

COURS DE PROBABILITES ET CALCUL STOCHASTIQUE

Olivier Lévêque, EPFL

Semestre d'hiver 2004-2005

Avertissement

Ce polycopié a été établi sur la base des notes du cours de probabilité et calcul stochastique donné dans le cadre du cycle d'études postgrades en ingénierie mathématique à l'EPFL. Il ne constitue en aucun cas une référence de calcul stochastique, étant donné les nombreuses lacunes et imprécisions qui le parsèment. Pour une exposition plus complète, le lecteur est renvoyé à la bibliographie.

Le but du cours est de donner une première idée du sujet à des étudiants provenant d'horizons différents (sur un court laps de temps de 30 périodes concentrées sur deux mois). Le niveau requis pour suivre le cours est celui d'un étudiant ayant accompli une formation d'ingénieur et déjà suivi un cours de base de probabilités.

Pour bénéficier pleinement de la lecture de ce polycopié, il est nécessaire d'effectuer en parallèle les exercices qui l'accompagnent. Ceux-ci sont disponibles sur la page web :

<http://lthiwww.epfl.ch/~leveque/Proba/>

Remerciements

Deux personnes ont fortement contribué à la rédaction de ce polycopié. Je désirerais remercier tout d'abord Raffi Aghayekian, étudiant de la volée 2002-2003, pour sa prise de notes appliquée qui a constitué la première version manuscrite de ces notes de cours. Mes remerciements vont ensuite à Erika Gindraux qui a patiemment dactylographié la première version de ces notes.

Un grand merci va également à tous les étudiants pour leur motivation à apprendre les probabilités et leurs nombreuses questions et commentaires qui ont contribué à l'élaboration de ces notes. Et finalement, merci à Benjamin Bergé pour avoir pris le temps de relire le manuscrit.

Table des matières

1	Rappels de probabilité	1
1.1	Espace de probabilité	1
1.2	Variable aléatoire	3
1.3	Loi d'une variable aléatoire	4
1.4	Espérance d'une variable aléatoire	6
1.5	Indépendance	9
1.5.1	Indépendance d'événements	9
1.5.2	Indépendance de tribus	9
1.5.3	Indépendance de variables aléatoires	10
1.6	Espérance conditionnelle	10
1.6.1	Conditionnement par rapport à un événement $B \in \mathcal{F}$	10
1.6.2	Conditionnement par rapport à une v.a. discrète Y (à valeurs dans D dénombrable)	11
1.6.3	Conditionnement par rapport à une v.a. Y continue? Cas général?	11
1.6.4	Conditionnement par rapport à une tribu \mathcal{G}	11
1.7	Convergences de suites de variables aléatoires	13
1.8	Processus aléatoire à temps discret	16
1.9	Vecteur aléatoire	21
2	Calcul stochastique	24
2.1	Processus aléatoire à temps continu	24
2.1.1	Processus à accroissements indépendants et stationnaires	24
2.1.2	Mouvement brownien standard	25
2.1.3	Processus gaussien	27
2.1.4	Processus de Markov	28
2.1.5	Martingale à temps continu	29
2.2	Intégrale de Riemann-Stieltjes	32
2.3	Variation quadratique	34
2.4	Intégrale stochastique (ou intégrale d'Itô)	38
2.5	Formules d'Itô	47
2.5.1	Formules de base	47
2.5.2	Processus d'Itô (ou "semi-martingale continue")	50
2.5.3	Retour à l'intégrale de Stratonovich	51
2.5.4	Formules d'Itô généralisées	53
2.5.5	Formule d'intégration par parties (IPP)	54
2.6	Equations différentielles stochastiques (EDS)	55
2.6.1	Equations homogènes en temps	55
2.6.2	Equations inhomogènes en temps	59

2.6.3	Equations linéaires	60
2.6.4	Solution faible	61
2.6.5	Martingale exponentielle	62
2.7	Théorème de Girsanov	63
2.8	Liens EDS \leftrightarrow EDP	69
2.8.1	Lien mouvement brownien \leftrightarrow équation(s) de la chaleur	69
2.8.2	Lien EDS \leftrightarrow EDP paraboliques (formule de Feynman-Kac)	71
2.9	Processus multidimensionnels	76
2.10	Analyse et simulation numériques des EDS	84
2.11	Pour conclure	89
Bibliographie		94

1 Rappels de probabilité

1.1 Espace de probabilité

Cours 1

Un espace de probabilité est un triplet $(\Omega, \mathcal{F}, \mathbb{P})$ où :

- Ω est un ensemble,
- \mathcal{F} est une tribu (ou σ -algèbre) sur Ω ,
- \mathbb{P} est une (mesure de) probabilité sur (Ω, \mathcal{F}) .

Définition 1.1.1. Une tribu (ou σ -algèbre) sur Ω est une famille \mathcal{F} de sous-ensembles de Ω (appelés “événements”) tels que

$$\begin{cases} i) & \emptyset \in \mathcal{F}, \\ ii) & A \in \mathcal{F} \Rightarrow A^c \in \mathcal{F}, \\ iii) & (A_n)_{n=1}^\infty \subset \mathcal{F} \Rightarrow \bigcup_{n=1}^\infty A_n \in \mathcal{F}. \end{cases}$$

En particulier : $A, B \in \mathcal{F} \Rightarrow A \cup B \in \mathcal{F}$.

De même, $A, B \in \mathcal{F} \Rightarrow A \cap B \in \mathcal{F}$ (cf. exercices pour d’autres propriétés).

Exemple 1.1.2. - Soit $\Omega = \{1, \dots, 6\}$. On peut définir plusieurs tribus sur Ω :

$\mathcal{F} = \mathcal{P}(\Omega) = \{\emptyset, \{1\}, \{2\}, \dots, \{1, 2\}, \dots, \Omega\}$ = tribu complète (la plus grande),
 $\mathcal{F}_0 = \{\emptyset, \Omega\}$ = tribu triviale (la plus petite) (\emptyset = évén. impossible, Ω = évén. arbitraire),
 $\mathcal{F}_1 = \{\emptyset, \{1\}, \{2, \dots, 6\}, \Omega\}$, $\mathcal{F}_2 = \{\emptyset, \{1, 3, 5\}, \{2, 4, 6\}, \Omega\}$, etc.

- Soit $\Omega = [0, 1]$ et I_1, \dots, I_n une famille d’intervalles formant une partition de Ω . La famille de sous-ensembles définie par

$$\mathcal{G} = \{\emptyset, I_1, I_2, \dots, I_1 \cup I_2, \dots, I_1 \cup I_2 \cup I_3, \dots, \Omega\}$$

est une tribu sur Ω .

Définition 1.1.3. Soit $\mathcal{A} = \{A_i, i \in I\}$ une famille de sous-ensembles de Ω . Alors la tribu engendrée par \mathcal{A} est la plus petite tribu sur Ω qui contient tous les sous-ensembles $A_i, i \in I$. Elle est notée $\sigma(\mathcal{A})$. (NB : l’ensemble I n’est pas forcément dénombrable.)

Exemple 1.1.4. Reprenons l’exemple 1.1.2.

- Soit $\Omega = \{1, \dots, 6\}$. Si $\mathcal{A}_1 = \{\{1\}\}$, alors $\sigma(\mathcal{A}_1) = \mathcal{F}_1$. Si $\mathcal{A}_2 = \{\{1, 3, 5\}\}$, alors $\sigma(\mathcal{A}_2) = \mathcal{F}_2$.
Et si $\mathcal{A} = \{\{1, 3, 5\}, \{1, 2, 3\}\}$, alors $\sigma(\mathcal{A}) = ?$ (cf. exercices).

- Soit $\Omega = [0, 1]$. Si $\mathcal{A} = \{I_1, \dots, I_n\}$, alors $\sigma(\mathcal{A}) = \mathcal{G}$.

Définition 1.1.5. Soit $\Omega = [0, 1]$. La tribu borélienne sur $[0, 1]$ est la tribu engendrée par la famille de sous-ensembles $\mathcal{A} = \{]a, b[: 0 \leq a < b \leq 1\} = \{\text{intervalles ouverts dans } [0, 1]\}$. Elle est notée $\mathcal{B}([0, 1])$. Elle contient un très grand nombre de sous-ensembles de $[0, 1]$, mais pas tous.

Remarque 1.1.6. - Pour Ω ensemble fini, on choisit souvent $\mathcal{F} = \mathcal{P}(\Omega)$.
- Pour $\Omega \subset \mathbb{R}$ ou $\Omega \subset \mathbb{R}^n$, on choisit souvent $\mathcal{F} = \mathcal{B}(\Omega)$.

Définition 1.1.7. Une sous-tribu de \mathcal{F} est une tribu \mathcal{G} telle que si $A \in \mathcal{G}$ alors $A \in \mathcal{F}$. On note $\mathcal{G} \subset \mathcal{F}$.

Exemple 1.1.8. Reprenons l'exemple 1.1.2. On a $\mathcal{F}_0 \subset \mathcal{F}_1 \subset \mathcal{F}$, $\mathcal{F}_0 \subset \mathcal{F}_2 \subset \mathcal{F}$, mais pas $\mathcal{F}_1 \subset \mathcal{F}_2$, ni $\mathcal{F}_2 \subset \mathcal{F}_1$.

Remarque importante :

Il est toujours vrai que $A \in \mathcal{G}$ et $\mathcal{G} \subset \mathcal{F} \Rightarrow A \in \mathcal{F}$.

Mais il est faux de dire que $A \subset B$ et $B \in \mathcal{F} \Rightarrow A \in \mathcal{F}$. Contre-exemple :

$$\{1\} \subset \{1, 3, 5\}, \{1, 3, 5\} \in \mathcal{F}_2, \quad \text{mais } \{1\} \notin \mathcal{F}_2.$$

Définition 1.1.9. Soit \mathcal{F} une tribu sur Ω . Une (mesure de) probabilité sur (Ω, \mathcal{F}) est une application $\mathbb{P} : \mathcal{F} \rightarrow [0, 1]$ telle que

$$\begin{cases} i) & \mathbb{P}(\emptyset) = 0 \text{ et } \mathbb{P}(\Omega) = 1, \\ ii) & (A_n)_{n=1}^\infty \subset \mathcal{F} \text{ disjoints (i.e. } A_n \cap A_m = \emptyset, \forall n \neq m) \Rightarrow \mathbb{P}(\cup_{n=1}^\infty A_n) = \sum_{n=1}^\infty \mathbb{P}(A_n). \end{cases}$$

En particulier : $A, B \in \mathcal{F}$ et $A \cap B = \emptyset \Rightarrow \mathbb{P}(A \cup B) = \mathbb{P}(A) + \mathbb{P}(B)$. De plus :

$$\begin{cases} i) & \text{Si } (A_n)_{n=1}^\infty \subset \mathcal{F}, A_n \subset A_{n+1} \text{ et } \cup_{n=1}^\infty A_n = A, \text{ alors } \lim_{n \rightarrow \infty} \mathbb{P}(A_n) = \mathbb{P}(A), \\ ii) & \text{Si } (A_n)_{n=1}^\infty \subset \mathcal{F}, A_n \supset A_{n+1} \text{ et } \cap_{n=1}^\infty A_n = A, \text{ alors } \lim_{n \rightarrow \infty} \mathbb{P}(A_n) = \mathbb{P}(A). \end{cases}$$

Pour d'autres propriétés, cf. exercices.

Exemple 1.1.10. Soient $\Omega = \{1, \dots, 6\}$, $\mathcal{F} = \mathcal{P}(\Omega)$. On définit :

- $\mathbb{P}_1(\{i\}) = \frac{1}{6} \forall i$ (mesure de probabilité associée à un dé équilibré).

Dans ce cas, on voit p. ex. que $\mathbb{P}_1(\{1, 3, 5\}) = \mathbb{P}_1(\{1\}) + \mathbb{P}_1(\{3\}) + \mathbb{P}_1(\{5\}) = \frac{1}{6} + \frac{1}{6} + \frac{1}{6} = \frac{1}{2}$.

- $\mathbb{P}_2(\{i\}) = 0 \forall i \leq 5$, $\mathbb{P}_2(\{6\}) = 1$ (mesure de probabilité associée à un dé pipé).

Définition 1.1.11. Soient $\Omega = [0, 1]$ et $\mathcal{F} = \mathcal{B}([0, 1])$. On appelle mesure de Lebesgue sur $[0, 1]$ la mesure de probabilité définie par

$$\mathbb{P}(]a, b[) = b - a, \quad \forall 0 \leq a < b \leq 1.$$

\mathbb{P} n'est définie a priori que sur les intervalles, mais est uniquement extensible à tout ensemble borélien $B \in \mathcal{B}([0, 1])$. Elle est notée $\mathbb{P}(B) = |B|$, $B \in \mathcal{B}([0, 1])$.

En utilisant la propriété (ii) ci-dessus, on déduit que pour tout $x \in [0, 1]$:

$$|\{x\}| = \lim_{n \rightarrow \infty} |]x - \frac{1}{n}, x + \frac{1}{n}[| = \lim_{n \rightarrow \infty} \frac{2}{n} = 0.$$

Généralisation à n dimensions : Soit $\Omega = [0, 1]^n$.

- Tribu borélienne : $\mathcal{B}(\Omega) = \sigma(\mathcal{A})$, où $\mathcal{A} = \{]a_1, b_1[\times]a_2, b_2[\times \cdots \times]a_n, b_n[, 0 \leq a_i < b_i \leq 1 \}$. \mathcal{A} est la famille des “rectangles” dans Ω .

- Mesure de Lebesgue : $\mathbb{P}(]a_1, b_1[\times]a_2, b_2[\times \cdots \times]a_n, b_n[) = (b_1 - a_1)(b_2 - a_2) \cdots (b_n - a_n)$. Comme dans le cas uni-dimensionnel, \mathbb{P} n’est définie a priori que sur certains ensembles (les rectangles), mais est uniquement extensible à tout $B \in \mathcal{B}(\Omega)$ (p.ex. B = disque, ovale...).

Notation : $\mathbb{P}(B) = |B|$, pour $B \in \mathcal{B}(\Omega)$.

1.2 Variable aléatoire

Définition 1.2.1. Soit $(\Omega, \mathcal{F}, \mathbb{P})$ un espace de probabilité. Une variable aléatoire (souvent abrégé v.a. par la suite) est une application $X : \Omega \rightarrow \mathbb{R}$ telle que

$$\{\omega \in \Omega : X(\omega) \in B\} = \{X \in B\} \in \mathcal{F}, \quad \forall B \in \mathcal{B}(\mathbb{R}).$$

Proposition 1.2.2. X est une v.a. ssi $\{\omega \in \Omega : X(\omega) \leq t\} \in \mathcal{F}, \forall t \in \mathbb{R}$.

Remarque 1.2.3. - X est aussi dite une fonction (ou v.a.) \mathcal{F} -mesurable.

- Si $\mathcal{F} = \mathcal{P}(\Omega)$, alors X est toujours \mathcal{F} -mesurable.

- Si $\Omega = \mathbb{R}$ et $\mathcal{F} = \mathcal{B}(\mathbb{R})$, alors X est dite une fonction borélienne.

Définition 1.2.4. Pour $A \subset \Omega$, on pose $1_A(\omega) = \begin{cases} 1 & \text{si } \omega \in A, \\ 0 & \text{sinon.} \end{cases}$

On vérifie que la v.a. 1_A est \mathcal{F} -mesurable ssi $A \in \mathcal{F}$.

Exemple 1.2.5. Soit $(\Omega, \mathcal{F}, \mathbb{P})$ l’espace de probabilité du dé équilibré (cf. exemple 1.1.10).

$$X_1(\omega) = \omega : \mathbb{P}(\{\omega \in \Omega : X_1(\omega) = i\}) = \mathbb{P}(\{i\}) = \frac{1}{6}.$$

$$X_2(\omega) = 1_{\{1,3,5\}}(\omega) : \mathbb{P}(\{\omega \in \Omega : X_2(\omega) = 1\}) = \mathbb{P}(\{1, 3, 5\}) = \frac{1}{2}.$$

Soit $\mathcal{F} = \mathcal{P}(\Omega)$. X_1 et X_2 sont toutes deux \mathcal{F} -mesurables.

Soit $\mathcal{F}_2 = \{\emptyset, \{1, 3, 5\}, \{2, 4, 6\}, \Omega\}$. Seule X_2 est \mathcal{F}_2 -mesurable ; X_1 ne l’est pas. En effet :

$$\{\omega \in \Omega : X_2(\omega) = 1\} = \{1, 3, 5\} \in \mathcal{F}_2 \quad \text{et} \quad \{\omega \in \Omega : X_2(\omega) = 0\} = \{2, 4, 6\} \in \mathcal{F}_2,$$

tandis que $\{\omega \in \Omega : X_1(\omega) = 1\} = \{1\} \notin \mathcal{F}_2$.

Définition 1.2.6. La tribu engendrée par une famille de v.a. $\{X_i, i \in I\}$ sur $(\Omega, \mathcal{F}, \mathbb{P})$ est définie par

$$\sigma(X_i, i \in I) = \sigma(\{X_i \in B\}, i \in I, B \in \mathcal{B}(\mathbb{R})) = \sigma(\{X_i \leq t\}, i \in I, t \in \mathbb{R}).$$

Exemple 1.2.7. Reprenons l'exemple précédent : $\sigma(X_1) = \mathcal{F} = \mathcal{P}(\Omega)$, $\sigma(X_2) = \mathcal{F}_2 \neq \mathcal{P}(\Omega)$.

Proposition 1.2.8. Si $g : \mathbb{R} \rightarrow \mathbb{R}$ est borélienne et $X : \Omega \rightarrow \mathbb{R}$ est une v.a., alors $g(X)$ est une v.a.

Démonstration. Soit $B \in \mathcal{B}(\mathbb{R})$. On a

$$\{\omega \in \Omega : g(X(\omega)) \in B\} = \{\omega \in \Omega : X(\omega) \in g^{-1}(B)\}.$$

Or $g^{-1}(B) = \{x \in \mathbb{R} : g(x) \in B\} \in \mathcal{B}(\mathbb{R})$, car g est borélienne. Comme X est une v.a., on en déduit que $\{\omega \in \Omega : X(\omega) \in g^{-1}(B)\} \in \mathcal{F}$, et donc finalement que $g(X)$ est une v.a. \square

Proposition 1.2.9. Toute fonction continue est borélienne (et pratiquement toute fonction discontinue l'est aussi!).

1.3 Loi d'une variable aléatoire

Définition 1.3.1. La loi d'une v.a. X est l'application $\mu_X : \mathcal{B}(\mathbb{R}) \rightarrow [0, 1]$ définie par

$$\mu_X(B) = \mathbb{P}(\{X \in B\}), \quad B \in \mathcal{B}(\mathbb{R}).$$

NB : $(\mathbb{R}, \mathcal{B}(\mathbb{R}), \mu_X)$ forme un nouvel espace de probabilité!

Exemple 1.3.2. - Soit $\Omega = \{1, \dots, 6\}$, $\mathcal{F} = \mathcal{P}(\Omega)$, $\mathbb{P}(\{i\}) = \frac{1}{6}$, $\forall i$.
 $X_1(\omega) = \omega$, $\mu_X(\{i\}) = \mathbb{P}(\{X = i\}) = \mathbb{P}(\{i\}) = \frac{1}{6}$.

- Soit $\Omega = \{1, \dots, 6\} \times \{1, \dots, 6\}$, $\mathcal{F} = \mathcal{P}(\Omega)$, $\mathbb{P}(\{(i, j)\}) = \frac{1}{36}$, $\forall (i, j) \in \Omega$.
 $X(\omega) = X(\omega_1, \omega_2) = \omega_1 + \omega_2$. On a alors, p.ex :
 $\mu_X(\{7\}) = \mathbb{P}(\{X = 7\}) = \mathbb{P}(\{(1, 6), (2, 5), (3, 4), (4, 3), (5, 2), (6, 1)\}) = 6 \times \frac{1}{36} = \frac{1}{6}$.

Fonction de répartition d'une variable aléatoire

Définition 1.3.3. La fonction de répartition d'une v.a. X est l'application $F_X : \mathbb{R} \rightarrow [0, 1]$ définie par

$$F_X(t) = \mathbb{P}(\{X \leq t\}) = \mu_X(]-\infty, t]), \quad t \in \mathbb{R}.$$

Proposition 1.3.4. La donnée de F_X équivaut à celle de μ_X .

Cette dernière proposition est à rapprocher de la proposition 1.2.2.

Deux types particuliers de variables aléatoires

A) Variable aléatoire discrète :

X prend ses valeurs dans un ensemble D dénombrable ($X(\omega) \in D, \forall \omega \in \Omega$). Dans ce cas, on a : $\sigma(X) = \sigma(\{X = x\}, x \in D)$, $p(x) = \mathbb{P}(\{X = x\}) \geq 0$ et $\sum_{x \in D} p(x) = \mathbb{P}(\{x \in D\}) = 1$. De plus, $F_X(t) = \sum_{x \in D: x \leq t} p(x)$.

B) Variable aléatoire continue :

$\mathbb{P}(\{X \in B\}) = 0$ si $|B| = 0$ (en part. $\mathbb{P}(\{X = x\}) = 0 \forall x$). Sous cette condition, le théorème de Radon-Nikodym assure l'existence d'une fonction borélienne $f_X : \mathbb{R} \rightarrow \mathbb{R}$ (appelée densité) telle que

$$f_X(x) \geq 0, \forall x \in \mathbb{R}, \quad \int_{\mathbb{R}} f_X(x) dx = 1 \quad \text{et} \quad \mathbb{P}(\{X \in B\}) = \int_B f_X(x) dx.$$

De plus, $F_X(t) = \int_{-\infty}^t f_X(x) dx$ et $F'_X(t) = f_X(t)$.

Exemple 1.3.5.

A) Loi binomiale $B(n, p)$, $n \geq 1, p \in [0, 1]$:

$$p(k) = \mathbb{P}(\{X = k\}) = \binom{n}{p} p^k (1-p)^{n-k}, \quad \text{pour } 0 \leq k \leq n,$$

$$\text{où } \binom{n}{p} = \frac{n!}{k!(n-k)!}.$$

B) Loi gaussienne $\mathcal{N}(\mu, \sigma^2)$, $\mu \in \mathbb{R}, \sigma > 0$:

$$\text{densité : } f_X(x) = \frac{1}{\sqrt{2\pi\sigma^2}} \exp\left(-\frac{(x-\mu)^2}{2\sigma^2}\right), \quad x \in \mathbb{R}.$$

Terminologie : - Si X suit une loi gaussienne (p.ex.), on écrit $X \sim \mathcal{N}(\mu, \sigma^2)$.

- Si X et Y suivent une même loi, on dit que X et Y sont identiquement distribuées (i.d.) et on note $X \sim Y$ (ne pas confondre avec $X \simeq Y$, qui veut dire que X est “à peu près égale” à Y).

1.4 Espérance d'une variable aléatoire

Cours 2

Construction de l'espérance (= intégrale de Lebesgue !)

On procède en trois étapes :

1. Soit $X(\omega) = \sum_{i=0}^{\infty} x_i 1_{A_i}(\omega)$, $x_i \geq 0$, $A_i \in \mathcal{F}$. On définit l'espérance de telles v.a. (dites simples) comme suit :

$$\mathbb{E}(X) = \sum_{i=0}^{\infty} x_i \mathbb{P}(A_i) \in [0, +\infty].$$

Attention ! $\mathbb{E}(X)$ peut prendre la "valeur" $+\infty$.

Exemples : - Si $X = 1_A$, $\mathbb{P}(A) = p$, alors $\mathbb{E}(X) = \mathbb{P}(A) = p$.

- Si $X = c 1_{\Omega} = \text{cte sur } \Omega$, alors $\mathbb{E}(X) = c$.

2. Soit X une v.a. \mathcal{F} -mesurable telle que $X(\omega) \geq 0$, $\forall \omega \in \Omega$. On pose

$$X_n(\omega) = \sum_{i=0}^{\infty} \frac{i}{2^n} 1_{\{\frac{i}{2^n} \leq X < \frac{i+1}{2^n}\}}(\omega).$$

Alors (X_n) est une suite croissante de v.a. qui tend vers X . On définit

$$\mathbb{E}(X) = \lim_{n \rightarrow \infty} \mathbb{E}(X_n) = \lim_{n \rightarrow \infty} \sum_{i=0}^{\infty} \frac{i}{2^n} \mathbb{P}\left(\left\{\frac{i}{2^n} \leq X < \frac{i+1}{2^n}\right\}\right) \in [0, +\infty].$$

3. Soit X une v.a. \mathcal{F} -mesurable quelconque. On pose

$$X(\omega) = X^+(\omega) - X^-(\omega) \quad \text{avec} \quad \begin{cases} X^+(\omega) = \max(0, X(\omega)) \geq 0, \\ X^-(\omega) = \max(0, -X(\omega)) \geq 0. \end{cases}$$

On a alors $|X(\omega)| = X^+(\omega) + X^-(\omega) \geq 0$.

- Si $\mathbb{E}(|X|) < \infty$, alors on définit $\mathbb{E}(X) = \mathbb{E}(X^+) - \mathbb{E}(X^-)$.

- Si $\mathbb{E}(|X|) = \infty$, alors on dit que $\mathbb{E}(X)$ n'est pas définie.

Terminologie : - Si $\mathbb{E}(X) = 0$, alors on dit que X est une v.a. centrée.

- Si $\mathbb{E}(|X|) < \infty$, alors on dit que X est une v.a. intégrable.

- Si $\mathbb{E}(X^2) < \infty$ alors on dit que X est une v.a. de carré intégrable.

- On dit que X est une v.a. bornée s'il existe une cte $K > 0$ telle que $|X(\omega)| \leq K$, $\forall \omega \in \Omega$.

Remarque 1.4.1. X bornée $\Rightarrow \mathbb{E}(X^2) < \infty \Rightarrow \mathbb{E}(|X|) < \infty$.

Proposition 1.4.2. Soient X une v.a. et $g : \mathbb{R} \rightarrow \mathbb{R}$ une fonction borélienne telle que $\mathbb{E}(|g(X)|) < \infty$. Alors

A) Si X est une v.a. discrète (à valeurs dans D dénombrable), alors

$$\mathbb{E}(g(X)) = \sum_{x \in D} g(x) \mathbb{P}(\{X = x\}).$$

B) Si X est une v.a. continue (avec densité f_X), alors

$$\mathbb{E}(g(X)) = \int_{\mathbb{R}} g(x) f_X(x) dx.$$

Ceci s'applique en particulier si $g(x) = x$.

Variance et covariance de variables aléatoires

Définition 1.4.3. Soient X, Y deux v.a. de carré intégrable. On pose

$$\begin{aligned} \text{Var}(X) &= \mathbb{E}((X - \mathbb{E}(X))^2) = \mathbb{E}(X^2) - (\mathbb{E}(X))^2 \geq 0 \\ \text{Cov}(X, Y) &= \mathbb{E}((X - \mathbb{E}(X))(Y - \mathbb{E}(Y))) = \mathbb{E}(XY) - \mathbb{E}(X)\mathbb{E}(Y) \end{aligned}$$

Terminologie : - Un événement $A \in \mathcal{F}$ est dit négligeable si $\mathbb{P}(A) = 0$.

- Un événement $A \in \mathcal{F}$ est dit presque sûr (souvent abrégé p.s.) si $\mathbb{P}(A) = 1$, i.e. si A^c est négligeable.

Exemple 1.4.4. Soit X une v.a. telle que $\mathbb{P}(\{X = c\}) = 1$. Alors on dit que $X = c$ presque sûrement (“ $X = c$ p.s.”)

Proposition 1.4.5. Si $(A_n)_{n=1}^\infty \subset \mathcal{F}$ est une famille d'événements négligeables (i.e. $\mathbb{P}(A_n) = 0 \forall n$), alors $\cup_{n=1}^\infty A_n$ est négligeable.

Démonstration. $\mathbb{P}(\cup_{n=1}^\infty A_n) \leq \sum_{n=1}^\infty \mathbb{P}(A_n) = 0$. □

Exemple 1.4.6. L'ensemble $A = [0, 1] \cap \mathbb{Q}$ est négligeable pour la mesure de Lebesgue, car \mathbb{Q} est dénombrable et $|\{x\}| = 0$ pour tout $x \in [0, 1]$.

Propriétés de l'espérance

Soient X, Y deux v.a. intégrables.

- Linéarité : $\mathbb{E}(cX + Y) = c\mathbb{E}(X) + \mathbb{E}(Y)$, $c \in \mathbb{R}$ et X, Y v.a. intégrables.
- Positivité : si $X \geq 0$ p.s., alors $\mathbb{E}(X) \geq 0$.
- Positivité stricte : si $X \geq 0$ p.s. et $\mathbb{E}(X) = 0$, alors $X = 0$ p.s.

- Monotonie : si $X \geq Y$ p.s., alors $\mathbb{E}(X) \geq \mathbb{E}(Y)$.

Inégalité de Cauchy-Schwarz

Soient X, Y deux v.a. de carré intégrable. Alors

$$\begin{cases} i) & XY \text{ est intégrable,} \\ ii) & (\mathbb{E}(|XY|))^2 \leq \mathbb{E}(X^2) \mathbb{E}(Y^2). \end{cases}$$

En posant $Y \equiv 1$, on trouve que $(\mathbb{E}(|X|))^2 \leq \mathbb{E}(X^2)$ (donc $\mathbb{E}(|X|) < \infty$ si $\mathbb{E}(X^2) < \infty$; cf. remarque 1.4.1).

Inégalité triangulaire

Soient X, Y deux v.a. intégrables. Alors

$$\mathbb{E}(|X + Y|) \leq \mathbb{E}(|X|) + \mathbb{E}(|Y|)$$

Inégalité de Jensen

Soient X une v.a. et $\varphi : \mathbb{R} \rightarrow \mathbb{R}$ une fonction borélienne et convexe telle que $\mathbb{E}(|\varphi(X)|) < \infty$. Alors

$$\varphi(\mathbb{E}(X)) \leq \mathbb{E}(\varphi(X)).$$

En particulier, $|\mathbb{E}(X)| \leq \mathbb{E}(|X|)$.

Démonstration. Vu que φ est convexe, on a $\varphi(x) = \sup_{a,b: ay+b \leq \varphi(y), \forall y \in \mathbb{R}} (ax + b)$ et donc :

$$\varphi(\mathbb{E}(X)) = \sup_{a,b:\dots} (a \mathbb{E}(X) + b) = \sup_{a,b:\dots} \mathbb{E}(aX + b) \leq \sup_{a,b:\dots} \mathbb{E}(\varphi(X)) = \mathbb{E}(\varphi(X)).$$

□

Exemple 1.4.7. Si $X = a$ ou b avec prob. $\frac{1}{2} - \frac{1}{2}$ et φ est convexe, alors $\varphi(\mathbb{E}(X)) = \varphi(\frac{a+b}{2}) \leq \frac{\varphi(a)+\varphi(b)}{2} = \mathbb{E}(\varphi(X))$.

Inégalité de Chebychev (ou Markov)

Soient X une v.a. et $\psi : \mathbb{R} \rightarrow \mathbb{R}_+$ telle que ψ est borélienne et croissante sur \mathbb{R}_+ , $\psi(a) > 0$ pour tout $a > 0$ et $\mathbb{E}(\psi(X)) < \infty$. Alors

$$\mathbb{P}(\{X \geq a\}) \leq \frac{\mathbb{E}(\psi(X))}{\psi(a)}, \quad \forall a > 0.$$

Démonstration. Du fait que ψ est croissante sur \mathbb{R}_+ , on a

$$\mathbb{E}(\psi(X)) \geq \mathbb{E}(\psi(X) 1_{\{X \geq a\}}) \geq \mathbb{E}(\psi(a) 1_{\{X \geq a\}}) = \psi(a) \mathbb{E}(1_{\{X \geq a\}}) = \psi(a) \mathbb{P}(\{X \geq a\}).$$

Comme $\psi(a) > 0$, ceci permet de conclure.

□

1.5 Indépendance

1.5.1 Indépendance d'événements

Définition 1.5.1. Deux événements A et $B \in \mathcal{F}$ sont indépendants si $\mathbb{P}(A \cap B) = \mathbb{P}(A) \mathbb{P}(B)$.
Attention ! Ne pas confondre : A et B sont disjoints si $A \cap B = \emptyset$ ($\Rightarrow \mathbb{P}(A \cup B) = \mathbb{P}(A) + \mathbb{P}(B)$).

Notation : Si A est indépendant de B , on note $A \perp B$ (de même pour les tribus et les v.a. ; voir plus bas).

Conséquence :

$$\begin{aligned}\mathbb{P}(A^c \cap B) &= \mathbb{P}(B \setminus (A \cap B)) = \mathbb{P}(B) - \mathbb{P}(A \cap B) \\ &= \mathbb{P}(B) - \mathbb{P}(A) \mathbb{P}(B) = (1 - \mathbb{P}(A)) \mathbb{P}(B) = \mathbb{P}(A^c) \mathbb{P}(B).\end{aligned}$$

De même, on a $\mathbb{P}(A \cap B^c) = \mathbb{P}(A) \mathbb{P}(B^c)$ et $\mathbb{P}(A^c \cap B^c) = \mathbb{P}(A^c) \mathbb{P}(B^c)$.

Définition 1.5.2. n événements $A_1, \dots, A_n \in \mathcal{F}$ sont indépendants si

$$\mathbb{P}(A_1^* \cap \dots \cap A_n^*) = \prod_{i=1}^n \mathbb{P}(A_i^*), \quad \text{où } A_i^* = \text{soit } A_i, \text{ soit } A_i^c.$$

Proposition 1.5.3. n événements $A_1, \dots, A_n \in \mathcal{F}$ sont indépendants ssi

$$\mathbb{P}(\cap_{i \in I} A_i) = \prod_{i \in I} \mathbb{P}(A_i), \quad \forall I \subset \{1, \dots, n\}.$$

Remarque 1.5.4. - Pour $n > 2$, la condition $\mathbb{P}(A_1 \cap \dots \cap A_n) = \mathbb{P}(A_1) \dots \mathbb{P}(A_n)$ ne suffit pas !
- L'indépendance de n événements telle que définie ci-dessus est plus forte que l'indépendance deux à deux ($A_i \perp A_j, \forall i \neq j$).

1.5.2 Indépendance de tribus

Définition 1.5.5. Une famille $(\mathcal{F}_1 \dots \mathcal{F}_n)$ de sous-tribus de \mathcal{F} est indépendante si

$$\mathbb{P}(A_1 \cap \dots \cap A_n) = \mathbb{P}(A_1) \dots \mathbb{P}(A_n), \quad \forall A_1 \in \mathcal{F}_1, \dots, A_n \in \mathcal{F}_n.$$

Proposition 1.5.6. $(\sigma(A_1), \dots, \sigma(A_n))$ est une famille de sous-tribus indépendantes ssi les événements (A_1, \dots, A_n) sont indépendants.

1.5.3 Indépendance de variables aléatoires

Définition 1.5.7. Une famille (X_1, \dots, X_n) de v.a. (\mathcal{F} -mesurables) est indépendante si $(\sigma(X_1), \dots, \sigma(X_n))$ est indépendante.

Proposition 1.5.8. (X_1, \dots, X_n) est une famille de v.a. indépendantes ssi les événements $\{X_1 \leq t_1\}, \dots, \{X_n \leq t_n\}$ sont indépendants $\forall t_1, \dots, t_n \in \mathbb{R}$.

En particulier : $X \perp Y$ ssi $\mathbb{P}(\{X \leq t, Y \leq s\}) = \mathbb{P}(\{X \leq t\}) \mathbb{P}(\{Y \leq s\})$, $\forall t, s \in \mathbb{R}$.

Exemple 1.5.9. Soit $\Omega = \{1 \dots 6\} \times \{1 \dots 6\}$, $\mathcal{F} = \mathcal{P}(\Omega)$, $\mathbb{P}(\{(i, j)\}) = \frac{1}{36}$. On pose

$$X(\omega) = X(\omega_1, \omega_2) = \omega_1, \quad Y(\omega) = Y(\omega_1, \omega_2) = \omega_2.$$

Calculons $\mathbb{P}(\{X = i\}) = \mathbb{P}(\{\omega \in \Omega : \omega_1 = i\}) = \mathbb{P}(\{(i, 1), \dots, (i, 6)\}) = \frac{1}{6} = \mathbb{P}(\{Y = j\})$. D'autre part, $\mathbb{P}(\{X = i, Y = j\}) = \mathbb{P}(\{(i, j)\}) = \frac{1}{36} = \frac{1}{6} \times \frac{1}{6} = \mathbb{P}(\{X = i\}) \mathbb{P}(\{Y = j\})$, donc X et Y sont indépendantes.

Exemple 1.5.10. Si $X(\omega) = c$, $\forall \omega \in \Omega$, alors $X \perp Y$, $\forall Y$ (une v.a. constante est indépendante de toute autre v.a.).

Exemple 1.5.11. Si $X \perp Y$ et g, h sont des fonctions boréliennes, alors $g(X) \perp h(Y)$; c'est vrai en particulier si $g = h$ (mais pas $X = Y$, bien sûr!). Cette propriété découle du fait que $g(X)$ est $\sigma(X)$ -mesurable (resp. que $h(Y)$ est $\sigma(Y)$ -mesurable) et de la définition d'indépendance pour les tribus.

Proposition 1.5.12. Soient $c \in \mathbb{R}$ et X, Y deux v.a. de carré intégrable. Si $X \perp Y$, alors

$$\mathbb{E}(XY) = \mathbb{E}(X) \mathbb{E}(Y) \quad \text{et} \quad \text{Var}(cX + Y) = c^2 \text{Var}(X) + \text{Var}(Y),$$

La première égalité dit que si deux v.a. sont indépendantes, alors elles sont décorrélées; la réciproque n'est pas vraie.

Proposition 1.5.13. Si $X \perp Y$ et X, Y sont deux v.a. continues possédant une densité conjointe $f_{X,Y}$ (i.e. $\mathbb{P}((X, Y) \in B) = \iint_B f_{X,Y}(x, y) dx dy$, $\forall B \in \mathcal{B}(\mathbb{R}^2)$), alors $f_{X,Y}(x, y) = f_X(x) f_Y(y)$, $\forall x, y \in \mathbb{R}$.

1.6 Espérance conditionnelle

Cours 3

1.6.1 Conditionnement par rapport à un événement $B \in \mathcal{F}$

Soit $A \in \mathcal{F}$: $\mathbb{P}(A|B) = \frac{\mathbb{P}(A \cap B)}{\mathbb{P}(B)}$, si $\mathbb{P}(B) \neq 0$.

Soit X une v.a. intégrable (i.e. $\mathbb{E}(|X|) < \infty$) : $\mathbb{E}(X|B) = \frac{\mathbb{E}(X 1_B)}{\mathbb{P}(B)}$, si $\mathbb{P}(B) \neq 0$.

1.6.2 Conditionnement par rapport à une v.a. discrète Y (à valeurs dans D dénombrable)

Soit $A \in \mathcal{F} : \mathbb{P}(A|Y) = \varphi(Y)$, où $\varphi(y) = \mathbb{P}(A|Y = y)$, $y \in D$.

Soit X une v.a. intégrable : $\mathbb{E}(X|Y) = \psi(Y)$, où $\psi(y) = \mathbb{E}(X|Y = y)$, $y \in D$.

Remarque 1.6.1. Il est important de voir que soit $\mathbb{P}(A|Y)$, soit $\mathbb{E}(X|Y)$ sont des v.a. !

Exemple 1.6.2. Soient (X_1, X_2) deux jets de dés indépendants : $\mathbb{E}(X_1 + X_2|X_2) = \psi(X_2)$, où

$$\begin{aligned}\psi(y) &= \mathbb{E}(X_1 + X_2|X_2 = y) = \frac{\mathbb{E}((X_1 + X_2) 1_{\{X_2=y\}})}{\mathbb{P}(\{X_2 = y\})} \\ &= \frac{\mathbb{E}((X_1 + y) 1_{\{X_2=y\}})}{\mathbb{P}(\{X_2 = y\})} = \frac{\mathbb{E}(X_1 + y) \mathbb{P}(\{X_2 = y\})}{\mathbb{P}(\{X_2 = y\})} = \mathbb{E}(X_1) + y,\end{aligned}$$

donc $\mathbb{E}(X_1 + X_2|X_2) = \mathbb{E}(X_1) + X_2$.

1.6.3 Conditionnement par rapport à une v.a. Y continue ? Cas général ?

$\mathbb{P}(A|Y) = \psi(Y)$ où $\psi(y) = \mathbb{P}(A|Y = y) = ???$ Problème : $\mathbb{P}(\{Y = y\}) = 0$.

Il vaut mieux généraliser la définition pour une tribu.

1.6.4 Conditionnement par rapport à une tribu \mathcal{G}

Soient $(\Omega, \mathcal{F}, \mathbb{P})$ un espace de probabilité ; X une v.a. \mathcal{F} -mesurable telle que $\mathbb{E}(|X|) < \infty$ et \mathcal{G} une sous-tribu de \mathcal{F} .

Théorème - Définition. Il existe une v.a. Z telle que $\mathbb{E}(|Z|) < \infty$ et

$$\begin{cases} i) & Z \text{ est une v.a. } \mathcal{G}\text{-mesurable,} \\ ii) & \mathbb{E}(XU) = \mathbb{E}(ZU), \forall \text{ v.a. } U \mathcal{G}\text{-mesurable et bornée.} \end{cases}$$

Z est notée $\mathbb{E}(X|\mathcal{G})$ et est appelée l'espérance conditionnelle de X par rapport à \mathcal{G} .

De plus, si Z_1 et Z_2 vérifient (i) et (ii), alors $Z_1 = Z_2$ p.s.

Remarque 1.6.3. A cause de cette dernière phrase, l'espérance conditionnelle n'est définie qu'à un "p.s." près, c'est-à-dire à un ensemble négligeable près (qu'on va peu à peu oublier...).

Définition 1.6.4. On pose $\mathbb{P}(A|\mathcal{G}) = \mathbb{E}(1_A|\mathcal{G})$ (probabilité conditionnelle par rapport à une tribu), ainsi que $\mathbb{E}(X|Y) = \mathbb{E}(X|\sigma(Y))$ (espérance conditionnelle par rapport à une variable aléatoire).

Remarque 1.6.5. Du fait que toute v.a. U $\sigma(Y)$ -mesurable et bornée s'écrit $U = g(Y)$ avec g borélienne et bornée, on peut définir de manière équivalente $\mathbb{E}(X|Y)$ comme la v.a. Z qui vérifie $\mathbb{E}(|Z|) < \infty$ et

$$\begin{cases} i) & Z \text{ est une v.a. } \sigma(Y)\text{-mesurable,} \\ ii) & \mathbb{E}(Xg(Y)) = \mathbb{E}(Zg(Y)), \forall \text{ fonction } g \text{ borélienne et bornée.} \end{cases}$$

Exemple 1.6.6. Soient X, Y deux v.a. aléatoires continues possédant une densité conjointe $f_{X,Y}$. Alors

$$\mathbb{E}(X|Y) = \psi(Y) \text{ p.s., } \quad \text{où } \psi(y) = \mathbb{E}(X|Y = y) = \frac{1}{f_Y(y)} \int_{\mathbb{R}} x f_{X,Y}(x, y) dx$$

$$\text{et } f_Y(y) = \int_{\mathbb{R}} f_{X,Y}(x, y) dx.$$

Démonstration. 1) Simplification : on admet que la fonction ψ définie ci-dessus est borélienne. Il en découle donc directement que $\psi(Y)$ est une v.a. $\sigma(Y)$ -mesurable.

2) Montrons que $\mathbb{E}(Xg(Y)) = \mathbb{E}(\psi(Y)g(Y))$, $\forall g$ borélienne bornée :

$$\begin{aligned} \mathbb{E}(\psi(Y)g(Y)) &= \int_{\mathbb{R}} \psi(y) g(y) f_Y(y) dy \\ &= \int_{\mathbb{R}} \left(\frac{1}{f_Y(y)} \int_{\mathbb{R}} x f_{X,Y}(x, y) dx \right) g(y) f_Y(y) dy \\ &= \int \int_{\mathbb{R}^2} x g(y) f_{X,Y}(x, y) dx dy = \mathbb{E}(Xg(Y)). \end{aligned}$$

□

Propriétés de l'espérance conditionnelle (démonstration : cf. exercices)

Soient X, Y deux v.a. intégrables.

- Linéarité : $\mathbb{E}(cX + Y|\mathcal{G}) = c\mathbb{E}(X|\mathcal{G}) + \mathbb{E}(Y|\mathcal{G})$ p.s.
- Positivité-monotonie : $X \geq Y$ p.s. $\Rightarrow \mathbb{E}(X|\mathcal{G}) \geq \mathbb{E}(Y|\mathcal{G})$ p.s.
- $\mathbb{E}(\mathbb{E}(X|\mathcal{G})) = \mathbb{E}(X)$.
- Si $X \perp \mathcal{G}$ alors $\mathbb{E}(X|\mathcal{G}) = \mathbb{E}(X)$ p.s.
- Si X est \mathcal{G} -mesurable, alors $\mathbb{E}(X|\mathcal{G}) = X$ p.s.
- Si Y est \mathcal{G} -mesurable et bornée, alors $\mathbb{E}(XY|\mathcal{G}) = \mathbb{E}(X|\mathcal{G})Y$ p.s.
- Si $\mathcal{H} \subset \mathcal{G} \subset \mathcal{F}$ alors $\mathbb{E}(\mathbb{E}(X|\mathcal{H})|\mathcal{G}) = \mathbb{E}(X|\mathcal{H}) = \mathbb{E}(\mathbb{E}(X|\mathcal{G})|\mathcal{H})$ p.s.

Proposition 1.6.7. (cf. exercices pour une démonstration simplifiée)

Si $X \perp \mathcal{G}$, Y est \mathcal{G} -mesurable et $\varphi : \mathbb{R} \times \mathbb{R} \rightarrow \mathbb{R}$ est borélienne et telle que $\mathbb{E}(|\varphi(X, Y)|) < \infty$, alors

$$\mathbb{E}(\varphi(X, Y)|\mathcal{G}) = \psi(Y), \quad \text{où } \psi(y) = \mathbb{E}(\varphi(X, y)).$$

Inégalité de Jensen (démonstration : cf. même propriété pour l'espérance)

Soient X une v.a., \mathcal{G} une sous-tribu de \mathcal{F} et $\varphi : \mathbb{R} \rightarrow \mathbb{R}$ convexe telle que $\mathbb{E}(|\varphi(X)|) < \infty$.

Alors

$$\varphi(\mathbb{E}(X|\mathcal{G})) \leq \mathbb{E}(\varphi(X)|\mathcal{G}) \text{ p.s.}$$

En particulier : $|\mathbb{E}(X|\mathcal{G})| \leq \mathbb{E}(|X||\mathcal{G})$ p.s.

Remarque 1.6.8. C'est très souvent en utilisant les propriétés mentionnées ci-dessus qu'on peut véritablement calculer une espérance conditionnelle. Le recours à la définition théorique n'est utile que dans certains cas délicats.

1.7 Convergences de suites de variables aléatoires

Soient $(X_n)_{n=1}^\infty$ une suite de v.a. et X une autre v.a., toutes définies sur $(\Omega, \mathcal{F}, \mathbb{P})$. Il y a plusieurs façons de définir la convergence de la suite (X_n) vers X , car $X : \Omega \rightarrow \mathbb{R}$ est une fonction.

Convergence en probabilité

$$X_n \xrightarrow[n \rightarrow \infty]{\mathbb{P}} X \quad \text{si} \quad \forall \varepsilon > 0, \lim_{n \rightarrow \infty} \mathbb{P}(\{\omega \in \Omega : |X_n(\omega) - X(\omega)| > \varepsilon\}) = 0.$$

Convergence presque sûre

$$X_n \xrightarrow[n \rightarrow \infty]{} X \text{ p.s.} \quad \text{si} \quad \mathbb{P}(\{\omega \in \Omega : \lim_{n \rightarrow \infty} X_n(\omega) = X(\omega)\}) = 1.$$

Remarque 1.7.1. Si $X_n \xrightarrow[n \rightarrow \infty]{} X$ p.s., alors $X_n \xrightarrow[n \rightarrow \infty]{\mathbb{P}} X$.

Le contraire est faux. Donnons un contre-exemple : soit une suite de “pile ou face” indépendants et équilibrés (p. ex. “PFPPPPFP...”). On définit

$$\begin{aligned} X_1 &= 1_P = \begin{cases} 1 & \text{si 1er résultat=pile,} \\ 0 & \text{sinon,} \end{cases} & X_2 = 1_F, & X_3 = 1_{PP}, & X_4 = 1_{PF}, \\ X_5 &= 1_{FP}, & X_6 = 1_{FF}, & X_7 = 1_{PPP}, & X_8 = 1_{PPF} \dots \end{aligned}$$

La suite (X_n) converge en probabilité vers 0 car $\mathbb{P}(|X_n| > \varepsilon) = \mathbb{P}(X_n = 1) \xrightarrow[n \rightarrow \infty]{} 0$.

Mais (X_n) ne converge pas p.s. vers 0, car pour un ω donné, on a p. ex.

$$(X_n(\omega)) = (1, 0, 0, 1, 0, 0, 0, 0, 0, 1, 0, 0 \dots 0 \dots 0, 1, 0, \dots),$$

Autrement dit, il y a toujours un “1” qui apparaît et la suite ne converge pas vers 0.

Critère pour la convergence presque sûre : $X_n \xrightarrow[n \rightarrow \infty]{} X$ p.s. ssi

$$\forall \varepsilon > 0, \mathbb{P}(\{\omega \in \Omega : |X_n(\omega) - X(\omega)| > \varepsilon \text{ pour une infinité de } n\}) = 0.$$

Revenons à l'exemple ci-dessus : $\mathbb{P}(\{X_n = 1 \text{ pour une infinité de } n\}) = 1 \neq 0$.

Convergence en moyenne (ou convergence L_1)

$$\lim_{n \rightarrow \infty} \mathbb{E}(|X_n - X|) = 0.$$

Convergence quadratique (ou convergence L_2)

$$\lim_{n \rightarrow \infty} \mathbb{E}(|X_n - X|^2) = 0$$

L'une ou l'autre de ces convergences implique la convergence en probabilité. En effet, on a p.ex. pour tout $\varepsilon > 0$ fixé :

$$\mathbb{P}(\{|X_n - X| > \varepsilon\}) \leq \frac{\mathbb{E}(|X_n - X|^2)}{\varepsilon^2} \xrightarrow[n \rightarrow \infty]{} 0.$$

(On a appliqué ici l'inégalité de Chebychev avec $\varphi(x) = x^2$.)

Lemme de Borel-Cantelli (B-C)

Soit (A_n) une suite d'événements de \mathcal{F} .

(a) Si $\sum_{n=1}^{\infty} \mathbb{P}(A_n) < \infty$, alors

$$\mathbb{P}(\{\omega \in \Omega : \omega \in A_n \text{ pour une infinité de } n\}) = 0.$$

(b) Si $(A_n)_{n=1}^{\infty}$ sont indépendants et $\sum_{n=1}^{\infty} \mathbb{P}(A_n) = \infty$, alors

$$\mathbb{P}(\{\omega \in \Omega : \omega \in A_n \text{ pour une infinité de } n\}) = 1.$$

Application : Si $\sum_{n=1}^{\infty} \mathbb{E}(|X_n - X|^2) < \infty$, alors $X_n \xrightarrow[n \rightarrow \infty]{} X$ p.s.

Démonstration.

$$\sum_{n=1}^{\infty} \mathbb{P}(\{|X_n - X| > \varepsilon\}) \leq \frac{1}{\varepsilon^2} \sum_{n=1}^{\infty} \mathbb{E}(|X_n - X|^2) < \infty, \quad \forall \varepsilon > 0 \text{ fixé.}$$

D'après B-C (a), on a alors $\forall \varepsilon > 0$ fixé,

$$\mathbb{P}(\{\omega \in \Omega : |X_n(\omega) - X(\omega)| > \varepsilon \text{ pour une infinité de } n\}) = 0.$$

Par le critère énoncé ci-dessus, $X_n \xrightarrow[n \rightarrow \infty]{} X$ p.s. □

Loi des grands nombres (LGN)

Soient $(X_n)_{n=1}^\infty$ une suite de v.a. indépendantes et identiquement distribuées (i.i.d.) et $S_n = X_1 + \dots + X_n$. Si $\mathbb{E}(|X_1|) < \infty$, alors

- (a) Loi faible : $\frac{S_n}{n} \xrightarrow[n \rightarrow \infty]{\mathbb{P}} \mathbb{E}(X_1)$.
- (b) Loi forte : $\frac{S_n}{n} \xrightarrow[n \rightarrow \infty]{} \mathbb{E}(X_1)$ p.s.

Remarque 1.7.2. Si on supprime l'hypothèse que $\mathbb{E}(|X_1|) < \infty$, alors les choses changent :

- (a) Si $\lim_{a \rightarrow \infty} a \mathbb{P}(|X_1| > a) = 0$, alors $\frac{S_n}{n} - \mathbb{E}(X_1 \cdot 1_{|X_1| \leq n}) \xrightarrow[n \rightarrow \infty]{\mathbb{P}} 0$.
- (b) Si $\mathbb{E}(|X_1|) = \infty$, alors $\frac{S_n}{n}$ diverge p.s. lorsque $n \rightarrow \infty$.

Exemple 1.7.3. Soit (X_n) une suite de pile (1) ou face (0) indépendants et équilibrés. Du fait que $\mathbb{E}(X_1) = \frac{1}{2} < \infty$, on a par la loi forte des grands nombres :

$$\frac{S_n}{n} \rightarrow \frac{1}{2} \text{ p.s., } \quad \text{i.e. } S_n \simeq \frac{n}{2}.$$

Question : quelle est la déviation moyenne de $S_n - \frac{n}{2}$?

Convergence en loi

Soient (X_n) une suite de v.a. (définies sur $(\Omega, \mathcal{F}, \mathbb{P})$) et (F_{X_n}) la suite des fonctions de répartition correspondantes ($F_{X_n}(t) = \mathbb{P}(\{X_n \leq t\})$, $t \in \mathbb{R}$). X_n converge en loi vers X (" $X_n \xRightarrow[n \rightarrow \infty]{} X$ ") si $\lim_{n \rightarrow \infty} F_{X_n}(t) = F_X(t)$, $\forall t \in \mathbb{R}$ point de continuité de F_X .

Théorème central limite (TCL)

Soit $(X_n)_{n=1}^\infty$ une suite de v.a. i.i.d. telle que $\mathbb{E}(X_1) = \mu \in \mathbb{R}$ et $\text{Var}(X_1) = \sigma^2 > 0$. Soit $S_n = X_1 + \dots + X_n$. Alors

$$\frac{S_n - n\mu}{\sqrt{n}\sigma} \xRightarrow[n \rightarrow \infty]{} Z \sim \mathcal{N}(0, 1), \quad \text{i.e. } \mathbb{P}\left(\frac{S_n - n\mu}{\sqrt{n}\sigma} \leq t\right) \xrightarrow[n \rightarrow \infty]{} \int_{-\infty}^t \frac{1}{\sqrt{2\pi}} e^{-\frac{x^2}{2}} dx = \Phi(t), \quad \forall t \in \mathbb{R}.$$

(NB : la fonction de répartition Φ de la loi $\mathcal{N}(0, 1)$ est continue sur tout \mathbb{R} .)

Exemple 1.7.4. Soit $(X_n)_{n=1}^\infty$ la suite de pile ou face indépendants et équilibrés considérée ci-dessus : $\mathbb{E}(X_1) = \frac{1}{2}$ et $\text{Var}(X_1) = \frac{1}{4}$, donc

$$\frac{S_n - \frac{n}{2}}{\sqrt{n}\frac{1}{2}} \Rightarrow Z \sim \mathcal{N}(0, 1), \quad \text{i.e. } S_n \simeq \frac{n}{2} + \frac{\sqrt{n}}{2} Z$$

La réponse à la question posée ci-dessus est donc que la déviation est d'ordre \sqrt{n} .

1.8 Processus aléatoire à temps discret

Cours 4

Marche aléatoire symétrique sur \mathbb{Z}

Soit $(X_n)_{n=1}^\infty$ une suite de v.a. i.i.d. telle que $\mathbb{P}(\{X_1 = +1\}) = \mathbb{P}(\{X_1 = -1\}) = \frac{1}{2}$.

Soient $S_0 = 0$, $S_n = X_1 + \dots + X_n$. Le processus $(S_n, n \in \mathbb{N})$ est appelé la marche aléatoire symétrique sur \mathbb{Z} .

Deux points de vue possibles :
$$\left\{ \begin{array}{l} (S_n, n \in \mathbb{N}) \text{ est une famille de v.a. à valeurs dans } \mathbb{Z}, \\ S : \left\{ \begin{array}{l} \Omega \rightarrow \{\text{suites } (z_n, n \in \mathbb{N}) \text{ à valeurs dans } \mathbb{Z}\}, \\ \omega \mapsto S(\omega) = (S_n(\omega), n \in \mathbb{N}) = \text{suite aléatoire.} \end{array} \right. \end{array} \right.$$

Remarque 1.8.1. $\mathbb{E}(X_1) = 0 \Rightarrow \mathbb{E}(S_n) = 0$.

$\text{Var}(X_1) = \mathbb{E}(X_1^2) = 1 \Rightarrow \text{Var}(S_n) = \sum_{i=1}^n \text{Var}(X_i) = n$.

Loi des grands nombres : $\frac{S_n}{n} \rightarrow 0$ p.s.

Théorème central limite : $\frac{S_n}{\sqrt{n}} \Rightarrow Z \sim \mathcal{N}(0, 1)$ (i.e. $S_n \simeq \sqrt{n} Z$).

Généralisations :

- Marche aléatoire asymétrique sur \mathbb{Z} : $S_n = X_1 + \dots + X_n$, X_i i.i.d., $\mathbb{P}(\{X_1 = +1\}) = 1 - \mathbb{P}(\{X_1 = -1\}) = p \neq 1/2$.
- Marche aléatoire à valeurs dans \mathbb{R} : $S_n = X_1 + \dots + X_n$, X_i i.i.d., $X_1 \sim N(0, 1)$ (p.ex.).
- Temps continu : mouvement brownien (voir section 2).

Chaîne de Markov

Processus aléatoire à temps discret $(M_n, n \in \mathbb{N})$ à valeurs dans un ensemble D dénombrable (appelé espace des états) tel que

$$\mathbb{P}(M_{n+1} = x_{n+1} | M_n = x_n, M_{n-1} = x_{n-1} \dots M_0 = x_0) = \mathbb{P}(M_{n+1} = x_{n+1} | M_n = x_n),$$

$\forall n \geq 1, x_0 \dots x_{n+1} \in D$ (c'est un processus qui "oublie" son histoire au fur et à mesure).

Chaîne de Markov stationnaire

Chaîne de Markov telle que $\mathbb{P}(M_{n+1} = y | M_n = x) = Q(x, y)$ pour tout $n \in \mathbb{N}$, $x, y \in D$.

$Q(x, y)$ est appelée la probabilité de transition de x à y .

La probabilité définie par $\mu(\{x\}) = \mathbb{P}(M_0 = x)$, $x \in D$, est appelée la loi initiale.

μ et Q caractérisent entièrement la chaîne de Markov (dans le cas stationnaire).

En effet, on vérifie que :

$$\mathbb{P}(M_n = x_n, \dots, M_0 = x_0) = \mu(x_0) Q(x_0, x_1) \cdots Q(x_{n-1}, x_n).$$

Exemple 1.8.2. La marche aléatoire symétrique ou asymétrique sur \mathbb{Z} est une chaîne de Markov stationnaire. Soit effectivement $S_0 = 0$, $S_n = X_1 + \dots + X_n$, avec X_i i.i.d. et $\mathbb{P}(\{X_1 = 1\}) = p = 1 - \mathbb{P}(\{X_1 = -1\})$:

$$\mathbb{P}(S_{n+1} = y | S_n = x, S_{n-1} = x_{n-1} \dots S_0 = x_0) = \frac{\mathbb{P}(S_{n+1} = y, S_n = x, \dots S_0 = x_0)}{\mathbb{P}(S_n = x, S_{n-1} = x_{n-1} \dots S_0 = x_0)} = (*)$$

Par définition, $S_{n+1} = X_1 + \dots + X_n + X_{n+1} = S_n + X_{n+1}$, i.e. $X_{n+1} = S_{n+1} - S_n$, donc

$$(*) = \frac{\mathbb{P}(X_{n+1} = y - x, S_n = x \dots S_0 = x_0)}{\mathbb{P}(S_n = x \dots S_0 = x_0)}.$$

Or $X_{n+1} \perp S_n, S_{n-1} \dots S_0$ (du fait que les (X_i) sont i.i.d.), donc

$$(*) = \frac{\mathbb{P}(X_{n+1} = y - x) \mathbb{P}(S_n = x \dots S_0 = x_0)}{\mathbb{P}(S_n = x \dots S_0 = x_0)} = \mathbb{P}(X_{n+1} = y - x) = Q(x, y),$$

où

$$Q(x, y) = \begin{cases} p & \text{si } y - x = +1, \\ 1 - p & \text{si } y - x = -1, \\ 0 & \text{sinon.} \end{cases}$$

On montre de la même manière que $\mathbb{P}(S_{n+1} = y | S_n = x) = Q(x, y)$. D'autre part, on a (mais ceci n'est pas nécessaire à la démonstration) :

$$\mu(\{x\}) = \mathbb{P}(S_0 = x) = \begin{cases} 1 & \text{si } x = 0, \\ 0 & \text{si } x \neq 0. \end{cases}$$

Martingale

Définition 1.8.3. Soit $(\Omega, \mathcal{F}, \mathbb{P})$ un espace de probabilité.

- Une filtration est une famille $(\mathcal{F}_n, n \in \mathbb{N})$ de sous-tribus de \mathcal{F} tq $\mathcal{F}_n \subset \mathcal{F}_{n+1}, \forall n \in \mathbb{N}$.
- Un processus aléatoire à temps discret $(X_n, n \in \mathbb{N})$ est adapté à la filtration $(\mathcal{F}_n, n \in \mathbb{N})$ si X_n est \mathcal{F}_n -mesurable $\forall n \in \mathbb{N}$.
- La filtration naturelle d'un processus $(X_n, n \in \mathbb{N})$ est donnée par $(\mathcal{F}_n^X, n \in \mathbb{N})$, où $\mathcal{F}_n^X = \sigma(X_i, 0 \leq i \leq n)$.

Définition 1.8.4. - Une sous-martingale (resp. sur-martingale) est un processus $(M_n, n \in \mathbb{N})$ adapté à une filtration $(\mathcal{F}_n, n \in \mathbb{N})$ tel que

$$\begin{cases} i) & \mathbb{E}(|M_n|) < \infty, \quad n \in \mathbb{N}, \\ ii) & \mathbb{E}(M_{n+1} | \mathcal{F}_n) \geq M_n \text{ p.s.} \quad (\text{resp. } \mathbb{E}(M_{n+1} | \mathcal{F}_n) \leq M_n \text{ p.s.}), \quad \forall n \in \mathbb{N}. \end{cases}$$

- Une martingale est un processus qui est à la fois une sous-martingale et une sur-martingale (i.e. un processus qui vérifie (i) et l'égalité au point (ii)).

Remarque 1.8.5. - La notion de martingale équivaut à celle de “jeu honnête” (s’il est toutefois possible de parler de jeu honnête lorsqu’on parle de jeu d’argent...) : à l’instant n (où l’on dispose de l’information \mathcal{F}_n), l’espérance de la somme possédée l’instant d’après est égale à la somme dont on dispose à l’instant présent.

- Les appellations “sous-martingale” et “sur-martingale” sont contre-intuitives : une sous-martingale est un processus qui a tendance à monter (jeu favorable), une sur-martingale est un processus qui a tendance à descendre (jeu défavorable).

Proposition 1.8.6. (*démonstration : cf. exercices*)

Soit $(M_n, n \in \mathbb{N})$ une martingale. Alors $\forall m, n \in \mathbb{N}$,

$$\begin{aligned}\mathbb{E}(M_{n+m}|\mathcal{F}_n) &= M_n \text{ p.s.} & \mathbb{E}(M_{n+1} - M_n|\mathcal{F}_n) &= 0 \text{ p.s.}, \\ \mathbb{E}(M_{n+1}) &= \mathbb{E}(\mathbb{E}(M_{n+1}|\mathcal{F}_n)) = \mathbb{E}(M_n) = \dots = \mathbb{E}(M_0).\end{aligned}$$

Exemple 1.8.7. Soit $(S_n, n \in \mathbb{N})$ une marche aléatoire symétrique à valeurs dans \mathbb{Z} ou \mathbb{R} , i.e. $S_0 = 0$, $S_n = X_1 + \dots + X_n$, avec X_i i.i.d. et $\mathbb{E}(X_1) = 0$. On pose

$$\mathcal{F}_0 = \{\emptyset, \Omega\} \text{ (la tribu triviale)} \quad \text{et} \quad \mathcal{F}_n = \sigma(X_i, 1 \leq i \leq n) \quad \forall n \geq 1.$$

Alors $(S_n, n \in \mathbb{N})$ est une martingale par rapport à $(\mathcal{F}_n, n \in \mathbb{N})$. En effet :

- (0) S_n est \mathcal{F}_n -mesurable, $\forall n \in \mathbb{N}$.
- (i) $\mathbb{E}(|S_n|) \leq \sum_{i=1}^n \mathbb{E}(|X_i|) < \infty$.
- (ii) $\mathbb{E}(S_{n+1}|\mathcal{F}_n) = \mathbb{E}(S_n|\mathcal{F}_n) + \mathbb{E}(X_{n+1}|\mathcal{F}_n) = S_n + \mathbb{E}(X_{n+1}) = S_n$ p.s.

(Dans cette série d’égalités, on a utilisé successivement : le fait que $S_{n+1} = S_n + X_{n+1}$ et la linéarité de l’espérance conditionnelle, le fait que S_n est \mathcal{F}_n -mesurable et que X_{n+1} est indépendante de \mathcal{F}_n , et finalement l’hypothèse que les v.a. X_n sont toutes centrées.)

Exemple 1.8.8. Si on reprend l’exemple ci-dessus en supposant cette fois que $\mathbb{E}(X_1) > 0$ (marche asymétrique), alors (S_n) devient une sous-martingale.

Proposition 1.8.9. Soit (M_n) une martingale par rapport à (\mathcal{F}_n) et $\varphi : \mathbb{R} \rightarrow \mathbb{R}$ convexe telle que $\mathbb{E}(|\varphi(M_n)|) < \infty$, $\forall n \in \mathbb{N}$. Alors $(\varphi(M_n))$ est une sous-martingale (par rapport à (\mathcal{F}_n)). En particulier, (M_n^2) est une sous-martingale.

Démonstration. Par l’inégalité de Jensen (pour l’espérance conditionnelle), on a

$$\mathbb{E}(\varphi(M_{n+1})|\mathcal{F}_n) \geq \varphi(\mathbb{E}(M_{n+1}|\mathcal{F}_n)) = \varphi(M_n) \quad \text{p.s.},$$

ce qui prouve la propriété (ii). Les autres vérifications sont triviales. □

Définition 1.8.10. Un temps d'arrêt par rapport à (\mathcal{F}_n) est une v.a. T à valeurs dans $\mathbb{N} \cup \{+\infty\}$ telle que $\{T \leq n\} \in \mathcal{F}_n, \forall n \in \mathbb{N}$.

Exemple 1.8.11. Si (X_n) est un processus adapté à (\mathcal{F}_n) , alors pour tout $a \in \mathbb{R}$, la v.a. T_a définie par

$$T_a = \inf\{n \in \mathbb{N} : X_n \geq a\},$$

est un temps d'arrêt, car

$$\{T_a \leq n\} = \{\exists i \in \{0, \dots, n\} \text{ tel que } X_i \geq a\} = \cup_{i=0}^n \{X_i \geq a\} \in \mathcal{F}_n,$$

car chaque événement $\{X_i \geq a\} \in \mathcal{F}_i \subset \mathcal{F}_n$, du fait que le processus (X_n) est adapté et que $i \leq n$.

Définition 1.8.12. Soit (X_n) un processus adapté à une filtration (\mathcal{F}_n) et T un temps d'arrêt (p.r. à (\mathcal{F}_n)). On pose

$$\begin{aligned} X_T(\omega) &= X_{T(\omega)}(\omega) = \sum_{n \in \mathbb{N}} X_n(\omega) 1_{\{T=n\}}(\omega), \\ \mathcal{F}_T &= \{A \in \mathcal{F} : A \cap \{T \leq n\} \in \mathcal{F}_n, \forall n \in \mathbb{N}\}. \end{aligned}$$

Remarque 1.8.13. Bien que sa définition soit peu explicite, la tribu \mathcal{F}_T représente simplement l'information dont on dispose au temps aléatoire T .

Théorème d'arrêt

Soient (M_n) une martingale par rapport à (\mathcal{F}_n) et T_1, T_2 deux temps d'arrêt tels que $0 \leq T_1(\omega) \leq T_2(\omega) \leq N < \infty, \forall \omega \in \Omega$. Alors

$$\mathbb{E}(M_{T_2} | \mathcal{F}_{T_1}) = M_{T_1} \text{ p.s.}, \quad \text{et donc} \quad \mathbb{E}(M_{T_2}) = \mathbb{E}(M_{T_1}).$$

En particulier si $0 \leq T(\omega) \leq N, \forall \omega \in \Omega$, alors $\mathbb{E}(M_T) = \mathbb{E}(M_0)$.

Remarque 1.8.14. Ce théorème reste vrai pour une sous-martingale ou une sur-martingale (avec les inégalités correspondantes).

Pour la démonstration, on aura besoin du lemme suivant :

Lemme 1.8.15. Y est une v.a. \mathcal{F}_T -mesurable ssi $Y 1_{\{T=n\}}$ est une v.a. \mathcal{F}_n -mesurable $\forall n$.

Démonstration. (théorème d'arrêt)

- Démontrons tout d'abord que si T est un temps d'arrêt tel que $0 \leq T(\omega) \leq N, \forall \omega \in \Omega$, alors

$$\mathbb{E}(M_N | \mathcal{F}_T) = \sum_{n=0}^N M_n 1_{\{T=n\}} = M_T \quad \text{p.s.} \quad (1)$$

Appelons Z le terme du milieu de l'équation ci-dessus. Pour vérifier que $\mathbb{E}(M_N|\mathcal{F}_T) = Z$, il faut vérifier les deux points de la définition de l'espérance conditionnelle :

(i) Z est \mathcal{F}_T -mesurable : effectivement, $Z 1_{\{T=n\}} = M_n 1_{\{T=n\}}$ est \mathcal{F}_n -mesurable $\forall n$, donc par le lemme, Z est \mathcal{F}_T -mesurable.

(ii) $\mathbb{E}(Z U) = \mathbb{E}(M_N U)$, \forall v.a. U \mathcal{F}_T -mesurable et bornée :

$$\begin{aligned} \mathbb{E}(Z U) &= \sum_{n=0}^N \mathbb{E}(M_n 1_{\{T=n\}} U) \stackrel{(a)}{=} \sum_{n=0}^N \mathbb{E}(\mathbb{E}(M_N|\mathcal{F}_n) 1_{\{T=n\}} U) \stackrel{(b)}{=} \sum_{n=0}^N \mathbb{E}(M_N 1_{\{T=n\}} U) \\ &= \mathbb{E}(M_N U), \end{aligned}$$

car (a) : (M_n) est une martingale et (b) : $1_{\{T=n\}} U$ est \mathcal{F}_n -mesurable par le lemme.

- En utilisant (1) avec $T = T_1$ et $T = T_2$ succesivement, on trouve donc que

$$M_{T_1} = \mathbb{E}(M_N|\mathcal{F}_{T_1}) \stackrel{(c)}{=} \mathbb{E}(\mathbb{E}(M_N|\mathcal{F}_{T_2})|\mathcal{F}_{T_1}) = \mathbb{E}(M_{T_2}|\mathcal{F}_{T_1}) \text{ p.s.},$$

où (c) découle du fait que $\mathcal{F}_{T_1} \subset \mathcal{F}_{T_2}$ (car $T_1 \leq T_2$). □

Intégrale stochastique discrète

Définition 1.8.16. *Un processus $(H_n, n \in \mathbb{N})$ est dit prévisible (par rapport à une filtration $(\mathcal{F}_n, n \in \mathbb{N})$) si $H_0 = 0$ et H_n est \mathcal{F}_{n-1} -mesurable $\forall n \geq 1$.*

Soit (H_n) un processus prévisible et (M_n) une martingale. On pose $I_0 = 0$ et

$$I_n = (H \cdot M)_n = \sum_{i=1}^n H_i (M_i - M_{i-1}), \quad n \geq 1.$$

I_n représente le gain effectué au temps n en appliquant la stratégie (H_n) sur le jeu (M_n) ; H_n représente la mise au temps n , qui ne doit logiquement dépendre que des observations passées (la tribu \mathcal{F}_{n-1}), sinon il y a délit d'initié (i.e. on a une information sur le futur du processus).

Proposition 1.8.17. *Si H_n est une v.a. bornée (i.e. $|H_n(\omega)| \leq K_n, \forall \omega \in \Omega$) $\forall n \geq 1$, alors le processus (I_n) est une martingale par rapport à (\mathcal{F}_n) .*

Démonstration. Il est clair que I_n est \mathcal{F}_n -mesurable $\forall n \in \mathbb{N}$. D'autre part, on a

$$\mathbb{E}(|I_n|) \leq \sum_{i=1}^n \mathbb{E}(|H_i \cdot (M_i - M_{i-1})|) \leq \sum_{i=1}^n K_i (\mathbb{E}(|M_i|) + \mathbb{E}(|M_{i-1}|)) < \infty,$$

$$\mathbb{E}(I_{n+1}|\mathcal{F}_n) = \mathbb{E}(I_n|\mathcal{F}_n) + \mathbb{E}(H_{n+1} (M_{n+1} - M_n)|\mathcal{F}_n) = I_n + H_{n+1} \mathbb{E}(M_{n+1} - M_n|\mathcal{F}_n) = I_n \text{ p.s.}$$

ce qui prouve les points (i) et (ii) de la définition d'une martingale. □

Remarque 1.8.18. Une réciproque à cette proposition existe : cf. exercices.

Application : Soit (M_n) la marche aléatoire symétrique sur \mathbb{Z} ($M_n = X_1 + \dots + X_n$, $\mathbb{P}(X_1 = +1) = \mathbb{P}(X_n = -1) = 1/2$). On pose

$$H_0 = 0, H_1 = 1, H_{n+1} = \begin{cases} 2H_n & \text{si } X_1 = \dots = X_n = -1, \\ 0 & \text{dès que } X_n = +1. \end{cases}$$

Alors $I_n = \sum_{i=1}^n H_i (M_i - M_{i-1}) = \sum_{i=1}^n H_i X_i$ est une martingale par la proposition ci-dessus (chacune des v.a. H_i est clairement bornée). En particulier $\mathbb{E}(I_n) = \mathbb{E}(I_0) = 0$. On définit le temps d'arrêt

$$T = \inf\{n \geq 1 : X_n = +1\}.$$

Il est alors facile de voir que $I_T = 1$, donc $\mathbb{E}(I_T) = 1 \neq \mathbb{E}(I_0) = 0$???

Effectivement, le théorème d'arrêt mentionné ci-dessus ne s'applique pas ici car T n'est pas borné (en d'autres mots, en appliquant la stratégie (H_n) , on est sûr de gagner un franc au temps T , mais le prix à payer est que ce temps T peut être arbitrairement grand, et donc qu'on risque de perdre toute notre fortune avant de gagner ce franc).

1.9 Vecteur aléatoire

Cours 5

Définition 1.9.1. Un vecteur aléatoire (de dimension $n \geq 1$) est une application

$$X : \begin{cases} \Omega \rightarrow \mathbb{R}^n \\ \omega \mapsto X(\omega) = (X_1(\omega), \dots, X_n(\omega)) \end{cases} \quad \text{telle que}$$

$$\{\omega \in \Omega : X(\omega) \in B\} \in \mathcal{F}, \quad \forall B \in \mathcal{B}(\mathbb{R}^n).$$

Proposition 1.9.2. X est un vecteur aléatoire ssi

$$\{\omega \in \Omega : X_1(\omega) \leq t_1, \dots, X_n(\omega) \leq t_n\} \in \mathcal{F}, \quad \forall t_1, \dots, t_n \in \mathbb{R}.$$

Remarque 1.9.3. Si X est un vecteur aléatoire, alors chaque X_i est une v.a. aléatoire, mais la réciproque n'est pas vraie.

Deux types particuliers de vecteurs aléatoires

A) Vecteur aléatoire discret : X prend ses valeurs dans un ensemble dénombrable D .

B) Vecteur aléatoire continu : $\mathbb{P}(\{X \in B\}) = 0$ si $|B| = 0$. Sous cette condition, le théorème de Radon-Nikodym assure l'existence d'une fonction borélienne $f_X : \mathbb{R}^n \rightarrow \mathbb{R}$ (appelée densité conjointe) telle que $f_X(x_1, \dots, x_n) \geq 0$,

$$\int_{\mathbb{R}^n} f_X(x_1, \dots, x_n) dx_1 \cdots dx_n = 1 \quad \text{et} \quad \mathbb{P}(\{X \in B\}) = \int_B f_X(x_1, \dots, x_n) dx_1 \cdots dx_n.$$

Remarque 1.9.4. A) X est un vecteur aléatoire discret ssi chaque X_i est une v.a. discrète.
 B) Si X est un vecteur aléatoire continu, alors chaque X_i est une v.a. aléatoire continue, mais la réciproque n'est pas vraie : soit U est une v.a. continue ; alors le vecteur aléatoire $X = (U, U)$ n'est pas un vecteur aléatoire continu, car pour $B = \{(x, y) \in \mathbb{R}^2 : x = y\}$, on a $\mathbb{P}(\{X \in B\}) = 1$ alors que $|B| = 0$.

Espérance et matrice de covariance d'un vecteur aléatoire

Soit X un vecteur aléatoire (tel que X_i est une v.a. de carré intégrable, $\forall i \in \{1, \dots, n\}$).
 Alors

$$\begin{cases} i) & \mathbb{E}(X) = (\mathbb{E}(X_1) \dots \mathbb{E}(X_n)), \\ ii) & \text{Cov}(X) = Q \text{ où } Q \text{ est une matrice } n \times n, Q_{ij} = \text{Cov}(X_i, X_j) = \mathbb{E}(X_i X_j) - \mathbb{E}(X_i) \mathbb{E}(X_j). \end{cases}$$

Proposition 1.9.5. *i) $Q_{ij} = Q_{ji}$ (i.e. Q est symétrique),
 ii) $\sum_{i,j=1}^n c_i c_j Q_{ij} \geq 0$, $\forall c_1 \dots c_n \in \mathbb{R}$ (i.e. Q est définie positive).*

Démonstration. La symétrie est claire. La seconde propriété découle du fait que

$$\sum_{i,j=1}^n c_i c_j Q_{ij} = \sum_{i,j=1}^n c_i c_j \text{Cov}(X_i, X_j) = \text{Cov} \left(\sum_{i=1}^n c_i X_i, \sum_{j=1}^n c_j X_j \right) = \text{Var} \left(\sum_{i=1}^n c_i X_i \right) \geq 0.$$

(On a utilisé ici la bilinéarité de la covariance.) □

Rappel : Une matrice symétrique Q est définie positive ssi toutes ses valeurs propres sont positives (i.e. ≥ 0).

Proposition 1.9.6. *Si $(X_1 \dots X_n)$ sont des v.a. indépendantes et de carré intégrable, alors $X = (X_1, \dots, X_n)$ est un vecteur aléatoire et $\text{Cov}(X)$ est une matrice diagonale (i.e. $\text{Cov}(X_i, X_j) = 0$, $\forall i \neq j$).*

Vecteur aléatoire gaussien

Convention : si $Y(\omega) \equiv c$, alors on dit par abus de langage que Y est une v.a. gaussienne (de moyenne c et de variance nulle) et on note $Y \sim \mathcal{N}(c, 0)$.

Définition 1.9.7. *Un vecteur aléatoire $X = (X_1, \dots, X_n)$ est dit gaussien si $\forall c_1, \dots, c_n \in \mathbb{R}$, la v.a. $c_1 X_1 + \dots + c_n X_n$ est une v.a. gaussienne.*

Exemple 1.9.8. Si X_1, \dots, X_n sont indépendantes et gaussiennes, alors $X = (X_1, \dots, X_n)$ est un vecteur gaussien.

Proposition 1.9.9. *Soit X un vecteur aléatoire gaussien. Alors les v.a. X_1, \dots, X_n sont indépendantes ssi la matrice $\text{Cov}(X)$ est diagonale.*

Remarque 1.9.10. - Il n'est pas obligatoire qu'un vecteur formé de deux v.a. gaussiennes X_1 et X_2 soit un vecteur gaussien.

- De la même manière, il n'est pas vrai que deux v.a. gaussiennes décorrélées sont toujours indépendantes (cf. exercices).

Notation : Si X est un vecteur gaussien de moyenne m et de (matrice de) covariance Q , on note $X \sim \mathcal{N}(m, Q)$.

Définition 1.9.11. Soit X un vecteur aléatoire gaussien de dimension n , moyenne m et covariance Q . X est dit dégénéré si $\text{rang}(Q) < n$.

Rappel : Une matrice symétrique Q vérifie :

$$\begin{aligned} \text{rang}(Q) < n \quad \text{ssi} \quad Q \text{ est non-inversible} \quad \text{ssi} \quad \det(Q) = 0 \\ \text{ssi} \quad Q \text{ admet au moins une valeur propre nulle.} \end{aligned}$$

NB : $\text{rang}(Q)$ = nombre de valeurs propres non-nulles.

Proposition 1.9.12. Soit X un vecteur aléatoire gaussien non-dégénéré de dimension n , moyenne m et covariance Q . Alors X un vecteur aléatoire continu dont la densité conjointe f_X est donnée par

$$f_X(x_1, \dots, x_n) = \frac{1}{\sqrt{(2\pi)^n \det Q}} \exp \left(-\frac{1}{2} \sum_{i,j=1}^n (x_i - m_i)(Q^{-1})_{ij}(x_j - m_j) \right).$$

NB : $\det(Q) \neq 0$ et Q^{-1} existe par hypothèse.

Proposition 1.9.13. (décomposition de Karhunen-Lowe)

Soit X un vecteur aléatoire gaussien de dimension n , moyenne m et covariance Q . Soit $k = \text{rang}(Q) \in \{1, \dots, n\}$. Alors il existe k v.a. U_1, \dots, U_k i.i.d. $\sim \mathcal{N}(0, 1)$ et des nombres réels $(\alpha_{ij})_{i,j=1}^{n,k}$ tels que

$$X_i = \sum_{j=1}^k \alpha_{ij} U_j, \quad \forall j \in \{1, \dots, n\}.$$

En mots, cette proposition dit qu'un vecteur gaussien dont la matrice de covariance est de rang k peut toujours s'exprimer en termes de k variables gaussiennes indépendantes. Cette décomposition correspond à la décomposition de la matrice Q dans la base de ses vecteurs propres (associés aux valeurs propres non-nulles).

2 Calcul stochastique

2.1 Processus aléatoire à temps continu

Définition 2.1.1. Un processus aléatoire à temps continu est une famille de v.a. $(X_t, t \in \mathbb{R}_+)$ définie sur $(\Omega, \mathcal{F}, \mathbb{P})$. On peut également le voir comme une fonction aléatoire :

$$X : \begin{cases} \Omega \rightarrow \{\text{fonctions de } \mathbb{R}_+ \text{ dans } \mathbb{R}\} \\ \omega \mapsto \{t \mapsto X_t(\omega)\} \end{cases}$$

Remarque 2.1.2. Pour caractériser une v.a. X , il suffit de donner sa loi $\mathbb{P}(X \leq x), \forall x \in \mathbb{R}$. Que faut-il pour caractériser un processus $(X_t, t \in \mathbb{R}_+)$?

- 1ère idée : donner $\mathbb{P}(X_t \leq x), \forall t \in \mathbb{R}_+, x \in \mathbb{R}$.

Ça n'est pas suffisant : si on demande p.ex. que $X_t \sim \mathcal{N}(0, t), \forall t \in \mathbb{R}_+$, alors $X_t^{(1)} = \sqrt{t}Y$, où $Y \sim N(0, 1)$ et $X_t^{(2)}$ = mouvement brownien (voir plus loin) satisfont tous deux la condition ci-dessus, mais ces deux processus ne se ressemblent pas du tout !

- 2ème idée : donner $\mathbb{P}(X_t \leq x, X_s \leq y), \forall t, s \in \mathbb{R}_+, x, y \in \mathbb{R}$.

Ça n'est pas suffisant non plus...

- n -ème idée : donner $P(X_t \leq x_1, \dots, X_{t_n} \leq x_n), \forall n \geq 1, t_1 \dots t_n \in \mathbb{R}_+, x_1 \dots x_n \in \mathbb{R}$. C'est suffisant, mais a-t-on toujours besoin de toutes ces informations pour caractériser un processus ?

Définition 2.1.3. Les v.a. $X_t - X_s, t > s \geq 0$, sont appelées des accroissements du processus (X_t) .

2.1.1 Processus à accroissements indépendants et stationnaires

(Indépendance) : $(X_t - X_s) \perp \mathcal{F}_s^X = \sigma(X_r, 0 \leq r \leq s), \forall t > s \geq 0$.

(Stationnarité) : $X_t - X_s \sim X_{t-s} - X_0, \forall t > s \geq 0$.

Pour de tels processus, donner la loi de $X_t - X_0, \forall t > 0$, ainsi que celle de X_0 suffit à caractériser entièrement le processus.

Citons encore une définition dont nous avons besoin pour définir le mouvement brownien.

Définition 2.1.4. Un processus (X_t) est appelé un processus à trajectoires continues (ou simplement processus continu) si $\mathbb{P}(\{\omega \in \Omega : t \mapsto X_t(\omega) \text{ est continue}\}) = 1$.

Attention : ne pas confondre “processus à temps continu” (=notion générale) et “processus continu” (=processus à trajectoires continues).

2.1.2 Mouvement brownien standard

Définition 2.1.5. (1ère caractérisation du mouvement brownien standard)

Un mouvement brownien standard (abrégé m.b.s.) est un processus aléatoire à temps continu $(B_t, t \in \mathbb{R}_+)$ tel que

$$\begin{cases} i) & B_0 = 0 \text{ p.s.}, \\ ii) & (B_t) \text{ est à accroissements indépendants et stationnaires}, \\ iii) & B_t \sim \mathcal{N}(0, t), \forall t > 0, \\ iv) & (B_t) \text{ est à trajectoires continues}. \end{cases}$$

Remarque 2.1.6. De cette définition, il suit que pour $t \geq s \geq 0$,

$$B_t - B_s \sim B_{t-s} \sim \mathcal{N}(0, t-s), \quad \text{i.e.} \quad \mathbb{E}(B_t - B_s) = 0 \quad \text{et} \quad \mathbb{E}((B_t - B_s)^2) = t-s.$$

En appliquant la loi des grands nombres, on trouve encore que $\frac{B_t}{t} \rightarrow 0$ p.s. lorsque $t \rightarrow \infty$. De plus, on a $\frac{B_t}{\sqrt{t}} \sim \mathcal{N}(0, 1)$, pour tout $t > 0$ (pas besoin par contre du TCL pour le prouver !). On voit donc que le m.b.s. est bien une généralisation à temps continu de la marche aléatoire symétrique sur \mathbb{Z} .

Il existe plusieurs manières de construire un mouvement brownien standard. Citons trois d'entre elles.

I. Construction de Kolmogorov

Un théorème d'existence général (le "théorème de Kolmogorov") implique qu'il existe un processus (B_t) qui vérifie (i) - (iii). A priori, il n'est pas clair que ce processus possède des trajectoires continues, mais un second théorème (baptisé "critère de continuité de Kolmogorov") permet de montrer que

$$\mathbb{E}(|B_t - B_s|^{2p}) \leq C_p |t - s|^p, \quad \forall p \geq 1, \quad \Rightarrow \quad \exists(\tilde{B}_t) \text{ tel que } \mathbb{P}(\tilde{B}_t = B_t) = 1, \forall t \in \mathbb{R},$$

et (\tilde{B}_t) est à trajectoires continues.

Le processus (\tilde{B}_t) est donc un m.b.s. Cette "construction" a le désavantage de ne pas être explicite, contrairement aux deux suivantes.

II. Principe d'invariance de Donsker

Soit $S_0 = 0$, $S_n = X_1 + \dots + X_n$, la marche aléatoire symétrique sur \mathbb{Z} (avec X_i i.i.d. et $\mathbb{P}(X_i = +1) = \mathbb{P}(X_i = -1) = \frac{1}{2}$).

Rappel : par le TCL, $\frac{S_n}{\sqrt{n}} \xrightarrow[n \rightarrow \infty]{} Z \sim \mathcal{N}(0, 1)$.

Soit $Y_t = S_{[t]} + (t - [t]) X_{[t]+1}$ (i.e. si $t = n + \varepsilon$, alors $Y_t = S_n + \varepsilon X_{n+1}$). On pose $B_t^{(n)} = \frac{Y_{nt}}{\sqrt{n}}$. Supposons pour simplifier que $nt \in \mathbb{N}$. Alors on a

$$B_t^{(n)} = \frac{S_{nt}}{\sqrt{n}} = \sqrt{t} \frac{S_{nt}}{\sqrt{nt}} \xrightarrow{n \rightarrow \infty} \sqrt{t} Z \sim \mathcal{N}(0, t), \quad \text{car } Z \sim \mathcal{N}(0, 1),$$

ce qui veut dire que

$$\mathbb{P}(B_t^{(n)} \leq x) \xrightarrow{n \rightarrow \infty} \mathbb{P}(B_t \leq x).$$

On peut montrer de manière similaire que $\forall m \geq 1, t_1, \dots, t_m \in \mathbb{R}_+, x_1, \dots, x_m \in \mathbb{R}$,

$$\mathbb{P}(B_{t_1}^{(n)} \leq x_1, \dots, B_{t_m}^{(n)} \leq x_m) \xrightarrow{n \rightarrow \infty} \mathbb{P}(B_{t_1} \leq x_1, \dots, B_{t_m} \leq x_m),$$

i.e. que la suite de processus $(B_t^{(n)})$ converge en loi vers le m.b.s. (B_t) .

III. Construction par série de Fourier

Soit $t \in [0, \pi]$. On pose

$$B_t = \frac{\sqrt{8}}{\pi} \sum_{n=1}^{\infty} \frac{\sin(nt)}{n} \xi_n,$$

où (ξ_n) est une suite de v.a. i.i.d. $\sim \mathcal{N}(0, 1)$. Alors B_t est une v.a. gaussienne (car c'est une somme de v.a. gaussiennes indépendantes), $\mathbb{E}(B_t) = 0$ et

$$\mathbb{E}(B_t^2) = \frac{8}{\pi^2} \sum_{n,m=1}^{\infty} \frac{\sin(nt)}{n} \frac{\sin(mt)}{n} \mathbb{E}(\xi_n \xi_m) = \frac{8}{\pi^2} \sum_{n=1}^{\infty} \frac{(\sin(nt))^2}{n^2} = t.$$

On peut également vérifier que le processus (B_t) ainsi défini a toutes les propriétés d'un m.b.s. A l'aide de cette troisième construction, effectuons un petit calcul (formel) :

$$\frac{\partial B_t}{\partial t} = \frac{\sqrt{8}}{\pi} \sum_{n=1}^{\infty} \frac{\partial}{\partial t} \left(\frac{\sin(nt)}{n} \right) \xi_n = \frac{\sqrt{8}}{\pi} \sum_{n=1}^{\infty} \cos(nt) \xi_n.$$

La v.a. $\frac{\partial B_t}{\partial t}$ est donc une v.a. gaussienne (car c'est une somme de v.a. gaussiennes indépendantes), mais

$$\mathbb{E} \left(\left(\frac{\partial B_t}{\partial t} \right)^2 \right) = \frac{8}{\pi^2} \sum_{n=1}^{\infty} (\cos(nt))^2 = \infty.$$

Il y a donc une contradiction, car une v.a. gaussienne ne peut pas avoir une variance infinie. En conclusion, la v.a. $\frac{\partial B_t}{\partial t}$ n'est pas bien définie, i.e. la fonction $t \mapsto B_t$ est continue mais pas dérivable. On reverra ceci de manière plus rigoureuse ci-après.

Remarque 2.1.7. Il est toutefois possible de définir la dérivée du mouvement brownien “au sens des distributions” ; l'objet représentant la dérivée est appelé “bruit blanc”.

Transformations du mouvement brownien standard (cf. exercices)

Soit $(B_t, t \in \mathbb{R}_+)$ un mouvement brownien standard. Alors les cinq processus ci-dessous sont également des mouvements browniens standard :

- 1) $B_t^{(1)} = -B_t, t \in \mathbb{R}_+$ (\leftrightarrow propriété de symétrie du mouvement brownien) ;
- 2) soit $T \in \mathbb{R}_+$ fixé : $B_t^{(2)} = B_{t+T} - B_T, t \in \mathbb{R}_+$ (\leftrightarrow stationnarité) ;
- 3) soit $T \in \mathbb{R}_+$ fixé : $B_t^{(3)} = B_T - B_{T-t}, t \in [0, T]$ (\leftrightarrow renversement du temps) ;
- 4) soit $a > 0$ fixé : $B_t^{(4)} = \frac{1}{\sqrt{a}} B_{at}, t \in \mathbb{R}_+$ (\leftrightarrow loi d'échelle) ;
- 5) $B_t^{(5)} = t B_{\frac{1}{t}}, t > 0$ et $B_0^{(5)} := 0$ (\leftrightarrow inversion du temps).

A cause de la propriété 4), on appelle le mouvement brownien standard un “fractal aléatoire” ou encore un “processus auto-similaire”, i.e. un processus qui garde le même aspect à différentes échelles.

2.1.3 Processus gaussien

Définition 2.1.8. *Un processus gaussien est un processus $(X_t, t \in \mathbb{R}_+)$ tel que $(X_{t_1}, \dots, X_{t_n})$ est un vecteur aléatoire gaussien $\forall n \geq 1$ et $t_1, \dots, t_n \in \mathbb{R}_+$. Ceci revient à dire que $c_1 X_{t_1} + \dots + c_n X_{t_n}$ est une v.a. gaussienne $\forall n \geq 1, t_1, \dots, t_n \in \mathbb{R}_+$ et $c_1, \dots, c_n \in \mathbb{R}$.*

Pour un vecteur aléatoire (pas forcément gaussien), on définit encore :

- La fonction $m : \mathbb{R}_+ \rightarrow \mathbb{R}$ donnée par $m(t) = \mathbb{E}(X_t)$ et appelée la moyenne du processus.
- La fonction $K : \mathbb{R}_+ \times \mathbb{R}_+ \rightarrow \mathbb{R}$ donnée par $K(t, s) = \text{Cov}(X_t, X_s)$ et appelée la covariance du processus.

Proposition 2.1.9. *(valide pour tout processus aléatoire)*

K est symétrique : $K(t, s) = K(s, t), \forall s, t \in \mathbb{R}_+$.

K est définie positive : $\sum_{i,j=1}^n c_i c_j K(t_i, t_j) \geq 0, \forall n \geq 1, c_1, \dots, c_n \in \mathbb{R}, t_1, \dots, t_n \in \mathbb{R}_+$.

La démonstration de cette proposition est identique à celle donnée pour un vecteur aléatoire.

Proposition 2.1.10. *(Kolmogorov)*

Etant donné $m : \mathbb{R}_+ \rightarrow \mathbb{R}$ et $K : \mathbb{R}_+ \times \mathbb{R}_+ \rightarrow \mathbb{R}$ symétrique et définie positive, il existe un processus gaussien $(X_t, t \in \mathbb{R}_+)$ de moyenne m et de covariance K . De plus, m et K caractérisent entièrement le processus (X_t) .

Proposition 2.1.11. *(2ème caractérisation du mouvement brownien standard)*

Un mouvement brownien standard $(B_t, t \in \mathbb{R}_+)$ est un processus à trajectoires continues et gaussien avec moyenne $m(t) = 0$ et covariance $K(t, s) = t \wedge s = \min(t, s)$.

Démonstration. Il faudrait vérifier que $c_1 B_{t_1} + \dots + c_n B_{t_n}$ est une v.a. gaussienne. Vérifions seulement que $B_t + B_s$ est gaussienne : $B_t + B_s = (B_t - B_s) + 2B_s$, c'est donc une v.a. gaussienne, car la somme de deux v.a. gaussiennes indépendantes est encore gaussienne. Soit maintenant $t \geq s \geq 0$:

$$\begin{aligned} m(t) &= \mathbb{E}(B_t) = 0, \\ K(t, s) &= \text{Cov}(B_t, B_s) = \mathbb{E}(B_t B_s) = \mathbb{E}((B_t - B_s) B_s) \\ &= \mathbb{E}((B_t - B_s) B_s) + \mathbb{E}(B_s^2) = 0 + s, \end{aligned}$$

car $(B_t - B_s) \perp B_s$. On a donc $K(t, s) = \min(t, s)$. □

2.1.4 Processus de Markov

Cours 6

Processus $(X_t, t \in \mathbb{R}_+)$ tel que

$$\mathbb{E}(f(X_t) | \mathcal{F}_s^X) = \mathbb{E}(f(X_t) | X_s) \quad \text{p.s.},$$

pour tout $t > s \geq 0$ et pour toute fonction $f : \mathbb{R} \rightarrow \mathbb{R}$ borélienne et bornée (NB : $\mathcal{F}_s^X = \sigma(X_r, r \leq s)$).

En particulier, si $f(x) = 1_B(x)$ avec $B \in \mathcal{B}(\mathbb{R})$, alors $\mathbb{P}(X_t \in B | \mathcal{F}_s^X) = \mathbb{P}(X_t \in B | X_s)$.

Proposition 2.1.12. *Le mouvement brownien standard $(B_t, t \in \mathbb{R}_+)$ est un processus de Markov.*

Démonstration. En utilisant le fait que $(B_t - B_s) \perp B_s$, B_s est \mathcal{F}_s^B -mesurable et la proposition 1.6.7, on obtient

$$\mathbb{E}(f(B_t) | \mathcal{F}_s^B) = \mathbb{E}(f(B_t - B_s + B_s) | \mathcal{F}_s^B) = \psi(B_s) \quad \text{p.s.},$$

où $\psi(y) = \mathbb{E}(f(B_t - B_s + y))$. Un calcul identique montre que $\mathbb{E}(f(B_t) | B_s) = \psi(B_s)$ p.s. et donc la propriété de Markov est démontrée. □

Remarque 2.1.13. De la démonstration ci-dessus, on déduit que

$$\mathbb{E}(f(B_t) | \mathcal{F}_s^B) = \psi(B_s) \quad \text{p.s.}, \quad \text{où } \psi(y) = \mathbb{E}(f(X + y)),$$

avec $X \sim \mathcal{N}(0, t - s)$. Ceci montre qu'une expression a priori compliquée faisant intervenir une espérance conditionnelle d'une fonction du mouvement brownien peut se réduire à une simple espérance d'une fonction d'une loi gaussienne.

2.1.5 Martingale à temps continu

Définition 2.1.14. Soit $(\Omega, \mathcal{F}, \mathbb{P})$ un espace de probabilité.

- Une filtration est une famille $(\mathcal{F}_t, t \in \mathbb{R}_+)$ de sous-tribus de \mathcal{F} tq $\mathcal{F}_s \subset \mathcal{F}_t, \forall t \geq s \geq 0$.
- Un processus (X_t) est adapté à (\mathcal{F}_t) si X_t est \mathcal{F}_t -mesurable $\forall t \geq 0$.
- La filtration naturelle d'un processus $(X_t, t \in \mathbb{R}_+)$ est donnée par $(\mathcal{F}_t^X, t \in \mathbb{R}_+)$, où $\mathcal{F}_t^X = \sigma(X_s, 0 \leq s \leq t)$.

Définition 2.1.15. Un processus (M_t) adapté à (\mathcal{F}_t) tel que

$$\begin{cases} i) & \mathbb{E}(|M_t|) < \infty, & \forall t \geq 0, \\ ii) & \mathbb{E}(M_t | \mathcal{F}_s) = M_s \text{ p.s., } & \forall t > s \geq 0, \end{cases}$$

est appelé une martingale (à temps continu). On définit de manière similaire une sous-martingale et une sur-martingale (à temps continu), avec les inégalités correspondantes.

Proposition 2.1.16. Le mouvement brownien standard $(B_t, t \in \mathbb{R}_+)$ est une martingale par rapport à sa filtration naturelle $(\mathcal{F}_t^B, t \in \mathbb{R}_+)$.

Démonstration. i) Par Cauchy-Schwarz, on a $\mathbb{E}(|B_t|) \leq \sqrt{\mathbb{E}(B_t^2)} = \sqrt{t} < \infty, \quad \forall t \geq 0$.

ii) $\forall t > s \geq 0$, on a $\mathbb{E}(B_t | \mathcal{F}_s^B) = \mathbb{E}(B_t - B_s | \mathcal{F}_s^B) + \mathbb{E}(B_s | \mathcal{F}_s^B) = \mathbb{E}(B_t - B_s) + B_s = B_s$. \square

Remarque 2.1.17. Pour des raisons techniques, on a parfois besoins “d’augmenter” la filtration naturelle (\mathcal{F}_t^B) . On ne rentrera pas dans ces détails dans ce cours.

Proposition 2.1.18. (démonstration : cf. exercices)

Les processus suivants sont des martingales par rapport à (\mathcal{F}_t^B) :

$$\begin{cases} i) & M_t = B_t^2 - t, \\ ii) & N_t = \exp(B_t - \frac{t}{2}). \end{cases}$$

Théorème de Lévy (3ème caractérisation du mouvement brownien standard)

Soit (X_t) un processus à trajectoires continues, adapté à une filtration (\mathcal{F}_t) et tel que

$$\begin{cases} i) & (X_t) \text{ est une martingale par rapport à } (\mathcal{F}_t), \\ ii) & (X_t^2 - t) \text{ est une martingale par rapport à } (\mathcal{F}_t), \end{cases}$$

Alors (X_t) est un mouvement brownien standard.

Définition 2.1.19. Soit $(\mathcal{F}_t, t \in \mathbb{R}_+)$ une filtration.

- Un temps d’arrêt T par rapport à (\mathcal{F}_t) est une v.a. T à valeurs dans $\mathbb{R}_+ \cup \{+\infty\}$ telle que $\{T \leq t\} \in \mathcal{F}_t, \forall t \in \mathbb{R}_+$.
- On définit

$$\mathcal{F}_T = \{A \in \mathcal{F} : A \cap \{T \leq t\} \in \mathcal{F}_t, \forall t \in \mathbb{R}_+\}.$$

- Si (X_t) est un processus adapté à la filtration (\mathcal{F}_t) , on pose $X_T(\omega) = X_{T(\omega)}(\omega)$.

Théorème d'arrêt

Soient (M_t) une martingale continue (par rapport à (\mathcal{F}_t)) et T_1, T_2 deux temps d'arrêt tels que $0 \leq T_1(\omega) \leq T_2(\omega) \leq K < \infty, \forall \omega \in \Omega$. Alors

$$\mathbb{E}(M_{T_2} | \mathcal{F}_{T_1}) = M_{T_1} \quad \text{p.s.} \quad \text{et donc} \quad \mathbb{E}(M_{T_2}) = \mathbb{E}(M_{T_1}).$$

En particulier, si $0 \leq T(\omega) \leq K, \forall \omega$, alors $\mathbb{E}(M_T) = \mathbb{E}(M_0)$.

Remarque 2.1.20. - Bien que l'énoncé soit en tout point identique au théorème dans le cas discret, la démonstration du théorème à temps continu est beaucoup plus laborieuse !

- Le théorème est encore valide pour une sous-martingale et une sur-martingale (avec les inégalités correspondantes). C'est ce que nous allons utiliser pour démontrer les inégalités de Doob ci-dessous.

Terminologie : Une martingale (M_t) est dite de carré intégrable si $\mathbb{E}(M_t^2) < \infty, \forall t \in \mathbb{R}_+$.

Inégalités de Doob

Soit (M_t) une martingale (par rapport à une filtration (\mathcal{F}_t)) continue, de carré intégrable et telle que $M_0 = 0$ p.s. Alors

$$\left\{ \begin{array}{l} a) \quad \mathbb{P}(\sup_{0 \leq s \leq t} |M_s| \geq \lambda) \leq \frac{\mathbb{E}(|M_t|)}{\lambda}, \quad \forall t > 0, \lambda > 0, \\ b) \quad \mathbb{E}(\sup_{0 \leq s \leq t} |M_s|^2) \leq 4 \mathbb{E}(|M_t|^2), \quad \forall t > 0. \end{array} \right.$$

Remarque 2.1.21. Ces inégalités sont précieuses, car elles permettent de borner (en espérance) le supremum d'une martingale sur un intervalle par quelque chose qui ne dépend que de la valeur de la martingale à la fin de l'intervalle.

Démonstration. (inégalités de Doob)

a) Par l'inégalité de Jensen, le processus $(|M_t|, t \in \mathbb{R}_+)$ est une sous-martingale. On définit

$$T_1 = \inf\{s \geq 0 : |M_s| \geq \lambda\} \wedge t, \quad T_2 = t.$$

On voit alors que $0 \leq T_1 \leq T_2 = t$. En appliquant le théorème d'arrêt ci-dessus à la sous-martingale $(|M_t|)$, on trouve donc que

$$|M_{T_1}| \leq \mathbb{E}(|M_t| | \mathcal{F}_{T_1}).$$

En multipliant par $1_{\{|M_{T_1}| \geq \lambda\}}$, on trouve encore

$$|M_{T_1}| 1_{\{|M_{T_1}| \geq \lambda\}} \leq \mathbb{E}(|M_t| 1_{\{|M_{T_1}| \geq \lambda\}} | \mathcal{F}_{T_1}),$$

car $1_{\{|M_{T_1}| \geq \lambda\}}$ est \mathcal{F}_{T_1} -mesurable. De là, on déduit que

$$\lambda \mathbb{P}(\{|M_{T_1}| \geq \lambda\}) = \mathbb{E}(\lambda 1_{\{|M_{T_1}| \geq \lambda\}}) \leq \mathbb{E}(|M_{T_1}| 1_{\{|M_{T_1}| \geq \lambda\}}) \leq \mathbb{E}(|M_t| 1_{\{|M_{T_1}| \geq \lambda\}}).$$

Soit maintenant $M^* = \sup_{0 \leq s \leq t} |M_s|$: il importe de remarquer que

$$\{|M_{T_1}| \geq \lambda\} = \{M^* \geq \lambda\}.$$

(Effectivement, chacun des deux événements ci-dessus correspond à l'événement "le processus $|M_t|$ a dépassé la valeur λ dans l'intervalle $[0, t]$ "). Ceci implique que

$$\lambda \mathbb{P}(\{M^* \geq \lambda\}) \leq \mathbb{E}(|M_t| 1_{\{M^* \geq \lambda\}}) \leq \mathbb{E}(|M_t|) \quad (2)$$

et donc l'inégalité a).

b) Pour simplifier, supposons que l'on sait que $\mathbb{E}((M^*)^2) < \infty$ (ça n'est pas clair a priori). Alors du fait que (cf. exercices)

$$g(M^*) = \int_0^\infty g'(x) 1_{\{M^* \geq x\}} dx,$$

on a par l'inégalité de gauche dans (2),

$$\mathbb{E}((M^*)^2) = \int_0^\infty 2x \mathbb{P}(\{M^* \geq x\}) dx \leq \int_0^\infty 2 \mathbb{E}(|M_t| 1_{\{M^* \geq x\}}) dx = 2 \mathbb{E}(|M_t| M^*).$$

L'inégalité de Cauchy-Schwarz permet finalement de conclure que

$$\mathbb{E}((M^*)^2) \leq 2 \sqrt{\mathbb{E}(|M_t|^2)} \sqrt{\mathbb{E}(M^*)^2},$$

autrement dit,

$$\sqrt{\mathbb{E}((M^*)^2)} \leq 2 \sqrt{\mathbb{E}(M_t^2)} \quad \text{i.e.} \quad \mathbb{E}((M^*)^2) \leq 4 \mathbb{E}(M_t^2),$$

ce qui termine la démonstration. □

2.2 Intégrale de Riemann-Stieltjes

Intégrale de Riemann

Soit $t > 0$ et $f : \mathbb{R}_+ \rightarrow \mathbb{R}$ continue. On définit

$$\int_0^t f(s) ds = \lim_{n \rightarrow \infty} \sum_{i=1}^n f(s_i^{(n)}) (t_i^{(n)} - t_{i-1}^{(n)}),$$

où $0 = t_0^{(n)} < t_1^{(n)} < \dots < t_n^{(n)} = t$, $s_i^{(n)} \in [t_{i-1}^{(n)}, t_i^{(n)}]$ et $(t_0^{(n)}, \dots, t_n^{(n)})$ est une suite de partitions de $[0, t]$ telle que $\lim_{n \rightarrow \infty} \max_{1 \leq i \leq n} |t_i^{(n)} - t_{i-1}^{(n)}| = 0$.

Fonction (localement) à variation bornée

Fonction $g : \mathbb{R}_+ \rightarrow \mathbb{R}$ telle que $\forall t > 0$,

$$\sup \sum_{i=1}^n |g(t_i) - g(t_{i-1})| < \infty,$$

où le supremum est pris sur toutes les partitions (t_0, \dots, t_n) de $[0, t]$, avec n arbitraire.

Remarque 2.2.1. - Si g est croissante, alors g est à variation bornée, car

$$\sum_{i=1}^n |g(t_i) - g(t_{i-1})| = \sum_{i=1}^n (g(t_i) - g(t_{i-1})) = g(t) - g(0)$$

est indépendant de la partition choisie, donc

$$\sup \sum_{i=1}^n |g(t_i) - g(t_{i-1})| = g(t) - g(0) < \infty.$$

- Si g est la différence de deux fonctions croissantes, alors g est à variation bornée. (La démonstration est similaire à ce qui précède.)

- Si g est continûment dérivable, alors g est à variation bornée, car

$$\begin{aligned} \sum_{i=1}^n |g(t_i) - g(t_{i-1})| &= \sum_{i=1}^n \left| \int_{t_{i-1}}^{t_i} g'(s) ds \right| \leq \sum_{i=1}^n \sup_{s \in [t_{i-1}, t_i]} |g'(s)| (t_i - t_{i-1}) \\ &\leq \sup_{s \in [0, t]} |g'(s)| \sum_{i=1}^n (t_i - t_{i-1}) = \sup_{s \in [0, t]} |g'(s)| t \end{aligned}$$

est indépendant de la partition choisie, donc

$$\sup \sum_{i=1}^n |g(t_i) - g(t_{i-1})| \leq \sup_{s \in [0, t]} |g'(s)| t < \infty.$$

Intégrale de Riemann-Stieltjes

Soit $t > 0$, $f : \mathbb{R}_+ \rightarrow \mathbb{R}$ continue et $g : \mathbb{R}_+ \rightarrow \mathbb{R}$ à variation bornée. On définit

$$\int_0^t f(s) dg(s) = \lim_{n \rightarrow \infty} \sum_{i=1}^n f(s_i^{(n)}) (g(t_i^{(n)}) - g(t_{i-1}^{(n)})),$$

où $0 = t_0^{(n)} < t_1^{(n)} < \dots < t_n^{(n)} = t$, $s_i^{(n)} \in [t_{i-1}^{(n)}, t_i^{(n)}]$ et $(t_0^{(n)}, \dots, t_n^{(n)})$ est une suite de partitions de $[0, t]$ telle que $\lim_{n \rightarrow \infty} \max_{1 \leq i \leq n} |t_i^{(n)} - t_{i-1}^{(n)}| = 0$.

Proposition 2.2.2. - Si f est continue et g est continûment dérivable, alors

$$\int_0^t f(s) dg(s) = \int_0^t f(s) g'(s) ds.$$

- Si f et g sont continues et à variation bornée, alors

$$\begin{aligned} \int_0^t f(s) df(s) &= \frac{f(t)^2}{2} - \frac{f(0)^2}{2}, \\ \int_0^t f(s) dg(s) &= f(t)g(t) - f(0)g(0) - \int_0^t g(s) df(s). \end{aligned}$$

Passons des fonctions déterministes aux processus aléatoires

Soit $(H_t, t \in \mathbb{R}^+)$ un processus à trajectoires continues.

Soit $(V_t, t \in \mathbb{R}^+)$ un processus (à trajectoires) à variation bornée.

On définit

$$\left(\int_0^t H_s dV_s \right) (\omega) = \int_0^t H_s(\omega) dV_s(\omega)$$

Cette intégrale est définie trajectoire par trajectoire (i.e. “ ω par ω ”).

Remarque 2.2.3. Si la courbe (V_t) est l'évolution du prix d'un actif financier et (H_t) est la stratégie d'investissement sur cet actif, alors $\int_0^t H_s dV_s$ représente le gain effectué sur l'intervalle $[0, t]$. L'intégrale ci-dessus est donc bien définie et l'on peut se demander pourquoi on a besoin de parler d'intégrale “stochastique” ?

Il se trouve qu'en mathématiques financières, les processus d'évolution des prix sont mieux représentés par des mouvements browniens (ou des processus plus compliqués encore). Or le mouvement brownien (standard) a des trajectoires continues mais pas à variation bornée (voir plus loin). Comment définir alors $\int_0^t H_s dB_s$?

Une fausse solution : si on suppose (H_t) à variation bornée, alors on pourrait définir

$$\int_0^t H_s dB_s = H_t B_t - H_0 B_0 - \int_0^t B_s dH_s.$$

Le terme de droite est bien défini, car (H_t) est à variation bornée et (B_t) est continu. Mais comment définir alors $\int_0^t B_s dB_s$???

2.3 Variation quadratique

Cours 7

Variation quadratique du mouvement brownien standard

Soit $(B_t, t \in \mathbb{R}_+)$ un mouvement brownien standard (par convenance, on notera parfois $B_t = B(t)$). Pour $t > 0$, on définit

$$\langle B \rangle_t^{(n)} = \sum_{i=1}^{2^n} \left(B\left(\frac{it}{2^n}\right) - B\left(\frac{(i-1)t}{2^n}\right) \right)^2.$$

Proposition - Définition. Pour $t > 0$,

$$\lim_{n \rightarrow \infty} \langle B \rangle_t^{(n)} = t \quad \text{p.s.}$$

et on définit la variation quadratique du mouvement brownien standard $\langle B \rangle_t$ comme étant donnée par cette limite (par convention, on pose également $\langle B \rangle_0 = 0$).

Démonstration. Soient $X_i = B\left(\frac{it}{2^n}\right) - B\left(\frac{(i-1)t}{2^n}\right)$, $1 \leq i \leq 2^n$ (t et n fixés).

Les v.a. X_i sont i.i.d. $\sim N(0, t/2^n)$ et $\langle B \rangle_t^{(n)} = \sum_{i=1}^{2^n} X_i^2$. De plus, on a

$$\begin{aligned} \mathbb{E}(\langle B \rangle_t^{(n)}) &= \sum_{i=1}^{2^n} \mathbb{E}(X_i^2) = \sum_{i=1}^{2^n} \frac{t}{2^n} = t, \\ \text{Var}(\langle B \rangle_t^{(n)}) &= \sum_{i=1}^{2^n} \text{Var}(X_i^2) = \sum_{i=1}^{2^n} (\mathbb{E}(X_i^4) - \mathbb{E}(X_i^2)^2) \\ &= \sum_{i=1}^{2^n} \left(3 \frac{t^2}{4^n} - \frac{t^2}{4^n} \right) = \frac{t^2}{2^{n-1}}. \end{aligned}$$

En conséquence,

$$\sum_{n=1}^{\infty} \mathbb{P}(|\langle B \rangle_t^{(n)} - t| > \varepsilon) \leq \frac{1}{\varepsilon^2} \sum_{n=1}^{\infty} \text{Var}(\langle B \rangle_t^{(n)}) = \frac{1}{\varepsilon^2} \sum_{n=1}^{\infty} \frac{t^2}{2^{n-1}} < \infty.$$

Par le lemme de Borel-Cantelli (a) et le critère donné pour la convergence presque sûre, on a donc

$$\langle B \rangle_t^{(n)} \xrightarrow[n \rightarrow \infty]{} t \quad \text{p.s.}$$

□

Corollaire 2.3.1.

$$\lim_{n \rightarrow \infty} \sum_{i=1}^{2^n} \left| B\left(\frac{it}{2^n}\right) - B\left(\frac{(i-1)t}{2^n}\right) \right| = \infty \quad \text{p.s.}$$

donc le processus (B_t) n'est pas à variation bornée.

Démonstration. Supposons par l'absurde que $\mathbb{P}(\lim_{n \rightarrow \infty} \sum_{i=1}^n |B(\frac{it}{2^n}) - B(\frac{(i-1)t}{2^n})| < \infty) > 0$. Alors en reprenant la notation de la démonstration précédente, on a

$$\langle B \rangle_t = \lim_{n \rightarrow \infty} \langle B \rangle_t^{(n)} = \lim_{n \rightarrow \infty} \sum_{i=1}^{2^n} X_i^2 \leq \left(\lim_{n \rightarrow \infty} \max_{1 \leq i \leq 2^n} |X_i| \right) \left(\lim_{n \rightarrow \infty} \sum_{i=1}^{2^n} |X_i| \right).$$

Vu que le processus (B_t) est continu, il est uniformément continu sur $[0, t]$, et donc

$$\lim_{n \rightarrow \infty} \max_{1 \leq i \leq 2^n} |X_i| = 0 \quad \text{p.s.}$$

D'un autre côté, on a supposé que

$$\mathbb{P} \left(\lim_{n \rightarrow \infty} \sum_{i=1}^{2^n} |X_i| < \infty \right) > 0.$$

De la première inégalité, on conclut donc que

$$\mathbb{P}(\langle B \rangle_t = 0) > 0,$$

ce qui est clairement en contradiction avec le fait que $\langle B \rangle_t = t$ p.s., démontré ci-dessus. \square

Remarque 2.3.2. Le fait que le processus (B_t) n'est pas à variation bornée implique que ses trajectoires ne sont pas dérivables.

Remarque 2.3.3. En appliquant la proposition 2.1.18 (i), on remarque que le processus $(B_t^2 - \langle B \rangle_t, t \in \mathbb{R}_+)$ est une martingale. Ceci n'est pas un hasard. Cette propriété va même nous servir à définir la variation quadratique d'une martingale continue de carré intégrable.

Théorème de décomposition de Doob

Soit $(X_t, t \in \mathbb{R}_+)$ une sous-martingale continue (par rapport à une filtration $(\mathcal{F}_t, t \in \mathbb{R}_+)$). Alors il existe un unique processus $(A_t, t \in \mathbb{R}_+)$ croissant, continu et adapté à $(\mathcal{F}_t, t \in \mathbb{R}_+)$ tel que $A_0 = 0$ et $(X_t - A_t, t \in \mathbb{R}_+)$ est une martingale.

Application : définition de la variation quadratique d'une martingale

Soit (M_t) une martingale continue de carré intégrable (par rapport à une filtration (\mathcal{F}_t)). Alors (M_t^2) est une sous-martingale et donc, d'après le théorème ci-dessus, il existe un unique processus (A_t) croissant, continu et adapté à (\mathcal{F}_t) tel que $A_0 = 0$ et $(M_t^2 - A_t)$ est une martingale. On note $A_t = \langle M \rangle_t$ et on appelle ce processus la variation quadratique de (M_t) .

Exemple 2.3.4. On a vu que $\langle B \rangle_t = t$. Noter que dans ce cas, la variation quadratique est un processus déterministe.

Remarque 2.3.5. - Noter qu'on a toujours par définition : $\mathbb{E}(\langle M \rangle_t) = \mathbb{E}(M_t^2) - \mathbb{E}(M_0^2)$.
- Du fait que le processus $(\langle M \rangle_t)$ est croissant, c'est un processus à variation bornée (i.e. la variation quadratique d'une martingale est un processus à variation bornée).
- Si (M_t) est une martingale continue à variation bornée, alors $M_t = M_0$ pour tout $t > 0$ (i.e. M_t est constante).

Proposition 2.3.6. *Si (M_t) est une martingale continue de carré intégrable, alors on a :*

$$\langle M \rangle_t^{(n)} = \sum_{i=1}^{2^n} \left(M\left(\frac{it}{2^n}\right) - M\left(\frac{(i-1)t}{2^n}\right) \right)^2 \xrightarrow[n \rightarrow \infty]{\mathbb{P}} \langle M \rangle_t, \quad \forall t \geq 0.$$

Remarque 2.3.7. - A priori, la convergence n'a lieu qu'en probabilité (alors qu'elle a lieu presque sûrement dans le cas où (M_t) est un mouvement brownien standard).
- Bien que la variation quadratique soit une quantité aléatoire, la proposition ci-dessus illustre le fait qu'elle est une généralisation de la notion de variance pour des processus aléatoires.

Définition 2.3.8. *Soient (M_t) , (N_t) deux martingales continues de carré intégrable. On définit la covariation quadratique de (M_t) et (N_t) par*

$$\langle M, N \rangle_t = \frac{1}{4} (\langle M + N \rangle_t - \langle M - N \rangle_t).$$

Proposition 2.3.9. *(démonstration : cf. exercices)
Le processus $(M_t N_t - \langle M, N \rangle_t)$ est une martingale.*

Proposition 2.3.10. *Si (M_t) et (N_t) sont deux martingales continues de carré intégrable, alors on a :*

$$\langle M, N \rangle_t^{(n)} = \sum_{i=1}^{2^n} \left(M\left(\frac{it}{2^n}\right) - M\left(\frac{(i-1)t}{2^n}\right) \right) \left(N\left(\frac{it}{2^n}\right) - N\left(\frac{(i-1)t}{2^n}\right) \right) \xrightarrow[n \rightarrow \infty]{\mathbb{P}} \langle M, N \rangle_t, \quad \forall t \geq 0.$$

Remarque 2.3.11. De la même manière que ci-dessus, on voit que la covariation quadratique est en quelque sorte une généralisation de la covariance pour des processus aléatoires. En particulier, elle est bilinéaire et on a également la proposition suivante.

Proposition 2.3.12. *(cf. proposition 1.5.12)
Soient (M_t) , (N_t) deux martingales continues de carré intégrable indépendantes et soit $c \in \mathbb{R}$. Alors*

$$\langle M, N \rangle_t = 0 \quad \text{et} \quad \langle cM + N \rangle_t = c^2 \langle M \rangle_t + \langle N \rangle_t.$$

La première égalité vient du fait que si (M_t) et (N_t) sont deux martingales indépendantes de carré intégrable, alors $(M_t N_t)$ est une martingale (mais la démonstration de ce fait est plus longue qu'on pourrait imaginer).

Changement de temps

Définition 2.3.13. On appelle mouvement brownien standard par rapport à une filtration (\mathcal{F}_t) un mouvement brownien standard (B_t) adapté à (\mathcal{F}_t) et tel que

$$(B_t - B_s) \perp \mathcal{F}_s, \quad \forall t > s \geq 0.$$

On est maintenant en mesure de donner une version différente du théorème de Lévy.

Théorème de Lévy (seconde version)

Soit (M_t) une martingale continue de carré intégrable par rapport à une filtration (\mathcal{F}_t) telle que

$$\langle M \rangle_t = t \quad \text{p.s.}, \quad \forall t \in \mathbb{R}_+.$$

Alors (M_t) est un mouvement brownien standard par rapport à (\mathcal{F}_t) .

Proposition 2.3.14. Soit (M_t) une martingale continue de carré intégrable (par rapport à une filtration (\mathcal{F}_t)) et telle que

$$\lim_{t \rightarrow \infty} \langle M \rangle_t = \infty \quad \text{p.s.}$$

Soit $T(s) = \inf\{t \geq 0 : \langle M \rangle_t \geq s\}$. On pose

$$\mathcal{G}_s = \mathcal{F}_{T(s)} \quad \text{et} \quad B_s = M_{T(s)}, \quad s \in \mathbb{R}_+.$$

Alors le processus (B_s) est un mouvement brownien standard par rapport à (\mathcal{G}_s) .

Démonstration. On utilise la version du théorème de Lévy mentionnée ci-dessus. Par continuité de l'application $s \mapsto T(s)$ (qui provient de la continuité de $t \mapsto \langle M \rangle_t$), le processus (B_s) est continu. D'autre part, pour $s_2 > s_1 \geq 0$, on a, par le théorème d'arrêt :

$$\mathbb{E}(B_{s_2} - B_{s_1} | \mathcal{G}_{s_1}) = \mathbb{E}(M_{T(s_2)} - M_{T(s_1)} | \mathcal{F}_{T(s_1)}) = 0.$$

(NB : ici, $T(s_2)$ n'est pas forcément borné (i.e. on n'est pas sûr que $T(s_2)(\omega) \leq K, \forall \omega \in \Omega$), mais le théorème d'arrêt s'applique malgré tout). Le processus (B_s) est donc une martingale p.r. à (\mathcal{F}_s) . Elle est également de carré intégrable et

$$\mathbb{E}(B_{s_2}^2 - B_{s_1}^2 | \mathcal{G}_{s_1}) = \mathbb{E}(M_{T(s_2)}^2 - M_{T(s_1)}^2 | \mathcal{F}_{T(s_1)}).$$

Par définition de la variation quadratique (et en utilisant à nouveau le théorème d'arrêt), on a

$$\mathbb{E}(M_{T(s_2)}^2 - \langle M \rangle_{T(s_2)} | \mathcal{F}_{T(s_1)}) = M_{T(s_1)}^2 - \langle M \rangle_{T(s_1)},$$

i.e.

$$\mathbb{E}(B_{s_2}^2 - B_{s_1}^2 | \mathcal{G}_{s_1}) = \mathbb{E}(\langle M \rangle_{T(s_2)} - \langle M \rangle_{T(s_1)} | \mathcal{F}_{T(s_1)}) = \mathbb{E}(s_2 - s_1 | \mathcal{F}_{T(s_1)}) = s_2 - s_1.$$

Donc le processus $(B_s^2 - s)$ est une martingale par rapport à (\mathcal{G}_s) , i.e. $\langle B \rangle_s = s$ p.s., et par le théorème de Lévy, (B_s) est un mouvement brownien standard par rapport à (\mathcal{G}_s) . \square

Corollaire 2.3.15. *Toute martingale continue de carré intégrable telle que $\lim_{t \rightarrow \infty} \langle M \rangle_t = \infty$ p.s. s'écrit $M_t = B(\langle M \rangle_t)$, où (B_t) est un mouvement brownien standard.*

On voit donc qu'une martingale se comporte généralement comme un mouvement brownien (en ce qui concerne l'allure des trajectoires). En ce sens, le mouvement brownien est "la" martingale typique.

Attention ! Du corollaire ci-dessus, on pourrait avoir envie de conclure que toute martingale est un processus gaussien. Ça n'est évidemment pas vrai, car le changement de temps $t \mapsto \langle M \rangle_t$ est lui-même aléatoire (en général), et donc la variable aléatoire $B(\langle M \rangle_t)$ n'est pas nécessairement gaussienne. Il est par contre vrai que si $(\langle M \rangle_t)$ est un processus déterministe, alors (M_t) est un processus gaussien.

2.4 Intégrale stochastique (ou intégrale d'Itô)

Soit $(B_t, t \in \mathbb{R}_+)$ un mouvement brownien standard par rapport à $(\mathcal{F}_t, t \in \mathbb{R}_+)$.

Soit $T > 0$ fixé (\neq temps d'arrêt). Notre but est de construire l'intégrale stochastique

$$\left(\int_0^t H_s dB_s, t \in [0, T] \right)$$

pour un processus (H_t) vérifiant certaines propriétés. Pour cela, on procède en plusieurs étapes (cf. construction de l'espérance).

Remarque 2.4.1. En mathématiques financières, T représente l'horizon du marché (désormais, on travaillera toujours à horizon fini, ce qui simplifie les choses), (B_t) représente l'évolution du prix d'un actif, (H_t) la stratégie d'investissement sur cet actif et $\int_0^t H_s dB_s$ la gain réalisé au temps t avec la stratégie H . De la même manière qu'à temps discret, (H_t) doit être "prévisible" pour qu'il n'y ait pas délit d'initié. Nous allons voir une manière simple de traduire cette propriété de prévisibilité à temps continu.

Première étape

Définition 2.4.2. *Un processus simple prévisible (par rapport à une filtration (\mathcal{F}_t)) est un processus $(H_t, t \in [0, T])$ tel que*

$$H_t = \sum_{i=1}^n X_i 1_{[t_{i-1}, t_i]}(t), \quad t \in [0, T],$$

où $0 = t_0 < t_1 < \dots < t_n = T$ forme une partition de $[0, T]$ et X_i est une v.a. $\mathcal{F}_{t_{i-1}}$ -mesurable et bornée $\forall i \in \{1, \dots, n\}$.

On voit donc que sur l'intervalle $]t_{i-1}, t_i]$, la valeur du processus (H_t) est déterminée par l'information $\mathcal{F}_{t_{i-1}}$, d'où le nom de “prévisible” (l'appellation “simple” vient quant à elle du fait que le processus ne prend qu'un nombre fini de valeurs (aléatoires!)). On pose

$$(H \cdot B)_T \equiv \int_0^T H_s dB_s = \sum_{i=1}^n X_i (B_{t_i} - B_{t_{i-1}}).$$

Cette définition de l'intégrale stochastique pour des processus simples prévisibles est unique (mais ça demande un petit travail de vérification) et l'intégrale est linéaire en H , i.e.

$$((cH + K) \cdot B)_T = c(H \cdot B)_T + (K \cdot B)_T.$$

Remarque 2.4.3. Cette intégrale représente en quelque sorte l’“aire sous la courbe $t \rightarrow H_t$ ”, lorsqu'on “mesure” les intervalles $]t_{i-1}, t_i]$ à l'aide du brownien (B_t) . Ici, il faut faire attention quand on parle de mesure car 1) la “mesure” de l'intervalle $B_{t_i} - B_{t_{i-1}}$ peut être négative; 2) cette “mesure” ne vérifie pas non plus l'axiome (ii) de la définition 1.1.9.

Proposition 2.4.4. *On a les égalités suivantes :*

$$\mathbb{E}((H \cdot B)_T) = 0 \quad \text{et} \quad \mathbb{E}((H \cdot B)_T^2) = \mathbb{E} \left(\int_0^T H_s^2 ds \right).$$

La seconde égalité ci-dessus est appelée l'isométrie d'Itô.

Remarquer que si $H(t) \equiv 1$, alors on retrouve l'égalité $\mathbb{E}(B_T^2) = T$.

Démonstration. En utilisant la linéarité de l'espérance et en introduisant un conditionnement à l'intérieur de l'espérance, on obtient

$$\begin{aligned} \mathbb{E}((H \cdot B)_T) &= \sum_{i=1}^n \mathbb{E}(X_i (B_{t_i} - B_{t_{i-1}})) \\ &= \sum_{i=1}^n \mathbb{E} \left(\mathbb{E}(X_i (B_{t_i} - B_{t_{i-1}}) | \mathcal{F}_{t_{i-1}}) \right) \\ &= \sum_{i=1}^n \mathbb{E} \left(X_i \mathbb{E}(B_{t_i} - B_{t_{i-1}} | \mathcal{F}_{t_{i-1}}) \right) \\ &= \sum_{i=1}^n \mathbb{E} \left(X_i \mathbb{E}(B_{t_i} - B_{t_{i-1}}) \right) = 0, \end{aligned}$$

où on a utilisé le fait que X_i est $\mathcal{F}_{t_{i-1}}$ -mesurable et $(B_{t_i} - B_{t_{i-1}}) \perp \mathcal{F}_{t_{i-1}}$. Pour l'isométrie, on calcule

$$\begin{aligned} \mathbb{E}((H \cdot B)_T^2) &= \sum_{i,j=1}^n \mathbb{E} \left(X_i X_j (B_{t_i} - B_{t_{i-1}}) (B_{t_j} - B_{t_{j-1}}) \right) \\ &= \sum_{i=1}^n \mathbb{E} \left(X_i^2 (B_{t_i} - B_{t_{i-1}})^2 \right) + 2 \sum_{\substack{i,j=1 \\ i < j}}^n \mathbb{E} \left(X_i X_j (B_{t_i} - B_{t_{i-1}}) (B_{t_j} - B_{t_{j-1}}) \right). \end{aligned}$$

En introduisant à nouveau un conditionnement à l'intérieur des espérances, on obtient

$$\begin{aligned}
\mathbb{E}((H \cdot B)_T^2) &= \sum_{i=1}^n \mathbb{E} \left(\mathbb{E}(X_i^2 (B_{t_i} - B_{t_{i-1}})^2 | \mathcal{F}_{t_{i-1}}) \right) \\
&\quad + 2 \sum_{\substack{i,j=1 \\ i < j}}^n \mathbb{E} \left(\mathbb{E}(X_i X_j (B_{t_i} - B_{t_{i-1}}) (B_{t_j} - B_{t_{j-1}}) | \mathcal{F}_{t_{j-1}}) \right) \\
&= \sum_{i=1}^n \mathbb{E} \left(X_i^2 \mathbb{E}((B_{t_i} - B_{t_{i-1}})^2) \right) + 2 \sum_{\substack{i,j=1 \\ i < j}}^n \mathbb{E} \left(X_i X_j (B_{t_i} - B_{t_{i-1}}) \mathbb{E}(B_{t_j} - B_{t_{j-1}}) \right) \\
&= \sum_{i=1}^n \mathbb{E}(X_i^2) (t_i - t_{i-1}) + 0,
\end{aligned}$$

car $\mathbb{E}((B_{t_i} - B_{t_{i-1}})^2) = t_i - t_{i-1}$. D'autre part, on a

$$\mathbb{E} \left(\int_0^t H_s^2 ds \right) = \mathbb{E} \left(\sum_{i=1}^n \int_{t_{i-1}}^{t_i} H_s^2 ds \right) = \mathbb{E} \left(\sum_{i=1}^n X_i^2 (t_i - t_{i-1}) \right),$$

donc l'isométrie est vérifiée. \square

Remarque 2.4.5. Pour être tout à fait exact, l'isométrie d'Itô dit encore que si H et K sont deux processus simples prévisibles, alors

$$\mathbb{E}((H \cdot B)_T (K \cdot B)_T) = \mathbb{E} \left(\int_0^T H_s K_s ds \right).$$

La vérification est similiaire à celle effectuée ci-dessus, quoiqu'un peu plus technique.

Deuxième étape

Soit $(H_t, t \in [0, T])$ un processus simple prévisible comme défini ci-dessus et $t \in [0, T]$. On pose

$$(H \cdot B)_t \equiv \int_0^t H_s dB_s = ((H 1_{[0,t]}) \cdot B)_T = \sum_{i=1}^n X_i (B_{t_i \wedge t} - B_{t_{i-1} \wedge t}).$$

A nouveau, cette intégrale est linéaire en H et on a par la proposition précédente :

$$\mathbb{E}((H \cdot B)_t) = 0 \quad \text{et} \quad \mathbb{E}((H \cdot B)_t^2) = \mathbb{E}(((H 1_{[0,t]}) \cdot B)_T^2) = \mathbb{E} \left(\int_0^T H_s^2 1_{[0,t]}(s) ds \right) = \mathbb{E} \left(\int_0^t H_s^2 ds \right).$$

De plus, en utilisant la remarque 2.4.5, on peut encore calculer

$$\text{Cov}((H \cdot B)_t, (K \cdot B)_s) = \mathbb{E}((H \cdot B)_t (K \cdot B)_s) = \mathbb{E} \left(\int_0^{t \wedge s} H_r K_r dr \right).$$

Remarque 2.4.6. Si $t \in]t_{k-1}, t_k]$, alors

Cours 8

$$(H \cdot B)_t = \sum_{i=1}^{k-1} X_i (B_{t_i} - B_{t_{i-1}}) + X_k (B_t - B_{t_{k-1}}).$$

Proposition 2.4.7. *Le processus $((H \cdot B)_t, t \in [0, T])$ est une martingale continue de carré intégrable.*

Démonstration. Par la remarque ci-dessus et la continuité de (B_t) , le processus $((H \cdot B)_t)$ est clairement continu à l'intérieur des intervalles $]t_{k-1}, t_k[$. Aux points limites, il est aisé de vérifier que le processus reste continu également. D'autre part, l'isométrie montrée plus haut dit que $((H \cdot B)_t)$ est un processus de carré intégrable, donc par l'inégalité de Cauchy-Schwarz,

$$\mathbb{E}(|(H \cdot B)_t|) \leq \sqrt{\mathbb{E}((H \cdot B)_t^2)} < \infty.$$

De plus, si on suppose que $t \in]t_{k-1}, t_k]$, alors

$$\begin{aligned} \mathbb{E}((H \cdot B)_T | \mathcal{F}_t) &= \sum_{i=1}^{k-1} \mathbb{E}(X_i (B_{t_i} - B_{t_{i-1}}) | \mathcal{F}_t) + \mathbb{E}(X_k (B_{t_k} - B_{t_{k-1}}) | \mathcal{F}_t) \\ &\quad + \sum_{i=k+1}^n \mathbb{E}(X_i (B_{t_i} - B_{t_{i-1}}) | \mathcal{F}_t) \\ &= \sum_{i=1}^{k-1} X_i (B_{t_i} - B_{t_{i-1}}) + X_k \mathbb{E}(B_{t_k} - B_{t_{k-1}} | \mathcal{F}_t) \\ &\quad + \sum_{i=k+1}^n \mathbb{E}(X_i \mathbb{E}(B_{t_i} - B_{t_{i-1}} | \mathcal{F}_{t_{i-1}}) | \mathcal{F}_t) \\ &= \sum_{i=1}^{k-1} X_i (B_{t_i} - B_{t_{i-1}}) + X_k (B_t - B_{t_{k-1}}) + 0 = (H \cdot B)_t, \end{aligned}$$

par la remarque ci-dessus. Le processus $((H \cdot B)_t)$ est donc une martingale car pour $t \geq s$, on a

$$\mathbb{E}((H \cdot B)_t | \mathcal{F}_s) = \mathbb{E}(\mathbb{E}((H \cdot B)_T | \mathcal{F}_t) | \mathcal{F}_s) = \mathbb{E}((H \cdot B)_T | \mathcal{F}_s) = (H \cdot B)_s.$$

□

D'après l'inégalité de Doob (b), on a donc :

$$\mathbb{E} \left(\sup_{t \in [0, T]} (H \cdot B)_t^2 \right) \leq 4 \mathbb{E} \left((H \cdot B)_T^2 \right) = 4 \mathbb{E} \left(\int_0^T H_s^2 ds \right).$$

Troisième étape

On étend maintenant l'intégrale $(H \cdot B)$ par continuité à l'ensemble

$$\mathcal{H}_T = \left\{ (H_t, t \in [0, T]) : H \text{ est adapté, continu à gauche, limité à droite} \right. \\ \left. \text{et tel que } \mathbb{E} \left(\int_0^T H_s^2 ds \right) < \infty \right\}.$$

Cet ensemble est un espace vectoriel normé et complet (ou espace de Banach), muni de la norme $\|\cdot\|_{T,1}$ définie par

$$\|H\|_{T,1}^2 = \mathbb{E} \left(\int_0^T H_s^2 ds \right).$$

Remarque 2.4.8. Un processus adapté et continu à gauche est l'équivalent à temps continu d'un processus prévisible à temps discret : la valeur d'un tel processus à l'instant t ne peut pas différer trop de sa valeur à l'instant $t - \varepsilon$, instant où il est $\mathcal{F}_{t-\varepsilon}$ -mesurable car adapté. Il existe toutefois une notion plus générale de processus prévisible à temps continu, mais nous n'entrerons pas dans ces détails dans ce cours.

D'autre part, on entend par "limité à droite" que le processus peut être discontinu à droite, mais doit tout de même avoir une limite depuis la droite (i.e. le "futur"). Cette restriction est essentiellement technique ; elle assure que le processus soit suffisamment régulier.

On a maintenant besoin du lemme suivant.

Lemme 2.4.9. $\forall H \in \mathcal{H}_T$, il existe une suite $(H^{(n)})$ de processus simples prévisibles tels que

$$\mathbb{E} \left(\int_0^T (H_s^{(n)} - H_s)^2 ds \right) \xrightarrow{n \rightarrow \infty} 0.$$

Conséquence : Du fait que la suite $(H^{(n)})$ converge, c'est également une suite de Cauchy :

$$\mathbb{E} \left(\int_0^T (H_s^{(n)} - H_s^{(m)})^2 ds \right) \xrightarrow{n,m \rightarrow \infty} 0.$$

Et donc, on a

$$\begin{aligned} \mathbb{E} \left(\sup_{t \in [0, T]} ((H^{(n)} \cdot B)_t - (H^{(m)} \cdot B)_t)^2 \right) &= \mathbb{E} \left(\sup_{t \in [0, T]} ((H^{(n)} - H^{(m)}) \cdot B)_t^2 \right) \\ &\leq 4 \mathbb{E} \left(\int_0^T (H_s^{(n)} - H_s^{(m)})^2 ds \right) \xrightarrow{n,m \rightarrow \infty} 0. \end{aligned}$$

La suite de processus $((H^{(n)} \cdot B))$ est donc une suite de Cauchy dans l'espace de Banach \mathcal{M}_T défini par

$$\mathcal{M}_T = \{(M_t, t \in [0, T]) \text{ martingale continue de carré intégrable telle que } M_0 = 0\}$$

et muni de la norme $\|\cdot\|_{T,2}$ définie par

$$\|M\|_{T,2}^2 = \mathbb{E} \left(\sup_{t \in [0, T]} M_t^2 \right).$$

Le fait que \mathcal{M}_T soit complet (i.e. que toute suite de Cauchy dans \mathcal{M}_T converge dans \mathcal{M}_T) implique que la suite $((H^{(n)} \cdot B))$ converge dans \mathcal{M}_T . Il existe donc un élément $(H \cdot B) \in \mathcal{M}_T$ tel que

$$\mathbb{E} \left(\sup_{t \in [0, T]} ((H^{(n)} \cdot B)_t - (H \cdot B)_t)^2 \right) \xrightarrow{n \rightarrow \infty} 0. \quad (3)$$

Nous avons ainsi défini l'application linéaire et continue

$$\begin{cases} \mathcal{H}_T \rightarrow \mathcal{M}_T, \\ H \mapsto (H \cdot B). \end{cases}$$

Remarque 2.4.10. - (3) implique en particulier que pour tout $\varepsilon > 0$,

$$\mathbb{P} \left(\sup_{t \in [0, T]} |(H^{(n)} \cdot B)_t - (H \cdot B)_t| > \varepsilon \right) \xrightarrow{n \rightarrow \infty} 0,$$

donc la suite $((H^{(n)} \cdot B))$ converge uniformément sur $[0, T]$ en probabilité vers $(H \cdot B)$.

- Attention : la construction de l'intégrale stochastique ci-dessus implique qu'elle n'est définie qu'à un ensemble négligeable près (cf. définition de l'espérance conditionnelle). Plus important : l'intégrale stochastique n'est pas définie trajectoire par trajectoire, i.e. " ω par ω " : pour un $\omega \in \Omega$ donné, il est impossible de dire ce que vaut $(H \cdot B)_T(\omega)$ si on ne connaît que les trajectoires $t \mapsto H_t(\omega)$ et $t \mapsto B_t(\omega)$; aussi étrange que cela puisse paraître, il faut connaître les processus (H_t) et (B_t) en entier !

Propriétés de l'intégrale stochastique

- Linéarité : $((cH + K) \cdot B)_t = c(H \cdot B)_t + (K \cdot B)_t$ p.s.
- Espérance nulle et isométrie : $\mathbb{E}((H \cdot B)_t) = 0$ et $\text{Cov}((H \cdot B)_t, (K \cdot B)_s) = \mathbb{E}(\int_0^{t \wedge s} H_r K_r dr)$.
- $(H \cdot B)$ est une martingale continue de carré intégrable telle que

$$\mathbb{E} \left(\sup_{t \in [0, T]} (H \cdot B)_t^2 \right) \leq 4 \mathbb{E} \left(\int_0^T H_s^2 ds \right).$$

Les propriétés ci-dessus ont déjà été vérifiées lors de la seconde étape pour des processus H simples prévisibles (on vérifie le cas général en utilisant la continuité). Citons encore une propriété qui n'a pas été mentionnée précédemment et qui est plus délicate à démontrer, mais très utile en pratique :

$$\langle (H \cdot B) \rangle_t = \int_0^t H_s^2 ds,$$

ou plus généralement,

$$\langle (H \cdot B), (K \cdot B) \rangle_t = \int_0^t H_s K_s ds.$$

Ainsi, on peut calculer facilement la variation quadratique d'une intégrale stochastique, ou la covariation quadratique de deux intégrales stochastiques, sans devoir se référer à la définition de (co)variation quadratique.

Remarque 2.4.11. Soit $(H_t, t \in [0, T]) \in \mathcal{H}_T$, qu'on suppose de plus à trajectoires continues. On pose

$$H_t^{(n)} = \sum_{i=1}^n H_{t_{i-1}^{(n)}} 1_{]t_{i-1}^{(n)}, t_i^{(n)}]}(t),$$

où $0 = t_0^{(n)} < t_1^{(n)} < \dots < t_n^{(n)} = t$ et $\lim_{n \rightarrow \infty} \max_{1 \leq i \leq n} |t_i^{(n)} - t_{i-1}^{(n)}| = 0$. Alors on a

$$(H^{(n)} \cdot B)_T = \sum_{i=1}^n H_{t_{i-1}^{(n)}} (B_{t_i^{(n)}} - B_{t_{i-1}^{(n)}}) \xrightarrow[n \rightarrow \infty]{\mathbb{P}} (H \cdot B)_T \equiv \int_0^T H_s dB_s.$$

Mais : 1) La convergence a lieu seulement en probabilité (on peut en fait montrer que si on exige que la convergence ait lieu presque sûrement pour tout processus continu H , alors le processus B doit être à variation bornée).

2) On n'a plus le choix du point $s_i^{(n)} \in [t_{i-1}^{(n)}, t_i^{(n)}]$, comme c'était le cas pour l'intégrale de Riemann-Stieltjes ; il faut maintenant choisir systématiquement le point à gauche de l'intervalle ($s_i^{(n)} = t_{i-1}^{(n)}$), sinon la limite change.

Illustration de ce dernier point : intégrale de Stratonovich

On pose

$$(H \circ B)_T \equiv \int_0^T H_s \circ dB_s = \lim_{n \rightarrow \infty} \sum_{i=1}^n \frac{1}{2} (H_{t_{i-1}^{(n)}} + H_{t_i^{(n)}}) (B_{t_i^{(n)}} - B_{t_{i-1}^{(n)}}),$$

où la limite est une limite en probabilité. On peut montrer que (cf. exercices) :

$$(H \circ B)_T = (H \cdot B)_T + \frac{1}{2} \langle H, B \rangle_T$$

où $(H \cdot B)_T$ est l'intégrale définie au sens d'Itô et

$$\langle H, B \rangle_T = \lim_{n \rightarrow \infty} \sum_{i=1}^n (H_{t_i^{(n)}} - H_{t_{i-1}^{(n)}}) (B_{t_i^{(n)}} - B_{t_{i-1}^{(n)}}),$$

où la limite est à nouveau une limite en probabilité (NB : si H et B sont des martingales continues de carré intégrable, alors la définition de $\langle H, B \rangle_T$ ci-dessus coïncide avec celle de covariation quadratique donnée plus haut).

Ainsi, si au lieu de considérer le point à gauche de l'intervalle, on choisit la moyenne entre le point à gauche et le point à droite, alors la valeur de l'intégrale change. De plus, le processus $((H \circ B)_t)$ n'est pas une martingale. Par contre, on verra que l'intégrale de Stratonovich présente des avantages du point de vue du calcul différentiel.

Intégrale stochastique par rapport à une martingale

On peut définir $\int_0^t H_s dM_s$ de la même manière que $\int_0^t H_s dB_s$ si on suppose que M est une *martingale continue de carré intégrable*.

1. Pour (H_t) un processus simple donné par $H_t = \sum_{i=1}^n X_i 1_{[t_{i-1}, t_i]}(t)$, on pose

$$(H \cdot M)_t \equiv \int_0^t H_s dM_s = \sum_{i=1}^n X_i (M_{t_i \wedge t} - M_{t_{i-1} \wedge t})$$

et on vérifie que

$$\mathbb{E}((H \cdot M)_t) = 0, \quad \mathbb{E}((H \cdot M)_t^2) = \mathbb{E} \left(\int_0^t H_s^2 d\langle M \rangle_s \right).$$

Remarquer que le terme ds apparaissant dans l'isométrie pour $(H \cdot B)_t$ a logiquement été remplacé par $d\langle M \rangle_s$ pour tenir compte de la variation quadratique $\langle M \rangle_s$ de la martingale qui n'est pas forcément égale à s .

D'autre part, on vérifie que le processus $((H \cdot M)_t, t \in [0, T])$ est également une martingale continue de carré intégrable.

2. L'extension de l'intégrale à un processus (H_t) adapté, continu à gauche, limité à droite et tel que

$$\mathbb{E} \left(\int_0^T H_s^2 d\langle M \rangle_s \right) < \infty$$

est identique à celle effectuée pour $\int_0^t H_s dB_s$.

Donnons ici la liste des **propriétés de** $(H \cdot M)_t \equiv \int_0^t H_s dM_s$:

Linéarité : $((cH + K) \cdot M)_t = c(H \cdot M)_t + (K \cdot M)_t$ et $(H \cdot (M + N))_t = (H \cdot M)_t + (H \cdot N)_t$.

$\mathbb{E}((H \cdot M)_t) = 0$, $\mathbb{E}((H \cdot M)_t^2) = \mathbb{E}(\int_0^t H_s^2 d\langle M \rangle_s)$.

$\text{Cov}((H \cdot M)_t, (K \cdot N)_t) = \mathbb{E}(\int_0^t H_s K_s d\langle M, N \rangle_s)$.

$\langle (H \cdot M)_t \rangle = \int_0^t H_s^2 d\langle M \rangle_s$.

$\langle (H \cdot M), (K \cdot N) \rangle_t = \int_0^t H_s K_s d\langle M, N \rangle_s$.

“Règle” utile à retenir : si $X_t = \int_0^t H_s dM_s$ (i.e. $dX_t = H_t dM_t$), alors $d\langle X \rangle_t = H_t^2 d\langle M \rangle_t$.

Noter que si $M_t = \int_0^t K_s dB_s$, alors

$$X_t = \int_0^t H_s K_s dB_s \quad \text{et} \quad \langle X \rangle_t = \int_0^t H_s^2 K_s^2 ds.$$

De même, vu que la propriété de martingale (continue de carré intégrable) est préservée, on peut encore définir $(L \cdot X)_t \equiv \int_0^t L_s dX_s$ etc. etc.

Remarque 2.4.12. On a maintenant défini une application linéaire et continue

$$\begin{cases} \tilde{\mathcal{H}}_T \times \mathcal{M}_T \rightarrow \mathcal{M}_T, \\ (H, M) \mapsto (H \cdot M), \end{cases}$$

où $\tilde{\mathcal{H}}_T$ est l'ensemble des processus (H_t) adaptés, continus à gauche, limités à droite et tels que $\mathbb{E}(\int_0^T H_s^2 d\langle M \rangle_s) < \infty$.

2.5.1 Formules de base

Jusqu'à présent, on a défini l'intégrale stochastique, mais il nous manque encore des règles de calcul. La première règle de calcul est la suivante.

Théorème 2.5.1. *Soit $(B_t, t \in \mathbb{R}^+)$ un mouvement brownien standard par rapport à $(\mathcal{F}_t, t \in \mathbb{R}^+)$ et $f \in C^2(\mathbb{R})$ (i.e. f, f' et f'' sont des fonctions continues). On suppose de plus que*

$$\mathbb{E} \left(\int_0^t (f'(B_s))^2 ds \right) < \infty, \quad \forall t > 0. \quad (4)$$

Alors pour tout $t > 0$,

$$f(B_t) - f(B_0) = \int_0^t f'(B_s) dB_s + \frac{1}{2} \int_0^t f''(B_s) ds \quad \text{p.s.} \quad (5)$$

Remarque 2.5.2. - Vu que $(f'(B_t), t \geq 0)$ est un processus continu et adapté à (\mathcal{F}_t) et que la condition (4) est vérifiée, $\int_0^t f'(B_s) dB_s$ est bien définie (et $\int_0^t f''(B_s) ds$ l'est également car l'application $s \mapsto f''(B_s)$ est continue).

- Le second terme du membre de droite dans (5) (absent dans les règles de calcul différentiel "classique") est appelé terme d'Itô; il résulte la variation quadratique non-nulle du mouvement brownien.

Démonstration.(idée principale)

$$f(B_t) - f(B_0) = \sum_{i=1}^n (f(B_{t_i^{(n)}}) - f(B_{t_{i-1}^{(n)}})),$$

où $0 = t_0^{(n)} < t_1^{(n)} < \dots < t_n^{(n)} = t$ est une suite de partitions de $[0, t]$ telle que

$$\lim_{n \rightarrow \infty} \max_{1 \leq i \leq n} |t_i^{(n)} - t_{i-1}^{(n)}| = 0.$$

Par un développement de Taylor classique, on a

$$f(y) - f(x) = f'(x)(y - x) + \frac{1}{2} f''(x)(y - x)^2 + r(y - x),$$

où $\lim_{h \rightarrow 0} \frac{r(h)}{h^2} = 0$ (on note aussi $r(h) = o(h^2)$). Donc

$$\begin{aligned} f(B_t) - f(B_0) &= \sum_{i=1}^n \left(f'(B_{t_{i-1}^{(n)}}) (B_{t_i^{(n)}} - B_{t_{i-1}^{(n)}}) + \frac{1}{2} f''(B_{t_{i-1}^{(n)}}) (B_{t_i^{(n)}} - B_{t_{i-1}^{(n)}})^2 + r_i^{(n)} \right) \\ &\xrightarrow[n \rightarrow \infty]{\mathbb{P}} \int_0^t f'(B_s) dB_s + \frac{1}{2} \int_0^t f''(B_s) d\langle B \rangle_s + 0, \end{aligned}$$

ce qui permet de conclure, vu que $\langle B \rangle_s = s$. □

Remarque 2.5.3. Cette “démonstration” évite un très grand nombre de détails techniques concernant la convergence des divers termes. Notamment, il faut choisir les points $B_{t_{i-1}^{(n)}}$ (et non $B_{t_i^{(n)}}$ ou autre chose) pour le développement de Taylor, ce qui n’apparaît pas clairement ci-dessus. Remarquer d’autre part que dans le cas où $f''(x) = 1$, on a déjà vu que

$$\sum_{i=1}^n 1 (B_{t_i^{(n)}} - B_{t_{i-1}^{(n)}})^2 \xrightarrow[n \rightarrow \infty]{\mathbb{P}} \langle B \rangle_t = \int_0^t 1 d\langle B \rangle_s.$$

Exemple 2.5.4. Soit $f(x) = x$; la formule (5) s’écrit alors

$$B_t - B_0 = \int_0^t 1 dB_s + 0,$$

autrement dit, rien de nouveau (ceci correspond au calcul différentiel classique).

Exemple 2.5.5. Soit $f(x) = x^2$; on a d’après (5) :

$$B_t^2 - B_0^2 = \int_0^t 2B_s dB_s + \frac{1}{2} \int_0^t 2 ds = 2 \int_0^t B_s dB_s + t,$$

autrement dit : $B_t^2 - t = 2 \int_0^t B_s dB_s$ est une martingale, ce qu’on avait déjà montré d’une autre manière (noter que la condition (4) est vérifiée car $\mathbb{E}(\int_0^t (2B_s)^2 ds) = \int_0^t 4s ds = 2t^2 < \infty$).

Exemple 2.5.6. Soit $f(x) = e^x$: $f'(x) = f''(x) = e^x$, et donc

$$e^{B_t} - 1 = \int_0^t e^{B_s} dB_s + \frac{1}{2} \int_0^t e^{B_s} ds.$$

Noter qu’ici aussi la condition (4) est vérifiée, car

$$\mathbb{E} \left(\int_0^t (e^{B_s})^2 ds \right) = \int_0^t \mathbb{E}(e^{2B_s}) ds = \int_0^t e^{2s} ds = \frac{e^{2t} - 1}{2} < \infty.$$

Si on pose $X_t = e^{B_t}$, alors la formule ci-dessus devient :

$$X_t - 1 = \int_0^t X_s dB_s + \frac{1}{2} \int_0^t X_s ds.$$

Autrement dit, on a trouvé une équation (intégrale) pour le processus (X_t) . Sous forme différentielle, celle-ci s’écrit

$$dX_t = X_t dB_t + \frac{1}{2} X_t dt \quad \text{et} \quad X_0 = 1.$$

Vu que $\frac{dB_t}{dt}$ n’existe pas, “on ne divise pas par dt ”. Noter que la forme différentielle ci-dessus n’a aucun sens en tant que telle et ne constitue qu’une notation pour la forme intégrale donnée

plus haut. L'esprit humain cependant a plus de facilité à raisonner avec des différentielles, c'est pourquoi celle-ci est largement répandue.

Sous une forme ou l'autre, l'équation ci-dessus est notre premier exemple d'équation différentielle stochastique. Il se trouve que c'est aussi un cas particulier de l'équation de Black & Scholes, largement utilisée en mathématiques financières pour décrire l'évolution du prix d'un actif (noter que la solution $X_t = e^{B_t}$ suit une loi log-normale et est toujours à valeurs positives).

Voyons maintenant une version légèrement plus élaborée de la formule (5).

Théorème 2.5.7. *Soient (B_t) un mouvement brownien standard et $f \in C^{1,2}(\mathbb{R}_+ \times \mathbb{R})$ (i.e $f, \frac{\partial f}{\partial t}, \frac{\partial f}{\partial x}, \frac{\partial^2 f}{\partial x^2}$ sont continues) telle que*

$$\mathbb{E} \left(\int_0^t \left(\frac{\partial f}{\partial x}(s, B_s) \right)^2 ds \right) < \infty, \quad \forall t > 0.$$

Alors pour tout $t > 0$,

$$f(t, B_t) - f(0, B_0) = \int_0^t \frac{\partial f}{\partial t}(s, B_s) ds + \int_0^t \frac{\partial f}{\partial x}(s, B_s) dB_s + \frac{1}{2} \int_0^t \frac{\partial^2 f}{\partial x^2}(s, B_s) ds \quad \text{p.s.}$$

Démonstration. La démonstration reprend le schéma de celle du théorème 2.5.1 esquissée ci-dessus. Le développement de Taylor utilisé est :

$$f(u, y) - f(t, x) = \frac{\partial f}{\partial t}(t, x)(u - t) + \frac{\partial f}{\partial x}(t, x)(y - x) + \frac{\partial^2 f}{\partial x^2}(t, x)(y - x)^2 + r(u - t, y - x),$$

ce qui explique l'apparition du terme supplémentaire dans la formule. \square

Remarque 2.5.8. Vu que la fonction $t \mapsto t$ est une fonction à variation bornée, il n'y a pas besoin d'effectuer un développement de la fonction f à l'ordre 2 en t .

Exemple 2.5.9. Soit $f(t, x) = x^2 - t$: $\frac{\partial f}{\partial t} = -1$, $\frac{\partial f}{\partial x} = 2x$, $\frac{\partial^2 f}{\partial x^2} = 2$, donc $\frac{\partial f}{\partial t} + \frac{1}{2} \frac{\partial^2 f}{\partial x^2} = 0$ et

$$B_t^2 - t = \int_0^t 2B_s dB_s.$$

On retrouve donc la même formule que ci-dessus.

Exemple 2.5.10. - Soit $f(t, x) = e^{x - \frac{t}{2}}$: $\frac{\partial f}{\partial t} = -\frac{1}{2}f$, $\frac{\partial f}{\partial x} = f$, $\frac{\partial^2 f}{\partial x^2} = f$, donc à nouveau $\frac{\partial f}{\partial t} + \frac{1}{2} \frac{\partial^2 f}{\partial x^2} = 0$ et

$$e^{B_t - \frac{t}{2}} - 1 = \int_0^t e^{B_s - \frac{s}{2}} dB_s.$$

Le processus $(Z_t = e^{B_t - \frac{t}{2}})$ est donc une martingale (appelée la martingale exponentielle associée au mouvement brownien standard). De plus, il satisfait l'équation différentielle stochastique :

$$Z_t - 1 = \int_0^t Z_s dB_s, \quad \text{i.e.} \quad dZ_t = Z_t dB_t \quad \text{et} \quad Z_0 = 1.$$

Remarque 2.5.11. On sait que le processus $X_t = e^{B_t}$ est une sous-martingale. En utilisant l'une ou l'autre des formules d'Itô ci-dessus, on a vu qu'il y a deux manières de l'écrire :

$$X_t = 1 + \int_0^t e^{B_s} dB_s + \frac{1}{2} \int_0^t e^{B_s} ds$$

ou

$$X_t = e^{t/2} + \int_0^t e^{\frac{t-s}{2}} e^{B_s} dB_s.$$

Or selon le théorème de décomposition de Doob, toute sous-martingale X_t continue s'écrit de manière unique comme $M_t + A_t$, où M_t est une martingale continue et A_t est un processus croissant continu adapté tel que $A_0 = 0$. L'exemple ci-dessus ne contredit-il donc pas le théorème ? Non, car le processus $\int_0^t e^{\frac{t-s}{2}} e^{B_s} dB_s$ n'est pas une martingale (à cause de la dépendance en t de l'intégrand).

2.5.2 Processus d'Itô (ou “semi-martingale continue”)

Définition 2.5.12. *Un processus d'Itô est un processus (X_t) pouvant se décomposer comme $(X_t = M_t + V_t)$, où :*

$$\begin{cases} (M_t) \text{ est une martingale continue de carré intégrable (p.r. à une filtration } (\mathcal{F}_t)), \\ (V_t) \text{ est un processus continu à variation bornée, adapté à } (\mathcal{F}_t) \text{ et tel que } V_0 = 0. \end{cases}$$

Remarque 2.5.13. Le nom “semi-martingale” vient simplement du fait que le processus (X_t) est composé pour moitié d'une martingale.

Exemple 2.5.14. D'après le théorème de décomposition de Doob, toute sous-martingale (resp. sur-martingale) continue de carré intégrable est un processus d'Itô (car un processus croissant (resp. décroissant) est à variation bornée).

Exemple 2.5.15. Soit (X_t) le processus défini par

$$X_t = X_0 + \int_0^t H_s dB_s + \int_0^t K_s ds,$$

où (H_t) est continu adapté et tel que $E(\int_0^t H_s^2 dB_s) < \infty$ pour tout $t \geq 0$ et (K_t) est continu et adapté. (X_t) un processus d'Itô.

Exemple 2.5.16. Soit $f \in C^2(\mathbb{R})$ vérifiant la condition (4). Alors $(f(B_t))$ est un processus d'Itô, car par la formule (5),

$$f(B_t) = f(B_0) + \int_0^t f'(B_s) dB_s + \frac{1}{2} \int_0^t f''(B_s) ds,$$

i.e. $(f(B_t))$ est la somme d'une martingale continue de carré intégrable ($M_t = f(B_0) + \int_0^t f'(B_s) dB_s$) et d'un processus continu à variation bornée ($V_t = \frac{1}{2} \int_0^t f''(B_s) ds$) tel que $V_0 = 0$.

Définition 2.5.17. Pour $t \geq 0$, la variation quadratique du processus d'Itô ($X_t = M_t + V_t$) est définie par

$$\langle X \rangle_t = \langle M \rangle_t,$$

et pour deux processus d'Itô ($X_t = M_t + V_t$) et ($Y_t = N_t + U_t$), on pose

$$\langle X, Y \rangle_t = \langle M, N \rangle_t.$$

Remarque 2.5.18. - Remarquer que si (X_t) est à variation bornée, alors $X_t = M_0 + V_t$ (rappelons ici que si (M_t) est une martingale continue à variation bornée, alors $M_t = M_0$ pour tout $t > 0$) et donc, $\langle X \rangle_t = 0$ selon la définition ci-dessus. De même, $\langle X, Y \rangle_t = 0$, quel que soit (Y_t) (ceci vient de l'inégalité de Cauchy-Schwarz : $|\langle X, Y \rangle_t| \leq \sqrt{\langle X \rangle_t \langle Y \rangle_t}$).

- Si d'autre part (X_t) et (Y_t) sont indépendants (mais pas forcément à variation bornée), alors $\langle X, Y \rangle_t = 0$.

Définition 2.5.19. Soit $(X_t = M_t + V_t)$ un processus d'Itô et (H_t) un processus continu, adapté à (\mathcal{F}_t) et tel que

$$\mathbb{E} \left(\int_0^t H_s^2 d\langle X \rangle_s \right) \equiv \mathbb{E} \left(\int_0^t H_s^2 d\langle M \rangle_s \right) < \infty.$$

On pose

$$(H \cdot X)_t \equiv \int_0^t H_s dX_s = \int_0^t H_s dM_s + \int_0^t H_s dV_s.$$

Remarquer qu'une intégrale stochastique par rapport à un processus d'Itô (X_t) est la somme d'une intégrale stochastique "pure" et d'une intégrale de Riemann-Stieltjes.

2.5.3 Retour à l'intégrale de Stratonovich

On peut maintenant définir l'intégrale de Stratonovich de la manière suivante. Etant donnés (H_t) et (X_t) deux processus d'Itô tels que

$$\mathbb{E} \left(\int_0^t H_s^2 d\langle X \rangle_s \right) < \infty,$$

on pose

$$\int_0^t H_s \circ dB_s \equiv (H \circ B)_t = (H \cdot B)_t + \frac{1}{2} \langle H, B \rangle_t.$$

On a la proposition suivante.

Proposition 2.5.20. *Soient (B_t) un mouvement brownien standard et $f \in C^3(\mathbb{R})$ telle que*

$$\mathbb{E} \left(\int_0^t (f'(B_s))^2 ds \right) < \infty \quad \text{et} \quad \mathbb{E} \left(\int_0^t (f''(B_s))^2 ds \right) < \infty.$$

Alors

$$f(B_t) - f(B_0) = \int_0^t f'(B_s) \circ dB_s \quad \text{p.s.}$$

Ceci montre bien le côté “agréable” du calcul de Stratonovich par rapport au calcul d’Itô ; l’absence de terme supplémentaire dans la formule ci-dessus permet en effet d’utiliser des règles de calcul classiques sans se poser de questions (faire attention toutefois qu’on a besoin d’hypothèses a priori plus fortes sur f). On a p.ex.

$$\int_0^t B_s \circ dB_s = \frac{B_t^2 - B_0^2}{2}.$$

Démonstration. Soit $g = f'$. Par la définition ci-dessus, on a :

$$\int_0^t g(B_s) \circ dB_s = \int_0^t g(B_s) dB_s + \frac{1}{2} \langle g(B), B \rangle_t.$$

Or par hypothèse $g \in C^2(\mathbb{R})$ et $\mathbb{E}(\int_0^t (g'(B_s))^2 ds) < \infty$, donc par la formule d’Itô,

$$g(B_t) = g(B_0) + \int_0^t g'(B_s) dB_s + \frac{1}{2} \int_0^t g''(B_s) ds = M_t + V_t,$$

donc

$$\langle g(B), B \rangle_t = \langle M, B \rangle_t = \int_0^t g'(B_s) 1 d\langle B, B \rangle_s = \int_0^t g'(B_s) ds.$$

Autrement dit,

$$\begin{aligned} \int_0^t g(B_s) \circ dB_s &= \int_0^t g(B_s) dB_s + \frac{1}{2} \int_0^t g'(B_s) ds \\ &= \int_0^t f'(B_s) dB_s + \frac{1}{2} \int_0^t f''(B_s) ds = f(B_t) - f(B_0), \end{aligned}$$

par l’application inverse de la formule d’Itô. Ceci conclut la démonstration. \square

2.5.4 Formules d'Itô généralisées

Les deux formules qui suivent sont des généralisations des théorèmes 2.5.1 et 2.5.7, respectivement.

Théorème 2.5.21. *Soit (M_t) une martingale continue de carré intégrable et $f \in C^2(\mathbb{R})$ telle que*

$$\mathbb{E} \left(\int_0^t (f'(M_s))^2 d\langle M \rangle_s \right) < \infty, \quad \forall t > 0.$$

Alors

$$f(M_t) - f(M_0) = \int_0^t f'(M_s) dM_s + \frac{1}{2} \int_0^t f''(M_s) d\langle M \rangle_s \quad \text{p.s.}$$

Remarquer que seul le terme ds de la formule (5) a été remplacé par le terme $d\langle M \rangle_s$ pour tenir compte de la variation quadratique de la martingale M .

Théorème 2.5.22. *Soient (M_t) martingale continue de carré intégrable, (V_t) un processus continu à variation bornée et $f \in C^{1,2}(\mathbb{R} \times \mathbb{R})$ telle que*

$$\mathbb{E} \left(\int_0^t \left(\frac{\partial f}{\partial x}(V_s, M_s) \right)^2 d\langle M \rangle_s \right) < \infty, \quad \forall t > 0.$$

Alors pour tout $t > 0$,

$$\begin{aligned} f(V_t, M_t) - f(V_0, M_0) &= \int_0^t \frac{\partial f}{\partial t}(V_s, M_s) dV_s + \int_0^t \frac{\partial f}{\partial x}(V_s, M_s) dM_s \\ &\quad + \frac{1}{2} \int_0^t \frac{\partial^2 f}{\partial x^2}(V_s, M_s) d\langle M \rangle_s \quad \text{p.s.} \end{aligned}$$

Remarquer que dans le cas particulier où $f(t, x) = g(t + x)$, la formule ci-dessus se récrit

$$\begin{aligned} g(V_t + M_t) - g(V_0 + M_0) &= \int_0^t g'(V_s + M_s) dV_s + \int_0^t g'(V_s + M_s) dM_s \\ &\quad + \frac{1}{2} \int_0^t g''(V_s + M_s) d\langle M \rangle_s \quad \text{p.s.} \end{aligned}$$

ou encore, en posant $X_t = M_t + V_t$:

$$g(X_t) - g(X_0) = \int_0^t g'(X_s) dX_s + \frac{1}{2} \int_0^t g''(X_s) d\langle X \rangle_s \quad \text{p.s.}$$

Il ne faut cependant pas oublier ici que la première intégrale est la somme de deux intégrales de natures différentes !

2.5.5 Formule d'intégration par parties (IPP)

Soient (X_t) , (Y_t) deux processus d'Itô. Alors pour tout $t \geq 0$, on a

$$X_t Y_t - X_0 Y_0 = \int_0^t X_s dY_s + \int_0^t Y_s dX_s + \langle X, Y \rangle_t \quad \text{p.s.,}$$

qu'on écrit encore sous forme différentielle

$$d(X_t Y_t) = X_t dY_t + Y_t dX_t + d\langle X, Y \rangle_t.$$

Démonstration. En utilisant la formule d'Itô généralisée ci-dessus, on trouve

$$\begin{aligned} (X_t + Y_t)^2 - (X_0 + Y_0)^2 &= 2 \int_0^t (X_s + Y_s) d(X_s + Y_s) + \langle X + Y \rangle_t, \\ (X_t - Y_t)^2 - (X_0 - Y_0)^2 &= 2 \int_0^t (X_s - Y_s) d(X_s - Y_s) + \langle X - Y \rangle_t, \end{aligned}$$

En soustrayant les deux égalités ci-dessus, on obtient

$$4(X_t Y_t - X_0 Y_0) = 4 \left(\int_0^t X_s dY_s + \int_0^t Y_s dX_s \right) + \langle X + Y \rangle_t - \langle X - Y \rangle_t.$$

Combiné avec la définition de covariation quadratique, ceci donne finalement le résultat. \square

Remarque 2.5.23. Si (X_t) ou (Y_t) est à variation bornée, alors par la remarque 2.5.18, $\langle X, Y \rangle_t = 0$ et on retrouve la formule d'intégration par parties classique.

Exemple 2.5.24. Soient $X_t = e^{B_t}$ et $Y_t = e^{-\frac{t}{2}}$. On a vu que $dX_t = \frac{1}{2}X_t dt + X_t dB_t$; d'autre part, on a $dY_t = -\frac{1}{2}Y_t dt$. Posons maintenant $Z_t = X_t Y_t = e^{B_t - \frac{t}{2}}$. Par la formule d'intégration par parties, on obtient :

$$Z_t - 1 = \int_0^t X_s dY_s + \int_0^t Y_s dX_s + 0,$$

car $\langle X, Y \rangle_t = 0$ ((Y_t) est à variation bornée). On a donc

$$Z_t - 1 = \int_0^t X_s \left(-\frac{1}{2}Y_s\right) ds + \int_0^t Y_s \left(\frac{1}{2}X_s\right) ds + \int_0^t Y_s X_s dB_s = \int_0^t Z_s dB_s,$$

i.e. $dZ_t = Z_t dB_t$ et $Z_0 = 1$, qui est l'équation qu'on avait trouvée plus haut.

Exemple 2.5.25. Soient $X_t = \int_0^t H_s dB_s$ et $Y_t = \int_0^t K_s dB_s$. On a

$$X_t Y_t = \int_0^t (H_s Y_s + K_s X_s) dB_s + \int_0^t H_s K_s ds.$$

2.6 Equations différentielles stochastiques (EDS)

Cours 10

2.6.1 Equations homogènes en temps

On a vu que $X_t = e^{B_t}$ (aussi appelé “mouvement brownien géométrique”) est solution de l’équation

$$X_t - 1 = \frac{1}{2} \int_0^t X_s ds + \int_0^t X_s dB_s,$$

qu’on écrit encore sous forme différentielle :

$$dX_t = \frac{1}{2} X_t dt + X_t dB_t \quad \text{et} \quad X_0 = 1.$$

De la même façon, on peut voir que $X_t = e^{(\mu - \frac{\sigma^2}{2})t + \sigma B_t}$ (avec $\mu \in \mathbb{R}$ et $\sigma > 0$ fixés) est solution de

$$dX_t = \mu X_t dt + \sigma X_t dB_t \quad \text{et} \quad X_0 = 1. \quad (6)$$

Cette équation est appelée l’équation de Black & Scholes ; μ est appelé le coefficient de dérive (il traduit la tendance générale du processus) et σ le coefficient de diffusion (il traduit la variabilité ou encore la “volatilité” du processus). Cette équation, ou des généralisations de celle-ci, sont couramment utilisées en mathématiques financières pour décrire l’évolution des prix des actifs.

De manière plus générale, on considère le problème suivant. Etant donnés (B_t) un mouvement brownien standard, $x_0 \in \mathbb{R}$ et $f, g : \mathbb{R} \rightarrow \mathbb{R}$, existe-t-il un processus (X_t) qui vérifie

$$dX_t = f(X_t) dt + g(X_t) dB_t \quad \text{et} \quad X_0 = x_0?$$

(dans l’exemple précédent, $f(x) = \mu x$ et $g(x) = \sigma x$). Pour répondre à cette question, on a besoin de la définition suivante.

Définition 2.6.1. Une fonction $f : \mathbb{R} \rightarrow \mathbb{R}$ est dite (globalement) lipschitzienne s’il existe $K \geq 0$ tel que

$$|f(y) - f(x)| \leq K|y - x|, \quad \forall x, y \in \mathbb{R}.$$

Remarque 2.6.2. - Si f est lipschitzienne, alors f est uniformément continue (et donc continue) sur \mathbb{R} .

- Si f est continûment dérivable et f' est bornée, alors f est lipschitzienne. En effet :

$$|f(y) - f(x)| = \left| \int_x^y f'(z) dz \right| \leq \sup_{z \in \mathbb{R}} |f'(z)| \cdot |y - x|.$$

Théorème 2.6.3. Soient $(B_t, t \in \mathbb{R}_+)$ un mouvement brownien standard (p.r. à une filtration $(\mathcal{F}_t, t \in \mathbb{R}_+)$), $x_0 \in \mathbb{R}$ et f, g lipschitziennes. Alors il existe un unique processus $(X_t, t \in \mathbb{R}_+)$ continu et adapté à $(\mathcal{F}_t, t \in \mathbb{R}_+)$ tel que

$$X_t = x_0 + \int_0^t f(X_s) ds + \int_0^t g(X_s) dB_s \quad \text{p.s.,} \quad \forall t \in \mathbb{R}_+. \quad (7)$$

De plus, $\mathbb{E}(\sup_{0 \leq t \leq T} X_t^2) < \infty$ pour tout $T > 0$.

Remarque 2.6.4. - La solution (X_t) de l'équation ci-dessus est également appelée une solution forte de l'équation (par opposition au concept de solution faible présenté au paragraphe 2.6.4).

- On appelle $f(X_t)$ le terme de dérive de l'équation et $g(X_t)$ le terme de diffusion.

Démonstration.(idée principale)

On définit

$$\mathcal{X}_T = \left\{ (X_t, t \in [0, T]) \text{ continu et adapté à } (\mathcal{F}_t) \text{ tel que } \mathbb{E}(\sup_{0 \leq t \leq T} X_t^2) < \infty \right\}.$$

Noter que l'espace vectoriel \mathcal{X}_T muni de la norme $\|X\|_{T,2}^2 = \mathbb{E}(\sup_{0 \leq t \leq T} X_t^2)$ est complet. Pour trouver le processus $X \in \mathcal{X}_T$ solution de (7), on utilise la méthode classique dite "méthode d'itération de Picard", i.e. on définit une suite de processus $(X^{(n)})$ de manière récursive :

$$X_t^{(0)} = x_0, \quad X_t^{(n+1)} = x_0 + \int_0^t f(X_s^{(n)}) ds + \int_0^t g(X_s^{(n)}) dB_s.$$

Il se trouve que la suite $(X^{(n)})$ est une suite de Cauchy dans \mathcal{X}_T , donc elle converge car \mathcal{X}_T est complet, et on montre que la limite de la suite est solution de (7). De plus, on montre que si (X_t) et (Y_t) sont des solutions de (7), alors $X_t = Y_t$ p.s. pour tout $t \in \mathbb{R}_+$.

Pour chaque étape, on a recours à l'estimation centrale suivante (qui se démontre en utilisant notamment l'inégalité de Doob (b) et l'isométrie d'Itô) ; si on pose

$$\phi(Y)_t = x_0 + \int_0^t f(Y_s) ds + \int_0^t g(Y_s) dB_s,$$

alors

$$\mathbb{E} \left(\sup_{0 \leq t \leq T} (\phi(Y)_t - \phi(Z)_t)^2 \right) \leq K \mathbb{E} \left(\int_0^T (Y_s - Z_s)^2 ds \right).$$

□

Remarque 2.6.5. Il est important de voir que le schéma d'itération de Picard ci-dessus, même s'il est explicite, se rapproche très lentement de la solution (X_t) . Il est donc fortement déconseillé de l'appliquer pour trouver une approximation numérique de la solution (pour un schéma numérique efficace, se référer à la section 2.10).

Proposition 2.6.6. *Le processus (X_t) solution de l'équation (7) est un processus d'Itô.*

Démonstration. On peut décomposer la solution comme $X_t = M_t + V_t$, où

$$M_t = x_0 + \int_0^t g(X_s) dB_s \quad \text{et} \quad V_t = \int_0^t f(X_s) ds.$$

Du fait que $X \in \mathcal{X}_T$ et que g est lipschitzienne, l'intégrale stochastique $\int_0^t g(X_s) dB_s$ est bien définie et (M_t) est une martingale continue de carré intégrable. D'autre part, du fait que $t \mapsto (X_t)$ et f sont des fonctions continues, le processus (V_t) est continûment dérivable, donc continu et à variation bornée. Le processus (X_t) est donc un processus d'Itô. \square

Exemple 2.6.7. L'équation de Black & Scholes

$$dX_t = \mu X_t dt + \sigma X_t dB_t \quad \text{et} \quad X_0 = x_0,$$

avec $\mu \in \mathbb{R}$ et $\sigma, x_0 > 0$ fixés, est un exemple d'équation admettant une unique solution (X_t) , car $f(x) = \mu x$ et $g(x) = \sigma x$ sont linéaires, donc lipschitziennes (pour la méthode de résolution, voir exercices).

Exemple 2.6.8. Soit $a, x_0 \in \mathbb{R}$ et $\sigma > 0$ fixés. On considère l'EDS

$$dX_t = -aX_t dt + \sigma dB_t \quad \text{et} \quad X_0 = x_0. \tag{8}$$

Les fonctions $f(x) = -ax$ et $g(x) \equiv \sigma$ sont lipschitziennes, donc il existe un unique processus (X_t) solution de l'équation ci-dessus. Ce processus est appelé le processus d'Ornstein-Uhlenbeck.

Proposition 2.6.9. *La solution de l'EDS (8) est donnée par*

$$X_t = e^{-at} x_0 + \int_0^t e^{-a(t-s)} dB_s.$$

Démonstration. Plutôt que de vérifier que le processus ci-dessus est magiquement solution de (8) (noter que ceci a déjà été fait aux exercices!), on présente ici la méthode de résolution de l'équation, qui est une adaptation de la méthode dite de "variation de la constante" au cas stochastique. Soit (Φ_t) le processus (déterministe) solution de l'équation différentielle ordinaire (et homogène) :

$$d\Phi_t = -a\Phi_t dt \quad \text{et} \quad \Phi_0 = 1.$$

Il est connu que $\Phi_t = e^{-at}$. Ecrivons maintenant $X_t = \Phi_t Y_t$. On a alors par la formule d'intégration par parties :

$$dX_t = \Phi_t dY_t + (d\Phi_t) Y_t + d\langle \Phi, Y \rangle_t = \Phi_t dY_t - a\Phi_t Y_t dt + 0$$

car le processus (Φ_t) est à variation bornée. D'un autre côté, on a

$$dX_t = -aX_t dt + \sigma dB_t.$$

En identifiant les deux équations (et en se rappelant que $X_t = \Phi_t Y_t$), on trouve l'EDS suivante pour (Y_t) :

$$\Phi_t dY_t = \sigma dB_t \quad \text{i.e.} \quad dY_t = \frac{\sigma}{\Phi_t} dB_t \quad (\text{et } Y_0 = x_0),$$

donc

$$Y_t = Y_0 + \int_0^t \frac{\sigma}{\Phi_s} dB_s = x_0 + \sigma \int_0^t e^{as} dB_s.$$

Ceci implique finalement que

$$X_t = e^{-at} x_0 + \sigma \int_0^t e^{-a(t-s)} dB_s.$$

□

Remarquer encore que si $x_0 = 0$ et $a > 0$, alors

$$\mathbb{E}(X_t) = 0 \quad \text{et} \quad \mathbb{E}(X_t^2) = \sigma^2 \frac{1 - e^{-2at}}{2a} \xrightarrow{t \rightarrow \infty} \frac{\sigma^2}{2a},$$

i.e. le processus d'Ornstein-Uhlenbeck (avec $a > 0$) est un processus qui ne grandit pas indéfiniment (comme c'est le cas pour le mouvement brownien), mais se stabilise autour de la valeur 0 avec une variance donnée $(\frac{\sigma^2}{2a})$.

Exemple 2.6.10. Considérons l'EDS

$$dX_t = \sin(X_t) dt + \cos(X_t) dB_t \quad \text{et} \quad X_0 = 0.$$

Ici, $f(x) = \sin(x)$ et $g(x) = \cos(x)$ sont lipschitziennes (car f' et g' sont bornées), donc il existe une unique solution à l'équation, mais quelle est-elle ? Il se trouve que même si la formulation de l'équation est assez simple, il n'existe pas de méthode de résolution analytique ! On voit donc l'intérêt de simuler numériquement la solution d'une telle équation. Noter que d'autre part, il existe des méthodes analytiques pour estimer le comportement asymptotique de la solution.

Exemple 2.6.11. Soit $a > 0$. On considère l'EDS

$$dX_t = \frac{a}{X_t} dt + dB_t \quad \text{et} \quad X_0 = 1.$$

La fonction $f(x) = \frac{a}{x}$ n'est pas lipschitzienne en 0 ($f'(0) = -\infty$). Toutefois, il existe une unique solution (forte) à l'équation (appelée processus de Bessel). La raison intuitive pour laquelle ceci est vérifié est que sitôt que le processus X_t se rapproche de 0 (l'endroit où $f(x)$ n'est pas lipschitzienne), il est fortement repoussé vers le haut par le terme de dérive $\frac{a}{X_t}$ (noter toutefois que si a est petit, des choses étranges commencent à se produire).

Exemple 2.6.12. Soit $a \in \mathbb{R}$. On considère l'EDS

$$dX_t = aX_t dt + \sqrt{X_t} dB_t \quad \text{et} \quad X_0 = 0.$$

La fonction $g(x) = \sqrt{x}$ n'est pas lipschitzienne en 0 ($g'(0) = \infty$) et il n'existe pas de solution forte, mais une solution dite "faible" (voir paragraphe 2.6.4), appelée processus de Feller.

Exemple 2.6.13. Considérons l'EDS

$$dX_t = \text{sgn}(X_t) dB_t \quad \text{et} \quad X_0 = 0.$$

La fonction

$$g(x) = \text{sgn}(x) = \begin{cases} +1 & \text{si } x > 0, \\ 0 & \text{si } x = 0, \\ -1 & \text{si } x < 0, \end{cases}$$

est discontinue en 0 et donc n'est pas lipschitzienne. L'équation n'admet pas de solution forte, mais seulement une solution faible (à nouveau, voir paragraphe 2.6.4).

2.6.2 Equations inhomogènes en temps

On considère le problème suivant. Etant donnés (B_t) un mouvement brownien standard, $x_0 \in \mathbb{R}$ et $f, g : \mathbb{R}_+ \times \mathbb{R} \rightarrow \mathbb{R}$, existe-t-il un processus (X_t) qui vérifie

$$dX_t = f(t, X_t) dt + g(t, X_t) dB_t \quad \text{et} \quad X_0 = x_0?$$

Pour répondre à cette question, on a besoin de la définition suivante.

Définition 2.6.14. Une fonction $f : \mathbb{R}_+ \times \mathbb{R} \rightarrow \mathbb{R}$ est dite lipschitzienne en x s'il existe une constante $K > 0$ telle que

$$|f(t, y) - f(t, x)| \leq K|y - x|, \quad \forall t \in \mathbb{R}_+, x, y \in \mathbb{R}.$$

Théorème 2.6.15. Si $f, g : \mathbb{R}_+ \times \mathbb{R} \rightarrow \mathbb{R}$ sont continues (conjointement en t et en x) et lipschitziennes en x , alors il existe un unique processus (X_t) solution de l'équation

$$X_t = x_0 + \int_0^t f(s, X_s) ds + \int_0^t g(s, X_s) dB_s \quad \text{p.s.,} \quad \forall t \in \mathbb{R}_+.$$

A nouveau, le processus (X_t) est appelé une solution forte de l'équation (et c'est un processus d'Itô).

2.6.3 Equations linéaires

Soient $x_0 \in \mathbb{R}$ et $a, \sigma : \mathbb{R}_+ \rightarrow \mathbb{R}$ continues et bornées. On appelle linéaire une EDS de la forme

$$dX_t = a(t) X_t dt + \sigma(t) dB_t \quad \text{et} \quad X_0 = x_0. \quad (9)$$

Remarque 2.6.16. On pourrait avoir envie plutôt d'appeler "équation linéaire" une équation du type

$$dX_t = a(t) X_t dt + \sigma(t) X_t dB_t \quad \text{et} \quad X_0 = x_0,$$

qui constitue une généralisation de l'équation de Black & Scholes (6). Il se trouve que le fait de "multiplier" X_t et dB_t dans le terme de droite implique que la solution de l'équation n'est pas un processus gaussien et qu'on préfère appeler linéaire une équation dont la solution est un processus gaussien (mais ceci est arbitraire).

Définition 2.6.17. Soit $(\Phi_t, t \in \mathbb{R}_+)$ la solution de l'équation différentielle ordinaire

$$d\Phi_t = a(t) \Phi_t dt \quad \text{et} \quad \Phi_0 = 1.$$

Remarquer que $\Phi_t = \exp \left(\int_0^t a(s) ds \right)$.

Proposition 2.6.18. L'équation (9) admet une unique solution forte (X_t) donnée par

$$X_t = \Phi_t x_0 + \int_0^t \frac{\Phi_t}{\Phi_s} \sigma(s) dB_s \quad t \in \mathbb{R}_+.$$

Démonstration. Il est clair que $f(t, x) = a(t)x$ et $g(t, x) = \sigma(t)$ sont continues en (t, x) et lipschitziennes en x , donc l'équation (9) admet une unique solution. La méthode de résolution est du même type que celle utilisée pour l'équation (8). On écrit tout d'abord $X_t = \Phi_t Y_t$, d'où on déduit que

$$dX_t = \Phi_t dY_t + a(t) \Phi_t Y_t dt + 0 = a(t) X_t dt + \sigma(t) dB_t.$$

De là, on tire que

$$\Phi_t dY_t = \sigma(t) dB_t, \quad \text{i.e.} \quad Y_t = x_0 + \int_0^t \frac{\sigma(s)}{\Phi_s} dB_s,$$

et donc

$$X_t = \Phi_t x_0 + \int_0^t \frac{\Phi_t}{\Phi_s} \sigma(s) dB_s.$$

□

Exemple 2.6.19. (pont brownien ; cf. exercices) Considérons l'EDS

$$dX_t = -\frac{X_t}{1-t} dt + dB_t, \quad 0 < t < 1, \quad \text{et} \quad X_0 = 0.$$

Ici, la fonction $a(t) = -\frac{1}{1-t}$ n'est pas continue en $t = 1$, mais l'équation ci-dessus admet tout de même une unique solution forte jusqu'en $t = 1$. Noter que $X_1 = 0$ p.s. et que $\mathbb{E}(X_t^2) = t(1-t)$.

2.6.4 Solution faible

Soient $x_0 \in \mathbb{R}$ et $f, g : \mathbb{R} \rightarrow \mathbb{R}$. On considère l'EDS

$$dX_t = f(X_t) dt + g(X_t) dB_t \quad \text{et} \quad X_0 = x_0. \quad (10)$$

On donne la définition suivante, qui peut paraître étrange au premier abord.

Définition 2.6.20. Une solution faible de l'équation (10) est un processus continu (X_t) tel que les processus (M_t) et (N_t) définis respectivement par

$$M_t = X_t - X_0 - \int_0^t f(X_s) ds \quad \text{et} \quad N_t = M_t^2 - \int_0^t g(X_s)^2 ds$$

sont des martingales.

Remarque 2.6.21. Le mouvement brownien standard (B_t) a disparu de la définition de solution faible ! Ainsi, une solution faible d'une EDS est une solution en loi, mais plus du tout une solution "trajectorielle" de l'équation (10).

La justification de cette définition est donnée ci-dessous.

Proposition 2.6.22. Supposons f, g continues, g bornée et supposons encore que l'équation (10) admette une solution forte (X_t) . Alors (X_t) est une solution faible de (10).

Démonstration. Vu que g est bornée, il est clair que

$$M_t = X_t - X_0 - \int_0^t f(X_s) ds = \int_0^t g(X_s) dB_s$$

est une martingale continue de carré intégrable. De plus, la variation quadratique de (M_t) est donnée par $\langle M \rangle_t = \int_0^t g(X_s)^2 ds$, donc le processus (N_t) défini par

$$N_t = M_t^2 - \int_0^t g(X_s)^2 ds \quad (11)$$

est également une martingale. □

Remarque 2.6.23. - Une EDS peut admettre une solution faible, mais pas de solution forte ; il existe donc plus souvent une solution faible qu’une solution forte.

- La question de l’unicité de la solution faible est par contre plus délicate ; il faut préciser ce qu’on entend par “unique”.

Exemple 2.6.24. La solution faible de l’EDS

$$dX_t = aX_t dt + \sqrt{X_t} dB_t \quad \text{et} \quad X_0 = 0$$

est un processus continu (X_t) tel que les processus (M_t) et (N_t) définis par

$$M_t = X_t - a \int_0^t X_s ds \quad \text{et} \quad N_t = M_t^2 - \int_0^t X_s ds$$

sont des martingales.

Exemple 2.6.25. La solution faible de l’EDS

$$dX_t = \text{sgn}(X_t) dB_t \quad \text{et} \quad X_0 = 0.$$

est un processus continu (X_t) tel que

$$(X_t - 0 = X_t) \quad \text{et} \quad (X_t^2 - \int_0^t 1 ds = X_t^2 - t) \quad \text{sont des martingales.}$$

Par le théorème de Lévy, la solution faible de l’équation ci-dessus est donc un mouvement brownien standard ! (mais qui n’est pas le mouvement brownien standard (B_t) ; on ne peut pas remplacer X_t par B_t dans l’équation ci-dessus...)

2.6.5 Martingale exponentielle

Cours 11

Définition 2.6.26. Soit M une martingale continue de carré intégrable telle que $M_0 = 0$ et $\langle M \rangle_t \leq Kt$ pour tout $t \in \mathbb{R}_+$. On définit la martingale exponentielle Y associée à M par

$$Y_t = \exp \left(M_t - \frac{\langle M \rangle_t}{2} \right), \quad t \in \mathbb{R}_+.$$

Remarque 2.6.27. Noter que la condition $\langle M \rangle_t \leq Kt$ n’impose pas forcément que la variation quadratique $(\langle M \rangle_t)$ soit un processus déterministe.

Proposition 2.6.28. Le processus (Y_t) défini ci-dessus satisfait l’EDS

$$Y_t = 1 + \int_0^t Y_s dM_s \quad (\text{i.e.} \quad dY_t = Y_t dM_t \quad \text{et} \quad Y_0 = 1),$$

et (Y_t) est donc une martingale.

Démonstration. (idée principale)

Remarquer que

$$Y_t = f(\langle M \rangle_t, M_t), \quad \text{où} \quad f(t, x) = \exp(x - \frac{t}{2})$$

et $\frac{\partial f}{\partial t} = -\frac{1}{2}f$, $\frac{\partial f}{\partial x} = \frac{\partial^2 f}{\partial x^2} = f$. En appliquant le théorème 2.5.22 (et en passant à dessein sous silence la condition d'intégrabilité énoncée dans le théorème en question!), on trouve donc :

$$Y_t - Y_0 = \int_0^t (-\frac{1}{2}Y_s) d\langle M \rangle_s + \int_0^t Y_s dM_s + \frac{1}{2} \int_0^t Y_s d\langle M \rangle_s = \int_0^t Y_s dM_s.$$

□

Remarque 2.6.29. La condition $\langle M \rangle_t \leq Kt$, même si elle n'est pas utilisée explicitement ci-dessus, a toute son importance (elle permet de justifier l'utilisation de la formule d'Itô). Noter qu'il est possible de montrer que le processus (Y_t) est une martingale sous une condition plus faible encore.

2.7 Théorème de Girsanov

Comme vu précédemment, la solution de l'équation

$$dX_t = f(t, X_t) dt + g(t, X_t) dB_t \quad \text{et} \quad X_0 = x_0,$$

n'est en général pas une martingale (sous la probabilité \mathbb{P}). La question que l'on se pose ici est de savoir s'il existe une autre mesure de probabilité $\tilde{\mathbb{P}}$ sous laquelle le processus (X_t) soit une martingale.

L'application en mathématiques financières est la suivante : pour évaluer le prix d'une option sur un actif, on a besoin de la propriété de martingale. Pourtant, le prix d'un actif donné n'est en général pas une martingale, mais affiche une tendance à la hausse ou à la baisse. C'est pourquoi on désire définir une nouvelle probabilité sous laquelle celui-ci soit une martingale, de manière à pouvoir effectuer des calculs.

Première étape : définition du changement de probabilité

Soit (M_t) une martingale par rapport à (\mathcal{F}_t) , continue, de carré intégrable telle que $M_0 = 0$ et $\langle M \rangle_t \leq Kt$ pour tout $t \in \mathbb{R}_+$. On pose (Y_t) la martingale exponentielle associée à (M_t) (voir paragraphe 2.6.5). On rappelle ici que $Y_t > 0$ pour tout $t \in \mathbb{R}_+$.

On se place à horizon fini $T > 0$ (ce qui simplifie considérablement les choses) et on définit

$$\tilde{\mathbb{P}}_T(A) = \mathbb{E}(1_A Y_T), \quad A \in \mathcal{F}.$$

Le triplet $(\Omega, \mathcal{F}, \tilde{\mathbb{P}}_T)$ forme un nouvel espace de probabilité. En effet :

$$\begin{cases} \tilde{\mathbb{P}}_T(A) \geq 0, & \forall A \in \mathcal{F}, \quad \text{car } Y_T > 0, \\ \tilde{\mathbb{P}}_T(\Omega) = \mathbb{E}(Y_T) = \mathbb{E}(Y_0) = 1, & \text{car } (Y_t) \text{ est une martingale,} \\ \text{et } \tilde{\mathbb{P}}_T(\cup_{n=1}^{\infty} A_n) = \mathbb{E}(\sum_{n=1}^{\infty} 1_{A_n} Y_T) = \sum_{n=1}^{\infty} \tilde{\mathbb{P}}_T(A_n), \end{cases}$$

si $(A_n)_{n=1}^{\infty}$ est une famille disjointe d'événements dans \mathcal{F} .

On montre d'autre part que si X est une v.a. telle que $\mathbb{E}(|X Y_T|) < \infty$, alors

$$\tilde{\mathbb{E}}_T(X) = \mathbb{E}(X Y_T).$$

Finalement, noter que $\mathbb{P}(A) = 0$ si et seulement si $\tilde{\mathbb{P}}_T(A) = 0$ (on dit que les deux mesures de probabilité \mathbb{P} et $\tilde{\mathbb{P}}_T$ sont "équivalentes").

Deuxième étape : martingales sous \mathbb{P} et martingales sous $\tilde{\mathbb{P}}_T$

Lemme 2.7.1. *Si Z est \mathcal{F}_t -mesurable et telle que $\mathbb{E}(|Z Y_T|) < \infty$, alors $\tilde{\mathbb{E}}_T(Z) = \mathbb{E}(Z Y_T)$.*

Démonstration. Du fait que (Y_t) est une martingale et que Z est \mathcal{F}_t -mesurable, on a

$$\tilde{\mathbb{E}}_T(Z) = \mathbb{E}(Z Y_T) = \mathbb{E}(Z Y_t).$$

□

Lemme 2.7.2. *Le processus (X_t) est une martingale sous $\tilde{\mathbb{P}}_T$ ssi $(X_t Y_t)$ est une martingale sous \mathbb{P} .*

Démonstration. On montre seulement que si $(X_t Y_t)$ est une martingale sous \mathbb{P} , alors (X_t) est une martingale sous $\tilde{\mathbb{P}}_T$ (pour la réciproque, voir exercices). Supposons donc que $(X_t Y_t)$ est une martingale sous \mathbb{P} , i.e.

$$\begin{cases} \text{(i)} & \forall t \in [0, T], \quad X_t Y_t \text{ est } \mathcal{F}_t\text{-mesurable et } \mathbb{E}|X_t Y_t| < \infty, \\ \text{(ii)} & \forall 0 \leq s \leq t \leq T, \quad \mathbb{E}(X_t Y_t Z) = \mathbb{E}(X_s Y_s Z), \quad \forall Z \text{ } \mathcal{F}_s\text{-mesurable et bornée.} \end{cases}$$

De là, on déduit que

(i) Pour tout $t \in [0, T]$, $X_t = \frac{X_t Y_t}{Y_t}$ est \mathcal{F}_t -mesurable et $\tilde{\mathbb{E}}_T(|X_t|) = \mathbb{E}(|X_t| Y_T) = \mathbb{E}(|X_t| Y_t) < \infty$.

(ii) Soit $0 \leq s \leq t \leq T$ et Z \mathcal{F}_s -mesurable et bornée. Par le lemme 2.7.1 et la condition (ii) ci-dessus, on a

$$\tilde{\mathbb{E}}_T(X_t Z) = \mathbb{E}(X_t Z Y_t) = \mathbb{E}(X_s Z Y_s) = \tilde{\mathbb{E}}_T(X_s Z)$$

i.e. $\tilde{\mathbb{E}}_T(X_t | \mathcal{F}_s) = X_s$, i.e. (X_t) est une martingale sous $\tilde{\mathbb{P}}_T$.

□

Théorème 2.7.3. (*Girsanov*)

Soit (Z_t) est une martingale continue de carré intégrable sous \mathbb{P} . Alors $(Z_t - \langle M, Z \rangle_t)$ est une martingale sous $\tilde{\mathbb{P}}_T$.

Démonstration. (idée principale)

Posons $A_t = \langle M, Z \rangle_t$. Pour montrer que $(Z_t - A_t)$ est une martingale sous $\tilde{\mathbb{P}}_T$, il suffit de montrer par le lemme 2.7.2 que $((Z_t - A_t) Y_t)$ est une martingale sous \mathbb{P} . Par la formule d'intégration par parties, on a :

$$(Z_t - A_t) Y_t - (Z_0 - 0) Y_0 = \int_0^t (Z_s - A_s) dY_s + \int_0^t Y_s dZ_s - \int_0^t Y_s dA_s + \langle Z - A, Y \rangle_t$$

Du fait que Y et Z sont des martingales sous \mathbb{P} , les deux premiers termes $\int_0^t (Z_s - A_s) dY_s$ et $\int_0^t Y_s dZ_s$ sont également des martingales sous \mathbb{P} . On aura donc montré le résultat si on montre que les deux derniers termes s'annulent, i.e.

$$- \int_0^t Y_s dA_s + \langle Z - A, Y \rangle_t = 0. \quad (12)$$

Or $dY_t = Y_t dM_t$, donc $dM_t = \frac{1}{Y_t} dY_t$ (noter que $Y_t > 0$) et

$$dA_t = d\langle M, Z \rangle_t = \frac{1}{Y_t} d\langle Y, Z \rangle_t.$$

Ceci implique que

$$\int_0^t Y_s dA_s = \int_0^t d\langle Y, Z \rangle_s = \langle Y, Z \rangle_t.$$

De l'autre côté, on sait que $\langle Z - A, Y \rangle_t = \langle Z, Y \rangle_t$, car A est à variation bornée. L'équation (12) est donc bien vérifiée et le théorème est démontré. \square

Troisième étape : application aux EDS

A) Soient (B_t) un mouvement brownien standard, $x_0 \in \mathbb{R}$ et $f : \mathbb{R}_+ \times \mathbb{R} \rightarrow \mathbb{R}$ une fonction continue, lipschitzienne en x et bornée (i.e. $|f(t, x)| \leq K_1$ pour tous t, x). On considère (X_t) la solution de l'équation

$$dX_t = f(t, X_t) dt + dB_t \quad \text{et} \quad X_0 = x_0.$$

Sous quelle mesure de probabilité $\tilde{\mathbb{P}}_T$ le processus $(X_t, t \in [0, T])$ est-il une martingale ?

Soit $M_t = - \int_0^t f(s, X_s) dB_s$; (M_t) est une martingale continue de carré intégrable (sous \mathbb{P}) et on vérifie que

$$\langle M \rangle_t = \int_0^t f(s, X_s)^2 ds \leq K_1^2 t$$

Soit (Y_t) la martingale exponentielle associée à (M_t) et $\tilde{\mathbb{P}}_T$ la probabilité définie par $\tilde{\mathbb{P}}_T(A) = \mathbb{E}(1_A Y_T)$.

Proposition 2.7.4. (i) (X_t) est une martingale sous $\tilde{\mathbb{P}}_T$.
(ii) (X_t) est même un mouvement brownien standard sous $\tilde{\mathbb{P}}_T$!

Démonstration. (i) Du fait que (B_t) est une martingale sous \mathbb{P} , le théorème de Girsanov dit que $(B_t - \langle M, B \rangle_t)$ est une martingale sous $\tilde{\mathbb{P}}_T$. Or

$$B_t - \langle M, B \rangle_t = B_t + \int_0^t f(s, X_s) ds = X_t - X_0,$$

donc le premier point est vérifié.

(ii) D'autre part, on a par définition :

$$\langle X \rangle_t = \langle B \rangle_t = t,$$

car X est un processus d'Itô dont la partie martingale est égale à B . Or on peut montrer que la variation quadratique de X est la même sous \mathbb{P} et sous $\tilde{\mathbb{P}}_T$, donc par le théorème de Lévy, X est un mouvement brownien standard sous $\tilde{\mathbb{P}}_T$ (remarquer que X est à trajectoires continues). \square

B) Soient (B_t) un mouvement brownien standard, $x_0 \in \mathbb{R}$ et $f, g : \mathbb{R}_+ \times \mathbb{R} \rightarrow \mathbb{R}$ deux fonctions continues, lipschitziennes en x et telles que $|f(t, x)| \leq K_1$ et $|g(t, x)| \geq K_2$ pour tous t, x . On considère (X_t) la solution de l'équation

$$dX_t = f(t, X_t) dt + g(t, X_t) dB_t \quad \text{et} \quad X_0 = x_0.$$

Remarque 2.7.5. L'hypothèse $|g(t, x)| \geq K_2$ ci-dessus assure que le processus X est "non-dégénéré", autrement dit qu'il a l'allure d'un mouvement brownien et pas celle d'une fonction à variation bornée (ce qui serait le cas si $g(t, x) = 0$ p.ex.). Sans cette hypothèse, il n'est plus possible de transformer le processus X en martingale en changeant la probabilité de référence.

Soit $M_t = - \int_0^t \frac{f(s, X_s)}{g(s, X_s)} dB_s$. On vérifie que $\langle M \rangle_t \leq \frac{K_1^2}{K_2^2} t$ et on pose (Y_t) la martingale exponentielle associée à (M_t) et $\tilde{\mathbb{P}}_t$ la probabilité définie par $\tilde{\mathbb{P}}_T(A) = \mathbb{E}(1_A Y_T)$.

Proposition 2.7.6. (i) (X_t) est une martingale sous $\tilde{\mathbb{P}}_T$ et (ii) $\langle X \rangle_t = \int_0^t g(s, X_s)^2 ds$.

Démonstration. (i) Soit $C_t = \int_0^t g(s, X_s) dB_s$; (C_t) est une martingale sous \mathbb{P} , donc

$$C_t - \langle M, C \rangle_t = \int_0^t g(s, X_s) dB_s + \int_0^t f(s, X_s) ds = X_t - X_0$$

est une martingale sous $\tilde{\mathbb{P}}_T$.

(ii) A nouveau, l'égalité découle de la définition de variation quadratique. \square

En général, le processus (X_t) n'est donc pas un mouvement brownien standard sous $\tilde{\mathbb{P}}_T$. Mais peut-on exhiber un processus qui soit un mouvement brownien standard sous $\tilde{\mathbb{P}}_T$? C'est le cas en effet : soit $\tilde{B}_t = B_t + \int_0^t \frac{f(s, X_s)}{g(s, X_s)} ds$.

Proposition 2.7.7. (i) (\tilde{B}_t) est une martingale sous $\tilde{\mathbb{P}}_t$
(ii) $\langle \tilde{B} \rangle_t = t$, donc (\tilde{B}_t) est un mouvement brownien standard sous $\tilde{\mathbb{P}}_T$.

Démonstration. (i) Du fait que (B_t) est une martingale sous \mathbb{P} ,

$$B_t - \langle M, B \rangle_t = B_t + \int_0^t \frac{f(s, X_s)}{g(s, X_s)} ds = \tilde{B}_t$$

est une martingale sous $\tilde{\mathbb{P}}_t$.

(ii) clair (on utilise à nouveau le théorème de Lévy). \square

Remarque importante : de la définition de (\tilde{B}_t) , il découle que

$$dX_t = g(t, X_t) d\tilde{B}_t.$$

Ceci montre d'une autre manière que le processus (X_t) est une martingale sous $\tilde{\mathbb{P}}_T$.

Exemple 2.7.8. (modèle de Black & Scholes)

Soient $\mu \in \mathbb{R}$, $\sigma, x_0 > 0$ fixés. On considère l'EDS

$$dX_t = \mu X_t dt + \sigma X_t dB_t \quad \text{et} \quad X_0 = x_0.$$

La fonction $g(x) = \sigma x$ ne satisfait pas l'hypothèse de non-dégénérescence mentionnée plus haut (car $g(0) = 0$), cependant, si on définit

$$M_t = - \int_0^t \frac{\mu X_s}{\sigma X_s} dB_s = -\frac{\mu}{\sigma} B_t,$$

alors on voit que (M_t) est une martingale sous \mathbb{P} et que $\langle M \rangle_t = \frac{\mu^2}{\sigma^2} t$. Soit (Y_t) la martingale associée à (M_t) et $\tilde{\mathbb{P}}_T(A) = \mathbb{E}(1_A Y_T)$.

Par le théorème de Girsanov, le processus (X_t) est une martingale sous $\tilde{\mathbb{P}}_T$. D'autre part, le processus (\tilde{B}_t) défini par

$$\tilde{B}_t = B_t + \frac{\mu}{\sigma} t$$

est un mouvement brownien standard sous $\tilde{\mathbb{P}}_T$ et $dX_t = \sigma X_t d\tilde{B}_t$.

Application : formule de Black & Scholes

Définition 2.7.9. Une option d'achat européenne est le droit d'acheter un actif (X_t) à un temps T futur au prix K fixé à l'avance.

Au temps T (aussi appelé la “maturité” de l’option), la valeur de l’option est donc donnée par

$$Z_T = (X_T - K)^+.$$

Supposons maintenant que l’actif (X_t) soit décrit par le modèle de Black & Scholes de l’exemple 2.7.8 (avec $\mu \in \mathbb{R}$, $\sigma > 0$ fixés) :

$$dX_t = \mu X_t dt + \sigma X_t dB_t \quad \text{et} \quad X_0 = x_0 > 0$$

et soit (\mathcal{F}_t) la filtration à laquelle X est adapté. Sous cette hypothèse, on peut montrer qu’il existe $z_0 \in \mathbb{R}_+$ et une stratégie d’investissement H sur le processus X (continue et adaptée à (\mathcal{F}_t)) tels que

$$Z_T = z_0 + \int_0^T H_s dX_s.$$

(Ceci est une conséquence du “théorème de représentation des martingales”).

Question : A quel prix une banque doit-elle vendre l’option d’achat Z_T au temps initial $t = 0$ de manière à être sûre de pouvoir payer Z_T au temps T ?

Réponse : Etant donné ce qui précède, elle doit vendre l’option au prix z_0 . Mais comment calculer ce prix ?

Calcul de z_0 :

a) On a vu que (X_t) est une martingale sous la probabilité $\tilde{\mathbb{P}}_T$ définie dans l’exemple 2.7.8. Donc le processus (M_t) défini par

$$M_t = \int_0^t H_s dX_s \quad t \in [0, T],$$

est également une martingale sous $\tilde{\mathbb{P}}_T$, d’où on tire que

$$z_0 = \tilde{\mathbb{E}}_T(Z_T) = \tilde{\mathbb{E}}_T((X_T - K)^+)$$

b) On sait d’autre part que (cf. exemple 2.7.8)

$$dX_t = X_t d\tilde{B}_t \quad \text{i.e.} \quad X_T = x_0 e^{\sigma \tilde{B}_T - \frac{\sigma^2}{2} T},$$

ce qui nous mène à la formule de Black & Scholes :

$$z_0 = \tilde{\mathbb{E}}_T((x_0 e^{\sigma \tilde{B}_T - \frac{\sigma^2}{2} T} - K)^+) = \int_{\mathbb{R}} (x_0 e^{\sigma y - \frac{\sigma^2}{2} T} - K)^+ f_T(y) dy,$$

où f_T est la densité de la loi $\mathcal{N}(0, T)$. Remarquer que de manière surprenante, le coefficient de dérive μ n’entre pas en compte dans le calcul du prix z_0 !

Le but de cette section est de montrer qu'on peut représenter les solutions d'équations aux dérivées partielles (EDP) classiques à l'aide de processus aléatoires.

Attention ! Les théorèmes d'analyse cités ci-dessous sont légèrement inexacts et donc parsemés de guillemets : voir la remarque 2.8.1, qui s'applique également aux résultats cités après.

2.8.1 Lien mouvement brownien \leftrightarrow équation(s) de la chaleur

A) Soit $(B_t, t \in \mathbb{R}_+)$ un mouvement brownien standard. On définit $(B_t^x, t \in \mathbb{R}_+)$, le mouvement brownien partant au temps $t = 0$ du point $x \in \mathbb{R}$ par

$$B_t^x = B_t + x, \quad t \in \mathbb{R}_+.$$

Résultat d'analyse (équation de la chaleur progressive)

Soient $T > 0$ et $u_0 \in C(\mathbb{R})$. Il existe alors une "unique" fonction $u \in C^{1,2}(\mathbb{R}_+ \times \mathbb{R})$ qui vérifie

$$\begin{cases} \frac{\partial u}{\partial t}(t, x) = \frac{1}{2} \frac{\partial^2 u}{\partial x^2}(t, x), & \forall (t, x) \in \mathbb{R}_+ \times \mathbb{R}, \\ "u(0, x) = u_0(x)", & \forall x \in \mathbb{R}. \end{cases} \quad (13)$$

Remarque 2.8.1. - La solution de l'équation de la chaleur (13) n'est en réalité pas unique. Toutefois, si on impose une (très) faible condition supplémentaire sur la solution u , alors u est unique.

- Si u_0 n'est pas C^2 , alors la solution en $t = 0$ ne peut être C^2 en x . Pour être tout à fait exact, on devrait remplacer la condition $u(0, x) = u_0(x)$ par

$$\lim_{t \downarrow 0} u(t, x) = u_0(x).$$

Lemme 2.8.2. Soit u la solution de l'équation (13). Pour tout $(T, x) \in \mathbb{R}_+ \times \mathbb{R}$ fixé, le processus $(u(T - t, B_t^x), t \in [0, T])$ est une martingale.

Démonstration. Par la formule d'Itô (et le fait que $dB_s^x = dB_s$, $d\langle B^x \rangle_s = ds$), on a

$$\begin{aligned} & u(T - t, B_t^x) - u(T, x) \\ &= \int_0^t -\frac{\partial u}{\partial t}(T - s, B_s^x) ds + \int_0^t \frac{\partial u}{\partial x}(T - s, B_s^x) dB_s^x + \frac{1}{2} \int_0^t \frac{\partial^2 u}{\partial x^2}(T - s, B_s^x) d\langle B^x \rangle_s \\ &= \int_0^t \left(-\frac{\partial u}{\partial t}(T - s, B_s^x) + \frac{1}{2} \frac{\partial^2 u}{\partial x^2}(T - s, B_s^x) \right) ds + \int_0^t \frac{\partial u}{\partial x}(T - s, B_s^x) dB_s \\ &= \int_0^t \frac{\partial u}{\partial x}(T - s, B_s^x) dB_s, \end{aligned}$$

car u satisfait l'équation (13) (on oublie ici de vérifier la condition technique permettant d'appliquer la formule d'Itô!). Le processus $(u(T-t, B_t^x), t \in [0, T])$ est “donc” une martingale. \square

Proposition 2.8.3. *Soit u la solution de l'équation (13). Pour tout $(T, x) \in \mathbb{R}_+ \times \mathbb{R}$, on a*

$$u(T, x) = \mathbb{E}(u_0(B_T^x)). \quad (14)$$

Démonstration. Par le lemme précédent et la condition initiale dans (13), on trouve que

$$u(T, x) = \mathbb{E}(u(T, B_0^x)) = \mathbb{E}(u(0, B_T^x)) = \mathbb{E}(u_0(B_T^x)).$$

\square

Remarque 2.8.4. - Le résultat ci-dessus nous donne donc une représentation probabiliste de la solution de l'équation de la chaleur (13). Noter qu'on peut le reformuler de manière plus “classique” :

$$u(T, x) = \int_{\mathbb{R}} u_0(y) f_T(y-x) dy,$$

où f_T est la densité de la loi $\mathcal{N}(0, T)$. La fonction $K_T(x, y) = f_T(y-x)$ est également appelée “noyau de Green de l'équation de la chaleur” en analyse.

- Remarquer qu'il est possible de montrer dans l'autre sens (mais c'est beaucoup plus long) que la fonction u définie par (14) est solution de l'équation (13).

B) Soit $(B_t, t \in \mathbb{R}_+)$ un mouvement brownien standard. On définit $(B_t^{t_0, x_0}, t \in [t_0, \infty[)$ le mouvement brownien partant au temps $t_0 \geq 0$ du point $x_0 \in \mathbb{R}$ par

$$B_t^{t_0, x_0} = B_t - B_{t_0} + x_0, \quad t \geq t_0.$$

Résultat d'analyse (équation de la chaleur rétrograde)

Soient $T > 0$ et $h \in C(\mathbb{R})$. Il existe alors une “unique” fonction $u \in C^{1,2}([0, T] \times \mathbb{R})$ qui vérifie

$$\begin{cases} \frac{\partial u}{\partial t}(t, x) + \frac{1}{2} \frac{\partial^2 u}{\partial x^2}(t, x) = 0, & \forall (t, x) \in [0, T] \times \mathbb{R}, \\ “u(T, x) = h(x)”, & \forall x \in \mathbb{R}. \end{cases} \quad (15)$$

Lemme 2.8.5. *Soit u la solution de l'équation (15). Pour tout $(t_0, x_0) \in [0, T] \times \mathbb{R}$ fixé, le processus $(u(t, B_t^{t_0, x_0}), t \in [t_0, T])$ est une martingale.*

Démonstration. Par la formule d'Itô (utilisée sur l'intervalle de temps $[t_0, t]$), on a

$$\begin{aligned} u(t, B_t^{t_0, x_0}) - u(t_0, x_0) &= \int_{t_0}^t \frac{\partial u}{\partial t}(s, B_s^{t_0, x_0}) ds + \int_{t_0}^t \frac{\partial u}{\partial x}(s, B_s^{t_0, x_0}) dB_s^{t_0, x_0} + \frac{1}{2} \int_{t_0}^t \frac{\partial^2 u}{\partial x^2}(s, B_s^{t_0, x_0}) ds \\ &= \int_{t_0}^t \frac{\partial u}{\partial x}(s, B_s^{t_0, x_0}) dB_s, \end{aligned}$$

car u satisfait l'équation (15). Le processus $(u(t, B_t^{t_0, x_0}), t \in [t_0, T])$ est “donc” une martingale. \square

Proposition 2.8.6. *Soit u la solution de l'équation (15). Pour tout $(t_0, x_0) \in [0, T] \times \mathbb{R}$, on a*

$$u(t_0, x_0) = \mathbb{E}(h(B_T^{t_0, x_0})).$$

Démonstration. Par le lemme précédent et la condition terminale dans (15), on trouve que

$$u(t_0, x_0) = \mathbb{E}(u(t_0, B_{t_0}^{t_0, x_0})) = \mathbb{E}(u(T, B_T^{t_0, x_0})) = \mathbb{E}(h(B_T^{t_0, x_0})).$$

\square

2.8.2 Lien EDS \leftrightarrow EDP paraboliques (formule de Feynman-Kac)

Soient $t_0 \in \mathbb{R}_+$, $x_0 \in \mathbb{R}$, $(B_t, t \in \mathbb{R}_+)$ un mouvement brownien standard et $f, g : \mathbb{R}_+ \times \mathbb{R} \rightarrow \mathbb{R}$ deux fonctions conjointement continues en (t, x) et lipschitziennes en x . On pose $(X_t^{t_0, x_0}, t \in [t_0, \infty[)$ la solution de l'EDS

$$dX_t = f(t, X_t) dt + g(t, X_t) dB_t \quad \text{et} \quad X_{t_0} = x_0. \quad (16)$$

On suppose de plus qu'il existe une constante $K > 0$ telle que

$$|g(t, x)| \geq K, \quad \forall (t, x) \in \mathbb{R}_+ \times \mathbb{R}.$$

Cette dernière hypothèse est l'hypothèse de “diffusion non-dégénérée” effectuée à la section 2.7.

Résultat d'analyse (EDP parabolique)

Soient $h \in C(\mathbb{R})$ et $T > 0$. Etant donné les hypothèses effectuées ci-dessus sur f et g , il existe une “unique” fonction $u \in C^{1,2}([0, T] \times \mathbb{R})$ qui vérifie

$$\begin{cases} \frac{\partial u}{\partial t}(t, x) + f(t, x) \frac{\partial u}{\partial x}(t, x) + \frac{1}{2} g^2(t, x) \frac{\partial^2 u}{\partial x^2}(t, x) = 0, & \forall (t, x) \in [0, T] \times \mathbb{R}, \\ “u(T, x) = h(x)”, & \forall x \in \mathbb{R}. \end{cases} \quad (17)$$

Lemme 2.8.7. *Soit u la solution de l'équation (17). Pour tout $(t_0, x_0) \in [0, T] \times \mathbb{R}$ fixé, le processus $(u(t, X_t^{t_0, x_0}), t \in [t_0, T])$ est une martingale.*

Démonstration. Par la formule d'Itô (utilisée sur l'intervalle de temps $[t_0, t]$), on a

$$\begin{aligned} & u(t, X_t^{t_0, x_0}) - u(t_0, x_0) \\ &= \int_{t_0}^t \frac{\partial u}{\partial t}(s, X_s^{t_0, x_0}) ds + \int_{t_0}^t \frac{\partial u}{\partial x}(s, X_s^{t_0, x_0}) dX_s^{t_0, x_0} + \frac{1}{2} \int_{t_0}^t \frac{\partial^2 u}{\partial x^2}(s, X_s^{t_0, x_0}) d\langle X^{t_0, x_0} \rangle. \end{aligned}$$

En utilisant le fait que X est solution de (16), on trouve donc que

$$\begin{aligned}
& u(t, X_t^{t_0, x_0}) - u(t_0, x_0) \\
&= \int_0^t \left(\frac{\partial u}{\partial t}(s, X_s^{t_0, x_0}) + \frac{\partial u}{\partial x}(s, X_s^{t_0, x_0}) f(s, X_s^{t_0, x_0}) + \frac{1}{2} \frac{\partial^2 u}{\partial x^2}(s, X_s^{t_0, x_0}) g(s, X_s^{t_0, x_0})^2 \right) ds \\
&\quad + \int_0^t \frac{\partial u}{\partial x}(s, X_s^{t_0, x_0}) g(s, X_s^{t_0, x_0}) dB_s \\
&= \int_0^t \frac{\partial u}{\partial x}(s, X_s^{t_0, x_0}) g(s, X_s^{t_0, x_0}) dB_s,
\end{aligned}$$

car u satisfait l'équation (17). Le processus $(u(t, X_t^{t_0, x_0}), t \in [t_0, T])$ est “donc” une martingale. \square

Proposition 2.8.8. (*formule de Feynman-Kac*)

Soit u la solution de l'équation (17). Pour tout $(t_0, x_0) \in [0, T] \times \mathbb{R}$, on a

$$u(t_0, x_0) = \mathbb{E}(h(X_T^{t_0, x_0})).$$

Démonstration. Par le lemme précédent et la condition terminale dans (17), on trouve que

$$u(t_0, x_0) = \mathbb{E}(u(t_0, X_{t_0}^{t_0, x_0})) = \mathbb{E}(u(T, X_T^{t_0, x_0})) = \mathbb{E}(h(X_T^{t_0, x_0}))$$

\square

Remarque 2.8.9. La formule donnée ci-dessus est l'une des nombreuses versions de la formule dite de Feynman-Kac (prononcer “Kats”). Une autre version de cette formule est donnée en exercice, qui correspond plus à la formule de Feynman-Kac connue des physiciens.

Propriété de Markov et processus de diffusion

On donne ici une nouvelle version de la propriété de Markov énoncée au paragraphe 2.1.4.

A) Soit tout d'abord $(B_t^{t_0, x_0})$ un mouvement brownien partant du point $x_0 \in \mathbb{R}$ au temps $t_0 \in \mathbb{R}_+$. On vérifie que si $t \geq s \geq 0$, alors

$$B_t^{s, B_s^{0, x}} = B_t^{0, x}, \quad (18)$$

En effet, on a par définition :

$$B_t^{s, B_s^{0, x}} = B_t - B_s + B_s^{0, x} = B_t - B_s + B_s - B_0 + x = B_t - B_0 + x = B_t^{0, x}.$$

La propriété (18) entraîne que

$$\mathbb{E}(h(B_t^{0, x}) | \mathcal{F}_s^{B^{0, x}}) = \mathbb{E}(h(B_t^{0, x}) | B_s^{0, x}), \quad \forall h \in C_b(\mathbb{R}).$$

Noter que la propriété (18) traduit bien la même idée que la propriété de Markov ci-dessus : le mouvement brownien parti du point x au temps 0 est le même que celui qu'on sait être passé au point $B_s^{0,x}$ au temps s . Ainsi, l'évolution du mouvement brownien après le temps s ne dépend pas de l'histoire du processus avant s (donc en particulier, pas de la valeur x), mais seulement de la valeur du processus au temps s .

B) Soit maintenant $(X_t^{t_0, x_0})$ la solution de l'EDS (16). On peut vérifier que pour tout $t \geq s \geq 0$, on a également

$$X_t^{s, X_s^{0,x}} = X_t^{0,x},$$

ce qui implique que

$$\mathbb{E}(h(X_t^{0,x}) | \mathcal{F}_s^{X^{0,x}}) = \mathbb{E}(h(X_t^{0,x}) | X_s^{0,x}), \quad \forall h \in C_b(\mathbb{R}).$$

La solution d'une EDS est donc un processus de Markov. On l'appelle également un processus de diffusion ou plus simplement une diffusion.

Générateur d'une diffusion

A) Diffusion homogène en temps

Soient $x \in \mathbb{R}$, $(B_t, t \in \mathbb{R})$ un mouvement brownien standard et $f, g : \mathbb{R} \rightarrow \mathbb{R}$ lipschitziennes. On pose X la solution de l'EDS

$$dX_t = f(X_t) dt + g(X_t) dB_t \quad \text{et} \quad X_0 = x.$$

Soit A l'opérateur linéaire différentiel de $C^2(\mathbb{R})$ dans $C(\mathbb{R})$ défini par

$$Av(x) = f(x) v'(x) + \frac{1}{2} g(x)^2 v''(x), \quad v \in C^2(\mathbb{R}).$$

Alors les propriétés suivantes sont vérifiées (cf. exercices) :

- (i) si v' est bornée, alors $v(X_t) - \int_0^t Av(X_s) ds$ est une martingale,
- (ii) $\lim_{t \downarrow 0} \mathbb{E} \left(\frac{1}{t} (v(X_t) - v(x)) \right) = Av(x),$
- (iii) $\lim_{t \downarrow 0} \mathbb{E} \left(\frac{1}{t} (X_t - x) \right) = f(x) \quad \text{et} \quad \lim_{t \downarrow 0} \mathbb{E} \left(\frac{1}{t} (X_t - x)^2 \right) = g(x)^2.$

B) Diffusion inhomogène en temps

Soient $x \in \mathbb{R}$, $(B_t, t \in \mathbb{R})$ un mouvement brownien standard et $f, g : \mathbb{R}_+ \times \mathbb{R} \rightarrow \mathbb{R}$ conjointement continues en (t, x) et lipschitziennes en x . On pose X la solution de l'EDS

$$dX_t = f(t, X_t) dt + g(t, X_t) dB_t \quad \text{et} \quad X_0 = x.$$

Pour $t \in \mathbb{R}_+$ fixé, on pose A_t l'opérateur linéaire différentiel de $C^2(\mathbb{R})$ dans $C(\mathbb{R})$ défini par

$$A_t v(x) = f(t, x) v'(x) + \frac{1}{2} g(t, x)^2 v''(x), \quad v \in C^2(\mathbb{R}).$$

Pour $u \in C^{1,2}(\mathbb{R}_+ \times \mathbb{R})$, on note encore

$$A_t u(t, x) = f(t, x) \frac{\partial u}{\partial x}(t, x) + \frac{1}{2} g(t, x)^2 \frac{\partial^2 u}{\partial x^2}(t, x).$$

De la même manière que ci-dessus, on montre que si $\frac{\partial u}{\partial x}$ est bornée, alors

$$u(t, X_t) - \int_0^t \left(\frac{\partial u}{\partial s}(s, X_s) + A_s u(s, X_s) \right) ds \quad \text{est une martingale.}$$

En particulier, si $\frac{\partial u}{\partial t} + A_t u = 0$, alors le processus $u(t, X_t)$ est une martingale.

Application : évaluation d'options européennes

Soient $x_0 \geq 0$, $(B_t, t \in \mathbb{R})$ un mouvement brownien standard et $f, g : \mathbb{R}_+ \times \mathbb{R} \rightarrow \mathbb{R}$ conjointement continues en (t, x) et lipschitziennes en x . Supposons de plus qu'il existe $K_1, K_2 > 0$ telles que

$$|f(t, x)| \leq K_1 \quad \text{et} \quad |g(t, x)| \leq K_2, \quad \forall (t, x) \in \mathbb{R}_+ \times \mathbb{R}.$$

On modélise l'évolution du prix d'un actif X par l'EDS

$$dX_t = f(t, X_t) dt + g(t, X_t) dB_t \quad \text{et} \quad X_0 = x_0.$$

Une option européenne d'échéance $T > 0$ sur l'actif (X_t) est une v.a. donnée par $Z_T = h(X_T)$ pour une certaine fonction $h \in C(\mathbb{R})$ (Z_T représente la valeur de l'option au temps T).

Soient (\mathcal{F}_t) la filtration à laquelle le processus X est adapté et $\tilde{\mathbb{P}}_T$ la probabilité sous laquelle le processus X est une martingale. Par ce qu'on a vu précédemment, il existe également un m.b.s. (\tilde{B}_t) sous $\tilde{\mathbb{P}}_T$ tel que

$$dX_t = g(t, X_t) d\tilde{B}_t.$$

On suppose ici (pour simplifier) que Z_T peut s'écrire comme $Z_T = Z_t + \int_t^T H_s dX_s$, où H est une stratégie d'investissement continue et adaptée à (\mathcal{F}_t) et Z_t est une v.a. positive \mathcal{F}_t -mesurable. De là, on déduit que

$$\tilde{\mathbb{E}}_T(Z_T | \mathcal{F}_t) = Z_t + \tilde{\mathbb{E}}_T \left(\int_0^T H_s dX_s - \int_0^t H_s dX_s \middle| \mathcal{F}_t \right) = Z_t,$$

car (X_t) est une martingale sous $\tilde{\mathbb{P}}_T$. Par la propriété de Markov, on a donc

$$Z_t = \tilde{\mathbb{E}}_T(h(X_T) | \mathcal{F}_t) = \tilde{\mathbb{E}}_T(h(X_T) | X_t) = z(t, X_t),$$

où

$$z(t, x) = \mathbb{E}_T(h(X_T) | X_t = x) = \tilde{\mathbb{E}}_T(h(X_T^{t,x})). \quad (19)$$

Par ce qu'on vient de voir plus haut sur le lien entre EDS et EDP, ceci signifie que la fonction $z(t, x)$ est solution de l'équation :

$$\frac{\partial z}{\partial t}(t, x) + \frac{1}{2} g(t, x)^2 \frac{\partial^2 z}{\partial x^2}(t, x) = 0. \quad \text{et} \quad z(T, x) = h(x) \quad (20)$$

Au temps $t < T$, la valeur de l'option Z_t est donc donnée par $z(t, X_t)$, où z est solution de l'équation ci-dessus.

Remarque 2.8.10. - Le terme de dérive f est absent de l'équation pour z (de la même façon que μ est absent de la formule de Black & Scholes plus classique donnée à la section précédente).

- Dans le cas particulier où $f(t, x) = \mu x$, $g(t, x) = \sigma x$ (attention : $g(t, 0) = 0$, mais ça ne pose pas de problème ici), on a

$$\frac{\partial z}{\partial t}(t, x) + \frac{1}{2} \sigma^2 x^2 \frac{\partial^2 z}{\partial x^2}(t, x) = 0 \quad \text{et} \quad z(T, x) = h(x).$$

En résolvant cette équation (ou en repartant directement de la formule (19) et en utilisant l'expression trouvée pour la solution $X_T^{t,x}$ de l'EDS), on trouve donc que

$$z(t, x) = \int_{\mathbb{R}} h(x e^{\sigma y - \frac{\sigma^2}{2}(T-t)}) f_{T-t}(y) dy,$$

où f_{T-t} est la densité de la loi $\mathcal{N}(0, T-t)$.

On peut finalement déterminer quelle est la stratégie (H_t) de couverture de l'option, en utilisant la formule d'Itô pour $Z_T = z(T, X_T)$:

$$\begin{aligned} Z_T - Z_t &= z(T, X_T) - z(t, X_t) \\ &= \int_t^T \frac{\partial z}{\partial t}(s, X_s) ds + \int_t^T \frac{\partial z}{\partial x}(s, X_s) dX_s + \frac{1}{2} \int_t^T \frac{\partial^2 z}{\partial x^2}(s, X_s) d\langle X \rangle_s \\ &= \int_t^T \frac{\partial z}{\partial x}(s, X_s) dX_s, \end{aligned}$$

car $d\langle X \rangle_s = g(s, X_s)^2 ds$ et z satisfait l'équation (20). D'autre part, on sait que

$$Z_T - Z_t = \int_t^T H_s dX_s.$$

En comparant les deux expressions, on trouve finalement que

$$H_s = \frac{\partial z}{\partial x}(s, X_s).$$

Donnons tout d'abord une liste de définitions.

Définition 2.9.1. - Un mouvement brownien standard à n dimensions est une famille $(B^{(1)}, \dots, B^{(n)})$ de n mouvements browniens standard par rapport à une même filtration (\mathcal{F}_t) , supposés de plus indépendants les uns des autres.

- Une martingale à n dimensions est une famille $(M^{(1)}, \dots, M^{(n)})$ de n martingales par rapport à la même filtration (\mathcal{F}_t) .

- Un processus d'Itô à n dimensions est une famille $(X^{(1)}, \dots, X^{(n)})$ de n processus d'Itô adaptés à la même filtration (\mathcal{F}_t) .

Notation : On pose $\underline{B}_t = (B_t^{(1)}, \dots, B_t^{(n)})$, $\underline{M}_t = (M_t^{(1)}, \dots, M_t^{(n)})$ et $\underline{X}_t = (X_t^{(1)}, \dots, X_t^{(n)})$.

Intégrale stochastique multidimensionnelle

Soit $\underline{B} = (B^{(1)}, \dots, B^{(m)})$ un mouvement brownien standard à m dimensions, adapté à (\mathcal{F}_t) , et $H = (H^{(i,j)})_{i,j=1}^{n,m}$ une famille de processus continus et adaptés à (\mathcal{F}_t) telle que

$$\mathbb{E} \left(\int_0^t (H_s^{(i,j)})^2 ds \right) < \infty, \quad \forall t \in \mathbb{R}_+, i \in \{1, \dots, n\}, j \in \{1, \dots, m\}.$$

On définit

$$M_t^{(i)} = \sum_{j=1}^m \int_0^t H_s^{(i,j)} dB_s^{(j)}, \quad i = 1, \dots, n.$$

et $\underline{M}_t = (M_t^{(1)}, \dots, M_t^{(n)})$. Le processus (\underline{M}_t) est une martingale continue de carré intégrable à n dimensions. On note encore de manière plus concise : $\underline{M}_t = \int_0^t H_s d\underline{B}_s$.

Le processus \underline{M} modélise (par exemple) les fluctuations de n actifs $M^{(1)}, \dots, M^{(n)}$ dues à m sources aléatoires indépendantes $B^{(1)}, \dots, B^{(m)}$. Ceci permet de rendre compte des corrélations des actifs :

$$\begin{aligned} \langle M^{(i)}, M^{(k)} \rangle_t &= \sum_{j,l=1}^m \langle (H^{(i,j)} \cdot B^{(j)}), (H^{(k,l)} \cdot B^{(l)}) \rangle_t \\ &= \sum_{j,l=1}^m \int_0^t H_s^{(i,j)} H_s^{(k,l)} d\langle B^{(j)}, B^{(l)} \rangle_s = \sum_{j=1}^m \int_0^t H_s^{(i,j)} H_s^{(k,j)} ds, \end{aligned}$$

$$\text{car } \langle B^{(j)}, B^{(l)} \rangle_s = \begin{cases} s, & \text{si } i = k, \\ 0, & \text{si } i \neq k, \end{cases} \quad (B^{(i)} \perp B^{(k)} \text{ si } i \neq k).$$

Formule d'Itô multidimensionnelle

Soit $\underline{X} = (X^{(1)}, \dots, X^{(n)})$ un processus d'Itô à n dimensions et $f \in C^2(\mathbb{R}^n)$ telle que

$$\mathbb{E} \left(\int_0^t \left(\frac{\partial f}{\partial x_i}(\underline{X}_s) \right)^2 d\langle X^{(i)} \rangle_s \right) < \infty, \quad \forall t \in \mathbb{R}_+, i \in \{1, \dots, n\}.$$

Alors

$$f(\underline{X}_t) - f(\underline{X}_0) = \sum_{i=1}^n \int_0^t \frac{\partial f}{\partial x_i}(\underline{X}_s) dX_s^{(i)} + \frac{1}{2} \sum_{i,k=1}^n \int_0^t \frac{\partial^2 f}{\partial x_i \partial x_k}(\underline{X}_s) d\langle X^{(i)}, X^{(k)} \rangle_s.$$

Exemple 2.9.2. Considérons le cas $n = 2$: on a déjà vu cette formule avec $X_t^{(1)} = t$, $X_t^{(2)} = B_t$ ou $X_t^{(1)} = V_t$, $X_t^{(2)} = M_t$ avec V à variation bornée et M martingale.

Exemple 2.9.3. Pour n quelconque et $\underline{X} = \underline{B}$, on a $\langle B^{(i)}, B^{(k)} \rangle_t = \begin{cases} t, & \text{si } i = k, \\ 0, & \text{si } i \neq k, \end{cases}$ donc

$$f(\underline{B}_t) - f(\underline{B}_0) = \sum_{i=1}^n \int_0^t \frac{\partial f}{\partial x_i}(\underline{B}_s) dB_s^{(i)} + \frac{1}{2} \int_0^t \Delta f(\underline{B}_s) ds,$$

où $\Delta f(\underline{x}) = \sum_{i=1}^n \frac{\partial^2 f}{\partial x_i^2}(\underline{x})$. De la formule ci-dessus, on déduit que le processus

$$f(\underline{B}_t) - f(\underline{B}_0) - \frac{1}{2} \int_0^t \Delta f(\underline{B}_s) ds$$

est une martingale (du moment que la condition $\mathbb{E} \left(\int_0^t \left(\frac{\partial f}{\partial x_i}(\underline{B}_s) \right)^2 ds \right) < \infty$ est vérifiée $\forall i$).

Donc en particulier, si f est harmonique (i.e. $\Delta f(\underline{x}) = 0$), alors $f(\underline{B}_t)$ est une martingale. D'autre part, si f est sous-harmonique (i.e. $\Delta f(\underline{x}) \geq 0$), alors $f(\underline{B}_t)$ est une sous-martingale (d'où l'origine de l'expression contre-intuitive "sous-martingale" pour un processus qui a tendance à monter).

EDS multidimensionnelles

Soient $\underline{x}_0 \in \mathbb{R}^n$, \underline{B} un mouvement brownien standard à m dimensions, $\underline{f} : \mathbb{R}_+ \times \mathbb{R}^n \rightarrow \mathbb{R}^n$ une fonction continue en (t, \underline{x}) et lipschitzienne en \underline{x} , i.e.

$$\|\underline{f}(t, \underline{x}) - \underline{f}(t, \underline{y})\| \leq K \|\underline{x} - \underline{y}\|, \quad \forall t \in \mathbb{R}_+, \underline{x}, \underline{y} \in \mathbb{R}^n,$$

et $\underline{g}^{(1)}, \dots, \underline{g}^{(m)} : \mathbb{R}_+ \times \mathbb{R}^n \rightarrow \mathbb{R}^n$ continues en (t, \underline{x}) et lipschitziennes en \underline{x} . Alors il existe une unique solution forte (\underline{X}_t) à l'équation

$$d\underline{X}_t = \underline{f}(t, \underline{X}_t) dt + \sum_{j=1}^m \underline{g}^{(j)}(t, \underline{X}_t) dB_t^{(j)}, \quad \underline{X}_0 = \underline{x}_0. \quad (21)$$

Rappelons qu'une solution forte est par définition un processus continu et adapté à la même filtration que le mouvement brownien \underline{B} , qui vérifie de plus l'équation intégrale correspondant à l'équation ci-dessus.

Exemple 2.9.4. Dans le cas où $n = 2$ et $m = 1$, nous avons déjà vu un exemple d'EDS multidimensionnelle (ex. 3, série 10) :

$$\begin{aligned} dX_t &= -\frac{1}{2}X_t dt - Y_t dB_t, & X_0 &= 1, \\ dY_t &= -\frac{1}{2}Y_t dt + X_t dB_t, & Y_0 &= 0. \end{aligned}$$

Ici $\underline{f}(t, x, y) = (-\frac{1}{2}x, -\frac{1}{2}y)$ et $\underline{g}^{(1)}(t, x, y) = (-y, x)$ sont des fonctions linéaires, donc lipschitziennes ; il existe donc une unique solution forte (X_t, Y_t) à l'équation ci-dessus (analysée dans l'exercice 3 de la série 10).

Exemple 2.9.5. Considérons encore le cas $n = 2$ et $m = 1$:

$$\begin{aligned} dX_t &= Y_t dt, & X_0 &= 1, \\ dY_t &= -X_t dt + X_t dB_t, & Y_0 &= 0. \end{aligned}$$

Cette équation, qui paraît simple à résoudre au premier abord, ne l'est pas tant que ça en réalité !

Exemple 2.9.6. Le modèle de Black & Scholes à plusieurs dimensions est le suivant :

$$dX_t^{(i)} = \mu_i X_t^{(i)} dt + \sum_{j=1}^m \sigma_{ij} X_t^{(i)} dB_t^{(j)}, \quad X_0^{(i)} = x_0^{(i)}, \quad i \in \{1, \dots, n\}.$$

Remarquer que les n équations ci-dessus sont découplées (contrairement aux deux exemples précédents), ce qui rend nettement plus facile la recherche de la solution.

EDS multidimensionnelles “linéaires”

Soient $\underline{x}_0 \in \mathbb{R}^n$, \underline{B} un mouvement brownien standard à m dimensions, $A = (a_{ik})$ une matrice $n \times n$ et $\Sigma = (\sigma_{ij})$ une matrice $n \times m$. On considère l'équation

$$d\underline{X}_t = A\underline{X}_t dt + \Sigma d\underline{B}_t, \quad \underline{X}_0 = \underline{x}_0.$$

Ici, $f^{(i)}(t, \underline{x}) = \sum_{k=1}^n a_{ik} x_k$ et $g^{(i,j)}(t, \underline{x}) = \sigma_{ij}$ (NB : $g^{(i,j)}$ est la i^e composante du vecteur $\underline{g}^{(j)}$).

Pour expliciter la solution de cette équation, on a besoin de la solution de l'équation homogène. Soit donc (Φ_t) le processus (déterministe) à valeurs dans l'espace des matrices $n \times n$, solution de

$$d\Phi_t = A\Phi_t dt, \quad \Phi_0 = Id,$$

où Id est la matrice identité. Suivant la matrice A considérée, le processus Φ peut être difficile à déterminer. Il s'écrit toutefois sous la forme synthétique

$$\Phi_t = \exp(tA) = \sum_{l \geq 0} \frac{(tA)^l}{l!}.$$

Remarquer d'autre part que $\Phi_{t+s} = \Phi_t \Phi_s$ et $\Phi_{-t} = \Phi_t^{-1}$.

Proposition 2.9.7. *La solution \underline{X} de l'EDS "linéaire" ci-dessus est donnée par*

$$\underline{X}_t = \Phi_t \underline{x}_0 + \int_0^t \Phi_{t-s} \Sigma d\underline{B}_s$$

Démonstration. Ecrivons $\underline{X}_t = \Phi_t \underline{Y}_t$; on a alors, du fait que Φ est un processus à variation bornée (i.e. chacune des entrées de la matrice est un processus à variation bornée) :

$$\begin{aligned} d\underline{X}_t &= (d\Phi_t) \underline{Y}_t + \Phi_t d\underline{Y}_t + 0 = A\Phi_t \underline{Y}_t dt + \Phi_t d\underline{Y}_t \\ &= A\underline{X}_t dt + \Sigma d\underline{B}_t. \end{aligned}$$

On en déduit donc que

$$d\underline{Y}_t = \Phi_t^{-1} \Sigma d\underline{B}_t, \quad \underline{Y}_0 = \underline{x}_0,$$

d'où

$$\underline{X}_t = \Phi_t \left(\underline{x}_0 + \int_0^t \Phi_s^{-1} \Sigma d\underline{B}_s \right) = \Phi_t \underline{x}_0 + \int_0^t \Phi_{t-s} \Sigma d\underline{B}_s.$$

□

Exemple 2.9.8. Lorsque $n = m = 1$, l'équation ci-dessus devient :

$$dX_t = aX_t dt + \sigma dB_t,$$

dont la solution est le processus d'Ornstein-Uhlenbeck vu précédemment.

Exemple 2.9.9. Considérons le cas $n = 2$ et $m = 1$:

$$\begin{aligned} dX_t &= Y_t dt, & X_0 &= 1, \\ dY_t &= -X_t dt + \sigma dB_t, & Y_0 &= 0. \end{aligned}$$

Ici, $A = \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix}$ et $\Sigma = \begin{pmatrix} 0 \\ \sigma \end{pmatrix}$. Formellement, ce système des deux équations du premier ordre peut se récrire comme une seule équation du second ordre :

$$\frac{d^2 X_t}{dt^2} = -X_t + \sigma \frac{dB_t}{dt},$$

qui décrit le mouvement d'un ressort perturbé par un bruit blanc.

Vecteur de dérive, matrice de diffusion et solution faible

Revenons au cas général. Le processus \underline{X} solution forte de l'EDS (21) est appelé une diffusion (c'est aussi un processus de Markov). On vérifie que :

(i) Le processus $M_t^{(i)} = X_t^{(i)} - \int_0^t f^{(i)}(s, \underline{X}_s) ds$ est une martingale pour tout $i \in \{1, \dots, n\}$; le vecteur $\underline{f}(t, \underline{x})$ est appelé vecteur de dérive du processus \underline{X} .

(ii) Du fait que

$$\begin{aligned} \langle X^{(i)}, X^{(k)} \rangle_t &= \langle M^{(i)}, M^{(k)} \rangle_t = \sum_{j,l=1}^m \int_0^t g^{(i,j)}(s, \underline{X}_s) g^{(k,l)}(s, \underline{X}_s) d\langle B^{(j)}, B^{(l)} \rangle_s \\ &= \int_0^t \sum_{j=1}^m g^{(i,j)}(s, \underline{X}_s) g^{(k,j)}(s, \underline{X}_s) ds, \end{aligned}$$

on trouve que le processus

$$N_t^{(i,k)} = M_t^{(i)} M_t^{(k)} - \int_0^t \sum_{j=1}^m g^{(i,j)}(s, \underline{X}_s) g^{(k,j)}(s, \underline{X}_s) ds$$

est une martingale pour tout $i, k \in \{1, \dots, n\}$. La matrice $G(t, \underline{x})$ définie par

$$G^{(i,k)}(t, \underline{x}) = \sum_{j=1}^m g^{(i,j)}(t, \underline{x}) g^{(k,j)}(t, \underline{x})$$

est appelé la matrice de diffusion du processus \underline{X} (noter que $G(t, \underline{x}) = g(t, \underline{x}) g(t, \underline{x})^T$).

Finalement, un processus \underline{X} tel que les processus $M^{(i)}$ et $N^{(i,k)}$ définis ci-dessus sont des martingales pour tout $i, k \in \{1, \dots, n\}$ est appelé une solution faible de l'EDS (21).

Existence d'une mesure martingale

Ci-dessous, on recherche ici des conditions suffisantes sur \underline{f} et G garantissant l'existence d'une mesure de probabilité $\tilde{\mathbb{P}}_T$ sous laquelle les processus $X^{(1)}, \dots, X^{(n)}$ soient simultanément des martingales sur l'intervalle $[0, T]$.

Remarque 2.9.10. Comme on l'a vu plus haut, l'existence d'une mesure martingale permet de fixer un prix et une stratégie de couverture à des options en mathématiques financières. Cette dernière affirmation n'est toutefois pas tout à fait exacte. Pour que cela soit possible, il faut encore que la mesure martingale soit unique (sinon, on voit bien que le prix et la stratégie de couverture dépendront de la mesure martingale choisie). Nous n'entrerons cependant pas plus loin dans les détails dans ce cours.

Voici donc les hypothèses supplémentaires à effectuer sur \underline{f} et G ; il existe $K_1, K_2 > 0$ telles que

$$\begin{aligned} \text{(i)} \quad & \|\underline{f}(t, \underline{x})\| \leq K_1, \quad \forall t \in \mathbb{R}_+, \underline{x} \in \mathbb{R}^n, \\ \text{(ii)} \quad & \sum_{i,k=1}^n G^{(i,k)}(t, \underline{x}) \xi_i \xi_k \geq K_2 \|\underline{\xi}\|^2, \quad \forall t \in \mathbb{R}_+, \underline{x}, \underline{\xi} \in \mathbb{R}^n. \end{aligned}$$

Si l'hypothèse (ii) est satisfaite, on dit que la diffusion \underline{X} est non-dégénérée.

Remarque 2.9.11. Du fait que $G(t, \underline{x}) = g(t, \underline{x}) g(t, \underline{x})^T$, il est toujours vrai que

$$\sum_{i,k=1}^n G^{(i,k)}(t, \underline{x}) \xi_i \xi_k \geq 0, \quad \forall t, \underline{x}, \underline{\xi}. \quad (22)$$

L'hypothèse (ii) assure que pour un point (t, \underline{x}) donné, on a de plus l'inégalité stricte :

$$\sum_{i,k=1}^n G^{(i,k)}(t, \underline{x}) \xi_i \xi_k > 0, \quad \forall \underline{\xi} \neq 0. \quad (23)$$

Etant donné que (22) est vérifié, la condition (23) peut se reformuler de plusieurs manières équivalentes :

- (a) toutes les valeurs propres de $G(t, \underline{x})$ sont positives,
- (b) $\det G(t, \underline{x}) > 0$,
- (c) $G(t, \underline{x})$ est inversible,
- (d) $\text{rang } G(t, \underline{x}) = n$.

Ceci nous amène à la remarque suivante : du fait que $G(t, \underline{x}) = g(t, \underline{x}) g(t, \underline{x})^T$ et que $g(t, \underline{x})$ est une matrice $n \times m$, on sait que $\text{rang } G(t, \underline{x}) \leq m \wedge n$, donc que

$$\text{si } m < n, \text{ alors } \text{rang } G(t, \underline{x}) < n, \quad \forall (t, \underline{x}), \quad (24)$$

auquel cas l'hypothèse (ii) ci-dessus ne peut être satisfaite.

Proposition 2.9.12. *Sous les hypothèses (i) et (ii) effectuées ci-dessus, il existe une mesure martingale $\tilde{\mathbb{P}}_T$ (i.e. une mesure de probabilité sous laquelle tous les processus $X^{(1)}, \dots, X^{(n)}$ sont simultanément des martingales).*

Démonstration. On définit les processus et fonctions suivants :

$$\begin{aligned} Z_t^{(i)} &= \sum_{j=1}^m \int_0^t g^{(i,j)}(s, \underline{X}_s) dB_s^{(j)}, \quad i \in \{1, \dots, n\}, \\ h^{(i)}(t, \underline{x}) &= \sum_{k=1}^n f^{(k)}(\underline{x}) (G^{-1})^{(k,i)}(t, \underline{x}), \quad i \in \{1, \dots, n\}, \\ M_t &= - \sum_{i=1}^n \int_0^t h^{(i)}(s, \underline{X}_s) dZ_s^{(i)}. \end{aligned}$$

Remarquer que $Z^{(i)}$ est une martingale sous \mathbb{P} pour tout $i \in \{1, \dots, n\}$, de même que M . On peut vérifier que $\langle M \rangle_t \leq Kt$, et on définit Y et $\tilde{\mathbb{P}}_T$ comme à la section 2.7. Par le théorème de Girsanov, on sait donc que pour tout $l \in \{1, \dots, n\}$, $Z_t^{(l)} - \langle M, Z^{(l)} \rangle_t$ est une martingale sous $\tilde{\mathbb{P}}_T$. Calculons ce processus :

$$\begin{aligned}
M_t &= - \sum_{i,k=1}^n \sum_{j=1}^m \int_0^t f^{(k)}(s, \underline{X}_s) (G^{-1})^{(k,i)}(s, \underline{X}_s) g^{(i,j)}(s, \underline{X}_s) dB_s^{(j)} \\
\langle M, Z^{(l)} \rangle_t &= - \sum_{i,k=1}^n \sum_{j=1}^m \int_0^t f^{(k)}(s, \underline{X}_s) (G^{-1})^{(k,i)}(s, \underline{X}_s) g^{(i,j)}(s, \underline{X}_s) g^{(l,j)}(s, \underline{X}_s) ds \\
&= - \sum_{i,k=1}^n \int_0^t f^{(k)}(s, \underline{X}_s) (G^{-1})^{(k,i)}(s, \underline{X}_s) G^{(i,l)}(s, \underline{X}_s) ds \\
&= - \sum_{k=1}^n \int_0^t f^{(k)}(s, \underline{X}_s) \delta_{kl} ds = - \int_0^t f^{(l)}(s, \underline{X}_s) ds
\end{aligned}$$

Donc

$$Z_t^{(l)} - \langle M, Z^{(l)} \rangle_t = \sum_{j=1}^m \int_0^t g^{(i,j)}(s, \underline{X}_s) dB_s^{(j)} + \int_0^t f^{(l)}(s, \underline{X}_s) ds = X_t^{(l)} - X_0^{(l)}$$

est une martingale sous $\tilde{\mathbb{P}}_T$ pour tout $l \in \{1, \dots, n\}$ et la proposition est démontrée. \square

Exemple 2.9.13. Appliquons le résultat ci-dessus au modèle de Black & Scholes multidimensionnel (cf. exemple 2.9.6). Dans ce cas, la matrice de diffusion G ne dépend que de \underline{x} et est donnée par

$$G^{(i,k)}(\underline{x}) = \sum_{j=1}^m \sigma_{ij} \sigma_{kj} x_i x_k,$$

La diffusion \underline{X} est donc non-dégénérée du moment que la matrice G_0 donnée par $G_0^{(i,k)} = \sum_{j=1}^m \sigma_{ij} \sigma_{kj}$ est de rang n (le fait que rang $G(\underline{x})$ soit strictement plus petit que n lorsque l'une des composantes de \underline{x} est nulle ne pose pas de problème ici, car on peut montrer facilement qu'aucune des composantes de la diffusion \underline{X} ne touche le point 0 au cours du temps).

Etant donné la remarque (24), on voit que rang $G_0 = n$ n'est possible que si $m \geq n$. On peut en fait montrer que la condition $m \geq n$ est nécessaire pour l'existence d'une mesure martingale (et que rang $G_0 = n$ est une condition nécessaire et suffisante).

Il est possible de généraliser tous les résultats de la section 2.8 au cas multidimensionnel. Toutefois, notre but ici est de considérer un autre type d'équation aux dérivées partielles, dite elliptique (pour le cas unidimensionnel, se référer à l'exercice 3, série 9).

Résultat d'analyse (EDP elliptique)

Soient D un domaine ouvert borné dans \mathbb{R}^n , ∂D sa frontière, $\overline{D} = D \cup \partial D$ et $h \in C(\partial D)$. Alors il existe une unique fonction $u \in C^2(D) \cap C(\overline{D})$ telle que

$$\begin{cases} \Delta u(\underline{x}) = 0, & \forall \underline{x} \in D, \\ u(\underline{x}) = h(\underline{x}), & \forall \underline{x} \in \partial D. \end{cases} \quad (25)$$

Noter que la solution u est une fonction à valeurs dans \mathbb{R} (et non dans \mathbb{R}^n).

Soit (\underline{B}_t^x) un mouvement brownien à n dimensions partant au temps $t = 0$ du point $\underline{x} \in D$. Soit également $\tau = \inf\{t > 0 : \underline{B}_t^x \notin D\}$, le premier temps de sortie du domaine D (remarquer que τ est un temps d'arrêt et que $\underline{B}_\tau^x \in \partial D$).

Proposition 2.9.14. *La solution de l'équation (25) s'écrit :*

$$u(\underline{x}) = \mathbb{E}(h(\underline{B}_\tau^x)), \quad \forall \underline{x} \in D.$$

Démonstration. On montre tout d'abord que le processus $(u(\underline{B}_t^x), t \geq 0)$ est une martingale (unidimensionnelle). En utilisant la formule d'Itô multidimensionnelle, on a pour tout $t < \tau$ (condition qui assure que $\underline{B}_s^x \in D$ pour tout $s \leq t$) :

$$\begin{aligned} u(\underline{B}_t^x) - u(\underline{B}_0^x) &= \sum_{i=1}^n \int_0^t \frac{\partial u}{\partial x_i}(\underline{B}_s^x) d\underline{B}_s^{x,(i)} + \frac{1}{2} \int_0^t \Delta u(\underline{B}_s^x) ds \\ &= \sum_{i=1}^n \int_0^t \frac{\partial u}{\partial x_i}(\underline{B}_s^x) d\underline{B}_s^{x,(i)}, \end{aligned}$$

étant donné que u satisfait (25). Le processus $(u(\underline{B}_t^x))$ est bien une martingale, donc en appliquant le théorème d'arrêt (et en laissant passer quelques détails), on obtient :

$$u(\underline{x}) = \mathbb{E}(u(\underline{B}_0^x)) = \mathbb{E}(u(\underline{B}_\tau^x)) = \mathbb{E}(h(\underline{B}_\tau^x)),$$

où on a utilisé la condition de bord de l'équation (25) (ce qui est rendu possible par le fait que $\underline{B}_\tau^x \in \partial D$). \square

2.10 Analyse et simulation numériques des EDS

Considérons l'EDS (unidimensionnelle)

$$dX_t = f(t, X_t) dt + g(t, X_t) dB_t, \quad X_0 = x_0.$$

Sous des hypothèses générales (p. ex. f, g continues en (t, x) et lipschitziennes en x), on sait qu'il existe une unique solution forte (X_t) à une telle équation. Pourtant, dès que f et g sont un peu compliquées, on ne connaît pas d'expression analytique pour la solution. On comprend donc l'intérêt de développer des méthodes pour simuler numériquement la solution de telles équations.

Schéma d'Euler

Le schéma présenté ici est l'adaptation au cas aléatoire du schéma d'Euler classique pour les équations différentielles ordinaires (EDO). Son analyse diffère toutefois en plusieurs aspects (voir plus loin).

Supposons donc qu'il existe une unique solution forte à l'équation ci-dessus (sans quoi la méthode numérique décrite ci-dessous est vouée à l'échec). Pour $t > r$, on a donc l'égalité

$$X_t = X_r + \int_r^t f(s, X_s) ds + \int_r^t g(s, X_s) dB_s.$$

En supposant que t est proche de r , on peut utiliser la continuité des fonctions $s \mapsto f(s, X_s)$ et $g \mapsto g(s, X_s)$ pour approximer la relation ci-dessus par

$$X_t \simeq X_r + f(r, X_r) (t - r) + g(r, X_r) (B_t - B_r).$$

(Noter qu'on a choisi à dessein le point à gauche de l'intervalle de façon à approximer l'intégrale d'Itô). Soient maintenant $T \in \mathbb{R}_+$ et $N \in \mathbb{N}$ fixés; en posant $\Delta t = \frac{T}{N}$, $t = (n+1) \Delta t$ et $r = n \Delta t$, on trouve donc

$$X_{(n+1)\Delta t} \simeq X_{n\Delta t} + f(n\Delta t, X_{n\Delta t}) \Delta t + g(n\Delta t, X_{n\Delta t}) (B_{(n+1)\Delta t} - B_{n\Delta t}), \quad (26)$$

où $B_{(n+1)\Delta t} - B_{n\Delta t}$ est une v.a. $\mathcal{N}(0, \Delta t)$ indépendante de $\mathcal{F}_{n\Delta t}$ (rappel : (\mathcal{F}_t) est la filtration à laquelle sont adaptés les processus (B_t) et (X_t)). De là, on déduit le schéma numérique suivant :

$$X_0^{(N)} = x_0, \quad X_{(n+1)\Delta t}^{(N)} = X_{n\Delta t}^{(N)} + f(n\Delta t, X_{n\Delta t}^{(N)}) \Delta t + g(n\Delta t, X_{n\Delta t}^{(N)}) \xi_{n+1} \sqrt{\Delta t}, \quad 0 \leq n < N,$$

où $(\xi_n)_{n=1}^N$ est une suite de v.a. i.i.d. $\sim N(0, 1)$ (de telle sorte à ce que les v.a. $\xi_{n+1} \sqrt{\Delta t}$ et $(B_{(n+1)\Delta t} - B_{n\Delta t})$ soient identiquement distribuées). Ceci définit le schéma numérique $X^{(N)}$

aux instants $n\Delta t$. Pour définir le schéma sur l'intervalle $[0, T]$ tout entier, on “relie les points entre eux”, i.e. on définit pour $t \in]n\Delta t, (n+1)\Delta t[$:

$$X_t^{(N)} = X_{n\Delta t}^{(N)} + (t - n\Delta t) (X_{(n+1)\Delta t}^{(N)} - X_{n\Delta t}^{(N)}).$$

Ainsi, $(X_t^{(N)}, t \in [0, T])$ définit une suite de processus stochastiques.

Noter que la démarche ci-dessus est équivalente à celle décrite au paragraphe 2.1.2 concernant l'approximation du mouvement brownien par une marche aléatoire (qui correspond au cas particulier $f \equiv 0$ et $g \equiv 1$).

Remarque 2.10.1. Attention : une erreur courante en analyse numérique consiste à confondre l'approximation (26) avec le schéma numérique défini une ligne plus bas.

Convergences

Une fois défini le schéma numérique $X^{(N)}$, on peut se demander combien celui-ci est proche de la solution X . Du fait qu'on a affaire à des processus aléatoires, on a le choix entre plusieurs notions de convergence.

A) Convergence en loi

De la même manière que dans le cas $f \equiv 0$ et $g \equiv 1$ étudié au paragraphe 2.1.2, on peut montrer que sous les hypothèses effectuées,

$$\mathbb{P}(X_{t_1}^{(N)} \leq x_1, \dots, X_{t_m}^{(N)} \leq x_m) \xrightarrow{N \rightarrow \infty} \mathbb{P}(X_{t_1} \leq x_1, \dots, X_{t_m} \leq x_m),$$

pour tout $m \geq 1$, $t_1, \dots, t_m \in [0, T]$, $x_1, \dots, x_m \in \mathbb{R}$. Ainsi, la suite de processus $(X^{(N)})$ converge en loi vers X .

Remarque 2.10.2. Pour obtenir la convergence en loi, on peut aussi considérer dans le schéma numérique une suite de v.a. $(\xi_n)_{n=1}^N$ i.i.d. telles que $\mathbb{P}(\xi_n = +1) = \mathbb{P}(\xi_n = -1) = \frac{1}{2}$, ou n'importe quelle autre suite de v.a. i.i.d. de moyenne nulle et de variance unité. Ça n'est par contre pas vrai pour ce qui va suivre.

Remarque 2.10.3. S'il existe une unique solution faible à l'équation, alors la suite $X^{(N)}$ définie ci-dessus converge également en loi vers la solution faible.

B) Convergence en moyenne

La convergence en loi ne dit rien sur la distance entre les trajectoires du schéma numérique et celles de la solution. Pour estimer cette distance, on a besoin de propositions comme celle qui suit.

Proposition 2.10.4. *Sous les hypothèses effectuées (et quelques hypothèses techniques additionnelles), il existe une constante $C_T > 0$ telle que*

$$\mathbb{E}(\sup_{0 \leq t \leq T} |X_t^{(N)} - X_t|) \leq \frac{C_T}{\sqrt{N}}.$$

Cette proposition est donnée sans démonstration. La technique de démonstration est similaire à celle utilisée pour montrer l'existence et l'unicité de la solution de l'équation originale (noter toutefois que les processus approximant la solution sont complètement différents dans les deux cas ; le schéma d'itération de Picard ne permet en aucun cas d'obtenir une estimation aussi précise que celle donnée ci-dessus).

Remarque 2.10.5. - A cause du facteur $\frac{1}{\sqrt{N}}$, on dit que le schéma d'Euler est d'ordre $\frac{1}{2}$ pour les EDS. Remarquer que pour les EDO, le schéma d'Euler est par contre d'ordre 1, i.e.

$$\sup_{0 \leq t \leq T} |X_t^{(N)} - X_t| \leq \frac{C_T}{N}.$$

La perte de précision dans le cas aléatoire vient essentiellement du facteur $\sqrt{\Delta t}$ présent dans le terme “intégrale d'Itô” du schéma. Nous verrons plus loin comment remédier à ce problème (cf. schéma de Milstein).

- La constante C_T dépend des constantes de Lipschitz de f et g et augmente exponentiellement avec T (ce qui est également vrai pour les EDO).

Généralisations et intérêt pratique

- Pour $p \geq 1$, on a

$$\mathbb{E}(\sup_{0 \leq t \leq T} |X_t^{(N)} - X_t|^p) \leq \frac{C_T}{N^{p/2}}.$$

(NB : Ceci ne veut pas dire que l'ordre de convergence est meilleur !)

- Pour des diffusions multidimensionnelles, l'ordre de convergence reste le même (i.e. est indépendant de la dimension de la diffusion).

L'intérêt pratique de la simulation des EDS est le suivant : on a vu qu'il est possible de représenter la solution d'une EDP classique à l'aide de la solution d'une EDS, au moyen d'une formule du type :

$$u(t, x) = \mathbb{E}(h(X_T^{t,x})).$$

Pour simuler numériquement la solution d'une EDP dont on ne connaît pas l'expression analytique, on a donc deux possibilités : soit utiliser des méthodes classiques sans passer par la formule de représentation ci-dessus ; soit simuler numériquement le processus $X^{t,x}$ jusqu'au temps T , puis approximer l'espérance ci-dessus. Pour ce faire, une solution simple est d'utiliser la loi des grands nombres, i.e. d'approximer l'espérance par la moyenne de M

réalisations indépendantes de la variable aléatoire (avec M grand) ; ceci implique de répéter M fois la simulation du processus X sur l'intervalle $[t, T]$!

Il faut également noter que l'approximation de l'espérance ci-dessus par une moyenne pose des problèmes car la moyenne possède elle-même une variance non-négligeable. Pour remédier à ce problème, des techniques dites “de réduction de variance” sont utilisées, mais toujours de manière ad hoc.

Malgré tout, cette méthode de simulation des EDP par des processus aléatoires présente un certain intérêt, surtout lorsque la dimension de la diffusion (i.e. du vecteur x) est grande, ce qui arrive fréquemment en mathématiques financières (où la dimension du vecteur x est égale au nombre d'actifs à prendre en considération, qui est généralement de l'ordre d'une dizaine en pratique). Donnons deux raisons pour justifier cet intérêt :

- Les méthodes classiques (déterministes) de résolution des EDP deviennent de plus en plus difficiles à mettre en pratique lorsque la dimension spatiale augmente. On peut penser par exemple à la méthode des éléments finis qui fonctionne parfaitement en dimension 2 et 3 (étant donné qu'elle a été développée principalement pour la mécanique des fluides), mais qui devient quasiment impossible à implémenter au-delà de la dimension 3. Par contre, les schémas de simulation des EDS sont faciles à implémenter en dimension supérieure à 3.
- L'ordre de convergence des méthodes déterministes diminue fortement lorsque la dimension augmente, alors que celui obtenu avec les méthodes stochastiques ne dépend pas de la dimension, comme déjà mentionné plus haut.

Schéma de Milstein

Ce schéma est un raffinement du schéma d'Euler ; sa convergence est d'ordre 1.

Repartons de l'équation de base. Pour $t > r$, on a :

$$X_t = X_r + \int_r^t f(s, X_s) ds + \int_r^t g(s, X_s) dB_s$$

Si on suppose maintenant que $g \in C^{1,2}(\mathbb{R}_+ \times \mathbb{R})$, alors on peut développer $g(s, X_s)$ en utilisant la formule d'Itô :

$$\begin{aligned} g(s, X_s) &= g(r, X_r) + \int_r^s \frac{\partial g}{\partial r}(u, X_u) du + \int_r^s \frac{\partial g}{\partial x}(u, X_u) dX_u + \frac{1}{2} \int_r^s \frac{\partial^2 g}{\partial x^2}(u, X_u) d\langle X \rangle_u \\ &= g(r, X_r) + \int_r^s \frac{\partial g}{\partial t}(u, X_u) du + \int_r^s \frac{\partial g}{\partial x}(u, X_u) f(u, X_u) du \\ &\quad + \int_r^s \frac{\partial g}{\partial x}(u, X_u) g(u, X_u) dB_u + \frac{1}{2} \int_r^s \frac{\partial^2 g}{\partial x^2}(u, X_u) g(u, X_u)^2 du. \end{aligned}$$

Ainsi on trouve que

$$g(s, X_s) = g(r, X_r) + \int_r^s \frac{\partial g}{\partial x}(u, X_u) g(u, X_u) dB_u + \int_r^s H_u du,$$

où H est un certain processus continu. On peut donc approximer $g(s, X_s)$ par

$$g(s, X_s) \simeq g(r, X_r) + \frac{\partial g}{\partial x}(r, X_r) g(r, X_r) (B_s - B_r) + H_r (s - r).$$

En introduisant cette approximation dans l'équation pour X , on trouve :

$$\begin{aligned} X_t \simeq & X_r + f(r, X_r) (t - r) + g(r, X_r) (B_t - B_r) + \frac{\partial g}{\partial x}(r, X_r) g(r, X_r) \int_r^t (B_s - B_r) dB_s \\ & + H_r \int_r^t (s - r) dB_r. \end{aligned}$$

Or un calcul rapide montre que

$$\int_r^t (B_s - B_r) dB_s = \frac{(B_t - B_r)^2 - (t - r)}{2}.$$

et que

$$\int_r^t (s - r) dB_s \sim \mathcal{N} \left(0, \int_r^t (s - r)^2 ds = \frac{(t - r)^3}{3} \right).$$

Ainsi, le premier terme est d'ordre $t - r$ tandis que le second est d'ordre $(t - r)^{3/2}$. Si donc $t = (n + 1)\Delta t$ et $r = n\Delta t$, le second terme sera négligeable par rapport au premier. C'est pourquoi on ne garde que le premier terme dans l'approximation, qui devient :

$$X_t \simeq X_r + f(r, X_r) (t - r) + g(r, X_r) (B_t - B_r) + \frac{\partial g}{\partial x}(r, X_r) g(r, X_r) \frac{(B_t - B_r)^2 - (t - r)}{2}.$$

Ceci nous amène à définir le schéma numérique suivant :

$$\begin{aligned} X_0^{(N)} &= x_0, \quad X_{(n+1)\Delta t}^{(N)} = X_{n\Delta t}^{(N)} + f(n\Delta t, X_{n\Delta t}^{(N)}) \Delta t + g(n\Delta t, X_{n\Delta t}^{(N)}) \xi_{n+1} \sqrt{\Delta t} \\ &+ \frac{\partial g}{\partial x}(n\Delta t, X_{n\Delta t}^{(N)}) g(n\Delta t, X_{n\Delta t}^{(N)}) \left(\frac{\xi_{n+1}^2 - 1}{2} \right) \Delta t, \end{aligned}$$

où $(\xi_n)_{n=1}^N$ est une suite de v.a. i.i.d. $\sim N(0, 1)$.

Proposition 2.10.6. *Sous beaucoup d'hypothèses (principalement f , g et les dérivées de g continues en (t, x) et lipschitziennes en x), on a*

$$\mathbb{E} \left(\sup_{0 \leq t \leq T} |X_t^{(N)} - X_t| \right) \leq \frac{C_T}{N}.$$

Le schéma de Milstein est donc un schéma d'ordre 1 (mais qui nécessite un développement de second ordre faisant intervenir la formule d'Itô!).

Dans cette section, nous passons en revue un certain nombre de points laissés de côté dans le cours.

Convergence des martingales

Soit M une martingale continue. On peut se poser la question de savoir sous quelles conditions M_t converge lorsque t tend vers l'infini. Il y a plusieurs réponses possibles à cette question. Voici un exemple de réponse.

Proposition 2.11.1. *Si*

$$\mathbb{E}(\sup_{t \geq 0} |M_t|^2) < \infty, \quad (27)$$

alors il existe une v.a. M_∞ telle que

$$\mathbb{E}(|M_t - M_\infty|^2) \xrightarrow{t \rightarrow \infty} 0, \quad \text{et} \quad M_t = \mathbb{E}(M_\infty | \mathcal{F}_t). \quad (28)$$

Remarque 2.11.2. - Le mouvement brownien standard B ne satisfait pas la condition (27).
- La condition plus faible $\mathbb{E}(\sup_{t \geq 0} |M_t|) < \infty$ ne suffit pas à garantir l'existence d'une v.a. M_∞ telle que $\mathbb{E}(|M_t - M_\infty|) \xrightarrow{t \rightarrow \infty} 0$ et $M_t = \mathbb{E}(M_\infty | \mathcal{F}_t)$.

Seconde version du théorème d'arrêt

Soient M une martingale continue vérifiant (27) et $0 \leq \tau_1 \leq \tau_2 \leq \infty$ deux temps d'arrêt. Alors

$$\mathbb{E}(M_{\tau_2} | \mathcal{F}_{\tau_1}) = M_{\tau_1}. \quad (29)$$

(Pour la démonstration, on utilise le fait que $M_\tau = \mathbb{E}(M_\infty | \mathcal{F}_\tau)$ pour tout temps d'arrêt τ , ce qu'il faut bien sûr démontrer !)

Remarque 2.11.3. Pour que l'égalité (29) soit vérifiée, on voit donc qu'il est nécessaire d'effectuer une hypothèse soit sur les temps d'arrêt τ_1 et τ_2 , soit sur la martingale elle-même (sans quoi on arrive à une contradiction avec des martingales du type "quitte ou double" évoquées au chapitre 1).

Martingale arrêtée

Proposition 2.11.4. *Soient M une martingale continue et τ un temps d'arrêt. Alors le processus N défini par $N_t = M_{t \wedge \tau}$ est également une martingale continue.*

Ceci se démontre aisément (remarquer qu'après le temps d'arrêt τ , le processus N est constant). En particulier, posons $\tau_a = \inf\{t > 0 : |M_t| \geq a\}$, avec $a > 0$ fixé. On voit que

$$\mathbb{E}(\sup_{t \geq 0} |M_{t \wedge \tau_a}|^2) \leq a^2 < \infty.$$

Par la proposition 2.11.1, on obtient donc que

$$\mathbb{E}(M_{\tau_2 \wedge \tau_a} | \mathcal{F}_{\tau_1}) = M_{\tau_1 \wedge \tau_a}.$$

pour tout $0 \leq \tau_1 \leq \tau_2 \leq \infty$. En choisissant $\tau_1 = 0$ et $\tau_2 = \infty$, on trouve donc que $\mathbb{E}(M_{\tau_a}) = \mathbb{E}(M_0)$, i.e. le théorème d'arrêt s'applique ici même si le temps d'arrêt τ_a n'est pas borné et qu'on ne fait aucune hypothèse sur M . Cette propriété a été utilisée en particulier dans l'exercice 3 de la série 9.

Martingale locale

Définition 2.11.5. Une martingale locale est un processus M tel qu'il existe une suite croissante de temps d'arrêt $0 \leq \tau_1 \leq \dots \leq \tau_n \leq \dots$ avec $\tau_n \xrightarrow{n \rightarrow \infty} \infty$ et $(M_{t \wedge \tau_n})$ martingale pour tout $n \geq 1$.

Donc si M est une martingale locale, on ne sait pas a priori que $\mathbb{E}(|M_t|) < \infty$, ni que $\mathbb{E}(M_t | \mathcal{F}_s) = M_s$. Il faut arrêter le processus M au temps d'arrêt τ_n pour retrouver la propriété de martingale (remarquer qu'on peut toujours se restreindre à la suite de temps d'arrêt $\tau_n = \inf\{t > 0 : |M_t| \geq n\}$).

Remarque 2.11.6. - Par la proposition 2.11.4, toute martingale est une martingale locale.
- Réciproquement, si M est une martingale locale et $\mathbb{E}(\sup_{0 \leq s \leq t} |M_s|) < \infty$ pour tout $t > 0$, alors M est une martingale (par contre, la condition $\mathbb{E}(|M_t|) < \infty$ pour tout $t > 0$ ne suffit pas).

Même si la définition de martingale locale paraît un peu plus compliquée à saisir au premier abord, nous allons voir que cette notion permet de simplifier grandement les énoncés des théorèmes, car on n'a plus besoin d'aucune hypothèse d'intégrabilité sur M . Dans ce qui va suivre, on considère toujours des martingales locales continues.

Variation quadratique d'une martingale locale continue

Définition 2.11.7. La variation quadratique d'une martingale locale continue M est l'unique processus A croissant, continu et adapté tel que $A_0 = 0$ et $(M_t^2 - A_t)$ est une martingale locale continue. On note $A_t = \langle M \rangle_t$.

Troisième version du théorème de Lévy

Soit M une martingale locale continue telle que $\langle M \rangle_t = t$ pour tout $t > 0$. Alors M est un mouvement brownien standard.

Intégrale stochastique

Soient M une martingale locale continue et H un processus continu adapté. Moyennant quelques précautions, il est possible de définir l'intégrale stochastique $N_t = \int_0^t H_s dM_s$. Le processus N ainsi défini est également une martingale locale continue. Ceci se démontre en considérant la suite de temps d'arrêt

$$\tau_n = \inf\{t > 0 : |M_t| \geq n \quad \text{ou} \quad \int_0^t H_s^2 d\langle M \rangle_s \geq n\}.$$

et en remarquant que le processus $(N_{t \wedge \tau_n})$ est une martingale pour tout $n \geq 1$.

Remarque 2.11.8. On n'a plus besoin d'une condition supplémentaire sur H pour s'assurer que l'intégrale stochastique est bien définie !

Semi-martingale continue

Une semi-martingale continue est un processus X qui s'écrit $X = M + V$, où M est une martingale locale continue et V est un processus continu adapté à variation bornée tel que $V_0 = 0$.

Remarque 2.11.9. Attention : le terme "semi-martingale continue" a déjà été utilisé (abusivement) dans ce cours pour désigner un processus d'Itô.

Formules d'Itô

Voyons ici deux cas particuliers de la formule d'Itô, où l'avantage de considérer des martingales locales apparaît clairement.

a) Soient M une martingale locale continue et $f \in C^2(\mathbb{R})$. Alors

$$f(M_t) - f(M_0) = \int_0^t f'(M_s) dM_s + \frac{1}{2} \int_0^t f''(M_s) d\langle M \rangle_s.$$

Remarquer que $f(M)$ est une semi-martingale continue, étant la somme d'une martingale locale continue et d'un processus continu adapté à variation bornée.

b) Soient \underline{B} un m.b.s. à n dimensions et $f \in C^2(\mathbb{R}^n)$. Alors

$$f(\underline{B}_t) - f(\underline{B}_0) = \sum_{i=1}^n \int_0^t \frac{\partial f}{\partial x_i}(\underline{B}_s) dB_s^{(i)} + \frac{1}{2} \int_0^t \Delta f(\underline{B}_s) ds.$$

Remarquer que les deux formules ci-dessus sont valides sans aucune condition d'intégrabilité supplémentaire !

Martingale exponentielle

Proposition 2.11.10. *Soit M une martingale locale continue telle que $M_0 = 0$. Alors le processus Y défini par $Y_t = \exp(M_t - \frac{\langle M \rangle_t}{2})$ est une martingale locale continue. Si de plus $\langle M \rangle_t \leq Kt$ pour tout $t > 0$, alors Y est une martingale.*

Démonstration. La première affirmation découle de la formule d'Itô (a) ci-dessus. Pour démontrer la seconde, posons

$$Z_t = e^{2(M_t - \langle M \rangle_t)}.$$

Par la formule d'Itô, on voit que Z est une martingale locale continue. De plus,

$$Y_t^2 = e^{2M_t - \langle M \rangle_t} = Z_t e^{\langle M \rangle_t}$$

Soit maintenant (τ_n) une suite de temps d'arrêt telle que $(Y_{t \wedge \tau_n})$ est une martingale pour tout $n \geq 1$ (cette suite existe car Y est une martingale locale). Alors par l'inégalité de Doob, on a

$$\mathbb{E}(\sup_{0 \leq s \leq t} |Y_{s \wedge \tau_n}|^2) \leq 4 \mathbb{E}(Y_{t \wedge \tau_n}^2) \leq 4 \mathbb{E}(Z_{t \wedge \tau_n} e^{Kt}) \leq 4e^{Kt},$$

par l'hypothèse et le fait que $\mathbb{E}(Z_{t \wedge \tau_n}) = 1$ pour tout $n \geq 1$. Donc par l'inégalité de Cauchy-Schwarz,

$$\mathbb{E}(\sup_{0 \leq s \leq t} |Y_s|)^2 \leq \mathbb{E}(\sup_{0 \leq s \leq t} |Y_s|^2) = \lim_{n \rightarrow \infty} \mathbb{E}(\sup_{0 \leq s \leq t} |Y_{s \wedge \tau_n}|^2) \leq 4e^{Kt},$$

d'où Y est une martingale par la remarque 2.11.6. □

Remarque 2.11.11. On peut montrer que Y est une martingale sous une condition plus faible, dite de Novikov :

$$\mathbb{E}\left(e^{\frac{\langle M \rangle_t}{2}}\right) < \infty, \quad \forall t > 0.$$

Martingales associée au mouvement brownien standard en dimension n

- Pour $n \geq 2$, on a

$$\|\underline{B}_t\| - \|\underline{B}_0\| = \sum_{i=1}^n \int_0^t \frac{B_s^{(i)}}{\|\underline{B}_s\|} dB_s^{(i)} + \frac{1}{2} \int_0^t \frac{n-1}{\|\underline{B}_s\|} ds,$$

où l'intégrale stochastique du membre de droite est une martingale (on peut même montrer que c'est un mouvement brownien standard en utilisant la troisième version du théorème de Lévy citée plus haut).

- Pour $n = 3$, on a

$$\frac{1}{\|\underline{B}_t\|} - \frac{1}{\|\underline{B}_0\|} = \sum_{i=1}^3 \int_0^t \frac{B_s^{(i)}}{\|\underline{B}_s\|^3} dB_s^{(i)} + 0,$$

(car $\Delta(\frac{1}{\|x\|}) = 0$ pour $x \neq 0$). Dans ce cas cependant, l'intégrale stochastique à droite n'est pas une martingale ; c'est seulement une martingale locale.

Noter qu'on a de manière plus générale : $M_t = \log(\|\underline{B}_t\|)$ est une martingale locale lorsque $n = 2$ et $M_t = \frac{1}{\|\underline{B}_t\|^{n-2}}$ est une martingale locale lorsque $n \geq 3$ (malgré le fait que $\mathbb{E}(|M_t|) < \infty$ pour tout $t > 0$ dans ces deux cas).

Remarque 2.11.12. Dans les deux derniers exemples, f n'est pas dérivable en 0, mais ça n'est pas grave car $\mathbb{P}(\exists t > 0 : \underline{B}_t = 0) = 0$ lorsque $n \geq 2$. Pour la dimension $n = 1$, les choses changent, comme le montre le paragraphe suivant.

Formule de Tanaka

Soit B un m.b.s. de dimension 1. Par une application naïve de la formule d'Itô, on devrait alors avoir, au vu de ce qui précède :

$$|B_t| - |B_0| = \int_0^t \text{sgn}(B_s) dB_s + 0?$$

Ça n'est cependant pas possible, car le théorème de Lévy implique que l'intégrale stochastique dans le membre de droite est un m.b.s., tandis que $|B_t|$ n'est clairement pas un m.b.s. (en particulier parce que ce processus est toujours positif). Le problème ici est que $f(x) = |x|$ n'est pas dérivable en 0 et que le m.b.s. B_t repasse régulièrement en 0 (alors que ça n'est pas le cas en dimension supérieure).

Posons donc

$$L_t = \lim_{\varepsilon \downarrow 0} \frac{1}{4\varepsilon} |\{s \in [0, t] : |B_s| < \varepsilon\}|,$$

où $|A|$ est la mesure (de Lebesgue) de l'ensemble de A . L_t représente en quelque sorte le temps passé par le m.b.s. B au point $x = 0$ pendant l'intervalle de temps $[0, t]$ (même si on pourrait penser a priori que ce temps est nul, ça n'est pas le cas !). On peut démontrer alors la formule étrange suivante :

$$|B_t| - |B_0| = \int_0^t \text{sgn}(B_s) dB_s + \frac{1}{2} L_t.$$

Une justification “à la main” de cette formule est que si $f(x) = |x|$, alors $f'(x) = \text{sgn}(x)$ et “ $f''(x) = \delta_0(x)$ ”, ainsi le terme d'Itô devient “ $\frac{1}{2} \int_0^t \delta_0(B_s) ds$ ”, qui représente bien la moitié du temps passé par B en $x = 0$ sur l'intervalle $[0, t]$.

Bibliographie

1) Rappels de probabilités

- * N. Bouleau, “Probabilités de l’ingénieur”, Hermann, 1986.
- * R. Durrett, “The essentials of probability”, The Duxbury Press, 1994.
- * S. M. Ross, “Initiation aux probabilités”, PPUR, 1987.
- * Virtual laboratories in probability and statistics : <http://www.math.uah.edu/stat/>.

2) Théorie des probabilités

- * P. Billingsley, “Probability and measure”, Wiley, 1995.
- * R. Durrett, “Probability : theory and examples”, The Duxbury Press, 1996.

3) Introduction au calcul stochastique et aux mathématiques financières

- * D. Lamberton, B. Lapeyre, “Introduction au calcul stochastique appliqué à la finance”, Ellipses, 1997. (existe aussi en anglais)
- * Th. Mikosch, “Elementary stochastic calculus with finance in view”, World Scientific, 1998.
- * B. Øksendal, “Stochastic differential equations. An introduction with applications” (corrected 5th ed.), Springer Verlag, 2000.

Ces trois ouvrages sont recommandés pour leur côté clair et introductif!

4) Calcul stochastique avancé

- * R. Durrett, “Stochastic calculus. A practical introduction”, CRC Press 1996.
- * I. Karatzas, S. E. Shreve, “Brownian motion and stochastic calculus”, Springer Verlag, 1991.
- * Ph. Protter, “Stochastic integration and differential equations. A new approach”, Springer Verlag, 1990.
- * D. Revuz, M. Yor, “Continuous martingales and Brownian motion”, Springer Verlag, 1999.

5) Simulation numérique des EDS

- * P. Kloeden, E. Platen, “Numerical solution of stochastic differential equations”, Springer Verlag, 1992. (un second tome existe aussi avec exercices corrigés)

Remarque. La présente liste n’est bien sûr de loin pas exhaustive !