Homework 7

Xiaokang Liu

26 October 2018

Contents

1 The likelihood function

From the normal mixture example, we know that the likelihood function for the data is

$$f(x|\theta) = \prod_{i=1}^{n}[\delta\phi(x_i, \mu_1, \sigma_1^2) + (1-\delta)\phi(x_i, \mu_2, \sigma_2^2)]$$

where $\phi()$ is the density function of a normal variable. Suppose that prior for μ_1 and μ_2 are $N(0, 100)$, that the prior for $1/\sigma_1^2$ and $1/\sigma_2^2$ are $\Gamma(a, b)$ with shape $a = 0.5$ and scale $b = 10$, and for δ we use noninformative prior, the unifrom distribution. Further, all the priors are independent. For simplicity, let's write the prior densities for $\mu_1, \mu_2, \sigma_1^2, \sigma_2^2$ are f_1, f_2, f_3 and f_4, respectively. Thus, the joint posterior distribution for those five parameters is

$$p(\theta|x) \propto f(x|\theta)f_1(\mu_1)f_2(\mu_2)f_3(\sigma_1^2)f_4(\sigma_2^2)$$

where $\theta = (\mu_1, \mu_2, \sigma_1^2, \sigma_2^2)$. For applying Gibbs sampling, we have to get the full conditional distribution for each parameter, they are:

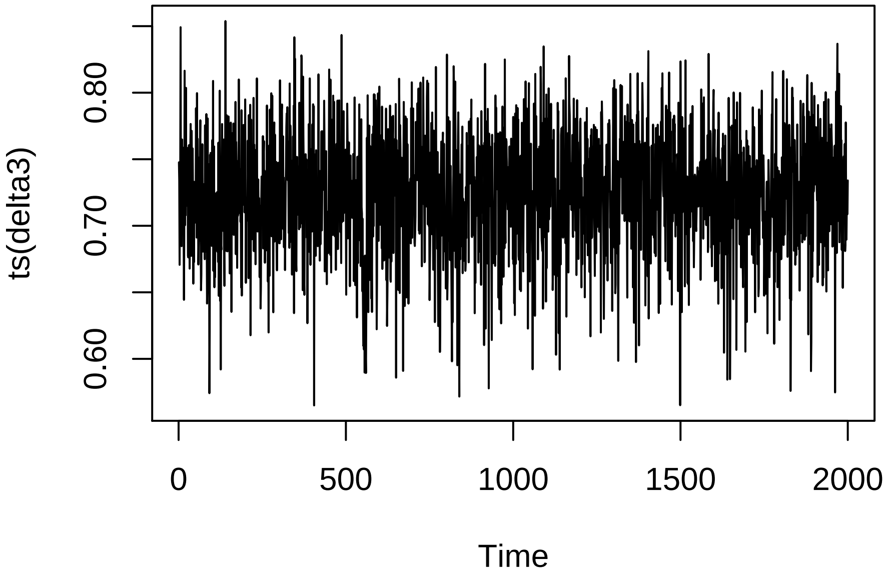
$$\begin{aligned}
p(\delta|x, \mu_1, \mu_2, \sigma_1^2, \sigma_2^2) &\propto f(x|\theta) \\
p(\mu_1|x, \delta, \mu_2, \sigma_1^2, \sigma_2^2) &\propto f(x|\theta)f_1(\mu_1) \\
p(\mu_2|x, \delta, \mu_1, \sigma_1^2, \sigma_2^2) &\propto f(x|\theta)f_2(\mu_2) \\
p(\sigma_1^2|x, \delta, \mu_2, \mu_1, \sigma_2^2) &\propto f(x|\theta)f_3(\sigma_1^2) \\
p(\sigma_2^2|x, \delta, \mu_2, \mu_1, \sigma_1^2) &\propto f(x|\theta)f_4(\sigma_2^2).
\end{aligned}$$

Or we can directly use the relationship:

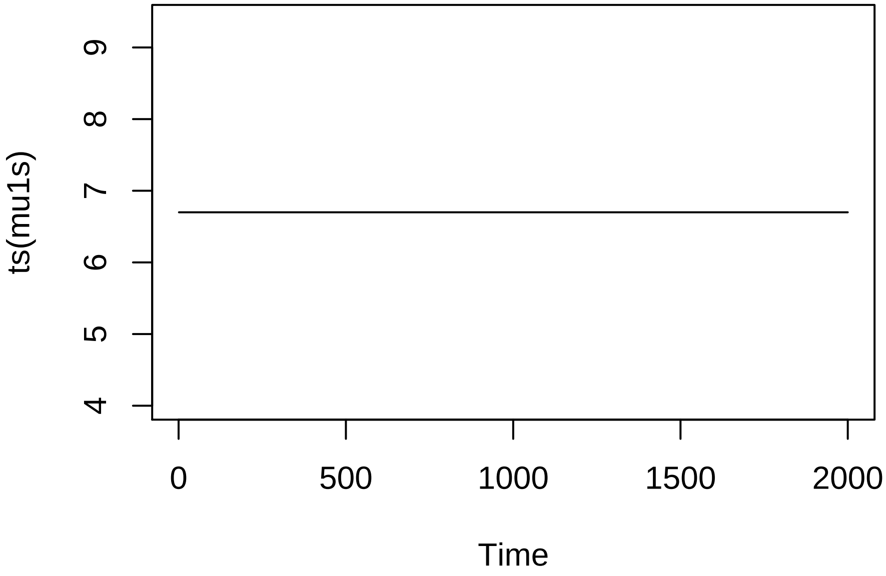
$$\begin{aligned}
p(\delta|x, \mu_1, \mu_2, \sigma_1^2, \sigma_2^2) &\propto f(x|\theta) \\
p(\mu_1|x, \delta, \mu_2, \sigma_1^2, \sigma_2^2) &\propto f(x|\theta) \\
p(\mu_2|x, \delta, \mu_1, \sigma_1^2, \sigma_2^2) &\propto f(x|\theta) \\
p(\sigma_1^2|x, \delta, \mu_2, \mu_1, \sigma_2^2) &\propto f(x|\theta) \\
p(\sigma_2^2|x, \delta, \mu_2, \mu_1, \sigma_1^2) &\propto f(x|\theta).
\end{aligned}$$

Then we can conduct Gibbs sampling step by step. For δ, we can just use the rejection sampling method from the example to sample from the conditional density. For all other methods, we can use rejection sampling function *arms* in R-package *HI*. The codes for applying the algorithm is like follows:

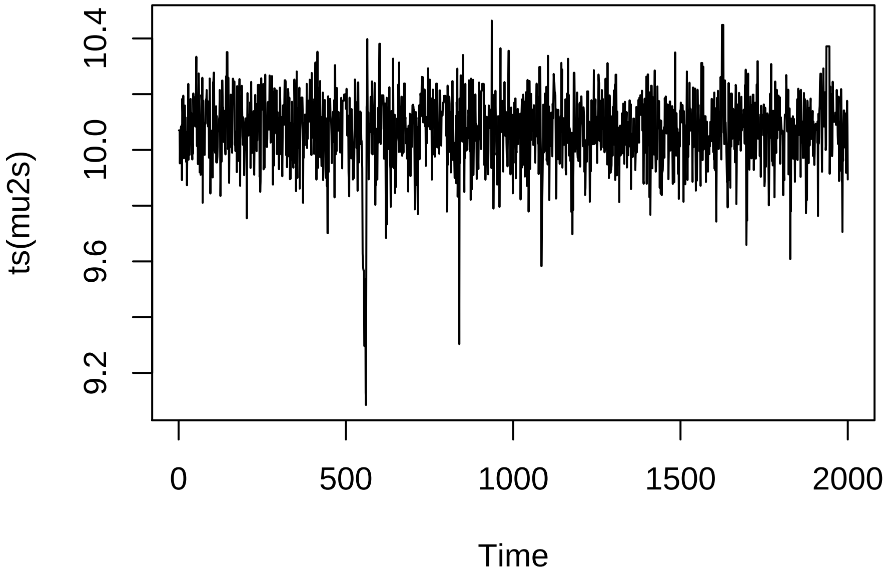
```
### Generate data: delta=0.7, mu_1=7, mu_2=10, sigma_1=sigma_2=0.5
delta <- 0.7
n <- 100
set.seed(254)
u <- rbinom(n, prob = delta, size = 1)
x <- rnorm(n, ifelse(u == 1, 7, 10), 0.5)


mylike <- function(delta,mu1,mu2,sigma1,sigma2,x) {
#  prod(delta * dnorm(x, mu1, sqrt(sigma1)) + (1 - delta) * dnorm(x, mu2, #sqrt(sigma2)))*dnor
  prod(delta * dnorm(x, mu1, sqrt(sigma1)) + (1 - delta) * dnorm(x, mu2, sqrt(sigma2)))*dnorm(
}

mymcmc5 <- function(niter, init, x){
  res <- matrix(nrow = niter, ncol = 5)
  for (i in 1:niter){
  init.delta <- init[1]
  init.mu1 <- init[2]
  init.mu2 <- init[3]
  init.sigma1 <- init[4]
  init.sigma2 <- init[5]

  f0 <- function(x0)  log(mylike(x0,init.mu1,init.mu2,init.sigma1,init.sigma2,x))
  new.delta <- res[i,1] <- arms(init.delta, f0, function(x0) (x0>0)*(x0<1), 1)

  f1 <- function(x1)  log(mylike(new.delta,x1,init.mu2,init.sigma1,init.sigma2,x))
  new.mu1 <- res[i,2] <- arms(init.mu1, f1, function(x1) (x1>-50)*(x1<100), 1)
  #new.mu1 <- res[i,2] <- arms(rnorm(1,6,1), f1, function(x1) (x1>-50)*(x1<100), 1)

  f2 <- function(x2)  log(mylike(new.delta,new.mu1,x2,init.sigma1,init.sigma2,x))
  new.mu2 <- res[i,3] <- arms(init.mu2, f2, function(x2) (x2>-50)*(x2<100),1)

  f3 <- function(x3)  log(mylike(new.delta,new.mu1,new.mu2,x3,init.sigma2,x))
  new.sigma1 <- res[i,4] <- arms(init.sigma1, f3, function(x3) (x3>0)*(x3<5), 1)

  f4 <- function(x4)  log(mylike(new.delta,new.mu1,new.mu2,new.sigma1,x4,x))
  new.sigma2 <- res[i,5] <- arms(init.sigma2, f4, function(x4) (x4>0)*(x4<5),1)

  new.par <- c(new.delta,new.mu1,new.mu2,new.sigma1,new.sigma2)
  init <- new.par
  }
  res
}

niter <- 3000
init <- c(0.5,5.7,9,0.2,0.2)
res <- mymcmc5(niter, init, x)
res <- res[-c(1:1000),]
delta3 <- res[,1]
mu1s <- res[,2]
mu2s <- res[,3]
sigma1s <- res[,4]
sigma2s <- res[,5]

plot(ts(delta3))
```



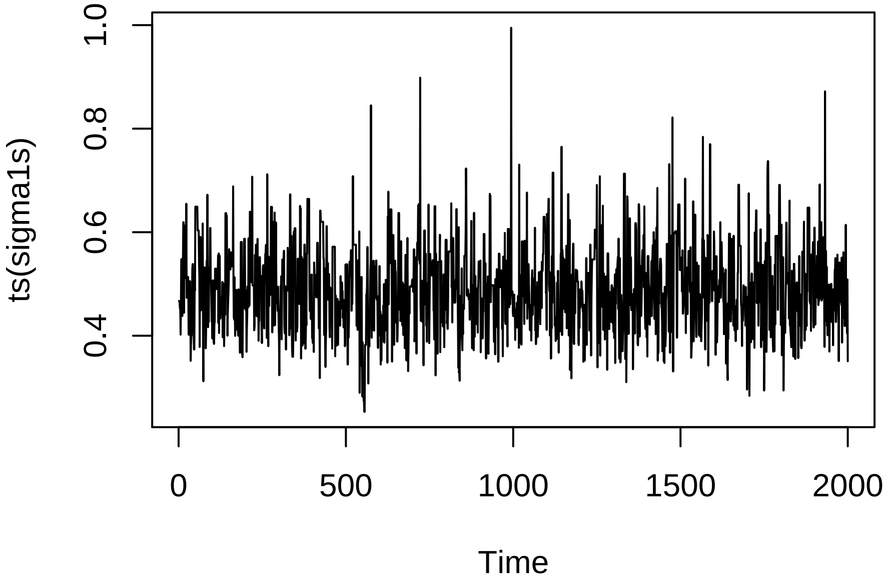
```
plot(ts(mu1s))
```



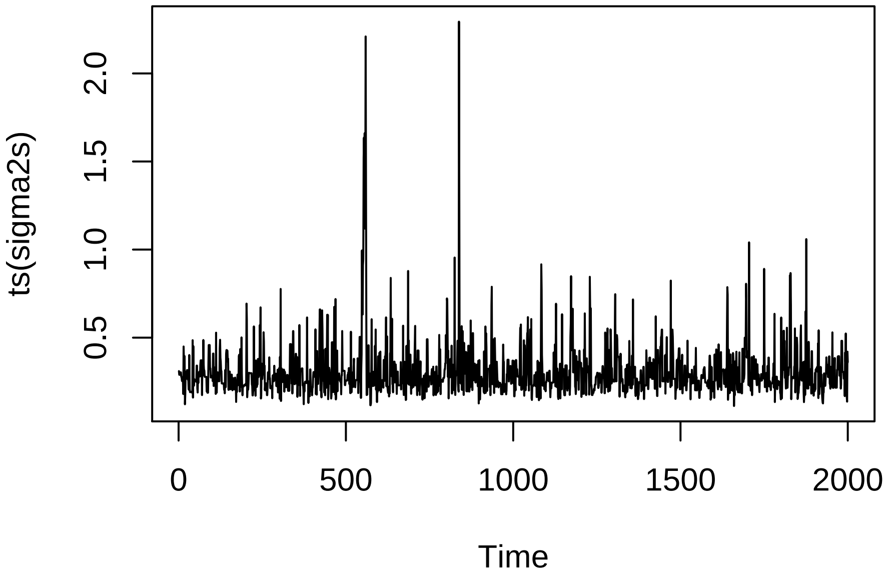
```r
plot(ts(mu2s))
```



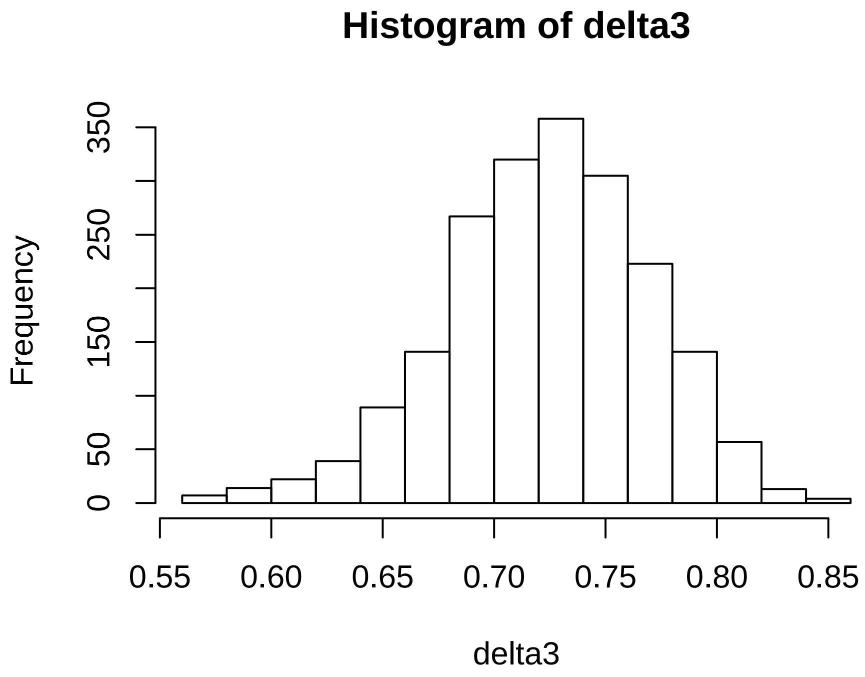
```r
plot(ts(sigma1s))
```



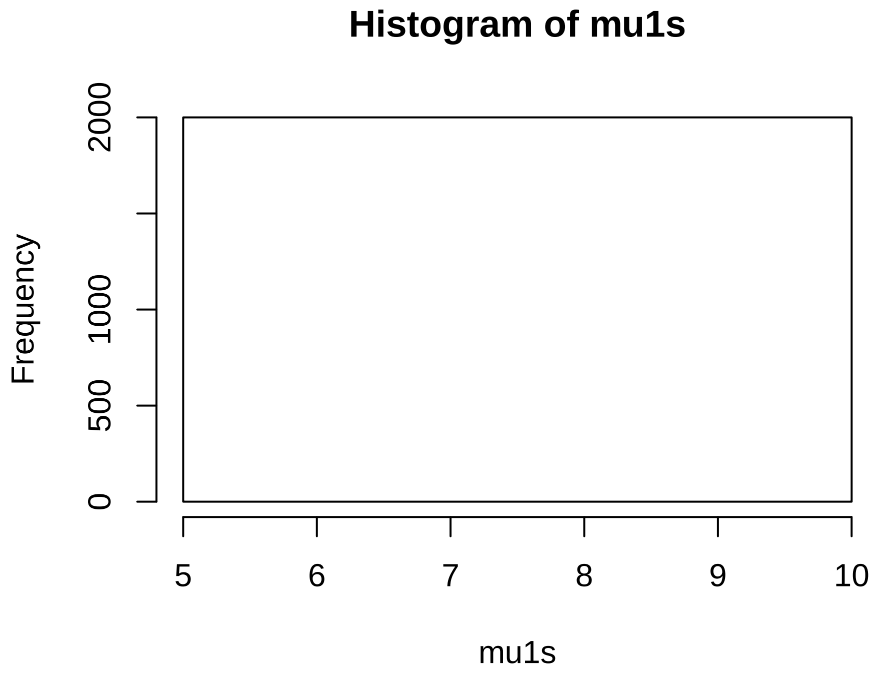
```r
plot(ts(sigma2s))
```



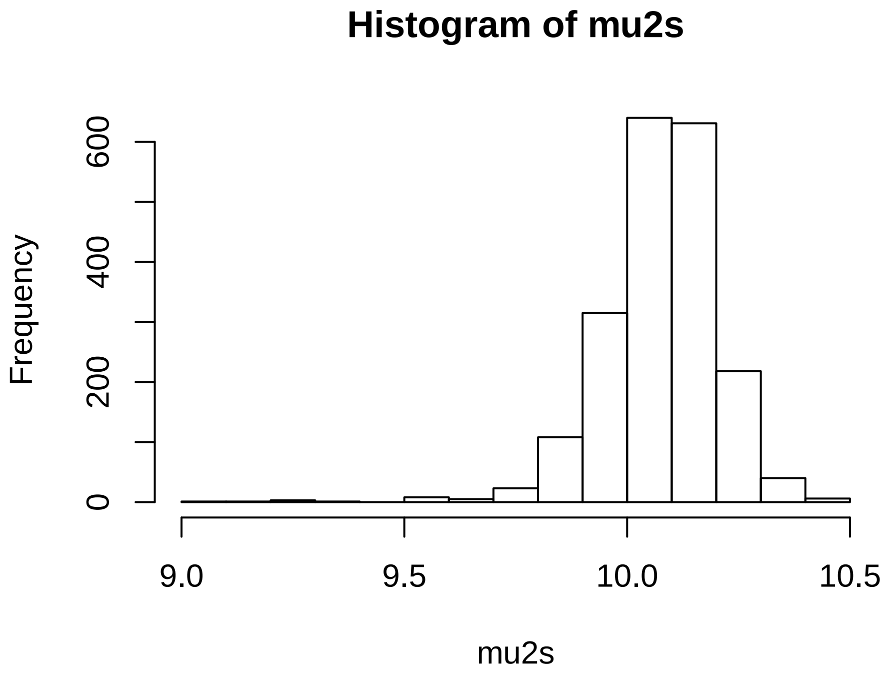
```
hist(delta3)
```



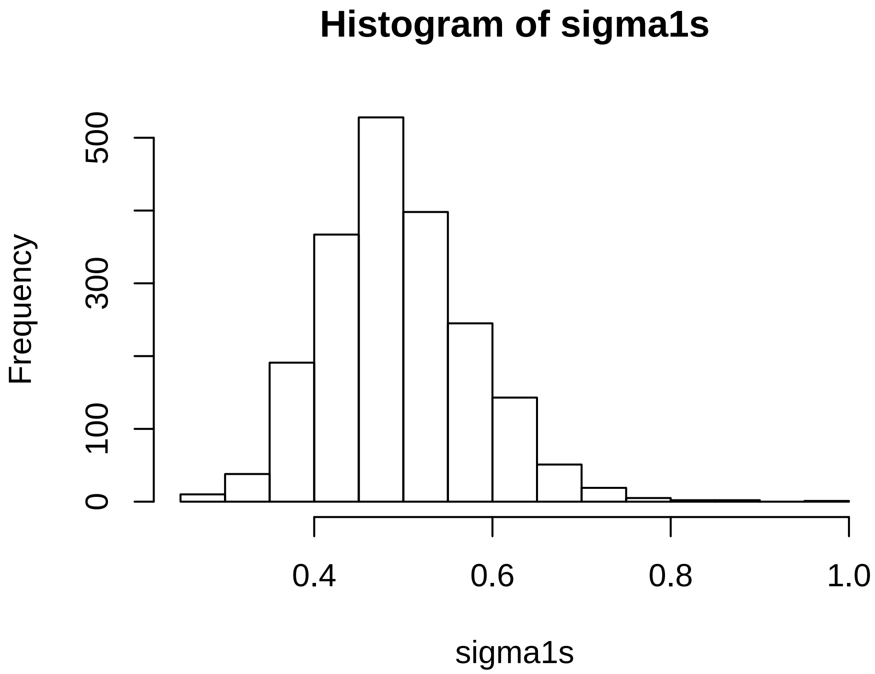
```
hist(mu1s)
```



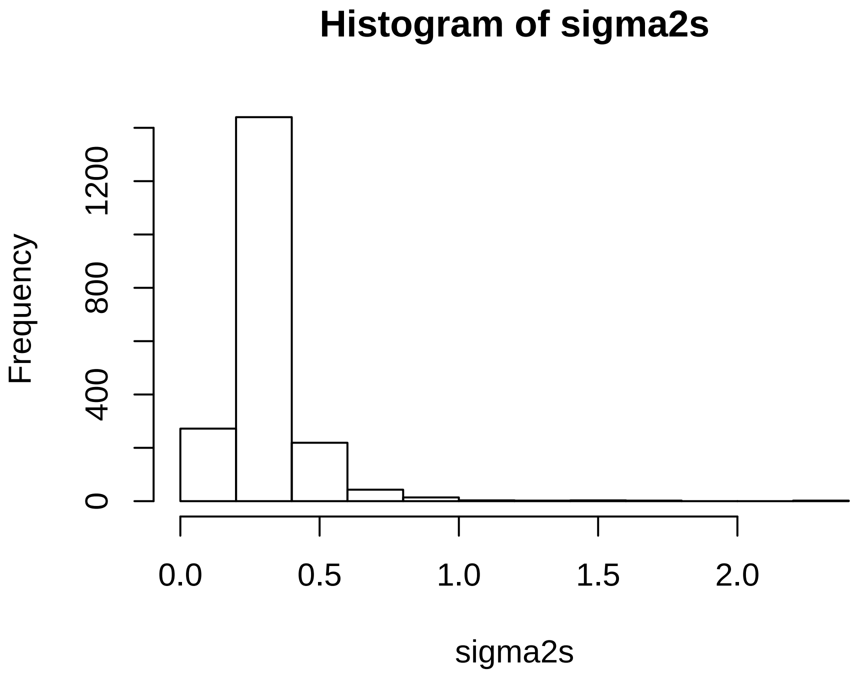
```
hist(mu2s)
```



```
hist(sigma1s)
```



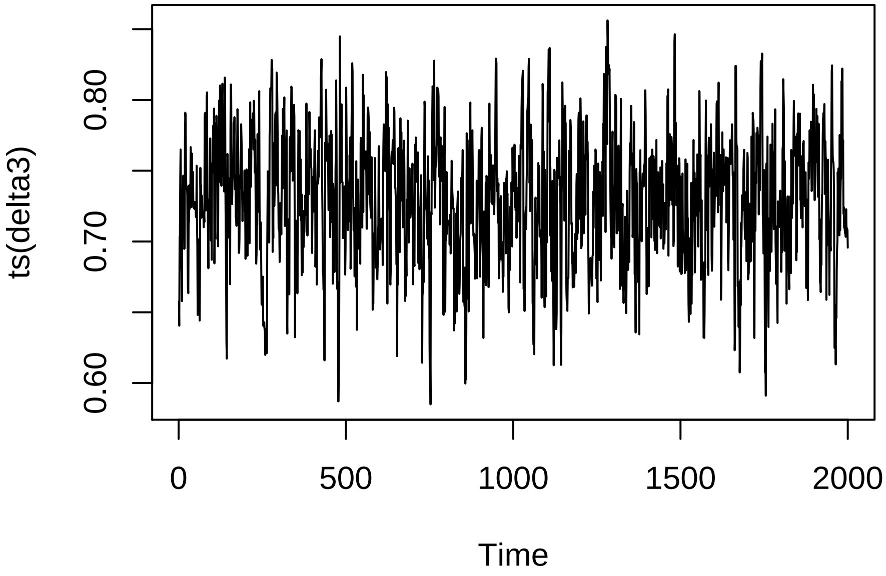
```
hist(sigma2s)
```



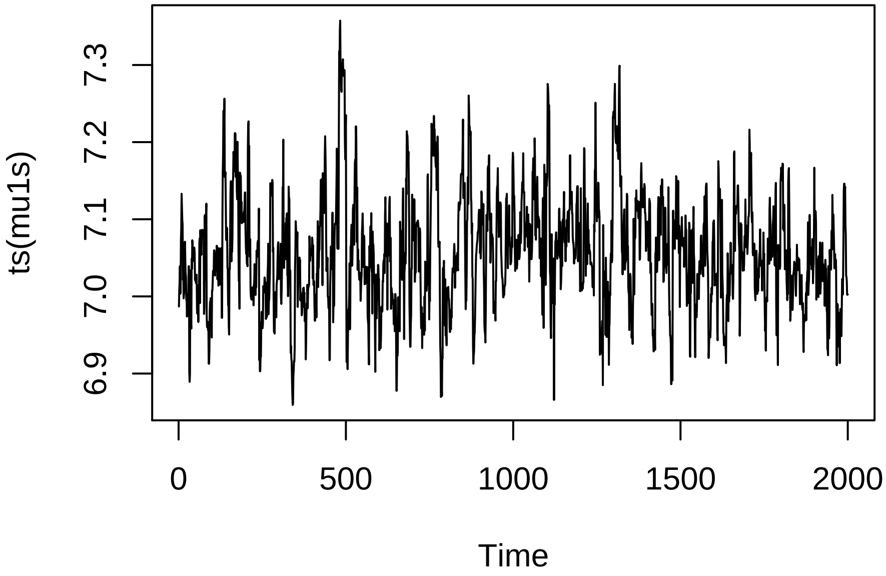
Or we can directly use the function *arms* to sample five parameters simultaneously, since the underlying method used for multivariate sampling in *arms* is Gibbs sampling.

```
f <- function(xa) log(mylike(xa[1],xa[2],xa[3],xa[4],xa[5],x))
init <- c(0.5,5.7,9,0.2,0.2)
res <- arms(init,f,function(xa) (xa[1]>0)*(xa[1]<1)*(xa[2]>-50)*(xa[2]<100)*(xa[3]>-50)*(xa[3]

res <- res[-c(1:1000),]
```

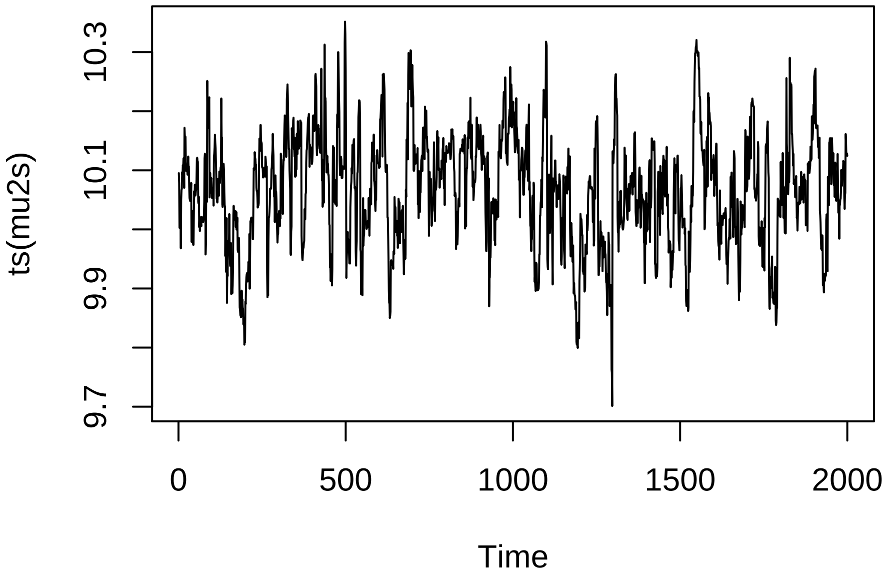
```
delta3 <- res[,1]
mu1s <- res[,2]
mu2s <- res[,3]
sigma1s <- res[,4]
sigma2s <- res[,5]

plot(ts(delta3))
```



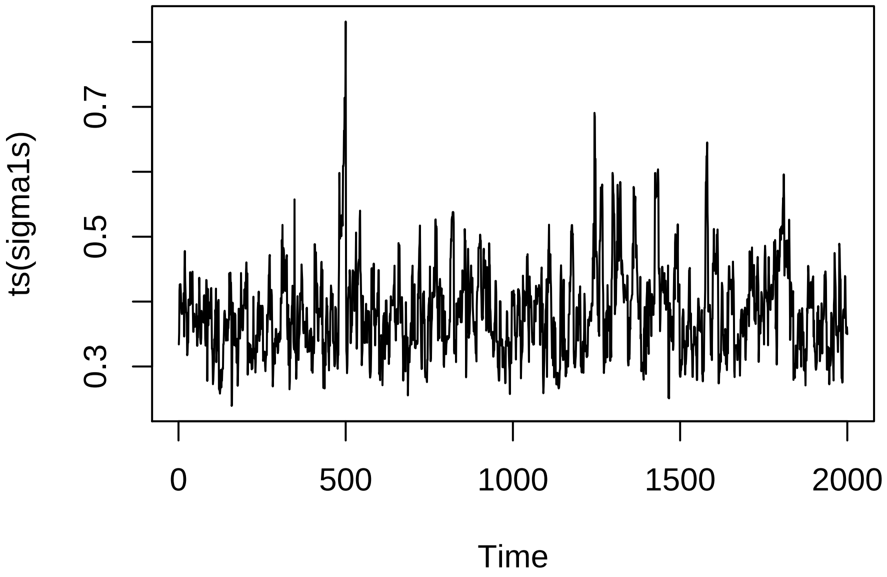
```
plot(ts(mu1s))
```



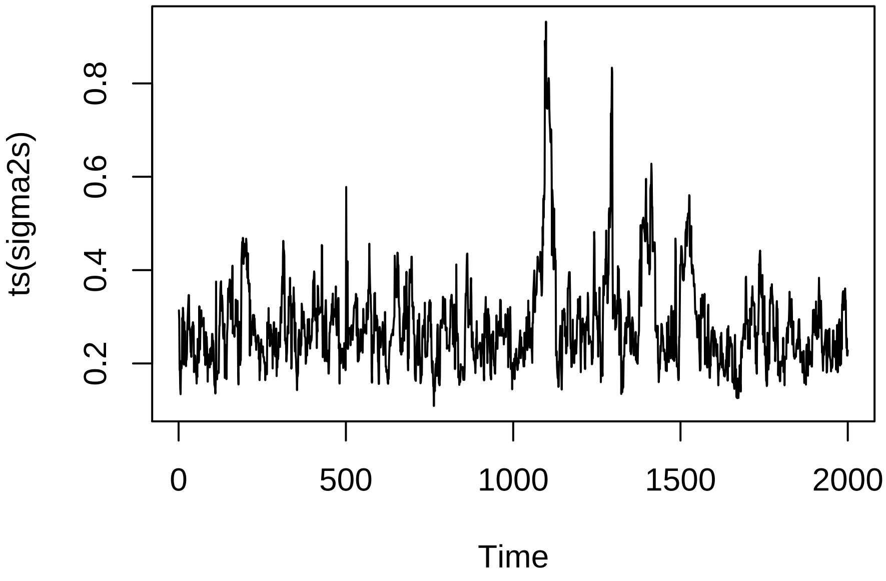
```
plot(ts(mu2s))
```



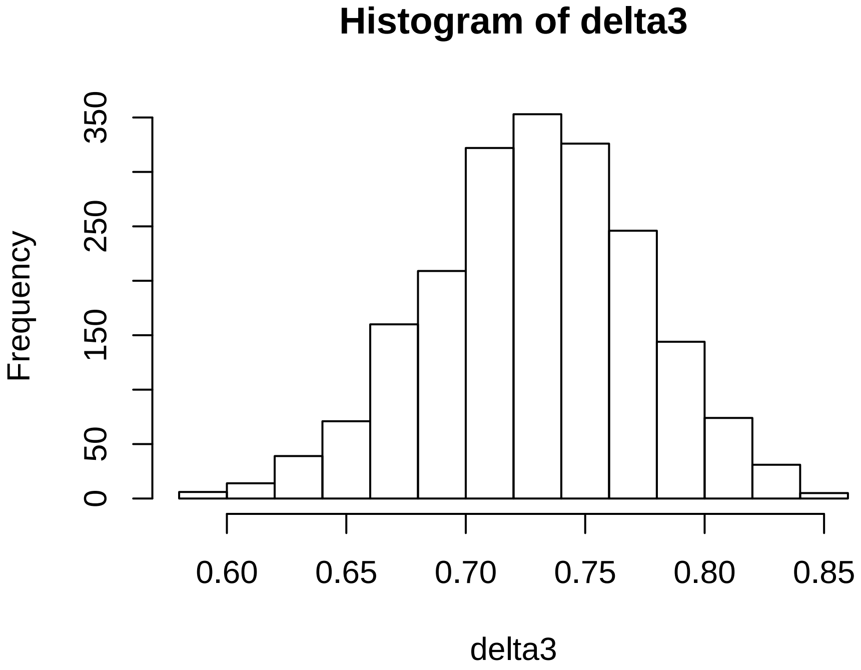
```r
plot(ts(sigma1s))
```



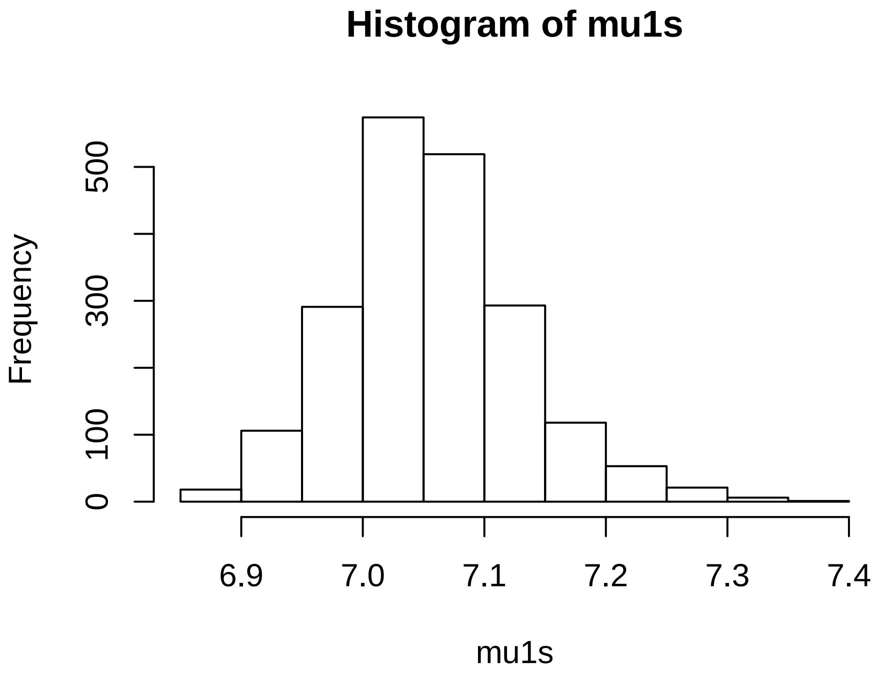
```r
plot(ts(sigma2s))
```



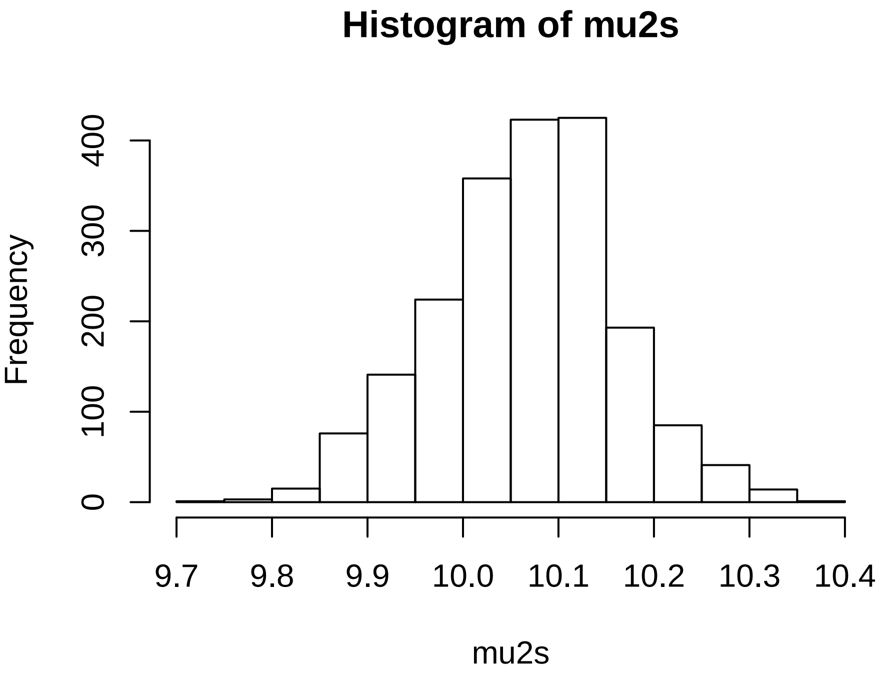
```r
hist(delta3)
```



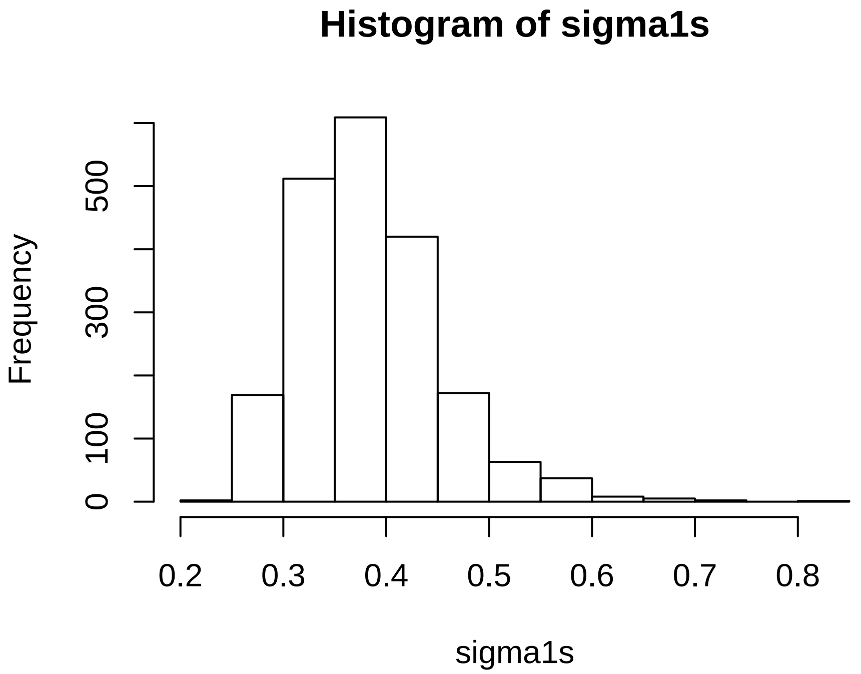
```r
hist(mu1s)
```



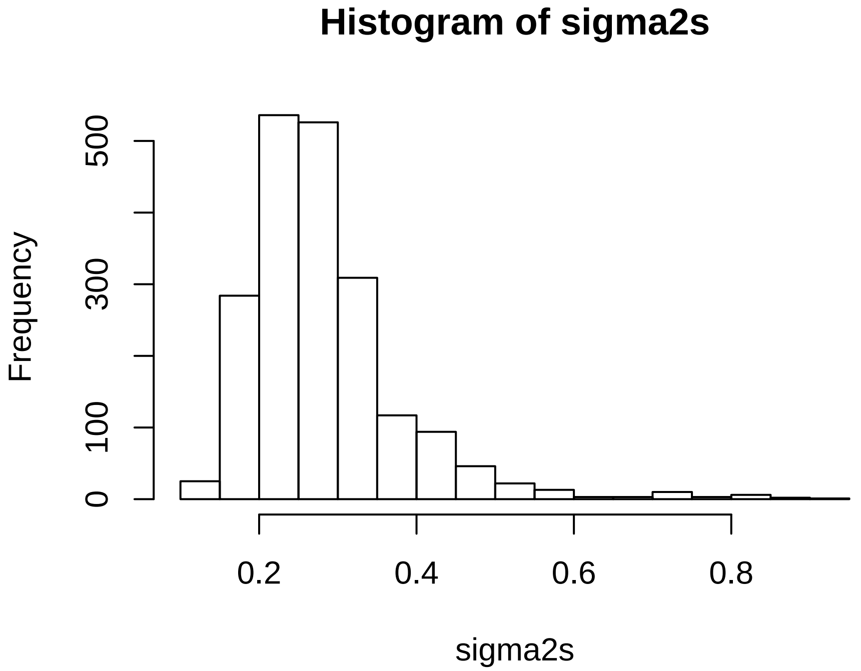
```
hist(mu2s)
```



```
hist(sigma1s)
```



```
hist(sigma2s)
```


Histogram of sigma2s
Frequency
sigma2s