

Tutorial 01, Hilary Term

Research Methods for Political Science (PO3600)

Stefan Müller

23 January 2018

Trinity College Dublin

<http://muellerstefan.net/research-methods>

1. Recap
2. Plotting relationships between variables
3. Calculating correlations
4. Computing t-tests and confidence intervals

Monday, 29th of January, 11-12am, Room ARTS 3020

- Go to the course page on Blackboard and download the `norris.sav` dataset
- Create descriptive statistics for `Co2_2001` and `Power2001`.
- What do these variables mean and how are the variables distributed?

- Use Co2_2001 and compute a 95% confidence interval for the mean of this variable. Interpret the result.
- Compute a 90% confidence intervals and compare it with the 95% confidence intervals. Which one is wider and why?

Recap: Important formulas

Variance of sample: average squared deviation from the mean)

$$\sigma^2 = \frac{\sum_{i=1}^n (x_i - \bar{x})^2}{n - 1} \quad (1)$$

Standard deviation of the mean: square root of the variance

$$\sigma = \sqrt{\sigma^2} \quad (2)$$

Standard error of the mean: expected value of the standard deviation of means of several samples

$$SE = \frac{s}{\sqrt{n}} \quad (3)$$

Recap: Confidence interval

$$CI = \bar{x} \pm z \times SE \quad (4)$$

Recap: Confidence interval

$$CI = \bar{x} \pm z \times SE \quad (4)$$

Confidence level determined through z . 95% confidence interval has a z -score of 1.96, 90% has one of 1.645 and 99% has a z score of 2.58.

Hypothesis: The average CO2 emissions in 2001 does not equal 20 metric tons per capita.

$$H_0 : \bar{x} = 20$$

$$H_1 : \bar{x} \neq 20$$

- Compute a one-sample t-test. Compare the results from the confidence interval and the t-test. What information is revealed by each?

Recap: One-sample t-test

$$t = \frac{\bar{x} - \mu}{\sigma / \sqrt{n}}, \quad (5)$$

Recap: One-sample t-test

$$t = \frac{\bar{x} - \mu}{\sigma / \sqrt{n}}, \quad (5)$$

- \bar{x} : sample mean
- μ : assumed population mean
- n : sample size
- σ : standard deviation

If the absolute value of the t-test statistics is greater than the critical value, then the difference is significant. Otherwise it is not!

Two-sample t-test

- Use the variables Co2_2001 and OECD.
- Compute and interpret a two-sample *independent* t-test.
- Use the variables Co2_1998 and Co2_2001.
- Compute and interpret a *paired* sample t-test.

Recap: Two-sample t-test

$$t = \frac{\bar{x}_A - \bar{x}_B}{\sqrt{S^2/n_A + S^2/n_B}} \quad (6)$$

S^2 is an estimator of the *common variance* of the two samples.

$$S^2 = \frac{\sum (x - \bar{x}_A)^2 + \sum (x - \bar{x}_B)^2}{n_A + n_B - 2} \quad (7)$$