1. Problem 10.1 in BDA –
   (a): approximate standard error is $4/\sqrt{100} = .4$
   (b): need to have sample size 1000 to make approx std error equal .1
   (c): the value of the N(8,4) density at the 2.5%ile (0.16) is .0146 so the approx std err is $
   \sqrt{0.25 \times 0.975/(0.0146 + 0.0146 + 100)} = 1.07$
   (d): need to have sample size 11,500 to make this approx std err equal .1

2. 5.8 in BDA: We need to show that $p(\theta, \mu, \tau|y)$ has finite integral. In this case we can do some
   of the required integration analytically
   $$
   \int \int \int p(\theta, \mu, \tau|y)d\theta \, d\mu \, d\tau = \int \int \int p(\theta|\mu, \tau, y)d\theta \, p(\mu|\tau, y)d\mu \, p(\tau|y)d\tau = \int p(\tau|y)d\tau
   $$
   so that $p(\theta, \mu, \tau|y)$ is proper if and only if $p(\tau|y)$ is a proper univariate distribution. The
   display above uses the results about $\theta|\mu, \tau$ and $\mu|\tau, y$ that are given in the text. As for the
   rest of the problem
   - For $\tau$ near zero, the posterior distribution is a constant multiple of the prior. So we
     need only check integrability of the prior distribution. An important point to note is
     that $p(\tau)$ being infinite at $\tau = 0$ is not the problem; if $p(\tau) = 1/\sqrt{\tau}$ then the posterior
     would behave fine near zero.
   - The argument from (a) show that if $p(\tau) = 1$ then the posterior is integrable near zero.
     Now we need to check the behavior as $\tau$ gets large. All we need is an upper bound for
     each term. The exponential term is clearly less than one. We can rewrite the remaining
     terms as $(\sum_{j=1}^{J} \prod_{k\neq j} (\theta_k^2 + \tau^2))^{-1/2}$. For $\tau$ larger than one we make this quantity bigger
     by dropping all of the $\sigma^2$ to yield $(J \tau^{2(J-1)})^{-1/2}$. An upper bound on $p(\tau|y)$ for $\tau$ large
     is $J^{-1/2}/\tau^{J-1}$. This upper bound is integrable if $J > 2$ so that $p(\tau|y)$ is integrable if
     $J > 2$. (For $J = 2$ we would actually need to find a lower bound and show that it is not
     integrable but the argument is similar).
   - For part (c) several people suggested a proper prior for $\tau$. Others noted that $p(\tau) \propto \tau^{-1/2}$
     would work fine in the arguments of parts (a) and (b). That would work. A key point
     is that the data provides very little information about $\tau$ if $J = 2$ (only a poor point
     estimate is possible for $\tau$). I would probably do a two-sample analysis and not worry
     about the hierarchical structure.

3. Poisson/Gamma model: Without having graded this problem yet it is hard to know where
   to place the emphasis in my comments. So here are more details than I usually provide.
   (a) Unnormalized joint posterior density
   $$
   p(\alpha, \beta, \theta|y) \propto \prod_{i=1}^{n} \left[ \frac{e^{-\theta_i} \theta_i^{\beta_i} \beta^{\alpha} \alpha^{-1} e^{-\theta \beta}}{y_i! \Gamma(\alpha)} \right]
   $$
   (b) Conditional posterior distribution is obtained by looking at the above ignoring terms
   that don't include $\theta$ ...
   $$
   p(\theta|\alpha_i, \beta, y) \propto \prod_{i=1}^{n} \left[ e^{-\theta_i} \theta_i^{\beta_i} \alpha^{-1} e^{-\theta \beta} \right]
   $$
   which we recognize as independent Gamma($\alpha + y_i, \beta + 1$) distns
(c) I seemed to have caused confusion here. It is generally most straightforward to just integrate out $\theta$ as shown here. On occasion, especially for the normal-normal case as on pg 139 of the text, it is actually easier to use the ratio $p(\theta, \alpha, \beta|y)/p(\theta|y, \alpha, \beta)$.

$$p(\alpha, \beta|y) \propto \int \cdots \int \prod_{i=1}^{n} \left[ \frac{e^{-\theta y_i} \beta^{\alpha y_i - 1} e^{-\theta \beta}}{y_i! \Gamma(\alpha)} \right] d\theta_1 \cdots d\theta_n$$

$$\propto \prod_{i=1}^{n} \left[ \frac{\beta^{\alpha y_i + y} \Gamma(\alpha + y)}{(\beta + 1)^{\alpha + y} \Gamma(\alpha) y_i!} \right]$$

Note that this is actually a negative binomial distn (if $\alpha$ is an integer) corresponding to the number of failures occurring while waiting for $\alpha$ successes where the probability of success is $\beta/(\beta + 1)$.

(d) Following the hint we consider $\alpha = c \beta$ and then let $\beta \to \infty$. Let's consider a single term of the product which after substituting looks like: $(\beta/(\beta + 1))^c \beta(1/(\beta + 1))^y(y_i + c\beta - 1)/(y_i + c\beta - 2) \cdots (c\beta)$. The first term converges to a constant $e^{-c}$ and the rest converges to $c^y$; this means the limit is constant and the density is not integrable along the direction chosen. Thus we don't get a proper posterior distn.

(e) This follows from transformation of variables. I expect you can do this, if not please talk to me.

4. Poisson/Gamma model (cont'd)

(a) Simulations should yield a $1000 \times 12$ matrix with column 1 containing simulated $\alpha$'s, column 2 containing simulated $\beta$'s, column 3 containing simulated $\theta_1$'s, columns 4-12 containing simulated $\theta_2 - \theta_{10}$'s.

(b) If we go back to street 1, then the simulated $\theta_1$'s are plausible values for the true rate there. Thus to simulate a predictive distribution we generate a new Poisson variate for each of the 1000 simulated $\theta_1$'s.

(c) Here we need to do something different. This is a new street; the only information we have is based on the simulated $\alpha$'s and $\beta$'s. First simulate a new $\theta$ for each of the 1000 $\alpha, \beta$ pairs in columns 1 and 2. Then simulate Poisson counts for these $\theta$'s. These counts should be considerably more variable than the counts in (g) because there is no information about the $\theta$ for the street.

5. Project - I have not heard from or spoken to about half the class. Please contact me to discuss project ideas if you have not done so.