DATA130004: Homework 3

董晴园 14300680173

2017.10.28

1. Exercise 5.6: Consider the antithetic variate approach for Monte Carlo integration of
$$\theta = \int_0^1 e^x dx.$$
Compute $Cov(e^U, e^{1-U})$ and $Var(e^U + e^{1-U})$, where $U \sim$ Uniform(0,1). What is the percent reduction in variance of $\hat{\theta}$ that can be achieved using antithetic variates (compared with simple MC)?

$$\begin{aligned} Cov(e^U, e^{1-U}) &= E(e^U e^{1-U}) - E(e^U)E(e^{1-U}) \\ &= E(e) - (e-1)^2 = -e^2 + 3e - 1 \end{aligned}$$

$$\begin{aligned} Var(e^U + e^{1-U}) &= Var(e^U) + Var(e^{1-U}) + 2Cov(e^U, e^{1-U}) \\ &= e^2 - 1 - 2(e-1)^2 + 2(-e^2 + 3e - 1) \\ &= -3e^2 + 10e - 5 \end{aligned}$$

Using antithetic variates,
$$\hat{\theta} = \frac{1}{m} \sum_{j=1}^{m/2} (e^{u_j} + e^{1-u_j}), \quad u_j \sim Unif(0,1).$$
The variance of $\hat{\theta}$ is
$$Var(\hat{\theta}) = \frac{1}{2m} Var(e^U + e^{1-U}) = \frac{-3e^2 + 10e - 5}{2m}.$$
While using simple MC,
$$\hat{\theta_m} = \frac{1}{m} \sum_{i=1}^{m} e^{u_i}, \quad u_i \sim Unif(0,1).$$
The variance of $\hat{\theta_m}$ is
$$Var(\hat{\theta_m}) = \frac{1}{m} Var(e^U) = \frac{2e-2}{2m}.$$
Therefore, the percent reduction in variance is
$$\begin{aligned} \frac{Var(\hat{\theta_m}) - Var(\hat{\theta})}{Var(\hat{\theta_m})} \times 100\% &= \frac{2e - 2 - 3e^2 + 10e - 5}{2e-2} \\ &= \frac{3e^2 - 8e + 3}{2e - 2} \approx 99.54\%. \end{aligned}$$

2. Exercise 5.9: The Rayleigh density is

$$f(x) = \frac{x}{\sigma^2} e^{-x^2/(2\sigma^2)}, \quad x \geq 0, \sigma > 0.$$

Implement a function to generate samples from a Rayleigh(σ) distribution, using antithetic variables. What is the percent reduction in variance of $\frac{X+X'}{2}$ compared with $\frac{X_1+X_2}{2}$ for independent $X1, X2$?

We can use antithetic variates with Inverse Transformation methods to generate random samples from the Rayleigh(σ) distribution. First we have to compute

$$F^{-1}(U) = \big(-2\sigma^2 ln(1-U)\big)^{\frac{1}{2}}.$$

Using antithetic variates to generate m samples,

$$\begin{aligned} X_j &= F^{-1}(U_j), \quad 1 \leq j \leq \frac{m}{2} \\ X'_j &= F^{-1}(1-U_j), \quad \frac{m}{2} + 1 \leq j \leq m \end{aligned}$$

```
Sample.Ray <- function(sigma, R = m, antithetic = TRUE) {
  u <- runif(R/2)
  if (!antithetic) v <- runif(R/2) else
    v <- 1 - u
    u <- c(u, v)
    F_x <- numeric(length(u))
  for (i in 1:length(u)) {
    F_x[i] <- (-2*sigma^2*log(1-u[i]))^(1/2)
  }
  F_x
}
```

The percent reduction in variance of $\frac{X+X'}{2}$ compared with $\frac{X_1+X_2}{2}$ is

$$\frac{-Cov(X,X')}{Var(X)}$$

```
m<-10000
X<-numeric(m)
sigma<-2
X<-Sample.Ray(sigma)

x1<- numeric(10000)
x2<- numeric(10000)
u<-runif(10000)
for(i in 1:length(u)){
  x1[i]<- (-2*sigma^2*log(1-u[i]))^(1/2)
  x2[i]<- (-2*sigma^2*log(u[i]))^(1/2)
}
```

```
print(cov(x1,x2))
print(var(X))
-cov(x1,x2)/var(X)
```

So, when $\sigma = 2$, $Cov(X, X') \approx 1.625$, $Var(X) \approx 1.6695$, and the percent reduction in variance is approximately 97.34%.

3. Exercise 5.10: Use Monte Carlo integration with antithetic variables to estimate

$$\int_0^1 \frac{e^{-x}}{1+x^2} dx,$$

and find the approximate reduction in variance as a percentage of the variance without variance reduction.

Assume $U_i \sim$ Unif(0,1). Using simple MC method to estimate,

$$\hat{\theta_m} = \frac{1}{m} \sum_{i=1}^{m} = \frac{e^{-U_i}}{1+U_i^2},$$

while using antithetic variates method,

$$\hat{\theta} = \frac{1}{m} \sum_{i=1}^{m/2} \left(\frac{e^{-U_i}}{1+U_i^2} + \frac{e^{-(1-U_i)}}{1+(1-U_i)^2} \right).$$

```
MC.Int <- function(m, antithetic = TRUE) {
  u <- runif(m/2)
  if (!antithetic) v <- runif(m/2) else
    v <- 1 - u
  u <- c(u, v)
  g <- exp(-u)/(1+u^2)
  cdf=mean(g)
  cdf
}

m <- 1000
MC1 <- MC2 <- numeric(m)
for (i in 1:m) {
  MC1[i] <- MC.Int(m, anti = FALSE)
  MC2[i] <- MC.Int(m)
}
print(sd(MC1))
print(sd(MC2))
print((var(MC1) - var(MC2))/var(MC1))
```

Therefore, the approximate reduction in variance is 96.63%.

4. Use Monte Carlo method to approximate the fraction of a d-dimensional hypersphere which lies in the inscribed d-dimensional hypercube. In class, we have discussed $d = 2$ case. Now, try dimensions $d = 3, 4, ..., 10$. Hint: use `apply` function.

   (a) Derive the formula for the EXACT values for the above problem for each d-dimension.

$$fraction(d) = \frac{\pi^{\frac{d}{2}} r^d}{\Gamma(\frac{d}{2}+1)} / (2r)^d, \quad d = 2, 3, ...10.$$

```
exact<- numeric(9)
for(d in 2:10){
  exact[d-1]<- ((pi^(d/2))/(gamma((d/2)+1)))/(2^d)
}
exact
```

   We get the EXACT values of fraction for each dimention (from $d = 2$ to $d = 10$):

```
[1] 0.785398163 0.523598776 0.308425138 0.164493407 0.080745512
[6] 0.036912234 0.015854344 0.006442400 0.002490395
```

   (b) Using the previous results, approximate the value of π. Find the number of points used which approximates π to the 4th digit for each d. Set the random seed with `set.seed(123)` at the beginning of your R code.

   For $d = 1, ..., 10$ case, we can use the following formula to approximate.

```
N=100000
set.seed(123)
fraction<- numeric(9)
Pi<- numeric(9)
for(d in 2:10){
  X=list(d)
  for(i in 1:d){
    X[[i]]=runif(N,-1,1)
  }
  x=do.call(cbind,X)
  fraction[d-1]<- sum(apply(x,MARGIN = 1,function(x){
    ifelse(sum(x^2)<1,1,0)
    }))/N
  Pi[d-1]<-(fraction[d-1]*(2^d)*gamma(d/2+1))^(2/d)
}
print(fraction)
print(Pi)
```

   And here we get the approximate value of the fraction and π for each dimension:

```
[1] 0.78658  0.52371  0.30833  0.16352  0.08094  0.03730  0.01571  0.00650  0.00236
[1] 3.146320 3.142038 3.141108 3.134143 3.144113 3.150987 3.134418 3.147813 3.107983
```