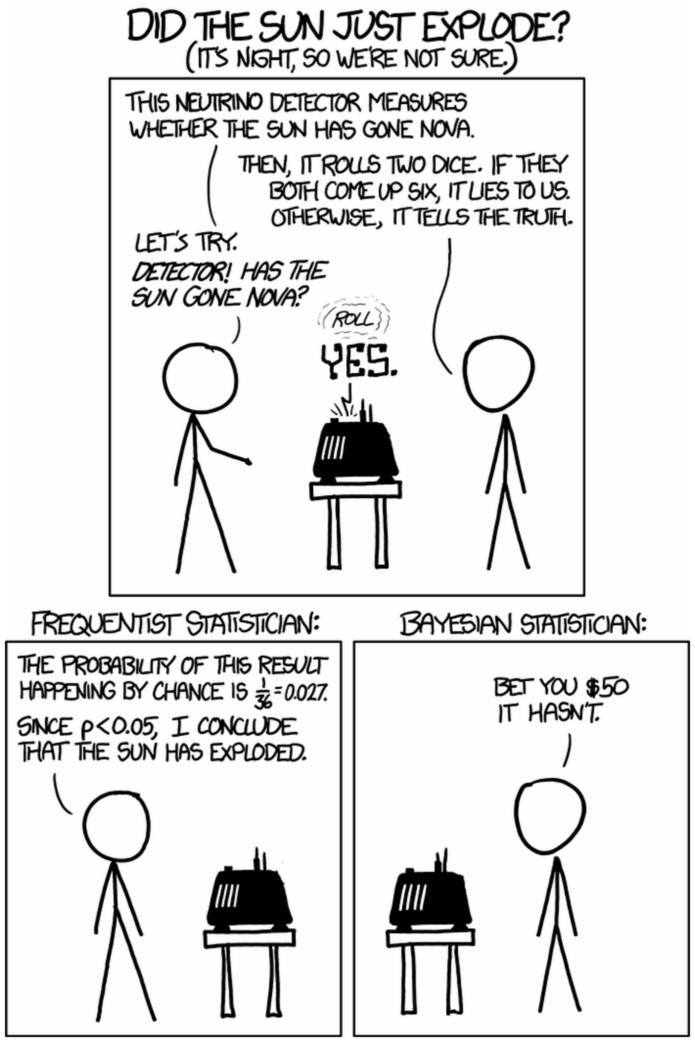
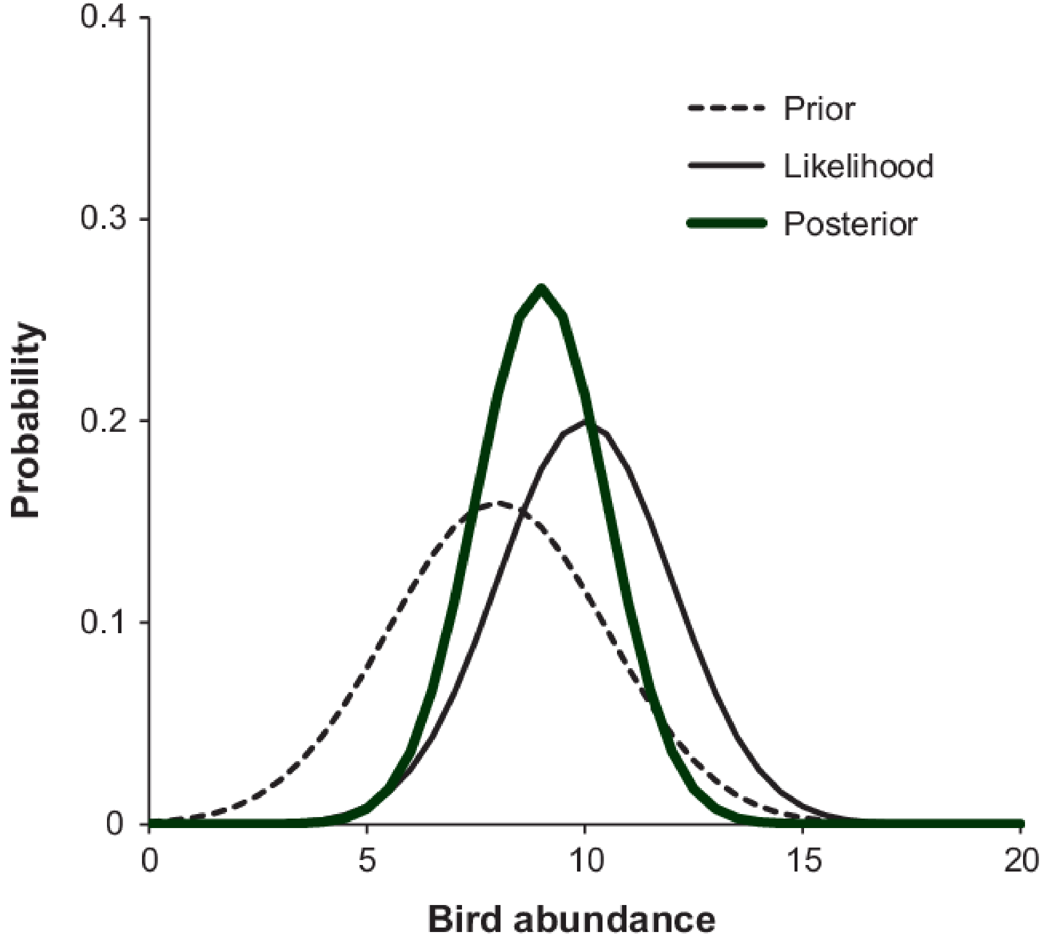
Bayesian statistics and priors





No class next week

Office hours (Stone 307A):
MF 0930-1100
W 1100-1300

We’re going to talk about a LOT of stuff today!

We won’t master it all (I can’t explain it all perfectly off the cuff)

The key idea is to BEGIN to think about how priors and our data interact to inform our final estimate in a Bayesian analysis

These are concepts we’ll revisit over and over again through the semester

Outline

1. **BAYES THEOREM**
   - *The technical backbone*

2. **STARTING SIMPLE**
   - *What is a prior? silly examples!*
   - *A back of the envelope wildlife example... beta priors*

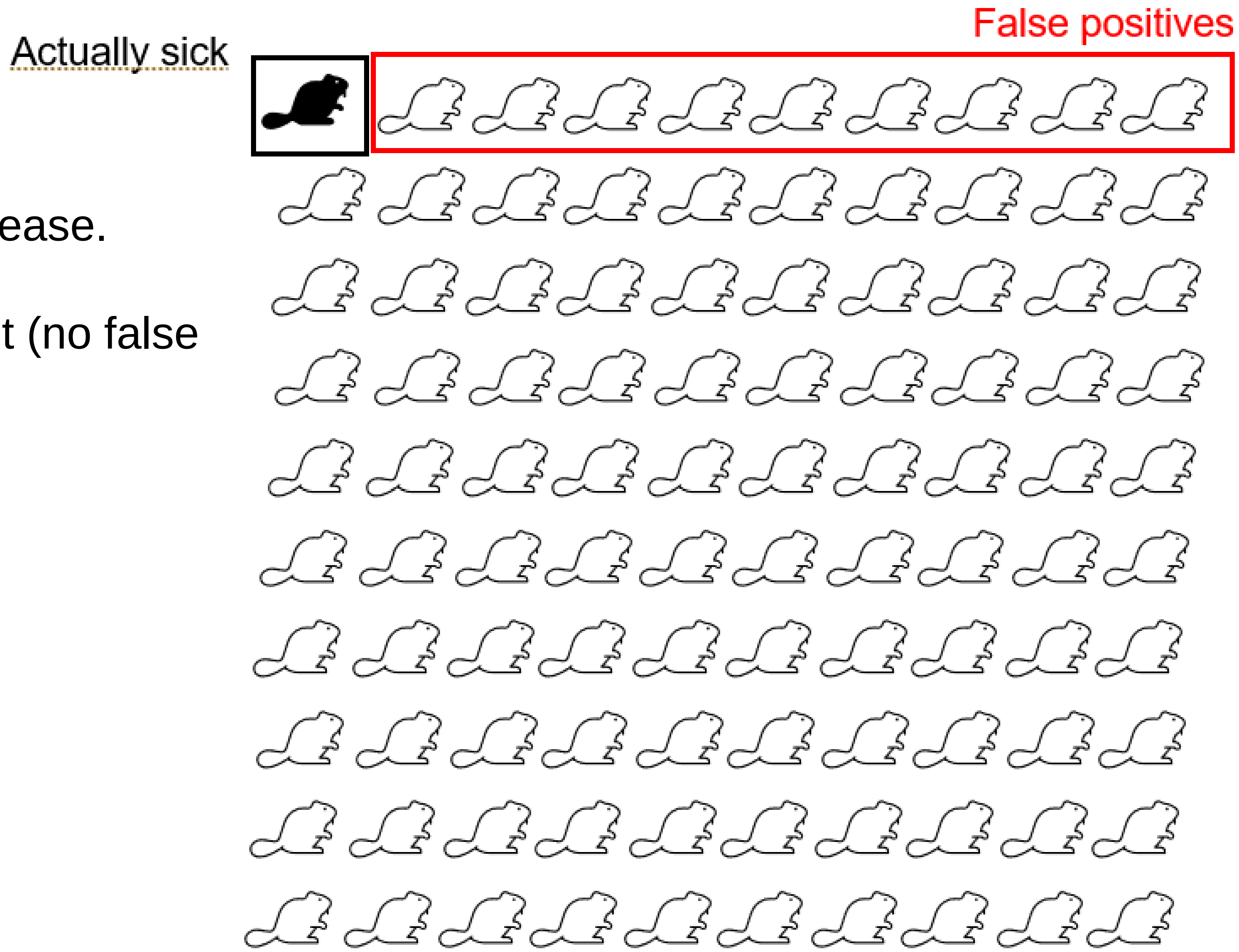
3. **ADDING COMPLEXITY (also in future weeks)**
   - *Key criticisms, and why they lead to problems*
   - *Unexpectedly bad priors (Northrup & Gerber, 2018) and philosophy*

BAYES THEOREM

$$P(A|B) = \frac{P(B|A) \times P(A)}{P(B)}$$

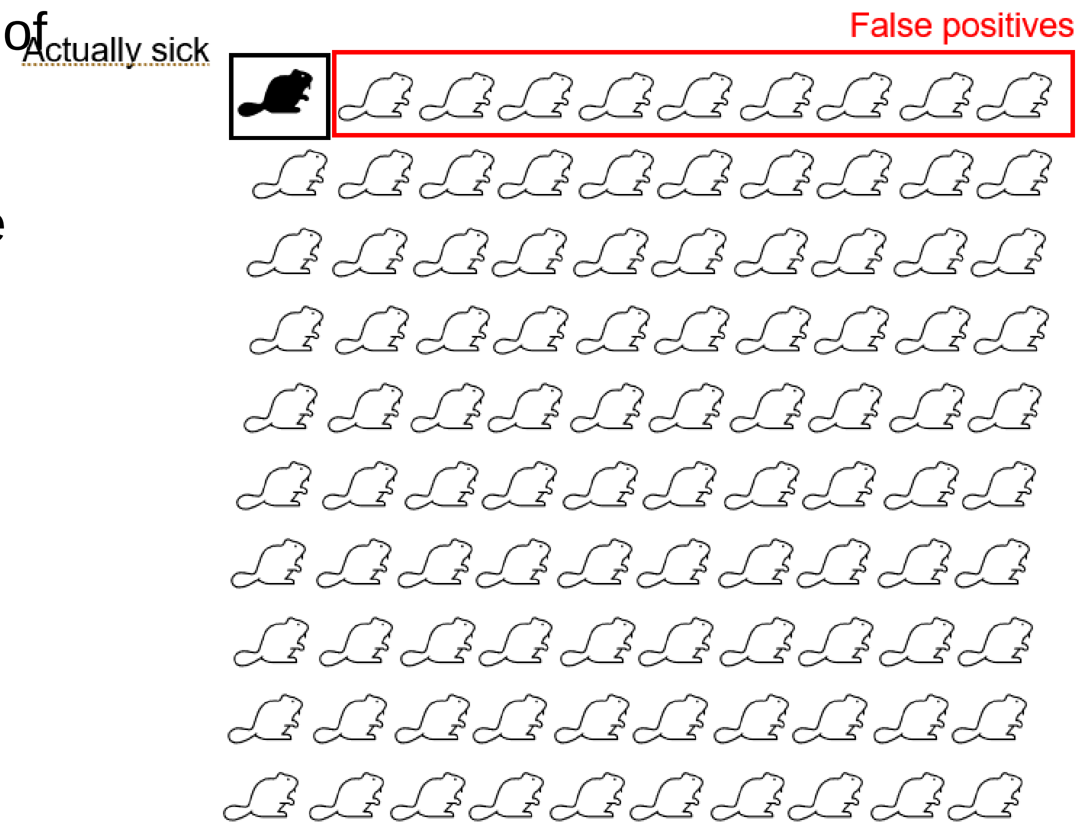
BAYES THEOREM: a quick example from disease ecology

- Imagine we test 100 beavers for a disease.
- If the disease is present, we observe it (no false negatives!)
- There is an 9% false positive rate…



BAYES THEOREM: a quick example from disease ecology

- Imagine we test 100 beavers for a disease. 1% of beavers are sick.
- If the disease is present, we observe it (no false negatives!)
- There is an ~9% false positive rate…

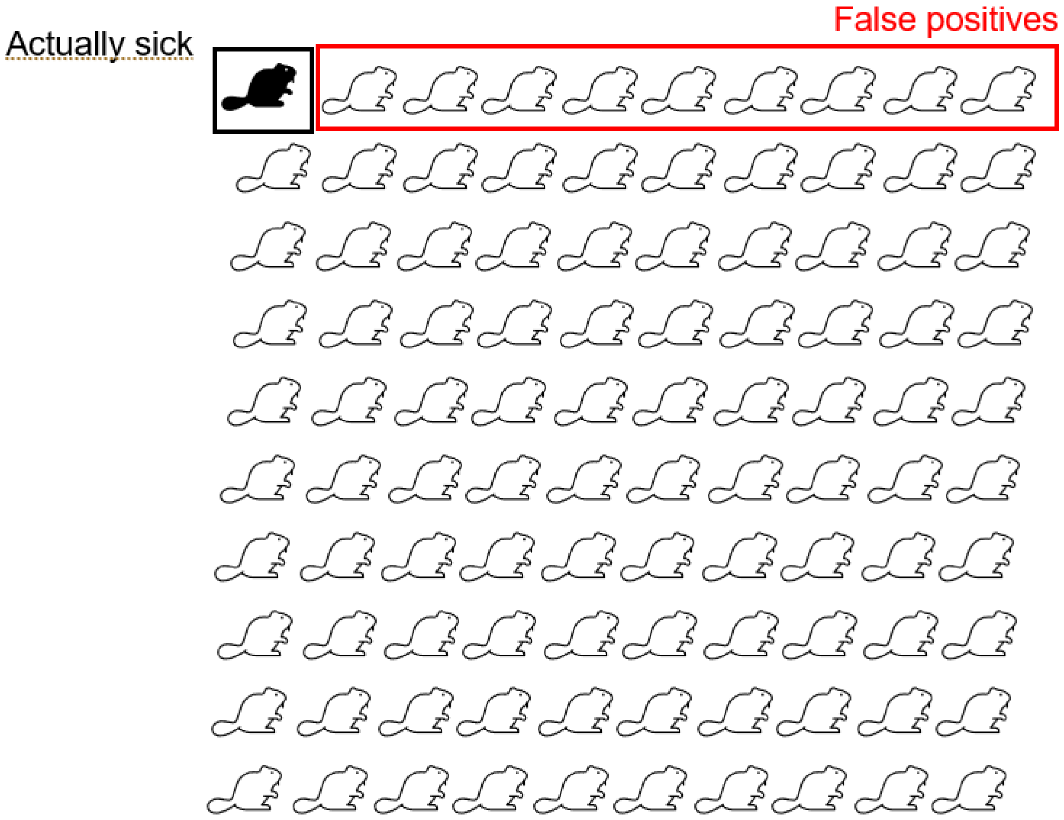


$$P(Disease|+) = \frac{P(+|Disease) \times P(Disease)}{P(+)}$$

BAYES THEOREM: a quick example from disease ecology

- Imagine we test 100 beavers for a disease. 1% of beavers are sick.
- If the disease is present, we observe it (no false negatives!)
- There is an ~9% false positive rate…

$$P(D|+) = \frac{P(+|D) \times P(D)}{P(+)}$$



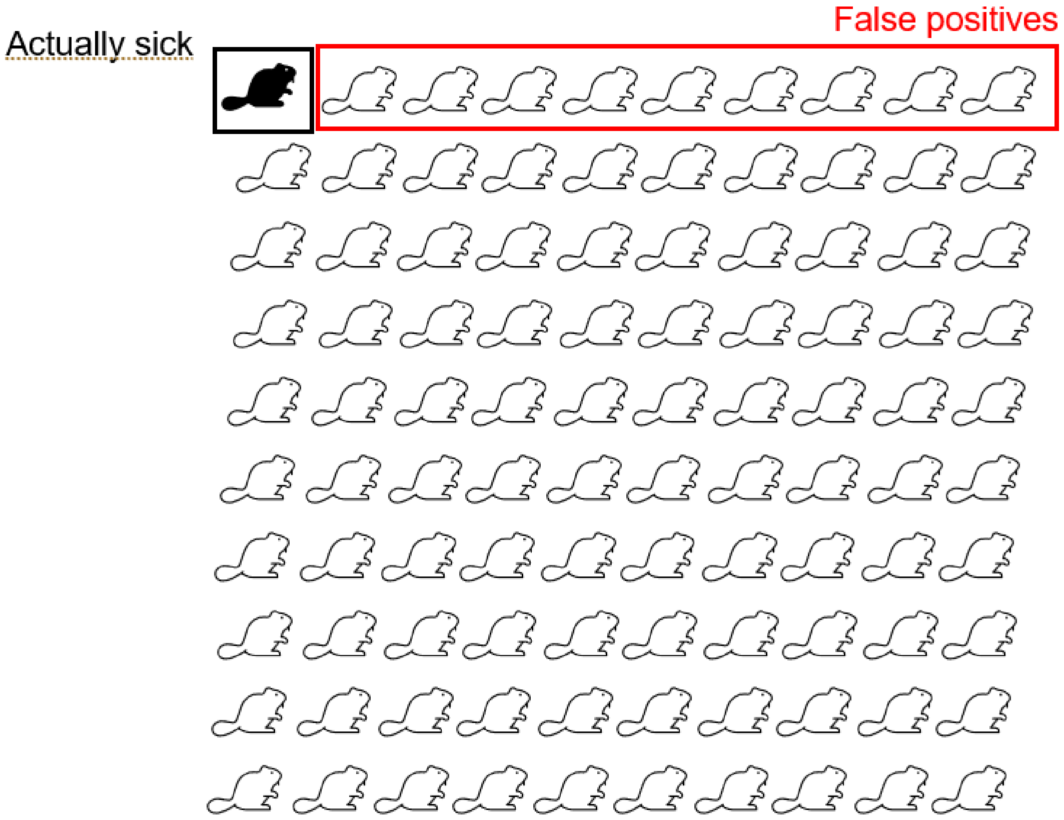
$$P(D|+) \approx 0.1$$

BAYES THEOREM: a quick example from disease ecology

- Imagine we test 100 beavers for a disease. 1% of beavers are sick.
- If the disease is present, we observe it (no false negatives!)
- There is an ~9% false positive rate…

$$P(D|+) = \frac{P(+|D) \times P(D)}{P(+)}$$

$$P(D|+) = \frac{1 \times 0.01}{(1 * 0.01) + (0.09 * .99)}$$



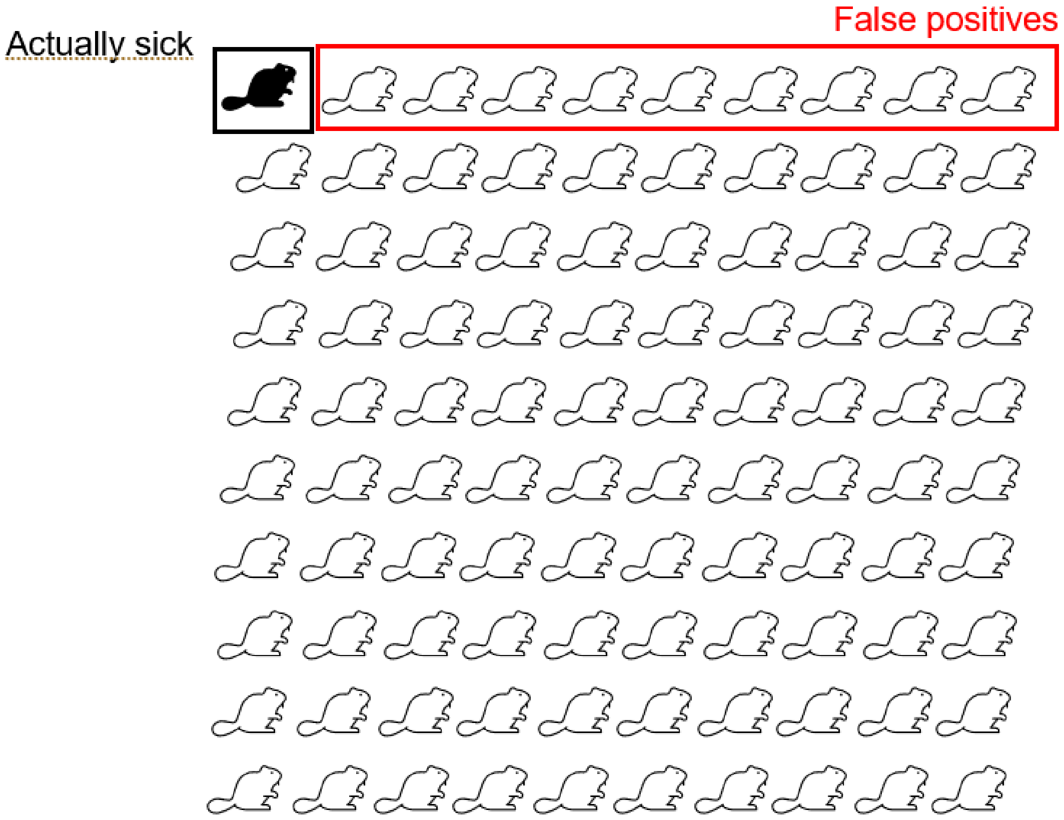
$$P(D|+) \approx 0.1$$

BAYES THEOREM: a quick example from disease ecology

- Imagine we test 100 beavers for a disease. 1% of beavers are sick.
- If the disease is present, we observe it (no false negatives!)
- There is an ~9% false positive rate…

$$P(D|+) = \frac{P(+|D) \times P(D)}{P(+)}$$

$$P(D|+) = \frac{1 \times 0.01}{(1 * 0.01) + (0.09 * .99)}$$



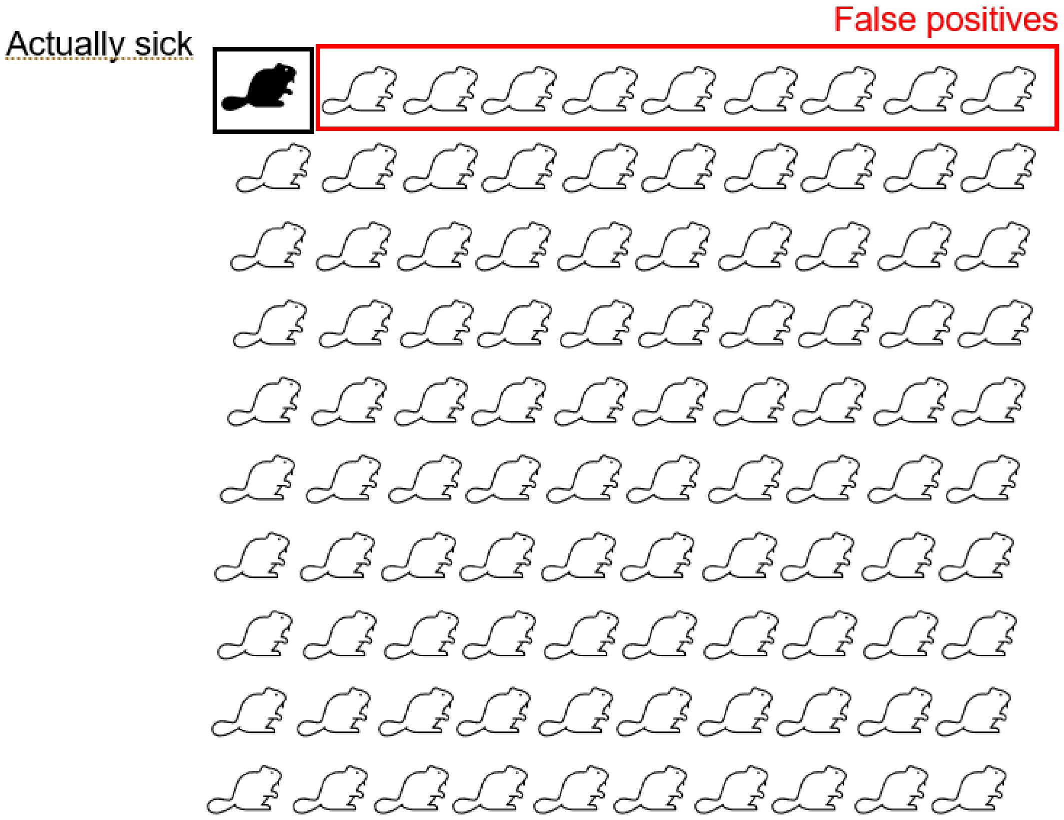
$$P(D|+) \approx 0.1$$

BAYES THEOREM: a quick example from disease ecology

- Imagine we test 100 beavers for a disease. 1% of beavers are sick.
- If the disease is present, we observe it (no false negatives!)
- There is an ~9% false positive rate…

$$P(D|+) = \frac{P(+|D) \times P(D)}{P(+)}$$

$$P(D|+) = \frac{1 \times 0.01}{(1 * 0.01) + (0.09 * .99)}$$



$$P(D|+) \approx 0.1$$

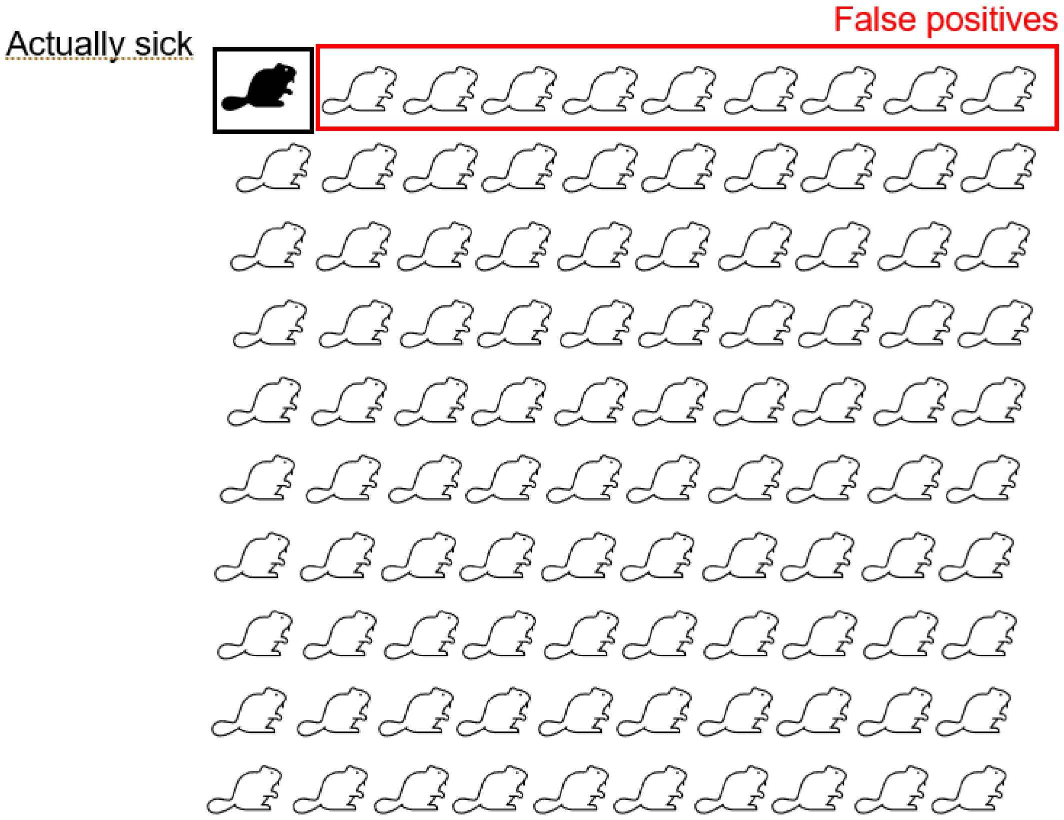
BAYES THEOREM: a quick example from disease ecology

- Imagine we test 100 beavers for a disease. 1% of beavers are sick.
- If the disease is present, we observe it (no false negatives!)
- There is an ~9% false positive rate…

$$P(D|+) = \frac{P(+|D) \times P(D)}{P(+)}$$

$$P(D|+) = \frac{1 \times 0.01}{(1 * 0.01) + (0.09 * .99)}$$

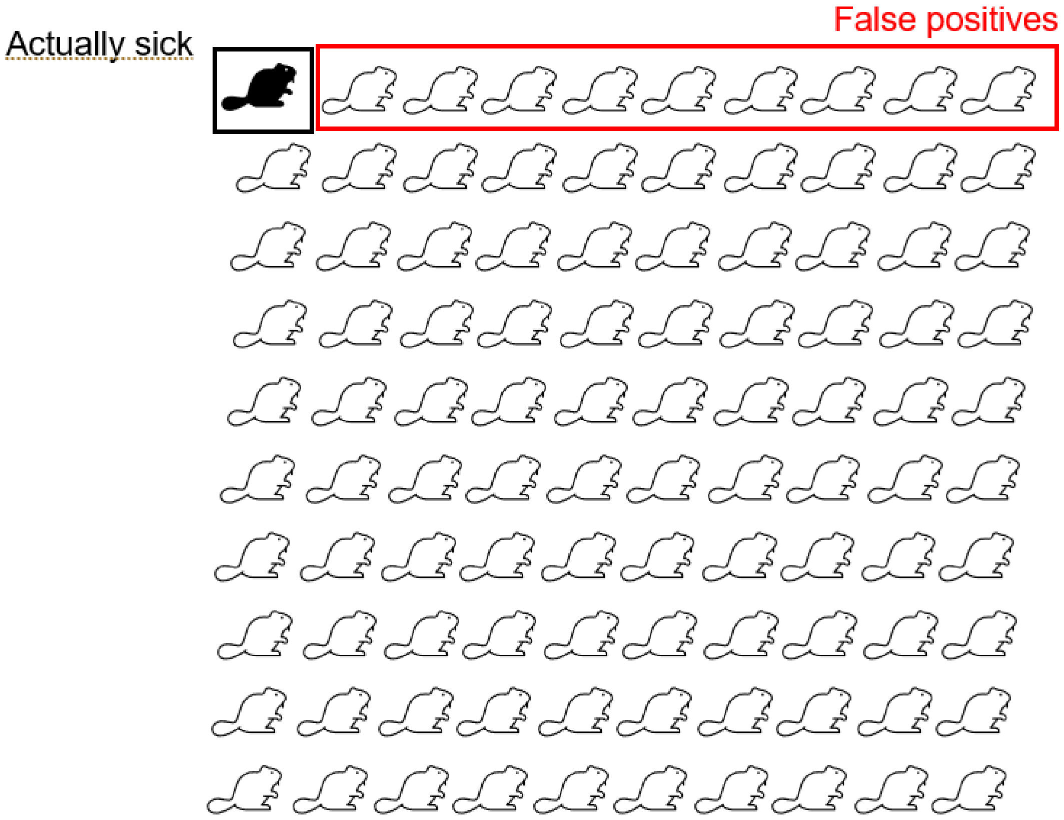
$$P(D|+) = \frac{0.01}{0.01 + 0.09}$$



$$P(D|+) \approx 0.1$$

BAYES THEOREM: a quick example from disease ecology

- Imagine we test 100 beavers for a disease. 1% of beavers are sick.
- If the disease is present, we observe it (no false negatives!)
- There is an ~9% false positive rate…



$$P(D|+) = \frac{P(+|D) \times P(D)}{P(+)}$$

$$P(D|+) = \frac{1 \times 0.01}{(1 * 0.01) + (0.09 * .99)}$$

$$P(D|+) = \frac{0.01}{0.01 + 0.09}$$

$$P(D|+) \approx 0.1$$

BAYES THEOREM

$$P(\boldsymbol{\theta}|y) = \frac{P(y|\boldsymbol{\theta}) \times P(\boldsymbol{\theta})}{P(y)}$$

BAYES THEOREM

$$P(\theta|y) = \frac{P(y|\theta) \times P(\theta)}{P(y)}$$

θ: parameter(s) of interest
y: data

BAYES THEOREM

Likelihood [these are the estimates you'd get from lm()]

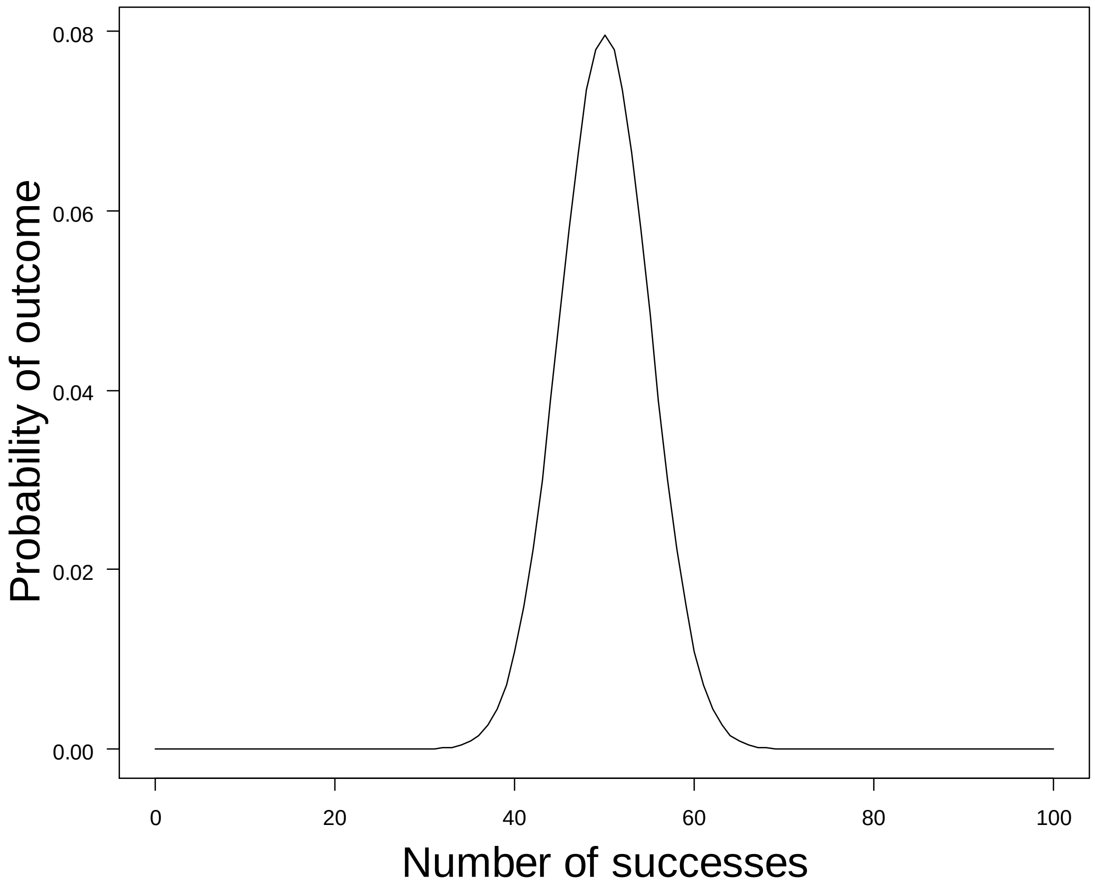
$$P(\theta|y) = \frac{\boxed{P(y|\theta)} \times P(\theta)}{P(y)}$$

θ: parameter(s) of interest

y: data

What is a likelihood?

- Probability for getting a result for a given value of parameter(s)



Flipping a coin 100 times with 50:50 chance of heads

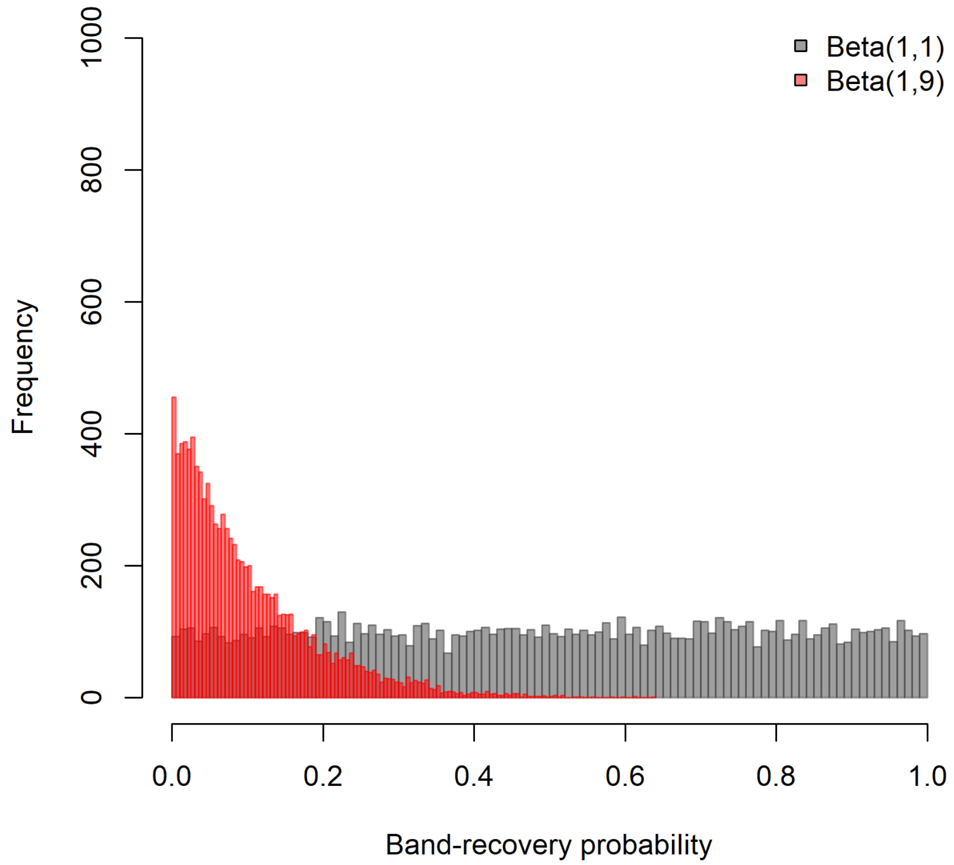
BAYES THEOREM

$$P(\theta|y) = \frac{\overset{\text{Likelihood}}{\boxed{P(y|\theta)}} \times \overset{\text{Prior}}{\boxed{P(\theta)}}}{P(y)}$$

θ: parameter of interest
y: data

What is a prior?

- Your prior belief in the distribution of a parameter.



BAYES THEOREM

Likelihood Prior

$$P(\theta|y) = \frac{\boxed{P(y|\theta)} \times \boxed{P(\theta)}}{\boxed{P(y)}}$$

Marginal

θ: parameter of interest
y: data

What is marginal?

- The marginal probability of the data!

- i.e., the sum of the probability of the different ways you could observe the data

BAYES THEOREM

$$\underbrace{\boxed{P(\theta|y)}}_{\text{Posterior}} = \frac{\overbrace{\boxed{P(y|\theta)}}^{\text{Likelihood}} \times \overbrace{\boxed{P(\theta)}}^{\text{Prior}}}{\underbrace{\boxed{P(y)}}_{\text{Marginal}}}$$

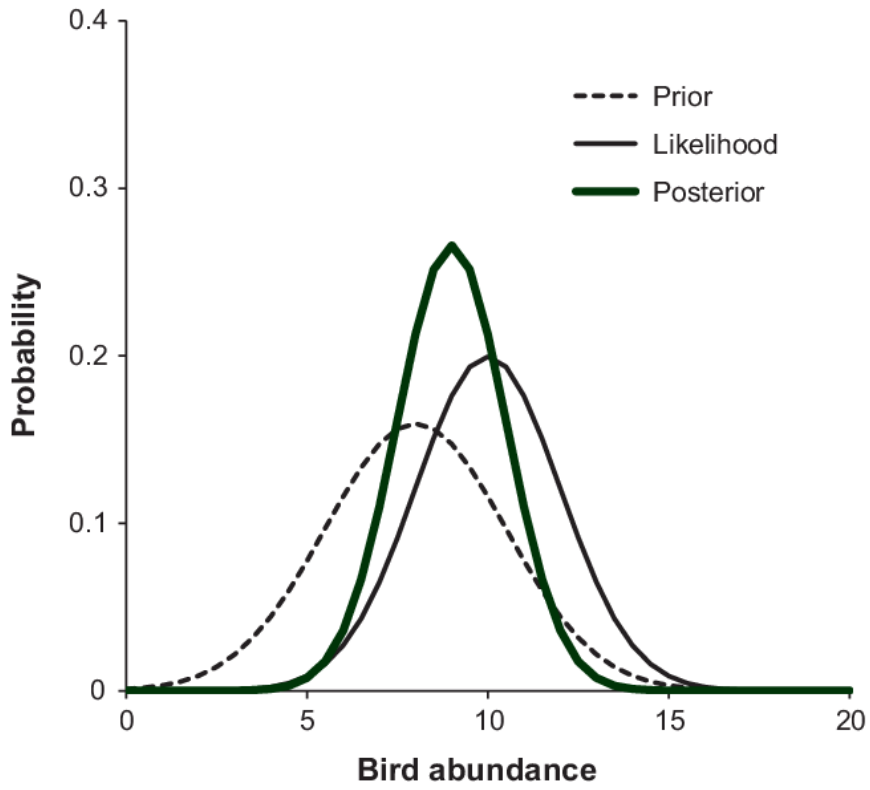
θ: parameter of interest

y: data

What is a posterior?

- The distribution of a parameter given your prior belief and the data.

Our estimate



BAYES THEOREM: let's work through some fun, simple e.g.'s

The case of the credulous crackpot

$$P(\theta|y) = \frac{P(y|\theta) \times P(\theta)}{P(y)}$$

BAYES THEOREM: let's work through a silly, intuitive example

$$P(\text{bigfoot}|+) = \frac{P(+|\text{bigfoot}) \times P(\text{bigfoot})}{P(+)}$$

Figure 1 Map of Bigfoot encounters from Washington, Oregon and California used in the analyses. Points represent visual/auditory detection, and foot symbols represent coordinates where footprint data were available. Shading indicates topography, with lighter values representing lower elevations.

Lozier **(2009)** *Journal of Biogeography*

$$P(\theta|y) = \frac{P(y|\theta) \times P(\theta)}{P(y)}$$

BAYES THEOREM: let's work through a silly, intuitive example

$$P(\text{bigfoot}|+) = \frac{P(+|\text{bigfoot}) \times \boxed{P(\text{bigfoot})}}{P(+)}$$

Figure 1 Map of Bigfoot encounters from Washington, Oregon and California used in the analyses. Points represent visual/auditory detection, and foot symbols represent coordinates where footprint data were available. Shading indicates topography, with lighter values representing lower elevations.

Lozier **(2009)** *Journal of Biogeography*

$$P(\theta|y) = \frac{P(y|\theta) \times P(\theta)}{P(y)}$$

BAYES THEOREM: let’s work through a silly, intuitive example

Prior

$$P(\text{bigfoot}|+) = \frac{P(+|\text{bigfoot}) \times \boxed{P(\text{bigfoot})}}{P(+)}$$

Figure 1 Map of Bigfoot encounters from Washington, Oregon and California used in the analyses. Points represent visual/auditory detection, and foot symbols represent coordinates where footprint data were available. Shading indicates topography, with lighter values representing lower elevations.

Prior: the probability of an event pre-data

Lozier **(2009)** *Journal of Biogeography*

$$P(\theta|y) = \frac{P(y|\theta) \times P(\theta)}{P(y)}$$

BAYES THEOREM: let's work through a silly, intuitive example

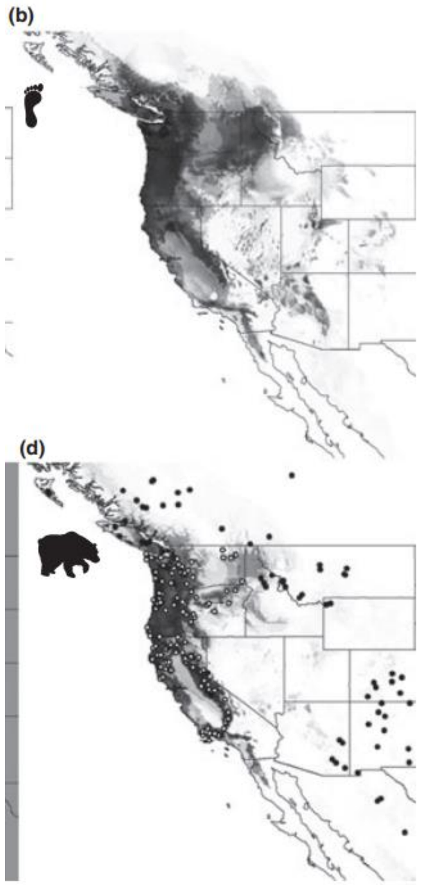
$$P(\text{bigfoot}|+) = \mathbf{0} = \frac{P(+|\text{bigfoot}) \times \mathbf{0}}{P(+)}$$

Bigfoot does not exist, zero divided by whatever = 0… analysis complete

$$P(\theta|y) = \frac{P(y|\theta) \times P(\theta)}{P(y)}$$

BAYES THEOREM: what about the marginal?

$$P(\text{bigfoot}|+) = \frac{P(+|\text{bigfoot}) \times \boldsymbol{P(\text{bigfoot})}}{\boldsymbol{P(+)}}$$

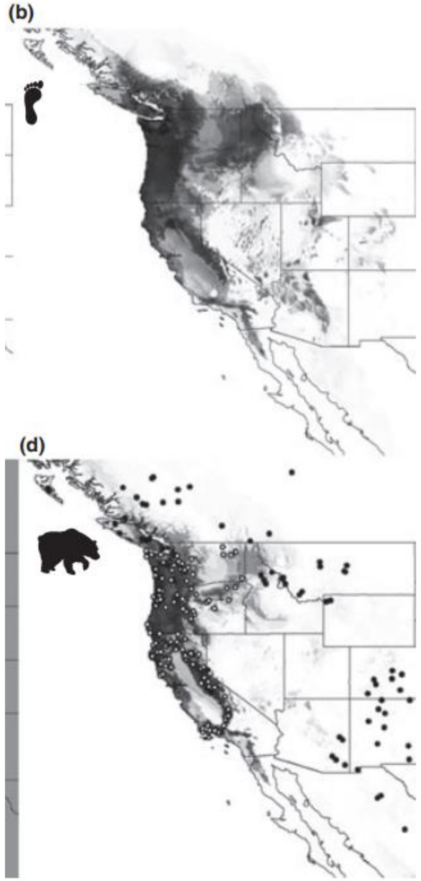


Lozier **(2009)** *Journal of Biogeography*

$$P(\theta|y) = \frac{P(y|\theta) \times P(\theta)}{P(y)}$$

BAYES THEOREM: what about the marginal?

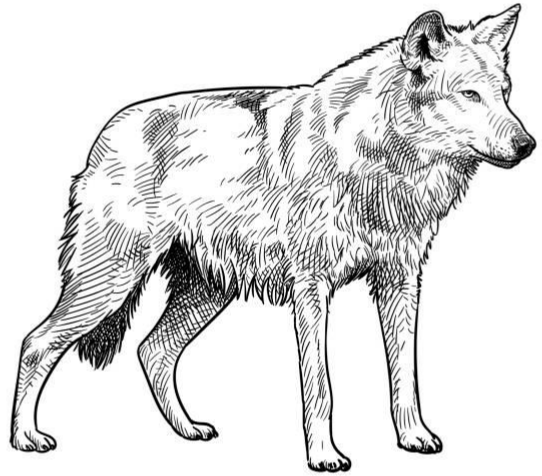
$$P(\text{bigfoot}|+) = \frac{P(+|\text{bigfoot}) \times \boldsymbol{P(\textbf{bigfoot})}}{\boldsymbol{P(+\cap \textbf{bigfoot}) + P(+\cap \textbf{bear})}}$$



Lozier **(2009)** *Journal of Biogeography*

$$P(\theta|y) = \frac{P(y|\theta) \times P(\theta)}{P(y)}$$

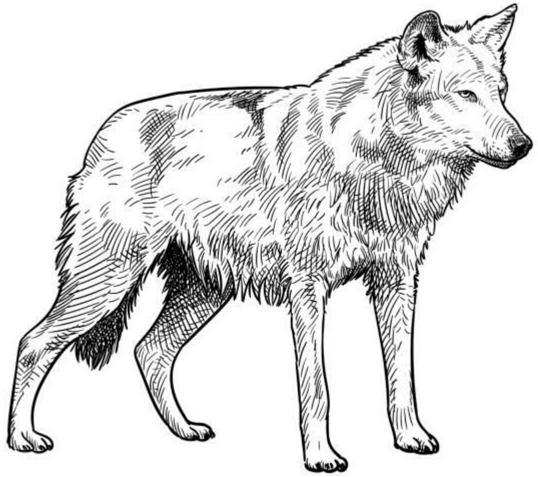
The credulous crackpot strikes again!



$$P(\theta|y) = \frac{P(y|\theta) \times P(\theta)}{P(y)}$$

BAYES THEOREM: let's work through a better example

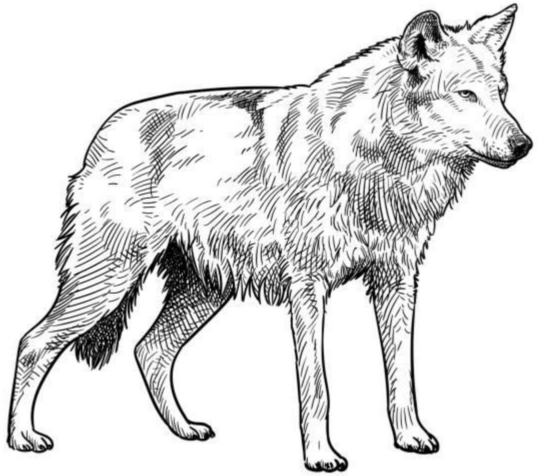
$$P(\text{wolf}|+) = \frac{P(+|\text{wolf}) \times P(\text{wolf})}{P(+)}$$



Lozier **(2009)** *Journal of Biogeography*

BAYES THEOREM: let's work through a better example

$$P(\text{wolf}|+) = \frac{P(+|\text{wolf}) \times \overbrace{\boxed{P(\text{wolf})}}^{\text{Prior}}}{\underbrace{\boxed{P(+)}}_{\text{Marginal}}}$$

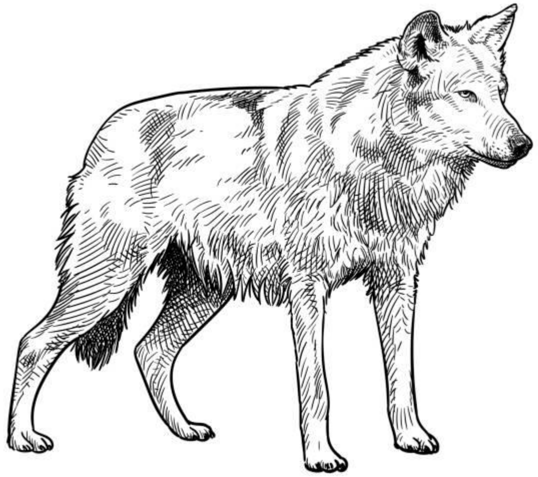


BAYES THEOREM: let's work through a better example

Prior

$$P(\text{wolf}|+) = \mathbf{0} = \frac{P(+|\text{wolf}) \times \boxed{P(\text{wolf})}}{\boxed{P(+)}}$$

Marginal



Prior: probability of the event

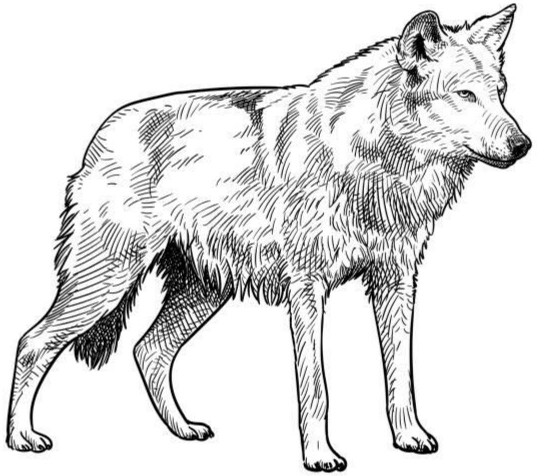
Marginal: the probability of the data…

BAYES THEOREM: let's make some assumptions

Prior

$$P(\text{wolf}|+) = \mathbf{0} = \frac{P(+|\text{wolf}) \times \boxed{P(\text{wolf})}}{\boxed{P(+\cap\ \text{wolf}) + P(+\cap\ \text{coyote})}}$$

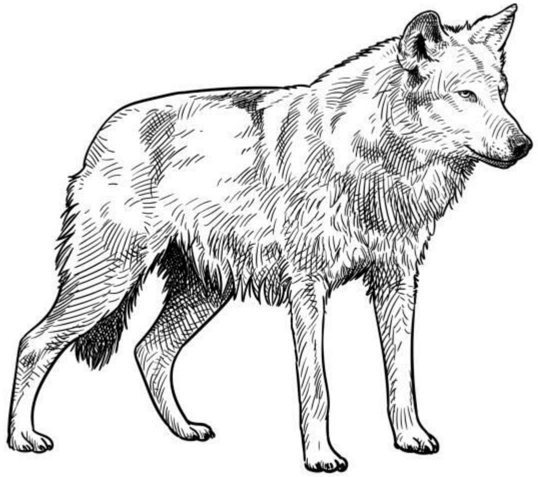
Marginal



We'll assume they actually saw a large wild canid (i.e., either coyote or wolf)

BAYES THEOREM: let’s make some assumptions

$$P(\text{wolf}|+) = \mathbf{0} = \frac{P(+|\text{wolf}) \times \overset{\text{Prior}}{\boxed{P(\text{wolf})}}}{\underset{\text{Marginal}}{\boxed{P(+\cap\ \text{wolf}) + P(+\cap\ \text{coyote})}}}$$



Prior: let’s imagine a 2:1 coyote:wolf ratio

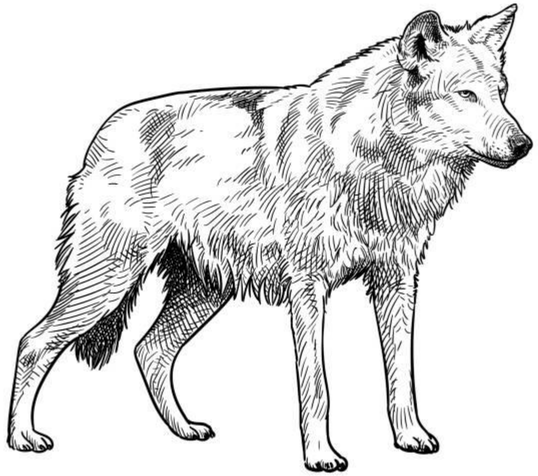
Marginal: if wolf, wolf, if coyote, 50:50

BAYES THEOREM: let's make some assumptions

Prior

$$P(\text{wolf}|+) = \mathbf{0.5} = \frac{P(+|\text{wolf}) \times \boxed{\mathbf{0.33}}}{\boxed{\mathbf{0.33 + 0.33}}}$$

Marginal



Prior: let's imagine a 2:1 coyote:wolf ratio

Marginal: if wolf -> wolf, if coyote, 50:50

BAYES THEOREM: let's make some assumptions

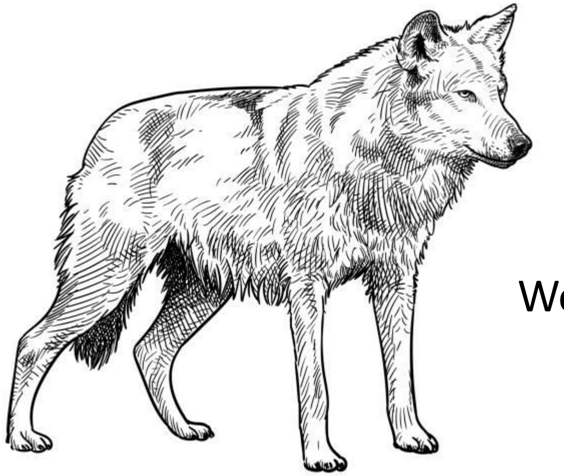
Prior

$$P(\text{wolf}|+) = \mathbf{0.5} = \frac{P(+|\text{wolf}) \times \boxed{\mathbf{0.33}}}{\boxed{\mathbf{0.33 + 0.33}}}$$

Marginal

Prior: let's imagine a 2:1 coyote:wolf ratio

Marginal: if wolf -> wolf, if coyote, 50:50

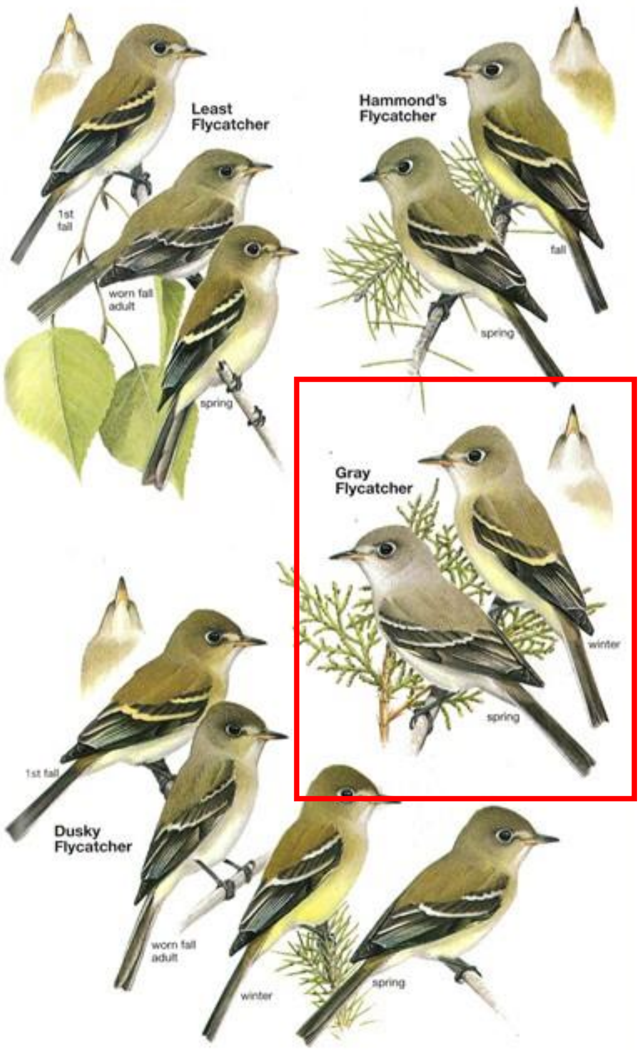


Wolf

Misidentified as wolf (coyote)

Coyote

BAYES THEOREM: a complex/realistic example (*Empid*. ppp.)



$$P(\text{gray}|+) = \frac{P(+|\text{gray}) \times P(\text{gray})}{P(+)}$$

$$P(A|B) = \frac{P(B|A) \times P(A)}{P(B)}$$

Hammond’s

Gray

Dusky

Least

I’m kidding, we’re not making up that many numbers

There are three key points here:

1. **The math works.**

There are three key points here:

1. **The math works.**

2. **This is how ‘we’ naturally think about things! We routinely (i.e., all the freaking time) use prior information!**

There are three key points here:

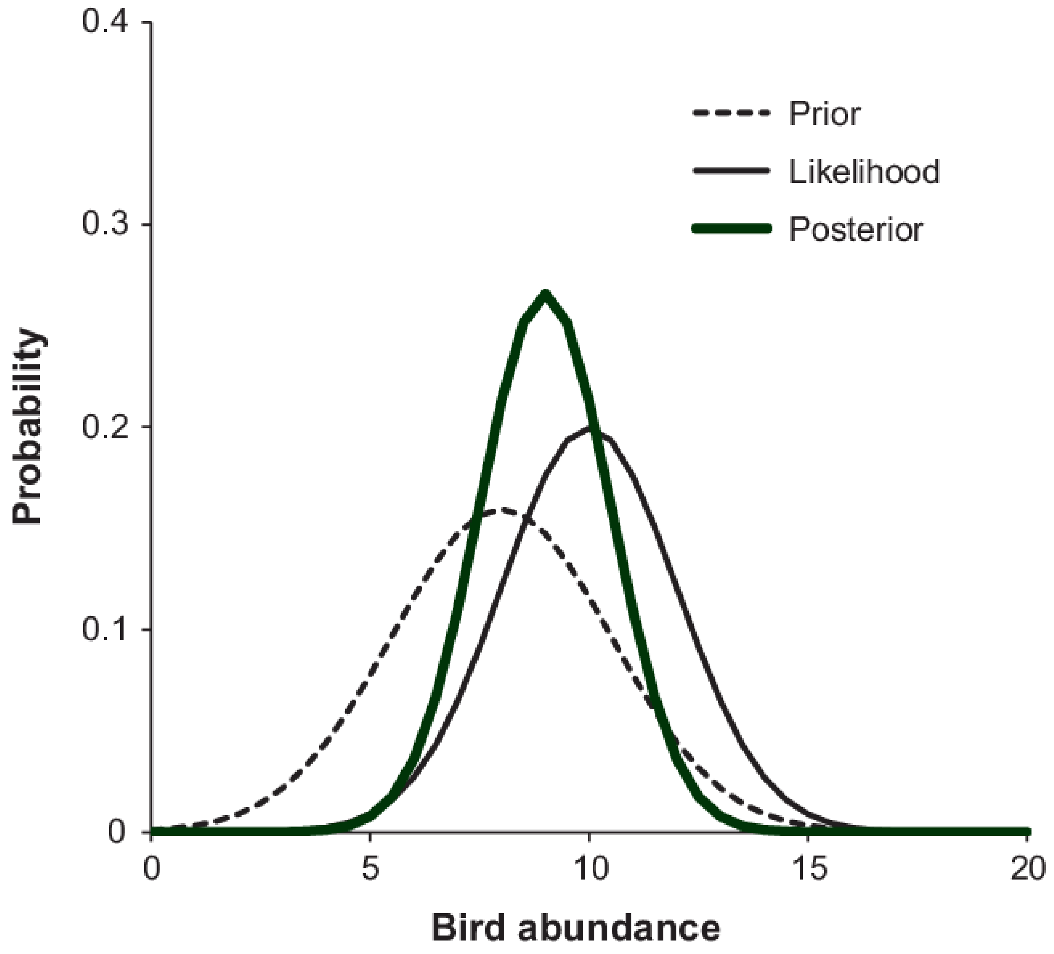
1. **The math works.**

2. **This is how ‘we’ naturally think about things! We routinely use prior information!**

3. **We’re using fixed (known) values. This starts to get really complicated with uncertainty (we’re going to need help from JAGS and Stan)!**

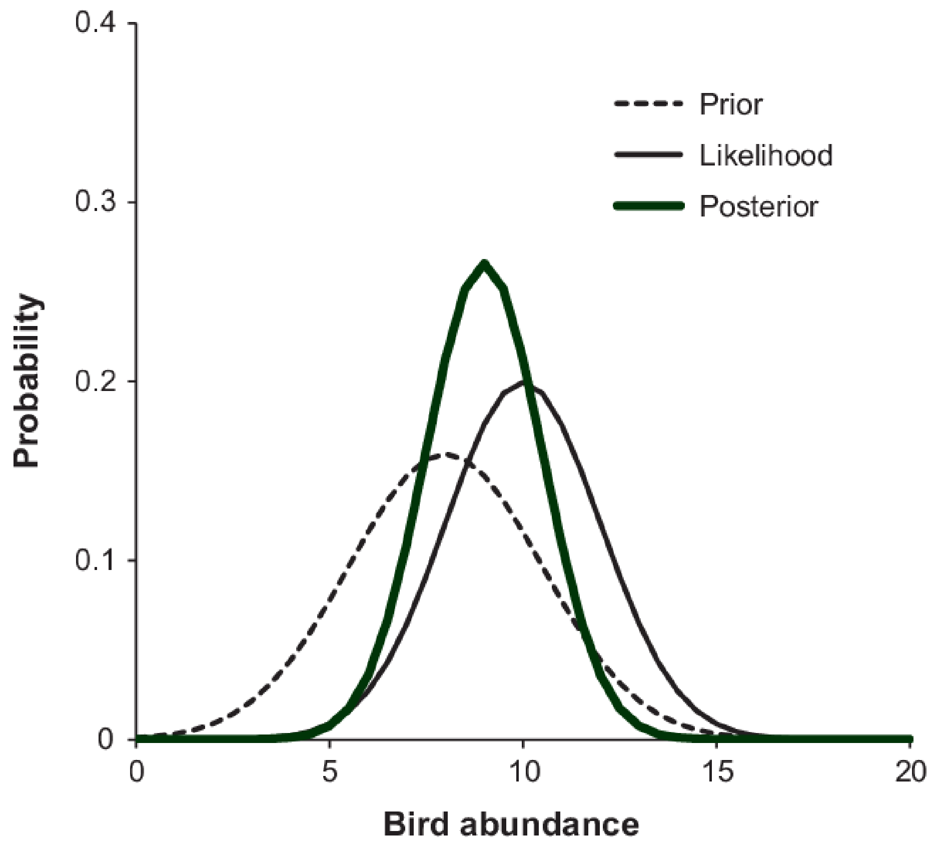
A huge non-Bayesian fear is that priors affect posteriors (inference)

- The distribution of a parameter given your prior belief and the data.



How humans learn might not be the way to go ☺

A huge non-Bayesian fear is that priors affect posteriors (inference)



Priors do affect posteriors, this is not an irrational fear

BAYES THEOREM: a response to a (the?) key criticism

We’ll use ‘uninformative’ priors for everything. These priors will not affect our inference.

Then you can’t criticize us!

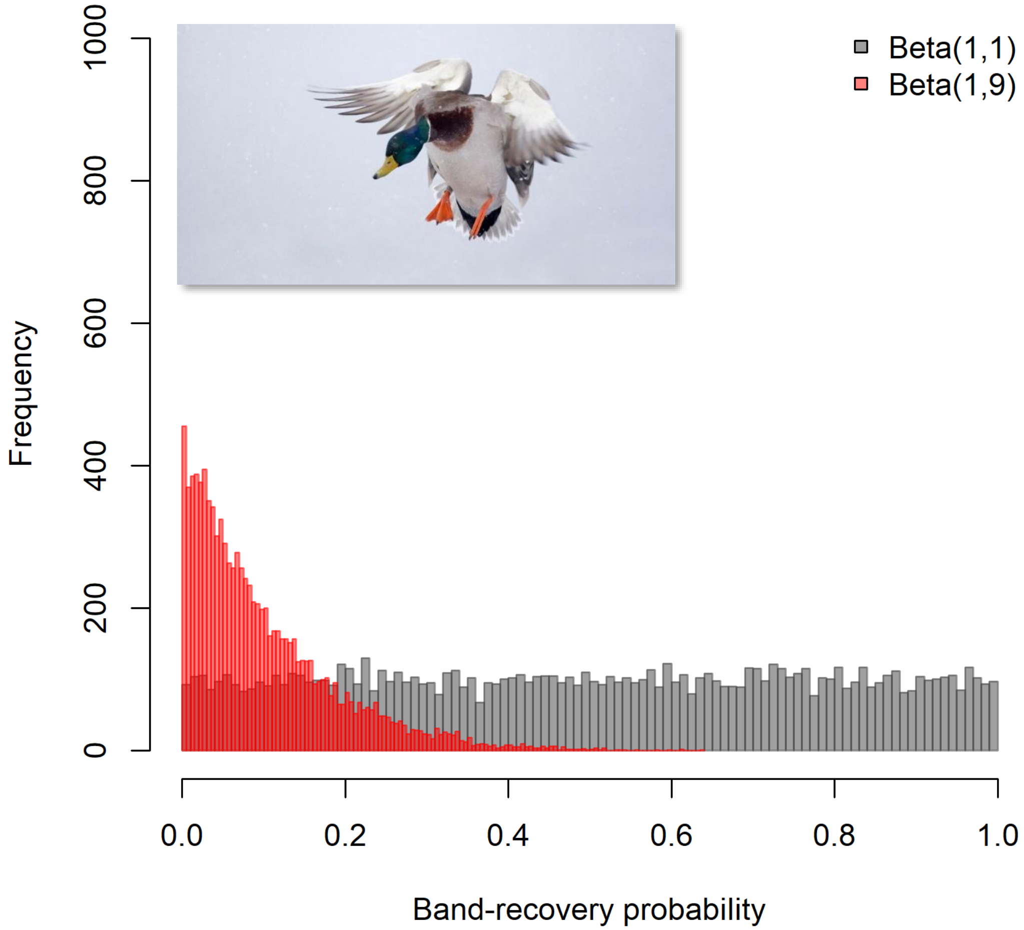
BAYES THEOREM: a response to a (the?) key criticism

We’ll use ‘uninformative’ priors for everything. These priors will not affect our inference.

Then you can’t criticize us!

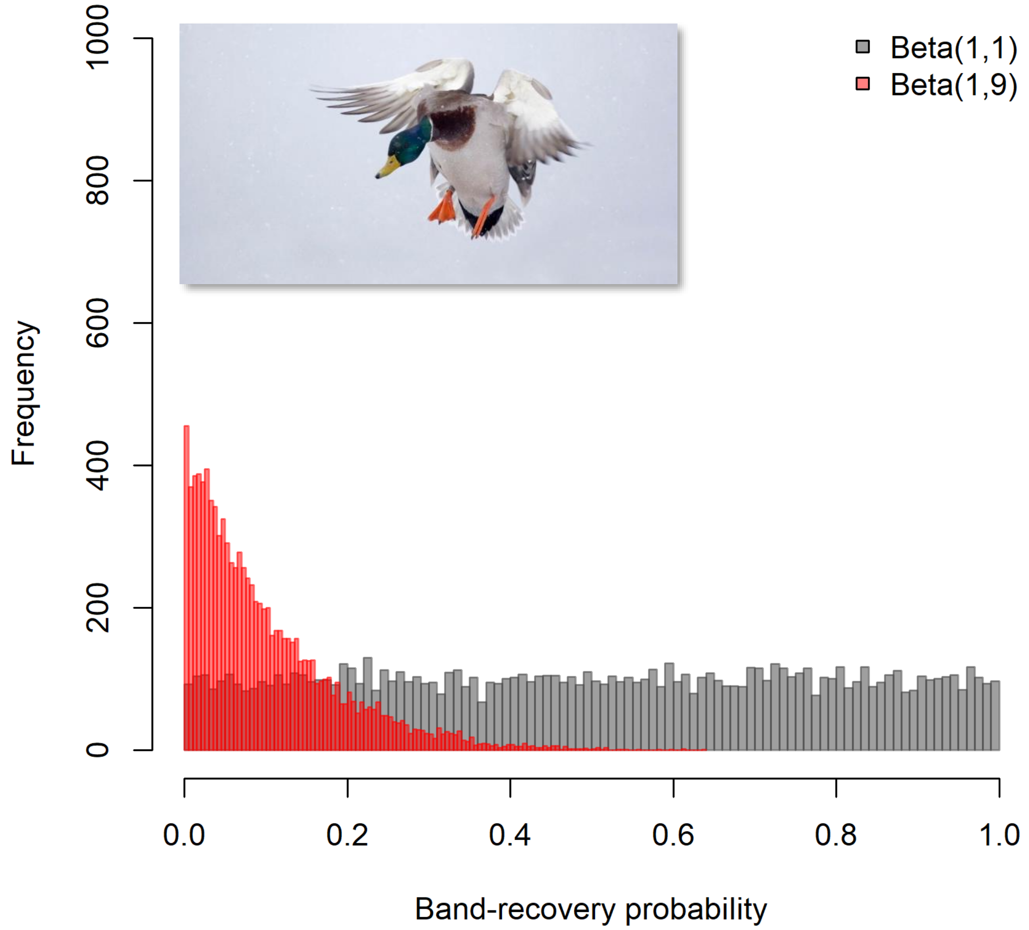
<u>This was a key mistake, and honestly it’s kind of bizarre</u>

‘Uninformative*’ priors



***Priors are <u>never ‘uninformative.’</u>**

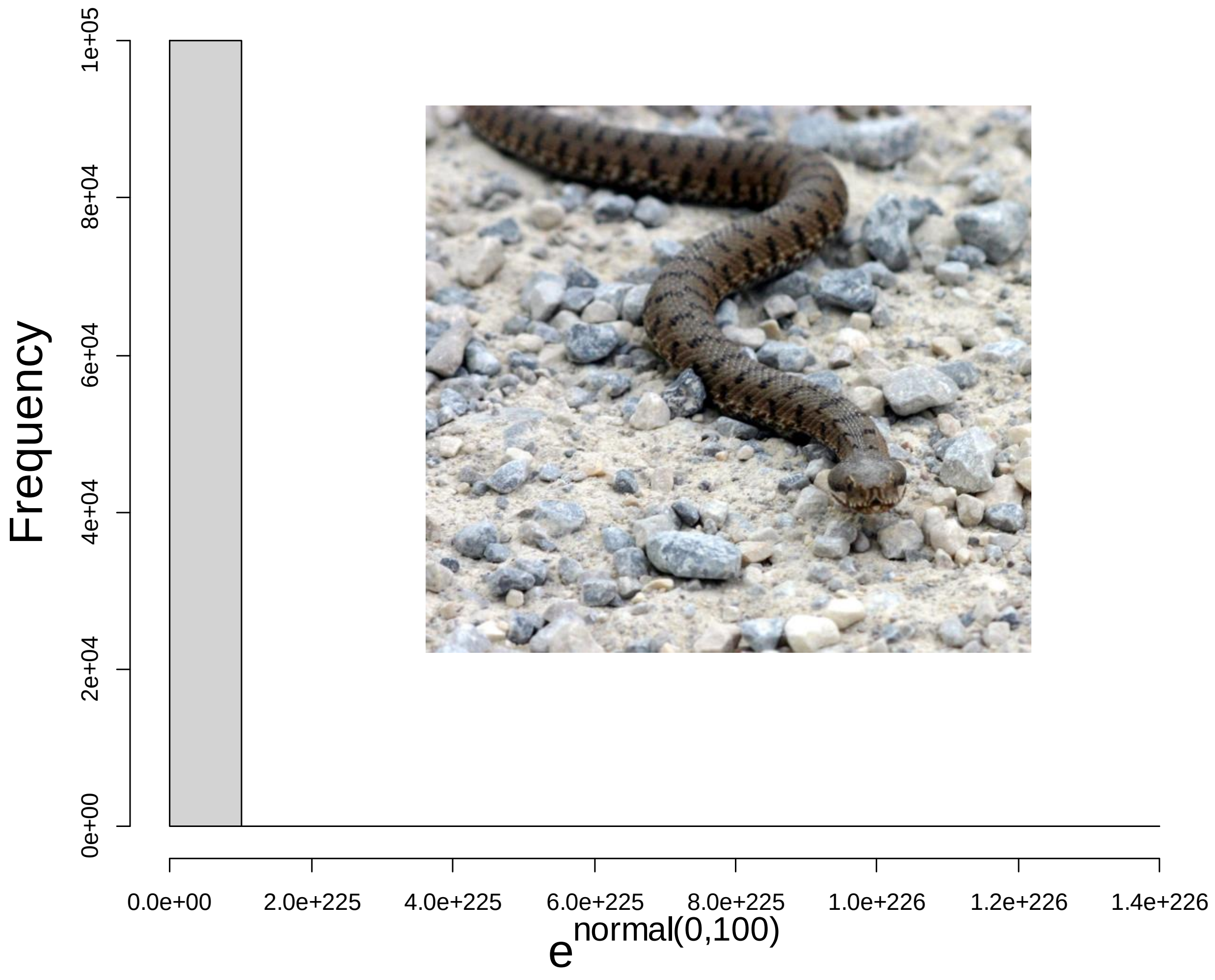
If you disagree…

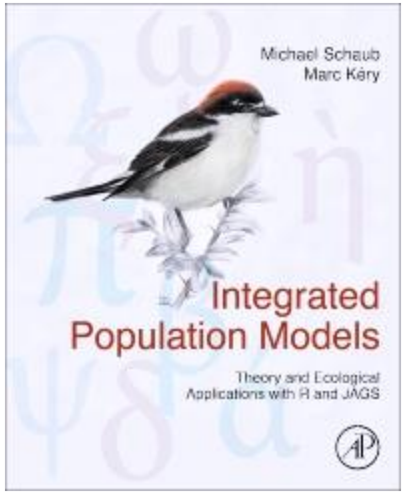


Then describe a prior without providing any information.

Sometimes priors are way too ‘uninformed’

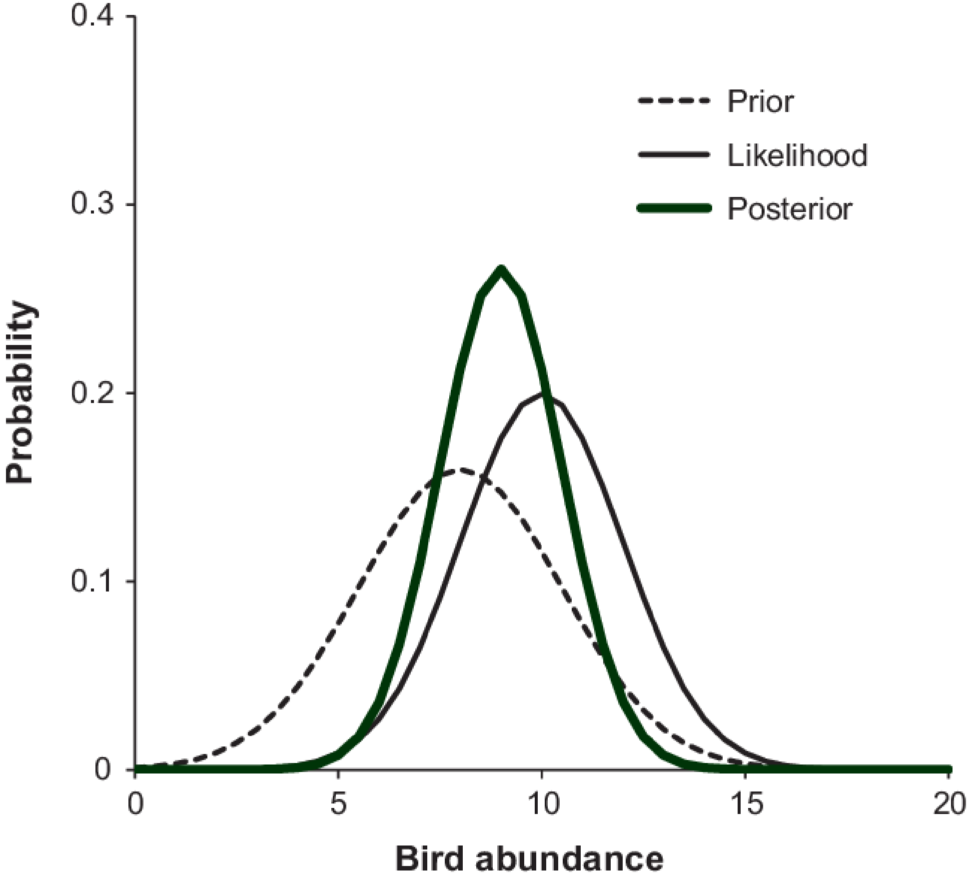
Sometimes priors are waaaay too ‘uninformed’

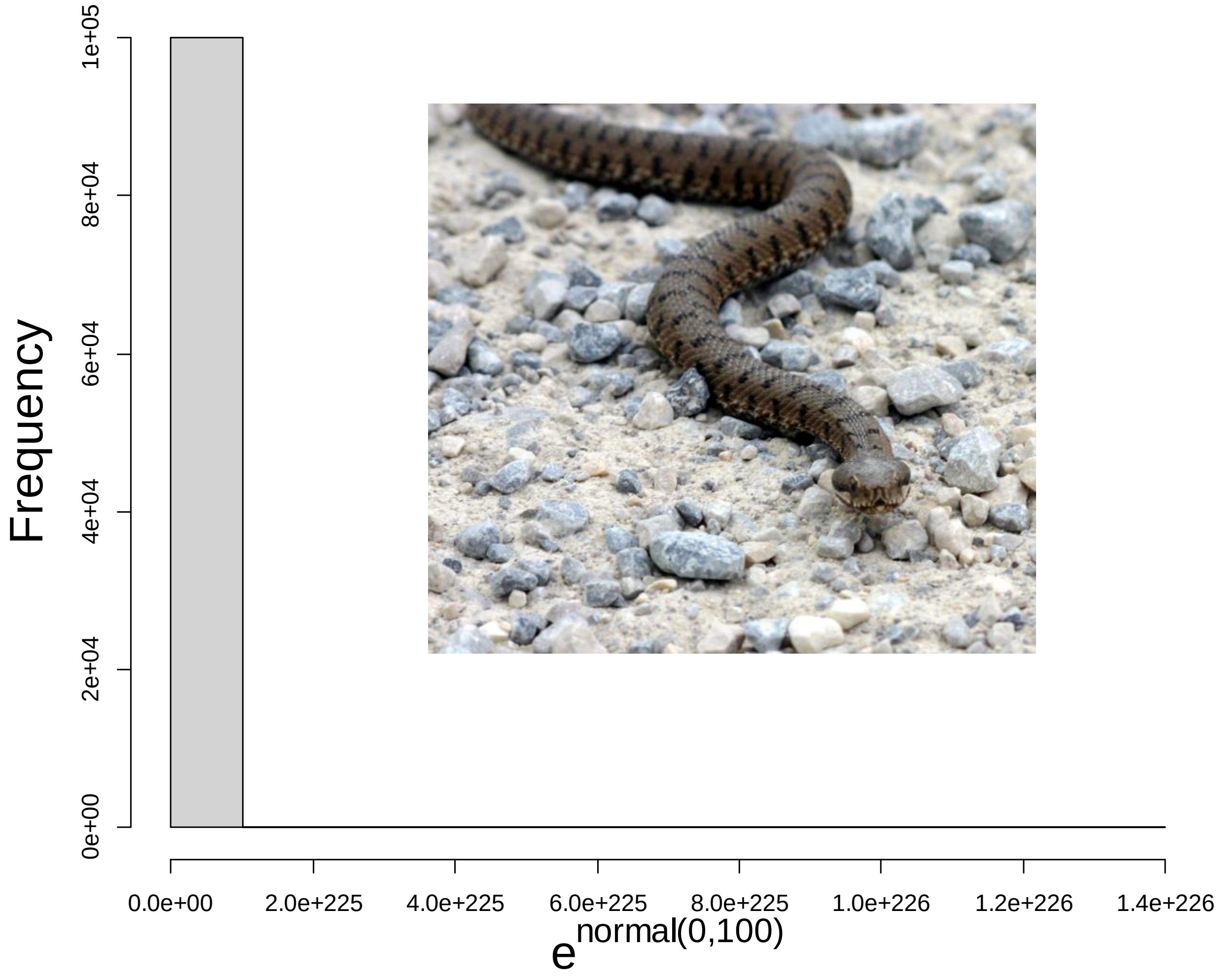




Earth’s diameter is 12,756 km

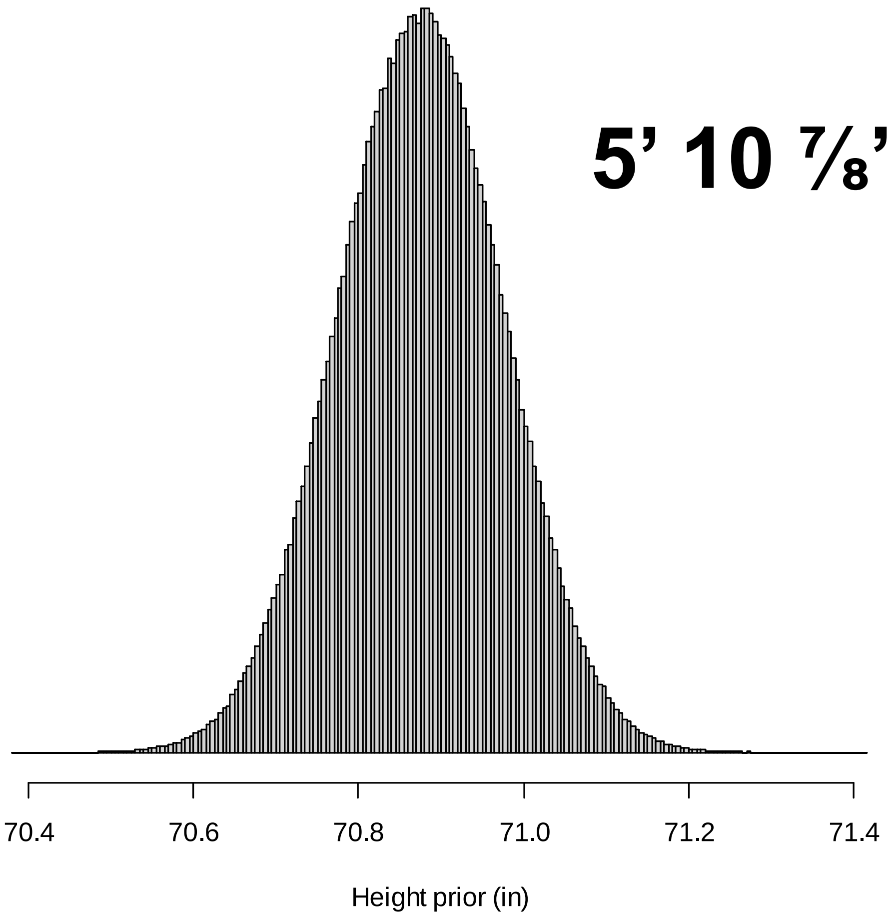
Starting simple: what should our goal be with a prior?





How tall is a man you’ve never met before?

An example: how tall is a man you've never met before?



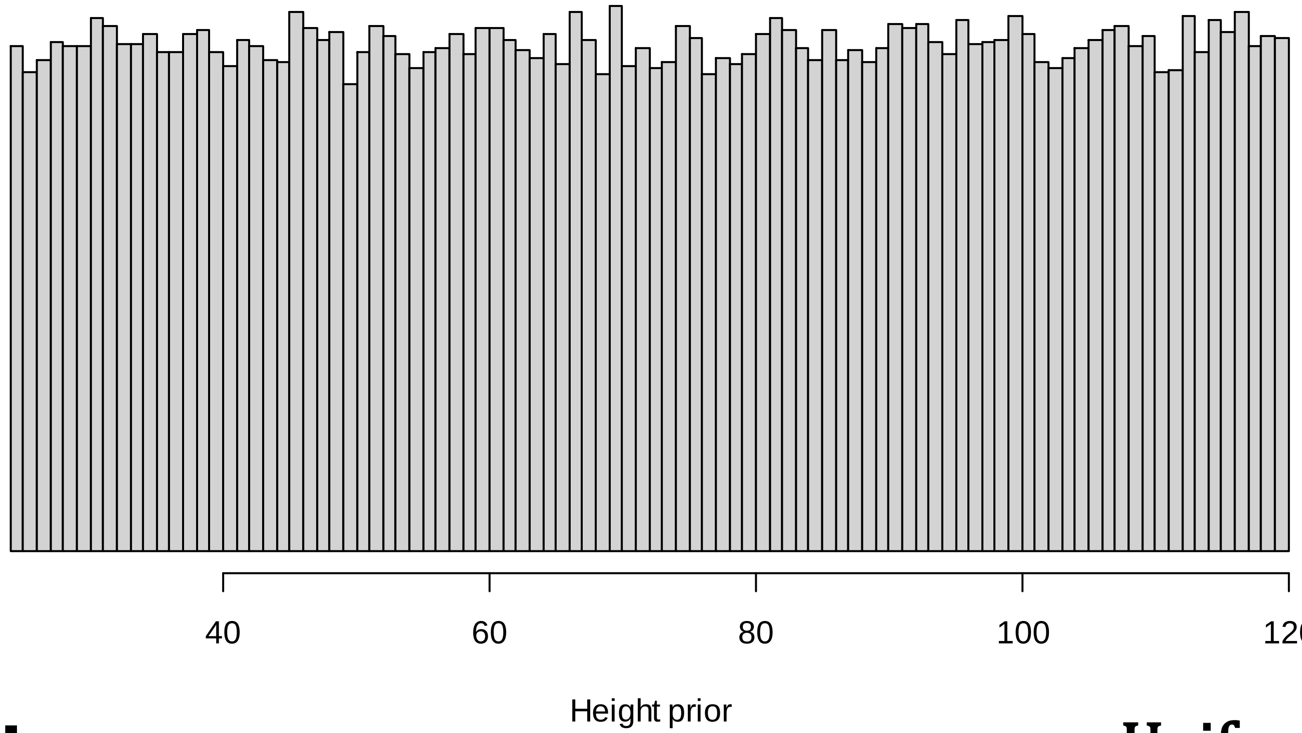
5' 10 ⅞" ± ⅓"

'Very informative'

$\text{Normal}(\mu = 70.875, \sigma = 0.1)$

An example: how tall is a man you've never met before?

Equally likely that they're between 2' and 10' tall

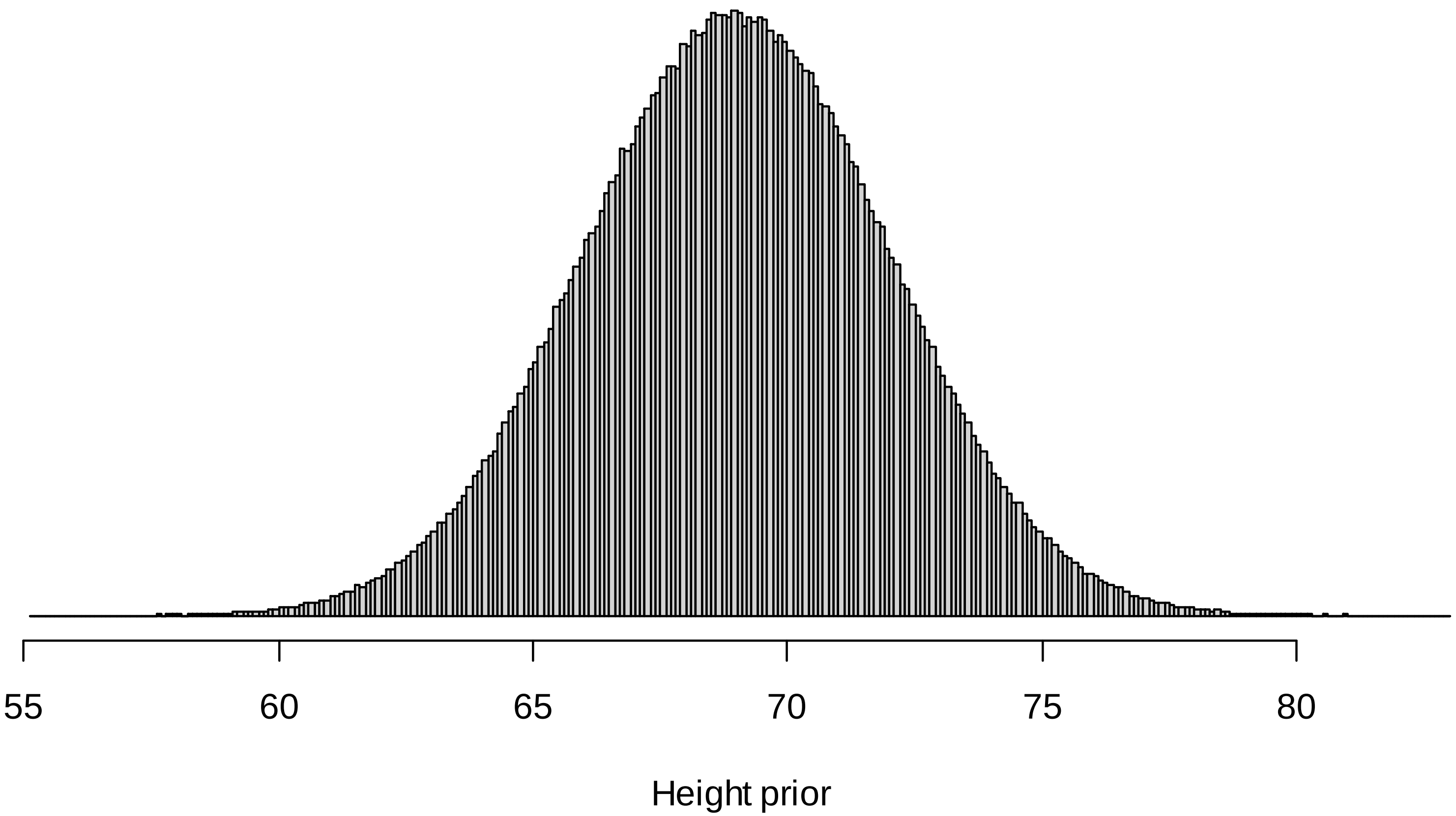


'Uninformative'

$\text{Uniform}(24, 120)$

An example: how tall is a man you've never met before?

5' 10" ± 12"



'Biologically reasonable!'

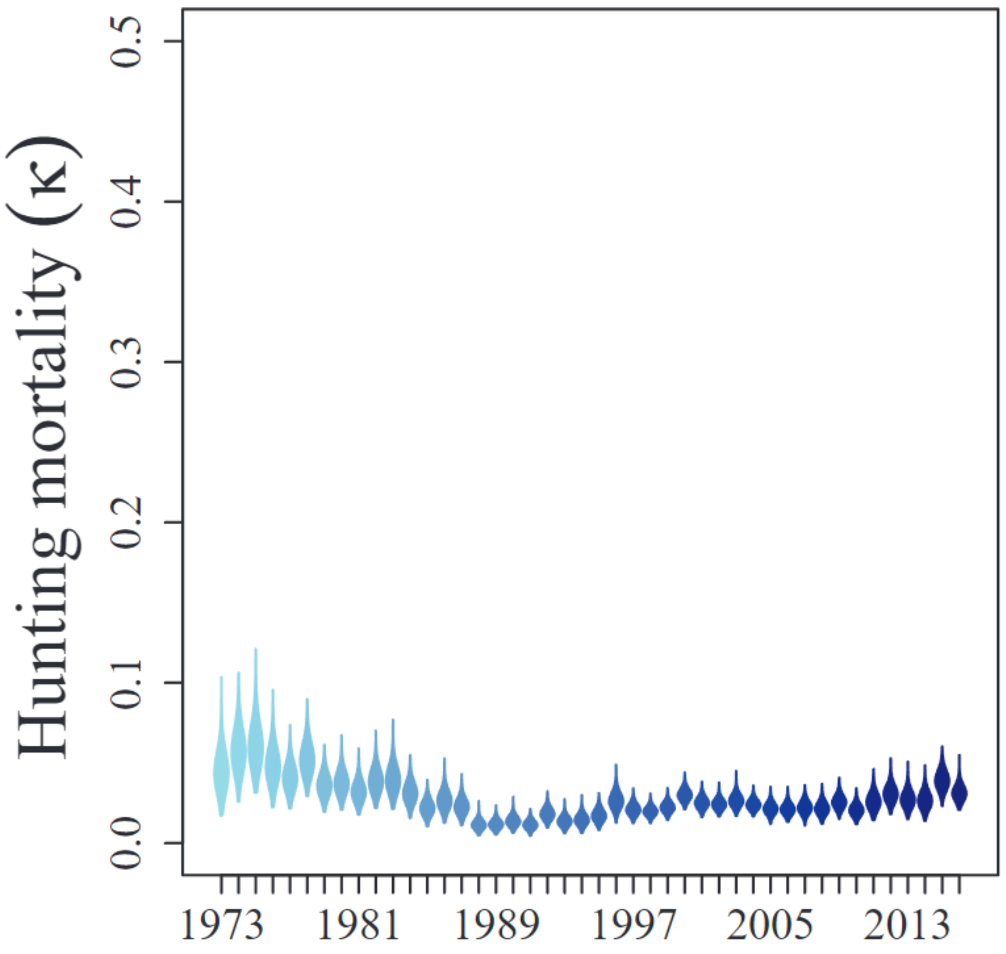
$\mathrm{Normal}(\mu = 70, \sigma = 3)$

An example: how tall is a man you’ve never met before?

‘Please stop talking to me.’

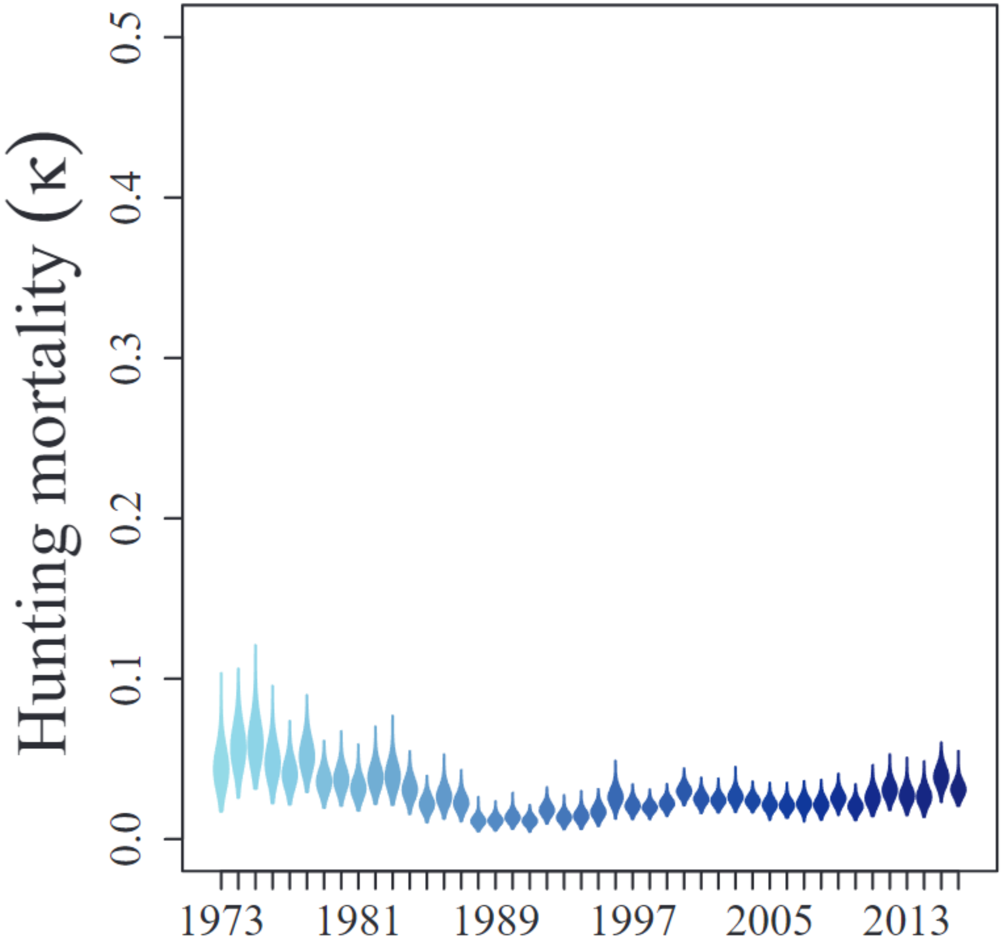
Also an acceptable answer.

Our goal: ‘Biologically informative’ or ‘reasonably vague’ priors



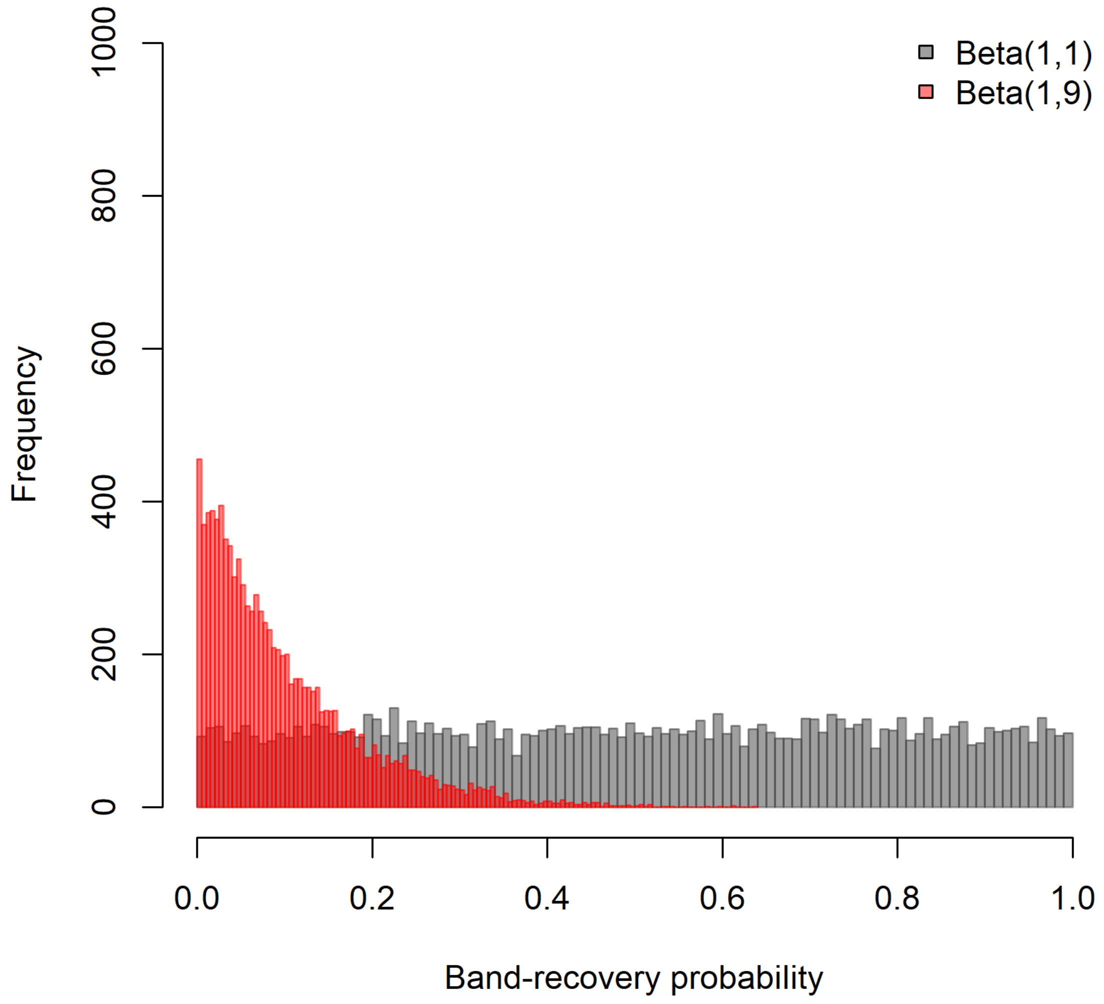
What is the band-recovery probability for adult female blue-winged teal?

Our goal: ‘Biologically informative’ or ‘reasonably vague’ priors

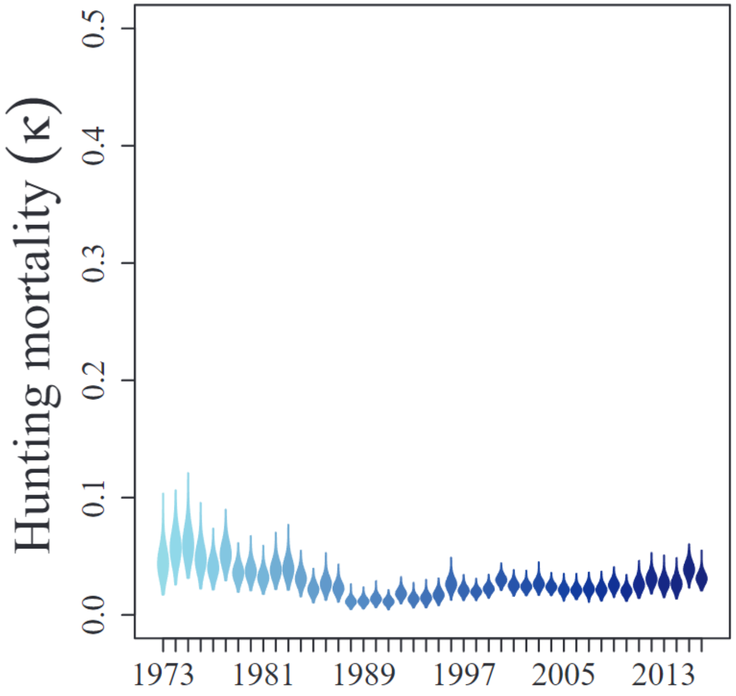


What if you knew a change to hunting regulations was coming that might increase harvest above historic levels?

Our goal: ‘Biologically informative’ or ‘reasonably vague’ priors



What is the band-recovery probability for adult female blue-winged teal?



Prior ‘guidelines’

1. Your prior should be based on the literature or existing knowledge if possible

Prior ‘guidelines’

1. Your prior should be based on the literature or existing knowledge if possible

2. It should be vague enough not to drive the analysis (e.g., prior-posterior overlap)

Prior ‘guidelines’

1. Your prior should be based on the literature or existing knowledge if possible

2. It should be vague enough not to drive the analysis (e.g., prior-posterior overlap)

3. You should not *a priori* believe in impossible values (i.e., priors can be too ‘uninformative’).

Prior ‘guidelines’

1. Your prior should be based on the literature or existing knowledge if possible

2. It should be vague enough not to drive the analysis (e.g., prior-posterior overlap)

3. You should not *a priori* believe in impossible values (i.e., priors can be too ‘uninformative’)

***I recognize that this is <u>incredibly subjective</u>!**

Let’s actually do something!

Beta-binomial models: our first Bayesian analysis

$$y \sim \text{binomial}(n, f)$$

$$f \sim \text{beta}(\alpha, \beta)$$

Beta-binomial models: our first Bayesian analysis

$$y \sim \text{binomial}(n, f)$$

$$f \sim \text{beta}(\alpha, \beta)$$

n: number of marked and released adult male mallards

The binomial distribution

$$y \sim \text{binomial}(n, f)$$

$$f \sim \text{beta}(\alpha, \beta)$$

y: the number of recoveries of banded adult male mallards

Beta-binomial models: our first Bayesian analysis

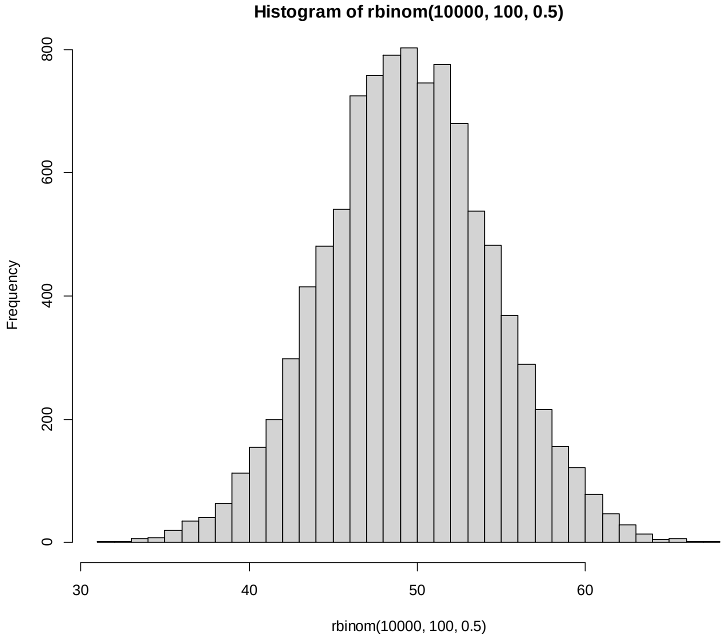
$$y \sim \text{binomial}(n, {\color{red}f})$$

$${\color{red}f} \sim \text{beta}(\alpha, \beta)$$

f: band-recovery probability (the parameter we want to estimate!)

The binomial distribution (and Bernoulli trials)

$$y \sim \text{Bernoulli}(0.5)$$



$$y \sim \text{binomial}(100, 0.5)$$

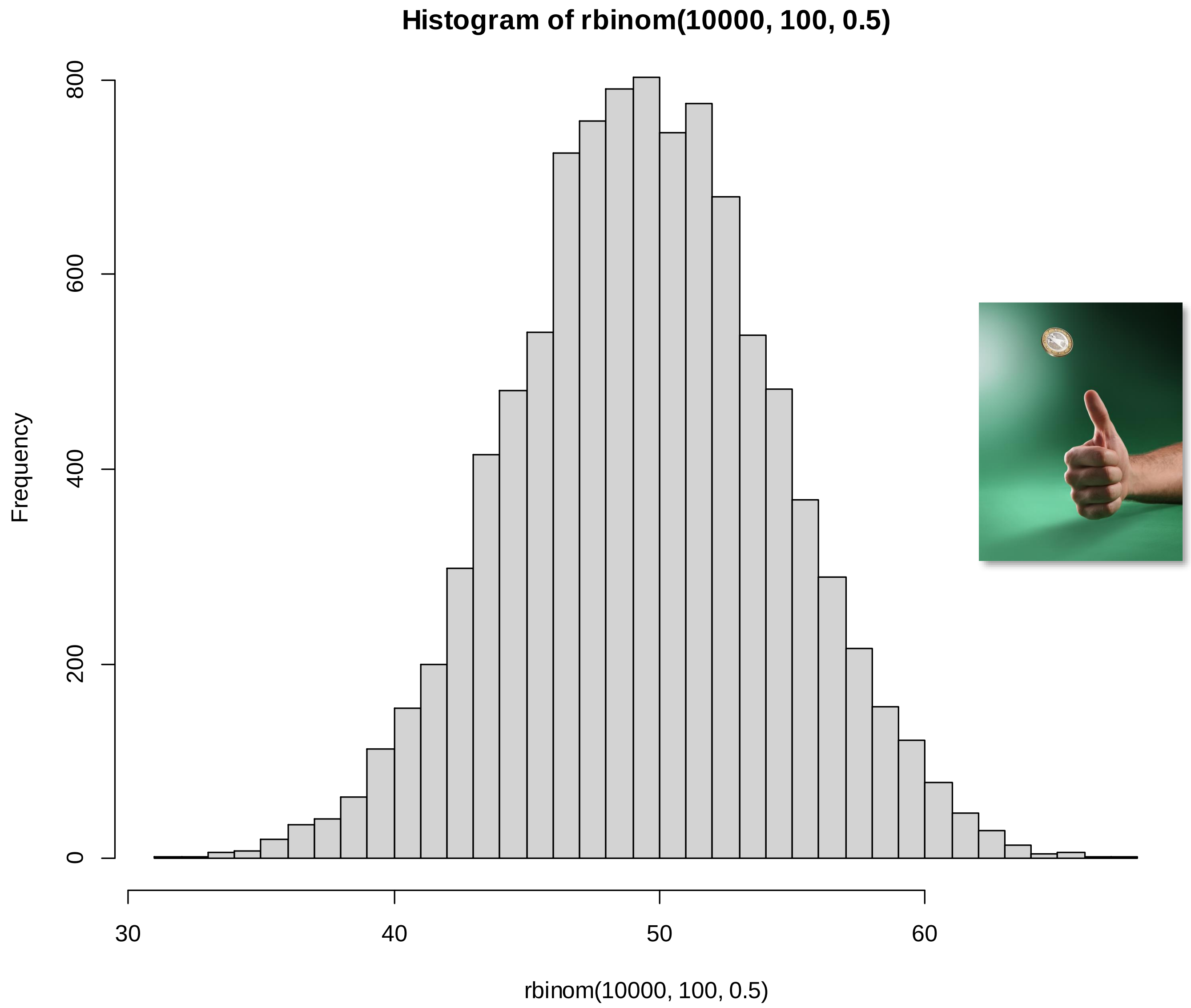
The binomial distribution (flipping coins)

$$y \sim \mathrm{binomial}(n, p)$$

$$y \sim \mathrm{binomial}(100, 0.5)$$

$$E(y) \sim n \times p$$

$$V(y) \sim n \times p \times (1 - p)$$



The binomial distribution

$$y \sim \text{binomial}(n, p)$$

$$y \sim \text{binomial}(40, 0.5)$$

$$E(y) \sim n \times p$$

$$V(y) \sim n \times p \times (1 - p)$$

The beta distribution

$$y \sim \text{binomial}(n, f)$$

$$f \sim \text{beta}(\alpha, \beta)$$

The beta distribution as a conjugate prior for the binomial dist.

Number of heads/wins

Number of flips/games

Prob. of heads/win

$$y \sim \text{binomial}(n, p)$$

$$p \sim \text{beta}(\alpha, \beta)$$

Prior for number of tails/losses

Prior for number of heads/wins

Conjugacy!*

$$y \sim \text{binomial}(n, f)$$

$$f \sim \text{beta}(\alpha, \beta)$$

Example [edit]

The form of the conjugate prior can generally be determined by inspection of the probability density or probability mass function of a distribution. For example, consider a random variable which consists of the number of successes s in n Bernoulli trials with *unknown* probability of success q in [0,1]. This random variable will follow the binomial distribution, with a probability mass function of the form

$$p(s) = \binom{n}{s} q^s (1-q)^{n-s}$$

The usual conjugate prior is the beta distribution with parameters (α, β):

$$p(q) = \frac{q^{\alpha-1}(1-q)^{\beta-1}}{\mathrm{B}(\alpha, \beta)}$$

where α and β are chosen to reflect any existing belief or information ($\alpha = 1$ and $\beta = 1$ would give a uniform distribution) and $\mathrm{B}(\alpha, \beta)$ is the Beta function acting as a normalising constant.

If we sample this random variable and get s successes and $f = n - s$ failures, then we have

$$\begin{aligned}
P(s, f \mid q = x) &= \binom{s+f}{s} x^s (1-x)^f, \\
P(q = x) &= \frac{x^{\alpha-1}(1-x)^{\beta-1}}{\mathrm{B}(\alpha, \beta)}, \\
P(q = x \mid s, f) &= \frac{P(s, f \mid x) P(x)}{\int P(s, f \mid y) P(y) dy} \\
&= \frac{\binom{s+f}{s} x^{s+\alpha-1}(1-x)^{f+\beta-1} / \mathrm{B}(\alpha, \beta)}{\int_{y=0}^{1} \left(\binom{s+f}{s} y^{s+\alpha-1}(1-y)^{f+\beta-1} / \mathrm{B}(\alpha, \beta) \right) dy} \\
&= \frac{x^{s+\alpha-1}(1-x)^{f+\beta-1}}{\mathrm{B}(s+\alpha, f+\beta)},
\end{aligned}$$

which is another Beta distribution with parameters $(\alpha + s, \beta + f)$. This posterior distribution could

We can do this by hand!!

$$y \sim \text{binomial}(n, f)$$

Data model

$$f \sim \text{beta}(\alpha, \beta)$$

Prior

$$\hat{f} \sim \text{beta}(\alpha + y, \beta + (n - y))$$

Posterior

A quick example

Data model

$$y \sim \text{binomial}(n, f)$$

$$f \sim \text{beta}(\alpha, \beta)$$

Prior

$$\hat{f} \sim \text{beta}(\alpha + y, \beta + (n - y))$$

Posterior

A quick example

$$y \sim \text{binomial}(n, f)$$

$$137 \sim \text{binomial}(1000, f)$$ Data

$$f \sim \text{beta}(1, 1)$$ Prior

$$\hat{f} \sim \text{beta}(1 + 137, 1 + (863))$$ Posterior

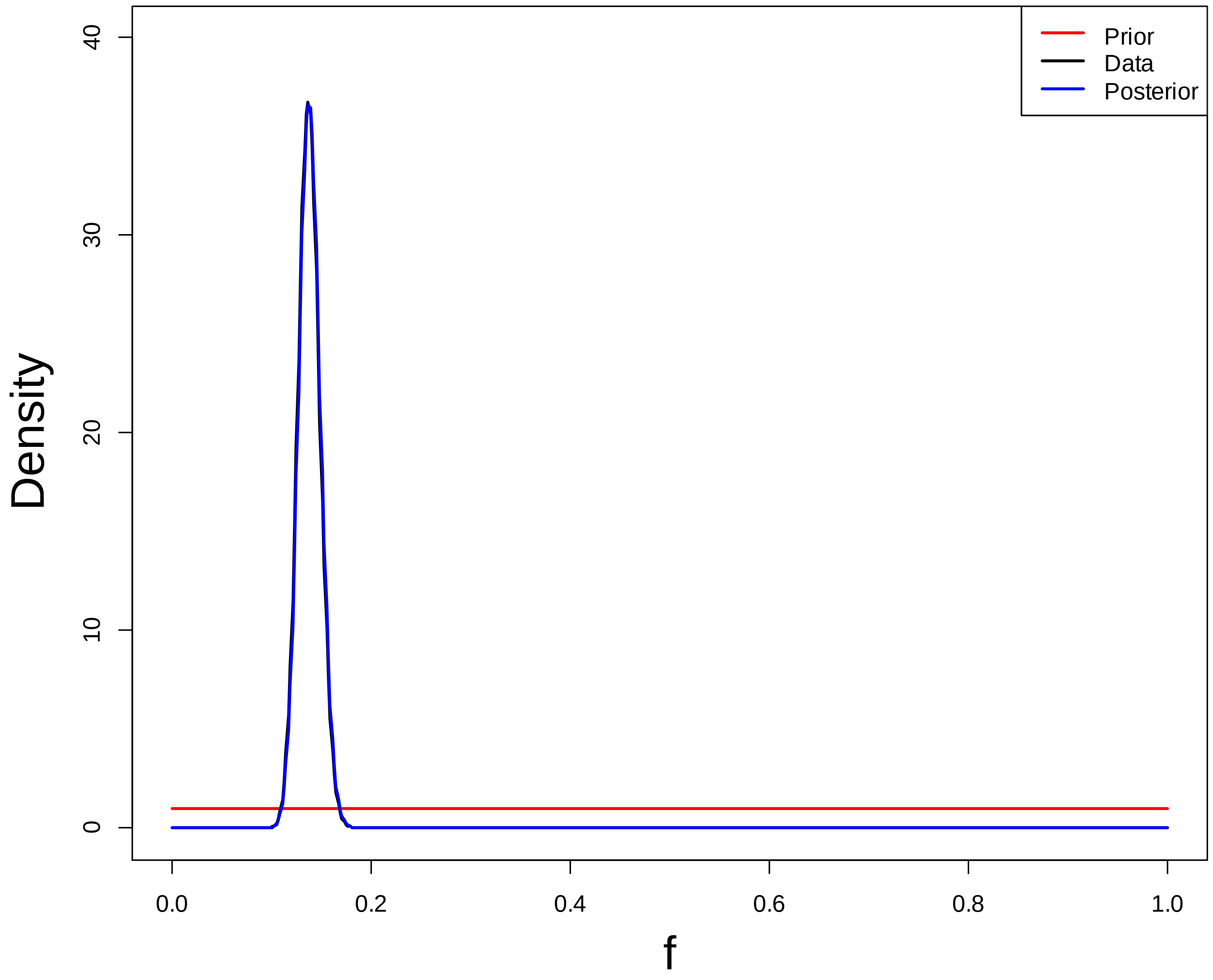
$$\hat{f} \sim \text{beta}(\alpha + y, \beta + (n - y))$$

A quick example

$$137 \sim \text{binomial}(1000, f)$$

$$f \sim \text{beta}(1, 1)$$

$$\hat{f} \sim \text{beta}(138, 864)$$



Playing with fire…

Weak (uniform b/w 0 and 1)

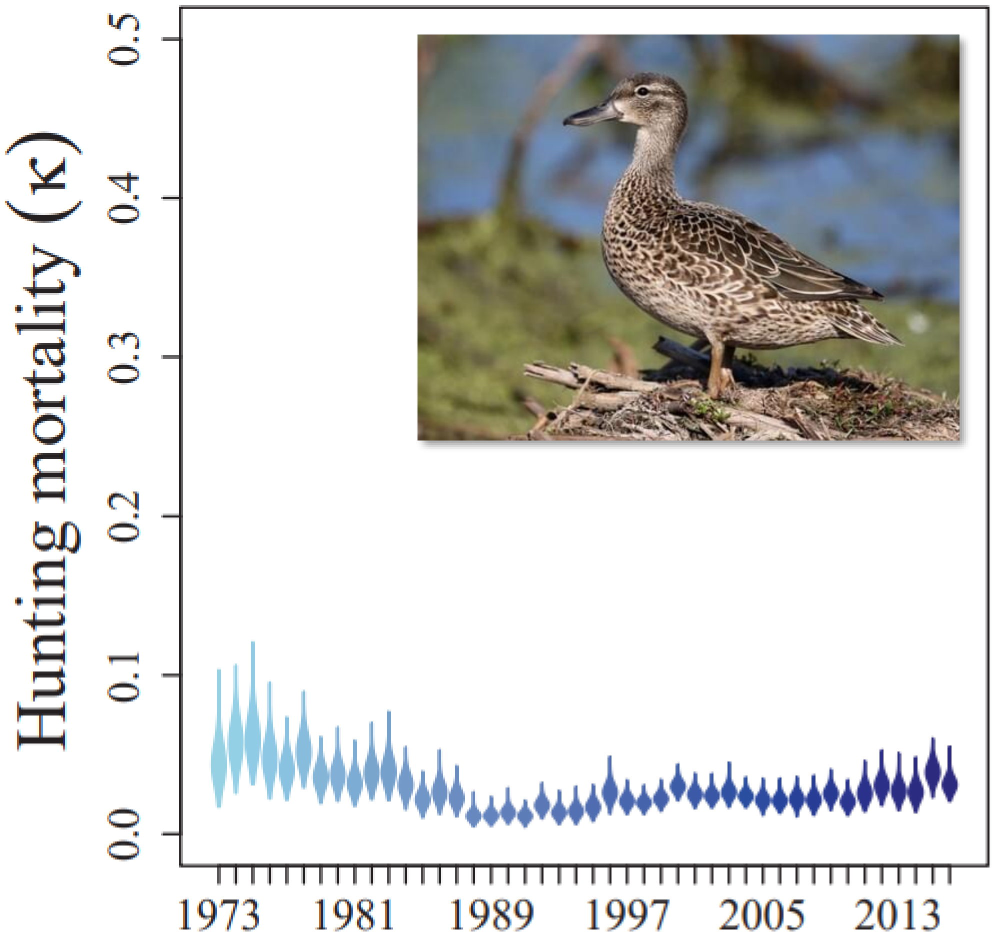
$$f \sim \text{beta}(1, 1)$$

Moderate

$$f \sim \text{beta}(1, 9)$$

Strong

$$f \sim \text{beta}(5, 95)$$



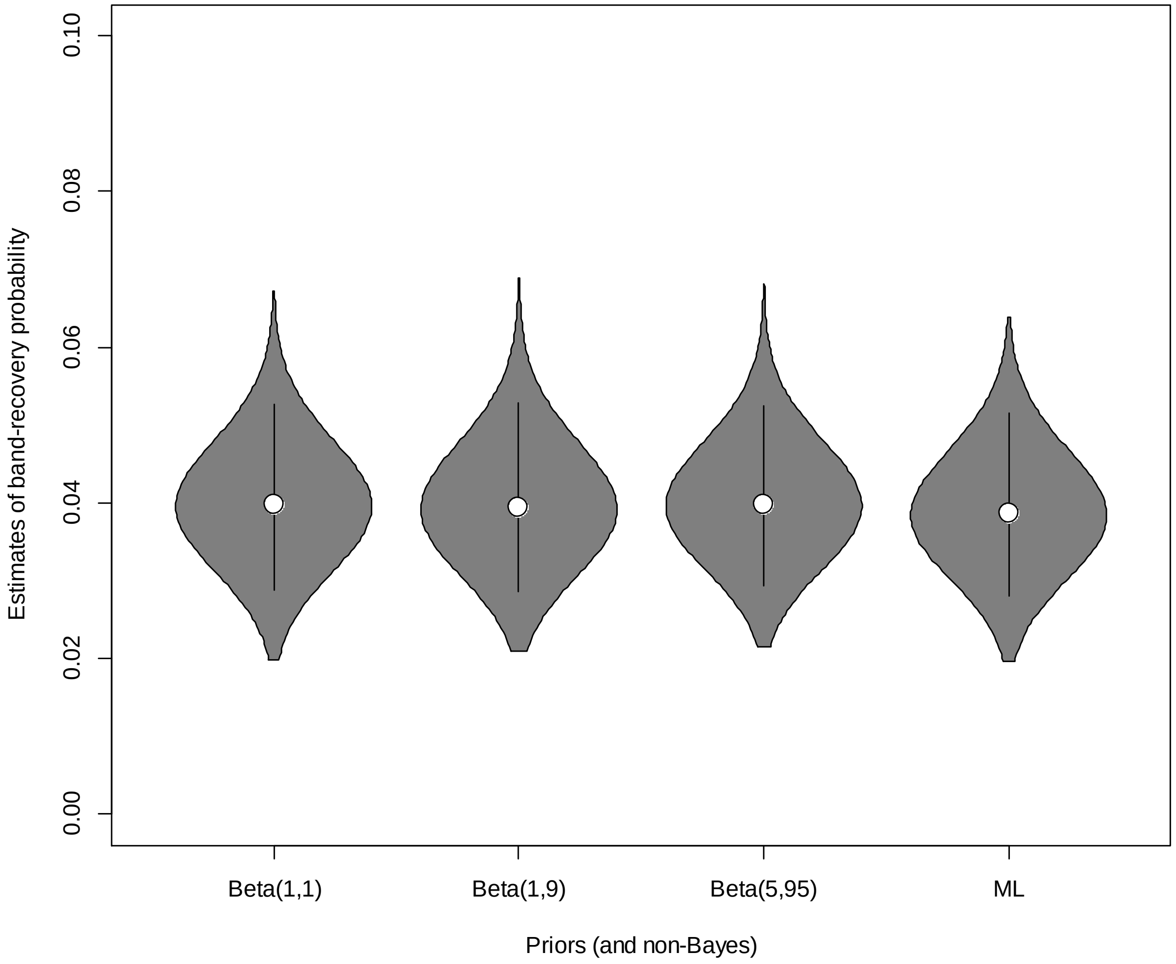
Let's try this again with some more substantial prior information

We just generated posterior distributions!

- They are vectors of length 10000 (samples)
- They contain means, medians, and CREDIBLE intervals
- We can extract those with simple R functions (Lines 80-85)

So what happened (Lines 88-97)?

We have strong data (and excellent prior knowledge)!



We marked 1000 ducks! The means vary by hundredths of 1 percent

What if we REALLY had strong prior knowledge

Let's assumed we're convinced we'll shoot 50 out of 1000 teal

Try beta(50,950) as a 'strong' prior

What if our prior knowledge was bad?

Let’s assume we’re convinced that 7%, or 7 out of 100, of teal get shot!

Try beta(7,93) as a ‘strong’ prior

What if our prior knowledge was bad?

Let’s assume we’re convinced that 7%, or 7 out of 100, of teal get shot!

Try beta(70,930) as a ‘strong’ prior

Moment-matching?!

Imagine that we have an estimate of average teal band-recovery probability.
It's reported in a manuscript as 0.0487 (sd = 0.01).

How would we use that in a beta prior?

Moment-matching?!

If we have an estimate with μ=0.0487 and σ=0.01

$$\alpha = \left(\frac{1-\mu}{\sigma^2} - \frac{1}{\mu}\right)\mu^2$$

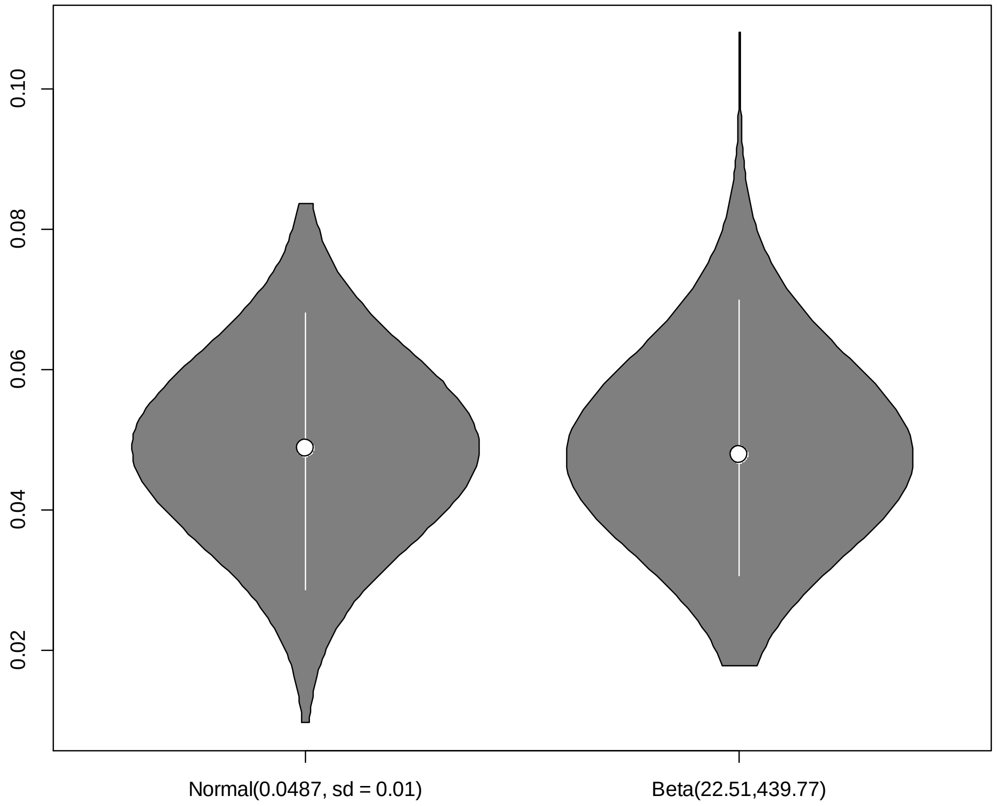
$$\beta = \alpha\left(\frac{1}{\mu} - 1\right)$$

Moment-matching?!

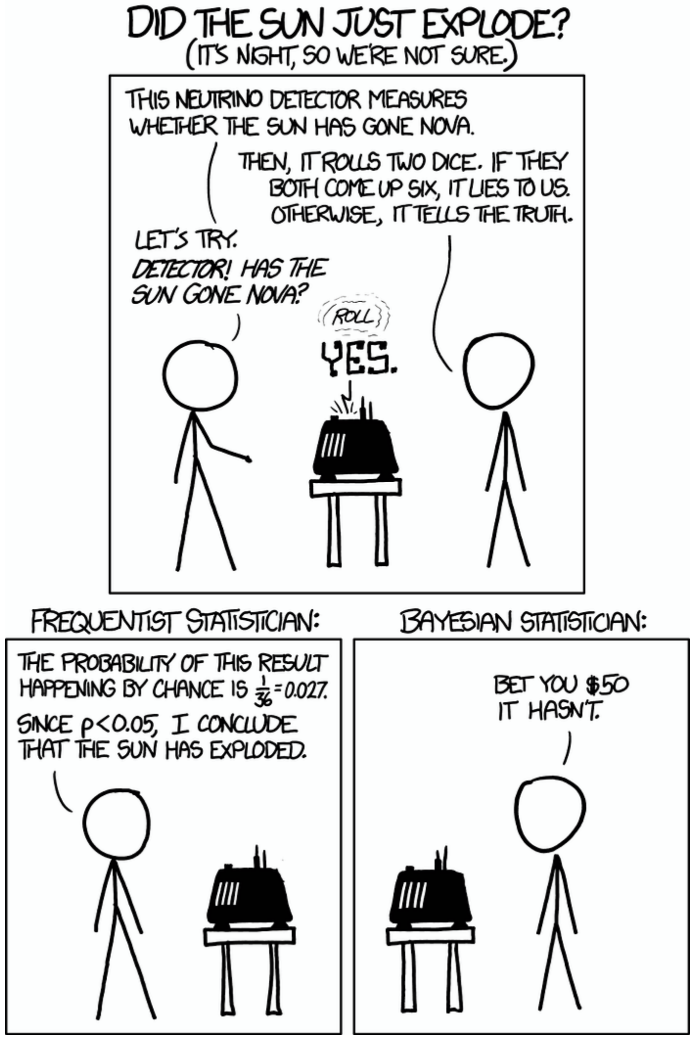
If we have an estimate with μ=0.0487 and σ=0.01

$$\alpha = \left(\frac{1-\mu}{\sigma^2} - \frac{1}{\mu}\right)\mu^2$$

$$\beta = \alpha\left(\frac{1}{\mu} - 1\right)$$



Wrapping up…



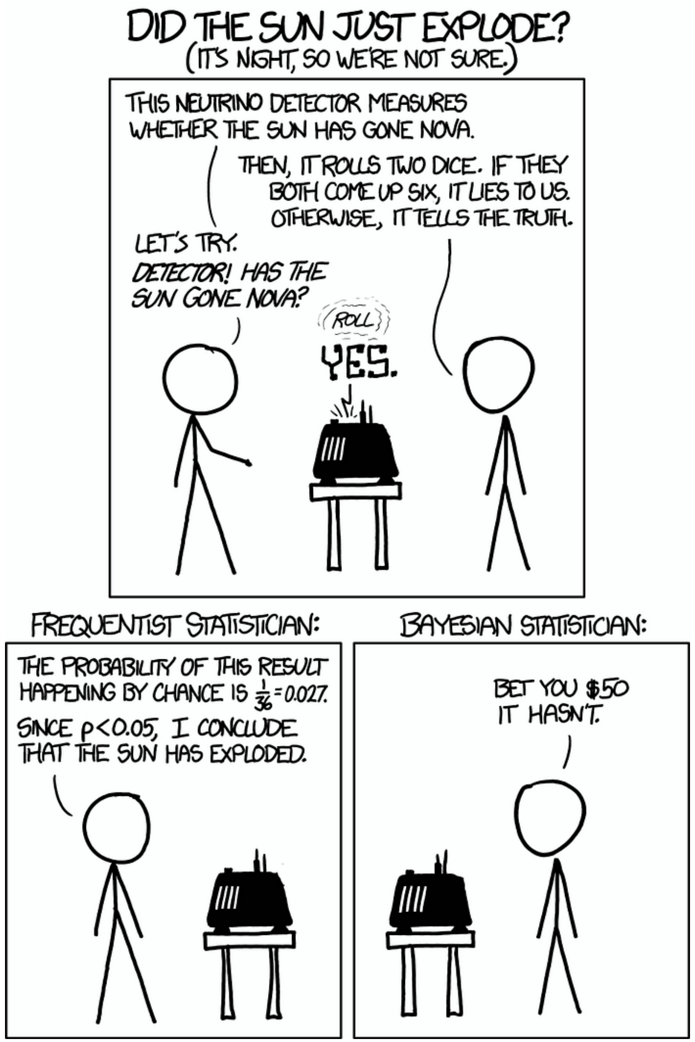
P(sun explosion) = 0.00000000000…000001

$$P(\text{sun explosion}|\text{YES}) = \frac{P(+|\text{sun explosion}) \times P(\text{sun explosion})}{P(+)}$$

Key take-home: Priors are just that, our prior belief

Key take-home: they are subjective, and cannot be uninformative

Key take-home: strong priors can save us from ‘bad’ data!



Key take-home: they can also hide real signals in data!

Imagine you were convinced the right prior for our duck harvest example was beta(500,500)

Key take-home: we can perform prior sensitivity analyses!!!

- If our prior sensitivity analyses reveal little difference in inference between extremely vague and extremely informative priors, then our conclusions may be robust (given our model assumptions).

Summary of summaries:

1. **Priors are just that, our prior belief. They are 'objectively subjective', <u>cannot be uninformative,</u> and should be carefully thought about based upon existing knowledge***

2. **Strong priors can save us from 'bad' data, or they can hide real signals in data!**

3. **Prior sensitivity analysis can allow us to determine how much the strength of our beliefs impacts inference**

***sorry to be a grumpy old man here! This is important!**

Reading

https://github.com/thomasriecke/ST595