For this assignment, we used code from: http://sites.stat.psu.edu/~mharan/MCMCtut/MCMCchpt.R

Problem 1

For this problem, we have β_0=5, λ=.4, σ_i=1 for all i, and n=24000. Let $f(x; \mu, \sigma, \lambda)$ be the pdf of a random variable with distribution $EMG(\mu, \sigma, \lambda)$.
The prior for β_1 has distribution $N(0, 10^2)$, and we denote its density by $\pi(\beta_1)$. We obtain the joint posterior density below (up to a normalizing constant).
Since β_1 is the only random parameter, this is also the full conditional density of β_1 (up to a normalizing constant).

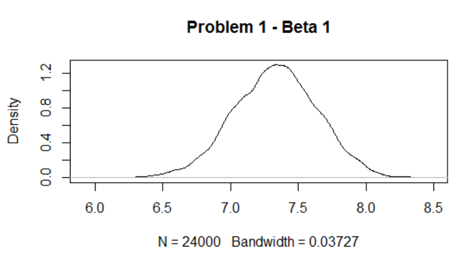
$$\pi(\beta_1|\boldsymbol{Y}) \propto L(\boldsymbol{Y}|\beta_1) * \pi(\beta_1), \qquad where\ L(\boldsymbol{Y}|\beta_1) = \left(\prod_{i=1}^{n} f(Y_i; \beta_0 + \beta_1 X_i, \sigma_i, \lambda)\right)$$

We apply the following Metropolis-Hastings algorithm:

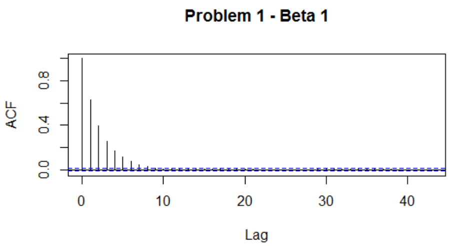
1. Pick the starting value for the Markov Chain: $\beta_1^1 = 7.3$. We chose this value based on several trial MCMC runs.
2. Pick a new value β_1^{i+1} based on the previous value β_1^i.
    a. Propose a candidate β_1^* from the distribution $N(\beta_1^i, {\tau_1}^2)$. Based on several trial MCMC runs, we chose $\tau_1 = .9$.
    b. Since the proposal distribution in Step 2(a) is symmetric, we have a Metropolis update. We compute the following acceptance probability (in log scale):
    $$\alpha(\beta_1^i, \beta_1^*) = \min\left(\frac{\pi(\beta_1^*|\boldsymbol{Y})}{\pi(\beta_1^i|\boldsymbol{Y})}, 1\right)$$
    c. With probability $\alpha(\beta_1^i, \beta_1^*)$, accept the candidate β_1^*, and set $\beta_1^{i+1} = \beta_1^*$. Otherwise, reject the candidate, and set $\beta_1^{i+1} = \beta_1^i$.
3. Repeat Step 2 (n-1) times to produce a Markov Chain of length n.

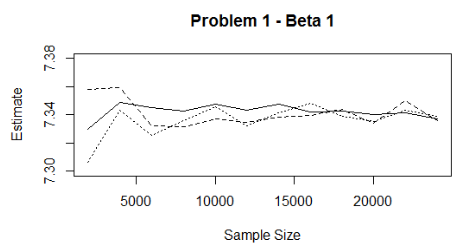
We ran the MCMC algorithm for 24000 iterations, and the acceptance rate for β_1 proposals was 39%.
Our estimate of the posterior expectation of β_1 was 7.333, and the associated MCMC standard error was .005.
We also obtain a 95% credible interval for β_1 based on the 2.5th and 97.5th percentiles of the samples, and have $\beta_1 \in (6.715, 7.934)$.

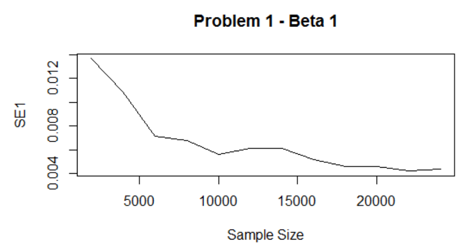
Below, we plot a smoothed estimate of the posterior pdf of β_1.

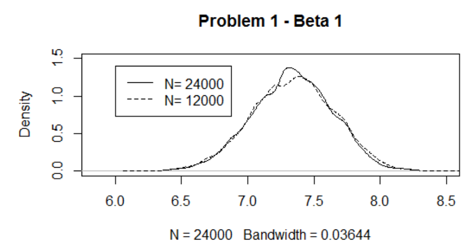


We verify the accuracy of the MCMC approximation based on several factors. First, we have an effective sample size of 5355, which is greater than 5000. Next, the plot of the autocorrelation function looks reasonable. Third, for different starting values ((5, 7.3, 10) for β_1), the estimates still converge as run size increases. Fourth, the standard error of the estimate declines as run size increases. Finally, the density plots at full run size and half run size look similar.









Problem 2

For this problem, we have σ_i=1 for all i, and n=70000. Let $f(x;\mu,\sigma,\lambda)$ be the pdf of a random variable with distribution $EMG(\mu,\sigma,\lambda)$.
The prior for β_0 has distribution $N(0,10^2)$. The prior for β_1 has distribution $N(0,10^2)$. The prior for λ has distribution $Gamma(.01,100)$; that is, the mean is .01(100)=1, and the variance is .01(100^2)=100. We denote these prior densities by $\pi(\beta_0)$, $\pi(\beta_1)$, and $\pi(\lambda)$.
We obtain the joint posterior density below (up to a normalizing constant).

$$\pi(\beta_0,\beta_1,\lambda|\boldsymbol{Y}) \propto L(\boldsymbol{Y}|\beta_0,\beta_1,\lambda) * \pi(\beta_0) * \pi(\beta_1) * \pi(\lambda), \qquad where\ L(\boldsymbol{Y}|\beta_0,\beta_1,\lambda) = \left(\prod_{i=1}^{n} f(Y_i;\beta_0+\beta_1 X_i,\sigma_i,\lambda)\right)$$

We obtain the following full conditional densities (up to a normalizing constant).

$$\pi(\beta_0,\beta_1|\boldsymbol{Y},\lambda) \propto L(\boldsymbol{Y}|\beta_0,\beta_1,\lambda) * \pi(\beta_0) * \pi(\beta_1), \qquad where\ L(\boldsymbol{Y}|\beta_0,\beta_1,\lambda) = \left(\prod_{i=1}^{n} f(Y_i;\beta_0+\beta_1 X_i,\sigma_i,\lambda)\right)$$

$$\pi(\lambda|\boldsymbol{Y},\beta_0,\beta_1) \propto L(\boldsymbol{Y}|\beta_0,\beta_1,\lambda) * \pi(\lambda), \qquad where\ L(\boldsymbol{Y}|\beta_0,\beta_1,\lambda) = \left(\prod_{i=1}^{n} f(Y_i;\beta_0+\beta_1 X_i,\sigma_i,\lambda)\right)$$

We apply the following Metropolis-Hastings algorithm:

1. Pick the starting values for the Markov Chain: $(\beta_0^1,\beta_1^1,\lambda^1) = (2.4,3.5,.8)$. We chose these values based on several trial MCMC runs.
2. Pick new values $(\beta_0^{i+1},\beta_1^{i+1},\lambda^{i+1})$ based on the previous values $(\beta_0^i,\beta_1^i,\lambda^i)$.
   a. Pick $(\beta_0^{i+1},\beta_1^{i+1})$ based on $(\beta_0^i,\beta_1^i,\lambda^i)$.
      i. Propose a candidate (β_0^*,β_1^*) from the bivariate Normal distribution with mean (β_0^i,β_1^i) and covariance matrix $\Sigma = \begin{bmatrix} {\tau_0}^2 & \rho\tau_0\tau_1 \\ \rho\tau_0\tau_1 & {\tau_1}^2 \end{bmatrix}$.
      Based on many trial MCMC runs, we chose $\tau_0 = .2$, $\tau_1 = .25$, and $\rho = -.8$.
      ii. Since the proposal distribution in Step 2(a)(i) is symmetric, we have a Metropolis update. We compute the following acceptance probability (in log scale):

$$\alpha\left((\beta_0^i,\beta_1^i),(\beta_0^*,\beta_1^*)\right) = \min\left(\frac{\pi\left(\beta_0^*,\beta_1^*|\boldsymbol{Y},\lambda^i\right)}{\pi\left(\beta_0^i,\beta_1^i|\boldsymbol{Y},\lambda^i\right)},1\right)$$

      iii. With probability $\alpha\left((\beta_0^i,\beta_1^i),(\beta_0^*,\beta_1^*)\right)$, accept the candidate (β_0^*,β_1^*), and set $(\beta_0^{i+1},\beta_1^{i+1}) = (\beta_0^*,\beta_1^*)$.
      Otherwise, reject the candidate, and set $(\beta_0^{i+1},\beta_1^{i+1}) = (\beta_0^i,\beta_1^i)$.

b. Pick λ^{i+1} based on $(\beta_0^{i+1}, \beta_1^{i+1}, \lambda^i)$.

i. Propose a candidate λ^* from the distribution $Gamma\left(\frac{\lambda^i}{\tau_\lambda}, \tau_\lambda\right)$. That is, the mean is λ^i, and the variance is $\lambda^i \tau_\lambda$.

We denote this proposal density as $q(\lambda^*|\lambda^i)$. Based on many trial MCMC runs, we chose $\tau_\lambda = .02$.

ii. This is a Metropolis-Hastings update, and we compute the following acceptance probability (in log scale):

$$\alpha(\lambda^i, \lambda^*) = \min\left(\frac{\pi(\lambda^*|\boldsymbol{Y}, \beta_0^{i+1}, \beta_1^{i+1})q(\lambda^i|\lambda^*)}{\pi(\lambda^i|\boldsymbol{Y}, \beta_0^{i+1}, \beta_1^{i+1})q(\lambda^*|\lambda^i)}, 1\right)$$

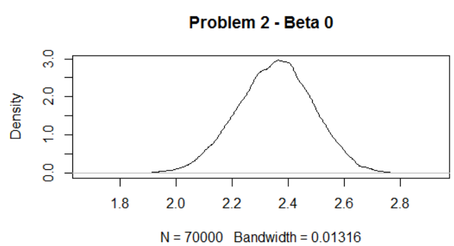
iii. With probability $\alpha(\lambda^i, \lambda^*)$, accept the candidate λ^*, and set $\lambda^{i+1} = \lambda^*$. Otherwise, reject the candidate, and set $\lambda^{i+1} = \lambda^i$.

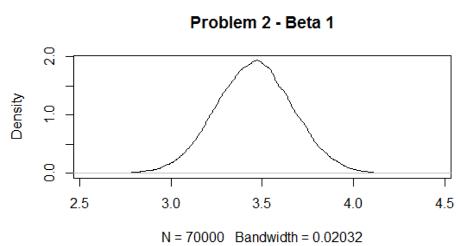
3. Repeat Step 2 (n-1) times to produce a Markov Chain of length n.

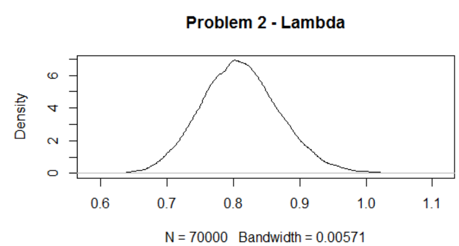
We ran the MCMC algorithm for 70000 iterations. The acceptance rate was 38% both for (β_0, β_1) proposals and for λ proposals.
In the table below, we report the posterior means, MCMC standard errors, and posterior 95% credible intervals.
We approximate the correlation between β_0 and β_1 to be -.78.

Parameter	Estimate	MCMC Standard Error	(.025,.975) Sample Quantiles
β_0	2.3532	.0016	(2.0812,2.6127)
β_1	3.4577	.0027	(3.0464,3.8740)
λ	.8108	.0009	(.6990,.9342)

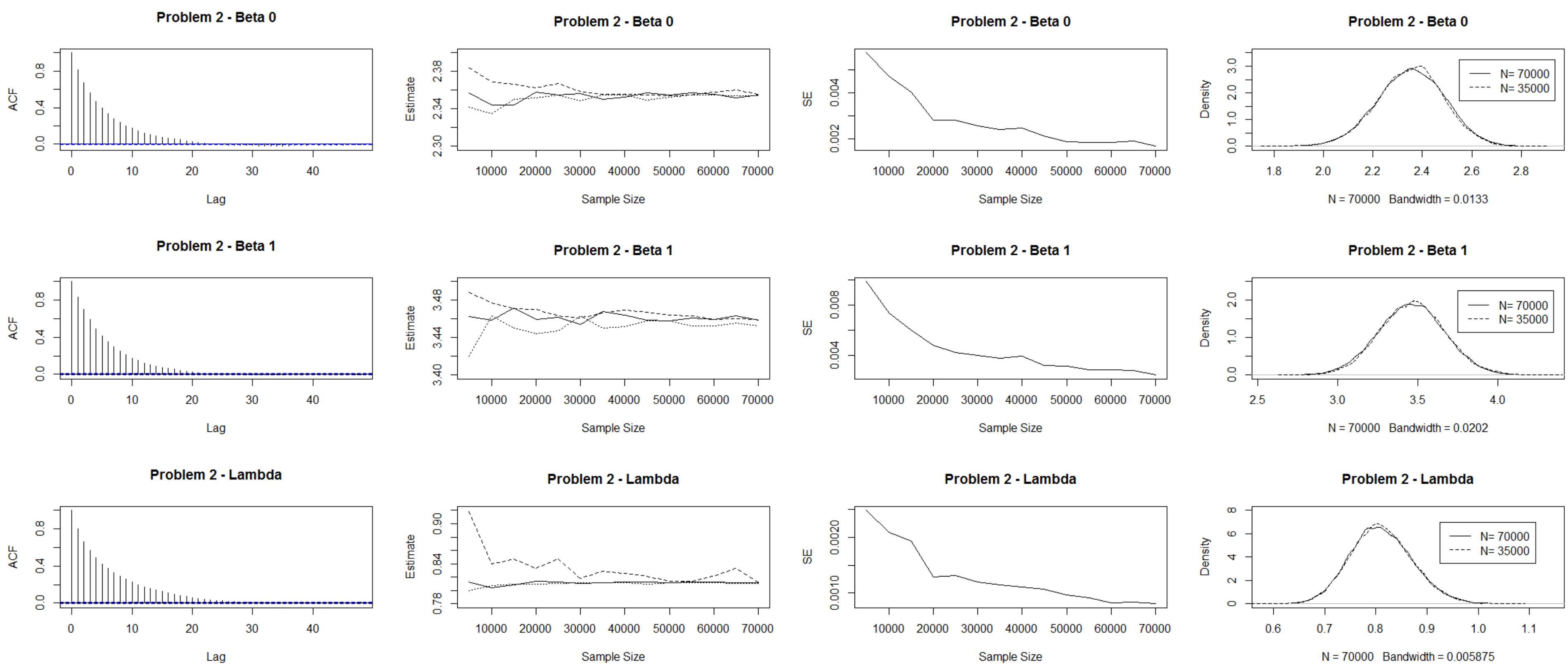
Below, we provide approximate density plots for the marginal distributions of β_0, β_1, and λ.







We verify the accuracy of the MCMC approximation based on several factors. First, the effective sample sizes for β_0, β_1, λ are 6407, 6230, 5478 respectively, which are greater than 5000. Next, the autocorrelation function plots look reasonable. Third, for different starting values ((0, 2.4, 5) for β_0, (1, 3.5, 6) for β_1, and (.3, .8, 1.3) for λ), the estimates still converge as run size increases. Fourth, all the standard errors decline as run size increases. Finally, the density plots at full run size and half run size look similar.



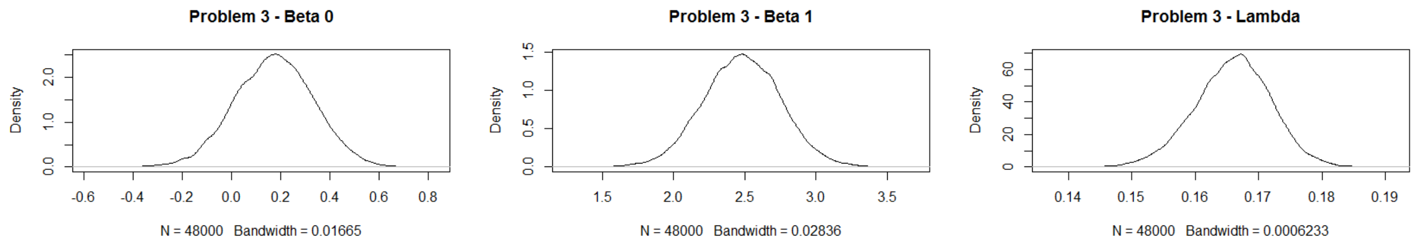
Problem 3

For this problem, we used the same MCMC algorithm as we did in Problem 2, but modified the inputs based on trial runs.
We now have n=48000. For the initial values, we have: $(\beta_0^1, \beta_1^1, \lambda^1) = (.15, 2.5, .16)$. For the tuning parameters, we have: $(\tau_0, \tau_1, \rho, \tau_\lambda) = (.22, .38, -.8, .001)$.

We ran the MCMC algorithm for 48000 iterations. The acceptance rate was 40% for (β_0, β_1) proposals and 44% for λ proposals.
In the table below, we report the posterior means, MCMC standard errors, and posterior 95% credible intervals.
We approximate the correlation between β_0 and β_1 to be -.83.

Parameter	Estimate	MCMC Standard Error	(.025,.975) Sample Quantiles
β_0	.17250	.00218	(-.14118,.48037)
β_1	2.48000	.00377	(1.94696,3.01322)
λ	.16594	.00007	(.15347,.17740)

Below, we provide approximate density plots for the marginal distributions of β_0, β_1, and λ.



We verify the accuracy of the MCMC approximation based on several factors. First, the effective sample sizes for β_0, β_1, λ are 5411, 5436, 7973 respectively, which are greater than 5000. Next, the autocorrelation function plots look reasonable. Third, for different starting values ((-1, .15, 1) for β_0, (0, 2.5, 5) for β_1, and (.05, .16, .3) for λ), the estimates still converge as run size increases. Fourth, all the standard errors decline as run size increases. Finally, the density plots at full run size and half run size look similar.