

# TMA4300 MCMC Practice

Author 1, Author 2, Author 3

*The document is intended as a tool for students to practice coding up their own MCMC sampler. Students are welcome to work together in groups, and to ask questions as they progress in this worksheet. When coding, students should take turns:*

1. *coding up various components of the samplers, and*
2. *looking over and double checking other group members' code as it is being written.*

*Make sure your code is well commented and has understandable variable and function names.*

## Problem setup

Assume we are interested in the response, relative to a baseline, of a patient after being assigned to one of two groups. We have  $n$  patients in a treatment group that receive medication for a condition, and  $n$  other patients in a control group that receive a placebo. Measurements are obtained for the  $n$  patients in each group, denoted by  $Y_{Ti}$  and  $Y_{Ci}$  for the treatment and control groups respectively, and for  $i = 1, 2, \dots, n$ .

The responses are modeled as, for  $G \in \{T, C\}$  denoting the patient group,

$$Y_{Gi} = \mu_G + \epsilon_{Gi},$$

where we model  $\mu_G$  as a Gaussian latent effect with prior  $\mu_T, \mu_C \mid \nu^2 \stackrel{iid}{\sim} N(0, \nu^2)$ , we assume Gaussian error,  $\epsilon_{Ti}, \epsilon_{Ci} \mid \sigma^2 \stackrel{iid}{\sim} N(0, \sigma^2)$  for  $i = 1, \dots, n$ , and we place inverse gamma hyperpriors on both  $\sigma^2$  and  $\nu^2$  so that  $\sigma^2, \nu^2 \stackrel{iid}{\sim} \text{Inv-Gamma}(\alpha, \beta)$ . Hence, our hyperpriors have density:

$$p(\sigma^2) = \frac{\beta^\alpha}{\Gamma(\alpha)} (1/\sigma^2)^{\alpha+1} \exp\{-\beta/\sigma^2\}$$
$$p(\nu^2) = \frac{\beta^\alpha}{\Gamma(\alpha)} (1/\nu^2)^{\alpha+1} \exp\{-\beta/\nu^2\}.$$

Let  $\mathbf{Y}_T = Y_{T1}, \dots, Y_{Tn}$  and  $\mathbf{Y}_C = Y_{C1}, \dots, Y_{Cn}$ .

## 1

Draw a conditional dependency graph for  $\mathbf{Y}_T, \mathbf{Y}_C, \mu_T, \mu_C, \sigma^2$ , and  $\nu^2$ . (Why can we ignore the  $\epsilon$  terms here?)

## 2

Give an expression that the posterior density,  $p(\mu_T, \mu_C, \sigma^2, \nu^2 \mid \mathbf{Y}_T, \mathbf{Y}_C)$ , is proportional to, removing unnecessary multiplicative constants. Make sure to factor out the final result into contributions due to the likelihood, latent effects, and hyperpriors.

### 3

Find the conditional density  $p(\nu^2 \mid \mu_T, \mu_C, \sigma^2, \mathbf{Y}_T, \mathbf{Y}_C)$  and simplify it, removing unnecessary multiplicative constants and combining exponential terms. Name the distribution and provide its parameters.

Do the same for  $p(\sigma^2 \mid \mu_T, \mu_C, \nu^2, \mathbf{Y}_T, \mathbf{Y}_C)$  (and, optionally, for  $p(\mu_T, \mu_C \mid \sigma^2, \nu^2, \mathbf{Y}_T, \mathbf{Y}_C)$ ).

### 4

Despite the fact that we could use a Gibbs step for every parameter due to the closed forms of the conditional distributions, assume we wish to take a bivariate Gaussian Metropolis step for the parameters  $\mu_T$  and  $\mu_C$  with step variance  $\tau^2 I_2$  for  $2 \times 2$  identity matrix  $I_2$ . We will let  $\mathbf{x} = (\sigma^2, \nu^2, \mu_T, \mu_C)$  be the current state of our MCMC sampler so that  $\mathbf{x}^j$  is the  $j$ -th element of  $\mathbf{x}$ .

What is the **log** acceptance probability of the proposal  $\mathbf{y} = (\sigma^2, \nu^2, \tilde{\mu}_T, \tilde{\mu}_C)$  where  $\mathbf{y}^{3,4} \sim N_2(\mathbf{x}^{3,4}, \tau^2 I_2)$ ?

### 5

Now for the fun part: we will code up our own hybrid Metropolis-within-Gibbs MCMC sampler. We will let  $\mathbf{x} = (\sigma^2, \nu^2, \mu_T, \mu_C)$  so that  $\mathbf{x}_i^j$  is the  $j$ -th element of the  $i$ -th sample from our MCMC sampler. The MCMC algorithm will proceed as follows:

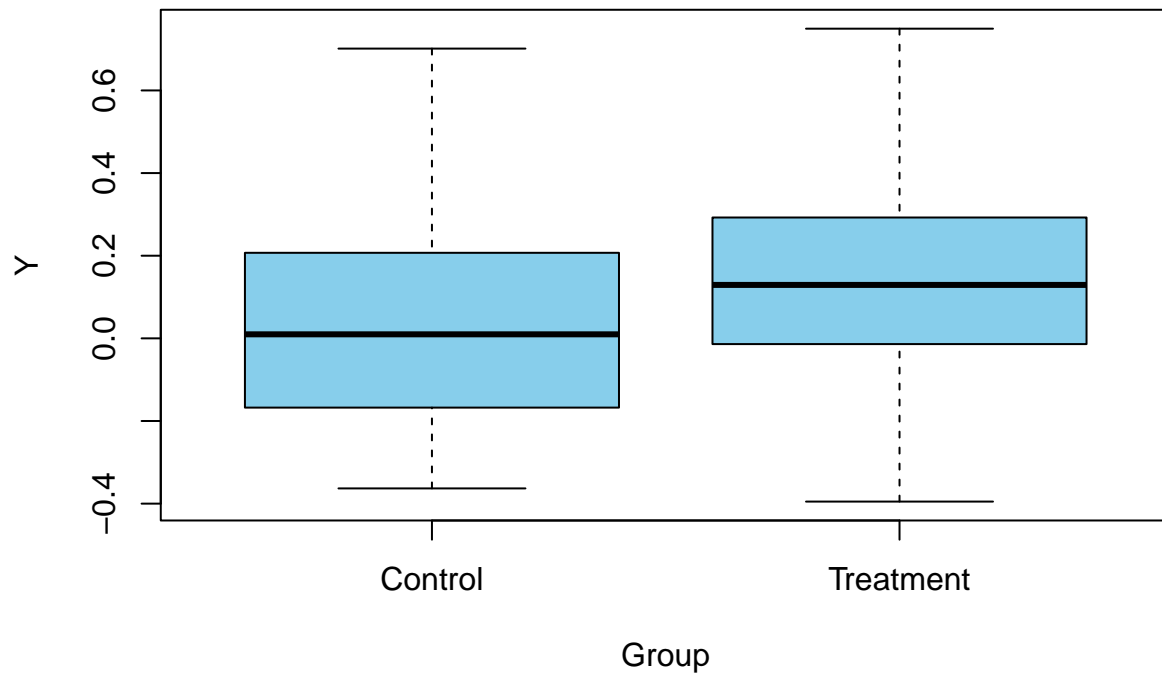
1. Set  $\mathbf{x} \leftarrow \mathbf{x}_0 \leftarrow (\sigma^2, \nu^2, \mu_T, \mu_C)$  to a reasonable initial value
2. For  $i$  in 1 to  $M$  iterations:
  - (a) Gibbs step: set  $x^1$  to a random draw from the conditional  $x^1 \mid \mathbf{x}^{-1}, \mathbf{Y}_T, \mathbf{Y}_C$
  - (b) Gibbs step: set  $x^2$  to a random draw from the conditional  $x^2 \mid \mathbf{x}^{-2}, \mathbf{Y}_T, \mathbf{Y}_C$
  - (c) MH step: propose  $\mathbf{y}^{3,4} \sim N_2(\mathbf{x}^{3,4}, \tau^2 I_2)$ . Accept the proposal with the probability from problem 4, setting  $\mathbf{x}^{3,4} \leftarrow \mathbf{y}^{3,4}$  if accept, or  $\mathbf{x}^{3,4} \leftarrow \mathbf{x}_{i-1}^{3,4}$  if reject.
  - (d) Update current state: set  $\mathbf{x}_i \leftarrow \mathbf{x}$
3. return  $\mathbf{x}_0, \dots, \mathbf{x}_M$

Assume  $M = 100000$  and remove 10000 samples due to burn-in. We will let the inverse gamma hyperprior parameters be  $\alpha = 2$  and  $\beta = 0.05$ . Assume the following are our observations:

```
library(invgamma)

# simulate data based on true parameters
alpha = 2
beta = 0.05
set.seed(1)
sigma2 = rinvgamma(1, alpha, beta)
nu2 = rinvgamma(1, alpha, beta)
muT = rnorm(1, sd=sqrt(nu2))
muC = rnorm(1, sd=sqrt(nu2))
n=100
YT = rnorm(n, muT, sd=sqrt(sigma2))
YC = rnorm(n, muC, sd=sqrt(sigma2))

# make dataset and a boxplot of the responses in the 2 groups
dat = data.frame(Group=c(rep("Treatment", n), rep("Control", n)), Y=c(YT, YC))
boxplot(Y ~ Group, data=dat, col="skyblue")
```



Construct your MCMC sampler. Decide on a reasonable starting value and a step size resulting in reasonably good mixing. Keep track of your acceptance probability, and report it. Make traceplots and histograms of all variables, and compare estimates of the mean of each distribution to the true values generated above. Has the chain converged? What do the results imply about the treatment? Does it significantly improve (increase) patient outcomes relative to the placebo?