Course 52558: Problem Set 3
Due May 19th, 2009

1. **Normal Approximation for the Bioassay example**.

   Consider the Bioassay example we encountered in class 3 (see the Handout section in the website for the note and figures).

   (a) Derive the analytic form of the information matrix and the normal approximation variance for this example/

   (b) In addition to performing a normal approximation to the joint posterior, we can also approximate any posterior estimand using a normal distribution (this is often called the delta method). Expand the posterior distribution of LD50, $-\alpha/\beta$ (see note), as a Taylor series around the posterior mode and thereby derive the asymptotic posterior median and standard deviation. Compare to the normal approximation figure in the website.

2. **Echangeability and Mixture of iid distributions**

   (a) Suppose the distribution of $\theta = (\theta_1, \ldots, \theta_{2J})$ can be written as a mixture of iid components

   $$P(\theta) = \int \prod_{j=1}^{2J} P(\theta_j \mid \phi) P(\phi) d\phi$$

   Prove that the covariances $cov(\theta_i, \theta_j)$ are all non-negative.

   (b) Suppose it is known a-priori that the $2J$ parameters are clustered into groups with exactly half being drawn from $N(1, 1)$, and the other half being drawn from a $N(-1, 1)$ distribution, be we have do not know which parameter comes from which distribution. Are $\theta_1, \ldots, \theta_{2J}$ exchangeable under this prior distribution?

   (c) Use (a) to show that the distribution in (b) cannot be written as a mixture of iid components.

   (d) Continuing (b), why not take the limit as $J \to \infty$ and get a counter-example to De-Finetti's theorem?

3. **Discrete Mixture Models**

If $P_m(\theta)$ for $m = 1, \ldots, M$, are conjugate prior densities for the sampling model $y \mid \theta$, show that the class of finite mixture prior densities given by

$$P(\theta) = \sum_{m=1}^{M} \lambda_m P_m(\theta)$$

is also a conjugate class, where the λ_m's are nonnegative weights that sum to 1.

(a) The above can provide a useful flexible extention of the natural conjugate family. As an example, use the mixture form to create a bimodal prior density for the normal mean, that is thought to be near 1, with a standard deviation of 0.5, but has a small probability of being near -1, with the same standard deviation. If the variance of each observation $y_1, \ldots, y_{10}$ is known to be 1, and their observed mean is -0.25, derive your posterior distribution for the mean, making a sketch of both prior and posterior densities. Be careful: the prior and posterior mixture proportions are different.

(b) Is the above a hierarchical model? Explain.

4. **Beta-blockers Meta Analysis**

In this question we will perform a hierarchical Bayesian analysis aimed at combining information from multiple experiment that try to estimate the effectiveness of beta-blockers at reducing mortality rate from heart attacks. The data appears in a handout in the webpage along with and some results that you can use for comparison. For each experiment j, y_{0j} will denote the number of deaths out of n_{0j} people in the control group and y_{1j} will denote the number of deaths out of n_{1j} people in the treated group. We will perform estimation on the log odds ratio parameter for each experiment

$$\theta_j = log \frac{p_{1j}}{1 - p_{1j}} - log \frac{p_{0j}}{1 - p_{0j}}$$

where p_{0j} is the probability of death in the control group and similarly for p_{1j}.

Relatively simple Bayesian meta-analysis (combination of experiments from different sources) is possible using the normal hierarchical model we used in class to analyze the SAT score improvements. Our approach here will be based on empirical logits, where for each study j we use the estimate

$$y_j = log \frac{y_{1j}}{n_{1j} - y_{1j}} - log \frac{y_{0j}}{n_{0j} - y_{0j}}$$

with approximate sampling variance

$$\sigma_j^2 = \frac{1}{y_{1j}} + \frac{1}{n_{1j} - y_{1j}} + \frac{1}{y_{0j}} + \frac{1}{n_{0j} - y_{0j}}$$

(columns 4 and 5 in the data table in the handout).

(a) Plot the posterior density of τ over an appropriate range that includes essentially all of the posterior density.

(b) Produce graphs that display how the posterior means and standard deviations of the θ_j's depend on τ.

(c) Produce a scatterplot of the crude effect estimates (each y_j) vs. the posterior median effect estimates of the 22 studies. Verify that the studies with smallest sample sizes are 'shrunk' the most toward the mean.

(d) Draw simulations from the posterior distribution of a new treatment effect, θ_{J+1}. Plot a histogram of the simulations.

(e) Given the above simulations, draw simulated outcomes from replications of a hypothetical experiment with 100 persons in each of the treated and control groups. Plot a histogram of the simulations of the crude estimated treatment effect in the new experiment.