36-463/663: Multilevel & Hierarchical Models
Fall 2016
HW09 – Due Thu 10 Nov 2016

Announcements

- Please turn in hw as a single pdf to blackboard.
- Reading:
  - This week: **Lynch**, Ch 4, primarily.
  - Next week (and a bit more):
    - **Lynch:** I will be looking at some topics from Chapter 9 of Lynch. Where R programming is concerned, pay closer attention to the lecture notes than to Lynch.
    - **Gelman & Hill:** As we finish up with our digression into Lynch we will start looking at portions of G&H Chapters 16, 17, 21 and 24, using examples from Chapters 12–15.
    - More details about the reading will come in later HW assignments.
    - Next week will also be the beginning of the JAGS part of the course!

Exercises

1. Let $x_1, \ldots, x_n$, $n = 169$, be an iid sample from a Normal distribution $N(\mu, \sigma_0^2)$ with *unknown* mean μ and *known* variance $\sigma_0^2 = 144$. Suppose the mean of the x's is $\overline{x} = 100$. Find and plot the posterior distribution for μ, using at least three different $N(\mu_0, \tau_0^2)$ prior distributions. When does the prior make the largest difference in the outcome—when the prior mean varies substantially from the sample mean, or when the prior variance is small or large? (This is a simplified version of Lynch, Ch 3, p. 74, #5). *Hints:*

   - The *likelihood* is going to be

   $$L(\mu) \propto f(\overline{x}|\mu, \sigma_0^2) = \frac{1}{\sqrt{2\pi\sigma_0^2/n}} \exp\left\{-\frac{(\overline{x}-\mu)^2}{2\sigma_0^2/n}\right\}$$

   with $n = 169$ and $\sigma_0^2 = 144$. The *prior distribution* for μ is going to be

   $$f(\mu|\mu_0, \tau_0) = \frac{1}{\sqrt{2\pi\tau_0^2}} \exp\left\{-\frac{(\mu-\mu_0)^2}{2\tau_0^2}\right\}$$

   The *posterior distribution* for μ will also be $N(\mu^*, \sigma^{*2})$ Using formulas from Lynch or class notes, figure out what the new parameters μ^* and σ^{*2} will be, in terms of $\overline{x}$, σ_0^2, μ_0 and τ_0^2.

- To answer the question, try some values of μ_0 that move away from $\overline{x} = 100$, with $\tau_0^2 = 81$ or so. Then try to make τ^2 smaller and see what happens.
- For graphing, imitate some of the ideas in the R handouts for week10, using the `dnorm()` function in R as appropriate.

2. Suppose we have two groups, with n_1 observations $x_{11}, x_{12}, \ldots, x_{1n_1}$ in the first group and n_2 observations $x_{21}, x_{22}, \ldots, x_{2n_2}$ in the second group. We wish to model this data with the following normal model

$$\begin{array}{rcl} \text{Level 1:} \;\; x_{ij} & \stackrel{iid}{\sim} & N(\mu_i, \sigma^2), \;\; j = 1, \ldots, n_i, \;\; i = 1, 2 \\ \text{Level 2:} \;\; \mu_i & \stackrel{iid}{\sim} & N(\mu_0, \tau^2), \;\; i = 1, 2 \\ \text{Level 3:} \;\; \mu_0 & \sim & N(0, 1) \\ 1/\sigma^2 & \sim & Gamma(0.001, 0.001) \\ 1/\tau^2 & \sim & Gamma(0.001, 0.001) \end{array}$$

Thus, μ_1 is the mean of group 1, μ_2 is the mean of group 2, and μ_0 is the overall mean [the x_{ij}'s could be the student scores in two classrooms, μ_1 and μ_2 are the two classroom effects, and μ_0 is the overall school effect, for example].

The data for the two groups are

$$\begin{array}{ll} \text{Group 1:} & 1.81, \; 0.52, \; -0.53, \; -1.60, \; -1.74 \\ \text{Group 2:} & 2.90, \; 0.38, \; 2.76, \; -0.13 \end{array}$$

(a) Use the following R code for the WinBUGS model and data.list, create an approporiate initial-values function, and use `rube()` to estimate (point estimates and CI's!) μ_0, μ_1, μ_2, σ^2 and τ^2. *[NOTE: The code below is contained in the file hw09.r, for your convenience.]*

```
M2 <- "model {

  for (i in 1:n) {
    x[i] ~ dnorm(mu[gp[i]],sig2inv)
  }

  for (j in 1:2) {
    mu[j] ~ dnorm(mu0,tau2inv)
  }

  mu0 ~ dnorm(0,1)

  sig2inv ~ dgamma(0.001,0.001)
  tau2inv ~ dgamma(0.001,0.001)
```

```
  sig2 <- 1/sig2inv
  tau2 <- 1/tau2inv

}"

data.list.2 <- list(x=c(1.81, 0.52, -0.53, -1.60, -1.74,
                         2.90, 0.38, 2.76, -0.13),
                    gp=c(rep(1,5),rep(2,4)),
                    n=9)
```

Inspect the plots from `p3()` and comment on anything unusual.

(b) Suppose Group 1 was the treatment group and Group 2 was the control group, in a randomized controlled experiment. Did the treatment work? Why or why not?

3. There is a certain intersection in town that has many accidents, even though the city installed a traffic light. Some students watched activity at that intersection during the noon hour on each of 16 successive days. They found that the number of cars that ran a red light at the intersection on each day was:

$$3, 7, 1, 7, 6, 1, 5, 4, 6, 2, 11, 3, 1, 6, 4, 4$$

Suppose we model this data by assuming that each day's count is an independent observation of a Poisson rv. with pmf

$$f(x|\lambda) = \frac{\lambda^x e^{-\lambda}}{x!}$$

for $x = 0, 1, 2, \ldots$, and we assume a Gamma distribution as the prior distribution for λ,

$$f(\lambda|\alpha, \beta) = \frac{\beta^\alpha}{\Gamma(\alpha)} \lambda^{\alpha-1} e^{-\beta\lambda}$$

with parameters $\alpha = 1$ and $\beta = 1$.

Follow the recipe in lecture #20 ("bayes and MCMC", week 11; lecture #21 may also be helpful) to estimate λ by MCMC [you did this on HW08 by direct calculation, so you will know if your estimate is reasonable!]. Include the following in your answer:

(a) Listings of

   i. The JAGS code for this model
   ii. The R code you used to create the data.list
   iii. The R code you used to define the initial-values function.
   iv. The `rube()` command you used to create the MCMC simulation.

(b) A point estimate and a 95% CI for λ, from the posterior distribution.

(c) The graph(s) for λ from the `p3()` function.