STP598-Assignment 3

Hao Wang
hwang306

October 24, 2017

1 Question 4

Monte Carlo integration

$$\theta = \int_0^{0.5} e^{-x} dx$$

let $f(x) = \frac{1}{0.5-0}$ then

$$\theta = (0.5 - 0) \int_0^{0.5} e^{-x} \frac{1}{0.5 - 0} dx = 0.5E[e^{-x}]$$

then

$$\hat{\theta} = 0.5Eg(X) = 0.5E(e^{-x})$$

The sample mean variance is

$$Var(\hat{\theta}) = 0.25/m * Var(g(x)) = \frac{0.25}{m} Var(e^{-x})$$

```
# generate unif(0, 0.5)
m <- 10000
theta <- numeric(m)
x <- runif(m, min = 0, max = 0.5)
theta.hat <- mean(exp(-x))*0.5
theta.hat
```

```
## [1] 0.3930036
```

```
# variance of theta.hat
for (i in 1:m) {
  theta[i] <- exp(-x[i])
}
variance <- 0.25/m * var(theta)
variance
```
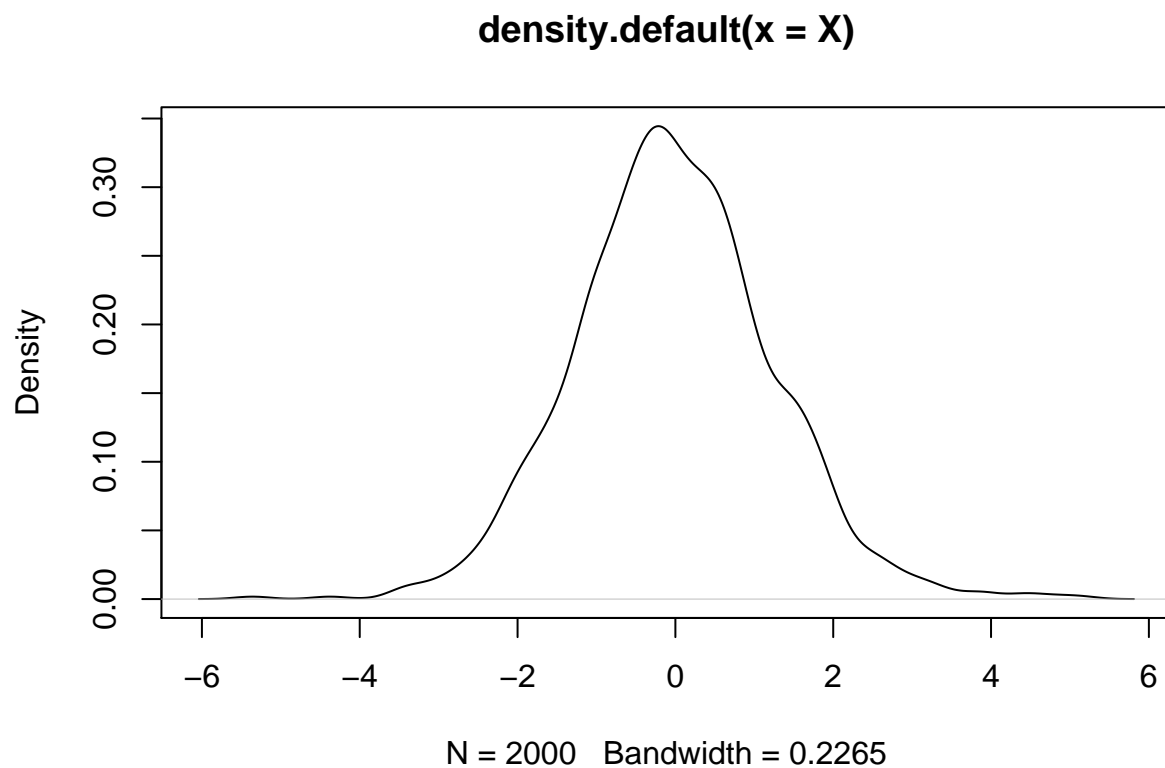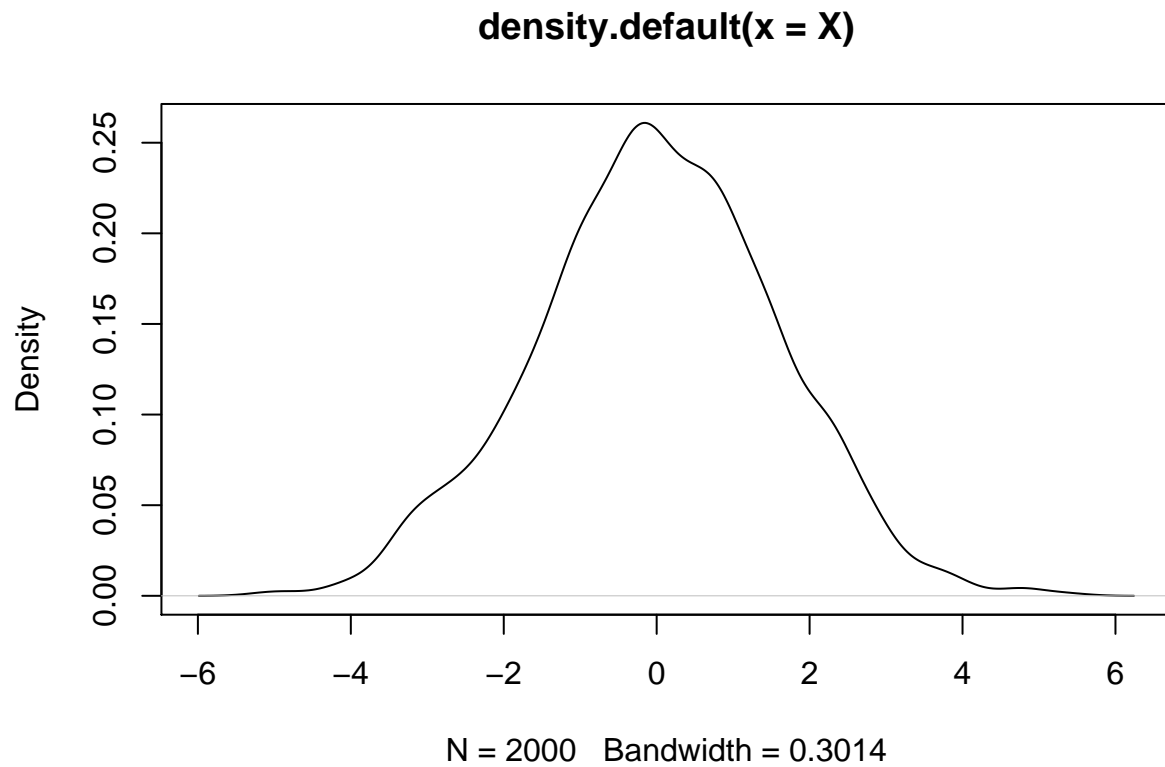
```
## [1] 3.196495e-07
```

2 Question 5

mixture of normal distribution

```r
library(MASS)
# write function
loc.mix.0 <- function(n, p, mu1, mu2, Sigma1, Sigma2) {
#generate sample from BVN location mixture
X <- matrix(0, n, 2)
for (i in 1:n) {
k <- rbinom(1, size = 1, prob = p)
if (k)
X[i,] <- mvrnorm(1, mu = mu1, Sigma1) else
X[i,] <- mvrnorm(1, mu = mu2, Sigma2)
}
return(X)
}

X <- loc.mix.0(1000, 0.75, 0, 0 ,1, 3)
plot(density(X))
```
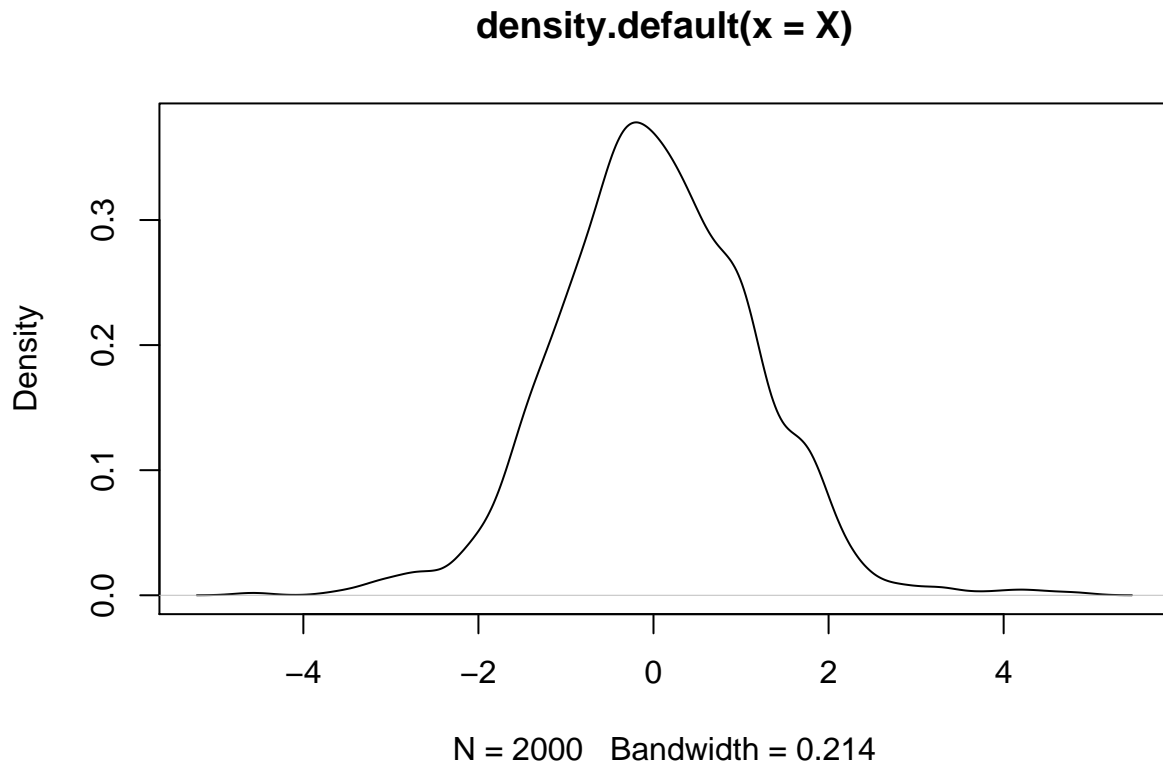
density.default(x = X)

N = 2000   Bandwidth = 0.2265

```
X <- loc.mix.0(1000, 0.1, 0, 0 ,1, 3)
plot(density(X))
```

density.default(x = X)

N = 2000   Bandwidth = 0.3014

```
X <- loc.mix.0(1000, 0.9, 0, 0 ,1, 3)
plot(density(X))
```

density.default(x = X)

N = 2000   Bandwidth = 0.214

I changed multiple p values, this does not look like a bimodal distribution. I think it is because the mu1 = mu2 = 0. The parts with highest probability are overlapping, thus the mixture looks like a normal distribution.

3   Question 6
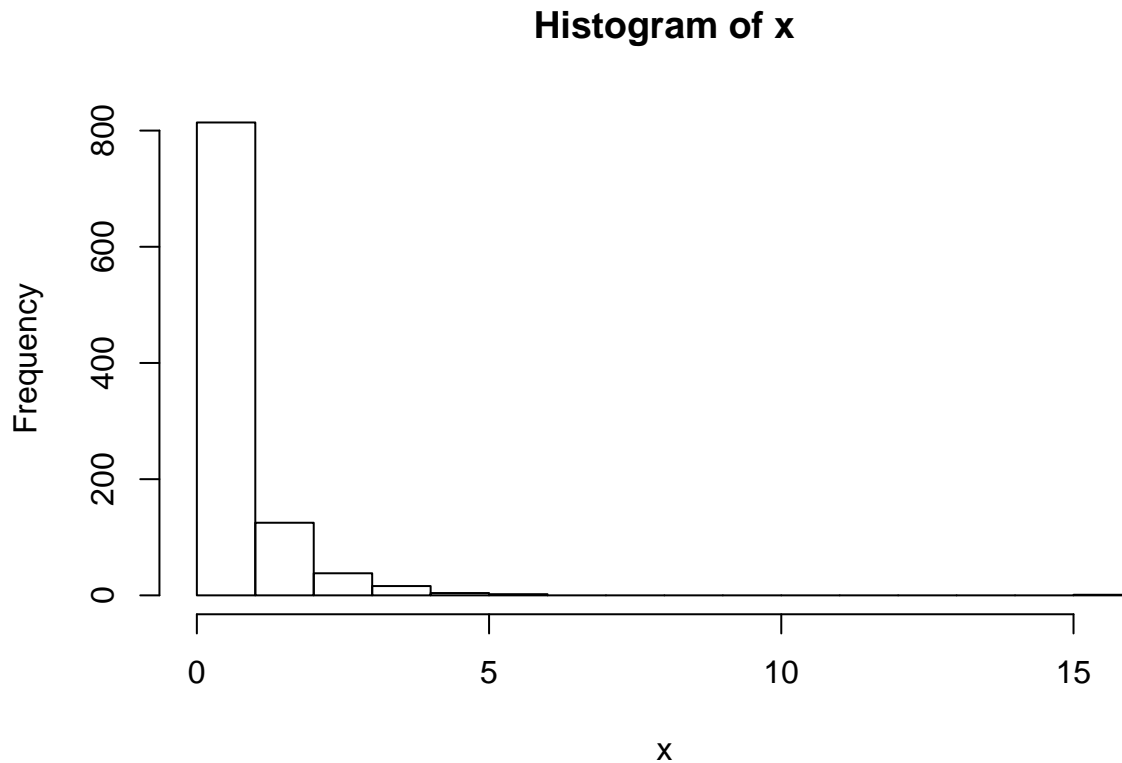
Gamma-Exponetial mixture.

$$\Lambda \sim \Gamma(r, \beta)$$
$$(Y|\Lambda = \lambda) \sim exp(\Lambda)$$

```r
# generate distributions of Lambda
n <- 1000 #1000 random variables
r <- 4
beta <- 2
lambda <- rgamma(n, r, beta)

#now apply the sample of lambda as the exponetial
```

```
x <- rexp(n, lambda)
hist(x)
```

Histogram of x

4   Question 7

requires to generate 500 random obsevations of the 3-dimensional MVN, using Choleski factorization method.

write a function of factorization method

```
rmvn.Choleski <-
function(n, mu, Sigma){
# generate n random vectors from MVN(mu, Sigma)
# dimension is inferred from mu and Sigma
d <- length(mu)
Q <- chol(Sigma) # Choleski factorization of Sigma
Z <- matrix(rnorm(n*d), nrow = n, ncol = d)
X <- Z %*% Q + matrix(mu, n, d, byrow = TRUE)
X
}
```

```
# write input mean matrix and covriance matrix

mu <- c(0,1,2)
Sigma <- matrix(c(1, -0.5, 0.5, -0.5, 1, -0.5, 0.5, -0.5, 1),
                nrow = 3, byrow = TRUE)

X <- rmvn.Choleski(500, mu, Sigma)
pairs(X)
```
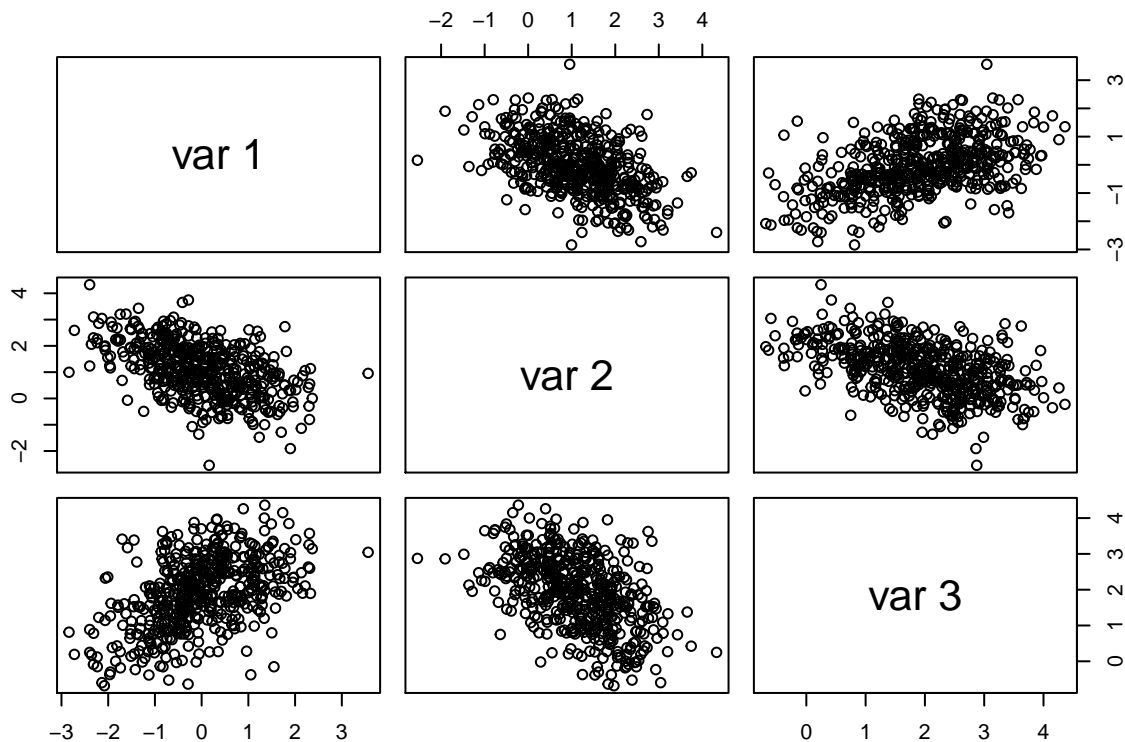

From the graph, it satisfies the theoretical expectation. The joint distribution of each pair of marginal distributions is theoretically bivariate normal. var1 and var2 are negatively correlated, and var1 and var 3 are postively correlated.