10

21

7

38

Name: Solutions

Bayesian Statistics, 22S:138
Midterm 2, 2009

1. Reconsider the "Dyes" example from homework 6, in which the observed data were the yields of dyestuff y_{ij} for 5 samples from each of 6 randomly chosen batches of raw material. The subscript $i, i = 1, \ldots, 6,$ indicates the batch, and the subscript $j, j = 1, \ldots, 5$, identifies the sample within the batch. In the WinBUGS example and in homework 6, it was assumed that the within-batch precision τ^2_{with} of dyestuff yield was the same for all batches, but that each batch i had its own population mean μ_i. This model can be expressed as:
Model 1
Likelihood:

3

$$y_{ij}|\mu_i, \tau^2_{with} \sim \cancel{N(\mu_i, \tau^2_{with})} \quad N(\theta, \tau^2_{with,i})$$

Second stage:

3

$$\cancel{\mu_i|\theta, \tau^2_{betw} \sim N(\theta, \tau^2_{betw})}$$

$$\tau^2_{with,i} \,\Big|\, \alpha, \beta \sim G(\alpha, \beta)$$

Third stage:

4

$$\theta \sim N(0, 0.000001) \quad \text{Keep}$$

$$\cancel{\tau^2_{with} \sim G(0.001, 0.001)}$$

$$\cancel{\tau^2_{betw} \sim G(20, 20000)}$$

$$\left.\begin{array}{l} \alpha \sim G(.01, .01) \\ \beta \sim G(.01, .01) \end{array}\right\} \text{ or similar}$$

Suppose that instead, we believed that the means were the same for all batches. However, we now relax the assumption of equal within-batch precisions, and allow each batch to have its own precision $\tau^2_{with,i}$.

(a) Write the following changes into the specification of Model 1.

i. Change the likelihood to assume a common mean for all batches but to allow for individual precisions $\tau^2_{with,i}$.

ii. To the second stage, add a semi-conjugate prior for the $\tau^2_{with,i}$s. The parameters of this semi-conjugate prior should be unknown parameters that will be estimated. Cross out any items that are no longer needed in the second stage.

iii. Complete the model specification by changing, adding and/or crossing out items in the third stage as appropriate. If you add any new prior(s) at the third stage, make them vague but proper.

10

(b) Using your new model, write the expression to which the joint posterior distribution of all unknown model ~~parameters given~~ the observed data is proportional.

6

$$\prod_{i=1}^{6}\left\{\sqrt{\tau^2_{with,i}}^{\,5}\exp\left[-\frac{\tau^2_{with,i}\sum_{j=1}^{5}(y_{ij}-\theta)^2}{2}\right](\tau^2_{with,i})^{\alpha-1}e^{-\beta\tau^2_{with,i}}\right\}$$

$$\times\exp\left[-\frac{(\theta-0)^2}{2\times 10000}\right]\alpha^{.01-1}e^{-.01\alpha}\beta^{.01-1}e^{-.01\beta}$$

(c) Derive the posterior full conditional for $\tau^2_{with,2}$. If possible, identify it as a standard probability distribution.

$$p(\tau^2_{with,2}\mid\alpha,\beta,y,\theta,\tau^2_{with(-2)})\propto$$

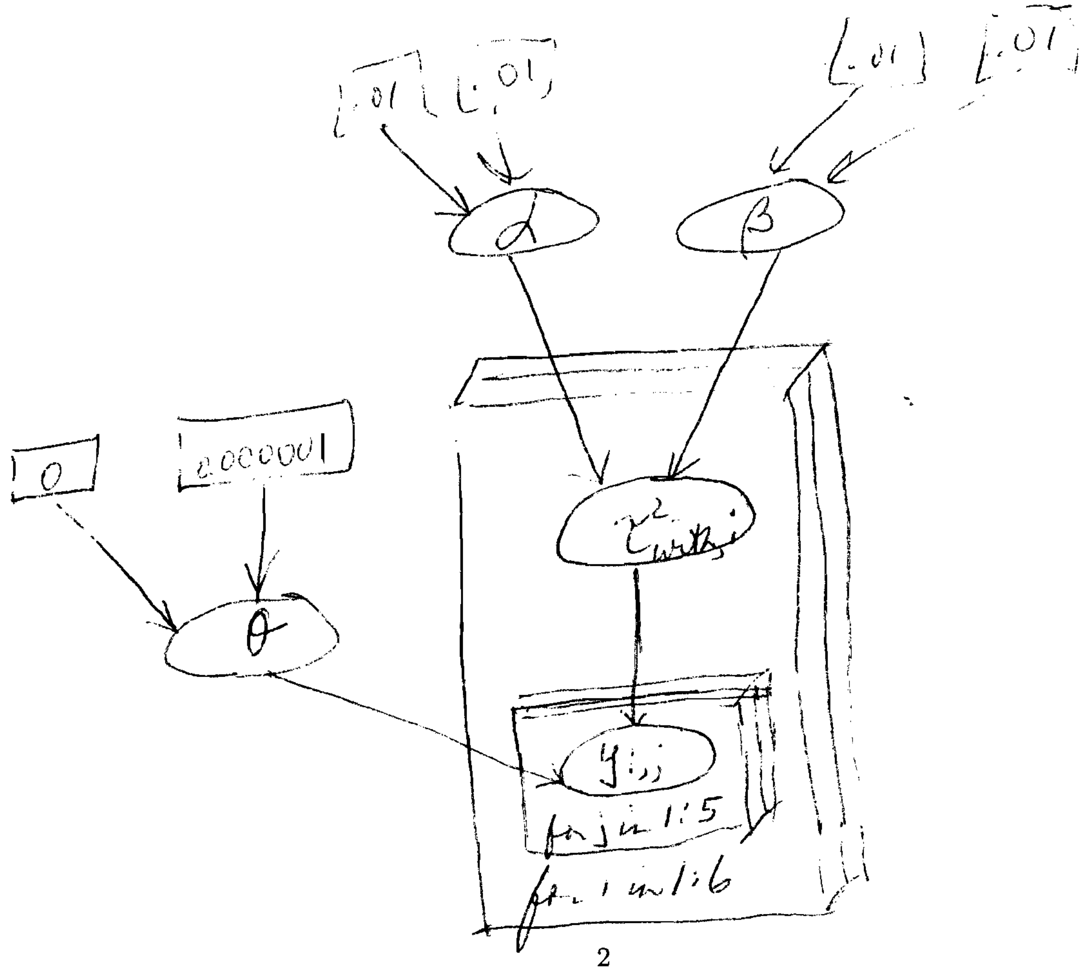
5

$$(\tau^2_{with,2})^{\alpha+\frac{5}{2}-1}\exp\left[-\left[\beta+\frac{\sum_{j=1}^{5}(y_{2j}-\theta)^2}{2}\right]\tau^2_{with,2}\right]$$

: Gamma $\left(\alpha+\frac{5}{2},\ \beta+\frac{\sum(y_{2j}-\theta)^2}{2}\right)$

(d) Draw a directed graph of your new model.

10



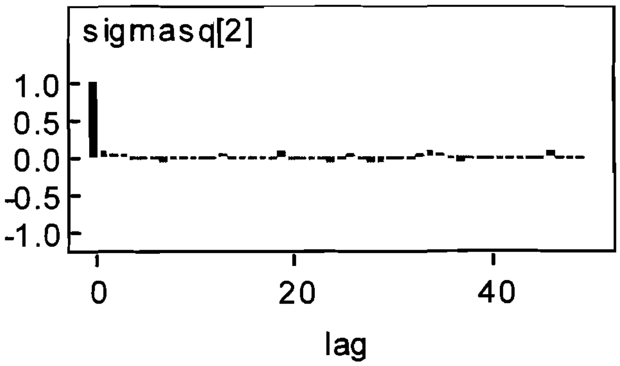
2. Below is a plot obtained from WinBUGS output.

   (a) What do the heights of the bars represent? (Explain in a sentence or two.).

   correlation between sampled values "lag" iterations apart

   (b) Would you expect the MCMC sampler on which this plot is based to converge slowly or quickly? Briefly explain.

   Quickly. The autocorrelation becomes negligible immediately – at lag 1.



3. Below is a line of output from WinBUGS. Which value gives you information about how accurately the posterior mean of this parameter can be estimated based on the MCMC samples run so far? Give a numeric answer, and then explain briefly.

   MC error = 135.0 It is the autocorrelation-adjusted standard error of the estimated posterior mean

node	mean	sd	MC error	2.5%	median	97.5%	start	sample
sigmasq[1]	4673.0	3354.0	135.0	1347.0	3649.0	12920.0	501	500