

## Lab assignment 9

The file *interexp.dat* contains data from an experiment that was interrupted before all the data could be gathered. Of interest was the difference in reaction times of experimental subjects when they were given stimulus *A* versus stimulus *B*. Each subject is tested under one of the two stimuli on their first day of participation in the study, and is tested under the other stimulus at some later date. Unfortunately the experiment was interrupted before it was finished, leaving the researchers with 26 subjects with both *A* and *B* responses, 15 subjects with only *A* responses and 17 subjects with only *B* responses. We assume the data generating function is a multivariate normal distribution with parameters  $\boldsymbol{\theta} = (\theta_A, \theta_B)'$  and  $\Sigma$ . We assume Jeffreys' prior for the parameters.

The data is given at

<http://www.stat.washington.edu/hoff/Book/Data/hwdata/interexp.dat>

- a) Using only the complete data (the set of subjects with no missing observations), compute a posterior mean for  $\theta_A - \theta_B$  as well as a 95% posterior interval for  $\theta_A - \theta_B$ .
- b) Replace each missing value with sample mean. Using the imputed data, compute a posterior mean for  $\theta_A - \theta_B$  as well as a 95% posterior interval for  $\theta_A - \theta_B$ .
- c) Implement a Gibbs sampler that approximates the joint distribution of the parameters and the missing data. Compute a posterior mean for  $\theta_A - \theta_B$  as well as a 95% posterior interval for  $\theta_A - \theta_B$ . Compare these results with the results from a) and b) and discuss.