

N° d'ordre : 29/2013-M/MT

RÉPUBLIQUE ALGÉRIENNE DÉMOCRATIQUE ET POPULAIRE
MINISTÈRE DE L'ENSEIGNEMENT SUPÉRIEUR ET DE LA RECHERCHE SCIENTIFIQUE

UNIVERSITÉ DES SCIENCES ET TECHNOLOGIE DE HOUARI BOUMEDIENNE
FACULTÉ DES MATHÉMATIQUES



MÉMOIRE
PRÉSENTÉ POUR L'OBTENTION DU DIPLÔME DE MAGISTER.

EN MATHÉMATIQUES
SPÉCIALITÉ : PROBABILITÉ & STATISTIQUE

Par: **BENMENSOUR Madina**

Sujet

Statistique des trajectoires pour des modèles
de diffusion en finance

Soutenu publiquement le 03/07/2013, devant le jury composé de :

Mr. Abdelkader TATACHAK	Maître de Conférences/A	à l'USTHB	Président.
Mr. Kamal BOUKHETALA	Professeur	à l'USTHB	Directeur de Mémoire.
Mr. Rachid OUAFI	Maître de Conférences/A	à l'USTHB	Examineur.

Table des matières

Introduction Générale	i
1 Concept général de processus et calcul stochastique	1
1.1 Introduction	1
1.2 Principaux aspects et définitions des processus stochastiques	2
1.2.1 Processus stochastique	2
1.2.2 La σ -algèbre et la Filtration	2
1.2.3 Processus Gaussien	3
1.2.4 Processus adapté	3
1.2.5 Processus stochastiques particuliers	4
1.3 Le mouvement Brownien	5
1.3.1 L'histoire du mouvement brownien	5
1.3.2 Passage à une description en temps continu	6
1.3.3 Définition d'un mouvement Brownien	7
1.3.4 Processus de Wiener et bruit blanc	8
1.3.5 Propriétés d'un mouvement brownien	8
1.3.6 Abandon de l'intégrale de Riemann	9
1.3.7 Simulation d'une trajectoire du processus de Wiener (Approximation du mouvement brownien par une marche aléatoire)	11
1.4 Introduction aux Equations différentielles stochastiques	12

1.4.1	Introduction à l'intégration stochastique	13
1.4.2	Intégrale stochastique de fonctions en escaliers	17
1.4.3	Propriétés d'une intégrale stochastique	18
1.4.4	Intégrale stochastique de fonctions appartenant à L^2	18
1.4.5	Différentielle stochastique	19
1.5	Calcul d'Itô	20
1.5.1	Formule d'Itô	20
1.6	Processus de Markov et processus de diffusion	22
1.6.1	Processus de Markov	22
1.6.2	Processus de diffusion	24
1.7	Conclusion	33
2	Equations différentielles stochastiques	34
2.1	Introduction	34
2.2	Equations différentielles de type linéaire	35
2.2.1	Résolution des EDS de type linéaire	35
2.2.2	Solutions générales de quelques EDS de type linéaire	39
2.2.3	Exemple d'application d'une EDS linéaire	40
2.2.4	Espérance et moment d'ordre deux d'un processus de diffusion	41
2.3	Résolution des équations différentielles stochastiques réductibles d'Itô	43
2.4	Existence et Unicité des solutions des équations différentielles stochastiques	48
2.5	L'Approximation des EDS par discrétisation du temps	49
2.5.1	Développement de Taylor Stochastique	49
2.5.2	La convergence	51
2.5.3	Les schémas du discretisation	52
2.5.4	Simulation des EDS	55

2.6	Conclusion	59
3	Modèle de Black-Scholes et estimation des paramètres	60
3.1	Introduction	60
3.2	Présentation du modèle de Black & Scholes	60
3.2.1	Les conditions du modèle	62
3.2.2	Formule de Black-Scholes	66
3.2.3	Le modèle de Black et Scholes en pratique	71
3.2.4	Modèle de Black-Scholes pour la valeur d'un titre	76
3.3	Estimation des paramètres du modèle de Black-Scholes	82
3.3.1	La méthode historique basées sur la fonction de vraisemblance	82
3.3.2	L'estimation des paramètres des EDS utilisant deux fonctions critères	91
3.4	Conclusion	109
4	Conclusion générale	110

Introduction Générale

Dans le monde d'aujourd'hui, la finance joue un rôle des plus importants et est parfois à l'origine de crises mondiales. Il apparaît alors important que la finance soit basée sur des modèles solides permettant d'évaluer les risques et les prix.

Dans une économie de marché moderne, les agents désirent choisir le niveau de risque approprié à la gestion de leurs transactions. C'est pourquoi, ils cherchent à se couvrir contre les événements défavorables sur les marchés financiers qui redistribuent les risques.

Les actifs financiers sont régis par des comportements aléatoires qui traduisent la complexité du monde économiques et politique. La mesure et la gestion des risques sont ainsi devenues des enjeux majeurs pour les opérateurs des marchés financiers, et intéressent les chercheurs du laboratoire de probabilités et modèles aléatoires. Les outils mathématiques que ces derniers développent offrent une modélisation et des méthodes quantitatives adaptées à la description et contrôle des risques financiers ,en particulier ceux liés aux problèmes des options.

La modélisation log- normale de l'actif sous-jacent est également un biais puisque les événements à chocs ne sont pas pris en compte. Si on tient compte de ces événements rares, on peut alors montrer que le prix des options incluant ces chocs à un prix supérieur aux prix des options évaluées par le modèle de Black et Scholes. Les actifs contingents sont donc sous évalués par le modèle de Black et Scholes.

L'étude la plus reconnue dans le monde de la mathématique financière est certainement le modèle d'évaluation des options de Black-Scholes (1973), qui a connu un succès immédiat auprès des chercheurs et des professionnels de la finance, tant le besoin était grand d'outil de gestion des risques. Le modèle de Black- Scholes permet d'évaluer le prix d'une option dans le cas de non arbitrage, en se basant sur l'hypothèse que l'évolution d'un titre suit un mouvement brownien géométrique et que la volatilité est constante.

Le modèle et la formule de Black-Scholes s'est imposée comme référence depuis 1973, dans le

calcul d'option. Malgré ses défauts, ce modèle connaît ce succès car il possède de nombreux avantages : sa simplicité d'application et de formule, son importante utilisation par les opérateurs du marché.

Cependant, le premier problème, auquel on se trouve confondu, et qui concerne les modèles en général est celui de l'estimation des paramètres. Ce problème, purement technique, doit être correctement traité si l'on souhaite disposer d'un instrument fiable de gestion de portefeuille.

Ce mémoire porte sur l'estimation des paramètres du modèle de Black et Scholes (Brownien géométrique) : volatilité et taux d'intérêt, deux méthodes d'estimation ont été présentées l'une est basée sur la Statistique de Chi-deux, l'autre est basée sur la Statistique de Kolmogorov-Smirnov ainsi l'estimation du maximum de vraisemblance (MLE), [26,27].

Dans cette étude, nous allons commencer par présenter au premier chapitre les principaux aspects et définitions des processus stochastiques, plus particulièrement le mouvement Brownien. Ce dernier constitue la base de la théorie classique des EDS et comme sa trajectoire est non différentiable alors il était indispensable de se tourner vers une nouvelle description mathématique qui est l'intégration stochastique. Concernant ce dernier on présentera les propriétés élémentaires de l'intégrale stochastique après l'avoir définis. Notons que cette intégrale n'obéit pas aux règles de calcul différentiel classique étant donné que le processus de Wiener est à variation non bornée. Ceci explique le passage vers un nouveau calcul, dit calcul d'Itô, basé sur la formule d'Itô qui permet de différentier une fonction deux fois continûment différentiable. Nous terminons ce chapitre par donner un petit aperçu sur le processus de diffusion qui est solution d'une EDS et sur sa continuité et on montrera que le processus de Wiener n'est qu'un cas particulier du processus de diffusion.

Dans le deuxième chapitre, nous parlerons de la méthode de résolution des équations différentielles stochastiques basée essentiellement sur le calcul d'Itô, les équations de type linéaire seront traitées en premier lieu en donnant les solutions exactes de divers EDS, suivies par un exemple d'application. Par ailleurs, nous allons voir comment peut-on passer d'une EDS s'écrivant en fonction du processus de diffusion à une EDS décrite en fonction de l'espérance et de la variance de ce processus si l'objectif principal est d'évaluer ces deux paramètres. Nous passerons dans le même chapitre aux méthodes de résolutions des EDS de type non linéaire (réductible). Ces méthodes consistent à transformer l'EDS en question, en

se basant sur la formule d'Itô, en équation explicitement intégrable.

Enfin dans ce chapitre nous présentons trois schèmes d'approximations (Euler, Milstein et Taylor d'ordre 1.5), en se basant sur le développement de Taylor stochastique, car la résolution des équations différentielles stochastiques d'une manière explicite est rare dans la pratique, pour surmonter ce problème en élaborant des schèmes d'approximations. Nous terminons ce chapitre par un exemple d'application simple, considérons le cas du mouvement brownien géométrique.

Dans le troisième chapitre, nous présentons le modèle de Black & Scholes qui reste le plus utilisé dans le domaine de l'ingénierie financière, ainsi la formule de Black et Scholes qui a été développée pour des actifs dont le comportement est celui du mouvement Brownien géométrique et qui permet de calculer le prix d'une option. Donnons l'exemple d'un call européen, nous allons calculer ce prix en appliquant la formule de Black & Scholes puis la méthode de Monte-Carlo.

Ensuite, nous allons voir comme peut-on utiliser des idées développées dans le modèle de Black-Scholes dans des modèles d'assurance-vie. Nous prenons comme exemple le prix de la garantie plancher qui est donnée en fonction des options de vente (puts).

Nous passerons dans la même section à présenter l'importance de la volatilité avec une simple estimation de la volatilité historique et implicite.

La section 2 de ce chapitre sera consacrée à l'étude des séries de rendement logarithmique, ensuite, on traite les techniques d'estimation du modèle de Black & Scholes, nous commençons par un bref aperçu d'une méthode courante (la méthode historique basées sur la fonction de vraisemblance) et deux autres méthodes basées sur une fonction critère, ces méthodes consiste à contourner l'estimation de la fonction de densité de transition entièrement pour estimer simultanément les paramètres du modèle. La première utilise une fonction critère basée sur la statistique de Chi-deux, la seconde, propose l'estimation des paramètres du modèle à l'aide d'une fonction critère basée sur la statistique du Kolmogorov-Smirnov, chaque méthode suivie par une application.

Nous finirons ce travail par une estimation des paramètres des EDS non linéaires et qui n'ont pas de solution explicite, alors nous a permis d'utiliser l'ensemble des schèmes d'approximations suscités, nous utilisons la Statistique de Chi-deux ainsi la Statistique du Kolmogorov-Smirnov pour l'estimation des paramètres et nous terminerons avec une conclusion générale.

Chapitre 1

Concept général de processus et calcul stochastique

1.1 Introduction

Ce chapitre est consacré à quelques notions de base du calcul stochastique, nous présentons de façon classique un rappel sur les principaux aspects et définitions des processus stochastiques qui sont des outils largement utilisés en pratique. On y trouve des applications dans plusieurs domaines notamment en physique, en mécanique, en écologie et bien entendu dans le domaine économique et financier. La modélisation de phénomènes en temps continu faisant intervenir le hasard, tels le cours des produits financiers (actifs, options, taux par exemple) par des processus stochastiques. La plus ancienne modélisation par un processus stochastique à temps continu a été faite sur un phénomène biologique, il s'agit du mouvement Brownien qui constitue la base de la théorie classique des équations différentielles stochastiques, ce processus sera présenté à la deuxième section et comme sa trajectoire est non différentiable alors nous présentons l'intégration stochastique et le calcul d'Itô qui base sur une formule dit "*formule d'Itô*". Dans la dernière section de ce chapitre on parlera du processus de diffusion et on montrera que le mouvement Brownien standard n'est qu'un cas particulier du processus de diffusion.

1.2 Principaux aspects et définitions des processus stochastiques

Dans cette partie, nous commençons par présenter des notions élémentaires des processus stochastiques afin de les introduire aux processus de diffusion et aux processus d'Itô. Ces processus servent à modéliser plusieurs phénomènes réels notamment les modèles d'actifs financiers.

1.2.1 Processus stochastique

Tout processus décrivant l'évolution chronologique d'un phénomène aléatoire est dit processus stochastique ou processus aléatoire, noté $X(t, \omega)$, X_t ou $X(t)$.

$X(t, \omega)$ est défini comme suit :

$$\begin{aligned} X: \quad \mathbb{R}^+ \times \Omega &\longrightarrow \mathbb{R} \\ (t, \omega) &\longrightarrow X(t, \omega) \end{aligned} \quad (1.1)$$

Cette application est mesurable lorsque l'on munit $\mathbb{R}^+ \times \Omega$ de la tribu $B(\mathbb{R}^+) \times A$. $B(\mathbb{R}^+)$ est la tribu des Boréliens.

·Pour t fixé, $X(t) = X(t, \cdot)$ est une variable aléatoire sur un espace de probabilité (Ω, A, P) .

·Pour ω fixé, $X(\cdot, \omega)$ est une réalisation ou une trajectoire du processus stochastique.

Par ailleurs, on peut définir les processus aléatoires en faisant appel à la notion dite '*Filtration*'.

1.2.2 La σ -algèbre et la Filtration

Soit (Ω, A, P) un espace de probabilité.

La σ -algèbre représente la famille d'évènements auxquels il est possible d'associer des probabilités.

La filtration est une famille croissante de sous σ -algèbre de A , noté généralement $(F_t, t \geq 0)$.

La tribu F_t n'est qu'une description mathématique de la somme de l'information dont on dispose à l'instant t . Cette information nous permet d'attribuer des probabilités cohérentes aux évènements pouvant intervenir.

1.2.3 Processus Gaussien

Définition 1.1

Un vecteur de v.a. (X_1, \dots, X_n) est dit gaussien si toute combinaison linéaire de ses composantes est une variable aléatoire gaussienne.

Les deux propositions suivantes très utiles dans la manipulation des vecteurs gaussiens.

Proposition 1.1

Si le vecteur (X_1, X_2) est gaussien, les v.a. X_1 et X_2 sont indépendantes si et seulement si $cov(X_1; X_2) = 0$.

Proposition 1.2

Tout vecteur de variables aléatoires gaussiennes indépendantes est un vecteur gaussien.

Remarque 1.1

Il est possible de trouver des variables aléatoires gaussiennes non indépendantes de covariance nulle.

La version continue des vecteurs gaussiens sont les processus gaussiens.

Définition 1.2

Un processus $(X_t)_{t \geq 0}$ est appelé processus gaussien si pour tout n et tout $t_1 \dots t_n$ le vecteur $(X_{t_1}, \dots, X_{t_n})$ est gaussien.

1.2.4 Processus adapté

Un processus adapté est celui pour lequel une description probabiliste est réalisable. Mathématiquement cela est traduit comme suit :

Un processus $X(t)$ est dit adapté à la filtration $(F_t, t \geq 0)$ si la variable aléatoire $X(t)$, t fixé, est mesurable pour la σ -algèbre F_t .

1.2.5 Processus stochastiques particuliers

Soit, $X(t)$ un processus stochastique et t_1, t_2, \dots, t_k ($k \in \mathbb{N}$) une succession de valeurs du temps.

Processus stationnaire

On considère les instants $t_1 + \tau, t_2 + \tau, \dots, t_k + \tau, \forall \tau > 0$.

Le processus $X(t)$ est dit stationnaire lorsque les variables aléatoires $[X(t_1), X(t_2), \dots, X(t_k)]$ et $[X(t_1 + \tau), X(t_2 + \tau), \dots, X(t_k + \tau)]$ ont la même distribution de probabilité. Cela signifie qu'à priori, quel que soit l'horizon considéré on met la même distribution de probabilité sur l'état du système (même moyenne et même variabilité).

Processus à accroissement indépendants

Un processus $X(t)$ est à accroissement indépendants si $\forall 0 < t_1 < t_2 < \dots < t_k, k \in \mathbb{N}$, les variables $X(t_k) - X(t_{k-1}), \dots, X(t_2) - X(t_1), X(t_1) - X(0)$ sont indépendantes. Cela s'interprète par le fait que les états décrits par ce type de processus évoluent dans un univers où les variations passées n'influencent pas les variations futures.

Processus à accroissement stationnaires

Considérons à présent deux instants θ_1 et θ_2 , supposons $\theta_1 \leq \theta_2$ et notons F la fonction de répartition d'une variable aléatoire.

Le processus $X(t)$ est à accroissement stationnaire si la distribution associée à l'accroissement de la variable sur l'intervalle $[\theta_1, \theta_2]$ ne dépend que de la durée écoulée ($\theta_2 - \theta_1$).

En terme d'équation on écrit :

$$F(\varepsilon, \theta_1, \theta_2) = P\{X(\theta_2) - X(\theta_1) < \varepsilon\} = F^*(\varepsilon, \theta_2 - \theta_1) \quad (1.2)$$

Processus homogène

On appelle fonction de distribution de transition pour un processus aléatoire :

$$F(s, x; t, B) = P(X_t \in B / X_s \leq x) = \int_B p(s, x; t, y) dy \quad \text{ou} \quad s < t$$

$p(s, x; t, y)$ est dite fonction de densité de transition.

B est un sous ensemble de $B_{\mathbb{R}}$ ($B_{\mathbb{R}}$ est une σ -algèbre)

Définition 1.3

On dit que le processus est homogène si :

$$F(s, x; t, B) = F(0, x; t - s, B) = P(X_{t-s} \in B / X_0 = x) \quad \text{pour tout } t \text{ et } s, (t > s).$$

Ceci veut dire que la probabilité ne dépend que de l'instant considéré t et non pas de la position de l'intervalle par rapport à l'axe temporel.

Pour avoir plus de détail sur les processus stochastiques voir [6].

L'un des processus stochastiques à temps continu les plus importants et les plus utilisés est le mouvement Brownien, c'est le plus célèbre des processus stochastiques et sans doute le plus central en théorie des probabilités et le premier processus qui ait été introduit pour modéliser un phénomène réel autre qu'un jeu de hasard.

Le mouvement Brownien est un exemple de processus stochastique à accroissement indépendant et stationnaire.

1.3 Le mouvement Brownien**1.3.1 L'histoire du mouvement brownien**

Le mouvement brownien est le nom donné aux trajectoires irrégulières du pollen en suspension dans l'eau, observé par le *botaniste Robert Brown* en 1828. En 1877, *Delsaux* explique que ce mouvement "aléatoire", dû aux chocs successifs entre le pollen et les molécules d'eau, entraîne la dispersion ou diffusion du pollen dans l'eau (changements incessants de direction). Le champ d'application du mouvement brownien est beaucoup plus vaste que l'étude des particules microscopiques en suspension et inclut la modélisation du prix des

actions, du bruit thermique dans les circuits électriques, du comportement limite des problèmes de files d'attente et des perturbations aléatoires dans un grand nombre de systèmes physiques, biologiques ou économiques.

En 1900 : *Louis Bachelier* dans sa thèse "Théorie de la spéculation" modélise les courbes de la bourse comme des processus à accroissements indépendants et gaussiens.

C'est *A.Einstein* qui, en 1905, a donné la première explication physique claire de ce phénomène, il a obtenu la densité de probabilité de transition du mouvement brownien à partir de la théorie moléculaire de la chaleur. Des vérifications expérimentales directes ont été effectuées en 1908 par *J.Perrin*.

Le premier traitement mathématique rigoureux est dû à *Norbert Wiener* (1923,1924), entre autre démonstration de l'existence du brownien.

A partir de 1925, *Nobert Wiener* jeta les fondements mathématiques du mouvement brownien, puis *Paul Lévy* étudia ses propriétés analytiques fines.

Depuis 1950, de nombreux travaux lui ont été consacrés, dont l'un des plus importants est le développement du calcul stochastique par *Kiyoshi Itô*.

Un mouvement brownien généralement noté B pour *Brown* ou W pour *Wiener*.

De manière générale, toutes les quantités observées expérimentalement ou variables macroscopiques, sont accompagnées de fluctuations dues au mouvement thermique des degrés de liberté microscopiques de la matière. Dans la plupart des cas, les fluctuations des variables macroscopiques sont extrêmement petites par rapport à leurs valeurs moyennes et peuvent être négligées. Cependant, puisqu'elles reflètent les mouvements à l'échelle microscopique dans le système considéré, leur analyse est importante pour l'étude de celui-ci.

Les grandeurs financières comme le prix d'une action cotée en Bourse fluctuent de façon analogue au mouvement brownien, sous l'effet cumulé des achats et ventes réalisées sur le marché par des nombreux opérateurs.

1.3.2 Passage à une description en temps continu

Le mouvement brownien a joué un grand rôle historique en mathématiques : c'est en effet pour représenter la position d'une particule brownienne qu'un processus stochastique a été construit, pour la première fois. Clairement, pour pouvoir associer à un phénomène physique un processus aléatoire, il est nécessaire de passer de la description en temps discret imposée

par l'expérience à une description en temps continu. Supposons par exemple que l'on observe au microscope une particule brownienne pendant un intervalle de temps $0 \leq t \leq T$, et que l'on enregistre en fonction du temps la projection de sa position sur un axe x . Répétant N fois les observations au cours du temps, on obtient N valeurs de la coordonnée de la particule, $x(t_1), x(t_2), \dots, x(t_N)$.

La valeur $x(t)$ de la coordonnée de la particule brownienne au temps t est une réalisation d'une variable aléatoire et chacune des séries observées $x(t_j)$ est un échantillon d'un ensemble statistique. Si l'on pouvait effectuer une observation continue, on obtiendrait une fonction aléatoire du temps ou processus stochastique $X(t)$ avec t comme paramètre continu. En pratique, on effectue les observations à des temps discrets $t_1 < t_2 < \dots < t_N$, obtenant ainsi un ensemble de N nombres, $x(t_1), x(t_2), \dots, x(t_N)$. La description mathématique par un processus en temps continu se fait en prenant N très grand et des intervalles entre les temps d'observation de plus en plus petits.

1.3.3 Définition d'un mouvement Brownien

Le processus $X(t)$, $t \in [0, T]$ est dit un mouvement brownien si $X(t)$ satisfait aux conditions :

- i.** $X(0) = 0$. Cette condition exprime que l'on impose à la perturbation de ne pas agir aléatoirement à l'origine des temps.
- ii.** $\forall t \in [0, T]$, la distribution de $X(t)$ est normale. (le choix de normalité est une conséquence du théorème central limite).
- iii.** L'espérance de $X(t)$ est nulle $\forall t \in [0, T]$ (en moyenne l'erreur par rapport au modèle déterministe est nulle).
- iv.** Le processus $X(t)$ est de variance égale au temps, alors le processus $X(t)$ est appelé processus de Wiener ou mouvement brownien standard (MBS), noté $W(t)$ ou $B(t)$, qui satisfait :

$$E(W(t)) = 0 \quad \text{et} \quad E(W^2(t)) = t$$

pour la démonstration voir [4]

1.3.4 Processus de Wiener et bruit blanc

En pratique et en particulier en économie, la brusque hausse ou baisse d'un paramètre quelconque décrivant ainsi des perturbations inattendues est modélisé au moyen d'un bruit blanc qui est en relation avec le processus de Wiener. En effet, on ne peut jamais calculer la valeur $U_t(w)$ d'un bruit blanc $U(t)$ mais par contre, il est possible de le rencontrer comme intégrant sous la forme :

$$\int_s^t A(\alpha)U_\alpha d\alpha.$$

$U_\alpha d\alpha$ est remplacé formellement par la différentielle stochastique d'un processus à accroissement indépendants et stationnaire dW_α dont la mesure positive associée est la mesure de **Lebesgue** et l'intégrale devient "une intégrale stochastique de Wiener", que nous présenterons dans la section 4

$$\int_s^t A(\alpha)U_\alpha d\alpha = \int_s^t A(\alpha)dW_\alpha \quad (2.1)$$

1.3.5 Propriétés d'un mouvement brownien

Le processus Brownien possède des propriétés très importantes qu'on peut trouver dans beaucoup de phénomènes réels.

a. Nous supposons dans la suite que les processus de Wiener sont continus car si nous considérons les processus prévisibles (la connaissance du passé permet la description du présent) qui sont des processus continus à gauche ou contrairement, les processus optionnels qui sont continus à droite, le support mathématique sera infiniment complexe.

b. La variance d'un processus de Wiener croit proportionnellement au temps. En effet, le processus $X(t)$ étant à accroissement indépendant (la covariance est nulle) et à accroissement stationnaire, on a alors :

$$\begin{aligned} \text{Var}[X(t+s)] &= \text{Var} [X(t) + (X(t+s) - X(t))] \\ &= \text{Var} [X(t)] + \text{Var} [X(t+s) - X(t)] = \text{Var} [X(t)] + \text{Var} [X(s)] \end{aligned}$$

c. Un processus de Wiener possède des trajectoires presque sûrement à variation non bornée. Presque toutes les trajectoires sont partout non différentiables.

En effet, une description correspondante à certaines exigences indispensables ne pouvait se faire au moyen de l'intégrale ordinaire de fonctions classiques, généralement à variations bornées. Pour cela, nous allons mettre en exergue les motivations mathématiques qui ont conduit à la création d'une nouvelle intégrale (intégrale stochastique) et donc l'abandon de l'intégral de Riemann.

1.3.6 Abandon de l'intégrale de Riemann

Considérons un modèle mathématique devant décrire l'état $X(t)$ d'un système au cours du temps. Supposons qu'à tout instant, les altérations soient fonction d'influences prévisibles et d'autres aléatoires. Mathématiquement, cette hypothèse conduit à l'équation différentielle :

$$dX(t) = a(x, t)dt + b(x, t)dt \quad (2.2)$$

où $a(x, t)$ décrit les fluctuations prévisibles et $b(x, t)$ les fluctuations aléatoires.

On suppose l'état du système connu à l'instant initial $t = 0$. On pose $X(0) = x_0$.

Proposition 1.3

Si les perturbations aléatoires non prévisibles sont décrites au moyen des équations différentielles ordinaires qui respectent des conditions réalistes alors elles sont de variance nulle et donc déterministes.

Preuve

Soit, les processus à variation cumulé : $u(x, t) = \int_0^t b(x, s)ds$, où le $b(x, t)$ satisfait aux conditions suivantes :

i. $b(x, t)$ indépendante de $b(x, s)$ pour tout $t \neq s$.

ii. $b(x, t)$ est de carré intégrable.

iii. $E[b(x, t)] = 0$

iv. $E[b^2(x, t)] \leq k$

Le processus u existe toujours car il vérifie la condition ii et iii et de la propriété de linéarité de l'intégrale, on montre que $E[u(x, t)] = \int_0^t E[b(x, s)]ds = 0$.

Pour le calcul de la variance, il faut passer par la définition de l'intégrale de Riemann. On a :

$$Var[u(x, t)] = E[u^2(x, t)] = E\left[\int_0^t b(x, s)ds\right]^2 = \lim_{v \rightarrow 0} E\left[\sum_{i=1}^n b(x, t_i^*) \cdot (t_{i+1} - t_i)\right]^2$$

Où : v est la norme de subdivision ($v = \text{Max}_i(t_{i+1} - t_i)$).

t_i^* est un point intermédiaire de l'intervalle $[t_i, t_{i+1}[$.

En développant le carré, on obtient :

$$Var[u(x, t)] = \lim_{v \rightarrow 0} E\left[\sum_{i=1}^n \sum_{j=1}^n b(x, t_i^*) b(x, t_j^*) \cdot (t_{i+1} - t_i)(t_{j+1} - t_j)\right]$$

De l'hypothèse **i**. $E[b(x, t_i^*) b(x, t_j^*)] = 0 \quad \forall i \neq j$ et la variance devient donc :

$$Var[u(x, t)] = \lim_{v \rightarrow 0} E\left[\sum_{i=1}^n b^2(x, t_i^*) \cdot (t_{i+1} - t_i)^2\right] = \lim_{v \rightarrow 0} \left[\sum_{i=1}^n E[b^2(x, t_i^*)] \cdot (t_{i+1} - t_i)^2\right]$$

Comme $E[b^2(x, t_i^*, \omega)]$ est bornée (Hypothèse **iv**.), on a

$Var[u(t)] \leq k \lim_{v \rightarrow 0} [\sum_{i=1}^n (t_{i+1} - t_i)^2]$, où k est un majorant pour les variances de b .

En choisissant une subdivision en n des intervalles de longueurs égales à (t/n) , on a

$$\lim_{n \rightarrow \infty} \left[\sum_{i=1}^n (t/n)^2\right] = \lim_{n \rightarrow \infty} [n \cdot (t/n)^2] = 0$$

Ce qui montre que l'intégrale de Riemann se révèle inutilisable.

Il est donc indispensable de se tourner vers une nouvelle description mathématique qui rendra possible la description d'un univers aléatoire.

1.3.7 Simulation d'une trajectoire du processus de Wiener (Approximation du mouvement brownien par une marche aléatoire)

On peut approcher un processus de Wiener W_t sur un intervalle de temps fini à une chaîne de Markov, en l'occurrence une marche aléatoire [13].

Considérons une succession de n pas de longueur Δt , fonction de la finesse du graphisme désiré, effectuée aléatoirement et indépendamment tantôt vers la gauche et tantôt vers la droite.

La variable position après n pas qu'on veut étudier peut être vue comme la somme de n variables aléatoires indépendantes :

$$W_{n\Delta t} = \sum_{k=1}^n (W_{k\Delta t} - W_{(k-1)\Delta t}) = \sum_{k=1}^n U_k, \quad U_k \sim N(0, \Delta t)$$

On déduit que $W_{n\Delta t}$ est une promenade aléatoire : $S_n = U_1 + U_2 + \dots + U_n \Rightarrow S_n = S_{n-1} + U_n$ (formule récurrente).

Donc pour simuler la trajectoire du processus Brownien sur l'intervalle $[T_0, T]$, il suffit de simuler une variable normale de moyenne nulle et de variance $\Delta t = \frac{T-T_0}{n}$ et les points de la trajectoire de la chaîne sont ensuite reliés par des droites affines.

Avec les mêmes démarches suivies dans la simulation d'une trajectoire d'un mouvement Brownien on peut simuler un flux de trajectoires.

Les trajectoires browniennes sont simulées pour $n = 1000$.

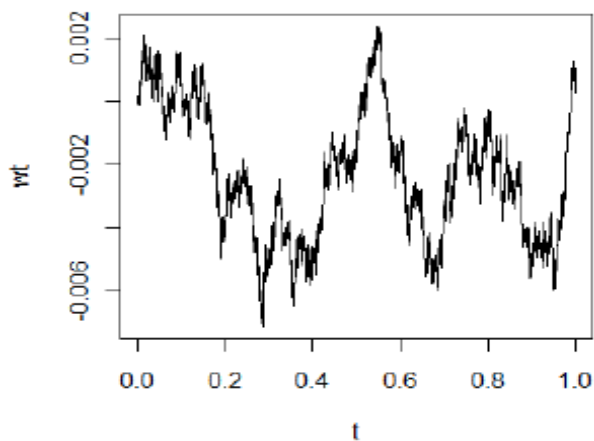


Figure 1.1- *Simulation d'une trajectoire Brownienne sur $[0,1]$*

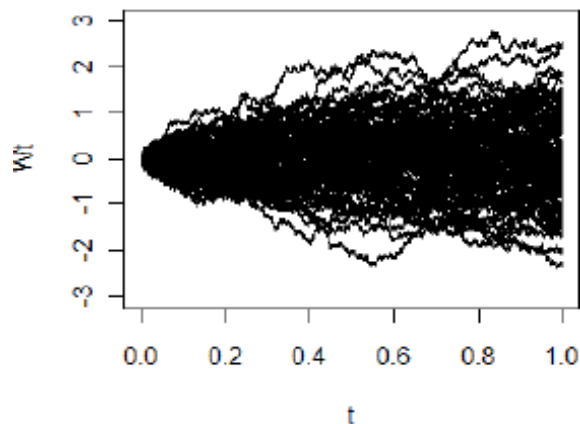


Figure 1.2—*Simulation de 100 trajectoires Browniennes sur $[0,1]$*

Nous pouvons simuler facilement une trajectoire du processus de Wiener $\{W_t, t \geq 0\}$ dans l'intervalle de temps $[T_0, T]$ avec un pas $\Delta t = \frac{T-T_0}{n}$ à l'aide de la fonction *BMN* et la fonction *BMNF* permet de simuler un flux brownienne standard. [30] et [31].

```
R> BMN(N = 1000, t0 = 0, T = 1, C = 1)
```

```
R> BMNF(N = 1000, M = 100, t0 = 0, T = 1, C = 1)
```

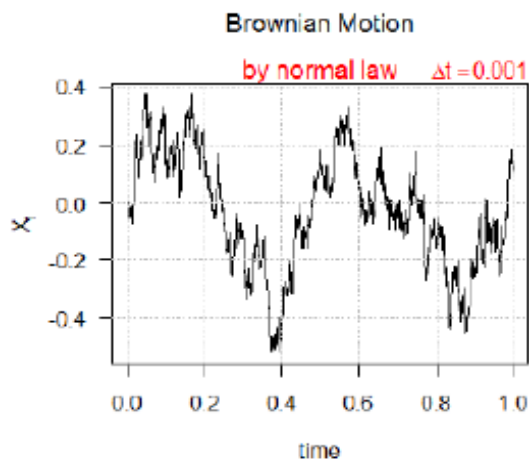


Figure 1.3-*Trajectoire brownienne simulée a partir d'une distribution gaussienne*

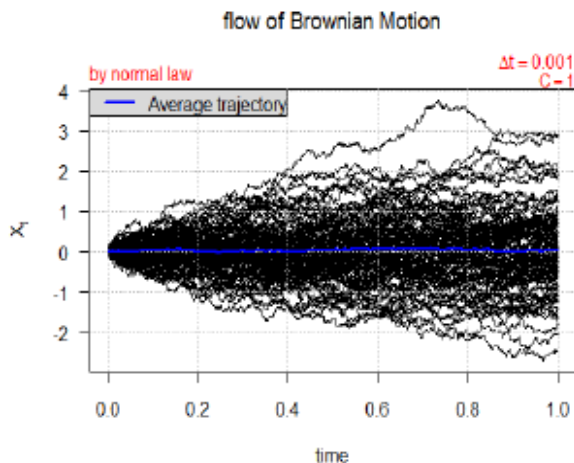


Figure 1.4-*Flux de trajectoires brownienne simulées a partir d'une distribution gaussienne*

En conclusion, le mouvement Brownien permet de modéliser d'une manière assez convaincante des phénomènes qui n'ont pas une description déterministe et ce en conservant une attitude réaliste. Cette modélisation est traduite par l'utilisation d'un outil puissant comme les équations différentielles stochastiques que nous allons présenter dans la section qui suit.

1.4 Introduction aux Equations différentielles stochastiques

Les équations différentielles stochastiques (EDS) est un outil mathématique conçu dans le but de modéliser de nombreux phénomènes pratiques dans différents domaines. L'inclusion des effets aléatoires dans ces équations conduit à deux catégories distinctes d'équations dont la solution est un processus aléatoire possédant une trajectoire différentiable ou non

différentiable, respectivement, ce qui nécessite différentes méthodes d'analyse. La première classe d'équations survient quand l'équation différentielle ordinaire possède des coefficients aléatoires, une valeur initiale aléatoire ou imposée par un processus stochastique assez régulier. Les équations de ce type sont appelées « Equations différentielles aléatoires » et se résolvent trajectoire par trajectoire comme dans les équations différentielles ordinaires (cas déterministe). Les chemins aléatoires des solutions et qui sont des processus sont donc des fonctions différentiables. Quant à la deuxième classe, les perturbations introduites dans les équations différentielles stochastiques sont des processus aléatoires irréguliers qui représentent un bruit blanc gaussien et sont appelées EDS de type Itô et c'est ce type d'équations que seront considérées dans cette étude. Dans le but de donner un sens à ces équations il est nécessaire de faire un rappel sur la notion d'intégrale stochastique. Le calcul stochastique d'Itô a permis d'étendre la notion de différentiation aux processus stochastique.

1.4.1 Introduction à l'intégration stochastique

Lorsque nous avons présenté les propriétés élémentaires du processus de Wiener, la construction intuitive que nous avons exposée était fondée sur les marches aléatoires particulières dont les accroissements prenaient deux valeurs seulement avec une espérance nulle et une variance proportionnelle à la durée séparant deux dates. Toutefois, les paramètres d'espérance et de variance par unité de temps sont constants dans le cas du mouvement brownien. Ceci constitue une restriction importante lorsque l'on cherche par exemple à modéliser la dynamique des taux d'intérêt dans laquelle il semble raisonnable, au vu des constatations empiriques, d'introduire une force de rappel induisant une espérance de variation négative quand les taux sont élevés et une espérance de variation positive quand ils sont faibles. De la même façon, pour ce qui concerne la variance par unité de temps, il est difficile de la supposer indépendante du niveau des taux. En particulier, lorsque les taux sont très faibles, on s'attend à observer une variance plus faible que quand ces taux sont très élevés. La définition d'un processus d'Itô suppose l'existence de fonctions $a(t, x)$ et $b(t, x)$ qui dépendent à la fois du niveau atteint par le processus considéré et de la date considérée. Il s'agit donc de bons candidats pour modéliser les dynamiques des variables financières comme les prix, les rentabilités, les taux d'investissement interne et les taux de change. Il est toutefois important de noter que le mouvement brownien général est un cas particulier de processus d'Itô.

Dans cette section, nous présentons comment est construit l'intégrale stochastique et le lemme d'Itô, encore appelé théorème fondamental du calcul stochastique.

Pour cela, nous considérons un phénomène pour lequel nous pouvons décrire son évolution au moyen des équations de type :

$$dX_t = a(t, X_t)dt + b(t, X_t)\varepsilon_t dt \quad (3.1)$$

où dX_t est la variation du phénomène X sur $[t, t + \Delta t]$ et $a(t, X_t)dt$ est une différentielle classique et $b(t, X_t)\varepsilon_t dt$ est un terme perturbateur à définir (terme de diffusion).

Les ε_t sont des variables aléatoires gaussiennes centrées réduites (ε_t est un bruit blanc).

Proposition 1.4

Lorsque $a = 0$ et $b = 1$, ε_t n'est autre qu'une dérivation formelle du processus de Wiener $W(t)$ qu'on notera W_t .

Preuve

Pour un processus aléatoire $(X_t)_{t \geq 0}$ stationnaire au sens faible, on définit une densité spectrale de ce processus par :

$$\begin{aligned} S(\omega) &= \int_{\mathbb{R}} C(k)e^{-2i\pi\omega k} dk \\ &= \int_{\mathbb{R}} C(k) \cos(2\pi\omega k) dk \end{aligned}$$

Où $C(k)$ est la fonction de covariance du processus $(X_t)_{t \geq 0}$.

*Un processus ε_t est qualifié de bruit blanc gaussien si :

- $E(\varepsilon_t) = 0$ et $E(\varepsilon_t^2) = \sigma^2 \quad \forall t > 0$
- ε_t et ε_s sont indépendants $\forall t \neq s$.
- $\varepsilon_t \sim N(0, \sigma^2)$.

Par analogie avec le bruit blanc en temps discret, défini comme une suite de variables aléatoires, centrées, du second ordre et indépendantes, on cherche à définir $(\varepsilon_t)_{t \geq 0}$ comme un processus stochastique vérifiant $\forall t > 0$ et $\forall k > 0$:

$$-E(\varepsilon_t) = 0.$$

$$-E(\varepsilon_t \varepsilon_{t+k}) = C(k) = \delta_0(k), \quad \forall \delta_0 \text{ est la mesure de Dirac en } 0.$$

Un tel processus n'existe pas. Son idéalisation est la dérivée d'un mouvement brownien standard. Si pour $h > 0$ fixé, on considère le processus

$$\varepsilon_t = \frac{W_{t+h} - W_t}{h}$$

Il est facile de montrer grâce aux propriétés de mouvement brownien que :

$$1-E(\varepsilon_t^h) = 0.$$

$$2-E(\varepsilon_t^h \varepsilon_s^h) = \frac{1}{h} \max(0, 1 - \frac{|t-s|}{h}) = c(t-s).$$

Démonstration

$$1-E(\varepsilon_t^h) = 0. \text{ évident.}$$

2- On distingue deux cas :

*1^{er} cas : $t \geq s+h$ c'est à dire les accroissements $W_{t+h} - W_t$ et $W_{s+h} - W_s$ sont indépendants :

$$E(\varepsilon_t^h \varepsilon_s^h) = \frac{1}{h^2} E[(W_{t+h} - W_t)(W_{s+h} - W_s)] = 0$$

*2^{ème} cas : $s \leq t \leq s+h$ c'est à dire les accroissements $W_{t+h} - W_t$ et $W_{s+h} - W_s$ ne sont pas indépendants :

$$\begin{aligned} E(\varepsilon_t^h \varepsilon_s^h) &= \frac{1}{h^2} E[(W_{t+h} - W_t)(W_{s+h} - W_s)] \\ &= \frac{1}{h^2} E[(W_{t+h} - W_t) + (W_t - W_s)(W_{t+h} - W_{s+h})(W_{s+h} - W_t)] \\ &= \frac{1}{h^2} E[W_{s+h} - W_t]^2 \\ &= \frac{1}{h^2} (s+h-t) \\ &= \frac{1}{h} \left(1 + \frac{|s-t|}{h} \right) \end{aligned}$$

Donc

$$E(\varepsilon_t^h \varepsilon_s^h) = \frac{1}{h} \max(0, 1 - \frac{|t-s|}{h}) = c(t-s).$$

Quand $h \rightarrow 0$

$$S_{\varepsilon_t}(\omega) = \int_{-h}^h \frac{1}{h} \left(1 - \frac{|k'|}{h}\right) \cos(2\pi\omega k') dk' = \left(\frac{\sin 2\pi\omega h}{\pi\omega h}\right)^2$$

C'est-à-dire $E(\varepsilon_t^h \varepsilon_s^h)$ converge vers $\delta_0(k)$, $k = (t-s)$. Il est donc clair que la dérivée formelle $\frac{dW_t}{dt}$ a les propriétés d'un bruit blanc gaussien. Ce qui justifie l'affirmation concernant l'idéalisation du bruit blanc.

Donc on peut écrire formellement : $\varepsilon_t = \frac{dW_t}{dt}$.

L'équation différentielle **(3.1)** peut être exprimée en forme intégrale :

$$X_t(\omega) = X_{t_0}(\omega) + \int_{t_0}^t a(s, X_s) ds + \int_{t_0}^t b(s, X_s) \varepsilon_s ds \quad (3.2)$$

En effet, dans le cas particulier où $a = 0$ et $b = 1$, ε_t n'est autre qu'une dérivation formelle du processus de Wiener de moyenne nulle et de variance $|t-t_0|$. Par conséquent **(3.2)** s'écrit :

$$X_t(\omega) = X_{t_0} + \int_{t_0}^t a(s, X_s) ds + \int_{t_0}^t b(s, X_s) dW_s \quad (3.3)$$

Le problème dans cette équation est qu'il n'est pas possible de définir une différentielle trajectoire par trajectoire de $W(t)$ au sens classique de l'analyse car le mouvement Brownien n'admet pas de dérivée. En effet, un processus est dérivable en moyenne quadratique lorsque l'on peut trouver $W'(t)$ tel que :

$$E \left[\frac{W(t+h) - W(t)}{h} - W'(t) \right]_{h \rightarrow 0}^2 \rightarrow 0$$

$$\text{Or } E [W(t+h) - W(t)]^2 = E [W(h) - W(0)]^2 = E [W(h)]^2 = h$$

$$E \left[\frac{W(t+h) - W(t)}{h} \right]_{h \rightarrow 0}^2 = \frac{1}{h^2} \cdot h = \frac{1}{h} \rightarrow \infty$$

D'où la non dérivabilité en moyenne quadratique du processus $W(t)$ et donc il est impossible d'interpréter l'expression **(3.3)** sans introduire la notion d'intégrale stochastique.

1.4.2 Intégrale stochastique de fonctions en escaliers

Considérons l'intégrale d'une fonction aléatoire f sur un intervalle de temps $[0, 1]$ que nous noterons $I(f)$ où $I(f) = \int_0^1 f(s)dW_s$. [8]

Nous limitons la construction de l'intégrale aux fonctions en escaliers et pour lesquelles les définitions ci-dessous ont un sens.

Définition 1.4

Soit f est une fonction de $[\alpha, \beta] \times \Omega$ dans \mathbb{R} . $f(t) \in L^2[\alpha, \beta]$ si et seulement si :

- f est adapté à la filtration $\{F_t\}$.

- $\int_{\alpha}^{\beta} f^2(t)dt < 1$ avec une probabilité 1.

Cette définition signifie que nous limitons l'univers à celui des fonctions ayant un sens probabiliste et pour lesquelles le calcul des paramètres probabilistes usuels est possible.

Définition 1.5

La fonction f étant une fonction en escaliers. On peut définir f sur $[\alpha, \beta]$ en découpant cette intervalle de la façon classique : $\alpha = t_0 < t_1 < \dots < t_n = \beta$ tel que $f(t) = f_i, t \in [t_i, t_{i+1}[$

Nous signalons que les f_i sont des variables aléatoires (deux réalisations successives du processus donneront des valeurs différentes). Par contre, les instants t_i sont déterministes.

Définition 1.6

On appelle intégrale stochastique d'une fonction en escaliers l'expression

$$\int_{\alpha}^{\beta} f(t)dW(t) = \sum_{k=0}^{n-1} f(t_k) [W(t_{k+1}) - W(t_k)] \quad (3.4)$$

où f est mesurable

Remarque 1.2

Lorsque $f(t) = W(t) = W_t$, l'intégrale d'Itô aura la particularité suivante

$$\int_0^t W_s dW(s) = \frac{1}{2}W_t^2 - \frac{1}{2}t \quad \text{avec une probabilité 1}$$

1.4.3 Propriétés d'une intégrale stochastique

a. La moyenne d'une intégrale stochastique est toujours nulle :

$$E \left[\int_{\alpha}^{\beta} f(t) dW(t) \right] = 0 \quad (3.5)$$

b. La variance d'une intégrale stochastique se calcule au moyen d'une intégrale ordinaire

$$E \left[\int_{\alpha}^{\beta} f(t) dW(t) \right]^2 = E \left[\int_{\alpha}^{\beta} f^2(t) dt \right] \quad (3.6)$$

Nous démontrons par la suite cette dernière propriété puisque la première est immédiate, étant donné la définition d'une intégrale stochastique à partir de laquelle nous obtenons :

$$E \left[\left[\int_{\alpha}^{\beta} f(t) dW(t) \right]^2 \right] = E \left[\sum_{k=1}^{n-1} f(t_k) [W(t_{k+1}) - W(t_k)] \right]^2$$

Le processus $W(t)$ étant à accroissement indépendant, les doubles produits sont d'espérance nulle. On obtient :

$$E \left[\left[\int_{\alpha}^{\beta} f(t) dW(t) \right]^2 \right] = E \left[\sum_{k=1}^{n-1} f^2(t_k) [W(t_{k+1}) - W(t_k)]^2 \right]$$

Lorsque $f(t)$ est indépendante du mouvement brownien standard W_t , l'espérance du produit est égale au produit des espérance, et par conséquent

$$\begin{aligned} E \left[\left[\int_{\alpha}^{\beta} f(t) dW(t) \right]^2 \right] &= \sum_{k=1}^{n-1} E[f^2(t_k)] E [W(t_{k+1}) - W(t_k)]^2 \\ &= \sum_{k=1}^{n-1} E[f^2(t_k)] [t_{k+1} - t_k] = E \left[\int_{\alpha}^{\beta} f^2(t) dt \right] \quad (\text{Integrale de Riemann}) \end{aligned}$$

1.4.4 Intégrale stochastique de fonctions appartenant à L^2

Si la fonction $f(t, \omega)$ est non en escaliers, on peut trouver pour toute fonction appartenant à L^2 une suite de fonctions en escaliers $f^n(t, \omega)$ qui converge dans L^2 vers $f(t, \omega)$:

$$\left(\int_{\alpha}^{\beta} [f^n(t) - f(t)]^2 dt \xrightarrow{n \rightarrow \infty} 0 \right)$$

Définition 1.7

On définit l'intégrale stochastique de $f \in L^2$ par :

$$\int_{\alpha}^{\beta} f(t)dW(t) = \lim_{n \rightarrow \infty} \int_{\alpha}^{\beta} f^n(t)dW(t) \quad (3.7)$$

Le calcul de la moyenne et de la variance de cette intégrale est similaire à celui d'une intégrale stochastique dans le cas d'une fonction en escaliers.

Remarque 1.3

On peut montrer que pour toute fonction continûment différentiable $f : [0, T] \rightarrow R$ on a :

$$\int_0^T f(t)dW(t) = f(T)W(T) - \int_0^T f'(t)W(t)dt \quad (3.8)$$

Considérons les fonctions en escaliers définies pour tout $t \in [t_j^n, t_{j+1}^n]$ pour $j = 0, 1, 2, \dots, n-1$

$$\begin{aligned} \int_0^T f(t)dW(t) &= \lim_{n \rightarrow \infty} \int_0^T f^n(t)dW(t) = \lim_{n \rightarrow \infty} \left[\sum_{j=1}^{n-1} f^n(t_j^n) [W(t_{j+1}^n) - W(t_j^n)] \right] \\ &= \lim_{n \rightarrow \infty} \left[f^n(T)W(T) - f^n(0)W(0) - \sum_{j=0}^{n-1} W(t_{j+1}^n) [f^n(t_{j+1}^n) - f^n(t_j^n)] \right] \\ &= f(T)W(T) - \int_0^T W(t)df(t) \text{ (convergence en moyenne quadratique)} \end{aligned}$$

Ce résultat reste toujours valide si la fonction f est aléatoire à condition qu'elle soit continûment différentiable, F_t -mesurable et ayant des incréments indépendantes.

1.4.5 Différentielle stochastique

Parallèlement à l'intégrale stochastique, on peut introduire un processus stochastique qui permet la définition formelle de la différentielle stochastique avec laquelle on peut modéliser les variations instantanées aléatoires, soit :

$$Y(t) = \int_0^t f(s)dW(s)$$

Définition 1.8

On définit la différentielle stochastique $dF(t) = f(t)dW(t)$, lorsque l'on peut écrire :

$$\int_0^t f(s)dW(s) = F(t) \quad (3.9)$$

Etant donné la nature du processus de Wiener qui est un processus à variation non bornée, l'intégrale stochastique n'obéit pas aux règles de calcul différentiel classique et donc il est nécessaire d'introduire le calcul d'Itô.

1.5 Calcul d'Itô

Le calcul d'Itô est basé sur une formule dit «Formule d'Itô». Elle permet de différentier une fonction deux fois continûment différentiables comme on va le voir par la suite.

1.5.1 Formule d'Itô

Considérons un processus $X(t)$ (X est un processus stochastique et non une variable déterministe) vérifiant :

$$dX(t) = e(t)dt + f(t)dW(t), \quad X_0 = x_0 \quad (4.1)$$

Par définition de la différentielle stochastique, $X(t)$ peut s'écrire :

$$X(t) = X(0) + \int_0^t e(s)ds + \int_0^t f(s)dW(s) \quad (4.2)$$

La différentielle stochastique permet aussi la différentiation de fonctions de processus, soit $U(t, X_t) = Y_t$ grâce à la formule d'Itô.

Théorème 1.1

Soit $Y_t = U(t, X_t)$ un processus stochastique où $U : [0, T] \times \mathbb{R} \rightarrow \mathbb{R}$ est une fonction deux fois continûment dérivable et soient U'_x, U'_t, U''_{xx} ses dérivées partielles.

Le processus X_t satisfait l'équation :

$$X_t - X_s = \int_s^t e(u)du + \int_s^t f(u)dW_u(u) \quad (4.3)$$

Alors

$$Y_t - Y_s = \int_s^t (U'_t(u, X_u) + e(u)U'_x(u, X_u) + \frac{1}{2}f^2(u)U''_{xx}(u, X_u)) du + \int_s^t f(u)U'_x(u, X_u)dW_u(u) \quad (4.4)$$

Avec une probabilité 1 pour tout $0 \leq s \leq t \leq T$

L'équation (4.4) peut être mise sous la forme différentielle :

$$dY(t, X(t)) = \left[U'_t(t, X_t) + e(t)U'_x(t, X_t) + \frac{1}{2}f_t^2 U''_{xx}(t, X_t) \right] dt + f_t U'_x(t, X_t) dW_t(t)$$

pour la démonstration de cette formule, voir [10].

Remarque 1.4

Dans la théorie classique, la différentielle totale de la fonction $Y_t = U(t, X_t)$ est

$$\begin{aligned} dY(t) &= U'_x(t, X_t)dx + U'_t(t, X_t)dt \quad \text{or} \quad dX_t = e_t dt + f_t dW_t \\ \implies dY(t) &= (U'_t(t, X_t) + e_t U'_x(t, X_t))dt + f_t U'_x(t, X_t) dW_t \end{aligned}$$

En comparant cette dernière expression à celle d'Itô, on déduit qu'il y'a un terme supplémentaire dit «terme correcteur» égal à $\frac{1}{2}U''_{xx} f^2 dt$. Ce terme permet de calculer l'écart-type du mouvement brownien qui est de l'ordre de la racine carré du temps. Pour obtenir une variation de l'ordre du temps, il faut passer aux terme du 2^{ème} ordre.

La formule d'Itô ne suffit pas pour résoudre tous les problèmes réels. Elle permet parfois de les mettre en équation, rarement d'en trouver la solution explicite. La résolution des équations différentielles stochastiques est bien souvent impossible. Cependant, on est amené à utiliser des méthodes numériques afin de déterminer des solutions approchées. Pour cela, il faut à priori savoir si l'EDS possède une solution (propriété d'existence de la solution) et vérifiant de préférence la propriété d'unicité, sachant bien sûr sa valeur initial. Il y'a lieu de noter que sous certaines hypothèses que nous exposerons dans l'étude sur l'existence et l'unicité des solutions des EDS, que la solution d'une équation différentielle stochastique est un processus de diffusion auquel nous donnerons un petit aperçu dans la prochaine section.

1.6 Processus de Markov et processus de diffusion

1.6.1 Processus de Markov

Un processus de Markov est un processus aléatoire où son évolution future ne dépend pas de son passé, mais uniquement de l'état présent. lorsque le processus est discret, on parle des chaînes de markov. Autrement dit, soit une séquence quelconque d'instants $0 < t_1 < t_2 < \dots < t_n < t_{n+1}$ auxquels un processus prend les valeurs arbitraires $x_1, x_2, \dots, x_n \in \mathbb{R}$.

Ce processus est markovien si la probabilité conditionnelle au temps t_{n+1} vérifie :

$$P[X(t_{n+1}) \in B / X(t_1) = x_1, X(t_2) = x_2, \dots, X(t_n) = x_n] = P[X(t_{n+1}) \in B / X(t_n) = x_n]$$

Où B est un sous ensemble de $B_{\mathbb{R}}$ ($B_{\mathbb{R}}$ est une σ - algèbre).

Par extention, on parle de processus de Markov du $n^{i\grave{e}me}$ ordre l'orsque l'état futur du processus dépend des n états précédents.

propriétés

$$-E[X_{n+1} / X_1 = x_1, X_2 = x_2, \dots, X_n = x_n] = E[X_{n+1} / X_n = x_n]$$

$$-p[X_1 = x_1, \dots, X_n = x_n, X_{n+1} = x_{n+1}] = p(X_{n+1} = x_{n+1} / X_n = x_n) \dots p(X_2 = x_2 / X_1 = x_1) \cdot p(X_1 = x_1)$$

Cette dernière propriété montre que la statistique du processus est déterminée par les probabilités de transition et la densité de probabilité initiale.

Généralement, les probabilités de transition sont données par :

$$P(s, x; t, B) = P(X(t) \in B / X(s) = x) \quad \text{où} \quad s < t \quad B \in B_{\mathbb{R}}$$

Pour des valeurs fixes (s, x) et une valeur non fixe t , $P(s, x; t, \cdot)$ est une fonction de probabilité (mesure) dans la σ -algèbre. Sous ces condition $p(s, x; t, \cdot)$ est dit densité de transition avec

$$P(s, x; t, B) = \int_B p(s, x; t, y) dy$$

On dit que le processus de Markov à temps continu est homogène si la densité de transition $p(s, x; t, y)$ dépend seulement de la différence $(t - s)$ et non pas des instants s et t . Par

simplification, on peut écrire :

$$\begin{aligned} p(s, x; t, y) &= p(0, x; t - s, y) = p(x; t - s, y) \\ P(s, x; t, B) &= P(0, x; t - s, B) = P(x; t - s, B), \quad B \in B_{\mathbb{R}} \end{aligned}$$

Cependant, les densités de transition ne sont toutefois pas arbitraires. Elles doivent satisfaire l'équation de Chapman Kolmogorov [3], défini comme suit :

$$p(s, x; t, y) = \int_{-\infty}^{+\infty} p(s, x; \tau, z) \cdot p(\tau, z; t, y) dz \quad (5.1) \quad \text{pour tout } s \leq \tau \leq t \text{ et } x, y \in \mathbb{R}$$

Pour les probabilités de transition, ces équations prennent la forme

$$P(s, x, t, B) = \int_{-\infty}^{+\infty} P(\tau, z; t, B) \cdot P(s, x; \tau, dz), \quad B \in B_{\mathbb{R}}$$

L'intégrale ci-dessus est une intégrale de Riemann Stieltjes.

Cas particulier (processus de Wiener)

Le processus de Wiener est un cas particulier des processus de Markov homogène où la densité de transition est donnée par :

$$p(s, x; t, y) = \frac{1}{\sqrt{2\Pi(t-s)}} \exp\left(-\frac{(y-x)^2}{2(t-s)}\right)$$

Nous rappelons que le processus de Wiener est un processus de Markov à accroissements indépendants vérifiant :

- $W(0) = 0$ avec probabilité 1.
- $E(W(t)) = 0$
- $Var(W(t) - W(s)) = t - s$ pour tout $0 \leq s \leq t$

La densité de transition du processus de Wiener satisfait l'équation de Chapman Kolmogorov (5.1). Dés lors

$$\int_{-\infty}^{+\infty} p(s, x; \tau, z) p(\tau, z; t, y) dz = \int_{-\infty}^{+\infty} \frac{1}{2\Pi\sqrt{(\tau-s)(t-\tau)}} \exp\left(-\frac{1}{2}\left(\frac{(z-x)^2}{(\tau-s)} + \frac{(y-z)^2}{(t-\tau)}\right)\right) dz$$

En posant $\mu = \left(z - \frac{x(t-\tau)+y(\tau-s)}{t-s} \right) \sqrt{\frac{t-s}{(\tau-s)(t-\tau)}}$, On a :

$$\begin{aligned} \int_{-\infty}^{+\infty} p(s, x; \tau, z) p(\tau, z; t, y) dz &= \frac{1}{\sqrt{2\Pi}(t-s)} \exp\left(-\frac{(y-x)^2}{2(t-s)}\right) \int_{-\infty}^{+\infty} \frac{1}{\sqrt{2\Pi}} \exp\left(-\frac{1}{2}\mu^2\right) d\mu \\ &= p(s, x; t, y).1 \end{aligned}$$

La densité de transition du processus de Wiener est évidemment une fonction de lissage (Smooth Function) pour $t > s$ et par conséquent les dérivées partielles de cette densité satisferont les équations différentielles partielles :

$$\frac{dp}{dt} - \frac{1}{2} \frac{d^2p}{dy^2} = 0 \quad (s, x) \text{ fixé (Avancé)} \quad (5.2)$$

$$\frac{dp}{ds} + \frac{1}{2} \frac{d^2p}{dx^2} = 0 \quad (t, y) \text{ fixé (Retardé)} \quad (5.3)$$

1.6.2 Processus de diffusion

Description générale

Un processus stochastique à temps continu $(X_t)_{t \geq 0}$ est dit de diffusion s'il est markovien et à trajectoires presque sûrement continues.

Plusieurs phénomènes physiques, biologiques, économiques et sociaux ont été modélisés par un processus de diffusion, citons par exemple : les mouvements des molécules dans les particules soumises à des interactions, systèmes de communication avec bruit et la croissance de la population.

Probabilités de premier passage, temps moyen d'absorption, distributions de temps d'occupation et les distributions stationnaires sont des fonctions qui peuvent être calculées par un processus de diffusion unidimensionnel, tandis que le processus de diffusion multidimensionnel est utilisé dans la formulation, le calcul et la résolution des équations différentielles de deuxième degré.

Considérons un processus de diffusion $\{X_t, t \geq 0\}$ où l'espace des états est un intervalle I (I peut être fermé, ouvert, fini ou infini). Pour tout point z dans I , on définit T_z comme

la variable aléatoire représentant l'instant de premier passage de processus par la valeur z . Dans le cas où z n'est jamais atteinte, on met par convention $T_z = \infty$. Le processus est dit régulier si :

$$P\{T_z < \infty / X_0 = x\} > 0, \forall x \in I$$

Processus de diffusion unidimensionnel

Tous les processus de diffusion satisfont les conditions suivantes $\forall s \geq 0$:

$$\lim_{t \rightarrow s} \frac{1}{(t-s)} P\{|X_t - x| > \varepsilon / X_s = x\} = 0, \forall x \in I, \forall \varepsilon > 0$$

$$\lim_{t \rightarrow s} \frac{1}{(t-s)} \int_{|y-x| > \varepsilon} p(s, x; t, y) dy = 0, \forall x \in I, \forall \varepsilon > 0 \quad (5.4)$$

Cette relation signifie que la probabilité d'avoir des grandes transitions (supérieur à ε) pour un intervalle de temps suffisamment petit est nulle (empêche le processus de diffusion d'avoir des sauts instantanés). Ce fait est du à la continuité des trajectoires de processus de diffusion. Tous les processus de diffusion existants possèdent deux conditions de base décrivant la moyenne et la variance des transitions infinitésimales respectivement :

$$\lim_{t \rightarrow s} \frac{1}{(t-s)} \int_{|y-x| \leq \varepsilon} (y-x) p(s, x; t, y) dy = a(s, x), \forall x \in I, \forall \varepsilon > 0 \quad (5.5)$$

$$\lim_{t \rightarrow s} \frac{1}{(t-s)} \int_{|y-x| \leq \varepsilon} (y-x)^2 p(s, x; t, y) dy = b^2(s, x), \forall x \in I, \forall \varepsilon > 0 \quad (5.6)$$

La fonction $a(s, x)$ est appelée coefficient de changement ou de tendance (drift) du processus de diffusion et $b(s, x)$ son coefficient de diffusion à l'instant s et à l'état x

L'expression (5.5) implique que $a(s, x)$ peut s'écrire :

$$a(s, x) = \lim_{t \rightarrow s} \frac{1}{(t-s)} E[X_t - X_s / X_s = x] \quad (5.7)$$

$a(s, x)$ est le taux de changement instantané de la moyenne du processus sachant $X_s = x$ (la moyenne infinitésimales)

On définit le taux de changement instantané des variances du processus sachant $X_s = x$ (variance infinitésimale) par :

$$b^2(s, x) = \lim_{t \rightarrow s} \frac{1}{(t-s)} E[(X_t - X_s)^2 / X_s = x] \quad (5.8)$$

Donc on peut définir un processus de diffusion par un processus Markovien $(X_t)_{t \geq 0}$ vérifiant les relations (5.4), (5.5) et (5.6) pour des fonctions bien définies a et b

Remarque 1.5

Etant donné que les densités de transition $p(s, x; t, y)$ ne doivent pas nécessairement exister, les expressions (5.4), (5.5) et (5.6) seront exprimées en terme de probabilités de transition $P(s, x; t, y)$

Processus de diffusion multidimensionnel

soit $X = \{X_t, t \geq 0\}$ le vecteur des processus stochastiques définis sur un espace de probabilités (Ω, A, P) , à valeurs dans \mathbb{R}^d où X_t est σ^d -mesurable pour tout $t \geq 0$ (σ^d est la σ -algèbre des sous ensembles de Lebesgue de \mathbb{R}^d).

Les probabilités de transition du processus X sont définies par :

$$P(s, x; t, A) = P(X_t \in A / X_s = x) \quad \forall 0 \leq s < t, x \in \mathbb{R}^d \text{ et } A \in \sigma^d$$

Le processus aléatoire X est un processus de diffusion si pour tout $\varepsilon > 0$, $s \geq 0$ et $x \in \mathbb{R}^d$, les conditions ci-dessous sont remplies :

$$\lim_{t \rightarrow s} \frac{1}{(t-s)} \int_{|y-x| > \varepsilon} p(s, x; t, y) dy = 0 \quad (5.9)$$

$$\lim_{t \rightarrow s} \frac{1}{(t-s)} \int_{|y-x| \leq \varepsilon} (y-x) p(s, x; t, y) dy = a(s, x) \quad (5.10)$$

$$\lim_{t \rightarrow s} \frac{1}{(t-s)} \int_{|y-x| \leq \varepsilon} (y-x)(y-x)^T p(s, x; t, y) dy = B(s, x) B^T(s, x) \quad (5.11)$$

Où a est le vecteur dérive de dimension d et $D = BB^T$ est la matrice des coefficients de diffusion ($d \times d$) qui est symétrique et définie positive.

Le terme général de vecteur dérive a est défini par :

Pour $i, j = 1, \dots, d$

$$\lim_{t \rightarrow s} \frac{1}{(t-s)} E[X_t^i - X_s^i / X_s^i = x = (x^1, x^2, \dots, x^d)] = a^i(s, x)$$

Le terme général de la matrice de diffusion D est défini par :

$$d^{i,j}(s, x) = \lim_{t \rightarrow s} \frac{1}{(t-s)} E[(X_t^i - X_s^i)(X_t^j - X_s^j) / X_s = x] \quad \forall i \neq j$$

En outre, si le vecteur dérive a et la matrice de diffusion D sont des fonctions régulières alors les probabilités de transition $P(s, x; t, A)$ ont des fonctions de densités $p(s, x; t, y)$ qui satisfont l'équation de Kolmogorov avancée :

$$\frac{dp}{dt} + \sum_{i=1}^d \frac{d}{dy_i} \{a^i(t, y)p(s, x; t, y)\} - \frac{1}{2} \sum_{i,j=1}^d \frac{d^2}{dy_i dy_j} \{d^{i,j}(t, y)p(s, x; t, y)\} = 0 \quad (s, x) \text{ fixé} \quad (5.12)$$

Avec la condition initial $\lim_{t \rightarrow s} p(s, x; t, y) = p(0, x; t-s, y) = \delta(x-y)$ où δ est la fonction Delta-Dirac définie sur \mathbb{R}^d par : $\delta(s) = 0 \quad \forall s \neq 0$ tel que $\int_{-\infty}^{+\infty} f(s)\delta(s)ds = f(0)$ pour toute fonction continue f à l'instant $s = 0$. La densité p est donc une solution fondamentale de l'équation différentielle parabolique (5.12).

On définit l'opérateur elliptique L par :

$$L\mu(s, x) = \sum_{i=1}^d a^i(s, x) \frac{d\mu}{dx_i}(s, x) + \frac{1}{2} \sum_{i,j=1}^d d^{i,j}(s, x) \frac{d^2}{dx_i dx_j} \mu(s, x)$$

L'équation de Kolmogorov retardé exprimée en fonction de cet opérateur est donnée par :

$$\frac{d\mu}{ds} + L\mu = 0, \text{ vérifiant } \mu(s, x) = p(s, x; t, y) \text{ pour } t \text{ et } y \text{ fixés.}$$

Cette équation admet aussi la solution :

$$\mu(s, x) = E(f(X_t) / X_s = x) = \int_{\mathbb{R}^d} f(y)p(s, x; t, y)dy,$$

appelée formule de Kolmogorov, ce qui correspond à la condition $\lim_{t \rightarrow s} \mu(s, x) = f(x)$ pour toute fonction continue $f : \mathbb{R}^d \rightarrow \mathbb{R}$.

Pour $x \in \varphi \subset \mathbb{R}^d$, où φ est un domaine borné, l'équation : $\frac{d\mu}{ds} + L\mu = -1$ avec $\mu(s, x) = 0 \forall x \in \varphi$, satisfait l'équation :

$$\mu(s, x) = E(\tau_{s,x}) - s$$

où $\tau_{s,x} = \text{Inf}\{t \geq s : X_t(\omega) \in \varphi \text{ avec } X_s(\omega) = x\}$ est l'instant du premier passage du processus au domaine φ sachant l'état initial (s, x) .

Cas particulier (processus de Wiener)

Le processus de Wiener est un processus de diffusion avec $a(s, x) = 0$ et $b(s, x) = 1$.

$$a(s, x) = \lim_{t \rightarrow s} \frac{1}{(t-s)} E[W_t - W_s / W_s = x] = \lim_{t \rightarrow s} \frac{1}{(t-s)} E[W_t - W_s] = \lim_{t \rightarrow s} \frac{1}{(t-s)} 0 = 0.$$

$$b(s, x) = \lim_{t \rightarrow s} \frac{1}{(t-s)} E[(W_t - W_s)^2 / W_s = x] = \lim_{t \rightarrow s} \frac{1}{(t-s)} E[(W_t - W_s)^2] = \lim_{t \rightarrow s} \frac{1}{(t-s)} (t-s) = 1$$

Lorsque a et b sont des fonction lissées (Smooth function), la densité de transition $p(s, x, t, y)$ d'un processus brownien standard satisfait les équations différentielles partielles déjà établis en (5.2) et (5.3). Il s'agit des équations dit "Kolmogorov Forward Equation" qui caractérisent l'évolution future du processus sachant l'état final (t, y) et celles qui expriment son évolution passé, sachant l'état initial (s, x) , sont appelées "Kolmogorov Backward Equation". Ces équations sont données par :

$$\frac{dp}{dt} + \frac{d}{dy} [a(t, y).p] - \frac{1}{2} \frac{d^2}{dy^2} [b^2(t, y).p] = 0 \quad (s, x) \text{ fixé (5.13)}$$

$$\frac{dp}{ds} + a(s, x) \frac{dp}{dx} + \frac{1}{2} b^2(s, x) \frac{d^2p}{dx^2} = 0 \quad (t, y) \text{ fixé (5.14)}$$

Les deux équations (5.13) et (5.14) sont connues par Fokker-Plank Equation.

Continuité d'un processus de diffusion

La relation (5.4) permet au chemin aléatoire du processus de diffusion d'avoir un bon comportement. Effectivement, il y'a des fonctions continues du temps bien quelles ne soient pas nécessairement différentiables comme on va le montrer par la suite dans le cas d'un processus de Wiener.

On va définir la continuité du processus stochastique X_t à t de différentes manières :

1. Continuité avec probabilité 1 (Continuité presque sûre)

$$\lim_{s \rightarrow t} P(\{\omega \in \Omega : |X(s, \omega) - X(t, \omega)| = 0\}) = 1$$

2. Continuité en moyenne quadratique

$$E(X_t^2) < \infty \quad \text{et} \quad \lim_{s \rightarrow t} E(|X_s - X_t|^2) = 0$$

3. Continuité en probabilité (Continuité stochastique)

$$\lim_{s \rightarrow t} P(\{\omega \in \Omega : |X(s, \omega) - X(t, \omega)| \geq \varepsilon\}) = 0 \quad \text{pour tout} \quad \varepsilon > 0$$

4. Continuité en distribution

$$\lim_{s \rightarrow t} F_s(x) = F_t(x)$$

Il existe une relation entre ces continuités, on a **1** et **2** \Rightarrow **3** \Rightarrow **4** et en plus **3** peut être mis sous la forme d'une métrique :

$$\lim_{s \rightarrow t} E \left(\frac{|X_s - X_t|}{1 + |X_s - X_t|} \right) = 0$$

Cette dernière limite provient du fait que

$$P(\{\omega \in \Omega : |X(s, \omega) - X(t, \omega)| \geq \varepsilon\}) = P(\{\omega \in \Omega : |X(s, \omega) - X(t, \omega)|^2 \geq \varepsilon^2\})$$

Or d'après l'inégalité de Chebyshev : $P(\{\omega; |X(\omega)|^2 \geq a\}) \leq \frac{1}{a} E(X^2) \quad \forall a > 0$, on trouve

$$P(\{\omega \in \Omega : |X(s, \omega) - X(t, \omega)|^2 \geq \varepsilon^2\}) \leq \frac{1}{\varepsilon^2} E(|X_s - X_t|^2)$$

Le concept de continuité d'un processus de diffusion entraîne la continuité du chemin aléatoire.

Il existe un critère découvert par Kolmogorov impliquant la continuité de la trajectoire du processus stochastique à temps continu $X = \{X_t, t \in T\}$, pour des constantes positives α, β, c et h tel que :

$$E(|X_t - X_s|^\alpha) \leq c|t - s|^{1+\beta}$$

pour tout $s, t \in T$ et $|t - s| \leq h$. On peut montrer que le processus de Wiener standard vérifie cette égalité avec $\alpha = 4, \beta = 1$ et $c = 3$.

Soit le processus de Wiener $W = \{W_t, t \geq 0\}$. l'accroissement du processus de Wiener $W_t - W_s \approx N(0, |t - s|)$. On admet le résultat suivant :

pour $X \approx N(\mu, \sigma^2)$ et $n = 0, 1, 2, \dots$,

$$E((X - \mu)^n) = \left\{ \begin{array}{ll} 0 & \text{si } n = 2k + 1 \\ 1.3.5 \dots (2k - 1)\sigma^{2k} & \text{si } n = 2k \end{array} \right\}$$

Dans notre cas $\sigma^2 = |t - s|$ et $k = 2 \Rightarrow E(|W_t - W_s|^4) = 1.3.\sigma^4 = 3|t - s|^2$ et par conséquent, le processus de Wiener est presque sûrement à trajectoires continues.

Lorsque le processus de Wiener est de dimension $d, W_t = (W_t^1, W_t^2, \dots, W_t^d)$, alors

$$E(|W_t - W_s|^4) = (d^2 + 2d)|t - s|^2$$

Propriétés d'un processus de diffusion

Les coefficients infinitésimaux

Généralement, les fonctions $a(s, x)$ et $b^2(s, x)$ sont continues pour tout s et x , et pour un processus régulier la fonction $b^2(s, x)$ est positive pour x dans I . Si le processus est homogène les fonctions $a(s, x) = a(x)$ et $b^2(s, x) = b^2(x)$ sont indépendantes de s . La classification de toutes les diffusions possibles par la prescription des coefficients infinitésimaux $a(s, x)$ et $b^2(s, x)$ dépend de la définition de mouvement qu'ils représentent. En plus des relations infinitésimal (5.7) et (5.8), la relation des moments infinitésimaux d'ordre supérieur est

vraie :

$$\lim_{t \rightarrow s} \frac{1}{(t-s)} E(|X_t - X_s|^r / X_s = x) = 0, \quad r = 3, 4, \dots \quad (5.15)$$

Transformation élémentaire d'un processus de diffusion

Théorème 1.2

Soit $\{X_t, t \geq 0\}$ un processus de diffusion régulier où l'espace des états est un intervalle I dont les extrémités sont les points l et r , on suppose que $a(x)$ et $b^2(x)$ sont les coefficients infinitésimaux de X_t . Soit g une fonction strictement monotone sur I avec une dérivée seconde continue $g''(x)$ pour $l < x < r$. D'où le processus $Y_t = g[X_t]$ est un processus de diffusion régulier sur l'intervalle dont les extrémités sont $g(l)$ et $g(r)$, et Y_t a comme coefficients infinitésimaux

$$\begin{aligned} a_Y(y) &= \frac{1}{2} b^2(x) g'(x) + a(x) g'(x) \\ b_Y^2(y) &= b^2(x) [g'(x)]^2 \\ \text{où } y &= g(x) \end{aligned}$$

Preuve

On considère le cas où la fonction g est strictement croissante car la preuve est similaire dans le cas où elle est strictement décroissante.

Comme g est deux fois continûment différentiable, la formule de Taylor [6.7] donne :

$$g(x + \Delta x) = g(x) + \Delta x g'(x) + \frac{1}{2} (\Delta x)^2 g''(x) + \frac{1}{2} (\Delta x)^2 [g''(\xi) - g''(x)] \text{ où } x \leq \xi \leq x + \Delta x$$

En remplaçant x par X_t et Δx par $\Delta X = X_{t+h} - X_t$, on aura :

$$g(X_{t+h}) = g(X_t) + (\Delta X) g'(X_t) + \frac{1}{2} (\Delta X)^2 g''(X_t) + \frac{1}{2} (\Delta X)^2 [g''(\xi(\omega)) - g''(X_t)]$$

Où $\xi(\omega)$ lie entre X_t et X_{t+h} et ω signifie la réalisation particulière de processus qui rend $\xi(\omega)$. Comme $X_t = x$ donc $g(x) = g(X_t) = T(y) = y$ et l'équation précédente devient :

$$Y_{t+h} - Y_t = \Delta X g'(X_t) + \frac{1}{2} (\Delta X)^2 g''(X_t) + \frac{1}{2} (\Delta X)^2 [g''(\xi(\omega)) - g''(X_t)] \quad (5.16)$$

En divisant sur h et en passant à la moyenne puis à la limite, on trouve :

$$\lim_{h \rightarrow 0} \frac{1}{h} E[Y_{t+h} - Y_t | Y_t = y] = a(x)g'(x) + \frac{1}{2}b^2(x)g''(x) + \frac{1}{2} \lim_{h \rightarrow 0} \frac{1}{h} E[(\Delta X)^2 \{g''(\xi(\omega)) - g''(X_t)\}]$$

La continuité de g'' assure la convergence de $g''(\xi(\omega))$ à $g''(X_t)$, donc :

$$\lim_{h \rightarrow 0} \frac{1}{h} E[(\Delta X)^2 \{g''(\xi(\omega)) - g''(X_t)\}] = 0$$

ça signifie que la moyenne infinitésimale de Y_t vaut :

$$a_Y(y) = \lim_{h \rightarrow 0} \frac{1}{h} E[Y_{t+h} - Y_t | Y_t = y] = a(x)g'(x) + \frac{1}{2}b^2(x)g''(x)$$

En calculant le carré de (5.16), on trouve :

$$[Y_{t+h} - Y_t]^2 = (\Delta X)^2 [g'(X_t)]^2 + R_h$$

Où R_h est le reste et qui contient $(\Delta X)^3$ comme terme d'ordre maximum.

D'après (5.14) $\lim_{h \rightarrow 0} \frac{1}{h} E[|\Delta X|^r / X_t = x] = 0$ pour tout $r \geq 3$. Par la suite la variance infinitésimale de Y_t est :

$$b_Y^2(y) = \lim_{h \rightarrow 0} \frac{1}{h} E[\{Y_{t+h} - Y_t\}^2 / Y_t = y] = b^2(x)[g'(x)]^2$$

La condition (1) de $Y_t = g[X_t]$ est déjà vérifiée, donc le processus Y_t est un processus de diffusion.

1.7 Conclusion

Le mouvement brownien joue un rôle central dans la théorie des processus aléatoires, d'une part parce que, dans de nombreux problèmes appliqués, le mouvement brownien sert à modéliser les erreurs ou les perturbations aléatoires, et d'autre part parce que le mouvement brownien ou les processus de diffusion qui en découlent permettent de construire des modèles simples sur lesquels des calculs peuvent être faits. Le calcul stochastique, ou calcul d'Itô, du nom d'un des pionniers en ce domaine, est en fait un calcul d'intégrale par rapport au mouvement brownien. Ce dernier étant une fonction qui n'est pas à variation finie, cette notion d'intégrale n'est pas usuelle et sa définition en est probabiliste. Elle permet en particulier de définir la notion d'équation différentielle stochastique qui est une équation obtenue par la perturbation aléatoire d'une équation différentielle ordinaire. Dans le chapitre qui suit nous présenterons ces équations dont ses solutions sont des processus de diffusion.

Chapitre 2

Equations différentielles stochastiques

2.1 Introduction

Ce chapitre est consacré aux équations différentielles stochastiques, qui motivèrent les premiers travaux d'Itô sur l'intégrale stochastique, nous traitons les différents types des équations différentielles stochastiques basons essentiellement sur le calcul d'Itô. Après, nous traitons en détail le cas lipchitzien, dans lequel des résultats forts d'existence et d'unicité des solutions peuvent être obtenus. Toujours dans ce cadre lipchitzien, nous montrons que la solution d'une équation différentielle stochastique est généralement un processus de Markov et dans la majorité des cas un processus de diffusion.

De manière informelle, on appelle équation différentielle stochastique une équation différentielle ordinaire perturbée par un terme stochastique (bruit aléatoire) le but de ces équations est de fournir un modèle mathématique. Partons d'une équation différentielle ordinaire de la forme

$$X'(t) = a(X(t))$$

soit encore sous forme différentielle,

$$dX_t = a(X_t)dt$$

Si l'on prend en compte les perturbations aléatoires, on ajoute un terme de bruit, qui sera de la forme $b dW_t$, où W désigne un mouvement brownien et b est pour l'instant une constante qui correspond à l'intensité du bruit. On arrive à une équation différentielle "stochastique"

de la forme

$$dX_t = a(X_t)dt + bW_t$$

ou encore sous forme intégrale, la seule qui ait un sens mathématique.

$$X_t = X_0 + \int_0^t a(X_s)ds + bW_t$$

On généralise cette équation en autorisant b à dépendre de l'état à l'instant t :

$$dX_t = a(X_t)dt + b(X_t)dW_t$$

soit sous forme intégrale,

$$X_t = X_0 + \int_0^t a(X_s)ds + \int_0^t b(X_s)dW_s$$

Remarquons que le sens donné à cette équation dépend de la théorie de l'intégrale stochastique développée dans le chapitre précédent.

2.2 Equations différentielles de type linéaire

Comme avec des équations ordinaires linéaires, la solution générale d'une équation stochastique linéaire peut être trouvée explicitement. La méthode de solution implique également un facteur d'intégration ou une solution fondamentale d'une équation homogène associée.

2.2.1 Résolution des EDS de type linéaire

La méthode de résolution d'une équation différentielle stochastique (EDS) nécessite l'introduction de la solution correspondante à l'équation différentielle homogène. Nous considérons ici les équations de type scalaire ayant la forme générale

$$dX_t = [a_1(t)X_t + a_2(t)] dt + [b_1(t)X_t + b_2(t)] dW_t \quad (1.1)$$

Où a_1, a_2, b_1 et b_2 sont des fonctions de temps, mesurables au sens de Lebesgue et bornées sur $[0, T]$.

On définit le processus X_t pour tout $t_0 \leq t \leq T$, où X_{t_0} est F_{t_0} mesurable, ainsi que le processus de Wiener W_t (W_t est F_t mesurable).

Lorsque $a_2(t) \equiv 0$ et $b_2(t) \equiv 0$, l'expression **(1.1)** se réduit à l'équation différentielle stochastique homogène et linéaire :

$$dX_t = a_1(t)X_t dt + b_1(t)X_t dW_t \quad \text{(1.2)}$$

Quand $b_1(t) \equiv 0$, le bruit apparaît comme un terme additif et l'équation **(1.1)** devient :

$$dX_t = [a_1(t)X_t + a_2(t)] dt + b_2(t)dW_t \quad \text{(1.3)}$$

Dans ce cas là l'EDS est dit linéaire dans un sens limité (Linear in the narrow sense) et l'équation homogène obtenue à partir de **(1.3)** est une équation différentielle ordinaire

$$\frac{dX_t}{dt} = a_1(t)X_t \quad \text{(1.4)}$$

où la solution fondamentale est donnée par : $\phi_{t,t_0} = \exp\left(\int_{t_0}^t a_1(s)ds\right)$

En appliquant la formule d'Itô **(I.4.4)** à la transformation $U(t, x) = \phi_{t,t_0}^{-1}x$ on obtient :

$$d(\phi_{t,t_0}^{-1}X_t) = \left[\frac{d\phi_{t,t_0}^{-1}}{dt} X_t + (a_1(t)X_t + a_2(t)) \phi_{t,t_0}^{-1} \right] dt + b_2(t)\phi_{t,t_0}^{-1} dW_t$$

Or

$$\frac{d\phi_{t,t_0}^{-1}}{dt} = -\phi_{t,t_0}^{-1}a_1(t) \Rightarrow d\phi_{t,t_0}^{-1}X_t = (a_2(t)\phi_{t,t_0}^{-1}) dt + b_2(t)\phi_{t,t_0}^{-1} dW_t$$

Cette dernière expression peut se mettre sous la forme intégrale :

$$\phi_{t,t_0}^{-1}X_t = \phi_{t_0,t_0}^{-1}X_{t_0} + \int_{t_0}^t (a_2(s)\phi_{s,t_0}^{-1})ds + \int_{t_0}^t b_2(s)\phi_{s,t_0}^{-1} dW_s$$

Sous la condition $\phi_{t_0, t_0} = 1$, la solution de **(1.3)** est

$$X_t = \phi_{t, t_0} \left(X_{t_0} + \int_{t_0}^t (a_2(s) \phi_{s, t_0}^{-1}) ds + \int_{t_0}^t b_2(s) \phi_{s, t_0}^{-1} dW_s \right) \quad (1.5)$$

où

$$\phi_{t, t_0} = \exp \left(\int_{t_0}^t a_1(s) ds \right) \quad (1.6)$$

Le processus X_t **(1.5)** est une combinaison linéaire de variables aléatoires gaussiennes car $\int_{t_0}^t f(s) dW_s$ est gaussienne pour toute fonction déterministe f ; en plus, X_{t_0} est supposé déterministe (constante) ou une variable gaussienne. Dès lors, X_t est un processus gaussien où la moyenne et le moment d'ordre deux satisfont les équations différentielles ordinaires. Toute fois, dans le cas général la solution de **(1.1)** n'est pas toujours gaussienne.

Proposition 2.1

La solution fondamentale de **(1.2)** est donné par :

$$\phi_{t, t_0} = \phi_{t_0, t_0} \exp \left[\int_{t_0}^t \left(a_1(s) - \frac{1}{2} b_1^2(s) \right) ds + \int_{t_0}^t b_1(s) dW_s \right] \quad (1.7)$$

où ϕ_{t_0, t_0} est généralement égale 1

Preuve

La solution fondamentale **(1.6)** d'une équation linéaire au sens limité satisfait l'équation différentielle ordinaire $d \ln \phi_{t, t_0} = a_1(t) dt$. En utilisant cette définition, la formule d'Itô **(1.4.4)** appliqué à la transformée de la solution fondamentale $\ln \phi_{t, t_0}$ de **(1.2)** sera :

$$\begin{aligned} d(\ln \phi_{t, t_0}) &= \left(a_1(t) \phi_{t, t_0} \phi_{t, t_0}^{-1} - \frac{1}{2} b_1^2(t) \phi_{t, t_0}^2 \phi_{t, t_0}^{-2} \right) dt + b_1(t) \phi_{t, t_0} \phi_{t, t_0}^{-1} dW_t \\ &= \left(a_1(t) - \frac{1}{2} b_1^2(t) \right) dt + b_1(t) dW_t \\ \Rightarrow \ln \phi_{t, t_0} &= \int_{t_0}^t \left(a_1(s) - \frac{1}{2} b_1^2(s) \right) ds + \int_{t_0}^t b_1(s) dW_s \end{aligned}$$

Lorsque $b_1(t) \equiv 0$, **(1.7)** se réduit à **(1.6)**.

Proposition 2.2

Le processus $\phi_{t,t_0}^{-1}X_t$ qui est fonction de la solution X_t de l'équation générale (1.1) possède une forme explicite de l'intégrale stochastique de la forme :

$$\phi_{t,t_0}^{-1}X_t = X_{t_0} + \int_{t_0}^t [a_2(s) - b_1(s)b_2(s)] \phi_{s,t_0}^{-1} ds + \int_{t_0}^t [b_2(s)\phi_{s,t_0}^{-1}] dW_s$$

Similairement, on applique la formule d'Itô à ϕ_{t,t_0}^{-1} et on obtient :

$$d\phi_{t,t_0}^{-1} = (-a_1(t) + b_1^2(t))\phi_{t,t_0}^{-1} dt - b_1(t)\phi_{t,t_0}^{-1} dW_t \quad (1.8)$$

Ce qui implique que la solution de l'équation générale (1.1) est donné par :

$$X_t = \phi_{t,t_0} \left[X_{t_0} + \int_{t_0}^t [a_2(s) - b_1(s)b_2(s)] \phi_{s,t_0}^{-1} ds + \int_{t_0}^t [b_2(s)\phi_{s,t_0}^{-1}] dW_s \right] \quad (1.9)$$

où ϕ_{t,t_0} est donné par l'expression (1.8).

Preuve

Les deux terme ϕ_{t,t_0}^{-1} et X_t possèdent des différentielles stochastiques incluant le même processus de Wiener W_t , ce qui implique que la différentielle de leur produit (voir [15]) :

$$U(X_t^1, X_t^2) = X_t^1 X_t^2 \quad \text{avec} \quad X_t^1 = \phi_{t,t_0}^{-1} \text{ et } X_t^2 = X_t \quad \text{donne :}$$

$$\begin{aligned} d\phi_{t,t_0}^{-1}X_t &= \phi_{t,t_0}^{-1}([a_1(t)X_t + a_2(t)] dt + [b_1(t)X_t + b_2(t)] dW_t) \\ &\quad + X_t((-a_1(t) + b_1^2(t))\phi_{t,t_0}^{-1} dt - b_1(t)\phi_{t,t_0}^{-1} dW_t) - b_1(t)\phi_{t,t_0}^{-1} [b_1(t)X_t + b_2(t)] dt \\ &= [a_2(t) - b_1(t)b_2(t)] \phi_{t,t_0}^{-1} dt + b_2(t)\phi_{t,t_0}^{-1} dW_t \end{aligned}$$

En intégrant sur l'intervalle de temps $[t_0, t]$ et en supposant que $\phi_{t_0,t_0} = 1$, on obtient :

$$\phi_{t,t_0}^{-1}X_t = X_{t_0} + \int_{t_0}^t [a_2(s) - b_1(s)b_2(s)] \phi_{s,t_0}^{-1} ds + \int_{t_0}^t [b_2(s)\phi_{s,t_0}^{-1}] dW_s.$$

Corollaire 2.1

Dans le cas d'une équation différentielle stochastique de type $dX_t = aX_t dt + bX_t dW_t$ on applique l'équation (1.9) avec $a_1(t) \equiv a, b_1(t) \equiv b$ et $a_2(t) \equiv b_2(t) \equiv 0$ et la solution est donné dans ce cas par :

$$X_t = X_0 \exp \left[\left(a - \frac{1}{2}b^2 \right) t + bW_t \right]$$

Ce corollaire est important lorsqu'on compare cette solution exacte du processus de diffusion à celle approchée grâce aux méthodes numériques conçues pour la résolution des EDS, présenté à la dernière section de ce chapitre

2.2.2 Solutions générales de quelques EDS de type linéaire

Nous exposons dans cette partie la solution générale de diverses équations différentielles stochastiques linéaires que le bruit, présenté par le mouvement brownien, soit additif ou multiplicatif.

Nous rappelons que la solution générale d'une équation linéaire de la forme :

$$dX_t = (a_1(t)X_t + a_2(t)) dt + (b_1(t)X_t + b_2(t)) dW_t$$

est donné par l'expression $X_t = \phi_{t,t_0} \left[X_{t_0} + \int_{t_0}^t [a_2(s) - b_1(s)b_2(s)] \phi_{s,t_0}^{-1} ds + \int_{t_0}^t [b_2(s)\phi_{s,t_0}^{-1}] dW_s \right]$

où la solution fondamentale est $\phi_{t,t_0} = \exp \left[\int_{t_0}^t (a_1(s) - \frac{1}{2}b_1^2(s)) ds + \int_{t_0}^t b_1(s) dW_s \right]$

$$\mathbf{1-} a_1(t) \equiv -\alpha, b_2(t) \equiv \sigma, a_2(t) \equiv b_1(t) \equiv 0.$$

$$X_t = e^{-\alpha t} \left(X_0 + \sigma \int_0^t e^{\alpha s} dW_s \right) \quad (\mathbf{1.10})$$

$$\mathbf{2-} a_1(t) \equiv a, a_2(t) \equiv b, b_1(t) \equiv 0, b_2(t) \equiv c$$

$$X_t = e^{at} \left(X_0 + \frac{b}{a}(1 - e^{-at}) + c \int_0^t e^{-as} dW_s \right) \quad (\mathbf{1.11})$$

$$\mathbf{3-}a_1(t) \equiv a(t), a_2(t) \equiv b(t), b_1(t) \equiv c(t), b_2(t) \equiv 0$$

$$X_t = \phi_{t,t_0} \left[X_{t_0} + \int_{t_0}^t \phi_{s,t_0}^{-1} b(s) ds + \int_{t_0}^t c(s) \phi_{s,t_0}^{-1} dW_s \right] \quad (\mathbf{1.12}) \quad \text{où } \phi_{t,t_0} = \exp \left(\int_{t_0}^t a(s) ds \right)$$

$$\mathbf{4-}a_1(t) \equiv a, b_1(t) \equiv b, a_2(t) \equiv b_2(t) \equiv 0$$

$$X_t = X_0 \exp \left[\left(a - \frac{1}{2} b^2 \right) t + bW_t \right] \quad (\mathbf{1.13})$$

$$\mathbf{5-}a_1(t) \equiv a, a_2(t) \equiv c, b_1(t) \equiv b, b_2(t) \equiv d$$

$$X_t = \phi_t \left[X_0 + (c - bd) \int_0^t \phi_s^{-1} ds + d \int_0^t \phi_s^{-1} dW_s \right] \quad (\mathbf{1.14})$$

$$\text{où } \phi_t = \exp \left[\left(a - \frac{1}{2} b^2 \right) t + bW_t \right]$$

$$\mathbf{6-}a_1(t) \equiv a(t), b_1(t) \equiv b(t), a_2(t) \equiv b_2(t) \equiv 0$$

$$X_t = X_0 \exp \left[\int_0^t \left(a(s) - \frac{1}{2} b^2(s) \right) ds + \int_0^t b(s) dW_s \right] \quad (\mathbf{1.15})$$

2.2.3 Exemple d'application d'une EDS linéaire

Equation de Langevin

Suite aux indications d'Einstein sur le mouvement brownien durant la première décennie de ce siècle, des tentatives ont été faites par Langevin et autres pour décrire le comportement dynamique d'une particule dans un fluide soumis à l'agitation thermique via les équations différentielles. Langevin a établi pour l'accélération de la particule l'équation différentielle :

$$\frac{dX_t}{dt} = -aX_t + b\varepsilon_t \quad , \quad a \text{ et } b > 0$$

Cette expression n'est autre que la somme de la force frictionnée retardée qui dépend de la vitesse de la particule et de la force de la molécule exprimée par le bruit blanc ε_t ainsi que de l'intensité b qui est indépendante de la vitesse.

L'équation de Langevin décrite précédemment peut se mettre sous la forme stochastique $dX_t = -aX_t dt + b dW_t$. Cette dernière est un cas particulier de **(1.10)** avec $a_1(t) \equiv -a$, $b_2(t) \equiv b$ et $a_2(t) \equiv b_2(t) \equiv 0$ et donc la solution générale représentant le processus d'Einstein Uhlenbeck [13] est donnée par : $X_t = e^{-at} X_0 + b e^{-at} \int_0^t e^{as} dW_s$

En appliquant l'expression **(1.3.8)** avec $f(t) = e^{at}$, l'intégrale stochastique d'Itô sera de la forme :

$$\int_0^t e^{as} dW_s = e^{at} W_t - a \int_0^t e^{as} W_s ds$$

a et b sont des paramètres inconnus à estimer.

Il existe diverses méthodes d'estimation des fonctions ou des paramètres inconnus qui interviennent dans une EDS, comme les méthodes paramétriques et non paramétriques. Notons que les méthodes paramétriques sont plus précises par rapport à la méthode des moments car elles se basent sur la méthode du maximum de vraisemblance (voir [14] et [32],[19]). Le calcul du moment d'ordre un et deux du processus de diffusion X_t , que nous allons présenter dans la prochaine section, est un résultat utilisé pour l'estimation des paramètres inconnus.

2.2.4 Espérance et moment d'ordre deux d'un processus de diffusion

Dans certains cas on ne peut trouver la forme explicite du processus de diffusion X_t ou bien l'objectif de l'étude des EDS n'est pas de simuler la trajectoire du processus mais seulement sa loi, alors on est amené à chercher l'espérance et la variance du processus de diffusion au moyen des EDS.

Proposition 2.3

L'espérance et le moment d'ordre deux du processus X_t , ayant la forme différentielle stochastique (1.1) satisfirent les équations différentielles ordinaires suivantes :

$$\frac{dm(t)}{dt} = a_1(t)m(t) + a_2(t) \quad (1.16) \quad \text{où } m(t) = E(X_t)$$

$$\frac{dP(t)}{dt} = (2a_1(t) + b_1^2(t))P(t) + 2m(t)(a_2(t) + b_1(t)b_2(t)) + b_2^2(t) \quad (1.17)$$

$$\text{où } P(t) = E(X_t^2)$$

Preuve

L'équation (1.16) est obtenue en introduisant l'espérance à la forme intégrale de (1.1) et en utilisant la propriété (I.3.5) d'une intégrale stochastique (la moyenne d'une intégrale stochastique est nulle).

L'équation (1.17) résulte de l'application de la formule d'Itô (I.4.4) afin d'obtenir une équation différentielle stochastique du processus X_t^2 .

$$\begin{aligned} d(X_t^2) &= \left[(a_1 X_t + a_2) 2X_t + \frac{1}{2} (b_1 X_t + b_2)^2 2 \right] dt + (b_1 X_t + b_2) 2X_t dW_t \\ &= (2a_1 X_t^2 + 2a_2 X_t + b_1^2 X_t^2 + 2b_1 b_2 X_t + b_2^2) dt + (2b_1 X_t^2 + 2b_2 X_t) dW_t \end{aligned}$$

En passant à l'intégrale, puis en introduisant l'espérance à l'expression ci-dessus, le dernier terme s'annule.

Comme $P(t)dt = \frac{dP(t)}{dt}dt = \frac{dE(X_t^2)}{dt}dt = E(dX_t^2)$, alors

$$\begin{aligned} E(dX_t^2) &= [2a_1 E(X_t^2) + 2a_2 E(X_t) + b_1^2 E(X_t^2) + 2b_1 b_2 E(X_t) + b_2^2] dt + 0 \\ \Rightarrow P(t) &= (2a_1(t) + b_1^2(t))P(t) + 2(a_2(t) + b_1 b_2)m(t) + b_2^2(t) \quad \text{où } m(t) = E(X_t) \end{aligned}$$

Lorsque l'équation différentielle stochastique est linéaire au sens limité (1.3), l'équation (1.16) reste la même, par contre l'équation (1.17) devient

$$\frac{dP(t)}{dt} = 2a_1(t)P(t) + 2a_2(t)m(t) + b_2^2(t)$$

Exemple 2.2

Dans l'exemple pratique de Langevin, on peut estimer les paramètres a et b en calculant la moyenne et le moment d'ordre deux de X_t , donnés par :

$$m(t) = m(0)e^{-at} \quad \text{et} \quad P(t) = P(0)e^{-2at} + \frac{b^2}{2a}(1 - e^{-2at})$$

Les deux expressions de $m(t)$ et $P(t)$ sont obtenues en résolvant les équations **(1.16)** et **(1.17)** dans ce cas particulier. En effet, $m'(t) = -am(t)$ et $P'(t) = -2aP(t) + b^2$ sont des équations différentielles ordinaires de premier ordre.

On rappelle que la solution générale d'une équation de la forme : $\frac{dx(t)}{dt} = \alpha x(t) + \beta(t)$ est donné par $x(t) = e^{\alpha t} \left(x_0 + \int_0^t e^{-\alpha s} \beta(s) ds \right)$ où $x_0 = x(0)$ et β est une fonction continue en t .

Lorsque $\beta(t) \equiv \beta$ (indépendante de t), la solution est $x(t) = x_0 e^{\alpha t} - \frac{\beta}{\alpha}(1 - e^{\alpha t})$.

2.3 Résolution des équations différentielles stochastiques réductibles d'Itô

Certaines équations différentielles non linéaires de type

$$dY_t = a(t, Y_t)dt + b(t, Y_t)dW_t \quad \text{(2.1)}$$

Peuvent s'écrire sous la forme linéaire :

$$dX_t = (a_1(t)X_t + a_2(t))dt + (b_1(t)X_t + b_2(t))dW_t \quad \text{(2.2)} \quad \text{avec } X_t = U(t, Y_t)$$

Si $\frac{dU}{dy}(t, y) \neq 0$, alors $y = V(t, x)$ est dit l'inverse local de $x = U(t, y)$ (Théorème d'inversion des fonctions) où $x = U(t, V(t, x))$ et $y = V(t, U(t, y))$.

Considérons le cas où $a_1(t) \equiv b_1(t) \equiv 0$, $a_2(t) \equiv \alpha(t)$, $b_2(t) \equiv \beta(t)$. La proposition ci après est une condition suffisante permettant la transformation de l'équation réductible **(2.1)** dans ce cas particulier.

Proposition 2.4

Soit la fonction $\gamma(t, y)$, défini par

$$\gamma(t, y) = \frac{1}{b(t, y)} \frac{db}{dt}(t, y) - b(t, y) \frac{d}{dy} \left(\frac{a(t, y)}{b(t, y)} - \frac{1}{2} \frac{db}{dy}(t, y) \right) \quad (2.3)$$

Si $\frac{d\gamma}{dy}(t, y) = 0$ alors l'équation différentielle stochastique non linéaire (2.1) peut être transformée en équation différentielle stochastique explicitement intégrable de la forme :

$$dX_t = \alpha(t)dt + \beta(t)dW_t \quad (2.4)$$

et cela en effectuant la transformation $x = U(t, y)$ avec

$$U(t, y) = c \exp \left(\int_0^t \gamma(s, y) ds \right) \left(\int_0^y \frac{1}{b(t, z)} dz \right) \quad (2.5)$$

où c est une constante arbitraire.

Preuve

La solution Y_t de (2.1) est de la forme $Y_t = V(t, X_t)$ où X_t est donnée en (2.2). Il en résulte de la formule d'Itô (I.4.4) que la différentielle de la transformée U est

$$dU(t, y) = \left[\frac{dU}{dt}(t, y) + a(t, y) \frac{dU}{dy}(t, y) + \frac{1}{2} b^2(t, y) \frac{d^2U}{dy^2}(t, y) \right] dt + \left[b(t, y) \frac{dU}{dy}(t, y) \right] dW_t$$

Comme $X_t = U(t, Y_t)$, alors par identification à l'équation différentielle linéaire (2.4), on arrive aux égalités suivantes :

$$\frac{dU}{dt}(t, y) + a(t, y) \frac{dU}{dy}(t, y) + \frac{1}{2} b^2(t, y) \frac{d^2U}{dy^2}(t, y) = \alpha(t) \quad (2.6)$$

$$b(t, y) \frac{dU}{dy}(t, y) = \beta(t) \quad (2.7)$$

En dérivant par rapport à y , les deux égalités ci-dessus deviennent :

$$\frac{d^2U}{dt dy}(t, y) = - \frac{d}{dy} \left(a(t, y) \frac{dU}{dy}(t, y) + \frac{1}{2} b^2(t, y) \frac{d^2U}{dy^2}(t, y) \right) \quad \text{et} \quad \frac{d}{dy} \left(b(t, y) \frac{dU}{dy}(t, y) \right) = 0$$

En plus, lorsqu'on dérive (2.7) par rapport à t , on obtient :

$$b(t, y) \frac{d^2 U}{dt dy}(t, y) + \frac{db}{dt}(t, y) \frac{dU}{dy}(t, y) = \beta'(t)$$

En supposant que $b(t, y) \neq 0$ et en remplaçant $\frac{d^2 U}{dt dy}(t, y)$ dans l'expression de $\beta'(t)$, à savoir que $\frac{dU}{dy}(t, y) = \frac{1}{b(t, y)}\beta(t)$, on obtient l'expression ci-dessous :

$$\beta'(t) = \beta(t)b(t, y) \left[\left(\frac{1}{b^2(t, y)} \right) \frac{db}{dt}(t, y) - \frac{d}{dy} \left(\frac{a(t, y)}{b(t, y)} \right) + \frac{1}{2} \frac{d^2 b}{dy^2}(t, y) \right] = \beta(t)\gamma(t, y)$$

Proposition 2.5

L'équation différentielle stochastique non linéaire homogène :

$$dY_t = a(Y_t)dt + b(Y_t)dW_t \quad (2.8)$$

Peut être transformé en équation linéaire ayant la forme

$$dX_t = (a_1 X_t + a_2)dt + (b_1 X_t + b_2)dW_t \quad (2.9)$$

Ce changement de la forme de l'équation initiale (2.8) se fait en effectuant la transformation $X_t = U(Y_t)$, définit par

$$U(Y_t) = \left\{ \begin{array}{ll} C \exp \left(b_1 \int_{y_0}^{y_t} \frac{1}{b(z)} dz \right) & \text{si } b_1 \neq 0 \\ b_2 \int_{y_0}^{y_t} \frac{1}{b(z)} dz + C & \text{si } b_1 = 0 \end{array} \right\} \quad (2.10)$$

Preuve

Comme U est une fonction indépendante du temps, alors $\frac{dU(y)}{dt} = 0$ et dans ce cas les deux égalités (2.6) et (2.7) auront la forme :

$$a(y) \frac{dU}{dy}(y) + \frac{1}{2} b^2(y) \frac{d^2 U}{dy^2}(y) = a_1 U(y) + a_2 \quad (2.11)$$

$$b(y) \frac{dU}{dy}(y) = b_1 U(y) + b_2 \quad (2.12)$$

En supposant que $b(y) \neq 0$ et $b_1 \neq 0$, il s'ensuit de (2.12) que

$$U(y) = C \exp \left(b_1 \int_{y_0}^y \frac{1}{b(z)} dz \right) - \frac{b_2}{b_1} \quad (2.13)$$

Où C est une constante arbitraire.

Maintenant, posons $B(y) = \int_{y_0}^y \frac{1}{b(z)} dz$, on a $\frac{dU}{dy}(y) = b_1 \frac{dB}{dy}(y) C \exp(b_1 B(y))$

$$\text{où} \quad \frac{dB}{dy}(y) = \frac{d}{dy} \int_{y_0}^y \frac{1}{b(z)} dz = \frac{1}{b(y)}$$

On a :

$$\frac{d^2 U}{dy^2}(y) = b_1 \frac{d^2 B}{dy^2}(y) C \exp(b_1 B(y)) + b_1^2 \left(\frac{dB(y)}{dy} \right)^2 C \exp(b_1 B(y))$$

$$\text{où} \quad \frac{d^2 B}{dy^2}(y) = -\frac{db}{dy}(y)/b^2(y)$$

$$\Rightarrow a(y) \frac{dU}{dy}(y) = b_1 \frac{a(y)}{b(y)} V(y) \quad \text{et} \quad b^2(y) \frac{d^2 U}{dy^2}(y) = -b_1 \left(\frac{db(y)}{dy} - b_1 \right) V(y)$$

$$\text{où} \quad V(y) = C \exp(b_1 B(y))$$

L'expression (2.11) implique que $V(y) [b_1 A(y) + \frac{1}{2} b_1^2 - a_1] = a_2 - a_1 \frac{b_2}{b_1}$

$$\text{Où} \quad A(y) = \frac{a(y)}{b(y)} - \frac{1}{2} \frac{db(y)}{dy}$$

En dérivant à présent l'expression ci-dessus, on aura $\frac{d}{dy} (V(y) [b_1 A(y) + \frac{1}{2} b_1^2 - a_1]) = 0$

$$\Rightarrow \frac{d}{dy} V(y) \left[b_1 A(y) + \frac{1}{2} b_1^2 - a_1 \right] + b_1 \frac{d}{dy} A(y) V(y) = 0 \quad \text{où} \quad \frac{d}{dy} V(y) = \frac{b_1}{b(y)} C \exp(b_1 B(y))$$

$$\Rightarrow \frac{b_1}{b(y)} \exp(b_1 B(y)) \left[b_1 A(y) + \frac{1}{2} b_1^2 - a_1 \right] + b_1 \frac{d}{dy} A(y) \exp(b_1 B(y)) = 0$$

En multipliant cette dernière expression par $\frac{b(y)}{b_1} \exp(-b_1 B(y))$ et en dérivant encore une

fois on obtient le résultat suivant :

$$b_1 \frac{d}{dy} A(y) + \frac{d}{dy} \left(b(y) \frac{d}{dy} A(y) \right) = 0 \quad (2.14)$$

Cette expression est vérifiée si $\frac{d}{dy} A(y) = 0$ où $b_1 = -\frac{d}{dy} \left(b(y) \frac{d}{dy} A(y) \right) / \frac{d}{dy} A(y)$

Si $b_1 \neq 0$ la transformation adéquate U est donné par $U(y) = C \exp \left(b_1 \int_{y_0}^y \frac{1}{b(z)} dz \right)$

Sinon $U(y) = b_2 \int_{y_0}^y \frac{1}{b(z)} dz + C$

Exemple 2.3

Soit l'EDS non linéaire :

$$dY_t = -\frac{1}{2} \exp(-2Y_t) dt + \exp(-Y_t) dW_t$$

La solution générale de cette équation est donnée par : $Y_t = \ln(\exp(Y_0) + W_t)$ pour $t \in [0, T[$ où $T = T(Y_0(\omega)) = \text{Min}\{t \geq 0 : W_t(\omega) = -\exp(Y_0(\omega))\}$ est le temps d'arrêt de la trajectoire.

Dans ce cas particulier $a(y) = -\frac{1}{2} \exp(-2y)$ et $b(y) = \exp(-y)$

En effet, comme l'expression (2.14) est vérifiée pour toute valeur de b_1 car $A(y) = 0$, alors on peut choisir $b_1 = 0$ et $b_2 = 1$ et donc l'expression (2.12) implique : $b(y) \frac{dU}{dy}(y) = 1$

En remplaçant l'expression de $U(y)$ dans (2.11), a_1 et a_2 seront nuls ($a_1 = a_2 = 0$).

Dés lors, l'EDS non linéaire plus haut se réduit en EDS linéaire :

$$dX_t = dW_t \quad \text{avec} \quad X_t = U(Y_t) = \exp(Y_t)$$

Cette équation linéaire possède la solution $X_t = X_0 + W_t$ (voir l'équation 1.10) $\Rightarrow \exp(Y_t) = \exp(Y_0) + W_t$ et donc la solution de l'EDS non linéaire sera : $Y_t = \ln(\exp(Y_0) + W_t)$ pour $t \in [0, T[$.

Pour plus de détails sur la résolution des EDS linéaires et non linéaires, voir [2], [5] et [7]

2.4 Existence et Unicité des solutions des équations différentielles stochastiques

Lorsque la solution de l'EDS ne change pas pour une autre valeur du processus de Wiener, alors on parle de solution forte. Dans le cas contraire, c'est-à-dire que si on trouve une solution correspondante à un processus de Wiener bien particulier, alors elle est dite solution faible. Certaines équations différentielles stochastiques admettent des solutions faibles et non fortes. Considérons l'EDS

$$dX_t = a(t, X_t)dt + b(t, X_t)dW_t \quad (3.1)$$

s'écrivant sous la forme d'une équation intégrale stochastique :

$$X_t = X_0 + \int_{t_0}^t a(s, X_s)ds + \int_{t_0}^t b(s, X_s)dW_s \quad (3.2)$$

La solution $\{X_t, t \in [t_0, T]\}$ de (3.2) devrait assurer la signification des deux intégrales, où plus précisément leurs existences. Cela est vérifié si pour tout a et b définis par $a(t, \omega) = b(t, \omega) = 0, (0 \leq t < t_0)$, les fonctions $\sqrt{|a|}$ et b sont mesurables pour $t_0 \leq t \leq T$.

Une solution (forte) de l'EDS (3.2), est un processus continu tel que, sous certaines conditions sur a et b , l'équation différentielle (3.1) a une unique solution (comme pour une équation différentielle ordinaire). Donc la solution de (3.1) est un processus X continu tel que les intégrales $\int_0^t a(s, X_s)ds$ et $\int_0^t b(s, X_s)dW_s$ ont un sens et l'égalité (3.2) est satisfaite pour tout t , avec

$$\int_0^t |a(s, X_s)| ds < \infty \quad \text{et} \quad \int_0^t |b(s, X_s)| dW_s < \infty \quad \text{p.s}$$

Théorème d'existence et d'unicité

On suppose que

–Les fonctions a et b sont continues.

–Il existe k tel que pour tout $t \in [0, T], x \in \mathbb{R}, y \in \mathbb{R}$.

$$(1) |a(t, x) - a(t, y)| + |b(t, x) - b(t, y)| \leq k |x - y|.$$

$$(2) |a(t, x)^2 + |b(t, x)|^2 \leq k^2(1 + |x|^2).$$

–La condition initiale X_0 est indépendante de $(W_t, t \geq 0)$ et est carré intégrable.

Alors il existe une unique solution de (3.2) à trajectoires continues pour $t \leq T$.

Cette solution vérifie :

$$E \left(\sup_{0 \leq t \leq T} |X_t|^2 \right) < \infty$$

L'unicité signifie que si $(X_t)_{0 \leq t \leq T}$ et $(Y_t)_{0 \leq t \leq T}$ sont deux solutions de (3.2),

Alors :

$$\forall 0 \leq t \leq T, \quad \text{p.s} \quad X_t = Y_t$$

Nos principales références pour cette section sont [2],[7],[9] et [13].

La résolution des équations différentielles stochastiques (EDS) d'une manière explicite est rare dans la pratique. Cependant, il existe de nombreuses méthodes numériques pour surmonter ce problème en élaborant des schémas d'approximations.

2.5 L'Approximation des EDS par discrétisation du temps

2.5.1 Développement de Taylor Stochastique

Le développement de Taylor stochastique est une généralisation du développement de Taylor déterministe et il est essentiel pour l'approximation par discrétisation du temps des équations différentielles stochastiques. La formule de Taylor stochastique est basé sur une application itérative de la formule d'Itô et est appelée "Formule d'Itô-Taylor stochastique" que nous présenterons d'une manière assez simple. [20].

Considérons la forme intégrale d'une équation différentielle stochastique

$$X_t = X_{t_0} + \int_{t_0}^t a(X_s)ds + \int_{t_0}^t b(X_s)dW_s \quad (4.1)$$

Pour $t \in [t_0, T]$, où il est supposé que les fonctions a et b sont "suffisamment" lissées au voisinage de X_{t_0} . Si nous appliquons la formule d'Itô **(I.4.4)** pour les fonctions a et b , supposées homogènes, nous obtenons :

$$X_t = X_{t_0} + a(X_{t_0}) \int_{t_0}^t ds + b(X_{t_0}) \int_{t_0}^t dW_s + R \quad (4.2)$$

$$\begin{aligned} R = & \int_{t_0}^t \int_{t_0}^s L^0 a(X_z) dz ds + \int_{t_0}^t \int_{t_0}^s L^1 a(X_z) dW_z ds \\ & + \int_{t_0}^t \int_{t_0}^s L^0 b(X_z) dz dW_s + \int_{t_0}^t \int_{t_0}^s L^1 b(X_z) dW_z dW_s \end{aligned} \quad (4.3)$$

Où les opérateurs L^0 et L^1 sont définis par :

$$L^0 = a \frac{\partial}{\partial x} + \frac{1}{2} b^2 \frac{\partial^2}{\partial x^2}$$

$$L^1 = b \frac{\partial}{\partial x}$$

Ceci est le développement de Taylor le plus simple, où la formule d'Itô n'est utilisée qu'une seule fois. L'intégrale déterministe dans le développement de Taylor **(4.2)** est égale à la longueur de l'intervalle discrétisé $t - t_0$, et l'intégrale stochastique est normalement distribué $N(0, t - t_0)$.

On peut pousser ce développement d'intégrants des intégrales multiples dans le reste R , les intégrales multiples avec des intégrants constants apparaissent. Par exemple si nous utilisons la formule d'Itô sur l'intégrant $L^1 b(X_z)$ dans **(4.2)** nous obtenons l'expression suivante

$$X_t = X_{t_0} + a(X_{t_0}) \int_{t_0}^t ds + b(X_{t_0}) \int_{t_0}^t dW_s + L^1 b(X_{t_0}) \int_{t_0}^t \int_{t_0}^s dW_z dW_s + \bar{R} \quad (4.4)$$

où le reste \bar{R} est une somme d'intégrales multiples avec des intégrants non-constants.

$$\begin{aligned} \bar{R} = & \int_{t_0}^t \int_{t_0}^s L^0 a(X_z) dz ds + \int_{t_0}^t \int_{t_0}^s L^1 a(X_z) dW_z ds \\ & + \int_{t_0}^t \int_{t_0}^s L^0 b(X_z) dz dW_s + \int_{t_0}^t \int_{t_0}^s \int_{t_0}^z L^0 L^1 b(X_\mu) d\mu dW_z dW_s \\ & + \int_{t_0}^t \int_{t_0}^s \int_{t_0}^z L^1 L^1 b(X_\mu) dW_\mu dW_z dW_s \end{aligned}$$

Dans la formule (4.3) le développement d'Itô-Taylor est utilisé pour obtenir des approximations par discrétisation du temps avec des différents degrés de précision. De la même manière, nous pouvons obtenir des approximations de Taylor plus précises en incluant plusieurs intégrales stochastiques multiples dans le développement de Taylor, parce que ces intégrales contiennent des informations supplémentaires sur la trajectoire du processus stochastique.

2.5.2 La convergence

Afin d'obtenir une mesure de la quantité d'erreurs introduites dans l'approximation par discrétisation du temps, deux définitions de la convergence sont présentées par la suite. La distinction entre les deux définitions se rapporte si l'approximation du processus stochastique en temps continu (4.1) est relative aux approximations des trajectoires pour tout t , ou il se rapproche juste les moments ou quelques propriétés probabilistes de (4.1).

Pour calculer l'erreur due à l'approximation des trajectoires $\{Y_t^\delta\}$, avec une taille maximale des pas δ , d'un processus d'Itô $\{X_t\}$, en utilisant le critère de l'erreur absolue $\varepsilon = E(|X_T - Y_T^\delta|)$

où l'erreur est exprimée comme l'espérance de la valeur absolue de la différence entre le processus d'Itô et l'approximation à l'instant T .

Définition 2.1 (convergence forte)

L'erreur absolue pour une approximation par discrétisation du temps généralisée Y^δ , avec un pas maximal δ , donné par : $\varepsilon(\delta) = E(|X_T - Y_T^\delta|)$

on dit que Y_T^δ converge fortement vers X_T si $\lim_{\delta \rightarrow 0} \varepsilon(\delta) = 0$

Dans une analyse numérique qu'elle soit déterministe ou stochastique, on opte pour le schéma ayant un ordre de convergence le plus élevé. dans ce type de convergence, on dit que l'approximation généralisée Y_T^δ converge dans un sens fort avec un ordre $\gamma > 0$ à l'instant T s'il existe une constante positive C , indépendante de δ , et $\delta_0 > 0$ tel que :

$$\varepsilon(\delta) \leq C\delta^\gamma \quad \forall \delta \in (0, \delta_0)$$

où δ_0 est la taille maximale des pas.

Dans beaucoup de situations pratiques nous n'avons pas besoin d'une telle convergence

forte comme l'approximation des trajectoires considérée ci-dessus. Par exemple, nous pouvons seulement intéresser au calcul des moments, des probabilités ou d'autres fonctions du processus d'Itô. Puisque les exigences pour une telle simulation ne sont pas les mêmes que pour le cas des approximations des trajectoires, c'est naturel et pratique de classer ces approximations séparément. Pour ce but nous définissons le concept de convergence faible.

Définition 2.2 (convergence faible) :

L'approximation par discrétisation du temps généralisée Y^δ , correspondante au temps de discrétisation $(\tau)_\delta$ et soit $g : \mathbb{R}^d \rightarrow \mathbb{R}$ une fonction continûment différentiable. On dit que Y_T^δ converge au sens faible vers X_T si

$$\lim_{\delta \rightarrow 0} | E(g(X_T)) - E(g(Y_T^\delta)) | = 0$$

On dit que l'approximation généralisée Y_T^δ converge faiblement avec un ordre de convergence $\beta > 0$ vers X_T , quand $\delta \rightarrow 0$ si

$\forall g \in C_p^l(\mathbb{R}^d, \mathbb{R})$ avec $l = 2(\beta + 1)$, $\exists D > 0$ (D indépendante de δ) et δ_0 , tel que :

$$| E(g(X_T)) - E(g(Y_T^\delta)) | \leq D\delta^\beta \quad \forall \delta \in (0, \delta_0)$$

où $C_p^l(\mathbb{R}^d, \mathbb{R})$ est l'espace des fonctions $g : \mathbb{R}^d \rightarrow \mathbb{R}$, l fois continûment différentiables.

Dans [20] il est démontré que les critères de la convergence forte et faible mènent au développement des différents schémas de discrétisation. Comme nous le verrons par la suite un schéma de discrétisation donné a généralement différents ordres de convergence à savoir ces deux critères de convergence.

2.5.3 Les schémas du discretisation

Approximations fortes de Taylor

Les approximations en temps discret sont basées sur le développement de Taylor stochastique quant à la convergence au sens fort, d'où l'appellation "approximation forte de Taylor". Nous décrivons les schémas forts de Taylor d'une manière succincte dans le cas général d'un processus d'Itô d'une dimension satisfaisant l'EDS de type Itô (4.1)

Schéma d'Euler stochastique : Il s'agit de l'approximation forte de Taylor la plus simple avec un ordre de convergence de 0.5, alors qu'il vaut 1 dans le cas déterministe. Nous présentons la forme de cette approximation comme suit

$$Y_{n+1} = Y_n + a(Y_n)\Delta_n + b(Y_n)\Delta W_n \quad (4.5)$$

où $\Delta_n = \int_{\tau_n}^{\tau_{n+1}} dt = \tau_{n+1} - \tau_n$ est le pas de discrétisation ou la longueur de l'intervalle de discrétisation $[\tau_n, \tau_{n+1}]$ et $\Delta W_n = W_{\tau_{n+1}} - W_{\tau_n}$ sont les incréments de processus de Wiener. Les ΔW_n sont des variables aléatoires gaussiennes indépendantes, de moyenne nulle et de variance Δ_n .

Schéma de Milstein : Ce schéma d'approximation est proposé par Milstein, son ordre de convergence est supérieur à celui d'Euler, il est égal à 1. Lorsque le processus d'Itô et de Wiener sont de dimension égale à 1, alors l'approximation de Milstein en terme d'intégrale d'Itô n'est autre que celle d'Euler plus un terme, soit

$$Y_{n+1} = Y_n + a(Y_n)\Delta_n + b(Y_n)\Delta W_n + \frac{1}{2}b(Y_n)b'(Y_n)\{(\Delta W_n)^2 - \Delta_n\} \quad (4.6)$$

Où la prime désigne la dérivée par rapport à la variable d'état. On voit aisément, que le schéma de Milstein se coïncide avec celui d'Euler si le terme de diffusion de b est indépendant de la variable d'état, parce que le dernier terme dans (4.6) part. Dû au fait que l'intégrale multiple peut être exprimé comme

$$\int_{t_0}^t \int_{t_0}^s dW_z dW_s = \frac{1}{2}\{(\Delta W_n)^2 - \Delta_n\}$$

Le schéma de Milstein semble correspondre avec le développement de Taylor stochastique (4.4).

voir [20] pour plus de détails.

Schéma fort de Taylor d'ordre 1.5 : Les intégrales stochastiques multiples fournissent des informations supplémentaires concernant la trajectoire de processus de Wiener dans un sous intervalle de discrétisation. La nécessité de les inclure dans des schémas d'ordre supérieur constitue la différence fondamentale entre l'analyse numérique des EDS et des

EDO. Grâce au développement de Taylor stochastique dans un cas particulier où $d = m = 1$ et $a(x_t) = a(x)$ et $b(x_t) = b(x)$, on obtient le schéma fort de Taylor d'ordre 1.5 :

$$\begin{aligned}
Y_{n+1} = & Y_n + a\Delta_n + b\Delta W_n + \frac{1}{2}bb'\{(\Delta W_n)^2 - \Delta_n\} + a'b\Delta Z_n + \frac{1}{2}(aa' + \frac{1}{2}b^2a'')\Delta_n^2 \\
& + (ab' + \frac{1}{2}b^2b'')\{\Delta W_n\Delta_n - \Delta Z_n\} + \frac{1}{2}b(bb'' + (b')^2)\{\frac{1}{3}(\Delta W_n)^2 - \Delta_n\}\Delta W_n \quad (4.7)
\end{aligned}$$

Où a et b sont évalués en fonction de Y_n et la variable aléatoire ΔZ_n est défini par :

$$\Delta Z_n = \int_{\tau_n}^{\tau_{n+1}} \int_{\tau_n}^s dW_s ds \quad \text{où } \Delta Z_n \approx N(0, \frac{1}{3}\Delta_n^3) \quad \text{et } E(\Delta Z_n \Delta W_n) = \frac{1}{2}\Delta_n^2$$

Le dernier terme dans (4.7) représente l'intégral d'Itô triplet $I_{(1,1,1)} = \frac{1}{2}\{\frac{1}{3}(\Delta W_n)^2 - \Delta_n\}\Delta W_n$

Approximations faibles de Taylor

Comme pour les approximations fortes, l'ordre de convergence désiré détermine où doit le développement de Taylor être tronqué. Cependant, le critère de la convergence faible ne concerne que les aspects probabilistes de la trajectoire et pas la trajectoire elle-même. Par conséquent, pour un certain degré de convergence, le nombre de termes demandé de développement est moins pour le cas de la convergence faible que pour le cas de la convergence forte, si on désire un certain degré de convergence.

Par exemple on peut montrer, que l'approximation d'Euler atteint un ordre de la convergence faible $\beta = 1.0$, alors qu'il n'atteint qu'un ordre $\gamma = 0.5$ dans la convergence forte.

Schème faible de Taylor d'ordre 2.0 : Le schéma faible de Taylor d'ordre 2.0 est donnée par :

$$\begin{aligned}
Y_{n+1} = & Y_n + a\Delta_n + b\Delta W_n + \frac{1}{2}bb'\{(\Delta W_n)^2 - \Delta_n\} + a'b\Delta Z_n + \frac{1}{2}(aa' + \frac{1}{2}b^2a'')\Delta_n^2 \\
& + (ab' + \frac{1}{2}b^2b'')\{\Delta W_n\Delta_n - \Delta Z_n\}
\end{aligned}$$

Le schéma faible de Taylor d'ordre 2.0 est plus simple que le schéma fort de Taylor d'ordre 1.5, même si le degré de convergence est plus élevé.

2.5.4 Simulation des EDS

Comme les solutions explicites des équations différentielles stochastiques n'existent que pour un nombre limité de cas, des méthodes numériques de la solution doivent être utilisées. Différentes approches numériques ont été proposées, comme les approximations par chaînes de Markov qui consiste à discrétiser simultanément le temps et l'espace des variables, mais avec de telle approche, on ne peut aboutir à la solution recherchée qu'après un nombre très important d'itérations. Cependant pour des buts de simulation nous allons utiliser des approximations par discrétisation du temps qu'ils ont été présentés dans ce chapitre. En choisissant une longueur suffisamment petite partie du sous-intervalle Δ , les schémas de discrétisation ci-dessus peuvent être utilisés pour générer des observations discrètes d'un système à temps continu.

Exemple 2.4 : Considérons le cas d'un mouvement brownien géométrique

$$dX_t = aX_t dt + bX_t dW_t \quad \forall t \in [t_0, T] \quad \text{avec } X_0 \in \mathbb{R}$$

D'après le **corollaire 2.1** la solution de cette EDS est donnée par

$$X_t = X_0 \exp\left\{\left(a + \frac{1}{2}b^2\right)t + bW_t\right\} \quad \forall t \in [t_0, T] \quad \text{et } \forall W = \{W_t, t \geq 0\}$$

La connaissance de la solution explicite nous donne la possibilité de comparer les schémas de discrétisation avec la solution exacte et de calculer l'erreur.

*Pour simuler une trajectoire par l'approximation d'Euler du mouvement brownien géométrique nous commençons simplement par la valeur initiale $Y_0 = X_0$ et continuons récursivement à générer les valeurs de la manière suivante :

$$Y_{n+1} = Y_n + aY_n \Delta_n + bY_n \Delta W_n \quad \forall n = 0, 1, 2, \dots$$

$$\text{où } \Delta W_n \approx N(0, \Delta_n) \quad \text{avec } \Delta_n = \Delta = \tau_{n+1} - \tau_n$$

*L'approximation de Milstein du mouvement brownien géométrique est donnée par :

$$Y_{n+1} = Y_n + aY_n\Delta_n + bY_n\Delta W_n + \frac{1}{2}b^2Y_n((\Delta W_n)^2 - \Delta_n)$$

$$\text{où } \Delta W_n \approx N(0, \Delta_n) \text{ avec } \Delta_n = \Delta = \tau_{n+1} - \tau_n$$

*L'approximation de Taylor d'ordre 1.5 est la suivante :

$$\begin{aligned} Y_{n+1} = & Y_n + aY_n\Delta_n + bY_n\Delta W_n + \frac{1}{2}b^2Y_n((\Delta W_n)^2 - \Delta_n) + abY_n\Delta Z_n + \frac{1}{2}(a^2Y_n)\Delta_n^2 \\ & + (aY_nb)\{\Delta W_n\Delta_n - \Delta Z_n\} + \frac{1}{2}b^3Y_n\{\frac{1}{3}(\Delta W_n)^2 - \Delta_n\}\Delta W_n \end{aligned}$$

$$\text{où } \Delta W_n \approx N(0, \Delta_n) \text{ avec } \Delta_n = \Delta = \tau_{n+1} - \tau_n \text{ et } \Delta Z_n \approx N(0, \frac{1}{3}\Delta_n^3)$$

Pour la comparaison, la trajectoire de la solution exacte est obtenue en évaluant le processus d'Itô aux instants de discrétisation :

$$X_{\tau_n} = X_0 \exp\left\{\left(a + \frac{1}{2}b^2\right)\tau_n + b \sum_{i=1}^n \Delta W_{i-1}\right\}$$

Pour plus de détails et d'exemples d'applications voir [30], [31] et [32]

Dans la Figure 2.1. le processus exact aussi bien que l'approximation d'Euler, de Milstein et de Taylor d'ordre 1.5 qui sont tracées pour différentes longueur de l'intervalle Δ . On voit aisément que les approximations deviennent mieux avec l'augmentation du nombre des sous intervalles. En outre on voit, comme prévu, que le schéma de Taylor fournit une meilleure approximation que le schéma d'Euler et le schéma de Milstein.

La Figure 2.2. montre le tracé d'une trajectoire de la solution exacte du mouvement brownien géométrique et l'approximation correspondante donnée par le schéma d'Euler, de Milstein et de Taylor avec $a = 1$, $b = 0.5$ et $X_0 = 1$ sur $[0, 1]$

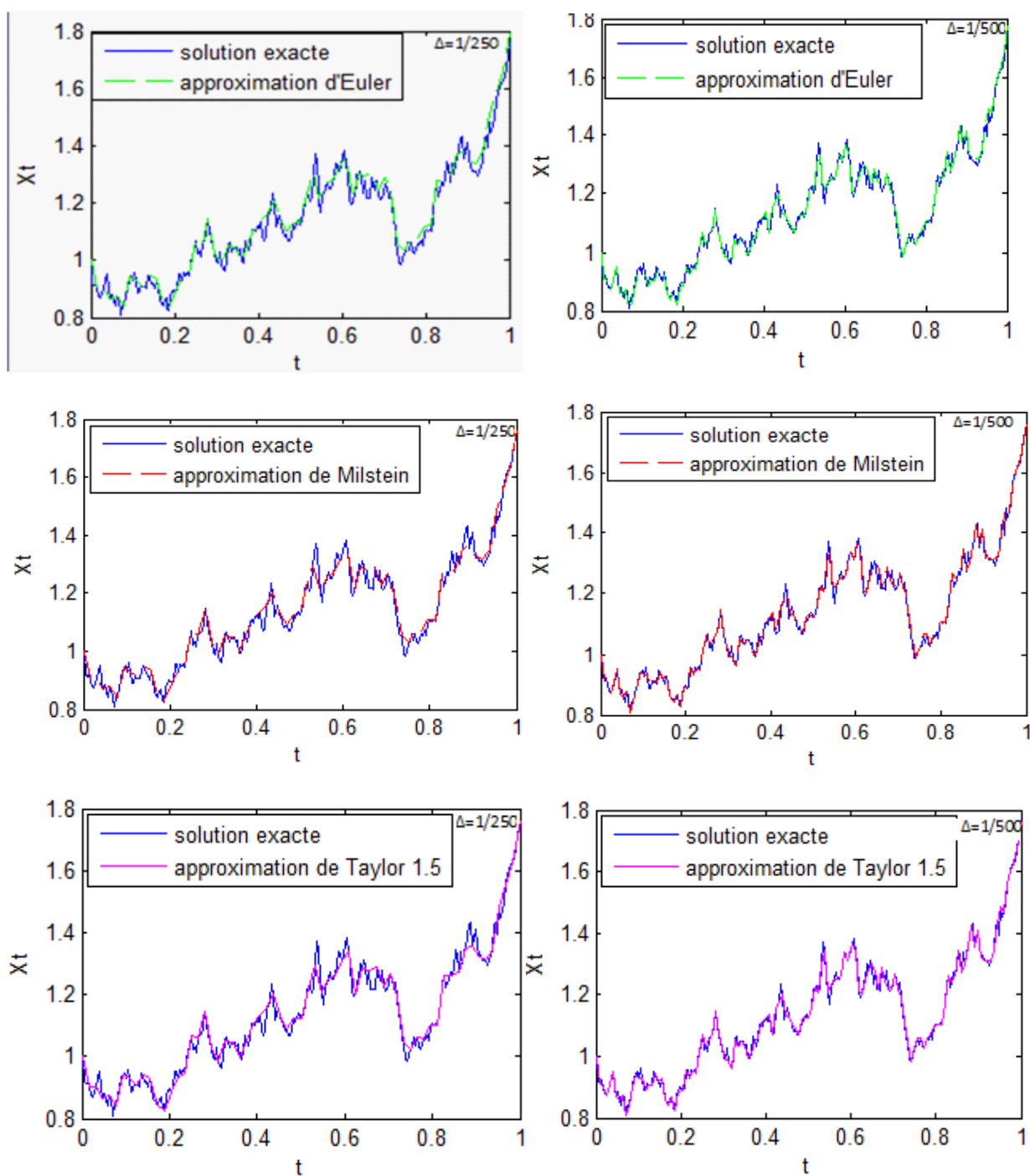


Figure 2.1-La solution exacte et les 3 schémas d'approximation d'une trajectoire de Mouvement Brownien géométrique pour différentes longueur de l'intervalle Δ

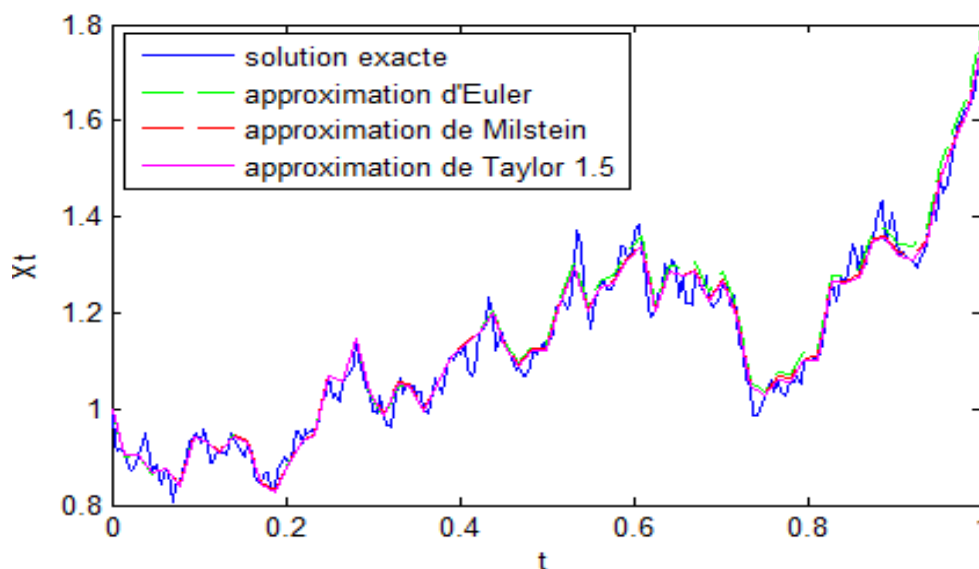


Figure 2.2- *La Comparaison de la solution exacte et les 3 schémas d'approximation d'une trajectoire de Mouvement Brownien géométrique*

La discrétisation exacte et selon les trois schémas d'approximation sont relativement proches graphiquement, l'erreur d'approximation selon les trois schémas est présentée dans le tableau suivant :

Le pas de discrétisation Δ	$\Delta=1/250$	$\Delta=1/500$
L'erreur selon le schéma d'Euler	0.4809	0.4243
L'erreur selon le schéma de Milstein	0.4373	0.3989
L'erreur selon le schéma de Taylor 1.5	0.4326	0.3871

Tableau 2.1- *l'erreur d'approximation selon les 3 schémas d'approximation*

La méthode d'approximation forte de Taylor d'ordre 1.5 donne une erreur d'approximation égale 0.4326 pour un pas $1/250$ et 0.3871 pour un pas de $1/500$, les deux autres schémas montre des résultats moins précises, on voit que le schéma de Milstein donne meilleure précision que cel d'Euler, mais le schéma de Milstein pose des problèmes pratiques de mise en œuvre dès lors que l'on s'intéresse à des phénomènes de dimension supérieure ou égale à 2, c'est-à-dire dès que l'on cherche à simuler des vecteurs aléatoires, voir [20] et [25].

2.6 Conclusion

Le but des équations différentielles stochastiques est de fournir un modèle mathématique pour une équation différentielle perturbée par un bruit aléatoire. Partons d'une équation différentielle ordinaire, ces équations permettent de modéliser des trajectoires aléatoires, telles des cours boursiers ou les mouvements de particules soumises à des phénomènes de diffusion. Une solution de telle équation est dite processus de diffusion (ou diffusion).

Or, en général, on ne peut pas les résoudre explicitement. Par exemple, la plupart des modèles de taux d'investissement interne conduisent à des équations qui n'ont pas de solutions explicites. De plus, même si on trouve une solution, celle-ci peut être trop complexe pour être simulée directement. Il existe de très nombreuses méthodes numériques de résolution des équations différentielles : plus ou moins précises, plus ou moins coûteuses en temps de calcul, plus ou moins adaptées à tel ou tel type d'équation. Toutes ces méthodes ont été implémentées et les meilleures sont proposées dans les logiciels de calcul standard. Dans notre travail nous avons présenté seulement les méthodes les plus simples, qui sont les schémas d'approximation de d'Euler, de Milstein et de Taylor d'ordre 1.5. les trois méthodes donne des bon résultats mais parmi les trois celle de Taylor qui donne des résultats plus précis.

Chapitre 3

Modèle de Black-Scholes et estimation des paramètres

3.1 Introduction

Dans ce chapitre nous introduirons le modèle de Black- Scholes qui reste le plus utilisé dans le domaine de l'ingénierie financière, le modèle dépend a priori de deux paramètres endogènes : la tendance μ (espérance du rendement instantané du sous-jacent) et la volatilité σ (écart type du rendement instantané du sous-jacent). Cependant, le premier problème, auquel on se trouve confronté, et qui concerne les modèles en général est celui de l'estimation des paramètres. Dans le premier paragraphe de ce chapitre, nous présentons le modèle de Black- Scholes ainsi ses principales propriétés. Le second, est exclusivement consacré aux techniques d'estimation, nous commençons par un bref aperçu d'une méthode courante, et nous illustrons, ensuite, deux méthodes d'estimation.

3.2 Présentation du modèle de Black & Scholes

Le modèle de Black-Scholes (du nom de Fischer Black et Myron Scholes) est un modèle de l'évaluation et la couverture d'une option de type européen sur une action de dividende et est utilisé en mathématiques financières afin d'estimer en théorie la valeur d'une option financière. La formule de Black et Scholes qui a été développée pour des actifs dont le comportement est celui du Mouvement Brownien Géométrique et qui permet de calculer le prix d'une option, a apporté en 1973 une contribution majeure à la théorie financière qui a eu un impact énorme, non seulement au niveau théorique mais également au niveau pratique

puisque le modèle de Black & Scholes est, sans aucun doute, le premier modèle théorique d'évaluation à avoir été utilisé de manière aussi intensive par les professionnels à des fins d'évaluation, de spéculation ou simplement de couverture.

Le modèle de Black & Scholes est, à l'origine, un modèle à deux actifs : l'un risqué, l'autre pas. Typiquement, l'actif risqué est une action (l'action sous-jacente à l'option) tandis que l'actif non risqué s'apparente à une obligation. À l'instant t , le prix de l'obligation est R_t et le prix de l'action est S_t .

-L'évolution de l'obligation est relativement simple puisqu'on suppose que :

$$dR_t = r_t R_t dt \Rightarrow \int \frac{dR_t}{R_t} = \int r_t dt \Rightarrow \log R_t = \int_0^t r_s ds$$

D'où, on obtient :

$$R_t = R_0 e^{\int_0^t r_s ds}$$

Où, $r_t \geq 0$ représente le taux d'investissement interne. Nous supposons toujours que $R_0 = 1$.

-Le prix de l'action $\{S_t\}_{t \geq 0}$ est régi par l'équation différentielle stochastique

$$dS_t = S_t(\mu_t dt + \sigma_t dW_t), \quad S_0 > 0 \quad \text{donné,}$$

Où, $\sigma_t \geq 0$ représente la volatilité et μ_t est un paramètre réel désigne le rendement instantané de S . Et bien évidemment $\{W_t\}_{t \geq 0}$ est un mouvement brownien standard.

Dans le modèle de Black-Scholes originel, les paramètres r , μ et σ sont des constantes. On a dans ce cas

$$R_t = \exp(rt)$$

et

$$dS_t = \mu S_t dt + \sigma S_t dW_t \iff S_t = S_0 + \int_0^t \mu S_s ds + \int_0^t \sigma S_s dW_s \quad (1.1)$$

Avec $S_0 = s_0 > 0$ donné

L'équation (1.1) est une équation différentielle stochastique de type linéaire (présenté au chapitre 2, Corollaire 2.1 avec $S_t = X_t$, $\mu = a$ et $\sigma = b$).

Alors la solution S_t de l'équation (1.1) est donné par :

$$S_t = s_0 \exp[\sigma W_t + (\mu - \frac{1}{2}\sigma^2)t]$$

3.2.1 Les conditions du modèle

Le modèle de Black & Scholes est fondé sur une série d'hypothèses plus ou moins vérifiées dans la réalité :

- Le prix de l'actif sous-jacent suit un mouvement brownien géométrique avec une volatilité σ constante est connue à l'avance et une dérivée μ constante : $dS_t = \mu S_t dt + \sigma S_t dW_t$, où W_t est un mouvement brownien (processus de Wiener)

- La vente à découvert d'actifs avec utilisation du produit de la vente est autorisée
- Il n'y a ni taxes ni coûts de transaction
- Les actifs sont parfaitement divisibles
- L'action ne verse pas de dividende
- Il n'y a pas d'opportunités d'arbitrage sur le marché
- Les actifs en traités de manière continue sur le marché
- Le taux sans risque est constant et identique sur toutes les maturités

Ce modèle s'adapte assez bien aux conditions réelles des marchés financiers malgré quelques écarts. Par exemple, l'hypothèse centrale du modèle, à savoir la continuité du temps, n'est pas réalisée. On ne s'étonnera donc pas de déceler des écarts entre les résultats du modèle et ce que l'on peut observer sur les marchés.

Exemple de simulation d'un mouvement brownien géométrique

Comme le prix de l'actif sous-jacent suit le mouvement brownien géométrique : $dS_t = \mu S_t dt + \sigma S_t dW_t$, $S_0 = s_0$

Deux cas peuvent être présentés

Premier cas :

Si nous avons seulement besoin de S_T comme cela est le cas lors de la tarification de plusieurs droits contingents, alors nous allons profiter du fait que nous connaissons la solution de cette équation différentielle stochastique :

$$\begin{aligned} S_T &= s_0 \exp \left[\sigma W_T + \left(\mu - \frac{1}{2} \sigma^2 \right) T \right] \\ &= s_0 \exp \left[\sigma T^{1/2} Z + \left(\mu - \frac{1}{2} \sigma^2 \right) T \right] \end{aligned}$$

Telque Z est une variables aléatoires de loi normale centrée et réduite.

Nous ne simulerons donc que la valeur de S au temps T et non pas toute sa trajectoire.

Pour $\mu = 1, \sigma = 0.5, s_0 = 1$ et $t = 0 \longrightarrow S = s_0 = 1$, quand on change les valeurs de t on obtient des autres valeurs de S par exemple :

$t = 0.5 \longmapsto S = 1.048998; t = 1 \longmapsto S = 4.158062\dots$

Deuxième cas :

Si nous devons connaître la trajectoire de S comme c'est le cas pour la tarification de l'option à barrière, alors nous utiliserons la méthode la plus élémentaire : la méthode d'Euler stochastique développée dans le chapitre précédent. Le principe en est le suivant, il s'agit de simuler l'équation :

$$\left\{ \begin{array}{l} dS_t = \mu S_t dt + \sigma S_t dW_t \\ S_0 = s_0 \end{array} \right\}$$

Nous nous fixons un pas de discrétisation du temps Δt . Nous pouvons alors construire un processus à temps discret $(X_n)_{n \geq 0}$ approximant l'équation différentielle stochastique aux instants $n\Delta t$, en posant :

$$\left\{ \begin{array}{l} X_0 = s_0 = x_0 \\ X_{n+1} - X_n = \mu X_n \Delta t + \sigma X_n (W_{(n+1)\Delta t} - W_{n\Delta t}) \end{array} \right\}$$

La Simulation d'une trajectoire du mouvement brownien géométrique avec $\mu = 1$ et $\sigma = 0.5$ sur l'intervalle $[T_0, T] = [0, 1]$ et pour $n = 10000$ est représentée à la Figure **3.1**

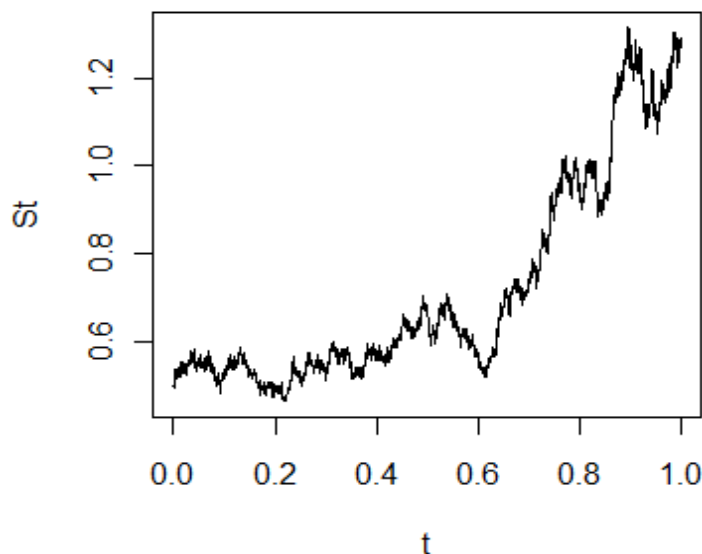


Figure 3.1- *La Simulation d'une trajectoire du mouvement Brownien géométrique*

Utilisant le R-package `SimDiffProc` et `SimDiffProcGUI`

Nous pouvons utiliser la fonction `GBM` pour simuler une seule trajectoire du mouvement brownien géométrique, sur un intervalle de temps $[T_0, T]$ avec un pas $\Delta_t = \frac{T_0 - T}{n}$, et pour un flux de trajectoires en utilisant la fonction `GBMF`. [30] et [31].

```
R> GBMF(N = 1000, M = 100, t0 = 0, T = 1, x0 = 1, theta = 1, sigma = 0.5).
```

La Figure 3.2 illustre 100 échantillons d'un actif ($S_t = X_t$), avec $\mu = \text{theta} = 1$, $\sigma = 0.5$ et $t \in [0, 1]$, le prix de l'actif à la date $t = 0$ est $s_0 = x_0 = 1$ pour $n = 1000$.

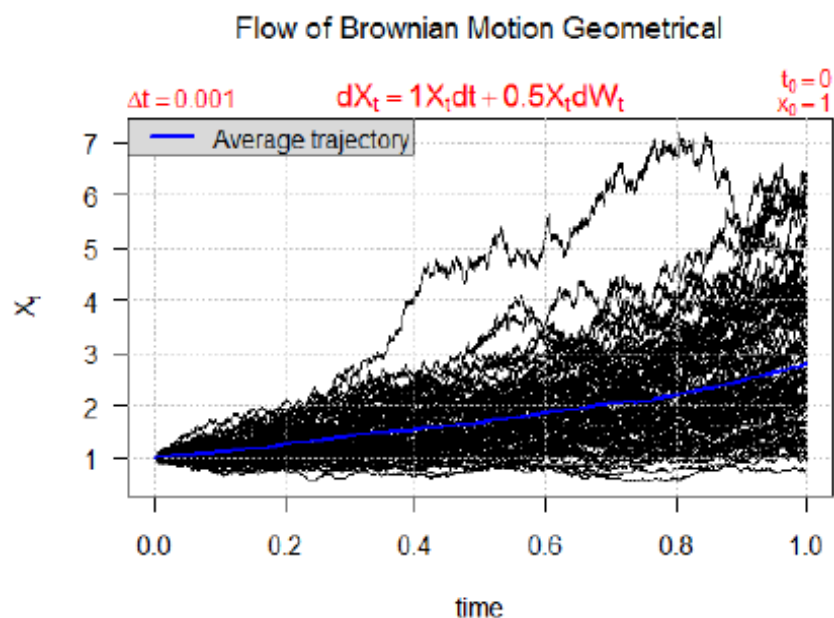


Figure 3.2-100 trajectoires d'un actif

A la date d'échéance $t = T$, le prix de l'actif est une variable aléatoire donné par la densité :

$$f(x) = \frac{\exp\left(\frac{-(\log(x/s_0) - (\mu - \sigma^2/2)T)^2}{2\sigma^2 T}\right)}{x\sigma\sqrt{2\pi T}}, \quad \text{pour } x > 0$$

avec $f(x) = 0$ pour $x \leq 0$, pour confirmer on donne la Figure 1.3, qui illustre l'histogramme où on prend la valeur finale de l'actif $S(T)$ pour 10^5 échantillons, et la courbe présente la

densité $f(x)$.

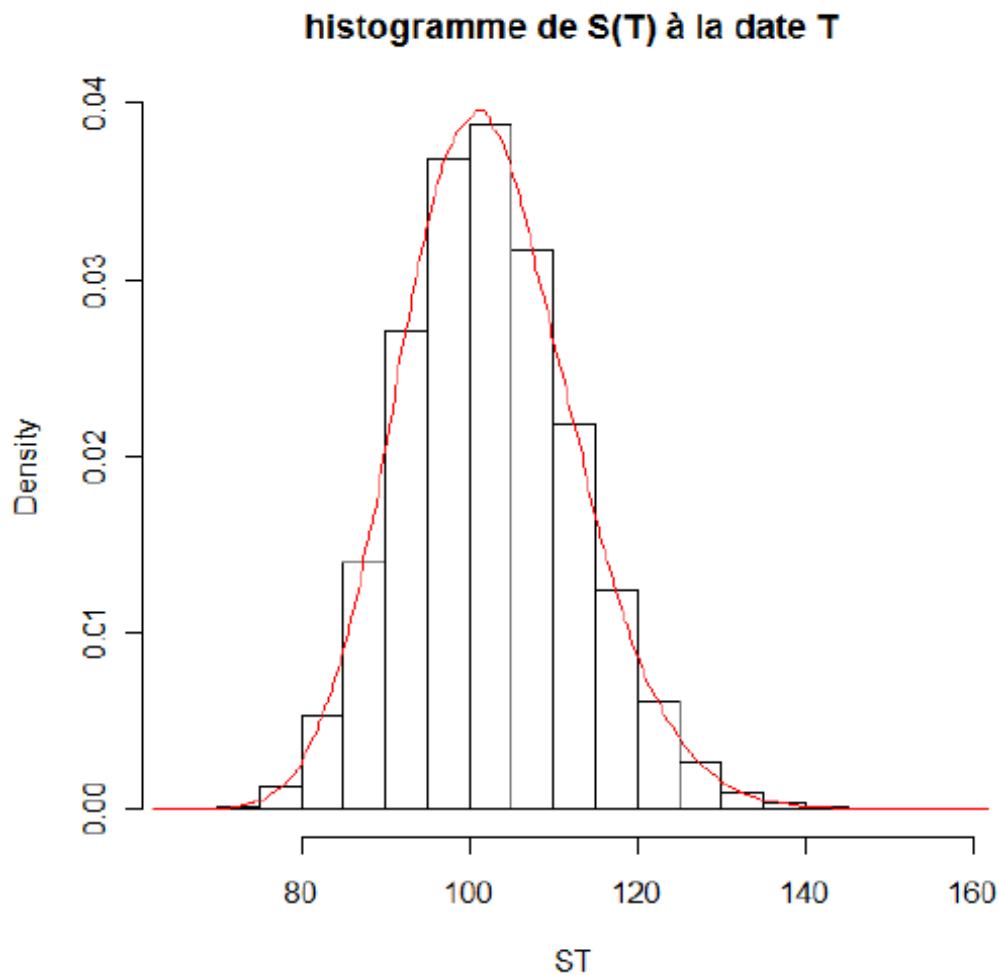


Figure 3.3—l'histogramme de S_T et la courbe de la densité $f(x)$.

3.2.2 Formule de Black-Scholes

La formule de Black-Scholes permet de calculer la valeur théorique d'une option à partir des cinq données suivantes :

- S_0 la valeur actuelle de l'action sous-jacente.
- T date d'exercice avant l'échéance (exprimé en années).
- K le prix d'exercice fixé par l'option.
- r le taux d'intérêt sans risque.

– σ la volatilité du prix de l'action.

Le prix théorique d'une option d'achat (call), qui donne le droit mais pas l'obligation d'acheter l'actif S à la valeur K à la date T , est caractérisé par son payoff :

$$(S_T - K)^+ = \max(S_T - K; 0)$$

Le prix de l'option est donné par l'espérance sous probabilité risque neutre du payoff terminal actualisé $C = E[\text{Payoff} \times e^{-rT}]$, soit la formule de Black-Scholes :

$$C_0(S_0, K, r, T, \sigma) = S_0 N(d_1) - K e^{-rT} N(d_2)$$

De même, le prix théorique d'une option de vente (put), de payoff :

$$(K - S_T)^+ = \max(K - S_T; 0)$$

est donné par :

$$P_0(S_0, K, r, T, \sigma) = -S_0 N(-d_1) + K e^{-rT} N(-d_2)$$

avec

- N la fonction de répartition de la loi normale centrée réduite $N(0, 1)$, c'est à dire

$$N(x) = \int_{-\infty}^x \frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2}u^2} du$$

$$-d_1 = \frac{1}{\sigma\sqrt{T}} \left[\ln\left(\frac{S_0}{K}\right) + \left(r + \frac{1}{2}\sigma^2\right)T \right]$$

$$-d_2 = d_1 - \sigma\sqrt{T}$$

Exemples d'application

1) La formule de Black-Schole pour un call Européen

Prenons l'exemple d'un call européen, sur un sous-jacent de valeur initiale $s_0 = 100$, le prix d'exercice K étant lui aussi égal à 100. Le taux sans risque r est de 10%, le temps avant l'échéance de l'option $T = 90/365$ en années, la volatilité du sous-jacent σ est de 20%.

Le prix de cette option call par la formule de Black et Scholes est $C_0(s_0, K, r, T, \sigma) = 5.24980$,

quand on change les valeurs de T on obtient des autres valeurs de call C_0

2) Simulation d'un call Européen par la méthodes de Monte-Carlo

La méthode dite "de Monte-Carlo" s'appuie sur l'argument simple de la loi des grands nombres. En effectuant un nombre suffisant de tirages aléatoires, la moyenne empirique convergera vers la moyenne théorique.[28]

Dans l'exemple d'un call européen, on peut par exemple tester la convergence de cette méthode vers le prix donné par l'application de la formule de Black et Scholes. La seule variable aléatoire à simuler est le cours du sous-jacent à la date d'exercice S_t . Pour approcher le prix "Black et Scholes", il faut donc simuler le tirage aléatoire d'une variable log-normale. On effectue l'approximation du prix de ce call par la méthode de Monte-Carlo sous la condition de risque $\mu = r$, autrement dit, nous pouvons reproduire la valeur de l'option par fixer $\mu = r$ dans le modèle de l'actif et calculer la moyenne de payoff de tous les trajectoires de l'actif; voici les résultats obtenus en fonction du nombre de tirages (gardons les mêmes paramètres que l'exemple précédent).

Nombre de tirages	Résultat	Ecart avec Black et Schole
1 000	5.21623	-3.35%
10 000	5.23246	-1.73%
100 000	5.25443	0.46%
1 000 000	5.24623	-0.35%

Table 3.1- *Approximation du prix de l'option par la méthode de Monte – Carlo*

Comme le montrent les résultats ci-dessus, la convergence n'est pas exceptionnelle : en effet même 1000000 de tirages ne permettent pas d'obtenir une précision inférieure à 0,30%. Certains cas nécessitent toutefois l'utilisation de méthodes de Monte-Carlo. En effet, sur certaines options complexes (par exemple dans le cas d'une option américaine de type "look-back"), il n'est pas toujours possible de trouver une formule fermée permettant de calculer le prix de manière analytique.

3) Les options et les contrats d'assurance vie [33]

Les idées développées dans le modèle de Black-Scholes peuvent être utilisées dans des modèles d'assurance-vie. Les prestations relatives aux contrats d'assurance vie peuvent être représentées financièrement par des options de vente (puts). Un contrat d'assurance vie en

unités de compte (UC) est un contrat qui garantit un montant indexé sur un actif financier. Nous disons d'un tel contrat qu'il est muni d'une garantie plancher lorsque celui-ci verse en cas de décès le maximum entre un montant garanti (par exemple le montant initial investi) et la valeur de l'UC. De ce fait, l'engagement de l'assureur à la date t au cas de décès (qui sera une date aléatoire) équivaut à : $\max(K, S_t) = (S_t + \max(K - S_t, 0))$, où K correspond au montant de la garantie plancher et S_t la valeur de l'UC au moment t . Par exemple, l'assuré paie une prime unique de $100000DA$ avec une garantie sur capital de 100% il sera donc sûr à échéance de détenir au minimum une somme de $K = 100000DA$ et cela même si la valeur boursière de son portefeuille est inférieure à la prime initiale investie. Nous constatons alors que l'engagement à la date t du réassureur en cas de décès sur ce type de produit est équivalent au flux terminal d'une option de vente européenne (put).

Supposons que $T = T_a$ corresponde à la durée de vie restante d'un assuré âgé de a années aujourd'hui et que celle-ci soit indépendante de la valeur de l'UC S .

On aura que $1 - F_{T_a}(t) = P[T_a > t]$ représentera la probabilité que l'assuré survive dans les t prochaines années.

De même $F_{T_a}(t) = P[T_a < t]$ représente la probabilité qu'il décède dans les t prochaines années.

La distribution de T_a dépend de l'âge, du sexe, du pays, de la santé,

Applications numériques

L'engagement de l'assureur est égal une série d'options financiers de vente détenues par l'assuré pondérées par les probabilités de mortalité. Le strike de chacun des puts est le capital garanti et le sous-jacent (prix et volatilité). Nous prendrons un strike et un prix du sous-jacent égal à 100 pour l'application numérique, l'écart-type du log-rendement annualisé est de ce fait de 20%, les taux sans risque des différentes maturités sont de 3,5%.

La formule exacte qui permet de calculer la garantie plancher est donnée par : Prix de la garantie plancher = $\sum_{i=1}^T \frac{d_{x+i}}{l_x} put(i+1)$ où d_x et l_x sont extraits de la table TD 99 (table de mortalité algérienne 1997-1999).

Remarque 3.1

La construction des tables de mortalité utilisables dans le secteur des assurances nécessite des données très détaillées et les plus fiables possible, caractéristiques qui font précisément

défaut dans les pays en voie de développement en général.

Nous allons faire varier certains paramètres du modèle tel que l'âge de l'assuré, la volatilité ou encore le taux sans risque afin d'analyser brièvement la sensibilité de la garantie plancher à ces paramètres.

Illustration des resultants : Une garantie plancher en cas de décès pour un assuré âgé de 45 ans à 75 ans souscrivant à un contrat valable jusqu'à l'âge 75 ans.

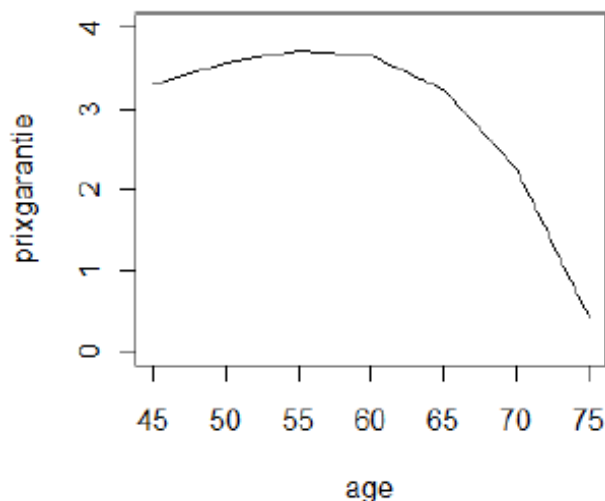


Figure 3.4-Evolution du prix de la garantie plancher pour un assuré âgé de x années pour un contrat valable jusqu'à 75 ans

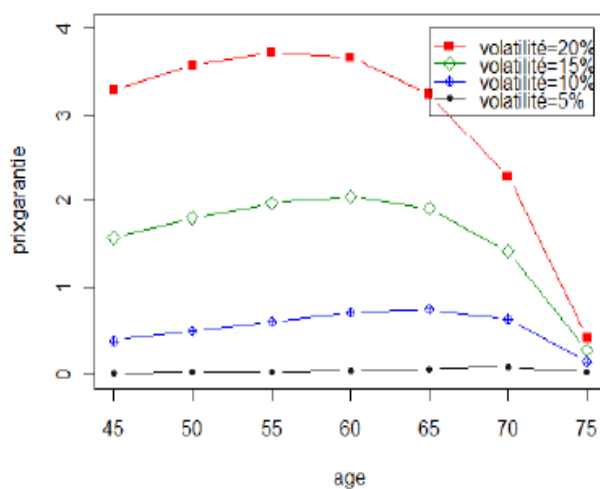


Figure 3.5-Impact de la volatilité sur le prix de la garantie plancher

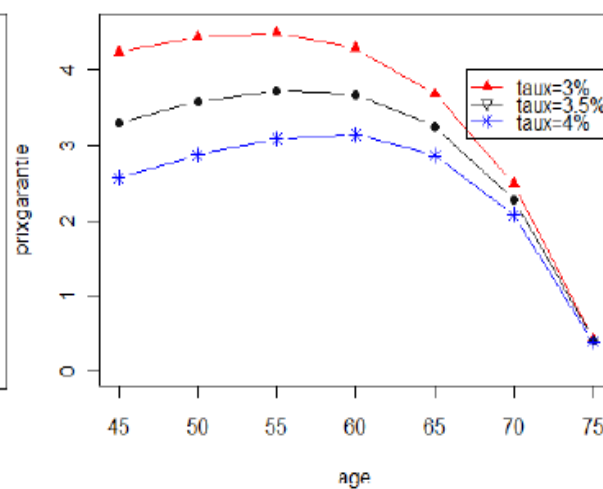


Figure 3.6-Impact du taux sans risque sur le prix de la garantie plancher

La Figure 3.4 montre que le prix de la garantie plancher a connu une augmentation légère jusqu'à l'âge 55 ans après il diminue en fonction de l'âge de l'assuré. Nous remarquons aussi que le prix de la garantie plancher augmente avec la volatilité et diminue avec le taux d'intérêt d'après la Figure 3.5 et la Figure 3.6 respectivement. La garantie plancher étant une série d'options de vente pondérées par les probabilités de mortalité, ce résultat est alors connu de la théorie des options.

3.2.3 Le modèle de Black et Scholes en pratique

Le modèle de Black & Scholes repose sur un principe fondamental : le sous-jacent d'une option traité couramment sur les marchés indique implicitement le prix d'une option d'achat. Parmi les paramètres du modèle (durée à l'échéance T , prix d'exercice K , taux d'intérêt sans risque r , prix du sous-jacent S et volatilité σ), seule cette dernière n'est pas facilement observable sur les marchés et est difficile à évaluer.

Importance de la volatilité

Lorsqu'un produit financier suit la loi

$$dS_t = \mu S_t dt + \sigma S_t dW_t, \quad S_0 = s_0$$

On a coutume de dire que σ représente la volatilité de ce produit. Intuitivement σ représente l'écart-type de $\frac{dS}{S}$ et est lié au risque de l'actif. Il est à remarquer que le coefficient μ (appreciation rate) n'entre pas dans la formule d'évaluation de Black-Scholes.

La volatilité est une mesure statistique du comportement irrégulier d'une action dans le passé et de comment est projeté son comportement (erratique) dans le futur. La volatilité est le degré d'amplitude de la variation dans le prix du sous-jacent. La volatilité est le facteur le plus important dans le prix d'une option. Plus la volatilité est importante et plus chère est l'option (aussi bien pour un call (option d'achat) que pour un put (option de vente)). Il y a deux types de volatilité que s'applique sur l'option :

La volatilité historique : mesure les variations, les fluctuations erratiques d'une action ou d'une matière première telles qu'elles l'ont été dans le passé.

La volatilité implicite : mesure les variations, les fluctuations erratiques d'une action ou d'une matière première telles qu'elles sont anticipées dans le futur. La volatilité future d'un sous-jacent est inconnue et doit être estimée

Les investisseurs peuvent comparer la volatilité implicite à la volatilité historique et faire leur propre opinion sur la volatilité future du sous-jacent. La volatilité joue un rôle crucial, et c'est l'élément le plus insaisissable, confus et difficile à appréhender par tous les participants de marché. Le calcul de la volatilité est un problème difficile pour l'application mathématique.

Estimation de la volatilité à partir de données historiques

Pour estimer empiriquement la volatilité du prix d'une action, on doit observer le prix de l'action en question dans des intervalles de temps fixe(ex : chaque jour, chaque semaine ou chaque mois).

On définit alors :

n : Nombre d'observations ;

S_i : Le prix de l'action à la fin du i^{eme} intervalle

τ : Longueur de l'intervalle de temps en années

et soit :

$$u_i = \log \left(\frac{S_i}{S_{i-1}} \right) \quad \text{pour } i = 1, \dots, n$$

L'estimateur usuel, s , de la déviation standard des u_i est donné par :

$$s = \sqrt{\frac{1}{n-1} \sum_{i=1}^n (u_i - \bar{u})^2}$$

ou

$$s = \sqrt{\frac{1}{n-1} \sum_{i=1}^n u_i^2 - \frac{1}{n(n-1)} \left(\sum_{i=1}^n u_i \right)^2}$$

Où \bar{u} est la moyenne des u_i .

Or d'après ce qui précède, la déviation standard des u_i est $\sigma\sqrt{\tau}$. Donc la variable s estime $\sigma\sqrt{\tau}$. d'où on peut estimer σ par σ^* tel que : $\sigma^* = \frac{s}{\sqrt{\tau}}$

On peut facilement montrer que l'erreur standard de cette estimation est approximativement

$$\frac{\sigma^*}{\sqrt{2n}}.$$

Le choix de la valeur n n'est cependant pas facile, mais plus le nombre de données est grand, meilleure sera l'approximation. Cependant, on sait maintenant que σ change au cours du temps et de vieilles données peuvent ne pas être révélatrice de la tendance future de la volatilité. Par exemple dans *Hull (1999) [36]*, il est conseillé de prendre les données entre 90 et 180 jours. Cependant, si on cherche à estimer la volatilité afin de calculer le prix d'exercice d'une option sur deux ans, il faut prendre en considérations les données des deux dernières années. Le prochain tableau illustre les résultats de ces calculs pour un titre.

Soit S_t le prix d'un titre au temps t . S_t est simulé pour une tendance $\mu = 1$ et d'une volatilité $\sigma = 0.5$.

Nombre de données (n)	Volatilité historique	Erreur standard
100	0.4329	0.0306
500	0.4861	0.0153
1 000	0.4984	0.0113
2 500	0.5065	0.0072
5 000	0.5001	0.0050
10 000	0.5003	0.0035

Tableau 3.2-Volatilité historique et erreur standard du titre S_t en fonction de n

Les résultats de tableau **3.2** montre que l'erreur standard diminue avec l'augmentation de nombre de données n et que la volatilité historique proche de celle fixée pour la simulation de S_t

Calcul de la volatilité implicite

Définition 3.1

Selon de la *Bruslerie (1988) [35]* : La volatilité implicite est le résultat d'une équation liant le prix au marché d'une option à ses déterminants dans le cadre d'un modèle d'évaluation donné. Elle peut s'interpréter comme étant une estimation contemporaine de la variabilité moyenne future de l'actif sous-jacent par le marché au cours de la vie de l'option. La volatilité implicite serait donc une estimation actuelle ayant un contenu prospectif, toujours selon de la *Bruslerie (1988) [35]*.

Modèle d'évaluation

Traditionnellement, le modèle de Black-Scholes a été utilisé pour estimer la volatilité implicite d'une option donnée. Comme nous l'avons déjà vu, les auteurs de ce modèle supposent que cette variable est constante durant la vie de l'option, ce qui n'est vrai dans la réalité. L'utilisation d'un modèle d'évaluation d'options pour estimer la volatilité implicite suppose le modèle connu et bien spécifié, c'est à dire que le prix que dégage le modèle est le prix de l'option sur le marché. Donc, la mesure de la volatilité implicite est extrêmement tributaire du choix du modèle d'évaluation d'options que nous choisissons pour l'estimer. Afin de minimiser l'erreur de l'estimation, il faudrait donc choisir un modèle qui reflète la réalité de notre actif sous-jacent.

Dans le modèle de Black-Scholes qu'on a vu à la section 1 du présent chapitre, la valeur d'une option call européenne est donnée par :

$$C(S, K, r, \tau, \sigma) = SN(d_1) - Ke^{-r\tau}N(d_2) \quad (1.2)$$

Où τ est le temps qui reste à l'option avant son échéance T

La valeur C est donnée en fonction de six paramètres, dont σ est le seul paramètre qui ne soit pas directement observable. Si on connaît σ on peut calculer le prix de l'option. Inversement si on connaît la valeur de l'option, on peut déduire la valeur de σ car $\frac{\partial C}{\partial \sigma} > 0$. Etant donnée que la valeur de l'option à l'instant t est comprise entre $\max(S - Ke^{-r\tau}, 0)$ et S , alors il existe une et une seule valeur de σ telle que la valeur théorique de l'option soit égale à la valeur réel sur le marché de cette même option. La valeur trouvée, est ce qu'on appelle la volatilité implicite.

A titre d'exemple, on suppose que la valeur d'un call sur un titre qui ne paye pas de dividendes soit 1.765 tel que

$$S_0 = 21, K = 20, r = 0.1 \text{ et } \tau = 0.25.$$

La volatilité implicite est la valeur de σ , telle que si on la remplace dans l'équation (1.2), on retrouve la valeur $C = 1.765$. Malheureusement on ne peut inverser l'équation (1.2) pour avoir une expression de σ en fonction de S_0, K, r, τ et C . Cependant, un procédé de recherche itératif, tel que l'algorithme de Raphson, peut être utilisé pour retrouver la valeur de la volatilité implicite σ . Par exemple, on commence en essayant la valeur $\sigma = 0.15$. Ce qui donne à C la valeur 1.624, ce qui est en dessous de la valeur visée de 1.765. Comme C

est une fonction croissante par rapport à σ , il nous faut prendre une valeur plus grande pour σ . On peut essayer comme deuxième valeur $\sigma = 0.25$ ce qui donne à C la valeur 1.927, ce qui est au-dessus de la valeur recherchée. Donc la valeur recherchée σ se trouve entre 0.15 et 0.25. Par un procédé itératif, on divise l'intervalle $[0.15, 0.25]$ en deux jusqu'à arriver à la valeur recherchée. Dans cet exemple, la volatilité implicite est 0.201 ou 20.1% par an.

Le modèle de Black-Scholes permet d'obtenir le niveau de volatilité implicite à partir de tous les autres paramètres, d'où son utilisation intensive sur les marchés financiers.

Cependant, le modèle de Black et Scholes ne permet pas de modéliser précisément le monde réel. L'expérience montre qu'en réalité la volatilité dépend du prix d'exercice et de la maturité donnant naissance à une surface de volatilité.

La formule de Black&Scholes ne donne un résultat explicite que lorsque les coefficients r et σ sont déterministes. Lorsque la volatilité est aléatoire, cela implique une incomplétude du marché et le prix de l'option est donné à ce moment là par une espérance conditionnelle.

Contrairement à l'hypothèse du modèle de Black & Scholes, la surface de volatilité n'est jamais plate. On pourra noter que :

- Souvent, pour une maturité donnée, la volatilité implicite par rapport au prix d'exercice (strike) a une forme de sourire (appelé le smile de volatilité)
- Le smile n'est souvent pas symétrique sur le marché des actions. Cela est dû au fait que les acteurs de marché sont plus sensibles au risque de baisse qu'au risque de hausse de

Le prix du titre (Le prix de l'action) S_t est modélisé par un mouvement brownien géométrique

$$S_t = s_0 e^{(\mu - \frac{\sigma^2}{2})t + \sigma W_t} \quad (1.3)$$

où W_t est un mouvement brownien standard et S_t satisfait l'équation différentielle stochastique

$$dS_t = \mu S_t dt + \sigma S_t dW_t, \quad S_0 = s_0.$$

Remarque 3.2

Le modèle contient deux paramètres inconnus : μ et σ . Il faudra donc les estimer. De plus, en pratique, on observe les données à des intervalles de longueurs fixes h . Par exemple, à toutes les 5 secondes, à tous les jours, toutes les semaines, etc.

Série de rendement

En général il n'est pas très intéressant de voir juste le prix d'un investissement. Du point de vue d'un investisseur, le rendement de l'investissement est beaucoup plus intéressant. Principalement à cause qu'un investisseur insiste plus sur le gain relatif réalisable, plutôt que sur le prix nominal de l'investissement, mais aussi parce que le rendement comme indice de changement du prix relatif permet des comparaisons entre compagnies, titres boursiers et monnaies. En plus, du fait que les rendements sont généralement stationnaires, une propriété que ne possèdent pas les prix actuels des titres. Dans le monde de la finance, le concept de rendement n'est pas défini de manière claire. Soit S_t le prix d'un titre au temps t . Le rendement à l'instant t peut être défini par :

Définition 3.2 (Le taux de rendement arithmétique)

$$X_{1,t} = \frac{S_t - S_{t-1}}{S_{t-1}}$$

ou par

Définition 3.3 (Le taux de rendement géométrique)

$$X_{2,t} = \ln \left(\frac{S_t}{S_{t-1}} \right)$$

Les deux rendements sont liés par la formule suivante, du moins pour les rendements quotidiens

$$X_{1,t} = e^{X_{2,t}} - 1 \approx X_{2,t}$$

Le taux de rendement géométrique est aussi dit rendement composé. Tout au long de ce chapitre, nous utiliserons le taux de rendement géométrique. Il sera notée par X_t et fera référence au rendement. Le taux de rendement géométrique est employé car il est le plus utilisé dans les diverses recherches, ce qui permet de comparer les résultats obtenus, et parce que le taux de rendement géométrique permet de relier les modèles en temps discret et ceux en temps continu.

Dans ce qui suit, on verra que μ et σ s'interprètent facilement en termes des rendements

$$X_j = X_j(h) = \ln(S_{jh}) - \ln(S_{(j-1)h}), \quad j \geq 1.$$

On commencera donc par trouver la loi de ces rendements.

Loi jointe des rendements

Proposition 3.1

Pour tout $h > 0$ fixé, les rendements :

$$X_j = \ln(S_{jh}) - \ln(S_{(j-1)h}), \quad j \geq 1$$

sont indépendants, identiquement distribués, et

$$X_j \sim N\left(\mu h - \frac{\sigma^2}{2}h, \sigma^2 h\right)$$

Preuve

Pour $1 \leq j \leq n$, on a

$$\begin{aligned} \frac{S_{jh}}{S_{(j-1)h}} &= \frac{s_0 \exp\left[\left(\mu - \frac{\sigma^2}{2}\right)jh + \sigma W_{jh}\right]}{s_0 \exp\left[\left(\mu - \frac{\sigma^2}{2}\right)(j-1)h + \sigma W_{(j-1)h}\right]} \\ &= \exp\left[\left(\mu - \frac{\sigma^2}{2}\right)h + \sigma\{W_{jh} - W_{(j-1)h}\}\right] \end{aligned}$$

En appliquant le logarithme de chaque côté,

$$\begin{aligned} X_j &= \ln(S_{jh}) - \ln(S_{(j-1)h}) \\ &= \left(\mu - \frac{\sigma^2}{2}\right)h + \sigma\{W_{jh} - W_{(j-1)h}\} \\ &\sim N\left(\mu h - \frac{\sigma^2}{2}h, \sigma^2 h\right) \end{aligned}$$

Les incréments $W_{jh} - W_{(j-1)h}$ sont indépendants, les X_j sont donc aussi indépendants.

Résultats empiriques

Une contrainte très importante du modèle de Black Scholes est que les rendements $X_j, 1 \leq j \leq n$ sont indépendants et de même loi gaussienne. On peut tester ces hypothèses.

-Test d'indépendance : Il existe des tests statistiques afin de vérifier si les rendements $X_j, 1 \leq j \leq n$, supposés de même loi, forment un bruit blanc, c'est-à-dire sont indépendants, que l'on connaisse ou non la loi des rendements. Pour cela, on utilise le test de *Box Ljung*. (si la quantité $Q = n(n+2) \sum_{k=1}^K \frac{\hat{\rho}_k^2}{n-k}$ de *Box Ljung* inférieure à $\chi_{(K)}^2(1-\alpha)$ on accepte l'hypothèse d'indépendance des rendements au seuil α tel que $\hat{\rho}_k$ = l'autocorrélation d'ordre k et $K = \frac{n}{5}$).

-Test de Normalité : Si les rendements $X_j, 1 \leq j \leq n$ sont indépendants, on peut tester l'hypothèse de loi gaussienne. Il existe plusieurs tests de normalité, le test le plus classique est de *Jarque et Bera* (si la quantité $S = \frac{n}{6}\beta_1 + \frac{n}{24}(\beta_2 - 3)^2$ de *Jarque et Bera* supérieure à $\chi_{(1-\alpha)}^2(2)$ on rejette l'hypothèse de normalité des résidus au seuil α tel que β_1 est le coefficient de Skewness et β_2 est le coefficient de Kurtosis).

Ou bien on accepte les deux hypothèses (d'indépendance et de la normalité) si la p-value est supérieur à α (α est fixée à 5%)

Nous simulons les rendements $X_j, 1 \leq j \leq n$, les paramètres du modèle sont fixés respectivement par $\mu = 1$ et $\sigma = 0.5$.

La sensibilité de ces résultats (ces hypothèses) par rapport au choix de h et de n est représentée dans le tableau suivant :

n	le test de la normalité		S	le test d'indépendance	
	h	p-value			p-value
100	1,00	0.1663	3.5876	1,00	0.321
	0.5	0.2109	3.1131	0.5	0.407
	0.01	0.4285	1.6948,	0.01	0.531
	0.001	0.5467	1.2079	0.001	0.601
	0.0001	0.7798	0.4975	0.0001	0.689
	0.00001	0.8458	0.335	0.00001	0.699
1000	1,00	0.3059	2.3692	1,00	0.332
	0.5	0.5062	1.3617	0.5	0.418
	0.01	0.5642	1.1446	0.01	0.543
	0.001	0.6654	0.8148	0.001	0.624
	0.0001	0.7382	0.607	0.0001	0.742
	0.00001	0.9382	0.1275	0.00001	0.769
10000	1,00	0.3877	1.8949	1,00	0.350
	0.5	0.5841	1.0753	0.5	0.421
	0.01	0.6668	0.8105	0.01	0.564
	0.001	0.7302	0.6289	0.001	0.666
	0.0001	0.9067	0.1958	0.0001	0.786
	0.00001	0.9558	0.0916	0.00001	0.801
100000	1,00	0.5246	1.2904	1,00	0.351
	0.5	0.5937	1.0426	0.5	0.529
	0.01	0.7321	0.6236	0.01	0.692
	0.001	0.7851	0.4839	0.001	0.798
	0.0001	0.9217	0.163	0.0001	0.912
	0.00001	0.9933	0.0134	0.00001	0.928

Tableau 3.3- *le test de la normalité et d'indépendance des rendements X_j*

D'après les résultats de tableau **3.3** nous constatons que plus le nombre de simulation n augmente (de 100 à 100 000) et le pas h diminue (de 1 à 0.00001), plus la $p - value$ pour les deux tests augmente (plus proche de 1), ce qui vient confirmer notre intuition. En effet, l'indépendance et la normalité des rendements X_j augmentent proportionnellement avec le nombre de simulation n et le contraire pour le pas h .

La Figure 1.8 illustre l'évolution d'une série en log-rendements.

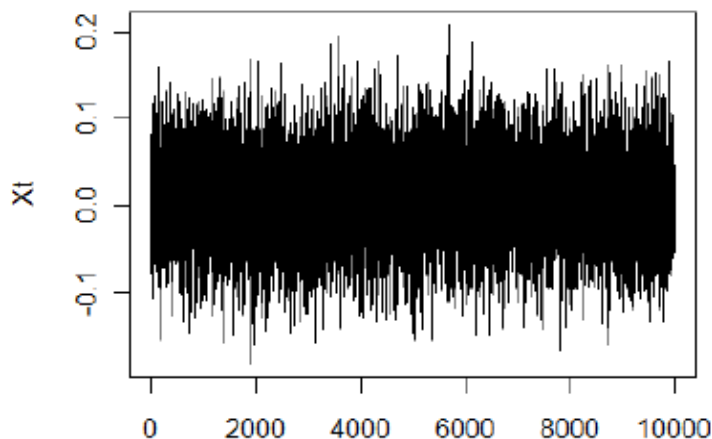


Figure 3.8-processus des rendements X_j simulé avec les paramètres de modèle $\mu = 1, \sigma = 0.5$ et pour $n = 10000$ et $h = 0.01$

Simulation du mouvement brownien géométrique à partir des rendements

Voici une application directe de la proposition **3.1**.

Pour simuler S_h, \dots, S_{nh} , on procède de la façon suivante :

–On génère des variables normales Z_1, \dots, Z_n , indépendantes, de moyenne 0 et de variance 1.

–Pour $1 \leq j \leq n$, on pose $X_j = \mu h - \frac{\sigma^2}{2}h + \sigma\sqrt{h}Z_j$.

–Pour $1 \leq j \leq n$, on pose $S_{jh} = s_0 \exp\left(\sum_{k=1}^j X_k\right)$.

La Figure **3.9** présente la simulation d'un mouvement brownien géométrique S_t à partir des rendements X_j , pour $n = 1000$ et $h = 0.1$ et pour des paramètres du modèle $\mu = 1, \sigma = 0.5$.

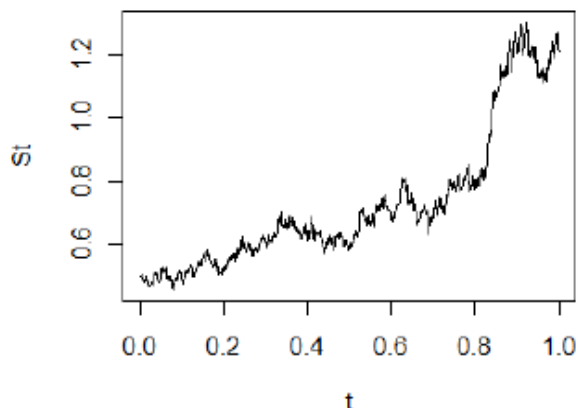


Figure 3.9-*Simulation d'un mouvement brownien géométrique à partir des rendements*

Nous nous intéressons principalement au modèle de Black & Scholes, alors le paragraphe qui suit développe l'estimation des paramètres de ce modèle brownien géométrique où dans sa première section, nous présentons la méthode historique basée sur la fonction de vraisemblance (la méthode du maximum de vraisemblance) suivie par deux autres méthodes à la 2 section.

3.3 Estimation des paramètres du modèle de Black-Scholes

3.3.1 La méthode historique basées sur la fonction de vraisemblance

La méthode du maximum de vraisemblance sera utilisée pour estimer les paramètres μ et σ du modèle brownien géométrique (1.3) utilisé pour les prix. On peut montrer qu'elle donne presque toujours les résultats les plus précis. On aurait aussi pu utiliser la méthode des moments.

La méthode des moments : apparier les moments de la population et ceux de la distribution pour obtenir des équations ; il faut évidemment apparier autant de moments qu'il y a de paramètres. Dans notre cas, il s'agit donc d'apparier les deux premiers moments de population,

soit $\bar{x} = \frac{\sum_{i=1}^n x_i}{n}$ et $s^2 = \frac{\sum_{i=1}^n (x_i)^2}{n}$ à ceux de la distribution de X_j (sans égard au j , puisqu'on a fait l'hypothèse "i.i.d" plus tôt). On obtient alors les paramètres μ et σ^2 en fonction des moments de la population.

Dans le cas gaussien, les deux donnent les mêmes résultats. Par contre, les résultats diffèrent en général.

Principe du maximum de vraisemblance

Soit X_1, \dots, X_n des observations dont la loi jointe possède une densité $f(\cdot; \theta)$, où le paramètre θ appartient à un ensemble $\Omega \in \mathbb{R}^k$. La fonction de vraisemblance $L(\theta)$ est définie par

$$L(\theta) = f(x_1, \dots, x_n; \theta) \quad (2.1)$$

La méthode du maximum de vraisemblance (MLE) est de prendre comme estimation de θ , la valeur $\hat{\theta}$ qui maximise la fonction de vraisemblance L , c'est-à-dire

$$\hat{\theta} = \arg \max_{\tilde{\theta} \in \Omega} L(\tilde{\theta})$$

où Ω définit les contraintes que $\tilde{\theta}$ doit respecter, le cas échéant.

Remarque 3.3

Comme la fonction de vraisemblance peut s'exprimer sous forme de produit, il est souvent plus facile de maximiser la log-vraisemblance (logarithme de la fonction de vraisemblance). On obtient alors

$$\hat{\theta} = \arg \max_{\tilde{\theta} \in \Omega} \ln(L(\tilde{\theta}))$$

De façon équivalente, on peut minimiser $-\ln(L)$. Dans ce cas, on a

$$\hat{\theta} = \arg \min_{\tilde{\theta} \in \Omega} -\ln(L(\tilde{\theta}))$$

Il faut se rappeler que les points optimaux d'une fonction sur un ensemble sont ceux dont

le gradient est nul. Sinon il faut regarder sur la frontière de l'ensemble.

Si $\hat{\theta}$ est à l'intérieur de Ω , alors on doit avoir, pour tout $1 \leq i \leq k$,

$$\frac{\partial}{\partial \theta_i} L(\theta) \Big|_{\theta=\hat{\theta}} = 0$$

ou encore

$$\frac{\partial}{\partial \theta_i} \ln(L(\theta)) \Big|_{\theta=\hat{\theta}} = 0$$

Estimation des paramètres

Utilisons la méthode du maximum de vraisemblance pour estimer les paramètres de notre modèle **(1.3)**. Nous avons les données suivantes :

Prix $S_0 = s_0, S_h = s_1, \dots, S_{jh} = s_j, \dots, S_{nh} = s_n, \quad 1 \leq j \leq n,$

d'où on extrait les rendements $x_j = \ln(s_j) - \ln(s_{j-1}), \quad 1 \leq j \leq n.$

Proposition 3.2

Les estimations de μ et σ , par la méthode du maximum de vraisemblance, dépendent seulement des rendements x_1, \dots, x_n , et sont données par

$$\hat{\mu} = \frac{\bar{x}}{h} + \frac{s_x^2}{2h} \quad \text{et} \quad \hat{\sigma} = \frac{s_x}{\sqrt{h}}$$

où

$$\bar{x} = \frac{1}{n} \sum_{j=1}^n x_j \quad \text{et} \quad s_x^2 = \frac{1}{n} \sum_{j=1}^n (x_j - \bar{x})^2$$

Nous avons utilisé s et s_x pour, respectivement, le 2^e moment et le 2^e moment centré.

Preuve

D'après la proposition 1, les rendements forment un bruit blanc gaussien (i.e des observations indépendantes de même loi). La fonction de densité jointe des rendements, étant donné S_0 , est le produit des densités gaussiennes

$$L(\mu, \sigma) = \prod_{j=1}^n \frac{1}{\sqrt{2\pi\sigma^2 h}} \exp \left[-\frac{(x_j - h(\mu - \frac{\sigma^2}{2}))^2}{2\sigma^2 h} \right]$$

que l'on veut maximiser, pour $\mu \in \mathbb{R}$ et $\sigma > 0$, ce qui revient à minimiser

$$\begin{aligned} g(\mu, \sigma) &= -\ln(L(\mu, \sigma)) \\ &= \frac{n}{2} \ln(2\pi h) + n \ln(\sigma) + \frac{1}{2\sigma^2 h} \sum_{j=1}^n \left(x_j - h\mu + \frac{\sigma^2 h}{2}\right)^2 \end{aligned}$$

On voit que

$$\frac{\partial}{\partial \mu} g = \frac{n}{\sigma^2} \left(\mu h - \frac{\sigma^2 h}{2} - \bar{x}\right)$$

$$\frac{\partial}{\partial \sigma} g = \frac{n}{\sigma} - \frac{1}{\sigma^3 h} \sum_{j=1}^n \left(x_j - h\mu + \frac{\sigma^2 h}{2}\right)^2 - \frac{n}{\sigma} \left(\mu h - \frac{\sigma^2 h}{2} - \bar{x}\right)$$

Les deux dérivées sont nulles pour $\hat{\mu}$ et $\hat{\sigma}$ donnés par

$$\begin{aligned} \hat{\mu} &= \frac{\bar{x}}{h} + \frac{(\hat{\sigma})^2}{2} = \frac{\bar{x}}{h} + \frac{s_x^2}{2h} \\ \hat{\sigma} &= \sqrt{\frac{1}{nh} \sum_{j=1}^n (x_j - \bar{x})^2} = \frac{s_x}{\sqrt{h}} \end{aligned}$$

Remarque 3.4

En pratique, on utilise surtout

$$\frac{1}{(n-1)h} \sum_{j=1}^n (x_j - \bar{x})^2$$

pour estimer σ^2 . La raison est que cet estimateur est très proche de $\frac{s_x^2}{h}$ et c'est un estimateur sans biais de σ^2 (voir remarque suivante). Par exemple, avec logiciel R, la fonction `std(x)`

donne $\sqrt{\frac{1}{(n-1)} \sum_{j=1}^n (x_j - \bar{x})^2}$ est non pas $\sqrt{\frac{1}{n} \sum_{j=1}^n (x_j - \bar{x})^2}$.

Remarque 3.5

Estimation sans biais : Supposons qu'on ait un échantillon X_1, X_2, \dots, X_n de la distribution de la variable aléatoire X . On utilise l'estimateur $u(X_1, X_2, \dots, X_n)$ pour un paramètre θ de la distribution. Alors l'estimateur $u(X_1, X_2, \dots, X_n)$ est dit sans biais si

$$E(u(X_1, X_2, \dots, X_n)) = \theta$$

Ci-haut, l'estimateur $\hat{\sigma} = \frac{s_{\sigma}}{\sqrt{h}}$ est un estimateur biaisé, chose qui se vérifie facilement.

Remarque 3.6

Lorsque l'on utilisera la méthode du maximum de vraisemblance dans le cas non gaussien, on trouvera rarement des expressions analytiques.

Il faudra donc minimiser ou maximiser "numériquement". Dans ce cas, il faut faire très attention aux contraintes sur les paramètres à estimer.

Par exemple, dans notre cas $\sigma > 0$.

Résultats empiriques

Pour estimer les paramètres du modèle de Black-Scholes en appliquant cette méthode historique, il suffit de suivre les étapes suivantes :

-Nous simulons les prix S_j $1 \leq j \leq n$ pour les vrais paramètres $\mu = 1$, $\sigma = 0.5$ et $s_0 = 1$, car on sait bien que les prix sont modélisés par un mouvement brownien géométrique dont sa solution est connue.

-A partir de ces données, nous calculons les rendements $X_j = \log\left(\frac{s_j}{s_{j-1}}\right)$.

-Nous vérifions la normalité et l'indépendance des rendements X_j (nous avons vérifié ces deux hypothèses à la section précédente pour différentes valeurs de n).

Les résultats sont présentés dans le tableau suivant, on fait varier le nombre de simulation n ($n = 100, 1000, 5000, 10\ 000, 100\ 000, 1\ 000\ 000$).

Le nombre de simulation n	L'estimation de μ	Intervalle de confiance pour μ	L'estimation de σ	Intervalle de confiance pour σ
100	0.9659	[0.8742 1.0576]	0.4654	[0.3640 0.6452]
1 000	0.9790	[0.9496 1.0084]	0.4734	[0.4351 0.5188]
5 000	0.9924	[0.9784 1.0065]	0.5069	[0.4877 0.5275]
10 000	0.9991	[0.9893 1.0089]	0.5013	[0.4878 0.5156]
100 000	1.0010	[0.9979 1.0041]	0.5007	[0.4964 0.5052]
1 000 000	1.0002	[0.9992 1.0012]	0.5002	[0.4988 0.5016]

Tableau 3.4- les résultats d'estimation de μ et σ par la méthode de vraisemblance

D'un point de vue global, la méthode historique qui est une méthode du maximum de vraisemblance basé sur la normalité des rendements offre des bons résultats. Nous constatons que les valeurs de tous les estimateurs se trouvent dans la fourchette des régions de confiance de leurs paramètres respectifs, et que la longueur des intervalles de confiances de chaque paramètre diminue respectivement avec l'augmentation de nombre de simulation n , ce qui confirme que les estimations de μ et σ par la méthode historique, dépendent seulement de n observations des rendements (X_1, \dots, X_n) .

Précision des estimations

On aimerait bien savoir quelle est la loi des erreurs d'estimation. Autrement dit, comment se comportent $\sqrt{n}(\hat{\mu} - \mu)$ et $\sqrt{n}(\hat{\sigma} - \sigma)$.

Lorsque la méthode du maximum de vraisemblance est appliquée à des données indépendantes et identiquement distribuées, comme c'est le cas ici, on sait que la loi jointe des erreurs est approximativement une loi normale multivariée.

Loi gaussienne multivariée : Soit $\mu \in \mathbb{R}^p$ et $A \in \mathbb{R}^{p \times p}$, où la matrice A est symétrique ($A = A^T$), définie positive ($x^T A x > 0$, pour tout $x \in \mathbb{R}^p, x \neq 0$).

Un vecteur aléatoire $X = (X_1, X_2, \dots, X_p)^T$ est distribuée selon une loi gaussienne de paramètres $\mu \in \mathbb{R}^p$ et $A \in \mathbb{R}^{p \times p}$, que l'on note par $X \sim N_p(\mu, A)$, si la densité de X est donnée

par

$$f(x) = \frac{1}{(2\pi)^{\frac{p}{2}} |A|^{\frac{1}{2}}} e^{-\frac{1}{2}(x-\mu)^T A^{-1}(x-\mu)}, \quad x \in \mathbb{R}^p$$

où $|A| = \det(A)$.

Représentation d'un vecteur aléatoire gaussien

Proposition 3.3

Si B une matrice $p \times d$, $\mu \in \mathbb{R}^d$ et si Y est un vecteur composé de p variables de loi $N(0, 1)$, c'est-à-dire $Y \sim N_p(0, I)$, alors

$$\mu + BY \sim N_d(\mu, BB^T).$$

Remarque 3.7

La proposition 3.3 permet de générer un vecteur gaussien $X \sim N_p(\mu, A)$, à partir de p lois gaussiennes centrées réduites. En effet, si A est une matrice $p \times p$ symétrique et définie positive, on peut trouver une matrice R triangulaire supérieure telle que $R^T R = A$ (décomposition de Cholesky). On pose alors $B = R^T$. Dans ce cas

$$X = \mu + R^T Y \sim N_p(\mu, A),$$

si $Y \sim N_p(0, I)$.

Proposition 3.4

Supposons que $X \sim N_p(\mu, A)$. Alors $E(X) = \mu$, et pour $1 \leq i, j \leq p$, on a

$$(E\{(x - \mu)(x - \mu)^T\})_{ij} = \text{cov}(X_i, X_j) = A_{ij}.$$

De plus, si $\lambda \in \mathbb{R}^p$, alors

$$\lambda^T X = \sum_{j=1}^p \lambda_j X_j \sim N(\lambda^T \mu, \lambda^T A \lambda)$$

Remarque 3.8

La matrice A est dite matrice de covariance de X . La loi d'un vecteur aléatoire gaussien est donc uniquement déterminée par sa moyenne et sa matrice de covariance.

Proposition 3.5

Pour n assez grand, on a

$$\begin{pmatrix} \sqrt{n}(\hat{\mu} - \mu) \\ \sqrt{n}(\hat{\sigma} - \sigma) \end{pmatrix} \approx N_2(0, V)$$

où $V = \frac{\sigma^2}{2} \begin{pmatrix} \frac{2}{h} + \sigma^2 & \sigma \\ \sigma & 1 \end{pmatrix}$ qui est estimée par

$$\hat{V} = \frac{s_x^2}{2h} \begin{pmatrix} \frac{2}{h} + \frac{s_x^2}{h} & \frac{s_x}{\sqrt{h}} \\ \frac{s_x}{\sqrt{h}} & 1 \end{pmatrix}$$

En particulier, on en déduit que

$$\sqrt{n}(\hat{\mu} - \mu) / \sqrt{\hat{V}_{11}} \approx N(0, 1) \quad \text{et} \quad \sqrt{n}(\hat{\sigma} - \sigma) / \sqrt{\hat{V}_{22}} \approx N(0, 1).$$

Preuve.

On a, $\sqrt{n}(\bar{x} - \mu + \frac{\sigma^2 h}{2}) \approx N(0, \sigma^2 h)$, et $\sqrt{n}(s_x - \sigma\sqrt{h}) \approx N(0, \frac{\sigma^2 h}{2})$. De plus, les deux statistiques sont asymptotiquement indépendantes.

Donc $\sqrt{n}(\frac{s_x}{\sqrt{h}} - \sigma) \approx N(0, \frac{\sigma^2}{2})$

De plus,

$$\sqrt{n}(\hat{\mu} - \mu) = \frac{1}{h} \sqrt{n}(\bar{x} - \mu h + \frac{\sigma^2 h}{2}) + \frac{1}{2h} \sqrt{n}(s_x^2 - \sigma^2 h).$$

On voit donc que $\sqrt{n}(\hat{\mu} - \mu) \approx N(0, \frac{\sigma^2}{h} + \frac{\sigma^4}{2})$, et la covariance limite entre les deux statistiques est $\frac{\sigma^3}{2}$.

Résultats empiriques

n	p-value pour μ	p-value pour σ
100	0.5302	0.5817
1 000	0.5774	0.8067
10 000	0.9518	0.9691
100 000	0.9730	0.9762

Tableau 3.5-*Précision des estimations par rapport au choix de n pour les paramètres μ et σ*

Dans le tableau **3.5** nous avons présenté les résultats de la précision des estimations par la méthode historique de paramètres μ et σ . Nous avons testé la normalité des estimateurs à l'aide de test de **Kolmogorov-Smirnov**, remarquons que la **p-value** pour μ est de 0.5302 pour $n = 100$ et de 0.9730 presque 1 pour $n = 100000$ et de même pour σ la **p-value** = 0.5817 pour $n = 100$ et de 0.9762 pour $n = 100000$, donc la proposition **3.5** a été vérifié. La loi des erreurs d'estimation pour μ et σ est asymptotiquement normale centrée est de variance V .

La difficulté lors d'utilisation des approches basées sur l'estimateur du maximum de vraisemblance (MLE) est que la fonction de densité de transition du processus de diffusion sous-jacent est souvent inconnue.

Une réponse à ce problème consiste à essayer de calculer une approximation numérique de la fonction de densité de transition. Dans notre travail nous allons présenter une approche (présentée par Hurn et Lindsay (1997) et étendu par McDonald et Sandal (1998)) qui consiste à contourner l'estimation de la fonction de densité de transition entièrement pour estimer simultanément les paramètres du modèle. Cette approche implique la simulation numérique des ensembles de données, en utilisant des schèmes d'approximations et en correspondant ces distributions simulées aux données expérimentales en utilisant la statistique de χ^2 de la qualité de l'ajustement (the χ^2 goodness-of-fit statistic) présentée par Hurn et Lindsay (1997) et la statistique du Kolmogorov-Smirnov présentée par McDonald et Sandal (1998).

3.3.2 L'estimation des paramètres des EDS utilisant deux fonctions critères

L'estimation simultanée des paramètres des EDS linéaires et non linéaires est un sujet intéressant, notamment pour évaluer l'importance relative des composantes déterministes et stochastiques du processus étudié. Pour cela nous avons concentré sur deux méthodes qui estime simultanément les paramètres des EDS, nous avons d'abord utilisé une fonction critère fondée sur la statistique du chi-deux, puis examiné l'utilité d'une fonction critère basée sur la statistique de Kolmogorov-Smirnov.

Nous nous intéressons principalement au modèle de Black & Scholes, pour cela nous concentrons dans cette section sur l'équation différentielle stochastique linéaire d'Itô suivante :

$$dX_t = \mu X_t dt + \sigma X_t dW_t \quad (2.2)$$

Où sa solution est donnée par

$$X_t = x_0 \exp \left(\left(\mu - \frac{1}{2} \sigma^2 \right) t + \sigma W_t \right)$$

L'Estimation des paramètres à l'aide d'une fonction critère basée sur la Statistique de Chi-deux

Les moindres carrés classiques et le maximum de vraisemblance sont insuffisants pour l'estimation simultanée des paramètres, alors que cette approche vise à fournir une méthode numérique, basée sur le principe du maximum de vraisemblance, pour l'estimation simultanée des deux paramètres de l'EDS (2.2). Cette méthode consiste à faire correspondre la distribution des données expérimentales avec un ensemble de données simulées générées par une expérience de Monte Carlo. L'adéquation entre les deux distributions est évaluée au moyen d'une statistique de chi-deux de la qualité d'ajustement (the χ^2 goodness-of-fit statistic) conduisant à une fonction de confiance calculée à partir d'une fonction gamma incomplète. Un algorithme d'optimisation numérique optimise alors le choix des paramètres pour maximiser cette fonction. Cette disposition est potentiellement importante en pratique depuis dans beaucoup d'applications, telle que la modélisation du comportement de cours des actions [11].

L'estimation des paramètres

La méthode que nous adoptons pour estimer simultanément les coefficients de la dérive et de diffusion exige un ensemble de m observations pour chaque pas de temps t_k ($1 \leq k \leq n$). L'hypothèse maintenue est que les données observées ont été générées par la même équation différentielle stochastique d'Itô **(2.2)** pour m réalisations du processus stochastique. Au cœur de la méthode d'estimation est une expérience de Monte Carlo qui génère un ensemble de données simulées, pour la meilleure estimation actuelle des paramètres inconnus et m processus de Wiener indépendants. Un algorithme numérique optimise la différence entre les données expérimentales et les données simulées pour les divers choix de paramètres du modèle. Nous avons déjà noté que pour produire des données simulées les m réalisations du processus stochastique sont nécessaires. Notez, cependant, qu'une fois généré ces processus sont stockés et utilisés pour évaluer la fonction cible pour chaque nouveau choix de paramètres du modèle par un algorithme d'optimisation.

En essence, les paramètres du modèle doivent être choisis afin de maximiser la qualité d'ajustement entre les deux distributions.

La construction de la fonction cible doit inclure les caractéristiques de la distribution entière et non seulement ceux de la moyenne. lorsque $t = t_k$ ($1 \leq k \leq n$), c_k mesure la confiance dans l'hypothèse que les données simulées et expérimentales sont tirées de la même distribution. Un niveau de confiance pour l'hypothèse que les deux ensembles de données sont tirés de la même distribution est mesuré par $\prod_{k=1}^n c_k$.

Soit X est $N(0, 1)$, alors X^2 est $\Gamma(1/2, 1/2)$ (gamma distribué). Supposons maintenant que X_1, \dots, X_r sont des r variables aléatoires indépendantes de lois $N(0, 1)$, alors la variable aléatoire $Y = X_1^2 + X_2^2 + \dots + X_r^2$ est $\Gamma(r/2, 1/2)$ et, particulièrement

$$p(Y \leq \chi^2) = \frac{1}{2^{\frac{r}{2}} \Gamma(r/2)} \int_0^{\chi^2} x^{\frac{r}{2}-1} e^{-\frac{x}{2}} dx = \frac{1}{\Gamma(r/2)} \int_0^{\frac{\chi^2}{2}} y^{\frac{r}{2}-1} e^{-y} dy$$

Soit y_1, \dots, y_m la simulation du m valeurs expérimentales Y_1, \dots, Y_m à $t = t_k$. Le test de chi-deux est largement acceptée la statistique pour tester l'hypothèse nulle, à savoir, que les deux ensembles de données sont tirés de la même distribution. Chaque membre des deux ensembles de données est assigné (attribué) à l'une des $r+1$ cellules. Soient n_j et N_j le nombre

d'éléments dans la cellule j pour les données simulées et expérimentales respectivement, alors la statistique de chi-deux pour les données au moment t_k a la valeur

$$\chi_k^2 = \sum_{j=1}^{r+1} \frac{(N_j - n_j)^2}{N_j + n_j}$$

Une valeur élevée de χ_k^2 indique que l'hypothèse nulle est peu probable, c'est-à-dire il y a une forte probabilité que les deux distributions à t_k sont différentes alors qu'une petite valeur soutient l'hypothèse nulle (une grande valeur de χ_k^2 indique que l'hypothèse nulle est fortement rejetée). Depuis $\sum_{j=1}^{r+1} N_j = \sum_{j=1}^{r+1} n_j = m$ alors seulement r du N_s et n_s sont indépendants. Le test compte sur le fait que la statistique $X_j(N_j - n_j)/\sqrt{N_j + n_j}$ est approximée vers la $N(0, 1)$ à condition que $N_j \geq 5$ et le nombre de cellules est grand ou le nombre d'éléments dans chaque cellule est grand. Notre confiance en hypothèse que les données simulées et expérimentales sont compatibles avec la même distribution à $t = t_k$ est

$$c_k = 1 - p(Y \leq \chi_k^2) = \frac{1}{\Gamma(r/2)} \int_{\frac{\chi_k^2}{2}}^{\infty} y^{\frac{r}{2}-1} e^{-y} dy$$

Cette idée est maintenant prolongée (étendue) pour tous l'ensemble de n points et mène à la fonction cible

$$\Psi = \prod_{k=1}^n c_k = \prod_{k=1}^n \left(\frac{1}{\Gamma(r/2)} \int_{\frac{\chi_k^2}{2}}^{\infty} y^{\frac{r}{2}-1} e^{-y} dy \right)$$

Cette fonction doit être maximisée par le choix des paramètres du modèle.

En fait, c'est beaucoup plus pratique de minimiser

$$\phi = -\log \Psi = -\sum_{k=1}^n \log c_k$$

Résultats empiriques

Les tableaux suivants présentent les résultats des expériences visant à tester l'efficacité de notre procédure d'estimation. Pour chaque expérience trois paramètres doivent être fixés, à

savoir, le nombre de données ($n = 50, 200, 500$), le nombre d'observations pour chaque t_k ($m = 20, 50, 100$) et le nombre de cellules.

Les tableaux **3.6** et **3.7** donnent les résultats d'estimation pour 500 réplifications et les tableaux **3.8** et **3.9** pour 1000 réplifications.

Dans notre travail, chaque cellule ait contenu le même nombre de valeurs expérimentales, 5 éléments au premier essai (les tableaux 3.6 et 3.8) et 10 éléments au deuxième essai (les tableaux 3.7 et 3.9), en fait, la statistique du chi-deux ne devrait pas être utilisée lorsque certaines cellules contiennent moins de 5 éléments en moyenne.

Pour générer les données observées, utilisant la solution exacte de notre EDS, avec les valeurs actuels des coefficients $\mu = 1$ et $\sigma = 0.5$, ont été simulé sur l'intervalle 0 à 1 et pour $x_0 = 0.5$. Ces données ont été alors utilisées comme des données expérimentales, tel que n représente le nombre de données expérimentale c'est-à-dire la taille de l'échantillon. Le processus a été répété pour chaque permutation de m et fréquence de cellule. Les paramètres de la dérive et de diffusion du l'EDS produisant ces données ont été alors estimés 500 fois au premier essai et 1000 fois au deuxième essai avec des différentes réalisations des ensembles du m processus stochastiques exigés. Les moyennes et les écart-types de μ et σ obtenus à partir de ces reproductions, sont présentés au tableaux 3.6 à 3.9 et soutiennent un certain nombre de conclusions, la méthode de simplex de Nelder et Mead [24] a été alors utilisé pour estimer μ et σ utilisant au départ les valeurs de $\mu = 1.3$ et $\sigma = 0.4$

Le nombre de donnes n	Le nombre de simulations m	La moyenne des paramètres	L'écart type des paramètres
50	20	$\mu=1.091$ $\sigma=0.539$	$\mu=0.095$ $\sigma=0.101$
	50	$\mu=1.021$ $\sigma=0.527$	$\mu=0.081$ $\sigma=0.098$
	100	$\mu=0.991$ $\sigma=0.511$	$\mu=0.052$ $\sigma=0.045$
200	20	$\mu=1.101$ $\sigma=0.543$	$\mu=0.092$ $\sigma=0.099$
	50	$\mu=1.035$ $\sigma=0.531$	$\mu=0.072$ $\sigma=0.097$
	100	$\mu=1.009$ $\sigma=0.517$	$\mu=0.049$ $\sigma=0.053$
500	20	$\mu=1.113$ $\sigma=0.551$	$\mu=0.090$ $\sigma=0.091$
	50	$\mu=1.087$ $\sigma=0.526$	$\mu=0.054$ $\sigma=0.051$
	100	$\mu=1.002$ $\sigma=0.511$	$\mu=0.046$ $\sigma=0.041$

Tableau 3.6- 500 réplifications pour 5 éléments dans chaque cellule

Le nombre de donnes n	Le nombre de simulations m	La moyenne des paramètres	L'écart type des paramètres
50	20	$\mu=1.102$ $\sigma=0.482$	$\mu=0.185$ $\sigma=0.156$
	50	$\mu=1.054$ $\sigma=0.489$	$\mu=0.099$ $\sigma=0.098$
	100	$\mu=0.987$ $\sigma=0.512$	$\mu=0.064$ $\sigma=0.051$
200	20	$\mu=1.113$ $\sigma=0.483$	$\mu=0.191$ $\sigma=0.160$
	50	$\mu=1.012$ $\sigma=0.520$	$\mu=0.091$ $\sigma=0.103$
	100	$\mu=0.991$ $\sigma=0.514$	$\mu=0.068$ $\sigma=0.057$
500	20	$\mu=1.091$ $\sigma=0.487$	$\mu=0.202$ $\sigma=0.169$
	50	$\mu=1.054$ $\sigma=0.525$	$\mu=0.063$ $\sigma=0.046$
	100	$\mu=1.004$ $\sigma=0.507$	$\mu=0.053$ $\sigma=0.042$

Tableau 3.7- 500 répliques pour 10 éléments dans chaque cellule

Le nombre de donnes n	Le nombre de simulations m	La moyenne des paramètres	L'écart type des paramètres
50	20	$\mu=1.101$ $\sigma=0.479$	$\mu=0.093$ $\sigma=0.112$
	50	$\mu=1.031$ $\sigma=0.521$	$\mu=0.075$ $\sigma=0.085$
	100	$\mu=0.991$ $\sigma=0.509$	$\mu=0.050$ $\sigma=0.048$
200	20	$\mu=1.098$ $\sigma=0.539$	$\mu=0.087$ $\sigma=0.10$
	50	$\mu=1.034$ $\sigma=0.521$	$\mu=0.071$ $\sigma=0.084$
	100	$\mu=1.007$ $\sigma=0.509$	$\mu=0.042$ $\sigma=0.054$
500	20	$\mu=1.098$ $\sigma=0.523$	$\mu=0.087$ $\sigma=0.090$
	50	$\mu=1.037$ $\sigma=0.514$	$\mu=0.051$ $\sigma=0.049$
	100	$\mu=1.012$ $\sigma=0.502$	$\mu=0.047$ $\sigma=0.039$

Tableau 3.8- 1000 réplifications pour 5 éléments dans chaque cellule

Le nombre de données n	Le nombre de simulations m	La moyenne des paramètres	L'écart type des paramètres
50	20	$\mu=1.103$ $\sigma=0.532$	$\mu=0.179$ $\sigma=0.153$
	50	$\mu=1.064$ $\sigma=0.491$	$\mu=0.093$ $\sigma=0.089$
	100	$\mu=1.020$ $\sigma=0.508$	$\mu=0.051$ $\sigma=0.049$
200	20	$\mu=1.126$ $\sigma=0.484$	$\mu=0.189$ $\sigma=0.159$
	50	$\mu=1.009$ $\sigma=0.519$	$\mu=0.071$ $\sigma=0.093$
	100	$\mu=0.993$ $\sigma=0.507$	$\mu=0.054$ $\sigma=0.050$
500	20	$\mu=1.258$ $\sigma=0.561$	$\mu=0.197$ $\sigma=0.198$
	50	$\mu=1.048$ $\sigma=0.520$	$\mu=0.061$ $\sigma=0.068$
	100	$\mu=1.002$ $\sigma=0.509$	$\mu=0.044$ $\sigma=0.042$

Tableau 3.9- 1000 réplifications pour 10 éléments dans chaque cellule

Les résultats présentés dans les 4 tableaux précédents montrent qu'il y a une légère amélioration entre les 500 réplifications et les 1000 réplifications, cependant, que pour les 1000 on a que doublé le temps d'exécution

Le nombre de cellules utilisées dans la comparaison des données expérimentales et les simulations de Monte Carlo semble d'être critique pour l'estimation des coefficients. Même lorsque le nombre de données de Monte Carlo, m , dans l'expérience est de 20, l'utilisation de 4 cellules au lieu de 2 réduit (de façon spectaculaire dans certains cas) les écart-types des coefficients. Pendant que l'effet est moins prononcé si on augmente le nombre de données de Monte Carlo, il reste vrai que les coefficients de la dérive et de diffusion sont résolus plus précisément lorsque en augmentant le nombre des cellules.

Pour une taille d'échantillon donnée, l'augmentation de nombre d'essais, m , améliore à la fois les estimations de paramètre et leurs erreurs standards. Cependant, la plupart de cette

amélioration se produit (arrive) lorsque m est augmenté de 20 à 50.

L'Augmentation de nombre de données, n , quand les données se réfèrent aux mêmes intervalles de temps de production peu ou pas d'amélioration des résultats d'estimation. Il est clair que 500 points sont à peine meilleurs que 50 points sur le même intervalle, mais exige 10 fois plus de calcul.

McDonald et Sandal (1998) ont noté que le test de χ^2 de la qualité d'ajustement est seulement valide asymptotiquement et donc un grand nombre d'ensembles de données sont nécessaires pour que cette méthode est fiable. Ainsi, ils ont prolongé (étendu) la méthode du Hurn et Lindsay(1997) en utilisant la distribution asymptotique de la statistique du Kolmogorov-Smirnov à deux échantillons comme une fonction objective.

L'estimation des paramètres utilisant une fonction critère basée sur la Statistique de Kolmogorov-Smirnov

La statistique de Kolmogorov-Smirnov adaptée à un problème à deux échantillons fournit une base pour une autre méthode de la qualité d'ajustement (goodness-of-fit) pour l'estimation des paramètre de l'EDS (2.2). Comme pour le test de chi-deux le test de la qualité d'ajustement de Kolmogorov-Smirnov à deux échantillons est utilisé pour comparer les fonctions de distribution empiriques des deux échantillons. Dans notre travail l'un de ces échantillons est produit comme si on l'observe à partir d'une EDS entièrement indiqué et l'autre est généré de la même EDS, mais en prenant en considération que les coefficients sont inconnus. Dans la pratique l'estimation de paramètre exige l'un des échantillons d'être observé et l'autre d'être généré par une EDS qui est utilisée pour modéliser le processus des données .

Pour deux échantillons aléatoires de tailles m et n , leurs statistiques d'ordre sont respectivement : $X(1), X(2), \dots, X(m)$ et $Y(1), Y(2), \dots, Y(n)$

Leurs fonctions de distribution empiriques respectivement , désigné par $S_m(x)$ et $S_n(x)$, sont définis comme suit :

$$S_m(x) = \begin{cases} 0 & \text{si } x < x(1) \\ \frac{k}{m} & \text{si } x(k) \leq x < x(k+1) \text{ pour } k = 1, 2, \dots, m-1 \\ 1 & \text{si } x \geq x(m) \end{cases}$$

$$S_n(x) = \left\{ \begin{array}{ll} 0 & \text{si } x < y(1) \\ \frac{k}{n} \text{ si } y(k) \leq x < y(k+1) & \text{pour } k = 1, 2, \dots, n-1 \\ 1 & \text{si } x \geq y(n) \end{array} \right\}$$

Le test statistique de Kolmogorov-Smirnov à deux échantillons :

$$D_{m,n}^* = \max_x |S_m(x) - S_n(x)|$$

C'est le maximum en valeur absolue de la différence entre les deux distributions empiriques. Cette statistique peut être utilisé pour tester l'hypothèse (nulle) que les distributions de la population sont identiques et, donc, que les deux échantillons ont été prélevés de la même population.

La statistique de Kolmogorov-Smirnov à deux échantillons a une distribution asymptotique nulle donnée par

$$\lim_{m,n \rightarrow \infty} p \left(\sqrt{\frac{mn}{m+n}} D_{m,n} \leq D \right) = L(D) \quad (2.3)$$

Où

$$L(D) = 1 - 2 \sum_{i=1}^{\infty} (-1)^{i-1} e^{-2i^2 D^2} \quad (2.4)$$

Une grande valeur de D , et donc une petite valeur de $L(D)$, indique que l'hypothèse nulle est peu probable d'être vrai, tandis que des petites valeurs de D soutiennent l'hypothèse nulle. Dans notre travail nous sommes concernés aux données répliquées des séries temporelles. Cela offre la possibilité d'évaluer la statistique de Kolmogorov-Smirnov pour chaque période de temps, D_t . Parce que nous sommes face à un processus stochastique que nous assumons d'être modélisé d'une manière adéquate par une équation de la forme générale de l'équation (2.2), nous devons estimer les paramètres de la dérive et de diffusion pour que la série temporelle entière soit prise en considération (de sorte que toute la série temporelle est prise en compte). Nous le faisons dans un premier temps par analogie avec l'estimation du maximum de vraisemblance, prenant comme notre fonction critère le produit des statistiques de Kolmogorov-Smirnov calculées à chaque instant. En utilisant la distribution asymptotique

nulle, on obtient

$$\phi = \prod_{t=1}^T L(D_t)$$

Où T est la longueur de la série temporelle.

Etant donné un ensemble d'observations ou des observations simulées donnant lieu à $S_m(x)$, cette fonction critère est maximisée par rapport aux paramètres de la dérive et de diffusion de l'EDS qui est utilisé dans l'évaluation de $S_n(x)$ pour chaque pas de temps.

Résultats empiriques

L'estimation des paramètres de la dérive et de diffusion de l'EDS **(2.2)** utilisant la fonction critère ci-dessus de Kolmogorov-Smirnov. m réalisations de T points de données observés ($m = (50, 100)$, $T = (20, 50)$) pour la valeur initiale $x_0 = 0.5$ et les vrais paramètres $\mu = 1$ et $\sigma = 0.5$. La méthode de simplex de Nelder et Mead (1965) a été alors utilisé pour estimer μ et σ utilisant au départ les valeurs de $\mu = 1.3$ et $\sigma = 0.4$, la valeur initiale $x_0 = 0.1$, le nombre de reproductions simulées $n = (50, 100)$ et la longueur des séries temporelles $T = (20, 50)$.

L'estimation de paramètre a été portée sur 500 fois puis 1000 fois, les moyennes et les écart-types obtenus à partir de ces estimations sont présentées dans les tableaux **3.10** et **3.11**.

Il convient de noter que, à cause de la complexité de calcul associée à $L(D)$, dans l'équation **(2.3)**, nous avons utilisé l'approximation

$$L(D) \approx 1 - e^{-2D^2}$$

Qui est la distribution asymptotique du test unilatéral à deux échantillons décrit par Gibbons (1985). Cette approximation réduit le temps de calcul pris lors de l'utilisation de la formule de $L(D)$ de l'équation **(2.4)**

Le nombre de données m	Le nombre de simulations n	la longueur des séries temporelles T	La moyenne des paramètres	L'écart type des paramètres
50	50	20	$\mu=1.102$ $\sigma=0.485$	$\mu=0.101$ $\sigma=0.105$
		50	$\mu=1.053$ $\sigma=0.491$	$\mu=0.077$ $\sigma=0.078$
	100	20	$\mu=1.085$ $\sigma=0.522$	$\mu=0.078$ $\sigma=0.086$
		50	$\mu=1.014$ $\sigma=0.505$	$\mu=0.049$ $\sigma=0.043$
100	50	20	$\mu=1.098$ $\sigma=0.523$	$\mu=0.099$ $\sigma=0.10$
		50	$\mu=1.037$ $\sigma=0.517$	$\mu=0.074$ $\sigma=0.076$
	100	20	$\mu=1.071$ $\sigma=0.531$	$\mu=0.079$ $\sigma=0.086$
		50	$\mu=1.009$ $\sigma=0.513$	$\mu=0.043$ $\sigma=0.044$

Tableau 3.10- pour 500 réplifications

Le nombre de données m	Le nombre de simulations n	la longueur des séries temporelles T	La moyenne des paramètres	L'écart type des paramètres
50	50	20	$\mu=1.098$ $\sigma=0.479$	$\mu=0.098$ $\sigma=0.110$
		50	$\mu=1.061$ $\sigma=0.487$	$\mu=0.075$ $\sigma=0.070$
	100	20	$\mu=0.985$ $\sigma=0.520$	$\mu=0.075$ $\sigma=0.080$
		50	$\mu=1.010$ $\sigma=0.493$	$\mu=0.045$ $\sigma=0.040$
100	50	20	$\mu=1.075$ $\sigma=0.481$	$\mu=0.092$ $\sigma=0.096$
		50	$\mu=0.991$ $\sigma=0.512$	$\mu=0.073$ $\sigma=0.053$
	100	20	$\mu=1.023$ $\sigma=0.481$	$\mu=0.069$ $\sigma=0.84$
		50	$\mu=1.005$ $\sigma=0.492$	$\mu=0.040$ $\sigma=0.041$

Tableau 3.11- pour 1000 réplifications

Les résultats de tableau 3.10 et 3.11 révèlent un léger biais dans les estimations de μ et σ au cours des 500 et 1000 répétitions.

Pour un nombre de réalisation donnée, m , l'augmentation de nombre de reproduction, n , améliore à la fois les moyennes des estimations de paramètre et leurs écart-types.

Les résultats d'estimation pour un nombre de réalisation $m = 50$ et 100 presque les mêmes. Nous remarquons aussi que les résultats d'estimation de la dérive et de diffusion pour $T = 50$ est meilleurs par rapport à celle donnée pour $T = 20$.

En conclusion l'estimation des paramètres des EDS linéaires (2.2) utilisant les deux fonctions critère basées sur les statistiques de Chi-deux et de Kolmogorov-Smirnov donne des résultats satisfaisants, l'inconvénient c'est qu'elles exigent beaucoup de calcul et de données surtout pour celle basée sur la statistique de Chi-deux.

La méthode historique basée sur l'estimation de maximum de vraisemblance donne des bons

résultats par rapport à ces deux méthodes et de même le temps d'exécution est moins. L'inconvénient de cette méthode d'estimation c'est que la fonction de vraisemblance doit être connue par contre pour les deux autres ce n'est pas nécessaire.

Dans notre travail nous avons concentré sur le modèle de Black et Scholes car c'est le modèle le plus couramment utilisé pour les mathématiques financières, alors que beaucoup de systèmes nous ne pouvons pas les modéliser par ce modèle de Black et Scholes, il existe des systèmes qui n'ont pas la forme de l'EDS linéaire (2.2). Cependant, il existe d'autres EDS linéaire et non linéaires pour modéliser ces systèmes.

les équations spécifiques que nous examinons sont les suivantes :

$$dX_t = X_t(1 - X_t)dt + \sigma X_t^2 dW_t \quad (2.5)$$

$$dX_t = \mu X_t(1 - X_t)dt + \sigma X_t dW_t \quad (2.6)$$

$$dX_t = \mu X_t(1 - X_t)dt + 0.2\sigma X_t(1 - X_t)dW_t \quad (2.7)$$

$$dX_t = X_t(1 - X_t)^\mu dt + 0.5X_t^{2\sigma} dW_t \quad (2.8)$$

Où μ et σ sont des coefficients constants à estimer en tant que paramètres.

Ces équations ont été choisies parce qu'elles sont de la forme fréquemment utilisée pour la modélisation des systèmes de ressource renouvelables. L'obtention des estimations ponctuelles des paramètres de la dérive et de diffusion de tels modèles est importante dans les travaux pratiques car ces estimations fournissent des moyens pour tester les hypothèses sur l'état du système et l'importance relative des influences stochastiques sur ce système.

L'estimation des paramètres d'une EDS exige sa solution ou son approximation. Compte tenu de la difficulté associée à la découverte des solutions sous forme explicites pour beaucoup de EDS non linéaires, nous nous basons sur trois méthodes d'approximations des équations (2.5) à (2.8) ; à savoir le schéma d'Euler, le schéma de Milstein et le schéma de Taylor fort décrites dans la dernière section de 2^{ème} chapitre.

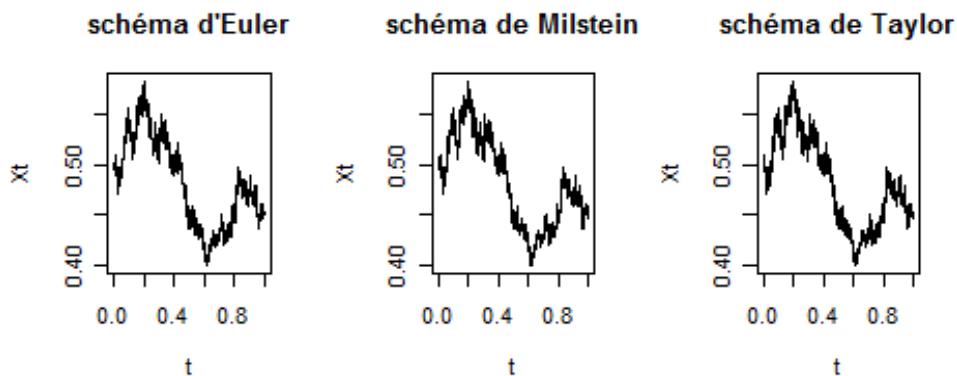


Figure 3.10-les 3 schémas pour le modèle (2.5)

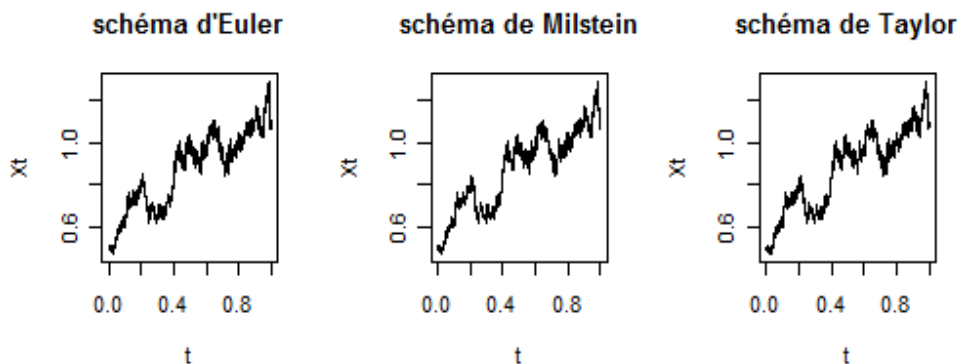


Figure 3.11-les 3 schémas pour le modèle (2.6)

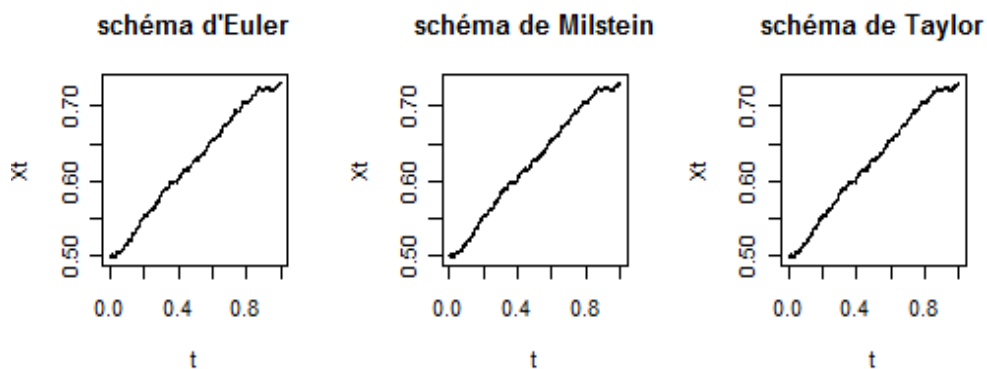


Figure 3.12-les 3 schémas pour le modèle (2.7)

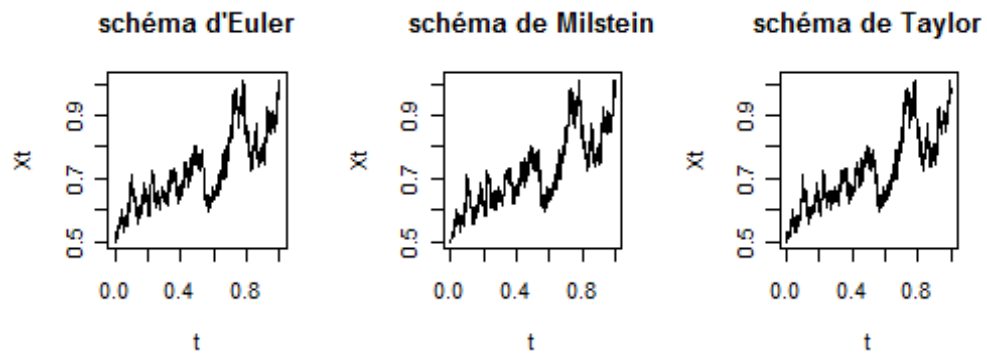


Figure 3.13-les 3 schémas pour le modèle (2.8)

Résultats empiriques

Les résultats des tableaux 3.12 à 3.15 présentent l'estimation des paramètres de modèle (2.5) à (2.8) par les deux méthodes présentées auparavant pour 500 réplifications.

Pour chaque modèle les valeurs actuelles des coefficients, $\mu = 1$ et $\sigma = 0.5$, ont été simulé sur l'intervalle 0 à 1 et pour $x_0 = 0.5$.

Pour la méthode basée sur la Statistique de Chi-deux, le nombre de données $n = 50$, le nombre d'observations pour chaque t_k , $m = 100$ et 5 éléments dans chaque cellule.

Pour la méthode basée sur la Statistique de Kolmogorov-Smirnov, Le nombre de données $m = 50$, Le nombre de simulations $n = 100$ et $T = 50$.

La fonction Critère		La moyenne des paramètres	L'écart type des paramètres
Basée sur la χ^2	Euler	$\sigma=0.515$	$\sigma=0.053$
	Milslein	$\sigma=0.514$	$\sigma=0.050$
	Taylor	$\sigma=0.517$	$\sigma=0.051$
Basée sur la K-S	Euler	$\sigma=0.517$	$\sigma=0.052$
	Milslein	$\sigma=0.520$	$\sigma=0.049$
	Taylor	$\sigma=0.518$	$\sigma=0.041$

Tableau 3.12- les résultats d'estimation pour le modèle (2.5)

La fonction Critère		La moyenne des paramètres	L'écart type des paramètres
Basée sur la χ^2	Euler	$\mu=1.123$ $\sigma=0.534$	$\mu=0.098$ $\sigma=0.057$
	Milslein	$\mu=1.123$ $\sigma=0.531$	$\mu=0.098$ $\sigma=0.054$
	Taylor	$\mu=1.098$ $\sigma=0.529$	$\mu=0.081$ $\sigma=0.052$
Basée sur la K-S	Euler	$\mu=1.110$ $\sigma=0.545$	$\mu=0.091$ $\sigma=0.063$
	Milslein	$\mu=1.097$ $\sigma=0.532$	$\mu=0.085$ $\sigma=0.050$
	Taylor	$\mu=1.103$ $\sigma=0.529$	$\mu=0.068$ $\sigma=0.048$

Tableau 3.13-les résultats d'estimation pour le modèle (2.6)

La fonction Critère		La moyenne des paramètres	L'écart type des paramètres
Basée sur la χ^2	Euler	$\mu=0.991$ $\sigma=0.551$	$\mu=0.075$ $\sigma=0.091$
	Milslein	$\mu=1.085$ $\sigma=0.534$	$\mu=0.074$ $\sigma=0.087$
	Taylor	$\mu=1.023$ $\sigma=0.541$	$\mu=0.062$ $\sigma=0.087$
Basée sur la K-S	Euler	$\mu=1.098$ $\sigma=0.521$	$\mu=0.071$ $\sigma=0.065$
	Milslein	$\mu=1.072$ $\sigma=0.519$	$\mu=0.065$ $\sigma=0.051$
	Taylor	$\mu=1.012$ $\sigma=0.514$	$\mu=0.052$ $\sigma=0.045$

Tableau 3.14-les résultats d'estimation pour le modèle (2.7)

La fonction Critère		La moyenne des paramètres	L'écart type des paramètres
Basée sur la χ^2	Euler	$\mu=1.089$ $\sigma=0.514$	$\mu=0.087$ $\sigma=0.074$
	Milslein	$\mu=1.081$ $\sigma=0.516$	$\mu=0.082$ $\sigma=0.063$
	Taylor	$\mu=1.068$ $\sigma=0.512$	$\mu=0.059$ $\sigma=0.051$
Basée sur la K-S	Euler	$\mu=1.051$ $\sigma=0.508$	$\mu=0.073$ $\sigma=0.065$
	Milslein	$\mu=1.049$ $\sigma=0.510$	$\mu=0.069$ $\sigma=0.054$
	Taylor	$\mu=1.033$ $\sigma=0.504$	$\mu=0.045$ $\sigma=0.042$

Tableau 3.15-les résultats d'estimation pour le modèle (2.8)

Le temps d'exécution pour la méthode basée sur la Statistique de Kolmogorov-Smirnov et moins par rapport à la méthode basée sur la Statistique de Chi-deux de même de la méthode basée sur la statistique de K-S donne des écarts types plus petits que ceux donnés par la méthode basée sur la statistique de Chi-deux pour les 4 EDS, par exemple pour la 4^{ème} EDS le schéma d'Euler donne un écart type de $\mu = 0.087$ pour Chi-deux et 0.073 pour K-S.

Les résultats d'estimation des paramètres pour les 4 modèles sont plus précis dans le cas où nous avons approximé l'EDS par le schéma de Taylor, par contre il y a pas une grande différence entre les deux autres schémas, par exemple pour la 4^{ème} EDS, dans le cas de la méthode basée sur la statistique de Chi-deux, la moyenne et l'écart type de μ est de 1.089 et 0.087 pour le schéma d'Euler, 1.081 et 0.082 pour le schéma de Milstein et 1.068 et 0.059 pour le schéma de Taylor.

3.4 Conclusion

Dans ce chapitre nous avons présenté quelques détails du modèle de Black et Scholes qui reste le plus utilisé dans le domaine de l'ingénierie financière. Ainsi sa formule qui a permis de répondre à la question de l'évolution du prix d'une option européenne. Pourtant définie en 1973, elle reste très utilisée de nos jours ; c'est une référence incontournable sur le marché des options. Pour cela, l'estimation de paramètres de ce modèle est un sujet intéressants en sois. Nous avons montré deux techniques d'estimation simule simultanément les paramètres du modèle l'une basée sur la statistique de Kolmogorov-Smirnov la seconde base sur une statistique de Chi-deux. La méthode de simplex de Nelder et Mead a été alors utilisé pour optimiser le chois de paramètres pour les deux méthodes.

La conclusion porte sur la comparaison des résultats obtenus par les deux méthodes et la méthode historique.

La méthode historique, basée sur l'estimation de maximum de vraisemblance, donne des bons résultats par apport à ces deux méthodes pour le modèle de de Black et Scholes qui suit la loi de mouvement brownien géométrique. Comparant les deux autres méthodes, celle de Kolmogorov-Smirnov donnent meilleurs résultats que la méthode de Chi-deux. En se basant sur certains critères, le critère qui minimise les variances des estimateurs et le temps d'exécution.

Chapitre 4

Conclusion générale

L'objet de cette thèse était de présenter et surtout d'estimer les paramètres du modèle de Black & Scholes qui reste le plus utilisé dans le monde de la finance.

Pour cela, nous avons choisi dans un premier temps d'estimer les paramètres du modèle par la méthode historique, qui utilise la propriété d'indépendance et de la normalité des rendements logarithmiques des actifs financiers, cette méthode basée sur l'estimation du maximum de vraisemblance car la loi des rendements est connue. Le modèle de Black & Scholes et ses variantes sont largement utilisés par les professionnels, pour cela, l'estimation de ses paramètres mérite qu'on s'intéresse à d'autres techniques plus adaptées : méthode estime simultanément les paramètres du modèle sans tenir compte de la distribution.

Les deux méthodes impliquent la simulation numérique des ensembles de données et en correspondant ces distributions simulées aux données expérimentales en utilisant la statistique de Chi-deux de la qualité de l'ajustement (the Chi-2 goodness-of-fit statistic) pour la première méthode et la statistique du Kolmogorov-Smirnov à deux échantillons pour la deuxième méthode. Les résultats d'estimation sont satisfaisants mais exigent beaucoup de calcul surtout pour la méthode basée sur la statistique de Chi-deux.

Dans le but de voir et vérifier l'efficacité d'estimation par ces deux méthodes, nous avons choisi et estimé les paramètres de quatre EDS non linéaires qui n'ont pas de solution explicite, nous avons eu l'occasion d'utiliser des schémas d'approximation tels que le schéma de Taylor d'ordre 1.5, le schéma d'Euler et le schéma de Milstein.

La conclusion porte sur la comparaison des résultats obtenus par les deux méthodes d'estimation, toutes ces deux méthodes donnent de bons résultats mais exigent beaucoup de calcul,

la méthode basée sur la statistique de Kolmogorov-Smirnov donne de meilleurs résultats que celle basée sur la statistique de Chi-deux, de même l'estimation des paramètres utilisant cette dernière méthode est meilleure pour l'EDS linéaire par rapport aux celles non linéaires.

Bibliographie

- [1] B. Bercu, D. Chafaï (2007). Modélisation stochastique et simulation, Dunod.
- [2] I.I.Gihman,A.V.Skorohod (1972).Stochastic differential equation.Springer-Verlag.
- [3] I.I.Gihman,A.V.Skorohod (1973).The theory of stochastic processes.Springer- Verlag,Berlin.
- [4] S.Karlin,H.M.Taylor (1975).A first cours in stochastic processes. Academic Press.New York.
- [5] A.Friedman (1975). Stochastic differential equations and applications.Academic Prress.
- [6] I.I.Gihman,A.V .Skorohod(1980).Introduction à la théorie des processus aléatoires.Moscou,éditions Mir.
- [7] N.Ikeda,S.Watanabe (1981). Stochastic differential equations and diffusion processes. North Holland Publishing company- Kodousha Itd-Tokio.
- [8] R.J.Elliott(1982).Stochastic Calculus and Applications.Springer–Verlag–New York.
- [9] L.Karatzas,S.E.Shreve (1988). Brownian motion and stochastic calculs (ISNN). Springer-Verlag.
- [10] P.E.Klooaden,E.Platen (1991).Numerical solution of stochastic differential equations. Springer.
- [11] E Black and M. Scholes (1973). The pricing of options and corporate liabilities, Journal of Political Economy 81, 637-659.
- [12] N.Limnios (2005). Introduction à la méthode de Monte Carlo, Université de technologie de Compiègne.
- [13] A.Tverto,R.Winther(1998).Introduction to partial differential equations. A computational approach.Springer-Verlag–New York.

- [14] K.Boukhetala (1984). Estimation des paramètres de dérivées d'une équation différentielle stochastique (EDS) linéaire contrôlée. Thèse de magister, USTHB.
- [15] J.K.Boukhetala (1998). Application des processus de diffusion et problème d'échantillonnage. Thèse de Doctorat d'état, USTHB.
- [16] Nils Berglund (2003). Introduction aux intégrales stochastiques, France.
- [17] Olivier Lévêque (2004-2005). Cours de probabilités et calcul stochastique.
- [18] McDonald, D., Sandal, L.(1998). Estimating the parameters of stochastic differential equations using a criterion function based on the Kolmogorov-Smirnov statistic. Technical Report, Norwegian School of Economics and Business Administration.
- [19] K.Boukhetala (1995). Identification and simulation of a communication system. MMR Review. (ISSN1114398), Vol.4 pp55-79.
- [20] Kloeden, P.E., Platen, E. and Schurz, H. (1994). Numerical Solution of SDE Through Computer Experiments, Springer-Verlag, Berlin.
- [21] Hurn, A.S. and Lindsay, K.A. (1997). Estimating the Parameters of Stochastic Differential Equations by Monte Carlo Methods. Mathematics and Computers in Simulation, 43(3-6), 495-501.
- [22] J.JANSSEN, M. SAIB et K.TAOUS. Techniques d'Estimation pour le Modèle de Black et Scholes.
- [23] Y.A. Kutoyants (1984). Parameter Estimation for Stochastic Processes, Heldermann Verlag, Berlin.
- [24] Nelder, J.A. and Mead, R. (1965). A simplex method for function minimisation. The Computer Journal, 7, 308-313.
- [25] P.E.Kloeden, E.Platten (1999). Numerical Solution of Stochastic Differential Equations, Springer-Verlag, Berlin.
- [26] Lo, A. (1988). Maximum likelihood estimation of generalized Itô processes with discretely sampled data. Econom. Theory 4, 231-247.
- [27] Kessler, M.(1997). Estimation of an ergodic diffusion from discrete observations. Scand. J. Statist. 24, 211-229.

- [28] N.Limnios (2005), Introduction à la méthode de Monte Carlo, Université de technologie de Compiègne.
- [29] Nabil Saïmi (2011). Estimation de la volatilité et filtrage non linéaire. Mémoire pour l'obtention du grade de M.Sc, Université du Québec à Trois-Rivières.
- [30] Boukhetala K, Guidoum A (2011). **Sim.DiffProc** : Simulation of Diffusion Processes. R package version 2.0. <http://CRAN.R-project.org/package=Sim.DiffProc>.
- [31] Boukhetala K, Guidoum A (2011). **Sim.DiffProcGUI** : Graphical User Interface for Simulation of Diffusion Processes. R package version 2.0. <http://CRAN.R-project.org/package=Sim.DiffProcGUI>.
- [32] Boudali Khalida (2005). Aspects numériques et simulation dans les équations différentielles stochastiques. Thèse de Magister, encadré par K.Boukhetala, USTHB.
- [33] Sijlamassi Mohamed, Ouaknine Yoram (2004). Valorisation par les deflateurs stochastiques. L'émoire d'actuariat, ENSAE.
- [34] Gibbons, J.D (1985). Nonparametric Statistical Inference (2nd Ed.), Marcel Dekker, New York.
- [35] Hubert de la Bruslerie (1988). Volatilité implicite des options, outil de mesure ou objet de prise de position ?. Analyse Financière, n 74, p 36-45.
- [36] Hull,J,C (1999). Options Futures and other derivatives, Fourth edition, New Jersey