University of Glasgow | Institute of Neuroscience & Psychology

CENTRE FOR COGNITIVE NEUROIMAGING

Benefits of learning and teaching bootstrap methods

Guillaume Rousselet
@robustgar
https://garstats.wordpress.com

"Trying to use a survey of 3,000 people to estimate tiny differences in sex ratios: this makes about as much sense as using a bathroom scale to weigh a feather, when that feather is resting loosely in the pouch of a kangaroo that is vigorously jumping up and down."

Andrew Gelman, 2018, The Failure of Null Hypothesis Significance Testing When Studying Incremental Changes, and What to Do About It. *Personality and Social Psychology Bulletin*

Teaching benefits

introduce or consolidate:

- **key frequentist concepts (sampling distributions, SE, confidence intervals…)**
- **inferential statistics**
- **experimental design (how do I bootstrap my data?)**
- **robust statistics**
- **simulations**
- **R skills (including graphical representations)**
- **dealing with distributions of plausible population values -> Bayes**

Special Section: Using Simulation to Convey Statistical Concepts
Tutorial

Advances in Methods and Practices in Psychological Science
January-March 2021, Vol. 4, No. 1, pp. 1–10

DOI: 10.1177/2515245920911881
www.psychologicalscience.org/AMPPS
SAGE

The Percentile Bootstrap: A Primer With Step-by-Step Instructions in R

Guillaume A. Rousselet[1], Cyril R. Pernet[2], and Rand R. Wilcox[3]
[1]Institute of Neuroscience & Psychology, University of Glasgow; [2]Centre for Clinical Brain Sciences, Neuroimaging Sciences, University of Edinburgh; and [3]Department of Psychology, University of Southern California

https://journals.sagepub.com/doi/full/10.1177/2515245920911881

PsyArXiv Preprints   Submit a Preprint   Search   Donate   Sign Up   Sign In

A practical introduction to the bootstrap: a versatile method to make inferences by using data-driven simulations

AUTHORS
Guillaume Rousselet, Dr Cyril Pernet, Rand R. Wilcox

https://psyarxiv.com/h8ft7/

The American Statistician ›
Volume 69, 2015 - Issue 4: Special Issue on Statistics and the Undergraduate Curriculum

Articles

What Teachers Should Know About the Bootstrap: Resampling in the Undergraduate Statistics Curriculum

Tim C. Hesterberg

Pages 371-386 | Received 01 Dec 2014, Accepted author version posted online: 11 Sep 2015, Published online:29 Dec 2015

Download citation https://doi.org/10.1080/00031305.2015.1089789 Check for updates

https://www.tandfonline.com/doi/full/10.1080/00031305.2015.1089789

CURRENT PROTOCOLS in Neuroscience
A Wiley Brand

UNIT

A Guide to Robust Statistical Methods in Neuroscience

Rand R. Wilcox, Guillaume A. Rousselet

First published: 22 January 2018 | https://doi.org/10.1002/cpns.41 | Citations: 35

https://currentprotocols.onlinelibrary.wiley.com/doi/abs/10.1002/cpns.41

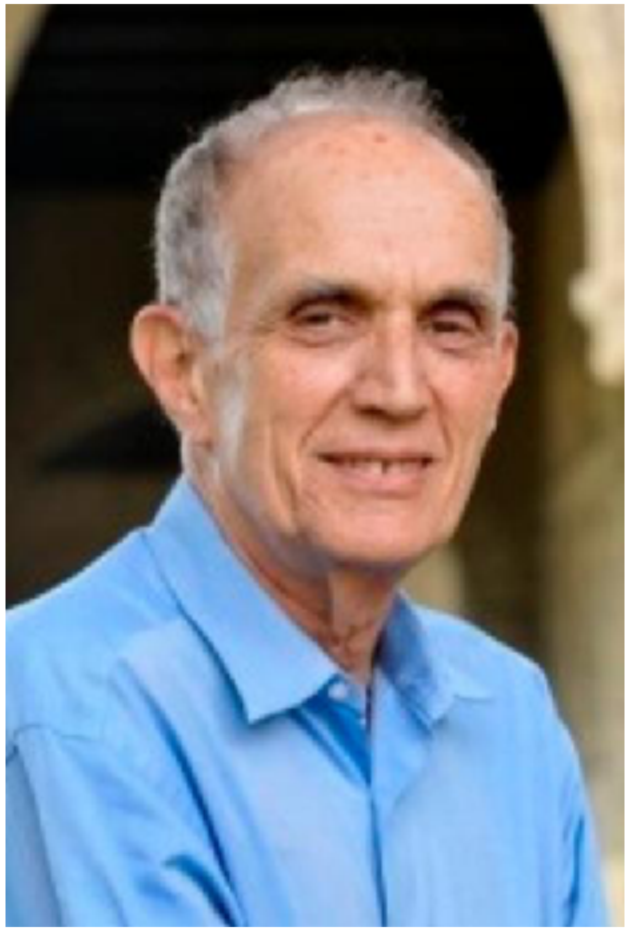
The Annals of Statistics
1979, Vol. 7, No. 1, 1–26

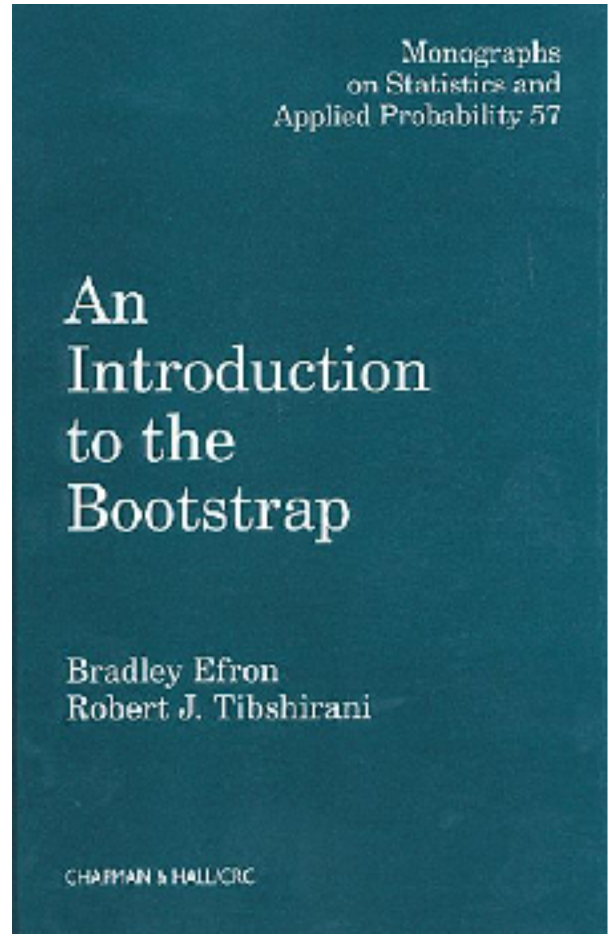
THE 1977 RIETZ LECTURE

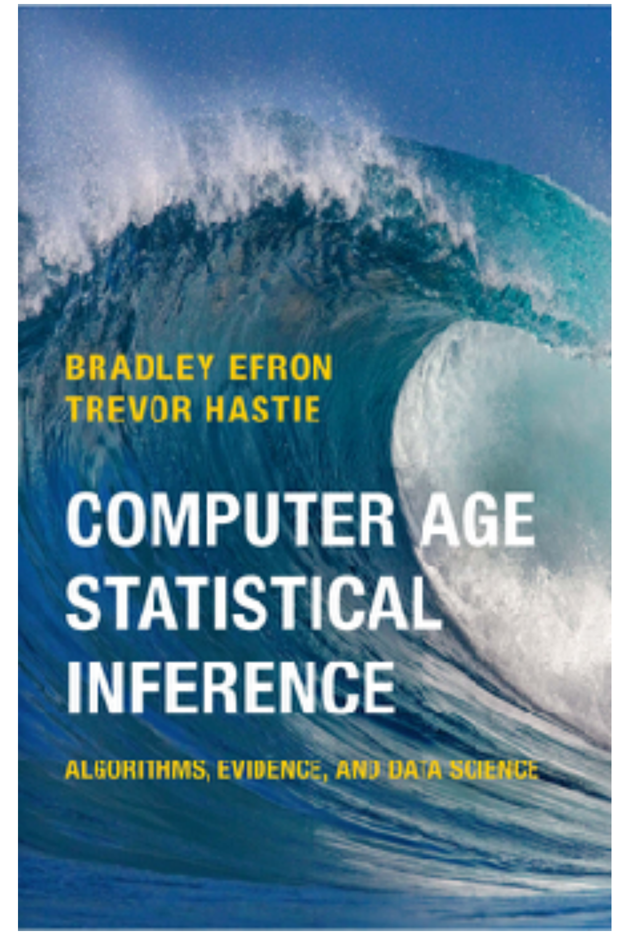
BOOTSTRAP METHODS: ANOTHER LOOK AT THE JACKKNIFE

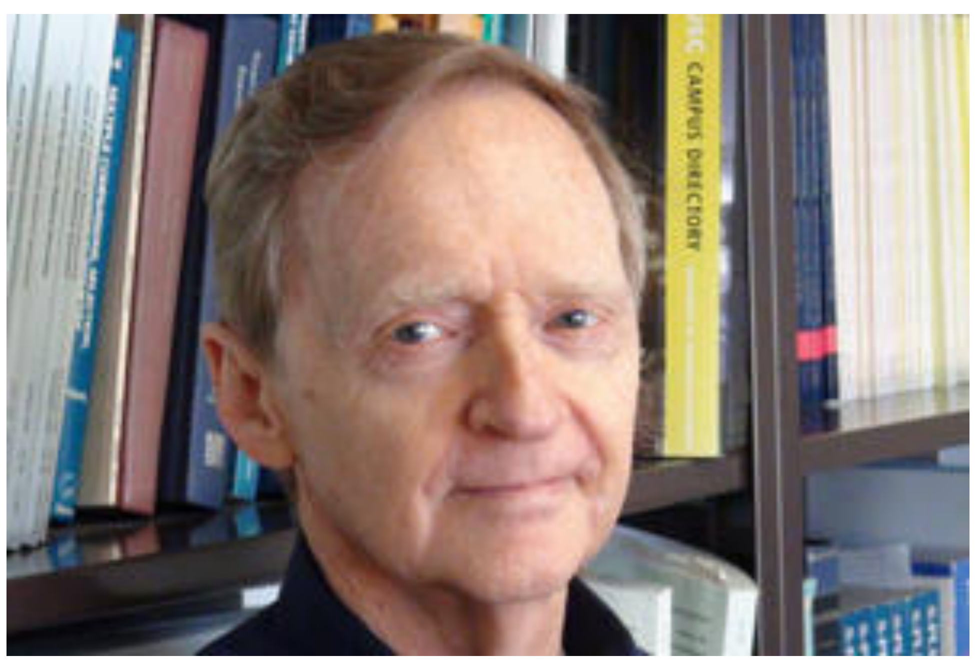
BY B. EFRON
Stanford University

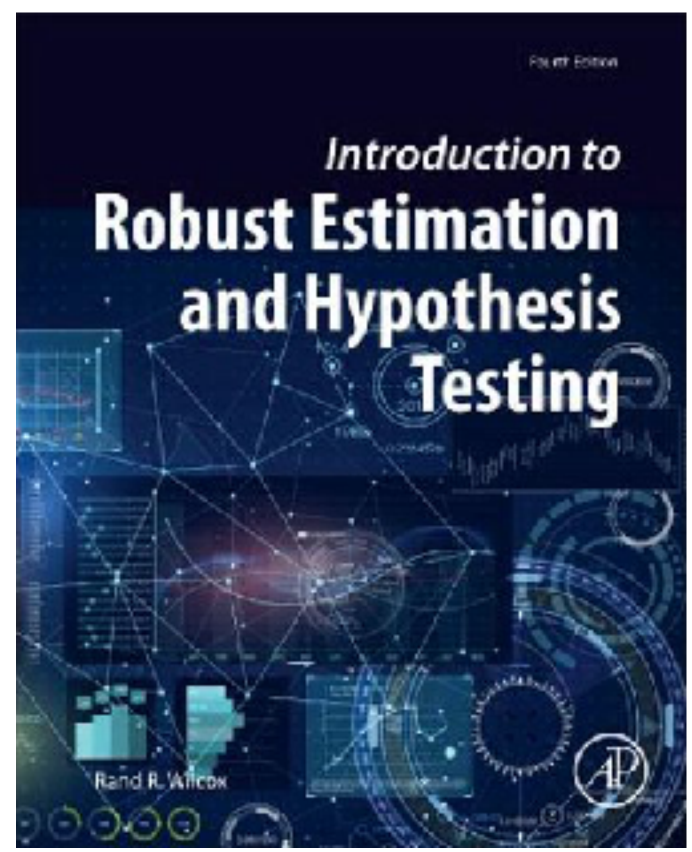
We discuss the following problem: given a random sample $\mathbf{X} = (X_1, X_2, \ldots, X_n)$ from an unknown probability distribution F, estimate the sampling distribution of some prespecified random variable $R(\mathbf{X}, F)$, on the basis of the observed data $\mathbf{x}$. (Standard jackknife theory gives an approximate mean and variance in the case $R(\mathbf{X}, F) = \theta(\hat{F}) - \theta(F)$, θ some parameter of interest.) A general method, called the "bootstrap," is introduced, and shown to work satisfactorily on a variety of estimation problems. The jackknife is shown to be a linear approximation method for the bootstrap. The exposition proceeds by a series of examples: variance of the sample median, error rates in a linear discriminant analysis, ratio estimation, estimating regression parameters, etc.



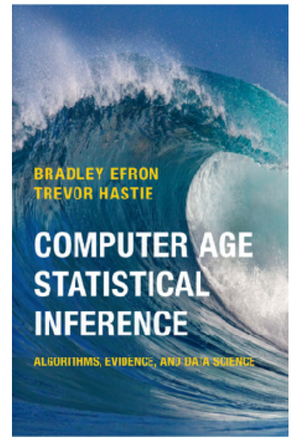

Monographs
on Statistics and
Applied Probability 57
An
Introduction
to the
Bootstrap
Bradley Efron
Robert J. Tibshirani
CHAPMAN & HALL/CRC


BRADLEY EFRON
TREVOR HASTIE
COMPUTER AGE
STATISTICAL
INFERENCE
ALGORITHMS, EVIDENCE, AND DATA SCIENCE


CAMPUS DIRECTORY


Introduction to
Robust Estimation
and Hypothesis
Testing
Rand R. Wilcox

Part II Early Computer-Age Methods



~~The~~ bootstrap?

percentile bootstrap	bootstrap-t	hierarchical bootstrap
fractional-random-weight boot	BCa bootstrap	AA bootstrap
observed imposed bootstrap	smooth bootstrap	wild bootstrap

Bootstrap demo:
[1] sampling without replacement
[2] sampling with replacement
[3] bootstrap sampling

R implementation

```
n <- 6
samp <- 1:n
sample(samp, size=n, replace=TRUE)
```

3 bootstrap samples

set.seed(21) # reproducible example

nboot <- 3

matrix(sample(samp, size = n*nboot, replace = TRUE), nrow = nboot, byrow = TRUE)

```
[1] 5 2 5 2 6 6 1 2 6 6 5 6 1 4 2 1 4 2
```

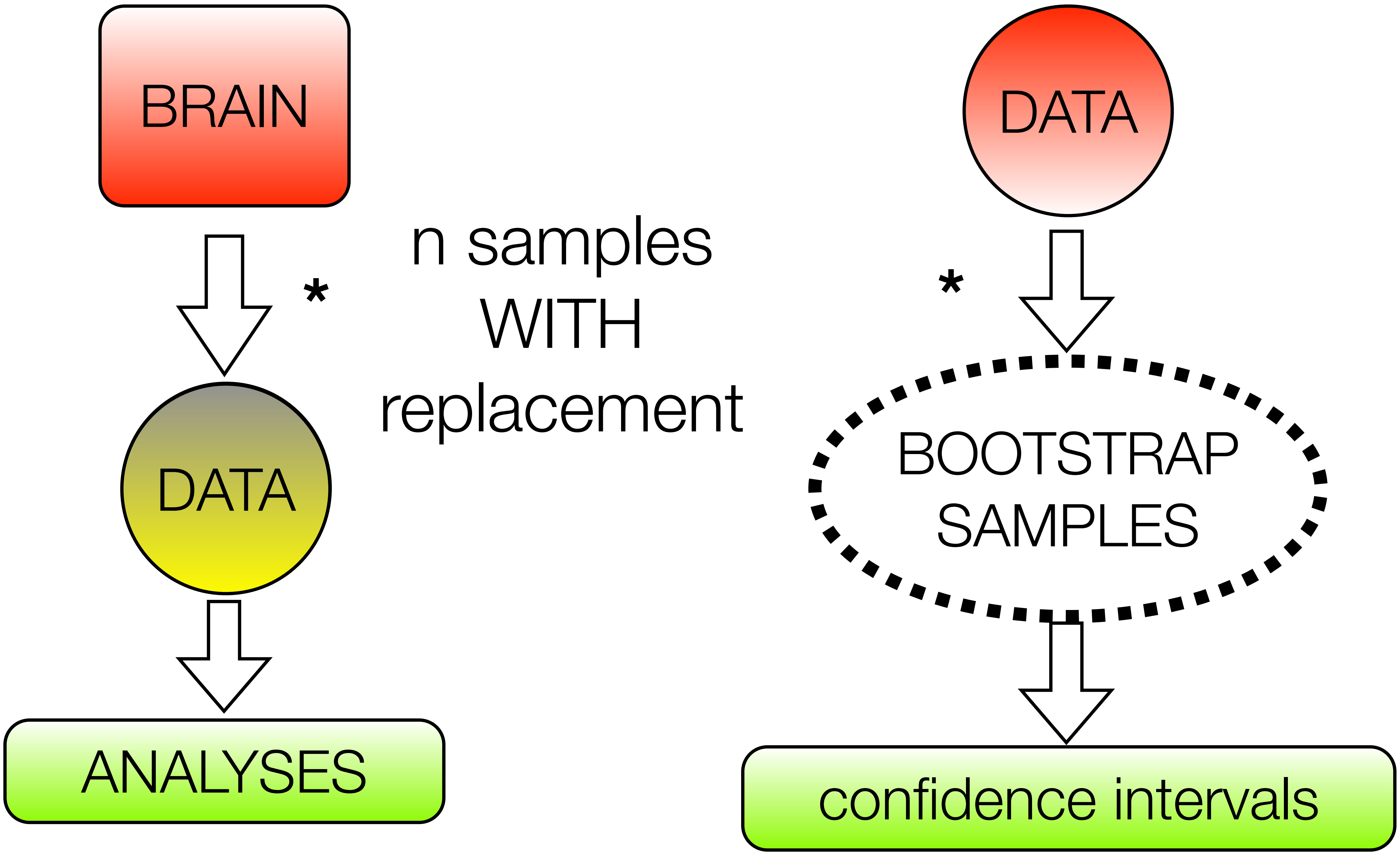
```
     [,1] [,2] [,3] [,4] [,5] [,6]
[1,]    5    2    5    2    6    6
[2,]    1    2    6    6    5    6
[3,]    1    4    2    1    4    2
```

Bootstrap: central idea

- "The bootstrap is a computer-based method for assigning measures of accuracy to statistical estimates." Efron & Tibshirani, 1993

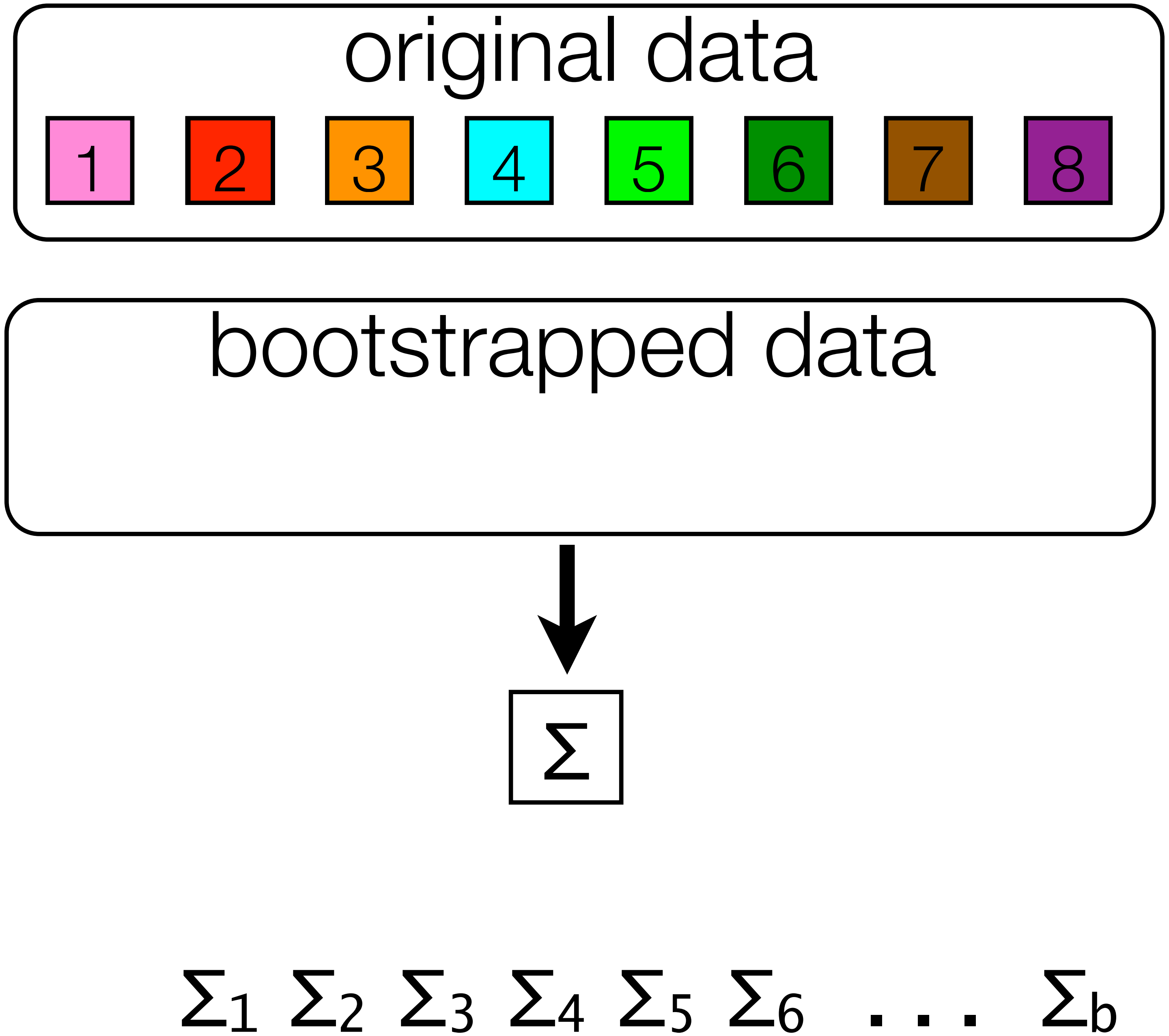
- "The central idea is that it may sometimes be better to draw conclusions about the characteristics of a population strictly from the sample at hand, rather than by making perhaps unrealistic assumptions about the population." Mooney & Duval, 1993

bootstrap philosophy



n samples WITH replacement

Percentile bootstrap: general recipe



(1) sample WITH replacement n observations

(2) compute estimate
e.g. sum, trimmed mean

(3) repeat (1) & (2) b times

(4) get 1-alpha confidence interval

R implementation

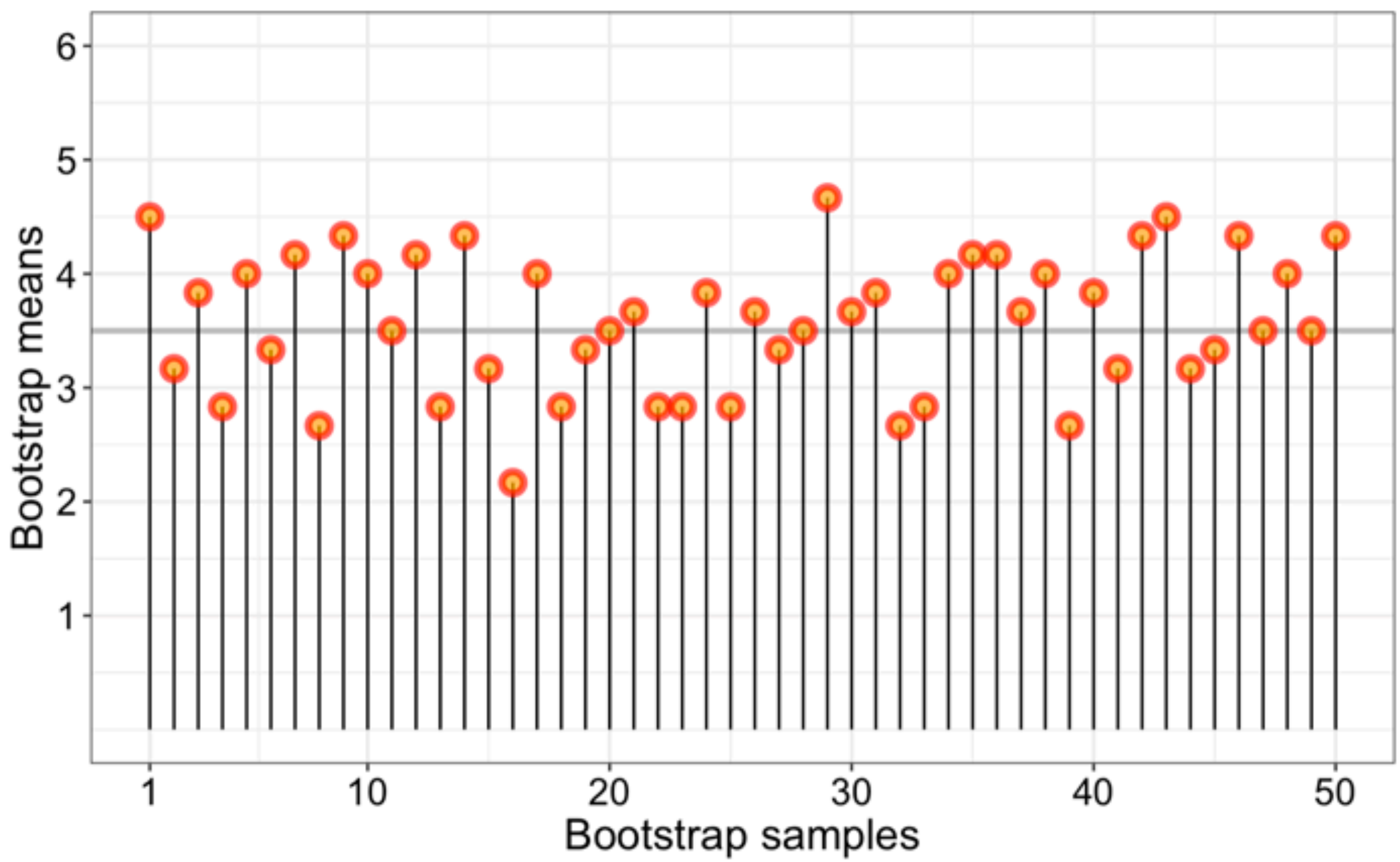
Loop

```
set.seed(21) # reproducible results
nboot <- 1000 # number of bootstrap samples

# declare vector of results
boot.m <- vector(mode = "numeric", length = nboot)

for(B in 1:nboot){
  boot.samp <- sample(samp, size = n, replace = TRUE) # sample with replacement
  boot.m[B] <- mean(boot.samp) # save bootstrap means
}
```

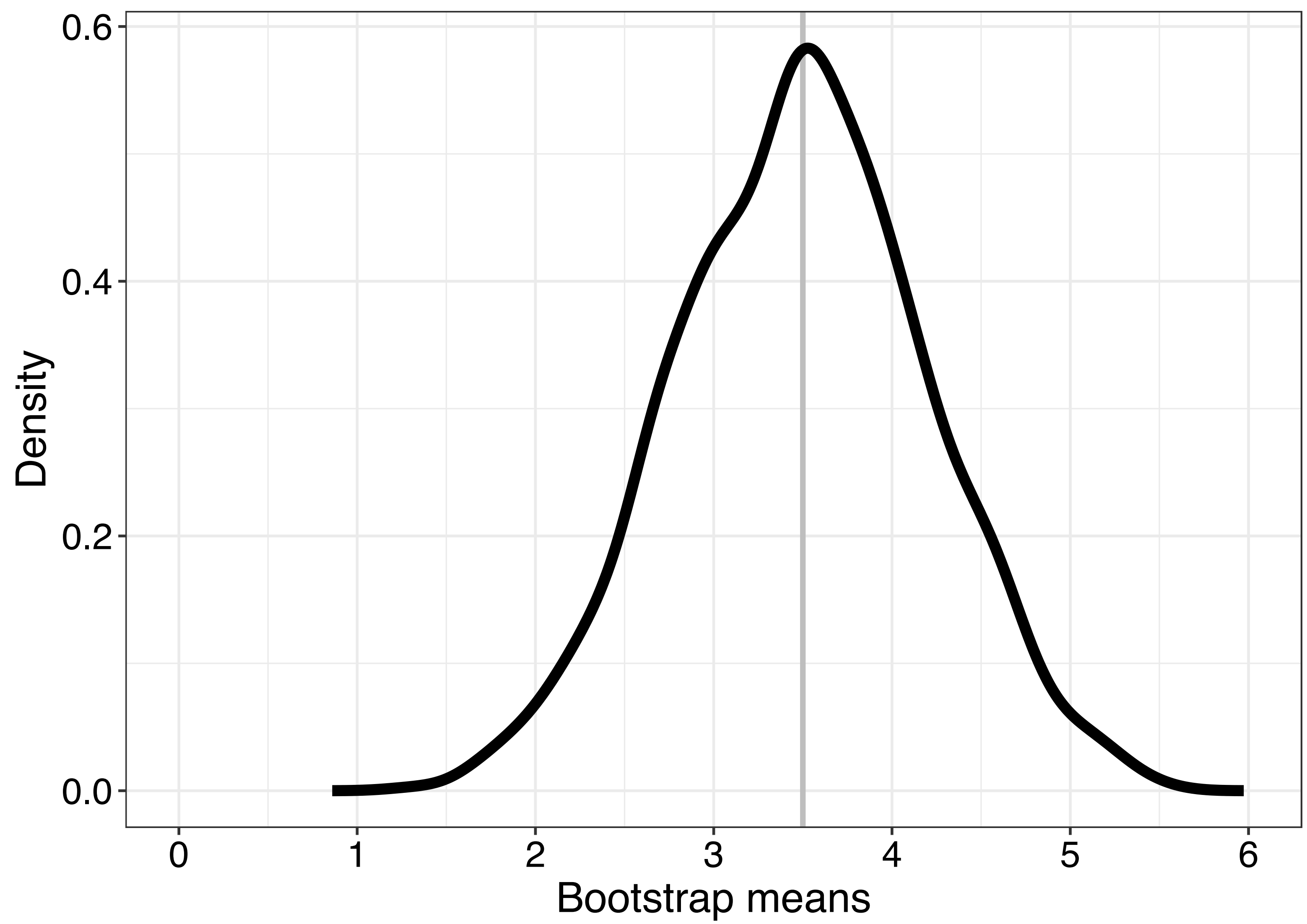
Matrix method

```
set.seed(21)
boot.m <- apply(matrix(sample(samp, size = n*nboot, replace = TRUE), nrow = nboot), 1, mean)
```

First 50 bootstrap means



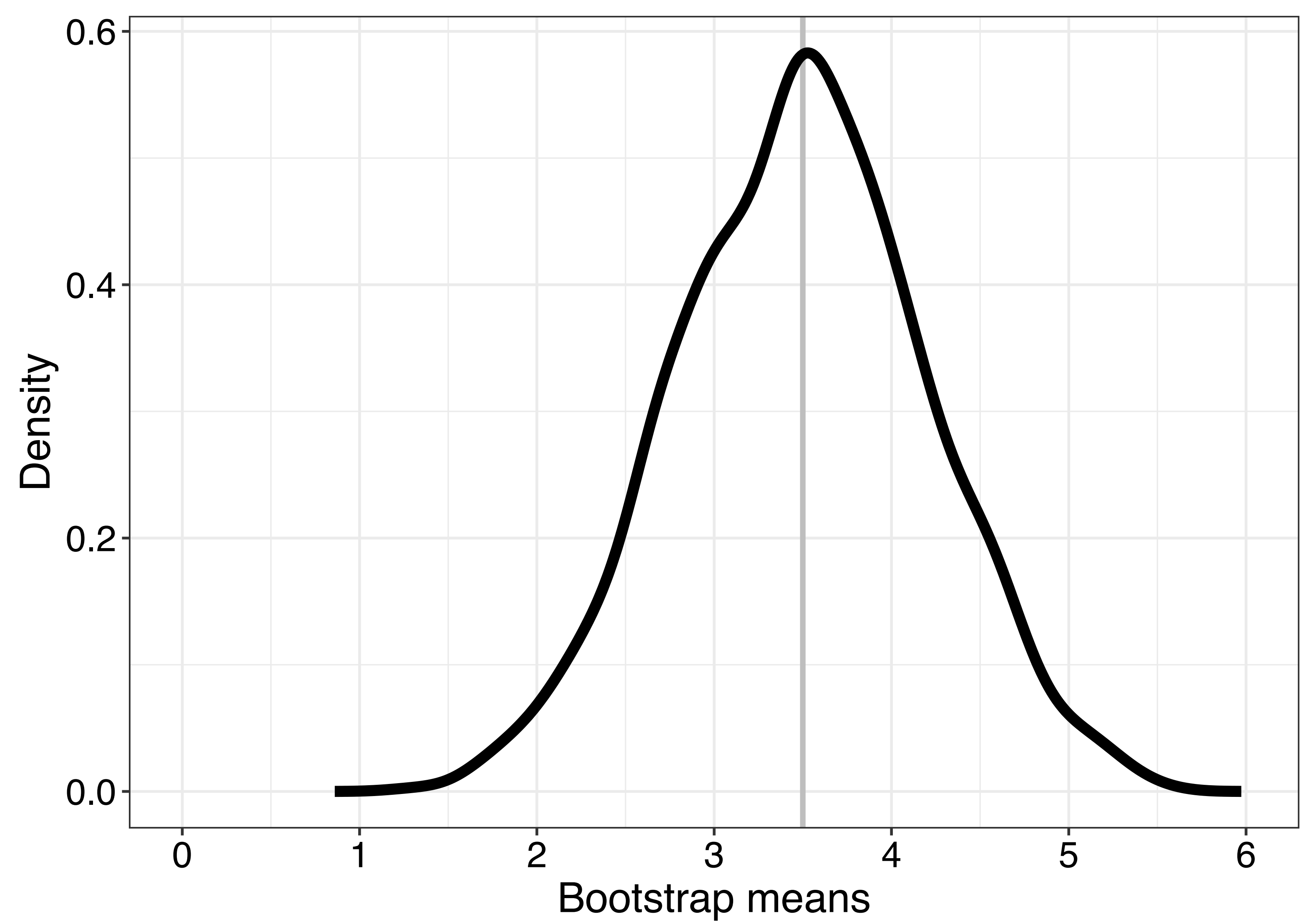
Density plot of 1,000 bootstrap estimates



Density plot of 1,000 bootstrap estimates

Bootstrap sampling distribution
Can be used to compute:

- SE estimate
- bias estimate
- confidence interval
- p value



More about bias and bootstrap bias estimation…

Meta-Psychology

Current Archives About

Home / Archives / Vol. 4 (2020) / Original articles

Reaction Times and other Skewed Distributions

Problems with the Mean and the Median

Guillaume A Rousselet
Institute of Neuroscience and Psychology, College of Medical, Veterinary and Life Sciences, University of Glasgow
https://orcid.org/0000-0003-0006-8729

Rand R Wilcox
Dept. of Psychology, University of Southern California
https://orcid.org/0000-0002-2524-2976

DOI: https://doi.org/10.15626/MP.2019.1630

Keywords: mean, median, trimmed mean, quantile, sampling, bias, bootstrap, estimation, skewness

PDF

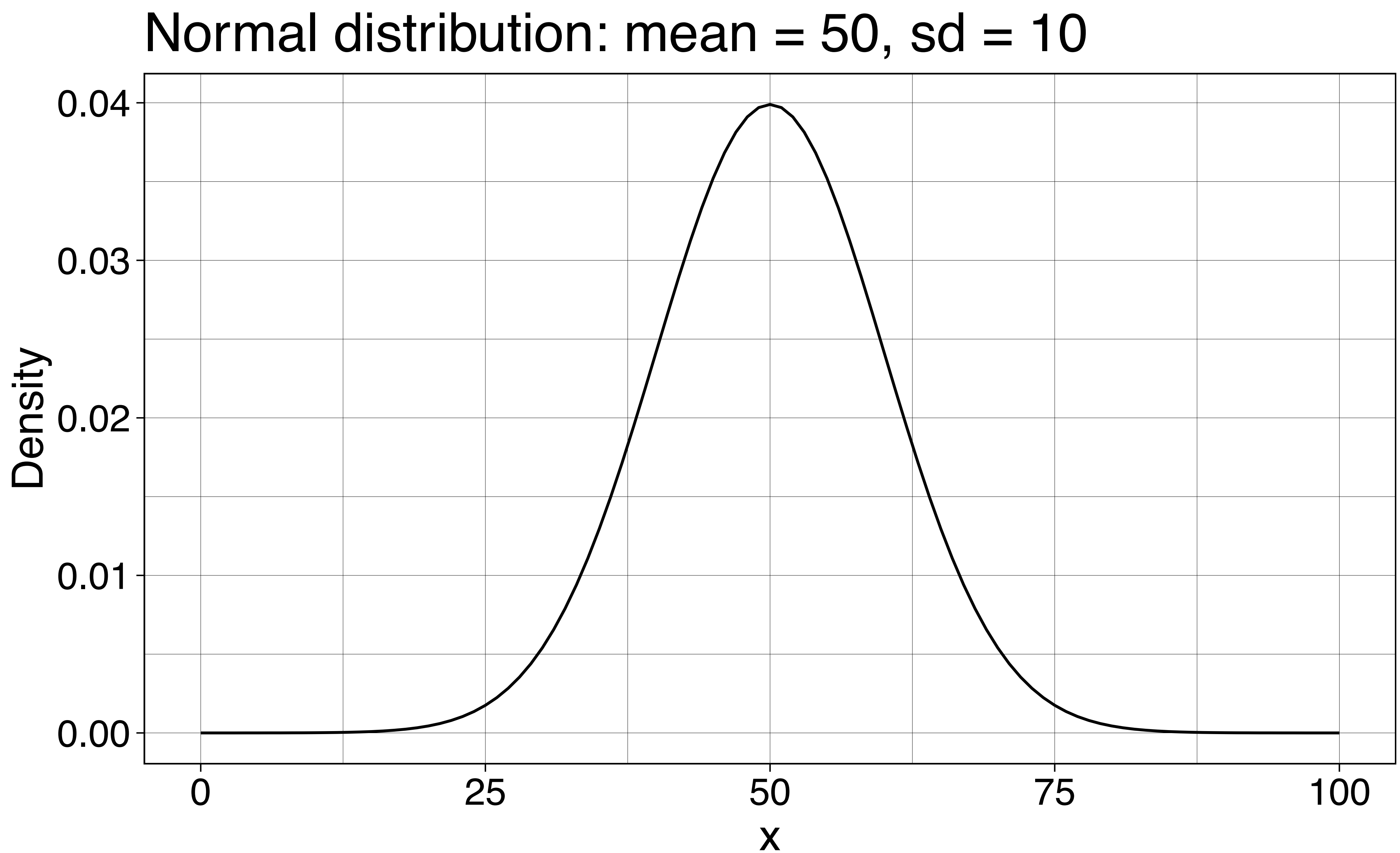
Published
2020-05-31

Issue
Vol. 4 (2020)

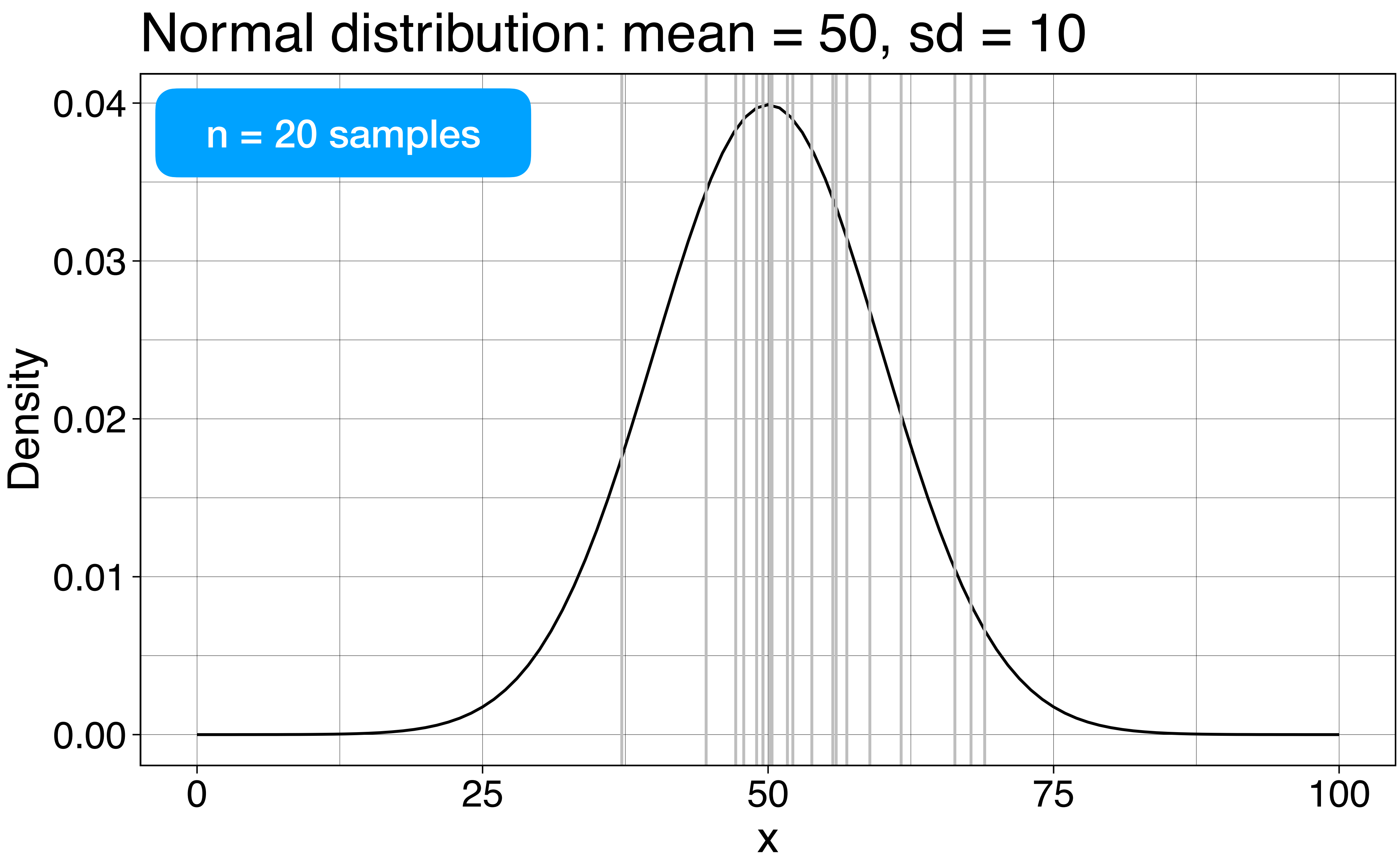
Section
Original articles

License

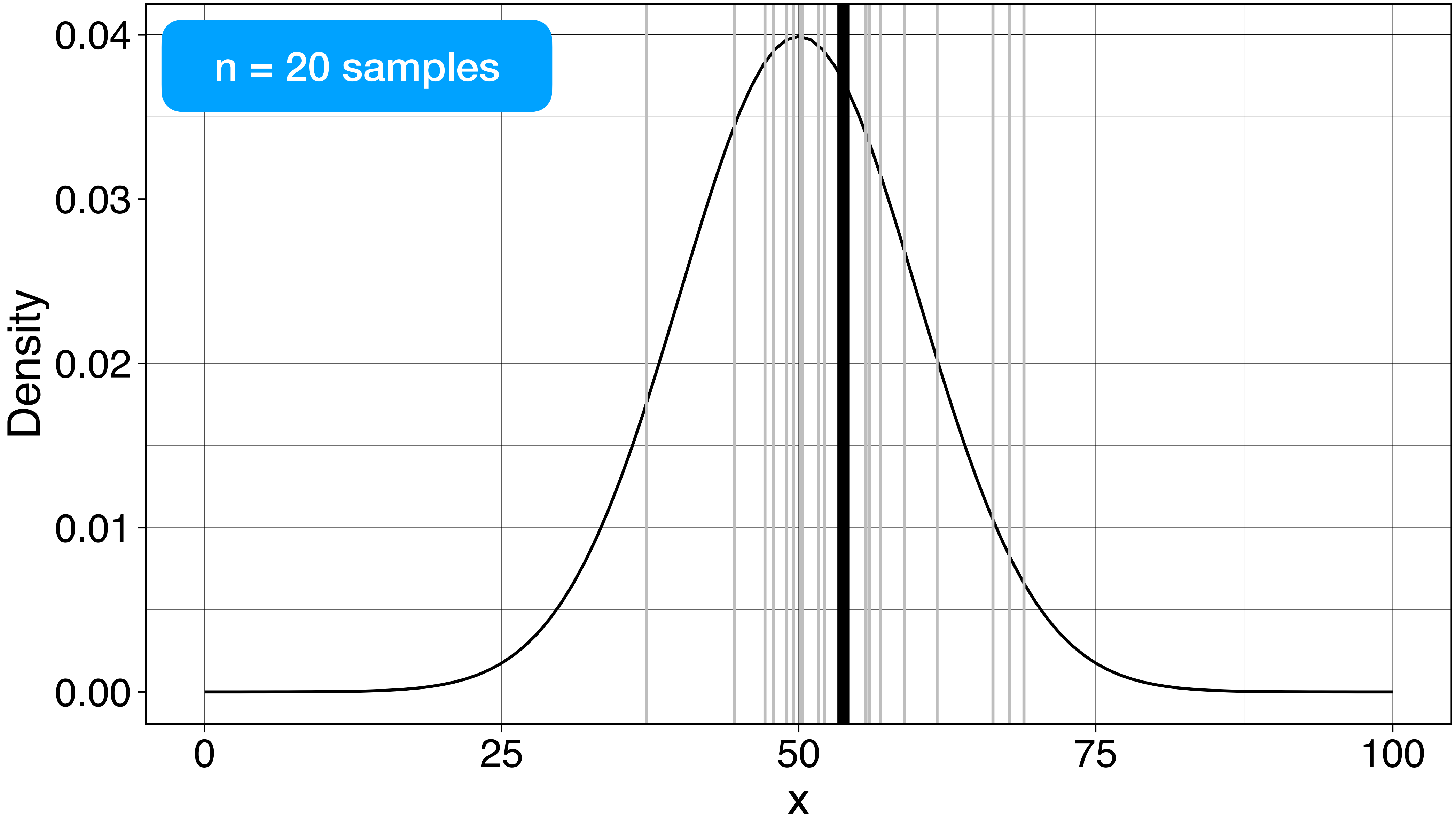
https://open.lnu.se/index.php/metapsychology/article/view/1630



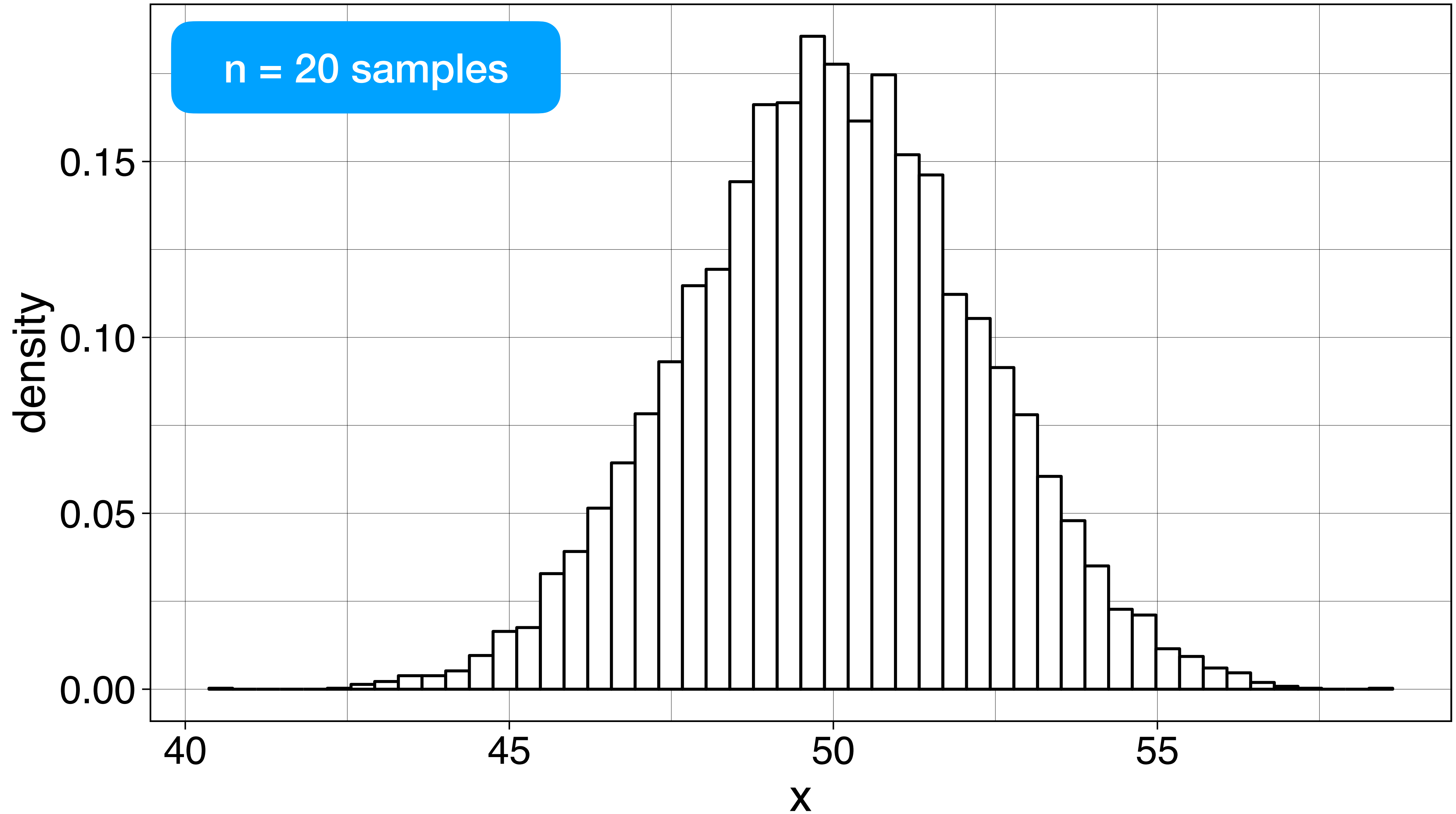
Normal distribution: mean = 50, sd = 10
Density
0.04
0.03
0.02
0.01
0.00
0
25
50
75
100
x


Normal distribution: mean = 50, sd = 10
n = 20 samples
0.04
0.03
0.02
0.01
0.00
Density
0
25
50
75
100
x

Sample mean = 53.8 , sd = 8



10000 sample means: mean = 50 , sd = 2.26



Standard error (of the mean) = sd(sampling distribution) =

sd(population) / $\sqrt{n}$ = 10 / $\sqrt{20}$ = 2.24

sd(sample) / $\sqrt{n}$

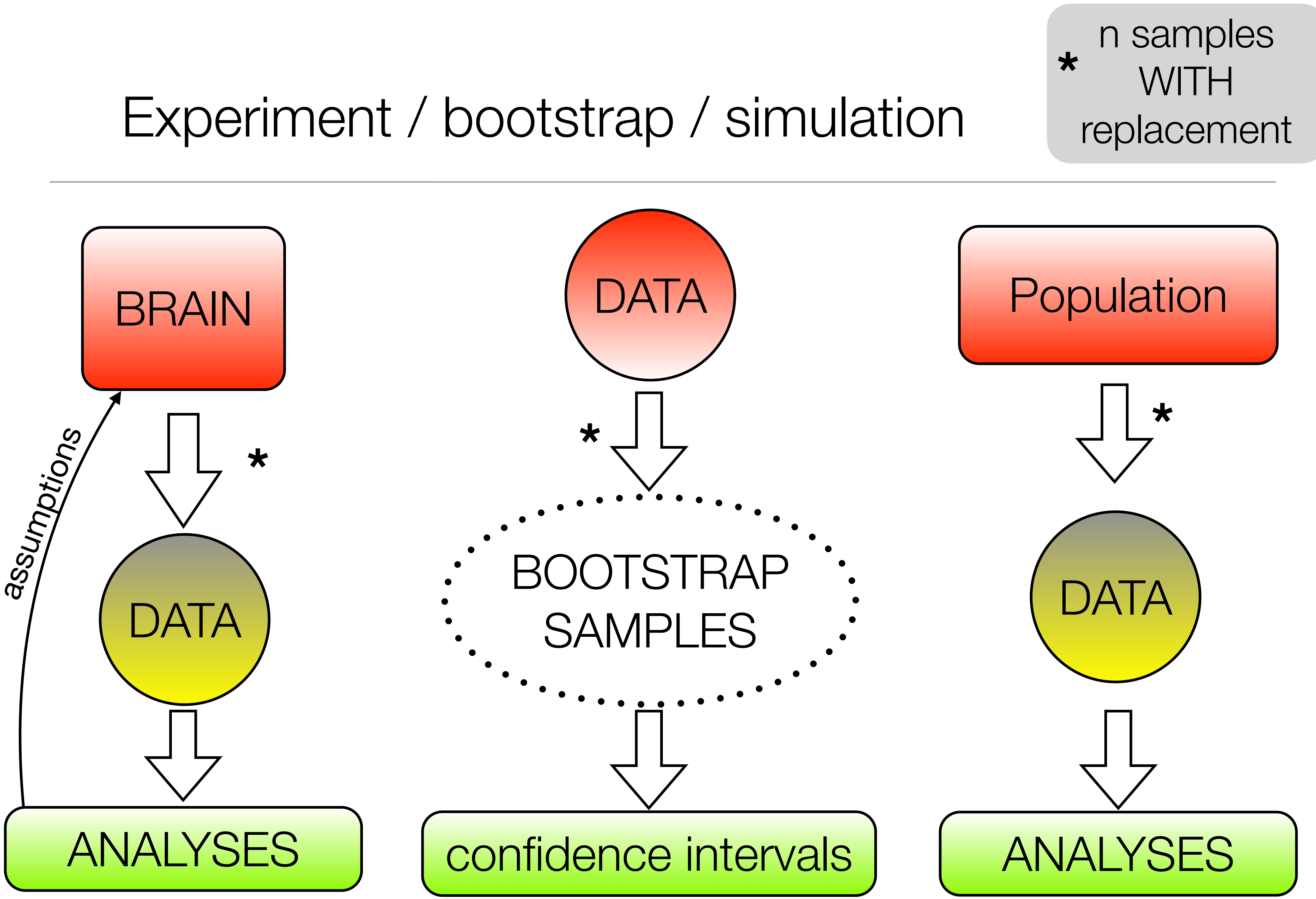
Bootstrap estimation of the sampling distribution:

```r
nboot <- 1000 # number of bootstrap samples
# samp # random sample from population
n <- length(samp) # sample size

# declare vector of results
boot.m <- vector(mode = "numeric", length = nboot)

for(B in 1:nboot){ # bootstrap loop
  # sample with replacement
  boot.samp <- sample(samp, size = n, replace = TRUE)
  # save bootstrap means
  boot.m[B] <- mean(boot.samp)
}
```

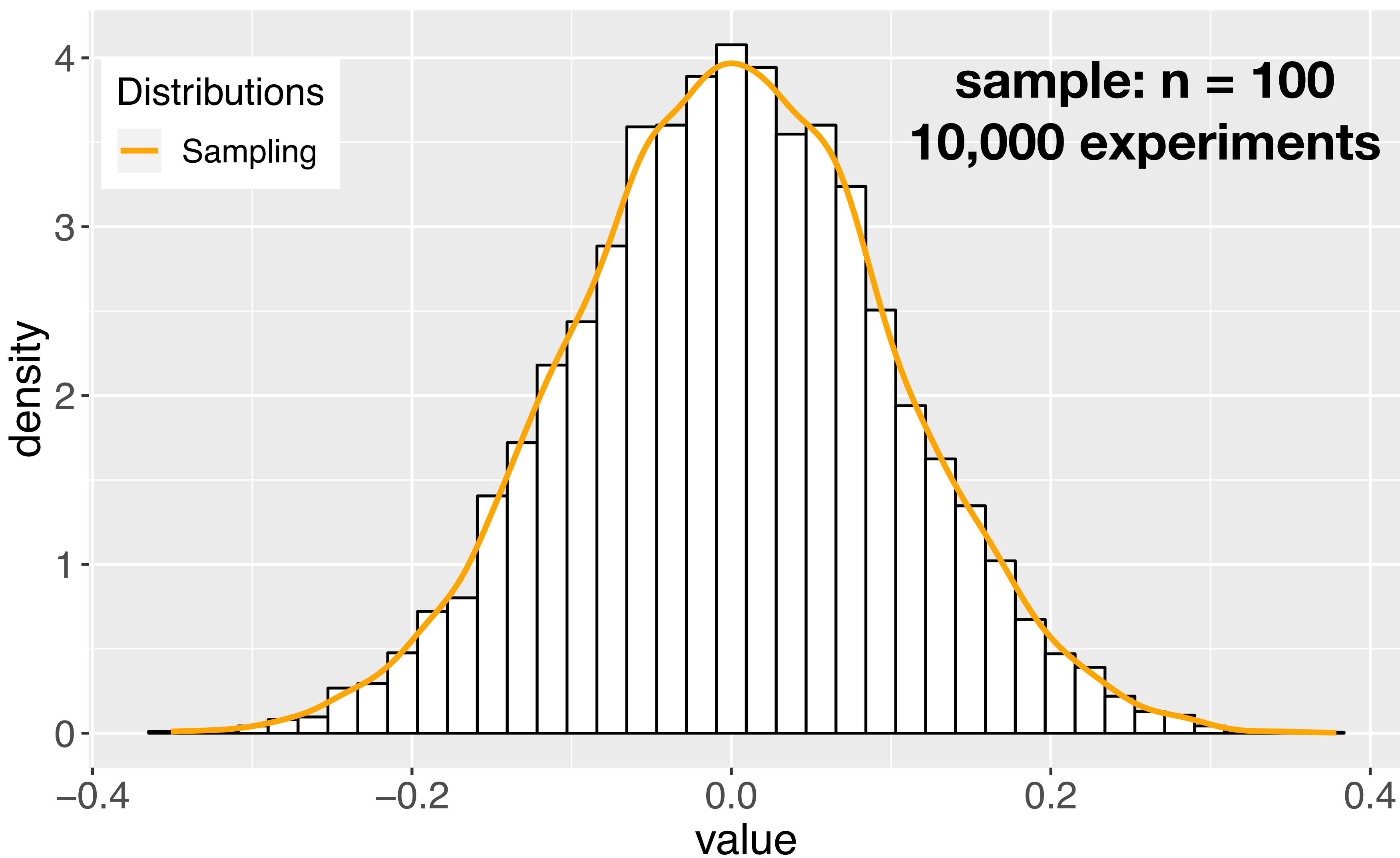
Simulation of sampling distribution:

```r
nsim <- 1000 # number of simulation iterations
n <- 20 # sample size
pop <- rnorm(100000, mean = 0, sd = 1)

# declare vector of results
sim.m <- vector(mode = "numeric", length = nsim)

for(S in 1:nsim){ # simulation loop
  # sample with replacement from population
  sim.samp <- sample(pop, size = n, replace = TRUE)
  # save simulated sample means
  sim.m[S] <- mean(sim.samp)
}
```

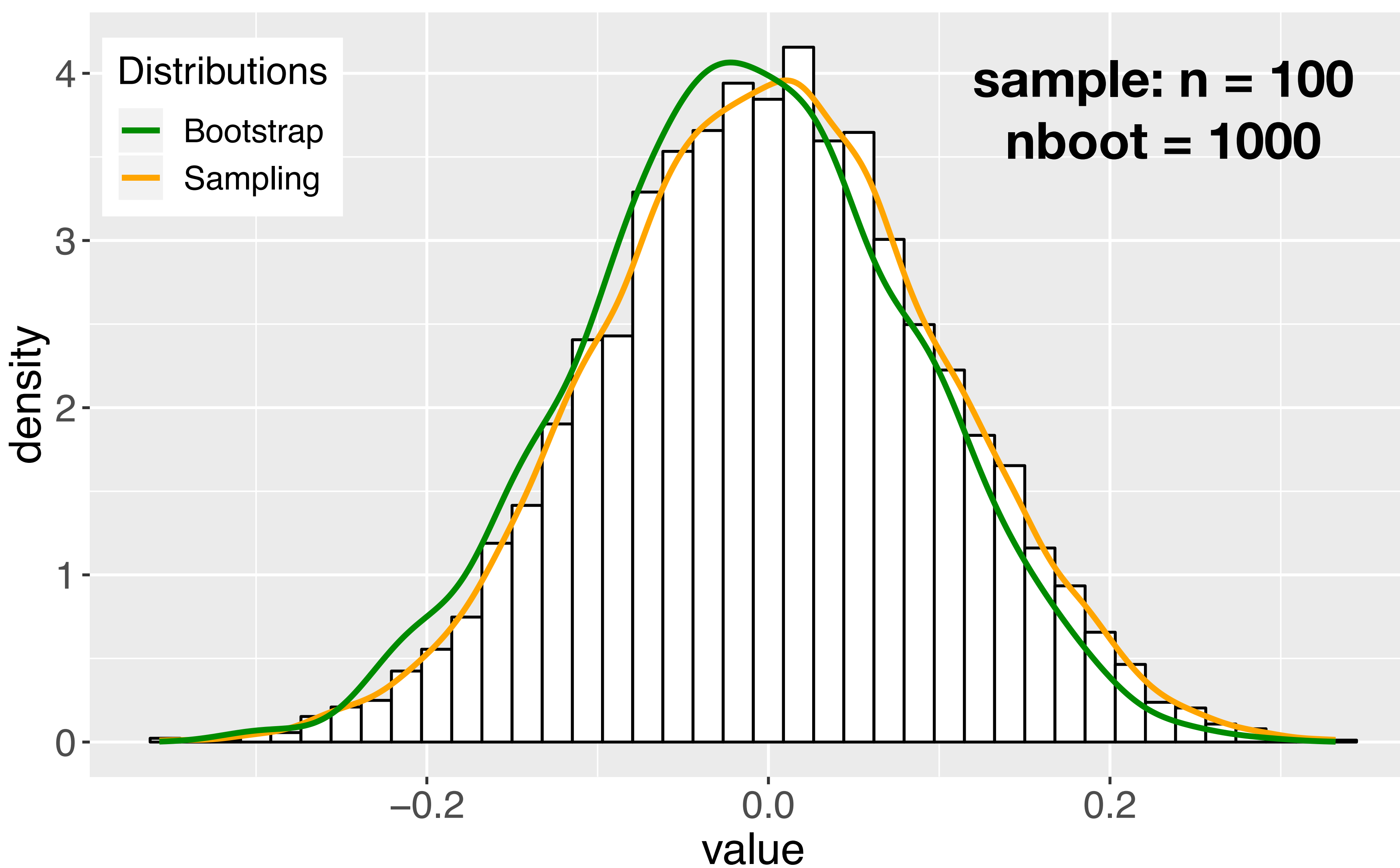
Experiment / bootstrap / simulation



Sampling distribution of the mean



Bootstrap sampling distribution of the mean

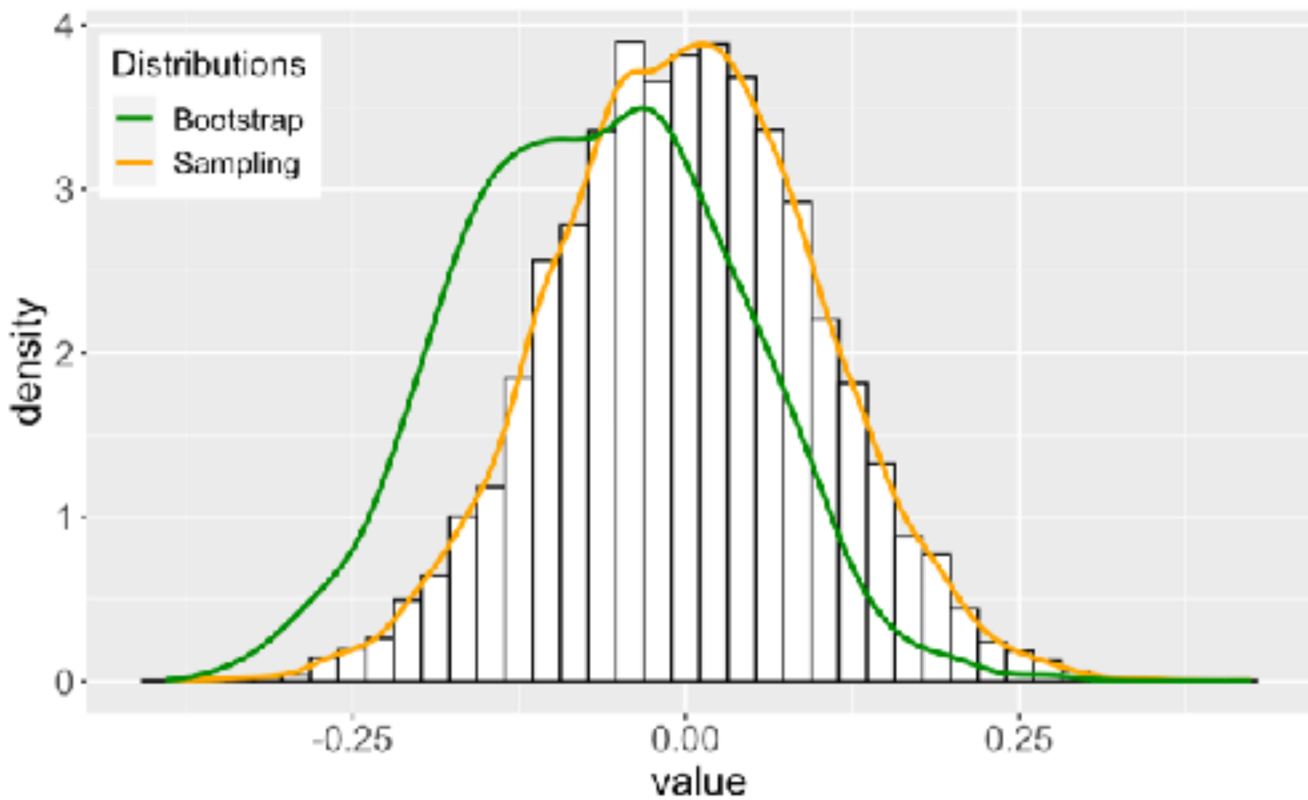


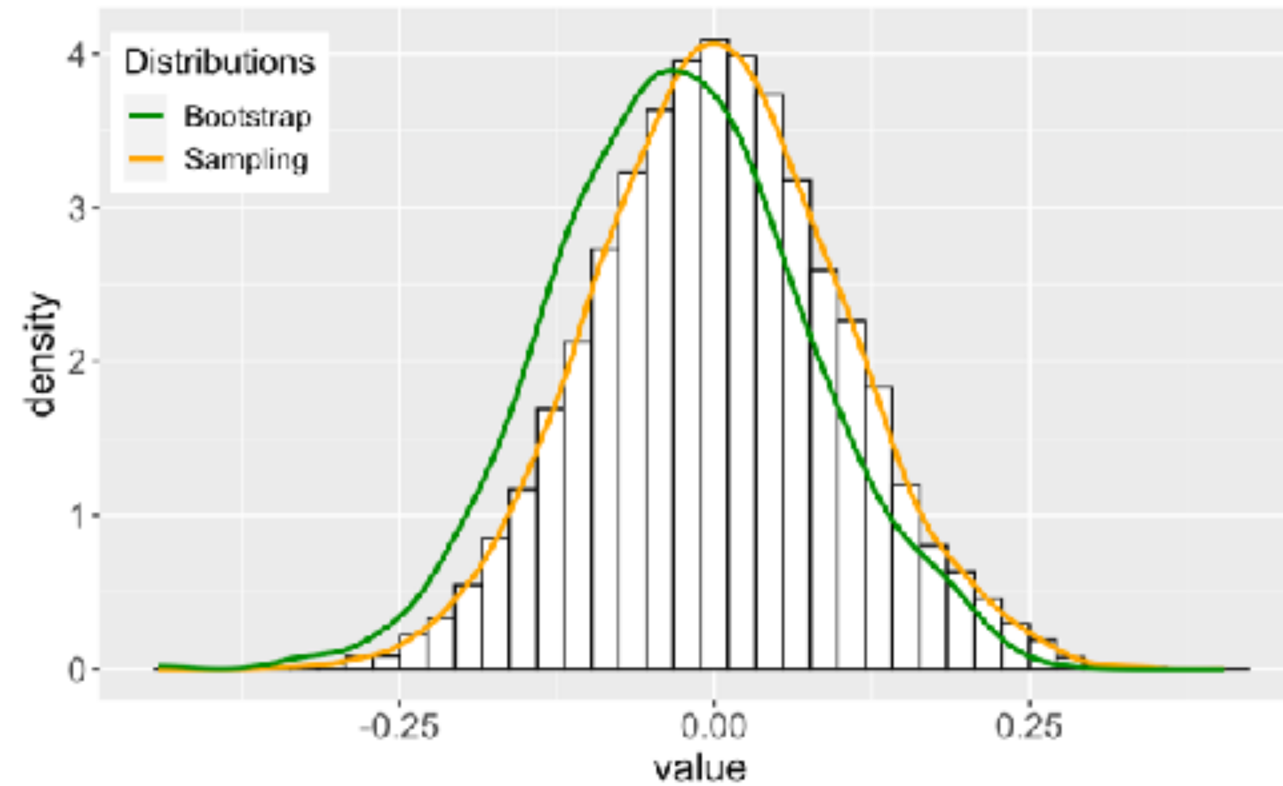
bootstrap SEM = SD(boot means)

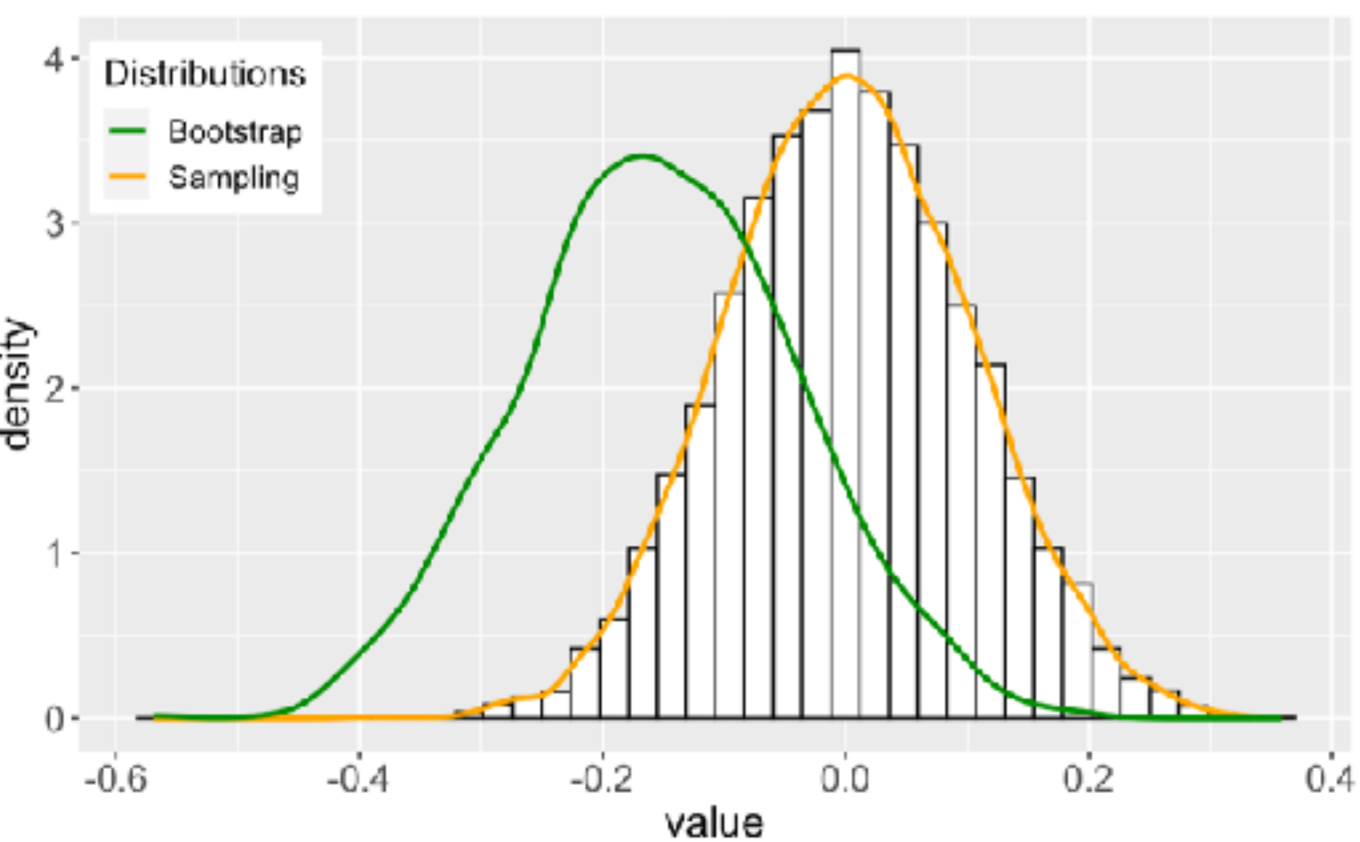
based on one sample only…

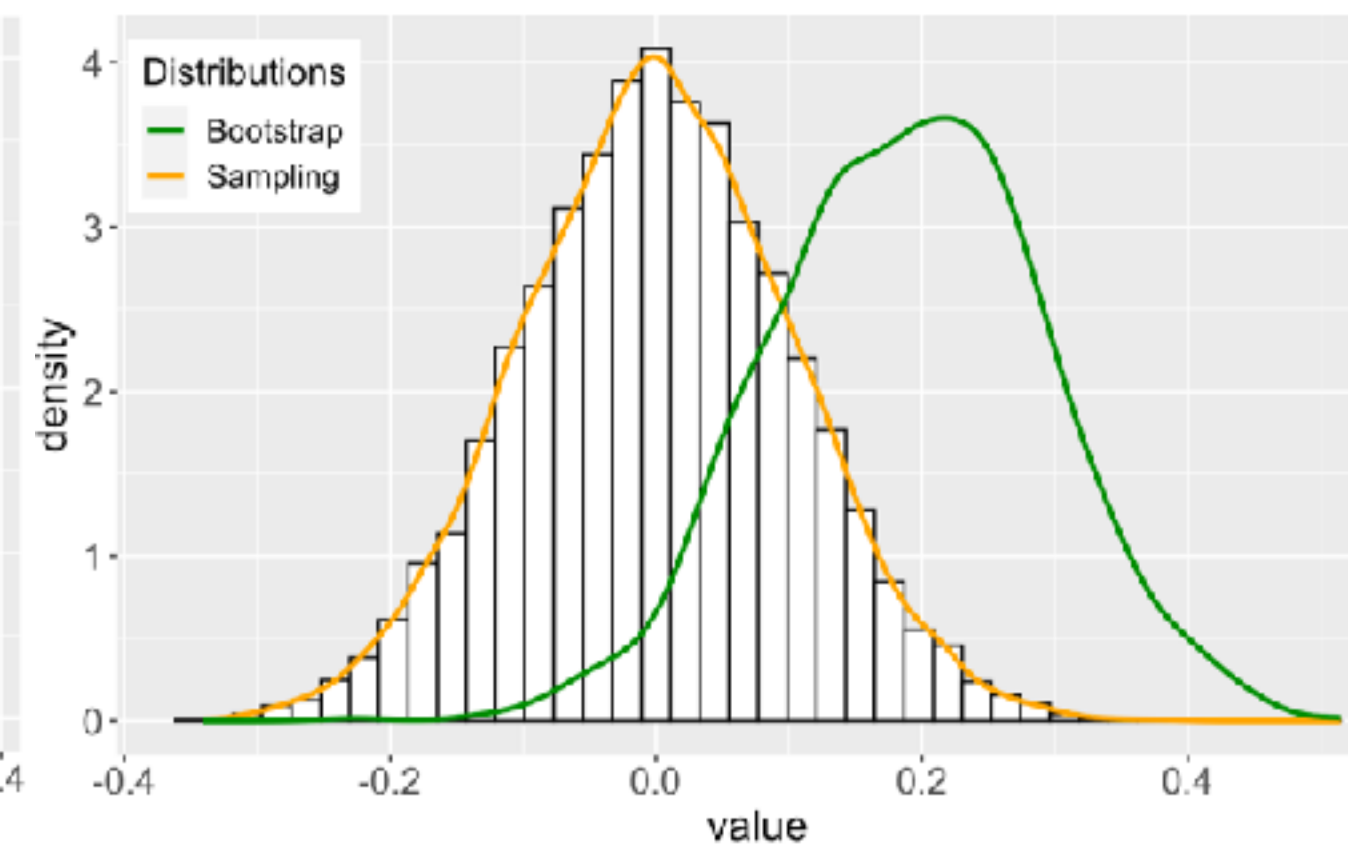
More samples…

Sampling = 10000 experiments; Bootstrap = 1 experiment

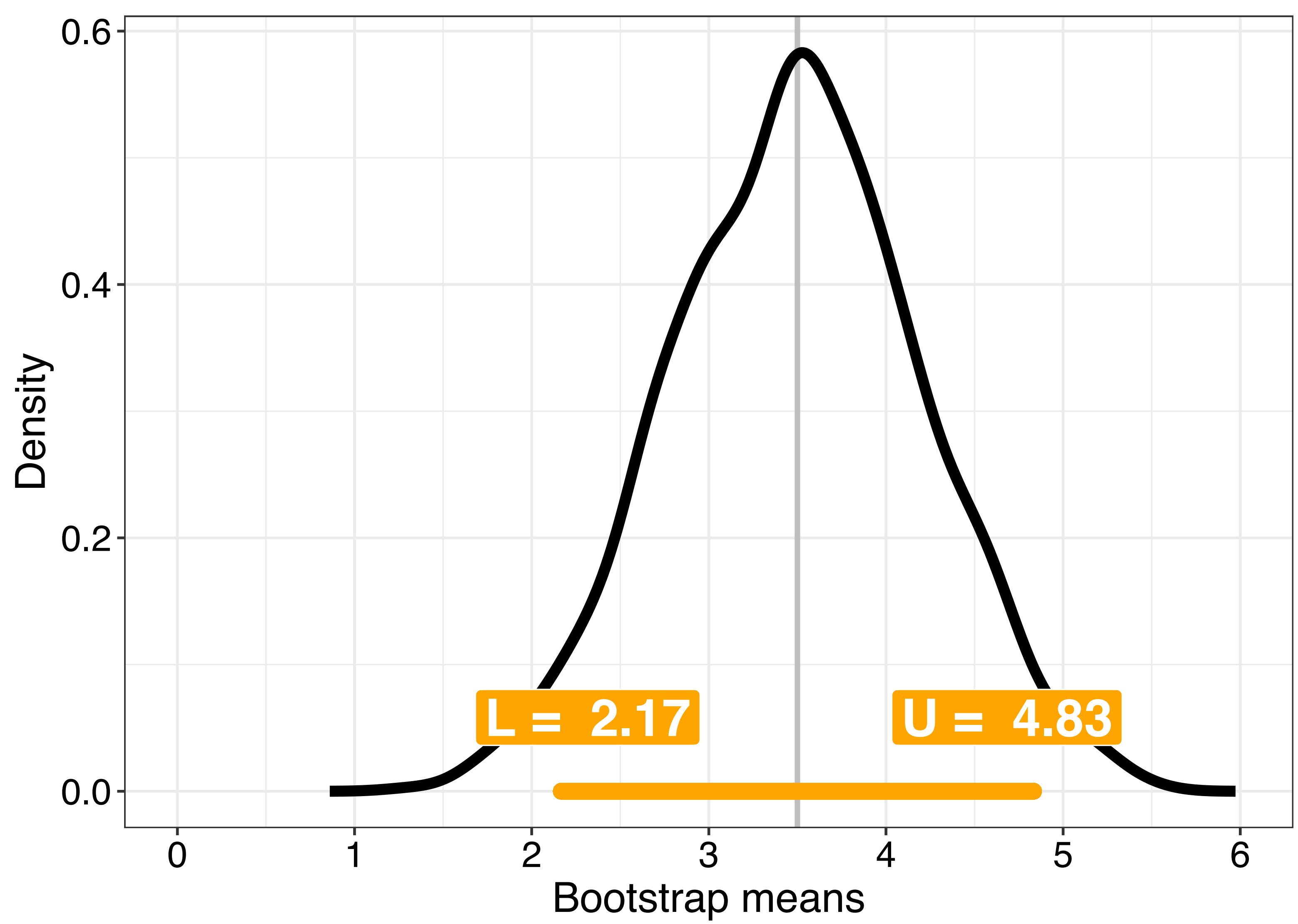






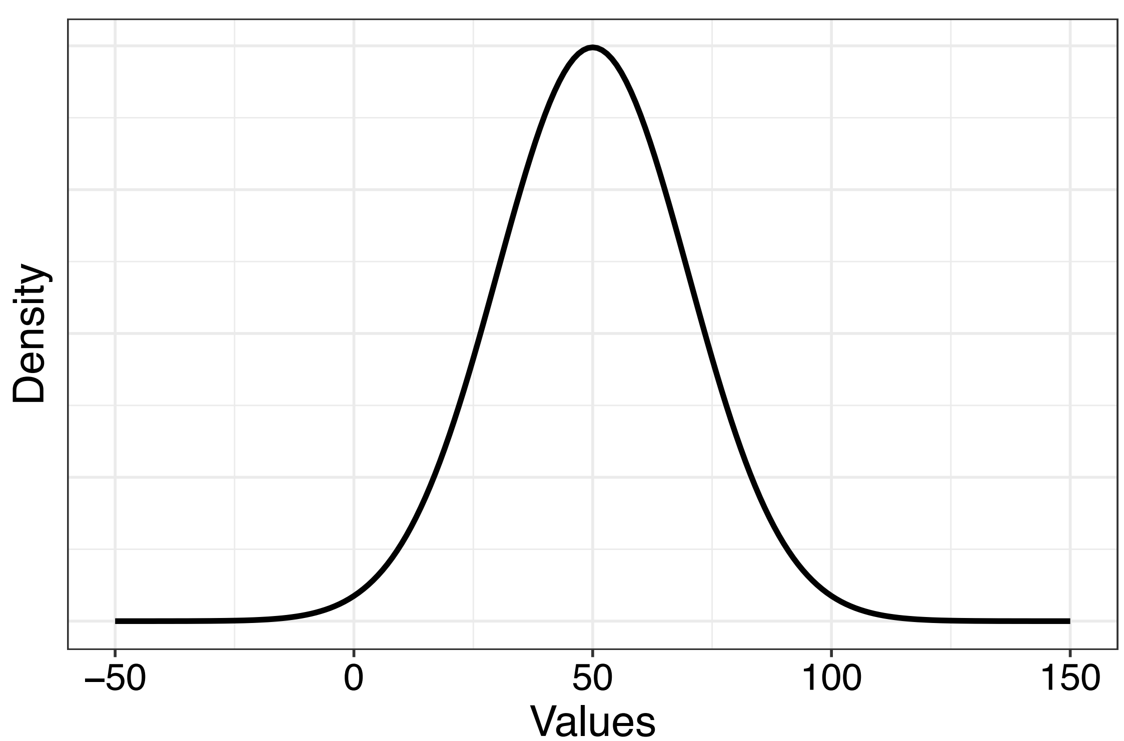


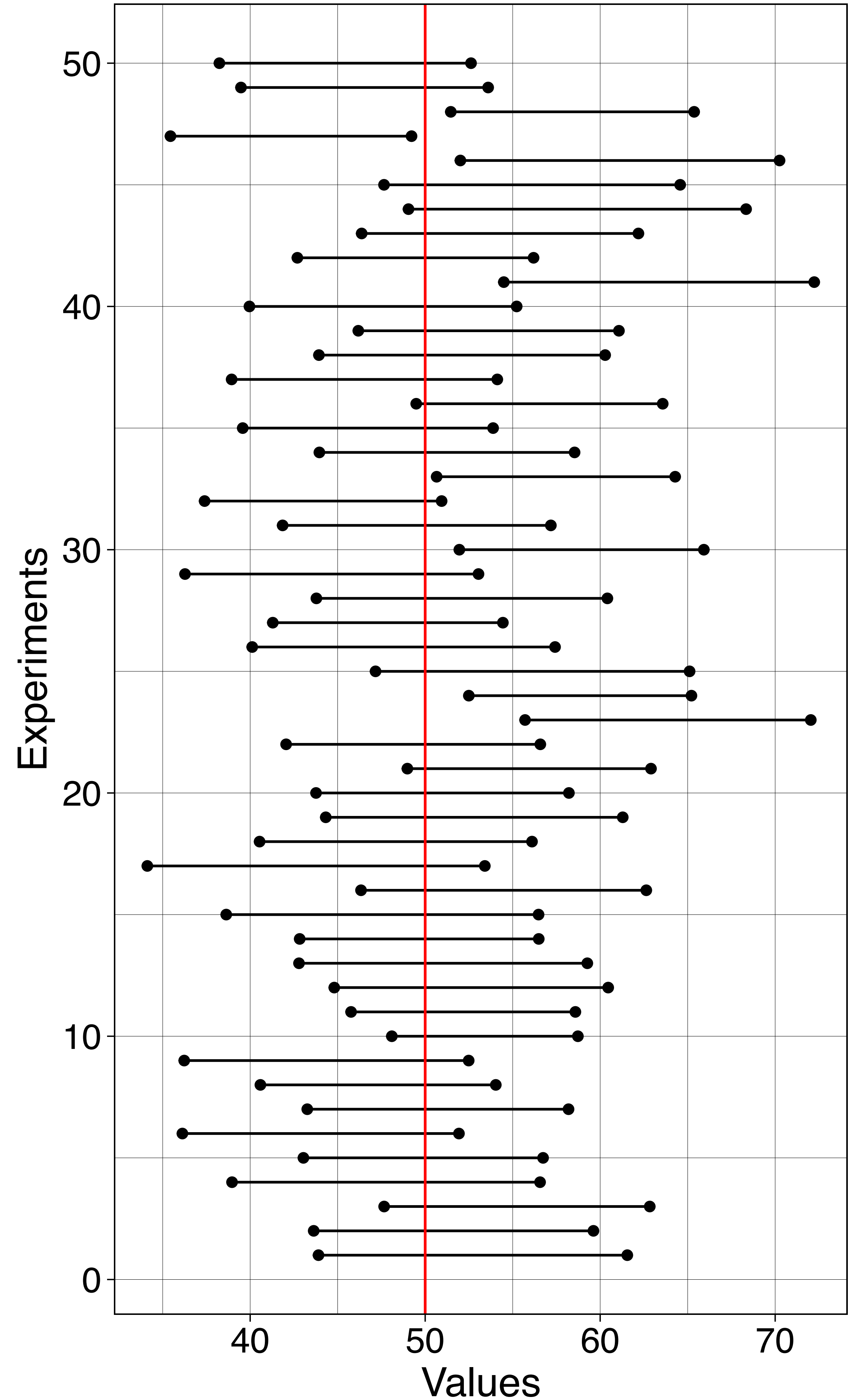
Bootstrap confidence interval




```r
alpha <- 0.05
ci <- quantile(boot.m, probs = c(alpha/2, 1-alpha/2))
```

50 experiments = 50 confidence intervals



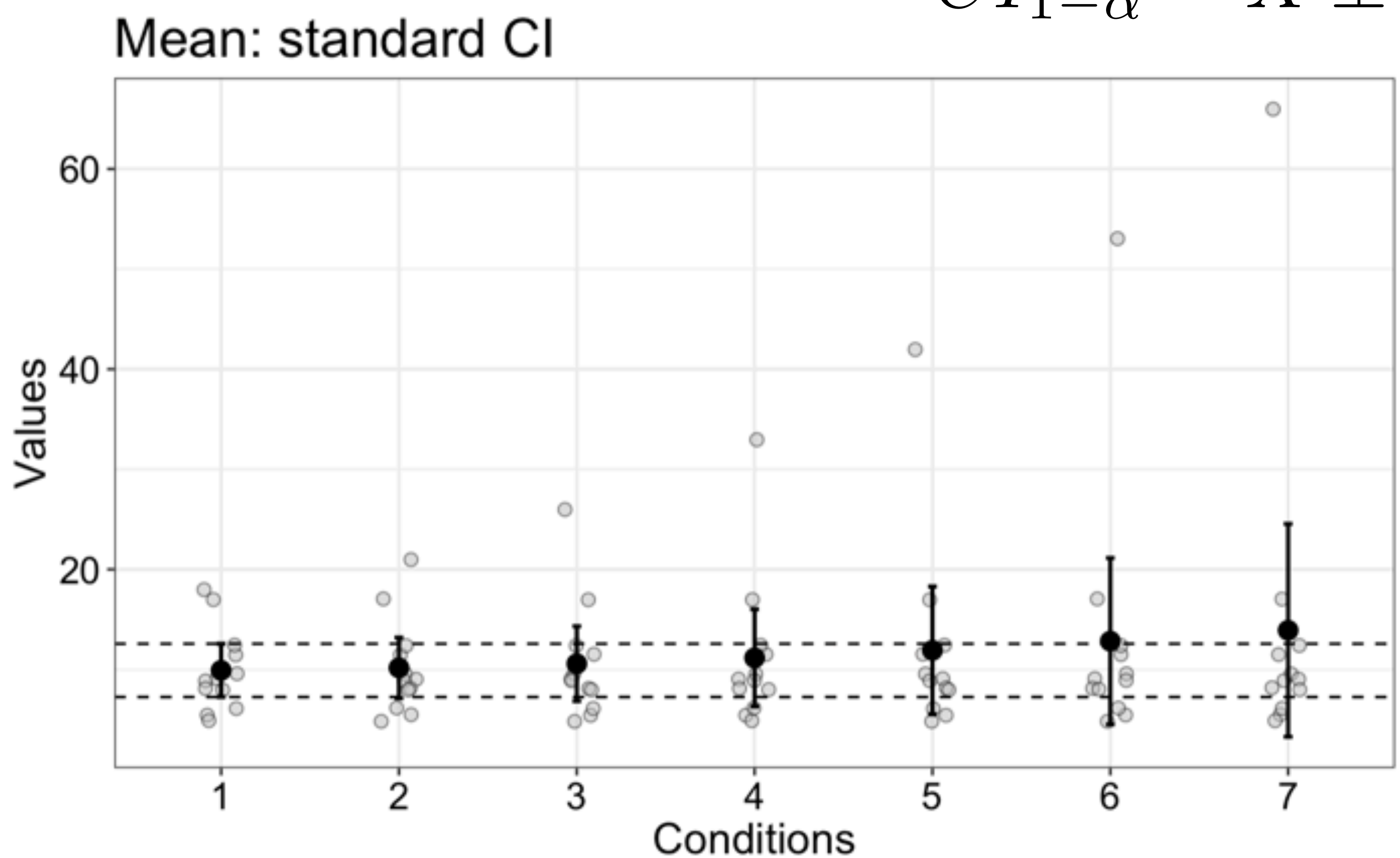


Strengths of the **bootstrap + robust estimates**

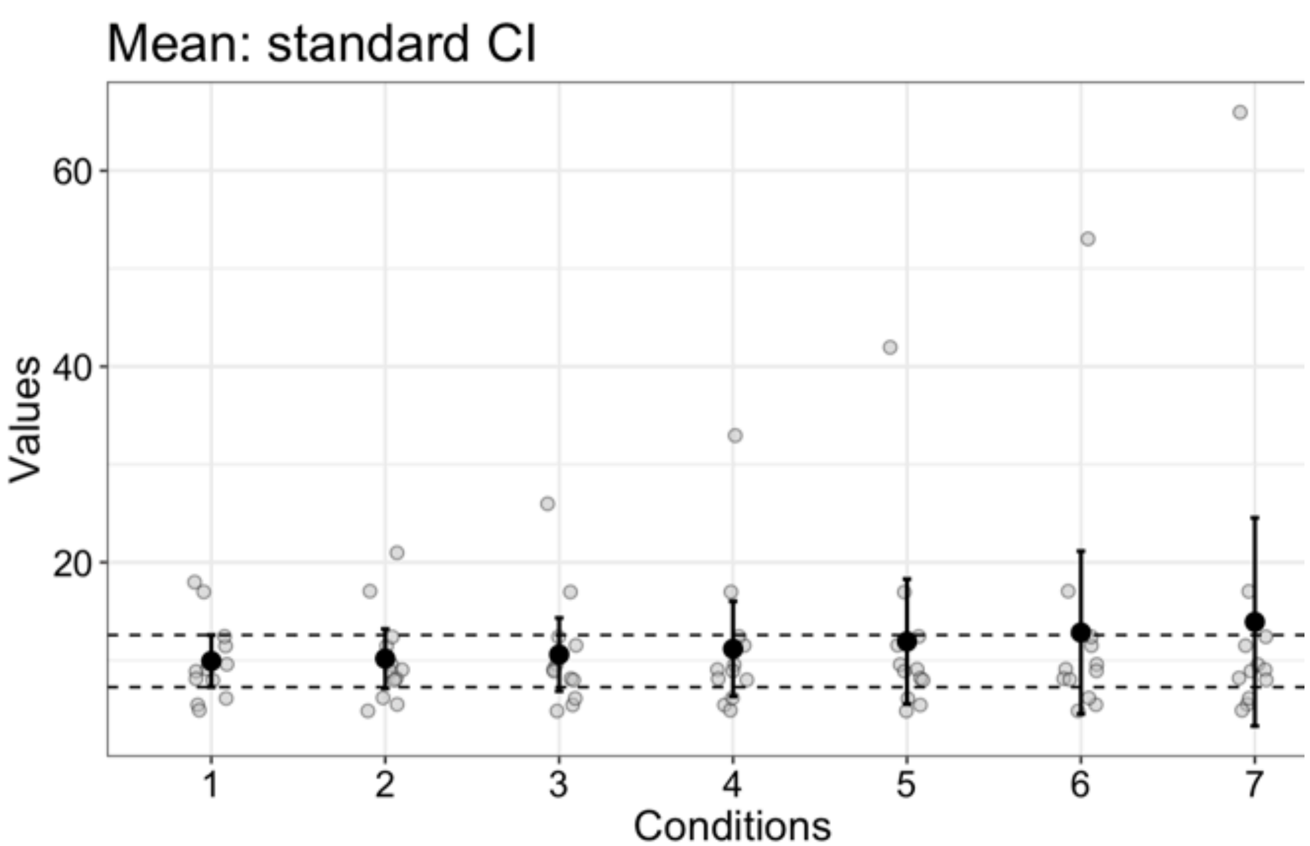
- Robust to heteroscedasticity
- Robust to non-normality
- Robust to outliers
- Confidence intervals can be computed for any statistics
- But no obvious best method...

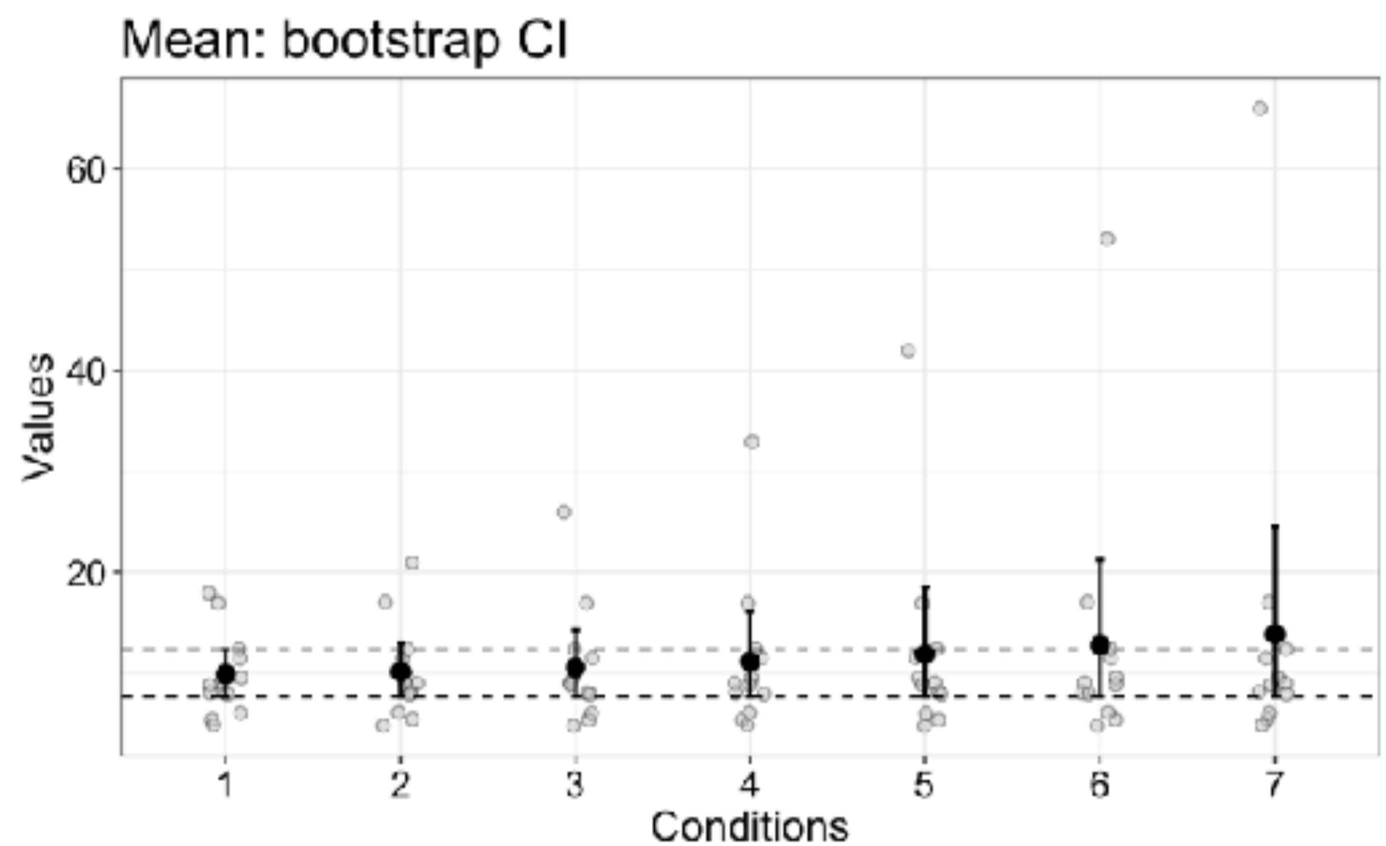
The bootstrap alone is not robust

$$CI_{1-\alpha} = \bar{X} \pm t_{\alpha/2} \frac{s}{\sqrt{n}}$$

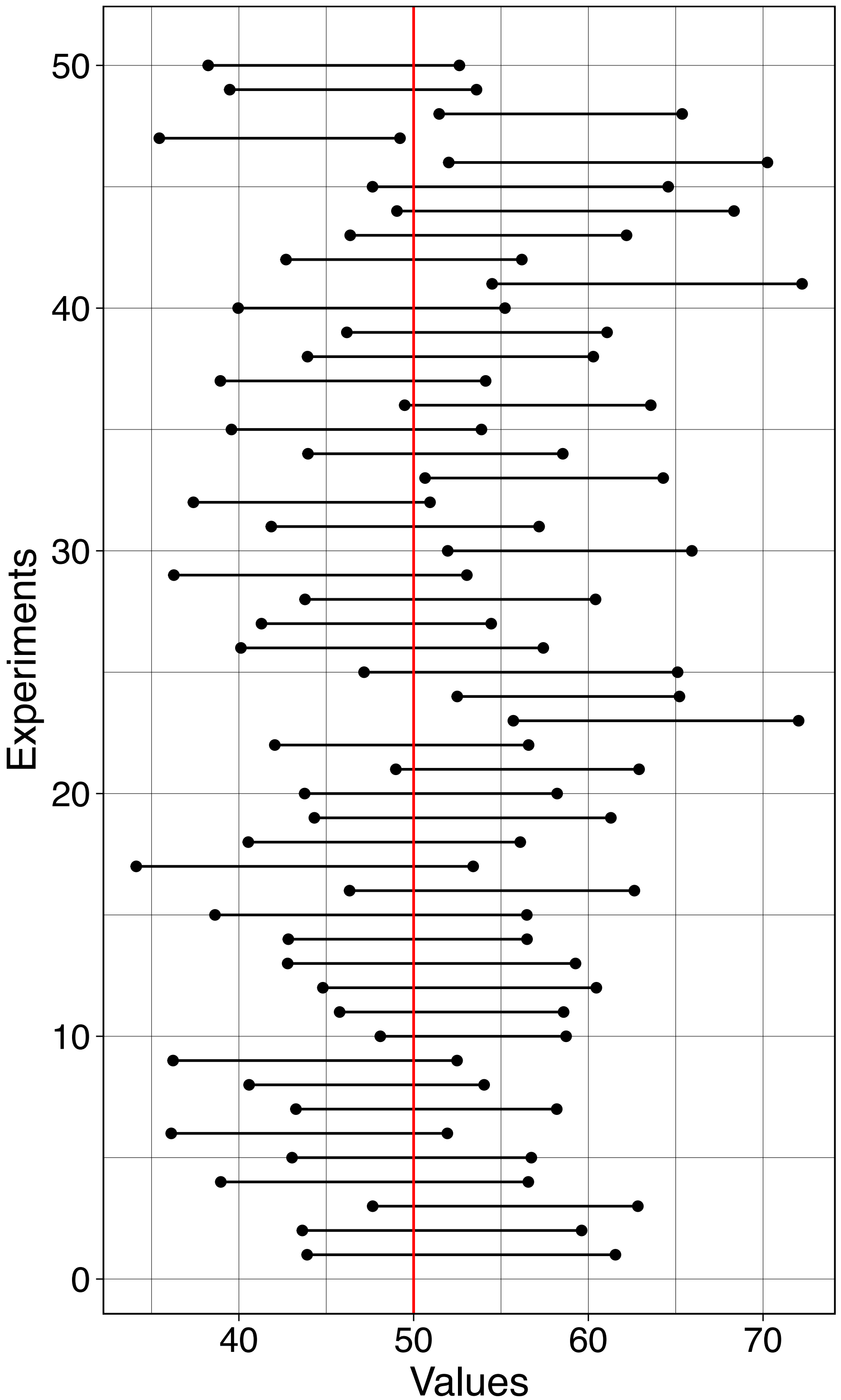


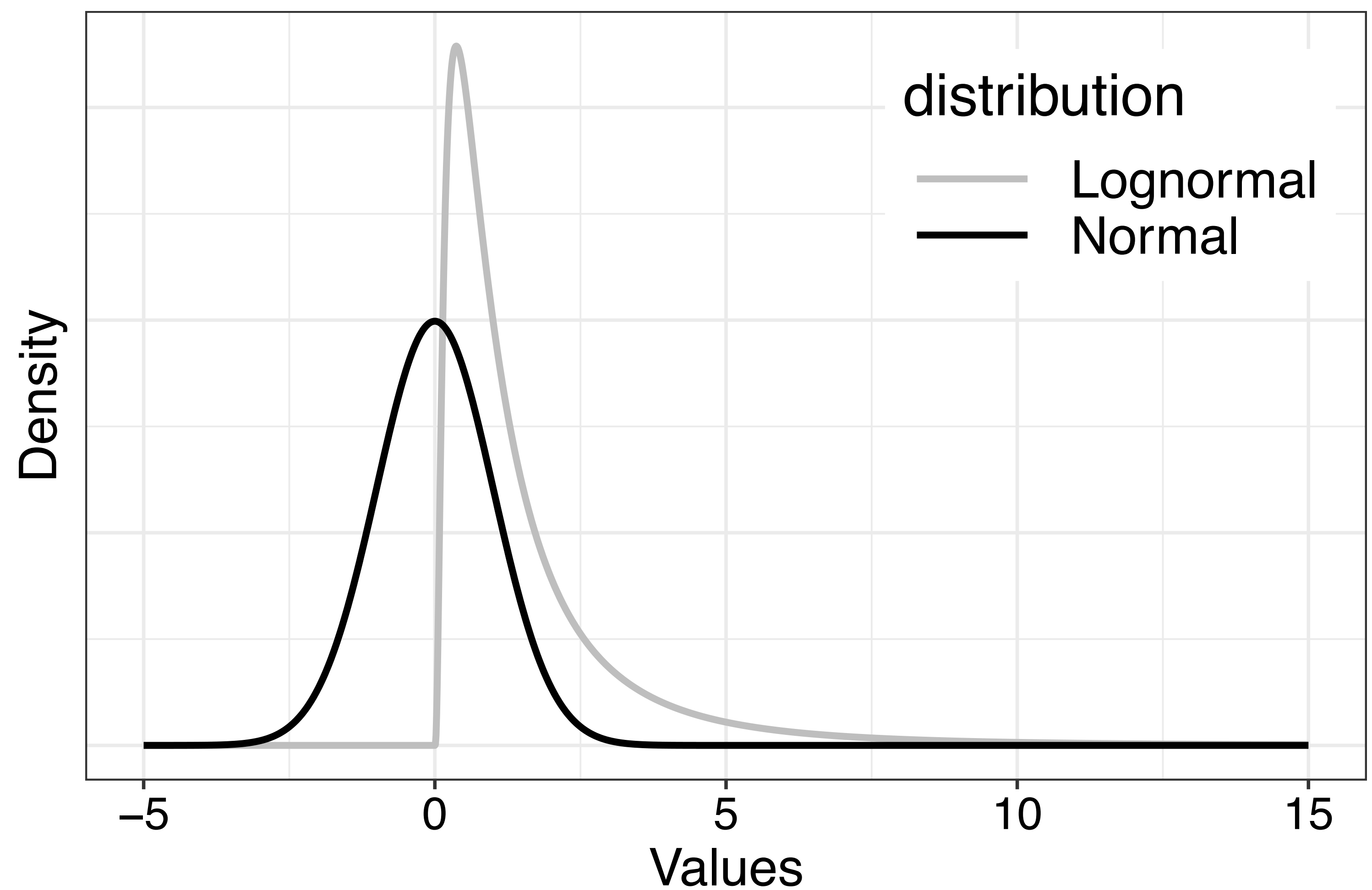
The bootstrap alone is not robust





Confidence interval coverage: expected level?





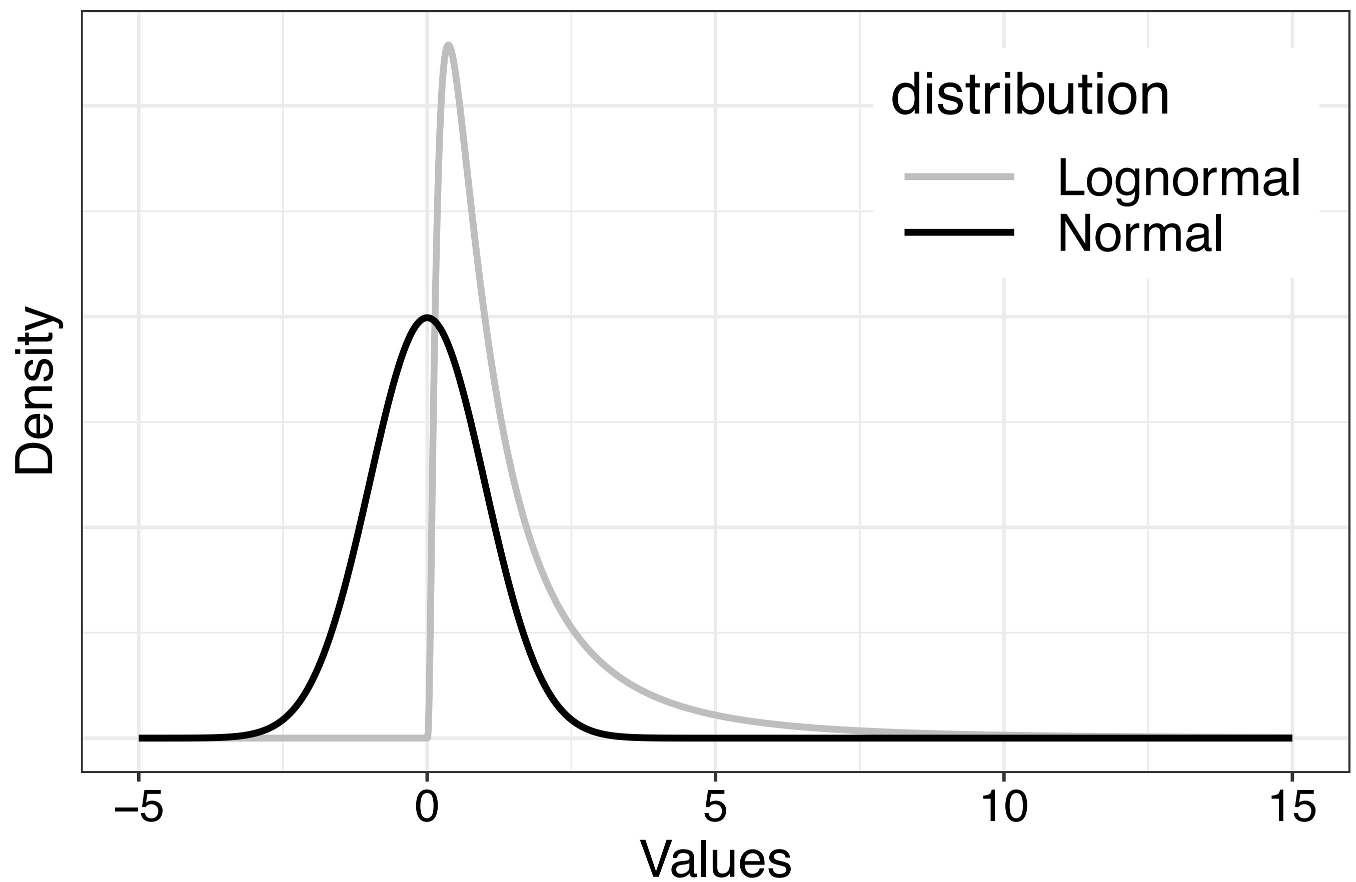
Coverage simulation: when is a 95% confidence interval not a 95% confidence interval?

```r
# Define parameters
nsim <- 10000 # simulation iterations
nsamp <- 30 # sample size
alpha.val <- 0.05
pop <- rnorm(1000000) # define population
pop.m <- mean(pop) # population mean

# declare matrices of results
ci.cov.norm <- matrix(0, nrow = nsim, ncol = 2)

for(S in 1:nsim){ # simulation loop

  # random sample from population
  samp <- sample(pop, nsamp, replace = TRUE)

  # mean + t-test ------------------------------
  ci <- t.test(samp, mu = pop1.m, conf.level = 1-alpha.val)$conf.int
  # CI includes population value?
  ci.cov.norm[S,1] <- ci[1]<pop.m && ci[2]>pop.m

  # mean + percentile bootstrap ----------------
  ci <- onesampb(samp, est=mean, nboot=nboot, trim=0, alpha=alpha.val)$ci
  # CI includes population value?
  ci.cov.norm[S,2] <- ci[1]<pop.m && ci[2]>pop.m

}

apply(ci.cov.norm, 2, mean) # average across simulations for each method
```

0.95

0.93

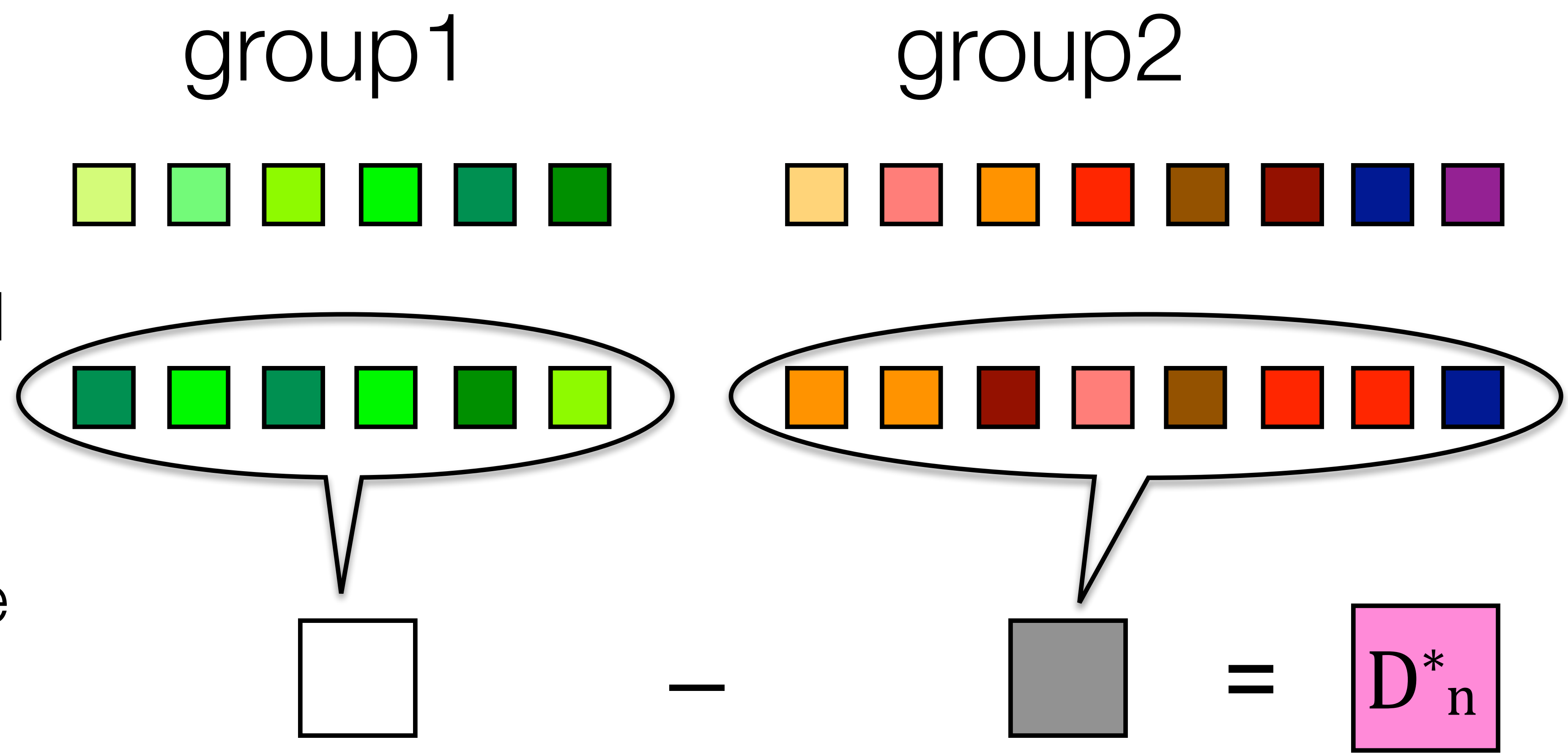
Coverage simulation



	Normal	Log-normal
t-test (mean)	95.2%	88.5%
boot. (mean)	93.6%	87.8
t-test (20% tm)	94.4%	93.4%
pboot. (20% tm)	94.4%	94.4%

n = 30
20,000 iterations
nboot = 2,000

Percentile bootstrap: general recipe



(1) sample WITH replacement n observations

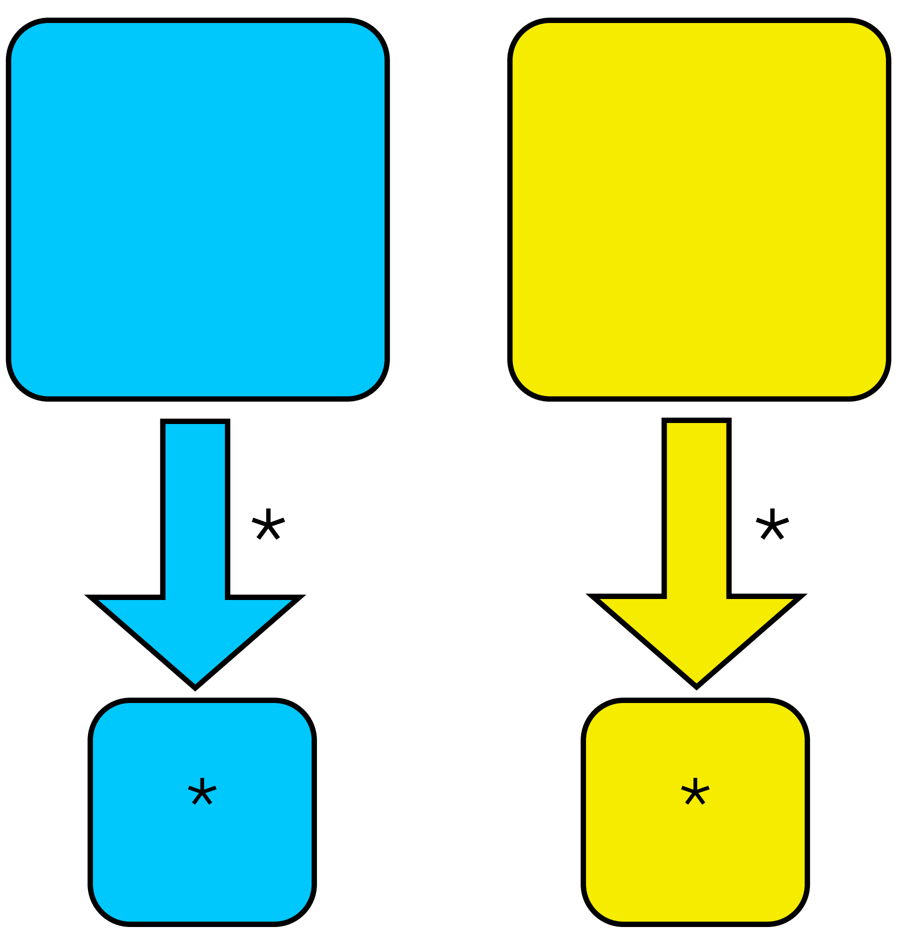
(2) compute estimate e.g. trimmed mean & get difference

repeat (1) & (2) b times
get CI at the 1-alpha level

resampling strategies:
follow the data acquisition process

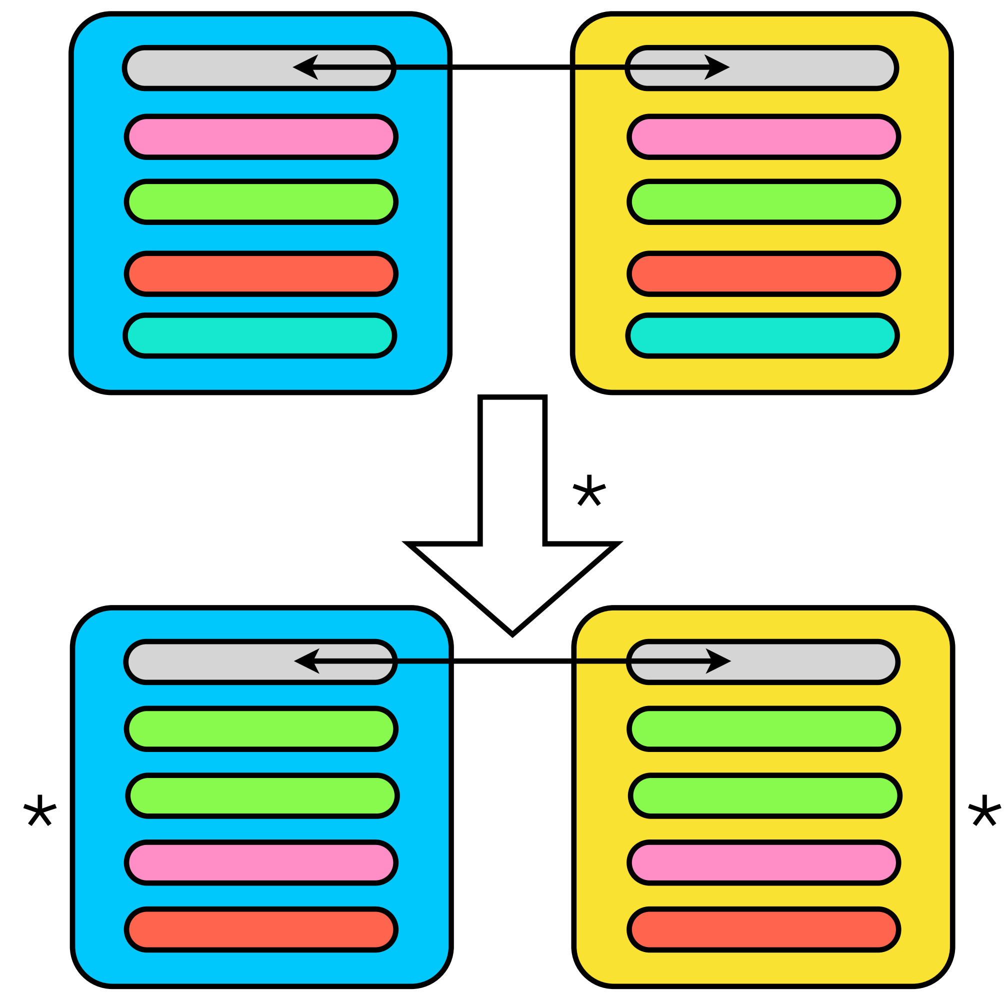
independent sets:

- 2 conditions in single-subject analyses
- 2 groups of subjects, e.g. patients vs. controls



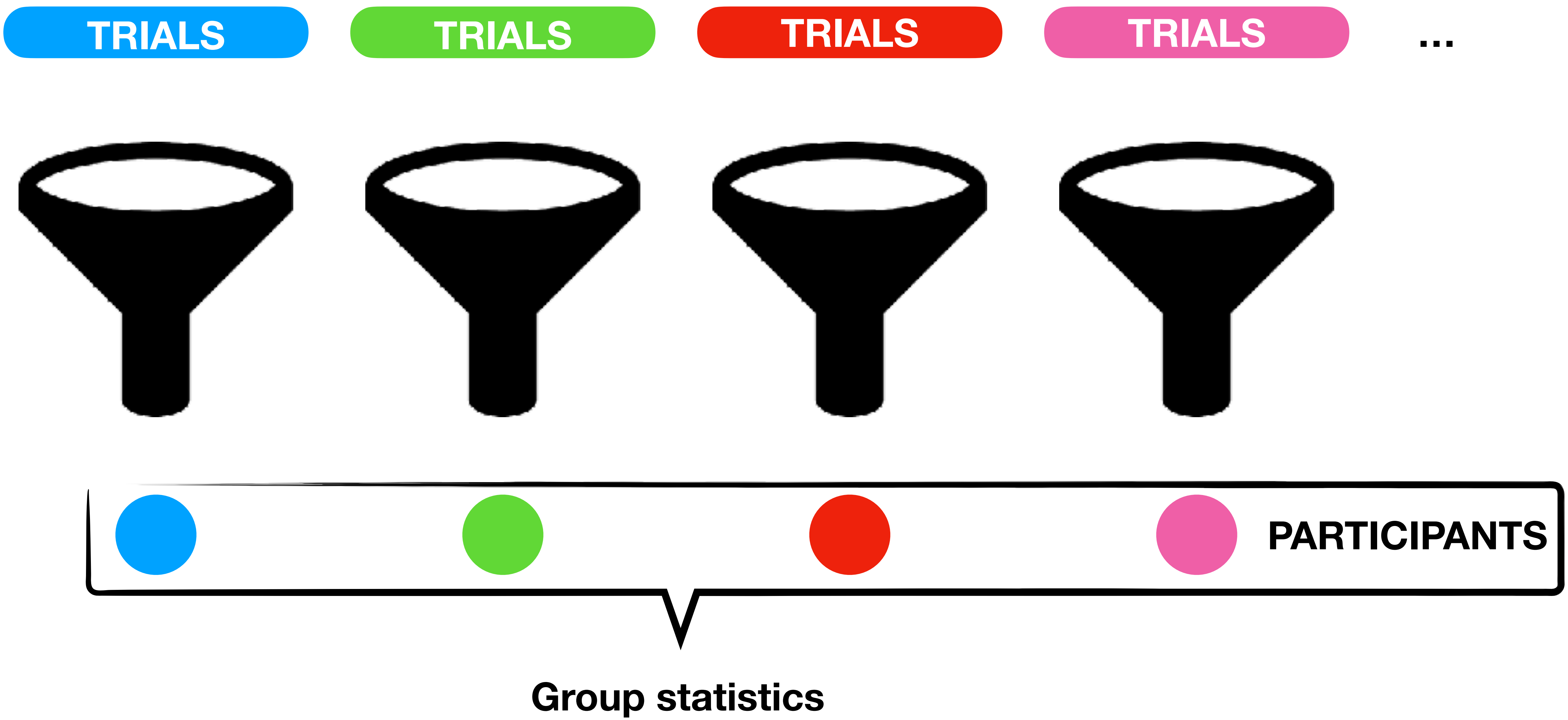
dependent sets:

- 2 conditions in group analyses
- correlations
- linear regression



Hierarchical bootstrap

raw trials/samples: RT, correct/incorrect (0/1), Likert scale, MCQ…



Questions?

Teaching benefits

introduce or consolidate:

- **key frequentist concepts (sampling distributions, SE, confidence intervals…)**
- **inferential statistics**
- **experimental design (how do I bootstrap my data?)**
- **robust statistics**
- **simulations**
- **R skills (including graphical representations)**
- **dealing with distributions of plausible population values -> Bayes**