Take Home Exam

STAT515 Ling Zhang

1(a)

The joint distribution of Y_i and β_1 is

$$f(Y,\beta_1|X)=\left(\prod_{i=1}^{n}\frac{\lambda}{2}\exp(\frac{\lambda}{2}(2(\beta_0+\beta_1X_i)+\lambda\sigma^2-2Y_i))(\frac{2}{\pi}\int_{\frac{(\beta_0+\beta_1X_i)+\lambda\sigma^2-Y_i}{\sqrt{2}\sigma}}^{\infty}e^{-t^2}\,dt)\right)\left(\frac{1}{\sqrt{20\pi}}e^{-\frac{\beta_1{}^2}{20}}\right)$$

Given $\beta_0=5,\lambda=0.4,\sigma_i=1$ for all i and sample Y and X, we have the posterior distribution $h(\beta_1|X,Y)$ as follows:

$$c*h(\beta_1|X,Y)=f(Y,\beta_1|X)=\left(\prod_{i=1}^{n}\frac{1}{5}\exp(\frac{1}{5}(2(5+\beta_1X_i)+\frac{2}{5}-2Y_i))(\frac{2}{\pi}\int_{\frac{(5+\beta_1X_i)+\frac{2}{5}-Y_i}{\sqrt{2}}}^{\infty}e^{-t^2}\,dt)\right)\left(\frac{1}{\sqrt{20\pi}}e^{-\frac{\beta_1{}^2}{20}}\right)$$

To do MCMC in a lazy way, we just use $f(Y,\beta_1|X)|_Y$ in M-H Algorithm because $h(\beta_1|X,Y)=\frac{f(Y,\beta_1|X)}{c}$, c is a constant of β_1. Now I give the Metropolis-Hastings algorithm:

Step1. Choose the symmetric proposal N (mean, sd) = N (x, 1) and random initial value of: $\beta_{10}=0,4\ and\ 8$.

Step2. Get the density value of $f(Y_i|\beta_1,X)$ for Y_i from the given R code and get joint density of $f(Y,\beta_1|X)$.

Step3. Given $\beta_{1\text{n}}$, generate β' from proposal $\beta'\sim N(\beta_{1\text{n}},1)$ and calculate $f(Y,\beta'|X)$

Step4. Calculate acceptance probability: $\alpha=\min\{1,\frac{f(Y,\beta'|X)}{f(Y,\beta_{1\text{n}}|X)}\}$, generate uniform random variable u from U (0, 1).

Step5. If u>= $\alpha(x,y)$, set $\beta_{1(\text{n}+1)}=\beta_{1\text{n}}$, else set $\beta_{1(\text{n}+1)}=\beta'$.

Step6. Keep doing from step1 to step5 for N times. Then we get N MCMC samples from the posterior distribution.

By following this algorithm, we can get MCMC samples of certain size.

(b)

From the algorithm from (a) we generate N=50000 samples from posterior distribution. From the given function "bm" from "batchmeans", we get that their sample mean and standard error are 7.341001 and 0.00320226.

(c)

By using R command "quantile", we get the 95% credible interval for β_1 based on my sample as follows:

[6.718232, 7.932105]

(d) & (e)

The following first plot consists of three density plots from different initial values of $\beta_1=0,4\ and\ 8$. Color black, red and green in each plot are respectively for three initial values. From the plot we can see they are almost the same. We can see from the following Plot2 and Plot3 that our estimate is very stable, its standard error is smaller than 0.01 after sample size is larger than 30000. Estimates from different initial values convergent to the same place. Also, autocorrelation is not that serious. Autocorrelation respectively truncated after 15, 14 and 2 lags. The estimate effective sample sizes for different initial values are 9670.716, 10521.03 and 10591.74. They are much larger than 5000, which is pretty good. The information above can show that my approximation is very good and accurate. All the mentioned plots are as follows:

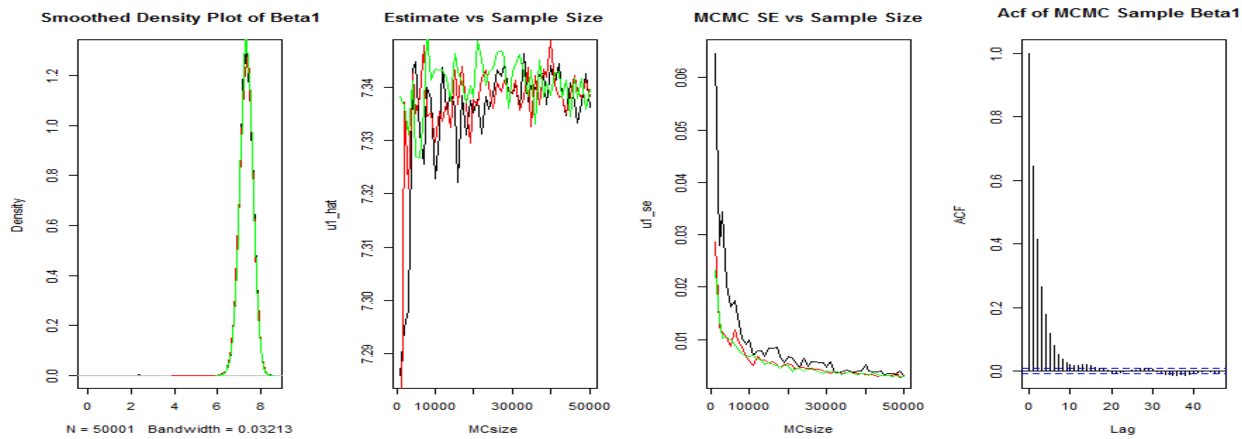

Fig.1. Smoothed Density Plots of Beta1 & Supporting Plot for the Accuracy of Our Approximations for Data #1

2(a)

Given the prior distribution of β_0, β_1 *and* λ and conditional distribution of Y, we have the following joint distribution of Y_i, β_0, β_1 *and* λ

$$f(Y, \beta_0, \beta_1, \lambda|X) = \left(\prod_{i=1}^{n} \frac{\lambda}{2} \exp(\frac{\lambda}{2}(2(\beta_0 + \beta_1 X_i) + \lambda\sigma^2 - 2Y_i))(\frac{2}{\pi}\int_{\frac{(\beta_0+\beta_1 X_i)+\lambda\sigma^2-Y_i}{\sqrt{2}\sigma}}^{\infty} e^{-t^2}\,dt)\right)$$
$$* \left(\frac{1}{\sqrt{20\pi}} e^{-\frac{{\beta_0}^2}{20}}\right)\left(\frac{1}{\sqrt{20\pi}} e^{-\frac{{\beta_1}^2}{20}}\right)\left(\frac{1}{\Gamma(0.01)100^{0.01}}\lambda^{0.01-1} e^{-\frac{\lambda}{100}}\right)$$

Given sample Y and X, we can get the full conditional distribution of β_0, β_1 *and* λ, by dividing the above joint distribution by marginal distribution of all other parameters, which is a constant function of β_0, β_1 *or* λ. So we have

$$c_0(\beta_0) * h(\beta_0|X, Y, \beta_1, \lambda) = f(Y, \beta_0, \beta_1, \lambda|X)$$
$$c_1(\beta_1) * h(\beta_1|X, Y, \beta_0, \lambda) = f(Y, \beta_0, \beta_1, \lambda|X)$$
$$c_2(\lambda) * h(\lambda|X, Y, \beta_0, \beta_1) = f(Y, \beta_0, \beta_1, \lambda|X)$$

Here $c_i(.)$ is responding marginal distribution of all other parameters. To do MCMC in a lazy way, for all the above cases, we just use $f(Y, \beta_0, \beta_1, \lambda|X)|_{Y,\beta_1,\lambda,X}$, $f(Y, \beta_0, \beta_1, \lambda|X)|_{Y,\beta_0,\lambda,X}$ and $f(Y, \beta_0, \beta_1, \lambda|X)|_{Y,\beta_1,\beta_0,X}$ in M-H Algorithm because all the above three conditional distribution are all proportional to the joint distribution by a constant.

Now I give the following "Variable at a time" Metropolis-Hastings algorithm:

Step1. Choose initial values of β_0, β_1 *and* λ: $\beta_{00} = 5, \beta_{10} = 7$ *and* $\lambda_0 = 0.4$; $\beta_{00} = 3, \beta_{10} = 4$ *and* $\lambda_0 = 0.3$; $\beta_{00} = 1, \beta_{10} = 1$ *and* $\lambda_0 = 0.2$; the symmetric proposal N (x, 0.7) for β_0 *and* β_1 and symmetric proposal U (max {0, x-0.3}, x+0.3) for λ_0. We choose these proposals with given variances because they give best ESS after trying different times. Initial values are randomly chosen in the domain of the distributions.

Step2. Get the density value of $f(Y_i|\beta_0, \beta_1, \lambda, X)$ for Y_i from the given R code and get joint density of $f(Y, \beta_0, \beta_1, \lambda|X)$, where $\beta_0 \sim N(0, 10), \beta_1 \sim N(0, 10)$ *and* $\lambda \sim Gamma(0.01, 100)$.

Step3. Given β_{0n}, generate ${\beta_0}'$ from proposal ${\beta_0}' \sim N(\beta_{0n}, 1)$ using the most up to date other parameters.

Step4. Calculate the acceptance probability: $\alpha = \min\{1, \frac{f(Y,\beta_1,\lambda,{\beta_0}'|X)}{f(Y,\beta_1,\lambda,\beta_{0n}|X)}\}$, generate uniform random variable u from U (0, 1). If u>= α, set $\beta_{0(n+1)} = \beta_{0n}$, else set $\beta_{0(n+1)} = {\beta_0}'$.

Step5. Given β_{1n}, generate $\beta_1{}'$ from proposal $\beta_1{}' \sim N(\beta_{1n}, 1)$ using the most up to date other parameters.

Step6. Calculate the acceptance probability: $\alpha = \min\{1, \frac{f(Y,\beta_0,\lambda,\beta_1{}'|X)}{f(Y,\beta_0,\lambda,\beta_{1n}|X)}\}$, generate uniform random variable u from U (0, 1). If u>= α, set $\beta_{1(n+1)} = \beta_{1n}$, else set $\beta_{1(n+1)} = \beta_1{}'$.

Step7. Given λ_n, generate λ' from proposal $\lambda' \sim \mathrm{U}\ (\max\{0, \mathrm{x} - 0.2\}, \mathrm{x} + 0.2)$using the latest other parameters.

Step8. Calculate the acceptance probability: $\alpha = \min\{1, \frac{f(Y,\beta_0,\lambda',\beta_1|X)}{f(Y,\beta_0,\lambda_n,\beta_1|X)}\}$, generate uniform random variable u from U (0, 1). If u>= α, set $\lambda_{n+1} = \lambda_n$, else set $\lambda_{n+1} = \lambda'$.

Step9. Keep doing from step1 to step8 for N times. Then we get N MCMC samples from the posterior distribution.

By following this algorithm, we can get MCMC samples of certain size.

(b)

By implementing above algorithm in R, we get the following results:

	Posterior Means and SEs	95% Credible Intervals
β_0	2.346868 (0.002528495)	(2.067457 2.605493)
β_1	3.45782 (0.003535598)	(3.052606 3.874303)
λ	0.8030586 (0.0006032448)	(0.6952092 0.9255032)

Table.1. Posterior Means, Standard Errors and 95%Credible Intervals of β_0, β_1 and λ for Data #2

(c)

It is easy to find out in R that our approximation of the correlation between β_0 and β_1 is -0.7592774.

(d) & (e)

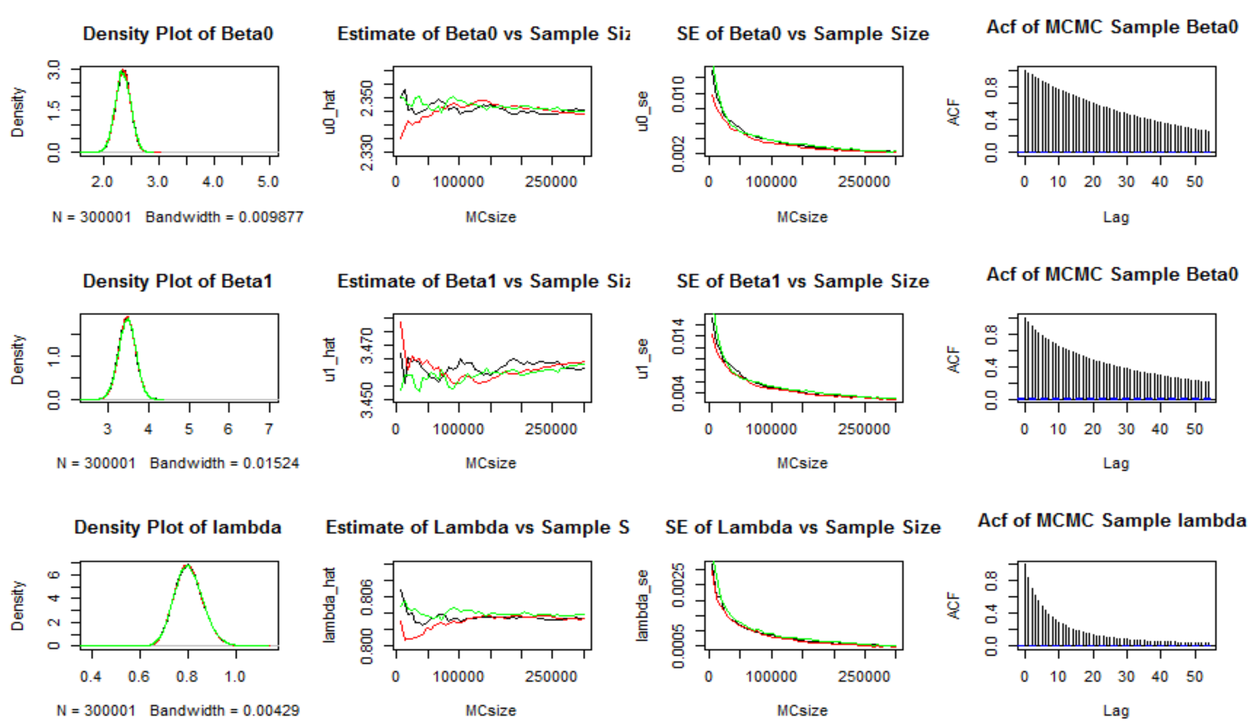

Fig.2. Smoothed Density Plot of Parameters & Supporting Plots for the Accuracy of Our Approximations for Data #2

The above 1st, 5th, 9th plots are about approximated density of β_0, β_1 and λ. Each of them consists of three approximated density from three different initial values of β_0, β_1 and λ: $\beta_{00} = 5, \beta_{10} = 7$ and $\lambda_0 = 0.4$; $\beta_{00} = 3, \beta_{10} = 4$ and $\lambda_0 = 0.3$; $\beta_{00} = 1, \beta_{10} = 1$ and $\lambda_0 = 0.2$. Color black, red and green in each plot are respectively for three sets of initial values. For each plot, we can see that approximated density from three different initial values are almost the same, which means these approximated density of three parameters are not affected by different initial values. So our approximation is accurate and stable.

The above 2nd, 3rd, 6th, 7th, 10th and 11th plots are about estimated means and standard errors of three parameters. Each of them also consists of estimates from three different initial values stated before. We set sample size as 300000. From above plots we can see that our estimates are very stable after 300000 draws, their standard error are all smaller than 0.01. Estimates from different initial values convergent to the same place. Although autocorrelation plot 4th and 8th are not that good, they are still results after carefully selection of variances of normal proposals of β_0 and β_1. Autocorrelation plot 12th of λ is good. The estimate effective sample sizes for different initial values are all larger than 5000, which is pretty good. The information above can show that my approximation is very good and accurate.

3(a)

By implementing same algorithm as 2(a) to Dataset#3 in R, we get the following results:

	Posterior Means and SEs	95% Credible Intervals
β_0	0.1547138 (0.006130513)	(-0.1783731 0.4578731)
β_1	2.466491 (0.009827102)	(1.936550 3.013678)
λ	0.1613539 (0.0001729936)	(0.1509812 0.1723439)

Table.2. Posterior Means, Standard Errors and 95%Credible Intervals of β_0, β_1 and λ for Data #3

(b) & (c)

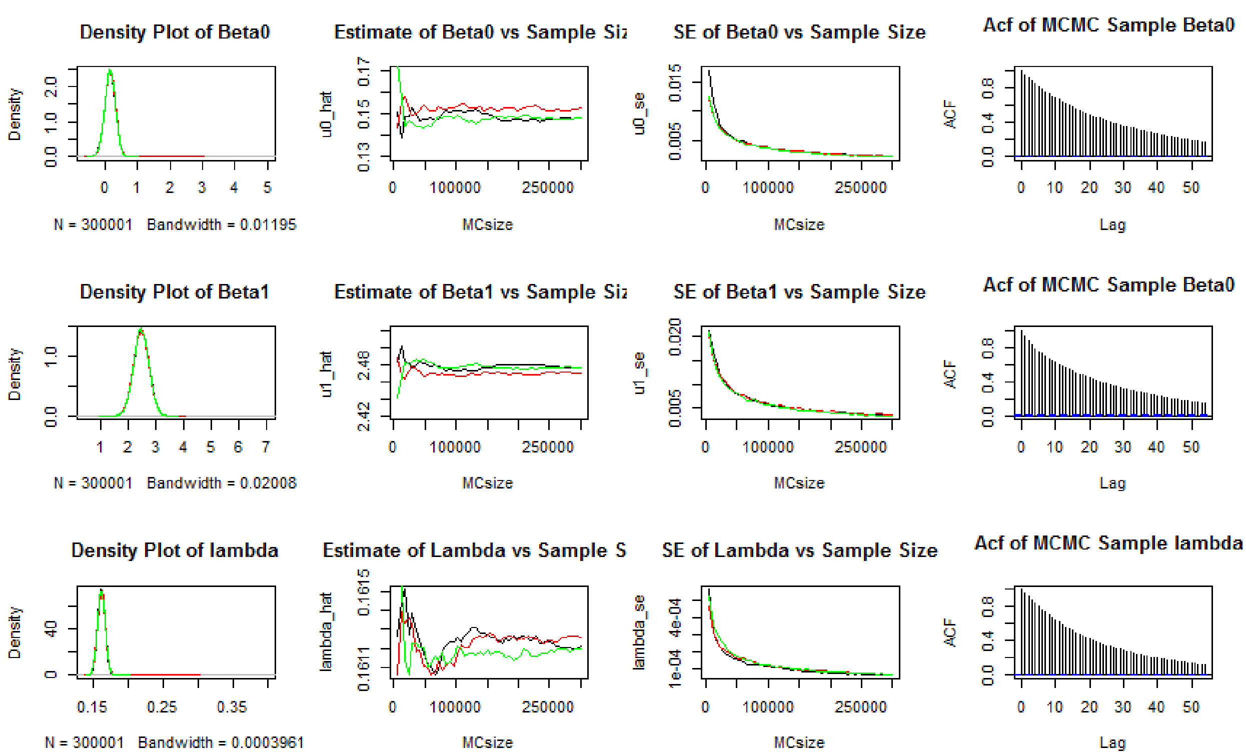

Fig.3. Smoothed Density Plot of Parameters & Supporting Plots for the Accuracy of Our Approximations for Data #3

I didn't modified my MCMC algorithm because my previous algorithm works pretty well for Data #3, which is showed as follows:

The above 1st, 5th, 9th plots are about approximated density of β_0, β_1 and λ. Each of them consists of three approximated density from three different initial values of β_0, β_1 and λ: $\beta_{00} = 5, \beta_{10} = 7$ and $\lambda_0 = 0.4$; $\beta_{00} = 3, \beta_{10} = 4$ and $\lambda_0 = 0.3$; $\beta_{00} = 1, \beta_{10} = 1$ and $\lambda_0 = 0.2$. Color black, red and green in each plot are respectively for three sets of initial values. For each plot, we can see that approximated density from three different initial values are almost the same, which means these approximated density of three parameters are not affected by different initial values. So our approximation is accurate and stable.

The above 2nd, 3rd, 6th, 7th, 10th and 11th plots are about estimated means and standard errors of three parameters. Each of them also consists of estimates from three different initial values stated before. We set sample size as 300000. From above plots we can see that our estimates are very stable after 300000 draws, their standard error are all smaller than 0.005. Estimates from different initial values convergent to the same place. Although autocorrelation plot 4th and 8th are not that good, they are still results after carefully selection of variances of normal proposals of β_0 and β_1. Autocorrelation plot 12th of λ is good. The estimate effective sample sizes for different initial values are all larger than 6000, which is very good. The information above can show that my approximation is very good and accurate.