

STATISTIQUES DES PROCESSUS

COURS DE MASTER 2

Lionel Truquet

2011-2012

Table des matières

1	Quelques généralités sur les processus stochastiques	7
1.1	Définition. Notation	7
1.2	Les lois fini-dimensionnelles et le théorème fondamental de Kolmogorov	8
1.3	Versions d'un processus et continuité des trajectoires	9
1.4	Les processus Gaussiens	10
1.5	Temps d'arrêt d'un processus	11
1.6	Estimateur du Maximum de Vraisemblance (EMV)	12
2	Le mouvement brownien	15
2.1	Définition et existence	15
2.2	Construction par marche aléatoire	18
2.3	Statistique sur un mouvement brownien avec drift	21
2.3.1	L'estimation par maximum de vraisemblance	21
2.3.2	Méthode des moindres carrés : le modèle linéaire gaussien	22
2.4	Variation quadratique et variation totale	24
2.5	Exercices	25
3	Processus de Poisson	31
3.1	Le processus de Poisson simple	31
3.1.1	Estimation de l'intensité par maximum de vraisemblance	34
3.2	Le processus de Poisson non homogène	35
3.2.1	Définition et propriétés fondamentales	35
3.2.2	Estimation par maximum de vraisemblance	38
3.2.3	Test de Laplace	40
3.3	Le processus de Poisson composé	42
3.4	Exercices	42
4	Processus de Markov à sauts	47
4.1	Définition générale d'un processus de Markov	47
4.2	Les processus de sauts	48
4.2.1	La propriété de Markov simple	49
4.2.2	La propriété de Markov forte	50
4.3	Description dynamique d'un processus de Markov de saut	51
4.4	Processus minimal et simulation d'un processus de Markov de saut	54
4.5	Générateur et taux de transition	56
4.6	L'estimateur du maximum de vraisemblance	58
4.6.1	Calcul informel de la vraisemblance	58
4.6.2	Définition de la mesure dominante	59
4.7	Quelques exemples	60

4.8	Exercices	61
5	Intégrale stochastique et modèles de diffusion	63
5.1	Martingales à temps continu	63
5.1.1	Définition et exemples	63
5.1.2	Inégalité maximale de Doob. Convergence	64
5.1.3	Le théorème d'arrêt	65
5.2	La notion d'intégrale stochastique	66
5.2.1	Motivation	66
5.3	Construction de l'intégrale stochastique	67
5.3.1	L'intégrale stochastique en tant que processus	71
5.4	La formule d'Itô	72
5.5	Introduction aux équations différentielles stochastiques	75
5.5.1	Existence et unicité des solutions (coefficients Lipschitz)	75
5.6	Exemples	79
5.6.1	Processus d'Ornstein-Uhlenbeck	79
5.6.2	Le mouvement brownien géométrique	80
5.6.3	Un modèle pour étudier la suspension d'un véhicule	80
5.7	Le théorème de Girsanov et le calcul de la vraisemblance d'une diffusion	81
5.7.1	Martingale exponentielle	81
5.7.2	Le théorème de Girsanov	83
5.7.3	Application au calcul de la vraisemblance d'une diffusion	83
5.7.4	L'EMV en pratique	86
5.8	Exercices	87

Bibliographie

- [1] Gallardo, L. (2008) *Mouvement brownien et calcul d'Itô*. Hermann.
- [2] Coccozza-Thivent, C. (1997) *Processus stochastiques et fiabilité des systèmes*. Springer.
- [3] Basawa, I. V., Prakasa Rao, B.L.S (1980) *Statistical inference for stochastic processes*. Academic Press.

Chapitre 1

Quelques généralités sur les processus stochastiques

1.1 Définition. Notation

Définition 1 Soit I un sous-ensemble de \mathbb{R}_+ . On appelle processus stochastique (resp. processus stochastique adapté) la donnée de $(\Omega, \mathcal{F}, (X_t)_{t \in I}, \mathbb{P})$ (resp. $(\Omega, \mathcal{F}, (\mathcal{F}_t)_{t \in I}, (X_t)_{t \in I}, \mathbb{P})$ où

- $(\Omega, \mathcal{F}, \mathbb{P})$ est un espace probabilisé.
- $(\mathcal{F}_t)_{t \in I}$ est une famille croissante de sous-tribus de \mathcal{F} (on parle de filtration).
- Pour tout $t \in I$, $X_t : (\Omega, \mathcal{F}, \mathbb{P}) \rightarrow (E, \mathcal{B}(E))$ est une variable aléatoire (resp. X_t est variable aléatoire \mathcal{F}_t -mesurable et on dira que $(X_t)_t$ est (\mathcal{F}_t) -adapté).

Remarques

1. Un processus stochastique modélise l'état d'un système aléatoire au cours du temps. E est appelé espace d'état. Dans ce cours on se restreindra à $E = \mathbb{R}^d$, $I = [0, T]$ ou $I = [0, +\infty[$. Lorsque $I = \mathbb{N}$, on parle de processus à temps discret. S'il n'y pas d'ambiguïté, on notera un processus stochastique simplement par $(X_t)_{t \in I}$.
2. Si pour tout $t \in I$, une variable aléatoire X_t est donnée, l'exemple le plus simple de filtration est la filtration naturelle

$$\mathcal{F}_t = \sigma(X_s / s \leq t), \quad t \in I.$$

La filtration naturelle est la plus petite filtration par rapport à laquelle $(X_t)_{t \in I}$ est adapté.

3. Si $\omega \in \Omega$, l'application $t \mapsto X_t(\omega)$ est appelée trajectoire. Une trajectoire est donc un élément de E^I .

Exemple. Le processus de Poisson $(N_t)_{t \in \mathbb{R}_+}$ peut être défini à partir d'une suite i.i.d $(\xi_n)_{n \in \mathbb{N}^*}$ de variables aléatoires toutes de loi exponentielle de paramètre $\lambda > 0$ (la suite étant définie sur un espace probabilisé $(\Omega, \mathcal{F}, \mathbb{P})$). On pose $S_0 = 0$ et $S_n = \xi_1 + \dots + \xi_n$ si $n \geq 1$. Le processus de Poisson est alors défini par :

$$N_t = \sum_{n=1}^{+\infty} \mathbf{1}_{S_n \leq t}.$$

Les trajectoires de N sont constantes par morceaux et ne croissent que par sauts d'amplitude 1. On peut alors choisir la filtration naturelle du processus. Ce processus intervient naturellement dans les problèmes de comptage (suite d'arrivées à un guichet, une suite de pannes...). Les variables ξ_n représentent le délai entre deux arrivées ou observations. On verra plus en détail ce processus au Chapitre 3.

1.2 Les lois fini-dimensionnelles et le théorème fondamental de Kolmogorov

Définition 2 Pour un processus stochastique $(X_t)_{t \in I}$, les lois fini-dimensionnelles sont les lois des vecteurs du type $(X_{t_1}, \dots, X_{t_n})$ où $t_1 < t_2 < \dots < t_n$.

Proposition 1 Si $(t_1, \dots, t_n) \in I^n$ vérifie $t_1 < \dots < t_n$, soit μ_{t_1, \dots, t_n} la loi de $(X_{t_1}, \dots, X_{t_n})$. Alors la famille des lois fini-dimensionnelles vérifient la propriété dite de cohérence i.e :

$$\mu_{t_1, \dots, t_n}(A_1 \times \dots \times A_n) = \mu_{t_1, \dots, t_{i-1}, t_{i+1}, \dots, t_n}(A_1 \times \dots \times A_{i-1} \times A_{i+1} \times \dots \times A_n),$$

pour tout $n \geq 2$, $t_1 < \dots < t_n$ dans I , $i = 1, \dots, n$, $A_i = E$ et $A_j \in \mathcal{B}(E)$ pour $j \neq i$.

On s'intéresse alors au problème réciproque. Étant donnée une famille $\{\mu_{t_1, \dots, t_n} / (t_1, \dots, t_n) \in I, t_1 < \dots < t_n\}$ de mesures de probabilités telles que pour $t_1 < \dots < t_n$, μ_{t_1, \dots, t_n} soit une mesure de probabilité sur $(E^n, \mathcal{B}(E^n))$, peut-on construire un processus stochastique qui possède cette famille pour lois fini-dimensionnelles? Lorsque $E = \mathbb{R}^d$, nous allons voir que oui. Introduisons d'abord la tribu cylindrique sur E^I . On appelle cylindre tout sous-ensemble C de E^I de la forme

$$C = \{x \in E^I / x(t_1) \in A_1, \dots, x(t_n) \in A_n\},$$

pour $n \in \mathbb{N}^*$, des éléments t_1, \dots, t_n de I , et des boréliens A_1, \dots, A_n . La tribu cylindrique est la tribu $\sigma(\mathcal{C})$ où \mathcal{C} est l'ensemble des cylindres. On la note également $\mathcal{B}(E)^{\otimes I}$. On peut remarquer que si pour tout $t \in I$, p_t désigne l'application définie sur E^I par $p_t(x) = x_t$ pour tout $x \in E^I$ (projection sur la t -ième coordonnée), alors

$$\mathcal{B}(E)^{\otimes I} = \sigma(p_t / t \in I).$$

La tribu cylindrique est donc la plus petite tribu qui rend ces projections mesurables.

Théorème 1 (Théorème de Kolmogorov) On note $S_I = \cup_{n \geq 1} \{(t_1, \dots, t_n) \in I^n / t_1 < \dots < t_n\}$. Soit $(\mu_J)_{J \in S_I}$ une famille cohérente de mesures de probabilités sur les produits cartésiens de E . Alors il existe un processus adapté $(\Omega, \mathcal{F}, (\mathcal{F}_t)_{t \in I}, (X_t)_{t \in I}, \mathbb{P})$ tel que

$$\mathbb{P}_{X_J} = \mu_J, \quad \forall J \in S_I.$$

Il est construit tel que :

1. $\Omega = E^I$.
2. $X_t : \omega \mapsto X_t(\omega) = \omega(t)$ (t -ième coordonnée).
3. $(\mathcal{F}_t)_{t \in I}$ est la filtration naturelle associée à $(X_t)_{t \in I}$.
4. $\mathcal{F} = \mathcal{B}(E)^{\otimes I}$.

Le processus X s'appelle le processus canonique associé à $(\mu_J)_{J \in S_I}$. De plus la mesure de probabilité \mathbb{P} est unique sur cet espace.

Remarque. Le théorème précédent simplifie énormément la construction d'un processus. Il suffit de disposer d'une famille cohérente de probabilités (les lois fini-dimensionnelles). Cependant, il s'agit juste d'une première étape dans la construction. En effet, on cherchera plutôt à avoir de la régularité sur les trajectoires du processus (continuité ou continuité à droite) ce qui n'est pas du tout garanti par le théorème précédent. D'autres problèmes peuvent également ce produire : rien ne garantit par exemple la mesurabilité de l'application T défini par $T(\omega) = \inf \{t : X_t(\omega) > 0\}$.

1.3 Versions d'un processus et continuité des trajectoires

Définition 3 On dit qu'un processus stochastique Y est une version de X (on dit aussi modification) si

$$\mathbb{P}(X_t = Y_t), \quad t \in I.$$

Proposition 2 Si Y est une version de X , alors X et Y ont les mêmes lois fini-dimensionnelles.

La preuve est laissée en exercice.

Définition 4 On dit que deux processus X et Y sont indistinguables si $\mathbb{P}(\forall t \in I, X_t = Y_t) = 1$.

Remarque. Si X et Y sont indistinguables, alors X et Y sont versions l'un de l'autre. La réciproque est fautive en général. Par exemple, sur un espace probabilisé $(\Omega, \mathcal{F}, \mathbb{P})$, considérons une variable aléatoire U de loi uniforme sur $[0, 1]$. Alors si $X_t = 0$ pour tout $t \in [0, 1]$ et $Y_t = \mathbb{1}_{U=t}$, on peut vérifier que X et Y sont versions l'un de l'autre mais pas indistinguables.

Exercice. Montrer que lorsque $I = \mathbb{N}$, ces deux notions coïncident.

Définition 5 – On dit que X est un processus continu si p.s, $t \mapsto X_t(\omega)$ est une fonction continue sur I .

– On dit que X est càd-làg si p.s, $t \mapsto X_t(\omega)$ est une fonction continue à droite avec limite à gauche.

Proposition 3 Deux versions càd-làg d'un même processus sont indistinguables.

Preuve

$$\begin{aligned} \mathbb{P}(\{\forall t \in I, X_t = Y_t\}^c) &= \mathbb{P}(\exists t \in I / X_t \neq Y_t) \\ &= \mathbb{P}(\exists t \in \mathbb{Q} / X_t \neq Y_t) \\ &\leq \sum_{t \in \mathbb{Q}} \mathbb{P}(X_t \neq Y_t) \\ &= 0. \square \end{aligned}$$

Le théorème suivant donne un critère pour montrer l'existence d'un version continue.

Théorème 2 (théorème de continuité des trajectoires de Kolmogorov) Soit $I = [0, T]$ ou $I = \mathbb{R}_+$ et $(X_t)_{t \in I}$ un processus stochastique à valeurs dans \mathbb{R}^d . Supposons l'existence de constantes $\alpha, \beta, C > 0$ telles que :

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

Alors, il existe une version \tilde{X} de X qui est à trajectoires continues (p.s). De plus, sur tout intervalle $[0, u]$ de I , les trajectoires sont Hölderiennes d'exposant γ pour tout $\gamma < \frac{\alpha}{\beta}$, c'est à dire que pour presque tout ω , il existe $K_\omega \geq 0$ tel que

$$|\tilde{X}_t(\omega) - \tilde{X}_s(\omega)| \leq K_\omega |t - s|, \quad 0 \leq s, t \leq u.$$

Remarque. L'hypothèse du théorème garantit seulement que $\lim_{h \rightarrow 0} X_{t+h} = X_t$ en probabilité.

Idée de la démonstration. On peut se limiter à $I = [0, 1]$. On utilise l'ensemble des dyadiques $D = \cup_{m \geq 0} D_m$ où

$$D_m = \left\{ \frac{k}{2^m} / k = 0, \dots, 2^m \right\}.$$

On considère alors

$$M_\gamma = \sup_{s, t \in D, s \neq t} \frac{|X_t - X_s|}{|t - s|^\gamma}.$$

On montre en utilisant l'hypothèse que $\mathbb{E} \left(M_\gamma^\beta \right) < +\infty$ (c'est la partie technique). Ceci permet de déduire que $M_\gamma < +\infty$ p.s. On utilise alors le théorème de prolongement des applications uniformément continues sur une partie dense (ici D est dense dans $[0, 1]$) ce qui permet de définir un processus \tilde{X} qui est à trajectoires continue p.s et qui coïncident avec X sur D . Ce processus \tilde{X} est la version de X satisfaisant les propriétés annoncées. \square

1.4 Les processus Gaussiens

Définition 6 *Un processus aléatoire à valeurs dans \mathbb{R}^d est dit Gaussien si toutes ses lois finidimensionnelles sont Gaussiennes.*

Deux fonctions vont servir à décrire la loi du processus : $m : I \rightarrow \mathbb{R}$ définie par $m(t) = \mathbb{E}(X_t)$ pour tout $t \in I$ est appelée la moyenne du processus X et la fonction $\Gamma : I \times I \rightarrow \mathbb{R}$ définie par $\Gamma(s, t) = \text{Cov}(X_s, X_t)$ est appelée la covariance du processus X .

On peut remarquer que la covariance d'un processus vérifie la propriété de positivité suivante : pour tout $J = \{t_1, \dots, t_n\} \in S_I$ et tout n -uplet $(a_1, \dots, a_n) \in \mathbb{R}^n$, on a

$$\sum_{i, j=1}^n a_i a_j \Gamma(t_i, t_j) \geq 0.$$

En effet, il suffit d'observer que

$$\sum_{i, j=1}^n a_i a_j \Gamma(t_i, t_j) = \mathbb{E} \left(\left| \sum_{i=1}^n a_i X_{t_i} \right|^2 \right).$$

Théorème 3 *Soit $m : I \rightarrow \mathbb{R}$ et $\Gamma : I \times I \rightarrow \mathbb{R}$ deux fonctions telles que Γ satisfasse la propriété de positivité ci-dessus. Alors il existe un processus Gaussien X tel que pour tout $J = \{t_1, \dots, t_n\} \in S_I$:*

$$(X_{t_1}, \dots, X_{t_n}) \sim \mathcal{N}_n(m_J, \Gamma_J),$$

où $m_J = (m(t_1), \dots, m(t_n))$ et $\Gamma_J = [\Gamma(t_i, t_j)]_{1 \leq i, j \leq n}$.

Preuve Il est évident que la famille des lois normales de moyenne m_J et de covariance Γ_J pour $J \in S_I$ constitue une famille cohérente de probabilités. Le théorème de Kolmogorov assure le résultat. \square

Exemples

- Lorsque $I = \mathbb{N}$, l'exemple le plus simple est l'exemple d'une suite ξ i.i.d de variables aléatoires Gaussiennes centrées réduites.
- A partir de l'exemple précédent, on peut considérer une marche aléatoire (indexée par \mathbb{N}). On définit $X_0 = 0$ puis $X_{n+1} = X_n + \xi_{n+1}$ pour $n \geq 0$. La moyenne de ce processus est nulle et la covariance Γ est définie par $\Gamma(i, j) = \min(i, j)$ pour $i, j \in \mathbb{N}$. C'est un mouvement Brownien à temps discret.
- On verra au prochain chapitre un processus Gaussien indexé par $I = \mathbb{R}_+$ et dont la fonction de covariance est donnée par $\Gamma(s, t) = \min(s, t)$ pour $s, t \geq 0$. Il s'agit du mouvement Brownien.
- On peut construire lorsque $I = \mathbb{R}_+$ une famille i.i.d $\{X_t/t \in I\}$ de variables aléatoires toutes de loi $\mathcal{N}(0, 1)$. Par contre les trajectoires sont très irrégulières : l'application $(\omega, t) \mapsto X_t(\omega)$ n'est même pas mesurable par rapport à la tribu produit $\mathcal{F} \otimes \mathcal{B}(I)$.

Exemple d'utilisation. En théorie du signal, on cherche à localiser une cible (bâtiment, avion). Un émetteur couplé à un récepteur émet un signal $(s(t))_{t \geq 0}$ déterministe qui est réfléchi par la cible. Le récepteur reçoit un signal atténué, retardé et bruité

$$Y_t = as(t - t_0) + B_t.$$

On cherche à estimer $t_0 = \frac{2d}{c}$ le temps d'aller-retour. Le bruit B est supposé Gaussien (pas forcément i.i.d).

1.5 Temps d'arrêt d'un processus

Définition 7 Soit $(\mathcal{F}_t)_{t \in I}$ une filtration et $\mathcal{F}_\infty = \sigma(\cup_{t \in I} \mathcal{F}_t)$ sa tribu terminale. Une variable aléatoire $\tau : \Omega \rightarrow I \cup \{+\infty\}$ est appelée (\mathcal{F}_t) -temps d'arrêt si pour tout $t \in I$, $\{\tau \leq t\} \in \mathcal{F}_t$. On pose alors

$$\mathcal{F}_\tau = \{A \in \mathcal{F}_\infty / \forall t \in I, A \cap \{\tau \leq t\} \in \mathcal{F}_t\}.$$

\mathcal{F}_τ est une tribu appelée tribu des événements antérieurs à τ .

Remarque. Si $\Omega = E^I$ et pour $(\omega, t) \in \Omega \times I$, $X_t(\omega) = \omega(t)$, $\mathcal{F}_t = \sigma(X_s/s \leq t)$ alors il est possible de montrer que

$$\mathcal{F}_\tau = \sigma(X_{t \wedge \tau}/t \in I).$$

Proposition 4 Soient σ et τ deux (\mathcal{F}_t) -temps d'arrêt. Alors

1. τ est \mathcal{F}_τ -mesurable.
2. $\min(\sigma, \tau)$, $\max(\sigma, \tau)$ et $\sigma + \tau$ sont des (\mathcal{F}_t) -temps d'arrêt.
3. Si $\sigma \leq \tau$, alors $\mathcal{F}_\sigma \subset \mathcal{F}_\tau$.

Preuve.

1. Il suffit de montrer que $\{\tau \leq s\} \in \mathcal{F}_\tau, \forall s \in I$. Si $t \in I$, on a

$$\{\tau \leq s\} \cap \{\tau \leq t\} = \{\tau \leq \min(s, t)\} \in \mathcal{F}_{\min(s, t)} \subset \mathcal{F}_t.$$

2. Soit $t \in I$. Les trois variables aléatoires sont des temps d'arrêt à cause des décompositions suivantes :

$$\{\min(\sigma, \tau) \leq t\} = \{\sigma \leq t\} \cap \{\tau \leq t\} \in \mathcal{F}_t,$$

$$\{\max(\sigma, \tau) > t\} = \{\sigma > t\} \cap \{\tau > t\} \in \mathcal{F}_t,$$

$$\{\sigma + \tau > t\} = \cup_{0 \leq r \leq t, r \in \mathbb{Q}} (\{\sigma > r\} \cap \{\tau > t - r\}) \cup (\{\sigma = 0\} \cap \{\tau > t\}) \in \mathcal{F}_t. \square$$

3. Si $\sigma \leq \tau$ et $A \in \mathcal{F}_\sigma$ alors

$$A \cap \{\tau \leq t\} = A \cap \{\sigma \leq t\} \cap \{\tau \leq t\} \in \mathcal{F}_t.$$

Donc $A \in \mathcal{F}_\tau. \square$

Définition 8 *Le processus X est dit mesurable si l'application $(t, \omega) \rightarrow X_t(\omega)$ est mesurable en tant qu'application de $(I \times \Omega, \mathcal{B}(I) \otimes \mathcal{F})$ vers \mathbb{R}^d .*

Proposition 5 1. *Si X est un processus mesurable et σ est une variable aléatoire à valeurs dans I alors*

$$X_\sigma : \omega \rightarrow X_{\sigma(\omega)}(\omega)$$

est une variable aléatoire.

2. *Si X est un processus adapté continu à droite et si τ est un (\mathcal{F}_t) -temps d'arrêt alors X_τ est \mathcal{F}_τ mesurable.*

Exemple. Si A est une partie de \mathbb{R}^d et X un processus continu, considérons

$$\tau_A = \inf\{t > 0 / X_t \in A\}.$$

On peut montrer que si A est fermé alors τ_A est un (\mathcal{F}_t) -temps d'arrêt. En revanche si A est ouvert, on peut montrer que τ_A est seulement un (\mathcal{F}_t^+) -temps d'arrêt avec

$$\mathcal{F}_t^+ = \cap_{\epsilon > 0} \mathcal{F}_{t+\epsilon}.$$

On pourra faire un dessin. Ceci amène alors à considérer plutôt la filtration élargie (\mathcal{F}_t^+) .

1.6 Estimateur du Maximum de Vraisemblance (EMV)

Soit (G, \mathcal{G}) un espace mesurable quelconque et $X : (\Omega, \mathcal{A}, \mathbb{P}) \rightarrow (G, \mathcal{G})$ une variable aléatoire. Supposons que \mathbb{P}_X appartienne à une famille de mesures de probabilité $\mathcal{P} = \{P_\theta / \theta \in \Theta\}$. On dira qu'une mesure μ sur (G, \mathcal{G}) domine la famille \mathcal{P} si pour tout $A \in \mathcal{G}$:

$$\mu(A) = 0 \quad \Rightarrow \quad \forall \theta \in \Theta, \quad P_\theta(A) = 0.$$

Le théorème de Radon-Nikodym assure l'existence d'une fonction $x \mapsto p(x, \theta)$ définie sur G et telle que pour tout $\theta \in \Theta$:

$$P_\theta(A) = \int_A p(x, \theta) \mu(dx), \quad A \in \mathcal{G}.$$

La mesure μ dominante n'est pas unique (car $\mu + \nu$ est encore une mesure dominante pour n'importe quelle autre mesure ν sur (G, \mathcal{G})). Cependant si μ_1 et μ_2 sont deux mesures dominantes, alors les densités notées $p_1(\cdot, \theta)$ et $p_2(\cdot, \theta)$ respectivement peuvent différer uniquement d'un facteur indépendant de θ . En effet si

$$d\mu_1 = f_1 d(\mu_1 + \mu_2), \quad d\mu_2 = f_2 d(\mu_1 + \mu_2)$$

(en utilisant toujours le théorème de Radon-Nikodym), alors si $A \in \mathcal{G}$:

$$P_\theta(A) = \int_A p_1(x, \theta) \mu_1(dx) = \int_A p_1(x, \theta) f_1(x) (\mu_1(dx) + \mu_2(dx))$$

et

$$P_\theta(A) = \int_A p_2(x, \theta) \mu_2(dx) = \int_A p_2(x, \theta) f_2(x) (\mu_1(dx) + \mu_2(dx)).$$

Pour une mesure dominante μ donnée, l'estimateur du maximum de vraisemblance du paramètre θ_0 (où $\mathbb{P}_X = P_{\theta_0}$) est défini par

$$\hat{\theta}(X) = \arg \max_{\theta \in \Theta} p(X, \theta).$$

La définition est valide si le maximum est bien atteint en un point. On supposera (ou on essaiera de supposer) que pour deux valeurs $\theta_1 \neq \theta_2$ alors $P_{\theta_1} \neq P_{\theta_2}$ (hypothèse d'identifiabilité).

L'exemple de n observations indépendantes. Lorsque on a un n -uplet de variables aléatoires réelles i.i.d $X = (X_1, \dots, X_n)$ et que la densité de X_1 appartient à une famille $\{f_\theta/\theta \in \Theta \subset \mathbb{R}^d\}$, on peut choisir μ comme la mesure de Lebesgue sur \mathbb{R}^n et on aura :

$$\hat{\theta}(X_1, \dots, X_n) = \arg \max_{\theta \in \Theta} \prod_{i=1}^n f_\theta(X_i).$$

Un exemple d'observations dépendantes. Il s'agit d'un exemple d'un chaîne de Markov à temps discret. A l'aide d'une suite i.i.d $(\xi_j)_{j \geq 1}$ de variables aléatoires $\mathcal{N}(0, 1)$, on définit $X_0 = x$ pour $x \in \mathbb{R}$ et

$$X_{j+1} = aX_j + \xi_{j+1}, \quad j = 0, \dots, n-1.$$

Le vecteur (X_1, \dots, X_n) est un vecteur Gaussien dont la densité est donnée par

$$f_{X_1, \dots, X_n}(x_1, \dots, x_n) = f_{X_n|X_{n-1}}(x_n|x_{n-1}) \times f_{X_{n-1}|X_{n-2}}(x_{n-1}|x_{n-2}) \times \dots \times f_{X_2|X_1}(x_2|x_1) \times f_{X_1}(x_1).$$

Pour $j = 1, \dots, n-1$, on a $X_{j+1}|X_j = x_j \sim \mathcal{N}(ax_j, 1)$ et donc

$$f_{X_{j+1}|X_j}(x_{j+1}|x_j) = \frac{1}{\sqrt{2\pi}} \exp\left(-\frac{(x_{j+1} - ax_j)^2}{2}\right).$$

L'estimateur du maximum de vraisemblance du paramètre $\theta = a$ est défini par

$$\hat{a}(X_1, \dots, X_n) = \arg \max_{a \in \Theta} \frac{1}{(2\pi)^{n/2}} \exp\left(-\frac{1}{2} \left(\sum_{j=1}^{n-1} (x_{j+1} - ax_j)^2 + (x_1 - ax)^2 \right)\right).$$

EMV et bijection. Supposons qu'il soit difficile de déterminer la loi de X que l'on suppose appartenir à une famille $\{P_\theta/\theta \in \Theta\}$ mais que pour une bijection bi-mesurable $g : (\tilde{G}, \tilde{\mathcal{G}}) \rightarrow (G, \mathcal{G})$, il existe une variable aléatoire

$$Y : (\Omega, \mathcal{F}, \mathbb{P}) \rightarrow (\tilde{G}, \tilde{\mathcal{G}}),$$

telle que $X = g(Y)$. Alors on a $\hat{\theta}(X) = \hat{\theta}(Y)$. En effet, on a pour une fonction h :

$$\begin{aligned}\mathbb{E}(h(Y)) &= \mathbb{E}(h \circ g^{-1}(X)) \\ &= \int h(g^{-1}(x)) p(x, \theta) \mu(dx) \\ &= \int h(y) p(g(y), \theta) \mu_{g^{-1}}(dy).\end{aligned}$$

où $\mu_{g^{-1}}$ est la mesure image : $\mu_{g^{-1}}(A) = \mu(g(A))$ pour $A \in \tilde{\mathcal{G}}$. On a bien

$$\hat{\theta}(X) = \arg \max_{\theta \in \Theta} p(X, \theta) = \arg \max_{\theta \in \Theta} p(g(Y), \theta) = \hat{\theta}(Y).$$

Cette observation sera en particulier utile pour les processus de sauts.

Chapitre 2

Le mouvement brownien

Ce processus stochastique trouve ses origines dans divers domaines, en tant que phénomène naturel et objet mathématique. Ce processus est mis en évidence par le botaniste Brown en 1827, qui observe les mouvements incessants et désordonnés d'une particule en suspension dans un liquide, puis par Bachelier en 1900 qui étudie les fluctuations du cours de la bourse et aussi par Einstein en 1905 qui étudie les impulsions transmises à une particule par les molécules du milieu, soumises à agitation thermique. En 1923, Wiener bâtit un modèle de processus dont les trajectoires sont continues avec vitesse infini en tout point. Depuis le mouvement brownien a beaucoup été étudié et est à la base de nombreuses modélisations.

2.1 Définition et existence

On rappelle qu'une variable aléatoire X est dite indépendante d'une tribu \mathcal{F} lorsque X est indépendante de toute variable aléatoire $Y : \Omega \rightarrow \mathbb{R}^k$ mesurable par rapport à \mathcal{F} .

Définition 9 *Un processus $B = (\Omega, \mathcal{F}, (\mathcal{F}_t)_{t \geq 0}, (B_t)_{t \geq 0}, \mathbb{P})$ à valeurs réelles est appelé mouvement Brownien si :*

1. $B_0 = 0$ p.s.
2. Si $0 \leq s \leq t$, alors la variable aléatoire $B_t - B_s$ est indépendante de \mathcal{F}_s et suit une loi normale $\mathcal{N}(0, t - s)$.
3. B est un processus continu.

La proposition suivante est une conséquence immédiate de la définition précédente.

Proposition 6 *Tout mouvement brownien est un processus à accroissements indépendants et stationnaires (PAIS) : pour $t_0 \leq t_1 \leq \dots \leq t_n$, les variables aléatoires $B_{t_k} - B_{t_{k-1}}$, $k = 1, \dots, n$ sont indépendantes et leur loi respective ne dépend que de la différence $t_k - t_{k-1}$, $k = 1, \dots, n$.*

Théorème 4 *Soit B un mouvement brownien. Alors B est un processus Gaussien centré (i.e de moyenne nulle) et dont la fonction de covariance est donnée par*

$$\text{Cov}(B_s, B_t) = s \wedge t, \quad s, t \in \mathbb{R}_+. \quad (2.1)$$

Inversement, tout processus Gaussien, continu, centré et dont la fonction de covariance est donnée par (2.1) est un mouvement brownien en considérant la filtration naturelle du processus (i.e $\mathcal{F}_t = \sigma(X_s/0 \leq s \leq t)$, $t \geq 0$).

Preuve.

- Si B est un mouvement brownien alors pour $t_1 \leq t_2 \leq \dots \leq t_n$, le vecteur des accroissements

$$(B_{t_1}, B_{t_2} - B_{t_1}, \dots, B_{t_n} - B_{t_{n-1}})$$

est un vecteur gaussien centré car il est formé de variables aléatoires gaussiennes indépendantes. Ainsi le vecteur $(B_{t_1}, \dots, B_{t_n})$ est aussi un vecteur gaussien centré, car image du vecteur précédent par l'application linéaire $f : \mathbb{R}^n \rightarrow \mathbb{R}^n$ définie par

$$f(x) = (x_1, x_1 + x_2, \dots, x_1 + x_2 + \dots + x_n), \quad x \in \mathbb{R}^n.$$

Donc B est un processus gaussien. L'indépendance de accroissements entraîne que

$$\text{Cov}(B_t - B_s, B_s) = \mathbb{E}(B_t - B_s)\mathbb{E}(B_s) = 0$$

pour $s \leq t$. On en déduit que $\text{Cov}(B_t, B_s) = \text{Var}(B_s) = s$. Par symétrie de s et de t , la fonction de covariance a bien l'expression annoncée.

- Inversement, soit B un processus gaussien, continu, centré et dont la covariance est donnée par (2.1). Posons

$$\mathcal{F}_t = \sigma(B_s / 0 \leq s \leq t), \quad t \geq 0.$$

Alors on a $\text{Var}(B_0) = 0$ donc $B_0 = 0$ p.s. Soient $t \geq s \geq 0$ et $r_1, \dots, r_k \leq s$. Alors le vecteur $(B_t - B_s, B_{r_1}, \dots, B_{r_k})$ est un vecteur gaussien. De plus

$$\text{Cov}(B_t - B_s, B_{r_i}) = \text{Cov}(B_t, B_{r_i}) - \text{Cov}(B_s, B_{r_i}) = r_i - r_i = 0, \quad 1 \leq i \leq k.$$

Ainsi $B_t - B_s$ est indépendante du vecteur $(B_{r_1}, \dots, B_{r_k})$. Ceci suffit à affirmer que $B_t - B_s$ est indépendante de la tribu \mathcal{F}_s . De plus si $s \leq t$, le couple (B_s, B_t) est gaussien donc la variable aléatoire $B_t - B_s$ suit une loi normale de moyenne 0 et de variance

$$\text{Var}(B_t - B_s) = \text{Var}(B_t) + \text{Var}(B_s) - 2 \text{Cov}(B_s, B_t) = t - s.$$

B vérifie bien la définition d'un mouvement brownien. \square

Nous n'avons pas encore prouvé l'existence du mouvement brownien. On peut la justifier en appliquant les résultats vus au Chapitre 1.

Théorème 5 *Posons $I = \mathbb{R}_+$. Il existe une mesure de probabilité \mathbb{P} sur l'espace probabilisable $(\mathbb{R}^I, \mathcal{B}(\mathbb{R})^{\otimes I})$ tel que le processus des coordonnées $(X_t)_{t \in I}$ soit un mouvement brownien (pour sa filtration naturelle).*

Preuve. D'après le théorème précédent, il suffit de prouver l'existence d'un processus gaussien, continu, centré et dont la covariance est donnée par (2.1). On va utiliser le résultat du Chapitre 1 sur l'existence des processus gaussiens. Il suffit en fait de vérifier que la fonction $\Gamma : I^2 \rightarrow \mathbb{R}$, définie par $\Gamma(s, t) = s \wedge t$ pour $(s, t) \in I^2$, satisfait la condition de positivité. Soit $0 \leq t_1 \leq t_2 \leq \dots \leq t_n$ et $u_1, u_2, \dots, u_n \in \mathbb{R}$. Alors il est possible de montrer (en utilisant une récurrence sur n) que

$$\sum_{i,j=1}^n u_i u_j (t_i \wedge t_j) \geq t_1 (u_1 + u_2 + \dots + u_n)^2,$$

ce qui implique la condition de positivité. Ainsi il existe bien un processus gaussien noté \tilde{B} , centré et dont la covariance est donnée par (2.1).

Nous montrons ensuite qu'il est possible de trouver une version continu de ce processus ce qui

achèvera la preuve du théorème. A cet effet, nous allons utiliser le théorème de continuité de Kolmogorov (cf Chapitre 1). Le processus \tilde{B} vérifie les hypothèses de ce théorème en posant $\beta = 4$, $\alpha = 1$ et $C = 3$. En effet, si $t \geq s$, la variable aléatoire $\tilde{B}_t - \tilde{B}_s$ a même loi que $\sqrt{t-s}X$ ou $X \sim \mathcal{N}(0, 1)$ et on a donc

$$\mathbb{E} \left(\left(\tilde{B}_t - \tilde{B}_s \right)^4 \right) = (t-s)^2 \mathbb{E} (X^4) = 3(t-s)^3.$$

Le théorème est prouvé. \square

La proposition suivante donne des propriétés d'invariance du mouvement brownien. Ces propriétés expriment le fait que modulo le centrage et la multiplication par un scalaire, les changements de temps $t \mapsto \frac{at+b}{ct+d}$ laisse la loi du mouvement brownien invariante.

Proposition 7 *Soit $B = (\Omega, \mathcal{F}, (\mathcal{F}_t)_{t \geq 0}, (B_t)_{t \geq 0}, \mathbb{P})$ un mouvement brownien. Soient s et c deux nombres réels avec $s > 0$ et $c \neq 0$. Alors les processus suivants sont aussi des mouvements browniens.*

1. $B^{(1)} = -B$ (avec la même filtration).
2. $B^{(2)} = B_{\cdot+s} - B_s$ (avec la filtration naturelle associée).
3. $B^{(3)} = cB_{\frac{\cdot}{c^2}}$ (avec la filtration naturelle associée).
4. $B_t^{(4)} = tB_{\frac{1}{t}}$ si $t > 0$ et $B_0^{(4)} = 0$ (avec la filtration naturelle associée).

Preuve. Nous laissons cette preuve en exercice, sauf pour le processus $X = B^{(4)}$ pour lequel nous montrons la continuité en 0, ce qui n'est pas immédiat. On admet que $(X_t)_{t \geq 0}$ et $(B_t)_{t > 0}$ ont les mêmes lois fini-dimensionnelles. De plus, il est clair que les trajectoires de \tilde{B} sont continues sur $]0, +\infty[$. Ainsi, pour tout $N \in \mathbb{N}^*$, on a p.s

$$\sup \left\{ |X_t|/0 < t \leq \frac{1}{N} \right\} = \sup \left\{ |X_t|/0 < t \leq \frac{1}{N}, t \in \mathbb{Q} \right\}.$$

Ainsi

$$\left\{ \lim_{t \rightarrow 0^+} X_t = 0 \right\} = \bigcap_{k \in \mathbb{N}^*} \bigcup_{N \in \mathbb{N}^*} \left\{ \sup \left\{ |X_t|/0 < t \leq \frac{1}{N}, t \in \mathbb{Q} \right\} < \frac{1}{k} \right\}.$$

Les variables aléatoires $Y_N = \sup \{ |X_t|/0 < t \leq \frac{1}{N}, t \in \mathbb{Q} \}$ et $Z_N = \sup \{ |B_t|/0 < t \leq \frac{1}{N}, t \in \mathbb{Q} \}$ ont même loi (le vérifier). Ceci entraîne en utilisant les propriétés de continuité supérieure et inférieure de la mesure \mathbb{P} :

$$\begin{aligned} \mathbb{P} \left(\lim_{t \rightarrow 0^+} X_t = 0 \right) &= \lim_{k \rightarrow \infty} \lim_{N \rightarrow \infty} \mathbb{P} \left(Y_N < \frac{1}{k} \right) \\ &= \lim_{k \rightarrow \infty} \lim_{N \rightarrow \infty} \mathbb{P} \left(Z_N < \frac{1}{k} \right) \\ &= \mathbb{P} \left(\lim_{t \rightarrow 0^+} B_t = 0 \right) \\ &= 1. \square \end{aligned}$$

Proposition 8 (loi des grands nombres) *Un mouvement brownien B vérifie*

$$\lim_{t \rightarrow \infty} \frac{B_t}{t} = 0, \text{ p.s.}$$

Preuve. Il s'agit d'une conséquence immédiate de la preuve précédente car nous avons montré que

$$\lim_{t \rightarrow 0^+} tB_{\frac{1}{t}} = 0, \text{ p.s.} \square$$

Remarque. On peut aussi obtenir cette loi des grands nombres directement en écrivant

$$\frac{B_t}{t} = \frac{1}{t} \left(\sum_{i=1}^{[t]} (B_i - B_{i-1}) + B_t - B_{[t]} \right)$$

et en utilisant la loi des grands nombres pour les variables aléatoires i.i.d. Mais il est nécessaire de contrôler le reste

$$\sup_{t > 0} \frac{|B_t - B_{[t]}|}{t},$$

ce qui pourra se faire en utilisant l'inégalité de Doob pour les martingales (cf Chapitre ?).

La définition du mouvement brownien se généralise à la dimension $d \geq 1$.

Définition 10 *Un processus $B = (B^{(1)}, \dots, B^{(d)})$ à valeurs dans \mathbb{R}^d est un mouvement brownien d -dimensionnel si et seulement si les processus coordonnées $B^{(1)}, \dots, B^{(d)}$ sont des mouvements brownien réels, indépendants entre eux.*

Exercice. On pourra montrer qu'un processus B à valeurs dans \mathbb{R}^d est un mouvement brownien si et seulement si les deux propriétés suivantes sont réalisées :

- $B_0 = 0$ p.s,
- le vecteur $B_t - B_s$ est indépendant de \mathcal{F}_s et de loi $\mathcal{N}_d(0, (t-s)I_d)$ pour $s < t$.

2.2 Construction par marche aléatoire

On peut approcher la distribution du mouvement brownien à l'aide d'une marche aléatoire. Pour simplifier on se restreindra à $I = [0, 1]$. On se donne une suite $(\xi_i)_{i \in \mathbb{N}^*}$ i.i.d de variables aléatoires telles que

$$\mathbb{P}(\xi_1 = -1) = \mathbb{P}(\xi_1 = 1) = \frac{1}{2}.$$

On pose alors

$$Z_n(t) = \frac{1}{\sqrt{n}} \left(\sum_{i=1}^{[nt]} \xi_i + (nt - [nt]) \xi_{[nt]+1} \right), \quad t \in I, \quad (2.2)$$

en adoptant la convention $\sum_{i=1}^0 = 0$. Z_n est donc une variable aléatoire à valeurs dans $F = \mathcal{C}(I, \mathbb{R})$ muni de la tribu cylindrique. On peut en fait montrer que la tribu des cylindres sur F coïncide avec la tribu borélienne $\mathcal{B}(F)$ (tribu engendrée par les ouverts) lorsque F est muni de la norme infini :

$$\mathcal{B}(F) = \sigma(\{B_\infty(f, r) / f \in F, r > 0\}).$$

On a posé

$$B_\infty(f, r) = \left\{ g \in F / \sup_{t \in I} |g(t) - f(t)| < r \right\}.$$

Ainsi \mathbb{P}_{Z_n} (loi de Z_n sous \mathbb{P}) est une mesure de probabilité sur $(F, \mathcal{B}(F))$ au même titre que d'ailleurs que \mathbb{P}_B lorsque B est un mouvement brownien (on parle alors de mesure de Wiener pour désigner la loi du mouvement brownien).

Les Figures 2.2, 2.3 et 2.4 montrent des trajectoires obtenues pour Z_N lorsque $N = 10, 100, 3000$ respectivement. Le théorème de Donsker (1951) assure que \mathbb{P}_{Z_n} converge en loi vers la mesure de Wiener, au sens où

$$\lim_{n \rightarrow \infty} \mathbb{E}(h(Z_n)) = \mathbb{E}(h(B)),$$

pour toute fonction $h : F \rightarrow \mathbb{R}$ continue bornée. En fait ce théorème de convergence vaut aussi pour des suites ξ plus générales.

Théorème 6 *Soit ξ une suite i.i.d de variables aléatoires centrées et de variance 1. Alors la suite de processus $(Z_n)_{n \in \mathbb{N}}$ (où Z_n est défini en (2.2)) converge en loi vers la mesure de Wiener.*

En particulier, on est assuré de la convergence des lois fini-dimensionnelles :

$$\lim_{n \rightarrow \infty} (Z_n(t_1), \dots, Z_n(t_k)) = (B(t_1), \dots, B(t_k)), \text{ en loi,}$$

pour tout $t_1 < t_2 < \dots, t_k$. Il suffit d'appliquer le théorème de Donsker en considérant des fonctions de la forme $f \rightarrow h(f(t_1), \dots, f(t_k))$ où $h : \mathbb{R}^k \rightarrow \mathbb{R}$ est continue bornée. Cependant la convergence des lois fini-dimensionnelles ne suffit pas à montrer la convergence en loi.

Exemple d'application. En posant $h(f) = \sup_{t \in I} |f(t)|$ et $S_i = \xi_1 + \dots + \xi_i$, le théorème de Donsker assure que

$$\lim_{n \rightarrow \infty} \frac{1}{\sqrt{n}} \max_{0 \leq i \leq n} S_i = \sup_{t \in I} |B_t|, \text{ en loi.}$$

On peut chercher la loi limite à partir de celle de la marche aléatoire symétrique en utilisant le principe dit de réflexion. On pose $M_n = \max_{0 \leq i \leq n} S_i$.

Lemme 1 *On a pour tout entier naturel a*

$$\mathbb{P}(M_n \geq a) = 2\mathbb{P}(S_n > a) + \mathbb{P}(S_n = a).$$

Preuve. Lorsque $a = 0$, on a $M_n \geq S_0 = 0$ et $\mathbb{P}(S_n > 0) = \mathbb{P}(S_n < 0)$ par symétrie. On obtient alors

$$\mathbb{P}(M_n \geq a) = 2\mathbb{P}(S_n > a) + \mathbb{P}(S_n = a) = 1.$$

Supposons $a > 0$. On

$$\begin{aligned} \mathbb{P}(M_n \geq a) - \mathbb{P}(S_n = a) &= \mathbb{P}(M_n \geq a, S_n < a) + \mathbb{P}(M_n \geq a, S_n > a) \\ &= \mathbb{P}(M_n \geq a, S_n < a) + \mathbb{P}(S_n > a). \end{aligned}$$

Il suffit donc de prouver que

$$\mathbb{P}(M_n \geq a, S_n < a) = \mathbb{P}(M_n \geq a, S_n > a). \quad (2.3)$$

Pour cela, il suffit de remarquer que chaque trajectoire de la marche a la même probabilité d'apparition ($1/2^n$). De plus si une trajectoire vérifie $S_n > a$, on peut lui faire correspondre une trajectoire qui vérifie $S_n < a$: il suffit pour cela de remplacer S_j par $a - (S_j - a)$ dès que $j \geq k$ où k est le premier indice pour lequel $S_k = a$ (voir Figure 2.1). Parmi les trajectoires qui franchissent le seuil a , on définit ainsi une bijection entre celles qui vérifient $S_n > a$ et celles qui vérifient $S_n < a$. On a donc bien l'égalité (2.3) et le lemme est prouvé. \square

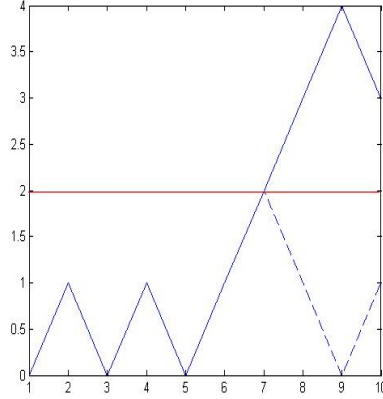


FIGURE 2.1: Principe de réflexion

Proposition 9 La loi de $\sup_{t \in I} B_t$ est égale à celle de $|B_1|$, c'est à dire

$$\mathbb{P}\left(\sup_{t \in I} B_t \leq \alpha\right) = 2 \int_0^\alpha \frac{1}{\sqrt{2\pi}} \exp\left(-\frac{x^2}{2}\right) dx,$$

pour tout $\alpha \geq 0$.

Preuve. On applique le lemme précédent avec $a_n = [\alpha\sqrt{n}] + 1$. On a ainsi les égalités :

$$\mathbb{P}\left(\frac{M_n}{\sqrt{n}} > \alpha\right) = \mathbb{P}(M_n \geq a_n) = 2\mathbb{P}(S_n \geq a_n) + \mathbb{P}(S_n = a_n).$$

Comme $\alpha\sqrt{n} \leq a_n \leq \alpha\sqrt{n} + 1$, on a l'encadrement

$$2\mathbb{P}\left(\frac{S_n}{\sqrt{n}} \geq \alpha + \frac{1}{\sqrt{n}}\right) \leq \mathbb{P}\left(\frac{M_n}{\sqrt{n}} > \alpha\right) \leq 2\mathbb{P}\left(\frac{S_n}{\sqrt{n}} \geq \alpha\right) + \mathbb{P}\left(\alpha \leq \frac{S_n}{\sqrt{n}} \leq \alpha + \frac{1}{\sqrt{n}}\right).$$

En utilisant le théorème central limite et le lemme de Slutsky (voir le Lemme 2), on a

$$\mathbb{P}\left(\alpha \leq \frac{S_n}{\sqrt{n}} \leq \alpha + \frac{1}{\sqrt{n}}\right) = \mathbb{P}\left(\frac{S_n}{\sqrt{n}} - \frac{1}{\sqrt{n}} \leq \alpha\right) - \mathbb{P}\left(\frac{S_n}{\sqrt{n}} < \alpha\right) \xrightarrow{n \rightarrow \infty} \Phi(\alpha) - \Phi(\alpha) = 0,$$

où Φ désigne la fonction de répartition de la gaussienne centrée réduite. En réitérant ce type d'argument, on voit que les deux bornes de l'encadrement précédent converge vers la même limite. De plus

$$\lim_{n \rightarrow \infty} \mathbb{P}\left(\frac{M_n}{\sqrt{n}} > \alpha\right) = 2(1 - \Phi(\alpha)),$$

ce qui prouve la proposition. \square

Remarque. On en déduit également la loi de la variable aléatoire $\sup_{0 \leq s \leq t} B_s$ pour tout $t \in \mathbb{R}_+$. En effet en utilisant la propriété d'invariance du mouvement brownien (cf Proposition 7), nous avons les égalités en loi suivantes

$$\sup_{0 \leq s \leq t} B_s = \sqrt{t} \sup_{0 \leq s \leq t} B_{\frac{s}{t}} = \sqrt{t} \sup_{s \in I} B_s = \sqrt{t} |B_1| = |B_t|.$$

Application à la détection de rupture. Le théorème de Donsker permet de considérer des tests pour détecter des phénomènes de rupture de la loi de n variables aléatoires indépendantes. Par exemple on peut chercher à savoir si la moyenne de l'échantillon (X_1, \dots, X_n) change au cours du temps. Soient $m \in \mathbb{R}$ et $\sigma^2 \in \mathbb{R}_+^*$. Ghorbanzadeh et Picard [2] proposent de tester les deux hypothèses suivantes.

H_0 : Les variables aléatoires X_1, \dots, X_n sont i.i.d de loi F_1 , de moyenne m et de variance σ^2 .

H_1 : Il existe $\tau \in]0, 1[$ tel que si $k \in \{1, \dots, [n\tau]\}$, X_1, \dots, X_k sont i.i.d de loi F_1 et $X_{[n\tau]+1}, \dots, X_n$ soient i.i.d de loi F_2 toutes de moyenne $m' > m$ et de variance σ^2 .

La statistique de test utilisée est $\lambda_n^* = \sup_{t \in [0,1]} Z_n(t) = \frac{1}{\sqrt{n}} \max_{0 \leq i \leq n} S_i$ où on pose ici $\xi_i = \frac{X_i - m}{\sigma}$, $1 \leq i \leq n$. L'hypothèse H_0 sera acceptée si λ_n^* n'est pas trop élevé. Sous H_0 , λ_n^* converge en loi vers $|B_1|$. La zone de rejet (de H_0) pour ce test est alors :

$$\{\lambda_n^* \geq z_\alpha\},$$

où $\mathbb{P}(|B_1| \geq z_\alpha) = \alpha$. z_α est donc le quantile d'ordre $1 - \alpha/2$ de la gaussienne centrée réduite. Lorsque $F_2 = F_{2,n}$ est une légère modification de F_1 , Ghorbanzadeh et Picard appliquent cette procédure pour détecter une rupture du taux de lymphocytes T4 chez des patients HIV-positifs. Ils étudient également la puissance de ce test (c'est à dire la fonction $(\tau, F_1, F_2) \rightarrow \mathbb{P}_{\tau, F_1, F_2}(\lambda_n^* \geq z_\alpha)$).

Le test présenté ici est unilatéral (on suppose que la moyenne ne peut qu'augmenter). On peut proposer un test bilatéral (on suppose alors $m' \neq m$ dans H_1), en remplaçant λ_n^* par $\sup_{t \in [0,1]} |Z_n(t)|$ et en utilisant la loi limite de $\sup_{t \in I} |B_t|$ qui est également une loi dont on connaît l'expression.

Lorsque la moyenne m est inconnue, un test de rupture est parfois basé sur la statistique de test

$$\beta_n^* = \sup_{t \in I} |\tilde{Z}_n(t)/q(t)|,$$

où $\xi_i = \frac{X_i}{\sigma}$ pour $1 \leq i \leq n$, $\tilde{Z}_n(t) = Z_n(t) - tZ_n(1)$ et $q(t)$ est un poids qui permet d'obtenir une bonne puissance si il est correctement choisi (le livre de Csörgö & Horváth [1] étudient les asymptotiques pour ce type de test). Lorsque $q(t) = 1$, on peut voir que sous H_0 , β_n^* converge vers $\sup_{t \in I} |B_t - tB_1|$ à l'aide du théorème de Donsker. Le processus $W_t = B_t - tB_1$, $t \in I$ est appelé pont brownien (il s'agit d'un processus gaussien, centré, continu et de covariance donnée par $\Gamma(s, t) = \min(s, t) - st$ pour $0 \leq s, t \leq 1$). Cette loi limite intervient également dans le test de Kolmogorov-Smirnov pour tester l'adéquation de la loi d'un échantillon à une distribution donnée.

2.3 Statistique sur un mouvement brownien avec drift

Soit X le processus défini par

$$X_t = \mu t + \sigma B_t, \quad t \in \mathbb{R}_+,$$

où B est un mouvement brownien. On suppose que X est observé aux temps $t_1 < t_2 < \dots < t_n$.

2.3.1 L'estimation par maximum de vraisemblance

X est un processus gaussien et nous allons déterminer l'estimateur du maximum de vraisemblance du couple (μ, σ) . Pour cela on pose $Y_1 = \frac{X_{t_1}}{\sqrt{t_1}}$ et

$$Y_j = \frac{X_{t_{j+1}} - X_{t_j}}{\sqrt{t_{j+1} - t_j}}, \quad j = 1, \dots, n-1.$$

Les variables Y_1, \dots, Y_n sont indépendantes et nous pouvons calculer le maximum de vraisemblance à partir de cet échantillon (qui est l'image de l'échantillon initial par une fonction bijective). La loi du vecteur $Y = (Y_1, \dots, Y_n)$ a une densité par rapport à la mesure de Lebesgue sur \mathbb{R}^n et qui est donnée par

$$f_Y(y) = \prod_{i=1}^n f_{Y_i}(y_i), \quad y \in \mathbb{R}^n.$$

Posons $s_1 = t_1$ et $s_j = t_j - t_{j-1}$ si $j = 2, \dots, n$. Comme $Y_i \sim \mathcal{N}(\mu\sqrt{s_i}, \sigma^2)$ pour $1 \leq i \leq n$, on obtient une fonction de vraisemblance L définie par

$$L(\mu, \sigma) = \frac{1}{(2\pi\sigma^2)^{n/2}} \exp\left(-\frac{1}{2\sigma^2} \sum_{i=1}^n (Y_i - \mu\sqrt{s_i})^2\right),$$

pour $(\mu, \sigma) \in \mathbb{R} \times \mathbb{R}_+^*$. La log-vraisemblance ℓ est alors définie par

$$\ell(\mu, \sigma) = \ln L(\mu, \sigma) = -\frac{n}{2} \ln(2\pi\sigma^2) - \sum_{i=1}^n \frac{(Y_i - \mu\sqrt{s_i})^2}{2\sigma^2}$$

Les dérivées partielles sont alors

$$\frac{\partial \ell}{\partial \mu}(\mu, \sigma^2) = -\frac{1}{\sigma^2} \sum_{i=1}^n \sqrt{s_i} (Y_i - \mu\sqrt{s_i}),$$

$$\frac{\partial \ell}{\partial \sigma^2}(\mu, \sigma^2) = -\frac{n}{2\sigma^2} + \frac{1}{2\sigma^4} \sum_{i=1}^n (Y_i - \mu\sqrt{s_i})^2,$$

ce qui donne pour estimateurs

$$\hat{\mu} = \frac{X_{t_n}}{t_n}, \quad \hat{\sigma}^2 = \frac{1}{n} \sum_{i=1}^n (Y_i - \hat{\mu}\sqrt{s_i})^2.$$

Au niveau de la convergence, lorsque $t_n \rightarrow \infty$, la loi des grands nombres pour le mouvement brownien assure la convergence presque sûre de $\hat{\mu}$ vers μ car

$$\hat{\mu} = \mu + \sigma \frac{B_{t_n}}{t_n}.$$

De plus $\hat{\mu} \sim \mathcal{N}\left(\mu, \frac{\sigma^2}{t_n}\right)$. On dispose également de la loi de $\hat{\sigma}^2$ mais nous allons d'abord comparer ces estimateurs à ceux obtenus par moindres carrés dans le modèle linéaire gaussien (ce qui permettra de déduire facilement la loi de $\hat{\sigma}^2$ en appliquant le théorème de Cochran).

2.3.2 Méthode des moindres carrés : le modèle linéaire gaussien

Posons $\varepsilon_1 = \frac{B_{t_1}}{\sqrt{s_1}}$ et pour $i = 2, \dots, n$, $\varepsilon_i = \frac{B_{t_i} - B_{t_{i-1}}}{\sqrt{s_i}}$. Le vecteur $Y = (Y_1, \dots, Y_n)'$ admet la décomposition

$$Y = H\mu + \sigma\varepsilon,$$

où $H = (\sqrt{s_1}, \dots, \sqrt{s_n})'$. L'estimateur des moindres carrés $\hat{\mu}_C$ de μ est défini par

$$\hat{\mu}_C = \arg \min_{\mu \in \mathbb{R}} \|Y - H\mu\|,$$

et est égal à

$$\hat{\mu}_c = (H'H)^{-1} H'Y = \frac{X_{t_n}}{t_n},$$

Remarquons que $H\hat{\mu}_c$ s'interprète comme la projection orthogonale de Y de \mathbb{R}^n sur le sous-espace vectoriel engendré par le vecteur H . Notons $Q = t_n^{-1}HH'$ la matrice de projection associée. On retrouve l'estimateur du maximum de vraisemblance. Pour σ^2 , l'estimateur des moindres carrés est défini par

$$\hat{\sigma}_c^2 = \frac{1}{n-1} \|Y - H\hat{\mu}_c\|^2 = \frac{n}{n-1} \hat{\sigma}^2.$$

De plus,

$$Y - H\hat{\mu}_c = Y - QY = \sigma(I - Q)\varepsilon.$$

Le théorème suivant est utile.

Théorème 7 (Cochran) *Soit U un vecteur gaussien centré réduit à valeurs dans \mathbb{R}^n et P la matrice de la projection orthogonale sur un sous-espace vectoriel F de dimension k . Alors les vecteurs aléatoires PU et $U - PU$ sont indépendantes et $\|PU\|^2$ suit la loi $\chi^2(k)$ (loi du Khi-deux à k degrés de liberté).*

Preuve du théorème. Les vecteurs PU et $U - PU$ sont gaussiens et il suffit de montrer qu'ils sont décorrélés pour affirmer qu'ils sont bien indépendants. Mais en utilisant les égalités $P' = P$ et $P^2 = P$, on obtient

$$\text{Cov}(PU, U - PU) = \mathbb{E}(U'P'U) - \mathbb{E}(U'P'PU) = \mathbb{E}(U'PU) - \mathbb{E}(U'PU) = 0,$$

ce qui prouve l'indépendance. De plus, si $\{f_1, \dots, f_k\}$ est une base orthonormée de F , on sait que

$$\|PU\|^2 = \sum_{j=1}^k (U'f_j)^2.$$

De plus, pour $1 \leq j \leq k$, $U'f_j \sim \mathcal{N}(0, 1)$ et $\text{Cov}(U'f_j, U'f_\ell) = 0$ si $\ell \neq j$. Il s'agit donc bien d'une loi $\chi^2(k)$. \square

En appliquant ce théorème à $P = Q$ et $U = \varepsilon$, on déduit que les estimateurs $\hat{\mu}$ et $\hat{\sigma}^2$ sont indépendants. En posant $P = I - Q$ (matrice de projection sur l'orthogonal de $\text{Vect}(H)$) et $U = \varepsilon$, on en déduit que $\hat{\sigma}^2 \sim \frac{\sigma^2}{n} \chi^2(n-1)$. De plus $\hat{\sigma}_c$ est l'estimé ajusté de fa Cov con à être sans biais. L'un ou l'autre de ces deux estimateurs converge en probabilité lorsque $n \rightarrow \infty$ puisque pour $\epsilon > 0$, on a

$$\begin{aligned} \mathbb{P}(|\hat{\sigma}^2 - \sigma^2| > \epsilon) &= \mathbb{P}\left(\left|\sum_{i=1}^n (Z_i^2 - 1)\right| > n\epsilon\sigma^2\right) \\ &\leq \frac{1}{n\epsilon^2\sigma^4} \mathbb{E}((Z_1^2 - 1)^2), \end{aligned}$$

en considérant des variables aléatoires Z_1, \dots, Z_n i.i.d $\mathcal{N}(0, 1)$ et en appliquant l'inégalité de Tchebychev.

2.4 Variation quadratique et variation totale

Proposition 10 Soit Δ une subdivision $0 = t_0 < t_1 < \dots < t_p = T$ de l'intervalle $[0, T]$ et soit $|\Delta| = \max_{i=0, \dots, p-1} (t_{i+1} - t_i)$ le pas de cette subdivision. Alors

$$\lim_{|\Delta| \rightarrow 0} \sum_{i=0}^{p-1} |B_{t_{i+1}} - B_{t_i}|^2 = T, \quad \text{dans } \mathbb{L}^2.$$

On dit que T est la variation quadratique du mouvement brownien sur $[0, T]$.

Preuve. Posons $V = \sum_{i=0}^{p-1} |B_{t_{i+1}} - B_{t_i}|^2$. Alors $\mathbb{E}(V) = T$ et

$$\mathbb{E}(|V - T|^2) = \text{Var}(V) = 2 \sum_{i=0}^{p-1} |t_{i+1} - t_i|^2 \leq 2T|\Delta|,$$

ce qui prouve la proposition. \square

En revanche, la variation dite totale est infinie. Cette propriété aura son importance lorsque nous construirons l'intégrale stochastique. Si $f : [0, T] \rightarrow \mathbb{R}$ est une fonction et $\Delta = (t_0, \dots, t_p)$ est une subdivision de $[0, T]$, notons

$$V_f^\Delta(0, T) = \sum_{i=0}^{p-1} |f(t_{i+1}) - f(t_i)|$$

et

$$V_f(0, T) = \sup \{ V_f^\Delta(0, T) / \Delta \text{ subdivision de } [0, T] \}.$$

$V_f(0, T)$ est appelé la variation totale de f sur $[0, T]$.

Proposition 11 On a

$$V_B(0, T) = +\infty, \text{ p.s.}$$

Preuve. Il suffit de remarquer que si Δ est une subdivision de $[0, T]$ alors

$$\begin{aligned} \sum_{i=0}^{p-1} |B_{t_{i+1}} - B_{t_i}|^2 &\leq \sup_i |B_{t_{i+1}} - B_{t_i}| \times \sum_{i=0}^{p-1} |B_{t_{i+1}} - B_{t_i}| \\ &\leq \sup_i |B_{t_{i+1}} - B_{t_i}| \times V_B(0, T). \end{aligned}$$

Lorsque $|\Delta| \rightarrow 0$ alors $\sup_i |B_{t_{i+1}} - B_{t_i}| \rightarrow 0$ p.s à cause de la continuité uniforme des trajectoires sur $[0, T]$. En revanche pour une suite de subdivision Δ_n dont le pas tend vers 0, on peut supposer quitte à considérer une sous-suite que la somme des carrés des accroissements converge p.s vers $T > 0$. D'après les inégalités précédentes on a forcément $V_B(0, T) = +\infty$ p.s. \square

Remarque. La preuve de la proposition précédente montre également que si la variation totale d'une fonction continue est finie alors sa variation quadratique est nulle. Une fonction f de classe \mathcal{C}^1 sur un intervalle $[0, T]$ est à variation finie et on peut montrer que la variation totale de f sur $[0, T]$ vaut $\int_0^T |f'(s)| ds < \infty$. Le mouvement brownien a donc des trajectoires beaucoup plus irrégulières. D'ailleurs le taux d'accroissement $\frac{B_{t+h} - B_t}{h} \sim \mathcal{N}(0, \frac{1}{h})$ converge vers $+\infty$ en probabilité ce qui empêche les trajectoires d'être dérivables au point t (on peut même montrer que p.s, les trajectoires du mouvement brownien sont non dérivables en tout point).

2.5 Exercices

EXERCICE 1 Sans utiliser le théorème de Donsker, montrer que les lois fini-dimensionnelles de la marche aléatoire convergent vers les lois fini-dimensionnelles correspondantes du mouvement brownien (on pourra plutôt montrer la convergence en loi des accroissements). Le lemme de Slutsky s'avèrera utile (voir [3] pour une preuve de ce résultat).

Lemme 2 Soient $(X_n)_{n \in \mathbb{N}}$ et $Y = (Y_n)_{n \in \mathbb{N}}$ deux suites de variables aléatoires à valeurs dans \mathbb{R}^k et \mathbb{R}^ℓ respectivement. On suppose que la suite $(X_n)_n$ converge en loi vers une variable aléatoire X et que $(Y_n)_n$ converge en probabilité vers une constante $y \in \mathbb{R}^\ell$. Posons $Z_n = (X_n, Y_n)$ pour $n \in \mathbb{N}$. Alors la suite $(Z_n)_n$ converge en loi vers le couple (X, y) . En particulier lorsque $k = \ell$ ou $\ell = 1$, on a les convergences en loi suivantes :

$$\lim_{n \rightarrow \infty} (X_n + Y_n) = X + y, \quad \lim_{n \rightarrow \infty} X_n Y_n = X y.$$

EXERCICE 2 Construire une suite de fonctions continues $(z_n)_n$ de $[0, 1]$ dans \mathbb{R} , convergente point par point vers 0 avec $\sup_{t \in [0, 1]} |z_n(t)| = 1$ pour tout $n \in \mathbb{N}$. En déduire que la suite de mesures $(\delta_{z_n})_n$ converge vers δ_0 au sens des loi fini-dimensionnelles mais pas au sens de la convergence étroite sur $E = \mathcal{C}^0([0, 1], \mathbb{R})$.

EXERCICE 3 Soit $(S_i)_{i \in \mathbb{N}}$ la marche aléatoire symétrique et $M_n = \max_{0 \leq i \leq n} S_i$, $n \in \mathbb{N}$. En utilisant le principe de réflexion, montrer que si $(a, b) \in \mathbb{Z} \times \mathbb{N}$ vérifie $a \leq b$ alors

$$\mathbb{P}(M_n \geq b, S_n < a) = \mathbb{P}(S_n < a - 2b).$$

En déduire que pour tout $t \geq 0$, la loi du couple $(B_t, \sup_{s \leq t} B_s)$ (où B est un mouvement brownien) a pour densité la fonction $f : \mathbb{R}^2 \rightarrow \mathbb{R}_+$ définie par

$$f(x, y) = (2/\pi t^3)^{1/2} (2y - x) \exp\left(-\frac{(2y - x)^2}{2t}\right) \mathbf{1}_{y > 0, x \leq y}, \quad (x, y) \in \mathbb{R}^2.$$

On pourra appliquer le théorème de Donsker et obtenir une expression de la densité pour $t = 1$.

EXERCICE 4 Si U est un pont Brownien, montrer que $(U_{1-t})_{t \in [0, 1]}$ est aussi un pont Brownien et que le processus $\left((t+1)U_{\frac{t}{t+1}}\right)_{t \geq 0}$ est un mouvement Brownien.

EXERCICE 5 Soient X_1, \dots, X_n , n variables aléatoires indépendantes et toutes à valeurs réelles. On suppose que pour $1 \leq i \leq n$, X_i a une densité $x \mapsto f(x, \theta_i)$ par rapport à la mesure de Lebesgue où $\theta_i \in \Theta$ est un paramètre inconnu. On considère les deux hypothèses suivantes :

H_0 : $\theta_i = \theta^*$, $1 \leq i \leq n$.

H_1 : Il existe $k \in \{1, \dots, n-1\}$ tel que $\theta_1 = \dots = \theta_k = \theta^*$ et $\theta_{k+1} = \dots = \theta_n = \theta^{**} \neq \theta^*$,

où θ^* et θ^{**} sont deux paramètres non connus. Un test classique est basée sur la statistique du rapport de vraisemblance

$$T_{n,k} = \frac{\sup_{\theta \in \Theta} L_n(\theta)}{\sup_{\theta_1 \in \Theta} L_k(\theta) \times \sup_{\theta_2 \in \Theta} L_{n,k}(\theta)}.$$

où $L_k(\theta) = \prod_{i=1}^k f(X_i, \theta)$ et $L_{n,k}(\theta) = \prod_{i=k+1}^n f(X_i, \theta)$ si $1 \leq k \leq n$. Lorsque les observations sont gaussiennes avec une variance σ^2 connue, montrer que si $1 \leq k < n$,

$$-2 \ln(T_{n,k}) = \left(\frac{1}{\sigma \sqrt{\frac{k}{n}(1 - \frac{k}{n})}} \max_{1 \leq k \leq n} |S_k - \frac{k}{n} S_n| \right)^2,$$

où pour $1 \leq i \leq n$, $S_i = X_1 + \dots + X_i$.

De fa Cov con générale, l'idée de ce test basé sur un rapport de vraisemblance est de rejeter H_0 si $Z_n = \max_{1 \leq k < n} -2 \ln(T_{n,k})$ n'est pas trop élevé.

EXERCICE 6 Soit $(B_t)_{0 \leq t \leq 1}$ un mouvement brownien. Pour $t \in [0, 1]$, on définit $\mathcal{G}_t = \sigma(B_1, B_s : s \leq t)$.

1. Montrer que si $0 \leq s \leq t \leq 1$, on a

$$\mathbb{E}(B_t - B_s | \mathcal{G}_s) = \frac{t-s}{1-s} (B_1 - B_s).$$

Indication : si $u_d < \dots < u_1 = s < t < 1$ on a toujours $\mathbb{E}(B_t - B_s | B_1, B_{u_1}, \dots, B_{u_d})$ de la forme

$$\alpha B_{u_d} + \sum_{j=2}^d \beta_j (B_{u_j} - B_{u_{j-1}}) + \beta (B_1 - B_{u_1}).$$

2. Montrer qu'en revanche, le processus β défini par

$$\beta_t = B_t - \int_0^t \frac{B_1 - B_s}{1-s} ds, \quad 0 \leq t \leq 1,$$

est un \mathcal{G}_t -mouvement brownien, indépendant de B_1 .

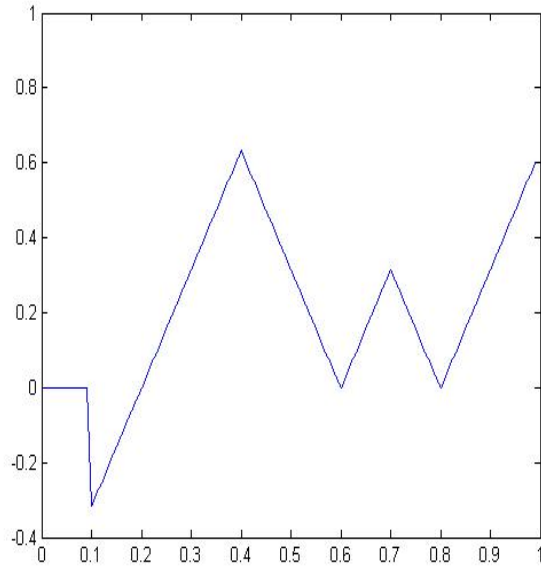


FIGURE 2.2: $t \rightarrow X_{10}(t)$

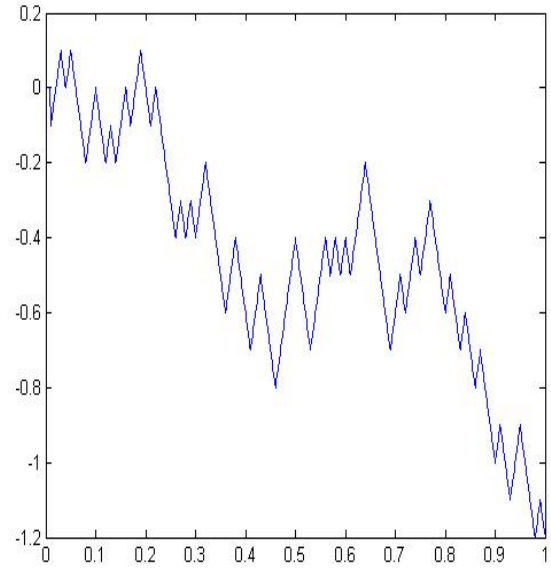


FIGURE 2.3: $t \rightarrow X_{100}(t)$

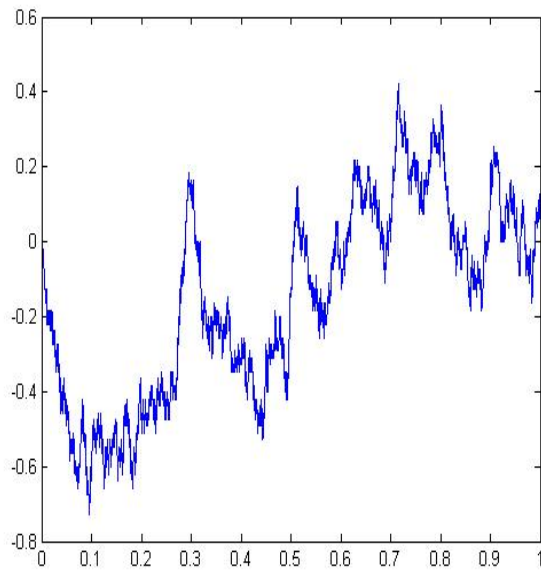


FIGURE 2.4: $t \rightarrow X_{3000}(t)$

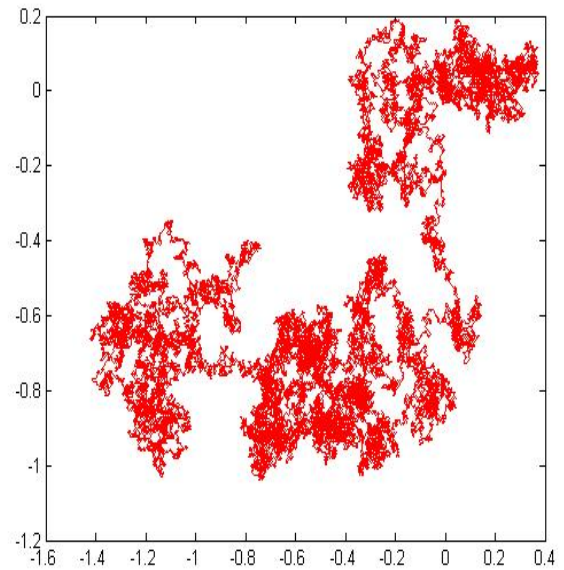


FIGURE 2.5: Approximation de la trajectoire d'un brownien 2D

Bibliographie

- [1] Csorgo, M., Horvath, L. *Limit Theorems in Change-Point Analysis*. John Wiley & Sons.
- [2] Ghorbanzadeh, D., Picard, D. (1991) *Étude asymptotique et pratique du comportement de deux tests de détection de rupture*. Statistique et analyse de données, tome 16, n°3, 63-84.
- [3] Van der Vaart, A.W. (2000) *Asymptotic statistics*. Cambridge University Press.

Chapitre 3

Processus de Poisson

3.1 Le processus de Poisson simple

Définition 11 Soit $(X_t)_{t \geq 0}$ un processus stochastique à valeurs réelles. On dit $(X_t)_{t \geq 0}$ est un processus de comptage si pour \mathbb{P} -presque tout $\omega \in \Omega$, la trajectoire $t \mapsto X_t(\omega)$ est croissante par sauts de 1, continue à droite et telle que $X_0 = 0$ p.s.

Définition 12 Un processus de comptage $(N_t)_{t \geq 0}$ est appelé un processus de Poisson simple si :

1. $N_0 = 0$ p.s.
2. le processus est à accroissements indépendants, c'est à dire $N_{t+s} - N_s$ est indépendante de $\sigma(N_u : u \leq s)$ pour tous $t, s \geq 0$.
3. Pour tous s et t positifs, la variable aléatoire $N_{t+s} - N_s$ suit une loi de Poisson de paramètre λt où λ est un réel strictement positif.

Remarque

- Les accroissements d'un processus de Poisson sont aussi stationnaires car la loi de $N_{t+s} - N_s$ est la même pour toute les valeur de s .
- Le paramètre λ est appelé l'intensité du processus (on rappelle que la moyenne d'une loi de Poisson de paramètre λ vaut λ).

Des exemples de modèles

- N_t peut représenter le nombre de clients arrivés à un guichet avant l'instant t (ou le nombre d'appels vers un central téléphonique). On peut également s'intéresser au nombre de pannes d'une machine à l'instant t , la machine étant réparé immédiatement après chaque panne.
- Ce processus a aussi été aussi utilisé pour modéliser le nombre de désintégrations de particules radioactives dans un intervalle de temps donné. Plus précisément si on considère un n -uplet $(X_1^{(n)}, \dots, X_n^{(n)})$ de variables aléatoires toutes indépendantes entre elles et toutes de loi exponentielle de paramètre λ/n (représentant les durée de vies des particules), alors le réarrangement croissant de ces variables aléatoires converge en loi vers la suite des temps de sauts d'un processus de Poisson d'intensité λ (voir [2] page 247).

Proposition 12 Soit N un processus de Poisson simple. Alors N est un processus stationnaire d'événements rares :

- $\mathbb{P}(N_{t+h} - N_t = 1) = \lambda h + o(h)$.
- $\mathbb{P}(N_{t+h} - N_t > 1) = o(h)$.

Preuve. On a

$$\begin{aligned}\mathbb{P}(N_{t+h} - N_t = 1) &= \lambda h \exp(-\lambda h) \\ &= \lambda h (1 + o(1)) \\ &= \lambda h + o(h).\end{aligned}$$

De même, on a

$$\begin{aligned}\mathbb{P}(N_{t+h} - N_t > 1) &= 1 - \lambda h \exp(-\lambda h) - \exp(-\lambda h) \\ &= o(h).\square\end{aligned}$$

Proposition 13 *Si $0 = T_0 < T_1 < \dots < T_n < \dots$ sont les instants de sauts du processus de Poisson simple alors*

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

De plus les instants $(T_n - T_{n-1})_{n \geq 1}$ d'inter-arrivées sont des variables aléatoires indépendantes et toutes de loi exponentielle de paramètre λ . La densité de (T_1, \dots, T_n) est donnée par

$$f_{(T_1, \dots, T_n)}(t_1, \dots, t_n) \lambda^n \exp(-\lambda t_n) \mathbb{1}_{0 < t_1 < \dots < t_n}.$$

Preuve. Soient $t_1, h_1, \dots, t_n, h_n$ des nombres réels positifs tels que

$$0 < t_1 < t_1 + h_1 < t_2 < t_2 + h_2 < \dots < t_n < t_n + h_n.$$

On a alors en utilisant l'indépendance et la loi des accroissements

$$\begin{aligned}&\mathbb{P}(t_1 < T_1 < t_1 + h_1, \dots, t_n < T_n < t_n + h_n) \\ &= \mathbb{P}(N_{t_1} = 0, N_{t_1+h_1} - N_{t_1} = 1, N_{t_2} - N_{t_1+h_1} = 0, \dots, N_{t_n} - N_{t_{n-1}+h_{n-1}} = 0, N_{t_n+h_n} - N_{t_n} \geq 1) \\ &= \exp(-\lambda t_n) \lambda^n h_1 \dots h_{n-1} (1 - \exp(-\lambda h_n)).\end{aligned}$$

En divisant par le produit $h_1 \dots h_n$ et en faisant tendre h vers 0, on obtient la loi annoncée pour le vecteur (T_1, \dots, T_n) . La formule du changement de variables permet alors de montrer que la densité du vecteur $(T_i - T_{i-1})_{1 \leq i \leq n}$ est bien celle d'un n -uplet de variables aléatoires indépendantes toutes de loi exponentielle de paramètre λ . \square

Remarque. Ce résultat permet d'envisager une définition du processus de Poisson à partir d'une suite i.i.d ξ de variables aléatoires toutes de loi exponentielle de paramètre λ . En effet en posant $T_n = \sum_{i=1}^n \xi_i$ le processus N défini par $N_t = \sum_{n \geq 1} \mathbb{1}_{T_n \leq t}$ si $t \in \mathbb{R}_+$ a la même loi qu'un processus de Poisson (si ce processus de Poisson existe vraiment). Il faut donc vérifier que le processus obtenu avec cette suite i.i.d de variables aléatoires vérifient bien la définition du processus de Poisson. On pourra consulter l'ouvrage [2] pour les détails de cette vérification.

Proposition 14 *Conditionnellement à $\{N_t = n\}$, le vecteur (T_1, \dots, T_n) a même loi que les statistiques d'ordre de n variables aléatoires indépendantes et toutes de loi uniforme sur l'intervalle $[0, t]$.*

Preuve. Si U_1, \dots, U_n sont des variables aléatoires i.i.d toutes de loi uniforme sur $[0, t]$, alors le vecteur des statistiques d'ordre $(U_{(1)}, \dots, U_{(n)})$ admet pour densité la fonction $f : \mathbb{R}^n \rightarrow \mathbb{R}_+$ définie par

$$f(t_1, \dots, t_n) = \frac{n!}{t^n} \mathbb{1}_{0 < t_1 < \dots < t_n \leq t}.$$

En utilisant la loi du $(n+1)$ -uplet (T_1, \dots, T_{n+1}) , on obtient pour tout borélien $B \in \mathcal{B}(\mathbb{R}^n)$

$$\begin{aligned} \mathbb{P}((T_1, \dots, T_n) \in B, N_t = n) &= \mathbb{P}((T_1, \dots, T_n) \in B, T_n \leq t < T_{n+1}) \\ &= \int \lambda^{n+1} \exp(-\lambda t_{n+1}) \mathbb{1}_B(t_1, \dots, t_n) \mathbb{1}_{0 < t_1 < \dots < t_n \leq t < t_{n+1}} dt_1 \cdots dt_{n+1} \\ &= \int \lambda^n \exp(-\lambda t) \mathbb{1}_B(t_1, \dots, t_n) \mathbb{1}_{0 < t_1 < \dots < t_n \leq t} dt_1 \cdots dt_n. \end{aligned}$$

Étant donné que $\mathbb{P}(N_t = n) = \exp(-\lambda t) \frac{\lambda^n t^n}{n!}$, le résultat est obtenue en divisant. \square

Au niveau du comportement asymptotique de N , nous avons les deux résultats suivants.

Théorème 8 1. On a $\lim_{t \rightarrow \infty} \frac{N_t}{t} = \lambda$ p.s.
2. On a de plus la convergence en loi

$$\sqrt{\frac{t}{\lambda}} \left(\frac{N_t}{t} - \lambda \right) \xrightarrow{t \rightarrow +\infty} \mathcal{N}(0, 1).$$

Preuve.

1. Nous avons l'encadrement $T_{N_t} \leq t < T_{N_t+1}$ p.s et donc

$$\frac{N_t}{T_{N_t+1}} \leq \frac{N_t}{t} \leq \frac{N_t}{T_{N_t}}. \quad (3.1)$$

Il est facile de montrer que $\lim_{t \rightarrow \infty} N_t = +\infty$ (exercice). De plus, comme $T_n = \sum_{i=1}^n (T_i - T_{i-1})$, la loi des grands nombres assure que $\lim_{n \rightarrow \infty} \frac{T_n}{n} = \frac{1}{\lambda}$ p.s (on rappelle que les intervalles entre les sauts sont i.i.d de loi exponentielle de paramètre λ). Il est alors facile d'en déduire qu'avec probabilité 1, les bornes de l'encadrement 3.1 tendent vers λ lorsque n tend vers l'infini.

2. Une première solution est d'utiliser le critère des fonctions caractéristiques (exercice). Une autre solution consiste à utiliser le théorème de la limite centrale en utilisant la décomposition

$$\sqrt{\frac{t}{\lambda}} \left(\frac{N_t}{t} - \lambda \right) = \sqrt{\frac{t}{\lambda}} \left(\frac{N_{[t]}}{[t]} - \lambda \right) + A_t.$$

– Pour le premier terme, on a

$$\sqrt{\frac{t}{\lambda}} \left(\frac{N_{[t]}}{[t]} - \lambda \right) = \sqrt{\frac{t}{[t]}} \sqrt{\frac{[t]}{\lambda}} \left(\frac{\sum_{i=1}^{[t]} (N_i - N_{i-1})}{[t]} - \lambda \right).$$

On obtient alors la convergence en loi de ce premier terme vers la loi $\mathcal{N}(0, 1)$ en utilisant le théorème central limite et le lemme de Slutsky (car $\lim_{t \rightarrow \infty} t/[t] = 1$).

– Pour le deuxième terme, on a

$$\begin{aligned} |A_t| &= \sqrt{\frac{t}{\lambda}} \left| \frac{N_{[t]}}{[t]} - \frac{N_t}{t} \right| \\ &\leq \sqrt{\frac{t}{\lambda}} \left(\frac{N_t - N_{[t]}}{[t]} + \frac{N_t}{[t]t} \right). \end{aligned}$$

On pourra montrer que le premier de deux termes majorants tend vers 0 en probabilité en utilisant l'inégalité de Markov. Comme le deuxième tend vers 0 p.s en utilisant le premier point du théorème, on conclut que $\lim_{t \rightarrow \infty} A_t = 0$ en probabilité.

D'après ce qui précède et en utilisant le lemme de Slutsky, on obtient la convergence annoncée. \square

3.1.1 Estimation de l'intensité par maximum de vraisemblance

Nous allons considérer trois cas d'échantillonnages possibles.

1. Considérons d'abord le cas où le processus est observé jusqu'à l'instant t . La variable aléatoire $\frac{N_t}{t}$ est en fait l'estimateur du maximum de vraisemblance du paramètre λ . Dans ce chapitre, nous calculerons la vraisemblance de fa Cov con informelle; le calcul sera justifié au chapitre suivant. L'idée est que lorsqu'on observe $N_t = n$, les temps de sauts T_1, \dots, T_n détermine complètement la trajectoire du processus sur $[0, t]$. Nous définirons donc la vraisemblance en calculant la loi de $(T_1, \dots, T_{N_t}, N_t)$. Ceci peut être fait en utilisant la Proposition 14. En effet, la loi de $(T_1, \dots, T_n) | N_t = n$ a pour densité

$$f : (t_1, \dots, t_n) \rightarrow \frac{n!}{t^n} \mathbf{1}_{0 < t_1 < \dots < t_n \leq t}.$$

La fonction de vraisemblance sera alors définie pour une observation (t_1, \dots, t_n, n) par

$$L(t_1, \dots, t_n, n, \lambda) = f(t_1, \dots, t_n) \mathbb{P}(N_t = n) = \lambda^n \exp(-\lambda t).$$

Le maximum de vraisemblance est alors $\hat{\lambda}(\omega) = \frac{n}{t} = \frac{N_t(\omega)}{t}$. On a vu que cet estimateur était consistant et asymptotiquement gaussien. De plus l'utilisation du lemme de Slutsky permet de montrer la convergence en loi

$$\sqrt{\frac{t^2}{N_t}} \left(\frac{N_t}{t} - \lambda \right) \xrightarrow{t \rightarrow \infty} \mathcal{N}(0, 1).$$

Un intervalle de confiance de niveau asymptotique $1 - \alpha$ est alors

$$I_n = \left[-\sqrt{\frac{N_t}{t^2}} q_{1-\frac{\alpha}{2}} + \frac{N_t}{t}, \sqrt{\frac{N_t}{t^2}} q_{1-\frac{\alpha}{2}} + \frac{N_t}{t} \right],$$

où $q_{1-\frac{\alpha}{2}}$ est le quantile d'ordre $1 - \frac{\alpha}{2}$ de la $\mathcal{N}(0, 1)$.

Remarque. Il existe un moyen de construire des intervalles de confiance de niveau exact lorsqu'on observe une loi de Poisson. Ce type d'intervalle ne sera pas détaillé ici et on pourra se référer au livre [1] (propositions 3.2 et B2) pour des détails.

2. On observe uniquement les temps de sauts T_1, \dots, T_n du processus. Dans ce cas la loi de (T_1, \dots, T_n) est donnée par la Proposition 13 et a pour densité

$$(t_1, \dots, t_n) \rightarrow \lambda^n \exp(-\lambda t_n) \mathbf{1}_{0 < t_1 < \dots < t_n},$$

par rapport à la mesure de Lebesgue sur \mathbb{R}^n . Il est alors facile de voir que le maximum de vraisemblance est donné par $\hat{\lambda} = \frac{n}{T_n} = \frac{N_{T_n}}{T_n}$. $\hat{\lambda}$ converge p.s vers λ (voir la preuve du Théorème 8). De plus, en décomposant $T_n = \sum_{i=0}^{n-1} (T_{i+1} - T_i)$ (somme de v.a indépendantes toutes de loi exponentielle de paramètre λ), une utilisation du théorème limite centrale permet d'obtenir la convergence en loi

$$\lim_{n \rightarrow +\infty} \sqrt{n} \left(\lambda \frac{T_n}{n} - 1 \right) = \mathcal{N}(0, 1).$$

A titre d'exercice, on pourra construire un intervalle de confiance de niveau donné pour λ à partir de cette convergence.

3. Supposons enfin que l'on observe le processus uniquement aux temps $t_1 < \dots < t_n$. On calcule alors le maximum de vraisemblance en utilisant l'indépendance des accroissements. La vraisemblance associée à l'échantillon $(N_{t_1}, N_{t_2} - N_{t_1}, \dots, N_{t_n} - N_{t_{n-1}})$ est alors donnée par

$$L(d_1, \dots, d_n, \lambda) = \exp(-\lambda t_n) \lambda^{\sum_{i=1}^n d_i} \frac{t_1^{d_1} \dots (t_n - t_{n-1})^{d_n}}{d_1! \dots d_n!}.$$

On obtient alors

$$\hat{\lambda} = \arg \max_{\lambda > 0} \{ \exp(-\lambda t_n) \lambda^{N_{t_n}} \} = \frac{N_{t_n}}{t_n}.$$

Remarquons que l'estimateur est identique à celui qui aurait été obtenu si toute la trajectoire avait été observée sur $[0, t_n]$. Ceci n'est pas étonnant car N_t est une statistique exhaustive du modèle (elle contient toute l'information).

3.2 Le processus de Poisson non homogène

3.2.1 Définition et propriétés fondamentales

Définition 13 Soit $\lambda : \mathbb{R}_+ \rightarrow \mathbb{R}_+$ une fonction localement intégrable et $\Lambda : \mathbb{R}_+ \rightarrow \mathbb{R}_+$ la fonction définie par

$$\Lambda(t) = \int_0^t \lambda(s) ds, \quad t \in \mathbb{R}_+.$$

On dit qu'un processus de comptage $(N_t)_{t \geq 0}$ est appelé un processus de Poisson d'intensité λ si

1. $N_0 = 0$ p.s.
2. Le processus est à accroissements indépendants.
3. $N_t - N_s$ suit une loi de Poisson de paramètre $\Lambda(t) - \Lambda(s)$, $0 \leq s < t$.

Remarques

- On a $\mathbb{E}(N_t) = \text{Var}(N_t) = \Lambda(t)$. De plus Λ est une fonction continue.
- Le processus de Poisson simple correspond à une intensité constante au cours du temps (i.e $\lambda(t) = \lambda$ pour tout $t \geq 0$).

Construction. La construction d'un processus de Poisson non homogène peut se faire à partir d'un processus de poisson simple. Si \tilde{N} est un processus de Poisson simple d'intensité 1, il est facile de vérifier que le processus \tilde{N} défini par $\tilde{N}_t = N_{\Lambda(t)}$ est un processus de Poisson non homogène d'intensité λ . Notons

$$\Lambda^{-1}(t) = \inf\{x \geq 0 : \Lambda(x) \geq t\}, \quad t \geq 0,$$

la pseudo-inverse continue à gauche de Λ (qui correspond à l'inverse si Λ est strictement croissante et de limite $+\infty$ en $+\infty$). On adopte la convention $\inf \emptyset = +\infty$. On pose alors $S_n = \Lambda^{-1}(T_n)$ pour $n \in \mathbb{N}^*$, où T_n est le temps du n -ième saut de N . La définition de Λ^{-1} assure que

$$\Lambda(x) \geq t \Leftrightarrow x \geq \Lambda^{-1}(t).$$

Ainsi on a

$$\tilde{N}_t = \sum_{n \geq 1} \mathbf{1}_{S_n \leq t}, \quad t \geq 0.$$

Applications

- On peut considérer que l'intensité des arrivées à un guichet subit une modification brusque au cours du temps : $\lambda(t) = \lambda_1$ si $t \leq t_0$ et $\lambda(t) = \lambda_2$ si $t > t_0$.
- En fiabilité, lorsque un matériel subit une succession de pannes et lorsque chaque panne est réparée instantanément, on utilise souvent un processus de Weibull, c'est à dire une processus de Poisson non homogène d'intensité :

$$\lambda(s) = \frac{\beta}{\alpha} \left(\frac{s}{\alpha}\right)^{\beta-1}, \quad s \geq 0.$$

On peut alors montrer que le premier instant de saut $S_1 = \Lambda^{-1}(T_1)$ (voir le paragraphe sur la construction) a pour loi la loi de Weibull de paramètres α et β dont la densité est donnée par

$$f(t) = \frac{\beta}{\alpha} \left(\frac{t}{\alpha}\right)^{\beta-1} \exp\left(-\left(\frac{t}{\alpha}\right)^\beta\right) \mathbf{1}_{\mathbb{R}_+}(t).$$

La loi de Weibull est souvent utilisé en fiabilité car le taux de défaillance d'un matériel dont la durée de vie T a une densité sur \mathbb{R}_+ est défini par

$$\lim_{h \rightarrow 0^+} \frac{1}{h} \mathbb{P}(t < T \leq t + h | T > t)$$

et la valeur de ce dernier correspond bien à celui obtenu pour la loi de Weibull, c'est à dire $\lambda(t)$ (plus précisément le taux de défaillance est souvent une fonction affine du temps dans un graphe log-log, d'après des études empiriques). De plus si $\beta = 1$ (taux de panne constant autour du temps) on retrouve la loi exponentielle de paramètre $\lambda = \frac{1}{\alpha}$ et donc le processus de Poisson simple. Si $\beta > 1$, le taux de défaillance augmente avec le temps (usure du système) alors que si $\beta < 1$, le taux de défaillance diminue avec le temps (les éléments défectueux qui fragilisent le système tombent en panne rapidement). Les figures 3.1, 3.2 et 3.3 montrent des trajectoires dans les trois cas.

Proposition 15 *Supposons l'intensité λ continu au voisinage de t . Alors*

$$\mathbb{P}(N_{t+h} - N_t = 1) = \lambda(t)h + o(h),$$

$$\mathbb{P}(N_{t+h} - N_t \geq 2) = o(h).$$

Preuve. Il suffit de faire un développement limité comme pour le processus de Poisson simple. On a

$$\begin{aligned} \mathbb{P}(N_{t+h} - N_t = 1) &= \exp(\Lambda(t) - \Lambda(t+h)) (\Lambda(t+h) - \Lambda(t)) \\ &= (1 + o(1)) (\lambda(t)h + o(h)) \\ &= \lambda(t)h + o(h). \end{aligned}$$

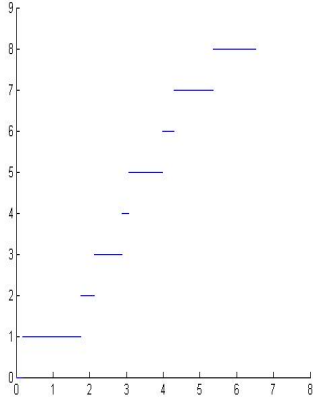


FIGURE 3.1: $\alpha = 1, \beta = 1$

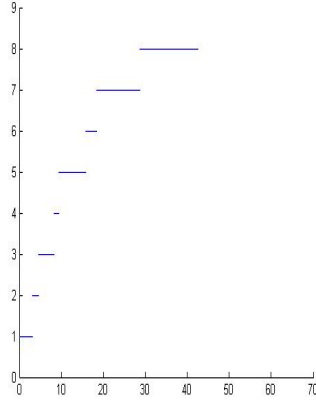


FIGURE 3.2: $\alpha = 1, \beta = 0.5$

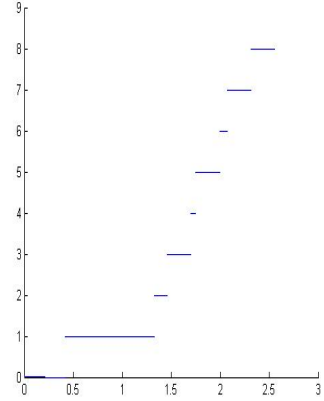


FIGURE 3.3: $\alpha = 1, \beta = 2$

On obtient ensuite

$$\begin{aligned}
 \mathbb{P}(N_{t+h} - N_t \geq 2) &= 1 - \mathbb{P}(N_{t+h} - N_t = 0) - \mathbb{P}(N_{t+h} - N_t = 1) \\
 &= 1 - \exp(\Lambda(t) - \Lambda(t+h)) - (\lambda(t)h + o(h)) \\
 &= \Lambda(t+h) - \Lambda(t) + o(\Lambda(t+h) - \Lambda(t)) - (\lambda(t)h + o(h)) \\
 &= o(h),
 \end{aligned}$$

en utilisant le fait que $\Lambda(t+h) - \Lambda(t) = \lambda(t)h + o(h) = O(h)$. \square

Nous pouvons également généraliser certains résultats sur la loi des temps de saut. **Pour simplifier nous supposons dans la suite de ce chapitre que l'intensité du processus vérifie toujours $\lim_{t \rightarrow +\infty} \Lambda(t) = +\infty$.**

Proposition 16 Soit $(T_n)_{n \geq 1}$ désigne les temps de sauts d'un processus de Poisson non homogène d'intensité λ . Soit $n \in \mathbb{N}^*$.

1. Pour $n \in \mathbb{N}^*$, la loi de (T_1, \dots, T_n) a pour densité

$$(t_1, \dots, t_n) \rightarrow \exp\left(-\int_0^{t_n} \lambda(u) du\right) \prod_{i=1}^n \lambda(t_i) \mathbb{1}_{0 < t_1 < \dots < t_n},$$

par rapport à la mesure de Lebesgue.

2. Soit $t < 0$. Alors la loi conditionnelle de $(T_1, \dots, T_n) | N_t = n$ a pour densité

$$(t_1, \dots, t_n) \rightarrow \frac{n!}{\Lambda(t)^n} \prod_{i=1}^n \lambda(t_i) \mathbb{1}_{0 < t_1 < \dots < t_n < t}.$$

Remarque. Rappelons que si X_1, \dots, X_n sont n variables aléatoires i.i.d de loi à densité f , le réarrangement croissant $(X_{(1)}, \dots, X_{(n)})$ de ces variables admet une loi dont la densité est

$$(y_1, \dots, y_n) \rightarrow n! \prod_{i=1}^n f(y_i) \mathbb{1}_{0 < y_1 < \dots < y_n}$$

par rapport à la mesure de Lebesgue. Ainsi la proposition ci-dessus assure que la loi conditionnelle de $(T_1, \dots, T_n) | N_t = n$ est celle du réarrangement croissant d'un n -échantillon dont la densité f est définie par $f(x) = \frac{\lambda(x)}{\Lambda(t)} \mathbb{1}_{[0, t]}(x)$.

Preuve

1. Pour le premier point, on se restreindra au cas d'une intensité continue. Dans ce cas Λ réalise un \mathcal{C}^1 difféomorphisme de \mathbb{R}_+^* sur lui même, d'inverse Λ^{-1} . D'après la construction du processus de Poisson non homogène, les temps $S_1 = \Lambda(T_1), \dots, S_n = \Lambda(T_n)$ sont les n premiers temps de saut d'un processus de Poisson homogène d'intensité 1. En appliquant la formule du changement de variable et le résultat de la Proposition 13, on a donc pour une fonction $\phi : \mathbb{R}^n \rightarrow \mathbb{R}$ mesurable bornée :

$$\begin{aligned} \mathbb{E}(\phi(T_1, \dots, T_n)) &= \mathbb{E}(\phi(\Lambda^{-1}(S_1), \dots, \Lambda^{-1}(S_n))) \\ &= \int \phi(\Lambda^{-1}(s_1), \dots, \Lambda^{-1}(s_n)) \exp(-s_n) \mathbf{1}_{0 < s_1 < \dots < s_n} ds_1 \cdots ds_n \\ &= \int \phi(t_1, \dots, t_n) \exp(-\Lambda(t_n)) \mathbf{1}_{0 < t_1 < \dots < t_n} \lambda(t_1) \cdots \lambda(t_n) dt_1 \cdots dt_n. \end{aligned}$$

Pour le cas général d'une intensité non nécessairement continue, on pourra consulter [1] p.40.

2. On en déduit le deuxième point exactement comme pour le processus de Poisson simple (voir la preuve de la Proposition 14). \square

3.2.2 Estimation par maximum de vraisemblance

Comme pour le cas du processus de Poisson simple, on peut utiliser la Proposition 16 pour calculer la fonction de vraisemblance du modèle. Nous nous placerons dans un modèle paramétrique où l'intensité du processus appartient à une famille $\{\lambda_\theta : \theta \in \Theta\}$. Nous supposerons ici uniquement le cas d'une trajectoire observée sur l'intervalle $[0, t]$. Si on observe $N_t = n$ et $(T_1, \dots, T_n) = (t_1, \dots, t_n)$, on a alors

$$L(t_1, \dots, t_n, n; \theta) = \prod_{i=1}^n \lambda_\theta(t_i) \exp(-\Lambda_\theta(t)).$$

Nous allons effectuer les calculs lorsque le processus est un processus de Weibull. Dans ce cas $\theta = (\alpha, \beta)$ et

$$\lambda_\theta(t) = \frac{\beta}{\alpha} \left(\frac{s}{\alpha}\right)^{\beta-1}, \quad s > 0.$$

Ainsi la log-vraisemblance s'écrit

$$\ell(t_1, \dots, t_n, n; \theta) = (\ln(\beta) - \beta \ln(\alpha)) n + (\beta - 1) \sum_{i=1}^n \ln(t_i) - \left(\frac{t}{\alpha}\right)^\beta.$$

L'unique point (α, β) qui annule les deux dérivées partielles est donné par

$$\frac{1}{\beta} = \ln(t) - \frac{1}{n} \sum_{i=1}^n \ln(t_i),$$

$$\ln(\alpha) = \ln(t) - \frac{1}{\beta} \ln(n).$$

On peut voir que ce point correspond au maximum de la vraisemblance en procédant comme suit. Posons $\lambda = \frac{1}{\alpha^\beta}$, $s = \sum_{i=1}^n \ln(t_i)$ et

$$g(\lambda, \beta) = n \ln(\beta) + n \ln(\lambda) + \beta s - \lambda t^\beta.$$

Pour $\beta > 0$ donné, la fonction concave $\lambda \rightarrow g(\lambda, \beta)$ atteint son maximum en $\lambda = nt^{-\beta}$, maximum qui vaut

$$h(\beta) = n \ln(\beta) + n \ln(n) - \beta n \ln(t) + \beta s - n.$$

La fonction h est concave sur $]0, +\infty[$ et atteint son maximum en $\beta = \frac{n}{n \ln(t) - s}$. On en déduit un maximum de la vraisemblance au point (α, β) annoncé. Ainsi le maximum de vraisemblance $\hat{\theta} = (\hat{\alpha}, \hat{\beta})$ est donné par

$$\frac{1}{\hat{\beta}} = \ln(t) - \frac{1}{N_t} \sum_{i=1}^{N_t} \ln(T_i),$$

$$\ln(\hat{\alpha}) = \ln(t) - \frac{1}{\hat{\beta}} \ln(N_t).$$

Nous n'étudierons pas les propriétés de convergence de ces estimateurs (on pourra consulter [1], Chapitre 3 pour cette étude). Il est néanmoins possible de donner des intervalles de confiance pour le paramètre β , en utilisant les propriétés que nous avons vues.

Proposition 17 *Dans le cas de l'observation d'un processus de Weibull sur $[0, t]$, la loi de $\frac{2n\beta}{\hat{\beta}}$ sachant $N_t = n$ ($n \neq 0$) est la loi du χ^2 à $2n$ degrés de liberté. Ainsi, chacun des intervalles I_n suivants est un intervalle de confiance de niveau γ , au sens où $\mathbb{P}(\beta \in I_n | N_t = n) = \gamma$:*

- $I_n = [0, \frac{\hat{\beta}}{2n} \chi_{\gamma}^2(2n)]$,
- $I_n = [\frac{\hat{\beta}}{2n} \chi_{1-\gamma}^2(2n), +\infty[$,
- $I_n = [\frac{\hat{\beta}}{2n} \chi_{\frac{1-\gamma}{2}}^2(2n), \frac{\hat{\beta}}{2n} \chi_{\frac{1+\gamma}{2}}^2(2n)]$.

Le nombre $\chi_{\gamma}^2(2n)$ désigne le quantile d'ordre γ de la loi du χ^2 à $2n$ degrés de liberté.

Preuve. En utilisant l'expression de $\hat{\beta}$ et la Proposition 16, on peut vérifier que la loi conditionnelle de $\frac{2n\beta}{\hat{\beta}} | N_t = n$ est la même que celle de

$$Z_n = 2n\beta - 2\beta \sum_{i=1}^n \ln(X_i),$$

où (X_1, \dots, X_n) est un n -échantillon de loi de densité $s \rightarrow \frac{\lambda(s)}{\Lambda(t)} \mathbf{1}_{[0,t]}(s)$ (en utilisant aussi le fait que la somme de n variables aléatoires correspond à la somme des statistiques d'ordre). En effectuant le changement de variables $u = \ln(x)$, on peut montrer que la fonction caractéristique de $\ln(X_1)$ est donnée par

$$\phi_{\ln(X_1)}(t) = \frac{\beta \exp(\ln(t)(is + \beta))}{\Lambda(t) \alpha^{\beta}(is + \beta)}.$$

En utilisant les propriétés d'indépendance des X_i et l'expression de l'intensité λ , on en déduit

$$\phi_{Z_n}(t) = \frac{1}{(1 - it)^n}.$$

La fonction caractéristique de Z_n est la même que celle de la loi du χ^2 à $2n$ degrés de liberté. Les propriétés des intervalles de confiance sont alors immédiates. \square

3.2.3 Test de Laplace

Supposons que l'on veuille tester l'homogénéité d'un processus de Poisson observé sur $[0, t]$. Nous allons considérer ce problème de fa Cov con non paramétrique en testant les deux hypothèses suivantes.

H_0 : N est un processus de Poisson simple.

H_1 : N est un processus de Poisson dont l'intensité est décroissante au cours du temps (resp. N est un processus de Poisson dont l'intensité est croissante au cours du temps).

Par décroissante, on veut dire décroissant au sens large et $\lambda(u) > \lambda(v)$ pour au moins un couple (u, v) tel que $0 < u < v < t$.

Proposition 18 *Supposons que $\int_0^{+\infty} \lambda(s)ds = +\infty$ et posons*

$$Z_t = \frac{\sum_{i=1}^n T_i - nt/2}{t\sqrt{n/12}}, \text{ sur } \{N_t = n\}.$$

1. Pour tester H_0 contre H_1 "l'intensité est décroissante", la région critique (rejet de H_0) peut être définie par

$$D = \{Z_t < q_\alpha\},$$

où q_α est le quantile d'ordre α de la loi $\mathcal{N}(0, 1)$.

2. Pour tester H_0 contre H_1 "l'intensité est croissante", la région critique (rejet de H_0) peut être définie par

$$D = \{Z_t > q_{1-\alpha}\},$$

où $q_{1-\alpha}$ est le quantile d'ordre $1 - \alpha$ de la loi $\mathcal{N}(0, 1)$.

Dans les deux cas, on obtient un test de niveau asymptotique α et dont la puissance tend vers 1.

Preuve. Pour justifier le niveau du test, on commence par déterminer la loi asymptotique de Z_t lorsque $t \rightarrow +\infty$. Comme la loi conditionnelle de $Z_t | N_t = n$ est aussi celle de la variable aléatoire

$$\frac{\sum_{i=1}^n X_i - nt/2}{t\sqrt{n/12}},$$

où les X_i sont i.i.d de loi uniforme sur $[0, t]$ (cf Proposition 14), le calcul de la fonction caractéristique entraîne que

$$\lim_{n \rightarrow +\infty} \mathbb{E}(\exp(iuZ_t) | N_t = n) = \exp(-u^2/2),$$

en utilisant le théorème limite centrale. Vu que sous H_0 , on a $\lim_{t \rightarrow +\infty} N_t = +\infty$, le théorème de convergence dominée assure que

$$\phi_{Z_t}(u) = \mathbb{E}(\mathbb{E}(\exp(iuZ_t) | N_t)) \rightarrow_{t \rightarrow +\infty} \exp(-u^2/2).$$

Dans les deux cas de figure, on a bien un test de niveau asymptotique α .

Examinons la puissance dans le cas où l'intensité décroissante sous H_1 (lorsque l'intensité est croissante, on a une preuve analogue que l'on pourra faire en exercice). Conditionnellement à $N_t = n$, Z_t a même loi (sous H_1) que

$$\frac{\sum_{i=1}^n X_i - nt/2}{t\sqrt{n/12}},$$

où la densité de X_1 est la fonction f donnée par

$$f(x) = \frac{1}{\Lambda(t)} \lambda(x) \mathbf{1}_{[0,t]}(x),$$

qui est une fonction décroissante. Ceci entraîne que $m = \mathbb{E}(X_1) < \frac{t}{2}$ (cf Lemme 3 ci-après appliqué à la densité de $\frac{X_1}{t}$) Ainsi

$$\mathbb{P}(D|N_t = n) = \mathbb{P}\left(\frac{\sum_{i=1}^n X_i - nm}{t\sqrt{n/12}} < q_\alpha + n\frac{(t/2) - m}{t\sqrt{n/12}}\right),$$

et la limite de cette dernière quantité vaut 1 lorsque n tend vers l'infini en utilisant le théorème limite centrale et le fait que $m < t/2$. Ainsi $\mathbb{P}(D) = \mathbb{E}(\mathbb{E}(\mathbf{1}_D|N_t))$ tend vers 1 par convergence dominée car sous l'hypothèse $\int_0^{+\infty} \lambda(s)ds = +\infty$, on a $\lim_{t \rightarrow +\infty} N_t = +\infty$ (ce qui peut se voir en reprenant la construction du processus non homogène à partir du processus de Poisson simple). \square

Lemme 3 Soit f une densité de probabilité sur $[0, 1]$, décroissante et telle qu'il existe s, t dans $]0, 1[$ avec $s < t$ et $f(s) > f(t)$, alors

$$\int_0^1 xf(x)dx < \frac{1}{2}.$$

Preuve. Soit $h : [0, 1] \rightarrow \mathbb{R}$ la fonction définie par

$$h(x) = \int_0^x f(u)du - x.$$

Soit également

$$c = \inf\{u \in [0, 1] : f(u) \leq 1\}.$$

c est bien définie car si $f(u) > 1$ pour $u \in [0, 1]$ alors $\int_0^1 f(u)du$ serait plus grande que 1 (et donc $c < 1$). De plus, $0 < c < 1$, car $c = 0$ entraînerait $f(u) = 1$ pour tout u . Si $0 \leq \alpha < \beta < c$, alors

$$h(\beta) - h(\alpha) = \int_\alpha^\beta f(u)du - (\beta - \alpha) > 0,$$

ce qui montre que h est strictement croissante sur $[0, c]$. On peut de même voir que h est strictement décroissante sur $[c, 1]$. On a donc $h(0) = h(1) = 0$ et $h(s) > 0$ si $s \in]0, 1[$. La fonction de répartition F de f vérifie donc $F(x) > x$ pour $x \in]0, 1[$. En utilisant la formule

$$\int_0^1 xf(x)dx = \int_0^1 (1 - F(u)) du,$$

on voit que

$$\int_0^1 xf(x)dx < \int_0^1 (1 - u)du = \frac{1}{2},$$

ce qui prouve le résultat. \square

Remarque. Sous les hypothèses du lemme précédent, on dit qu'une variable aléatoire de densité f est stochastiquement inférieure à une variable aléatoire de loi uniforme sur $[0, 1]$ au sens où la fonction de répartition de la première variable est plus grande que celle de la deuxième (ce qui est montré dans la preuve).

3.3 Le processus de Poisson composé

Définition 14 Soit $(X_n)_{n \geq 1}$ une suite de variables aléatoires i.i.d et N un processus de Poisson simple indépendant de cette suite. Pour $t \leq 0$, posons

$$S_t = \sum_{k=1}^{N_t} X_k.$$

On dit que S est un processus de Poisson composé.

Ruine d'une compagnie d'assurances. On peut modéliser le nombre de sinistres survenu entre le temps 0 et le temps t à l'aide d'un processus de Poisson simple (on pourrait également utiliser un processus de Poisson non homogène ce qui serait sans doute plus réaliste). Chacun des sinistres survenus à un coût (le premier a un coût X_1 , le deuxième X_2, \dots). Le processus de Poisson composé S associé représente l'évolution du coût des sinistres au cours du temps.

Remarque. Le mouvement brownien, le processus de Poisson simple et le processus de Poisson composé ont un point commun : ce sont des processus càd-làg dont les accroissements sont indépendants et stationnaires (pour le processus de Poisson composé, voir les TD). Un processus possédant ces propriétés est appelé un processus de Levy. De part leur structure, les processus de Levy présentent de nombreuses caractéristiques communes et leurs propriétés générales ont beaucoup été étudiées.

3.4 Exercices

EXERCICE 7 Soient N et M deux processus de Poisson simples indépendants. Montrer que la somme $N + M$ est encore un processus de Poisson.

EXERCICE 8 On suppose que des particules arrivent suivant un processus de Poisson simple d'intensité λ . On suppose que chaque particule qui arrive a une probabilité p de passer à travers une barrière. A l'instant t , on a observé M_t particules ayant traversé cette barrière. Quel est le maximum de vraisemblance du paramètre λ ?

EXERCICE 9 Soit X un processus de Poisson composé : $X_t = \sum_{j=1}^{N_t} Y_j$, avec N processus de Poisson simple de paramètre λ et indépendant de la suite Y qui est i.i.d.

1. Montrer que X est un processus de Levy.
2. Si une réalisation de X est observée sur l'intervalle $[0, t]$, quelle est la fonction de vraisemblance associée ?

EXERCICE 10 On reprend les notations du cours et on considère $(\hat{\alpha}, \hat{\beta})$ l'estimateur du maximum de vraisemblance associé à un processus de Weibull observé sur $[0, t]$. On rappelle les expressions :

$$\frac{1}{\hat{\beta}} = \ln(t) - \frac{1}{N_t} \sum_{i=1}^{N_t} \ln(T_i),$$

$$\ln(\hat{\alpha}) = \ln(t) - \frac{1}{\hat{\beta}} \ln(N_t).$$

1. Question préliminaire : pour deux variables aléatoires X et Y à valeurs réelles avec Y de carré intégrable, on définit la variance conditionnelle de Y sachant X par

$$\text{Var} (Y|X) = \mathbb{E} \left((Y - \mathbb{E}(Y|X))^2 |X \right) = \mathbb{E} (Y^2|X) - (\mathbb{E}(Y|X))^2.$$

Démontrer la relation

$$\text{Var} (Y) = \mathbb{E} (\text{Var} (Y|X)) + \text{Var} (\mathbb{E}(Y|X)).$$

2. Calculer la moyenne et la variance de la variable aléatoire $\frac{1}{\beta}$.

Indication : on pourra d'abord calculer la moyenne et la variance sachant N_t en utilisant les propriétés démontrées en cours.

Bibliographie

- [1] Coccozza-Thivent, C. *Processus stochastiques et fiabilité des systèmes*. Springer.
- [2] Paul S. Toulouse *Thèmes de probabilités et statistique*. Dunod.

Chapitre 4

Processus de Markov à sauts

4.1 Définition générale d'un processus de Markov

Définition 15 Notons $I = \mathbb{R}_+$ et $E = \mathbb{R}^d$. Un processus stochastique adapté $(\Omega, \mathcal{F}, (\mathcal{F}_t)_{t \in I}, (X_t)_{t \in I}, \mathbb{P})$ est un processus de Markov si pour toute fonction $f : E \rightarrow \mathbb{R}$ bornée, on a pour $0 \leq s \leq t$,

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

Exemples. Le mouvement brownien B muni de sa filtration naturelle est un processus de Markov. Prenons $A \in \mathcal{F}_s$ et notons par g_t la densité de la loi $\mathcal{N}(0, t)$. En utilisant les propriétés du mouvement brownien et en reprenant les notations de la définition ci-dessus, on a

$$\begin{aligned} \mathbb{E}(\mathbb{1}_A f(B_t)) &= \mathbb{E}(f(B_s + B_t - B_s)) \\ &= \int \mathbb{E}(\mathbb{1}_A f(B_s + y)) g_{t-s}(y) dy \\ &= \mathbb{E}(\mathbb{1}_A \int f(B_s + y) g_{t-s}(y) dy) \\ &= \mathbb{E}(\mathbb{1}_A h(B_s)). \end{aligned}$$

La troisième égalité est une conséquence du théorème de Fubini. La fonction $h = \int f(\cdot + y) g_{t-s}(y) dy$ étant bornée, la caractérisation de l'espérance conditionnelle entraîne que $h(B_s) = \mathbb{E}(f(B_t) | \mathcal{F}_s)$. En prenant $A \in \sigma(B_s)$, on voit qu'on a aussi $h(B_s) = \mathbb{E}(f(B_t) | B_s)$. De plus, on voit

$$\mathbb{E}(f(B_t) | B_s = z) = \int f(z + y) g_{t-s}(y) dy = \int f(y) g_{t-s}(y - z) dy.$$

Ceci entraîne que la loi conditionnelle de $B_t | B_s = z$ admet pour densité la fonction

$$z \mapsto g_{t-s}(y - z) = \frac{1}{\sqrt{2\pi(t-s)}} \exp\left(-\frac{(y-z)^2}{2(t-s)}\right).$$

Plus généralement, on peut montrer avec la même démonstration que ci-dessus que tout processus de Levy est un processus de Markov (donc en particulier le processus de Poisson simple et composé). Le processus de Poisson non homogène est également un processus de Markov. La différence par rapport aux processus de Levy est que les accroissements sont non stationnaires et la loi de $X_{s+h} | X_s$ peut dépendre de s . Lorsque la loi de $X_{s+h} | X_s$ est indépendante de s (ce qui revient à dire que la loi de $X_t | X_s$ ne dépend que de $t - s$), on dit que le processus de Markov est homogène, ce qui sera toujours le cas pour les processus de sauts que nous étudierons dans ce chapitre.

4.2 Les processus de sauts

A partir de maintenant, on se donne un ensemble E fini ou infini dénombrable.

Définition 16 Un processus $(X_t)_{t \geq 0}$ à valeurs dans E est appelé un processus de Markov de sauts de semi-groupe $(P_t)_{t \geq 0}$ si pour tout $s, t \geq 0$ et $y \in E$,

$$\mathbb{P}(X_{s+h} = y | \mathcal{F}_s) = P_h(X_s, y),$$

avec $\mathcal{F}_s = \sigma(X_u : u \leq s)$ et si ses trajectoires sont continues à droite.

Si X vérifie cette définition alors c'est un processus de Markov homogène. La matrice $P_t : E \times E \rightarrow [0, 1]$ est aussi appelée noyau de transition.

Proposition 19 (équations de Chapman-Kolmogorov) La famille de noyaux de transition $\{P_t : t \geq 0\}$ possède la propriété dite de semi-groupe :

$$P_{t+s} = P_t P_s, \quad s, t \geq 0.$$

Preuve. Soit $i, j \in E$. On a

$$\begin{aligned} P_{t+s}(i, j) &= \mathbb{P}(X_{t+s} = j | X_0 = i) \\ &= \sum_{k \in E} \mathbb{P}(X_{t+s} = j, X_t = k | X_0 = i) \\ &= \sum_{k \in E} \mathbb{P}(X_{t+s} = j | X_t = k, X_0 = i) \mathbb{P}(X_t = k | X_0 = i) \\ &= \sum_{k \in E} P_s(k, j) P_t(i, k) \\ &= (P_t P_s)(i, j). \quad \square \end{aligned}$$

Remarques

1. On peut montrer qu'un processus X continu à droite est un processus de Markov de sauts de semi-groupe P si et seulement si pour tout entier $n \geq 2$, pour tous $0 \leq t_1 < t_2 < \dots < t_n$ et pour tout $(i_1, \dots, i_n) \in E^n$,

$$\mathbb{P}(X_{t_n} = i_n | X_{t_{n-1}} = i_{n-1}, \dots, X_1 = i_1) = P_{t_n - t_{n-1}}(i_{n-1}, i_n).$$

L'argument utilise le théorème des classes monotones (pour le sens réciproque).

2. Les lois fini-dimensionnelles du processus sont complètement déterminées par la loi de X_0 et la famille $\{P_t : t \geq 0\}$. Plus précisément, on a pour tous $0 < t_1 < \dots < t_n$:

$$\mathbb{P}(X_0 = i_0, X_{t_1} = i_1, \dots, X_{t_n} = i_n) = \mathbb{P}(X_0 = i_0) P_{t_1}(i_0, i_1) P_{t_2 - t_1}(i_1, i_2) \cdots P_{t_n - t_{n-1}}(i_{n-1}, i_n).$$

De plus, on voit que pour toute suite strictement croissante $(t_i)_{i \in \mathbb{N}}$ de nombres réels, le processus Y défini par $Y_n = X_{t_n}$ est une chaîne de Markov. La chaîne est homogène de matrice de transition P_h lorsque $t_{n+1} - t_n = h$ pour tout $n \in \mathbb{N}$.

Donnons un exemple de construction d'un processus de sauts à l'aide d'une chaîne de Markov et d'un processus de Poisson. Comme nous le verrons vers la fin de ce chapitre, cette construction se révélera assez générale. Nous aurons besoin de l'exponentielle d'une matrice $A : E \times E \rightarrow \mathbb{R}$. Définissons la norme de A par

$$\|A\| = \sup_{i \in E} \sum_{j \in E} |A(i, j)|.$$

Lorsque $\|A\| < +\infty$, on définit l'exponentielle de la matrice A par

$$\exp(A) = \sum_{n=0}^{+\infty} \frac{A^n}{n!}.$$

Cette série est bien défini car on a $\sup_{i,j \in E} |A^n(i,j)| \leq \|A\|^n$ et donc

$$\exp(\|A\|) = \sum_{n=0}^{+\infty} \frac{\|A\|^n}{n!} \geq \sum_{n=0}^{+\infty} \frac{|A^n(i,j)|}{n!}.$$

Lorsque A et B commutent et vérifient $\|A\|, \|B\| < +\infty$, on a l'égalité $\exp(A+B) = \exp(A)\exp(B)$.

Proposition 20 Soit $(Z_n)_{n \in \mathbb{N}}$ une chaîne de Markov sur E de probabilité de transition P et indépendante d'un processus de Poisson simple N d'intensité λ . Posons $X_t = Z_{N_t}$ pour $t \geq 0$. Alors X est un processus de saut markovien de semi-groupe $P_t = \exp(\lambda t(P - I))$:

$$P_t(i,j) = \exp(-\lambda t) \sum_{n=0}^{+\infty} \frac{(\lambda t)^n P^n(i,j)}{n!}.$$

Preuve. Soient $0 = t_0 < t_1 < \dots < t_n$. On a pour $(i_0, \dots, i_n) \in E^{n+1}$,

$$\begin{aligned} & \mathbb{P}(X_{t_0} = i_0, \dots, X_{t_n} = i_n) \\ &= \sum_{h_1, \dots, h_n \in \mathbb{N}} \mathbb{P}(N_{t_1} = h_1, \dots, N_{t_n} = h_1 + \dots + h_n, Z_0 = i_0, Z_{h_1} = i_1, \dots, Z_{h_1 + \dots + h_n} = i_n) \\ &= \sum_{h_1, \dots, h_n \in \mathbb{N}} \mathbb{P}(Z_0 = i_0) P^{h_1}(i_0, i_1) \cdots P^{h_n}(i_{n-1}, i_n) \prod_{i=1}^n \exp(-\lambda(t_i - t_{i-1})) \frac{(\lambda(t_i - t_{i-1}))^{h_i}}{h_i!} \\ &= \mathbb{P}(Z_0 = i_0) P_{t_1}(i_0, i_1) \cdots P_{t_n - t_{n-1}}(i_{n-1}, i_n). \square \end{aligned}$$

Exemple. Le processus de Poisson composé associé à une suite $(X_n)_{n \in \mathbb{N}}$ i.i.d à valeurs dans \mathbb{Z} est de cette forme si on pose

$$Z_n = X_0 + \dots + X_n, \quad n \in \mathbb{N}.$$

4.2.1 La propriété de Markov simple

On notera \mathbb{P}_x la loi du processus de Markov lorsque $\mathbb{P}(X_0 = x) = 1$ pour $x \in E$. \mathbb{P}_x est une mesure sur $(E^{\mathbb{R}_+}, \mathcal{B}(E)^{\otimes \mathbb{R}_+})$. De plus \mathbb{E}_x désignera l'espérance sous \mathbb{P}_x . On se placera souvent sur l'espace dit canonique. Ceci revient à travailler sur $\Omega = \mathbb{R}_+$ avec $\mathbb{P} = \mathbb{P}_x$ et à considérer le processus des coordonnées défini par $X_t(\omega) = \omega_t$ pour $\omega \in E^{\mathbb{R}_+}$ et $t \in \mathbb{R}_+$. Si X est un processus de Markov de sauts défini sur un univers Ω et tel que $\mathbb{P}(X_0 = x) > 0$, alors \mathbb{P}_x correspond à $\mathbb{P}(\cdot | X_0 = x)$. Remarquons que $\mathbb{P}_x(X_t = i) = P_t(x, i)$ pour $i \in E$ et $t > 0$. On a alors pour $t_1 < t_2 < \dots < t_n$ et $(i_1, \dots, i_n) \in E^n$,

$$\mathbb{P}_x(X_{t_1} = i_1, \dots, X_{t_n} = i_n) = P_{t_1}(x, i_1) P_{t_2 - t_1}(i_1, i_2) \times \cdots \times P_{t_n - t_{n-1}}(i_{n-1}, i_n).$$

Proposition 21 Soit $f : E^{\mathbb{R}_+} \rightarrow \mathbb{R}$ une fonction mesurable bornée. Alors,

$$\mathbb{E}_x(f(X_{t+s}, s \geq 0) | \mathcal{F}_t) = \mathbb{E}_{X_t}(f(X_s, s \geq 0)).$$

Preuve. En utilisant un argument de classes monotones, il suffit de prouver le résultat pour

$$f(X_s, s \geq 0) = g(X_{s_1}, \dots, X_{s_n}),$$

pour des indices $0 \leq s_1 < \dots < s_d$. On peut même se contenter de prouver que

$$\mathbb{P}_x(X_{t+s_1} = i_1, \dots, X_{t+s_n} = i_n | X_{u_1} = j_1, \dots, X_{u_k} = j_k, X_t = k) = \mathbb{P}_k(X_{s_1} = i_1, \dots, X_{s_d} = i_d),$$

pour $u_1 < \dots < u_k < t$ et $(j_1, \dots, j_k) \in E^k$. En utilisant la formule des conditionnements successifs, on a

$$\begin{aligned} & \mathbb{P}_x(X_{t+s_1} = i_1, \dots, X_{t+s_n} = i_n | X_{u_1} = j_1, \dots, X_{u_k} = j_k, X_t = k) \\ &= P_{s_d-s_{d-1}}(i_{d-1}, i_d) \times \dots \times P_{s_2-s_1}(i_1, i_2) P_{s_1}(k, i_1) \\ &= \mathbb{P}_k(X_{s_1} = i_1, \dots, X_{s_d} = i_d). \square \end{aligned}$$

4.2.2 La propriété de Markov forte

La propriété de Markov forte généralise la propriété de Markov simple au cas des temps d'arrêt.

Proposition 22 Soit $f : (E^{\mathbb{R}_+}, \mathcal{B}(E)^{\otimes \mathbb{R}_+}) \rightarrow \mathbb{R}$ une fonction mesurable bornée. Alors

$$\mathbb{E}_x(f(X_{s+T}, s \geq 0) \mathbf{1}_{T < +\infty} | \mathcal{F}_T) = \mathbb{E}_{X_T}(f(X_s, s \geq 0)) \mathbf{1}_{T < +\infty}.$$

Preuve. Supposons d'abord que T ne prenne qu'un nombre au plus dénombrable de valeurs $\{t_1, t_2, \dots\} \cup \{+\infty\}$. Soit $A \in \mathcal{F}_T$. On rappelle alors que pour tout $i \in \mathbb{N}^*$, l'événement $A \cap \{T = t_i\}$ appartient à \mathcal{F}_{t_i} . Nous avons en utilisant la propriété de Markov simple :

$$\begin{aligned} \mathbb{E}_x(f(X_{s+T}, s \geq 0) \mathbf{1}_{T < +\infty} \mathbf{1}_A) &= \sum_{i=1}^{+\infty} \mathbb{E}_x(f(X_{s+t_i}, s \geq 0) \mathbf{1}_{A \cap \{T=t_i\}}) \\ &= \sum_{i=1}^{+\infty} \mathbb{E}_x(\mathbb{E}_x(f(X_s, s \geq 0) | \mathcal{F}_{t_i}) \mathbf{1}_{A \cap \{T=t_i\}}) \\ &= \sum_{i=1}^{+\infty} \mathbb{E}_x(\mathbb{E}_{X_{t_i}}(f(X_s, s \geq 0)) \mathbf{1}_{A \cap \{T=t_i\}}) \\ &= \mathbb{E}_x(\mathbb{E}_{X_T}(f(X_s, s \geq 0)) \mathbf{1}_{T < +\infty} \mathbf{1}_A). \end{aligned}$$

Le résultat annoncé suit alors de la caractérisation de l'espérance conditionnelle.

Passons maintenant au cas général. Toujours par un argument de classe monotone, il suffit de considérer des fonctions f telles que

$$f(X_s, s \geq 0) = g(X_{s_1}, \dots, X_{s_d}).$$

Considérons alors pour $n \in \mathbb{N}$, $T_n = \frac{[2^n T] + 1}{2^n}$. Alors $(T_n)_n$ est une suite décroissante de temps d'arrêt qui converge simplement vers T . Remarquons que $\{T = +\infty\} = \{T_n = +\infty\}$. Soit alors $A \in \mathcal{F}_T$. Comme $T_n \geq T$, on a $A \in \mathcal{F}_{T_n}$. On a alors en utilisant la convergence dominée et le fait que pour (t, ω) fixé, $X_{T_n+t}(\omega) = X_{T+t}(\omega)$ lorsque n est assez grand :

$$\begin{aligned} \mathbb{E}_x(g(X_{T+s_1}, \dots, X_{T+s_d}) \mathbf{1}_{T < +\infty} \mathbf{1}_A) &= \lim_{n \rightarrow +\infty} \mathbb{E}_x(g(X_{s_1+T_n}, \dots, X_{s_d+T_n}) \mathbf{1}_A \mathbf{1}_{T_n < +\infty}) \\ &= \lim_{n \rightarrow +\infty} \mathbb{E}_x(\mathbb{E}_{X_{T_n}}(g(X_{s_1}, \dots, X_{s_d})) \mathbf{1}_{T_n < +\infty} \mathbf{1}_A) \\ &= \mathbb{E}_x(\mathbf{1}_A \mathbb{E}_{X_T}(g(X_{s_1}, \dots, X_{s_d})) \mathbf{1}_{T < +\infty}). \end{aligned}$$

Le résultat est encore une fois une conséquence de la caractérisation de l'espérance conditionnelle. \square

4.3 Description dynamique d'un processus de Markov de saut

Soit X un processus de Markov de saut à valeurs dans un ensemble au plus dénombrable E . On définit les temps d'arrêt suivants :

$$T_0 = 0, \quad T_{n+1} = \inf\{t > T_n : X_t \neq X_{T_n}\}, \quad n \in \mathbb{N}.$$

On adopte les conventions suivantes.

- $\inf \emptyset = +\infty$.
- Si $T_n(\omega) = +\infty$ alors $T_{n+j}(\omega) = +\infty$ et $(T_{n+j} - T_{n+j-1})(\omega) = +\infty$ pour tout $j \geq 1$.
- Pour $i \in \mathbb{N}$, on pose $Z_i = X_{T_i}$. Si $T_j(\omega) < +\infty$ et $T_{j+1}(\omega) = +\infty$, on convient que $Z_{j+i}(\omega) = Z_j(\omega)$ pour tout $i \in \mathbb{N}$.

Enfin, pour $(i, j) \in \mathbb{N}^2$, on définit $Q(i, j) = \mathbb{P}_i(Z_1 = j)$ et $\lambda(i) = \frac{1}{\mathbb{E}_i(T_1)}$. Dans la suite, on conviendra qu'une variable aléatoire de loi exponentielle de paramètre 0 est égale à $+\infty$.

Remarque que la continuité à droite des trajectoires assure que $T_{j+1} - T_j > 0$ p.s pour tout $j \in \mathbb{N}$.

Théorème 9 Soit $(X_t)_{t \geq 0}$ un processus de Markov de saut sur E .

1. (Z_n) est une chaîne de Markov de transition Q (on parle de **chaîne immergée**).
2. Conditionnellement aux états Z_0, Z_1, \dots , les variables $T_1, T_2 - T_1, \dots, T_n - T_{n-1}$ sont indépendantes et de loi exponentielle de paramètres respectifs $\lambda(Z_0), \dots, \lambda(Z_{n-1})$. Ceci peut s'écrire

$$\mathbb{P}(T_1 > s_1, \dots, T_n - T_{n-1} > s_n | Z_0 = i_0, Z_1 = i_1, \dots) = \exp(-\lambda(i_0)s_1) \cdots \exp(-\lambda(i_{n-1})s_n). \quad (4.1)$$

Remarques.

1. Si $\lambda(i) = 0$, alors $\mathbb{E}_i(T_1) = +\infty$. D'après le théorème précédente, la loi de $T_1 | Z_0 = i$ suit une loi exponentielle de paramètre 0 ce qui entraîne $\mathbb{P}_i(T_1 = +\infty) = 1$. On en déduit facilement que si le processus passe par l'état i , alors il y reste pour toujours. On dit que i est une **trappe**.
2. On voit que conditionnellement aux états, la loi des n premiers délais entre les sauts ne dépend que des n premiers états. On a donc aussi

$$\mathbb{P}(T_1 > s_1, \dots, T_n - T_{n-1} > s_n | Z_0 = i_0, \dots, Z_{n-1} = i_{n-1}) = \exp(-\lambda(i_0)s_1) \cdots \exp(-\lambda(i_{n-1})s_n).$$

3. On peut reformuler le théorème en écrivant que pour tous $i, i_1, \dots, i_n \in E$ et $s_1, \dots, s_n \in \mathbb{R}_+$,

$$\begin{aligned} & \mathbb{P}_i(T_1 > s_1, Z_1 = i_1, \dots, T_n - T_{n-1} > s_n, Z_n = i_n) \\ &= \exp(-\lambda(i)s_1) Q(i, i_1) \cdots \exp(-\lambda(i_{n-1})s_n) Q(i_{n-1}, i_n). \end{aligned}$$

Preuve. On se place sur l'espace canonique. Les relations suivantes seront utiles (faire un dessin si besoin). Sur l'événement $\{T_n < +\infty\}$, on a

$$T_{n+1} = T_n + T_1 \circ \theta_{T_n}, \quad Z_{n+1} = Z_1 \circ \theta_{T_n}.$$

De plus, sur l'événement $\{T_1 > t\}$, on a

$$Z_1 \circ \theta_t = Z_1, \quad T_1 = t + T_1 \circ \theta_t.$$

La preuve se fait en trois étapes.

1. On commence par déterminer $\mathbb{P}_i(T_1 > t)$ pour $f \in \mathbb{R}_+$ et $i \in E$.
On va utiliser la propriété de Markov simple. On a

$$\begin{aligned}
\mathbb{P}_i(T_1 > t + s) &= \mathbb{P}_i(T_1 > t + s, T_1 > t) \\
&= \mathbb{E}_i(\mathbf{1}_{T_1 > t} \mathbb{E}_i(\mathbf{1}_{T_1 > t+s} | \mathcal{F}_t)) \\
&= \mathbb{E}_i(\mathbf{1}_{T_1 > t} \mathbb{E}_i(\mathbf{1}_{T_1 \circ \theta_t > s} | \mathcal{F}_t)) \\
&= \mathbb{E}_i(\mathbb{E}_{X_t}(\mathbf{1}_{T_1 > s}) \mathbf{1}_{T_1 > t}) \\
&= \mathbb{E}_i(\mathbb{E}_i(\mathbf{1}_{T_1 > s}) \mathbf{1}_{T_1 > t}) \\
&= \mathbb{P}_i(T_1 > t) \mathbb{P}_i(T_1 > s).
\end{aligned}$$

La fonction $f : \mathbb{R}_+ \rightarrow [0, 1]$ définie par $f(t) = \mathbb{P}_i(T > t)$ est continue à droite et vérifie $f(t+s) = f(t)f(s)$. On en déduit facilement que $f(t) = f(1)^t$ pour $t \in \mathbb{R}_+$ avec $0 < f(1) \leq 1$ et donc que T_1 suit une loi exponentielle de paramètre $-\ln f(1)$. Vu la définition de $\lambda(i)$, on a nécessairement $-\ln f(1) = \lambda(i)$. Ainsi T_1 suit une loi exponentielle de paramètre $\lambda(i)$. Au passage on voit que si $\lambda(i) = 0$, $T_1 = +\infty$ p.s.

2. On détermine la loi du couple (T_1, Z_1) sous \mathbb{P}_i .

– Si $\lambda(i) = 0$ alors $\mathbb{P}_i(T_1 = +\infty) = 1$. On a alors :

$$\mathbb{P}_i(T_1 > t, Z_1 = j) = \mathbb{P}_i(Z_1 = j) = Q(i, j) = Q(i, j) \mathbb{P}_i(T_1 > t).$$

Au passage, on peut remarquer que si $\lambda(i) = 0$, les conventions faites avant l'énoncé de ce théorème entraînent que

$$Q(i, j) = \mathbb{P}_i(Z_1 = j) = \mathbb{P}_i(Z_1 = j, T_1 = +\infty) = \mathbb{P}_i(Z_0 = j, T_1 = +\infty) = \mathbb{P}_i(Z_0 = j) = \mathbf{1}_{i=j}.$$

– Si maintenant, $\lambda(i) \neq 0$, alors $\mathbb{P}_i(T_1 < +\infty) = 1$ et $Q(i, i) = 0$. De plus, en utilisant la propriété de Markov simple, on a

$$\begin{aligned}
\mathbb{P}_i(T_1 > t, Z_1 = j) &= \mathbb{P}_i(T_1 > t, Z_1 \circ \theta_t = j) \\
&= \mathbb{E}_i(\mathbf{1}_{T_1 > t} \mathbb{E}_i(\mathbf{1}_{Z_1 \circ \theta_t = j} | \mathcal{F}_t)) \\
&= \mathbb{E}_i(\mathbf{1}_{T_1 > t} \mathbb{E}_{X_t}(\mathbf{1}_{Z_1 = j})) \\
&= \mathbb{E}_i(\mathbf{1}_{T_1 > t} \mathbb{E}_i(\mathbf{1}_{Z_1 = j})) \\
&= Q(i, j) \mathbb{P}_i(T_1 > t) \\
&= Q(i, j) \exp(-\lambda(i)t).
\end{aligned}$$

On retiendra que $\mathbb{P}_i(T_1 > t, Z_1 = j) = Q(i, j) \exp(-\lambda(i)t)$.

3. On détermine maintenant la loi du $2n$ -uplet $(T_1, Z_1, \dots, T_n - T_{n-1}, Z_n)$ sous \mathbb{P}_i . Nous avons déjà déterminé cette loi lorsque $n = 1$. Nous allons montrer que si n est un entier ≥ 2 , $i_1, \dots, i_n \in E$ et $s_1, \dots, s_n \in \mathbb{R}_+$, alors

$$\begin{aligned}
&\mathbb{P}_i(T_1 > s_1, Z_1 = i_1, \dots, T_n - T_{n-1} > s_n, Z_n = i_n) \\
&= \mathbb{P}_i(T_1 > s_1, Z_1 = i_1, \dots, T_{n-1} - T_{n-2} > s_{n-1}, Z_{n-1} = i_{n-1}) \exp(-\lambda(i_{n-1})s_n) Q(i_{n-1}, i_n).
\end{aligned} \tag{4.2}$$

Ainsi nous obtiendrons alors facilement la loi du $2n$ -uplet en itérant et (4.1) sera prouvée.

– Étudions tout d'abord le cas où les valeurs i_1, \dots, i_n de Z_1, \dots, Z_n vérifient $i_n \neq i_{n-1}$ et Ainsi, on a l'inclusion $\{Z_1 = i_1, \dots, Z_n = i_n\} \subset \{T_{n-1} < +\infty\}$. Nous allons appliquer la propriété de Markov forte. Par souci de simplicité, l'indicateur $\mathbf{1}_{T_{n-1} < +\infty}$ sera omis

(puisque sa valeur est toujours 1 sur l'événement considéré).

$$\begin{aligned}
& \mathbb{P}_i(T_1 > s_1, Z_1 = i_1, \dots, T_n - T_{n-1} > s_n, Z_n = i_n) \\
&= \mathbb{E}_i \left(\mathbb{1}_{T_1 > s_1, Z_1 = i_1, \dots, T_{n-1} - T_{n-2} > s_{n-1}} \mathbb{E}_i \left(\mathbb{1}_{T_1 \circ \theta_{T_{n-1}} > s_n} \mathbb{1}_{Z_1 \circ \theta_{T_{n-1}} = i_n} \middle| \mathcal{F}_{T_{n-1}} \right) \right) \\
&= \mathbb{E}_i \left(\mathbb{1}_{T_1 > s_1, Z_1 = i_1, \dots, T_{n-1} - T_{n-2} > s_{n-1}} \mathbb{E}_{Z_{n-1}} \left(\mathbb{1}_{T_1 > s_n} \mathbb{1}_{Z_1 = i_n} \right) \right) \\
&= \mathbb{P}_i(T_1 > s_1, Z_1 = i_1, \dots, T_{n-1} - T_{n-2} > s_{n-1}) Q(i_{n-1}, i_n) \exp(-\lambda(i_{n-1})s_n).
\end{aligned}$$

– Considérons maintenant des valeurs i, i_1, \dots, i_n telles que il existe un entier $m \in \{1, \dots, n-1\}$ tel que $i_{m-1} \neq i_m = i_{m+1} = \dots = i_n$ (on convient que $i_0 = i$). Ainsi on a l'inclusion

$$\{Z_1 = i_1, Z_2 = i_2, \dots, Z_n = i_n\} \subset \{T_m < +\infty\}.$$

En appliquant la propriété de Markov forte, on a

$$\begin{aligned}
& \mathbb{P}_i(T_1 > s_1, Z_1 = i_1, \dots, T_n - T_{n-1} > s_n, Z_n = i_n) \\
&= \mathbb{P}_i(T_1 > s_1, Z_1 = i_1, \dots, T_m - T_{m-1} > s_m, Z_m = i_m, T_{m+1} = +\infty) \\
&= \mathbb{P}_i(T_1 > s_1, \dots, Z_m = i_m, T_{m+1} - T_m = +\infty) \\
&= \mathbb{P}_i(T_1 > s_1, \dots, Z_m = i_m, T_1 \circ \theta_{T_m} = +\infty) \\
&= \mathbb{P}_i(T_1 > s_1, \dots, Z_m = i_m) \mathbb{P}_{i_m}(T_1 = +\infty).
\end{aligned}$$

Supposons d'abord $\lambda(i_m) = 0$. Alors $\mathbb{P}_{i_m}(T_1 = +\infty) = \mathbb{P}_{i_{n-1}}(T_1 = +\infty) = 1 = \exp(-\lambda(i_{n-1}))$ et $Q(i_{n-1}, i_n) = 1$. On obtient alors (4.2). Supposons ensuite que $\lambda(i_m) > 0$. Alors $\mathbb{P}_{i_m}(T_1 = +\infty) = 0$. Or, dans la deuxième partie de la preuve, on a obtenu que $Q(i_m, i_m) = 0$ et donc $Q(i_{n-1}, i_n) = 0$. L'égalité (4.2) est encore valable.

– Supposons enfin que $i = i_1 = \dots = i_n$. Si $\lambda(i) = 0$, alors $Q(i, i) = 1$ et on obtient d'après les conventions :

$$\mathbb{P}_i(T_1 > s_1, Z_1 = i_1, \dots, T_n - T_{n-1} > s_n, Z_n = i_n) = 1,$$

ce qui prouve encore (4.2). Si $\lambda(i) > 0$, $Q(i, i) = 0$ et (4.2) est toujours valable (on obtient 0 des deux côtés de l'équation).□

Pour un processus de Markov à saut, il peut se produire une accumulation de sauts dans un intervalle de temps fini. Ce phénomène correspond au cas où $\mathbb{P}(\lim_{n \rightarrow +\infty} T_n < +\infty) > 0$. On dit alors qu'il y a **explosion**. Ce cas de figure ne se produira pas sur les modèles que nous étudierons et rend l'étude des processus à saut plus compliquée. Nous donnons dans la suite deux conditions sous lesquelles l'explosion n'a pas lieu.

Proposition 23 *Les deux cas suivants garantissent qu'il n'y a pas explosion.*

1. $\sup_{i \in E} \lambda(i) < +\infty$ (condition vérifiée en particulier si E est fini).
2. La chaîne de Markov (Z_n) est récurrente.

Remarquer qu'on est aussi assuré de la non-explosion si les trajectoires $t \rightarrow X_t(\omega)$ ont des limites à gauche. On retiendra en particulier qu'il n'y a pas explosion lorsque E est fini.

Preuve. Nous allons utiliser le lemme suivant.

Lemme 4 *Soit $(U_n)_{n \geq 1}$ une suite de variables aléatoires indépendantes toutes de loi exponentielle et dont les paramètres sont respectivement $\lambda_1, \lambda_2, \dots$.*

1. Si $\sum_{k=1}^{+\infty} \frac{1}{\lambda_k} = +\infty$ alors $\mathbb{P}\left(\sum_{k=0}^{+\infty} U_k = +\infty\right) = 1$.
2. Si $\sum_{k=1}^{+\infty} \frac{1}{\lambda_k} < +\infty$ alors $\mathbb{P}\left(\sum_{k=0}^{+\infty} U_k < +\infty\right) = 1$.

Prouvons d'abord ce lemme. Dans le premier cas de figure, on a

$$\mathbb{E} \left(\sum_{k=1}^{+\infty} U_k \right) = \sum_{k=1}^{+\infty} \frac{1}{\lambda_k} < +\infty.$$

D'où $\sum_{k=1}^{+\infty} U_k < +\infty$ p.s.

Dans le deuxième cas de figure, on a

$$\begin{aligned} \mathbb{E} \left(\exp \left(- \sum_{k=1}^{+\infty} U_k \right) \right) &= \lim_{N \rightarrow +\infty} \mathbb{E} \left(\exp \left(- \sum_{k=1}^N U_k \right) \right) \\ &= \lim_{N \rightarrow +\infty} \prod_{k=1}^N \frac{1}{1 + \lambda_k^{-1}}. \end{aligned}$$

Il est facile de prouver que la dernière limite vaut 0 en passant au logarithme et en utilisant un développement limité (exercice). On en déduit que $\mathbb{E} \left(\exp \left(- \sum_{k=1}^{+\infty} U_k \right) \right) = 0$ et donc que

$$\sum_{k=1}^{+\infty} U_k = +\infty \text{ p.s.}$$

Passons maintenant à la preuve de la proposition. Nous avons

$$\begin{aligned} \mathbb{P} \left(\lim_{n \rightarrow +\infty} T_n = +\infty \right) &= \mathbb{P} \left(\sup_{n \geq 1} T_n < +\infty \right) \\ &= \mathbb{P} \left(\sum_{i=0}^{+\infty} (T_{i+1} - T_i) < +\infty \right) \\ &= \mathbb{E} \left(\mathbb{E} \left(\mathbf{1}_{\sum_{k=0}^{+\infty} T_{k+1} - T_k < +\infty} \mid \sigma(Z_0, Z_1, \dots) \right) \right) \\ &= \mathbb{P} \left(\sum_{k=1}^{+\infty} \frac{1}{\lambda(Z_{k-1})} < +\infty \right). \end{aligned}$$

1. Supposons tout d'abord que $A = \sup_{i \in E} \lambda(i) < +\infty$. Alors pour tout $i \in E$, $\frac{1}{\lambda(i)} \geq \frac{1}{A}$ et donc on a p.s

$$\sum_{k=1}^{+\infty} \frac{1}{\lambda(Z_{k-1})} = +\infty. \tag{4.3}$$

Ainsi $\mathbb{P}(\lim_{n \rightarrow +\infty} T_n = +\infty) = 1$.

2. Si la chaîne immergée est récurrente, toute trajectoire passe une infinité de fois par chaque point. On a donc encore (4.3) et la non-explosion est immédiate. \square

Dorénavant, nous supposons que le processus est non explosif.

4.4 Processus minimal et simulation d'un processus de Markov de saut

On conserve les notations de la section précédente. Pour simplifier les formulations, nous supposons dans la proposition suivante que $\lambda(i) > 0$ pour tout $i \in E$. (il n'y a pas de trappe).

Proposition 24 Soit X un processus de Markov à saut. Soient (Z_n) une chaîne de Markov de transition Q et (U_k) une suite i.i.d de variables aléatoires toutes de loi exponentielle de paramètre 1 (que l'on supposera indépendante de (Z_n)). Alors le processus \tilde{X} défini par

$$\tilde{X}_t = Z_n, \quad \text{si} \quad \sum_{k=1}^n \frac{U_k}{Z_{k-1}} \leq t < \sum_{k=1}^{n+1} \frac{U_k}{Z_{k-1}}, \quad n \in \mathbb{N},$$

est un processus de Markov de saut de même loi que X .

Preuve. Comme il n'y a pas explosion, on a $\lim_{n \rightarrow +\infty} T_n = +\infty$ p.s. De plus X vérifie

$$X_t = Z_n, \quad \text{si} \quad T_n \leq t < T_{n+1}, \quad n \in \mathbb{N}$$

et les variables aléatoires $U_n = \lambda(Z_{n-1})(T_n - T_{n-1})$, $n \geq 1$ sont bien i.i.d et toutes de loi exponentielle de paramètre 1 d'après les résultats de la section précédente. \square

Il reste maintenant à construire des processus de Markov de saut. Plus exactement, étant données une famille $\{\lambda(i) : i \in E\}$ de nombres réels positifs ainsi qu'une matrice de transition Q sur E vérifiant $Q(i, i) = 0$ si $\lambda(i) > 0$ et $Q(i, i) = 1$ si $\lambda(i) = 0$, peut-on construire un processus de Markov de saut X tel que $\lambda(i) = \frac{1}{\mathbb{E}_i(T_1)}$ et $Q(i, j) = \mathbb{P}_i(Z_1 = j)$ pour $(i, j) \in E^2$?

La proposition précédente nous incite à procéder ainsi :

- On commence par simuler une chaîne de Markov de transition Q (la chaîne immergée (Z_n)).
- Les états étant donnés par la suite (Z_n) , on simule conditionnellement à ces états une suite de variables indépendantes, toutes de loi exponentielle et dont les paramètres sont respectivement $\lambda(Z_0), \lambda(Z_1), \dots$. On obtient ainsi les délais $T_{i+1} - T_i$ entre les sauts ($i \geq 0$).

Il resterait néanmoins à vérifier que le processus ainsi construit vérifie bien la propriété de Markov. Ceci peut se justifier au moyen de calculs un peu laborieux. Nous admettrons le résultat. Lorsque $\sup_{i \in E} \lambda(i) < +\infty$, nous verrons un peu plus loin la possibilité de construire le processus à partir d'un processus de Poisson.

La construction évoquée ci-dessus offre aussi une méthode de simulation du processus. Nous donnons une deuxième méthode de simulation en simulant directement les couples $(T_{n+1} - T_n, Z_{n+1})$ conditionnellement à $Z_n = i$ avec $\lambda(i) > 0$ (sinon i est une trappe et on reste dans cet état pour toujours). Lorsque le processus est en l'état i , on simule des variables aléatoires indépendantes $(Y_k)_{k \neq i, k \in E}$ toutes de loi exponentielle et dont les paramètres sont donnés est $A(i, k) = \lambda(i)Q(i, k)$. La valeur de la plus petite de ces variables aléatoires donnera le nouveau délai avant le prochain saut et son indice sera la valeur de l'état suivant $Z_{n+1}(\omega)$. Avant de voir pourquoi tout ceci fonctionne, il faut remarquer que le minimum $\min_{k \neq i, j} Y_k$ suit une loi exponentielle de paramètre $\sum_{k \neq i, j} A(i, k) = \lambda(i) - A(i, j)$. On a alors

$$\begin{aligned} & \mathbb{P}(Z_{n+1} = j, T_{n+1} - T_n > t | Z_n = i) \\ &= \mathbb{P}\left(t < Y_j < \min_{k \neq i, j} Y_k\right) \\ &= \int \mathbb{1}_{0 \leq t < y < z} A(i, j) \exp(-A(i, j)y) (\lambda(i) - A(i, j)) \exp(-(-\lambda(i) - A(i, j))y) dy dz \\ &= Q(i, j) \exp(-\lambda(i)t). \end{aligned}$$

Ce procédé permet donc bien de simuler le processus.

4.5 Générateur et taux de transition

Définition 17 On appelle générateur la matrice $(A(i, j))_{(i, j) \in E^2}$ définie par

$$A(i, j) = \begin{cases} \lambda(i)Q(i, j) & \text{si } i \neq j, \\ -\lambda(i) & \text{si } i = j \end{cases}$$

Remarque. On a toujours $\sum_{j \in E} A(i, j) = 0$.

Théorème 10 Le semi-groupe $\{P_t, t \in \mathbb{R}_+\}$ vérifie pour $t \geq 0$:

$$\frac{dP_t}{dt} AP_t \quad (\text{équation dite backward}).$$

En particulier, nous avons

$$\begin{cases} P_t(i, j) = A(i, j)t + o(t) & \text{si } i \neq j, \\ P_t(i, i) = 1 + A(i, i)t + o(t) & \text{si } i = j \end{cases}$$

Preuve. La preuve va utiliser le lemme suivant.

Lemme 5

$$P_t(i, j) = \exp(-\lambda(i)t) \mathbf{1}_{i=j} + \sum_{k \neq i} \int_0^t \exp(-\lambda(i)(t-s)) A(i, k) P_s(k, j) ds.$$

Voyons comment ce lemme permet d'obtenir le résultat. Observons tout d'abord que

$$\sum_{k \neq i} A(i, k) P_s(k, j) \leq \sum_{k \neq i} \lambda(i)Q(i, k) \leq \lambda(i).$$

En utilisant le théorème de convergence dominée, le Lemme 5 permet d'en déduire que la fonction (en s) sous le signe intégral est continue. Ainsi $t \rightarrow P_t(i, j)$ est dérivable et

$$\begin{aligned} P'_t(i, j) &= -\lambda(i)P_t(i, j) + \exp(-\lambda(i)t) \exp(-\lambda(i)t) \sum_{k \neq i} A(i, k) P_t(k, j) \\ &= (AP_t)(i, j). \end{aligned}$$

Les développements limités en $t = 0$ s'obtiennent alors aisément. \square

Preuve du Lemme 5. On a

$$\begin{aligned} P_t(i, j) &= \mathbb{P}_i(X_t = j, T_1 > t) + \mathbb{P}_i(X_t = j, T_1 \leq t) \\ &= \mathbf{1}_{i=j} \mathbb{P}_i(T_1 > t) + \mathbb{P}_i(X_t = j, T_1 \leq t) \\ &= \mathbf{1}_{i=j} \exp(-\lambda(i)t) + \mathbb{P}_i(X_t = j, T_1 \leq t). \end{aligned}$$

Examinons

$$\mathbb{P}_i(X_t = j, T_1 \leq t) = \sum_{k \neq i} \mathbb{P}_i(Z_1 = k, T_1 \leq t, X_t = j).$$

On remarque que

$$\mathbb{P}_i(Z_1 = k, T_1 \leq t, X_t = j) = \mathbb{E}_i(\mathbf{1}_{Z_1=k} \mathbb{E}_i(g(T_1, X \circ \theta_{T_1}) \mathcal{F}_{T_1})),$$

où $g(s, X \circ \theta_{T_1}) = \mathbb{1}_{s \leq t} \mathbb{1}_{X_{t-s} \circ \theta_{T_1} = j}$. La propriété de Markov forte assure que

$$\begin{aligned} \mathbb{E}_i(g(s, X \circ \theta_{T_1}) | \mathcal{F}_{T_1}) &= \mathbb{E}_{Z_1}(\mathbb{1}_{X_{t-s} = j}) \mathbb{1}_{s \leq t} \\ &= P_{t-s}(Z_1, j) \mathbb{1}_{s \leq t}. \end{aligned}$$

Mais notons qu'ici, on n'est pas tout à fait dans le cadre de l'application de la propriété de Markov forte car s est remplacé par la variables aléatoire T_1 . Mais comme T_1 est \mathcal{F}_{T_1} -mesurable (et joue donc le rôle d'une constante dans le conditionnement), nous admettrons qu'une version plus générale de la propriété de Markov forte assure que :

$$\mathbb{E}_i(g(T_1, X \circ \theta_{T_1}) | \mathcal{F}_{T_1}) = P_{t-T_1}(Z_1, j) \mathbb{1}_{T_1 \leq t}.$$

La fin de la démonstration utilise l'expression de la loi du couple (T_1, Z_1) . Nous avons

$$\begin{aligned} \mathbb{P}_i(Z_1 = k, T_1 \leq t, X_t = j) &= \mathbb{E}_i(\mathbb{1}_{Z_1 = k} P_{t-T_1}(k, j) \mathbb{1}_{T_1 \leq t}) \\ &= \int_0^t \exp(-\lambda(i)s) P_{t-s}(k, j) \lambda(i) Q(i, k) ds \\ &= \int_0^t \exp(-\lambda(i)s) A(i, k) P_s(k, j) ds. \end{aligned}$$

Le lemme est ainsi prouvé. \square

Proposition 25 *On suppose $M = \sup_{i \in E} \lambda(i) < +\infty$. Alors*

$$P_t = \exp(tA) = \sum_{n=0}^{+\infty} \frac{t^n A^n}{n!}.$$

Preuve. On commence par remarquer que $Q_t = \exp(tA)$ est bien définie. En effet,

$$\sum_{j \in E} |A(i, j)| \leq \lambda(i) + \sum_{j \neq i} \lambda(i) Q(i, j) = 2\lambda(i),$$

ce qui entraîne

$$|A^n(i, j)| \leq (2M)^n.$$

On en déduit que $t \rightarrow Q_t$ est bien définie, qu'elle est dérivable et que $Q'_t = A Q_t$. La dérivée de $t \rightarrow Q_{-t} P_t$ est nulle ce qui entraîne $\exp(-tA) P_t = I$ (matrice identité correspondante au cas $t = 0$). On en déduit que $P_t = \exp(tA)$. \square

Corollaire 1 *Soit A une matrice sur $E \times E$ tel que $A(i, j) \geq 0$ et $\sum_{j \in E} A(i, j) = 0$ pour tout $i \in E$. Si $M = \sup_{i \in E} -A(i, i) < +\infty$ alors $T = \frac{A}{M} + I$ est une probabilité de transition. Dans ce cas, si (W_n) est une chaîne de Markov de transition T , alors le processus de X défini par $X_t = W_{N_t}$ est un processus de Markov de saut de générateur A .*

Preuve. Remarquons que $T(i, j) = \frac{A(i, j)}{M} \geq 0$ si $i \neq j$ et $T(i, i) = \frac{A(i, i)}{M} + 1 \geq 0$. Ainsi $T(i, j) \geq 0$. De plus,

$$\sum_{j \in E} T(i, j) = \frac{\sum_{j \neq i} A(i, j) + A(i, i)}{M} + 1 = 1.$$

T est donc bien une matrice de transition. D'après la Proposition 20, X est alors un processus de Markov tel que

$$P_t = \exp(Mt(T - I)) = \exp(tA). \square$$

Remarques

1. Si $\sup_{i \in E} -A(i, i) = +\infty$, il est possible de construire un processus de générateur A en considérant le processus minimal.
2. On peut remarquer que les sauts du processus ne coïncident pas avec les sauts du processus de Poisson N . En fait N contient plus de sauts que X : N peut avoir un saut alors que deux valeurs consécutives de la chaîne W peuvent coïncider.
3. Le semi-groupe $\{P_t : t \in \mathbb{R}_+\}$ n'est pas toujours calculable (si E est fini, on peut calculer l'exponentielle de matrice $\exp(tA)$).
4. En général, on définit le processus de Markov directement à l'aide d'un développement limité en utilisant les nombres $A(i, j)$ qui s'interprètent comme des taux de transition. Si $i \neq j$:

$$P_t(i, j) = A(i, j)t + o(t).$$

En pratique les nombres $A(i, j)$ ont une signification concrète.

4.6 L'estimateur du maximum de vraisemblance

4.6.1 Calcul informel de la vraisemblance

On suppose qu'une trajectoire de X est observée sur un intervalle $[0, t]$ avec $t > 0$. On suppose que le générateur $A = A_\theta$ est paramétré par un paramètre $\theta \in \mathbb{R}^d$. Comme nous l'avons déjà fait pour le processus de Poisson, nous allons étudier l'information disponible c'est à dire

- le nombre de sauts $N_t(\omega) = n \in \mathbb{N}$,
- les délais entre les sauts $S_1 = T_1, \dots, S_n = T_n - T_{n-1}$,
- les différents états du processus $Z_0(\omega), \dots, Z_n(\omega)$.

Intuitivement, le calcul du maximum de vraisemblance nécessitera le calcul des probabilités du type

$$\mathbb{P}(N_t = n, S_1 \in B_1, \dots, S_n \in B_n, Z_0 = i_0, \dots, Z_n = i_n),$$

où B_1, \dots, B_n sont des boréliens de \mathbb{R}_+ et i_0, \dots, i_n des points de E . Nous avons

$$\begin{aligned} & \mathbb{P}(N_t = n, S_1 \in B_1, \dots, S_n \in B_n, Z_0 = i_0, \dots, Z_n = i_n) \\ &= \mathbb{P}(T_{n+1} > t, T_n \leq t, S_1 \in B_1, \dots, S_n \in B_n, Z_0 = i_0, \dots, Z_n = i_n) \\ &= \mathbb{P}(Z_0 = i_0) \mathbb{P}_{i_0}(T_{n+1} > t, T_n \leq t, S_1 \in B_1, \dots, S_n \in B_n, Z_1 = i_1, \dots, Z_n = i_n) \\ &= \mathbb{P}(Z_0 = i_0) \mathbb{P}_{i_0}(T_{n+1} > t, T_n \leq t, S_1 \in B_1, \dots, S_n \in B_n | Z_1 = i_1, \dots, Z_n = i_n) \\ &\times Q_\theta(i_0, i_1) \cdots Q_\theta(i_{n-1}, i_n) \\ &= \mathbb{P}(Z_0 = i_0) Q_\theta(i_0, i_1) \cdots Q_\theta(i_{n-1}, i_n) \\ &\times \int_{B_1 \times \dots \times B_n} \exp\left(-\lambda_\theta(i_n)\left(t - \sum_{j=1}^n s_j\right)\right) \mathbb{1}_{\sum_{j=1}^n s_j \leq t} \prod_{\ell=0}^{n-1} A_\theta(i_\ell, i_{\ell+1}) \exp(-\lambda_\theta(i_\ell)s_{\ell+1}) ds_1 \cdots ds_n. \end{aligned}$$

On définit la vraisemblance par

$$\begin{aligned} L_\theta^{(n)}(s_1, \dots, s_n, i_1, \dots, i_n, n) &= \prod_{\ell=0}^{n-1} A_\theta(i_\ell, i_{\ell+1}) \exp(s_{\ell+1}(A_\theta(i_\ell, i_\ell) - A_\theta(i_n, i_n))) \\ &\times \exp(A_\theta(i_n, i_n)t) \mathbb{P}(Z_0 = i_0). \end{aligned}$$

On peut alors réécrire la log-vraisemblance sous la forme

$$\begin{aligned} & \ln L_\theta^{(n)}(s_1, \dots, s_n, i_1, \dots, i_n, n) \\ = & \sum_{i,j \in E} N_{i,j}^{(t)} \ln(A_\theta(i, j)) + \sum_{i \in E} (A_\theta(i, i) - A_\theta(i_n, i_n)) R_i^{(t)} + tA_\theta(i_n, i_n) + \ln(\mathbb{P}(Z_0 = i_0)), \end{aligned}$$

où $N_{i,j}^{(t)}$ désigne le nombre de transitions de i vers j avant t et $R_i^{(t)}$ désigne le temps passé en i dans $[0, t]$.

4.6.2 Définition de la mesure dominante

On notera par μ la mesure de comptage sur E et par λ la mesure de Lebesgue sur \mathbb{R} . On a vu que l'information disponible lorsqu'on observe $(X_s)_{0 \leq s \leq t}$ était complètement résumée par les variables aléatoires $(Z_0, Z_1, S_1, \dots, Z_{n-1}, S_{n-1}, Z_n)$ lorsque $N_t = n$. Le nombre de variables utilisées est donc également aléatoire et nous allons définir un espace mesurable adapté à cette situation. Plus précisément, on définit pour tout entier n non nul, l'espace produit

$$W_n = E \times \prod_{i=1}^n (E \times \mathbb{R}),$$

que l'on munit de la tribu produit notée \mathcal{G}_n (E étant muni de l'ensemble de ses parties et \mathbb{R} de la tribu des boréliens) ainsi que de la mesure produit

$$\sigma^{(n)} = \mu \otimes (\otimes_{i=1}^n \mu \otimes \lambda).$$

Pour $n = 0$, on pose $W_0 = E$ et on munit cet ensemble de la tribu des parties de E ainsi que de la mesure μ . On pose enfin $W = \cup_{n \in \mathbb{N}} W_n$ qui sera muni de la tribu

$$\mathcal{G} = \{A \subset W : \forall n \in \mathbb{N}, A \cap W_n \in \mathcal{G}_n\}$$

et de la mesure σ définie par

$$\sigma(A) = \sum_{n=0}^{+\infty} \sigma^{(n)}(A \cap W_n), \quad A \in \mathcal{G}.$$

L'application $Y : (\Omega, \mathcal{A}, \mathbb{P}) \rightarrow (W, \mathcal{G})$ définie par

$$Y(\omega) = (Z_0(\omega), Z_1(\omega), S_1(\omega), \dots, Z_n(\omega), S_n(\omega)) \quad \text{si } N_t(\omega) = n,$$

est alors une variable aléatoire. De plus, il existe une bijection f telle que $(X_s)_{0 \leq s \leq t} = f(Y)$. D'après une remarque énoncée à la fin du premier chapitre, il suffit de maximiser la vraisemblance

associée à Y pour déterminer l'EMV. Nous avons

$$\begin{aligned}
\mathbb{P}(Y \in A) &= \sum_{n \in \mathbb{N}} \mathbb{P}(Y \in A \cap W_n) \\
&= \sum_{n=0}^{+\infty} \mathbb{P}((Z_0, \dots, S_n) \in A \cap W_n, N_t = n) \\
&= \sum_{n=0}^{+\infty} \int_{A \cap W_n} L_\theta^{(n)}(\cdot, n) d\sigma^{(n)} \\
&= \sum_{n=0}^{+\infty} \int_A L_\theta^{(n)}(\cdot, n) \mathbb{1}_{W_n} d\sigma^{(n)} \\
&= \sum_{n=0}^{+\infty} \int_A g_\theta \mathbb{1}_{W_n} d\sigma^{(n)} \\
&= \int_A g_\theta d\sigma,
\end{aligned}$$

où l'on a posé

$$g_\theta = \sum_{k \in \mathbb{N}} L_\theta^{(k)}(\cdot, k) \mathbb{1}_{W_k}.$$

La vraisemblance est donc la fonction $\theta \rightarrow g_\theta(Y)$ et correspond bien au calcul qui a été effectué précédemment.

4.7 Quelques exemples

1. Un exemple simple en fiabilité. On suppose $E = \{0, 1\}$ et on pose

$$A = \begin{pmatrix} -\mu & \mu \\ \lambda & -\lambda \end{pmatrix}.$$

Ainsi $\lambda(0) = \mu$, $\lambda(1) = \lambda$ et $Q(0, 1) = Q(1, 0) = 1$. Le processus X associé correspond à l'état d'une machine au cours du temps. Si $X_t = 1$, la machine est en état de marche et si $X_t = 0$, la machine est en état de panne. On voit alors que les temps de fonctionnement sont i.i.d de loi exponentielle de paramètre μ et sont aussi indépendants des temps de panne qui sont i.i.d de loi exponentielle de paramètre λ (remarquer que sous \mathbb{P}_i la suite des états est déterministe, $i = 0, 1$). A titre d'exercice, on pourra vérifier que

$$P_t = \exp(tA) = \begin{pmatrix} \frac{\lambda}{\lambda+\mu} + \frac{\mu}{\lambda+\mu} \exp(-(\lambda+\mu)t) & \frac{\mu}{\lambda+\mu} - \frac{\mu}{\lambda+\mu} \exp(-(\lambda+\mu)t) \\ \frac{\lambda}{\lambda+\mu} - \frac{\lambda}{\lambda+\mu} \exp(-(\lambda+\mu)t) & \frac{\mu}{\lambda+\mu} + \frac{\lambda}{\lambda+\mu} \exp(-(\lambda+\mu)t) \end{pmatrix}.$$

On pourra aussi déterminer l'EMV du couple (μ, λ) .

2. Voici un exemple d'un processus à 5 états qui a été utilisé pour étudier l'évolution du SIDA. On considère un malade dont les 5 états peuvent être 1 infecté sans anticorps, 2 avec anticorps sans symptômes, 3 avec premiers symptômes, 4 état cliniquement avancé, 5 mort. Le générateur est donné par

$$A = \begin{pmatrix} -\alpha & \alpha & 0 & 0 & 0 \\ 0 & -\beta & \beta & 0 & 0 \\ 0 & 0 & -\gamma & \gamma & 0 \\ 0 & 0 & 0 & -\delta & \delta \\ 0 & 0 & 0 & 0 & 0 \end{pmatrix}.$$

En général, pour ce type de problème, on dispose de N trajectoires associées à N individus infectés. Si les états de ces individus sont indépendants, on peut calculer la vraisemblance en effectuant le produit des vraisemblances des N individus. L'asymptotique des estimateurs peut alors être étudiée lorsque $N \rightarrow +\infty$.

3. Processus de naissance et de mort. On suppose que X_t désigne le nombre d'individus dans une population à l'instant t . Le processus X est supposé markovien avec si $n \in \mathbb{N}^*$:

$$P_t(n, n+1) = \beta_n t + o(t), \quad P_t(n, n-1) = \alpha_n t + o(t).$$

On a alors $A(n, n-1) = \alpha_n$, $\beta_n = A(n, n+1)$, $\lambda(n) = \alpha_n + \beta_n$, $Q(n, n-1) = \frac{\alpha_n}{\alpha_n + \beta_n}$ et $Q(n, n+1) = \frac{\beta_n}{\alpha_n + \beta_n}$. L'état $n = 0$ est une trappe.

Par exemple, lorsque $\alpha_n = n\alpha$ et $\beta_n = n\beta$, α s'interprète comme le taux de mortalité et β comme le taux de natalité. On peut justifier le caractère non-explosif. Si $Z_0 = i$, alors $Z_k \leq k + i$ et $\lambda(Z_k) \leq (\alpha + \beta)(k + i)$ ce qui entraîne

$$\sum_{k=1}^{+\infty} \frac{1}{\lambda(Z_k)} = +\infty.$$

Cette condition entraîne la non-explosion comme le montre la preuve de la Proposition (23).

4.8 Exercices

EXERCICE 11 On considère une population de m individus. Au temps $t = 0$, on suppose qu'un individu vient d'être contaminé; les autres individus sont susceptibles d'être contaminé. De plus l'infection est supposée non curable. On admet que dans chaque intervalle de temps de longueur h , toute personne infectée peut contaminer une personne saine avec probabilité $\alpha h + o(h)$, $h \rightarrow 0^+$. Soient X_t le nombre d'individus infectés à l'instant $t \geq 0$ et T_i le temps nécessaire pour passer de i individus infectés à $i + 1$, pour $1 \leq i \leq m - 1$.

1. Montrer que X est un processus de Markov dont le générateur est donnée par

$$a_{i,j} = \begin{cases} (m-i)\alpha & \text{si } j = i+1, 1 \leq i \leq m-1, \\ -(m-i)\alpha & \text{si } i = j \in \{1, \dots, m-1\}, \\ 0, & \text{sinon.} \end{cases}$$

2. Soit T le temps nécessaire pour que toute la population soit contaminée. Calculer $\mathbb{E}(T)$ et $Var(T)$ (on utilisera les délais S_i entre les sauts ainsi que leur loi).

EXERCICE 12 *processus du télégraphe*

Soit Y une variable aléatoire à valeurs dans $\{-1, 1\}$, indépendante d'un processus de Poisson simple N d'intensité λ . On définit alors le processus X par

$$X_t = Y \times (-1)^{N_t}, \quad t \in \mathbb{R}_+.$$

Montrer que X est un processus de Markov à sauts dont les transitions sont données par

$$p_t(1, 1) = p_t(-1, -1) = e^{-\lambda t} \sum_{n=0}^{+\infty} \frac{(\lambda t)^{2n}}{(2n)!},$$

$$p_t(-1, 1) = p_t(1, -1) = e^{-\lambda t} \sum_{n=0}^{+\infty} \frac{(\lambda t)^{2n+1}}{(2n+1)!}.$$

Quel est alors le générateur A de X ?

Comment obtenir facilement l'estimateur du maximum de vraisemblance de λ si on dispose des observations $\{X_s/0 \leq s \leq t\}$?

EXERCICE 13 Soit X un processus de Markov de sauts à trois états et dont le générateur est donné par

$$A = \begin{pmatrix} -\lambda & \lambda & 0 \\ 0 & -\lambda & \lambda \\ \lambda & 0 & -\lambda \end{pmatrix}.$$

Donner l'expression du l'estimateur du maximum de vraisemblance du paramètre λ lorsque on suppose observé une trajectoire sur l'intervalle $[0, t]$.

EXERCICE 14 Soit X un processus de Markov de sauts à valeurs dans un espace E au plus dénombrable. En reprenant les notations du cours, on suppose que le générateur A vérifie $\sup_{i \in E} |A(i, i)| < +\infty$. Soit $B_t = \{T_2 \leq t\}$ l'événement "au moins deux sauts se produisent dans l'intervalle de temps $[0, t]$ ". Montrer que

$$\limsup_{t \rightarrow 0} \sup_{i \in E} \frac{\mathbb{P}_i(B_t)}{t} = 0.$$

Indication : on pourra utiliser pour $t \in \mathbb{R}_+$, l'inégalité $\mathbb{P}(X + Y \leq t) \leq \mathbb{P}(X \leq t) \cdot \mathbb{P}(Y \leq t)$ valable pour deux variables aléatoires X et Y indépendantes et à valeurs positives. De plus on rappelle que $1 - \exp(x) \leq -x$ pour $x \in \mathbb{R}$.

Chapitre 5

Intégrale stochastique et modèles de diffusion

L'objectif de ce chapitre est de présenter brièvement les notions de martingale à temps continu, d'intégrale stochastique et d'équation différentielle stochastique. Il ne sera pas question ici de présenter ces outils de la manière la plus générale et de mener une étude sérieuse de leurs propriétés mathématiques (se reporter au cours de J.C Breton [http : //perso.univ – rennes1.fr/jean – christophe.breton/Fichiers/processusM2.pdf](http://perso.univ-rennes1.fr/jean-christophe.breton/Fichiers/processusM2.pdf) pour un cours détaillé sur ce sujet ainsi que pour les preuves de certaines propriétés de ce chapitre).

5.1 Martingales à temps continu

5.1.1 Définition et exemples

Définition 18 *Un processus $(M_t)_{t \geq 0}$ est une martingale (resp. sous-martingale, resp. sur-martingale) pour la filtration $(\mathcal{F}_t)_{t \geq 0}$ si*

1. M_t est \mathcal{F}_t -mesurable (i.e processus adapté).
2. $\mathbb{E}(|M_t|) < +\infty, \forall t \geq 0$.
3. $\forall s \leq t, \mathbb{E}(M_t | \mathcal{F}_s) = M_s$ (resp. \geq, \leq).

Remarques

- Évidemment, une martingale est à la fois une sous et une sur-martingale.
- X est une sous-martingale si et seulement si $-X$ est une sur-martingale.
- Comme $\mathbb{E}(M_t) = \mathbb{E}(M_t | \mathcal{F}_s)$, on voit que l'espérance d'une martingale est constante au cours du temps (alors qu'elle est croissante pour une sous-martingale et décroissante pour une sur-martingale).

Proposition 26 *Si X est un processus à accroissements indépendants avec $\mathbb{E}(|X_t|) < +\infty$ pour tout $t \geq 0$, alors le processus M défini par $M_t = X_t - \mathbb{E}(X_t)$ est une martingale pour sa filtration naturelle.*

Preuve. Il n'y a que le point 3. de la définition à vérifier. Remarquons qu'il revient au même de montrer que pour $s \leq t, \mathbb{E}(M_t - M_s | \mathcal{F}_s) = 0$. En observant que pour $s \leq t$, la variable

$X_t - X_s$ est indépendante de \mathcal{F}_s , on a

$$\begin{aligned}\mathbb{E}(M_t - M_s | \mathcal{F}_s) &= \mathbb{E}(X_t - X_s - \mathbb{E}(X_t) + \mathbb{E}(X_s) | \mathcal{F}_s) \\ &= \mathbb{E}(X_t - X_s | \mathcal{F}_s) - \mathbb{E}(X_t - X_s) \\ &= \mathbb{E}(X_t - X_s) - \mathbb{E}(X_t - X_s) \\ &= 0. \square\end{aligned}$$

Exemples. Le mouvement brownien, $(N_t - \lambda t)_{t \geq 0}$ où N est un processus de Poisson simple de paramètre $\lambda > 0$ sont des martingales.

Proposition 27 *Si B est un mouvement brownien, les processus $(B_t^2 - t)_t$ et $(\exp(\sigma B_t - \frac{\sigma^2}{2}t))_t$ (pour $\sigma \in \mathbb{R}$) sont des martingales pour la filtration naturelle du mouvement brownien.*

Preuve. Les points 1. et 2. sont aisément vérifiés. Montrons le point 3.. L'argument principal consiste à utiliser l'indépendance et la stationnarité des accroissements du mouvement brownien. Pour le premier processus, nous avons

$$\begin{aligned}\mathbb{E}(B_t^2 - t | \mathcal{F}_s) &= \mathbb{E}((B_t - B_s)^2 - B_s^2 + 2B_s B_t - t | \mathcal{F}_s) \\ &= t - s + B_s^2 - t \\ &= B_s^2 - s.\end{aligned}$$

Pour le deuxième,

$$\begin{aligned}\mathbb{E}\left(\exp\left(\sigma B_t - \frac{\sigma^2}{2}t\right) | \mathcal{F}_s\right) &= \exp\left(\sigma B_s - \frac{\sigma^2}{2}t\right) \mathbb{E}(\exp(\sigma(B_t - B_s))) \\ &= \exp\left(\sigma B_s - \frac{\sigma^2}{2}t\right) \times \exp\left(\sigma^2 \frac{t-s}{2}\right) \\ &= \exp\left(\sigma B_s - \frac{\sigma^2}{2}s\right).\end{aligned}$$

Nous avons utilisé au passage l'expression de la transformée de Laplace de $B_t - B_s \sim \mathcal{N}(0, t-s)$. \square

Proposition 28 *Si M est une martingale (resp. sous-martingale) et $\phi : \mathbb{R} \rightarrow \mathbb{R}$ est une fonction convexe (resp. convexe croissante) et telle que $\mathbb{E}(|\phi(M_t)|) < +\infty$ pour tout $t \geq 0$, alors le processus $(\phi(M_t))_{t \geq 0}$ est une sous-martingale pour la même filtration.*

Preuve. Il s'agit d'une application de l'inégalité de Jensen conditionnelle. Dans les deux cas de figures, on a pour $s \leq t$,

$$\mathbb{E}(\phi(M_t) | \mathcal{F}_s) \geq \phi(\mathbb{E}(M_t | \mathcal{F}_s)) \geq \phi(M_s). \square$$

5.1.2 Inégalité maximale de Doob. Convergence

Dans la suite, on ne considérera que des (sous, sur) martingales continues à droite et avec limite à gauche. Ce cas correspond à celui de tous les exemples que nous verrons.

Théorème 11 (Inégalité de Doob) *Soient X est une martingale indexée par un intervalle I de \mathbb{R}_+ et $p \geq 1$ un nombre réel. Alors*

$$\lambda^p \mathbb{P}\left(\sup_{t \in I} |X_t| \geq \lambda\right) \leq \sup_{t \in I} \mathbb{E}(|X_t|^p).$$

Si de plus $p > 1$, alors

$$\mathbb{E} \left(\sup_{t \in I} |X_t|^p \right) \leq \left(\frac{p}{1-p} \right)^p \sup_{t \in I} \mathbb{E} (|X_t|^p).$$

Remarque. Si $(X_t)_{t \in I}$ avec $I = [0, M]$ est une martingale, alors le processus $(|X_t|^p)_{t \in I}$ est une sous-martingale lorsque $p \geq 1$ (voir Proposition 28). En utilisant la croissance de l'espérance, on a

$$\sup_{t \in I} \mathbb{E} (|X_t|^p) = \mathbb{E} (|X_M|^p).$$

Théorème 12 Pour une martingale X , les deux propriétés suivantes sont équivalentes.

1. $\lim_{t \rightarrow +\infty} X_t$ existe dans \mathbb{L}^1 .
2. Il existe $X_\infty \in \mathbb{L}^1$ tel que $X_t = \mathbb{E}(X_\infty | \mathcal{F}_t)$.

Sous l'une ou l'autre des conditions précédentes, la limite est aussi presque sûre (i.e $X_\infty = \lim_{t \rightarrow +\infty} X_t$ p.s.). De plus, si X est bornée dans \mathbb{L}^p pour un certain $p > 1$ (i.e $\sup_{t \geq 0} \|X_t\|_p < +\infty$) alors les propriétés précédentes sont valables et la convergence a lieu aussi dans \mathbb{L}^p (i.e $\lim_{t \rightarrow +\infty} \|X_t - X_\infty\|_p = 0$).

Remarque. L'implication 1. \Rightarrow 2. peut être prouvée facilement. En effet, on a

$$\begin{aligned} \|X_t - \mathbb{E}(X_\infty | \mathcal{F}_t)\|_1 &= \|\mathbb{E}(X_{t+h} | \mathcal{F}_t) - \mathbb{E}(X_\infty | \mathcal{F}_t)\|_1 \\ &\leq \|X_{t+h} - X_\infty\|_1. \end{aligned}$$

Lorsque h tend vers $+\infty$, cette borne converge vers 0. On a alors automatiquement $X_t = \mathbb{E}(X_\infty | \mathcal{F}_t)$ p.s.

5.1.3 Le théorème d'arrêt

La propriété de martingale se généralise au cas des temps d'arrêt.

Théorème 13 (théorème d'arrêt) Soient X une martingale, S et T deux temps d'arrêt bornés tel que $S \leq T$. Alors

$$X_S = \mathbb{E}(X_T | \mathcal{F}_S).$$

Proposition 29 Si X est une martingale et T est un temps d'arrêt, alors le processus X^T défini par $X_t^T = X_{T \wedge t}$ (où $t \wedge T = \min(t, T)$) est encore une martingale pour la même filtration que X .

En prenant $S = 0$, on voit que sous les hypothèses du théorème d'arrêt, $\mathbb{E}(X_T) = \mathbb{E}(X_0)$. Ce théorème peut se généraliser à des temps d'arrêt non bornés sous certaines conditions.

Exemple d'utilisation. Soit $a > 0$. Définissons le temps d'arrêt T_a par

$$T_a = \inf \{t > 0 : B_t \geq a\},$$

qui représente le premier instant où un mouvement brownien dépasse le seuil a . On peut vérifier que $T_a < +\infty$ p.s. En effet si $N > 0$, on a

$$\mathbb{P}(T_a > N) = \mathbb{P} \left(\sup_{0 \leq t \leq N} B_t < a \right) = \mathbb{P} (|B_N| \leq a) = \mathbb{P} \left(|B_1| \leq \frac{a}{\sqrt{N}} \right),$$

en utilisant la loi du sup du brownien sur un intervalle (loi que nous avons déterminée au Chapitre 2). On en déduit que

$$\mathbb{P}(T_a = +\infty) = \lim_{N \rightarrow +\infty} \mathbb{P}(T_a > N) = 0.$$

Ainsi, en utilisant la continuité des trajectoires, on en déduit que $B_{T_a} = a$ p.s. Nous allons utiliser le théorème d'arrêt afin de déterminer la loi de T_a . Soient $s \in \mathbb{R}$ et

$$M_t = \exp\left(sB_t - \frac{s^2}{2}t\right), \quad t \in \mathbb{R}_+.$$

Alors M est une martingale (en utilisant la filtration associée au brownien). Comme $t \wedge T_a$ (minimum de t est de T_a) est un temps d'arrêt borné, nous avons d'après le théorème d'arrêt $\mathbb{E}(M_{t \wedge T_a}) = 1$. Comme $M_{t \wedge T_a} \leq \exp(sa)$, le théorème de convergence dominée garantit que

$$1 = \mathbb{E}(M_{T_a}) = \lim_{t \rightarrow +\infty} \mathbb{E}(M_{t \wedge T_a}).$$

Mais l'égalité $\mathbb{E}(M_{T_a}) = 1$ s'écrit $\mathbb{E}\left(\exp\left(-\frac{s^2}{2}T_a\right)\right) = \exp(-sa)$. Ainsi la transformée de Laplace de T_a est donnée par

$$\mathbb{E}(\exp(-\lambda T_a)) = \exp(-\sqrt{2\lambda a}), \quad \lambda > 0.$$

Ainsi la loi de T_a est déterminée. On peut en fait montrer qu'il s'agit de la transformée de Laplace de la densité

$$x \mapsto \frac{a}{\sqrt{2\pi x^3}} \exp\left(-\frac{a^2}{2x}\right) \mathbb{1}_{x>0}.$$

5.2 La notion d'intégrale stochastique

5.2.1 Motivation

Nous présentons deux exemples qui illustrent l'intérêt de définir l'intégrale d'un processus par rapport à un mouvement brownien.

1. Considérons l'équation différentielle

$$N'_t = a_t N_t, \quad N_0 = x.$$

La fonction N peut par exemple représenter l'évolution du cours d'un actif au cours du temps; a_t est alors interprété comme le taux d'intérêt instantané du prix au temps t . Il s'agit d'une équation déterministe. Il serait naturel de supposer que a_t soit soumis à des effets aléatoires :

$$a_t = r_t + \text{"bruit"}.$$

On pourrait alors penser à écrire

$$N'_t = (r_t + Y_t) N_t,$$

où $\{Y_t : t \geq 0\}$ est famille de variables aléatoires i.i.d toutes de loi gaussienne $\mathcal{N}(0, \sigma^2)$. Cependant, les trajectoires $t \mapsto Y_t(\omega)$ sont très irrégulières. Une solution est alors d'utiliser les accroissements d'un mouvement brownien B . Si $t_{i+1} - t_i$ est proche de 0, on pourrait remplacer

$$N_{t_{i+1}} - N_{t_i} \approx a_{t_i} N_{t_i} (t_{i+1} - t_i) + N_{t_i} Y_{t_i} (t_{i+1} - t_i)$$

par

$$N_{t_{i+1}} - N_{t_i} \approx a_{t_i} N_{t_i} (t_{i+1} - t_i) + N_{t_i} (B_{t_{i+1}} - B_{t_i}).$$

En sommant sur une subdivision $0 = t_0 < t_1 < \dots < t_n = t$, on aurait

$$N_t - N_0 \approx \sum_{i=0}^{n-1} a_{t_i} N_{t_i} (t_{i+1} - t_i) + \sum_{i=0}^{n-1} N_{t_i} (B_{t_{i+1}} - B_{t_i}).$$

Si on fait tendre le pas de la subdivision vers 0, on aimerait que le processus N vérifie

$$N_t - N_0 = \int_0^t a_s N_s ds + \int_0^t N_s dB_s, \quad t \geq 0.$$

La notion d'intégrale par rapport au brownien reste à définir.

2. On peut étudier le mouvement d'une particule soumise à une infinité de chocs aléatoires en des temps très courts (par exemple dans un liquide) en écrivant pour h petit

$$X_{t+h} - X_t \approx b(t, X_t) h + \sigma(t, X_t) (B_{t+h} - B_t),$$

où B désigne un mouvement brownien de \mathbb{R}^3 , X_t désigne la position de la particule au temps t , la fonction b donne la direction du mouvement (paramètre de dérive ou drift) et le second terme est un terme de fluctuation avec une fonction σ appelée coefficient de diffusion (ou volatilité) qui dépend du temps et de la position (et est à valeurs dans l'ensemble de matrices réelles de taille $(3, 3)$). Remarquons que si X_t est indépendant de l'accroissement $B_{t+h} - B_t$, alors on considère que l'accroissement $X_{t+h} - X_t$ suit approximativement une loi gaussienne de moyenne $b(t, X_t) h$ et de variance $\sigma(t, X_t) \sigma(t, X_t)^T h$ (conditionnellement à X_t). Là aussi, la somme des accroissements le long d'une subdivision d'un intervalle de temps conduit à la notion d'intégrale par rapport au brownien.

5.3 Construction de l'intégrale stochastique

Nous avons vu lors du Chapitre 2 que les trajectoires du mouvement brownien n'étaient pas à variation bornée. Il n'est donc pas possible de définir l'intégrale d'un processus X par rapport au mouvement brownien en étudiant la convergence des sommes de Riemann $\sum_{i=0}^{n-1} X_{t_i}(\omega) (B_{t_{i+1}}(\omega) - B_{t_i}(\omega))$ pour $\omega \in \Omega$ fixé. L'idée développée par Itô est d'étudier la convergence dans $\mathbb{L}^2(\Omega)$ de ces sommes partielles.

Dans tout ce chapitre, on se donne un mouvement brownien continu

$$B = \left(\Omega, \mathcal{F}, (\mathcal{F}_t)_{t \geq 0}, (B_t)_{t \geq 0}, \mathbb{P} \right).$$

On considère également un intervalle $I = [0, T]$ où T est un réel strictement positif.

Définition 19 On dit que $X = (X_t)_{t \in I}$ est un processus élémentaire s'il existe une subdivision $0 = t_0 < t_1 < \dots < t_n = T$ de I ainsi que des variables aléatoires Z_i , \mathcal{F}_{t_i} -mesurable pour $i = 0, \dots, n-1$, telles que

$$X_t = \sum_{i=0}^{n-1} Z_i \mathbf{1}_{t_i \leq t < t_{i+1}}, \quad t \in I.$$

On notera E_2 l'ensemble des processus élémentaires de carré intégrable et E l'ensemble des processus élémentaires quelconques.

$X \in E$ est un processus dont les trajectoires sont constantes par morceaux. On définit l'intégrale de ce type de processus de la manière suivante. On peut remarquer que E est un espace vectoriel.

Définition 20 Si $X \in E$, on définit son intégrale stochastique par la formule :

$$\int_0^T X_s dB_s = \sum_{i=0}^{n-1} Z_i (B_{t_{i+1}} - B_{t_i}).$$

Proposition 30 1. Si $X, Y \in E$ et $\lambda \in \mathbb{R}$, alors

$$\int_0^T (X_s + \lambda Y_s) dB_s = \int_0^T X_s dB_s + \lambda \int_0^T Y_s dB_s.$$

2. Si $X \in E_2$, alors la moyenne et la variance de l'intégrale stochastique sont données par

$$\mathbb{E} \left(\int_0^T X_s dB_s \right) = 0, \quad \mathbb{E} \left[\left(\int_0^T X_s dB_s \right)^2 \right] = \mathbb{E} \left(\int_0^T X_s^2 ds \right).$$

Preuve

1. Il suffit de prendre une subdivision commune à X et à Y (par exemple la réunion des deux subdivisions associées). L'égalité est alors immédiate.
2. La preuve du second point utilise l'argument clé suivant. Pour tout $s < t$, l'accroissement $B_t - B_s$ est indépendant de \mathcal{F}_s . Pour le calcul de la moyenne, nous avons

$$\begin{aligned} \mathbb{E} \left(\int_0^T X_s dB_s \right) &= \sum_{i=0}^{n-1} \mathbb{E} (Z_i (B_{t_{i+1}} - B_{t_i})) \\ &= \sum_{i=0}^{n-1} \mathbb{E} (Z_i) \mathbb{E} (B_{t_{i+1}} - B_{t_i}) \\ &= 0. \end{aligned}$$

On a

$$\left(\int_0^T X_s dB_s \right)^2 = \sum_{i,j=0}^{n-1} Z_i Z_j (B_{t_{i+1}} - B_{t_i}) (B_{t_{j+1}} - B_{t_j}).$$

Si $i < j$, alors $B_{t_{j+1}} - B_{t_j}$ est indépendant du produit $Z_i Z_j (B_{t_{i+1}} - B_{t_i})$ qui est \mathcal{F}_{t_j} -mesurable. On en déduit que

$$\begin{aligned} \mathbb{E} \left(\left(\int_0^T X_s dB_s \right)^2 \right) &= \sum_{i=0}^{n-1} \mathbb{E} (Z_i^2 (B_{t_{i+1}} - B_{t_i})^2) \\ &= \sum_{i=0}^{n-1} \mathbb{E} (Z_i^2) \mathbb{E} ((B_{t_{i+1}} - B_{t_i})^2) \\ &= \sum_{i=0}^{n-1} \mathbb{E} (Z_i^2) (t_{i+1} - t_i) \\ &= \mathbb{E} \left(\int_0^T X_s^2 ds \right). \square \end{aligned}$$

Remarque. Si $X \in E_2$, alors $(\omega, t) \mapsto X_t(\omega)$ appartient à M_2 qui le sous-espace vectoriel de

$$\mathbb{L}^2(\Omega \times [0, T], \mathcal{F} \otimes \mathcal{B}([0, T]), d\mathbb{P} \otimes dt)$$

formé des processus adaptés. De plus $\sqrt{\mathbb{E} \left(\int_0^T X_s^2 ds \right)}$ représente la norme de X pour cet espace.

Corollaire 2 L'application $X \mapsto I(X) = \int_0^T X_t dB_t$ est une isométrie sur l'espace vectoriel E_2 muni de la norme de M_2 dans l'espace vectoriel $\mathbb{L}^2(\Omega, \mathcal{F}, \mathbb{P})$.

Preuve. I est linéaire et d'après la proposition précédente,

$$\|I(X)\|_{\mathbb{L}^2(\mathbb{P})}^2 = \mathbb{E} \left(\int_0^T X_s^2 ds \right). \square$$

La construction de l'intégrale stochastique est basée sur le résultat de densité suivant que nous admettrons.

Théorème 14 Si $X \in M_2$, alors il existe une suite $(X^{(n)})_n$ d'éléments de E_2 telle que

$$\lim_{n \rightarrow +\infty} \mathbb{E} \left(\int_0^T |X_s^{(n)} - X_s|^2 ds \right) = 0.$$

Corollaire 3 L'isométrie $I : X \mapsto \int_0^T X_s dB_s$ de E_2 dans $\mathbb{L}^2(\Omega)$ se prolonge de manière unique en une isométrie de M_2 dans $\mathbb{L}^2(\Omega)$. On parle d'isométrie d'Itô.

Preuve. Soient $X \in M_2$ et $(X^{(n)})_n$ une suite d'éléments de E_2 telle que

$$\lim_{n \rightarrow \infty} \mathbb{E} \left(\int_0^T (X_t - X_t^{(n)})^2 dt \right) = 0.$$

Alors si $p, q \in \mathbb{N}$, on a

$$\|I(X^{(p)}) - I(X^{(q)})\|_2 = \sqrt{\mathbb{E} \left(\int_0^T (X_t^{(p)} - X_t^{(q)})^2 dt \right)}.$$

On voit alors que la suite $(I(X^{(n)}))_n$ est une suite de Cauchy dans $\mathbb{L}^2(\Omega)$ qui est un espace vectoriel normé complet. Cette suite converge et on peut définir alors $I(X) = \lim_{n \rightarrow \infty} I(X^{(n)})$. Il est aisé de vérifier que cette limite ne dépend pas de la suite choisie. En effet, si $(Y^{(n)})_n$ est une autre suite qui approche X , la propriété d'isométrie montre que

$$\|I(X^{(n)}) - I(Y^{(n)})\|_2^2 = \mathbb{E} \left(\int_0^T (X_t^{(n)} - Y_t^{(n)})^2 dt \right).$$

Ainsi les deux suites $(I(X^{(n)}))_n$ et $(I(Y^{(n)}))_n$ possèdent la même limite dans $\mathbb{L}^2(\Omega)$. \square

Définition 21 Notons encore I le prolongement de l'isométrie évoqué dans le corollaire précédent. Si $X \in M_2$, la variable aléatoire $I(X)$ est notée également $\int_0^T X_s dB_s$ et est appelée intégrale stochastique de X par rapport à B sur l'intervalle $[0, T]$.

Exemple de calcul. On se propose de calculer $\int_0^T B_t dB_t$. Notons que cette intégrale existe bien car

$$\mathbb{E} \left(\int_0^T B_t^2 dt \right) = \int_0^T t dt = \frac{T^2}{2},$$

ce qui montre que B est bien dans M_2 . Utilisons l'approximation

$$X_t^{(n)} = \sum_{k=0}^{2^n-1} B_{\frac{kT}{2^n}} \mathbf{1}_{\frac{kT}{2^n} \leq t < \frac{(k+1)T}{2^n}}.$$

En effet

$$\begin{aligned} \mathbb{E} \left(\int_0^T (X_t^{(n)} - B_t)^2 dt \right) &= \sum_{k=0}^{2^n-1} \mathbb{E} \left(\int_{\frac{kT}{2^n}}^{\frac{(k+1)T}{2^n}} (B_{\frac{kT}{2^n}} - B_t)^2 dt \right) \\ &= \sum_{k=0}^{2^n-1} \int_{\frac{kT}{2^n}}^{\frac{(k+1)T}{2^n}} \left(-\frac{kT}{2^n} + t \right) dt \\ &= \frac{T^2}{2^{n+1}} \xrightarrow{n \rightarrow \infty} 0. \end{aligned}$$

On a donc

$$\begin{aligned} \int_0^T B_s dB_s &= \lim_{n \rightarrow +\infty} \int_0^T X_t^{(n)} dB_t \\ &= \lim_{n \rightarrow +\infty} \sum_{k=0}^{2^n-1} B_{\frac{kT}{2^n}} (B_{\frac{(k+1)T}{2^n}} - B_{\frac{kT}{2^n}}) \\ &= \frac{1}{2} \lim_{n \rightarrow +\infty} \sum_{k=0}^{2^n-1} \left(B_{\frac{(k+1)T}{2^n}}^2 - B_{\frac{kT}{2^n}}^2 - (B_{\frac{(k+1)T}{2^n}} - B_{\frac{kT}{2^n}})^2 \right) \\ &= \frac{1}{2} B_T^2 - \frac{1}{2} \lim_{n \rightarrow +\infty} \sum_{k=0}^{2^n-1} (B_{\frac{(k+1)T}{2^n}} - B_{\frac{kT}{2^n}})^2 \end{aligned}$$

En utilisant l'expression de la variation quadratique du mouvement brownien (Chapitre 2), la limite dans \mathbb{L}^2 de l'expression ci-dessus vaut $\frac{1}{2} (B_T^2 - T)$. On a donc

$$\int_0^T B_s dB_s = \frac{1}{2} (B_T^2 - T).$$

Cas particulier de l'intégrale de Wiener. Soit $h : \mathbb{R}_+ \rightarrow \mathbb{R}$ une fonction (déterministe) mesurable et de carré intégrable sur $[0, T]$, i.e

$$\int_0^T h(s)^2 ds < +\infty.$$

L'intégrale stochastique $\int_0^T h(s) dB_s$ est donc bien définie. On peut choisir des fonctions déterministes h_n pour $n \in \mathbb{N}$ de la forme

$$h_n = \sum_{i=0}^{p_n-1} c_i^{(n)} \mathbf{1}_{t_i^{(n)} \leq t < t_{i+1}^{(n)}}$$

et telles que

$$\lim_{n \rightarrow +\infty} \int_0^T |h_n(s) - h(s)|^2 ds = 0.$$

Remarquons que si h est continue, on peut choisir $c_i^{(n)} = h\left(s_i^{(n)}\right)$ où $t_i^{(n)} \leq s_i^{(n)} < t_{i+1}^{(n)}$ (avec un pas de subdivision qui tend vers 0 quand $n \rightarrow +\infty$). Ainsi,

$$M_T = \int_0^T h(s)dB_s = \lim_{n \rightarrow +\infty} \sum_{i=0}^{p_n-1} c_i^{(n)} \left(B_{t_{i+1}^{(n)}} - B_{t_i^{(n)}} \right)$$

est la limite dans $\mathbb{L}^2(\Omega)$ de variables aléatoires gaussiennes. M_T est donc une variable aléatoire gaussienne. Sa moyenne vaut 0 et sa variance est $\mathbb{E}(M_T^2) = \int_0^T h(s)^2 ds$. On parle d'intégrale de Wiener lorsque le processus à intégrer est déterministe.

Si $X \in M_2$, on peut définir plus généralement $\int_a^b X_s dB_s$ pour $0 \leq a < b \leq T$. On dispose alors des propriétés suivantes.

Proposition 31 *Si X et Y sont dans M_2 et $\lambda \in \mathbb{R}$, alors*

1. $\int_u^v X_s dB_s = \int_u^w X_s dB_s + \int_w^v X_s dB_s$, $0 \leq u < w < v \leq T$,
2. $\int_u^v (X_s + \lambda Y_s) dB_s = \int_u^v X_s dB_s + \lambda \int_u^v Y_s dB_s$,
3. $\mathbb{E}\left(\int_u^v X_s dB_s\right) = 0$ et $\text{Var}\left(\int_u^v X_s dB_s\right) = \mathbb{E}\left(\int_u^v X_s^2 ds\right)$,
4. $\int_u^v ZX_s dB_s = Z \int_u^v X_s dB_s$ si Z est \mathcal{F}_u -mesurable et $ZX \in M_2$.

Preuve. Il suffit d'approcher par des processus élémentaires et de passer à la limite. \square

5.3.1 L'intégrale stochastique en tant que processus

Proposition 32 *Soit $X \in M_2$.*

1. *Le processus $(I_t)_{0 \leq t \leq T}$ est une martingale de carré intégrable adaptée à la filtration brownienne.*
2. *Il existe une version continue du processus précédent.*

Preuve. On ne prouve que le premier point. Si $t \in \mathbb{R}_+$, I_t peut s'écrire comme une limite dans \mathbb{L}^2 d'une suite de variables aléatoires \mathcal{F}_t -mesurables. Quitte à considérer une sous-suite, on peut supposer que la convergence est presque sûre. Ceci entraîne que I_t est \mathcal{F}_t -mesurable. Pour vérifier la propriété de martingale, on considère $A \in \mathcal{F}_t$ et $h > 0$. On a

$$\begin{aligned} \mathbb{E}((I_{t+h} - I_t) \mathbb{1}_A) &= \mathbb{E}\left(\mathbb{1}_A \int_t^{t+h} X_s dB_s\right) \\ &= \mathbb{E}\left(\int_t^{t+h} \mathbb{1}_A X_s dB_s\right) \\ &= 0, \end{aligned}$$

en utilisant les propriétés de l'intégrale stochastique (Proposition 31). D'après la caractérisation de l'espérance conditionnelle, on a bien $\mathbb{E}(I_{t+h} | \mathcal{F}_t) = I_t$. \square

Remarque sur l'intégrale de Wiener. Lorsque $X_t = \int_0^t f(s)dB_s$ où f désigne une fonction localement de carré intégrable, on peut vérifier que le processus X est gaussien : si $t_1 < t_2 < \dots < t_n$, toute combinaison linéaire de variables X_{t_1}, \dots, X_{t_n} peut aussi s'écrire comme une intégrale de Wiener entre $t = 0$ et $t = t_n$. Pour le prouver, il suffit en fait d'observer que

$$\int_0^{t_i} f(s)dB_s = \int_0^{t_n} \mathbb{1}_{[0,t_i]}(s)f(s)dB_s, \quad i = 1, \dots, n.$$

En effet l'égalité précédente est vraie pour les fonctions en escalier et on peut passer à la limite. On peut alors utiliser les propriétés de linéarité de l'intégrale stochastique.

Extension de l'intégrale stochastique. On peut donner un sens à l'intégrale $\int_a^b X_s dB_s$ lorsque on a seulement $\mathbb{P}\left(\int_a^b X_s^2 ds < +\infty\right) = 1$. Pour un tel processus X mesurable et adapté à la filtration brownienne, on peut montrer qu'il existe une suite $(X^{(n)})_n$ d'éléments de E tels que

$$\lim_{n \rightarrow +\infty} \int_a^b \left(X_s^{(n)} - X_s\right)^2 ds = 0, \quad \mathbb{P} - \text{p.s.}$$

Dans ce cas, il existe une variable aléatoire Z telle que

$$\lim_{n \rightarrow +\infty} \int_a^b X_s^{(n)} dB_s = Z, \quad \text{en probabilité}$$

et Z ne dépend pas de la suite choisie (à l'égalité presque sûre près). On écrit $Z = \int_a^b X_s dB_s$. Lorsque si $X \in M_2$, on retrouve l'intégrale stochastique précédemment définie. Cependant, dans le cas général, le processus $\left(\int_0^t X_s dB_s\right)_{0 \leq t \leq T}$ n'est plus forcément une martingale. On parle alors de martingale locale, notion qui ne sera pas développée dans ce cours. Lorsque X est à trajectoires continues, on peut facilement déduire de cette construction que

$$\lim_{n \rightarrow +\infty} \sum_{i=0}^{m_n-1} X_{t_{n,i}} (B_{t_{n,i+1}} - B_{t_{n,i}}) = \int_a^b X_s dB_s,$$

en probabilité pour une suite de subdivisions de $[0, T]$ dont le pas tend vers 0. En effet, presque sûrement, les trajectoires du processus

$$X_t^{(n)} = \sum_{i=0}^{m_n-1} X_{t_{n,i}} \mathbb{1}_{t_{n,i} \leq t < t_{n,i+1}}$$

converge uniformément vers celles de X sur l'intervalle $[0, T]$.

5.4 La formule d'Itô

Définition 22 On dit que X est un processus d'Itô s'il est $(\mathcal{F}_t)_t$ -adapté et s'il s'écrit

$$X_t = X_0 + \int_0^t a_s ds + \int_0^t b_s dB_s, \quad t \in [0, T],$$

où a et b sont deux processus adaptés tels que $\int_0^T |a_s| ds < +\infty$ et $\int_0^T b_s^2 ds < +\infty$ presque sûrement.

Deux processus adaptés a et b vérifient les conditions d'intégrabilité de la définition précédente si leurs trajectoires sont continues. On écrira aussi $dX_t = a_t dt + b_t dB_t$.

Exemple. On a vu que $\int_0^t B_s dB_s = \frac{1}{2} (B_t^2 - t)$ ce qui entraîne que

$$B_t^2 = t + 2 \int_0^t B_s dB_s.$$

Ceci montre que $(B_t^2)_t$ est un processus d'Itô. Plus généralement, nous avons le résultat fondamental suivant qui donne la formule dite d'Itô.

Théorème 15 Soit X un processus d'Itô sur $[0, T] : dX_t = a_t dt + b_t dB_t$, si $t \in [0, T]$. Si $f : \mathbb{R} \times \mathbb{R}_+ \rightarrow \mathbb{R}$ est une fonction de classe \mathcal{C}^2 , alors le processus $(f(X_t, t))_{0 \leq t \leq T}$ est encore un processus d'Itô et on a presque sûrement

$$f(X_t, t) = f(X_0, 0) + \int_0^t \frac{\partial f}{\partial t}(X_s, s) ds + \int_0^t \frac{\partial f}{\partial x}(X_s, s) dX_s + \frac{1}{2} \int_0^t \frac{\partial^2 f}{\partial x^2}(X_s, s) b_s^2 ds.$$

Remarque. Nous admettrons la formule précédente appelée formule d'Itô mais nous allons expliquer brièvement l'origine du dernier terme. Prenons le cas du processus $X_t = f(B_t)$ où B est un mouvement brownien et f est une fonction \mathcal{C}^2 . On suppose que f, f' et f'' sont bornées. La formule de Itô donne

$$dX_t = f'(B_t) dB_t + \frac{1}{2} f''(B_t) dt.$$

Si maintenant $Y_t = f(A_t)$ où les trajectoires du processus A sont de classe \mathcal{C}^1 alors on sait d'après la formule de la moyenne que

$$dY_t = f'(A_t) dA_t,$$

avec $dA_t = A'_t dt$ et le deuxième terme est absent. Nous avons déjà évoqué au Chapitre 2 le fait les trajectoires de B étaient bien plus irrégulières que celles d'un processus \mathcal{C}^1 . Comparons le comportement infinitésimal des deux processus à l'aide de développements limités.

$$X_{t_{i+1}} - X_{t_i} = f'(B_{t_i}) (B_{t_{i+1}} - B_{t_i}) + \frac{1}{2} f''(B_{t_i}) (B_{t_{i+1}} - B_{t_i})^2 + \text{"reste"}.$$

En sommant sur une subdivision de $[0, t]$, on obtient

$$X_t - X_0 = \sum_i f'(B_{t_i}) (B_{t_{i+1}} - B_{t_i}) + \frac{1}{2} \sum_i f''(B_{t_i}) (B_{t_{i+1}} - B_{t_i})^2 + \text{"reste"}.$$

On a bien sûr la même chose pour Y . Au niveau du premier terme, on a

$$\begin{aligned} \sum_i f'(B_{t_i}) (B_{t_{i+1}} - B_{t_i}) &\rightarrow \int_0^t f'(B_s) dB_s, \\ \sum_i f'(A_{t_i}) (A_{t_{i+1}} - A_{t_i}) &\rightarrow \int_0^t f'(A_s) dA_s. \end{aligned}$$

Au niveau du deuxième terme, on voit que

$$\begin{aligned} \sum_i f''(A_{t_i}) (A_{t_{i+1}} - A_{t_i})^2 &\leq \sup_i |A_{t_{i+1}} - A_{t_i}| \sum_i |f''(A_{t_i})| |A_{t_{i+1}} - A_{t_i}| \\ &\leq \sup_i |A_{t_{i+1}} - A_{t_i}| \times V_{[0,t]}(A) \times \sup_{s \in [0,t]} |f''(A_s)| \end{aligned}$$

et cette dernière borne tend vers 0 lorsque $n \rightarrow +\infty$ ($V_{[0,t]}(A)$ désigne la variation totale de A sur $[0, t]$). En revanche le deuxième terme semblable pour X ne tend pas vers 0. C'est clair

intuitivement, car au niveau des espérances on peut voir en utilisant les propriétés du brownien que

$$\mathbb{E} \left(\sum_i f''(B_{t_i}) (B_{t_{i+1}} - B_{t_i})^2 \right) = \mathbb{E} \left(\sum_i f''(B_{t_i}) (t_{i+1} - t_i) \right),$$

quantité qui tend vers $\mathbb{E} \left(\int_0^t f''(B_s) ds \right)$ qui est la moyenne du deuxième terme de la formule d'Itô. On peut en fait montrer que

$$\lim \sum_i f''(B_{t_i}) (B_{t_{i+1}} - B_{t_i})^2 = \int_0^t f''(B_s) ds,$$

dans $\mathbb{L}^2(\Omega)$. Au niveau des "reste", on peut montrer que les deux quantités tendent vers 0 (pour X ou pour Y). On se rend compte sur cet exemple que l'irrégularité des trajectoires du mouvement brownien est à l'origine de l'apparition d'un terme supplémentaire pour la formule de la moyenne habituelle des fonctions \mathcal{C}^1 .

Dans le même ordre d'idée, on a la formule d'intégration par parties.

Proposition 33 Soit X et Y deux processus d'Itô dont les décompositions sont données par

$$\begin{aligned} dX_t &= a_t^{(1)} dt + b_t^{(1)} dB_t, \\ dY_t &= a_t^{(2)} dt + b_t^{(2)} dB_t. \end{aligned}$$

Alors le processus $(X_t Y_t)_{t \geq 0}$ est aussi un processus d'Itô et sa décomposition est donnée par

$$d(XY)_t = X_t dY_t + Y_t dX_t + b_t^{(1)} b_t^{(2)} dt.$$

Preuve. Il s'agit d'une conséquence de la formule d'Itô en utilisant la décomposition

$$X_t Y_t = \frac{1}{2} (X_t + Y_t)^2 - \frac{1}{2} X_t^2 - \frac{1}{2} Y_t^2,$$

et en appliquant la formule d'Itô à chaque terme. \square

Exemples.

1. $B_t^2 = t + 2 \int_0^t B_s dB_s$ est une formule déjà vue auparavant.
2. Soit $X_t = \exp(B_t - \frac{1}{2}t) = f(B_t, t)$. On obtient

$$\begin{aligned} dX_t &= \exp\left(B_t - \frac{1}{2}t\right) dB_t - \frac{1}{2} X_t dt + \frac{1}{2} X_t dt \\ &= X_t dB_t. \end{aligned}$$

3. Soit $X_t = tB_t^3$. On a

$$\begin{aligned} X_t &= \int_0^t B_s^3 ds + \int_0^t 3sB_s^2 dB_s + \frac{1}{2} \int_0^t 6sB_s ds \\ &= \int_0^t B_s^3 ds + 3 \int_0^t sB_s^2 dB_s + 3 \int_0^t sB_s ds. \end{aligned}$$

Remarquons qu'il revient au même de chercher d'abord la décomposition d'Itô de $Y_t = B_t^3$:

$$Y_t = \int_0^t 3B_t^2 dB_t + \int_0^t 3B_t dt.$$

En posant $X_t = tY_t$ et en appliquant de nouveau la formule d'Itô, on retrouve alors le même résultat.

Extension au cas multivarié. On considère un mouvement brownien B sur \mathbb{R}^d (les processus coordonnés sont indépendants et sont tous des browniens réels)

$$B = (B^{(1)}, \dots, B^{(d)}).$$

Un processus d'Itô X à valeurs dans \mathbb{R}^n est de la forme

$$\begin{cases} dX_t^{(1)} = u_t^{(1)} dt + \sum_{j=1}^d v_t^{(1,j)} dB_t^{(j)} \\ \dots \\ dX_t^{(n)} = u_t^{(n)} dt + \sum_{j=1}^d v_t^{(n,j)} dB_t^{(j)} \end{cases}$$

Les processus $u^{(i)}$ et $v^{(i,j)}$ satisfont les conditions d'intégrabilité

$$\int_0^t |u_s^{(i)}| ds, \quad \int_0^t (v_s^{(i,j)})^2 ds < +\infty.$$

On note plus simplement

$$dX_t = u_t dt + v_t dB_t.$$

La formule d'Itô s'écrit pour $Y_t = f(X_t, t)$

$$dY_t = \frac{\partial f}{\partial t}(X_t, t) dt + \sum_{i=1}^d \frac{\partial f}{\partial x_i}(X_t, t) dX_t^{(i)} + \frac{1}{2} \sum_{i,j=1}^d \frac{\partial^2 f}{\partial x_i \partial x_j}(X_t, t) d \langle X^{(i)}, X^{(j)} \rangle_t,$$

avec

$$d \langle X^{(i)}, X^{(j)} \rangle_t = \sum_{k=1}^d v_t^{(i,k)} v_t^{(j,k)} dt = (v_t v_t')_{ij} dt.$$

Remarque. On retrouve la formule d'intégrations par parties à l'aide de cette formule en posant

$$u_t = \begin{pmatrix} a_t^{(1)} \\ a_t^{(2)} \end{pmatrix}, \quad v_t = \begin{pmatrix} b_t^{(1)} & 0 \\ b_t^{(2)} & 0 \end{pmatrix}.$$

5.5 Introduction aux équations différentielles stochastiques

5.5.1 Existence et unicité des solutions (coefficients Lipschitz)

On considère l'équation différentielle stochastique (EDS) suivante

$$\begin{cases} dX_t = b(t, X_t) dt + \sigma(t, X_t) dB_t, \\ X_0 = x. \end{cases}$$

Le théorème suivant assure sous certaines conditions l'existence et l'unicité d'un processus solution de l'EDS. De tels processus sont aussi appelés processus de diffusion. Dans la suite, on notera toujours par $|\cdot|$ une norme arbitraire sur \mathbb{R}^n (par exemple la norme euclidienne) pour une valeur de n quelconque.

Théorème 16 Soit $T > 0$. On suppose qu'il existe une constante $C > 0$ telle que

1. $|\sigma(t, x)| + |b(t, x)| \leq C(1 + |x|), \quad \forall t \in [0, T], x \in \mathbb{R}^n,$
2. $|\sigma(t, x) - \sigma(t, y)| + |b(t, x) - b(t, y)| \leq C|x - y|, \quad x, y \in \mathbb{R}^n, \quad t \in [0, T].$

Alors il existe une solution $X = (X_t)_{t \in [0, T]}$ appartenant à $M_2([0, T])$ et une seule à indistinguishabilité près (i.e si X et Y sont deux solutions, alors $\mathbb{P}(X_t = Y_t : t \in [0, T]) = 1$).

Preuve. Nous considérerons uniquement le cas $n = 1$ (le cas $n > 1$ complique juste un peu les notations mais la preuve est identique). Sur $M_2([0, T])$, on définit une nouvelle norme :

$$\|X\|_\lambda^2 = \mathbb{E} \left(\int_0^T \exp(-\lambda t) |X_t|^2 dt \right).$$

Le réel positif λ sera précisé par la suite. On va appliquer le théorème du point fixe dans l'espace vectoriel normé complet $(M_2, \|\cdot\|_\lambda)$. La norme introduite est en fait équivalente à la norme usuelle sur M_2 (car $\exp(-\lambda t) \in [\exp(-\lambda T), 1]$). On définit l'application $\Psi : M_2 \rightarrow M_2$ par

$$\Psi(X)_t = x + \int_0^t b(s, X_s) ds + \int_0^t \sigma(s, X_s) dB_s, \quad t \in [0, T].$$

On peut voir que X est solution de l'EDS si et seulement si $\Psi(X)_t = X_t$, $t \in [0, T]$. On va montrer que Ψ est bien à valeurs dans M_2 et qu'elle est contractante.

- Soit $X \in M_2$ et $\Psi_1(X)_t = \int_0^t b(s, X_s) ds$, $t \in [0, T]$. D'après l'inégalité de Cauchy-Schwarz, on a

$$\Psi_1(X)_t^2 \leq t \int_0^t b(s, X_s)^2 ds.$$

En utilisant la condition de non-explosion, on en déduit que

$$\begin{aligned} \mathbb{E} \left(\int_0^T \Psi_1(X)_t^2 dt \right) &\leq \mathbb{E} \left(t \int_0^T \int_0^t b(s, X_s)^2 ds dt \right) \\ &\leq T^2 \mathbb{E} \left(\int_0^T b(s, X_s)^2 ds \right) \\ &\leq T^2 \mathbb{E} \left(\int_0^T (1 + |X_s|)^2 ds \right) < +\infty. \end{aligned}$$

Ceci montre que $\Psi_1(X) \in M_2$. De plus, si $X, Y \in M_2$,

$$\begin{aligned} \|\Psi_1(X) - \Psi_1(Y)\|_\lambda &= \mathbb{E} \left(\int_0^T \exp(-\lambda t) \left[\int_0^t (b(s, X_s) - b(s, Y_s)) ds \right]^2 dt \right) \\ &\leq T \mathbb{E} \left(\int_0^T \exp(-\lambda t) \int_0^t |b(s, X_s) - b(s, Y_s)|^2 ds dt \right) \\ &\leq TC^2 \mathbb{E} \left(\int_0^T \exp(-\lambda t) \int_0^t |X_s - Y_s|^2 ds dt \right) \\ &\leq TC^2 \left(-\frac{\exp(-\lambda T)}{\lambda} \int_0^T |X_s - Y_s|^2 ds + \int_0^T \frac{\exp(-\lambda t)}{\lambda} |X_t - Y_t|^2 dt \right) \\ &\leq \frac{TC^2}{\lambda} \|X - Y\|_\lambda^2. \end{aligned}$$

L'avant dernière inégalité est en fait une égalité et se prouve en utilisant une intégration par parties.

- Si $X \in M_2$, on pose $\Psi_2(X)_t = \int_0^t \sigma(s, X_s) dB_s$, $t \in [0, T]$. D'après la propriété d'isométrie d'Itô et l'hypothèse de non-explosion, on a

$$\mathbb{E} \left(\left| \int_0^t \sigma(s, X_s) dB_s \right|^2 \right) = \mathbb{E} \left(\int_0^t \sigma(s, X_s)^2 ds \right) \leq C^2 \mathbb{E} \left(\int_0^t (1 + |X_s|)^2 ds \right) < +\infty.$$

En intégrant l'espérance ci-dessus pour t variant de 0 à T , on montre que $\Psi_2(X) \in M_2$. De plus si $X, Y \in M_2$,

$$\begin{aligned} \|\Psi_2(X) - \Psi_2(Y)\|_\lambda &= \mathbb{E} \left(\int_0^T \exp(-\lambda t) \left[\int_0^t (\sigma(s, X_s) - \sigma(s, Y_s)) dB_s \right]^2 dt \right) \\ &\leq \int_0^T \exp(-\lambda t) \mathbb{E} \left(\int_0^t |\sigma(s, X_s) - \sigma(s, Y_s)|^2 ds \right) dt. \end{aligned}$$

En procédant comme pour Ψ_1 , on obtient

$$\|\Psi_2(X) - \Psi_2(Y)\|_\lambda \leq \frac{C^2}{\lambda} \|X - Y\|_\lambda^2.$$

Comme $\Psi = \Psi_1 + \Psi_2$, on déduit des calculs précédents que $\Psi(X) \in M_2$ si $X \in M_2$ et

$$\|\Psi(X) - \Psi(Y)\|_\lambda \leq \left(\frac{TC^2}{\lambda} + \frac{C^2}{\lambda} \right) \|X - Y\|_\lambda^2.$$

Si λ est suffisamment grand, on a $\frac{TC^2}{\lambda} + \frac{C^2}{\lambda} < 1$ et l'application Ψ est contractante. D'après le théorème du point fixe dans un e.v.n complet, il existe un et un seul $X \in M_2$ tel que $\Psi(X) = X$. Il reste à prouver que deux solutions X et Y sont indistinguables. On sait que $\|X - Y\|_\lambda = 0$ et donc par équivalence des normes, $\|X - Y\|_{M_2} = 0$. On a

$$\begin{aligned} & \sup_{0 \leq t \leq T} |X_t - Y_t| \\ & \leq \sup_{0 \leq t \leq T} \left| \int_0^t (b(s, X_s) - b(s, Y_s)) ds \right| + \sup_{0 \leq t \leq T} \left| \int_0^t (\sigma(s, X_s) - \sigma(s, Y_s)) dB_s \right| \\ & = A + B. \end{aligned}$$

Comme

$$A \leq \sqrt{TC} \sqrt{\int_0^T |X_s - Y_s|^2 ds},$$

on trouve pour tout $\epsilon > 0$,

$$\mathbb{P}(A > \epsilon) \leq \frac{1}{\epsilon} \mathbb{E}(A) \leq \frac{\sqrt{TC}}{\epsilon} \|X - Y\|_{M_2} = 0.$$

De même, en utilisant l'inégalité maximale de Doob

$$\mathbb{P}(B^2 > \epsilon) \leq \frac{1}{\epsilon} \mathbb{E} \left(\left| \int_0^T (\sigma(s, X_s) - \sigma(s, Y_s)) dB_s \right|^2 \right) \leq \frac{C^2}{\epsilon} \|X - Y\|_{M_2}^2 = 0.$$

On en déduit facilement que $\sup_{0 \leq t \leq T} |X_t - Y_t| = 0$ p.s, ce qui montre l'indistinguabilité. \square

Remarques

1. L'hypothèse 1. du théorème précédent est une condition de non explosion. On peut en fait remplacer (le vérifier) cette hypothèse par

$$\sup_{0 \leq t \leq T} |b(t, 0)| + \sup_{0 \leq t \leq T} |\sigma(t, 0)| < +\infty.$$

2. Sous les hypothèses du théorème précédent, si $X \in M_2$, alors le processus $\left(\int_0^t \sigma(s, X_s) dB_s \right)_{0 \leq t \leq T}$ est une martingale (car on intègre un processus de M_2 par rapport au brownien).
3. L'utilisation de la méthode du point fixe permet de construire une suite d'approximations successives de la solution X . On peut par exemple poser $X_t^{(0)} = 0$ pour $t \in [0, T]$ puis si $n \in \mathbb{N}$,

$$X_t^{(n+1)} = x + \int_0^t b(s, X_s^{(n)}) ds + \int_0^t \sigma(s, X_s^{(n)}) dB_s, \quad t \in [0, T].$$

Alors, $\lim_{n \rightarrow +\infty} \|X^{(n)} - X\|_{M_2} = 0$. En utilisant des arguments similaires à ceux utilisés à la fin de la preuve précédente, on peut aussi montrer que

$$\forall \epsilon > 0, \quad \lim_{n \rightarrow +\infty} \mathbb{P} \left(\sup_{0 \leq t \leq T} |X_t^{(n)} - X_t| > \epsilon \right) = 0.$$

4. Avec un peu plus d'effort, il est possible de montrer que la solution X du théorème précédent est unique tout court (sans se restreindre à l'unicité dans M_2).
5. La solution X du théorème précédent est appelée une solution forte, au sens où elle est adaptée par rapport à la filtration du mouvement brownien B donné au départ. Certaines EDS admettent des solutions dites faibles : il existe un couple (X, B) telle que l'EDS correspondante soit vérifiée (le brownien B n'est pas donné à l'avance ici). Nous n'envisagerons pas ce dernier cas de figure dans ce cours.
6. Si Z est une variable aléatoire de carré intégrable et indépendante de la tribu $\sigma(B_s : s \geq 0)$, on peut prouver l'existence et l'unicité d'une solution de l'EDS

$$\begin{cases} dX_s = b(s, X_s) ds + \sigma(s, X_s) dB_s \\ Y_0 = Z \end{cases} \quad (5.1)$$

La preuve est identique à celle du théorème précédent en changeant la tribu \mathcal{F}_t par

$$\mathcal{F}_t^Z = \sigma(Z, B_s : s \geq 0).$$

Voici une des propriétés fondamentales des solutions des EDS.

Proposition 34 *Sous les hypothèses du théorème précédent, l'unique solution de l'EDS*

$$\begin{cases} dX_t = b(t, X_t)dt + \sigma(t, X_t)dB_t, \\ X_0 = x \end{cases}$$

est un processus de Markov, homogène lorsque les coefficients b et σ ne dépendent pas de t .

Preuve. On rappelle que $(\mathcal{F}_t)_t$ désigne la filtration naturelle du mouvement brownien B . Tout d'abord, on peut remarquer que si $Y \in M_2([0, T])$ et $s, t \in \mathbb{R}_+$ sont tels que $s + t \leq T$ alors

$$\int_t^{t+s} Y_u dB_u = \int_0^s Y_{u+s} d\tilde{B}_u, \quad (5.2)$$

où $\tilde{B}_u = B_{t+u} - B_t$ définit un autre mouvement brownien indépendant de la tribu \mathcal{F}_t (le vérifier pour les processus élémentaires). Soit $t \in [0, T]$ fixé. Nous allons utiliser l'EDS suivante sur l'intervalle $[0, T - t]$:

$$\begin{cases} dY_s = b(t + s, Y_s) ds + \sigma(t + s, Y_s) d\tilde{B}_s \\ Y_0 = Z \end{cases} \quad (5.3)$$

où Z est une variable aléatoire de carré intégrable, indépendante de B . Il est alors possible de monter l'existence d'une fonction F mesurable par rapport à la tribu produit et telle que

$$Y_t(\omega) = F(s, t, Z(\omega), \omega), \text{ p.s.}$$

Soit $(X_s^{x,t})_{0 \leq s \leq T-t}$ la solution de l'EDS (5.3) lorsque la condition initiale $Z = x$ est déterministe. On a alors (p.s)

$$X_s^{x,t}(\omega) = F(s, t, x, \omega).$$

On peut remarquer que $F(s, t, x, \cdot)$ est une variable aléatoire indépendante de \mathcal{F}_t (car mesurable par rapport à la tribu $\sigma(\tilde{B}_s : s \geq 0)$ qui est indépendante de \mathcal{F}_t). Posons $Y_s = X_{t+s}^{x,0}$. En utilisant

(5.2), on peut montrer que Y vérifie l'équation (5.3) mais avec la condition initiale $Y_0 = X_t^{x,0}$. Par unicité de la solution, on a p.s

$$Y_s(\omega) = F\left(s, t, X_t^{x,0}(\omega), \omega\right).$$

Soit $h : \mathbb{R}^n \rightarrow \mathbb{R}$ est une fonction mesurable bornée. Alors d'après les propriétés des espérances conditionnelles, on a

$$\mathbb{E}\left(h\left(X_{t+s}^{x,0}\right) \mid \mathcal{F}_t\right) = \phi_{s,t}\left(X_t^{x,0}\right),$$

où $\phi_{s,t} : \mathbb{R}^d \rightarrow \mathbb{R}$ est la fonction définie par

$$\phi_{s,t}(y) = \mathbb{E}\left(h \circ F(s, t, y, \cdot)\right).$$

Ceci montre la propriété de Markov. Lorsque b et σ ne dépendent pas de t (cas homogène), on voit que $\phi_{s,t}$ ne dépend pas de t et donc que le processus de Markov $\left(X_s^{x,0}\right)_s$ est homogène. \square

5.6 Exemples

5.6.1 Processus d'Ornstein-Uhlenbeck

Ce processus vérifie l'EDS $dX_t = -\theta(X_t - \mu)dt + \sigma dB_t$. Au niveau des paramètres, μ représente un état d'équilibre, $\sigma > 0$ un degré de volatilité et $\theta > 0$ un coefficient de vitesse de retour à la moyenne. Cette EDS vérifie les hypothèses du théorème 16. En fait, on connaît l'expression de la solution. Si $Y_t = X_t \exp(\theta t)$ pour $t \geq 0$, une application de la formule d'Itô donne

$$dY_t = \theta \mu \exp(\theta t) dt + \sigma \exp(\theta t) dB_t.$$

On en déduit l'expression

$$X_t = X_0 \exp(-\theta t) + \mu(1 - \exp(-\theta t)) + \sigma \exp(-\theta t) \int_0^t \exp(\theta s) dB_s.$$

Si $X_0 = x$, on voit que X est un processus gaussien et

$$X_t \sim \mathcal{N}\left((x - \mu) \exp(-\theta t) + \mu, \sigma^2 \exp(-2\theta t) \int_0^t \exp(2\theta s) ds\right).$$

On peut voir que la loi de X_t converge vers la loi gaussienne $\mathcal{N}(\mu, \sigma^2)$ lorsque $t \rightarrow +\infty$.

En physique, une version tridimensionnelle de ce processus permet l'étude de la trajectoire d'une particule dans un milieu avec friction. Plus précisément la vitesse V_t de la particule au temps t vérifie l'EDS

$$dV_t = -\theta V_t + dB_t.$$

La force $F_t = -\theta V_t$ est la force du frottement du fluide ($\theta \in \mathbb{R}$) et l'accroissement du brownien dB_t s'interprète comme la résultante des chocs aléatoires des molécules de fluide environnantes. L'équation précédente est obtenue en appliquant le principe fondamental de la dynamique. La solution est unique et est donnée par

$$V_t = V_0 \exp(-\theta t) + \int_0^t \exp(-\theta(t-s)) dB_s.$$

5.6.2 Le mouvement brownien géométrique

En utilisant la formule d'Itô, on voit que le processus $X_t = x \exp\left(\sigma B_t + \left(\mu - \frac{\sigma^2}{2}\right)t\right)$ vérifie l'EDS

$$dX_t = \mu X_t dB_t + \sigma X_t dB_t.$$

En finance, ce processus définit le modèle de Black-Scholes. Il correspond au processus présenté en 5.2.1 lorsque $a_t = \mu$ pour tout t (taux d'intérêt constant). On peut remarquer que

$$\ln(X_t) = \ln(x) + \left(\mu - \frac{\sigma^2}{2}\right)t + \sigma B_t,$$

ce qui permet d'estimer les paramètres μ et σ^2 en utilisant le maximum de vraisemblance (voir Chapitre 2).

5.6.3 Un modèle pour étudier la suspension d'un véhicule

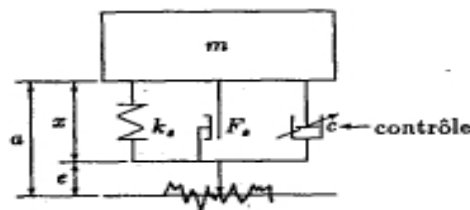


FIGURE 5.1: Suspension d'un véhicule (source : ref [1])

On considère un système de suspension d'un véhicule automobile (voir figure 5.1). Cet exemple est étudié dans [1]. L'objectif est de maximiser le confort de l'utilisateur en faisant en sorte que l'accélération a'' soit petite. En appliquant le principe fondamental de la dynamique, on a l'équation

$$\begin{aligned} ma_t'' &= mx_t'' + me_t'' \\ &= -u_t x_t' - k_t x_t - F_t \times \text{sign}(x_t'), \end{aligned}$$

où

- F est une force de frottement,
- k est un coefficient de raideur de l'amortisseur,
- u est lié au diamètre du trou qui laisse passer l'huile. Cette fonction peut être commandée en fonction des fonctions x et x' (position/vitesse) qui sont mesurées.

La composante e dont les variations sont dues au profil de la route est modélisée à l'aide d'un mouvement brownien $B : e'_t = \sigma B_t$. Les équivalents stochastiques des fonctions x, x' sont respectivement les processus Y et Z . On a alors l'EDS bi-dimensionnelle

$$\begin{cases} dY_t = Z_t dt \\ dZ_t = -\frac{1}{m}u(Y_t, Z_t)Y_t dt - \frac{k_t}{m}Y_t dt - \frac{F_t}{m}\text{sign}(Z_t)dt + \sigma dB_t. \end{cases}$$

La deuxième équation peut s'écrire sous la forme

$$dZ_t = b_2(Y_t, Z_t)dt + \sigma dB_t.$$

A une loi de commande u , on associe la fonction coût

$$J(u) = \lim_{T \rightarrow +\infty} \frac{1}{T} \int_0^T b_2^2(Y_t, Z_t)dt.$$

L'objectif est alors de trouver la commande u qui minimise J (problème directement lié au confort du conducteur).

5.7 Le théorème de Girsanov et le calcul de la vraisemblance d'une diffusion

Pour simplifier, on ne considère dans ce paragraphe que des processus d'Itô à valeurs réelles. Tout ce qui sera évoqué dans cette section se généralise à des processus à valeurs dans \mathbb{R}^d . Le théorème de Girsanov est un outil central pour calculer la vraisemblance des modèles de diffusion. Il permet via un changement de probabilité, d'utiliser la loi du mouvement brownien (la mesure de Wiener) pour établir une expression de la vraisemblance d'une diffusion X . Rappelons qu'en général, la loi d'un n -uplet $(X_{t_1}, \dots, X_{t_n})$ est inconnue car la transition du processus de Markov associé n'est pas disponible. On utilisera plutôt l'absolue continuité de la loi de la variable aléatoire $(X_s)_{0 \leq s \leq T}$ par rapport à une mesure dominante bien choisie.

5.7.1 Martingale exponentielle

On va considérer des mesures de probabilité Q sur (Ω, \mathcal{F}_T) définie par

$$\mathbb{Q}(A) = \int_A \exp\left(\int_0^T b_s dB_s - \frac{1}{2} \int_0^T b_s^2 ds\right) d\mathbb{P}, \quad A \in \mathcal{F}_T,$$

où b est un processus mesurable adapté tel que $\int_0^T b_s^2 ds < +\infty$ p.s. Posons

$$\mathcal{E}(b)_t = \exp\left(\int_0^t b_s dB_s - \frac{1}{2} \int_0^t b_s^2 ds\right), \quad 1 \leq t \leq T.$$

Pour que Q ait bien un sens, il faut que $\mathbb{E}(\mathcal{E}(b)_T) = 1$. On peut remarquer que si $(\mathcal{E}(b)_t)_{0 \leq t \leq T}$ est une martingale alors cette égalité a bien lieu car dans ce cas

$$\mathbb{E}(\mathcal{E}(b)_t) = \mathbb{E}(\mathcal{E}(b)_0) = \mathbb{E}(1) = 1.$$

En écrivant $\mathcal{E}(b)_t = \exp(X_t)$ et en appliquant la formule d'Itô, on obtient

$$d\mathcal{E}(b)_t = \mathcal{E}(b)_t b_t dB_t.$$

On a donc l'égalité

$$\mathcal{E}(b)_t = 1 + \int_0^t \mathcal{E}(b)_s b_s dB_s, \quad (5.4)$$

ce qui ne permet pas d'affirmer que ce processus vérifie la propriété de martingale. Pour montrer la propriété de martingale, il faudrait en fait prouver que

$$\mathbb{E} \left(\int_0^t \mathcal{E}(b)_s^2 b_s^2 ds \right) < +\infty,$$

ce qui n'est pas aisé et de plus non optimal. Nous admettrons que si $(Y_t)_t$ est un processus défini à l'aide d'une intégrale stochastique, alors il existe une suite croissante de temps d'arrêt $(T_n)_n$ de limite $+\infty$ et tels que $(Y_{t \wedge T_n})_t$ est une martingale (en fait ce type de temps d'arrêt permet d'approcher une intégrale stochastique par une intégrale stochastique de processus dans M_2). En appliquant cette propriété au processus $Y_t = \int_0^t \mathcal{E}(b)_s b_s dB_s$ avec l'égalité (5.4), on obtient alors pour tout n et $s < t$,

$$\mathbb{E}(\mathcal{E}(b)_{t \wedge T_n} | \mathcal{F}_s) = \mathcal{E}(b)_{s \wedge T_n}.$$

En utilisant la continuité des trajectoires du processus $\mathcal{E}(b)$ et le lemme de Fatou conditionnel, on obtient

$$\mathcal{E}(b)_s = \lim_{n \rightarrow +\infty} \mathbb{E}(\mathcal{E}(b)_{t \wedge T_n} | \mathcal{F}_s) \geq \mathbb{E}(\mathcal{E}(b)_t | \mathcal{F}_s).$$

Ainsi, le processus $(\mathcal{E}(b))_{0 \leq t \leq T}$ est une sur-martingale. On peut alors montrer que c'est une martingale si et seulement si $\mathbb{E}(\mathcal{E}(b)_T) = 1$ (voir exercice). Finalement, nous avons la proposition suivante.

Proposition 35 *Le processus $(\mathcal{E}(b)_t)_{0 \leq t \leq T}$ est une martingale si et seulement si $\mathbb{E}(\mathcal{E}(b)_T) = 1$. Dans ce cas, on a pour $t \in [0, T]$ et $A \in \mathcal{F}_t$,*

$$\mathbb{Q}(A) = \int_A \mathcal{E}(b)_t d\mathbb{P}.$$

Preuve. L'équivalence a été prouvée avant l'énoncé de la proposition. Montrons la conséquence. Si $0 \leq t \leq T$ et $A \in \mathcal{F}_t \subset \mathcal{F}_T$ alors en utilisant la propriété de martingale

$$\mathbb{Q}(A) = \int_A \mathcal{E}(b)_T d\mathbb{P} = \int_A \mathbb{E}(\mathcal{E}(b)_T | \mathcal{F}_t) d\mathbb{P} = \int_A \mathcal{E}(b)_t d\mathbb{P}. \square$$

Remarque. Montrer la condition $\mathbb{E}(\mathcal{E}(b)_T) = 1$ n'est pas aisée. Il existe des critères permettant de l'obtenir. La condition de Novikov

$$\mathbb{E} \left(\exp \left(\frac{1}{2} \int_0^T b_s^2 ds \right) \right) < +\infty$$

est une condition suffisante pour avoir $\mathbb{E}(\mathcal{E}(b)_T) = 1$. Nous admettrons également que sous les hypothèses du théorème 16, l'EDS

$$\begin{cases} dX_t = b(t, X_t) dt + dB_t, \\ X_0 = x \end{cases}$$

est telle que

$$\mathbb{E}(\mathcal{E}(b)_T) = \mathbb{E} \left(\exp \left(\int_0^T b(s, X_s) dB_s - \frac{1}{2} \int_0^T b(s, X_s)^2 ds \right) \right) = 1.$$

On pourra consulter la référence [2], Chapitre 6 et en particulier l'exemple 3(b) pour une preuve de ce résultat.

5.7.2 Le théorème de Girsanov

Voici maintenant le théorème fondamental permettant de calculer la vraisemblance d'une diffusion.

Théorème 17 *On reprend les notations de la sous-section précédente (en particulier celle de la probabilité \mathbb{Q}). On suppose que $(\mathcal{E}(b)_t)_{0 \leq t \leq T}$ est une martingale. Alors le processus*

$$\tilde{B}_t = B_t - \int_0^t b_s ds, \quad t \in [0, T]$$

est un $(\Omega, \mathcal{F}, (\mathcal{F}_t)_t, \mathbb{Q})$ mouvement brownien.

On voit donc que sous les hypothèses du Théorème 17, un brownien "drifté" redevient un brownien si on change la probabilité \mathbb{P} par une autre probabilité \mathbb{Q} qui est à densité par rapport à \mathbb{P} .

Preuve. Nous n'allons faire la preuve que dans le cas où b est un processus déterministe. Dans ce cas les intégrales stochastiques ont des lois gaussiennes et la démonstration est très simple. Pour une preuve générale, nous référons à [2], section 6.3. Pour prouver le théorème, il suffit de montrer que $\lambda \in \mathbb{R}$ et $s < t$,

$$\mathbb{E}_{\mathbb{Q}} \left(\exp \left(\lambda \left(\tilde{B}_t - \tilde{B}_s \right) \right) \middle| \mathcal{F}_s \right) = \exp \left(\frac{\lambda^2}{2} (t - s) \right), \quad (5.5)$$

où $\mathbb{E}_{\mathbb{Q}}$ désigne l'espérance sous la probabilité \mathbb{Q} . Si $A \in \mathcal{F}_s$, on a

$$\begin{aligned} & \mathbb{E}_{\mathbb{Q}} \left(\mathbf{1}_A \exp \left(\lambda \left(\tilde{B}_t - \tilde{B}_s \right) \right) \right) \\ &= \mathbb{E}_{\mathbb{P}} \left(\mathbf{1}_A \mathcal{E}(b)_t \exp \left(\lambda \left(\tilde{B}_t - \tilde{B}_s \right) \right) \right) \\ &= \mathbb{E}_{\mathbb{P}} \left(\mathbf{1}_A \mathcal{E}(b)_s \exp \left(\int_s^t (b_u + \lambda) dB_u \right) \right) \times \exp \left(-\frac{1}{2} \int_s^t b_u^2 du \right) \times \exp \left(-\lambda \int_s^t b_u du \right) \\ &= \mathbb{E}_{\mathbb{P}} \left(\mathbf{1}_A \mathcal{E}(b)_s \right) \mathbb{E}_{\mathbb{P}} \left(\exp \left(\int_s^t (b_u + \lambda) dB_u \right) \right) \times \exp \left(-\frac{1}{2} \int_s^t b_u^2 du \right) \times \exp \left(-\lambda \int_s^t b_u du \right) \\ &= \mathbb{Q}(A) \times \exp \left(\frac{1}{2} \int_s^t (b_u + \lambda)^2 du \right) \times \exp \left(-\frac{1}{2} \int_s^t b_u^2 du \right) \times \exp \left(-\lambda \int_s^t b_u du \right) \\ &= \mathbb{Q}(A) \exp \left(\frac{1}{2} \lambda^2 (t - s) \right). \end{aligned}$$

Noter que nous avons utilisé l'indépendance de $\int_s^t (b_u + \lambda) dB_u$ et de \mathcal{F}_s . D'après la caractérisation de l'espérance conditionnelle, (5.5) est prouvée. \square

5.7.3 Application au calcul de la vraisemblance d'une diffusion

1. On considère dans un premier temps le modèle diffusion suivant

$$\begin{cases} dX_t = b_{\theta}(t, X_t) dt + dB_t, \\ X_0 = x \end{cases}$$

où θ est un paramètre à valeurs dans un ensemble $\Theta \subset \mathbb{R}^k$. On suppose qu'une trajectoire de $X^{(T)} = (X_s)_{0 \leq s \leq T}$ est observée. Évidemment en pratique, c'est impossible. C'est pour cela que l'estimateur de vraisemblance que nous allons définir est purement théorique. Il

correspond à un cadre idéal où toute l'information antérieure à l'instant T est disponible. En pratique, des méthodes de discrétisation sont nécessaires. Posons

$$M_T = \exp \left(- \int_0^T b_\theta(s, X_s) dB_s - \frac{1}{2} \int_0^T b_\theta^2(s, X_s) ds \right),$$

et $d\mathbb{Q} = M_T d\mathbb{P}$ sur la tribu \mathcal{F}_T . Sous les hypothèses du Théorème 16, le processus $(W_t = B_t + \int_0^t b_\theta(s, X_s) ds)_{0 \leq t \leq T}$ est un mouvement brownien (sous la probabilité \mathbb{Q}) et

$$X_t = x + W_t, \quad 0 \leq t \leq T.$$

On a pour un borélien A de $\mathcal{C}([0, T], \mathbb{R})$,

$$\begin{aligned} \mathbb{P} \left(X^{(T)} \in A \right) &= \int_{X^{(T)} \in A} d\mathbb{P} \\ &= \int_{X^{(T)} \in A} M_T^{-1} d\mathbb{Q} \\ &= \int_{X^{(T)} \in A} \exp \left(\int_0^T b_\theta(s, X_s) dB_s + \frac{1}{2} \int_0^T b_\theta^2(s, X_s) ds \right) d\mathbb{Q} \\ &= \int_{X^{(T)} \in A} \exp \left(\int_0^T b_\theta(s, X_s) dX_s - \frac{1}{2} \int_0^T b_\theta^2(s, X_s) ds \right) d\mathbb{Q}. \end{aligned}$$

Sous \mathbb{Q} , $X_t = x + W_t$ est un mouvement brownien issu de x . Notons μ la loi de $(X_s)_{0 \leq s \leq T}$ sous \mathbb{Q} (ou la loi de $(B_s + x)_{0 \leq s \leq T}$ sous \mathbb{P} , ce qui revient au même). μ est donc une mesure sur $\mathcal{C}([0, T], \mathbb{R})$ et elle ne dépend pas du paramètre θ . On voit alors que

$$\mathbb{P} \left(X^{(T)} \in A \right) = \int_A \exp \left(\int_0^T b_\theta(s, y_s) dy_s(\alpha) - \frac{1}{2} \int_0^T b_\theta^2(s, y_s(\alpha)) ds \right) d\mu(\alpha).$$

Une mesure dominante pour le modèle est donc μ et nous définissons la vraisemblance L par

$$\ln L \left(X^{(T)}, \theta \right) = \int_0^T b_\theta(s, X_s) dX_s - \frac{1}{2} \int_0^T b_\theta^2(s, X_s) ds.$$

L'EMV est donc la maximiseur de cette fonction sur $\theta \in \Theta$.

2. On considère maintenant le modèle,

$$\begin{cases} dX_t = b_\theta(t, X_t) dt + \sigma_\theta(t, X_t) dB_t, \\ X_0 = x \end{cases}$$

Lorsque $X^{(T)}$ est observée, il n'y a en fait pas lieu de paramétriser le coefficient de diffusion σ . En effet on peut montrer que pour tout $t \in [0, T]$, on a la limite suivante en probabilité (en considérant une suite de subdivisions de $[0, t]$ dont le pas tend vers 0) :

$$\int_0^t \sigma_\theta(s, X_s)^2 ds = \lim_{n \rightarrow +\infty} \sum_{i=0}^{p_n-1} \left(X_{t_{i+1}^{(n)}} - X_{t_i^{(n)}} \right)^2. \quad (5.6)$$

On parle de la variation quadratique du processus X (comme pour le mouvement brownien). La connaissance de ces intégrales permet alors de trouver le paramètre θ (ou du moins les coordonnées de θ qui paramétrisent σ). Nous ne prouverons pas (5.6) mais notons que $X_t = A_t + Y_t$ où $A_t = x + \int_0^t b_\theta(s, X_s) ds$ est un processus dont les trajectoires sont

à variations finies (voir Chapitre 2). La partie intégrale stochastique $Y_t = \int_0^t \sigma_\theta(s, X_s) dB_s$ est alors la seule qui contribue à cette variation quadratique et on a

$$\int_0^t \sigma_\theta(s, X_s)^2 ds = \lim_{n \rightarrow +\infty} \sum_{i=0}^{p_n-1} \left(\int_{t_i^{(n)}}^{t_{i+1}^{(n)}} \sigma_\theta(s, X_s) dB_s \right)^2.$$

En utilisant la propriété d'isométrie, on peut remarquer l'égalité des moyennes :

$$\mathbb{E} \left(\int_0^t \sigma_\theta(s, X_s)^2 ds \right) = \mathbb{E} \left(\sum_{i=0}^{p_n-1} \left(\int_{t_i^{(n)}}^{t_{i+1}^{(n)}} \sigma_\theta(s, X_s) dB_s \right)^2 \right).$$

Au vu de ce qui précède, nous supprimons la dépendance en θ du coefficient de diffusion σ . Nous allons alors montrer que \mathbb{P}_{X_T} a une densité par rapport à la loi μ de la diffusion

$$\begin{cases} dY_t = \sigma(t, Y_t) dB_t, & 0 \leq t \leq T \\ Y_0 = x \end{cases}$$

Définissons

$$\tilde{B}_t = B_t + \int_0^t \frac{b_\theta(s, X_s)}{\sigma(s, X_s)} ds, \quad t \in [0, T],$$

de sorte que

$$dX_t = \sigma(t, X_t) d\tilde{B}_t.$$

En posant $d\mathbb{Q} = M_T d\mathbb{P}$ sur \mathcal{F}_T avec

$$M_T = \exp \left(- \int_0^T \frac{b_\theta(s, X_s)}{\sigma(s, X_s)} dB_s - \frac{1}{2} \int_0^T \frac{b_\theta^2(s, X_s)}{\sigma^2(s, X_s)} ds \right).$$

Sous réserve d'intégrabilité du quotient $\frac{b_\theta}{\sigma}$, on voit que sous la probabilité \mathbb{Q} , \tilde{B} est un mouvement brownien et que X suit la loi μ (même loi que Y sous \mathbb{P}). Si A est borélien de $\mathcal{C}([0, T], \mathbb{R})$,

$$\begin{aligned} \mathbb{P} \left(X^{(T)} \in A \right) &= \int_{\{X^T \in A\}} M_T^{-1} d\mathbb{Q} \\ &= \int_{\{X^T \in A\}} \exp \left(\int_0^T \frac{b_\theta(s, X_s)}{\sigma(s, X_s)} d\tilde{B}_s - \frac{1}{2} \int_0^T \frac{b_\theta^2(s, X_s)}{\sigma^2(s, X_s)} ds \right) d\mathbb{Q} \\ &= \int_A \exp \left(\int_0^T \frac{b_\theta(s, X_s)}{\sigma^2(s, X_s)} dX_s - \frac{1}{2} \int_0^T \frac{b_\theta^2(s, X_s)}{\sigma^2(s, X_s)} ds \right) d\mathbb{Q} \\ &= \int_A \exp \left(\int_0^T \frac{b_\theta(s, y_s)}{\sigma^2(s, y_s)} dy_s(\alpha) - \frac{1}{2} \int_0^T \frac{b_\theta^2(s, y_s(\alpha))}{\sigma^2(s, y_s(\alpha))} ds \right) d\mu(\alpha). \end{aligned}$$

La log-vraisemblance est alors donnée par

$$\ln L \left(X^{(T)}, \theta \right) = \int_0^T \frac{b_\theta(s, X_s)}{\sigma^2(s, X_s)} dX_s - \frac{1}{2} \int_0^T \frac{b_\theta^2(s, X_s)}{\sigma^2(s, X_s)} ds.$$

Nous n'avons pas justifié la validité du changement de probabilité (i.e $\mathbb{E}(M_T) = 1$). Sa validité est en fait assurée lorsque les hypothèses du théorème 16 sont satisfaites et lorsque $\delta = \inf_{s,x} \sigma(s, x) > 0$.

5.7.4 L'EMV en pratique

Nous n'étudierons pas dans ce cours les propriétés asymptotiques de l'EMV. On pourra se reporter à la référence [4] pour un étude détaillée de ces problèmes. Sous certaines hypothèses de régularité, $\hat{\theta}_T = \arg \max_{\theta} L(X^{(T)}, \theta)$ a de bonnes propriétés asymptotiques. Toutefois, l'observation de toute une trajectoire entre 0 et T n'est pas réaliste en pratique. Si on observe uniquement les variables $X_{t_1}, X_{t_2}, \dots, X_{t_n}$, il n'est pas possible en général de déterminer la loi du n -échantillon car la transition $\mathbb{P}(X_{t_{i+1}} | X_{t_i})$ est inconnue (ce sont ces transitions qui permettent de reconstruire la loi jointe en utilisant la propriété de Markov). De plus, dans le cas d'observations discrétisées, il n'est plus possible de connaître exactement le coefficient de diffusion et l'estimation de ses paramètres redevient alors d'actualité. On essaie alors de montrer que certaines versions discrétisées et/ou approchées de l'EMV conservent les mêmes propriétés asymptotiques. Mentionnons deux approches.

- La première idée consiste à discrétiser la vraisemblance. Ainsi, au lieu de maximiser la vraisemblance de la sous-section précédente, on maximisera

$$\sum_{i=2}^n \left\{ \frac{b_{\theta} \left(t_{i-1}^{(n)}, X_{t_{i-1}^{(n)}}, \theta \right)}{\sigma_{\theta}^2 \left(t_{i-1}^{(n)}, X_{t_{i-1}^{(n)}} \right)} \left(X_{t_i^{(n)}} - X_{t_{i-1}^{(n)}} \right) - \frac{1}{2} \frac{b_{\theta}^2 \left(t_{i-1}^{(n)}, X_{t_{i-1}^{(n)}}, \theta \right)}{\sigma_{\theta}^2 \left(t_{i-1}^{(n)}, X_{t_{i-1}^{(n)}} \right)} \left(t_i^{(n)} - t_{i-1}^{(n)} \right) \right\}.$$

- Une deuxième méthode consiste à utiliser une discrétisation du modèle (approximation par le schéma d'Euler). On suppose alors que les variables aléatoires $X_{t_1}, X_{t_2}, \dots, X_{t_n}$ satisfont l'équation

$$X_{t_{i+1}} - X_{t_i} = b_{\theta}(t_i, X_{t_i})(t_{i+1} - t_i) + \sigma_{\theta}(t_i, X_{t_i})(B_{t_{i+1}} - B_{t_i}).$$

Ce modèle approché est alors une chaîne de Markov (non homogène) dont la transition est une distribution gaussienne :

$$p_i(x_{t_{i+1}} | x_{t_i}) = \frac{1}{\sqrt{2\pi\sigma_{\theta}^2(t_i, X_{t_i})(t_{i+1} - t_i)}} \exp\left(-\frac{x_{t_{i+1}} - x_{t_i} - b_{\theta}(t_i, X_{t_i})(t_{i+1} - t_i)}{2\sigma_{\theta}^2(t_i, X_{t_i})(t_{i+1} - t_i)}\right).$$

On obtient alors facilement la densité jointe en utilisant la formule des conditionnements successifs. Bien sûr on n'obtiendra pas la vraisemblance associée au vrai modèle mais seulement une "pseudo-vraisemblance".

Bibliographie

- [1] Campillo, Bellizi, Bouc, R., Pardoux, E. (1988) *Contrôle optimal semi-actif de suspension de véhicule*. In A. Bensoussan and J.L. Lions editors, vol. 111 of Lecture Notes in Control and Information Sciences, p 689-699, INRIA. Springer Verlag.
- [2] Lipster, R. S., Shiryaev, A. N. *Statistics of random processes*. Vol I. Springer Verlag (2001), deuxième édition.
- [3] Oksendal, B. *Stochastic differential equations*. Springer.
- [4] Prakasa Rao, B. L. S. (1999) *Statistical Inference for Diffusion Type Processes*. Arnold, London and Oxford University Press, New York.

5.8 Exercices

EXERCICE 15

1. Si B est un mouvement Brownien, montrer que le processus X défini par $X_t = B_t^3 - 3tB_t$ est une martingale par rapport à $\sigma(B_s/s \leq t)$.
2. Si N est un processus de Poisson de paramètre λ , montrer que le processus Y défini par $Y_t = (N_t - \lambda t)^2 - \lambda t$ est une martingale par rapport à $\sigma(N_s/s \leq t)$.
3. Tout processus d'espérance constante est-il une martingale ?

EXERCICE 16 Montrer qu'une sur-martingale $(X_s)_{0 \leq s \leq T}$ est une martingale si et seulement si $\mathbb{E}(X_T) = \mathbb{E}(X_0)$.

EXERCICE 17 [intégrale de Stratonovich]

Si B est un mouvement Brownien, déterminer

$$\lim_{n \rightarrow \infty} \sum_j B_{s_j^{(n)}} \left(B_{t_{j+1}^{(n)}} - B_{t_j^{(n)}} \right), \quad s_j^{(n)} = \frac{t_j^{(n)} + t_{j+1}^{(n)}}{2}$$

dans \mathbb{L}^2 où $(t_j^{(n)})_j$ est une suite de subdivisions de $[0, T]$ dont le pas tend vers 0. Comparer avec l'intégrale stochastique $\int_0^T B_s dB_s$.

EXERCICE 18 (extrait de l'examen de 2008)

Étant donné un mouvement Brownien $(B_t)_{t \geq 0}$, on définit le processus $(X_t)_{t \geq 0}$ par :

$$X_t = \int_0^{\sqrt{t}} \sqrt{2s} dB_s.$$

1. Montrer que ce processus est gaussien. Calculer sa moyenne et sa variance.

2. Montrer que X est un mouvement Brownien.

EXERCICE 19 Soit X l'intégrale d'Itô définie $dX_t = V_t dB_t$, $0 \leq t \leq T$ avec $\mathbb{E} \left(\int_0^T V_s^2 ds \right) < \infty$. Montrer que le processus M défini par

$$M_t = X_t^2 - \int_0^t |V_s|^2 ds, \quad t \in [0, T],$$

est une martingale.

EXERCICE 20 Soit f une fonction de carré intégrable sur $[0, T]$. Montrer que le processus défini pour $t \in [0, T]$ par

$$M_t = \exp \left(\int_0^t f(s) dB_s - \frac{1}{2} \int_0^t f^2(s) ds \right),$$

est une martingale. On s'inspirera du cas $f \equiv 1$.

EXERCICE 21 En utilisant la formule de Itô, montrer que les processus suivants sont des martingales (B est un mouvement Brownien réel) :

1. $X_t = e^{\frac{1}{2}t} \cos(B_t)$.
2. $X_t = e^{\frac{1}{2}t} \sin(B_t)$.
3. $X_t = (B_t + t)e^{-B_t - \frac{1}{2}t}$.

EXERCICE 22 (extrait de l'examen de 2006) Montrer que le processus X défini par

$$X_t = \exp \left(\int_0^t a(s) ds \right) \left[x + \int_0^t b(s) \exp \left(- \int_0^s a(u) du \right) dB_s \right]$$

est solution de l'équation différentielle stochastique :

$$dX_t = a(t)X_t dt + b(t)dB_t$$

avec la condition initiale $X_0 = x$.

EXERCICE 23 Soit B un mouvement Brownien sur $[0, 1]$ et θ une fonction \mathcal{C}^1 sur $[0, 1]$, à valeurs réelles. Si η est le processus défini par

$$\eta_t = B_t + \theta(t), \quad t \in [0, 1].$$

Trouver la densité de μ_η par rapport à μ_B .

EXERCICE 24 Soit B un mouvement brownien et W le processus défini par

$$W_t = B_t + \mu t, \quad t \in \mathbb{R}_+,$$

où μ est un nombre réel. En appliquant le théorème de Girsanov et en utilisant l'exercice 3, trouver la loi de $\sup_{t \in [0, 1]} W_t$.

EXERCICE 25 On considère l'équation différentielle stochastique

$$\begin{cases} dX_t = f(t, X_t) dt + c(t)X_t dB_t \\ X_0 = x \end{cases}$$

où $f : \mathbb{R} \times \mathbb{R} \rightarrow \mathbb{R}$ et $c : \mathbb{R} \rightarrow \mathbb{R}$ sont deux fonctions continues.

1. On pose $F_t = \exp\left(-\int_0^t c(s)dB_s + \frac{1}{2}\int_0^t c^2(s)ds\right)$. On suppose que X est une solution de l'EDS ci-dessus et on pose $Y_t = F_t X_t$. En utilisant la formule d'Itô, montrer qu'une trajectoire $t \mapsto Y_t(\omega)$ vérifie l'équation différentielle déterministe

$$\frac{dY_t(\omega)}{dt} = F_t(\omega) \cdot f\left(t, F_t^{-1}(\omega)Y_t(\omega)\right); \quad Y_0 = x.$$

2. Utiliser la question précédente pour résoudre l'équation différentielle stochastique

$$dX_t = X_t^\gamma dt + \alpha X_t dB_t; \quad X_0 = x > 0,$$

où α et γ sont des nombres réels donnés.

EXERCICE 26 Soit B un mouvement brownien réel. On considère le processus d'Ornstein-Uhlenbeck solution de l'équation différentielle stochastique

$$dX_t = \theta X_t dt + dB_t; \quad X_0 = 0.$$

1. On suppose que $(X_s)_{0 \leq s \leq T}$ est observée pour un réel $T > 0$ donné. Donner une expression du maximum de vraisemblance du paramètre θ .
2. Démontrer que $X_t = \int_0^t \exp(\theta(t-s)) dB_s$ et calculer la covariance $\text{Cov}(X_t, X_s)$ pour tout $s, t \geq 0$.
3. On suppose que seules les coordonnées X_{t_1}, \dots, X_{t_n} sont observés, avec $0 < t_1 < t_2 < \dots < t_n$. Comment peut-on calculer la vraisemblance associée à ces observations ?

EXERCICE 27 (*EDS du pont Brownien*)

Soit B un mouvement Brownien. Montrer que le processus X défini par

$$X_t = (1-t) \int_0^t \frac{1}{1-s} dB_s, \quad 1 > T \geq t \geq 0,$$

satisfait l'équation différentielle stochastique

$$\begin{cases} dX_t = -\frac{X_t}{1-t} dt + dB_t \\ X_0 = 0 \end{cases}$$

Montrer aussi que X a la même loi que le processus $(B_t - tB_1)_{0 \leq t < 1}$.

EXERCICE 28 Montrer que le processus défini pour $t \geq 0$ par

$$X_t = \frac{\exp\left\{(rK - \frac{1}{2}\beta^2)t + \beta B_t\right\}}{x^{-1} + r \int_0^t \exp\left\{(rK - \frac{1}{2}\beta^2)s + \beta B_s\right\} ds},$$

avec $x > 0$ est solution de l'équation différentielle stochastique

$$dX_t = rX_t(K - X_t) dt + \beta X_t dB_t, \quad X_0 = x.$$