

2022년 2학기

# Computational Statistics

## HW#7

222STG10

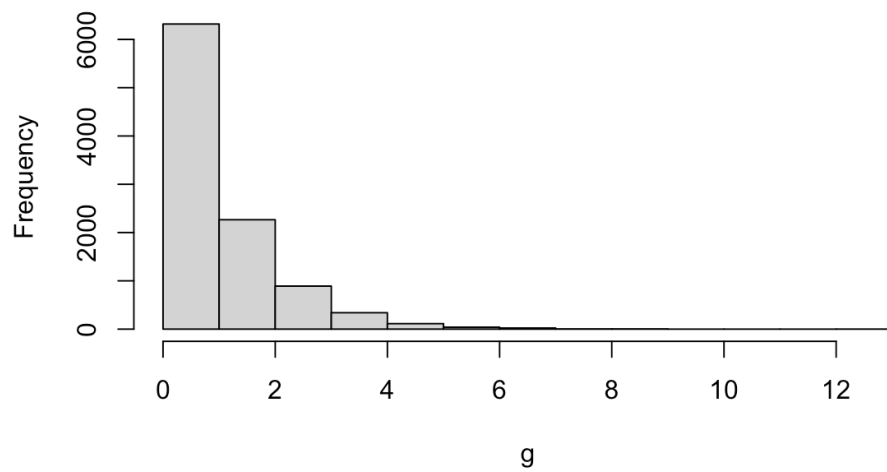
김희숙

## Problem

### 1. Gamma Deviates

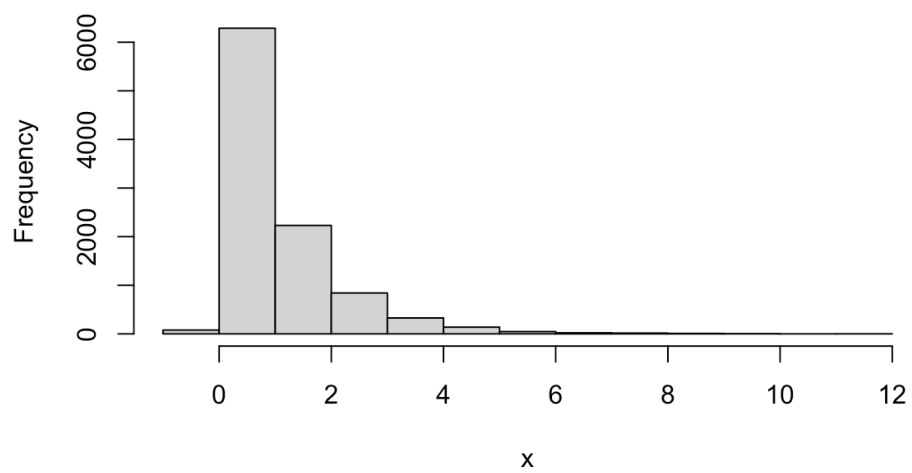
$r=1$ ;  $N = 10000$ ; `rgamma(N,r)`

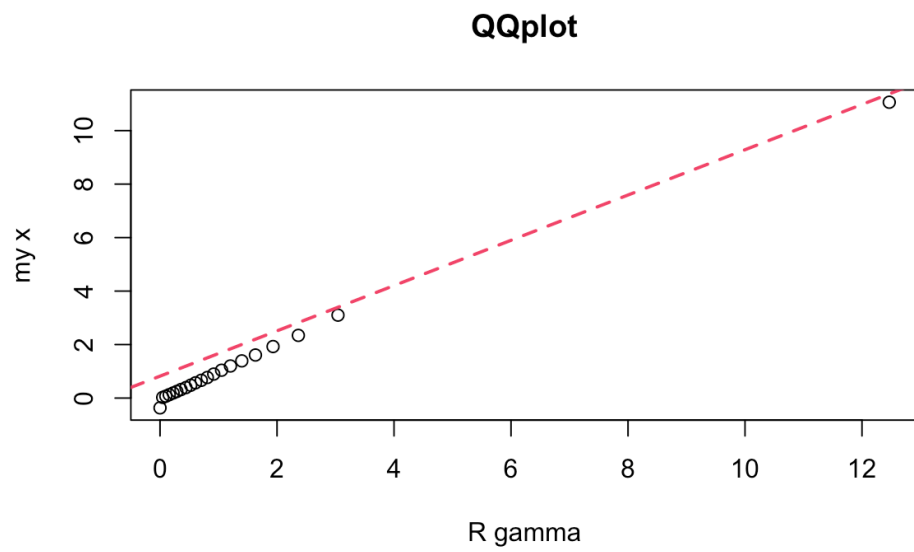
Histogram of g



알고리즘을 통해 발생시킨 r.v

Histogram of x

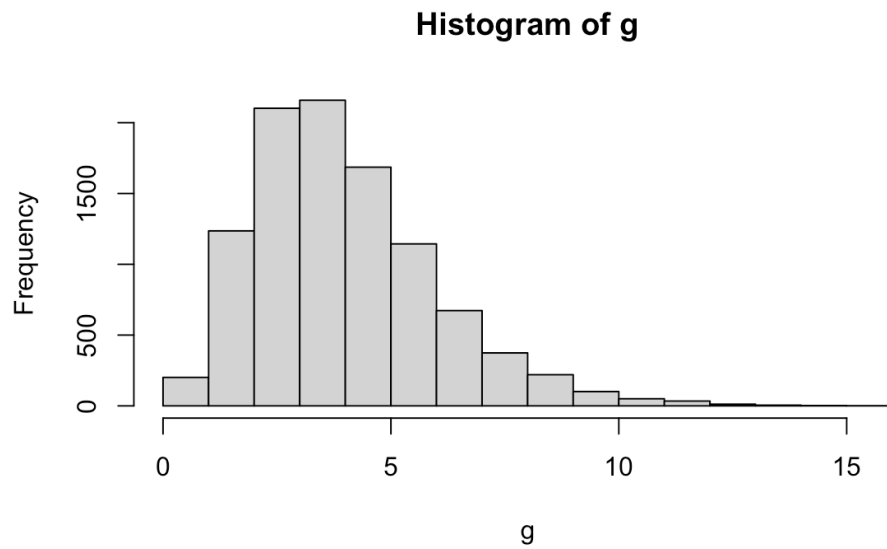




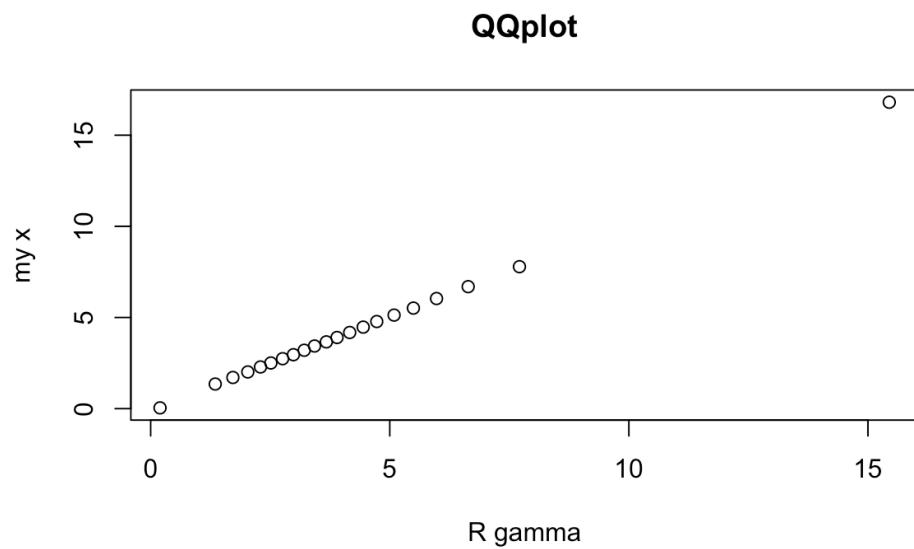
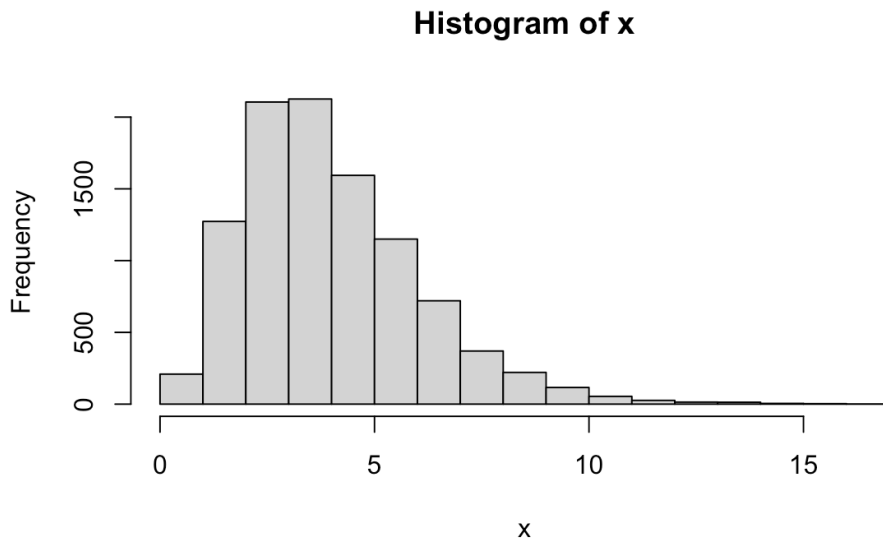
**0.9523%로 Accepted.**

감마 분포를 통해 발생시킨 값과 알고리즘을 통해 발생시킨 r.v의 분포가 유사하며 qqplot을 통해 본 두 r.v 를 보아 거의 동일한 값을 갖는 것을 알 수 있다.

**r=4; N = 10000; rgamma(N,r);**



알고리즘을 통해 발생시킨 r.v



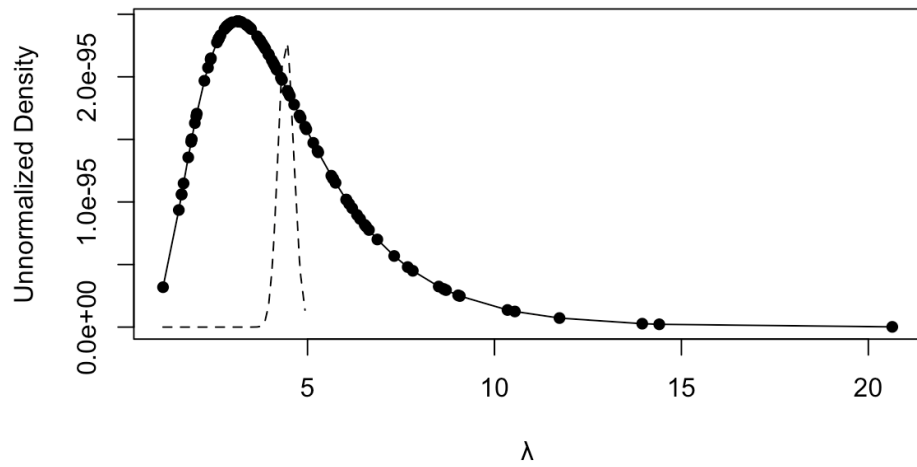
**0.9908%로 Accepted.**

r=1일 때와 마찬가지로 감마를 통해 발생시킨 r.v와 알고리즘을 통해 발생시킨 r.v가 비슷한 분포를 보인다. r=4일땐, 99%로 accepted 되는 것을 확인 할 수 있었다.

## 2. Sampling a Bayesian Posterior

# of total iteration	# of acceptance	# of rejection	acceptance ratio
34090	10000	24090	0.2933412

### Target and Envelope Function

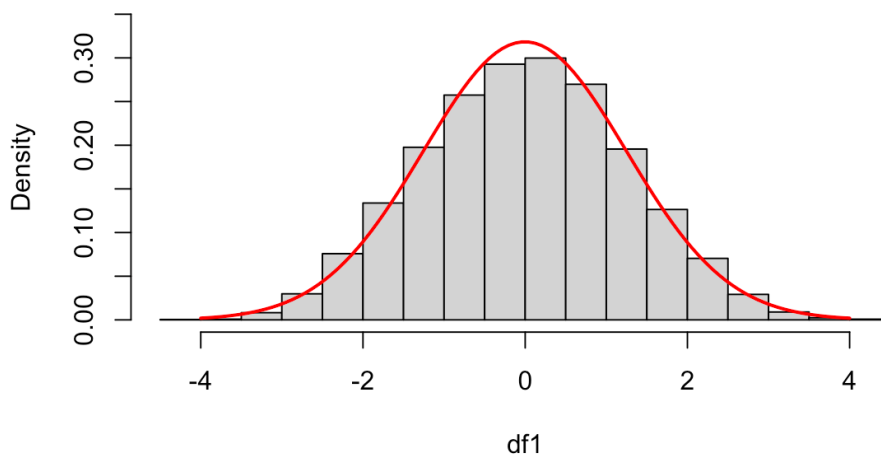


Unnormalized target(dotted)와 envelope(solid) 그래프를 그리면 위의 그림과 같다.

### 3. Slash Distribution

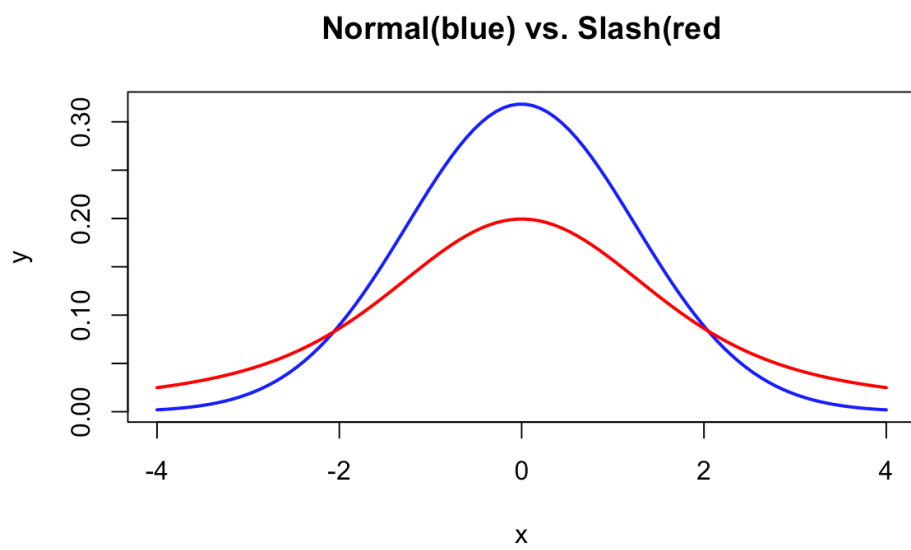
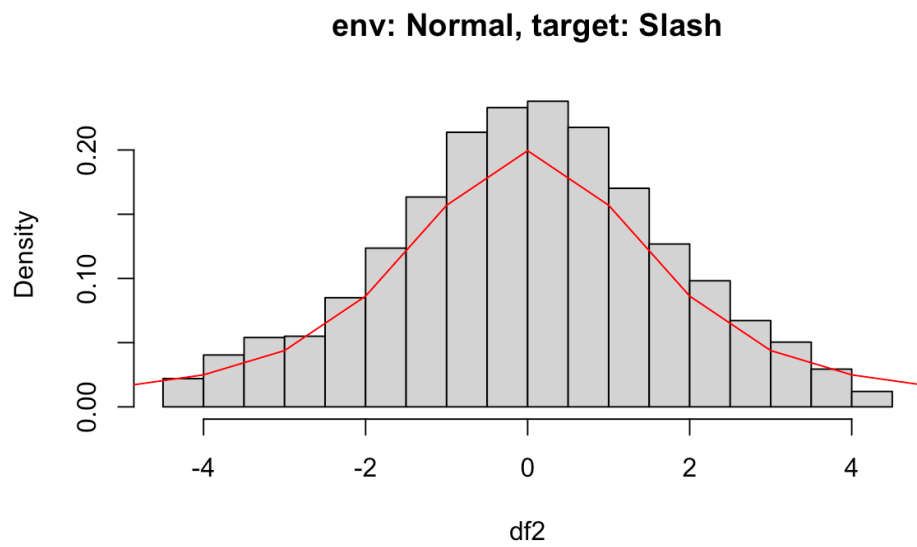
#### 1) env: Slash, target: Normal

#### env: Slash, target: Normal



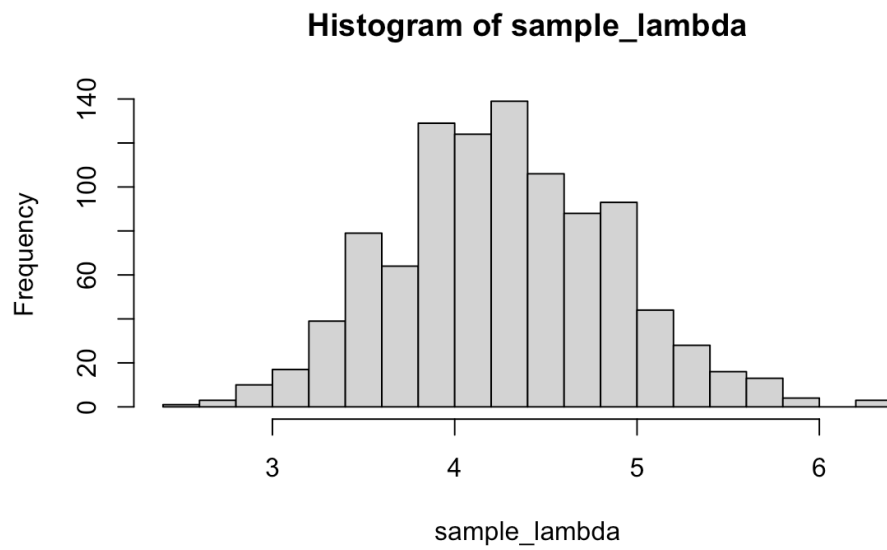
슬래시 분포 envelope을 사용하여 SIR를 통해 얻은 표준 정규로부터 대략적인 도면 히스토그램을 그려보면 위의 그림과 같으며, slash를 통한 샘플링이 낮은 density를 보인다.

2) env: Normal, target: Slash



Slash와 Normal의 그래프를 비교해보면 slash(red line)가 납작한 형태를 띈다.

#### 4. Bayesian Inference



mean(x)	mean(sample_lambda)
4.3	4.269356

Mean을 비교한 결과 유사한 값이 나온 것을 확인 할 수 있었다.

## Code appendix

```
# 1

# Gamma Deviates

set.seed(1108)

N = 10000 #sample size

z <- rnorm(N,mean=0,sd=1)

u <- runif(N,min=0,max=1)

gammaDev <- function(r,z,u,g){

  a = r - 1/3

  b = 1/sqrt(9*a)

  ty <- function(y){ a*(1+b*y)^3 }

  qy <- function(y){ exp(a*log(ty(y)/a)-ty(y)+a)}

  ey <- function(y){exp(-y^2/2)}

  x <- ty(z)

  result = rep(0,length(z))

  result[which(u<=qy(z)/ey(z))] = 1

  prob = sum(result==1)/length(result)

  return(prob)

}

r = 1

g <- rgamma(N,r)

hist(g)

gammaDev(r,z,u,g)

gq <- quantile(g, probs = seq(0, 1, by = 0.05))

xq <- quantile(x, probs = seq(0, 1, by = 0.05))

plot(gq, xq,

      xlab = "R gamma",

      ylab = "my x",

      main = "QQplot")

qqline(xq, col = 2,lwd=2,lty=2)

hist(x) #발생시킨 x

r = 4

g <- rgamma(N,r)

hist(g)

gammaDev(r,z,u,g)
```



```

gq <- quantile(g, probs = seq(0, 1, by = 0.05))
xq <- quantile(x, probs = seq(0, 1, by = 0.05))
plot(gq, xq,
     xlab = "R gamma",
     ylab = "my x",
     main = "QQplot")
hist(x) #발생시킨 x

# 2
# Sampling a Bayesian Posterior
set.seed(1200)

func2 <- function(n, x) {# n: sample 개수  x: observations

  niter <- 0
  reject_num <- 0
  accept_num <- 0
  samples <- c()

  while (accept_num < n) {
    y <- rlnorm(1, log(4), 0.5) # proposal density
    u <- runif(1) # sample for rejection
    value <- prod(dpois(x, y))/prod(dpois(x, 4.3))

    if (u > value) {
      reject_num <- reject_num + 1
    }
    else {
      samples <- append(samples, y)
      accept_num <- accept_num + 1
    }
    niter <- niter+1
  }
  cat(
    "# of total iteration:", niter, "\n",
    "# of acceptance:", accept_num, "\t", "# of rejection:", reject_num, "\n",
    "acceptance ratio:", accept_num/niter)
  return(samples)
}

x <- c(8, 3, 4, 3, 1, 7, 2, 6, 2, 7)
samples <- func2(10^4, x)
N = 100 #sample size

```

```

xx <- c(8, 3, 4, 3, 1, 7, 2, 6, 2, 7)
mu = log(4)
sig = 0.5
lamda.hat = mean(xx)
x <- rpois(N, lamda.hat)
u <- runif(N,min=0,max=1)

loglamda <- rlnorm(N, meanlog = mu, sdlog = sig)

flamda <- function(lamda){
  1/(lamda*sig*sqrt(2*pi)) * exp(-(log(lamda)-mu)^2/(2*sig^2))
}
lik <- function(lamda){
  (exp(-N*lamda) * lamda^(sum(x)))/(prod(factorial(x)))
  #prod(exp(-lamda)*lamda^x/gamma(x + 1))
}
elamda <- function(lamda){
  flamda(lamda)*lik(lamda=4.3)
}
qlamda <- function(lamda){
  flamda(lamda)*lik(lamda)
}
result = rep(0,length(loglamda))
result[which(u<(qlamda(loglamda)/elamda(loglamda)))] = 1 #accept
prob = sum(result==1)/length(result)

plot(loglamda, elamda(loglamda), xlab="λ", ylab="Unnormalized Density",pch=16, main = "Target
and Envelope Function")
lines(loglamda[order(loglamda)], elamda(loglamda)[order(loglamda)],pch=16)
lines(loglamda[order(loglamda)], qlamda(loglamda)[order(loglamda)],pch=16,lty=2)

# 3
# Slash Distribution
dslash <- function(y){
  ifelse(x==0, 1/(2*sqrt(2*pi)), (1-exp((-y^2)/2))/(y^2*sqrt(2*pi)))
}

m = 100000
n = 10000

# 1. env: Slash, target: Normal

```

```

set.seed(1)
x = sample(x=seq(-10, 10, 0.001), replace=T, size=m)

f = dnorm(x)
g = dslash(x)

IW = (f/g)/(sum(f/g))

df1 = sample(x, size=n, replace=T, prob=IW)

hist(df1, freq=FALSE, ylim=c(0, 0.35), main='env: Slash, target: Normal')

x = seq(-4, 4, length.out=100)
y = dnorm(x, mean(df1), sd(df1))
lines(x, y, col='red', lwd=2)

# 2. env: Normal, target: Slash
set.seed(1)
x = rnorm(m)
f = dslash(x)
g = dnorm(x)

IW = (f/g)/(sum(f/g))

df2 = sample(x, size=n, replace=T, prob=IW)

hist(df2, freq=FALSE, main='env: Normal, target: Slash', breaks=20)
x = seq(-10, 10)
y = dslash(x)
lines(x, y, col='red')

#####
x = seq(-4, 4, length.out=100)
y = dnorm(x, mean(df1), sd(df1))
plot(x, y, col='blue', lwd=2, type = 'l', main= "Normal(blue) vs. Slash(red)")

y2 = dslash(x)
lines(x, y2, col='red', lwd=2)

# 4
# Bayesian Inference

```

```
xx = c(8, 3, 4, 3, 1, 7, 2, 6, 2, 7)
lambda = rlnorm(10000, log(4), 0.5)

lik = c()
for (i in 1:length(lambda)) {
  lik = append(lik, prod(dpois(xx, lambda[i])))
}
IW = lik/sum(lik)
sample_lambda = sample(lambda, 1000, replace=TRUE, prob=IW)
hist(sample_lambda)
```