Lecture 4

Monte Carlo Simulation

Lecture Notes

by Jan Palczewski

with additions

by Andrzej Palczewski

We know from Discrete Time Finance that one can compute a fair price for an option by taking an expectation

$$\mathbb{E}_{\mathbb{Q}}\left(e^{-rT}X\right)$$

Therefore, it is important to have algorithms to compute the expectation of a random variable with a given distribution. In general, we cannot do it precisely – there is only an approximation on our disposal.

In previous lectures we studied methods to generate independent realizations of a random variable. We can use this knowledge to generate a sample from a distribution of X.

Main theoretical ingredients that allow to compute approximations of expectation from a sample of a given distribution are provided by the Law of Large Numbers and the Central Limit Theorem.

Let $X_1, X_2, \ldots$ be a sequence of independent identically distributed random variables with expected value μ and variance σ^2. Define the sequence of averages

$$Y_n = \frac{X_1 + X_2 + \cdots + X_n}{n}, \qquad n = 1, 2, \ldots.$$

(Law of Large Numbers) Y_n converges to μ almost surely as $n \to \infty$.

Let

$$Z_n = \frac{(X_1 - \mu) + (X_2 - \mu) + \cdots + (X_n - \mu)}{\sqrt{n}}.$$

(Central Limit Theorem) The distribution of Z_n converges to $N(0, \sigma^2)$.

Assume we have a random variable X with unknown expectation and variance

$$a = \mathbb{E}X, \qquad b^2 = \mathsf{Var}X.$$

We are interested in computing an approximation to a and possibly b.

Suppose we are able to take independent samples of X using a pseudo-random number generator.

We know from the strong law of large number, that the average of a large number of samples can give a good approximation to the expected value.

Therefore, if we let $X_1, X_2, \ldots, X_M$ denote a sample from a distribution of X, one might expect that

$$a_M = \frac{1}{M} \sum_{i=1}^{M} X_i$$

is a good approximation to a.

To estimate variance we use the following formula (notice that we divide by $M-1$ and not by M)

$$b_M^2 = \frac{\sum_{i=1}^{M} (X_i - a_M)^2}{M-1}.$$

Assessment of the precision of estimation of a and b^2

By the central limit theorem we know that $\sum_{i=1}^{M} X_i$ behaves approximately like a $\mathcal{N}(Ma, Mb^2)$ distributed random variable.
Therefore

$$a_M - a \text{ is approximately } \mathcal{N}(0, \frac{b^2}{M}).$$

In other words

$$a_M - a \sim \frac{b}{\sqrt{M}} Z,$$

where $Z \sim \mathcal{N}(0, 1)$.
Therefore a_M converges to a with speed $\mathcal{O}\left(\frac{b}{\sqrt{M}}\right)$.

More quantitative

$$\mathbb{P}\left(a - \frac{1.96b}{\sqrt{M}} \leq a_M \leq a + \frac{1.96b}{\sqrt{M}}\right) = 0.95.$$

This is equivalent to

$$\mathbb{P}\left(a_M - \frac{1.96b}{\sqrt{M}} \leq a \leq a_M + \frac{1.96b}{\sqrt{M}}\right) = 0.95.$$

Replacing the unknown b by the approximation b_M we see that the unknown expected value a lies in the interval

$$\left[a_M - \frac{1.96\, b_M}{\sqrt{M}}, a_M + \frac{1.96\, b_M}{\sqrt{M}}\right]$$

approximately with probability 0.95.

The interval above is called a 95 percent confidence interval.

Confidence intervals explained

If $Z \sim N(\mu, \sigma^2)$ then

$$\mathbb{P}(\mu - 1.96\sigma < Z < \mu + 1.96\sigma) = 0.95$$

because

$$\Phi(1.96) = 0.975$$

To construct a confidence interval with a confidence level α different from 0.95, we have to find a number A such that

$$\phi(A) = 1 - \frac{1-\alpha}{2}.$$

Then the α-confidence interval

$$\mathbb{P}(\mu - A\sigma < Z < \mu + A\sigma) = \alpha.$$

- The width of the confidence interval is a measure of the accuracy of our estimate.
- In 95% of cases the true value lies in the confidence interval.
- Beware! In 5% of cases it is outside the interval!
- The width of the confidence interval depends on two factors:
  - Number of simulations M
  - Variance of the variable X

$$\left[a_M - \frac{1.96\, b_M}{\sqrt{M}}, a_M + \frac{1.96\, b_M}{\sqrt{M}}\right]$$

1. The size of the confidence interval shrinks like the inverse square root of the number of samples. This is one of the main disadvantages of Monte Carlo method.
2. The size of the confidence interval is directly proportional to the standard deviation of X. This indicates that Monte Carlo method works the better, the smaller the variance of X is. This leads to the idea of variance reduction, which we shall discuss later.

Monte Carlo in a nutshell

To compute

$$a = \mathbb{E}X$$

we generate M independent samples for X and compute a_M.

In order to monitor the error we also approximate the variance by b_M^2.

$$\left[a_M - \frac{1.96\, b_M}{\sqrt{M}}, a_M + \frac{1.96\, b_M}{\sqrt{M}}\right]$$

Important! People often use confidence intervals with other confidence levels, e.g. 99% or even 99.9%. Then instead of 1.96 we have a different number.

Monte Carlo put into action

We can now apply Monte Carlo simulation for the computation of option prices.

We consider a European-style option $\psi(S_T)$ with the payoff function ψ depending on the terminal stock price.

We assume that under a risk-neutral measure the stock price S_t at $t \geq 0$ is given by

$$S_t = S_0 \exp\left((r - \frac{1}{2}\sigma^2)t + \sigma W_t\right).$$

Here W_t is a Brownian motion.

We know that $W_t \sim \sqrt{t}Z$ with $Z \sim \mathcal{N}(0,1)$.

We can therefore write the stock price at expiry T as

$$S_T = S_0 \exp\left((r - \frac{1}{2}\sigma^2)T + \sigma\sqrt{T}Z \right).$$

We can compute a fair price for $\psi(S_T)$ taking the discounted expectation

$$\mathbb{E}\left[e^{-rT}\psi\left(S_0 e^{(r-\frac{1}{2}\sigma^2)T+\sigma\sqrt{T}Z} \right) \right]$$

This expectation can be computed via Monte Carlo method with the following algorithm

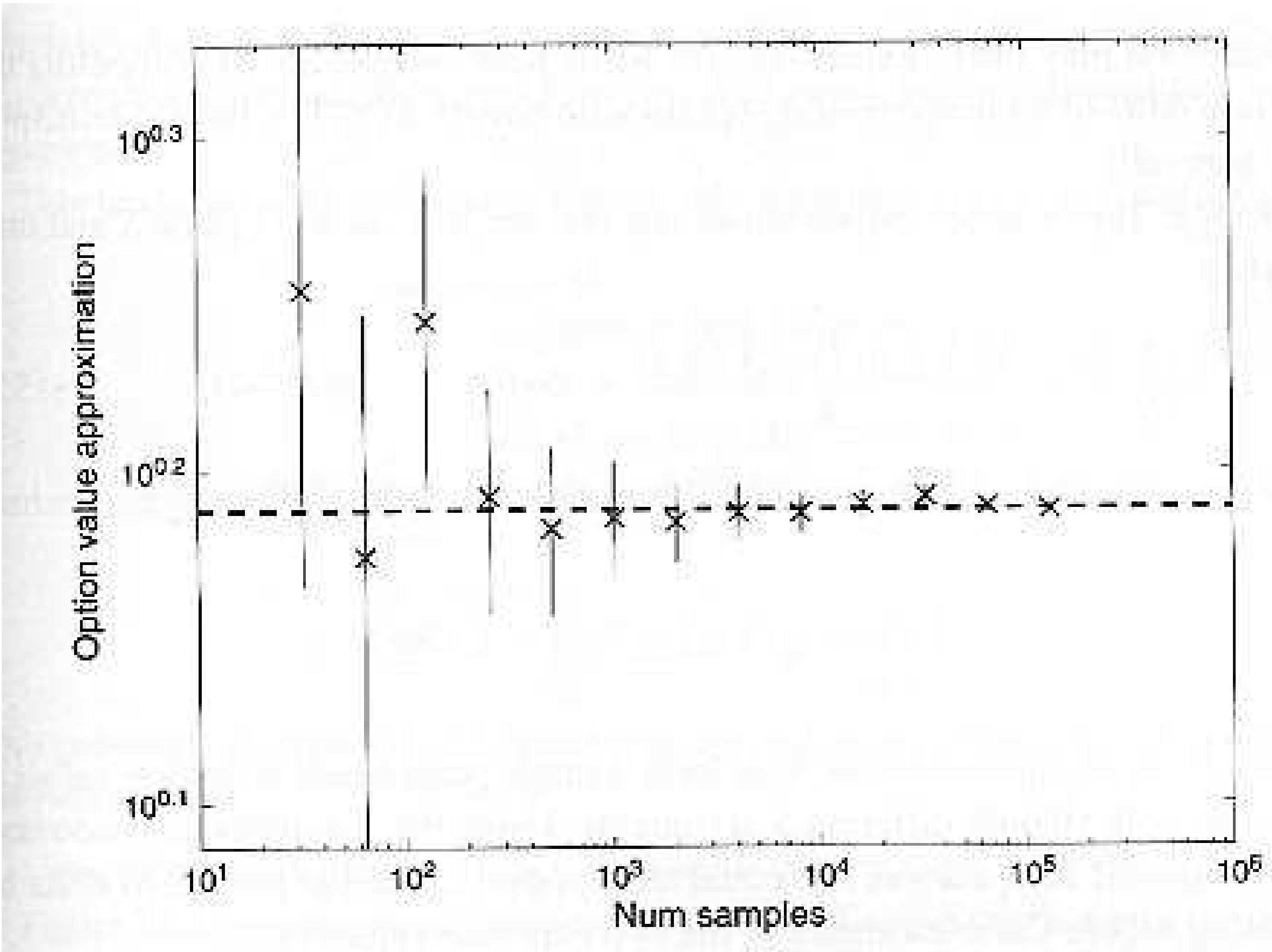
for $i = 1$ to M
    compute an $\mathcal{N}(0,1)$ sample ξ_i
    set $S_i = S_0 \exp\left(\left(r - \frac{1}{2}\sigma^2\right)T + \sigma\sqrt{T}\xi_i\right)$
    set $V_i = e^{-rT}\psi(S_i)$
end
set $a_M = \frac{1}{M}\sum_{i=1}^{M} V_i$

The output a_M provides an approximation of the option price. To assess the quality of our computation, we compute an approximation to the variance

$$b_M^2 = \frac{1}{M-1}\sum_i (V_i - a_M)^2$$

to construct a confidence interval.

Confidence intervals v. sample size



Variance Reduction Techniques

- Antithetic Variates
- Control Variate
- Importance Sampling (not discussed here)
- Stratified Sampling (not discussed here)

We have seen before that the size of the confidence interval is determined by the value of

$$\frac{\sqrt{\mathrm{Var}(X)}}{\sqrt{M}},$$

where M is the sample size.

We would like to find a method to decrease the width of the confidence interval by other means than increasing the sample size M.

A simple idea is to replace X with a random variable Y which has the same expectation, but a lower variance.

In this case we can compute the expectation of X via Monte Carlo method using Y instead of X.

Since the variance of Y is lower than the variance of X the results are better.

The question is how we can find such a random variable Y with a smaller variance that X and the same expectation.

There are two main methods to find Y:

- a method of antithetic variates
- a method of control variates

Antithetic Variates

The idea is as follows. Next to X consider a random variable Z which has the same expectation and variance as X but is negative correlated with X, i.e.

$$\text{cov}(X, Z) \leq 0.$$

Now take as Y the random variable

$$Y = \frac{X + Z}{2}.$$

Obviously $\mathbb{E}(Y) = \mathbb{E}(X)$. On the other hand

$$\begin{aligned} \text{Var}(Y) &= \text{cov}\left(\frac{X+Z}{2}, \frac{X+Z}{2}\right) \\ &= \frac{1}{4}\Big(\text{Var}(X) + \underbrace{2\,\text{cov}(X,Z)}_{\leq 0} + \underbrace{\text{Var}(Z)}_{=\text{Var}(X)}\Big) \leq \frac{1}{2}\text{Var}(X). \end{aligned}$$

With this we can reduce the variance by a factor of 2.

On the way to find Z...

In the extreme case, if we would know that $\mathbb{E}(X) = 0$, we could just take $Z = -X$. Then $Y = 0$ would give the right result, even deterministically.

The equality above would then trivially hold

$$\text{cov}(X, -X) = -\text{Var}(X).$$

But in general we don't know the expectation.

The expectation is what we want to compute, so this naive idea is not applicable.

But it puts us on the right track!

The following Lemma is a step further to identify a suitable candidate for Z and hence for Y.

Lemma. *Let X be an arbitrary random variable and f a monotonically increasing or monotonically decreasing function, then*

$$\textit{cov}(f(X), f(-X)) \leq 0.$$

Let us now consider the case of a random variable which is of the form $f(U)$, where U is a standard normal distributed random variable, i.e. $U \sim \mathcal{N}(0, 1)$.

The standard normal distribution is symmetric, and hence also $-U \sim \mathcal{N}(0, 1)$.

It then follows obviously that $f(U) \sim f(-U)$.

In particular they have the same expectation !

Therefore, in order to compute the expectation of $X = f(U)$, we can take $Z = f(-U)$ and define

$$Y = \frac{f(U) + f(-U)}{2}.$$

If we now assume that the map f is monotonically increasing, then we conclude from the previous Lemma that

$$\text{cov}\big(f(U), f(-U)\big) \leq 0$$

and we finally obtain

$$\mathbb{E}\left(\frac{f(U) + f(-U)}{2}\right) = \mathbb{E}(f(U)),$$

$$\text{Var}\left(\frac{f(U) + f(-U)}{2}\right) < \frac{1}{2}\text{Var}(f(U)).$$

Implementation

for $i = 1$ to M
  compute an $\mathcal{N}(0,1)$ sample ξ_i
  set $S_i^+ = S_0 \exp\left(\left(r - \frac{1}{2}\sigma^2\right)T + \sigma\sqrt{T}\xi_i\right)$
  set $S_i^- = S_0 \exp\left(\left(r - \frac{1}{2}\sigma^2\right)T - \sigma\sqrt{T}\xi_i\right)$
  set $V_i^+ = e^{-rT}\psi(S_i^+)$
  set $V_i^- = e^{-rT}\psi(S_i^-)$
  set $V_i = (V_i^+ + V_i^-)/2$
end
set $a_M = \frac{1}{M}\sum_{i=1}^M V_i$

The output a_M provides an approximation of the option price.

Example: European put: $(K - S_T)^+$ with $S_0 = 4,\ K = 5,\ \sigma = 0.3,\ r = 0.04,\ T = 1.$

Plain Monte Carlo

[1]	"Mean"	"1.02421105149"	
[1]	"Variance"	"0.700745604547"	
[1]	"Standard deviation"	"0.837105491887"	
[1]	"Confidence interval"	"1.00780378385"	"1.04061831913"

Antithetic Variates

[1]	"Mean"	"1.01995595996"	
[1]	"Variance"	"0.019493094166"	
[1]	"Standard deviation"	"0.139617671396"	
[1]	"Confidence interval"	"1.01721945360"	"1.02269246632"

Control Variates

Given that we wish to estimate $\mathbb{E}(X)$, suppose that we can somehow find another random variable Y which is close to X in some sense and has known expectation $\mathbb{E}(Y)$.

Then the random variable Z defined by

$$Z = X + \mathbb{E}(Y) - Y$$

obviously satisfies

$$\mathbb{E}(Z) = \mathbb{E}(X).$$

We can therefore obtain the desired value $\mathbb{E}(X)$ by running Monte Carlo simulation on Z.

In this context Y is called the control variate.

Since adding a constant to a random variable does not change its variance, we see that

$$\text{Var}(Z) = \text{Var}(X - Y).$$

Therefore, in order to get some benefit from this approach we would like $X - Y$ to have a small variance.

This is what we mean by ”close in some sense” from the previous slide.

There is in general no clear candidate for a control variate, this depends on the particular problem. Intuition is needed !

Fine-tuning the Control Variate

Given that we have a candidate Y for a control variate, we can define for any $\theta \in \mathbb{R}$

$$Z_\theta = X + \theta(\mathbb{E}(Y) - Y).$$

We still have $\mathbb{E}(Z) = \mathbb{E}(X)$, so we may apply Monte Carlo to Z_θ in order to compute $\mathbb{E}(X)$.

We have

$$\text{Var}(Z_\theta) = \text{Var}(X - \theta Y) = \text{Var}(X) - 2\theta\,\text{cov}(X, Y) + \theta^2 \text{Var}(Y)$$

We can consider this as a function of θ and look for the minimizer.

It is easy to see that the minimizer is given by

$$\theta_{min} = \frac{\mathrm{cov}(X,Y)}{\mathrm{Var}(Y)}$$

One can furthermore show that $\mathrm{Var}(Z_\theta) < \mathrm{Var}(X)$ if and only if $\theta \in (0, 2\,\theta_{min})$.

In general, however, the expression $\mathrm{cov}(X,Y)$ which is used to compute θ_{min} is not known.

The idea is to run Monte Carlo, where in a first step $\mathrm{cov}(X,Y)$ is computed via Monte Carlo method with lower accuracy and the approximation is then used in a second step in order to compute $\mathbb{E}(Z_\theta) = \mathbb{E}(X)$ by Monte Carlo method as indicated above.

Underlying asset as control variate

If $S(t)$ is an asset price then $\exp(-rt)S(t)$ is a martingale and

$$\mathbb{E}[\exp(-rT)S(T)] = S(0).$$

Suppose we are pricing an option on S with discounting pay-off Y. From independent replications S_i, $i = 1, \ldots, M$, we can form the control variate estimator

$$\frac{1}{M}\sum_{i=1}^{M}\Big(Y_i - \big(S_i - e^{rT}S(0)\big)\Big).$$

Hedge control variates

Because the payoff of a hedged portfolio has a lower standard deviation than the payoff of an unhedged one, using hedges can reduce the volatility of the value of the portfolio.

A delta hedge consists of holding $\Delta = \partial C / \partial S$ shares in the underlying asset, which is rebalanced at the discrete time intervals. At time T, the hedge consists of the savings account and the asset, which closely replicates the payoff of the option. This gives

$$C(t_0)\mathrm{e}^{r(T-t_0)} - \sum_{i=0}^{n} \Big(\frac{\partial C(t_i)}{\partial S} - \frac{\partial C(t_{i-1})}{\partial S} \Big) S_{t_i} \mathrm{e}^{r(T-t_i)} = C(T),$$

where $\partial C(t_{-1} / \partial S = \partial C(t_n) / \partial S = 0$.

After rearranging terms we get

$$C(t_0)\mathrm{e}^{r(T-t_0)}+\sum_{i=0}^{n-1}\frac{\partial C(t_i)}{\partial S}\left(S_{t_{i+1}}-S_{t_i}\mathrm{e}^{r(t_{i+1}-t_i)}\right)\mathrm{e}^{r(T-t_{i+1})}=C(T).$$

Because the term

$$CV=\sum_{i=0}^{n-1}\frac{\partial C(t_i)}{\partial S}\left(S_{t_{i+1}}-S_{t_i}\mathrm{e}^{r(t_{i+1}-t_i)}\right)\mathrm{e}^{r(T-t_{i+1})}$$

is a martingale due to the known property that option price is a value of the hedging portfolio, the mean of CV is zero.

We can use than CV as a control variate.

When we are pricing a call option on S with discounting payoff $C(T) = (S_T - K)^+$, then from independent replications S_T^j, CV^j, $j = 1, \ldots, M$, we can form the control variate estimator

$$C(t_0)\mathrm{e}^{r(T-t_0)} = \frac{1}{M}\sum_{j=1}^{M}\Big(C^j(T) - CV^j\Big).$$

Example: Average Asian option

Asian options are options where payoff depends on the average price of the underlying asset during at least some part of the life of the option. The payoff from the average call is

$$\max(S_{ave} - K, 0) = (S_{ave} - K)^+$$

and that from the average put is

$$\max(K - S_{ave}, 0) = (K - S_{ave})^+,$$

where S_{ave} is the average value of the underlying asset calculated over a predetermined averaging period.

Asian Options and Control Variates

The average call option traded on the market has a payoff of the form

$$\left(\frac{1}{n} \sum_{i=1}^{n} S_{t_i} - K \right)^+$$

Even in the most elementary Black-Scholes model, there is no analytic expression for this price.

On the other hand, for the corresponding geometric average Asian option

$$\left(\left(\prod_{i=1}^{n} S_{t_i} \right)^{\frac{1}{n}} - K \right)^+$$

there is an analytic formula similar to Black-Scholes formula.

Asian Options cont.

Intuition tells us that prices of the above options are similar. One can therefore use the geometric average Asian option price as a control variate to improve efficiency of computation of the price for the arithmetic average Asian option.

The price of geometric (continuous) average call option in the Black-Scholes model is given by the formula

$$e^{-\frac{1}{2}\left(r+\frac{\sigma^2}{6}\right)T} S_0\Phi(b_1) - e^{-rT} K\Phi(b_2),$$

where

$$b_1 = \frac{\log\frac{S_0}{K} + \frac{1}{2}\left(r + \frac{\sigma^2}{6}\right)T}{\frac{\sigma}{\sqrt{3}}\sqrt{T}},$$

$$b_2 = b_1 - \frac{\sigma}{\sqrt{3}}\sqrt{T}.$$

Algorithm

Let P be the price of the geometric average option in B-S model (where an analytic formula is available).
For $k = 1, 2, \ldots, M$ simulate sample paths $(S_{t_0}^{(k)}, S_{t_1}^{(k)}, \ldots, S_{t_n}^{(k)})$ and calculate

$$A_k = e^{-rT}\left(\frac{1}{n}\sum_{i=1}^{n} S_{t_i}^{(k)} - K\right)^{+}, \quad G_k = e^{-rT}\left(\left(\prod_{i=1}^{n} S_{t_i}^{(k)}\right)^{\frac{1}{n}} - K\right)^{+}$$

Then calculate $X_k = A_k - (G_k - P)$.
Price is estimated by

$$a_M = \frac{1}{M}\sum_{k=1}^{M} X_k$$

and the confidence interval ...

Asian option improved

It appears that the formula for the geometric (continuous) average call option gives the price which differs by almost 1% from the price obtained by MC simulations. This price discrepancy can be explained by the fact that in the MC simulation we use discrete geometric averaging and not continuous. Fortunately we can also derive in the Black-Scholes model an analytic formula for the geometric discrete average call option.

The price is given by the formula

$$e^{-rT+(r-\sigma^2/2)T\frac{N+1}{2N}+\sigma^2 T\frac{(N+1)(2N+1)}{12N^2}} S_0\Phi(b_1) - e^{-rT} K\Phi(b_2),$$

where

$$b_1 = \frac{\log\frac{S_0}{K} + (r-\sigma^2/2)T\frac{N+1}{2N} + \sigma^2 T\frac{(N+1)(2N+1)}{6N^2}}{\frac{\sigma}{N}\sqrt{\frac{T(N+1)(2N+1)}{6}}},$$

$$b_2 = \frac{\log\frac{S_0}{K} + (r-\sigma^2/2)T\frac{N+1}{2N}}{\frac{\sigma}{N}\sqrt{\frac{T(N+1)(2N+1)}{6}}}$$

and N is the number of sample points.

Other applications of the control variate technique:

- valuation of options with "strange" payoffs,
- American options,
- various path-dependent options.

Random walk construction

We focus on simulating values $(W(t_1), \ldots, W(t_n))$ at a fixed set of points $0 < t_1 < \cdots < t_n$.
Let $Z_1, \ldots, Z_n$ be independent standard normal random variables.
For a standard Brownian motion $W(0) = 0$ and subsequent values are generated as follows

$$W(t_{i+1}) = W(t_i) + \sqrt{t_{i+1} - t_i} Z_{i+1}, \quad i = 0, \ldots, n-1.$$

For log-normal stock prices and equidistributed time points $t_{i+1} - t_i = \delta t$, for $i = 0, \ldots, n-1$, we have

$$S_{t_{i+1}} = S_{t_i} \exp\left(\left(r - \frac{1}{2}\sigma^2\right)\delta t + \sigma\sqrt{\delta t}\xi_i \right),$$

where ξ_i is a sample from $\mathcal{N}(0,1)$.

Computing the Greeks

Finite Differences

For fixed h > 0, we estimate $\Delta = \frac{\partial\psi}{\partial x}$ by forward finite differences

$$\frac{1}{h}\Big(\mathbb{E}\big[\psi(S_T^{x+h})\big] - \mathbb{E}\big[\psi(S_T^{x})\big]\Big)$$

or by centered finite differences

$$\frac{1}{2h}\Big(\mathbb{E}\big[\psi(S_T^{x+h})\big] - \mathbb{E}\big[\psi(S_T^{x-h})\big]\Big),$$

where S_T^x denotes the value of S_T under the condition $S_0 = x$.

The both terms of these differences can be estimated by Monte Carlo methods.

For instance, for the Black-Scholes model

$$S_t^x = x \exp\Big((r - \sigma^2/2)t + \sigma W_t\Big),$$

Delta can be estimated by

$$\frac{1}{2hM} \sum_{i=1}^{M} \Bigg(\psi\Big((x+h)e^{(r-\sigma^2/2)T+\sigma\sqrt{T}\xi_i}\Big) - \psi\Big((x-h)e^{(r-\sigma^2/2)T+\sigma\sqrt{T}\xi_i}\Big)\Bigg),$$

where the ξ_i are standard normal variables.

General setting

Let a contingent claim have the form

$$Y(\theta) = \psi\big(X_1(\theta), \ldots, X_n(\theta)\big),$$

where θ is a parameter on which the instrument depends.
The derivative with respect to θ is the Greek parameter we are interested in.
The payoff of this instrument is $\alpha(\theta) = \mathbb{E}[Y(\theta)]$.
The forward and centered differences by which we calculate Greeks are

$$\hat{A}_F = h^{-1}\big(\alpha(\theta+h) - \alpha(\theta)\big), \quad \hat{A}_C = (2h)^{-1}\big(\alpha(\theta+h) - \alpha(\theta-h)\big).$$

In addition we have two methods of Monte Carlo simulations for the two values of α which appear in these differences. We can simulate each value of α by an independent sequence of random numbers (this will be denoted by an additional subscript i) or use common random numbers (this will be denoted by a subscript ii).

Hence in fact we have four estimators: $\hat{A}_{F,i}$, $\hat{A}_{F,ii}$, $\hat{A}_{C,i}$ and $\hat{A}_{C,ii}$.
To assess the accuracy of these estimators let us consider their bias and variance. Then we can combine these two values into a mean square error (MSE) which is given by

$$\text{MSE}(\hat{A}) = \text{Bias}^2(\hat{A}) + \text{Var}(\hat{A}).$$

For bias we have an approximation

$$\text{Bias}(\hat{A}) \approx bh^{\beta},$$

where $\beta = 1$ for forward differences and $\beta = 2$ for centered differences (possibly with different positive b).
For variance we have

$$\text{Var}(\hat{A}) \approx \sigma^2 / Mh^{\eta},$$

where M is the number of simulations, $\eta = 1$ for common random numbers and $\eta = 2$ for independent random numbers in simulating the difference of α's (possibly with different values of σ).

Assuming the following step size dependence on M

$$h \approx M^{-\gamma},$$

we find optimal γ

$$\gamma = 1/(2\beta + \eta).$$

Then the rate of convergence for our estimators is

$$\mathcal{O}\left(M^{-\frac{\beta}{2\beta+\eta}}\right).$$

As follows from these approximations, it is advised to use rather centered differences and common random numbers, i.e. the estimator $\hat{A}_{C,ii}$, as it produces smaller error and higher order of convergence.

Pathwise differentiation of the payoff

In cases where ψ is regular and we know how to differentiate S_T^x with respect to x, Delta can be computed as

$$\Delta = \frac{\partial}{\partial x}\mathbb{E}[\psi(S_T^x)] = \mathbb{E}\Big[\frac{\partial}{\partial x}\psi(S_T^x)\Big] = \mathbb{E}\Big[\psi'(S_T^x)\frac{\partial S_T^x}{\partial x}\Big]$$

For instance, in the Black-Scholes model, $\frac{\partial S_T^x}{\partial x} = \frac{S_T^x}{x}$ and

$$\Delta = \mathbb{E}\Big[\psi'(S_T^x)\frac{S_T^x}{x}\Big]$$

provided the derivative ψ' exists.

General setting

The differentiation of the payoff as a method of calculating Greeks is based on the following equality

$$\alpha'(\theta) = \frac{d}{d\theta}\mathbb{E}[Y(\theta)] = \mathbb{E}\Big[\frac{dY}{d\theta}\Big].$$

The interchange of differentiation and integration holds under the assumption that the differential quotient

$$h^{-1}\big(Y(\theta+h) - Y(\theta)\big)$$

is uniformly integrable.

The uniform integrability mentioned on the previous slide holds when

A1. $X_i'(\theta)$ exists with probability 1 for every $\theta \in \Theta$ and $i = 1, \ldots, n$, where Θ is a domain of θ,

A2. $\mathbb{P}(X_i(\theta) \in D_\psi) = 1$ for $\theta \in \Theta$ and $i = 1, \ldots, n$, where D_ψ is a domain on which ψ is differentiable,

A3. ψ is Lipschitz continuous,

A4. for all assets X_i we have

$$|X_i(\theta_1) - X_i(\theta_2)| \leq \kappa_i |\theta_1 - \theta_2|,$$

where κ_i is a random variable with $\mathbb{E}[\kappa_i] < \infty$.

Black-Scholes delta

$$\psi(S_T^x) = \mathrm{e}^{-rT}(S_T^x - K)^+$$

with

$$S_T^x = x\exp\Big((r - \sigma^2/2)T + \sigma\sqrt{T}Z\Big) \text{ and } Z \sim \mathcal{N}(0,1).$$

Then

$$\frac{d\psi}{dx} = \frac{d\psi}{dS_T^x} \times \frac{dS_T^x}{dx}.$$

For the first derivative we clearly have

$$\frac{d}{dS_T^x}(S_T^x - K)^+ = \begin{cases} 0 & \text{if } S_T^x < K, \\ 1 & \text{if } S_T^x > K. \end{cases}$$

This derivative fails to exists at $S_T = K$ but the event $S_T = K$ has probability 0.

Finally

$$\frac{d\psi}{dx} = \mathrm{e}^{-rT}\frac{S_T^x}{x}\mathbb{1}_{S_T^x>K}.$$

This is already the expression which is easy for Monte Carlo simulations.

But Gamma cannot be computed by this approach since even for vanilla options the payoff function in not twicely differentiable.

Likelihood ratio

We have to compute a derivative of $\mathbb{E}[\psi(S_T^x)]$.
Assume that we know the transition density

$$g(S_T^x) = g(x, S_T^x).$$

Then

$$\mathbb{E}[\psi(S_T^x)] = \int \psi(S_T^x) g(S_T^x) dS_T^x = \int \psi(S) g(x, S) dS$$

To calculate Δ we get

$$\begin{aligned}\Delta =& \frac{\partial}{\partial x}\mathbb{E}[\psi(S_T^x)] = \int \psi(S) \frac{\partial}{\partial x} g(x, S) dS = \\ =& \int \psi(S) \frac{\partial \log g}{\partial x} g(x, S) dS = \mathbb{E}\Big[\psi(S) \frac{\partial \log g}{\partial x}\Big].\end{aligned}$$

Black-Scholes gamma

When the transition function $g(x, S_T)$ which is the probability density function is smooth, all derivatives can be easily computed.

For the gamma we have

$$
\begin{aligned}
\Gamma =& \frac{\partial^2}{\partial x^2}\mathbb{E}[\psi(S_T^x)] = \int \psi(S)\left(\frac{\partial^2 \log g}{\partial x^2} g(x, S) + \frac{\partial \log g}{\partial x}\frac{\partial g}{\partial x}\right) dS = \\
=& \int \psi(S)\left(\frac{\partial^2 \log g}{\partial x^2} + \left(\frac{\partial \log g}{\partial x}\right)^2\right) g(x, S) dS = \\
=& \mathbb{E}\left[\psi(S)\left(\frac{\partial^2 \log g}{\partial x^2} + \left(\frac{\partial \log g}{\partial x}\right)^2\right)\right].
\end{aligned}
$$

For the log-normal distributions, we have

$$
\begin{aligned}
g(x, S_T) = {} & \frac{1}{\sqrt{2\pi\sigma^2 T}\, S_T} \times \\
& \exp\left(- \frac{\left(\log S_T - \log x - (r - \sigma^2/2)T\right)^2}{2\sigma^2 T} \right), \\
\frac{\partial \log g(x, S_T)}{\partial x} = {} & \frac{\log S_T - \log x - (r - \sigma^2/2)T}{x\sigma^2 T}, \\
\frac{\partial^2 \log g(x, S_T)}{\partial x^2} = {} & -\frac{1 + \log S_T - \log x - (r - \sigma^2/2)T}{x^2\sigma^2 T}.
\end{aligned}
$$

Putting this all together, we have the likelihood ratio method for the gamma of options in the Black-Scholes model.

A very nice survey of applications of Monte Carlo to pricing of financial derivatives may be found in

Boyle, Brodie, Glasserman (1997) ”Monte Carlo methods for security pricing”, *Journal of Economic Dynamics and Control* **21**.

There is also a fundamental monograph

Paul Glasserman *Monte Carlo Methods in Financial Engineering*, Springer 2004.