence Intervals

- For a given statistic calculated for a sample of observations (e.g. sample mean) the confidence interval (CI) is a range of values around that statistic that are believed to contain, with a certain probability (e.g. 95%), the true value of that statistic (i.e. the population value).

Confidence Intervals

- For a given statistic calculated for a sample of observations (e.g. sample mean) the confidence interval (CI) is a range of values around that statistic that are believed to contain, with a certain probability (e.g. 95%), the true value of that statistic (i.e. the population value).
- For short: we are 95% confident that the mean value of x is between 13 and 15

Calculating CI95

1. Estimate Mean: $\bar{x} = \frac{\sum x}{n}$

Calculating CI95

1. Estimate Mean: $\bar{x} = \frac{\sum x}{n}$
2. Estimate Standard Deviation: $\sigma = \sqrt{\frac{\sum (x - \bar{x})^2}{n-1}}$

Calculating CI95

1. Estimate Mean: $\bar{x} = \frac{\sum x}{n}$
2. Estimate Standard Deviation: $\sigma = \sqrt{\frac{\sum (x - \bar{x})^2}{n-1}}$
3. Estimate standard error of the mean: $sd(\bar{X}) = \frac{\sigma}{\sqrt{n}}$

Calculating CI95

1. Estimate Mean: $\bar{x} = \frac{\sum x}{n}$
2. Estimate Standard Deviation: $\sigma = \sqrt{\frac{\sum (x - \bar{x})^2}{n-1}}$
3. Estimate standard error of the mean: $sd(\bar{X}) = \frac{\sigma}{\sqrt{n}}$
4. Confidence Interval: CI95 = $\bar{x} \pm 1.96 * sd(\bar{X})$

Calculating CI95

1. Estimate Mean: $\bar{x} = \frac{\sum x}{n}$
2. Estimate Standard Deviation: $\sigma = \sqrt{\frac{\sum (x - \bar{x})^2}{n-1}}$
3. Estimate standard error of the mean: $sd(\bar{X}) = \frac{\sigma}{\sqrt{n}}$
4. Confidence Interval: CI95 = $\bar{x} \pm 1.96 * sd(\bar{X})$

Hands-on exercise:

Calculating CI95

1. Estimate Mean: $\bar{x} = \frac{\sum x}{n}$
2. Estimate Standard Deviation: $\sigma = \sqrt{\frac{\sum (x - \bar{x})^2}{n-1}}$
3. Estimate standard error of the mean: $sd(\bar{X}) = \frac{\sigma}{\sqrt{n}}$
4. Confidence Interval: CI95 = $\bar{x} \pm 1.96 * sd(\bar{X})$

Hands-on exercise:

- $\bar{x}$: 170

Calculating CI95

1. Estimate Mean: $\bar{x} = \frac{\sum x}{n}$
2. Estimate Standard Deviation: $\sigma = \sqrt{\frac{\sum (x - \bar{x})^2}{n-1}}$
3. Estimate standard error of the mean: $sd(\bar{X}) = \frac{\sigma}{\sqrt{n}}$
4. Confidence Interval: CI95 = $\bar{x} \pm 1.96 * sd(\bar{X})$

Hands-on exercise:

- $\bar{x}$: 170
- σ: 10

Calculating CI95

1. Estimate Mean: $\bar{x} = \frac{\sum x}{n}$
2. Estimate Standard Deviation: $\sigma = \sqrt{\frac{\sum (x - \bar{x})^2}{n-1}}$
3. Estimate standard error of the mean: $sd(\bar{X}) = \frac{\sigma}{\sqrt{n}}$
4. Confidence Interval: CI95 = $\bar{x} \pm 1.96 * sd(\bar{X})$

Hands-on exercise:

- $\bar{x}$: 170
- σ: 10
- $n = 30$

Calculating CI95

1. Estimate Mean: $\bar{x} = \frac{\sum x}{n}$
2. Estimate Standard Deviation: $\sigma = \sqrt{\frac{\sum (x - \bar{x})^2}{n-1}}$
3. Estimate standard error of the mean: $sd(\bar{X}) = \frac{\sigma}{\sqrt{n}}$
4. Confidence Interval: CI95 = $\bar{x} \pm 1.96 * sd(\bar{X})$

Hands-on exercise:

- $\bar{x}$: 170
- σ: 10
- $n = 30$
- We can calculate $sd(\bar{X})$

Calculating CI95

1. Estimate Mean: $\bar{x} = \frac{\sum x}{n}$
2. Estimate Standard Deviation: $\sigma = \sqrt{\frac{\sum (x - \bar{x})^2}{n-1}}$
3. Estimate standard error of the mean: $sd(\bar{X}) = \frac{\sigma}{\sqrt{n}}$
4. Confidence Interval: CI95 = $\bar{x} \pm 1.96 * sd(\bar{X})$

Hands-on exercise:

- $\bar{x}$: 170
- σ: 10
- $n = 30$
- We can calculate $sd(\bar{X})$
- Task: Estimate the 95 % confidence intervals

Solution

Solution

1. $sd(\bar{X}) = \frac{10}{\sqrt{30}} = 1.83$

Solution

1. $sd(\bar{X}) = \frac{10}{\sqrt{30}} = 1.83$
2. $CI = 170 \pm 1.96 * 1.83$

Solution

1. $sd(\bar{X}) = \frac{10}{\sqrt{30}} = 1.83$
2. $CI = 170 \pm 1.96 * 1.83$
3. $CI_{low} = 170 - 1.96 * 1.83 = 166.41$

Solution

1. $sd(\bar{X}) = \frac{10}{\sqrt{30}} = 1.83$
2. $CI = 170 \pm 1.96 * 1.83$
3. $CI_{low} = 170 - 1.96 * 1.83 = 166.41$
4. $CI_{high} = 170 + 1.96 * 1.83 = 173.58$

Furthering the interpretation

Let's reverse the example:

- We have a normally distributed population with $\bar{x}$: 170 and σ: 10.
- We then draw a sample with $n = 30$.
- Are we likely to find a sample mean of 172?

In-class Exercise

In-class Exercise

- Download ”parlgov elections.xlsx”
  `https://tinyurl.com/datamt2`
- Download Parlgov Elections Codebook
  `https://tinyurl.com/codebookmt2`
- You can work in pairs

In-class Exercise

- Load the Data
- What do they represent? What are the observations here?
- Describe the variables we have.
- How are missing values coded?
- Subset the data in order to select just the Irish elections.
- Arrange the data based on ”left_right” from Left to Right.
- Let’s take a random sample of size 30.
- Using that sample calculate the mean, standard deviation and standard error of the mean of ”vote_share” then calculate a CI95 for the mean.

In-class Exercise (from last tutorial)

- Re-code the ”election_type” variable into a numeric variable called ”election_type_num” where ”1” is ”parliament” and 2 is ”ep”. Assign them a label accordingly.
- Create a new binary variable (0,1) called ”left_right_binary” where ”right” = 1 and ”left” = 0 . How would you do that? (at least two ways!)
- ”Split” the data-set based on the ”election_type” variable (Hint: Data -> Split File). Now try to calculate the mean, standard deviation and Standard Error of the Mean for the ”vote_share” variable. What happened?
- Plot a histogram of your choice that conveys meaningful information.
- How do you get rid of the ”split”?