Stat 515 Take Home Exam

Abhishek Rao

April 29, 2015

1 Problem 1

Theory

We need to find

$$\pi(\beta_1|\mathbf{Y}, X) = \frac{f_j(Y, X, \beta_1)}{f_Y X(Y, X)} \propto f_j(Y, X, \beta_1)$$

Where $f_Y X())$ is the joint distribution of Y,X and beta1 and $f_j()$ be the distribution of Y,X. Let $f_j()$ be proportional to $h(\beta_1) \propto f_{Y|X,\beta_1}(Y|X, \beta_1) f_{prior}(\beta_1) = EMG(\beta_0 + \beta_1 X, \sigma_i, \lambda) N(0, 10)$ Let us use log scale to make calculations for likelihood easier, Let $h_l(\beta_1) = log(h(\beta_1))$ one can use the provided emg density function and get posterior distribution. $h_l(\beta_1) = \sum_{i=1}^{n} dexpgauss(Y_i, \mu = 5 + \beta_1 X_i, \sigma = 1, \lambda = 0.4, log = TRUE) + dnorm(\beta_1, mean = 0, sd = 10, log = TRUE)$

1.1 Part a: Algorithm

The Algorithm is given by

1. Choose starting values, for our case arbitrarily chosen as -10, 0, $10 = startvalue$. Repeat the below for each.
2. Let the length of MCMC chain, chainlength be $= 3.3123 \times 10^4$
3. Create an empty vector called chain with length of chainlength. Set $Chain(1) = startvalue$ obtained in step 1.
4. choose Proposal function, we have chosen Normal function with different variance $= \tau = 1, 2$
5. Prior distribution is Normal with mean 0 and variance 20.
6. Set Likelihood function. This is given by $EMG(Y_i, \mu = 5 + \beta_1 * Xi, \sigma = 1, \lambda = 0.4)$ where Y_i and X_i are given in the data.
7. Let $h_l(\beta_1)$ be the log of posterior distribution calculated. For our case it is sum of loglikelihood for each point + log prior.
8. Repeat for chainlength, let the current value of chain be $chain[i]$
   a) Sample y from the proposal
   b) Calculate $p = exp(h_l(y) - h_l(chain[i]))$. Also take a sample $U \sim Unfirom(0, 1)$
   c) if $U < p$ then $chain[i + 1] = y$. Else $chain[i + 1] = chain[i]$

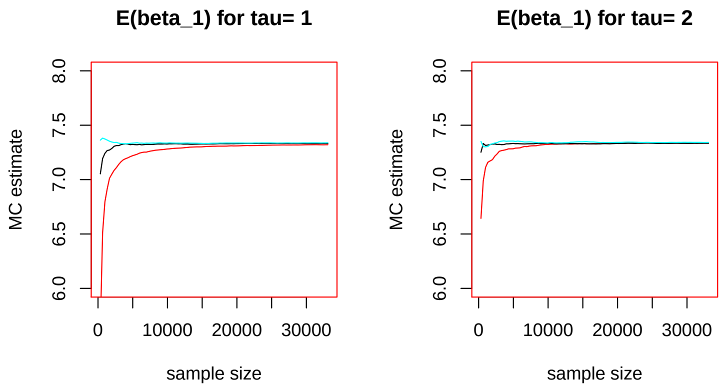

Figure 1: Plot of Estimate vs iteration number

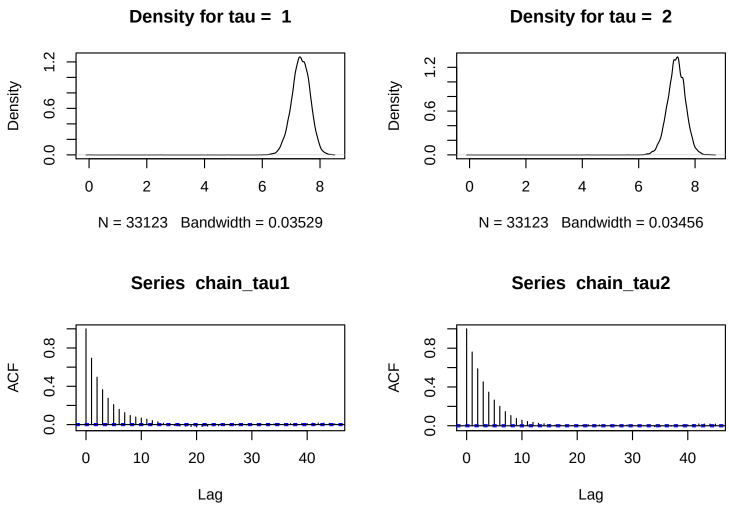

Figure 2: Density and ACF plots for Problem1

1.2 Estimates

The provided batchmeans code is used to find the value of estimates and MCSE for all 3 problems. The Output of the code gave Estimates of 7.3322 and 7.3343 and MCSE of 0.0046 and 0.0048 Respectievely for $\tau = 1, 2$ and sample size of 3.3123×10^4

1.3 Credible Interval

The 95pc credible interval is given by $(6.698, 7.9211)$ and $(6.7017, 7.9266)$ respectievely for $\tau = 1, 2$

1.4 Density

The Density of posterior distribution of β_0 is given in Figure 2. This was plotted using the default density() function in R.

1.5 Accuracy

If you observe the Estimate vs number of iterations in Figure1, they all seem to converge to the same value in all 6 cases. One can also check the autocorrelation plots in Figure2. They die down quickly. The acceptance rate is given by 0.3573 and 0.1867 respectievely for $\tau = 1, 2$. The first acceptance rate is good value for 1D distribution. All these along with low MCSE given in 1.2 gives strong indication that the Estimates are accurate.

2 Problem 2

2.1 Part a: Algorithm

Theory

Similar to previous problem, we get the posterior function as sum of log likelihood and prior. Prior is given by
$Prior(\beta_0, \beta_1, \lambda) = dnorm(\beta_0, mean = 0, sd = 10, log = TRUE) + dnorm(\beta_1, mean = 0, sd = 10, log = TRUE) + dgamma(lambda, shape = 0.001, scale = 100, log = TRUE)$
Let LL denote LogLikelihood, then
$LL(Y_i, X_i, \beta_0, \beta_1, \lambda) = \sum_{i=1}^{n} dexpgauss(Y_i, \mu = \beta_0 + \beta_1 * Xi, \sigma = 1, \lambda = \lambda, log = TRUE)$

$$h_l(\beta_0, \beta_1, \lambda) = LL(Y_i, X_i, \beta_0, \beta_1, \lambda) + Prior(\beta_0, \beta_1, \lambda)$$

Algorithm

We will use All at once MH algorithm. The Algorithm is given by

1. Choose 3 sets of starting values. Choose starting values as draws from prior. i.e. $\beta_0 \sim N(0, 10), \beta_1 \sim N(0, 10), \lambda \sim Gamma(0.01, 100)$.

2. Let the length of MCMC chain, chainlength be $= 3.3123 \times 10^4$

3. Create an empty vector called chain with dimension 3xchainlength. Set the first value from step 1

4. Choose proposal function $y \sim q(y|x)$ where x is the curent value in the chain and y is the proposed new value . This is independently done, same way for all parameter but with different variance. We have chosen Normal function with a base variance $= \tau = 0.3$. Throught trial and error the variance for proposal function for $(\beta_0, \beta_1, \lambda)$ was scaled by factor of $(1, 2, 0.3333)$. This was done by trial and error by keeping in check the different autocorrelation, acceptance and convergence of estimate.

5. Log prior functions are created. We add prior of each parameter to get the total prior.
$Prior_{log} = dnorm(\beta_0, mean = 0, sd = 10, log = TRUE) + dnorm(\beta_1, mean = 0, sd = 10, log = TRUE) + dgamma(\lambda, shape = 0.01, scale = 100)$.

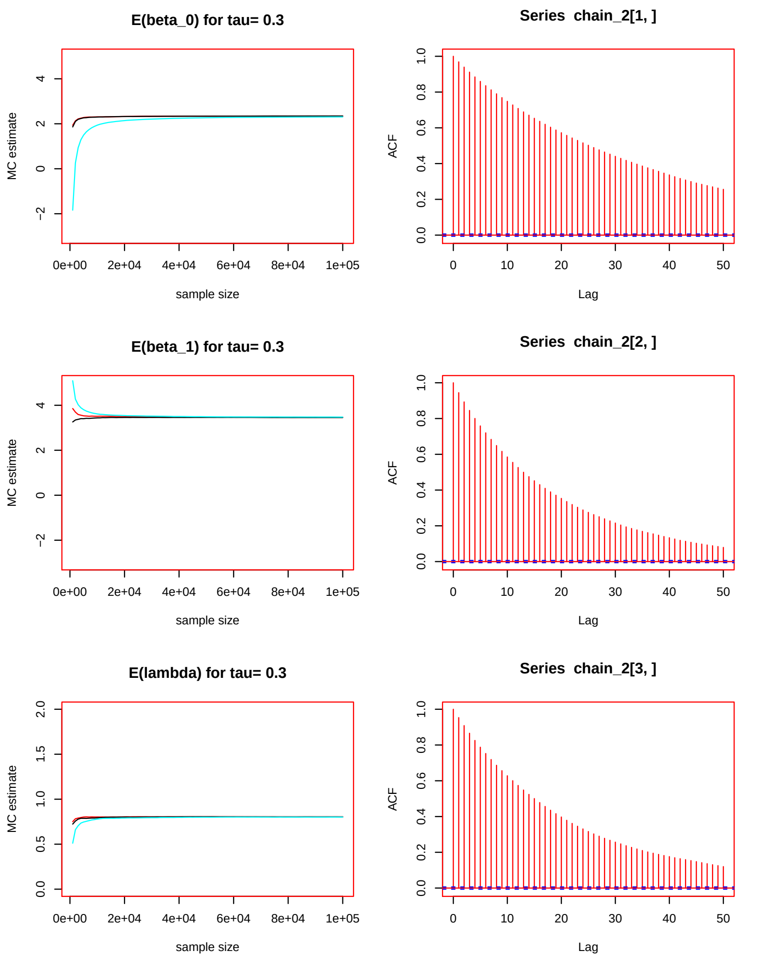

Figure 3: Estimates vs iterations and their corresponding autocorrelation for Problem 2

6. Set Likelihood function. This is given by $EMG(Y_i, mu = \beta_0 + \beta_1 \times Xi, \sigma = 1, \lambda)$ where Y_i and X_i are given in the data.
Note: If $\lambda < 0$ then then we need to rejct the sample, so we will return a very large negative value (-10^{9999}). So if we ever get a negative value it will be a reject update.

7. Let $h_l(\beta_0, \beta_1, \lambda)$ be the log of posterior distribution calculated. For our case it is sum of loglikelihood for each point + log prior.

8. Repeat for chainlength, let the current value of chain be $chain[i]$
   a) Sample 3 independent *y1,y2,y3*s from the proposal.
   b) Calculate $p = exp(h_l(y1, y2, y3) - h_l(chain[i]))$. Also take a sample $U \sim Unfirom(0, 1)$
   c) if $U < p$ then $chain[i+1] = y$. Else $chain[i+1] = chain[i]$

2.2 Estimates

Estimates are given in Table 1. Here CI stands for Credible interval. The number of samples is 10^5 and τ was found through trial and error to be 0.3.

2.3 correlation

The correlation between β_0, β_1 is given by -0.1811

Parameter	β_0	β_1	λ
Samples	10^5	10^5	10^5
Estimate	2.3446	3.4533	0.8031
MCSE	0.0054	0.0045	0.0013
CI start	(2.079)	(3.023)	(0.6951)
CI end	(2.609)	(3.879)	(0.9267)

Table 1: Estimates for Problem 2

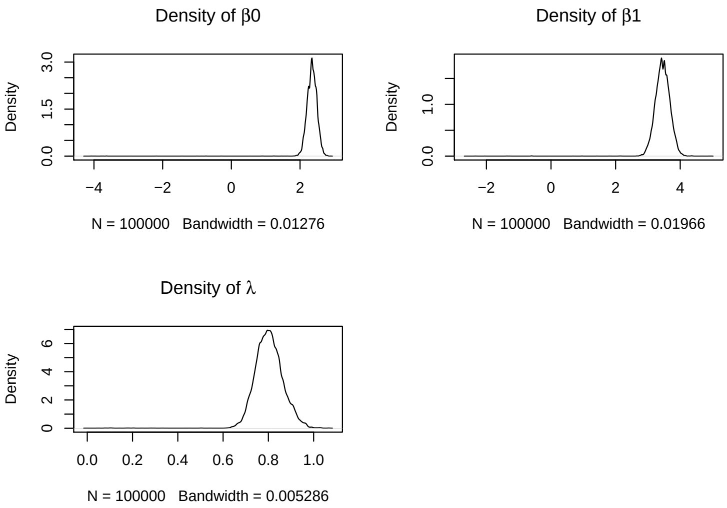

Figure 4: Density plots for Problem 2

2.4 Densities

The marginal densities of $\beta_0, \beta_1, \lambda$ are given in Figure 4

2.5 Reliability

Reliability seems to be good based on the observation that in Figure 3 the values seem to be converging. The autocorrelation seems to be steadily decreasing which is a good sign.. The MCSE errors in Table 1 are low. The acceptance rate is given by 0.0751 for $\tau = 0.3$ seems acceptable. The number of accepted samples given by $acceptanceRate * sampleSize =$ 7507.

3 Problem 3

Trying the same code on the dataset EMG3. We notice that the data, density(y) has a bimodal distribution. Thus we try to give more variance to β_0 as this seems to be bimodally distributed.

3.1 Estimates

Estimates are given in Table 2. Here CI stands for Credible interval. The number of samples is 10^5 and τ was found through trial and error to be 0.08. Note the batchmeans code was used that was provided to find estimate value and MCSE.

3.2 Densities

The marginal densities of $\beta_0, \beta_1, \lambda$ are given in Figure 4

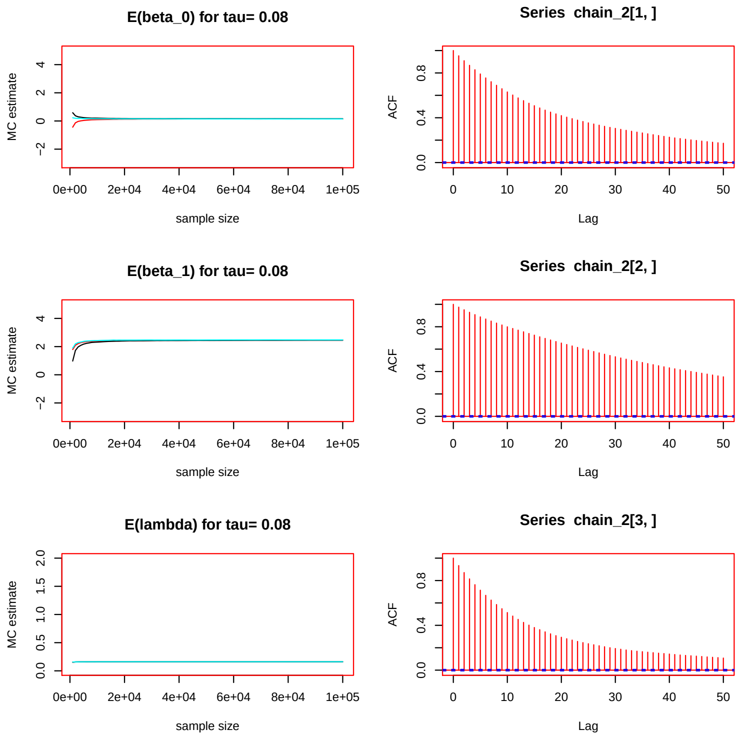

Figure 5: Plot of Estimate vs iteration number and the corresponding autocorrelation for Problem 3

Parameter	β_0	β_1	λ
Samples	10^5	10^5	10^5
Estimate	0.1572	2.4546	0.1611
MCSE	0.0049	0.0152	1.1264×10^{-4}
CI start	(-0.1653)	(1.926)	(0.1504)
CI end	(0.4668)	(3.009)	(0.1723)

Table 2: Estimates for Problem 3

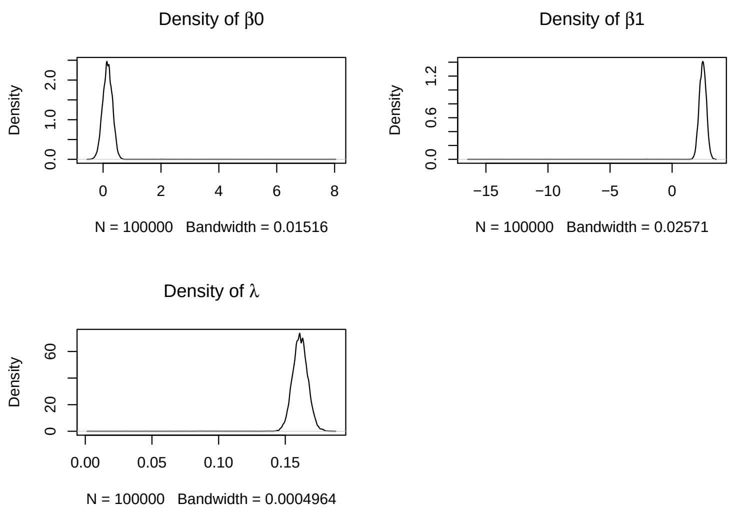

Figure 6: Density plots for Problem 3

3.3 Results and Changes

Realiability

Reliability seems to be good based on the observation that in Figure 5 the values seem to be converging. The MCSE errors in Table 2 are low. The acceptance rate is given by 0.0971 for $\tau = 0.08$ seems acceptable. The number of accepted samples given by $acceptanceRate * sampleSize = 9707$

Changes

The proposal function was changed and variance scaling for each proposal function was changed to $(4, 8, 0.125)$.