Bayes rule

Bayes rule

I've left talking about Bayes rule until now, because I think you can understand the concept of the Bayes factor without it, and because I wanted to emphasise the idea that the Bayes factor is a ratio of **two weighted averages**. However, now that we have this simple understanding, I'm hoping to deepen your understanding a bit by introducing Bayes rule. This deeper understanding of Bayes rule will also help use understand some of the topic we'll cover later in the course.

What is Bayes rule

Bayes rule follows straightforwardly from the axioms of conditional probability. In this sense, there's nothing particularly "Bayesian" about it in that both Frequentists and Bayesians can, and do, make use of the concept of conditional probability.

Conditional probability form

When you encounter Bayes rule in a frequentist context, it often takes the following form:

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

or

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

In this form, it typically explained by way of an example usually involving some kind of a test. In classic examples, the context if *often* a test for a rare disease. It is then shown that Bayes rule can be used to calculate the probability that the **positive** test indicates the **presence** of the disease [p(disease present| positive test)], by taking into account the **sensitivity** of the test [p(positive test | disease present)], the prevalence of the disease [p(disease)], and the probability of the test returning a positive result irrespective of the presence of the disease [p(positive)].

Bayes rule presented in this form is useful for thinking about evidence. The left side of the equation - $\frac{p(B|A)p(A)}{p(B)}$ - or more specifically, part of it - $\frac{p(B|A)}{p(B)}$ - can be read as representing the **evidence** the test provides or the presence of the disease. This **evidence** is then **weighted** by the **base rate** or the prior probability of the disease being present.

Proportional form

In the context of Bayesian inference, it is often given in a slightly different form:

$$p(A|B) \propto P(B|A) \cdot P(A)$$

or

$$p(\theta|Y) \propto \mathcal{L}(\theta|Y) \cdot p(\theta)$$

In this form it is usually read as "the posterior probability is proportional to the likelihood times the prior". The proportional form drops the denominator, which for a continuous parameter is given as:

$$p(Y) = \int_{\Theta} p(Y|\theta)p(\theta)d(\theta)$$

Integrals are generally difficult to work out, so they're often best avoided! We'll see in the section on parameter estimation that while it's not always possible to work out the posterior, we can just **draw samples from it** without needing to solve the integral.

Ratio form

Both of these forms, however, obscure the relationship between **Bayes** and **prediction**.

Following Rouder and Morey (2019), I think it's useful to present Bayes rule in the ratio form:

$$\frac{\pi(\theta|Y)}{\pi(\theta)} = \frac{p(Y|\theta)}{p(Y)}$$

The ratio form relates our "beliefs" about parameters $\frac{\pi(\theta|Y)}{\pi(\theta)}$ to probabilities about data $\frac{p(Y|\theta)}{p(Y)}$. Or put another way, it relates **beliefs** and **evidence** to **predictions**. To understand how this is the case, we'll examine the example given by Rouder and Morey (2019).

To explore this formula we'll first have to set two things. First, we'll need to set what our observation is—that is, our **data**. This will just be the number of heads (x) we've observed after n flips. The second thing we need to set if our **prior**. This is just the weights that we set in the previous section, and the **prior** represents our *"beliefs"* about plausible values for the parameter (in our case, the bias of the coin) **before** seeing the data (more on whether priors represent beliefs in the next section). We'll represent our prior with a **Beta** distribution, because this has some convenient mathematical properties (again, more on that in the next section). By changing the two parameters of the **Beta** distribution (α and β) you can assign more or less prior mass to the extreme (i.e., $\theta = 0$ and $\theta = 1$). When the values are the same, the distribution will be symmetrical and then they're different the distribution with be asymmetrical.

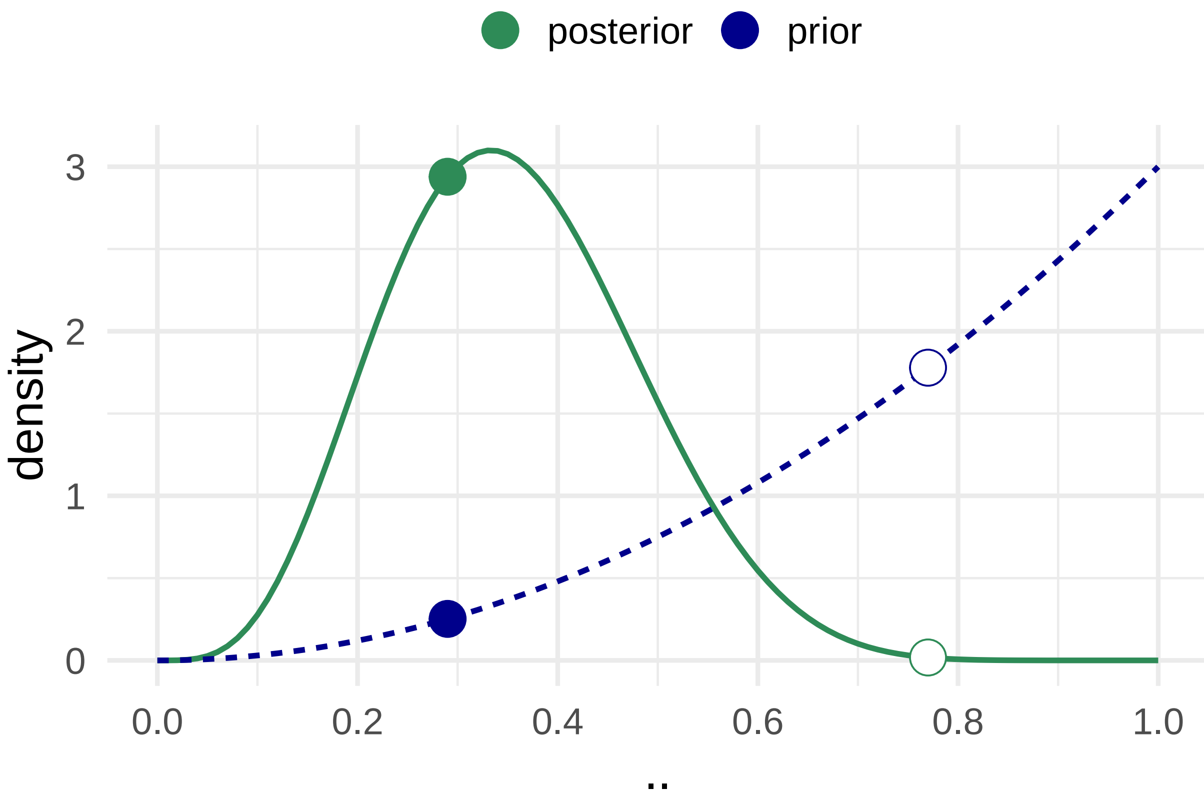
For our simple coin flip example, we'll just be able to calculate the posterior directly. This posterior represents what we believe about the parameter **after** seeing the data.

Our data is 2 heads in 10 flipsThe prior is a Beta(3, 1) distribution
The mean (expected value): $\theta = 0.75$
The mode (max probablity density): $\theta = 1$
The variance: 0.038

The posterior is a Beta(5, 9) distribution
The mean (expected value): $\theta = 0.36$
The mode (max probablity density): $\theta = 0.33$
The variance: 0.015

```
## Warning in grid.Call(C_textBounds, as.graphicsAnnot(x$label), x$x, x$y, :
## conversion failure on ' ' in 'mbcsToSbcs': dot substituted for <ce>

## Warning in grid.Call(C_textBounds, as.graphicsAnnot(x$label), x$x, x$y, :
## conversion failure on ' ' in 'mbcsToSbcs': dot substituted for <b8>

## Warning in grid.Call(C_textBounds, as.graphicsAnnot(x$label), x$x, x$y, :
## conversion failure on ' ' in 'mbcsToSbcs': dot substituted for <ce>
```

```
## Warning in grid.Call(C_textBounds, as.graphicsAnnot(x$label), x$x, x$y, :
## conversion failure on ' ' in 'mbcsToSbcs': dot substituted for <b8>

## Warning in grid.Call(C_textBounds, as.graphicsAnnot(x$label), x$x, x$y, :
## conversion failure on ' ' in 'mbcsToSbcs': dot substituted for <ce>

## Warning in grid.Call(C_textBounds, as.graphicsAnnot(x$label), x$x, x$y, :
## conversion failure on ' ' in 'mbcsToSbcs': dot substituted for <b8>

## Warning in grid.Call(C_textBounds, as.graphicsAnnot(x$label), x$x, x$y, :
## conversion failure on ' ' in 'mbcsToSbcs': dot substituted for <ce>

## Warning in grid.Call(C_textBounds, as.graphicsAnnot(x$label), x$x, x$y, :
## conversion failure on ' ' in 'mbcsToSbcs': dot substituted for <b8>

## Warning in grid.Call(C_textBounds, as.graphicsAnnot(x$label), x$x, x$y, :
## conversion failure on ' ' in 'mbcsToSbcs': dot substituted for <ce>

## Warning in grid.Call(C_textBounds, as.graphicsAnnot(x$label), x$x, x$y, :
## conversion failure on ' ' in 'mbcsToSbcs': dot substituted for <b8>

## Warning in grid.Call(C_textBounds, as.graphicsAnnot(x$label), x$x, x$y, :
## conversion failure on ' ' in 'mbcsToSbcs': dot substituted for <ce>

## Warning in grid.Call(C_textBounds, as.graphicsAnnot(x$label), x$x, x$y, :
## conversion failure on ' ' in 'mbcsToSbcs': dot substituted for <b8>

## Warning in grid.Call(C_textBounds, as.graphicsAnnot(x$label), x$x, x$y, :
## conversion failure on ' ' in 'mbcsToSbcs': dot substituted for <ce>

## Warning in grid.Call(C_textBounds, as.graphicsAnnot(x$label), x$x, x$y, :
## conversion failure on ' ' in 'mbcsToSbcs': dot substituted for <b8>

## Warning in grid.Call.graphics(C_text, as.graphicsAnnot(x$label), x$x, x$y, :
## conversion failure on ' ' in 'mbcsToSbcs': dot substituted for <ce>

## Warning in grid.Call.graphics(C_text, as.graphicsAnnot(x$label), x$x, x$y, :
## conversion failure on ' ' in 'mbcsToSbcs': dot substituted for <b8>
```

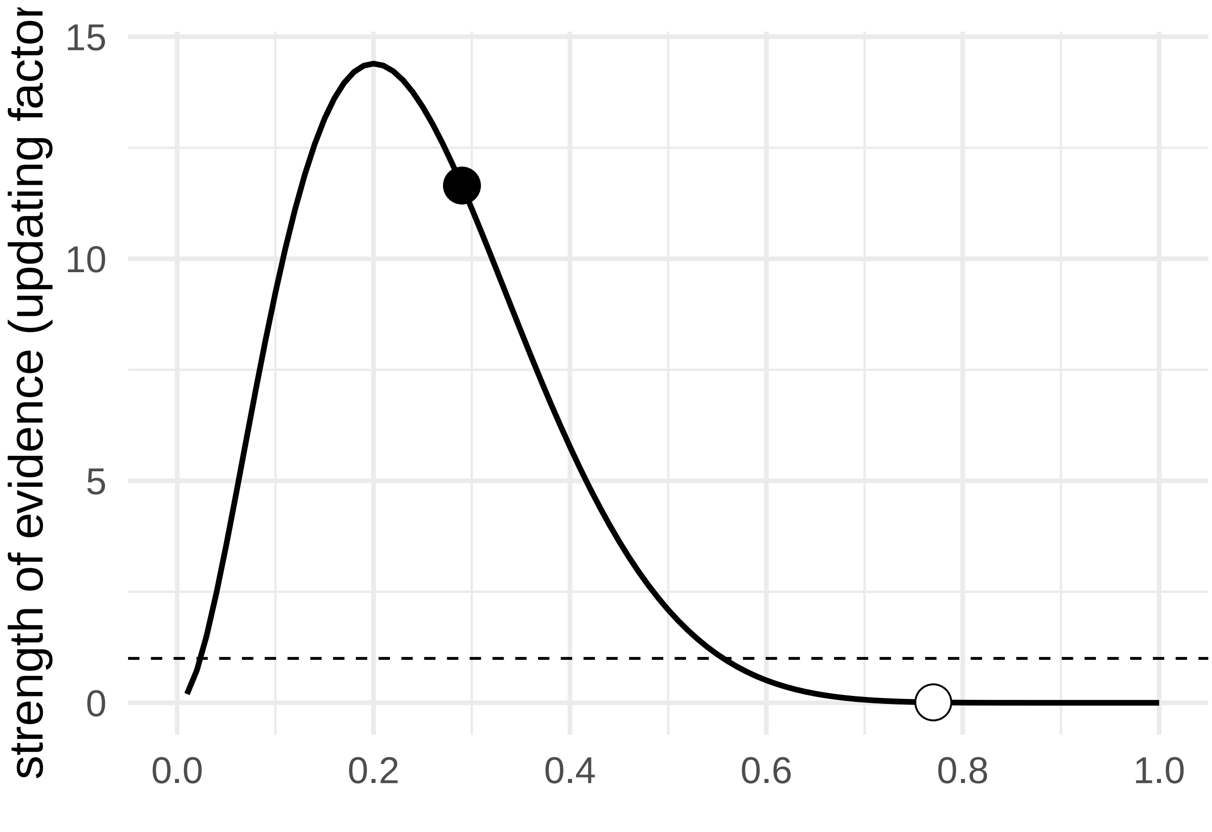


Once we plot the **prior** and the **posterior** together we'll see that for some values of θ seeing the data resulted in us *believing* that that value of θ is *more probable*. For other values, we now *believe* that that value of θ is *less probable* (in the plots, a value that is *less probable* after seeing the data is shown with empty point and a value that is *more probable* after seeing the data is shown with a filled point).

For each value of the parameter we can examine whether the data resulted in us believing that that value of the parameter is more or less probable. We can call this the **strength of evidence from the data about** θ. We can calculate this by just calculating the relative difference between the prior and the posterior—that is, by calculating $\frac{\pi(\theta|Y)}{\pi(\theta)}$.

```
## Warning in grid.Call(C_textBounds, as.graphicsAnnot(x$label), x$x, x$y, :
## conversion failure on ' ' in 'mbcsToSbcs': dot substituted for <ce>

## Warning in grid.Call(C_textBounds, as.graphicsAnnot(x$label), x$x, x$y, :
## conversion failure on ' ' in 'mbcsToSbcs': dot substituted for <b8>

## Warning in grid.Call(C_textBounds, as.graphicsAnnot(x$label), x$x, x$y, :
## conversion failure on ' ' in 'mbcsToSbcs': dot substituted for <ce>

## Warning in grid.Call(C_textBounds, as.graphicsAnnot(x$label), x$x, x$y, :
## conversion failure on ' ' in 'mbcsToSbcs': dot substituted for <b8>

## Warning in grid.Call(C_textBounds, as.graphicsAnnot(x$label), x$x, x$y, :
## conversion failure on ' ' in 'mbcsToSbcs': dot substituted for <ce>

## Warning in grid.Call(C_textBounds, as.graphicsAnnot(x$label), x$x, x$y, :
## conversion failure on ' ' in 'mbcsToSbcs': dot substituted for <b8>

## Warning in grid.Call(C_textBounds, as.graphicsAnnot(x$label), x$x, x$y, :
## conversion failure on ' ' in 'mbcsToSbcs': dot substituted for <ce>

## Warning in grid.Call(C_textBounds, as.graphicsAnnot(x$label), x$x, x$y, :
## conversion failure on ' ' in 'mbcsToSbcs': dot substituted for <b8>
```

```
## Warning in grid.Call(C_textBounds, as.graphicsAnnot(x$label), x$x, x$y, :
## conversion failure on ' ' in 'mbcsToSbcs': dot substituted for <ce>

## Warning in grid.Call(C_textBounds, as.graphicsAnnot(x$label), x$x, x$y, :
## conversion failure on ' ' in 'mbcsToSbcs': dot substituted for <b8>

## Warning in grid.Call(C_textBounds, as.graphicsAnnot(x$label), x$x, x$y, :
## conversion failure on ' ' in 'mbcsToSbcs': dot substituted for <ce>

## Warning in grid.Call(C_textBounds, as.graphicsAnnot(x$label), x$x, x$y, :
## conversion failure on ' ' in 'mbcsToSbcs': dot substituted for <b8>

## Warning in grid.Call(C_textBounds, as.graphicsAnnot(x$label), x$x, x$y, :
## conversion failure on ' ' in 'mbcsToSbcs': dot substituted for <ce>

## Warning in grid.Call(C_textBounds, as.graphicsAnnot(x$label), x$x, x$y, :
## conversion failure on ' ' in 'mbcsToSbcs': dot substituted for <b8>

## Warning in grid.Call.graphics(C_text, as.graphicsAnnot(x$label), x$x, x$y, :
## conversion failure on ' ' in 'mbcsToSbcs': dot substituted for <ce>

## Warning in grid.Call.graphics(C_text, as.graphicsAnnot(x$label), x$x, x$y, :
## conversion failure on ' ' in 'mbcsToSbcs': dot substituted for <b8>
```



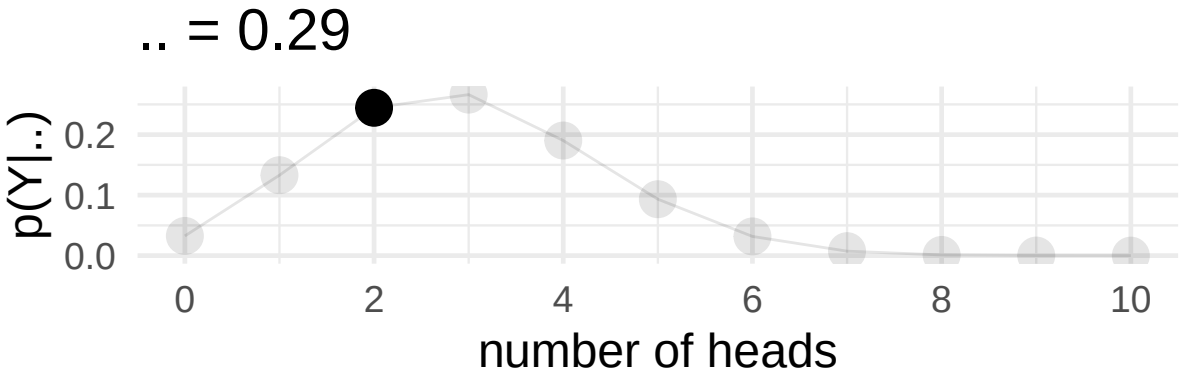
We can now turn our attention to the data and we can ask: "what is the probability of different observations **assuming** different values of θ?". This can be done with a simulation (like in our earlier examples); however, I know that it follows a **Binomial** distribution, so I can just generate it for different assumed values of θ.

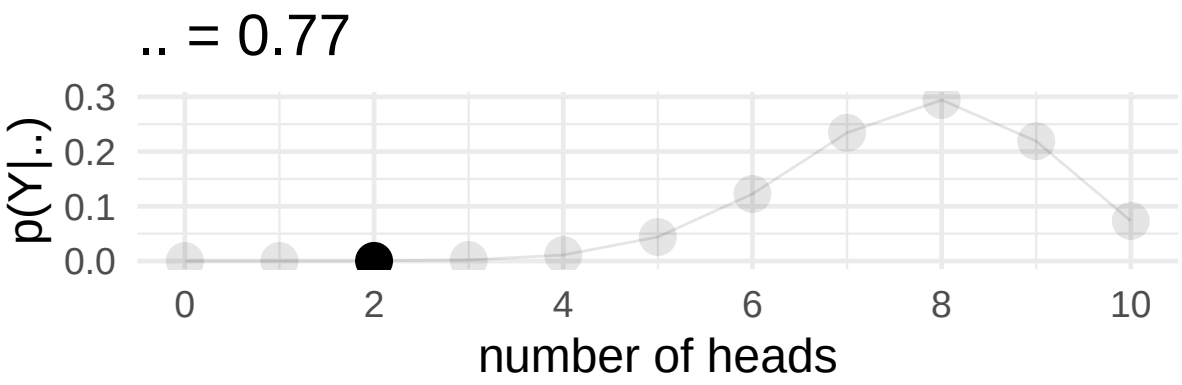
```
## Warning in grid.Call(C_textBounds, as.graphicsAnnot(x$label), x$x, x$y, :
## conversion failure on 'p(Y| )' in 'mbcsToSbcs': dot substituted for <ce>

## Warning in grid.Call(C_textBounds, as.graphicsAnnot(x$label), x$x, x$y, :
```

```
## conversion failure on 'p(Y| )' in 'mbcsToSbcs': dot substituted for <b8>

## Warning in grid.Call(C_textBounds, as.graphicsAnnot(x$label), x$x, x$y, :
## conversion failure on 'p(Y| )' in 'mbcsToSbcs': dot substituted for <ce>

## Warning in grid.Call(C_textBounds, as.graphicsAnnot(x$label), x$x, x$y, :
## conversion failure on 'p(Y| )' in 'mbcsToSbcs': dot substituted for <b8>

## Warning in grid.Call(C_textBounds, as.graphicsAnnot(x$label), x$x, x$y, :
## conversion failure on 'p(Y| )' in 'mbcsToSbcs': dot substituted for <ce>

## Warning in grid.Call(C_textBounds, as.graphicsAnnot(x$label), x$x, x$y, :
## conversion failure on 'p(Y| )' in 'mbcsToSbcs': dot substituted for <b8>

## Warning in grid.Call(C_textBounds, as.graphicsAnnot(x$label), x$x, x$y, :
## conversion failure on ' = 0.29' in 'mbcsToSbcs': dot substituted for <ce>

## Warning in grid.Call(C_textBounds, as.graphicsAnnot(x$label), x$x, x$y, :
## conversion failure on ' = 0.29' in 'mbcsToSbcs': dot substituted for <b8>

## Warning in grid.Call(C_textBounds, as.graphicsAnnot(x$label), x$x, x$y, :
## conversion failure on ' = 0.29' in 'mbcsToSbcs': dot substituted for <ce>

## Warning in grid.Call(C_textBounds, as.graphicsAnnot(x$label), x$x, x$y, :
## conversion failure on ' = 0.29' in 'mbcsToSbcs': dot substituted for <b8>

## Warning in grid.Call(C_textBounds, as.graphicsAnnot(x$label), x$x, x$y, :
## conversion failure on ' = 0.29' in 'mbcsToSbcs': dot substituted for <ce>

## Warning in grid.Call(C_textBounds, as.graphicsAnnot(x$label), x$x, x$y, :
## conversion failure on ' = 0.29' in 'mbcsToSbcs': dot substituted for <b8>

## Warning in grid.Call(C_textBounds, as.graphicsAnnot(x$label), x$x, x$y, :
## conversion failure on 'p(Y| )' in 'mbcsToSbcs': dot substituted for <ce>

## Warning in grid.Call(C_textBounds, as.graphicsAnnot(x$label), x$x, x$y, :
## conversion failure on 'p(Y| )' in 'mbcsToSbcs': dot substituted for <b8>

## Warning in grid.Call(C_textBounds, as.graphicsAnnot(x$label), x$x, x$y, :
## conversion failure on 'p(Y| )' in 'mbcsToSbcs': dot substituted for <ce>

## Warning in grid.Call(C_textBounds, as.graphicsAnnot(x$label), x$x, x$y, :
## conversion failure on 'p(Y| )' in 'mbcsToSbcs': dot substituted for <b8>

## Warning in grid.Call(C_textBounds, as.graphicsAnnot(x$label), x$x, x$y, :
## conversion failure on 'p(Y| )' in 'mbcsToSbcs': dot substituted for <ce>

## Warning in grid.Call(C_textBounds, as.graphicsAnnot(x$label), x$x, x$y, :
## conversion failure on 'p(Y| )' in 'mbcsToSbcs': dot substituted for <b8>

## Warning in grid.Call(C_textBounds, as.graphicsAnnot(x$label), x$x, x$y, :
## conversion failure on ' = 0.77' in 'mbcsToSbcs': dot substituted for <ce>

## Warning in grid.Call(C_textBounds, as.graphicsAnnot(x$label), x$x, x$y, :
## conversion failure on ' = 0.77' in 'mbcsToSbcs': dot substituted for <b8>

## Warning in grid.Call(C_textBounds, as.graphicsAnnot(x$label), x$x, x$y, :
## conversion failure on ' = 0.77' in 'mbcsToSbcs': dot substituted for <ce>

## Warning in grid.Call(C_textBounds, as.graphicsAnnot(x$label), x$x, x$y, :
## conversion failure on ' = 0.77' in 'mbcsToSbcs': dot substituted for <b8>

## Warning in grid.Call(C_textBounds, as.graphicsAnnot(x$label), x$x, x$y, :
## conversion failure on ' = 0.77' in 'mbcsToSbcs': dot substituted for <ce>
```

```
## Warning in grid.Call(C_textBounds, as.graphicsAnnot(x$label), x$x, x$y, :
## conversion failure on '  = 0.77' in 'mbcsToSbcs': dot substituted for <b8>

## Warning in grid.Call(C_textBounds, as.graphicsAnnot(x$label), x$x, x$y, :
## conversion failure on 'p(Y| )' in 'mbcsToSbcs': dot substituted for <ce>

## Warning in grid.Call(C_textBounds, as.graphicsAnnot(x$label), x$x, x$y, :
## conversion failure on 'p(Y| )' in 'mbcsToSbcs': dot substituted for <b8>

## Warning in grid.Call(C_textBounds, as.graphicsAnnot(x$label), x$x, x$y, :
## conversion failure on 'p(Y| )' in 'mbcsToSbcs': dot substituted for <ce>

## Warning in grid.Call(C_textBounds, as.graphicsAnnot(x$label), x$x, x$y, :
## conversion failure on 'p(Y| )' in 'mbcsToSbcs': dot substituted for <b8>

## Warning in grid.Call.graphics(C_text, as.graphicsAnnot(x$label), x$x, x$y, :
## conversion failure on 'p(Y| )' in 'mbcsToSbcs': dot substituted for <ce>

## Warning in grid.Call.graphics(C_text, as.graphicsAnnot(x$label), x$x, x$y, :
## conversion failure on 'p(Y| )' in 'mbcsToSbcs': dot substituted for <b8>

## Warning in grid.Call(C_textBounds, as.graphicsAnnot(x$label), x$x, x$y, :
## conversion failure on '  = 0.29' in 'mbcsToSbcs': dot substituted for <ce>

## Warning in grid.Call(C_textBounds, as.graphicsAnnot(x$label), x$x, x$y, :
## conversion failure on '  = 0.29' in 'mbcsToSbcs': dot substituted for <b8>

## Warning in grid.Call(C_textBounds, as.graphicsAnnot(x$label), x$x, x$y, :
## conversion failure on '  = 0.29' in 'mbcsToSbcs': dot substituted for <ce>

## Warning in grid.Call(C_textBounds, as.graphicsAnnot(x$label), x$x, x$y, :
## conversion failure on '  = 0.29' in 'mbcsToSbcs': dot substituted for <b8>

## Warning in grid.Call.graphics(C_text, as.graphicsAnnot(x$label), x$x, x$y, :
## conversion failure on '  = 0.29' in 'mbcsToSbcs': dot substituted for <ce>

## Warning in grid.Call.graphics(C_text, as.graphicsAnnot(x$label), x$x, x$y, :
## conversion failure on '  = 0.29' in 'mbcsToSbcs': dot substituted for <b8>

## Warning in grid.Call(C_textBounds, as.graphicsAnnot(x$label), x$x, x$y, :
## conversion failure on 'p(Y| )' in 'mbcsToSbcs': dot substituted for <ce>

## Warning in grid.Call(C_textBounds, as.graphicsAnnot(x$label), x$x, x$y, :
## conversion failure on 'p(Y| )' in 'mbcsToSbcs': dot substituted for <b8>

## Warning in grid.Call(C_textBounds, as.graphicsAnnot(x$label), x$x, x$y, :
## conversion failure on 'p(Y| )' in 'mbcsToSbcs': dot substituted for <ce>

## Warning in grid.Call(C_textBounds, as.graphicsAnnot(x$label), x$x, x$y, :
## conversion failure on 'p(Y| )' in 'mbcsToSbcs': dot substituted for <b8>

## Warning in grid.Call.graphics(C_text, as.graphicsAnnot(x$label), x$x, x$y, :
## conversion failure on 'p(Y| )' in 'mbcsToSbcs': dot substituted for <ce>

## Warning in grid.Call.graphics(C_text, as.graphicsAnnot(x$label), x$x, x$y, :
## conversion failure on 'p(Y| )' in 'mbcsToSbcs': dot substituted for <b8>

## Warning in grid.Call(C_textBounds, as.graphicsAnnot(x$label), x$x, x$y, :
## conversion failure on '  = 0.77' in 'mbcsToSbcs': dot substituted for <ce>

## Warning in grid.Call(C_textBounds, as.graphicsAnnot(x$label), x$x, x$y, :
## conversion failure on '  = 0.77' in 'mbcsToSbcs': dot substituted for <b8>
```

```
## Warning in grid.Call(C_textBounds, as.graphicsAnnot(x$label), x$x, x$y, :
## conversion failure on '  = 0.77' in 'mbcsToSbcs': dot substituted for <ce>

## Warning in grid.Call(C_textBounds, as.graphicsAnnot(x$label), x$x, x$y, :
## conversion failure on '  = 0.77' in 'mbcsToSbcs': dot substituted for <b8>

## Warning in grid.Call.graphics(C_text, as.graphicsAnnot(x$label), x$x, x$y, :
## conversion failure on '  = 0.77' in 'mbcsToSbcs': dot substituted for <ce>

## Warning in grid.Call.graphics(C_text, as.graphicsAnnot(x$label), x$x, x$y, :
## conversion failure on '  = 0.77' in 'mbcsToSbcs': dot substituted for <b8>
```





The next concept, $p(Y)$, or the **marginal probability**, is a slightly tricky concept: $p(Y)$ is the probability of observing our data independent of whatever value θ might take. Often this value is ignored, especially in the context of parameter estimation (as you'll see in later sections). In fact, this value isn't present in the "proportional" formulation of Bayes rule; however, understanding $p(Y)$ is extremely useful in the context of **Bayes factors**.

The **marginal probability** distribution/mass plot can be more readily conceptualised as the predictions a model ($\mathcal{M}_I$) makes about the data. We can generate this by seeing what data is predicted by each value of θ where θ itself has a probability distribution specified by $\pi(\theta)$. This concept is maybe easiest to understand when we consider a uniform prior where each value of θ is equally probably. Then we can ask, what is the probability of observing a specific outcome Y independent of the value of θ (or, averaged across all possible values of θ. This is just $\frac{1}{n}$, where n is the number of possible outcomes. In our coin flip example, there are 11 possible outcomes—0 heads, 1 head, 2 heads,… 10 heads. So $p(Y)$ would be $\frac{1}{11}$ for any outcome. Or phrased another way, we can say that, without knowing θ, but knowing that every value of θ is equally probably, we can predict that any observation, such as our specific observation, would occur with a probability of $\frac{1}{11}$. A very important thing to note about the *marginal probability distribution* is that it must sum to 1. We'll see in the example below, that for different priors ($\pi(\theta)$), the pattern see in the marginal distribution changes, but it always sums to 1. This means that when some observations become **more** probable, other observations must become **less** probable.

In the table below, you'll see how the **marginal probability** is calculated for each observation. The table just shows the calculation for our specific observation—that is, our $p(Y)$. Note that the accuracy of our estimate for $p(Y)$ depends on how many values of θ we average across. This means that for a uniform prior, the limit of our estimate will approach $\frac{1}{11}$ when the number of values of θ that we average across approaches infinity.

θ

X

N

$p(Y|\theta)$

$\pi(\theta)$

$p(Y, \theta, \pi)$

0.0

2

10

0.000

0.000

0

0.1

2

10

0.194

0.000

0

0.2

2

10

0.302

0.000

0

0.3

2

10

0.233

0.000

0

0.4

2

10
0.121
0.000
0
0.5
2
10
0.044
0.001
0
0.6
2
10
0.011
0.001
0
0.7
2
10
0.001
0.001
0
0.8
2
10
0.000
0.002
0
0.9
2
10
0.000
0.002
0
1.0
2

10

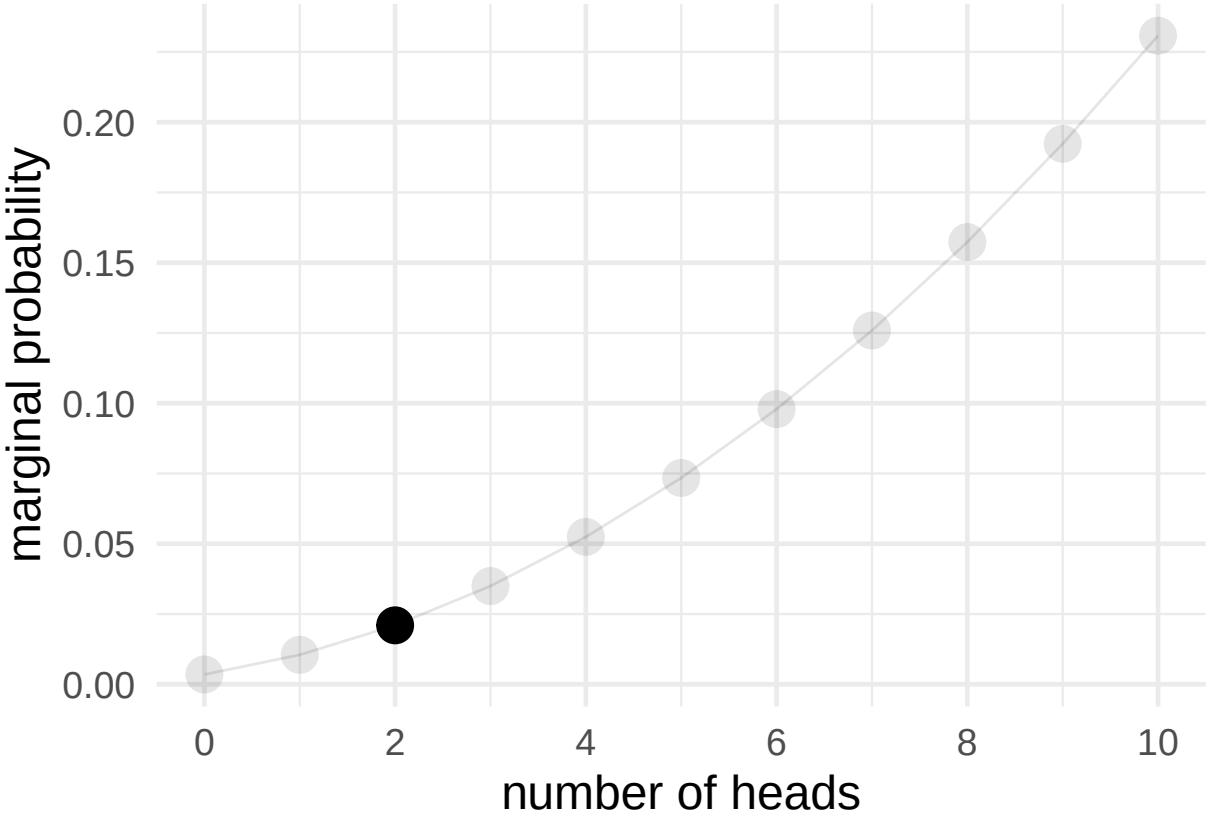
0.000

0.003

0

$p(Y) = 0.021$

The table just shows the marginal probability for our observation, but in the figure below we can plot the marginal distribution which considers every possible observation. This allows us to look of the entire range of possible observations and see which are more or less probable. These are the predictions our model makes.

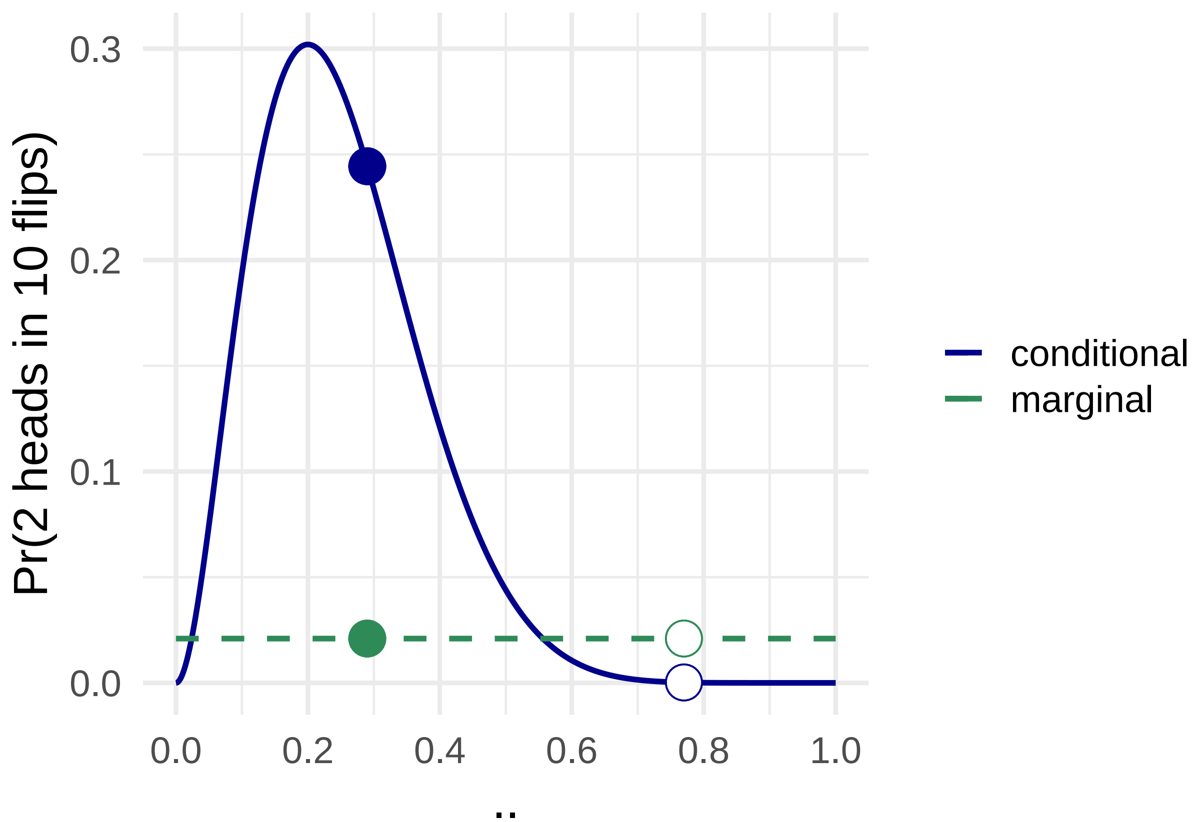


We can compare the marginal probability of our observation $p(Y)$ with the conditional probability $p(Y|\theta)$ — that is, conditional on a specific value of θ. The ratio of these two $\frac{p(Y|\theta)}{p(Y)}$ is the predictive accuracy for our data that gained by considering θ.

The following plot simply shows the conditional probability of the data give different values of the paramater (labelled **conditional**) and the marginal probability or the probability of the data irrespective of the value of the parameter (labelled **marginal**).

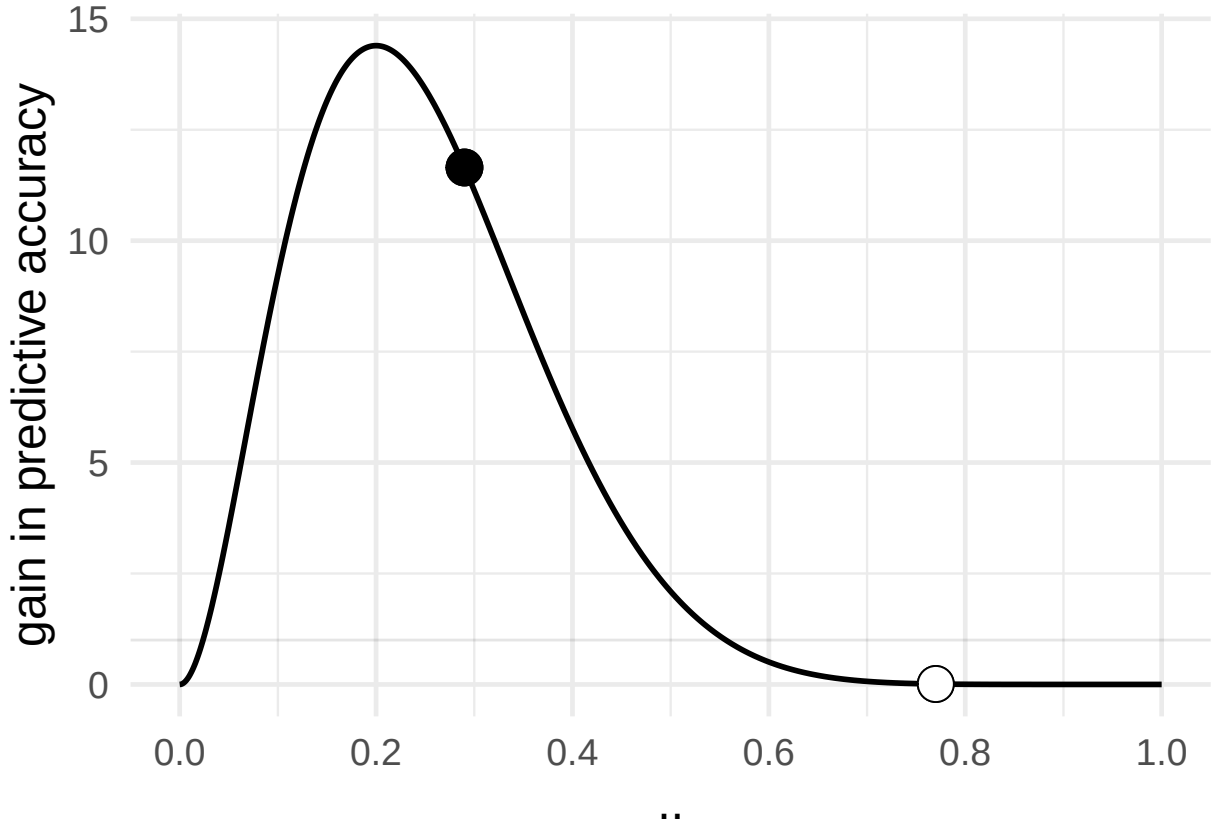
```
## Warning in grid.Call(C_textBounds, as.graphicsAnnot(x$label), x$x, x$y, :
## conversion failure on ' ' in 'mbcsToSbcs': dot substituted for <ce>

## Warning in grid.Call(C_textBounds, as.graphicsAnnot(x$label), x$x, x$y, :
## conversion failure on ' ' in 'mbcsToSbcs': dot substituted for <b8>

## Warning in grid.Call(C_textBounds, as.graphicsAnnot(x$label), x$x, x$y, :
## conversion failure on ' ' in 'mbcsToSbcs': dot substituted for <ce>

## Warning in grid.Call(C_textBounds, as.graphicsAnnot(x$label), x$x, x$y, :
## conversion failure on ' ' in 'mbcsToSbcs': dot substituted for <b8>
```

```
## Warning in grid.Call(C_textBounds, as.graphicsAnnot(x$label), x$x, x$y, :
## conversion failure on ' ' in 'mbcsToSbcs': dot substituted for <ce>

## Warning in grid.Call(C_textBounds, as.graphicsAnnot(x$label), x$x, x$y, :
## conversion failure on ' ' in 'mbcsToSbcs': dot substituted for <b8>

## Warning in grid.Call(C_textBounds, as.graphicsAnnot(x$label), x$x, x$y, :
## conversion failure on ' ' in 'mbcsToSbcs': dot substituted for <ce>

## Warning in grid.Call(C_textBounds, as.graphicsAnnot(x$label), x$x, x$y, :
## conversion failure on ' ' in 'mbcsToSbcs': dot substituted for <b8>

## Warning in grid.Call(C_textBounds, as.graphicsAnnot(x$label), x$x, x$y, :
## conversion failure on ' ' in 'mbcsToSbcs': dot substituted for <ce>

## Warning in grid.Call(C_textBounds, as.graphicsAnnot(x$label), x$x, x$y, :
## conversion failure on ' ' in 'mbcsToSbcs': dot substituted for <b8>

## Warning in grid.Call(C_textBounds, as.graphicsAnnot(x$label), x$x, x$y, :
## conversion failure on ' ' in 'mbcsToSbcs': dot substituted for <ce>

## Warning in grid.Call(C_textBounds, as.graphicsAnnot(x$label), x$x, x$y, :
## conversion failure on ' ' in 'mbcsToSbcs': dot substituted for <b8>

## Warning in grid.Call(C_textBounds, as.graphicsAnnot(x$label), x$x, x$y, :
## conversion failure on ' ' in 'mbcsToSbcs': dot substituted for <ce>

## Warning in grid.Call(C_textBounds, as.graphicsAnnot(x$label), x$x, x$y, :
## conversion failure on ' ' in 'mbcsToSbcs': dot substituted for <b8>

## Warning in grid.Call.graphics(C_text, as.graphicsAnnot(x$label), x$x, x$y, :
## conversion failure on ' ' in 'mbcsToSbcs': dot substituted for <ce>

## Warning in grid.Call.graphics(C_text, as.graphicsAnnot(x$label), x$x, x$y, :
## conversion failure on ' ' in 'mbcsToSbcs': dot substituted for <b8>
```




```
## Warning in grid.Call(C_textBounds, as.graphicsAnnot(x$label), x$x, x$y, :
## conversion failure on ' ' in 'mbcsToSbcs': dot substituted for <ce>

## Warning in grid.Call(C_textBounds, as.graphicsAnnot(x$label), x$x, x$y, :
## conversion failure on ' ' in 'mbcsToSbcs': dot substituted for <b8>

## Warning in grid.Call(C_textBounds, as.graphicsAnnot(x$label), x$x, x$y, :
## conversion failure on ' ' in 'mbcsToSbcs': dot substituted for <ce>

## Warning in grid.Call(C_textBounds, as.graphicsAnnot(x$label), x$x, x$y, :
## conversion failure on ' ' in 'mbcsToSbcs': dot substituted for <b8>

## Warning in grid.Call(C_textBounds, as.graphicsAnnot(x$label), x$x, x$y, :
## conversion failure on ' ' in 'mbcsToSbcs': dot substituted for <ce>

## Warning in grid.Call(C_textBounds, as.graphicsAnnot(x$label), x$x, x$y, :
## conversion failure on ' ' in 'mbcsToSbcs': dot substituted for <b8>

## Warning in grid.Call(C_textBounds, as.graphicsAnnot(x$label), x$x, x$y, :
## conversion failure on ' ' in 'mbcsToSbcs': dot substituted for <ce>

## Warning in grid.Call(C_textBounds, as.graphicsAnnot(x$label), x$x, x$y, :
## conversion failure on ' ' in 'mbcsToSbcs': dot substituted for <b8>

## Warning in grid.Call(C_textBounds, as.graphicsAnnot(x$label), x$x, x$y, :
## conversion failure on ' ' in 'mbcsToSbcs': dot substituted for <ce>

## Warning in grid.Call(C_textBounds, as.graphicsAnnot(x$label), x$x, x$y, :
## conversion failure on ' ' in 'mbcsToSbcs': dot substituted for <b8>

## Warning in grid.Call(C_textBounds, as.graphicsAnnot(x$label), x$x, x$y, :
## conversion failure on ' ' in 'mbcsToSbcs': dot substituted for <ce>
```

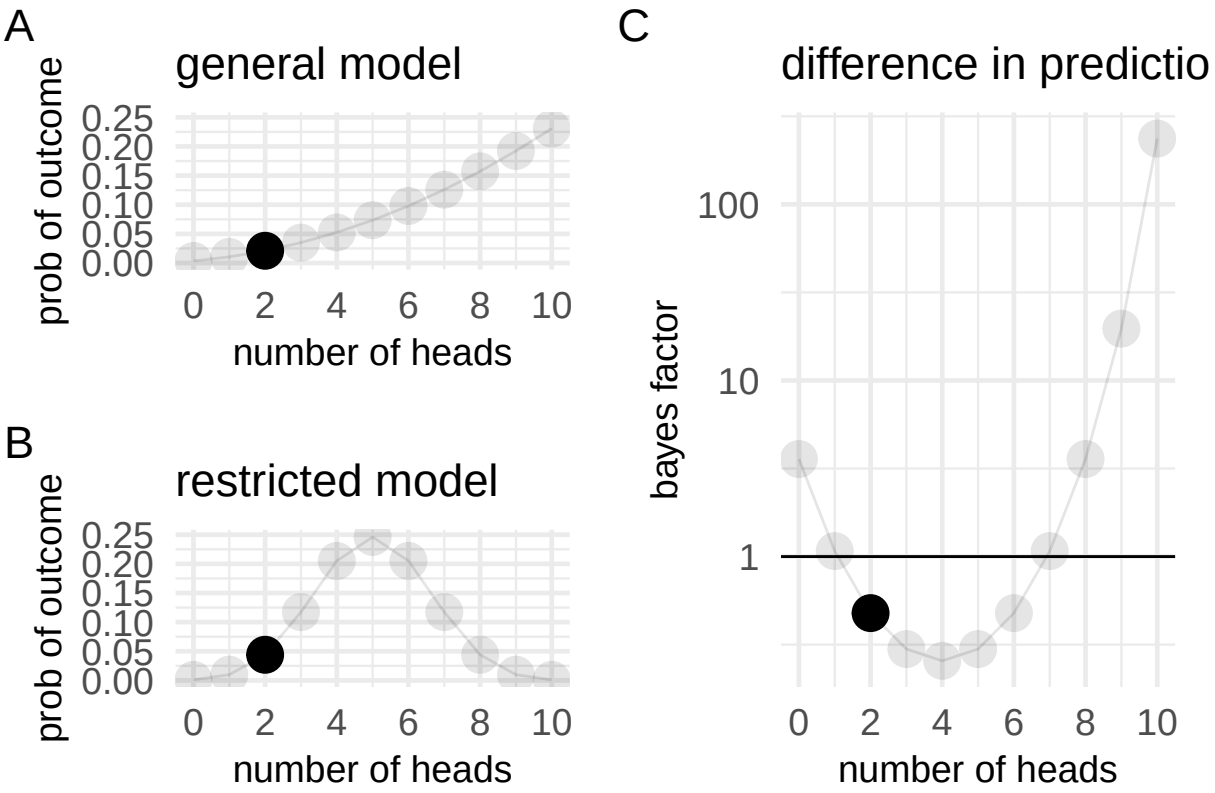
```
## Warning in grid.Call(C_textBounds, as.graphicsAnnot(x$label), x$x, x$y, :
## conversion failure on ' ' in 'mbcsToSbcs': dot substituted for <b8>

## Warning in grid.Call(C_textBounds, as.graphicsAnnot(x$label), x$x, x$y, :
## conversion failure on ' ' in 'mbcsToSbcs': dot substituted for <ce>

## Warning in grid.Call(C_textBounds, as.graphicsAnnot(x$label), x$x, x$y, :
## conversion failure on ' ' in 'mbcsToSbcs': dot substituted for <b8>

## Warning in grid.Call.graphics(C_text, as.graphicsAnnot(x$label), x$x, x$y, :
## conversion failure on ' ' in 'mbcsToSbcs': dot substituted for <ce>

## Warning in grid.Call.graphics(C_text, as.graphicsAnnot(x$label), x$x, x$y, :
## conversion failure on ' ' in 'mbcsToSbcs': dot substituted for <b8>
```



This plot **is just the same as the strength of evidence** for values of θ or the factor by which we update our beliefs about θ after observing the data. This fact is just represented by the equality in the ratio form of Bayes rule $\frac{\pi(\theta|Y)}{\pi(\theta)} = \frac{p(Y|\theta)}{p(Y)}$. This equation can now we read as meaning that the strength of evidence that we have for a parameter value is just the same as the gain in predictive accuracy.

Bayes factor

In this example, we've only considered one model defined by the prior we set at the beginning. However, marginal densities are particularly useful when we consider multiple models. In the next example, we plot the marginal density for our current model ($\mathcal{M}_1$; subplot **A**) and more restricted model where we no longer a probability distribution over every possible value of θ, but instead only consider one possible value, $\theta = 0.5$ ($\mathcal{M}_2$; subplot **A**). The difference in predictions the models make is shown in subplot **C**. This plot is just generated as the ratio $\frac{p(Y|\mathcal{M}_1)}{p(Y|\mathcal{M}_2)}$. Once we have our data in hand, we can see whether our data is better predicted by Model 1 or Model 2—this value is the **Bayes factor**.



heads	flips	Y	$p(Y\|\mathcal{M}_1)$	$p(Y\|\mathcal{M}_0)$	BF_{10}	BF_{01}
0	10	0 in 10	0.00	0.00	3.58	0.28
1	10	1 in 10	0.01	0.01	1.07	0.93
2	10	2 in 10	0.02	0.04	0.48	2.09
3	10	3 in 10	0.03	0.12	0.30	3.35
4	10	4 in 10	0.05	0.21	0.26	3.91
5	10	5 in 10	0.07	0.25	0.30	3.35
6	10	6 in 10	0.10	0.21	0.48	2.09
7	10	7 in 10	0.13	0.12	1.07	0.93
8	10	8 in 10	0.16	0.04	3.58	0.28
9	10	9 in 10	0.19	0.01	19.69	0.05
10	10	10 in 10	0.23	0.00	236.31	0.00