2019302030053-胡哲-第二次作业

第四章习题

5-3

首先利用简单蒙特卡洛方法计算模拟值，对于指数分布方法,利用rexp函数生成服从指数分布的随机数，再根据其中<=0.5的比例利用示性函数方法求估计值。

之后用SSE方法求其方差

```
#### 5-3 ####
m  <-  1e4
# 均匀分布取样
estimate.unif <- function () {
  g <- function (y) {
    exp(-y)
  }
  xs <- g(runif(m, 0, 0.5))
  var <- var(xs)/m # 计算方差

  theta.hat <- mean(xs) * 1/2
  return(data.frame(theta.hat, var))
}
# 指数分布取样
estimate.exp <- function () {
  # theta.hat <- pexp(0.5, rate <- 1) - pexp(0, rate <- 1)<-pexp(0.5,rate<-1)
  y <- rexp(m, rate <- 1) <= 0.5
  var <- var(y)/m # 注意在对逻辑数据做数值运算时，R默认将其转换为01
```

```
  theta.hat <- mean(y)
  return(data.frame(theta.hat, var))
}

estimate.unif()
estimate.exp()
# 指数分布抽样方差更高 这里用的是SSE
```

输出结果分别为:

```
  theta.hat          var
1 0.3927496 1.262505e-06 #均匀分布抽样
  theta.hat          var
1    0.3914 2.382299e-05 #指数分布抽样
```

可见指数分布方法计算所得方差更多，原因应该是指数分布的取样范围是[0,1],而均匀分布的取样范围是[0,0.5]

5-6

对偶变量法方差:

$Var(e^U + e^{1-U}) = Var(e^U) + Var(e^{1-U}) + 2Cov(e^U, e^{1-U})$

$= \frac{e^2-1}{2} - (e-1)^2 + \frac{e^2-1}{2} - (e-1)^2 + 2(3e - e^2 - 1)$

$\approx 0.2420 + 0.2420 - 0.4684$

$= 0.0156$

简单的蒙特卡洛模型的方差:

$Var(g(U)) = Var(e^U) \approx 0.2420$

所以方差减少的百分比为$1 - 0.0156/0.2420 \approx 93.6\%$

5-7

```
#### 5-6, 5-7 ####
m <- 1e4

estimate.5_7 <- function(){
x <- runif(m, min=0, max=1)
theta.hat <- mean(exp(x)) # the simple Monte Carlo method
x1 <- runif(m/2, min=0, max=1)
theta_.hat <- (mean(exp(x1))+mean(exp(1-x1)))/2 # antithetic variate approach
true <- exp(1)-1
c(theta.hat,theta_.hat,true)
}
estimate.5_7()
```

接着5-6用普通方法和对偶变量方法计算估计值，结果如下：

```
# 普通     对偶变量  #真值
1.723539 1.718548 1.718282
```

对偶变量法距离真值更近

5-11

$\hat{\theta}_1$ 和 $\hat{\theta}_2$ 是对立的，所以 $Var(\hat{\theta}_1)$=$Var(\hat{\theta}_2)$，$Corr(\hat{\theta}_1,\hat{\theta}_2)=-1$，所以 $Cov(\hat{\theta}_1,\hat{\theta}_2)=-Var(\hat{\theta}_2)$

$\hat{\theta}_c=c\hat{\theta}_1+(1-c)\hat{\theta}_2$

$Var(\hat{\theta}_c)=c^2Var(\hat{\theta}_1)+(1-c^2)Var(\hat{\theta}_2)+2c(1-c)Cov(\hat{\theta}_1,\hat{\theta}_2)$

$=4c^2Var(\hat{\theta}_2)-4cVar(\hat{\theta}_2)+Var(\hat{\theta}_2)$

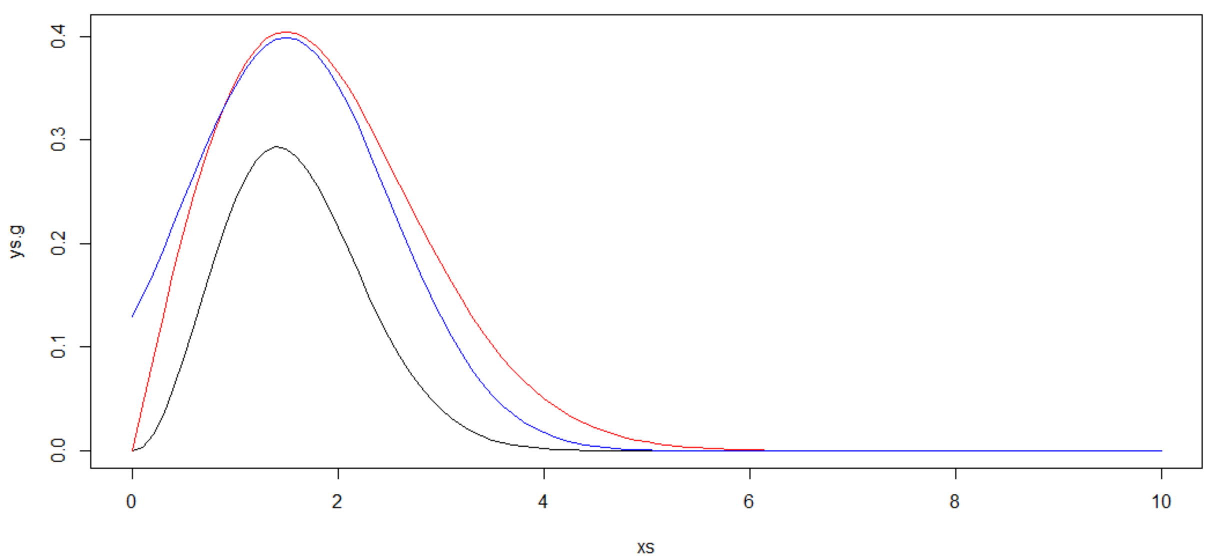
$=(4c^2-4c+1)Var(\hat{\theta}_2)$

所以，当$c=\frac{1}{2}$时，方差最小，原命题得证

5-13

分别生成参数为1.5的瑞利分布密度函数$f_1(X)$和正态分布$f_2(X)$作为近似的重要函数，绘制密度曲线图象查看哪个距离$f(x)$更近

```
#### 5-13 ####
library(bayesmeta) # 这个包可以生成瑞利分布
g = function (x) {
  x ^ 2 / sqrt(2*pi) * exp(-x^2/2)
}

xs = seq(0,10,0.1)

ys.g = g(xs)
ys.rayleigh = drayleigh(xs, scale = 1.5)
ys.norm = dnorm(xs, mean = 1.5)
lim = max(c(ys.g, ys.rayleigh, ys.norm))

plot(xs, ys.g, type = "l", ylim = c(0, lim))
lines(xs, ys.rayleigh, col="red", ylim = c(0, lim))
lines(xs, ys.norm, col="blue", ylim = c(0, lim))

# f1(x) = drayleigh(x, scale = 1.5)
# f2(x) = dnorm(x, mean = 1.5)

# f2 距离g更近，效果应该更好
```



上图中黑线是$f(x)$,红线是$f_1(X)$，蓝线是$f_2(X)$，可见$f_2(X)$距离f更近，效果应更好。

5-15

将[0,1] 分成K份。对每一个 j=0,1,..k-1,计算 $\hat{\theta^j} = \int_{j/k}^{j+1/k} e^{-x}/1 + x^2dx \quad \theta^= \sum_0^k \hat{\theta^j}$.

```
#### 5-15 ####
M<-1e4
N<-1000
k<-5
# 先利用逆变换方法生成密度函数为f_k (x) 的随机数
# 注意这里的(a,b)对应分的不同层的密度函数
inv_fun<-function(n,a,b){
u<-runif(n)
x<--log(exp(-a)-(exp(-a)-exp(-b))*u)
x
}

res3<-sapply(1:N,FUN = function(o){
x<-inv_fun(M,0,1)
M1<-mean((1-exp(-1))/(1+x^2)) #重要抽样
M2<-numeric(k)
for (j in 0:(k-1)){
 a<-j/k
 b<-(j+1)/k
 xj<-inv_fun(M/k,a,b) #生成随机数
 M2[j+1]<-mean((exp(-a)-exp(-b))/(1+xj^2)) #分层重要抽样，注意此时f/g的形式
}
 c(M1,sum(M2))
})

c(var(res3[1,]),var(res3[2,])) #ESE方法估计方差
```

具体细节见代码注释，估计方差如下：

```
# 重要抽样      分层重要抽烟
9.246587e-07 3.095436e-08
```

方差减小96.65%