1.

(i) 利用 Trapezoidal rule,求 $f_Z(z)=\frac{1}{C}[1+(\frac{z-0.25}{0.5})^2]^{-1.5}$ 積分從 0 到 0.7，

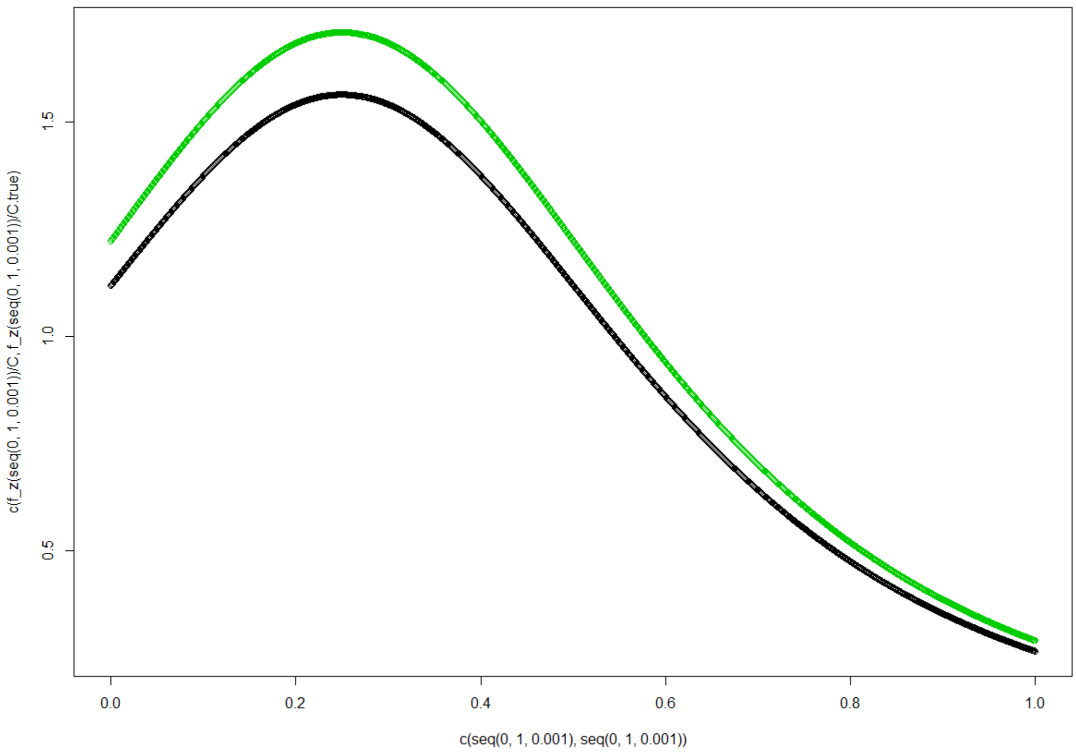
由$\int_0^1 f_Z(z)\,dz=1$,得 $\min f_Z(z)<C<\max f_Z(z)$ ，

取 $C=\frac{\min f_Z(z)+\max f_Z(z)}{2}=\frac{0.170677+1}{2}=0.58533$ ，(用 R 算 C=0.6396319)

接著 ,$|I_T-I|\le\frac{(0.7-0)^3 f_Z''(\tau)}{12n^2}+o(n^{-2})$,$\tau\in(0,0.7)$ 且 $|I_T-I|<0.0001$

得 ,$n>\sqrt{\frac{(0.7-0)^3 f_Z''(0.6)}{12\times 0.0001}}=37.32528$,take $\tau=0.6$ s.t. $f_Z''(\tau)>0$

故 ,理論上所需最小樣本數約為 38



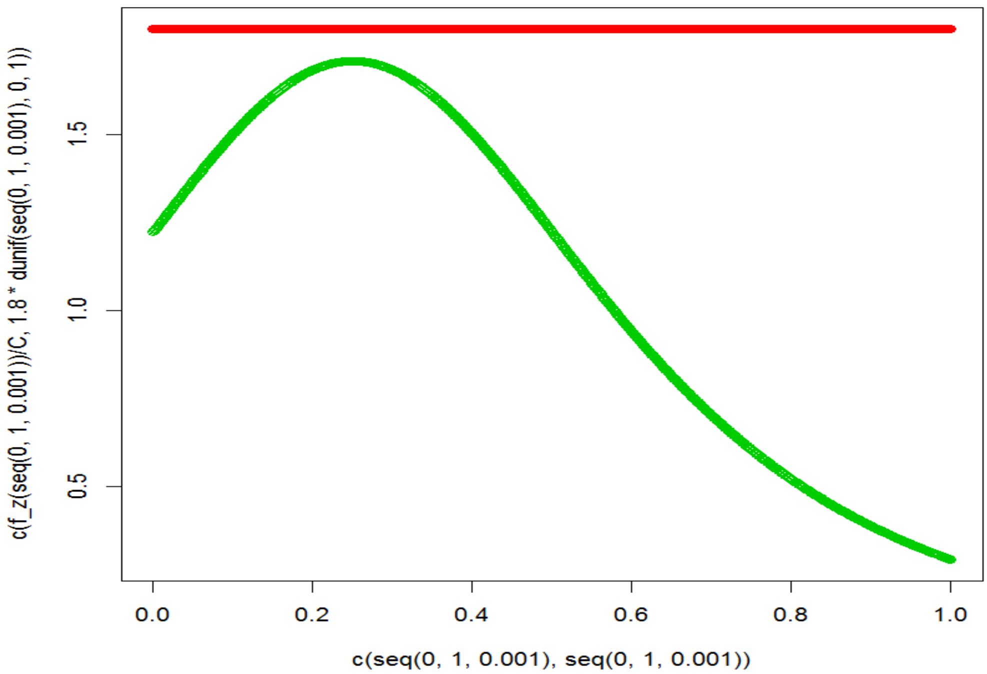
(圖為 $f_Z(z)$ density: 黑色為 R 算的 C ,綠色為估計的 C)

(ii)利用 empirical CDF 去估計 F_z(0.7)=P(Z < 0.7) 的值,

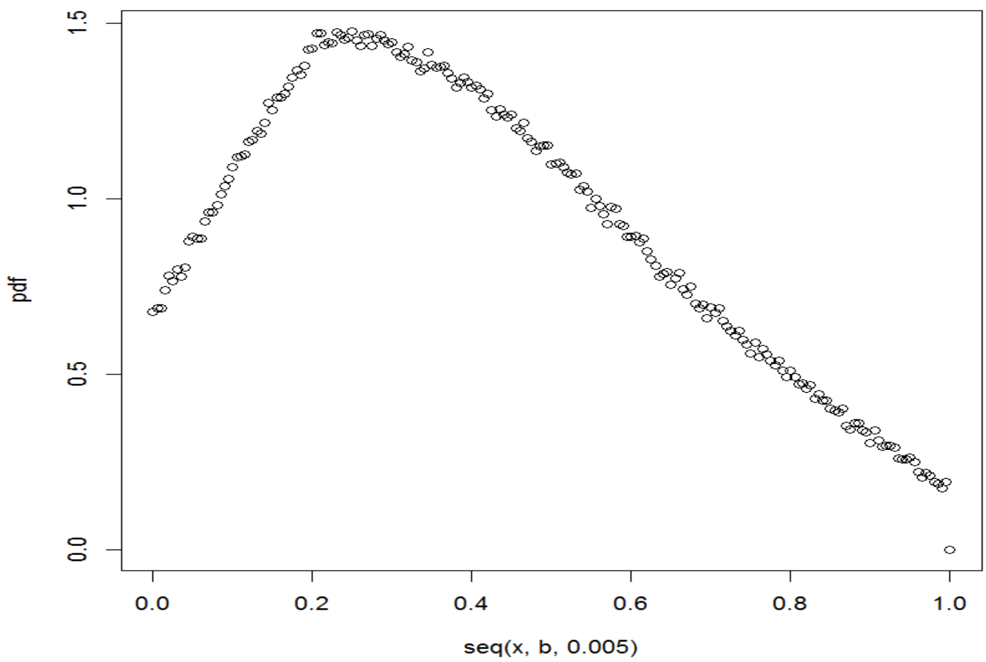
1. sample $z_1, z_2, \ldots, z_n$ iid from f_Z , $C=\frac{\min f_Z(z)+\max f_Z(z)}{2}=0.58533$

2. $F_n(0.7)=1/n\sum_{i=0}^{n} I(Z_i\le 0.7)$

首先利用 Rejection sampling 生成 $z_1, z_2, \ldots, z_n$
選擇 g ~ U(0,1)作為 envolope , e(x)=a*g(x) ,取 a=1.8 畫圖如下



(紅色為 envolope ; 黑色為 $f_Z(z)$)



(此為生成後$z1, z2, \ldots, zn$ 利用 empirical pdf 估計 $f_Z(z)$)

由於$F_n(0.7) = 1/n \sum_{i=0}^{n} I(Z_i \le 0.7)$ 為一個估計量，可以得到在樣本數為 10^5 時估計值為 0.87254 ,用 R 算的積分值為 0.872516089 ,
bias 為 8.608991e-05.
由$I(Z_1 \le z), I(Z_2 \le z), \ldots$ iid with mean $F_Z(z)$, variance $F_Z(z)(1 - F_Z(z))$,
根據中央極限定理 ,$\sqrt{n}(F_n(z) - F_z(z)) \xrightarrow{D} N\left(0, F_Z(z)(1 - F_Z(z))\right)$, with

$$F_n(z) = \frac{1}{n\sum_{i=0}^{n} I(Z_i \le z)} \ and\ S^2 = F_n(z)(1 - F_n(z)) \xrightarrow{p} F_Z(z)(1 - F_Z(z))$$

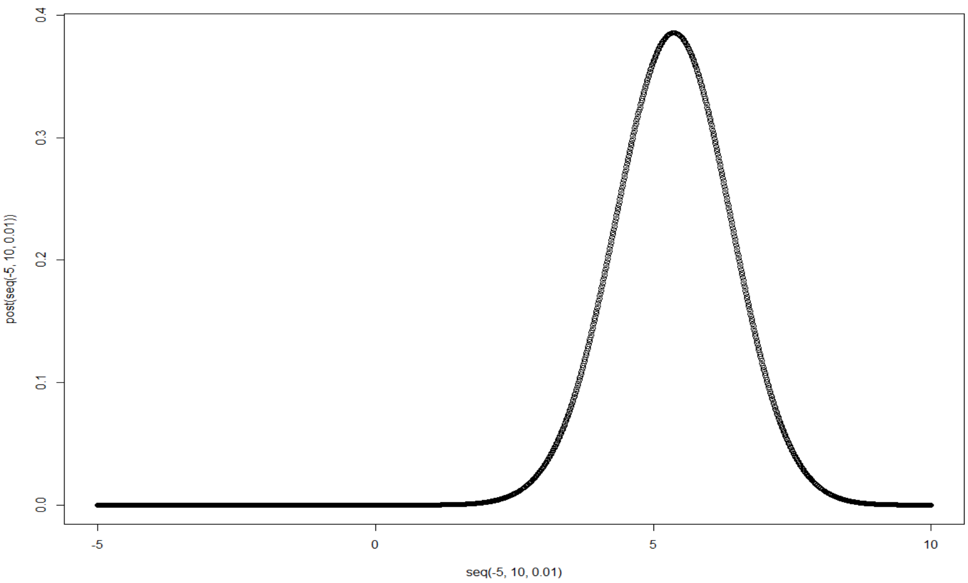
我們構造 level α 信賴區間 ,$[\ F_n(z) \pm \frac{S}{\sqrt{n}} Z_{\alpha/2}\ ]$

$$\Rightarrow |\ F_n(0.7) - F_Z(0.7)\ | \le \frac{S}{\sqrt{n}} Z_{\alpha/2} < 0.0001$$

$$\Rightarrow \sqrt{\frac{F_n(z)(1-F_n(z))}{n}} Z_{\alpha/2} \le 0.0001$$

取$z = 0.7$, $\alpha = 0.01$, $n \ge 10^8 \times F_n(0.7)(1 - F_n(0.7)) \times {Z_{0.005}}^2$
亦即我們有 99%信心水準 ,取出的樣本數滿足$|\ F_n(0.7) - F_Z(0.7)\ | < 0.0001$
模擬結果 n 約為 10^6.

2.估計 mean and variance of posterior density ,
由於此處無法理論推得 posterior density 屬於某個特定分配 ，
Step1.模擬 theta=5 ,X ~N(5,1)，先畫出 posterior density ,



(posterior density 給定 theta=5 ,X ~N(5,1)的畫圖)

因

$$\mathrm{E}(\theta|\mathrm{X}) = \int_{-\infty}^{\infty} \theta p(\theta|X) d\theta$$
$$= \int_{-\infty}^{\infty} \theta \frac{p(X|\theta)p(\theta)}{p(X)} d\theta$$

我們生成 x1,x2,…,xn ~N(5,1) ,取 n=10^4 ,可以利用 Simpson method 算出對於 theta=5 的估計值.

Step 2.由此我們每給定一個 theta ,可以算出相對應的$p(\theta|X)$,

利用 $\mathrm{I.hat} = \sum_\theta \theta p(\theta|X)$ 作為 $\mathrm{E}(\theta|\mathrm{X})$的估計值 ,

考慮樣本平均 $\frac{1}{n}\sum_{i=1}^{n}\sum_\theta \theta p(\theta|X)$,取 n=100 ,

重複 3 次得到估計值為 120.9677, -67.09211, 197.5573

發現對於 mean of posterior 的估計在正負值間跳動 ,

我們對於 theta 的先驗分布為 cauchy ,而理論上 cauchy 分布的 mean 是不存在的 ,而這裡即使給定了資料 X ,對於 theta 的 mean 估計仍沒有很好.

(附上程式碼部分)

1.

```
f_z<-function(z){ (1+( (z-0.25)/0.5 )^2 )^-1.5}
C.true<-integrate(f_z,0,1)$value
I_true<-integrate(f_z,0,0.7)$value /C.true
par(mfrow=c(1,1))
#est. C
C<-( max(f_z(seq(0,1,0.001)) )+min(f_z(seq(0,1,0.001)) ) )/2
plot(c(seq(0,1,0.001),seq(0,1,0.001) ) ,c( f_z(seq(0,1,0.001))/C ,
f_z(seq(0,1,0.001))/C.true ),
        col= c(rep(3,1001),rep(1,1001)) )
plot(seq(0,1,0.001), f_z(seq(0,1,0.001))/C )
#conv. rate
secdiff.f_z<-function(z){    60/C*( 1+(2*z-0.5)^2 )^(-3.5) *(2*z-0.5)^2 -
12/C*( 1+(2*z-0.5)^2 )^(-2.5) }
plot(seq(0,1,0.001),secdiff.f_z(seq(0,1,0.001)));lines(c(0.6,0.6),c(-22,5))
n<-sqrt( (0.7)^3*secdiff.f_z(0.6)/12* (10^4) )

windows();plot(x=c(seq(0,1,0.001),seq(0,1,0.001)),y=c(f_z(seq(0,1,0.001))/C,1.8
*dunif(seq(0,1,0.001),0,1)),
                        col =c(rep(3,1001),rep(2,1001)) )
c<-1.8 ;C*c
```

```
envo<-function(x,c){    f_z(x)/C / c }
acc.rej.exp<-function(n,c=1.8){
   u1<-runif(n,0,1)
   Y<-u1
   u2<-runif(n,0,1)
   X<-rep(0,n)
   N<-length(which(u2 <= envo(Y,c) ))
   for(i in 1:N){      X[i]<-Y[which(u2 <=envo(Y,c))][i]        }        #exp(-(Y-1)^2/2)
   while(N<n){
      uu1<-runif(n-N,0,1)
      Y<-uu1
      uu2<-runif(n-N,0,1)
      if(length(which(uu2<=envo(Y,c)))>0){
         for(i in 1:length(which(uu2<=envo(Y,c)))){ X[N+i]<-
Y[which(uu2<=envo(Y,c))][i] }}
      N<-N+length(which(uu2 <= envo(Y,c)))
   }
   return(X)
}
PDF<-function(a,b,n,band){
   #a=0;b=1;n=1000;band=0.2
   x<-a
   pdf<-rep(0,length(seq(x,b,0.005)))
   j<-1
   while(a<=b){
      I<-0
      for(i in 1:n){ if( abs(a-acc.rej.exp(1))<=band){ I<-I+1    }    }
      pdf[j]<-I/(2*n*band)
      a<-a+0.005
      j<-j+1
   }
   windows();plot(seq(x,b,0.005),pdf) #;lines(c(0,0),c(0,1)) ;lines(c(-
x,b),c(0.5,0.5))
   return(pdf)
}
PDF(0,1,5000,0.2)
#sample size n
cdf<-0 ;n<-10^6
```

```
I<-0
for(i in 1:n){ if(acc.rej.exp(1)<=0.7){ I<-I+1    }    }
cdf<-I/n
while(n<=cdf*(1-cdf)*qnorm(1-0.4,0,1)^2*10^8 ){
   if(n<10^6){n<-n*10}else{n<-n+10^5}
   cdf<-0
   I<-0
   for(i in 1:n){ if(acc.rej.exp(1)<=0.7){ I<-I+1    }    }
   cdf<-I/n
   cat("\n","The sample size is",n," ,bias is",abs(cdf-I_true),"\n")
}
```

2.

```
px.theta<-function(x,theta){ 1/sqrt(2*pi)*exp(-(x-theta)^2/2) }
p.theta<-function(theta){ 1/(pi*(1+theta^2)) }
px<-function(theta){ px.theta(x,theta)*p.theta(theta) }                #plot(seq(-
10,10,0.01),px(seq(-10,10,0.01)))    #plot p(x)
post<-function(theta){ px.theta(x,theta)*p.theta(theta)/I_Spx }    #p(theta|x)
(posterior)    #x<-rnorm(1,5,1);plot(seq(-10,10,0.01),post(seq(-
10,10,0.01)))#plot p(theta|x)
#f<-function(theta){    theta*post(theta) }
#theta*p(theta|x)
#plot p(theta|x)
theta<-5 ;x<-rnorm(1,theta,1)
a<-theta-10 ;b<-theta+10 ;n<-100
h<-seq(a,b,(b-a)/n)
if(n%%2==0){
   z1<-0;z2<-0
   for(i in 2:(n/2)){      z1<-z1+2*px(h[2*i-1])    }
   for(i in 1:(n/2)){      z2<-z2+4*px(h[2*i])    }
   I_Spx<-( px(h[1])+px(h[n+1])+z1+z2 )*(b-a)/n/3
   rm(z1,z2)
} else { print("n need to be even") }
plot(seq(-5,10,0.01),post(seq(-5,10,0.01))) ;lines(c(theta,theta),c(-
0.1,0.6));lines(c(seq(-5,10,0.01)[which(post(seq(-5,10,0.01))==max(post(seq(-
5,10,0.01))))],seq(-5,10,0.01)[which(post(seq(-5,10,0.01))==max(post(seq(-
5,10,0.01))))]),c(-0.1,max(post(seq(-5,10,0.01)))) )
#estmator
E<-rep(0,100)
```

```
for(k in 1:100){
   y<-c()
   for(j in -1000:1000){
      theta<-j
      x<-rnorm(1,theta,1)
      a<-theta-10 ;b<-theta+10 ;n<-100
      h<-seq(a,b,(b-a)/n)
      z1<-0;z2<-0
      for(i in 2:(n/2)){     z1<-z1+2*px(h[2*i-1])    }
      for(i in 1:(n/2)){     z2<-z2+4*px(h[2*i])    }
      I_Spx<-( px(h[1])+px(h[n+1])+z1+z2 )*(b-a)/n/3
      y<-c(y,theta*post(theta))
      }
   E[k]<-sum(y)
}
mean(E)
```