STAT 515. Take-home Exam. Spring 2015

Geng-Yuan Chen, 9-2637-4914

April 29, 2015

Problem 1. $Y_i \sim EMG(\beta_0 + \beta_1 X, \sigma_i, \lambda)$: $\beta_0 = 5, \lambda = 0.4, \sigma_i = 1, P(\beta_1) \sim N(\mu = 0, \sigma = 10)$

(a) Algorithm and Pseudocode:

1. Draw $\beta_1 \tilde{\sim} N(\mu, \sigma)$. $N(\cdot)$ is approximated from previous trials to make β_1 less bias.
2. Calculate the prior probability, $P(\beta_1)$, at this $\beta_1^{(0)}, \beta_1^{(0)} \sim N(0, 10)$
3. Calculate the likelihood function, $L(\beta_1^{(0)} \mid Y, X) = \prod f(\beta_1^{(0)} \mid Y, X)$, where $f(\cdot) \sim EMG(\cdot)$
4. Calculate the posterior distribution of initial guess. $log(\pi) \sim log(L(\beta_1^{(0)} \mid Y, X)) + log(P(\beta_1^{(0)}))$
5. Metropolis-Hasting Draw with random walk proposal (written in: MetropolisHastingDrawer(N, τ, IC)). N: realisations; IC: initial condition
   a. Generate a uniform random variable, J, with length N (to Judge accept or reject).
   b. Carry out the i^{th} draw, $\beta_1^{(i)}$, from the proposal, $q \sim N(\beta_1^{(i)}, \tau)$. In this case, $\tau = 0.3$.
   c. Calculate the posterior distribution of the i^{th} draw. $log(\pi^*) \sim log(L(\beta_1^{(i)} \mid Y, X)) + log(P(\beta_1^{(i)}))$, L: likelihood; P: prior distribution
   d. Set probability to accept this draw, $Prob\{accept\} = min\{\frac{\pi^*}{\pi^{(i-1)}}, 1\}$
   e. Decision: if $Prob\{accept\} > J$, then accept this draw; otherwise reject by assigning this draw as the old one.
   f. repeat step b-e until N realisations are done.
6. Estimate mean and its MCMCse via the functions, estvssamp, mcsevssamp, and bm.
7. Plot the residuals of the estimations. Compare this with $W \sim EMG(0, \sigma, \lambda)$
8. Use other parameters to check convergence, e.g. initial conditions, random seeds, proposal variance, ...etc.

(b-c) $E(\beta_1)$, MCMCse, and 95% confidence intervals of different trials

Group	I	II	III	IV	V
Seed	123456	**654321**	123456	123456	123456
IC	10	10	**5**	10	$\hat{\pi}$
τ	1	1	1	**10**	0.3
N	5000	5000	5000	5000	50000
$E(\beta_1)(s.e.)$	7.359 (0.009)	7.358 (0.009)	7.337 (0.011)	7.366 (0.023)	7.344 (0.004)
$95\%CI$	[6.733, 7.948]	[6.720, 7.980]	[6.694, 7.949]	[6.696, 7.893]	[6.739, 7.928]
ESS	1109	1088	1124	152	6276

These simulation results do not rely on those control variables (in bold), except for the tuning parameter, τ. To reduce the bias, the submitted code is just for group V, where the initial condition is drawn from the approximated stationary distribution, $\hat{\pi}$, and the sample size is assigned as 50000 to ensure ESS is greater than 5000. In this case, β_1 is assumed as a normal distribution with parameters $\mu = 7.3, \sigma = 0.3$.

(d) Density of $\hat{\beta}_1$ and its fitting results

The density of β_1 is shown in Figure 1a, and the estimated regression is carried out in Figure 1b. From regression, these 100 estimated residuals are collected (Figure 1c), and the density of which is compared with the simulated noise, $W_{100\times1} \sim EMG(\cdot)$. The match between 1c and 1d provides the first evidence of viability of this fitting result (although I did not consider the bias issue of residuals in this case).

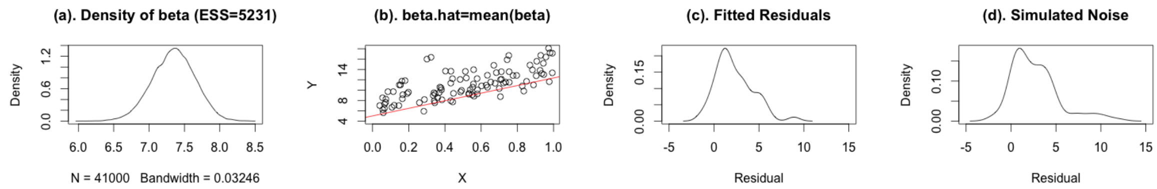

Figure 1: Result of $\hat{\beta}_1$

(e) Verification of this simulation
As shown the estimate of β_1 converges at 7.36 (Figure 2a), and its MCMCse converges to 0 as sample size grows (Figure 2b). The standard deviation of β_1 converges at 0.30 (Figure 2c), while the ACF plot show that the memory last about 15 lags (Figure 2d). Based on these evidences, this MCMC approximation is accurate.

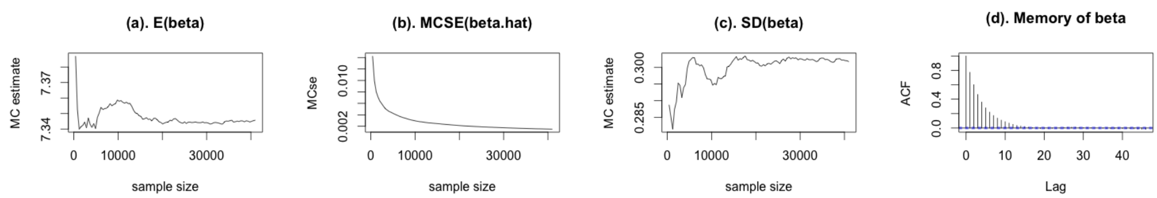

Figure 2: Convergence of $\hat{\beta}_1$

Problem 2. $Y_i \sim EMG(\beta_0 + \beta_1 X, \sigma_i, \lambda) : \sigma_i = 1, P(\beta_0), P(\beta_1) \sim N(\mu = 0, \sigma = 10), P(\lambda) \sim Gamma(0.01, 100)$

(a) Algorithm and Pseudocode: (let $\theta = \{\boldsymbol{\beta_0}, \boldsymbol{\beta_1}, \lambda\}$)

1. Do a pilot study to assign initial condition, $\theta^{\mathbf{0}} : \boldsymbol{\beta_0}, \boldsymbol{\beta_1}, \lambda$ (given in table).
2. Calculate the prior probability at this $\theta_{\mathbf{i}}^{(\mathbf{0})}$
3. Calculate the likelihood function, $L(\theta_{\mathbf{i}}^{(\mathbf{0})} \mid \mathbf{Y}, \mathbf{X}, \theta_{\mathbf{-i}}^{(\mathbf{0})}) = \prod \mathbf{f}(\theta^{(\mathbf{0})} \mid \mathbf{Y}, \mathbf{X}, \theta_{\mathbf{-i}}^{(\mathbf{0})})$, where $f(\cdot) \sim EMG(\cdot)$
4. Calculate the posterior distribution of initial guess. $log(\pi_i^{(0)}) \sim log(L(\theta_{\mathbf{i}}^{(\mathbf{0})} \mid \mathbf{Y}, \mathbf{X}, \theta_{\mathbf{-i}}^{(\mathbf{0})})) + \mathbf{log}(\mathbf{P}(\theta_{\mathbf{i}}^{(\mathbf{0})}))$
5. Metropolis-Hasting Draw with random walk proposal (written in: MetropolisHastingDrawer($N, \tau, \mathbf{IC}$)). N: realisations; IC: initial condition; τ: tuning parameters.
   a. Generate a uniform random variable, $J_{\beta_0}, J_{\beta_1}, J_\lambda$, with length N (to Judge accept or reject).
   b. Update $\beta_0^{(i)}$ by fixing $\beta_1^{(i-1)}$ and $\lambda^{(i-1)}$
   c. Carry out the i^{th} draw, β_0^*, from the proposal, $q_0 \sim N(\beta_0^{(i-1)}, \tau_0)$.
   d. Calculate the posterior distribution of the i^{th} draw. $log(\pi^*) \sim log(L(\beta_0^* \mid Y, X, \beta_1^{(i-1)}, \lambda^{(i-1)})) + log(P(\beta_0^*))$, L: likelihood; P: prior distribution
   e. Set probability to accept this draw, $Prob\{accept\} = min\{e^{log(\pi^*) - log(\pi^{(i-1)})}, 1\}$
   f. Decision: if $Prob\{accept\} > J_0$, then accept this draw; otherwise reject by assigning this draw as the old one.
   g. Repeat step b-f for $\beta_1^{(i)}$ by fixing $\beta_0^{(i)}$ and $\lambda^{(i-1)}$
   h. Repeat step b-f for $\lambda^{(i)}$ by fixing $\beta_0^{(i)}$ and $\beta_1^{(i)}$
   i. Repeat b-h until N realisations are done.
6. Estimate mean and its MCMCse via the functions, estvssamp, mcsevssamp, and bm
7. Use other parameters to check convergence, e.g. initial conditions, random seeds, proposal variance, sample size...etc.

(b) $E(\beta_1)$, MCMCse, and 95% confidence intervals of different trials

Group	I	II	III	IV	V
Seed	123456	**654321**	123456	123456	123456
IC	(4, 3, 1)	(4, 3, 1)	**(2.3, 3.5, 0.8)**	(4, 3, 1)	$\tilde{\pi}^a$
τ	(1, 1, 0.01)	(1, 1, 0.01)	(1, 1, 0.01)	**(10, 10, 0.1)**	(1, 1, 0.01)
N	5000	5000	5000	5000	**500000**
$E(\beta_0)$	2.359 (0.012)	2.384 (0.015)	2.354 (0.009)	2.371 (0.026)	2.346 (0.002)
$95\%CI$	[2.114, 2.620]	[2.174, 2.645]	[2.148, 2.537]	[2.186, 2.614]	[2.114, 2.571]
ESS	107	105	118	86	7085
$E(\beta_1)$	3.438 (0.016)	3.417 (0.017)	3.460 (0.144)	3.450 (0.036)	3.461 (0.002)
$95\%CI$	[3.074, 3.746]	[3.028, 3.737]	[3.139, 3.829]	[2.636, 3.862]	[3.104, 3.816]
ESS	134	130	133	80	9025
$E(\lambda)$	0.809 (0.006)	0.820 (0.008)	0.811 (0.006)	0.814 (0.006)	.803 (0.001)
$95\%CI$	[0.709, 0.930]	[0.695, 0.975]	[0.712, 0.920]	[0.706, 0.945]	[0.698, 0.919]
ESS	89	80	89	89	6111

^{a} : $\beta_0 \sim N(\mu = 2.3, \sigma = 0.12); \beta_1 \sim N(\mu = 3.5, \sigma = 0.18); \lambda \sim N(\mu = 0.79, \sigma = 0.053)$.

(c) Correlation between (β_0, β_1): in Group V setting, $\hat{\rho}(\hat{\beta_0}, \hat{\beta_1}) = -0.703$, with highly negative correlation.

(d) Density of $\hat{\beta_0}, \hat{\beta_1}, \hat{\lambda}$ and the fitting results (Figure 3).

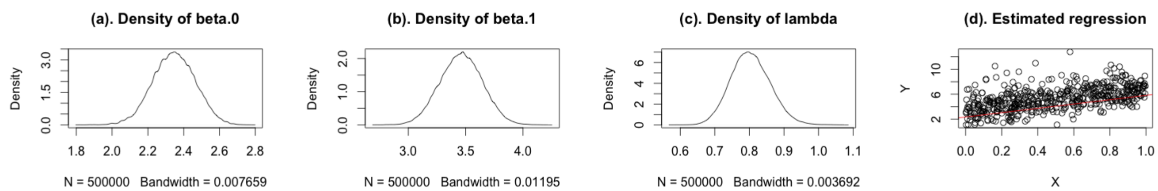

Figure 3: Result of $\hat{\beta_0}$, $\hat{\beta_1}$, and $\hat{\lambda}$

(e) Verification

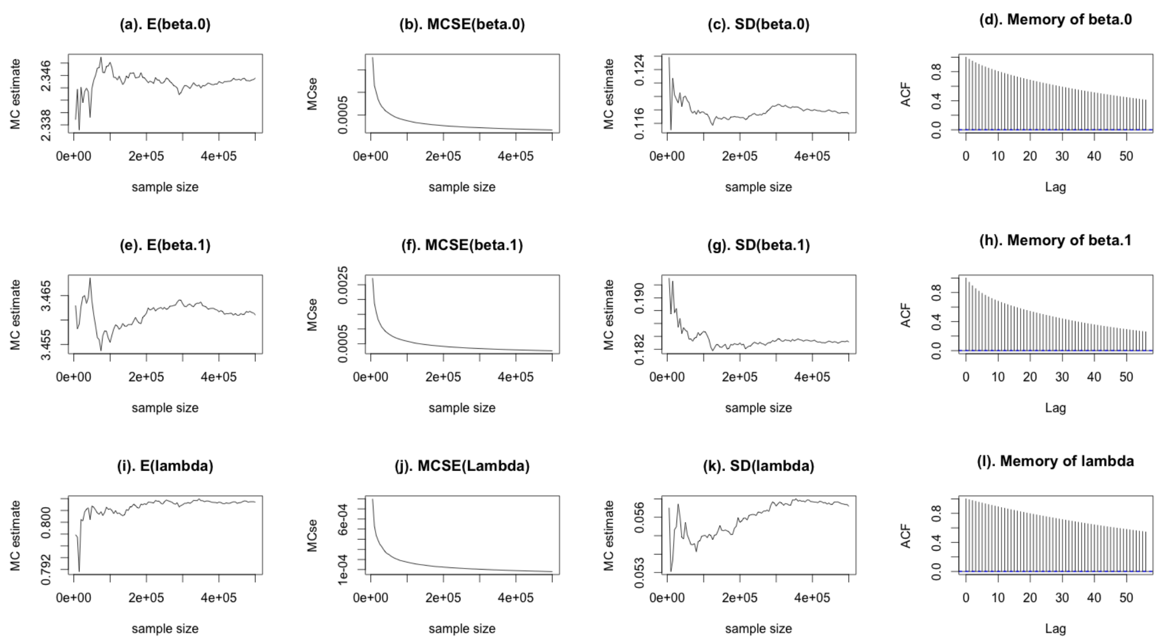

Figure 4: Convergence of $\hat{\theta}$

In this case, because I choose to use random walk proposal for λ, I have to set its tuning parameter small enough to prevent from a negative λ. This leads to a slower decay of ACF. Because the ACFs decays slowly, I have to use N = 500,000 to obtain more than 5,000 ESS (Figure 4). With this sample size, parameters converge as sample size grows. Therefore, it is a reasonable approximation, although not perfect.

Problem 3. Two-population data

(a) $E(\beta_1)$, MCMCse, and 95% confidence intervals of different trials

Group	G^b	$2-GM+G^{ab}$
Seed	123456	123456
IC	π	π
N	500000	500000
τ	(1, 1, 1)	(1, 1, 1)
$E(\beta_0)$	0.149 (0.002)	0.275 (0)
$95\%CI$	[-0.123, 0.420]	[0.275, 0.275]
ESS	9183	4505
$E(\beta_1)$	2.469 (0.003)	2.290 (0.001)
$95\%CI$	[2.006, 2.930]	[1.993, 2.578]
ESS	10625	33501
$E(\lambda)$	0.161 (0.0001)	0.161 (0.0001)
$95\%CI$	[0.151, 0.171]	[0.151, 0.172]
ESS	6127	6186

[a] : Gaussian-Mixture proposal with inter-peak distance 10 for β_0. τ is set as dispersion as each peak.

[b] : Proposal for λ follows Gamma(k,θ). $\tau = 1$ is set as θ

(b) Density of $\hat{\beta}_0, \hat{\beta}_1, \hat{\lambda}$ and the fitting results (Figure 5)

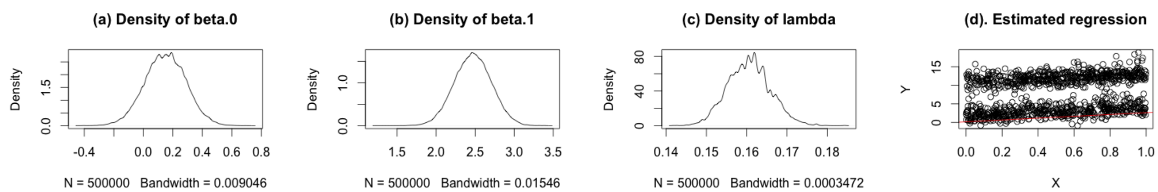

Figure 5: Result of $\hat{\beta}_0$, $\hat{\beta}_1$, and $\hat{\lambda}$

(c) Model comparison

The convergence of parameter estimates, standard deviations, MCMCse, and the decay of auto-correlation were used to contrast with different models. In each test, the τ was optimised by ESS index. From the first few trials, I noticed that lambda is close to 0, which results in inappropriateness of using the random walk proposal. To ensure the positive value of λ, I used the Gamma distribution as proposal for λ^* (the mean is $\lambda\theta$, which is set as centre; θ is set as 1, so it actually is an exponential proposal). Note that θ was tried as 0.5, 2, and 10, but $\theta = 1$ works better (Figure 6). β_0 could have two populations, thus I also try a Gaussian-mixture proposal (with equal weight, with offset 10). Unexpectedly, the 2-Gaussian Mixture model works poorly than conventional random walk (super slow ACF decay) (Figure 7). Therefore, I still sticked to use the conventional Gaussian Random Walk model for β_0 proposal. To summary, my best trial is: q_{β_0}, q_{β_1} $N(\cdot, \tau = 1)$, and $q_\lambda \sim Exp(\lambda)$, as coded, although my strategy may not be good enough.

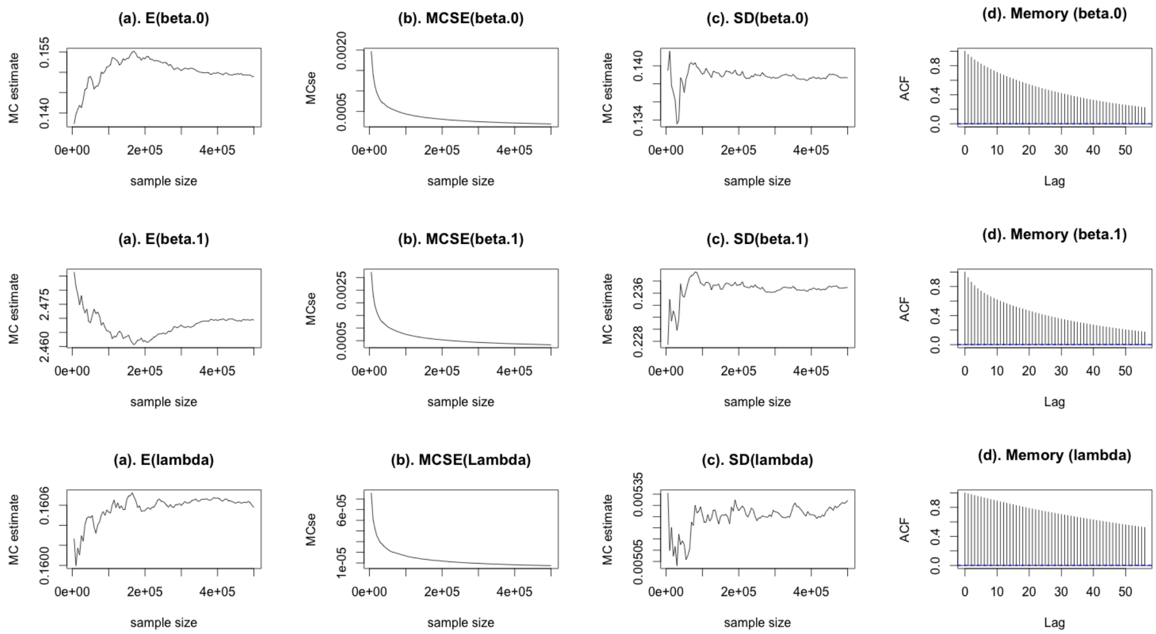

Figure 6: Gaussian Random Walk Proposal

Figure 7: Gaussian-Mixture Proposal