Statistical Modeling, Winter 2025

Homework 8

Please submit your answers to these questions via Canvas prior to class next Tuesday.

I will accept a pdf only. This means you can knit your `.Rmd` document into a pdf, or you can make a word document and insert relevant images of your code and plots.

Please name the file with your last name and hw number (e.g., elahi_hw1).

Complete the following:

- read chapter 14.1 in *Statistical Rethinking* (McElreath 2020).
- watch Gaussian Processes

Questions

From McElreath

1. Repeat the café robot simulation from the beginning of the chapter. This time, set `rho` to zero, so that there is no correlation between intercepts and slopes. How does the posterior distribution of the correlation reflect this change in the underlying simulation?
2. Fit this multilevel model to the simulated café data:

$$
\begin{aligned}
W_i &\sim \text{Normal}(\mu_i, \sigma) \\
\mu_i &= \alpha_{\text{cafe[i]}} + \beta_{\text{cafe[i]}} A_i \\
\alpha_{\text{cafe[i]}} &\sim \text{Normal}(\alpha, \sigma_\alpha) \\
\beta_{\text{cafe[i]}} &\sim \text{Normal}(\beta, \sigma_\beta) \\
\alpha &\sim \text{Normal}(5, 2) \\
\beta &\sim \text{Normal}(-1, 0.5) \\
\sigma, \sigma_\alpha, \sigma_\beta &\sim \text{Exponential}(1)
\end{aligned}
$$

   (a) Use WAIC to compare this model to model `m14.1` from the chapter, the one that uses a multivariate Gaussian prior. Explain the result.

   (b) Plot the mean slope vs the mean intercept from the posterior distribution of each model (i.e., replicate Fig 14.5 from SR2). Interpret the plot.

Project

It is nearly time for your project presentation! From here on out, you will answer/edit the following questions (which will eventually translate to slides in your presentation). You do not have to answer all of these now, but if you want feedback, you should do the best you can.

1. Give a **quick** overview of your awesome project.
2. Draw your DAG.
3. Write out your mathematical model (you can hand-write it and take a picture). Or you can try writing it in .Rmd. For example, the markdown syntax below will create the cafe model above:

```
$$
\begin{aligned}
W_i & \sim \text{Normal}(\mu_i, \sigma) \\
\mu_i & = \alpha_{\text{cafe[i]}} + \beta_{\text{cafe[i]}}A_i \\
\alpha_{\text{cafe[i]}} & \sim \text{Normal}(\alpha, \sigma_{\alpha}) \\
\beta_{\text{cafe[i]}} & \sim \text{Normal}(\beta, \sigma_{\beta}) \\
\alpha & \sim \text{Normal}(5, 2) \\
\beta & \sim \text{Normal}(-1, 0.5) \\
\sigma, \sigma_{\alpha}, \sigma_{\beta} & \sim \text{Exponential}(1) \\
\end{aligned}
$$
```

   This website is helpful.

4. Write out the model in `rethinking` code.
5. Diagnose your chains - you don't need traceplots for every parameter, but you should discuss any issues.
6. **ONE** plot of the posterior and prior distribution (together) for one important parameter.
7. Caterpillar plot of the most important parameters (e.g., the plot you get from using `plot(precis)`)
8. **ONE** plot of posterior predictions that help answer your question.
9. What are your conclusions?