MCMC:HW1

Qi Chen(qc586)

January 30, 2017

Problem 1

Use rejection sampling to sample from the density function

$$f(x) \propto (-\log x)^2\, x^3 (1-x)^2,\ 0 < x < 1 \tag{1}$$

Carefully detail the method you use, and provide a figure of the histogram of the samples you obtained. What is the approximate, or actual if you can find it, probability of acceptance.

Solution: For rejection sampling, we can write

$$f(x) = l(x)\, h(x) \tag{2}$$

where

$$h(x) = \frac{\Gamma(3+3)}{\Gamma(3)\,\Gamma(3)} x^{3-1} (1-x)^{3-1} = 30\, x^2 (1-x)^2 \tag{3}$$

is the Beta(3, 3) pdf. Then

$$l(x) = \frac{\mathcal{N}}{30} (-\log x)^2\, x \tag{4}$$

where $\mathcal{N}$ is the normalization constant. $l(x)$ is bounded by

$$M = l\left(e^{-2}\right) = \frac{4\,\mathcal{N}}{30\, e^2} = \frac{2\,\mathcal{N}}{15\, e^2} \tag{5}$$

As

$$\mathcal{N} = \frac{108\,000}{919} \tag{6}$$

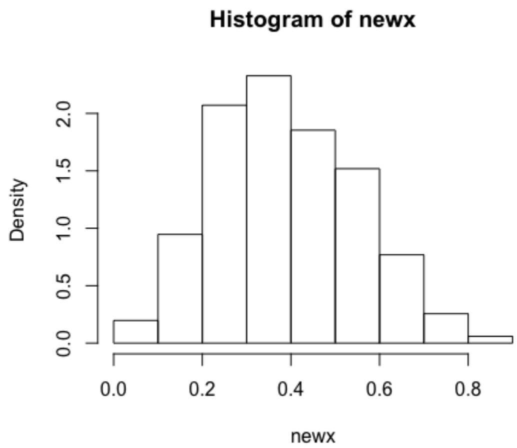
The rate of acceptance should be

$$P = 1/M \approx 0.471565 \tag{7}$$

The simulation result($N = 1000$) gives

$$P \approx 0.507 \tag{8}$$

The histogram is as follows:



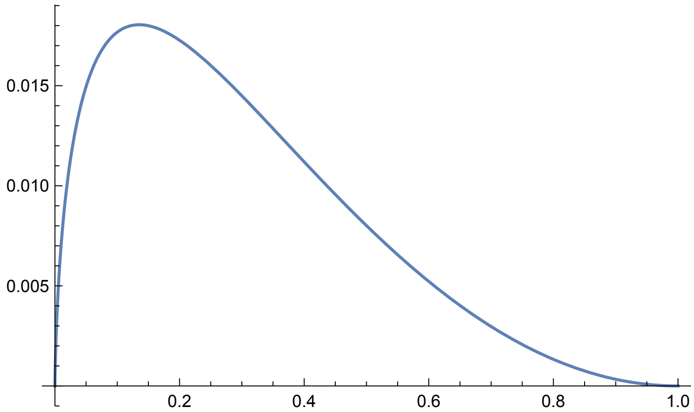

```
Integrate[30 x^2 (1 - x)^2, {x, 0, 1}]
```

1

```
Integrate[(-Log[x])^2 x^3 (1 - x)^2, {x, 0, 1}]
```

$$\frac{919}{108\,000}$$

```
Plot[1/30 (Log[x])^2 x, {x, 0, 1}]
```



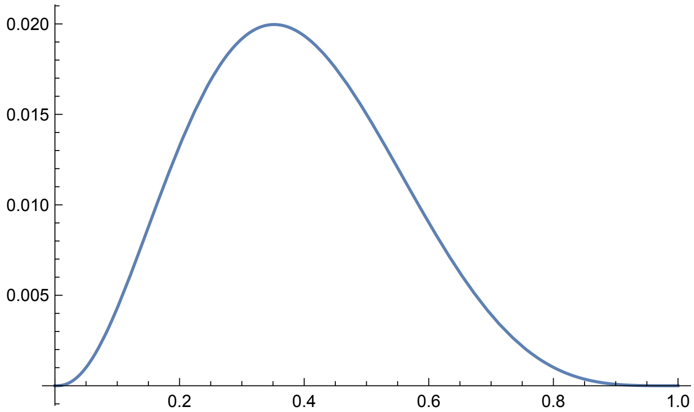
```
N[Exp[-2]]
```

0.135335

$$N\left[1\Big/\left(\frac{108\,000}{919}*4/(30\,e^2)\right)\right]$$

0.471565

```
Plot[(-Log[x])^2 x^3 (1 - x)^2, {x, 0, 1}]
```



Problem 2

Use Monte Carlo methods to evaluate the integral

$$I = \int_0^1 (-\log x)^2\, x^3 (1-x)^{5/2}\, dx \tag{9}$$

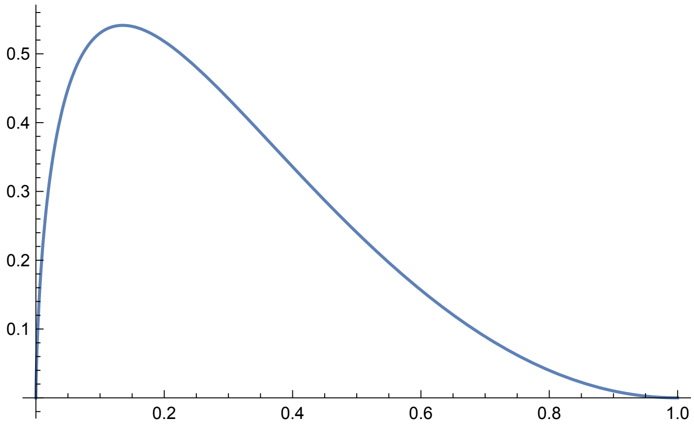
Describe in detail how you do this. Provide a graphical demonstration that your method has worked. If you fix the Monte Carlo sample size as $N = 1000$, is there a way to estimate the variance of your $\hat{I}_N$.

Solution:

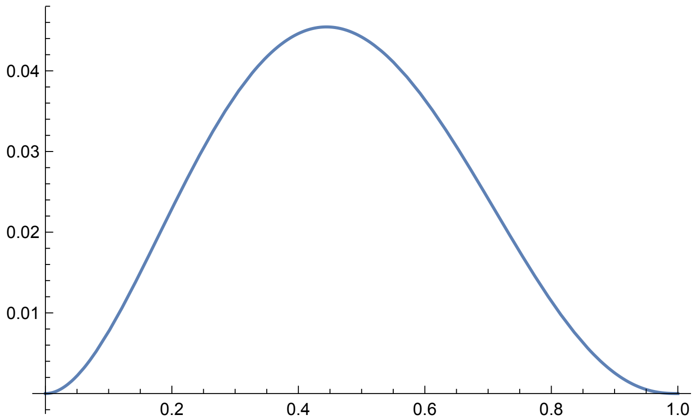
```
Integrate[(-Log[x])^2 x^3 (1 - x)^(5/2), {x, 0, 1}]
```

$$-\frac{32\,\pi^2}{9009} + \frac{64\,(7\,857\,230\,824 + 45\,045\,\mathrm{Log}[2]\,(-187\,111 + 90\,090\,\mathrm{Log}[2]))}{6\,093\,243\,231\,075}$$

```
Plot[(Log[x])^2 x, {x, 0, 1}]
```




```
Plot[x^2 (1 - x)^(5/2), {x, 0, 1}]
```



The exact result is obtained in R as 0.006587444 with absolute error < 4.1e-07. First we consider direct evaluation:

$$I = \int_0^1 g(x)\, f(x)\, dx \tag{10}$$

where

$$f(x) = \frac{\Gamma(3 + 7/2)}{\Gamma(3)\,\Gamma(7/2)} x^{3-1}(1-x)^{7/2-1} \tag{11}$$

$$g(x) = \frac{\Gamma(3)\,\Gamma(7/2)}{\Gamma(3 + 7/2)} (-\log x)^2\, x \tag{12}$$

The integral can be approximated by generating $X_1, X_2, \ldots, X_N \overset{i.i.d}{\sim} \text{Beta}(3, 7/2)$ and calculate the average:

$$\hat{I}_N = \frac{1}{N}\sum_{i=1}^{N} g(X_i) \approx 0.006744701 \tag{13}$$

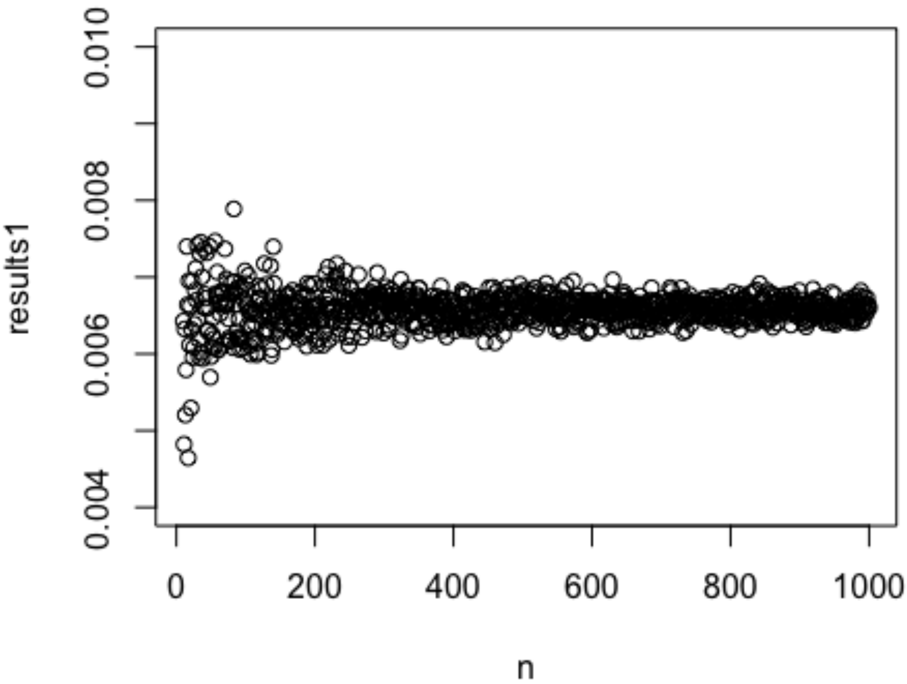
The variance is evaluated as

$$\text{Var}(\hat{I}_N) = \frac{\text{Var}(g(X_i))}{N} \tag{14}$$

which can be estimated by sample variance:

$$S(\hat{I}_N)^2 = \frac{S(g(X_i))^2}{N} = \frac{\sum_i \left(g(X_i) - \hat{I}_N\right)^2}{N(N-1)} \approx 1.198441\, e - 08 \tag{15}$$

The demonstration of variance and convergence is shown in the following graph with the MC result as a function of sample size



Alternatively, we could use rejection sampling as

$$f(x) = (-\log x)^2\, x^3 (1-x)^{5/2} = l(x)\, h(x) \tag{16}$$

$$h(x) = \frac{\Gamma(3+7/2)}{\Gamma(3)\,\Gamma(7/2)}\, x^{3-1} (1-x)^{7/2-1} \tag{17}$$

$$l(x) = \frac{\Gamma(3)\,\Gamma(7/2)}{\Gamma(3+7/2)} (-\log x)^2\, x \tag{18}$$

$$I = \int_0^1 f(x)\, dx = \int_0^1 \int_0^1 M^{-1}\, \mathbf{1}\, (u < l(x)/M)\, h(x).du\, dx = \int_0^1 \int_0^1 f(x,\, u)\, du\, dx \tag{19}$$

where

$$M = \max\,\{l(x)\} = \frac{\Gamma(3)\,\Gamma(7/2)}{\Gamma(3+7/2)}\, 4\, e^{-2} \tag{20}$$

The acceptance probability is given by

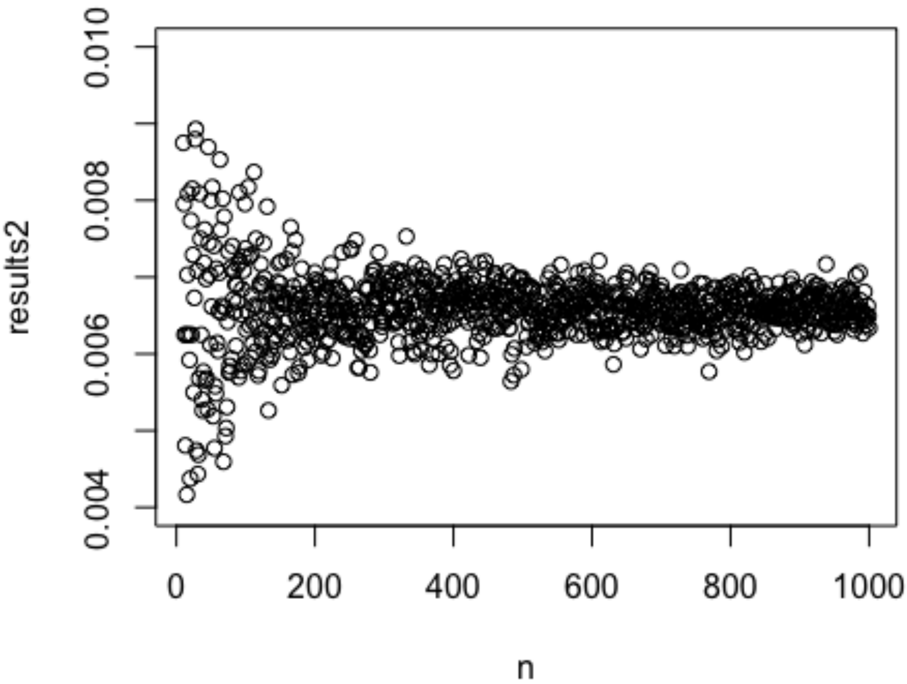
$$P = \frac{I}{M} \Rightarrow I = M \times P = \frac{\Gamma(3)\,\Gamma(7/2)}{\Gamma(3+7/2)}\, 4\, e^{-2} \times P \approx 0.006424227 \tag{21}$$

where P is obtained by numerical simulation.

The variance is evaluated as

$$\begin{aligned} \mathrm{Var}(\bar{I}_N) &= \frac{1}{N} \mathrm{Var}(\mathbf{1}\, (u < l(x)/M)) = \frac{1}{N}\left[E(\mathbf{1}\, (u < l(x)/M)) - E(\mathbf{1}\, (u < l(x)/M))^2\right] \\ &= \frac{1}{N}\left[\frac{I}{M} - \left(\frac{I}{M}\right)^2\right] = \frac{1}{N}\left(P - P^2\right) \approx 0.000249804 \end{aligned} \tag{22}$$

The demonstration of variance and convergence is shown in the following graph with the MC result as a function of sample size



As we can see, the variance is broader than the direct evaluation with beta distribution.

Problem 3

What is the acceptance probability when sampling a standard normal random variable with density

$$f(x) = \frac{1}{\sqrt{2\pi}} e^{-x^2/2} \tag{23}$$

using a Cauchy density as proposal; i.e.

$$h(x) = \frac{1}{\pi(1+x^2)} \tag{24}$$

when using rejection sampling. Verify this using simulation and plot a histogram of 1000 accepted samples.

Solution: We can write $f(x)$ as

$$f(x) = h(x)\, l(x) \tag{25}$$

where

$$l(x) = \sqrt{\frac{\pi}{2}}\, e^{-x^2/2}(1+x^2) \tag{26}$$

which is bounded by

$$M = l(\pm 1) = \sqrt{2\pi}\, e^{-1/2} \tag{27}$$

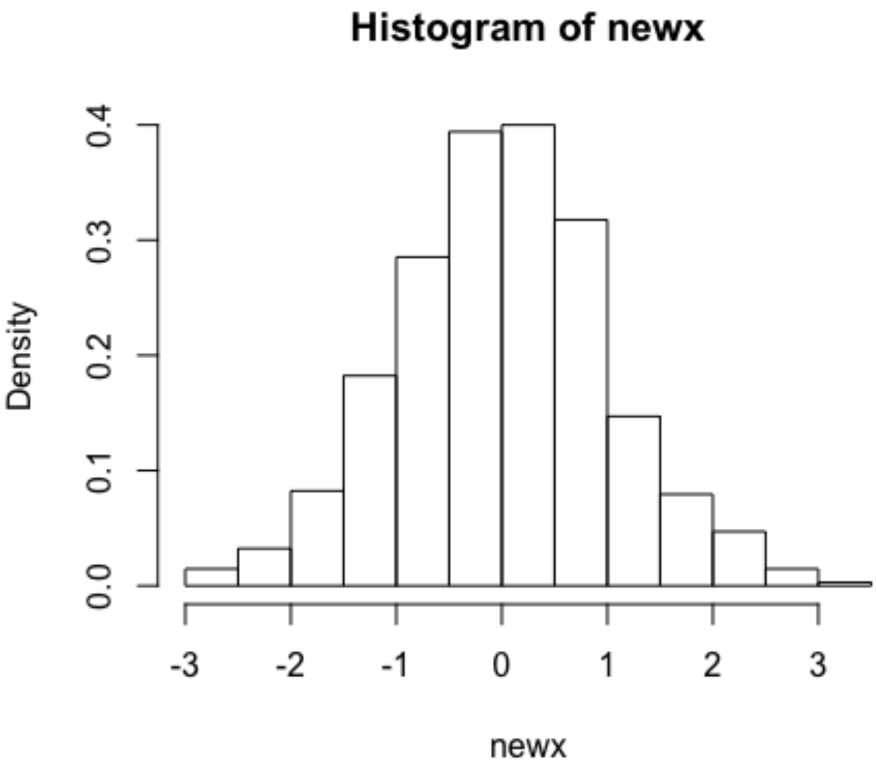
The acceptance probability is given by

$$P = 1/M = \frac{1}{\sqrt{2\pi}\, e^{-1/2}} \approx 0.657745 \tag{28}$$

The simulation result gives

$$P \approx 0.68 \tag{29}$$

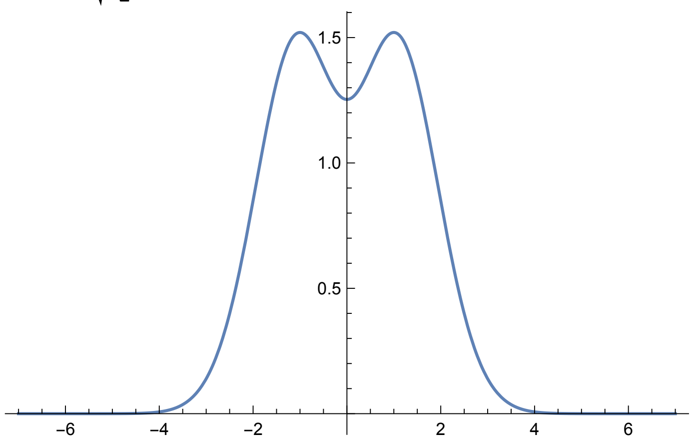
The histogram is as follows




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

1

```
Plot[Sqrt[π/2] e^(-x^2/2) (1 + x^2), {x, -7, 7}, PlotRange → Full]
```




```
N[1/(Sqrt[2 π] e^(-1/2))]
```

0.657745

Appendix: R-code

```
# Use rejection sampling to sample from the density function
# f(x)∝(-log x)^2 x^3(1-x)^2,0<x<1
```

```
n <- 1000
x <- rbeta(n, shape1=3, shape2=3)
u <- runif(n, min=0, max=1)

sum(u<exp(2)*x*log(x)**2/4)/n

newx <- x[u<exp(2)*x*log(x)**2/4]
hist(newx,freq=FALSE)

# Use Monte Carlo methods to evaluate the integral
# I=\!\(
# \*SubsuperscriptBox[\(∫\), \(0\), \(1\)]\(
#  \*SuperscriptBox[\((\(-log\)\ x)\), \(2\)]
#  \*SuperscriptBox[\(
#    \*SuperscriptBox[\(x\), \(3\)](1 - x)\), \(5/2\)] ⅆx\)\)

n<-1000

# analytical result
integrand <- function(x) {(log(x))**2*x**3*(1-x)**(5/2)}
integrate(integrand, lower = 0, upper = 1)

# Method I: sampling from beta distribution

x <- rbeta(n, shape1=3, shape2=7/2)
g <- gamma(3)*gamma(7/2)/gamma(3+7/2)*log(x)**2*x
I1 <- sum(g)/n
I1

# demonstration of convergence
I_N1 <- function(n){
  x <- rbeta(n, shape1=3, shape2=7/2)
  g <- gamma(3)*gamma(7/2)/gamma(3+7/2)*log(x)**2*x
  return(sum(g)/n)
}

n = 10:1000
results1 <- lapply(n,I_N1)
plot(n, results1, ylim=c(0.004, 0.01))

# variance estimation
varI1 = sum((g-mean(g))**2)/(n-1)/n
varI1

# Method II: rejection sampling

x <- rbeta(n, shape1=3, shape2=7/2)
u <- runif(n, min=0, max=1)

M=1/(sum(u<exp(2)*x*log(x)**2/4)/n)

I2 <- 4*gamma(3)*gamma(7/2)/(exp(2)*M*gamma(3+7/2))
I2

I_N2 <- function(n){
  x <- rbeta(n, shape1=3, shape2=7/2)
  u <- runif(n, min=0, max=1)
  M=1/(sum(u<exp(2)*x*log(x)**2/4)/n)
  return(4*gamma(3)*gamma(7/2)/(exp(2)*M*gamma(3+7/2)))
}

n = 10:1000
results2 <- lapply(n,I_N2)
plot(n, results2, ylim=c(0.004, 0.01))

# variance estimation
```

```
varI2 = 1/n*(1/M-(1/M)**2)
varI2

# What is the acceptance probability when sampling a standard normal random
variable with density
# f(x)=1/Sqrt[2π] E^(-x^2/2)
# using a Cauchy density as proposal; i.e.
# h(x)=1/π(1+x^2)
# when using rejection sampling. Verify this using simulation and plot a histogram
of 1000 accepted samples.

n <- 1000
x <- rcauchy(n, location=0, scale=1)
u <- runif(n, min=0, max=1)

sum(u < 1/2*exp(-(x**2-1)/2)*(1+x**2))/n

newx <- x[u<1/2*exp(-(x**2-1)/2)*(1+x**2)]
hist(newx,freq=FALSE)
```