Assignment 4

Mary Silva

3/13/2019

1A

Using $\theta_1 = 1.5$ and $\theta_2 = 2$ we draw a sample of size 1000 using the independence Metropolis Hastings algorithm with gamma distribution as the proposal density.

```
theta_1 = 1.5 # true value theta1
theta_2 = 2 # true value theta2
mean_z1 = sqrt(theta_2/theta_1)
mean_z2 = sqrt(theta_1/theta_2) + 1/(2*theta_2)

# hyperparams
b = 2.5
a = mean_z1*b
#M-H Algorithm
MH_alg1 = function(N){
  MH_samples = rep(NA, N)
  count = 0
  current_z = 1.0
  for(i in 1:N){
    curr_p = pdf_z(current_z)
    z_new = rgamma(1, a, b)
    p_new = pdf_z(z_new)

    accept = exp(p_new + dgamma(current_z,a,b,log = T) -
                   p_new - dgamma(z_new,a,b,log = T))
    if(runif(1) < accept){
      current_z = z_new
      count = count + 1
    }
    MH_samples[i] = current_z
  }
  return(list(MH_samples=MH_samples,count=count))
}
```

After trying several hyperparameters for different Gamma distributions, the best sample obtains a mean, $E(Z)$, of

```
## [1] 1.763996
```

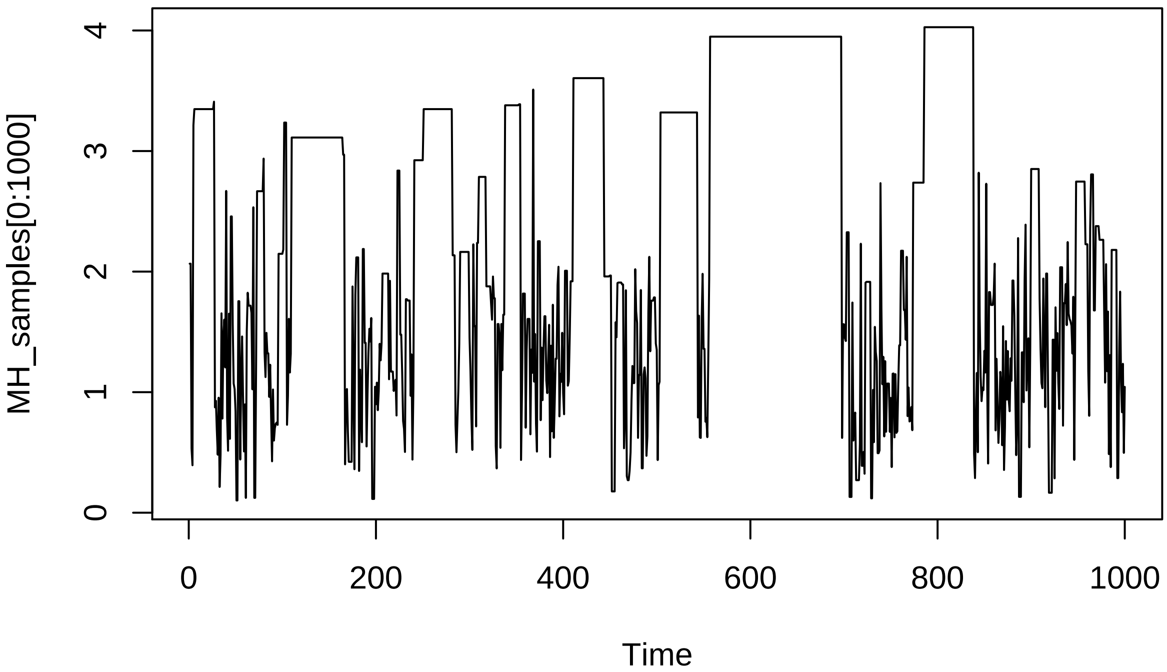
$E(1/Z)$

```
## [1] 1.537128
```

and an accuracy of

```
## [1] 0.513
```

The traceplot for the samples for Metropolis-Hastings is shown below:



1B

The density of $W = log(Z)$ is given by

$$f_W(w) \propto \exp\left\{-\frac{3}{2}w - \theta_1 exp\{w\} - \frac{\theta_2}{\exp(w)}\right\}\exp(w)$$

We draw a sample of size 1000 using the random-walk Metropolis algorithm with this density.

```
v = 0.01
MH_RW = function(N){
  N = N
  MH_RW = rep(NA, N)
  a_count = 0
  z_curr = 1.0
  for (i in 1:N) {
    p_curr = pdf_z2(z_curr)
    z_new = exp(log(z_curr) + rnorm(1,0,sqrt(v)))
    p_new = pdf_z2(z_new)
    acceptance = exp(p_new - p_curr)
    if(runif(1) < acceptance){
      z_curr = z_new
      a_count = a_count+1
    }
    MH_RW[i] = z_curr
  }
  return(list(MH_RW=MH_RW, a_count=a_count))
}
```

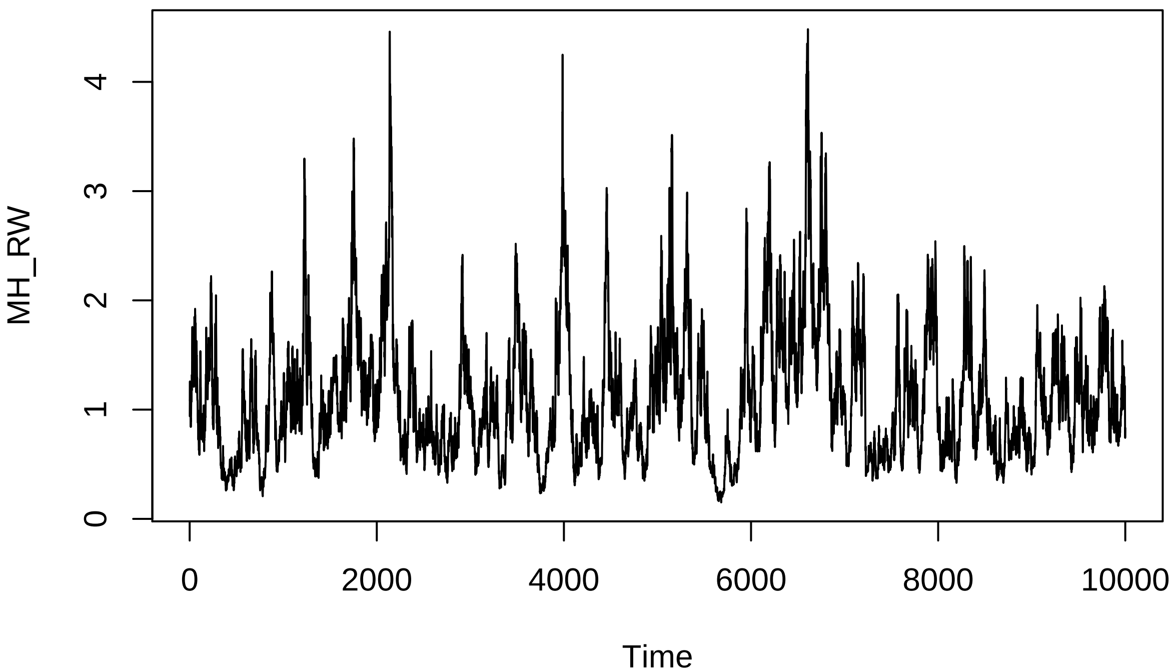
The mean for the samples, $E(W_{samples})$, is

```
## [1] 1.114488
```

And the accuracy is

```
## [1] 0.9363
```

If we use 10000 metropolis hastings random ralk samples, the traceplot is shown below



2A

$$x_i|\nu, \theta \sim Gamma(\nu, \theta)$$
$$\nu \sim Gamma(a, b)$$
$$\theta \sim Gamma(\alpha, \beta)$$

The joint posterior for θ and ν

$$\pi(\theta, \nu, \boldsymbol{x}) \propto \frac{\left(\prod_{i=1}^n x_i\right)^{\nu-1} \nu^{a-1} e^{-b\nu}}{(\Gamma(\nu))^n} \theta^{a+n\nu-1} \exp\left\{-\theta\left(\beta + \sum_{i=1}^n x_i\right)\right\}$$

The full conditionals:

$$\pi(\theta|\nu, \boldsymbol{x}) \propto \theta^{a+n\nu-1} \exp\left\{-\theta\left(\beta + \sum_{i=1}^n x_i\right)\right\}$$

thus, $\theta|\nu, \boldsymbol{x} \sim Gamma(n\nu, \beta + \sum x_i)$.

$$\pi(\nu|\theta, \boldsymbol{x}) \propto \theta^{n\nu} \frac{\left(\prod_{i=1}^n x_i\right)^{\nu-1} \nu^{a-1} e^{-b\nu}}{(\Gamma(\nu))^n}$$

which is not a recognizable distribution. We use a Metropolis within Gibbs algorithm to sample from the full conditionals, using a random walk proposal on $log(\nu)$. I tried various hyperparameters appropriate for this data.

```r
sample = NULL
sample$theta = rep(NA,N)
sample$nu = rep(NA,N)
alpha = 3
beta = 2
v = 0.05
theta_curr = 2
nu_curr = 3
set.seed(2)
for(i in 1:N){
  theta_curr = rgamma(1, n*nu_curr + alpha, beta + sum_x)
  nu_new = exp(log(nu_curr) + rnorm(1,0,sqrt(v)))
  pnu_curr = nu_condit(nu_curr, theta_curr)
  pnu_new = nu_condit(nu_new, theta_curr)
  accept = exp(pnu_new - pnu_curr)
  if(runif(1) < accept)
    nu_curr = nu_new

  sample$theta[i] = theta_curr
  sample$nu[i] = nu_new
}
```

The effective sample size for θ is

```
##     var1
## 37.15345
```

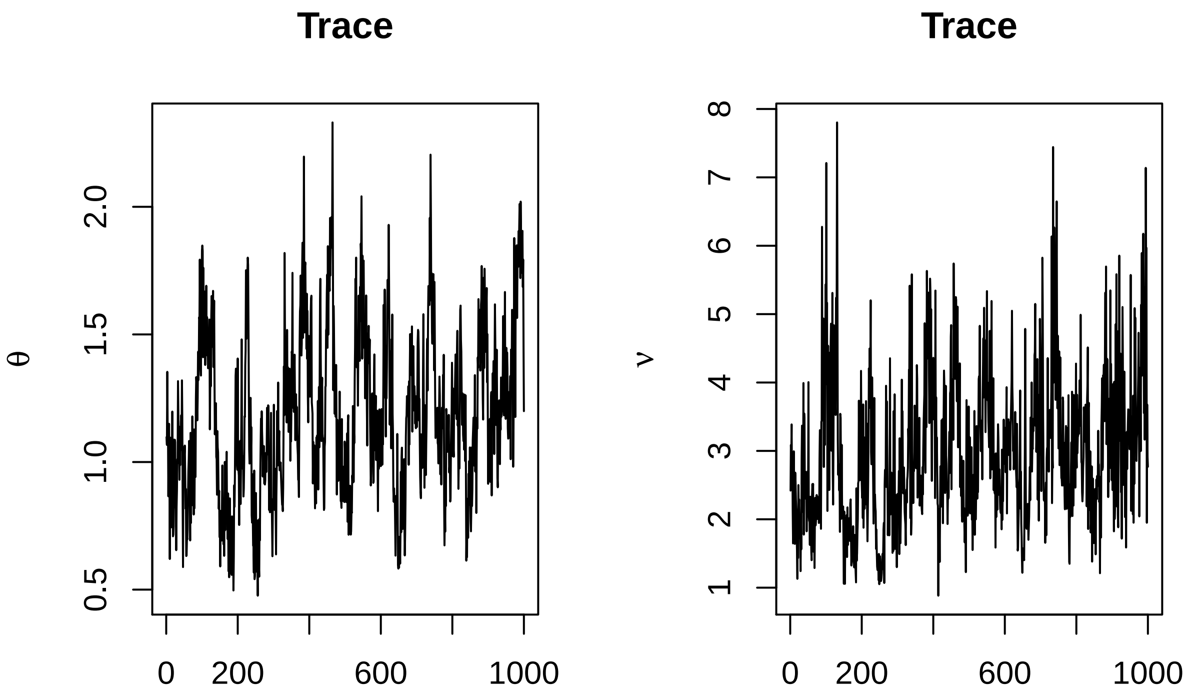
The effective sample size for ν is

```
##     var1
## 62.45584
```

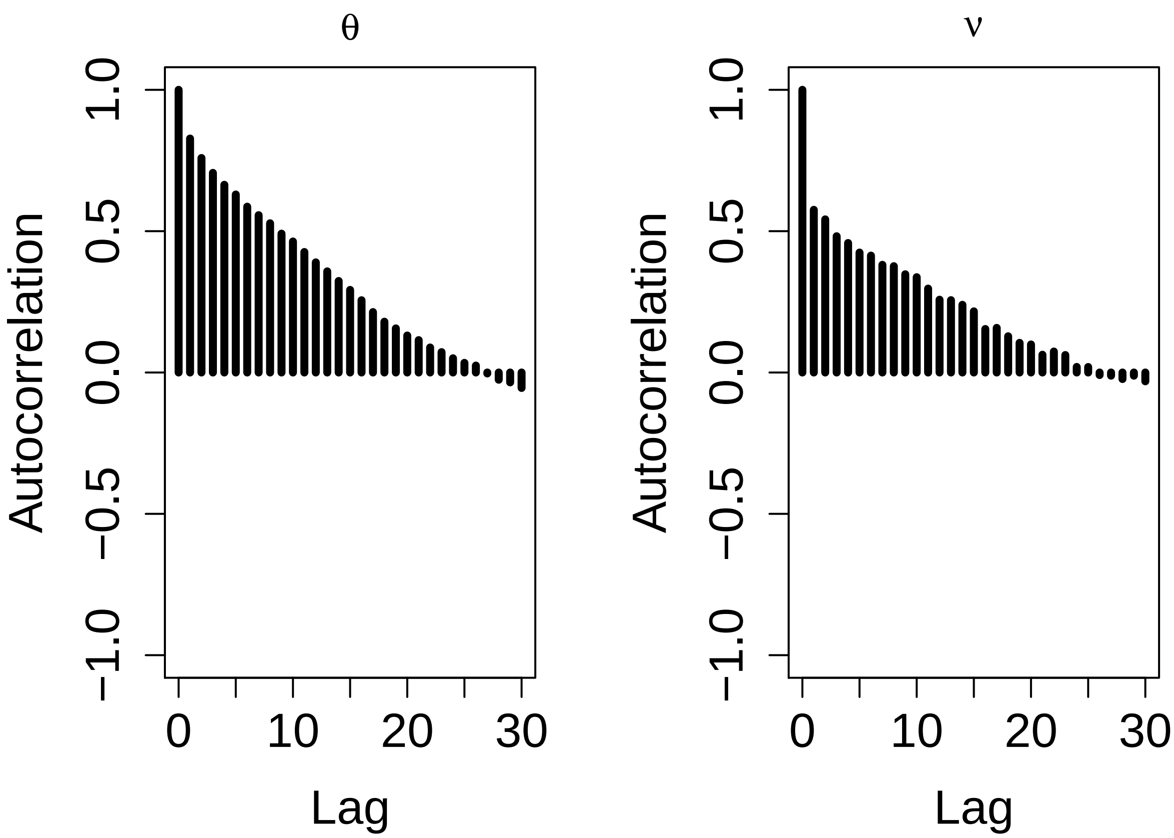
The table below summarizes the results

parameter	mean	95% Credible Interval
θ	1.13	(0.635,1.846)
ν	2.807	(1.669, 4.55)

The traceplots are below



The autocorrelation plots are below



2B

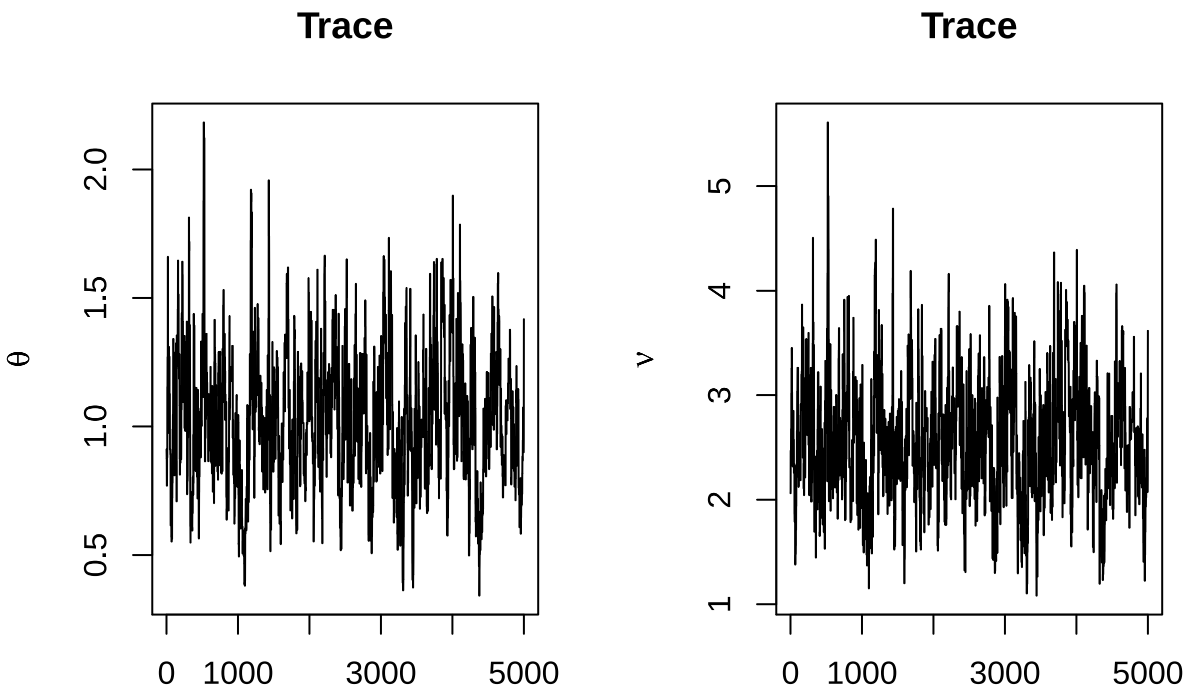
Now we develop a Metropolis-Hastings algorithm that jointly proposes $\log(\nu)$ and $\log(\theta)$ using a Gaussian random walk centered on the current value of the parameters. Tune the variance-covariance matrix of the proposal using a test run that proposes the parameters independently:

```
V = 0.05*diag(2)
theta_curr = 2
nu_curr = 3

for(i in 1:N_test){
  nu_new = exp(log(nu_curr) + rnorm(1,0,sqrt(V[1,1])))
  theta_new = exp(log(theta_curr) + rnorm(1,0, V[2,2]))
  p_curr = pcurr(nu_curr, theta_curr)
  p_new = pcurr(nu_new, theta_new)
  accept = exp(p_new - p_curr)
  if(runif(1) < accept){
    nu_curr = nu_new
    theta_curr = theta_new
  }
  sample$theta[i] = theta_curr
  sample$nu[i] = nu_curr
}
```

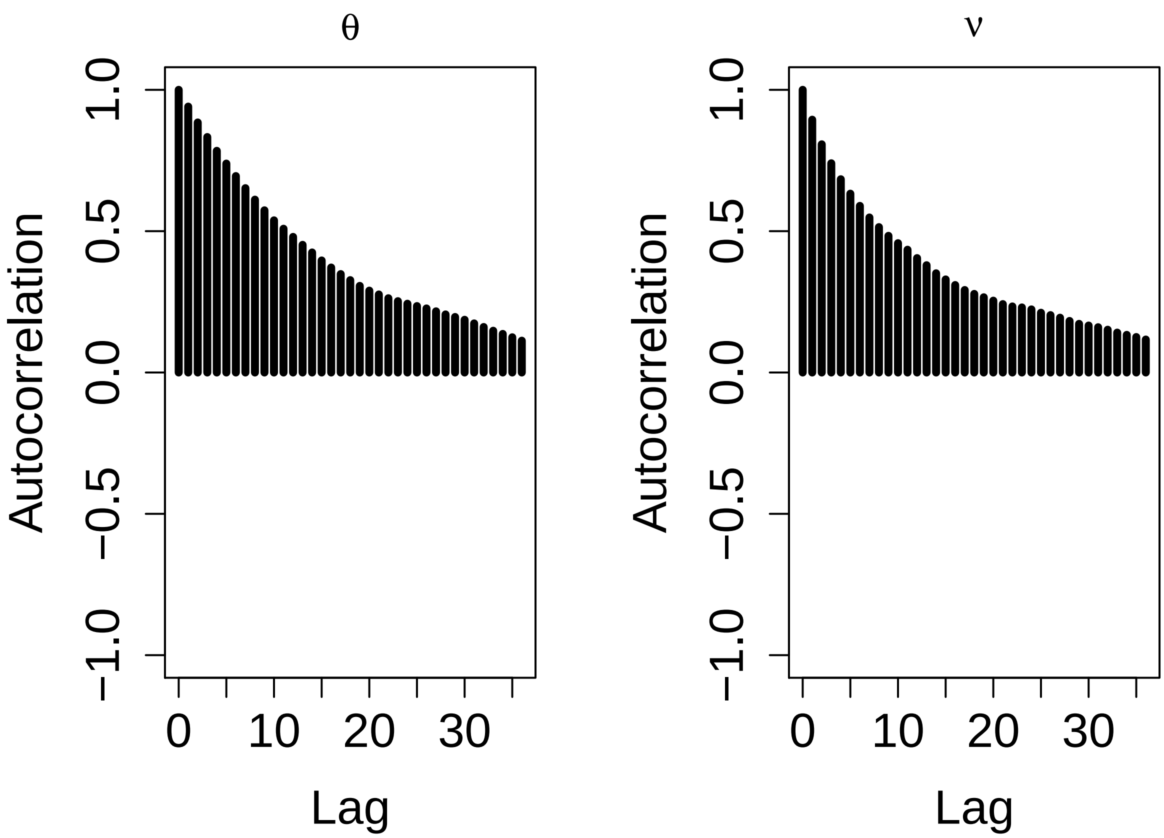
```
for(i in N_test+1:N){
  new = mvrnorm(1, c(log(nu_curr), log(theta_curr)), V)
  nu_new = exp(new[1])
  theta_new = exp(new[2])
  p_curr = pcurr2(nu_curr, theta_curr)
  p_new = pcurr2(nu_new, theta_new)
  acceptance = exp(p_new - p_curr)
  if(runif(1) < acceptance){
    nu_curr = nu_new
    theta_curr = theta_new
  }
  sample$theta[i] = theta_curr
  sample$nu[i] = nu_curr
}
```

parameter	mean	95% Credible Interval
θ	1.11	(0.59,1.79)
ν	2.807	(1.49, 4.38)

The trace plot for these samples are below



The corresponding autocorrelation plot is below



2C

Now we are going to develop a Metropolis algorithm that jointly proposes $\log \nu$ and $\log \theta$ using independent proposals based on Laplace approximation of the posterior distribution of $\log \nu$ and $\log \theta$.

We let $t = \log \theta$ and $v = \log \nu$, then the posterior becomes

$$\begin{aligned}
\pi(\theta, \nu | \boldsymbol{x}) &\propto \exp\left\{(\nu - 1)\sum_{i=1}^{n} \log x_i + (a-1)\log \nu - b\nu - n \log \Gamma(\nu)\right\} \\
&\times \exp\left\{(\alpha + n\nu - 1)\log \theta - \theta\left(\beta + \sum_{i=1}^{n} x_i\right)\right\} \\
\Rightarrow \pi(t, v|\boldsymbol{x}) &\propto \exp\left\{(e^v - 1)sum_{i=1}^n \log x_i + av - be^v - n\log\Gamma(e^v)\right\} \\
&\times \exp\left\{(a + ne^v)t - e^t\left(\beta + sum_{i=1}^n x_i\right)\right\}
\end{aligned}$$

Now, we let

$$h(t, v) = (e^v = 1)\sum_{i=1}^{n} x_i + av - be^v - n\log\Gamma(e^v)\exp\left\{(a + ne^v)t - e^t\left(\beta + sum_{i=1}^n x_i\right)\right\}$$

Then we use the definition of Laplace approximation

```
h = function(w) {
  a1 = (exp(w[2]) - 1) * sum_logx + 3 * w[2] - exp(w[2])
  return(-(a1 - n * lgamma(exp(w[2])) +
             (2 + n * exp(w[2])) * w[1] - exp(w[1]) * (2 + sum_x)))
}
laplace = optim(c(0,1), h, hessian = T)
```

The laplace maximum for the parameters are

```
## [1] 0.09685737 1.00165219
```

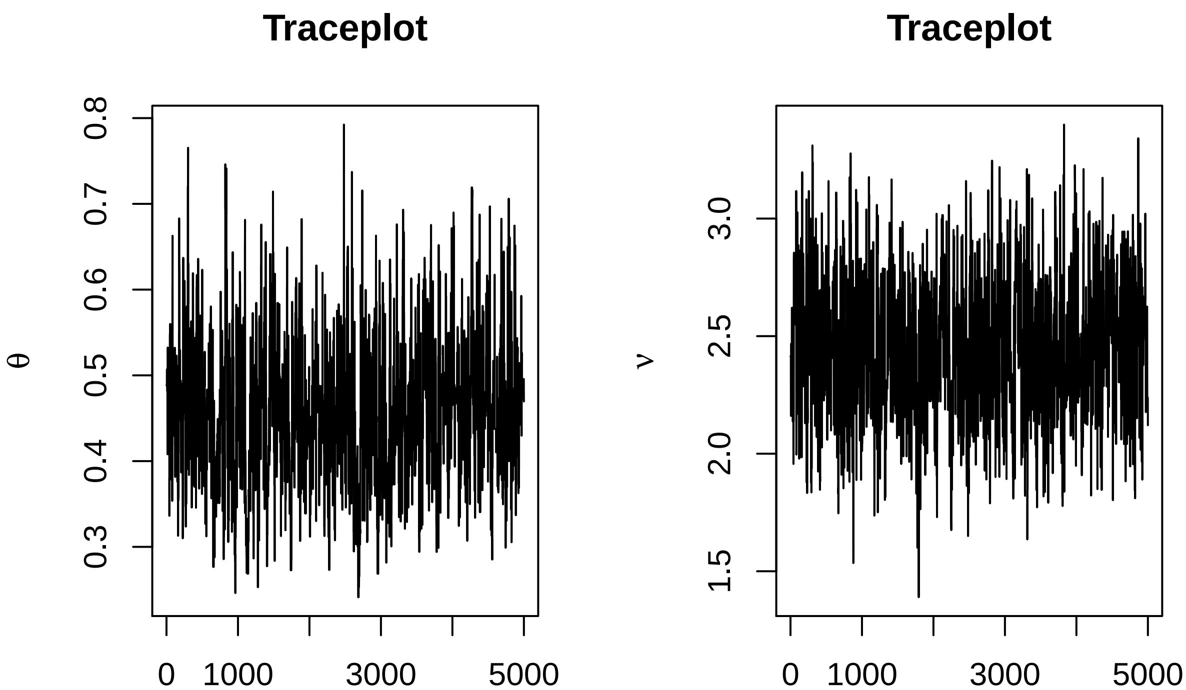
and the hessian obtained at the maximum is

```
##           [,1]      [,2]
## [1,]  70.06943 -68.06943
## [2,] -68.06943  85.06425
```

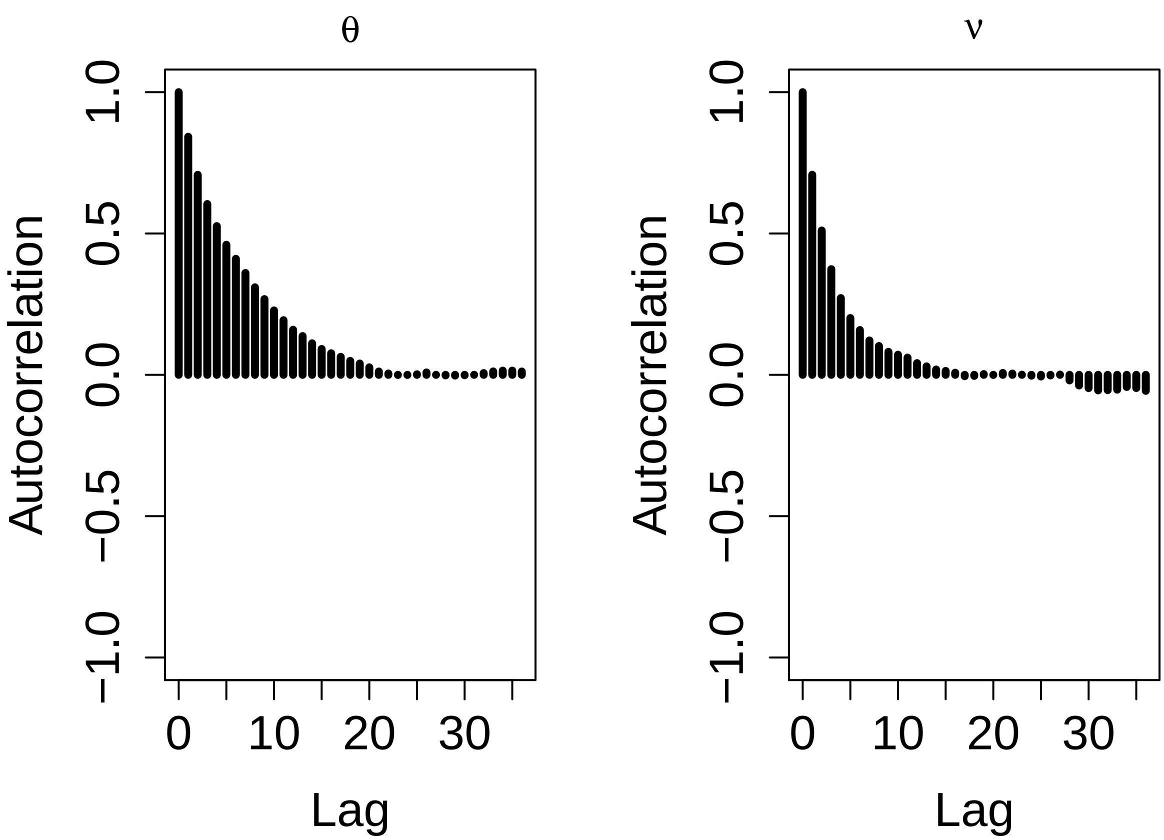
Now we update the variance-covariance matrix then resume the Metropolis sampling algorithm

```
for(i in N_test+1:N){
  nu_new = exp(log(nu_curr) + rnorm(1,0,sqrt(V[1,1])))
  theta_new = exp(log(theta_curr) + rnorm(1,0,sqrt(V[2,2])))
  p_curr = pcurr(nu_curr = nu_curr, theta_curr = theta_curr)
  p_new =pcurr(nu_curr = nu_new, theta_curr = theta_new)

  accept = exp(p_new - p_curr)
  if(runif(1) < accept){
    nu_curr = nu_new
    theta_curr = theta_new
  }
  sample$theta[i] = theta_curr
  sample$nu[i] = nu_curr
}
```

The corresponding traceplots are below



The corresponding autocorrelation plots are below



The effective sample size associated with θ is

```
##      var1
```

```
## 347.7253
```

The effective sample size associated with ν is

```
##     var1
## 855.5197
```

parameter	mean	95% Credible Interval
θ	1.05	(0.56,1.58)
ν	2.55	(1.49, 3.79)

3

Given the random effects model we have $(y_{ij}-(\beta+u_i)) \sim N(0,\tau^2)$, $u_i \sim N(0,\tau^2)$, and $\pi(\beta,\sigma^2,\tau^2) \propto (\sigma^2\tau^2)^{-1}$. Then the joint posterior is

$$\pi(u_i,\beta,\tau^2,\sigma^2|y) \propto (\tau^2)^{-\left(\frac{IJ}{2}+1\right)}(\sigma^2)^{-\left(\frac{I}{2}+1\right)} \exp\left\{-\frac{1}{2\tau^2}\sum_{ij}(y_{ij}-(\beta+u_i))^2 - \frac{1}{2\tau^2}\sum u_i^2\right\}$$

3A

i)

$$\begin{aligned}
\pi(u_i|y,\beta,\tau,\sigma^2) &\propto \exp\left\{-\frac{1}{2\tau^2}\sum\left[y_{ij}^2 - 2y_{ij}(\beta+u_i) + (\beta+u_i)^2\right] - \frac{1}{2\sigma^2}\sum u_i^2\right\} \\
&= \exp\left\{-\frac{1}{2\tau^2}\left[\sum(-2y_{ij}u_i) + \sum(2\beta u_i + u_i^2)\right] - \frac{1}{2\sigma^2}\sum u_i^2\right\} \\
&= \exp\left\{-\frac{1}{2\tau^2}\left[\sum u_i^2 - 2\sum u_i(y_{ij}-\beta)\right] - \frac{1}{2\sigma^2}\sum u_i^2\right\} \\
&= \exp\left\{-\frac{1}{2\tau^2}\left[\sum Ju_i^2 - 2\sum u_i(y_{ij}-\beta)\right] - \frac{1}{2\sigma^2}u_i^2\right\} \\
&= \exp\left\{-\frac{1}{2\tau^2\sigma^2}\left[J\sigma^2\sum u_i^2 - 2\sigma^2\sum u_i(y_{ij}-\beta) + \tau^2\sum u_i^2\right]\right\} \\
&= \exp\left\{-\frac{1}{2\tau^2\sigma^2}\left(u_i^2\left(J\sigma^2+\tau^2\right) - 2u_i\sum(y_{ij}-\beta)\right)\right\} \\
&= \exp\left\{-\frac{J\sigma^2+\tau^2}{2\sigma^2\tau^2}\left(u_i^2 - 2u_i\frac{\sum(y_{ij}-\beta)}{J\sigma^2+\tau^2}\right)\right\}
\end{aligned}$$

Therefore,

$$u_i|\cdot \sim N\left(\frac{\sum_j(y_{ij}-\beta)}{J\sigma^2+\tau^2}, \frac{\tau^2\sigma^2}{J\sigma^2+\tau^2}\right) = N\left(\left(\frac{J}{\tau^2}+\frac{1}{\sigma^2}\right)^{-1}\left(\frac{\sum_j(y_{ij}-\beta)}{\tau^2}\right), \left(\frac{J}{\tau^2}+\frac{1}{\sigma^2}\right)^{-1}\right)$$

I am lazy, so I am just going to skip to the end results so I don't have to type all my work :(

ii)

$$\beta|\cdot \sim N\left(\frac{\tau^2}{IJ}, \frac{\sum_{ij}(y_{ij}-u_i)}{IJ}\right) = N\left(\left(\frac{IJ}{\tau^2}\right)^{-1}\left(\frac{\sum_{ij}(y_{ij}-u_i)}{\tau^2}\right), \left(\frac{IJ}{\tau^2}\right)^{-1}\right)$$

iii)

$$\sigma^2|\cdot \sim IG\left(\frac{I}{2}, \frac{1}{2}\sum_i u_i^2\right)$$

iv)

$$\tau^2|\cdot \sim IG\left(\frac{IJ}{2}, \frac{1}{2}\sum_{ij}(y_{ij} - (\beta + u_i))^2\right)$$

3B

$$\pi(\beta, \tau^2, \sigma^2|y) \propto (\tau^2)^{-\left(\frac{I(J-1)}{2}+1\right)}(\sigma^2)^{-1}(J\sigma^2 + \tau^2)^{I/2}\exp\left\{-\frac{1}{2\tau^2}\sum_{ij}(y_{ij} - \beta)^2\right\}$$

$$\times \exp\left\{\frac{\sigma^2}{2\tau^2(J\sigma^2 + \tau^2)}\sum_i\left(\sum_j(y_{ij} - \beta)\right)^2\right\}$$

3C

$$\pi(\tau^2, \sigma^2|y) \propto (\tau^2)^{-\left(\frac{I(J-1)}{2}+1\right)}(\sigma^2)^{-1}\left(J\sigma^2 + \tau^2\right)^{\frac{I+1}{2}}\exp\left\{-\frac{1}{2\tau^2}\sum_{ij}y_{ij}^2\right\}$$

$$\times \exp\left\{\frac{\sigma^2}{2\tau^2(J\sigma^2 + \tau^2)}\sum_i\left(\sum_j y_{ij}^2\right)\right\}$$

$$\times \exp\left\{\frac{1}{2IJ(J\sigma^2 + \tau^2)}\left(\sum_{ij}y_{ij}\right)^2\right\}$$

4

The joint posterior is obtained by

$$\begin{aligned}\pi(\theta, \phi, m|\boldsymbol{y}) \propto& f(\boldsymbol{y}|\theta, \phi, m)\pi(\theta)\pi(\phi)\pi(m) \\ \propto& \theta^{\sum_{i=1}^m y_i + \alpha - 1}\exp\left\{-\theta\left(\beta + m\right)\right\}\phi^{\sum_{i=m+1}^n y_i + \gamma - 1}\exp\left\{-\phi(\delta + n - m)\right\}\end{aligned}$$

Then the full conditionals are as follows

$$\phi|m, \boldsymbol{y} \sim Gamma\left(\sum_{i=m+1}^n y_i + \gamma - 1, \delta + n - m\right)$$

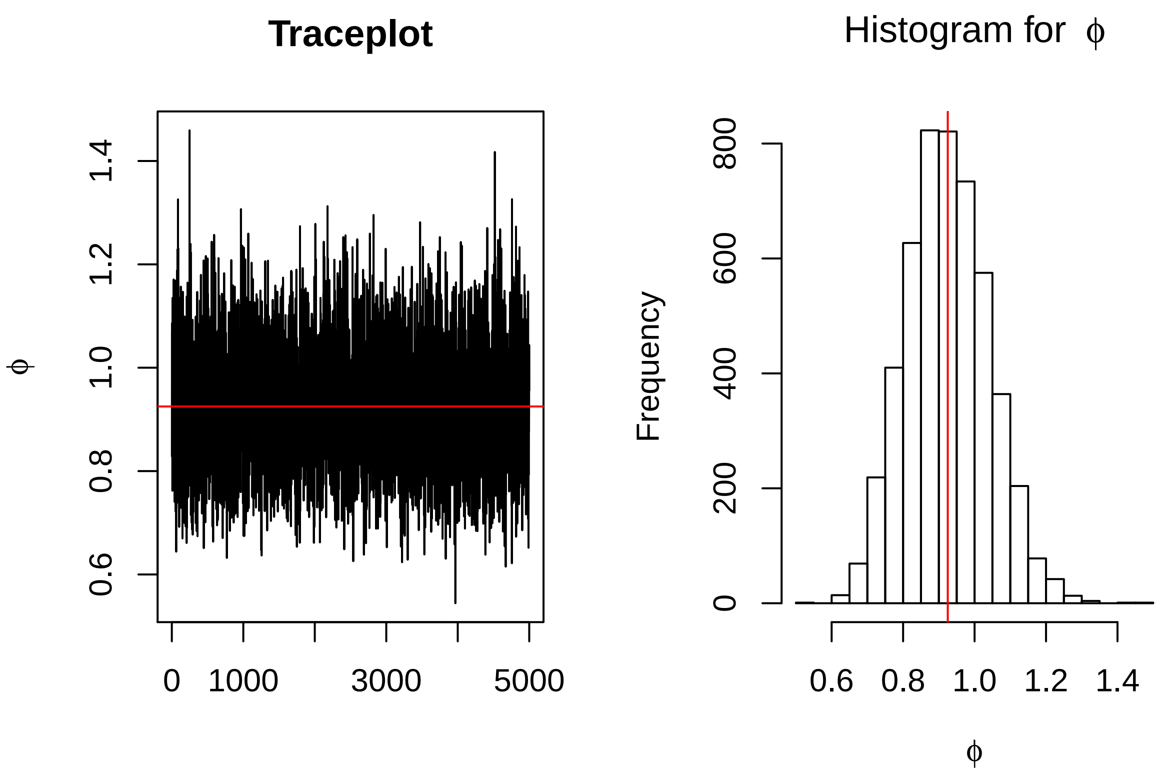
$$\theta|m, \boldsymbol{y} \sim Gamma\left(\sum_{i=1}^m y_i + \alpha - 1, \beta + m\right)$$

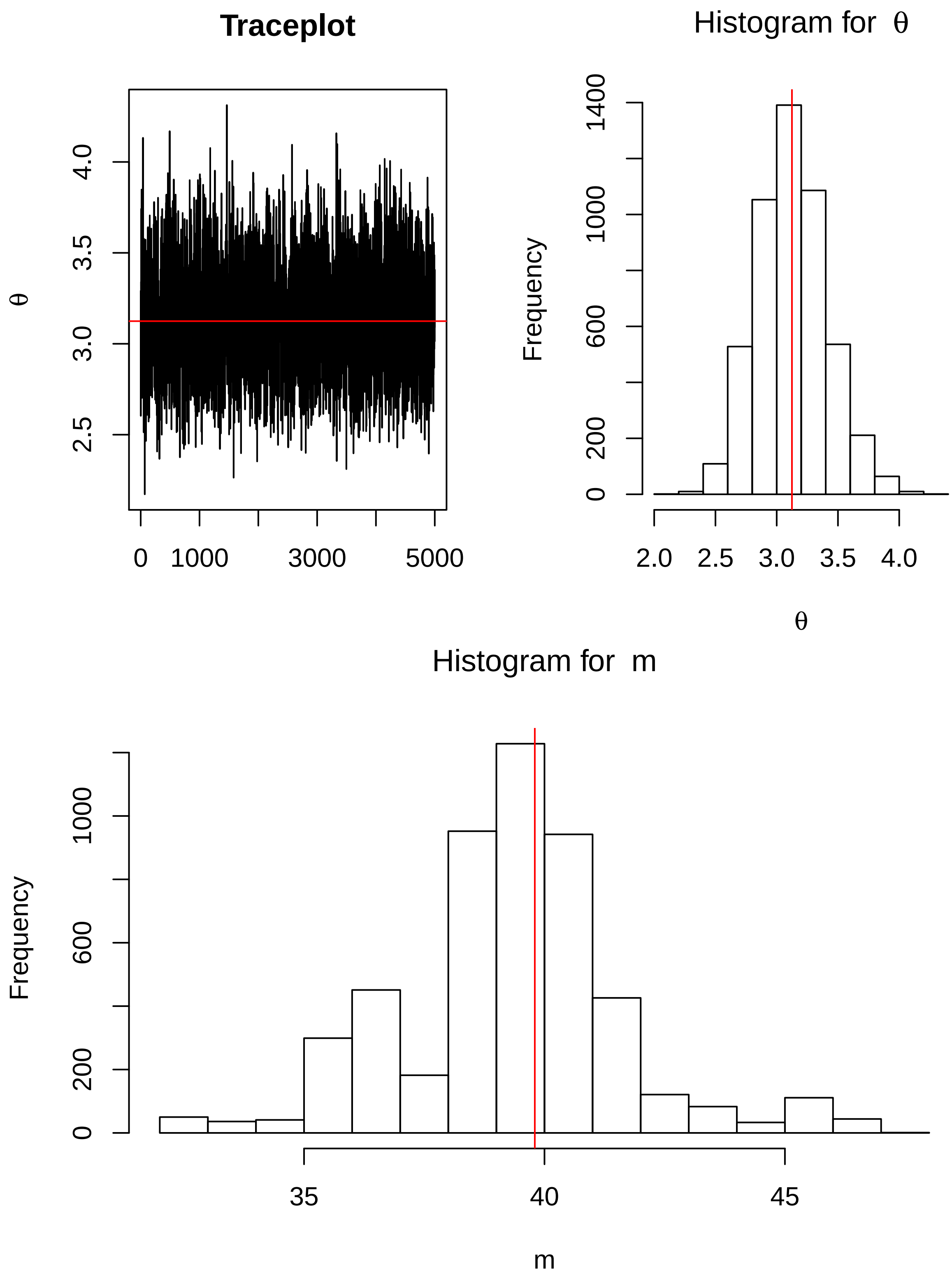
$$\pi(m|\theta, \phi, \boldsymbol{y}) \propto \theta^{\sum_{i=1}^m y_i + \alpha - 1}\exp\left\{-\theta(\beta + m)\right\}\phi^{\sum_{i=m+1}^n y_i + \gamma - 1}\exp\left\{-\phi(\delta + n - m)\right\}$$

We now use a Metropolis-within-Gibbs method because it converges better.

```
# Algorithm
for(i in 1:N){
  theta_curr <- rgamma(1, sum(y[1:m_curr]) + alpha, m_curr + beta)

  phi_curr <- rgamma(1, sum(y[-(1:m_curr)]) + gam, (n-m_curr + delta))

  m_new <- sample((1:n), 1, FALSE)

  p_curr <- lgamma(sum(y[1:m_curr]) + alpha) - (sum(y[1:m_curr]) + alpha)*log(m_curr + beta) +lgamma(su

  p_new <- lgamma(sum(y[1:m_new]) + alpha) - (sum(y[1:m_new]) + alpha)*log(m_new + beta) +lgamma(sum(y[
  # calculate acceptance probability and accept/reject acordingly
  accpt.prob <- exp(p_new - p_curr)
  if(runif(1) < accpt.prob)
  {
    m_curr <- m_new
  }
  # save the current draws
  sample_save$theta[i] <- theta_curr
  sample_save$phi[i] <- phi_curr
  sample_save$m[i] <- m_curr
}
```

Using $N = 50000$ samples and a burnin of 5000, the results are plotted below with the red line signifying the mean for the parameters




Traceplot
θ
2.5
3.0
3.5
4.0
0
1000
3000
5000
Histogram for θ
Frequency
0
200
600
1000
1400
2.0
2.5
3.0
3.5
4.0
θ
Histogram for m
Frequency
0
200
600
1000
35
40
45
m