

RÉPUBLIQUE DU CAMEROUN

Paix-Travail-Patrie

UNIVERSITÉ DE YAOUNDÉ 1

ÉCOLE NORMALE SUPÉRIEURE

DÉPARTEMENT DE MATHÉMATIQUES



REPUBLIC OF CAMEROON

Peace-Work-Fatherland

UNIVERSITY OF YAOUNDE 1

HIGHER TEACHER'S TRAINING COLLEGE

DEPARTMENT OF MATHEMATICS

ESTIMATION DES PARAMETRES DANS LES EDS PAR L'APPROXIMATION HERMITE-SPLINE-CUBIQUE

Mémoire rédigé en vue de l'obtention du diplôme de professeur de
l'Enseignement Secondaire deuxième grade (D.I.P.E.S II) en
mathématiques

Par

ONANA ASSOUGA Florent

Matricule : 16Y554

Licence en Mathématiques

Sous la direction de :

Dr TAKAM SOH Patrice

Chargé de Cours

Année Académique : 2020-2021

Dédicace

Ce mémoire est dédié à mon père :

ASSOUGA ASSOUGA TELESPHORE

Remerciements

*P*arvenu au terme de ce travail, je remercie en premier l'Eternel Dieu tout puissant et lui rends grâce pour toute sa miséricorde et ses grâces pour m'avoir donné santé, courage et audace pour surpasser les difficultés.

*Q*u'il me soit permis d'exprimer ma gratitude à mon directeur de mémoire : **Dr TAKAM SOH Patrice** pour avoir accepté de diriger ce travail, pour son soutien multiforme, pour sa disponibilité et son encadrement constant dans nos premiers pas de recherche.

*J*e remercie les enseignants de l'**Ecole Normale Supérieure de Yaoundé**, en particulier ceux du département de Mathématiques, pour leurs enseignements et leur accompagnement tout au long de notre formation académique et sociale.

*C*e travail est avant tout le fruit d'énormes sacrifices familiaux et je veux adresser des remerciements spéciaux à la famille et particulièrement à :

♡ Ma grand mère NDZIE ATANGANA Françoise pour l'éducation qu'elle a su nous donner, pour son amour sans faille et son support inconditionnel.

♡ Mes oncles MINFEGUE ASSOUGA Calvin, ONGUENE ASSOUGA Hilaire pour l'abnégation dont ils ont fait preuve pour nous garantir une scolarisation allant jusqu'à faire des sacrifices leur portant préjudice sur leurs propres objectifs ; aussi à OKOME OUMAROU Dominique pour le rôle de mère qu'elle a toujours su jouer à notre endroit.

♡ Mes frères et soeurs : ASSOUGA ELEBE Roland, NDZIE Emmanuel, FEUMENI Brayan, BELINGA ASSOUGA Armand, Rachelle, Marina NOUGA, Kengne Angés.

♡ Mes oncles et tantes : NGA ASSOUGA Marie Sylvie, NDZIE ASSOUGA Mirice, Marie Thérèse NOUGA, l'abbé Simon Pierre BITEGA.

♡ Mes amis : Sobjio Ludane, Tsogmo Cybèle, Landry Eteme, Nyemeck Stephane, Tsafac Cyrille, Nyambongo Arlette, Nicaise Ndel, Cabrel Ndontsop, Abdoul, Kihdze, Faissal, Djala, Fezeu, Ndayou, Ngassa, Nsen, Samiha, Tedonmegne, Ketcha, Nguefack, Kamdem.

Que tous ceux dont les noms ne sont pas mentionnés et qui ont contribué à la réalisation de ce travail, trouvent ici l'expression de ma profonde gratitude.

Déclaration sur l'honneur

Le présent travail est une œuvre originale du candidat et n'a été soumis nulle part ailleurs, en partie ou en totalité, pour une autre évaluation académique. Les contributions externes ont été dûment mentionnées et recensées en bibliographie

Signature du candidat

ONANA ASSOUGA Florent

Résumé

Il est question de proposer une approche d'estimation des paramètres dans les EDS dans le cas où la loi de transition n'est pas connue de façon exacte.

Pour cela, nous allons approximer cette densité de transition par l'expansion polynomiale d'Hermite qu'a proposé Aït-Sahalia puis la coupler avec l'approximation spline-cubique connue dans la littérature. A partir des approximations faites, les paramètres vont être obtenus en maximisant la vraisemblance approximée.

L'approche ainsi proposée va être appliquée sur un jeu de données simulées.

Mots clés :

vraisemblance, EDS, expansion polynomiale d'Hermite, Densité de transition, Spline-cubique.

Abstract

It is question of proposing an approach of estimation of the parameters of SDE in the case where the transition density in not known in an exact way.

For that, we will approximate this transition density by the series of Hermite polynomial expansion developed by Aït-Sahalia then couple it with the spline-cubic approximation well known in literature. From the approximations made, the parameters will be obtained by maximizing the approximate likelihood.

The approach thus proposed will be applied to a set simulated data.

Keywords :

likelihood, SDE, Hermite polynomial expansion, Transition density, Spline-cubic

Table des matières

Dédicaces	i
Remerciements	ii
Déclaration sur l'honneur	iii
Résumé	iv
Abstract	v
Table des figures	viii
Introduction générale	1
1 RAPPELS SUR LA THEORIE DES EDS	4
1.1 Processus stochastique en temps continu	4
1.1.1 Définition et exemple	4
1.1.2 Filtration-Adaptabilité et Prédictabilité	5
1.1.3 Martingale en temps continu	5
1.1.4 Le mouvement Brownien et ses propriétés	6
1.2 Processus d'Itô et formule d'Itô	7
1.3 Equations différentielles stochastiques	8
1.4 Processus de diffusion	9
1.4.1 Définition	9
1.4.2 Notion de générateur infinitésimal	11
1.4.3 Equations aux dérivées partielles de Fokker-Planck-Kolmogorov	11
2 METHODES D'ESTIMATION DES PARAMÈTRES DANS LES EDS PAR MAXIMUM DE VRAISEMBLANCE	13

2.1	Le cadre	13
2.2	Cas de vraisemblance exacte	16
2.2.1	Modèle 1 : Le modèle de Black-Scholes ou Mouvement Brownien Géométrique	16
2.2.2	Modèle 2 : Modèle Ornstein-Uhlenbeck ou Vasicek process	20
2.2.3	Modèle 3 : Le Modèle de Cox-Ingersoll-Ross	24
2.2.4	Modèle 4 : Le modèle CEV	27
2.3	Cas de vraisemblance approximée	28
3	APPROXIMATION DE LA DENSITE DE TRANSITION	30
3.1	Approximation de la densité de transition selon Aït-Sahalia	30
3.1.1	Première étape : Transformation du processus X_t	31
3.1.2	Deuxième étape : Pseudo-normalisation du processus $(Y_t)_{t \in [0, T]}$	33
3.1.3	Troisième étape : Approximation de la densité de transition p_Z	34
3.1.4	Expression explicite de la densité de transition \tilde{p}_x^{herm}	37
3.2	Approximation de la densité de transition en utilisant l'approximation Spline-Cubique	41
3.2.1	Première étape : Partition de l'intervalle de l'intervalle d'observation $[a; b]$	41
3.2.2	Deuxième étape : Interpolation d'Hermite de $p_X^l(\Delta, \cdot x_{i-1}; \theta)$ sur $[a_l; a_{l+1}]$	42
3.2.3	Troisième étape : Choix des pentes $p_l \ l = 0; 1; \dots; L$	43
3.2.4	Expression de la densité de transition $\tilde{p}_X^{SPL}(\Delta, \cdot x_{i-1}; \theta)$	44
4	SIMULATIONS	45
4.1	Simulations des densités de transition	46
4.2	Simulation des données	48
4.3	Estimation des paramètres par bootstrap	49
	IMPLICATIONS PEDAGOGIQUES	53
	Conclusion et perspectives	55
	Bibliographie	55
	Annexe	57

Table des figures

1.1	Une trajectoire du MBS	6
3.1	Courbes des 7 premiers polynômes d'Hermite	35
4.1	Courbes des densités de transition	47
4.2	Densités de transition mensuelles	48
4.3	Une simulation de la trajectoire du processus d'Ornstein-Uhlenbeck	49
4	Code R d'une trajectoire du MBS	60
5	Code R d'une trajectoire du processus Ornstein-Uhlenbeck	61

Introduction générale

En statistique, on s'intéresse le plus souvent à l'estimation des paramètres présents dans un modèle. Ce modèle peut être :

i) soit sous forme de régression : à l'exemple de

- Les modèles économétriques (modèles linéaires simples et multiples, modèles linéaires généralisés) :

$$Y = X\beta + \varepsilon \text{ où } \beta \text{ est le vecteur de paramètres à estimer} \quad (1)$$

- Les modèles de panels ou effets mixtes :

$$Y = X\beta + \lambda + \alpha + \varepsilon \text{ où } \beta, \lambda, \alpha \text{ sont les vecteurs de paramètres à estimer} \quad (2)$$

ii) soit sous récurrente :

$$X_t = \beta X_{t-1} + \alpha. \quad (3)$$

iii) soit sous forme d'équation différentielle (ordinaire ou stochastique) : à l'exemple de

- Les modèles dynamiques : On a par exemple les systèmes différentiels de la forme

$$\begin{cases} \frac{\partial}{\partial t} f(t, \alpha) = F(f(t, \alpha), t, \alpha) \\ f(t_0, \alpha) = f_0(\alpha) \end{cases}, \quad (4)$$

où f décrit un processus biologique solution de cette équation différentielle.

- Les équations aux dérivées partielles (Méthodes mécano-stat)

$$\frac{\partial C}{\partial t} = \rho C(1 - C) + \Delta.(D\Delta C) \quad (5)$$

où C décrit un processus d'évolution spatio-temporelle des cellules tumorales dans le cerveau comme des ondes progressives et ρ est un paramètre à estimer à

partir des observations spatio-temporelles.

- Les équations différentielles stochastiques (Modèles dynamiques stochastiques) :

$$dY_t = f(t, Y_t, \theta)dt + g(t, Y_t, \theta)dW_t \quad (6)$$

où f et g sont des fonctions bien définies et W_t est un processus à temps continu et θ le vecteur des paramètres à estimer en fonction des observations de Y_t .

Pour estimer ces paramètres, on a besoin des observations. Elles peuvent être de nature exogènes (x_i, y_i) à l'exemple des modèles du type i) ou alors elles peuvent être des trajectoires des solutions : c'est le cas des EDO et EDS.

Dans les modèles évoqués plus haut, on a besoin d'estimer les paramètres pour mieux calibrer le modèle. En effet, si on veut représenter une trajectoire de la solution (dans les modèles sous forme d'équation) il faut que l'on dispose des valeurs suffisamment précises des paramètres ; et ces valeurs précises sont obtenues par estimation statistique.

Le cas des modèles sous formes d'EDS semble constituer un centre d'intérêt ces dernières années dans plusieurs domaines variés. Il est généralement étudié en finance lorsqu'on veut par exemple modéliser le taux d'intérêt à l'instar du modèle de Bachelier. Ce modèle a ensuite été amélioré par Black-Scholes(1973), Vasicek(1977) et a même cessé d'être seulement étudié en finance pour atteindre de nouveaux champs notamment en économie (Black-Scholes, Cox and al, Merton), en biologie (Mc Adams and Arkim 1997), en génétique (Kimura and Ohta 1971), en chimie (Obuhov 1959), et en ingénierie (Pardoux and Pignol 1984).

Une problématique de ces modèles réside toujours au niveau de leur calibration et ceci implique donc l'estimation de ces paramètres dans ces EDS. Pour cela, plusieurs méthodes sont proposées dans la littérature et ces méthodes dépendent du fait que la loi de la densité de transition du processus soit connue ou pas. Nous allons nous intéresser dans ce travail à estimer les paramètres dans de tels modèles principalement par la méthode du maximum de vraisemblance lorsque la densité de transition est inconnue. Nous présenterons deux méthodes d'approximation : celle proposée par Aït-Sahalia (1) et (2) et la seconde par l'approximation spline-cubique.

Notre travail est organisé en quatre parties. Premièrement nous ferons un bref rappel sur les notions de la théories des EDS. Deuxièmement nous présenterons les différentes méthodes d'estimation des paramètres par maximum de vraisemblance dans les EDS notamment dans les cas exacte et approximé, cette étude se fera par des exemples concrets dans des modèles très connus. Troisièmement, nous présenterons deux méthodes d'approximation des

densités de transition : celle décrite par par Aït-Sahalia en utilisant l'expansion polynomiale d'Hermite et une deuxième qui est celle de l'approximation par les splines cubiques. Enfin, pour illustrer nos résultats, nous ferons des simulations numériques à partir du processus d'Ornstein-Uhlenbeck.

RAPPELS SUR LA THEORIE DES EDS

Dans ce chapitre, nous présentons certains éléments clés et résultats de l'analyse stochastique qui nous serviront dans la suite de notre travail.

1.1 Processus stochastique en temps continu

1.1.1 Définition et exemple

Définition 1

Un **processus stochastique en temps continu** sur un espace mesurable (Ω, \mathcal{A}) à valeurs (ou espace d'état) dans \mathbb{R} et indexé par $I = [0, T]$ est une application $X : I \times \Omega \rightarrow \mathbb{R}$ qui est $\mathcal{B}([0, T]) \otimes \mathcal{A}$ -mesurable.

Remarque 1.1.1. 1 Dans cette définition, si $t \in [0, T]$ est fixé, on a $X_t : \Omega \rightarrow \mathbb{R}$ qui est une variable aléatoire. Donc généralement, un processus stochastique en temps continu est souvent défini comme une collection de variables aléatoires.

2 $\forall \omega \in \Omega$, on a d'après la définition, $X(\cdot, \omega) : [0, T] \rightarrow \mathbb{R}$ qui est une application $\mathcal{B}([0, T])$ -mesurable : on l'appelle *trajectoire* ou *chemin* du processus.

Exemple 1.1.1 (Construction canonique de Levy). Posons $\Omega = C^k([0, T], \mathbb{R})$ l'espace des applications de classe C^k sur l'intervalle $[0, T]$. Soit $\omega : [0, T] \rightarrow \mathbb{R}$ un élément de Ω .

Pour tout $t \in [0, T]$, définissons l'application coordonnée : $X_t : \Omega \rightarrow \mathbb{R}$ par $X_t(\omega) = \omega(t)$. Prenons $\mathcal{A} = \sigma(X_t, t \in [0, T])$ la plus petite tribu rendant mesurables toutes les variables aléatoires X_t . On obtient ainsi une collection $(X_t)_{t \in [0, T]}$ de variables aléatoires qui est un processus stochastique sur (Ω, \mathcal{A}) à valeurs dans \mathbb{R} dit : **processus canonique de Levy**.

Définition 2

Soit $(X_t)_{t \in [0, T]}$ un processus stochastique en temps continu sur un espace probabilisé $(\Omega, \mathcal{A}, \mathbb{P})$. On dit que le processus $(X_t)_{t \in [0, T]}$ est continu si ses trajectoires sont des applications continues $\mathbb{P} - ps$; c'est à dire :

$$\forall t \in [0, T] \quad \mathbb{P}(\{\omega \in \Omega, X(\cdot, \omega) \text{ est continue en } t\}) = 1$$

Exemple 1.1.2. *Le processus de Levy est un processus continu.*

1.1.2 Filtration-Adaptabilité et Prédictabilité**Définition 3**

Soit $(\Omega, \mathcal{A}, \mathbb{P})$ une espace probabilisé.

- 1** On appelle **filtration en temps continu**, toute collection $\mathbb{F} = (\mathcal{F}_t)_{t \in [0, T]}$ de sous σ -algèbres de \mathcal{A} vérifiant la propriété d'accroissement de l'information suivante :

$$\mathcal{F}_s \subseteq \mathcal{F}_t \quad \forall s, t \in [0, T] \quad \text{tels que } s < t.$$

On dit que l'espace probabilisé $(\Omega, \mathcal{A}, \mathbb{P})$ est filtré et on note $(\Omega, \mathcal{A}, \mathbb{F}, \mathbb{P})$.

- 2** On dit qu'un processus $(X_t)_{t \in [0, T]}$ en temps continu est adapté à une filtration $\mathbb{F} = (\mathcal{F}_t)_{t \in [0, T]}$ si :

$$\forall t \in [0, T] \quad X_t \text{ est } \mathcal{F}_t - \text{mesurable}$$

- 3** On dit qu'un processus $(X_t)_{t \in [0, T]}$ en temps continu est prévisible par rapport à une filtration $\mathbb{F} = (\mathcal{F}_t)_{t \in [0, T]}$ si :

$$\forall t \in [0, T] \quad X_t \text{ est } \mathcal{F}_s - \text{mesurable} \quad \forall s \leq t$$

Exemple 1.1.3. *Tout processus stochastique $(X_t)_{t \in [0, T]}$ admet une filtration par rapport à laquelle il est adapté. En effet il suffit de considérer la famille de sous σ -algèbre $\mathcal{F}_t^X = \sigma(X_s, s \leq t) \forall t \in [0, T]$*

1.1.3 Martingale en temps continu

Soit $(\Omega, \mathcal{A}, \mathbb{F}, \mathbb{P})$ un espace probabilisé filtré $\mathbb{F} = (\mathcal{F}_t)_{t \in [0, T]}$.

Définition 4

Un processus stochastique à temps continu $(M_t)_{t \in [0;T]}$ est une **martingale** sur l'espace probabilisé filtré $(\Omega, \mathcal{A}, \mathbb{F}, \mathbb{P})$ si :

- 1 $(M_t)_{t \in [0;T]}$ est intégrable, ie : $\forall t \in [0; T], \mathbb{E}(M_t) < +\infty$.
- 2 $(M_t)_{t \in [0;T]}$ est adapté à la filtration $\mathbb{F} = (\mathcal{F}_t)_{t \in [0;T]}$.
- 3 $\forall 0 \leq s < t \leq T, \mathbb{E}(M_t | \mathcal{F}_s) = M_s$.

1.1.4 Le mouvement Brownien et ses propriétés**Définition 5**

Soit $(W_t)_{t \in [0;T]}$ un processus stochastique à temps continu sur un espace probabilisé $(\Omega, \mathcal{A}, \mathbb{P})$. On dit que $(W_t)_{t \in [0;T]}$ est un mouvement Brownien standard (MBS) ou processus de Wiener standard si :

- 1 $W_0 = 0, \mathbb{P} - ps$;
- 2 Les accroissements du processus sont indépendants, ie, pour tout entier $n \geq 2$ et tous instants $0 < t_1 < t_2 < \dots < t_n$, les accroissements $W_{t_1}, W_{t_2} - W_{t_1}, \dots, W_{t_n} - W_{t_{n-1}}$ du processus sont indépendants ;
- 3 Pour tous temps $0 \leq s < t$, l'accroissement $W_t - W_s$ est une variable aléatoire Gaussienne centrée de variance $t - s$;
- 4 Les trajectoires $W(\cdot, \omega) : [0, t] \rightarrow \mathbb{R}, \omega \in \Omega$, sont continues $\mathbb{P} - ps$.

Ci-après, une illustration du graphe de simulation d'une telle trajectoire :

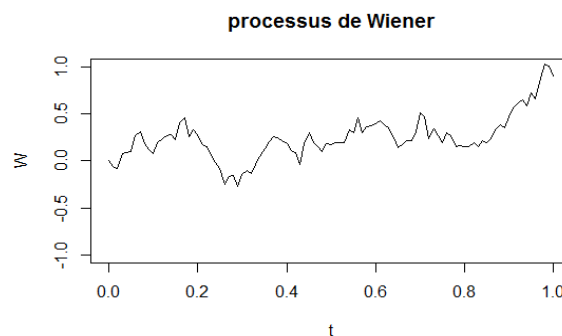


FIGURE 1.1 – Une trajectoire du MBS

Conséquence 6

Pour tous temps $0 \leq s < t$, l'accroissement $W_t - W_s$ du MBS est indépendant de la sous- σ -algèbre $\sigma(W_u, u \leq s)$.

Remarque 1.1.2. *Le mouvement Brownien est une martingale par rapport à sa filtration naturelle $\mathcal{F}_t^W = \sigma(W_s, s \leq t) \forall t \in [0, T]$. En effet :*

- $(\mathbb{E}(|W_t|))^2 \leq \mathbb{E}(|W_t|^2) = t < +\infty, \forall t \in [0, T]$.
- $\mathbb{E}(W_t | \mathcal{F}_s) = \mathbb{E}(W_t - W_s + W_s | \mathcal{F}_s) = \mathbb{E}(W_t - W_s | \mathcal{F}_s) + \mathbb{E}(W_s | \mathcal{F}_s) = 0 + W_s = W_s,$
 $\forall 0 \leq s < t \leq T$. Nous avons utilisé le fait que $W_t - W_s$ est indépendante de \mathcal{F}_s et aussi $\mathbb{E}(W_s | \mathcal{F}_s) = W_s$.

1.2 Processus d'Itô et formule d'Itô

Définition 7

Soit $(W_t)_{t \in [0; T]}$ un mouvement Brownien unidimensionnel sur un espace probabilisé $(\Omega, \mathcal{A}, \mathbb{P})$. Un processus d'Itô est un processus $(X_t)_{t \in [0; T]}$ sur $(\Omega, \mathcal{A}, \mathbb{P})$ de la forme :

$$X_t = X_0 + \int_0^t F_u du + \int_0^t G_u dW_u \quad (1.1)$$

où F_t et G_t sont deux processus adaptés à la filtration Brownienne appelés respectivement coefficient drift et coefficient de diffusion, satisfaisant :

$$\int_0^T |G_u|^2 du < +\infty \quad \text{et} \quad \int_0^T |F_u| du < +\infty \quad \mathbb{P} - ps \quad (1.2)$$

L'équation 1.1 peut encore s'écrire plus simplement :

$$dX_t = F_t dt + G_t dW_t \quad (1.3)$$

et l'équation 1.3 est appelée la différentielle stochastique du processus $(X_t)_{t \in [0; T]}$.

Théorème 8 (Lemme d'Itô)

Soit $(X_t)_{t \in [0;T]}$ un processus d'Itô. Si f est une fonction de classe $C^{1,2}([0, T] \times \mathbb{R})$, alors le processus $f(t, X_t)$ est également un processus d'Itô et sa différentielle stochastique est :

$$df(t, X_t) = \frac{\partial f(t, X_t)}{\partial t} dt + \frac{\partial f(t, X_t)}{\partial x} dX_t + \frac{1}{2} \frac{\partial^2 f(t, X_t)}{\partial x^2} d[X]_t \quad (1.4)$$

où $d[X]_t$ est la variation quadratique du processus $(X_t)_{t \in [0;T]}$ définie par :

$$d[X]_t = (dX_t).(dX_t) \quad (1.5)$$

avec la règle de multiplication suivante : $(dt)(dt) = (dt)(dW_t) = (dW_t)(dt) = 0$ et $(dW_t)(dW_t) = dt$

Démonstration. Voir (7) □

1.3 Equations différentielles stochastiques

Définition 9

- 1** Une **équation différentielle stochastique (EDS)** gouvernée par un mouvement Brownien unidimensionnel W_t est définie par la donnée d'une différentielle stochastique du type processus d'Itô avec condition initiale suivante :

$$\begin{cases} dX_t = b(t, X_t)dt + \sigma(t, X_t)dW_t \\ X_0 = x_0 \in \mathbb{R}, \quad 0 \leq t \leq T \end{cases}, \quad (1.6)$$

où $b(t, X_t)$ et $\sigma(t, X_t)$ sont deux processus stochastiques dépendant du temps t et de l'inconnue X_t .

- 2** Un processus X_t est une solution forte de l'EDS 1.6 si il est adapté par rapport à la filtration naturelle du mouvement Brownien W_t , si $X_0 = x_0$ et s'il existe deux processus $b(t, X_t)$ et $\sigma(t, X_t)$ satisfaisant les conditions 1.2 telle :

$$X_t = x_0 + \int_0^t b(t, X_t)dt + \int_0^t \sigma(t, X_t)dW_t \quad \forall t \in (0, T)$$

Théorème 10 (Théorème d'existence et d'unicité pour les EDS)

Soit $T > 0$ et $b(\cdot, \cdot) : [0, T] \times \mathbb{R} \rightarrow \mathbb{R}$ et $\sigma(\cdot, \cdot) : [0, T] \times \mathbb{R} \rightarrow \mathbb{R}_+$ deux fonctions mesurables satisfaisant :

$$|b(t, x)| + |\sigma(t, x)| \leq C(1 + |x|) \quad x \in \mathbb{R} \quad t \in [0, T] \quad (1.7)$$

$$|b(t, x) - b(t, y)| + |\sigma(t, x) - \sigma(t, y)| \leq |x - y| \quad x, y \in \mathbb{R} \quad t \in [0, T] \quad (1.8)$$

avec C et D des constantes. Soit Z une variable aléatoire indépendante de la σ -algèbre engendré par W_t , $t \geq 0$ et telle que $\mathbb{E}(|Z|^2) < \infty$. Alors l'EDS :

$$\begin{cases} dX_t = b(t, X_t)dt + \sigma(t, X_t)dW_t \\ X_0 = Z \quad 0 \leq t \leq T \end{cases}, \quad (1.9)$$

possède une unique solution X_t adaptée à la filtration engendrée par W_t et Z ; de plus :

$$\mathbb{E} \left(\int_0^T |X_t|^2 dt \right) < \infty$$

Démonstration. Voir (7)

□

1.4 Processus de diffusion

1.4.1 Définition

Définition 11

Un processus de diffusion est un processus stochastique solution d'une équation différentielle stochastique de la forme 1.6.

Définition 12 (Processus de Markov)

Un processus stochastique en temps continu $(X_t)_{t \in [0, T]}$ sur un espace probabilisé $(\Omega, \mathcal{A}, \mathbb{P})$, à valeurs dans \mathbb{R} est un **processus de Markov** si pour toute suite croissante d'instants $0 = t_0 < t_1 < \dots < t_n = s < t$, pour tous $x_0, x_1, \dots, x_n = x$, et pour tout borelien $A \in \mathcal{B}(\mathbb{R})$,

$$\mathbb{P}(X_t \in A \mid X_{t_0} = x_0, X_{t_1} = x_1, \dots, X_s = x) = \mathbb{P}(X_t \in A \mid X_s = x) \quad (1.10)$$

équivalent encore à :

$$\mathbb{E}(X_t \mid \mathcal{F}_s^X) = \mathbb{E}(X_t \mid \sigma(X_s)) \quad \text{pour tous instants } s < t. \quad (1.11)$$

avec $\mathcal{F}_s^X = \sigma(X_u; u \leq s)$

Remarque 1.4.1. Cette propriété est encore appelée propriété de Markov et elle traduit que la probabilité que le processus X_t prend des valeurs futures ne dépend que de la valeur présente du processus X_s ($s < t$) et non pas du passé.

Exemple 1.4.1. Le mouvement Brownien est un processus de Markov.

Théorème 13 (Propriété de Markov pour les processus d'Itô)

Soit $(X_t)_{t \in [0, T]}$ un processus solution d'une EDS homogène, ie, d'une EDS de la forme

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

Alors le processus $(X_t)_{t \in [0, T]}$ vérifie la propriété de Markov .

Démonstration. Voir (7)

□

1.4.2 Notion de générateur infinitésimal

Définition 14

Considérons une EDS telle que définie par l'équation 1.6 :

$$\begin{cases} dX_t = b(t, X_t)dt + \sigma(t, X_t)dW_t \\ X_0 = x_0 \in \mathbb{R}, \quad 0 \leq t \leq T \end{cases}$$

On appelle générateur infinitésimal de cette EDS, l'opérateur noté L et défini sur l'espace $C^{1,2}([0, T] \times \mathbb{R})$ par :

$$L_t f(t, x) = \frac{1}{2} \sigma^2(t, x) \frac{\partial^2 f}{\partial x^2}(t, x) + b(t, x) \frac{\partial f}{\partial x}(t, x), \quad (1.12)$$

pour $f \in C^{1,2}([0, T] \times \mathbb{R})$ et pour tout $(t, x) \in [0, T] \times \mathbb{R}$

Remarque 1.4.2. Cette notion est importante car elle est par exemple utile dans le Théorème de Feymann-Kac qui est un pont entre les équations différentielles stochastiques et les équations aux dérivées partielles déterministes.

1.4.3 Equations aux dérivées partielles de Fokker-Planck-Kolmogorov

Définition 15

Soit X_t un processus de diffusion solution d'une EDS de la forme 1.6. Soit $t, s \in [0, T]$ tels que $s < t$. On appelle densité de transition du processus X_t prenant la valeur $y \in \mathbb{R}$ à l'instant t sachant qu'il a une valeur $x \in \mathbb{R}$ à l'instant s le nombre noté : $p(t, y | s, x)$ et défini par :

$$p(t, y | s, x) = p(X_t = y | X_s = x). \quad (1.13)$$

Théorème 16

La densité de transition d'un processus de diffusion X_t solution de l'EDS 1.6 satisfait les équations aux dérivées partielles de Kolmogorov ; notamment :

— Kolmogorv forward equation

$$\frac{\partial}{\partial t} p(t, y | s, x) = -\frac{\partial}{\partial y} (b(t, y) p(t, y | s, x)) + \frac{1}{2} \frac{\partial^2}{\partial y^2} (\sigma^2(t, y) p(t, y | s, x)). \quad (1.14)$$

— Kolmogorv backward equation

$$-\frac{\partial}{\partial s} p(t, y | s, x) = b(t, x) \frac{\partial}{\partial x} p(t, y | s, x) + \frac{1}{2} \sigma^2(t, x) \frac{\partial^2}{\partial x^2} p(t, y | s, x). \quad (1.15)$$

Démonstration. Voir (7)

□

METHODES D'ESTIMATION DES PARAMÈTRES DANS LES EDS PAR MAXIMUM DE VRAISEMBLANCE

Dans tout ce chapitre, $(\Omega, \mathcal{A}, \mathbb{F}, \mathbb{P})$ est un espace probabilisé filtré et $(W_s)_{s \in [0, T]}$ un mouvement Brownien unidimensionnel adapté à la filtration \mathbb{F}

2.1 Le cadre

Nous supposons donc qu'on dispose des observations $(X_{t_i}^{obs})_{i=0}^n$ d'un processus $(X_t)_{t \in [0, T]}$ avec $T \geq 0$, observé aux instants $t_i = i\Delta$, avec $\Delta = \frac{T}{n}$, n étant fixé, qui évolue selon l'équation différentielle stochastique :

$$dX_t = b(X_t, \theta)dt + \sigma(X_t, \theta)dW_t$$

sur un espace probabilisé filtré où $\theta \in \Theta \subset \mathbb{R}^p$ qui est l'espace des paramètres (on le prendra ouvert et borné) et θ_0 est le vecteur des vrais paramètres à estimer. Les fonctions $\sigma : \mathbb{R} \times \Theta \rightarrow \mathbb{R}_+$ et $b : \mathbb{R} \times \Theta \rightarrow \mathbb{R}$ sont connues.

Notre objectif est donc de construire un estimateur $\hat{\theta}$ du vecteur des vrais paramètres θ_0 .

Il existe plusieurs méthodes d'estimation dans les EDS et donc plusieurs types d'estimateurs :

- L'estimateur du maximum de vraisemblance (**EMV**), ce qui suppose de calculer la vraisemblance de la famille d'observations.
- L'estimateur de minimum de contraste **EMC**, lorsque la vraisemblance est difficile à calculer, voir impossible à calculer ou difficile à exploiter; une alternative est de construire une pseudo-log-vraisemblance dont l'opposé sera dit **contraste** et dont on tirera des estimateurs.

- Les estimateurs empiriques **EMM**, fondés sur une méthode des moments, éventuellement méthode de moments généralisés.

Nous nous attarderons exclusivement sur la méthode d'estimation par maximum de vraisemblance.

Définition 1

- 1 On appelle **modèle statistique**, la donnée d'un ensemble d'observations \mathcal{H}^n et d'une famille q de lois probabilité définie sur \mathcal{H}^n . On note $MS = (\mathcal{H}^n, q)$.
- 2 Un modèle statistique $MS = (\mathcal{H}^n, q)$ est dit paramétré par Θ si $q = \{p_\theta\}_{\theta \in \Theta}$. On note $MS = (\mathcal{H}^n, \{p_\theta\}_{\theta \in \Theta})$. Si Θ est inclu dans un espace vectoriel de dimension finie, alors le modèle est dit **paramétrique**.

Remarque 2.1.1. Dans notre cas, $\mathcal{H} = \mathbb{R}$ et $\Theta \subset \mathbb{R}^p$; donc on a en face le modèle paramétrique $MS = (\mathbb{R}^n, \{p_\theta\}_{\theta \in \Theta})$

Définition 2

Dans un modèle paramétrique $(\mathcal{H}^n, \{p_\theta\}_{\theta \in \Theta})$, on appelle **vraisemblance** des observations $\underline{x}_n = (x_0, x_1, \dots, x_n)$, l'application :

$$\begin{aligned} L_n(\underline{x}_n; \cdot) : \Theta &\longrightarrow [0, 1] \\ \theta &\mapsto L_n(\underline{x}_n; \theta) = p_\theta(X_{t_0} = x_0, X_{t_1} = x_1, \dots, X_{t_n} = x_n). \end{aligned}$$

Définition 3 (EMV)

On appelle **estimateur du maximum de vraisemblance (EMV)**, des observations X_1, X_2, \dots, X_n tout élément $\hat{\theta}_n \equiv \hat{\theta}_n(X_1, X_2, \dots, X_n)$ appartenant à $\underset{\theta \in \Theta}{\operatorname{Argmax}} L_n(\theta)$

Remarque 2.1.2. Comme la fonction \ln est croissante, alors $\underset{\theta \in \Theta}{\operatorname{Argmax}} \ln(L_n)(\theta) = \underset{\theta \in \Theta}{\operatorname{Argmax}} L_n(\theta)$. Et donc, au lieu de chercher à maximiser la vraisemblance, on cherchera souvent à maximiser la log-vraisemblance qu'on notera ℓ_n .

Remarque 2.1.3. Sous les hypothèses d'un modèle statistique régulier, supposons l'estimateur du maximum de vraisemblance de θ existe et est unique; notons le $\hat{\theta}_n^{EMV}$. Alors $\hat{\theta}_n^{EMV}$ a de très bonne propriétés asymptotiques, notamment :

— *La consistance forte :*

$$\hat{\theta}_n^{EMV} \xrightarrow{ps} \theta_0 \text{ lorsque } n \rightarrow +\infty \quad (2.1)$$

— *La normalité asymptotique :*

$$\sqrt{n}(\hat{\theta}_n^{EMV} - \theta_0) \xrightarrow{\mathcal{L}} \mathcal{N}(0, I^{-1}(\theta_0)). \quad (2.2)$$

où $I(\theta_0)$ est la matrice information de Fisher au point θ_0 .

En considérant donc l'équation différentielle stochastique :

$$\begin{cases} dX_t = b(X_t, \theta)dt + \sigma(X_t, \theta)dW_t \\ X_{t_0} = x_0 \end{cases},$$

et en supposant que les fonctions b et σ sont suffisamment régulières et satisfont les conditions d'existence et d'unicité, notamment les conditions de Lipschitz et de croissance linéaire, on déduit du Théorème d'existence et d'unicité des solutions des EDS que cette équation admet une unique solution $(X_t)_{t \in [0, T]}$ qui en plus vérifie la propriété Markovienne. Grâce à cette propriété, la vraisemblance s'écrit :

$$\begin{aligned} L_n(\theta) &= L(\underline{x}_n, \theta) = p_\theta(X_{t_0} = x_0, X_{t_1} = x_1, \dots, X_{t_n} = x_n) \\ &= \prod_{i=1}^n P_\theta(X_{t_i} = x_i | X_{t_{i-1}} = x_{i-1}) \cdot p_\theta(X_0 = x_0). \end{aligned}$$

Et donc la log-vraisemblance s'écrit :

$$\ell_n(\theta) = \ln(L_n(\theta)) = \sum_{i=1}^n \ln p_\theta(X_{t_i} = x_i | X_{t_{i-1}} = x_{i-1}) + \ln p_\theta(X_0 = x_0).$$

Note 2.1.1. $p_\theta(X_{t_i} = x_i | X_{t_{i-1}} = x_{i-1})$ est la densité de probabilité de la variable aléatoire de transition $X_{t_i} | X_{t_{i-1}} = x_{i-1}$ au point x_i encore appelée **densité de transition**. On notera désormais cette densité de transition par $p_X(\Delta, x_i | x_{i-1}; \theta)$; c'est à dire :

$$p_\theta(X_{t_i} = x_i | X_{t_{i-1}} = x_{i-1}) \equiv p_X(\Delta, x_i | x_{i-1}; \theta) \quad (2.3)$$

Selon que la loi de transition soit connue ou approximée, on parlera de vraisemblance exacte ou de vraisemblance approximée.

L'estimateur $\hat{\theta}^{EMV}$ du vecteur des vrais paramètres θ_0 obtenu par la méthode du maximum

de vraisemblance sera alors, selon qu'on soit dans le cas exacte ou approximé, donné par :

$$\hat{\theta}^{EMV} = \underset{\theta \in \Theta}{\operatorname{argmax}} \left(\sum_{i=1}^n \ln p_X(\Delta, x_i | x_{i-1}; \theta) \right) \quad \text{cas exacte;}$$

$$\tilde{\theta}^{EMV} = \underset{\theta \in \Theta}{\operatorname{argmax}} \left(\sum_{i=1}^n \ln \hat{p}_X(\Delta, x_i | x_{i-1}; \theta) \right) \quad \text{cas approximé;}$$

où \hat{p}_X est la loi de transition approximée.

Ainsi pour trouver les estimateurs de maximum de vraisemblance, nous n'aurons qu'à trouver l'expression des densités de transition (de façon exacte ou approximée).

Par la suite, nous déterminerons donc $p_X(\Delta, x | x_0; \theta) = p_\theta(X_{t+\Delta} = x | X_t = x_0)$ pour un temps quelconque t et pour x, x_0 positifs, avant d'adapter à la formule pour trouver le maximum de vraisemblance comme énoncé.

2.2 Cas de vraisemblance exacte

Il s'agit du cas de vraisemblance où la loi de la transition est connue. Dans la suite, nous étudierons des modèles de base tels que : le modèle de Black-Scholes, le modèle de Ornstein-Uhlenbeck, le modèle de Cox-Ingersoll-Ross. Pour chaque modèle, nous déterminerons la solution exacte si possible de l'EDS associée, nous déterminerons la densité de transition et ensuite nous déduirons l'expression de la vraisemblance exacte; nous donnerons enfin les expressions des estimateurs du maximum de vraisemblance associées.

2.2.1 Modèle 1 : Le modèle de Black-Scholes ou Mouvement Brownien Géométrique

Définition 4

Le mouvement Brownien Géométrique est un processus stochastique souvent utilisé pour représenter l'évolution du prix d'un actif financier. Il est défini comme étant la solution de l'équation différentielle stochastique :

$$dX_t = \mu X_t dt + \sigma X_t dW_t \quad (2.4)$$

où $\theta = (\mu, \sigma)$ et $b(X_t, \theta) = \mu X_t$, $\sigma(X_t, \theta) = \sigma X_t$ sont respectivement le coefficient Drift et le coefficient de diffusion.

Interprétation

μ et σ sont respectivement appelés le drift et la volatilité du prix d'action et représentent respectivement la moyenne instantanée du taux d'évolution du prix et l'écart type instantané correspondant.

Proposition 5

- 1** La solution de l'EDS de Black-Scholes $\begin{cases} dX_t = \mu X_t dt + \sigma X_t dW_t \\ X_0 = x_0 \end{cases}$ est définie par :

$$X_t = x_0 \exp \left\{ \sigma W_t + \left(\mu - \frac{\sigma^2}{2} \right) t \right\}. \quad (2.5)$$

- 2** La distribution conditionnelle de $X_{t+\Delta} | X_t = x_0$ est une loi log-normale et la densité de transition $p_X(\Delta, x | x_0; \theta) = p_\theta(X_{t+\Delta} = x | X_t = x_0)$ est donnée par :

$$p_X(\Delta, x | x_0; \theta) = \frac{1}{\sqrt{2\pi\sigma^2 x^2 \Delta}} \exp \left(-\frac{\left(\ln \left(\frac{x}{x_0} \right) - \left(\mu - \frac{\sigma^2}{2} \right) \Delta \right)^2}{2\sigma^2 \Delta} \right). \quad (2.6)$$

Rappel 2.2.1. Si X et Y sont deux variables aléatoires réelles telles que $X = \exp(Y)$ et Y suit une loi normale de paramètre m et σ , c'est à dire $Y \sim \mathcal{N}(m, \sigma)$, alors X suit une loi log-normale et sa densité est donnée par :

$$f_X(x) = \begin{cases} \frac{1}{x} f_Y(\log x) & \text{si } x > 0 \\ 0 & \text{si } x \leq 0 \end{cases}$$

Démonstration. 1) On va appliquer la formule d'Itô au processus $f(X_t) = \ln(X_t)$. Par définition et par écriture, X_t est un processus d'Itô et la fonction $f : x \rightarrow \ln(x)$ est de classe C^2 ; donc d'après le Lemme d'Itô, $f(X_t)$ est un processus d'Itô et sa différentielle stochastique est :

$$\begin{aligned} df(X_t) &= d\ln(X_t) = \frac{1}{X_t} dX_t - \frac{1}{2X_t^2} d[X]_t \\ &= \frac{1}{X_t} (\mu X_t dt + \sigma X_t dW_t) - \frac{1}{2X_t^2} \sigma^2 X_t^2 dt. \\ &= \left(\mu - \frac{\sigma^2}{2} \right) dt + \sigma dW_t. \end{aligned} \quad (2.7)$$

En intégrant entre 0 et t l'égalité 2.7, on obtient :

$$\int_0^t df(X_s) = \int_0^t \left(\mu - \frac{\sigma^2}{2} \right) ds + \sigma \int_0^t dW_s \text{ ie } f(X_t) - f(X_0) = \left(\mu - \frac{\sigma^2}{2} \right) t + \sigma W_t$$

Ce qui implique que $\ln\left(\frac{X_t}{X_0}\right) = \left(\mu - \frac{\sigma^2}{2} \right) t + \sigma W_t$. D'où $X_t = x_0 \exp \left\{ \sigma W_t + \left(\mu - \frac{\sigma^2}{2} \right) t \right\}$.

2) On rappelle que $W_t \sim \mathcal{N}(0, t)$.

En intégrant à nouveau entre t et $t + \Delta$ l'égalité 2.7, on a :

$$\int_t^{t+\Delta} df(X_s) = \int_t^{t+\Delta} \left(\mu - \frac{\sigma^2}{2} \right) ds + \sigma \int_t^{t+\Delta} dW_s.$$

Donc $\ln\left(\frac{X_{t+\Delta}}{X_t}\right) = \left(\mu - \frac{\sigma^2}{2} \right) \Delta + \sigma(W_{t+\Delta} - W_t)$. On en déduit que $X_{t+\Delta} = X_t \exp \left\{ \sigma(W_{t+\Delta} - W_t) + \left(\mu - \frac{\sigma^2}{2} \right) \Delta \right\}$.

En supposant que $X_t = x_0$ et $x_0 > 0$, il vient que :

$$(X_{t+\Delta} | X_t = x_0) = \exp \left\{ \sigma(W_{t+\Delta} - W_t) + \left(\mu - \frac{\sigma^2}{2} \right) \Delta + \ln(x_0) \right\}.$$

Or $W_{t+\Delta} - W_t \sim \mathcal{N}(0, \Delta)$, il vient que

$$\sigma(W_{t+\Delta} - W_t) + \left(\mu - \frac{\sigma^2}{2} \right) \Delta + \ln(x_0) \sim \mathcal{N} \left(\left(\mu - \frac{\sigma^2}{2} \right) \Delta + \ln(x_0); \sigma^2 \Delta \right).$$

Il s'en suit du **Rappel** que la variable aléatoire $X_{t+\Delta} | X_t = x_0$ suit une loi Log-normale et que sa densité est donnée par :

$$p_X(\Delta, x | x_0; \theta) = \frac{1}{x} \frac{1}{\sqrt{2\pi\sigma^2\Delta}} \exp \left(- \frac{[\ln(x) - (\mu - \sigma^2/2)\Delta - \ln(x_0)]^2}{2\sigma^2\Delta} \right).$$

$$\text{D'où : } p_X(\Delta, x | x_0; \theta) = \frac{1}{\sqrt{2\pi\sigma^2 x^2 \Delta}} \exp \left(- \frac{\left(\ln\left(\frac{x}{x_0}\right) - \left(\mu - \frac{\sigma^2}{2} \right) \Delta \right)^2}{2\sigma^2 \Delta} \right)$$

□

Conséquence 6

Connaissant la densité des variables aléatoires de transition $X_{t_i} | X_{t_{i-1}}$ qu'on déduit de la proposition précédente, on déduit que l'expression de la Log-vraisemblance du modèle est donnée par :

$$\begin{aligned}\ell_n(\theta) &= \sum_{i=1}^n \ln p_\theta(X_{t_i} = x_i | X_{t_{i-1}} = x_{i-1}) \\ &= -\frac{n}{2} \ln(2\pi\sigma^2\Delta) - \frac{1}{2} \sum_{i=1}^n \ln x_i^2 - \frac{1}{2\sigma^2\Delta} \sum_{i=1}^n \left(\ln\left(\frac{x_i}{x_{i-1}}\right) - \left(\mu - \frac{\sigma^2}{2}\right)\Delta \right)^2.\end{aligned}$$

Théorème 7

Les estimateurs du maximum de vraisemblance μ^{EMV} et $(\sigma^2)^{EMV}$ sont donnés par :

$$(\hat{\sigma}^2)^{EMV} = \frac{1}{T} \sum_{i=1}^n \ln^2\left(\frac{X_{t_i}}{X_{t_{i-1}}}\right) - \frac{1}{nT} \left(\sum_{i=1}^n \ln\left(\frac{X_{t_i}}{X_{t_{i-1}}}\right) \right)^2 \quad (2.8)$$

$$\hat{\mu}^{EMV} = \frac{1}{T} \sum_{i=1}^n \ln\left(\frac{X_{t_i}}{X_{t_{i-1}}}\right) + \frac{1}{2} \left(\frac{1}{T} \sum_{i=1}^n \ln^2\left(\frac{X_{t_i}}{X_{t_{i-1}}}\right) - \frac{1}{nT} \left(\sum_{i=1}^n \ln\left(\frac{X_{t_i}}{X_{t_{i-1}}}\right) \right)^2 \right). \quad (2.9)$$

Démonstration. Pour trouver les estimateurs de μ et de σ^2 , nous allons d'abord trouver les estimateurs de σ^2 et de $\left(\mu - \frac{\sigma^2}{2}\right)\Delta$. Posons donc $m = \left(\mu - \frac{\sigma^2}{2}\right)\Delta$. Alors la Log-vraisemblance s'écrit encore :

$$\ell_n(\theta) = -\frac{n}{2} \ln(2\pi\sigma^2\Delta) - \frac{1}{2} \sum_{i=1}^n \ln(x_i^2) - \frac{1}{2\sigma^2\Delta} \sum_{i=1}^n \left[\ln\left(\frac{x_i}{x_{i-1}}\right) - m \right]^2.$$

En dérivant cette Log-vraisemblance par rapport à m et σ^2 , on obtient :

$$\begin{aligned}\frac{\partial \ell_n(\theta)}{\partial m} &= -\frac{1}{2\sigma^2\Delta} \sum_{i=1}^n (-2) \left[\ln\left(\frac{x_i}{x_{i-1}}\right) - m \right] = \frac{1}{\sigma^2\Delta} \left[\sum_{i=1}^n \ln\left(\frac{x_i}{x_{i-1}}\right) - nm \right] \\ \frac{\partial \ell_n(\theta)}{\partial \sigma^2} &= -\frac{n}{2\sigma^2} + \frac{1}{2\Delta\sigma^4} \sum_{i=1}^n \left[\ln\left(\frac{x_i}{x_{i-1}}\right) - m \right]^2\end{aligned}$$

Les valeurs \hat{m} et $\hat{\sigma}^2$ vérifient le système d'équations : $\frac{\partial \ell_n(\theta)}{\partial m} = \frac{\partial \ell_n(\theta)}{\partial \sigma^2} = 0$. Ce qui donne après

résolution : $\hat{m} = \frac{1}{n} \sum_{i=1}^n \ln \left(\frac{x_i}{x_{i-1}} \right)$; et :

$$\hat{\sigma}^2 = \frac{1}{T} \sum_{i=1}^n \left[\ln \left(\frac{x_i}{x_{i-1}} \right) - \hat{m} \right]^2 = \frac{1}{n\Delta} \sum_{i=1}^n \ln^2 \left(\frac{x_i}{x_{i-1}} \right) - \frac{n}{T} (\hat{m})^2 = \frac{1}{T} \sum_{i=1}^n \ln^2 \left(\frac{X_{t_i}}{X_{t_{i-1}}} \right) - \frac{1}{nT} \left(\sum_{i=1}^n \ln \left(\frac{X_{t_i}}{X_{t_{i-1}}} \right) \right)^2.$$

De l'expression $\hat{m} = \left(\mu - \frac{\sigma^2}{2} \right)^{EMV}$, on déduit que $\hat{\mu} = \hat{m} + \frac{1}{2} \hat{\sigma}^2$. Ou mieux encore :

$$\hat{\mu} = \frac{1}{n} \sum_{i=1}^n \ln \left(\frac{x_i}{x_{i-1}} \right) + \frac{1}{2} \left[\frac{1}{T} \sum_{i=1}^n \ln^2 \left(\frac{x_i}{x_{i-1}} \right) - \frac{1}{nT} \left(\sum_{i=1}^n \ln \left(\frac{x_i}{x_{i-1}} \right) \right)^2 \right]$$

□

2.2.2 Modèle 2 : Modèle Ornstein-Uhlenbeck ou Vasicek process

Définition 8

Le modèle de Vasicek est un processus stochastique utilisé pour modéliser le taux d'intérêt. Il est défini comme étant la solution de l'EDS suivante :

$$dX_t = \beta(\bar{\alpha} - X_t)dt + \sigma dW_t \quad (2.10)$$

où β , $\bar{\alpha}$, σ sont trois réels strictement positifs.

Interprétation

- $\bar{\alpha}$ représente la moyenne à long terme.
- β est la vitesse de retour à la moyenne $\bar{\alpha}$.
- σ est la volatilité.

Proposition 9

1 La solution de l'EDS de **Vasicek** $\begin{cases} dX_t = \beta(\bar{\alpha} - X_t)dt + \sigma dW_t \\ X_0 = x_0 \end{cases}$ est donnée par :

$$X_t = \bar{\alpha} + e^{-\beta t} \left(x_0 - \bar{\alpha} + \sigma \int_0^t e^{\beta u} dW_u \right) \quad (2.11)$$

2 Le processus X_t ci dessus est un processus Gaussien de moyenne $\bar{\alpha} + e^{-\beta t} (x_0 - \bar{\alpha})$ et de variance $\frac{\sigma^2}{2\beta} (1 - e^{-2\beta t})$ et la densité de transition est donnée par :

$$p_X(\Delta, x | x_0; \theta) = \left(\frac{\pi \sigma^2 (1 - e^{-2\beta \Delta})}{\beta} \right)^{-\frac{1}{2}} \exp \left\{ -\beta \frac{(x - \bar{\alpha} - (x_0 - \bar{\alpha})e^{-\beta \Delta})^2}{\sigma^2 (1 - e^{-2\beta \Delta})} \right\} \quad (2.12)$$

Rappel 2.2.2. W_t est le mouvement Brownien standard. Soit f une fonction appartenant à $L^2_{loc}(\mathbb{R}_+)$, c'est à dire $\int_0^T |f(x)|^2 ds < \infty, \forall T$. L'intégrale de Wiener définie par : $\int_0^t f(s) dW_s$ est un processus Gaussien centré de variance $\int_0^t f^2(s) ds$

Démonstration. 1) Posons $Y_t = X_t - \bar{\alpha}$. Alors, en utilisant le Lemme d'Itô, le processus Y_t est également un processus d'Itô et sa différentielle stochastique est : $dY_t = X_t$; c'est à dire : $dY_t = -\beta Y_t dt + \sigma dW_t$. On va appliquer la formule d'intégration par parties au processus $g(Y_t) = e^{\beta t} Y_t$. On a donc :

$$\begin{aligned} dg(Y_t) &= d(e^{\beta t} Y_t) = \beta e^{\beta t} Y_t dt + e^{\beta t} dY_t + \underbrace{d[e^{\beta \cdot}; Y_t]}_{=0} \text{ car } t \mapsto e^{\beta t} \text{ est à variation finie} \\ &= \beta e^{\beta t} Y_t dt + e^{\beta t} (-\beta Y_t dt + \sigma dW_t) = \sigma e^{\beta t} dW_t \end{aligned}$$

En intégrant la dernière égalité entre 0 et t , on obtient : $\beta e^{\beta t} Y_t - Y_0 = \sigma \int_0^t e^{\beta u} dW_u$. Donc $Y_t = e^{-\beta t} (Y_0 + \sigma \int_0^t e^{\beta u} dW_u)$.

D'où $X_t = \bar{\alpha} + e^{-\beta t} (x_0 - \bar{\alpha} + \sigma \int_0^t e^{\beta u} dW_u)$.

2) Grace au **Rappel**, comme la fonction $t \mapsto e^{\beta u}$ appartient à $L^2_{loc}(\mathbb{R}_+)$, on déduit que $\int_0^t e^{\beta u} dW_u$ est un processus Gaussien ; et donc que $X_t = \bar{\alpha} + e^{-\beta t} (x_0 - \bar{\alpha}) + \sigma e^{-\beta t} \left(\int_0^t e^{\beta u} dW_u \right)$ est un processus Gaussien de moyenne et de variance (en utilisant la propriété d'isométrie d'Itô) :

$$\mathbb{E}(X_t) = \bar{\alpha} + e^{-\beta t} (x_0 - \bar{\alpha}) \text{ et } \mathbb{V}ar(X_t) = (\sigma e^{-\beta t})^2 \int_0^t e^{2\beta u} du = \frac{\sigma^2}{2\beta} (1 - e^{-2\beta t}).$$

Pour obtenir la loi de la variable aléatoire de transition $X_{t+\Delta} | X_t$, on intègre l'égalité $de^{\beta t} Y_t =$

$\sigma e^{\beta t} dW_t$ entre t et $t + \Delta$; et on obtient :

$$X_{t+\Delta} = \bar{\alpha} + e^{-\beta\Delta} (X_t - \bar{\alpha}) + \sigma e^{-\beta(t+\Delta)} \left(\int_t^{t+\Delta} e^{\beta u} dW_u \right).$$

En supposant que $X_t = x_0$, on déduit que la variable de transition $X_{t+\Delta} | X_t = x_0$ suit une loi Gaussienne de moyenne : $\mathbb{E}(X_{t+\Delta} | X_t = x_0) = \bar{\alpha} + e^{-\beta\Delta} (x_0 - \bar{\alpha})$; et de variance :

$$\begin{aligned} \mathbb{V}ar(X_{t+\Delta} | X_t = x_0) &= \left(\sigma e^{-\beta(t+\Delta)} \right)^2 \int_t^{t+\Delta} e^{2\beta u} du \\ &= \left(\sigma e^{-\beta(t+\Delta)} \right)^2 \times \frac{1}{2\beta} \left(e^{2\beta(t+\Delta)} - e^{2\beta t} \right) = \frac{\sigma^2}{2\beta} \left(1 - e^{-2\beta\Delta} \right). \end{aligned}$$

Il s'en suit que la densité de transition est donnée par :

$$\begin{aligned} p_X(\Delta |, x_0; \theta) &= \frac{1}{\sqrt{2\pi \frac{\sigma^2}{2\beta} (1 - e^{-2\beta\Delta})}} \times \exp \left\{ -\frac{1}{2} \frac{(x - \bar{\alpha} - e^{-\beta\Delta} (x_0 - \bar{\alpha}))^2}{\frac{\sigma^2}{2\beta} (1 - e^{-2\beta\Delta})} \right\} \\ &= \left(\frac{\pi \sigma^2 (1 - e^{-2\beta\Delta})}{\beta} \right)^{-\frac{1}{2}} \times \exp \left\{ -\beta \frac{(x - \bar{\alpha} - (x_0 - \bar{\alpha})e^{-\beta\Delta})^2}{\sigma^2 (1 - e^{-2\beta\Delta})} \right\} \end{aligned}$$

□

Conséquence 10

A partir de la densité de transition donnée dans la proposition précédente, on déduit que la log vraisemblance du modèle est exprimée par :

$$\ell_n(\theta) = -\frac{n}{2} \ln(2\pi) - \frac{n}{2} \ln \left(\frac{\sigma^2 (1 - e^{-2\beta\Delta})}{2\beta} \right) - \left(\frac{\beta}{\sigma^2 (1 - e^{-2\beta\Delta})} \right) \sum_{i=1}^n (x_i - \bar{\alpha} - (x_{i-1} - \bar{\alpha})e^{-\beta\Delta})^2$$

Théorème 11

Les estimateurs du maximum de vraisemblance $(\bar{\alpha})^{EMV}$, $(\beta)^{EMV}$, $(\sigma)^{EMV}$ sont liés par les relations suivantes :

$$\hat{\alpha} = \frac{1}{n(1 - e^{-\hat{\beta}\Delta})} \sum_{i=1}^n (X_{t_i} - X_{t_{i-1}} e^{-\hat{\beta}\Delta}) \quad (2.13)$$

$$\hat{\beta} = -\frac{1}{\Delta} \ln \left(\frac{\sum_{i=1}^n (X_{t_{i-1}} - \hat{\alpha})(X_{t_i} - \hat{\alpha})}{\sum_{i=1}^n (X_{t_{i-1}} - \hat{\alpha})^2} \right) \quad (2.14)$$

$$\hat{\sigma}^2 = \frac{2\hat{\beta}}{n(1 - e^{-2\hat{\beta}\Delta})} \sum_{i=1}^n (X_{t_i} - X_{t_{i-1}} e^{-\hat{\beta}\Delta} - (1 - e^{-\hat{\beta}\Delta})\hat{\alpha})^2 \quad (2.15)$$

Démonstration. Posons $F = \frac{\sigma^2}{2\beta} (1 - e^{-2\beta\Delta})$ et déterminons donc les estimateurs du maximum de vraisemblance de F , $\bar{\alpha}$ et β .

La log-vraisemblance s'écrit : $\ell_n(\theta) = -\frac{n}{2} \ln(2\pi) - \frac{1}{2} \ln(F) - \frac{1}{2F} \sum_{i=1}^n (x_i - \bar{\alpha} - (x_{i-1} - \bar{\alpha})e^{-\beta\Delta})^2$.

En dérivant cette log-vraisemblance respectivement par rapport aux variables F , $\bar{\alpha}$ et β , on obtient :

$$\begin{aligned} \frac{\partial \ell_n(\theta)}{\partial \bar{\alpha}} &= \frac{1 - e^{-\beta\Delta}}{2F} \left(\sum_{i=1}^n (x_i - x_{i-1} e^{-\beta\Delta}) - n(1 - e^{-\beta\Delta})\bar{\alpha} \right); \\ \frac{\partial \ell_n(\theta)}{\partial \beta} &= -\frac{\Delta e^{-\beta\Delta}}{F} \left(\sum_{i=1}^n (x_{i-1} - \bar{\alpha})^2 - e^{-\beta\Delta} \sum_{i=1}^n (x_{i-1} - \bar{\alpha})(x_i - \bar{\alpha}) \right); \\ \frac{\partial \ell_n(\theta)}{\partial F} &= -\frac{n}{2F} + \frac{1}{2F^2} \sum_{i=1}^n (x_i - \bar{\alpha} - (x_{i-1} - \bar{\alpha})e^{-\beta\Delta})^2. \end{aligned}$$

En résolvant le système d'équations $\frac{\partial \ell_n(\theta)}{\partial \bar{\alpha}} = \frac{\partial \ell_n(\theta)}{\partial \beta} = \frac{\partial \ell_n(\theta)}{\partial F} = 0$, on obtient :

$$\begin{aligned} \hat{\alpha} &= \frac{1}{n(1 - e^{-\hat{\beta}\Delta})} \sum_{i=1}^n (x_i - x_{i-1} e^{-\hat{\beta}\Delta}) \\ \hat{\beta} &= -\frac{1}{\Delta} \ln \left(\frac{\sum_{i=1}^n (x_{i-1} - \hat{\alpha})(x_i - \hat{\alpha})}{\sum_{i=1}^n (x_{i-1} - \hat{\alpha})^2} \right) \end{aligned}$$

$$\hat{\sigma}^2 = \frac{2\hat{\beta}}{n(1 - e^{-2\hat{\beta}\Delta})} \times \sum_{i=1}^n \left(x_i - x_{i-1}e^{-\hat{\beta}\Delta} - (1 - e^{-\hat{\beta}\Delta})\hat{\alpha} \right)^2.$$

Et on a le résultat voulu en remplaçant les estimations par les estimateurs.

□

Remarque 2.2.1. *L'estimateur $\hat{\beta}$ est un estimateur consistant et asymptotiquement Gaussien et n'est défini que lorsque*

$$(X_{t_{i-1}} - \hat{\alpha})(X_{t_i} - \hat{\alpha}) > 0$$

2.2.3 Modèle 3 : Le Modèle de Cox-Ingersoll-Ross

Définition 12

Le processus de Cox-Ingersoll-Ross modélise également l'évolution des taux d'intérêt X_t et est défini comme étant une solution de l'équation différentielle stochastique ci-contre :

$$dX_t = \beta(\bar{\alpha} - X_t)dt + \sigma \sqrt{X_t}dW_t \quad (2.16)$$

où $\bar{\alpha}; \beta; \theta$ sont trois réels positifs.

Dans ce modèle, il n'y a pas de solution explicite de cette EDS. Mais on peut toujours trouver l'expression explicite de la densité de transition et par conséquent trouver une expression explicite de la log-vraisemblance pour trouver l'estimateur du maximum de vraisemblance exacte du vecteur des paramètres $\theta = (\bar{\alpha}; \beta; \theta)$.

En écrivant l'EDS sous forme intégrale entre t et $t + \Delta$, on a :

$$\int_t^{t+\Delta} dX_u = \beta \int_t^{t+\Delta} (\bar{\alpha} - X_u)du + \sigma \int_t^{t+\Delta} \sqrt{X_u}dw_u$$

Donc : $X_{t+\Delta} = X_t + \beta \int_t^{t+\Delta} (\bar{\alpha} - X_u)du + \sigma \int_t^{t+\Delta} \sqrt{X_u}dw_u$.

On en déduit que si $X_t = x_0$, alors : $X_{t+\Delta} = x_0 + \beta \int_t^{t+\Delta} (\bar{\alpha} - X_u)du + \sigma \int_t^{t+\Delta} \sqrt{X_u}dw_u$.

Ainsi :

Proposition 13

On a les différents résultats ci-après :

$$\mathbb{E}(X_{t+\Delta} | X_t = x_0) = x_0 e^{-\beta\Delta} + \bar{\alpha}(1 - e^{-\beta\Delta}) \quad (2.17)$$

$$\begin{aligned} \mathbb{V}ar(X_{t+\Delta} | X_t = x_0) &= (2\beta\bar{\alpha} + \sigma^2)(x_0 - \bar{\alpha})\Delta e^{-\beta\Delta} + \frac{\bar{\alpha}}{\beta}(2\beta\bar{\alpha} + \sigma^2) \\ &\quad + \left(x_0^2 - \frac{\bar{\alpha}}{\beta}(2\beta\bar{\alpha} + \sigma^2)\right)e^{-\beta\Delta} - \left((x_0 - \bar{\alpha})e^{-\beta\Delta} + \bar{\alpha}\right)^2 \end{aligned} \quad (2.18)$$

Démonstration. Commençons par déterminer $\mathbb{E}(X_{t+\Delta} | X_t = x_0)$.

De l'égalité $X_{t+\Delta} = X_t + \beta \int_t^{t+\Delta} (\bar{\alpha} - X_u) du + \sigma \int_t^{t+\Delta} \sqrt{X_u} dw_u$, en passant à l'espérance conditionnelle sachant l'évènement $\{X_t = x_0\}$, on déduit que :

$$\mathbb{E}(X_{t+\Delta} | X_t = x_0) = x_0 + \beta\bar{\alpha}\Delta - \beta \int_t^{t+\Delta} \mathbb{E}(X_u | X_t = x_0) du + \underbrace{\sigma \mathbb{E} \left[\left(\int_t^{t+\Delta} \sqrt{X_u} dw_u \right) | X_t = x_0 \right]}_{=0}$$

car l'intégrale d'Itô est d'espérance nulle.

$$\text{Donc : } \mathbb{E}(X_{t+\Delta} | X_t = x_0) = x_0 + \beta\bar{\alpha}\Delta - \beta \int_t^{t+\Delta} \mathbb{E}(X_u | X_t = x_0) du$$

$$\text{Autrement : } \mathbb{E}(X_{t+\Delta} | X_t = x_0) = x_0 + \beta\bar{\alpha}\Delta - \beta \int_0^\Delta \mathbb{E}(X_{t+\delta} | X_t = x_0) d\delta$$

Donc $\Delta \mapsto \mathbb{E}(X_{t+\Delta} | X_t = x_0)$ est solution de l'EDO de la forme $f' + \beta f = \beta\bar{\alpha}$ avec $f(0) = x_0$; dont la solution générale est donnée par : $f(t) = ke^{-\beta t} + \bar{\alpha}$ avec k réel. Or la condition initiale $f(0) = x_0$ implique que $k = x_0 - \bar{\alpha}$. D'où :

$$\mathbb{E}(X_{t+\Delta} | X_t = x_0) = (x_0 - \bar{\alpha})e^{-\beta\Delta} + \bar{\alpha} = x_0 e^{-\beta\Delta} + \bar{\alpha}(1 - e^{-\beta\Delta})$$

Pour ce qui est de la variance, déterminons d'abord l'EDS vérifiée par le processus X_t^2 . En utilisant le Lemme d'Itô au processus d'Itô X_t et à la fonction $x \mapsto x^2$ qui est C^2 , on peut écrire que :

$$\begin{aligned} dX_t^2 &= 2X_t dX_t + d[X]_t = 2X_t \left(\beta(\bar{\alpha} - X_t)dt + \sigma \sqrt{X_t} dw_t \right) + \sigma^2 X_t dt \\ &= \left(2X_t \beta(\bar{\alpha} - X_t) + \sigma^2 X_t \right) dt + 2\sigma X_t^{\frac{3}{2}} dw_t. \end{aligned}$$

En écrivant l'EDS précédente sous forme intégrale entre t et $t + \Delta$, on obtient :

$$\begin{aligned} X_{t+\Delta}^2 &= X_t^2 + \int_t^{t+\Delta} (2X_u\beta(\bar{\alpha} - X_u) + \sigma^2 X_u) du + 2\sigma \int_t^{t+\Delta} X_u^{\frac{3}{2}} dW_u \\ &= X_t^2 + \int_0^\Delta (2X_{t+\delta}\beta(\bar{\alpha} - X_{t+\delta}) + \sigma^2 X_{t+\delta}) d\delta + 2\sigma \int_0^\Delta X_{t+\delta}^{\frac{3}{2}} dW_{t+\delta}. \end{aligned}$$

En passant à l'espérance conditionnelle sachant l'évènement $\{X_t = x_0\}$, on déduit que :

$$\mathbb{E}(X_{t+\Delta}^2 | X_t = x_0) = x_0^2 + (2\beta\bar{\alpha} + \sigma^2) \int_0^\Delta \mathbb{E}(X_{t+\delta} | X_t = x_0) d\delta - 2\beta \int_0^\Delta \mathbb{E}(X_{t+\delta}^2 | X_t = x_0) d\delta + 0$$

$$\text{Or : } \int_0^\Delta \mathbb{E}(X_{t+\delta} | X_t = x_0) d\delta = \int_0^\Delta [(x_0 - \bar{\alpha})e^{-\beta\delta} + \bar{\alpha}] d\delta = \frac{x_0 - \bar{\alpha}}{\beta} (1 - e^{-\beta\Delta}) + \bar{\alpha}\Delta.$$

Il s'en suit que la fonction $\Delta \mapsto \mathbb{E}(X_{t+\Delta}^2 | X_t = x_0)$ est solution de l'EDO de la forme $f' = (2\beta\bar{\alpha} + \sigma^2)((x_0 - \bar{\alpha})e^{-\beta t} + \bar{\alpha}) - 2\beta f$ et $f(0) = x_0^2$ dont la solution générale est donnée par :

$$f(t) = (2\beta\bar{\alpha} + \sigma^2)(x_0 - \bar{\alpha})te^{-\beta t} + \frac{\bar{\alpha}}{\beta}(2\beta\bar{\alpha} + \sigma^2) + ke^{-\beta t}$$

où k est un réel déterminé par la condition initiale $f(0) = x_0^2$; donc $k = x_0^2 - \frac{\bar{\alpha}}{\beta}(2\beta\bar{\alpha} + \sigma^2)$.

$$\text{D'où : } \mathbb{E}(X_{t+\Delta}^2 | X_t = x_0) = (2\beta\bar{\alpha} + \sigma^2)(x_0 - \bar{\alpha})te^{-\beta t} + \frac{\bar{\alpha}}{\beta}(2\beta\bar{\alpha} + \sigma^2) + \left(x_0^2 - \frac{\bar{\alpha}}{\beta}(2\beta\bar{\alpha} + \sigma^2)\right)e^{-\beta t}$$

On déduit que :

$$\begin{aligned} \text{Var}(X_{t+\Delta} | X_t = x_0) &= \mathbb{E}(X_{t+\Delta}^2 | X_t = x_0) - (\mathbb{E}(X_{t+\Delta} | X_t = x_0))^2 \\ &= (2\beta\bar{\alpha} + \sigma^2)(x_0 - \bar{\alpha})\Delta e^{-\beta\Delta} + \frac{\bar{\alpha}}{\beta}(2\beta\bar{\alpha} + \sigma^2) + \left(x_0^2 - \frac{\bar{\alpha}}{\beta}(2\beta\bar{\alpha} + \sigma^2)\right)e^{-\beta\Delta} \\ &\quad - \left((x_0 - \bar{\alpha})e^{-\beta\Delta} + \bar{\alpha}\right)^2. \end{aligned}$$

□

Théorème 14

Sous l'hypothèse $2\beta\bar{\alpha} > \sigma^2$, le processus X_t est stationnaire. Plus précisément, la densité de transition existe sous une forme explicite. Le processus $Y_t = 2cX_t$, où $c = \frac{2\beta}{\sigma^2(1 - e^{-\beta\Delta})}$, a une distribution conditionnelle $Y_{t+\Delta} | Y_t = y_0$ qui suit une distribution de Khi-deux non-centrée de degré de liberté $\nu = \frac{4\beta\bar{\alpha}}{\sigma^2}$ et le paramètre de non centralisation est $y_0 e^{-\beta\Delta}$.

Démonstration. Voir (6)

□

La densité de transition du processus X_t peut être déduite à partir de la distribution conditionnelle $Y_{t+\Delta} | Y_t = y_0$; et on a le résultat suivant :

Proposition 15

$$p_X(\Delta, x | x_0; \theta) = c e^{-(u+v)} \left(\frac{u}{v} \right)^{\frac{q}{2}} \mathbb{I}_q(2\sqrt{uv})$$

où $u = cx_0 e^{-\beta\Delta}$, $v = cx$, $q = \frac{2\beta\bar{\alpha}}{\sigma^2} - 1$, et $\mathbb{I}_q(\cdot)$ est la fonction de BESSEL modifiée d'ordre q . Plus précisément :

$$\mathbb{I}_q = \sum_{k=0}^{\infty} \left(\frac{x}{2} \right)^{2k+q} \frac{1}{k! \Gamma(k+q+1)}; \quad \Gamma(z) = \int_0^{\infty} x^{z-1} e^{-x} dx, \quad z \in \mathbb{R}_+$$

Démonstration. Voir (6)

□

A partir de l'expression de la densité de transition obtenue grâce à la proposition précédente, on peut déduire l'expression de la log-vraisemblance et aussi les estimateurs du maximum de vraisemblance. Mais on peut d'abord décliner les observations suivantes sur le comportement de la distribution conditionnelle.

observation 2.2.1. Dans l'optique d'inferer la vraisemblance, bien que la distribution de Khi-deux non centrée soit une distribution de probabilité spéciale et bien connue, il n'est pas facile de la compiler. En effet, lorsque $\Delta \rightarrow 0$ alors $u \rightarrow \infty, v \rightarrow \infty$ (donc u et v explosent). Donc le problème de l'évaluation de la distribution de Khi-deux non-centrale réside dans l'évaluation de la fonction de Bessel en des arguments assez grands.

2.2.4 Modèle 4 : Le modèle CEV

Définition 16

Le modèle CEV (The Constant Elasticity of Variance) est le processus stochastique solution de l'EDS suivante :

$$dX_t = \beta(\bar{\alpha} - X_t) dt + \sigma X_t^p dW_t. \quad (2.19)$$

Donc le vecteur des paramètres de ce modèle est : $\theta = (\bar{\alpha}, \beta, \sigma, p)$ avec $p > 0, \beta > 0$ et $p > \frac{1}{2}$

Proposition 17

Pour ces paramètres, le modèle CEV n'admet pas de solution explicite et de densité de transition explicite bien que le théorème d'existence garantit l'existence de la solution.

Démonstration. Voir (6) □

2.3 Cas de vraisemblance approximée

Au regard de la section précédente, nous remarquons qu'il est d'autant plus difficile d'obtenir les estimateurs du maximum de vraisemblance si on n'a pas la solution exacte de l'EDS ou si on n'arrive pas à obtenir la distribution conditionnelle de la transition par d'autres moyens à l'exemple du modèle CIR précédent.

Dans la littérature de l'approximation de la densité de transition des processus des EDS, il y a plusieurs méthodes à l'instar de :

- **La méthode d'Euler** : qui utilise le schéma d'Euler classique pour les équations différentielles ordinaires déterministes.
- **La méthode de Eulerian** : qui utilise le schéma de Milstein.
- **La méthode de linéarisation locale**
- **La méthode de Kessler.**

Pour plus d'amples explications de ces méthodes d'approximation, se référer à (5).

Dans la suite nous présentons la méthode de Euler en ayant les mêmes considérations énoncées dans le cadre.

Si les coefficients drift et de diffusion de l'EDS sont constants sur chaque petit intervalle $[t, t + \Delta[$, alors le schéma d'Euler produit la discrétisation suivante :

$$X_{t+\Delta} - X_t = b(X_t, \theta) \Delta + \sigma(X_t, \theta) (W_{t+\Delta} - W_t).$$

Et on a le résultat suivant :

Proposition 18

les incréments $X_{t+\Delta} - X_t$ sont des variables aléatoires Gaussiennes indépendantes et :

$$X_{t+\Delta} - X_t \sim \mathcal{N}\left(b(X_t, \theta) \Delta; \sigma^2(X_t, \theta) \Delta\right)$$

De plus la densité de transition approximée s'écrit :

$$p_X^{EULER}(\Delta, x | x_0, \theta) = \frac{1}{2\pi\Delta\sigma^2(x_0, \theta)} \exp\left\{-\frac{1}{2} \frac{(x - x_0 - b(X_t, \theta) \Delta)^2}{\sigma^2(X_t, \theta) \Delta}\right\}$$

Démonstration. Elle découle du fait que $(W_{t+\Delta} - W_t) \sim \mathcal{N}(0; \Delta)$

□

Remarque 2.3.1. Cette approximation est bonne si Δ est suffisamment petit.

APPROXIMATION DE LA DENSITE DE TRANSITION

Nous présentons ici deux méthodes pour illustrer l'approximation la vraie densité de transition qui en générale est inconnue excepté quelques rares cas à l'instar de : Wong(1964), le modèle de Blac-Scholes(1973), Vasicek process(1978), et Cox-Ingersol-Ross(1985). La première est due à **Yacine Aït-Sahalia**(1998) qui a utilisé l'expansion polynomiale d'Her-mite; et la seconde méthode est la notre où nous utiliserons une approximation spline cu-bique.

Nous disposons des observations $(X_{t_i}^{obs})_{i=0}^N$ avec $t_i = t_0 + i\Delta$ où $\Delta = \frac{T}{N}$ N étant fixé; d'un processus $(X_t)_{t \in [0;T]}$ dont la dynamique est représentée par l'équation différentielle stochas-tique :

$$dX_t = b(X_t, \theta)dt + \sigma(X_t, \theta)dW_t \quad (3.1)$$

où $(W_t)_t$ est le mouvement brownien standard unidimensionnel; $b(\cdot, \cdot)$ et $\sigma(\cdot, \cdot)$ sont des fonc-tions bien connues; θ est le vecteur des vrais paramètres appartenant à l'ouvert borné $\Theta \subset \mathbb{R}^p$. Le processus est observé aux instants $t_i = i\Delta$ avec $i = 0, \dots, n$ où Δ est fixé.

L'objectif est de proposer un estimation de la densité de transition $p_X(\Delta, x | x_0; \theta)$ basée sur l'approximation de Sahalia où on essaye d'améliorer l'estimation en...

3.1 Approximation de la densité de transition selon Aït-Sahalia

Nous présentons ici la méthode décrite par Aït-Sahalia tout en précisant les hypothèses pour l'approximation et en ressortant les différentes étapes de façon succincte.

Définition 1

Le domaine D_X du processus $X : [0, T] \times \Omega \rightarrow \mathbb{R}$ est l' ensemble image de $[0,T] \times \Omega$ par X .

La méthode est basée sur les hypothèses suivantes :

Hypothèse 3.1.1 (Régularité des coefficients drift et de diffusion). *Les fonctions drift et de diffusion, $b(\cdot, \cdot) : \mathbb{R} \times \Theta \rightarrow \mathbb{R}$ et $\sigma(\cdot, \cdot) : \mathbb{R} \times \Theta \rightarrow \mathbb{R}_+$ respectivement, sont infiniment différentiables par rapport à la variable x sur D_X et deux fois continûment différentiables par rapport à la variable θ sur l'espace ouvert borné du vecteur de paramètres $\Theta \subset \mathbb{R}^p$.*

Hypothèse 3.1.2. *Cette hypothèse a deux volets :*

1 Si $D_X =]-\infty, +\infty[$, il existe une constante $c > 0$ telle que :

$$\sigma(x, \theta) > c > 0 \text{ pour tout } x \in D_X \text{ et } \theta \in \Theta.$$

2 Si $D_X =]0, +\infty[$, nous admettons la dégénérescence locale de la fonction $\sigma(\cdot, \cdot)$; c'est à dire :

— si $\sigma(0, \theta) = 0$ alors il existe des constantes ε_0, ω positives et ρ tels que

$$\sigma(x, \theta) \geq \omega x^\rho \text{ pour tout } 0 < x < \varepsilon_0 \text{ et } \theta \in \Theta.$$

— En dehors de 0, $\sigma(\cdot, \cdot)$ est non-dégénérée; c'est à dire :

Pour tout $\varepsilon > 0$, il existe une constante c_ε telle que :

$$\sigma(x, \theta) \geq c_\varepsilon > 0 \text{ pour tout } x \in [\varepsilon, +\infty[\text{ et } \theta \in \Theta.$$

3.1.1 Première étape : Transformation du processus X_t

Elle consiste à transformer le processus X_t en un processus Y_t dont le coefficient de diffusion est constant et égale à 1. Cette transformation s'appelle **la transformation de Lamperti**. Elle consiste à définir le processus $(Y_t)_{t \in [0, T]}$ par :

$$Y_t = \gamma(X_t, \theta) = \int_0^{X_t} \frac{du}{\sigma(u, \theta)}$$

où n'importe quelle primitive de la fonction $\frac{1}{\sigma}$ peut être choisie. Puisque le coefficient de Drift est strictement positif $\sigma > 0$ sur D_X , alors la fonction γ est bien définie, croissante et inversible.

En supposant que $D_X =]\underline{x}, \bar{x}[$, alors le domaine D_Y du processus Y est $D_Y =]\underline{y}, \bar{y}[$ où $\underline{y} = \lim_{x \rightarrow \underline{x}} \gamma(x, \theta)$ et $\bar{y} = \lim_{x \rightarrow \bar{x}} \gamma(x, \theta)$ (car γ est croissante). Afin que les notations soient simples,

l'espace du vecteur des paramètres Θ est choisi de sorte que le domaine D_Y du processus Y soit indépendant de $\theta \in \Theta$. Cette supposition n'est pas trop essentiel.

Proposition 2

L'équation différentielle stochastique vérifiée par le processus $(Y_t)_{t \in [0, T]}$ est :

$$dY_t = \mu_Y(Y_t, \theta)dt + dW_t \quad (3.2)$$

$$\text{où : } \mu_Y(Y_t, \theta) = \frac{b(\gamma^{-1}(Y_t, \theta); \theta)}{\sigma(\gamma^{-1}(Y_t, \theta); \theta)} - \frac{1}{2} \frac{\partial \sigma}{\partial x}(\gamma^{-1}(Y_t, \theta); \theta)$$

Démonstration. Par hypothèse, X_t est un processus d'Itô et la fonction $\gamma(., \theta)$ est de classe C^2 ; donc d'après le Lemme d'Itô, Y_t est un processus d'Itô et sa différentielle stochastique est :

$$\begin{aligned} dY_t &= d\gamma(X_t; \theta) = \frac{1}{\sigma(x_t; \theta)} dX_t - \frac{\frac{\partial \sigma}{\partial x}(X_t; \theta)}{2\sigma^2(X_t; \theta)} d[X]_t \\ &= \frac{1}{\sigma(x_t; \theta)} (b(X_t; \theta)dt + \sigma(X_t; \theta)) - \frac{\frac{\partial \sigma}{\partial x}(X_t; \theta)}{2\sigma^2(X_t; \theta)} \sigma^2(X_t; \theta)dt \\ &= \left(\frac{b(X_t; \theta)}{\sigma(X_t; \theta)} - \frac{1}{2} \frac{\partial \sigma}{\partial x}(X_t; \theta) \right) dt + dW_t \end{aligned}$$

On déduit le résultat en remplaçant $X_t = \gamma^{-1}(Y_t, \theta)$ □

Rappel 3.1.1. 1 Une fonction indéfiniment différentiable f a une croissance au plus polynomiale si il existe un entier $p > 0$ tel que $|y|^{-p}|f(y)|$ soit bornée dans un voisinage de l'infini;

- Si $p = 1$ on parle de croissance au plus linéaire.
- Si $p = 2$ on parle de croissance au plus quadratique

2 Une fonction indéfiniment différentiable f est dite à croissance au plus exponentielle si il existe $\lambda \geq 0$ tel que $e^{-\lambda|y|}|f(y)|$ soit bornée au voisinage de l'infini.

Hypothèse 3.1.3. Pour $\theta \in \Theta$, $\mu_Y(Y_t, \theta)$, $\frac{\partial \mu_Y}{\partial y}(y; \theta)$ et $\frac{\partial^2 \mu_Y}{\partial y^2}(y; \theta)$ sont à croissance au plus exponentielle au voisinage de l'infini et $\lim_{y \rightarrow \underline{y} \text{ ou } \bar{y}} g(y, \theta) < +\infty$ où

$$g(y; \theta) = -\frac{1}{2} \left(\mu_Y^2(Y_t, \theta) + \frac{\partial \mu_Y}{\partial y}(y; \theta) \right)$$

1 **Borne gauche**

i) Si $y = 0^+$, il existe des constantes ϵ_0 , k et α telles que $\forall 0 < y < \epsilon_0$ et $\theta \in \Theta$, $\mu_Y(Y_t, \theta) \geq ky^{-\alpha}$

avec $(\alpha > 1$ et $k > 0)$ ou $(\alpha = 1$ et $k \geq \frac{1}{2})$.

ii) Si $y = -\infty$, il existe des constantes $E_0 > 0$, $K > 0$ telles que : $\forall y \leq -E_0$ et $\theta \in \Theta$, $\mu_Y(Y_t, \theta) \geq Ky$.

2 Borne droite

i) $\bar{y} = +\infty$, il existe des constantes $E_0 > 0$, $K > 0$ telles que :

$$\forall y \geq E_0 \text{ et } \theta \in \Theta, \mu_Y(Y_t, \theta) \leq Ky.$$

ii) Si $\bar{y} = 0^-$, il existe des constantes ϵ_0 , k et α telles que

$$\forall 0 > y > -\epsilon_0 \text{ et } \theta \in \Theta, \mu_Y(Y_t, \theta) \leq ky^{-\alpha}$$

avec $(\alpha > 1$ et $k > 0)$ ou $(\alpha = 1$ et $k \geq \frac{1}{2})$.

Remarque 3.1.1. Le processus Y_t est meilleur que le processus X_t , mais on ne peut pas encore développer la densité conditionnelle de Y_t autour d'une densité normale.

3.1.2 Deuxième étape : Pseudo-normalisation du processus $(Y_t)_{t \in [0, T]}$

On rappelle que le processus $(Y_t)_{t \in [0, T]}$ est solution de l'EDS $dY_t = \mu_Y(Y_t, \theta)dt + dW_t$.

Pour $\Delta > 0$, $\theta \in \Theta$ et $y_0 \in \mathbb{R}$, définissons le processus $(Z_t)_{t \in [0, T]}$ par :

$$Z_t = Z(\Delta, Y_t | y_0; \theta) = \frac{Y_t - y_0}{\sqrt{\Delta}}.$$

Proposition 3

Les différentes relations entre les densités de transition des processus $(X_t)_{t \in [0, T]}$, $(Y_t)_{t \in [0, T]}$ et $(Z_t)_{t \in [0, T]}$ sont données par :

$$p_Z(\Delta, z | y_0; \theta) = \sqrt{\Delta} p_Y(\Delta, \sqrt{\Delta}z + y_0 | y_0; \theta) \quad (3.3)$$

$$p_Y(\Delta, y | y_0; \theta) = \sigma(\gamma^{-1}(y, \theta); \theta) p_X(\Delta, \gamma^{-1}(y, \theta) | \gamma^{-1}(y_0, \theta); \theta) \quad (3.4)$$

Rappel 3.1.2. Si U et V sont deux variables aléatoires réelles continues de densités respectives f_U et f_V telles que $V = h(U)$ avec h une fonction réelle inversible différentiable, alors les densités f_U et f_V sont liées par : $f_V(v) = \left(\frac{dh}{du}(u) \right)^{-1} \cdot f_U(h^{-1}(v))$

Démonstration. Pour donc montrer l'égalité (3.1), on applique le rappel précédent à la fonction h définie par $h(y) = \frac{y - y_0}{\sqrt{\Delta}}$ et on a bien $Z_t = h(Y_t)$.

Montrons l'égalité (3.2). On a :

$$\begin{aligned} p_Y(\Delta, y \mid y_0; \theta) &= \frac{\partial}{\partial y} \mathbb{P}(Y_{t+\Delta} \leq y \mid Y_t = y_0) \\ &= \frac{\partial}{\partial y} \mathbb{P}(X_{t+\Delta} \leq \gamma^{-1}(y, \theta) \mid Y_t = \gamma^{-1}(y_0, \theta)) \\ &= \frac{\partial}{\partial y} \left[\int_{-\infty}^{\gamma^{-1}(y, \theta)} p_X(\Delta, x \mid \gamma^{-1}(y_0, \theta); \theta) dx \right] \\ &= \frac{\partial}{\partial y} \gamma^{-1}(y, \theta) \times p_X(\Delta, \gamma^{-1}(y, \theta) \mid \gamma^{-1}(y_0, \theta); \theta) \end{aligned}$$

Or $\frac{\partial}{\partial y} \gamma^{-1}(y, \theta) = \frac{1}{\frac{\partial}{\partial x}(\gamma^{-1}(y, \theta))} = \sigma(\gamma^{-1}(y, \theta); \theta)$. On déduit donc le résultat voulu.

□

Conséquence 4

On déduit que :

$$p_Y(\Delta, y \mid y_0; \theta) = \Delta^{-\frac{1}{2}} p_Z\left(\Delta, \frac{y - y_0}{\sqrt{\Delta}} \mid y_0; \theta\right); \quad (3.5)$$

$$p_X(\Delta, x \mid x_0; \theta) = [\sigma(x, \theta)]^{-1} p_Y(\Delta, \gamma(x, \theta) \mid \gamma(x_0, \theta); \theta). \quad (3.6)$$

Et donc :

$$p_X(\Delta, x \mid x_0; \theta) = [\sigma(x, \theta)]^{-1} \Delta^{-\frac{1}{2}} p_Z\left(\Delta, \frac{y - y_0}{\sqrt{\Delta}} \mid y_0; \theta\right) \quad (3.7)$$

avec $\gamma(x_0, \theta) = y_0$

3.1.3 Troisième étape : Approximation de la densité de transition p_Z

Définition 5

Le j -ième polynôme d'Hermite H_j est le polynôme de degré j , de coefficient dominant $(-1)^j$ et est défini par :

$$H_j(z) = e^{\frac{z^2}{2}} \frac{d^j}{dz^j} \left(e^{-\frac{z^2}{2}} \right) \quad (3.8)$$

Remarque 3.1.2. Les sept premiers polynômes d'Hermite sont : $H_0(z) = 1$, $H_1(z) = -z$, $H_2(z) = z^2 - 1$, $H_3(z) = -z^3 + 3z$, $H_4(z) = z^4 - 6z^2 + 3$, $H_5(z) = -z^5 + 10z^3 - 15z$, et $H_6(z) = z^6 - 15z^4 + 45z^2 - 15$.

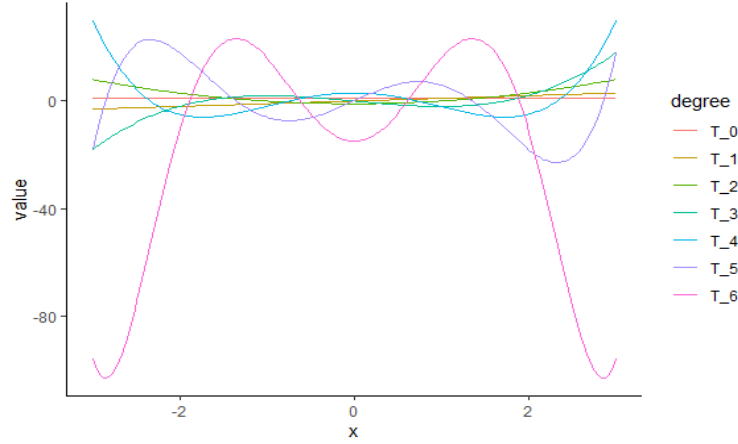


FIGURE 3.1 – Courbes des 7 premiers polynômes d'Hermite

Lemme 6

Les polynômes H_j satisfont :

i)

$$(-1)^{j+1}(j+1)H_j(z) = \frac{dH_{j+1}}{dz}(z) \quad \forall z \in \mathbb{R} \forall j \in \mathbb{N} \quad (3.9)$$

ii)

$$\int_{-\infty}^{+\infty} (2\pi)^{-\frac{1}{2}} e^{-\frac{w^2}{2}} H_j(w) H_k(w) dw = \begin{cases} j! & \text{si } j = k \\ 0 & \text{si } j \neq k \end{cases} \quad (3.10)$$

Démonstration. Voir annexe □

Grace à l'équation 3.4, on remarque qu'il nous suffit d'avoir l'expression estimée de la densité p_Z pour obtenir celle de p_X du départ.

En écrivant l'expansion polynomiale d'Hermite (démontrée en annexe) de la fonction $\frac{1}{\phi} p_Z$ où ϕ est la densité de la loi normale centrée réduite, ie $\phi(z) = \frac{1}{\sqrt{2\pi}} e^{-\frac{z^2}{2}}$, on a :

$$\begin{aligned} \frac{1}{\phi(z)} p_Z(\Delta, z | y_0; \theta) &= \sum_{j=0}^{+\infty} \frac{1}{j!} \int_{-\infty}^{+\infty} \frac{1}{\phi(s)} p_Z(\Delta, s | y_0; \theta) H_j(s) \phi(s) ds H_j(z) \\ &= \sum_{j=0}^{+\infty} \frac{1}{j!} \int_{-\infty}^{+\infty} p_Z(\Delta, s | y_0; \theta) H_j(s) ds H_j(z) \end{aligned}$$

On déduit que : $p_Z(\Delta, z | y_0; \theta) = \phi(z) \sum_{j=0}^{+\infty} \frac{1}{j!} \int_{-\infty}^{+\infty} p_Z(\Delta, s | y_0; \theta) H_j(s) ds H_j(z)$.

On pose donc

$$\eta_j(\Delta, y_0; \theta) = \frac{1}{j!} \int_{-\infty}^{+\infty} p_Z(\Delta, s | y_0; \theta) H_j(s) ds \quad (3.11)$$

et la densité de transition s'écrit alors :

$$p_Z(\Delta, z | y_0; \theta) = \phi(z) \sum_{j=0}^{+\infty} \eta_j(\Delta, y_0; \theta) H_j(z) \quad (3.12)$$

En prenant la somme partielle dans l'expansion précédente jusqu'au rang J , on notera donc :

$$p_Z^{(J)}(\Delta, z | y_0; \theta) = \phi(z) \sum_{j=0}^J \eta_j(\Delta, y_0; \theta) H_j(z) \quad (3.13)$$

où $\eta_j(\Delta, y_0; \theta)$ est toujours défini par l'équation 3.11. Et les relations entre $p_Z^{(J)}(\Delta, z | y_0; \theta)$, $p_X^{(J)}(\Delta, x | x_0; \theta)$ et $p_Y^{(J)}(\Delta, y | y_0; \theta)$ sont analogues à celles obtenues dans la **proposition 3**.

Remarque 3.1.3. On peut toujours calculer $\eta_0(\Delta, y_0; \theta)$ et :

$$\begin{aligned} \eta_0(\Delta, y_0; \theta) &= \frac{1}{0!} \int_{-\infty}^{+\infty} p_Z(\Delta, s | y_0; \theta) H_0(s) ds \\ &= \int_{-\infty}^{+\infty} p_Z(\Delta, s | y_0; \theta) ds \\ &= 1 \end{aligned}$$

Le Théorème suivant montre la convergence uniforme de $p_X^{(J)}$ vers p_X et ceci permet de justifier que $p_X^{(J)}$ est une estimation cohérente de p_X :

Théorème 7

Sous les hypothèses 3.1.1, 3.1.2 et 3.1.3, il existe $\bar{\Delta} > 0$ tel que pour tout $\Delta \in [0, \bar{\Delta}]$, $\theta \in \Theta$ et $(x, x_0) \in D_X^2$,

$$p_X^{(J)}(\Delta, x | x_0; \theta) \xrightarrow{J \rightarrow +\infty} p_X(\Delta, x | x_0; \theta)$$

De plus, cette convergence est uniforme pour $\theta \in \Theta$ et pour x_0 appartenant à un compact de D_X . Si $\sigma(\cdot, \cdot)$ est non-dégénérée, la convergence est davantage uniforme pour x appartenant à D_X tout entier. Si $\sigma(\cdot, \cdot)$ est dégénérée en 0, alors la convergence est uniforme en x sur tout intervalle de la forme $[\epsilon, +\infty[$ avec $\epsilon > 0$

Démonstration. Voir (1)

□

3.1.4 Expression explicite de la densité de transition \tilde{p}_x^{herm}

Lemme 8

Sous les hypothèses 3.1.1, 3.1.2 et 3.1.3, soit f une fonction telle que ses dérivées ont une croissance plus qu'exponentielle. Alors, pour tout $\Delta \in [0, \bar{\Delta}]$, $y_0 \in \mathbb{R}$ et $\theta \in \Theta$, il existe δ dans $[0, \Delta]$ tel que :

$$\mathbb{E}[f(Y_{t+\Delta}) | Y_t = y_0] = \sum_{k=1}^K \mathcal{L}_\theta^k(f)(y_0) \frac{\Delta^k}{k!} + \underbrace{\mathbb{E}[\mathcal{L}_\theta^{K+1}(f)(Y_{t+\delta}) | Y_t = y_0]}_{\text{Reste}} \frac{\Delta^{K+1}}{(K+1)!} \quad (3.14)$$

où \mathcal{L}_θ est le générateur infinitésimal du processus Y ; c'est à dire :

$$\mathcal{L}_\theta : f \mapsto \mu_Y(\cdot, \theta) \frac{\partial f}{\partial y}(\cdot) + \frac{1}{2} \frac{\partial^2 f}{\partial y^2}(\cdot) \quad \text{et} \quad \mathcal{L}_\theta^k(f) = \mathcal{L}_\theta^{k-1}(\mathcal{L}_\theta(f))$$

Aussi, il existe une constante P_K dépendant de K mais ne dépendant pas de f et de δ telle que :

$$|\mathbb{E}[\mathcal{L}_\theta^{K+1}(f)(Y_{t+\delta}) | Y_t = y_0]| \leq P_K$$

Démonstration. Posons $\varphi(s, y_0, \theta) = \mathbb{E}[f(Y_{t+s}) | Y_t = y_0]$ avec $s \in [0, \bar{\Delta}]$. φ est différentiable en s car $s \mapsto p_Y(s, y | y_0; \theta)$ l'est; en utilisant les Backward et Forward équations de Kolmogorov que vérifie la densité de transition du processus Y , on a :

$$\begin{aligned} \frac{\partial}{\partial s} \varphi(s, y_0, \theta) &= \int_{-\infty}^{+\infty} f(y) \frac{\partial}{\partial s} p_Y(s, y | y_0; \theta) dy \\ &= \int_{-\infty}^{+\infty} f(y) \left\{ \frac{\partial}{\partial y} \left(-\mu_Y(y, \theta) p_Y(s, y | y_0; \theta) + \frac{1}{2} \frac{\partial^2}{\partial y^2} p_Y(s, y | y_0; \theta) \right) \right\} dy \\ &= \underbrace{- \int_{-\infty}^{+\infty} f(y) \left\{ \frac{\partial}{\partial y} (\mu_Y(y, \theta) p_Y(s, y | y_0; \theta)) \right\} dy}_A + \underbrace{\frac{1}{2} \int_{-\infty}^{+\infty} f(y) \frac{\partial^2}{\partial y^2} p_Y(s, y | y_0; \theta) dy}_B \end{aligned}$$

En faisant des intégrations par parties de A et B , on obtient :

$$A = \int_{-\infty}^{+\infty} \frac{df}{dy}(y) \mu_Y(y, \theta) p_Y(s, y | y_0; \theta) dy \quad \text{et} \quad B = \int_{-\infty}^{+\infty} \frac{d^2 f}{dy^2} p_Y(s, y | y_0; \theta) dy$$

Donc :

$$\begin{aligned}\frac{\partial}{\partial s}\varphi(s, y_0, \theta) &= \int_{-\infty}^{+\infty} \left\{ \frac{df}{dy}(y)\mu_Y(y, \theta) + \frac{1}{2} \frac{d^2 f}{dy^2} \right\} p_Y(s, y | y_0; \theta) dy \\ &= \int_{-\infty}^{+\infty} \mathcal{L}_\theta(f)(y) p_Y(s, y | y_0; \theta) dy\end{aligned}$$

Donc :

$$\begin{aligned}\frac{\partial}{\partial s}\varphi(0, y_0, \theta) &= \int_{-\infty}^{+\infty} \mathcal{L}_\theta(f)(y) p_Y(0, y | y_0; \theta) dy = \int_{-\infty}^{+\infty} \mathcal{L}_\theta(f)(y) p_\theta(Y_t = y | Y_t = y_0) dy \\ &= \int_{-\infty}^{+\infty} \mathcal{L}_\theta(f)(y_0) p_\theta(Y_t = y | Y_t = y_0) dy \\ &= \mathcal{L}_\theta(f)(y_0) \underbrace{\int_{-\infty}^{+\infty} p_\theta(Y_t = y | Y_t = y_0) dy}_1.\end{aligned}$$

D'où : $\frac{\partial}{\partial s}\varphi(0, y_0, \theta) = \mathcal{L}_\theta(f)(y_0)$.

On montre par itération que : $\frac{\partial^i}{\partial s^i}\varphi(0, y_0, \theta) = \mathcal{L}_\theta^i(f)(y_0)$.

Appliquons le Théorème de Taylor à la fonction φ aux points $s = 0$ et $s = \Delta$, on obtient :

$$\varphi(\Delta, y_0, \theta) = \underbrace{\varphi(0, y_0, \theta)}_0 + \sum_{k=1}^K \frac{\partial^k}{\partial s^k}\varphi(0, y_0, \theta) \frac{\Delta^k}{k!} + Reste = \sum_{k=1}^K \mathcal{L}_\theta^k(f)(y_0) \frac{\Delta^k}{k!} + Reste$$

D'où le résultat □

Remarque 3.1.4. On n'impose pas au reste du développement de Taylor de converger vers 0 quand $K \rightarrow +\infty$ pour $\Delta \in [0, \bar{\Delta}]$ fixé, bien que \mathcal{L}_θ soit généralement non bornée.

La convergence du développement de Taylor doit être interpréter de la manière suivante :

$$\lim_{\Delta \rightarrow 0} \Delta^{-(K+1)} \left\{ \mathbb{E}[f(Y_{t+\Delta}) | Y_t = y_0] - \sum_{k=1}^K \mathcal{L}_\theta^k(f)(y_0) \frac{\Delta^k}{k!} \right\} = \frac{\mathcal{L}_\theta^{K+1}(f)(y_0)}{(K+1)!}$$

On rappelle que :

$$p_Z^{(j)}(\Delta, z | y_0; \theta) = \phi(z) \sum_{j=0}^J \eta_j(\Delta, y_0; \theta) H_j(z) \quad \text{avec} \quad \eta_j(\Delta, y_0; \theta) = \frac{1}{j!} \int_{-\infty}^{+\infty} p_Z(\Delta, s | y_0; \theta) H_j(s) ds$$

On a :

$$\begin{aligned}\eta_j(\Delta, y_0; \theta) &= \frac{1}{j!} \int_{-\infty}^{+\infty} p_Z(\Delta, s | y_0; \theta) H_j(s) ds \\ &= \frac{1}{j!} \int_{-\infty}^{+\infty} H_j(s) p_Y(\Delta, \sqrt{\Delta}s + y_0 | y_0; \theta) ds\end{aligned}$$

En effectuant le changement de variable $y = \sqrt{\Delta}s + y_0$, on obtient :

$$\begin{aligned}\eta_j(\Delta, y_0; \theta) &= \frac{1}{j!} \int_{-\infty}^{+\infty} H_j\left(\frac{y - y_0}{\sqrt{\Delta}}\right) p_Y(\Delta, y | y_0; \theta) dy \\ &= \frac{1}{j!} \mathbb{E}\left(H_j\left(\frac{Y_{t+\Delta} - y_0}{\sqrt{\Delta}}\right) \mid Y_t = y_0; \theta\right)\end{aligned}$$

En appliquant le **Lemme 6** précédent, on déduit que :

$$\begin{aligned}\eta_j^K(\Delta, y_0; \theta) &= \frac{1}{j!} \mathbb{E}\left(H_j\left(\frac{Y_{t+\Delta} - y_0}{\sqrt{\Delta}}\right) \mid Y_t = y_0; \theta\right) \\ &= \frac{1}{j!} \sum_{k=1}^K \mathcal{L}_\theta^k\left(H_j \circ \left(\frac{\cdot - y_0}{\sqrt{\Delta}}\right)\right)(y_0) \frac{\Delta^k}{k!} + \text{Reste}.\end{aligned}$$

Théorème 9

L'approximation de la densité de transition selon **Aït-Sahalia** est donnée par :

$$\begin{aligned}\tilde{p}_X^{\text{herm}}(\Delta, x | x_0; \theta) &\equiv p_X^{(J,K)}(\Delta, x | x_0; \theta) = [\sigma(x, \theta)]^{-1} \Delta^{-\frac{1}{2}} \phi\left(\frac{y - y_0}{\sqrt{\Delta}}\right) \\ &\quad \times \sum_{j=0}^J \eta_j^K(\Delta, y_0; \theta) H_j\left(\frac{y - y_0}{\sqrt{\Delta}}\right)\end{aligned}\quad (3.15)$$

Avec :

$$\eta_j^K(\Delta, y_0; \theta) = \frac{1}{j!} \sum_{k=1}^K \mathcal{L}_\theta^k\left(H_j \circ \left(\frac{\cdot - y_0}{\sqrt{\Delta}}\right)\right)(y_0) \frac{\Delta^k}{k!} + \text{Reste} \quad (3.16)$$

Démonstration. En se servant des résultats précédents. □

Remarque 3.1.5. $p_X^{(J,K)}$ est approximé de sorte qu'on ne retienne que dans tous les η_j $j = 0, \dots, J$ les termes d'ordre inférieur ou égale à $\Delta^{\frac{J}{2}}$. Cela revient à choisir K de sorte que le degré maximum de $p_Z^{(J,K)}$ en Δ soit $\frac{J}{2}$. Donc le choix de K est conditionné par celui de J et on peut simplement noter l'approximation $p_X^{(J,K)}$ par $\tilde{p}_X^{(J)}$. Il est plus intéressant de compiler les six premières approximations de $p_X^{(J)}$.

Illustration 3.1.1 (Expression de la densité approximée dans le modèle de Ornstein-Uhlenbeck). On a le processus stochastique $(X_t)_{t \in [0, T]}$ gouverné par l'équation différentielle stochastique :

$$dX_t = -\beta X_t dt + \sigma dW_t \quad \text{avec} \quad \theta = (\beta, \sigma).$$

Le processus vérifie bien les hypothèses.

1 Première étape : Transformation de Lamperti

On définit le processus $(Y_t)_{t \in [0, T]}$ par : $Y_t = \gamma(X_t; \theta) = \int_0^{X_t} \frac{du}{\sigma} = \frac{1}{\sigma} X_t$.

Donc $\gamma^{-1}(Y_t; \theta) = \sigma Y_t$, $D_X =]-\infty; +\infty[$, donc $D_Y =]-\infty; +\infty[$ et l'équation différentielle stochastique vérifiée par le processus $(Y_t)_{t \in [0, T]}$ est :

$$dY_t = -\beta Y_t dt + dW_t$$

Donc le générateur infinitésimal du processus $(Y_t)_{t \in [0, T]}$ est défini par :

$$\mathcal{L}_\theta(f)(y) = -\beta y \frac{\partial f}{\partial y}(y) + \frac{1}{2} \frac{\partial^2 f}{\partial y^2}(y). \quad (3.17)$$

Et $\mu_Y(y) = -\beta y$.

2 Pseudo-normalisation du processus $(Y_t)_{t \in [0, T]}$

On définit le processus $(Z_t)_{t \in [0, T]}$ par : $Z_t = \frac{Y_t - y_0}{\sqrt{\Delta}}$, $z = \frac{y - y_0}{\sqrt{\Delta}}$.

Donc la densité de transition du processus $(X_t)_{t \in [0, T]}$ est donnée par :

$$p_X(\Delta, x \mid x_0; \theta) = \frac{1}{\sigma} \Delta^{-\frac{1}{2}} p_Z\left(\Delta, \frac{y - y_0}{\sqrt{\Delta}} \mid y_0; \theta\right) \quad (3.18)$$

3 Approximation de la densité p_Z

On a montré que :

$$p_Z^{(J, K)}\left(\Delta, \frac{y - y_0}{\sqrt{\Delta}} \mid y_0; \theta\right) = \phi(z) \sum_{j=0}^J \eta_j^K(\Delta, y_0; \theta) H_j(Z)$$

Avec :

$$\eta_j^K(\Delta, y_0; \theta) = \frac{1}{j!} \sum_{k=1}^K \mathcal{L}_\theta^k \left(H_j \circ \left(\frac{\cdot - y_0}{\sqrt{\Delta}} \right) \right) (y_0) \frac{\Delta^k}{k!}$$

où dans les η_j^K , on ne retient que les termes d'ordre inférieur à $\Delta^{\frac{j}{2}}$.

Il est question de calculer les 6 premières approximations $\tilde{p}_Z^{(1)}, \tilde{p}_Z^{(2)}, \dots, \tilde{p}_Z^{(6)}$. En se

servant des équations obtenues dans : et le fait que $\mu_Y(y) = -\beta y$, on a :

$$\tilde{p}_Z^{(1)} = \phi \quad (3.19)$$

$$\tilde{p}_Z^{(2)} = \tilde{p}_Z^{(1)} + \phi \left[-\frac{\beta}{2} H_2 \right] \Delta \quad (3.20)$$

$$\tilde{p}_Z^{(3)} = \tilde{p}_Z^{(2)} - \phi \left[\frac{\beta^2 y_0}{2} H_1 \right] \Delta^{\frac{3}{2}} \quad (3.21)$$

$$\tilde{p}_Z^{(4)} = \tilde{p}_Z^{(3)} + \phi \left[\frac{\beta^2}{3} H_2 + \frac{\beta^2}{8} H_4 \right] \Delta^2 \quad (3.22)$$

$$\tilde{p}_Z^{(5)} = \tilde{p}_Z^{(4)} - \phi \left[-\frac{\beta^3 y_0}{6} H_1 - \frac{\beta^3 y_0}{4} H_3 \right] \Delta^{\frac{5}{2}} \quad (3.23)$$

$$\tilde{p}_Z^{(6)} = \tilde{p}_Z^{(5)} + \phi \left[\left\{ \frac{\beta^4 y_0^2}{8} - \frac{\beta^3}{6} \right\} H_2 - \frac{\beta^3}{6} H_4 - \frac{\beta^3}{48} H_6 \right] \Delta^3 \quad (3.24)$$

Il faut noter que les écritures ont été simplifiées; donc au lieu d'écrire par exemple $H_5(z)$ avec $z = \frac{y-y_0}{\sqrt{\Delta}}$, on a juste écrit H_5 ; ou alors, $\phi(z)$ a simplement été noté ϕ . On conclut donc que :

$$\begin{aligned} \tilde{p}_Z^{(6)}(\Delta, z | y_0; \theta) = & \phi + \phi \left[-\frac{\beta}{2} H_2 \right] \Delta - \phi \left[\frac{\beta^2 y_0}{2} H_1 \right] \Delta^{\frac{3}{2}} + \phi \left[\frac{\beta^2}{3} H_2 + \frac{\beta^2}{8} H_4 \right] \Delta^2 \\ & - \phi \left[-\frac{\beta^3 y_0}{6} H_1 - \frac{\beta^3 y_0}{4} H_3 \right] \Delta^{\frac{5}{2}} + \phi \left[\left\{ \frac{\beta^4 y_0^2}{8} - \frac{\beta^3}{6} \right\} H_2 - \frac{\beta^3}{6} H_4 - \frac{\beta^3}{48} H_6 \right] \Delta^3 \end{aligned} \quad (3.25)$$

Et

$$\tilde{p}_X^{(6)}(\Delta, x | x_0; \theta) = \frac{1}{\sigma} \Delta^{-\frac{1}{2}} \tilde{p}_Z^{(6)}(\Delta, z | y_0; \theta) \quad (3.26)$$

3.2 Approximation de la densité de transition en utilisant l'approximation Spline-Cubique

3.2.1 Première étape : Partition de l'intervalle de l'intervalle d'observation $[a; b]$

On supposera que les observations sont faites sur l'intervalle $[a; b]$ de sorte que $\min_{i=0, \dots, n} x_i \leq a < b \leq \max_{i=0, \dots, n} x_i$. On partitionne donc l'intervalle $[a; b]$ en L sous intervalles $([a_l; a_{l+1}])_{0 \leq l \leq L-1}$: $a_0 = a < a_1 < a_2 < \dots < a_{L-1} < a_L = b$ avec $L \in \mathbb{N}$, $a = \min_{i=0, \dots, n} x_i$, $b = \max_{i=0, \dots, n} x_i$, $h = \frac{b-a}{L}$ et $a_l = a_0 + lh$. Notre objectif est de trouver les approximations splines cubiques des fonctions $p_X(\Delta, \cdot | x_{i-1}; \theta)$ pour tout $i \in \{1, 2, \dots, n\}$. Nous supposons donc que cette

fonction est connue aux points a_0, a_1, \dots, a_L grace à l'approximation proposée par Haït Sahalia, c'est à dire la fonction $p_X(\Delta, \cdot | x_{i-1}; \theta)$ est connue aux points $\begin{pmatrix} a_0 \\ \tilde{p}_X^{herm}(\Delta, a_0 | x_{i-1}; \theta) \end{pmatrix}, \begin{pmatrix} a_1 \\ \tilde{p}_X^{herm}(\Delta, a_1 | x_{i-1}; \theta) \end{pmatrix}, \dots, \begin{pmatrix} a_{l-1} \\ \tilde{p}_X^{herm}(\Delta, a_{l-1} | x_{i-1}; \theta) \end{pmatrix}, \begin{pmatrix} a_L \\ \tilde{p}_X^{herm}(\Delta, a_L | x_{i-1}; \theta) \end{pmatrix}.$

Définition 10

Une spline cubique est une fonction $S : [a; b] \rightarrow \mathbb{R}$ de classe C^2 sur l'intervalle $[a; b]$ qui se réduit à un polynôme de degré 3 sur chaque sous intervalle $[a_l; a_{l+1}]$ de $[a; b]$.

Donc la spline cubique de la fonction $p_X(\Delta, \cdot | x_{i-1}; \theta)$ est la fonction par morceaux $\tilde{p}_X^{SPL}(\Delta, \cdot | x_{i-1}; \theta)$ de la forme :

$$\tilde{p}_X^{SPL}(\Delta, x | x_{i-1}; \theta) = \sum_{l=0}^{L-1} p_X^l(\Delta, x | x_{i-1}; \theta) \mathbb{1}_{[a_l; a_{l+1}]}(x)$$

où chaque morceau $p_X^l(\Delta, x | x_{i-1}; \theta)$ est un polynôme de degré 3 (une cubique).

3.2.2 Deuxième étape : Interpolation d'Hermite de $p_X^l(\Delta, \cdot | x_{i-1}; \theta)$ sur $[a_l; a_{l+1}]$

Considérons un seul sous intervalle $[a_l; a_{l+1}]$, on a noté $h = a_{l+1} - a_l$, et on cherche $p_X^l(\Delta, \cdot | x_{i-1}; \theta)$ sur $[a_l; a_{l+1}]$ un polynôme de degré 3 vérifiant :

$$p_X^l(\Delta, a_l | x_{i-1}; \theta) = \tilde{p}_X^{herm}(\Delta, a_l | x_{i-1}; \theta) \quad [p_X^l]'(\Delta, a_l | x_{i-1}; \theta) = p_l \quad (3.27)$$

$$p_X^l(\Delta, a_{l+1} | x_{i-1}; \theta) = \tilde{p}_X^{herm}(\Delta, a_{l+1} | x_{i-1}; \theta) \quad [p_X^l]'(\Delta, a_{l+1} | x_{i-1}; \theta) = p_{l+1} \quad (3.28)$$

Théorème 11

Il existe un unique polynôme $p_X^l(\Delta, \cdot | x_{i-1}; \theta)$ sur $[a_l; a_{l+1}]$ de degré 3 vérifiant les conditions 3.27 et 3.28 et il est défini par :

$$p_X^l(\Delta, x | x_{i-1}; \theta) = \tilde{p}_X^{herm}(\Delta, a_l | x_{i-1}; \theta) + (x - a_l)\delta y[a_l; a_{l+1}] + \frac{(x - a_l)(x - a_{l+1})}{h^2} [(p_{l+1} - \delta y[a_l; a_{l+1}])(x - a_l) - (p_l - \delta y[a_l; a_{l+1}])(x - a_{l+1})] \quad (3.29)$$

$$\text{avec } \delta y[a_l; a_{l+1}] = \frac{\tilde{p}_X^{herm}(\Delta, a_l | x_{i-1}; \theta) - \tilde{p}_X^{herm}(\Delta, a_{l+1} | x_{i-1}; \theta)}{h}$$

Démonstration. Notons par $\mathbb{R}_3[X]$ l'espace vectoriel des polynômes degré au plus 3 et considérons l'application :

$$\begin{aligned}\varphi : \mathbb{R}_3[X] &\rightarrow \mathbb{R}^4 \\ P &\mapsto (P(a_l); P(a_{l+1}); P'(a_l); P'(a_{l+1}))\end{aligned}$$

φ est une application linéaire injective entre deux espaces vectoriels de même dimension 4, donc φ est bijective et on déduit l'existence et l'unicité de $p_X^l(\Delta, \cdot | x_{i-1}; \theta)$ pour la donnée du quadruplet $(\tilde{p}_X^{herm}(\Delta, a_l | x_{i-1}; \theta); \tilde{p}_X^{herm}(\Delta, a_{l+1} | x_{i-1}; \theta); p_l; p_l + 1)$. On vérifie aisément que le polynôme donné dans le Théorème vérifie 3.27 et 3.28 \square

Définition 12

Le polynôme $p_X^l(\Delta, \cdot | x_{i-1}; \theta)$ ainsi obtenu est appelé **polynôme d'Hermite**.

3.2.3 Troisième étape : Choix des pentes $p_l \ l = 0; 1; \dots; L$

Pour construire la spline interpolante $\tilde{p}_X^{SPL}(\Delta, \cdot | x_{i-1}; \theta)$, il reste à déterminer les pentes $p_l \ l = 0; 1; \dots; L$ de sorte que $[\tilde{p}_X^{SPL}(\Delta, \cdot | x_{i-1}; \theta)]''$ soit continue; ie

$$[p_X^l(\Delta, a_l | x_{i-1}; \theta)]'' = [p_X^{l+1}(\Delta, a_l | x_{i-1}; \theta)]'' \quad l = 0, \dots, L-2 \quad (3.30)$$

En dérivant deux fois le polynôme d'Hermite $p_X^l(\Delta, \cdot | x_{i-1}; \theta)$ obtenue au Théorème précédent, on obtient :

$$\begin{aligned}[p_X^l(\Delta, a_l | x_{i-1}; \theta)]'' &= \frac{2}{h}(2p_{l+1} + p_l - 3\delta y[a_l; a_{l+1}]) \\ [p_X^l(\Delta, a_{l-1} | x_{i-1}; \theta)]'' &= -\frac{2}{h}(2p_l + p_{l+1} - 3\delta y[a_l; a_{l+1}])\end{aligned}$$

La condition 3.30 donne : $\frac{2}{h}(2p_{l+1} + p_l - 3\delta y[a_l; a_{l+1}]) = -\frac{2}{h}(2p_{l+1} + p_{l+2} - 3\delta y[a_{l+1}; a_{l+2}]) \quad l = 0, \dots, L-2$; ie :

$$p_l + 4p_{l+1} + p_{l+2} = 3(\delta y[a_l; a_{l+1}] + \delta y[a_{l+1}; a_{l+2}]) \quad l = 0, \dots, L-2 \quad (3.31)$$

Ceci donne un système de $(L-1)$ équations à $(L+1)$ inconnues. Nous ajouterons donc une nouvelle condition dite du type spline.

Définition 13

Une spline S est dite naturelle si $S''(a) = S''(b) = 0$

On complète donc le système 3.31 par les deux équations $p_0 = p_L = 0$.

Proposition 14

Pour la spline naturelle, les valeurs de p_l $l = 0, \dots, L$ sont données par : $p_0 = p_L = 0$ et $p = (p_1, p_2, \dots, p_{L-2})$ est solution du système linéaire :

$$Ap = C \quad (3.32)$$

$$\text{où } A = \text{Tridiag}(1, 4, 1) \text{ et } C = 3 \begin{pmatrix} \delta y[a_0; a_1] + \delta y[a_1; a_2] \\ \delta y[a_1; a_2] + \delta y[a_2; a_3] \\ \cdot \\ \cdot \\ \cdot \\ \delta y[a_{L-2}; a_{L-1}] + \delta y[a_{L-1}; a_L] \end{pmatrix}$$

3.2.4 Expression de la densité de transition $\tilde{p}_X^{SPL}(\Delta, \cdot \mid x_{i-1}; \theta)$

Au vue des résultats précédents, on a le Théorème suivant :

Théorème 15

La spline cubique naturelle $\tilde{p}_X^{SPL}(\Delta, \cdot \mid x_{i-1}; \theta)$ de la fonction $p_X(\Delta, \cdot \mid x_{i-1}; \theta)$ sur l'intervalle $[a, b]$, aux nœuds $\begin{pmatrix} a_0 \\ \tilde{p}_X^{herm}(\Delta, a_0 \mid x_{i-1}; \theta) \end{pmatrix}, \begin{pmatrix} a_1 \\ \tilde{p}_X^{herm}(\Delta, a_1 \mid x_{i-1}; \theta) \end{pmatrix}, \dots, \begin{pmatrix} a_{l-1} \\ \tilde{p}_X^{herm}(\Delta, a_{l-1} \mid x_{i-1}; \theta) \end{pmatrix}, \begin{pmatrix} a_L \\ \tilde{p}_X^{herm}(\Delta, a_L \mid x_{i-1}; \theta) \end{pmatrix}$ est donnée par :

$$\tilde{p}_X^{SPL}(\Delta, x \mid x_{i-1}; \theta) = \sum_{l=0}^{L-1} p_X^l(\Delta, x \mid x_{i-1}; \theta) \mathbb{1}_{[a_l; a_{l+1}]}(x) \quad (3.33)$$

où $p_X^l(\Delta, x \mid x_{i-1}; \theta)$ est le polynôme d'Hermite sur le sous intervalle $[a_l; a_{l+1}]$ dont les pentes p_l sont solutions de l'équation 3.32.

SIMULATIONS

Dans ce chapitre, nous étudions un modèle particulier qui est celui de Ornstein-Uhlenbeck dont nous implémenterons exclusivement la méthode obtenue par l'expansion polynomiale d'Hermite décrite au chapitre précédent. Nous rappelons l'équation différentielle stochastique vérifiée par ce processus :

$$\begin{cases} dX_t = -\beta X_t dt + \sigma dW_t \\ X_0 = x_0 \in \mathbb{R}, \quad 0 \leq t \leq T \end{cases}, \quad (4.1)$$

Le théorème d'existence et d'unicité nous garantit l'existence d'une solution car les coefficient Drift et de diffusion sont linéaires ; et la solution est donnée par :

$$X_t = e^{-\beta t} \left(x_0 + \sigma \int_0^t e^{\beta u} dW_u \right). \quad (4.2)$$

Remarque 4.0.1. *On a montré que :*

$$X_t \sim \mathcal{N}(x_0 e^{-\beta t}; \sigma^2 \frac{1 - e^{-2\beta t}}{2\beta}).$$

et également que le processus X est un processus Gaussien de covariance :

$$\mathbb{Cov}(X_t, X_{t+s}) = e^{-\beta s} \sigma^2 \frac{1 - e^{-2\beta t}}{2\beta}. \quad (4.3)$$

En effet, $X_{t+s} = e^{-\beta s} X_t + \sigma e^{-\beta(t+s)} \int_t^{t+s} e^{\beta u} dW_u$; donc $X_{t+s} - e^{-\beta s} X_t = \sigma e^{-\beta(t+s)} \int_t^{t+s} e^{\beta u} dW_u$ est

indépendante de X_t . Ainsi, puisque $X_{t+s} = X_{t+s} - e^{-\beta s} X_t + e^{-\beta s} X_t$, on peut écrire :

$$\begin{aligned}
 \mathbb{C}ov(X_t, X_{t+s}) &= \mathbb{C}ov(X_t, X_{t+s} - e^{-\beta s} X_t + e^{-\beta s} X_t) \\
 &= \underbrace{\mathbb{C}ov(X_t, X_{t+s} - e^{-\beta s} X_t)}_{=0} + \mathbb{C}ov(X_t, e^{-\beta s} X_t) \\
 &= e^{-\beta s} \mathbb{C}ov(X_t, X_t) \\
 &= e^{-\beta s} \mathbb{V}ar(X_t) = e^{-\beta s} \sigma^2 \frac{1 - e^{-2\beta t}}{2\beta}.
 \end{aligned}$$

4.1 Simulations des densités de transition

Il a été démontré au **chapitre 2** que la densité de transition exacte du processus est Gaussienne et est donnée par :

$$p_X(\Delta, x | x_0; \theta) = \left(\frac{\pi \sigma^2 (1 - e^{-2\beta \Delta})}{\beta} \right)^{-\frac{1}{2}} \exp \left\{ -\beta \frac{(x - x_0 e^{-\beta \Delta})^2}{\sigma^2 (1 - e^{-2\beta \Delta})} \right\}; \quad (4.4)$$

et au **chapitre 3** que la densité de transition approximée par les polynômes d'Hermite est :

$$\tilde{p}_X^{(6)}(\Delta, x | x_0; \theta) = \frac{1}{\sigma} \Delta^{-\frac{1}{2}} \tilde{p}_Z^{(6)}(\Delta, z | y_0; \theta) \quad (4.5)$$

avec :

$$\begin{aligned}
 \tilde{p}_Z^{(6)}(\Delta, z | y_0; \theta) &= \phi + \phi \left[-\frac{\beta}{2} H_2 \right] \Delta - \phi \left[\frac{\beta^2 y_0}{2} H_1 \right] \Delta^{\frac{3}{2}} + \phi \left[\frac{\beta^2}{3} H_2 + \frac{\beta^2}{8} H_4 \right] \Delta^2 \\
 &\quad - \phi \left[-\frac{\beta^3 y_0}{6} H_1 - \frac{\beta^3 y_0}{4} H_3 \right] \Delta^{\frac{5}{2}} + \phi \left[\left\{ \frac{\beta^4 y_0^2}{8} - \frac{\beta^3}{6} \right\} H_2 - \frac{\beta^3}{6} H_4 - \frac{\beta^3}{48} H_6 \right] \Delta^3
 \end{aligned} \quad (4.6)$$

Nous simulons les densités lorsque le processus est observé pour une taille d'échantillon $n=3000$ fixe et faisons varier l'intervalle d'observation de temps Δ ; pour $x_0 = 0, 1$ et pour le vecteur de paramètres $(\beta = 0, 01; \sigma = 0, 02)$.

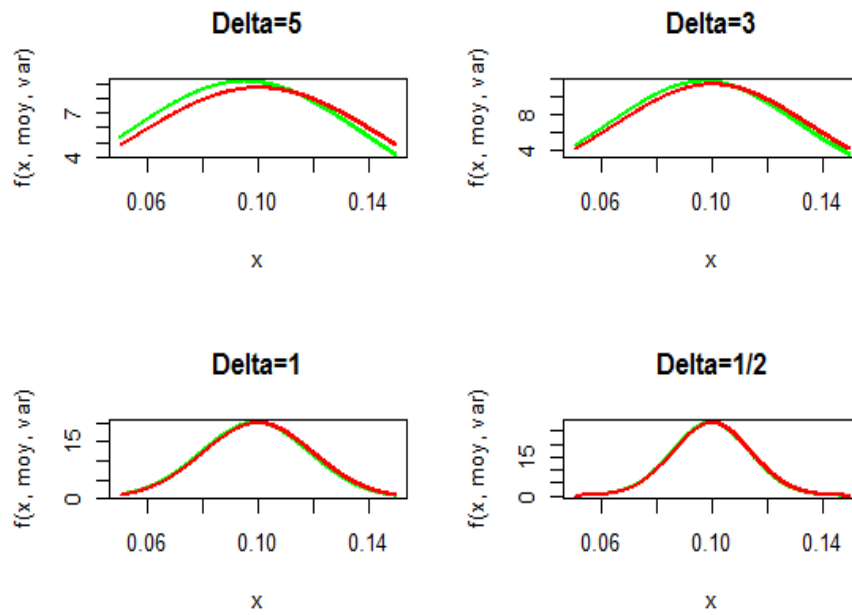


FIGURE 4.1 – Courbes des densités de transition

La courbe en vert représente la densité de transition exacte p_X et celle en rouge est la densité de transition approximée par l'expansion d'Hermite $\tilde{p}_Z^{(6)}$.

Nous avons tracer les densités de transitions lorsque le phénomène est observé chaque 5 ans ($\Delta = 5$), chaque 3 ans ($\Delta = 3$), annuellement ($\Delta = 1$) et semi-annuellement ($\Delta = 1/2$).

OBSERVATIONS

On remarque que les courbes de p_X et de $\tilde{p}_Z^{(6)}$ sont distinctes pour $\Delta = 5$ et pour $\Delta = 3$; et cette distinction est d'autant plus flagrante pour $\Delta = 5$ (grand). Mais pour $\Delta = 1/2$, on a du mal à distinguer les deux courbes car elles se superposent.

INTERPRETATION

Les densités de transition p_X (équation 4.4) et $\tilde{p}_Z^{(6)}$ (équation 4.5) coïncident lorsque Δ est petit : ceci justifie l'approximation de p_X par $\tilde{p}_Z^{(6)}$.

Remarque 4.1.1. *l'hypothèse de considérer Δ très petit n'est pas anodine car combien de temps il nous faudrait pour par exemple avoir 3000 observations si on observe un phénomène chaque trois ans ? Il nous en faudra 1000 ans, irréaliste et même impossible en coût de temps. Donc l'hypothèse que Δ soit assez petit est justifiée et l'approximation de p_X par $\tilde{p}_Z^{(6)}$ a donc tout son sens, comme en témoigne d'avantage le graphique des densités mensuelles suivant où les densités sont totalement confondues :*

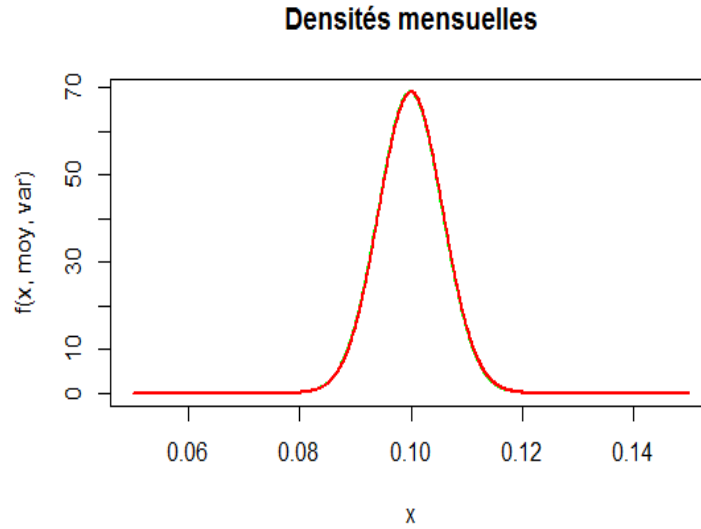


FIGURE 4.2 – Densités de transition mensuelles

4.2 Simulation des données

Discrétisation pour obtenir un AR(1)

Pour tout $T \in [0, T]$, on a :

$$X_{t+\Delta} = e^{-\beta\Delta}X_t + \sigma e^{-\beta(t+\Delta)} \int_t^{t+\Delta} e^{\beta u} dW_u. \quad (4.7)$$

En posant donc $t_i = i\Delta$, on a donc :

$$X_{(i+1)\Delta} = e^{-\beta\Delta}X_{i\Delta} + \sigma e^{-\beta(i+1)\Delta} \int_{i\Delta}^{(i+1)\Delta} e^{\beta u} dW_u. \quad (4.8)$$

Si on pose $Y_i = X_{i\Delta}$, alors Y_{i+1} est la somme de $e^{-\beta\Delta}Y_i$ et d'une variable aléatoire Gaussienne centrée de variance $\sigma^2 \frac{1 - e^{-2\beta\Delta}}{2\beta}$. On peut donc considérer l'équation AR(1) (ie modèle autorégressif ou ARMA(1,0) **ARMA** : modèle auto-régressif à moyenne mobile) :

$$Y_{i+1} = e^{-\beta\Delta}Y_i + \sigma \sqrt{\frac{1 - e^{-2\beta\Delta}}{2\beta}} \varepsilon_{i+1}. \quad (4.9)$$

où ε est un bruit blanc $BB(0, 1)$.

Simulation des données

Pour simuler les données, nous pouvons utiliser l'équation AR(1) 4.9 déjà démontré :

$$Y_{i+1} = e^{-\beta\Delta}Y_i + \sigma \sqrt{\frac{1 - e^{-2\beta\Delta}}{2\beta}}\varepsilon_{i+1}.$$

où ε est un bruit blanc $BB(0, 1)$.

C'est cette AR(1) qui nous permettra de simuler la trajectoire du processus et donc de pouvoir avoir des observations simulées.

Ci-après nous illustrons une trajectoire du processus pour $x_0 = 0.1, \beta = 0.01$ et $\sigma = 0.02$ dont le code R est en annexe :

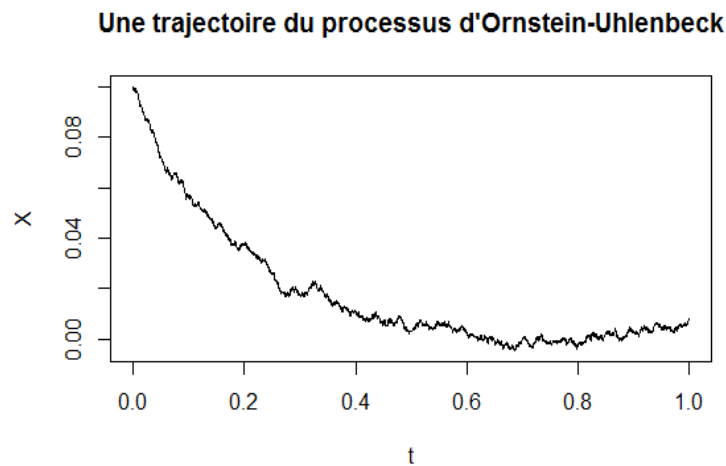


FIGURE 4.3 – Une simulation de la trajectoire du processus d'Ornstein-Uhlenbeck

Grace à ce jeu de données obtenu, nous estimerons donc les paramètres à partir de ces observations.

4.3 Estimation des paramètres par bootstrap

Le bootstrap est une méthode qui permet estimer la variance d'un estimateur.

A partir du jeu de données obtenu à la section précédente précédente, on estime le vecteur des paramètres (σ_1, β_1) , qui à son tour nous permettra de simuler un nouveau jeu de données dont on estimera à nouveau un vecteur de paramètres (σ_1, β_1) . La méthode donc consiste à :

- 1** On simule une trajectoire de données $Data_0 = (X_1, \dots, X_n)$
- 2** On estime le vecteur de paramètre θ_n^{herm} en utilisant la densité approximée par l'expansion polynomiale d'Hermite.

3 On répète les étapes 1 et 2 B fois pour obtenir $\theta_{n,1}^{herm}, \theta_{n,2}^{herm}, \dots, \theta_{n,B}^{herm}$.

4 On obtient :

$$\mathbb{V}ar_{boot} = \frac{1}{B} \sum_{b=1}^B \left(\theta_{n,b}^{herm} - \frac{1}{B} \sum_{r=1}^B \theta_{n,r}^{herm} \right)^2 \quad (4.10)$$

En appliquant la méthode Bootstrap décrite ci haut dans notre modèle pour $B = 100$. Nous obtenons les différentes estimations suivantes dont la première colonne représente les estimations de σ et la seconde celles de β .

	[,1]	[,2]
[1,]	0.01989535	2.102692e-01
[2,]	0.01954164	-1.338222e-03
[3,]	0.01973076	4.864529e-04
[4,]	0.01938051	5.234736e-01
[5,]	0.01901667	-1.638740e-03
[6,]	0.01884030	-3.900964e-03
[7,]	0.01780913	9.750412e-01
[8,]	0.01748105	3.786725e-03
[9,]	0.01738281	3.175013e-01
[10,]	0.01708003	4.591830e-03
[11,]	0.01680820	2.568264e-01
[12,]	0.01680911	2.165305e-03
[13,]	0.01665117	-7.769781e-03
[14,]	0.01647752	1.073023e+00
[15,]	0.01600632	-1.612986e-05
[16,]	0.01568814	5.228096e-02
[17,]	0.01574837	-2.961527e-03
[18,]	0.01641945	-6.324741e-01
[19,]	1.00000000	2.000000e+00
[20,]	0.88330871	2.843015e+00
[21,]	0.83921345	4.376035e-04
[22,]	0.82956010	-2.771360e-06
[23,]	0.84470618	6.130837e-04
[24,]	0.85012902	-3.552710e-06
[25,]	0.82930097	8.107472e-01
[26,]	0.77964965	1.064961e+00
[27,]	0.76195313	2.239690e-04
[28,]	0.71658932	1.560866e+00
[29,]	0.66714952	8.422451e-01
[30,]	0.63083299	9.098268e-01
[31,]	0.57539483	2.270505e+00
[32,]	0.51892769	1.340638e+00
[33,]	0.49596758	4.769170e-01
[34,]	0.46372293	1.487481e+00
[35,]	0.44251658	1.447637e-04
[36,]	0.42801606	1.451646e+00

```
[37,] 0.42318401 3.872427e-03
[38,] 0.41017649 1.596073e+00
[39,] 0.36244196 2.240764e+00
[40,] 0.33819143 1.023269e+00
[41,] 0.32016397 1.113201e+00
[42,] 0.31189673 -3.856216e-05
[43,] 0.30918144 1.231090e-03
[44,] 0.30862115 3.420838e-01
[45,] 0.29224588 1.248090e+00
[46,] 0.28701691 9.912374e-04
[47,] 0.28283222 7.521745e-01
[48,] 0.26993975 9.957476e-01
[49,] 0.25904011 6.614239e-01
[50,] 0.24948045 5.940870e-01
[51,] 0.23113721 1.722606e+00
[52,] 0.22591622 5.135229e-04
[53,] 0.22106265 3.695520e-04
[54,] 0.22290419 1.125608e-03
[55,] 0.22294787 -3.653413e-05
[56,] 0.22206100 3.710287e-03
[57,] 0.22292639 -6.172432e-05
[58,] 0.22054645 9.368842e-04
[59,] 0.21911651 2.521586e-04
[60,] 0.21244810 3.351641e-01
[61,] 0.20960642 8.628493e-03
[62,] 0.21104820 7.621863e-03
[63,] 0.21049067 -3.268811e-05
[64,] 0.20974287 -3.716573e-05
[65,] 0.20124273 8.890525e-01
[66,] 0.18667395 1.380773e+00
[67,] 0.18064441 4.023190e-04
[68,] 0.17100552 1.456657e+00
[69,] 0.15813155 1.279491e+00
[70,] 0.15187299 -7.349749e-05
[71,] 0.15142641 -1.157684e-05
[72,] 0.14298273 1.110214e+00
[73,] 0.13402660 8.358080e-01
```

```
[74,] 0.12579217 8.782475e-01
[75,] 0.81426317 -1.559602e+04
[76,] 1.00000000 2.000000e+00
[77,] 0.90646757 1.598897e+00
[78,] 0.79541197 2.351925e+00
[79,] 0.76810431 2.249998e-02
[80,] 0.76074813 2.194513e-04
[81,] 0.75281936 -8.041210e-06
[82,] 0.72270284 9.692536e-01
[83,] 0.70625685 3.648514e-04
[84,] 0.65326581 1.885622e+00
[85,] 0.63286104 1.442888e-04
[86,] 0.59645890 1.638096e+00
[87,] 0.56938606 6.589286e-04
[88,] 0.57434464 1.633315e-04
[89,] 0.54046835 2.134210e+00
[90,] 0.51605331 2.000847e-04
[91,] 0.51264267 3.208198e-04
[92,] 0.52427175 2.036929e-04
[93,] 0.51394602 2.065303e-04
[94,] 0.52112201 1.675279e-04
[95,] 0.52119208 2.181198e-01
[96,] 0.50660459 -9.822592e-06
[97,] 0.50845481 2.128634e-04
[98,] 0.50367978 1.529984e-04
[99,] 0.50576067 -4.678368e-05
[100,] 0.50081092 2.795757e-01
```

On peut donc calculer la variance de chaque paramètre et on obtient :

```
> Sd  
[1] 0.2807745 1559.6564432
```

Le biais des estimations vaut :

```
> Biais  
[1] 0.3640243 -155.6363502
```

Et l'erreur quadratique de chaque paramètre est :

```
> MSE  
[1] 2.113481e-01 2.456751e+06
```

IMPLICATIONS PEDAGOGIQUES

La rédaction de ce mémoire nous a donné l'opportunité de faire nos premiers pas dans la recherche scientifique et de développer les aptitudes à émettre des raisonnements scientifiques et à développer les capacités sur le plan humain telles que le vivre ensemble, le partage avec les collègues. Dans les lignes qui suivent, nous présentons les différents apports de la rédaction de notre mémoire qui montrent l'impact pédagogique dans notre profession d'enseignant.

Apport dans la mobilisation et la construction du savoir à enseigner

Pour mener à bien notre travail, il était impératif pour nous de comprendre une situation et sa modélisation mathématique par les équations différentielles stochastiques, ceci en vue d'obtenir des outils mathématiques (estimateurs, estimation) qui peuvent nous servir à prendre des décisions. On a donc été amené à effectuer des recherches dans des documents (livres et articles scientifiques mentionnés en bibliographie...) et sites web. Ces recherches peuvent être perçues comme ressources pédagogiques dont doit se servir un enseignant pour élaborer un cours : faire des recherches dans les livres-programmes et livres au programme afin d'entrer en possession des objectifs du cours, du savoir à enseigner.

Après avoir effectué les recherches dans les documents et articles scientifiques, nous nous sommes servis du contenu de ceux-ci pour élaborer ce document dans le strict respect des directives de rédaction d'un mémoire ; en effet, nous nous sommes servis du **Guide du mémoire**. Cette démarche fondamentale nous indique comment est ce qu'on doit se comporter lorsqu'on monte le cours à enseigner. C'est à dire, nous devons respecter les étapes d'une leçon selon l'APC et toutes les directives prescrites.

Apport dans la construction du savoir scientifique et du raisonnement scientifique

Le thème de ce mémoire nous a permis d'approfondir les notions de probabilités avancées et de découvrir encore plus ce champ des probabilités et statistiques qui pourras nous permettre de mieux enseigner ces notions en classes de Terminales (pour la probabilité) et en classe de Premières (pour les statistiques).

Apport dans l'initiation aux nouvelles technologies de l'information et de la communication

Les outils technologiques tels que l'ordinateur, les logiciels informatiques (WORD, LATEX, Geogebra, Mathlab, R) et internet ont été d'un appui considérable dans la rédaction de ce mémoire. Ces outils TIC pourront nous aider dans la confection de nos sujets d'évaluations et même de nos notes de cours, étant donné l'avènement du monde des TIC et la situation sanitaire qui pourrait contraindre à faire des cours à distance et en ligne. La familiarisation avec les outils TIC dont ce mémoire nous a offert nous sera donc d'une grande importance dans l'exercice de notre métier.

Conclusion et perspectives

Parvenu au terme de notre travail où il était question de comprendre et présenter deux approches d'estimation par maximum de vraisemblance notamment l'expansion polynomiale d'Hermite décrite par Aït-Sahalia et l'approximation spline-cubique, il ressort que l'approximation décrite par Aït-Sahalia prend en compte deux principaux paramètres que sont :

- l'intervalle d'observation du phénomène observé. Nous avons opté pour un intervalle de temps Δ que nous suggérons très petit.
- Le terme J qui est le nombre de terme utilisé dans l'expansion polynomiale d'Hermite de p_Z .

Nous sommes passés par une série de transformations du processus de départ X_t pour arriver à une approximation $p_X^{(J)}$ de p_X dont la convergence de la première vers la deuxième est assurée lorsque $J \rightarrow +\infty$. Avec en amont un rappel sur la théorie des EDS, nous avons donc succinctement présenté cette méthode et à l'aide d'un exemple, le processus d'Ornstein-Uhlenbeck, dont nous avons fait certaines illustrations de cette méthode et présenté les propriétés.

Il était également question pour nous de proposer une estimation de la densité de transition par la méthode des splines-cubiques dont nous nous sommes basés sur les travaux de Aït-Sahalia pour en proposer un : p_X^{SPL} obtenue à l'équation 3.33. Les questions qu'on peut se poser sont celles de savoir : Qu'en est-il de la convergence de p_X^{SPL} vers p_X ? Les estimateurs de maximum de vraisemblance obtenus à l'aide de p_X^{SPL} ont-ils de bonnes propriétés à l'instar de la consistance forte ? Ont-ils la propriété de la normalité asymptotique ? Là sont autant de questions qui pourront élargir notre champ de recherche et constituer de belles perspectives dans cette nouvelle école qu'est la recherche dont nous faisons nos premiers pas.

Bibliographie

- [1] Yacine Ait-Sahalia. Maximum-likelihood estimation of discretely sampled diffusions : A closed-form approach. *Center for Research in Security Prices (CRSP) Working Paper*, (467), 1998.
- [2] Yacine Ait-Sahalia. Transition densities for interest rate and other nonlinear diffusions. In *Quantitative Analysis In Financial Markets : Collected Papers of the New York University Mathematical Finance Seminar (Volume II)*, pages 1–34. World Scientific, 2001.
- [3] Jinyuan Chang, Song Xi Chen, et al. On the approximate maximum likelihood estimation for diffusion processes. *The Annals of Statistics*, 39(6) :2820–2851, 2011.
- [4] Jean-Pierre Demailly. *Analyse numérique et équations différentielles*. EDP sciences, 2021.
- [5] Stefano M Iacus. *Simulation and inference for stochastic differential equations : with R examples*. Springer Science & Business Media, 2009.
- [6] Alan Lindsay and Dominic Brecher. Results on the cev process, past and present. *Past and Present (March 9, 2010)*, 2010.
- [7] Bernt Oksendal. *Stochastic differential equations : an introduction with applications*. Springer Science & Business Media, 2013.

Annexe

Preuve du Lemme 6 : chapitre 3

Démonstration. On note $\omega(t) = e^{-\frac{t^2}{2}}$ et considère la mesure μ sur \mathbb{R} de densité ω par rapport à la mesure de Lebesgue, ie : $d\mu(t) = \omega(t)dt$. On note par $L^2(\omega)$ l'espace $L^2(\mathbb{R}, \mu)$. On munit $L^2(\omega)$ du produit scalaire $(\cdot | \cdot)_{L^2(\omega)}$ défini par $(f | g)_{L^2(\omega)} = \int_{\mathbb{R}} f(t)\overline{g(t)}e^{-\frac{t^2}{2}}dt$. On va montrer en un premier temps que l'espace des polynômes est dense dans $L^2(\omega)$ enfin de justifier l'existence des intégrales 3.10 .

1 Montrons que l'espace $L^2(\omega)$ contient tous les polynômes.

Il suffit de montrer que $L^2(\omega)$ contient tous les monômes m_n où $m_n(t) = t^n$.

Comme pour tout $n \in \mathbb{N}$, $t^2(t^{2n}e^{-\frac{t^2}{2}}) \rightarrow 0$ quand $|t| \rightarrow +\infty$, alors $\int_{\mathbb{R}} t^{2n}e^{-\frac{t^2}{2}}dt < +\infty$.

C'est à dire $(t^n | t^n)_{L^2(\omega)} = \int_{\mathbb{R}} t^{2n}e^{-\frac{t^2}{2}}dt < +\infty$; donc $m_n \in L^2(\omega)$.

2 Montrons que si $f \in L^2(\omega)$ alors la formule

$$F(z) = \int_{\mathbb{R}} e^{-2\pi itz} f(t)\omega(t)dt \quad (11)$$

définit une fonction holomorphe sur \mathbb{C} .

Posons $\varphi(t, z) = e^{-2\pi itz} f(t)\omega(t)$. Soit $B > 0$; on a :

$$\begin{aligned} |\varphi(t, z)| &= e^{2\pi t(Imz)} |f(t)|\omega(t) \\ &\leq e^{2\pi B|t|} |f(t)|\omega(t) \quad \text{pour } |Imz| \leq B \end{aligned}$$

Par l'inégalité de Cauchy-Schwartz, on a

$$\begin{aligned} \int_{\mathbb{R}} e^{2\pi B|t|} |f(t)|\omega(t)dt &\leq \left(\int_{\mathbb{R}} (e^{2\pi B|t|} \sqrt{\omega(t)})^2 dt \right)^{\frac{1}{2}} \left(\int_{\mathbb{R}} (|f(t)| \sqrt{\omega(t)})^2 dt \right)^{\frac{1}{2}} \\ &= \left(\int_{\mathbb{R}} e^{4\pi B|t|} e^{-\frac{t^2}{2}} dt \right)^{\frac{1}{2}} \times \|f\|_{L^2(\omega)}. \end{aligned}$$

Or $4\pi B|t| - \frac{t^2}{2} \leq -\frac{t^2}{4}$ pour $|t|$ assez grand.

Donc $\int_{\mathbb{R}} e^{2\pi B|t|} |f(t)| \omega(t) dt \leq \left(\int_{\mathbb{R}} e^{-\frac{t^2}{4}} dt \right)^{\frac{1}{2}} \times \|f\|_{L^2(\omega)} < +\infty$.

La version holomorphe du Théorème de dérivation sous le signe intégral peut alors être utilisée pour obtenir l'holomorphie de F sur la bande $\{z \in \mathbb{C} : |Im z| \leq B\}$ et donc sur \mathbb{C} puisque c'est vrai pour tout $B > 0$. De plus ce Théorème permet de dériver indéfiniment sous le signe intégral et on obtient donc :

$$F^{(n)}(z) = \int_{\mathbb{R}} (-2\pi i t)^n e^{-2\pi i t z} f(t) \omega(t) dt \quad \forall z \in \mathbb{C} \quad \forall n \in \mathbb{N}.$$

En particulier :

$$F^{(n)}(0) = (-2\pi i)^n \int_{\mathbb{R}} t^n f(t) \omega(t) dt \quad \forall n \in \mathbb{N}. \quad (12)$$

3 Montrons donc que l'ensemble des polynômes est dense dans $L^2(\omega)$.

Il suffit de montrer que si $f \in L^2(\omega)$ est orthogonale à tous les polynômes, alors $f = 0$.

Or, lorsque c'est le cas, on a en particulier $(f | m_n)_{L^2(\omega)} = 0$ pour tout $n \in \mathbb{N}$.

Et cela signifie que $\int_{\mathbb{R}} t^n f(t) \omega(t) dt = 0$ ie $F^{(n)}(0) = 0 \quad \forall n \in \mathbb{N}$.

Mais comme F est entière, on a :

$$F(z) = \sum_{n=0}^{\infty} \frac{F^{(n)}(0)}{n!} z^n \quad \forall z \in \mathbb{C}.$$

On obtient par conséquent $F \equiv 0$.

Or :

$$\begin{aligned} \int_{\mathbb{R}} |f(t)| \omega(t) dt &\leq \left(\int_{\mathbb{R}} (\sqrt{\omega(t)})^2 dt \right)^{\frac{1}{2}} \times \left(\int_{\mathbb{R}} (|f(t)| \sqrt{\omega(t)})^2 dt \right)^{\frac{1}{2}} \\ &= \left(\int_{\mathbb{R}} e^{-\frac{t^2}{2}} dt \right)^{\frac{1}{2}} \times \|f\|_{L^2(\omega)} \\ &< \infty \end{aligned}$$

Donc $f\omega \in L^1(\mathbb{R})$, déduit donc que F coïncide avec la transformée de Fourier de $f\omega$: donc $\hat{f}\omega = F = 0$.

L'injectivité de la transformation de Fourier entraîne alors que $f\omega = 0$. Il en résulte que $f = 0$.

4 On rappelle que $\omega^{(n)}(t) = H_n(t)\omega(t)$ où H_n est le n-ième polynôme d'Hermite. Montrons l'orthogonalité des H_n suivant le produit scalaire $(\cdot | \cdot)_{L^2(\omega)}$.

Soit $n, k \geq 1$ deux entiers. On a :

$$\begin{aligned} \int_{\mathbb{R}} t^k \omega^{(n)}(t) dt &= \left[t^k \omega^{(n-1)}(t) \right]_{+\infty}^{-\infty} - k \int_{\mathbb{R}} t^{k-1} \omega^{(n-1)}(t) dt \\ &= -k \int_{\mathbb{R}} t^{k-1} \omega^{(n-1)}(t) dt. \end{aligned}$$

Car $t^k \omega^{(n-1)}(t) = t^k H_{n-1}(t) e^{-\frac{t^2}{2}} \xrightarrow{|t| \rightarrow +\infty} 0$.

Par récurrence on montre que pour $k \leq n$:

$$\int_{\mathbb{R}} t^k \omega^{(n)}(t) dt = (-1)^k k! \int_{\mathbb{R}} \omega^{(n-k)}(t) dt \quad (13)$$

Or pour $k < n$, on a :

$$\int_{\mathbb{R}} \omega^{(n-k)}(t) dt = \left[\omega^{(n-k+1)}(t) \right]_{-\infty}^{+\infty} = 0$$

car $\omega^{(n-k+1)}(t) = H_{n-k+1}(t) e^{-\frac{t^2}{2}} \xrightarrow{|t| \rightarrow +\infty} 0$

Donc :

$$\int_{\mathbb{R}} t^k \omega^{(n)}(t) dt = 0 \quad \text{pour } k < n \quad (14)$$

Comme H_k est une combinaison linéaire de $1, t, t^2, \dots, t^k$, il résulte que

$$\int_{\mathbb{R}} H_k(t) \omega^{(n)}(t) dt = 0 \quad \text{pour } k < n$$

Cela signifie : $\int_{\mathbb{R}} H_k(t) H_n(t) \omega(t) dt = 0$ pour $k < n$

Par ailleurs, pour $k = n$ dans l'équation 13, on a :

$$\begin{aligned} \int_{\mathbb{R}} t^n \omega^{(n)}(t) dt &= (-1)^n n! \int_{\mathbb{R}} \omega(t) dt \\ &= (-1)^n n! \int_{\mathbb{R}} e^{-\frac{t^2}{2}} dt \\ &= (-1)^n n! \sqrt{2\pi} \end{aligned}$$

Puisque les polynômes H_k sont de degré k , ils sont linéairement indépendants. Il en résulte que les polynômes H_0, H_1, \dots, H_n engendrent l'espace des polynômes de degré inférieur ou égal à $n - 1$. Ainsi :

$$H_n(t) = (-1)^n t^n + \sum_{k=0}^{n-1} a_{n,k} H_k(t).$$

Comme on a vu que H_n est orthogonale à H_k pour $k \leq n-1$, on obtient :

$$\begin{aligned}\|H_n\|_{L^2(\omega)}^2 &= (-1)^n \int_{\mathbb{R}} t^n H_n(t) \omega(t) dt + (-1)^n \sum_{k=0}^{n-1} a_{n,k} \int_{\mathbb{R}} H_k(t) H_n(t) \omega(t) dt \\ &= (-1)^n \int_{\mathbb{R}} t^n \omega^{(n)}(t) dt + 0 \\ &= n! \sqrt{2\pi}\end{aligned}$$

□

On vient ainsi de montrer que l'ensemble des polynômes est dense dans le Hilbert $L^2(\omega)$ et que les polynômes d'Hermite sont une famille de polynômes orthogonaux dans cette espace; donc les polynômes d'Hermite normalisés $\left(\frac{H_n}{\|H_n\|_{L^2(\omega)}}\right)_{n \geq 0}$ forment une base Hilbertienne de l'espace $L^2(\omega)$, et on déduit que :

Proposition 1 (Expansion polynômiale d'Hermite)

Pour toute fonction f définie sur \mathbb{R} ,

$$\begin{aligned}f &= \sum_{n=0}^{\infty} \left(f \mid \frac{H_n}{\|H_n\|_{L^2(\omega)}} \right)_{L^2(\omega)} \frac{H_n}{\|H_n\|_{L^2(\omega)}} \\ &= \sum_{n=0}^{\infty} \frac{1}{n!} \int_{\mathbb{R}} f(t) H_n(t) \frac{1}{\sqrt{2\pi}} e^{-\frac{t^2}{2}} dt H_n \\ &= \sum_{n=0}^{\infty} \frac{1}{n!} \int_{\mathbb{R}} f(t) H_n(t) \phi(t) dt H_n\end{aligned}\tag{15}$$

avec $\phi(t) = \frac{1}{\sqrt{2\pi}} e^{-\frac{t^2}{2}}$ la densité de la loi normale centrée réduite.

Simulation d'une trajectoire du MBS

```
> set.seed(123)
> N<-100
> T<-1
> Delta<-T/N
> w<-numeric(N+1)
> t<-seq(0,T,length=N+1)
> for(i in 2:(N+1))
+ w[i]<-w[i-1]+rnorm(1)*sqrt(Delta)
> plot(t,w, type="l",main="processus de wiener",ylim = c(-1,1))
```

FIGURE 4 – Code R d'une trajectoire du MBS

Simulation d'une trajectoire du processus d'Ornstein-Uhlenbeck

```
> sigma<-0.02
> w<-numeric(N+1)
> t<-seq(0,T,length=N+1)
> for (i in 2:(N+1)) {
+   w[i]<-w[i-1]+rnorm(1)*sqrt(Delta)
+ }
> plot(t,w,type="l",main="wiener process",ylim=c(-1,1))
> x<-numeric(N)
> x[1]<-x
> for (i in 2:(N+1)) {
+   w<-rnorm(1)
+   x[i]<-(exp(-beta*Delta))*x[i-1]+sigma*sqrt((1-exp(-2*beta*Delta))/(2*beta))*w
+ }
> plot(t,x,type = "l",main="Une trajectoire du processus d'ornstein-uhlenbeck")
.
```

FIGURE 5 – Code R d'une trajectoire du processus Ornstein-Uhlenbeck

Code R pour la realisant du Bootstrap

```

library(mpoly)
library(polynom)

#Affectation des paramètres
x_0 <- 0.1
N <- 3000
sigma_0 <- 0.02
beta_0 <- 0.01
Delta <- 1/12
theta_0 <- c(sigma_0, beta_0)

#Simulation des données
Donnees <- function(theta, x_0=0.1, Delta=1/12, N=3000){
  sig <- theta[1]
  bet <- theta[2]
  b <- N-1; r <- c()
  r[1] <- x_0
  for (i in 1:b) {
    w <- rnorm(1)
    r[i+1] <- (exp(-bet*Delta))*r[i] + sig*sqrt((1-exp(-2*bet*Delta))/(2*bet))*w
  }
  return(r)
}

Data_0 <- Donnees(theta_0) #les x_i générés

#les fonctions préliminaires
f <- function(x, m=0, v=1){
  (1/(sqrt(2*v*pi)))*exp(-((x-m)^2)/(2*v))
}

h <- function(i){hermite(i)}
H1 <- as.function(h(1))
H2 <- as.function(h(2))
H3 <- as.function(h(3))
H4 <- as.function(h(4))
H6 <- as.function(h(6))

```

```

g_i <- function(x1,x2,sig,bet){
  z <- (x1-x2)/(sig*sqrt(Delta))
  M1 <- f(z)
  M2 <- M1*((bet/2)*H2(z))*Delta
  M3 <- M1*(-(bet^2*x_0)/(2*sig))*H1(z)*Delta^{1.5}
  M4 <- M1*((bet^2/3)*H2(z)) + (-(bet^2/8)*H4(z))*Delta^2
  M5 <- M1*(((bet^3*x_0)/(6*sig))*H1(z)+(((bet^3*x_0)/(4*sig))*H3(z))*Delta^{2.5}
  M6 <- M1*(-((bet^4*x_0^2)/(8*sig^2) -
(bet^3/6)*H2(z)+(bet^3/6)*H4(z)+(bet^3/48)*H6(z))*Delta^3
  result <- (Delta^{-1/2}/sig)*(M1+M2-M3+M4-M5+M6)
}

#Logarithme de la vraisemblance
log_vrais <- function(theta,data=Data_0){
  N<-length(data)
  sig<-theta[1]
  bet<-theta[2]
  t_approx <- c()
  for (i in 1:(N-1)) {
    j <- log(g_i(data[i+1],data[i],sig,bet))
    t_approx <- c(t_approx, j)
  }
  return(-sum(t_approx))
}

#theta_0=(0.02,0.01)
Data_0 <- Donnees(theta_0)

```



```

#estimations des paramètres par bootstrap

simb_boot<-function(B,Data_0,thetainit=c(1,2)){
  Res<-matrix(0,B,2)
  Res[1,<-nlm(log_vrais,thetainit,data=Data_0)$estimate
  for (i in 2:B) {
    traj<-Donnees(Res[i-1,])
    Res[i,<-nlm(log_vrais,thetainit,data=traj)$estimate
  }
  return(Res)
}

Mat=simb_boot(100,Data_0)
Sd=apply(Mat,2,sd)
Biais=apply(Mat,2,mean)-Mat[1,]
MSE=Biais^2+Sd^2

```

