Lecture 20 - Intro to Bayesian Statistics

Emily Josephs
November 10 2022

Quick Recap

1. Probability
2. Probability distributions
3. Likelihood

Quick Recap

1. Probability: Figuring out how likely it is that an event (or many events, combinations of events) will happen.

2. Probability distributions

3. Likelihood

Quick Recap

1. Probability: Figuring out how likely it is that an event (or many events, combinations of events) will happen.
2. Probability distributions: Calculating the probability of event(s) based on specific parameters.
3. Likelihood

Quick Recap

1. Probability: Figuring out how likely it is that an event (or many events, combinations of events) will happen.

2. Probability distributions: Calculating the probability of event(s) based on specific parameters.

3. Likelihood: Given a set of events, what are the most likely parameters?

Bird’s-eye view

Step 7: say something *quantitative* and *objective* about the effect of a predictor on a response.

Step 6: run model (use likelihood-based inference and an algorithm to estimate values of parameters of the model)

Step 5: build model (make choices about how we think our data are distributed and the nature of the relationship between predictor and response)

Step 4: know about prob. distributions and deterministic functions

Step 3: understand the concept of likelihood

Step 2: understand the rules of probability

Step 1: be able to do stuff in R

Bird's-eye view

Step 7: say something *quantitative* and *objective* about the effect of a predictor on a response.

Step 6: run model (use likelihood-based inference and an algorithm to estimate values of parameters of the model)

Step 5: build **Bayesian** model (make choices about how we think our data are distributed and the nature of the relationship between predictor and response)

Step 4: know about prob. distributions and deterministic functions

Step 3: understand the concept of likelihood

Step 2: understand the rules of probability

Step 1: be able to do stuff in R

Bird's-eye view

Step 7: say something *quantitative* and *objective* about the effect of a predictor on a response.

Step 6: run model (use likelihood-based inference and an algorithm to estimate values of parameters of the model)

Step 5: build **Bayesian** model (make choices about how we think our data are distributed and the nature of the relationship between predictor and response)

Step 4: know about prob. distributions and deterministic functions

Step 3.5: **learn about Bayes Theorem**

Step 3: understand the concept of likelihood

Step 2: understand the rules of probability

Deriving Bayes Theorem

Say we have a hypothesis H that the probability that a given coin flips heads is 0.5.

And we have observed data D of a large number of flips from that coin.

Deriving Bayes Theorem

We might be interested in the conditional probability of our data D given hypothesis H.

$$p(D \mid H) = \frac{P(H \cap D)}{P(H)}$$

We also know, based on how conditional probabilities work, the conditional probability of our hypothesis H given our data D.

$$p(H \mid D) = \frac{P(D \cap H)}{P(D)}$$

We can combine these two equations and rearrange.

Deriving Bayes Theorem

$p(H \mid D) = \frac{P(D \cap H)}{P(D)}$ and $p(D \mid H) = \frac{P(H \cap D)}{P(H)}$

$p(H \mid D) = \frac{P(D \cap H)}{P(D)}$ and $P(H \ \cap \ D) = p(D \mid H) \times P(H)$

$$p(H \mid D) = \frac{P(D \mid H) \times P(H)}{P(D)}$$

Bayes Theorem

$$p(H \mid D) = \frac{P(D \mid H) \times P(H)}{P(D)}$$

Bayes Theorem allows us to assess the probability of a hypothesis given some data, rather than the other way around, as in frequentist statistics.

Bayes Theorem

$$p(H \mid D) = \frac{P(D \mid H) \times P(H)}{P(D)}$$

$p(H \mid D)$ is the *posterior probability*

$p(D \mid H)$ is the *likelihood* of D given H

$p(H)$ is the *prior probability* of H

$p(D)$ is the *marginal likelihood* or *model evidence* of D

Bayes Theorem - Example

- Using data collected in 1975, we know that the proportion of people that develop thyroid cancer is 10^{-4}.
- The probability that a biopsy correctly identifies these people as having cancer is 0.9.
- The probability of a “false positive” (the test saying there was cancer when there wasn’t) is 0.001.

What is the probability that a person with a positive result actually has cancer?

Bayes Theorem - Example

$$p(H \mid D) = \frac{P(D \mid H) \times P(H)}{P(D)}$$

What’s H, and what’s D?

“H” is cancer, and “D” is the positive result.

$$p(\text{Cancer} \ \mid \ \text{PosRes}) = \frac{p(\text{PosRes} \ \mid \ \text{Cancer}) \ \times \ p(\text{Cancer})}{p(\text{PosRes})}$$

Bayes Theorem - Example

$$p(\text{Cancer} \mid \text{PosRes}) = \frac{p(\text{PosRes} \mid \text{Cancer}) \times p(\text{Cancer})}{p(\text{PosRes})}$$

$$p(\text{PosRes} \mid \text{Cancer}) =$$

$$p(\text{Cancer}) =$$

$$p(\text{PosRes}) =$$

Bayes Theorem - Example

$$p(\text{Cancer} \;\mid\; \text{PosRes}) = \frac{p(\text{PosRes} \mid \text{Cancer}) \;\times\; p(\text{Cancer})}{p(\text{PosRes})}$$

$p(\text{PosRes} \;\mid\; \text{Cancer}) =$ test sensitivity (0.9)

$p(\text{Cancer}) =$ cancer frequency (10^{-4})

$p(\text{PosRes}) =$???

Bayes Theorem - Example

What about $p(\mathrm{PosRes})$?

We can break all positive results down into positive results where one has cancer, and positive results where one doesn't have cancer.

$$\begin{aligned} p(\mathrm{PosRes}) = {} & p(\mathrm{Cancer}) \times p(\mathrm{PosRes} \mid \mathrm{Cancer}) \\ & + p(\mathrm{noCancer}) \times p(\mathrm{PosRes} \mid \mathrm{noCancer}) \end{aligned}$$

$p(\mathrm{PosRes}) =$ 10^{-4} $\times$ 0.9 $+$ 0.9999 $\times$ 0.001

$p(\mathrm{PosRes}) =$ 0.00109

Bayes Theorem - Example

$$p(\mathrm{Cancer} \ \mid \ \mathrm{PosRes}) = \frac{p(\mathrm{PosRes} \mid \mathrm{Cancer}) \times p(\mathrm{Cancer})}{p(\mathrm{PosRes})}$$

$p(\mathrm{PosRes} \ \mid \ \mathrm{Cancer}) =$ 0.9

$p(\mathrm{Cancer}) =$ 10^{-4}

$p(\mathrm{noCancer}) =$ 0.9999

$p(\mathrm{PosRes} \mid \mathrm{noCancer}) =$ 0.001

$p(\mathrm{PosRes}) =$ 0.00109

Bayes Theorem - Example

$p(\text{Cancer} \mid \text{PosRes}) = \frac{9 \times 10^{-5}}{0.00109}$

$p(\text{PosRes} \mid \text{Cancer}) =$ 0.9

$p(\text{Cancer}) =$ 10^{-4}

$p(\text{noCancer}) =$ 0.9999

$p(\text{PosRes} | \text{noCancer}) =$ 0.001

$p(\text{PosRes}) =$ 0.00109

Bayes Theorem - Example

$p(\text{Cancer} \ \mid \ \text{PosRes}) = \frac{9 \times 10^{-5}}{0.00109}$ = 0.08258

$p(\text{PosRes} \ \mid \ \text{Cancer}) =$ 0.9

$p(\text{Cancer}) =$ 10^{-4}

$p(\text{noCancer}) =$ 0.9999

$p(\text{PosRes} \mid \text{noCancer}) =$ 0.001

$p(\text{PosRes}) =$ 0.00109

Bayes'd & Confused

What's going on here?

We know the false positive rate is only 0.001, so if a person gets a "positive" diagnosis, shouldn't there be a 0.999 that they have cancer?

Yes, but think about the numbers!

Bayes'd & Confused

Out of 10^{7} people:

- there will be 1000 cancer victims
    - of whom 900 will be correctly diagnosed.

Of the remaining 9.999^{6} healthy people:

- there will be 9999 false positives.

So if there's a positive diagnosis, the probability that a person actually has cancer is $\frac{900}{900+9999}$ = 0.08258

Bayes Theorem - Goats and Doors

I’m a game show host, and you’re a contestant!

The game:

There are 3 doors. One hides $1,000,000, and there are goats behind the other two doors.

for the purposes of the game, we assume goats are not a pet that we’d be delighted to welcome into the family.

Bayes Theorem - Goats and Doors

I’m a game show host, and you’re a contestant!

The game:

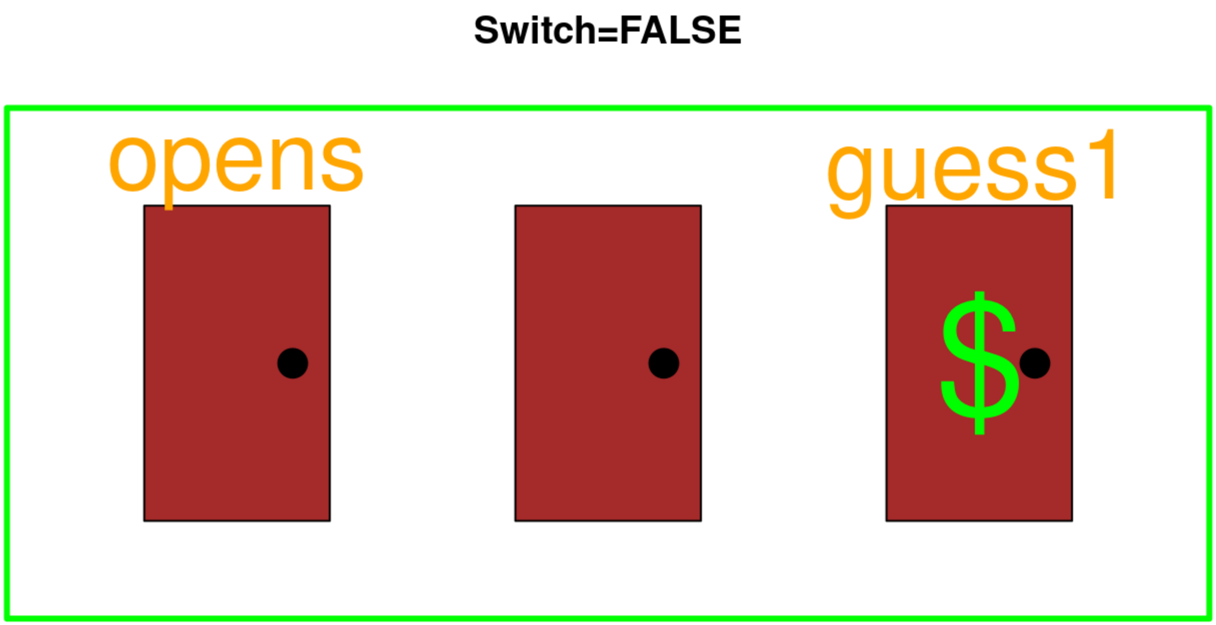
There are 3 doors. One hides $1,000,000, and there are goats behind the other two doors.

The rules:

- You guess a door, then I open one of the *other two doors*, and I *never reveal the money*.
- After I open the door, I ask whether you want to switch your guess to the *other* door.
- **Should you switch your guess?!** (POLL: A for switch and B for stay)

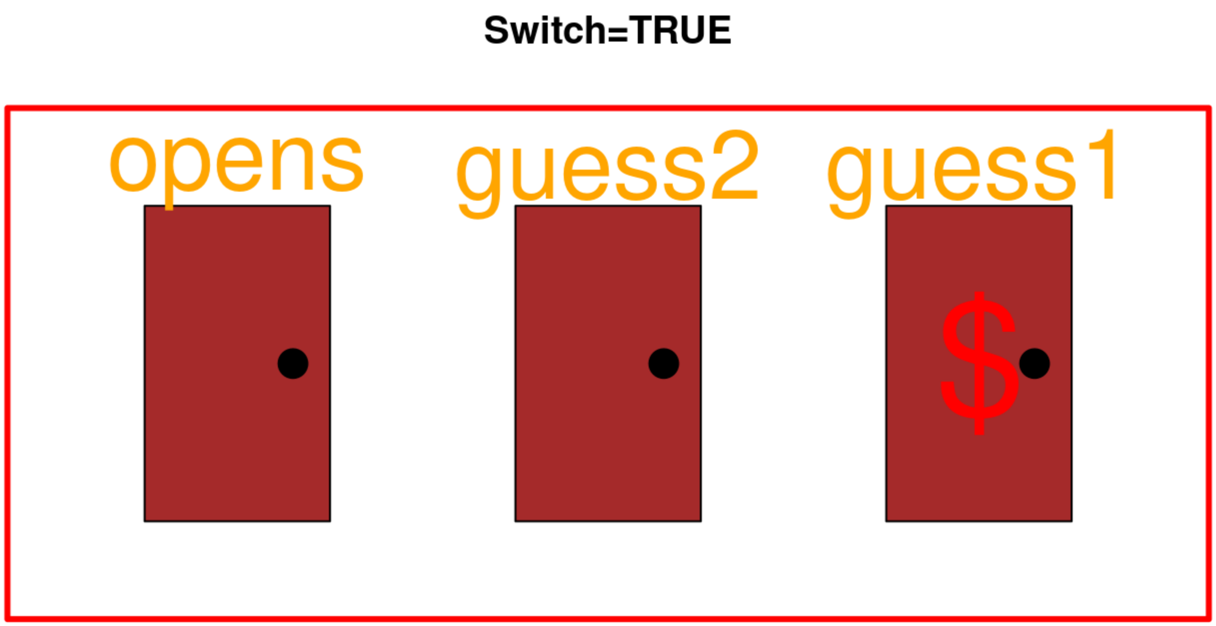
Goats & Doors (Monty Hall)

Goats & Doors (Monty Hall)



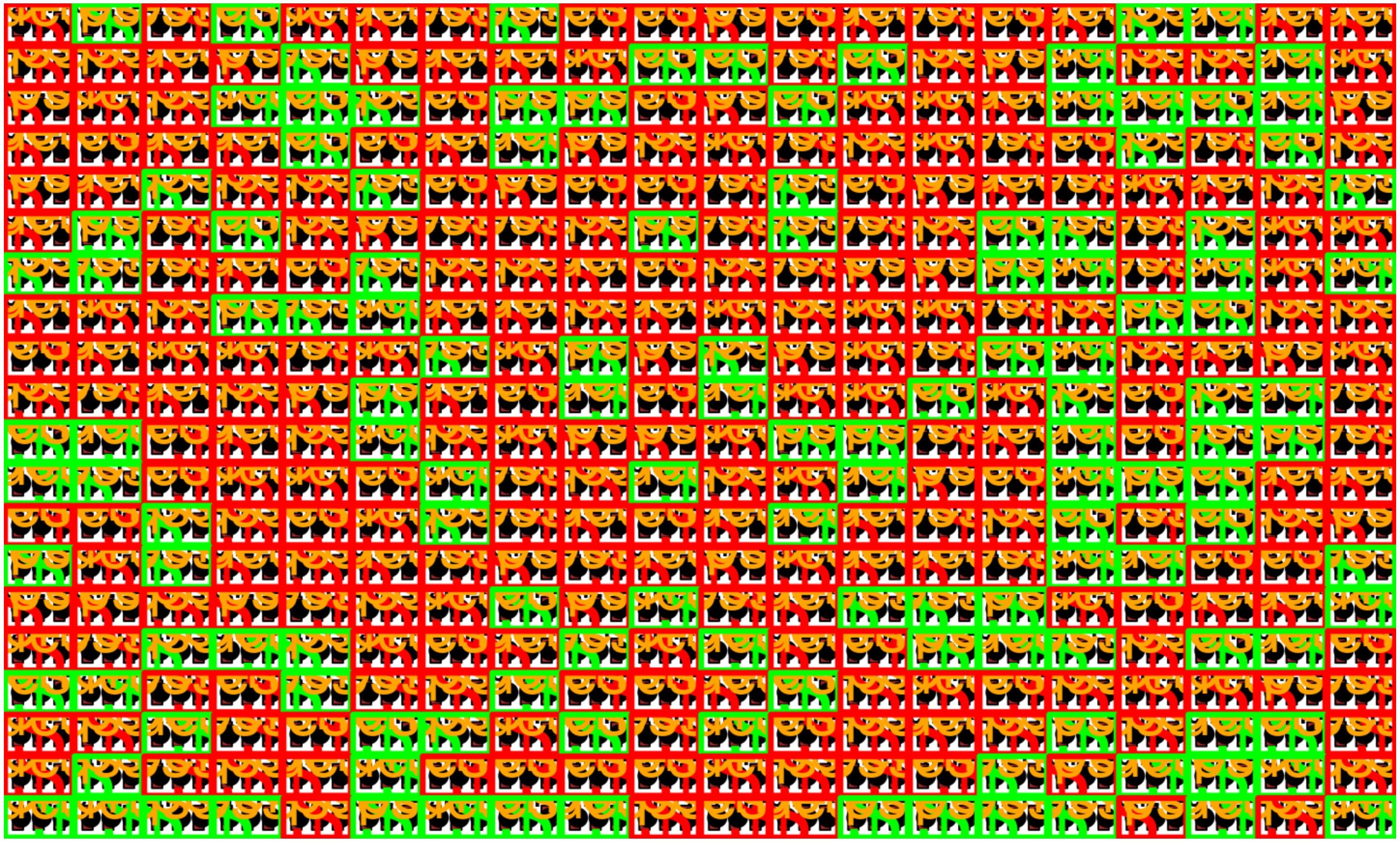

```
noswitch.results <- replicate(400,monty.hall.sim(switch=FALSE))
```

Goats & Doors (Monty Hall)

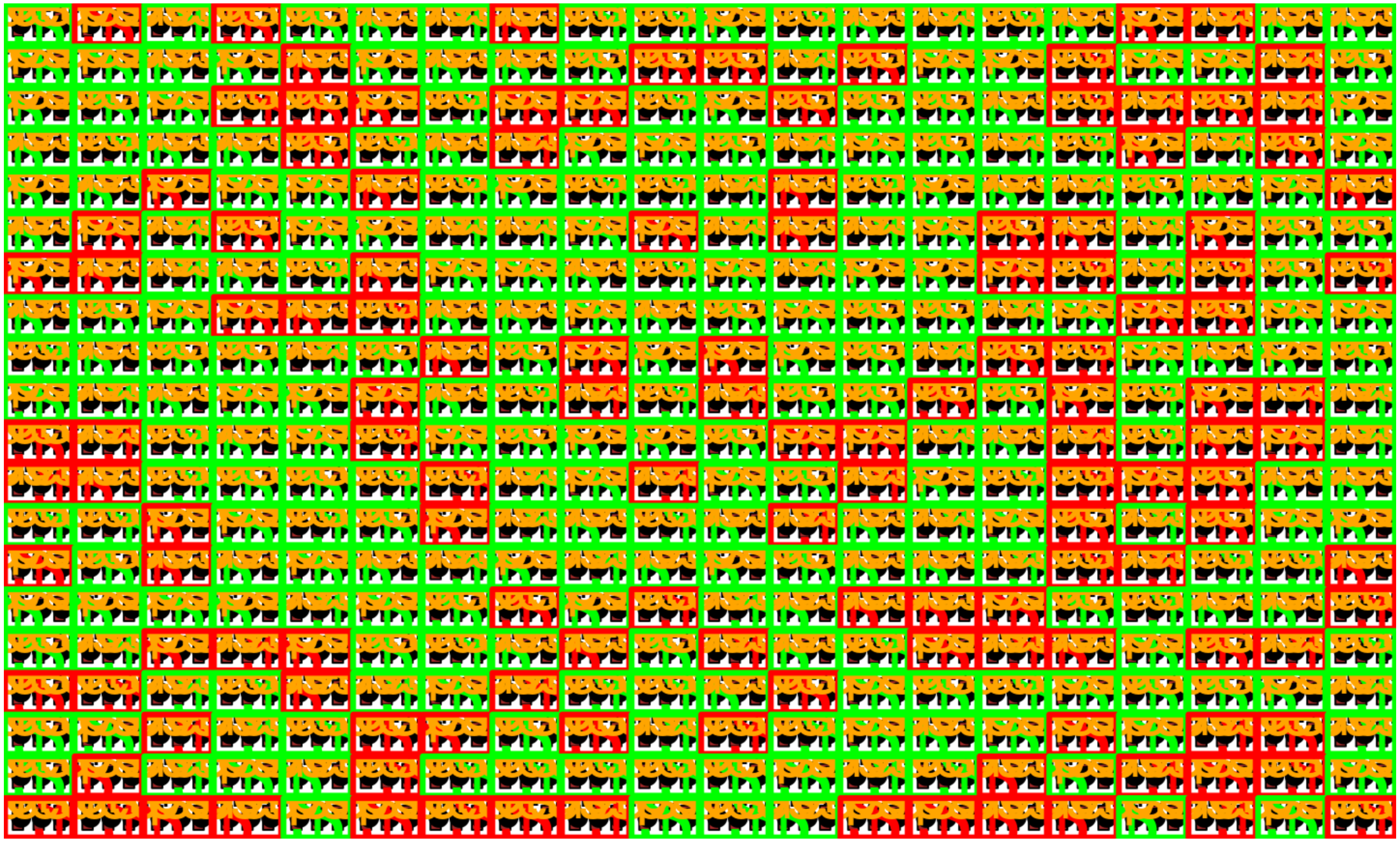



```
switch.results <- replicate(400,monty.hall.sim(switch=TRUE))
```

No switch p(win): 0.3375



Switch p(win): 0.6625



What's going on??

Let's use Bayes theorem to figure it out.

Say I pick Door 1 and Monty opens Door 3.

H1: Door 1 has money
H2: Door 2 has money
H3: Door 3 has money
D3 = Door 3 opened after picking Door 1

We want to compare $p(H1 \mid D3)$ with $p(H2 \mid D3)$ to decide whether to switch or stay.

Goats & Doors, with Bayes

$$p(H1 \mid D3) = \frac{p(D3 \ \mid \ H1) \ \times \ p(H1)}{p(D3)}$$

$$p(H2 \mid D3) = \frac{p(D3 \ \mid \ H2) \ \times \ p(H2)}{p(D3)}$$

$$p(D3 \ \mid \ H1) = \frac{1}{2}$$

$$p(H1) = \frac{1}{3}$$

$$p(D3) = ?$$

(D3 = Door 3 opened after picking Door 1)

Goats & Doors, with Bayes

D3 = Door 3 opened after picking Door 1

$$\begin{aligned} p(D3) = \; & p(H1) \times p(D3 \mid H1) + \\ & p(H2) \times p(D3 \mid H2) + \\ & p(H3) \times p(D3 \mid H3) \end{aligned}$$

$$p(D3 \mid H1) = \tfrac{1}{2}$$

$$p(D3 \mid H2) = 1$$

$$p(D3 \mid H3) = 0$$

$$p(H1) = p(H2) = p(H3) = \tfrac{1}{3}$$

$$p(D3) = \tfrac{1}{3}\left(\tfrac{1}{2} + 1 + 0\right)$$

Goats & Doors, with Bayes

$$p(H1 \mid D3) = \frac{p(D3 \mid H1) \times p(H1)}{p(D3)}$$

$$p(H1 \mid D3) = \frac{\frac{1}{2} \times \frac{1}{3}}{\frac{1}{2}} = \frac{1}{3}$$

and now compare to:

$$p(H2 \mid D3) = \frac{p(D3 \mid H2) \times p(H2)}{p(D3)}$$

Goats & Doors, with Bayes

$$p(H2 \mid D3) = \frac{p(D3 \mid H2) \times p(H2)}{p(D3)}$$

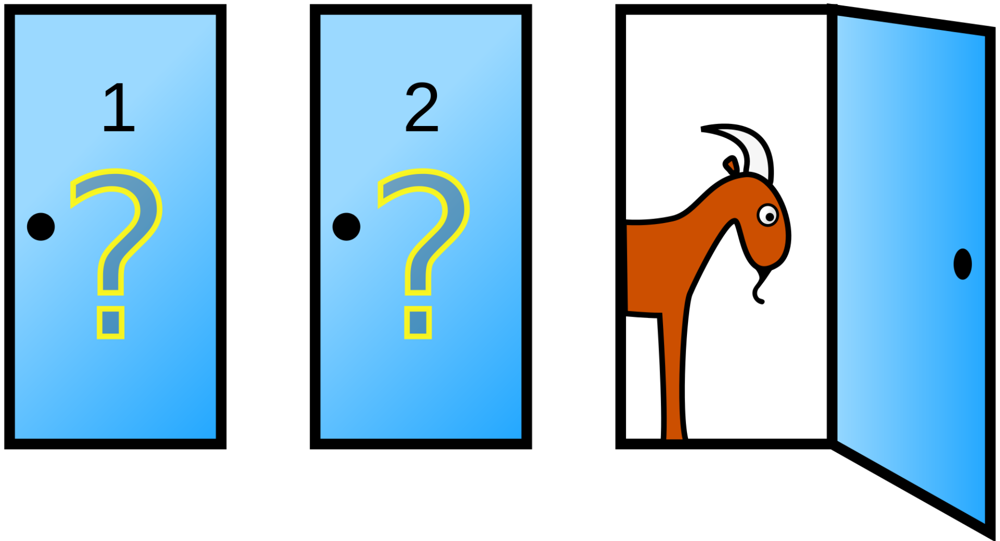
$p(D3 \mid H2) = 1$

$p(H2) = \text{ same as } p(H1) = (\frac{1}{3})$

$p(D3) =$ same as before (0.5)

so $p(H2 \mid D3) = \frac{1 \times \frac{1}{3}}{\frac{1}{2}} = \frac{2}{3}$

What door do you pick??

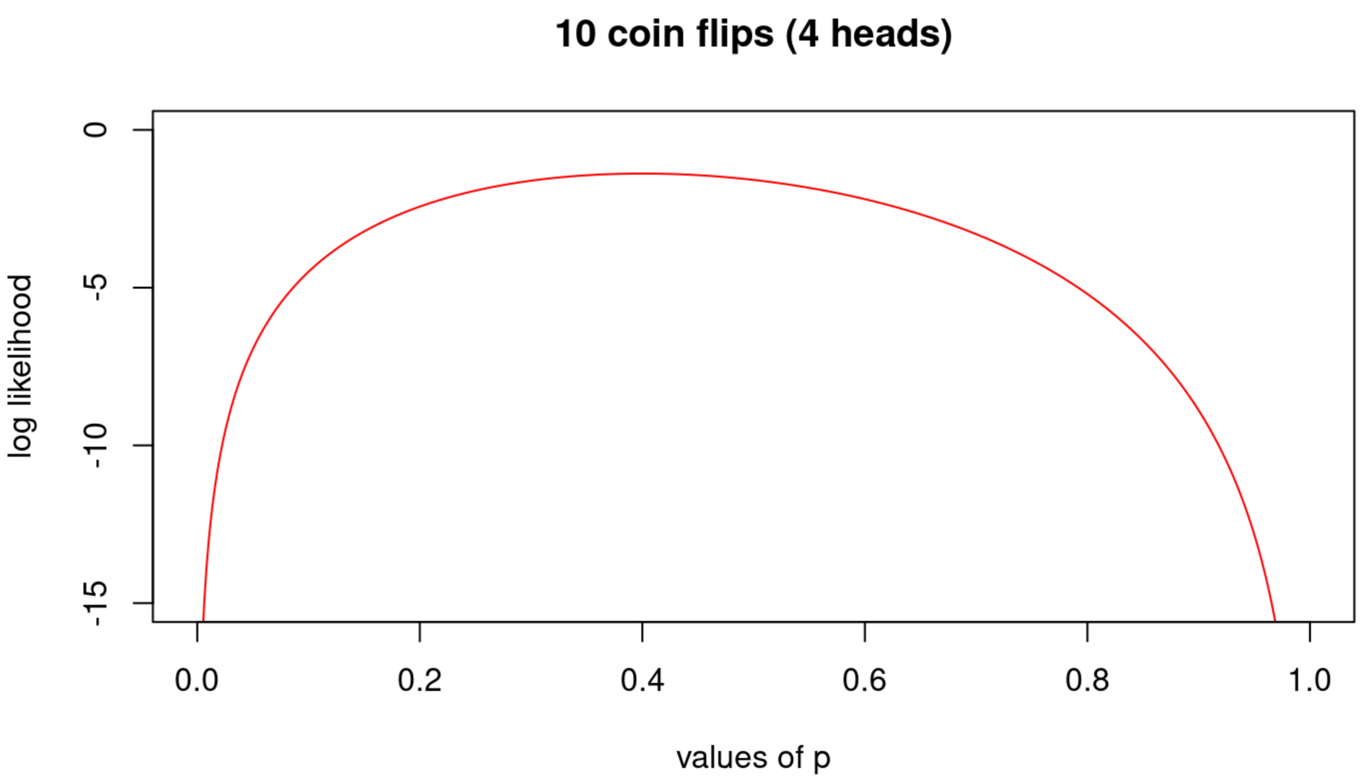


$$p(H1 \mid D3) = \frac{1}{3}$$

$$p(H2 \mid D3) = \frac{2}{3}$$

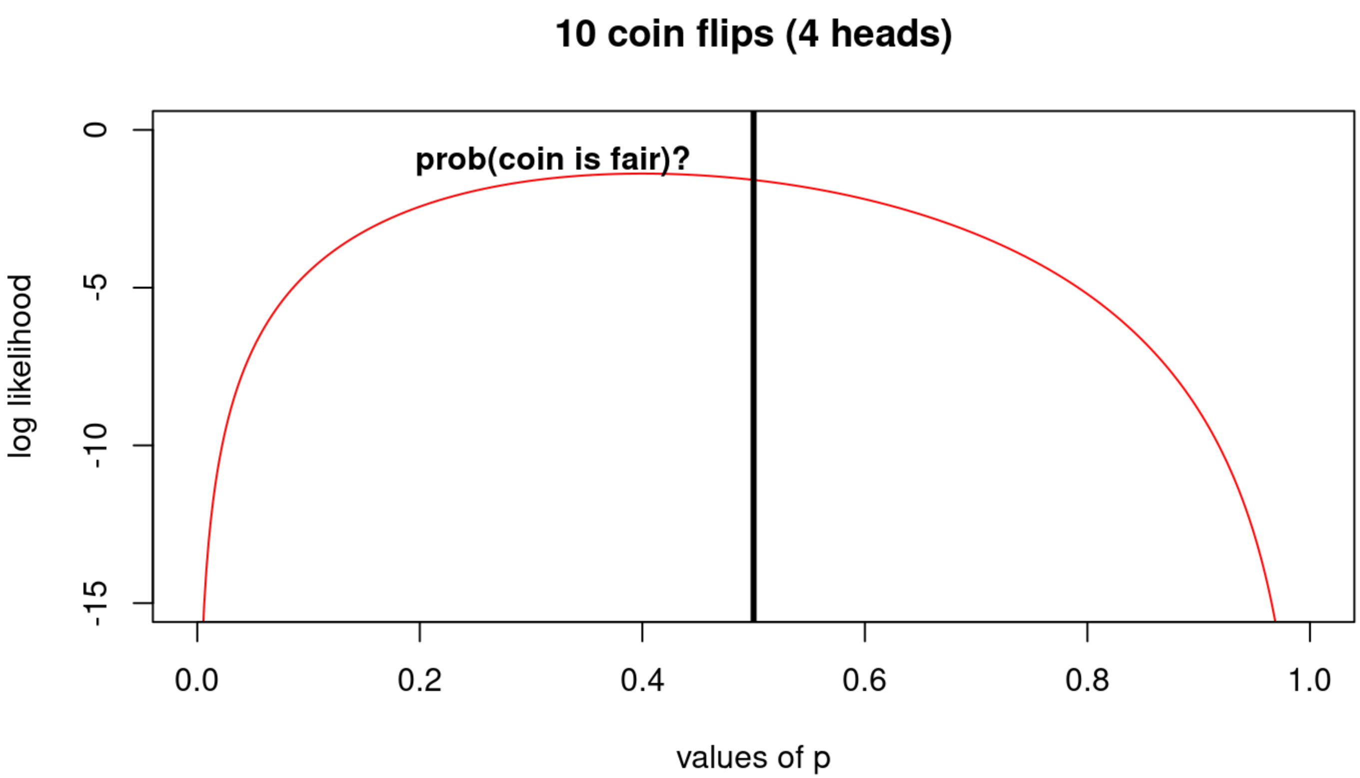
Again with the coins

We flip a coin 10 times. What is our likelihood of getting 4 heads if the coin is fair (p=0.5)?



Again with the coins

We flip a coin 10 times. What is our likelihood of getting 4 heads if the coin is fair (p=0.5)?



Prob(coin is fair)?

$$p(H \mid D) = \frac{p(H \mid D) \times p(H)}{p(D)}$$

$$p(p = 0.5 \mid \text{flips}) = \frac{p(\text{flips} | p{=}0.5) \times p(p{=}0.5)}{p(\text{flips})}$$

what is $p(p = 0.5)$?

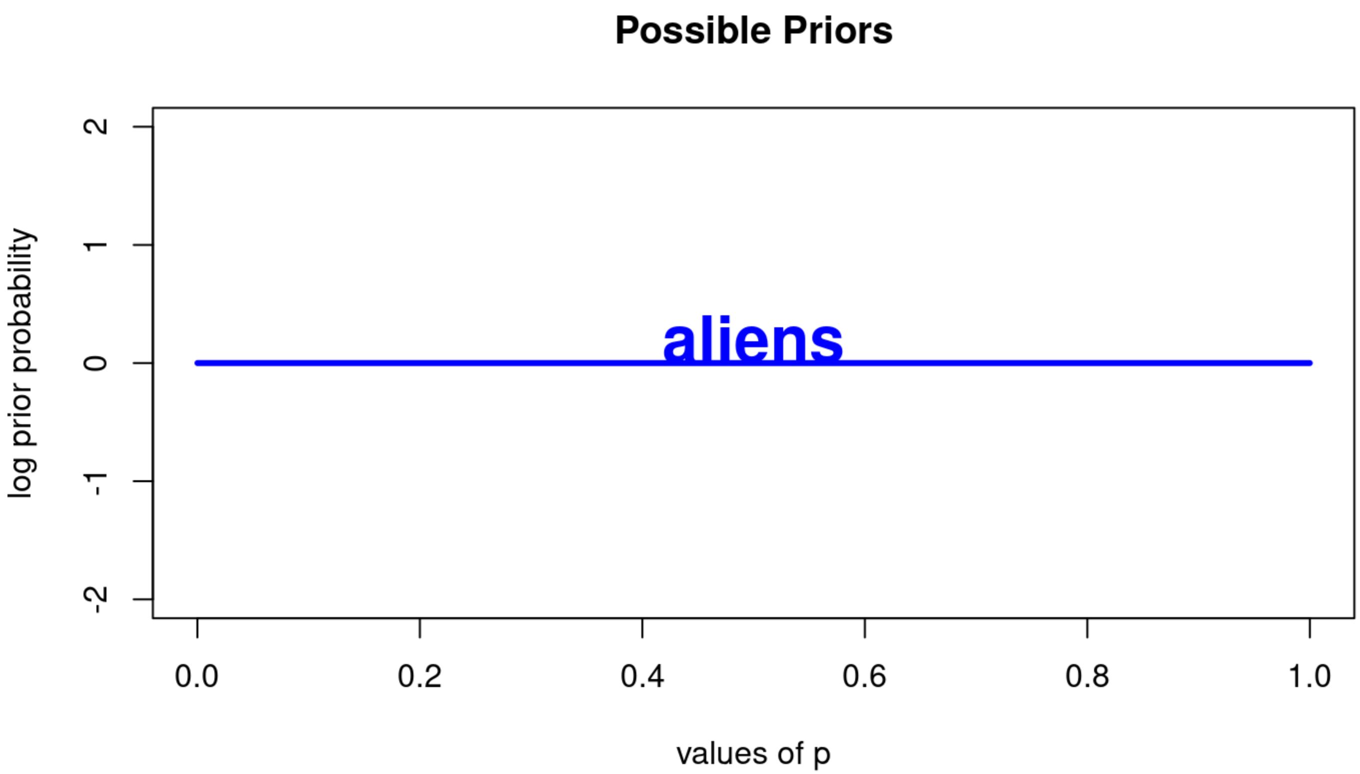
Prior probability distributions

We have to specify a prior probability of p=0.5.

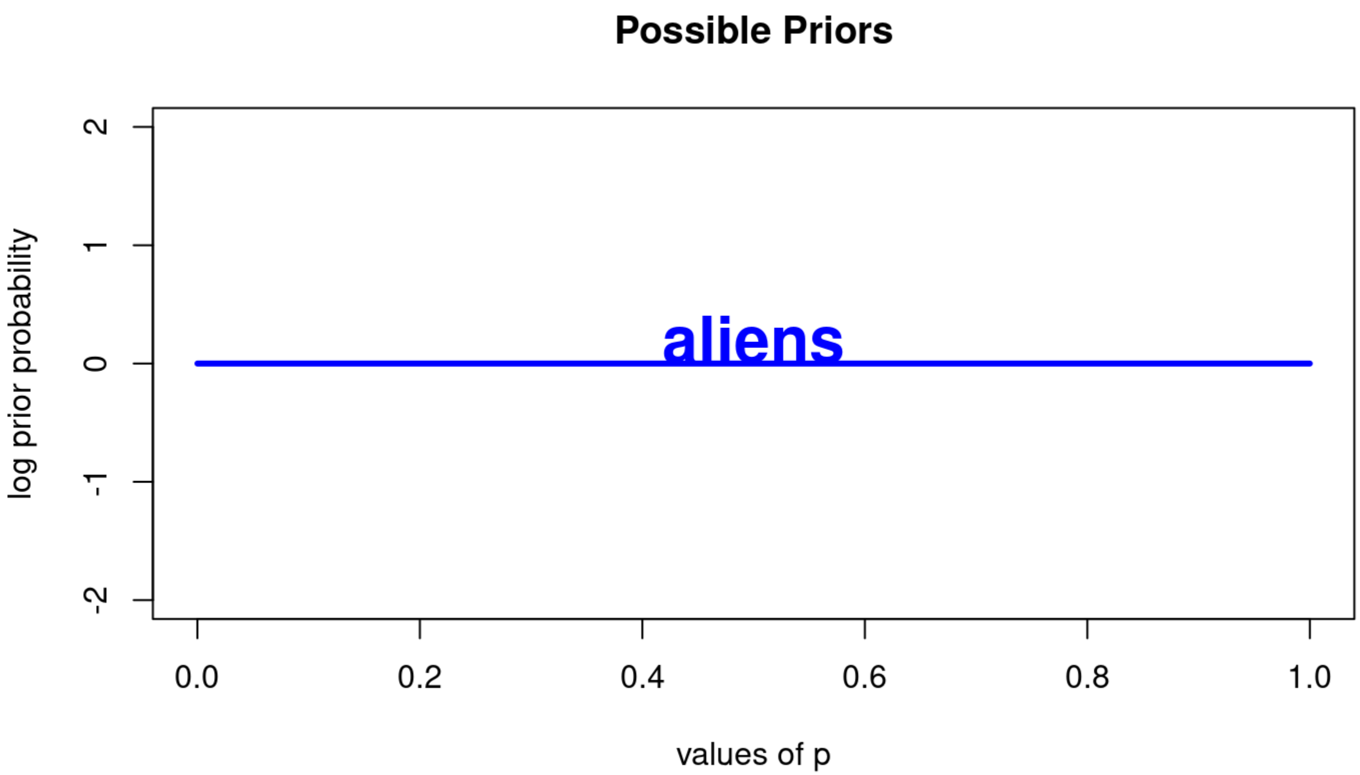
What prior should we specify if:

- we are aliens, and we’ve never seen a coin before?
- if we used to be professional coin-flippers, and this isn’t our first rodeo?
- if we know the coin-maker is a shifty dude?

Prior probability distributions

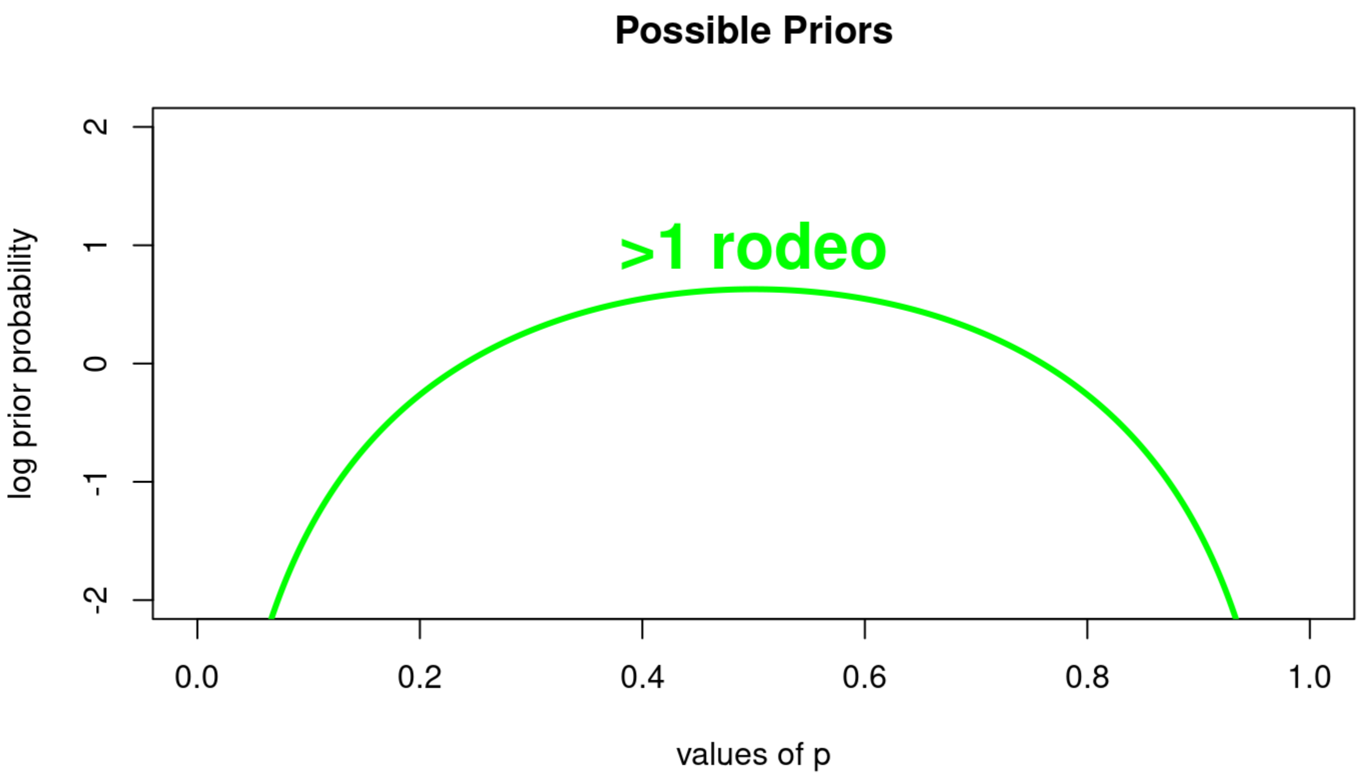


Prior probability distributions

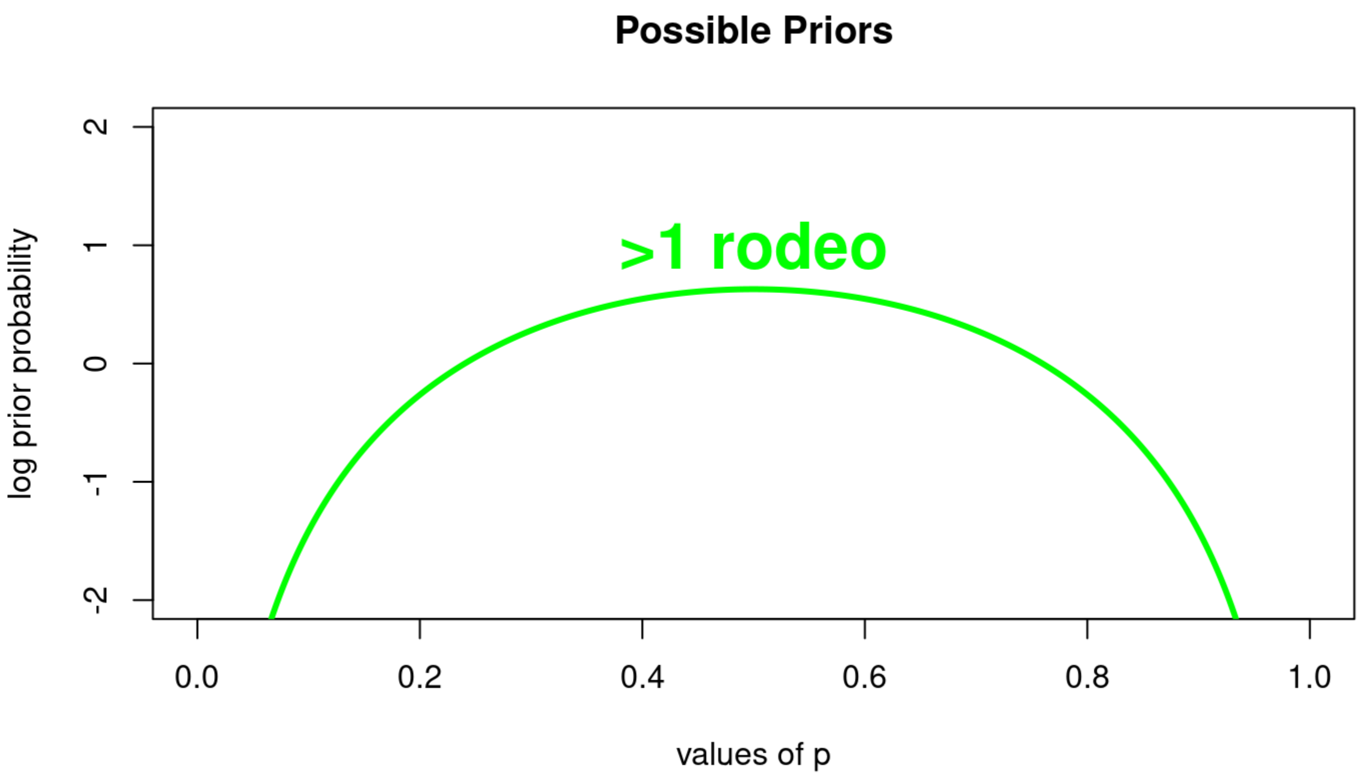


$$p \sim U(a = 0, b = 1)$$

Prior probability distributions

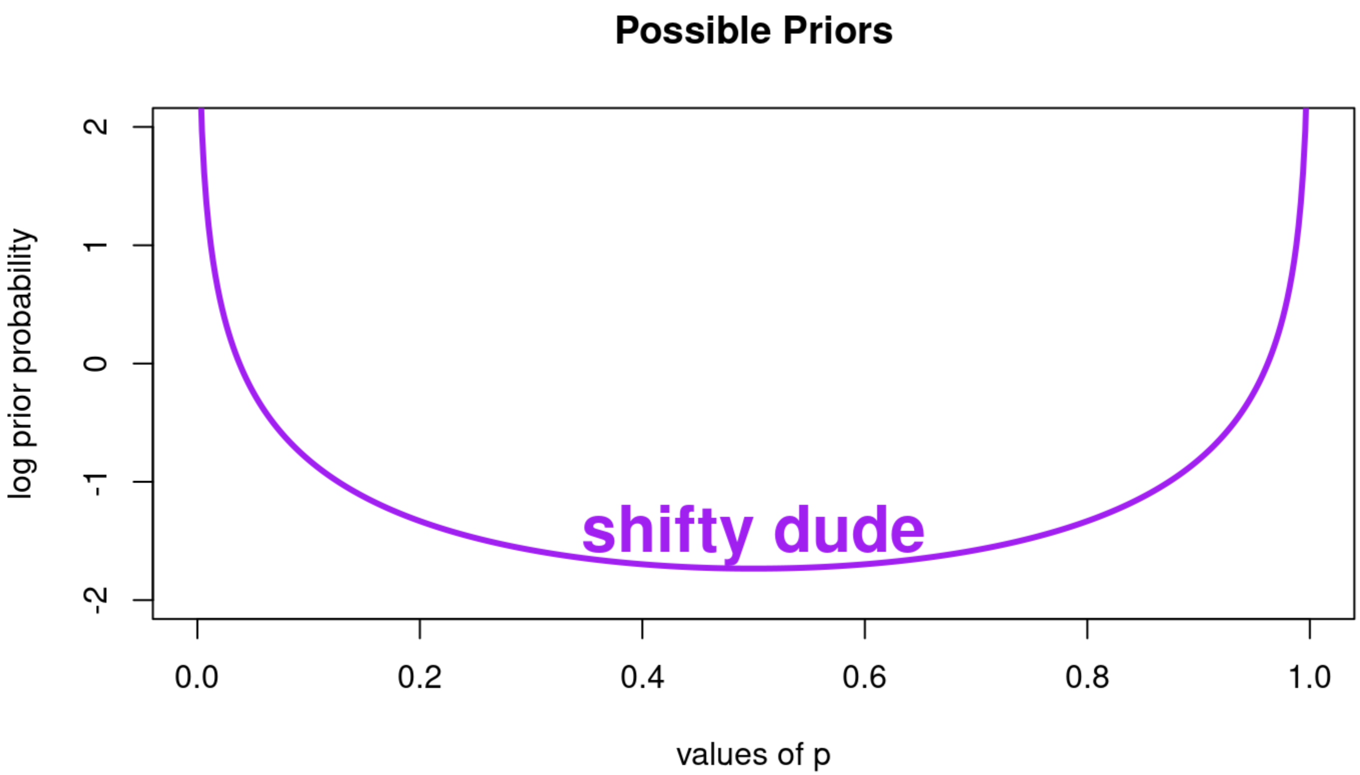


Prior probability distributions

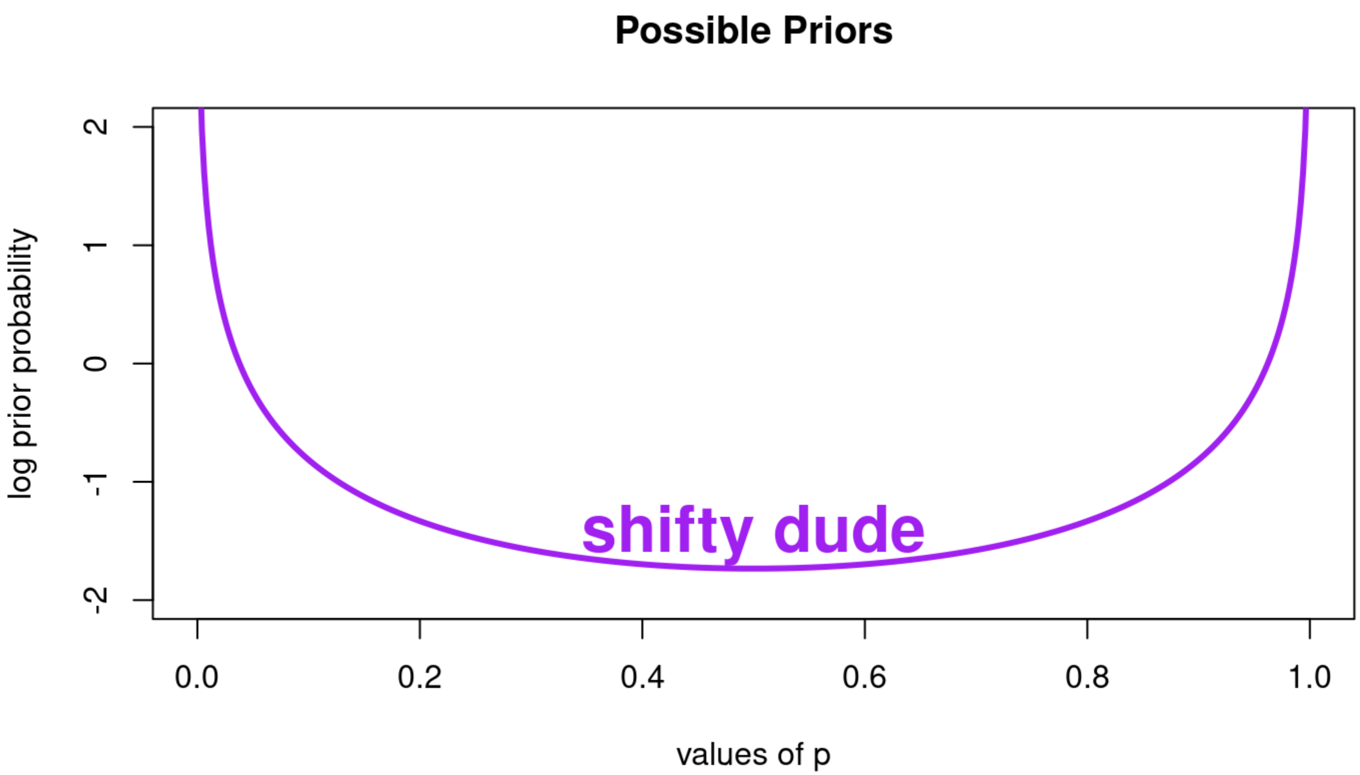


$$p \sim \mathrm{Beta}(\alpha = 3, \beta = 3)$$

Prior probability distributions

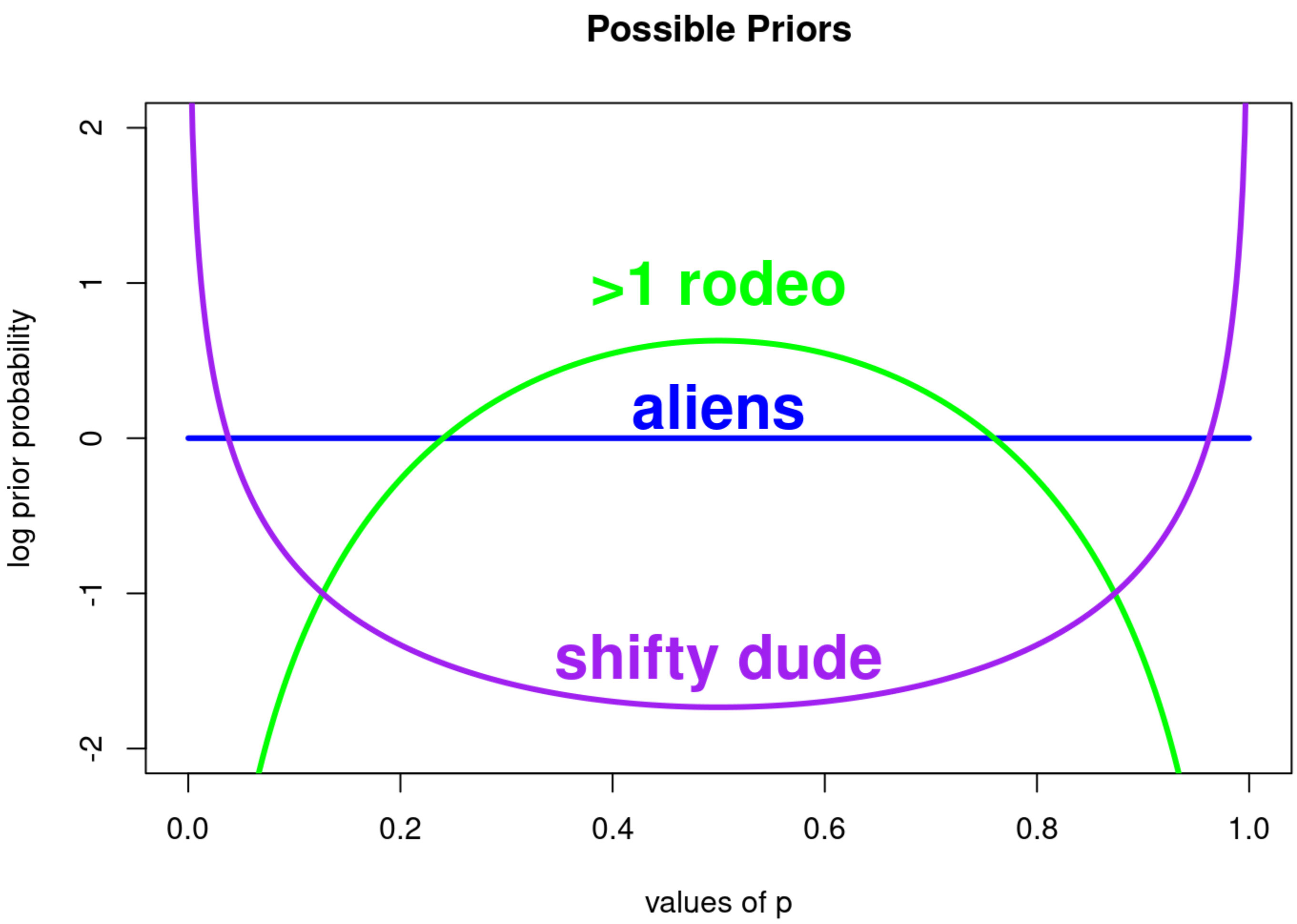


Prior probability distributions

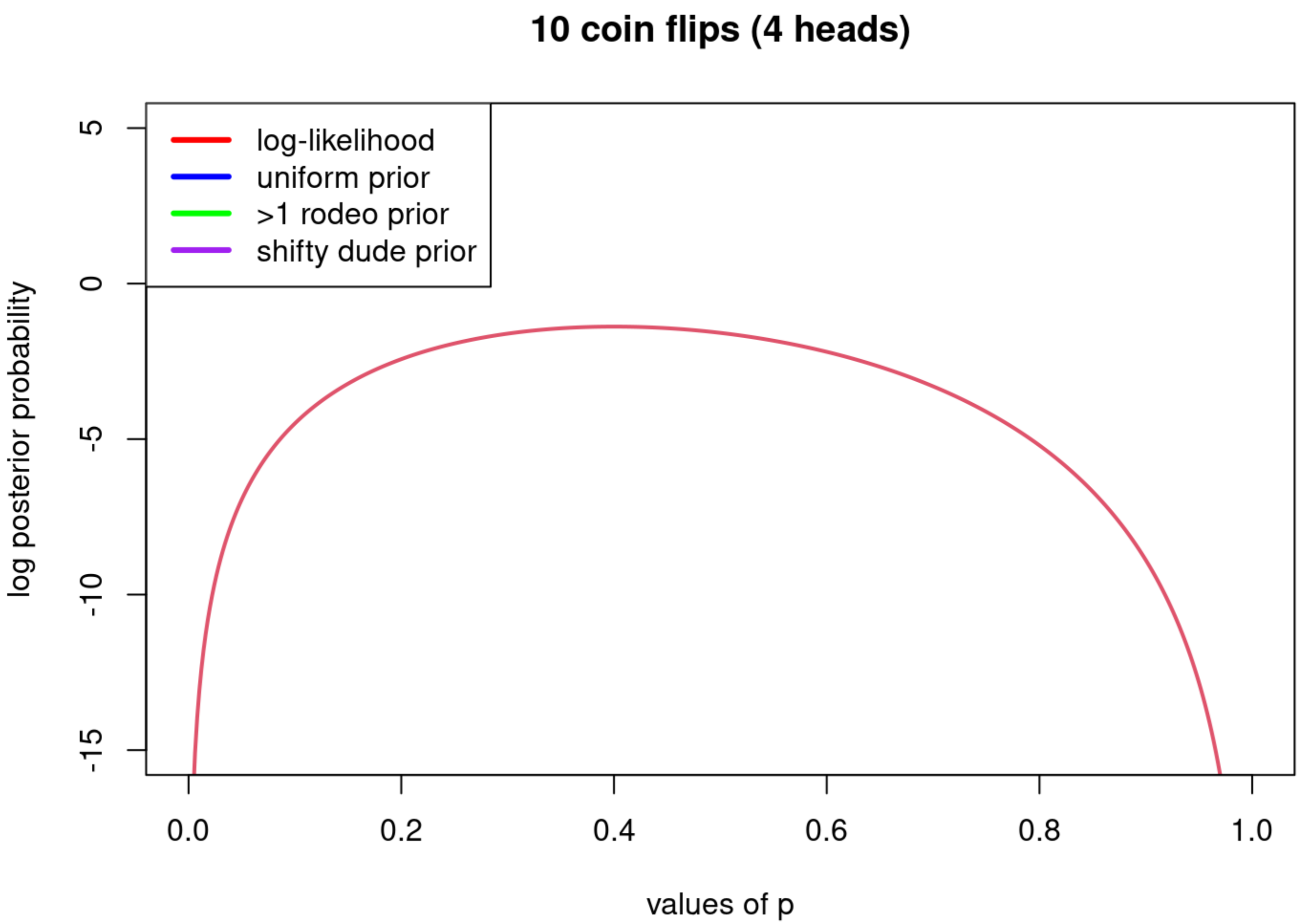


$$p \sim \mathrm{Beta}(\alpha = 0.1, \beta = 0.1)$$

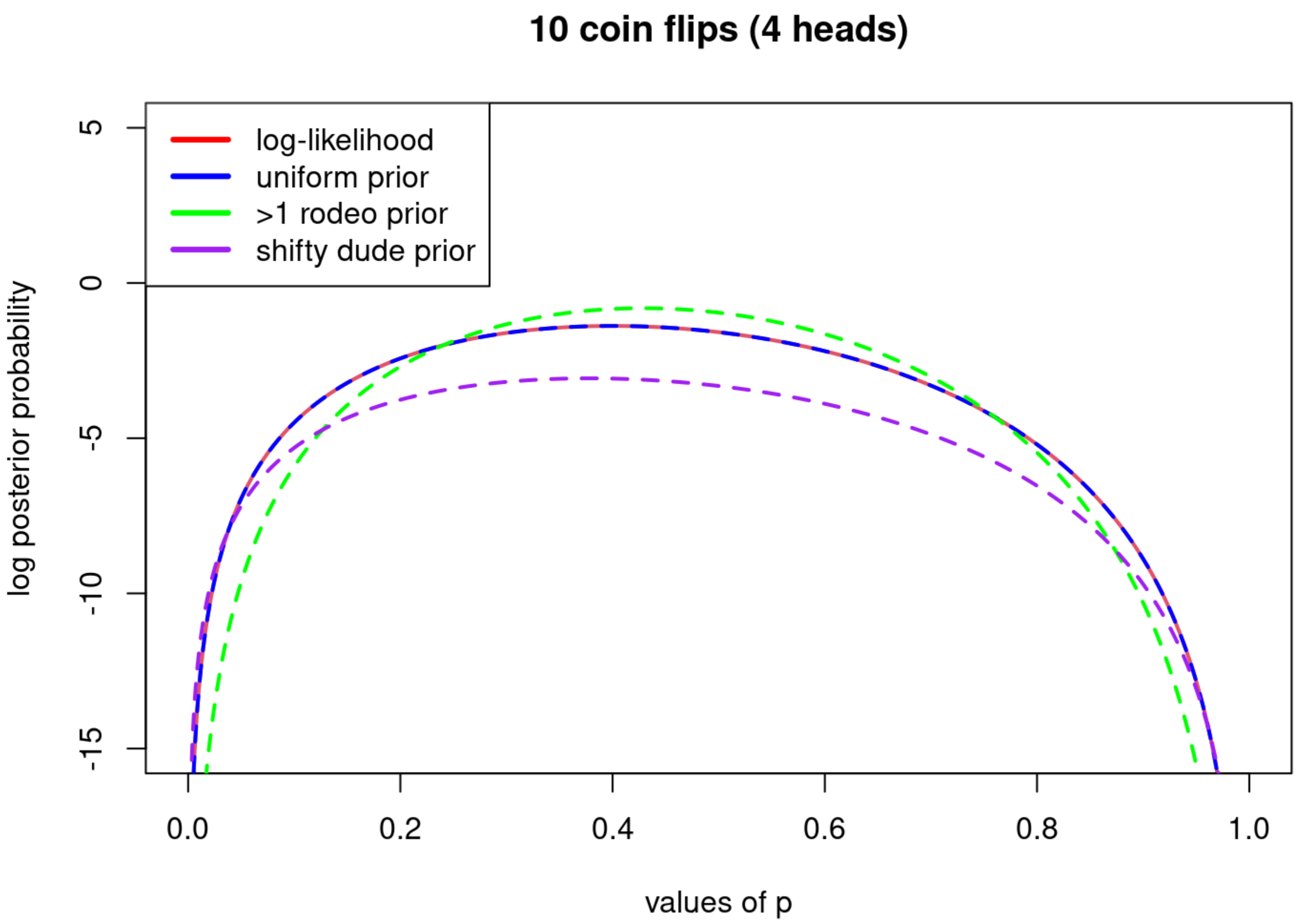
Prior probability distributions



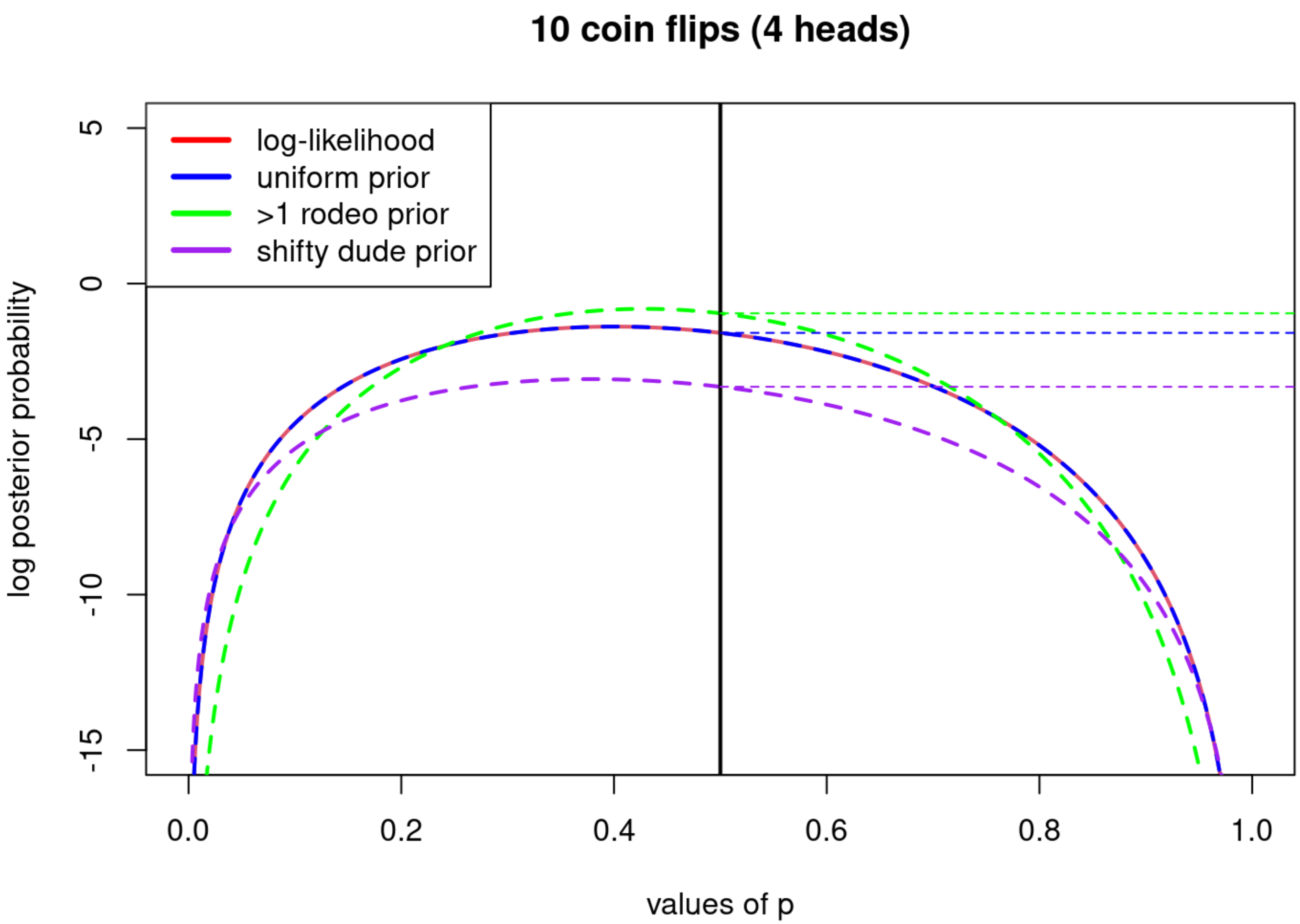
How do priors affect the posterior?



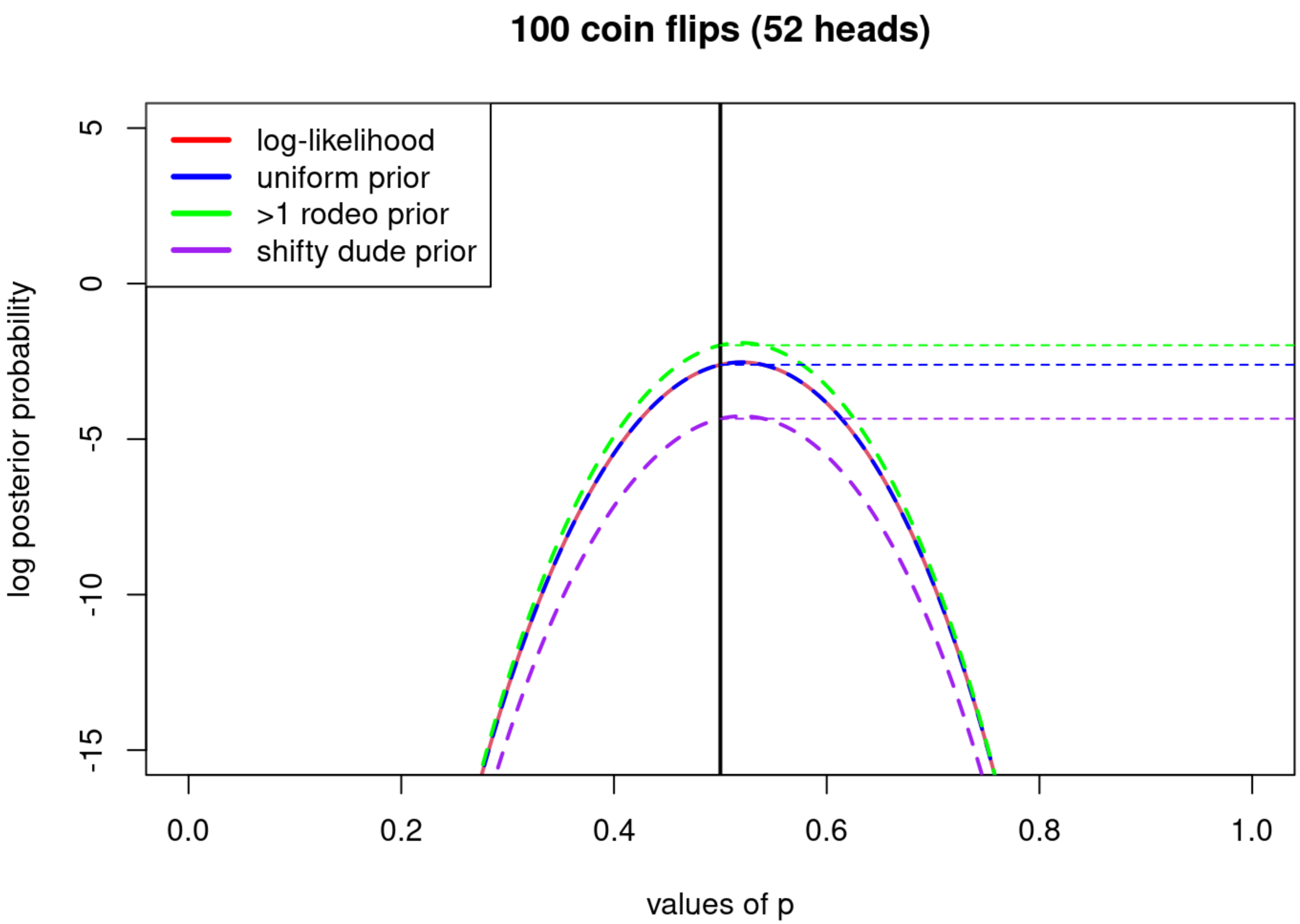
How do priors affect the posterior?



How do priors affect the posterior?



How do priors affect the posterior?

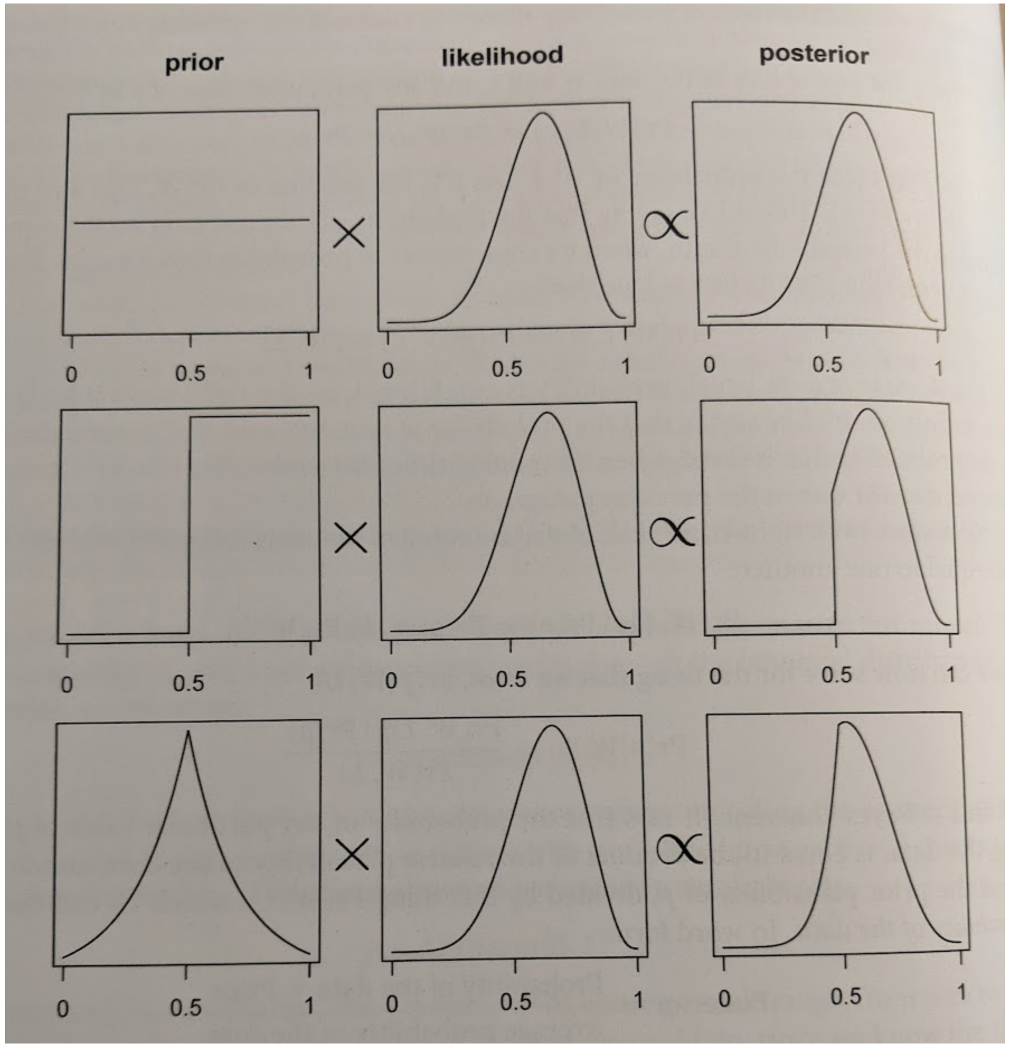


Prior to posterior

$$p(H \mid D) = \frac{p(D \mid H) \times p(H)}{p(D)}$$

All Bayes’ theorem is really doing is multiplying the likelihood by the prior (ignoring p(D))

Prior to posterior



Bayes Takeaways

$$p(H \mid D) = \frac{p(D \mid H) \times p(H)}{p(D)}$$

- the *prior* represents a belief based on previous information
- the *posterior probability* is an update of previous beliefs, based on new information
- the likelihood is the vehicle by which the data update the prior
- Bayes Theorem allows us to assess the probability of a hypothesis given some data (rather than the other way around)

Next time on Bayes of our lives

Doing inference in a Bayesian world!!