

# SYST/STAT 664: Homework Assignment 8

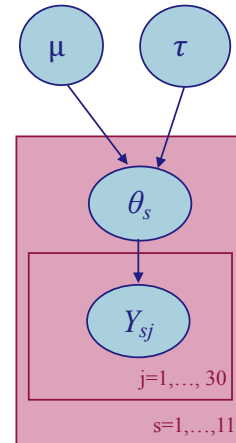
Due April 27, 2020

Homework is at the time and date indicated. Please make sure your name is on every page of the assignment, and it is clearly marked which question you are answering. Your response will be graded for correctness and clarity.

- Section 18.4 of Gelman, et al. (2013) analyzes data on reaction times for 11 non-schizophrenic and 6 schizophrenic subjects. The data set can be found at:

<http://www.stat.columbia.edu/~gelman/book/data/schiz.asc>

The first 11 rows are data for the non-schizophrenic subjects. (We analyzed data from one of these subjects in Unit 5.) Consider a hierarchical model for the natural logarithms of reaction times of the non-schizophrenic subjects. A plate diagram for the model is shown at the right.



- Following Gelman, et al., assume that the logarithms of the response times are independent normal random variables with person-specific mean  $\theta_s$  ( $s = 1, \dots, 11$ ).
- Following Gelman, et al., assume all the observations have the same standard deviation  $\sigma$ . Assume  $\sigma$  is known and equal to the average of the 11 sample standard deviations.
- The 11 means  $\theta_s$ ,  $s=1, \dots, 11$ , are independent and identically distributed normal random variables with mean  $\mu$  and standard deviation  $\tau$ .
- The parameters  $\mu$  and  $\tau$  are independent of each other.
- The unknown mean  $\mu$  has a normal distribution with mean 5.52 and standard deviation 0.22. This reflects a prior 95% credible interval of [162, 385] ms for the population average reaction time, which is consistent with the literature on reaction times.
- The inverse variance  $1/\tau^2$  of the  $\theta_s$  has a gamma distribution with shape  $\frac{1}{2}$  and scale 50. This reflects weak prior information focused on a value of 25 for  $1/\tau^2$ , or 0.2 for  $\tau$ . This is consistent with typical variability of reaction times.

Using the formulas in the Unit 5 notes and the posted examples as a guide, find the posterior distribution for each of the unknown parameters given the other parameters. *Remember that the hyperparameters of the posterior distribution will be formulas involving the other hyperparameters.*

- Use Gibbs sampling to draw 5000 samples from the posterior distribution of the parameters  $\mu$ ,  $\tau$ , and  $\theta_s$ ,  $s=1, \dots, 11$ . Find 95% posterior credible intervals for each of these parameters. You may use JAGS or you may directly implement a Gibbs sampler. Use the examples provided as a guide.
- Consider the following empirical Bayes model for the subject-specific means  $\theta_s$ ,  $s=1, \dots, 11$ .
  - Assume the observation standard deviation  $\sigma$  is known and equal to the average of the 11 sample standard deviations of the log reaction times of the 11 subjects.
  - Assume the means  $\theta_s$ ,  $s=1, \dots, 11$  are normally distributed with mean  $\mu$  equal to the grand mean of all the log reaction times and standard deviation  $\tau$  equal to the standard deviation of the eleven sample means  $\bar{x}_s$ ,  $s=1, \dots, 11$ .

Find the posterior distribution for  $\theta_s$ ,  $s=1, \dots, 11$ . Compare with your results in Problem 2.

- For this problem, you were asked to assume that the natural logarithms of the response times for each non-schizophrenic subject are independent and normally distributed with person-specific mean  $\theta_j$  ( $j = 1, \dots, 11$ ) and common variance  $\sigma^2$ . Discuss whether you think these assumptions are reasonable. Consider both theoretical arguments and analysis of the data.