Stat 8701, Spring 2003, Homework 5

The following data are taken from Bishop, Fienberg and Holland (*Discrete Multivariate Analysis*, MIT Press, 1975, Section 5.2.8, Exercise 4).

Active	Passive Participant					
Participant	R	S	T	U	V	W
R	—	1	5	8	9	0
S	29	—	14	46	4	0
T	0	0	—	0	0	0
U	2	3	1	—	38	2
V	0	0	0	0	—	1
W	9	25	4	6	13	—

The data are in the file

```
http://www.stat.umn.edu/~charlie/monkey.dat
```

and the R statements

```
foo <- scan(url("http://www.stat.umn.edu/~charlie/monkey.dat"))
bar <- matrix(foo, 6, 6, byrow = TRUE)
```

reads it and stuffs it into a matrix.

The data involve displays between squirrel monkeys (there are six monkeys labeled R through W). Each display has an active and a passive participant giving the two-way classification of the table. Since a monkey cannot display toward itself, the diagonal cells of the table are impossible (structural zeros in the jargon of categorical data analysis).

We wish to fit the following Bayesian model to these data (the same model for which Bishop, Fienberg, and Holland discuss frequentist analysis in their Section 5.2), called the *quasi-independence* model. We assume Poisson sampling, so the likelihood is

$$\prod_{i=1}^{6} \prod_{\substack{j=1 \\ j \neq i}}^{6} \mu_{ij}^{x_{ij}} e^{-\mu_{ij}} \tag{1}$$

where x_{ij} is the data (given in the table above) and the parameters μ_{ij} have the log-linear additive form

$$\log(\mu_{ij}) = \alpha + \beta_i + \gamma_j$$

where in order to obtain identifiability we must have the constraints

$$\sum_{i=1}^{6} \beta_i = 0 \tag{2}$$

and

$$\sum_{j=1}^{6} \gamma_j = 0. \tag{3}$$

Because the third row of the data is all zero, an improper prior would give an improper posterior. So we can't use a flat prior. The conjugate prior we use looks like a gamma with density proportional to

$$\prod_{i=1}^{6} \prod_{\substack{j=1 \\ j \neq i}}^{6} \mu_{ij}^{\epsilon} e^{-\delta \mu_{ij}} \tag{4}$$

but it isn't gamma because we do not consider the μ_{ij} the fundamental random variables but the α, β_i, and γ_j.

Thus the unnormalized posterior density is the product of (1) and (3) considered as a function of the α, β_i, and γ_j. The specific values of the hyperparameters we use are $\epsilon = 0.2$ and $\delta = 0.3$.

(a) Write a MCMC sampler for this 13-dimensional posterior distribution. Also write a description of your algorithm with sufficient detail that it could be reimplemented by an MCMC expert.

(b) Estimate the posterior mean of each of the 13 parameters, α, β_i, and γ_j.

(c) Estimate Monte Carlo standard error (MCSE) for each estimated quantity (the posterior mean of each of the 13 parameters). Try to get at least two significant figures (Monte Carlo standard error a hundredth the size of the estimated quantity). Estimate MCSE two different ways.

   (i) Use the method of batch means. Report your batch size and whether you use overlapping or nonoverlapping batch means.

   (ii) Use the method of Geyer (1992) for reversible Markov chains (this requires that you make your MCMC sampler reversible). Calculate the "big gamma" function $\Gamma_k = \gamma_{2k} + \gamma_{2k+1}$. where the "little gamma" function is the autocovariance function of the time series averaged to get some estimate. Let K be the largest integer such that $\Gamma_0, \ldots, \Gamma_K$ are all positive. Then estimate MCSE as

$$\sqrt{\frac{1}{n}\left(-\gamma_0 + 2\sum_{i=1}^{K} \Gamma_i\right)}$$