STAT 654 Problem Set 3

Nikko Cleri cleri@tamu.edu

April 24, 2022

All relevant R code is appended after the answers to all questions.

A

1. Deriving the full posteriors, we start with Bayes Theorem, which tells us that

$$P(\sigma^2, \mu|x) = \frac{P(x|\sigma^2, \mu)P(\sigma^2, \mu)}{\int\int P(x|\sigma^2, \mu)P(\sigma^2, \mu)\,\mathrm{d}\sigma^2\,\mathrm{d}\mu}$$

where $P(x|\sigma^2, \mu)$ is the data likelihood

$$\prod_{i=1}^{n} \frac{1}{\sqrt{2\pi\sigma^2}} e^{-\frac{1}{2}\left(\frac{x_i-\mu}{\sigma}\right)^2} = \frac{1}{(\sqrt{2\pi\sigma^2})^n} e^{-\frac{1}{2\sigma^2}\sum_{i=1}^{n}(x_i-\mu)^2}$$

assuming independence between μ and σ^2, we can say $P(\sigma^2, \mu) = P(\sigma^2)P(\mu)$. Using the definitions from the prompt we have

$$P(\sigma^2) \sim \mathrm{IG}(1,1) = \frac{1}{)(\sigma^2)^2} e^{-1/\sigma^2}$$
$$P(\mu) \sim \mathrm{N}(2,1) = \frac{1}{\sqrt{2\pi}} e^{1\frac{1}{2}(\mu-2)^2}$$

we then get the full posterior

$$P(\sigma^2, \mu|x) = \frac{\frac{1}{(\sqrt{2\pi\sigma^2})^n} e^{-\frac{1}{2\sigma^2}\sum_{i=1}^{n}(x_i-\mu)^2} \frac{1}{(\sigma^2)^2} e^{-1/\sigma^2} \frac{1}{\sqrt{2\pi}} e^{1\frac{1}{2}(\mu-2)^2}}{\int\int \frac{1}{(\sqrt{2\pi\sigma^2})^n} e^{-\frac{1}{2\sigma^2}\sum_{i=1}^{n}(x_i-\mu)^2} \frac{1}{(\sigma^2)^2} e^{-1/\sigma^2}\,\mathrm{d}\sigma^2\,\mathrm{d}\mu\, \frac{1}{\sqrt{2\pi}} e^{1\frac{1}{2}(\mu-2)^2}\,\mathrm{d}\sigma^2\,\mathrm{d}\mu}$$

For the full conditionals, we know that

$$\begin{aligned} P(\sigma^2|\mu, x) &\propto P(x|\sigma^2, \mu)P(\sigma^2, \mu) \\ &\propto \frac{1}{(\sqrt{2\pi\sigma^2})^n} e^{-\frac{1}{2\sigma^2}\sum_{i=1}^{n}(x_i-\mu)^2} \frac{1}{(\sigma^2)^2} e^{-1/\sigma^2} \frac{1}{\sqrt{2\pi}} e^{1\frac{1}{2}(\mu-2)^2} \end{aligned}$$

We can disregard constants in the proportionalities. In the following we treat x as a constant:

$$\begin{aligned} P(\sigma^2|\mu, x) &\propto \frac{1}{(\sigma^2)^{n/2}} e^{-\frac{1}{2\sigma^2}\sum_{i=1}^{n}(x_i-\mu)^2} \frac{1}{(\sigma^2)^2} e^{-1/\sigma^2} \\ &\propto \frac{1}{(\sigma^2)^{n/2+2}} e^{-\frac{1}{\sigma^2}\sum_{i=1}^{n}(x_i-\mu)^2} \end{aligned}$$

which is the kernel of an $\text{IG}(n/2+1, 1/2\sum_{i=1}^{n}(x_i-\mu)^2)$ distribution. For $P(\mu|\sigma^2,\mu)$ we have (working with proportionalities and dropping constants)

$$\begin{aligned} P(\mu|\sigma^2,\mu) &\propto P(x|\sigma^2,\mu)P(\sigma^2,\mu) \\ &\propto e^{-\frac{1}{2\sigma^2}\left[\sum_{i=1}^{n}(x_i-\mu)^2+\sigma^2(\mu-2)\right]} \end{aligned}$$

Completing the square gives us

$$P(\mu|\sigma^2,\mu) \propto e^{-\frac{1}{2}\frac{n+\sigma^2}{2\sigma^2}\left(\mu-\frac{\sum_{i=1}^{n}(x_i)+2\sigma^2}{n+2\sigma^2}\right)^2}$$

which is the kernel of a $\text{N}\left(\frac{\sum_{i=1}^{n}(x_i)+2\sigma^2}{n+2\sigma^2}, \frac{\sigma^2}{n+\sigma^2}\right)$ distribution.

2. The following plots show the posterior densities and the posterior MCMC samples for both parameters. Here we choose a true mean and variance for our generated data of 10 and 5, respectively. We also choose a number of iterations of 10000 and a burn-in of 5000. We find a posterior mean of each parameter $\mu \approx 5016$ and $\sigma^2 \approx 2.5 \times 10^7$. See the following figures.

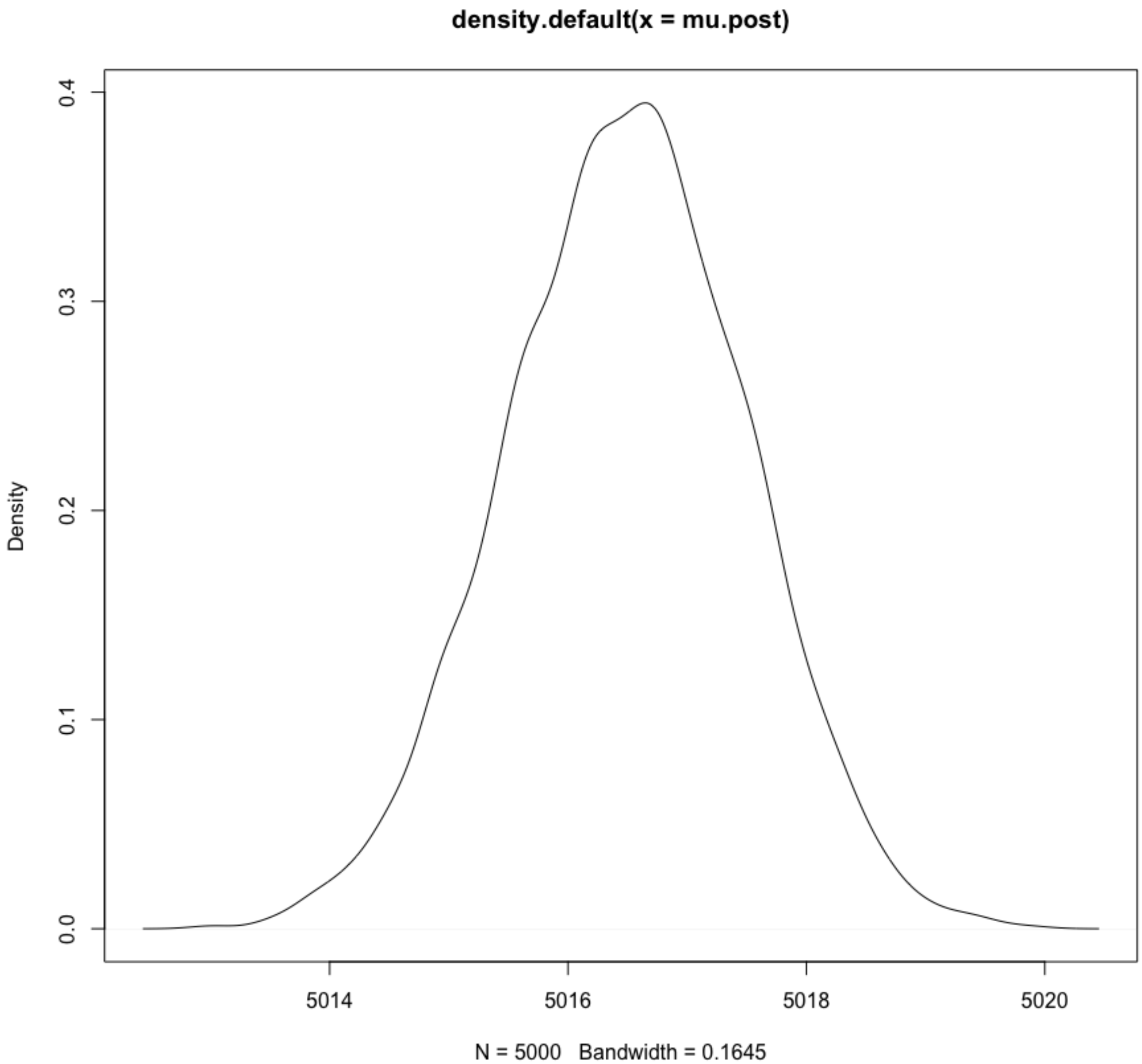

Figure 1: Kernel density of μ distribution post burn-in. We find a posterior mean of 5016.

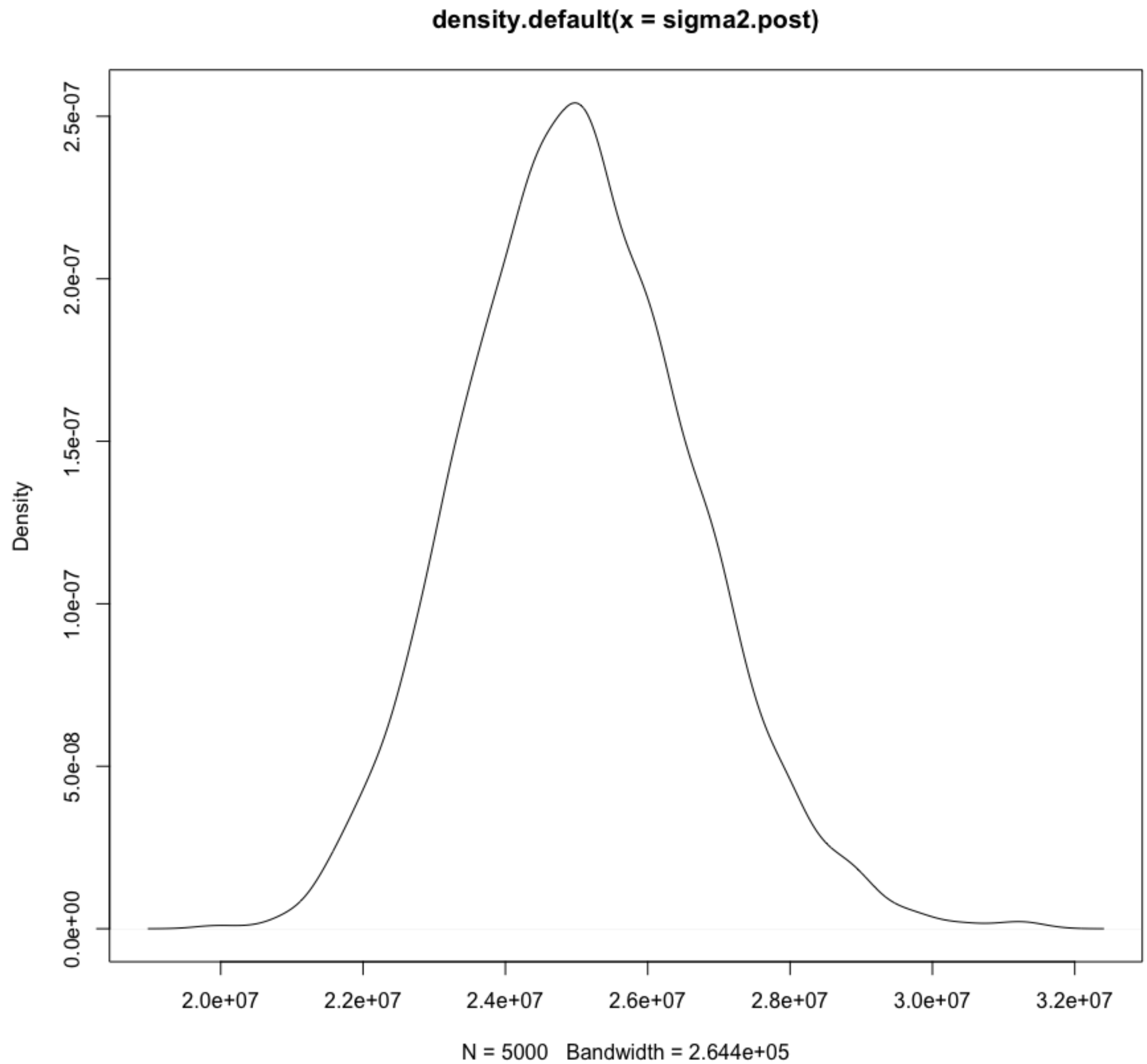

Figure 2: Kernel density of σ^2 distribution post burn-in. We find a posterior mean of 2.5×10^7.

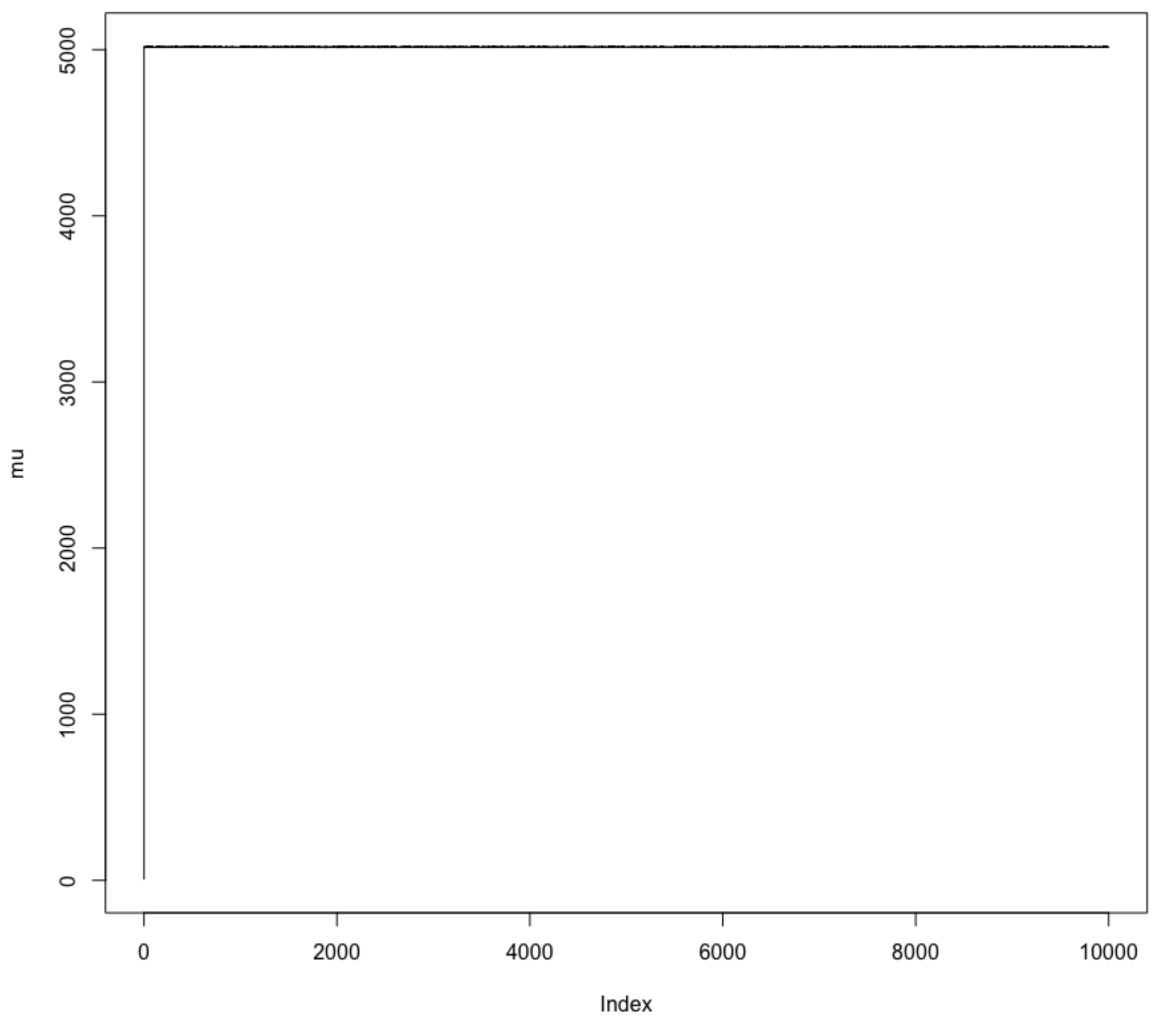

Figure 3: Full MCMC for μ.

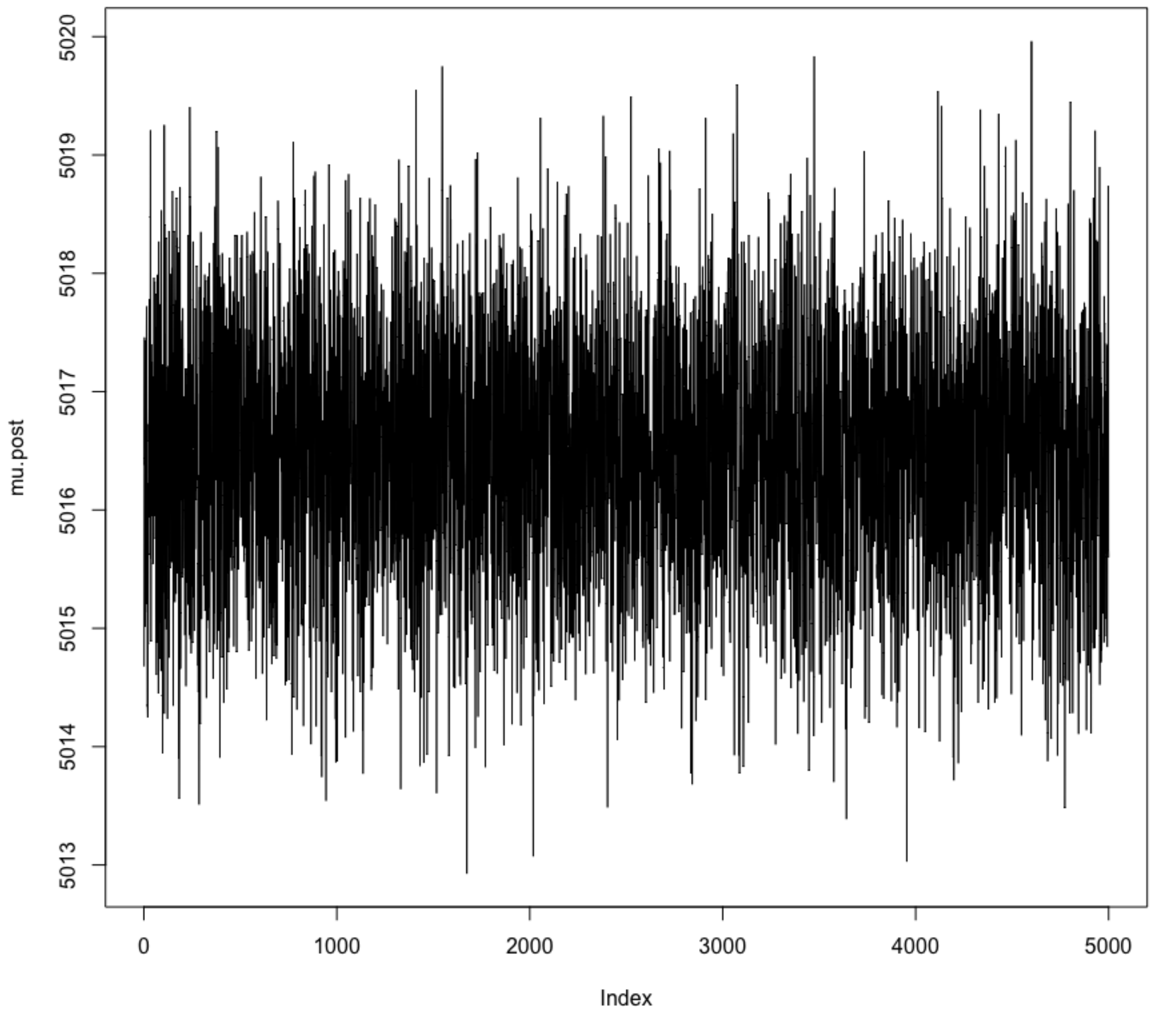

Figure 4: MCMC for μ post convergence. We see that the MCMC sampler has converged.

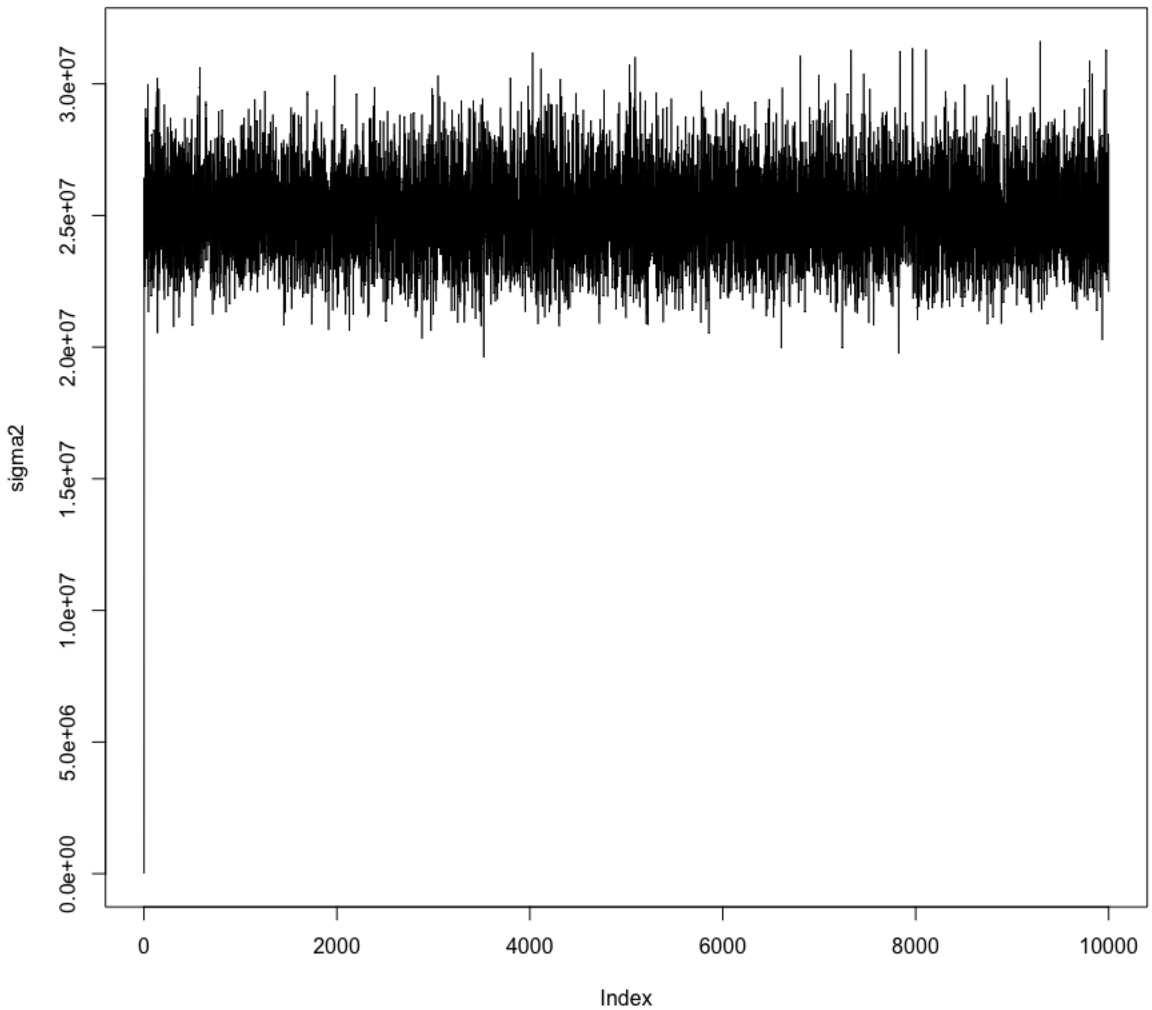

Figure 5: Full MCMC for σ^2.

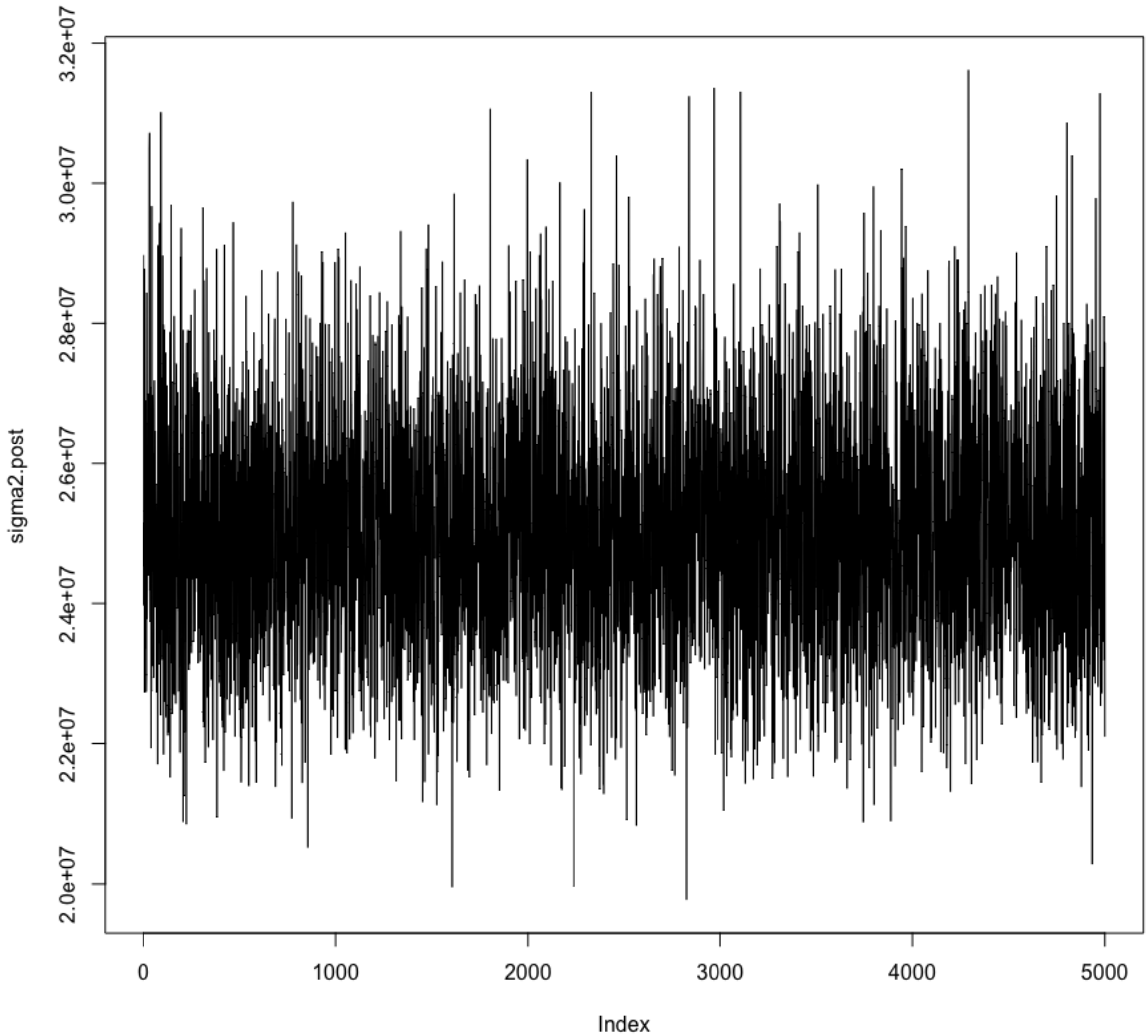

Figure 6: MCMC for σ^2 post convergence. We see that the MCMC sampler has converged.

```r
library(pscl)

## Generate Data
true.mu <- 10
true.sigma <- 5
n <- 500
y <- rnorm(n, true.mu, true.sigma)
ybar <- mean(y)
plot(density(y))

niter <- 10000  ## No. of Iterations.
burnin <- 5000  ## No. of burn-in samples (We will throw out this no. of
initial draws.)

##################
## Initialize ####
##################
mu <- c(rep(NA,niter))
sigma2 <- c(rep(NA,niter))
mu[1] <- 5  ## Initial Value
sigma2[1] <- 1  ## Initial Value

######################
## Gibbs Sampling ####
######################
for (i in 2:niter){
  mu[i] <- rnorm(1, sum(y)+2*sigma2[i-1]/(n + sigma2[i-1]), sigma2[i-1]/(n
+ sigma2[i-1]))
  sigma2[i] <- rigamma(1, n/2+1, sum((y - mu[i])^2)/2)
  print(i)
}

## Plot
plot(mu, type = "l")
mu.post <- mu[(burnin+1):niter]
plot(mu.post, type = "l")
plot(sigma2, type = "l")
sigma2.post <- sigma2[(burnin+1):niter]
plot(sigma2.post, type = "l")

plot(density(mu.post))
plot(density(sigma2.post))

mean(mu.post)
mean(sigma2.post)
```