Bayesian Calculations: Simulation of Coin Flipping

Andy Grogan-Kaylor

July 08, 2019

Contents

1 *Background*

Bayes Theorem allows us to state our *prior* beliefs, to calculate the *likelihood* of our data given those beliefs, and then to update those beliefs with data, thus arriving at a set of *posterior* beliefs. However, Bayesian calculations can be difficult to understand. This document attempts to provide a simple walkthrough of some Bayesian calculations.

2 *Bayes Rule*

A draft interactive version of these ideas can be found here.

Mathematically Bayes Theorem is as follows:

$$P(H|D) = \frac{P(D|H)P(H)}{P(D)}$$

In words, Bayes Theorem may be written as follows:

$$posterior = \frac{likelihood * prior}{data}$$

Our posterior beliefs are proportional to our prior beliefs, multiplied by the likelihood of those beliefs, given the data.

3 *This Example*

In this example, we provide an example of using Bayes Theorem to examine our conclusions about the proportion of heads when a coin is flipped 10 times.

Conventionally, we call this proportion that we are trying to estimate θ.

For the sake of simplicity, this example uses a relatively simple set of prior beliefs about 3 possible values for the proportion θ.

R code in this example is adapted and simplified from Kruschke (2011), p. 70

4 *Prior*

We set a simple set of prior beliefs, concerning 3 values of θ, the proportion of heads.

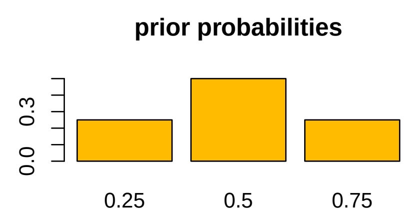
```
theta1 <- c(0.25, 0.5, 0.75)  # candidate parameter values

ptheta1 <- c(0.25, 0.5, 0.25)  # prior probabilities

ptheta1 <- ptheta1/sum(ptheta1)  # normalize
```

Our values of θ are *0.25*, *0.5* and *0.75*, with probabilities $P(\theta)$ of *0.25*, *0.5* and *0.25*.

```
barplot(names.arg = theta1, height = ptheta1,
    main = "prior probabilities", col = "#FFBB00")  # graph
```




```
myBayesianEstimates <- tibble(theta1, ptheta1)

pander(myBayesianEstimates)  # nice table
```

theta1	ptheta1
0.25	0.25
0.5	0.5
0.75	0.25

5 *The Data*

10 coin flips. 1 Heads. 9 Tails.

```
data1 <- c(1, 0, 0, 0, 0, 0, 0, 0, 0, 0)  # the data

data1_factor <- factor(data1, levels = c(0, 1),
    labels = c("T", "H"))

n_heads <- sum(data1 == 1)  # number of heads

n_tails <- sum(data1 == 0)  # number of tails

plot(x = seq(1,10), # x from 1 to 10
     y = rep(1,10), # y = 1
     col = data1_factor, # color by data
     cex = 3, # size
     lwd = 2, # line width
     axes = FALSE,
     main = "10 Coin Tosses",
     xlab = " ",
     ylab = " ")

text(x = seq(1,10), # x from 1 to 10
     y = rep(1,10), # y = 1
     labels = data1_factor) # label with data1_factor
```

6 *Likelihood*

The likelihood is the probability that a given value of θ would produce this number of heads.

The probability of multiple independent events A, B, C, etc. is $P(A, B, C, ...) = P(A) * P(B) * P(C) * ....$

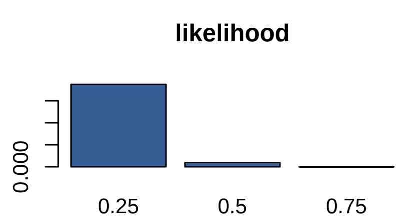
Therefore, in this case, the likelihood is $[P(heads)]^{\text{number of heads}}$ and multiply this by $[P(tails)]^{\text{number of tails}}$.

Thus:

$$\mathcal{L}(\theta) = \theta^{\text{number of heads}} * (1 - \theta)^{\text{number of tails}}$$

```
likelihood1 <- theta1^n_heads * (1 - theta1)^n_tails  # likelihood

barplot(names.arg = theta1, height = likelihood1,
    main = "likelihood", col = "#375E97")  # graph
```



At this point our estimates include not only a value of θ and $P(\theta)$, but also the likelihood, $\mathcal{L}(\theta)$.

```
myBayesianEstimates <- tibble(theta1, ptheta1,
    likelihood1)
```

```
pander(myBayesianEstimates)  # nice table
```

theta1	ptheta1	likelihood1
0.25	0.25	0.01877
0.5	0.5	0.0009766
0.75	0.25	2.861e-06

7 *Posterior*

We then calculate the denominator of Bayes theorem:

$$\Sigma[\mathcal{L}(\theta) * P(\theta)]$$

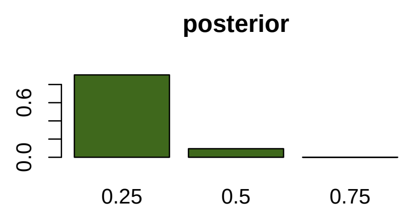
```
pdata1 <- sum(likelihood1 * ptheta1)  # normalize
```

We then use Bayes Rule to calculate the posterior:

$$P(H|D) = \frac{P(D|H)P(H)}{P(D)}$$

```
posterior1 <- likelihood1 * ptheta1/pdata1  # Bayes Rule
```

```
barplot(names.arg = theta1, height = posterior1,
    main = "posterior", col = "#3F681C")  # graph
```



Our estimates now include θ, $P(\theta)$, $\mathcal{L}(\theta)$ and $P(\theta|D)$.

```
myBayesianEstimates <- tibble(theta1, ptheta1,
    likelihood1, posterior1)
```

```
pander(myBayesianEstimates)  # nice table
```

theta1	ptheta1	likelihood1	posterior1
0.25	0.25	0.01877	0.9056
0.5	0.5	0.0009766	0.09423
0.75	0.25	2.861e-06	0.000138

8 *Credits*

Prepared by Andy Grogan-Kaylor `agrogan@umich.edu`, `www.umich.edu/~agrogan`.

Questions, comments and corrections are most welcome.