IYSE 6420 Fall 2020 Homework4

Xiao Nan
GT Account: nxiao30
GT ID: 903472104

1. Simple Metropolis: Normal Precision – Gamma.

Suppose $X = -2$ was observed from the population distributed as $\mathcal{N}(0, \frac{1}{\theta})$ and one wishes to estimate the parameter θ.
(Here θ is the reciprocal of the variance σ^2 and is called the precision parameter).
Suppose the analyst believes that the prior on θ is $\mathcal{G}a(\frac{1}{2}, 1)$.
Using Metropolis algorithm, approximate the posterior distribution and the Bayes' estimator of θ. As the proposal distribution, use gamma $\mathcal{G}a(\alpha, \beta)$with parameters α, β selected to ensure efficacy of the sampling (this may require some experimenting).

Likelihood

$$L(\theta) \propto \sqrt{\theta} e^{-\frac{\theta x^2}{2}}$$

Log-likelihood

$$l(\theta) = \frac{1}{2} log\theta - \frac{\theta x^2}{2}$$

$$l'(\theta) = \frac{1}{2\theta} - x^2$$

When $l'(\theta) = 0$, MLE

$$\hat{\theta} = \frac{1}{x^2}$$

If prior is $\theta \sim \mathcal{G}a(r, \lambda)$, posterior

$$\pi(\theta|x) \propto \theta^{r+\frac{1}{2}-1} e^{\lambda + \frac{x^2}{2}}$$

is gamma $\mathcal{G}a\left(r + \frac{1}{2}, \lambda + \frac{x^2}{2}\right)$
Bayes estimator for θ

$$\widehat{\theta_B} = \frac{r + \frac{1}{2}}{\lambda + \frac{x^2}{2}}$$

When $X = -2, r = \frac{1}{2}, \lambda = 1$, posterior is gamma $\mathcal{G}a(1,3)$, which is also exponential $\mathcal{E}(3)$

$$\pi(\theta \mid x = -2) \propto \theta^{\frac{1}{2}+\frac{1}{2}-1} \exp\left\{1 + \frac{(-2)^2}{2}\right\} = e^{-3\theta}$$

The bayes estimator is the mean of posterior

$$\frac{1}{\lambda} = \frac{1}{3}$$

Proposal $q(\theta' \mid \theta) = \mathcal{G}a(1,2.85)$

$$\gamma = \frac{\pi(\theta')q(\theta \mid \theta')}{\pi(\theta)q(\theta' \mid \theta)} = \frac{\theta'^{-\frac{1}{2}} e^{-\theta'} 2.85 e^{-2.85\theta}}{\theta^{-\frac{1}{2}} e^{-\theta} 2.85 e^{-2.85\theta'}} = \frac{\theta'^{-\frac{1}{2}} e^{-1.85\theta}}{\theta^{-\frac{1}{2}} e^{-1.85\theta'}}$$

Use $X = -2, \theta_0 = \frac{1}{2}$ we get bayes estimator = 0.3331, which is close to $\frac{1}{3}$
Code:

```
close all
clear all

rand('seed',1);
randn('seed',1);
x = -2; %data
theta = 0.5; % initial value
thetas =[theta]; %save all thetas.
%

tic
 for i = 1:100000
theta_prop = randn + x;  %N(x,1).
%----------------------------------------------------------------------
--
r = (theta_prop^(-0.5)*exp(-1.85*theta))/(theta^(-0.5)*exp(-
1.85*theta_prop));
%----------------------------------------------------------------------
--
rho = min(r ,1);
   if (rand < rho)
       theta = theta_prop;
   end
thetas =  [thetas theta];
end
toc
%Burn in 500
 thetas = thetas(500:end);
figure(1)
 hist(thetas, 50)
 mean(thetas)
  var(thetas)
```

2. Normal-Cauchy by Gibbs.

Assume that $y_1, y_2, \ldots, y_n$ is a sample from $\mathcal{N}(\theta, \sigma^2)$ distribution, and that the prior on θ is Cauchy $\mathcal{C}a(\mu, \tau)$

$$f(\theta \mid \mu, \tau) = \frac{1}{\pi} \cdot \frac{\tau}{\tau^2 + (\theta - \mu)^2}$$

Even though the likelihood for $y_1, y_2, \ldots, y_n$ simplifies by sufficiency arguments to a likelihood of $\bar{y} \sim \mathcal{N}(\theta, \frac{\sigma^2}{n})$, a closed form for the posterior is impossible and numerical integration is required.

The approximation of the posterior is possible by Gibbs sampler as well. Cauchy $\mathcal{C}a(\mu, \tau)$ distribution can be represented as a scale-mixture of normals:

$$[\theta] \sim \mathcal{C}a(\mu, \tau) \equiv [\theta \mid \lambda] \sim \mathcal{N}\left(\mu, \frac{\tau^2}{\lambda}\right), [\lambda] \sim \mathcal{G}a\left(\frac{1}{2}, \frac{1}{2}\right)$$

that is

$$\frac{\tau}{\pi(\tau^2 + (\theta - \tau)^2)} \propto \int_0^\infty \sqrt{\frac{\lambda}{2\pi\tau^2}} \exp\left\{-\frac{\lambda}{2\tau^2}(\theta - \mu)^2\right\} \cdot \lambda^{\frac{1}{2}-1} \exp\left\{-\frac{\lambda}{2}\right\} d\lambda$$

The full conditionals can be derived from the product of the densities for the likelihood and priors

$$[\bar{y} \mid \theta, \sigma^2] \sim \mathcal{N}\left(\theta, \frac{\sigma^2}{n}\right)$$
$$[\theta \mid \lambda] \sim \mathcal{N}\left(\mu, \frac{\tau^2}{\lambda}\right)$$
$$[\lambda] \sim \mathcal{G}a\left(\frac{1}{2}, \frac{1}{2}\right)$$

(a) Show that full conditionals are normal and exponential

$$[\theta \mid \bar{y}, \lambda] \sim \mathcal{N}\left(\frac{\tau^2}{\tau^2 + \lambda\sigma^2/n}\bar{y} + \frac{\lambda\sigma^2/n}{\tau^2 + \lambda\sigma^2/n}\mu, \frac{\tau^2 \cdot \sigma^2/n}{\tau^2 + \lambda\sigma^2/n}\right)$$
$$[\lambda \mid \bar{y}, \theta] \sim \mathcal{E}\left(\frac{\tau^2 + (\theta - \mu)^2}{2\tau^2}\right)$$

Let $\varphi = \frac{\sigma^2}{n}$

$$[\bar{y} \mid \theta, \sigma^2] \sim \mathcal{N}\left(\theta, \frac{\sigma^2}{n}\right) = \mathcal{N}(\theta, \varphi) = \sqrt{\frac{1}{2\pi\varphi}}\exp\{-\frac{1}{2}\frac{(x-\theta)^2}{\varphi}\}$$

$$joint \propto \sqrt{\frac{1}{2\pi\varphi}}\exp\{-\frac{1}{2}\frac{(x-\theta)^2}{\varphi}\}\int_0^\infty \sqrt{\frac{\lambda}{2\pi\tau^2}}\exp\left\{-\frac{\lambda}{2\tau^2}(\theta-\mu)^2\right\}\cdot \lambda^{\frac{1}{2}-1}\exp\left\{-\frac{\lambda}{2}\right\}d\lambda$$

$$joint \propto \sqrt{\frac{1}{\varphi}}\exp\{-\frac{1}{2}\frac{(\bar{y}-\theta)^2}{\varphi}\}\int_0^\infty \sqrt{\frac{1}{\tau^2}}\exp\left\{-\frac{\lambda}{2\tau^2}(\theta-\mu)^2 - \frac{\lambda}{2}\right\}d\lambda$$

$$[\,\theta \mid \bar{y}, \lambda\,] \propto \exp\left\{-\frac{1}{2}\frac{(\theta(\tau^2+\lambda\varphi)^2 - \tau^2\bar{y} - \lambda\varphi\mu)^2}{\tau^2\cdot\varphi(\tau^2+\lambda\varphi)}\right\}\}$$

$$[\,\theta \mid \bar{y}, \lambda\,] \propto \exp\left\{-\frac{1}{2}\frac{\tau^2+\lambda\varphi}{\tau^2\cdot\varphi}(\theta - \frac{\tau^2}{\tau^2+\lambda\varphi}\bar{y} - \frac{\lambda\varphi}{\tau^2+\lambda\varphi}\mu)^2\right\}\}$$

$$[\,\theta \mid \bar{y}, \lambda\,] \propto \exp\left\{-\frac{1}{2}\frac{\tau^2+\frac{\lambda\sigma^2}{n}}{\tau^2\cdot\frac{\sigma^2}{n}}(\theta - \frac{\tau^2}{\tau^2+\frac{\lambda\sigma^2}{n}}\bar{y} - \frac{\frac{\lambda\sigma^2}{n}}{\tau^2+\frac{\lambda\sigma^2}{n}}\mu)^2\right\}$$

So $[\theta \mid \bar{y}, \lambda] \sim \mathcal{N}\left(\frac{\tau^2}{\tau^2+\lambda\sigma^2/n}\bar{y} + \frac{\lambda\sigma^2/n}{\tau^2+\lambda\sigma^2/n}\mu, \frac{\tau^2\cdot\sigma^2/n}{\tau^2+\lambda\sigma^2/n}\right)$

From joint with respect to λ

$$[\lambda \mid \bar{y}, \theta] \propto exp\left(\frac{\tau^2 + (\theta-\mu)^2}{2\tau^2}\lambda\right)$$

So $[\lambda \mid \bar{y}, \theta] \sim \mathcal{E}\left(\frac{\tau^2+(\theta-\mu)^2}{2\tau^2}\right)$

b) Jeremy models the score on his IQ tests as $\mathcal{N}(\theta, \sigma^2)$ with $\sigma^2 = 90$ He places Cauchy $\mathcal{C}a(110, \sqrt{120})$ prior on θ.
In 10 random IQ tests Jeremy scores $y = [100, 106, 110, 97, 90, 112, 120, 95, 96, 109]$. The

average score is 103.5, which is the frequentist estimator of θ. Using Gibbs sampler described in (a) approximate the posterior mean and variance. Approximate 95% equi-tailed credible set by sample quantiles.

$$\bar{y} \sim \mathcal{N}(\theta, \sigma^2) \sim \mathcal{N}(\theta, 90)$$
$$[\theta] \sim \mathcal{C}a(\mu, \tau) \equiv [\theta \mid \lambda] \sim \mathcal{N}\left(\mu, \frac{\tau^2}{\lambda}\right)$$
$$[\theta] \sim \mathcal{C}a\left(110, \sqrt{120}\right) \equiv [\theta \mid \lambda] \sim \mathcal{N}\left(110, \frac{120}{\lambda}\right)$$

$$[\lambda] \sim \mathcal{G}a\left(\frac{1}{2}, \frac{1}{2}\right)$$
$$[\theta \mid \bar{y}, \lambda] \sim \mathcal{N}\left(\frac{120}{120 + \lambda 90/n} 103.5 + \frac{\lambda 90/n}{120 + \lambda 90/n} 110, \frac{120 \cdot 90/n}{120 + \lambda 90/n}\right)$$
$$[\lambda \mid \bar{y}, \theta] \sim \mathcal{E}\left(\frac{120 + (\theta - 110)^2}{2 \times 120}\right)$$
$$\theta_0 = [100, 106, 110, 97, 90, 112, 120, 95, 96, 109]$$

```
clear all
close all force
randn('state',4);

data = [100,106,110,97,90,112,120,95,96,109];
yhat = 103.5;
%
lendata=length(data);
sumdata=sum(data);

sigma2 = 90;
tau2 = 120;
mu = 110;
%
theta = 0;
thetas =[theta];
lambda = gamma(0.5);
lambdas=[lambda];
burn =1000;
ntotal = 10000 + burn;
tic
for i = 1: ntotal
  theta = (tau2/(tau2 + lambda * sigma2) * yhat + ...
    lambda * sigma2/(tau2 + lambda * sigma2) * mu) + ...
    sqrt(tau2 * sigma2/(tau2 + lambda *sigma2)) * randn;
  lambda =  exprnd( 1/((tau2 + (theta - mu)^2)/(2*tau2)));
thetas =[thetas theta];
lambdas =[lambdas lambda];
end
toc
%
 mean(thetas(burn+1:end))
 var(thetas(burn+1:end))
 hist(thetas(burn+1:end), 40)
 prctile(thetas(burn+1:end), 2.5)
 prctile(thetas(burn+1:end), 97.5)
```

Mean: 106.1853

Var: 55.3554

95% creditable set: (90.29, 120.08)