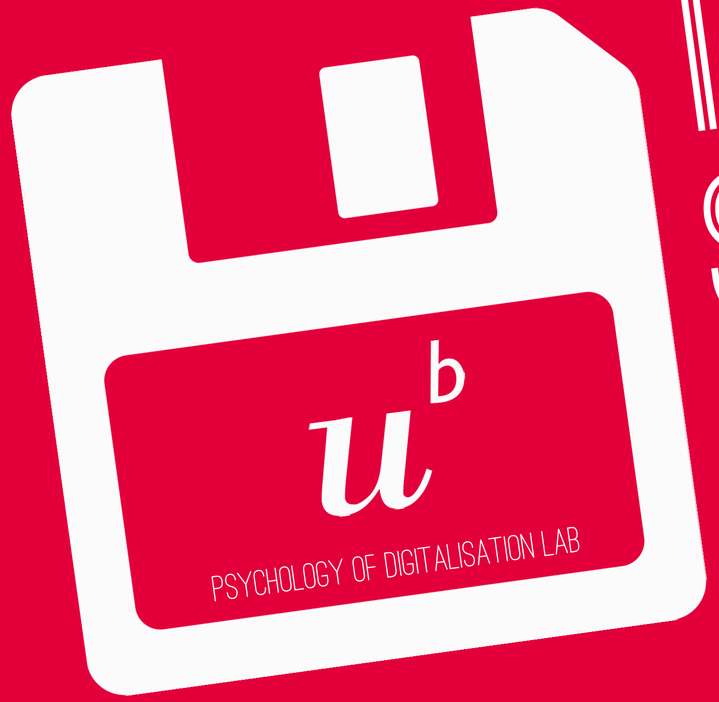


IMPROVING YOUR STATISTICAL INFERENCES

THROUGH MONTE CARLO SIMULATION STUDIES

PD DR. IAN HUSSEY & DR. JAMIE CUMMINS

EVERYTHING IS COMPUTER

Substantive Knowledge

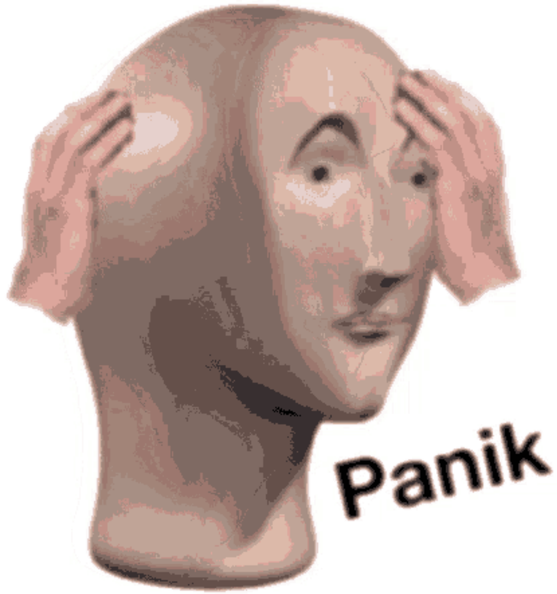
Meta-Science	Privacy and Security	Digitalisation
Psychology of Digitalisation: Computer is Everything (L)		
Digitalisation of Psychology: Everything is Computer (L)		
Psychology in Crisis? (S/MS)	Usable Security (S)	Influencer Marketing (S)
The Science of Behavior (S)	Escape Rooms: Cybersecurity Education at School (S)	How (not) to use AI in research (S/MS)

Technical Skills

Beginner	Intermediate	Advanced
R you Ready? Reproducible Data Preparation and Analyses in R [Beginner] (S/MS)	Transdisciplinary Data Science (L)	Improving your Statistical Inferences through Monte Carlo Simulations (S/MS)
	Estimating the Credibility of Past Research (S/MS)	
	Reproducible Data Processing and Visualisation in R [Intermediate] (S/MS)	

WHY ARE WE HERE

WHY ARE WE HERE



WHY AM I HERE

- I’m a user of stats, not a trained statistician or methodologist.
- I’m a user of code, not trained in software engineering.
- I use simulations to teach myself, and others, about quantitative methods to use them in research.

WHY AM I HERE

- Bad at math

$$\bar{x} = \frac{1}{n}\sum_{i=1}^{n} x_i$$

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

$$SE = \sqrt{\frac{s_1^2}{n_1} + \frac{s_2^2}{n_2}}$$

$$t = \frac{\bar{x}_1 - \bar{x}_2}{SE}$$

$$df = \frac{\left(\frac{s_1^2}{n_1} + \frac{s_2^2}{n_2}\right)^2}{\frac{\left(\frac{s_1^2}{n_1}\right)^2}{n_1 - 1} + \frac{\left(\frac{s_2^2}{n_2}\right)^2}{n_2 - 1}}$$

$$t = \frac{\bar{X} - \mu}{S/\sqrt{n}} \sim t_{n-1}$$

$$f(t) = \frac{\Gamma\left(\frac{\nu+1}{2}\right)}{\sqrt{\nu\pi}\,\Gamma\left(\frac{\nu}{2}\right)}\left(1 + \frac{t^2}{\nu}\right)^{-\frac{\nu+1}{2}}$$

$$\Gamma(n) = (n-1)!$$

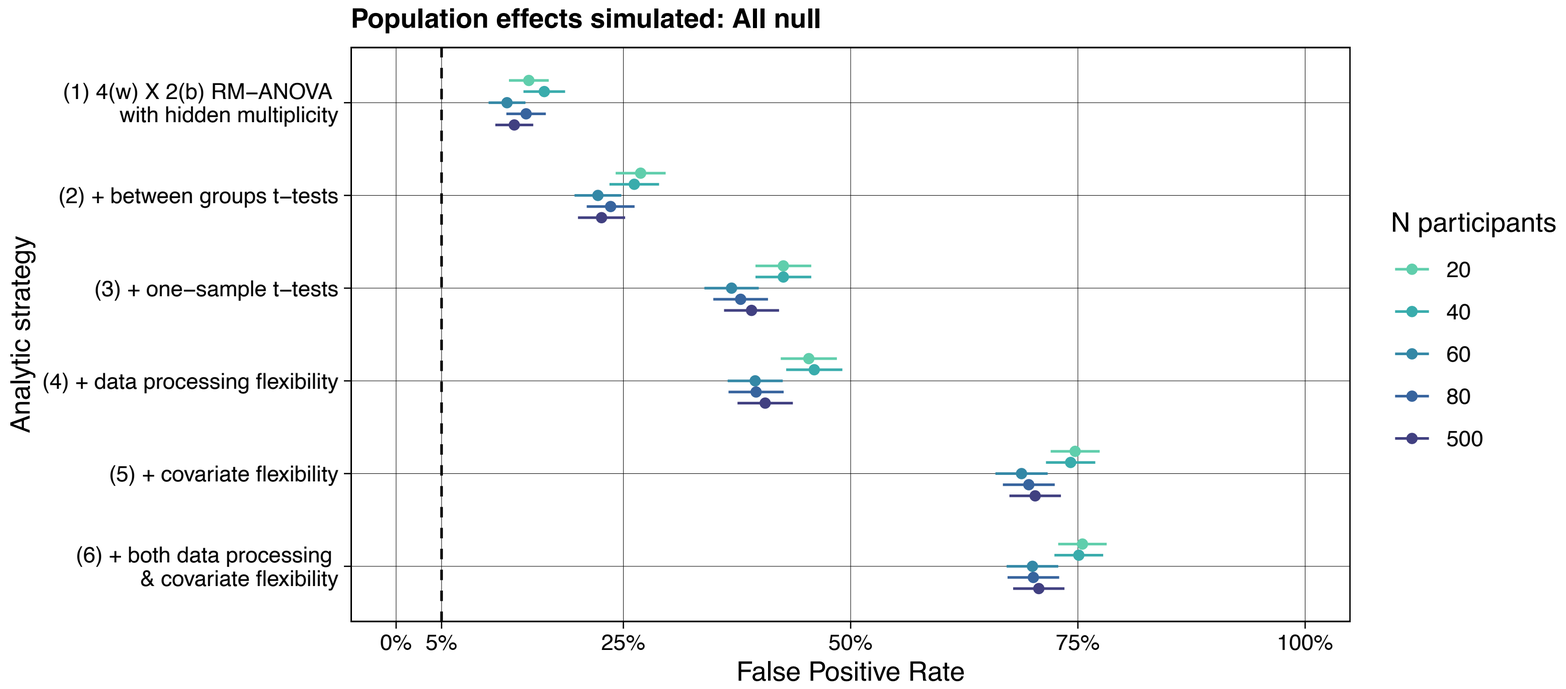
$$p = 2 \cdot P(T_{df} > |t|)$$

WHY AM I HERE

- Understanding the impact of *p*-hacking in the IRAP task

WHY AM I HERE

- Understanding the impact of *p*-hacking in the IRAP task



WHY ARE YOU HERE

- What do you want to get from this course?

- What is your existing skill level?
  - Programming languages
  - Confidence

- What career directions interest you?

GETTING TO KNOW ONE ANOTHER

GETTING TO KNOW ONE ANOTHER

- Aka “help Ian and Jamie integrate 🇨🇭”
- Are Badi-Pommes actually good?

GETTING TO KNOW ONE ANOTHER

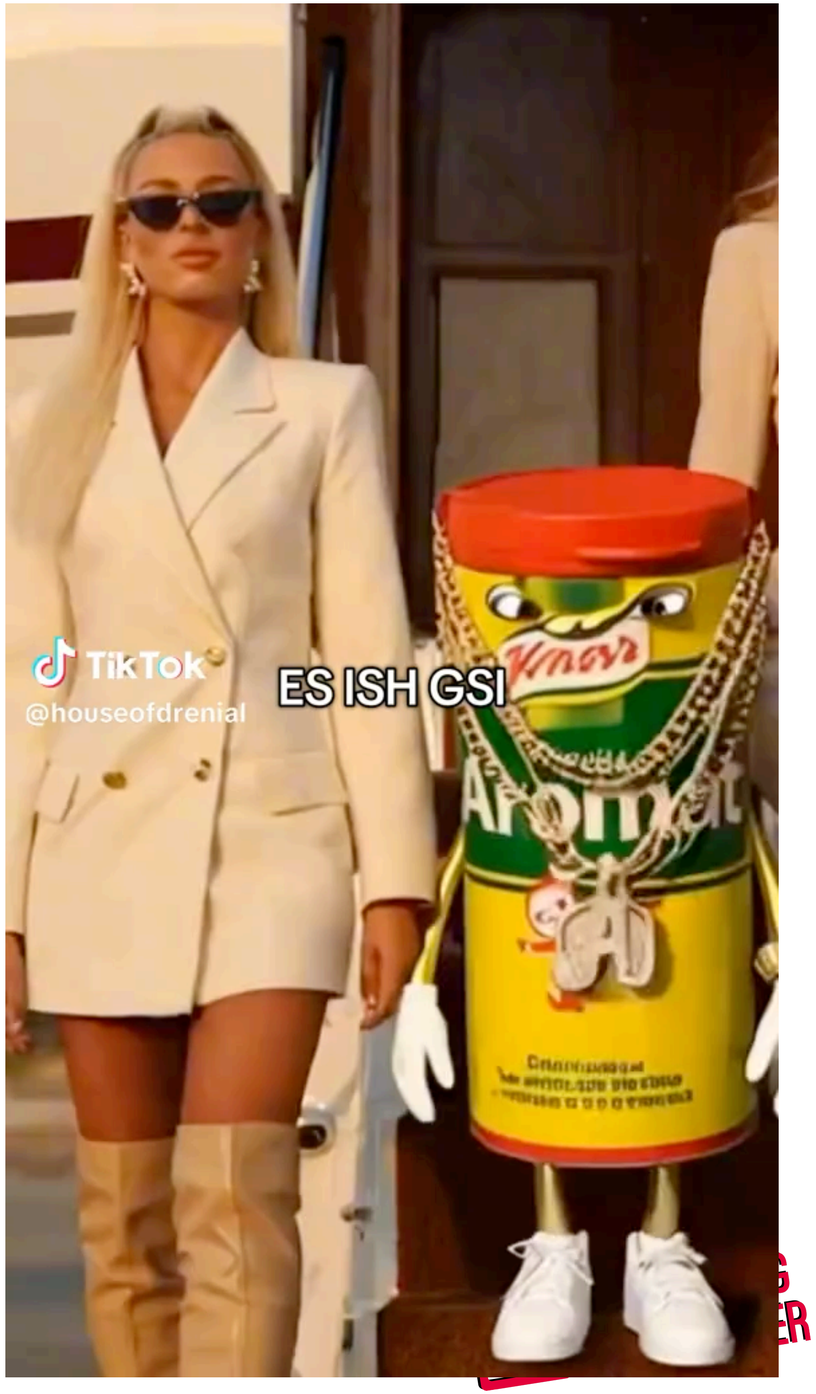
- I don’t know what a Rakete is. There, I said it. Plz explain

GETTING TO KNOW ONE ANOTHER

- What is the most Bünzli thing you have ever done?

GETTING TO KNOW ONE ANOTHER

- Quiz: Does this slap?



GETTING TO KNOW ONE ANOTHER

- Explain this man to Jamie

GETTING TO KNOW ONE ANOTHER

- Communication style
  - Wollen wir uns duzen?

- Preferred names, pronunciation, pronouns
- Accessibility & accommodations

- Languages where we care about mistakes
  - English ❌
  - R/tidyverse ✅

STRUCTURE

GOALS

- ‘Violating assumptions of statistical tests is bad’
  - How bad?
  - Under what circumstances?
  - What should I do instead?

- ‘How do I learn this new method without accidentally *p*-hacking?’
  - Students often learn a method only when the need to use it
  - Significant results are often a ‘stop signal’
    - Inevitably leads to *p*-hacking

GOALS

- Be able to plan, design, implement, interpret, and report simulations
- Gain deeper insight into statistical methods, how (not) to use them
- Acquire the conceptual and technical toolkit; be able learn about other statistical methods via simulation

- “Create the skillsets professors want to hire as PhD students”
  - Or data science etc.

- This course will get harder every year, because you keep impressing us

COMMUNICATION

- Via Slack
  - You have been sent an invite

SCHEDULE

SCHEDULE

#	Date	Topic	Focus
1	18.02.2026	Foundational concepts, writing functions	Learning to simulate
2	25.02.2026	Generating data	Learning to simulate
3	04.03.2026	Analysing data, everything is a linear model	Learning to simulate
4	11.03.2026	Mapping functions, creating experiments	Learning to simulate
5	18.03.2026	Summarizing results, understanding precision and power	Learning to simulate, Learning from simulations
6	25.03.2026	Planning and reporting simulations	Learning to simulate
7	01.04.2026	Understanding, using, misusing *p*-values	Learning from simulations
	08.04.2026	No class (Easter break)	
8	15.04.2026	Violating assumptions, testing parameters other than the mean	Learning from simulations
9	22.04.2026	Testing assumptions and modelling researcher behaviors	Learning from simulations
10	29.05.2026	How (not) to analyze an RCT	Learning from simulations
11	06.05.2026	*p*-hacking and data tampering	Learning from simulations
12	13.05.2026	Selection effects: regression to the mean, range restriction	Learning from simulations
13	20.05.2026	Meta-analysis and publication bias	Learning from simulations
14	27.05.2026	Causal modelling	Learning from simulations

SCHEDULE

SCHEDULE

#	Date	Topic	Focus
1	18.02.2026	Foundational concepts, writing functions	Learning to simulate
2	25.02.2026	Generating data	Learning to simulate
3	04.03.2026	Analysing data, everything is a linear model	Learning to simulate
4	11.03.2026	Mapping functions, creating experiments	Learning to simulate
5	18.03.2026	Summarizing results, understanding precision and power	Learning to simulate, Learning from simulations
6	25.03.2026	Planning and reporting simulations	Learning to simulate
7	01.04.2026	Understanding, using, misusing *p*-values	Learning from simulations
	08.04.2026	No class (Easter break)	
8	15.04.2026	Violating assumptions, testing parameters other than the mean	Learning from simulations
9	22.04.2026	Testi mpt ng researcher behav	Learning from simulations
10	29.05.2026	How (n e	simulations
11	06.05.2026	*p*-hacki ta	simulations
12	13.05.2026	Selection reg restriction	simulations
13	20.05.2026	Meta-analysis and publication bias	Learning from simulations
14	27.05.2026	Causal modelling	Learning from simulations

Vibes

EVERYTHING IS COMPUTER

REQUIREMENTS

- Weekly attendance (80% minimum)
- Laptop + recent version of R, RStudio, & {tidyverse}

- Working knowledge of R, RStudio, and dplyr/tidyr/ggplot2
  - Or willingness to acquire it
  - Catch-up material available

- Practice
  - Semi-weekly exercises
  - Self-guided practice

ASSESSMENT

- End of course assignment
  - Design, implement, and transparently report a simulation study with a meaningful research question not covered in class (e.g., following Siepe et al., 2024)
  - Define scope by discussing it with us one-to-one and using examples from previous years
  - Deadline: [date we all agree on] - Start early! Ask questions!
  - In English (preferably) or German (if necessary)

ASSESSMENT

- 3rd time I have run this course.
- This course is intended to be both difficult and rewarding.
- As my materials become better, my expectations for you rise too.

- I also acknowledge my weaknesses:
  - I’m really bad at putting weekly assignments online on time
  - So I did away with them

ASSESSMENT

- All assessments must be licensed CC BY 4.0
  - i.e., can be used or modified with attribution

- In the service of:
  - Our materials are open source, so will yours be
  - Building a cumulative knowledge base of simulation studies
  - Practicing Open Science skills

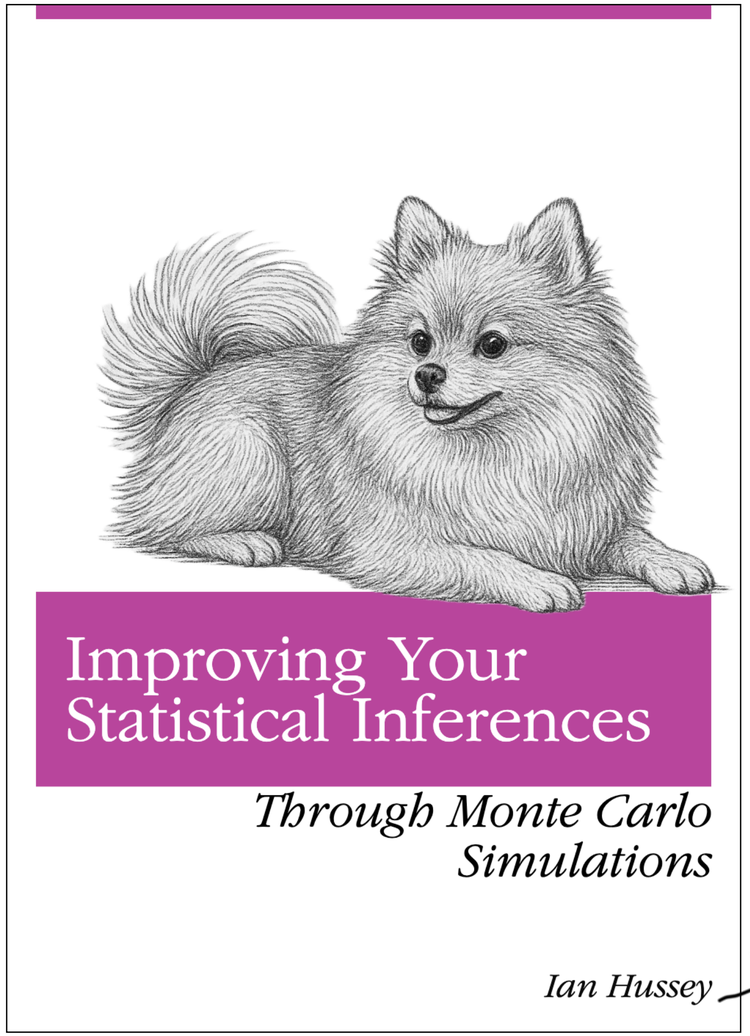
AI

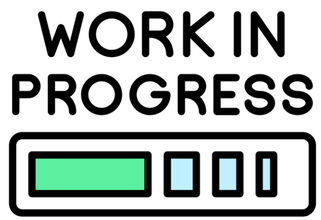
- You may use AI (Claude, Gemini, ChatGPT, etc.) within reason
- Make it help you learn, don’t outsource thinking and understanding to it.
  - “Fix this map() code”
  - “Teach me how to use map()”
- You take ultimate responsibility for your code and learning
- We reserve the right to ask you to explain your assignment’s code in an oral exam

IMPROVING OUR STATISTICAL INFERENCES

ianhussey.quarto.pub/improving-your-statistical-inferences-through-monte-carlo-simulations

TEXTBOOK

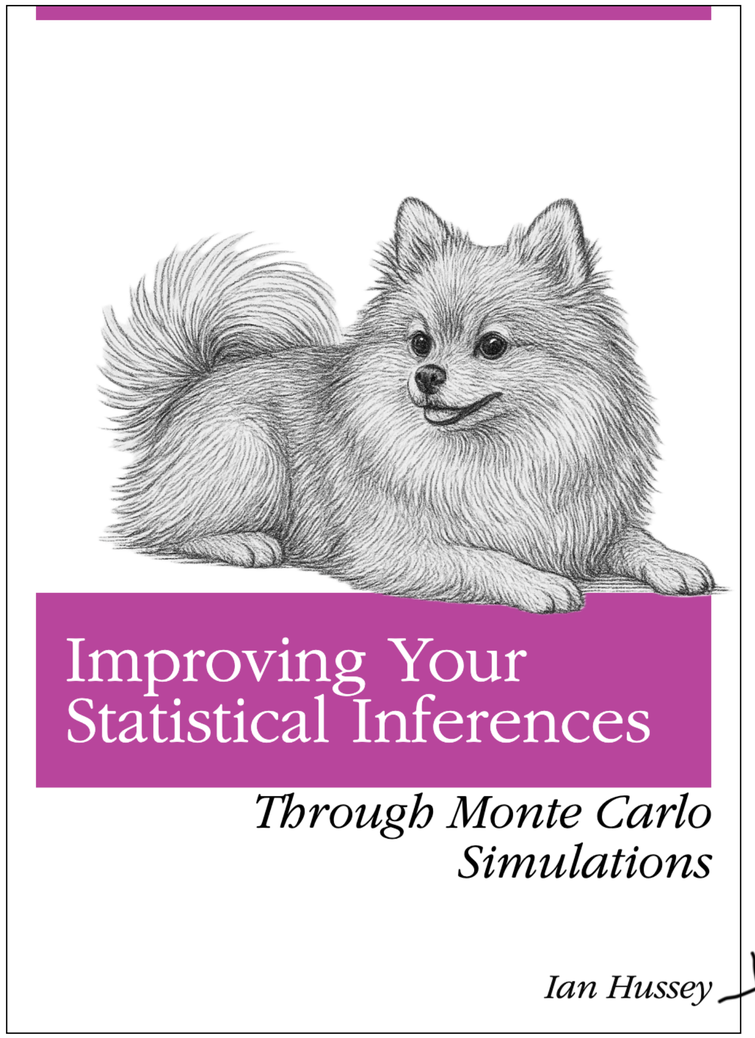




+ JAMIE

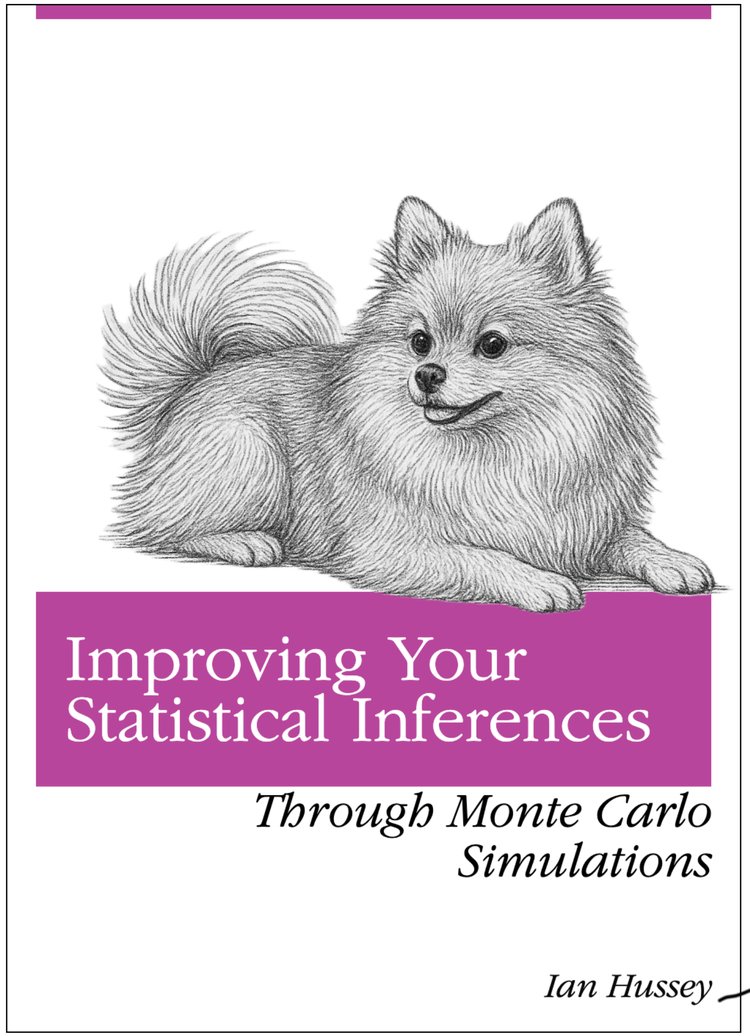
EVERYTHING IS COMPUTER

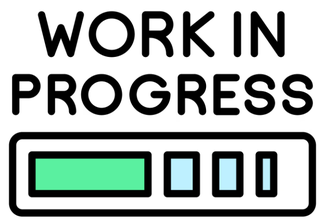
.pub/improving-your-statistical-inferences-thro



ianhussey.quarto.pub/improving-your-statistical-inferences-through-monte-carlo-simulations

TEXTBOOK



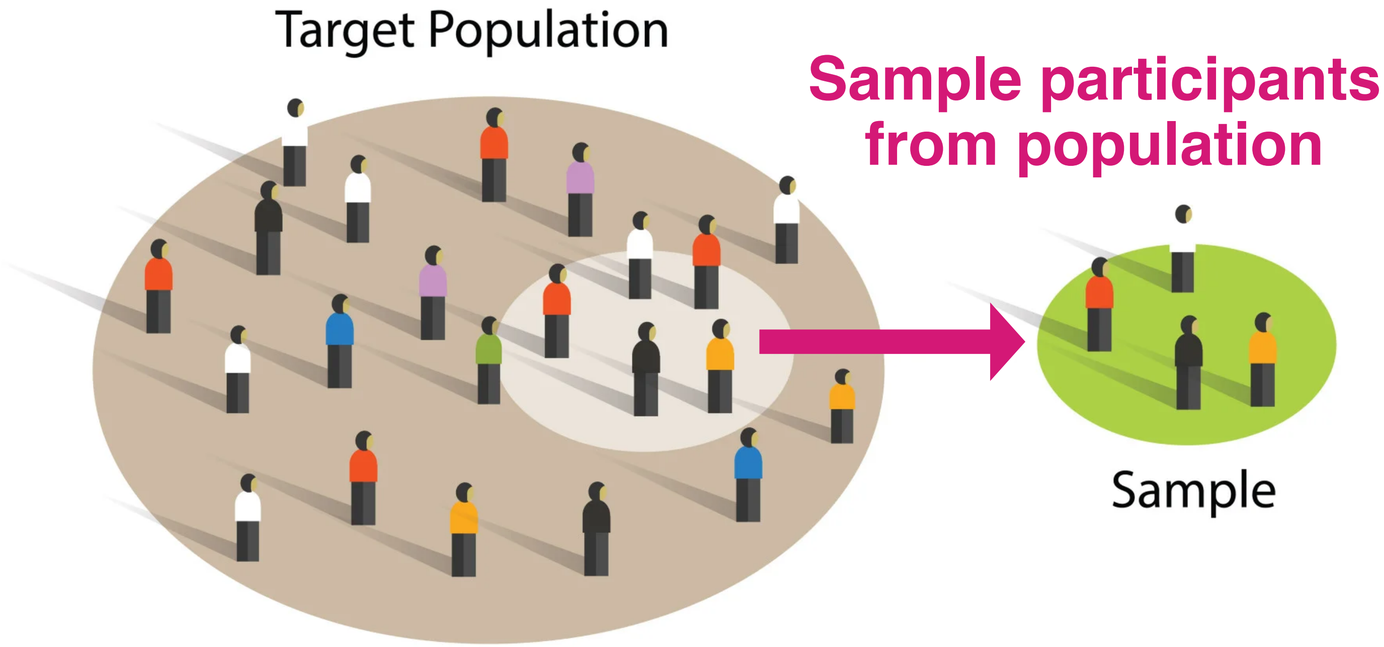


+ JAMIE

EVERYTHING IS COMPUTER

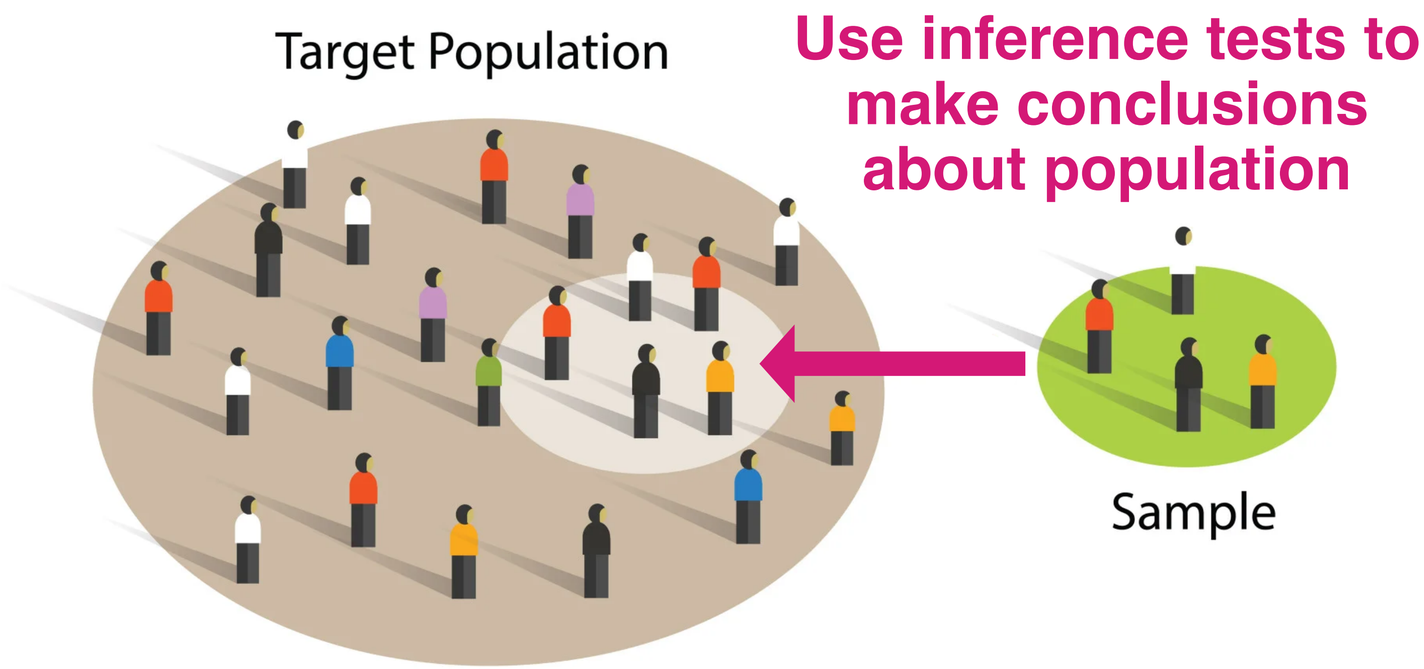
STATISTICAL INFERENCE

- Sample a small number of participants from a population
- Take measurements, maybe intervene
- Analyse data
  - Use inferential statistics to make inferences about <u>population</u> from <u>sample</u>



STATISTICAL INFERENCE

- Sample a small number of participants from a population
- Take measurements, maybe intervene
- Analyse data
  - Use inferential statistics to make inferences about <u>population</u> from <u>sample</u>



STATISTICAL INFERENCE

- What is a t-test used for?

STATISTICAL INFERENCE

- What is a t-test used for?
  - "Testing differences between groups”

STATISTICAL INFERENCE

- What is a t-test used for?
  - ~~"Testing differences between groups"~~ - imprecise, confusing
  - Making inferences about differences in population means

STATISTICAL INFERENCE

- What is a t-test used for?
  - ~~"Testing differences between groups"~~ - imprecise, confusing
  - Making inferences about differences in population means

- Do these means differ?
  - $M_{intervention} = 19.1$, $M_{control} = 21.8$, $p = .46$

STATISTICAL INFERENCE

- What is a t-test used for?
  - ~~"Testing differences between groups"~~ - imprecise, confusing
  - Making inferences about differences in population means

- Do these means differ?
  - $M_{intervention} = 19.1$, $M_{control} = 21.8$, $p = .46$
  - Yes: $17.1 \neq 21.8$

STATISTICAL INFERENCE

- What is a t-test used for?
  - ~~"Testing differences between groups"~~ - imprecise, confusing
  - Making inferences about differences in population means

- Do these means differ?
  - $M_{intervention} = 19.1$, $M_{control} = 21.8$, $p = .46$
  - Yes: $17.1 \neq 21.8$

  - Should we infer that the population means differ?
    - No, difference was not 'statistically significant' (inference rule)

STATISTICAL INFERENCE - REAL WORLD STUDIES

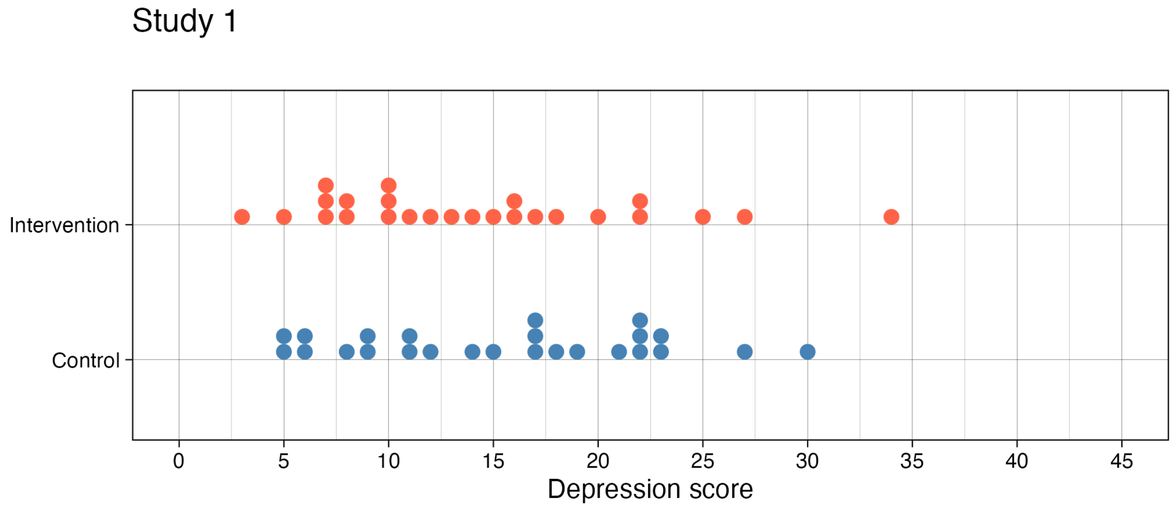
- Population values are unknown and unknowable

Population

Mean difference (Δμ) = ??

sampling

inference



STATISTICAL INFERENCE - REAL WORLD STUDIES

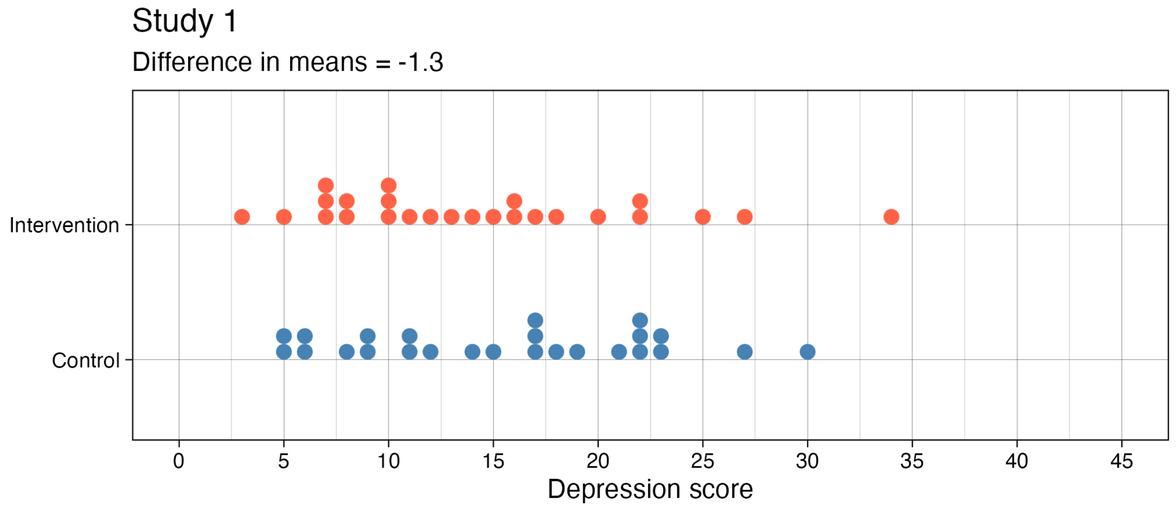
- Population values are unknown and unknowable
- But can be estimated

Population

Mean difference (Δμ) = ??

sampling

inference



STATISTICAL INFERENCE - REAL WORLD STUDIES

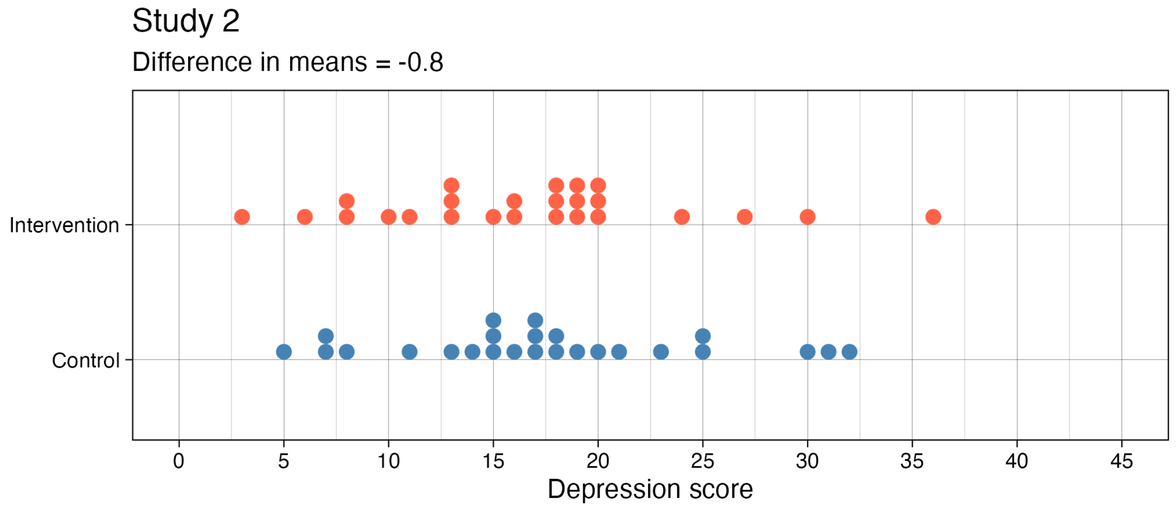
- Population values are unknown and unknowable
- But can be estimated

Population

Mean difference (Δμ) = ??

sampling

inference



STATISTICAL INFERENCE - REAL WORLD STUDIES

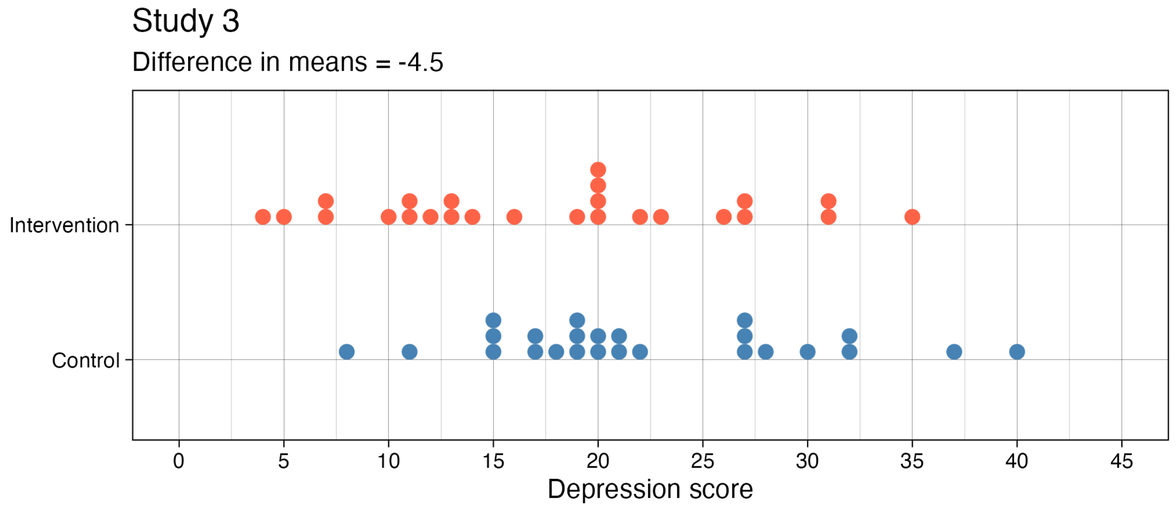
- Population values are unknown and unknowable
- But can be estimated

Population

Mean difference (Δμ) = ??

sampling

inference



STATISTICAL INFERENCE - REAL WORLD STUDIES

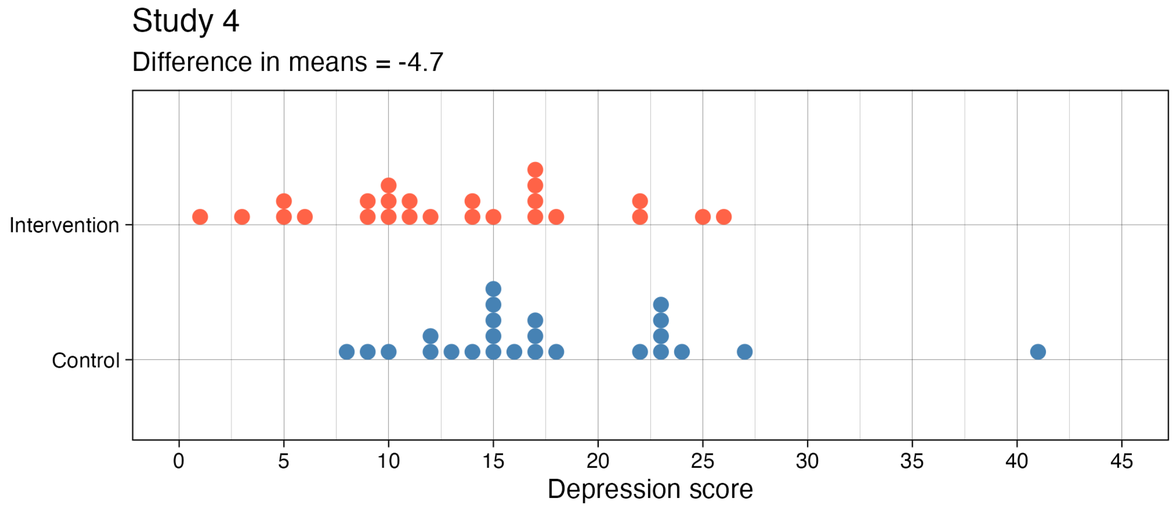
- Population values are unknown and unknowable
- But can be estimated / decided with knowable error rates

Population

Mean difference (Δμ) = ??

sampling

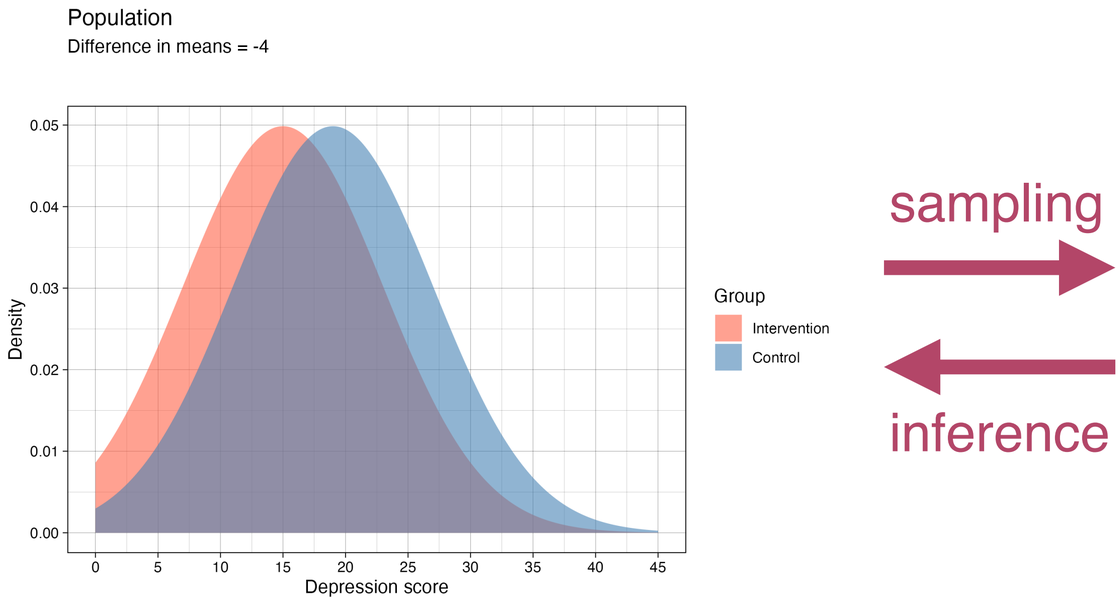
inference



SIMULATION STUDIES

SIMULATION STUDIES

- Population values are known
- The ability of different estimation and inference methods can be tested



sampling

inference

Studies on average:

Study 1
Difference in means = -1.3

Study 2
Difference in means = -0.8

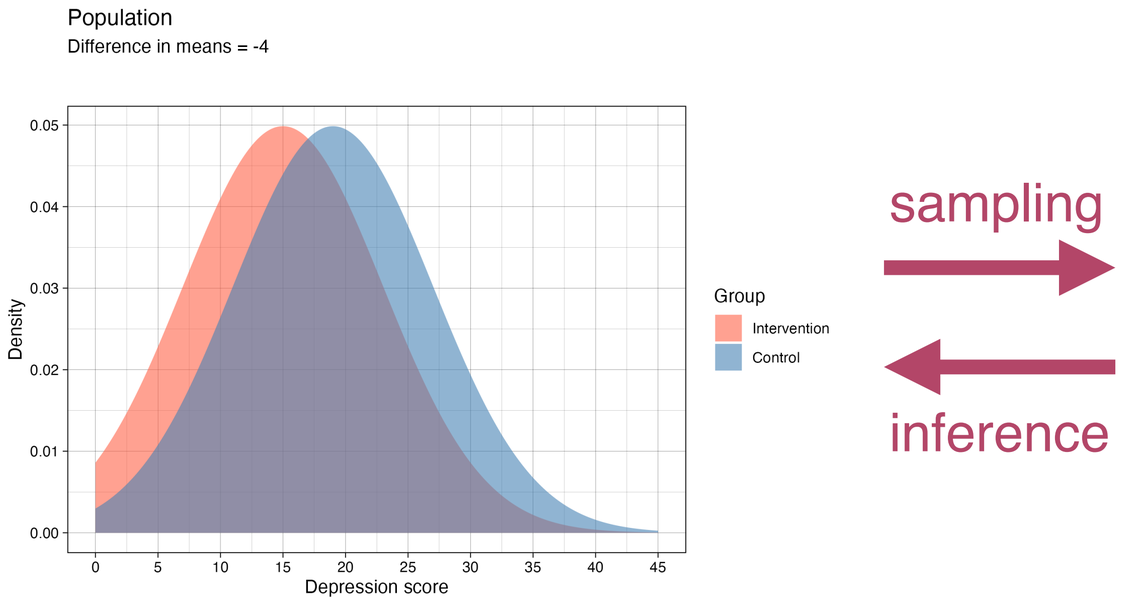
Study 3
Difference in means = -4.5

Study 4
Difference in means = -4.7

...

SIMULATION STUDIES

- Population values are known
- The ability of different estimation and inference methods can be tested
  - Setting alpha = .05 should mean 5% false positive rate. Does it?



sampling

inference

Studies on average:

Study 1
Difference in means = -1.3

Study 2
Difference in means = -0.8

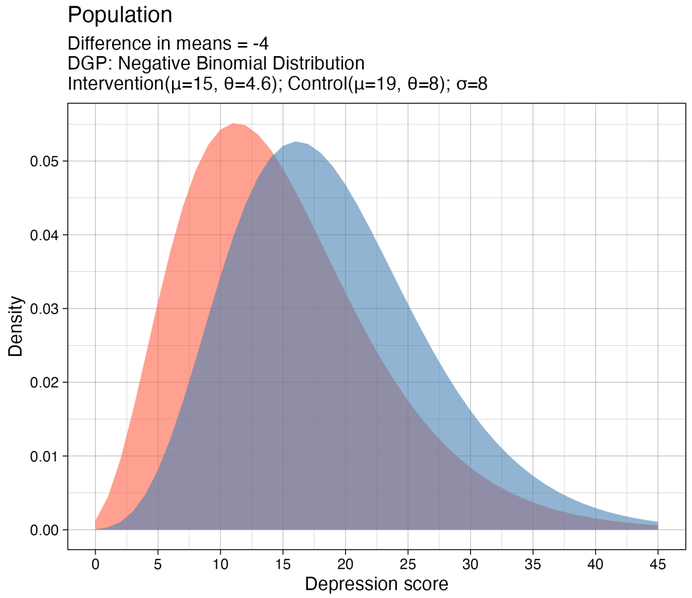
Study 3
Difference in means = -4.5

Study 4
Difference in means = -4.7

…

SIMULATION STUDIES

- Population values are known
- The ability of different estimation and inference methods can be tested
  - Under violations of assumptions, eg normality



sampling

inference

Studies on average:

Study 1
Difference in means = -1.3

Study 2
Difference in means = -0.8

Study 3
Difference in means = -4.5

Study 4
Difference in means = -4.7

...

SIMULATION STUDIES

Allow us to:

- Choose and vary the Ground Truth
- Observe how statistical methods perform under these variations
- Compare different methods/choices under the same ground truth
- Model researcher behavior as a series of rule-based decisions