Project 2 Monte Carlo

Maciej Szczutko

2025-01-27

Problem description

We are interested in estimating the following (called an option, with discounted payoff at time 1) with price given by the formula

$$I = e^{-r} E \left(A_n - K \right)_+,$$

where

$$A_n = \frac{1}{n} \sum_{i=1}^{n} S(i/n),$$

and

$$S(t) = S(0) \exp \left(\mu^* t + \sigma B(t) \right), \quad 0 \le t \le T$$

where $B(t)(0 \le t \le T)$ is Brownian motion.

European and Asian option

In the case n = 1, this is called a European call option; otherwise, it is called an Asian call option.

Used methods

1. Crude Monte Carlo estimator
2. Stratified estimator
3. Anthithetic variables
4. Control variates

Monte Carlo description

First we need to generate Brownian Motion n points, equally spaced sample on $[0, 1]$. We will use the fact that $\mathbf{B} = (B(1/n), B(2/n), \ldots, B(1))$ is a multivariate normal random variable $\mathcal{N}(\mathbf{0}, \mathbf{\Sigma})$ with the covariance matrix

$$\mathbf{\Sigma}(i, j) = \frac{1}{n} \min(i, j).$$

Table 1: MC estimators values for R=10000

1	2	4	8	16	32
12.235	9.725	8.397	7.532	7.204	6.913

Boxplot of Each Column
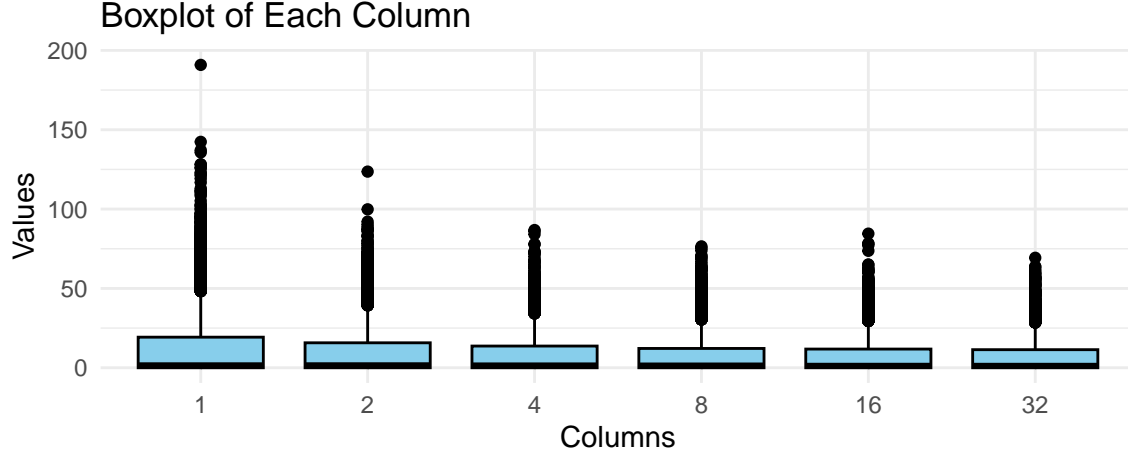
Theoretical calculation for European option using Black-Scholes formula

We can compare the simulation results for European option with theoretical calculation. The formula is

$$E(S(1) - K)_+ = S(0)\Phi(d_1) - Ke^{-r}\Phi(d_2),$$

where

$$d_1 = \frac{1}{\sigma}\left[\log\left(\frac{S(0)}{K}\right) + r + \frac{\sigma^2}{2}\right],$$

and

$$d_2 = d_1 - \sigma.$$

Using our setup parameters we obtain value 12.3359989. It is align with the MC simulation result. Even using $R = 10^4$ replication we have quite bias simulation.

Observation

For European option (instant pay off ?) we have higher expected value but with higher variability. For Asian option the variability is a bit lower but the estimated payoff is also lower. The value of pay off seems converge to some value as n increase. I would bet for some constants (e.g. 6/7) but I haven't perform simulation for extremely large n to confirm this.

Stratified sampling

Stratified sampling is a sampling technique in which a population is divided into smaller, homogeneous groups called strata based on shared characteristics (e.g., age, gender, income). In this example the strata is determined by the ellipses. Then, a random sample is taken from each stratum proportionally or equally to ensure all groups are represented in the final sample. This method improves accuracy and reduces sampling bias and variance.

The algorithm for generating sample from $N(0, \Sigma)$.

1. Perform Cholesky decomposition: $\boldsymbol{\Sigma} = \mathbf{A}\mathbf{A}^T$.
2. Sample $\boldsymbol{\xi} = (\xi_1, \ldots, \xi_n)^T$, where $\xi_i \sim \mathcal{N}(0, 1)$ i.i.d. Set

$$\mathbf{X} = (X_1, \ldots, X_n)^T = \left(\frac{\xi_1}{\|\boldsymbol{\xi}\|}, \ldots, \frac{\xi_n}{\|\boldsymbol{\xi}\|} \right)^T$$

3. Sample $U \sim \mathcal{U}(0,1)$. Set

$$D^2 = F_{\chi_n^2}^{-1} \left(\frac{i-1}{m} + \frac{1}{m} U \right)$$

4. Set $\mathbf{Z} = (Z_1, \ldots, Z_n) = (DX_1, \ldots, DX_n)$.
5. Return $\mathbf{B}^i = \mathbf{AZ}$.

Lets see how the algorithm works. The Σ is the same as before for Brownian motion.
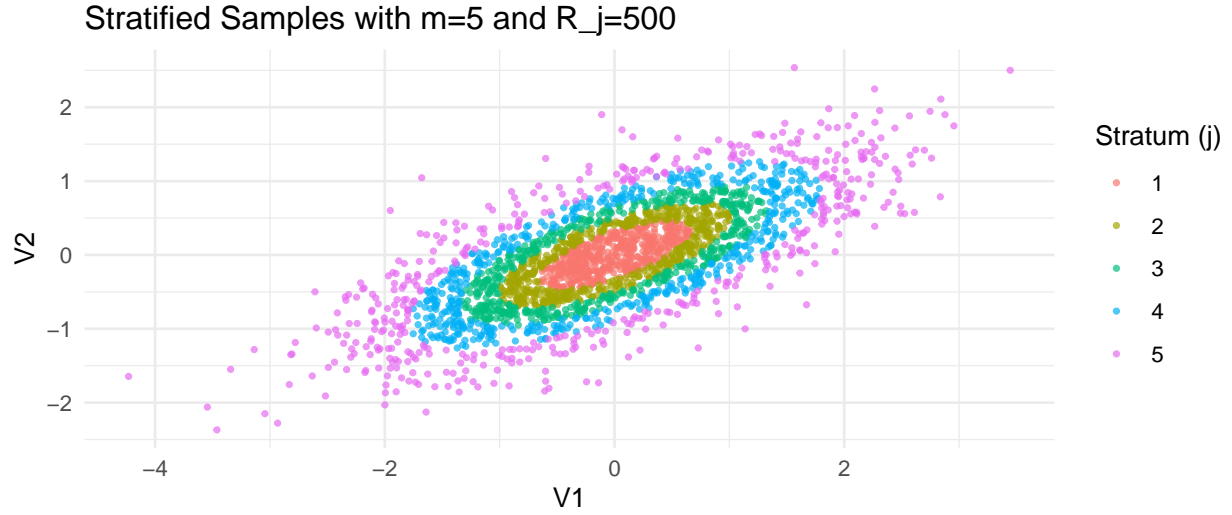

Figure 1: Starat viusalization

It looks that indeed procedure works as expected. Note that our stratas are ellipses, not the circle. But the assumption that each strata has probability $\frac{1}{m}$ still holds. Or change but since we apply the same linear transformation this will not impact calculation as the factor simplify at the end.

Estimator formulas

$$\hat{Y}_{R_j}^j = \frac{1}{R_j} \sum_{i=1}^{R_j} Y_i^j, \quad \hat{Y}_R^{\text{str}} = p_1 \hat{Y}_{R_1}^1 + \ldots + p_m \hat{Y}_{R_m}^m,$$

$$\text{Var}\,\hat{Y}_R^j = \frac{\sigma_j^2}{R_j}, \quad \text{Var}\,\hat{Y}_R^{\text{str}} = \sum_{j=1}^m p_j^2 \,\text{Var}\,\hat{Y}_R^j = \sum_{j=1}^m \frac{p_j^2}{R_j} \sigma_j^2.$$

Proportional scheme for n=1.

For stratified sampling we can use also the different number of replication for each stratum. In proportional scheme we simply take $R_j = p_j R$. For below simulation I will use $m = 5$ and $R = 10000$ (to be comparable with CRMC).

The estimator for $\hat{I} = 12.097266$. Calculated variance is 0.0266185. For MC the estimator variance is 0.0343468. So the variance is decreased even with simple choice of strata numbers.

Table 2: Comarison between MC and Strartified with differen allocation scheme.

n	Monte Carlo		Stratified Proportional		Stratified Optimal	
	$E(I)$	$Var(I)$	$E(I)$	$Var(I)$	$E(I)$	$Var(I)$
1	12.23521	0.03435	12.47133	0.02675	12.39387	0.01703
2	9.72451	0.02040	9.57401	0.01696	9.43218	0.01329
4	8.39666	0.01463	8.20635	0.01323	8.15068	0.01176
8	7.53158	0.01205	7.47103	0.01153	7.46856	0.01076
16	7.20404	0.01085	7.30583	0.01075	7.20869	0.01032
32	6.91305	0.00983	7.09645	0.01056	6.91351	0.00963

Optimal choice for $n = 1$.

Now we will use optimal allocation scheme. From theorem presented in the script we can take

$$R_j = \frac{p_j \sigma_j}{\sum_{i=1}^{m} p_i \sigma_i} R.$$

As we don't know the variance for each stratum (or it's to complex to calculate by hand for me) we can use following procedure.

1. Use proportional scheme to estimate σ_i.
2. Calculate R_j based using estimator for σ_i.
3. Sample again and estimate.

The estimated value for $I = 12.2002346$. The variance is 0.0169925.

$n \geq 2$ using proportional and best scheme

Now we can calculate the estimator for higher n. Proportion as before. The result are presented in the table.

Observation

The variance techniques works according to the theory. The variance is reduced when we compare to MC. Also changing the allocation scheme give some improvement. Here the result is presented for single R. I made some test and for higher R (e.g. 10^6) we obtain even better improvement ratio in terms of variance. For the MC method we have smaller bias for smaller number of replication R. I think it also highly depends on choosing strata for simulation. In practice I would use ordinary MC method.

I haven't analyse the m strata number influence here.

Antithetic estimator

The key idea here is to introduce dependency between elements in the sample. Analysing formula for variance for two random variables

$$\mathrm{Var}(X + Y) = \mathrm{Var}(X) + \mathrm{Var}(Y) + 2\,\mathrm{Cov}(X, Y)$$

should give us the intuition: if the covariance of X and Y is negative the overall variance should be reduced.

The estimator formula is exactly the same as for crude MC.

In this context we will use this method for European option. We will use $Z_{2i} = -Z_{2i-1}$ where Z_i is standard normal. For single generated value from brownian motion we apply following procedure.

Make the pair (Z_{2i-1}, Z_{2i}) and compute two stock prices:

$$S_{2i-1}(1) = S(0)\exp\left(\mu + \sigma Z_{2i-1}\right),$$
$$S_{2i}(1) = S(0)\exp\left(\mu + \sigma Z_{2i}\right).$$

The for each $S_i(1)$ calculate the payoff as before.

Same number of replication is used to compare with MC method.

The estimator value is 12.3813703. The variance is 0.0347665.

Observation. Variance is not reduce a lot.

Control Variate

This is another simple technique. The key idea here is the fact that we know the exact expected value of other random variable. Then we use following correction for crude MC estimator:

$$\hat{Y}_R^{\mathrm{CV}} = \frac{1}{R}\sum_{i=1}^{R}\left(Y_j - \beta\left(X_j - \mathbb{E}[X]\right)\right),$$

where β is a coefficient chosen to minimize the variance of the estimator. The optimal β is given by:

$$\beta = \frac{\mathrm{Cov}(Y, X)}{\mathrm{Var}\, X}$$

Usually we don't know the $Cov(Y, X)$. So we will use estimator from simulation.

In this context $X = B \sim N(0,1)$ will be used as control variate. Since $\mathbb{E}[B] = 0$ the formula become

$$\hat{Y}_R^{\mathrm{CV}} = \frac{1}{R}\sum_{j=1}^{R}\left(Y_j - \beta X_j\right).$$

The main difference between crude MC simulation is fact that we need to store sample from Brownian motion from each replicate.

Estimated value $I = 12.4111408$. The variance for estimator is 0.0343468. No significant improvement for European option compared to MC method. But we have lower bias comparing to previous method.

Higher number of replication $R = 1000000$

Lets check whether the difference in variance become more significant if we increase the number of replication.

R	$Var(\hat{I}^{MC})$	$Var(\hat{I}^{CV})$
10000	0.0343468	0.0343468
1000000	3.4257597×10^{-4}	3.4183318×10^{-4}

That's a bit odd. According to theory variance should be reduced. From theoretical calculation

$$\mathrm{Var}\left(\hat{Y}_R^{CV}\right) = \ldots = \mathrm{Var}\,\hat{Y}_R^{\mathrm{CMC}}\left(1 - \rho^2\right),$$

where ρ is correlation between pay of and Brownian motion. In this case variables should be highly correlated. Estimate for correlation is 0.8472165 (in case $R = 1000000$) thus I should observe significant reduction of variance (4 times less). Please note that estimated value is close to theoretical value even with lower bias. I didn't find any issue in implementation. I don't know how to comment the results.