Statistical Modeling, Winter 2025

Homework 1

Please submit your answers to these questions via Canvas prior to class next Tuesday. I prefer a pdf, or a .Rmd / .Qmd / .R file. Please name the file with your last name and hw number (e.g., elahi_hw1).

Complete the following readings:

- chapter 2.3-2.4 in *Statistical Rethinking* (McElreath 2020).
- chapter 2.3 in *Bayes Rules!* (Johnson, Ott, and Dogucu 2022).

We will discuss the reading, and your answers to the following questions, in class next time.

Questions

The purpose of these questions is to become comfortable with the basic approach in Bayesian modeling, grid approximation, and the binomial distribution.

1. More globe tossing. Compute and plot the grid approximate posterior distribution for each of the following sets of observations. In each case, assume a uniform prior for p.

    a. W, W, W

    b. W, W, W, L

    c. L, W, W, L, W, W, W

2. Now assume a prior for p that is equal to zero when $p < 0.5$ and is a positive constant when $p \geq 0.5$. Again compute and plot the grid approximate posterior distribution for each of the sets of observations in the problem just above.

3. For each variable Y below, determine whether Y is Binomial. If yes, use notation to specify this model and its parameters. If not, explain why the Binomial model is not appropriate for Y.

a. At a certain hospital, an average of 6 babies are born each hour. Let Y be the number of babies born between 9 a.m. and 10 a.m. tomorrow.

b. Tulips planted in fall have a 90% chance of blooming in spring. You plant 27 tulips this year. Let Y be the number that bloom.

c. Each time they try out for the television show Ru Paul's Drag Race, Alaska has a 17% probability of succeeding. Let Y be the number of times Alaska has to try out until they're successful.

d. Y is the amount of time that Henry is late to your lunch date.

e. Y is the probability that your friends will throw you a surprise birthday party even though you said you hate being the center of attention and just want to go out to eat.

f. You invite 60 people to your "π day" party, none of whom know each other, and each of whom has an 80% chance of showing up. Let Y be the total number of guests at your party.

4. Cuckoo birds are brood parasites, meaning that they lay their eggs in the nests of other birds (hosts), so that the host birds will raise the cuckoo bird hatchlings. Lisa is an ornithologist studying the success rate, π, of cuckoo bird hatchlings that survive at least one week. She is taking over the project from a previous researcher who speculated in their notes the following prior model for π:

```
library(tidyverse)
library(knitr)
success_rate_pi <- c(0.6, 0.65, 0.7, 0.75)
success_prior <- c(0.3, 0.4, 0.2, 0.1)
d <- tibble(success_rate_pi, success_prior)
d %>% kable()
```

success_rate_pi	success_prior
0.60	0.3
0.65	0.4
0.70	0.2
0.75	0.1

a. If the previous researcher had been more sure that a hatchling would survive, how would the prior model be different?

b. If the previous researcher had been less sure that a hatchling would survive, how would the prior model be different?

c. Lisa collects some data. Among the 15 hatchlings she studied, 10 survived for at least one week. What is the posterior model for π?

d. Lisa needs to explain the posterior model for π in a research paper for ornithologists, and can't assume they understand Bayesian statistics. Briefly summarize the posterior model in context.

References

Johnson, Alicia A., Miles Q. Ott, and Mine Dogucu. 2022. *Bayes Rules!: An Introduction to Applied Bayesian Modeling*. CRC Press.

McElreath, Richard. 2020. *Statistical Rethinking: A Bayesian Course with Examples in R and Stan*. CRC Press.