S 2016 STAT 775 HW # 8 Due May 5, 2016

1. a. For the genetic linkage model applied to the data $Y = (125, 18, 20, 34)$, implement the Metropolis algorithm (use a flat prior on θ). Use one long chain and plot θ^i versus i. Try several driver functions: (1) uniform on (0,1); (2) normal centered at the current point in the chain and standard deviation 0.01; (3) normal with same mean as in (2) and standard deviation 0.1; (4) normal with same mean as in (2) and standard deviation 0.5; (5) normal centered at 0.4 with standard deviation 0.1. Draw a histogram of the θ values and the true normalized likelihood in the same figure. Discuss the adequacy of the Metropolis estimate. How did you assess convergence? Did you toss out any initial values? Should you use the Metropolis or generalized Metropolis algorithm in part (5)?

   b. Repeat (a) for the data $Y = (14, 0, 1, 5)$.

   c. Compute both the posterior mean and standard deviation for both datasets. Compare to the results from the previous problem.

2. a. Consider the 1-way variance components model

$$Y_{ij} = \theta_i + \varepsilon_{ij},$$

where Y_{ij} is the jth observation from the ith group, θ_i is the effect of the ith group and ε_{ij} is the error, where $i = 1, \dots, K$ and $j = 1, \dots, J$. It is assumed that the ε_{ij} are an *iid* sample from $N(0, \sigma_\varepsilon^2)$ and the θ_i are an independent *iid* sample from $N(\mu, \sigma_\theta^2)$. Under the prior specification

$p(\sigma_\varepsilon^2, \sigma_\theta^2, \mu) = p(\sigma_\varepsilon^2)p(\sigma_\theta^2)p(\mu)$, with $p(\sigma_\theta^2) = IG(a_1, b_1)$, $p(\sigma_\varepsilon^2) = IG(a_2, b_2)$ and $p(\mu) = N(\mu_0, \sigma_0^2)$, show that

   i. $p(\mu \mid \theta, \sigma_\varepsilon^2, \sigma_\theta^2, Y) = N\left(\frac{\sigma_\theta^2\mu_0 + \sigma_0^2\Sigma\theta_i}{\sigma_\theta^2 + K\sigma_0^2}, \frac{\sigma_\theta^2\sigma_0^2}{\sigma_\theta^2 + K\sigma_0^2}\right)$

   ii. $p(\theta_i \mid \mu, \sigma_\varepsilon^2, \sigma_\theta^2, Y) = N\left(\frac{J\sigma_\theta^2}{J\sigma_\theta^2 + \sigma_\varepsilon^2}\bar{Y}_{i.} + \frac{\sigma_\varepsilon^2}{J\sigma_\theta^2 + \sigma_\varepsilon^2}\mu, \frac{\sigma_\theta^2\sigma_\varepsilon^2}{J\sigma_\theta^2 + \sigma_\varepsilon^2}\right)$,

   iii. $p(\sigma_\varepsilon^2 \mid \mu, \theta, \sigma_\theta^2, Y) = IG\left(a_2 + \frac{KJ}{2}, b_2 + \frac{1}{2}\sum_{i=1}^K \sum_{j=1}^J (Y_{ij} - \theta_i)^2\right.$

   iv. $p(\sigma_\theta^2 \mid \mu, \theta, \sigma_\varepsilon^2, Y) = IG\left(a_1 + \frac{K}{2}, b_1 + \frac{1}{2}\sum_{i=1}^K (\theta_i - \mu)^2\right)$,

   where $\bar{Y}_{i.} = \frac{1}{J}\sum_{j=1}^J Y_{ij}$ and $\theta = (\theta_1, \dots, \theta_K)$.

   b. Run the Gibbs sampler for the data below. Use one chain of length 75,000. Take $p(\mu) = N(0, 10^{12})$, $p(\sigma_\varepsilon^2) = IG(0, 0)$ and $p(\sigma_\theta^2) = IG(1, 1)$. For each θ_i, for σ_ε, and for σ_θ, plot the simulated value at iteration j versus j. Summarize each posterior marginal.

   c. Repeat (b) using the prior specification $p(\mu) = N(0, 10^{12})$, $p(\sigma_\varepsilon^2) = IG(0, 0)$ and $p(\sigma_\theta^2) = IG(0, 0)$. Does this specification violate the Hobert–Casella conditions? Describe what happens to the Gibbs sampler chain in this case.

   d. Repeat (b) with $p(\sigma_\varepsilon^2) = \textit{constant}$ and $p(\sigma_\mu^2) = \textit{constant}$.

Data Generated from a Table of Random Normal Deviates (Box and Tiao, 1973).

Batch	Observations				
1	7.298	3.846	2.434	9.566	7.990
2	5.220	6.556	0.608	11.788	−0.892
3	0.110	10.386	13.434	5.510	8.166
4	2.212	4.852	7.092	9.288	4.980
5	0.282	9.014	4.458	9.446	7.198
6	1.722	4.782	8.106	0.758	3.758

Hobert and Casella (1993) prove the following result: Consider the model:

$$Y_{ij} = \beta + \mu_i + \varepsilon_{ij},$$

where $i = 1, \dots, a$, $j = 1, \dots, n_i$, the μ_i are an *iid* sample from $N(0, \sigma_\mu^2)$, and the ε_{ij} are an *iid* sample from $N(0, \sigma_\varepsilon^2)$, with independent priors $p(\beta) \propto \textit{constant}$, $p(\sigma_\varepsilon^2) \propto (\sigma_\varepsilon^2)^{-(b+1)}$ and $p(\sigma_\mu^2) \propto (\sigma_\mu^2)^{-(c+1)}$. The resulting posterior distribution will be proper if and only if $a - 1 > -2c > 0$ and $N + 2c + 2b - 1 > 0$.

Problem 3

Let y_i be normally distributed with mean θ and variance 1. Suppose the prior distribution of θ is a t-distribution with 2 degrees of freedom, and center = 2 and scale =1. Three observations have been collected: $y_1 = 3.3$, $y_2 = 2.5$, and $y_3 = 4.2$. Use each the following methods (1) to simulate 100 values of θ from the posterior distribution; (2) to plot the histogram for these 100 values. Remember to check the convergence for methods 2 and 3.

(a) Method 1: Rejection sampling.
(b) Method 2: Gibbs sampling.
(c) Method 3: Metropolis algorithm (try two different transition functions or drivers for this algorithm).