Monte Carlo - 1st class

Anna Battauz

Bocconi University

1 Monte Carlo

1.1 Idea, problems, numerical issues. Efficiency and bias [G1.1].

The key idea of Monte Carlo methods. How to evaluate a Monte Carlo approximation.

1.2 The inverse function method (discrete and continuous variables) [G1.1].

The most important technique to sample whenever there exists a *generalized* distribution function

1.3 Simulating asset prices. CEV, Square-Root diffusions [G3.4].

Euler discretization when *exact* simulation is unfeasible

2 Monte Carlo: a very simple idea I

Goal: estimate

$$\begin{aligned} \alpha &= \int_0^1 f(x)\,dx \\ &= \mathbb{E}[f(X)] \text{ where } X \sim U(0;1) \end{aligned}$$

Draw n independent replications of $X_i \sim U(0;1)$
The sample mean

$$\widehat{\alpha}_n := \frac{1}{n}\sum_{i=1}^{n} f(X_i)$$

is the Monte Carlo (MC) estimator of α.

3 Monte Carlo: a very simple idea II

The Monte Carlo (MC) estimator i.e. the sample mean

$$\widehat{\alpha}_n := \frac{1}{n} \sum_{i=1}^{n} f\left(X_i\right)$$

is a *strongly consistent estimator* of $\alpha = \mathbb{E}\left[f\left(X\right)\right]$

$$\widehat{\alpha}_n \underset{n \to \infty}{\longrightarrow} \alpha \text{ by the Strong Law Large Numbers,}$$

that is also *unbiased*

$$\mathbb{E}\left[\widehat{\alpha}_n\right] = \mathbb{E}\left[f\left(X_i\right)\right] = \alpha.$$

4 Confidence Interval for the Monte Carlo estimator I

Denote with

$$\sigma_f = \sqrt{\mathbb{E}\left[\left(f\left(X_i\right)-\alpha\right)^2\right]} = \text{ standard deviation of } f\left(X_i\right).$$

The Central Limit Theorem (CLT) implies

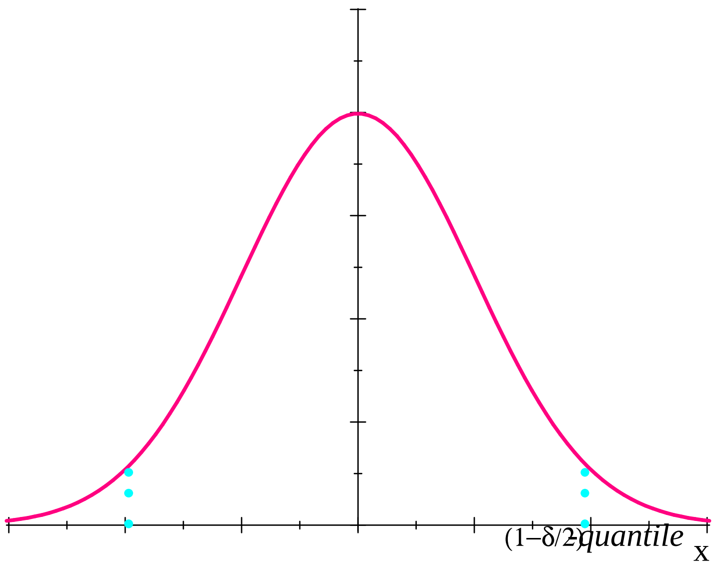
$$\frac{\widehat{\alpha}_n - \alpha}{\frac{\sigma_f}{\sqrt{n}}} \underset{n\to\infty}{\Rightarrow} N\left(0;1\right) \text{ in distribution}$$

The CLT allows to compute the confidence interval.

$$\left|\widehat{\alpha}_n - \alpha\right| : \text{ absolute error in the estimator.}$$

For a given δ (for instance $\delta = 5\%$ or $\delta = 1\%$) denote with

$z_{\frac{\delta}{2}}$ the $\left(1-\frac{\delta}{2}\right)$ – quantile of the $N\left(0;1\right)$ distribution



The $(1-\delta)-$confidence interval can be deduced from:

$$\mathbb{P}\left[\frac{|\widehat{\alpha}_n - \alpha|}{\frac{\sigma_f}{\sqrt{n}}} \leq z_{\frac{\delta}{2}}\right] \approx 1-\delta \text{ when } n \text{ is large enough}$$

implying

$$|\widehat{\alpha}_n - \alpha| \leq \frac{\sigma_f}{\sqrt{n}} z_{\frac{\delta}{2}} \text{ with probability } 1-\delta \text{ when } n \text{ is large enough}$$

Radius of the $(1-\delta)-$*confidence interval* is

$$\frac{\sigma_f}{\sqrt{n}} z_{\frac{\delta}{2}}$$

When $\delta = 5\%$ then $z_{\frac{\delta}{2}} = 1.96$.

5 Confidence Interval for the Monte Carlo estimator II

Problem: usually when $\alpha = \mathbb{E}\left[f\left(X\right)\right]$ is unknown, $\sigma_f = \sqrt{\mathbb{E}\left[\left(f\left(X_i\right) - \alpha\right)^2\right]}$ is unknown too.

Solution: replace σ_f with the sample standard deviation

$$\widehat{\sigma}_n = \sqrt{\frac{1}{n-1}\sum_{i=1}^{n}\left[\left(f\left(X_i\right) - \widehat{\alpha}_n\right)^2\right]}$$

The radius of the confidence interval is still

$$\frac{\widehat{\sigma}_n}{\sqrt{n}} z_{\frac{\delta}{2}}$$

For a more prudent estimation you can replace $z_{\frac{\delta}{2}}$ with the quantile of the *t-Student distribution* with $n-1$ degrees of freedom

6 Convergence of the Monte Carlo estimator

$$\text{Radius of Confidence Interval } \propto \frac{1}{\sqrt{n}}$$

implies that

$$\textit{rate of convergence} = O\left(\frac{1}{\sqrt{n}}\right)$$

Comparison: to estimate

$$\int_{[0;1]^d} f\left(\underline{x}\right) d\underline{x}$$

trapezoidal rules achieve a rate of convergence of $O\left(\sqrt[d]{\frac{1}{n^2}}\right)$.

Monte Carlo does a better job when $d > 4$ (although variance has to be controlled)

7 Example: Call option in the Black-Scholes model

$$c = \mathbb{E}\left[e^{-rT}\left(\underbrace{S_0 e^{\left(r-\frac{\sigma^2}{2}\right)T+\sigma\sqrt{T}Z}}_{S(T)} - K\right)^+\right]$$

To estimate c :

For $i = 1, ..., n$

generate $Z_i \sim N(0;1)$ i.i.d

$S_i(T) = S_0 e^{\left(r-\frac{\sigma^2}{2}\right)T+\sigma\sqrt{T}Z_i}$ *simulated terminal underlying price*

$c_i = e^{-rT}(S_i(T) - K)^+$ *simulated discounted terminal call payoff*

$$\widehat{c}_n := \frac{1}{n}\sum_{i=1}^{n} c_i$$ Monte Carlo estimator of the call option price

$\widehat{c}_n$ is *strongly consistent and unbiased estimator* of c

Confidence interval: sample standard deviation

$$\widehat{\sigma}_n = \sqrt{\frac{1}{n-1}\sum_{i=1}^{n}\left[(c_i - \widehat{c}_n)^2\right]}$$

Radius of confidence interval:

$$\frac{\widehat{\sigma}_n}{\sqrt{n}} z_{\frac{\delta}{2}} \text{ with } z_{\frac{\delta}{2}} = 1.96 \text{ for } 1-\delta = 95\%$$

Error in the estimate:

$$\widehat{c}_n - c$$

in this case can be computed (because c can be computed)

8 Example: the confidence interval for the call option

To provide the confidence interval for c : add the last two rows to the algorithm

For $i = 1, ..., n$

generate $Z_i \sim N(0;1)$ i.i.d

$S_i(T) = S_0 e^{\left(r - \frac{\sigma^2}{2}\right)T + \sigma\sqrt{T}Z_i}$ *simulated terminal underlying price*

$c_i = e^{-rT}(S_i(T) - K)^+$ *simulated discounted terminal call payoff*

$\widehat{c}_n = \frac{1}{n}\sum_{i=1}^n c_i$ *estimator for* c

$\widehat{\sigma}_n = \sqrt{\frac{1}{n-1}\sum_{i=1}^n \left[(c_i - \widehat{c}_n)^2\right]}$ *sample standard deviation*

$r_n = \frac{\widehat{\sigma}_n}{\sqrt{n}} 1.96$ *radius of the confidence interval*

9 Bias

The sample mean is an unbiased estimator for the mean.
Bias? only due to **discretization errors** (of the option payoff, of the underlying dynamics,...)

$$Bias\left(\widehat{\alpha}_n\right) = \mathbb{E}\left[\widehat{\alpha}_n\right] - \alpha$$

Reduce the Mean Square Error (MSE) as much as possible:

$$\begin{aligned} MSE\left(\widehat{\alpha}_n\right) &= \mathbb{E}\left[\left(\widehat{\alpha}_n - \alpha\right)^2\right] \\ &= \left(\mathbb{E}\left[\widehat{\alpha}_n\right] - \alpha\right)^2 + \mathbb{E}\left[\left(\widehat{\alpha}_n - \mathbb{E}\left[\widehat{\alpha}_n\right]\right)^2\right] \\ &= \left(Bias\left(\widehat{\alpha}_n\right)\right)^2 + Variance\left(\widehat{\alpha}_n\right) \end{aligned}$$

to control for both variance (radius of the confidence interval) and bias.
Usually variance matters more than bias (except for American options).

We are going to see for finite difference approximations for the Greeks:

h discretization mesh with $h \to 0$ to reduce the bias

n number of replications with $n \to \infty$ to reduce the variance

Reducing the MSE implies the rule of thumb $nh^2 \to \infty$

10 Efficiency I

If you have two different ways to simulate Z_i, which is the best?

`Method 1:` variance per replication V_1, time per replication τ_1

`Method 2:` variance per replication V_2, time per replication τ_2

Non trivial case: $V_1 < V_2$ but $\tau_1 > \tau_2$

T : total time at your disposal

11 Efficiency II

For $h = 1, 2:$
$n_h = \frac{T}{\tau_h}$ number of replications in T with `Method h`

$r_h = \frac{\sqrt{V_h}}{\sqrt{n_h}} z_{\frac{\delta}{2}}$ radius of confidence interval for `Method h`

Choose the method with smallest radius of confidence interval

$$\frac{\sqrt{V_h}}{\sqrt{n_h}} = \frac{\sqrt{V_h \tau_h}}{\sqrt{T}}$$

i.e. the smallest product

$$\underbrace{V_h}_{Var \text{ per replication}} \quad \underbrace{\tau_h}_{\text{Time per replication}}$$

If the time per replication is stochastic (e.g. barrier options), use the same rule replacing τ_h with the *expected* time per replication

12 The inverse function method I

$X \sim F_X$ distribution function, i.e. $\mathbb{P}[X \leq x] = F_X(x)$
If $\exists F_X^{-1} : [0;1] \to \Re$ the *inverse* function of F_X then

$$F_X^{-1}(U) \sim X \text{ where } U \text{ is } Uniform(0;1)$$

Proof: by the definition of inverse function we have

$$F_X^{-1}(u) \leq x \text{ iff } u \leq F_X(x) \tag{1}$$

since

$$F_X^{-1}(u) \leq x \Longleftrightarrow F_X\left(F_X^{-1}(u)\right) \leq F_X(x) \Longleftrightarrow u \leq F_X(x).$$

But then

$$\mathbb{P}\left[F_X^{-1}(U) \leq x\right] = \mathbb{P}[U \leq F_X(x)] = F_X(x)$$

i.e. $F_X^{-1}(U) \sim X$.

13 The generalized inverse function method I

If $\nexists F_X^{-1}?$ (for instance when F_X is constant on some interval)
$\rightsquigarrow$ **Generalized Inverse Distribution Function**

$$F_X^{-1}(u) = \inf\{t : F_X(t) \geq u\}$$

Still

$$F_X^{-1}(U) \sim X \text{ where } U \text{ is } Uniform\,(0;1)$$

Proof. If the equivalence (1) works, we conclude that

$$\mathbb{P}\left[F_X^{-1}(U) \leq x\right] = \mathbb{P}\left[U \leq F_X(x)\right] = F_X(x)$$

Let us now verify (1).

14 The generalized inverse function method II

1. Suppose $t^* = F_X^{-1}(u) \leq x$ then $F_X(t^*) \leq F_X(x)$ because F_X is increasing.
   But $F_X(t^*) \geq u$ because $t^* = F_X^{-1}(u) = \inf\{t : F_X(t) \geq u\}$.
   Hence $u \leq F_X(t^*) \leq F_X(x)$ i.e. $u \leq F_X(x)$.

2. Suppose $u \leq F_X(x)$.
   Then $x \in \{t : F_X(t) \geq u\}$ and therefore
   $F_X^{-1}(u) = \inf\{t : F_X(t) \geq u\} \leq x$

15 A couple of possible applications of the inverse function method

The inverse function method can be used to simulate

- continuous random variables
- discrete random variables

16 Simulating the exponential random variable

$$X \sim Exp\left(\frac{1}{\theta}\right) \quad \text{i.e. } F_X(x) = \begin{cases} 1 - e^{-\frac{x}{\theta}} \text{ for } x \geq 0 \\ 0 \text{ for } x < 0 \end{cases}$$

$$F_X^{-1}(u)?$$

$$\begin{aligned} u &= F_X(x) = 1 - e^{-\frac{x}{\theta}} \\ -\frac{x}{\theta} &= \ln(1-u) \\ x &= -\theta \ln(1-u) = F_X^{-1}(u) \end{aligned}$$

Take $U \sim Uniform(0;1)$ then

$$\begin{aligned} X &= -\theta \ln(1-U) \\ &\sim -\theta \ln(U) \sim Exp\left(\frac{1}{\theta}\right) \end{aligned}$$

17 Simulating a discrete random variable

$$X \ : \ \Omega \to \{x_1, ..., x_n\}$$
$$\mathbb{P}[X = x_i] \ = \ p_i \ \text{ with } i = 1, ..., n$$

Set

$$q_0 \ = \ 0$$
$$q_j \ = \ \sum_{i=1}^{j} p_i = \mathbb{P}[X \leq x_j] = F_X(x_j) \ \text{ for } j = 1, ..., n$$

- `Simulate` $U \sim Uniform(0;1)$
- `Find index` j `such that` $q_{j-1} < U \leq q_j$ `and assign to` X `the value` x_j

In this way

$$\mathbb{P}[X = x_j] = \mathbb{P}[q_{j-1} < U \leq q_j] = q_j - q_{j-1} = p_j \quad \text{for } j = 1, ..., n$$

18 Simulating lognormal paths: Euler Discretization

Lognormal security:

$$\frac{dS(t)}{S(t)} = \mu dt + \sigma dW(t)$$

where W is std B.M.
If the reference probability measure is the risk neutral one, then $\mu = r$ the riskless interest rate.

Conditionally exact discretization of S :
Fix $\Delta t = \frac{T}{M}$ and $t_j = \frac{jT}{M}$ for $j = 0, ..., M$

$$S(t_j)\big|_{S(t_{j-1})} = S(t_{j-1})\, e^{\left(\mu - \frac{\sigma^2}{2}\right)\Delta t + \sigma \Delta W(t_{j-1})}$$

where $\Delta W(t_{j-1}) \sim N(0; \Delta t)$ independent from $\mathcal{F}_{t_{j-1}}$
Hence

$$S(t_j)\big|_{S(t_{j-1})} = S(t_{j-1})\, e^{\left(\mu - \frac{\sigma^2}{2}\right)\Delta t + \sigma \sqrt{\Delta t} Z_j} \text{ with } Z_j \sim N(0; 1) \text{ i.i.d.}$$

is *(conditionally) exact in distribution.*

19 A discretely monitored Asian option I

Monitoring dates: $t_j = \frac{jT}{M}$ for $j = 0, ..., M$

Final payoff (European option):

$$\left(\underbrace{\frac{1}{M+1} \sum_{j=0}^{M} S\left(t_j\right)}_{A(T)} - K \right)^+$$

No-arbitrage initial price of the option:

$$\mathbb{E}^{\mathbb{Q}}\left[e^{-rT}\left(A\left(T\right) - K\right)^+\right]$$

20 A discretely monitored Asian option II

1. `Simulate` n `paths of` S `to evaluate` $A^i(T)$, `the average price of` S, `for` $i = 1, ..., n$
2. `The MC estimator for the initial price of the option is`

$$\frac{1}{n}\sum_{i=1}^{n} e^{-rT}\left(A^i\left(T\right) - K\right)^+$$

3. This is an *unbiased* estimator of the Asian option initial price, because paths are simulated with the exact conditional distribution.
4. If the option is monitored at any $t \in [0; T]$, i.e. if the final payoff is

$$\left(\frac{1}{T}\int_0^T S\left(t\right) dt - K\right)^+,$$

then the estimator is biased (discretization error of the payoff).

21 A discretely monitored Asian option III

$S(t_0) = S(0)$ `and` $A(t_0) = S(0)$ *initialize* S *and* A

`For` $i = 1, ..., n$

`for` $j = 1, ..., M$

`Simulate` $Z_j^i \sim N(0;1)$ `i.i.d.`

$S^i(t_j) = S^i(t_{j-1})\, e^{\left(r-q-\frac{\sigma^2}{2}\right)\Delta t + \sigma\sqrt{\Delta t} Z_j^i}$ *simulated underlying path in* $[t_{j-1}, t_j]$

$A^i(t_j) = A^i(t_{j-1}) + S^i(t_j)$ *sum of all simulated underlying prices in* $[t_0, t_j]$

$A^i(t_M) = \frac{A^i(t_M)}{M+1}$ *simulated underlying average price in* $[t_0, t_M]$

$\frac{1}{n}\sum_{i=1}^n e^{-rt_M}\left(A^i(t_M) - K\right)^+$ *Monte Carlo estimator of initial option price*

22 Euler discretization when the exact solution is difficult to be computed

$$\frac{dS(t)}{S(t)} = \mu(t)\,dt + \sigma(t)\,dW(t)$$

where W is std B.M. and $\mu(t), \sigma(t)$ are deterministic.

1. Fix M such that $\Delta t = \frac{T}{M}$ and $t_j = \frac{jT}{M}$ for $j = 0, ..., M$
2. *Freeze* coefficients $\mu(t), \sigma(t)$ at the beginning of each subinterval $[t_{j-1}; t_j]$ and consider them as *(locally) constant* coefficients:

$$S(t_j)|_{S(t_{j-1})} = S(t_{j-1})\, e^{\left(\mu(t_{j-1}) - \frac{\sigma^2(t_{j-1})}{2}\right)\Delta t + \sigma(t_{j-1})\sqrt{\Delta t} Z_j} \quad \text{with } Z_j \sim N(0;1) \text{ i.i.d.}$$

The discretization error (due to the freezing of the coefficients) reduces as $\Delta t \to 0$.
A similar useful approximation can be applied if $\mu(t), \sigma(t)$ are *stochastic*, by properly simulating $\mu(t_{j-1}), \sigma(t_{j-1})$ for $j = 1, ..., M$

23 Simulating square root diffusions (CIR for interest rates)

The CIR model for the short rate

$$dr(t) = \alpha(b - r(t))\,dt + \sigma\sqrt{r(t)}dW(t)$$

where
$\alpha > 0$ is the speed of convergence
$b > 0$ is the limit value for $r : \lim_{t\to\infty} r(t) = b$
$r(0) > 0 \Rightarrow r(t) \geq 0$ for all t
$2\alpha b \geq \sigma^2 \Rightarrow r(t) > 0$ for all t a.s.

To fit a Bond Term Structure:

$$B(0;T) = \mathbb{E}\left[e^{\int_0^T r(t)dt}\right] \quad \text{for } T = T_1, ..., T_K$$

take a deterministic coefficient $b = b(t)$.

24 Euler discretization for the CIR model

From the SDE

$$dr(t) = \alpha(b(t) - r(t))\,dt + \sigma\sqrt{r(t)}dW(t)$$

Fix M such that $\Delta t = \frac{T}{M}$ and $t_j = \frac{jT}{M}$ for $j = 0, ..., M$.
The discretization:

$$r(t_j) - r(t_{j-1}) = \alpha(b(t_{j-1}) - r(t_{j-1}))\,\Delta t + \sigma\sqrt{(r(t_{j-1}))^+}\sqrt{\Delta t}Z_j$$

follows with $Z_j \sim N(0;1)$ i.i.d. for $j = 1, ..., M$

The $(r(t_{j-1}))^+$ inside the argument of the square root in the diffusive coefficient avoids problems with possible negative sign of r (due to discretization errors).

This method is biased (discretization error)

25 Exact simulation of the CIR model with constant coefficients

For $u < t$

$$r(t)|_{r(u)} \sim \frac{\sigma^2 \left(1 - e^{-\alpha(t-u)}\right)}{4\alpha} \chi_{\nu}^{1,2}(\lambda)$$

$\chi_{\nu}^{1,2}(\lambda)$ non-central Chi-square distribution with

$$\nu = \frac{4b\alpha}{\sigma^2} \text{ degrees of freedom}$$

$$\lambda = \frac{4\alpha e^{-\alpha(t-u)}}{\sigma^2 \left(1 - e^{-\alpha(t-u)}\right)} r(u) \text{ non-centrality parameter}$$

26 How to sample a non-central Chi-Square distribution

1. If $\nu > 1$ then

$$\chi_{\nu}^{1,2}(\lambda) \sim \left(Z + \sqrt{\lambda}\right)^2 + \chi_{\nu-1}^2$$

   where $Z \sim N(0;1)$ and $\chi_{\nu-1}^2$ is a Chi-Square distribution (can be sampled within Excel)

2. If $\nu \leq 1$ the non-central Chi-Square is a Chi-Square with a random number degrees of freedom

$$N \sim Poisson\left(\frac{\lambda}{2}\right)$$

$$\chi_{\nu}^{1,2}(\lambda) \sim \chi_{\nu+2N}^{2}$$

Hence sample $N \sim Poisson\left(\frac{\lambda}{2}\right) \Rightarrow$ sample a $\chi_{\nu+2N}^{2}$ Chi-Square with $\nu+2N$ degrees of freedom, then:

$$r(t)|_{r(u)} \sim \frac{\sigma^2\left(1-e^{-\alpha(t-u)}\right)}{4\alpha}\chi_{\nu+2N}^{2}$$

27 Sampling the CEV model

The Constant Elasticity of Variance (CEV) model:

$$dS(t) = \mu S(t)\,dt + \sigma (S(t))^{\frac{\beta}{2}}\,dW(t)$$

i.e.

$$\frac{dS(t)}{S(t)} = \mu dt + \sigma (S(t))^{\frac{\beta-2}{2}}\,dW(t)$$

Typically $\beta < 2$ to describe the negative relation between volatility and current price of S.

$$X(t) = (S(t))^{2-\beta}$$

is a square-root.

Hence: sample the square root X, then compute $S(t) = (X(t))^{\frac{1}{2-\beta}}$