SYST/STAT 664: Homework Assignment 8

due April 16, 2018

Homework is due at class time on the date indicated. You may submit on paper or electronically via Blackboard. 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.

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

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

   The first 11 rows are data for the non-schizophrenic subjects. Gelman, et al. assume that the logarithms of the response times for each non-schizophrenic subject are independent and identically distributed normal random variables with person-specific mean θ_j ($j = 1,\ldots 11$) and common variance σ^2. Discuss whether you think this assumption is reasonable.

2. For this problem, we will assume that the logarithms of the response times for each non-schizophrenic subject are independent and identically distributed normal random variables with person-specific mean θ_j. Unlike Gelman, et al., we will assume the precision ρ_j may also depend on the subject. We will assume that the 11 means and precisions (θ_j, ρ_j) are independent and identically distributed normal-gamma random variables with center μ, precision multiplier k, shape α, and scale β. To do an empirical Bayes analysis, we need estimates of the hyperparameters μ, k, α, and β. To construct the estimates, first
   - To estimate the shape α and scale β, estimate the sample precisions $\hat{\rho}_1, \ldots, \hat{\rho}_{11}$ by calculating the sample variances and taking the inverses. Recall that the unknown precision has a gamma distribution with mean $\alpha\beta$ and variance $\alpha\beta^2$. Therefore, we can estimate β as the sample variance of precisions $\hat{\rho}_1, \ldots, \hat{\rho}_{11}$ divided by the sample mean of precisions $\hat{\rho}_1, \ldots, \hat{\rho}_{11}$. Then we can estimate α as the sample mean of the $\hat{\rho}_1, \ldots, \hat{\rho}_{11}$ divided by the estimate of β. That is, we are solving the following two equations for α and β:
     i. $\alpha\beta^2$ = sample variance of $\hat{\rho}_1, \ldots, \hat{\rho}_{11}$
     ii. $\alpha\beta$ = sample mean of $\hat{\rho}_1, \ldots, \hat{\rho}_{11}$
   - Estimate the center μ as the grand mean of all the log reaction times.
   - Estimate the precision multiplier k as follows. First, find the sample mean of the log reaction times for each of the 11 subjects. Then calculate the sample variance of these 11 values. Then take the inverse to get the sample precision. Divide this sample precision by the average of $\hat{\rho}_1, \ldots, \hat{\rho}_{11}$ to get your estimate of k.

   Using these empirical Bayes estimates of μ, k, α, and β, find the eleven posterior distributions of (θ_j, ρ_j). Make a table to show your eleven posterior sets of hyperparameters. The table will have eleven rows, one for each subject, and a column for each of the four posterior hyperparameters.

3. Find 95% credible intervals for each of the eleven means θ_j ($j = 1,\ldots 11$) and precisions ρ_j ($j = 1,\ldots 11$). Discuss your results, especially whether and how the distribution of reaction times differs across subjects.