Stratified Monte Carlo Integration

Rami El Haddad, Rana Fakhereddine, Christian Lécot

Université Saint-Joseph, Département de Mathématiques,
BP 11-514 Riad El Solh, Beyrouth 1107 2050, Liban
&
Université de Savoie, Laboratoire de Mathématiques, UMR 5127,
Campus scientifique, 73376 Le Bourget-du-Lac Cedex, France

Eighth IMACS Seminar on Monte Carlo Methods,
August 29 – September 2, 2011, Borovets, Bulgaria

Plan of the talk

Numerical integration

- $I := [0, 1)$,
- for $s \geq 1$, λ_s is the s-dimensional Lebesgue measure,
- $g : I^s \to \mathbb{R}$ is square-integrable.

We want to approximate

$$\mathcal{J} := \int_{I^s} g(x) d\lambda_s(x).$$

Monte Carlo approximation

- $\{U_1, \ldots, U_N\}$ i.i.d. random variables uniformly distributed over I^s,

$$X := \frac{1}{N}\sum_{\ell=1}^{N} g \circ U_\ell.$$

$$E[X] = \mathcal{J} \quad \text{and} \quad \mathrm{Var}(X) = \frac{\sigma^2(g)}{N},$$
where
$$\sigma^2(g) := \int_{I^s} \big(g(x)\big)^2 d\lambda_s(x) - \Big(\int_{I^s} g(x) d\lambda_s(x)\Big)^2.$$

Stratified sampling

- $\{D_1, \ldots, D_p\}$ a partition of I^s,
- $N_1, \ldots, N_p$ integers,
- $\{U_1^k, \ldots, U_{N_k}^k\}$ i.i.d. random variables uniformly distributed over D_k.

$$T_k := \frac{1}{N_k} \sum_{\ell=1}^{N_k} g \circ U_\ell^k, \quad T := \sum_{k=1}^{p} \lambda_s(D_k) T_k.$$

$$E[T] = \mathcal{J} \quad \text{and} \quad \mathrm{Var}(T) \leq \mathrm{Var}(X)^{1\,2}.$$

1. J.M. Hammersley, D.C. Handscomb, Monte Carlo Methods, Methuen, London (1964)
2. M. Evans, T. Swartz, Approximating Integrals via Monte Carlo and Deterministic Methods, Oxford University Press (2000)

Simple stratified sampling

- $\{D_1, \ldots, D_N\}$ a partition of I^s,

$$\lambda_s(D_1) = \cdots = \lambda_s(D_N) = \frac{1}{N},$$

- $\{V_1, \ldots, V_N\}$ independent random variables,
- V_ℓ uniformly distributed over D_ℓ.

$$Y := \frac{1}{N} \sum_{\ell=1}^{N} g \circ V_\ell.$$

$$E[Y] = \mathcal{J} \quad \text{and} \quad \mathrm{Var}(Y) \leq \mathrm{Var}(X)$$ [3] [4].

3. S. Haber, A modified Monte Carlo method, Math. Comput. 20, 361–368 (1966)

4. R.C.H. Cheng, T. Davenport, The problem of dimensionality in stratified sampling, Manage. Sci. 35, 1278–1296 (1989)

Monte Carlo approximation

- $A \subset I^s, \quad g := 1_A, \quad \mathcal{J} = \lambda_s(A),$
- $\{U_1, \dots, U_N\}$ i.i.d. random variables uniformly distributed over I^s,

$$X := \frac{1}{N} \sum_{\ell=1}^{N} 1_A \circ U_\ell.$$

$$\mathrm{Var}(X) = \frac{1}{N} \lambda_s(A)\big(1 - \lambda_s(A)\big) \leq \frac{1}{4N}.$$

ε-collars

$$\|x\|_\infty := \max_{1 \le i \le s} |x_i|.$$

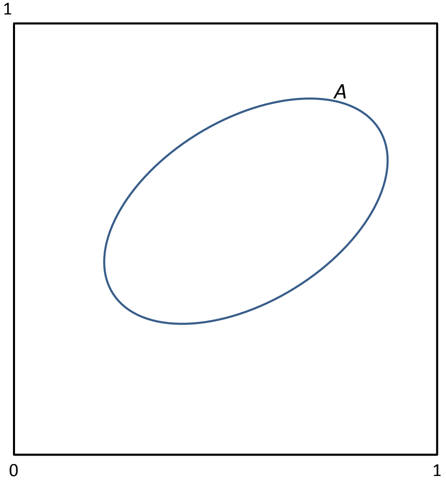
For $\varepsilon > 0$,

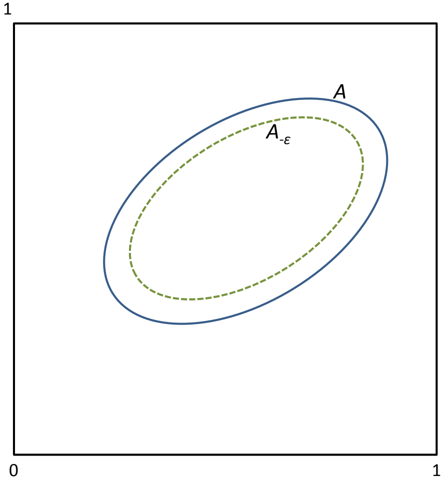
$$\begin{aligned} A_{-\varepsilon} &:= \{x \in I^s : \forall y \in I^s \setminus A \ \|x - y\|_\infty \ge \varepsilon\}, \\ A_\varepsilon &:= \{x \in I^s : \exists y \in A \ \|x - y\|_\infty < \varepsilon\}. \end{aligned}$$

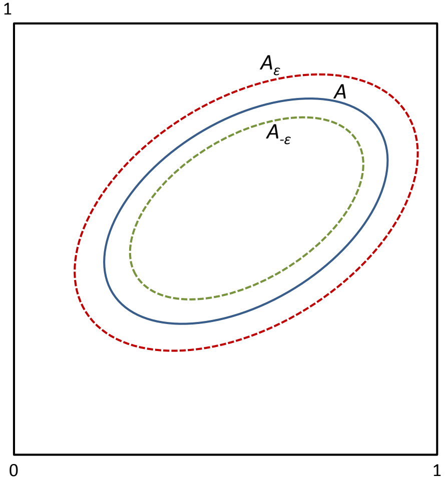
$$A_{-\varepsilon} \subset A \subset A_\varepsilon.$$

- Suppose there exists a nondecreasing $\gamma : [0, +\infty) \to [0, +\infty)$ such that

$$\forall \varepsilon > 0 \quad \max\left(\lambda_s(A_\varepsilon \setminus A), \lambda_s(A \setminus A_{-\varepsilon})\right) \le \gamma(\varepsilon),$$


1
A
0
1


1
A
A-ε
0
1


1
A_ε
A
$A_{-\varepsilon}$
0
1

Simple stratified sampling : Jordan measurable set

- $N = n^s$; for $k = (k_1, \dots, k_s)$ with $1 \le k_i \le n$,

$$C_k := \prod_{i=1}^{s} \left[\frac{k_i - 1}{n}, \frac{k_i}{n} \right),$$

- $\{V_k : 1 \le k_i \le n\}$ independent random variables,
- V_k uniformly distributed over C_k.

$$Y := \frac{1}{N} \sum_k 1_A \circ V_k.$$

Proposition 1. If

$$\forall \varepsilon > 0 \quad \max\left(\lambda_s(A_\varepsilon \setminus A), \lambda_s(A \setminus A_{-\varepsilon})\right) \le \gamma(\varepsilon),$$

then

$$\mathrm{Var}(Y) \le \frac{1}{2N} \gamma\left(\frac{1}{N^{1/s}}\right).$$

Simple stratified sampling : Jordan measurable set

Proof. We have

$$\mathrm{Var}(Y) \leq \frac{1}{4N^2} \#\{k : C_k \cap A \neq \emptyset \text{ and } C_k \not\subset A\}.$$

Since

$$\bigcup_{C_k \cap A \neq \emptyset \text{ and } C_k \not\subset A} C_k \subset A_{1/n} \setminus A_{-1/n},$$

we have

$$\frac{1}{N} \#\{k : C_k \cap A \neq \emptyset \text{ and } C_k \not\subset A\} \leq 2\gamma\left(\frac{1}{n}\right),$$

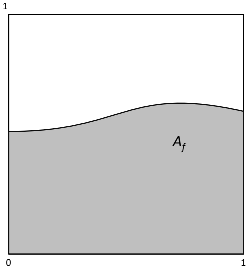
hence the result.

- for linear γ we obtain

$$\mathrm{Var}(Y) \leq \mathcal{O}\left(\frac{1}{N^{1+1/s}}\right).$$

Under the hypersurface

- $f : \bar{I}^{s-1} \to \bar{I}$ and $A_f := \{(x', x_s) \in I^s : x_s < f(x')\}$.



We want to approximate

$$\mathcal{I} := \int_{I^{s-1}} f(x') d\lambda_{s-1}(x') = \int_{I^s} 1_{A_f}(x) d\lambda_s(x).$$

Simple stratified sampling : under the hypersurface

- $N = n^s$; for $k = (k_1, \ldots, k_s)$ with $1 \leq k_i \leq n$,

$$C_k := \prod_{i=1}^{s} \left[\frac{k_i - 1}{n}, \frac{k_i}{n}\right),$$

- $\{V_k : 1 \leq k_i \leq n\}$ independent random variables,
- V_k uniformly distributed over C_k.

$$Y := \frac{1}{N} \sum_k 1_{A_f} \circ V_k.$$

Proposition 2. If f is of bounded variation $V(f)$ on $\bar{I}^s$, then

$$\mathrm{Var}(Y) \leq \left(\frac{s-1}{4} V(f) + \frac{1}{2}\right) \frac{1}{N^{1+1/s}}.$$

Simple stratified sampling : under the hypersurface

Proof. For $k = (k_1, \dots, k_s)$ with $1 \leq k_i \leq n$,

$$k' = (k_1, \dots, k_{s-1}) \quad \text{and} \quad C'_{k'} := \prod_{i=1}^{s-1} \left[\frac{k_i - 1}{n}, \frac{k_i}{n} \right).$$

We have :

$$\mathrm{Var}(Y) \leq \frac{1}{4N^2} \sum_{k'} \#\{k_s : C_{(k',k_s)} \cap A_f \neq \emptyset \text{ and } C_{(k',k_s)} \not\subset A_f\}$$

and

- if $C_{(k',k_s)} \cap A_f \neq \emptyset$ then $\exists x'_{k'} \in C'_{k'}$ such that $k_s < nf(x'_{k'}) + 1$,
- if $C_{(k',k_s)} \not\subset A_f$ then $\exists y'_{k'} \in C'_{k'}$ such that $nf(y'_{k'}) < k_s$.

Hence

$$\mathrm{Var}(Y) \leq \frac{1}{4N^2} \sum_{k'} \left(n \left(f(x'_{k'}) - f(y'_{k'}) \right) + 2 \right).$$

The result follows from Lemma 1.

Variation

Lemma 1.[5] Let f be a function of bounded variation $V(f)$ on $\overline{I}^s$. Let $n_1, \ldots, n_s$ be integers. For $k = (k_1, \ldots, k_s)$ with $1 \leq k_i \leq n_i$, denote $C_k := \prod_{i=1}^{s} \left[\frac{k_i - 1}{n_i}, \frac{k_i}{n_i} \right)$ and $x_k, y_k \in \overline{C_k}$. Then

$$\sum_k |f(x_k) - f(y_k)| \leq V(f) \prod_{i=1}^{s} n_i \sum_{i=1}^{s} \frac{1}{n_i}.$$

5. C. Lécot, Error bounds for quasi-Monte Carlo integration with nets. Math. Comput. 65, 179–187 (1996)

1. Stratified sampling
2. Integration of indicator functions
3. Numerical results

Measure of the unit ball : error

$$Q := \{x \in I^s : \|x\|_2 < 1\} \quad \text{then} \quad \lambda_s(Q) = \frac{\pi^{s/2}}{2^s \Gamma(\frac{s}{2} + 1)}.$$

We compare MC, stratified MC and QMC (using Faure sequence).

- $s = 2 : N = 10^2, 20^2, 30^2, \ldots, 400^2 = 160\,000$ points,
- $s = 3 : N = 10^3, 20^3, 30^3, \ldots, 200^3 = 8\,000\,000$ points,
- $s = 4 : N = 10^4, 12^4, 14^4, \ldots, 40^4 = 2\,560\,000$ points.

Suppose error $= \mathcal{O}(N^{-\alpha})$,

TABLE: Order α of the error

dimension	MC	stratified MC	QMC
2	0.36	0.66	0.78
3	0.49	0.52	0.80
4	0.18	0.60	0.68

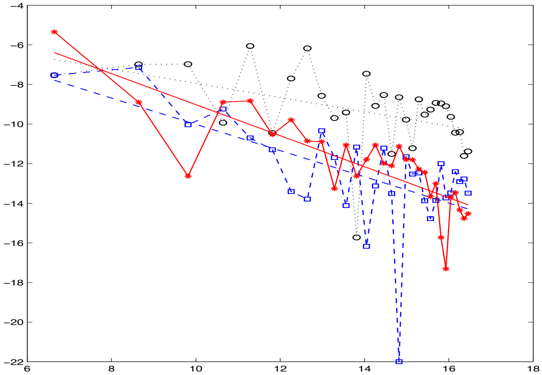


FIGURE: Error in the calculation of $\lambda_2(Q)$ and linear regression estimates of the error as a function of N (from 10^2 to 400^2). MC (black dotted lines), stratified MC (blue dashed lines) and QMC (red solid lines).

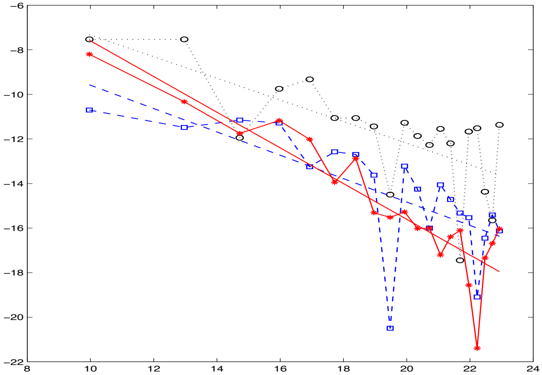


FIGURE: Error in the calculation of $\lambda_3(Q)$ and linear regression estimates of the error as a function of N (from 10^3 to 200^3). MC (black dotted lines), stratified MC (blue dashed lines) and QMC (red solid lines).

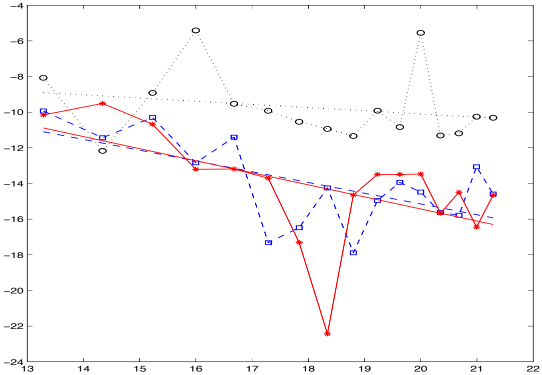


FIGURE: Error in the calculation of $\lambda_4(Q)$ and linear regression estimates of the error as a function of N (from 10^4 to 40^4). MC (black dotted lines), stratified MC (blue dashed lines) and QMC (red solid lines).

Measure of the unit ball : variance

- We compare MC and stratified MC,
- in order to estimate the variance, we replicate the quadrature independently M times and compute the sample variance,
- we use $M = 100, 200, \ldots, 1\,000$.

Suppose $\mathrm{Var} = \mathcal{O}(N^{-\beta})$,

TABLE: Order β of the variance

dimension	MC	stratified MC	theoretical bound
2	1.00	1.49	1.50
3	0.99	1.34	1.33
4	1.00	1.26	1.25

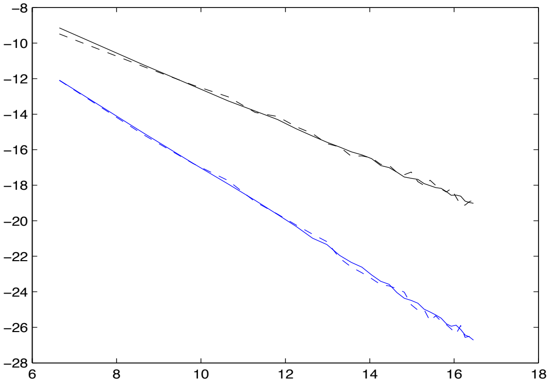


FIGURE: Sample variance of M independent copies of the approximation of $\lambda_2(Q)$ as a function of N (from 10^2 to 400^2). MC (black) and stratified MC (blue) ; $M = 100$ (dashed lines) and $M = 1\,000$ (solid lines).

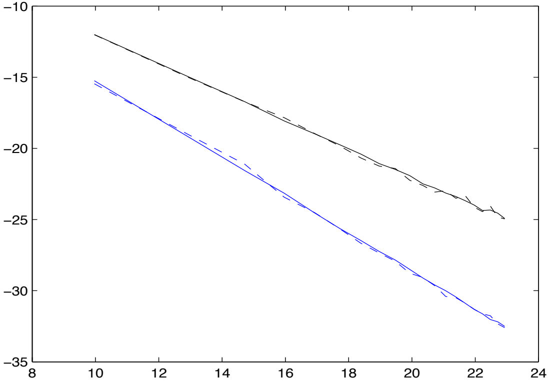


FIGURE: Sample variance of M independent copies of the approximation of $\lambda_3(Q)$ as a function of N (from 10^3 to 200^3). MC (black) and stratified MC (blue) ; $M = 100$ (dashed lines) and $M = 1\,000$ (solid lines).

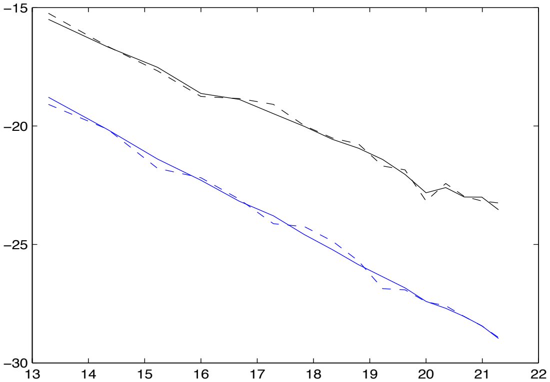


FIGURE: Sample variance of M independent copies of the approximation of $\lambda_4(Q)$ as a function of N (from 10^4 to 40^4). MC (black) and stratified MC (blue) ; $M = 100$ (dashed lines) and $M = 1\,000$ (solid lines).