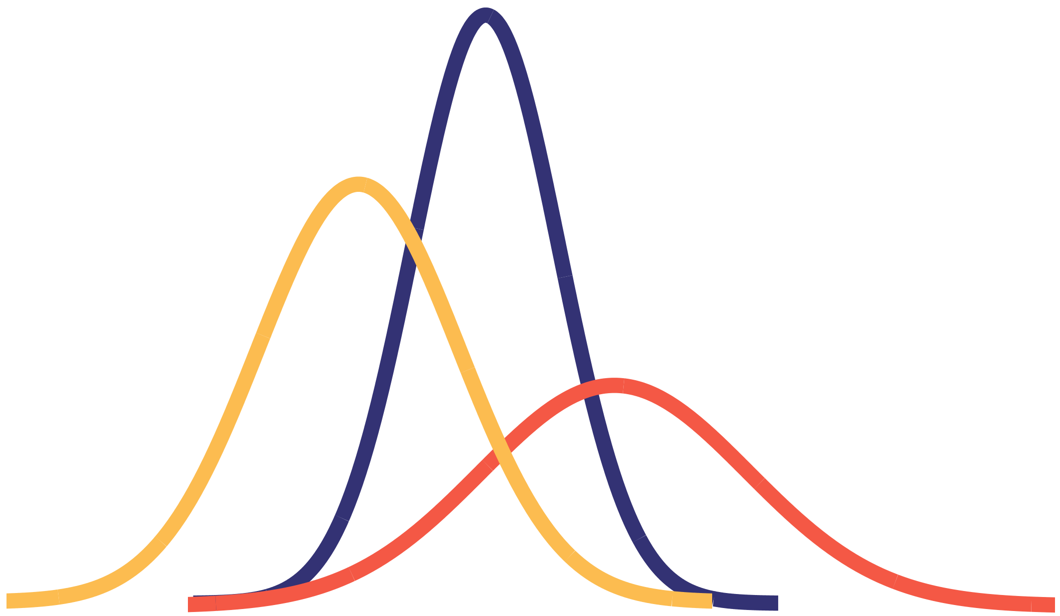
Hierarchical Bayesian modeling with applications for spatial environmental data science

UCLA IDRE – May 6, 2022



Instructors: Casey Youngflesh (Postdoc, Dept Eco and Evo Bio)
Sudipto Banerjee (Professor, Dept of Biostatistics)

Schedule

Time (PST)	Presentation title (speaker)
09:00 AM – 09:05 AM	Welcome and Introduction
09:05 AM – 12:00 PM	Intro to hierarchical Bayesian modeling using Stan (Instructor: Youngflesh)
12:00 PM – 01:00 PM	Lunch break
01:00 PM – 04:00 PM	Hierarchical Bayesian modeling for spatial data science (Instructor: Banerjee)

Software – Part 1

R

RStudio®

Stan

https://github.com/caseyyoungflesh/IDRE-Hierarchicial-Bayesian-Modeling

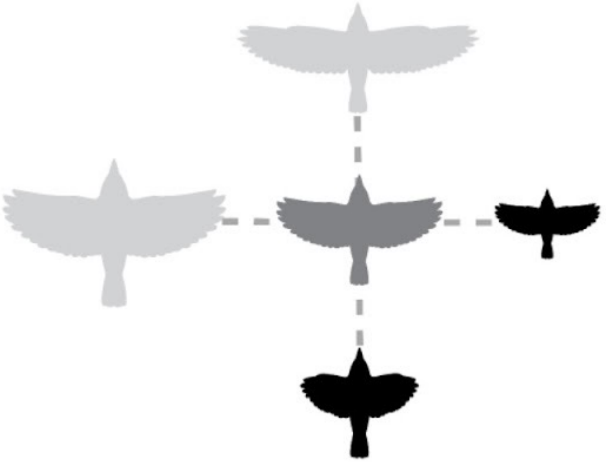
Goals

- Understand the basics of a principled approach to spatial environmental data science using a hierarchical Bayesian approach.
- Fit basic models using Stan in R
- Develop a basis of knowledge that will facilitate future learning

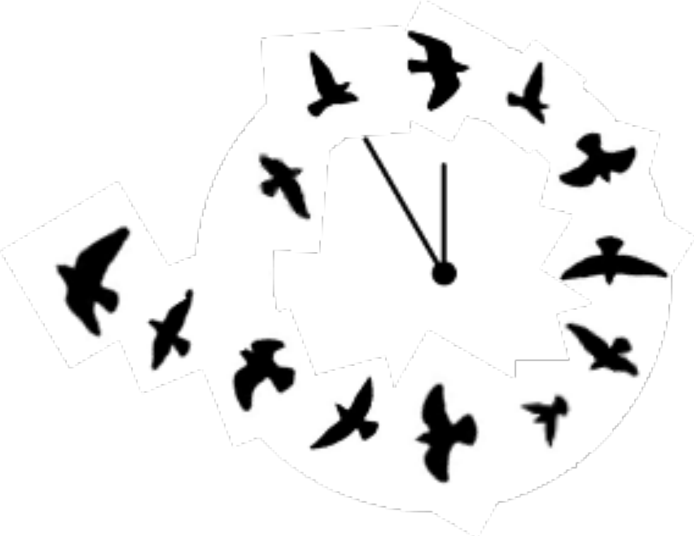

100% GUARANTEE
MONEY BACK
100% GUARANTEE

Casey Youngflesh

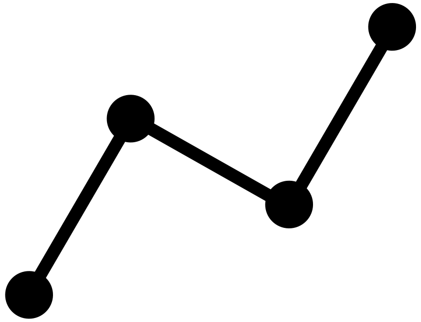
Postdoc, Dept. Ecology and Evolutionary Biology



Biodiversity



Phenology



Demographic Dynamics

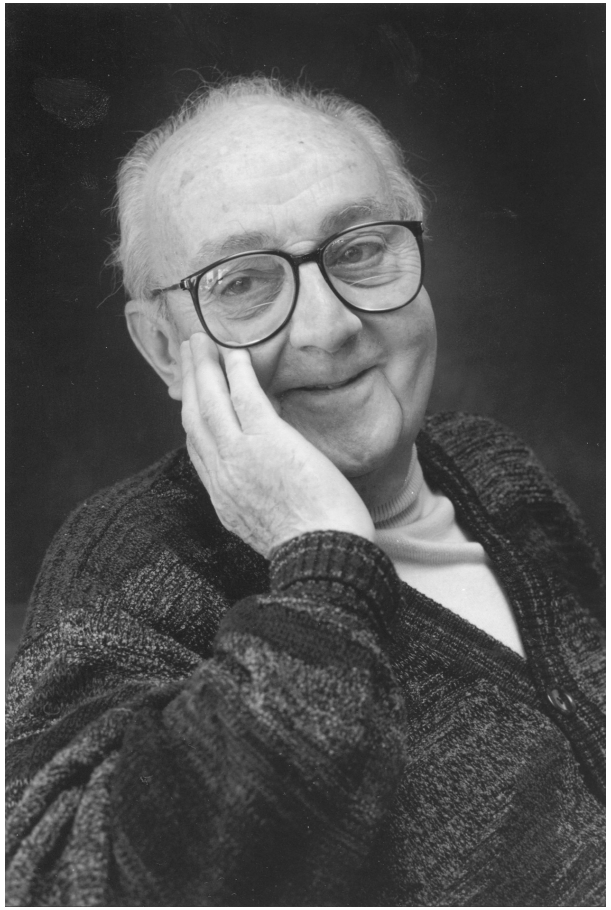
What is a model?

What is a model?

- Abstraction of some process of interest

What is a model?

- Abstraction of some process of interest
- Idealized representation of the data generating process



... all models are approximations. Essentially, all models are wrong, but some are useful. However, the approximate nature of the model must always be borne in mind....

-George Box

Scientific questions

↓

Models and Observations

↓

New insight (and uncertainty)

Why Bayesian?

Flexibility

Why Bayesian?

Flexibility

-multiple data sources
-multiple sources of uncertainty
-missing data
-unobserved quantities
-forecasting

The Bayesian worldview

The Bayesian worldview

- Observed quantities

The Bayesian worldview

- Observed quantities
- Unobserved quantities

The Bayesian worldview

- Observed quantities
- Unobserved quantities
    - parameters
    - latent states
    - missing data

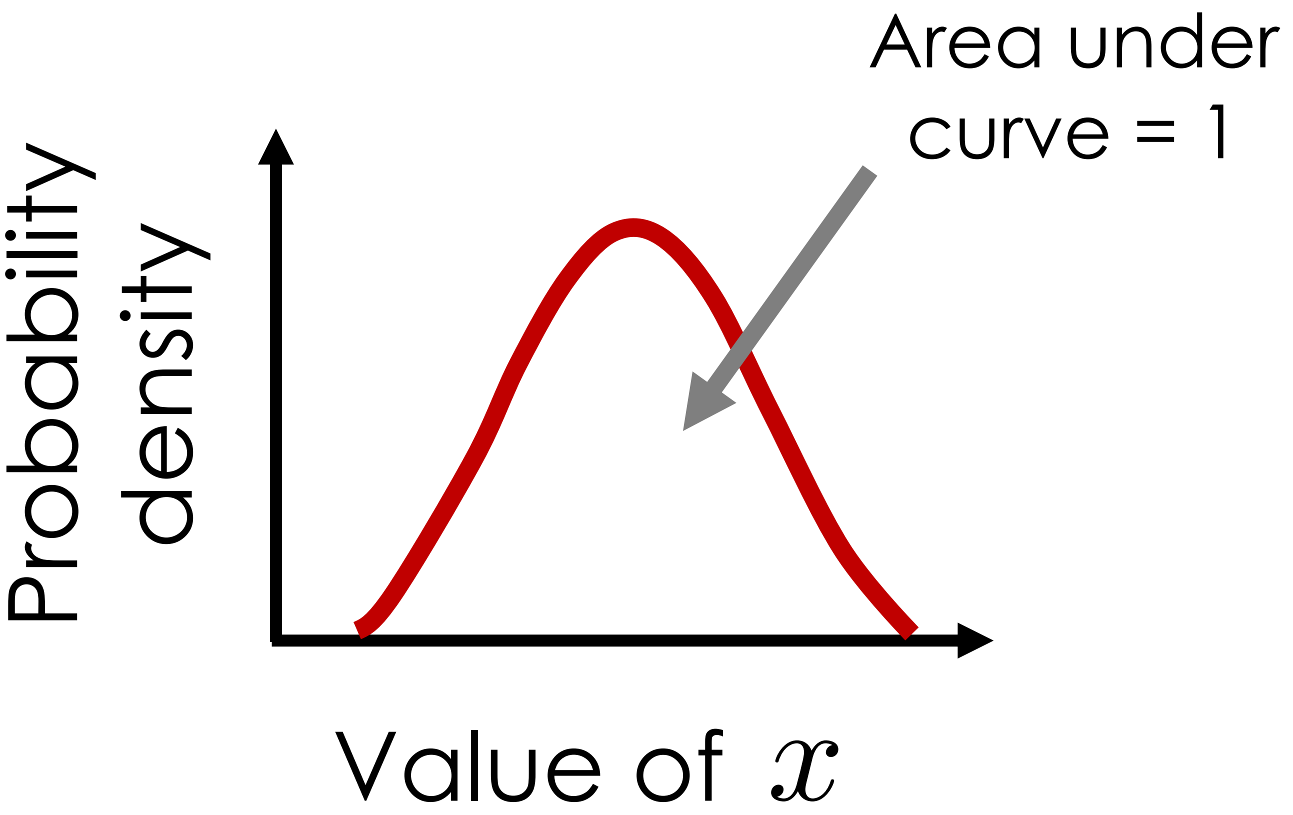
The Bayesian worldview

- Observed quantities
- Unobserved quantities
  - parameters
  - latent states
  - missing data

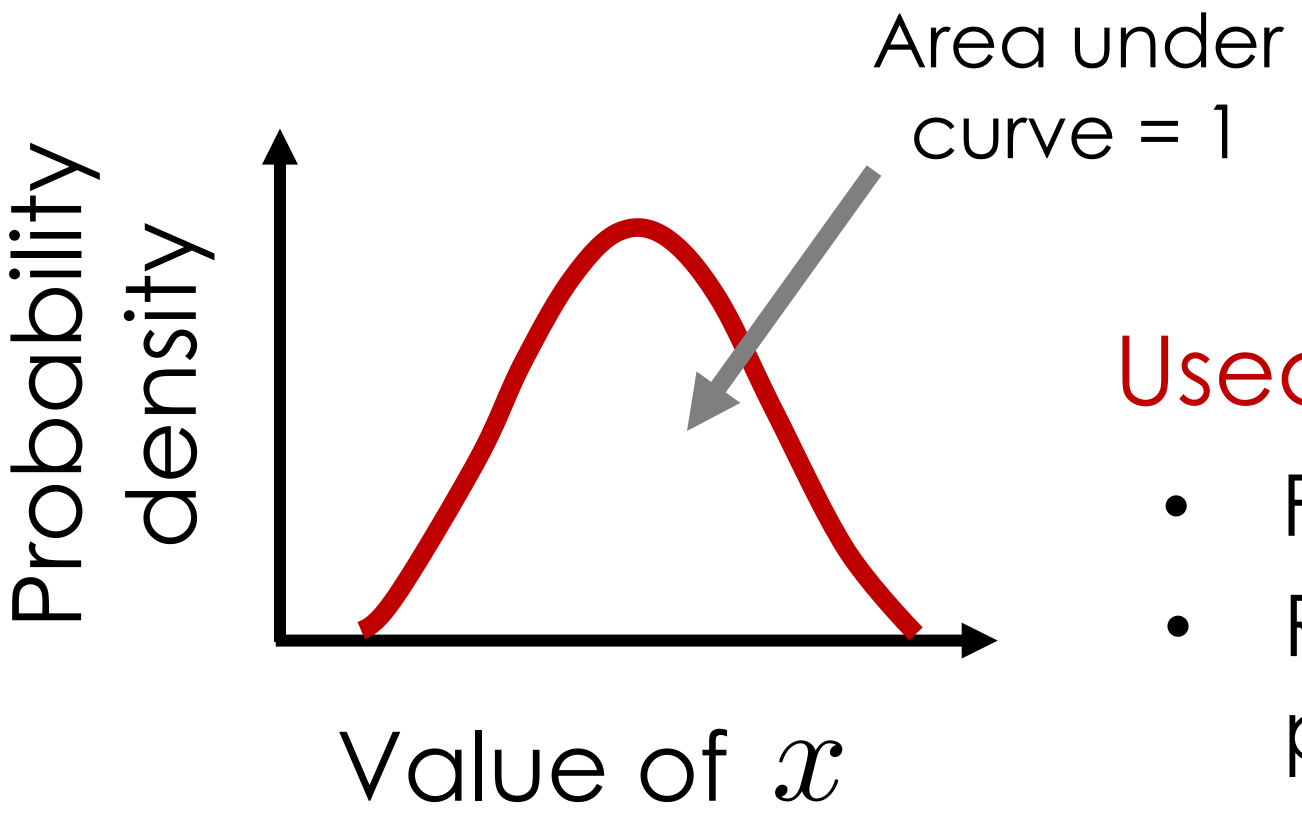
Governed by probability distributions

$$\underbrace{Pr}_{\text{Probability of}}(\overbrace{\theta}^{\text{Unobserved qtys}}|\underbrace{y}_{\text{Observed qtys}})$$

Probability distribution?



Probability distribution?



Used to:

- Fit models to data
- Represent uncertainty in parameters
- Portray prior information

Normal Probability Density Function (PDF)

$$f(x|\mu,\sigma) = \frac{1}{\sqrt{2\pi * \sigma^2}} * e^{-\frac{1}{2}*(\frac{x-\mu}{\sigma})^2}$$

```r
#fixed parameters
mu <- 4
sigma <- 2

#varying data
x <- seq(-5, 15, length.out = 1000)
```

1-pdf.R

Normal Probability Density Function (PDF)

$$f(x|\mu,\sigma) = \frac{1}{\sqrt{2\pi * \sigma^2}} * e^{-\frac{1}{2}*(\frac{x-\mu}{\sigma})^2}$$

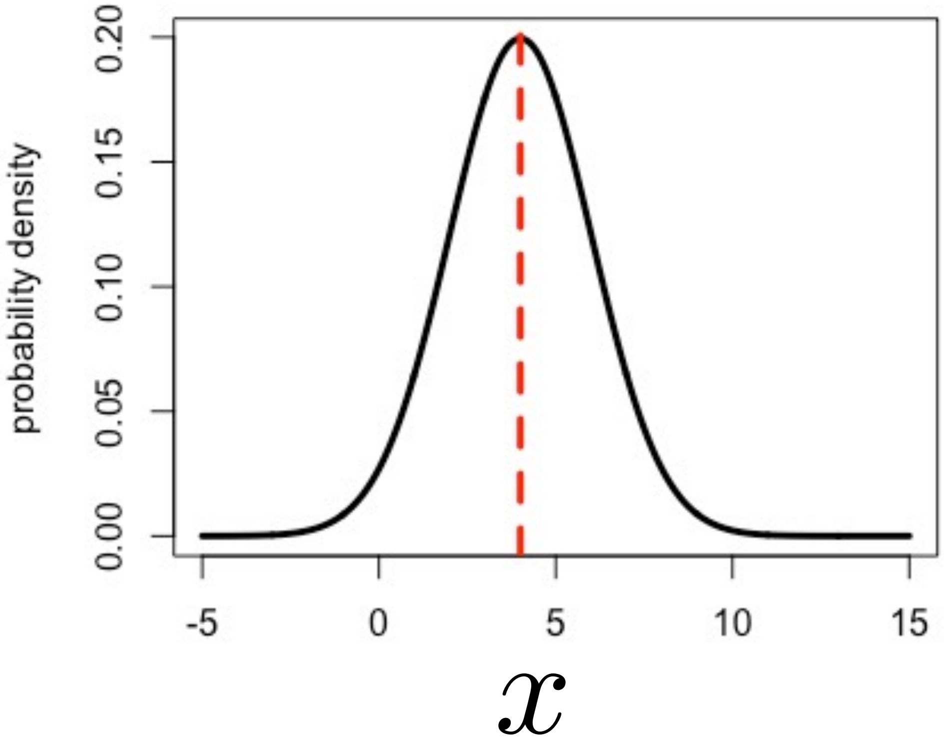
```
#fixed parameters
mu <- 4
sigma <- 2

#varying data
x <- seq(-5, 15, length.out = 1000)

#calculate density
pd <- dnorm(x, mu, sigma)
```

1-pdf.R

Normal Probability Density Function (PDF)

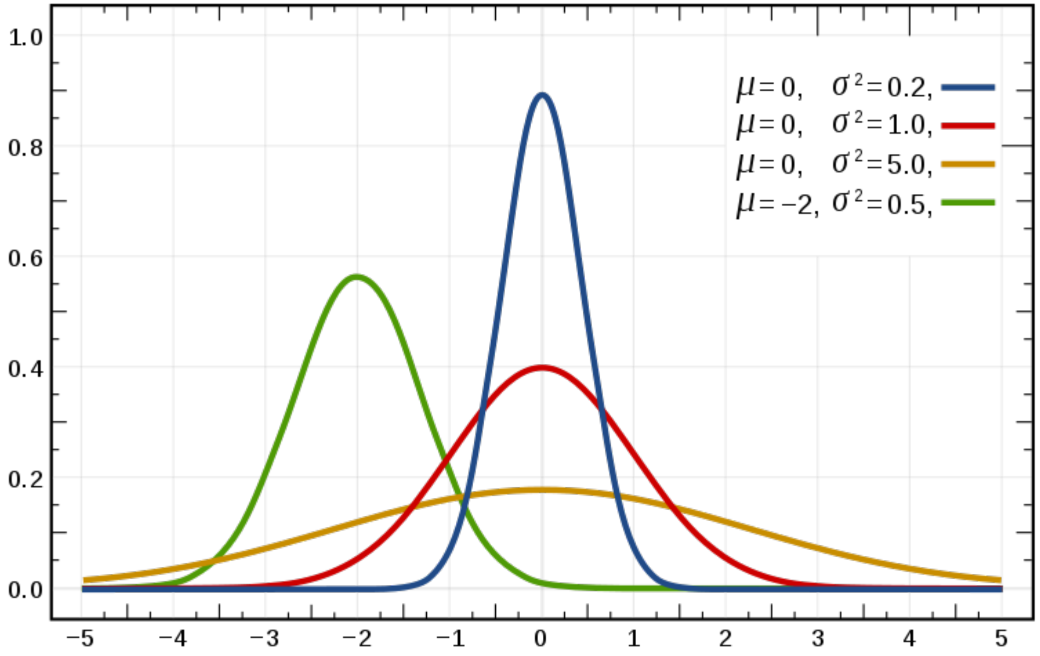
$$f(x|\mu,\sigma) = \frac{1}{\sqrt{2\pi * \sigma^2}} * e^{-\frac{1}{2}*(\frac{x-\mu}{\sigma})^2}$$

```r
#fixed parameters
mu <- 4
sigma <- 2

#varying data
x <- seq(-5, 15, length.out = 1000)

#calculate density
pd <- dnorm(x, mu, sigma)

#plot
plot(x, pd, type = 'l', lwd = 3,
     xlab = ‘x',
     ylab = 'probability density')
abline(v = mu, lwd = 3, lty = 2, col = 'red')
```



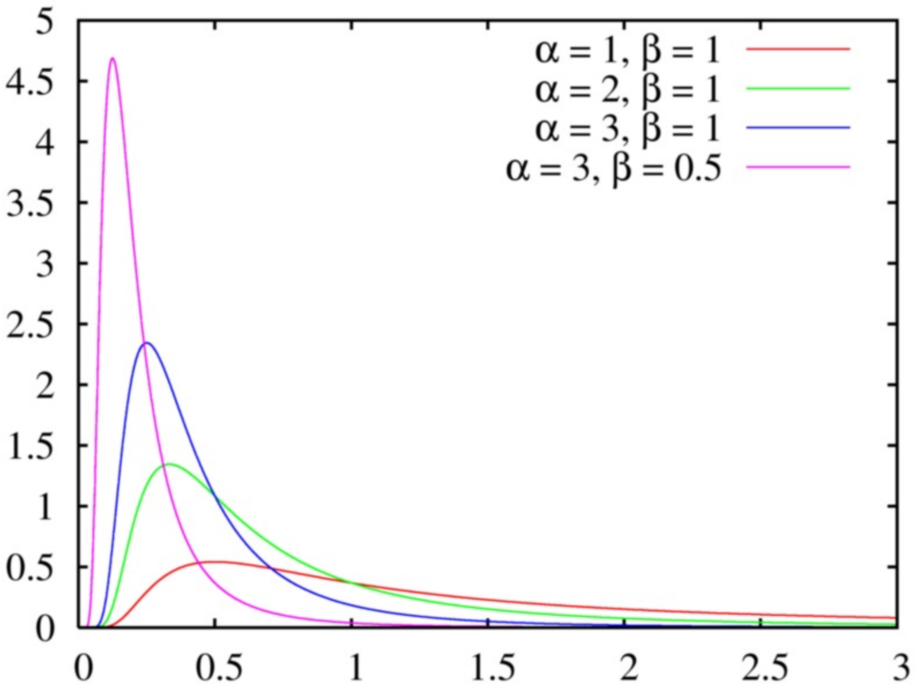
1-pdf.R

Normal distribution



$$f(x|\mu,\sigma) = \frac{1}{\sqrt{2\pi * \sigma^2}} * e^{-\frac{1}{2}*\left(\frac{x-\mu}{\sigma}\right)^2}$$

Gamma distribution



$$f(x|\alpha,\beta) = \frac{\beta^\alpha}{\Gamma(\alpha)} x^{\alpha-1} e^{-\beta x}$$

Distributions

http://www.math.wm.edu/~leemis/chart/UDR/UDR.html

Which distribution?

normal	➡	real numbers
lognormal	➡	non-neg real numbers
gamma	➡	non-neg real numbers
beta	➡	0 to 1 real numbers
binomial	➡	counts in 2 categories
Bernoulli	➡	0 or 1
Poisson	➡	counts

Which distribution?

normal	➡	real numbers
lognormal	➡	non-neg real numbers
gamma	➡	non-neg real numbers
beta	➡	0 to 1 real numbers
binomial	➡	counts in 2 categories
Bernoulli	➡	0 or 1
Poisson	➡	counts

- Presence or absence of a species in a collection of forests
- The number of plants in a $1km^2$ patch
- Mean summer temperatures across North America
- The mass of adult birds across their range
- Total above ground biomass in a $1km^2$ patch
- The probability that an individual survives to adulthood

Bayes Rule

$$Pr(\theta|y)$$

Bayes Rule

$$Pr(\theta|y) \qquad \underbrace{Pr(y|\theta)}_{\text{Likelihood}}$$

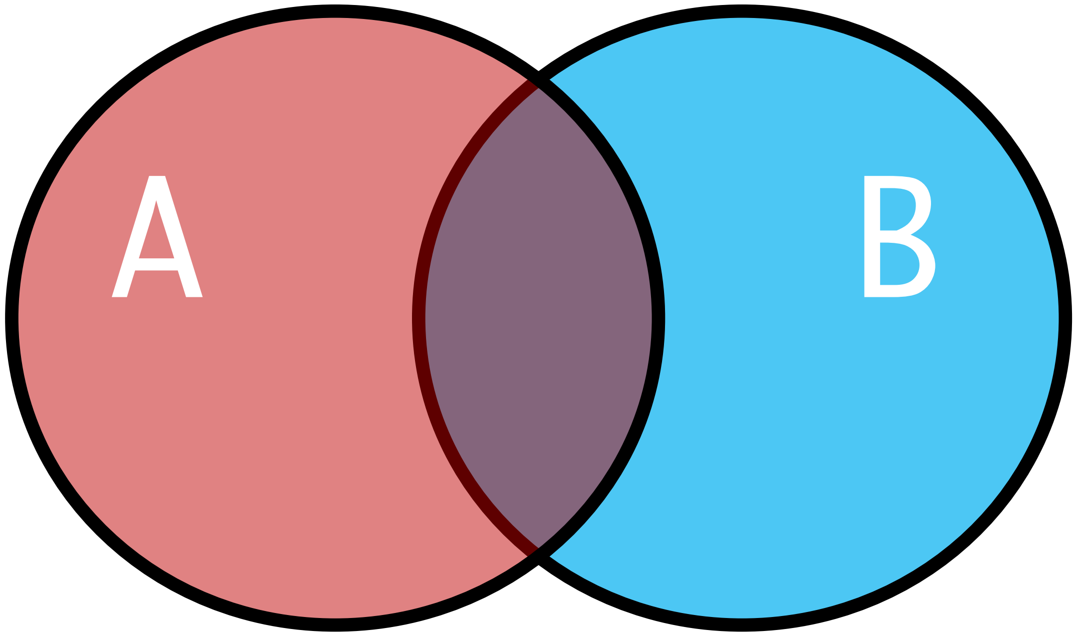
Bayes Rule

$$Pr(\theta|y) \qquad \underbrace{Pr(y|\theta)}_{\text{Likelihood}}\overbrace{Pr(\theta)}^{\text{Prior}}$$

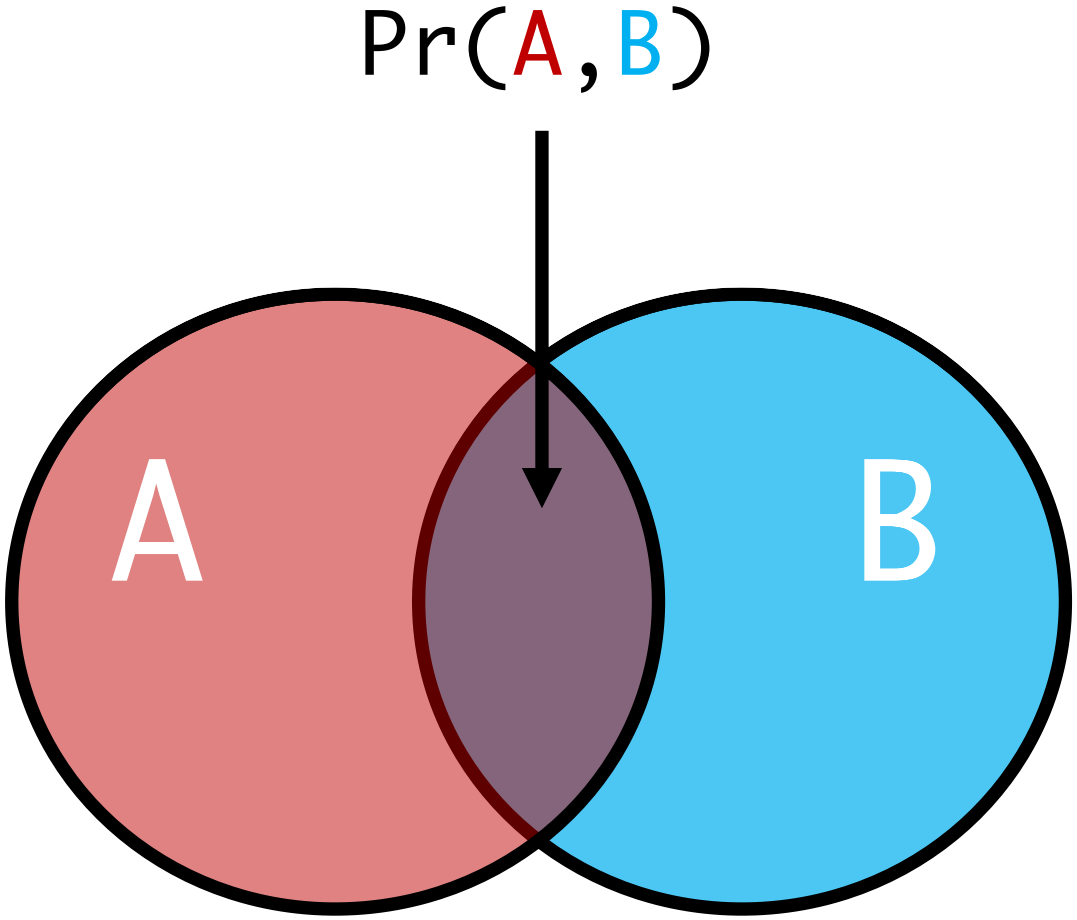
Bayes Rule

$$\overbrace{Pr(\theta|y)}^{\text{Posterior}} \overset{\text{‘Proportional to’}}{\propto} \underbrace{Pr(y|\theta)}_{\text{Likelihood}} \overbrace{Pr(\theta)}^{\text{Prior}}$$

Bayes Rule

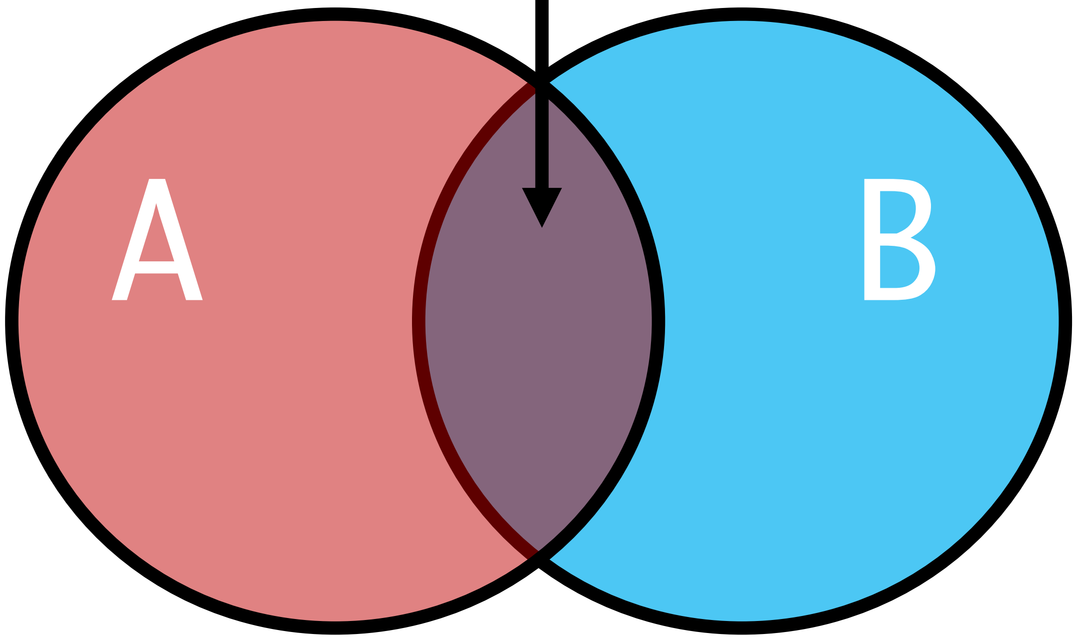


Bayes Rule



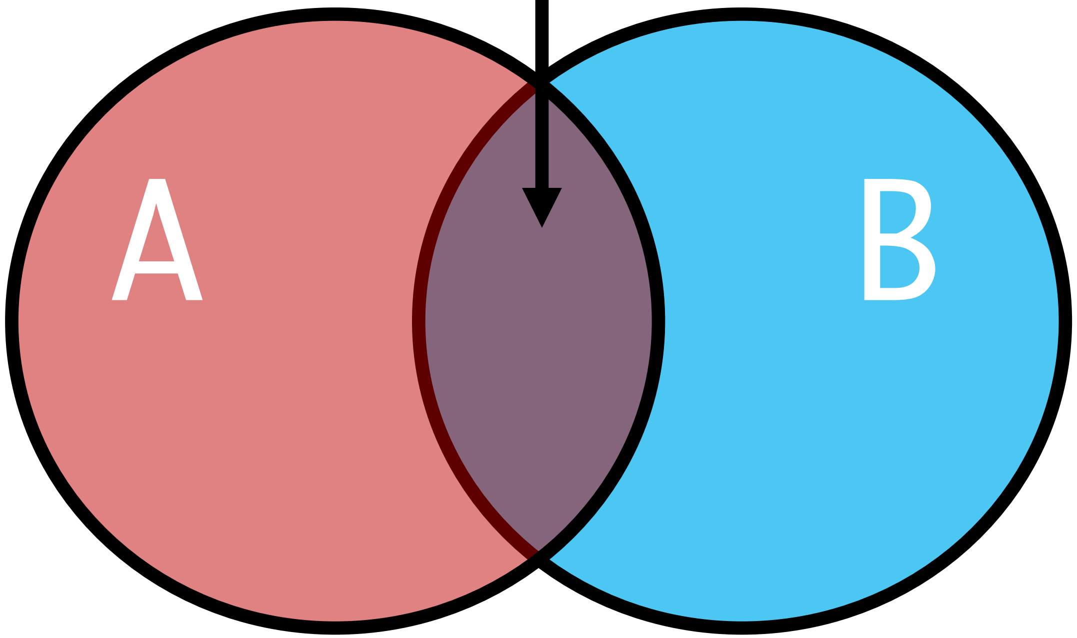
Bayes Rule

$$\Pr(A,B) = \Pr(B|A)\Pr(A)$$



Bayes Rule

$$\Pr(A|B)\Pr(B) = \Pr(A,B) = \Pr(B|A)\Pr(A)$$



Bayes Rule

$$\Pr(A|B)\Pr(B) = \Pr(A,B) = \Pr(B|A)\Pr(A)$$

Bayes Rule

$$\Pr(A|B)\Pr(B) = \Pr(A,B) = \Pr(B|A)\Pr(A)$$

!!!

Bayes Rule

$$\Pr(A|B)\Pr(B) = \Pr(A,B) = \Pr(B|A)\Pr(A)$$

$$\Pr(A|B) = \frac{\Pr(A,B)}{\Pr(B)}$$

Bayes Rule

$$\Pr(A|B)\Pr(B) = \Pr(A,B) = \Pr(B|A)\Pr(A)$$

$$\Pr(A|B) = \frac{\Pr(A,B)}{\Pr(B)}$$

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

Bayes Rule

$$\Pr(A|B)\Pr(B) = \Pr(A,B) = \Pr(B|A)\Pr(A)$$

$$\Pr(A|B) = \frac{\Pr(A,B)}{\Pr(B)}$$

$$\Pr(\theta \mid y) = \frac{\Pr(y \mid \theta)\Pr(\theta)}{\Pr(y)}$$

$$\underbrace{Pr(\theta|y)}_{\substack{\text{Probability of the} \\ \text{unobserved, given the} \\ \text{observed} \\ \text{(Posterior)}}} = \frac{Pr(y|\theta)Pr(\theta)}{Pr(y)}$$

$$\underbrace{Pr(\theta|y)}_{\substack{\text{Probability of the} \\ \text{unobserved, given the} \\ \text{observed} \\ \text{(Posterior)}}} = \frac{\overbrace{Pr(y|\theta)}^{\substack{\text{Probability of the observed,} \\ \text{given the unobserved} \\ \text{(Likelihood)}}}Pr(\theta)}{Pr(y)}$$

$$\underbrace{Pr(\theta|y)}_{\substack{\text{Probability of the}\\ \text{unobserved, given the}\\ \text{observed}\\ \text{(Posterior)}}} = \frac{\overbrace{Pr(y|\theta)}^{\substack{\text{Probability of the observed,}\\ \text{given the unobserved}\\ \text{(Likelihood)}}}\ \overbrace{Pr(\theta)}^{\substack{\text{Probability of the}\\ \text{unobserved}\\ \text{(Prior)}}}}{Pr(y)}$$

$$\underbrace{Pr(\theta|y)}_{\substack{\text{Probability of the} \\ \text{unobserved, given the} \\ \text{observed} \\ \text{(Posterior)}}} = \frac{\overbrace{Pr(y|\theta)}^{\substack{\text{Probability of the observed,} \\ \text{given the unobserved} \\ \text{(Likelihood)}}}\overbrace{Pr(\theta)}^{\substack{\text{Probability of the} \\ \text{unobserved} \\ \text{(Prior)}}}}{\underbrace{Pr(y)}_{\text{Probability of the observed}}}$$

Bayes Rule

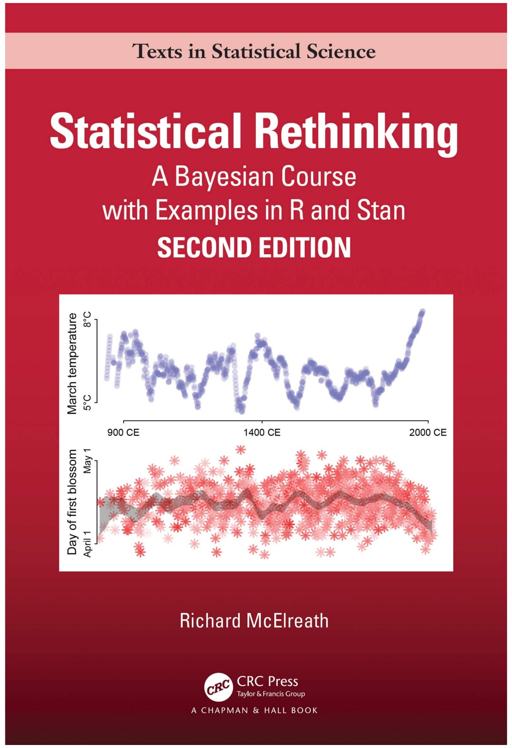
$$Pr(\theta|y) = \frac{Pr(y|\theta)Pr(\theta)}{Pr(y)}$$

We ignore!

$$Pr(\theta|y) \propto Pr(y|\theta)Pr(\theta)$$

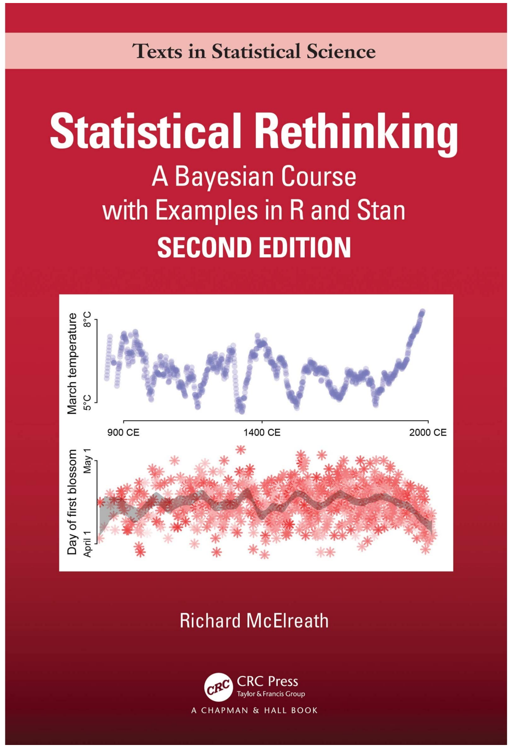
An example

'What proportion of the Earth is covered with water?'



An example

'What proportion of the Earth is covered with water?'



$$W \sim Binomial(N, p)$$

Observed quantities

$$W \sim Binomial(N, p)$$

Observed quantities

'Distributed as'

$$W \sim Binomial(N, p)$$

Observed quantities

‘Distributed as’

Likelihood

$$W \sim Binomial(N, p)$$

Observed quantities

'Distributed as'

Likelihood

$$W \sim Binomial(N, p)$$

$$p \sim Uniform(0, 1)$$

Observed quantities

'Distributed as'

Likelihood

$$W \sim Binomial(N, p)$$

Unobserved quantity (proportion water)

$$p \sim Uniform(0, 1)$$

Prior

$$Pr(W|N,p) = Binomial(W|N,p)$$
$$Pr(p) = Uniform(0,1)$$

Posterior

$$Pr(W|N,p) = Binomial(W|N,p)$$

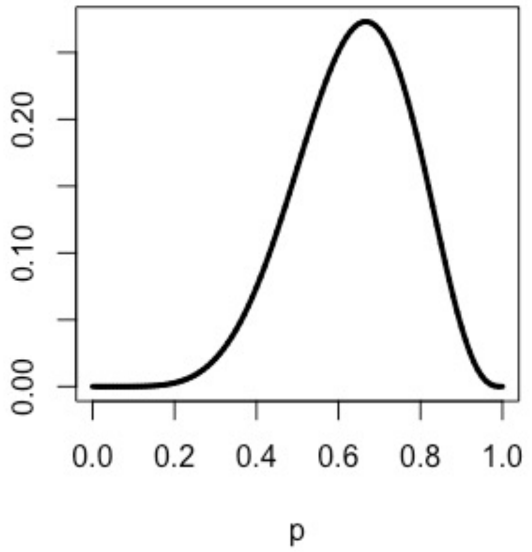
$$Pr(p) = Uniform(0,1)$$

$$Pr(p|W,N) \propto Binomial(W|N,p)Uniform(p)$$

```
W <- 6
N <- 9
pp <- seq(0, 1, length.out = 10000)
```

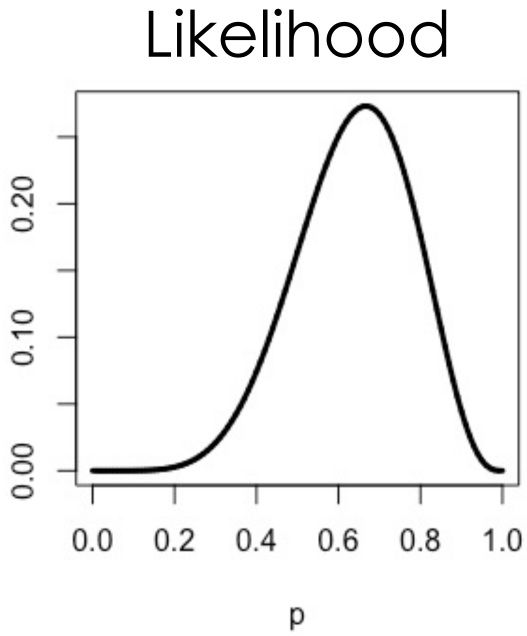
2-globe.R

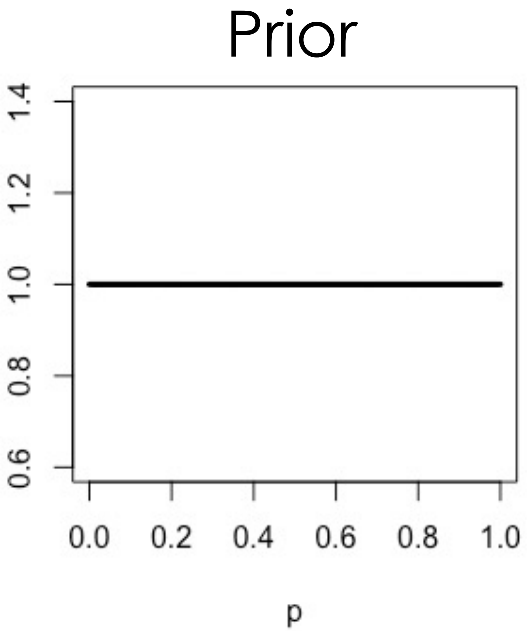
Likelihood




```r
W <- 6
N <- 9
pp <- seq(0, 1, length.out = 10000)
PrW <- dbinom(W,N,pp) #likelihood
```

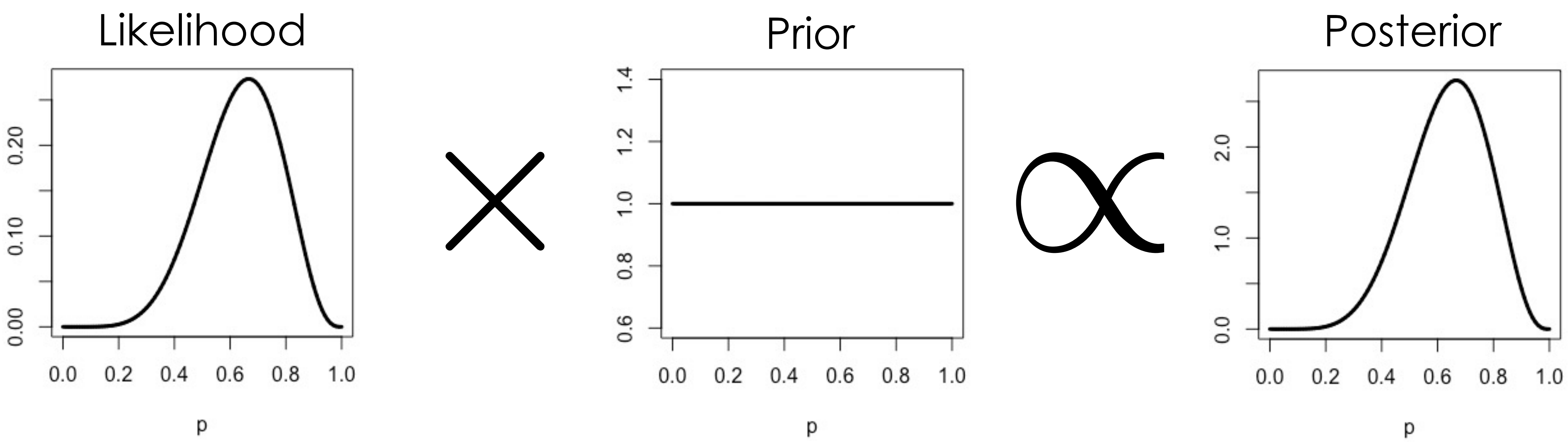
2-globe.R





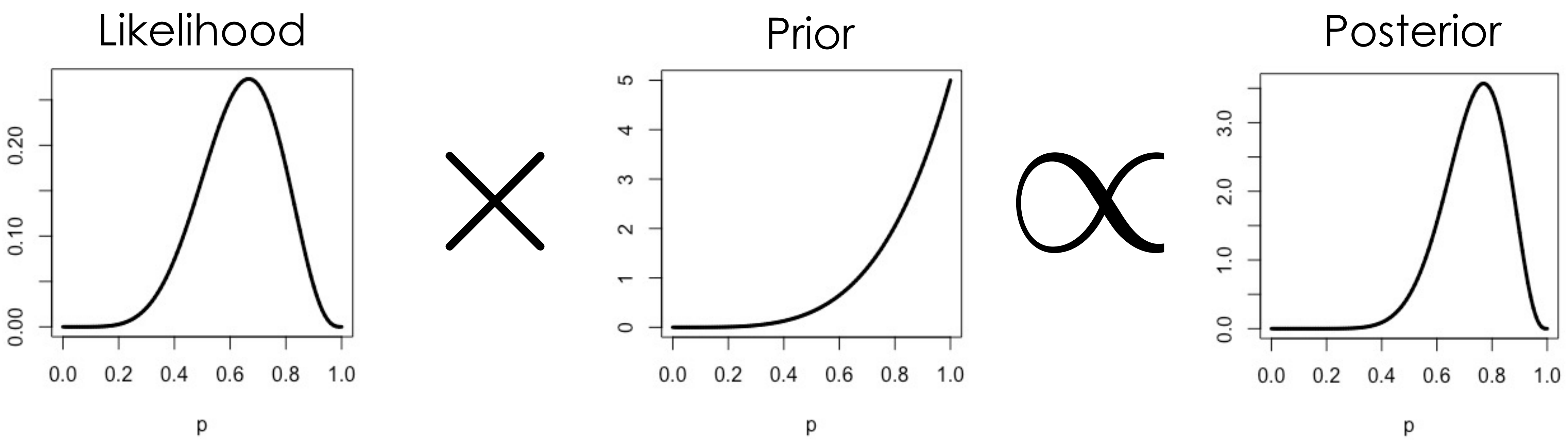

```
W <- 6
N <- 9
pp <- seq(0, 1, length.out = 10000)
PrW <- dbinom(W,N,pp) #likelihood
Prp <- dunif(pp,0,1) #prior
```

2-globe.R




```r
W <- 6
N <- 9
pp <- seq(0, 1, length.out = 10000)
PrW <- dbinom(W,N,pp) #likelihood
Prp <- dunif(pp,0,1) #prior
posterior_us <- PrW * Prp #posterior
posterior <- posterior_us / (diff(pp)[1] * sum(posterior_us))
```

2-globe.R




```r
W <- 6
N <- 9
pp <- seq(0, 1, length.out = 10000)
PrW <- dbinom(W,N,pp) #likelihood
Prp <- dbeta(pp,5,1) #prior
posterior_us <- PrW * Prp #posterior
posterior <- posterior_us / (diff(pp)[1] * sum(posterior_us))
```

2-globe.R

MCMC = **M**arkov **C**hain **M**onte **C**arlo

MCMC = **M**arkov **C**hain **M**onte **C**arlo

- Sample from probability distribution without knowing its true value

$$Pr(\theta|y)$$

MCMC = **M**arkov **C**hain **M**onte **C**arlo

- Sample from probability distribution without knowing its true value
- 'Wanders around' the distribution, spending more time in areas of high probability

MCMC Visualization

https://arogozhnikov.github.io/2016/12/19/markov_chain_monte_carlo.html