Bayesian Statistical Inference

Tuulia Malén
PhD student, the University of Turku
Turku PET Centre, group of Professor Lauri Nummenmaa: Human Emotion Systems Laboratory (aka Emotion Lab)
PhD project: The human dopamine system, funded by Päivikki and Sakari Sohlberg Foundation
Licenced psychologist, undergraduate student in Statistics

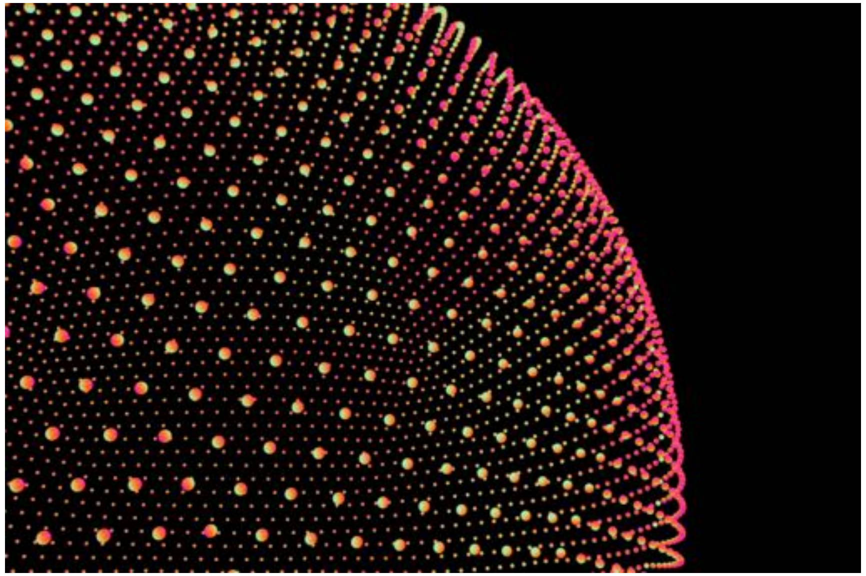
Contents

Statistical inference

”Statistical inference is the process of drawing conclusions about an ***underlying population*** *based on a* ***sample*** *or subset of the data.”*

Thomas D. Gauthier, Mark E. Hawley (2015). Introduction to Environmental Forensics (Third Edition).

Statistical inference = Translating data into understanding

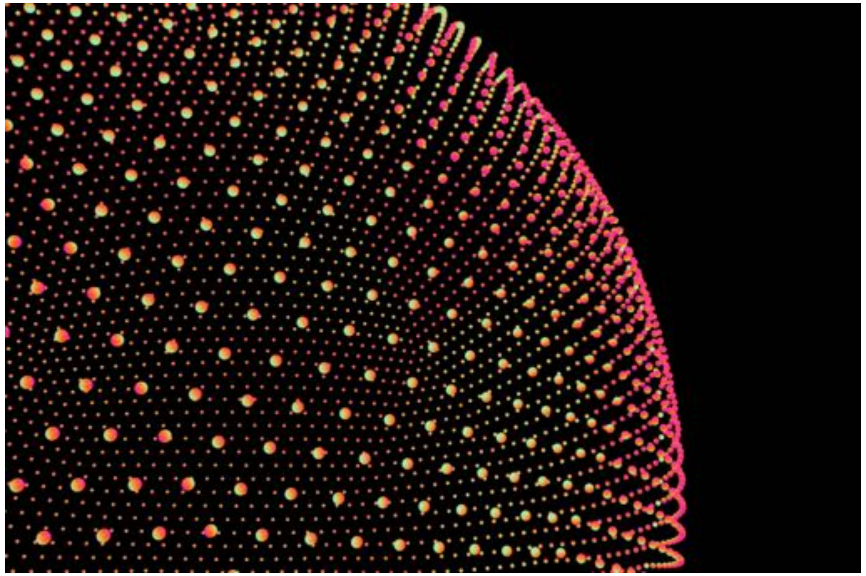


Data

Underlying population

Thomas D. Gauthier, Mark E. Hawley (2015). Introduction to Environmental Forensics (Third Edition).; PowerPoint image bank.

Statistical inference = Translating data into understanding

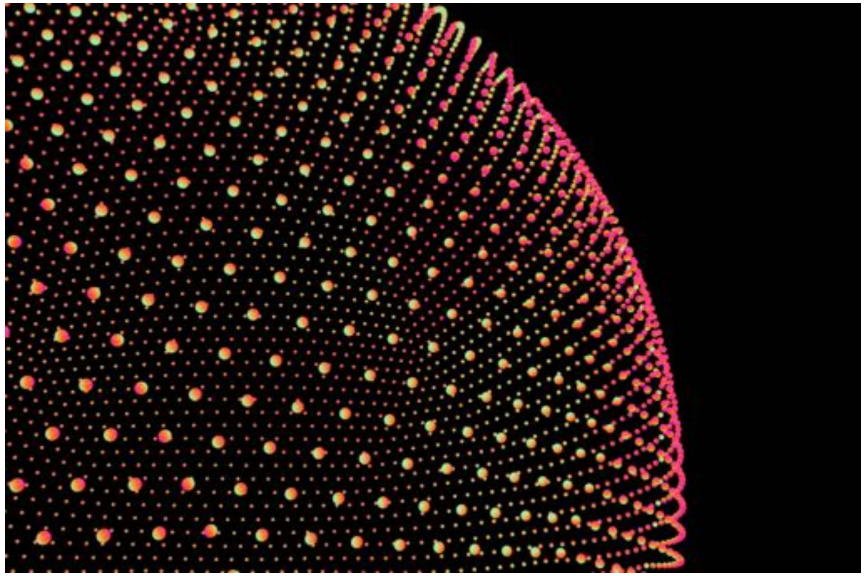


Example: How much people love statistics?

- Subset: not practical to ask everyone
- Sample bias: poll at stats dept. – generalize to others?
- Measuring variation: bad day?

Holopainen, M. & Pulkkinen P. (2012). *Tilastolliset menetelmät*. Sanoma Pro Oy.; PowerPoint image bank.

Statistical inference



$\neq$

→ Dealing with **uncertainty**:

Gap

The heart and soul of statistical inference

McElreath, R. (2018). Statistical rethinking: A Bayesian course with examples in R and Stan. Chapman and Hall/CRC.; PowerPoint image bank.

Describing uncertainty: Probability

0 ——— Likelihood of an event ——— 1

$$P(\text{green}) = \frac{1}{5} = 0.2$$

Holopainen, M. & Pulkkinen P. (2012). *Tilastolliset menetelmät*. Sanoma Pro Oy.; PowerPoint image bank.

Describing uncertainty: Probability

0 — Likelihood of an event — 1

$$P(\text{green}) = \frac{1}{5} = 0.2$$

Rarely that simple in research!
But a key concept that can be applied in various ways

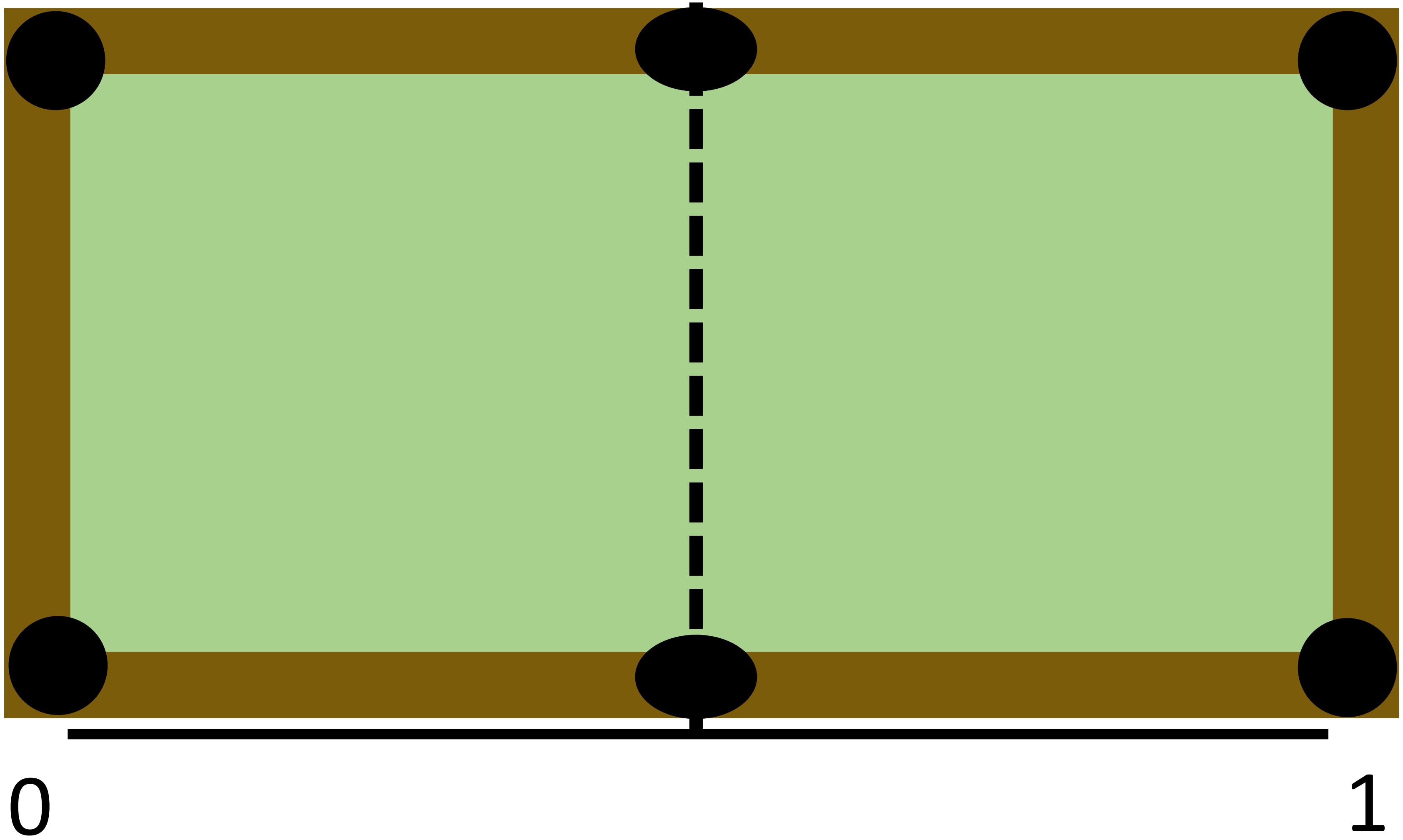
Holopainen, M. & Pulkkinen P. (2012). *Tilastolliset menetelmät*. Sanoma Pro Oy.; PowerPoint image bank.

Two approaches to statistical inference

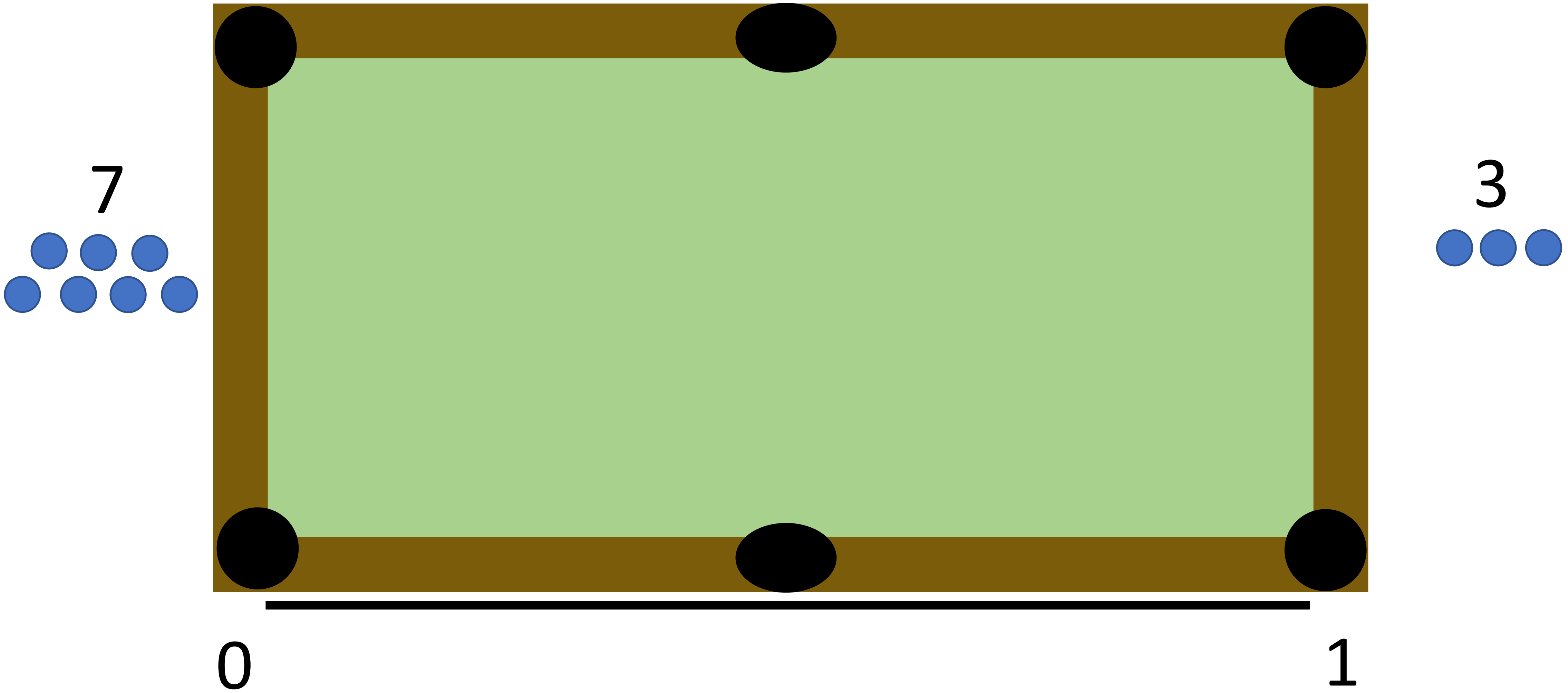
1. Bayesian
2. Frequentist

Gelman, A., Carlin, J. B., Stern, H. S., & Rubin, D. B. (1995-2020). Bayesian data analysis. Chapman and Hall/CRC.; PowerPoint image bank.

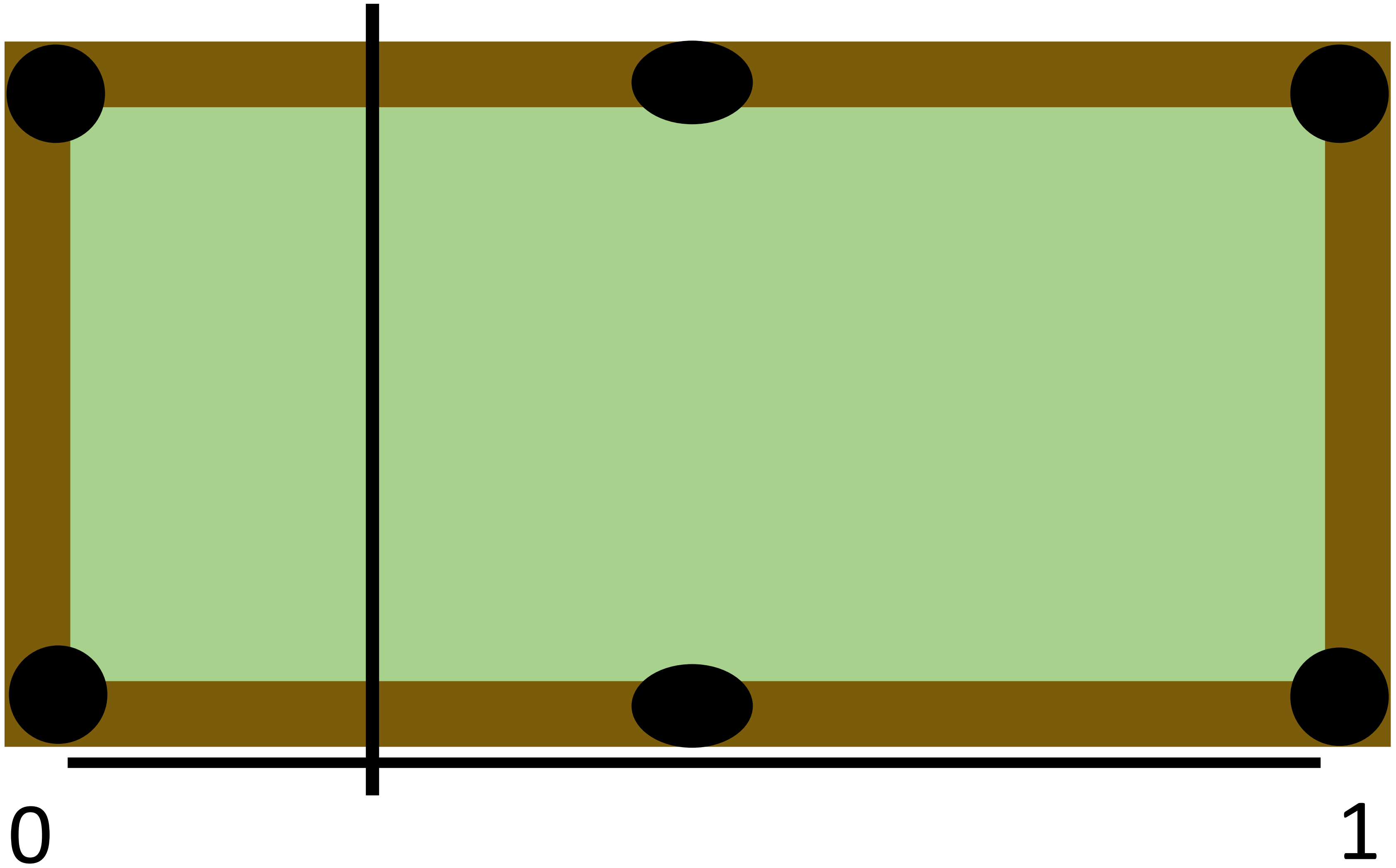
Thomas Bayes



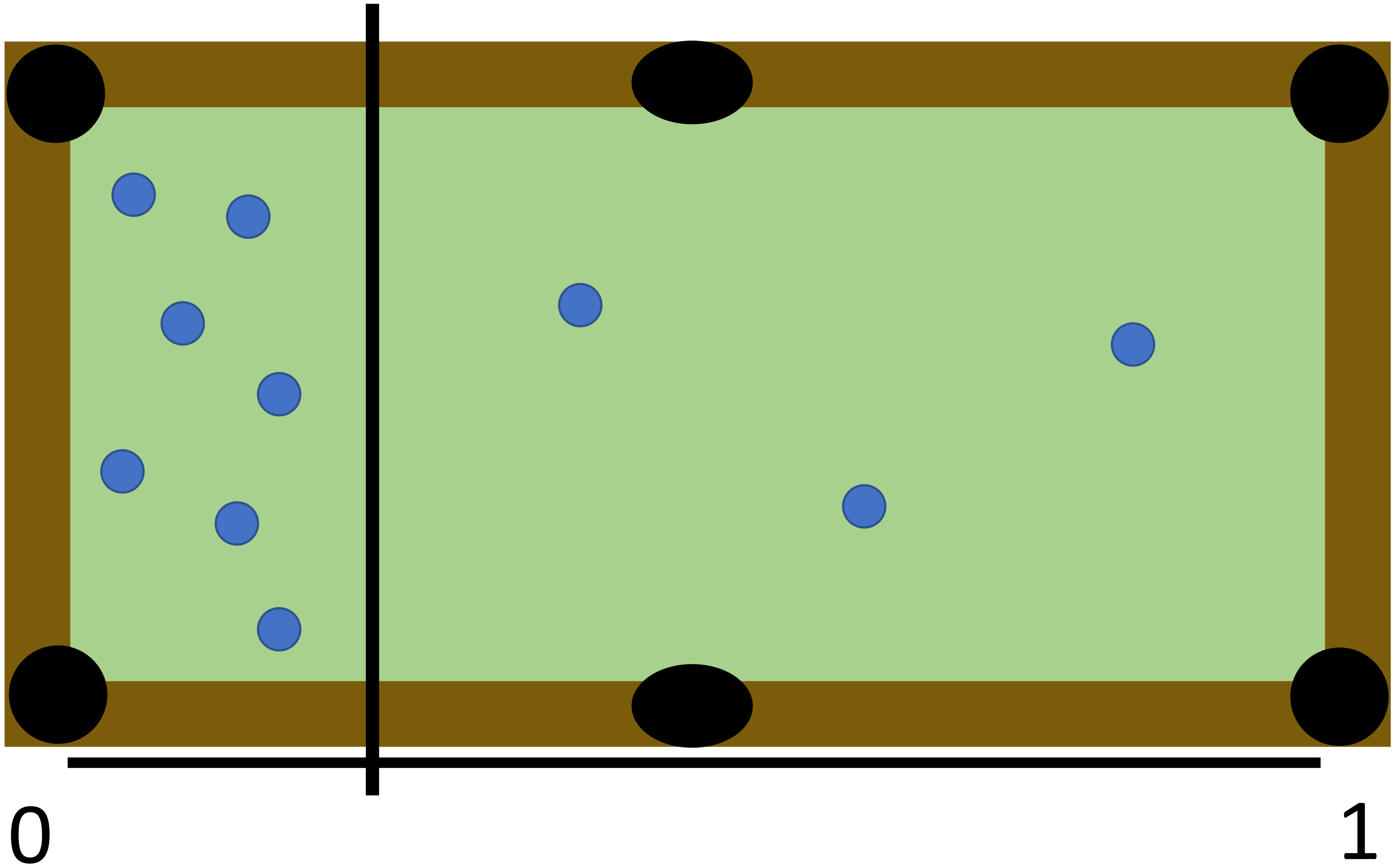
Gower, B. (1997). Scientific method. Routledge.



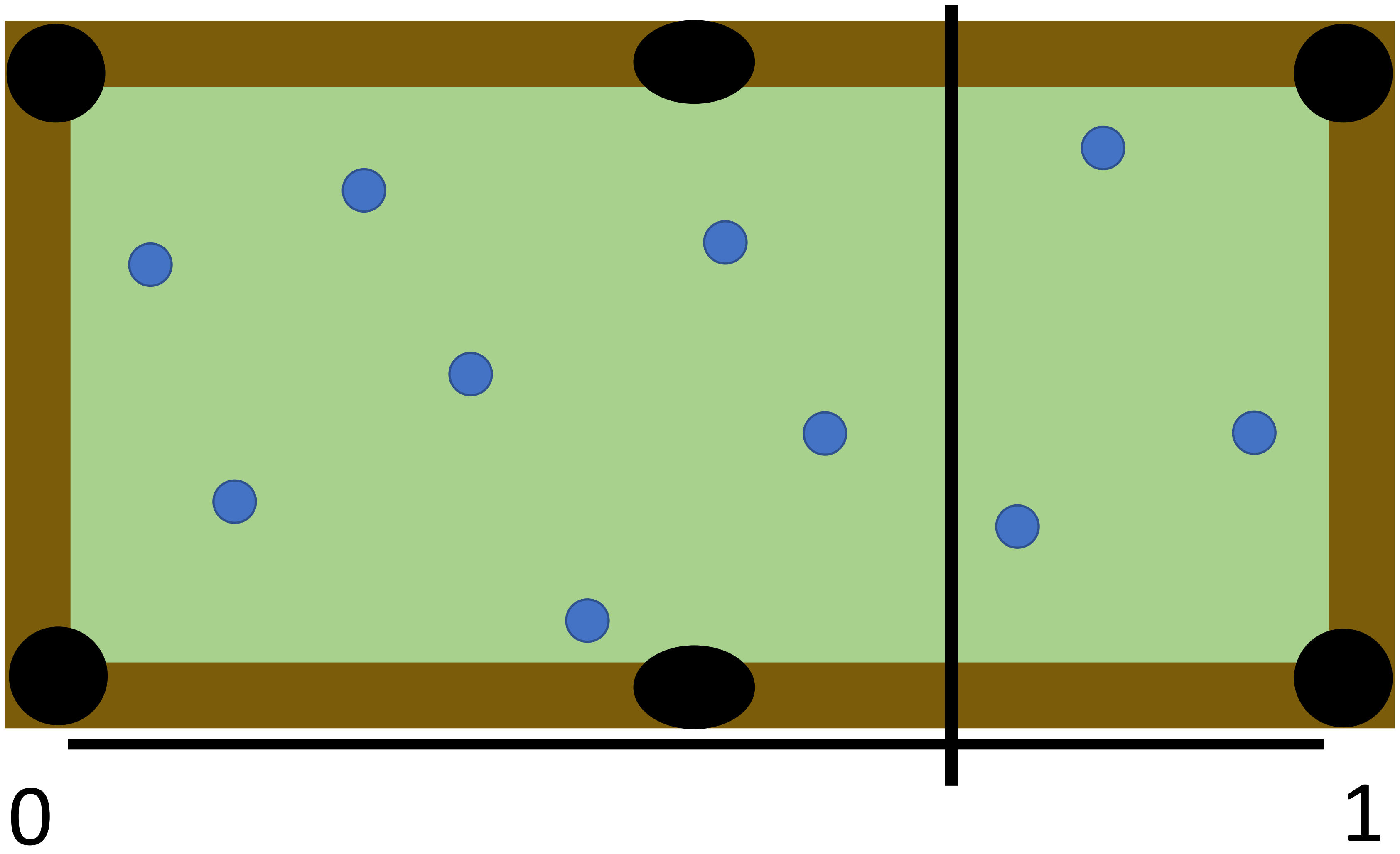
Gower, B. (1997). Scientific method. Routledge.



Gower, B. (1997). Scientific method. Routledge.

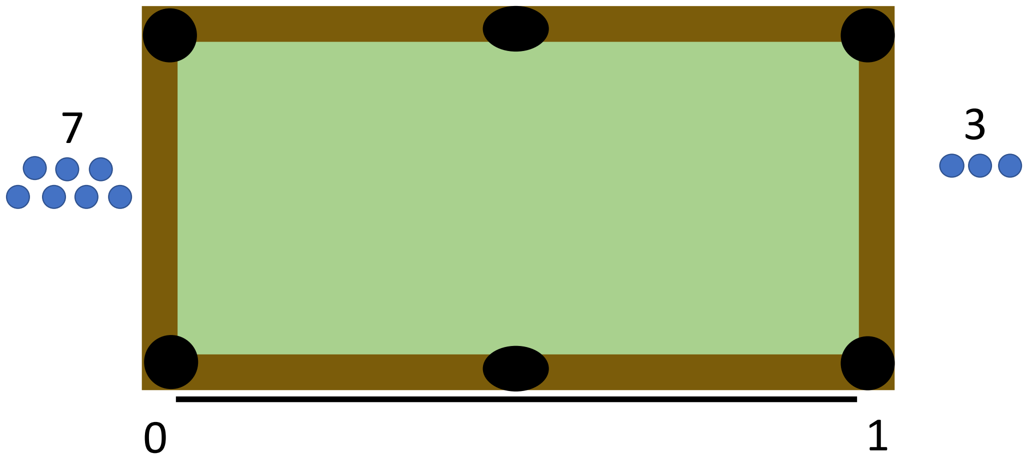


Gower, B. (1997). Scientific method. Routledge.



Gower, B. (1997). Scientific method. Routledge.

Where is the line?



1) What do we know prior to data?

- There is a line, somewhere from 0 to 1

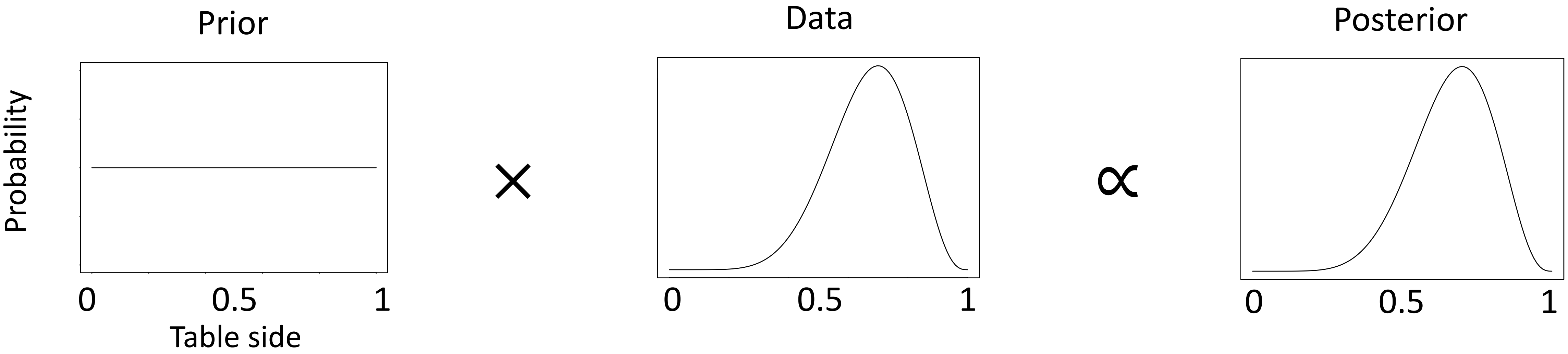
2) What do we observe?

- 7 on the left
- 3 on the right

Both the **prior** knowledge and the **data** (observations) affect our **posterior** understanding

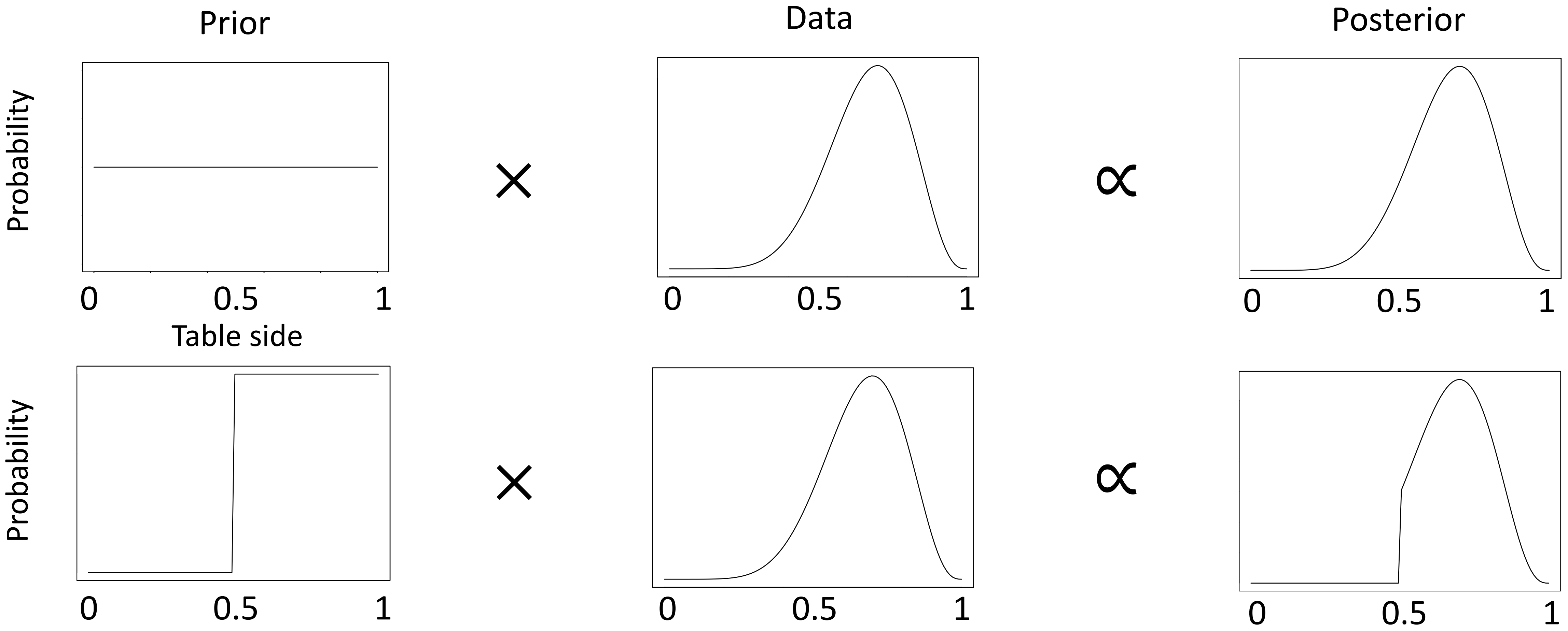
Gelman, A., Carlin, J. B., Stern, H. S., & Rubin, D. B. (1995-2020). Bayesian data analysis. Chapman and Hall/CRC.

Where is the line?



McElreath, R. (2018).; R Core Team (2020). R: A language and environment for statistical computing.

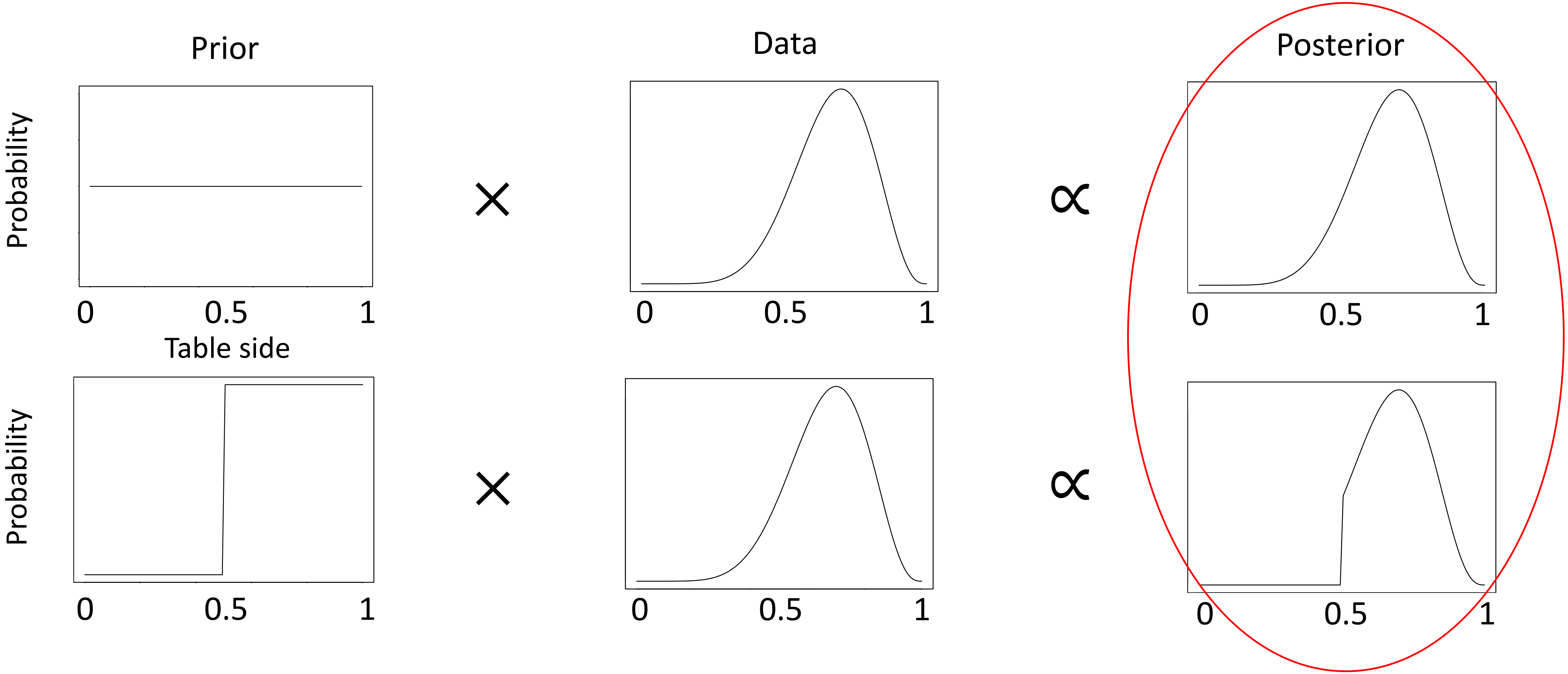
Where is the line?



McElreath, R. (2018).; R Core Team (2020). R: A language and environment for statistical computing.

Where is the line?

Can be an iterative process:
If we got more data,
we could use the posterior as the new pior



McElreath, R. (2018).; R Core Team (2020). R: A language and environment for statistical computing.

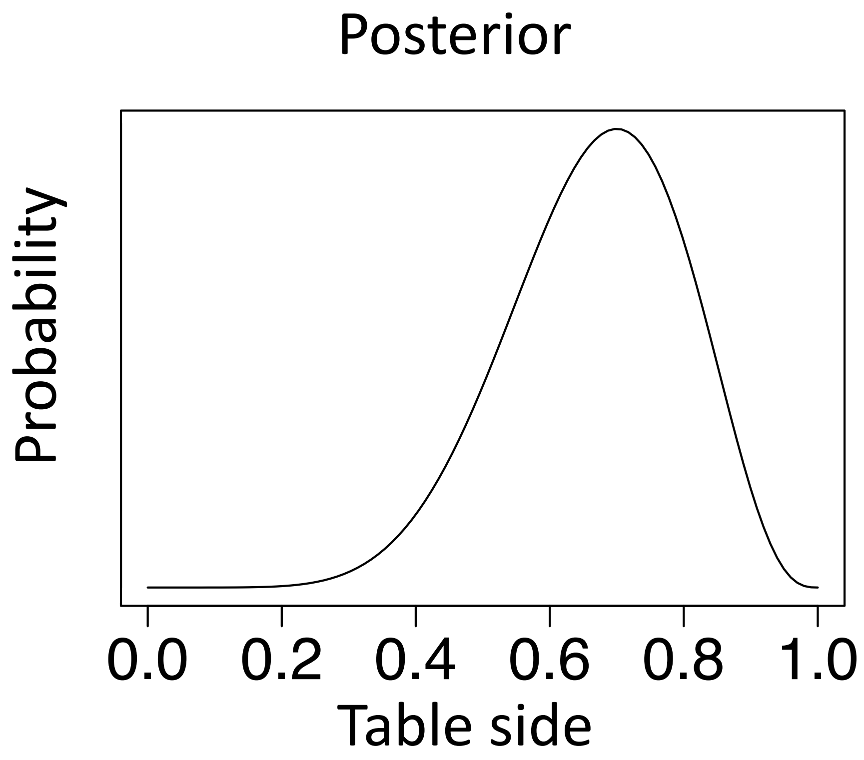
The result is a probability distribution for the parameter

Parameter= the unknown ground truth we try to solve

Probability for each possible location of the line
But the line is somewhere, we just don't know it

(50% probability one has a virus, is pregnant…)
What does it mean to be 50% likely pregnant?
One is or not

Subjective probability:
- Give parameter probabilities
- Get posterior probabilities back (updated with data)



Where is the probability?

Bayesian accepts **subjective**:
Given the data, what is the probability distribution of the parameter?

Data

Sample
Observable
Known
(everybody accepts probability here)

Parameter

Population
Unknown
Interest
’State of world’
Truth

McElreath, R. (2018).; Holopainen, M. & Pulkkinen P. (2012). *Tilastolliset menetelmät*. Sanoma Pro Oy.

Where is the probability?

Data

Sample
Observable
Known
(everybody accepts probability here)

Bayesian accepts **subjective**:
Given the data, what is the probability distribution of the parameter?

Frequentist does not:
Assuming the parameter (H0), how probable is the data?

Parameter

Population
Unknown
Interest
'State of world'
Truth

McElreath, R. (2018).; Holopainen, M. & Pulkkinen P. (2012). *Tilastolliset menetelmät*. Sanoma Pro Oy.

Example: Laugh

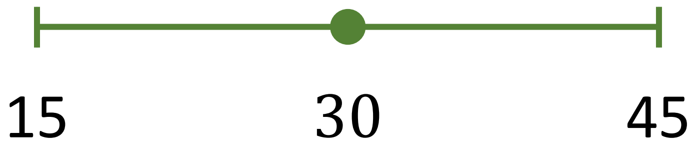
- How much people laugh?

- Sample n= 100 people
- Laughtermeter: Measure how long a subject laughs during a day

- Parameter: true mean laughing time in the population

Example: Laugh

Bayesian 95% credible interval



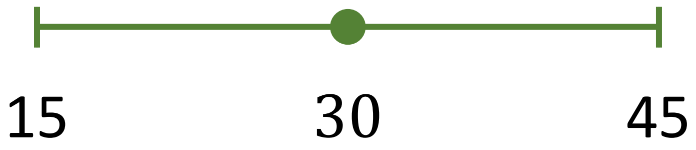
95% probability that people laugh 15-45 min a day
Most likely 30 min

McElreath, R. (2018).; Data: mock-up!

θ = true parameter value

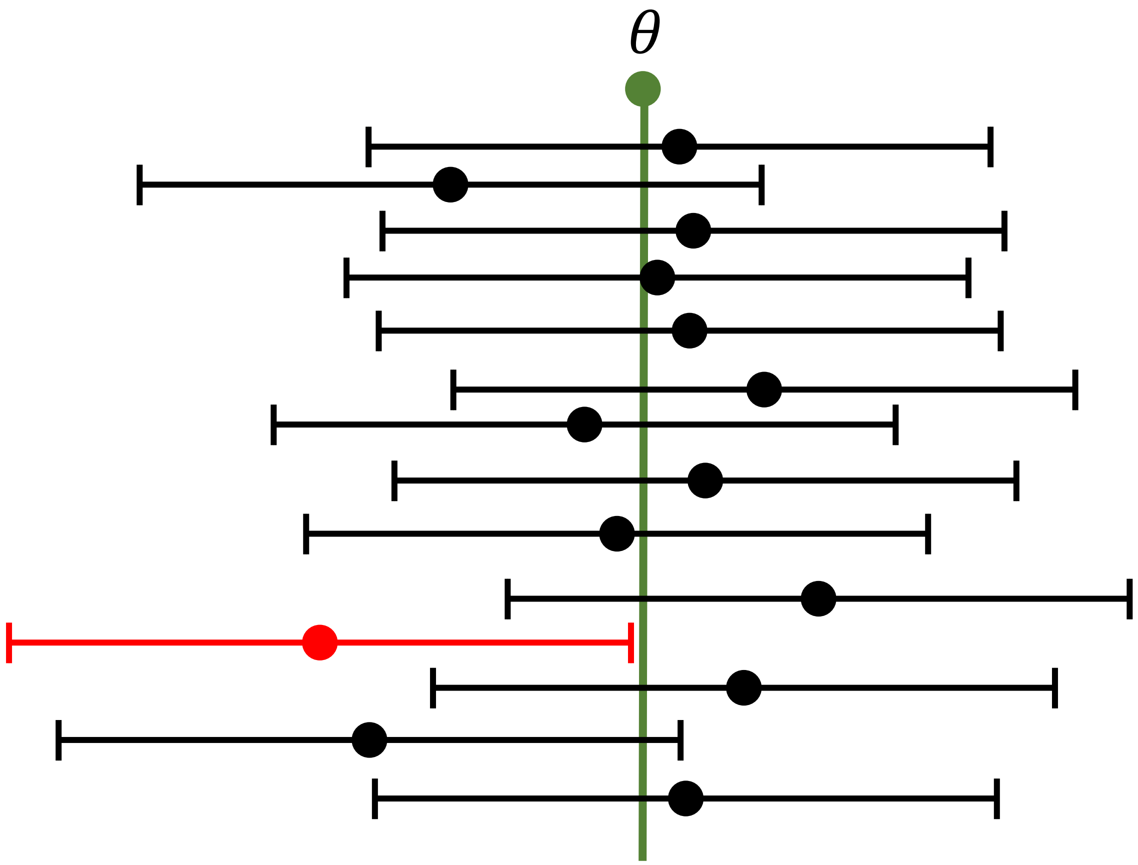
Example: Laugh

Bayesian 95% credible interval



95% probability that people laugh 15-45 min a day
Most likely 30 min

Frequentist 95% confidence interval



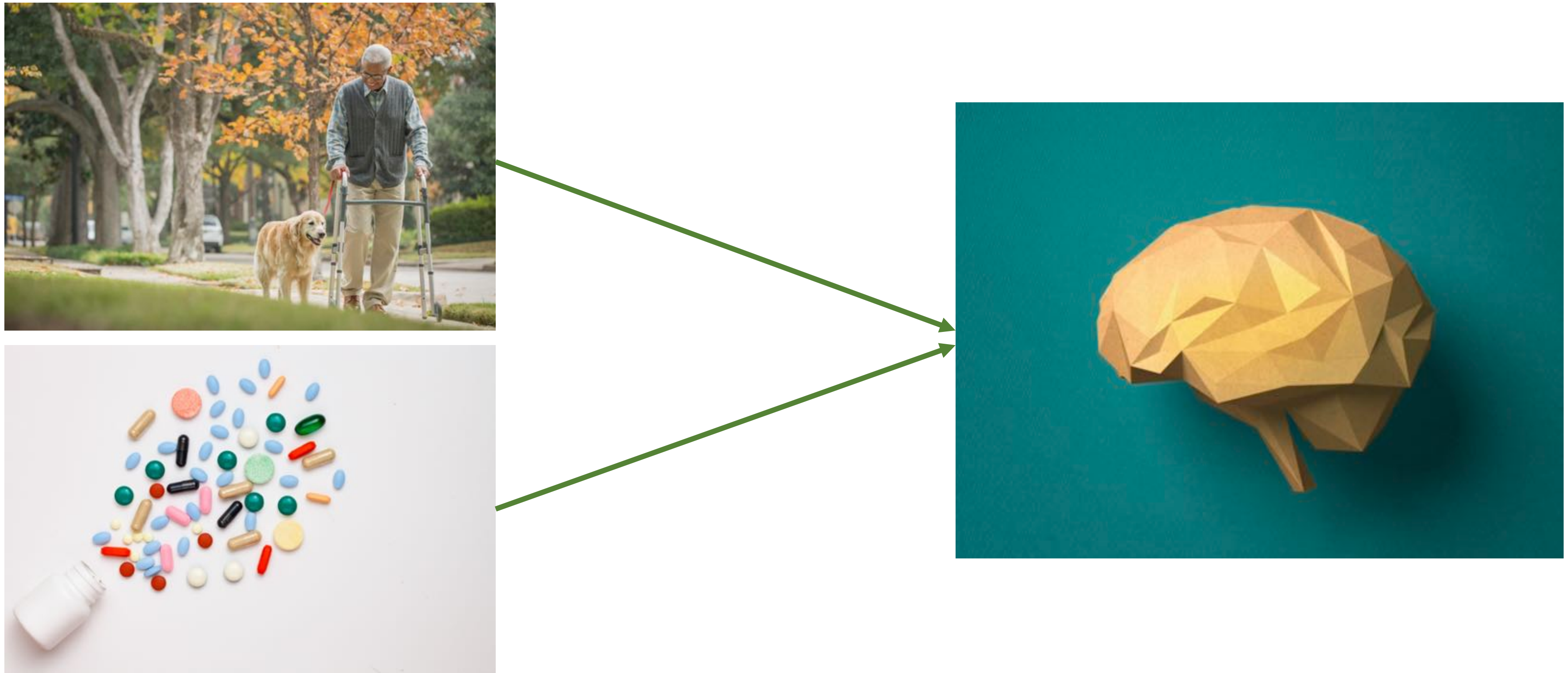
When we take many many samples and calculate confidence interval for each, 95% of the times the true mean laughing time is within the interval

McElreath, R. (2018).; Cumming, G. (2013). Intro Statistics 9 Dance of the p Values.; Holopainen, M. & Pulkkinen P. (2012).; Data: mock-up!

Contents

Usual interest

- How some variable is related to another



McElreath, R. (2018).; PowerPoint image bank.

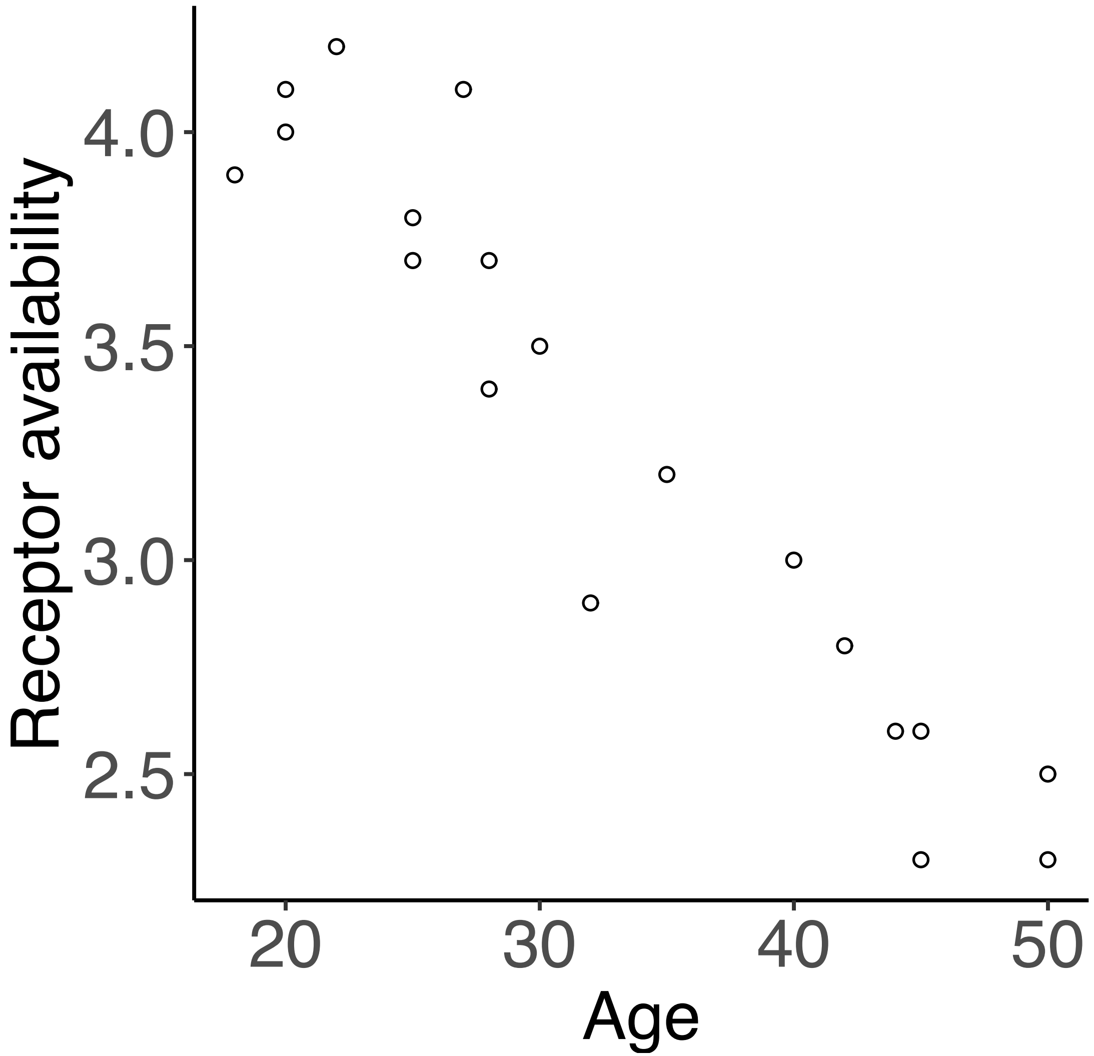
Example

How is **age** associated with (dopamine) **receptor availability** in a region of interest (ROI)?

Receptor availability (binding potential = bp)

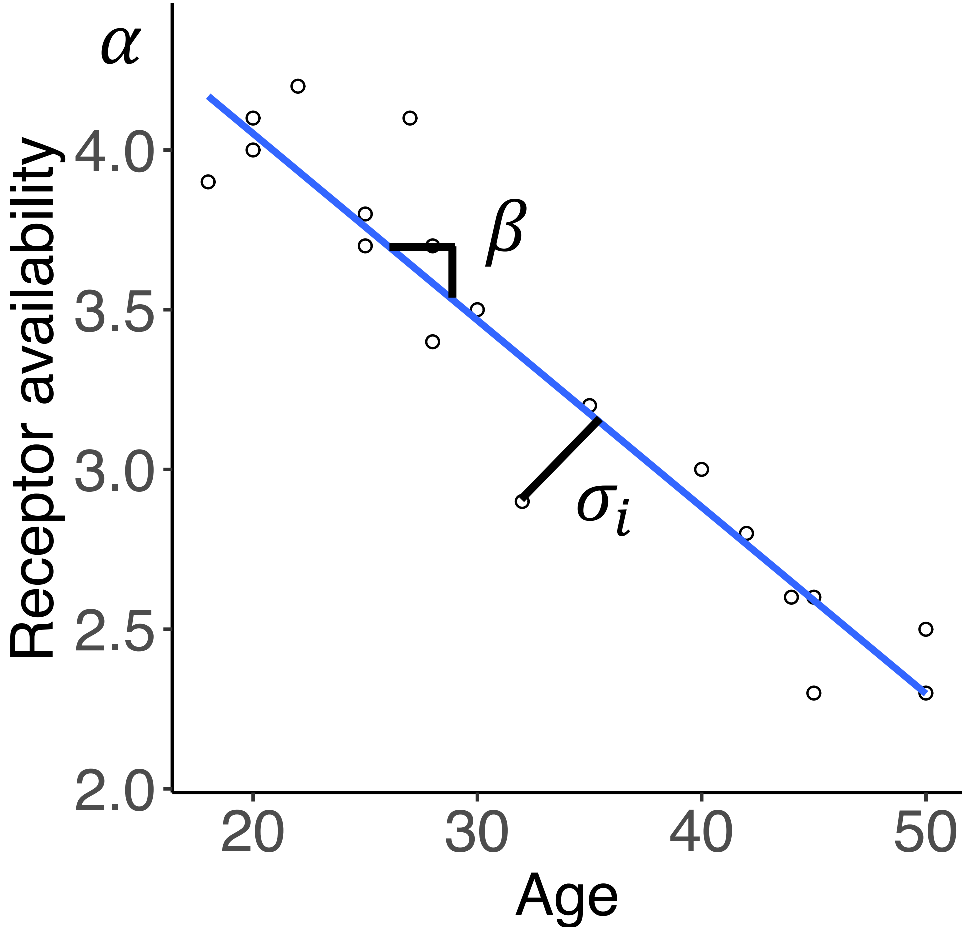
- Always positive value acquired from a PET image (1 estimate/ ROI for each subject)
- The density (and affinity) of available (unoccupied) dopamine receptors in a specific region → information about dopamine function
- $\text{Normal}(\mu, \sigma)$

Innis et al. (2007). Consensus nomenclature for in vivo imaging of reversibly binding radioligands. Journal of Cerebral Blood Flow & Metabolism, 27(9), 1533-1539.



- n= 20
- Age
- Receptor availability (bp) from PET
- 1 brain region

Linear regression: linear relation

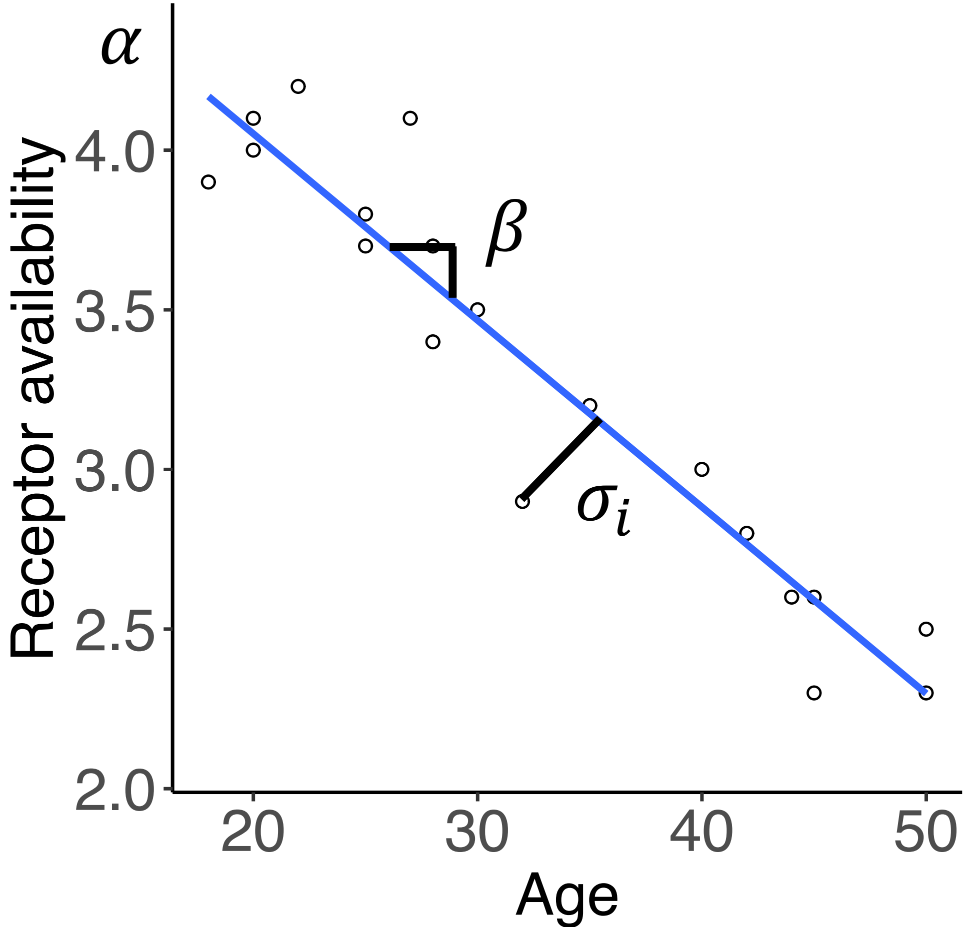


Parameters:
α = expected availability, when age = 0

β = regression coefficient (slope), steepness of the line, the change in availability when age increases 1 unit

σ_i = distance between an observation and regression line

Linear regression: linear relation



Parameters:
α = expected availability, when age = 0

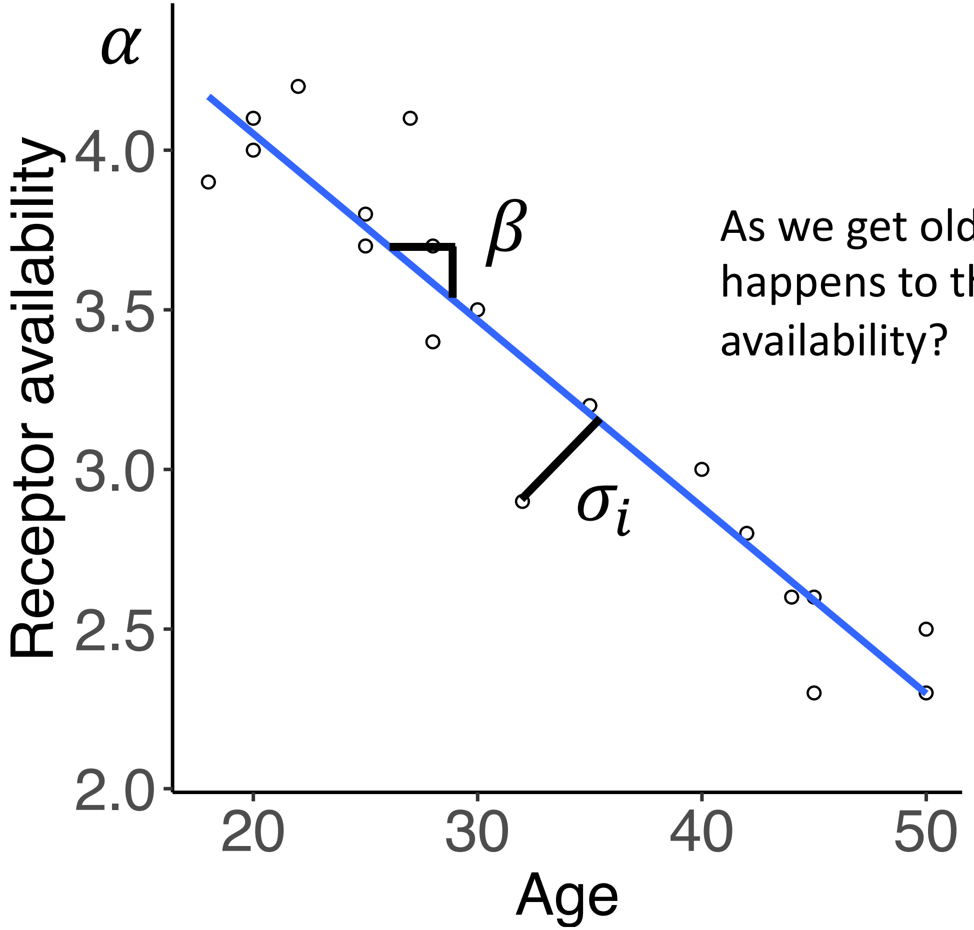
β = regression coefficient (slope), steepness of the line, the change in availability when age increases 1 unit

σ_i = distance between an observation and regression line

Linear model:
$\mu = \alpha + \beta \times \text{age}$

McElreath, R. (2018).; R Core Team (2020).; Data: mock-up!

Linear regression: linear relation



As we get older, what happens to the receptor availability?

Parameters:
α = expected availability, when age = 0

β = regression coefficient (slope), steepness of the line, the change in availability when age increases 1 unit

σ_i = distance between an observation and regression line

Linear model:
$\mu = \alpha + \beta \times \text{age}$

Let’s solve β the Bayesian way

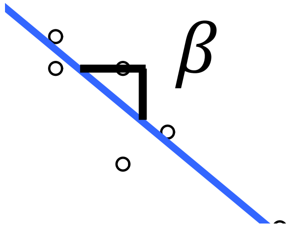
(= what happens to the receptor availability when we age)

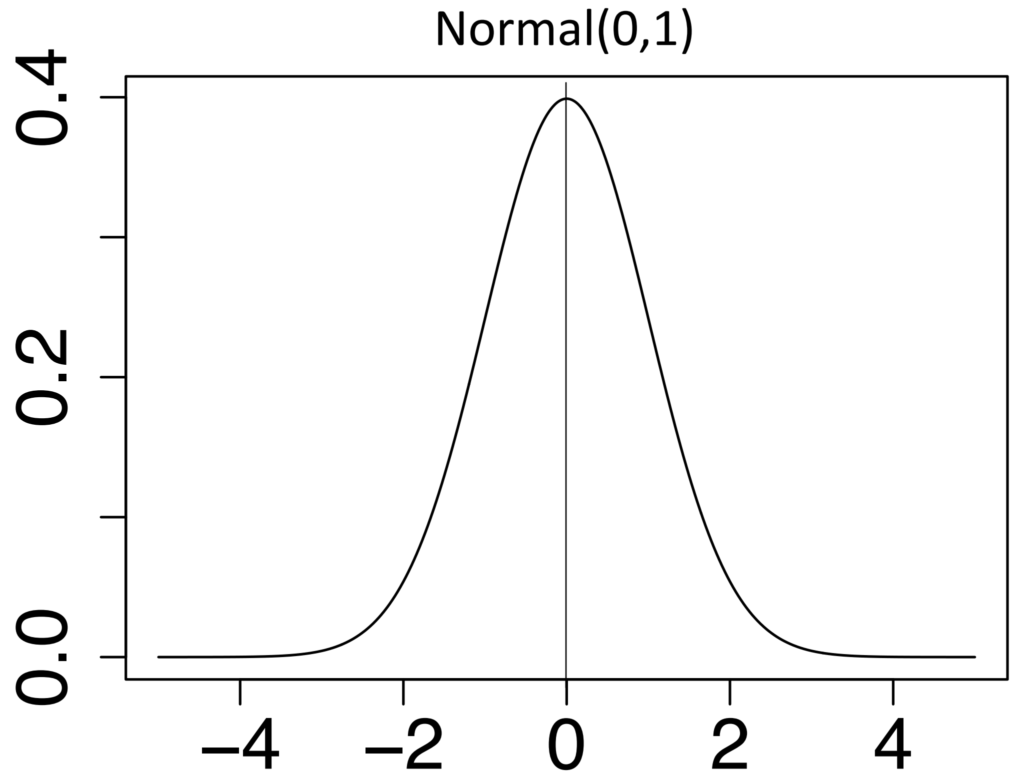
Tools

- brms-package in R

R Core Team (2020). Buerkner P-C, Buerkner MP-C (2016): Package ‘brms’.; Bürkner, P. C. (2017). brms: An R package for Bayesian multilevel models using Stan. Journal of statistical software, 80(1), 1-28.

Prior for each parameter



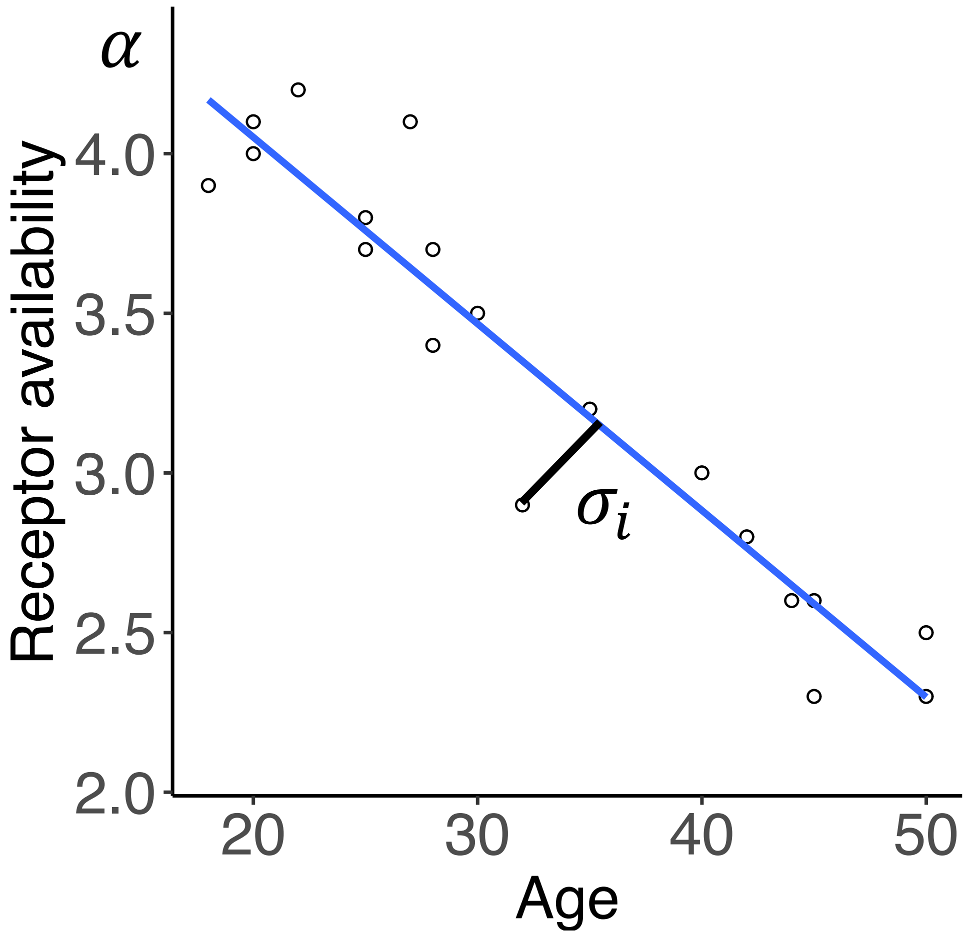


β = regression coefficient (slope)

- Mean is at ”zero-effect” (no association)
- The extremely high associations (steep line) are suggested less probable than weaker associations: does not push toward strong associations → if one found the evidence from data
- No values for β are excluded (no 0 probability given)
- Symmetrical: neither positive nor negative slope is given higher probability than the other

McElreath, R. (2018).; Buerkner P-C, Buerkner MP-C (2016).; Bürkner, P. C. (2017).; R Package Documentation. set_prior: Prior Definitions for 'brms' Models in brms: Bayesian Regression Models using 'Stan'; R Core Team (2020).

Prior for each parameter



Also α and σ are parameters of the model: need priors

For now brms package default priors (weakly informative)

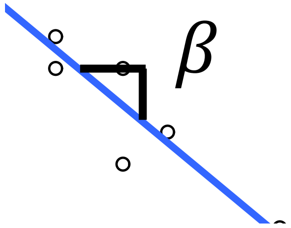
Good to be aware of all priors used and their fit in your data and population

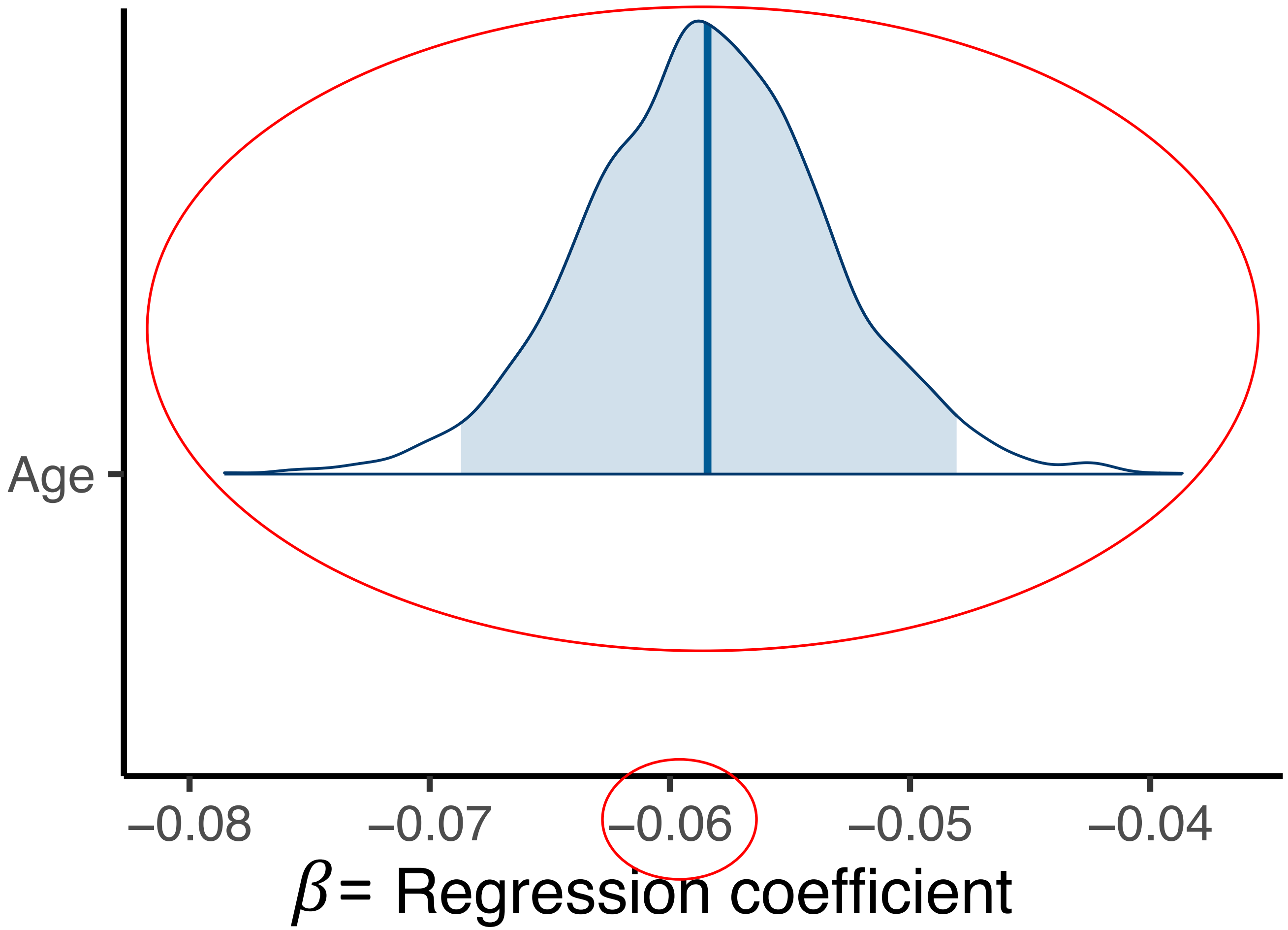
McElreath, R. (2018).; Buerkner P-C, Buerkner MP-C (2016); R Core Team (2020).

```
#### REQUIRED PACKAGES
library(rstan)
library(brms)

### STATISTICAL MODELING

prior <- c(set_prior("normal(0,1)", class = "b")) # prior for beta

model <- bp ~ age    # bp = receptor availability

fit <- brm(formula = model,
           data = data, # mock-up
           family = gaussian(), # receptor availability normally distributed
           prior = prior,
           warmup = 1000, iter = 2000, chains = 4, # sampling settings, see brms
           control = list(adapt_delta = 0.95)) # sampling settings, see brms
```

McElreath, R. (2018).; R Core Team (2020).; Buerkner P-C, Buerkner MP-C (2016).; Bürkner, P. C. (2017).; R Package Documentation. set_prior: Prior Definitions for 'brms' Models in brms: Bayesian Regression Models using 'Stan'; Bürkner, P. (2021). Rdocumentation. brms.

Result for β





Ageing 1 year,
The availability decreases 0.06 units
=
Ageing 10 years
The availability decreases 0.6 units

McElreath, R. (2018).; Bürkner, P. C. (2017).; R Core Team (2020).; Gabry J, Mahr T (2021). "bayesplot: Plotting for Bayesian Models." R package version 1.8.0.

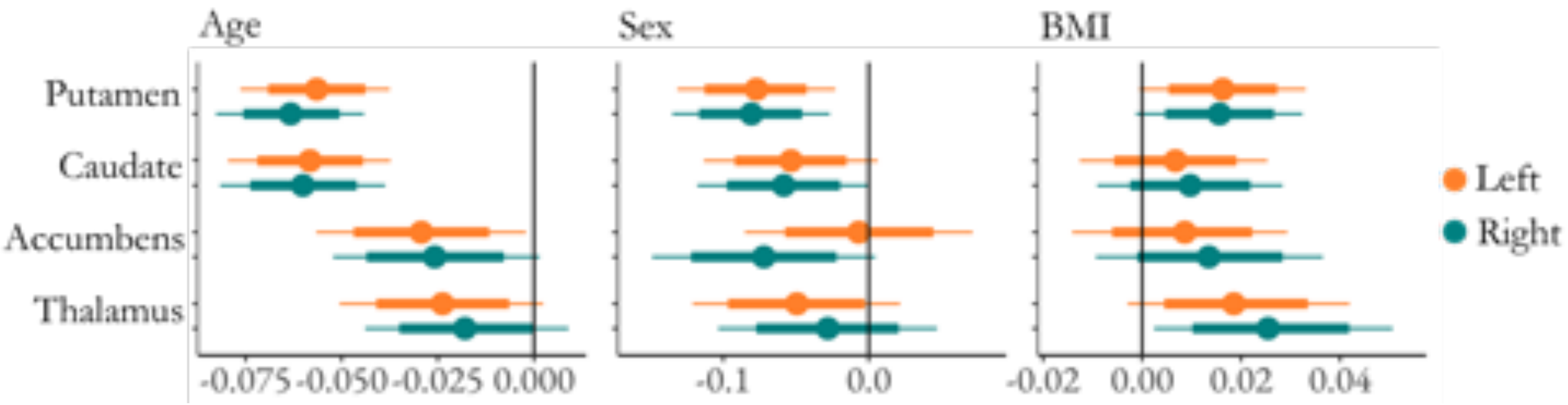
When?

- Simple enough setting
  - Computational methods (sampling from the posterior distribution)
  - Regional rather than voxel-level
  - Some predictors (e.g. age, sex…) but not too many

When?

Age, sex, BMI → dopamine receptor availability (preprint)

- Ageing 10 years and 5% decrease in the availability (putamen and caudate)
- Females had higher availability than males
- BMI: weak positive association with the availability
- Please see the preprint for more information about the scales to use (standardizing, log-transformation)



Bayesian methods not limited to linear regression

Malén T, Karjalainen T, Isojärvi J, Vehtari A, Bürkner P-C, Putkinen V, et al. (2021): Age and sex dependent variability of type 2 dopamine receptors in the human brain: A large-scale PET cohort. *bioRxiv*.2021.2008.2010.455776.; R Core Team (2020.; Gabry J, Mahr T (2021).; McElreath, R. (2018).

Why?

Interpretation intuitive

- Probabilities for parameter values (what we want)
  - Instead of having probability for the data in the light of only 1 assumed parameter value (null-hypothesis)
- One sample enough -> imaginary resampling not needed

To take home

- Prior and data form posterior

- If interested in how your prior affects the result, you can try different ones and compare your posteriors

- We have information about the world before we see the data and we can use it (as long as we accept the subjective probability)!

Further Learning Material

- Bayesian Data Analysis (Gelman, Carlin, Stern Rubin), Chapman Hall, 1995, 2003, 2013
  - http://www.stat.columbia.edu/~gelman/books

- McElreath, R. (2018). *Statistical rethinking: A Bayesian course with examples in R and Stan*. Chapman and Hall/CRC.
  - https://www.youtube.com/watch?v=4WVelCswXo4
  - Scripts and datasets

Acknowledgements

- **Tomi Karjalainen**, Data scientist, Polar Electro Oy, Turku PET Centre
- **Kari Auranen**, Professor, Department of Mathematics and Statistics, Department of Clinical Medicine, University of Turku
- **Pekka Nieminen**, University lecturer, Department of Mathematics and Statistics, University of Turku
- **Paul Bürkner**, Statistician, Junior Research Group Leader, Cluster of Excellence SimTech, University of Stuttgart
- **Aki Vehtari**, Associate professor, Aalto University, Helsinki
- **Vesa Putkinen**, Postdoctoral researcher, University of Turku
- **Birgitta Paranko**, Project researcher, University of Turku
- **Severi Santavirta**, MD, Doctoral candidate, University of Turku

References

- Thomas D. Gauthier, Mark E. Hawley (2015). Introduction to Environmental Forensics (Third Edition).
- McElreath, R. (2018). Statistical rethinking: A Bayesian course with examples in R and Stan. Chapman and Hall/CRC.
- Holopainen, M. & Pulkkinen P. (2012). *Tilastolliset menetelmät*. Sanoma Pro Oy.
- Gelman, A., Carlin, J. B., Stern, H. S., & Rubin, D. B. (1995-2020). Bayesian data analysis. Chapman and Hall/CRC.
- Gower, B. (1997). Scientific method. Routledge.
- Chick, S. E. (2006). Subjective probability and Bayesian methodology. *Handbooks in Operations Research and Management Science*, *13*, 225-257.
- Cumming, G. (2013). Intro Statistics 9 Dance of the p Values. Video. Retrieved from: https://www.youtube.com/watch?v=5OL1RqHrZQ8. Cited: 13.8.2021.
- Innis, R. B., Cunningham, V. J., Delforge, J., Fujita, M., Gjedde, A., Gunn, R. N., ... & Carson, R. E. (2007). Consensus nomenclature for in vivo imaging of reversibly binding radioligands. Journal of Cerebral Blood Flow & Metabolism, 27(9), 1533-1539.

R

- R Core Team (2020). R: A language and environment for statistical computing.
- Buerkner P-C, Buerkner MP-C (2016): Package ‘brms’.
- Bürkner, P. C. (2017). brms: An R package for Bayesian multilevel models using Stan. Journal of statistical software, 80(1), 1-28.
- R Package Documentation. set_prior: Prior Definitions for 'brms' Models in brms: Bayesian Regression Models using 'Stan’. Retrieved from: https://rdrr.io/cran/brms/man/set_prior.html. Cited: 16.8.2021.
- Bürkner, P. (2021). Rdocumentation. brms. Retrieved from: https://www.rdocumentation.org/packages/brms/versions/2.15.0. Cited: 17.8.2021.
- Gabry J, Mahr T (2021). “bayesplot: Plotting for Bayesian Models.” R package version 1.8.0, <URL:https://mc-stan.org/bayesplot/>.

Images

- PowerPoint image bank