Numerical methods: efficient Monte Carlo simulation

—

DU Financial Engineering

École Polytechnique & Sorbonne Université

—

G. Pagès

LPSM–SU

—

2025-2026

1. **Basic Monte Carlo simulation**
   - From random numbers to Monte Carlo simulation
   - Central Limit theorem, Law of Iterated Logarithm
   - Confidence intervals
   - Back to the simulation of Gaussian random vectors
2. **Variance reduction techniques**
   - Control variate: two random variables for one expectation
   - Dynamic control variate (regression)
   - Typical examples: Parity equations and option pricing
   - Antithetic (back to)
   - Importance sampling
   - Stratified sampling
3. **From the hedge side**
4. **Quasi-Monte Carlo methods**
   - Monte Carlo revisited
   - Uniformly distributed sequences
   - Sequences with low discrepancy
   - Pros & Cons (curse of dimensionality, etc)

Introduction: the Monte Carlo method

▷ Aim: Let $X : (\Omega, \mathcal{A}, \mathbb{P}) \to \mathbb{R}$ be an integrable random variable i.e. satisfying $\mathbb{E}\,|X| < +\infty$. Then the *expectation of X* exists

$$m_X = \mathbb{E}\,X = \int_\Omega X(\omega)\mathbb{P}(d\omega) \in \mathbb{R}.$$

Basic question: How to compute m_X?

▷ Method 1: transfer principle. Switch – or not – from the scenario space to the state space:

$$X : \underbrace{(\Omega, \mathcal{A}, \mathbb{P})}_{\textit{Scenarios space}} \longrightarrow \underbrace{\mathbb{R}}_{\textit{state space}}$$

▷ Distribution

$$\forall\, B \text{Borel set of } \mathbb{R} \quad \mathbb{P}_X(B) = \mathbb{P}\big(\{\omega \in \Omega : X(\omega) \in B\}\big).$$

– $\mathbb{P}$ is a probability on $(\Omega, \mathcal{A})$ and $\mathbb{P}_X$ is a probability (exercise!) on the state space $\mathbb{R}$.

– X random variable means: if B is a Borel set (say an interval) then $\{\omega \in \Omega : X(\omega) \in B\} \in \mathcal{A}$

– Notation : $\{X \in B\} := X^{-1}(B) := \{\omega \in \Omega : X(\omega) \in B\}$.

▷ Integration theory.

If $X = \sum_{i \in I} \alpha_i \mathbf{1}_{A_i}$ where $(A_i)_{i \in I}$ is a finite measurable partition of Ω

$$\mathbb{E}\, X = \sum_{i \in I} \alpha_i \mathbb{P}(A_i)$$

[It is Blaise Pascal's idea to multiply the value by its probability to define *expectation*]. We note that

$$A_i = \{\omega \in \Omega : X(\omega) = \alpha_i\} = X^{-1}(\{\alpha_i\}) = \{X = \alpha_i\}.$$

- Then

$$\begin{aligned}\mathbb{E}\,X &= \sum_{i\in I}\alpha_i\mathbb{P}(\{\omega\in\Omega : X(\omega)=\alpha_i\})\\ &= \sum_{i\in I}\alpha_i\mathbb{P}_X(\{\alpha_i\})\\ &= \int_{\mathbb{R}}\xi\mathbb{P}_X(d\xi).\end{aligned}$$

- In particular: if $X = \mathbf{1}_A$ i.e $X(\omega) = 1$ if $\omega\in A$ and $X(\omega) = 0$ if $\omega\notin A$. Then

$$\begin{aligned}\mathbb{E}\,X &= 1\times\mathbb{P}(A) + 0\times\mathbb{P}(A^c)\\ &= \mathbb{P}(A) = \mathbb{P}(\{X=1\}) = \mathbb{P}_X(\{1\}) = 1\times\mathbb{P}_X(\{1\}) + 0\times\mathbb{P}_X(\{0\}) = \int_{\mathbb{R}}\xi\mathbb{P}_X(d\xi).\end{aligned}$$

- Extends to more general random variables: one has the transfert property

$$\mathbb{E}\,X = \int_{\mathbb{R}}\xi\,\mathbb{P}_X(d\xi)$$

  with $\mathbb{P}_X = \mathbb{P}\circ X^{-1}$ distribution of X on $\mathbb{R}$.
- Both are integrals, one on Ω, the other on $\mathbb{R}$.

Examples I

X takes finitely or countably many values. Staying on the scenario space is possible

- Heads & Tails (once): $X \stackrel{d}{=} Bernoulli(\{0,1\},p)$ with $\mathbb{P}(\text{“}Head'') = p \in [0,1]$.

$$\mathbb{E}\,X = 1 \times \mathbb{P}(Head) + 0 \times \mathbb{P}(Tail) = p \cdot 1 + 0 \cdot (1-p) = p$$

- Head & Tails (n times): $Binomial(n,p)$,

$$\mathbb{E}\,X = n \cdot p$$

since $X = X_1 + \cdots + X_n$ using linearity of expectation... One also has

$$\mathbb{P}(X = k) = \binom{n}{k}\; p^k (1-p)^{n-k}.$$

Then, as $\binom{n}{k} = \frac{n}{k}\binom{n-1}{k-1}$

$$\mathbb{E}\,X = \sum_{k=0}^{n} k\,\mathbb{P}(X=k) = \sum_{k=1}^{n} k\,\binom{n}{k} p^k (1-p)^{n-k}$$

$$= np \sum_{k=1}^{n} \binom{n-1}{k-1} p^{k-1} (1-p)^{(n-1)-(k-1)} = np.1 \quad \text{by Newton's binomial formula.}$$

- First succes: *Geometric*$^*(p)$,

$$
\begin{aligned}
\mathbb{E}\,X &= \sum_{k\geqslant 1} k\,p(1-p)^{k-1}\\
&= p\sum_{k\geqslant 1} k(1-p)^{k-1} = p\left(\sum_{k\geqslant 0} x^k\right)'_{|x=1-p}\\
&= p\Big(\frac{1}{1-x}\Big)' = p\frac{1}{(1-(1-p))^2} = p\frac{1}{p^2} = \frac{1}{p}
\end{aligned}
$$

  by geometric formula. Here no smart trick like for the binomial distribution.

- More generally $\mathbb{P}(X = x_i) = p_i$, $i \in I$, I finite or countable

$$
\mathbb{E}\,X = \sum_{i\in I} p_i x_i
$$

More difficult with Head & Tails

- Let $S_n = X_1 + \cdots + X_n$ with X_k independent (...) identically distributed (i.i.d.) Bernoulli distributed with values ± 1.

$$\mathbb{P}(X_1 = 1) = p \in [0,1] \quad \mathbb{P}(X_1 = -1) = 1 - p \in [0,1].$$

so that $\mathbb{E}\, X = 1/p - 1.(1-p) = 2p - 1$.

- Assume $p = \frac{1}{2}$. Then $\mathbb{E}\, X_k = 0$, $\mathbb{E}\, S_n = 0$ and (...) $\mathbb{E}\, S_n^2 = n$. Let

$$a \in \mathbb{N},\ b \in \mathbb{Z}, \qquad \tau_b = \min\big\{k : S_k = b\big\}, \quad S_{\tau_b \wedge \tau_a}$$

- One has $\mathbb{E}\, S_{\tau_b \wedge \tau_a} = 0$ (stopped martingale) but, as $S_n^2 - n$ is also a martingale

$$\mathbb{E}\, \tau_b \wedge \tau_a = \mathbb{E}\, S^2_{\tau_b \wedge \tau_a} = a^2 \mathbb{P}(\tau_a < \tau_b) + b^2 \mathbb{P}(\tau_b < \tau_a).$$

- Why not use (repeated) simulation?

Examples II

▷ What happens if the law of X takes infinitely many values and $\mathbb{P}_X$ is absolutely continuous to the Lebesgue measure i.e.

$$\mathbb{P}_X(d\xi) = f_X(\xi)d\xi$$

▷ One has to do some work on the state space. If the *density* f_X is explicit then closed form/quadrature formula.

- Closed form: $f_X(\xi) = \lambda e^{-\lambda\xi}\mathbf{1}_{\{x\geqslant 0\}}$, $\lambda > 0$, then $\mathbb{E}\,X = 1/\lambda$,

$$\mathbb{E}\,X = \int_0^{+\infty} \xi\lambda e^{-\lambda\xi}d\xi = \cdots = \frac{1}{\lambda}$$

- Gaussian law: $f_X(\xi) = \exp\left(-\frac{\xi^2}{2}\right)/\sqrt{2\pi}$, $\mathbb{E}\,X = 0$.

$$\mathbb{E}\,X = \int_{-\infty}^{+\infty} \xi\exp\left(-\frac{\xi^2}{2}\right)/\sqrt{2\pi}d\xi = 0.$$

- Symmetry argument: f_X even, then $\mathbb{E}\,X = 0$, if $f_X(L-\xi) = f_X(\xi)$ then $\mathbb{E}\,X = L/2$, etc.

- Quadrature: If $X = \varphi(U)$, $U \stackrel{d}{=} \mathcal{U}([0,1])$,

$$\mathbb{E}\,X = \mathbb{E}\,\varphi(U) = \int_0^1 \varphi(u)du = \lim_n \frac{1}{n}\sum_{k=1}^{n} \varphi\Big(\frac{2k-1}{2n}\Big).$$

  among others “quadrature methods" to compute such integrals.

- What about this $\mathbb{E}\,\varphi(U)$?

Function of a random variable

- Discrete world: $X : (\Omega, \mathcal{A}, \mathbb{P}) \to \{x_i,\ i \in I\}$ ($\subset \mathbb{N}$ or $\mathbb{R}^d$ or $\dots$) with $\mathbb{P}(X = x_i) = p_i$, $i \in I$ ($I = \{1, \dots, N\}$ or $I = \mathbb{N}$, etc, at most countable).

  Let $g : \mathbb{R}$ or $\mathbb{R}^d \to \mathbb{R}$

$$\mathbb{E}\, g(X) = \sum_{i \in I} p_i g(x_i).$$

- Absolutely continuous world: $X : (\Omega, \mathcal{A}, \mathbb{P}) \to \mathbb{R}$ (or $\mathbb{R}^d$ or any vector space...) with density $f : \mathbb{R} \to \mathbb{R}$.

  Let $\varphi : \mathbb{R}$ or $\mathbb{R}^d \to \mathbb{R}$. Then

$$\mathbb{E}\, \varphi(X) = \int_{\mathbb{R}} \varphi(\xi) \underbrace{f(\xi) d\xi}_{= \mathbb{P}_X(d\xi)}.$$

Independence of random vectors: discrete world

- Discrete world: $X : (\Omega, \mathcal{A}, \mathbb{P}) \to \{x_i,\, i \in I\}$ and $Y : (\Omega, \mathcal{A}, \mathbb{P}) \to \{y_j,\, j \in J\}$ (or $\subset \mathbb{R}$ or $\mathbb{R}^d$) are independent if

$$\forall\, i \in I,\ j \in J, \quad \mathbb{P}(\{X = x_i\} \cap \{Y = y_j\}) = \mathbb{P}(\{X = x_i\}) \cdot \mathbb{P}(\{Y = y_j\})$$

or, equivalently, for every $\varphi, \psi : \mathbb{R}^d \to \mathbb{R}$ (for which the quantities below exist),

$$\mathbb{E}\,\varphi(X)\psi(Y) = \mathbb{E}\,\varphi(X) \cdot \mathbb{E}\,\psi(Y).$$

- Example. Roll of two dice. Let $\Omega = \{1, 2, 3, 4, 5, 6\}^2 = \{(i, j) : 1 \leqslant i, j \leqslant 6\}$ $\omega = (\omega_1, \omega_2)$

$$X(\omega) = \omega_1 \quad Y(\omega) = \omega_2$$

are independent since

$$\mathbb{P}(X = i,\ Y = j) = \frac{1}{36} = \frac{1}{6} \cdot \frac{1}{6} = \mathbb{P}(X = i) \cdot \mathbb{P}(Y = j),\ \forall\, i, j.$$

Independence of random vectors: Interpretation

- The identity

$$\forall\, i \in I,\ j \in J, \quad \mathbb{P}(\{X = x_i\} \cap \{Y = y_j\}) = \mathbb{P}(\{X = x_i\}) \cdot \mathbb{P}(\{Y = y_j\})$$

also reads

$$\forall\, i \in I,\ j \in J, \quad \underbrace{\mathbb{P}(\{X = x_i\} \,|\, \{Y = y_j\})}_{\text{conditional prob.}} := \frac{\mathbb{P}(\{X = x_i\} \cap \{Y = y_j\})}{\mathbb{P}(\{Y = y_j\})}$$

$$= \mathbb{P}(\{X = x_i\}).$$

Independence of random vectors: continuous world

- Absolutely continuous world: $X : (\Omega, \mathcal{A}, \mathbb{P}) \to \mathbb{R}$ (or $\mathbb{R}^d$ or any vector space...) with density $f : \mathbb{R} \to \mathbb{R}$ and $Y : (\Omega, \mathcal{A}, \mathbb{P}) \to \mathbb{R}$ (or $\mathbb{R}^d$ or any vector space...) with density g are independent if (X, Y) has a density such that

$$\underbrace{\mathbb{P}_{(X,Y)}(dx, dy)}_{\text{law of } (X,Y)} = f(x)g(y)dxdy$$

or, equivalently, for every $\varphi, \psi : \mathbb{R}^d \to \mathbb{R}$ (for which such quantities exist),

$$\mathbb{E}\,\varphi(X)\psi(Y) = \mathbb{E}\,\varphi(X) \cdot \mathbb{E}\,\psi(Y).$$

Independence of random vectors: Example

▷ Normal distribution $\mathcal{N}(0, I_d)$ on $\mathbb{R}^d$. Let $X = (X_1, \ldots, X_d) : (\Omega, \mathcal{A}, \mathbb{P}) \to \mathbb{R}^d$ with density

$$
\begin{aligned}
\mathbb{P}_X(dx) = \underbrace{\mathbb{P}(X \in dx)}_{\text{alternative notation}} &= e^{-\frac{|x|^2}{2}} \frac{dx}{(2\pi)^{d/2}} \\
&= e^{-\frac{x_1^2+\cdots+x_d^2}{2}} \frac{dx_1 \cdots dx_d}{\sqrt{2\pi} \cdots \sqrt{2\pi}} \\
&= \prod_{k=1}^{d} \underbrace{\frac{e^{-\frac{x_k^2}{2}}}{\sqrt{2\pi}}}_{\mathcal{N}(0;1)\text{ density}} \, dx_1 \cdots dx_d
\end{aligned}
$$

- However, if $\mathbb{P}_X(dx) = f(x_1, \ldots, x_d) dx_1 \cdots dx_d$, how to compute in general

$$\mathbb{E}\, \psi(X) = \int_{\mathbb{R}^d} \psi(x) f(x_1, \ldots, x_d) dx_1 \cdots dx_d.$$

even if ψ and the density f are explicit, e.g. if $d \geqslant 3$ or 4 ?