Name: Solutions

Bayesian Statistics, 22S:138
Midterm 2, 2011

1. A research study conducted in a driving simulator investigated how long (in seconds) it took drivers to see the simulated image of a deer approaching the roadway.

   Fifty subjects were recruited into the study. Ten were teenagers who were taking driver education classes. The others were licensed drivers. The licensed drivers were divided into 4 groups based on gender and length of time they had had a driver's license. Thus there were 5 groups of 10 subjects each.

   Each subject was instructed to "drive" in the driving simulator while wearing a microphone and to speak the word "deer" when he or she first saw a deer anywhere in the scene being projected. The time in seconds for the jth subject in group i to speak the word after the image of the deer appeared in the projected scene was recorded as data value x_{ij}.

   Since times have to be nonnegative, the researcher wants to use a gamma density at the first stage of his model. He believes that the populations represented by the 5 samples of drivers will have different by similar distributions for their members' times to seeing the image. Therefore, he chooses to use a gamma likelihood with the same shape parameter (equal to 4) but a different scale parameter for each group. His model is as follows.

$$\left.\begin{array}{lcl} \lambda_i & \sim & \text{dgamma}(\alpha,\beta) \\ x_{ij} & \sim & \text{dgamma}(4,\lambda_i) \quad \}\, 1 \le j \le 10 \end{array}\right\} 1 \le i \le 5$$

$$\begin{array}{lcl} \alpha & \sim & \text{dunif}(0,30) \\ \beta & \sim & \text{dunif}(0,20) \\ \text{ratio} & = & \alpha/\beta \end{array}$$

   OpenBUGS code and output for fitting this model are provided. Three chains were run for 1000 iterations.

   2 pts (a) On the OpenBUGS code, indicate which line or lines represent the third stage of the model. Put your name on the page of OpenBUGS code!

   (b) In this model, are the $x_{5,j}$s (the data values from group 5) considered exchangeable? Answer yes or no, and briefly justify your answer.

   Yes. They all share the same parameter λ_5, so permuting their j indices would not change the likelihood evaluation.

(c) The autocorrelation between samples drawn from the marginal posterior of the parameter α at a lag of 2 iterations apart is closest to (circle one)

i. 1.0

ii. 0.5

iii. 0.0

iv. -0.5

v. -1.0

vi. No output is provided that addresses this question.

(d) For the α parameter, the M.C. error is greater than 1/20th of the estimated posterior standard deviation. This suggests that the statistician should (circle all that apply):

i. discard a large number of burn-in iterations

ii. conclude that the model does not fit the data

iii. run the chains for more iterations

iv. none of the above

v. all of the above

(e) How many burnin-iterations would you discard? Justify your answer by referring to specific OpenBUGS output.

at least 500. We do not see the blue and green lines stabilizing before that.

(f) The estimated posterior mean for λ_1 is 1.547. If it were possible to get a frequentist estimate of λ_1 based on the data from group 1 subjects only, which of the following would you expect (circle one):

i. the frequentist estimate of this parameter will be larger than the posterior mean from the Bayesian hierarchical model

ii. the frequentist estimate of this parameter will be smaller than the posterior mean from the Bayesian hierarchical model

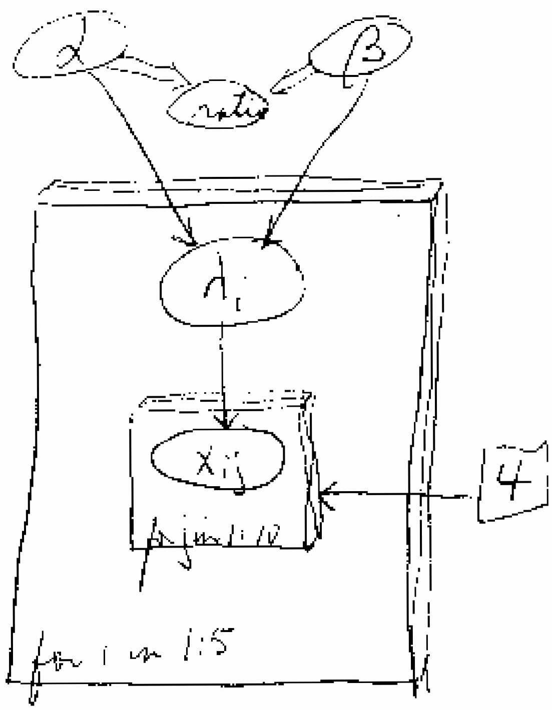
iii. the frequentist estimate of this parameter will be approximately equal to the posterior mean from the Bayesian hierarchical model

iv. there is no way to guess whether the frequentist estimate based on group 1 data will compare to the Bayesian estimate from the hierarchical model.

(g) Briefly justify your answer to part (f).

The posterior mean of λ_1 is the smallest of the estimated λs in the Bayesian analysis. It most likely was shrunk upward from the frequentist estimate toward the larger λs.

(h) Draw a directed graph of the model for the drivers data.



2. A farmer believes that the lengths in inches of ears of corn of a particular variety follow a normal density with unknown population mean μ and population variance σ^2. He performs a Bayesian analysis to estimate μ and σ^2.

   (a) He wishes to introduce no external information into his analysis. What is the conventional noninformative joint prior for μ and σ^2 for a normal likelihood?

   $$p(\mu, \sigma^2) \propto \frac{1}{\sigma^2}, \quad -\infty < \mu < \infty, \quad 0 < \sigma^2 < \infty$$

   (b) If the farmer measures 50 ears of corn and obtains a sample mean $\bar{x} = 7.2$ inches and a sample variance $s^2 = 1$ inch,

      i. What is the posterior marginal density of σ^2, that is $p(\sigma^2|\mathbf{x})$? Name the density and give numeric values of its parameters.

      $$IG\left(\frac{n-1}{2}, \frac{(n-1)s^2}{2}\right) =$$

      $$IG\left(\frac{49}{2}, \frac{49}{2}\right)$$

      ii. What is the posterior marginal density of μ, that is $p(\mu|\mathbf{x})$? Name the density and give numeric values of its parameters.

      $$t_{49}\left(7.2, \frac{1}{50}\right)$$

      iii. Write the line of R code that you could use to find a 95% posterior credible set for μ.

      ```
      7.2 + qt(c(0.025, 0.975), 49) * 1/sqrt(50)
      ```

   (c) Suppose the farmer had used the prior from part(a) but had been less industrious about gathering data: suppose he weighed only two (2) ears of corn, and both ears weighed exactly 7.2 ounces to the greatest accuracy attainable on his scale. Would Bayesian inference for μ and σ^2 have been possible? Briefly explain why or why not.

   Since the 2 data values were equal, s^2 would have been 0. In combination with the improper prior, this would have produced an improper posterior — $p(\sigma^2|x)$ would be $IG(\frac{1}{2}, 0)$ which is improper, and $p(\mu|X)$ would have 0 for its scale parameter. Thus no Bayesian inference would be possible.