Bayesian Inference PSTAT 215A
Homework 3 (of 4) RBG/Fall 2009

Assigned 29 Oct 2009; due 10 Nov 2009. *You may study the questions/answers with others, but what you hand in for credit must be your own work.*

1. Sensitivity analysis: Thirty-two students in a science classroom were randomly assigned to one of two study methods, A and B, so that $n_A = n_B = 16$ were assigned to each method. After several weeks of study, students were examined on the course material with an exam designed to give an average score of 75 with a variance of 100. The scores for the two groups are summarized by $\{\bar{y}_A = 75.2,_A s^2 = 7.3^2\}$ and $\{\bar{y}_B = 77.5, s_B^2 = 8.1^2\}$. Consider independent, conjugate normal prior distributions for each of μ_A and μ_B, with a sensible setting for μ_0 and σ_0^2 for both groups, which you should choose (and justify). For each of $(\kappa_0, \nu_0) \in \{(1,1),(2,2),(4,4),(8,8),(16,16),(32,32)\}$ (or more values), obtain $P(\mu_A < \mu_B | \mathbf{y}_A, \mathbf{y}_B)$ via MC. Plot this probability as a function of $(\kappa_0 = \nu_0)$. Describe how you might use this plot to convey the evidence that $\mu_A < \mu_B$ to people of a variety of prior opinions.

2. Suppose that we wish to use the Gibbs sampler to sample from a "distribution" with the following density
$$\pi(\theta_1, \theta_2) = \exp\left\{-\frac{1}{2}(\theta_1 - 1)^2(\theta_2 - 2)^2\right\}.$$
Write down the two full conditionals for this "distribution". How would you sample from these? *Note this is correct; there is no plus between the squared terms. The full conditionals are derivable, and they should have a standard form. However, you should ask yourself if the resulting sampler "makes any sense"?*

3. In this question we shall reconsider the data from Question 1 from Homework 2. Assume Poisson sampling models for the two groups as before, but now parameterize θ_A and θ_B as $\theta_A = \theta$ and $\theta_B = \theta \times \gamma$. Here, γ represents the relative rate θ_B/θ_A. Let $\theta \sim G(a_\theta, b_\theta)$ and let $\gamma \sim G(a_\gamma, b_\gamma)$.
   (a) Are θ_A and θ_B independent or dependent under this prior distribution? In what situations is such a joint prior justified?
   (b) Obtain the full conditional distribution of θ given $\mathbf{y}_A$, $\mathbf{y}_B$ and γ.
   (c) Obtain the form of the full conditional distribution of γ given $\mathbf{y}_A$, $\mathbf{y}_B$ and θ.
   (d) Set $(a_\theta, b_\theta) = (2, 1)$. For each of $a_\gamma = b_\gamma \in \{8, 16, 32, 64, 128\}$, run a Gibbs sampler of at least 5,000 iterations and obtain an estimate of $\mathbb{E}\{\theta_B - \theta_A | \mathbf{y}_A, \mathbf{y}_B\}$. Describe the effects of the prior distribution for γ on this expectation.

4. Suppose that we wish to use the Metropolis Hastings algorithm to generate from $\pi(\theta) = \mathcal{N}(\theta; 0, \sigma^2)$ using the proposal $q(\theta, \phi) = \mathcal{N}(\phi; a\theta, \tau^2)$. What is the corresponding acceptance function $\alpha(\theta, \phi)$? For what value of a would this particular sampler never reject?

5. A data file on the course web page contains measurements of the amount of rain (in acre–feet) produced from 26 pairs of clouds. One set was "seeded" to try to increase the level of rainfall, and the other was not.

   (a) Explore the data (in R) and explain why a normal sampling model would be inappropriate for these data.

   (b) Consider using $G(\alpha_s, \beta_s)$ and $G(\alpha_u, \beta_u)$ sampling models for seeded and unseeded clouds, respectively, where the data are otherwise independent and identically distributed. Use independent exponential priors for the unknown parameters $\{\alpha_s, \beta_s, \alpha_u, \beta_u\}$ with a rate parameter of $\lambda = 5$. Develop a random–walk Metropolis–Hastings (RWMH) algorithm (i.e., proceeding coordinate–wise via *Metropolis–within–Gibbs*) to sample from the posterior distribution of the parameters. *Some hints:*

      i. *Work in log–space or you will have overflow problems.*

      ii. *The posterior for the α parameters will be on a different scale than β ones. So you must carefully tune your starting values proposal mechanisms for these values, and carefully choose the burn–in B. Be sure to comment on how you proceeded in your writeup. Also remember that these values must be positive.*

   (c) Show that the β parameters may be sampled with a "Gibbs step", since the full conditionals for β_u and β_s are familiar densities.

   (d) Develop a sampler that exploits the result from part (c). Compare its performance to the fully RWMH one you developed in part (b) in terms of trace plots and estimates of autocorrelation and effective sample size.

   (e) Report on a posterior summary statistic (or statistics) of your choosing, that you approximated from the MCMC samples, in order to support or refute a claim that the level of rainfall is higher for seeded clouds.