Computational Statistics Homework 3

吴嘉骜 21307130203

2023 年 10 月 14 日

1 ex 5.11

If $\hat{\theta}_1$ and $\hat{\theta}_2$ are unbiased estimators of θ, and $\hat{\theta}_1$ and $\hat{\theta}_2$ are antithetic, we derived that $c^* = 1/2$ is the optimal constant that minimizes the variance of $\hat{\theta}_c = c\hat{\theta}_1 + (1-c)\hat{\theta}_2$. Derive c^* for the general case. That is, if $\hat{\theta}_1$ and $\hat{\theta}_2$ are unbiased estimators of θ, find the value c^* that minimizes the variance of $\hat{\theta}_c = c\hat{\theta}_1 + (1-c)\hat{\theta}_2$ in equation (5.11). (c^* will be a function of the variances and the covariance of the estimators.)

Solution:

If $\hat{\theta}_1$ and $\hat{\theta}_2$ are any unbiased estimators of θ, we have

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

This is a quadratic function of c, and the minimum is achieved at

$$c^* = \frac{Var(\hat{\theta}_2) - Cov(\hat{\theta}_1, \hat{\theta}_2)}{Var(\hat{\theta}_1) + Var(\hat{\theta}_2) - 2Cov(\hat{\theta}_1, \hat{\theta}_2)}.$$

2

Consider estimating the rare event probability $P(Z > 10)$ where $Z \sim N(0,1)$.

(a) Apply the simple Monte Carlo method, record your results when the sample size is $10^3, 10^4, \ldots$

(b) Apply the change of measure using $Y \sim N(10,1)$, record your results when the sample size is $10^3, 10^4, \ldots$

Solution:

The R code is as follows:

```
# Rare event P(Z>10) where Z N(0,1)
# (a) Simple Monte Carlo
SMC <- function(N){
  X <- rnorm(N)
  mean(X>10)
}

# (b) Change of measure
CoM <- function(N){
```

```
  Y <- rnorm(N, mean=10, sd=1)
  mean((Y>10)*exp(50-10*Y))
}


Ns <- c(1e3,1e4,1e5,1e6,1e7)
results <- matrix(NA, nrow=length(Ns), ncol=3)
colnames(results) <- c("Sample Size", "Simple Monte Carlo", "Change of Measure")

for (i in 1:length(Ns)){
  n <- Ns[i]
  results[i,] <- c(n, SMC(n), CoM(n))
}
results_df <- as.data.frame(results)
results_df
```

The results are as follows:

```
> results_df
  Sample Size Simple Monte Carlo Change of Measure
1       1e+03                  0      9.366175e-24
2       1e+04                  0      7.532025e-24
3       1e+05                  0      7.633165e-24
4       1e+06                  0      7.642613e-24
5       1e+07                  0      7.625574e-24
```

From the results, we can see that the simple Monte Carlo method cannot give a good estimation of the rare event probability, because the probability $P(Z > 10),\ Z \sim N(0, 1)$ is too small, and a much larger sample size is needed to get a good estimation. By the way, my hardware cannot afford a sample size larger than 10^9.

However, the change of measure method can present a satisfactory estimation, because $P(Y > 10),\ Y \sim N(10, 1)$ is much larger, and exponential function is continuous, giving small but non-zero values for us to calculate.