UNIVERSITÉ PARIS

MIDO - Master 1 Mathématiques & Applications, 2025–2026

Monte Carlo Methods

Final Exam - 06/01/2026

2H00 – DOCUMENTS AND CALCULATOR PROHIBITED

Part I	Part II	Total
...... / 10	...... / 20	...... / 30

TO BE FOLDED

LASTNAME : ....

FIRST NAME : ....

Most questions are independent. Answers need to be written on this document. There is sufficient space allotted to each question to accommodate a proper answer. If needed, add extra pages in the booklet.

PART 1: Advanced Monte Carlo methods

...... /**10**

This part covers advanced Monte Carlo estimators and variance reduction techniques (antithetic variates, control variates, stratified sampling). These tools are exploited in Part 2 to analyse Multilevel Monte Carlo (MLMC) estimators.

Question 1: Antithetic variates (2 pts)

Let $U \sim \mathcal{U}(0,1)$ and $\mathcal{I} = \mathbb{E}[g(U)]$. We consider the *standard* Monte Carlo estimator and its *antithetic* version, i.e.,

$$\hat{I}_N^{\mathrm{MC}} = \frac{1}{N}\sum_{i=1}^{N} g(U_i), \quad \text{and} \quad \hat{I}_N^{\mathrm{ant}} = \frac{1}{2N}\sum_{i=1}^{N}\Big(g(U_i) + g(1-U_i)\Big),$$

where the elements of $(U_i)_{i=1}^N$ are i.i.d. $\mathcal{U}(0,1)$.

(a) Provide a sufficient condition on g under which $\mathrm{Var}(\hat{I}_N^{\mathrm{ant}}) \leq \mathrm{Var}(\hat{I}_N^{\mathrm{MC}})$. (Justify your answer.)

(b) Let $g_1(x) = e^x$, $\quad g_2(x) = (x-0.5)^2$, $\quad g_3(x) = \sin(4\pi x)$. For each function g_j, $j = 1,2,3$, determine if antithetic sampling reduces the variance when compared to standard Monte Carlo. Justify briefly your answers.

...... /**2**

TO BE FOLDED

Question 2: Control variates (2 pts)

Let $X \sim \mathcal{U}(0,1)$. We want to estimate $\mu = \mathbb{E}[X^3]$ and consider the control variate $h_0(X) = X^2$, with $\mathbb{E}[h_0(X)] = 1/3$.

(a) Define the control variate estimator to estimate μ using h_0, and show that it is unbiased for any choice of the control coefficient $\beta \in \mathbb{R}$.

(b) Compute $\beta^\star$, the optimal control coefficient that minimizes the variance of the control variate estimator.

...... /**2**

Question 3: Stratified sampling on $[0,1]$ with non-uniform intra-stratum variances (4 pts)

For $X \sim \mathcal{U}(0,1)$ and $g(x) = e^x$, the goal is to estimate

$$\mathcal{I} = \mathbb{E}[g(X)] = \int_0^1 e^x \, dx = e - 1.$$

Resorting to *stratified sampling* with K equal strata:

$$D_k = \left[\tfrac{k-1}{K}, \tfrac{k}{K}\right), \qquad k = 1, \ldots, K.$$

means, within each stratum D_k, sampling N_k i.i.d. points $X_{k,1}, \ldots, X_{k,N_k} \sim \mathcal{U}(D_k)$. The resulting quantities

$$Y_{k,j} = g(X_{k,j}) = e^{X_{k,j}}, \quad \bar{Y}_k = \frac{1}{N_k} \sum_{j=1}^{N_k} Y_{k,j}, \quad \mathcal{I}_{\text{str}} = \frac{1}{K} \sum_{k=1}^{K} \bar{Y}_k$$

deliver the stratified estimator as $\mathcal{I}_{\text{str}}$.

(a) For each stratum D_k, compute explicitly

$$\mu_k = \mathbb{E}[Y_{k,1} \mid X_{k,1} \in D_k], \qquad \sigma_k^2 = \mathrm{Var}(Y_{k,1} \mid X_{k,1} \in D_k)$$

in terms of K and k, and show that σ_k^2 *depends* on k (i.e. the intra-stratum variance varies across strata).

(b) Assuming independence *between and within* strata simulation, express the variance of the stratified estimator $\mathcal{I}_{\text{str}}$ in terms of (N_k) and (σ_k^2).

(c) Using the formula for the optimal allocation in stratified sampling, give the explicit expression of the optimal sample size N_k in stratum k as a function of N, k, and K.

...... /4

(d) Write an R function delivering the standard Monte Carlo estimator approximating $\mathscr{I}$. This function should be called `MC(N)`, take as input the total number of samples N, and return both the estimate and an estimate of its variance.

(e) Write an R function implementing the stratified sampling estimator with proportional allocations that estimates $\mathscr{I}$. This function should be called `Stratified_prop(N,K)`, take as input the total number of samples N and the number of strata K, and return the estimate and an estimate of its variance.

...... /2

PART 2: Multilevel Monte Carlo (MLMC)

...... /**20**

This part studies why *Multilevel Monte Carlo* (MLMC) can be much more efficient than plain Monte Carlo by using numerical approximations (grids, time steps, etc.). The goal remains to approximate $\mathscr{I} = \mathbb{E}[Y]$ when simulating Y is costly or impossible. Instead, one produces Z_ℓ at increasing complexity *levels*, $\ell = 0, 1, 2, \ldots, L$.

- In **standard Monte Carlo**, the level L is set and N i.i.d. copies of Z_L are simulated and averaged. The root mean squared error (RMSE) is of order ε when
  - the approximation level L is of order ε,
  - $N = O(\varepsilon^{-2})$ samples control the *Monte Carlo variance*.

  However, the level L may be such that sampling is too expensive (e.g., of order 2^L or ε^3).
- **MLMC** avoids this cost explosion by combining cheap simulations at cruder levels with only a few expensive simulations at the finer levels, exploiting
  - a *telescoping identity* that writes the finest-level expectation as a sum of differences between consecutive levels,
  - the fact that the variance of these differences $Z_\ell - Z_{\ell-1}$ typically decreases very fast with ℓ.

  This allows MLMC to reach the same accuracy at a much lower cost.

Asymptotic notations

- For functions $f(\varepsilon)$ and $g(\varepsilon)$, we denote $f(\varepsilon) = O\big(g(\varepsilon)\big)$ when $\varepsilon \approx 0$ if there exist constants $C > 0$ and $\varepsilon_0 > 0$ such that $|f(\varepsilon)| \le C|g(\varepsilon)| \quad$ for all $0 < \varepsilon \le \varepsilon_0$.
- For sequences (a_ℓ) and (b_ℓ), we denote $a_\ell \asymp b_\ell$ if $a_\ell = O\big(b_\ell\big)$ and $b_\ell = O\big(a_\ell\big)$.

Question 1: Standard Monte Carlo (2 pts)

Let $X \sim f(x)$ be square-integrable with $\mathbb{E}[X] = \mu$. Define the Monte Carlo estimator

$$\hat{\mu}_N = \frac{1}{N}\sum_{i=1}^{N} X_i, \quad X_i \text{ i.i.d. } f$$

(a) Show that $\mathrm{Var}(\hat{\mu}_N) = \mathrm{Var}(X)/N$ and that in order to achieve $\mathrm{MSE}(\hat{\mu}_N) = \mathbb{E}[\hat{\mu}_N - \mu)^2] = O(\varepsilon^2)$, standard Monte Carlo requires $N = O(\varepsilon^{-2})$ samples. (1.5 pt)

(b) If the cost per sample is C, derive the total computational cost. (0.5 pt)

...... /**2**

Question 2: Building a deterministic quadrature hierarchy (2 pts)

For $p \geq 1$, define

$$Q_\ell = \frac{1}{2^\ell} \sum_{k=1}^{2^\ell} \left(\frac{k}{2^\ell}\right)^p, \qquad \ell \geq 0.$$

(a) Compute $\mathscr{I} = \int_0^1 x^p \, dx$. (0.5 pt)

(b) Show that $Q_\ell - \mathscr{I} = O(2^{-\ell})$. (0.5 pt)

(c) Using Faulhaber's formula

$$\sum_{k=1}^{N} k^p = \frac{N^{p+1}}{p+1} + \frac{N^p}{2} + O(N^{p-1}),$$

show that $|Q_\ell - Q_{\ell-1}| = O(2^{-\ell})$ and conclude that (Q_ℓ) is a Cauchy sequence. (1 pt)

...... /**2**

Question 3: Telescoping identity (1 pt)

Show that, for L^1 random variables (Y_ℓ) ,

$$\mathbb{E}[Y_L] = \mathbb{E}[Y_0] + \sum_{\ell=1}^{L} \mathbb{E}[Y_\ell - Y_{\ell-1}].$$

...... /**1**

Question 4: Monte Carlo version of the telescoping sum (3 pts)

Let $(Z_\ell)_{\ell=0}^L$ be random variables such that $\mathbb{E}[Z_\ell] = Q_\ell$. Define the multilevel estimator

$$\hat{Y}_L = \frac{1}{N_0}\sum_{i=1}^{N_0} Z_0^{(i)} + \sum_{\ell=1}^{L} \frac{1}{N_\ell}\sum_{i=1}^{N_\ell}\big(Z_\ell^{(i)} - Z_{\ell-1}^{(i)}\big).$$

with $(N_\ell)_\ell$ a non-increasing sequence. Assuming $\mathrm{Var}(Z_\ell - Z_{\ell-1}) = O(2^{-2\ell})$

(a) Show that $\mathbb{E}[\hat{Y}_L] = Q_L$. (0.5 pt)

(b) Show that, if $N_\ell \to \infty$ for all $\ell = 0, \ldots, L$, then $\hat{Y}_L \to Q_L$ in L^2 and deduce that the convergence also holds in L^1. (1 pt)

(c) Discuss how the choice of (N_ℓ) impact variance and convergence rate. (0.5 pt)

(d) Compute $\mathrm{Var}(\hat{Y}_L)$ when samples are coupled, i.e., the same $Z_\ell^{(i)}$ are used in $(Z_\ell^{(i)} - Z_{\ell-1}^{(i)})$ and $(Z_{\ell+1}^{(i)} - Z_\ell^{(i)})$, and identify cross-level covariances. (.5 pt)

(e) Recompute $\mathrm{Var}(\hat{Y}_L)$ when level terms are independent and compare. (0.5 pt)

...... /**3**

Question 5: Variance and cost analysis (3 pts)

Define $D_\ell^{(i)} = Z_\ell^{(i)} - Z_{\ell-1}^{(i)}$, $v_\ell = \mathrm{Var}(D_\ell)$, c_ℓ cost per sample, and covariance $\gamma_{\ell,m} = \mathrm{Cov}(D_\ell, D_m)$, assuming further that $\gamma_{\ell,m} = 0$ for $|\ell - m| > 1$.

(a) Show that (1 pt)

$$\mathrm{Var}(\hat{Y}_L) = \sum_{\ell=0}^{L} \frac{v_\ell}{N_\ell} + 2\sum_{\ell=0}^{L-1} \frac{\gamma_{\ell,\ell+1}}{\max(N_\ell, N_{\ell+1})}.$$

(b) In the independent-level case, i.e. when $\gamma_{\ell,m} = 0$ for all $\ell \neq m$, using the Lagrange multiplier method, show that the optimal allocation $(N_0^\star, N_1^\star, \ldots, N_L^\star)$ that minimises $\sum_{\ell=0}^{L} N_\ell c_\ell$ under the constraint $\mathrm{Var}(\hat{Y}_L) \leq V_0$ satisfies

$$N_\ell^\star = \sqrt{v_\ell / c_\ell}\left(\frac{1}{V_0}\sum_{m=0}^{L}\sqrt{v_m c_m}\right).$$

Deduce that the minimal cost is

$$C_{\min} = {}^1/_{V_0}\left(\sum_{\ell=0}^{L}\sqrt{v_\ell c_\ell}\right)^2.$$

...... /**3**

Question 6: Optimal MLMC complexity (3 pts)

Assume the independent-level variance formula from Question 5, and toy scalings $v_\ell \asymp 2^{-2\ell}$ and $c_\ell \asymp 2^\ell$.

(a) Using the expression for $C_{\min}$ from Question 5, show that to achieve $\mathrm{Var}(\hat{Y}_L) \le V_0$ with $V_0 \asymp \varepsilon^2$, the total MLMC cost satisfies $C_{\mathrm{MLMC}} \asymp \varepsilon^{-2}$. (1 pt)

(b) Consider now a standard Monte Carlo method at the finest level L, chosen so that both the bias and the Monte Carlo variance are of order $O(\varepsilon^2)$. Assuming $|\mathbb{E}[Z_L] - \mathscr{I}| \asymp 2^{-L}$ and $c_L \asymp 2^L$, compute the resulting level cost C_{SL} as a function of ε, and compare it to C_{MLMC}. (2 pts)

...... /**3**

Question 7: Rigorous MLMC error bound (2.5 pts)

Assume the bias satisfies $|\mathbb{E}[Y_L] - \mathscr{I}| \le B\,2^{-L}$ for some constant $B > 0$.

(a) Show that choosing $L = \lceil \log_2(B/\varepsilon) \rceil$ ensures $|\mathbb{E}[Y_L] - \mathscr{I}| \le \varepsilon$. (0.5 pt)

(b) Using the independent-level variance expression from Question 5 and the scaling assumptions $\mathrm{Var}(Z_\ell - Z_{\ell-1}) \asymp 2^{-2\ell}$, derive an upper bound for $\mathrm{Var}(\hat{Y}_L)$ in terms of the sequence $(N_\ell)_{\ell=0}^{L}$. (1 pt)

(c) Show that choosing $(\ell = 0, \dots, L)$ $N_\ell \asymp 2^{-\ell}\,\varepsilon^{-2}$ yields $\mathrm{Var}(\hat{Y}_L) = O(\varepsilon^2)$. (0.5 pt)

(d) Combine the bias and variance bounds to conclude that $\mathrm{MSE}(\hat{Y}_L) = \mathbb{E}\big[(\hat{Y}_L - \mathscr{I})^2\big] = O(\varepsilon^2)$. (0.5 pt)

...... /**2.5**

Question 8: Debiasing Monte Carlo estimators (3.5 pts)

Let $(Y_\ell)_{\ell\geq 0}$ be biased estimators of μ with $Y_0 = 0$ and $\mathbb{E}[Y_\ell] \to \mu$ as $\ell \to \infty$.

(a) Justify the telescoping identity (0.5 pt)

$$\mu = \sum_{\ell=1}^{\infty} \mathbb{E}[Y_\ell - Y_{\ell-1}].$$

(b) Under suitable conditions, establish unbiasedness for the single-term estimator $\hat{Y} = {}^{Y_L - Y_{L-1}}/_{p_L}$, when L is a random variable on $\mathbb{N}^*$ with probability mass function $(p_\ell)_{\ell\geq 1}$. (1 pt)

(c) Assume $\mathrm{Var}(Y_\ell - Y_{\ell-1}) = O(2^{-2\ell})$ and $p_\ell \propto 2^{-\alpha\ell}$. Determine the values of α for which $\mathrm{Var}(\hat{Y}) < \infty$. (1 pt)

(d) If the cost of computing Y_ℓ is $O(2^\ell)$, compute the expected cost of $\hat{Y}$ in terms of α, and discuss the variance–cost trade-off when choosing α. Compare qualitatively this debiasing approach with the MLMC strategy (in particular in terms of asymptotic cost for a given RMSE). (1 pt)

...... /**3.5**