



2018 Fall

Computational Statistics HW#8

182STG18 이하경

I. Description

< Variance Reduction Techniques >

Monte Carlo Simulation 을 이용해 적분 값을 추정할 때, 단순 추정치 $\hat{\mu}_{MC} = \sum_{i=1}^n h(X_i)$ 이외에도 추정치의 분산 및 표준 오차를 줄이기 위한 다양한 방법들이 존재한다. HW7에서는 그 중 관심 event 의 발생 확률이 매우 작은 경우 유용하게 사용 가능한 Important Sampling 방법에 대해 다루었다. HW8에서는 다음의 두 방법에 대해 소개한다.

- Antithetic Sampling

동일한 분포를 가지며 n 개의 sample로부터 추정된 두 개의 추정치 $\hat{\mu}_1, \hat{\mu}_2$ 가 서로 음의 상관관계를 가질 때, $\hat{\mu}_{AS} = (\hat{\mu}_1 + \hat{\mu}_2)/2$ 의 분산은 아래와 같다.

$$\text{Var}\{\hat{\mu}_{AS}\} = \frac{1}{4}(\text{Var}\{\hat{\mu}_1\} + \text{Var}\{\hat{\mu}_2\}) + \frac{1}{2}\text{Cov}\{\hat{\mu}_1, \hat{\mu}_2\} = \frac{(1 + \rho)\sigma^2}{2n}$$

Antithetic Sampling (AS)는 n 개의 sample로부터 $\hat{\mu}_1 = \sum_{i=1}^n h_1(X_i)/n$ 을 얻었을 때 동일한 sample을 재사용해 이와 음의 상관관계가 있는 $\hat{\mu}_2 = \sum_{i=1}^n h_2(X_i)/n$ 을 손쉽게 계산함으로써, $2n$ 개의 sample로부터 얻을 수 있는 $\hat{\mu} = \sum_{i=1}^{2n} h(X_i)/2n$ 의 분산 $\sigma^2/2n$ 보다 더 작은 분산을 가져 정밀한 추정을 가능하게 할 수 있는 방법이다.

h_1 과 h_2 가 둘다 동시에 증가하거나 감소(monotone increasing or decreasing)할 경우, random variable X_i 에 대해 $\text{Cov}\{h_1(X_i), h_2(X_i)\} \geq 0$ 이고, 이 경우 위에서 제시된 $\hat{\mu}_1$ 과 $\hat{\mu}_2$ 또한 $\text{Cov}\{\hat{\mu}_1, \hat{\mu}_2\} \geq 0$ 이다.

F 가 X_i 들이 따르는 CDF라고 할 때 $F(X_i) = U_i$, $U_i \sim U(0,1)$ 에서 $X_i = F^{-1}(U_i)$ 으로 표현할 수 있다. 따라서 다음과 같이 식을 변형 가능하며 monotone 성질을 잃지 않는다.

$$\hat{\mu}_1 = \frac{\sum_{i=1}^n h_1(X_i)}{n} = \frac{\sum_{i=1}^n h_1(F^{-1}(U_i))}{n}$$

또한 $U_i \sim U(0,1)$ 일 경우 $1 - U_i \sim U(0,1)$ 또한 성립하고 만약 $h_1(F^{-1}(U_i))$ 가 U_i 에 대해 단조 증가할 경우 $h_1(F^{-1}(1 - U_i))$ 는 단조 감소하며 역의 경우도 동일하다. 따라서 $\hat{\mu}_1$ 과 음의 상관관계에 있고 동일한 분포를 가지는

$$\hat{\mu}_2 = \frac{\sum_{i=1}^n h_2(X_i)}{n} = \frac{\sum_{i=1}^n h_1(F^{-1}(1 - U_i))}{n}$$

을 얻을 수 있고 이를 토대로 적분 값 μ 의 추정치로서 분산이 작은 $\hat{\mu}_{AS}$ 을 얻을 수 있다.

$$* \hat{\mu}_{AS} = \frac{\hat{\mu}_1 + \hat{\mu}_2}{2} = \frac{[\sum_{i=1}^n h_1(F^{-1}(U_i)) + \sum_{i=1}^n h_1(F^{-1}(1 - U_i))]}{2n}$$

- Control Variates

Control Variates (CV) 방법은 알려지지 않은 적분 값 $\mu = E[h(X)]$ 을 추정하기 위해 이와 관련있고 analytical solution이 존재하는 알려진 추정치 $\theta = E[c(Y)]$ 를 이용하는 방법이다. MC 시뮬레이션을 통해 추정한 $\hat{\mu}_{MC}$ 와 $\hat{\theta}_{MC}$ 가 서로 관계가 존재하고, $\hat{\theta}_{MC}$ 의 실제 값 θ 을 직접 계산할 수 있을 때 $\hat{\theta}_{MC}$ 과 θ 의 차이와 일정한 비율로 $\hat{\mu}_{MC}$ 을 조정해준다는 의미를 가지고 있다. 예를 들어 $\hat{\theta}_{MC}$ 가 θ 에 비해 큰 값이 나왔을 경우 만약 $\hat{\mu}_{MC}$, $\hat{\theta}_{MC}$ 이 양의 상관관계를 가진다면 $\hat{\mu}_{MC}$ 역시 실제 값보다 큰 값이라고 여겨 하향 조정하고, 두 값이 음의 상관관계를 가진다면 반대로 $\hat{\mu}_{MC}$ 을 상향 조정한다는 것이다.

$$(X_i, Y_i): \text{ indep. pairs of r.v.'s of a sample size } n \quad \hat{\mu}_{MC} = \sum_{i=1}^n h(X_i)/n, \hat{\theta}_{MC} = \sum_{i=1}^n c(Y_i)/n, \quad \text{Corr}\{h(X_i), c(Y_i)\} \neq 0$$

$$* \hat{\mu}_{CV} = \hat{\mu}_{MC} + \lambda(\hat{\theta}_{MC} - \theta)$$

$$\text{Var}\{\hat{\mu}_{CV}\} = \text{Var}\{\hat{\mu}_{MC}\} + \lambda^2 \text{Var}\{\hat{\theta}_{MC}\} + 2\lambda \text{Cov}\{\hat{\mu}_{MC}, \hat{\theta}_{MC}\}$$

$$\rightarrow \min_{\lambda} \text{Var}\{\hat{\mu}_{CV}\} = \text{Var}\{\hat{\mu}_{MC}\} - \frac{\text{Cov}\{\hat{\mu}_{MC}, \hat{\theta}_{MC}\}^2}{\text{Var}\{\hat{\mu}_{MC}\}} \leq \text{Var}\{\hat{\mu}_{MC}\} \quad \text{where } \hat{\lambda} = \underset{\lambda}{\text{argmin}} \text{Var}\{\hat{\mu}_{CV}\} = -\frac{\text{Cov}\{\hat{\mu}_{MC}, \hat{\theta}_{MC}\}}{\text{Var}\{\hat{\mu}_{MC}\}}$$

위에 식에서 확인할 수 있듯이 CV 추정치는 특정한 λ 값을 추정하여 계산할 경우 단순한 MC 추정치보다 분산을 더 줄일 수 있다.

(ex) ISCV (Importance Sampling with Control Variates)

Variance Reduction Technique 중 하나인 Importance Sampling 에서, unstandardized weights $w^*(X) = f(X)/g(X)$ 와 관심 함수 $h(X)$ 에 대해 $h(X)w^*(X)$ 는 서로 Correlation 이 존재하고, $f(X)$ 와 $g(X)$ 는 가능한 한 서로 가깝길 기대하므로 $E[w^*(X)] = 1$ 이다. 따라서 다음과 같이 Control Variates 전략을 이용하여 좀 더 향상된 추정을 할 수 있다.

$$\hat{\mu}_{ISCV} = \hat{\mu}_{IS} + \lambda(\bar{w}^* - 1), \quad \bar{w}^* = \frac{\sum_{i=1}^n w^*(X_i)}{n}$$

이때 $\hat{\lambda} = -\text{Cov}\{\hat{\mu}_{IS}, \bar{w}^*\} / \text{Var}\{\bar{w}^*\}$ 으로 추정 가능하며 이를 이용해 \bar{w}^* 과 w^* 의 기댓값 1의 차이만큼에 비례하여 $\hat{\mu}_{IS}$ 을 조정할 수 있다.

II. Implementation

1. Option Pricing with Control Variates (Example 6.13)

Information of an Asian Call Option				
Current Price of underlying stock	Strike Price	Volatility	Maturity	risk-free rate of return
S_0	K	σ	N	r
100	102	0.3	50	0.05

payoff = $(\bar{S}_T - K)^+$, \bar{S}_T = the arithmetic mean stock price during the holding period (until T)

$$S_{t+1} = S_t \cdot \exp\left\{\frac{r - \sigma^2/2}{365} + \frac{\sigma Z_t}{\sqrt{365}}\right\}, \quad Z_t \sim N(0, 1)$$

$$\text{the fair price of the option } A = \exp\left\{-\frac{rT}{365}\right\} \cdot \max\{\bar{S}_T - K, 0\}, \quad \bar{A} = \sum_{i=1}^n A_i = \hat{\mu}_{MC}$$

위와 같은 옵션에서 \bar{S} 을 계산할 때 산술평균 $\sum_{t=1}^T S_t / T$ 가 아닌 기하평균 $(\prod_{t=1}^T S_t)^{\frac{1}{T}}$ 을 사용할 경우, 옵션의 적정가격 A 의 기댓값은 analytic solution θ 가 존재한다. ($\theta = 1.82796$)

Goal

MC Simulation 을 통해 \bar{S} 에 산술평균을 이용한 $\hat{\mu}_{MC}$ 와 기하평균을 이용한 $\hat{\theta}_{MC}$ 을 각각 generate 한 후, Control Variates 전략을 이용하여 $\hat{\mu}_{CV} = \hat{\mu}_{MC} + \lambda(\hat{\theta}_{MC} - \theta)$ 을 계산해보고 분산을 비교해본다.

* 일반적으로 $\hat{\mu}_{MC}$ 와 $\hat{\theta}_{MC}$ 는 매우 강한 양의 상관관계가 있으므로 적절한 λ 값으로 $\lambda = -1$ 과 $\hat{\lambda} = -\frac{\text{Cov}\{\hat{\mu}_{MC}, \hat{\theta}_{MC}\}}{\text{Var}\{\hat{\mu}_{MC}\}}$

두 가지를 이용해보고 차이를 확인한다.

sample size n 은 100000, simulation 의 반복 횟수는 100 으로 설정하였다.

Result

	$\hat{\mu}_{MC}$	$\hat{\theta}_{MC}$	$\hat{\mu}_{CV} (\lambda = -1)$	$\hat{\mu}_{CV} (\hat{\lambda} = -1.0228)$
mean	1.86263	1.81501	1.87559	1.87588
sd	0.01169	0.01142	0.00030	0.00015

10000 개의 $\hat{\mu}_{MC}$ 와 $\hat{\theta}_{MC}$ 에서 $\text{Corr}\{\hat{\mu}_{MC}, \hat{\theta}_{MC}\} = 0.99993$ 으로, 강한 양의 상관관계가 있음을 확인하였다. $\hat{\theta}_{MC}$ 값이 실제 값 closed form equation 을 이용해 계산한 $\theta = 1.82796$ 보다 다소 작게 추정되었으며 $\hat{\mu}_{MC}$ 역시 실제 기댓값보다 작게 계산되었을 것이라고 예상하므로 조정한 $\hat{\mu}_{CV}$ 은 계산된 $\hat{\mu}_{MC}$ 보다 큰 값을 가진다. 두 가지 λ 값에 대한 CV 추정치의 표준오차는 MC 추정치보다 확연히 작았으며 CV 추정치의 분산을 가장 작게 하는 $\hat{\lambda} = -\frac{\text{Cov}\{\hat{\mu}_{MC}, \hat{\theta}_{MC}\}}{\text{Var}\{\hat{\mu}_{MC}\}} = -1.0228$ 을 이용한 경우 표준오차가 0.00015 으로 MC 추정치의 약 100 분의 1 에 그치지 않는다.

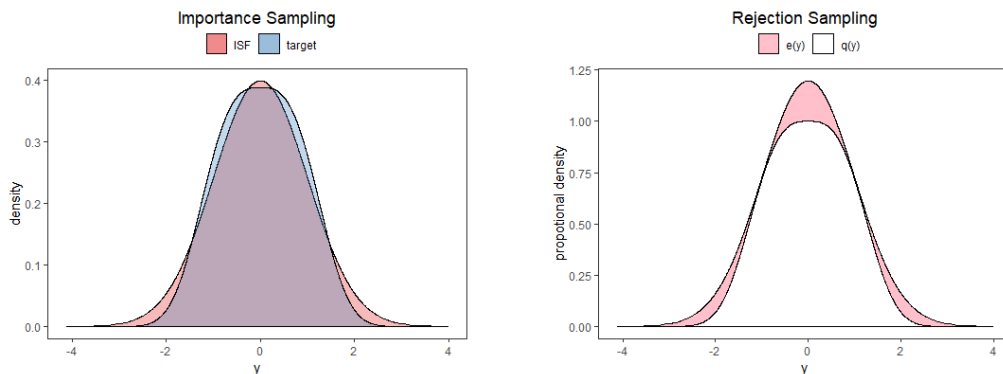
2. Estimation using Various Sampling Strategies (Problem 6.3)

proportional density of X, $q(X) = f(X)/c = \exp\{-|X|^3/3\}$ and $\sigma^2 = E[X^2] = E[h(X)]$

Goal

- Importance Sampling with standardized weights 을 이용해 기댓값 $\sigma^2 = E[X^2]$ 을 추정한다.
- Rejection Sampling 을 이용해 추정해본다.
- (b)에서 생성된 표본에 대해 Phillippe & Robert 의 Importance-weighted averaging

$$\frac{\sum_{i=1}^{n-1} (X_{[i+1]} - X_{[i]}) h(X_{[i]}) q(X_{[i]})}{\sum_{i=1}^{n-1} (X_{[i+1]} - X_{[i]}) q(X_{[i]})}, \quad X_{[i]}: \text{ordered sample of } X_i\text{'s}$$
 을 이용하여 추정치를 계산해본다.
- (b)와 (c)에서의 시뮬레이션을 여러 번 반복해보고 두 가지 추정 방법을 비교해본다.



왼쪽의 그래프는 $\int_{-\infty}^{\infty} q(X) dx / \text{Const.} = 1$ 이 되게 하는 비례상수 $C=2.5758$ 을 이용하여 조정한 target density f 을 그려본 결과이다. IS 추정 시 해당 density 보다 두꺼운 꼬리를 가지는 Standard Normal density $N(0, 1)$ 을 적절한 ISF (or envelope)으로 이용하였다.

Rejection Sampling 의 경우 envelope 는 target $f=q/c$ 보다 모든 x 에 대해 항상 커야 하므로 standard normal density 에 비례상수 2.57 보다 좀 더 큰 적절한 상수 3 을 곱하여 envelope 로 이용하였다. 따라서 Rejection Sampling 은 다음의 절차에 따라 진행되었다.

- 1) Sample $Y \sim N(0, 1)$ & $U \sim N(0, 1)$
- 2) Reject Y if $U \geq \frac{q(Y)}{3 \cdot e(Y)}$, $e(Y)$ = density of $N(0, 1)$
- 3) Otherwise keep Y & Set $X=Y$

Result (a)-(c).

	IS	RS	P.R
estimate	0.77659	0.77494	0.77576

$n=10000$ 의 sample 에 대해 각각의 추정치를 계산한 결과 세 가지 방법을 이용했을 때 추정치가 비슷하였다. Rejection Sampling 의 경우 10000 개의 초기 sample 중 유지된 sample 은 총 8548 개로, acceptance rate=85.48%으로 계산되었다.

Result (d). $\hat{\sigma}_{RS}^2$ & $\hat{\sigma}_{P.R}^2$ 의 비교

of Simulation iterations = 1000

	mean	sd	comp.time
RS	0.77646	0.00957	0.00401
Phillipe & Robert	0.77526	0.00146	0.00644
(+) IS	0.77697	0.00834	0.00389

시뮬레이션의 반복횟수를 1000 번으로 하여 추정치들의 평균값과 표준오차, 소요 시간을 비교해보았다. Rejection Sampling 방법으로 생성된 sample X 에 대해 단순 평균으로 구한 $\hat{\sigma}_{RS}^2$ 와 Phillipe and Robert 방법으로 구한 $\hat{\sigma}_{P.R}^2$ 의 표준오차를 비교해보면 P.R 추정치가 계산 시간은 조금 더 소요되었지만 훨씬 더 작은 표준오차를 보인다. IS 에 대해서도 1000 번의 시뮬레이션을 반복한 결과 표준오차와 소요시간은 RS 방법과 비슷하였다.

3. Testing the Hypotheses from a Poisson Model (Problem 6.6)

$$H_0: \lambda = 2 \quad \text{v.s.} \quad H_a: \lambda > 2$$

\bar{X} 가 Poisson(λ)의 25 개의 관측치로부터 계산된 표본평균일 때, CLT 에 의해 $\bar{X} \sim N\left(2, \sqrt{2/25}\right)$ 으로 근사 가능하다. 따라서 검정통계량 $Z = (\bar{X} - 2)/\sqrt{2/25}$ 을 계산하여 $Z \geq 1.645$ 일 경우 H_0 를 기각한다. ($Z_\alpha = 1.645$, $\alpha = 0.05$)

$$\text{Type I error Rate} = P\{Z > 1.645 \text{ under } H_0\} = E[h(Z)], \quad h(Z) = \begin{cases} 1 & \text{if } Z > 1.645 \\ 0 & \text{if } Z \leq 1.645 \end{cases}$$

$$\text{Type II error Rate} = P\{Z > Z_0 \text{ under } H_a\} = 1 - P\{Z \leq Z_0\} = E[k(Z)], \quad Z_0 = (b - \bar{X})/\sqrt{s/25}$$

Goal

a. 5 가지 Monte-Carlo Approach 을 이용해 Size of the test = $P\{\text{Type I Error}\} = P\{\text{Reject } H_0 \mid H_0 \text{ is true}\}$ 을 추정하고, 95% CI 을 계산한다. 분산 축소 측면에서 방법 별 이점을 비교해보고 IS Approach 간의 차이 또한 비교해본다. IS 추정치들의 경우 envelope 로 rejection Region 의 threshold 2.4653 을 이용한다.

b. Power of the test = $P\{\text{Type II Error}\} = P\{\text{Reject } H_0 \mid H_a \text{ is true}\}$ 을 추정하고 $\lambda \in [2.2, 4]$ 의 5 개의 값에 대해 power curve 및 각 경우의 confidence band 를 그려보고 각 Approach 의 결과에 대해 비교한다.

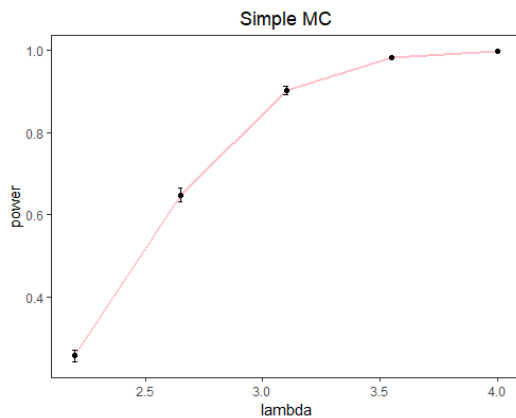
Result (a). estimates of the size of the test α

method	mean	sd	lower CI	upper CI
MC	0.05564	0.00729	0.041	0.070
AS	0.05156	0.00684	0.038	0.065
unstandardized IS*	0.05089	0.00254	0.04606	0.05575
standardized IS	0.05385	0.00497	0.04443	0.06407
ISCV	0.05041	0.00247	0.04559	0.05515

sample size $n=1000$ 에 대해 1000 번의 simulation 을 반복하여 모형의 size 추정치들의 분포를 구하였다. 값들의 평균이 전부 $\alpha = 0.05$ 에 근접했으며 표준오차는 $ISCV < IS^* < IS < AS < MC$ 순으로 Importance Sampling 과 Control Variates 를 동시에 이용할 수 있는 ISCV 가 단순한 MC 추정치에 비해 가장 분산을 크게 감소시킬 수 있었다.

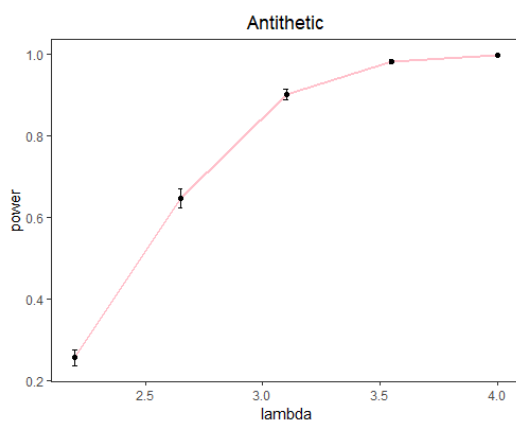
AS 추정의 경우 본 시뮬레이션에서 sample size 를 500 개만 사용하면서도 추정치들의 평균이 MC 추정치보다 $\alpha = 0.05$ 에 더 가까웠고, 분산도 어느 정도 줄어들어 효과적인 방법이다.

IS 추정 시 \bar{X} 에 대한 Importance Sampling Function 으로 $N(2.6453, 2.6453/25)$ 을 이용하였으며 standardized IS 추정치보다 unstandardized IS 추정치의 분산이 더 작았다. $h(\bar{X})w^*(\bar{X})$ 와 $w^*(\bar{X})$ 의 상관관계가 강할 경우 standardized IS 추정치가 더 유용하나 이 경우는 unstandardized 추정치의 분산이 더 작으므로 더 유용하다.

Result (b). estimates of the power of the test where $\lambda \in [2.2, 4]$ 

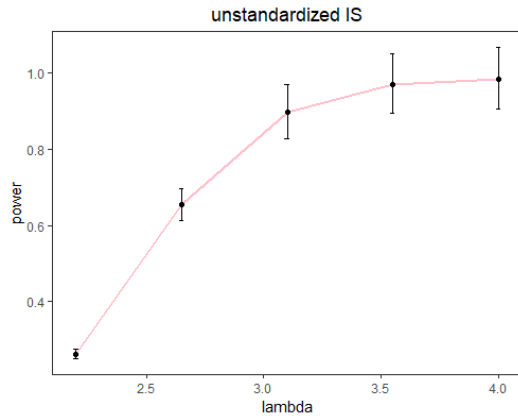
Simple Monte-Carlo

lambda	mean	sd	lower CI	upper CI
2.20	0.25620	0.00757	0.24090	0.27113
2.65	0.64752	0.00877	0.63023	0.66466
3.10	0.90282	0.00474	0.89276	0.91187
3.55	0.98278	0.00156	0.97946	0.98582
4.00	0.99778	0.00036	0.99696	0.99841

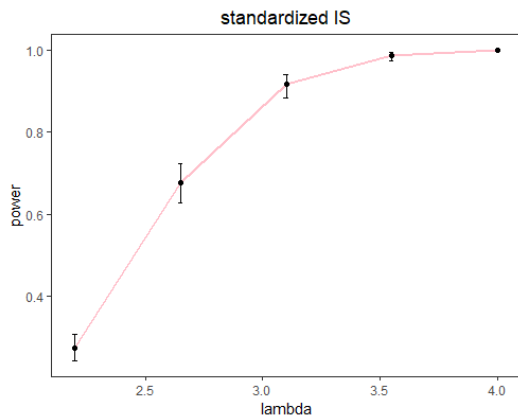


Antithetic Sampling

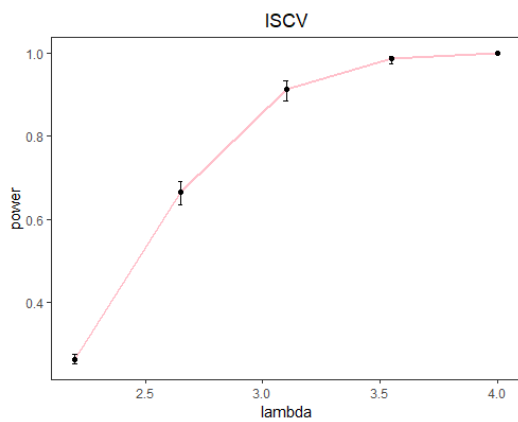
lambda	mean	sd	lower CI	upper CI
2.20	0.25618	0.01062	0.23506	0.27580
2.65	0.64719	0.01232	0.62423	0.67103
3.10	0.90254	0.00663	0.88922	0.91471
3.55	0.98272	0.00218	0.97792	0.98656
4.00	0.99778	0.00050	0.99665	0.99858



unstandardized IS				
lambda	mean	sd	lower CI	upper CI
2.20	0.26199	0.00615	0.24997	0.27452
2.65	0.65485	0.02171	0.61371	0.69751
3.10	0.89659	0.03686	0.82638	0.97072
3.55	0.96924	0.03940	0.89376	1.05135
4.00	0.98401	0.04167	0.90502	1.06775



standardized IS				
lambda	mean	sd	lower CI	upper CI
2.20	0.27475	0.01595	0.24278	0.30697
2.65	0.67733	0.02492	0.62711	0.72311
3.10	0.91679	0.01525	0.88290	0.94016
3.55	0.98644	0.00520	0.97221	0.99293
4.00	0.99831	0.00118	0.99520	0.99941



IS with CV				
lambda	mean	sd	lower CI	upper CI
2.20	0.26361	0.00571	0.25242	0.27499
2.65	0.66458	0.01359	0.63592	0.69053
3.10	0.91194	0.01148	0.88529	0.93183
3.55	0.98539	0.00467	0.97344	0.99160
4.00	0.99817	0.00113	0.99526	0.99933

[2.2, 4] 구간의 5 개의 λ 에 따른 모형의 power curve 와 $n=1000$, simulation 횟수 1000 번의 평균 및 표준오차, 95% CI를 계산한 결과이다. λ 의 값이 커질수록 모든 그래프에서 power의 추정 값이 1에 가까워지며 unstandardized IS 추정치를 이용하였을 때를 제외하고 신뢰구간의 폭은 λ 가 커질수록 감소하였다. power 추정의 경우 size 추정과 다르게 unstandardized IS 방법이 다른 방법들에 비해 안정적이지 못하였다. $\lambda = 4$ 을 기준으로 표준오차는 $AS < MC < IS < ISCV < IS^*$ 순으로 작았다.

4. Option Pricing with Various Techniques (Problem 6.7)

Information of a European Call Option				
Current Price of underlying stock	Strike Price	Volatility	Maturity	risk-free rate of return
S_0	K	σ	N	r
50	52	0.5	30	0.05

Goal

- a. 위의 European Call Option 의 적정 가격 $C = \exp\{-rT/365\} \cdot \max\{S_{30} - K, 0\}$ 을 계산해보고 2.10 이 적절한 가격인지 확인한다.
- b. 위의 Option 이 Asian Call Option 이라고 가정할 때, simple MC 방법을 이용해 옵션의 적정가격 A 에 대해 $E[A]$ 을 추정해본다.

$$\widehat{E}[A] = \hat{\mu}_{MC} = \sum_{i=1}^n A_i, \quad A_i = \exp\{-rT/365\} \cdot \max\{\bar{S} - K, 0\}$$

- c. (b)에서 Control Variates 방법을 이용하여 $\widehat{E}[A] = \hat{\mu}_{CV} = \hat{\mu}_{MC} + \hat{\lambda}(\hat{\theta}_{MC} - \theta)$ 을 계산해본다.
- d. (b)에서 Antithetic Sampling 방법을 이용하여 $\widehat{E}[A] = \hat{\mu}_{AS}$ 을 계산해본다.
- e. 시뮬레이션을 이용해 (b), (c), (d)의 추정치들의 분포를 서로 비교해본다.

Result (a). European Call Price

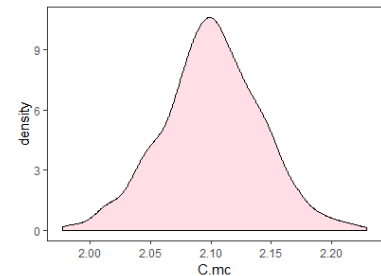
1) by changing sample size n

n=10000	n=50000	n=100000
2.09672	2.10061	2.10628

2) by iterating 1000 simulations (n=10000)

mean	sd
2.10110	0.04084

sample size n 을 바꿔가며 n 개의 평균으로 적정가격 C 의 단순 MC 추정치를 계산한 결과 제시된 모두 2.10 과 거의 비슷한 값이 계산되었다. 1 번의 시뮬레이션에서는 추정치의 분포를 알 수 없으므로 sample size 를 10000 으로 고정하고 1000 번의 시뮬레이션을 반복해본 결과 표준오차는 0.04 정도를 보였다.



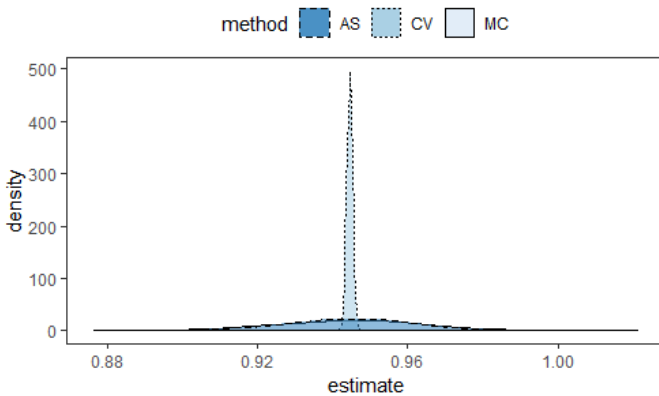
Result (b)-(d) & €. Asian Call Price

$$\lambda = -1$$

$$\hat{\lambda} = -0.96855$$

	naive MC	CV	CV	AS
mean	0.94484	0.94458	0.94459	0.94469
sd	0.02009	0.00078	0.00132	0.01707

(이어서)



$n=10000$, simulation 반복횟수를 1000 번으로 하여 1000 개의 추정치들의 분포를 확인하였다. AS의 경우는 S_t 를 생성하기 위해 sampling 해야 하는 random variable $Z \sim N(0, 1)$ 에 10000 개의 절반인 5000 개만을 사용하였다.

추정치들의 값들이 모두 약 0.945 주변으로 대칭적으로 분포해 있으며 위의 표에서 표준오차를 보면 $MC > AS > CV$ 추정치 순으로 작아진다. 특히 CV 추정치가 다른 두 가지 방법을 이용했을 때에

비해 10분의 1 이상 작아져 그래프에서도 상당히 좁은 구간 내에 분포하고 있음을 확인할 수 있다.

III. Discussion

4 가지 예제를 통해 MC Simulation 을 이용한 적분 계산에서 다양한 Variance Reduction Technique 의 효과를 확인하였다. 대부분의 경우에서 naive (simple) MC estimator $\hat{\mu}_{MC} = \frac{1}{n} \sum_{i=1}^n h(X_i)$ 에 비해 작게는 2 배에서 크게는 10 배 이상으로 분산을 효과적으로 줄이면서도 더 정확한 값을 추정할 수 있었다.

IS, CV 는 MC 와 동일한 sample size n 개의 sample 에서 서로 다른 계산 방법으로 추정치를 구하지만 Antithetic Sampling 은 추출해야 하는 sample 의 크기를 반으로 줄인다. 본 과제의 경우 AS 방법이 다른 분산 축소 전략에 비해 분산을 크게 줄이지는 못했지만 계산 시간 등의 면에서 효율적인 방법일 수 있다. Implementation 1-4 모두에서 Control Variates 방법으로 추정하였을 때 표준오차가 단순 추정에 비해 매우 줄어들어 상당히 좋은 분산 축소 방법이라고 판단하였다.

Simulation 을 통해 서로 독립적이고 동일한 분포를 따르는 sample 을 구성하였을 때, 단순히 이들의 평균으로 기댓값을 추정하는 것보다 더 정밀히 추정이 가능한 다양한 전략들을 예제에 직접 적용해 봄으로써 계산 과정을 좀 더 이해할 수 있었다.

[Appendix] R Code**# Implementation 1**

```

S0 = 100 ; K = 102 ; sig = 0.3 ; N = 50 ; r = 0.05 ; n = 100000
# analytic sol. of theta
c3 = 1 + 1/N
c2 = sig * ( c3*N/1095 * (1+1/(2*N)) )^0.5
c1 = 1/c2 * ( log(S0/K) + (c3*N/730)*(r-sig^2/2) + c3*sig^2*N/1095*(1+1/(2*N)) )
theta = S0 * pnorm(c1) * exp( -N*(r+c3*sig^2/6)*(1-(1/N))/730 ) - K * pnorm(c1-c2) * exp(-r*N/365)

# simul.
mycomp1.1 <- function(S0, K, sig, N, r, n=100000) {
  St = matrix(nrow = n, ncol = N) ; A = numeric(n) ; A.geo = numeric(n)
  for (i in 1:n) {
    Zt <- rnorm(N)
    St[i,] <- S0 * cumprod(exp((r-sig^2/2)/365 + sig*Zt/sqrt(365)))
    A[i] <- exp(-r*N/365) * max(0, mean(St[i,])-K)
    A.geo[i] <- exp(-r*N/365) * max(0, prod(St[i,])^(1/N)-K)
  }
  return(c(mu.mc = mean(A), theta.mc = mean(A.geo)))
}

nsim = 100
cal1 <- data.frame(mu.mc = numeric(nsim), theta.mc = numeric(nsim))
for (i in 1:nsim) { cal1[i,] <- mycomp1.1(100, 102, 0.3, 50, 0.05) }
cal1 <- cal1 %>% mutate(mu.cv1 = mu.mc - (theta.mc - theta), mu.cv2 = mu.mc - cov(mu.mc, theta.mc)/var(theta.mc) * (theta.mc - theta))
apply(cal1, 2, mean) ; apply(cal1, 2, sd)
cor(cal1$mu.mc, cal1$theta.mc)

```

Implementation 2

```

qfunc <- function(x) { exp(-(abs(x))^3/3) }
const <- integrate(qfunc, -Inf, Inf)$value
q <- seq(-4, 4, length = 100)
plot(q, 3*dnorm(q), type = "l", col = "blue") ; lines(q, qfunc(q), type = "l")

# a) IS w/ std.weights
n = 10000
a <- rnorm(n) ; usiw <- qfunc(a)/(dnorm(a)) ; siw <- usiw/sum(usiw)
mu.is <- sum(a^2*siw)

ggplot() + scale_fill_brewer("", palette = "Set1") + theme_test() + labs(x = "y", y = "density", title = "Importance Sampling") +
  theme(plot.title = element_text(hjust = 0.5), legend.position = "top") +
  geom_area(aes(a, dnorm(a), fill = "ISF"), alpha = 0.3, color = "black") +
  geom_area(aes(a, qfunc(a)/const, fill = "target"), alpha = 0.3, color = "black")

for (i in 1:1000) { a <- rnorm(n) ; usiw <- qfunc(a)/(dnorm(a)) ; siw <- usiw/sum(usiw) ; mu.is[i] <- sum(a^2*siw) }
mean(mu.is) ; sd(mu.is)

# b) RS
mycomp2.2 <- function(n = 10000) {
  y <- rnorm(n) ; u <- runif(n)
  keep <- (u <= qfunc(y)/(3*dnorm(y))) ; x <- y[keep]
  return(x)
}
b <- mycomp2.2() ; mu.rs <- mean(b^2)

```

```
ggplot() + scale_fill_manual("", values = c("e(y)" = "pink", "q(y)" = "white")) + theme_test() +
  labs(x = "y", y = "propotional density", title = "Rejection Sampling") + theme(plot.title = element_text(hjust = 0.5), legend.position = "top") +
  geom_area(aes(a, 3*dnorm(a), fill = "e(y)", color = "black")) + geom_area(aes(a, qfunc(a), fill = "q(y)", color = "black"))
```

c) Philippe and Robert

```
mycomp2.3 <- function(x.rs) {
  c1 <- sum(diff(sort(x.rs))*sort(x.rs)[-length(x.rs)]^2*qfunc(sort(x.rs)[-length(x.rs)]))
  c2 <- sum(diff(sort(x.rs))*qfunc(sort(x.rs)[-length(x.rs)]))
  return(c1/c2)
}
mu.pnr <- mycomp2.3(b)
```

d)

```
nsim = 1000 ; cal2.4 <- data.frame(a = numeric(nsim), b = numeric(nsim), c = numeric(nsim))
for (i in 1:nsim) {
  x <- rnorm(n) ; usiw <- qfunc(x)/(dnorm(x)) ; siw <- usiw/sum(usiw)
  x.rs <- mycomp2.2() ; cal2.4[i,] <- c(sum(x^2*siw), mean(x.rs^2), mycomp2.3(x.rs))
}

apply(cal2.4, 2, mean) ; apply(cal2.4, 2, sd)
system.time( for (i in 1:nsim) { x <- rnorm(n) ; usiw <- qfunc(x)/(dnorm(x)) ; siw <- usiw/sum(usiw) ; sum(x^2*siw) } )
system.time( for (i in 1:nsim) { x.rs <- mycomp2.2() ; mean(x.rs^2) } )
system.time( for (i in 1:nsim) mycomp2.3(x.rs) )
```

Implementation 3

a. size

```
myh <- function(z) { ifelse(z >= 1.645, 1, 0) }

mycomp3.1 <- function(n = 1000, nsim = 1000) {
  lambda1.is <- 2 + 1.645*sqrt(2/25)

  cal <- data.frame(matrix(numeric(nsim)*5, ncol = 5))
  colnames(cal) <- c("MC", "AS", "IS.ustd", "IS.std", "wbar")

  for (i in 1:nsim) {
    x <- matrix(rpois(25*n, 2), ncol = 25)
    z <- apply(x, 1, function(x) (mean(x)-2)/sqrt(2/25))

    mu1 <- mean( myh(z) ) # naive MC
    mu2 <- sum( myh(z[1:(n/2)]) + myh(-z[1:(n/2)]) ) / n # AS

    x.is <- matrix(rpois(25*n, lambda1.is), ncol = 25)
    xbar.is <- apply(x.is, 1, mean)
    z.is <- apply(x.is, 1, function(x) (mean(x)-2)/sqrt(2/25))

    usiw <- dnorm(xbar.is, 2, sqrt(2/25)) / dnorm(xbar.is, lambda1, sqrt(lambda1/25))
    siw <- usiw / sum(usiw)

    mu3.1 <- mean( myh(z.is)*usiw ) # IS w/ unstd.weights
    mu3.2 <- sum( myh(z.is)*siw ) # IS w/ std.weights

    cal[i,] <- c(mu1, mu2, mu3.1, mu3.2, mean(usiw))
  }
  cal <- cal %>% transmute(MC, AS, IS.ustd, IS.std,
    ISCV = IS.ustd - cov(IS.ustd, wbar)/var(wbar)*(wbar-1), wbar)
  return(cal)
}
```

```

table3.1 <- mycomp3.1()
info3.1 <- data.frame(mean = apply(table3.1, 2, mean), sd = apply(table3.1, 2, sd),
                      lower = apply(table3.1, 2, function(x) sort(x)[25]),
                      upper = apply(table3.1, 2, function(x) sort(x)[975]))[-6,]

ggplot(table3.1) + theme_test() + scale_color_brewer("", palette = "Set1") +
  geom_density(aes(MC, color = "MC")) + geom_density(aes(AS, color = "AS")) +
  geom_density(aes(IS.ustd, color = "IS*")) + geom_density(aes(IS.std, color = "IS")) + geom_density(aes(ISCv, color = "ISCv"))

# b. power
mycomp3.2 <- function(n = 1000, nsim = 1000, lambda) {
  b <- 2 + 1.645*sqrt(2/25)
  lambda.is <- lambda + 1.645*sqrt(2/25)

  cal <- data.frame(matrix(numeric(nsim)*5, ncol = 5))
  colnames(cal) <- c("MC", "AS", "IS.ustd", "IS.std", "wbar")

  for (i in 1:nsim) {

    x <- matrix(rpois(n*25, lambda), ncol = 25) # under H1: lambda = 2.2 > 2
    z <- apply(x, 1, function(x) (b-mean(x))/(sd(x)/5))

    mu1 <- mean( 1 - pnorm(z) )
    mu2 <- sum( 1 - pnorm(z[1:(n/2)]) + pnorm(-z[1:(n/2)]) ) / n

    x.is <- matrix(rpois(n*25, lambda.is), ncol = 25)
    xbar.is <- apply(x.is, 1, mean) ; z.is <- apply(x.is, 1, function(x) (b-mean(x))/(sd(x)/5))

    usiw <- dnorm(xbar.is, lambda, sqrt(lambda/25)) / dnorm(xbar.is, lambda.is, sqrt(lambda.is/25))
    siw <- usiw / sum(usiw)

    mu3.1 <- mean( (1-pnorm(z.is))*usiw )
    mu3.2 <- sum( (1-pnorm(z.is))*siw )

    cal[i,] <- c(mu1, mu2, mu3.1, mu3.2, mean(usiw))
    if (i %% 10 == 0) print(i)
  }
  cal <- cal %>% transmute(MC, AS, IS.ustd, IS.std,
                          ISCV = IS.ustd - cov(IS.ustd, wbar)/var(wbar)*(wbar-1), wbar)

  return(cal)
}

mylambda <- list() ; l <- seq(2.2, 4, length = 5)
for (i in 1:2) { mylambda[[i]] <- mycomp3.2(lambda = l[i]) }
for (i in 3:5) { mylambda[[i]] <- mycomp3.2(lambda = l[i]) }

table3.2 <- data.frame(mylambda[[1]], mylambda[[2]], mylambda[[3]], mylambda[[4]], mylambda[[5]]) %>% select(-6*(1:5))
table3.2 <- table3.2[c(seq(1, 21, by = 5), seq(2, 22, by = 5), seq(3, 23, by = 5), seq(4, 24, by = 5), seq(5, 25, by = 5))]
colnames(table3.2) <- c(paste("MC", 1:5, sep = ""), paste("AS", 1:5, sep = ""),
                      paste("IS.ustd", 1:5, sep = ""), paste("IS.std", 1:5, sep = ""), paste("ISCv", 1:5, sep = ""))

table3.2.1 <- data.frame(lambda = rep(l, 5),
                        mean = apply(table3.2, 2, mean), sd = apply(table3.2, 2, sd),
                        lower = apply(table3.2, 2, function(x) sort(x)[25]),
                        upper = apply(table3.2, 2, function(x) sort(x)[975]))

```

```
ggplot(table3.2.1[21:25,]) + geom_line(aes(lambda, mean), color = "pink", size = 1) +
  geom_point(aes(lambda, mean)) + geom_errorbar(aes(lambda, ymin = lower, ymax = upper), size = 0.5, width = 0.02) +
  labs(y = "power", title = "ISCV") + theme_test() + theme(plot.title = element_text(hjust = 0.5))
```

Implementation 4

```
S0 = 50 ; K = 52 ; sig = 0.5 ; N = 30 ; r = 0.05
```

```
# a.
```

```
n = 10000 ; nsim = 1000 ; C.mc <- c()
for (i in 1:nsim) {
  z <- rnorm(n) ; ST <- S0 * exp( (r-sig^2/2)*N/365 + sig*z*sqrt(N/365) )
  C <- exp(-r*N/365) * ifelse(ST>=K, ST-K, 0) ; C.mc[i] <- mean(C)
}
mean(C.mc) ; sd(C.mc)
ggplot() + geom_density(aes(C.mc), fill = "pink", alpha = 0.5) + theme_test()
```

```
# b.
```

```
A.cv <- data.frame(mu.mc = numeric(nsim), theta.mc = numeric(nsim))
for (i in 1:nsim) { A.cv[i,] <- mycomp1.1(50, 52, 0.5, 30, 0.05) }

c3 = 1 + 1/N
c2 = sig * ( c3*N/1095 * (1+1/(2*N)) )^0.5
c1 = 1/c2 * ( log(S0/K) + (c3*N/730)*(r-sig^2/2) + c3*sig^2*N/1095*(1+1/(2*N)) )
theta = S0 * pnorm(c1) * exp( -N*(r+c3*sig^2/6)*(1-(1/N))/730 ) - K * pnorm(c1-c2) * exp(-r*N/365)

A.cv <- A.cv %>% mutate(mu.cv1 = mu.mc - (theta.mc - theta), mu.cv2 = mu.mc - cov(mu.mc, theta.mc)/var(theta.mc) * (theta.mc - theta))
apply(A.cv, 2, mean) ; apply(A.cv, 2, sd)
```

```
# c & d.
```

```
mycomp4 <- function(S0, K, sig, N, r, n = 10000) {
  St = matrix(nrow = n, ncol = N) ; St.as = matrix(nrow = n, ncol = N)
  A.mean = numeric(n) ; A.geo = numeric(n) ; A.as = numeric(n/2)
  for (i in 1:(n/2)) {
    Zt <- rnorm(N)
    St.as[i,] <- S0 * cumprod(exp((r-sig^2/2)/365 + sig*Zt/sqrt(365))) ; St.as[n/2+i,] <- S0 * cumprod(exp((r-sig^2/2)/365 - sig*Zt/sqrt(365)))
  }
  for (i in 1:n) { A.as[i] <- exp(-r*N/365) * max(0, mean(St.as[i,])-K) }
  return(mean(A.as))
}
```

```
nsim = 1000
table4 <- data.frame(mu.as = numeric(nsim), mu.mc = numeric(nsim), theta.mc = numeric(nsim))
for (i in 1:nsim) {
  table4[i, 1] <- mycomp4(50, 52, 0.5, 30, 0.05, n = 10000) ; table4[i, 2:3] <- mycomp1.1(50, 52, 0.5, 30, 0.05, n = 10000)
}
table4 <- table4 %>% mutate(mu.cv1 = mu.mc - (theta.mc - theta),
  mu.cv2 = mu.mc - cov(mu.mc, theta.mc)/var(mu.mc) * (theta.mc - theta))
apply(table4, 2, mean) ; apply(table4, 2, sd)
```

```
ggplot(table4) + theme_test() + scale_fill_brewer("method", palette = "Blues", direction = -1) +
  scale_linetype_manual("method", values = c(MC = 1, AS = 2, CV = 3)) +
  labs(x = "estimate") + theme(legend.position = "top") +
  geom_density(aes(mu.mc, linetype = "MC", fill = "MC"), alpha = 0.5) +
  geom_density(aes(mu.cv1, linetype = "CV", fill = "CV"), alpha = 0.5) +
  geom_density(aes(mu.as, linetype = "AS", fill = "AS"), alpha = 0.5)
```

