Research Methods in Political Science II, Homework Assignment

Due: Noon on Thursday, December 15, 2016.

How to submit: Send by Slack's direct message

File names: Please make sure that all files (including data sets) have *your name in the file name* so that I can easily tell who created the files.

1. rm2-hw04-YourName.pdf (I will no longer accept html files.)
2. rm2-hw04-YourName-XXX.stan (replace XXX with appropriate words)

Note: Please include R codes in your write up. So you should use R markdown (or R markdown notebook).

1. Professor Masahiko Asano at Takushoku Univeristy asked 20 of his students to drink a cup coffee and evaluate its taste. Coffee was provided either by Starbucks or by Tully's. Each student had only one cup without knowing whether it was Starbucks' coffee or Tully's and gave a score between 0 and 100 to the cup. Using the data set `coffee-1.csv`, which is posted on Slack, examine if there is difference between two companies' coffee.

   (a) Compare the sample statistics for two groups.

   (b) Visually compare two groups.

   (c) Build a statistical model to examine this data set. You may assume that the score is normally distributed.

   (d) Translate your statistical model into a Stan model

   (e) Fit the Stan model with RStan.

   (f) Evaluate the fit.

   (g) Is the taste different between two companies? What is your conclusion based on the evidence?

2. Professor Masahiko Asano at Takushoku Univeristy asked 10 of his students to drink two cups of coffee and evaluate their taste. Coffee was provided by Starbucks and Tully's. Each student had one cup from each coffee shop without knowing which was Starbucks' or Tully's and gave a score between 0 and 100 to each cup. Using the data set `coffee-2.csv`, which is posted on Slack, examine if there is difference between two companies' coffee.

   (a) Compare the sample statistics for two groups.

   (b) Visually compare two groups.

   (c) Build a statistical model to examine this data set. You may assume that the difference in score is normally distributed.

   (d) Translate your statistical model into a Stan model

   (e) Fit the Stan model with RStan.

(f) Evaluate the fit.

(g) Is the taste different between two companies? What is your conclusion based on the evidence?

3. Suppose we would like to examine the following statistical model.

$$
\begin{aligned}
Y_n &\sim \text{Bernoulli}(\theta_n), \\
\theta_n &= \text{logit}^{-1}(\beta_1 + \beta_2 x_n + \beta_3 z_n), \\
\beta_1 &\sim \text{Normal}(0, 100), \\
\beta_2 &\sim \text{Normal}(0, 10), \\
\beta_3 &\sim \text{Normal}(0, 10),
\end{aligned}
$$

where $n = 1, 2, \ldots, N$. I translated this model into a Stan model, which is posted on Slack as `hw4-bern-notrun.stan`. This Stan model has some errors. Please correct the errors and explain what changes you make and why for each correction (you might want to add comments in the Stan file). If you can, vectorize the model. Before submitting your Stan file, please make sure that it is compiled without an error.

4. Complete Q2. of Day 4's in-class exercises.