IPRES Tutorial 10: Sampling and Confidence Intervals

Milan Babic

m.babic@uva.nl
Department of Political Science - Political Economy and Transnational Governance

03 May 2019

Former UvA student **Rami Sidky** detained and tortured in Egypt

- Rami **studied in Amsterdam** in 2014-2015.
- He was **friendly, open, warm and kind** and loved playing guitar.
- On 8 May 2018 he **was arrested** because of a song critical of the Egyptian president.
- A song that he **neither played, produced nor performed**.
- Remains arbitrarily detained in the notorious Tora prison for nearly 1 year now.
- Rami has no access to his lawyer.
- He has been tortured, sleeps on a concrete and dirty floor in a small cell shared with 5 others.
- His psychological and physical state is deteriorating to such an extent that **his life is now directly at risk**.
- **Please sign and share the Amnesty International petition** for his release now:

www.amnesty.nl/rami

Agenda

1.) Check-in and Recap

2.) Samples, Sampling and Sampling Issues

3.) Confidence Intervals

4.) Wrap-up and Questions

m.babic@uva.nl

Check-in (also Google doc)

m.babic@uva.nl

Where we are

Part 2:

- Prerequisites of quantitative research: Experiments and comparative case studies (**Tutorial 7**)
- Descriptive Statistics I: LoM and data hands-on (**Tutorials 8**)
- Descriptive Statistics II: Spreads (**Tutorial 9**)
- Sampling and Confidence Intervals (**Tutorial 10**)
- Inferential Statistics (**Tutorial 11**)
- Limits of numbers and ethics/assignment prep. (**Tutorial 12**)

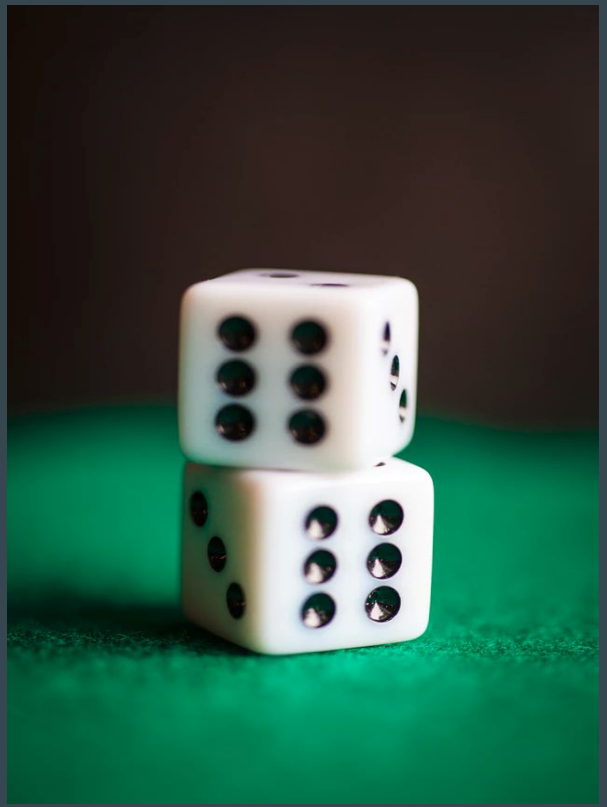
m.babic@uva.nl

Recap

Last week's exercise

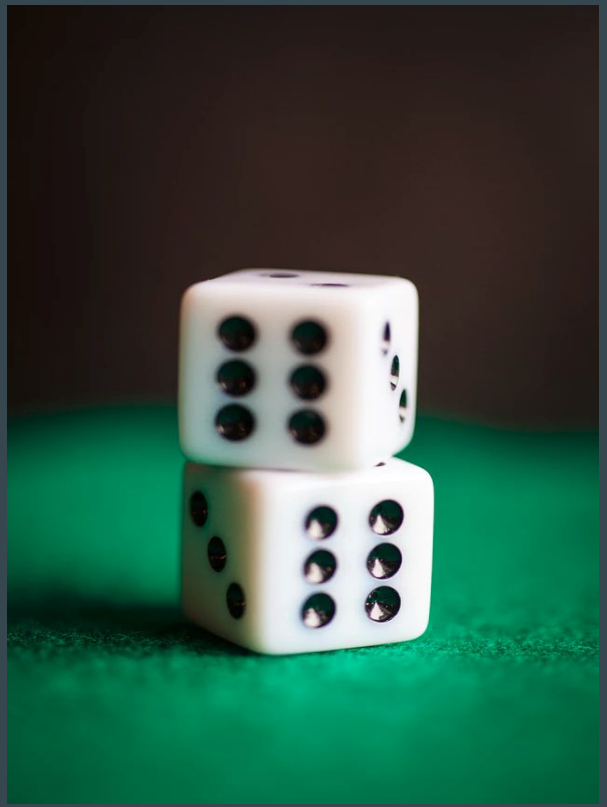
m.babic@uva.nl

Samples and Sampling



m.babic@uva.nl

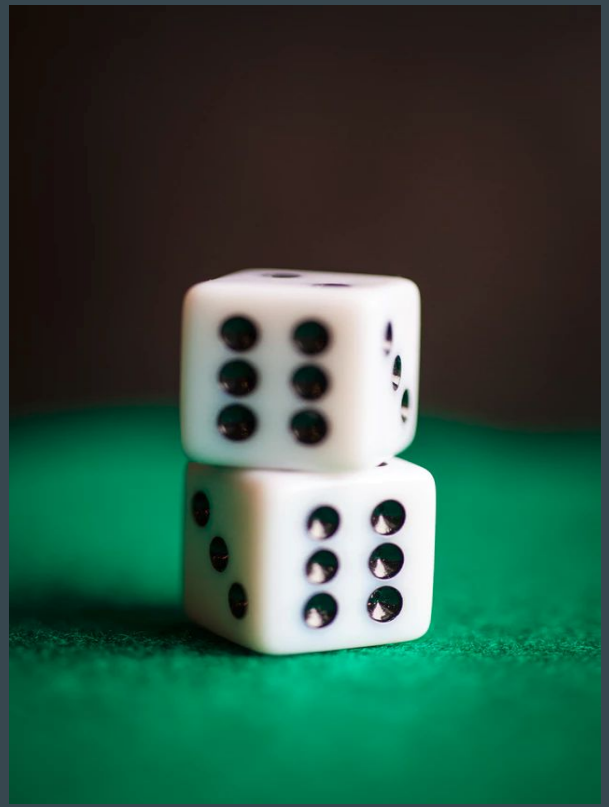
Samples and Sampling



- What is sampling again?
  - why do we sample?
- Types of sampling?
- Types of sampling error?

m.babic@uva.nl

Samples and Sampling



- What is sampling again?
  - why do we sample?
- Types of sampling?
- Types of sampling error?

Check slides no. 11-19 from Tutorial 6!

m.babic@uva.nl

Central Limit Theorem (important)

- Distribution of sample means ≈ normal distribution (BUT: Less variability than samples!)
- Mean of sample means ≈ true population mean

m.babic@uva.nl

Central Limit Theorem (important)

- Distribution of sample means ≈ normal distribution (BUT: Less variability than samples!)
- Mean of sample means ≈ true population mean
- Standard error: $SE = \frac{SD}{\sqrt{n}}$
    - why “error”?

m.babic@uva.nl

Central Limit Theorem (important)

- Distribution of sample means ≈ normal distribution (BUT: Less variability than samples!)
- Mean of sample means ≈ true population mean
- Standard error: $SE = \frac{SD}{\sqrt{n}}$
    - why “error”?
    - be aware: SD = SD of *population*!

m.babic@uva.nl

Central Limit Theorem (important)

- Why do you need this?
  - Testing whether your sample mean is different due to chance or not
    - Foundation of inferential statistics

- How? ⟶ Exercise

m.babic@uva.nl

In-class exercise

- We know that students have a mean disposable income of 1000 Euros per month. Standard deviation is 150.
- The 50 of you have a mean income of 1050 Euros.
    - Is this due to chance or are UvA students filthy rich?

m.babic@uva.nl

In-class exercise

- We know that students have a mean disposable income of 1000 Euros per month. Standard deviation is 150.
- The 50 of you have a mean income of 1050 Euros.
    - Is this due to chance or are UvA students filthy rich?
- Use the CLT:
    - population mean is known: 1050
    - Standard Error: SD/sqrt(n) = 150/sqrt(50) = 21.21

m.babic@uva.nl

In-class exercise

- We know that students have a mean disposable income of 1000 Euros per month. Standard deviation is 150.
- The 50 of you have a mean income of 1050 Euros.
    - Is this due to chance or are UvA students filthy rich?
- Use the CLT:
    - population mean is known: 1050
    - Standard Error: SD/sqrt(n) = 150/sqrt(50) = 21.21
- Then: Z-score
    - Z = (sample mean – pop mean) / SE
    - Z = (1050 – 1000) / 21.21 = **2.36**

m.babic@uva.nl

In-class exercise

- We know that students have a mean disposable income of 1000 Euros per month. Standard deviation is 150.
- The 50 of you have a mean income of 1050 Euros.
    - Is this due to chance or are UvA students filthy rich?
- Use the CLT:
    - population mean is known: 1050
    - Standard Error: SD/sqrt(n) = 150/sqrt(50) = 21.21
- Then: Z-score
    - Z = (sample mean – pop mean) / SE
    - Z = (1050 – 1000) / 21.21 = **2.36**
    - **Check area under the curve for Z-score: it is 0.49086**

m.babic@uva.nl

In-class exercise

- We know that students have a mean disposable income of 1000 Euros per month. Standard deviation is 150.
- The 50 of you have a mean income of 1050 Euros.
    - Is this due to chance or are UvA students filthy rich?
- Use the CLT:
    - population mean is known: 1050
    - Standard Error: SD/sqrt(n) = 150/sqrt(50) = 21.21
- Then: Z-score
    - Z = (sample mean – pop mean) / SE
    - Z = (1050 – 1000) / 21.21 = **2.36**
    - **Check area under the curve for Z-score: it is 0.49086**
    - **Calculate 0.5 - 0.49086 = 0.00914 = 0.9% probability of chance - you are filthy rich!**

m.babic@uva.nl

More exercises: Handout

First: on your own

Then: in pairs or more

m.babic@uva.nl

Confidence Intervals

- Sample means & their distributions give us the best possible idea of the population distribution
  - However: how sure can we be?
    - Literally: *how* sure (number)

m.babic@uva.nl

Confidence Intervals

- Sample means & their distributions give us the best possible idea of the population distribution
  - However: how sure can we be?
    - Literally: *how* sure (number)

- Standard: 95%-confidence interval (i.e. reducing the chance of error to 5%)

$$95\%\,CI = sample\,mean - \left(1.96 \times \frac{SD}{\sqrt{n}}\right)$$

m.babic@uva.nl

Wrap-up

- What we did *not* do today:
  - Talk in-depth about different types of sampling - Learn it from the book and Tutorial 6
  - Use t-tables in addition to normal tables - Same logic, just for non-normal distributions
  - Looked at proportions in addition to means - Same logic, just for categorical variables
  - Used Excel - Your task to exercise at home, also for next week
  - **Everything else that is left in the lecture and book, so check it!**

m.babic@uva.nl

Sources (all pictures CC license)

Image Wave Emoji: https://upload.wikimedia.org/wikipedia/commons/thumb/4/48/Emojione_1F44B.svg/2000px-Emojione_1F44B.svg.png

Image dice: Image dice: Photo by Heather Gill on Unsplash

m.babic@uva.nl