Homework 6

STA-360, Fall 2020

Due at 5:00 PM Standard Eastern Time on TBD

General instructions for homeworks: Please follow the uploading file instructions according to the syllabus. You will give the commands to answer each question in its own code block, which will also produce plots that will be automatically embedded in the output file. Each answer must be supported by written statements as well as any code used. Your code must be completely reproducible and must compile. Syllabus: (https://github.com/resteorts/modern-bayes/blob/master/syllabus/syllabus-sta602-spring19.pdf)

Advice: Start early on the homeworks and it is advised that you not wait until the day of. While the professor and the TA's check emails, they will be answered in the order they are received and last minute help will not be given unless we happen to be free.

Commenting code Code should be commented. See the Google style guide for questions regarding commenting or how to write code `https://google.github.io/styleguide/Rguide.xml`. No late homework's will be accepted.

1. *Censored Gibbs sampling* (20 points) Please complete the censored Gibbs sampling problem from Module 8 (Part II). You may assume that $a = 1, b = 1, r = 10$. You may use any code that was given to you in class or you may choose to write your own code. If you use the code from class, please cite and reference this.

   (a) (5 points) Write code to produce traceplots and running average plots for the censored values for 40 iterations. Do these diagnostic plots suggest that you have run the sampler long enough? Explain.

   (b) (5 points) Now run the chain for 10,000 iterations and update your diagnostic plots (traceplots and running average plots). Report your findings for both traceplots and the running average plots for θ and the censored values. Do these diagnostic plots suggest that you have run the sampler long enough? Explain.

   (c) (5 points) Give plots of the estimated density of $\theta \mid \cdots$ and $z_9 \mid \cdots$. Be sure to give brief explanations of your results and findings. (Present plots for 10,000 iterations).

   (d) (5 points) Finally, let's suppose that $r = 10, a = 1, b = 100$. Do the posterior densities in part (c) change for $\theta \mid \cdots$ and $z_9 \mid \cdots$? Do the

associated posterior densities change when $r = 10, a = 100, b = 1$? Please provide plots and an explanation to back up your answer. (Use 10,000 iterations for the Gibbs sampler).

2. *Gibbs sampling* (30 pts) Consider the following Exponential model for observation(s) $\mathbf{x} = (\mathbf{x_1}, \dots, \mathbf{x_n})$.[1]:

$$p(x|a, b) = ab \exp(-abx)I(x > 0),$$

where the x_i are assumed to be iid for $i = 1, \dots n$. and suppose the prior is

$$p(a, b) = \exp(-a - b)I(a, b > 0).$$

You want to sample from the posterior $p(a, b|x_{1:n})$. You may assume that $a = 0.25, b = 0.25$. **This problem will not be graded and is left as an exercise.**

(a) (5 pts) Find the conditional distributions needed for implementing a Gibbs sampler.

(b) (5 pts) Code up your own Gibbs sampler using part (a).

(c) (10 pts) Plot a histogram or a density estimate of the estimated posterior using (b).

(d) (10 pts) How do you know that your estimated posterior in (c) is reliable? Provide at least two convergence diagnostic plots and commentary.

(e) (Extra credit: 5 points) Plot 95 percent credible intervals on your histogram in part (b) and give commentary.

---

[1]Please note that in the attached data there are 40 observations, which can be found in data-exponential.csv.