第二次上机作业

斯蓬

2024-10-24

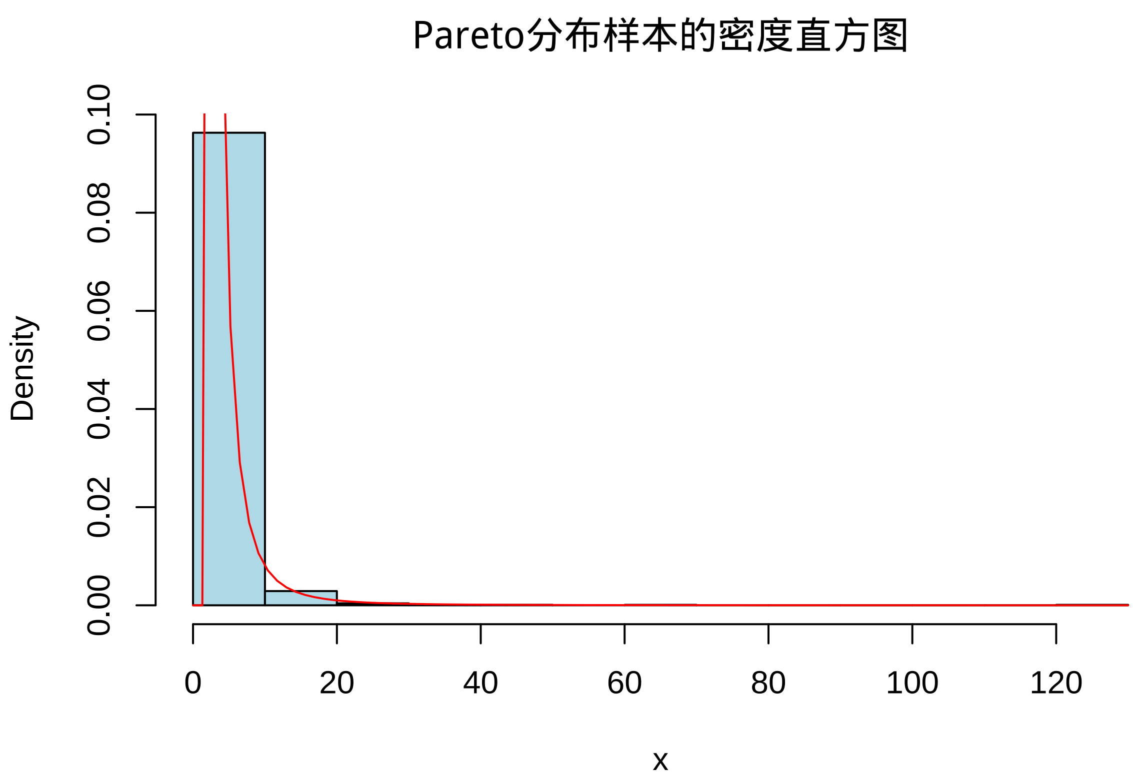
目录

1 第一章课后习题编程部分

1.1 Task1.1

使用逆变换抽样法

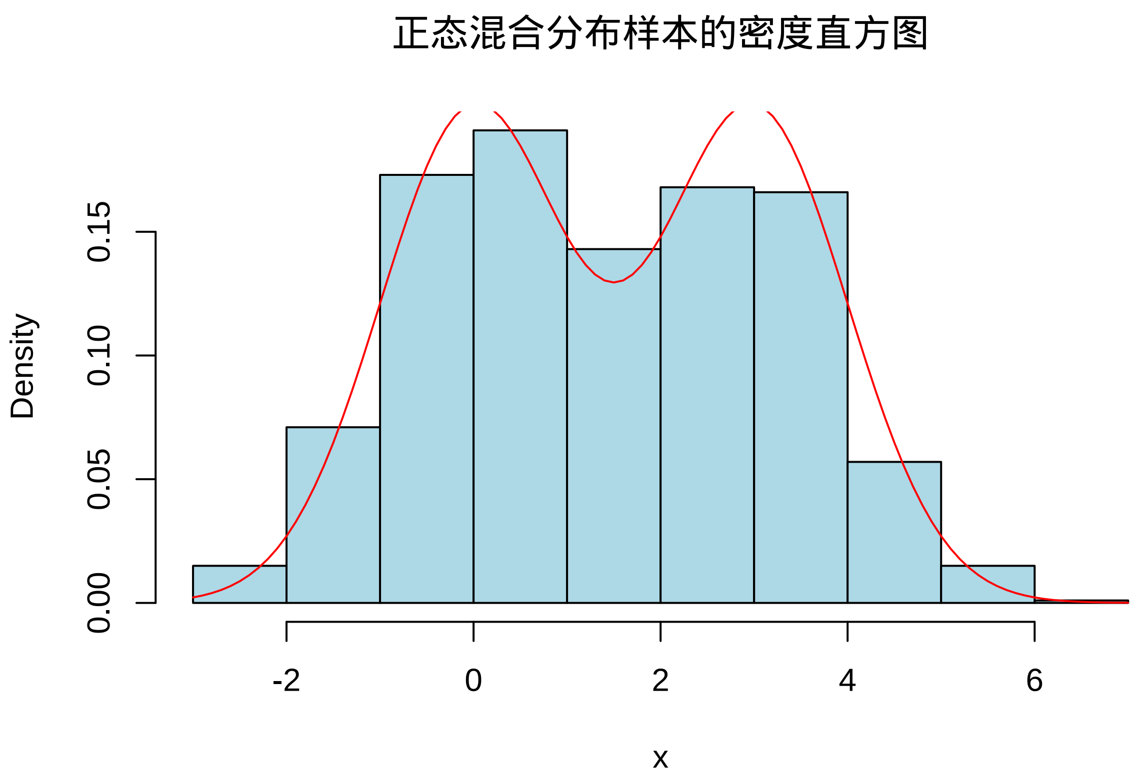
```
library(VGAM)

# 设置参数
a <- 2
b <- 2
n <- 1000  # 样本量

# 逆变换法生成 Pareto 分布随机样本
u <- runif(n)  # 生成 [0,1] 均匀分布的随机数
x <- b / (u^(1/a))  # 逆变换公式

# 绘制密度直方图和使用 VGAM 包中的 dpowerpareto 函数绘制 Pareto 分布密度曲线
hist(x, probability = TRUE, main = "Pareto 分布样本的密度直方图", col = "lightblue")
curve(dpareto(x, shape = a, scale = b), col = "red", add = TRUE)
```



1.2 Task1.2

```
# 设置参数
n <- 1000
p1 <- 0.5  # N(0, 1) 的权重
mu1 <- 0  # N(0, 1) 的均值
sigma1 <- 1  # N(0, 1) 的标准差
mu2 <- 3  # N(3, 1) 的均值
sigma2 <- 1  # N(3, 1) 的标准差

# 生成正态混合分布的随机样本
z <- rbinom(n, 1, p1)  # 生成 0/1 变量
```

```
x <- z * rnorm(n, mu1, sigma1) + (1 - z) * rnorm(n, mu2, sigma2)

# 绘制密度直方图和密度曲线
hist(x, probability = TRUE, main = " 正态混合分布样本的密度直方图", col = "lightblue")
curve(p1 * dnorm(x, mu1, sigma1) + (1 - p1) * dnorm(x, mu2, sigma2), col = "red", add = TRUE)
```

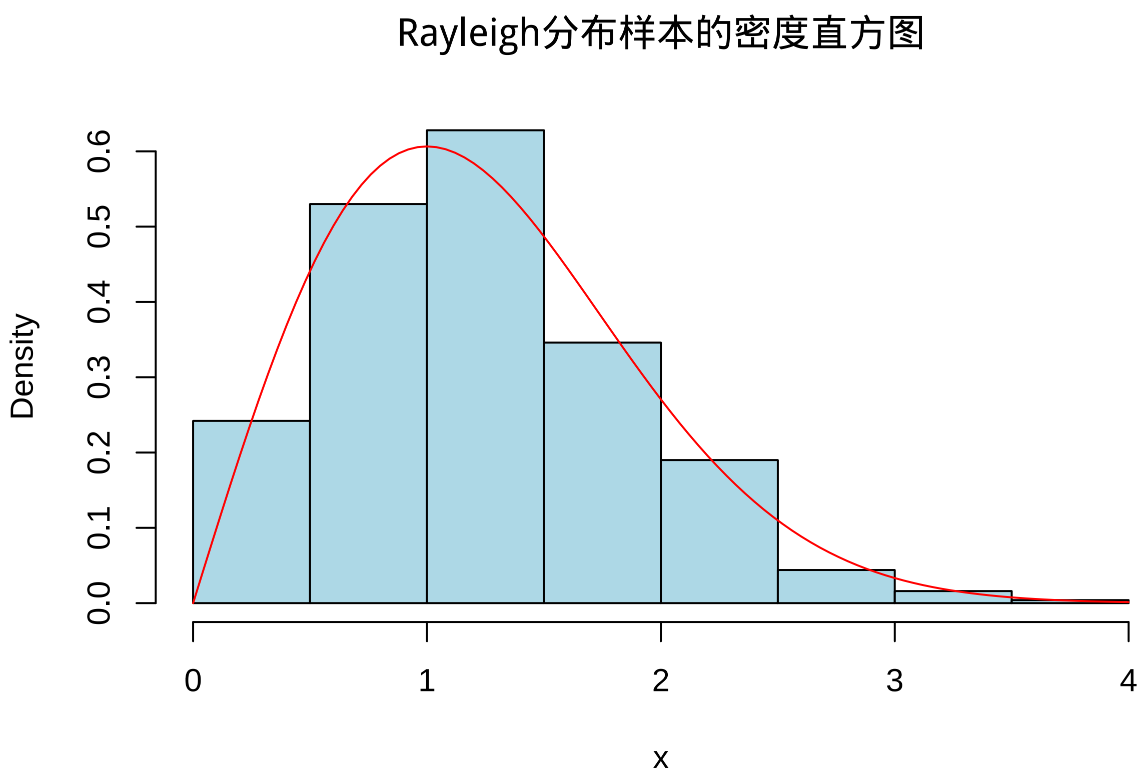


从图中我们可以发现，直方图和曲线大致拟合。当 n 趋向于无穷的时候，直方图所表示的应该和曲线一致。

1.3 Task1.3

```
# 设置参数
sigma <- 1
```

```r
n <- 1000  # 样本量

# 生成 Rayleigh 分布的随机样本
x <- sqrt(-2 * sigma^2 * log(runif(n)))

# 绘制密度直方图和密度曲线
hist(x, probability = TRUE, main = "Rayleigh 分布样本的密度直方图", col = "lightblue")
curve((x / sigma^2) * exp(-x^2 / (2 * sigma^2)), col = "red", add = TRUE)
```

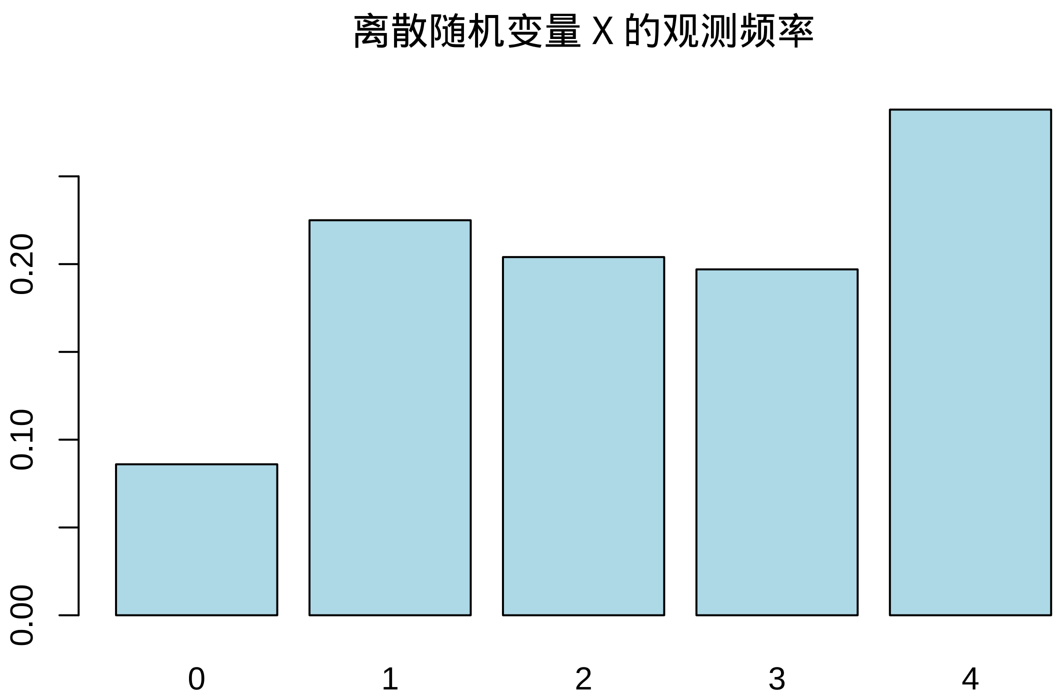


1.4 Task1.4

```
# 定义累积概率函数 (CDF)
cdf <- c(0.1, 0.3, 0.5, 0.7, 1.0)  # 对应的累计概率

# 自定义函数生成随机变量 X
generate_random_variable <- function(n) {
  x <- numeric(n)  # 存储生成的随机变量
  for (i in 1:n) {
    u <- runif(1)  # 生成一个 [0, 1] 均匀分布的随机数
    # 使用累积概率分布来判断随机变量的取值
    if (u <= cdf[1]) {
      x[i] <- 0
    } else if (u <= cdf[2]) {
      x[i] <- 1
    } else if (u <= cdf[3]) {
      x[i] <- 2
    } else if (u <= cdf[4]) {
      x[i] <- 3
    } else {
      x[i] <- 4
    }
  }
  return(x)  # 返回生成的随机变量
}

# 生成 1000 个随机变量样本
n <- 1000
samples <- generate_random_variable(n)

# 打印观测频率
observed_freq <- table(samples) / n
print(observed_freq)
```

```
## samples
##     0     1     2     3     4
## 0.086 0.225 0.204 0.197 0.288
```

```
# 绘制频率直方图
barplot(observed_freq, main = " 离散随机变量 X 的观测频率", col = "lightblue")
```

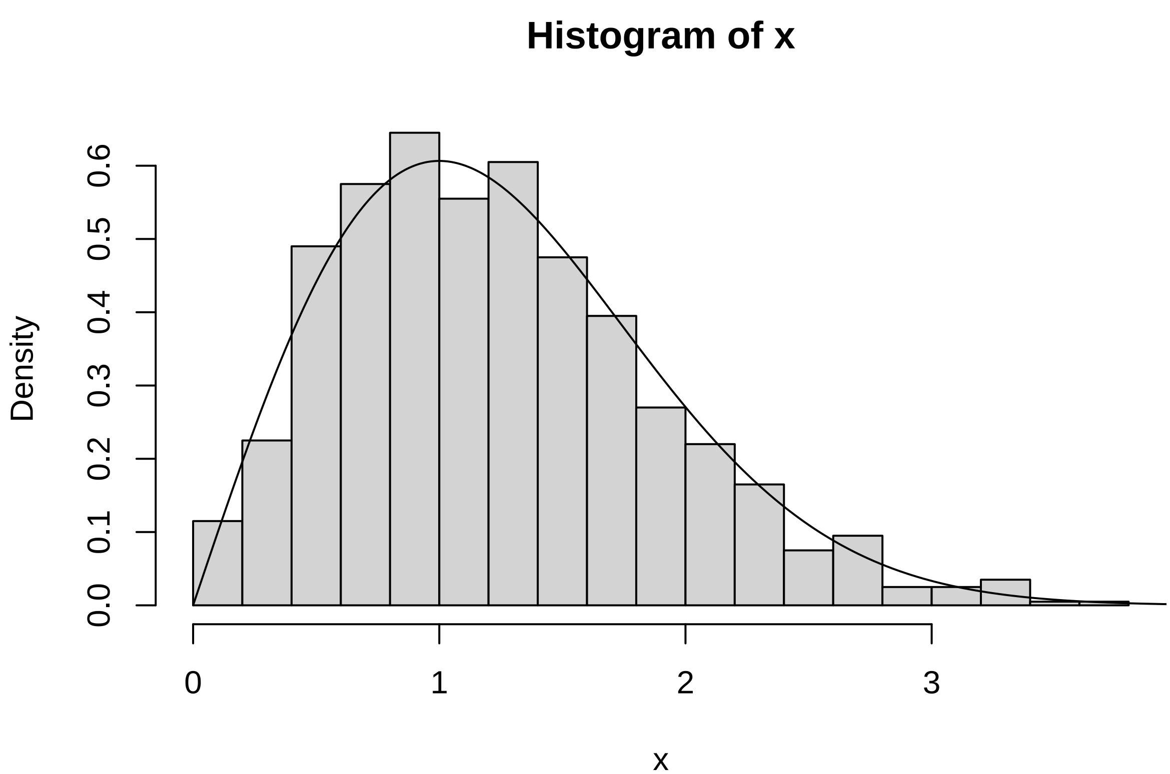


从图中我们可以看出，我们通过算法来实现的离散随机变量和给出的概率是符合的。

2 第一章例题复现部分

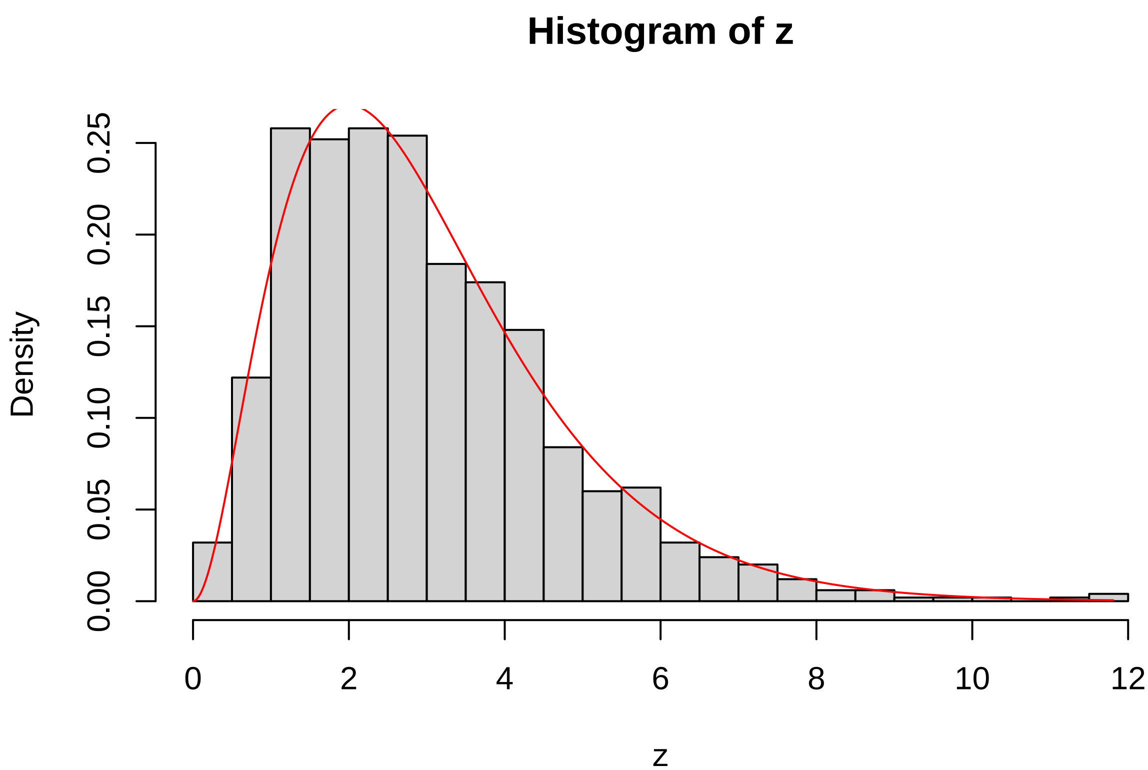
2.1 1.1.1

```
set.seed(220810332)
n = 1000
sigma=1
r = matrix(runif(n),n,1)
x =sqrt(-2*sigma^2*log(1-r))
x0=seq(0,4,0.01)
fx=x0/(sigma^2)*exp(-x0^2/(2*sigma^2))
hist(x,20,freq=F)
lines(x0,fx,'l')
```



2.2 1.1.5

```
set.seed(220810332)
func_h = function(x,alpha){
  if(alpha < 1 ){
    print("ERROR_PARAMETER") # 当 alpha<1 时该抽样不适用
  }
  else (exp(alpha-1)/(alpha^(alpha-1))*x^(alpha-1)
        *exp(-(alpha-1)/alpha*x))}
n = 1000
m = 0
i = 0;alpha = 3
z = c()
repeat{
  m = m+1;
  x = rexp(1,1/alpha)
  y = runif(1)
  ty = func_h(x,alpha)
  if (y<=ty){
    z = c(z,x)
    i = i+1}
  if (i == n) break}
x0=seq(0,max(z),0.01)
fx=dgamma(x0,alpha,1)
hist(z,30,freq=F)
lines(x0,fx,'l',col='red')
```



2.3 1.2.1

```
set.seed(220810332);
X=c(1,2.5,3.5,5,6);
prob = c(0.1,0.2,0.3,0.2,0.2)
Fx=cumsum(prob)
n = 10000;xi=NULL
for (i in 1:n){
  r=runif(1,0,1)
  if (r<=prob[1]){
    xi[i]=X[1]
  }else{xi[i]=X[min(which(Fx>=r))]}}
```

```
x=sample(c(1,2.5,3.5,5,6),size=n,replace=TRUE,prob=prob)
table(xi)/n
```

```
## xi
##      1    2.5    3.5      5      6
## 0.1006 0.2038 0.2962 0.1986 0.2008
```

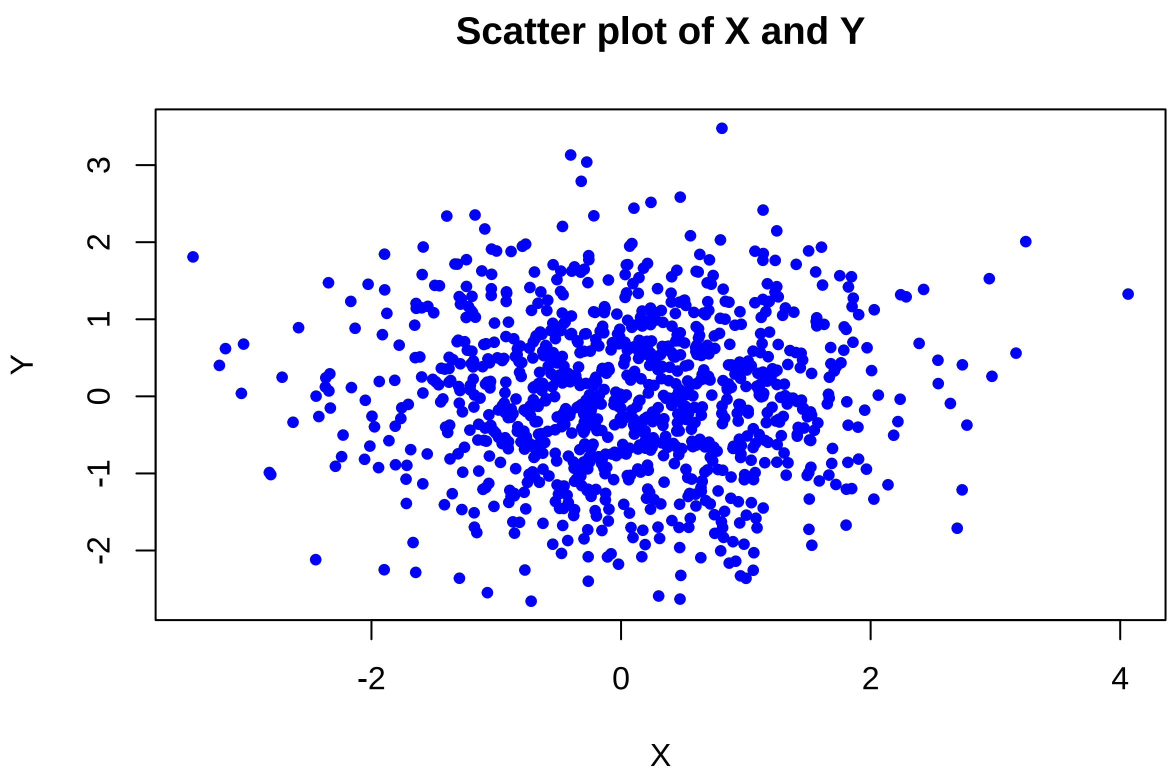
```
table(x)/n
```

```
## x
##      1    2.5    3.5      5      6
## 0.1009 0.1960 0.3051 0.2044 0.1936
```

3 第二章课后习题编程部分

3.1 Task2.1

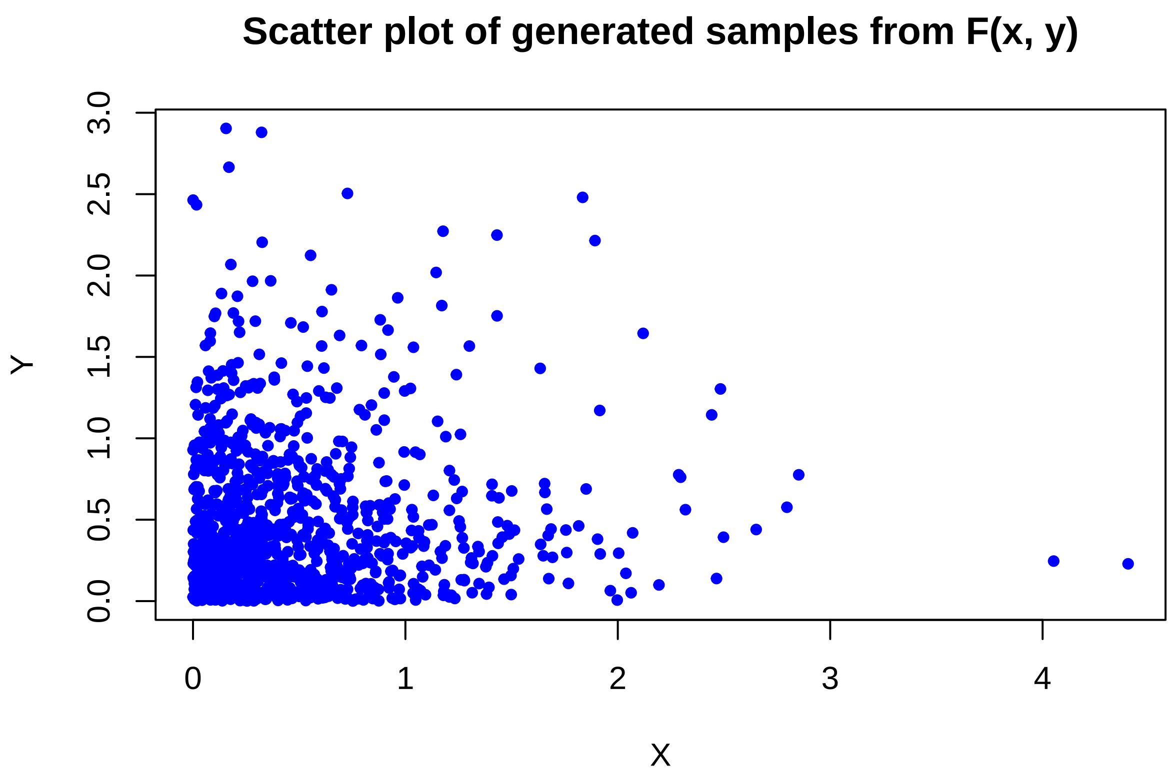
```
# 设置随机样本数量
n <- 1000

# 生成  和 R 的随机样本
alpha <- runif(n, 0, 2 * pi)  # 从 U(0, 2）均匀分布中采样
R <- rexp(n, rate = 0.5)  # 从 Exp(0.5) 指数分布中采样

# 计算 X 和 Y
X <- sqrt(R) * cos(alpha)
Y <- sqrt(R) * sin(alpha)

# 生成 X 和 Y 的散点图
plot(X, Y, main = "Scatter plot of X and Y",
     xlab = "X", ylab = "Y", col = "blue", pch = 20)
```



3.2 Task2.2

```
# 设置样本数量
n <- 1000

# 定义二维分布的联合概率密度函数（PDF）
F_joint <- function(x, y) {
  1 - exp(-2 * x) - exp(-2 * y) + exp(-2 * (x + y))
}

# 生成从二维分布中采样的随机数
x <- rexp(n, rate = 2)  # 从边缘分布 Exp(2) 中采样
```

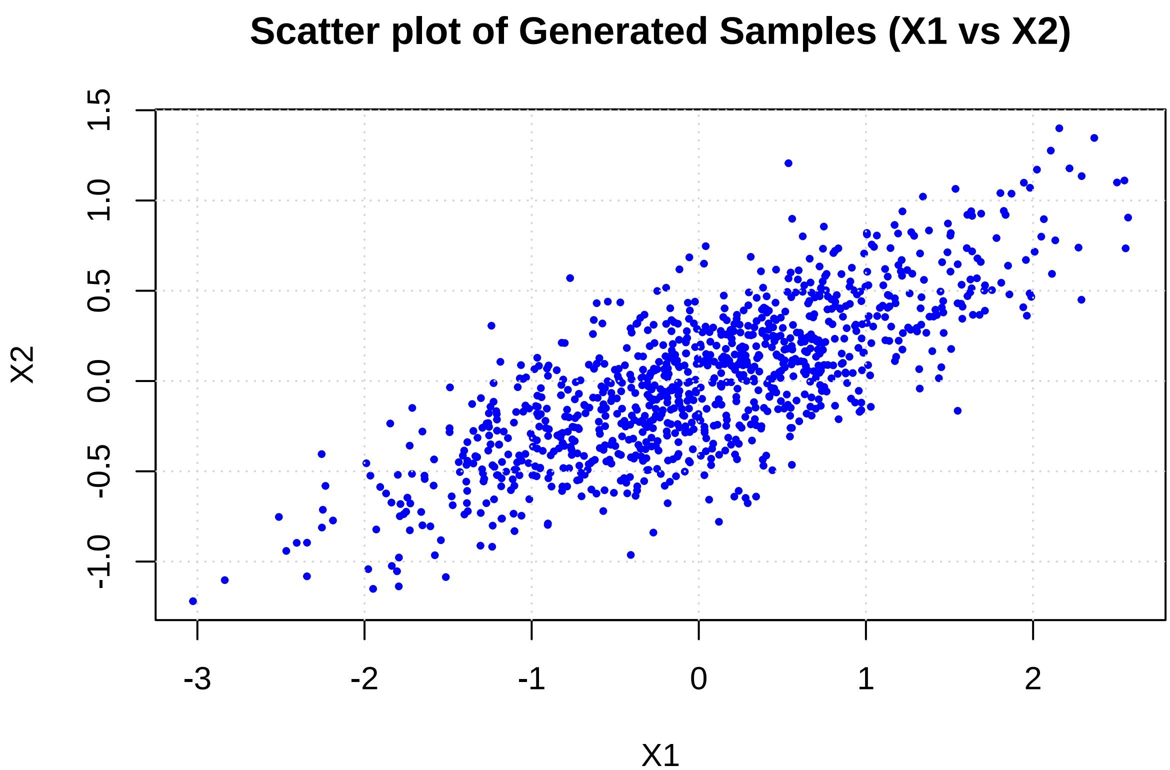
```
y <- rexp(n, rate = 2)  # 从边缘分布 Exp(2) 中采样

# 绘制二维分布的散点图
plot(x, y, main = "Scatter plot of generated samples from F(x, y)",
     xlab = "X", ylab = "Y", col = "blue", pch = 20)
```



4 第二章例题复现部分

4.1 2.1.4

```r
set.seed(220810332)
rep=1000
xi=matrix(0,rep,2)
mu1=mu2=0
sigma11=1
sigma22=0.2
rho=0.8
sigma12=sigma21=rho*sqrt(sigma11*sigma22)
sigma_c=sigma11-sigma12*sigma22^(-1)*sigma21
for (i in 1:rep){
  x2=rnorm(1,mu2,sigma22^0.5)
  mu_c=mu1+sigma12*sigma22^(-1)*(x2-mu2)
  x1=rnorm(1,mu_c,sigma_c^0.5)
  xi[i,]=c(x1,x2)
}
plot(xi[, 1], xi[, 2],
     main = "Scatter plot of Generated Samples (X1 vs X2)",
     xlab = "X1", ylab = "X2",
     col = "blue", pch = 20,
     cex = 0.6)## 绘制散点图来展示结果
grid()
```


Scatter plot of Generated Samples (X1 vs X2)
X2
1.5
1.0
0.5
0.0
-0.5
-1.0
-3
-2
-1
0
1
2
X1