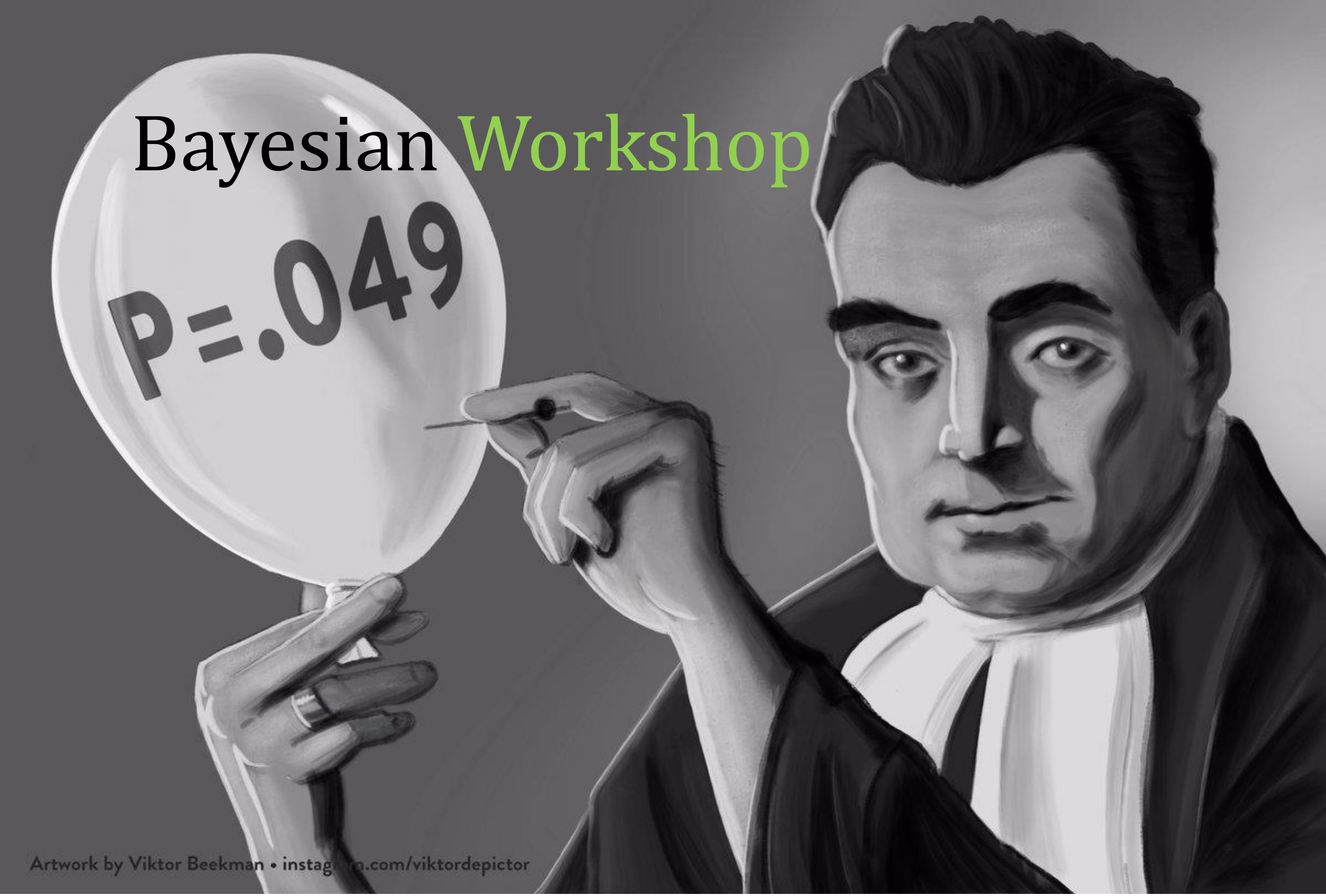
Bayesian Workshop



5th meeting[01.06.2018] **Refresher session!**

Today's Game plan

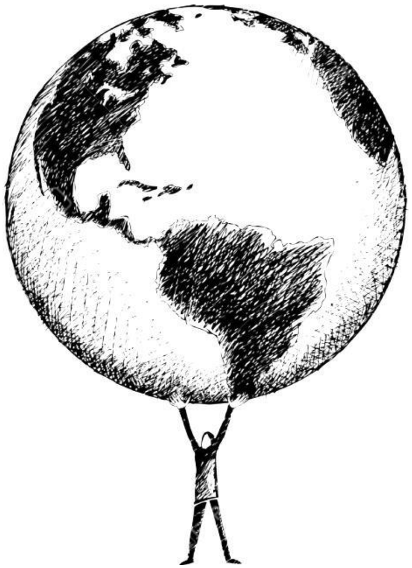
- Intro to **Bayesian Inference** reading
  McElreath (2015): "Statistical Rethinking"
  [Chap 2, up to & including section 2.3.4, p.19 – 37]

- Video lectures
  Daniël Lakens: "Improving your statistical inferences"
  Lecture 2.1 [16mins] **Likelihoods**
  Lecture 2.2 [14mins] **Binomial Bayesian Inference**
  Lecture 2.3 [11mins] **Bayesian Thinking**

- Announcement of **upcoming events!**
  OSF homepage: https://osf.io/hcm7p/wiki/home/

Small Worlds & Large Worlds

2 frames of statistical modeling



SMALL WORLD

model

self-contained
logical consistency

LARGE WORLD

reality

broader context in which
one deploys a model

model assumptions ***may or may not*** approximate reality

The garden of forking data

Jorge Luis Borges

The garden of forking data

- In order to make good inference about what actually happened, it helps to consider everything that could have happened.

- Bayesian analysis → garden of forking data
  → alternative sequences of events

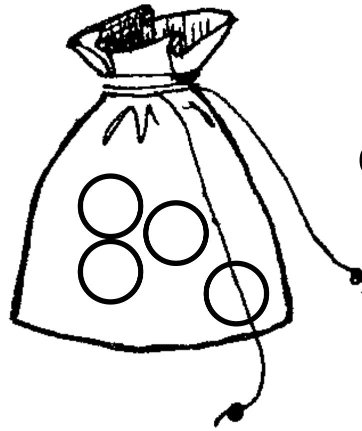
- As we learn what did happen, some of these alternative sequences are pruned.

→ In the end, what remains is only what is logically consistent with our knowledge.

Counting possibilities

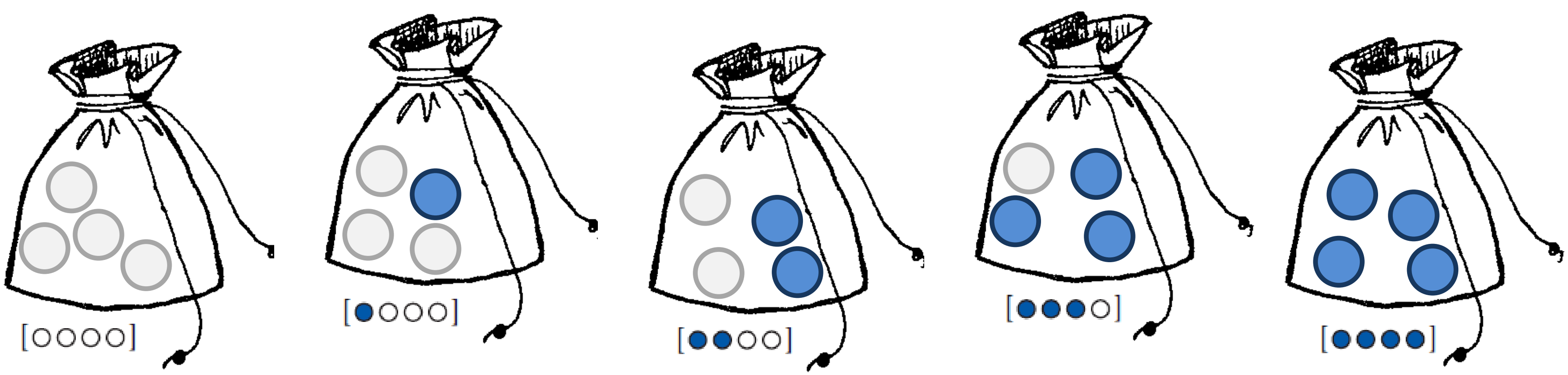
- Probability theory
- Marble example

Given:



either blue white

- All possibilities consistent with what we know about the bag
  → 5 *conjectures*

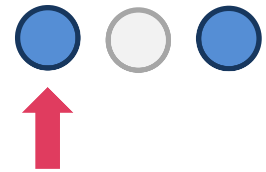


GOAL: Determine which conjecture is the most plausible, given some **evidence**.

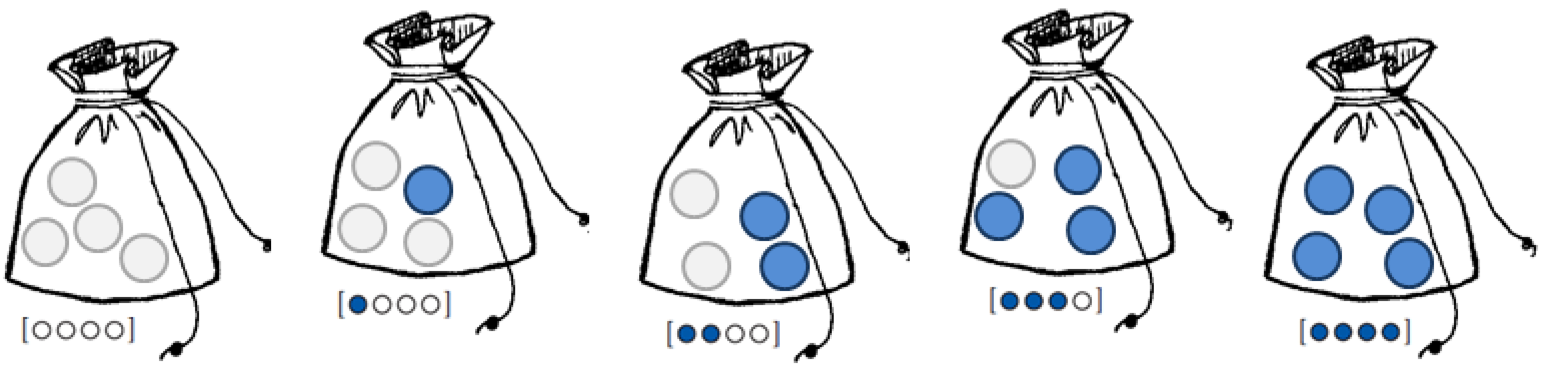
Counting possibilities

Evidence (i.e. data)**:**

Draw 3 marbles from the bag (with replacement)



1st draw

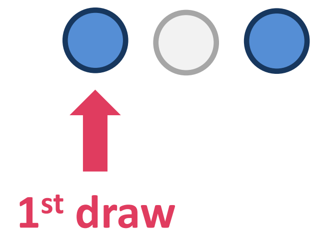


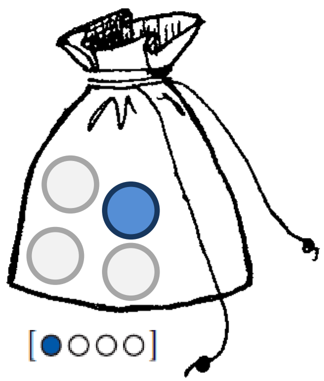
GOAL: Determine which conjecture is the most plausible, given some **evidence.**

Counting possibilities

Evidence (i.e. data):

Draw 3 marbles from the bag (with replacement)

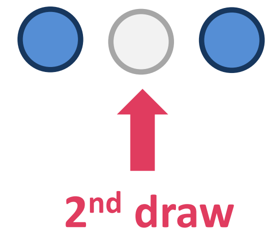


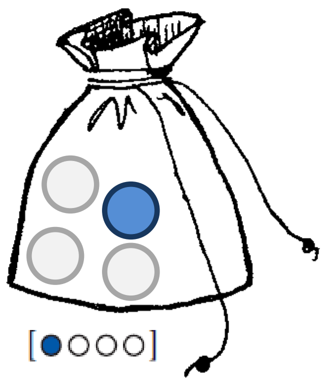


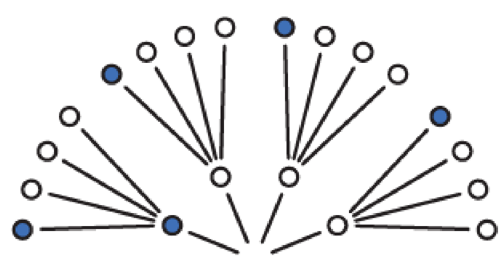
Counting possibilities

Evidence (i.e. data):

Draw 3 marbles from the bag (with replacement)



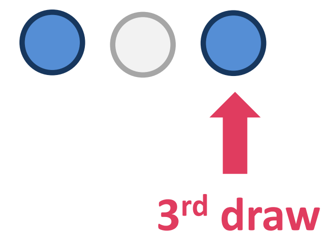


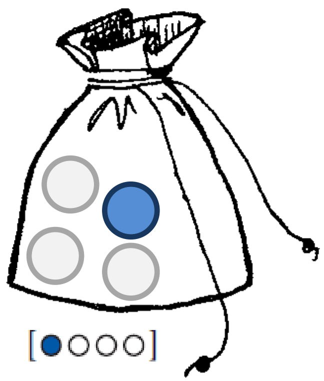


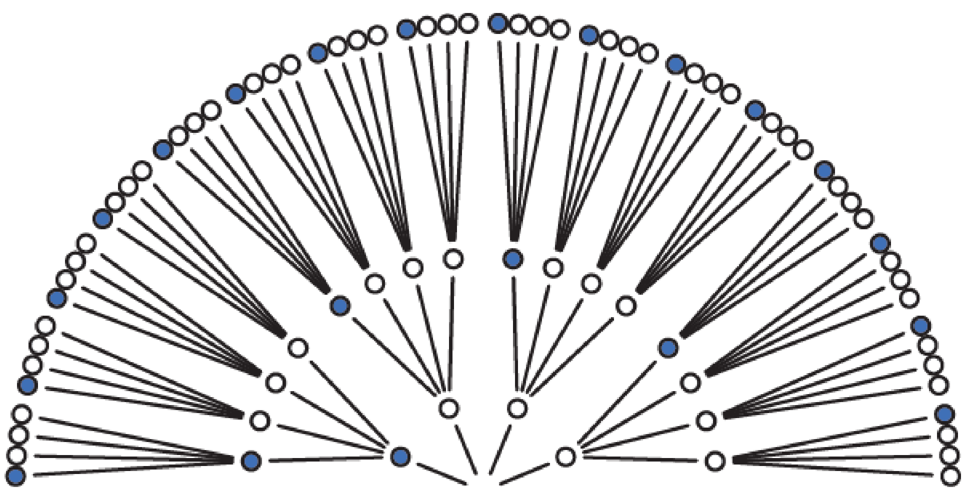
Counting possibilities

Evidence (i.e. data):

Draw 3 marbles from the bag (with replacement)



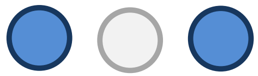


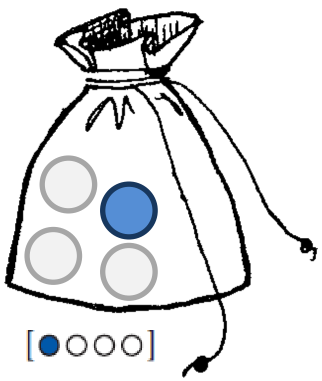


Counting possibilities

Evidence (i.e. data):

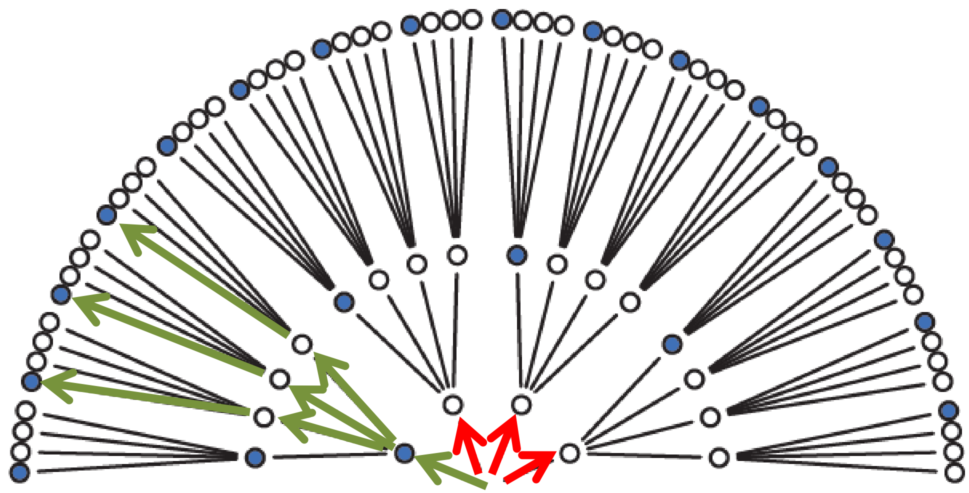
Draw 3 marbles from the bag (with replacement)





64 possible paths:

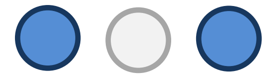
- Some are **logically consistent** with our data
- Some can be **eliminated**

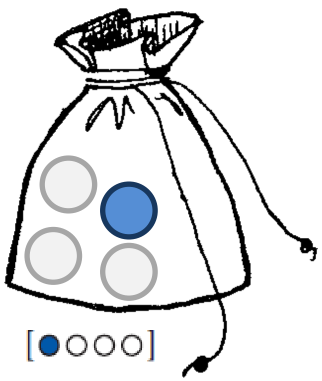


Counting possibilities

Evidence (i.e. data):

Draw 3 marbles from the bag (with replacement)

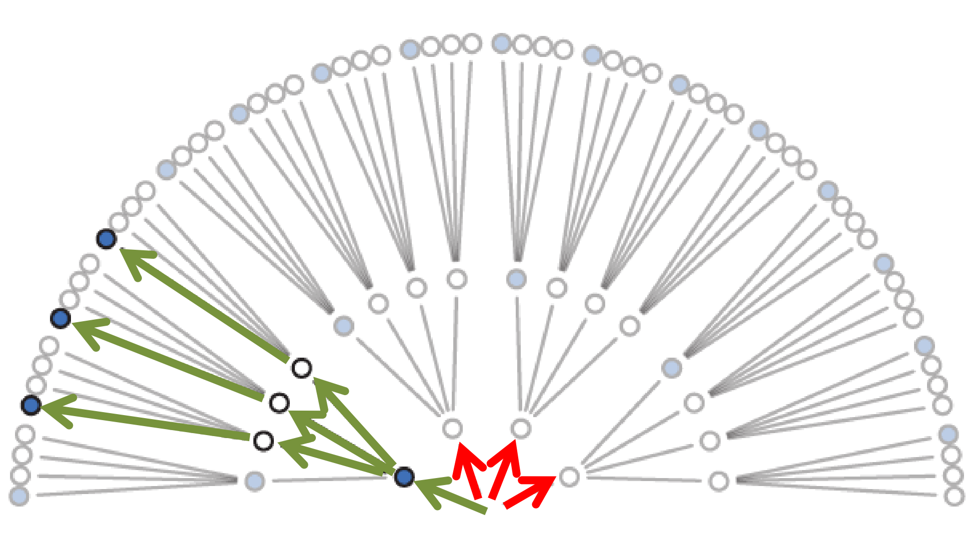




64 possible paths:

- Some are **logically consistent** with our data
- Some can be **eliminated**

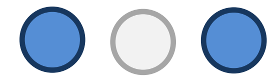
3 paths remain

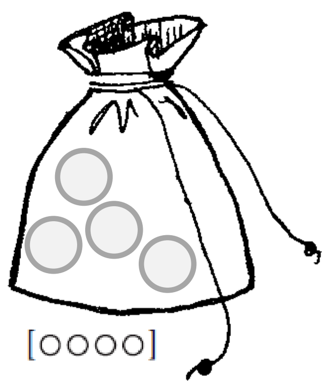


Counting possibilities

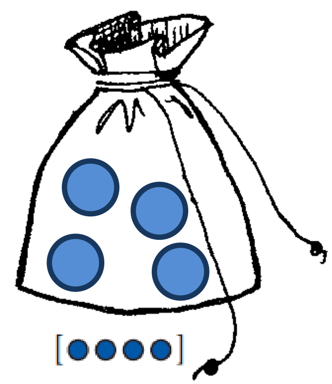
Evidence (i.e. data):

Draw 3 marbles from the bag (with replacement)





0 paths

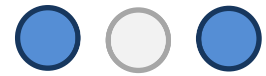


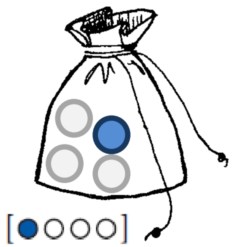
0 paths

Counting possibilities

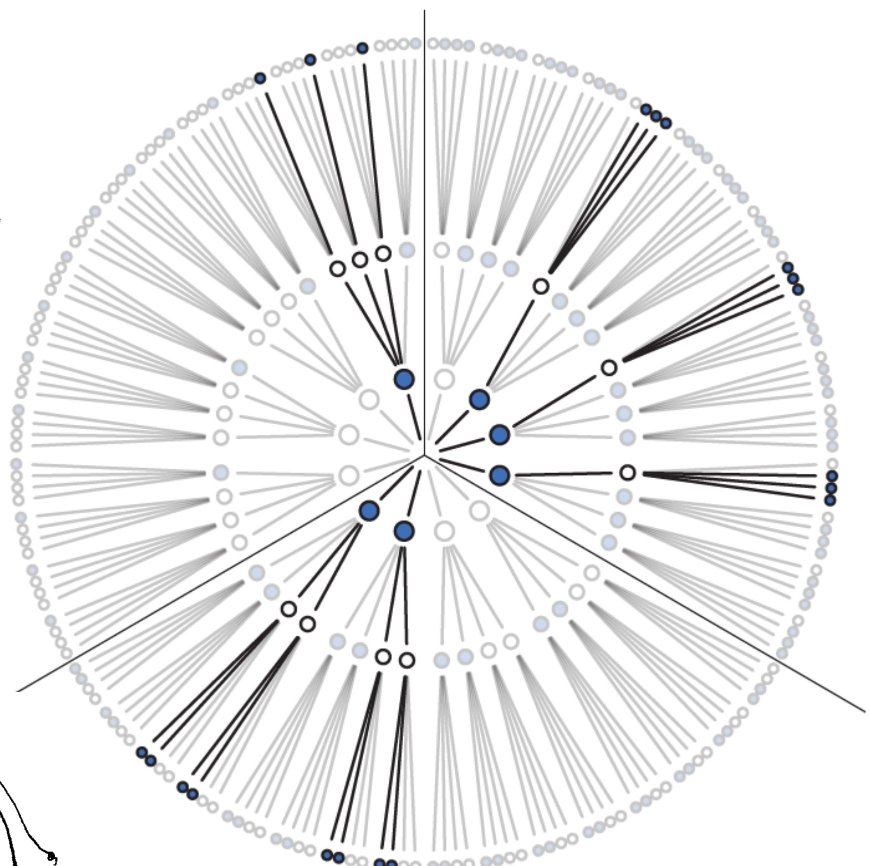
Evidence (i.e. data):

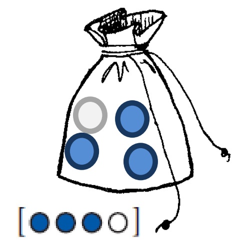
Draw 3 marbles from the bag (with replacement)



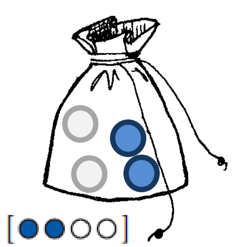


3 paths





9 paths



8 paths

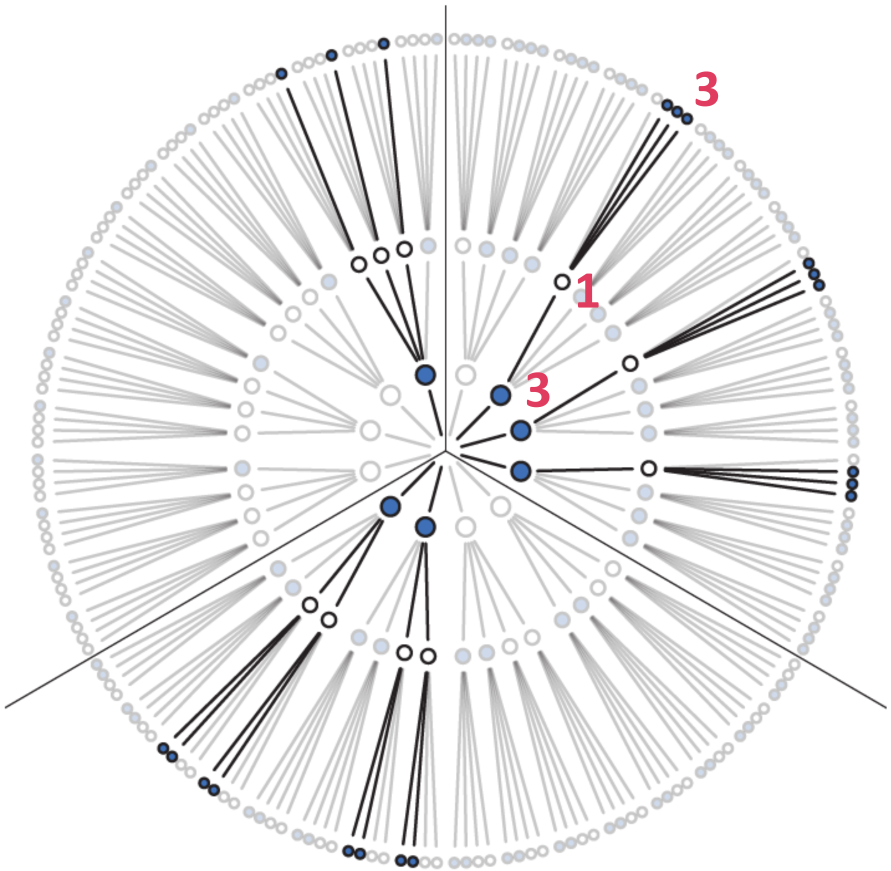
Conjecture	Ways to produce ●○●
[○○○○]	$0 \times 4 \times 0 = 0$
[●○○○]	$1 \times 3 \times 1 = 3$
[●●○○]	$2 \times 2 \times 2 = 8$
[●●●○]	$3 \times 1 \times 3 = 9$
[●●●●]	$4 \times 0 \times 4 = 0$

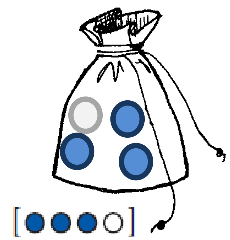
Counting possibilities

Evidence (i.e. data):

Draw 3 marbles from the bag (with replacement)

● ○ ●





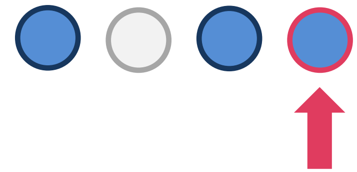
9 paths

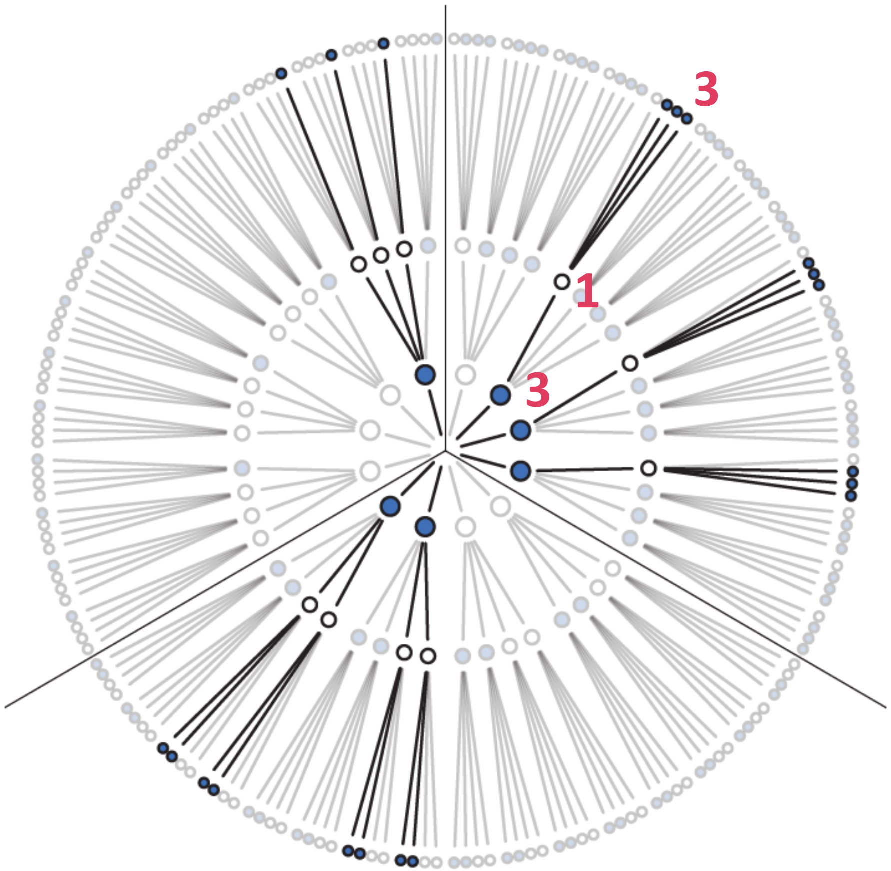
Conjecture	Ways to produce ●○●
[○○○○]	$0 \times 4 \times 0 = 0$
[●○○○]	$1 \times 3 \times 1 = 3$
[●●○○]	$2 \times 2 \times 2 = 8$
[●●●○]	$3 \times 1 \times 3 = 9$
[●●●●]	$4 \times 0 \times 4 = 0$

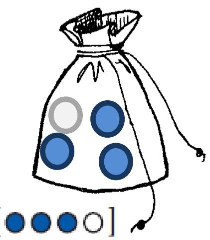
Counting possibilities

Evidence (i.e. data):

Draw 3 marbles from the bag (with replacement)







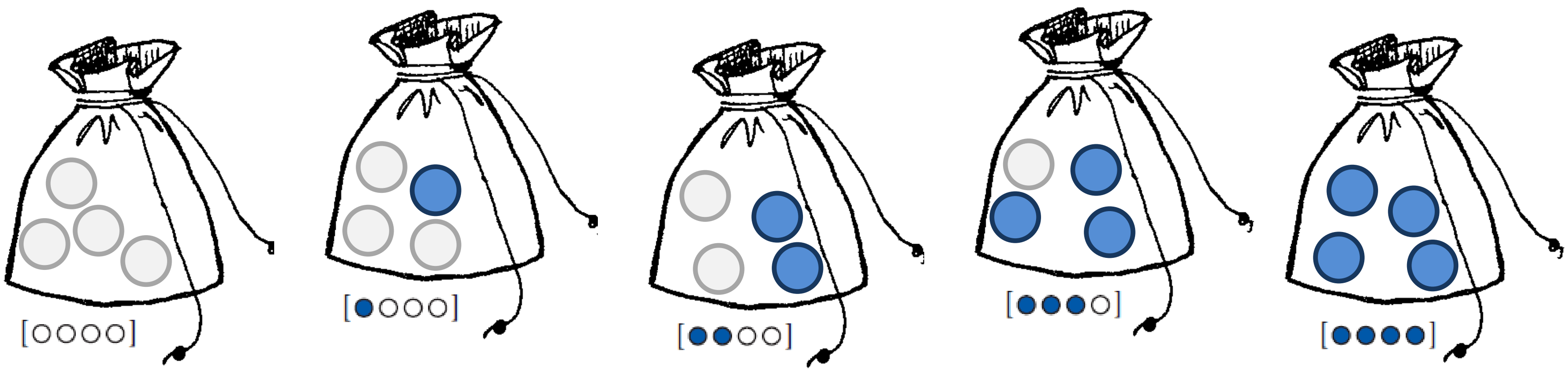
9 paths

→ update by multiplying new count by old count

Conjecture	Ways to produce ●○●
[○○○○]	$0 \times 4 \times 0 = 0$
[●○○○]	$1 \times 3 \times 1 = 3$
[●●○○]	$2 \times 2 \times 2 = 8$
[●●●○]	$3 \times 1 \times 3 = 9$ x 3
[●●●●]	$4 \times 0 \times 4 = 0$

Counting possibilities

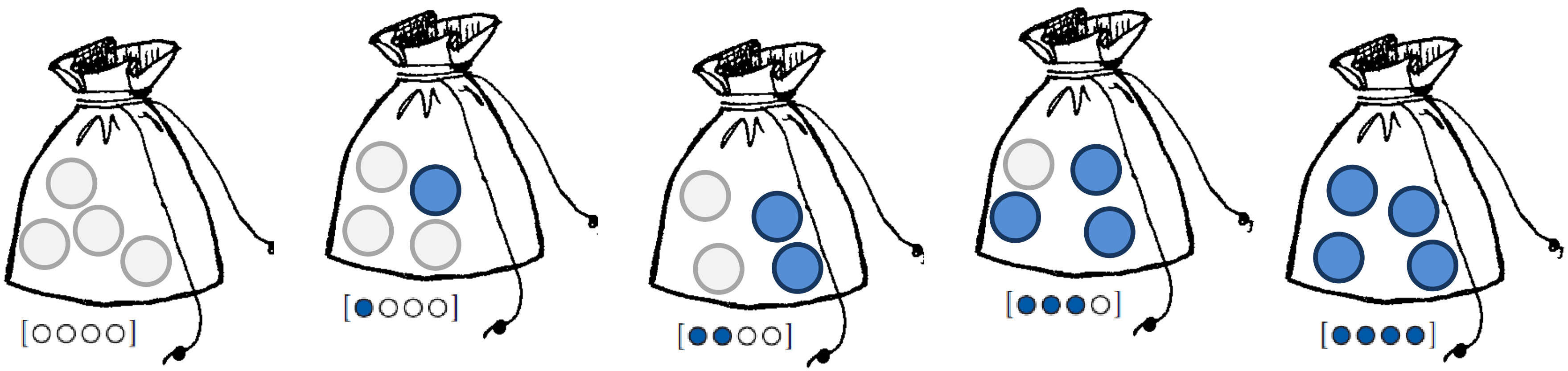
Principle of indifference: When there is no reason to say that one conjecture is more plausible than another, weigh all of the conjectures equally.



1 1 1 1 1

Counting possibilities

Prior information: Each bag contains at least 1 blue & 1 white marble.
Blue marbles are rare.



0 3 2 1 0

Conjecture	Prior count	Factory count	New count
[○○○○]	0	0	$0 \times 0 = 0$
[●○○○]	3	3	$3 \times 3 = 9$
[●●○○]	16	2	$16 \times 2 = 32$
[●●●○]	27	1	$27 \times 1 = 27$
[●●●●]	0	0	$0 \times 0 = 0$

Relative plausibility of each conjecture in terms of **raw counts.**

Counting possibilities

Conjecture	Prior count	Factory count	New count
[○○○○]	0	0	$0 \times 0 = 0$
[●○○○]	3	3	$3 \times 3 = 9$
[●●○○]	16	2	$16 \times 2 = 32$
[●●●○]	27	1	$27 \times 1 = 27$
[●●●●]	0	0	$0 \times 0 = 0$

Relative plausibility of each conjecture in terms of **raw counts.**

Can compute these plausibilities as ***proportions*** *(standardized).*

$$\text{plausibility of } p \text{ after } D_{\text{new}} = \frac{\overset{16}{\text{ways } p \text{ can produce } D_{\text{new}}} \overset{\mathbf{x}}{\times} \overset{2}{\text{prior plausibility } p}}{\underset{(9+32+27)}{\text{sum of products}}}$$

[●○○○] **9/68 ~ 0.13**

[●●○○] **32/68 ~ 0.47**

[●●●○] **27/68 ~ 0.40**

SUM to 1.

Bayesian Updating

Probability theory terminology:

PARAMETERS: Represent the different *conjectures* for causes or explanations of the data.

LIKELIHOOD x **PRIOR PROBABILITY**

POSTERIOR PROBABILITY

$$\text{plausibility of } p \text{ after } D_{\text{new}} = \frac{\text{ways } p \text{ can produce } D_{\text{new}} \times \text{prior plausibility } p}{\text{sum of products}}$$

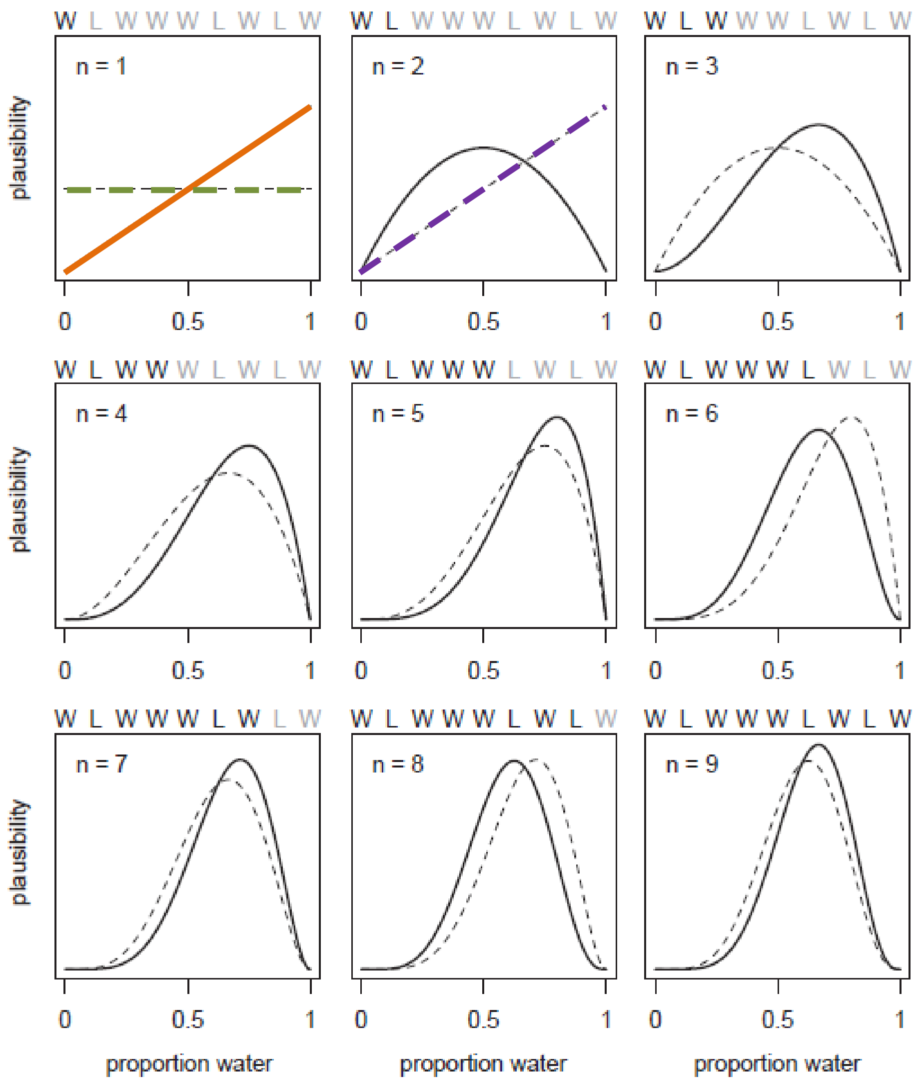
BAYESIAN UPDATING:

- Bayesian model begins with one set of plausibilities (*prior probabilities*)
- Model updates in light of new data (*likelihood*)
- This produces *posterior probabilities*

→ Every updated set of plausibilities (*posteriors*) becomes the initial plausibilities (*priors*) for the next observation

Bayesian Updating

How a Bayesian model learns



uniform (uninformative) prior

- Each toss of the globe produces an observation of water (W) or land (L).
- With each new observation, the plausibility (i.e. estimate of the proportion of water) is updated.
- **Priors:** dashed curves
- **Posteriors:** solid curves

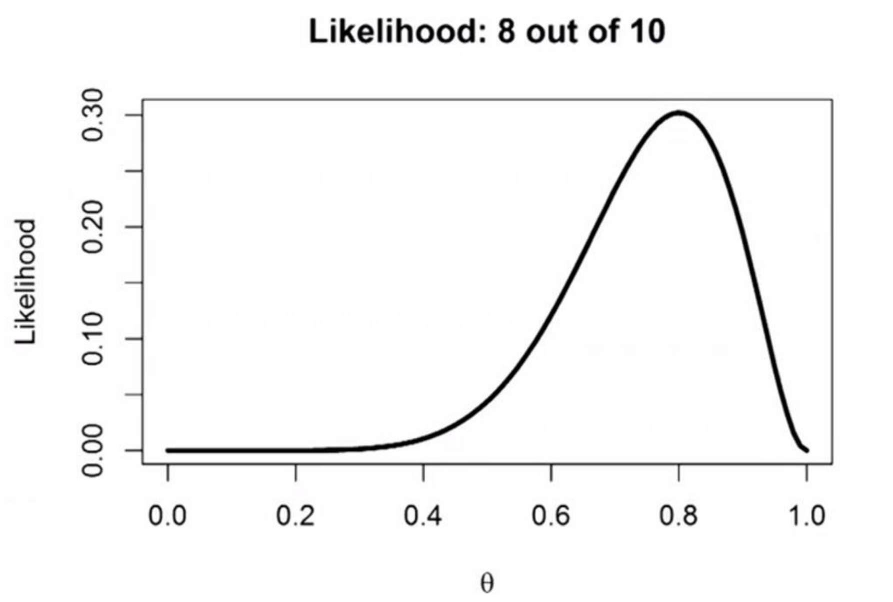
Bayesian Statistics

Key concepts

LIKELIHOOD: A likelihood gives you the function of a parameter given the data.

"So when you've observed some data, you can plot the accompanying likelihood function, and you can check how likely each hypothesis you may have is."

You flip 8 out of 10 heads. The likelihood of $\theta = 0.8$ is 0.30.



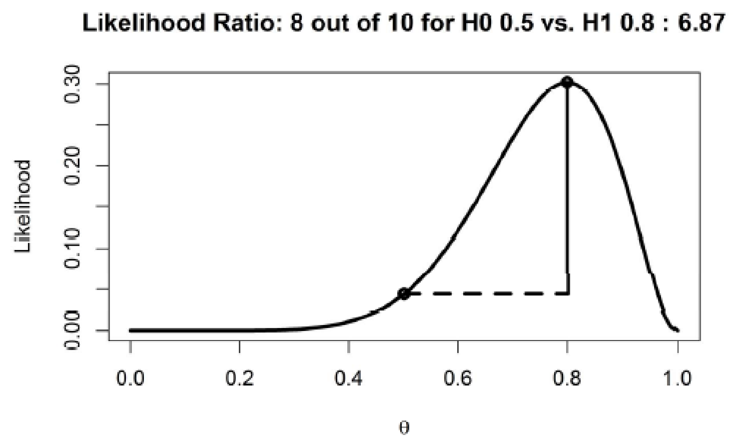
Key concepts

LIKELIHOOD RATIO: *"We can use the likelihood under the **null hypothesis** and the likelihood under the **alternative hypothesis**, and calculate the likelihood ratio."*

→ *We are taking the **relative evidence** of the one hypothesis (e.g. **H0**) and the **relative evidence** of the other hypothesis (**H1**), and calculating the **odds** of one over the other:*

$$\frac{\text{odds(H1)}}{\text{odds(H0)}} \qquad \frac{L(\text{H1})}{L(\text{H0})} \qquad \frac{L(\theta=0.8)}{L(\theta=0.5)}$$

We can use the likelihood under H0 and H1 to calculate the **likelihood ratio**



→ **Given the observed data, we have support for the hypothesis that the coin 'unfair'.**

Bayesian Statistics

Key concepts

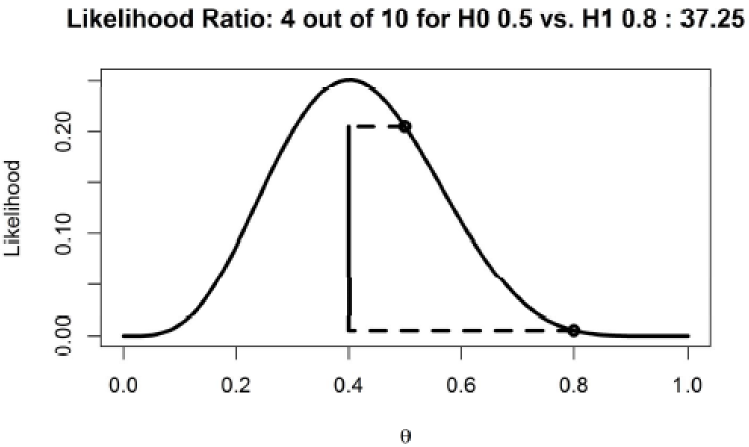
LIKELIHOOD RATIO: *"We can use the likelihood under the **null hypothesis** and the likelihood under the **alternative hypothesis**, and calculate the likelihood ratio."*

→ *We are taking the **relative evidence** of the one hypothesis (e.g. **H0**) and the **relative evidence** of the other hypothesis (**H1**), and calculating the **odds** of one over the other:*

$$\frac{\text{odds(H1)}}{\text{odds(H0)}} \qquad \frac{L(H1)}{L(H0)} \qquad \frac{L(\theta=0.8)}{L(\theta=0.5)}$$

Given a different observation: e.g., 4 out of 10 flips are HEADS

→ Now, we have support for the hypothesis that the coin is 'fair'.



LIKELIHOODS & LIKELIHOOD RATIOS

2 important notes!

Likelihood ratios of **8** and **32** are moderately strong and strong evidence.

(Royall, 1997)

Likelihoods are **relative** evidence for H1 vs. H0. H0 and H1 might be unlikely.

Bayesian Statistics

Key concepts

PRIORS & POSTERIORS (Bayesian updating):
*"You have some **prior belief**, & you have some **data**, & you combine these into a **posterior belief**."*

Prior Belief + Data = Posterior Belief

We can calculate **posterior odds** that the alternative hypothesis is true (given the data), compared to the probability that the null hypothesis is true (given the data).

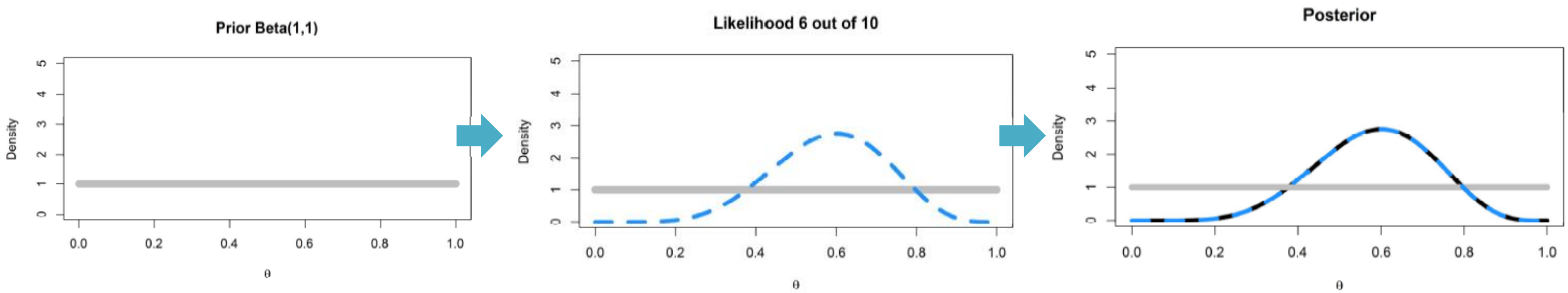
Posterior odds:

$$\frac{P(\mathrm{H1}|\mathrm{D})}{P(\mathrm{H0}|\mathrm{D})} = \frac{P(\mathrm{D}|\mathrm{H1})}{P(\mathrm{D}|\mathrm{H0})} \times \frac{P(\mathrm{H1})}{P(\mathrm{H0})}$$

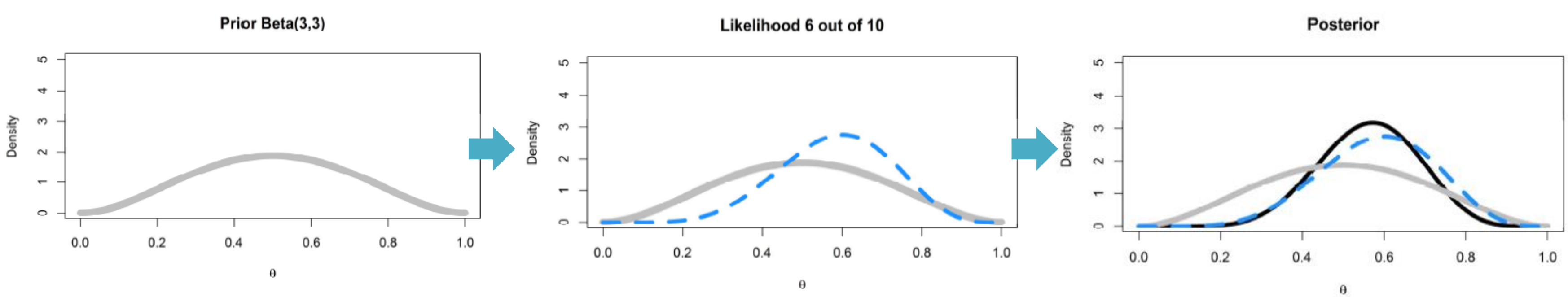
$$Posterior \ = \ Likelihood\ Ratio \ \times \ Prior$$

Bayesian Statistics
Key concepts

PRIORS & POSTERIORS (Bayesian updating): *Examples*



*with a uniform (uninformative) prior, only the likelihood influences our posterior beliefs.



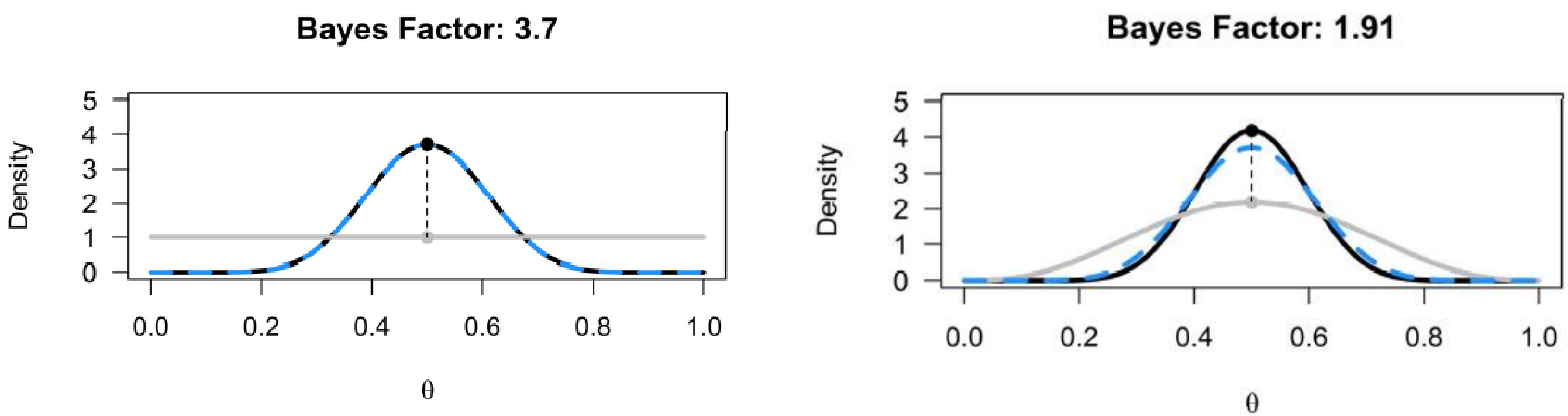
*given a non-uniform prior, the posterior is a combination of the likelihood + the prior.

Bayesian Statistics
Key concepts

BAYES FACTORS: *Relative evidence of one model compared to another.*

→ We can compare the **prior distribution** to the **posterior distribution** to see how much our beliefs have changed by collecting some data.

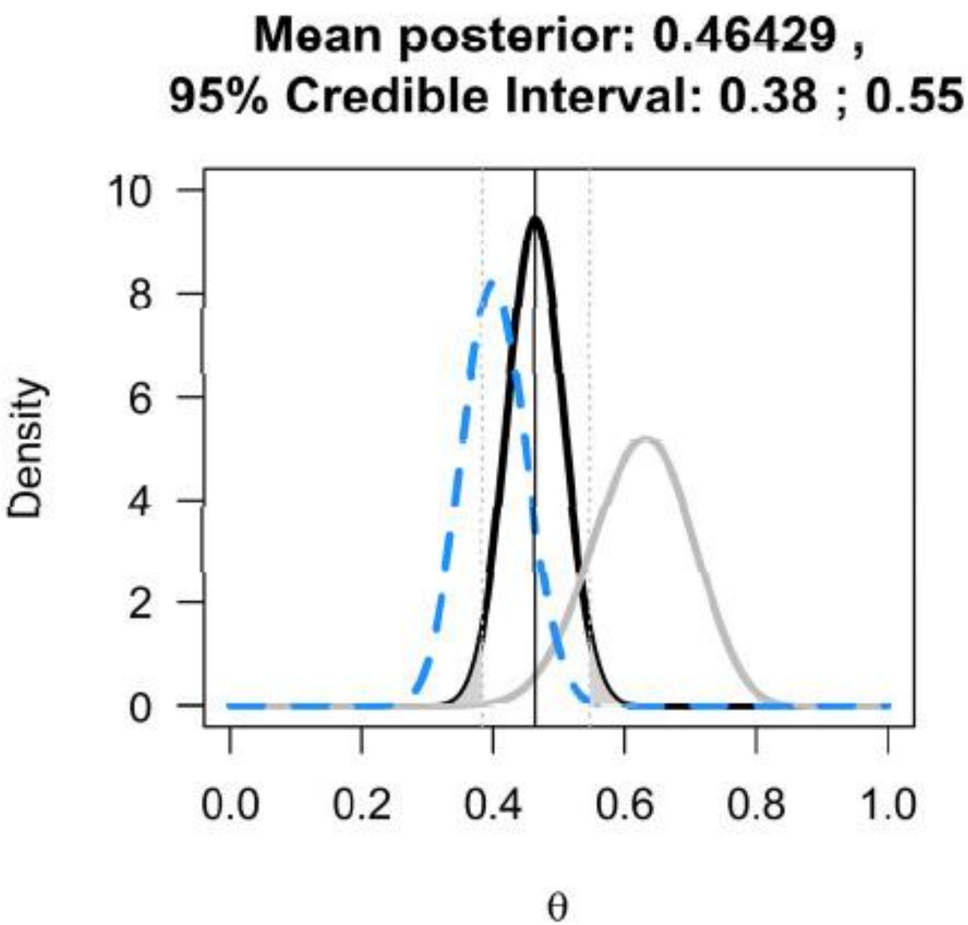
→ **Larger the BF, larger the change in belief** (from prior to posterior)



After looking at the data, θ=0.5 has become **3.70** or **1.91** times more likely, depending on the prior.

Bayesian Statistics
Key concepts

BAYES FACTORS: *Relative evidence of one model compared to another.*

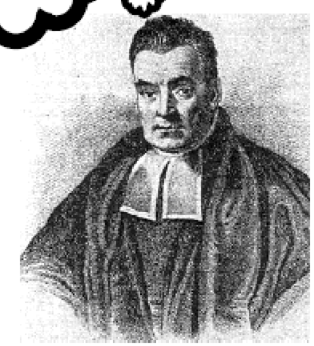


BF = 1 to 3
Inconclusive evidence

BF > 3, BF < 1/3
substantial

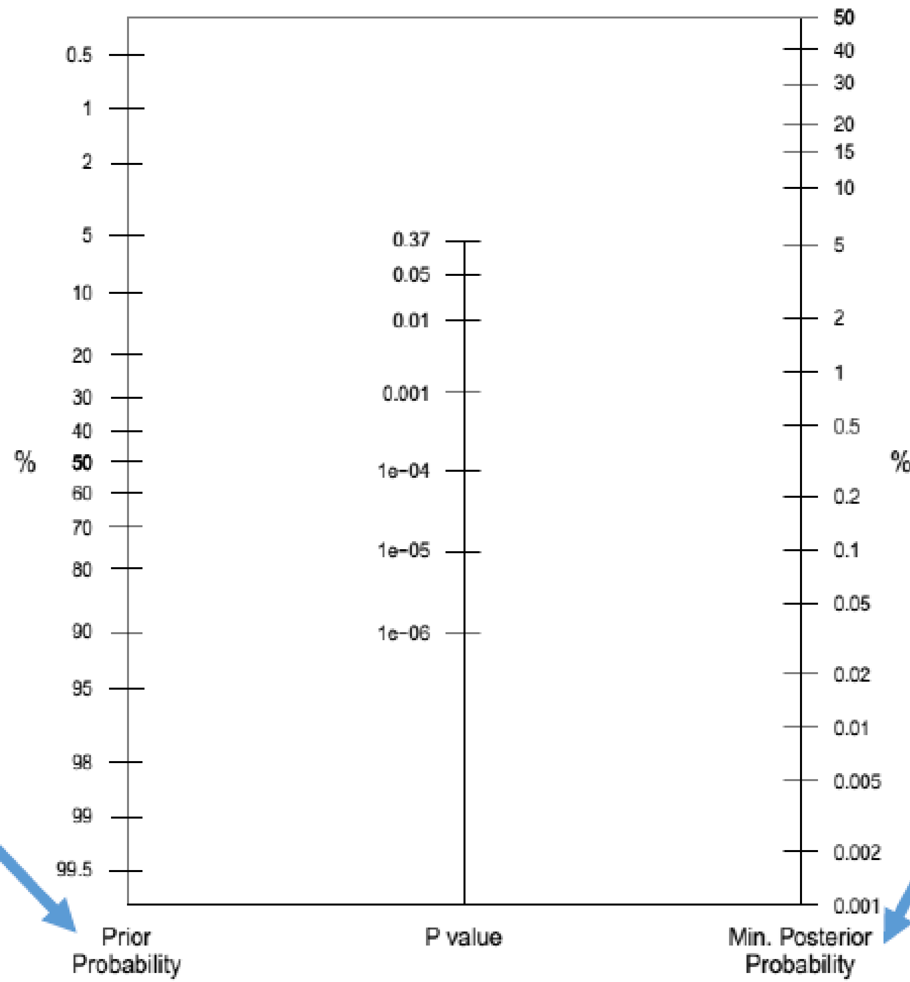
BF > 10, BF < 1/10
STRONG

Bayesian THINKING



outside the context of formal
Bayesian statistics

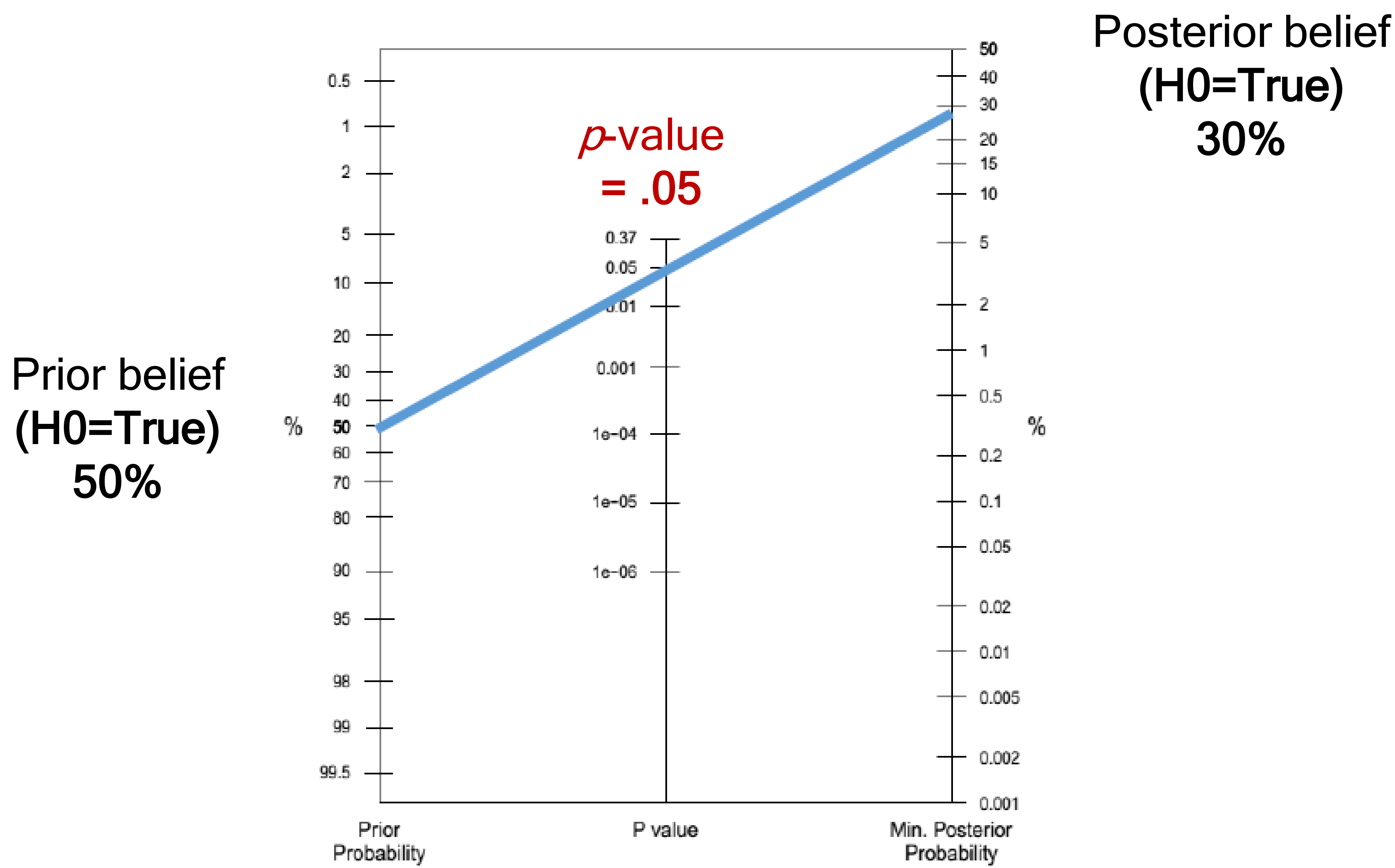
Bayesian THINKING



Prior Probability H0 is True

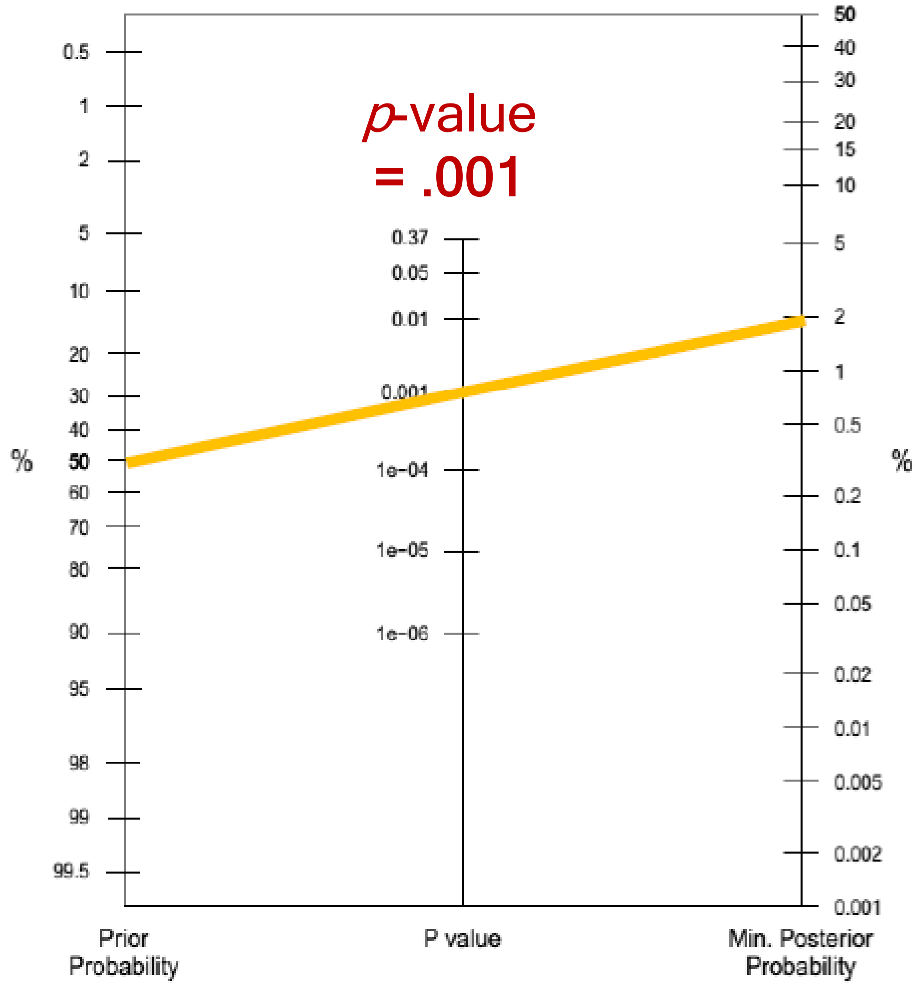
Posterior Probability H0 is True

Bayesian THINKING



Bayesian THINKING

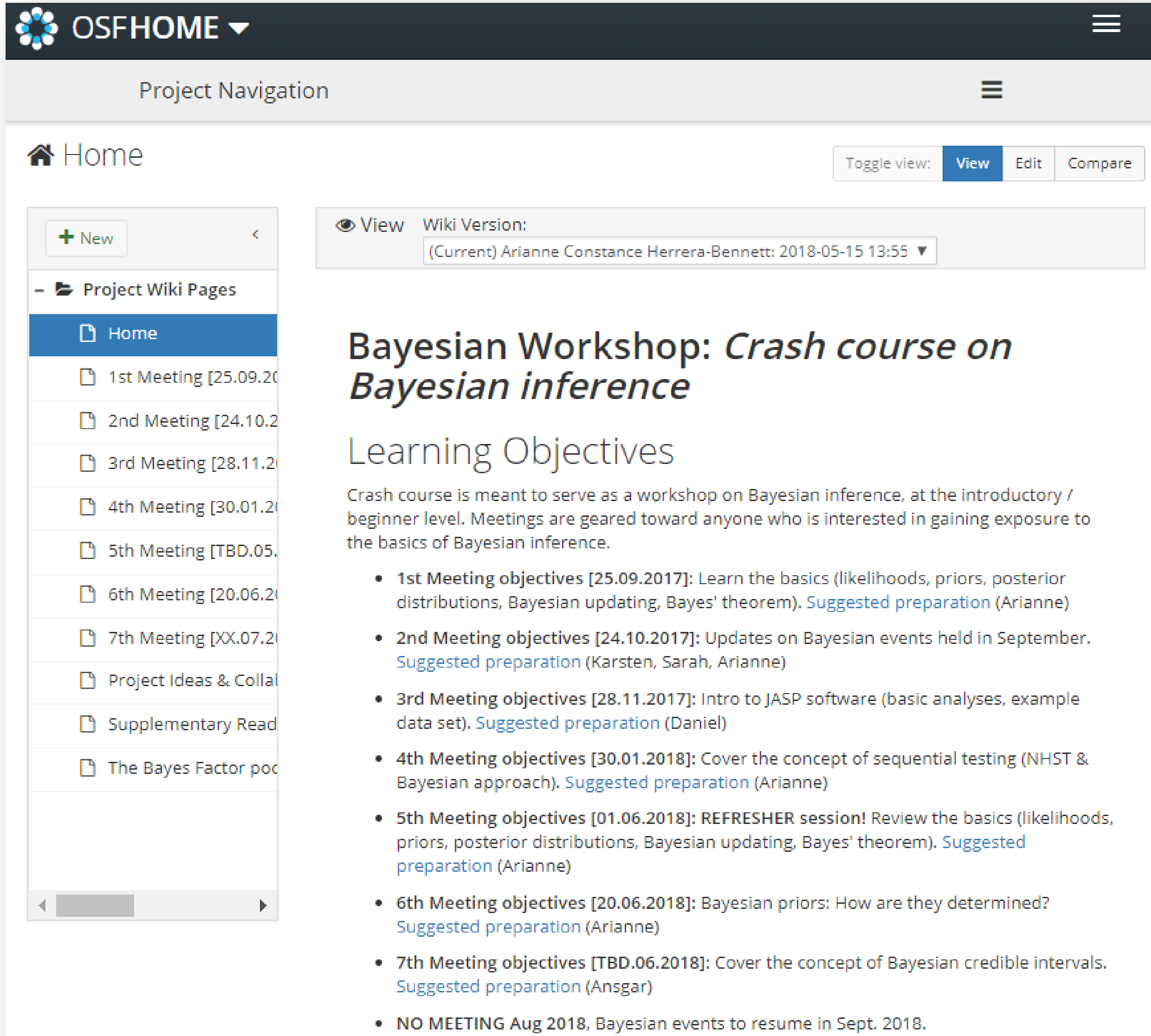
Prior belief
(H0=True)
50%



Posterior belief
(H0=True)
2%

Bayesian THINKING

Taking prior probabilities into account is often smart thinking.



REASON initiatives

Sarah Bichler &
Arianne Herrera-Bennett
(+ Daniel Sommerhoff &
Ansgar Opitz)

https://osf.io/hcm7p/wiki/home/

Upcoming **Guest Talk** & **2-day Workshop** on Bayesian analysis

Bayesian Events & News

Upcoming

- **Guest Talk (24.09.2018, LMU, Munich):** Mark Andrews & Thom Baguley offer a talk on the general issues surrounding teaching Bayesian data analysis to social scientists, what they've learned from their experience, and how they think things will evolve in the future. **[any & all welcome]**
- **2-day Bayesian Workshop (25-26.09.2018, LMU, Munich):** Day 1 *"Bayes for beginners"* aims to be a general introduction to Bayesian data analysis and how it differs from the more familiar classical approaches to data analysis. Day 2 *"Doing Bayesian data analysis"* aims to provide a solid theoretical and practical foundation for real-world Bayesian data analysis in psychology and social sciences. **[details about workshop capacity / sign-up to be announced shortly]**

Ongoing

- ***"The Bayes Factor"*: New podcast**, interviewing the people behind Bayesian statistics and other hot methodological issues in psychological reearch, hosted by JP de Ruiter (@jpderuiter) & Alex Etz (@alxetz). First season interview lineup just announced [5.11.2017]: click here to see!
- ***Episodes:*** See "The Bayes Factor podcast" wiki tab for more info on all posted episodes.

Past

- **Free Seminar (26.09.2017, Munich)** *'Bayesian Networks: Artificial Intelligence for Research, Analytics, and Reasoning'*: In this seminar, we illustrate how scientists in many fields of study - rather than only computer scientists - can employ Bayesian networks as a very practical form of A.I. for exploring complex problems.
- **ESRC funded conference workshop (29.09.2017, Nottingham)** *'Bayesian Data Analysis in the Social Sciences Curriculum'*: Conference considers how and why we should aim to bring Bayesian methods into the statistics curriculum in the social sciences.

REASON initiatives

Sarah Bichler &
Arianne Herrera-Bennett
(+ Daniel Sommerhoff &
Ansgar Opitz)

https://osf.io/hcm7p/wiki/home/

Upcoming **Guest Talk** & **2-day Workshop** on Bayesian analysis

Guest Talk (Monday, Sept. 24th, 2018)
Mark Andrews & Thom Baguley offer a talk on the general issues surrounding teaching Bayesian data analysis to social scientists, what they've learned from their experience, and how they think things will evolve in the future.

Workshop Day 1 (Tuesday, Sept. 25th, 2018) – *"Bayes for beginners"*
This workshop aims to be a general introduction to Bayesian data analysis and how it differs from the more familiar classical approaches to data analysis.

Workshop Day 2 (Wednesday, Sept. 26th, 2018) – *"Doing Bayesian data analysis"*
This workshop aims to provide a solid theoretical and practical foundation for real-world Bayesian data analysis in psychology and social sciences.

REASON initiatives

Sarah Bichler &
Arianne Herrera-Bennett
(+ Daniel Sommerhoff &
Ansgar Opitz)

Mark Andrews
@xmjandrews
Thom Baguley
@seriousstats

http://www.priorexposure.org.uk/

References

Contents borrowed from:

McElreath, R. (2012). Rethinking: statistical Rethinking book package. *R package version*, *1*.

Daniel Laken's Coursera MOOC: *Improving your statistical inferences.*