

Sébastien ROLAND

Mathématiques et Statistique de la Finance

Calcul Stochastique
Statistique des Processus

Première Edition (Sept. 2005).

Sébastien ROLAND © 2005

Préface

L'essor important de la théorie financière moderne, telle qu'elle fut développée à partir des résultats novateurs de Samuelson (1961-1965), Merton (1969-1971-1973-1976), Black et Scholes (1973) s'est pour une large part réalisé grâce au développement important de la théorie des processus stochastiques. Et sur ces trois dernières décennies, leur emploi s'est considérablement accru pour décrire l'évolution des fluctuations des marchés financiers, tant d'un point de vue de l'évaluation des actifs optionnels que de la gestion des risques. Dans le même temps, les progrès importants de l'informatique ont permis d'une part de rendre opérationnelles des méthodes numériques et statistiques parfois exigeantes en termes de ressources et d'autre part de constituer de volumineuses bases de données d'historiques des actifs financiers et des marchés optionnels. Finalement, la combinaison de ces deux effets place au centre des considérations actuelles et futures des risk managers les questions de la compréhension et de l'exploitation de ces informations.

L'objectif de ces notes est alors de deux ordres. Il vise à fournir une présentation autonome de la théorie des processus stochastiques et des méthodes statistiques connexes et de montrer comment ces notions et outils interagissent dans des problématiques d'évaluation et de gestion des risques, auxquels les praticiens de la finance sont constamment confrontés.

Le but de ces notes est alors:

- d'expliquer les motivations pour l'emploi des processus stochastiques pour la modélisation des grandeurs économiques et financières.
- de motiver la nécessité de l'équipement mathématique introduit pour la manipulation des processus stochastiques.
- de fournir les résultats mathématiques principaux des théories abordées sans entrer dans des détails techniques fastidieux.
- de présenter des exemples concrets d'emploi des processus stochastiques en évaluation et gestion des risques.

Sébastien ROLAND
Paris (France), Septembre 2005.

Mail: sebastien.roland@gmail.com.

⁰Version: May 3, 2006.

Sommaire

Préface	iii
I Calcul Stochastique	ix
1 Notions Générales de Probabilités	1
1.1 Eléments de Théorie de la Mesure	1
1.1.1 σ -algèbres et Mesures	1
1.1.2 Absolue Continuité et Densité	3
1.2 Variables Aléatoires	4
1.2.1 Espace de Probabilité	4
1.2.2 Quelle Caractérisation pour $(\Omega, \mathcal{F}, \mathbb{P})$?	6
1.2.3 Fonction Caractéristique	7
1.3 Processus Stochastiques	8
1.3.1 Fonctions Aléatoires	9
1.3.2 Filtrations et Histoires	10
1.3.3 Espérance Conditionnelle	13
1.3.4 Temps Aléatoire	13
1.3.5 Martingales et Propriété de Markov	14
2 Processus de Wiener, Poisson et Lévy	17
2.1 Processus de Wiener	17
2.1.1 Loi Gaussienne	17
2.1.2 Mouvement Brownien	19
2.1.3 Version Multidimensionnelle	28
2.2 Processus de Poisson	30
2.2.1 Loi Exponentielle	30
2.2.2 Loi de Poisson	31
2.2.3 Processus de Poisson	32
2.2.4 Processus de Poisson Composé	34
2.3 Processus de Lévy	35
2.3.1 Mesures Aléatoires	36
2.3.2 Marches Aléatoires	38
2.3.3 Propriété de Markov	41
A Propriétés Empiriques des Séries Financières	43
A.1 Du Temps Continu au Temps Discret	43
A.2 Faits Stylisés des Séries de Rendements	44

3	Calcul Stochastique	47
3.1	Intégrales Stochastiques	47
3.1.1	Mouvement Brownien	47
3.1.2	Cadre des Semimartingales	53
3.1.3	Processus de Poisson et Lévy	57
3.2	Variation Quadratique	59
3.2.1	Lien avec la Volatilité Réalisée	59
3.2.2	Covariance Quadratique	61
3.3	Formule d'Itô	63
3.3.1	Cas du Mouvement Brownien	63
3.3.2	Cas des Processus d'Itô	65
3.3.3	Cas des Processus de Poisson	68
3.4	Changement de Mesure	72
3.4.1	Construction	72
3.4.2	Cas du Processus de Wiener	74
3.4.3	Cas des Processus de Poisson	75
3.4.4	Cas des Modèles Mixtes	78
B	Evaluation Risque-Neutre	79
B.1	Règles d'Evaluation	79
B.2	Mesures Martingales Equivalentes	82
4	Equations aux Dérivées Partielles	85
4.1	Equations Différentielles Stochastiques	85
4.1.1	Cas Brownien	85
4.1.2	Cas Mixte Brownien-Poisson	87
4.2	Propriété de Markov	89
4.3	Equations aux Dérivées Partielles	90
4.3.1	Cas Brownien	90
4.3.2	Cadre Multidimensionnel	93
4.3.3	Cas Mixte Brownien-Poisson	93
4.4	Modèles de Taux d'Intérêt	94
C	Simulation des EDS-S	97
C.1	Simulation Exacte	97
C.1.1	Cas du Mouvement Brownien	97
C.1.2	Cas des Processus à Sauts	99
C.2	Méthodes de Discrétisation	102
C.2.1	Schéma d'Euler	102
C.2.2	Schéma de Milstein	103
<hr/>		
II	Statistique des Processus	107
<hr/>		
5	Inférence Statistique	109
5.1	Contexte	109
5.1.1	Estimations MV et MMG	109
5.2	Quelques Exemples	113
5.2.1	Mouvement Brownien Géométrique	113

5.2.2	Processus d'Ornstein-Uhlenbeck	114
5.2.3	Equation Différentielle Stochastique	116
6	Filtrage Linéaire Gaussien	119
6.1	Contexte	119
6.2	Filtrage en Temps Discret	119
6.2.1	Cadre	119
6.2.2	Normalité A Priori-A Posteriori	120
6.2.3	Equations Récursives	122
6.2.4	Interpolation: Lissage	125
6.2.5	Extrapolation: Prédiction	126
6.3	Filtrage en Temps Continu	127
6.3.1	Cadre	127
6.3.2	Structure des Innovations	128
6.3.3	Résultat Général	130
6.3.4	Equations Récursives	135
6.3.5	Interpolation: Lissage	137
6.3.6	Extrapolation: Prédiction	140
7	Filtrage de Chaînes de Markov	143
7.1	Contexte	143
7.2	Structure des Probabilités	144
7.3	Résultat Général	146
7.4	Approche Alternative	147
8	Algorithme Espérance-Maximisation	153
8.1	Cadre	153
8.2	Inférence de Modèles Latents	154
8.2.1	Inférence Statistique	155
8.3	Estimation d'une Chaîne de Markov	157

Part I

Calcul Stochastique

1

Notions Générales de Probabilités

L'objectif de ce Chapitre est d'introduire l'ensemble des concepts mathématiques qui seront utiles dans la suite de la présentation du calcul stochastique en finance. Après avoir présenté les concepts de mesures, de σ -algèbres, de filtrations et de martingales, nous aborderons les processus stochastiques en temps continu.

1.1. Éléments de Théorie de la Mesure

1.1.1 σ -algèbres et Mesures

La notion de mesure est une extension naturelle des concepts plus familiers de distance, d'aire et de volume à un cadre plus abstrait. Considérons un ensemble E , que nous supposons la plupart du temps être \mathbb{R}^d (pour $d \geq 1$). On dira qu'une mesure μ sur E associe à certains ensembles $A \subset E$, que nous nommerons alors des ensembles mesurables, un nombre (pouvant être infini) positif $\mu(A) \in [0, \infty]$ que nous appellerons la mesure de A . Par analogie avec les notions d'aire ou de volume, il est naturel de poser que l'ensemble vide \emptyset a pour mesure 0, soit $\mu(\emptyset) = 0$. De même, si A et B sont des ensembles mesurables disjoints, l'ensemble $A \cup B$ est aussi mesurable et sa mesure est définie comme $\mu(A \cup B) = \mu(A) + \mu(B)$; il s'agit de la propriété d'additivité. Par ailleurs, pour une suite $(A_n)_{n \geq 1}$ d'ensembles mesurables disjoints, nous avons:

$$\mu\left(\bigcup_{n \geq 1} A_n\right) = \sum_{n \geq 1} \mu(A_n)$$

Cette propriété est parfois connue sous le nom de σ -additivité. Par ailleurs, grâce à la propriété d'additivité, nous avons aussi:

$$\mu(A^c) = \mu(E) - \mu(A)$$

où A^c est le complémentaire de A . Pour conclure sur ces remarques, on dira que le domaine de définition d'une mesure \mathcal{E} sur E est une collection d'ensembles de E qui:

- contient l'élément nul: $\emptyset \in \mathcal{E}$,
- est stable par passage à l'union:

$$A_n \in \mathcal{E}, (A_n)_{n \geq 1} \text{ disjoints} \Rightarrow \bigcup_{n \geq 1} A_n \in \mathcal{E}$$

- contient le complémentaire de tout élément: $\forall A \in \mathcal{E}, A^c \in \mathcal{E}$.

Une telle collection d'ensemble est alors appelée une σ -algèbre. D'où la définition.

Définition 1.1.1 (σ -algèbre) Soit Ω un ensemble non-nul et soit \mathcal{F} une collection de sous-ensembles de Ω . Alors, \mathcal{F} est une σ -algèbre si:

1. L'élément nul \emptyset appartient à \mathcal{F} ,
2. Si A appartient à \mathcal{F} , alors son complémentaire $A^c = \Omega - A$ appartient à \mathcal{F} ,
3. Si la séquence $(A_n)_{n \geq 1}$ appartient à \mathcal{F} , alors $\bigcup_{n \geq 1} A_n$ appartient à \mathcal{F} .

Un ensemble mesurable est alors un élément d'une σ -algèbre. Une première question que nous sommes en droit de nous poser est: comment pouvons-nous construire des σ -algèbres?

Exemple 1.1.2 Si Ω représente la droite réelle \mathbb{R} , les opérations de réunion et d'image par complémentaire d'intervalles forment une σ -algèbre.

Exemple 1.1.3 Si E est doté d'une topologie, c'est-à-dire qu'il est possible de définir une notion de convergence, ce qui est le cas pour $E = \mathbb{R}^d$, la σ -algèbre générée par tous les ouverts est appelée la tribu borelienne $\mathcal{B}(E)$. Ce cas est d'un intérêt particulier car la définition de mesures sur E assure que tous les ensembles ouverts ou fermés sont mesurables.

Jeu de Pile ou Face

Supposons que Ω représente les huit résultats possibles d'un jeu de pile ou face avec trois pièces. Si nous nous intéressons plus particulièrement au résultat du jeu après le lancer de la première pièce, alors les ensembles:

$$A_P = \{PPP, PPF, PFP, P, FF\}, \quad A_F = \{FPP, FPF, FFP, FFF\}$$

sont connus suivant l'information disponible (*resolved by the information*). Par ailleurs, les ensembles \emptyset et Ω sont eux aussi connus, soit $\mathcal{F}_0 = \{\emptyset, \Omega\}$. De fait, les ensembles connus dès lors que le résultat de la première pièce est disponible forment la σ -algèbre:

$$\mathcal{F}_1 = \{\emptyset, \Omega, A_P, A_F\}$$

Dit autrement, la σ -algèbre \mathcal{F}_1 représente l'information *accumulée* par observation du résultat du lancer de la première pièce.

Nous pouvons à présent définir une mesure.

Définition 1.1.4 (Mesure) Soit \mathcal{E} une σ -algèbre de sous-ensembles de E . (E, \mathcal{E}) est appelée un ensemble mesurable. Une mesure (positive) sur (E, \mathcal{E}) est une fonction:

$$\begin{aligned} \mu : \mathcal{E} &\rightarrow [0, \infty] \\ A &\mapsto \mu(A) \end{aligned}$$

telle que:

1. $\mu(\emptyset) = 0$,
2. Pour toute suite d'ensemble disjoints $(A_n)_{n \geq 1} \in \mathcal{E}$:

$$\mu\left(\bigcup_{n \geq 1} A_n\right) = \sum_{n \geq 1} \mu(A_n)$$

Un élément $A \in \mathcal{E}$ est appelé un ensemble mesurable et $\mu(A)$ sa mesure.

Un exemple bien connu de mesure est la mesure de Lebesgue sur \mathbb{R}^d : elle correspond à la notion (en dimension d) de volume:

$$\lambda(A) = \int_A dx, \quad A \in \mathcal{E}$$

Et de façon plus générale et pour reprendre la définition précédente, il est possible pour toute fonction continue $\rho: \mathbb{R}^d \rightarrow \mathbb{R}^+$ de définir une mesure sur \mathbb{R}^d comme:

$$\begin{aligned} \mu: \mathcal{B}(\mathbb{R}^d) &\rightarrow [0, \infty] \\ A &\mapsto \int_A \rho(x) dx = \int \mathbf{1}_A \rho(x) dx \end{aligned}$$

La fonction ρ est appelée la densité de μ par rapport à la mesure de Lebesgue λ .

Soit μ une mesure sur E , $\mu(E)$ n'est pas nécessairement finie. Une mesure μ sur (E, \mathcal{E}) est dite finie si $\mu(A) < \infty$, pour tout ensemble mesurable $A \subset E$. La quantité $\mu(E)$ est habituellement appelée la masse de μ . Par exemple, une masse de Dirac est une mesure finie de masse 1. Une mesure de masse 1 est appelée une mesure de probabilité et alors (E, \mathcal{E}, μ) est un espace de probabilité. Toutefois, toutes les mesures que nous pouvons rencontrer en pratique ne sont pas toujours des mesures finies. Un exemple est la mesure de Lebesgue sur \mathbb{R}^d qui est de masse infinie. Une notion plus flexible de mesure est alors donnée par une mesure de Radon.

Définition 1.1.5 (Mesure de Radon) Soit $E \subset \mathbb{R}^d$. Une mesure de Radon sur (E, \mathcal{B}) où \mathcal{B} est la tribu borélienne est une mesure μ telle que pour tout compact (sous-ensemble borné et fermé) mesurable $B \in \mathcal{B}$, $\mu(B) < \infty$.

Pour exemple, la mesure de Lebesgue sur \mathbb{R} est une mesure de Radon: la longueur de tout intervalle borné est finie. De même, les mesures de Dirac et toute combinaison linéaire de mesures de Dirac sont des mesures de Radon.

1.1.2 Absolue Continuité et Densité

Considérons un ensemble mesurable (E, \mathcal{E}) doté des mesures μ_1 et μ_2 . La question naturelle que nous nous posons est: comment ces deux mesures peuvent-elles être comparées? Une idée naturelle est de comparer μ_1 et μ_2 en observant le ratio de vraisemblance $\mu_2(A)/\mu_1(A)$ pour plusieurs définitions de l'ensemble A . Mais cela n'est possible que si $\mu_2(A) = 0$ chaque fois que $\mu_1(A) = 0$. Cette observation motive alors pour la définition suivante.

Définition 1.1.6 (Absolue continuité) Une mesure μ_2 est dite absolument continue par rapport à la mesure μ_1 et on note $\mu_2 \ll \mu_1$ si pour tout ensemble mesurable A :

$$\mu_1(A) = 0 \Rightarrow \mu_2(A) = 0$$

La propriété d'absolue continuité peut par ailleurs être caractérisée sous la forme suivante.

Théorème 1.1.7 (Radon-Nikodym) Soit $\mu_2 \ll \mu_1$. Alors, il existe une fonction mesurable $Z : E \rightarrow [0, \infty]$ telle que pour tout ensemble mesurable A :

$$\mu_2(A) = \int_A Z d\mu_1 = \mu_1(\mathbf{1}_A Z)$$

La fonction Z est appelée la densité ou dérivée de Radon-Nikodym de μ_2 par rapport à μ_1 et on note $Z = \frac{d\mu_2}{d\mu_1}$. Pour toute fonction intégrable par rapport à μ_2 , nous avons :

$$\mu_2(f) = \int_E f d\mu_2 = \mu_1(fZ) = \int_E fZ d\mu_1$$

Preuve. Selon la Définition 1.1.4, la preuve est complète si μ_2 est une mesure, c'est-à-dire si $\mu_2(\emptyset) = 0$ et si la propriété de σ -additivité est vérifiée. ■

Le résultat précédent s'interprète de la sorte: si μ_2 est absolument continue par rapport à μ_1 , alors l'intégrale par rapport à μ_2 est une intégrale pondérée par rapport à μ_1 , le poids étant donnée par la densité Z . Il peut par ailleurs se réécrire en fonction du signe espérance. Soit, pour une variable aléatoire X non-négative:

$$\mathbb{E}^{\mu_2}[X] = \mathbb{E}^{\mu_1}[ZX]$$

ou pour une variable aléatoire Y non-négative:

$$\mathbb{E}^{\mu_1}[Y] = \mathbb{E}^{\mu_2}\left[\frac{Y}{Z}\right]$$

Remarque 1.1.8 (Mesures équivalentes) Lorsque μ_1 est absolument continue par rapport à μ_2 et que μ_2 est absolument continue par rapport à μ_1 , on dit que μ_1 et μ_2 sont des mesures équivalentes, c'est-à-dire comparables.

Remarque 1.1.9 (Absolue continuité et évaluation d'actifs) Le cadre d'évaluation dit par arbitrage suppose l'existence d'une nouvelle mesure, dite risque-neutre ou martingale, sous laquelle le prix des actifs actualisés est une martingale. Le théorème de Radon-Nikodym est alors central dans cette approche, cf. Chapitre 3 et Section 3.4.

1.2. Variables Aléatoires

1.2.1 Espace de Probabilité

Considérons l'ensemble Ω , appelé ensemble des possibles (ou des scénarios), équipé de la σ -algèbre \mathcal{F} . Une mesure de probabilité sur (Ω, \mathcal{F}) est une mesure

positive finie \mathbb{P} de masse totale 1. Le triplet $(\Omega, \mathcal{F}, \mathbb{P})$ est alors appelé un espace de probabilité. Et un ensemble $A \in \mathcal{F}$, que nous appellerons un évènement, est alors un ensemble de scénarios auquel une probabilité peut être assignée. Une mesure de probabilité attribue ainsi une probabilité, nombre compris entre 0 et 1, à chaque évènement:

$$\begin{aligned} \mathbb{P} : \mathcal{F} &\rightarrow [0, 1] \\ A &\mapsto \mathbb{P}(A) \end{aligned}$$

Il est par ailleurs possible de définir une propriété d'absolue continuité et d'équivalence pour des mesures de probabilités: deux probabilités \mathbb{P} et \mathbb{Q} définies sur (Ω, \mathcal{F}) sont équivalentes si elles définissent les mêmes ensembles d'évènements à probabilité nulle:

$$\mathbb{P} \sim \mathbb{Q} \Rightarrow [\forall A \in \mathcal{F}, \mathbb{P}(A) = 0 \Leftrightarrow \mathbb{Q}(A) = 0]$$

Une variable aléatoire X à valeurs dans E est une fonction mesurable:

$$X : \Omega \rightarrow E$$

où $(\Omega, \mathcal{F}, \mathbb{P})$ est un espace de probabilité. Un élément $\omega \in \Omega$ est appelé un scénario ou un évènement. Ainsi $X(\omega)$ représente le résultat de la variable aléatoire si le scénario ω se réalise et est appelé la réalisation de X pour le scénario ω . La loi de X est la mesure de probabilité définie sur E comme:

$$\mu_X(A) = \mathbb{P}(X \in A)$$

Si $\mu_X = \mu_Y$ alors les variables X et Y sont dites identiques en loi et on note $X \stackrel{d}{=} Y$.

Changement de Mesure pour une Variable Aléatoire Gaussienne

Nous montrons par la suite une application du Théorème 1.1.7. Soit X une variable aléatoire gaussienne définie sur l'espace de probabilité $(\Omega, \mathcal{F}, \mathbb{P})$. En termes de mesure, cela équivaut à:

$$\mu_X(A) = \mathbb{P}(X \in A) = \int_A \phi(x) dx, \quad A \in \mathcal{F}$$

où:

$$\phi(x) = \frac{1}{\sqrt{2\pi}} e^{-\frac{x^2}{2}}, \quad x \in \mathbb{R}$$

est la distribution de probabilité de la loi normale centrée réduite $\mathcal{N}(0, 1)$. Soit par ailleurs θ une constante positive et nous définissons $Y = X + \theta$, qui est alors, sous \mathbb{P} , une variable aléatoire gaussienne de moyenne θ et de variance 1. Nous cherchons à présent à expliciter une nouvelle mesure de probabilité, soit $\tilde{\mathbb{P}}$, sous laquelle Y est une variable gaussienne de moyenne 0 et de variance 1. On notera que l'on ne cherche pas à soustraire θ de Y , mais plutôt à donner plus de poids aux trajectoires ω pour lesquels $Y(\omega)$ est positive et moins de poids aux trajectoires ω pour lesquels $Y(\omega)$ est négative. Ainsi, nous changeons la

distribution de Y sans affecter sa définition. A cette fin, nous considérons la v.a.:

$$Z(\omega) = \exp\left(-\theta X(\omega) - \frac{1}{2}\theta^2\right), \quad \omega \in \Omega$$

qui possède deux propriétés remarquables qui lui assurent de pouvoir être utilisée comme dérivée de Radon-Nikodym et ainsi obtenir une nouvelle probabilité $\tilde{\mathbb{P}}$. Nous avons:

$$\mathbb{E}[Z] = 1, \quad Z(\omega) > 0, \quad \forall \omega \in \Omega$$

Dès lors, par application du Théorème 1.1.7 de Radon-Nikodym, nous définissons:

$$\tilde{\mathbb{P}}(A) = \int_A Z(\omega) d\mathbb{P}(\omega), \quad A \in \mathcal{F}$$

et Z a la propriété que si $X(\omega)$ est positif, alors $Z(\omega) < 1$, ce qui se traduit par le fait que $\tilde{\mathbb{P}}$ attribue moins de poids que \mathbb{P} aux ensembles pour lesquels X est positif, recentrant de fait la variable aléatoire Y . Il est par ailleurs possible de montrer que:

$$\tilde{\mathbb{P}}(Y \leq a) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^a e^{-\frac{y^2}{2}} dy$$

et ainsi Y est bien une variable gaussienne centrée sous la mesure de probabilité $\tilde{\mathbb{P}}$.

1.2.2 Quelle Caractérisation pour $(\Omega, \mathcal{F}, \mathbb{P})$?

La définition que nous avons précédemment donnée d'un ensemble de probabilité $(\Omega, \mathcal{F}, \mathbb{P})$ peut, en première analyse, sembler excessive. En effet, d'un point de vue probabiliste et statistique, la définition d'une variable aléatoire X s'accompagne de la spécification de sa distribution de probabilité μ_X et toutes les opérations: somme, espérance, variance, etc. réalisées sur des variables aléatoires se résument à des opérations analytiques: intégration, convolution, etc. faites sur leurs distributions de probabilités. Pour tout cela, un espace de probabilité ne semble pas nécessaire.

Cependant, comme le triplet $(\Omega, \mathcal{F}, \mathbb{P})$ est à la base de la théorie (moderne) des processus stochastiques, il est nécessaire de comprendre pourquoi et comment ce concept est pertinent.

D'un point de vue de la modélisation, le point de départ est un ensemble d'observations qui prend ses valeurs dans un ensemble E , sur lequel nous souhaitons construire un processus stochastique. Dans la plupart des cas, il est naturel d'identifier Ω avec E , c'est-à-dire d'identifier l'aléa ω avec son observé. Dans l'exemple d'un jeu de dés, $\Omega = \{1, 2, 3, 4, 5, 6\}$ et donc l'ensemble des possibles correspond à l'ensemble des réalisations et donc $X(\omega) = \omega$. Ce choix de Ω est qualifié de canonique pour la variable aléatoire X : le résultat de l'aléa est caractérisé par l'aléa lui-même. Dès lors, comme cette construction est propre à X , il n'est pas possible de définir sur le même espace de probabilité une autre variable aléatoire qui soit indépendante de X : X est la seule source d'incertitude pour tout le modèle. Ainsi, si nous souhaitons prendre en compte

des modèles dotés d'une structure suffisamment riche, nous devons distinguer Ω - l'ensemble des scénarios sources de l'aléa - de E , l'ensemble des valeurs prises par les variables aléatoires. Une structure riche doit se comprendre dans le sens où le système étudié ne peut s'expliquer par une seule variable aléatoire, mais par plusieurs et que les liaisons entre chacune de ces sources d'aléa peuvent être complexes.

Pour résumer ces arguments, nous retiendrons les commentaires suivants. La description probabiliste d'une variable aléatoire X peut se ramener à la connaissance de sa distribution μ_X seulement si X est la seule source d'aléa. Dans ce cas, un processus stochastique peut être construit à partir d'une construction canonique de la variable aléatoire X . Dans les autres cas, c'est-à-dire dès lors que nous nous intéressons à une deuxième variable aléatoire qui ne soit pas une fonction déterministe de X , la probabilité \mathbb{P} contient plus d'information sur X que sa seule distribution. En particulier, elle traduit une information sur la dépendance de X avec toutes les autres variables aléatoires du modèle: la spécification de \mathbb{P} englobe la spécification des distributions jointes de toutes les variables aléatoires construites sur Ω .

1.2.3 Fonction Caractéristique

La fonction caractéristique d'une variable aléatoire est la transformée de Fourier de sa distribution. Ce concept est important car nombre de propriétés probabilistes d'une variable aléatoire correspondent à des propriétés analytiques de sa fonction caractéristique.

Définition 1.2.1 (Fonction caractéristique) *La fonction caractéristique d'une \mathbb{R}^d -variable aléatoire X est la fonction $\Phi_X : \mathbb{R}^d \rightarrow \mathbb{R}$ définie par:*

$$\Phi_X(z) = \mathbb{E} [e^{iz \cdot X}] = \int_{\mathbb{R}^d} e^{iz \cdot x} d\mu_X(x), \quad \forall z \in \mathbb{R}^d$$

La fonction caractéristique spécifie complètement une loi: deux variables aléatoires ayant la même fonction caractéristique sont identiquement distribuées. Nous donnons dans le Tableau 1.1 l'expression de la fonction caractéristique pour quelques lois usuelles.

Distribution	Densité	Fonction caractéristique
Exponentielle $\mathcal{E}(\lambda)$	$\mu(x) = \lambda e^{-\lambda x}$	$\Phi(z) = \frac{\lambda}{\lambda - iz}$
Normale $\mathcal{N}(m, \sigma)$	$\mu(x) = \frac{1}{\sqrt{2\pi}\sigma} \exp\left(-\frac{(x-m)^2}{2\sigma^2}\right)$	$\Phi(z) = \exp(imz - \frac{1}{2}\sigma^2 z^2)$
Gamma $\mathcal{G}(\lambda, c)$	$\mu(x) = \frac{\lambda^c}{\Gamma(c)} x^{c-1} e^{-\lambda x}$	$\Phi(z) = \frac{1}{(1 - i\lambda^{-1}z)^c}$
Cauchy $\mathcal{C}(m, c)$	$\mu(x) = \frac{c}{\pi[(x-m)^2 + c^2]}$	$\Phi(z) = \exp(imz - c z)$

Table 1.1: Fonction caractéristique de lois usuelles.

Une propriété importante de cette fonction est qu'elle relie les moments d'une variable aléatoire à sa dérivée, rendant ainsi possible leurs calculs, de façon analytique ou numérique.

Proposition 1.2.2 (Fonction caractéristique et moments) *La variable aléatoire X possède des moments à tous les ordres si et seulement si $z \mapsto \Phi_X(z)$ est C^∞ en $z = 0$. Alors les moments de X sont reliés aux dérivées de Φ_X par:*

$$m_n = \mathbb{E}[X^n] = \frac{1}{i^n} \frac{\partial^n \Phi_X}{\partial z^n}(0), \quad n \geq 1$$

Si les $(X_i, i = 1, \dots, n)$ sont des variables aléatoires indépendantes, la fonction caractéristique de $S_n = X_1 + \dots + X_n$ est le produit de leurs fonctions caractéristiques:

$$\Phi_{S_n}(z) = \prod_{i=1}^n \Phi_{X_i}(z)$$

Remarque 1.2.3 (Conditions de moments) *La fonction caractéristique permet d'obtenir les moments théoriques d'une variable aléatoire à tous les ordres. Cette dérivation se trouve être particulièrement utile dans le cadre de la construction d'un estimateur des moments généralisés (MMG) pour l'obtention d'un vecteur des moments, cf. Section 5.1.1*

Nous concluons ce paragraphe en présentant une fonction qui est reliée de façon importante à la fonction caractéristique, il s'agit de la fonction génératrice des moments.

Définition 1.2.4 (Fonction génératrice des moments) *La fonction génératrice des moments d'une \mathbb{R}^d -variable aléatoire X est la fonction $M_X : \mathbb{R}^d \rightarrow \mathbb{R}$ définie par:*

$$M_X(u) = \mathbb{E}[e^{uX}], \quad \forall u \in \mathbb{R} \quad (1.1)$$

Alors que la fonction caractéristique est toujours définie (mais pas toujours explicite), la fonction génératrice des moments n'est pas toujours définie: l'intégrale 1.1 peut ne pas converger pour certaines valeurs de u . Quand M_X est définie, elle est reliée à Φ_X par:

$$M_X(u) = \Phi_X(-iu)$$

Par ailleurs, les moments d'ordre n d'une variable aléatoire X sont dérivés comme:

$$m_n = \mathbb{E}[X^n] = \frac{\partial^n M_X}{\partial u^n}(0), \quad n \geq 1$$

1.3. Processus Stochastiques

Un processus stochastique¹ est une famille $(X_t)_{t \geq 0}$ de variables aléatoires indexées par le temps. L'indice de temps t peut soit être discret, soit continu². Pour chaque réalisation de l'aléa ω , la trajectoire $X.(\omega) : t \rightarrow X_t(\omega)$ définit une fonction du temps, appelée la trajectoire du processus. De fait, un processus stochastique peut aussi être vu comme une suite de fonctions aléatoires: variables aléatoires à valeurs dans des espaces de fonctions.

¹Stochastique doit être compris dans le sens d'aléatoire.

²Dans ce document, nous nous intéresserons plus particulièrement au cas continu.

Si nous souhaitons interpréter l'indice t comme le temps, il est nécessaire de prendre en compte le fait que les événements deviennent moins incertains au fur et à mesure que plus d'information devient disponible au cours du temps. Pour caractériser cette idée, il est indispensable de pouvoir décrire comment l'information est progressivement révélée. Cela est réalisé en introduisant la notion importante de filtration, qui va ainsi permettre de définir les notions d'information passée et de non-anticipation et par suite permettre de classer les processus en fonction de ces propriétés.

1.3.1 Fonctions Aléatoires

Pour pouvoir définir les processus stochastiques comme des variables aléatoires de fonctions, il est nécessaire de définir des mesures sur des espaces de fonctions. L'espace de fonction le plus simple est dans doute l'espace de toutes les fonctions $f : [0, T] \rightarrow \mathbb{R}^d$. Mais il est aussi possible et intéressant de considérer un espace qui permet de travailler avec des fonctions discontinues. La classe des fonctions càdlàg est alors adéquate pour traiter ce point.

Définition 1.3.1 (Fonction càdlàg) Une fonction $f : [0, T] \rightarrow \mathbb{R}^d$ est dite càdlàg si elle est continue à droite et possède des limites à gauche. Pour tout $t \in [0, T]$, les limites:

$$f(t_-) = \lim_{s \rightarrow t, s < t} f(s), \quad f(t_+) = \lim_{s \rightarrow t, s > t} f(s)$$

existent et $f(t) = f(t_+)$.

On remarquera que toute fonction continue est une càdlàg, mais que les fonctions càdlàg peuvent présenter des discontinuités. Si t est un point de discontinuité, nous notons par:

$$\Delta f(t) = f(t) - f(t_-)$$

le saut de f en t . Cependant, une fonction càdlàg ne peut pas sauter de façon trop importante: elle n'admet qu'un nombre fini de sauts sur tout intervalle compact, c'est-à-dire que l'ensemble $\{t \in [0, T], f(t) \neq f(t_-)\}$ est dénombrable. Ainsi, pour $\varepsilon > 0$, le nombre de discontinuités supérieures à ε (grands sauts) est fini. Si cela n'était pas le cas, les discontinuités auraient une limite à droite ou à gauche, ce qui contredit la définition d'une fonction càdlàg. En conséquence, le nombre de discontinuités inférieures à ε (petits sauts) peut être infini. Un exemple de fonction càdlàg est donné par:

$$f(t) = g(t) + \sum_{i=0}^{n-1} f_i \mathbf{1}_{[t_i; t_{i+1}]}(t)$$

où on interprète la fonction g comme la composante continue de f à laquelle est ajoutée une composante discontinue de sauts.

Remarque 1.3.2 (Càdlàg ou càglàd?) Dans l'exemple précédent, la fonction f est continue à droite à l'instant de saut t_i car nous avons défini $f(t_i) :=$

$f(t_{i+})$, c'est-à-dire que f est égale à la valeur après le saut. Nous aurions pu aussi définir f comme $f(t_i) := f(t_{i-})$, c'est-à-dire qu'à l'instant t_i du saut, la fonction f est égale à la valeur avant le saut. Dans ce cas, f aurait été continue à gauche avec limites à droites, soit càglàd. La distinction est importante. En effet, si f est càdlàg et qu'un saut survient à la date t , $f(t)$ n'est pas prévisible par observation de sa trajectoire jusqu'en t , ce qui traduit bien le caractère discontinu et soudain d'un saut. Par contre, si f est càglàd, le processus décrit par f est dit prédictible.

1.3.2 Filtrations et Histoires

L'interprétation de l'indice t comme une variable de temps introduit un aspect dynamique qui nécessite d'être pris en compte et impose la nécessité de définir correctement les notions d'information, de causalité et de prédictibilité dans le contexte d'un modèle stochastique.

Dans un cadre dynamique, lorsque le temps passe, l'information dont dispose l'observateur (modélisateur) devient progressivement plus importante. Ainsi, des quantités qui étaient jugées aléatoires à la date $t = 0$ peuvent changer de statut à une date $t > 0$ si leur valeur est (partiellement) révélée par l'information disponible à la date t . En reprenant les objets que nous avons précédemment définis et en particulier l'espace de probabilité $(\Omega, \mathcal{F}, \mathbb{P})$, il apparaît nécessaire de lui ajouter une dépendance en temps pour prendre en compte ces nouvelles caractéristiques. Le concept de filtration réalise cet objectif.

Définition 1.3.3 (Filtration) Une filtration ou flux d'information sur $(\Omega, \mathcal{F}, \mathbb{P})$ est une famille croissante de σ -algèbres $(\mathcal{F}_t)_{t \geq 0} : \forall t \geq s \geq 0, \mathcal{F}_s \subseteq \mathcal{F}_t \subseteq \mathcal{F}$.

Exemple 1.3.4 (Fonctions continues) Soit $\Omega = C_0[0, T]$, l'ensemble des fonctions continues sur $[0, T]$ prenant la valeur 0 à la date $t = 0$. Soit ω^* l'une de ces fonctions, que nous observons jusqu'à la date $t \in [0, T]$, c'est-à-dire que nous connaissons la valeur de $\omega^*(s)$ pour $s \in [0, t]$, mais que nous ne la connaissons pas pour $s \in]t, T]$.

Jeu de Pile ou Face (bis)

Nous reprenons l'exemple du jeu de pile ou face, cf. Section 1.1. Nous cherchons à présent à étudier les σ -algèbres engendrées par la connaissance du résultat du lancer de plusieurs pièces (et non plus uniquement la première). Soit les ensembles:

$$A_{PP} = \{PPP, PPF\}, \quad A_{PF} = \{PFP, PFF\}$$

$$A_{FP} = \{FPP, FPF\}, \quad A_{FF} = \{FFP, FFF\}$$

et nous définissons la σ -algèbre \mathcal{F}_2 comme l'information accumulée après le lancer des deux premières pièces de la sorte:

$$\mathcal{F}_2 = \left\{ \begin{array}{l} \emptyset, \Omega, A_P, A_F, A_{PP}, A_{PF}, A_{FF}, A_{FP}, A_{PP}^c, A_{PF}^c, A_{FF}^c, A_{FP}^c \\ A_{PP} \cup A_{FP}, A_{PP} \cup A_{FF}, A_{PF} \cup A_{FP}, A_{PF} \cup A_{FF} \end{array} \right\} \quad (1.2)$$

Si, à présent, nous considérons le résultat après le lancer des trois pièces, il n'y a plus aucune incertitude et la σ -algèbre \mathcal{F}_3 correspond à tous les sous-ensembles de Ω ; par ailleurs $\mathcal{F}_0 = \{\emptyset, \Omega\}$. Ainsi, nous avons construit quatre σ -algèbres $\mathcal{F}_0, \mathcal{F}_1, \mathcal{F}_2$ et \mathcal{F}_3 indexées par le temps. Lorsque le temps passe, nous obtenons une meilleure précision des résultats du jeu de pile ou face, c'est-à-dire que pour $n < m$, nous avons $\mathcal{F}_m \subset \mathcal{F}_n$. La collection $\{\mathcal{F}_0, \mathcal{F}_1, \mathcal{F}_2, \mathcal{F}_3\}$ est un exemple typique de filtration (dans le cas discret). La Définition 1.3.3 généralise cette notion dans le cas continu.

La filtration \mathcal{F}_t s'interprète alors comme l'information connue à la date t et qui s'accroît avec le temps. Un espace de probabilité $(\Omega, \mathcal{F}, \mathbb{P})$ équipé d'une filtration est appelé un espace de probabilité filtré et on le note $(\Omega, \mathcal{F}, (\mathcal{F}_t)_{t \geq 0}, \mathbb{P})$. Comme le temps change l'information disponible et donc la connaissance des événements, la probabilité d'occurrence d'un événement aléatoire devrait alors changer avec le temps. Cependant, au lieu de changer la mesure de probabilité \mathbb{P} avec le temps, nous la conservons inchangée et modélisons l'impact de l'information en conditionnant par rapport au flux d'information \mathcal{F}_t .

Le flux d'information étant représenté par la filtration \mathcal{F}_t , il faut alors distinguer les quantités qui sont connues étant donné l'information disponible et celles qui sont toujours considérées comme aléatoires à la date t . Un événement $A \in \mathcal{F}_t$ est un événement tel qu'étant donné l'information \mathcal{F}_t à la date t , l'observateur peut décider si oui ou non A s'est réalisé. De façon similaire, une variable aléatoire \mathcal{F}_t -mesurable est une variable aléatoire dont la valeur sera révélée à la date t . Enfin, un processus dont la valeur à la date t est révélée par l'information \mathcal{F}_t est qualifié de non-anticipatif ou d'adapté. Nous commençons par définir ces notions dans le cadre d'une variable aléatoire.

Définition 1.3.5 (Histoire d'une variable aléatoire) *Soit X une variable aléatoire à valeurs dans Ω . La σ -algèbre générée par X , notée $\sigma(X)$, est l'ensemble des parties de Ω de la forme $\{X \in A\}$ pour tout $A \in \mathcal{B}(\mathbb{R})$.*

Nous pouvons à nouveau illustrer notre propos à l'aide du jeu de pile ou face avec trois pièces, cf. Section 1.1, en l'enrichissant pour mettre en avant la notion de mesurabilité.

Jeu de Pile ou Face (ter)

Nous définissons l'évolution de la variable aléatoire S comme:

$$S_{n+1}(\omega) = \begin{cases} 2S_n(\omega) & \text{si } \omega_{n+1} = F \\ \frac{1}{2}S_n(\omega) & \text{si } \omega_{n+1} = P \end{cases}, n \in \mathbb{N}$$

Il s'agit d'un modèle simple de description de l'évolution du prix d'un actif financier; on parle aussi de modèle binomial (cf. Cox-Ross-Rubinstein, 1979); Soit:

$$\begin{aligned} S_2(FFF) &= S_2(FFP) = 16 \\ S_2(FPF) &= S_2(FPP) = S_2(PFF) = S_2(PFP) = 4 \\ S_2(PPF) &= S_2(PPP) = 1 \end{aligned}$$

Soit B l'ensemble contenant le chiffre 16, alors $\{S_2 \in B\} = \{FFF, FFP\} = A_{FF}$. Par suite, $A_{FF} \in \sigma(S_2)$. Nous pouvons faire de même pour les valeurs 4 et 1 et nous avons alors respectivement $(A_{FP} \cup A_{PF}) \in \sigma(S_2)$ et $A_{PP} \in \sigma(S_2)$. Par suite, si $B = [4, 16]$, alors $(A_{FF} \cup A_{FP} \cup A_{PF}) \in \sigma(S_2)$. Par ailleurs, tout ensemble de $\sigma(S_2)$ est contenu dans la σ -algèbre \mathcal{F}_2 définie par 1.2, qui est l'information après les deux premiers lancers. D'un autre côté, A_{FP} et A_{PF} sont contenus dans \mathcal{F}_2 et seulement leur union appartient à $\sigma(S_2)$. Cela est dû au fait que la connaissance des deux premiers lancers nous permet de faire la différence entre "face suivi de pile" et "pile suivi de face", alors que la seule connaissance de S_2 ne nous le permet pas. Ainsi, \mathcal{F}_2 contient suffisamment d'information pour déterminer la valeur de S_2 et plus encore. Nous dirons alors que S_2 est \mathcal{F}_2 -mesurable.

Cela nous amène alors à la notion de mesurabilité.

Définition 1.3.6 (Mesurabilité) *Soit X une variable aléatoire à valeurs dans Ω et soit \mathcal{F} une σ -algèbre de sous-ensembles de Ω . Si tout ensemble de $\sigma(X)$ est aussi contenu dans \mathcal{F} , alors la variable aléatoire X est dite \mathcal{F} -mesurable.*

Dit autrement, la variable aléatoire X est \mathcal{F} -mesurable si et seulement si l'information contenue dans \mathcal{F} est suffisante pour déterminer la valeur de X . Par ailleurs, si X est \mathcal{F} -mesurable, alors pour toute fonction borélienne f , la variable aléatoire $f(X)$ est également \mathcal{F} -mesurable. Nous étendons cette définition aux processus stochastiques.

Définition 1.3.7 (Processus adapté) *Un processus $(X_t)_{t \geq 0}$ est dit non-anticipatif par rapport à la structure d'information $(\mathcal{F}_t)_{t \geq 0}$ ou \mathcal{F}_t -adapté si pour tout $t \geq 0$, la valeur de X_t est révélée à la date t : la variable aléatoire X_t est \mathcal{F}_t -mesurable.*

Si la seule information disponible pour décrire un processus X sont ses valeurs passées, alors l'information est résumée par l'histoire (ou filtration naturelle) de X .

Définition 1.3.8 (Filtration naturelle) *L'histoire d'un processus X est représentée par le flux d'information $(\mathcal{F}_t^X)_{t \geq 0}$ où \mathcal{F}_t^X est la σ -algèbre générée par les valeurs passées du processus, complétée des ensembles nuls:*

$$\mathcal{F}_t^X = \sigma(X_s, s \in [0, t]) \vee \mathcal{N}, \quad t \geq 0$$

Il est possible de voir \mathcal{F}_t^X comme toute l'information qu'il a été possible d'extraire de l'observation des trajectoires de X entre les dates 0 et t .

Remarque 1.3.9 (Théorie du filtrage) *La filtration naturelle d'un processus X , soit \mathcal{F}^X , revêt une importance particulière dans la théorie du filtrage. Par filtrage, il faut comprendre estimation (à variance minimale au sens des moindres carrés) des valeurs d'un processus Y non-observé (l'état) par les seules observations du processus X (la mesure) lorsqu'il existe une relation fonctionnelle reliant X à Y , cf. Partie II.*

1.3.3 Espérance Conditionnelle

Etant donné un espace de probabilité $(\Omega, \mathcal{F}, \mathbb{P})$ et une variable aléatoire X , la meilleure prédiction de X est donnée par son espérance $\mathbb{E}[X]$. Mais si nous disposons d'une information additionnelle sur les réalisations des événements possibles de $(\Omega, \mathcal{F}, \mathbb{P})$, il est alors possible de donner une meilleure prédiction de X . La meilleure prédiction, au sens de la norme L^2 , est donnée par l'espérance conditionnelle.

Définition 1.3.10 (Espérance conditionnelle) *Soit une variable aléatoire X définie sur $(\Omega, \mathcal{F}, \mathbb{P})$ et une sous σ -algèbre $\mathcal{G} \subset \mathcal{F}$. La variable aléatoire Z est appelée l'espérance conditionnelle de X sachant \mathcal{G} , notée $Z = \mathbb{E}[X|\mathcal{G}]$ et est unique si:*

1. Z est \mathcal{G} -mesurable.
2. $\mathbb{E}[YX] = \mathbb{E}[YZ]$ pour toute variable aléatoire Y qui est \mathcal{G} -mesurable.

L'espérance conditionnelle a plusieurs propriétés intéressantes.

Proposition 1.3.11 *L'espérance conditionnelle d'une v.a. X sur $(\Omega, \mathcal{F}, \mathbb{P})$ satisfait:*

1. (Linéarité) $\mathbb{E}[\alpha_1 X_1 \pm \alpha_2 X_2] = \mathbb{E}[\alpha_1 X_1] \pm \mathbb{E}[\alpha_2 X_2]$.
2. (Suffisance) Si Y est \mathcal{G} -mesurable, alors $\mathbb{E}[YX|\mathcal{G}] = Y\mathbb{E}[X|\mathcal{G}]$.
3. (Loi des espérances itérées) Soit $\mathcal{G}_0 \subset \mathcal{G}_1 \subset \mathcal{F}$, alors $\mathbb{E}[X|\mathcal{G}_0] = \mathbb{E}[\mathbb{E}[X|\mathcal{G}_1]|\mathcal{G}_0]$ et en particulier $\mathbb{E}[X] = \mathbb{E}[X|\mathcal{G}]$ pour tout sous σ -algèbre \mathcal{G} .
4. (Inégalité de Jensen) Si $\varphi(x)$ est une fonction convexe en la variable x et X est intégrable, alors $\mathbb{E}[\varphi(X)|\mathcal{G}] \geq \varphi(\mathbb{E}[X|\mathcal{G}])$.

1.3.4 Temps Aléatoire

Il arrive assez souvent que l'on doit prendre en compte dans la modélisation des événements qui arrivent à des dates aléatoires. Un temps aléatoire est alors une variable aléatoire positive $T \geq 0$ qui représente l'instant où un événement doit se produire. Soit $\tau \in [0, \infty]$ l'instant de survenance de cet événement. Une question naturelle que nous pouvons nous poser est: la connaissance de la filtration $(\mathcal{F}_t)_{t \geq 0}$ est-elle suffisante pour savoir si un événement est arrivé ($\tau \leq T$) ou pas ($\tau > T$)? Si la réponse est positive (suffisance de la filtration), alors le temps aléatoire τ est appelé un temps aléatoire non-anticipatif ou temps d'arrêt. On dit ainsi que τ est un \mathcal{F}_t -temps d'arrêt si:

$$\{T \leq t\} \in \mathcal{F}_t, \forall t > 0$$

On remarquera que si T_1 et T_2 sont deux temps d'arrêt, alors $T_1 \wedge T_2 = \min\{T_1, T_2\}$ est aussi un temps d'arrêt. Etant donné un temps d'arrêt τ et un processus $(X_t)_{t \geq 0}$ non-anticipatif, nous pouvons définir un nouveau processus $X_{t \wedge \tau}$ qui est le processus X stoppé en τ , comme:

$$X_{t \wedge \tau} = X_t \text{ si } t < \tau, \quad X_{t \wedge \tau} = X_\tau \text{ si } t \geq \tau$$

Remarque 1.3.12 (Défaut) *Les temps aléatoires permettent la modélisation d'évènements rares ou extrêmes, c'est-à-dire des évènements à probabilité faible. Leur emploi est ainsi particulièrement utile dans le cadre des modèles de défaillance (entreprise) ou de défaut (insolvabilité d'un émetteur d'actifs financiers).*

1.3.5 Martingales et Propriété de Markov

Considérons à présent un espace de probabilité $(\Omega, \mathcal{F}, \mathbb{P})$ et un flux d'information $(\mathcal{F}_t)_{t \geq 0}$. La notion de martingale va permettre de préciser l'idée de prise en compte de l'information et de conditionnement par rapport à une filtration. Elle va en outre permettre de modéliser le caractère aléatoire d'un phénomène, ainsi que son évolution dans le temps. Finalement, nous définirons la notion connexe de processus (stochastique) markovien.

Définition 1.3.13 (Martingale) *Un processus càdlàg $(X_t)_{t \geq 0}$ est une martingale si X est non-anticipatif (\mathcal{F}_t -adapté), $\mathbb{E}[|X_t|]$ est finie pour tout $t \geq 0$ et:*

$$\mathbb{E}[X_s | \mathcal{F}_t] = X_t, \quad \forall s > t$$

De façon similaire, nous définissons une sur-martingale et une sous-martingale X comme:

$$\mathbb{E}[X_s | \mathcal{F}_t] \geq X_t, \quad \forall s > t$$

$$\mathbb{E}[X_s | \mathcal{F}_t] \leq X_t, \quad \forall s > t$$

si le processus n'a respectivement pas tendance à diminuer (soit plutôt à augmenter) et à augmenter (soit plutôt à diminuer).

Remarque 1.3.14 (Martingales et analyse) *Si nous devons faire une analogie entre les probabilités et l'analyse, nous pourrions dire qu'une martingale correspond à une fonction constante, tandis qu'une sur-martingale et une sous-martingale s'assimilent à une fonction respectivement croissante et décroissante.*

Dit autrement, la meilleure prédiction de la valeur future d'une martingale est sa valeur présente. Un exemple typique de martingale est le processus de Wiener $(W_t)_{t \geq 0}$ ou mouvement brownien, cf. Section 2.1. Nous pouvons maintenant nous demander comment construire une martingale. Nous présentons un exemple typique de construction.

Exemple 1.3.15 *Soit une variable aléatoire H dont la valeur est uniquement connue à la date T , on dit alors que T est \mathcal{F}_T -mesurable, avec $\mathbb{E}[|H|] < \infty$. Le processus $(M_t)_{t \geq 0}$ défini par $M_t = \mathbb{E}[H | \mathcal{F}_t]$ est une martingale.*

La définition d'une martingale a une conséquence importante. Une martingale est d'espérance constante: $\forall t \in [0, T], \mathbb{E}[X_t] = \mathbb{E}[X_0]$. Toutefois, un processus sans moyenne n'est pas forcément une martingale.

Remarque 1.3.16 (Martingales et jeux de hasard) *Grâce à cette interprétation, il est possible de dire que si X_t représente la richesse d'un joueur à la*

date t , alors une martingale (respectivement sur-martingale, sous-martingale) correspond à un jeu équitable (respectivement favorable, défavorable), du point de vue du joueur (et l'inverse pour son adversaire, par exemple la banque au casino). Ce que nous pouvons traduire de la sorte:

$$\mathbb{E}[X_n - X_{n-1} | \mathcal{F}_{n-1}] = 0$$

où $\Delta X_n = X_n - X_{n-1}$ est le gain après le n -ième jeu. Soit, même après avoir observé le jeu jusqu'au $(n-1)$ -ième jeu, le gain espéré est toujours nul.

Nous introduisons par ailleurs la définition d'une martingale de carré intégrable comme une martingale X qui vérifie la condition d'intégrabilité suivante:

$$\mathbb{E} \left[\int_0^t |X_t|^2 dt \right] < \infty, \quad t \in [0, T]$$

Processus de Markov

Nous définissons à présent un processus de Markov (ou markovien).

Définition 1.3.17 (Processus de Markov) Soit $(\mathcal{F}_t)_{t \geq 0}$ une filtration composée de σ -algèbres de \mathcal{F} et $(X_t)_{t \geq 0}$ un processus adapté. Si, pour toute fonction borélienne non-négative f , il existe une autre fonction borélienne g telle que:

$$\mathbb{E}[f(X_t) | \mathcal{F}_s] = g(X_s)$$

alors, on dit que X est un processus de Markov. Soit, en explicitant la dépendance en temps:

$$\mathbb{E}[f(t, X_t) | \mathcal{F}_s] = f(s, X_s)$$

Remarque 1.3.18 (Processus de Markov et EDP) La définition précédente implique que $f(t, x)$ satisfait une équation aux dérivées partielles (EDP). Cette équation fournit un moyen commode de déterminer $f(s, x)$ par la connaissance de $f(t, x)$. L'EDP d'évaluation dite de Black-Scholes-Merton (1973) est un exemple de cette technique, cf. Chapitre 4.

2

Processus de Wiener, Poisson et Lévy

Ce Chapitre se concentre sur la présentation des processus stochastiques en temps continu élémentaires pour la construction de modèles financiers plus complexes de description des actifs et d'évaluation de produits conditionnels. Une attention toute particulière sera ainsi apportée à la présentation des processus de Wiener, de Poisson et de Lévy.

2.1. Processus de Wiener

Le processus de Wiener, ou mouvement brownien, est un exemple typique de processus stochastique à trajectoires continues. Il sera par la suite employé pour modéliser des processus de type marche aléatoire en temps continu.

2.1.1 Loi Gaussienne

Une variable aléatoire Y est dite suivre une loi gaussienne à valeurs dans \mathbb{R} de paramètres de moyenne m et de dispersion $\sigma > 0$ et est notée $\mathcal{N}(m, \sigma)$ si sa densité de probabilité s'écrit sous la forme:

$$f_Y(y) = \frac{1}{\sigma\sqrt{2\pi}} \exp\left[-\frac{(y-m)^2}{2\sigma^2}\right] \mathbf{1}_{y \in \mathbb{R}}$$

La fonction de répartition de Y est donnée par:

$$F_Y(y) = \mathbb{P}(Y \leq y) = \mathbb{P}\left(Y \leq \frac{y-\mu}{\sigma}\right) = \Phi\left(\frac{y-\mu}{\sigma}\right), \quad \forall y \in \mathbb{R}$$

où Φ est la fonction de répartition de la loi normale réduite ou $\mathcal{N}(0, 1)$, donnée par:

$$\Phi(y) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^y \exp\left(-\frac{u^2}{2}\right) \mathbf{1}_{y \in \mathbb{R}} du$$

Nous introduisons par ailleurs deux quantités importantes dans la caractérisation de la loi normale, cf Section 1.2.3. Les transformées de Fourier et de

Laplace d'une loi normale $X \sim \mathcal{N}(m, \sigma^2)$ sont respectivement données par:

$$\mathbb{E}[e^{itX}] = \exp\left(imt - \frac{1}{2}\sigma^2 t^2\right), \quad \mathbb{E}[e^{\alpha X}] = \exp\left(m\alpha + \frac{1}{2}\sigma^2 \alpha^2\right)$$

La loi normale est reliée de façon importante à une autre loi largement rencontrée dans l'étude du mouvement brownien (géométrique): la loi log-normale. Une variable aléatoire X est dite suivre une loi log-normale à valeurs dans $]0, \infty[$ de paramètres de moyenne m et de dispersion $\sigma > 0$ si $Y = \log(X)$ suit la loi $\mathcal{N}(m, \sigma)$. La densité et la fonction de répartition de X sont respectivement données par:

$$f_X(x) = \frac{1}{\sigma\sqrt{2\pi}} \frac{1}{x} \exp\left[-\frac{1}{2}\left(\frac{\log(x) - m}{\sigma}\right)^2\right] \mathbf{1}_{]0, \infty[}(x)$$

$$F_X(x) = \Phi\left(\frac{\log(x) - m}{\sigma}\right), \quad \forall x > 0$$

Et on a $\mathbb{E}[X] = e^{m+\sigma^2/2}$ et $Var[X] = e^{2m+\sigma^2}(e^{\sigma^2} - 1)$.

Finalement, une propriété intéressante de la loi normale est sa stabilité pour l'opération de convolution: si Y_1 et Y_2 sont des variables aléatoires gaussiennes indépendantes de paramètres (m_i, σ_i) pour $i = (1, 2)$, alors la variable aléatoire $Y_1 + Y_2$ est encore une loi normale de paramètres $(m_1 + m_2, \sigma_1 + \sigma_2)$.

Propriétés de Skewness et de Kurtosis

La Figure 2.1 représente les densités de probabilités de variables aléatoires gaussiennes diverses.

La première densité est celle de la loi normale centrée réduite $\mathcal{N}(0, 1)$: nous notons la forme en cloche classique de cette distribution et le fait qu'elle soit centrée autour de sa moyenne. Les autres densités présentent des propriétés bien différentes et caractéristiques des propriétés statistiques des séries financières, cf. Annexe A. La seconde densité a une forme évasée, avec un sommet large et relativement plat: elle est illustrative du phénomène de queues épaisses. Celui-ci témoigne de l'importance des réalisations situées dans les bandes extérieures de la distribution et qui ont donc les plus faibles probabilités de survenance. Ce phénomène est caractéristique d'une non-normalité de la série de rendements et d'événements extrêmes. Les troisième et quatrième densités caractérisent deux autres phénomènes de la distribution gaussienne liés à ses moments centrés d'ordre trois et quatre, respectivement le skewness et le kurtosis :

- Le skewness est une mesure de l'asymétrie d'une distribution autour de sa moyenne. Si $skew < 0$, la série de données est étalée sur la gauche, tandis que si $skew > 0$, la série est étalée sur la droite. Le skewness de la loi gaussienne a une valeur de 0.

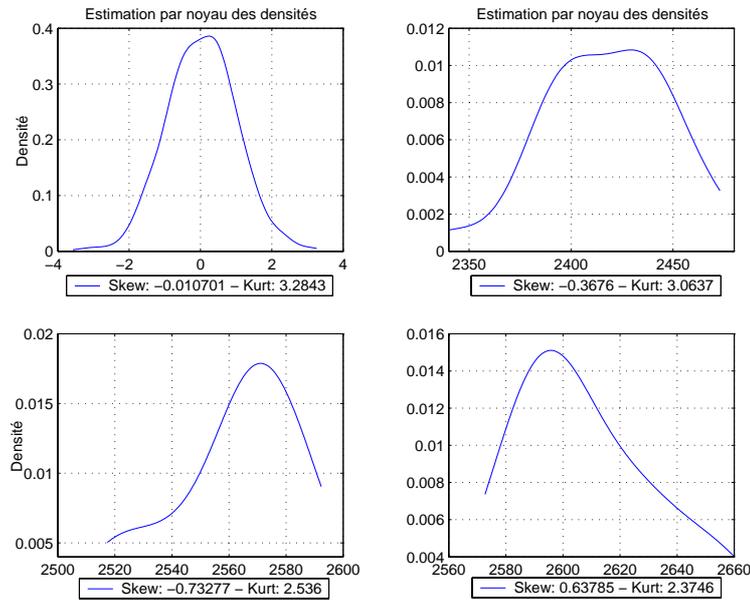


Figure 2.1: Estimation par noyau de séries de rendements.

- Le kurtosis fournit une mesure de la sensibilité d'une distribution aux points aberrants. Une valeur positive indique une distribution peu sensible aux points aberrants et se caractérise par une représentation relativement piquée. À l'inverse, une valeur négative est caractéristique d'une distribution relativement plate et donc plus sensible aux points aberrants et qui présentera alors un phénomène de queues épaisses. Le kurtosis de la loi gaussienne a une valeur de 3.

Remarque 2.1.1 (Extrêmes et queues épaisses) *Les événements extrêmes sont tels que leur probabilité d'occurrence est faible mais leur réalisation peut être importante (krach boursier, désastre naturel). Ils se situent ainsi dans les queues de distribution. Pour les prendre en compte, la loi gaussienne n'est vraisemblablement pas la plus adaptée. On lui préférera des lois dites α -stables ou à queues épaisses telles que les loi de Gumbel ou de Fréchet.*

2.1.2 Mouvement Brownien

La construction du mouvement brownien en dimension un peut se faire en considérant l'exemple de la trajectoire d'une particule soumise à des chocs aléatoires (marche aléatoire). Le mouvement brownien apparaît ainsi comme la loi limite atteinte par la particule¹.

¹Il s'agit de la construction historique du mouvement brownien telle que découverte par le biologiste anglais Robert Wiener à la fin du 19^{ème} siècle pour décrire le mouvement aléatoire de grains de pollen.

Soit une particule qui est sujette à une séquence de perturbations aléatoires et identiquement distribuées ξ_1, ξ_2, \dots suivant une distribution de Bernoulli avec $\mathbb{P}[\xi_i = \pm 1] = \frac{1}{2}$ de taille $\delta > 0$ sur des intervalles de temps de taille $h > 0$. Après la survenance de n mouvements aléatoires de cette particule, sa position est donnée par $h \sum_{k=1}^n \xi_k(\omega)$ et de façon plus générale, la position de cette particule à la date $t > 0$ est:

$$S_t(\omega) = h \sum_{k=1}^{[t/\delta]} \xi_k(\omega), \quad 0 < t < \infty$$

La Figure 2.2 présente une marche aléatoire sur cinq étapes. Le processus S est à trajectoires continues à droite et à incréments stationnaires et indépendants (du fait de l'indépendance des variables aléatoires ξ_j). Alors:

$$\mathbb{E}[S_t] = 0$$

$$\text{Var}[S_t] = h^2 \frac{t}{\delta} \simeq \frac{h^2}{\delta} t$$

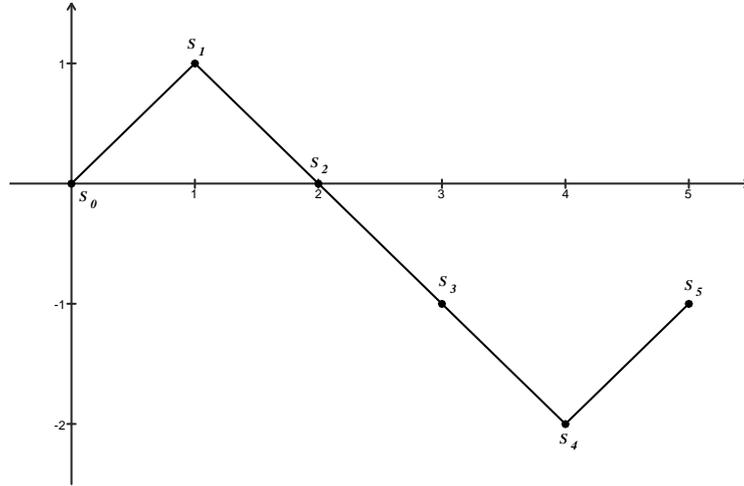


Figure 2.2: Cinq étapes d'une marche aléatoire.

Nous souhaitons à présent obtenir une version continue de ce processus, c'est-à-dire observer le comportement de S à la limite quand $h \downarrow 0$ et $\delta \downarrow 0$. Mais en même temps nous souhaitons conserver une variance positive et finie pour ce nouveau processus. Si nous prenons $h^2 = \sigma^2 \delta$ pour une constante finie $\sigma > 0$ et si $\delta_n = 1/n$ et $h_n = \sigma/\sqrt{n}$, alors:

$$S_t^{(n)}(\omega) \triangleq \frac{\sigma}{\sqrt{n}} \sum_{k=1}^{[nt]} \xi_k(\omega), \quad 0 \leq t \leq \infty, \quad n \geq 1$$

Nous introduisons à présent une nouvelle quantité: la variation quadratique, qui est reliée au concept de volatilité réalisée, cf. Section 3.2, notée $[\cdot, \cdot]$, et qui est définie de la sorte.

Définition 2.1.2 (Variation quadratique) Soit $f(t)$ une fonction définie pour $t \in [0, T]$. La variation quadratique de f jusqu'en T est:

$$[f, f]_T = \lim_{\|\Pi\| \rightarrow 0} \sum_{j=0}^{n-1} [f(t_{j+1}) - f(t_j)]^2$$

où $\Pi = \{t_0, t_1, \dots, t_n\}$ est une partition de $[0, T]$ avec $0 = t_0 < t_1 < \dots < t_n = T$.

Pour la marche aléatoire $S_t^{(n)}$, nous avons:

$$\begin{aligned} [S^{(n)}, S^{(n)}]_t &= \sum_{k=1}^{[nt]} \left[S^{(n)}\left(\frac{k}{n}\right) - S^{(n)}\left(\frac{k-1}{n}\right) \right]^2 \\ &= \sum_{k=1}^{[nt]} \left[\frac{\sigma}{\sqrt{n}} \xi_k \right]^2 = \sum_{k=1}^{[nt]} \frac{\sigma^2}{n} = \sigma^2 t \end{aligned}$$

On remarquera qu'il s'agit d'un calcul trajectoire par trajectoire et non pas sur toutes les trajectoires possibles (à l'inverse du calcul de la variance) et qu'il doit donc en principe dépendre de la trajectoire particulière sur laquelle nous calculons la variation quadratique. Toutefois, pour la marche aléatoire $S_t^{(n)}$, nous obtenons toujours $\sigma^2 t$.

Distribution Limite

Il est par ailleurs possible de montrer que la distribution limite de $S_t^{(n)}$ converge vers la distribution d'une variable aléatoire gaussienne de moyenne 0 et de variance $\sigma^2 t$. Le théorème de la limite centrée (*central limit*) précise cette approximation.

Théorème 2.1.3 (TCL) Pour $t \geq 0$ et lorsque $n \rightarrow +\infty$, nous avons:

$$\left\{ M_t^{(n)} \right\}_{n=1}^{\infty} \xrightarrow{\mathcal{L}} \mathcal{N}(0, t), \quad M_t^{(n)}(\omega) \triangleq \frac{1}{\sqrt{n}} \sum_{k=1}^{[nt]} \xi_k(\omega)$$

Idée. Il est possible d'identifier une distribution par sa fonction génératrice des moments, cf. Section 1.2.3. Pour la loi normale $\mathcal{N}(0, t)$, sa densité s'écrit:

$$f(x) = \frac{1}{\sqrt{2\pi t}} e^{-\frac{x^2}{2t}}$$

et sa fonction génératrice des moments:

$$\varphi(u) = \int_{\mathbb{R}} e^{ux} f(x) dx = e^{\frac{1}{2}u^2 t}$$

De suite, la fonction génératrice des moments de $M_t^{(n)}$ est:

$$\varphi_n(u) = \mathbb{E} \left[e^{u M_t^{(n)}} \right] = \mathbb{E} \left[\exp \left(\frac{u}{\sqrt{n}} \sum_{k=1}^{[nt]} \xi_k \right) \right] = \mathbb{E} \left[\prod_{k=1}^{[nt]} \exp \left(\frac{u}{\sqrt{n}} \xi_k \right) \right] \quad (2.1)$$

Et comme les variables aléatoires ξ sont indépendantes, 2.1 se réécrit:

$$\prod_{k=1}^{[nt]} \mathbb{E} \left[\exp \left(\frac{u}{\sqrt{n}} \xi_k \right) \right] = \prod_{k=1}^{[nt]} \left(\frac{1}{2} e^{\frac{u}{\sqrt{n}}} + \frac{1}{2} e^{-\frac{u}{\sqrt{n}}} \right) = \left(\frac{1}{2} e^{\frac{u}{\sqrt{n}}} + \frac{1}{2} e^{-\frac{u}{\sqrt{n}}} \right)^{[nt]}$$

Nous devons à présent montrer que lorsque $n \rightarrow \infty$:

$$\varphi_n(u) = \left(\frac{1}{2} e^{\frac{u}{\sqrt{n}}} + \frac{1}{2} e^{-\frac{u}{\sqrt{n}}} \right)^{[nt]} \rightarrow \varphi(u) = e^{\frac{1}{2}u^2t}$$

ou de façon équivalente:

$$\log \varphi_n(u) = [nt] \log \left(\frac{1}{2} e^{\frac{u}{\sqrt{n}}} + \frac{1}{2} e^{-\frac{u}{\sqrt{n}}} \right) \rightarrow \log \varphi(u) = \frac{1}{2}u^2t$$

Soit alors le changement de variable $x = \frac{1}{\sqrt{n}}$ tel que:

$$\lim_{n \rightarrow \infty} \log \varphi_n(u) = t \lim_{x \downarrow 0} \frac{1}{x^2} \log \left(\frac{1}{2} e^{ux} + \frac{1}{2} e^{-ux} \right)$$

Cette limite peut être calculée par la règle de L'Hôpital en notant que:

$$\frac{\partial}{\partial x} \log \left(\frac{1}{2} e^{ux} + \frac{1}{2} e^{-ux} \right) = \frac{\frac{u}{2} e^{ux} - \frac{u}{2} e^{-ux}}{\frac{1}{2} e^{ux} + \frac{1}{2} e^{-ux}}, \quad \frac{\partial}{\partial x} x^2 = 2x$$

Soit:

$$\lim_{n \rightarrow \infty} \log \varphi_n(u) = t \lim_{x \downarrow 0} \frac{1}{2x} \frac{\frac{u}{2} e^{ux} - \frac{u}{2} e^{-ux}}{\frac{1}{2} e^{ux} + \frac{1}{2} e^{-ux}} = \frac{t}{2} \lim_{x \downarrow 0} \frac{1}{x} \left(\frac{u}{2} e^{ux} - \frac{u}{2} e^{-ux} \right)$$

où $\lim_{x \downarrow 0} \left(\frac{1}{2} e^{ux} + \frac{1}{2} e^{-ux} \right) = 1$. Une autre application de la règle de L'Hôpital avec:

$$\frac{\partial}{\partial x} \left(\frac{u}{2} e^{ux} - \frac{u}{2} e^{-ux} \right) = \frac{u}{2} e^{ux} + \frac{u^2}{2} e^{-ux}$$

donne:

$$\lim_{n \rightarrow \infty} \log \varphi_n(u) = \frac{t}{2} \lim_{x \downarrow 0} \left(\frac{u}{2} e^{ux} + \frac{u^2}{2} e^{-ux} \right) = \frac{1}{2}u^2t$$

ce qui est le résultat attendu. ■

Le théorème précédent nous permet alors d'établir le corollaire suivant.

Corollaire 2.1.4 *Nous avons:*

1. Pour $t > 0$, la suite $\left\{ S_t^{(n)} \right\}_{n=1}^{\infty}$ converge en loi vers une variable aléatoire gaussienne de moyenne 0 et de variance $\sigma^2 t$, soit:

$$\left\{ S_t^{(n)} \right\}_{n=1}^{\infty} \xrightarrow{\mathcal{L}} W_t \sim \mathcal{N}(0, \sigma^2 t)$$

2. Pour $m \geq 1$ et $0 = t_0 < t_1 < \dots < t_{m-1} < t_m < \infty$, la suite de vecteurs aléatoires:

$$\left\{ S_{t_1}^{(n)}, S_{t_2}^{(n)} - S_{t_1}^{(n)}, \dots, S_{t_m}^{(n)} - S_{t_{m-1}}^{(n)} \right\}_{n=1}^{\infty}$$

converge en loi vers un vecteur de variables aléatoires i.i.d.:

$$(W_{t_1}, W_{t_2} - W_{t_1}, \dots, W_{t_m} - W_{t_{m-1}})$$

avec:

$$W_{t_j} - W_{t_{j-1}} \sim \mathcal{N}(0, \sigma^2(t_j - t_{j-1})), \quad j \in [1, \infty]$$

Finalement, le processus $S^{(n)} = (S_t^{(n)})_{t \geq 0}$ converge en loi, quand $n \rightarrow \infty$, vers un processus $W = (W_t)_{t \geq 0}$ dont les propriétés sont résumées dans la définition.

Définition 2.1.5 (Mouvement brownien) *Un processus adapté $W = (W_t)_{t \geq 0}$ à valeurs dans \mathbb{R} est un mouvement brownien standard si:*

1. Pour $0 \leq s < t < \infty$, $W_t - W_s$ est indépendant de \mathcal{F}_s , i.e.: les incréments de W sont indépendants du passé.
2. Pour $0 < s < t$, $W_t - W_s$ est une variable aléatoire gaussienne de moyenne nulle et de variance $(t - s)$.
3. Pour $\omega \in \Omega$, l'application $t \mapsto W_t(\omega)$ est continue.

Remarque 2.1.6 *Le mouvement brownien est dit commencé à la valeur 0 si $\mathbb{P}(W_0 = 0) = 1$.*

La Figure 2.3 présente des trajectoires typiques de processus de Wiener. Nous énonçons à présent plusieurs propriétés importantes du mouvement brownien.

Proposition 2.1.7 *Le mouvement brownien est une martingale, soit:*

1. $\forall t \geq 0, \mathbb{E}[W_t] < \infty$.
2. $\forall t \geq 0, W_t$ est (\mathcal{F}_t) -mesurable.
3. $\forall s > t \geq 0$ tel que $s < t, \mathbb{E}[W_t | \mathcal{F}_s] = W_s$.

Preuve. Pour $0 \leq s \leq t$, nous avons:

$$\begin{aligned} \mathbb{E}[W_t | \mathcal{F}_s] &= \mathbb{E}[W_t - W_s + W_s | \mathcal{F}_s] \\ &= \mathbb{E}[W_t - W_s | \mathcal{F}_s] + \mathbb{E}[W_s | \mathcal{F}_s] \\ &= \mathbb{E}[W_t - W_s] + W_s = W_s \end{aligned}$$

du fait de la mesurabilité de W_s , ce qui conclut la preuve. ■

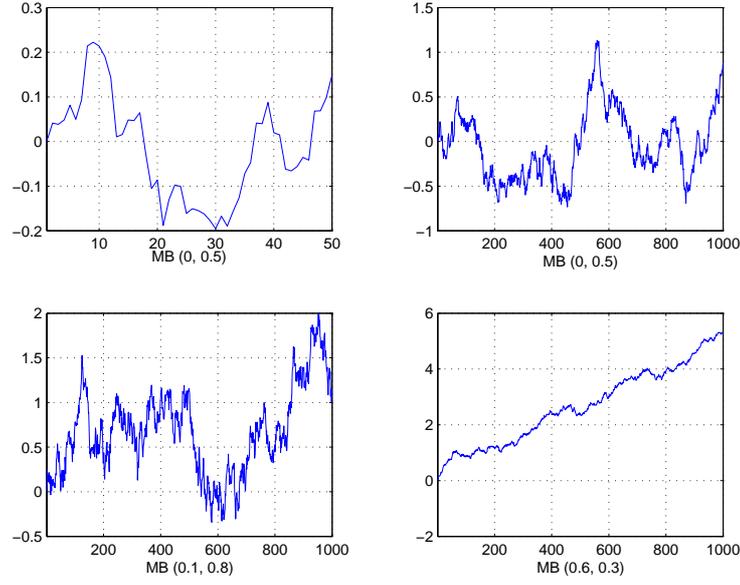


Figure 2.3: Trajectoires de processus de Wiener.

Variation Quadratique

Nous nous intéressons dans cette section au calcul de la variation quadratique, cf. Définition 2.1.2, pour un mouvement brownien.

Théorème 2.1.8 *Soit W un mouvement brownien. Alors $[W, W]_T = T$ pour tout $T > 0$.*

Preuve. Soit $\Pi = \{t_0, \dots, t_n\}$ une partition de $[0, T]$. Nous définissons:

$$Q_\Pi = \sum_{j=0}^{n-1} (W_{t_{j+1}} - W_{t_j})^2$$

qui est une variable aléatoire. Nous voulons montrer que $\mathbb{E}[Q_\Pi] = T$ et $\text{Var}[Q_\Pi] \rightarrow 0$ quand $\|\Pi\| \rightarrow 0$. Notant que:

$$\mathbb{E} \left[(W_{t_{j+1}} - W_{t_j})^2 \right] = \text{Var} [W_{t_{j+1}} - W_{t_j}] = t_{j+1} - t_j$$

nous avons:

$$\mathbb{E}[Q_\Pi] = \sum_{j=0}^{n-1} \mathbb{E} \left[(W_{t_{j+1}} - W_{t_j})^2 \right] = \sum_{j=0}^{n-1} (t_{j+1} - t_j) = T$$

comme attendu. Par ailleurs:

$$\begin{aligned} & \text{Var} \left[(W_{t_{j+1}} - W_{t_j})^2 \right] \\ &= \mathbb{E} \left[(W_{t_{j+1}} - W_{t_j})^2 - (t_{j+1} - t_j)^2 \right]^2 \\ &= \mathbb{E} \left[(W_{t_{j+1}} - W_{t_j})^4 \right] - 2(t_{j+1} - t_j)^2 \mathbb{E} \left[(W_{t_{j+1}} - W_{t_j})^2 \right] + (t_{j+1} - t_j)^2 \end{aligned}$$

Et comme:

$$\begin{aligned} \mathbb{E} \left[(W_{t_{j+1}} - W_{t_j})^4 \right] &= 3(t_{j+1} - t_j)^2 \\ \text{Var} \left[(W_{t_{j+1}} - W_{t_j})^2 \right] &= 2(t_{j+1} - t_j)^2 \end{aligned}$$

il vient:

$$\begin{aligned} \text{Var} [Q_{\Pi}] &= \sum_{j=0}^{n-1} \text{Var} \left[(W_{t_{j+1}} - W_{t_j})^2 \right] = \sum_{j=0}^{n-1} 2(t_{j+1} - t_j)^2 \\ &\leq \sum_{j=0}^{n-1} 2 \|\Pi\| (t_{j+1} - t_j) = 2 \|\Pi\| T \rightarrow 0, \quad \|\Pi\| \rightarrow 0 \end{aligned}$$

ce qui conclut la preuve. ■

Remarque 2.1.9 (Calcul stochastique) *La plupart des fonctions admettent des dérivées continues, ainsi leur variation quadratique est nulle. Dès lors, la variation quadratique n'est jamais considérée dans le "calcul ordinaire". D'un autre côté, les trajectoires du mouvement brownien ne sont pas différentiables par rapport au temps, il n'admet donc pas de dérivées. Le concept de variation quadratique permet alors de donner du sens à une "dérivée trajectoire par trajectoire" et sera ainsi très utile pour le calcul stochastique, cf. Section 3.3.*

De façon informelle, nous avons:

$$dW_t dW_t = dt, \quad dW_t dt = 0, \quad dt dt = 0$$

ce qui équivaut à dire que sur $[0, T]$, le mouvement brownien accumule T unités de variation quadratique. Par ailleurs, comme $[W, W]_{T_1} = T_1$ et $[W, W]_{T_2} = T_2$, alors $[W, W]_{T_2} - [W, W]_{T_1} = T_2 - T_1$. La variation quadratique est alors la source de la *volatilité*.

Exemple 2.1.10 (Mouvement brownien géométrique) *Comme rappelé en Section 2.1.1, il existe une relation forte entre les lois normale et log-normale dans le sens où si $W \sim \mathcal{N}(0, 1)$, alors $S = \exp(a + bW) \sim \mathcal{LN}(a, b)$. Sur cette base, nous définissons le mouvement brownien géométrique comme:*

$$S_t = S_0 \exp \left(\left(\alpha - \frac{1}{2} \sigma^2 \right) t + \sigma W_t \right) \quad (2.2)$$

où $\alpha \in \mathbb{R}$ et $\sigma > 0$. Il s'agit du modèle de Black-Scholes-Merton (1973). Si nous voulons calculer la variation quadratique de 2.2 sur l'intervalle $[T_1, T_2]$, nous définissons la partition $\Pi = \{T_1 = t_1 < \dots < t_n = T_2\}$ et nous calculons les rendements logarithmiques:

$$\log \frac{S_{t_{j+1}}}{S_{t_j}} = \left(\alpha - \frac{1}{2} \sigma^2 \right) (t_{j+1} - t_j) + \sigma (W_{t_{j+1}} - W_{t_j}), \quad j = 1, \dots, n-1$$

La somme des carrés des logs-rendements, parfois appelée la volatilité réalisée, est:

$$\begin{aligned} \sum_{j=1}^{n-1} \left(\log \frac{S_{t_{j+1}}}{S_{t_j}} \right)^2 &= \left(\alpha - \frac{1}{2} \sigma^2 \right)^2 \sum_{j=1}^{n-1} (t_{j+1} - t_j)^2 + \sigma^2 \sum_{j=1}^{n-1} (W_{t_{j+1}} - W_{t_j})^2 \\ &\quad + 2\sigma \left(\alpha - \frac{1}{2} \sigma^2 \right) \sum_{j=1}^{n-1} (W_{t_{j+1}} - W_{t_j}) (t_{j+1} - t_j) \end{aligned} \quad (2.3)$$

Considérons la partie droite de 2.3. Le premier terme vaut $(\alpha - \frac{1}{2}\sigma^2)^2$ fois la variation quadratique de t , qui vaut 0. Le second terme vaut σ^2 fois la variation quadratique du mouvement brownien sur $[T_1, T_2]$, soit $T_2 - T_1$. Le troisième terme est égal à $2\sigma(\alpha - \frac{1}{2}\sigma^2)$ fois la variation quadratique croisée de W et de t , qui est nulle. Nous concluons alors:

$$\frac{1}{T_2 - T_1} \sum_{j=1}^{n-1} \left(\log \frac{S_{t_{j+1}}}{S_{t_j}} \right)^2 \rightarrow \sigma^2, \quad \|\Pi\| \rightarrow 0 \quad (2.4)$$

L'identité 2.4 permet ainsi de calculer σ à partir des observations.

Propriété de Markov

Dans cette section, nous montrons que le mouvement brownien est un processus markovien, cf. Définition 1.3.17, et dérivons sa probabilité de transition.

Proposition 2.1.11 *Le mouvement brownien est un processus markovien, soit:*

$$\mathbb{E} [f(W_t) | \mathcal{F}_s^W] = g(W_s), \quad 0 < t < s$$

où f et g sont des fonctions boréliennes et \mathcal{F}^W est la filtration naturelle de W .

Preuve. On note:

$$\mathbb{E} [f(W_t) | \mathcal{F}_s^W] = \mathbb{E} [f(W_t - W_s) + W_s | \mathcal{F}_s^W]$$

Comme la variable aléatoire $W_t - W_s$ est indépendante de \mathcal{F}_s et que W_s est \mathcal{F}_s^W -mesurable, la règle d'indépendance s'applique. Ainsi, si nous notons $g(x) = \mathbb{E} [f(W_t - W_s + x)]$ et comme $W_t - W_s$ est gaussien $\mathcal{N}(0, t - s)$, nous avons alors:

$$g(x) = \frac{1}{\sqrt{2\pi(t-s)}} \int_{-\infty}^{+\infty} f(\omega + x) e^{-\frac{\omega^2}{2(t-s)}} d\omega \quad (2.5)$$

Par suite, si nous posons $x = W_s$, la règle d'indépendance nous permet de conclure. ■

La propriété de Markov du mouvement brownien peut être explicitée de façon plus précise en notant que par le changement de variable $\tau = t - s$ et $y = \omega + x$, 2.5 s'écrit:

$$g(x) = \frac{1}{\sqrt{2\pi\tau}} \int_{-\infty}^{+\infty} f(y) e^{-\frac{(y-x)^2}{2\tau}} dy$$

et nous définissons la probabilité de transition d'un mouvement brownien par:

$$p(\tau, x, y) = \frac{1}{\sqrt{2\pi\tau}} e^{-\frac{(y-x)^2}{2\tau}}$$

Nous avons alors:

$$g(x) = \int_{-\infty}^{+\infty} f(y) p(\tau, x, y) dy$$

et on conclut finalement:

$$\mathbb{E}[f(W_t) | \mathcal{F}_s^W] = \int_{-\infty}^{+\infty} f(y) p(\tau, W_s, y) dy$$

Cette relation s'interprète de la sorte: conditionnellement à la filtration \mathcal{F}_s , la densité conditionnelle de W_s est $p(\tau, W_s, y)$, soit une densité gaussienne de moyenne W_s et de variance $\tau = t - s$. Par ailleurs, la propriété de Markov du mouvement brownien est traduite dans le fait que la seule information contenue dans \mathcal{F}_s qui est pertinente est W_s .

Nous concluons finalement cette présentation du mouvement brownien standard en donnant deux propriétés importantes de ses trajectoires²:

- $t \rightarrow W_t(\omega)$ n'est à variation finie sur aucun intervalle,
- $t \rightarrow W_t(\omega)$ n'est dérivable en aucun point.

Et en donnant le théorème de caractérisation d'un mouvement brownien dû à Lévy.

Théorème 2.1.12 (Lévy) *Soit $(M_t)_{t \geq 0}$ une martingale telle que $M_0 = 0$, à trajectoires continues avec $[M, M]_t = 0$ pour tout $t \geq 0$. Alors M_t est un mouvement brownien.*

²Ces deux propriétés sont importantes pour la dérivation du calcul stochastique présenté au Chapitre 3.

2.1.3 Version Multidimensionnelle

Nous abordons à présent le mouvement brownien dans un cadre multidimensionnel, c'est-à-dire vectoriel. Grâce aux propriétés des vecteurs gaussiens, les définitions suivantes sont relativement faciles à obtenir.

Définition 2.1.13 (Mouvement brownien multidimensionnel) *Le mouvement brownien standard W de dimension d est une famille de variables aléatoires:*

$$\left\{ W_t = \left(W_t^{(1)}, \dots, W_t^{(d)} \right) : t \geq 0 \right\}$$

où $W_t^{(1)}, \dots, W_t^{(d)}$ représentent des mouvements browniens standard indépendants construits sur l'espace probabilisé filtré $(\Omega, \mathcal{F}, \mathbb{P})$.

L'utilité de pouvoir disposer d'une version vectorielle du mouvement brownien est évidente. Toutefois, l'hypothèse d'indépendance des chocs entre chacun des mouvements browniens est assez restrictive. C'est pourquoi nous aimerions construire un mouvement brownien multidimensionnel dont les composantes sont corrélées. Soit:

$$\Gamma = \begin{bmatrix} 1 & \gamma_{12} & \dots & \gamma_{1d} \\ \gamma_{12} & 1 & \dots & \gamma_{2d} \\ \vdots & \vdots & \ddots & \vdots \\ \gamma_{1d} & \gamma_{2d} & \dots & 1 \end{bmatrix}$$

et $W = (W^{(1)}, \dots, W^{(d)})$ un mouvement brownien standard multidimensionnel dont toutes les composantes sont indépendantes. Si nous posons:

$$\begin{pmatrix} B^{(1)} \\ \vdots \\ B^{(d)} \end{pmatrix} = \Gamma \begin{pmatrix} W^{(1)} \\ \vdots \\ W^{(d)} \end{pmatrix} = \begin{pmatrix} \sum_{k=1}^d \gamma_{1k} W^{(k)} \\ \vdots \\ \sum_{k=1}^d \gamma_{ik} W^{(k)} \\ \vdots \\ \sum_{k=1}^d \gamma_{dk} W^{(k)} \end{pmatrix}$$

Alors:

$$\begin{aligned} \text{cov} \left[B_t^{(i)}, B_t^{(j)} \right] &= t \sum_{k=1}^d \gamma_{ik} \gamma_{jk} \\ \text{cor} \left[B_t^{(i)}, B_t^{(j)} \right] &= \frac{\sum_{k=1}^d \gamma_{ik} \gamma_{jk}}{\left(\sum_{k=1}^d \gamma_{ik}^2 \right)^{1/2} \left(\sum_{k=1}^d \gamma_{jk}^2 \right)^{1/2}} \end{aligned}$$

Nous venons de montrer qu'il est possible de construire un mouvement brownien B dont les composantes sont corrélées à partir d'un mouvement brownien standard W . Plus précisément, si $B = \Gamma W$, alors nous savons trouver la matrice de corrélations de B .

Remarque 2.1.14 *Soit Σ une matrice symétrique réelle de taille (d, d) définie positive et m un vecteur de \mathbb{R}^d . Alors, la loi gaussienne sur \mathbb{R}^d $\mathcal{N}(m, \Sigma)$ existe*

et est égale à:

$$\phi_X(x) = \frac{1}{(2\pi)^{d/2} \sqrt{\det \Sigma}} \exp\left(-\frac{1}{2}(x-m)^\top \Sigma^{-1}(x-m)\right)$$

Finalement, le Théorème 2.1.12 de Lévy admet une version multivariée qui permet de caractériser un vecteur de mouvements browniens.

Théorème 2.1.15 (Lévy - Multidimensionnel) Soit $(M_t^i)_{t \geq 0, i \geq 1}$ un vecteur de martingales tel que $M_0^i = 0$, à trajectoires continues avec $[M^i, M^i]_t = 0$ pour tout $t \geq 0$ et $i \geq 1$. Si par ailleurs $[M^i, M^j]_t = 0, i \neq j$, alors les $(M_t^i)_{t \geq 0, i \geq 1}$ sont des mouvements browniens indépendants.

Exemple 2.1.16 (Actifs corrélés) En reprenant l'Exemple 2.1.10, nous définissons deux actifs S^1 et S^2 suivant chacun un MBG, soit:

$$\frac{dS_t^1}{S_t^1} = a_1 dt + b_1 dW_t^1$$

$$\frac{dS_t^2}{S_t^2} = a_2 dt + b_2 \left[\rho dW_t^1 + \sqrt{1-\rho^2} dW_t^2 \right]$$

où W_t^1 et W_t^2 sont des mouvements browniens indépendants and $\rho \in (-1, 1)$. Nous définissons:

$$W_t^3 = \rho W_t^1 + \sqrt{1-\rho^2} W_t^2$$

Par suite, W^3 est une martingale continue avec $W_0^3 = 0$ et:

$$\begin{aligned} dW_t^3 dW_t^3 &= \rho^2 dW_t^1 dW_t^1 + 2\rho\sqrt{1-\rho^2} dW_t^1 dW_t^2 + (1-\rho^2) dW_t^2 dW_t^2 \\ &= \rho^2 dt + (1-\rho^2) dt = dt \end{aligned}$$

soit $[W^3, W^3]_t = t$. Une application du Théorème 2.1.12 de Lévy nous permet alors de conclure que W^3 est un mouvement brownien et nous pouvons écrire:

$$\frac{dS_t^2}{S_t^2} = a_2 dt + b_2 dW_t^3$$

Il est par ailleurs possible de montrer que les mouvements browniens W^1 et W^3 sont corrélés. En effet, par application de la règle de multiplication, cf. Proposition 3.2.8, nous avons:

$$\begin{aligned} d(W_t^1 W_t^3) &= W_t^1 dW_t^3 + W_t^3 dW_t^1 + dW_t^1 dW_t^3 \\ &= W_t^1 dW_t^3 + W_t^3 dW_t^1 + \rho dt \end{aligned}$$

Et en intégrant, il vient:

$$W_t^1 W_t^3 = \int_0^t W_s^1 dW_s^3 + \int_0^t W_s^3 dW_s^1 + \rho t$$

et comme les intégrales d'Itô sont nulles, nous avons:

$$\mathbb{E}[W_t^1 W_t^3] = \rho t$$

et ρ est alors la corrélation entre les mouvements browniens W^1 et W^3 .

2.2. Processus de Poisson

Le processus de Poisson est un exemple typique de processus stochastique à trajectoires discontinues et sera par la suite employé pour construire des modèles à sauts.

2.2.1 Loi Exponentielle

Les propriétés des variables aléatoires exponentielles et de leur somme jouent un rôle essentiel dans la définition des processus de Markov à sauts. Une variable aléatoire Y est dite suivre une loi exponentielle à valeurs dans $[0, \infty[$ de paramètre $\lambda > 0$ et est notée $\mathcal{E}(\lambda)$ si sa densité de probabilité s'écrit sous la forme:

$$f_Y(y) = \lambda e^{-\lambda y}, \quad y \in \mathbb{R}^+$$

et $\mathbb{E}[Y] = \frac{1}{\lambda}$. La fonction de répartition de Y est alors donnée par:

$$F_Y(y) = \mathbb{P}(Y \leq y) = 1 - e^{-\lambda y}, \quad \forall y \in [0, \infty[$$

La distribution exponentielle a la propriété suivante remarquable: si T est une variable aléatoire exponentielle, alors:

$$\mathbb{P}(T > t + s | T > t) = \frac{\int_{t+s}^{+\infty} \lambda e^{-\lambda y} dy}{\int_t^{+\infty} \lambda e^{-\lambda y} dy} = \mathbb{P}(T > s) = e^{-\lambda s}, \quad \forall t, s > 0$$

Cette propriété s'interprète de la sorte: si T est assimilée à un temps aléatoire, cf. Section 1.3.4, la distribution de $T - t$ connaissant $T > t$ est la même que la distribution de T elle-même: il s'agit de la propriété d'absence de mémoire. La loi exponentielle est la seule distribution dotée de cette propriété.

Définition 2.2.1 (Absence de mémoire) Soit $T \geq 0$ une variable aléatoire telle que:

$$\mathbb{P}(T > t + s | T > t) = \mathbb{P}(T > s), \quad \forall t, s > 0$$

Alors T suit une loi exponentielle.

Remarque 2.2.2 (Défaut d'un actif) Supposons que la survenance du défaut d'un actif financier (par exemple une obligation) soit modélisée par une loi exponentielle de moyenne $1/\lambda$. Supposons en outre que nous ayons déjà attendu t unités de temps. Nous nous intéressons à présent à la probabilité de ne pas avoir à attendre s unités de temps supplémentaires avant le défaut: il s'agit de la distribution exponentielle. La propriété d'absence de mémoire traduit par ailleurs le fait qu'avoir déjà attendu s unités de temps ne modifie en rien la distribution de probabilité du temps d'attente qu'il reste encore à courir avant le défaut.

2.2.2 Loi de Poisson

Une variable aléatoire N est dite suivre une loi de Poisson de paramètre $\lambda > 0$ si

$$\forall n \in \mathbb{N}, \mathbb{P}(N = n) = e^{-\lambda} \frac{\lambda^n}{n!}$$

Il existe un lien particulier entre la loi de Poisson et la somme de variables aléatoires exponentielles indépendantes

Proposition 2.2.3 *Si $(\tau_i)_{i \geq 1}$ sont des variables aléatoires exponentielles indépendantes de paramètre λ , alors pour tout $t > 0$, la variable aléatoire:*

$$N_t = \inf \left\{ n \geq 1, S_n = \sum_{i=1}^n \tau_i > t \right\}$$

suit une loi de Poisson de paramètre $\lambda t > 0$, soit $P(N_t = n) = e^{-\lambda t} \frac{(\lambda t)^n}{n!}$, $n \in \mathbb{N}$.

Preuve. Pour $n = 0$, nous avons:

$$P(N_t = 0) = P(S_1 > t) = P(\tau_1 > t) = e^{-\lambda t}$$

Et pour $n \geq 1$, nous avons $N_t \geq n$ si et seulement si n sauts se sont produits jusqu'au temps t , i.e.: $S_n \leq t$, le temps du n -ième saut, est au plus égal à t . Soit:

$$P(N_t \geq n) = P(S_n \leq t) = \int_0^t \frac{(\lambda s)^{n-1}}{(n-1)!} \lambda e^{-\lambda s} ds$$

De même:

$$P(N_t \geq n+1) = P(S_{n+1} \leq t) = \int_0^t \frac{(\lambda s)^n}{n!} \lambda e^{-\lambda s} ds$$

soit par une intégration par parties:

$$P(N_t \geq n+1) = -\frac{(\lambda t)^n}{n!} e^{-\lambda t} + P(N_t \geq n)$$

Soit:

$$P(N_t = n) = P(N_t \geq n) - P(N_t \geq n+1) = \frac{(\lambda t)^n}{n!} e^{-\lambda t}$$

ce qui conclut la preuve. ■

Une autre propriété intéressante de la loi de Poisson est sa stabilité pour l'opération de convolution: si Y_1 et Y_2 sont des variables aléatoires de Poisson indépendantes de paramètres λ_1 et λ_2 , alors $Y_1 + Y_2$ suit une loi de Poisson de paramètre $\lambda_1 + \lambda_2$.

2.2.3 Processus de Poisson

La définition et les propriétés d'un processus de Poisson s'obtiennent par analogie des résultats obtenus précédemment pour la loi de Poisson.

Définition 2.2.4 (Processus de Poisson) Soit $(\tau_i)_{i \geq 1}$ une suite de variables aléatoires exponentielles et indépendantes de paramètre λ et $T_n = \sum_{i=1}^n \tau_i$. Le processus $(N_t)_{t \geq 0}$ défini par:

$$N_t = \sum_{n \geq 1} \mathbf{1}_{T_n \leq t}$$

est un processus de Poisson d'intensité $\lambda > 0$.

La Figure 2.4 représente une trajectoire typique d'un processus de Poisson.

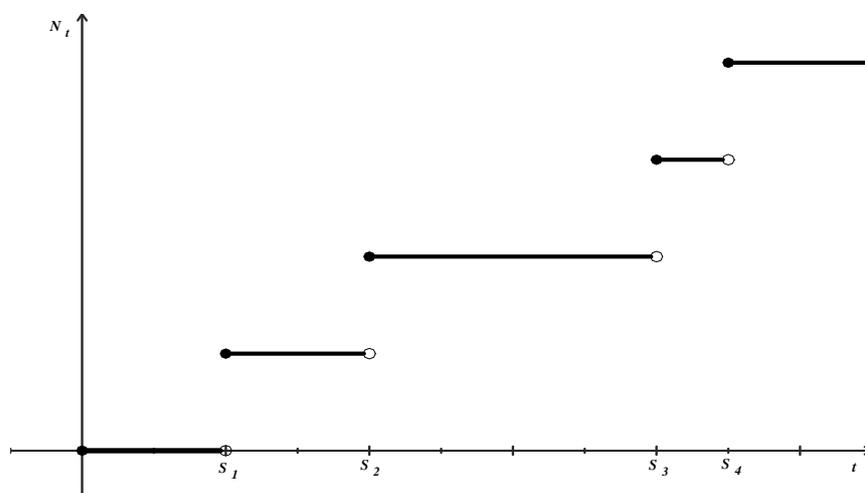


Figure 2.4: Trajectoire d'un processus de Poisson.

Processus de Comptage

Un processus de Poisson peut aussi être défini comme un processus de comptage: il compte le nombre de temps aléatoires $(T_n)_{n \geq 1}$ qui surviennent entre 0 et t , où $(T_n - T_{n-1})_{n \geq 1}$ est une séquence de variables aléatoires exponentielles indépendantes et identiquement distribuées, ce qui s'exprime sous la forme:

$$N_t = \begin{cases} 0 & , \text{ si } 0 \leq t < T_1 \\ 1 & , \text{ si } T_1 \leq t < T_2 \\ \vdots & \\ n & , \text{ si } T_n \leq t < T_{n+1} \\ \vdots & \end{cases}$$

ou de façon résumée:

$$N_t = \sum_{n \geq 1} \mathbf{1}_{T_n \leq t} = \# \{n \geq 1, t \leq T_n\} \quad (2.6)$$

où $\# \{A\} = \text{card}(A)$. Comme nous l'avons vu précédemment, si les temps aléatoires $(T_n)_{n \geq 1}$ sont construits comme la somme partielle de variables aléatoires exponentielles, alors N_t est un processus de Poisson. De façon générale, les $(T_n)_{n \geq 1}$ peuvent avoir une distribution et une structure de dépendance quelconques. Nous avons alors le résultat important suivant.

Lemme 2.2.5 *Soit N_t un processus de comptage à accroissements indépendants et stationnaires. Alors N_t est un processus de Poisson.*

Remarque 2.2.6 (Intensité du processus de Poisson) *Etant donné qu'entre la survenue de deux sauts consécutifs, le temps moyen d'attente est de $1/\lambda$, les sauts arrivent avec un taux moyen de λ par unité de temps. Soit N_t a l'intensité λ .*

Processus de Poisson Compensé

Soit la version *centrée* du processus de Poisson N_t :

$$M_t = N_t - \lambda t$$

$(M_t)_{t \geq 0}$ est alors doté de la propriété de martingale:

$$\mathbb{E}[M_t | \mathcal{F}_s] = M_s, \quad \forall t > s$$

Le processus $(M_t)_{t \geq 0}$ est appelé processus de Poisson compensé et $(\lambda t)_{t \geq 0}$ (qui n'est pas un processus) le compensateur de $(N_t)_{t \geq 0}$: il s'agit de la quantité qui soustraite à N permet de retrouver la propriété de martingale. La Figure 2.5 représente un tel processus.

De façon remarquable, nous pouvons observer que le processus de Poisson compensé M n'est pas un processus de comptage et que par ailleurs, sa version *réduite* partage les mêmes deux premiers moments que le mouvement brownien. En effet:

$$\mathbb{E} \left[\frac{M_t}{\lambda} \right] = 0, \quad \text{Var} \left[\frac{M_t}{\lambda} \right] = t$$

La similitude entre ces deux processus peut-être rendue plus précise grâce au résultat suivant:

$$\frac{M_t}{\lambda} \xrightarrow[\infty]{\lambda} W_t, \quad t \in [0, T]$$

qui est une sorte de théorème de la limite centrée sur l'espace fonctionnelle $D([0, T])$.

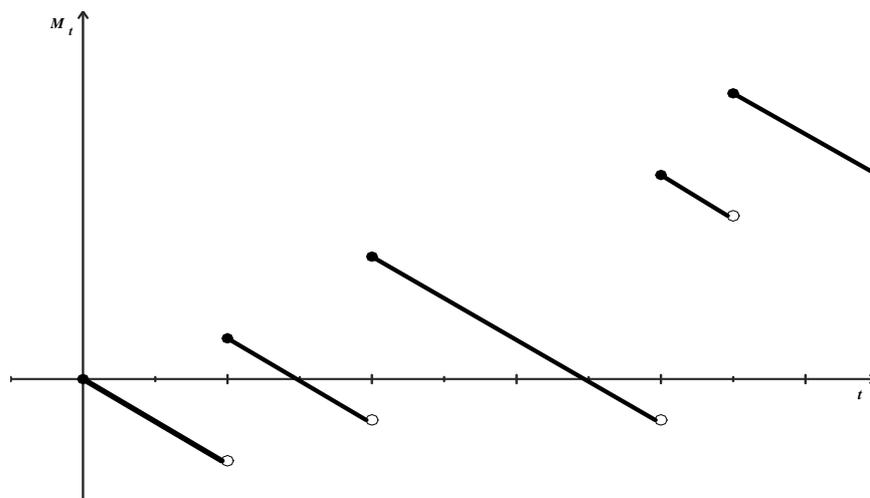


Figure 2.5: Trajectoire d'un processus de Poisson compensé.

2.2.4 Processus de Poisson Composé

Le processus de Poisson composé joue un rôle essentiel dans la construction des processus mixtes Wiener-Poisson, cf. Section 4.1.2. Il va permettre d'en représenter la partie sauts. En guise d'introduction, on remarquera que lorsqu'un processus de Poisson saute, il le fait de façon unitaire. L'intérêt d'un processus de Poisson composé est de prendre en compte des sauts de taille aléatoire. La Figure 2.6 présente une trajectoire typique d'un processus de Poisson composé où les Y_i sont les tailles aléatoires des sauts.

Définition 2.2.7 (Processus de Poisson composé) *Un processus de Poisson composé d'intensité $\lambda > 0$ et de distribution de tailles des sauts f est un processus stochastique $(X_t)_{t \geq 0}$ défini par:*

$$X_t = \sum_{i=1}^{N_t} Y_i, \quad t \geq 0$$

où les tailles des sauts Y_i sont des variables aléatoires i.i.d. de distribution f d'espérance $\bar{\mu} = \mathbb{E}[Y_i]$ et $(N_t)_{t \geq 0}$ est un processus de Poisson d'intensité λ , indépendant de $(Y_i)_{i \geq 1}$.

Nous avons alors les propriétés suivantes pour un processus de Poisson composé:

1. Les tailles des sauts $(Y_i)_{i \geq 1}$ sont i.i.d. de distribution f .
2. Les trajectoires de X sont des fonctions càdlàg constantes par morceaux.
3. Les instants de sauts $(T_i)_{i \geq 1}$ ont la même loi que les instants de sauts du processus de Poisson N_t : sommes de variables aléatoires i.i.d. exponentielles de paramètre λ .

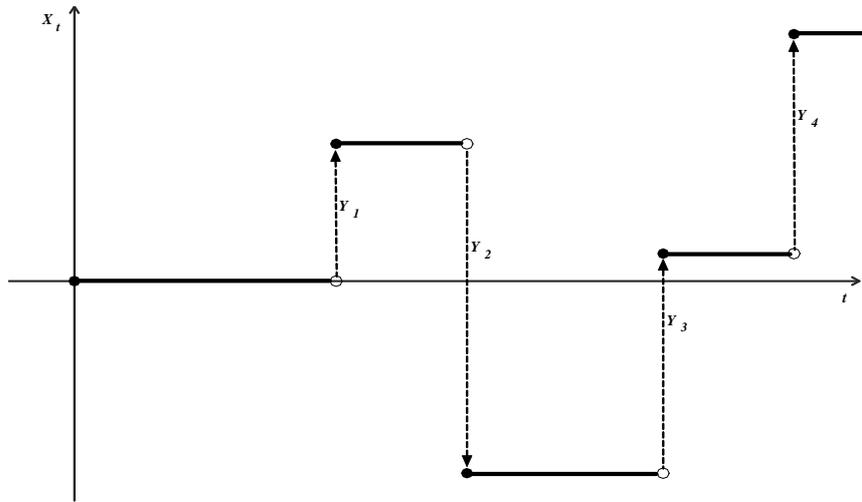


Figure 2.6: Trajectoire d'un processus de Poisson composé.

Par suite, un processus de Poisson composé compensé est défini de la sorte:

$$\begin{aligned}\bar{X}_t &= X_t - \lambda \mathbb{E}[Y_0] t \\ &= X_t - \lambda \bar{\mu} t\end{aligned}$$

Pour conclure, la Figure 2.7 présente une trajectoire d'un processus de Poisson composé avec des sauts gaussiens. L'intensité des sauts λ varie entre 10 et 1000.

2.3. Processus de Lévy

Les processus de Lévy sont des processus continus, à accroissements indépendants et stationnaires, analogues aux marches aléatoires discrètes qui sont des sommes de variables aléatoires indépendantes et identiquement distribuées. Ils fournissent un cadre général pratique pour la construction de processus stochastiques en temps continu dont les processus de Wiener et de Poisson sont des cas particuliers. Par ailleurs, comme nous allons le voir, ils constituent des blocs essentiels dans la construction de processus de Lévy, ces derniers étant en effet la *superposition* d'un processus de Wiener et de processus de Poisson indépendants.

Remarque 2.3.1 (Processus de Lévy à taux d'activité fini) *Dans la suite, nous ne présentons pas les processus de Lévy en toute généralité. Nous ne discutons qu'une classe particulière parmi cette famille, celle des processus de Lévy à taux d'activité fini; ces processus admettent un nombre fini de sauts sur tout intervalle fini de temps.*

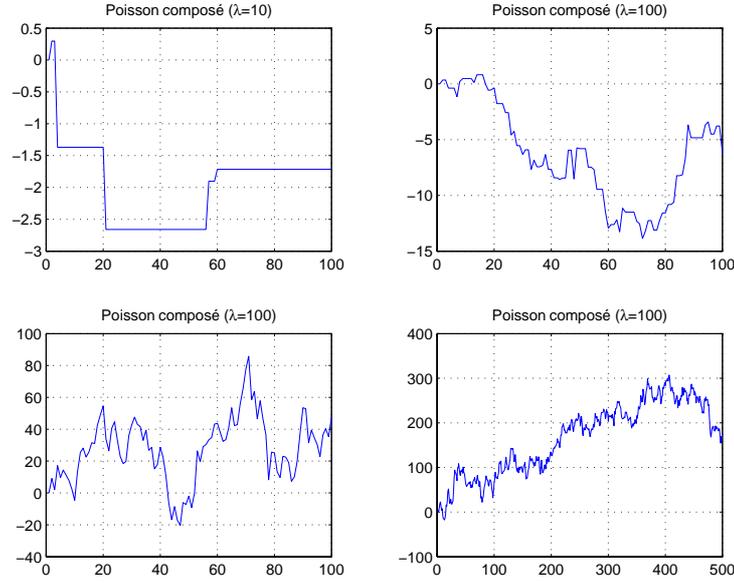


Figure 2.7: Trajectoires de processus de Poisson composé (sauts gaussiens).

2.3.1 Mesures Aléatoires

Un processus de Poisson $(N_t)_{t \geq 0}$ a été défini comme un processus de comptage. De façon plus précise, si T_1, T_2, \dots est une séquence d'instants de sauts de N , alors N_t , cf. 2.6, représente le nombre de sauts survenus entre les instants 0 et t , soit:

$$N_t = \# \{i \geq 1 : T_i \in [0, t]\}, \quad t \geq 0$$

et de façon similaire:

$$N_t - N_s = \# \{i \geq 1 : T_i \in]s, t]\}, \quad t > s$$

Cette procédure de comptage définit une *mesure*, notée J , sur $[0, \infty[$. Dès lors, pour tout ensemble mesurable $A \subset \mathbb{R}^+$, nous posons:

$$J_N(\omega, A) = \# \{i \geq 1 : T_i \in A\} \quad (2.7)$$

Comme la mesure $J_N(\omega, \cdot)$ dépend de ω , il s'agit alors d'une mesure aléatoire. Par ailleurs, l'intensité λ du processus de Poisson détermine la valeur moyenne de la mesure aléatoire, soit $\mathbb{E}[J_N(A)] = \lambda |A|$ où $|A|$ est la mesure de Lebesgue de A . On dit alors que J_N est la mesure de sauts associée au processus de Poisson N et ce dernier peut être exprimé en fonction de la mesure aléatoire, sous la forme:

$$N_t(\omega) = J_N(\omega, [0, t]) = \int_{[0, t]} J_X(\omega, ds)$$

De la même façon, nous pouvons associer une mesure aléatoire au processus de Poisson compensé M_t , $t \geq 0$, de la sorte:

$$M(\omega, A) = N(\omega, A) - \int_A \lambda dt = M(\omega, A) - \lambda |A|$$

Cette construction peut être généralisée dans de nombreuses directions.

Mesure de Poisson

La mesure J définie par 2.7 spécifie une mesure aléatoire de comptage sur \mathbb{R}^+ telle que $\mathbb{E}[J(A)] = \lambda |A|$ pour tout ensemble mesurable $A \subset \mathbb{R}^+$. Il est possible d'étendre ce cadre en remplaçant \mathbb{R}^+ par un ensemble $E \subset \mathbb{R}^d$ et la mesure de Lebesgue par une mesure de Radon μ sur E . On parle alors de mesure de Poisson.

Définition 2.3.2 (Mesure de Poisson) Soit $E \subset \mathbb{R}^d$ et μ une mesure de Radon sur (E, \mathcal{E}) . Une mesure de Poisson aléatoire sur E d'intensité μ est une mesure aléatoire:

$$J : \Omega \times \mathcal{E} \rightarrow \mathbb{N} \\ (\omega, A) \mapsto J(\omega, A)$$

telle que:

1. Pour tout ensemble mesurable $A \subset E$, $J(\cdot, A) = J(A)$ est une variable aléatoire de paramètre $\mu(A)$, soit:

$$\mathbb{P}(J(A) = k) = e^{-\mu(A)} \frac{(\mu(A))^k}{k!}, \quad k \in \mathbb{N}$$

2. Pour tous les ensembles disjoints A_1, \dots, A_n , les variables aléatoires $J(A_1), \dots, J(A_n)$ sont indépendantes.

De façon équivalente, il est possible de définir une mesure de Poisson comme un processus de comptage associé à une séquence aléatoire de marques (indépendantes et identiquement distribuées) à valeurs dans E . Ainsi, il existe $\{X_n(\omega), n \geq 1\}$ tel que:

$$J(\omega, A) = \sum_{n \geq 1} \mathbf{1}_A(X_n(\omega)), \quad \forall A \subset \mathcal{E}$$

ou encore:

$$J = \sum_{n \geq 1} \delta_{X_n}$$

Pareillement à la définition du processus de Poisson compensé, nous définissons la mesure de Poisson compensée \tilde{J} en soustrayant de J sa mesure d'intensité, soit:

$$\tilde{J}(A) = J(A) - \mu(A)$$

et nous avons:

$$\mathbb{E}[\tilde{J}(A_n)] = 0, \quad \text{Var}[\tilde{J}(A_n)] = \mu(A_n)$$

Processus à Marques

Comme nous l'avons vu précédemment, une mesure aléatoire de Poisson sur $[0, T] \times \mathbb{R}^d$ peut être vue comme un processus de comptage, soit:

$$J(\omega, \cdot) = \sum_{n \geq 1} \delta_{(T_n(\omega), Y_n(\omega))}$$

pour une séquence aléatoire $(T_n, Y_n)_{n \geq 1}$ à valeurs dans $[0, T] \times \mathbb{R}^d$. Il est alors possible de définir des structures de dépendance plus riches que celles déjà rencontrées. Nous avons alors la définition d'un processus à marques.

Définition 2.3.3 (Processus à marques) *Un processus à marques sur $(\Omega, \mathcal{F}, \mathbb{P})$ est une séquence $(T_n, Y_n)_{n \geq 1}$ telle que:*

1. $(T_n)_{n \geq 1}$ est une séquence non-anticipative de temps aléatoires avec $T_n \xrightarrow[n]{\infty} \infty$.
2. $(T_n)_{n \geq 1}$ est une séquence de variables aléatoires à valeurs dans E .
3. La valeur de Y_n est \mathcal{F}_{T_n} -mesurable.

Les processus à marques ne possèdent pas la propriété d'indépendance propre aux mesures de Poisson; ainsi pour A_1 et A_2 disjoints, $J([0, t] \times A_1)$ et $J([0, t] \times A_2)$ ne sont pas nécessairement indépendants, ni distribués suivant une loi de Poisson. De fait, ces processus permettent de prendre en compte des distributions et des structures de dépendance quelconques. Finalement, toute mesure de Poisson ne vérifie pas forcément $T_n \rightarrow \infty$ et de suite, toute mesure de Poisson ne peut pas être représentée comme un processus à marques; seules celles vérifiant la condition $\mu([0, t] \times \mathbb{R}^d) < \infty$ le pourront.

2.3.2 Marches Aléatoires

Nous entamons cette Section en donnant la définition d'un processus de Lévy.

Définition 2.3.4 (Processus de Lévy) *Un processus càdlàg $(X_t)_{t \geq 0}$ sur $(\Omega, \mathcal{F}, \mathbb{P})$ à valeurs dans \mathbb{R}^d tel que $X_0 = 0$ est un processus de Lévy s'il possède les propriétés suivantes :*

1. *Incréments indépendants: pour tout suite croissante de dates t_0, \dots, t_n , les variables aléatoires $X_{t_0}, X_{t_1} - X_{t_0}, \dots, X_{t_n} - X_{t_{n-1}}$ sont indépendantes.*
2. *Incréments stationnaires: la loi de $X_{t+h} - X_t$ ne dépend pas de t .*
3. *Continuité stochastique: $\forall \varepsilon > 0, \lim_{\varepsilon \rightarrow \infty} \mathbb{P}(|X_{t+h} - X_t| \geq \varepsilon) = 0$.*

La définition précédente peut être précisée si nous essayons de comprendre la construction d'un tel objet étape par étape. Si nous simulons un processus de Lévy (c'est-à-dire si nous simulons les incréments indépendants tels que définis dans la première condition de la proposition) sur une grille régulière $\Delta, 2\Delta, \dots$,

(Δ représente le pas de temps) nous obtenons alors une marche aléatoire. En effet, si nous définissons $S_n(\Delta) \equiv X_{n\Delta}$, nous pouvons écrire $S_n(\Delta) = \sum_{k=0}^{n-1} Y_k$ où les variables aléatoires $Y_k = X_{(k+1)\Delta} - X_{k\Delta}$ sont i.i.d. de distribution identique à celle de X_Δ . Si nous répétons cette procédure pour chaque intervalle de taille Δ de la grille, nous voyons bien que spécifier un processus de Lévy est équivalent à spécifier une famille de marches aléatoires $\{S_n(\delta)\}_{\delta=\Delta, 2\Delta, \dots}$. Si nous choisissons $\Delta = t/n$, alors la distribution de $X_t = S_n(\Delta)$ est la même que celle de la somme de n variables aléatoires de distribution celle de $X_{t/n}$. Une distribution ayant cette propriété de divisibilité est dite indéfiniment divisible.

Définition 2.3.5 (Distribution indéfiniment divisible) *Une distribution de probabilité F est dite indéfiniment divisible sur \mathbb{R}^d si pour $n \geq 2$, il existe n variables aléatoires Y_1, \dots, Y_n telle que $Y_1 + \dots + Y_n$ a pour distribution F .*

La propriétés d'indéfinie divisibilité concerne la stabilité de la distribution pour l'opération de convolution. Les distributions indéfiniment divisibles sont celles qui sont stables par cette opération. Dès lors, si X est un processus de Lévy, alors la distribution de X_t pour tout $t > 0$ est indéfiniment divisible. Cela ajoute alors une contrainte au choix de la distribution de X_t . Les exemples classiques de distributions indéfiniment divisibles sont: la loi normale, la loi gamma, les lois α -stables et la loi de Poisson. Et par exemple, si $X \sim \mathcal{N}(\mu, \sigma^2)$, alors nous pouvons écrire $X = \sum_{k=0}^{n-1} Y_k$ où les variables aléatoires Y_k sont i.i.d. de loi $\mathcal{N}(\mu/n, \sigma^2/n)$.

Processus de Poisson Composé

A partir de la Définition 2.2.7 d'un processus de Poisson composé X , les propriétés suivantes peuvent être déduites:

1. Les trajectoires de X sont des fonctions càdlàg constantes par morceaux.
2. Les tailles de sauts $(Y_i)_{i \geq 1}$ sont i.i.d. de distribution f .
3. Les instants de sauts $(T_i)_{i \geq 1}$ ont la même loi que le processus de Poisson N .

Nous avons alors la proposition importante suivante qui établit que les processus de Poisson composé sont les seuls processus de Lévy à trajectoires constantes par morceaux.

Proposition 2.3.6 *$(X_t)_{t \geq 0}$ est un processus de Poisson composé si et seulement si $(X_t)_{t \geq 0}$ est un processus de Lévy et que ses trajectoires sont des fonctions constantes par morceaux.*

Comme une fonction càdlàg peut être relativement bien approchée par une fonction constante par morceaux, il est attendu que les processus de Lévy peuvent être bien approchés par des processus de Poisson composés, ce qui permet, en première approche, d'en apprendre beaucoup sur ces processus de façon relativement aisée.

Mesure de Sauts Associée

Nous souhaitons à présent employer la notion de mesure aléatoire pour étudier le comportement des sauts d'un processus de Poisson composé. Nous avons alors la proposition suivante qui établit que la mesure J_X est une mesure de Poisson.

Proposition 2.3.7 *Soit $(X_t)_{t \geq 0}$ un processus de Poisson composé d'intensité λ et de distribution des sauts f . Sa mesure de saut J_X est une mesure de Poisson aléatoire sur $\mathbb{R}^d \times]0, \infty[$ d'intensité la mesure $\mu(dx \times dt) = \nu(dx) dt = \lambda f(x) dt$.*

Cette proposition suggère une interprétation de la mesure de Lévy d'un processus de Poisson composé comme le nombre moyen de sauts par unité de temps. En fait, cette interprétation n'est pas propre aux seuls processus de Poisson composé, mais s'applique à tous les processus de Lévy, ce que nous résumons par la définition suivante.

Définition 2.3.8 (Mesure de Lévy) *Soit $(X_t)_{t \geq 0}$ un processus de Lévy sur \mathbb{R}^d . La mesure ν sur \mathbb{R}^d définie par:*

$$\nu(A) = \mathbb{E}[\#\{t \in [0, 1]; \Delta X_t \neq 0, \Delta X_t \in A\}], \quad A \in \mathcal{B}(\mathbb{R}^d)$$

est appelée la mesure de Lévy de X : $\nu(A)$ est le nombre attendu, par unité de temps, de sauts dont la taille appartient à A .

Le résultat précédent implique que tout processus de Poisson composé peut être représenté sous la forme équivalente suivante:

$$X_t = \sum_{s \in [0, t]} \Delta X_s = \int_{[0, t] \times \mathbb{R}^d} x J_X(ds \times dx) \quad (2.8)$$

où J_X est une mesure de Poisson aléatoire d'intensité la mesure $\nu(dx) dt$. Comme un processus de Poisson composé a presque sûrement un nombre fini de sauts sur l'intervalle $[0, t]$, l'intégrale stochastique (cf. Section 3.1) définie par 2.8 est une somme finie et il n'y a pas de problèmes de convergence.

Nous donnons dans le Tableau 2.1 les définitions des mesures de Lévy de quelques spécifications couramment rencontrées dans la pratique.

Il s'agit pour toutes ces spécifications de modèles à taux d'activité fini, cf. Remarque 2.3.1, ce qui est équivalent à la condition suivante:

$$\int_{\mathbb{R}^d} \nu(dx) = \lambda < \infty$$

Décomposition d'un Processus de Lévy

Nous venons de voir que tout processus de Lévy X_t^0 peut se représenter sous la forme 2.8 pour une mesure de Poisson aléatoire d'intensité $\nu(dx) dt$. De la même façon, étant donné un mouvement brownien avec dérive $bt + \sigma W_t$,

Processus	Mesure de Lévy ν
<i>Pas de sauts</i> $bt + \sigma W_t$	-
<i>Taux d'activité fini</i> Merton (1976)	$\lambda \frac{1}{\sqrt{2\pi\sigma_j^2}} \exp\left(-\frac{(x-\alpha)^2}{2\sigma_j^2}\right)$
Kou (2002)	$\lambda \frac{1}{2\eta} \exp\left(-\frac{ x-k }{\eta}\right)$
Eraker (2003)	$\lambda \frac{1}{\eta} \exp\left(-\frac{x}{\eta}\right)$

Table 2.1: Caractéristiques de quelques processus de Lévy.

indépendant de X_t^0 , la somme $X_t = X_t^0 + bt + \sigma W_t$ définit un nouveau processus de Lévy, dont la décomposition est donnée par:

$$X_t = bt + \sigma W_t + \sum_{s \in [0, t]} \Delta X_s = bt + \sigma W_t + \int_{[0, t] \times \mathbb{R}^d} x J_X(dx \times ds) \quad (2.9)$$

où J_X est une mesure de Poisson aléatoire d'intensité la mesure $\nu(dx) dt$.

Remarque 2.3.9 (Processus de Lévy à taux d'activité infini) *Tant que la mesure ν est finie, la décomposition 2.9 est vérifiée. Mais ν n'est pas nécessairement finie. En effet, si X a un nombre infini de sauts sur $[0, T]$, la somme des sauts devient infinie et sa convergence n'est assurée qu'en imposant des conditions sur la mesure ν .*

En conséquence, la décomposition 2.9 implique que tout processus de Lévy est une combinaison d'un mouvement brownien avec dérive et d'une somme pouvant être infinie de processus de Poisson composés indépendants.

2.3.3 Propriété de Markov

Une propriété importante des processus de Lévy est la propriété de Markov. Elle établit que conditionnellement à X_t , l'évolution du processus après la date t est indépendante de son passé. Soit, pour une variable aléatoire Y et une filtration $\mathcal{F}_t^X = \sigma(X_s, s \in [0, t])$:

$$\mathbb{E}[Y | \mathcal{F}_t] = \mathbb{E}[Y | X_t]$$

Par ailleurs, le noyau ou probabilité de transition du processus X se définit comme:

$$P_{s,t}(x, B) = \mathbb{P}(X_t \in B | X_s = x), \quad \forall B \in \mathcal{B} \quad (2.10)$$

Par suite, la propriété de Markov implique la relation suivante entre les noyaux de transition:

$$P_{s,u}(x, B) = \int_{\mathbb{R}^d} P_{s,t}(x, dy) P_{t,u}(y, B)$$

qui est dénommée équation de Chapman-Kolmogorov. A partir de la définition 2.10, il est possible de montrer que les processus de Lévy sont homogènes en

temps et en espace, soit:

$$P_{s,t}(x, B) = P_{0,t-s}(0, B - x)$$

Nous introduisons aussi la notion de générateur infinitésimal pour un processus de Lévy. Ces opérateurs sont importants car le calcul d'espérance de fonctionnelles de processus de Lévy peut se réduire au calcul d'équations aux dérivées partielles (intégro-différentielles) mettant en jeu le générateur infinitésimal, cf. Chapitre 4.

Proposition 2.3.10 *Soit $(X_t)_{t \geq 0}$ un processus de Lévy de triplet (b, σ, ν) . Alors le générateur infinitésimal de X est défini pour toute fonction mesurable f comme:*

$$\begin{aligned} \mathcal{L}f(x) &= \sum_{i=1}^d b_i \frac{\partial}{\partial x_i} f(x) + \frac{1}{2} \sum_{i,j=1}^d \sigma_{ij} \frac{\partial^2}{\partial x_i \partial x_j} f(x) \\ &\quad + \int_{\mathbb{R}^d} \left(f(x+y) - f(y) - \sum_{i=1}^d y_i \frac{\partial}{\partial x_i} f(x) \mathbf{1}_{|y| \leq 1} \right) \nu(dy) \end{aligned}$$

A

Propriétés Empiriques des Séries Financières

L'une des principales motivations pour l'emploi de processus stochastiques en modélisation financière est de prendre en compte les faits empiriques observés sur les séries de rendements des actifs financiers. Toutefois, alors que les données à disposition des opérateurs sont observées de façon discrète, les processus stochastiques se proposent de fournir une vision continue de ces grandeurs. Nous nous interrogeons alors dans cette Annexe sur cet apparent paradoxe et discutons ainsi des propriétés statistiques des séries de rendements.

Cadre

Les observations à destination des opérateurs consistent en des séries de trajectoires du processus de prix des actifs financiers et c'est à partir de l'évolution de ces trajectoires que le profit ou la perte d'un investisseur sont déterminés.

Dans la suite, nous supposons que S_t décrit le prix, relevé à la date $t \geq 0$, d'un actif financier - une action, un taux d'intérêt, un taux de change ou un indice de marché - et $X_t = \log(S_t)$ est son logarithme. Etant donné une échelle de temps Δ (ou fréquence d'observation) le rendement est donné par:

$$r_t(\Delta) = X_{t+\Delta} - X_t$$

La Figure A.1 présente les rendements de l'Eurostoxx 50 de juin 2003 à mai 2004. Le choix du paramètre Δ est effectivement important car il indique avec quelle précision nous mesurons les données à disposition: une valeur faible (resp. élevée) de Δ indique que nous employons une échelle de temps fine (resp. grossière) (ou rapprochée, ou resp. large).

A.1. Du Temps Continu au Temps Discret

La modélisation statistique de X_t que nous privilégions adopte une vision continue de l'évolution de cette grandeur en travaillant avec l'équipement des processus stochastiques. Nous nous interrogeons à présent sur les motivations

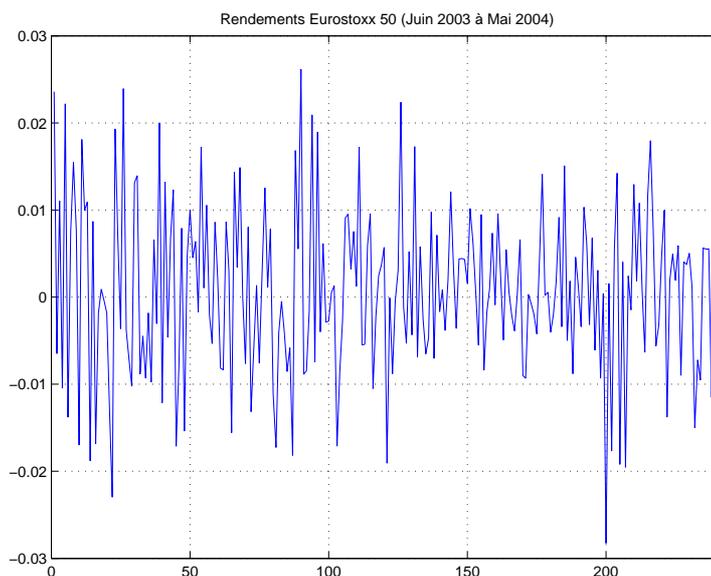


Figure A.1: Rendements de l'Eurostoxx 50 de juin 2003 à mai 2004.

pour employer des modèles en temps continu pour représenter des données discrètes. En effet, si nous considérons des données observées à une fréquence Δ , une approche usuelle de type séries temporelles voudrait que l'on adopte une représentation discrète du processus stochastique. Toutefois, pour une même série de rendements financiers, il est possible d'envisager des applications qui nécessitent des horizons temporels différents (jour, mois, année, etc.) et comme la Section suivante le montre, les propriétés statistiques des séries financières sont très sensibles à l'échelle de temps Δ . De fait, un modèle qui se montre satisfaisant pour décrire les rendements pour une échelle Δ_1 peut se montrer incapable de fournir les mêmes résultats pour une échelle Δ_2 . Les modèles de séries temporelles de type ARIMA ou GARCH témoignent souvent de ce manque de stabilité dans leur représentation. Et il est beaucoup trop lourd de définir un modèle différent pour chaque fréquence d'observation. A contrario, un modèle stochastique en temps continu incorpore dans sa définition les propriétés statistiques des séries de rendements pour toutes les fréquences d'observation possibles. Ainsi, toutes les propriétés discrètes possèdent une limite dès lors que l'échelle de temps devient fine.

A.2. Faits Stylisés des Séries de Rendements

Du fait de l'hétérogénéité des données financières existantes, il n'est pas possible d'établir de façon précise la liste des propriétés quantitatives qui permettraient a priori de distinguer les qualités respectives de deux représentations paramétriques données. De fait, l'observation des faits stylisés passe par la reconnaissance d'un dénominateur commun entre les propriétés de séries fi-

nancières pouvant être observées sur des marchés différents ou établis selon des caractéristiques différentes. Les propriétés statistiques des séries de rendements les plus remarquables sont les suivantes :

1. Queues épaisses: la distribution des rendements présente un phénomène de queue épaisse avec un indice de kurtosis¹ positif.
2. Absence d'autocorrélation: l'autocorrélation entre les séries de rendements est la plupart du temps non significative ,sauf pour des échelles de temps très fines, pour lesquelles des effets microéconomiques peuvent jouer.
3. Asymétrie entre gain et perte: on observe plus fréquemment des mouvements importants à la baisse que des mouvements à la hausse.
4. Normalité par agrégation: l'augmentation de l'échelle de temps Δ se traduit par une distribution des rendements de plus en plus proche de la loi normale. En particulier l'allure de la distribution n'est pas la même pour différentes échelles de temps: le phénomène de queues épaisses est moins prononcé quand Δ augmente.
5. Volatilité en grappes: des mouvements importants tendent à être suivis de mouvements importants, mais de signe opposés et de la même façon pour des mouvements faibles.
6. Effet de levier: certaines mesures de la volatilité sont négativement corrélées avec les rendements calculés pour un actif financier.

¹Une distribution se caractérise par son kurtosis, c'est-à-dire la taille de ses queues de distribution. Les distributions à queues épaisses sont dites leptokurtiques et celles à queues fines de platokurtiques.

3

Calcul Stochastique

L'objectif principal de ce Chapitre est de présenter certains résultats utiles sur l'intégration et le calcul stochastiques, sans toutefois entrer dans de trop nombreux détails. Et alors qu'il est habituel d'employer et de développer l'équipement des semimartingales pour obtenir ces résultats dans le cadre des processus mixtes de diffusions et à sauts, nous avons recours à une approche plus pragmatique de ces notions, ce qui nous permet de ne pas trop entrer dans des développements de trop techniques et donc difficiles.

3.1. Intégrales Stochastiques

Nous débutons cette Section en définissant sur $(\Omega, \mathcal{F}, \mathbb{P})$, de façon directe, l'intégrale stochastique lorsque l'intégrateur, un processus S qui est adapté à la filtration \mathcal{F} , est successivement le mouvement brownien et un processus de sauts pur. Si nous considérons un processus X qui est \mathcal{F} -adapté, notre objectif est de définir l'intégrale stochastique:

$$I_t(X) = \int_0^t X_u dS_u \quad (3.1)$$

pour $t \in [0, T]$ et d'étudier ses propriétés en tant que processus indexé par le temps t . Comme nous l'avons vu précédemment, la difficulté par rapport au *calcul ordinaire* réside dans le fait que les processus de Wiener, de Poisson et de Lévy ne sont pas différentiables par rapport au temps et qu'il est alors nécessaire de donner du sens à 3.1 dans ce contexte.

3.1.1 Mouvement Brownien

Soit $T > 0$, nous cherchons à définir la quantité:

$$\int_0^T X_t dW_t \quad (3.2)$$

où l'intégrateur $(W_t)_{t \geq 0}$ est un mouvement brownien défini sur une filtration $(\mathcal{F}_t)_{t \geq 0}$ et l'intégrand $(X_t)_{t \geq 0}$ est un processus \mathcal{F}_t -adapté quelconque. Le problème que nous rencontrons lorsque nous cherchons à donner du sens à l'intégrale d'Itô 3.2 provient du fait que les trajectoires du mouvement brownien

ne sont pas différentiables par rapport au temps. Dans le cadre du calcul *ordinaire*, si $g(t)$ est une fonction différentiable, alors:

$$\int_0^T X_t dg(t) = \int_0^T X_t g'(t) dt$$

peut être définie comme une intégrale classique au sens de Lebesgue. Ce n'est pas le cas du mouvement brownien pour lequel W' ne peut pas être explicité.

Recours aux Processus Simples

Pour contourner ce problème de non-différentiabilité, il est possible de procéder en deux étapes. Dans un premier temps, l'intégrale d'Itô est défini pour des intégrands simples, qui sont en fait des fonctions constantes par morceaux. Puis, dans un second temps, elle est étendue à des intégrands non simples par passage à la limite. Soit $\Pi = \{t_0, t_1, \dots, t_n\}$ une partition de $[0, T]$ et nous supposons que le processus X_t est constant en t dans chaque sous-intervalle $[t_j, t_{j+1})$. Un tel processus est qualifié de processus simple. La Figure 3.1 présente une trajectoire d'un tel processus.

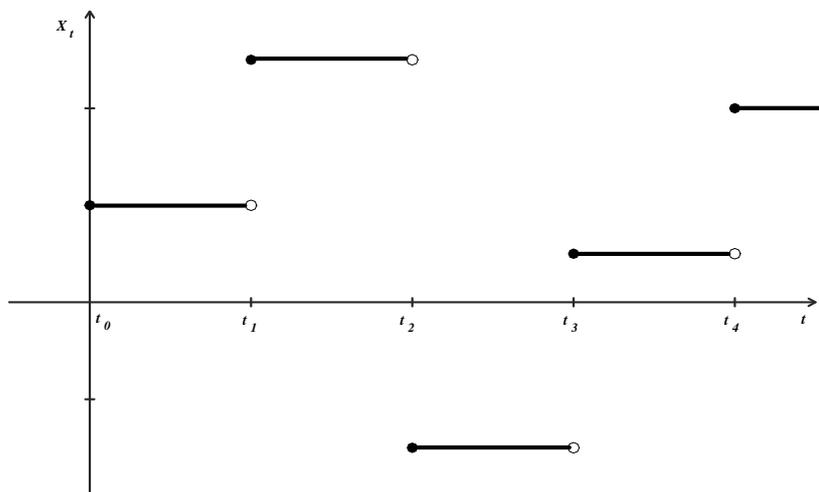


Figure 3.1: Trajectoire d'un processus simple.

Il est possible de définir la relation existante entre les processus W_t et X_t intervenant dans la définition de l'intégrale d'Itô 3.2 dans le sens suivant. Supposons que W_t représente le prix d'un actif financier à la date t . Soit par ailleurs, t_0, t_1, \dots, t_{n-1} , les dates où une transaction dans l'actif sous-jacent intervient. Par suite, $X_{t_0}, X_{t_1}, \dots, X_{t_{n-1}}$ représentent les positions détenues dans cet actif. Le gain lié à la détention de ces parts est alors égale à:

$$\begin{aligned} I_t &= X_{t_0} [W_t - W_{t_0}], \quad t \in [0, t_1] \\ I_t &= X_{t_0} W_{t_1} + X_{t_1} [W_t - W_{t_1}], \quad t \in [t_1, t_2] \\ I_t &= X_{t_0} W_{t_1} + X_{t_1} [W_{t_2} - W_{t_1}] + X_{t_2} [W_t - W_{t_1}], \quad t \in [t_2, t_3] \end{aligned}$$

Soit, de façon générale, pour $t \in [t_k, t_{k+1}]$:

$$I_t = \sum_{j=0}^{k-1} X_{t_j} [W_{t_{j+1}} - W_{t_j}] + X_{t_k} [W_t - W_{t_k}] \quad (3.3)$$

Et le processus 3.3 représente l'intégrale d'Itô du processus simple X_t , ce que nous résumons, par passage à la limite lorsque $\|\Pi\| \rightarrow 0$, par l'identité suivante:

$$I_t = \int_0^t X_u dW_u$$

Si $t = t_n = T$, nous disposons alors, par 3.3, d'une définition de l'intégrale d'Itô 3.2 en termes de processus simples. Cette définition est par ailleurs valide pour tout $t \in [0, T]$ et pas uniquement pour la borne supérieure d'intégration T .

Propriétés de l'Intégrale

Nous avons défini l'intégrale 3.3 comme le gain lié à la détention de la martingale W_t . Et comme une martingale n'a ni tendance à augmenter ni à diminuer, il est attendu que I_t possède aussi cette propriété.

Théorème 3.1.1 *L'intégrale d'Itô 3.3 est une martingale.*

Preuve. Il nous faut montrer $\mathbb{E}[I_t | \mathcal{F}_s] = I_s$ pour tout $0 \leq s \leq t \leq T$. Supposons que les indices de temps s et t sont dans des sous-intervalles disjoints de Π , i.e.: $s \in [t_l, t_{l+1})$ et $t \in [t_k, t_{k+1})$ pour $t_l < t_k$. Soit, de 3.3, il vient:

$$\begin{aligned} I_t &= \sum_{j=0}^{l-1} X_{t_j} [W_{t_{j+1}} - W_{t_j}] + X_{t_l} [W_{t_{l+1}} - W_{t_l}] \\ &\quad + \sum_{j=l+1}^{k-1} X_{t_j} [W_{t_{j+1}} - W_{t_j}] + X_{t_k} [W_t - W_{t_k}] \end{aligned} \quad (3.4)$$

Et nous prenons alors l'espérance conditionnelle de chaque côté de 3.4. Nous avons:

$$\mathbb{E} \left[\sum_{j=0}^{l-1} X_{t_j} [W_{t_{j+1}} - W_{t_j}] | \mathcal{F}_s \right] = \sum_{j=0}^{l-1} X_{t_j} [W_{t_{j+1}} - W_{t_j}]$$

du fait que chaque terme $X_{t_j} [W_{t_{j+1}} - W_{t_j}]$ est \mathcal{F}_s -mesurable. Et:

$$\mathbb{E} [X_{t_l} [W_{t_{l+1}} - W_{t_l}] | \mathcal{F}_s] = X_{t_l} \mathbb{E} [W_{t_{l+1}} | \mathcal{F}_s - W_{t_l}] = X_{t_l} [W_s - W_{t_l}]$$

Par suite:

$$\begin{aligned} \mathbb{E} [X_{t_j} (W_{t_{j+1}} - W_{t_j}) | \mathcal{F}_s] &= \mathbb{E} [\mathbb{E} [X_{t_j} (W_{t_{j+1}} - W_{t_j}) | \mathcal{F}_{t_j}] | \mathcal{F}_s] \\ &= \mathbb{E} [X_{t_j} \mathbb{E} [W_{t_{j+1}} | \mathcal{F}_{t_j}] - W_{t_j} | \mathcal{F}_s] \\ &= \mathbb{E} [X_{t_j} (W_{t_j} - W_{t_j}) | \mathcal{F}_s] = 0 \end{aligned}$$

du fait que W est une martingale. Soit:

$$\mathbb{E} \left[\sum_{j=l+1}^{k-1} X_{t_j} [W_{t_{j+1}} - W_{t_j}] | \mathcal{F}_s \right] = 0$$

Finalement:

$$\begin{aligned} \mathbb{E} [X_{t_k} (W_t - W_{t_k}) | \mathcal{F}_s] &= \mathbb{E} [\mathbb{E} [X_{t_k} (W_t - W_{t_k}) | \mathcal{F}_{t_k}] | \mathcal{F}_s] \\ &= \mathbb{E} [X_{t_k} \mathbb{E} [W_t | \mathcal{F}_{t_k}] - W_{t_k} | \mathcal{F}_s] \\ &= \mathbb{E} [X_{t_k} (W_{t_k} - W_{t_k}) | \mathcal{F}_s] = 0 \end{aligned}$$

ce qui conclut la preuve. ■

Comme I_t est une martingale et $I_0 = 0$, nous avons $\mathbb{E} [I_t] = 0$ pour tout $t \geq 0$. Par suite, $\text{Var} [I_t] = \mathbb{E} [I_t^2]$, quantité que nous pouvons encore préciser.

Théorème 3.1.2 (Isométrie) *L'intégrale d'Itô 3.3 satisfait:*

$$\mathbb{E} [I_t] = 0, \quad \mathbb{E} [I_t^2] = \mathbb{E} \left[\int_0^t X_u^2 du \right], \quad t \geq 0$$

Preuve. Nous posons $X_j = X_{t_j}$, $D_j = W_{t_{j+1}} - W_{t_j}$ pour $j = 0, \dots, k-1$ et $D_k = W_t - W_{t_k}$, telle que 3.3 devient $I_t = \sum_{j=0}^k X_j D_j$ et alors:

$$I_t^2 = \sum_{j=0}^k X_j^2 D_j^2 + 2 \sum_{0 \leq i < j \leq k} X_i X_j D_i D_j$$

Comme pour $i < j$, la variable aléatoire $X_i X_j D_i$ est \mathcal{F}_j -mesurable, D_j est indépendant de \mathcal{F}_j et $\mathbb{E} [D_j] = 0$, il vient:

$$\mathbb{E} [X_i X_j D_i D_j] = \mathbb{E} [X_i X_j D_i] \mathbb{E} [D_j] = 0$$

Par suite, comme $\mathbb{E} [D_j^2] = t_{j+1} - t_j$ pour $j = 0, \dots, k-1$ et $\mathbb{E} [D_k^2] = t - t_k$, nous avons:

$$\begin{aligned} \mathbb{E} [I_t^2] &= \sum_{j=0}^k \mathbb{E} [X_j^2 D_j^2] = \sum_{j=0}^k \mathbb{E} [X_j^2] \mathbb{E} [D_j^2] \\ &= \sum_{j=0}^{k-1} \mathbb{E} [X_j^2] (t_{j+1} - t_j) + \mathbb{E} [X_k^2] (t - t_k) \end{aligned} \quad (3.5)$$

Finalement, comme X_i^2 est constant sur $[t_i, t_{i+1})$, nous avons $X_i^2 (t_{i+1} - t_i) = \int_{t_i}^{t_{i+1}} X_u^2 du$, et 3.5 se réécrit:

$$\mathbb{E} [I_t^2] = \sum_{j=0}^{k-1} \mathbb{E} \left[\int_{t_j}^{t_{j+1}} X_u^2 du \right] + \mathbb{E} \left[\int_{t_k}^t X_u^2 du \right] = \mathbb{E} \left[\int_0^t X_u^2 du \right]$$

ce qui est le résultat attendu. ■

Finalement, nous pouvons aussi nous intéresser à la variation quadratique, cf. Définition 2.1.2, de l'intégrale d'Itô 3.3.

Théorème 3.1.3 *La variation quadratique jusqu'en t de l'intégrale d'Itô 3.3 est égale à:*

$$[I, I]_t = \int_0^t X_u^2 du, \quad t \geq 0$$

Preuve. Sur chaque sous-intervalle $[t_j, t_{j+1})$ pour $j = 0, \dots, k-1$, la variation quadratique accumulée par l'intégrale d'Itô est égale à:

$$\sum_{i=0}^{m-1} [I_{s_{i+1}} - I_{s_i}]^2 = \sum_{i=0}^{m-1} [X_{t_j} (W_{s_{i+1}} - W_{s_i})]^2 = X_{t_j}^2 \sum_{i=0}^{m-1} (W_{s_{i+1}} - W_{s_i})^2 \quad (3.6)$$

Quand $m \rightarrow \infty$, la quantité $\sum_{i=0}^{m-1} (W_{s_{i+1}} - W_{s_i})^2$ converge vers la variation quadratique accumulée par le mouvement brownien W entre les dates t_j et t_{j+1} , qui est égale à $t_{j+1} - t_j$. Par suite, à la limite, 3.6 est égale à:

$$X_{t_j}^2 (t_{j+1} - t_j) = \int_{t_j}^{t_{j+1}} X_u^2 du \quad (3.7)$$

du fait que X_u^2 est constant sur $[t_j, t_{j+1})$. Si nous répétons cette opération pour tout sous-intervalle $[t_j, t_{j+1})$ pour $j = 0, \dots, k-1$ et pour le sous-intervalle $[t_k, t)$ et que nous sommes les termes de la forme 3.7, nous obtenons le résultat annoncé. ■

Cas d'Intégrands Généraux

Nous souhaitons à présent pouvoir définir l'intégrale d'Itô pour des processus X_t quelconques, c'est-à-dire des processus qui peuvent varier avec le temps et aussi *sauter*. Ils doivent néanmoins vérifier une condition dite de carré intégrable:

$$\mathbb{E} \left[\int_0^T X_t^2 dt \right] < \infty \quad (3.8)$$

Afin de pouvoir donner du sens à $\int_0^T X_t dW_t$, nous approchons le processus X_t par des processus simples. La Figure 3.2 présente une telle approximation. Celle-ci devient d'autant plus précise que la partition considérée est fine.

A la limite, lorsque $n \rightarrow \infty$, nous devrions avoir la *convergence* suivante:

$$\lim_{n \rightarrow \infty} \mathbb{E} \left[\int_0^T |X_t^n - X_t|^2 dt \right] = 0 \quad (3.9)$$

où X_t^n est la version approchée par un processus simple du processus stochastique X_t . Pour chaque X_t^n , l'intégrale d'Itô correspondante a déjà été définie comme étant $\int_0^t X_u^n dW_u$. Il est ainsi possible de définir l'intégrale d'Itô pour un intégrand quelconque de la sorte:

$$\int_0^t X_u dW_u = \lim_{n \rightarrow \infty} \int_0^t X_u^n dW_u, \quad t \in [0, T] \quad (3.10)$$

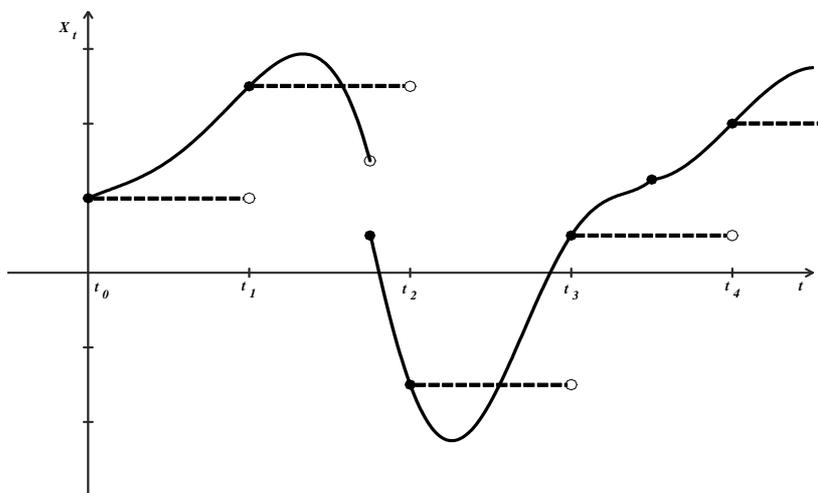


Figure 3.2: Approximation d'une intégrale par un processus simple.

et cette intégrale hérite des propriétés de l'intégrale d'Itô définie pour des processus simples.

Théorème 3.1.4 Soit X_t un processus adapté satisfaisant 3.8. Alors l'intégrale $I_t = \int_0^t X_u dW_u$ définie par 3.10 vérifie les propriétés suivantes:

1. (Mesurabilité) Pour tout $t \geq 0$, I_t est \mathcal{F}_t -mesurable.
2. (Martingale) Pour tout $t \geq 0$, I_t est une martingale.
3. (Isométrie) $\mathbb{E}[I_t] = 0$ et $\mathbb{E}[I_t^2] = \mathbb{E}\left[\int_0^t X_u^2 du\right]$
4. (Variation quadratique) $[I, I](t) = \int_0^t X_u^2 du$.
5. (Linéarité) Si $I_t^i = \int_0^t X_u^i dW_u$ pour $i = (1, 2)$ et $\alpha \in \mathbb{R}$, alors:

$$\alpha (I_t^1 \pm I_t^2) = \int_0^t \alpha (X_u^1 \pm X_u^2) dW_u$$

Nous donnons à présent un exemple de calcul d'une intégrale d'Itô.

Exemple 3.1.5 Nous nous intéressons au calcul de $\int_0^T W_t dW_t$. A cette fin, pour une constante n grande, nous approchons l'intégrand $X_t = W_t$ par le processus simple:

$$X_t^n = \begin{cases} W_0 = 0 & \text{si } 0 \leq t < \frac{T}{n} \\ W\left(\frac{T}{n}\right) & \text{si } \frac{T}{n} \leq t < \frac{2T}{n} \\ \vdots & \vdots \\ W\left(\frac{(n-1)T}{n}\right) & \text{si } \frac{(n-1)T}{n} \leq t < T \end{cases},$$

Par suite:

$$\begin{aligned} \int_0^T W_t dW_t &= \lim_{n \rightarrow \infty} \int_0^T X_t^n dW_t \\ &= \lim_{n \rightarrow \infty} \sum_{j=0}^{n-1} W\left(\frac{jT}{n}\right) \left[W\left(\frac{(j+1)T}{n}\right) - W\left(\frac{jT}{n}\right) \right] \end{aligned} \quad (3.11)$$

Nous posons $W_j = W\left(\frac{jT}{n}\right)$ et en prévision du calcul de 3.11, nous calculons:

$$\begin{aligned} \frac{1}{2} \sum_{j=0}^{n-1} (W_{j+1} - W_j)^2 &= \frac{1}{2} \sum_{j=0}^{n-1} W_{j+1}^2 - \sum_{j=0}^{n-1} W_j W_{j+1} + \frac{1}{2} \sum_{j=0}^{n-1} W_j^2 \\ &= \frac{1}{2} \sum_{k=0}^n W_k^2 - \sum_{j=0}^{n-1} W_j W_{j+1} + \frac{1}{2} \sum_{j=0}^{n-1} W_j^2 \\ &= \frac{1}{2} W_n^2 + \frac{1}{2} \sum_{k=0}^{n-1} W_k^2 - \sum_{j=0}^{n-1} W_j W_{j+1} + \frac{1}{2} \sum_{j=0}^{n-1} W_j^2 \\ &= \frac{1}{2} W_n^2 + \sum_{j=0}^{n-1} W_j^2 - \sum_{j=0}^{n-1} W_j W_{j+1} \\ &= \frac{1}{2} W_n^2 + \sum_{j=0}^{n-1} W_j (W_j - W_{j+1}) \end{aligned} \quad (3.12)$$

De 3.12, il vient alors:

$$\sum_{j=0}^{n-1} W_j (W_{j+1} - W_j) = \frac{1}{2} W_n^2 - \frac{1}{2} \sum_{j=0}^{n-1} (W_{j+1} - W_j)^2$$

Soit par substitution dans 3.11 et par passage à la limite:

$$\int_0^T W_t dW_t = \frac{1}{2} W_T^2 - \frac{1}{2} [W, W]_T = \frac{1}{2} W_T^2 - \frac{1}{2} T \quad (3.13)$$

A comparer avec le résultat obtenu par le calcul ordinaire, pour g différentiable avec $g(0) = 0$:

$$\int_0^T g(t) dg(t) = \int_0^T g(t) g'(t) dt = \frac{1}{2} g^2(t) \Big|_0^T = \frac{1}{2} g^2(T)$$

3.1.2 Cadre des Semimartingales

L'objectif de la théorie de l'intégration stochastique et partant de sa représentation la plus connue, l'intégrale stochastique, est de donner une approximation raisonnable de 3.1 pour une classe suffisamment large de processus. Il nous faut alors définir les processus stochastiques qui pourront servir d'intégrands sur une famille particulière de processus adaptés. La construction précédente de

l'intégrale d'Itô fournit une approche générale pour donner du sens à l'intégrale stochastique 3.1 lorsque l'intégrateur est le mouvement brownien et les intégrands des processus simples. Un développement similaire pourrait être menée pour définir l'intégrale stochastique lorsque l'intégrateur est un processus de Poisson ou plus généralement un processus de Lévy. Au contraire, nous présentons une approche générale. L'équipement des semimartingales va être l'outil adéquat pour répondre à ces questions.

Soit pour $t \geq 0$, l'intégrale stochastique:

$$\mathcal{I}_t = \int_0^t \phi_u dS_u \quad (3.14)$$

où ϕ et S sont des processus \mathcal{F} -adaptés. Nous commençons par définir les processus simples.

Définition 3.1.6 (Processus simple) *Un processus stochastique $\phi_t, t \in [0, T]$, est appelé un processus simple s'il peut être représenté sous la forme:*

$$\phi_t = \phi_0 \mathbf{1}_{t=0} + \sum_{i=0}^{n-1} \phi_i \mathbf{1}_{[t_i, t_{i+1})}(t) \quad (3.15)$$

où $\Pi = \{0 = t_0 < t_1 < \dots < t_n = T\}$ est une partition de $[0, T]$ et chaque ϕ_i est une variable aléatoire bornée dont la valeur est révélée en t_i , i.e.: \mathcal{F}_{t_i} -mesurable.

De la même façon que nous avons défini l'intégrale d'Itô $\int_0^t X_u dW_u$ comme le gain associé à la détention de parts X_u dans l'actif martingale W_u , il est possible de donner du sens à 3.14 par la même analogie: ϕ est la stratégie d'investissement (*trading strategy*) dans l'actif sous-jacent martingale S (*underlying martingale*). Nous avons alors la propriété intéressante suivante.

Proposition 3.1.7 *Soit $S_t, t \in [0, T]$, une martingale. Alors pour tout processus prédictif simple ϕ , l'intégrale stochastique $\mathcal{I}_t = \int_0^t \phi_u dS_u$ est une martingale.*

Preuve. La preuve peut se faire en suivant celle développée dans le cadre de l'intégrale d'Itô, Théorème 3.1.1, en utilisant le fait que les ϕ sont \mathcal{F} -mesurables. ■

Une autre utilisation des intégrales stochastiques peut être de construire de nouveaux processus stochastiques à partir d'existants. Ceci peut se révéler utile si nous cherchons à décrire des modèles stochastiques complexes, par exemple des modèles mixtes.

Proposition 3.1.8 *Soit $X_t, t \in [0, T]$, un processus càdlàg non-anticipatif et σ_t et ϕ_t deux processus simples. Alors $S_t = \int_0^t \sigma_u dX_u$ est un processus càdlàg non-anticipatif et:*

$$\int_0^t \phi_u dS_u = \int_0^t \phi_u \sigma_u dX_u$$

Remarque 3.1.9 En termes de notations, la relation $S_t = \int_0^t \sigma_u dX_u$ est souvent résumée sous la forme $dS_t = \sigma_t dX_t$. Par suite, si $dS_t = \sigma_t dX_t$ et $dY_t = \phi_t dS_t$, alors nous avons $dY_t = \phi_t \sigma_t dX_t$. On parle alors de règle de composition des processus stochastiques.

Semimartingales

Jusqu'à présent, nous n'avons requis aucune propriété particulière sur l'intégrateur S . En effet, 3.14 est définie pour tout processus S càdlàg. Si nous nous rappelons de la construction de l'intégrale d'Itô, le passage de l'intégrale définie pour des intégrands simples à celle définie pour des intégrands généraux nécessite la validation de la condition de convergence 3.9. Dans le cadre présent, cette propriété de *stabilité*, à savoir $\phi^n \rightarrow \phi$ alors $\int \phi^n dS \rightarrow \int \phi dS$, nécessite de définir la notion de semimartingale.

Définition 3.1.10 (Semimartingale) *Le processus stochastique càdlàg non-anticipatif S est une semimartingale si l'intégrale stochastique d'un processus simple par rapport à S :*

$$\phi = \phi_0 \mathbf{1}_{t=0} + \sum_{i=0}^{n-1} \phi_i \mathbf{1}_{[t_i, t_{i+1})} \longmapsto \int_0^T \phi dS = \phi_0 S_0 + \sum_{i=0}^{n-1} \phi_i (S_{t_{i+1}} - S_{t_i}) \quad (3.16)$$

vérifie la condition de continuité:

$$\sup_{t \in [0, T]} |\phi_t^n - \phi_t| \xrightarrow[n \rightarrow \infty]{} 0 \Rightarrow \int_0^T \phi^n dS \xrightarrow[n \rightarrow \infty]{\mathbb{P}} \int_0^T \phi dS$$

La Définition 3.1.10 semble relativement difficile à vérifier en pratique. En effet, comment s'assurer qu'un processus donné vérifie bien la condition de continuité précédente? Toutefois, nous avons le *corollaire* suivant qui nous assure que la plupart des processus stochastiques rencontrés dans la pratique sont des semimartingales.

Corollaire 3.1.11 (Semimartingale) *Nous avons:*

1. *Tout processus à variation finie est une semimartingale.*
2. *Toute martingale de carré intégrable est une semimartingale.*

Nous faisons à présent trois remarques importantes, qui nous seront utiles pour le reste de la présentation. Les processus classiques suivants sont des semimartingales :

- Le processus de Wiener en temps que martingale de carré intégrable.
- Le processus de Poisson en temps que processus à variation finie.
- Tous les processus de Lévy car comme le montre la décomposition de Lévy-Itô, tout processus de Lévy peut se décomposer en une somme d'une martingale de carré intégrable (partie diffusion) et un processus à variation finie (partie sauts).

Intégrale Stochastique pour des Processus Càdlàg

La Définition 3.1.10 nous indique que, dans un certain sens, l'intégrale stochastique 3.14 peut être approchée par un processus simple. Toutefois, cette construction est relativement limitée. En effet, si nous définissons l'intégrand S comme $S_t = \int_0^t \sigma_u dX_u$ où X est un processus adapté non-anticipatif pouvant sauter (construction par associativité), il est nécessaire d'étendre la définition précédente à une classe plus large de processus intégrateurs.

Proposition 3.1.12 (Somme de Riemann) *Soit S une semimartingale, ϕ un processus càdlàg et $\Pi = \{0 = t_0 < \dots < t_n = T\}$ une partition de $[0, T]$. Alors:*

$$\phi_{t_0} S_{t_0} + \sum_{j=0}^{n-1} \phi_{t_j} (S_{t_{j+1}} - S_{t_j}) \xrightarrow[n \rightarrow \infty]{\mathbb{P}} \int_0^t \phi_{u-} dS_u \quad (3.17)$$

La définition 3.17 de l'intégrale stochastique est proche de la définition d'une intégrale comme une somme de Riemann. Toutefois, on notera que la variation de S est multipliée par la valeur de ϕ relevée à la limite gauche de l'intervalle (ϕ_{u-}), ce qui assure que le processus ainsi défini est bien non-anticipatif. Dès lors, la construction précédente fournit une construction de l'intégrale stochastique par des sommes de Riemann non-anticipatives.

Nous avons alors les propriétés suivantes liées à l'intégrale stochastique. À l'instar de l'intégrale d'Itô, elles peuvent être établis dans un premier temps pour des processus simples, puis par passage à la limite pour des intégrands càdlàg généraux.

Théorème 3.1.13 *Soit ϕ un processus càdlàg non-anticipatif et S une semimartingale. Alors l'intégrale stochastique $\mathcal{I}_t = \int_0^t \phi_u dS_u$ vérifie les propriétés suivantes:*

1. (Mesurabilité) Pour tout $t \geq 0$, \mathcal{I}_t est \mathcal{F}_t -mesurable.
2. (Semimartingale) Pour tout $t \geq 0$, \mathcal{I}_t est une semimartingale.
3. (Isométrie) $\mathbb{E}[\mathcal{I}_t] = 0$ et $\mathbb{E}[\mathcal{I}_t^2] = \mathbb{E}\left[\int_0^t \phi_u^2 du\right]$
4. (Variation quadratique) $[\mathcal{I}, \mathcal{I}]_t = \int_0^t \phi_u^2 du$.
5. (Linéarité) Si $\mathcal{I}_t^i = \int_0^t \phi_u^i dS_u$ pour $i = (1, 2)$ et $\alpha \in \mathbb{R}$, alors:

$$\alpha (\mathcal{I}_t^1 \pm \mathcal{I}_t^2) = \int_0^t \alpha (\phi_u^1 \pm \phi_u^2) dS_u$$

Cas du Mouvement Brownien

Bien que nous ayons déjà défini l'intégrale stochastique dans le cas où l'intégrand est le mouvement brownien, cf. Section 3.1.1, nous la définissons à nouveau, mais cette fois grâce à l'équipement des semimartingales. Comme le processus de Wiener est une semimartingale, l'intégrale stochastique par rapport à

ce dernier se définit comme en 3.16 et 3.17. Ainsi, étant donné un processus prédictif ϕ :

$$\phi_t = \phi_0 \mathbf{1}_{t=0} + \sum_{i=0}^{n-1} \phi_i \mathbf{1}_{[t_i, t_{i+1})}(t), \quad t \in [0, T]$$

l'intégrale d'Itô se définit comme:

$$\mathcal{I}_t(\phi) = \int_0^T \phi_t dW_t = \sum_{i=0}^{n-1} \phi_i (W_{t_{i+1}} - W_{t_i})$$

3.1.3 Processus de Poisson et Lévy

A partir du cadre des semimartingales, nous pouvons à présent définir l'intégrale stochastique lorsque l'intégrateur est un processus de sauts pur, qu'il soit de Poisson ou de Lévy.

Recours aux Fonctions Simples

Nous avons défini les mesures aléatoires de Poisson en Section 2.3.1. Nous en rappelons ici les principales propriétés. Soit J une mesure de Poisson sur $[0, T] \times \mathbb{R}^d$ d'intensité $\mu(dt, dx)$. La mesure de Poisson compensé correspondante \tilde{J} est défini comme la version centrée de J : $\tilde{J}(A) = J(A) - \mu(A) = J(A) - \mathbb{E}[J(A)]$.

Par analogie avec les processus prédictifs simples, nous définissons les fonctions simples $\phi : \Omega \times [0, T] \times \mathbb{R}^d \rightarrow \mathbb{R}$, comme:

$$\phi(t, y) = \sum_{i=0}^{n-1} \sum_{j=0}^{m-1} \phi_{ij} \mathbf{1}_{[t_i, t_{i+1})}(t) \mathbf{1}_{A_j}(y)$$

où $\Pi = \{0 = t_0 < \dots < t_n = T\}$ est une partition de $[0, T]$, les $(\phi_{ij})_{i,j}$ sont des variables aléatoires bornées \mathcal{F}_{t_i} -mesurables et les $(A_j)_j$ des sous-espaces disjoints de \mathbb{R}^d . L'intégrale stochastique d'une mesure de Poisson pour un intégrand ϕ quelconque:

$$\mathcal{I}_t(\phi) = \int_{[0, T] \times \mathbb{R}^d} \phi(t, y) J(dt, dy)$$

se définit alors comme la variable aléatoire:

$$\int_0^T \int_{\mathbb{R}^d} \phi(t, y) M(dt, dy) = \sum_{i,j=0}^{n-1, m-1} \phi_{ij} J([t_i, t_{i+1}) \times A_j)$$

soit encore:

$$\mathcal{I}_t(\phi) = \int_{[0, T] \times \mathbb{R}^d} \phi(t, y) J(dt, dy) = \sum_{i,j=0}^{n-1, m-1} \phi_{ij} (J_{t_{i+1}}(A_j) - J_{t_i}(A_j)) \quad (3.18)$$

Un exemple important pour notre propos est le cas où la mesure de Poisson J décrit les instants et tailles des sauts d'un processus stochastique S_t . C'est-à-dire que J est la mesure de sauts définie comme:

$$J = J_S(\omega, \cdot) = \sum_{t \in [0, T]}^{\Delta S_t \neq 0} \delta_{(t, \Delta S_t)}$$

Il s'agit par exemple du cas d'un processus de Lévy (ou d'un processus de Poisson composé). De fait, si S est un processus de Lévy de mesure de Lévy ν , alors le mesure de sauts J_S est une mesure de Poisson d'intensité $\mu(dt, dx) = \nu(dx)dt$. Par suite, l'intégrale stochastique 3.18 est une somme de termes faisant intervenir les tailles et instants de sauts de S , soit:

$$\int_0^T \int_{\mathbb{R}^d} \phi(s, y) J(ds, dy) = \sum_{t \in [0, T]}^{\Delta S_t \neq 0} \phi(t, \Delta S_t)$$

Exemple 3.1.14 (Processus de Poisson) Dans le cas où S est un processus de Poisson alors $\phi(\omega, s, y) = \psi_s(\omega) y$ où $\psi = \sum_i \psi_i \mathbf{1}_{[t_i, t_{i+1})}$ est constante entre deux sauts successifs. Par suite, l'intégrale par rapport à J_S est juste une intégrale stochastique par rapport à S :

$$\int_0^T \int_{\mathbb{R}^d} \phi(s, y) J(ds, dy) = \int_0^T \int_{\mathbb{R}^d} \psi_s y J(ds, dy) = \sum_{t \in [0, T]}^{\Delta S_t \neq 0} \psi_t \Delta S_t = \int_0^T \psi_t dS_t$$

et peut s'interpréter comme dans le cas de l'intégrale d'Itô comme le gain lié à la stratégie ψ dont les transactions dans l'actif S ont lieu aux dates $(t_i)_{i \geq 1}$.

De façon similaire, nous dérivons aussi l'intégrale stochastique compensée:

$$\begin{aligned} \int_0^T \int_{\mathbb{R}^d} \phi(t, y) \tilde{J}(dt, dy) &= \sum_{i,j=0}^{n-1, m-1} \phi_{ij} \tilde{J}([t_i, t_{i+1}) \times A_j) \\ &= \sum_{i,j=0}^{n-1, m-1} \phi_{ij} [J([t_i, t_{i+1}) \times A_j) - \mu([t_i, t_{i+1}) \times A_j)] \end{aligned}$$

Et nous avons les propriétés suivantes.

Proposition 3.1.15 Soit ϕ une fonction simple et \tilde{J} une mesure de Poisson compensée d'intensité la mesure $\mu(dx, dt) = \nu(dx)dt$. Alors, pour $t \geq 0$, nous avons:

1. (Isométrie) $\mathbb{E} [\mathcal{I}_t^2(\phi)] = \int_0^t \int_{\mathbb{R}^d} \phi^2(u, y) \mu(du, dy)$.
2. (Martingale) $\mathcal{I}_t(\phi) = \int_0^t \int_{\mathbb{R}^d} \phi(u, y) \tilde{J}(du, dy)$ est une martingale de carré intégrable.

3.2. Variation Quadratique

L'objectif d'un modèle stochastique qui décrit l'évolution des marchés financiers est de pouvoir modéliser le comportement aléatoire des nombreuses variables qui affectent la dynamique des prix. D'un point de vue statistique, une telle représentation doit alors permettre de pouvoir capturer la notion de volatilité du marché et partant de volatilité stochastique. L'équipement introduit précédemment va alors permettre de donner du sens aux concepts statistiques de variance et de covariance d'un point de vue aléatoire.

3.2.1 Lien avec la Volatilité Réalisée

Le concept de volatilité réalisée a été initialement définie dans le contexte de la mesure de la volatilité. Etant donné un processus X observé à intervalles de temps discret sur la grille $\Pi = \{0 = t_0 < t_1 < \dots < t_n = T\}$, la volatilité réalisée est définie comme:

$$V_X(\Pi) = \sum_{t_i \in \Pi} (X_{t_{i+1}} - X_{t_i})^2$$

Si nous réécrivons chaque terme de la somme:

$$(X_{t_{i+1}} - X_{t_i})^2 = X_{t_{i+1}}^2 - X_{t_i}^2 - 2X_{t_i}(X_{t_{i+1}} - X_{t_i})$$

la volatilité réalisée peut alors s'écrire sous la forme d'une suite de Riemann:

$$X_T^2 - X_0^2 - 2 \sum_{t_i \in \Pi} X_{t_i} (X_{t_{i+1}} - X_{t_i}) \quad (3.19)$$

Si nous considérons maintenant que X est une semimartingale, cf. Définition 3.1.10, avec $X_0 = 0$ et que nous définissons le processus $X_- = (X_{t-})_{t \in [0, T]}$, on peut montrer que la somme de Riemann 3.19 converge uniformément en probabilité vers la variable aléatoire:

$$[X, X]_T := |X_T|^2 - 2 \int_0^T X_{u-} dX_u$$

appelée la variation quadratique de X sur $[0, T]$, qui est une variable aléatoire et pas un nombre à l'instar de la variance. Finalement, si nous répétons cette procédure sur l'intervalle $[0, t]$, nous pouvons définir le processus de variation quadratique.

Définition 3.2.1 (Variation quadratique) *Le processus de variation quadratique d'une semimartingale $(X_t)_{t \geq 0}$ est un processus càdlàg non-anticipatif défini par:*

$$[X, X]_t = |X_t|^2 - 2 \int_0^t X_{u-} dX_u, \quad t \in [0, T]$$

et satisfait les propriétés suivantes:

1. $[\cdot, \cdot]_t$ est un processus aléatoire croissant.

2. Les sauts de $[X, X]$ sont reliés aux sauts de X selon $\Delta [X, X]_t = |\Delta X_t|^2$.
3. Si X est continu et ses trajectoires sont à variation finie, alors $[X, X] = 0$.

Contrairement à la variance, la variation quadratique n'est pas définie à partir du signe espérance: il s'agit d'une propriété de trajectoire. Nous nous intéressons à présent à des exemples de processus dont la variation quadratique peut être explicitement calculée.

Exemple 3.2.2 (Processus de Wiener) Soit $B_t = \sigma W_t$, où W est un mouvement brownien standard, alors $[B, B]_t = \sigma^2 t$, cf. Section 2.1.2 pour plus de détails. Le cas du mouvement brownien est très particulier. En effet, la variation quadratique est égale à la variance du processus, ce qui donne une explication au terme de volatilité réalisée.

Exemple 3.2.3 Soit le processus stochastique défini par $X_t = \int_0^t \sigma_s dW_t$ où W_t est un mouvement brownien et $(\sigma_t)_{t \in [0, T]}$ un processus càdlàg. Alors:

$$[X, X]_t = \int_0^t \sigma_s^2 ds$$

Et en particulier si $[X, X] = 0$, alors $X = 0$ presque sûrement.

Proposition 3.2.4 (Processus de Poisson et de comptage)

1. Si N est un processus de Poisson, alors par définition $[N, N]_t = N_t$, $t \in [0, T]$.
2. Si N_t est un processus de comptage d'instantants de sauts T_i et de taille des sauts Z_i , alors

$$X_t = \sum_{i=1}^{N_t} Z_i \Rightarrow [X, X]_t = \sum_{i=1}^{N_t} |Z_i|^2 = \sum_{0 \leq s \leq t} |\Delta X_s|^2, \quad t \in [0, T]$$

Exemple 3.2.5 (Mesure de Poisson) Soit J une mesure aléatoire de Poisson sur $[0, T] \times \mathbb{R}^d$ d'intensité la mesure $\mu(dt, dx)$ et ψ une fonction simple $\psi : [0, T] \times \mathbb{R}^d \mapsto \mathbb{R}$. Nous définissons alors le processus X comme l'intégrale de ψ par rapport à J , soit:

$$X_t = \int_0^t \int_{\mathbb{R}^d} \psi(s, y) J(ds, dy)$$

Alors, la variation quadratique de X est donnée par:

$$[X, X]_t = \int_0^t \int_{\mathbb{R}^d} |\psi(s, y)|^2 J(ds, dy)$$

Exemple 3.2.6 (Processus de Lévy) Si X est un processus de Lévy de triplet (b, σ, ν) , sa variation quadratique est donnée par:

$$[X, X]_t = \sigma^2 t + \sum_{s \in [0, t]} |\Delta X_s|^2 = \sigma^2 t + \int_{[0, t]} \int_{\mathbb{R}} y^2 J_X(ds, dy)$$

3.2.2 Covariance Quadratique

Le concept de volatilité réalisée a un analogue dans le cas multidimensionnel, qui peut être défini de la sorte: étant donnés deux processus X et Y observés à dates discrètes sur une grille $\Pi = \{0 = t_0 < t_1 < \dots < t_{n+1} < T\}$, nous définissons la covariance comme:

$$\sum_{t_i \in \Pi} (X_{t_{i+1}} - X_{t_i}) (Y_{t_{i+1}} - Y_{t_i})$$

A nouveau, si nous réécrivons chaque terme de la somme sous la forme:

$$X_{t_{i+1}}Y_{t_{i+1}} - X_{t_i}Y_{t_i} - Y_{t_i}(X_{t_{i+1}} - X_{t_i}) - X_{t_i}(Y_{t_{i+1}} - Y_{t_i})$$

qui peut à nouveau se réécrire comme une suite de Riemann:

$$X_T Y_T - X_0 Y_0 - \sum_{t_i \in \pi} \{Y_{t_i}(X_{t_{i+1}} - X_{t_i}) + X_{t_i}(Y_{t_{i+1}} - Y_{t_i})\} \quad (3.20)$$

Quand X et Y sont des semimartingales, 3.20 converge en probabilité vers:

$$X_T Y_T - X_0 Y_0 - \int_0^T X_{t-} dY_t - \int_0^T Y_{t-} dX_t$$

qui est une variable aléatoire, appelée la covariance quadratique de X et Y sur $[0, T]$.

Définition 3.2.7 (Covariance quadratique) *Etant données deux semimartingales X et Y , le processus de covariance quadratique $[X, Y]$ est la semimartingale définie par:*

$$[X, Y]_t = X_t Y_t - X_0 Y_0 - \int_0^t X_{s-} dY_s - \int_0^t Y_{s-} dX_s, \quad t \in [0, T] \quad (3.21)$$

et satisfait les propriétés suivantes:

1. $[\cdot, \cdot]_t$ est un processus aléatoire croissant.
2. (Identité polaire) $[X, Y] = \frac{1}{4}([X + Y, X + Y] - [X - Y, X - Y])$
3. Si ϕ et ψ sont des processus intégrables, alors:

$$\left[\int \phi dX, \int \psi dY \right]_t = \int_0^t \phi \psi d[X, Y]$$

Le résultat suivant est juste une reformulation de 3.21. Il va se révéler particulièrement utile dans la suite lors de la dérivation du calcul stochastique.

Proposition 3.2.8 (Règle de multiplication) *Si X et Y sont des semimartingales, alors:*

$$X_t Y_t = X_0 Y_0 + \int_0^t X_{s-} dY_s + \int_0^t Y_{s-} dX_s + [X, Y]_t, \quad t \in [0, T] \quad (3.22)$$

De façon importante, nous remarquons que la covariance quadratique entre deux processus est nulle si l'un est continu et l'autre ne possède pas d'intégrale d'Itô: c'est le cas par exemple des processus mixtes Wiener-Poisson, cf. Section 4.1.2. De façon générale, la covariance quadratique est non nulle si (i) les processus possèdent tous les deux une partie en dW ou (ii) les processus ont des sauts simultanés. Nous avons alors le résultat.

Lemme 3.2.9 *Soit W_t un mouvement brownien et M_t un processus de Poisson (composé) compensé par rapport à la même filtration \mathcal{F}_t . Alors $[W, M]_t = 0$ pour $t \geq 0$.*

Nous donnons finalement quelques exemples où la covariance quadratique peut être calculée de façon explicite. Ceux-ci seront utiles pour la dérivation du calcul stochastique.

Exemple 3.2.10 (Processus de Wiener) *Soit $B_t^1 = \sigma^1 W_t^1$ et $B_t^2 = \sigma^2 W_t^2$ où W^1 et W^2 sont des mouvements browniens standard de corrélation ρ , alors $[B^1, B^2]_t = \rho \sigma^1 \sigma^2 t$.*

Exemple 3.2.11 *Soit $X_t = \int_0^t \sigma_s^1 dW_s^1$ et $Y_t = \int_0^t \sigma_s^2 dW_s^2$ où les σ^i sont des processus prédictifs et W^i des mouvements browniens standard corrélés tel que $\text{Cov}[W_t^1, W_t^2] = \rho t$, alors la covariance quadratique de X et Y est $[X, Y]_t = \int_0^t \rho \sigma_s^1 \sigma_s^2 ds$.*

Exemple 3.2.12 (Mesure de Poisson) *Soit J une mesure aléatoire de Poisson sur $[0, T] \times \mathbb{R}$ et W un processus de Wiener indépendant de J . Soit:*

$$X_t^i = X_0^i + \int_0^t \phi_s^i dW_s + \int_0^t \int_{\mathbb{R}^d} \psi^i(s, y) J(ds, dy), \quad i = 1, 2$$

alors la covariance quadratique $[X^1, X^2]$ est égale à:

$$[X^1, X^2] = \int_0^t \phi_s^1 \phi_s^2 ds + \int_0^t \int_{\mathbb{R}^d} \psi^1(s, y) \psi^2(s, y) J(ds, dy)$$

Le résultat précédent nous permet alors la spécialisation suivante, basée sur le concept de processus mixte (Wiener-Poisson), cf. Section 4.1.2.

Lemme 3.2.13 *Soit $X_t^i = X_0^i + R_t^i + I_t^i + J_t^i$ pour $i = 1, 2$ un processus mixte où $R_t^i = \int_0^t a_s^i ds$, $I_t^i = \int_0^t b_s^i dW_s$ où W_t est un processus de Wiener et J_t^i est un processus de sauts pur. Alors $(X_t^i)^c = X_0^i + R_t^i + I_t^i$ et:*

$$[X^1, X^2]_t = [(X^1)^c, (X^2)^c]_t + [J^1, J^2]_t = \int_0^t b_s^1 b_s^2 ds + \sum_{0 \leq s \leq t} (\Delta X_s^1 \Delta X_s^2)$$

3.3. Formule d'Itô

La définition de l'intégrale stochastique permet la dérivation d'un calcul, c'est-à-dire d'un ensemble de règles qui vont permettre de réaliser des opérations algébriques faisant intervenir cette intégrale, à l'instar de ce qui existe pour l'intégrale de Lebesgue.

Dans le cas présent, nous allons voir qu'une règle de dérivation en chaîne (*chain rule*) peut se généraliser depuis les règles du calcul ordinaire. Soit les fonctions $f : \mathbb{R} \rightarrow \mathbb{R}$ et $\psi : [0, \infty) \rightarrow \mathbb{R}$ de classe \mathcal{C}^1 . La règle de dérivation en chaîne habituelle donne:

$$(f(\psi(t)))' = f'(\psi(t))\psi'(t)$$

ou de façon équivalente:

$$f(\psi(t)) = f(\psi(0)) + \int_0^t f'(\psi(s))d\psi(s) \quad (3.23)$$

L'égalité 3.23 est vraie même si ψ est seulement à variations finies (et pas nécessairement différentiable) dès lors que nous considérons l'intégrale de droite dans le sens d'une intégrale de Stieltjes, qui est une généralisation de l'intégrale de Riemann, i.e.:

$$\int_a^b f(x)d\alpha(x) \approx \sum_{i=0}^{n-1} f(\xi_i)[\alpha(x_{i+1}) + \alpha(x_i)], \max(x_{i+1} - x_i) \rightarrow 0$$

où $a = x_0 < \dots < x_n = b$ est une partition de $[a, b]$.

Si nous posons $f(x) = x^2$, alors 3.23 devient:

$$\psi_t^2 - \psi_0^2 = 2 \int_0^t \psi_s d\psi_s$$

Par ailleurs, lorsque X est une semimartingale, nous avons depuis 3.22 que:

$$X_t^2 - X_0^2 = 2 \int_0^t X_{s-} dX_s + [X, X]_t$$

Nous retrouvons alors un résultat déjà souligné dans la Remarque 2.1.9 qui est qu'en générale $[X, X]_t \neq 0$ et donc les règles de calcul ordinaire ne peuvent être utilisées en l'état pour des processus stochastiques. Dans la suite, nous présentons la formule d'Itô qui permet d'obtenir un résultat du type 3.23 quand X est une semimartingale.

3.3.1 Cas du Mouvement Brownien

La relation 3.23 n'est plus vérifiée si nous remplaçons $\psi(\cdot)$ par $W(\omega)$ étant donné que la trajectoire $t \mapsto W_t(\omega)$ est à variation infinie pour presque tout $\omega \in \Omega$ (cf. Section 2.1). Nous devons alors ajouter un terme de correction de second ordre pour conserver l'égalité. Ce résultat est connu sous le nom de formule d'Itô et nous la dérivons par la suite.

Théorème 3.3.1 (Formule d'Itô - MB) Soit $f(t, x) : \mathbb{R}^2 \rightarrow \mathbb{R}$ de classe \mathcal{C}^2 et $(W_t)_{t \geq 0}$ un mouvement brownien. Alors pour $t \geq 0$:

$$f(t, W_t) - f(0, W_0) = \int_0^t f_s(s, W_s) ds + \int_0^t f_x(s, W_s) dW_s + \frac{1}{2} \int_0^t f_{xx}(s, W_s) ds \quad (3.24)$$

Soit, sous forme différentielle:

$$df(t, W_t) = f_t(t, W_t) dt + f_x(t, W_t) dW_t + \frac{1}{2} f_{xx}(t, W_t) dt$$

Idée. Soit la fonction $f(t, x)$ dépendantes des variables (t, x) , une application du théorème de Taylor à l'ordre 2 nous donne:

$$\begin{aligned} f(t_{i+1}, x_{i+1}) - f(t_i, x_i) &= f_t(t_i, x_i)(t_{i+1} - t_i) + f_x(t_i, x_i)(x_{i+1} - x_i) \\ &+ \frac{1}{2} f_{xx}(t_i, x_i)(x_{i+1} - x_i)^2 + f_{tx}(t_i, x_i)(t_{i+1} - t_i)(x_{i+1} - x_i) \\ &+ \frac{1}{2} f_{tt}(t_i, x_i)(t_{i+1} - t_i)^2 + \text{termes d'ordre supérieur} \end{aligned} \quad (3.25)$$

pour une partition $\Pi = \{0 = t_0 < \dots < t_n = t\}$ de $[0, t]$. Soit, si nous posons $x_i = W_i$ et prenons la somme de 3.25, il vient:

$$\begin{aligned} f(t, W_t) - f(0, W_0) &= \sum_{i=0}^{n-1} [f(t_{i+1}, W_{i+1}) - f(t_i, W_i)] \\ &= \sum_{i=0}^{n-1} f_t(t_i, W_i)(t_{i+1} - t_i) + \sum_{i=0}^{n-1} f_x(t_i, W_i)(W_{i+1} - W_i) \\ &+ \frac{1}{2} \sum_{i=0}^{n-1} f_{xx}(t_i, W_i)(W_{i+1} - W_i)^2 + \sum_{i=0}^{n-1} f_{tx}(t_i, W_i)(t_{i+1} - t_i)(W_{i+1} - W_i) \\ &+ \frac{1}{2} \sum_{i=0}^{n-1} f_{tt}(t_i, W_i)(t_{i+1} - t_i)^2 + \text{termes d'ordre supérieur} \end{aligned} \quad (3.26)$$

Lorsque nous passons à la limite dans 3.26, soit $\|\Pi\| \rightarrow 0$, le premier terme converge vers l'intégrale de Lebesgue, soit:

$$\lim_{\|\Pi\| \rightarrow 0} \sum_{i=0}^{n-1} f_t(t_i, W_i)(t_{i+1} - t_i) = \int_0^t f_t(s, W_s) ds$$

le second terme converge vers l'intégrale d'Itô, soit:

$$\lim_{\|\Pi\| \rightarrow 0} \sum_{i=0}^{n-1} f_x(t_i, W_i)(W_{i+1} - W_i) = \int_0^t f_x(s, W_s) dW_s$$

Par ailleurs, le troisième terme converge lui aussi vers une intégrale de Lebesgue:

$$\lim_{\|\Pi\| \rightarrow 0} \frac{1}{2} \sum_{i=0}^{n-1} f_{xx}(t_i, W_i)(W_{i+1} - W_i)^2 = \frac{1}{2} \int_0^t f_{xx}(s, W_s) ds$$

car nous pouvons remplacer, à la limite et non de façon exacte, le terme $(W_{i+1} - W_i)^2$ par $t_{i+1} - t_i$. Nous voulons à présent montrer que les quatrième et cinquième termes de 3.26 convergent vers 0 quand $\|\Pi\| \rightarrow 0$. A cette fin, nous notons pour le quatrième terme que:

$$\begin{aligned} & \lim_{\|\Pi\| \rightarrow 0} \left| \sum_{i=0}^{n-1} f_{tx}(t_i, W_i) (t_{i+1} - t_i) (W_{i+1} - W_i) \right| \\ & \leq \lim_{\|\Pi\| \rightarrow 0} \sum_{i=0}^{n-1} |f_{tx}(t_i, W_i)| (t_{i+1} - t_i) |W_{i+1} - W_i| \\ & \leq \lim_{\|\Pi\| \rightarrow 0} \max_{0 \leq k \leq n-1} |W_{k+1} - W_k| \lim_{\|\Pi\| \rightarrow 0} \sum_{i=0}^{n-1} |f_{tx}(t_i, W_i)| (t_{i+1} - t_i) \\ & = 0 \cdot \int_0^T f_{tx}(s, W_s) ds = 0 \end{aligned}$$

Un raisonnement analogue peut alors être mené pour le cinquième terme. ■

Exemple 3.3.2 Si nous voulons calculer $\int_0^t W_s dW_s$, il nous suffit de poser $f(x) = x^2$ et d'appliquer 3.24. Nous obtenons alors:

$$\int_0^t W_s dW_s = \frac{1}{2} (W_t^2 - t)$$

qui est 3.13 dérivée dans l'Exemple 3.1.5 de façon directe.

3.3.2 Cas des Processus d'Itô

Nous étendons dans cette Section la formule d'Itô au cas de processus plus généraux que le mouvement brownien. Il s'agit des processus d'Itô. A l'exception des processus de sauts pur, tous les processus stochastiques (continus) sont des processus d'Itô.

Définition 3.3.3 (Processus d'Itô) Soit $(W_t)_{t \geq 0}$ un mouvement brownien et $(\mathcal{F}_t)_{t \geq 0}$ une filtration associée. Un processus d'Itô est un processus stochastique de la forme:

$$X_t = X_0 + \int_0^t a_u du + \int_0^t b_u dW_u \tag{3.27}$$

où X_0 est une constante et a_u, b_u sont des processus stochastiques \mathcal{F} -adaptés.

Afin de comprendre la volatilité associée à un processus d'Itô, nous devons déterminer à quel taux il accumule la variation quadratique dans le temps.

Lemme 3.3.4 La variation quadratique jusqu'en t d'un processus d'Itô est donnée par:

$$[X, X]_t = \int_0^t b_u^2 du, \quad t \geq 0 \tag{3.28}$$

La relation 3.28 peut par ailleurs se réécrire:

$$dX_t dX_t = b_t^2 dt$$

Ainsi, à chaque date t , le processus X_t accumule la variation quadratique au taux b_t^2 par unité de temps. Cette variation est uniquement due à la variation quadratique de l'intégrale d'Itô $I_t = \int_0^t b_u dW_u$, l'intégrale ordinaire $R_t = \int_0^t a_u du$ ayant une variation quadratique nulle, elle ne participe pas à la variation quadratique de X . Nous pouvons à présent définir l'intégrale stochastique par rapport à un processus d'Itô.

Définition 3.3.5 Soit $(X_t)_{t \geq 0}$ un processus d'Itô et ϕ_t un processus adapté. L'intégrale au sens d'Itô, pour $t \geq 0$, se définit comme:

$$\int_0^t \phi_u dX_u = \int_0^t \phi_u a_u du + \int_0^t \phi_u b_u dW_u$$

Ce qui nous permet de généraliser la formule d'Itô au cas où X est un processus d'Itô.

Théorème 3.3.6 (Formule d'Itô - Processus d'Itô) Soit $(X_t)_{t \geq 0}$ un processus d'Itô et $f(t, x) : \mathbb{R}^2 \rightarrow \mathbb{R}$ une fonction de classe \mathcal{C}^2 . Alors, pour $t \geq 0$:

$$f(t, X_t) - f(0, X_0) = \int_0^t f_s(s, X_s) ds + \int_0^t f_x(s, X_s) dX_s + \frac{1}{2} \int_0^t f_{xx}(s, X_s) d[X, X]_s$$

Soit:

$$\begin{aligned} f(t, X_t) - f(0, X_0) &= \int_0^t f_s(s, X_s) ds + \int_0^t f_x(s, X_s) b_s dW_s \\ &\quad + \int_0^t f_x(s, X_s) a_s ds + \frac{1}{2} \int_0^t f_{xx}(s, X_s) b_s^2 ds \end{aligned} \quad (3.29)$$

Exemple 3.3.7 (Mouvement brownien géométrique) Soit le processus d'Itô:

$$X_t = \int_0^t \left(\alpha_s - \frac{1}{2} \sigma_s^2 \right) ds + \int_0^t \sigma_s dW_s$$

Alors:

$$dX_t = \left(\alpha_t - \frac{1}{2} \sigma_t^2 \right) dt + \sigma_t dW_t$$

Et:

$$dX_t dX_t = \sigma_t^2 dW_t dW_t = \sigma_t^2 dt$$

Soit $S_t = f(X_t)$ avec $f(x) = S_0 e^x$. Une application de la formule d'Itô donne:

$$\frac{dS_t}{S_t} = \alpha_t dt + \sigma_t dW_t$$

soit:

$$S_t = S_0 \exp \left(\int_0^t \left(\alpha_s - \frac{1}{2} \sigma_s^2 \right) ds + \int_0^t \sigma_s dW_s \right)$$

Le résultat suivant permet de caractériser la distribution d'une intégrale d'Itô.

Théorème 3.3.8 (Distribution de l'intégrale d'Itô) *Soit $(W_t)_{t \geq 0}$ un mouvement brownien et σ_t une fonction déterministe. On définit $I_t = \int_0^t \sigma_s dW_s$. Alors pour $t \geq 0$, la variable aléatoire I_t est gaussienne d'espérance 0 et de variance $\int_0^t \sigma_s^2 ds$.*

Preuve. Comme I_t est une intégrale d'Itô, cf. Section 3.1.1, avec $I_0 = 0$, l'isométrie, cf. Théorème 3.1.4, donne:

$$\mathbb{E}[I_t] = 0, \text{Var}[I_t] = \mathbb{E}[I_t^2] = \int_0^t \sigma_s^2 ds$$

Il nous faut à présent montrer que la variable aléatoire I_t est gaussienne, ce qui peut être fait en ayant recours à sa caractérisation par la fonction génératrice des moments, cf. Section 1.2.3. Nous voulons montrer que:

$$\mathbb{E}[e^{uI_t}] = \exp\left(\frac{1}{2}u^2 \int_0^t \sigma_s^2 ds\right), u \in \mathbb{R}$$

ce qui est équivalent à montrer que:

$$\mathbb{E}\left[\exp\left(uI_t - \frac{1}{2}u^2 \int_0^t \sigma_s^2 ds\right)\right] = 1$$

soit:

$$\mathbb{E}\left[\exp\left(\int_0^t u\sigma_s dW_s - \frac{1}{2} \int_0^t (u\sigma_s)^2 ds\right)\right] = 1 \quad (3.30)$$

Or le processus:

$$\exp\left(\int_0^t u\sigma_s dW_s - \frac{1}{2} \int_0^t (u\sigma_s)^2 ds\right)$$

est une martingale, ce qui valide 3.30 et conclut la preuve. ■

Exemple 3.3.9 (Modèle de Vasicek) *Le modèle de Vasicek pour le taux d'intérêt de court-terme R_t s'écrit:*

$$dR_t = (a - bR_t) dt + \sigma dW_t \quad (3.31)$$

où a, b, σ sont des constantes positives. Une solution de 3.31 peut être explicitée sous la forme:

$$R_t = e^{-bt} R_0 + \frac{a}{b} (1 - e^{-bt}) + \sigma e^{-bt} \int_0^t e^{bs} dW_s \quad (3.32)$$

Pour vérifier ceci, nous utilisons la formule d'Itô pour:

$$f(t, x) = e^{-bt} R_0 + \frac{a}{b} (1 - e^{-bt}) + \sigma e^{-bt} x$$

avec $X_t = \int_0^t e^{bs} dW_s$. Soit:

$$\begin{aligned} df(t, X_t) &= f_t(t, X_t) dt + f_x(t, X_t) dX_t + \frac{1}{2} f_{xx}(t, X_t) dX_t dX_t \\ &= (a - bf(t, X_t)) dt + \sigma dW_t \end{aligned}$$

ce qui prouve que $f(t, X_t)$ satisfait 3.31. Par ailleurs, le Théorème 3.3.8 nous apprend que le terme $\int_0^t e^{bs} dW_s$ dans 3.32 est gaussien d'espérance nulle et de variance:

$$\int_0^t e^{2bs} ds = \frac{1}{2b} (e^{2bt} - 1)$$

Exemple 3.3.10 (Modèle de Cox-Ingersoll-Ross) Le modèle CIR pour R_t satisfait:

$$dR_t = (a - bR_t) dt + \sigma \sqrt{R_t} dW_t \quad (3.33)$$

Dans ce cas, une solution à 3.33 ne peut pas être déterminée, mais sa distribution de probabilité pour $t \geq 0$ peut l'être.

3.3.3 Cas des Processus de Poisson

Les processus de Poisson sont des processus admettant des discontinuités. Il est alors possible de comprendre la formule de changement de variable par une analogie avec les fonctions constantes par morceaux. A cette fin, considérons la fonction $x : [0, T] \rightarrow \mathbb{R}$ admettant un nombre fini de discontinuités aux dates $(t_i)_{i=1}^n$ mais qui est régulière sur chaque intervalle $[t_i, t_{i+1})$. Une telle fonction peut être représentée sous la forme:

$$x_t = \int_0^t b_s ds + \sum_{i, t_i \leq t} \Delta x_i, \quad \Delta x_i = x_{t_i} - x_{t_i-} \quad (3.34)$$

Soit à présent une fonction $f : \mathbb{R} \rightarrow \mathbb{R}$ régulière sur chaque intervalle $[t_i, t_{i+1})$. Par suite, nous pouvons appliquer la formule 3.23 à $f(x)$ et nous avons:

$$f(x_{t_{i+1}-}) - f(x_{t_i}) = \int_{t_i}^{t_{i+1}-} f'(x_t) x'_t dt = \int_{t_i}^{t_{i+1}-} f'(x_t) b_t dt$$

En chaque point de discontinuité, $f(x_t)$ a un saut égal à:

$$f(x_{t_i}) - f(x_{t_i-}) = f(x_{t_i-} + \Delta x_i) - f(x_{t_i-})$$

Par suite, la variation de f sur $[0, t]$ est égale à:

$$\begin{aligned} f(x_t) - f(x_0) &= \sum_{i=0}^n [f(x_{t_{i+1}}) - f(x_{t_i})] \\ &= \sum_{i=0}^n [f(x_{t_{i+1}}) - f(x_{t_{i+1}-}) + f(x_{t_{i+1}-}) - f(x_{t_i})] \\ &= \sum_{i=1}^{n+1} [f(x_{t_{i-}} + \Delta x_i) - f(x_{t_{i-}})] + \sum_{i=0}^n \int_{t_i}^{t_{i+1}-} f'(x_t) b_t dt \end{aligned}$$

Nous obtenons alors le résultat suivant.

Proposition 3.3.11 Soit x une fonction de classe \mathcal{C}^1 constante par morceaux telle que:

$$x_t = \int_0^t b_s ds + \sum_{i, t_i \leq t} \Delta x_i, \quad \Delta x_i = x_{t_i} - x_{t_{i-}}$$

Alors pour toute fonction $f : \mathbb{R} \rightarrow \mathbb{R}$ de classe \mathcal{C}^1 , nous avons:

$$f(x_t) - f(x_0) = \int_0^t f'(x_{t-}) b_t dt + \sum_{i=1}^{n+1} [f(x_{t_{i-}} + \Delta x_i) - f(x_{t_{i-}})]$$

Cas des Mesures de Poisson

La dérivation précédente va permettre de traiter le cas de processus de sauts pur. A cette fin, considérons le processus stochastique $(X_t)_{t \geq 0}$ dont les trajectoires sont de la forme 3.34, soit:

$$X_t(\omega) = X_0 + \int_0^t b_s(\omega) ds + \sum_{i=1}^{N_t(\omega)} \Delta X_i(\omega)$$

où les $\Delta X_i = X_{t_i} - X_{t_{i-}}$ représentent les instants de sauts et N_t est le nombre aléatoire de sauts survenus en t . La Proposition 3.3.11 s'applique et nous avons alors:

$$f(X_t) - f(X_0) = \int_0^t b_s f'(x_{s-}) ds + \sum_{0 \leq s \leq t}^{\Delta X_s \neq 0} [f(X_{s-} + \Delta X_s) - f(X_{s-})]$$

Cette formule de changement de variable est valide quelle que soit la structure probabiliste du processus X . Si nous précisons cette structure, en donnant les distributions suivies par les incréments $t_{i+1} - t_i$ (exponentielle, par exemple) et par ΔX_i (densité ν , gaussienne, par exemple), nous définissons alors un processus de Poisson composé, auquel nous pouvons associer une mesure aléatoire sur $[0, T] \times \mathbb{R}$ qui décrit les tailles et instants des sauts. Soit:

$$J_X = \sum_{n \geq 1} \delta_{(t_n, \Delta X_{t_n})}$$

qui est une mesure de Poisson d'intensité $\mu(dt, dy) = \lambda \nu(dy) dt$. Et nous avons alors:

$$\sum_{0 \leq s \leq t}^{\Delta X_s \neq 0} [f(X_{s-} + \Delta X_s) - f(X_{s-})] = \int_0^t \int_{\mathbb{R}} [f(X_{s-} + y) - f(X_{s-})] J_X(ds, dy)$$

Nous avons alors le résultat suivant.

Théorème 3.3.12 (Formule d'Itô - Mesure de Poisson) Soit X_t un processus donné par:

$$X_t = \int_0^t b_s ds + \sum_{i=1}^{N_t} Y_i, \quad t \geq 0$$

où b_s est un processus non-anticipatif, N_t un processus de comptage donnant le nombre de sauts survenus sur $[0, t]$ aux dates $(t_n)_{n \geq 1}$ et Y_i la taille du i -ème saut. Alors la mesure de Poisson associée aux sauts de X est donnée par:

$$J_X = \sum_{n \geq 1} \delta_{(t_n, Y_n)}$$

Si $f(t, x) : \mathbb{R}^2 \rightarrow \mathbb{R}$ est une fonction de classe \mathcal{C}^2 . Alors pour $t \geq 0$:

$$\begin{aligned} f(t, X_t) - f(0, X_0) &= \int_0^t [f_s(s, X_{s-}) + f_x(s, X_{s-}) b_s] ds \\ &+ \sum_{n \geq 1} [f(s, X_{s-} + \Delta X_s) - f(s, X_{s-})] \end{aligned} \quad (3.35)$$

ou encore:

$$\begin{aligned} f(t, X_t) - f(0, X_0) &= \int_0^t [f_s(s, X_{s-}) + f_x(s, X_{s-}) b_s] ds \\ &+ \int_0^t \int_{\mathbb{R}} [f(s, X_{s-} + y) - f(s, X_{s-})] J_X(ds, dy) \end{aligned}$$

Cas des Processus Mixtes

Nous considérons à présent le cas des processus mixtes (Wiener-Poisson), cf. Section 4.1.2, que nous écrivons sous la forme:

$$X_t = X_0 + R_t + I_t + J_t \quad (3.36)$$

avec $R_t = \int_0^t a_s ds$ une intégrale ordinaire et $I_t = \int_0^t b_s dW_s$ une intégrale d'Itô et nous dénotons par $X_t^c = X_0 + R_t + I_t$ la partie continue de X_t . La formule d'Itô appliquée à X_t^c , cf. Théorème 3.3.6, pour une fonction $f(x)$ de classe \mathcal{C}^2 , nous donne:

$$\begin{aligned} df(X_t^c) &= f'(X_t^c) dX_t^c + \frac{1}{2} f''(X_t^c) dX_t^c dX_t^c \\ &= f'(X_t^c) b_t dW_t + f'(X_t^c) a_t dt + \frac{1}{2} f''(X_t^c) b_t^2 dt \end{aligned} \quad (3.37)$$

Entre les dates de sauts de J , l'analogie de 3.37 reste vraie:

$$\begin{aligned} df(X_t) &= f'(X_t) dX_t + \frac{1}{2} f''(X_t) dX_t dX_t \\ &= f'(X_t) b_t dW_t + f'(X_t) a_t dt + \frac{1}{2} f''(X_t) b_t^2 dt \\ &= f'(X_t) dX_t^c + \frac{1}{2} f''(X_t^c) dX_t^c dX_t^c \end{aligned} \quad (3.38)$$

Lorsqu'un saut survient dans X , soit de X_{t-} à X_t , il y a aussi un saut dans $f(X)$, de $f(X_{t-})$ à $f(X_t)$. Dès lors, lorsqu'on intègre 3.38 entre 0 et t , il est nécessaire de rajouter tous les sauts qui ont eu lieu entre ces deux dates. Soit le résultat suivant.

Théorème 3.3.13 (Formule d'Itô - Processus mixte) Soit $(X_t)_{t \geq 0}$ un processus mixte et $f(t, x) : \mathbb{R}^2 \rightarrow \mathbb{R}$ une fonction de classe C^2 . Alors pour $t \geq 0$:

$$\begin{aligned} f(t, X_t) - f(0, X_0) &= \int_0^t f_s(s, X_s) dX_s^c + \frac{1}{2} \int_0^t f_{xx}(X_s) dX_s^c dX_s^c \\ &+ \sum_{0 \leq s \leq t} [f(s, X_{s-} + \Delta X_s) - f(s, X_{s-})] \end{aligned} \quad (3.39)$$

Soit :

$$\begin{aligned} f(t, X_t) - f(0, X_0) &= \int_0^t [f_s(s, X_s) + f_x(s, X_s) b_s] ds \\ &+ \frac{1}{2} \int_0^t f_{xx}(s, X_s) \sigma_s^2 ds + \int_0^t f_x(s, X_s) \sigma_s dW_s \\ &+ \sum_{0 \leq s \leq t} [f(s, X_{s-} + \Delta X_s) - f(X_{s-})] \end{aligned}$$

ou sous forme différentielle :

$$\begin{aligned} df(t, X_t) &= f_t(t, X_t) dt + b_t f_x(t, X_t) dt + \frac{1}{2} \sigma_t^2 f_{xx}(t, X_t) dt \\ &+ \sigma_t f_x(t, X_t) dW_t + [f(X_{t-} + \Delta X_t) - f(X_{t-})] \end{aligned}$$

Exemple 3.3.14 (Processus de Poisson Géométrique) Soit le processus :

$$S_t = S_0 \exp(N_t \log(1 + \sigma) - \lambda \sigma t) = S_0 e^{-\lambda \sigma t} (1 + \sigma)^{N_t}$$

où $\sigma > -1$ est une constante. Nous posons $S_t = S_0 f(X_t)$ avec :

$$X_t = N_t \log(1 + \sigma) - \lambda \sigma t$$

tel que $X_t^c = -\lambda \sigma t$ et $J_t = N_t \log(1 + \sigma)$. Une application de la formule d'Itô nous donne :

$$\begin{aligned} S_t &= f(X_t) \\ &= f(X_0) - \lambda \sigma \int_0^t f'(X_u) du + \sum_{0 \leq u \leq t} [f(X_u) - f(X_{u-})] \\ &= S_0 - \lambda \sigma \int_0^t S_u du + \sum_{0 \leq u \leq t} [S_u - S_{u-}] \end{aligned}$$

Si un saut survient à la date u , alors $S_u = (1 + \sigma) S_{u-}$, soit :

$$S_u - S_{u-} = \sigma S_{u-}$$

d'où :

$$S_u - S_{u-} = \sigma S_{u-} \Delta N_u$$

Ce qui nous amène à :

$$\sum_{0 \leq u \leq t} [S_u - S_{u-}] = \sum_{0 \leq u \leq t} \sigma S_{u-} \Delta N_u = \sigma \int_0^t S_{u-} dN_u$$

Soit finalement:

$$S_t = S_0 - \lambda \sigma \int_0^t S_{u-} du + \sigma \int_0^t S_{u-} dN_u = S_0 + \sigma \int_0^t S_{u-} dM_u$$

où est la version compensée du Processus de Poisson N , soit $M_u = N_u - \lambda u$.

3.4. Changement de Mesure

Un processus stochastique X (de Wiener, de Poisson ou de Lévy) peut être considéré comme une variable aléatoire définie sur l'espace des trajectoires possibles Ω , doté d'une σ -algèbre \mathcal{F} qui renseigne quels événements sont mesurables, c'est-à-dire quelles trajectoires sont susceptibles de se réaliser. La distribution de probabilité de X définit une mesure de probabilité \mathbb{P}^X sur cet espace de trajectoires. Considérons à présent un second processus Y et notons \mathbb{P}^Y sa distribution sur l'espace Ω . Nous nous intéressons dans cette Section aux conditions sous lesquelles les mesures \mathbb{P}^X et \mathbb{P}^Y sont équivalentes, dans le sens où elles définissent le même ensemble de scénarios possibles, soit:

$$\mathbb{P}^X(A) = 1 \Leftrightarrow \mathbb{P}^Y(A) = 1, \forall A \in \mathcal{F}$$

et les mesures \mathbb{P}^X et \mathbb{P}^Y sont alors dites *équivalentes* (comparables), dans le sens où les processus stochastiques X et Y définissent le même ensemble d'évolutions possibles. La construction d'un nouveau processus sur le même ensemble de trajectoires, mais en affectant de nouvelles probabilités aux événements est appelée un changement de mesure.

3.4.1 Construction

A partir de l'espace de probabilité $(\Omega, \mathcal{F}, \mathbb{P})$ et d'une variable aléatoire Z non nulle et vérifiant $\mathbb{E}[Z] = 1$, le théorème de Radon-Nikodym, cf. Théorème 1.1.7, permet de définir une nouvelle mesure de probabilité $\tilde{\mathbb{P}}$ telle que:

$$\tilde{\mathbb{P}}(A) = \int_A Z(\omega) d\mathbb{P}(\omega), \forall A \in \mathcal{F} \quad (3.40)$$

Dès lors, une variable aléatoire X sur $(\Omega, \mathcal{F}, \mathbb{P})$ dispose de deux opérations d'espérance: l'une sous la mesure originale \mathbb{P} , notée \mathbb{E} , et une seconde sous $\tilde{\mathbb{P}}$ dénotée $\tilde{\mathbb{E}}$ ou $\mathbb{E}^{\tilde{\mathbb{P}}}$. On passe de l'une à l'autre par la relation:

$$\tilde{\mathbb{E}}[X] = \mathbb{E}[ZX]$$

et Z est appelée la dérivée de Radon-Nikodym de $\tilde{\mathbb{P}}$ par rapport à \mathbb{P} pour \mathcal{F} , notée:

$$Z = \frac{d\tilde{\mathbb{P}}}{d\mathbb{P}} \Big|_{\mathcal{F}} \quad (3.41)$$

Il est par ailleurs possible de définir le processus de dérivée de Radon-Nikodym de la sorte:

$$Z_t = \mathbb{E}[Z | \mathcal{F}_t], \quad t \in [0, T]$$

qui est martingale par la loi des espérances itérées:

$$\mathbb{E}[Z|\mathcal{F}_s] = \mathbb{E}[\mathbb{E}[Z|\mathcal{F}_t]|\mathcal{F}_s] = \mathbb{E}[Z|\mathcal{F}_s] = Z_s, \quad 0 \leq s \leq t \leq T$$

Cette dernière relation permet d'établir le lemme suivant.

Lemme 3.4.1 *Soit Y une variable aléatoire \mathcal{F}_t -mesurable. Alors pour $0 \leq s \leq t \leq T$:*

$$\begin{aligned} \tilde{\mathbb{E}}[Y] &= \mathbb{E}[YZ_t] \\ \tilde{\mathbb{E}}[Y|\mathcal{F}_s] &= \frac{1}{Z_s} \mathbb{E}[YZ_t|\mathcal{F}_s] \end{aligned}$$

Exponentielle Stochastique

Nous introduisons à présent un processus stochastique central dans la dérivation du changement de mesure. Il s'agit de l'exponentielle stochastique, dite de Doléans-Dade. Nous l'établissons pour un processus mixte de la forme:

$$X_t = X_0 + \int_0^t a_s ds + \int_0^t b_s dW_s + \sum_{i=0}^{N_t} Y_i \quad (3.42)$$

où a_s est un processus càdlàg non-anticipatif, W_t un mouvement brownien, N_t un processus de comptage donnant le nombre de sauts survenus sur l'intervalle $[0, t]$ et Y_i la taille du i -ème saut. Nous notons aussi $X_t^c = X_0 + \int_0^t a_s ds + \int_0^t b_s dW_s$ la partie continue de X .

Proposition 3.4.2 (Exponentielle de Doléans-Dade) *Soit $(X_t)_{t \geq 0}$ un processus mixte de la forme 3.42. Alors, il existe un unique processus càdlàg Z_t pour $t \geq 0$ tel que:*

$$dZ_t = Z_{t-} dX_t, \quad Z_0 = 1$$

et Z est donné par:

$$Z_t = \exp\left(X_t^c - \frac{1}{2}[X^c, X^c]_t\right) \prod_{0 \leq s \leq t} (1 + \Delta X_s)$$

Preuve. Soit le processus continu:

$$Y_t = \exp\left(\int_0^t a_s ds + \int_0^t b_s dW_s - \frac{1}{2} \int_0^t b_s^2 ds\right) = \exp\left(X_t^c - \frac{1}{2}[X^c, X^c]_t\right)$$

Par application de la formule d'Itô, nous avons:

$$dY_t = Y_t dX_t^c = Y_{t-} dX_t^c$$

Soit par ailleurs le processus discontinu:

$$K_t = \prod_{0 \leq s \leq t} (1 + \Delta X_s)$$

Il est clair que $Z_t = Y_t K_t$. Si X a un saut à la date t , alors $K_t = K_{t-} (1 + \Delta X_t)$ et:

$$\Delta K_t = K_t - K_{t-} = K_{t-} \Delta X_t$$

Comme Y et K ne partagent pas de points de (dis)continuité, $[Y, K]_t = 0$. La règle de multiplication, cf. Proposition 3.2.8, nous permet alors d'obtenir:

$$\begin{aligned} Z_t &= Y_t X_t \\ &= Y_0 K_0 + \int_0^t K_{s-} dY_s + \int_0^t Y_{s-} dK_s \\ &= 1 + \int_0^t Y_{s-} K_{s-} dX_s^c + \sum_{0 \leq s \leq t} Y_{s-} K_{s-} \Delta X_s \\ &= 1 + \int_0^t Y_{s-} K_{s-} dX_s = 1 + \int_0^t Z_{s-} dX_s \end{aligned}$$

ce qui est le résultat attendu. ■

La variable aléatoire Z est appelée l'exponentielle stochastique ou exponentielle de Doléans-Dade de X et est dénotée $Z = \mathcal{E}(X)$. Nous montrons à présent une propriété importante.

Proposition 3.4.3 *Soit $(X_t)_{t \geq 0}$ un processus mixte et une martingale. Alors, son exponentielle stochastique $Z_t = \mathcal{E}(X)_t$ est aussi une martingale.*

3.4.2 Cas du Processus de Wiener

Dans le cas du mouvement brownien, le changement de mesure est connu sous le nom de théorème de Girsanov. Il va permettre de modifier l'espérance du processus de Wiener.

Théorème 3.4.4 (Girsanov - MB) *Soit $(W_t)_{t \geq 0}$ un mouvement brownien standard défini sur $(\Omega, \mathcal{F}, \mathbb{P})$ et soit $(\mathcal{F}_t)_{t \geq 0}$, pour $t \in [0, T]$, une filtration associée à ce mouvement brownien. Soit $(\theta_t)_{t \geq 0}$ un processus adapté, nous définissons alors:*

$$Z_t = \exp \left(- \int_0^t \theta_u dW_u - \frac{1}{2} \int_0^t \theta_u^2 du \right), \quad Z_0 = 1 \quad (3.43)$$

$$\widetilde{W}_t = W_t + \int_0^t \theta_u du \quad (3.44)$$

Alors $\mathbb{E}[Z_T] = 1$ et sous $\widetilde{\mathbb{P}}$, donnée par 3.40, le processus $(\widetilde{W}_t)_{t \geq 0}$ est un brownien.

Preuve. Le théorème de caractérisation de Lévy, cf. Théorème 2.1.12, d'un mouvement brownien nous permet de dire qu'une martingale, commençant à la date 0 à la valeur 0, à trajectoires continues et de variation quadratique égale à t à chaque date t , est un mouvement brownien. Or $[\widetilde{W}, \widetilde{W}]_t = [W, W]_t = t$, ce qui conclut la preuve. ■

Le Théorème 3.4.4 peut être généralisé dans le cas où X est, sous \mathbb{P} , un processus d'Itô:

$$X_t = X_0 + \int_0^t a_s^{\mathbb{P}} ds + \int_0^t b_s dW_s \quad (3.45)$$

Proposition 3.4.5 (Girsanov - Itô) Soit $(X_t)_{t \geq 0}$ un processus d'Itô de la forme 3.45 défini sur $(\Omega, \mathcal{F}, \mathbb{P})$. Alors la dérivée de Radon-Nikodym 3.41 est donnée par:

$$Z_t = \exp \left(- \int_0^t \frac{a_s^{\mathbb{P}} - a_s^{\mathbb{Q}}}{b_s} dW_s - \frac{1}{2} \int_0^t \left(\frac{a_s^{\mathbb{P}} - a_s^{\mathbb{Q}}}{b_s} \right)^2 ds \right), \quad Z_0 = 1$$

avec $b_s^{\mathbb{P}} = b_s^{\mathbb{Q}} = b_s$. Alors, sous $\tilde{\mathbb{P}}$, le processus X_t satisfait:

$$X_t = X_0 + \int_0^t a_s^{\mathbb{Q}} ds + \int_0^t b_s d\tilde{W}_s$$

où \tilde{W}_t est un $(\mathcal{F}, \tilde{\mathbb{P}})$ -mouvement brownien.

Le théorème de Girsanov montre que la dérive et la volatilité jouent un rôle très différent dans la spécification d'un processus d'Itô. En effet, un changement de dérive équivaut à donner une nouvelle pondération aux probabilités d'occurrence de certaines trajectoires, tandis qu'un changement de volatilité produit un modèle totalement différent de celui d'origine: on dit alors que les modèles obtenus sont singuliers, car non comparables.

Remarque 3.4.6 De façon informelle, nous notons que Z donnée par 3.43 peut se réécrire:

$$Z_t = Z_0 - \int_0^t Z_u \theta_u W_u, \quad Z_t = \mathcal{E}(-\theta \cdot W)_t, \quad dZ_t = -Z_t \theta_t dW_t$$

3.4.3 Cas des Processus de Poisson

Le cas des processus de Poisson amène à traiter de façon séparée les versions simple et composé de ces processus, chacune produisant un changement de mesure différent.

Processus de Poisson

Dans le cas d'un processus de Poisson, le résultat suivant établit qu'un changement de mesure équivaut à un changement de l'intensité du processus.

Théorème 3.4.7 (Changement intensité - Poisson) Soit $(N_t)_{t \geq 0}$ un processus de Poisson d'intensité $\lambda^{\mathbb{P}} > 0$ défini sur $(\Omega, \mathcal{F}, \mathbb{P})$, $M_t = N_t - \lambda t$ sa version compensée et $\lambda^{\mathbb{Q}}$ une constante positive quelconque. Nous définissons alors:

$$Z_t = \exp \left((\lambda^{\mathbb{P}} - \lambda^{\mathbb{Q}}) t \right) \left(\frac{\lambda^{\mathbb{Q}}}{\lambda^{\mathbb{P}}} \right)^{N_t}, \quad Z_0 = 1 \quad (3.46)$$

Alors $\mathbb{E}[Z_T] = 1$ et sous $\tilde{\mathbb{P}}$, donnée par 3.40, le processus N_t est Poisson d'intensité $\lambda^{\mathbb{Q}}$.

Preuve. Nous calculons la fonction génératrice des moments de N_t sous $\tilde{\mathbb{P}}$, soit $\forall u \in \mathbb{R}$:

$$\begin{aligned}\tilde{\mathbb{E}}[e^{u \cdot N_t}] &= \mathbb{E}[e^{u \cdot N_t} Z_t] = \exp((\lambda^{\mathbb{P}} - \lambda^{\mathbb{Q}})t) \mathbb{E}\left[\exp(u \cdot N_t) \left(\frac{\lambda^{\mathbb{Q}}}{\lambda^{\mathbb{P}}}\right)^{N_t}\right] \\ &= \exp((\lambda^{\mathbb{P}} - \lambda^{\mathbb{Q}})t) \mathbb{E}\left[\exp\left(u + \log \frac{\lambda^{\mathbb{Q}}}{\lambda^{\mathbb{P}}}\right) N_t\right] \\ &= \exp((\lambda^{\mathbb{P}} - \lambda^{\mathbb{Q}})t) \exp\left(\lambda^{\mathbb{P}} t \left(\exp\left(u + \log \frac{\lambda^{\mathbb{Q}}}{\lambda^{\mathbb{P}}}\right) - 1\right)\right) \\ &= \exp(\lambda^{\mathbb{Q}} t (e^u - 1))\end{aligned}$$

qui est la fonction génératrice des moments d'une loi de Poisson d'intensité $\lambda^{\mathbb{Q}}$.

■

Le théorème précédent s'interprète alors de la sorte. Le changement d'intensité d'un processus de Poisson équivaut à donner une nouvelle pondération aux probabilités de réalisation de certaines trajectoires, mais celles-ci ne sont pas changées.

Remarque 3.4.8 *De façon informelle, nous notons que Z donnée par 3.46 peut se réécrire:*

$$Z_t = Z_0 - \int_0^t Z_{u-} \psi_u M_u, \quad Z_t = \mathcal{E}(-\psi \cdot M)_t, \quad dZ_t = -Z_t \psi_t dM_t$$

avec $\psi = \frac{\lambda^{\mathbb{P}} - \lambda^{\mathbb{Q}}}{\lambda^{\mathbb{P}}}$. Par ailleurs, le processus:

$$\tilde{M}_t = M_t + \lambda \psi t = N_t - \lambda(1 - \psi)t$$

est une $(\mathcal{F}, \tilde{\mathbb{P}})$ -martingale.

Processus de Poisson Composé

Dans le cas du processus de Poisson composé, le changement de mesure va permettre de changer l'intensité et l'amplitude des sauts.

Théorème 3.4.9 (Changement intensité et amplitude - Poisson) *Soit $(X_t)_{t \geq 0}$ un processus de Poisson composé défini sur $(\Omega, \mathcal{F}, \mathbb{P})$ et donné par:*

$$X_t = \sum_{i=0}^{N_t} Y_i, \quad t \geq 0 \quad (3.47)$$

où $(N_t)_{t \geq 0}$ est un processus de Poisson d'intensité $\lambda^{\mathbb{P}} > 0$ et les $(Y_i)_{i \geq 1}$ sont des variables aléatoires i.i.d. de densité $f^{\mathbb{P}}$, sous \mathbb{P} . Nous définissons alors:

$$Z_t = \exp((\lambda^{\mathbb{P}} - \lambda^{\mathbb{Q}})t) \prod_{i=0}^{N_t} \frac{\lambda^{\mathbb{Q}} f^{\mathbb{Q}}(Y_i)}{\lambda^{\mathbb{P}} f^{\mathbb{P}}(Y_i)}, \quad Z_0 = 1$$

Alors $\mathbb{E}[Z_T] = 1$ et sous $\tilde{\mathbb{P}}$, donnée par 3.40, le processus X_t est un processus de Poisson composé de la forme 3.47 avec N_t un processus de Poisson d'intensité $\lambda^{\mathbb{Q}}$ et les (Y_i) des variables aléatoires i.i.d. de densité $f^{\mathbb{Q}}$.

Idée. Nous devons montrer que, sous $\tilde{\mathbb{P}}$, la fonction génératrice des moments de X correspond à celle d'un processus de Poisson d'intensité λ et d'amplitude \tilde{f} . Nous devons ainsi montrer:

$$\tilde{\mathbb{E}}[e^{uX_t}] = \exp(\lambda^{\mathbb{Q}}t(\varphi^{\mathbb{Q}}(u) - 1)), \quad \forall u \in \mathbb{R} \quad (3.48)$$

où:

$$\varphi^{\mathbb{Q}}(u) = \int_{\mathbb{R}} e^{u \cdot y} f^{\mathbb{Q}}(y) dy$$

A cette fin, nous introduisons le processus:

$$A_t = \exp(uX_t - \lambda^{\mathbb{Q}}t(\varphi^{\mathbb{Q}}(u) - 1))$$

Alors 3.48 est équivalent à montrer que $A_t Z_t$ est une \mathbb{P} -martingale. La règle de multiplication, cf. Proposition 3.2.8, nous donne:

$$A_t Z_t = 1 + \int_0^t A_{s-} dZ_s + \int_0^t Z_{s-} dA_s + [A, Z]_t$$

Par suite, l'intégrale $\int_0^t A_{s-} dZ_s$ est une martingale du fait que l'intégrand soit continu à gauche et que Z soit une martingale. Par ailleurs, nous avons:

$$\begin{aligned} & \int_0^t Z_{s-} dA_s + [A, Z]_t \\ &= \int_0^t Z_{s-} dA_s^c + \sum_{0 \leq s \leq t} Z_{s-} \Delta A_s + [A, Z]_t \\ &= -\lambda^{\mathbb{Q}}(\varphi^{\mathbb{Q}}(u) - 1) \int_0^t A_{s-} Z_{s-} ds + \sum_{0 \leq s \leq t} A_{s-} Z_{s-} (e^{u\Delta X_s} - 1) \\ &+ \sum_{0 \leq s \leq t} A_{s-} Z_{s-} \Delta V_s - \sum_{0 \leq s \leq t} A_{s-} Z_{s-} \Delta H_s - \sum_{0 \leq s \leq t} A_{s-} Z_{s-} (e^{u\Delta X_s} - 1) \end{aligned}$$

où:

$$\Delta V_t = e^{u\Delta X_t} \frac{\lambda^{\mathbb{Q}} f^{\mathbb{Q}}(\Delta X_t)}{\lambda^{\mathbb{P}} f^{\mathbb{P}}(\Delta X_t)} = e^{u\Delta X_t} \Delta H_t$$

soit:

$$\int_0^t Z_{s-} dA_s + [A, Z]_t = \int_0^t A_{s-} Z_{s-} d(V_s - \lambda^{\mathbb{Q}}_s \varphi^{\mathbb{Q}}(u) ds) - \int_0^t A_{s-} Z_{s-} d(H_s - \lambda^{\mathbb{Q}}_s) \quad (3.49)$$

Comme les processus $V_s - \lambda^{\mathbb{Q}}_s \varphi^{\mathbb{Q}}(u) ds$ et $H_s - \lambda^{\mathbb{Q}}_s$ sont des martingales et que les intégrands sont continus à gauche, 3.49 est une martingale et par suite le processus $A_t Z_t$ est une martingale telle que $\mathbb{E}[A_t Z_t] = 1$, ce qui implique:

$$1 = \mathbb{E}[A_t Z_t] = \exp(-\lambda^{\mathbb{Q}}t(\varphi^{\mathbb{Q}}(u) - 1)) \mathbb{E}[e^{uX_t} Z_t]$$

et comme $\widetilde{\mathbb{E}} [e^{uX_t}] = \mathbb{E} [e^{uX_t} Z_t]$, nous retrouvons bien le résultat 3.48 attendu. ■

A l'instar du processus de Poisson, le changement de mesure pour un processus de Poisson composé permet une plus grande flexibilité. Il est en effet possible de changer à la fois l'intensité (aspect temps) et l'amplitude (aspect taille) des sauts par une modification de la mesure de sauts associée au processus de Poisson composé.

3.4.4 Cas des Modèles Mixtes

Nous pouvons à présent généraliser les résultats précédents au cas où X est un modèle mixte de la forme 3.42. Nous définissons alors les processus stochastiques suivants:

$$Z_t^W = \exp \left(- \int_0^t \frac{a_s^{\mathbb{P}} - a_s^{\mathbb{Q}}}{b_s} dW_s - \frac{1}{2} \int_0^t \left(\frac{a_s^{\mathbb{P}} - a_s^{\mathbb{Q}}}{b_s} \right)^2 ds \right)$$

$$Z_t^M = \exp \left((\lambda^{\mathbb{P}} - \lambda^{\mathbb{Q}}) t \right) \prod_{i=0}^{N_t} \frac{\lambda^{\mathbb{Q}} f^{\mathbb{Q}}(Y_i)}{\lambda^{\mathbb{P}} f^{\mathbb{P}}(Y_i)}$$

et:

$$Z_t = Z_t^W Z_t^M \tag{3.50}$$

Nous avons alors le théorème suivant de changement de mesure.

Théorème 3.4.10 *Sous $\widetilde{\mathbb{P}}$, donnée par 3.40 et 3.50, nous avons:*

1. $\widetilde{W}_t = W_t + \int_0^t b_s^{-1} (a_s^{\mathbb{P}} - a_s^{\mathbb{Q}}) ds$ est un $(\mathcal{F}, \widetilde{\mathbb{P}})$ -mouvement brownien,
2. N_t est un processus de Poisson Poisson (composé) d'intensité $\lambda^{\mathbb{Q}}$,
3. $(Y_i)_{i \geq 1}$ sont des variables aléatoires i.i.d. de densité $f^{\mathbb{Q}}$.

B

Evaluation Risque-Neutre

Dans cette Annexe, nous explorons les concepts fondamentaux en finance d'évaluation risque-neutre (*risk-neutral pricing*), d'absence d'arbitrage (*absence of arbitrage*) et leur lien avec les mesures martingales équivalentes (*equivalent martingale measures*).

B.1. Règles d'Evaluation

Soit un marché financier dont les évolutions possibles sur l'intervalle $[0, T]$ sont décrites par l'espace des scénarios (Ω, \mathcal{F}) : \mathcal{F} représente alors toutes les évolutions possibles des prix entre les dates 0 et T . Supposons en outre que les actifs sous-jacents puissent être décrits par un processus càdlàg non-anticipatif de la forme:

$$\begin{aligned} S : [0, T] \times \Omega &\mapsto \mathbb{R}^{d+1} \\ (t, \omega) &\mapsto (S_t^0(\omega), \dots, S_t^d(\omega)) \end{aligned}$$

où $S_t^i(\omega)$ représente la valeur de l'actif i à la date t pour le scénario ω et où S_t^0 est le *numéraire*, c'est-à-dire l'actif de référence utilisé pour réaliser les conversions monétaires entre tous les actifs présents sur le marché: c'est un facteur de normalisation. En pratique, le numéraire est pris comme étant le facteur d'actualisation, soit:

$$S_t^0 = \exp(rt) \text{ ou } S_t^0 = \exp\left(\int_0^t r_s ds\right)$$

Remarque B.1.1 (Numéraire) *Le numéraire est l'unité monétaire dans laquelle des actifs financiers différents sont dénommés. Habituellement, le numéraire représente une monnaie nationale et il est possible de changer de numéraire en passant d'une monnaie à une autre. L'utilisation du numéraire trouve alors tout son intérêt dans le cas de modèles de taux de change ou de modèles à taux d'intérêt stochastique.*

Actif Conditionnel

Un actif conditionnel (ou contingent) de maturité $T > 0$ peut être représenté en spécifiant son payoff terminal $H(\omega)$ pour le scénario ω . Comme H est uniquement révélé en T , c'est une variable aléatoire \mathcal{F}_T -mesurable, $H : \Omega \rightarrow \mathbb{R}$. Un

Type	Call	Put
<i>vanilla</i>	$H^c = \max(S_T^i - K, 0)$	$H^c = \max(K - S_T^i, 0)$
<i>barrier</i>	—	—
<i>lookback</i>	$H^c = \max(S_T^i - \min_{t \in [0, T]} S_t^i, 0)$	$H^p = \min(\max_{t \in [0, T]} S_t^i - S_T^i, 0)$
<i>asian</i>	$H^c = \max\left(\sum_{t \in [0, T]} S_t^i - K, 0\right)$	$H^c = \max\left(K - \sum_{t \in [0, T]} S_t^i, 0\right)$

Table B.1: Payoffs d'options usuelles.

exemple typique de payoff est donné par une option européenne. Ces options sont parfois appelées *vanilles* (*vanilla*) car leur fonction payoff ne dépend que de la valeur terminale de l'actif sous-jacent. Les options dont le payoff dépend de la trajectoire suivie par l'actif sous-jacent sont dénommées dépendantes de la trajectoire (*path-dependent*) ou exotiques (*exotic*). Les plus connues sont les options barrières (*barrier*), les options (*lookback*) et les options asiatiques (*Asian*). Nous donnons dans le Tableau B.1 la forme de leur payoff.

Remarque B.1.2 (Options européennes, américaines et bermudéennes)

Les options européennes (European) spécifient dans leur contrat une date d'échéance à laquelle l'option doit être exercée. A contrario, une option dont le détenteur peut choisir de l'exercer à n'importe quelle date, y compris la date d'exercice, est dénommée américaine (American). Une option intermédiaire entre les options européenne et américaine et l'option bermudéenne (Bermudan): elle autorise un exercice anticipé, mais seulement à certaines dates définies de façon contractuelle.

Règle d'évaluation

Le problème que nous rencontrons à présent est celui de l'évaluation. La question que nous nous posons est: comment attribuer une notion de valeur à un actif contingent H ? Une règle d'évaluation (*pricing rule*) ou opérateur d'évaluation est une procédure qui attribue à chaque actif contingent H une valeur $\Pi_t(H)$ à chaque date t avant la date de maturité. Pour être une règle d'évaluation, $\Pi_t(H)$ doit vérifier certaines propriétés:

1. (*Mesurabilité*) Le processus $\Pi_t(H)$ doit être calculé avec l'information disponible à la date t , i.e.: $\Pi_t(H)$ est \mathcal{F}_t -mesurable.
2. (*Positivité*). Un actif de payoff positif doit naturellement avoir une valeur positive:

$$H(\omega) \geq 0, \forall \omega \in \Omega \Rightarrow \Pi_t(H) \geq 0, \forall t \in [0, T]$$

3. (*Linéarité*). La valeur d'un portefeuille d'actifs contingents est donnée par la somme des valeurs de chacun de ces actifs:

$$\Pi_t\left(\sum_{i \in I} H_i\right) = \sum_{i \in I} \Pi_t(H_i) \quad (\text{B.1})$$

Lotterie

Pour tout évènement $A \in \mathcal{F}$, la variable aléatoire $\mathbf{1}_A$ représente le payoff d'un actif contingent qui paye 1 à la date $T > 0$ si A se réalise et 0 sinon: on parle de lotterie. La valeur $\Pi_t(1)$ représente la valeur aujourd'hui d'une unité monétaire de nominal 1 que l'on reçoit en T , c'est-à-dire le facteur d'escompte du temps:

$$\Pi_t(1) = e^{-r(T-t)}$$

Nous définissons à présent la quantité $\mathbb{Q} : \mathcal{F} \rightarrow \mathbb{R}$ par:

$$\mathbb{Q}(A) = \frac{\Pi_0(\mathbf{1}_A)}{\Pi_0(1)} = e^{rT} \Pi_0(\mathbf{1}_A) \quad (\text{B.2})$$

$\mathbb{Q}(A)$ est alors la valeur d'une lotterie de nominal $\exp(rT)$ sur l'évènement A . Mais c'est aussi plus que ça. Par la positivité et linéarité de l'opérateur Π , nous avons:

1. $0 \leq \mathbb{Q}(A) \leq 1$ du fait que $0 \leq \mathbf{1}_A \leq 1$.
2. Si A et B sont disjoints, alors $\mathbb{Q}(A \cup B) = \mathbb{Q}(A) + \mathbb{Q}(B)$.

Si nous étendons à présent la propriété de linéarité B.1 à des sommes infinies, alors $\mathbb{Q}(A)$ définit une mesure de probabilité, cf. Définition 1.1.4, sur l'espace des scénarios (Ω, \mathcal{F}) . Ainsi, à partir d'une règle d'évaluation H , nous avons construit une mesure de probabilité \mathbb{Q} pour l'ensemble des scénarios. L'inverse est aussi possible. Si nous posons $H = \sum_i c_i \mathbf{1}_{A_i}$, la linéarité de Π implique $\Pi_0(H) = \mathbb{E}^{\mathbb{Q}}[H]$. Nous avons alors la relation suivante:

$$\Pi_0(H) = e^{-rT} \mathbb{E}^{\mathbb{Q}}[H] \quad (\text{B.3})$$

De fait, il existe une correspondance univoque entre les règles d'évaluation linéaires Π vérifiant les propriétés sus-citées et les mesures de probabilités \mathbb{Q} définies sur les scénarios (Ω, \mathcal{F}) . Cette correspondance est donnée par la relation:

$$\Pi_0(H) = e^{-r(T-t)} \mathbb{E}^{\mathbb{Q}}[H] \text{ et } \mathbb{Q}(A) = e^{r(T-t)} \Pi_0(\mathbf{1}_A)$$

La relation B.3 est parfois appelée formule d'évaluation risque-neutre (*risk-neutral pricing formula*): la valeur d'un actif contingent est donnée par la valeur actualisée de son espérance sous \mathbb{Q} . Nous faisons à présent trois remarques importantes.

Remarque B.1.3 (Linéarité ou non-linéarité de H) *La dérivation précédente repose de façon importante sur la propriété de linéarité de H . Si H vérifie une condition de non-linéarité, il n'est plus possible de dériver la même probabilité \mathbb{Q} .*

Remarque B.1.4 (Nature de \mathbb{Q}) *Il est important de noter que la mesure \mathbb{Q} n'a rien à voir avec la mesure objective/statistique de probabilité \mathbb{P} qui décrit les probabilités d'occurrence des scénarios sur (Ω, \mathcal{F}) . Nous n'avons d'ailleurs jamais défini clairement \mathbb{P} . La mesure \mathbb{Q} est appelée une mesure d'évaluation*

(pricing measure) ou une mesure risque-neutre (risk-neutral measure). Elle ne doit pas être interprétée en termes de scénarios: on ne peut pas ainsi dire que $\mathbb{Q}(A)$ représente la probabilité d'occurrence de A dans le monde financier, mais doit être analysée en termes d'une lotterie qui aurait pour support l'évènement A .

Remarque B.1.5 (Nature économétrique de \mathbb{P} et \mathbb{Q}) D'un point de vue économétrique, les différences entre les mesures respectivement objective \mathbb{P} et risque-neutre \mathbb{Q} sont apparentes. La mesure \mathbb{P} peut être obtenue par une analyse des séries temporelles des actifs sous-jacents, tandis qu'une inférence menée sur \mathbb{Q} ne prend en compte que les prix des actifs contingents disponibles à la date $t = 0$.

Finalement, il est possible de modifier l'expression B.3 pour rendre la règle d'évaluation Π consistente en temps, c'est-à-dire \mathcal{F}_t -mesurable, soit:

$$\Pi_t(H) = e^{-r(T-t)} \mathbb{E}^{\mathbb{Q}}[H | \mathcal{F}_t] \quad (\text{B.4})$$

B.2. Mesures Martingales Equivalentes

Jusqu'à présent, nous n'avons supposé aucune hypothèse sur la distribution des actifs sous-jacents. Supposons désormais que nous ayons une connaissance sur les probabilités d'occurrence des scénarios sur (Ω, \mathcal{F}) , représentée par une mesure de probabilité \mathbb{P} . Il s'agit soit d'une mesure objective/statistique sur les scénarios ou d'une mesure subjective que le modélisateur se donne a priori. Nous nous interrogeons maintenant sur les conditions que doit vérifier la règle d'évaluation et partant la mesure \mathbb{Q} définie par B.2 et B.4 pour être compatible avec la visions statistique de l'évolution du marché donnée par \mathbb{P} .

Définition de la MME

Pour cela, considérons un marché financier où un agent évalue la valeur des actifs conditionnels selon la règle d'évaluation B.4 pour une mesure de probabilité \mathbb{Q} . Considérons de plus un évènement A tel que $\mathbb{P}(A) > 0$ où \mathbb{P} représente la vision subjective de cet agent des probabilités d'occurrence des scénarios sur (Ω, \mathcal{F}) et une option qui paie une unité monétaire si l'évènement A se réalise. Comme cet évènement est considéré par l'agent impossible ($\mathbb{P}(A) = 0$), il lui accorde une valeur monétaire nulle. Toutefois, la règle d'évaluation B.3 affecte une valeur à cette option:

$$\Pi_0(\mathbf{1}_A) = e^{-rT} \mathbb{E}^{\mathbb{Q}}[\mathbf{1}_A] = e^{-rT} \mathbb{Q}(A)$$

De fait, on dira que la mesure de probabilité \mathbb{Q} est cohérente avec la vision du monde de l'agent si $\mathbb{Q}(A) = 0$. Inversement, si $\mathbb{Q}(A) > 0$, alors l'option de payoff $\mathbf{1}_A$ est de valeur nulle et si $\mathbb{P}(A) > 0$ l'acquisition de cette option pour un prix nul conduirait à un gain certain, c'est-à-dire un arbitrage. Ce qui n'est pas possible et donc nécessairement $\mathbb{P}(A) = 0$ aussi. Dès lors, la compatibilité de la règle d'évaluation \mathbb{Q} avec le modèle stochastique \mathbb{P} nécessite que les mesures

\mathbb{P} et \mathbb{Q} soient équivalentes, c'est-à-dire qu'elles définissent les mêmes ensembles des possibles sur (Ω, \mathcal{F}) , ce qui se traduit par la condition suivante:

$$\mathbb{P} \sim \mathbb{Q} \Leftrightarrow [\mathbb{Q}(A) = 0 \Leftrightarrow \mathbb{P}(A) = 0, \forall A \in \mathcal{F}] \quad (\text{B.5})$$

Propriété de Martingale

Considérons à nouveau l'actif financier S^i . Deux stratégies (équivalentes) sont possibles pour un agent :

1. détenir cet actif jusqu'à la date T , le vendre à cette date et récupérer la valeur terminale pour un payoff S_T^i , soit une richesse finale $e^{r(T-t)} S_t^i$.
2. vendre cet actif à la date $t < T$, la somme résultante étant investie jusqu'en T au taux d'intérêt sans risque r , l'agent récupère alors une richesse finale égale à $e^{r(T-t)} S_t^i$.

Ces deux stratégies ont la même valeur à la date T , elles doivent donc être aussi égales pour toute date t , ce que nous résumons sous la forme:

$$\mathbb{E}^{\mathbb{Q}} [S_T^i | \mathcal{F}_t] = \mathbb{E}^{\mathbb{Q}} [e^{r(T-t)} S_t^i | \mathcal{F}_t] = e^{r(T-t)} S_t^i$$

Soit en divisant par le numéraire $S_T^0 = e^{rT}$, nous avons:

$$\mathbb{E}^{\mathbb{Q}} \left[\frac{S_T^i}{S_T^0} | \mathcal{F}_t \right] = \frac{S_t^i}{S_t^0} \text{ i.e.: } \mathbb{E}^{\mathbb{Q}} [\tilde{S}_T^i | \mathcal{F}_t] = \tilde{S}_t^i \quad (\text{B.6})$$

Dès lors, l'actif actualisé $\tilde{S}_t^i = e^{-rt} S_t^i$ est une martingale par rapport à la mesure de probabilité \mathbb{Q} . Nous avons alors la définition suivante d'une mesure martingale équivalente (MME) ou mesure risque-neutre.

Définition B.2.1 (MME) Sur $(\Omega, \mathcal{F}, \mathbb{P})$, la mesure de probabilité \mathbb{Q} est une MME si:

1. les mesures \mathbb{P} et \mathbb{Q} sont équivalentes, i.e.: elles vérifient B.5.
2. sous \mathbb{Q} , l'actif actualisé $e^{-rt} S_t^i$ est une martingale, i.e.: il vérifie B.6.

4

Equations aux Dérivées Partielles

Le cadre d'évaluation des actifs financiers peut s'envisager selon deux approches. La première, dite par simulation, utilise des simulations de Monte-Carlo pour approcher la valeur espérée des actifs financiers. La seconde, dite numérique, résout une équation aux dérivées partielles. Nous abordons dans ce Chapitre cette seconde approche et montrons comment elle peut être reliée à l'approche *sous espérance* dite aussi risque-neutre, cf. Annexe B. Ceci sera par ailleurs l'occasion d'introduire un équipement central dans l'étude des processus stochastiques, les équations différentielles stochastiques.

4.1. Equations Différentielles Stochastiques

4.1.1 Cas Brownien

Dans le cas classique où l'incertitude est portée par un mouvement brownien $(W_u)_{u \geq 0}$, une équation différentielle stochastique (par la suite EDS) est une équation de la forme:

$$dX_u = a(u, X_u) du + b(u, X_u) dW_u \quad (4.1)$$

où $a(u, X_u)$ et $b(u, X_u)$ sont des fonctions données appelées respectivement la *tendance* et la *volatilité*. Par ailleurs, une condition initiale de la forme $X_t = x$ où $t \geq 0$ et $x \in \mathbb{R}$ est spécifiée. Nous cherchons à présent à expliciter le processus X_T , pour $T \geq t$ tel que:

$$X_T = X_t + \int_t^T a(u, X_u) du + \int_t^T b(u, X_u) dW_u, \quad X_t = x \quad (4.2)$$

Il est possible de montrer que sous des hypothèses peu restrictives sur les fonctions $a(u, X_u)$ et $b(u, X_u)$, il existe une solution unique au processus X_T . Toutefois, ce processus peut être difficile (voire impossible) à déterminer de façon explicite du fait qu'il apparaît des deux côtés de l'équation définissant X_T , cf. 4.2.

Remarque 4.1.1 (Mesurabilité de la solution) *Nous noterons par ailleurs que la solution X_T à la date T est bien \mathcal{F}_T -mesurable. Cela étant dû au fait que comme $X_t = x$ est une constante, tout ce qui est nécessaire pour déterminer X_T est la connaissance de la trajectoire suivie par le mouvement brownien W entre les dates t et T .*

Bien que les EDS sont en général difficiles à résoudre dans le cas général (multivarié), une classe d'EDS univarié peut être résolue de façon explicite. Soit l'équation différentielle:

$$dX_u = (\alpha_u + \beta_u X_u) du + (\gamma_u + \sigma_u X_u) dW_u$$

où $\alpha_u, \beta_u, \gamma_u$ et σ_u sont des fonctions qui ne dépendent pas du temps. Nous présentons trois exemples classiques issus de cette famille d'EDS.

Exemple 4.1.2 (Mouvement brownien géométrique) *L'équation satisfaite par un MBG est donnée par:*

$$dS_u = \alpha S_u du + \sigma S_u dW_u$$

soit, dans les notations de 4.1, $a(u, x) = \alpha x$ et $b(u, x) = \sigma x$. Cette équation admet pour une condition initiale $S_0 > 0$, la solution:

$$S_t = S_0 \exp \left(\left(\alpha - \frac{1}{2} \sigma^2 \right) t + \sigma W_t \right)$$

De même pour $T \geq t$:

$$S_T = S_0 \exp \left(\left(\alpha - \frac{1}{2} \sigma^2 \right) T + \sigma W_T \right)$$

Soit:

$$\frac{S_T}{S_t} = \exp \left(\left(\alpha - \frac{1}{2} \sigma^2 \right) (T - t) + \sigma (W_T - W_t) \right)$$

Si finalement, la condition initiale est donnée en date t et non 0 et que $S_t = x$, alors:

$$S_T = x \exp \left(\left(\alpha - \frac{1}{2} \sigma^2 \right) (T - t) + \sigma (W_T - W_t) \right)$$

et l'on retrouve bien le fait que S_T ne dépend que de la trajectoire de W entre t et T .

Exemple 4.1.3 (Modèle de Hull-White) *Soit l'EDS:*

$$dR_u = (a_u - b_u R_u) du + \sigma_u dW_u \quad (4.3)$$

soit, dans les notations de 4.1, $a(u, x) = a_u - b_u x$ et $b(u, x) = \sigma$ pour une condition initiale $R_t = r$. En notant que:

$$d \left(e^{\int_0^u b_s ds} R_u \right) = e^{\int_0^u b_s ds} (a_u du + \sigma_u dW_u) \quad (4.4)$$

et par intégrant des deux côtés de 4.4 sur $[t, T]$, il vient:

$$R_T = r e^{-\int_0^t b_s ds} + \int_t^T e^{-\int_0^u b_s ds} a_u du + \int_t^T e^{-\int_0^u b_s ds} \sigma_u dW_u$$

ce qui est une formule explicite pour R_T .

Exemple 4.1.4 (Modèle de Cox-Ingersoll-Ross) Soit l'EDS:

$$dR_u = (a - bR_u) du + \sigma\sqrt{R_u}dW_u \quad (4.5)$$

Cette équation n'admet pas de solution explicite, même si nous savons que, pour une condition initiale $R_t = r$ donnée, elle admet une solution unique. Sa solution peut être approchée en ayant recours à des schémas de discrétisation et de simulation, cf. Annexe C.

Remarque 4.1.5 (Modélisation des taux d'intérêt) Il est possible de montrer que le modèle 4.3 produit des trajectoires qui peuvent être négatives, tandis que le modèle 4.5 n'en produit jamais. En effet, quand R_u approche 0, le terme de volatilité $\sigma\sqrt{R_u}dW_u$ approche lui aussi 0, dès lors R_u est uniquement guidé par $a > 0$. Ceci sera d'intérêt si R_u représente la valeur d'un taux d'intérêt, qui n'est bien évidemment jamais négatif.

4.1.2 Cas Mixte Brownien-Poisson

Dans le cas où l'incertitude est portée par un mouvement brownien $(W_u)_{u \geq 0}$ et un processus de sauts pur (*pure jump process*) $(J_u)_{u \geq 0}$, une équation différentielle stochastique à sauts (par la suite EDS-S) est de la forme:

$$dX_u = a(u, X_u) du + b(u, X_u) dW_u + c(u, X_u) dJ_u \quad (4.6)$$

où les fonctions $a(u, X_u)$, $b(u, X_u)$ et $c(u, X_u)$ sont données et respectivement nommées la tendance et la volatilité de la partie martingale et l'amplitude de la partie discontinue. De façon plus précise, le processus de sauts pur admet la représentation:

$$J_u = \sum_{i=1}^{N_u} Y_i$$

où $(N_u)_{u \geq 0}$ est un processus de comptage d'intensité $\lambda > 0$ et les variables aléatoires $(Y_i)_{i \geq 1}$ sont i.i.d. de distribution f . La mesure de sauts associée à ce processus s'écrit $\mu(dx, dt) = \nu(dx) dt = \lambda f(x) dt$. Par ailleurs, une condition initiale de la forme $X_t = x \in \mathbb{R}$ pour $t \geq 0$ est donnée. Nous cherchons à présent à expliciter le processus X_T pour $T \geq t$ tel que:

$$X_T = X_t + \int_t^T a(u, X_u) du + \int_t^T b(u, X_u) dW_u + \sum_{t \leq u \leq T} c(u, X_u) \Delta J_u, \quad X_t = x \quad (4.7)$$

Dit autrement et pour simplifier la présentation, les processus mixtes que nous distinguons par la suite sont continus à droite et de la forme:

$$X_t = X_0 + R_t + I_t + J_t \quad (4.8)$$

où X_0 est une condition initiale non aléatoire. Le processus:

$$R_t = \int_0^t a_s ds$$

est une intégrale au sens de Riemann pour un processus adapté a_s quelconque. Le processus:

$$I_t = \int_0^t b_s dW_s$$

est un intégrale au sens d'Itô pour un processus adapté b_s . Par ailleurs, la partie continue du processus X_t se définit comme une EDS de la forme 4.1, soit:

$$X_t^c = X_0 + R_t + I_t = X_0 + \int_0^t a_s ds + \int_0^t b_s dW_s$$

Finalement, le processus J_t est un processus de sauts pur, adapté, continu à droite avec $J_0 = 0$. C'est-à-dire que $J_t = \lim_{s \downarrow t} J_s$, soit si J a un saut à la date t , alors J_t est la valeur de J immédiatement après le saut, tandis que J_{t-} est sa valeur juste avant le saut. Nous supposons finalement que J_t ne saute pas à la date initiale, a seulement un nombre fini de sauts sur chaque intervalle ferme $(0, T]$ et qu'il est constant entre deux instants de sauts consécutifs. Les processus de Poisson simple et composé ont cette propriété, mais pas un processus de Poisson compensé. Par suite, comme X_t est continu à droite et adapté et les processus R_t et I_t sont continus, la version continue à gauche de X_t est:

$$X_{t-} = X_0 + R_t + I_t + J_{t-}$$

et le saut de X à la date t est noté:

$$\Delta X_t = X_t - X_{t-}$$

Exemple 4.1.6 (Modèle à sauts) *Un modèle élémentaire comportant des sauts s'écrit:*

$$dS_t = \alpha S_t dt + \sigma S_{t-} dM_t$$

où $(M_t)_{t \geq 0}$ est la version compensée d'un processus de Poisson N_t d'intensité $\lambda > 0$. En intégrant cette équation, on obtient:

$$\begin{aligned} S_t &= S_0 \exp(\alpha t + N_t \log(1 + \sigma) - \lambda \sigma t) \\ &= S_0 e^{(\alpha - \lambda \sigma)t} (1 + \sigma)^{N_t} \end{aligned}$$

Exemple 4.1.7 (Modèle mixte) *Soit l'EDS-S:*

$$dS_t = \alpha S_t dt + \sigma S_t dW_t + S_{t-} (Q_t - \beta \lambda t) \quad (4.9)$$

où $(W_t)_{t \geq 0}$ est un mouvement brownien, $(N_t)_{t \geq 0}$ un processus de Poisson d'intensité $\lambda > 0$ et on définit le processus de Poisson composé:

$$Q_t = \sum_{i=1}^{N_t} Y_i$$

où les Y_i sont des variables aléatoires i.i.d et $\beta = \mathbb{E}[Y_i]$. Nous voulons montrer que:

$$S_t = S_0 \exp\left(\left(\alpha - \beta \lambda - \frac{1}{2} \sigma^2\right)t + \sigma W_t\right) \prod_{i=1}^{N_t} (1 + Y_i)$$

A cette fin, nous définissons le processus continu:

$$X_t = S_0 \exp \left(\left(\alpha - \beta\lambda - \frac{1}{2}\sigma^2 \right) t + \sigma W_t \right)$$

et le processus de sauts pur:

$$J_t = \prod_{i=1}^{N_t} (1 + Y_i)$$

Alors, $S_t = X_t J_t$ et nous montrons que cette équation est solution de 4.9. La formule d'Itô pour un processus continu donne:

$$dX_t = (\alpha - \beta\lambda) X_t dt + \sigma X_t dW_t$$

Et à la date de survenue du i -ème saut, nous avons:

$$\Delta J_t = J_t - J_{t-} = J_{t-} Y_i = J_{t-} \Delta Q_t$$

Soit:

$$dJ_t = J_{t-} dQ_t$$

La règle de multiplication donne alors:

$$S_t = X_t J_t = S_0 + \int_0^t X_{s-} dJ_s + \int_0^t J_s dX_s + [X, J]_t$$

Comme J est un processus de sauts pur et X est continu, $[X, J]_t = 0$. Nous obtenons alors:

$$S_t = X_t J_t = S_0 + \int_0^t X_{s-} J_{s-} dQ_s + (\alpha - \beta\lambda) \int_0^t J_s X_s ds + \sigma \int_0^t J_s X_s dW_s$$

Soit sous forme différentielle:

$$\begin{aligned} dS_t &= d(X_t J_t) = X_{t-} J_{t-} dQ_t + (\alpha - \beta\lambda) J_t X_t dt + \sigma J_t X_t dW_t \\ &= (\alpha - \beta\lambda) S_t dt + \sigma S_t dW_t + S_{t-} dQ_t \end{aligned}$$

4.2. Propriété de Markov

Considérons à nouveau l'EDS 4.1 ou l'EDS-S 4.7. Soit $t \in [0, T]$ et $h(y)$ une fonction borélienne. On note:

$$g(t, x) = \mathbb{E}^{t,x} [h(X_T)] \quad (4.10)$$

l'espérance de $h(X_T)$ où X_T est la solution de 4.1 ou 4.7 pour une condition initiale $X_t = x$. Ainsi, X_T est différente suivant la date et la valeur de la condition initiale, ce que traduit la dépendance en t et x dans la notation $\mathbb{E}^{t,x}$ de 4.10.

Calcul de $g(t, x)$

Dans les cas où nous disposons d'une solution explicite pour X_T , cf. Exemples 4.1.2 et 4.1.3, sa distribution de probabilité est elle-aussi complètement spécifiée et il est possible de calculer $g(t, x)$ de façon directe. A contrario, lorsque X_T n'est pas solvable, cf. Exemple 4.1.4, il faut recourir à des méthodes numériques pour approcher la valeur de $g(t, x)$. Une approche possible est d'employer une discrétisation en temps de X_u , par exemple par un schéma d'Euler, cf. Annexe C, de simuler un grand nombre de trajectoires de ce processus, puis de prendre la moyenne de toutes les réalisations de $h(X_T)$: il s'agit du principe de la méthode de Monte-Carlo.

Nous énonçons à présent le résultat important de cette Section.

Théorème 4.2.1 *Soit $X_u, u \geq 0$, la solution d'une EDS de la forme 4.1 ou une EDS-S donnée par 4.7 avec une condition initiale donnée à la date 0. Alors, pour $t \in [0, T]$:*

$$\mathbb{E}[h(X_T) | \mathcal{F}_t] = g(t, X_t)$$

Ce théorème suggère deux observations. La première est que basée sur l'information \mathcal{F}_t , le calcul de l'espérance conditionnelle du processus X_T revient à calculer son espérance en T lorsque le processus a commencé en t , soit $g(t, x) = \mathbb{E}^{t,x}[h(X_T)]$ où $X_t = x$. La seconde est que bien que le processus X_u soit observée entre les dates $u \in [0, t]$, la seule information d'intérêt pour calculer $\mathbb{E}[h(X_T) | \mathcal{F}_t]$ est la valeur de X_t . Ceci signifie alors que X_u est un processus de Markov, cf. Définition 1.3.17.

Corollaire 4.2.2 *Les solutions d'EDS ou d'EDS-S sont des processus markoviens.*

4.3. Equations aux Dérivées Partielles

La résolution de $g(t, x)$ dans 4.10 peut se faire en ayant recours à une méthode de simulation dite de Monte-Carlo. Toutefois, cette stratégie converge lentement et ne donne la solution que pour un couple (t, x) donné. Les équations aux dérivées partielles vont permettre d'exhiber une relation entre $g(t, x)$ et certaines de ses dérivées, à même de proposer une solution numérique à la fois rapide (dans le cas univarié) et efficace dans le sens où il sera possible d'évaluer $g(t, x)$ pour toutes les valeurs de (t, x) possibles de façon simultanée. Le théorème de Feynman-Kac établit un lien entre les équations différentielles stochastiques et les équations aux dérivées partielles.

4.3.1 Cas Brownien

Nous énonçons le théorème de Feynman-Kac dans le cas d'une EDS de la forme 4.1.

Théorème 4.3.1 (Feynman-Kac) *Soit l'EDS:*

$$dX_u = a(u, X_u) du + b(u, X_u) dW_u \tag{4.11}$$

Soit $h(y)$ une fonction borélienne quelconque. Pour $T > 0$ et $t \in [0, T]$, nous définissons:

$$g(t, x) = \mathbb{E}^{t,x} [h(X_T)] \quad (4.12)$$

Alors $g(t, x)$ est dite satisfaire une équation aux dérivées partielles (EDP) donnée par:

$$g_t(t, x) + a(t, x)g_x(t, x) + \frac{1}{2}b^2(t, x)g_{xx}(t, x) = 0 \quad (4.13)$$

pour la condition terminale:

$$g(T, x) = h(x), \quad \forall x$$

La preuve du théorème de Feynman-Kac dépend du lemme suivant.

Lemme 4.3.2 Soit $(X_u)_{u \geq 0}$ la solution d'une EDS de la forme 4.11 et $g(t, x)$ donné par 4.12. Alors le processus stochastique $g(t, X_t)$, $t \in [0, T]$ est une martingale.

Preuve. Pour $0 \leq s \leq t \leq T$ donné, le Théorème 4.2.1 implique:

$$\mathbb{E}[h(X_T) | \mathcal{F}_s] = g(s, X_s)$$

$$\mathbb{E}[h(X_T) | \mathcal{F}_t] = g(t, X_t)$$

Soit:

$$\begin{aligned} \mathbb{E}[g(t, X_t) | \mathcal{F}_t] &= \mathbb{E}[\mathbb{E}[h(X_T) | \mathcal{F}_t] | \mathcal{F}_s] \\ &= \mathbb{E}[h(X_T) | \mathcal{F}_s] = g(s, X_s) \end{aligned}$$

en conditionnant de façon itérée. ■

Feynman-Kac. Soit X_t la solution de l'EDS 4.11 démarrant en 0. Comme $g(t, X_t)$ est une martingale, le terme en dt dans l'équation $dg(t, X_t)$ doit être nul; sinon $g(t, X_t)$ aurait une tendance soit à augmenter, soit à baisser s'il était respectivement positif ou négatif. Nous avons:

$$\begin{aligned} dg(t, X_t) &= g_t dt + g_x dX + \frac{1}{2}g_{xx} dX dX \\ &= g_t dt + ag_x dt + bg_x dW + \frac{1}{2}b^2 g_{xx} dt \\ &= \left(g_t + bg_x + \frac{1}{2}b^2 g_{xx} \right) dt + bg_x dW \end{aligned}$$

Soit en annulant le terme en dt , il vient:

$$g_t(t, x) + a(t, x)g_x(t, x) + \frac{1}{2}b^2(t, x)g_{xx}(t, x) = 0$$

en tout point (t, x) qui peut être atteint par (t, X_t) . ■

Nous illustrons ce théorème par un résultat complémentaire, voire connexe.

Théorème 4.3.3 (Feynman-Kac actualisé) Soit $(X_u)_{u \geq 0}$ la solution d'une EDS de la forme 4.11 et $h(y)$ une fonction borélienne quelconque. Pour $T > 0$ et $t \in [0, T]$ donnés, nous définissons:

$$f(t, x) = \mathbb{E}^{t, x} \left[e^{-r(T-t)} h(X_T) \right] \quad (4.14)$$

Alors $f(t, x)$ satisfait l'EDP suivante:

$$f_t(t, x) + a(t, x) f_x(t, x) + \frac{1}{2} b^2(t, x) f_{xx}(t, x) = r f(t, x) \quad (4.15)$$

pour la condition terminale:

$$f(T, x) = h(x), \quad \forall x$$

Preuve. Soit:

$$f(t, X_t) = \mathbb{E} \left[e^{-r(T-t)} h(X_T) \mid \mathcal{F}_t \right]$$

Si nous souhaitons reproduire la preuve du théorème de Feynman-Kac, il faudrait montrer que $f(t, X_t)$ est une martingale, mais du fait de la présence de la variable de conditionnement t dans le facteur d'actualisation, $f(t, X_t)$ ne possède pas la propriété de martingale:

$$\mathbb{E} [f(t, X_t) \mid \mathcal{F}_s] = \mathbb{E} \left[e^{-r(T-t)} h(X_T) \mid \mathcal{F}_s \right] \neq f(s, X_s) = \mathbb{E} \left[e^{-r(T-s)} h(X_T) \mid \mathcal{F}_s \right]$$

En notant que:

$$e^{-rt} f(t, X_t) = \mathbb{E} \left[e^{-rT} h(X_T) \mid \mathcal{F}_t \right]$$

il est à présent possible d'appliquer la règle de conditionnement itéré pour montrer que $e^{-rt} f(t, X_t)$ est une martingale. A cette fin, on note:

$$d(e^{-rt} f(t, X_t)) = e^{-rt} \left(-rf + f_t + af_x + \frac{1}{2} b^2 f_{xx} \right) dt + e^{-rt} b f_x dW$$

où en annulant le terme en dt , il vient 4.15. ■

Exemple 4.3.4 (Option sur MBG) Soit $h(S_T)$ la fonction payoff d'une option expirant à la date T et écrite sur un sous-jacent $(S_u)_{u \geq 0}$ dont la dynamique satisfait:

$$\frac{dS_u}{S_u} = r du + \sigma dW_u$$

Le prix risque-neutre de cette option à la date t est donnée par:

$$V_t = \mathbb{E} \left[e^{-r(T-t)} h(S_T) \mid \mathcal{F}_t \right]$$

Comme S est markovien et que la fonction payoff h ne dépend pas du temps, il existe une fonction $v(t, x)$ telle que $V_t = v(t, S_t)$. Par ailleurs $v(t, x)$ doit satisfaire l'EDP 4.15. Il s'agit de l'EDP d'évaluation dite de Black-Scholes-Merton (1973):

$$v_t(t, x) + rxv_x(t, x) + \frac{1}{2} \sigma^2 x^2 v_{xx}(t, x) = rv(t, x)$$

4.3.2 Cadre Multidimensionnel

Nous présentons dans cette Section les versions multidimensionnelles des théorèmes de Feynman-Kac simple et actualisé. Il s'agit des cas où le nombre d'équations aux dérivées partielles ainsi que le nombre de mouvements browniens entrant dans la définition de ces équations peut être supérieur à 1 et non nécessairement du même nombre.

Soit $W_u = (W_u^1, W_u^2)_{u \geq 0}$ un mouvement brownien bivarié, i.e.: un vecteur formé de deux mouvements browniens indépendants. Et nous considérons le système d'EDS:

$$dX_u^1 = a_1(u, X_u^1, X_u^2) du + b_{11}(u, X_u^1, X_u^2) dW_u^1 + b_{12}(u, X_u^1, X_u^2) dW_u^2$$

$$dX_u^2 = a_2(u, X_u^1, X_u^2) du + b_{21}(u, X_u^1, X_u^2) dW_u^1 + b_{22}(u, X_u^1, X_u^2) dW_u^2$$

La solution de ce système d'équations, pour un vecteur de conditions initiales $X_t^1 = x_1$ et $X_t^2 = x_2$ donné, est un processus de Markov. Soit par ailleurs $h(y_1, y_2)$ une fonction borélienne. Associé au triplet (t, x_1, x_2) pour $t \in [0, T]$, nous définissons:

$$g(t, x_1, x_2) = \mathbb{E}^{t, x_1, x_2} [h(X_T^1, X_T^2)] \quad (4.16)$$

$$f(t, x_1, x_2) = \mathbb{E}^{t, x_1, x_2} [e^{-r(T-t)} h(X_T^1, X_T^2)] \quad (4.17)$$

Alors, on peut montrer que 4.16 et 4.17 vérifient respectivement:

$$g_t + a_1 g_{x_1} + a_2 g_{x_2} + \frac{1}{2} (b_{11}^2 + b_{12}^2) g_{x_1 x_1} + (b_{11} b_{21} + b_{12} b_{22}) g_{x_1 x_2} + \frac{1}{2} (b_{21}^2 + b_{22}^2) g_{x_2 x_2} = 0$$

$$f_t + a_1 f_{x_1} + a_2 f_{x_2} + \frac{1}{2} (b_{11}^2 + b_{12}^2) f_{x_1 x_1} + (b_{11} b_{21} + b_{12} b_{22}) f_{x_1 x_2} + \frac{1}{2} (b_{21}^2 + b_{22}^2) f_{x_2 x_2} = 0$$

sujets aux conditions terminales:

$$g(T, x_1, x_2) = f(T, x_1, x_2) = h(x_1, x_2), \quad \forall (x_1, x_2)$$

4.3.3 Cas Mixte Brownien-Poisson

Nous énonçons le théorème de Feynman-Kac dans le cas d'une EDS-S de la forme 4.6.

Théorème 4.3.5 (Feynman-Kac) *Soit l'EDS:*

$$dX_u = a(u, X_u) du + b(u, X_u) dW_u + c(u, X_u) dJ_u \quad (4.18)$$

Soit $h(y)$ une fonction borélienne quelconque. Pour $T > 0$ et $t \in [0, T]$, nous définissons:

$$g(t, x) = \mathbb{E}^{t, x} [h(X_T)] \quad (4.19)$$

Alors $g(t, x)$ est dite satisfaire une équation aux dérivées partielles intégrodifférentielle (EDP-ID) donnée par:

$$g_t(t, x) + a(t, x)g_x(t, x) + \frac{1}{2}b^2(t, x)g_{xx}(t, x) + \lambda \int_{\mathbb{R}} [g(t, x + c(t, y)) - g(t, x)] \nu(dy) = 0 \quad (4.20)$$

pour la condition terminale:

$$g(T, x) = h(x), \quad \forall x$$

Preuve. La démonstration du Théorème de Feynman-Kac dans le cas d'un processus mixte se fait de façon similaire au cas d'une EDS, cf. Preuve 4.3.1. ■

Nous avons aussi la version actualisée de ce théorème.

Théorème 4.3.6 (Feynman-Kac actualisée) Soit $(X_u)_{u \geq 0}$ la solution d'une EDS-S de la forme 4.18 et $h(y)$ une fonction borélienne quelconque. Pour $T > 0$ et $t \in [0, T]$ donnés, nous définissons:

$$f(t, x) = \mathbb{E}^{t, x} \left[e^{-r(T-t)} h(X_T) \right]$$

Alors $f(t, x)$ satisfait l'EDP-ID suivante:

$$f_t(t, x) + a(t, x)f_x(t, x) + \frac{1}{2}b^2(t, x)f_{xx}(t, x) + \lambda \int_{\mathbb{R}} [f(t, x + c(t, y)) - f(t, x)] \nu(dy) = rf(t, x) \quad (4.21)$$

pour la condition terminale:

$$f(T, x) = h(x), \quad \forall x$$

4.4. Modèles de Taux d'Intérêt

Nous illustrons dans cette Section l'équipement précédemment introduit dans le cadre de modèles de taux d'intérêt. Les modèles qui décrivent l'évolution du taux d'intérêt R_t sont parfois qualifiés de modèles de *taux court* du fait que R_t représente le taux d'emprunt à court-terme. Le processus d'actualisation (*discount process*) est donnée par:

$$D_t = e^{-\int_0^t R_s ds}, \quad t \geq 0$$

et nous notons le processus de prix de marché (*money market account price process*) par:

$$\frac{1}{D_t} = e^{\int_0^t R_s ds}, \quad t \geq 0$$

Il s'agit de la valeur à la date t d'une unité monétaire investie dans le marché monétaire à la date 0 et continuellement actualisée au taux d'intérêt de court-terme R_s . Il vient alors:

$$dD_t = -R_t D_t dt, \quad d\left(\frac{1}{D_t}\right) = \frac{R_t}{D_t} dt$$

Une obligation zéro-coupon (*zero-coupon bond*) est un contrat promettant de payer une certaine valeur *faciale*, que nous prendrons par la suite égale à 1, à une certaine échéance T fixée. Le prix B_t^T de cet actif, pour $t \in [0, T]$, est donnée par:

$$D_t B_t^T = \mathbb{E}[D_T | \mathcal{F}_t]$$

Nous avons alors la formule d'évaluation du prix d'une obligation zéro-coupon suivante:

$$B_t^T = \mathbb{E}\left[e^{-\int_t^T R_s ds} | \mathcal{F}_t\right] \quad (4.22)$$

Par ailleurs, nous pouvons aussi définir le rendement (*yield*) entre les dates t et T comme:

$$Y_t^T = -\frac{1}{T-t} \log B_t^T$$

soit encore:

$$B_t^T = e^{-Y_t^T (T-t)}$$

Comme R satisfait une EDS de la forme 4.1, c'est un processus de Markov et alors:

$$B_t^T = f(t, R_t)$$

pour une fonction $f(t, r)$ quelconque en les variables t et r . Afin de déterminer l'EDP satisfaite par la fonction $f(t, r)$, nous appliquons la stratégie suivante:

- Déterminer la martingale,
- Prendre sa version différentielle,
- Annuler le terme en dt de cette martingale.

Dans notre cas, la martingale d'intérêt est $D_t B_t^T = D_t f(t, R_t)$. Sa différentielle vérifie:

$$\begin{aligned} d(D_t B_t^T) &= f(t, R_t) dD_t + D_t df(t, R_t) \\ &= D_t \left[-R f dt + f_t dt + f_r dR + \frac{1}{2} f_{rr} dR dR \right] \\ &= D_t \left[-R f + f_t + a f_r + \frac{1}{2} b^2 f_{rr} \right] dt + D_t b f_r dW \end{aligned}$$

ce qui amène à l'EDP suivante:

$$f_t(t, r) + a f_r(t, r) + \frac{1}{2} b^2 f_{rr}(t, r) = r f(t, r) \quad (4.23)$$

avec la condition terminale:

$$f(T, r) = 1, \quad \forall r \quad (4.24)$$

du fait que la valeur de l'obligation à maturité est égale à sa valeur *faciale*.

Exemple 4.4.1 (Modèle de Hull-White) Dans ce cadre, le taux d'intérêt R satisfait:

$$dR_t = (a_t - b_t R_t) dt + \sigma_t dW_t$$

Si nous nous intéressons à évaluer une obligation zéro-coupon, l'équation 4.23 devient:

$$f_t(t, r) + (a_t - b_t r) f_r(t, r) + \frac{1}{2} \sigma_t^2 f_{rr}(t, r) = r f(t, r) \quad (4.25)$$

Il est possible de montrer que la solution de 4.25 est de la forme:

$$f(t, r) = e^{-rC(t, T) - A(t, T)} \quad (4.26)$$

pour des fonctions déterministes $C(t, T)$ et $A(t, T)$ à déterminer telles que $C(T, T) = A(T, T) = 0$ d'après la condition terminale 4.24. En substituant 4.26 dans 4.25, on a:

$$C(t, T) = \int_t^T e^{-\int_t^s b_u du} ds$$

$$A(t, T) = \int_t^T \left(a_s C(s, T) - \frac{1}{2} \sigma_s^2 C^2(s, T) \right) ds$$

Exemple 4.4.2 (Option sur obligation) Supposons que nous souhaitons évaluer le prix d'une option d'achat de maturité T_1 écrite sur une obligation zéro-coupon de maturité T_2 avec $0 \leq t \leq T_1 < T_2$. D'après les résultats précédents, nous savons que la fonction $f(t, r)$, qui donne le prix d'une obligation zéro-coupon en fonction de t et r , satisfait une EDP de la forme 4.23 avec la condition terminale 4.24. La valeur du call à la date t est:

$$c(t, R_t) = \mathbb{E} \left[e^{-\int_t^{T_1} R_s ds} (f(T_1, R_{T_1}) - K)^+ | \mathcal{F}_t \right]$$

$$= \frac{1}{D_t} \mathbb{E} \left[D_{T_1} (f(T_1, R_{T_1}) - K)^+ | \mathcal{F}_t \right]$$

La valeur actualisée du call vérifie:

$$D_t c(t, R_t) = \mathbb{E} \left[D_{T_1} (f(T_1, R_{T_1}) - K)^+ | \mathcal{F}_t \right]$$

et c'est une martingale dont la forme différentielle s'écrit:

$$d(D_t c(t, R_t)) = c(t, R_t) dD_t + D_t dc(t, R_t)$$

$$= D \left[-Rcdt + c_t dt + c_r dR + \frac{1}{2} c_{rr} dRdR \right]$$

$$= D \left[-Rc + c_t + ac_r + \frac{1}{2} b^2 c_{rr} \right] dt + D b c_r dW$$

Soit en annulant le terme en dt , nous obtenons l'EDP d'évaluation:

$$c_t(t, r) + b(t, r) c_r(t, r) + \frac{1}{2} b^2(t, r) c_{rr}(t, r) = r c(t, r)$$

sujette à la condition terminale:

$$c(T_1, r) = (f(T_1, r) - K)^+, \quad \forall r$$

C

Simulation des EDS-S

Cette Annexe présente des méthodes de discrétisation en temps des équations différentielles stochastiques à sauts (EDS-S). Nous verrons dans un premier temps que l'existence d'une solution exacte pour un processus stochastique est plus l'exception que la règle et qu'il est alors important de pouvoir disposer de schémas d'approximation performants des EDS-S afin de pouvoir réaliser par la suite des simulations à faible perte: c'est-à-dire dont les trajectoires discrètes ne s'éloignent pas trop des vraies trajectoires continues du processus. Nous présentons ainsi le schéma de discrétisation d'Euler qui est le plus simple et étudions aussi des méthodes pour améliorer son efficacité, ce qui nous amène au schéma de Milstein.

C.1. Simulation Exacte

L'exploitation d'un système de processus stochastiques nécessite de pouvoir simuler des réalisations de ce dernier, c'est-à-dire de simuler des trajectoires. Dans ce cadre, il est important de distinguer le cas exact du cas approché. Le cadre dit exact de simulation correspond à la situation où la distribution jointe des valeurs simulées coïncide avec la distribution jointe en temps continu du processus, sur une grille temporelle de simulation ; on dit aussi que le processus admet une solution de forme forte. Sinon, on parle de solution de forme faible et d'un schéma approché de simulation.

C.1.1 Cas du Mouvement Brownien

Soit $(W_t)_{t \geq 0}$ un mouvement brownien et $(X_t)_{t \geq 0}$ un processus de Markov tel que:

$$dX_t = \mu dt + \sigma dW_t$$

Si nous posons $S_t = S_0 \exp(X_t)$, alors d'après le lemme d'Itô, cf. Théorème 3.3.6, il vient:

$$dS_t = S_t \left[\left(\mu - \frac{\sigma^2}{2} \right) dt + \sigma dW_t \right]$$

ou de façon équivalente, en utilisant à nouveau la formule d'Itô sur la fonction $\log(S_t)$:

$$d \log(S_t) = \left(\mu - \frac{\sigma^2}{2} \right) dt + \sigma dW_t \tag{C.1}$$

Une intégration directe de C.1 pour tout $u < t$ donne:

$$S_t = S_u \exp \left[\left(\mu - \frac{\sigma^2}{2} \right) (t - u) + \sigma (W_t - W_u) \right]$$

Par ailleurs, comme les incréments de $(W_t)_{t \geq 0}$ sont indépendants et distribués de façon gaussienne, il est possible de dériver une procédure récursive simple pour simuler des valeurs de S pour des dates $0 = t_0 < t_1 < \dots < t_n = T$, sous la forme:

$$S_{t_{i+1}} = S_{t_i} \exp \left[\left(\mu - \frac{\sigma^2}{2} \right) (t_{i+1} - t_i) + \sigma \sqrt{t_{i+1} - t_i} Z_{t_{i+1}} \right] \quad (\text{C.2})$$

où les variables aléatoires Z_1, \dots, Z_n sont des variables aléatoires gaussiennes standard. La méthode de simulation C.2 est alors exacte dans le sens où les variables S_{t_1}, \dots, S_{t_n} qu'elle produit ont la distribution jointe du processus S aux dates t_1, \dots, t_n et la méthode ne produit pas alors de biais de discrétisation. Nous présentons successivement une procédure générale de discrétisation sur une grille du temps et un algorithme de simulation basé sur C.2.

Algorithme C.1.1 (Discrétisation du temps) *Nous cherchons à représenter un processus stochastique $(S_t)_t$ pour $t \in [t_0, T]$ par une trajectoire discrète, que nous notons:*

$$\{(t_0, S_0), \dots, (t_i, S_i), \dots, (t_{N-1}, S_{N-1})\}$$

avec $N \in \mathbb{N}$ le nombre de points de la trajectoire. Notons Δ le pas de mesure entre deux observations. Nous avons alors:

$$\Delta = \frac{T - t_0}{N - 1} \text{ et } t_i = \Delta(i + 1)$$

Algorithme C.1.2 (Processus de Wiener)

► Simulation de $(S_{t_1}, \dots, S_{t_n})$ pour n dates fixées t_1, \dots, t_n sur l'intervalle $[t_0, T]$.

S_t est un mouvement brownien de dérive μ et de volatilité σ .

- Simuler n variables aléatoires i.i.d. N_1, \dots, N_n suivant une loi $\mathcal{N}(0, 1)$.
- Calculer $\Delta S_k = \sigma \sqrt{t_{k+1} - t_k} N_k + \mu (t_{k+1} - t_k)$.

► Une trajectoire est alors donnée par:

$$S_{t_i} = \sum_{k=1}^i \Delta S_k$$

La Figure C.1 présente la simulation d'un processus de Wiener à l'aide de l'Algorithme C.1.2. Il nous permet de juger de deux choses. D'une part, du phénomène du passage du temps continu au temps discret et d'autre part de la convergence de la trajectoire discrète vers la trajectoire continue. L'un des problèmes lors du passage du temps continu au temps discret est que nous perdons de l'information. Nous pouvons en effet montrer que pour des processus moins *simples* que le mouvement brownien géométrique, la perte d'information s'accompagne aussi de sa déformation.

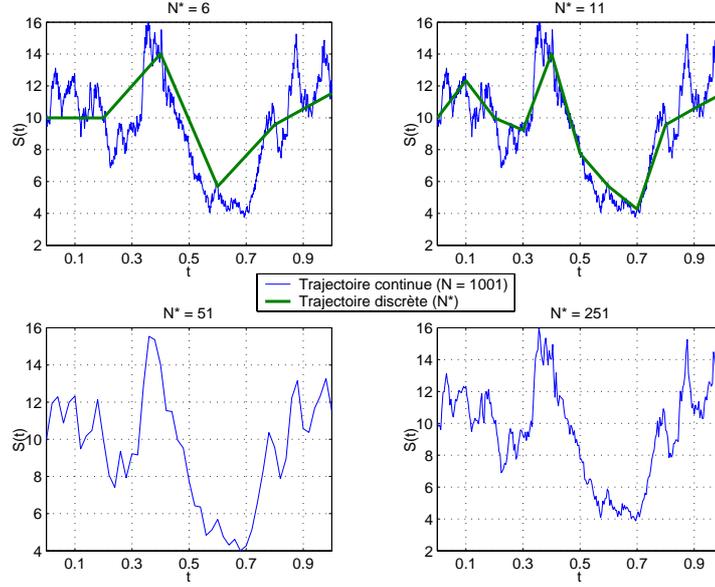


Figure C.1: Passage du temps continu au temps discret pour un mouvement brownien.

C.1.2 Cas des Processus à Sauts

Soit $(W_t)_{t \geq 0}$ un mouvement brownien standard et $(J_t)_{t \geq 0}$ un processus de sauts pur, indépendants l'un de l'autre. Le processus de Markov $(X_t)_{t \geq 0}$ est dit suivre processus mixte, cf. Section 4.1.2, s'il est solution de l'équation:

$$dS_t = S_{t-} [\mu dt + \sigma dW_t + dJ_t] \quad (\text{C.3})$$

où μ et σ sont des constantes réelles et où en particulier J_t est donné par:

$$J_t = \sum_{j=1}^{N_t} (Y_j - 1) \quad (\text{C.4})$$

où les $(Y_j)_{j \geq 1}$ sont des variables aléatoires de distribution quelconque et N_t est un processus de comptage. On distingue des dates d'arrivée ou de survenance:

$$0 < \tau_1 < \tau_2 < \dots$$

et le terme:

$$N_t = \sup \left\{ n : \sum_{i=1}^n \tau_i < t \right\} \quad (\text{C.5})$$

compte le nombre d'arrivées dans l'intervalle $[0, t]$. Et aussi des amplitudes qui sont décrites par les réalisations $Y_j - 1$ si $t = \tau_j$ et 0 si t ne coïncide avec

aucune des dates τ_j . Pour clarifier les expressions, nous suivons les convention usuelles et supposons que le processus S est continu à droite, soit:

$$S_t = \lim_{u \downarrow t} S_u$$

n'inclut pas l'effet d'un saut. Pour prendre en compte la possible survenance d'un saut, nous notons S_{t-} qui est la limite à gauche:

$$S_{t-} = \lim_{u \uparrow t} S_u$$

Une solution de l'équation C.3 est donnée par:

$$S_t = S_0 e^{\left(\mu - \frac{\sigma^2}{2}\right)t + \sigma W_t} \prod_{j=1}^{N_t} Y_j \quad (\text{C.6})$$

Nous étudions à présent deux stratégies pour la simulation de trajectoires d'un modèle mixte. Dans la première, nous ne prenons pas en compte la possible occurrence des sauts et simulons à des dates $0 = t_0 < t_1 < \dots < t_{n-1} < t_n$, tandis que dans le second cas, nous simulons explicitement les instants de sauts $(\tau_i)_{i \geq 1}$ qui peuvent alors ne pas correspondre avec les dates $(t_i)_{i \geq 1}$.

Simulation à Dates Fixes

Pour simuler aux dates t_1, \dots, t_n , nous utilisons la solution de forme forte suivante, dérivée à partir de C.6 de façon récursive:

$$S_{t_{i+1}} = S_{t_i} e^{\left(\mu - \frac{\sigma^2}{2}\right)(t_{i+1} - t_i) + \sigma(W_{t_{i+1}} - W_{t_i})} \prod_{j=N_{t_i}+1}^{N_{t_{i+1}}} Y_j \quad (\text{C.7})$$

Nous nous intéressons dans un premier temps à la simulation de la partie sauts de C.7 telle que définie par C.4. La simulation de la partie diffusion se réalise dans le même cadre que l'Algorithme C.1.2 pour un mouvement brownien.

Algorithme C.1.3 (Processus de Poisson composé - dates fixes)

► Simulation de $(J_{t_1}, \dots, J_{t_n})$ pour n dates fixées t_1, \dots, t_n sur l'intervalle $[t_0, T]$.

J_t est un processus de Poisson composé d'intensité λ et de distribution des sauts m .

- Simuler une variable aléatoire N_i suivant une loi $\mathcal{P}(\lambda(t_{i+1} - t_i))$.
 N_i représente alors le nombre total de sauts survenus dans l'intervalle $[t_i, t_{i+1})$.
- Instants de sauts.
Simuler N_i variables aléatoires i.i.d. U_j distribuées selon $\mathcal{U}(t_i, t_{i+1})$.
- Tailles des sauts.
Simuler N_i variables aléatoires i.i.d. Y_j distribuées selon m .

► Une trajectoire est alors donnée par:

$$J_{t_i} = \sum_{j=1}^{N_i} \mathbf{1}_{U_j < t_i} Y_j$$

Simulation aux Instants de Sauts

La stratégie précédente ne prend pas en compte les instants où le processus S saute. Au contraire, il génère un nombre total de sauts dans chaque intervalle $(t_i, t_{i+1}]$ sous l'hypothèse que le nombre de sauts suit une distribution de Poisson. Mais il est aussi possible d'adopter une approche où l'on simule explicitement les dates auxquelles le processus S saute. Pour ce faire, nous remarquons qu'entre deux occurrences successives d'un saut, le processus S évolue comme un mouvement brownien géométrique ordinaire - du fait que nous avons supposé l'indépendance entre W et J . De fait, conditionnellement aux instants de sauts τ_1, τ_2, \dots , nous avons:

$$S_{\tau_{j+1}-} = S_{\tau_j} e^{\left(\mu - \frac{\sigma^2}{2}\right)(\tau_{j+1} - \tau_j) + \sigma(W_{\tau_{j+1}} - W_{\tau_j})} \quad (\text{C.8})$$

qui représente le processus S lorsqu'aucun saut ne s'est produit et:

$$S_{\tau_{j+1}} = S_{\tau_{j+1}-} J_{j+1} \text{ où } J_{j+1} = \sum_{i=1}^{N_{\tau_{j+1}}} Y_i \quad (\text{C.9})$$

qui prend en compte l'occurrence d'un saut. Nous nous intéressons alors à la simulation de la partie sauts de C.9. La simulation de la partie diffusion C.8 se réalise dans le même cadre que l'Algorithme C.1.2 pour un mouvement brownien.

Algorithme C.1.4 (Processus de Poisson composé - instants de sauts)

► Simulation de $(J_{t_1}, \dots, J_{t_n})$ pour n dates fixées t_1, \dots, t_n sur l'intervalle $[t_0, T]$.

J_t est un processus de Poisson composé d'intensité λ et de distribution des sauts m .

- Initialiser $k = 0$.
- Tant que $\sum_{j=1}^k T_j < t_i$, faire :
 - Poser $k = k + 1$,
 - Simuler $\tau_j \sim \text{Exp}(\lambda(t_{j+1} - t_j))$,
 - Simuler $Y_j \sim m(\cdot)$.

► Une trajectoire est alors donnée par:

$$J_{t_i} = \sum_{j=1}^{N_{t_i}} Y_j, \text{ où, } N_{t_i} = \sup \left\{ k : \sum_{i=1}^k \tau_i < t \right\}$$

Après avoir envisagé deux stratégies possibles de simulation de la partie sauts d'un modèle mixte, nous pouvons à présent simuler toute sa trajectoire. On remarquera que la différence entre les procédures décrites par les Algorithmes C.1.3 et C.1.4 est que dans le premier cas on se donne explicitement les dates (t_1, \dots, t_n) d'observations des trajectoires, tandis que dans le second cas on ne dispose que des instants de survenance des sauts (τ_1, \dots) . De façon générale, l'Algorithme C.1.3 fournit la meilleure approximation car il tient compte de l'information sur la distribution de probabilité de N_t , ce que ne prend pas en compte l'Algorithme C.1.4 qui l'estime de façon empirique. Nous présentons alors un algorithme général de simulation d'un modèle mixte pour des trajectoires calculées à dates fixes.

Algorithme C.1.5 (Modèle mixte - EDS-S)

► Simulation de $(S_{t_1}, \dots, S_{t_n})$ pour n dates fixées t_1, \dots, t_n sur l'intervalle $[t_0, T]$.

S_t est un processus mixte, i.e.: somme d'un mouvement brownien de dérive μ et de volatilité σ et d'un processus de Poisson composé d'intensité

λ

et de distribution des sauts m .

- Partie diffusion. Simulation selon l'Algorithme C.1.2.
- Partie sauts. Simulation selon l'Algorithme C.1.3.

► Une trajectoire est alors donnée par:

$$S_{t_i} = \mu t_i + \sum_{k=1}^i \Delta S_k + \sum_{j=1}^N \mathbf{1}_{U_j < t_i} Y_j$$

Remarque C.1.6 La simulation exacte de processus mixtes est rendue possible sous l'hypothèse d'indépendance des processus W et J et celle-ci devient plus délicate dès lors que les dates de survenance des sauts suivent une distribution paramétrée sur la trajectoire du processus.

C.2. Méthodes de Discrétisation

Dès lors qu'il n'est plus possible de dériver une solution de forme forte d'un processus stochastique, nous avons recours aux méthodes dites de discrétisation. Nous commençons par présenter et discuter le schéma d'Euler, qui est la méthode la plus simple pour la simulation approchée d'une diffusion. Par la suite, nous présentons un prolongement de cette stratégie, ce qui nous amène à la méthode de Milstein.

C.2.1 Schéma d'Euler

Nous considérons un processus de Markov $(X_t)_{t \geq 0}$ solution de:

$$dX_t = a(X_t) dt + b(X_t) dW_t \tag{C.10}$$

où $(W_t)_{t \geq 0}$ est un mouvement brownien standard, pour une condition initiale fixé $X_0 = x_0$ et où les fonctions $a(\cdot)$ et $b(\cdot)$ sont lipschitziennes¹. Nous notons par \widehat{X} une version discrétisée (donc approchée) de X . L'approximation d'Euler, pour les dates $t_0 < t_1 < \dots < t_n$ est donnée par $\widehat{X}_0 = X_0$ et de façon récursive pour $i = 0, 1, \dots, n-1$:

$$\widehat{X}_{t_{i+1}} = \widehat{X}_{t_i} + a(\widehat{X}_{t_i})(t_{i+1} - t_i) + b(\widehat{X}_{t_{i+1}}) \sqrt{t_{i+1} - t_i} Z_{t_{i+1}} \quad (\text{C.11})$$

où les variables aléatoires Z_1, \dots, Z_n sont des variables aléatoires i.i.d. suivant une loi normale centrée réduite.

Algorithme C.2.1 (Discrétisation d'Euler d'une diffusion)

► Simulation de $(X_{t_1}, \dots, X_{t_n})$ pour n dates fixées t_1, \dots, t_n sur l'intervalle $[t_0, T]$.

- Initialiser $\widehat{X}_{t_0} = X_0$.
 - Calculer les valeurs $a_{t_i} = a(\widehat{X}_{t_i})$ et $b_{t_i} = b(\widehat{X}_{t_i})$.
 - Calculer $\Delta \widehat{X}_{t_i} = a_{t_i}(t_{i+1} - t_i) + b_{t_i} \sqrt{t_{i+1} - t_i} N_{t_i}$.
- Une trajectoire est alors donnée par:

$$\widehat{X}_{t_{i+1}} = \widehat{X}_{t_i} + \Delta \widehat{X}_{t_i}$$

C.2.2 Schéma de Milstein

Le schéma de discrétisation d'Euler repose sur une approximation de premier ordre de l'équation C.10. En effet, sur l'intervalle $[t, t+h]$, nous avons:

$$X_{t+h} = X_t + \int_t^{t+h} a(X_u) du + \int_t^{t+h} b(X_u) dW_u \quad (\text{C.12})$$

De C.11 et C.12, il est clair que le schéma d'Euler résulte de l'approximation suivante:

$$\int_t^{t+h} a(X_u) du \approx a(X_t) h$$

et:

$$\int_t^{t+h} b(X_u) dW_u \approx b(X_t) [W_{t+h} - W_t] \quad (\text{C.13})$$

C'est-à-dire qu'une intégrale, sur l'intervalle $[t, t+h]$, est approchée par sa valeur en t . Dès lors, il est possible d'améliorer la qualité du schéma en proposant une approximation plus fine du terme de diffusion, comme par exemple en augmentant le nombre et le degré des termes intervenant dans le développement de l'intégrale. Ainsi, une application de la formule d'Itô à la fonction

¹La fonction numérique f est lipschitzienne de rapport k sur un intervalle I si pour tout $(x, y) \in I^2$ on a $|f(x) - f(y)| \leq k|x - y|$. Par ailleurs, toute fonction lipschitzienne sur un intervalle I est continue.

$b(X_t)$ donne:

$$\begin{aligned} db(X_t) &= b'(X_t) dX_t + \frac{1}{2} b''(X_t) b^2(X_t) dt \\ &= \mu_b(X_t) dt + \sigma_b(X_t) dW_t \end{aligned}$$

Et une application du schéma d'Euler au processus $b(X_u)$ pour $t \leq u \leq t+h$ produit:

$$\begin{aligned} b(X_u) &\approx b(X_t) + \mu_b(X_t)[u-t] + \sigma_b(X_t)[W_u - W_t] \\ &= b(X_t) + \left(b'(X_t) a(X_t) + \frac{1}{2} b''(X_t) b^2(X_t) \right) [u-t] \\ &\quad + b'(X_t) b(X_t) [W_u - W_t] \end{aligned} \quad (\text{C.14})$$

Soit, si nous omettons les termes en $O(u-t)$, C.14 se simplifit sous la forme:

$$b(X_u) \approx b(X_t) + b'(X_t) b(X_t) [W_u - W_t] \quad (\text{C.15})$$

Nous obtenons alors une nouvelle approximation de C.13 en remplaçant $b(X_u)$ par C.15:

$$\begin{aligned} \int_t^{t+h} b(X_u) dW_u &\approx \int_t^{t+h} b(X_t) + b'(X_t) b(X_t) [W_u - W_t] dW_u \\ &= b(X_t) [W_{t+h} - W_t] \\ &\quad + b'(X_t) b(X_t) \left[\int_t^{t+h} [W_u - W_t] dW_u \right] \end{aligned} \quad (\text{C.16})$$

Pour être utilisable, l'expression C.16 nécessite que la dernière intégrale (entre crochets) soit simplifiée. Il est possible de vérifier que cette intégrale est la solution de l'équation:

$$Y_t = \frac{1}{2} [W_t^2 - t^2]$$

ce qui nous permet de la simplifier, soit en substituant les expressions:

$$\int_t^{t+h} [W_u - W_t] dW_u = \frac{1}{2} \left([W_{t+h} - W_t]^2 - t^2 \right)$$

Finalement, le résultat précédent permet d'améliorer la discrétisation de premier ordre de C.12 fourni par Euler et que nous rappelons:

$$X_{t+h} \approx X_t + a(X_t) h + b(X_t) [W_{t+h} - W_t]$$

par l'ajout d'un terme de correction de second ordre qui est le schéma de Milstein:

$$\begin{aligned} X_{t+h} &\approx X_t + a(X_t) h + b(X_t) [W_{t+h} - W_t] \\ &\quad + \frac{1}{2} b'(X_t) b(X_t) \left([W_{t+h} - W_t]^2 - h \right) \end{aligned}$$

Nous avons alors l'algorithme de simulation suivant dû à Milstein.

Algorithme C.2.2 (Discrétisation de Milstein d'une diffusion)

► Simulation de $(X_{t_1}, \dots, X_{t_n})$ pour n dates fixées t_1, \dots, t_n sur l'intervalle $[t_0, T]$.

- Initialiser $\widehat{X}_{t_0} = X_0$ et poser $\Delta = t_{i+1} - t_i$
- Calculer les valeurs $a_{t_i} = a(\widehat{X}_{t_i})$, $b_{t_i} = b(\widehat{X}_{t_i})$, $b'_{t_i} = b'(\widehat{X}_{t_i})$.
- Calculer $\Delta W_{t_i} = \sqrt{\Delta} N_{t_i}$.

► Une trajectoire est alors donnée par:

$$\widehat{X}_{t_{i+1}} = \widehat{X}_{t_i} + a_{t_i} \Delta + b_{t_i} \Delta W_{t_i} + \frac{1}{2} b_{t_i} b'_{t_i} ([\Delta W_{t_i}]^2 - \Delta)$$

Etude Comparative

Il peut être intéressant d'étudier les qualités respectives des deux schémas de discrétisation, Euler et Milstein. Un bon exemple est fourni par la discrétisation du processus de Cox-Ingersoll-Ross, $dr_t = (a - br_t) dt + \sigma \sqrt{r_t} dW_t$. Ce processus a la particularité d'avoir une probabilité nulle d'atteindre des valeurs négatives, c'est-à-dire:

$$\mathbb{P}[X(t) \leq 0] = 0, \forall t$$

Cependant, les simulations numériques du processus de CIR nécessitent de recourir à une discrétisation de son équation. Et dans ce cas, le processus discrétisé a une probabilité non nulle de devenir négatif. Notre étude porte alors sur la comparaison des schémas d'Euler et de Milstein en ce qui concerne les points aberrants, c'est-à-dire les points atteignant zéro dans une trajectoire.

Lorsque la volatilité est faible (paramètre σ_v), nous observons que les processus générés selon Euler et Milstein sont très proches. Mais lorsque l'on augmente la volatilité (de 0.3 à 0.5 puis à 0.7), des différences sont observées et en particulier sur les points les plus éloignés de la moyenne. Les résultats sont reportés dans le Tableau C.1.

% σ_v	Moyenne		Ecart-type		Points à zéro		Chemins à zéro	
	Euler	Milstein	Euler	Milstein	Euler	Milstein	Euler	Milstein
0.1	0.0903	0.0902	0.0107	0.0107	0	0	0	0
0.3	0.0897	0.0897	0.0321	0.0322	0.6	0	6	0
0.5	0.0911	0.0904	0.0527	0.0543	2.6	0.5	90	20.8
0.7	0.0945	0.091	0.0721	0.0769	6.8	2.1	100	87.6
0.9	0.1001	0.092	0.0905	0.0982	10.8	5.8	100	99.6

Table C.1: Simulations de points et de chemins touchant zéro.

Et lorsque la volatilité est très importante, le nombre de points touchant zéro augmente bien plus rapidement pour le schéma d'Euler que pour le schéma de Milstein. De façon plus précise, nous nous sommes intéressés à compter ce nombre de points à zéro sur un grand nombre de trajectoires et aussi le nombre de trajectoires dont au moins un des points devient négatif. Un exemple de

trajectoires d'un processus de CIR simulé par Euler et Milstein est donné sur la Figure C.2.

Nous constatons alors que le schéma de Milstein produit de meilleurs résultats que le schéma d'Euler, étant donné que les pourcentages moyens de points à zéro dans une trajectoire et de chemins touchant zéro sont plus faibles dans le cas de l'emploi du schéma de Milstein. Toutefois, dès lors que σ_v devient très grand (au delà de 0.7, voir de 1.0), les deux schémas produisent le même type de résultats.

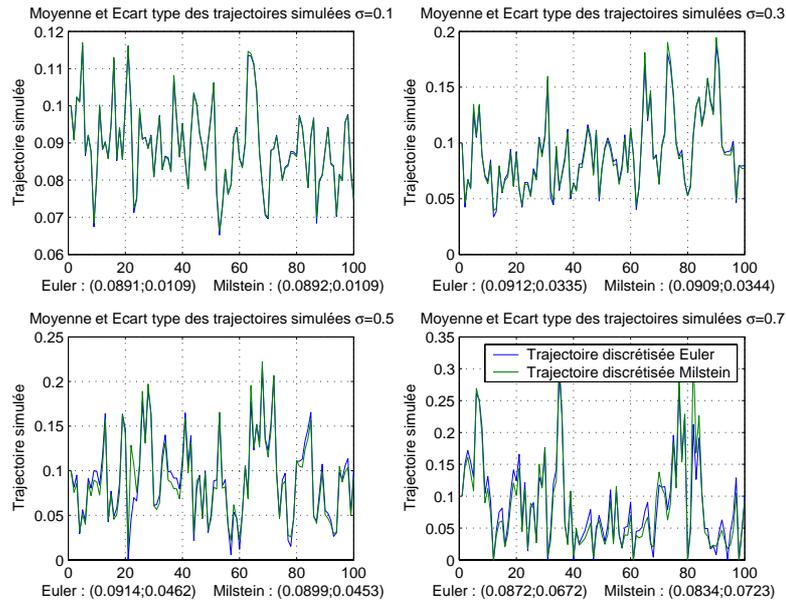


Figure C.2: Effet de la volatilité σ_v sur les trajectoires simulées d'un processus CIR.

Part II

Statistique des Processus

5

Inférence Statistique

5.1. Contexte

Toutes les procédures d'inférence statistique sont basées sur le même principe de sélection des paramètres d'un modèle dans le but d'optimiser un certain critère calculé à partir de données discrètes, soit $(x_t, t = 1, \dots, T)$. Dans le cadre du maximum de vraisemblance, ce critère est la fonction de vraisemblance, tandis que pour la méthode des moments (généralisés), le critère est construit comme la différence entre certains moments théoriques de la distribution et leurs équivalents empiriques. Dans la plupart des cas, le problème d'optimisation n'admet pas de solution analytique et il faut alors recourir à des méthodes numériques.

5.1.1 Estimations MV et MMG

Nous présentons par la suite les principes de la méthode du maximum de vraisemblance (MV ou *maximum likelihood*, soit ML) et de la méthode des moments généralisés (MMG ou *generalized method of moments*, soit GMM)

Principe MV

La méthode du maximum de vraisemblance (MV) est l'approche la plus couramment employée pour estimer un modèle paramétrique. Son principe est le suivant. Etant donné une fonction $f(x; \theta)$ pour la densité de la (des) variable(s) d'état du système dépendante d'un vecteur $\theta \in \Theta$ de paramètres et un vecteur d'observations $(x_t, t = 1, \dots, T)$, on choisit le vecteur θ des paramètres du modèle qui maximise la vraisemblance que les observations ont bien été générées depuis le modèle théorique choisi. Soit alors le critère:

$$\theta^* = \arg \max_{\theta \in \Theta} \prod_{t=1}^T f(x_t; \theta)$$

Ce qui est équivalent à maximiser la fonction de log-vraisemblance:

$$l(\theta) = \sum_{t=1}^T l_t(\theta) = \sum_{t=1}^T \ln f(x_t; \theta)$$

Si la densité $f(\cdot|\theta)$ est connue de façon analytique, il est alors possible de dériver une équation pour les solutions maximales de $l(\theta)$ en différentiant l et en résolvant cette équation numériquement. Toutefois, il se peut que pour nombre de processus et en particulier les modèles mixtes, cf. Section 4.1.2, la fonction de vraisemblance ne soit pas connue explicitement, mais nécessite le calcul de fonctions spéciales qui ne peuvent être évaluées que numériquement. Par conséquent, dans ces cas, même la fonction de vraisemblance doit être évaluée numériquement.

Exemple 5.1.1 (MBG) *Nous reprenons l'exemple du mouvement brownien géométrique:*

$$\frac{dX_t}{X_t} = \mu dt + \sigma dW_t, \quad X_{t_0} = x_{t_0}$$

Nous avons précédemment vu que ce processus admettait la solution:

$$X_t = x_{t_0} \exp\left(\left(\mu - \frac{1}{2}\sigma^2\right)(t - t_0) + \sigma(W_t - W_{t_0})\right)$$

et que sa distribution limite était log-normale. Soit, pour un vecteur d'observations $(x_t)_{t \geq 0}$, la densité f du modèle est telle que:

$$\ln f(x_t; \theta) = -\frac{1}{2} \ln(2\pi\sigma^2\Delta) - \frac{1}{2\sigma^2\Delta} \left[\ln \frac{x_t}{x_{t-1}} - \left(\mu - \frac{1}{2}\sigma^2\right)\Delta \right]^2$$

où Δ est le pas de temps entre deux observations successives.

Exemple 5.1.2 (MBG & sauts gaussiens) *Supposons que le processus $(X_t)_{t \geq 0}$ satisfait à l'EDS-S, dite de Merton (1976), suivante:*

$$dX_t = \mu dt + \sigma dW_t + \sum_{i=1}^{N_t} Y_i \quad (5.1)$$

où $(W_t)_{t \geq 0}$ est un mouvement brownien et $(N_t)_{t \geq 0}$ un processus de Poisson d'intensité $\lambda > 0$. La taille des sauts $(Y_i)_{i \geq 1}$ est supposée suivre une distribution normale, soit $Y_i \sim \mathcal{N}(\beta, \eta)$. Une application de la formule d'Itô à $S_t = S_0 \exp(X_t)$ donne:

$$\frac{dS_t}{S_{t-}} = \left(\mu - \frac{\sigma^2}{2}\right) dt + \sigma dW_t + \sum_{i=1}^{N_t} (\exp(Y_i) - 1)$$

La probabilité de transition de X peut être spécifiée de façon explicite. Une discrétisation en temps de 5.1 selon un schéma d'Euler nous permet de dériver la solution approchée:

$$Y_\Delta = X_\Delta - X_0 = \mu\Delta + \sigma W_\Delta + \int_0^\Delta Y_s dN_s$$

où $W_\Delta \sim \mathcal{N}(0, \Delta)$ et:

$$P(N_\Delta = n; \theta) = \frac{(\lambda\Delta)^n}{n!} e^{-\lambda\Delta}$$

Il s'ensuit que les variables aléatoires Z_Δ sont i.i.d. Par suite:

$$P(X_\Delta \leq x | X_0 = x_0, \Delta; \theta) = \sum_{n=0}^{+\infty} P(X_\Delta \leq x | X_0 = x_0, \Delta, N_\Delta = n; \theta) P(N_\Delta = n; \theta)$$

Conditionnellement à l'évènement $N_\Delta = n$, il s'est produit exactement n dates de sauts, $\tau_i, i = 1, \dots, n$, entre 0 et Δ tel que $dN_{\tau_i} = 1$, soit:

$$\int_0^\Delta Y_s dN_s = \sum_{i=1}^n Y_{\tau_i}$$

qui est la somme de n termes de sauts indépendants. Et comme $J_t \sim \mathcal{N}(\beta, \eta)$, alors la probabilité de transition de X_Δ sachant X_0 est donnée par:

$$\begin{aligned} f(x|x_0, \Delta; \theta) &= \sum_{n=0}^{+\infty} f(x|x_0, \Delta, N_\Delta = n; \theta) P(N_\Delta = n; \theta) \\ &= \sum_{n=0}^{+\infty} \frac{e^{-\lambda\Delta} (\lambda\Delta)^n}{n! \sqrt{2\pi} \sqrt{n\eta + \Delta\sigma^2}} \exp\left(-\frac{(x - x_0 - \mu\Delta - n\beta)^2}{2(n\eta + \Delta\sigma^2)}\right) \end{aligned}$$

Il est finalement possible de proposer une approximation de premier ordre en Δ de cette densité, sous la forme:

$$\begin{aligned} f^{(1)}(x|x_0, \Delta; \theta) &= \frac{(1 - \lambda\Delta)}{\sqrt{2\pi} \sqrt{\Delta\sigma^2}} \exp\left(-\frac{(x - x_0 - \mu\Delta)^2}{2\Delta\sigma^2}\right) \\ &\quad + \frac{\lambda\Delta}{\sqrt{2\pi} \sqrt{n\eta + \Delta\sigma^2}} \exp\left(-\frac{(x - x_0 - \mu\Delta - n\beta)^2}{2(n\eta + \Delta\sigma^2)}\right) \end{aligned}$$

ce qui rend l'estimation par maximum de vraisemblance plus aisée.

Remarque 5.1.3 Le modèle précédent, un mouvement brownien géométrique avec une partie de sauts gaussiens est l'un des rares exemples où il est possible de complètement spécifier la vraisemblance du modèle sous-jacent. En dehors de ce cadre, on préférera au maximum de vraisemblance une autre méthode d'inférence, par exemple la méthode des moments.

Principe MMG

Alors qu'il est parfois difficile d'obtenir la fonction de vraisemblance d'un processus et en particulier d'un processus mixte, cf. Section 4.1.2, sous une forme explicite, il est presque toujours possible de calculer les moments théoriques en fonction des paramètres θ du modèle: ils sont en effet obtenus par différentiation de la fonction caractéristique, cf. Section 1.2.3. Il est alors possible de construire des estimateurs de la méthode des moments en égalisant les moments théoriques aux moments empiriques et obtenir ainsi une estimation des paramètres du modèle. La méthode de moments généralisés (MMG) consiste à choisir le vecteur de paramètres θ qui réalise - dans le sens des moindres carrés - un vecteur de moyennes empiriques:

$$\theta^* = \arg \max_{\theta \in \Theta} M^\top(\theta) WM(\theta)$$

pour le vecteur de moments:

$$M(\theta) = \langle f_i(x_i) \rangle - \mathbb{E}[f_i(x_i) | \theta], \quad i = 1, \dots, m$$

où les $f_i(\cdot)$ sont un ensemble de quantités choisies pour produire des conditions de moments et W une matrice symétrique définie positive, dite de pondération. Les moments généralisés $\langle f_i(x_i) \rangle$ peuvent représenter la moyenne empirique de n'importe quelle quantité du modèle, c'est-à-dire les moments habituels (de tout ordre) et les cumulants, mais aussi des probabilités empiriques ($f_i = \mathbf{1}_{[a,b]}$) ou des quantiles.

Etant donné que les conditions de moments d'un estimateur GMM ne sont pas indépendantes les unes des autres, il est probable que les erreurs sur les différents moments soient corrélées. Il n'est pas alors souhaitable de considérer une matrice de pondération W qui soit diagonale : de façon idéale, W devrait être choisie en fonction des données. Un choix optimal de matrice de pondération est détaillé par Hansen.

Nous présentons par ailleurs une méthode alternative à l'utilisation de la fonction caractéristique pour obtenir des conditions de moments pour un processus mixte du type 5.1. A cette fin, pour un processus stochastique X , nous définissons son générateur infinitésimal \mathcal{A} par son action sur des fonctions $f(\delta, x, x_0)$ mesurables par:

$$\begin{aligned} \mathcal{A}f(\Delta, x, x_0) &= \frac{\partial f(\Delta, x, x_0)}{\partial \Delta} + \mu \frac{\partial f(\Delta, x, x_0)}{\partial x} + \frac{1}{2} \sigma^2 \frac{\partial^2 f(\Delta, x, x_0)}{\partial x^2} \\ &\quad + \lambda \mathbb{E}_Y [f(\Delta, x + Y, x_0) - f(\Delta, x, x_0)] \end{aligned}$$

cf. Proposition 2.3.10 et où \mathbb{E}_Y est l'espérance par rapport à la distribution des sauts Y . Et nous évaluons une espérance conditionnelle par un développement de Taylor, soit:

$$\mathbb{E}[f(\Delta, X_\Delta, x_0) | X_0 = x_0] = \sum_{k=0}^K \frac{\Delta^k}{k!} \mathcal{A}^k f(\delta, x, x_0)|_{\delta=0, x=x_0} + O(\Delta^{K+1}) \quad (5.2)$$

qui est une série de Taylor, qui peut être calculée de façon exacte dans certains cas.

Exemple 5.1.4 (MBG & sauts gaussiens) *Nous reprenons l'Exemple 5.1.2 et appliquons 5.2 à la fonction $f(\delta, x, x_0) = (x - x_0)^i$ pour $i = 1, \dots, 4$, ce qui dans ce cas conduit à des expressions explicites. Nous avons alors:*

$$\mathbb{E}[Y_\Delta] = \mathbb{E}[X_\Delta - X_0] = \Delta(\mu + \beta\lambda)$$

et en notant:

$$M(\Delta, r) = \mathbb{E}[(Y_\Delta - \mathbb{E}[Y_\Delta])^r], \quad r > 1$$

il vient:

$$M(\Delta, 2) = \Delta(\sigma^2 + (\beta^2 + \eta)\lambda)$$

$$M(\Delta, 3) = \Delta\lambda\beta(\beta^2 + 3\eta)$$

$$M(\Delta, 4) = \Delta(\beta^4\lambda + 6\beta^2\eta\lambda + 3\eta^2\lambda) + 3\Delta^2(\sigma^2 + (\beta^2 + \eta)\lambda)^2$$

5.2. Quelques Exemples

5.2.1 Mouvement Brownien Géométrique

Soit le processus stochastique $(X_t)_{t \geq 0}$ solution de:

$$dX_t = \mu X_t + \sigma X_t dW_t, \quad X_{t_0} = x_0 \quad (5.3)$$

où $(W_t)_{t \geq 0}$ est un mouvement brownien, $\mu \in \mathbb{R}$ et $\sigma > 0$ des constantes. Une solution à 5.3 est donnée par:

$$X_t = x_0 \exp \left(\left(\mu - \frac{1}{2} \sigma^2 \right) (t - t_0) + \sigma (W_t - W_{t_0}) \right)$$

Et X_t admet comme distribution limite, pour tout $t \geq s$:

$$\ln X_t - \ln X_s \sim \mathcal{N} \left(\left(\mu - \frac{1}{2} \sigma^2 \right) (t - s), \sigma^2 (t - s) \right)$$

Les paramètres $\theta = (\mu, \sigma)$ de cette distribution peuvent alors être estimées soit par maximum de vraisemblance, soit par la méthode des moments généralisés.

Maximum de Vraisemblance

Si nous considérons la suite des observations $(x_t, t = 1 \dots T)$ du processus X_t entre les dates $[1, T]$ comme une série temporelle dont les observations sont mesurées à intervalle de temps constant Δ , la log-vraisemblance pour l'observation t est alors donnée par:

$$l_t = -\frac{1}{2} \ln(2\pi) - \frac{1}{2} \ln(\sigma^2 \Delta) - \frac{1}{2} \frac{\left[\ln \frac{x_t}{x_{t-1}} - \left(\mu - \frac{1}{2} \sigma^2 \right) \Delta \right]^2}{\sigma^2 \Delta}$$

Remarque 5.2.1 (Pas de discrétisation Δ) *Le choix du pas de temps Δ revêt une importance capitale dans la procédure d'estimation. S'il est trop élevé, il va entraîner un biais dû à un manque de précision dans la description d'un processus continu par son équivalent discret. S'il est trop petit, un problème de consistance de l'estimateur de θ peut être rencontré (effet de bruit). En général, Δ est choisi selon des considérations liées aux données à disposition. Ainsi, si les données sont mensuelles, $\Delta = 1/12$, trimestrielles $\Delta = 1/4$, quotidiennes $\Delta = 1/252$ (nombre de jours de trading).*

La Figure 5.1 permet de juger de l'effet du pas de discrétisation sur la *qualité* d'une trajectoire discrète par rapport à son équivalent continu. Il représente un mouvement brownien géométrique avec $x_0 = 10, \mu = 0.8, \sigma = 0.02$, pour $\Delta = 1/(N^* - 1)$. Cette Figure est intéressante car elle permet de juger de deux choses: d'une part du phénomène de passage du temps continu au temps discret et d'autre part de la convergence de la trajectoire discrète vers la trajectoire continue.

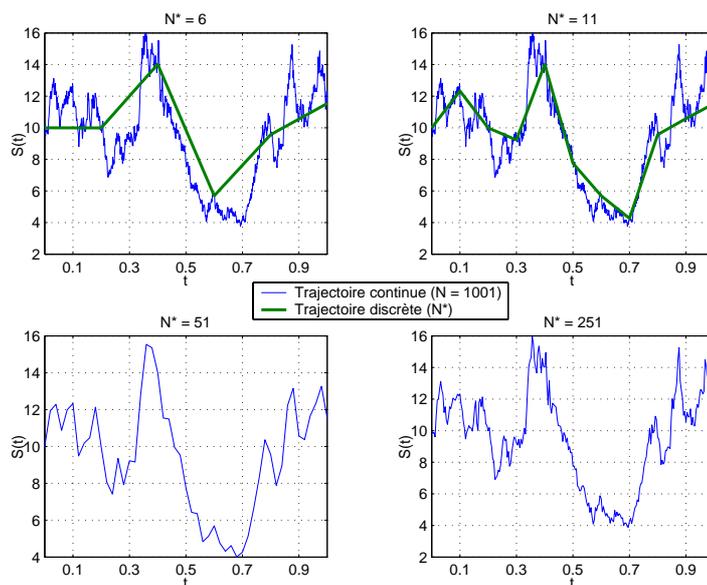


Figure 5.1: Passage du temps continu au temps discret pour un MBG.

Méthode des Moments Généralisés

Nous formons le vecteur des moments théoriques à partir des moments d'ordre 1 et 2 de X . A cette fin, notons:

$$w_t = \ln \frac{x_t}{x_{t-1}} - \left(\mu - \frac{1}{2} \sigma^2 \right) \Delta$$

et on calcule alors:

$$\mathbb{E}[w_t] = 0, \quad \mathbb{E}[w_t^2 - \sigma^2 \Delta] = 0$$

Remarque 5.2.2 (Nombre de moments) *Il n'existe pas de règles a priori pour le nombre de moments à prendre en compte dans une estimation MMG. A minima, les moments d'ordre 1 et 2 doivent être pris en compte. Et s'il est possible de disposer des moments jusqu'à des ordres élevés, il ne peut être que profitable pour la précision de l'estimation de les prendre en compte. En effet, au plus nous disposons de conditions de moments sur le processus, au mieux nous capturons sa description statistique, ce qui in extenso améliore l'estimation.*

5.2.2 Processus d'Ornstein-Uhlenbeck

Un processus $(X_t)_{t \geq 0}$ est dit suivre l'équation d'Ornstein-Uhlenbeck s'il vérifie:

$$dX_t = (a - bX_t) dt + \sigma dW_t, \quad X_{t_0} = x_0 \quad (5.4)$$

où $(W_t)_{t \geq 0}$ est un mouvement brownien et $a, b, \sigma > 0$ sont des constantes. Le processus 5.4 admet comme solution:

$$X_t = x_0 e^{-a(t-t_0)} + b \left(1 - e^{-a(t-t_0)}\right) + \sigma \int_{t_0}^t e^{-a(s-t)} dW_s$$

Et alors le processus X est du type gaussien tel que pour $t \geq s$:

$$X_t - e^{-a(t-s)} X_s \sim \mathcal{N} \left(b \left(1 - e^{-a(t-s)}\right), \sigma^2 \frac{1 - e^{-2a(t-s)}}{2a} \right)$$

A nouveau, nous pouvons estimer les paramètres $\theta = (a, b, \sigma)$ soit par le maximum de vraisemblance, soit par la méthode des moments.

Maximum de Vraisemblance

Si nous considérons la suite des observations $(x_t, t = 1 \dots T)$ du processus X_t entre les dates $[1, T]$ comme une série temporelle dont les observations sont mesurées à intervalle de temps constant Δ , la log-vraisemblance pour l'observation t est alors donnée par:

$$l_t = -\frac{1}{2} \ln(2\pi) - \frac{1}{2} \ln \left(\sigma^2 \frac{1 - e^{-2a\Delta}}{2a} \right) - \frac{1}{2} \frac{[x_t - e^{-a\Delta} x_{t-1} - b(1 - e^{-a\Delta})]^2}{\sigma^2 \frac{1 - e^{-2a\Delta}}{2a}}$$

La Figure 5.2 représente les fonctions de densité des estimateurs du maximum de vraisemblance pour $T = 1000$. Les valeurs des paramètres sont $a = 0.8$, $b = 0.1$ et $\sigma = 0.02$.

Nous observons l'influence de la valeur de Δ (noté ici k) sur la vitesse de convergence des estimateurs, surtout pour le coefficient b . Toutefois, il n'existe pas de relation monotone entre la valeur de k et la vitesse de convergence. En effet, nous obtenons de très bons résultats pour σ quand $k = 0.1$, alors que ceux-ci sont très mauvais pour le paramètre a .

Ainsi, dans le cas du processus OU, une valeur faible de k n'améliore pas forcément l'estimation des paramètres, contrairement à ce que nous aurions pu nous attendre. Une autre explication réside dans le fait que la plus faible valeur de k que nous avons testé, à savoir 0.05, est peut-être encore trop élevée pour améliorer sensiblement les estimations.

Méthode des Moments Généralisés

Si nous posons:

$$w_t = x_t - e^{-a\Delta} x_{t-1} - b(1 - e^{-a\Delta})$$

alors:

$$\mathbb{E}[w_t] = 0, \mathbb{E} \left[w_t^2 - \sigma^2 \frac{1 - e^{-2a\Delta}}{2a} \right] = 0, \mathbb{E}[w_t x_{t-1}] = 0$$

La Figure 5.3 correspond à la même étude de Monte-Carlo que celle reportée sur la Figure 5.2, mais pour la méthode des moments généralisés. Il semblerait que dans ce cas l'estimateur MMG ne soit pas plus efficace que l'estimateur MV.

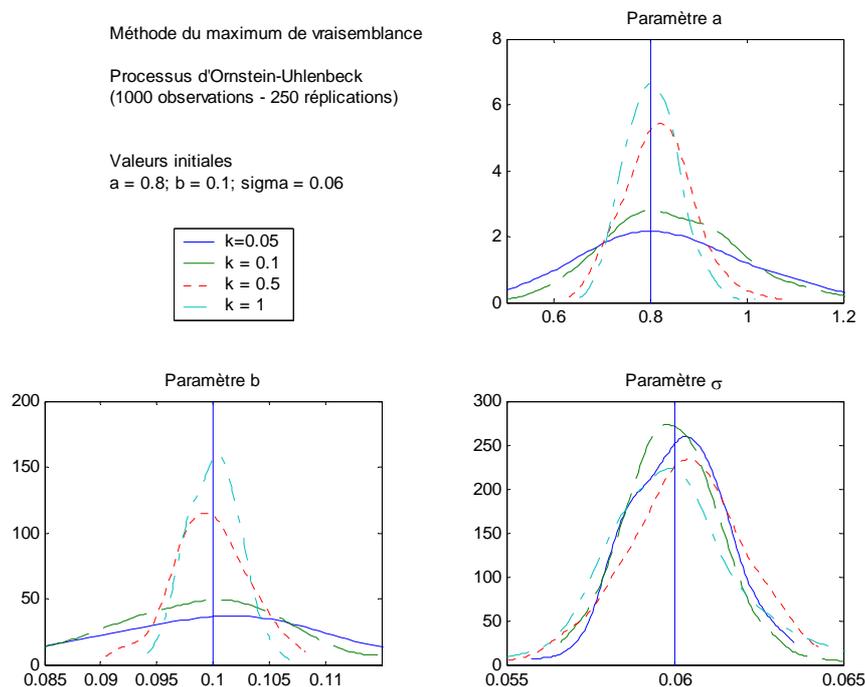


Figure 5.2: Densité empirique des estimateurs de la méthode du MV pour OU.

5.2.3 Equation Différentielle Stochastique

Considérons l'équation différentielle stochastique:

$$dX_t = \mu(t, \theta, X) dt + \sigma(t, \theta, X) dW_t, \quad X_{t_0} = x_0 \quad (5.5)$$

où $(W_t)_{t \geq 0}$ est un mouvement brownien, $\mu(t, X)$ et $\sigma(t, X)$ des fonctions non-anticipatives et $\theta \in \Theta$ un vecteur de paramètres dont dépendent les fonctions μ et σ . Dans le cadre général, 5.5 n'admet pas de solution explicite. Il est alors nécessaire de recourir à un schéma de discrétisation, cf. Annexe C par exemple le schéma d'Euler:

$$X_{t+1} = X_t + \mu(t, \theta, X) \Delta + \sigma(t, \theta, X) \sqrt{\Delta} \varepsilon_{t+1}$$

où $\varepsilon_{t+1} \sim \mathcal{N}(0, 1)$ et Δ le pas de discrétisation. Il est à nouveau possible de s'intéresser à l'estimation des paramètres θ en ayant recours soit au maximum de vraisemblance ou la méthode des moments généralisés.

Maximum de Vraisemblance

Si nous considérons la suite des observations $(x_t, t = 1 \dots T)$ du processus X_t entre les dates $[1, T]$ comme une série temporelle dont les observations sont

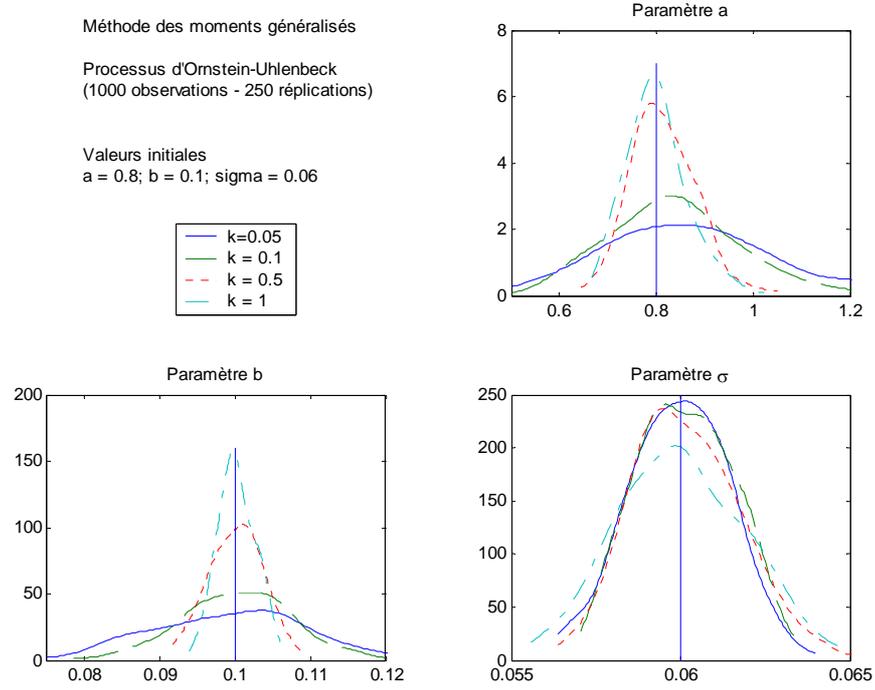


Figure 5.3: Densité empirique des estimateurs de la MMG pour OU.

mesurées à intervalle de temps constant Δ , la log-vraisemblance approchée pour l'observation t est alors donnée par:

$$l_t = -\frac{1}{2} \ln(2\pi) - \frac{1}{2} \ln(\sigma^2(t, \theta, x) \Delta) - \frac{1}{2} \frac{[x_{t+1} - x_t - \mu(t, \theta, x) \Delta]^2}{\sigma^2(t, \theta, x) \Delta}$$

Méthode des Moments Généralisés

Nous introduisons:

$$w_{t+1} = x_{t+1} - x_t - \mu(t, \theta, x) \Delta$$

Et nous formons alors le vecteur suivant de conditions de moments:

$$\begin{cases} \mathbb{E}[w_{t+1}] = 0 \\ \mathbb{E}[w_{t+1}^2 - \sigma^2(t, \theta, x) \Delta] = 0 \\ \mathbb{E}[w_{t+1} x_t] = 0 \\ \mathbb{E}[(w_{t+1}^2 - \sigma^2(t, \theta, x) \Delta) x_t] = 0 \\ \vdots \end{cases}$$

6

Filtrage Linéaire Gaussien

6.1. Contexte

Sur l'espace de probabilité (Ω, \mathcal{F}, P) , nous définissons le couple de processus stochastiques $(Y_t, X_t)_{t \geq 0}$ où Y_t est observé, tandis que X_t est non-observable; on dit encore latent. A chaque date $t \geq 0$, nous cherchons à produire une estimation de X_t à partir des observations de Y_t relevées jusqu'en t , soit $\mathcal{F}_t^Y = \sigma(Y_s, s \in [0, t])$. Ce problème d'estimation de X à partir de Y revêt le vocable de *filtrage*.

Il est bien connu, en statistiques, que la meilleure estimation, au sens des moindres carrées, de X_t depuis $(Y_s, s \in [0, t])$ est la moyenne a posteriori $m_t = \mathbb{E}[X_t | \mathcal{F}_t^Y]$. De fait, le problème du filtrage se ramène au calcul de l'espérance conditionnelle:

$$m_t = \mathbb{E}[X_t | \mathcal{F}_t^Y]$$

6.2. Filtrage en Temps Discret

6.2.1 Cadre

Sur l'espace de probabilité (Ω, \mathcal{F}, P) , nous considérons le couple partiellement observé de processus stochastiques $(Y, X) = (Y_t, X_t)_{t \geq 0}$ (ou $t = 0, \Delta, 2\Delta, \dots$) où Δ représente le pas de temps) qui satisfait au système récursif d'équations:

$$Y_{t+1} = A_0(t, Y) + A_1(t, Y) X_t + B_1(t, Y) \varepsilon_{t+1}^1 + B_2(t, Y) \varepsilon_{t+1}^2 \quad (6.1)$$

$$X_{t+1} = a_0(t, Y) + a_1(t, Y) X_t + b_1(t, Y) \varepsilon_{t+1}^1 + b_2(t, Y) \varepsilon_{t+1}^2 \quad (6.2)$$

où ε_t^1 et ε_t^2 sont des variables aléatoires $\mathcal{N}(0, 1)$ i.i.d. Nous notons $\mathcal{F}_t^Y = \sigma(Y_s, s \in [0, t])$ pour $t \geq 0$ la σ -algèbre générée par la variable aléatoire Y . En outre les fonctions A_i, B_i, a_i, b_i pour $i = (0, 1)$ sont non-anticipatives, c'est-à-dire \mathcal{F}^Y -mesurables.

La théorie du filtrage consiste à produire une estimation de la variable non observable X_t à partir des observations $Y_0^t = (Y_0, \dots, Y_t)$. Soit la distribution

de probabilité $F_{Y_0^t}(x) = P(X_t \leq x | \mathcal{F}_t^Y)$. Nous cherchons à produire une estimation optimale, dans le sens des moindres carrés, de X_t , qui est alors donnée par l'espérance conditionnelle:

$$m_t = \mathbb{E}[X_t | \mathcal{F}_t^Y] \quad (6.3)$$

et dont la variance est égale à:

$$\gamma_t = \mathbb{E}[(X_t - m_t)^2 | \mathcal{F}_t^Y] \quad (6.4)$$

Sous l'hypothèse de normalité de la distribution a priori $P(X_0 | Y_0)$ de X_0 et de la distribution jointe du couple (Y, X) , le problème du filtrage admet une solution.

6.2.2 Normalité A Priori-A Posteriori

La distribution de X_0 est supposée être gaussienne. Nous allons à présent montrer que la distribution a posteriori de X_t , soit $F_{Y_0^t}(x) = P(X_t \leq x | \mathcal{F}_t^Y)$, est gaussienne, de moyenne m_t et de variance γ_t ; elle se note alors $X_t \sim \mathcal{N}(m_t, \gamma_t)$.

Théorème 6.2.1 (Normalité a priori - a posteriori) *Sous 6.1-6.2, la séquence (Y, X) est conditionnellement gaussienne et alors les distributions conditionnelles:*

$$P(X_0 \leq x_0, \dots, X_t \leq x_t | \mathcal{F}_t^Y)$$

sont gaussiennes pour tout $t = 0, 1, \dots$

Preuve. Nous allons uniquement montrer la normalité de $P(X_t \leq a_t | \mathcal{F}_t^Y)$. La preuve se fera par récurrence. Nous supposons ainsi que la distribution $F_{Y_0^t}(x_t) = P(X_t \leq x_t | \mathcal{F}_t^Y)$ est gaussienne telle que $X_t \sim \mathcal{N}(m_t, \gamma_t)$. Sous 6.1-6.2, la distribution conditionnelle:

$$P(Y_{t+1} \leq y, X_{t+1} \leq x | \mathcal{F}_t^Y | X_t = z)$$

est gaussienne, de vecteur d'espérance:

$$\mathbf{A}_0 + \mathbf{A}_1 X_t = \begin{pmatrix} A_0 + A_1 X_t \\ a_0 + a_1 X_t \end{pmatrix}$$

et de matrice de variance-covariance:

$$\mathbf{B} = \begin{pmatrix} B & c \\ c & b \end{pmatrix} = \begin{pmatrix} B_1^2 + B_2^2 & B_1 b_1 + B_2 b_2 \\ B_1 b_1 + B_2 b_2 & b_1^2 + b_2^2 \end{pmatrix}$$

Soit $\nu_t = (Y_t, X_t)$ de fonction caractéristique conditionnelle (cf. Partie I, Section 2.1):

$$\mathbb{E}(\exp[iu^\top \nu_{t+1}] | \mathcal{F}_t^Y) = \exp\left[iu^\top (\mathbf{A}_0 + \mathbf{A}_1 X_t) - \frac{1}{2} u^\top \mathbf{B} u | \mathcal{F}_t^Y\right] \quad (6.5)$$

Comme nous avons supposé que X_t était gaussien $\mathcal{N}(m_t, \gamma_t)$, nous avons:

$$\mathbb{E}(\exp[iu^\top \mathbf{A}_1 X_t] | \mathcal{F}_t^Y) = \exp\left[iu^\top \mathbf{A}_1 m_t - \frac{1}{2} u^\top (\mathbf{A}_1 \gamma_t \mathbf{A}_1^\top) u\right]$$

ce qui combiné avec 6.5 donne:

$$\mathbb{E}(\exp[iu^\top \nu_{t+1}] | \mathcal{F}_t^Y) = \exp\left[iu^\top (\mathbf{A}_0 + \mathbf{A}_1 m_t) - \frac{1}{2} u^\top \mathbf{B} u - \frac{1}{2} u^\top (\mathbf{A}_1 \gamma_t \mathbf{A}_1^\top) u\right]$$

Et par récurrence, les distributions conditionnelles:

$$P(Y_{t+1} \leq x, X_{t+1} \leq a | \mathcal{F}_t^Y) \quad (6.6)$$

sont gaussiennes. Nous considérons à présent le vecteur:

$$\eta = [X_{t+1} - \mathbb{E}(X_{t+1} | \mathcal{F}_t^Y)] - C [Y_{t+1} - \mathbb{E}(Y_{t+1} | \mathcal{F}_t^Y)] \quad (6.7)$$

pour une matrice C quelconque. Par 6.6, il est conditionnellement gaussien. Il s'ensuit alors par 6.7 que:

$$\begin{aligned} & \mathbb{E}(\exp[iu^\top X_{t+1}] | \mathcal{F}_t^Y, Y_{t+1}) \\ &= \mathbb{E}(\exp(iu^\top [\mathbb{E}(X_{t+1} | \mathcal{F}_t^Y) + C(Y_{t+1} - \mathbb{E}(Y_{t+1} | \mathcal{F}_t^Y)) + \eta]) | \mathcal{F}_t^Y, Y_{t+1}) \\ &= \exp(iu^\top [\mathbb{E}(X_{t+1} | \mathcal{F}_t^Y) + C(Y_{t+1} - \mathbb{E}(Y_{t+1} | \mathcal{F}_t^Y))]) \mathbb{E}(\exp[iu^\top \eta | \mathcal{F}_t^Y, Y_{t+1}]) \end{aligned}$$

est conditionnellement gaussien et donc la distribution conditionnelle $P(X_{t+1} \leq x | \mathcal{F}_t^Y)$ est gaussienne, ce qui conclut la preuve. ■

Après avoir démontré la normalité de la distribution a posteriori de X_t , nous allons à présent dériver une expression pour l'espérance et la variance conditionnelle de $(X|Y, \mathcal{F}_t^Y)$. D'un point de vue du filtrage, ces équations prennent en compte l'effet de l'apprentissage et traduisent aussi de combien les estimations de X s'améliorent lorsque le temps passe avec la disponibilité de nouvelles informations.

Théorème 6.2.2 (Normal correlation) *Soit la séquence gaussienne (Y, X) . La distribution conditionnelle de (Y, X) est gaussienne, d'espérance et de variance-covariance:*

$$m_Y = \mathbb{E}(Y), \quad m_X = \mathbb{E}(X)$$

$$D_{YY} = \text{cov}(Y, Y), \quad D_{YX} = \text{cov}(Y, X), \quad D_{XX} = \text{cov}(X, X)$$

Alors l'espérance et la variance de $X|Y$ sont données par:

$$\mathbb{E}(X|Y) = \mathbb{E}(X) + D_{YX} (\gamma_{YY})^{-1} [Y - \mathbb{E}(Y)]$$

$$\text{cov}(X, X|Y) = D_{XX} - D_{YX} (D_{YY})^{-1} (D_{YX})^\top$$

Preuve. Soit le vecteur:

$$\eta = (X - m_X) + C(Y - m_Y) \quad (6.8)$$

où la matrice C est choisie telle que $\mathbb{E}[\eta(Y - m_Y)] = 0$, ce qui assure l'indépendance conditionnelle de η et de Y . Si une telle matrice existe, alors par la propriété d'indépendance gaussienne, i.e.: $\text{cov}(Y, \eta) = 0$, elle vérifie:

$$D_{YX} + CD_{YY} = 0$$

Et si D_{YY} est une matrice définie positive, soit inversible, alors:

$$C = -D_{YX}(D_{YY})^{-1}$$

Le vecteur $\eta = (Y, X)$ est gaussien et il en sera alors de même de (η, Y) , du fait de l'indépendance. Sa fonction caractéristique conditionnelle est donnée par:

$$\begin{aligned} & \mathbb{E}[\exp(iu_1^\top \eta + iu_2^\top Y) | Y] \\ &= \mathbb{E}[\exp(iu_1^\top [(X - m_X) + C(Y - m_Y)] + iu_2^\top Y) | Y] \end{aligned} \quad (6.9)$$

Et du fait de la normalité du couple (Y, X) , 6.9 entraîne que les composantes de (η, Y) sont indépendantes et donc $\mathbb{E}(\eta|Y) = \mathbb{E}(\eta) = 0$, ce qui avec 6.8 donne:

$$\mathbb{E}(\eta|Y) = (\mathbb{E}(X|Y) - m_X) - D_{YX}(D_{YY})^{-1}(Y - m_Y) = 0$$

ce qui est le résultat annoncé pour la moyenne. Finalement, notant que $X - \mathbb{E}(X|Y) = \eta$ et du fait de l'indépendance entre Y et η , nous avons:

$$\text{cov}(X, X|Y) = \mathbb{E}(\eta\eta^\top | Y) = \mathbb{E}(\eta\eta^\top)$$

ce qui conclut la preuve. ■

La version conditionnelle de ce Théorème est évidente. Soit le couple (Y, X) conditionnellement gaussien sur l'espace de probabilité (Ω, \mathcal{F}, P) et soit \mathcal{F}^Y une sous σ -algèbre de \mathcal{F} générée par les observations (Y_0, \dots, Y_t) . Si le couple (Y, X) est gaussien d'espérances $\mathbb{E}[Y|\mathcal{F}^Y]$ et $\mathbb{E}[X|\mathcal{F}^Y]$ et de covariances $d_{11} = \text{cov}(X, X|\mathcal{F}^Y)$, $d_{12} = \text{cov}(X, Y|\mathcal{F}^Y)$ et $d_{22} = \text{cov}(Y, Y|\mathcal{F}^Y)$, alors le Théorème 6.2.2 nous donne:

$$\mathbb{E}[X|Y, \mathcal{F}^Y] = \mathbb{E}[X|\mathcal{F}^Y] + d_{12}d_{22}^{-1}[Y - \mathbb{E}[Y|\mathcal{F}^Y]]$$

$$\text{cov}(X, X|Y, \mathcal{F}^Y) = d_{11} - d_{12}^2 d_{22}^{-1}$$

6.2.3 Equations Récursives

Nous pouvons à présent déduire le système récursif d'équations de filtrage pour m_t et γ_t .

Théorème 6.2.3 (Filtre - Discret) Les paramètres m_t et γ_t définis par 6.3-6.4 satisfont, pour $t \geq 0$, le système:

$$m_{t+1} = [a_0(t, Y) + a_1(t, Y) m_t] + [c(t, Y) + a_1(t, Y) \gamma_t A_1(t, Y)] \\ \times [B(t, Y) + A_1^2(t, Y) \gamma_t]^{-1} [Y_{t+1} - A_0(t, Y) - A_1(t, Y) m_t] \quad (6.10)$$

$$\gamma_{t+1} = [a_1(t, Y) \gamma_t A_1(t, Y) + b(t, Y)] - [c(t, Y) + a_1(t, Y) \gamma_t A_1(t, Y)]^2 \\ \times [B(t, Y) + A_1^2(t, Y) \gamma_t]^{-1} \quad (6.11)$$

avec:

$$B(t, Y) = B_1^2(t, Y) + B_2^2(t, Y) \\ b(t, Y) = b_1^2(t, Y) + b_2^2(t, Y) \\ c(t, Y) = B_1(t, Y) b_1(t, Y) + B_2(t, Y) b_2(t, Y)$$

Preuve. La distribution conditionnelle de (Y_{t+1}, X_{t+1}) est gaussienne de vecteur d'espérance donné par:

$$m_Y = \mathbb{E}(Y_{t+1} | \mathcal{F}_t^Y) = A_0 + A_1 m_t \\ m_X = \mathbb{E}(X_{t+1} | \mathcal{F}_t^Y) = a_0 + a_1 m_t$$

Soit, en utilisant 6.1-6.2 et les équations précédentes, nous avons:

$$Y_{t+1} - \mathbb{E}(Y_{t+1} | \mathcal{F}_t^Y) = A_1 [X_t - m_t] + B_1 \varepsilon_1(t+1) + B_2 \varepsilon_2(t+1) \\ X_{t+1} - \mathbb{E}(X_{t+1} | \mathcal{F}_t^Y) = a_1 [X_t - m_t] + b_1 \varepsilon_1(t+1) + b_2 \varepsilon_2(t+1)$$

ce qui permet de dériver la matrice de variance-covariance:

$$D_{YY} = \text{cov}(Y_{t+1}, Y_{t+1} | \mathcal{F}_t^Y) = B + A_1^2 \gamma_t \\ D_{YX} = \text{cov}(Y_{t+1}, X_{t+1} | \mathcal{F}_t^Y) = c + a_1 \gamma_t A_1 \\ D_{XX} = \text{cov}(X_{t+1}, X_{t+1} | \mathcal{F}_t^Y) = b + a_1^2 \gamma_t$$

En outre, grâce au Théorème 6.2.2, nous avons:

$$\mathbb{E}(X_{t+1} | Y_{t+1}, \mathcal{F}_t^Y) = \mathbb{E}(X_{t+1} | \mathcal{F}_t^Y) + D_{YX} (D_{YY})^{-1} [Y_{t+1} - \mathbb{E}(Y_{t+1}, \mathcal{F}_t^Y)] \\ \text{cov}(X_{t+1}, X_{t+1} | Y_{t+1}, \mathcal{F}_t^Y) = D_{XX} - D_{YX} (D_{YY})^{-1}$$

Soit, par substitution, nous obtenons les équations récurrentes annoncées. ■

Nous présentons par ailleurs les équations précédentes pour un cadre légèrement différent. Soit le couple partiellement observée (Y, X) satisfaisant le système suivant:

$$Y_t = \tilde{A}_0(t-1, Y) + \tilde{A}_1(t-1, Y) X_t + \tilde{B}_1(t-1, Y) \varepsilon_1(t) + \tilde{B}_2(t-1, Y) \varepsilon_2(t) \quad (6.12)$$

$$X_{t+1} = a_0(t, Y) + a_1(t, Y) X_t + b_1(t, Y) \varepsilon_1(t+1) + b_2(t, Y) \varepsilon_2(t+1) \quad (6.13)$$

Les équations récurrentes 6.10-6.11 ont été établies pour un vecteur (Y_{t+1}, X_{t+1}) , ce qui n'est pas la forme satisfaite par 6.12-6.13. Néanmoins, par le changement de variable:

$$Y_{t+1} = \tilde{A}_0(t, Y) + \tilde{A}_1(t, Y) [a_0(t, Y) + a_1(t, Y) X_t + b_1(t, Y) \varepsilon_1(t+1) + b_2(t, Y) \varepsilon_2(t+1)] + \tilde{B}_1(t, Y) \varepsilon_1(t+1) + \tilde{B}_2(t, Y) \varepsilon_2(t+1)$$

et en posant:

$$\begin{aligned} A_0 &= \tilde{A}_0 + \tilde{A}_1 a_0 & A_1 &= \tilde{A}_1 a_1 \\ B_1 &= \tilde{A}_1 b_1 + \tilde{B}_1 & B_2 &= \tilde{A}_1 b_2 + \tilde{B}_2 \end{aligned} \quad (6.14)$$

le couple (Y, X) satisfait les équations 6.1-6.2 sous le changement de temps 6.14 et (m_t, γ_t) satisfait alors aux équations de filtrage 6.10-6.11.

Remarque 6.2.4 (Système état-mesure) *Le système d'équations 6.12-6.13 est parfois qualifié de représentation état-mesure (state-space). Et les équations de filtrage associées sont connues comme étant les équations ou filtre de Kalman-Bucy (1969).*

Nous concluons par un résultat utile pour la suite. Le théorème suivant donne une représentation alternative de Y_t pour $t \geq 0$, cf. 6.1, en fonction des processus (m_t, γ_t) .

Théorème 6.2.5 (Innovations) *Il existe un processus gaussien $\bar{\varepsilon}_t$ pour $t \geq 0$ tel que:*

$$\mathbb{E}[\bar{\varepsilon}_t] = 0, \quad \mathbb{E}[\bar{\varepsilon}_t \bar{\varepsilon}_s] = t - s$$

et alors:

$$Y_{t+1} = A_0(t, Y) + A_1(t, Y) m_t + [B(t, Y) + A_1^2(t, Y) \gamma_t]^{1/2} \bar{\varepsilon}_{t+1} \quad (6.15)$$

Preuve. Soit:

$$\begin{aligned} \bar{\varepsilon}_{t+1} &= [B(t, Y) + A_1^2(t, Y) \gamma_t]^{-1/2} \\ &\times [A_1(t, Y) (X_t - m_t) + B_1(t, Y) \varepsilon_{t+1}^1 + B_2(t, Y) \varepsilon_{t+1}^2] \end{aligned} \quad (6.16)$$

Il est alors possible de montrer que la distribution conditionnelle $P(\bar{\varepsilon}_{t+1} \leq x | \mathcal{F}_t^Y)$ est gaussienne et de vérifier les conditions de moment pour $\bar{\varepsilon}_t$, ce qui conclut la preuve. ■

Il est par ailleurs possible de montrer l'égalité suivante:

$$\mathcal{F}_t^Y = \mathcal{F}_t^{Y_0, \bar{\varepsilon}}, \quad t \geq 0$$

Par suite, le processus 6.16 est naturellement appelé le processus des innovations (à comparer avec la définition donnée en Section 6.3.2).

6.2.4 Interpolation: Lissage

Pour le couple partiellement observé (Y, X) , l'interpolation est le problème de construire une estimation optimale, au sens des moindres carrés, de X_s à partir des observations $Y_0^t = (Y_0, \dots, Y_t)$ pour $t \geq s$. A cette fin, nous définissons:

$$m(s, t) = \mathbb{E} [X_s | \mathcal{F}_t^Y], \quad \gamma(s, t) = \mathbb{E} [(X_s - m(s, t))^2 | \mathcal{F}_t^Y]$$

qui sont respectivement l'espérance et la variance de la distribution $P(X_s \leq x | \mathcal{F}_t^Y)$ pour $t \geq s$. Ces quantités permettent de juger de combien l'interpolation (sur t quand s est fixé) améliore la précision des estimations quand de nouvelles données sont disponibles.

Nous introduisons par ailleurs pour $t \geq s$ et $\varphi_s^s = 1$:

$$\begin{aligned} \varphi_s^t &= \prod_{u=s}^{t-1} \{a_1(t, Y) - [c(t, Y) + a_1(t, Y) \gamma(u, s) A_1(t, Y)] \\ &\quad \times [B(t, Y) + A_1^2(t, Y) \gamma(u, s)]^{-1} A_1(t, Y)\} \end{aligned}$$

Nous pouvons à présent dériver les équations d'interpolation (ou de lissage) satisfaites par $m(s, t)$ et $\gamma(s, t)$ sous l'hypothèse que $P(X_s \leq x | \mathcal{F}_s^Y)$ est gaussienne $\mathcal{N}(m_s, \gamma_s)$.

Théorème 6.2.6 (Lisseur - Discret) *Sous 6.1-6.2, si $P(X_s \leq x | \mathcal{F}_s^Y) \sim \mathcal{N}(m_s, \gamma_s)$, alors $m(s, t)$ et $\gamma(s, t)$ pour $t \geq s$ vérifient:*

$$\begin{aligned} m(s, t+1) &= m(s, t) + \gamma(s, t) \varphi_s^t A_1(t, Y) [B(t, Y) + A_1^2(t, Y) \gamma_t]^{-1} \\ &\quad \times [Y_{t+1} - A_0(t, Y) - A_1(t, Y) m_t] \end{aligned} \quad (6.17)$$

$$\begin{aligned} \gamma(s, t+1) &= \gamma(s, t) - \gamma(s, t) \varphi_s^t A_1(t, Y) [B(t, Y) + A_1^2(t, Y) \gamma_t]^{-1} \\ &\quad \times [A_1(t, Y) \varphi_s^t \gamma(s, t)] \end{aligned} \quad (6.18)$$

avec $m(t, t) = m_t$ et $\gamma(t, t) = \gamma_t$.

Preuve. Soit la distribution conditionnelle $P(X_s \leq x, Y_s \leq y | \mathcal{F}_{t-1}^Y)$ gaussienne. Les paramètres de cette distribution peuvent être obtenues par le Théorème 6.2.2. Il vient:

$$m(s, t) = \mathbb{E} [X_s | Y_t, \mathcal{F}_{t-1}^Y] = \mathbb{E} [X_s | \mathcal{F}_{t-1}^Y] + d_{12} d_{22}^{-1} [Y_t - \mathbb{E} [Y_t | \mathcal{F}_{t-1}^Y]]$$

où:

$$\begin{aligned} d_{12} &= \text{cov}(X_s, Y_t | \mathcal{F}_{t-1}^Y) = \gamma(s, t-1) \varphi_s^{t-1} A_1(t-1, Y) \\ d_{22} &= \text{cov}(Y_t, Y_t | \mathcal{F}_{t-1}^Y) = A_1^2(t-1, Y) \gamma_{t-1} + B(t-1, Y) \end{aligned}$$

Nous détaillons le calcul de d_{12} . Nous avons:

$$\mathbb{E} [Y_t | \mathcal{F}_{t-1}^Y] = A_0(t-1, Y) + A_1(t-1, Y) m_{t-1}$$

et:

$$\begin{aligned}
& \mathbb{E} [Y_t - \mathbb{E} [Y_t | \mathcal{F}_{t-1}^Y] | X_s, \mathcal{F}_{t-1}^Y] \\
&= \mathbb{E} \{ [A_1(t-1, Y)(X_{t-1} - m_{t-1}) \\
&+ B_1(t-1, Y)\varepsilon_t^1 + B_2(t-1, Y)\varepsilon_t^2] | X_s, \mathcal{F}_{t-1}^Y \} \\
&= \mathbb{E} [A_1(t-1, Y)(X_{t-1} - m_{t-1}) | X_s, \mathcal{F}_{t-1}^Y] \\
&= (X_s - m(s, t-1)) \varphi_s^{t-1} A_1(t-1, Y)
\end{aligned}$$

Par suite:

$$\begin{aligned}
d_{12} &= \text{cov}(X_s, Y_t | \mathcal{F}_{t-1}^Y) \\
&= \mathbb{E} [(X_s - m(s, t-1))(Y_t - \mathbb{E}[Y_t | \mathcal{F}_{t-1}^Y]) | \mathcal{F}_{t-1}^Y] \\
&= \mathbb{E} [(X_s - m(s, t-1))^2 \varphi_s^{t-1} A_1(t-1, Y) | \mathcal{F}_{t-1}^Y] \\
&= \gamma(s, t-1) \varphi_s^{t-1} A_1(t-1, Y)
\end{aligned}$$

En s'appuyant toujours sur le Théorème 6.2.2, nous avons:

$$\gamma(s, t) = \text{cov}(X_s, X_s | Y_t, \mathcal{F}_{t-1}^Y) = d_{11} - d_{12}^2 d_{22}^+$$

où:

$$d_{11} = \text{cov}(X_s, X_s | \mathcal{F}_{t-1}^Y) = \gamma(s, t-1)$$

ce qui conclut la preuve. ■

6.2.5 Extrapolation: Prédiction

L'extrapolation consiste en l'estimation des processus X_t et Y_t à partir des observations $Y_0^s = (Y_0, \dots, Y_s)$ pour $t > s$. D'un point de vue statistique, on parle plus généralement de prédiction. De 6.10 et 6.15, il vient:

$$\begin{aligned}
m_{t+1} &= [a_0(t, Y) + a_1(t, Y)m_t] + [c(t, Y) + a_1(t, Y)\gamma_t A_1(t, Y)] \\
&\times [B(t, Y) + A_1^2(t, Y)\gamma_t]^{-1/2} \bar{\varepsilon}_{t+1}
\end{aligned} \tag{6.19}$$

$$Y_{t+1} = A_0(t, Y) + A_1(t, Y)m_t + [B(t, Y) + A_1^2(t, Y)\gamma_t]^{1/2} \bar{\varepsilon}_{t+1} \tag{6.20}$$

Nous introduisons alors les quantités:

$$n_1(t, s) = \mathbb{E}[X_t | \mathcal{F}_s^Y], \quad n_2(t, s) = \mathbb{E}[Y_t | \mathcal{F}_s^Y], \quad t > s$$

qui sont les estimations optimales, au sens des moindres carrés, de X_t et Y_t suivant Y_0^s . Comme $n_1(t, s) = \mathbb{E}[\mathbb{E}[X_t | \mathcal{F}_t^Y] | \mathcal{F}_s^Y] = \mathbb{E}[m_t | \mathcal{F}_s^Y]$ et $\mathbb{E}[\bar{\varepsilon}_{t+1} | \mathcal{F}_s^Y] = 0$, il est clair que les équations satisfaites par $n_1(t, s)$ et $n_2(t, s)$ peuvent être obtenues en prenant l'espérance conditionnelle $\mathbb{E}[\cdot | \mathcal{F}_s^Y]$ de chaque côté de 6.19 et 6.20.

Pour s'assurer que les distributions conditionnelles $P(X_t \leq x, Y_t \leq y | \mathcal{F}_s^Y)$ soient gaussiennes et de pouvoir calculer simultanément $n_1(t, s)$ et $n_2(t, s)$, il est nécessaire que:

$$a_0(t, Y) = a_0(t) + a_1(t) Y_t, \quad a_1(t, Y) = a_1(t) \quad (6.21)$$

$$A_0(t, Y) = A_0(t) + A_1(t) Y_t, \quad A_1(t, Y) = A_1(t) \quad (6.22)$$

Nous avons alors le théorème suivant, dit d'extrapolation.

Théorème 6.2.7 (Prédicteur - Discret) *Sous 6.19-6.20 et 6.21-6.22, $n_1(t, s)$ et $n_2(t, s)$, $t > s$, vérifient:*

$$n_1(t+1, s) = a_0(t) + a_1(t) n_1(t, s) + a_2(t) n_2(t, s) \quad (6.23)$$

$$n_2(t+1, s) = A_0(t) + A_1(t) n_1(t, s) + A_2(t) n_2(t, s) \quad (6.24)$$

avec $n_1(s, s) = m_s$ et $n_2(s, s) = Y_s$.

6.3. Filtrage en Temps Continu

De façon formelle, les équations de filtrage en temps continu peuvent être obtenus à partir des équations *discrètes* en passant à la limite, soit $\Delta \rightarrow 0$. Néanmoins, cette approche ne se trouve pas être suffisamment rigoureuse et nous dérivons dans la suite la preuve complète.

6.3.1 Cadre

Sur l'espace de probabilité (Ω, \mathcal{F}, P) , soit $(Y, X) = (Y_t, X_t)_{0 \leq t \leq T}$ un couple partiellement observé de processus stochastiques satisfaisant le système:

$$dY_t = [A_0(t, Y) + A_1(t, Y) X_t] dt + B(t, Y) dW_2(t) \quad (6.25)$$

$$dX_t = [a_0(t, Y) + a_1(t, Y) X_t] dt + b_1(t, Y) dW_1(t) + b_2(t, Y) dW_2(t) \quad (6.26)$$

où W_t^1 et W_t^2 sont des mouvements browniens indépendants et les fonctions A_i, B, a_i, b_i pour $i = (0, 1)$ sont non-anticipatives, c'est-à-dire \mathcal{F}^Y -mesurables, avec $\mathcal{F}_t^Y = \sigma(Y_s, s \in [0, t])$ la σ -algèbre générée par les observations de Y .

Le problème du filtrage est le même que dans le cas discret. Nous cherchons à produire une estimation de X_t à partir des observations $Y_0^t = (Y_0, \dots, Y_t)$, i.e.: \mathcal{F}_t^Y . A cette fin, nous considérons la distribution de probabilité $F_{Y_0^t}(x) = P(X_t \leq x | \mathcal{F}_t^Y)$ et dérivons une estimation optimale de X_t , au sens des moindres carrés, donnée par la moyenne a posteriori:

$$m_t = \mathbb{E}(X_t | \mathcal{F}_t^Y) \quad (6.27)$$

dont la variance est donnée par:

$$\gamma_t = \mathbb{E} \left[(X_t - m_t)^2 | \mathcal{F}_t^Y \right] \quad (6.28)$$

Le problème du filtrage repose largement sur l'hypothèse de distribution conditionnelle gaussienne du couple (Y, X) . En effet, sous l'hypothèse de normalité de la distribution conditionnelle a priori $P(X_0 \leq a | Y_0)$ de X_0 et de la distribution jointe de (Y, X) depuis 6.25-6.26, le problème du filtrage peut être résolu. En première étape, nous présentons certains résultats qui seront utiles dans la dérivation des équations de filtrage.

6.3.2 Structure des Innovations

Soit $Y = (Y_t, \mathcal{F}_t)$ ¹ un processus d'Itô satisfaisant:

$$dY_t = A_t(Y) dt + B_t(Y) dW_t \quad (6.29)$$

où les fonctions $A_t(Y)$ et $B_t(Y)$ sont non-anticipatives. Nous avons alors le théorème fondamental suivant, que nous donnons sans le démontrer.

Théorème 6.3.1 (Théorème de représentation des martingales) *Soit $X = (x_t, \mathcal{F}_t^Y)$ une martingale, i.e.: pour tout t , x_t est \mathcal{F}_t^Y -mesurable et pour tout $0 \leq s \leq t \leq T$, $\mathbb{E}[x_t | \mathcal{F}_s^Y] = x_s$. Alors X admet la représentation:*

$$x_t = x_0 + \int_0^t f_s(\omega) dW_s$$

où le processus $(f_s(\omega), \mathcal{F}_s^Y)$ est \mathcal{F}^Y -adapté.

Le théorème suivant met en avant le rôle central des processus d'innovations. En première analyse, le processus W est qualifié de processus d'innovation pour le processus Y si $\mathcal{F}_t^W = \mathcal{F}_t^Y$, c'est-à-dire si les filtrations (informations) engendrées par Y et W coïncident.

Théorème 6.3.2 *Soit $Y = (Y_t, \mathcal{F}_t)$ un processus d'Itô satisfaisant 6.29. Alors, il existe:*

1. Une fonction mesurable \bar{A} telle que:

$$\bar{A}_t(Y) = \mathbb{E}(A_t(\omega) | \mathcal{F}_t^Y)$$

2. Un processus de Wiener $\bar{W} = (\bar{W}_t, \mathcal{F}_t^Y)$ tel que:

$$B_t(Y) \bar{W}_t = Y_t - \int_0^t \bar{A}_s(Y) ds \quad (6.30)$$

Et le processus Y satisfait:

$$dY_t = \bar{A}_t(Y) dt + B_t(Y) d\bar{W}_t \quad (6.31)$$

¹La notation $Y = (Y_t, \mathcal{F}_t)$ se lit: le processus Y_t est \mathcal{F}_t -adapté.

Preuve. De 6.29 et 6.30, nous pouvons écrire que:

$$\overline{W}_t = W_t + \int_0^t B_s^{-1}(Y) [A_s(Y) - \overline{A}_s(Y)] ds \quad (6.32)$$

Par suite, une application de la formule d'Itô nous donne:

$$\begin{aligned} \exp[iu(\overline{W}_t - W_s)] &= 1 + iu \int_s^t \exp[iu(\overline{W}_u - W_s)] dW_u \\ &\quad + iu \int_s^t \exp[iu(\overline{W}_u - \overline{W}_s)] B_u^{-1}(Y) [A_u(Y) - \overline{A}_u(Y)] du \\ &\quad - \frac{u^2}{2} \int_s^t \exp[iu(\overline{W}_u - W_s)] du \end{aligned} \quad (6.33)$$

Et comme W est un mouvement brownien:

$$\mathbb{E} \left[\int_s^t \exp[iu(\overline{W}_u - W_s)] dW_u | \mathcal{F}_s^Y \right] = 0$$

Par suite:

$$\begin{aligned} &\mathbb{E} \left[\int_s^t \exp[iu(\overline{W}_u - \overline{W}_s)] B_u^{-1}(Y) [A_u(Y) - \overline{A}_u(Y)] du | \mathcal{F}_s^Y \right] \\ &= \mathbb{E} \left[\int_s^t \exp[iu(\overline{W}_u - \overline{W}_s)] B_u^{-1}(Y) \mathbb{E} [[A_u(Y) - \overline{A}_u(Y)] | \mathcal{F}_u^Y] du | \mathcal{F}_s^Y \right] = 0 \end{aligned}$$

Par passage à l'espérance conditionnelle de chaque côté de 6.33, nous avons:

$$\mathbb{E} [\exp[iu(\overline{W}_t - W_s)] | \mathcal{F}_s^Y] = 1 - \frac{u^2}{2} \int_s^t \mathbb{E} [\exp[iu(\overline{W}_u - W_s)] | \mathcal{F}_s^Y] du$$

qui admet comme solution:

$$\mathbb{E} [\exp[iu(\overline{W}_t - W_s)] | \mathcal{F}_s^Y] = \exp \left[-\frac{u^2}{2} (t - s) \right]$$

Ainsi, le processus $\overline{W} = (\overline{W}_t, \mathcal{F}_t^Y)$ est un mouvement brownien et 6.31 découle directement par substitution de 6.30, ce qui conclut la preuve. ■

De 6.29 et 6.31, nous pouvons écrire que:

$$\overline{W}_t = W_t + \int_0^t B_s^{-1}(Y) [A_s(Y) - \overline{A}_s(Y)] ds \quad (6.34)$$

est un (\mathcal{F}^Y, P) -mouvement brownien. Par ailleurs, la terminologie *processus d'innovations* apparaît plus clair. La quantité $[A_s(Y) - \overline{A}_s(Y)]$ représente la part d'information nouvelle, soit l'innovation, apportée par la connaissance d'une nouvelle observation du processus de mesure Y . Le processus \overline{W} est alors la version corrigée du processus non-observable W par notre connaissance plus fine de la tendance de Y .

Remarque 6.3.3 (Apprentissage de la tendance) *Le théorème précédent amène deux remarques importantes. D'une part, la théorie du filtrage gaussien se résume à un apprentissage de la tendance d'un processus d'Itô. D'autre part, le filtrage ne permet pas d'apprendre sur la volatilité du processus. Les raisons à cela sont de deux ordres. La première est que la variation quadratique d'un processus d'Itô est égale à sa "volatilité", cf. Section 3.2. Dès lors, il est possible d'estimer sans biais cette volatilité à partir d'observations, ce qui rend inutile l'apprentissage. La seconde raison est plus profonde et s'explique par les origines probabilistes de la théorie du filtrage. Sous l'approche dite de la probabilité de référence, un changement absolument continu de probabilité permet de dériver l'équation 6.34 satisfaite par le processus d'innovations. Or le théorème de Girsanov, s'il permet de modifier la tendance d'un processus d'Itô, ne permet pas d'en changer la volatilité, cf. Section 3.4.*

Nous caractérisons à présent la filtration engendrée par le processus des innovations \overline{W} .

Lemme 6.3.4 *Soit \overline{W}_t défini par 6.34, alors $\mathcal{F}^Y = \mathcal{F}^{Y_0, \overline{W}}$.*

Et nous concluons cette section en donnant un résultat qui sera important dans la suite de la présentation et qui est une application directe des Théorèmes 6.3.1 et 6.3.2.

Lemme 6.3.5 *Soit $X = (x_t, \mathcal{F}_t^Y)$ une martingale. Alors:*

$$x_t = x_0 + \int_0^t f_s(\omega) d\overline{W}_s$$

où le processus $(f_s(\omega), \mathcal{F}_s^Y)$ est \mathcal{F}^Y -adapté et \overline{W} est donné par 6.30 ou 6.34.

6.3.3 Résultat Général

Une solution de 6.25 est donnée par:

$$Y_t = Y_0 + \int_0^t A_s ds + \int_0^t B_s dW_2(s) \quad (6.35)$$

où $A_s = [A_0(s, Y) + A_1(s, Y) X_s]$. De même, le processus 6.26 admet comme solution:

$$X_t = X_0 + \int_0^t [a_0 + a_1 X_s] ds + \int_0^t [b_1 dW_1(s) + b_2 dW_2(s)] \quad (6.36)$$

Et si nous dénotons par:

$$h_t = X_t, \quad H_s = [a_0 + a_1 X_s], \quad x_t = \int_0^t [b_1 dW_1(s) + b_2 dW_2(s)]$$

Alors, 6.36 peut se réécrire comme:

$$h_t = h_0 + \int_0^t H_s ds + x_t \quad (6.37)$$

Nous donnons à présent le résultat général du filtrage, à savoir la dérivation de $\pi_t(h) = \mathbb{E}(h_t | \mathcal{F}_t^Y)$, dans le cadre de la théorie des innovations.

Théorème 6.3.6 Soit le couple partiellement observé (Y, h) , admettant la représentation donnée par 6.35-6.37. Pour $t \in [0, T]$, l'estimateur $\pi_t(h) = \mathbb{E}(h_t | \mathcal{F}_t^Y)$ de $h_t | \mathcal{F}_t^Y$ satisfait:

$$\begin{aligned} \pi_t(h) &= \pi_0(h) + \int_0^t \pi_s(H) ds \\ &+ \int_0^t \{ \pi_s(D) + B_s^- [\pi_s(hA) - \pi_s(h)\pi_s(A)] \} d\overline{W}_2(s) \end{aligned} \quad (6.38)$$

où:

$$\overline{W}_2(t) = \int_0^t B_s^{-1} [dY_s - \pi_s(A)] ds$$

et:

$$D_t = \frac{1}{dt} d[x, W]_t$$

Preuve. En prenant l'espérance conditionnelle de chaque côté de 6.37, nous avons:

$$\pi_t(h) = \mathbb{E}(h_0 | \mathcal{F}_t^Y) + \mathbb{E} \left(\int_0^t H_s ds | \mathcal{F}_t^Y \right) + \mathbb{E}(x_t | \mathcal{F}_t^Y) \quad (6.39)$$

Nous voulons à présent modifier la partie droite de 6.39. Pour cela, nous avons recours à une propriété de représentation de l'espérance conditionnelle comme une martingale de carré intégrable et à la structure des innovations (cf. Théorème 6.3.2 et 6.3.5). Grâce au Théorème 6.3.2, nous savons que 6.35 satisfait la représentation:

$$dY_t = \pi_t(A) dt + B_t(Y) d\overline{W}_2(t) \quad (6.40)$$

où $\overline{W}_2 = (\overline{W}_2(t), \mathcal{F}_t^Y)$ est un mouvement brownien. ■

Les lemmes suivants sont des applications directes du Lemme 6.3.5.

Lemme 6.3.7 Le processus $(\mathbb{E}(h_0 | \mathcal{F}_t^Y), \mathcal{F}_t^Y)_t$ est une martingale de carré intégrable de représentation:

$$\mathbb{E}(h_0 | \mathcal{F}_t^Y) = \pi_0(h) + \int_0^t g_s^h d\overline{W}_2(s)$$

Lemme 6.3.8 Le processus $(\mathbb{E}(x_t | \mathcal{F}_t^Y), \mathcal{F}_t^Y)_t$ est une martingale de carré intégrable de représentation:

$$\mathbb{E}(x_t | \mathcal{F}_t^Y) = \int_0^t g_s^x d\overline{W}_2(s)$$

Le lemme suivant permet d'interchanger les signes espérance et intégration.

Lemme 6.3.9 Soit α_t un processus stochastique à variations finies et \mathcal{F}^Y une sous σ -algèbre de \mathcal{F} . Alors:

$$\mathbb{E} \left(\int_0^t \alpha_s ds | \mathcal{F}^Y \right) = \int_0^t \mathbb{E}(\alpha_s | \mathcal{F}^Y) ds$$

Preuve. Soit λ une variable aléatoire bornée \mathcal{F}^Y -mesurable. Le théorème de Fubini nous donne:

$$\begin{aligned}\mathbb{E} \left[\lambda \int_0^t \alpha_s ds \right] &= \int_0^t \mathbb{E} [\lambda \alpha_s] ds = \int_0^t \mathbb{E} [\lambda \mathbb{E} [\alpha_s | \mathcal{F}^Y]] ds \\ &= \mathbb{E} \left[\lambda \int_0^t \mathbb{E} [\alpha_s | \mathcal{F}^Y] ds \right]\end{aligned}$$

Par ailleurs:

$$\mathbb{E} \left[\lambda \int_0^t \alpha_s ds \right] = \mathbb{E} \left[\lambda \mathbb{E} \left[\int_0^t \alpha_s ds | \mathcal{F}^Y \right] \right]$$

D'où:

$$\mathbb{E} \left[\lambda \int_0^t \mathbb{E} [\alpha_s | \mathcal{F}^Y] ds \right] = \mathbb{E} \left[\lambda \mathbb{E} \left[\int_0^t \alpha_s ds | \mathcal{F}^Y \right] \right]$$

ce qui conclut la preuve. ■

Lemme 6.3.10 *Le processus $\left(\mathbb{E} \left(\int_0^t H_s ds | \mathcal{F}_t^Y \right) - \int_0^t \pi_s(H) ds, \mathcal{F}_t^Y \right)_t$ est une martingale de carré intégrable de représentation:*

$$\mathbb{E} \left(\int_0^t H_s ds | \mathcal{F}_t^Y \right) - \int_0^t \pi_s(H) ds = \int_0^t g_s^H d\overline{W}_2(s)$$

Suite du Théorème 6.3.6. Grâce aux lemmes précédents, nous pouvons à présent spécifier la forme de 6.39:

$$\pi_t(h) = \pi_0(h) + \int_0^t \pi_s(H) ds + \int_0^t g_s d\overline{W}_2(s) \quad (6.41)$$

où:

$$g_s = g_s^h + g_s^x + g_s^H$$

Il nous reste à montrer que $g_s = \pi_s(D) + [\pi_s(hA) - \pi_s(h)\pi_s(A)]B_s^{-1}$. A cette fin, notons $y_t = \int_0^t g_s d\overline{W}_2(s)$ et $z_t = \int_0^t \lambda_s d\overline{W}_2(s)$ pour un processus λ quelconque et calculons:

$$\mathbb{E}(y_t z_t) = \mathbb{E} \left(\int_0^t g_s \lambda_s ds \right)$$

avec $\int_0^t (dW_s)^2 = \int_0^t ds$. Il est par ailleurs possible d'obtenir une expression alternative pour $\mathbb{E}(y_t z_t)$. De 6.41 et comme $y_t = \int_0^t g_s d\overline{W}_2(s)$, il vient:

$$y_t = \pi_t(h) - \pi_0(h) - \int_0^t \pi_s(H) ds$$

Et comme:

$$\mathbb{E}(\pi_0(h) z_t) = \mathbb{E}(\pi_0(h) \mathbb{E}(z_t | \mathcal{F}_0^Y)) = 0$$

du fait de l'indépendance entre h et \overline{W}_2 , le théorème de Fubini, cf. Lemme 6.3.9, nous donne:

$$\begin{aligned}\mathbb{E}\left(\left[\int_0^t \pi_s(H) ds\right] z_t\right) &= \int_0^t \mathbb{E}(\pi_s(H) z_t) ds \\ &= \int_0^t \mathbb{E}(\pi_s(H) \mathbb{E}(z_t | \mathcal{F}_s^Y)) ds = \int_0^t \mathbb{E}(\pi_s(H) z_s) ds\end{aligned}$$

soit:

$$\begin{aligned}\mathbb{E}(y_t z_t) &= \mathbb{E}(\pi_t(h) z_t) - \int_0^t \mathbb{E}(\pi_s(H) z_s) ds \\ &= \mathbb{E}(\mathbb{E}(h_t | \mathcal{F}_t^Y) z_t) - \int_0^t \mathbb{E}(\mathbb{E}(H_s | \mathcal{F}_s^Y) z_s) ds \\ &= \mathbb{E}\left(h_t z_t - \int_0^t z_s H_s ds\right)\end{aligned}\tag{6.42}$$

De 6.40, nous avons:

$$\overline{W}_2(t) = \int_0^t B_s^{-1} [dY_s - \pi_s(A)] ds = W_2(t) + \int_0^t B_s^{-1} [A_s - \pi_s(A)] ds$$

et comme $z_t = \int_0^t \lambda_s d\overline{W}_2(s)$, nous obtenons:

$$z_t = \tilde{z}_t + \int_0^t \lambda_s B_s^{-1} [A_s - \pi_s(A)] ds\tag{6.43}$$

où $\tilde{z}_t = \int_0^t \lambda_s dW_2(s)$. Finalement, de 6.42-6.43, il vient:

$$\begin{aligned}\mathbb{E}(y_t z_t) &= \mathbb{E}\left(h_t z_t - \int_0^t z_s H_s ds\right) = \mathbb{E}\left(h_t \tilde{z}_t - \int_0^t \tilde{z}_s H_s ds\right) \\ &\quad + \mathbb{E}\left(h_t \int_0^t \lambda_s B_s^{-1} [A_s - \pi_s(A)] ds - \int_0^t \left(\int_0^s \lambda_u B_u^+ [A_u - \pi_u(A)] du\right) H_s ds\right)\end{aligned}\tag{6.44}$$

Le premier terme de la partie droite de 6.44 donne grâce à 6.37 que:

$$\mathbb{E}\left(h_t \tilde{z}_t - \int_0^t \tilde{z}_s H_s ds\right) = \mathbb{E}\left(\tilde{z}_t \left[h_t - h_0 - \int_0^t H_s ds\right]\right) = \mathbb{E}(\tilde{z}_t x_t)\tag{6.45}$$

et comme $x_t y_t = m_t + \langle x, y \rangle_t$ pour tous processus croissants prédictifs (x, y) et une martingale m , il s'en suit que 6.45 est de la forme:

$$\mathbb{E}(\tilde{z}_t x_t) = \mathbb{E}(\langle \tilde{z}, x \rangle_t)$$

ce qui depuis $\tilde{z}_t = \int_0^t \lambda_s dW_2(s)$ donne:

$$\langle \tilde{z}, x \rangle_t = \int_0^t \lambda_s d\langle x, W_2 \rangle_s = \int_0^t \lambda_s D_s$$

où $D_t = \frac{1}{dt} d[x, W_2]_t$. Dès lors:

$$\mathbb{E} \left(h_t \tilde{z}_t - \int_0^t \tilde{z}_s H_s ds \right) = \mathbb{E} \left(\int_0^t \lambda_s D_s ds \right) = \mathbb{E} \left(\int_0^t \lambda_s \pi_s(D) ds \right) \quad (6.46)$$

Le second terme de la partie droite de 6.44 donne après application des transformations $h_t = h_s + [h_t - h_s]$ et $h_s = \pi_s(h)$:

$$\begin{aligned} & \mathbb{E} \left(h_t \int_0^t \lambda_s B_s^{-1} [A_s - \pi_s(A)] ds \right) \quad (6.47) \\ &= \mathbb{E} \left(\int_0^t \lambda_s B_s^{-1} [h_s A_s - h_s \pi_s(A)] ds \right) + \mathbb{E} \left(\int_0^t \lambda_s [h_t - h_s] B_s^{-1} [A_s - \pi_s(A)] ds \right) \\ &= \mathbb{E} \left(\int_0^t \lambda_s B_s^{-1} [\pi_s(hA) - \pi_s(h) \pi_s(A)] ds \right) + \mathbb{E} \left(\int_0^t \lambda_s [h_t - h_s] B_s^{-1} [A_s - \pi_s(A)] ds \right) \end{aligned}$$

Nous remarquons par ailleurs depuis 6.37 que:

$$h_t - h_s = \int_s^t H_u du + (x_t - x_s)$$

Et comme $\mathbb{E}(x_t - x_s | \mathcal{F}_s) = 0$, la seconde espérance dans 6.47 devient:

$$\begin{aligned} & \mathbb{E} \left(\int_0^t \lambda_s [h_t - h_s] B_s^{-1} [A_s - \pi_s(A)] ds \right) \\ &= \mathbb{E} \left(\int_0^t \lambda_s [x_t - x_s] B_s^{-1} [A_s - \pi_s(A)] ds \right) \\ &+ \mathbb{E} \left(\int_0^t \lambda_s \left(\int_s^t H_u du \right) B_s^{-1} [A_s - \pi_s(A)] ds \right) \\ &= \mathbb{E} \left(\int_0^t \left[\int_0^s \lambda_u B_u^+ [A_u - \pi_u(A)] du \right] H_s ds \right) \end{aligned}$$

Retournant au calcul du second terme de la partie droite de 6.44, nous avons alors:

$$\begin{aligned} & \mathbb{E} \left(h_t \int_0^t \lambda_s B_s^{-1} [A_s - \pi_s(A)] ds \right) \\ &= \mathbb{E} \left(\int_0^t \lambda_s B_u^+ [\pi_s(hA) - \pi_s(h) \pi_s(A)] ds \right) \\ &+ \mathbb{E} \left(\int_0^t \left[\int_0^s \lambda_u B_u^+ [A_u - \pi_u(A)] du \right] H_s ds \right) \quad (6.48) \end{aligned}$$

Et depuis 6.44, 6.46, 6.48, nous concluons:

$$\mathbb{E}(y_t z_t) = \mathbb{E} \left(\int_0^t \lambda_s ([\pi_s(D) + B_s^{-1} [\pi_s(hA) - \pi_s(h) \pi_s(A)]] ds) \right)$$

ce qui est le résultat attendu. ■

6.3.4 Equations Récursives

A partir du Théorème 6.3.6, de la spécification donnée par 6.25-6.26 et en notant que:

$$[x, W_2]_t = \int_0^t b_2 ds$$

$$\pi_s(hA) = \mathbb{E}(h_s A_s | \mathcal{F}_s^Y) = \mathbb{E}(X_s (A_0 + A_1 X_s) | \mathcal{F}_s^Y) = A_0 m_s + A_1 \mathbb{E}(X_s^2 | \mathcal{F}_s^Y)$$

et:

$$\pi_s(h) \pi_s(A) = \mathbb{E}(X_s | \mathcal{F}_s^Y) \mathbb{E}((A_0 + A_1 X_s) | \mathcal{F}_s^Y) = A_0 m_s + A_1 m_s^2$$

alors:

$$\pi_s(hA) - \pi_s(h) \pi_s(A) = A_1 [\mathbb{E}(X_s^2 | \mathcal{F}_s^Y) - m_s^2] = A_1 \gamma_s$$

De 6.38, nous obtenons alors:

$$m_t = m_0 + \int_0^t [a_0 + a_1 m_s] ds + \int_0^t \left\{ b_2 + \frac{A_1 \gamma_s}{B_s} \right\} d\bar{W}_2(s) \quad (6.49)$$

où:

$$\bar{W}_2(s) = \int_0^s B_s^{-1} [dY_s - (A_0 + A_1 m_s)] ds$$

Par ailleurs, si nous dénotons $\pi_t(h) = \mathbb{E}(h_t | \mathcal{F}_t^Y)$ l'estimateur de $h_t = X_t$, il est possible d'exprimer la variance de cet estimateur sous la forme $\omega_t(h) = \mathbb{E}[(h_t - \pi_t)^2 | \mathcal{F}_t^Y] = \mathbb{E}(h_t^2 | \mathcal{F}_t^Y) - \pi_t^2 = \delta_t - \pi_t^2$. Une application de la formule d'Itô à 6.36 donne alors:

$$X_t^2 = X_0^2 + \int_0^t [2X_s (a_0 + a_1 X_s) + (b_1^2 + b_2^2)] ds + \int_0^t 2X_s (b_1 dW_1(s) + b_2 dW_2(s))$$

Et si nous posons:

$$\tilde{h}_t = X_t^2, \tilde{H}_s = 2X_s (a_0 + a_1 X_s) + (b_1^2 + b_2^2), \tilde{x}_t = \int_0^t 2X_s (b_1 dW_1(s) + b_2 dW_2(s))$$

il vient:

$$\tilde{h}_t = \tilde{h}_0 + \int_0^t \tilde{H}_s ds + \tilde{x}_t \quad (6.50)$$

A nouveau, grâce au Théorème 6.3.6 appliqué à 6.50 et en notant que:

$$[\tilde{x}, W_2]_t = \int_0^t 2X_s b_2 ds$$

$$\pi_s(\tilde{h}A) = A_0 \delta_s + A_1 \mathbb{E}(X_s^3 | \mathcal{F}_s^Y)$$

et:

$$\pi_s(\tilde{h}) \pi_s(A) = \delta_s (A_0 + A_1 m_s)$$

il vient:

$$\begin{aligned} \delta_t &= \delta_0 + \int_0^t [2a_0 m_s + 2a_1 \delta_s + (b_1^2 + b_2^2)] ds \\ &+ \int_0^t \{2m_s b_2 + [A_0 \delta_s + A_1 \mathbb{E}(X_s^3 | F_s^Y) - \delta_s (A_0 + A_1 m_s)] B^{-1}\} d\overline{W}_2(s) \end{aligned}$$

soit encore:

$$\begin{aligned} \delta_t &= \delta_0 + \int_0^t [2a_0 m_s + 2a_1 \delta_s + (b_1^2 + b_2^2)] ds \\ &+ \int_0^t B^{-1} \{2m_s b_2 B + A_1 [\mathbb{E}(X_s^3 | F_s^Y) - \delta_s m_s]\} d\overline{W}_2(s) \end{aligned} \quad (6.51)$$

Par suite, une application de la formule d'Itô à 6.49 donne:

$$\begin{aligned} m_t^2 &= m_0^2 + \int_0^t \left(2m_s [a_0 + a_1 m_s] + \left[\frac{b_2 B + A_1 \gamma_s}{B} \right]^2 \right) ds \\ &+ \int_0^t \left\{ 2m_s \frac{b_2 B + A_1 \gamma_s}{B} \right\} d\overline{W}_2(s) \end{aligned} \quad (6.52)$$

En calculant $\gamma_t = \delta_t - m_t^2$ depuis 6.51-6.52, nous obtenons:

$$\begin{aligned} \gamma_t &= \gamma_0 + \int_0^t \left[2a_1 \gamma_s + b_1^2 + b_2^2 - \left(\frac{b_2 B + A_1 \gamma_s}{B} \right)^2 \right] ds \\ &+ \int_0^t \frac{A_1}{B} \{ \mathbb{E}(X_s^3 | F_s^Y) - \delta_s m_s - 2m_s \gamma_s \} d\overline{W}_2(s) \end{aligned}$$

Et comme la distribution conditionnelle de $X_s | \mathcal{F}_s^Y$ est gaussienne, nous avons:

$$\mathbb{E}(X_s^3 | \mathcal{F}_s^Y) = 3m_s \gamma_s - 2m_s^3 = \delta_s m_s + 2m_s \gamma_s$$

ce qui amène à:

$$\gamma_t = \gamma_0 + \int_0^t \left[2a_1 \gamma_s + b_1^2 + b_2^2 - \left(\frac{b_2 B + A_1 \gamma_s}{B} \right)^2 \right] ds$$

Nous rassemblons les résultats précédents dans le théorème suivant qui donne les équations de filtrage dans le cadre continu.

Théorème 6.3.11 (Filtre - Continu) *Les paramètres m_t et γ_t définis par 6.27 et 6.28 satisfont, pour $t \in [0, T]$:*

$$dm_t = [a_0 + a_1 m_t] dt + \frac{b_2 B + A_1 \gamma_t}{B^2} [dY_t - (A_0 + A_1 m_t) dt] \quad (6.53)$$

$$\dot{\gamma}_t = 2a_1 \gamma_t + b_1^2 + b_2^2 - \left(\frac{b_2 B + A_1 \gamma_t}{B} \right)^2 \quad (6.54)$$

Nous présentons à présent quelques exemples d'utilisation du Théorème 6.3.11 et dérivons des solutions explicites pour les équations 6.53 et 6.54.

Exemple 6.3.12 (Filtre) *Supposons que le couple partiellement observé (Y, X) satisfait:*

$$dY_t = (\nu + \lambda X_t) dt + dW_t^2, \quad Y_0 = 0 \quad (6.55)$$

$$dX_t = \alpha(\delta - X_t) dt + dW_t^1, \quad X_0 = 0 \quad (6.56)$$

et que nous observons $(Y_s, s \in [0, t])$ et on définit $\mathcal{F}_t^Y = \sigma(Y_s, s \in [0, t])$ pour $t \in [0, T]$. Si on note $m_t = \mathbb{E}[X_t | \mathcal{F}_t^Y]$ et $\gamma_t = \mathbb{E}[(X_t - \gamma_t)^2 | \mathcal{F}_t^Y]$, respectivement le filtre et la variance du filtre, une application du Théorème 6.3.11 nous donne alors:

$$m_t = \alpha(\delta - m_t) dt - \lambda \gamma_t(\nu + \lambda m_t) dt + \lambda \gamma_t dY_t, \quad m_0 = 0 \quad (6.57)$$

et:

$$\frac{d\gamma_t}{dt} = -\lambda^2 \gamma_t^2 - 2\alpha \gamma_t + 1, \quad \gamma_0 = 0 \quad (6.58)$$

Le filtre m_t est obtenue comme solution de 6.57. Notant:

$$\Phi_t = \exp\left(-\alpha t - \lambda^2 \int_0^t \gamma_s ds\right)$$

la solution est donnée par:

$$m_t = \Phi_t \left[\lambda \int_0^t \gamma_s \Phi_s^{-1} (dY_s - \nu ds) + \alpha \delta \int_0^t \Phi_s^{-1} ds \right] \quad (6.59)$$

Tandis que γ_t est donnée par la solution de 6.58, sous la forme:

$$\gamma_t = \frac{b \exp(2bt) - d}{\lambda^2 \exp(2bt) + d} - \frac{\alpha}{\lambda^2} \quad (6.60)$$

avec $b = \sqrt{\alpha^2 + \lambda^2}$, $d = (b - \alpha) / (b + \alpha)$.

6.3.5 Interpolation: Lissage

Soit le couple partiellement observé de processus stochastiques (Y, X) satisfaisant 6.26-6.25. Considérons que la distribution a priori $P(X_0 \leq x | Y_0)$ est gaussienne telle que $X_0 \sim \mathcal{N}(m_0, \gamma_0)$. Alors la distribution conditionnelle $P(X_s \leq x | \mathcal{F}_t^Y)$, $s \leq t$, est gaussienne de paramètres:

$$m(s, t) = \mathbb{E}[X_s | \mathcal{F}_t^Y], \quad \gamma(s, t) = \mathbb{E}[(X_s - m(s, t))^2 | \mathcal{F}_t^Y]$$

Ces équations nous permettent d'établir une échelle de précision des estimations des valeurs de X_t avec l'augmentation de la taille de l'échantillon d'observations, c'est-à-dire lorsque t augmente et s est fixé. Nous introduisons les quantités φ_s^t et q_s^t comme:

$$\frac{d\varphi_s^t}{dt} = [a(t, Y) - \gamma(t, s) d(t, Y)] \varphi_s^t, \quad \varphi_s^s = 1 \quad (6.61)$$

avec:

$$\begin{aligned} a(t, Y) &= a_1(t, Y) - c(t, Y) B^{-1}(t, Y) A_1(t, Y) \\ d(t, Y) &= A_1^2(t, Y) B^{-1}(t, Y) \end{aligned}$$

et:

$$\begin{aligned} q_s^t &= \int_s^t (\varphi_s^u)^{-1} [a_0(u, Y) du + [c(u, Y) + \gamma(u, s) A_1(u, Y)]] \\ &\quad \times B^{-1}(u, Y) [dY_u - A_0(u, Y) du] \end{aligned}$$

Alors il est possible de montrer par application de la formule d'Itô que:

$$m_{X_s}(t, s) = \mathbb{E}[X_t | X_s, \mathcal{F}_t^Y] = \varphi_s^t (X_s + q_s^t) \quad (6.62)$$

et:

$$\gamma_{X_s}(t, s) = \mathbb{E}[(X_t - m_{X_s}(t, s))^2 | X_s, \mathcal{F}_t^Y]$$

Nous avons alors le lemme important suivant.

Lemme 6.3.13 *Soit $0 \leq s \leq t \leq T$, alors:*

$$m_t = \varphi_s^t [m(s, t) + q_s^t] \quad (6.63)$$

$$\gamma_t = \gamma(t, s) + (\varphi_t^s)^2 \gamma(s, t) \quad (6.64)$$

Preuve. Nous avons:

$$m_t = \mathbb{E}[X_t | \mathcal{F}_t^Y] = \mathbb{E}[\mathbb{E}[X_t | X_s, \mathcal{F}_t^Y] | \mathcal{F}_t^Y] = \mathbb{E}[m_{X_s}(t, s) | \mathcal{F}_t^Y] \quad (6.65)$$

Et, de 6.62, il vient:

$$\mathbb{E}[m_{X_s}(t, s) | \mathcal{F}_t^Y] = \mathbb{E}[\varphi_s^t [X_s + q_s^t] | \mathcal{F}_t^Y] = \varphi_s^t (m(s, t) + q_s^t)$$

qui utilisé avec 6.65 donne la représentation 6.63. Par ailleurs, comme:

$$\mathbb{E}[(X_t - m_{X_s}(t, s))(m_{X_s}(t, s) - m_t) | X_s, \mathcal{F}_t^Y] = 0$$

il s'ensuit:

$$\begin{aligned} \gamma_t &= \mathbb{E}[(X_t - m_t)^2 | \mathcal{F}_t^Y] \\ &= \mathbb{E}\{[(X_t - m_{X_s}(t, s)) + (m_{X_s}(t, s) - m_t)] \\ &\quad \times [(X_t - m_{X_s}(t, s)) + (m_{X_s}(t, s) - m_t)] | \mathcal{F}_t^Y\} \\ &= \mathbb{E}\left\{\mathbb{E}[(X_t - m_{X_s}(t, s))^2 | X_s, \mathcal{F}_t^Y] | \mathcal{F}_t^Y\right\} \\ &\quad + \mathbb{E}[(m_{X_s}(t, s) - m_t)^2 | \mathcal{F}_t^Y] \\ &= \gamma(t, s) + \mathbb{E}[(m_{X_s}(t, s) - m_t)^2 | \mathcal{F}_t^Y] \end{aligned}$$

Notant que:

$$m_{X_s}(t, s) - m_t = \varphi_s^t (X_s + q_s^t) - \varphi_s^t (m(s, t) + q_s^t) = \varphi_s^t (X_s - m(s, t))$$

nous avons alors:

$$\begin{aligned} & \mathbb{E} \left[(m_{X_s}(t, s) - m_t)^2 \mid \mathcal{F}_t^Y \right] \\ &= \varphi_s^t \mathbb{E} \left[(m_{X_s}(t, s) - m_t)^2 \mid \mathcal{F}_t^Y \right] \\ &= (\varphi_s^t)^2 \mathbb{E} \left[(X_s - m(s, t))^2 \mid \mathcal{F}_t^Y \right] \\ &= (\varphi_s^t)^2 \gamma(s, t) \end{aligned}$$

ce qui conclut la preuve. ■

Les représentations 6.63 et 6.64 nous permettent alors de dériver les équations satisfaites par $m(s, t)$ et $\gamma(s, t)$, ce qui va permettre de comprendre comment ces paramètres changent avec l'évolution du temps t lorsque s est fixé.

Théorème 6.3.14 (Lisseur - Continu) *Sous 6.25-6.26, les paramètres $m(t, s)$ et $\gamma(t, s)$, pour $0 \leq s \leq t \leq T$, satisfont:*

$$\begin{aligned} m(s, t) &= m_s + \int_s^t \gamma(s, u) \varphi_s^u A_1(u, Y) B^{-1}(u, Y) \\ &\quad \times (dY_u - [A_0(u, Y) + A_1(u, Y) m_u] du) \end{aligned} \quad (6.66)$$

$$\gamma(s, t) = \left(1 + \gamma_s \int_s^t (\varphi_s^u)^2 A_1^2(u, Y) B^{-1}(u, Y) du \right)^{-1} \gamma_s \quad (6.67)$$

Preuve. De 6.63, nous avons:

$$m(s, t) = (\varphi_s^t)^{-1} m_t - q_s^t \quad (6.68)$$

où $(\varphi_s^t)^{-1}$ est dérivé depuis 6.61 pour $t \geq s$:

$$\frac{d\varphi_s^t}{dt} = -(\varphi_s^t)^{-1} [a(t, Y) - \gamma(t, s) d(t, Y)], \quad (\varphi_s^s)^{-1} = 1 \quad (6.69)$$

Alors, de 6.53, 6.68, 6.69 et par la formule d'Itô, il vient:

$$\begin{aligned} m(s, t) &= m_s + \int_s^t (\varphi_s^u)^{-1} (\gamma_u - \gamma(u, s)) A_1(u, Y) B^{-1}(u, Y) \\ &\quad \times (dY_u - [A_0(u, Y) + A_1(u, Y) m_u] du) \end{aligned} \quad (6.70)$$

Mais de 6.64, nous avons:

$$(\varphi_s^u)^{-1} (\gamma_u - \gamma(u, s)) = \gamma(s, u) \varphi_s^u$$

ce qui par substitution dans 6.70 donne la représentation 6.66. Par ailleurs, de 6.64:

$$\gamma(s, t) = (\varphi_s^t)^{-2} (\gamma_t - \gamma(t, s)) \quad (6.71)$$

Alors, en différenciant 6.71 et en utilisant 6.54, 6.69, il vient:

$$\frac{d\gamma(s,t)}{dt} = -\gamma(s,t)^2 (\varphi_s^t)^2 A_1^2(t, Y) B^{-1}(t, Y)$$

dont une solution est donnée par 6.67, ce qui conclut la preuve. ■

A titre d'illustration, nous reprenons l'exemple précédent.

Exemple 6.3.15 (Lisseur) Sous 6.55-6.56, nous nous intéressons à présent à $m(s,t) = \mathbb{E}[X_t | \mathcal{F}_s^Y]$, $\gamma(s,t) = \mathbb{E}[(X_t - m(s,t))^2 | \mathcal{F}_s^Y]$ et $n(s,t) = \mathbb{E}[X_t^2 | \mathcal{F}_s^Y]$ pour $s \leq t$, où $m(s,t)$ est appelée le lisseur et $\gamma(s,t)$ sa variance. Une application du Théorème 6.3.14 nous donne:

$$m(s,t) = m_s - \alpha \int_s^t \varphi_s^u \gamma(s,u) (dY_u - (\nu + \lambda m_u) du), \quad s \leq t$$

et:

$$\gamma(s,t) = \left[1 + \lambda^2 \gamma_s \int_s^t (\varphi_s^u)^2 du \right]^{-1} \gamma_s, \quad s \leq t$$

où φ_s^t , $t \geq s$, est solution de:

$$\frac{d\varphi_s^t}{dt} = (-\alpha - \lambda^2 \gamma(t,s)), \quad \varphi_s^s = 1, \quad t \geq s \quad (6.72)$$

Pour $t > s$, $\gamma(t,s) = \text{Var}(X_t | \mathcal{F}_t^Y, X_s)$ satisfait:

$$\frac{d\gamma(t,s)}{dt} = -2\alpha\gamma(t,s) + 1 - \lambda^2 \gamma(t,s)^2, \quad \gamma(s,s) = 0$$

qui d'après l'Exemple 6.3.12 admet comme solution:

$$\gamma(t,s) = \frac{b \exp(2b(t-s)) - d}{\lambda^2 \exp(2b(t-s)) + d} - \frac{\alpha}{\lambda^2}, \quad t > s$$

Par suite, une solution à 6.72 est donnée par:

$$\varphi_s^t = \exp\left(-b \int_s^t \frac{\exp(2b(u-s)) - d}{\exp(2b(u-s)) + d} du\right) = \frac{(1+d) \exp(b(t-s))}{\exp(2b(t-s)) + d}$$

6.3.6 Extrapolation: Prédiction

Les équations dite d'extrapolation permettent de calculer des estimations optimales (au sens des moindres carrés) de la variable X_t à partir des observations $Y_0^s = (Y_0, \dots, Y_s)$ pour $s \leq t \leq T$. Il s'agit alors d'un problème de prédiction et on note:

$$m_X(t,s) = \mathbb{E}[X_t | \mathcal{F}_s^Y], \quad m_Y(t,s) = \mathbb{E}[Y_t | \mathcal{F}_s^Y], \quad s \leq t$$

Ces équations sont aussi appelées équations backward. De ces équations, nous pouvons juger, à une date $s \leq t$ fixée, de combien les prédictions de X_t se

détériorient lorsque t augmente. Par ailleurs, nous supposons à présent que l'équation d'état satisfait:

$$dX_t = [a_0(t) + a_1(t) X_t] dt + b_1(t, Y) dW_t^1 + b_2(t, Y) dW_t^2 \quad (6.73)$$

où $a_0(t)$ et $a_1(t)$ sont des fonctions déterministes du temps. Cette restriction à une classe de processus moins riche que la spécification 6.26 permet d'assurer que les distributions a posteriori $P(X_t \leq x | \mathcal{F}_s^Y)$ pour $t > s$ sont bien gaussiennes.

Soit φ_s^t , $t \geq s$, solution de l'équation:

$$\frac{d\varphi_s^t}{dt} = a_1(t) \varphi_s^t, \quad \varphi_s^s = 1$$

Nous avons alors le résultat important suivant.

Théorème 6.3.16 (Prédicteur - Continu) *Sous 6.25-6.73, le paramètre $m_X(t, s)$, pour $0 \leq s \leq t \leq T$, satisfait:*

$$\frac{dm_X(t, s)}{dt} = a_0(t) + a_1(t) m_X(t, s), \quad m_X(s, s) = m_s$$

où (m_s, γ_s) est donné par 6.53 et 6.54. Soit aussi:

$$\begin{aligned} m_X(t, s) = m_X(t, 0) + \int_0^s \varphi_u^t [c(u, Y) + \gamma_u A_1(u, Y)] B^{-1}(u, Y) \\ \times (dY_u - [A_0(u, Y) + A_1(u, Y) m_u]) du \end{aligned} \quad (6.74)$$

avec:

$$m_X(t, 0) = \varphi_0^t \left[m_0 + \int_0^t (\varphi_0^s)^{-1} a_0(s) ds \right]$$

La dérivation de $m_Y(t, s)$ est par la suite claire. Nous avons:

$$m_Y(t, s) = A_0(t, Y) + A_1(t, Y) m_X(t, s)$$

A titre d'illustration, nous reprenons l'exemple précédent.

Exemple 6.3.17 (Prédicteur) *Sous 6.55-6.56, nous avons $a_1(t) = -\alpha$, soit:*

$$\varphi_s^t = \exp(-\alpha(t-s))$$

D'où pour $t \geq s$:

$$m_X(t, s) = m_X(t, 0) + \int_0^s e^{-\alpha(t-u)} (1 + \lambda \gamma_u) (dY_u - (\nu + \lambda m_u) du)$$

avec:

$$m_X(t, 0) = e^{-\alpha t} [m_0 + \delta (1 - e^{\alpha t})]$$

7

Filtrage de Chaînes de Markov

7.1. Contexte

Le présent Chapitre s'intéresse à la question du filtrage pour un couple de processus stochastiques $(Y_t, X_t)_{t \geq 0}$ où le processus non-observable X est un processus de Markov pouvant prendre un nombre fini ou dénombrable d'états - on parle alors d'une chaîne de Markov - et où le processus observable Y satisfait l'équation:

$$dY_t = A_t(X_t, Y) dt + B_t(Y) dW_t \quad (7.1)$$

où $(W_t)_{t \geq 0}$ est un mouvement brownien et A, B sont des fonctions non-anticipatives vérifiant des conditions (de croissance, Lipschitz) telle qu'il existe une unique solution Y_t à 7.1.

Supposons que les réalisations $Y_0^t = (Y_s, s \in [0, t])$ du processus observable Y sont connues pour tout $t \in [0, T]$. Le problème du filtrage du processus non-observable X se ramène à la construction d'estimations des valeur de X_t à partir de Y_0^t . D'un point de vue statistique, le critère adéquat est la probabilité a posteriori:

$$\pi_x(t) = P(X_t = x | \mathcal{F}_t^Y), \quad x \in \mathcal{X}$$

De façon intéressante, on remarquera que la connaissance de $\pi_x(t)$ permet le calcul de nombreuses estimations connexes à la valeur de X_t . En particulier, l'espérance conditionnelle:

$$\mathbb{E}[X_t | \mathcal{F}_t^Y] = \sum_{x \in \mathcal{X}} x \pi_x(t)$$

et l'estimation $x_t(Y)$ obtenue comme:

$$\max_x P(X_t = x | \mathcal{F}_t^Y) = \pi_{x_t(Y)}(t)$$

qui est alors l'estimation de la valeur en laquelle la probabilité a posteriori est maximale.

7.2. Structure des Probabilités

Nous introduisons les notations suivantes:

$$p_x(t) = P(X_t = x), \quad x \in \mathcal{X}$$

$$p_{xy}(t, s) = P(X_t = x, X_s = y), \quad 0 \leq s \leq t \leq T$$

et notons $\lambda_{yx}(t)$ la densité des probabilités de transition de y à x à la date t . Nous avons alors le résultat important suivant sur la structure des probabilités de transition que nous caractérisons à l'aide de l'équation de Kolmogorov.

Lemme 7.2.1 *Les probabilités de transition $p_{xy}(t, s)$ satisfont l'équation de Kolmogorov:*

$$p_{xy}(t, s) = \delta(x, y) + \int_s^t \mathcal{L}p_{xy}(u, s) du \quad (7.2)$$

où:

$$\mathcal{L}p_{xy}(u, s) = \sum_{\gamma \in \mathcal{X}} \lambda_{\gamma x}(u) p_{\gamma y}(u, s)$$

avec $\delta(x, y)$ l'opérateur de Kronecker tel que:

$$\delta(x, y) = \begin{cases} 0, & \text{si } x \neq y \\ 1, & \text{si } x = y \end{cases}$$

Et les probabilités $p_x(t)$ satisfont:

$$p_x(t) = p_x(0) + \int_0^t \mathcal{L}p_x(u) du \quad (7.3)$$

où:

$$\mathcal{L}p_x(u) = \sum_{\gamma \in \mathcal{X}} \lambda_{\gamma x}(u) p_\gamma(u)$$

Preuve. Soit $s = t_0 < t_1 < \dots < t_n = t$. Du fait que X est markovien, nous avons:

$$\begin{aligned} p_{xy}(t_{j+1}, s) &= P(X_{t_{j+1}} = x | X_s = y) \\ &= \mathbb{E}[P(X_{t_{j+1}} = x | X_{t_j}, X_s = y) | X_s = y] \\ &= \mathbb{E}[P(X_{t_{j+1}} = x | X_{t_j} | X_s = y)] \end{aligned}$$

ou aussi:

$$p_{xy}(t_{j+1}, s) = \sum_{\gamma \in \mathcal{X}} p_{x\gamma}(t_{j+1}, t_j) p_{\gamma y}(t_j, s) \quad (7.4)$$

Notant:

$$r_{x\gamma}(t_{j+1}, t_j) = p_{x\gamma}(t_{j+1}, t_j) - \delta(x, \gamma) - \lambda_{\gamma x}(t_j)(t_{j+1} - t_j)$$

alors, depuis 7.4, il vient que:

$$\begin{aligned} & p_{xy}(t_{j+1}, s) \\ &= \sum_{\gamma \in \mathcal{X}} [\delta(x, \gamma) + \lambda_{\gamma x}(t_j)(t_{j+1} - t_j) + r_{x\gamma}(t_{j+1}, t_j)] p_{\gamma y}(t_j, s) \\ &= p_{xy}(t_j, s) + \left(\sum_{\gamma \in \mathcal{X}} \lambda_{\gamma x}(t_j) p_{\gamma y}(t_j, s) \right) (t_{j+1} - t_j) + \sum_{\gamma \in \mathcal{X}} r_{x\gamma}(t_{j+1}, t_j) p_{\gamma y}(t_j, s) \end{aligned}$$

Soit encore:

$$\begin{aligned} p_{xy}(t_{j+1}, s) - \delta(x, y) &= \sum_{j=0}^{n-1} [p_{xy}(t_{j+1}, s) - p_{xy}(t_j, s)] \\ &= \int_s^t \sum_{\gamma \in \mathcal{X}} \lambda_{\gamma x}(t_j(u)) p_{\gamma y}(t_j(u), s) du \end{aligned}$$

ce qui conclut la preuve. ■

Le lemme suivant sera important dans la dérivation des équations de filtrage.

Lemme 7.2.2 *For tout $x \in \mathcal{X}$, nous posons:*

$$z_t^x = \delta(x, X_t) - \delta(x, X_0) - \int_0^t \lambda_{X_u x}(u) du \quad (7.5)$$

Le processus stochastique $Z^x = (z_t^x, \mathcal{F}_t)$ est une martingale de carré intégrable.

Preuve. Soit $t > s$ pour $t \in [0, T]$, alors:

$$z_t^x = z_s^x + \left[\delta(x, X_t) - \delta(x, X_s) - \int_s^t \lambda_{X_u x}(u) du \right]$$

Soit par passage à l'espérance conditionnelle:

$$\mathbb{E}[z_t^x | \mathcal{F}_s] = z_s^x + \mathbb{E} \left[\delta(x, X_t) - \delta(x, X_s) - \int_s^t \lambda_{X_u x}(u) du | \mathcal{F}_s \right]$$

Par suite, du fait du caractère markovien du processus X_t et de 7.2, il vient:

$$\begin{aligned} & \mathbb{E} \left[\delta(x, X_t) - \delta(x, X_s) - \int_s^t \lambda_{X_u x}(u) du | \mathcal{F}_s \right] \\ &= \mathbb{E} \left[\delta(x, X_t) - \delta(x, X_s) - \int_s^t \lambda_{X_u x}(u) du \right] \\ &= p_{xX_s}(t, s) - \delta(x, X_s) - \int_s^t \sum_{\gamma \in \mathcal{X}} \lambda_{\gamma x}(u) p_{\gamma X_s}(u, s) = 0 \end{aligned}$$

ce qui conclut la preuve. ■

7.3. Résultat Général

Nous pouvons à présent dériver l'équation de filtrage satisfaite par $\pi_x(t)$.

Théorème 7.3.1 *La probabilité a posteriori $\pi_x(t)$, $x \in \mathcal{X}$, satisfait le système d'équations:*

$$\pi_x(t) = \pi_x(0) + \int_0^t \mathcal{L}\pi_x(u) du + \int_0^t \pi_x(u) B_u^{-1}(Y) [A_u(x, Y) - \bar{A}_u(Y)] d\bar{W}_u \quad (7.6)$$

où:

$$\mathcal{L}\pi_x(u) = \sum_{\gamma \in \mathcal{X}} \lambda_{\gamma x}(u) \pi_\gamma(t)$$

$$\bar{A}_u(Y) = \sum_{\gamma \in \mathcal{X}} A_u(\gamma, Y) \pi_\gamma(u)$$

et $\bar{W} = (\bar{W}, \mathcal{F}_t^Y)$ est un mouvement brownien tel que:

$$\bar{W}_t = \int_0^t B_u^{-1}(Y) [dY_u - \bar{A}_u(Y) du] \quad (7.7)$$

Preuve. Grâce au Lemme 7.2.2, nous avons:

$$\delta(x, X_t) = \delta(x, X_0) - \int_0^t \lambda_{X_u x}(u) du + z_t^x$$

où $Z^x = (z_t^x, \mathcal{F}_t)$ est une martingale de carré intégrable. On notera par ailleurs que comme les processus Z^x et W sont indépendants, $\langle z^x, W \rangle_t = 0$ pour $t \in [0, T]$. Notant $h_t = \delta(x, Y_t)$, une application du Théorème 6.3.6 à h donne:

$$\pi_t^x(\delta) = \pi_0^x(\delta) + \int_0^t \pi_s^x(\lambda) ds + \int_0^t B_u^{-1}(Y) [\pi_s^x(\delta A) - \pi_s^x(\delta) \pi_s^x(A)] d\bar{W}_u$$

où:

$$\pi_t^x(\delta) = \mathbb{E}[\delta(x, X_t) | \mathcal{F}_t^Y] = \pi_x(t)$$

$$\pi_s^x(\lambda) = \mathbb{E}[\lambda_{X_s x}(s) | \mathcal{F}_s^Y] = \sum_{\gamma \in \mathcal{X}} \lambda_{\gamma x}(s) \pi_\gamma(s) = \mathcal{L}\pi_x(s)$$

$$\pi_t^x(\delta A) = \mathbb{E}[\delta(x, X_s) A_s(X_s, Y) | \mathcal{F}_s^Y] = A_s(X_s, Y) \pi_x(s)$$

$$\pi_s^x(A) = \mathbb{E}[A_s(X_s, Y) | \mathcal{F}_s^Y] = \bar{A}_s(Y) = \sum_{\gamma \in \mathcal{X}} A_s(x, Y) \pi_x(s)$$

ce qui est le résultat attendu. ■

Remarque 7.3.2 (Filtre de Wonham) *La récursion 7.6 est appelée filtre de Wonham (1965). C'est la prédiction optimale, au sens des moindres carrés, d'une chaîne de Markov admettant un nombre fini d'états. On notera que sa dérivation est proche de celle du filtre de Kalman et on pourra comparer 6.38 avec 7.6 pour s'en assurer.*

Exemple 7.3.3 (Modèle à régimes) Supposons que X soit une chaîne de Markov à valeurs dans l'alphabet $\mathcal{X} = (x_1, \dots, x_m)$ où m est le nombre total d'états et de matrice de transition $Q = (q_{ij}) \in \mathbb{R}^{m \times m}$. Les probabilités de transition s'écrivent alors:

$$p_{ij}(s) = P(X_{t+s} = x_j | X_s = x_i)$$

et:

$$p_{ij}(s) = \begin{cases} 1 - q_i \times s + o(s) & \text{si } i = j \\ q_{ij} \times s + o(s) & \text{si } i \neq j \end{cases}$$

où $q_i = \sum_{i \neq j} q_{ij}$. Nous supposons par ailleurs que le processus X est mesuré par l'observation du processus Y sous la forme:

$$dY_t = X_t dt + \sigma_t dW_t, \quad Y_0 = 0$$

Nous définissons alors:

$$p(t) = (p_1(t), \dots, p_m(t)) \in \mathbb{R}^{1 \times m}$$

avec:

$$p_i(t) = P(X_t = x_i | X_s), \quad p_i(0) = p_0^i, \quad s \leq t, \quad i = 1, \dots, m$$

Une application du Théorème 7.3.1 nous permet d'écrire, pour $i = 1, \dots, m$:

$$dp_i(t) = \sum_{j=1}^m p_j(t) q_{ji} dt - \sigma_t^{-2} \bar{\alpha}(t) [x_i - \bar{\alpha}(t)] p_i(t) dt + \sigma_t^{-2} [x_i - \bar{\alpha}(t)] p_i(t) dY_t$$

où:

$$\bar{\alpha}(t) = \sum_{i=1}^m x_i p_i(t)$$

Notant:

$$A(t) = \text{diag}(x_1 - \bar{\alpha}(t), \dots, x_m - \bar{\alpha}(t))$$

le filtre de Wonham se réécrit alors:

$$dp_t = Q p_t dt + \sigma_t^{-2} A(t) p_t [dY_t - \bar{\alpha}(t) dt]$$

7.4. Approche Alternative

La dérivation précédente des équations de filtrage est basée sur la théorie des innovations, comme développée en Section 6.3.2 et le processus 7.7 est alors qualifié de processus des innovations. Une approche alternative existe. Elle est basée sur l'utilisation du théorème de Girsanov et utilise de façon importante la formule de Bayes: on parle de l'approche par la probabilité de référence. Dans la suite, nous la développons au travers d'un exemple.

Nous commençons par donner la forme de la formule de Bayes conditionnelle que nous emploierons dans la suite de la présentation.

Définition 7.4.1 (Bayes) Soit le changement de probabilité $Z_T = \frac{dP_0}{dP}$ donné par le théorème de Radon-Nikodym, alors pour toute \mathcal{F} -mesurable variable aléatoire X , nous avons:

$$\mathbb{E}_0[X|\mathcal{F}_t] = \frac{1}{\mathbb{E}[Z_t|\mathcal{F}_t]} \mathbb{E}[Z_t X|\mathcal{F}_t], \quad t \in [0, T]$$

Soit $Y = (Y_t, \mathcal{F}_t)$ un processus d'Itô satisfaisant, pour $t \in [0, T]$:

$$dY_t = \mu_t dt + \sigma dW_t \quad (7.8)$$

où $(W_t)_{t \geq 0}$ est un mouvement brownien et $\sigma > 0$ est la *volatilité* de Y . Nous supposons par ailleurs que $\mu = (\mu_t, \mathcal{F}_t)$, la *tendance* de Y , est une chaîne de Markov stationnaire, indépendante de W admettant d états $B = (b_1, \dots, b_d) \in \mathbb{R}^d$. Alors μ s'écrit sous la forme:

$$\mu_t = BX_t \quad (7.9)$$

où $X = (X_t, \mathcal{F}_t)$ est une chaîne de Markov stationnaire, indépendante de W à valeurs dans l'alphabet $\{e_1, \dots, e_d\}$, où les e_i sont les vecteurs unitaires de \mathbb{R}^d .

Remarque 7.4.2 La décomposition 7.9 va permettre de spécifier les équations de filtrage pour X et pas seulement pour ses probabilités de transition, comme précédemment.

Par ailleurs, X est caractérisé par sa matrice de transition $Q \in \mathbb{R}^{d \times d}$ où:

$$Q_{ij} = \lim_{t \rightarrow 0} \frac{1}{t} P(X_t = e_i | X_0 = e_j), \quad i \neq j$$

est le taux de transition de e_i à e_j et $\lambda_j = -Q_{jj} = \sum_{k=1}^d Q_{jk}$ est le taux auquel on quitte e_j . D'où, dans l'état e_j , le temps d'attente pour le prochain saut est distribué suivant une loi exponentielle de paramètre λ_j et Q_{jk}/λ_j est la probabilité que la chaîne atteigne e_k en quittant e_j . Finalement, nous supposons que seules les observations de Y sont disponibles et on note $\mathcal{F}_t^Y = \sigma(Y_s, s \in [0, t])$ pour $t \in [0, T]$ la filtration engendrée par Y .

Soit l'équation différentielle stochastique:

$$Z_t = \exp \left(- \int_0^t \sigma^{-1} B_s X_s dW_s - \frac{1}{2} \int_0^t (\sigma^{-1} B_s X_s)^2 ds \right), \quad Z_0 = 1 \quad (7.10)$$

Comme le terme $(\sigma^{-1} B_t X_t)_t$ est borné, le processus Z est alors une martingale de carré intégrable. Le théorème de Radon-Nikodym nous permet alors de définir une nouvelle mesure de probabilité \tilde{P} absolument continue par rapport à P sur \mathcal{F} telle que:

$$Z_T = \frac{d\tilde{P}}{dP} \Big|_{\mathcal{F}_T}$$

On notera par la suite $\tilde{\mathbb{E}}$ l'espérance prise sous la probabilité \tilde{P} tel que $\tilde{\mathbb{E}}[X] = \mathbb{E}[Z_T X]$. Par ailleurs, une application du théorème de Girsanov nous assure que le processus :

$$\tilde{W}_t = W_t + \int_0^t \sigma^{-1} B_s X_s ds \quad (7.11)$$

est un (\mathcal{F}, \tilde{P}) -mouvement brownien. De 7.8 et 7.11, il vient $Y_t = \sigma \tilde{W}_t$, ce qui implique $\mathcal{F}^{\tilde{W}} = \mathcal{F}^Y$ et nous permet de pouvoir utiliser le théorème de représentation des martingales pour toute \mathcal{F}^Y -martingale, cf. Théorème 6.3.1 et Lemme 6.3.5.

Contrairement à l'approche par la théorie des innovations, nous introduisons dans la suite deux équations de filtrage: le filtre conditionnel (*conditional* ou *normalized*) et le filtre non-conditionnel (*unconditional* ou *unnormalized*). Pour la suite, on note $\Lambda_t = \mathbb{E}[Z_t | \mathcal{F}_t^Y]$.

Définition 7.4.3 Les filtres conditionnel et non-conditionnel pour $X = (X_t, \mathcal{F}_t)$ sont respectivement définis par:

$$\pi_t = \mathbb{E}[X_t | \mathcal{F}_t^Y] \text{ et } \Pi_t = \tilde{\mathbb{E}}[Z_t^{-1} X_t | \mathcal{F}_t^Y]$$

La formule de Bayes donne la relation:

$$\Pi_t = \frac{\tilde{\mathbb{E}}[Z_t^{-1} X_t | \mathcal{F}_t^Y]}{\tilde{\mathbb{E}}[Z_t^{-1} | \mathcal{F}_t^Y]} = \frac{\mathbb{E}[Z_t Z_t^{-1} X_t | \mathcal{F}_t^Y]}{\mathbb{E}[Z_t | \mathcal{F}_t^Y]} = \Lambda_t^{-1} \pi_t \quad (7.12)$$

De 7.12, il est clair que la dérivation de π ou Π repose largement sur la connaissance du processus $Z_t^{-1}, t \in [0, T]$. Nous dérivons par la suite sa représentation.

Lemme 7.4.4 Soit $Z_t, t \in [0, T]$, donné par 7.10. Alors:

$$Z_t^{-1} = 1 + \int_0^t Z_s^{-1} \sigma^{-2} B_s X_s dY_s \quad (7.13)$$

Preuve. Notant $\theta_s = \sigma^{-1} B_s X_s$, soit:

$$A_t = - \int_0^t \theta_s dW_s - \frac{1}{2} \int_0^t \theta_s^2 ds$$

Alors pour $f(x) = e^{-x}$ avec $x = A$, la formule d'Itô nous donne:

$$\begin{aligned} dZ_t^{-1} &= df(A_t) = f'(A_t) dA_t + \frac{1}{2} f''(A_t) dA_t dA_t \\ &= -e^{-A_t} \left(-\theta_t dW_t - \frac{1}{2} \theta_t^2 dt \right) + \frac{1}{2} e^{-A_t} \theta_t^2 dt \\ &= Z_t^{-1} \theta_t (dW_t + \theta_t dt) \end{aligned}$$

Soit, d'après 7.11 et comme $Y_s = \sigma \tilde{W}_s$, il vient:

$$dZ_t^{-1} = Z_t^{-1} \theta_t d\tilde{W}_t = Z_t^{-1} (\sigma^{-2} B_t X_t) dY_s$$

ce qui conclut la preuve. ■

Nous avons alors la caractérisation suivante pour le filtre non-conditionnelle Π .

Théorème 7.4.5 *Le filtre non-conditionnel $\Pi_t = \tilde{\mathbb{E}} [Z_t^{-1} X_t | \mathcal{F}_t^Y]$, $t \in [0, T]$, satisfait:*

$$\Pi_t = \Pi_0 + \int_0^t Q \Pi_s ds + \int_0^t \sigma^{-2} \Pi_s B_s dY_s \quad (7.14)$$

Preuve. Une représentation équivalent pour X_t est donnée par:

$$X_t = X_0 + \int_0^t Q X_s ds + \int_0^t dX_s$$

où M est une martingale à sauts. La règle de multiplication donne alors:

$$Z_t^{-1} X_t = Z_0 X_0 + \int_0^t Q Z_s^{-1} X_s ds + \int_0^t \sigma^{-2} B_s X_s Z_s^{-1} X_s Y_s ds$$

Soit par passage à l'espérance conditionnelle:

$$\begin{aligned} \mathbb{E} [Z_t^{-1} X_t | \mathcal{F}_t^Y] &= \mathbb{E} [Z_0 X_0 | \mathcal{F}_t^Y] + \mathbb{E} \left[\int_0^t Q Z_s^{-1} X_s ds | \mathcal{F}_t^Y \right] \\ &\quad + \mathbb{E} \left[\int_0^t \sigma^{-2} B_s X_s Z_s^{-1} X_s Y_s ds | \mathcal{F}_t^Y \right] \end{aligned} \quad (7.15)$$

une application du théorème de Fubini permet d'interchanger les signes espérance et intégration dans 7.15 ,ce qui donne le résultat attendu. ■

Nous nous intéressons à présent à la dérivation de Λ_t^{-1} où $\Lambda_t = \mathbb{E} [Z_T | \mathcal{F}_t^Y]$.

Lemme 7.4.6 *Soit Z_t , $t \in [0, T]$ donné par 7.10. Alors:*

$$\Lambda_t^{-1} = \tilde{\mathbb{E}} [Z_t^{-1} | \mathcal{F}_t^Y] = 1 + \int_0^t \sigma^{-2} B_s \Pi_s dY_s \quad (7.16)$$

Preuve. Une application de la formule de Bayes conditionnelle pour $P_0 = \tilde{P}$ et $V = Z_t^{-1}$ donne:

$$\tilde{\mathbb{E}} [Z_t^{-1} | \mathcal{F}_t^Y] = (\mathbb{E} [Z_t | \mathcal{F}_t^Y])^{-1}$$

Par suite, $\Lambda_t^{-1} = \tilde{\mathbb{E}} [Z_t^{-1} | \mathcal{F}_t^Y]$. Par ailleurs, d'après le Lemme 7.4.4, Z_t^{-1} satisfait:

$$Z_t^{-1} = 1 + \int_0^t Z_s^{-1} \sigma^{-2} B_s X_s dY_s$$

soit par passage à l'espérance conditionnelle:

$$\tilde{\mathbb{E}} [Z_t^{-1} | \mathcal{F}_t^Y] = 1 + \tilde{\mathbb{E}} \left[\int_0^t Z_s^{-1} \sigma^{-2} B_s X_s dY_s | \mathcal{F}_t^Y \right]$$

et par application du Lemme 6.3.9, il vient:

$$\Lambda_t^{-1} = 1 + \int_0^t \sigma^{-2} B_s \tilde{\mathbb{E}} [Z_s^{-1} X_s | \mathcal{F}_s^Y] dY_s = 1 + \int_0^t \sigma^{-2} B_s \Pi_s dY_s$$

ce qui conclut la preuve. ■

Nous concluons en dérivant la représentation du filtre π .

Théorème 7.4.7 *Le filtre conditionnel $\pi_t = \mathbb{E} [X_t | \mathcal{F}_t^Y]$, $t \in [0, T]$, satisfait:*

$$\pi_t = \pi_0 + \int_0^t Q \pi_s ds + \int_0^t \sigma^{-1} B_s \pi_s (1 - \pi_s) d\bar{W}_s \quad (7.17)$$

où le processus:

$$\bar{W}_t = \tilde{W}_t - \int_0^t \sigma^{-1} B_s \pi_s ds \quad (7.18)$$

est une (\mathcal{F}^Y, P) -mouvement brownien, dénommé le processus des innovations.

Preuve. Il s'agit d'une application directe de la règle de multiplication. ■

Remarque 7.4.8 *Comme $Y_t = \sigma \tilde{W}_t$, 7.18 se réécrit:*

$$\bar{W}_t = \sigma^{-1} Y_t - \int_0^t \sigma^{-1} B_s \pi_s ds$$

Soit, depuis 7.17, il vient:

$$\pi_t = \pi_0 + \int_0^t Q \pi_s ds + \int_0^t \sigma^{-2} B_s \pi_s (1 - \pi_s) (dY_s - B_s \pi_s ds)$$

8

Algorithme Espérance-Maximisation

L'algorithme Espérance-Maximisation (*expectation-maximization*) ou EM est un procédé générale pour déterminer l'estimateur du maximum de vraisemblance des paramètres d'une distribution de probabilité. Il se trouve par ailleurs particulièrement adapté dans le cadre de l'inférence statistique pour des modèles partiellement observés car il permet de séparer les étapes de filtrage et d'estimation, ce qui conduit à une procédure itérative efficace.

8.1. Cadre

Soit P^θ une famille de distributions toutes équivalentes à P^{θ^*} pour tout $\theta, \theta^* \in \Theta$ avec $\theta \neq \theta^*$. Soit par ailleurs $L_{Y,X}^{\mathcal{F}}(\alpha)$ la fonction de vraisemblance jointe du couple partiellement observé de variables aléatoires (Y, X) dépendant de α et restreinte à la filtration $\mathcal{F} = \mathcal{F}_T^Y$. L'estimateur du maximum de vraisemblance $\hat{\theta}$ de θ est alors défini par le problème d'optimisation suivant:

$$\hat{\theta} = \arg \max_{\theta \in \Theta} \log L_{(Y,X)}^{\mathcal{F}}(\theta)$$

ce qui est équivalent à maximiser la log-vraisemblance conditionnelle:

$$\log L_{(Y,X)}^{\mathcal{F}}(\theta) = \mathbb{E}^{P^\theta} [\log L_{(Y,X)}(\theta) | \mathcal{F}_T^Y]$$

L'algorithme EM procède en deux moments. Au cours d'une première étape, il s'agit de déterminer la valeur espérée de la log-vraisemblance des observations étant donné la filtration observée et les estimations présentes des paramètres. Ainsi, à la $(n+1)$ -ème étape de l'algorithme, étant donné les estimations θ^n , nous calculons:

$$Q(\theta, \theta^n) = \mathbb{E}^{P^{\theta^n}} \left[\log \frac{L_{(Y,X)}(\theta)}{L_{(Y,X)}(\theta^n)} | \mathcal{F}_T^Y \right] \quad (8.1)$$

où θ sont les nouveaux paramètres que nous optimisons et qui doivent être tels qu'ils augmentent la valeur de Q . L'évaluation (analytique ou numérique) de cette espérance consiste en l'étape E de l'algorithme.

La seconde étape, nommée M, consiste à maximiser l'espérance 8.1 et à déterminer la nouvelle estimation optimisée, soit:

$$\theta^{n+1} = \arg \max_{\theta \in \Theta} Q(\theta, \theta^n)$$

où le maximum de θ est obtenu comme la solution de:

$$\frac{\partial}{\partial \theta} Q(\theta, \theta^n) = \frac{\partial}{\partial \theta} \mathbb{E}^{P^{\theta^n}} \left[\log \frac{L_{(Y,X)}(\theta)}{L_{(Y,X)}(\theta^n)} | \mathcal{F}_T^Y \right] = 0 \quad (8.2)$$

Ces deux étapes sont répétées autant de fois que nécessaires, c'est-à-dire jusqu'à ce que la différence entre deux estimations successives de θ soit suffisamment faible pour pouvoir considérer que le maximum (local) de la fonction de vraisemblance a été atteint. Chaque itération garantit par ailleurs que la fonction de log-vraisemblance est bien augmentée et l'algorithme est assurée de converger vers un maximum local. En effet, comme:

$$L_{(Y,X)}(\theta) = \log \frac{dP_\theta^Y}{dP_{\theta^*}^Y} \text{ et } \frac{dP_\theta^Y}{dP_{\theta^*}^Y} = \mathbb{E}^{P^{\theta^*}} \left[\frac{dP_\theta}{dP_{\theta^*}} | \mathcal{F}_T^Y \right]$$

alors:

$$L_{(Y,X)}(\theta^n) - L_{(Y,X)}(\theta) = \log \frac{dP_{\theta^n}}{dP_\theta} = \log \mathbb{E}^{P^\theta} \left[\frac{dP_{\theta^n}}{dP_\theta} | \mathcal{F}_T^Y \right]$$

et par l'inégalité de Jensen:

$$L_{(Y,X)}(\theta^n) - L_{(Y,X)}(\theta) \geq \mathbb{E}^{P^\theta} \left[\frac{dP_{\theta^n}}{dP_\theta} | \mathcal{F}_T^Y \right]$$

l'égalité n'étant vraie que dans le cas où $\theta^n = \theta$.

Remarque 8.1.1 *S'il est possible de permuter les signes dérivée et espérance dans 8.2, alors la solution de θ vérifie:*

$$\mathbb{E}^{P^{\theta^n}} \left[\frac{\partial}{\partial \theta} \log \frac{L_{(Y,X)}(\theta)}{L_{(Y,X)}(\theta^n)} | \mathcal{F}_T^Y \right] = 0$$

8.2. Inférence de Modèles Latents

Nous considérons le système d'équations différentielles stochastiques suivant:

$$\begin{aligned} \frac{dS_t}{S_t} &= \mu_t dt + \sigma dW_t^2, \quad S_0 = s_0 \\ d\mu_t &= \alpha(v - \mu_t) dt + \beta dW_t^1, \quad \mu_0 \in \mathbb{R} \end{aligned}$$

où W_t^1 et W_t^2 sont des mouvement browniens indépendants, $\sigma > 0$ et $\alpha > 0, \beta > 0$ et v sont des paramètres inconnus à estimer. Par ailleurs, le processus μ est non-observable, tandis que X est observé sur $t \in [0, T]$. Nous appliquons le changement de variable suivant:

$$X_t = \frac{1}{\beta} (\mu_t - \mu_0), \quad Y_t = \frac{1}{\sigma} \left(\log \frac{S_t}{S_0} \right) + \frac{\sigma}{2} t$$

Alors, le processus latent devient:

$$dX_t = \alpha(\delta - X_t) dt + dW_t^1, \quad X_0 = 0$$

avec $\delta = v/\beta$ et le processus observé vérifie:

$$dY_t = (\nu + \lambda X_t) dt + dW_t^2, \quad Y_0 = 0$$

où $\lambda = \beta/\sigma$ et $\nu = \mu_0/\sigma$. Soit alors $\theta = (a, \delta, \lambda, \nu)$ le vecteur de paramètres à estimer. Le but de cet exemple est de produire une estimation de θ à partir des observations (Y_0, \dots, Y_T) ; soit alors la filtration $\mathcal{F}_T^Y = \sigma(Y_s, s \in [0, T])$.

Remarque 8.2.1 Cette transformation a permis d'éliminer les paramètres de volatilité (σ, β) des équations satisfaites par le couple (Y, X) et par ailleurs de faire en sorte que la loi du processus $(Y_t)_{t \leq T}$ soit équivalente à celle du mouvement brownien.

8.2.1 Inférence Statistique

Afin de pouvoir utiliser l'algorithme EM, il est nécessaire de calculer la vraisemblance conditionnelle des observations. A cette fin, considérons le couple $(Y, X) = (Y, X_t)_{t \geq 0}$ comme des observations sous la probabilité P paramétrée par $\theta = (\alpha, \delta, \lambda, \nu)$. Nous souhaitons à présent savoir sous quelle nouvelle mesure, notée P^n , donnée par le nouveau jeu de paramètres $\theta_n = (\alpha_n, \delta_n, \lambda_n, \nu_n)$, nos observations sont les plus probables (vraisemblables). Le rapport de vraisemblance est donnée par la dérivée de Radon-Nikodym:

$$\frac{dP^n}{dP} \Big|_{\mathcal{F}_T^Y} = L_T \quad (8.3)$$

et notant $\tau_n(X_t) = \nu_n + \lambda_n X_t$ et $v_n(X_t) = \alpha_n(\delta_n - X_t)$, nous avons:

$$\begin{aligned} L_T = & \exp \left(\int_0^T (\tau_n(X_t) - \tau(X_t)) dW_t - \frac{1}{2} \int_0^T (\tau_n(X_t) - \tau(X_t))^2 dt \right) \\ & \times \exp \left(\int_0^T (v_n(X_t) - v(X_t)) dB_t - \frac{1}{2} \int_0^T (v_n(X_t) - v(X_t))^2 dt \right) \end{aligned}$$

Soit:

$$\begin{aligned} \log \frac{dP^n}{dP} = & \left\{ \int_0^T (\nu_n + \lambda_n X_t) dY_t - \frac{1}{2} \int_0^T (\nu_n + \lambda_n X_t)^2 dt \right\} \\ & + \left\{ \int_0^T \alpha_n (\delta_n - X_t) dX_t - \frac{1}{2} \int_0^T \alpha_n^2 (\delta_n - X_t)^2 dt \right\} + f(\theta) \end{aligned}$$

où $f(\theta)$ est un terme qui ne dépend pas de θ_n .

Remarque 8.2.2 (Equivalence des mesures) L'algorithme EM en temps continu ne peut pas être utilisé pour estimer les paramètres de volatilité. En effet, si (σ, β) entraînent dans le calcul du ratio de vraisemblance 8.3, alors à chaque mise à jour de ces paramètres, de (σ, β) à (σ_n, β_n) , les mesures ne seraient plus équivalentes car les conditions du théorème de Girsanov ne seraient plus satisfaites, ce qui résulte en une perte d'information.

Par suite, la maximisation de la fonction de vraisemblance $\mathbb{E} [\log L_T | \mathcal{F}_T^Y]$ donne:

$$\mathbb{E} \left[\frac{\partial}{\partial \alpha_n} l_T | \mathcal{F}_T^Y \right] = \mathbb{E} \left[\int_0^T (\delta_n - X_t) dX_t - \int_0^T \alpha_n (\delta_n - X_t)^2 dt | \mathcal{F}_T^Y \right] = 0 \quad (8.4)$$

$$\mathbb{E} \left[\frac{\partial}{\partial \delta_n} l_T | \mathcal{F}_T^Y \right] = \mathbb{E} \left[\alpha_n X_T - \int_0^T \alpha_n^2 (\delta_n - X_t) dt | \mathcal{F}_T^Y \right] = 0 \quad (8.5)$$

et:

$$\mathbb{E} \left[\frac{\partial}{\partial \nu_n} l_T | \mathcal{F}_T^Y \right] = \mathbb{E} \left[Y_T - \int_0^T (\nu_n + \lambda_n X_t) dt | \mathcal{F}_T^Y \right] = 0 \quad (8.6)$$

$$\mathbb{E} \left[\frac{\partial}{\partial \lambda_n} l_T | \mathcal{F}_T^Y \right] = \mathbb{E} \left[\int_0^T X_t dY_t - \int_0^T (\nu_n + \lambda_n X_t) X_t dt | \mathcal{F}_T^Y \right] = 0 \quad (8.7)$$

Posons les notations suivantes:

$$m(t, T) = \mathbb{E} [X_t | \mathcal{F}_T^Y], \quad n(t, T) = \mathbb{E} [X_t^2 | \mathcal{F}_T^Y] \quad (8.8)$$

et comme $\int_0^T X_t dX_t = \frac{1}{2} X_T^2 - \frac{T}{2}$, alors 8.4 devient:

$$\begin{aligned} \delta_n m(T, T) - \frac{1}{2} n(T, T) + \frac{T}{2} - \alpha_n \delta_n^2 T \\ + 2\alpha_n \delta_n \int_0^T m(t, T) dt - \alpha_n \int_0^T n(t, T) dt = 0 \end{aligned} \quad (8.9)$$

et 8.5 donne:

$$m(T, T) - \alpha_n \delta_n T + \alpha_n \int_0^T m(t, T) dt = 0 \quad (8.10)$$

En multipliant 8.10 par δ_n et en soustrayant le résultat de 8.9 produit:

$$-\frac{1}{2} n(T, T) + \frac{T}{2} + \alpha_n \delta_n \int_0^T m(t, T) dt - \alpha_n \int_0^T n(t, T) dt = 0 \quad (8.11)$$

De 8.10 et 8.11, nous dérivons alors les valeurs mises à jour de (α_n, δ_n) , soit:

$$\alpha_n = \frac{-\frac{T}{2} n(T, T) + \frac{T^2}{2} + m(T, T) \int_0^T m(t, T) dt}{T \int_0^T n(t, T) dt - \left(\int_0^T m(t, T) dt \right)^2} \quad (8.12)$$

$$\delta_n = \frac{m(T, T)}{T \alpha_n} + \frac{1}{T} \int_0^T m(t, T) dt \quad (8.13)$$

De la même façon, 8.6 peut se réécrire:

$$Y_T - \nu_n T + \lambda_n \int_0^T m(t, T) dt = 0$$

et 8.7 comme:

$$\mathbb{E} \left[\int_0^T X_t dY_t | \mathcal{F}_T^Y \right] - \nu_n \int_0^T m(t, T) dt + \lambda_n \int_0^T n(t, T) dt = 0$$

qui peut être simplifié grâce au lemme suivant.

Lemme 8.2.3 *Nous avons:*

$$\begin{aligned} \mathbb{E} \left[\int_0^T X_t dY_t \right] &= \int_0^T m_t dY_t - \nu \cdot \left[\int_0^T (m_t - m(t, T)) dt \right] \\ &\quad - \lambda \cdot \left[\int_0^T (m_t^2 + m_t m(t, T)) dt \right] \end{aligned}$$

Preuve. Comme:

$$\bar{W}_t = W_t + \int_0^t \lambda (X_s - m_s) ds$$

est un mouvement brownien, nous avons alors:

$$d\bar{W}_t = dY_t - (\nu + \lambda m_t) dt$$

Dès lors:

$$\begin{aligned} \mathbb{E} \left[\int_0^T X_t dY_t \right] &= \mathbb{E} \left[\int_0^T X_t d\bar{W}_t + \int_0^T (\nu + \lambda m_t) X_t dt \right] \\ &= \mathbb{E} \left[\int_0^T X_t d\bar{W}_t \right] + \int_0^T (\nu + \lambda m_t) m(t, T) dt \\ &= \int_0^T m_t d\bar{W}_t + \int_0^T (\nu + \lambda m_t) m(t, T) dt \end{aligned}$$

ce qui est, après substitution, le résultat annoncé. ■

Nous obtenons alors les valeurs mises à jour de (ν_n, λ_n) , soit:

$$\nu_n = \frac{1}{T} Y_T - \frac{\lambda_n}{T} \int_0^T m(t, T) dt \quad (8.14)$$

$$\lambda_n = \frac{T \int_0^T m_t dY_t - Y_T \int_0^T m_t dt}{T \int_0^T (m_t m(t, T) + n(t, T) - m_t^2) dt - \int_0^T m_t dt \int_0^T m(t, T) dt} \quad (8.15)$$

Remarque 8.2.4 (Filtrage) *Le calcul des estimations 8.12, 8.13, 8.14, 8.15 nécessite la connaissance des filtres et lisseurs suivants:*

$$m_t = \mathbb{E} [X_t | \mathcal{F}_t^Y], \quad m(t, T) = \mathbb{E} [X_t | \mathcal{F}_T^Y], \quad n(t, T) = \mathbb{E} [X_t^2 | \mathcal{F}_T^Y]$$

qui ont été précédemment dérivés dans les Exemples 6.3.12 et 6.3.15.

8.3. Estimation d'une Chaîne de Markov

Nous reprenons le modèle développé en Section 7.4, soit:

$$dY_t = \mu_t dt + \sigma dW_t$$

où μ est une chaîne de Markov telle que:

$$\mu_t = BX_t$$

avec Q sa matrice de transition. Comme σ peut être obtenu sans biais à partir des données, le problème d'estimation se réduit à celui de l'inférence du vecteur B et de la matrice Q . A cette fin, nous introduisons les quantités suivantes: le temps d'occupation O^i dans l'état e_i , le nombre de sauts N^{ij} de l'état e_i à l'état e_j , $i \neq j$, et l'intégrale de niveau G^i , soit:

$$O_t^i = \int_0^t X_s^i ds, \quad N_t^{ij} = \int_0^t X_{s-}^i dX_s^j, \quad G_t^i = \int_0^t \sigma^{-2} X_s^i dY_s \quad (8.16)$$

Comme dans la Section 8.2, nous considérons O^i, N^{ij}, G^i comme des observations sous la probabilité P paramétrée par B et Q et nous voulons déterminer sous quelle nouvelle mesure, notée P^* , donnée par les nouveaux paramètres B^* et Q^* , ces observations sont les plus vraisemblables. Sur l'intervalle $[0, t]$, la mesure P^* est donnée par:

$$\frac{dP^*}{dP} |_{\mathcal{F}_t} = L_t = L_t^Q L_t^B$$

où le changement de probabilité de Q à Q^* s'écrit:

$$L_t^B = \exp \left(\int_0^t \sum_{i,j=1, i \neq j}^d (Q_{ij} - Q_{ij}^*) X_s^i ds \right) \prod_{i,j=1, i \neq j}^d \left(\frac{Q_{ij}^*}{Q_{ij}} \right)^{N_t^{ij}}$$

et celui de B à B^* satisfait:

$$L_t^B = \exp \left(\int_0^t \sigma^{-2} (B_s^* - B_s) X_s (dY_s - B_s X_s ds) - \frac{1}{2} \int_0^t \sigma^{-2} (B_s^* - B_s)^2 X_s^2 ds \right)$$

La maximisation de la fonction de vraisemblance $\mathbb{E} [\ln L_t | \mathcal{F}_t^Y]$ donne alors:

$$Q_{ij}^* = \frac{\Pi_t(N^{ij})}{\Pi_t(O^i)}, \quad B_i^* = \sigma^2 \frac{\Pi_t(G^i)}{\Pi_t(O^i)} \quad (8.17)$$

où $\Pi_t(X) = \tilde{\mathbb{E}} [Z_t^{-1} X_t | \mathcal{F}_t^Y]$ est le filtre non-conditionnel, cf. Théorème 7.4.5.

Remarque 8.3.1 Les estimations 8.17 peuvent être indifféremment exprimées en fonction de π ou de Π . Toutefois, il n'est pas possible en général d'obtenir une solution pour $\pi_t(HX)$ quand $H_t = \{1, O^i, N^{ij}, G^i\}$, tandis qu'il est possible d'explicitement $\Pi_t(HX)$ et on note alors:

$$\Pi_t(H) = \mathbf{1}\Pi(HX)$$

Index

- équation
 - aux dérivées partielles, 92
 - différentielle stochastique, 87
 - à sauts, 89
- actif
 - conditionnel, 83
- algorithme
 - espérance-maximisation, 151
- calcul stochastique, 32
- changement
 - de mesure, 77
- Chapman-Kolmogorov
 - équation de, 48
- covariance
 - quadratique
 - semimartingale, 66
- discrétisation, 99, 104
 - du temps, 100
- distribution
 - indéfiniment divisible, 45
- Doléans-Dade
 - exponentielle de, 78
- espérance
 - conditionnelle, 21
- espace
 - de probabilité, 14
 - filtré, 20
- Euler
 - schéma d', 104
 - pour une diffusion, 104
- exponentielle
 - stochastique, 77
- extrapolation
 - cas continu, 141
 - cas discret, 127
- Feynman-Kac
 - théorème actualisé de, 93
 - processus mixte, 95
- théorème de, 92
 - multidimensionnel, 94
 - processus mixte, 95
- filtrage, 21, 121
 - cas continu, 128
 - cas discret, 121
 - probabilité de référence, 131
 - théorème fondamental, 132
 - théorie des innovations, 129
- filtration, 18, 19
- filtre
 - cas continu, 137
 - cas discret, 124
- fonction
 - aléatoire, 18
 - càdlàg, 18
 - càglàd, 19
 - caractéristique, 16
 - constante par morceaux, 73
 - discontinue, 18
 - génératrice des moments, 17
 - payoff, 83
 - simple, 62
- formule d'Itô
 - mesure de Poisson, 74
 - mouvement brownien, 68
 - processus d'Itô, 71
 - processus mixte, 75
- générateur infinitésimal, 48
- Girsanov
 - théorème de
 - processus d'Itô, 79
 - processus de Poisson, 80
 - processus de Poisson composé, 81
 - processus de Wiener, 79
 - processus mixte, 82

- inférence statistique, 111
- innovations
 - cas continu, 129
 - cas discret, 125
- intégrale
 - d'Itô, 53
 - stochastique, 53, 59
 - compensée, 63
- interpolation
 - cas continu, 138
 - cas discret, 126
- isométrie
 - intégrale d'Itô, 56
 - intégrale de Lévy, 63
- Jensen
 - inégalité de, 22
- Kalman-Bucy
 - filtre de, 125
- kurtosis, 27
- Lévy
 - théorème de, 35
 - multidimensionnel, 36
- lissage
 - cas continu, 138
 - cas discret, 126
- lisseur
 - cas continu, 140
 - cas discret, 126
- loi
 - de Poisson, 38
 - des espérances itérées, 22
 - exponentielle, 37
 - propriété d'absence de mémoire, 37
 - log-normale, 25
 - normale, 25
- méthode
 - des moments généralisés, 113
 - du maximum de vraisemblance, 111
- marche aléatoire, 27, 45
- Markov
 - chaîne de, 155
- martingale, 22
 - de carré intégrable, 23
 - théorème de représentation, 129
- mesure, 12
 - équivalence, 14, 77
 - absolue continuité, 13
 - aléatoire, 43
 - de Lévy, 46
 - de Lebesgue, 13
 - de Poisson, 44
 - de Poisson compensée, 44
 - de probabilité, 13, 14
 - de Radon, 13
 - de sauts, 43
 - martingale équivalente, 86
 - masse, 13
 - risque-neutre, 86
- Milstein
 - schéma de, 104
 - pour une diffusion, 106
- modèle
 - à régimes, 146
 - latent, 152
- numéraire, 83
- option, 83
- prédicteur
 - cas continu, 141
 - cas discret, 128
- prédiction
 - cas continu, 141
 - cas discret, 127
- probabilité
 - de référence, 147
- processus
 - à marques, 44
 - d'Itô, 70
 - de comptage, 39
 - de Lévy, 45
 - décomposition, 47
 - propriété de Markov, 47
 - de Markov, 23
 - de Poisson, 39
 - compensé, 40
 - simulation à dates fixes, 102
 - simulation aux instants de sauts, 103
 - de Poisson composé, 41
 - compensé, 42

- de Wiener, 30
 - multidimensionnel, 35
 - probabilité de transition, 34
 - propriété de Markov, 34
 - simulation, 100
 - variation quadratique, 32
- des innovations, 130
 - cas continu, 129
 - cas discret, 125
- mixte, 89
 - simulation, 103
- simple, 54, 60
- stochastique, 18
 - adapté, 21
 - histoire d'un, 21
 - latente, 121
 - mesurable, 21
- variation quadratique, 28
 - intégrale d'Itô, 56
 - processus d'Itô, 70
 - semimartingale, 64
- volatilité
 - réalisée, 28, 33, 64
- vraisemblance, 111
- Wonham
 - filtre de, 146
- règle
 - d'évaluation, 84
 - de composition, 60
 - de dérivation en chaîne, 67
 - de multiplication, 66
- Radon-Nikodym
 - dérivée de, 13, 77
 - processus de dérivée de, 77
 - théorème de, 13
- Riemann
 - intégrale de, 68
 - somme de, 61
 - suite de, 64, 66
- semimartingale, 60
- σ -additivité, 11
- σ -algèbre, 12
- skewness, 27
- temps
 - aléatoire, 22, 39
- théorème
 - de la limite centrée, 29
 - de normalité a priori - a posteriori, 122
 - de représentation des martingales, 129
 - de Radon-Nikodym, 13
 - des corrélations normales, 123
- variable
 - aléatoire, 15
 - histoire d'une, 20