MCMC:HW3

Qi Chen(qc586)

March 1, 2017

Problem 1

Suppose

$$\pi(a, b) \propto a^4 b^6 e^{-a-b-3ab} \tag{1}$$

for $a,\ b > 0$ is required to be sampled. Find

$$\pi(a \mid b) \text{ and } \pi(b \mid a) \tag{2}$$

and hence find a transition density $p(a_{n+1}, b_{n+1} \mid a_n, b_n)$ which has π as the stationary density. Implement the chain and use the output to evaluate the integral

$$I = \int\int a\, b\, \pi(a, b)\, da\, db \tag{3}$$

Solution: According to the definition of the marginal density,

$$\begin{aligned}
\pi(a \mid b) &= \frac{\pi(a, b)}{\pi(b)} \\
&= \frac{a^4 b^6 e^{-a-b-3ab}}{\int_0^\infty a^4 b^6 e^{-a-b-3ab}\, da} \\
&= \frac{a^4 b^6 e^{-a-b-3ab}}{e^{-b} b^6 \int_0^\infty a^4 e^{-(1+3b)a}\, da} \\
&= \frac{a^4 b^6 e^{-a-b-3ab}}{e^{-b} b^6 \frac{24}{(1+3b)^5}} \\
&= \frac{a^4 e^{-(1+3b)a}(1+3b)^5}{24}
\end{aligned} \tag{4}$$

We can check that

$$\int_0^\infty \pi(a \mid b)\, da = 1 \tag{5}$$

```
Integrate[a^4 e^-(1+3 b) a, {a, 0, ∞}, Assumptions → {b > 0}]
```

$$\frac{24}{(1+3b)^5}$$

$$\begin{aligned}
\pi(b \mid a) &= \frac{\pi(a, b)}{\pi(a)} \\
&= \frac{a^4 b^6 e^{-a-b-3ab}}{\int_0^\infty a^4 b^6 e^{-a-b-3ab}\, db} \\
&= \frac{b^6 e^{-b-3ab}}{\int_0^\infty b^6 e^{-(1+3a)b}\, db} \\
&= \frac{b^6 e^{-b-3ab}}{\frac{720}{(1+3a)^7}} \\
&= \frac{b^6 e^{-b(1+3a)}(1+3a)^7}{720}
\end{aligned} \tag{6}$$

We can check that

$$\int_0^\infty \pi(b \mid a)\, db = 1 \tag{7}$$

```
Integrate[b^6 e^-(1+3 a) b, {b, 0, ∞}, Assumptions → {a > 0}]
```

$$\frac{720}{(1+3\,a)^7}$$

We want to find the transition density $p(a_{n+1}, b_{n+1} \mid a_n, b_n)$ such that

$$\pi(a_{n+1}, b_{n+1}) = \int p(a_{n+1}, b_{n+1} \mid a_n, b_n)\, \pi(a_n, b_n)\, da_n\, db_n \tag{8}$$

Consider the marginal density

$$\pi(a_{n+1}) = \int p(a_{n+1} \mid a_n)\, \pi(a_n)\, da_n \tag{9}$$

Obviously,

$$\begin{aligned} p(a_{n+1} \mid a_n) &= \int \pi(a_{n+1} \mid u)\, \pi(u \mid a_n)\, du \\ &= \int \frac{a_{n+1}{}^4\, e^{-(1+3\,u)\,a_{n+1}} (1+3\,u)^5}{24} \frac{u^6\, e^{-u(1+3\,a_n)} (1+3\,a_n)^7}{720}\, du \end{aligned} \tag{10}$$

satisfies the stationary condition. Similarly,

$$\pi(b_{n+1}) = \int p(b_{n+1} \mid b_n)\, \pi(b_n)\, da_n \tag{11}$$

$$\begin{aligned} p(b_{n+1} \mid b_n) &= \int \pi(b_{n+1} \mid u)\, \pi(u \mid b_n)\, du \\ &= \int \frac{b_{n+1}{}^6\, e^{-b_{n+1}(1+3\,u)} (1+3\,u)^7}{720} \frac{u^4\, e^{-(1+3\,b_n)\,u} (1+3\,b_n)^5}{24}\, du \end{aligned} \tag{12}$$

So in principle we can sample from the marginal transition density, but the integrals are not easy to evaluate. Alternatively, We can verify that

$$\begin{aligned} & p(a_{n+1}, b_{n+1} \mid a_n, b_n) \\ &= \pi(a_{n+1} \mid b_{n+1})\, \pi(b_{n+1} \mid a_n) \\ &= \frac{a_{n+1}{}^4\, e^{-(1+3\,b_{n+1})\,a_{n+1}} (1+3\,b_{n+1})^5}{24} \frac{b_{n+1}{}^6\, e^{-b_{n+1}(1+3\,a_n)} (1+3\,a_n)^7}{720} \end{aligned} \tag{13}$$

satisfies Eq. (8) as

$$\begin{aligned} & \int p(a_{n+1}, b_{n+1} \mid a_n, b_n)\, \pi(a_n, b_n)\, da_n\, db_n \\ &= \int \pi(a_{n+1} \mid b_{n+1})\, \pi(b_{n+1} \mid a_n)\, \pi(a_n, b_n)\, da_n\, db_n \\ &= \int \pi(a_{n+1} \mid b_{n+1})\, \pi(b_{n+1} \mid a_n)\, \pi(a_n)\, da_n \\ &= \pi(a_{n+1} \mid b_{n+1})\, \pi(b_{n+1}) \\ &= \pi(a_{n+1}, b_{n+1}) \end{aligned} \tag{14}$$

In evaluating the integral, notice that

$$\begin{aligned} I &= \int\int a\, b\, \pi(a, b)\, da\, db \\ &= \frac{\int\int a\, b\, \mathrm{kernel}(a, b)\, da\, db}{\int\int \mathrm{kernel}(a, b)\, da\, db} \end{aligned} \tag{15}$$

where

$$\mathrm{kernel}(a, b) = a^4\, b^6\, e^{-a-b-3\,a\,b} \tag{16}$$

The exact value is given by

$$I \approx 1.43461 \tag{17}$$

```
I2 = Integrate[a^4 b^6 e^-a-b-3 a b, {a, 0, ∞}, {b, 0, ∞}]
```

$$\frac{-43\,842 - 43\,813\, e^{1/3}\, \mathrm{ExpIntegralEi}\left[-\frac{1}{3}\right]}{177\,147}$$

```
I1 = Integrate[a b a^4 b^6 e^(-a-b-3 a b), {a, 0, ∞}, {b, 0, ∞}]
```

$$-\frac{2\left(541842+506573\,e^{1/3}\,\text{ExpIntegralEi}\left[-\frac{1}{3}\right]\right)}{1594323}$$

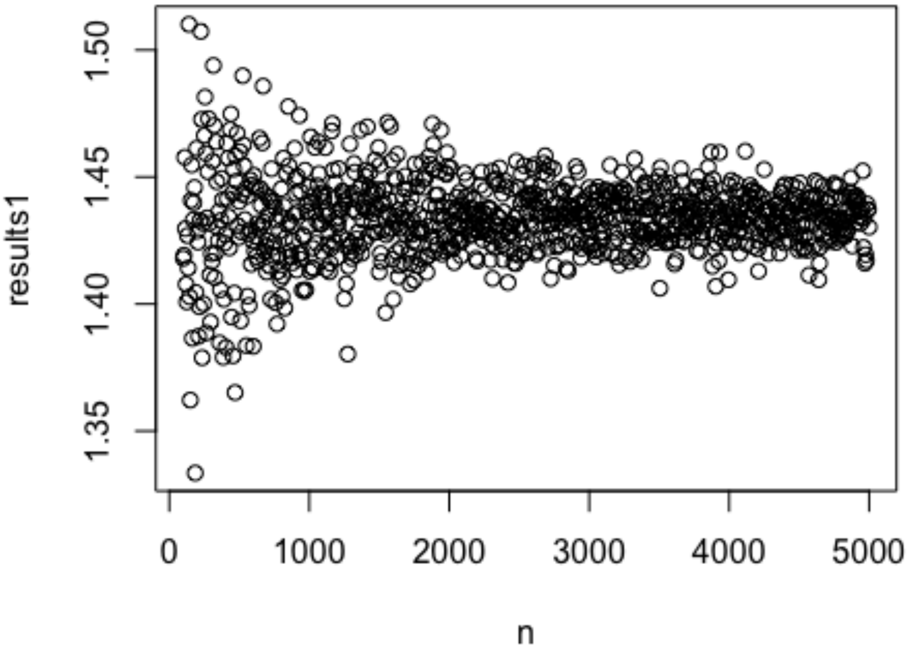
```
N[I1/I2]
```

1.43461

By use of the transition density Eq. (13), the MCMC simulation results for $n = 1000$ gives

$$I \approx 1.435733 \tag{18}$$

The convergence is monitored as follows:



Problem 2

Suppose we wish to sample from

$$\pi(x) \propto \frac{\exp\left(-\frac{1}{2}x^2\right)}{1+x^2} \tag{19}$$

using a Metropolis-Hastings algorithm with density $q(x' \mid x) = N\left(x' \mid x, \sigma^2\right)$ for some σ as the proposal density.

Describe what you think might be the problems encountered if (i) σ is too small and (ii) σ is too big. Run the algorithm with such σ to verify your conclusions.

Without using any theory, find what you think is a suitable σ and run the algorithm with this σ to evaluate

$$I = \int x\,\pi(x)\,dx \tag{20}$$

What happens if instead you use $q(x' \mid x) = N\left(x' \mid -x, \sigma^2\right)$.

Solution: The pdf of our choice of easy to sample density is

$$q(x' \mid x) = \frac{1}{\sqrt{2\pi}\,\sigma}\,e^{-\frac{(x'-x)^2}{2\sigma^2}} \tag{21}$$

According to the Metropolis-Hastings algorithm,

$$p(x' \mid x) = \min\left\{1, \frac{\pi(x')\, q(x \mid x')}{\pi(x)\, q(x' \mid x)}\right\} q(x' \mid x) + (1 - r(x))\, \mathbf{1}\,(x' = x)$$
$$= r(x)\, \tilde{q}(x' \mid x) + (1 - r(x))\, \mathbf{1}\,(x' = x) \tag{22}$$

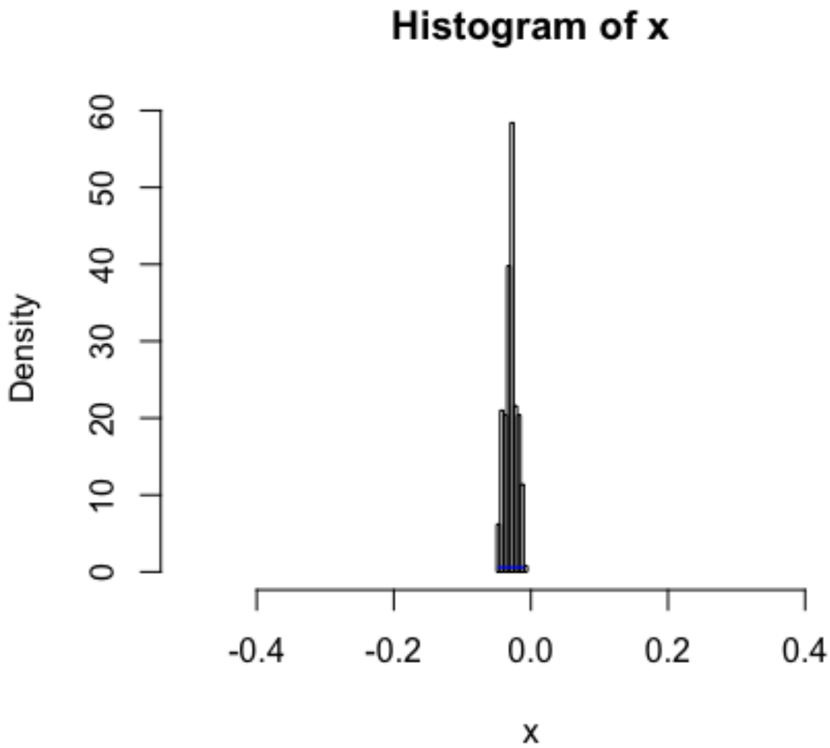
$$\tilde{q}(x' \mid x) = \frac{\alpha(x', x)\, q(x' \mid x)}{r(x)} \tag{23}$$

$$r(x) = \int \alpha(x', x)\, q(x' \mid x)\, dx' \tag{24}$$

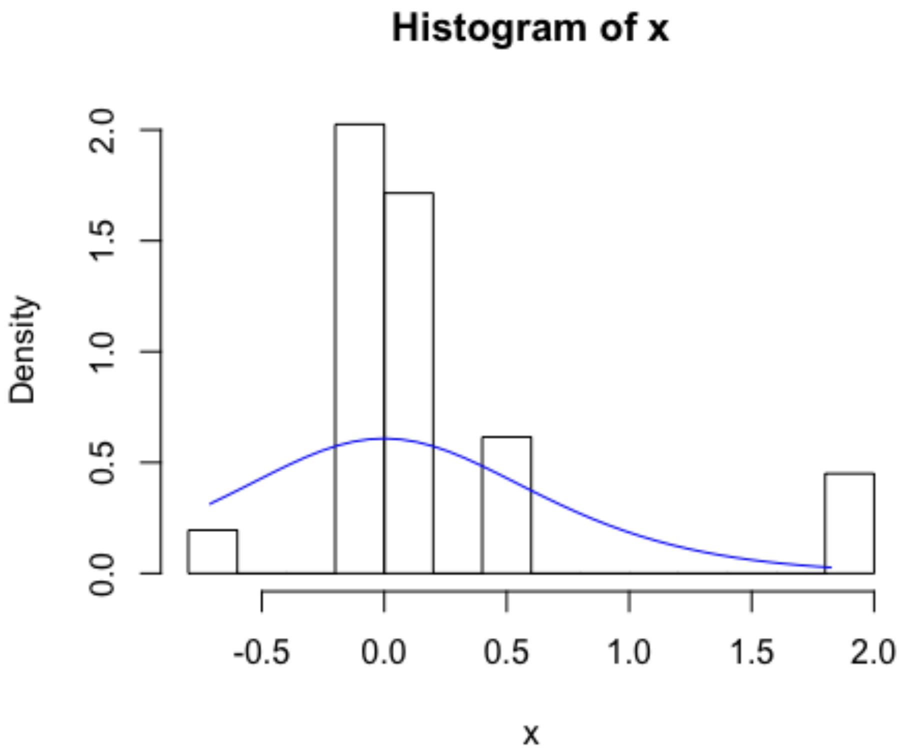
$$\alpha(x', x) = \min\left\{1, \frac{\pi(x')\, q(x \mid x')}{\pi(x)\, q(x' \mid x)}\right\} = \min\left\{1, \frac{\pi(x')}{\pi(x)}\right\} \tag{25}$$

Since $\tilde{q}(x' \mid x)$ and $\mathbf{1}\,(x' = x)$ are two densities, we sample x' from $\tilde{q}(x' \mid x)$ with density $r(x)$ else we sample x' from $1\,(x' = x)$ with probability $1 - r(x)$. But we do not know $r(x)$ as it involves an integral. The trick is we sample x' from $q(x' \mid x)$, u from $U(0, 1)$. If $u < \alpha(x', x)$, x' automatically comes from $\tilde{q}(x' \mid x)$; otherwise we set $x' = x$.

In our problem, if σ is too small, then we can only generate x' too close to x from $q(x' \mid x)$, which will not be able to cover the whole support. As a result, the sampling is too close to the initial value. This is demonstrated in the following histogram for the choice of $\sigma = 100$



On the other hand, if σ is too big, then we will generate x' that has a sparse distribution in $(-\infty, \infty)$ from $q(x' \mid x)$. Among all the x', only those $|x'| \lesssim |x|$ will be likely to succeed, for those not succeed, we set $x' = x$, so the histogram of x will be very sparse with high peaks around a few values, which means the chain does not move very well. This is demonstrated in the following histogram for the choice of $\sigma = 100$



The exact value of the integral is 0 by symmetry

$$I = \int x\,\pi(x)\,dx \propto \int_{-\infty}^{\infty} x\,\frac{\exp\left(-\frac{1}{2}x^2\right)}{1+x^2}\,dx = 0 \tag{26}$$

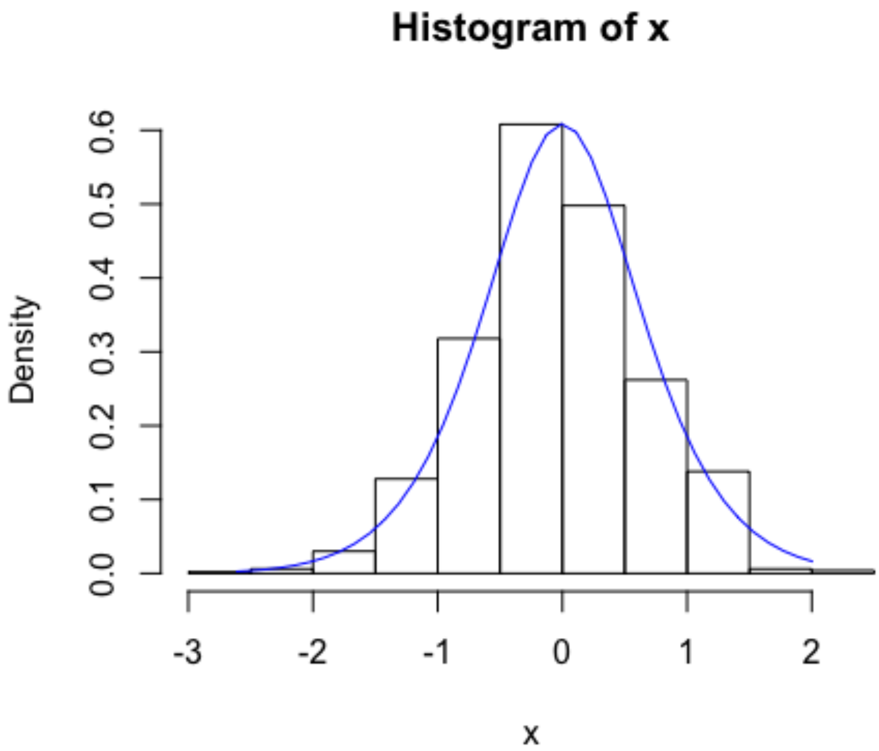
```
N[Integrate[x^2 Exp[-1/2 x^2]/(1 + x^4), {x, -∞, ∞}]/
  Integrate[Exp[-1/2 x^2]/(1 + x^4), {x, -∞, ∞}]]
```

```
0.396347
```

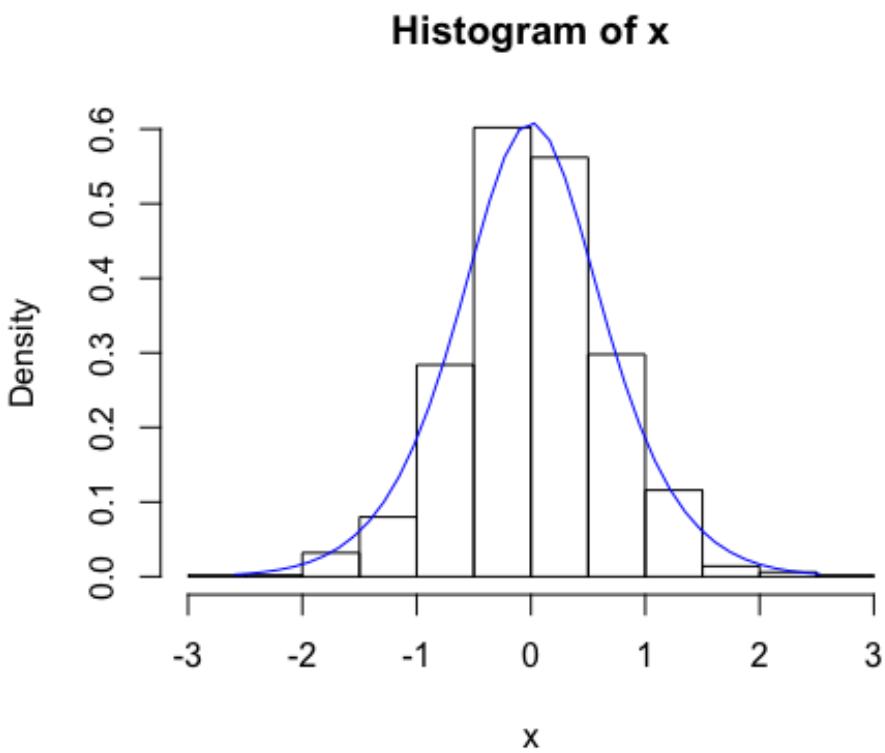
By choosing $\sigma = 1$, we find

$$I = -0.04562668 \tag{27}$$

The sampling histogram for $n = 1000$ compared with pdf is given as follows:



If we choose $q(x' \mid x) = N(x' \mid -x, \sigma^2)$, the sampling becomes more symmetric around $x = 0$ according to the following histogram:



This is due to the even nature of $\pi(x)$, x and $-x$ should have equal probability of being accepted in the choice of

$q(x' \mid x) = N(x' \mid -x, \sigma^2)$. The integral is evaluated to be

$$I \approx 0.004607338 \tag{28}$$

Problem 3

Suppose a posterior density is given by

$$f(\theta \mid \text{data}) \propto e^{\theta a} e^{-n e^{\theta}} e^{-\frac{1}{2}\theta^2} \tag{29}$$

for some $a > 0$ and n integer, and $-\infty < \theta < +\infty$. In fact this is a Poisson model with mean e^{θ} and standard normal prior for θ.

Find a Markov chain for sampling from f and implement it.

Solution: We want to find $p(\theta_{n+1} \mid \theta_n)$ such that

$$f(\theta_{n+1} \mid \text{data}) = \int p(\theta_{n+1} \mid \theta_n) f(\theta_n \mid \text{data}) \, d\theta_n \tag{30}$$

By Metropolis-Hastings algorithm, we can choose

$$q(x) = \frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2}x^2} \tag{31}$$

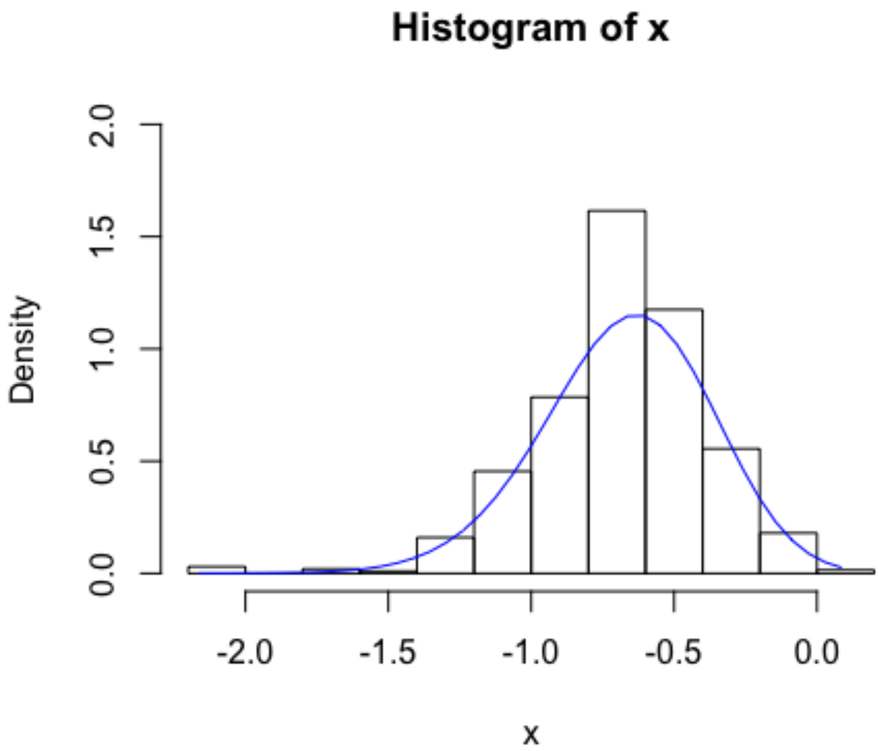
$$\begin{aligned} p(x' \mid x) &= \min\left\{1, \frac{\pi(x')\, q(x)}{\pi(x)\, q(x')}\right\} q(x') + (1 - r(x))\, \mathbf{1}\,(x' = x) \\ &= r(x)\, \tilde{q}(x' \mid x) + (1 - r(x))\, \mathbf{1}\,(x' = x) \end{aligned} \tag{32}$$

$$\tilde{q}(x' \mid x) = \frac{\alpha(x', x)\, q(x')}{r(x)} \tag{33}$$

$$r(x) = \int \alpha(x', x)\, q(x')\, dx' \tag{34}$$

$$\alpha(x', x) = \min\left\{1, \frac{\pi(x')\, q(x)}{\pi(x)\, q(x')}\right\} \tag{35}$$

Here we choose $a = 10$, $n = 20$ for MCMC sampling. The histogram $n = 1000$ compared with the pdf is given as follows:



Appendix: R-code

```
# problem 1
n<-1000
a<-array(dim=n)
b<-array(dim=n)
# initialize
a[1]<-1
b[1]<-1
# MCMC
for(i in 2:n){
  b[i]<-rgamma(1,shape=7,scale=1/(1+3*a[i-1]))
  a[i]<-rgamma(1,shape=5,scale=1/(1+3*b[i]))
}
I <- sum(a*b)/n
I
# integral vs n
I_N<- function(n){
  a<-array(dim=n)
  b<-array(dim=n)
  # initialize
  a[1]<-1
  b[1]<-1
  # MCMC
  for(i in 2:n){
    b[i]<-rgamma(1,shape=7,scale=1/(1+3*a[i-1]))
    a[i]<-rgamma(1,shape=5,scale=1/(1+3*b[i]))
  }
  return(sum(a*b)/n)
}
```

```
n = seq.int(100,5000,5)
results1 <- lapply(n,I_N)
plot(n, results1)

# problem 2
n<-1000
# pi(x)
pi <- function(x){
  return(exp(-1/2*x**2)/(1+x**2))
}
# normalized density
f <- function(x){
  result=integrate(pi,-Inf,Inf)$val
  return(pi(x)/result)
}
# q(x1|x2)
q <- function(x1,x2,sigma){
  return(dnorm(x1,x2,sigma))
}
# alpha(x1,x2)
alpha <- function(x1,x2,sigma){
  return(min(1,pi(x1)*q(x2,x1,sigma)/(pi(x2)*q(x1,x2,sigma))))
}

x<-array(dim=n)
u<-array(dim=n)
x[1]=-0.01
u[1]=0.1
sigma=1
for(i in 2:n){
  x[i]<-rnorm(1,x[i-1],sigma)
  u[i]<-runif(1,0,1)
  if(u[i]>=alpha(x[i],x[i-1],sigma)){
    x[i]=x[i-1]
  }
}
I <- sum(x)/n
I
hist(x,freq=FALSE)
xfit<-seq(min(x),max(x),length=40)
yfit<-lapply(xfit,f)
lines(xfit, yfit, col="blue")

# problem 3
n<-1000
# pi(x)
pi <- function(x){
  a<-10
  m<-20
  return(exp(a*x)*exp(-m*exp(x))*exp(-1/2*x**2))
}
# normalized density
library(cubature)
f <- function(x){
  result=integrate(pi,-50,50)$val
  return(pi(x)/result)
}
# q(x1,x2)
q <- function(x1,x2){
  #return(dnorm(x1,x2,1))
  return(dnorm(x1,0,1))
}
# alpha(x1,x2)
alpha <- function(x1,x2){
  return(min(1,pi(x1)*q(x2,x1)/(pi(x2)*q(x1,x2))))
}
```

```
x<-array(dim=n)
u<-array(dim=n)
x[1]=0
u[1]=0.1
for(i in 2:n){
  x[i]<-rnorm(1,0,1)
  u[i]<-runif(1,0,1)
  if(u[i]>=alpha(x[i],x[i-1])){
    x[i]=x[i-1]
  }
}
I <- sum(x)/n
I
hist(x,freq=FALSE,ylim=c(0,2))
xfit<-seq(min(x),max(x),length=40)
yfit<-sapply(xfit,f)
lines(xfit, yfit, col="blue")
```