International Journal of Advanced Statistics and Probability, 3 (1) (2015) 8-14
www.sciencepubco.com/index.php/IJASP
©Science Publishing Corporation
doi: 10.14419/ijasp.v3i1.3773
Research Paper

Determining the optimum confidence interval based on the hybrid Monte Carlo method and its application in financial calculations

Kianoush Fathi Vajargah

Department of Statistics, Islamic Azad University, North branch, Tehran, Iran
E-mail: k_fathi@iau-tnb.ac.ir

Copyright © 2014 Kianoush Fathi vajargah. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

Abstract

The accuracy of Monte Carlo and quasi-Monte Carlo methods decreases in problems of high dimensions. Therefore, the objective of this study was to present an optimum method to increase the accuracy of the answer. As the problem gets larger, the resulting accuracy will be higher. In this respect, this study combined the two previous methods, QMC and MC, and presented a hybrid method with efficiency higher than that of those two methods.

Keywords: *Central Limit Theorem; Confidence Interval; Efficiency; Hybrid Method; Pricing Model.*

1. Introduction

In high-dimension problems, the quasi-Monte Carlo method usually yields estimates more favorable than Monte Carlo method does, but the superiority of quasi-Monte Carlo decreases as the problem gets larger. The hybrid sequences are used to solve the above problem in very high dimensions. Hybrid sequences are the result of combining pseudorandom vectors with low-discrepancy sequences. The low-discrepancy sequences are members of quasi-Monte Carlo's sequences. Several famous researchers have used hybrid methods for computations in quasi-Monte Carlo methods of high dimensions. For instance, these sequences were introduced by Spanier in 1995, and the hybrid sequence was applied by Okten in 1996 and 2004. In this study, an optimum interval confidence was presented using the central limit theorem.

2. Monte Carlo method

For a given population L, a parameter, such asθ, is estimated. In the Monte Carlo method, an estimate detector S(x) is first determined, in which x is a random variable with density function$f_x(x)$.
The estimate detector should have the following two conditions:
A: The estimate detector should be unbiased.
$E[S(x)] = 0$
B: The estimate detector should have definite variance.
$\mathrm{var}(S(x)) = \sigma^2$
Regarding the random samples $X_1 \dots . N_N$ of the function, density of $f_x(x)$ is used.
$\hat{\theta}_N(X_1 \dots . N_N) = \frac{1}{N}\sum_{n=1}^{N} S(X_n)$
As it is known:
$\mathrm{var}(\hat{\theta}_N) = \frac{\sigma^2}{N} < \infty, E(\hat{\theta}_N) = \theta$
The above estimate detector is introduced as the Monte Carlo estimator.

2.1. Estimation of the convergence rate of Monte Carlo

The central limit theorem defines the measure of statistical characteristics of the convergence rate of Monte Carlo.

Theorem 1: *for large Ns:* $\varepsilon_N [f] \approx \sigma N^{-\frac{1}{2k}}$

Where, k is the standard normal random variable, and σ is the square root of the variance of f and is constant.
In other words:

$\lim_{n\to\infty} \text{pr}\,(a < \frac{\sqrt{N}}{\sigma}\varepsilon_N < b) => \int_a^b (2\pi)^{-\frac{1}{2}} e^{-\frac{x^2}{2}} dx$

The above equation denotes that the error in Monte Carlo integration is $O(N^{-\frac{1}{2}})$.

Proof: It is assumed that Xis have normal distribution, thus: $Z_i = \frac{(f(x_i) - \bar{f})}{\sigma}$

$$E(z_i) = 0 \qquad E(z_i^2) = \int \frac{(f(x_i) - \bar{f})^2}{\sigma} dx = 1 \quad E[z_i z_j] = 0 \quad if \quad i \neq j$$

The later equality results from the independence of Xis.
Now, the following set is assumed:

$$Q_N = N^{-1} \sum_{i=1}^{N} z_i = \sigma^{-1} \varepsilon_N$$

$$E\left[Q_N^2\right]^{\frac{1}{2}} = E\left[N^{-2} (\sum_{i=1}^{w} z_i)^2\right]^{+\frac{1}{2}}$$

$$= N^{-1} \left\{ E\left[\sum_{i=1}^{N} z_i^2\right] + E\left[\sum_{i=1}^{N} \sum_{i \neq j} z_i z_j\right] \right\}^{\frac{1}{2}}$$

$$N^{-1} \left\{ \sum_{i=1}^{N} 1 + o \right\}^{\frac{1}{2}} = N^{-\frac{1}{2}}$$

Therefore, $E(\varepsilon_N^2) = \sigma N^{-\frac{1}{2}}$, and consequently, RMSE would be $O(\sigma N^{-\frac{1}{2}})$.

2.2. The quasi-Monte Carlo method

It is a stochastic method with low-discrepancy sequences (Halton, Sobol, and Faure sequences) and (t, m, s)-nets.
In mathematics, low-discrepancy sequences in which the subsequences of $X_1, \ldots, X_N$ have a small difference with one another for all values of N, the difference of a sequence is small when the proportion of points in the sequence that fall into a given set of B is close to the proportion of the size of the B, which occurs in the case of a normal distribution on average (not for a certain sample).
The conventional randomized quasi-Monte Carlo method [2] is the conversion of the quasi-Monte Carlo method into a variance reduction method with the random accuracy of p_n.

3. The hybrid method

The main objective of Monte Carlo and quasi-Monte Carlo methods are to examine the following integral with the distribution function (x):

$I = \int_{[0.1]^s} f(x) d(x)$

Which is estimated using the following estimator?

$\hat{I} = \frac{1}{N} \sum_{k=1}^{N} f(x^{(k)})$

Where, appropriate $x^{(k)}$s of s-dimensional vectors are chosen, and these vectors are random points distributed uniformly. If the S dimension is large, and a subset of the variable d, such as $i_1, i_2, i_3, \ldots i_d$, can be identified, the hybrid sequence is defined as follows:

It is assumed that $x^k = (q^k, X^k$ is the s-dimensional sequence obtained from interconnected vectors X^k and q^k, as $q^k, k \geq 1$ is a Monte-Carlo s-dimensional sequence, and $X^k, k \geq 1$, are independent random variables with uniform distribution on $(0,1)^{s-d}$. The above defined x^k is a hybrid sequence [7].

3.1. The central limit theorem for the hybrid estimator

Prior to the proof of central limit theorem for hybrid sequences, the following variables are defined:
$Y_k = f\left(q_1^{(k)}, \dots, q_d^{(k)}, X_{d+1}^{(k)}, \dots X_s^{(k)}\right)$
It is assumed:
$\mu_k = E[Y_k]$, $\sigma_k^2 = \mathrm{V}ar(Y_k)$, and $s_N^2 = Var(\theta_m)N^2 = \sigma_1^2 + \cdots + \sigma_N^2$
The central limit theorem:
1) The estimator θ_m has been assumed to be distributed normally.
2) Variance of assumptions is theoretically clear.
3) The estimator has a variance smaller than that in the assumed Monte Carlo method [1].

Theorem 3.1.1: *It is assumed that f is the upper limit of $[0,1]^s$ and then functions* $(x_1, \dots, x_d) = \int_{[0,1]^{s-d}} (f(x_1, \dots x_d, X_{d+1}, \dots, X_s))^2 dX_{d+1} \dots dX_s$
And
$(x_1, \dots, x_d) = \left(\int_{[0,1]^{s-d}} f(x_1, \dots, x_d, X_{d+1}, \dots, X_s) dX_{d+1} \dots dX_s\right)^2$
Is Riemann integrable.
1) The distribution of the normalized set of $\frac{\sum_{K=1}^N Y_k - \sum_{k=1}^N \mu_k}{s_N}$ inclines toward the standard normal distribution.
2) Then,
$s_N^2/N \to L = \int_{[0,1]^s} f(x)^2 dx - \int_{[0,1]^d} \left(\int_{[0,1]^{s-d}} f(x,y) dx\right)^2 dy;$
(1) The hybrid method always results in the reduction of the variance of the standard Monte Carlo method, and the reduction applies by $\frac{\int_{[0,1]^s} f(x)^2 dx - \int_{[0,1]^d} \left(\int_{[0,1]^{s-d}} f(x,y) dx\right)^2 dy}{\int_{[0,1]^s} f(x)^2 dx - \left(\int_{[0,1]^s} f(x) dx\right)^2} \leq 1$
Argument: The variance of Y_k equals:
$\sigma_k^2 = \int_{[0,1]^{s-d}} \left(f(q_1^{(k)}, \dots q_d^{(k)}, X_{d+1}, \dots, X_s)\right)^2 dX_{d+1} \dots dX_s$
$- \left(\int_{[0,1]^{s-d}} f\left(q_1^{(k)}, \dots, q_d^{(k)}, X_{d+1}, \dots, X_d\right) dX_{d+1} \dots dX_s\right)^2$
When f is bounded and Y_n is bounded too and yielded from the same result, the adequacy shows that $s_N \to \infty$ if $N \to \infty$. Lindeberg condition, which guarantees and proves the central limit theorem for independent random variables with different distributions. However, according to the theory of uniform distribution of sequences, when h and g is Riemann integrable:
$\frac{1}{N} \sum_{k=1}^N g\left(q_1^{(k)}, \dots, q_d^{(k)}\right) \to \int_{[0,1]^d} f(x)^2 \mathrm{d}x$
And
$\frac{1}{N} \sum_{k=1}^N h\left(q_1^{(k)}, \dots, q_d^{(k)}\right) \to \int_{[0,1]^d} h(y) dy = \int_{[0,1]^d} \left(\int_{[0,1]^{s-d}} f(x,y) dx\right)^2 dy,$
Proof of the claim 2:

Regarding the later claim, it should be note that $s_N/N \to \int_{[0,1]^s} f(x)^2 dx - \int_{[0,1]^d} \left(\int_{[0,1]^{s-d}} f(x,y) dx\right)^2 dy$ when$N \to \infty$, as $\sigma^2 = \int_{[0,1]^s} f(x)^2 dx - \left(\int_{[0,1]^s} f(x) dx\right)^2$, variance of f(x) for uniform distribution of x is on$(0,1)^s$. In fact, the reduced variance obtained from $\int_{[0,1]^d} \left(\int_{[0,1]^{s-d}} f(x,y) dx\right)^2 dy \geq \left(\int_{[0,1]^d} \int_{[0,1]^{s-d}} f(x,y) dx dy\right)^2$ is taken into account.

4. Confidence interval

It is assumed that $X_1, \dots, X_n$ is an n-tuple random sample of a distribution of the probability density function $\{f_\theta(x) : : \theta \in \Theta$, as$\subset \theta \mathrm{R}$. Furthermore, it is assumed that $\mathcal{L}1, \dots, \mathcal{L}n$ is an interval of real numbers. For example,
$\left(\bar{X} - \frac{s}{\sqrt{n}} t_{1-\alpha/2}(n-1), \bar{X} + \frac{s}{\sqrt{n}} t_{1-\alpha/2}(n-1)\right)$ Is a confidence interval for the mean population?
Definition: The interval $I(X_1, \dots, X_n)$ is a family of confidence intervals [9] at the level of $(1-\alpha)$ for θ if $P_\theta(\theta \in \mathcal{L}1, \dots, \mathcal{L}n) = 1-\alpha$ per $\Theta \ni \theta$.

4.1. The shortest confidence interval

For every Q, axial quantity if a and b are selected in a way that the observed length of confidence interval decreased to minimum, such a length of confidence interval would be the shortest confidence interval on the basis of the Q axial quantity.

Definition: It is assumed that $Q = q(X_1, \dots, X_n; \theta)$ is a function of the random sample and the desired unknown parameter, that is, θ. If the distribution of Q does not depend on θ, Q is determined as an axial quantity.

Numerical results: In this section, the hybrid method is compared with Monte Carlo and quasi-Monte Carlo numerical methods, when these methods are used for pricing stock. The problem in question is the following integral:

$$I = \int_{[0,1]^s} f(x)dx$$

It is assumed that $x^{(k)}$, $k = 1, \dots$ is a sequence of independent random variables with the same distribution as that of $U(0,1)^s$ and $i = d+1, \dots, s$; $x^{(k)}$, $k = 1, \dots$ is a sequence of independent random variables with the same distribution as that of $U(0,1)$; $x^{(k)} = (q_1^{(k)}, \dots, q_d^{(k)}, x_{d+1}^{(k)}, \dots, x_s^{(k)})$ is the kth element of the s-dimensional hybrid sequence with a certain d-dimensional component; $u^{(k,i)}$ is the kth element of the real hybrid sequence, as the certain component is the ith sequence of real random quasi-Monte Carlo, and the remaining (s-d) is the component sampled from the distribution of $U(0,1)^{s-d}$; the estimators are defined as follows:

$\theta = \frac{1}{NM}\sum_{k=1}^{NM} f(X^{(k)}) - Mc,$

$\theta_{mixed} = \frac{1}{NM}\sum_{k=1}^{NM} f(x^{(k)}) - Mixed(padding\ QMC\ \ by\ Mc)$

$\theta_{Rmixed} = \frac{1}{M}\sum_{i=1}^{M}(\frac{1}{N}\sum_{k=1}^{N} f\left(u^{(k,i)}\right)\ - Randomized\ mixed(pandding\ RQmc\ by\ Mc)$

It should be noted that θ_{mixed} is a biased estimator. The efficiency $\varepsilon(\theta)$ of the estimator θ is defined as follows:

$\varepsilon(\theta) = (var(\theta) + (E[\theta - I])^2 t)^{-1}$

t is the time of computation, and $E[\theta - I]$ is the computed bias for the estimator θ_{Rmixed}.

Example 1: (Geometric Asian option pricing):

The price of Geometric Asian options [8] can be computed precisely, while, the arithmetic Asian options cannot be priced precisely by the formula. In simulation, firstly, a sequence of the asset price $s_0, s_1, \dots\dots, s_k$ is generated and is a subject of Ito process with ds $= \mu s dt + \sigma s dx$. The interest function is defined as follows:

h(So,S1,...,Sk)=max(G(S0,S1,…SK)-F(0))

Where, G (S0, S1…SK)= $(\pi_{i=0}^{k} S_i)^{1/(k+1)}$ is the geometric mean of the asset price, and F is the spot price. The price of this option is the expected value of $E[e^{-rT}h(S_0, S_1, \dots, S_K)]$ that have been estimated using simulation techniques. In this formula, r is the risk-free interest rate, and T is the expiration time when the final price of SK is observed. The price of the option is estimated using Monte Carlo, quasi-Monte Carlo, and randomized hybrid methods. In this problem, k corresponds to the dimension of the problem. In this numerical example, the dimensions of certain parts of the hybrid sequence d are 32 and 64, and other constants are as follows:

S0=500,$\sigma = 0.1$ r = μ = 0.1, K = 256, F = 5, T = 128

Firstly, S0 and SK, and then, $S_{k/2}, S_{k/4}, S_{3k/4}, S_{k/8}, S_{3k/8}, S_{5k/8}, S_{7k/8}$ are calculated.

5. Analysis

Figures 1, 2, 3, and 4 respectively show the time of computations, bias, efficiency, and width of the confidence interval. The efficiency figure shows the excessive effects of the methods. In general, the randomized hybrid method has the most favorable efficiency than the Monte Carlo method. According to the central limit theorem, the variance of the hybrid estimator is low, and consequently, the efficiency of this method is high as shown in Figure 3. As the variance and confidence interval have a direct relation, the confidence interval becomes shorter. According to Figure 4, this is the shortest confidence interval and consequently, an optimum confidence interval for the hybrid method.

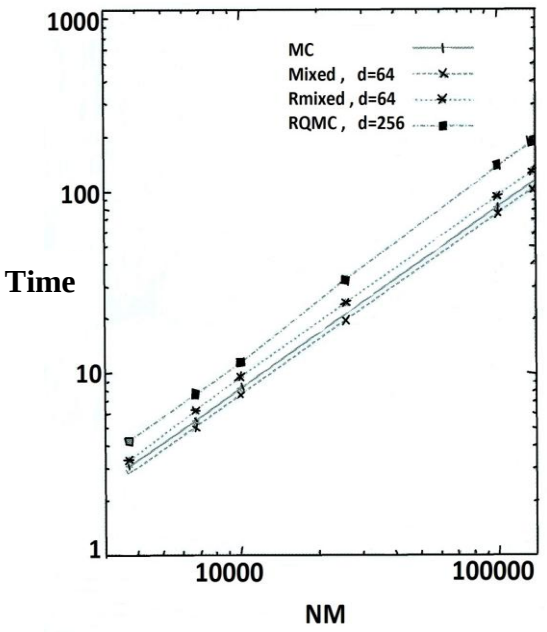


Fig. 1: The Time of Computations for the Example of Geometric Asian Option Pricing

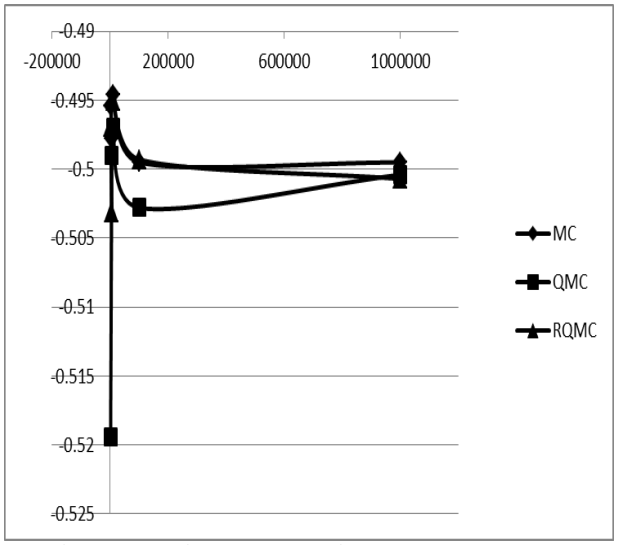


Fig. 2: The Bias of the Methods for the Example of Geometric Asian Option Pricing efficiency

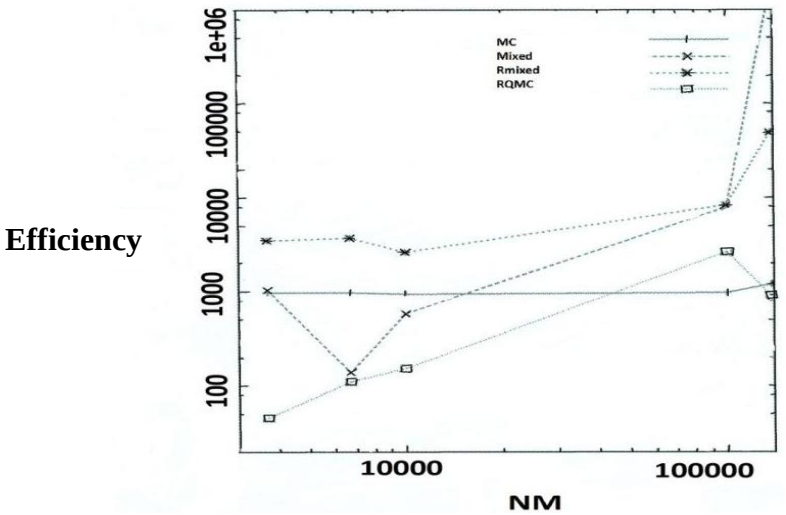


Fig. 3: Methods for the Example of Geometric Asian Option Pricing

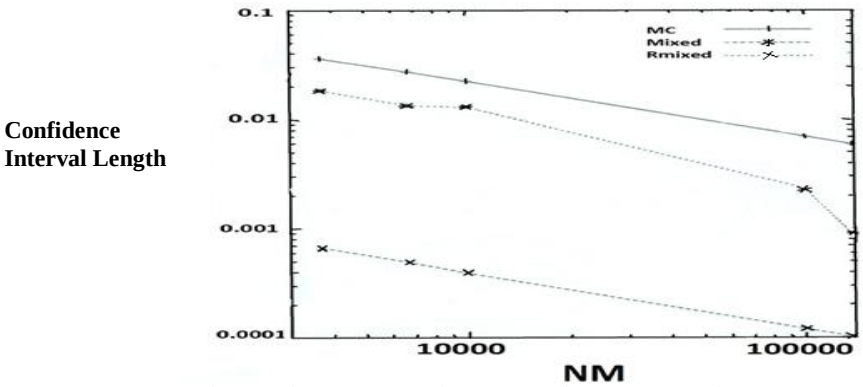


Fig. 4: The Width of the Confidence Interval of the Methods for the Example of Geometric Asian Option Pricing

Example 2: The European call option allows the owner to buy the asset at the given time T (expiration or effective date) for the determined price K (effective price). Therefore, the efficiency of European call is
Max $(0, S(T) - K)$.
The Asian option is the option that the efficiency depending on the mean asset.
The efficiency for mean purchase price with the execution price K is the $(\bar{S}(T) - K)^+$ in which $\bar{S}(T)$ is the mean price in$[0, T]$.
If S (t) indicates the price at the time t, a widely used model in financial mathematics is as follows:
$\frac{dS}{S} = \mu dt + \sigma dB$
Where, dB is the standard Brownian motion, μ is the annual expected growth rate, and σ denotes changes (standard deviation of return). Using the Ito lemma, the above stochastic differential equation is solved as follows:
$S(t) = S_0 \left(\left(\mu - \frac{1}{2}\sigma^2 \right) t + \sigma dB \right)$
Or
$S(t) = S_0 \left(\left(\mu - \frac{1}{2}\sigma^2 \right) t + \sigma\sqrt{t}.z \right)$
Therefore, the model follows the geometric Brownian motion, and the final price is obtained within the framework of Black-Scholes model under the risk-neutral measure as follows:
$C(S_t, \mathrm{t}) = \mathrm{e}^{-\mathrm{r}(\mathrm{T}-\mathrm{t})}\mathrm{E}_\mathrm{t}[\mathrm{payoff}(S_\mathrm{T})]$
Thus, for the European call option:
$\mathrm{C} = \mathrm{e}^{-\mathrm{r}(\mathrm{T}-\mathrm{t})}\,\mathrm{E}_\emptyset \left(\mathrm{S}(0)\mathrm{e}^{\left(\mathrm{r}-\sigma^2/2\right)\mathrm{T}+\mathrm{Z}\sigma\sqrt{\mathrm{T}}} - \mathrm{K} \right)^+$
For the Asian purchase option:
$\mathrm{C} = \mathrm{e}^{-\mathrm{rT}}\mathrm{E}_{\mathrm{Z}\sim\mathrm{N}(0,1)} \left[\left(\frac{1}{\mathrm{n}}\sum_{\mathrm{j}=1}^{\mathrm{n}} \mathrm{X}_\mathrm{j} \right) - \mathrm{K} \right]^+$
We assume the evaluation of option pricing through Monte Carlo, quasi-Monte Carlo, and hybrid methods in practice:
$\mathrm{S}_0 = 100\,, \mathrm{K} = 100\,, \mathrm{T} = 1\,, \mathrm{r} = 5\%\,, \sigma = 0.10\,, \mathrm{N} = 1000$
The following tables show the results of the evaluation of European and Asian call option through different methods and compare them with the precise answer (meanwhile, the Sobol sequences have been used for generation of uniform samples in QMC method).

Table 1: The Table for the Evaluation of European Option through Different Methods

	Price	Standard deviation	Time of computation
Answer of Black-Scholes model	5.4713	-	-
Answer of MC method	5.6539	0.0334	6
Answer of QMC method	5.3367	0.0246	8
Answer of hybrid method	5.4485	0.0042	5

An Asian option with the following parameters is assumed to be evaluated:

Table 2: The Table for the Evaluation of Asian Option through Different Methods, N = 1000,$S_0 = 100$, K = 100 , T = 1 , R = 5% , Σ = 0.10

	Price	Standard deviation	Time of computation
Answer of Black-Scholes model	3.8352	-	-
Answer of MC method	3.3182	0.1348	7
Answer of QMC method	3.1675	0.1295	6
Answer of hybrid method	3.1778	0.1291	7

6. Conclusion

In this study, the efficiency of the hybrid method was compared with that of MC and QMC methods in financial problems of high dimensions. Having used the central limit theorem, the optimum confidence intervals for the integral were obtained. The numerical results were discussed finally when the hybrid method and its random versions were used for problems of pricing models. The numerical analysis shows that the hybrid method (combination of Monte Carlo and quasi-Monte Carlo) can considerably improve the efficiency in problems of high dimensions for the sample size, especially large sample size. Moreover, the strategy of hybrid method is perfect in terms of efficiency. This occurs when the variance of the hybrid method is reduced according to the central limit theorem, and this reduction makes the confidence interval shorter and improves the efficiency significantly.

References

[1] Giray Okten, Bruno Tuffin, Vadim Burago, 22 March 2006, A central limit theorem and improved error bounds for a hybrid-Monte Carlo sequence with applications in computational finance, Department of Mathematics, Florida State University, Tallahassee, FL 32306 USA.

[2] Pierre L'Ecuyer, 23 July 2009, Quasi-Monte Carlo methods with applications in finance, this article is published with open access at Springerlink.com.

[3] Peter Winker, Jenny X. Li, September 27, 2000, Time Series Simulation with Quasi Monte Carlo Methods, Departments of Mathematics and Economics, The Pennsylvania State University, University Park, PA, 16802, USA.

[4] R. Caflisch, M. Morokoff, A. Owen, Valuation of mortgage backed securities using brownian bridges to reduce effective dimension, J. Comput. Finance 1 (1) (1997) 27–46.

[5] Kahn, H., Marshall, A.W. (1953) Methods of reducing sample size in Monte Carlo computations, J.Oper.Res.Soc.Amer. 1, 263–271. http://dx.doi.org/10.1287/opre.1.5.263.

[6] R. E. Caflisch, Monte Carlo and quasi-Monte Carlo methods, Acta Numerica vol. 7, Cambridge University Press, 1998, pp. 49-1.

[7] tJ. Spanier, Quasi-Monte carlo methods for particle transport problems, in: H. Niederreiter, P.J.-S. Shiue (Eds.), Monte Carlo and Quasi-Monte Carlo Methods in Scientific Computing, Lecture Notes in Statistics, vol. 106, Springer, Berlin,1995, pp. 121–148. http://dx.doi.org/10.1007/978-1-4612-2552-2_6.

[8] trL. Clewlow, C. Strickland, Implementing Derivatives Models,Wiley, NewYork

[9] Parsiyan, A. Fundamentals of mathematic statistics. Tehran, Tehran University.