

ÉCOLE NATIONALE DE LA STATISTIQUE
ET DE L'ADMINISTRATION ÉCONOMIQUE



Projet de Simulation et Monte Carlo

Réduction de variance par la méthode de Romberg

C. Lacombe, L. Chabannes et T. Laugel

2012-2013

Table des matières

1	Méthode de Monte Carlo	3
1.1	Schéma d'Euler	3
1.2	Méthode de Monte-Carlo : principaux résultats	4
1.3	Complexité	4
2	Méthode de réduction de variance de Romberg	5
2.1	Principe	5
2.2	Réduction de la variance	5
2.3	Théorème Central Limite	6
2.4	Complexité	6
3	Applications	8
3.1	Introduction	8
3.1.1	Génération de nombres	8
3.1.2	Erreur quadratique moyenne	9
3.2	Exemple théorique	9
3.2.1	Contexte	9
3.2.2	Estimation par la méthode de Monte Carlo	10
3.2.3	Estimation par la méthode de Romberg	12
3.2.4	Comparaison des deux méthodes	12
3.3	Application aux options asiatiques	13
3.3.1	Contexte	13
3.3.2	Le cas du call asiatique : Théorie	14
3.3.3	Application des deux méthodes	14
3.3.4	Comparaison des deux méthodes	15

Partie 1

Méthode de Monte Carlo

1.1 Schéma d'Euler

L'utilisation des méthodes de Monte Carlo et de Romberg Statistique pour l'approximation numérique de solutions d'équations différentielles stochastiques repose sur les schémas d'Euler (parfois appelés "méthode d'Euler-Maruyama"), que nous allons maintenant présenter. Cette méthode consiste en la construction d'un estimateur de la solution d'une équation différentielle par la discrétisation de celle-ci suivant un pas choisi. A partir d'une subdivision de $[0, T]$, on évalue la solution en chaque point de séparation de la subdivision. Si l'on considère une équation différentielle de solution le processus $(X_t)_{0 \leq t \leq T}$ de \mathbb{R}^d de la forme :

$$dX_t = b(X_t)dt + \sigma(X_t)dW_t \quad (1.1)$$

avec $X_0 \in \mathbb{R}^d$, W un mouvement brownien de dimension q et b et σ deux fonctions de classe C^1 respectivement à valeurs dans \mathbb{R}^d et $\mathbb{R}^{d \times q}$ satisfaisant l'inéquation suivante :

$$\forall (x, y) \in \mathbb{R}^d, |b(x) - b(y)| + |\sigma(x) - \sigma(y)| \leq C_T |x - y| \text{ avec } C_T = \text{Constante}(T) \quad (1.2)$$

On obtient alors la discrétisation d'Euler X^n de X au pas $\frac{T}{N}$ de ce processus, vérifiant l'équation :

$$X_{t_{n+1}}^N = X_{t_n}^N + b(X_{t_n}^N) \frac{T}{N} + \sigma(X_{t_n}^N)(W_{t_{n+1}}^N - W_{t_n}^N) \quad (1.3)$$

avec $t_n = \frac{nT}{N}$. On montre qu'avec la méthode d'Euler-Maruyama, ce processus vérifie :

$$\forall p > 1, \mathbb{E} \left(\sup_{t \in [0, T]} |X_t - X_t^n|^p \right) \leq \frac{K_p(T)}{n^{p/2}} \text{ avec } K_p(T) > 0 \quad (1.4)$$

Le processus discret ainsi construit va constituer une approximation de notre solution ; en effet, en définissant la convergence en loi stable d'un processus X^n vers X , notée $X^n \xrightarrow{\text{stable}} X$ comme la convergence :

$$\mathbb{E}(Uh(X_n)) \longrightarrow \mathbb{E}(Uh(X)) \quad (1.5)$$

pour toute fonction h bornée continue et toute variable aléatoire bornée U .

On a alors le résultat suivant :

$$\sqrt{n}(X^n - X) \xrightarrow{\text{stable}} U \quad (1.6)$$

avec U un processus de \mathbb{R}^d .

1.2 Méthode de Monte-Carlo : principaux résultats

Intéressons nous maintenant à la convergence de cet estimateur, et en particulier au comportement de $n^\alpha \epsilon_n$, avec $\alpha \in [0, 1]$. Le but final de la méthode étant l'approximation numérique de $\mathbb{E}(f(X_T))$, on définit deux types d'erreurs pour notre algorithme : on définit ainsi tout d'abord *l'erreur de discrétisation* comme le biais :

$$\epsilon_n = \mathbb{E}(f(X_T^n) - f(X_T)) \quad (1.7)$$

pour une fonction f donnée.

Afin d'évaluer l'erreur de notre résultat final, nous définissons également *l'erreur statistique*, dont l'expression est donnée ci-dessous :

$$\hat{f}^{n,N} - \mathbb{E}(f(X_T)) \quad (1.8)$$

avec $\hat{f}^{n,N} = \frac{1}{N} \sum_{i=1}^N f(X_{T,i}^n)$ une approximation de l'espérance de N copies indépendantes $f(X_{T,i}^n)_{1 \leq i \leq N}$ de $f(X_T^n)$.

On montre alors les résultats suivants :

Si f est une fonction de \mathbb{R}^d vérifiant :

$$\forall (x, y) \in \mathbb{R}^d, |f(x) - f(y)| \leq C(1 + |x|^p + |y|^p)|x - y|, p > 0 \quad (1.9)$$

où C est une constante positive, et si $\mathbb{P}(X_T \notin D_F) = 0$, avec D_F l'ensemble des points de \mathbb{R}^d où f est différentiable, alors on a :

$$\lim_{n \rightarrow \infty} \sqrt{n} \epsilon_n = 0 \quad (1.10)$$

On en déduit, lorsque plus généralement $\lim_{n \rightarrow \infty} n^\alpha \epsilon_n = C_f(t, \alpha)$ et lorsque $\mathbb{P}(X_T \notin D_F) = 0$, avec D_F l'ensemble des points de \mathbb{R}^d où f est différentiable :

$$n^\alpha \left(\frac{1}{n^{2\alpha}} \sum_{i=1}^{n^{2\alpha}} f(X_{T,i}^n) - \mathbb{E}(f(X_T)) \right) \xrightarrow{\text{stable}} \sigma_1 G + C_f(T, \alpha) \quad (1.11)$$

avec $\sigma_1^2 = \mathbb{V}(f(X_T))$ et $G \sim \mathcal{N}(0, 1)$.

1.3 Complexité

Nous nous intéressons maintenant à la complexité de l'algorithme de Monte Carlo classique. Le but de la méthode Romberg Statistique étant de concurrencer sur ce plan l'algorithme habituel, il convient de présenter celle-ci ainsi que les méthodes que nous avons utilisées pour la mesurer.

Le résultat précédent (1.11) prouve le comportement de l'erreur statistique lorsque l'on utilise la méthode de Monte Carlo classique : pour une erreur d'ordre $\frac{1}{n^\alpha}$, l'effort nécessaire minimal de computation de l'algorithme de Monte Carlo est atteint en $N = n^{2\alpha}$. On déduit de ce résultat la formule de la complexité de l'algorithme ; l'algorithme de Monte Carlo optimal a une complexité donnée par :

$$C_{MC} = C \times (nN) = C \times n^{2\alpha+1} \quad (1.12)$$

où C est une constante positive.

Nous utiliserons ce résultat pour comparer la complexité de l'algorithme de Monte Carlo classique avec celle de la méthode de Romberg Statistique, que nous allons maintenant présenter.

Partie 2

Méthode de réduction de variance de Romberg

2.1 Principe

Quand on estime $\mathbb{E}(X)$ où X est une variable aléatoire, la méthode de Monte Carlo implique une erreur globale d'ordre $\frac{\sigma}{\sqrt{n}}$, σ étant l'écart type de X et n le nombre de simulations. Pour diminuer cette erreur, il est possible d'appliquer des méthodes de réduction de variance par variable de contrôle. Celles-ci consistent en l'utilisation pour l'estimation de $\mathbb{E}(X)$ d'une variable $Q = X - Y + \mathbb{E}(Y)$, tel que $\mathbb{E}(Q) = \mathbb{E}(X)$. La variance de Q est alors plus faible mais on augmente par ailleurs le nombre de simulations puisqu'il faut également estimer $\mathbb{E}(X - Y)$.

Dans le cadre de la méthode de Romberg Statistique, on utilise la variable de contrôle $f(X_T^m)$ et on estime donc l'espérance de la variable $Q = f(X_T^n) - f(X_T^m) + \mathbb{E}(f(X_T^m))$ où f vérifie (1.9). En pratique, il s'agit tout d'abord d'estimer la quantité $Q_1 = \mathbb{E}(f(X_T^m))$ qui nous est inconnue puis d'estimer la quantité $Q_2 = \mathbb{E}[f(X_T^n) - f(X_T^m)]$ et cela de façon indépendante. Il est donc nécessaire de discrétiser deux fois le processus de diffusion X_T en utilisant des pas différents ($\frac{T}{n}$ et $\frac{T}{m}$ où $m < n$) et cela un nombre N_n de fois ce qui permet d'estimer Q_2 . Pour évaluer Q_1 , on discrétise X_T également en utilisant un pas $\frac{T}{m}$ mais on utilise des chemins browniens différents et on estime l'espérance en répétant cette discrétisation N_m fois.

La quantité qu'il faut calculer est donc :

$$V_n = \frac{1}{N_m} \sum_{i=1}^{N_m} f(X_{T,i}^m) + \frac{1}{N_n} \sum_{i=1}^{N_n} [f(X_{T,i}^n) - f(X_{T,i}^m)] \quad (2.1)$$

2.2 Réduction de la variance

En utilisant cette méthode, on obtient une variance de Q plus faible que la variance obtenue dans le cadre d'une méthode de Monte Carlo classique. En effet, sous les hypothèses précédentes :

$$\mathbb{V}(Q) = O\left(\frac{1}{m}\right). \quad (2.2)$$

Preuve : Pour f α -lipschitzienne,

$$\begin{aligned}\sigma_Q &= ||Q - \mathbb{E}(Q)|| \\ &\leq ||f(X_t^n) - f(X_t^m)|| \\ &\leq \alpha \left(\sup_{t \in [0, T]} ||X_t - X_t^n|| + \sup_{t \in [0, T]} ||X_t - X_t^m|| \right)\end{aligned}$$

On utilise alors la proposition (1.4) ce qui implique $\sigma_Q \leq K \times \frac{1}{\sqrt{m}}$ où K dépend de la constante associée à (1.4) et de α . \square

2.3 Théorème Central Limite

Comme dans le cas de la méthode de Monte Carlo classique, un théorème central limite permet de déterminer l'ordre de l'erreur globale pour des valeurs de m , N_m et N_n que l'on suppose dépendantes de n .

On pose $m = n^\beta$, $N_m = n^{\gamma_1}$ et $N_n = n^{\gamma_2}$ où $\beta \in [0, 1]$, $(\gamma_1, \gamma_2) \in]1, \infty[^2$.

Pour f une fonction de \mathbb{R}^d vérifiant :

$$\forall (x, y) \in \mathbb{R}^d, |f(x) - f(y)| \leq C(1 + |x|^p + |y|^p)|x - y|, p > 0 \quad (2.3)$$

où C est une constante positive.

On suppose que $\mathbb{P}(X_T \notin D_F) = 0$, avec D_F l'ensemble des points de \mathbb{R}^d où f est différentiable et que pour $\alpha \in [\frac{1}{2}, 1]$,

$$\lim_{n \rightarrow \infty} n^\alpha \epsilon_n = C_f(T, \alpha) \quad (2.4)$$

Alors, pour $\gamma_1 = 2\alpha$ et $\gamma_2 = 2\alpha - \beta$:

$$n^\alpha (V_n - \mathbb{E}(f(X_T))) \xrightarrow{\text{stable}} \sigma_2 G + C_f(T, \alpha) \quad (2.5)$$

avec $\sigma_2^2 = \mathbb{V}(f(X_T)) + \mathbb{V}(\nabla f(X_T) \cdot U_T)$

2.4 Complexité

Dans son article, Kebaier s'attache à déterminer les nombres m , N_m et N_n qui donnent une complexité minimale à l'algorithme. Initialement, la complexité de la méthode de Romberg vaut $C_{RS} = K \times (mN_m + (n + m)N_n)$ puisqu'on simule :

- N_m fois les m pas de discrétisation pour calculer $\mathbb{E}(f(X_t^m))$
- N_n fois les m pas de discrétisation associé à X_T^m
- N_n fois les n pas de discrétisation associé à X_T^n , les deux dernières étapes ayant pour but de calculer $\mathbb{E}(f(X_T^n) - f(X_T^m))$.

En utilisant les paramètres donnés dans le théorème central limite, on obtient une complexité $C_{RS} = K \times [n^{\beta+2\alpha} + (n + n^\beta)n^{2\alpha-\beta}]$.

Cette complexité est minimisée pour $\beta = \frac{1}{2}$.

Preuve :

On pose $f_\alpha(\beta) = [n^{\beta+2\alpha} + (n + n^\beta)n^{2\alpha-\beta}]$.

$\forall \beta \in [0, 1], f'_\alpha(\beta) = \ln(n).n^{2\alpha}.(n^\beta - n^{1-\beta})$.

Cela implique que f_α est minimisée pour $\beta = \frac{1}{2}$. \square

Pour cette valeur de β , on obtient alors une complexité de :

$$C_{RS} = K \times n^{\frac{2\alpha+1}{2}}$$

ce qui est plus faible que la complexité de la méthode de Monte Carlo classique.

Partie 3

Applications

3.1 Introduction

3.1.1 Génération de nombres

Pour simuler nos mouvements Browniens ainsi que pour initialiser les paramètres de nos différentes applications, nous avons utilisé les fonctions *runif* et *rnorm* implémentées dans R.

Ces fonctions s'appuient sur le générateur de nombres pseudo-aléatoires Mersenne-Twister que nous allons présenter ici. Considérons un nombre entier X dont la décomposition binaire est $\sum_{i=0}^{\omega-1} x_i 2^i$, X est représenté dans l'ordinateur par un vecteur de bits $(x_{\omega-1}, \dots, x_0)$.

Soit $Y = \sum_{i=0}^{\omega-1} y_i 2^i$, on note \oplus et \otimes les opérateurs définis par :

$$- X \oplus Y = \sum_{i=0}^{\omega-1} (x_i + y_i[2]).2^i$$

$$- X \otimes Y = \sum_{i=0}^{\omega-1} (x_i \times y_i[2]).2^i$$

On définit également la fonction A :

$$\begin{aligned} A(x) &= (x \gg 1) \oplus 0 \text{ si } x[2] = 0 \\ &= (x \gg 1) \oplus a \text{ si } x[2] = 1 \end{aligned}$$

où $(x \gg k)$ est le vecteur x décalé de k bits vers la droite, $(x \ll k)$ est le vecteur x décalé de k bits vers la gauche.

Le processus de Mersenne Twister se fonde sur deux opérations :

Une opération de récurrence :

$$X_{k+n} = X_{k+m} \oplus A((X_{k+1} \otimes \underline{M_r}) \oplus (X_k \otimes \overline{M_r})) \quad (3.1)$$

où $\underline{M_r}$ est le vecteur composé des r premiers bits égaux à 1, le reste étant nul, $\overline{M_r}$ le vecteur composé des r derniers bits égaux à 1, le reste étant nul et $m \leq n$.

Une opération de tempering :

Elle consiste à mélanger les bits du vecteur obtenu par l'opération de récurrence par le processus suivant :

$$\begin{aligned}
X_{k+n} &\leftarrow X_{k+n} \bigoplus (X_{k+n} \gg u) \\
X_{k+n} &\leftarrow X_{k+n} \bigoplus ((X_{k+n} \ll s) \otimes b) \\
X_{k+n} &\leftarrow X_{k+n} \bigoplus ((X_{k+n} \ll t) \otimes c) \\
X_{k+n} &\leftarrow X_{k+n} \bigoplus (X_{k+n} \gg l)
\end{aligned}$$

On utilise ensuite la fonction de sortie $g : X_k \rightarrow \frac{X_k + 0.5}{2^\omega}$ pour obtenir un réel dans $[0,1]$.

Dans le générateur implémenté dans R, le Mersenne Twister 19937, les paramètres suivants sont utilisés :

- Pour la récurrence : $\omega=32$, $n=624$, $r=31$, $m=397$ et $a=2567483615$
- Pour le tempering : $u=11$, $s=7$, $t=15$, $l=18$ et $b=2636928640$

Ces paramètres permettent en fait de maximiser la période du générateur qui vaut alors $2^{19937} - 1$.

3.1.2 Erreur quadratique moyenne

Afin de comparer les deux méthodes, on utilise la méthodologie proposée par Broadie et Detemple¹. Cette dernière repose sur le fait que pour un ensemble de paramètres de la diffusion, un des deux algorithmes donnera de meilleurs résultats.

L'erreur de mesure est donnée par l'erreur quadratique moyenne :

$$RMS = \sqrt{\frac{1}{M} \sum_{i=1}^M (\mathbb{E}g_\alpha(Z_T) - \text{Valeur simulée})^2} \quad (3.2)$$

La valeur simulée sera différente pour les deux méthodes de Monte Carlo classique et Romberg Statistique.

3.2 Exemple théorique

Dans cette partie, nous allons appliquer les méthodes de Monte Carlo classique et Romberg Statistique à un exemple théorique simple afin de les comparer.

3.2.1 Contexte

On considère la diffusion bi-dimensionnelle $Z=(X,Y)$ suivant le système d'équations différentielles suivant :

Soit $(W_t)_t$ un mouvement Brownien standard,

$$\begin{cases} dX_t = -\frac{X_t}{2dt} - Y_t dW_t \\ dY_t = -\frac{Y_t}{2dt} + X_t dW_t \end{cases} \quad (3.3)$$

1. M. Broadie, J. Detemple, *Recent advances in numerical methods for pricing derivative securities*, 1997

La solution du système d'équations différentielles, avec la condition initiale $Z_0 = (\cos(\theta), \sin(\theta))$, est donc :

$$Z_t = (\cos(\theta + W_t), \sin(\theta + W_t)) \quad (3.4)$$

On suppose également que $\theta \in [0, 2\pi]$ de manière à avoir Z qui décrit le cercle unité.

Preuve :

On souhaite vérifier que $Z_t = (\cos(\theta + W_t), \sin(\theta + W_t))$ est solution du système d'équations différentielles énoncé plus haut.

On applique la formule d'Ito² et on obtient :

$$\begin{aligned} dX_t &= -\sin(\theta + W_t)dW_t - \frac{1}{2}\cos(\theta + W_t)dt \\ &= -Y_t dW_t - \frac{1}{2}Y_t dt \end{aligned}$$

et,

$$\begin{aligned} dY_t &= \cos(\theta + W_t)dW_t - \frac{1}{2}\sin(\theta + W_t)dt \\ &= X_t dW_t - \frac{1}{2}X_t dt \end{aligned} \quad \square$$

L'objectif des méthodes de Monte Carlo classique et Romberg Statistique est d'estimer $\mathbb{E}(f_\alpha(Z_t))$ avec $f_\alpha : z = (x, y) \rightarrow ||z|^2 - 1|^{2\alpha}$.

3.2.2 Estimation par la méthode de Monte Carlo

Dans cette section, on cherche à approximer $\mathbb{E}(f_\alpha(Z_t))$ par $\hat{f}_\alpha^{n,N} = \frac{1}{N} \sum_{i=1}^N f_\alpha(Z_{T,i}^n)$.

Soit Z^n le schéma d'Euler associé à Z , $\alpha \in [\frac{1}{2}, 1]$, $G \sim \mathcal{N}(0, 1)$, alors on a :

$$n^\alpha \mathbb{E}(f_\alpha(Z_T^n) - f_\alpha(Z_T)) \rightarrow (2T)^\alpha \mathbb{E}|G|^{2\alpha}, \quad t \geq 0 \quad (3.5)$$

Dans la suite, on pose : $(2T)^\alpha \mathbb{E}|G|^{2\alpha} = C_{f_\alpha}(T, \alpha)$.

Preuve :

$$\begin{aligned} n_\alpha \mathbb{E}(f_\alpha(Z_t^n) - f_\alpha(Z_t)) &= n^\alpha \mathbb{E}(|Z_t^n|^2 - 1|^{2\alpha} - |Z_t|^2 - 1|^{2\alpha}) \\ &= n^\alpha \mathbb{E}||Z_t^n|^2 - 1|^{2\alpha} \text{ car } Z \text{ est sur le cercle unité} \\ &= n^\alpha \mathbb{E}|(Z_t + Z_t^n)(Z_t^n - Z_t)|^{2\alpha} \\ &= \mathbb{E}|(Z_t + Z_t^n)\sqrt{n}(Z_t^n - Z_t)|^{2\alpha} \end{aligned}$$

Or on a : $Z_t + Z_t^n \rightarrow 2Z_t$ et $\sqrt{n}(Z_t^n - Z_t) \xrightarrow{\text{stable}} \tilde{U}_t$ selon le Théorème (6).

On obtient donc que :

$$n_\alpha \mathbb{E}(f_\alpha(Z_t^n) - f_\alpha(Z_t)) \rightarrow 2^{2\alpha} \mathbb{E}|Z_t \tilde{U}_t|^{2\alpha}$$

2. Soit $dX_t = \mu_t dt + \sigma_t dB_t$ avec B_t un mouvement Brownien. Alors la formule d'Ito donne :

$$df(X_t, t) = \frac{\partial f}{\partial t}(X_t, t)dt + \frac{\partial f}{\partial x}(X_t, t)dX_t + \frac{1}{2} \frac{\partial^2 f}{\partial x^2}(X_t, t)\sigma_t^2 dt$$

Avec $\tilde{U} = (\tilde{U}^1, \tilde{U}^2)$ qui vérifie :

$$\begin{cases} d\tilde{U}_t^1 = -\frac{1}{2}\tilde{U}_t^1 dt - \tilde{U}_t^2 dW_t + \frac{1}{\sqrt{2}}X_t d\tilde{B}_t \\ d\tilde{U}_t^2 = -\frac{1}{2}\tilde{U}_t^2 dt + \tilde{U}_t^1 dW_t + \frac{1}{\sqrt{2}}Y_t d\tilde{B}_t \end{cases} \quad (3.6)$$

et \tilde{B} un mouvement Brownien standard indépendant de W .

La solution de ce système d'équations différentielles est :

$$\begin{cases} \tilde{U}_t^1 = \frac{1}{\sqrt{2}}X_t \tilde{B}_t \\ \tilde{U}_t^2 = \frac{1}{\sqrt{2}}Y_t \tilde{B}_t \end{cases} \quad (3.7)$$

On a également le résultat suivant :

$$n^\alpha \left(\frac{1}{n^{2\alpha}} \sum_{i=1}^{n^{2\alpha}} f_\alpha(Z_{T,i}^n) - \mathbb{E}(f_\alpha(Z_t)) \right) \Rightarrow \sigma \bar{G} + C_{f_\alpha}(T, \alpha) \quad (3.8)$$

avec $\sigma^2 = \mathbb{V}(f_\alpha(Z_t))$ et $\bar{G} \sim \mathcal{N}(0, 1)$.

Cependant, afin d'obtenir les paramètres optimaux donnés dans les théorèmes précédents, on considère $g_\alpha(x, y) = f_\alpha(x, y) + x$ au lieu de $f_\alpha(x)$. En effet, cela nous permet d'avoir une variance de la fonction non nulle et une erreur de discrétisation de l'ordre $\frac{1}{n^\alpha}$.

Avec un $\alpha = \frac{1}{2}$, le résultat obtenu précédemment reste valable :

$$\sqrt{n}\mathbb{E} \left(f_{\frac{1}{2}}(Z_T^n) - f_{\frac{1}{2}}(Z_T) \right) \rightarrow \sqrt{2T}\mathbb{E}|G| \quad , \quad t \geq 0 \quad (3.9)$$

car $\sqrt{n}(X_T^n - X_T) \rightarrow 0$

On adapte maintenant les résultats précédents à notre cas particulier :

$$\begin{aligned} C_{g_{\frac{1}{2}}}(T, \frac{1}{2}) &= C_{f_{\frac{1}{2}}}(T, \frac{1}{2}) \\ &= \sqrt{2T}\mathbb{E}|G| \quad \text{avec } G \in \mathcal{N}(0, 1) \\ &= \sqrt{2T} \int_0^\infty \frac{2}{\sqrt{2\pi}} y e^{-\frac{y^2}{2}} dy \\ &= \sqrt{2T} \sqrt{\frac{2}{\pi}} \int_0^\infty y e^{-\frac{y^2}{2}} dy \\ &= \sqrt{2T} \sqrt{\frac{2}{\pi}} \\ &= 2\sqrt{\frac{T}{\pi}} \end{aligned}$$

Puis, $\sigma^2 = \frac{1}{2} (1 + \cos(2\theta)e^{-2T}) - \cos^2(\theta)e^{-T}$ (c.f. Annexe A)

Ce qui nous donne donc :

$$\sqrt{n} \left(\frac{1}{n} \sum_{i=1}^n g_{\frac{1}{2}}(Z_{T,i}^n) - \mathbb{E}(g_{\frac{1}{2}}(Z_T)) \right) \Rightarrow \sigma \bar{G} + C_{g_{\frac{1}{2}}}(T, \frac{1}{2})$$

3.2.3 Estimation par la méthode de Romberg

Dans cette section, on cherche à approximer $\mathbb{E}(g_\alpha(Z_t))$ par :

$$\frac{1}{\sqrt{n}} \sum_{i=1}^{\sqrt{n}} \left[g_\alpha(Z_{T,i}^n) - g_\alpha(Z_{T,i}^{\sqrt{n}}) \right] + \frac{1}{n} \sum_{i=1}^n g_\alpha(\hat{Z}_{T,i}^{\sqrt{n}})$$

On applique le théorème Central limite de Romberg Statistique et on obtient :

$$\sqrt{n}(V_n - \mathbb{E}g_{\frac{1}{2}}(Z_T)) \Rightarrow \sigma_2 \tilde{G} + 2\sqrt{\frac{2}{\pi}}$$

avec V_n défini par l'équation (2.1) et :

$$\begin{aligned} \sigma_2^2 &= \mathbb{V}f(Z_T) + \tilde{\mathbb{V}}(\nabla f(Z_T).U_T) \\ &= \sigma^2 + \tilde{\mathbb{V}}(\nabla f(Z_T).U_T) \end{aligned}$$

3.2.4 Comparaison des deux méthodes

Calculons la valeur exacte de $\mathbb{E}(g_\alpha(Z_T))$:

$$\begin{aligned} \mathbb{E}(g_\alpha(Z_T)) &= \mathbb{E}(X_T) \\ &= \mathbb{E}(\cos(\theta + W_T)) \\ &= \mathbb{E}[\operatorname{Re}(e^{i\theta + iW_T})] \\ &= \operatorname{Re}[\mathbb{E}(e^{i\theta + iW_T})] \\ &= \operatorname{Re}[e^{i\theta} \mathbb{E}(e^{iW_T})] \\ &= \operatorname{Re}[e^{i\theta} e^{-\frac{T}{2}}] \\ &= \cos(\theta) e^{-\frac{T}{2}} \end{aligned}$$

Afin de comparer les deux méthodes, on teste les deux algorithmes pour un grand échantillon de paramètres choisis au hasard. On génère donc, au hasard, $M = 100$ valeurs pour $Z_0 = (X_0, Y_0)$. Ensuite, pour chacun des deux algorithmes, on calcule la vitesse, comme le nombre de valeurs simulées calculées par seconde, et l'erreur de mesure, comme le RMS. On obtient alors le couple de points $(RMS^{(MC, M=100, n)}, Vitesse^{(MC, M=100, n)})$. Puis on itère ce processus pour $n \in \{60, 90, 120, 180, 220, 250\}$. On obtient ainsi la figure 3.1.

On remarque ainsi que la courbe obtenue par la méthode Romberg Statistique est au-dessus de celle obtenue par la méthode de Monte Carlo classique. Ainsi, pour une erreur ϵ donnée, le nombre de valeurs calculées par seconde avec la méthode Romberg Statistique est supérieur au nombre de valeurs calculées par seconde avec la méthode de Monte Carlo classique.

On peut donc conclure que la méthode Romberg Statistique est plus rapide que la méthode de Monte Carlo classique.

Nous avons également appliqué l'algorithme à l'exemple théorique en utilisant une suite à discrétance faible de Sobol pour simuler nos mouvements browniens. Cette suite permet d'avoir des nombres répartis plus uniformément que dans le cas d'un générateur aléatoire comme le Mersenne-Twister. La figure 3.2 présente les résultats pour les mêmes valeurs de M et n que précédemment.

L'utilisation d'une suite à discrétance faible vient nuancer les résultats précédemment établis. En effet, on constate qu'à vitesse donnée, l'erreur obtenue par la méthode Romberg Statistique est beaucoup plus importante que celle obtenue par la méthode de Monte Carlo classique.

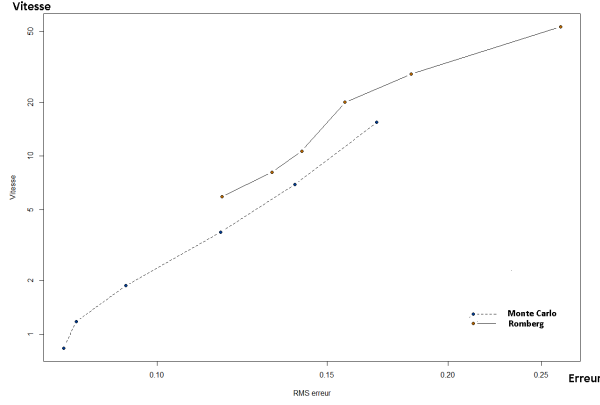


FIGURE 3.1 – Évolution de la vitesse et du RMS en fonction de n , pour $M=100$

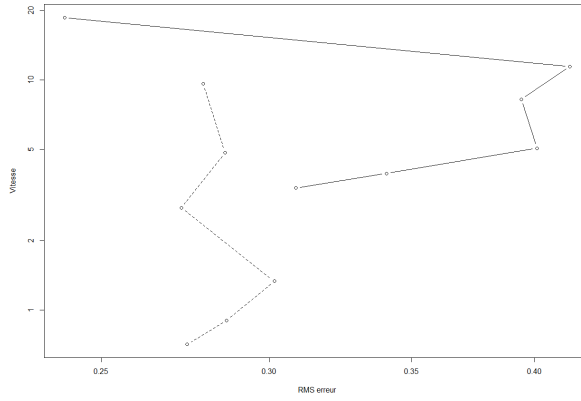


FIGURE 3.2 – Évolution de la vitesse et du RMS en fonction de n , pour $M=100$ avec suite de Sobol

3.3 Application aux options asiatiques

Dans cette partie, nous allons appliquer les méthodes de Monte Carlo classique et Romberg Statistique à la valorisation des options asiatiques puis comparer les résultats.

3.3.1 Contexte

On cherche à déterminer le prix d'une option asiatique. Ce type d'option est caractérisé par la dépendance de sa valeur par rapport à l'intégrale des prix du sous-jacent. Le pay-off peut s'exprimer de la manière suivante :

$$\Pi(S) = e^{-rT} \mathbb{E}(f(S_T, I_T)) \quad (3.10)$$

où f est la fonction qui détermine le type de l'option et $I_T = \frac{1}{T} \int_0^T S_u du$

Par exemple :

- si $f(x, y) = (y - K)_+$, l'option est un call asiatique
- si $f(x, y) = (K - y)_+$, l'option est un put asiatique

La diffusion S_t suit le modèle de Black-Scholes :

$$\frac{dS_t}{S_t} = rdt + \sigma dW_t \text{ avec } t \in [0, T], \quad T > 0 \quad (3.11)$$

En résolvant cette équation (utilisation de la formule d'Itô), on trouve que :

$$S_t = S_0 \exp\left((r - \frac{\sigma^2}{2})t + \sigma W_t\right) \quad (3.12)$$

En simulant des mouvements Browniens, on est donc capable d'évaluer la valeur du sous-jacent à l'instant t .

3.3.2 Le cas du call asiatique : Théorie

Dans la suite, nous avons voulu comparer les performances des deux méthodes dans le cas d'un call asiatique (i.e. $f(x, y) = (y - K)_+$).

Nous avons été confronté à un problème d'ordre théorique pour calculer la valeur réelle d'un tel call pour des paramètres σ , r , T et S_0 donnés. En effet, dans la littérature financière, nous n'avons pas réussi à trouver une formule analytique donnant la valeur exacte du call asiatique. De nombreuses approximations existent mais il nous a semblé que les étudier en détail pour déterminer laquelle utiliser sortait du cadre du projet. La valeur exacte du call étant nécessaire pour le calcul du RMS, nous avons donc décidé de fixer le strike à $K=0$.

En effet, dans ce cas, nous sommes capables de déterminer la valeur exacte de l'option :

$$\Pi(S) = \frac{S_0}{rT} (1 - e^{-rT}) \quad (3.13)$$

Preuve :

$$\begin{aligned} \Pi(S) &= e^{-rT} \mathbb{E}\left(\frac{1}{T} \int_0^T S_u du\right) \\ &= e^{-rT} \mathbb{E}\left(\frac{1}{T} \int_0^T e^{ru} \tilde{S}_u du\right) \text{ avec } \tilde{S}_u \text{ le prix du sous-jacent actualisé} \\ &= e^{-rT} \frac{1}{T} \int_0^T e^{ru} \mathbb{E}(\tilde{S}_u) du \text{ par le théorème de Fubini} \end{aligned}$$

En utilisant $\mathbb{E}(\tilde{S}_u) = S_0$ et $\int_0^T e^{ru} du = \frac{e^{rT}-1}{r}$, on en déduit le résultat final. \square

3.3.3 Application des deux méthodes

Si l'on connaît l'expression de S_t de par la résolution de Black-Scholes, la valeur de l'intégrale I_T nous est inconnue. Il s'agit donc d'approximer cette quantité avant d'utiliser les deux méthodes présentées précédemment. L'étape de discrétisation de (X_T) dans la méthode de Romberg théorique, est ici remplacée par l'évaluation de l'intégrale. Kebaier choisit d'utiliser la méthode des trapèzes pour l'estimer :

$$I_T^n = \frac{\delta}{T} \sum_{k=1}^n S_{t_{k-1}} \left(1 + \frac{r\delta}{2} + \sigma \frac{W_{t_k} - W_{t_{k-1}}}{2}\right) \quad (3.14)$$

où $\delta = \frac{T}{n}$ et $t_k = \delta k$.

Ce schéma possède l'avantage de converger en $O(\frac{1}{n})$ ce qui nous permet d'appliquer les deux méthodes.

Romberg Statistique :

On estime la quantité suivante :

$$E_n = \frac{e^{-rT}}{n^2} \sum_{i=1}^{n^2} I_T^{n^\beta} + \frac{e^{-rT}}{n^{2-2\beta}} \sum_{i=1}^{n^{2-2\beta}} (I_T^n - I_T^{n^\beta}) \quad (3.15)$$

pour un β que nous allons préciser par la suite. En effet, le théorème central limite de la méthode Romberg Statistique adapté au cas des options asiatiques énonce que :

Pour f une fonction de \mathbb{R}^2 vérifiant :

$$\forall (x, y_1, y_2) \in \mathbb{R}^3, |f(x, y_1) - f(x, y_2)| \leq C(1 + |x|^p + |y_1|^p + |y_2|^p)|y_1 - y_2| \quad (3.16)$$

où $p > 0$ et $C = \text{cte}$, Si $\mathbb{P}(S_T, I_T \notin D_F) = 0$, avec D_F l'ensemble des points de \mathbb{R}^{+2} où la dérivée de f par rapport à y existe, alors :

$$n(E_n - \mathbb{E}(I_T)) \xrightarrow{\text{stable}} \hat{\sigma}_2 G + \mathbb{E}(g(I_T)\chi_T) \quad (3.17)$$

où $\hat{\sigma}_2^2 = \mathbb{V}(I_T) + \mathbb{V}(g(I_T)\chi_T)$, G une variable suivant une loi normale $\mathcal{N}(0, 1)$, g la fonction définie par $g : x \rightarrow \frac{x}{|x|}$ et χ un processus limite tel que :

$$\chi_t = \frac{\sigma}{2\sqrt{3}} \int_0^t S_s dB_s \quad (3.18)$$

où B est un mouvement Brownien.

Ce Théorème Central Limite implique que la complexité associée à la méthode de Romberg Statistique appliquée à la valorisation d'un call asiatique est de :

$$C_{RS} = K \times (n^{\beta+2} + (n + n^\beta)n^{2-2\beta}) \quad (3.19)$$

On minimise cette expression ce qui donne finalement, pour $\beta \approx \frac{1}{3}$, $C_{RS} \approx K \times n^{\frac{7}{3}}$.

Monte Carlo :

On estime la quantité suivante :

$$\frac{e^{-rT}}{N} \sum_{i=1}^N (I_T^n) \quad (3.20)$$

Pour comparer cette méthode à celle de Romberg, il faut prendre $N = n^2$.

La complexité associée à cette méthode est alors de $C_{MC} = K \times n^3$.

3.3.4 Comparaison des deux méthodes

Afin de comparer les deux méthodes, on teste les deux algorithmes pour un grand échantillon de paramètres choisis au hasard. On génère donc, au hasard, $M = 200$ valeurs pour :

– $r \sim \mathcal{U}_{[0,0.1]}$;

- $\sigma \sim \mathcal{U}_{[0.05,0.7]}$;
- $S_0 \sim \mathcal{U}_{[20,40]}$;
- $T \sim \mathcal{U}_{[0.05,5]}$.

Ensuite, pour chacun des deux algorithmes, on calcule la vitesse et l'erreur de mesure. On obtient alors le couple de points $(RMS^{(MC,M=200,n)}, Vitesse^{(MC,M=200,n)})$. Puis on itère ce processus 5 fois, pour $n \in \{10, 20, 60, 120, 180\}$. On obtient ainsi la figure 3.3.

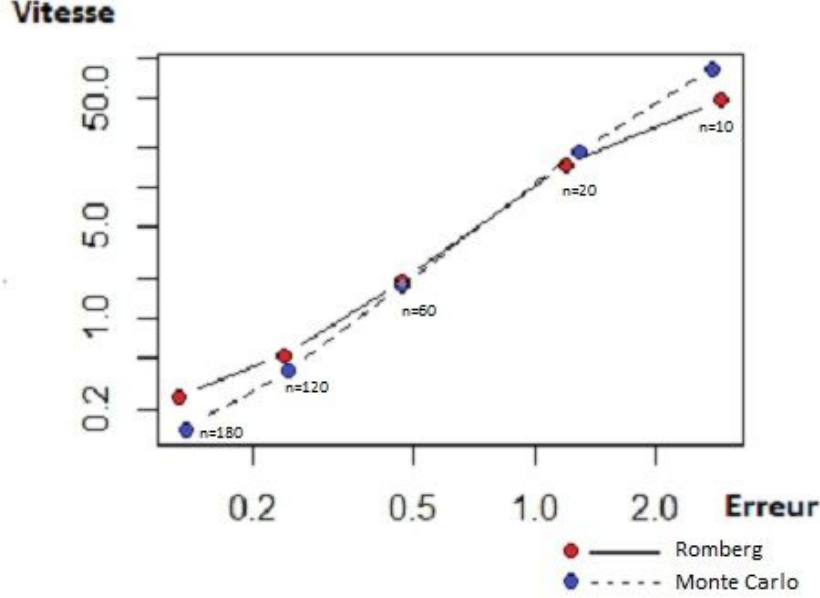


FIGURE 3.3 – Évolution de la vitesse et du RMS en fonction de n , pour $M=200$

On remarque alors que pour une erreur ϵ grande (n petit) le nombre de valeurs calculées par seconde avec la méthode de Monte Carlo classique est supérieur au nombre de valeurs calculées par seconde avec la méthode Romberg Statistique. A l'inverse, pour une erreur petite (n grand), le nombre de valeurs calculées par seconde avec la méthode Romberg Statistique est supérieur au nombre de valeurs calculées par seconde avec la méthode de Monte Carlo classique.

Par ailleurs, que l'erreur soit grande (n petit) ou petite (n grand), la différence entre les deux courbes est moindre.

Or, nous recherchons à avoir l'erreur la plus petite possible, et donc avoir des n grands.

On peut donc, encore une fois, conclure que la méthode Romberg Statistique est plus rapide que la méthode de Monte Carlo classique pour un n grand.

Annexe A : Calcul de la variance

On souhaite calculer $\mathbb{V}(\cos(\theta + W_T))$ pour notre exemple théorique.

$$\mathbb{V}(\cos(\theta + W_T)) = \mathbb{E}(\cos^2(\theta + W_T)) - [\mathbb{E}(\cos(\theta + W_T))]^2$$

On a :

$$\begin{aligned} \mathbb{E}(\cos^2(\theta + W_T)) &= \frac{1}{\sqrt{2\pi}} \int_{\mathbb{R}} \cos^2(\theta + \sqrt{t}x) e^{-\frac{x^2}{2}} dx \\ &= \frac{1}{2} + \frac{1}{2\sqrt{2\pi}} \int_{\mathbb{R}} \cos(2\theta + 2\sqrt{t}x) e^{-\frac{x^2}{2}} dx \\ &= \frac{1}{2} + \frac{1}{2\sqrt{2\pi}} \operatorname{Re} \left[\int_{\mathbb{R}} e^{2i\theta + 2i\sqrt{t}x} e^{-\frac{x^2}{2}} dx \right] \\ &= \frac{1}{2} + \frac{1}{2\sqrt{2\pi}} \operatorname{Re} \left[e^{2i\theta} \int_{\mathbb{R}} e^{2i\sqrt{t}x - \frac{x^2}{2}} dx \right] \\ &= \frac{1}{2} + \frac{1}{2\sqrt{2\pi}} \operatorname{Re} \left[e^{2i\theta} \mathbb{E}(e^{2i\sqrt{t}\xi}) \right] \quad \text{avec } \xi \sim \mathcal{N}(0, 1) \\ &= \frac{1}{2} + \frac{1}{2\sqrt{2\pi}} \operatorname{Re} \left[e^{2i\theta} \phi_{\xi}(2\sqrt{t}) \right] \quad \text{avec } \phi_{\xi}(u) = e^{-\frac{u^2}{2}} \\ &= \frac{1}{2} + \frac{1}{2} \operatorname{Re} \left[e^{2i\theta} e^{-2t} \right] \\ &= \frac{1}{2} + \frac{1}{2} e^{-2t} \cos(2\theta) \\ &= \frac{1}{2} (1 + \cos(2\theta) e^{-2t}) \end{aligned}$$

Par ailleurs,

$$\begin{aligned} [\mathbb{E}(\cos(\theta + W_T))]^2 &= \frac{1}{\sqrt{2\pi}} \int_{\mathbb{R}} \cos(\theta + \sqrt{t}x) e^{-\frac{x^2}{2}} dx \\ &= \frac{1}{\sqrt{2\pi}} \operatorname{Re} \left[\int_{\mathbb{R}} e^{i\theta + i\sqrt{t}x} e^{-\frac{x^2}{2}} dx \right] \\ &= \frac{1}{\sqrt{2\pi}} \operatorname{Re} \left[e^{i\theta} \int_{\mathbb{R}} e^{i\sqrt{t}x - \frac{x^2}{2}} dx \right] \\ &= \operatorname{Re} \left[e^{i\theta} \mathbb{E}(e^{i\sqrt{t}\xi}) \right] \quad \text{avec } \xi \sim \mathcal{N}(0, 1) \\ &= \operatorname{Re}(e^{i\theta} e^{-\frac{t}{2}}) \\ &= \cos(\theta) e^{-\frac{t}{2}} \end{aligned}$$

On en déduit donc que : $\mathbb{V}(\cos(\theta + W_T)) = \frac{1}{2} (1 + \cos(2\theta) e^{-2T}) - \cos^2(\theta) e^{-T}$ \square