SC-HW1

ID : 111024517 Name : 鄭家豪

due on 02/27

(1) Weibull distribution:

- Sampling by inverse CDF method:

  $F(x) = 1 - \exp(-(x/\theta)^{\beta}) \sim U(0,1)$, by Probability intergral transformation.

  Do some calculation,we have

  $F^{-1}(y) = \theta \times (\ln \frac{1}{1-y})^{1/\beta}$,where $Y \sim U(0,1)$.

- Algorithm:

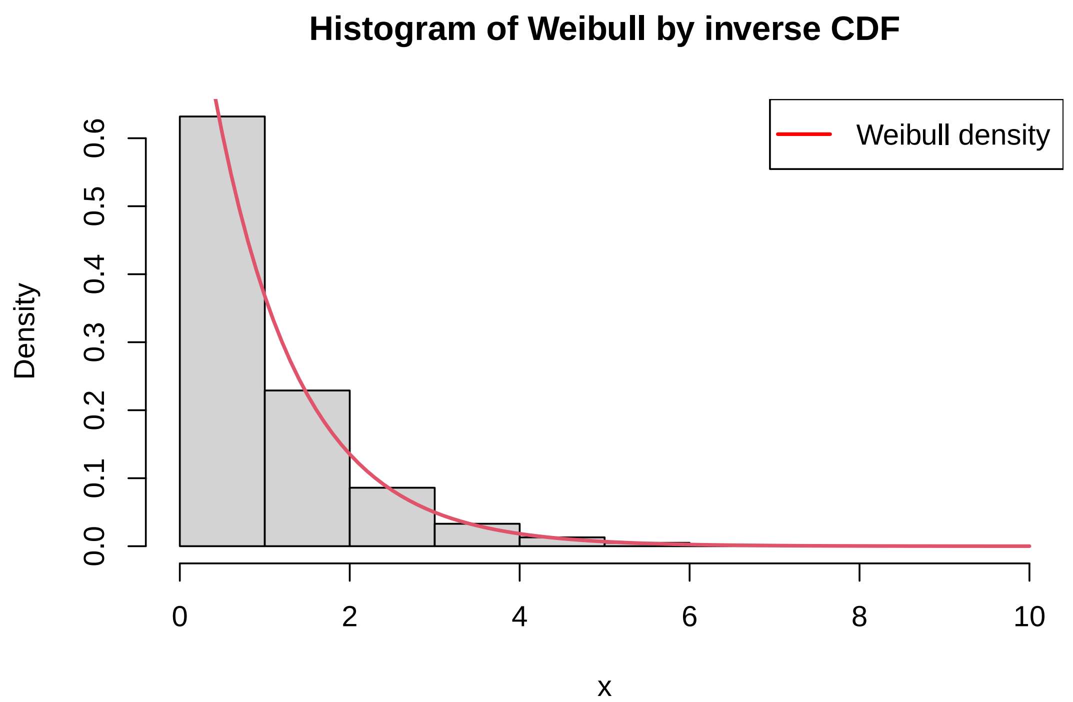
  step 1: fixed (θ, β),generate $U \sim U(0,1)$

  step 2: $X = F^{-1}(U) \sim \text{Weibull}(\theta, \beta)$

```
Weibull <- function(n,scale,shape){
  x <- c()
  set.seed(1)
  for(i in 1:n){
    x[i] <- scale*(log(1/(1-runif(1))))^(1/shape)
  }
  hist(x,freq = F,
       main="Histogram of Weibull by inverse CDF")
  curve(dweibull(x,shape,scale),col=2, lwd=2, add=T)
  legend("topright",legend = c("Weibull density"),
       col=c("red"), lty=c(1), pch=c(-1), lwd=2)
  table <- data.frame("efficiency(empirical)" = length(x)/n,
                      "efficiency(theoretical)" = 1)
  return(kable(table,row.names = F))
}
```

- n=1000;(θ, β) = (1,1) (i.e. Exp(1)):

```
Weibull(1000,1,1)
```



efficiency.empirical.	efficiency.theoretical.
1	1

由於這裡採用，Weibull random variable 由 Uniform random variable 藉由 inverse CDF method 轉換得來的，所以 empirical and theoretical efficiency 皆等於 1，都不受 parameters (θ, β) 所影響。由以上的例子驗證此結論。

(2) Pareto distribution:

- Sampling by inverse CDF method:
  $F(x) = \int_0^x \frac{\beta}{\theta}(1+t/\theta)^{-(\beta+1)}dt = 1-(1+x/\theta)^{-\beta} \sim U(0,1)$, by Probability intergral transformation.

  Do some calculation,we have
  $F^{-1}(y) = \theta \times [(\frac{1}{1-y})^{1/\beta} - 1]$,where $Y \sim U(0,1)$.

- Algorithm:

  step 1: fixed (θ, β),generate $U \sim U(0,1)$

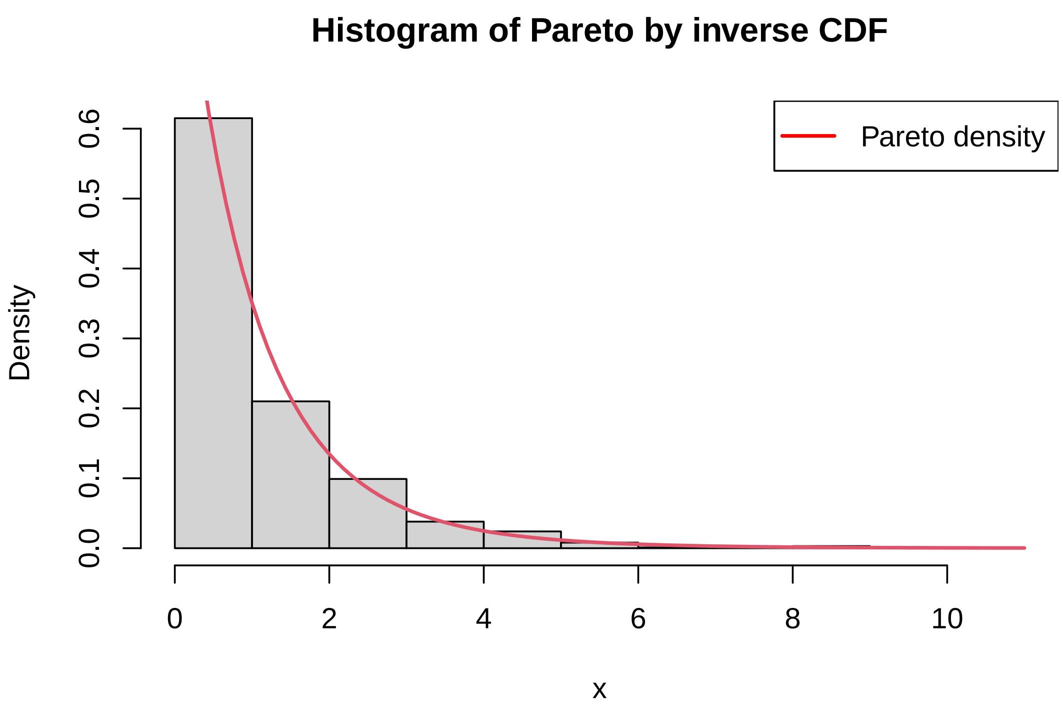
  step 2: $X = F^{-1}(U) \sim \text{Pareto}(\theta, \beta)$

```
pareto <- function(n,scale,shape){
  x <- c()
  set.seed(2)
```

```
  for(i in 1:n){
    x[i] <- scale*(1/(1-runif(1))^(1/shape)-1)
  }
  f = function(x,scale,shape){
    y <- c()
    for(i in 1:length(x)){
      y[i] <- shape/(scale*(1+x[i]/scale)^(shape+1))
    }
    return(y)
  }
  hist(x,freq = F,
       main="Histogram of Pareto by inverse CDF")
  curve(f(x,scale,shape), col=2, lwd=2, add=T)
  legend("topright",legend = c("Pareto density"),
       col=c("red"), lty=c(1), pch=c(-1), lwd=2)
  table <- data.frame("efficiency(empirical)" = length(x)/n,
                          "efficiency(theoretical)" = 1)
  return(kable(table,row.names = F))
}
```

- n=1000;(θ,β) = (10,10) :

```
pareto(1000,10,10)
```



efficiency.empirical.	efficiency.theoretical.
1	1

由於這裡採用，Pareto random variable 由 Uniform random variable 藉由 inverse CDF method 轉換得來的，所以 empirical and theoretical efficiency 皆等於 1，都不受 parameters (θ, β) 所影響。由以上的例子，驗證此結論。

(3) skewed distribution I:

- Rejection sampling:

Do some simplification to $f(x)$:

$$f(x) = \begin{cases} \dfrac{2r^2}{1+r^2}\phi_1(x) = f_1(x) & \text{if } x \geq 0 \\ \dfrac{2}{1+r^2}\phi_2(x) = f_2(x) & \text{if } x < 0 \end{cases} \text{, where } \begin{cases} \phi_1 : \text{pdf of } N(0, r^2) \\ \phi_2 : \text{pdf of } N(0, 1/r^2) \end{cases}$$

For any x:

$$\frac{f_1(x)}{f_2(x)} = r^2 \times \frac{1/r}{r} \times \exp(-\frac{x^2}{2}(\frac{1}{r^2} - r^2)) = \begin{cases} \geq 1 \text{ if } r \geq 1 \\ \in (0,1) \text{ if } r \in (0,1) \end{cases}$$

Set:

$$\text{proposal } g(x) = \begin{cases} \phi_1(x) \text{ , if } r \geq 1 \\ \phi_2(x) \text{ , if } r \in (0,1) \end{cases} \text{, } x \in \mathbb{R}$$

$$M = M(r) = \begin{cases} \dfrac{2r^2}{1+r^2} \text{ , if } r \geq 1 \\ \dfrac{2}{1+r^2} \text{ , if } r \in (0,1) \end{cases}$$

* Algorithm:

step 1: fixed r,generate $Y \sim g$ and $U \sim U(0,1)$, Y and U are indep;

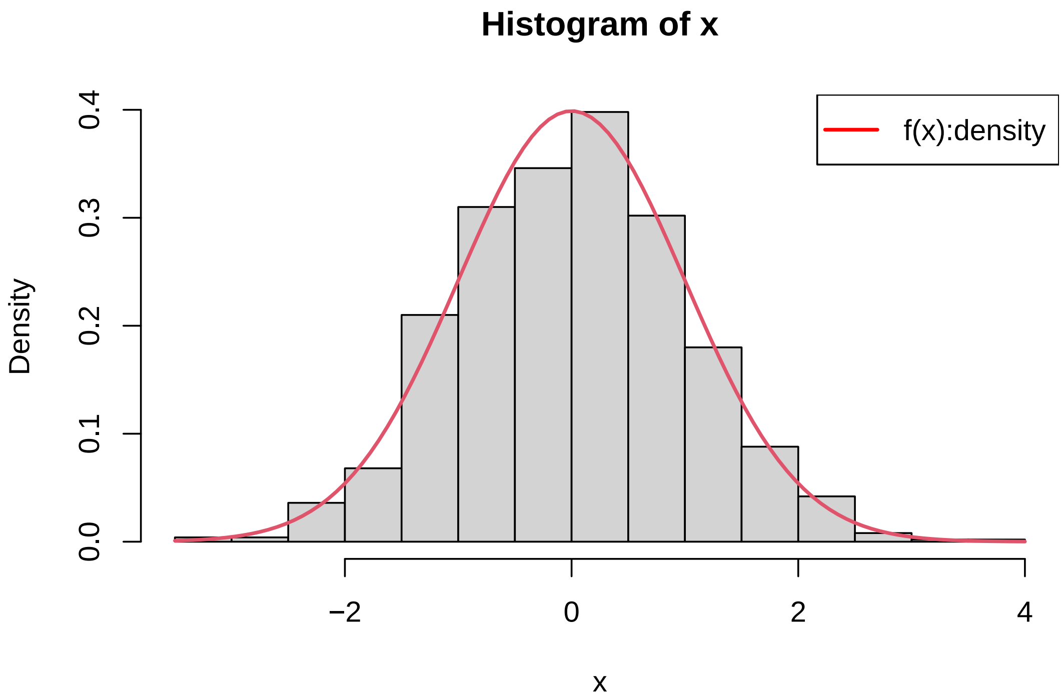
step 2: let X=Y if $U \leq \dfrac{f(y)}{M(r) \times g(y)}$,and reject the draw otherwise.

```
rej_3 <- function(n,r){
  M = function(r){
    z = 0
    if (r >=1){ z = 2*r^2/(1+r^2)}
    else { z = 2/(1+r^2)}
    return(z)
  }
  f = function(x,r){
```

```r
    y = c()
    for (i in 1:length(x)){
      if (x[i] <0) {y[i] <- 2/(1+r^2)*dnorm(x[i],mean = 0,sd = 1/r)}
      else{y[i] <- 2*r^2/(1+r^2)*dnorm(x[i],mean = 0,sd = r)}
    }
    return(y)
  }
  g = function(x,r){
    y <- c()
    if (r >=1){y <- dnorm(x,mean = 0,sd = r)}
    else{y <- dnorm(x,mean = 0,sd = 1/r)}
    return(y)
  }
  set.seed(3)
  y <-c()
  if (r >= 1){y <- rnorm(n,mean = 0,sd = r)}
  else{y <- rnorm(n,mean = 0,sd = 1/r)}
  u <- runif(n,0,1)
  x = y[u <= f(y,r)/(M(r)*g(y,r))]
  hist(x, 20, probability=T)
  curve(f(x,r), col=2, lwd=2, add=T)
  legend("topright",legend = c("f(x):density"),
       col=c("red"), lty=c(1), pch=c(-1), lwd=2)
  table <- data.frame("efficiency(empirical)" = length(x)/n,
                      "efficiency(theoretical)" = 1/M(r))
  return(kable(table,row.names = F))
}
```

- n=1000 ; $r = 1$:

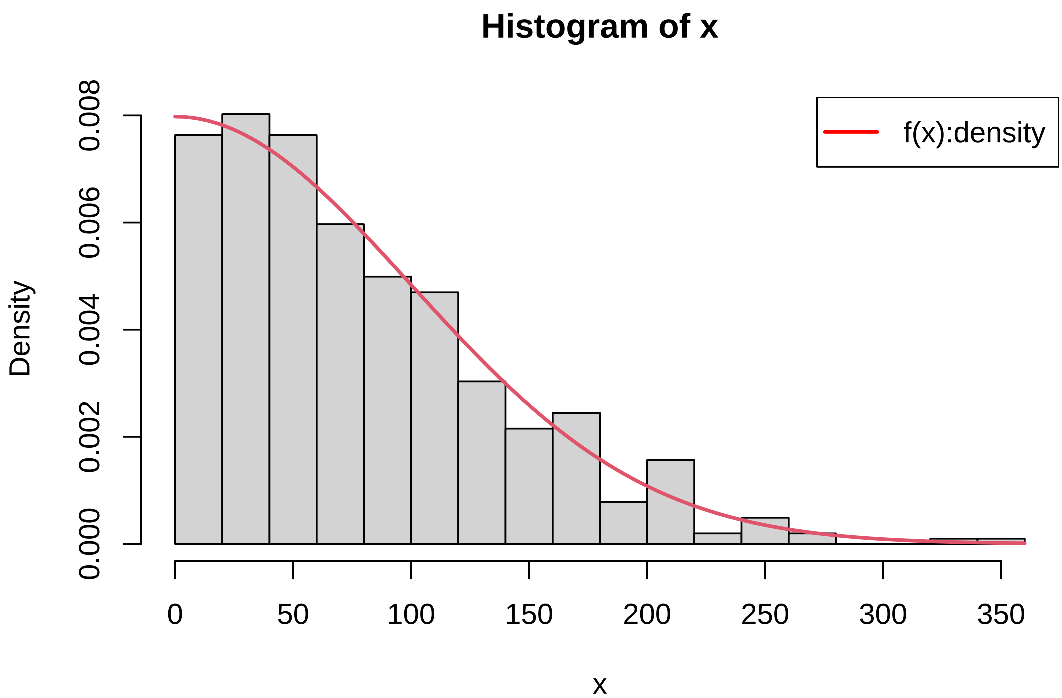
```r
rej_3(1000,1)
```



efficiency.empirical.	efficiency.theoretical.
1	1

- n=1000 ; $r = 100$:

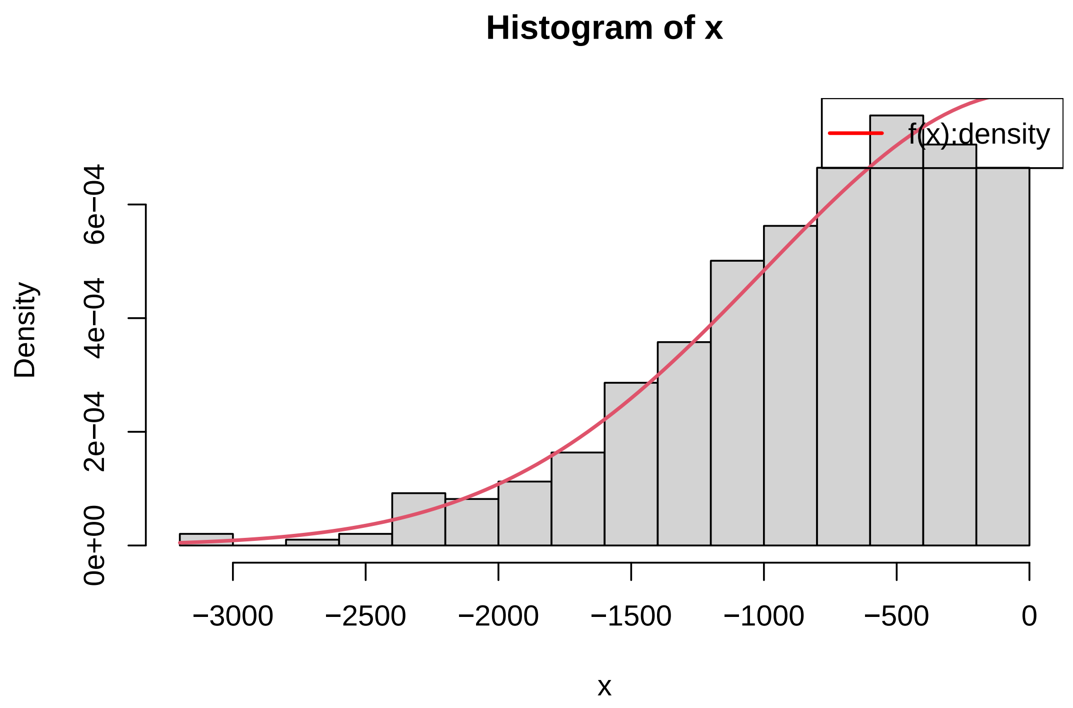
```
rej_3(1000,100)
```



efficiency.empirical.	efficiency.theoretical.
0.511	0.50005

- n=1000 ; $r = 0.001$:

```
rej_3(1000,0.001)
```



efficiency.empirical.	efficiency.theoretical.
0.489	0.5000005

由於 M = M(r) ，為一個 v 的函數，故 theoretical efficiency = 1/M = 1/M(r)，其 efficiency 會受 parameter r 影響。另外，以上任取三個不同的 r ，從所得到的模擬結果，可以觀察出其 empirical efficiency 一致於 theoretical efficiency。由以上的例子，驗證此結論。

(4) skewed distribution II:

- Rejection sampling:

$$
\begin{aligned}
f(x) &= 2h(x)G(\alpha x) \\
&\leq 2h(x) \times 1 \ (\because G(\alpha x) \leq 1, \forall x)
\end{aligned}
$$

Set:

$M = 2$

proposal $g(x) = h(x)$: pdf of the t_v distribution

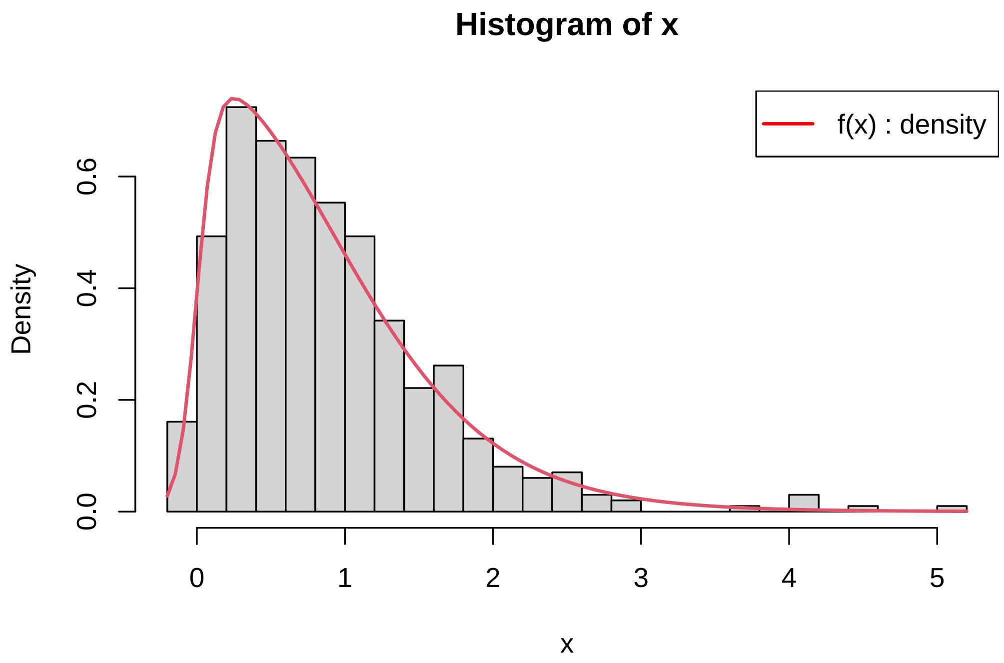
- Algorithm:

  step 1: generate $Y \sim t_v$ and $U \sim U(0,1)$, Y and U are indep;

  step 2: let X=Y if $U \leq \frac{f(y)}{2 \times g(y)}$,and reject the draw otherwise.

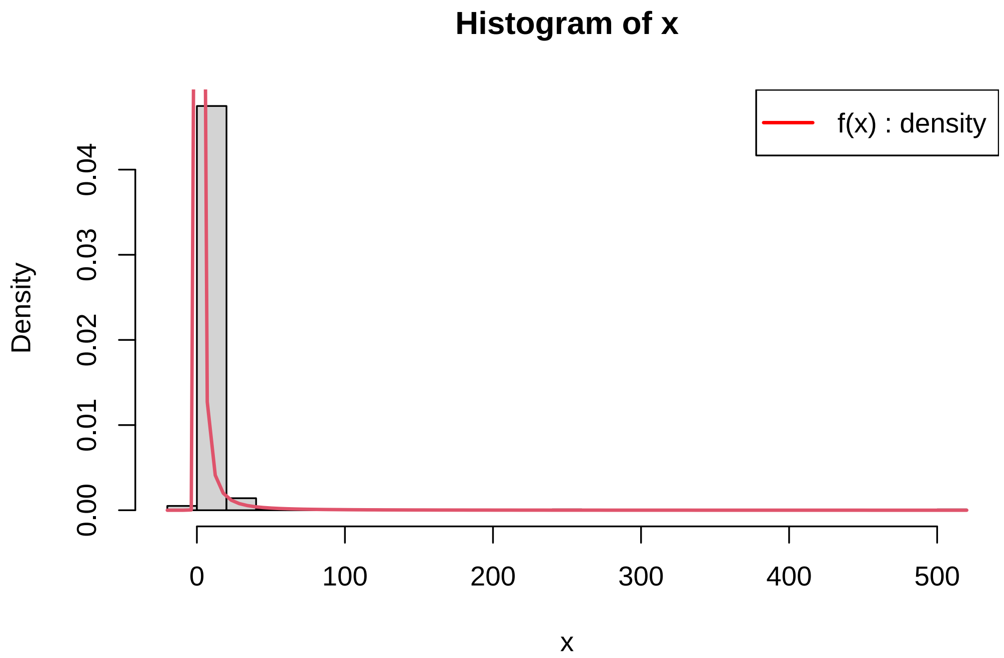
```
rej_4 <- function(n,alpha,v){
  M=2
  f = function(x,alpha,v){
    y = c()
    for(i in 1:length(x)){
      y[i] <- 2*dt(x[i],df = v)*pt(alpha*x[i],df = v)
    }
    return(y)
  }
  #g = dt(x,v)
  set.seed(4)
  u = runif(n,0,1)
  y <- rt(n,df = v)
  x = y[u <= f(y,alpha,v)/(M*dt(x = y,df = v))]
  hist(x, 20, probability=T)
  curve(f(x,alpha,v), col=2, lwd=2, add=T)
  legend("topright",legend = c("f(x) : density"),
       col=c("red"), lty=c(1), pch=c(-1), lwd=2)
  table <- data.frame("efficiency(empirical)" = length(x)/n,
                      "efficiency(theoretical)" = 1/M)
  return(kable(table,row.names = F))
}
```

- n=1000 ; $(\alpha, v) = (10, 10)$:



efficiency.empirical.	efficiency.theoretical.
0.497	0.5

- n=1000 ; $(\alpha, v) = (99, 1)$:



efficiency.empirical.	efficiency.theoretical.
0.495	0.5

由於 M = 2 ，為一個常數，故 theoretical efficiency = 1/M = 0.5，其 efficiency 不受 parameter (α, v) 影響。另外，以上是任取 (α, v) 所得到的模擬結果，可以觀察出其 empirical efficiency 一致於 theoretical efficiency。由以上的例子，驗證此結論。

(5) 2-dimension:

- Gibbs sampling:

We firstly calculate $f(x|y)$ and $f(y|x)$:

$$
\begin{aligned}
&f(x) = \int_0^1 f(x,y)dy = 2(1-x)\int_0^1 (1-y)(1-xy)^{-3}dy = 1.\\
&f(y) = \text{similar to above} = 1.\\
&f(x|y) = 2(1-x)(1-y)(1-xy)^{-3}, \forall x \in (0,1) \text{ for fixed } y.\\
&f(y|x) = 2(1-x)(1-y)(1-xy)^{-3}, \forall y \in (0,1) \text{ for fixed } x.
\end{aligned}
$$

Consider the rejection sampling:

By probability intergral transformation,

$$
\begin{aligned}
Z_X = 1 - F_{X|y}(x) &= \int_x^1 f(t|y)dt\\
&= 2(1-y)[\frac{1}{2y}(1-t)(1-ty)^{-2}|_x^1 + \frac{1}{2y}\int_x^1 (1-ty)^{-2}dt]\\
&= 2(1-y)[\frac{-(1-x)}{2y(1-xy)^2} + \frac{(1-x)}{2y(1-y)(1-xy)}]\\
&= \frac{-(1-x)(1-y)}{y(1-xy)^2} + \frac{1-x}{y(1-xy)}\\
&= (\frac{1-x}{1-xy})^2 \sim U(0,1).\\
&\text{Similarly,} Z_Y = 1 - F_{Y|x}(y) = (\frac{1-y}{1-xy})^2 \sim U(0,1).
\end{aligned}
$$

Do some calculation,we have

$$
\begin{aligned}
X|y &= \frac{\sqrt{Z_X}-1}{y\sqrt{Z_X}-1}.\\
Y|x &= \frac{\sqrt{Z_Y}-1}{x\sqrt{Z_Y}-1}.
\end{aligned}
$$

- Algorithm:

  step 1: Set $x^{(0)}, y^{(0)}$ be the initial value ,both $\in (0,1)$.

  step 2: t=t+1, by inverse CDF method,draw

$$X \sim f(x|y^{(t-1)}), \text{denoted as } x^{(t)}$$

$$Y \sim f(y|x^{(t-1)}), \text{denoted as } y^{(t)}$$

step 3: repeat step 2 until converge.

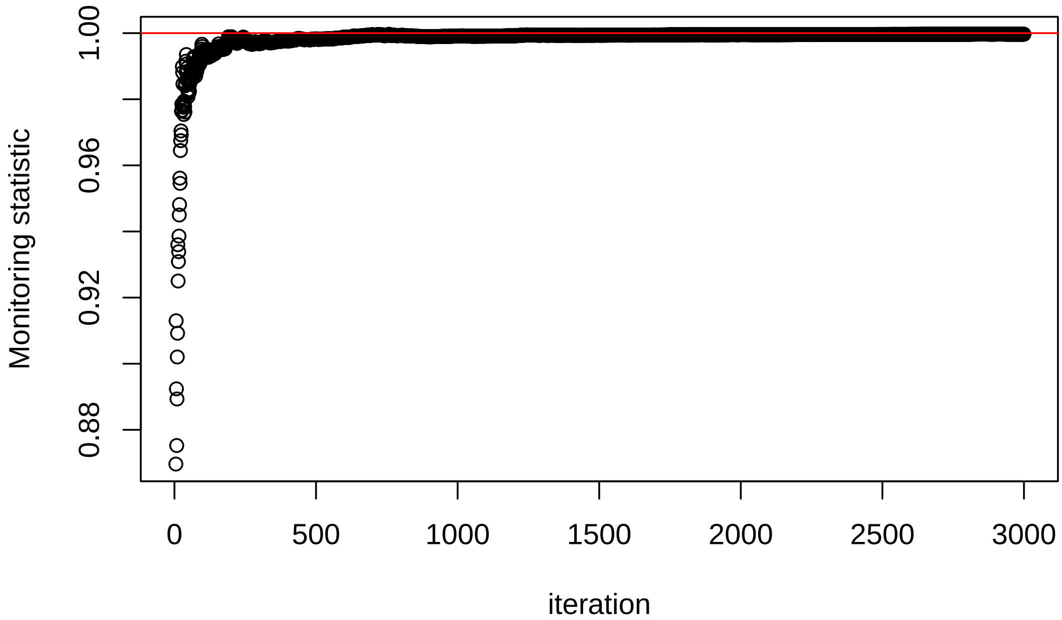
- Check covergence:

1. Set iteration:3000, $x^{(0)} = 0.5, y^{(0)} = 0.5$,generating the samples (X,Y)

```
gibb <- function(init, n.iter){
  xx <- matrix(init,1,2)
  x <- xx[1,1]
  for (i in 1:(2*n.iter)){
    zy <- runif(1,0,1)
    zx <- runif(1,0,1)
    y <- (sqrt(zy)-1)/(sqrt(zy)*x-1)
    xx <- rbind(xx,c(x,y))
    x <- (sqrt(zx)-1)/(sqrt(zx)*y-1)
    xx <- rbind(xx,c(x,y))
  }
  return(xx[2*(1:n.iter)-1,])
}
set.seed(231)
n.iter <- 3000
init <- c(0.5,0.5)
test <- gibb(init=init, n.iter=n.iter)
```

2. Check Monitoring statistic $\sqrt{\hat{R_n}}$:

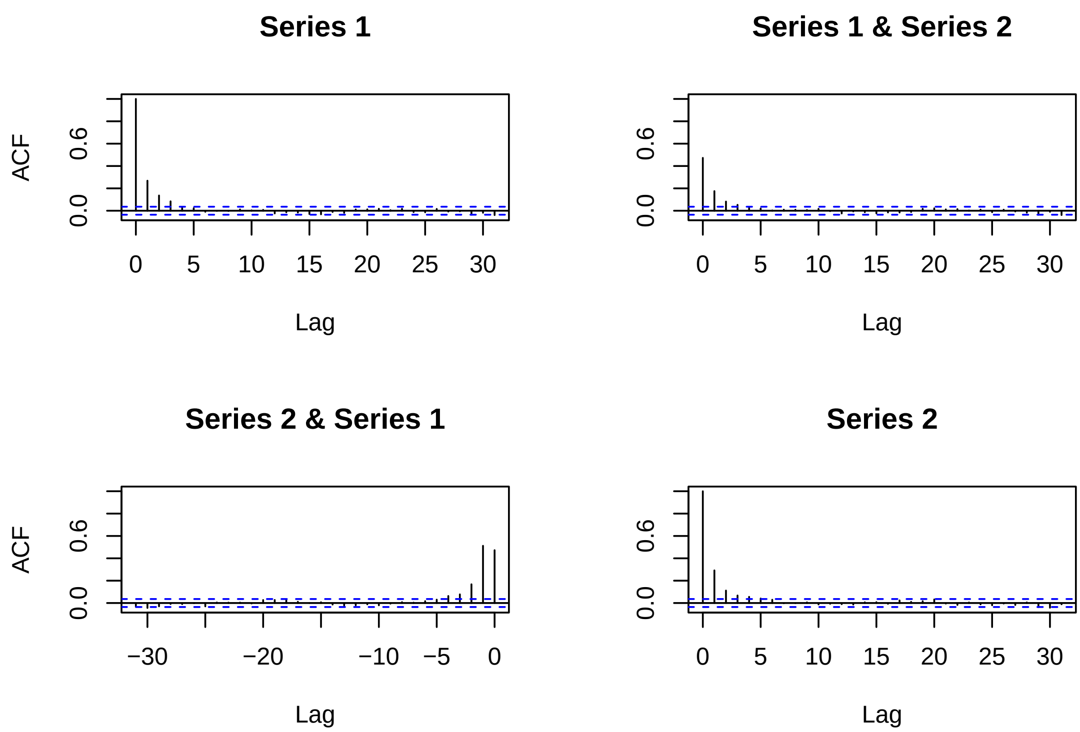
```
gelman.Rhat <- function(mc.draw){
  n.iter <- dim(mc.draw)[1]
  b1 <- w <-rep(0, n.iter)
  for (i in 5:n.iter){
    b1[i]<-var(apply(mc.draw[1:i,],2,mean)) # b1 = b_n/n
    w[i]<-mean(apply(mc.draw[1:i,],2,var))  # within variance: w_n
  }
  Rhat = b1/w+((1:n.iter)-1)/(1:n.iter)
  return(Rhat)
}
plot(gelman.Rhat(test),xlab ="iteration",ylab = "Monitoring statistic")
abline(h=1,col=c("red"))
```



由上可觀察出，$\sqrt{\hat{R}_n} \xrightarrow{\text{n is large}} 1$，根據理論，所有的抽樣服從一個穩健分布。

- Check serial correlations:

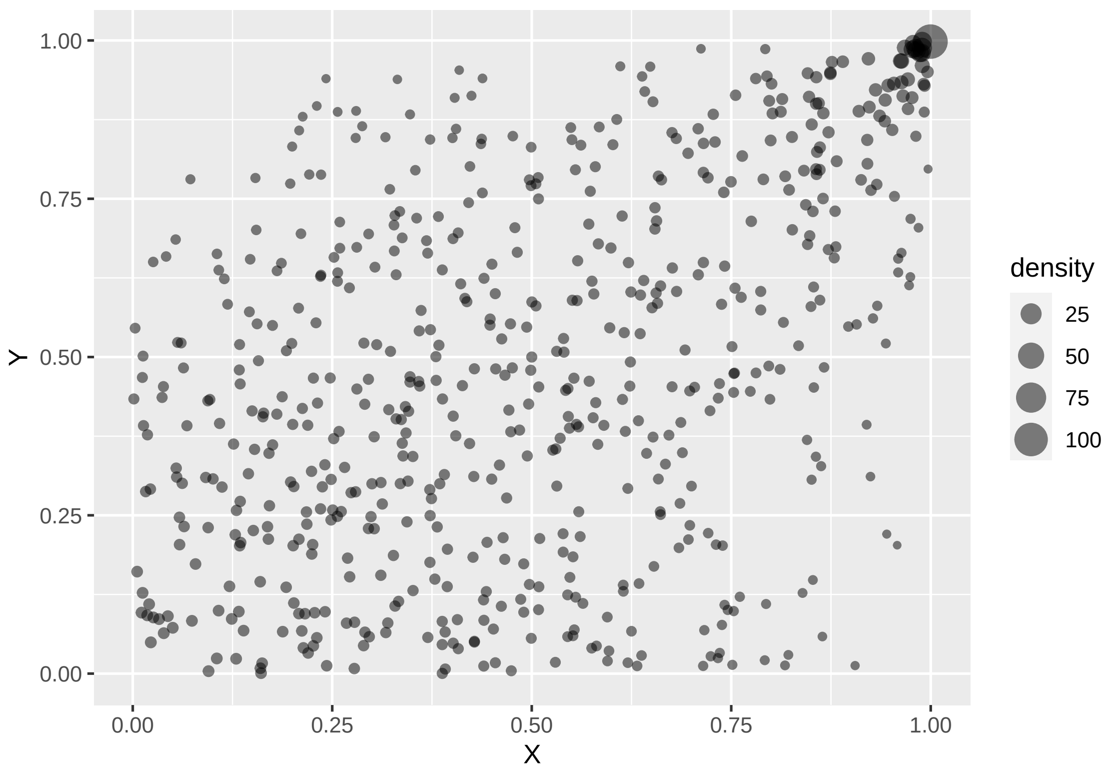
```
acf(test)
```



由上可以看出，大約每 5 個 lag 其 autocorrelation 趨於 0，於是這裡從 3000 個模擬出來的樣本中，抽取 $(X_1, Y_1), (X_6, Y_6), \ldots., (X_{2996}, Y_{2996})$ 來觀察其 Bubble plot 與原始的 $f(x, y)$ plot 做個對比:

(i) Bubble plot

```
library(ggplot2)
f2 <- function(x,y){
  2*(1-x)*(1-y)*(1-x*y)^(-3)
}
index <- as.integer(seq(1,2996,length=600))
result <- data.frame("X"=test[index,1],
                     "Y"=test[index,2],
                     "density"=f2(test[index,1],
                                  test[index,2]))
ggplot(result, aes(x = X, y =Y, size = density)) +
  geom_point(alpha=0.5) +
  scale_size() +
  scale_alpha()
```



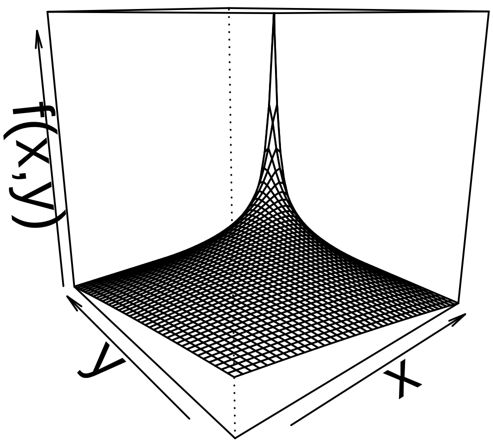
可觀察出，以”f(x,y)“的大小作為點的形狀大小，可以發現點座標越靠近”X=1 和 Y=1” 時，其 f(x,y) 會非常大。

(ii) density plot

```
f <- function(x){
  2*(1-x[1])*(1-x[2])*(1-x[1]*x[2])^(-3)
}
grid1 <- seq(0.05,1,length=40)
```

```r
xy = expand.grid(x=grid1, y=grid1)
z = apply(xy,1,f)

par(mfcol=c(1,1))
persp(grid1, grid1, matrix(z,40,40), theta= -40,
      zlab="f(x,y)", xlab="x", ylab="y",cex.lab=2)
```



可觀察出模擬結果 (Bubble plot) 與原始的 f(x,y) plot 有一定程度上的相似性。

最後，由於這裡採用的抽樣樣本是用 inverse CDF method 轉換得來的，所以 empirical and theoretical efficiency 皆等於 1。