AMS 206 – Homework Assignment #6
due Friday, March 14

Required Problem

1. The data file **hearing.txt** is from an experiment to calibrate word lists used to measure the hearing ability of subjects. The four word lists had been designed so that they should be equally difficult to perceive, but were designed for normal-hearing subjects in an environment without background noise. The data in this experiment were collected in the presence of a noisy background. Each column is a word list, and each row is a subject. The entry is their score on that list. So each subject was tested on all four lists. We will do a simplified version of a two-way ANOVA model. We'll use a normal likelihood for each score, with a fixed variance of 36, and a mean that depends on both the subject and on the list. We'll put normal priors on each subject effect θ_h (centered at a joint mean score μ with variance 36) and on each list effect ϕ_j (centered at zero with variance 9). The prior for μ is normal with mean 30 and variance 4. Thus the hierarchical model is:

$$\begin{aligned} y_{hj} &\sim N(\theta_h + \phi_j, 36) \\ \theta_h &\sim N(\mu, 36) \\ \phi_j &\sim N(0, 9) \\ \mu &\sim N(30, 4) \end{aligned}$$

for $h = 1, \ldots, n = 24$ and $j = 1, \ldots, k = 4$. Note that the joint likelihood is thus

$$f(\mathbf{y}|\boldsymbol{\theta}, \boldsymbol{\phi}) = (2\pi 36)^{-nk/2} \exp\left\{ -\frac{1}{2 * 36} \sum_{h=1}^{n} \sum_{j=1}^{k} (y_{hj} - \theta_h - \phi_j)^2 \right\}$$

(a) Show that the complete conditional distribution for a θ_h is

$$f(\theta_h|\boldsymbol{\theta_{-h}}, \boldsymbol{\phi}, \mu, \mathbf{y}) \propto \exp\left\{ -\frac{k+1}{2 * 36} \left(\theta_h - \frac{\mu + \sum_{j=1}^{k}(y_{hj} - \phi_j)}{k+1} \right)^2 \right\}$$

and recognize this as something that can be sampled with a Gibbs step.

(b) Find the complete conditional distribution for a ϕ_j and recognize it as something that can be sampled with a Gibbs step.

(c) Find the complete conditional distribution for μ and recognize it as something that can be sampled with a Gibbs step.

(d) Fit the model with MCMC. Show your trace plots for mu, for at least three θ_j's, and for at least two ϕ_j's of your choice. Remove burn-in as appropriate.

(e) Make a plot comparing the maximum likelihood estimates of the θ_h's to your estimated posterior means of the θ_h's. Use `abline(0,1)` to add the $y = x$ line to your plot. Comment on what you see. How does this Bayesian analysis compare to a simple frequentist (maximum likelihood) one?

(f) Of interest to the researchers is whether the lists have the same level of difficulty. Plot histograms of the posterior for all four ϕ_j's. Construct 95% credible intervals for each ϕ_j and see if they include zero. What can you conclude about the lists?

Optional Problem

2. Expand the above model so that the variance is unknown, and put an inverse-gamma(1,1) prior on it. In particular, use the model:

$$
\begin{aligned}
y_{hj} &\sim N(\theta_h + \phi_j, \sigma^2) \\
\theta_h|\sigma^2 &\sim N(\mu, \sigma^2) \\
\phi_j|\sigma^2 &\sim N(0, \frac{\sigma^2}{4}) \\
\mu|\sigma^2 &\sim N(30, \frac{\sigma^2}{9}) \\
\sigma^2 &\sim \Gamma^{-1}(1, 1)
\end{aligned}
$$

Fit this model compare the results to the above model and to the MLE's.