Take Home Final Stat 515

Mauricio Nascimento[a)]

(Dated: 29 April 2015)

1. (a) First let's write the joint distribution, defining $f(y_i|\beta_0, \beta_1, X_i, \sigma_i, \lambda)$ the pdf for each y_i and $\pi(\beta_1)$ the prior for the parameter. We have that the joint distribution is $f(\mathbf{y}, \beta_1) = \pi(\beta_1) \prod_i f(y_i|\beta_0, \beta_1, X_i, \sigma_i, \lambda)$. Since this the distribution include an integral I obted to write the posterior as been proportional to the joint, this way it is possible to use the code at the webpage to calculate the part of the posterior. So the posterior for β_1 is the following.

$$\pi(\beta_1|.) \propto \pi(\beta_1) \prod_i f(y_i|\beta_0, \beta_1, X_i, \sigma_i, \lambda) = h(\beta_1)$$

. Where |. mean given everything else. Since most of the time calculating this quantity is not easy computationally let's use it in log. So

$$log(\pi(\beta_1|.)) \propto log(\pi(\beta_1)) \sum_i log(f(y_i|\beta_0, \beta_1, X_i, \sigma_i, \lambda))$$

This way the algorithm to simulate from the posterior is the following.

i. Define $x[1] = 0$, this value is the prior mean of β_1
ii. Propose $y \sim N(x[t-1], 0.5)$
iii. Calculate the log of the acceptance probability defined as $log(\alpha) = min(0, h(y) - h(x[t-1])$
iv. Sample u from a uniform in $(0, 1)$
v. If $log(u) <= log(alpha)$ define $x[t] = y$, otherwise $x[t] = x[t-1]$, than go back to step (ii) until you have the number of samples that you want.

(b) Let's sample 32000 sample from the posterior distribution, since we are starting the chain in a different point that is not the stationary distribution we are doing a burn in with the first 2000, this remove the part of the chain that has not converged yet.
After this the expected posterior mean and the MCMC standard deviation is the following.

```
## $est
## [1] 7.335632
```

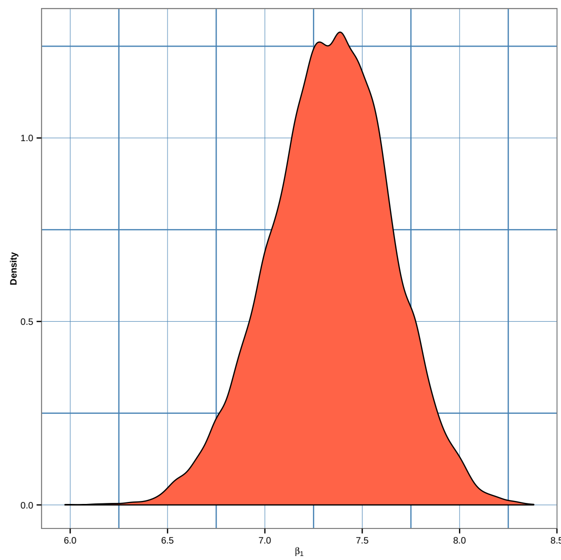

FIG. 1. Posterior Density Question 1

```
##
## $se
## [1] 0.004122128
```

(c) The 95% credible interval is the following.

```
##     2.5%    97.5%
## 6.710689 7.922338
```

(d) The estimated posterior distribution can be found on FIG.1.

(e) To evaluate the accuracy of the algorithm let's first analyse the cumulative mean, the cumulative MCMC standard error and the MCMC when we start in different values.(FIG.2)
We can see in the first plot that doesn't matter the begging state the chain is going converging to the same distribution. The second plot show that the mean increases until it reaches the posterior mean that is close to 7. The last plot shows that the MCMC standard error is decreasing to 0 as the sample size increases, this happens really fast.

[a)] Electronic mail: mfn120@psu.edu.; Also at Statistic Department, Pennsylvania State University.

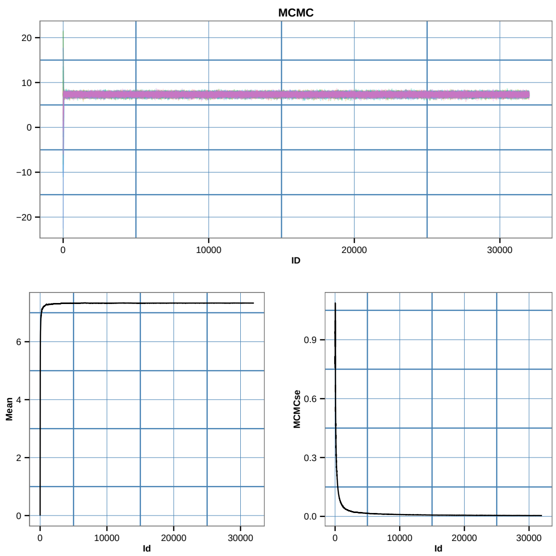

FIG. 2. Chains, Posterior mean and MCMCse by iteration Question 1

2. (a) Let‘s do the same thing as the first question and calculate the joint distribution. Defining $f(y_i|\beta_0, \beta_1, X_i, \sigma_i, \lambda)$ the pdf for each y_i and $\pi(\beta_1)$ the prior for the β_1, $\pi(\beta_0)$ the prior for the β_0, $\pi(\lambda)$ the prior for the λ. We have that the joint distribution is $f(\mathbf{y}, \beta_0, \beta_1, \lambda) = \pi(\beta_0)\pi(\beta_1)\pi(\lambda)\prod_i f(y_i|\beta_0, \beta_1, X_i, \sigma_i, \lambda)$.
Since this the distribution include an integral I opted to write the posterior as been proportional to the joint, this way it is possible to use the code at the webpage to calculate the part of the posterior. So the posterior for β_1 is

$$\pi(\beta_1|.) \propto \pi(\beta_1)\prod_i f(y_i|\beta_0, \beta_1, X_i, \sigma_i, \lambda) = h_2(\beta_1)$$

, β_0 is

$$\pi(\beta_0|.) \propto \pi(\beta_0)\prod_i f(y_i|\beta_0, \beta_1, X_i, \sigma_i, \lambda) = h_1(\beta_0)$$

and λ is

$$\pi(\lambda|.) \propto \pi(\lambda)\prod_i f(y_i|\beta_0, \beta_1, X_i, \sigma_i, \lambda) = h_3(\lambda)$$

. Where |. mean given everything else. Since most of the time calculating this quantity is not easy computationally let‘s use it in log. This way the algorithm to simulate from the posterior is the following.

i. Define x as been the values for β_0, y as been the values for β_1 and z as been the values for λ.
ii. Define $x[1] = 0$, $y[1] = 0$ and $z[1] = 0.1$. These values are the prior mean of β_0 and β_1 and an arbitrary value for λ.
iii. Propose $w \sim N(x[t-1], 1)$
iv. Calculate the log of the acceptance probability defined as $log(\alpha) = min(0, h_1(w) - h_1(x[t-1])$
v. Sample u from a uniform in $(0, 1)$
vi. If $log(u) <= log(alpha)$ define $x[t] = w$, otherwise $x[t] = x[t-1]$
vii. Propose $w \sim N(y[t-1], 1)$
viii. Calculate the log of the acceptance probability defined as $log(\alpha) = min(0, h_2(w) - h_2(x[t-1])$
ix. Sample u from a uniform in $(0, 1)$
x. If $log(u) <= log(alpha)$ define $y[t] = w$, otherwise $y[t] = y[t-1]$
xi. Propose $w \sim Gamma(z[t-1]^2, 1/z[t-1])$, this way the mean of the proposal is $z[t-1]$ and the variance is 1.
xii. Calculate the log of the acceptance probability defined as $log(\alpha) = min(0, h_1(w) - h_1(x[t-1] + Gamma(z[t-1], w^2, 1/w) - Gamma(w, z[t-1]^2, 1/z[t-1]))$
xiii. Sample u from a uniform in $(0, 1)$
xiv. If $log(u) <= log(alpha)$ define $z[t] = w$, otherwise $z[t] = z[t-1]$
xv. Than go back to step (ii) until you have the number of samples that you want.

(b) TABLE I

	Posteiror mean	MCMC Standard Deviation	2.5%	97.5%
Beta0	1.60372	0.00178	1.37983	1.81907
Beta1	3.56576	0.00267	3.20899	3.92416
Lambda	0.51671	0.00027	0.46320	0.57408

TABLE I. Summary of the posterior, Question 2

(c) An approximation to the correlation between β_0 and β_1 is the sample correlation. This is showed following.

```
## [1] -0.8147203
```

(d) (FIG.3)

(e) To evaluate the accuracy of the algorithm let‘s first analyse the cumulative mean, the cumulative MCMC standard error and the MCMC when we start in different values.(FIG.4)
We can see that in the first plot the chain start in different values but it converges to the stationary distribution. The second plot show us that the mean is also converging to the posterior mean and the MCMC standard error is decreasing, after 10000 it so near 0 that we can‘t see any more difference.
We can also see the ESS, it was 4914 to β_0, 5636 to β_1 and 11235 to λ.

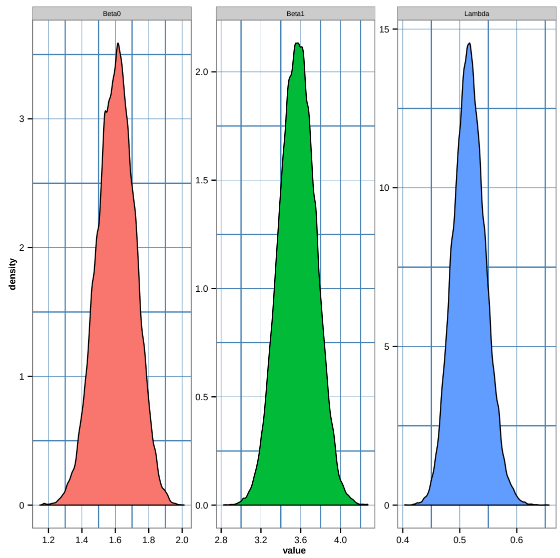

FIG. 3. Posterior Density Question 2

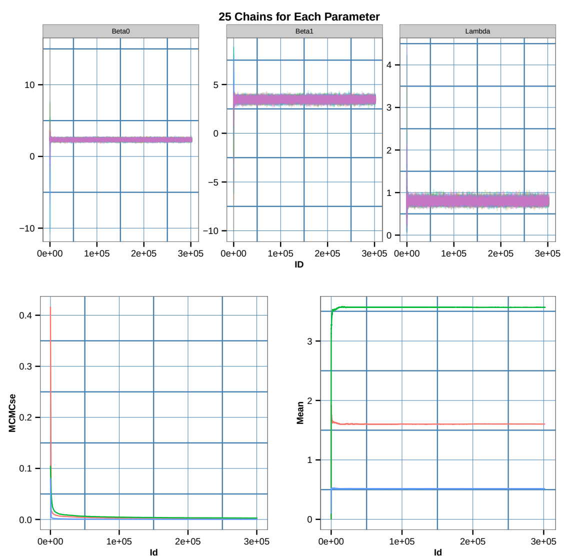

FIG. 4. Chains, Posterior mean and MCMCse by iteration

3. (a) TABLE II

	Posteiror mean	Standard Deviation	2.5%	97.5%
Beta0	0.14932	0.00173	-0.17711	0.45840
Beta1	2.47328	0.00278	1.93347	3.01449
Lambda	0.16116	0.00010	0.15028	0.17203

TABLE II. Summary of the posterior, Question 3

(b) (FIG.5)

(c) None modification was made, some cleaver way of dealing with the increase sample size must be found. To do the MCMC on the second question, it took 5 to 6 minutes while on the third one it took 10 to 12 minutes, even though we were sampling the same number of parameters.

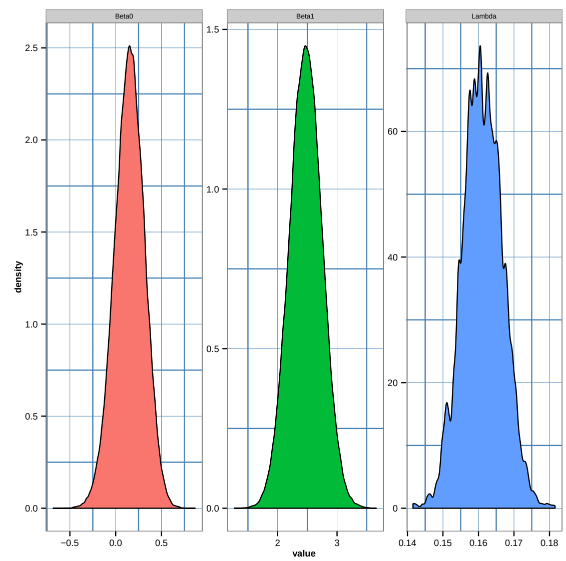

FIG. 5. Posterior Density Question 3