

STATISTICAL RETHINKING WINTER 2022

HOMEWORK, WEEK 1

What is homework? Your completed answers to the prompts below should contain all the code necessary to repeat your calculations.

When is homework due? Homework is due each FRIDAY before the relevant discussion of the solutions. So for this first homework assignment, it'll be due on January 14 2022. You are welcome to work in groups. Just please turn in your individual completed answers. If ever for any reason you are late with homework, it's okay. Life is complicated. Turn it in when you finish it. The learning comes from doing it. I want you to do it.

Where is homework due? Upload your homework at the link provided via email. Please name the file with your name and the course week. The preferred file format is PDF or a plain text file (.Rmd or .R or .py or .jl). Please do NOT turn in a Microsoft Word document. Please convert it to PDF first.

WEEK 1 PROBLEMS. For your own good, it would be helpful to review the EASY problems at the end of Chapters 1, 2 and 3. The answers are in the solutions guide.

1. Suppose the globe tossing data (Chapter 2) had turned out to be 4 water and 11 land. Construct the posterior distribution, using grid approximation. Use the same flat prior as in the book.

2. Now suppose the data are 4 water and 2 land. Compute the posterior again, but this time use a prior that is zero below $p = 0.5$ and a constant above $p = 0.5$. This corresponds to prior information that a majority of the Earth's surface is water.

3. For the posterior distribution from 2, compute 89% percentile and HPDI intervals. Compare the widths of these intervals. Which is wider? Why? If you had only the information in the interval, what might you misunderstand about the shape of the posterior distribution?

4. **OPTIONAL CHALLENGE.** Suppose there is bias in sampling so that Land is more likely than Water to be recorded. Specifically, assume that 1-in-5 (20%) of Water samples are accidentally recorded instead as "Land". First, write a generative simulation of this sampling process. Assuming the true proportion of Water is 0.70, what proportion does your simulation tend to produce instead? Second, using a simulated sample of 20 tosses, compute the unbiased posterior distribution of the true proportion of water.