Monte Carlo Methods Midterm

Kylie Taylor

3/28/2019

Suppose observations y_{ij} for i =1,. . . ,n and for j = 1,. . . ,m arise given (v_i) accroding to

$log(y_{ij}) = \mu + log(\frac{v_i}{v_i+x_{ij}}) + \sigma\epsilon ij$

where (ϵij) are iid standard normal

$\epsilon ij \sim iidN(0,1) \propto e^{-\frac{1}{2}x^2}$

The $log(v_i)$ are random parameters and i.i.d. from a normal distribution with mean ν and variance ϕ^2

$f(log(v_i)) \sim iidN(\nu,\phi^2) \propto \frac{1}{\sqrt{2\pi\phi^2}}e^{-\frac{(logv_i-\nu)^2}{2\phi^2}}$

The parameters of the model are $\theta = (\mu,\sigma,\nu,\phi)$.

The prior for μ is normal with mean 0 and variance ξ^2.

$f(\mu) \sim N(0,\xi^2) = \frac{1}{\sqrt{2\pi\xi^2}}e^{-\frac{\mu^2}{2\xi^2}}$

The prior to ν is normal with mean 0 and variance ψ^2.

$f(\nu) \sim N(0,\psi^2) = \frac{1}{\sqrt{2\pi\psi^2}}e^{-\frac{\nu^2}{2\psi^2}}$

The prior for $\eta = \frac{1}{\phi^2}$ is Gamma with parameters (c,d).

$f(\eta) = \frac{1}{\phi^2} \sim Ga(c,d) \propto \eta^{c-1}e^{-d\eta}$

The prior for $\lambda = \frac{1}{\sigma^2}$ is Gamma with parameters (a,b).

$f(\lambda) = \frac{1}{\sigma^2} \sim Ga(a,b) \propto \lambda^{a-1}e^{-b\lambda}$

The data are for n = 86 and m = 7 for each i, it is that $x_{ij} = j$.

The prior settings, i.e. (a, b, ψ, ξ, c, d)

For this midterm, I will do the following steps:

1) Download and use the data, where I am given x_{ij}, i = 1,. . . ,86 and j = 1,. . . ,7
2) Pick a fixed (μ,σ,ν,ϕ) and generate $v_i = (v_1), ..., (v_{86})$ from $logN(\nu,\phi^2)$
3) Generate $log(y_{ij})$ from $N(\mu + log(\frac{v_i}{v_i+x_{ij}}),\sigma^2)$
4) Code the Markov Chain of the conditional densities $f(\mu|...), f(\nu|...), f(\eta|...), f(\lambda|...)$ and $f(v_i|...)$
5) Report the output of the chain.
6) Find an estimate for v_{87}

1) Data

The first step is to make sure the data is downloaded correctly. Since I do not know any information about the seven variables in this dataset, I cannot say that the summary statistics reveal anything alarming.

Table 1: Table continues below

V1	V2	V3	V4
Min. : 1.00	Min. :0.0370	Min. :0.03500	Min. :0.0180
1st Qu.:22.25	1st Qu.:0.1710	1st Qu.:0.09725	1st Qu.:0.0745
Median :43.50	Median :0.2580	Median :0.16550	Median :0.1110
Mean :43.50	Mean :0.3115	Mean :0.19230	Mean :0.1455
3rd Qu.:64.75	3rd Qu.:0.3912	3rd Qu.:0.23775	3rd Qu.:0.1837
Max. :86.00	Max. :0.9640	Max. :0.91100	Max. :0.5070

V5	V6	V7	V8
Min. :0.0110	Min. :0.01200	Min. :0.01100	Min. :0.01100
1st Qu.:0.0540	1st Qu.:0.05125	1st Qu.:0.03875	1st Qu.:0.03200
Median :0.0735	Median :0.07700	Median :0.06350	Median :0.04450
Mean :0.1011	Mean :0.09319	Mean :0.08303	Mean :0.05827
3rd Qu.:0.1330	3rd Qu.:0.11000	3rd Qu.:0.09500	3rd Qu.:0.07500
Max. :0.2960	Max. :0.40900	Max. :0.41400	Max. :0.28600

The summary statistics of the 7 variables, V2-V8 (V1 is a count), reveal that every observation is bounded between 0 and 1.

2) Generating v_i from fixed (μ, σ, ν, ϕ)

In order to generate samples $(v_1), ..., (v_{86})$ from $v_i \sim logN(\nu, \phi^2)$ I need to specify my parameters (μ, σ, ν, ϕ). I will specifiy these parameters by sampling from their joint densities, which are written above.

I will generate an estimate for μ by randomly generating 1,000,000 observations from $f(\mu) \sim N(0, \xi^2)$, then taking the average of all those samples. In order to do this, I must pick a ξ^2 value that I deem appropriate, which I chose to be 0.1. This makes the estimate for $\mu = 0.00014$.

I will generate an estimate for ν using the same process as I did for μ. I will generate 1,000,000 random samples from $f(\nu) \sim N(0, \psi^2)$, then take the average of all samples and use that as my fixed ν. Since ψ^2 is not already specified, I will chose $\psi^2 = 0.05$. This makes the estimate of ν=0.000024.

Following the same steps, I will generate an estimate for σ^2. Recall, $f(\lambda) = \frac{1}{\sigma^2} \sim Ga(a, b)$. This means I will be randomly sampling $f(\lambda)$, then solving for σ^2. In order to do this, I must chose an (a, b) so I can generate 1,000,000 samples from $f(\lambda)$. I pick $a = 2$ and $b = 1$. This gave me an average estimate of $\lambda = 2.0$. This makes $\sigma = \sqrt{\frac{1}{\lambda^2}} = \sqrt{0.25}$, or $\sigma^2 = 0.25$.

Lastly, I will pick a ϕ^2 by randomly sampling from $f(\eta) = \frac{1}{\phi^2} \sim Ga(c, d)$. The (c, d) that I chose are $c = 3$, and $d = 2$. This gives me an estimate for $\eta = 6.0$ averaged over 1,000,000 samples. This means that $\phi = \sqrt{\frac{1}{\eta^2}} = \sqrt{0.02778}$, or that $\phi^2 = 0.02778$

Stated above, $log(v_i) \sim N(\nu, \phi^2)$ was given, this means that the distribution of v_i is a log-normal, since it is simply $e^{log(v_i)}$. This is why I will now generate estimates for $v_i = (v_1), ..., (v_{86})$ from $logN(\nu, \phi^2)$ using the fixed parameters that I generated above. I will sample from $logN(\nu, \phi^2)$ 10,000 times, then take the average v_i across all the generated samples and use that to continue my analysis. This leaves me with a 86 x 1 matrix with each entry as the bootstrapped sample for v_i for every $i = 1, .., 86$.

3) Generating $log(y_{ij})$

I will now generate $log(y_{ij})$ from $N(\mu+log(\frac{v_i}{v_i+x_{ij}}), \sigma^2)$. I will be using the fixed parameters I determined earlier, the estimates for v_i, an 86x1 matrix, and the data, x_{ij}, and 86x7 matrix. I will sample $N(\mu+log(\frac{v_i}{v_i+x_{ij}}), \sigma^2)$ 100,000 times, then take the average of all entries to have the final 86x7 matrix of $log(y_{ij})$ that I will use for implementing the Markov Chain and further analysis.

4) Code Markov Chain

Now that I have my generated $log(y_{ij})$'s, I now must code the Markov chain of the conditional posterior denisties, $f(\mu|...), f(\nu|...), f(\eta|...), f(\lambda|...)$ and $f(v_i|...)$. My goal is to have Monte Carlo estimates for all of my parameters and for each v_i, that I can use to fit $log(y_{ij})$ and ultimately determine a value for v_{87}.

In order to do this, I must use a target density, which is $f(\theta)$, where $\theta = (y_{ij}, v_i, \mu, \nu, \lambda, \eta)$. This means that there will be 6 conditional densities.

We were given the conditional density for $log(y_{ij})$ in the beginning, as

$f(log(y_{ij})|v_i, \mu, \nu, \lambda, \eta) \sim N(\mu + log(\frac{v_i}{v_i+x_{ij}}), \sigma^2)$.

This will be used to solve for other conditional posterior densities.

To begin, I will define the joint density,

$f(y_{ij}, v_i, \mu, \nu, \lambda, \eta) = f(log(y_{ij})|v_i, \mu, \nu, \lambda, \eta) * f(v_i|\mu, \nu, \lambda, \eta) * f(\mu) * f(\nu) * f(\lambda) * f(\eta)$

The first conditional posterior density I will find is

$f(\lambda|y_{ij}, v_i, \mu, \nu, \eta) \propto f(\lambda) \prod_{i,j} f(log(y_{ij})|v_i, \mu, \nu, \lambda, \eta)$

since $\lambda = \frac{1}{\sigma^2}$, the only other density where λ shows up is in $f(log(y_{ij})|v_i, \mu, \nu, \lambda, \eta)$. This makes

$f(\lambda|y_{ij}, v_i, \mu, \nu, \eta) \propto \lambda^{a-1} e^{-b\lambda} \prod_{i,j} \frac{1}{\sqrt{2\pi\sigma^2}} e^{-\frac{(log(y_{ij})-\mu-log(\frac{v_i}{v_i+x_{ij}}))^2}{2\sigma^2}}$

$\propto \lambda^{a-1} e^{-b\lambda} \lambda^{\frac{nm}{2}} e^{-\frac{\lambda}{2}\sum_{i,j}(log(y_{ij})-\mu-log(\frac{v_i}{v_i+x_{ij}}))^2}$

$\propto \lambda^{a+\frac{nm}{2}-1} e^{-b\lambda-\frac{\lambda}{2}\sum_{i,j}(log(y_{ij})-\mu-log(\frac{v_i}{v_i+x_{ij}}))^2}$

Therefore $f(\lambda|y_{ij}, v_i, \mu, \nu, \eta) \sim Ga(a + \frac{nm}{2}, b + \frac{1}{2}\sum_{i,j}(log(y_{ij}) - \mu - log(\frac{v_i}{v_i+x_{ij}})))$

The next conditional posterior density I will calculate is

$f(\eta|y_{ij}, log(v_i), \mu, \nu, \lambda) \propto f(\eta) \prod_{i=1}^{n} f(log(v_i)|\mu, \nu, \lambda, \eta)$

Here, $\eta = \frac{1}{\phi^2}$, which only shows up in the conditional density $f(logv_i|\mu, \nu, \lambda, \eta)$. I only take the product over all the individuals, i, since the seven traits, j, do not enter these densities. I will use the density given for $log(v_i) \sim N(\nu, \phi^2)$ instead of $v_i \sim logN(\nu, \phi^2)$ when solving for this conditional density, since it makes calculations more simple.

$f(\eta|log(y_{ij}), log(v_i), \mu, \nu, \lambda) \propto \eta^{c-1} e^{-d\eta} \prod_{i=1}^{n} \frac{1}{\sqrt{2\pi\phi^2}} e^{-\frac{(logv_i-\nu)^2}{2\phi^2}}$

$\propto \eta^{c-1} e^{-d\eta} \prod_{i=1}^{n} \eta^{\frac{1}{2}} e^{-\frac{\eta}{2}(log(v_i)-\nu)^2}$

$\propto \eta^{c-1} e^{-d\eta} \eta^{\frac{n}{2}} e^{-\frac{\eta}{2}\sum_{i=1}^{n}(log(v_i)-\nu)^2}$

$\propto \eta^{c+\frac{n}{2}-1} e^{-d\eta-\frac{\eta}{2}\sum_{i=1}^{n}(log(v_i)-\nu)^2}$

Therefore $f(\eta|y_{ij}, log(v_i), \mu, \nu, \lambda) \sim Ga(c + \frac{n}{2}, d + \frac{1}{2}\sum_{i=1}^{n}(log(v_i) - \nu)^2)$

The next conditional posterior density I will calculate is

$f(\mu|y_{ij},v_i,\nu,\lambda,\eta) \propto f(\mu)\prod_{i,j} f(log(y_{ij})|v_i,\mu,\nu,\lambda,\eta)$

This is the form of the conditional density, because μ only shows up in $f(\mu)$ and $f(log(y_{ij})|v_i,\mu,\nu,\lambda,\eta)$.

$f(\mu|y_{ij},v_i,\nu,\lambda,\eta) \propto \frac{1}{\sqrt{2\pi\xi^2}}e^{-\frac{\mu^2}{2\xi^2}}\prod_{i,j}\frac{1}{\sqrt{2\pi\sigma^2}}e^{-\frac{(log(y_{ij})-\mu-log(\frac{v_i}{v_i+x_{ij}}))^2}{2\sigma^2}}$

$\propto e^{-\frac{\mu^2}{2\xi^2}}e^{-\frac{\lambda}{2}\sum_{i,j}(log(y_{ij})-\mu-log(\frac{v_i}{v_i+x_{ij}}))^2}$

I will call $log(y_{ij})-log(\frac{v_i}{v_i+x_{ij}})=K_{ij}$ and $\frac{1}{\xi^2}=\tau$ which gives

$\propto e^{-\frac{1}{2}(\tau\mu^2-\lambda\sum_{i,j}(K_{ij}-\mu)^2)}$

$\propto e^{-\frac{1}{2}(\tau\mu^2-\lambda\sum_{i,j}(K_{ij}^2-2\mu K_{ij}+\mu^2))}$

$\propto e^{-\frac{1}{2}(\tau\mu^2-\lambda\sum_{i,j}K_{ij}^2+2\lambda\mu\sum_{i,j}K_{ij}-\lambda\mu^2 nm)}$

$\propto e^{-\frac{1}{2}(\mu^2(\tau-\lambda nm)+\mu(2\lambda\sum_{i,j}K_{ij})-\lambda\sum_{i,j}K_{ij}^2)}$

$\propto e^{-\frac{(\tau-\lambda nm)}{2}(\mu^2+\mu\frac{(2\lambda\sum_{i,j}K_{ij})}{(\tau-\lambda nm)}-\frac{\lambda\sum_{i,j}K_{ij}^2}{(\tau-\lambda nm)})}$

$\propto e^{-\frac{(\tau-\lambda nm)}{2}(\mu^2-\mu\frac{(2\lambda\sum_{i,j}K_{ij})}{(\lambda nm-\tau)}+\frac{\lambda\sum_{i,j}K_{ij}^2}{(\lambda nm-\tau)^2})}$

$\propto e^{-\frac{(\tau-\lambda nm)}{2}(\mu-\frac{\lambda\sum_{i,j}K_{ij}}{(\lambda nm-\tau)})^2}$

Therefore $f(\mu|y_{ij},v_i,\nu,\lambda,\eta)\sim N(\frac{\lambda\sum_{i,j}K_{ij}}{(\lambda nm-\tau)},\frac{1}{(\tau-\lambda nm)})$

substituting back in $log(y_{ij})-log(\frac{v_i}{v_i+x_{ij}})=K_{ij}$ and $\frac{1}{\xi^2}=\tau$,

$f(\mu|y_{ij},v_i,\nu,\lambda,\eta)\sim N(\frac{\lambda\sum_{i,j}(log(y_{ij})-log(\frac{v_i}{v_i+x_{ij}}))}{(\lambda nm-\frac{1}{\xi^2})},\frac{1}{(\frac{1}{\xi^2}-\lambda nm)})$

The next conditional posterior densitu I will calculate is

$f(\nu|y_{ij},v_i,\mu,\lambda,\eta)\propto f(\nu)\prod_{i=1}^n f(log(v_i)|\mu,\nu,\lambda,\eta)$

I will again use $log(v_i)\sim N(\nu,\phi^2)$ for simplicity of the calculation, keeping in mind ν is the parameter I'm interested in.

$f(\nu|y_{ij},v_i,\mu,\lambda,\eta)\propto\frac{1}{\sqrt{2\pi\psi^2}}e^{-\frac{\nu^2}{2\psi^2}}\prod_{i=1}^n\frac{1}{\sqrt{2\pi\phi^2}}e^{-\frac{(logv_i-\nu)^2}{2\phi^2}}$

$\propto e^{-\frac{\nu^2}{2\psi^2}}\prod_{i=1}^n e^{-\frac{\eta}{2}(log(v_i)-\nu)^2}$

$\propto e^{\frac{1}{2}(\frac{1}{\psi^2}\nu^2-\eta\sum_i(log(v_i)-\nu)^2)}$

I will set $\frac{1}{\psi^2}=A$.

$\propto e^{-\frac{1}{2}(A\nu^2-\eta\sum_i(log(v_i)^2-2log(v_i)\nu+\nu^2))}$

$\propto e^{-\frac{1}{2}(A\nu^2-\eta\sum_i(log(v_i)^2)+2\eta\nu\sum_i log(v_i)-\eta n\nu^2))}$

$\propto e^{-\frac{1}{2}(\nu^2(A-\eta n)-\nu(2\eta\sum_i log(v_i))-\eta\sum_i(log(v_i)^2)))}$

$\propto e^{-\frac{(A-\eta n)}{2}(\nu^2-\nu(\frac{2\eta\sum_i log(v_i)}{\eta n-A})+\frac{\eta\sum_i log(v_i)^2}{(\eta n-A)^2})}$

$\propto e^{-\frac{(A-\eta n)}{2}(\nu^2-\frac{\eta\sum_i log(v_i)}{(\eta n-A)})^2}$

Therefore, $f(\nu|y_{ij},v_i,\mu,\lambda,\eta)\sim N(\frac{\eta\sum_i log(v_i)}{\eta n-A},\frac{1}{A-\eta n})$

The last conditional posterior density I will calculate is for

$f(v_i|log(y_{ij}, \mu, \nu, \lambda, \eta) \propto f(v_i|\mu, \nu, \lambda, \eta) \prod_{i,j} f(log(y_{ij})|v_i, \mu, \nu, \lambda, \eta)$

I will be using the distribution for v_i, not for $log(v_i)$, because contained within $log(y_{ij})$ is both $log(v_i)$ and $log(v_i + x_{ij})$, which would casue some issues if I only find the conditional on $log(vi)$. Recall that $v_i \sim logN(\nu, \phi^2)$. I only take the product across j = 1,. . . ,m of $log(y_{ij})$ since v_i is only indexed across i = 1,. . . ,n, where m=7 and n=86.

This renders a conditional density of

$f(v_i|log(y_{ij}, \mu, \nu, \lambda, \eta) \propto \frac{1}{v_i\sqrt{2\pi\phi^2}} e^{-\frac{(logv_i-\nu)^2}{2\phi^2}} \prod_j \frac{1}{\sqrt{2\pi\sigma^2}} e^{-\frac{(log(y_{ij})-\mu-log(\frac{v_i}{v_i+x_{ij}}))^2}{2\sigma^2}}$

$\propto \frac{1}{v_i} e^{-\frac{\eta}{2}(logv_i-\nu)^2} e^{-\frac{\lambda}{2}\sum_j (y_{ij}-\mu-log(v_i)+log(v_i+x_{ij})^2}$

$\propto \frac{1}{v_i} e^{-\frac{1}{2}(\eta(logv_i-\nu)^2-\lambda\sum_j (y_{ij}-\mu-log(v_i)+log(v_i+x_{ij})^2)}$

The form of this conditional density makes me believe that is should follow some log Normal distribution, but I am unable to find the exact parameters of the conditional distribution, since I am unable to isolate v_i in the above equation. This means that I will have to implement a Metropolis step in order to sample from this conditional density, since I cannot sample directly from it. I will need to implement the Metropolis Hastings algorithm seperately for the 86 v_i estimates, since v_i is i.i.d..

Gibbs Sampler with MH for all other conditionals

Now that I have found the conditional densities, I must take samples from them, so I can estimate $f(log(y_{ij}))$. I will implement a Gibbs Sampler to estimate the conditional densities. The Gibbs Sampler can be implemented since there are more than 2 parameters in this model, and I am able to sample directly from all the conditional densities besides $f(v_i|log(y_{ij}, \mu, \nu, \lambda, \eta)$, where I will introduce a Metropolis step.

Some information on why we can use a Gibbs Sampler follows. In a Gibbs framework the we define the transition density $(\lambda, \eta, \nu, \mu, v_1, ..., v_{86}) \rightarrow (\lambda', \eta', \nu', \mu', v_1', ..., v_{86}')$ by $P(\lambda', \eta', \nu', \mu', v_1', ..., v_{86}'|\lambda, \eta, \nu, \mu, v_1, ..., v_{86})$

$P(\lambda', \eta', \nu', \mu', v_1', ..., v_{86}'|\lambda, \eta, \nu, \mu, v_1, ..., v_{86}) = f(\lambda'|\eta, \nu, \mu, v_1, ..., v_{86}) * f(\eta'|\lambda', \nu, \mu, v_1, ..., v_{86}) * f(\nu'|\lambda', \eta', \mu, v_1, ..., v_{86}) * f(\mu'|\lambda', \eta', \nu', v_1, ..., v_{86}) * f(v_1'|\lambda', \eta', \nu', \mu', v_2, ..., v_{86}) * ...f(v_{86}'|\lambda', \eta', \nu', \mu', v_1', ..., v_{85}')$

I will then take N many iterations of the transition densities. We want a transition density that will give us a stationary density $f(\lambda', \eta', \nu', \mu', v_1', ..., v_{86}') = \int P(\lambda', \eta', \nu', \mu', v_1', ..., v_{86}'|\lambda, \eta, \nu, \mu, v_1, ..., v_{86}) f(\lambda, \eta, \nu, \mu, v_1, ..., v_{86})$. I will simplify this to $f(\theta', v_i') = \int\int P(\theta', v_i'|\theta, v_i) f(\theta, v_i) = \int\int f(v_i'|\theta', v_i) f(\theta'|\theta, v_i) f(\theta, v_i) d\theta dv_i$

Using $f(\theta'|\theta, v_i) = \frac{f(\theta', \theta, v_i)}{f(\theta, v_i)}$, I can now simplify to

$f(\theta', v_i') = \int\int f(v_i'|\theta', v_i) f(\theta', \theta, v_i) d\theta dv_i$

$f(\theta', v_i') = \int f(v_i'|\theta', v_i) f(\theta', v_i) dv_i = \int \frac{f(v_i', \theta', v_i)}{f(\theta', v_i)} f(\theta', v_i) dv_i$

$\int f(v_i', \theta', v_i) dv_i = f(\theta', v_i')$

This confirms that the transition density does render a stationary density, $f(\theta', v_i')$.

We may get stuck when trying to do these calculations in the Gibbs framework since v_i cannot be sampled in that framework. This means to sample from $f(v_i|log(y_{ij}, \mu, \nu, \lambda, \eta)$, we will need a transition density that satisfies $f(v_i)q(v_i'|v_i) = f(v_i')q(v_i|v_i')$. This means $p(v_i'|v_i)$ shows up in the form $p(v_i'|v_i) = \alpha(v_i, v_i')q(v_i') + (1 - r(v_i))1(v_i' = v_i)$.

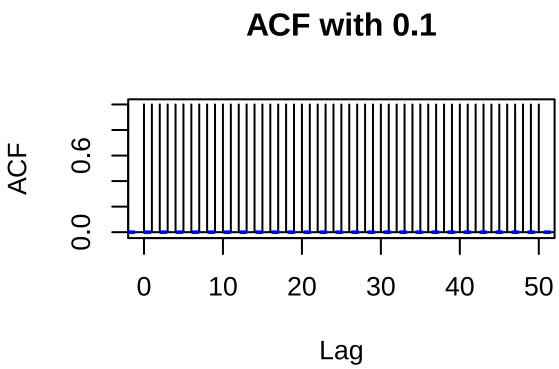
Plugging in for $p(v_i'|v_i)$, we get $f(v_i)\alpha(v_i, v_i')q(v_i') = f(v_i')\alpha(v_i, v_i')q(v_i)$. In the Metropolis Hastings algorithm, $\alpha = min(1, \frac{f(v_i')q(v_i|v_i')}{f(v_i)q(v_i'|v_i)})$.

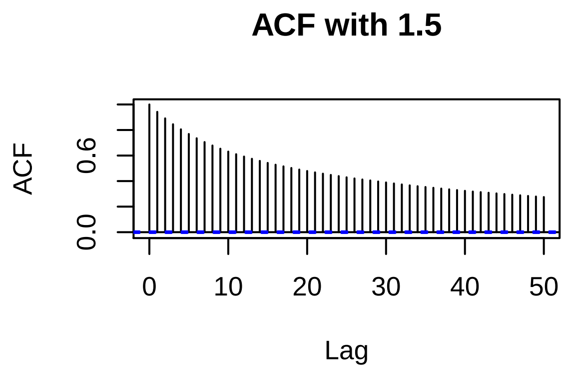
The algorithm I will take to implement the Metropolis Hastings step is 1) Decide current state $v_{i,n}$ (this will be the entries in my generated v_i matrix) 2) Take a proposal state such that $q(v_i'|v_i) \sim logNormal(v_i'|v_i, \beta^2)$ 3) Take $u \sim U(0,1)$ 4) If $u < \alpha(v_i, v_i')$, then $v_{i,n+1} = v_i'$, else $v_{i,n+1} = v_{i,n}$

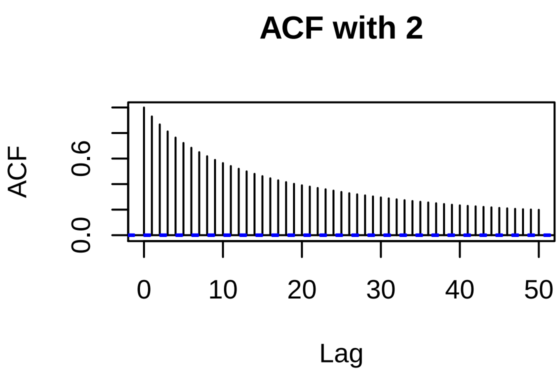
I will implement this MH step for v_i for every $i = 1, ..., 86$. To obtain the estimate for each v_i, I simply will take the mean of all the respective v_i generated from my MH algorithm. I iterated both the Gibbs Sampler and Metropolis Hastings step 110,000 times with a 10,000 sample burnin, to render 100,000 generated estimates for each parameter.

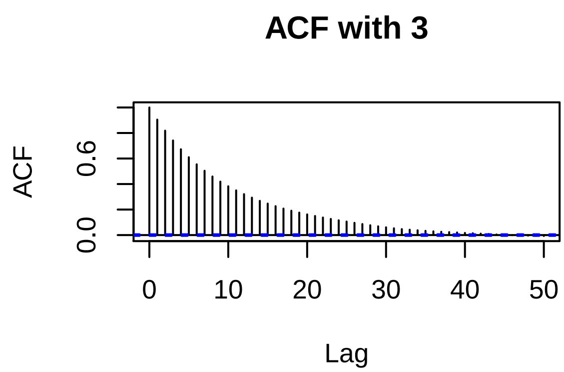
5) Report output of chain

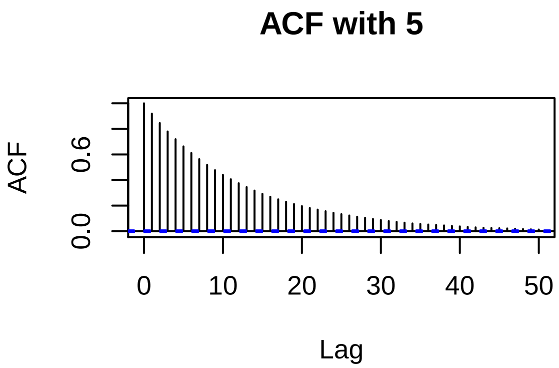
In order to determine which β^2 I want to include in the proposal density of the MH step, $q(v_i'|v_i) \sim logNormal(v_i'|v_i, \beta^2)$, I tested several different values and picked the one that had the lowest, or quickest converging ACF. A plot of the ACF's for β^2 = (0.1, 1.5, 2, 3, 5, 10, 50). I am using the Metropolis Hastings estimate for v_1 in the plots below. Since each v_i is i.i.d., I assumed that a β^2 for v_1 will also be reasonable for the other 85 v_i's.

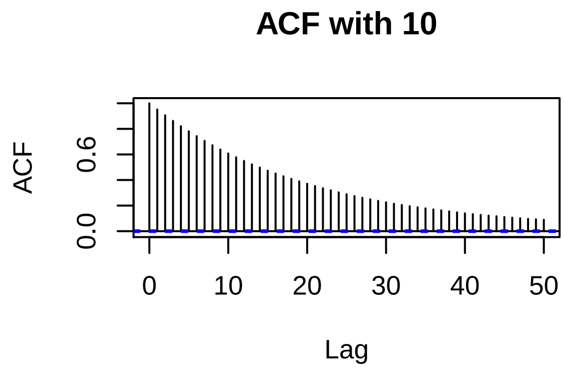


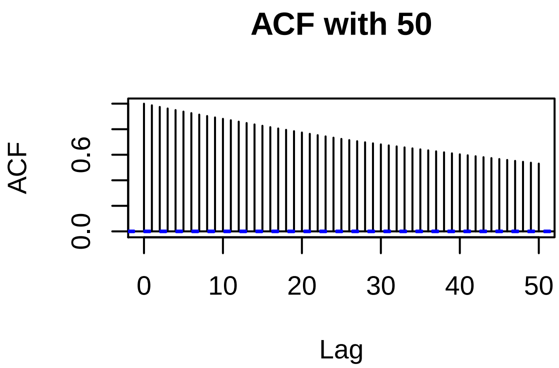




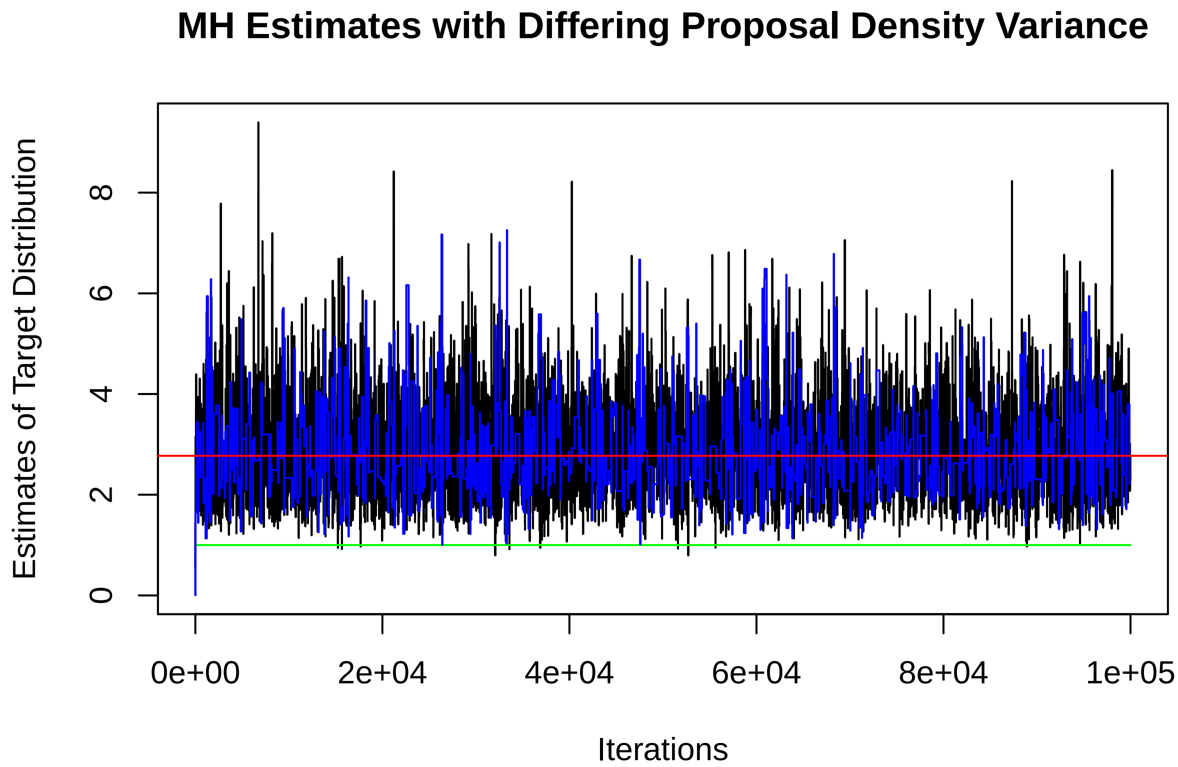




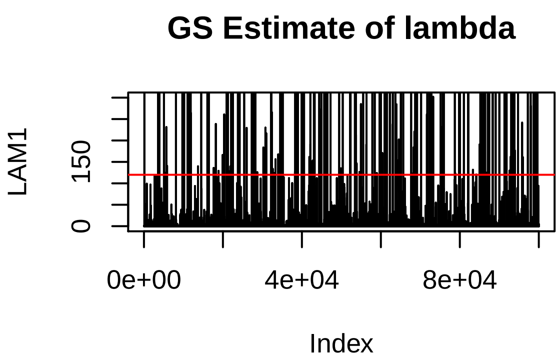


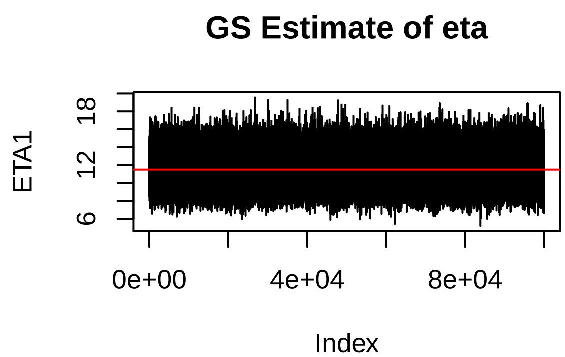


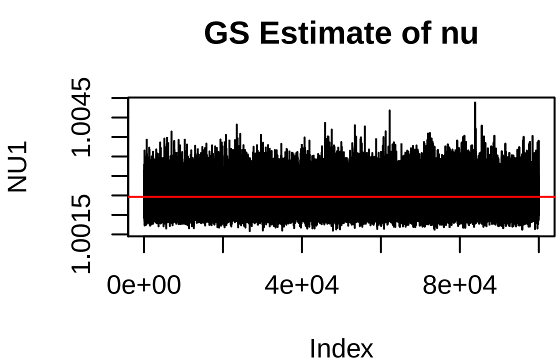
The ACF plots above reveal that a $\beta^2 = 3$ has the least autocorrelated samples of v_i, since the ACF function converges to zero the fastest compared to other β^2's. The graph below has lines of three MH estimates of v_1 for β^2's equal to 0.1, 3, and 50. We see that MH estimates from proposals with very low and very large variances are persistent and highly correlated to past samples. The MH estimate with an appropriately chosen proposal density variance converges to the true value of v_1 much faster. The MH estimate of v_1 is resembled by the red horizontal line on the graph.

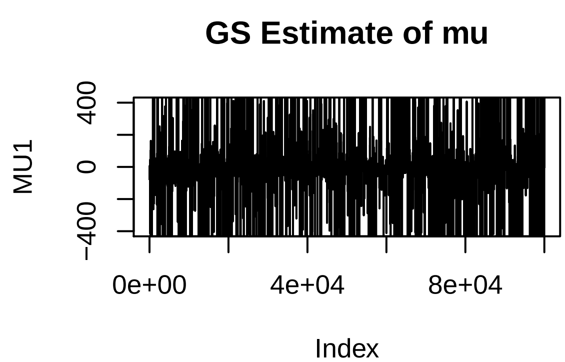


Below, I graphed the convergence of the estimates for the four other parameters, λ, η, ν, μ, estimated using the Gibbs Sampler. The horizontal red line is the estimate value of the respective parameters.









To finalize the output from the chain, below is a table of the Monte Carlo estimates for each parameter in the model. The estimates were calculated by the average of the 100,000 generated estimates from either the

Gibbs Sampler or Metropolis Hastings step for each respective parameter.

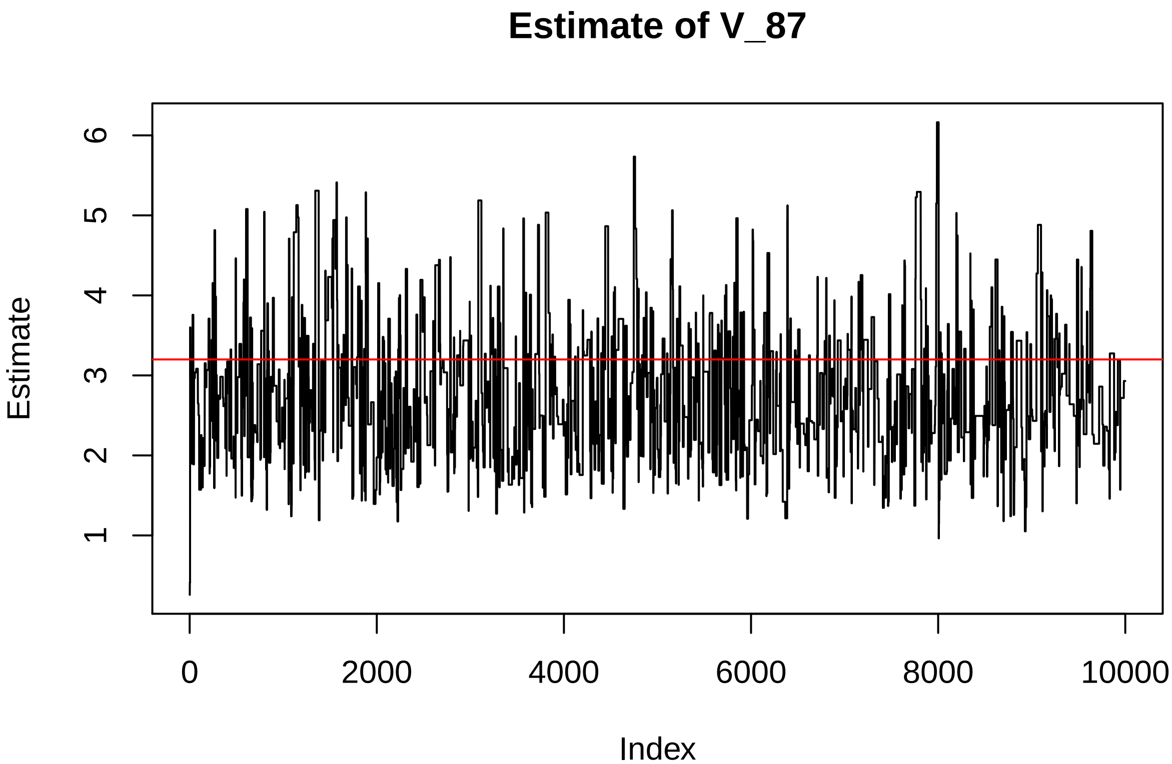
```
##          LAM1          ETA1           NU1           MU1              
##  1.199456e+02  1.149280e+01  1.002464e+00 -3.993270e+06  3.059693e+00 
##                                                                       
##  2.862872e+00  2.859581e+00  2.813779e+00  2.843475e+00  2.866993e+00 
##                                                                       
##  2.841031e+00  2.841610e+00  2.846661e+00  2.839789e+00  2.872932e+00 
##                                                                       
##  2.883032e+00  2.844536e+00  2.860234e+00  2.841633e+00  2.860492e+00 
##                                                                       
##  2.848366e+00  2.878490e+00  2.853686e+00  2.864761e+00  2.843102e+00 
##                                                                       
##  2.873618e+00  2.833961e+00  2.852434e+00  2.833910e+00  2.845875e+00 
##                                                                       
##  2.855804e+00  2.853000e+00  2.843076e+00  2.830444e+00  2.851942e+00 
##                                                                       
##  2.831745e+00  2.842067e+00  2.860739e+00  2.855515e+00  2.856475e+00 
##                                                                       
##  2.861674e+00  2.848602e+00  2.866302e+00  2.835882e+00  2.852899e+00 
##                                                                       
##  2.843300e+00  2.863989e+00  2.856650e+00  2.851324e+00  2.860309e+00 
##                                                                       
##  2.857020e+00  2.167454e+01  2.837052e+00  1.847006e+01  2.832940e+00 
##                                                                       
##  2.837146e+00  2.833774e+00  2.851262e+00  2.864572e+00  2.861129e+00 
##                                                                       
##  2.855566e+00  2.840053e+00  2.826487e+00  2.861927e+00  2.836985e+00 
##                                                                       
##  2.864352e+00  2.837557e+00  2.848588e+00  2.850235e+00  2.844762e+00 
##                                                                       
##  2.856183e+00  2.830151e+00  2.833596e+00  2.858889e+00  2.840102e+00 
##                                                                       
##  2.871495e+00  2.856291e+00  2.845871e+00  2.861714e+00  2.843539e+00 
##                                                                       
##  2.844165e+00  2.846672e+00  2.842949e+00  2.799516e+00  2.853721e+00 
##                                                                       
##  2.861624e+00  2.852210e+00  2.851662e+00  2.844858e+00  2.849888e+00
```

6) Estimating v__87

Two pieces of information given was that $y_{87,1} = 0.257$ for $x_{87,1} = 1$. This will help to determine the value of $v_[87]$, along with the Monte Carlo estimates that were found above.

Since I have found the estimates for $(\lambda, \mu, \eta, \nu)$, I do not need to re-run the Gibbs Sampler again on all the parameters to find v_87. I will use the Metropolis Hastings step to generate 100,000 samples of v_{87}, then average over those samples to find the estimate for v_{87}. In order to implement the Metropolis Hastings step, I appended an extra row to the $loy(y_{ij})$ matrix and to the $x_{ij} = j$ vector, to make a place to store the estimate.

I will run the MH algorithm for the v_i's, using the estimated for the parameters $(\lambda, \mu, \eta, \nu)$, and find an estimate of v_{87} given the other fixed parameters, $y_{87,1} = 0.257$ for $x_{87,1} = 1$, too. I specified a β^2 of the proposal density as 3 as well.



This process rendered me an estimate for $v_{87} = 2.798$ for an idividual with the parameters specified above. As we can see, under this proposal density variance, the estimate for $v_8 7$ does not appear to be highly correlated with past estimates.