Homework 3 Problem 3

William Matzko

February 18, 2021

Let θ be the probability of a coin toss yielding heads and $\mathcal{D}$ be the results of a set of coin tosses. We can asses the probability θ given a set of data $\mathcal{D}$ using Bayes' theorem:

$$P(\theta|\mathcal{D}) = P(\mathcal{D}|\theta)\frac{P(\theta)}{P(\mathcal{D})}$$

First, we suppose that $\mathcal{D} = \{H\}$. What is the probability that the coin gives heads? To answer this, we first compute $P(\mathcal{D}|\theta)$ using the binomial distribution

$$P(N,k,\theta) = \binom{N}{k}(1-\theta)^{N-k}\theta^k$$

with $N = k = 1$. This yields

$$P(1,1,\theta) = \theta$$

We may assume that all probabilities of heads are equally likely, so $P(\theta) = c$ where c is some constant. Lastly, we compute $P(\mathcal{D})$ as

$$P(\mathcal{D}) = \int_0^1 P(\theta)P(\mathcal{D}|\theta)d\theta = c\int_0^1 \theta d\theta = \frac{c}{2}$$

Putting this all together, we see

$$P(\theta|\mathcal{D}) = 2\theta$$

The plot of this curve is generated in the corresponding Python script.

Now we suppose that $\mathcal{D} = \{H, T\}$ and we ask the same question. This time, we have $N = 2$ and $k = 1$ in our binomial distribution:

$$P(2,1,\theta) = 2 * (1-\theta)\theta$$

As before, we take $P(\theta) = c$ and calculate $P(\mathcal{D})$ as

$$P(\mathcal{D}) = \int_0^1 = P(\theta)P(\mathcal{D}|\theta)d\theta = 2c\int_0^1 (\theta - \theta^2)d\theta = c/3$$

It thus follows

$$P(\theta|\mathcal{D}) = 6(1-\theta)\theta$$

The plot for this is again generated in the corresponding Python code.

Finally, we assume that instead of flat prior $P(\theta) = c$, we have some Gaussian prior that accounts for our preconceived notion that most coins are fair. We apply a Gaussian prior

$$P(\theta) \propto \exp\{-(\theta - 0.5)^2/0.1\}$$

Here you should technically have computed the normalization of the prior. However, in practice, it does not matter. Also, given my notation screw-up, I can't complain.

and take $\mathcal{D} = \{H, T\}$. $P(\mathcal{D}|\theta)$ is the same as above, but now we have

$$P(\mathcal{D}) = 2\int_0^1 \exp\{-(\theta - 0.5)^2/0.1\}\theta(1-\theta)d\theta \approx 0.655117$$

This integral is a little tedious to evaluate by hand, so I computed it numerically. We thus have

$$P(\theta|\mathcal{D}) = \frac{2\theta(1-\theta)\exp\{-(\theta - 0.5)^2/0.1\}}{0.6551171}$$

which is again graphed in the corresponding Python script.