Bayesian Statistics and Hierarchical Bayesian Modeling for Psychological Science

Lecture 04

Lei Zhang

Social, Cognitive and Affective Neuroscience Unit (SCAN-Unit)

Department of Cognition, Emotion, and Methods in Psychology

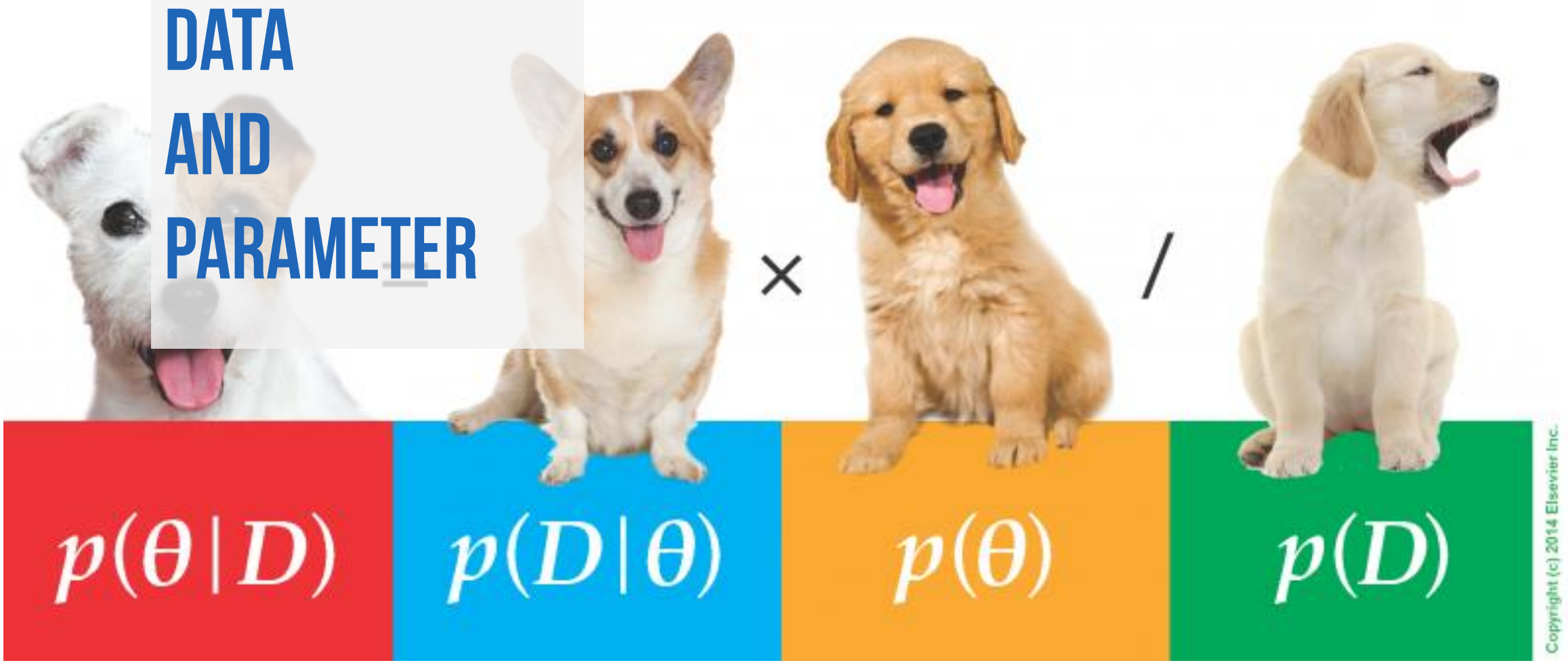
lei.zhang@univie.ac.at
lei-zhang.net
@lei_zhang_lz

https://github.com/lei-zhang/BayesCog_Wien

universität wien
Fakultät für Psychologie

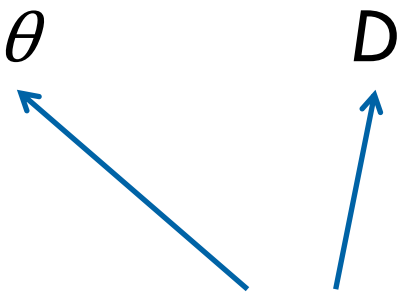
Bayesian warm-up?

LINKING DATA AND PARAMETER



Linking Data and Parameter

cognitive model
statistics
computing



$$p(A|B) = \frac{p(B|A)p(A)}{p(B)}$$

Linking Data and Parameter

cognitive model
statistics
computing

$$p(\theta|D) = \frac{p(D|\theta)p(\theta)}{p(D)}$$

Linking Data and Parameter

cognitive model
statistics
computing

Likelihood
How plausible is the data given our parameter is true?

Prior
How plausible is our parameter before observing the data?

$$p(\theta|D) = \frac{p(D|\theta)p(\theta)}{p(D)}$$

Posterior
How plausible is our parameter given the observed data?

Evidence
How plausible is the data under all possible parameters?

What is $p(\text{Data}|\vartheta)$

- This is the "Model"
- Data is fixed, ϑ varies
- Not a probability distribution
  - the sum is not "one"

L(theta | Data)

Probability of coin landing heads up, θ	Number of heads, X 0	1	2	Total
0.0	1.00	0.00	0.00	**1.00**
0.2	0.64	0.32	0.04	**1.00**
0.4	0.36	0.48	0.16	**1.00**
0.6	0.16	0.48	0.36	**1.00**
0.8	0.04	0.32	0.64	**1.00**
1.0	0.00	0.00	1.00	**1.00**
Total	**2.20**	**1.60**	**2.20**	

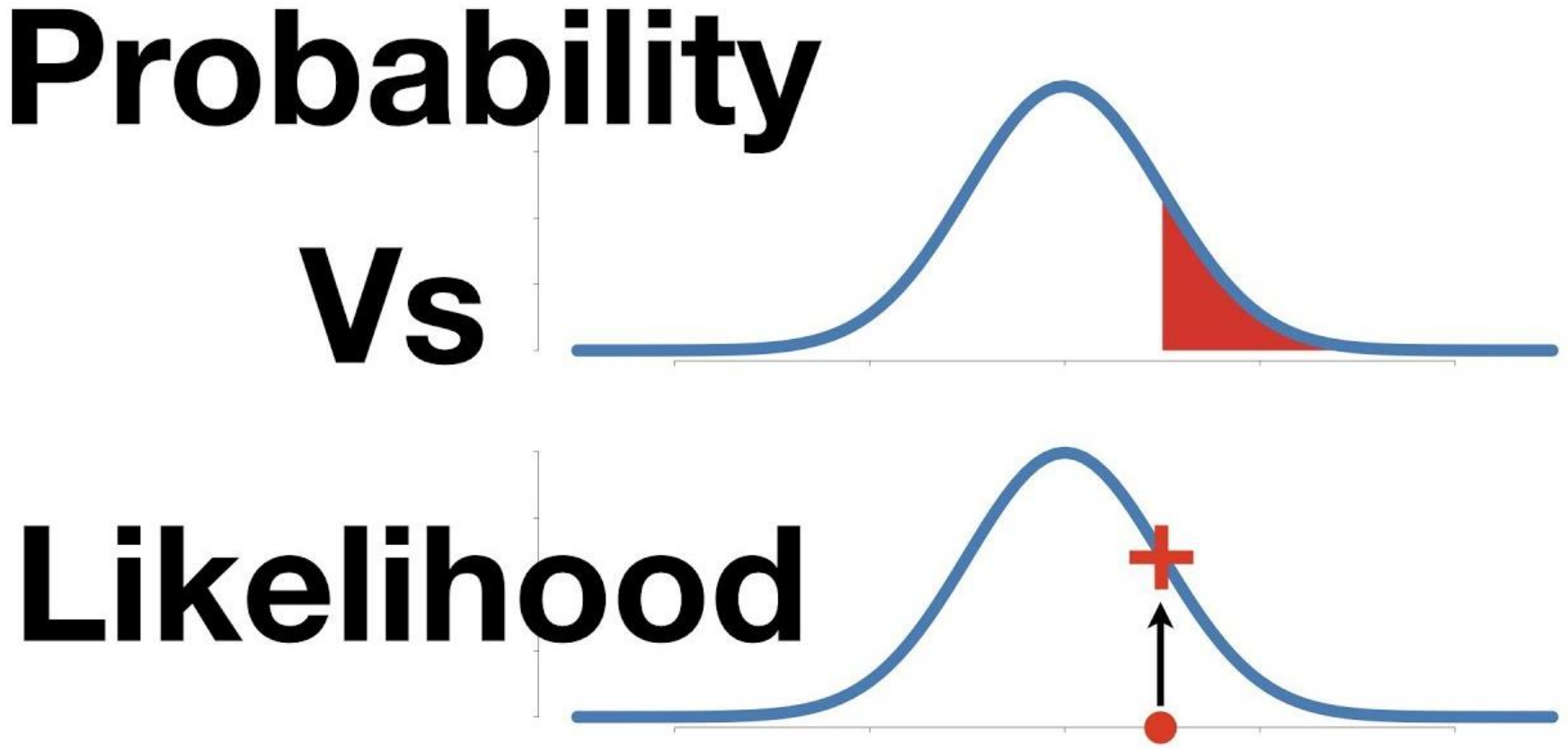
$$Pr(X=0|\theta)=Pr(T,T|\theta)=Pr(T|\theta)\times Pr(T|\theta)=(1-\theta)^2$$

$$Pr(X=1|\theta)=Pr(H,T|\theta)+Pr(T,H|\theta)=2\times Pr(T|\theta)\times Pr(H|\theta)=2\theta(1-\theta)$$

$$Pr(X=2|\theta)=Pr(H,H|\theta)=Pr(H|\theta)\times Pr(H|\theta)=\theta^2.$$

Lambert (2018)

Watch this video!

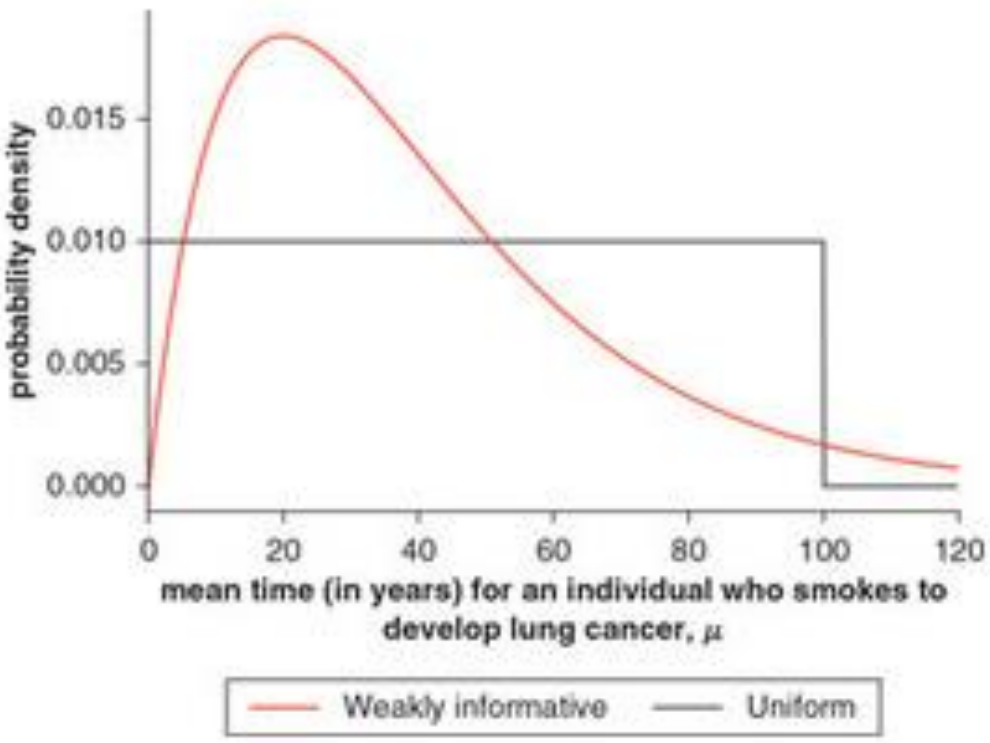


StatQuest with Josh Starmer
250K subscribers

https://youtu.be/pYxNSUDSFH4

What is $p(\vartheta)$?

cognitive model
statistics
computing



Lambert (2018)

What is *p*(*Data*)?

cognitive model
statistics
computing

discrete parameters

$$p(\theta \mid D) = \frac{p(D \mid \theta)\, p(\theta)}{\sum_{\theta^*} p(D \mid \theta^*)\, p(\theta^*)}$$

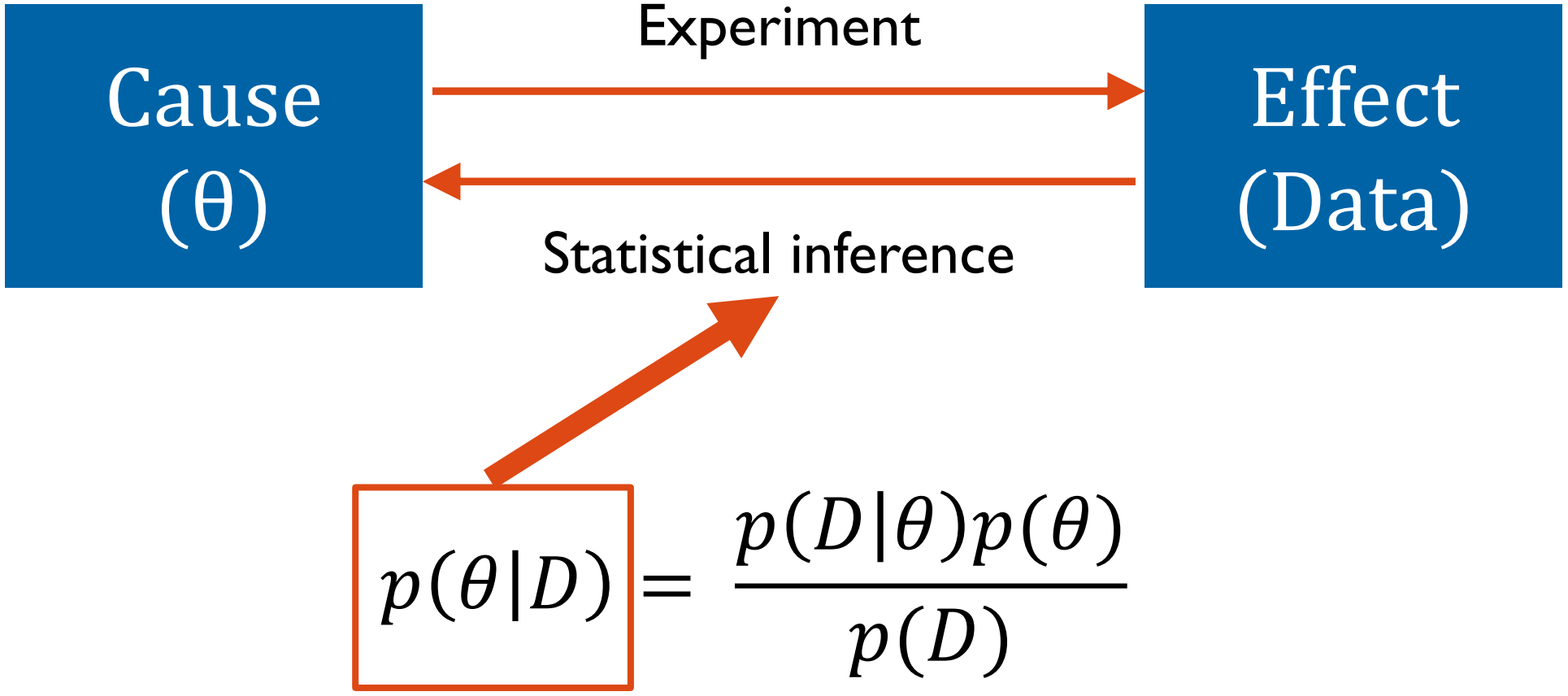
$$p(\theta|D) = \frac{p(D|\theta)p(\theta)}{p(D)}$$

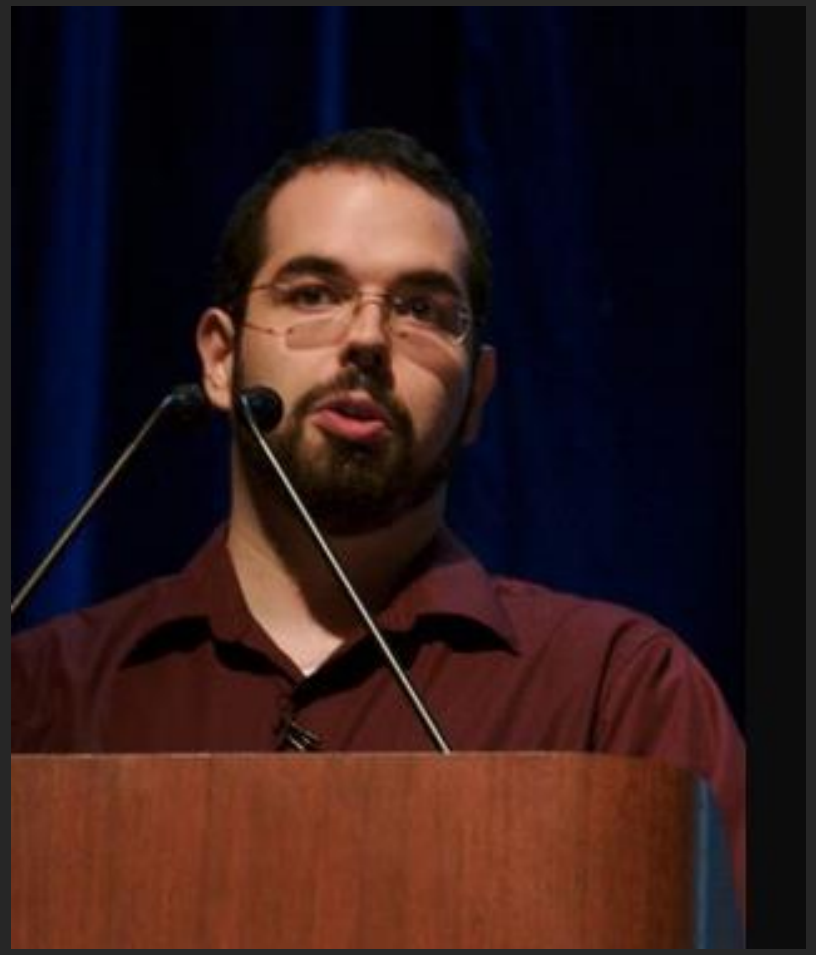
continuous parameters

$$p(\theta \mid D) = \frac{p(D \mid \theta)\, p(\theta)}{\int p(D \mid \theta^*)\, p(\theta^*)\, d\theta^*}$$

Why the Bayes' theorem is important?

cognitive model
statistics
computing





“Probability is orderly opinion and inference from data is nothing other than the revision of such opinion in the light of relevant new information.”

Eliezer S. Yudkowsky

BINOMIAL MODEL

Binomial Model

cognitive model
statistics
computing

- You are curious how much of the surface is covered in water.
- You will toss the globe up in the air.
- You will record whether or not the surface under your right index finger is water (W) or land (L).
- You might observe: W L W W W L W L W
- → 6/9 = 0.666667?
- Is it right? If not, what to do next?

McElreath (2016)

Steps of (Bayesian) Modeling?

cognitive model
statistics
computing

A data story

Think about how the data might arise.
It can be *descriptive* or even *causal*.

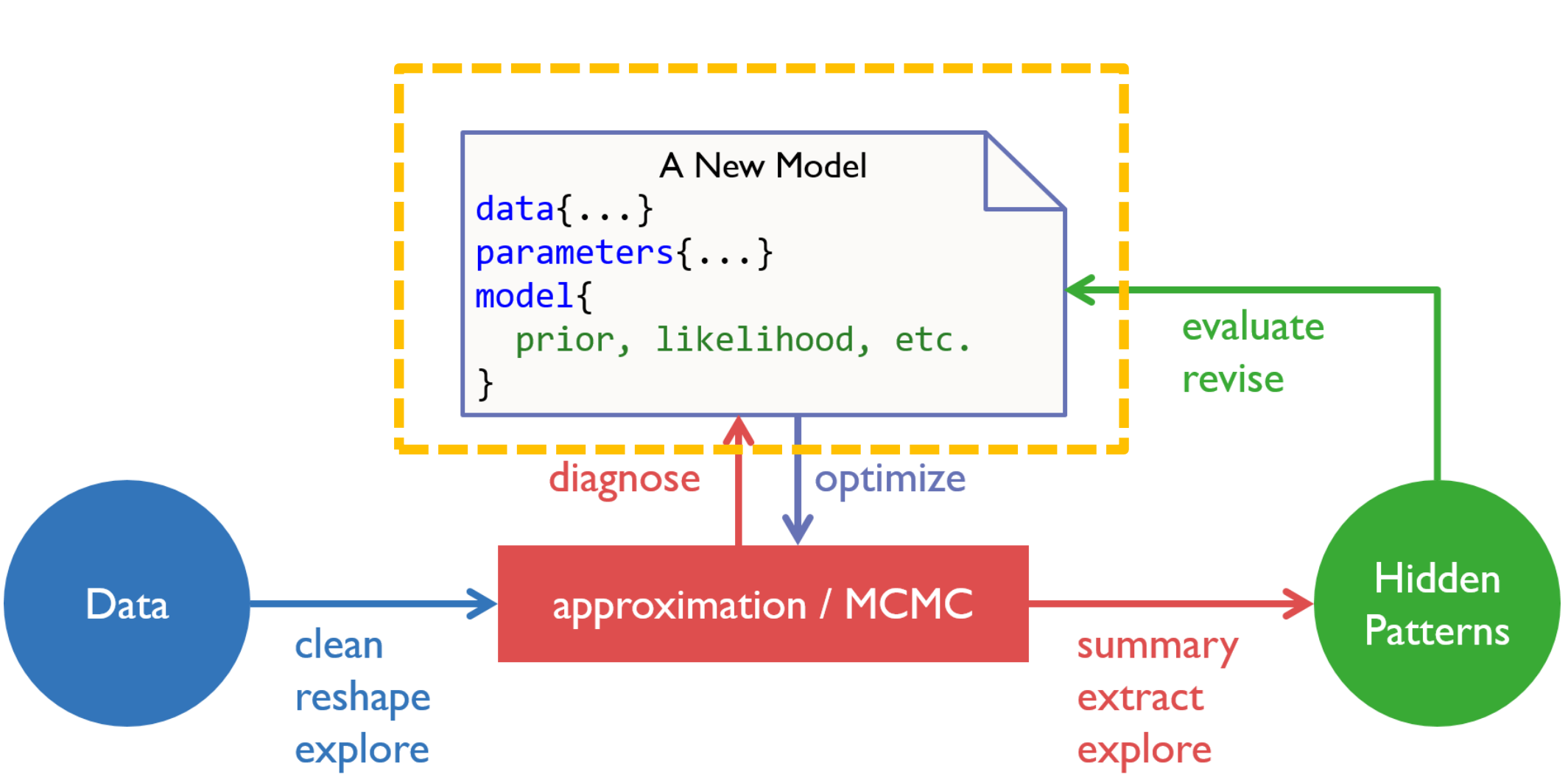
Update

Educate your model by feeding it with data.
Bayesian Update:
update the prior, in light of data, to produce posterior
the updated posterior then becomes the prior of next update

Evaluate

Compare model with reality.
Revise your model.

McElreath (2016)


cognitive model
statistics
computing
A New Model
data{...}
parameters{...}
model{
prior, likelihood, etc.
}
evaluate
revise
diagnose
optimize
Data
clean
reshape
explore
approximation / MCMC
summary
extract
explore
Hidden
Patterns

A Data Story of the Globe

cognitive model
statistics
computing

- The true proportion of water covering the globe is ϑ.
- A single toss of the globe has a probability ϑ of producing a water (W) observation.
- It has a probability $(1 - \vartheta)$ of producing a land (L) observation.
- Each toss of the globe is independent of the others.

McElreath (2016)

Components of a Model

cognitive model
statistics
computing

think about the likelihood function (of Binomial):

$$p(\theta \mid D) = \frac{p(D \mid \theta)\, p(\theta)}{\int p(D \mid \theta^*)\, p(\theta^*)\, d\theta^*}$$

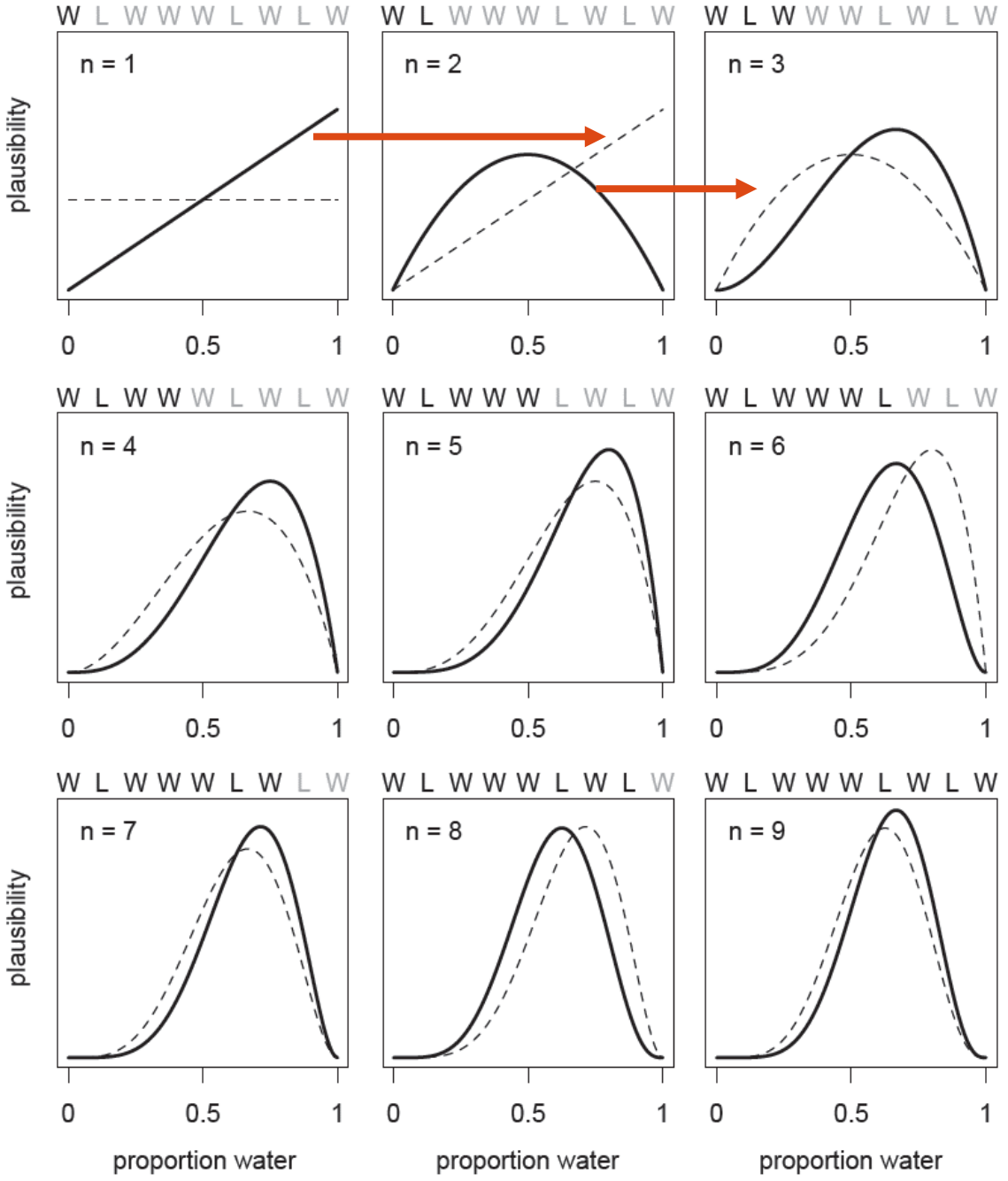
$$p(w \mid N, \theta) = \binom{N}{w} \theta^w (1-\theta)^{N-w}$$

N: total number of observations
w: number of water
ϑ : proportion of water

} known (data)

unknown (parameter)

Update



cognitive model

statistics

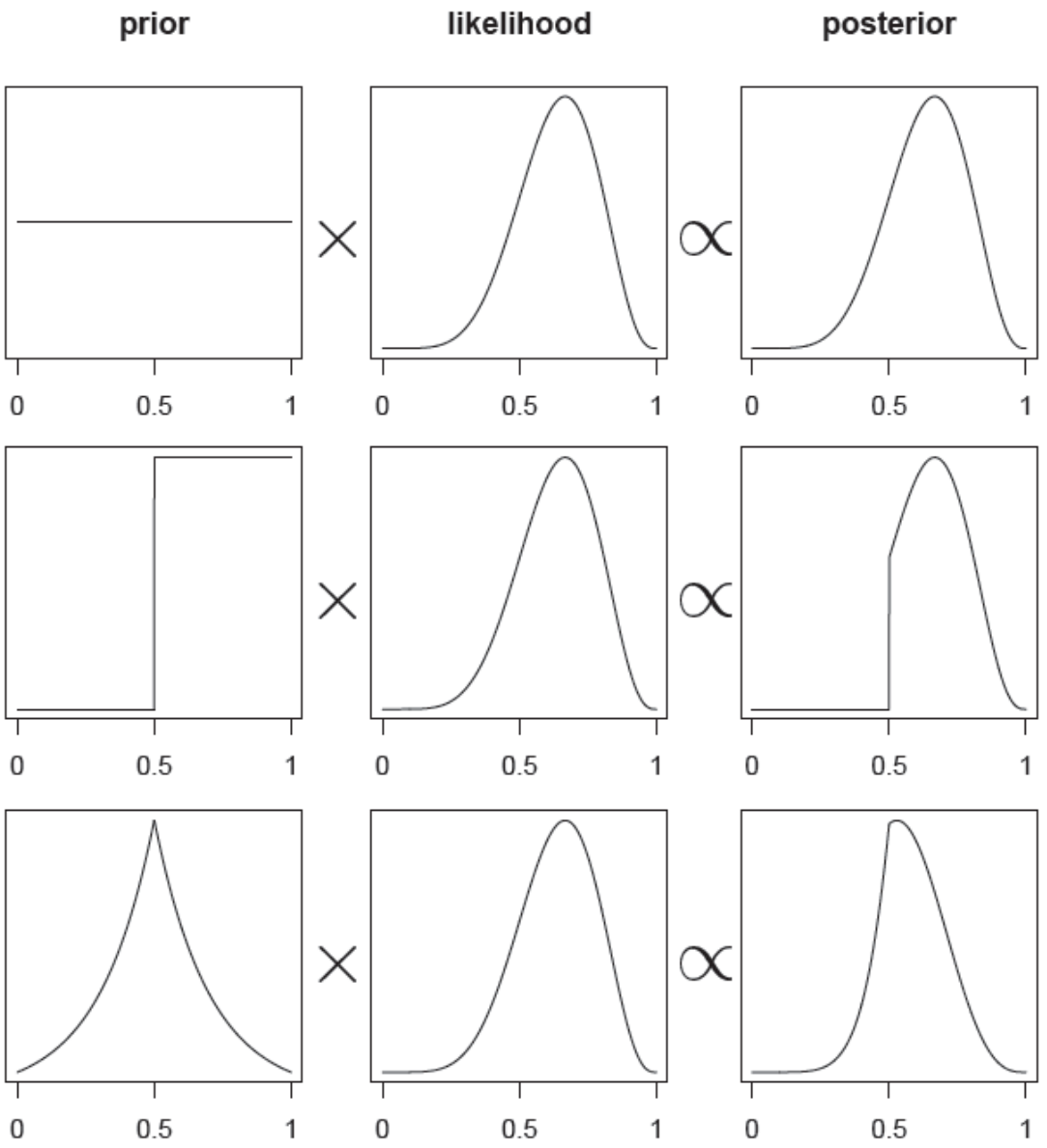
computing

- order doesn't matter
- 2/3 is most likely
- others are not ruled out

McElreath (2016)

Impact of Prior

cognitive model
statistics
computing



McElreath (2016)


ANY
QUESTIONS
?

Happy Computing!