Wprowadzenie do symulacji i metod Monte Carlo
Projekt 2

Zuzanna Brzóska, anonymous second person

22 maja 2023

Spis treści

1 Wprowadzenie

Celem niniejszego projektu jest oszacowanie opcji kupna o cenie wyrażonej poprzez

$$I = e^{-r}\mathbb{E}((A_n - K)_+), \tag{1}$$

gdzie

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

a $S(t)$ jest geometrycznym ruchem Browna $GBM(\mu, \sigma)$ opisanym wzorem

$$S(t) = S(0)\exp\left(\left(r - \frac{\sigma^2}{2}\right)t + \sigma B(t)\right), \quad 0 \le t \le T.$$

Tutaj $B(t)$ $(0 \le t \le T)$ stanowi ruch Browna.
Jeżeli $n = 1$ to (1) nazywamy Europejską opcją kupna, a w przeciwnym przypadku - Azjatycką opcją kupna.
Dla Europejskiej opcji kupna istnieje wzór, zwany formułą Blacka-Scholesa, pozwalający obliczyć dokładną jej wartość:

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

gdzie

$$d_1 = \frac{1}{\sigma}\left[\log\frac{S(0)}{K} + r + \frac{\sigma^2}{2}\right], \quad d_2 = d_1 - \sigma$$

oraz Φ jest dystrybuantą standardowego rozkładu normalnego.
Dla branych w projekcie wartości parametrów, czyli: $r = 0.05$, $\sigma = 0.25$, $S(1) = 100$ i $K = 100$, I obliczone za pomocą Blacka-Scholesa wynosi około 12.336.

2 Crude Monte Carlo Estimator

Weźmy $Y = k(V)$, gdzie V ma rozkład jednostajny na B (w tym projekcie $B = [0, 1]$). Wówczas $\mathbb{E}[Y] = I$. Ponadto, niech $V_1, \ldots, V_R$ stanowią niezależne kopie V o rozkładzie jednostajnym $\mathcal{U}(B)$. Wtedy estymator

$$\hat{Y}_R^{CMC} = \frac{1}{R}\sum_{j-1}^{R} Y_j = \frac{1}{R}\sum_{j-1}^{R} k(V_j)$$

nazywany jest estymatorem Crude Monte Carlo.
W naszym przypadku $k(V) = e^{-r}(A_n - K)_+$. Musimy zatem wygenerować R razy zmienne $B(t)$ z ruchu Browna w punktach $t = 1/n, ..., 1$.

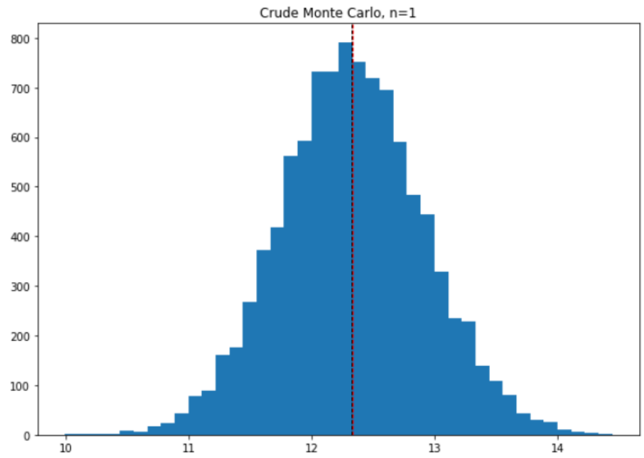
Dla opcji europejskiej (czyli $n = 1$) musimy więc wygenerować R razy wartość ruchu Browna $B(1)$, który jest zwyczajnie zmienną ze standardowego rozkładu normalnego $N(0, 1)$. W tej sytuacji wzór na funkcję k sprowadza się do $k(B_i) = e^{-r}(S(1) - K)_+$ dla $i = 1, ..., R$.

Dla opcji azjatyckiej (gdzie $n > 1$) odcinek $[0, 1]$ dzielimy na n równych części, co może być interpretowane jako analiza ceny co pewien okres np. $n = 12$ - co miesiąc, $n = 4$ - co kwartał. Potrzebujemy więc wygenerować R razy ruch Browna dla n punktów. W tym celu wygenerujemy R zmiennych z n-wymiarowego rozkładu normalnego $N(0, \Sigma)$ o macierzy kowariancji

$$\Sigma(i, j) = Cov(B(t_i), B(t_j)) = min(t_i, t_j)$$

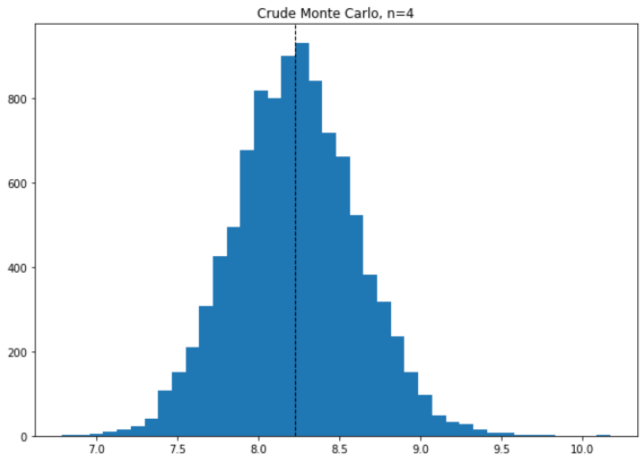
dla $(t_1, ..., t_n) = (1/n, ..., 1)$.
Do obliczeń użyjemy $R = 1000$ i powtórzymy je 10000 razy. Na przedstawionych poniżej histogramach czarną przerywaną linią zaznaczona jest średnia wartość estymatora obliczona w symulacji, natomiast dla $n = 1$ czerwona przerywana linia oznacza wyliczoną wcześniej teoretyczną wartość I.

Opcja europejska, R = 1000, 10000 symulacji

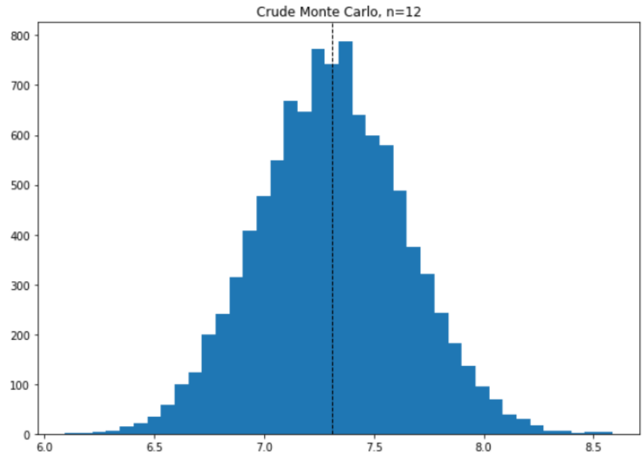


Dla powyższej symulacji średnia wartość estymatora wyniosła 12.33 i jak widać na histogramie niemal nakłada się z faktyczną wartością opcji (przerywane linie). Wariancja próbkowa w tym przypadku wyniosła 0.34.

Opcja azjatycka, n = 4, R = 1000, 10000 symulacji



Opcja azjatycka, n = 12, R = 1000, 10000 symulacji



Dla symulacji estymatora dla opcji azjatyckiej z $n = 4$ średnia oszacowana wartość wynosi 8.23, a średnia dla $n = 12$ to 7.31. Wariancje dla tych symulacji wyniosły odpowiednio 0.14 i 0.11. Średnie wartości estymatorów porównamy później z wartościami estymatorów w metodzie warstw (inaczej: stratified estimator), ponieważ formuła Blacka-Scholesa daje nam teoretyczny wynik tylko dla $n = 1$.

Wariancja tego estymatora wynosi $Var(\hat{Y}_R^{CMC}) = \frac{1}{R}\sum_{j=1}^{R} Var(Y_j)$, co zgadza się z wynikami symulacji podczas zmiany R, które zawarte są w poniższych tabelach.

Zależność wariancji estymatora CMC od R dla n=1

R	średnia wartość estymatora	wariancja
100	12.39	3.3392
1000	12.35	0.3708
10000	12.33	0.0343

Zależność wariancji estymatora CMC od R dla n=4

R	średnia wartość estymatora	wariancja
100	8.25	1.4693
1000	8.25	0.1519
10000	8.23	0.0140

Zależność wariancji estymatora CMC od R dla n=12

R	średnia wartość estymatora	wariancja
100	7.31	1.2355
1000	7.31	0.1099
10000	7.31	0.0110

Możemy zauważyć, że wraz ze wzrostem n, średnia wartość estymatora się zmniejsza. Ponadto, otrzymane wyniki są lepsze dla większej liczby obserwacji.

3 Stratified Estimator

Tym razem w celu estymacji $\mathbb{E}[Y] = I$ przyjmijmy, że $A^1, \dots, A^m$ są rozłącznymi zbiorami, dla których zachodzi $\mathbb{P}[Y \in \bigcup_{j=1}^m A^j] = 1$. Owe rozbicie definiuje warstwy $W^j = \{Y \in A^j\}$, gdzie prawdopodobieństwem wylosowania warstwy jest $p_j = \mathbb{P}[Y \in A^j]$, a $I^j = \mathbb{E}[Y|Y \in A^j]$ $(j = 1, \dots, m)$. Mamy

$$\mathbb{E}[Y] = p_1 I^1 + \ldots + p_m I^m.$$

Wobec tego, w poszczególnych warstwach będziemy generować zmienne $Y^j =_d (Y|Y \in A_j)$. Estymator I^j definiujemy poprzez

$$\hat{Y}^j_{R_j} = \frac{1}{R_j} \sum_{i=1}^{R_j} Y_i^{R_j},$$

z kolei estymator warstwowy będzie postaci

$$\hat{Y}^{str}_R = p_1 \hat{Y}^1_{R_1} + \ldots + p_m \hat{Y}^m_{R_m}.$$

Zakładając, że $Var(Y^j) = \sigma_j^2$, wariancja otrzymanego estymatora wynosi

$$Var(\hat{Y}^{str}_R) = \sum_{j=1}^m \frac{p_j^2}{R_j} \sigma_j^2$$

Istotnym krokiem będzie właściwy dobór liczby replikacji R_j w warstwach. Mianowicie:
(a) w przypadku *alokacji proporcjonalnej* (używanej, gdy σ_j^2 nie jest znane) mamy

$$R_j = R p_j$$

i przyjmujemy $p_j = 1/m$;
(b) dla zadanych warstw (gdzie p_j oraz σ_j^2 są znane) *optymalnym* doborem będzie

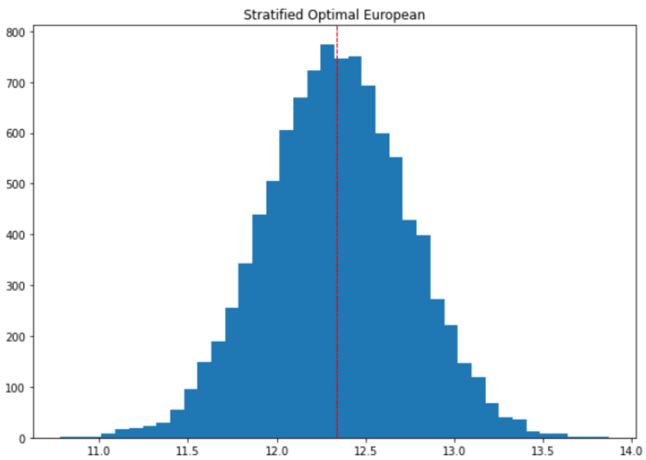
$$R_j = \frac{p_j \sigma_j}{\sum_{i=1}^m p_i \sigma_i} R.$$

Tutaj z kolei najpierw należy wyznaczyć σ_j, które obliczamy jako próbkowe odchylenia standardowe zmiennych Y^j przy użyciu alokacji proporcjonalnej.

3.1 opcja Europejska

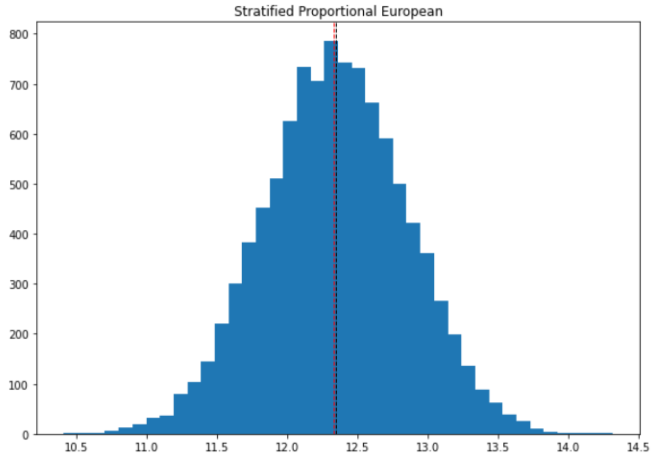
Analizę wyników zaczniemy od opcji Europejskiej, poniżej znajdują się histogramy wartości estymatorów.

Opcja europejska, wariant optimal, R = 1000, 10000 symulacji



Zauważamy pokrycie się linii przerywanej z ciągłą. Oznacza to, że średnia wartość estymatora jest równa wartości rzeczywistej.

Opcja europejska, wariant proportional, R = 1000, 10000 symulacji



W tym wariancie obserwujemy już drobne odstępstwo. Badane wartości są sobie nadal bardzo bliskie, ale można zauważyć, że linie przebiegają obok siebie.

Sprawdzamy zachowanie estymatorów dla zmiany ilości losowań R.

Zależność wariancji estymatora Stratified optimal od R dla n=1

R	średnia wartość estymatora	wariancja
100	12.33	1.6934
1000	12.32	0.1480
10000	12.34	0.0164

Zależność wariancji estymatora Stratified proportional od R dla n=1

R	średnia wartość estymatora	wariancja
100	12.34	2.3947
1000	12.35	0.2529
10000	12.34	0.0252

W obu wariantach średnie wartości estymatora są niemalże takie same, a otrzymana wariancja maleje wraz ze wzrostem R. Ponadto, owe wyniki dla wariancji w przypadku optimal są średnio około 2.5 razy niższe, a w przypadku proportional średnio około 1.5 razy niższe niż odpowiednie wartości uzyskane za pomocą CMC.

W przypadku estymatora warstwowego, możemy również przeanalizować wpływ wyboru ilości warstw na wyniki.

Zależność wariancji estymatora Stratified optimal od m dla n=1

R	m	wariancja
1000	5	0.1729
1000	10	0.1589
1000	50	0.1523
10000	5	0.0165
10000	10	0.0144
10000	50	0.0155

W owym wariancie dla ustalonego $R = 1000$ wariancja maleje wraz ze wzrostem m. Gdy R ulega zwiększeniu do 10000, uzyskana wariancja najpierw maleje (wraz ze wzrostem m z 5 do 10), po czym dla $m = 50$ przyjmuje większą wartość.

Zależność wariancji estymatora Stratified proportional od m dla n=1

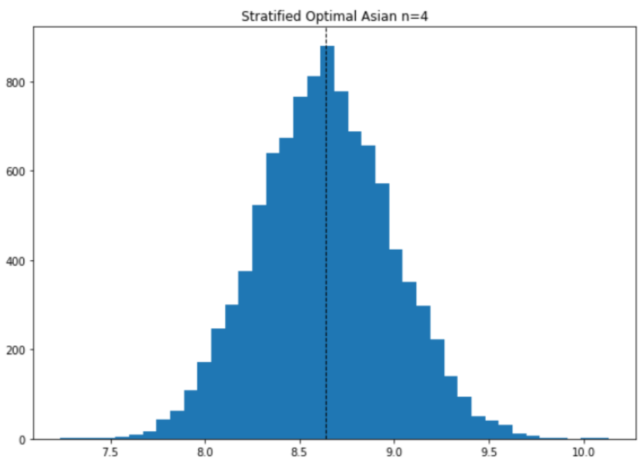
R	m	wariancja
1000	5	0.2573
1000	10	0.2672
1000	50	0.2404
10000	5	0.0286
10000	10	0.0239
10000	50	0.0249

Bez względu na R, dla coraz większego m wariancja ponownie najpierw maleje, z później rośnie.

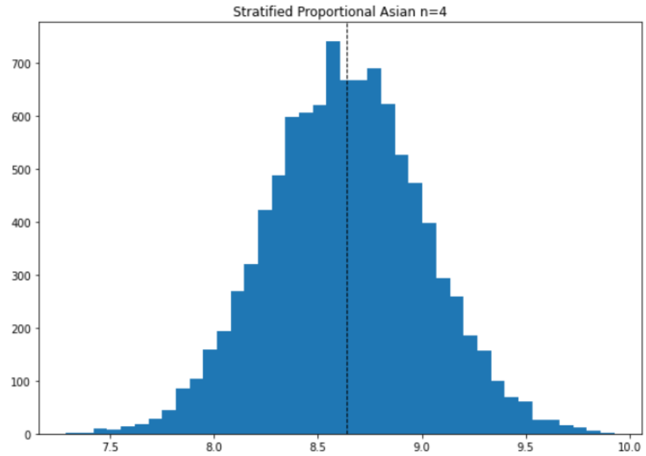
3.2 opcja Azjatycka

Podobnie jak poprzednio, opcję Azjatycką rozpatrzymy dla $n = 4$ oraz $n = 12$.

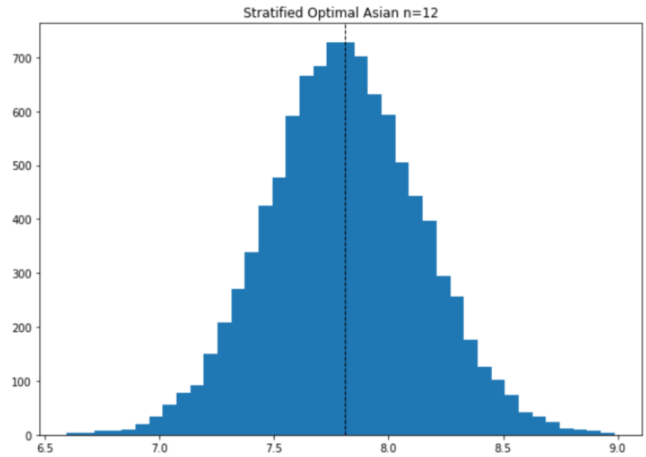
Opcja azjatycka, wariant optimal, n=4, R = 1000, 10000 symulacji



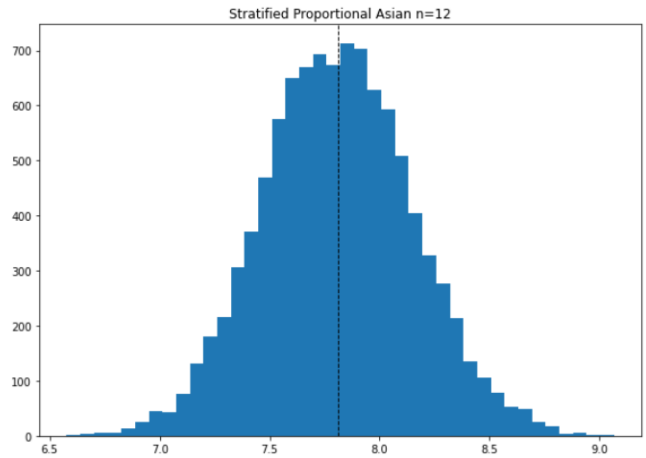
Opcja azjatycka, wariant proportional, n=4, R = 1000, 10000 symulacji



Opcja azjatycka, wariant optimal, n=12, R = 1000, 10000 symulacji



Opcja azjatycka, wariant proportional, n=12, R = 1000, 10000 symulacji



Dla ustalonego n wyniki uzyskane w poszczególnych wariantach są podobne. Mianowicie zarówno dla proportional i optimal, gdy $n = 4$ mamy nieco powyżej 8.5, a gdy n wzrasta do 12 mamy nieco poniżej 8.

Możemy także porównać wariancję dla różnych R z wynikami dla estymatora CMC.

Zależność wariancji estymatora Stratified optimal od R dla n=4

R	średnia wartość estymatora	wariancja
100	8.68	1.4552
1000	8.64	0.1211
10000	8.64	0.0123

Zależność wariancji estymatora Stratified optimal od R dla n=12

R	średnia wartość estymatora	wariancja
100	7.83	1.4022
1000	7.81	0.1073
10000	7.80	0.0114

Zależność wariancji estymatora Stratified proportional od R dla n=4

R	średnia wartość estymatora	wariancja
100	8.65	1.4045
1000	8.64	0.1389
10000	8.64	0.0145

Zależność wariancji estymatora Stratified proportional od R dla n=12

R	średnia wartość estymatora	wariancja
100	7.82	1.1081
1000	7.81	0.1228
10000	7.81	0.0112

Dla estymatora Stratified optimal oraz estymatora Stratified proportional, gdy $n = 4$ wyniki wariancji zamieszczone w tabeli prawie pokrywają się z wartościami uzyskanymi w przypadku estymatora CMC, przy czym średnia wartość estymatora jest teraz wyższa. Gdy $n = 12$ wariancja jest bliska (a dla $R = 10000$ prawie identyczna) wariancji otrzymanej w metodzie CMC dla owej liczby n. Tym razem średnia wartość estymatora ponownie jest wyższa.

Sprawdzimy również jak na wyniki w tym wypadku wpływa wybór ilości warstw m.

Zależność wariancji estymatora Stratified optimal od m dla n=4

R	m	wariancja
1000	5	0.1271
1000	10	0.1221
1000	50	0.1186
10000	5	0.0134
10000	10	0.0115
10000	50	0.0119

Zależność wariancji estymatora Stratified optimal od m dla n=12

R	m	wariancja
1000	5	0.1096
1000	10	0.1046
1000	50	0.1238
10000	5	0.0114
10000	10	0.0107
10000	50	0.0119

Zależność wariancji estymatora Stratified proportional od m dla n=4

R	m	wariancja
1000	5	0.1407
1000	10	0.1331
1000	50	0.1522
10000	5	0.0134
10000	10	0.0134
10000	50	0.0142

Zależność wariancji estymatora Stratified proportional od m dla n=12

R	m	wariancja
1000	5	0.1153
1000	10	0.1239
1000	50	0.1223
10000	5	0.0110
10000	10	0.0116
10000	50	0.0120

Wraz ze wzrostem warstw dla estymatora Stratified optimal ($n = 4$) zauważamy spadek wariancji przy ustalonym $R = 1000$, natomiast w wariancie proportional, gdy $R = 10000$ zaobserwować można łagodny wzrost wartości wariancji. W pozostałych przypadkach nie ma wyraźnie zarysowanej tendencji.

4 Antithetic Estimator

Inną metodą symulacji $I = \mathbb{E}[Y]$ jest metoda antytetyczna, która nie zakłada niezależności replikacji. Rozważymy ją jedynie dla opcji europejskiej. Weźmy parzystą liczbę R zmiennych losowych $Y_1, \ldots, Y_R$ i załóżmy, że pary (Y_{2i-1}, Y_{2i}), $i = 1, \ldots, \frac{R}{2}$ są niezależne i pochodzą z tego samego rozkładu, oraz rozkład brzegowy zmiennych losowych $(Y_i)_{i=1}^{R}$ jest taki sam jak Y.

Wówczas estymator antytetyczny będzie postaci

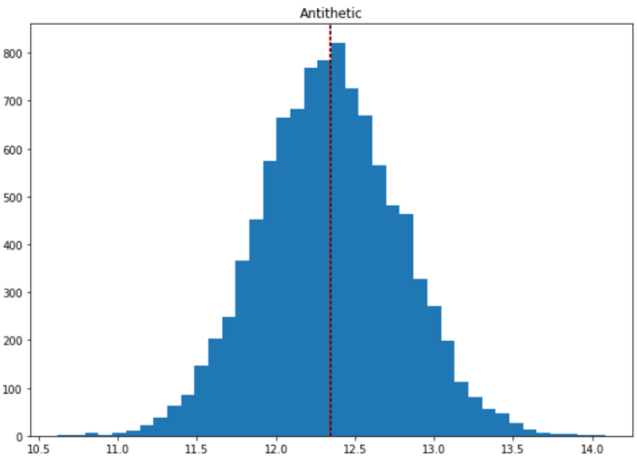
$$\hat{Y}_R^{anth} = \frac{1}{R} \sum_{i=1}^{R} Y_i.$$

Wtedy jego wariancja wynosi

$$Var(\hat{Y}_R^{anth}) = \frac{\frac{R}{2} Var(Y_1 + Y_2)}{R^2} = \frac{1}{R} Var(Y)(1 + corr(Y_1, Y_2)),$$

zatem przy $corr(Y_1, Y_2) < 0$ wariancja estymatora zostanie zmniejszona w stosunku do estymatora CMC. W projekcie jako nasze pary zmiennych losowych weźmiemy (Z_{2i-1}, Z_{2i}), przy czym $Z_{2i} = -Z_{2i-1}$, a $Z_{2i-1}, i = 1, \ldots, \frac{R}{2}$ są iid i pochodzą z rozkładu normalnego $N(0, 1)$. Funkcja k jest taka jak wcześniej, a przy obliczaniu $S(1)$ zamiast zmiennych $B(1)_i$ bierzemy zmienne Z.

Opcja europejska, R = 1000, 10000 symulacji



W symulacji ponownie średnia wartość estymatora pokrywa się niemal idealnie z wartością teoretyczną, co widać jako pokrycie się linii na powyższym histogramie. Wariancja w tym przypadku powinna wyjść dużo mniejsza niż dla estymatora CMC, ponieważ tak zdefiniowane zmienne Z jak powyżej mają współczynnik korelacji równy -1. I faktycznie, teraz wariancja wynosi 0.19, czyli prawie dwukrotnie mniej niż przy CMC, zatem obserwujemy znaczną poprawę przy zachowanej dokładności wyniku. Również tak samo jak poprzednio, im większe R, tym mniejsza wariancja, ponieważ tak samo licząc wariancję mamy wyraz $\frac{1}{R}$.

Zależność wariancji estymatora antytetycznego od R

R	średnia wartość estymatora	wariancja
100	12.36	1.9512
1000	12.35	0.1950
10000	12.34	0.0180

5 Control Variate Estimator

Ostatnią metodą jest control variate method czyli metoda zmiennych kontrolnych. Opiera się ona na jednoczesnym symulowaniu drugiej zmiennej losowej X, dla której wartość oczekiwana jest znana. Wobec tego otrzymujemy niezależne pary postaci $(Y_1, X_1), \ldots, (Y_R, X_R)$. Załóżmy, że $\hat{X}_R = \frac{1}{R}\sum_{i=1}^{R} X_i$, wówczas

$$\hat{Y}_R^{CV} = \hat{Y}_R^{CMC} + c(\hat{X}_R - \mathbb{E}[X]),$$

przy czym

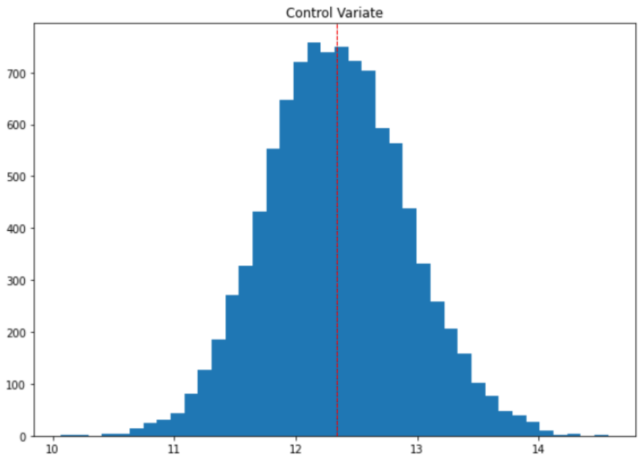
$$c = -\frac{Cov(Y, X)}{Var(X)}$$

i c jest dobrane tak, aby wariancja była minimalna.

Ponownie rozpatrujemy jedynie europejską opcję, gdzie $B(1)$ o standardowym rozkładzie normalnym stanowić będzie zmienną losową X. Skoro $X\ N(0,1)$ to wiemy, że $E[X] = 0$ oraz $Var(X) = 1$. Dlatego dla takiego wyboru zmiennej X wzór na estymator sprowadza się do:

$$\hat{Y}_R^{CV} = \hat{Y}_R^{CMC} - Cov(Y, X)(\hat{X}_R)$$

Opcja europejska, R = 1000, 10000 symulacji



Również w tym wypadku średnia wartość estymatora wynosi 12.34 i widać, że zgadza się z wartością teoretyczną. Natomiast wariancja wynosi tutaj 0.33 czyli jedynie minimalnie mniej niż wariancja przy CMC. Licząc wariancję teoretycznie otrzymujemy

$$\begin{aligned} Var(\hat{Y}_R^{CV}) = & \frac{Var(Y + cX)}{R} = \\ & \frac{1}{R}(VarY + c^2 Var(X) + 2cCov(X,Y)) = \\ & \frac{1}{R}\left(Var(Y) - \frac{Cov(X,Y)}{Var(X)}\right). \end{aligned} \tag{2}$$

U nas $Var(X) = 1$, zatem $Var(\hat{Y}_R^{CV}) = \frac{1}{R}(Var(Y) - Cov(X,Y))$, co wskazuje na to, że im bardziej zmienne X oraz Y są niezależne, tym bardziej wariancja będzie zbliżona do tej otrzymywanej przy użyciu estymatora CMC. W naszym wypadku zmienne X_i losowane są niezależnie od Y_i, więc faktycznie nie powinny mieć wysokiej korelacji, także symulacja zgadza się z oczekiwaniami. Ponownie widzimy wyraźnie element $\frac{1}{R}$ przy wzorze na wariancję estymatora, co można zaobserwować również w wynikach symulacji przedstawionych poniżej. Zauważamy też, że niezależnie od R, wariancja wciąż jest tylko minimalnie mniejsza niż ta przy estymacji metodą CMC.

Zależność wariancji estymatora CV od R

R	średnia wartość estymatora	wariancja
100	12.42	3.4478
1000	12.34	0.3294
10000	12.33	0.0321

Literatura

[1] Tomasz Rolski, *Symulacje stochastyczne i teoria Monte Carlo*, `http://www.math.uni.wroc.pl/~rolski/Zajecia/sym.pdf`

[2] Projekt, *Introduction to simulations and Monte Carlo methods. Project nr 2*, `https://moodle.math.uni.wroc.pl/pluginfile.php/17722/mod_resource/content/2/2022_monte_carlo_project2.pdf`