Exercises 5: Hierarchical models and data augmentation

Hierarchical models and shrinkage

The data set in "mathtest.csv" shows the scores on a standardized math test from a sample of 10th-grade students at 100 different U.S. urban high schools, all having enrollment of at least 400 10th-grade students. (A lot of educational research involves "survey tests" of this sort, with tests administered to all students being the rare exception.)

Let θ_i be the underlying mean test score for school i, and let y_{ij} be the score for the jth student in school i. Starting with the "mathtest.R" script, you'll notice that the extreme school-level averages $\bar{y}_i$ (both high and low) tend to be at schools where fewer students were sampled.

1. Explain briefly why this would be.

2. Fit a normal hierarchical model to these data via Gibbs sampling:

$$\begin{aligned} y_{ij} &\sim \mathrm{N}(\theta_i, \sigma^2) \\ \theta_i &\sim \mathrm{N}(\mu, \tau^2) \end{aligned}$$

   Decide upon sensible priors for the unknown model parameters (μ, σ^2, τ^2).

3. Suppose you use the posterior mean $\hat{\theta}_i$ from the above model to estimate each school-level mean θ_i. Define the shrinkage coefficient κ_i as

$$\kappa_i = \frac{\bar{y}_i - \hat{\theta}_i}{\bar{y}_i},$$

   which tells you how much the posterior mean shrinks the observed sample mean. Plot this shrinkage coefficient for each school as a function of that school's sample size, and comment.

Data augmentation

Read the following paper:

> "Bayesian Analysis of Binary and Polychotomous Response Data." James H. Albert and Siddhartha Chib. *Journal of the American Statistical Association*, Vol. 88, No. 422 (Jun., 1993), pp. 669-679

The surefire way to get this paper is via access to JStor through the UT Library. (I have no comment on whether others might have sneakily posted copies of the PDF on the open web.)