Homework 08

Math 315, Fall 2018

Due: Nov. 12 by 4 p.m.

Instructions: Complete the following problems and submit them by 4 p.m. on the due date. Please make sure that your solution is neatly written, clearly organized, and stapled (if there are multiple pages).

Problem 1

In 1986, the space shuttle Challenger exploded during takeoff, killing the seven astronauts aboard. The explosion was a result of an O-ring failure, a splitting of a rubber ring that seals parts of the ship together. The accident was believed to be caused by the unusually cold weather (31°F) at the time of the launch, as there is reason to believe that the O-ring failure probabilities increase as the temperature decreases.

Data collected about O-rings on other shuttle launches include: temperature at flight time (in °F) and failure/success (0/1) of the O-ring for the flight at the given temperature. The temperatures range from 53°F to 81°F, with the majority of O-ring failures at the lower temperatures. Notice that the temperature at the Challenger launch was significantly lower those studied in the given data.

```
challenger <- read.csv("https://aloy.rbind.io/data/challenger.csv")
```

(a) Write down an expression for the likelihood and link function. See the model presented on page 293 of the textbook for an example.

(b) Fit the model you specified in part (a) using the priors: $\alpha \sim \mathcal{N}(0, 10)$ and $\beta \sim \mathcal{N}(0, 5)$. Fir this model using HMC via `map2stan()`, with the additional arguments `iter = 4000, warmup = 1000, chains = 2`. Be sure to run this model as a separate R script and save your results as outlined in class. Report the output from `precis()` as a **formatted** table, and explain how you know that the Markov chains produced by `map2stan()` converged to the stationary distribution.

(c) Interpret the slope coefficient estimate in the context of the problem.

(d) Calculate a 97% prediction interval for the probability of an O-ring failure at 31°F, the temperature at the time the Challenger launched.

(e) Are you worried about the effective number of independent draws you have from the posterior distribution for each parameter? Why or why not?

(f) Standardize `temp` and refit your logistic regression model using the same prior on α and $\beta \sim \mathcal{N}(0, 2)$. Report the output from `precis()` as a **formatted** table. What do you notice about the effective number of independent draws you have from the posterior distribution for each parameter?

Problem 2

Read the description of problem 10H3 in the textbook and complete the following:

(a) Fit the model specified in 10H3 (a) using HMC via `map2stan()`. Be sure to run this model as a separate R script and save your results as outlined in class. Report the output from `precis()` as a **formatted** table, and explain how you know that the Markov chains produced by `map2stan()` converged to the stationary distribution.

(b) Interpret each slope coefficient estimate in the context of the problem.

(c) Create two plots displaying the posterior predictions:

(i) Plot the *probability of success* (on the y-axis) against the cases (on the x-axis). Add 89% credible intervals to the predicted probabilities. Additionally, add the emprical probabilities (`y/n`) to this plot using a different plotting color and shape.

(ii) Plot the *number of successes* (on the y-axis) against the cases (on the x-axis). Add 89% credible intervals to the predicted number of successes. Additionally, add the observed number of successes to this plot using a different plotting color and shape.

(d) Fit a model with an interaction between the pirate's size and age (immature or adult) using HMC via `map2stan()`. Report the output from `precis()` as a **formatted** table, and explain how you know that the Markov chains produced by `map2stan()` converged to the stationary distribution.

(e) Compare this model to the previous model using WAIC. Report your findings.