HW04 統算.

※6 In the rejection sampling method, for the selected density $g(x)$, which is easily sampling, we need to choose a positive constant α s.t. $e(x) \geq f(x)$, with $e(x) = \alpha g(x)$

(a) N: The number of the required iteration to successfully generate X.
Show that $N \sim$ geometric distribution.

(b) Verify $E(N) = \alpha$.

pf. Let A denote the event that " if $U \leq \frac{f(Y)}{e(Y)}$, set $X = Y$ "

consider
$$P(A \mid Y=y) = P\left(U \leq \frac{f(Y)}{e(Y)} \mid Y=y\right)$$
$$= P\left(U \leq \frac{f(y)}{e(y)}\right) \quad (\because Y \perp\!\!\!\perp U)$$
$$= \frac{f(y)}{e(y)} \quad \left(\because \frac{f(y)}{e(y)} \leq 1\right)$$
$$= \frac{f(y)}{\alpha g(y)}$$

$$\Rightarrow P(A) = \int_{-\infty}^{\infty} P(A \mid Y=y)\, g(y)\, dy$$
$$= \int_{-\infty}^{\infty} \frac{f(y)}{\alpha g(y)} g(y)\, dy = \frac{1}{\alpha}\int_{-\infty}^{\infty} f(y)\,dy = \frac{1}{\alpha}$$

i.e. the probability of one sample Y being accepted is $\frac{1}{\alpha}$

Note that: for each sample Y, we have with prob. $\frac{1}{\alpha}$ successfully generate X.

That is, N denote the number of "fail" until getting one "success".

Hence, $N \sim geo(\frac{1}{\alpha})$ $\quad \left(P(N=n) = (1-\frac{1}{\alpha})^{n-1}\frac{1}{\alpha}\right)$

(b) $E(N) = \alpha$ (the mean of $geo(\frac{1}{\alpha})$).

(4)

首先用 R package 算出 f(x)=x^(-1/3)+x/10 從 0 到 1 的積分值=1.55

我們利用 Monte Carlo method with U(0,1) ,得到一個估計量 I_U ,

(e.g 樣本數 10000 下估計值=1.54964)

接下來 ,重複生成 1000 次 ,建構 95%信賴區間=[1.54362   1.551501]

```
f<-function(x){x^(-1/3)+x/10}            #積分 f(x)=x^(-1/3)+x/10
a<-0;b<-1;n<-1e04;N<-1000               #從 0 到 1
I<-integrate(f,0,1)$value               #用 R 算出積分值為 1.55

#Monte Carlo method
F_u<-function(n){ (b-a)/n*sum(f(runif(n,0,1))) }
I_U<-F_u(n)                             #利用 U(0,1)算出的積分= 1.54964
confi.of.I_U<-function(N,alpha){
   x<-rep(0,N)
   for(i in 1:N){    x[i]<-x[i]+F_u(n) }
   I.bar<-sum(x)/N
   I.sd<-sd(x)
   cat("The 95% confidence interval for estimating I is [",
         I.bar-I.sd/sqrt(N)*qnorm(1-alpha),"",
         I.bar+I.sd/sqrt(N)*qnorm(1-alpha),"]","\n")}
confi.of.I_U(100,0.05)
               #建構  95% confidence interval for estimating I is [ 1.54362
1.551501 ]
```

(5)

比較兩個不同的 p(x)對於 Monte Carlo method 估計量的 sample variance ,

1.對於 p_1(x) ~ U(0,1) ,理論上的 variance=0.0007208333 ,

樣本數 1000 利用 S^2 估計得 sample variance=0.00078835

2.對於 p_2(x)=2/3*x^(-1/3) ,理論上的 var=1.99999999999978e-06 ,

樣本數 1000 利用 S^2 估計得 sample variance=1.77499754521021e-06

```
#p_1(x) ~ U(0,1)
var1<-(3+3/25+1/300-(31/20)^2)/n
x1<-rep(0,N)
for(i in 1:N){ x1[i]<-x1[i]+F_u(n)}
var.hat1<-sd(x1)^2;rm(x1)            #理論上的 var=0.72/n
                                       利用 S^2 估計得 var.hat=0.603/n
```

```
#p_2(x)=2/3*x^(-1/3)
var2<-(9/4+3/20+9/2000-(31/20)^2)/n
F_2<-function(n){ sum(3/2+3/20*(runif(n,0,1))^(4/3))/n }
I_2<-F_2(n)                           #利用 p_2(x)算出的積分= 1.565765
x2<-rep(0,N)
for(i in 1:N){ x2[i]<-x2[i]+F_2(n)}
var.hat2<-sd(x2)^2;rm(x2)             #理論上的 var=0.002/n
                                        利用 S^2 估計得 var.hat= 0.0020029/n
```

(7)

利用 monte carlo 估計 E(X^4) ,X~N(0,1) ,由 R package 算出 E(X^4)=3 ,

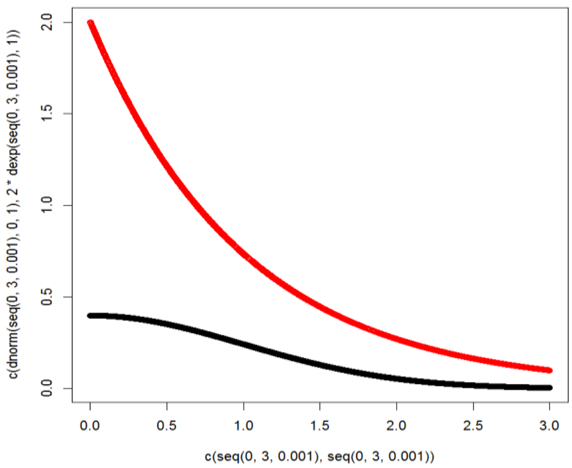
1. with p(x)~U(0,1) , 選擇 b=10 使得目標 f 函數值 < 1e-19
   重複生成 100 次 ,得到算出積分= 2.987094 ,建構 95%信賴區間為
   [2.9586575 3.015532]
2. with p(x)~exp(1), 重複生成 100 次 ,得到算出積分= .006623499 ,建構 95%信賴區間為[2.99874603 3.01450096]

```
#est. E(X^4) by monte carlo
f<-function(x){ 1/sqrt(2*pi)*(x^4)*exp(-x^2/2) }
I<-integrate(f,-Inf,Inf)$value                    #利用 R 算出積分=3
plot(x=seq(0.01,5,0.01),y=f(seq(0.01,5,0.01)))

#with U(0,1)
n<-1e04;b<-10 #integrate(f,10,Inf)
I_Muni<-2*sum(f(runif(n,0,b)))*b/n
confi.of.I_Muni<-function(N,alpha){
  x<-rep(0,N)
  for(i in 1:N){   x[i]<-x[i]+2*sum(f(runif(n,0,b)))*b/n }
  I.bar<-sum(x)/N
  I.sd<-sd(x)
  cat("The 95% confidence interval for estimating I is [",
        I.bar-I.sd/sqrt(N)*qnorm(1-alpha),"",
        I.bar+I.sd/sqrt(N)*qnorm(1-alpha),"]","\n",I.bar)
}
confi.of.I_Muni(100,0.05)

#with exp(1)
n<-1e04;lambda<-1
Exp<-function(n,lambda){
```

```
    X<--log(runif(n,0,1))/lambda
    return(X)
}
fex<-function(x){f(x)/exp(-x) }
I_Mexp<-2*sum(fex(Exp(n,1)))/n
confi.of.I_Mexp<-function(N,alpha){
    x<-rep(0,N)
    for(i in 1:N){    x[i]<-x[i]+2*sum(fex(rexp(n,1)))/n }
    I.bar<-sum(x)/N
    I.sd<-sd(x)
    cat("The 95% confidence interval for estimating I is [",
            I.bar-I.sd/sqrt(N)*qnorm(1-alpha),"",
            I.bar+I.sd/sqrt(N)*qnorm(1-alpha),"]","\n")
}
confi.of.I_Mexp(100,0.05)
```

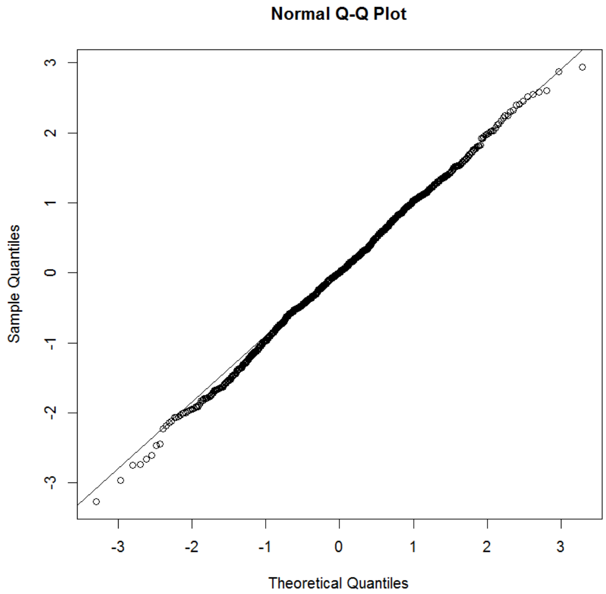
3. with p(x)~N(0,1),

(a)首先生成 N(0,1) by rejection sampling N(0,1) with g(x)~exp(1) , e(x)=c*g(x) with g(x)=exp(-x) 且選擇 c=2 使得 e(x)>p(x) (如下圖)



(黑色為 N(0,1) ;紅色為 2*exp(1))

接著成功生成 10000 個隨機變數~N(0,1) ,



(由 QQ-plot 判斷生成的 N(0,1))

(b) 重複生成 100 次 ,得到算出積分= 3.017778408 ,建構 95%信賴區間為
[3.00298434 3.032572474]