Applications of Monte Carlo Methods on Financial Engineering

Chuan-Hsiang Han

IMA, University of Minnesota
and
Ford Motor Company

December 20 & 27, 2004

Outline

- I. Basic Monte Carlo Methods
  1. Revision of probability theory
  2. Methods to reduce the variance
  3. Applications in computational finance
- II. Monte Carlo Methods for Diffusion Equations
  1. Simulation of diffusion processes
  2. Variance reduction under Black-Scholes
  3. Ergodicity and stochastic volatility models
- III. Methods for Variance Reduction (Next Monday)
- IV. Challenges from Financial Engineering (Next Monday)

Main References

1. J.-P. Fouque, G. Papanicolaou, R. Sircar, Derivatives in Financial Markets with Stochastic Volatility, Cambridge University Press, 2000.

2. Paul Glasserman, Monte Carlo Methods in Financial Engineering, Springer Verlag, 2003.

3. B. Lapeyre, E. Pardoux, R. Sentis, Introduction to Monte Carlo Methods for Transport and Diffusion Equations, Oxford University Press, 2003.

Revision of probability theory

Theorem 1 *The strong law of large numbers Let* $(X_i, i \geq 1)$ *be a sequence of independent random variables following the same distribution as a random variable* X*. We assume that* $\mathbf{E}(|X|) < +\infty$*. Then, for almost every* ω*:*

$$\mathbf{E}(X) = \lim_{n \to \infty} \frac{1}{n}(X_1(\omega) + X_2(\omega) + \cdots + X_n(\omega)).$$

Notations:

1. Sample mean: $\frac{1}{n}(X_1 + X_2 + \cdots + X_n)$
2. Error: $\epsilon_n = \mathbf{E}(X) - \frac{1}{n}(X_1 + X_2 + \cdots + X_n)$.

In finance, we are interested in calculating the quantity

$$\mathbf{E}\{h(X)\},$$

where h is called the payoff function.

Theorem 2 *The central limit theorem Let* $(X_i, i \geq 1)$ *be a sequence of independent random variables following the same distribution as a random variable* X*. We assume that* $\mathbf{E}(X^2) < +\infty$*. We denote by* σ^2 *the variance of* X*, then*

$$\frac{\sqrt{n}}{\sigma}\epsilon_n \textit{ converges in probability to } G,$$

G *being a random variable with a reduced centred Gaussian distribution.*

We often describe the error of Monte Carlo method by giving the standard deviation of ϵ_n, that is, $\sigma/\sqrt{n}$.

Confidence interval

Suppose X is a $\mathcal{R}^d$-valued random variable with distribution $f(x)dx$, we want to evaluate

$$I = \int_{\mathcal{R}^d} x f(x) dx = \mathbf{E}(X).$$

Let X_i be independent realizations under the distribution of X, we can approximate I by sample mean $\bar{I}_N = \frac{1}{N}\sum_{i=1}^{N} X_i$.
Define sample or empirical standard deviation $\bar{\sigma}$ by

$$\bar{\sigma} = \sqrt{\frac{1}{N-1}\sum_{i=1}^{N}(X_i - \bar{I}_N)^2}.$$

A 95% confidence interval for I is $\left[\bar{I}_N - 2\frac{\bar{\sigma}}{\sqrt{N}}, \bar{I}_N + 2\frac{\bar{\sigma}}{\sqrt{N}}\right]$.

A remark on convergence rate for integration

To calculate $\int_{[0,1]^d} f(x)dx$ by simple trapezoidal rule, the convergence rate is $O(n^{-2/d})$, at least for twice continuously differentiable f.
To calculate $\int_{[0,1]^d} f(x)dx$ by the sample mean of $\mathbf{E}\{f(U)\}$, with U uniformly distributed between 0 and 1. The convergence rate is $O(n^{-1/2})$ as long as f is integrable.

So, Monte Carlo methods are attractive in evaluating integrals in high dimensions.
And, we focus on reduce the sample variance $\bar{\sigma}^2$. (In contrast to Quasi-Monte-Carlo methods)

Example 1

Evaluate a call option

$$C = \mathbf{E}\left\{\max\left(e^{\beta G} - K, 0\right)\right\}$$

with $\beta = 1.0$ and $K = 1.0$.

N	estimated value	95% confidence interval
100	0.8947	[0.5933, 1.1960]
1,000	0.8724	[0.7281, 1.0168]
10,000	0.8758	[0.8370, 0.9546]
100,000	0.8939	[0.8813, 0.9065]

The exact value is 0.8871 based on the Black-Scholes formula

$$C = e^{\beta^2/2}\mathcal{N}\left(\beta - \frac{\log K}{\beta}\right) - K\mathcal{N}\left(-\frac{\log K}{\beta}\right)$$

Example 1: Cont.

Evaluate a put option

$$P = \mathbf{E}\left\{\max\left(K - e^{\beta G}, 0\right)\right\}$$

with $\beta = 1.0$ and $K = 1.0$.

N	estimated value	95% confidence interval
100	0.2079	[0.1512, 0.2646]
1,000	0.2419	[0.2230, 0.2608]
10,000	0.2343	[0.2284, 0.2402]
100,000	0.2391	[0.2372, 0.2410]

The exact value is 0.2384 based on the Black-Scholes formula

$$P = K\mathcal{N}\left(\frac{\log K}{\beta}\right) - e^{\beta^2/2}\mathcal{N}\left(\frac{\log K}{\beta} - \beta\right)$$

Methods to reduce variance

Aim: to reduce the sample variance $\bar{\sigma}^2$ in order to reduce the rate of convergence of Monte Carlo methods $\bar{\sigma}/\sqrt{n}$.

Common reduction of variance techniques include:

1. Importance sampling: $\mathbf{E}(X) = \mathbf{E}(Y)$
2. Control Variates (or variables): $\mathbf{E}(f(X)) = \mathbf{E}(f(X) - \lambda h(X)) + \lambda \mathbf{E}(h(X))$, where $\mathbf{E}(h(X))$ has an explicit solution.
3. Antithetic variables: $\int_0^1 f(x)dx = \frac{1}{2}\int_0^1 (f(x) + f(1-x))\, dx$
4. Stratification: $\mathbf{E}(g(X)) = \sum_{i=1}^m \mathbf{E}\left(1_{\{X\in D_i\}} g(X)\right)$

Importance Sampling

Let X be a $\mathcal{R}$-valued random variable with distribution $f(x)dx$.

$$\begin{aligned} \mathbf{E}(g(X)) &= \int_{\mathbf{R}} g(x)f(x)dx \\ &= \int_{\mathbf{R}} \frac{g(x)f(x)}{\tilde{f}(x)}\tilde{f}(x)dx = \mathbf{E}\left(\frac{g(Y)f(Y)}{\tilde{f}(Y)}\right), \end{aligned}$$

where $\tilde{f} > 0$ is a density function and Y follows the distribution $\tilde{f}(x)dx$.

Can we always find

$$Var\left(\frac{g(Y)f(Y)}{\tilde{f}(Y)}\right) < Var(g(X))?$$

Importance Sampling: Cont.

Ideally, the variance of $\frac{g(Y)f(Y)}{\tilde{f}(Y)}$ can be diminished by choosing

$$\tilde{f}(x) = (g(x)f(x))/(\mathbf{E}(g(X)))$$

because

$$Var\left(\frac{g(Y)f(Y)}{\tilde{f}(Y)}\right) = \int_{\mathcal{R}} \frac{g^2(x)f^2(x)}{\tilde{f}^2(x)}dx - \mathbf{E}(g(X))^2.$$

But $\mathbf{E}(g(X))$ is unknown and is exactly what we want to compute. Any approximation of this quantity might be helpful.

Example 2

Evaluate a put option

$$P = \mathbf{E}\left\{\left(1 - e^{\beta G}\right)^+\right\} = \int_{\mathcal{R}} \frac{(1-e^{\beta x})^+}{\beta|x|}\beta|x|e^{-x^2/2}\frac{dx}{\sqrt{2\pi}}$$

Change of variables $x = \sqrt{y}$ over $\mathcal{R}^+$ and $x = -\sqrt{y}$ over $\mathcal{R}^-$ we deduce

$$\begin{aligned} P &= \int_0^{+\infty} \frac{(1-e^{\beta\sqrt{y}})^+ + (1-e^{-\beta\sqrt{y}})^+}{\sqrt{2\pi}\sqrt{y}} e^{-y/2}\frac{dy}{2} \\ &= \mathbf{E}\left(\frac{(1-e^{\beta\sqrt{Y}})^+ + (1-e^{-\beta\sqrt{Y}})^+}{\sqrt{2\pi}\sqrt{Y}}\right), \end{aligned}$$

where Y is exponentially distributed with parameter $1/2$.

Example 2: Cont.

N	estimated value	95% confidence interval
100 (IS)	0.249	[0.239, 0.260]
100	0.208	[0.151, 0.265]
1,000 (IS)	0.239	[0.235,0.243]
1,000	0.242	[0.223, 0.261]
10,000(IS)	0.238	[0.237, 0.239]
10,000	0.234	[0.228, 0.240]

The exact value is 0.23842 based on the Black-Scholes formula

Control Variates

$\mathbf{E}(f(X)) = \mathbf{E}(f(X) - \lambda h(X)) + \lambda \mathbf{E}(h(X))$, where $\mathbf{E}(h(X))$ has an explicit solution and λ is a constant as a control parameter.

$$Var\left[(f(X) - \lambda h(X)) + \lambda \mathbf{E}(h(X))\right] = \sigma_f^2 - 2\lambda\sigma_f\sigma_h\rho_{fh} + \lambda^2\sigma_f^2$$
$$< Var[f(X)],$$

provided $\lambda^2\sigma_f < 2\lambda\sigma_h\rho_{fh}$.
The optimal $\lambda^* = \frac{\sigma_f}{\sigma_h}\rho_{fh} = \frac{COV[f(X),h(X)]}{Var(f(X))}$ results in

$$\frac{Var\left[(f(X) - \lambda^* h(X)) + \lambda^* \mathbf{E}(h(X))\right]}{Var[f(X)]} = 1 - \rho_{fh}^2.$$

Example 3

Evaluate a call option $C = \mathbf{E}\left\{\max\left(e^{\beta G} - K, 0\right)\right\}$.

By Call-Put Parity:

$$C - P = \mathbf{E}\left(e^{\beta G} - K\right) = e^{\beta^2/2} - K,$$

we define a control variate estimator with $\lambda = -1$

$$C = \left[C + \left(e^{\beta G} - K\right)\right] - \mathbf{E}\left(e^{\beta G} - K\right) = P + e^{\beta^2/2} - K$$

We already observed in Example 1 that variance of P is much less than variance of C.

Example 3: Cont.

N	estimated value	95% confidence interval
100 (CV)	0.8566	[0.7999, 0.9133]
100	0.8947	[0.5933, 1.1960]
1,000 (CV)	0.8906	[0.8717, 0.9095]
1,000	0.8724	[0.7281, 1.0168]
10,000(CV)	0.8830	[0.8771, 0.8889]
10,000	0.8758	[0.8370, 0.9546]
100,000(CV)	0.8878	[0.8859, 0.8897]
100,000	0.8939	[0.8813, 0.9065]

The exact value is 0.8871.

Antithetic variables

Assume we want to calculate $\int_0^1 f(x)dx = \mathbf{E}(f(U))$.
By symmetry

$$\int_0^1 f(x)dx = \frac{1}{2}\int_0^1 (f(x) + f(1-x))\, dx$$

we double the size of trials of sample mean

$$\begin{aligned} I_{2n} &= \frac{1}{n}\left(\frac{1}{2}(f(U_1) + f(1-U_1)) + \cdots + \frac{1}{2}(f(U_n) + f(1-U_n))\right) \\ &= \frac{1}{2\,n}\left(f(U_1) + f(1-U_1) + \cdots + f(U_n) + f(1-U_n)\right) \end{aligned}$$

This reduces sample variance at most by 2.

Stratification

Assume we wan to calculate $I = \mathbf{E}(g(X))$. Decompose

$$I = \sum_{i=1}^{m} \mathbf{E}\left(1_{\{X \in D_i\}} g(X)\right) = \sum_{i=1}^{m} \mathbf{E}\left(g(X) | X \in D_i\right) \mathbf{P}(X \in D_i)$$

and approximate I by

$$\tilde{I} = \sum_{i=1}^{m} \tilde{I}_i p_i,$$

where $\tilde{I}_i$ is a sample mean of $I_i = \mathbf{E}\left(g(X) | X \in D_i\right)$ with n_i independent trials and $p_i = \mathbf{P}(X \in D_i)$.
Since samples I_i are assumed independent, the variance of $\tilde{I}$ becomes

$$Var(\tilde{I}) = \sum_{i=1}^{m} p_i^2 \frac{\sigma_i^2}{n_i}.$$

Stratification: Cont

Under the constraint of total trials $n = \sum_{i=1}^{m} n_i$, each n_i is optimally chosen as

$$n_i = n \frac{p_i \sigma_i}{\sum_{i=1}^{m} p_i \sigma_i}.$$

Hence, the minimum of the variance of $\tilde{I}$ becomes

$$\frac{1}{n} \left(\sum_{i=1}^{m} p_i \sigma_i \right)^2,$$

which is less than $Var(\tilde{I}) = \sum_{i=1}^{m} p_i^2 \frac{\sigma_i^2}{n_i}$.