

Université Assane Seck de Ziguinchor (U.A.S.Z)



U.F.R des Sciences et Technologies (URF-ST)

Département de Mathématiques

Mémoire de Master

Domaine : Sciences et Technologies
Mention : Mathématiques et Applications
Spécialité : Mathématiques Appliquées
Option : Statistique

TITRE :

Quasi-Monte-Carlo : un intermédiaire entre les méthodes déterministes et les méthodes Monte-Carlo

Présenté par

Bernard MANGA

Sous la direction de : Dr Emmanuel Nicolas CABRAL

Sous la supervision de : Pr Alassane DIEDHIOU

Membres du Jury :

PRESIDENT :

Alassane DIEDHIOU, Professeur (UFR-ST/U.A.S.Z)

DIRECTEUR :

Emmanuel N. CABRAL, Maître Assistant (UFR-ST/U.A.S.Z)

EXAMINATEURS :

Mouhamadou Samsidy GOUDIABY, Maître Assistant (UFR-ST/U.A.S.Z)

Clément MANGA, Maître Assistant (UFR-ST/U.A.S.Z)

Soutenu le 11 février 2021

Dédicaces

À mes très chers parents,

À Charlot MANGA,

À Combé MANGA,

À Zackarie MANGA,

À Landing SAGNA,

À Djihemo SAMBOU,

À Mariama SAMBOU,

À toute la promotion de M.P.I,

À tous mes sœurs et frères en ALLAH,

Remerciements

Je tiens à remercier le Tout Puissant pour toute la volonté et le courage qu'Il m'a donnés pour l'achèvement de ce travail.

La réalisation de ce mémoire a été possible grâce au concours de plusieurs personnes à qui je voudrais témoigner toute ma reconnaissance.

Je voudrais tout d'abord exprimer ma profonde reconnaissance à Mr Emmanuel Nicolas CABRAL pour avoir accepté de diriger ce travail. Je tiens à le remercier de m'avoir accordé sa confiance en faisant preuve d'une très grande disponibilité à mon égard. Ses conseils, encouragements et remarques minutieuses ont été déterminants et je réalise aujourd'hui à quel point travailler avec lui a été enrichissant.

Je tiens à remercier l'Université Assane SECK, l'UFR des Sciences et Technologies, le Département de Mathématiques ainsi que tous les professeurs qui nous ont initié et donné les bases de l'étude mathématique pour exceller dans ce domaine.

J'exprime ma gratitude envers le président du jury, le professeur Alassane DIEDHIOU, qui a accepté de présider ce jury et d'apporter sa contribution pour la présentation de ce mémoire. J'exprime aussi ma gratitude en vers les professeurs Mouhamadou Samsidy GOU-DIABY, Clément MANGA qui ont bien accepté de m'examiner et d'apporter leurs expertises.

Mes plus vifs remerciements à ma très chère mère, mon très cher père, à mes frères et mes sœurs et, en particulier, à mon frère Galbert MANGA qui m'a soutenu moralement et matériellement tout au long de mes études.

Mes remerciements vont également à l'endroit de mes frères doctorants Marcel Sihin-toé BADIANE et Pape Modou SARR et de mes promotionnaires pour leur soutien indéfectible.

Résumé

Dans le domaine Mathématiques le calcul analytique d'intégrations de certaines fonctions numériques est difficile voir impossible. Pour contourner cette difficulté, nous recourons à des méthodes d'approximations déterministes telles que la méthode des Rectangles, la méthode des Trapèzes, la méthode de Simpson, ... Mais ces méthodes sont très lentes en dimension élevée. Pour résoudre ce problème, nous utilisons une autre méthode (aléatoire), inconnue à la dimension, qui permet de trouver une valeur approximative et converge rapidement. Cette méthode est dite de Monte-Carlo. La méthode de Monte-Carlo qui utilise des suites aléatoires, est l'une des méthodes numériques les plus polyvalentes et les plus utilisées. Son taux de convergence, étant de $o(\frac{1}{\sqrt{n}})$, est indépendant de la dimension, ce qui montre que la méthode Monte-Carlo peut être très robuste, mais aussi lente. Dans ce mémoire nous présentons un rappel des méthodes susmentionnées pour les problèmes d'intégration et quelques techniques de réduction de la variance relative à Monte-Carlo. Par ailleurs l'accélération de la convergence pour la méthode de Monte-Carlo est atteinte en utilisant des suites à discrétion faible qui sont une alternative déterministe à des suites aléatoires ou pseudo-aléatoires. La méthode résultante, appelée Quasi-Monte-Carlo est un intermédiaire entre les méthodes déterministes et la méthode de Monte-Carlo. La méthode Quasi-Monte-Carlo est basée sur l'idée que les techniques de Monte-Carlo aléatoires peuvent souvent être améliorées par le remplacement de la source de nombres aléatoires avec une suite déterministe distribuée plus uniformément. Son taux de convergence est de l'ordre de $o(\frac{(\log n)^d}{n})$, d étant la dimension. Nous présentons également cette méthode pour l'approximation d'intégrale en utilisant des suites à discrétion faible telles que : Van Der Corput, Halton, Hammersley, Sobol et Faure. Nous terminons par une comparaison de ces différentes méthodes pour montrer l'idée selon laquelle la méthode Quasi-Monte-Carlo est un intermédiaire entre les méthodes déterministes et la méthode de Monte-Carlo.

Table des matières

Introduction générale	6
1 NOTIONS PRELIMINAIRES	7
1.1 Quelques outils probabilistes	7
1.1.1 Notions de vecteur aléatoire	7
1.1.2 Vecteur aléatoire discret	7
1.1.3 Vecteur aléatoire à densité	7
1.1.4 Cas du vecteur aléatoire gaussien	7
1.1.5 Loi conjointe et loi marginale d'un vecteur aléatoire	8
1.1.6 Espérance et matrice de variance covariance d'un vecteur aléatoire	8
1.2 Simulation de variable aléatoire	9
1.2.1 Méthode d'inversion	9
1.2.2 Méthode de "Box-Muller"	11
1.2.3 Méthode de rejet	12
1.3 Simulation d'un vecteur aléatoire	13
1.3.1 Vecteur aléatoire à loi jointe connue	13
1.3.2 Vecteur aléatoire à loi jointe inconnue	14
2 QUELQUES METHODES DETERMINISTES	15
2.1 Intégrale simple ($d = 1$)	15
2.1.1 Méthode des Rectangles	15
2.1.2 Méthode des Trapèzes	17
2.1.3 Méthode de Simpson	18
2.2 Intégration multiple ($d \geq 2$)	21
2.2.1 Intégration double	21
2.2.2 Intégration Triple	22
2.3 Conclusion	23
3 GENERALITES SUR LES METHODES DE MONTE-CARLO	24
3.1 Introduction	24
3.2 Simulation de Monte-Carlo : Calcul d'intégrale	24
3.2.1 Description de la méthode	24
3.2.2 Exemple d'un estimateur de π	25
3.3 Propriétés de l'estimateur \hat{I}_n	26
3.3.1 Qualités de l'estimateur	26
3.3.2 Exemple (Estimation de π)	27
3.3.3 Erreurs dues à la méthode de Monte-Carlo	27
3.4 Techniques de Réduction de la Variance	28
3.4.1 Variable de contrôle	28
3.4.2 Echantillonnage préférentiel	29
3.4.3 Variable antithétique	29
3.4.4 Stratification	29

3.5	Conclusion	30
4	LES METHODES QUASI-MONTE-CARLO (QMC)	31
4.1	Introduction	31
4.2	Principe de la méthode	31
4.3	Notion de discrédance	32
4.4	Les suites de QMC	33
4.4.1	Les suites de Van Der Corput	33
4.4.2	Les suites de Halton	34
4.4.3	Les suites de Hammersley	36
4.4.4	Les suites de Sobol	36
4.4.5	Les suites de Faure	39
4.5	Conclusion	40
5	MISE EN OEUVRE PRATIQUE ET COMPARAISON	42
5.1	Introduction	42
5.2	Comparaison entre les méthodes déterministes et la méthode de Monte-Carlo	42
5.2.1	Erreur d'approximation en dimension 1	43
5.2.2	Erreur d'approximation en dimension 2	45
5.2.3	Erreur d'approximation en dimension 3	48
5.3	Erreur d'approximation par la méthode QMC	49
5.4	Comparaison entre les méthodes QMC et Monte-Carlo en dimension d	50
5.5	Conclusion	53
	Conclusion générale	54
	Bibliographie	55

Introduction générale

La résolution de nombreux problèmes scientifiques nécessite de calculer des sommes, des intégrales, ou encore de résoudre des équations ou des problèmes d'optimisation. Les techniques de calculs directs, encore appelé techniques analytiques, sont très vite dépassées par la complexité des modèles, elles nécessitent souvent des hypothèses trop fortes, de sorte qu'on ne puisse pas les appliquer, ou alors comme dans le cas de calcul de sommes, le nombre d'opérations requis peut être trop important pour être réalisé en un temps raisonnable. On doit alors nécessairement faire appel à des méthodes d'approximation. Cependant, celles-ci requièrent également des hypothèses fortes, bien que moins fortes que pour les méthodes analytiques.

Dans notre cas, nous nous intéressons à l'approximation des quantités du type

$$I = \int_C \varphi(x)f(x)dx \tag{0.0.1}$$

où $f, \varphi : C \subset \mathbb{R}^d \rightarrow \mathbb{R}$, avec $d \in \mathbb{N}^*$ et I une intégrale dont le calcul analytique s'avère difficile voir impossible.

Parmi les méthodes d'approximation figurent les méthodes déterministes (telles que la méthode des Rectangles, la méthode des Trapèzes et méthode de Simpson), la méthode de Monte-Carlo et celle de Quasi-Monte-Carlo.

Dans la suite de notre travail, après avoir introduit quelques notions de probabilités au premier chapitre, nous présenterons quelques méthodes déterministes (deuxième chapitre). Nous parlerons ensuite du principe de la méthode de Monte-Carlo (troisième chapitre) et des méthodes Quasi-Monte-Carlo (quatrième chapitre). Nous terminerons au cinquième chapitre par une mise en oeuvre pratique et une comparaison de ces différentes méthodes.

Chapitre 1

NOTIONS PRELIMINAIRES

1.1 Quelques outils probabilistes

1.1.1 Notions de vecteur aléatoire

Definition 1.1.1 Soit $(\Omega, \mathcal{A}, \mathbb{P})$ un espace probabilisé. On appelle variable aléatoire réelle (v.a.r) toute application $X : \Omega \rightarrow \mathbb{R}$ telle que pour tout intervalle I de \mathbb{R} , $X^{-1}(I) \in \mathcal{A}$.

Definition 1.1.2 On appelle vecteur aléatoire de dimension n un n -uplet (X_1, \dots, X_n) où chaque X_i est une variable aléatoire réelle définie sur le même espace.

1.1.2 Vecteur aléatoire discret

Definition 1.1.3 Soit $X = (X_1, \dots, X_n)$ un vecteur aléatoire. On dit que X est un vecteur aléatoire discret s'il prend ses valeurs dans un sous-ensemble dénombrable de \mathbb{R}^n c'est-à-dire les X_1, \dots, X_n sont des variables aléatoires discrètes.

1.1.3 Vecteur aléatoire à densité

Definition 1.1.4 Un vecteur aléatoire $X = (X_1, \dots, X_n)$ est à densité s'il existe une fonction positive Lebesgue-intégrable $f_X : \mathbb{R}^n \rightarrow \mathbb{R}$ telle que $\mathbb{P}(X \in A) = \int_A f_X(x_1, \dots, x_n) dx_1 \cdots dx_n, \forall A \in \mathcal{B}(\mathbb{R}^n)$.

1.1.4 Cas du vecteur aléatoire gaussien

Nous allons voir à présent un exemple particulièrement important de vecteur aléatoire. Si $x, y \in \mathbb{R}^n$, on note leur produit scalaire $\langle x, y \rangle$.

Definition 1.1.5 Un vecteur aléatoire $X = (X_1, \dots, X_n) : \Omega \rightarrow \mathbb{R}^n$ est un vecteur aléatoire gaussien si les variables aléatoires $\langle a, X \rangle$ suivent des lois normales, pour tout $a \in \mathbb{R}^n$.

Propriété 1.1.1 1. Si $X = (X_1, \dots, X_n)$ est un vecteur aléatoire gaussien alors les X_i sont des variables aléatoires gaussiennes pour chaque $i = 1, \dots, n$.

2. Si $A : \mathbb{R}^n \rightarrow \mathbb{R}^n$ est une application linéaire, le vecteur $AX : \Omega \rightarrow \mathbb{R}^n$ est un vecteur gaussien.

Remarque 1.1.1 La réciproque de la première propriété est fausse.

1.1.5 Loi conjointe et loi marginale d'un vecteur aléatoire

Comme pour les variables aléatoires, un vecteur aléatoire induit naturellement une mesure de probabilité sur $(\mathbb{R}^n, \mathcal{B}(\mathbb{R}^n))$.

Definition 1.1.6 Soit $X = (X_1, \dots, X_n)$ un vecteur aléatoire sur un espace probabilisé $(\Omega, \mathcal{A}, \mathbb{P})$. On appelle fonction de répartition conjointe de X la fonction $F_X : \mathbb{R}^n \rightarrow [0, 1]$ définie par

$$F_X(x) = \mathbb{P}(X \leq x), \forall x \in \mathbb{R}^n.$$

La densité f de X existe si F_X est n fois différentiable et on a

$$f(x_1, x_2, \dots, x_n) = \frac{\partial^n F_X(x_1, x_2, \dots, x_n)}{\partial x_1 \partial x_2 \cdots \partial x_n}.$$

Propriété 1.1.2 Les fonctions de répartition conjointes possèdent des propriétés tout à fait analogues à celles des fonctions de répartition. La fonction de répartition conjointe F_X d'un vecteur aléatoire X satisfait :

1. $\lim_{x_i \rightarrow -\infty} F_X(x_i) = 0, \lim_{x_i \rightarrow +\infty} F_X(x_i) = 1 \quad \forall i = 1, \dots, n;$
2. si $x \leq y$ alors $F_X(x) \leq F_X(y)$.

Definition 1.1.7 Soit $X = (X_1, \dots, X_n)$ un vecteur aléatoire. Alors les fonctions de répartition $F_{X_i}, i = 1, \dots, n$ sont appelées fonctions de répartition marginales de F_X . Il est possible de caractériser simplement l'indépendance de variables aléatoires en termes de leur fonction de répartition conjointe.

Lemme 1.1.1 La famille (X_1, \dots, X_n) de variables aléatoires est indépendante si et seulement si $F_{(X_1, \dots, X_n)}(x_1, \dots, x_n) = F_{X_1}(x_1) \times \dots \times F_{X_n}(x_n), \forall (x_1, \dots, x_n) \in \mathbb{R}^n$.

1.1.6 Espérance et matrice de variance covariance d'un vecteur aléatoire

Espérance d'un vecteur aléatoire

Soit $X = (X_1, \dots, X_n) \in \mathbb{R}^n$ un vecteur aléatoire (discret ou continu).

Definition 1.1.8 L'espérance d'un vecteur aléatoire de dimension n est le vecteur de \mathbb{R}^n constitué des espérances de chacune des coordonnées i.e. $\mathbb{E}(X) = (\mathbb{E}(X_1), \dots, \mathbb{E}(X_n))$.

Proposition 1.1.1 Soit $b \in \mathbb{R}^m$ et $A \in \mathcal{M}_{m,n}(\mathbb{R})$. L'espérance du vecteur aléatoire $Y = AX + b \in \mathbb{R}^m$ est $\mathbb{E}(Y) = A\mathbb{E}(X) + b$.

En particulier on a $\mathbb{E}(X_1 + X_2 + \dots + X_n) = \mathbb{E}(X_1) + \mathbb{E}(X_2) + \dots + \mathbb{E}(X_n)$.

Covariance et corrélation

Soit $X = (X_1, \dots, X_n) \in \mathbb{R}^n$ un vecteur aléatoire (discret ou continu).

La covariance de X_i et X_j est donnée par :

$$\begin{aligned} Cov(X_i, X_j) &= \mathbb{E}((X_i - \mathbb{E}(X_i))(X_j - \mathbb{E}(X_j))) \\ &= \mathbb{E}(X_i X_j) - \mathbb{E}(X_i)\mathbb{E}(X_j). \end{aligned}$$

Le coefficient de corrélation linéaire entre X_i et X_j est donné par

$$\rho(X_i, X_j) = \frac{Cov(X_i, X_j)}{\sqrt{Var(X_i)Var(X_j)}} \in [-1, 1].$$

Proposition 1.1.2 Soit $X = (X_1, \dots, X_n) \in \mathbb{R}^n$ un vecteur aléatoire (discret ou continu). Si X_i et X_j sont indépendantes alors $\text{Cov}(X_i, X_j) = \mathbb{E}(X_i X_j) - \mathbb{E}(X_i)\mathbb{E}(X_j) = 0$ car $\mathbb{E}(X_i X_j) = \mathbb{E}(X_i)\mathbb{E}(X_j)$.

Attention! La réciproque est en général fausse.

Matrice de variance covariance

Definition 1.1.9 Soit $X = (X_1, \dots, X_n)$ un vecteur aléatoire de dimension n . On appelle matrice de variance covariance de X , la matrice carrée de taille $n \times n$, noté Σ_X , dont les coefficients $(\Sigma_{ij})_{1 \leq i, j \leq n}$ sont donnés par $\Sigma_{ij} = \text{Cov}(X_i, X_j)$. En notation matricielle, la matrice de variance covariance d'un vecteur aléatoire est

$$\Sigma_X = \begin{pmatrix} \text{Var}(X_1) & \text{Cov}(X_1, X_2) & \cdots & \text{Cov}(X_1, X_n) \\ \text{Cov}(X_2, X_1) & \text{Var}(X_2) & \cdots & \cdots \\ \cdots & \cdots & \cdots & \cdots \\ \cdots & \cdots & \cdots & \cdots \\ \cdots & \cdots & \cdots & \cdots \\ \text{Cov}(X_n, X_1) & \cdots & \text{Cov}(X_n, X_{n-1}) & \text{Var}(X_n) \end{pmatrix}.$$

Remarque 1.1.2 Si les coordonnées de X sont indépendantes, alors Σ_X est une matrice diagonale.

Proposition 1.1.3 Soit $X = (X_1, \dots, X_n) \in \mathbb{R}^n$ un vecteur aléatoire. On suppose que chacune des X_i admet un moment d'ordre 2, alors on a

$$\text{Var}(X_1 + \cdots + X_n) = \text{Var}(X_1) + \cdots + \text{Var}(X_n) + 2 \sum_{1 \leq i < j \leq n} \text{Cov}(X_i, X_j).$$

Plus généralement, si A désigne la matrice de variance-covariance du vecteur aléatoire $X = (X_1, \dots, X_n)$, alors pour tout $\lambda = (\lambda_1, \dots, \lambda_n) \in \mathbb{R}^n$

$$\text{Var}(\lambda_1 X_1 + \cdots + \lambda_n X_n) = (\lambda_1, \dots, \lambda_n) A (\lambda_1, \dots, \lambda_n)^T.$$

1.2 Simulation de variable aléatoire

Definition 1.2.1 Soit X une variable aléatoire de fonction de répartition F . On appelle inverse généralisée de F ou fonction quantile de F la fonction F^* ou F^{-1} définie pour tout $u \in]0, 1[$ par :

$$F^{-1}(u) = \inf \left\{ x \in \mathbb{R} : F(x) \geq u \right\}.$$

De façon générale lorsque, $0 < \alpha < 1$, $q_{1-\alpha} = F^{-1}(1 - \alpha)$ est le quantile d'ordre $(1 - \alpha)$ de F .

Remarque 1.2.1 Si F est inversible la fonction quantile coïncide avec l'inverse classique.

1.2.1 Méthode d'inversion

La méthode de simulation par inversion repose sur le résultat suivant.

Proposition 1.2.1 Soient $U \rightsquigarrow \mathcal{U}_{[0,1]}$, F une fonction de répartition et F^{-1} son inverse généralisée alors :

1. la v.a $X = F^{-1}(U)$ a pour fonction de répartition F ;

2. si X a pour fonction de répartition F et si F est continue alors la v.a $F(X)$ suit une loi uniforme sur $[0, 1]$.

Preuve.

1. Soit $X = F^{-1}(U)$ avec $U \rightsquigarrow \mathcal{U}_{[0,1]}$ et $x \in \mathbb{R}$.
 Montrons que $\forall u \in [0, 1]$ on a $F^{-1}(u) \leq x \Leftrightarrow u \leq F(x)$
 \Leftarrow Par définition de $F^{-1}(u)$, si $u \leq F(x)$ alors $F^{-1}(u) \leq x$.
 \Rightarrow si $F^{-1}(u) \leq x$, alors $\forall \varepsilon > 0$ on a $F^{-1}(u) < x + \varepsilon$, donc par définition de $F^{-1}(u)$ $u \leq F(x + \varepsilon)$, F étant continue à droite. On déduit que $u \leq F(x)$
 Par ailleurs puisque $U \rightsquigarrow \mathcal{U}_{[0,1]}$, on a

$$\mathbb{P}(X \leq x) = \mathbb{P}(F^{-1}(U) \leq x) = \mathbb{P}(U \leq F(x)) = F(x).$$

2. On applique (1), la variable $Y = F^{-1}(U)$ a même loi que X , donc la v.a $F(X)$ a même loi que $F(Y) = F \circ F^{-1}(U)$, or d'après (1) pour tout $u \in [0, 1]$ on a

$$F^{-1}(u) \leq F^{-1}(u) \Rightarrow u \leq (F \circ F^{-1})(u).$$

Réciproquement, $\forall u \in [0, 1]$ et $\forall \varepsilon > 0$, on a d'après (1).

$$F^{-1}(u) - \varepsilon < F^{-1}(u) \Rightarrow F(F^{-1}(u) - \varepsilon) < u$$

Comme $u \in [0, 1]$ et que F est continue en $F^{-1}(u) > 0$ le passage à la limite quand $\varepsilon \rightarrow 0$ donne $F \circ F^{-1}(u) < u$. Ainsi, on a $\forall u \in [0, 1]$, $F \circ F^{-1}(u) = u$

En revenant à la v.a $F(Y)$, on a $F(Y) = F \circ F^{-1}(U) = U$ p.s.

■

La méthode d'inversion s'applique lorsqu'on connaît une forme explicite de la fonction de répartition F et sa fonction inverse généralisée F^{-1} . Son algorithme est le suivant :

1. Générer une v.a U uniforme dans $[0, 1]$
2. Retourner $X = F^{-1}(U)$

Prenons les deux exemples suivants.

1. La loi uniforme sur $[a, b]$

Soit X une variable aléatoire qui suit une loi uniforme sur $[a, b]$, sa fonction de répartition est $F(x) = \frac{x-a}{b-a}$ et son inverse généralisé est $F^{-1}(u) = (b-a)u + a$

Pour simuler cette variable aléatoire il suffit de générer des nombres aléatoires u d'une variable aléatoire U qui suit une loi uniforme sur $[0, 1]$ et on déduit les réalisations suivantes de X

$$x = (b-a)u + a.$$

2. La loi exponentielle

Soit X une v.a exponentielle, sa fonction de répartition est $F(x) = (1 - \exp(-\lambda x)) \mathbb{1}_{x>0}$.

La fonction inverse de F est : $F^{-1}(u) = -\lambda^{-1} \ln(1-u)$.

Remarque : Si la variable aléatoire U est uniforme sur $[0, 1]$, $(1-U)$ l'est aussi, on obtient une réalisation de X en posant $X = -\lambda^{-1} \ln(U)$.

1.2.2 Méthode de "Box-Muller"

Elle est basée sur le théorème suivant.

Théorème 1.2.1 (de Box-Muller) Soient X et Y deux variables aléatoires réelles indépendantes suivant chacune la loi $\mathcal{N}(0, 1)$. On établit les égalités en loi suivantes :

$$X \stackrel{\mathcal{L}}{=} \sqrt{-2 \ln(U)} \cos(2\pi V)$$

$$Y \stackrel{\mathcal{L}}{=} \sqrt{-2 \ln(U)} \sin(2\pi V).$$

où $U, V \text{ i.i.d. } \rightsquigarrow \mathcal{U}_{[0,1]}$

Preuve. Soient X et Y deux variables aléatoires réelles indépendantes suivant chacune la loi $\mathcal{N}(0, 1)$. En coordonnées polaires X et Y s'écrivent :

$$X = R \cos(A) \tag{1.2.1}$$

$$Y = R \sin(A), \tag{1.2.2}$$

où R (module) est une v.a ayant pour support \mathbb{R}_+ et A (angle) est une v.a ayant pour support $[0, 2\pi]$.

Notons que $R^2 = X^2 + Y^2 \rightsquigarrow \chi_{(2)}^2 = \varepsilon(\frac{1}{2})$ et H la f.d.r de R , on a :

$$H(r) = \mathbb{P}(R \leq r) = \mathbb{P}(R^2 \leq r^2) \text{ comme } R \rightsquigarrow \varepsilon(\frac{1}{2}) \text{ on a}$$

$$H(r) = 1 - e^{-\frac{r^2}{2}} \text{ d'où } H^{-1}(u) = \sqrt{-2 \ln(1-u)}.$$

On en déduit directement l'égalité en loi suivante (cf. à la Proposition 1.2.1)

$$R \stackrel{\mathcal{L}}{=} \sqrt{-2 \ln(U)} \quad \text{où} \quad U \rightsquigarrow \mathcal{U}_{[0,1]} \tag{1.2.3}$$

Considérons la transformation $(X, Y) \longrightarrow (R, A)$ et notons f la densité de (X, Y) , soit

$$f(x, y) = \frac{1}{2\pi} e^{-\frac{x^2+y^2}{2}}, \forall (x, y) \in \mathbb{R}^2.$$

D'après la formule de changement de variables, la densité du couple (R, A) est donnée par :

$$g(r, a) = f(r \cos(a), r \sin(a)) |J|,$$

où $|J|$ est le déterminant de la matrice Jacobienne donnée par :

$$|J| = \begin{vmatrix} \cos(a) & -r \sin(a) \\ \sin(a) & r \cos(a) \end{vmatrix} = r$$

$$g(r, a) = \frac{1}{2\pi} \left[r e^{-\frac{r^2}{2}} \right], \quad \forall r \geq 0, \quad a \in [0, 2\pi].$$

Le terme entre crochet n'est autre que la densité h de la v.a R . En effet, on vérifie que

$$h(r) = H'(r) = r e^{-\frac{r^2}{2}}, \quad \forall r \geq 0.$$

Déterminons la densité l de la v.a A .

$$l(a) = \int_0^{+\infty} g(r, a) dr = \frac{1}{2\pi}.$$

On en déduit que R et A sont indépendantes car $g(r, a) = l(a) \times h(r)$ et que $A \rightsquigarrow \mathcal{U}_{[0,2\pi]}$. La fonction de répartition L de A s'écrit

$$L(a) = \frac{1}{2\pi} \text{ et donc } a = L^{-1}(v) = 2\pi v.$$

Ce qui permet d'écrire l'égalité en loi

$$A \stackrel{\mathcal{L}}{=} 2\pi V \quad \text{où} \quad V \rightsquigarrow \mathcal{U}_{[0,1]}. \quad (1.2.4)$$

En confrontant les égalités (1.2.1), (1.2.2), (1.2.3) et (1.2.4) on obtient le résultat ■

Pour simuler les variables aléatoires X et Y on peut procéder comme suit :

1. générer 2 valeurs u et v d'une variable aléatoire $\mathcal{U}_{[0,1]}$,
2. calculer x et y ainsi

$$\begin{aligned} x &= \sqrt{-2 \ln(u)} \cos(2\pi v) \\ y &= \sqrt{-2 \ln(u)} \sin(2\pi v). \end{aligned}$$

1.2.3 Méthode de rejet

On veut simuler une variable de densité f (dans \mathbb{R}^d) et on suppose qu'il existe une loi de densité g facile à simuler telle que

$$\forall x \in \mathbb{R}^d, f(x) \leq kg(x). \quad (1.2.5)$$

avec k une constante. Posons $\alpha(x) = \frac{f(x)}{kg(x)}$ sur l'ensemble $\{x, g(x) > 0\}$.

Proposition 1.2.2 *Soit $(X_n, U_n)_{n \geq 1}$ une suite de variable i.i.d. avec pour tout n , X_n est de loi de densité g et est indépendante de U_n , qui est de loi uniforme sur $[0, 1]$. Soit $T = \inf\{n, U_n \leq \alpha(X_n)\}$ (T est un entier aléatoire). Soit $X = X_T$. Alors la v.a. X ainsi simulée est de loi de densité f .*

Preuve. Remarquons que $\alpha(X_n)$ est défini pour presque tout ω .

En effet, $\mathbb{P}(X_n \in \{x : g(x) = 0\}) = 0$. Soit $\varphi \in C_b^+(\mathbb{R}^d)$. Nous avons $\varphi(X_T) = \sum_{n=1}^{+\infty} \varphi(X_n) \mathbb{1}_{n=T}$ (pour tout ω , cette somme contient au plus un terme non nul).

$$\begin{aligned} \mathbb{E}(\varphi(X)) &= \sum_{n=1}^{+\infty} \mathbb{E}(\varphi(X_n) \mathbb{1}_{n=T}) \\ &= \sum_{n=1}^{+\infty} \mathbb{E}(\varphi(X_n) \prod_{i=1}^{n-1} \mathbb{1}_{U_i > \alpha(X_i)} \times \mathbb{1}_{U_n \leq \alpha(X_n)}) \\ (\text{par indépendance des } (X_i, U_i)) &= \sum_{n=1}^{+\infty} \mathbb{E}(\varphi(X_n) \mathbb{1}_{U_n \leq \alpha(X_n)}) \prod_{i=1}^{n-1} \mathbb{E}(\mathbb{1}_{U_i > \alpha(X_i)}). \end{aligned}$$

Calculons, pour tout n ,

$$\begin{aligned} \mathbb{E}(\varphi(X_n) \mathbb{1}_{U_n \leq \alpha(X_n)}) &= \int_{x \in \mathbb{R}^d} \left(\int_{u \in [0,1]} \mathbb{1}_{u \leq \alpha(x)} du \right) \varphi(x) g(x) dx \\ &= \int_{x \in \mathbb{R}^d} \frac{f(x)}{kg(x)} \varphi(x) g(x) dx \\ &= \frac{1}{k} \int_{x \in \mathbb{R}^d} \varphi(x) f(x) dx. \end{aligned}$$

Nous en déduisons, pour tout n ,

$$\begin{aligned}\mathbb{E}(\mathbb{1}_{U_n > \alpha(X_n)}) &= 1 - \mathbb{E}(\mathbb{1}_{U_n \leq \alpha(X_n)}) \\ &= 1 - \frac{1}{k}.\end{aligned}$$

Remarquons que l'inégalité $f \leq kg$ implique

$$\int_{\mathbb{R}^d} f(x) dx \leq k \int_{\mathbb{R}^d} g(x) dx,$$

donc $k \geq 1$. Nous avons donc

$$\begin{aligned}\mathbb{E}(\varphi(X)) &= \sum_{n=1}^{+\infty} \int_{\mathbb{R}^d} \varphi(x) f(x) dx \times \frac{1}{k} \left(1 - \frac{1}{k}\right)^{n-1} \\ (\text{somme géométrique}) &= \int_{\mathbb{R}^d} \varphi(x) f(x) dx \times \frac{1}{k} \times \frac{1}{1 - \left(1 - \frac{1}{k}\right)} \\ &= \int_{\mathbb{R}^d} \varphi(x) f(x) dx.\end{aligned}$$

D'où le résultat. ■

Pour simuler X selon f on peut utiliser la procédure suivante :

étape 1 : simuler X selon g , soit x la valeur obtenue,

étape 2 : simuler Y selon $U_{[0,1]}$ de manière indépendante de X , soit y la valeur obtenue,

étape 3 : comparer y à $\alpha(x) = \frac{f(x)}{kg(x)}$. Si $y < \alpha(x)$, on accepte x comme valeur simulée de X selon f , sinon on rejette x et on recommence à l'étape 1.

1.3 Simulation d'un vecteur aléatoire

1.3.1 Vecteur aléatoire à loi jointe connue

Soit $X = (X_1, X_2, \dots, X_n)$ un vecteur aléatoire.

1. Si les X_i sont indépendantes pour $i = 1, \dots, n$, pour simuler ce vecteur aléatoire on peut simuler séparément X_1, X_2, \dots, X_{n-1} et X_n selon leurs lois marginales respectives.
2. Si les X_i sont liées pour $i = 1, \dots, n$, pour simuler $X = (X_1, X_2, \dots, X_n)$ on procède comme suit :
 - on simule X_1 selon sa loi marginale, soit x_1 la valeur trouvée,
 - on simule X_2 selon sa loi conditionnelle à $X_1 = x_1$, soit x_2 la valeur trouvée,
 - ...
 - enfin on simule X_n selon sa loi conditionnelle à $X_1 = x_1, X_2 = x_2, \dots, X_{n-1} = x_{n-1}$, soit x_n la valeur trouvée.

Cas particulier d'un vecteur gaussien.

La simulation d'un vecteur gaussien peut se faire selon la méthode ci-dessus. Cependant l'analyse des propriétés de la loi normale multidimensionnelle conduit à adapter une autre méthode plus facile à mettre en œuvre.

Proposition 1.3.1 Soit μ un vecteur de \mathbb{R}^n et Σ une matrice carré de taille n symétrique positive et soit $X = (X_1, \dots, X_n)$ un vecteur aléatoire de loi $\mathcal{N}(0, I_d)$ et A une matrice carré d'ordre n tel que $AA^T = \Sigma$, alors le vecteur $Y = AX + \mu$ est un vecteur gaussien de moyenne μ et de matrice de variance covariance Σ .

Preuve. Y est une transformation linéaire de X donc gaussienne puisque X est gaussienne. De plus

$$\mathbb{E}(Y) = \mathbb{E}(AX + \mu) = A\mathbb{E}(X) + \mu = \mu.$$

$$\text{Var}(Y) = \text{Var}(AX + \mu) = \text{Var}(AX) = A\text{Var}(X)A^T = AA^T = \Sigma.$$

■

Ainsi, pour simuler $Y \rightsquigarrow \mathcal{N}(\mu, \Sigma)$, on procède comme suit :

étape 1 : simuler $X \rightsquigarrow \mathcal{N}(0, I_d)$,

étape 2 : calculer $Y = AX + \mu$.

1.3.2 Vecteur aléatoire à loi jointe inconnue

Dans certaines applications on ne cherche pas à simuler un vecteur aléatoire selon sa loi conjointe (inconnue) mais à simuler seulement ses composantes selon leurs lois marginales respectives tout en respectant un niveau donné de liaison probabiliste entre ses composantes.

Notion de Copule(divariée)

Definition 1.3.1 (*analytique*)

Une copule est une fonction $C : [0, 1]^2 \rightarrow [0, 1]$ vérifiant :

1. $\forall 0 \leq u, v \leq 1, C(u, 0) = C(0, v) = 0$
2. $\forall 0 \leq u, v \leq 1, C(u, 1) = u$ et $C(1, v) = v$
3. C est 2-croissante si $0 \leq u_1 \leq u_2 \leq 1$ et $0 \leq v_1 \leq v_2 \leq 1$ alors

$$C(u_1, v_1) - C(u_1, v_2) - C(u_2, v_1) + C(u_2, v_2) \geq 0.$$

Definition 1.3.2 (*pratique*)

On dit qu'une fonction $C : [0, 1]^2 \rightarrow [0, 1]$ est une copule si C est la fonction de répartition d'un couple (U, V) tel que U et V suivent chacune une loi $\mathcal{U}_{[0,1]}$ i.e

$$C(u, v) = \mathbb{P}(U \leq u, V \leq v); U, V \rightsquigarrow \mathcal{U}_{[0,1]}.$$

Théorème 1.3.1 (Sklar, 1959) Soit (X, Y) un couple aléatoire de f.d.r F avec X de f.d.r F_X et Y de f.d.r F_Y alors \exists une copule C telle que

$$\forall (x, y) \in \mathbb{R}^2, F(x, y) = C(F_X(x), F_Y(y)).$$

De plus, si F_X et F_Y sont continues, le copule C est unique et définie par :

$$\forall 0 \leq u, v \leq 1, C(u, v) = F(F_X^{-1}(u), F_Y^{-1}(v)),$$

où F_X^{-1} et F_Y^{-1} sont les inverses généralisées de X et Y .

D'après Sklar, le couple (X, Y) est complètement caractérisé par la donnée de sa copule et de ses lois marginales.

Considérons C, F_X et F_Y données et supposons qu'on veuille simuler une réalisation de (X, Y) , nous pouvons procéder en deux étapes :

étape 1 : simuler (X, Y) de lois marginales uniformes et de copule C ,

étape 2 : calculer $X = F_X^{-1}(U)$ et $Y = F_Y^{-1}(V)$.

Cette procédure peut être généralisée facilement en dimension supérieure où égale 2.

Chapitre 2

QUELQUES METHODES DETERMINISTES

Supposons pour simplifier que f défini dans l'intégrale (0.0.1) soit la densité uniforme sur le pavé $C = [0, 1]^d$, c'est-à-dire que

$$I = \int_{[0,1]^d} \varphi(x) dx.$$

2.1 Intégrale simple ($d = 1$)

Considérons que la fonction φ est suffisamment régulière. L'idée est de subdiviser l'intervalle $[0, 1]$ en n sous-intervalles de même longueur $h = x_{i+1} - x_i = \frac{1}{n}$ (h est le pas de la subdivision). On pose $x_i = \frac{i}{n}$, $i = 0, \dots, n$.

2.1.1 Méthode des Rectangles

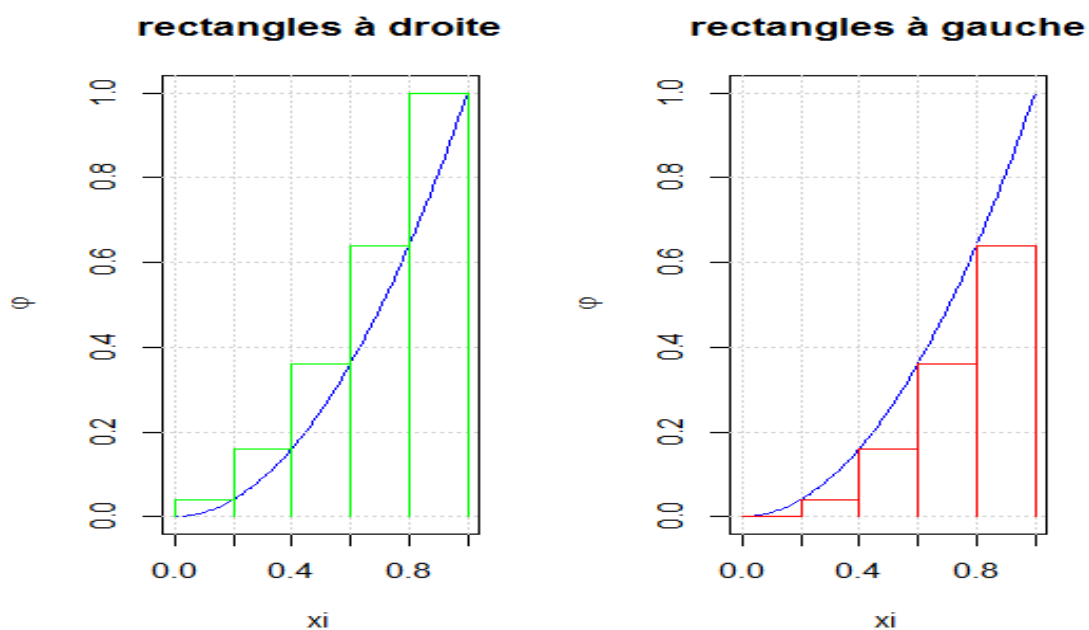


FIGURE 2.1 – Méthode des Rectangles.

Elle consiste à remplacer φ par la valeur qu'elle prend sur un bord, à gauche ou à droite de l'intervalle $[x_i, x_{i+1}]$.

Sur chacun des intervalles $[x_i, x_{i+1}]$ (pour $0 \leq i \leq n-1$) on remplace l'aire sous la courbe par l'aire du rectangle dont les dimensions sont $(x_{i+1} - x_i)$ et $\varphi(x_i)$ et on a donc

$$\int_{x_i}^{x_{i+1}} \varphi(x) dx \approx (x_{i+1} - x_i) \varphi(x_i) = h \varphi(x_i).$$

L'approximation de $\int_0^1 \varphi(x)$ notée R_n est donnée par

$$\int_0^1 \varphi(x) = \sum_{i=0}^{n-1} \int_{x_i}^{x_{i+1}} \varphi(x) dx \approx h \sum_{i=0}^{n-1} \varphi(x_i) = R_n. \quad (2.1.1)$$

$R_n^g = h \sum_{i=0}^{n-1} \varphi(x_i)$ est l'approximation de I par la méthode des rectangles à gauche.

Une approximation de I par la méthode des rectangles à droite est donnée par :

$$R_n^d = h \sum_{i=1}^n \varphi(x_i).$$

Evaluation de l'erreur

Proposition 2.1.1 Si φ est de classe C^1 sur $[0, 1]$ et $M_1 = \sup_{x \in [0,1]} |\varphi'(x)|$, on a :

$$\left| \int_0^1 \varphi(x) dx - R_n^g \right| \leq \frac{M_1}{2n}, \quad \left| \int_0^1 \varphi(x) dx - R_n^d \right| \leq \frac{M_1}{2n}.$$

Preuve. Supposons que φ est de classe C^1 sur $[x_k, x_{k+1}]$, $\forall x \in [x_k, x_{k+1}]$, d'après le théorème des accroissement finis

$$\begin{aligned} |\varphi(x) - \varphi(x_k)| &\leq M_1(x - x_k) \\ \int_{x_k}^{x_{k+1}} |\varphi(x) - \varphi(x_k)| dx &\leq M_1 \int_{x_k}^{x_{k+1}} (x - x_k) dx \\ \left| \int_{x_k}^{x_{k+1}} [\varphi(x) - \varphi(x_k)] dx \right| &\leq M_1 \int_{x_k}^{x_{k+1}} (x - x_k) dx \\ \left| \int_{x_k}^{x_{k+1}} \varphi(x) dx - \int_{x_k}^{x_{k+1}} \varphi(x_k) dx \right| &\leq M_1 \left[\frac{1}{2} x^2 - x x_k \right]_{x_k}^{x_{k+1}} \\ \left| \int_{x_k}^{x_{k+1}} \varphi(x) dx - (x_{k+1} - x_k) \varphi(x_k) \right| &\leq \frac{M_1}{2} (x_{k+1} - x_k)^2 \quad \text{or} \quad h = x_{k+1} - x_k = \frac{1}{n} \\ \left| \sum_{k=0}^{n-1} \int_{x_k}^{x_{k+1}} \varphi(x) dx - \sum_{k=0}^{n-1} (x_{k+1} - x_k) \varphi(x_k) \right| &\leq \frac{M_1}{2} \sum_{k=0}^{n-1} h^2 \\ \left| \int_0^1 \varphi(x) dx - R_n^g \right| &\leq \frac{M_1}{2n}. \end{aligned}$$

■

Dans le cas où la fonction φ est de classe C^1 , la méthode des rectangles sur la subdivision régulière $\{\frac{1}{n}, \dots, \frac{n-1}{n}, 1\}$ à une précision de $o(1/n)$.

2.1.2 Méthode des Trapèzes

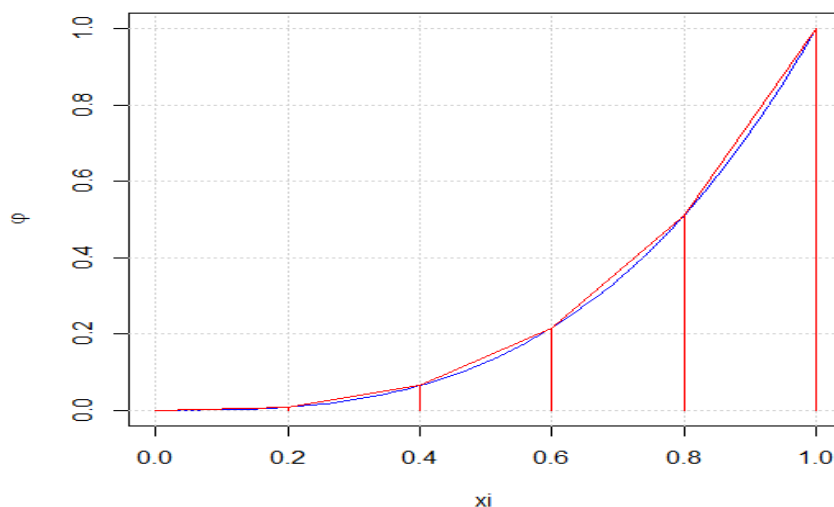


FIGURE 2.2 – méthode des Trapèzes.

Ici on remplace $\int_{x_i}^{x_{i+1}} \varphi(x)dx$ par l'aire du trapèze de hauteur $x_{i+1} - x_i$ et dont les bases sont $\varphi(x_i)$ et $\varphi(x_{i+1})$. On fait donc l'approximation $\int_{x_i}^{x_{i+1}} \varphi(x) \approx \frac{h}{2} (\varphi(x_i) + \varphi(x_{i+1}))$
L'approximation de I , notée T_n , est donnée par :

$$\int_{[0,1]} \varphi(x)dx \approx \frac{h}{2} \sum_{i=0}^{n-1} (\varphi(x_i) + \varphi(x_{i+1})) = T_n.$$

Pour éviter de calculer deux fois $\varphi(x_i)$ pour $1 \leq i \leq n - 1$, on peut réécrire cette approximation par :

$$T_n = \frac{h}{2} \left[\varphi(0) + \varphi(1) + 2 \sum_{i=1}^{n-1} \varphi(x_i) \right]. \quad (2.1.2)$$

Evaluation de l'erreur

Proposition 2.1.2 Si φ est de classe C^2 sur $[0, 1]$ et $M_2 = \sup_{x \in [0,1]} |\varphi''(x)|$, on a :

$$\left| \int_0^1 \varphi(x)dx - T_n \right| \leq \frac{M_2}{12n^2}$$

Preuve. L'idée est de majorer $\int_{x_i}^{x_{i+1}} \varphi(x)dx - \frac{\varphi(x_i) + \varphi(x_{i+1})}{2}(x_{i+1} - x_i)$

Considérons x_{i+1} comme une variable et étudions la fonction définie par

$$g(x) = \int_{x_i}^x \varphi(t)dt - \frac{\varphi(x_i) + \varphi(x)}{2}(x - x_i), \quad \forall x \in [x_{i+1}, x_i]$$

Nous remarquons que $g(x_i) = 0$ donc g est dérivable et on a

$$g'(x) = \varphi(x) - \frac{\varphi'(x)}{2}(x - x_i) - \frac{\varphi(x_i) + \varphi(x)}{2}, \quad \forall x \in [x_i, x_{i+1}]$$

$g'(x_i) = 0$ donc g' est dérivable et on a

$$g''(x) = -\frac{\varphi''(x)}{2}(x - x_i) + \varphi'(x) - \frac{\varphi'(x)}{2} - \frac{\varphi'(x)}{2} = -\frac{\varphi''(x)}{2}(x - x_i), \quad \forall x \in [x_i, x_{i+1}]$$

$$\begin{aligned} |g'(x)| &= \int_{x_i}^x |g''(t)| dt \leq \int_{x_i}^x \frac{M_2}{2}(t - x_i) dt \\ &\leq \frac{M_2}{4}(x - x_i)^2 \\ |g(x)| &= \int_{x_i}^x |g'(t)| dt \leq \int_{x_i}^x \frac{M_2}{4}(t - x_i)^2 dt \\ &\leq \frac{M_2}{12}(x - x_i)^3 \\ \left| \int_{x_i}^x \varphi(t) dt - \frac{\varphi(x_i) + \varphi(x)}{2}(x - x_i) \right| &\leq \frac{M_2}{12}(x - x_i)^3 \end{aligned}$$

En remplaçant x par x_{i+1} on obtient :

$$\begin{aligned} \left| \int_{x_i}^{x_{i+1}} \varphi(x) dx - \frac{\varphi(x_i) + \varphi(x_{i+1})}{2}(x_{i+1} - x_i) \right| &\leq \frac{M_2}{12}(x_{i+1} - x_i)^3 \quad \text{or } h = x_{i+1} - x_i \\ \left| \sum_{i=0}^{n-1} \int_{x_i}^{x_{i+1}} \varphi(x) dx - \frac{h}{2} \sum_{i=0}^{n-1} [\varphi(x_i) + \varphi(x_{i+1})] \right| &\leq \frac{M_2}{12} \sum_{i=0}^{n-1} h^3 \quad \text{or } h = \frac{1}{n} \\ \left| \int_0^1 \varphi(x) dx - T_n \right| &\leq \frac{M_2}{12n^2}. \end{aligned}$$

■

2.1.3 Méthode de Simpson

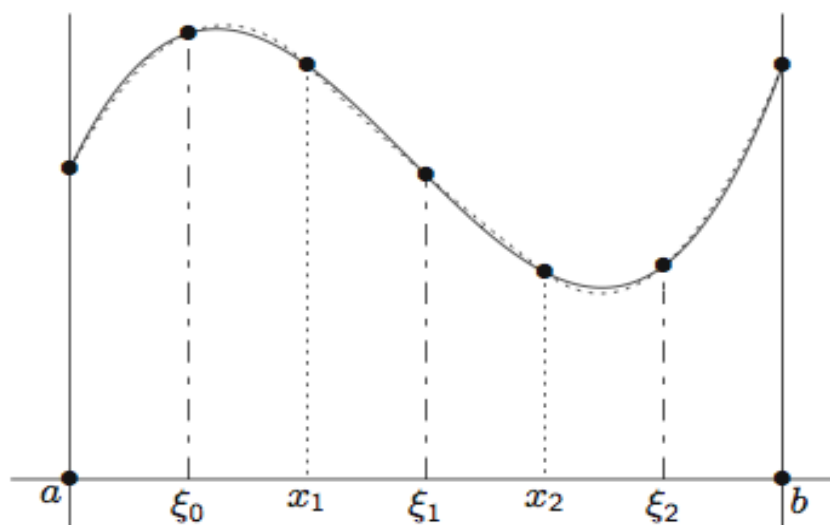


FIGURE 2.3 – Méthode de Simpson.

Definition 2.1.1 Soient $n \in \mathbb{N}^*$ et $(x_i, y_i)_{i \in \llbracket 0, n \rrbracket}$ avec $(x_i, y_i) \in \mathbb{R}^2$ et les x_i distincts deux à deux. Le polynôme d'interpolation de Lagrange associé aux $n + 1$ points $(x_i, y_i)_{i \in \llbracket 0, n \rrbracket}$, noté P_n , est donné par

$$P_n(x) = \sum_{i \neq j}^n y_i L_i(x), \text{ avec } L_i(x) = \prod_{i \neq j}^n \frac{x - a_j}{a_i - a_j}, i = 1, \dots, n \forall x \in \mathbb{R}. \quad (2.1.3)$$

Principe

Cette méthode consiste à remplacer φ sur chaque segment $[x_i, x_{i+1}]$ de la subdivision par son polynôme de Lagrange P_i de degré inférieur ou égal à 2 qui prend les mêmes valeurs que φ aux extrémités et au milieu ξ_i de ce segment.

La valeur approchée de I par la méthode de Simpson est donnée par :

$$S_n = \frac{h}{6} \sum_{i=0}^{n-1} \left[\varphi(x_i) + \varphi\left(\frac{x_i + x_{i+1}}{2}\right) + \varphi(x_{i+1}) \right]. \quad (2.1.4)$$

Lemme 2.1.1 Soit $P \in \mathbb{R}_2[X]$ et $c < d$, alors

$$\int_c^d P(t) dt = \frac{d-c}{6} \left[P(c) + P(d) + 4P\left(\frac{c+d}{2}\right) \right].$$

Preuve. On fait le changement de variable $t = (d-c)u + c$ pour se ramener à une intégrale de 0 à 1.

$$\int_c^d P(t) dt = (d-c) \int_0^1 P((d-c)u + c) du.$$

Les polynômes de Lagrange aux points 0, 1 et $\frac{1}{2}$ sont :

$$P_1(x) = 2x^2 - 3x + 1, P_2(x) = 2x^2 + x \text{ et } P_3(x) = 4x - 4x^2;$$

leurs intégrales entre 0 et 1 valent respectivement $\frac{1}{6}$, $\frac{1}{6}$ et $\frac{4}{6}$.

Soit $Q(u) = P((d-c)u + c)$, Q est un polynôme de degré inférieur ou égal à 2, on a donc

$$Q = Q(0)P_0 + Q(1)P_1 + Q\left(\frac{1}{2}\right)P_2 \text{ et } \int_0^1 Q(u) du = \frac{1}{6} [Q(0) + Q(1) + 4Q\left(\frac{1}{2}\right)].$$

Comme $Q(0) = P(c)$, $Q(1) = P(d)$ et $Q\left(\frac{1}{2}\right) = P\left(\frac{c+d}{2}\right)$, on obtient le résultat. ■

Evaluation de l'erreur

Proposition 2.1.3 Si φ est de classe C^4 sur $[0, 1]$ et $M_4 = \sup_{x \in [0, 1]} |\varphi^{(4)}(x)|$, alors pour tout entier n non nul

$$\left| S_n - \int_{[0, 1]} \varphi(x) dx \right| \leq \frac{M_4}{2880n^4}.$$

Preuve. On pose $x_i = \alpha - \frac{h}{2}$, $\xi_i = \alpha$ et $x_{i+1} = \alpha + \frac{h}{2}$. Montrons que

$$\left| \int_{\alpha - \frac{h}{2}}^{\alpha + \frac{h}{2}} (\varphi(t) - P_i(t)) dt \right| \leq \frac{h^5}{2880} M_4,$$

où P_i est la fonction polynomiale de degré 2 égale à φ aux points x_i , ξ_i , et x_{i+1} .

Posons $g(h) = \int_{\alpha-\frac{h}{2}}^{\alpha+\frac{h}{2}} (\varphi(t) - P(t)) dt$. On a

$$g(h) = \int_{\alpha-\frac{h}{2}}^{\alpha+\frac{h}{2}} \varphi(t) dt - \frac{h}{6} \left[\varphi\left(\alpha + \frac{h}{2}\right) + \varphi\left(\alpha - \frac{h}{2}\right) + 4\varphi(\alpha) \right].$$

Calculons ses dérivées jusqu'à l'ordre 3.

$$g'(h) = \frac{1}{3} \left[\varphi\left(\alpha + \frac{h}{2}\right) + \varphi\left(\alpha - \frac{h}{2}\right) - 2\varphi(\alpha) \right] - \frac{h}{12} \left[\varphi'\left(\alpha + \frac{h}{2}\right) + \varphi'\left(\alpha - \frac{h}{2}\right) \right]$$

$$g'(0) = 0$$

$$g''(h) = \frac{1}{12} \left[\varphi\left(\alpha + \frac{h}{2}\right) - \varphi\left(\alpha - \frac{h}{2}\right) \right] - \frac{h}{24} \left[\varphi''\left(\alpha + \frac{h}{2}\right) + \varphi''\left(\alpha - \frac{h}{2}\right) \right]$$

$$g''(0) = 0$$

$$g^{(3)}(h) = \frac{h}{48} \left[\varphi^{(3)}\left(\alpha + \frac{h}{2}\right) + \varphi^{(3)}\left(\alpha - \frac{h}{2}\right) \right]$$

$$g^{(3)}(0) = 0.$$

La fonction φ étant de classe C^4 , d'après l'inégalité des accroissements finis, on a :

$$\left| g^{(3)}(h) \right| \leq \frac{h^2}{48} M_4.$$

En intégrant entre 0 et h , on obtient :

$$\int_0^h \left| g^{(3)}(t) \right| dt \leq \int_0^h \frac{t^2}{48} M_4 dt$$

$$\left| g^{(2)}(h) \right| \leq \frac{h^3}{144} M_4$$

$$\int_0^h \left| g^{(2)}(t) \right| dt \leq \int_0^h \frac{t^3}{144} M_4 dt$$

$$\left| g^{(1)}(h) \right| \leq \frac{h^4}{576} M_4$$

$$\int_0^h \left| g^{(1)}(t) \right| dt \leq \int_0^h \frac{t^4}{576} M_4 dt$$

$$\left| g(h) \right| \leq \frac{h^5}{2880} M_4.$$

En notant P la fonction polynômiale par morceaux égale à P_i sur chaque segment $[x_i, x_{i+1}]$, on obtient :

$$\begin{aligned} \left| S_n - \int_{[0,1]} \varphi(x) dx \right| &\leq \sum_{i=0}^{n-1} \left| \int_{\xi_i-\frac{h}{2}}^{\xi_i+\frac{h}{2}} (\varphi(t) - P(t)) dt \right| \\ &\leq \sum_{i=0}^{n-1} \frac{h^5}{2880} M_4 \\ &\leq n \frac{h^5}{2880} M_4 \\ &\leq \frac{M_4}{2880n^4} \quad \text{car } h = \frac{1}{n}. \end{aligned}$$

■

De façon générale, si φ est de classe C^s , une méthode adaptée à la subdivision régulière $\{\frac{1}{n}, \dots, \frac{n-1}{n}, 1\}$ permettra d'obtenir une erreur en $o(\frac{1}{n^s})$.

2.2 Intégration multiple ($d \geq 2$)

2.2.1 Intégration double

Focalisons-nous sur la plus simple des méthodes déterministes, à savoir celle des rectangles. Si l'on considère

$$I = \int_0^1 \int_0^1 \varphi(x, y) dx dy.$$

Soit $D = [0, 1]^2$ un domaine régulier du plan et $\varphi : D \rightarrow \mathbb{R}$ une fonction continue.

Pour approcher l'intégrale de φ on va procéder comme en dimension 1 par découpage. Pour $n \geq 1$ entier donné, on peut découper le plan en des carrés de largeur $\frac{1}{n}$ et on a

$$C_{k,l} = \left\{ (x, y) \in \mathbb{R}^2, \frac{k}{n} \leq x < \frac{k+1}{n}, \frac{l}{n} \leq y < \frac{l+1}{n}, \right\}$$

où $k, l = 0, \dots, n-1$. On évalue la fonction φ au point $M_{k,l} = (\frac{k}{n}, \frac{l}{n}) \in C_{k,l}$. Comme la surface du carré $C_{k,l}$ est $\frac{1}{n^2}$ ainsi on a :

$$I = \int_0^1 \int_0^1 \varphi(x, y) dx dy = \sum_{k,l=0}^{n-1} \int_{k/n}^{(k+1)/n} \int_{l/n}^{(l+1)/n} \varphi(x, y) dx dy.$$

L'approximation numérique de I est donnée par :

$$R_n = \frac{1}{n^2} \sum_{k,l=0}^{n-1} \varphi(k/n, l/n).$$

L'erreur est bornée par

$$|I - R_n| \leq \sum_{k,l=0}^{n-1} \int_{k/n}^{(k+1)/n} \int_{l/n}^{(l+1)/n} |\varphi(x, y) - \varphi(k/n, l/n)| dx dy.$$

Par le théorème des accroissements finis (on rappelle que φ est à valeurs dans \mathbb{R}), pour tout couple (k, l) , il existe $C_{k,l} \in [k/n, (k+1)/n] \times [l/n, (l+1)/n]$ tel que :

$$|\varphi(x, y) - \varphi(k/n, l/n)| = \frac{\partial \varphi}{\partial x}(C_{k,l})(x - k/n) + \frac{\partial \varphi}{\partial y}(C_{k,l})(y - l/n) \leq \frac{1}{n} \|\nabla \varphi(C_{k,l})\|_1.$$

En notant $M_1 = \sup_{0 \leq x, y \leq 1} \|\nabla \varphi(x, y)\|_1$, on obtient donc

$$|I - R_n| \leq \frac{M_1}{n}.$$

Notons par ailleurs que souvent un changement de variable s'impose. Il est d'usage de procéder à un changement de variable par le biais des coordonnées polaires en posant $x = r \cos(\theta)$ et $y = r \sin(\theta)$. Dans ce cas, l'élément de surface $dx dy$ se transforme en $r dr d\theta$. La formule de changement de variable s'écrit alors :

$$I = \int_D \varphi(x, y) dx dy = \int_{(r,\theta) \in [0, \sqrt{2}] \times [0, \frac{\pi}{2}[} \varphi(r \cos(\theta), r \sin(\theta)) r dr d\theta.$$

Théorème 2.2.1 (Fubini) Soit $f : D \subset \mathbb{R}^2 \longrightarrow \mathbb{R}$ continue sur un domaine élémentaire du type

$$D = \{(x, y) \in \mathbb{R}^2, a \leq x \leq b, \phi(x) \leq y \leq \psi(x)\},$$

avec ϕ, ψ continues et $\phi \leq \psi$.

Alors la fonction $g : x \in [a, b] \longmapsto \int_{\phi(x)}^{\psi(x)} f(x, y) dy$ est continue et on a :

$$\int \int_D f(x, y) dx dy = \int_a^b \left(\int_{\phi(x)}^{\psi(x)} f(x, y) dy \right) dx.$$

2.2.2 Intégration Triple

L'intégrale d'une fonction à 3 variables $\varphi : [0, 1]^3 \longrightarrow \mathbb{R}$ sur un domaine de \mathbb{R}^3 se définit selon les mêmes principes que ci-dessus. Pour $n \geq 1$, l'espace se découpe en cubes

$$C_{k,l,m} = \left\{ (x, y, z) \in \mathbb{R}^3, \frac{k}{n} \leq x < \frac{k+1}{n}, \frac{l}{n} \leq y < \frac{l+1}{n}, \frac{m}{n} \leq z < \frac{m+1}{n} \right\}.$$

où $k, l, m = 0, \dots, n-1$. Sur chaque cube, on échantillonne la fonction φ au point $M_{k,l,m} = \left(\frac{k}{n}, \frac{l}{n}, \frac{m}{n}\right) \in C_{k,l,m}$. Le volume de chaque petit cube $C_{k,l,m}$ de côté $\frac{1}{n}$ est $\frac{1}{n^3}$. L'approximation numérique de I est donnée par :

$$S_n(\varphi) = \frac{1}{n^3} \sum_{M_{k,l,m} \in [0,1]^3} \varphi(M_{k,l,m}).$$

D'autres systèmes de coordonnées, à l'instar des coordonnées cylindriques et sphériques, sont également utilisés pour résoudre ce type d'intégrale. En coordonnées cylindriques le volume élémentaire devient $dv = \rho d\rho d\theta dz$, ce qui donne :

$$\int \int \int_{(x,y,z) \in D} \varphi(x, y, z) dx dy dz = \int \int \int_{(\rho,\theta,z) \in Q} \varphi(\rho \cos(\theta), \rho \sin(\theta), z) \rho d\rho d\theta dz, \quad (2.2.1)$$

où $Q = \{(\rho, \theta, z) \in]0, \sqrt{2}] \times]0, \frac{\pi}{2}] \times [0, 1]\}$

En coordonnées sphérique le volume élémentaire devient $dv = r^2 \sin(\theta) dr d\psi d\theta$, ce qui donne :

$$\begin{aligned} \int \int \int_{(x,y,z) \in D} \varphi(x, y, z) dx dy dz = \\ \int \int \int_{(r,\psi,\theta) \in Q} \varphi(r \cos(\psi) \sin(\theta), r \sin(\psi) \sin(\theta), r \cos(\theta)) r^2 \sin(\theta) dr d\psi d\theta. \end{aligned} \quad (2.2.2)$$

où $Q = \{(r, \psi, \theta) \in]0, \sqrt{3}] \times]0, \frac{\pi}{2}] \times [0, \frac{\pi}{2}]\}$

De façon générale le problème de l'évaluation des intégrales multiples est étroitement lié à la difficulté de l'évaluation numérique d'un très grand nombre de points pour la fonction considérée. Par exemple, dans un espace à trois dimensions, si on utilise 30 points dans chacune des directions, il est nécessaire de calculer la fonction pour 30^3 points. La situation s'aggrave très rapidement quand la dimension de l'espace augmente. Le calcul des intégrales multiples est néanmoins possible dans un certain nombre de cas.

Le cas où la fonction possède une symétrie importante, on peut ramener le calcul de l'intégrale de dimension d à celui d'une intégrale de dimension $d' \ll d$ en utilisant un changement de coordonnées où les méthodes précédentes peuvent encore s'appliquer. Considérons le cas où

d est impair : on définit les coordonnées hypersphériques de la manière suivante :

$$\begin{aligned}
x_1 &= \rho \sin(\phi_1) \sin(\phi_2) \sin(\phi_3) \dots \sin(\phi_{d-3}) \sin(\phi_{d-2}) \sin(\phi_{d-1}) \\
x_2 &= \rho \sin(\phi_1) \sin(\phi_2) \sin(\phi_3) \dots \sin(\phi_{d-3}) \sin(\phi_{d-2}) \cos(\phi_{d-1}) \\
x_3 &= \rho \sin(\phi_1) \sin(\phi_2) \sin(\phi_3) \dots \sin(\phi_{d-3}) \cos(\phi_{d-2}) \\
x_4 &= \rho \sin(\phi_1) \sin(\phi_2) \sin(\phi_3) \dots \cos(\phi_{d-3}) \\
&\dots \\
x_{d-1} &= \rho \sin(\phi_1) \cos(\phi_2) \\
x_d &= \rho \cos(\phi_1),
\end{aligned}$$

où les variables ϕ_i varient de 0 à π pour $1 \leq i \leq d-2$ et la variable ϕ_{d-1} varie entre 0 et 2π , et ρ est positive ou nulle.

Pour d pair, on définit les coordonnées hypersphériques de la manière suivante :

$$\begin{aligned}
x_1 &= \rho \sin(\phi_1) \sin(\phi_2) \sin(\phi_3) \dots \sin(\phi_{d-3}) \sin(\phi_{d-2}) \sin(\phi_{d-1}) \\
x_2 &= \rho \sin(\phi_1) \sin(\phi_2) \sin(\phi_3) \dots \sin(\phi_{d-3}) \sin(\phi_{d-2}) \cos(\phi_{d-1}) \\
x_3 &= \rho \sin(\phi_1) \sin(\phi_2) \sin(\phi_3) \dots \sin(\phi_{d-3}) \cos(\phi_{d-2}) \\
x_4 &= \rho \sin(\phi_1) \sin(\phi_2) \sin(\phi_3) \dots \cos(\phi_{d-3}) \\
&\dots \\
x_{d-1} &= \rho \sin(\phi_1) \cos(\phi_2) \\
x_d &= \rho \cos(\phi_2) \cos(\phi_1).
\end{aligned}$$

Pour évaluer une intégrale à d dimensions, il est nécessaire de calculer le jacobien de la transformation des coordonnées cartésiennes aux coordonnées hypersphériques, pour d impair on a

$$J = \frac{\partial(x_1 x_2 \dots x_d)}{\partial(\rho \phi_1 \phi_2 \dots \phi_d)} = \rho^{d-1} \times \begin{vmatrix} \sin(\phi_1) \sin(\phi_2) \dots \sin(\phi_{d-1}) & \cos(\phi_1) \dots \sin(\phi_{d-1}) & \dots & \sin(\phi_1) \dots \sin(\phi_{d-1}) \\ \sin(\phi_1) \sin(\phi_2) \dots \cos(\phi_{d-1}) & \cos(\phi_1) \dots \cos(\phi_{d-1}) & \dots & -\sin(\phi_1) \dots \sin(\phi_{d-1}) \\ \dots & \dots & \dots & \dots \\ \cos(\phi_1) & -\sin(\phi_1) & 0 & 0 \end{vmatrix}.$$

Par récurrence, on montre que le jacobien J se réexprime sous la forme

$$J = \rho^{d-1} \prod_{j=1}^{d-1} \sin^{d-1-j}(\phi_j) = \prod_{j=1}^{d-1} \sin(\phi_{d-1-j})^j.$$

On peut vérifier que la formule convient aussi pour d pair.

2.3 Conclusion

Les méthodes usuelles de quadrature (méthode de Simpson, méthodes de Trapèze, méthodes des rectangles...) sont efficaces pour le calcul d'intégrales à une dimension et se généralisent facilement à plusieurs dimensions lorsque le domaine d'intégration est simple (hypercube, par exemple). De façon générale, si φ est de classe C^s sur $[0, 1]^d$, alors une méthode déterministe adaptée à cette régularité et faisant n appels à la fonction φ permettra d'atteindre une vitesse en $o(n^{-s/d})$. Cette vitesse qui s'effondre avec d est symptomatique du fléau de la dimension. A contrario, comme nous allons le voir, la méthode de Monte-Carlo, avec une vitesse en $\frac{1}{\sqrt{n}}$, est insensible à la dimension et peut donc s'avérer plus avantageuse dès que l'on travaille en dimension élevée où sur une fonction irrégulière.

Chapitre 3

GENERALITES SUR LES METHODES DE MONTE-CARLO

3.1 Introduction

Lorsque la résolution mathématique d'un problème donné n'est pas possible analytiquement, on fait appel à des méthodes d'approximation. Cela permet de modéliser des situations très complexes dont le calcul analytique s'avère difficile voire impossible. Parmi ces méthodes d'approximation figure celle de Monte-Carlo. Une application classique des méthodes de Monte-Carlo est le calcul des quantités du type

$$I = \mathbb{E}[\varphi(X)] = \int \varphi(x)f(x)dx, \quad (3.1.1)$$

où $\varphi : \mathbb{R}^n \rightarrow \mathbb{R}$ est une fonction donnée et X un vecteur aléatoire (V.a) de densité f suivant laquelle on sait simuler. Dans ce contexte, l'estimateur Monte-Carlo de base est défini par :

$$\hat{I}_n = \frac{1}{n} \sum_{i=1}^n \varphi(X_i), \quad (3.1.2)$$

où les X_i sont générées de façon indépendante et identiquement distribuée (i.i.d) selon f . Dans ce chapitre, nous allons présenter les propriétés de cet estimateur et nous donnerons aussi une manière d'améliorer sa précision grâce à des techniques de réduction de la variance.

3.2 Simulation de Monte-Carlo : Calcul d'intégrale

3.2.1 Description de la méthode

L'idée sous-jacente aux méthodes de Monte-Carlo consiste à représenter la quantité que l'on cherche à calculer comme l'espérance d'une distribution et d'utiliser un générateur de nombre aléatoire pour construire un estimateur empirique de cette quantité.

Considérons une fonction h (d'une ou plusieurs variables) intégrable et $I = \int_C h(x)dx$ où $C \subset \mathbb{R}^d$. On suppose qu'il n'est pas possible d'évaluer analytiquement I et on cherche en conséquence à l'approximer numériquement.

Soit f une densité de probabilité ayant comme support C alors on peut réécrire

$$\begin{aligned} I &= \int_C \varphi(x)f(x)dx && \text{où } \varphi(x) = \frac{h(x)}{f(x)} \\ I &= \mathbb{E}[\varphi(X)] && \text{où } X \text{ loi de densité } f. \end{aligned}$$

Une approximation de I est donnée par :

$$I \simeq \widehat{I}_n = \frac{1}{n} \sum_{i=1}^n \varphi(X_i).$$

Remarque 3.2.1 *Toute intégrale peut s'écrire comme une espérance mathématique d'une certaine fonction φ d'un V.a X*

On peut calculer cette approximation de Monte-Carlo en utilisant l'algorithme suivant :

- générer n v.a i.i.d selon $f : x_1, \dots, x_n$,
- déterminer les quantités $\varphi(x_1), \dots, \varphi(x_n)$,
- calculer la moyenne arithmétique de ces quantités donnant

$$I \simeq \widehat{I}_n = \frac{1}{n} \sum_{i=1}^n \varphi(x_i).$$

Remarque 3.2.2 *L'approximation de Monte-Carlo de I n'est pas unique. En effet on peut écrire $I = \mathbb{E}[\varphi^*(X)]$, où $\varphi^*(x) = \frac{h(x)}{f^*(x)}$ et f^* une autre densité de X ayant pour support C et donc $\widehat{I}_n^* = \frac{1}{n} \sum_{i=1}^n \varphi^*(x_i)$ définit une approximation de Monte-Carlo de I .*

En fait, il y a autant d'approximations que de densités de probabilité sur le support C . La méthode de Monte-Carlo est une méthode d'approximation au sens statistique du terme, son fondement mathématique est la loi des grands nombres.

Proposition 3.2.1 (loi des grandes nombres(LGN)) *Etant donné un n -échantillon de v.a.r (X_1, \dots, X_n) i.i.d tel que $\mathbb{E}(X_1) = m < \infty$. alors*

$$\frac{1}{n} \sum_{i=1}^n X_i \xrightarrow[n \rightarrow \infty]{p.s} \mathbb{E}(X_1) = m.$$

Proposition 3.2.2 *Si $\mathbb{E}[\varphi(X)] < \infty$, alors*

$$\widehat{I}_n = \frac{1}{n} \sum_{i=1}^n \varphi(X_i) \xrightarrow[n \rightarrow \infty]{p.s} I.$$

Prenons un exemple élémentaire de l'estimation de π .

3.2.2 Exemple d'un estimateur de π

Supposons que (X, Y) suive la loi uniforme sur le carré $C = [0, 1] \times [0, 1]$ et que $\psi(x, y) = \mathbb{1}_{x^2+y^2 \leq 1}$.

En notant $D = \{(x, y) \in C, x^2 + y^2 \leq 1\}$ le quart de disque unité et λ la mesure de Lebesgue de \mathbb{R}^2 , on a :

$$I = \iint_C \mathbb{1}_D(x, y) dx dy = \lambda(D) = \frac{\pi}{4}.$$

Ainsi on génère n points (X_i, Y_i) uniformément dans C et, d'après la LGN, on a :

$$\widehat{I}_n = \frac{1}{n} \sum_{i=1}^n \mathbb{1}_D(X_i, Y_i) \xrightarrow[n \rightarrow \infty]{p.s} \frac{\pi}{4} \Leftrightarrow 4 \times \widehat{I}_n \xrightarrow[n \rightarrow \infty]{p.s} \pi.$$

On dispose donc d'un estimateur Monte-Carlo pour la constante π .

3.3 Propriétés de l'estimateur \widehat{I}_n

3.3.1 Qualités de l'estimateur

Absence de biais

L'estimateur \widehat{I}_n est sans biais, pour tout n , ce qui signifie que la moyenne de \widehat{I}_n est exactement I . Effet,

$$\begin{aligned}\mathbb{E}[\widehat{I}_n] &= \mathbb{E}\left[\frac{1}{n} \sum_{i=1}^n \varphi(X_i)\right] \\ &= \frac{1}{n} \sum_{i=1}^n \mathbb{E}[\varphi(X_i)] \text{ par linéarité} \\ &= \frac{1}{n} (n \times \mathbb{E}[\varphi(X_1)]) \text{ car les v.a sont i.i.d} \\ &= \mathbb{E}[\varphi(X_1)] = I.\end{aligned}$$

Convergence de \widehat{I}_n

Par construction, \widehat{I}_n converge en probabilité, puisque $\widehat{I}_n \xrightarrow[n \rightarrow \infty]{p.s.} I$ (LGN) et donc $\mathbb{P}(\lim_{n \rightarrow \infty} \widehat{I}_n = I) = 1$.

La normalité asymptotique

Proposition 3.3.1 (Théorème central limite(TCL)) Avec les notations précédentes, si $\mathbb{E}[\varphi^2(X)] < \infty$ alors

$$\sqrt{n}(\widehat{I}_n - I) \xrightarrow[n \rightarrow \infty]{\mathcal{L}} \mathcal{N}(0, \sigma^2),$$

avec $\sigma^2 = \text{Var}[\varphi(X)] = \mathbb{E}[\varphi(X)^2] - \mathbb{E}[\varphi(X)]^2 = \int \varphi(x)^2 f(x) dx - I^2$.

Ainsi, lorsque n est grand, notre estimateur suit à peu près une loi normale. Avec un abus de notation flagrant, on a $\widehat{I}_n \approx \mathcal{N}(0, \sigma^2/n)$ c'est-à-dire que \widehat{I}_n tend vers I avec une vitesse en $o(\frac{1}{\sqrt{n}})$. Plus précisément, le TCL doit permettre de construire des intervalles de confiance(IC). Cependant l'écart-type σ est lui aussi inconnu Car

$$\text{Var}(\varphi(X)) = \mathbb{E}[\varphi(X)^2] - I^2.$$

En effet, on peut approximer I^2 par \widehat{I}_n^2 et quant à $J = \mathbb{E}[\varphi(X)^2]$ on peut utiliser le même principe de la méthode de Monte-Carlo pour l'approximer. L'approximation de J est donnée par :

$$\widehat{J}_n = \frac{1}{n} \sum_{i=1}^n \varphi^2(X_i) \text{ où les } X_i \text{ i.i.d } \rightsquigarrow f$$

Une approximation de σ^2 peut être donnée par :

$$\widehat{\sigma}_n^2 = \frac{1}{n} \sum_{i=1}^n \varphi^2(X_i) - \widehat{I}_n^2. \quad (3.3.1)$$

Ce qui fournit une estimation de la précision de l'approximation de la méthode de Monte-Carlo. La Méthode de MC en fournit justement un estimateur basé sur le même échantillon

X_1, \dots, X_n i.i.d $\rightsquigarrow f$.

D'après la LGN et la Proposition 3.3.1 on a :

$$\widehat{\sigma}_n^2 = \frac{1}{n} \sum_{i=1}^n \varphi^2(X_i) - \widehat{I}_n^2 \xrightarrow[n \rightarrow \infty]{p.s.} \sigma^2. \quad (3.3.2)$$

Le lemme de Slutsky implique que

$$\sqrt{n} \left(\frac{\widehat{I}_n - I}{\widehat{\sigma}_n} \right) \xrightarrow[n \rightarrow \infty]{\mathcal{L}} \mathcal{N}(0, 1).$$

3.3.2 Exemple (Estimation de π)

Dans ce cas, la variance vaut tout simplement

$$\begin{aligned} \sigma^2 &= \text{Var}(\varphi(X)) \\ &= \mathbb{E}[\varphi^2(X, Y)] - \mathbb{E}[\varphi(X, Y)]^2 \quad \text{or} \quad \varphi(X, Y) = 1_{X^2+Y^2} \quad \text{et} \quad \mathbb{E}[1_{X^2+Y^2}^2] = E[1_{X^2+Y^2}] = I \\ &= I - I^2 \\ &= I(1 - I) \\ &= \frac{\pi}{4} \left(1 - \frac{\pi}{4} \right). \end{aligned}$$

Cette variance est estimée par $\widehat{\sigma}_n^2 = \widehat{I}_n - \widehat{I}_n^2$ et d'après ce qui précède

$$\sqrt{n} \left(\frac{\widehat{I}_n - I}{\sqrt{\widehat{I}_n - \widehat{I}_n^2}} \right) \xrightarrow[n \rightarrow \infty]{\mathcal{L}} \mathcal{N}(0, 1).$$

Ce qui permet de définir un intervalle de confiance (IC) asymptotique de niveau $(1 - \alpha)$.

Proposition 3.3.2 (Intervalle de Confiance) Soit $\alpha \in (0, 1)$ fixé. Un IC de niveau asymptotique $(1 - \alpha)$ est

$$\left[\widehat{I}_n - \Phi^{-1}(1 - \alpha/2) \frac{\widehat{\sigma}_n}{\sqrt{n}}; \widehat{I}_n + \Phi^{-1}(1 - \alpha/2) \frac{\widehat{\sigma}_n}{\sqrt{n}} \right],$$

où $\Phi^{-1}(1 - \alpha/2)$ désigne le quantile d'ordre $1 - \alpha/2$ de la loi normale centrée réduite.

3.3.3 Erreurs dues à la méthode de Monte-Carlo

Lors de l'utilisation de l'approche de Monte-Carlo, deux types d'effets sont présents dans un échantillon aléatoire : celui lié à l'ensemble des valeurs qui représente l'effet ensembliste et celui lié à leur séquence (effet séquentiel).

- Effet ensembliste (erreurs indésirables) : ce sont des erreurs dues à l'ensemble des valeurs produites par la loi uniforme de la méthode de Monte-Carlo. Ceci est caractérisé par la différence entre l'échantillon généré et la distribution théorique.
- Effet séquentiel (erreurs désirables) : Ce sont des erreurs dues à la suite de cet ensemble de valeurs. Ceci est caractérisé par l'ordre dans lequel les valeurs de la loi uniforme se présentent.

La méthode de Monte-Carlo présente une vitesse de convergence relativement lente. En effet, du fait de la normalité asymptotique, on a

$$|\widehat{I}_n - I| < \Phi^{-1}(1 - \alpha/2) \frac{\sigma}{\sqrt{n}},$$

avec une probabilité égale à $(1 - \alpha)$.

L'erreur maximale d'approximation de I par \widehat{I}_n de la méthode de Monte-Carlo est définie par :

$$M_n = \Phi^{-1}(1 - \alpha/2) \frac{\sigma}{\sqrt{n}}.$$

Cette erreur se réduit au fur et à mesure que n augmente mais cette réduction est assez lente, de l'ordre $\frac{\sigma}{\sqrt{n}}$.

Nous pouvons voir que la vitesse de convergence de la méthode de Monte-Carlo est de l'ordre de $\frac{\sigma}{\sqrt{n}}$. Pour améliorer cette méthode, il existe de nombreuses techniques, dites de réduction de variance, qui cherchent à diminuer la valeur de σ^2 sans augmenter significativement le temps de calcul.

3.4 Techniques de Réduction de la Variance

3.4.1 Variable de contrôle

Soit $I = E[\varphi(X)] = \int_C \varphi(x)f(x)dx$ où f est une densité de probabilité d'un V.a X ayant pour support $C \subset \mathbb{R}^d$.

On sait que $\widehat{I}_n = \frac{1}{n} \sum_{i=1}^n \varphi(X_i)$, où les X_i sont i.i.d de densité f , est l'estimateur standard de Monte-Carlo. Remarquons que l'on peut écrire I sous la forme

$$\begin{aligned} I &= E[\varphi(X) - h(X) + h(X)] \\ &= E[\varphi(X) - h(X)] + E[h(X)]. \end{aligned}$$

En choisissant h proche de φ tel que $\mathbb{E}[h(X)]$ peut se calculer analytiquement, soit θ sa valeur et en notant $J = E[\varphi(X) - h(X)]$, l'approximation de J est :

$$\widehat{J}_n = \frac{1}{n} \sum_{i=1}^n (\varphi(X_i) - h(X_i)) , \text{ où les } X_i \text{ i.i.d selon } f.$$

Une autre approximation de I est donnée par :

$$\begin{aligned} \widetilde{I}_n &= \widehat{J}_n + \theta \\ \widetilde{I}_n &= \frac{1}{n} \sum_{i=1}^n (\varphi(X_i) - h(X_i)) + \theta. \end{aligned}$$

Et comme θ est une constante, on a :

$$Var(\widetilde{I}_n) = Var(\widehat{J}_n) = \frac{1}{n} Var[\varphi(X) - h(X)].$$

h étant proche de φ , $Var(\widetilde{I}_n)$ deviendra petite et inférieure à $Var(\widehat{I}_n)$, ce qui fournit un estimateur de I plus précis que \widehat{I}_n .

3.4.2 Echantillonnage préférentiel

On cherche toujours à calculer $I = E[\varphi(X)] = \int_C \varphi(x)f(x)dx$.

Pour tout densité $\tilde{f} > 0$, on peut écrire

$$I = E[\varphi(X)] = \int_C \varphi(x)f(x)dx = \int_C \frac{\varphi(x)f(x)}{\tilde{f}(x)}\tilde{f}(x)dx,$$

$$I = E[\varphi(X)] = E\left[\frac{\varphi(Y)f(Y)}{\tilde{f}(Y)}\right] \text{ où } Y \text{ est un V.a de densité } \tilde{f}.$$

Nous pouvons donc approcher I par :

$$\tilde{I}_n = \frac{1}{n} \sum_{i=1}^n \frac{\varphi(Y_i)f(Y_i)}{\tilde{f}(Y_i)} \text{ où les } Y_i \text{ sont i.i.d selon } \tilde{f}. \quad (3.4.1)$$

Pour choisir \tilde{f} tel que $Var(\tilde{I}_n) \leq Var(\hat{I}_n)$, il suffit de trouver une fonction \tilde{f}_1 proche de $|f \times \varphi|$ et on peut déterminer \tilde{f} par :

$$\tilde{f} = \frac{\tilde{f}_1}{\int \tilde{f}_1(x)dx}.$$

3.4.3 Variable antithétique

On suppose qu'il existe A tel que X et $X - A$ est la même loi. Comme exemples nous pouvons citer la loi uniforme sur $[0, 1]$ (avec $A = 1$), les lois normales $\mathcal{N}(\mu, \sigma^2)$ (avec $A = 2\mu$), les lois de Laplace de paramètre $\lambda > 0$ (de densité $f(x) = \frac{\lambda}{2} e^{-\lambda|x|}$ avec $A = 0$).

Les variables antithétiques aussi appelées variables complémentaires reposent sur ce principe : faire une exécution en utilisant les nombres aléatoires (X_1, \dots, X_n) pour obtenir un échantillon d'observations aléatoires (x_1, \dots, x_n) , puis calculer l'estimateur

$$\tilde{I}_n = \frac{1}{2n} \sum_{i=1}^n (\varphi(X_i) + \varphi(A - X_i)) \text{ où les } X_i \text{ i.i.d selon } f,$$

qui est un estimateur sans biais et plus précis que l'estimateur standard \hat{I}_n car les variables X et $A - X$ sont négativement collérées.

3.4.4 Stratification

La méthode de Stratification est basée sur le principe suivant :

1. partitionner le domaine d'intégration C en J sous-domaines $\Omega_j : C = \bigcup_{j \in J} \Omega_j$. Dans ce

cas on peut réécrire I :

$$I = \mathbb{E}[\varphi(X)] = \sum_{j \in J} \mathbb{P}(\Omega_j) I_j$$

avec pour tout $j \in J$, $I_j = \mathbb{E}[\varphi(X)/X \in \Omega_j]$ et $\sum_{j \in J} \mathbb{P}(\Omega_j) = 1$,

2. pour $j \in J$ simuler n_j points $(X_1^{(j)}, \dots, X_{n_j}^{(j)})$ dans chaque Ω_j avec une probabilité $\mathbb{P}(\Omega_j)$.

L'estimateur est alors :

$$\tilde{I}_n = \sum_{j=1}^J \frac{\mathbb{P}(\Omega_j)}{n_j} \sum_{i=1}^{n_j} \varphi(X_i^{(j)}). \quad (3.4.2)$$

La variance estimée est :

$$\widehat{Var}(\tilde{I}_n) = \sum_{j=1}^J \frac{\mathbb{P}^2(\Omega_j) \tilde{\sigma}_j^2}{n_j}$$

$$\text{avec } \tilde{\sigma}_j^2 = \frac{1}{n_j} \sum_{i=1}^{n_j} \varphi^2(X_{ij}) - \left(\frac{1}{n_j} \sum_{i=1}^{n_j} \varphi(X_{ij}) \right)^2.$$

3.5 Conclusion

Les méthodes de Monte-Carlo ont comme grand avantage, par rapport aux méthodes d'analyse numérique classiques, d'avoir une vitesse de convergence en $o(\frac{1}{\sqrt{n}})$ (pour un échantillon de n points), donc indépendante de la dimension du problème. Néanmoins, il doit exister des suites de nombres telles que la convergence soit plus rapide, en supprimant l'aspect aléatoire. Ceci conduit à s'intéresser aux méthodes dites de Quasi-Monte-Carlo.

Chapitre 4

LES METHODES QUASI-MONTE-CARLO (QMC)

4.1 Introduction

Reprenons le problème d'intégration par rapport à la loi uniforme, autrement dit l'estimation de

$$I = \int_{[0,1]^d} \varphi(x) dx.$$

Dans ce chapitre on cherche à estimer l'intégrale d'une fonction φ définie sur $[0, 1]^d$. La première idée serait d'utiliser la méthode de Monte-Carlo standard. Cependant, il est fréquent que, pour des raisons de coût, la taille de l'échantillon des réalisations de φ soit limitée. La précision obtenue n'est alors plus toujours satisfaisante. La méthode de Quasi-Monte-Carlo (QMC) ayant (nous allons le voir) une meilleure vitesse de convergence que Monte-Carlo standard, pour un même nombre de points, sa précision devrait être meilleure. Cependant, cette vitesse de convergence dépend de la dimension. Plus précisément, lorsque la dimension est trop élevée, la précision de Monte-Carlo est en théorie meilleure que celle de QMC. Nous discutons dans ce chapitre des méthodes alternatives à celles de Monte-Carlo connues sous le nom de méthodes QMC. Celles-ci diffèrent de la méthode de Monte-Carlo classique dans le sens où elles ne sont pas basées sur la génération de suites pseudo-aléatoires mais sur des ensembles de points engendrés de façon déterministe et possédant des propriétés fortes d'équirépartition.

4.2 Principe de la méthode

Une des façons d'améliorer les méthodes de type Monte-Carlo est de renoncer au caractère aléatoire des tirages et de trier les points de façon "plus ordonnée", i.e d'essayer de retirer les erreurs liées à l'ensemble des valeurs générées par la méthode de Monte-Carlo. En fait cette méthode approche I par :

$$\int_{[0,1]^d} \varphi(x) dx \approx \frac{1}{n} \sum_{i=1}^n \varphi(x_i),$$

où $(x_i)_{i>0} \subset [0, 1]^d$, est choisie de façon plus uniformément distribuée qu'un ensemble de points aléatoires, on parle dans ce cas de la méthode de QMC. Comme nous voulons qu'il ait asymptotiquement convergence quand $n \rightarrow +\infty$, la suite $(x_i)_{i>0}$ doit être équirépartie, d'où la notion de discrépance qui sera décrite ci-après. Les méthodes QMC peuvent être décrites comme des versions déterministes des méthodes de Monte-Carlo, dans le sens où

les échantillons aléatoires dans les méthodes de Monte-Carlo sont remplacés par des points déterministes bien choisis. L'intérêt particulier des méthodes QMC porte sur les mesures d'uniformité de la distribution et sur le développement des suites bien réparties.

4.3 Notion de discrédance

Soit un ensemble de points $(x_1, \dots, x_n) \in [0, 1]^d$, et un sous-ensemble $G \subset [0, 1]^d$. On définit la fonction $S_n(G)$ par le nombre de points $x_i \in G, \forall i = 1, \dots, n$. Pour chaque $\gamma = (\gamma_1, \dots, \gamma_d) \in [0, 1]^d$, nous introduisons une région rectangulaire $G_\gamma = [0, \gamma_1] \times \dots \times [0, \gamma_d]$ avec le volume d -dimensionnel $\gamma_1 \times \dots \times \gamma_d$. La discrédance des points x_1, \dots, x_n est donnée par :

$$D(x_1, \dots, x_n) = \sup_{\gamma \in [0, 1]^d} \left| \frac{S_n(G_\gamma)}{n} - \gamma_1 \times \dots \times \gamma_d \right|.$$

Une autre définition simple et basique de la discrédance est donnée pour comprendre en quoi elle répond à nos attentes de remplissage de l'espace. La discrédance mesure l'écart entre une distribution de points uniforme et une distribution de points donnée ; en d'autres termes elle permet de mesurer l'irrégularité de la distribution.

Si $d = 1$ on considère \widehat{F}_n la fonction de répartition empirique des points x_1, \dots, x_n , on définit la discrédance par :

$$D_n^*(X) = \sup_{x \in [0, 1]} \left| \widehat{F}_n(x) - F_U(x) \right|,$$

où F_U est la fonction de répartition de la loi uniforme sur $[0, 1]$ et $X = (X_1, \dots, X_n)$.

Il existe de nombreux résultats de majoration de la discrédance. Le résultat le plus fameux est connu sous le nom d'inégalité de Koksma-Hlawka aussi appelé théorème de bornes qui donne une définition précise de l'erreur absolue liée à la discrédance.

Definition 4.3.1 Une suite $\xi = (\xi_n)_{n \geq 1}$ de $[0, 1]^d$ est à faible discrédance si elle vérifie :

$$D_n^*(\xi) = o((\log n)^d/n).$$

Definition 4.3.2 (Variation de Hardy-Krausse) La variation au sens de Hardy-Krausse d'une fonction $\varphi : [0, 1]^d \rightarrow \mathbb{R}$ de classe C^d est

$$V(\varphi) = \sum_{j=1}^d \sum_{i_1 < \dots < i_j} \int_{[0, 1]^j} \left| \frac{\partial^j \varphi}{\partial x_{i_1} \dots \partial x_{i_j}}(x_{i_1}, \dots, x_{i_j}) \right| dx_{i_1} \dots dx_{i_j},$$

avec x_{i_1}, \dots, x_{i_j} le point de \mathbb{R}^d dont toutes les coordonnées valent 1 sauf celles de rangs x_{i_1}, \dots, x_{i_j} .

Ainsi, lorsque $d = 1$, on a tout simplement

$$V(\varphi) = \int_0^1 |\varphi'(x)| dx = \|\varphi'\|_1.$$

Pour $d = 2$, on a :

$$V(\varphi) = \int_0^1 \left| \frac{\partial \varphi}{\partial x_1}(x_1, 1) \right| dx_1 + \int_0^1 \left| \frac{\partial \varphi}{\partial x_2}(1, x_2) \right| dx_2 + \int \int_{[0, 1]^d} \left| \frac{\partial^2 \varphi}{\partial x_1 \partial x_2}(x_1, x_2) \right| dx_1 dx_2.$$

Il est clair qu'en dimension supérieure, estimer la variation de Hardy-Krausse devient vite inextricable. Néanmoins, cette notion intervient de façon cruciale pour majorer la qualité d'un estimateur basé sur une séquence $(\xi_n)_{n \geq 1}$.

Théorème 4.3.1 (Inégalité de Koksma-Hlawka) Pour tout fonction $\varphi : [0, 1]^d \mapsto \mathbb{R}$ et toute suite $(\xi_n)_{n \geq 1}$ de $[0, 1]^d$, on a

$$\left| \frac{1}{n} \sum_{i=1}^n \varphi(\xi_i) - \int_{[0,1]^d} \varphi(x) dx \right| \leq V(\varphi) \times D_n^*(\xi),$$

où $V(\varphi)$ est la variation de φ au sens de Hardy-Krause et $D_n^*(\xi)$ la discrédance de ξ à l'ordre n .

Preuve. Voir Caflish (1998), page 27. ■

4.4 Les suites de QMC

Definition 4.4.1 Soit $b \geq 2$ un entier, la fonction radicale inverse à base b est définie par :

$$\phi_b(n) = \sum_{j=0}^m a_j(n) b^{-j-1} \quad n \geq 0,$$

où a_j est la position du bit en partant de droite $\forall j = 0, \dots, m$, avec $n = a_0 + a_1 b^1 + \dots + a_m b^m$ et

$$m = \begin{cases} 1 + \log_b(n) & \text{si } n \neq 0 \\ 1 & \text{sinon} \end{cases}.$$

4.4.1 Les suites de Van Der Corput

Nous introduisons dans cette section une classe spécifique de suite à discrédance faible en dimension 1, appelée suite de Van der Corput.

Tout entier positif n possède une représentation unique (appelée représentation en base b) comme une combinaison linéaire de puissances positives de b à coefficients dans $\{0, \dots, b-1\}$

$$n = \sum_{j=0}^{\infty} a_j(n) b^j. \tag{4.4.1}$$

Cette décomposition comporte un nombre fini de termes non nuls. La fonction radicale inverse $\phi_b(n)$ transforme n en un point dans $[0, 1]$ qui est appelée suite de Van Der Corput en base b donnée par :

$$\xi_n = \phi_b(n) = \sum_{j=0}^{\infty} a_j(n) b^{-j-1}.$$

La discrédance de la suite est donnée par :

$$D_n^*(\xi) = o\left(\frac{\ln(n)}{n}\right). \tag{4.4.2}$$

Exemple : Suites de Van Der corput en dimension 1

On se fixe un nombre entier, par exemple 2, et on part de la suite des entiers naturels que l'on traduit en base 2. En appliquant la décomposition (4.4.1) et la fonction radicale inverse à base 2, nous établissons le tableau suivant :

n	en base 2	$\phi_b(n)$
1	1	$\frac{1}{2}$
2	10	$\frac{1}{4}$
3	11	$\frac{3}{4}$
4	100	$\frac{1}{8}$
5	101	$\frac{5}{8}$
6	110	$\frac{3}{8}$
7	111	$\frac{7}{8}$
...

Ce qui retraduit la suite de Van Der Corput de base 2

$$(\xi_n)_{n \geq 1} = \left(\frac{1}{2}, \frac{1}{4}, \frac{3}{4}, \frac{1}{8}, \frac{5}{8}, \frac{3}{8}, \frac{7}{8}, \dots \right).$$

4.4.2 Les suites de Halton

Elles correspondent à la généralisation des suites de Van der Corput en dimension d supérieur à 1. La première dimension de la suite de Halton est la suite de Van der Corput à base 2 et la seconde dimension de Halton est la suite de van der Corput à base 3. La dimension d de la suite de Halton est le d^{eme} nombre premier, qu'utilise la suite de Van Der Corput comme base. On considère d nombre premiers entre eux b_1, \dots, b_d et on construit les d suites de Van der Corput $\xi^{(1)}, \dots, \xi^{(d)}$ de bases respectives b_1, \dots, b_d . La suite de Halton associé est alors

$$\xi = (\xi_n)_{n \geq 1} = ((\xi_n^{(1)}, \dots, \xi_n^{(d)}))_{n \geq 1}.$$

La discrédance de la suite de Halton vérifie :

$$D_n^*(\xi) \leq \frac{1}{n} \prod_{i=1}^d \frac{b_i \ln(b_i n)}{\ln(b_i)} = o\left(\frac{(\ln(n))^d}{n}\right). \quad (4.4.3)$$

Exemple

Les 10 premiers nombres de la suite de Halton dans les trois premières dimensions sont donnés dans le tableau suivant :

n	dim=1 (base=2)	dim=2 (base=3)	dim=3 (base=5)
1	1/2	1/3	1/5
2	1/4	2/3	2/5
3	3/4	1/9	3/5
4	1/8	4/9	4/5
5	5/8	7/9	1/25
6	3/8	2/9	6/25
7	7/8	5/9	11/25
8	1/16	8/9	16/25
9	9/16	1/27	21/25
10	5/16	10/27	2/25

La suite de Halton associé est :

$$\xi = \left(\left(\frac{1}{2}, \frac{1}{3}, \frac{1}{5} \right), \left(\frac{1}{4}, \frac{2}{3}, \frac{2}{5} \right), \left(\frac{3}{4}, \frac{1}{9}, \frac{3}{5} \right), \left(\frac{1}{8}, \frac{4}{9}, \frac{4}{5} \right), \left(\frac{5}{8}, \frac{7}{9}, \frac{1}{25} \right), \left(\frac{3}{8}, \frac{2}{9}, \frac{6}{25} \right), \left(\frac{7}{8}, \frac{5}{9}, \frac{11}{25} \right), \left(\frac{1}{16}, \frac{8}{9}, \frac{16}{25} \right), \dots \right).$$

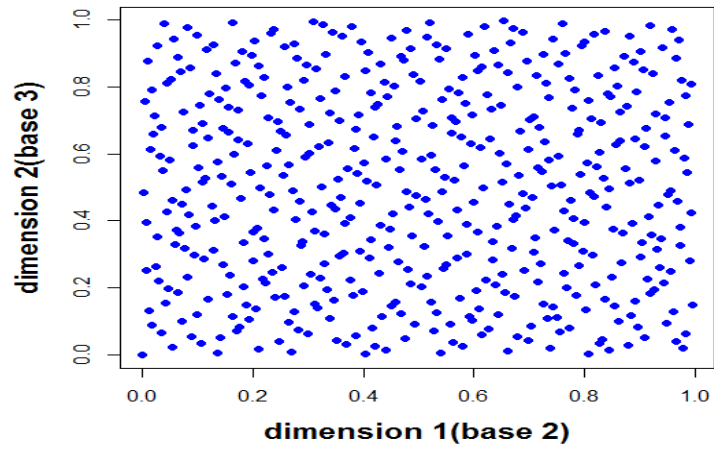


FIGURE 4.1 – Représentation graphique des 500 premiers nombres de la suite de Halton.

Les deux figures suivantes représentent la projection des points de la suite de Halton dans le carré unité pour $n = 1000$.

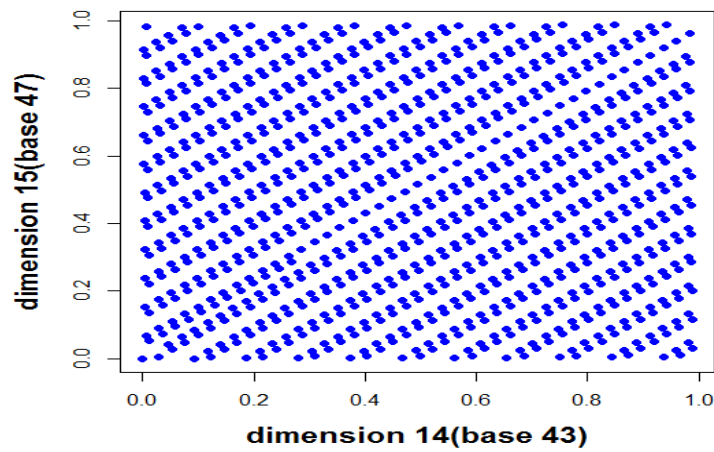


FIGURE 4.2 – Les 1000 nombres de la suite de Halton en dimension 14 et 15.

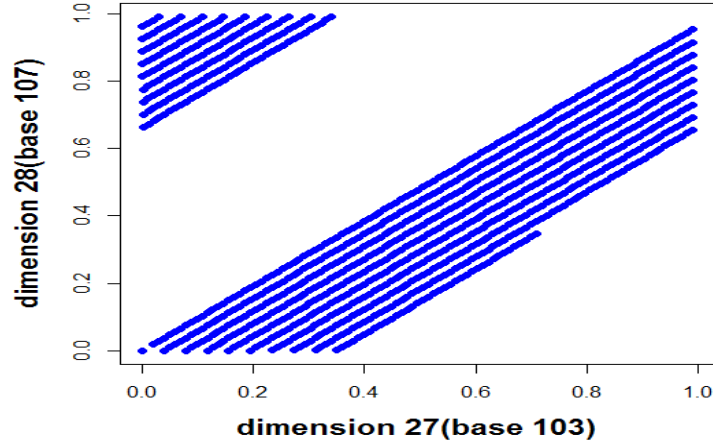


FIGURE 4.3 – Les 1000 nombres de la suite de Halton en dimension 27 et 28.

Malgré l'uniformité de la distribution dans le carré unité, ces suites montrent une certaine faiblesse qui se pose dans les dimensions élevées. En effet, lorsque l'on projette les points d'une telle suite dans le carré unité, nous pouvons observer des diagonales décalées à chaque itération, qui laissent de la place à de nombreuses zones lacunaires dans le carré unité. Ce phénomène résulte de la régularité des sous-suites de longueur b des suites de Van der Corput. Le problème avec les suites de Halton réside du fait que les points dans les dimensions successives sont fortement corrélés.

4.4.3 Les suites de Hammersley

Une suite de Hammersley en dimension d est construite à partir d'un terme dépendant du nombre de points et d'une suite de Halton en dimension $(d-1)$. Soit $b_1, \dots, b_{d-1} \in \{2, 3, 5, \dots\}$ des entiers positifs premiers entre eux deux à deux. L'ensemble de points N de Hammersley en dimension $d > 2$ est :

$$\xi_n = \left(\frac{n}{N}, \phi_{b_1(n)}, \dots, \phi_{b_{d-1}(n)} \right); n = 0, \dots, N - 1, \quad (4.4.4)$$

où $\phi_{b_i}, i = 1, \dots, d-1$ sont la fonction radicales inverses à base b_i . La discrépance de la suite de Hammersley est telle que :

$$D_n^*(\xi) \leq c_d \left(\frac{(\ln(N))^{d-1}}{N} \right). \quad (4.4.5)$$

Il est impossible de rajouter des points supplémentaires à ces suites sans pour autant perturber la discrépance. Dans le cas où le nombre de points à générer n'est pas connu à l'avance, il est déconseillé d'utiliser une suite de Hammersley. Aussi, ces suites ont perdu l'aspect itératif des suites de Halton qui permet de rajouter des points facilement. Par contre l'avantage des ensembles de Hammersley sur les suites de Halton est le taux de convergence de la borne pour la discrépance qui est légèrement plus rapide car elle dépend de $(d-1)$ bases au lieu de d .

4.4.4 Les suites de Sobol

Les nombres de directions

Definition 4.4.2 *Un ensemble de nombres de directions v_i est défini comme étant*

$$v_i = \frac{m_i}{2^i}, i = 1, \dots, w. \quad (4.4.6)$$

Les m_i peuvent être choisis librement à condition que chaque m_i soit impair et inférieur à 2^i .

Par exemple, $n = 5$ est 101 dans la base 2, on peut prendre $v_1 = 0.1$, $v_2 = 0.11$, $v_3 = 0.111$.

Le code Gray

Definition 4.4.3 *Le code de Gray (voir Sobol, 1967) est défini comme suit :*

$$G(j) = j \oplus \lfloor \frac{j}{2} \rfloor = (\dots j_3 j_2 j_1)_2 \oplus (\dots j_4 j_3 j_2)_2, \quad (4.4.7)$$

où \oplus désigne l'addition modulo, appelé aussi le "ou" exclusif.

Le polynôme primitif

Definition 4.4.4 *Soit X une variable aléatoire et b_i , $i = 1, \dots, l - 1$ les coefficients prenant pour valeurs 0 ou 1. Un polynôme primitif est défini comme suit :*

$$P(X) = X^l + b_1 X^{l-1} + b_2 X^{l-2} + \dots + b_{l-1} X + 1, \quad (4.4.8)$$

avec le $\text{pgcd}(b_i, i = 1, \dots, l - 1) = 1$ et $P(X)$ est irréductible sur \mathbb{Z}_2 .

La suite de Sobol est l'une des suites à faible discrédance standard et la plus utilisée dans les applications QMC. Les suites de Sobol de dimension d utilisent le nombre premier 2 comme base. La suite de Sobol en $d = 1$ est la suite de van der Corput en base 2, et dans les grandes dimensions, la suite est présentée comme des permutations de la suite en première dimension. Sa construction est basée sur les polynômes primitifs dans le domaine \mathbb{Z}_2 définis en (4.4.8) et les opérations définies en (4.4.7). Les permutations générées à partir de ces polynômes irréductibles permettent de développer certaines corrélations. Les permutations de la suite de Sobol dépendent des nombres de direction définis en (4.4.6).

Pour construire une suite de Sobol en dimension d , nous procédons comme suit :

1. Choisir un entier n au hasard, par exemple, $n = 2$. Ce nombre définit le point de départ de la suite.
2. Calculer le Code Gray (4.4.7) de n , soit $G(n)$.
3. Transformer $G(n)$ à sa représentation binaire ($G(2) = 3 = 1 * 2^1 + 1 * 2^0 = 11$).
4. Sommer bit par bit les nombres de directions associés aux chiffres de $G(n)$ dans la base binaire, dans cet exemple on compte de droite à gauche les bits de $G(n)$ différents de zéro qui sont v_1 et v_2 , par conséquent $y(2; d) = 0.1 \oplus 0.11 = 0.01$.
5. Transformer le nombre obtenu en un nombre décimal en multipliant chaque bit de cette représentation binaire par 2^l , où l est la position du bit dans la partie décimale du nombre, et sommer de gauche à droite les termes. Le résultat est le premier terme de la suite de Sobol en dimension d :

$$\phi_2^d(n) = 0 * 2^{-1} + 1 * 2^{-2} = 0.250.$$

La discrédance de la suite de Sobol est telle que :

$$D_n^*(\phi_2^d) \leq c_d \frac{(\log(n))^d}{n} + o\left(\frac{(\log(n))^{d+1}}{n}\right). \quad (4.4.9)$$

Exemple

Ce tableau illustre la suite de Sobol dans les 3 premières dimensions en base deux.

n	dim=1 (base=2)	dim=2 (base=2)	dim=3 (base=2)
1	1/2	1/2	1/2
2	1/4	1/4	3/4
3	3/4	3/4	1/4
4	3/8	3/8	5/8
5	7/8	7/8	1/8
6	5/8	1/8	3/8
7	1/8	5/8	7/8
8	3/16	5/16	5/16

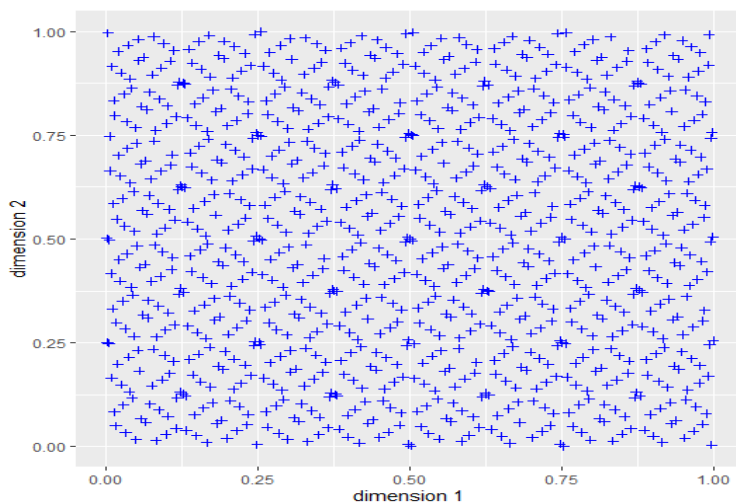


FIGURE 4.4 – Représentation graphique des 1000 premiers nombres de la suite Sobol en dimension 1 et 2.

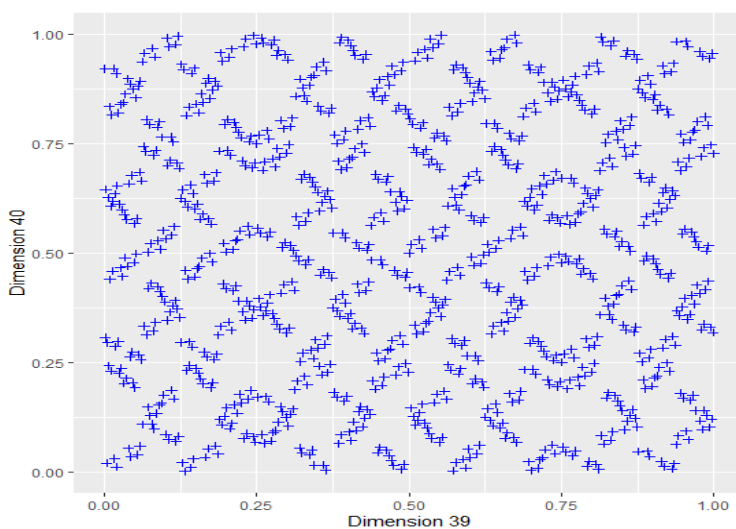


FIGURE 4.5 – Représentation graphique des 1000 premiers nombres de la suite Sobol en dimension 39 et 40.

4.4.5 Les suites de Faure

Soient d et b deux entiers avec b premier et $1 \leq d \leq b$. On choisit d entiers distincts r_1, \dots, r_d compris entre 1 et b . On écrit pour tout $n \in \mathbb{N}$, son développement (4.4.1) en base b . On note alors

$$y_j^{(i)}(n) = \sum_{l=j-1}^{\infty} \binom{l}{j-1} r_i^{l-j+1} a_l(n) \bmod b.$$

où $\binom{l}{j-1} = \frac{l!}{(j-1)!(l-j+1)!}$ et

$$\phi_b^i(n) = \sum_{j=0}^{\infty} y_j^{(i)}(n) b^{-j} \in [0, 1]. \quad (4.4.10)$$

Puisque les termes $a_l(n)$ sont nuls à partir d'un certain rang alors les sommes précédentes sont finies. La suite de Faure en base b est la suite

$$\phi^d = \{\phi_b^1(n), \dots, \phi_b^d(n), n \geq 0\} \subset [0, 1]^d. \quad (4.4.11)$$

La discrédance de ces suites vérifie l'inégalité :

$$D_n^*(\phi^d) \leq C_d \frac{(\log n)^d}{n} + o\left(\frac{(\log n)^{d-1}}{n}\right), \quad (4.4.12)$$

pour tout $n \geq 1$, où

$$C_d = \begin{cases} \frac{3}{16(\log 2)^2} & \text{si } b = d = 2 \\ \frac{1}{d!} \left(\frac{b-1}{2-\log b}\right)^d & \text{si } d \geq 2 \text{ et } b \geq d \text{ impair premier} \end{cases}.$$

Exemple

Ce tableau illustre la suite de Faure dans les 3 premières dimensions en base 3.

n	dim=1 (base=3)	dim=2 (base=3)	dim=3 (base=3)
1	1/3	1/3	1/3
2	2/3	2/3	2/3
3	1/9	4/9	7/9
4	4/9	7/9	1/9
5	7/9	1/9	4/9
6	2/9	8/9	5/9
7	5/9	2/9	8/9
8	8/9	5/9	2/9

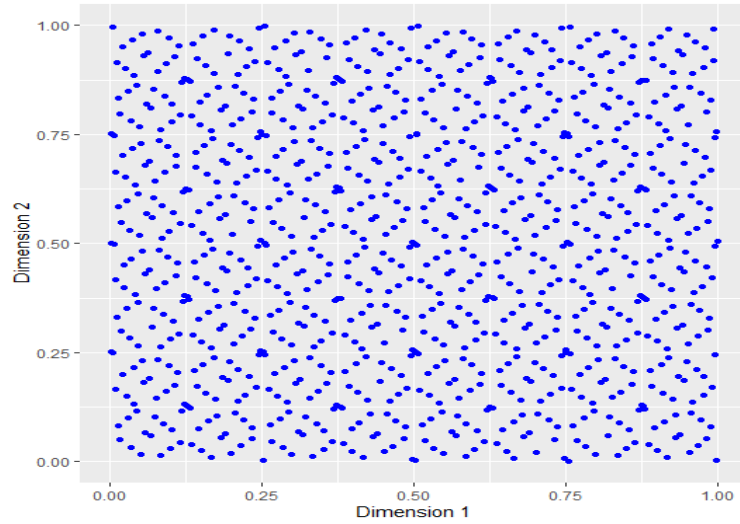


FIGURE 4.6 – Représentation graphique des 1000 premiers nombres de la suite de Faure en dimension 1 et 2.

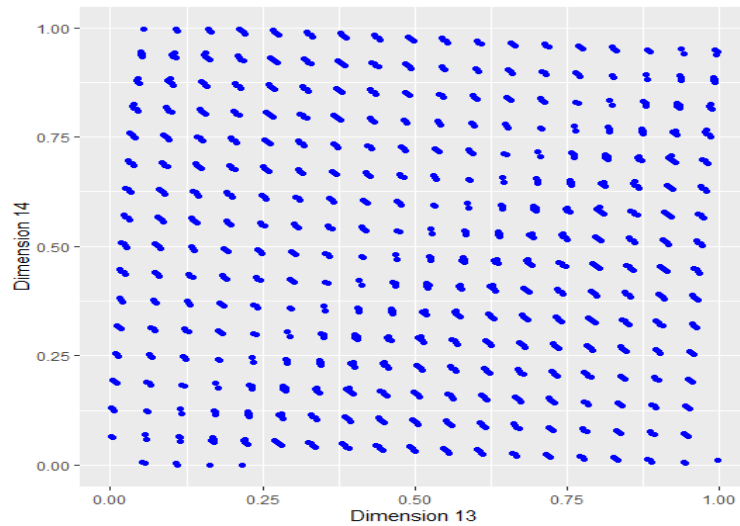


FIGURE 4.7 – Représentation graphique des 1000 premiers nombres de la suite de Faure en dimension 13 et 14.

Les problèmes des suites de Faure surviennent avec les dimensions élevées car elles exigent un nombre premier assez grand comme base et provoquent plus de corrélation entre les composantes voisines de la suite.

4.5 Conclusion

Revenons au problème d'intégration

$$I = \int_{[0,1]^d} \varphi(x) dx$$

et que l'on se donne une suite $(\xi_n)_{n \geq 1}$ à discrétance faible, par exemple une suite de Halton basée sur les d premiers nombres premiers, alors l'estimateur

$$\hat{I}_n = \frac{1}{n} \sum_{i=1}^n \varphi(\xi_i)$$

est un estimateur Quasi-Monte-Carlo de l'intégrale I , dont l'erreur (déterministe) est donc en $o((\log n)^d/n)$. Dans ce chapitre, nous avons abordé la méthode d'intégration numérique de Quasi-Monte-Carlo (QMC) qui peut être considérée comme une version déterministe de la méthode de Monte-Carlo. Nous avons donné les fondements et propriétés théoriques de l'approche Quasi Monte-Carlo. Elle repose sur un échantillonnage déterministe du domaine d'intégration de la fonction étudiée par des suites de points uniformes, appelées suites à discrétion faible. L'utilisation de ces suites garantit un taux de convergence asymptotique théorique d'un ordre minimal, ce qui est nettement meilleur que le taux de convergence de la méthode de Monte-Carlo. Le principal inconvénient de la méthode QMC est que sa vitesse de convergence dépend de la dimension d du problème. Autrement dit, elle perd de son efficacité lorsque la dimension du problème augmente, contrairement à la méthode de Monte-Carlo dont la vitesse de convergence ne dépend que du nombre de simulations réalisées.

Chapitre 5

MISE EN OEUVRE PRATIQUE ET COMPARAISON

5.1 Introduction

À travers ce chapitre nous allons voir l'utilisation des notions et principaux résultats obtenus dans les chapitres précédents par une application numérique. Dans notre cas, l'application portera sur l'intégrale suivante :

$$I(k, d) = \int_{[0,1]^d} \varphi(x) dx = \int_{[0,1]^d} \left(\prod_{j=1}^d \frac{k\pi}{2} \sin(k\pi x_j) \right) dx_1 \cdots dx_d, \quad (5.1.1)$$

où k et d deux entiers naturels non nuls. En premier lieu, nous allons estimer cette intégrale en utilisant les méthodes déterministes et de Monte-Carlo en dimension $d = 1, 2, 3$ et faire une comparaison de ces méthodes. en second lieu nous allons générer des ensembles de points engendrés de façon déterministe avec les suites de Halton, Sobol et Faure pour estimer l'intégrale $I(k, d)$ et comparer l'efficacité ces suites pour une dimension d fixé. Cette comparaison d'efficacité nous permettra de choisi la suite optimal et enfin de pouvoir comparer les méthodes QMC et Monte-Carlo en dimension d .

5.2 Comparaison entre les méthodes déterministes et la méthode de Monte-Carlo

Commençons par calculer la valeur exacte de l'intégrale (5.1.1).

Les variables $X_j, j = 1, \dots, d$ étant séparables, l'intégrale du produit dans (5.1.1) peut s'écrire comme produit d'intégrales, ce qui donne pour tout $d \in \mathbb{N}^*$:

$$\begin{aligned} I(k, d) &= \left(\int_0^1 \frac{k\pi}{2} \sin(k\pi x) dx \right)^d \\ &= \left(\left[-\frac{1}{2} \cos(k\pi x) \right]_0^1 \right)^d \\ &= \left(\frac{1 - (-1)^k}{2} \right)^d \\ I(k, d) &= \begin{cases} 1 & \text{si } k \text{ impair} \\ 0 & \text{sinon } k \text{ pair} \end{cases} \end{aligned}$$

Soit I_{ap} l'approximation de l'intégrale (5.1.1). On s'intéresse ici à l'erreur absolue d'approximation de $I(k, d)$ par I_{ap} , c'est à dire

$$\text{Erreur absolue} = |I_{ap} - I(k, d)|.$$

5.2.1 Erreur d'approximation en dimension 1

$$I(k, 1) = \int_0^1 \frac{k\pi}{2} \sin(k\pi x) dx.$$

Les approximations de $I(k, 1)$ sont données par :

- Méthode des rectangles : $I_{ap} = \frac{k\pi}{2n} \sum_{i=1}^n \sin\left(\frac{k\pi i}{n}\right)$.
- Méthode des trapèzes : $I_{ap} = \frac{k\pi}{4n} \sum_{i=1}^n \left[\sin\left(\frac{k\pi(i-1)}{n}\right) + \sin\left(\frac{k\pi i}{n}\right) \right]$.
- Méthode de Simpson : $I_{ap} = \frac{k\pi}{12n} \sum_{i=1}^n \left[\sin\left(\frac{k\pi(i-1)}{n}\right) + 4 \sin\left(\frac{k\pi(2i-1)}{2n}\right) + \sin\left(\frac{k\pi i}{n}\right) \right]$.
- Méthode de Monte-Carlo : on a

$$\begin{aligned} I(k, 1) &= \int_{\mathbb{R}} \frac{k\pi}{2} \sin(k\pi x) \mathbb{1}_{[0,1]}(x) dx \\ &= \mathbb{E} \left[\frac{k\pi}{2} \sin(k\pi X) \right], \text{ où } X \rightsquigarrow \mathcal{U}_{[0,1]}. \end{aligned}$$

$$\text{Donc } I_{ap} = \frac{k\pi}{2n} \sum_{i=1}^n \sin(k\pi x_i) \text{ où } x_1, \dots, x_n \text{ des valeurs simulées de } X.$$

Les résultats obtenus par la simulation de cette intégrale pour $k = 1, 2$ et pour $n = 1000$ par les méthodes déterministes et de Monte-Carlo sont établis dans les deux tableaux (Table 5.1) et (Table 5.2).

Méthodes	I_{ap}	Erreur absolue
Valeur exacte de I	1	0
Rectangles	0.9999992	0.0000008
Trapèzes	0.9999992	0.0000008
Simpson	1	0
Monte-Carlo	1.014756	0.014756

TABLE 5.1 – Approximations et erreurs absolues de I par I_{ap} pour $k = 1$

Nous constatons, pour $n = 1000$, qu'avec la méthode de Simpson on retrouve la valeur exacte de $I(1, 1)$ et que la méthode des rectangles et celle des trapèzes ont la même approximation avec une erreur égale à 8.10^{-7} . Ces méthodes déterministes approximent mieux $I(1, 1)$ que la méthode de Monte-Carlo qui a une erreur de $1.47568 \cdot 10^{-2}$.

Nous donnons une représentation graphique des erreurs absolues des différentes méthodes en fonction de n pour $k = 1$.

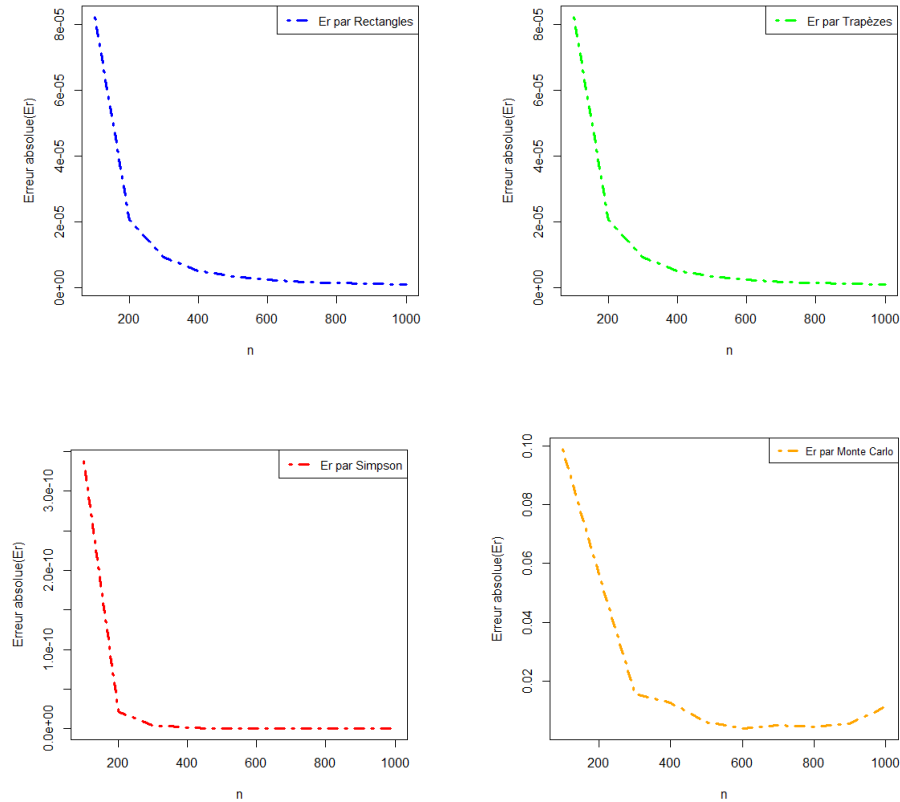


FIGURE 5.1 – Représentation graphique des erreurs en fonction de n .

Pour $k = 1$ nous remarquons que l'erreur pour la méthode de Simpson converge plus rapidement vers 0 quand n augmente ensuite vient celles des méthodes des rectangles et des trapèzes qui convergent en même temps vers 0 quand n augmente et enfin vient celle de Monte-Carlo qui converge lentement vers 0 lorsque n augmente.

Méthodes	I_{ap}	Erreur absolue
Valeur exacte de I	0	0
Rectangles	$-9.284163 \cdot 10^{-17}$	$9.284163 \cdot 10^{-17}$
Trapèzes	$-0.1089229 \cdot 10^{-17}$	$0.1089229 \cdot 10^{-17}$
Simpson	$-0.02.27698 \cdot 10^{-17}$	$0.0227698 \cdot 10^{-17}$
Monte-Carlo	0.1759462	0.1759462

TABLE 5.2 – Approximations et erreurs absolues de I par I_{ap} pour $k = 2$

Pour $k = 2$ nous remarquons que les erreurs obtenues avec les méthodes déterministes sont très faibles de l'ordre de 10^{-17} par rapport à celle de Monte-Carlo qui est de l'ordre de 10^{-1} . Ainsi nous constatons que les erreurs obtenus pour $k = 2$ sont largement inférieur à celles obtenues pour $k = 1$.

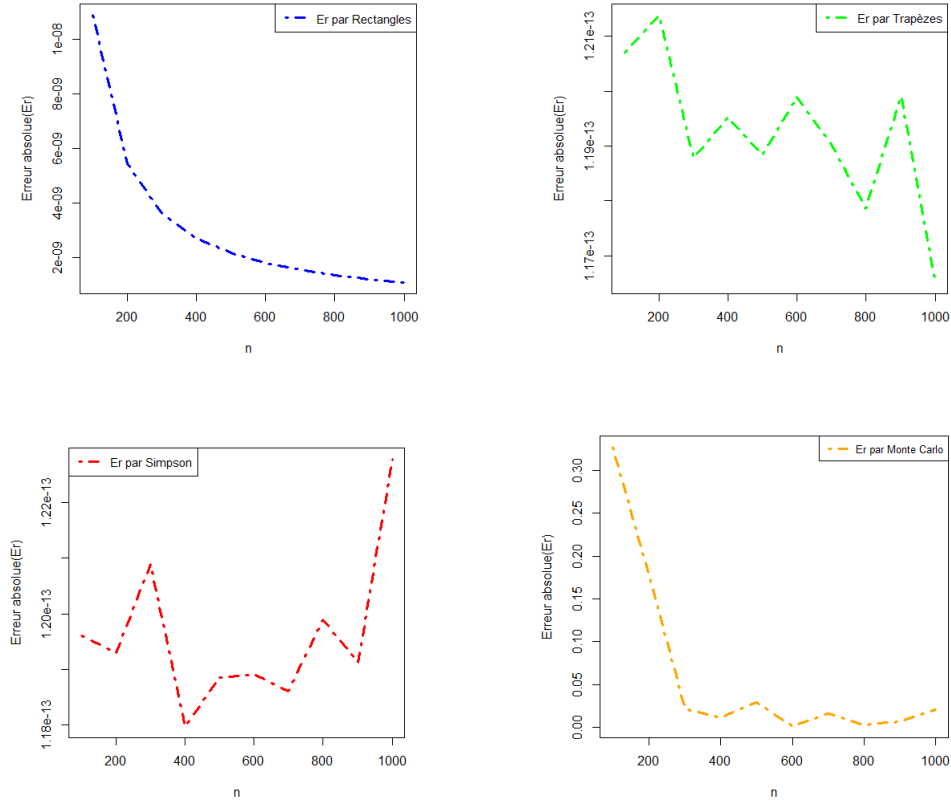


FIGURE 5.2 – Représentation graphique des erreurs en fonction de n .

La Figure (5.2) nous montrent que les erreurs fournies par les méthodes déterministes sont plus faibles que celle de Monte-Carlo. C'est qui confirme qu'en dimension 1 les méthodes déterministes sont préférables que la méthode de Monte-Carlo.

5.2.2 Erreur d'approximation en dimension 2

$$I(k, 2) = \int_0^1 \int_0^1 \frac{k^2 \pi^2}{4} \sin(k\pi x) \sin(k\pi y) dx dy.$$

En dimension 2 les approximations de $I(k, 2)$ sont données par :

- Méthode des rectangles : $I_{ap} = \frac{k^2 \pi^2}{4n^2} \sum_{i,j=1}^n \sin\left(\frac{k\pi i}{n}\right) \sin\left(\frac{k\pi j}{n}\right)$.
- Méthode des trapèzes : $I_{ap} = \frac{k^2 \pi^2}{8n^2} \sum_{i,j=1}^n \left[\sin\left(\frac{k\pi(i-1)}{n}\right) \sin\left(\frac{k\pi(j-1)}{n}\right) + \sin\left(\frac{k\pi i}{n}\right) \sin\left(\frac{k\pi j}{n}\right) \right]$.
- Méthode de Simpson :
 $I_{ap} = \frac{k^2 \pi^2}{24n^2} \sum_{i,j=1}^n \left[\sin\left(\frac{k\pi(i-1)}{n}\right) \sin\left(\frac{k\pi(j-1)}{n}\right) + 4 \sin\left(\frac{k\pi(2i-1)}{2n}\right) \sin\left(\frac{k\pi(2j-1)}{2n}\right) + \sin\left(\frac{k\pi i}{n}\right) \sin\left(\frac{k\pi j}{n}\right) \right]$.
- Méthode de Monte-Carlo

$$\begin{aligned} I(k, 2) &= \int_{\mathbb{R}} \int_{\mathbb{R}} \frac{k^2 \pi^2}{2^2} \sin(k\pi x) \sin(k\pi y) \mathbb{1}_{[0,1]^2}(x, y) dx dy \\ &= \mathbb{E} \left[\frac{k^2 \pi^2}{4} \sin(k\pi X) \sin(k\pi Y) \right], \text{ où } (X, Y) \rightsquigarrow \mathcal{U}_{[0,1]^2}. \end{aligned}$$

Donc $I_{ap} = \frac{k^2 \pi^2}{4n} \sum_{i=1}^n \sin(k\pi x_i) \sin(k\pi y_i)$ où les (x_i, y_i) sont les valeurs simulées de (X, Y) .

Les résultats obtenus par la simulation de cette intégrale pour $k = 1, 2$ et pour $n = 1000$ par les méthodes ci-dessus sont établis dans les deux tableaux ci-après (Table 5.3) et (Table 5.4).

Méthodes	I_{ap}	Erreur absolue
Valeur exacte de I	1	0
Rectangles	0.9999984	0.0000016
Trapèzes	0.9999984	0.0000016
Simpson	1	0
Monte-Carlo	1.01394	0.013946

TABLE 5.3 – Approximations et erreurs absolues de I par I_{ap} pour $k = 1$.

Méthodes	I_{ap}	Erreur absolue
Valeur exacte de I	0	0
Rectangles	$-11.78233 \cdot 10^{-17}$	$11.78233 \cdot 10^{-17}$
Trapèzes	$7.083331 \cdot 10^{-18}$	$7.083331 \cdot 10^{-18}$
Simpson	$2.877846 \cdot 10^{-18}$	$2.877846 \cdot 10^{-18}$
Monte-Carlo	-0.02328906	0.02328906

TABLE 5.4 – Approximations et erreurs absolues de I par I_{ap} pour $k = 2$.

Les résultats obtenus dans les tableaux Table 5.3 et Table 5.4. nous montrent qu'en dimension 2, pour $k = 1$, la méthode de Simpson approxime mieux l'intégrale car pour $n = 1000$ on retrouve la valeur exacte de I , ensuite vient la méthode des trapèzes et des rectangles qui ont une même approximation et enfin vient la méthode de Monte-Carlo. Pour $k = 2$, les méthodes déterministes approximent mieux I que celle de Monte-Carlo dans l'ordre décrit ci-dessus, mais nous remarquons qu'en dimension 2 les erreurs d'approximations sont légèrement plus importantes pour les méthodes déterministes et moyen importante pour Monte-Carlo que celles obtenues en dimension 1 pour les mêmes méthodes.

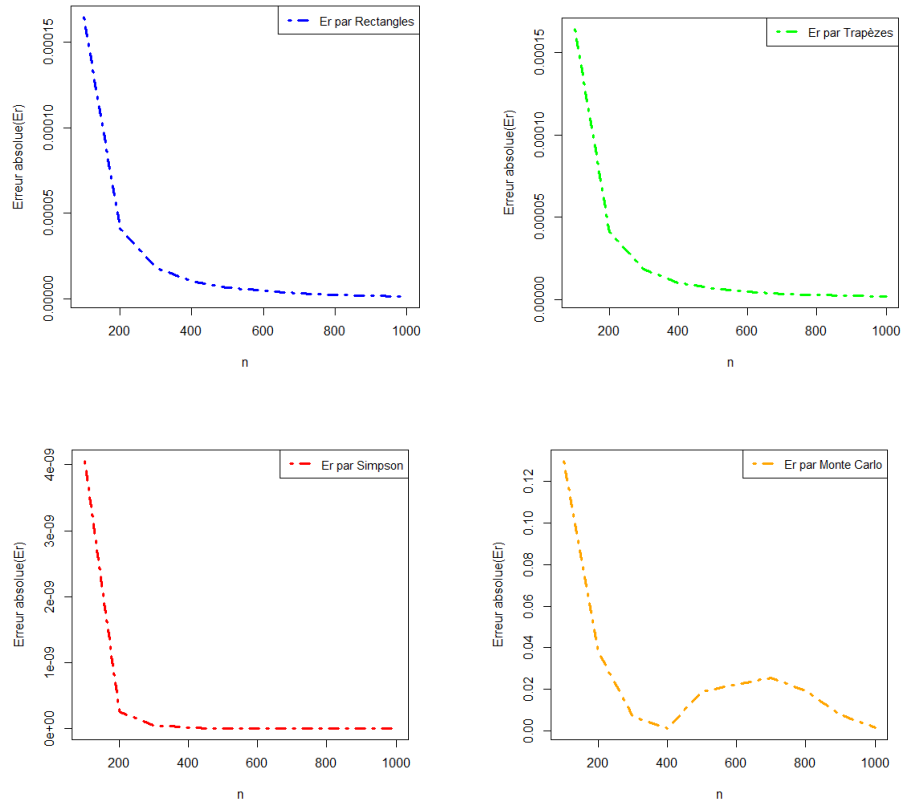


FIGURE 5.3 – Représentation graphique des erreurs en fonction de n en dimension 2.

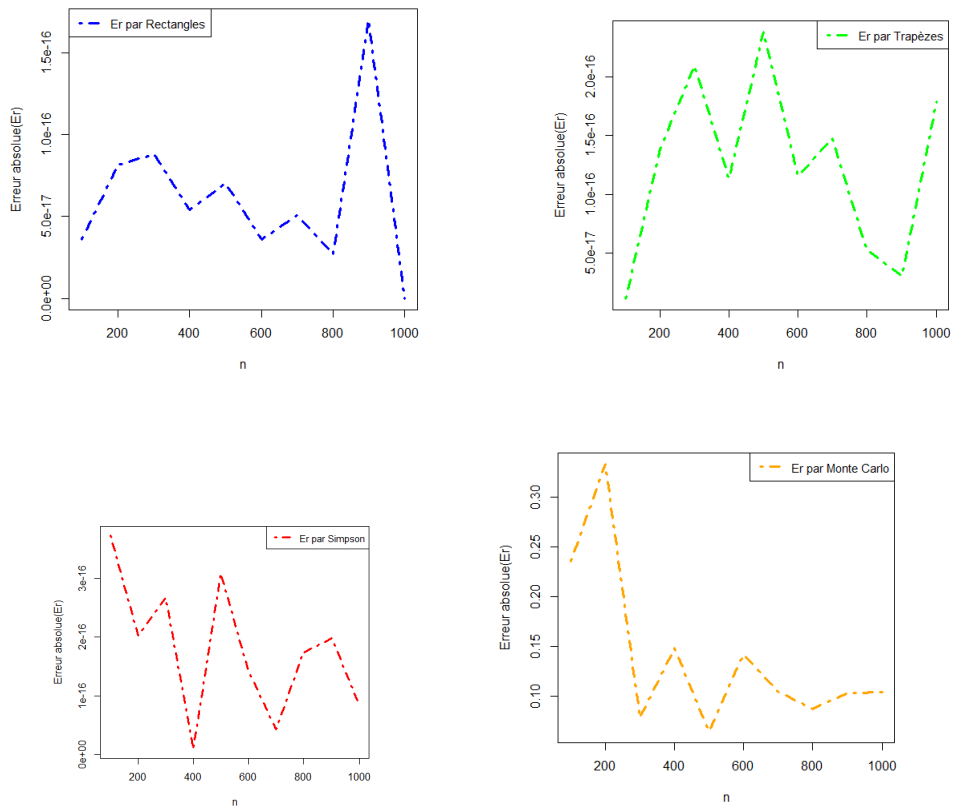


FIGURE 5.4 – Représentation graphique des erreurs en fonction de n pour $k = 2$.

La Figure (5.3) nous montre que pour les méthodes déterministes l'erreur d'approximation converge rapidement vers 0 quand n augmente contrairement à celle de Monte-Carlo qui converge lente et pour la Figure (5.4) nous remarquons que les erreurs dues aux méthodes déterministes sont plus faibles que celle de Monte-Carlo.

5.2.3 Erreur d'approximation en dimension 3

$$I(k, 3) = \int_0^1 \int_0^1 \int_0^1 \varphi(x, y, z) dx dy dz$$

$$\text{avec } \varphi(x, y, z) = \frac{k^3 \pi^3}{8} \sin(k\pi x) \sin(k\pi y) \sin(k\pi z).$$

En dimension 3 les approximations de $I(k, 3)$ sont données par :

- Méthode des rectangles :

$$I_{ap} = \frac{1}{n^3} \sum_{i,j,l=1}^n \varphi\left(\frac{i}{n}, \frac{j}{n}, \frac{l}{n}\right).$$

- Méthode des trapèzes :

$$I_{ap} = \frac{1}{2n^3} \sum_{i,j,l=1}^n \left[\varphi\left(\frac{i-1}{n}, \frac{j-1}{n}, \frac{l-1}{n}\right) + \varphi\left(\frac{i}{n}, \frac{j}{n}, \frac{l}{n}\right) \right].$$

- Méthode de Simpson :

$$I_{ap} = \frac{1}{6n^3} \sum_{i,j,l=1}^n \left[\varphi\left(\frac{i-1}{n}, \frac{j-1}{n}, \frac{l-1}{n}\right) + 4\varphi\left(\frac{2i-1}{2n}, \frac{2j-1}{2n}, \frac{2l-1}{2n}\right) + \varphi\left(\frac{i}{n}, \frac{j}{n}, \frac{l}{n}\right) \right].$$

- Méthode de Monte-Carlo

$$\begin{aligned} I(k, 3) &= \int_{\mathbb{R}} \int_{\mathbb{R}} \int_{\mathbb{R}} \varphi(x, y, z) \mathbb{1}_{[0,1]^3}(x, y, z) dx dy dz \\ &= \mathbb{E} \left[\varphi(X, Y, Z) \right], \text{ où } (X, Y, Z) \rightsquigarrow \mathcal{U}_{[0,1]^3}. \end{aligned}$$

$$\text{Donc } I_{ap} = \frac{1}{n} \sum_{i=1}^n \varphi(x_i, y_i, z_i) \text{ où les } (x_i, y_i, z_i) \text{ sont les valeurs simulées de } (X, Y, Z).$$

On prend $n = 100$, on obtient le résultat suivant

Méthodes	I_{ap}	Erreur absolue
Valeur exacte de I	1	0
Rectangles	0.9997533	$2.467 \cdot 10^{-4}$
Trapèzes	0.9997533	$2.467 \cdot 10^{-4}$
Simpson	1	0
Monte-Carlo	1.009747	0.009747

TABLE 5.5 – Approximations et erreurs absolues de I par I_{ap} pour $k = 1$.

Méthodes	I_{ap}	Erreur absolue
Valeur exacte de I	1	0
Rectangles	$6.919417 \cdot 10^{-18}$	$6.919417 \cdot 10^{-18}$
Trapèzes	$5.969932 \cdot 10^{-18}$	$5.969932 \cdot 10^{-18}$
Simpson	$-4.233584 \cdot 10^{-18}$	$4.233584 \cdot 10^{-18}$
Monte-Carlo	-0.0124036	0.0124036

TABLE 5.6 – Approximations et erreurs absolues de I par I_{ap} pour $k = 2$.

Pour $k = 1$ et $k = 2$ nous remarquons que pour $n = 100$, les méthodes déterministes approximent mieux $I(k, 3)$ que celle de Monte-Carlo.

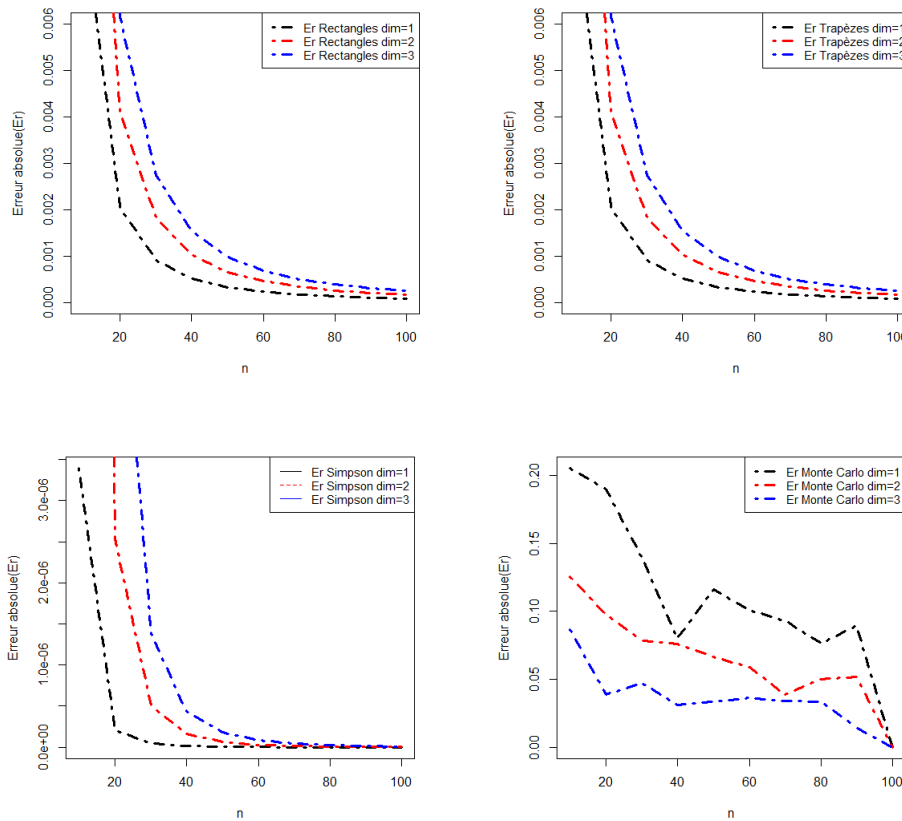


FIGURE 5.5 – Comparaison des erreurs en dimension 1, 2, 3 pour $k = 1$.

la figure 5.7 nous montre que pour les méthodes déterministes plus la dimension augmente plus l'erreur est élevée mais pour la méthode de Monte-Carlo plus la dimension augmente plus l'erreur est faible. De plus les méthodes déterministes utilisent plus de temps pour générer les résultats obtenus. C'est ce qui fait en dimension élevée les méthodes déterministes sont défavorisées par rapport à celle de Monte-Carlo.

5.3 Erreur d'approximation par la méthode QMC

Dans cette section nous générons des suites de Halton, de Sobol et de Faure pour estimer $I(k, d)$ pour $n = 1000$ en faisant varier la dimension d . Les résultats obtenus des erreurs d'approximation pour ces trois suites sont établis dans le tableau suivant, pour $k = 1$.

Les Suites	Dimension	Base	Estimation de I	Erreur absolue
Halton	1	2	1.000482	0.000482
	4	7	0.9996695	0.0003305
	10	31	1.054771	0.034771
	12	43	1.011341	0.011341
	15	53	0.9605606	0.03944
Sobol	1	2	1.001036	0.001036
	4	2	0.998146	0.001854
	10	2	1.067709	0.067709
	12	2	1.124253	0.124253
	15	2	1.710759	0.710759
Faure	1	2	1.001136	0.001136
	4	5	1.003463	0.003463
	10	11	1.040292	0.040292
	12	13	1.525488	0.525488
	15	17	3.49891	2.49891

TABLE 5.7 – Approximation de $I(1, d)$ par les suites de Halton, de Sobol et de Faure.

Le tableau 5.7 nous montre que l'erreur fournie par la suite de Halton est la plus faible, les résultats obtenus par la suite de Sobol et celle de Faure sont plutôt bons malgré le problème qu'on peut avoir dès que la dimension est élevée.

5.4 Comparaison entre les méthodes QMC et Monte-Carlo en dimension d

Dans cette section nous utilisons la suite de Halton pour l'approximation de $I(k, d)$ par la méthode QMC afin de la comparer avec celle de Monte-Carlo.

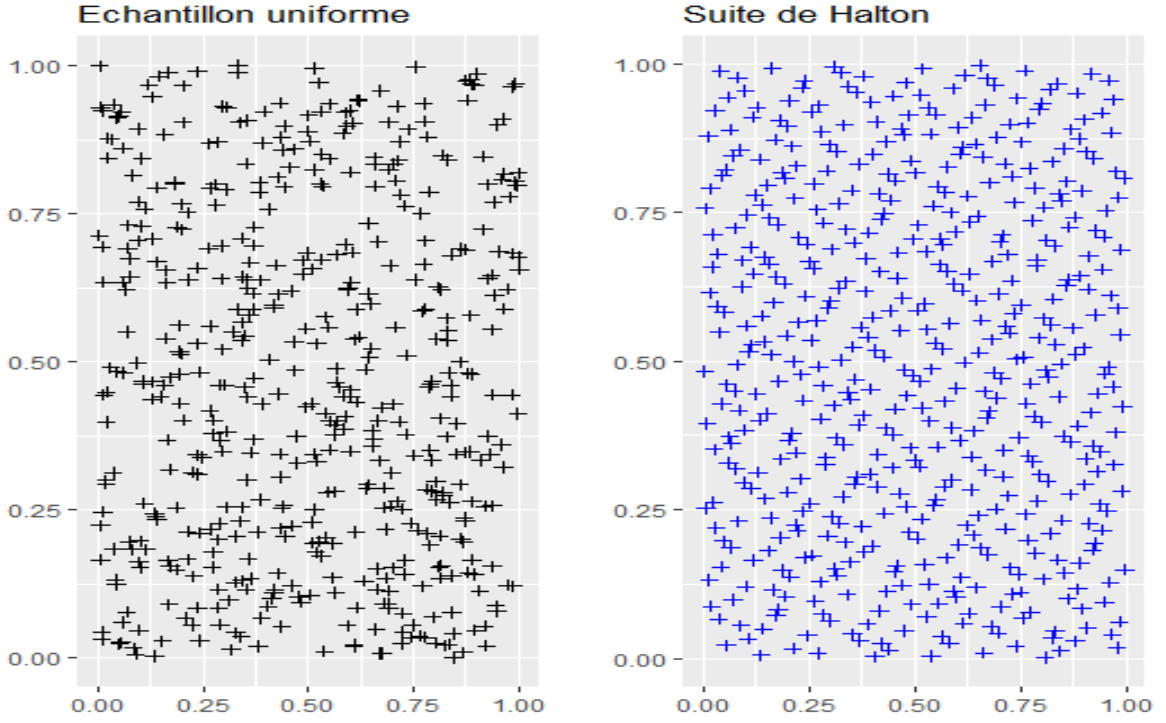


FIGURE 5.6 – Echantillons de 500 points d’une suite uniforme et d’une suite de Halton.

Dans la figure 5.6 nous constatons que la suite de Halton est plus uniformément distribuée que celle d’une loi uniforme.

Dans cette partie nous souhaitons réaliser une comparaison entre la méthode QMC et la méthode de Monte-Carlo en utilisant comme critère de comparaison le *Mean Square Error* (MSE) donné par la formule suivante :

$$MSE = \frac{1}{n} \sum_{i=1}^n (\text{erreur}_i)^2$$

Les estimateurs de $I(k, d)$ pour les deux méthodes sont donnés par :

- La méthode de Monte-Carlo, on a

$$\begin{aligned} I(k, d) &= \int_{[0,1]^d} \varphi(x_1, \dots, x_d) dx_1 \cdots dx_d \\ &= \int_{\mathbb{R}^d} \varphi(x_1, \dots, x_d) \mathbb{1}_{[0,1]^d}(x_1, \dots, x_d) dx_1 \cdots dx_d \\ &= \mathbb{E} \left[\varphi(X_1, \dots, X_d) \right], \text{ où } (X_1, \dots, X_d) \text{ sont générés selon la loi uniforme } \mathcal{U}_{[0,1]^d}. \end{aligned}$$

Dans ce contexte l’estimateur Monte-Carlo de base est défini par :

$$\hat{I}_n = \frac{1}{n} \sum_{i=1}^n \varphi(X_{1i}, \dots, X_{di}) \text{ où } (X_{1i}, \dots, X_{di}) \text{ i.i.d. } \rightsquigarrow \mathcal{U}_{[0,1]^d}.$$

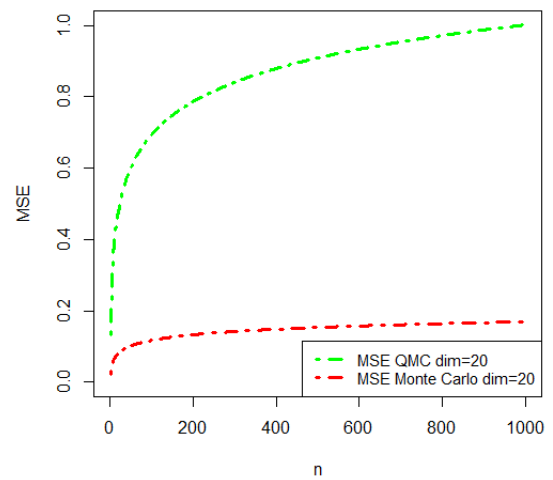
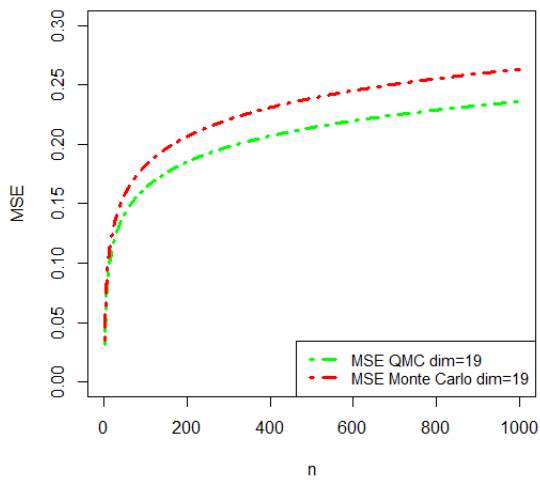
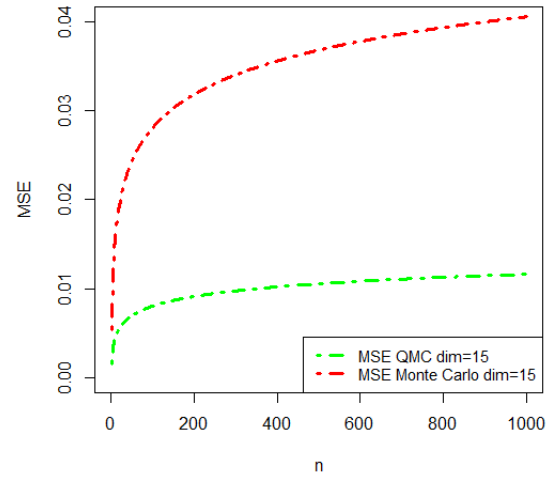
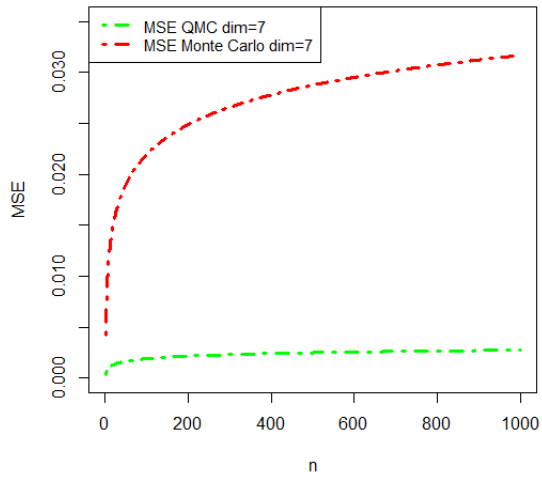
De plus \hat{I}_n est sans biais et d’après la LGN il est fortement convergent i.e

$$\hat{I}_n = \frac{1}{n} \sum_{i=1}^n \varphi(X_{1i}, \dots, X_{di}) \xrightarrow[n \rightarrow \infty]{p.s.} I(k, d)$$

- La méthode QMC : en se donnant une suite $(\xi_n^d)_{n \geq 1}$ de Halton basée sur d premiers nombres premiers, alors

$$\tilde{I}_n = \frac{1}{n} \sum_{i=1}^n \varphi(\xi_i^d)$$

est un estimateur QMC de l'intégrale $I(k, d)$.



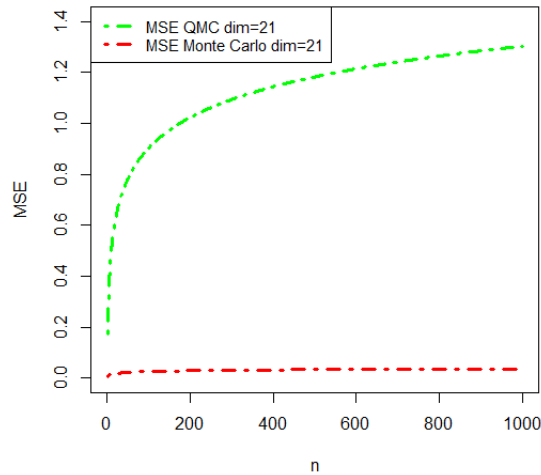


FIGURE 5.7 – Comparaison des méthodes MC et QMC par le MSE.

La Figure 5.7 nous montre qu'en dimension inférieure où égale à 19 la méthode de QMC approxime mieux l'intégrale $I(k, d)$ car son MSE est plus faible que celui de Monte-Carlo. Mais, plus la dimension augmente plus le MSE de la méthode de QMC est élevé alors que celui de Monte-Carlo diminue. En effet comme le montrent les deux derniers graphiques de la figure ci-avant, dès que la dimension est strictement supérieure à 19, le MSE de la méthode de Monte-Carlo devient plus faible. C'est ce qui justifie qu'en dimension supérieure la méthode de Monte-Carlo est plus performante que celle de QMC mais en dimension intermédiaire c'est celle de QMC qui l'emporte.

5.5 Conclusion

Dans ce chapitre, on peut retenir que les méthodes numériques déterministes approximent mieux l'intégrale $I(k, d)$ pour les petites dimensions car ayant la meilleure précision. Mais plus la dimension augmente plus l'erreur d'approximation déterministe augmente contrairement à celle Monte-Carlo. Malgré le caractère aléatoire de ses échantillons la mise en oeuvre pratique nous a permis de constater que la méthode de Monte-Carlo est meilleure en dimension élevée. Quant à sa comparaison avec la méthode QMC générée en particulier par une suite de Halton, elle fournit une plus grande erreur quadratique moyenne (MSE) en dimension intermédiaire et une plus faible MSE en dimension élevée.

En outre, les différentes comparaisons de ces méthodes peuvent être étendues par le choix des fonctions φ de $I(k, d)$ plus générales que celle proposée dans la mise en oeuvre pratique.

Conclusion générale et perspectives

Les méthodes usuelles de quadrature (méthode des Rectangles, méthode des Trapèzes, méthode de Simpson,...) sont efficaces pour le calcul d'intégrales dans les petites dimensions. De façon générale, une méthode déterministe adaptée permet d'atteindre une vitesse en $o(n^{-s/d})$ (où la fonction étudiée est de classe C^s sur $[0, 1]^d$). La méthode de Monte-Carlo est une méthode d'approximation par introduction de procédés aléatoires. Elle est l'une des techniques de simulation offrant la possibilité de calculer des intégrales en dimension élevée. Elle a comme grand avantage, par rapport aux autres méthodes d'approximation numériques, d'avoir une vitesse de convergence en $o(\frac{1}{\sqrt{n}})$ (pour un échantillon aléatoire de n points), donc indépendante de la dimension d du problème. Son inconvénient majeur est la lenteur de sa vitesse de convergence qui peut être améliorée par l'utilisation des techniques de réduction de la variance. Afin de remédier à ce problème de lenteur, nous faisons appel à la méthode numérique non aléatoire appelée méthode QMC, qui consiste à remplacer les procès aléatoires par des suites déterministes à discrédance faible. Nous avons introduit les outils théoriques nécessaires à la bonne compréhension des mécanismes de l'intégration QMC en insistant sur les différences fondamentales avec la méthode de Monte-Carlo. Nous avons notamment étudié les suites de Van Der Corput, de Halton, de Hammersley, de Sobol et de Faure, qui font partie de cette famille de suites à discrédance faible. L'avantage majeur de la méthode QMC est qu'elle converge plus rapidement que la méthode de Monte-Carlo avec une vitesse en $o(\frac{(\log(n))^d}{n})$. Cependant, elle comporte un inconvénient qui n'existe pas avec la méthode de Monte-Carlo : sa vitesse de convergence dépend de la dimension d du problème. Autrement dit, la méthode QMC perd de son efficacité lorsque la dimension augmente. Les résultats obtenus dans la partie pratique confortent bien la théorie.

Par ailleurs, en guise de perspectives, des pistes peuvent encore être explorées aussi bien d'un point de vue théorique que numérique. On peut étudier par exemple, le couplage entre méthodes déterministes et de Monte-Carlo ou encore l'introduction de paramètres aléatoires dans les ensembles de points déterministes pour améliorer les performances de la méthode QMC en grande dimension.

Bibliographie

- [1] Amal A . (2019). Calcul d'Intégrales par la Méthode de Monte-Carlo. *Mémoire de fin de cycle de Master*. Université D'Adrar, Algérie
- [2] Béjia R., Faure H. (1979). Discrédance de la suite de Van der Corput. Séminaire Delange-Pisot-Poitou. *Théorie des nombres*, 19, N.1, exp. N.13, 1-14.
- [3] Braaten E., Weller G. (1979). An improved low-discrepancy sequence for multidimensional quasi-Monte-Carlo integration. *Journal of Computational Physics* 33, 2, 249–258.
- [4] Caflish. R.E. (1998). Monte-Carlo and Quasi-Monte-Carlo Methods. *Acta Num.*,1-49.
- [5] Delacroix F. (2010). Intégrales multiples. *Support de cours*. École des Mines de Douai.
- [6] El Merouani M.(2019). Méthodes de réduction de la variance. *Support de cours*, Université Abdelmalek Essaâdi, Maroc.
- [7] Faure H. (1982). Discrédance de suites associées à un système de numération (en dimension s). *Acta Arithmetica* 41, 337-351.
- [8] Gauthier G. (2009). Les suites à discrédance faible, simulation Monte-Carlo. *HEC Montréal*.
- [9] Hammersley J. (1960). Monte-carlo methods for solving multivariate problems. *Annals of the New York Academy of Sciences* 86 , 844–874.
- [10] Lemieux C., L'Ecuyer P. (2001). On the use of Quasi-Monte-Carlo Methods in computational finance. Computational Science - ICCS 2001 (part I. *Lecture Notes in Computer Science*, vol.2073, Springer, 607-618.
- [11] Medouakh F. Z. (2018). Méthodes de Monte-Carlo. *PhD thesis*.
- [12] Nawel D., Linda M. (2013). Les méthodes Monte-Carlo, Quasi Monte-Carlo et leurs applications. *Mémoire de fin de cycle de Master*, Université de Béjaïa, Algérie.
- [13] Niederreiter H. (1992). *Random number generation and Quasi-Monte-Carlo methods*. Society for industrial and Applied Mathematics (SIAM), Philadelphia, PA.
- [14] Ourbih M. (2012). Méthodes de Monte-Carlo. *Support de cours*, Université de Béjaïa, Algérie.
- [15] Rubino G., Tuffin B. (2007). Simulations et méthodes de Monte-Carlo. In : *Techniques de l'Ingénieur*. No AF600.
- [16] Sobol I. (1967). On the distribution of points in a cube and the approximate evaluation of integrals. *Zhurnal Vychislitel'noi Matematiki i Matematicheskoi Fiziki* 7, 4 , 784–802.
- [17] Tuffin B. (2010). *La simulation de Monte-Carlo*. Hermes Science Publications.
- [18] Tuffin B. (1998). Variance reduction order using good lattice points in Monte-Carlo methods. *Computing* 61, 4 , 371-378.