

Processus Stochastiques II

Céline Esser et Yvik Swan

Année académique 2020–2021

Table des matières

1	Généralités sur les processus stochastiques	3
1.1	Exemples introductifs	3
1.1.1	Marches aléatoires	3
1.1.2	Problème de la ruine	4
1.1.3	Arbres de Galton-Watson	5
1.1.4	Chaînes de Markov	5
1.1.5	Processus de Poisson	7
1.1.6	Mouvement brownien	8
1.2	Définition et loi d'un processus stochastique	9
1.3	Égalités de processus	12
1.4	Théorème de consistance de Kolmogorov	14
2	Processus stochastiques à temps discret et martingales	16
2.1	Filtration et martingales	16
2.2	Transformées de martingales	19
2.3	Temps d'arrêt et martingales arrêtées	21
2.4	Inégalités de Doob	25
2.5	Convergence presque sûre des martingales	29
2.6	Convergence L^p , $p > 1$	32
2.7	Convergence L^1	32
3	Processus stochastiques à temps continu	36
3.1	Régularité des trajectoires	36
3.2	Processus gaussiens	41
3.3	Propriétés en loi de processus stochastiques	42
4	Définition et construction du mouvement Brownien	44
4.1	Introduction et motivation	44
4.2	Définition du mouvement Brownien	44

4.3	Construction du mouvement Brownien	46
4.3.1	Base de Haar et système de Faber-Schauder	46
4.3.2	Construction du mouvement Brownien	49
5	Propriétés en loi du mouvement Brownien	54
5.1	Mesure de Wiener	54
5.2	Propriétés immédiates	56
5.3	Propriété de Markov forte et temps d'atteinte	58
5.4	Théorème de Donsker	63
6	Martingales en temps continu	67
6.1	Définitions et propriétés	67
6.2	Martingales et Mouvement Brownien	70
7	Propriétés trajectorielles du mouvement Brownien	72
7.1	Temps d'atteinte	72
7.2	Zéros du mouvement Brownien	75
7.3	Non-dérivabilité	77
7.4	Variation non-bornée	78
8	L'intégrale d'Itô	82
8.1	Construction de l'intégrale d'Itô	83
8.2	Cas déterministe	87
8.3	Un premier exemple	89
8.4	L'intégrale d'Itô comme processus stochastique	90
8.5	Extension de la classe des intégrants	91
8.6	La formule d'Itô	92
A	Modes de convergence et le théorème central limite	94
B	Vecteurs gaussiens	97
C	Normes L^p et inégalités de Hölder	101
	References	104

Chapitre 1

Généralités sur les processus stochastiques

1.1 Exemples introductifs

1.1.1 Marches aléatoires

Une marche aléatoire est un processus qui décrit un chemin consistant en une succession d'étapes aléatoires sur un espace d'états donné. La marche aléatoire sur \mathbb{Z} permet de modéliser la marche d'un promeneur complètement ivre qui sort du bar et se déplace dans une rue droite (infinie) bordée de lampadaires à intervalles réguliers. A chaque lampadaire, il s'arrête pour se reposer. Comme il est ivre, lorsqu'il repart, il a oublié d'où il venait et choisit donc aléatoirement un sens ou l'autre. Retournera-t-il au bar ?

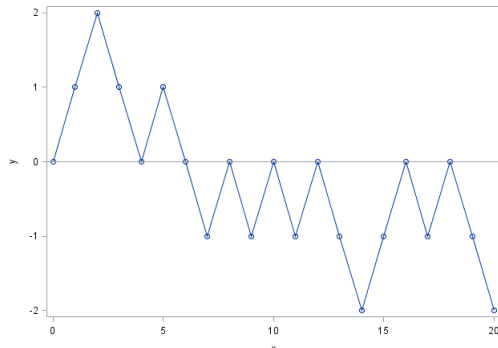


FIGURE 1.1 – Une réalisation de 20 pas de la marche du promeneur ivre sur une rue.

On peut maintenant imaginer que notre promeneur est perdu dans le quartier quadrillé de Manhattan (on suppose le quartier infini). A chaque croisement, il fait une pause et repart ensuite dans une des quatre directions au hasard, avec une chance que quatre pour chacune.

Si on souhaite modéliser ces deux situations sur l'espace \mathbb{Z}^d , on considère une suite $(Y_n)_{n \in \mathbb{N}}$ de variables aléatoires indépendantes et identiquement distribuées, pouvant prendre avec la même probabilité les $2d$ valeurs

$$(\pm 1, 0, \dots, 0), (0, \pm 1, 0, \dots, 0), \dots, (0, \dots, 0, \pm 1).$$

On pose $X_0 = 0$, c'est-à-dire on suppose que notre marcheur démarre à l'originie. Pour tout $n \geq 0$, on pose $X_{n+1} = X_n + Y_n$ qui représente la position au temps $n + 1$. Le déplacement entre X_n et X_{n+1} est donné par la valeur de Y_n .

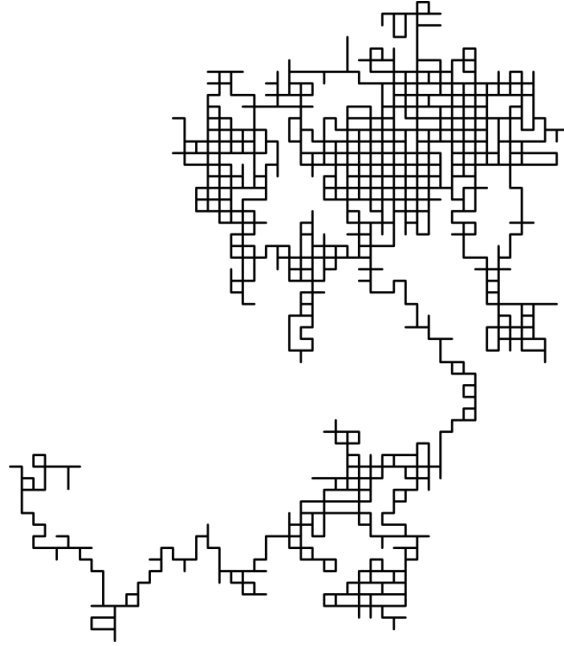


FIGURE 1.2 – Une réalisation de la marche du promeneur ivre dans Manhattan.

La marche aléatoire sur \mathbb{Z}^d a été beaucoup étudiée et est bien comprise. Par exemple, on sait que la marche revient presque sûrement en 0 si et seulement si $d = 1$ ou $d = 2$. Dans ce cas, on peut se demander le temps moyen pris par la marche pour revenir en 0. Les outils développés dans ce cours permettent d’aborder ce genre de questions.

1.1.2 Problème de la ruine

Le problème de la ruine du joueur est une variation du problème précédent, dans lequel on autorise la marche à être biaisée. Il fournit un modèle de l’évolution de la richesse d’une personne qui gagne sa vie en jetant une pièce truquée et en faisant des paris de 1 euro à chaque lancer. S’il obtient pile, ce qui arrive avec une probabilité p , la banque lui donne 1 euro. S’il obtient face, il donne 1 euro à la banque.

Afin de modéliser cette situation, on considère une suite $(Y_n)_{n \in \mathbb{N}}$ de variables aléatoires indépendantes et identiquement distribuées, dont la distribution de probabilité commune est donnée par

$$\mathbb{P}[Y_n = 1] = p \quad \text{et} \quad \mathbb{P}[Y_n = -1] = q = 1 - p.$$

Soit X_0 la fortune initiale de notre parieur et pour tout $n \geq 1$, notons X_n son argent après n lancers. Ainsi, on a

$$X_n = X_0 + \sum_{j=1}^n Y_j.$$

Notons que X_n peut prendre une valeur négative ou arbitrairement grande. Avant de commencer, le parieur peut se fixer un niveau A pour lequel il décide d’arrêter de jouer s’il a gagné A euros. Il s’arrête également s’il a perdu toute sa fortune. On considère donc le temps aléatoire

$$T = \min \{n \geq 0 : X_n \in \{0, A\}\}$$

auquel le parieur arrêtera de jouer. Peut-on affirmer que ce temps est presque sûrement fini ? Si oui, on a donc $X_T = 0$ ou $X_T = A$. Il est alors intéressant de pouvoir estimer par exemple la durée moyenne du jeu, c'est-à-dire $\mathbb{E}[T]$, ainsi que la probabilité de gagner, c'est-à-dire $\mathbb{P}[X_T = A]$.

1.1.3 Arbres de Galton-Watson

Le processus de Galton-Watson est le processus le plus simple pour décrire l'évolution d'une population. Il a été introduit pour étudier la probabilité d'extension d'un nom de famille dans un arbre généalogique.

Soit X_0 le nombre initial d'individus et X_n le nombre d'individus à l'instant n . On suppose que chaque individu donne naissance, juste avant de mourir, à un nombre aléatoire Z_{n+1}^k de descendants, pour $1 \leq k \leq X_n$, de façon indépendante des autres et du passé, selon une loi de reproduction identique. On suppose donc que

$$X_{n+1} = \sum_{k=1}^{X_n} Z_{n+1}^k$$

si $X_n \geq 1$, et $X_{n+1} = 0$ si $X_n = 0$, où les variables $(Z_{n+1}^k)_{n,k}$ sont indépendantes et identiquement distribuées.

On montrera que si l'espérance commune des variables Z_{n+1}^k est strictement inférieure à 1, alors la population s'éteint presque sûrement.

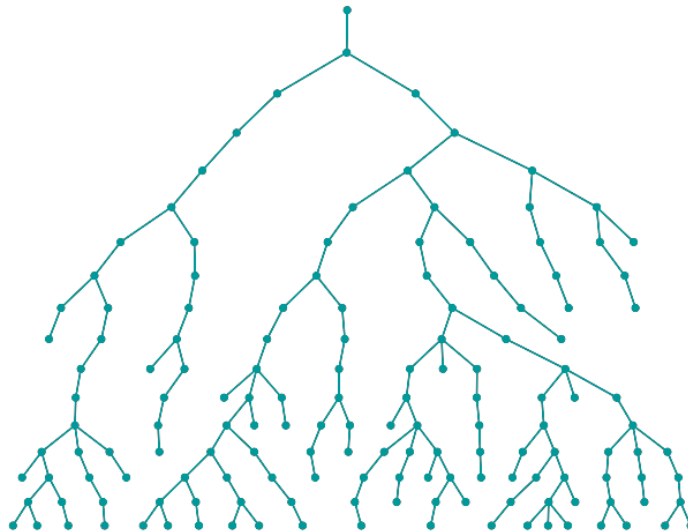


FIGURE 1.3 – Une réalisation du processus de Galton-Watson, où $X_0 = 1$. Le temps s'écoule de haut en bas et la distribution du nombre de descendants est une loi binomiale $\text{Bin}(3, 0.4)$.

1.1.4 Chaînes de Markov

Les trois exemples précédents partagent une propriété importante : Leur évolution future ne dépend que de la valeur présente, indépendamment du passé. Autrement dit, on a

$$\mathbb{P}[X_{n+1} = x \mid X_n = x_n, \dots, X_0 = x_0] = \mathbb{P}[X_{n+1} = x \mid X_n = x_n]$$

quels que soient n, x_0, \dots, x_{n+1} . Cette dernière propriété s'appelle la *propriété de Markov*. De manière générale, on peut définir une *chaîne de Markov* de la manière suivante : Une suite $(X_n)_{n \geq 0}$ de variables aléatoires à valeurs dans un espace discret $S = \{s_j : j \in E\}$ est une *chaîne de Markov* si pour tout $j \in E$

$$\mathbb{P}[X_{n+1} = s_j \mid X_0, X_1, \dots, X_n] = \mathbb{P}[X_{n+1} = s_j \mid X_n] \quad \forall n \geq 0.$$

La chaîne de Markov est *homogène* si ces probabilités ne dépendent pas de n , c'est-à-dire si

$$\mathbb{P}[X_{n+1} = s_j \mid X_n = s_i] = \mathbb{P}[X_1 = s_j \mid X_0 = s_i] = p_{ij} \quad \forall i, j \in E.$$

Les états et les probabilités de transitions peuvent donc être représentés dans un graphe (pas nécessairement connexe et potentiellement infini) appelé *le graphe de transition* :

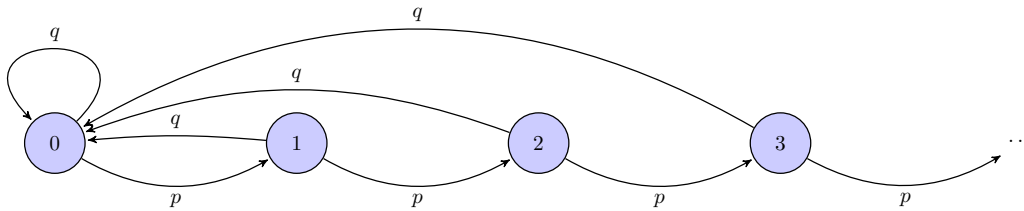
- les sommets du graphe sont les points de S ;
- on met une arête orientée $(s_i \rightarrow s_j)$ du sommet s_i vers le sommet s_j si $p_{ij} > 0$, et on indique la valeur de p_{ij} sur cette arête.

Les probabilités de transition $\{p_{ij}, i, j \in E\}$ sont rassemblées dans une matrice de dimension $\dim E \times \dim E$ (potentiellement infinie)

$$\mathbf{P} = \begin{pmatrix} p_{00} & p_{01} & p_{02} & \dots \\ p_{10} & p_{11} & p_{12} & \dots \\ p_{30} & p_{31} & p_{32} & \dots \\ \vdots & \vdots & \vdots & \ddots \end{pmatrix}$$

appelée la *matrice de transition* de la chaîne. La distribution de X_0 est appelée la *distribution initiale*.

Par exemple, considérons un processus à valeurs dans \mathbb{N} de graphe de transition



avec $q = 1 - p$, c'est-à-dire de matrice de transition

$$\mathbf{P} = \begin{pmatrix} q & p & 0 & 0 & \dots \\ q & 0 & p & 0 & \dots \\ q & 0 & 0 & p & \dots \\ \vdots & \vdots & \vdots & \vdots & \ddots \end{pmatrix}.$$

On considère la distribution initiale triviale $\mathbf{u} = (1, 0, 0, \dots)$. Ce processus compte le nombre de succès consécutifs lors d'une répétition d'une expérience de Bernoulli de probabilité de succès p .

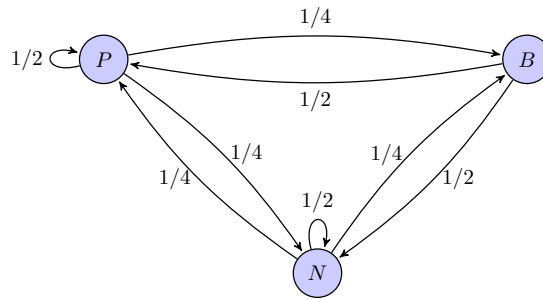
Présentons à présent un deuxième exemple à espace d'états finis. D'après Kemeny, Snell et Thompson¹, le pays d'Oz a de nombreux atouts, mais le beau temps n'en fait pas partie ! Il n'y

1. J. G. Kemeny, J. L. Snell, G. L. Thompson, *Introduction to Finite Mathematics*, 3rd ed. (Englewood Cliffs, NJ : Prentice-Hall, 1974).

a jamais deux belles journées d'affilée. Si une journée est belle, on risque tout autant d'avoir de la neige que de la pluie le lendemain. S'il y a de la neige ou de la pluie, on a une chance sur deux d'avoir la même chose le lendemain ; s'il y a du changement par rapport à la neige ou à la pluie, la moitié du temps on a une belle journée le lendemain. Cela décrit une chaîne de Markov de matrice de transition

$$\mathbf{P} = \begin{matrix} & \begin{matrix} P & B & N \end{matrix} \\ \begin{matrix} P \\ B \\ N \end{matrix} & \begin{pmatrix} 1/2 & 1/4 & 1/4 \\ 1/2 & 0 & 1/2 \\ 1/4 & 1/4 & 1/2 \end{pmatrix} \end{matrix}$$

S'il fait beau aujourd'hui, alors la distribution initiale est $\mathbf{u} = (0 \ 1 \ 0)$. Le graphe de transition de cette chaîne de Markov est donné par



Les chaînes de Markov jouent un grand rôle à la fois dans la théorie et dans les applications : Leur structure est suffisamment riche pour modéliser de nombreux phénomènes aléatoires et suffisamment simple pour permettre de répondre à de multiples questions sur le comportement de ces phénomènes. Elles ne font pas l'objet de ce cours, mais le comportement Markovien interviendra lors de l'étude du mouvement brownien.

1.1.5 Processus de Poisson

Les processus présentés précédemment étaient formés d'une suite de variables aléatoires définies sur le même espace probabilisé. On parle de processus stochastiques à temps discret. On pourrait également considérer une famille de variables aléatoires définies sur le même espace probabilisé et indexée par un ensemble continu. On parle de processus stochastiques à temps continu.

Le processus de Poisson est un exemple d'un tel processus. Il permet de modéliser des phénomènes d'arrivées aléatoires isolés dans le temps tels que des appels téléphoniques reçus par une centrale, des arrivées de clients devant un guichet, des visites sur une page Internet,... Une méthode de construction du processus de Poisson est la suivante : On considère une suite $(W_n)_{n \in \mathbb{N}_0}$ de variables aléatoires indépendantes et identiquement distribuées selon une loi exponentielle de paramètre $\lambda > 0$. On pose

$$T_n = \sum_{i=1}^n W_i$$

et

$$N(t) = \sup\{n \in \mathbb{N} : T_n \leq t\}.$$

Alors $\{N(t), t \geq 0\}$ est appelé un *processus de Poisson* de taux λ . Les temps $0 = T_0 \leq T_1 \leq$

$T_2 \leq \dots \leq T_j \leq \dots$ sont les *instants d'arrivées*. On a

$$N(t) = j \iff T_j \leq t < T_{j+1}.$$

Les *temps inter-arrivées* (temps entre deux arrivées successives) sont donnés par les variables $W_i, i \geq 1$.

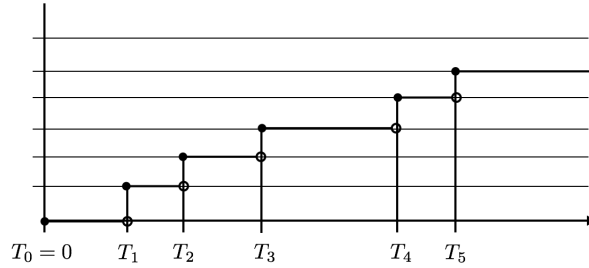


FIGURE 1.4 – Trajectoire d'un processus de Poisson et instants d'arrivées.

Ce processus est un exemple fondamental de processus Markoviens qui donnent une version continue des chaînes de Markov : On a

$$\mathbb{P}big[N(t+u) = j \mid N(v), 0 \leq v \leq t] = \mathbb{P}[N(t+u) = j \mid N(t)]$$

pour tous $t, u \geq 0$, et tout $j \in \mathbb{N}$.

Ce processus possède d'autres propriétés remarquables et classiques pour des processus stochastiques. Par exemple, il s'agit d'un processus à *accroissements indépendants*, c'est-à-dire pour tout $p \geq 2$ et tous $0 \leq t_1 < t_2 < \dots < t_p$, les variables aléatoires

$$N(t_2) - N(t_1), \dots, N(t_p) - N(t_{p-1})$$

sont indépendantes. Les accroissements du processus de Poisson sont également *stationnaires*, c'est-à-dire pour tout $h > 0$ et tout $t \geq 0$,

$$N(t+h) - N(t) \stackrel{\mathcal{L}}{=} N(h) - N(0).$$

On a en fait $N(t+h) - N(t) \sim \text{Pois}(\lambda h)$. Notons finalement que les réalisations de ce processus ne sont pas continues puisqu'il s'agit d'un processus croissant à valeurs dans \mathbb{N} .

1.1.6 Mouvement brownien

Si on change convenablement l'échelle de temps et l'échelle d'espace d'une marche aléatoire dans \mathbb{Z} , en effectuant un zoom de plus en plus précis, on obtient un processus stochastique continu qui évolue continuellement dans le temps et qui admet une distribution unique. Ce processus limite s'appelle le mouvement brownien. L'étude du mouvement brownien sera un des objectifs principaux de ce cours. Notons que cet objet mathématique permet de modéliser des phénomènes naturels et a été étudié initialement pour expliquer le mouvement irrégulier et aléatoire d'une particule de pollen en suspension dans l'eau.

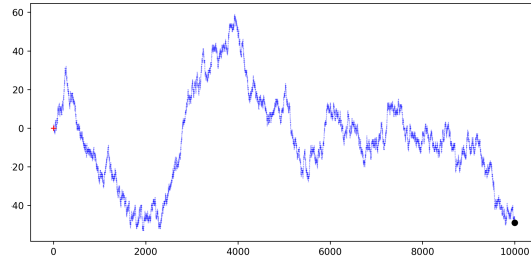


FIGURE 1.5 – Réalisation d'une marche aléatoire.

1.2 Définition et loi d'un processus stochastique

Les processus stochastiques permettent de modéliser des quantités qui évoluent aléatoirement au cours du temps. Nous désignons par $(\Omega, \mathcal{F}, \mathbb{P})$ un espace probabilisé et par I un ensemble quelconque d'indices. L'indice $t \in I$ réfère donc en général à une position dans le temps.

Définition 1.2.1 Un *processus stochastique* $X = \{X(t) : t \in I\}$ sur $(\Omega, \mathcal{F}, \mathbb{P})$ est une famille de variables aléatoires définies sur $(\Omega, \mathcal{F}, \mathbb{P})$ à valeurs réelles indexée par I .

Si l'ensemble I est fini, le processus est un vecteur aléatoire. Si I est dénombrable, le processus est une suite de variables aléatoires et on parle de processus à *temps discret*. Si $I \subseteq \mathbb{R}$ est un intervalle, on dit que le processus est à *temps continu*. Mentionons également que si $I \subseteq \mathbb{R}^n$ est un intervalle, on parle de *champ aléatoire*. Dans ce cours, on s'intéressera en général aux processus indexés par $I = \mathbb{N}$ ou $I = [0, +\infty[$.

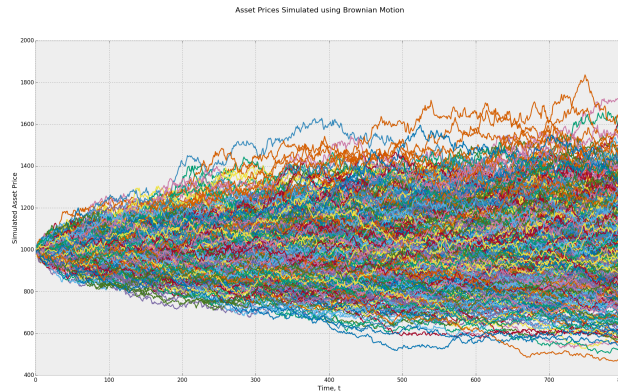


FIGURE 1.6 – Différentes trajectoires d'un processus, chaque ligne représente une réalisation particulière du processus, i.e. une trajectoire $t \mapsto X(t, \omega)$; à chaque instant t , on a des réalisations de la variable aléatoire réelle $X(t, \cdot)$

Un processus dépend de deux paramètres : $t \in I$ et $\omega \in \Omega$; une illustration est donnée à la Figure 1.6.

- Pour tout $t \in I$ fixé, $X(t, \cdot) : \omega \in \Omega \mapsto X(t, \omega)$ est une variable aléatoire. Afin d'alléger les notations, nous noterons souvent cette variable X_t .
- Pour tout $\omega \in \Omega$ fixé, $X(\cdot, \omega) : t \in I \mapsto X(t, \omega)$ est une fonction à valeurs réelles, appelée *trajectoire du processus*. L'étude des propriétés des trajectoires d'un processus est un des

objectifs de la théorie des processus stochastiques.

En regardant un processus stochastique comme une application à valeurs dans \mathbb{R}^I , à savoir

$$X : \Omega \rightarrow \mathbb{R}^I : \omega \mapsto X(\cdot, \omega),$$

il est naturel de vouloir définir sa loi. Pour cela, nous devons introduire quelques résultats sur les σ -algèbres produits. Considérons d'abord le cas où $I = \{1, \dots, n\}$. Rappelons que le produit $\mathbb{R}^n = \mathbb{R} \times \dots \times \mathbb{R}$ peut être muni de la σ -algèbre engendrée par les rectangles

$$B_1 \times \dots \times B_n$$

où $B_k \in \mathbb{B}$ pour tout $k \in I$. C'est la plus petite σ -algèbre rendant les projections

$$\pi_k : \mathbb{R}^n \rightarrow \mathbb{R} : (x_1, \dots, x_n) \mapsto x_k$$

Borel-mesurables. Dans le cas des produits infinis, on a considéré les cylindres à la place des rectangles.

Définition 1.2.2 On appelle *cylindre* tout ensemble de la forme

$$\prod_{t \in I} B_t \quad \text{où } B_t \in \mathbb{B} \text{ et } B_t \neq \mathbb{R} \text{ pour un nombre fini d'indices.}$$

La σ -algèbre produit sur \mathbb{R}^I est la σ -algèbre engendrée par les cylindres. On la note \mathbb{B}^I .

Le résultat suivant permet de justifier la définition ci-dessus.

Proposition 1.2.3 La σ -algèbre produit sur \mathbb{R}^I est la plus petite σ -algèbre rendant les projections

$$\pi_{t_0} : (\mathbb{R}^I, \mathbb{B}^I) \rightarrow (\mathbb{R}, \mathbb{B}) : (x_t)_{t \in I} \mapsto x_{t_0}, \quad t_0 \in I$$

mesurables.

Démonstration : Si $B \in \mathbb{B}$, on a

$$\pi_{t_0}^{-1}(B) = \prod_{t \in I} A_t \quad \text{où } A_u = \begin{cases} B & \text{si } t = t_0, \\ \mathbb{R} & \text{sinon.} \end{cases}$$

En particulier, $\pi_{t_0}^{-1}(B)$ est un cylindre et appartient donc à \mathbb{B}^I . Ainsi, \mathbb{B}^I rend bien les projections mesurables. De plus, si \mathcal{A} est une σ -algèbre sur \mathbb{R}^I qui rend les projections mesurables, alors pour tout cylindre $C = \prod_{t \in I} B_t$ où $B_t \neq \mathbb{R}$ si $t \in \{t_1, \dots, t_n\}$, on a

$$C = \bigcap_{k=1}^n \pi_{t_k}^{-1}(B_{t_k}) \in \mathcal{A}.$$

Ainsi \mathcal{A} contient donc tous les cylindres, ce qui suffit. ■

Proposition 1.2.4 Si $X = \{X(t) : t \in I\}$ est un processus stochastique sur $(\Omega, \mathcal{F}, \mathbb{P})$, alors l'application

$$X : (\Omega, \mathcal{F}) \rightarrow (\mathbb{R}^I, \mathbb{B}^I) : \omega \mapsto X(\cdot, \omega)$$

est mesurable.

Démonstration : Soit un cylindre $C = \prod_{t \in I} B_t$ où $B_t \neq \mathbb{R}$ si $t \in \{t_1, \dots, t_n\}$. Il suffit de montrer que $X^{-1}(\prod_{t \in I} B_t) \in \mathcal{F}$. Or, on sait que

$$X^{-1}(C) = \bigcap_{k=1}^n X_{t_k}^{-1}(B_{t_k})$$

et puisque chaque X_{t_k} est une variable aléatoire, $X_{t_k}^{-1}(B_{t_k}) \in \mathcal{F}$. Cela permet de conclure. ■

On en tire que tout processus stochastique définit une mesure de probabilité sur l'espace mesurable $(\mathbb{R}^I, \mathbb{B}^I)$.

Définition 1.2.5 Soit $X = \{X(t) : t \in I\}$ un processus stochastique sur $(\Omega, \mathcal{F}, \mathbb{P})$. On note \mathbb{P}_X la mesure image de \mathbb{P} par X , c'est-à-dire la mesure de probabilité définie sur $(\mathbb{R}^I, \mathbb{B}^I)$ par

$$\mathbb{P}_X[B] = \mathbb{P}[X^{-1}(B)] \quad \forall B \in \mathbb{B}^I.$$

On appelle \mathbb{P}_X la *loi de probabilité* du processus stochastique X .

Remarque 1.2.6 On peut dès lors se demander si toute mesure de probabilité sur $(\mathbb{R}^I, \mathbb{B}^I)$ provient d'un processus stochastique. Plus précisément, étant donnée une mesure de probabilité $\tilde{\mathbb{P}}$ sur $(\mathbb{R}^I, \mathbb{B}^I)$, existe-t-il un espace probabilisé $(\Omega, \mathcal{F}, \mathbb{P})$ et un processus X défini sur cet espace tel que $\mathbb{P}_X = \tilde{\mathbb{P}}$? On peut répondre à cette question par l'affirmative en considérant l'espace probabilisé $(\mathbb{R}^I, \mathbb{B}^I, \tilde{\mathbb{P}})$ et le processus X défini par

$$X(t, \omega) = \omega(t)$$

pour tous $t \in I$, $\omega \in \mathbb{R}^I$. En effet, pour tout $B \in \mathbb{B}^I$, on a

$$\tilde{\mathbb{P}}_X[B] = \tilde{\mathbb{P}}[\{\omega \in \mathbb{R}^I : X(\cdot, \omega) \in B\}] = \tilde{\mathbb{P}}[\{\omega \in \mathbb{R}^I : \omega \in B\}] = \tilde{\mathbb{P}}[B].$$

Définition 1.2.7 Les *lois fini-dimensionnelles* d'un processus $X = \{X(t) : t \in I\}$ sont données par l'ensemble des lois des vecteurs aléatoires

$$(X_{t_1}, \dots, X_{t_n}), \quad n \in \mathbb{N}_0, t_1, \dots, t_n \in I.$$

Les distributions fini-dimensionnelles de X sont donc caractérisées par la famille de fonctions de répartition $\{F_{t_1, \dots, t_n}^X : n \in \mathbb{N}_0, t_1, \dots, t_n \in I\}$ donnée par

$$F_{t_1, \dots, t_n}^X(x_1, \dots, x_n) = \mathbb{P}[X_{t_1} \leq x_1, \dots, X_{t_n} \leq x_n]$$

pour tous $n \in \mathbb{N}_0$, $t_1, \dots, t_n \in I$ et $(x_1, \dots, x_n) \in \mathbb{R}^n$.

En particulier, si $|I| < \infty$, les lois fini-dimensionnelles sont les lois de toutes les projections de la loi du vecteur $X = (X_1, \dots, X_{|I|})$.

Définition 1.2.8 Soient $X = \{X(t) : t \in I\}$ et $Y = \{Y(t) : t \in I\}$ deux processus stochastiques définis respectivement sur les espaces probabilisés $(\Omega, \mathcal{F}, \mathbb{P})$ et $(\tilde{\Omega}, \tilde{\mathcal{F}}, \tilde{\mathbb{P}})$. On dit que X et Y possèdent *les mêmes lois fini-dimensionnelles* (on dit aussi qu'ils sont égaux au sens des lois fini-dimensionnelles) si pour tout $n \in \mathbb{N}_0$ et tous $t_1, \dots, t_n \in I$, les vecteurs aléatoires $(X_{t_1}, \dots, X_{t_n})$ et $(Y_{t_1}, \dots, Y_{t_n})$ sont de mêmes lois.

Autrement dit, on demande que

$$\mathbb{P}\left[(X_{t_1}, \dots, X_{t_n}) \in B\right] = \tilde{\mathbb{P}}\left[(Y_{t_1}, \dots, Y_{t_n}) \in B\right]$$

pour tous $n \in \mathbb{N}_0$, $t_1, \dots, t_n \in I$ et $B \in \mathbb{B}^n$.

Proposition 1.2.9 La loi d'un processus stochastique X est entièrement caractérisée par ses lois fini-dimensionnelles.

Démonstration : Considérons deux processus stochastiques $X = \{X(t) : t \in I\}$ et $Y = \{Y(t) : t \in I\}$ définis respectivement sur les espaces probabilisés $(\Omega, \mathcal{F}, \mathbb{P})$ et $(\tilde{\Omega}, \tilde{\mathcal{F}}, \tilde{\mathbb{P}})$, et possédant les mêmes lois fini-dimensionnelles. Nous allons montrer que $\mathbb{P}_X = \tilde{\mathbb{P}}_Y$. Soit $C = \prod_{t \in I} B_t$ un cylindre de \mathbb{R}^I , où $B_t \neq \mathbb{R}$ si $t \in \{t_1, \dots, t_n\}$. Alors, par hypothèse, on a

$$\mathbb{P}_X[C] = \mathbb{P}\left[X_{t_1} \in B_{t_1}, \dots, X_{t_n} \in B_{t_n}\right] = \tilde{\mathbb{P}}\left[Y_{t_1} \in B_{t_1}, \dots, Y_{t_n} \in B_{t_n}\right] = \tilde{\mathbb{P}}_Y[C].$$

Puisque l'ensemble des cylindres de \mathbb{R}^I est un π -système (i.e. stable par intersection finie) qui engendre \mathbb{B}^I , le lemme de la classe monotone permet d'affirmer que $\mathbb{P}_X = \tilde{\mathbb{P}}_Y$. ■

1.3 Égalités de processus

Nous venons d'étudier l'égalité de processus stochastiques au sens des lois. On peut introduire d'autres façon pour des processus d'être égaux.

Définition 1.3.1 Soient $X = \{X(t) : t \in I\}$ et $Y = \{Y(t) : t \in I\}$ deux processus stochastiques.

1. On dit que X et Y ont la *même loi* s'ils ont les mêmes lois fini-dimensionnelles, c'est-à-dire pour tous $n \in \mathbb{N}_0$ et $t_1, \dots, t_n \in I$, on a

$$F_{t_1, \dots, t_n}^X = F_{t_1, \dots, t_n}^Y.$$

On écrit $X \stackrel{\mathcal{L}}{=} Y$.

2. On dit que X et Y ont sont des *versions* ou *modifications* l'un de l'autre s'ils sont définis sur le même espace probabilisé et si

$$\mathbb{P}[X(t) = Y(t)] = 1, \quad \forall t \in I.$$

3. On dit que X et Y sont *indistinguables* s'ils sont définis sur le même espace probabilisé $(\Omega, \mathcal{F}, \mathbb{P})$ et s'il existe un ensemble $N \subseteq \Omega$ qui est \mathbb{P} -négligeable et tel que pour tout $\omega \in \Omega \setminus N$

$$X(t, \omega) = Y(t, \omega), \quad \forall t \in I. \quad a$$

a. De manière un peu abusive (parce que $\{X(t) = Y(t), \forall t \in I\}$ n'est pas nécessairement un événement), on écrit souvent $\mathbb{P}[X(t) = Y(t), \forall t \in I] = 1$.

Remarque 1.3.2 Les processus X et Y sont des versions l'un de l'autre si pour tout $t \in I$, presque sûrement, $X(t) = Y(t)$: Autrement dit, pour tout $t \in I$, il existe un ensemble $N_t \subseteq \Omega$ \mathbb{P} -négligeable tel que $X(t, \omega) = Y(t, \omega)$ pour tout $\omega \in \Omega \setminus N_t$. Pour obtenir que X et Y sont indistinguables, il faut que presque sûrement, pour tout $t \in I$, $X(t) = Y(t)$: L'ensemble \mathbb{P} -négligeable qui intervient ici ne dépend donc pas de $t \in I$.

Il est facile de voir que les notions d'égalité de processus stochastiques s'ordonnent de la façon suivante :

$$\text{indistinguables} \implies \text{versions} \implies \text{même loi.}$$

Si I est un ensemble dénombrable, deux processus sont indistinguables si et seulement s'ils sont des versions l'un de l'autre. En effet, si X et Y sont des modifications l'un de l'autre, on a

$$\mathbb{P}(\exists t \in I : X(t) \neq Y(t)) = \mathbb{P}\left(\bigcup_{t \in I} \{X(t) \neq Y(t)\}\right) \leq \sum_{t \in I} \mathbb{P}(X(t) \neq Y(t)) = 0$$

d'où $\mathbb{P}(X(t) = Y(t), \forall t \in I) = 1$.

En revanche, lorsque I est non-dénombrable, deux processus peuvent être des versions l'un de l'autre sans pour autant être indistinguables. En fait, toutes les implications données précédemment sont strictes comme illustré dans les exemples ci-dessous.

Exemple 1.3.3 Soit $Z \sim \mathcal{N}(0, 1)$. Posons

$$X(t) = Z \text{ et } Y(t) = -Z$$

pour tout $t \in I$. Alors X et Y ont la même loi, mais

$$\mathbb{P}(X(t) = Y(t)) = \mathbb{P}(2Z = 0) = 0, \quad \forall t \in I.$$

Par conséquent, X et Y ne sont pas des versions l'un de l'autre.

Exemple 1.3.4 Considérons l'espace probabilisé $([0, 1], \mathbb{B}([0, 1]), \lambda)$ et $I = [0, 1]$. Notons $D = \{(x, x) : x \in [0, 1]\}$ la diagonale de $[0, 1] \times [0, 1]$. Posons

$$X(t, \omega) = 0 \text{ et } Y(t, \omega) = \mathbf{1}_D(t, \omega)$$

pour tous $t \in I, \omega \in [0, 1]$. Pour tout $t \in I$, on a

$$\lambda(\{\omega \in [0, 1] : X(t, \omega) = Y(t, \omega)\}) = \lambda([0, 1] \setminus \{t\}) = 1.$$

Par conséquent, X et Y sont des versions l'un de l'autre. Par contre, X et Y ne sont pas indistinguables puisque

$$\lambda(\{\omega \in [0, 1] : X(t, \omega) = Y(t, \omega), \forall t \in I\}) = 0.$$

Dans l'exemple ci-dessus, on observe que les trajectoires de X sont continues alors que celles de Y ne le sont pas. En fait, c'est ce qu'il manque pour obtenir l'implication inverse.

Proposition 1.3.5 *Supposons que $I \subseteq \mathbb{R}$ est un intervalle et que les processus X et Y sont des versions l'un de l'autre. Si les trajectoires de X et Y sont presque sûrement continues à droite (ou à gauche), alors X et Y sont indistinguables.*

Démonstration : Notons Ω_1 l'événement de probabilité 1 sur lequel les trajectoires de X et Y sont continues à droite. Soit D une partie dénombrable dense de I . Pour tout $t \in D$, on a

$$\mathbb{P}(X(t) = Y(t)) = 1$$

et par conséquent, puisque D est dénombrable, on a

$$\mathbb{P}\left(\bigcap_{t \in D} \{X(t) = Y(t)\}\right) = 1.$$

Notons Ω_2 l'événement de probabilité 1 donné ci-dessus. Si $t \notin D$, il existe une suite décroissante $(t_n)_{n \in \mathbb{N}}$ de D telle que $t_n \rightarrow t$. En utilisant la continuité à droite des trajectoires, on a $X(t_n, \omega) \rightarrow X(t, \omega)$ et $Y(t_n, \omega) \rightarrow Y(t, \omega)$ pour tout $\omega \in \Omega_1$. On en tire que pour tout $\omega \in \Omega_1 \cap \Omega_2$, $X(t, \omega) = Y(t, \omega)$, d'où la conclusion puisque $\Omega_1 \cap \Omega_2$ est un événement de probabilité 1. ■

1.4 Théorème de consistance de Kolmogorov

Il y a plusieurs façons de “spécifier” un processus stochastique. Une première approche consiste à le construire explicitement, c'est-à-dire l'écrire en termes de variables aléatoires plus simples. C'est ce que nous avons fait dans la Section 1.1.

Une deuxième façon de spécifier un processus stochastique consiste à imposer ses distributions fini-dimensionnelles (ce qui revient bien à spécifier sa loi, grâce à la Proposition 1.2.9). Dans ce cas, la question suivante doit être résolue : Étant données des lois fini-dimensionnelles, existe-t-il un espace probabilisé et un processus sur cet espace ayant ces lois fini-dimensionnelles ? Pour se fixer les idées, rappelons ce qui se passe pour une loi sur \mathbb{R} . Se donner une loi de probabilité revient à se donner une fonction de répartition. Ainsi la question devient : Étant donnée une fonction $F : \mathbb{R} \rightarrow [0, 1]$ qui satisfait aux contraintes (i) F est croissante, (ii) F est continue à droite, (iii) $\lim_{x \rightarrow -\infty} F(x) = 0$ et $\lim_{x \rightarrow +\infty} F(x) = 1$, existe-t-il un espace probabilisé $(\Omega, \mathcal{F}, \mathbb{P})$ et une variable aléatoire $X : \Omega \rightarrow \mathbb{R}$ tels que $\mathbb{P}[X \leq x] = F(x)$? Ce cas simple peut se traiter via le lemme suivant.

Lemme 1.4.1 *Soit F une fonction satisfaisant les propriétés (i) - (iii) et soit U une variable aléatoire uniforme sur $[0, 1]$. Définissons $F^{-1}(x) := \inf \{t : F(t) \geq x\}$ l'inverse généralisé de F . Alors la variable $F^{-1}(U)$ est de loi F .*

Démonstration : Le cas où G est strictement croissante sur \mathbb{R} est facile ; pour le cas général voir [9]. ■

Dans notre cas, nous considérons pour tout $n \in \mathbb{N}_0$ et tous $t_1, \dots, t_n \in I$, une mesure de probabilité $\mathbb{P}_{t_1, \dots, t_n}$ définie sur \mathbb{R}^n . Afin de pouvoir être des candidats pour être les lois fini-dimensionnelles d'un processus, ces mesures de probabilité doivent naturellement vérifier les hypothèses suivantes : Pour tous $n \in \mathbb{N}_0$, $t_1, \dots, t_n \in I$,

— si π est une permutation de $\{1, \dots, n\}$, alors

$$\mathbb{P}_{t_1, \dots, t_n} [A_1 \times \dots \times A_n] = \mathbb{P}_{\pi(t_1), \dots, \pi(t_n)} [A_{\pi(t_1)} \times \dots \times A_{\pi(t_n)}]$$

pour tous $A_1, \dots, A_n \in \mathcal{F}$,

— si $t_{n+1} \in I$, alors

$$\mathbb{P}_{t_1, \dots, t_{n+1}} [A \times \mathbb{R}] = \mathbb{P}_{t_1, \dots, t_n} [A]$$

Ces deux hypothèses sont appelées les *conditions de consistance* et mènent au résultat suivant, admis sans preuve.

Théorème 1.4.2 (Théorème d'existence de Kolmogorov) *Considérons pour tout $n \in \mathbb{N}_0$ et tous $t_1, \dots, t_n \in I$ une mesure de probabilité $\mathbb{P}_{t_1, \dots, t_n}$ définie sur \mathbb{R}^n . Supposons que la famille $\mathcal{Q} = \{\mathbb{P}_{t_1, \dots, t_n} : n \in \mathbb{N}_0, t_1, \dots, t_n \in I\}$ satisfait les conditions de consistances. Alors il existe un espace probabilisé $(\Omega, \mathcal{F}, \mathbb{P})$ et un processus stochastique X sur $(\Omega, \mathcal{F}, \mathbb{P})$ qui admet \mathcal{Q} comme famille de lois fini-dimensionnelles.*

Le problème du théorème d'existence de Kolmogorov est que l'espace Ω construit est beaucoup trop "gros" et bien souvent on a besoin de processus qui ont certaines propriétés (continuité, différentiabilité, ...) qui peuvent vivre sur des espaces beaucoup plus petits.

Chapitre 2

Processus stochastiques à temps discret et martingales

La notion de martingale est fondamentale dans le calcul des probabilités. Ces processus permettent de modéliser des jeux équitables, c'est-à-dire des jeux pour lesquels l'espérance de gain à un temps donné est égal au gain déjà accumulé. On peut garder en tête par exemple la marche aléatoire sur \mathbb{Z} . Les martingales ont de nombreuses propriétés qui les rendent très utiles dans l'étude de processus plus généraux.

2.1 Filtration et martingales

Fixons un espace mesurable (Ω, \mathcal{F}) . Une *sous- σ -algèbre* de \mathcal{F} est une sous-famille $\mathcal{F}_1 \subseteq \mathcal{F}$ qui est également une σ -algèbre.

Définition 2.1.1 Une suite de sous- σ -algèbres $(\mathcal{F}_n)_{n \in \mathbb{N}_0}$ de \mathcal{F} est une *filtration* si $\mathcal{F}_n \subseteq \mathcal{F}_{n+1}$ pour tout $n \in \mathbb{N}_0$.

Intruitivement, \mathcal{F}_n contient les événements qui peuvent survenir jusqu'à l'instant n . Il faut donc comprendre \mathcal{F}_n comme l'information dont on dispose au temps n : Plus le temps passe, plus nous disposons d'information. Par exemple, considérons un jeu de hasard qui est répété un grand nombre de fois : \mathcal{F}_n est l'information dont nous disposons après n parties et la suite de sous- σ -algèbres est bien croissante si nous gagnons (ou à tout le moins ne perdons pas) d'informations avec le temps.

Définition 2.1.2 Soient $(\Omega, \mathcal{F}, \mathbb{P})$ un espace probabilisé et $(\mathcal{F}_n)_{n \in \mathbb{N}_0}$ une filtration. Une suite de variables aléatoires $(Z_n)_{n \in \mathbb{N}_0}$ est *adaptée à la filtration* $(\mathcal{F}_n)_{n \in \mathbb{N}_0}$ si Z_n est \mathcal{F}_n -mesurable pour tout $n \in \mathbb{N}_0$.

On a typiquement $\mathcal{F}_n = \sigma(W_1, \dots, W_n)$ où les W_i sont des variables aléatoires, et la suite $(Z_n)_{n \in \mathbb{N}_0}$ adaptée à la filtration $(\mathcal{F}_n)_{n \in \mathbb{N}_0}$ est de la forme

$$Z_n = f(W_1, \dots, W_n)$$

pour une fonction f Borel-mesurable.

Introduisons à présent la notion de martingale. Une martingale peut se voir comme la généralisation du jeu équitable (i.e. un pari dont l'espérance de gain est nulle) : $(Z_n)_n$ est une martingale pour $(\mathcal{F}_n)_n$ si quels que soient les résultats des n premières parties, la fortune attendue après le $n+1^{\text{ème}}$ jeu est exactement ce qu'elle était avant. En d'autres termes, quelle que soit l'historique du jeu le gain du joueur est, en moyenne, nul.

Dans la suite, nous considérons un espace probabilisé $(\Omega, \mathcal{F}, \mathbb{P})$ fixé.

Définition 2.1.3 Une suite $(Z_n)_{n \in \mathbb{N}_0}$ de variables aléatoires est une *martingale* pour la filtration $(\mathcal{F}_n)_{n \in \mathbb{N}_0}$ si

1. $(Z_n)_{n \in \mathbb{N}_0}$ est adapté à $(\mathcal{F}_n)_{n \in \mathbb{N}_0}$,
2. $\mathbb{E}[|Z_n|] < +\infty$ pour tout $n \in \mathbb{N}_0$,
3. $\mathbb{E}[Z_{n+1} | \mathcal{F}_n] = Z_n$ pour tout $n \in \mathbb{N}_0$.

En prenant l'espérance de part et d'autre de la dernière condition de la Définition 2.1.3, on déduit

$$\mathbb{E}[Z_{n+1}] = \mathbb{E}[Z_n]$$

et donc, en particulier, $\mathbb{E}[Z_n] = \mathbb{E}[Z_1]$ pour tout $n \in \mathbb{N}_0$. On appelle $\mathbb{E}[Z_1]$ la *moyenne de la martingale*. Notons qu'on a également

$$\begin{aligned} \mathbb{E}[Z_{n+2} | \mathcal{F}_n] &= \mathbb{E}[\mathbb{E}[Z_{n+2} | \mathcal{F}_{n+1} \cup \mathcal{F}_n] | \mathcal{F}_n] \\ &= \mathbb{E}[\mathbb{E}[Z_{n+2} | \mathcal{F}_{n+1}] | \mathcal{F}_n] \\ &= \mathbb{E}[Z_{n+1} | \mathcal{F}_n] \\ &= Z_n \end{aligned}$$

et donc, en répétant,

$$\mathbb{E}[Z_m | \mathcal{F}_n] = Z_n, \quad \forall m \geq n.$$

Les martingales sont donc des processus constant en espérance. On peut également définir la notion de martingale sans filtration préalable.

Définition 2.1.4 A toute suite de variables aléatoires $Z = (Z_n)_{n \in \mathbb{N}_0}$, on peut associer sa *filtration naturelle*, donnée par

$$\mathcal{F}_n^Z = \sigma(Z_1, \dots, Z_n).$$

On dit que $(Z_n)_{n \in \mathbb{N}_0}$ est une *martingale* si elle est une martingale pour sa filtration naturelle, c'est-à-dire si

1. $\mathbb{E}[|Z_n|] < \infty$ pour tout $n \in \mathbb{N}_0$,
2. $\mathbb{E}[Z_{n+1} | Z_1, \dots, Z_n] = Z_n$ pour tout $n \in \mathbb{N}_0$.

Soit $(\mathcal{G}_n)_{n \in \mathbb{N}_0}$ une filtration. Notons que si $(Z_n)_{n \in \mathbb{N}}$ est une martingale pour $(\mathcal{G}_n)_{n \in \mathbb{N}_0}$, alors c'est nécessairement une martingale pour \mathcal{F}_n^Z : en effet, on a

$$\mathcal{F}_n^Z \subseteq \mathcal{G}_n$$

puisque $(Z_n)_{n \in \mathbb{N}}$ est adapté à $(\mathcal{G}_n)_{n \in \mathbb{N}_0}$ (la filtration naturelle est la plus petite filtration adaptée). De plus, en utilisant la propriété d'emboîtement de l'espérance conditionnelle, l'hypothèse que $(Z_n)_{n \in \mathbb{N}}$ est une martingale pour $(\mathcal{G}_n)_{n \in \mathbb{N}_0}$ et la mesurabilité de Z_n par rapport à \mathcal{F}_n^Z , on a

$$\mathbb{E}[Z_{n+1} | \mathcal{F}_n^Z] = \mathbb{E}[\mathbb{E}[Z_{n+1} | \mathcal{G}_n] | \mathcal{F}_n^Z] = \mathbb{E}[Z_n | \mathcal{F}_n^Z] = Z_n.$$

Présentons à présent quelques exemples simples de martingales.

Exemple 2.1.5 1. Soient $X_i, i \in \mathbb{N}_0$, des variables aléatoires indépendantes d'espérance nulle. Alors

$$Z_n = \sum_{i=1}^n X_i, \quad n \in \mathbb{N}_0$$

est une martingale.

2. Soient $X_i, i \in \mathbb{N}_0$, des variables aléatoires indépendantes d'espérance 1. Alors

$$Z_n = \prod_{i=1}^n X_i, \quad n \in \mathbb{N}_0$$

est une martingale.

3. Soient $X_i, i \in \mathbb{N}_0$, des variables aléatoires iid d'espérance nulle et de variance σ^2 . Alors

$$Z_n = \left(\sum_{i=1}^n X_i \right)^2 - n\sigma^2, \quad n \in \mathbb{N}_0$$

est une martingale.

4. Soient $X_i, i \in \mathbb{N}_0$, des variables aléatoires indépendantes d'espérance nulle et $(\mathcal{F}_n)_{n \in \mathbb{N}_0}$ leur filtration naturelle. Pour chaque $n \in \mathbb{N}_0$, considérons une variable aléatoire B_n bornée et mesurable pour \mathcal{F}_{n-1} ; on peut voir B_n comme un pari sur X_n . La fortune totale à l'instant n est $W_0 = 0$ et

$$W_n = \sum_{j=1}^n B_j X_j.$$

Alors, la suite $(W_n)_{n \in \mathbb{N}_0}$ est une martingale.

Exemple 2.1.6 [Martingale de Doob] Soit Y une variable aléatoire telle que $\mathbb{E}[|Y|] < \infty$. Soit \mathcal{F}_n une filtration et

$$Z_n = \mathbb{E}[Y | \mathcal{F}_n].$$

Alors Z_n est une martingale dans la filtration \mathcal{F}_n . En effet, les deux premiers points sont immédiats et pour le troisième, il suffit de remarquer que

$$\mathbb{E}[Z_{n+1} | \mathcal{F}_n] = \mathbb{E}[\mathbb{E}[Y | \mathcal{F}_{n+1}] | \mathcal{F}_n] = \mathbb{E}[Y | \mathcal{F}_n] = Z_n$$

car $\mathcal{F}_n \subset \mathcal{F}_{n+1}$. On dit que $(Z_n)_n$ est une martingale de Doob.

A la place de considérer un jeu équitable, on peut aussi s'intéresser à des jeux qui sont favorables au joueur, ou au contraire défavorables.

Définition 2.1.7 Une suite de variables aléatoires $(Z_n)_{n \in \mathbb{N}_0}$ forme une *sous-martingale* pour la filtration $(\mathcal{F}_n)_{n \in \mathbb{N}_0}$ si

1. $(Z_n)_{n \in \mathbb{N}_0}$ est adapté à $(\mathcal{F}_n)_{n \in \mathbb{N}_0}$,
2. $\mathbb{E}[|Z_n|] < +\infty$,
3. $\mathbb{E}[Z_{n+1} | \mathcal{F}_n] \geq Z_n$.

Si, dans le dernier point, \geq est remplacé par \leq , on parle de *sur-martingale*.

Bien entendu, $(Z_n)_{n \in \mathbb{N}_0}$ est une sous-martingale si et seulement si $(-Z_n)_{n \in \mathbb{N}_0}$ est une sur-martingale. De plus, un processus adapté à la filtration $(\mathcal{F}_n)_{n \in \mathbb{N}_0}$ est une martingale si et seulement si c'est à la fois un sous-martingale et une sur-martingale. Pour cette raison, nous nous intéresserons parfois uniquement à des propriétés sur les sous-martingales, les propriétés correspondantes des sur-martingales et des martingales pourront alors s'en déduire aisément.

Une sous-martingale (resp. sur-martingale) est donc un processus croissant (resp. décroissant) en moyenne conditionnelle et donc en moyenne. En effet, il suit de la tower property des espérances conditionnelles que

$$\mathbb{E}[Z_{n+1}] \geq \mathbb{E}[Z_n]$$

pour une sous-martingale, et de manière plus générale,

$$\mathbb{E}[Z_m | \mathcal{F}_n] \geq Z_n, \quad \forall m \geq n.$$

De même, si $(Z_n)_{n \in \mathbb{N}_0}$ est une sur-martingale, alors

$$\mathbb{E}[Z_m | \mathcal{F}_n] \leq Z_n, \quad \forall m \geq n.$$

Comme dans la Définition 2.1.4, on peut parler de sous-martingale (resp. de sur-martingale) sans spécifier une filtration : il s'agit alors d'une sous-martingale (resp. une sur-martingale) pour la filtration naturelle.

2.2 Transformées de martingales

Le résultat suivant est très utilisé et permet de construire facilement des sous-martingales à partir de (sous-)martingales.

Proposition 2.2.1

1. Soient $(Z_n)_{n \in \mathbb{N}_0}$ une martingale pour la filtration $(\mathcal{F}_n)_{n \in \mathbb{N}_0}$ et f une fonction convexe Borel-mesurable. Si $\mathbb{E}[|f(Z_n)|] < +\infty$ pour tout $n \in \mathbb{N}_0$, alors $(f(Z_n))_{n \in \mathbb{N}_0}$ est une sous-martingale pour la filtration $(\mathcal{F}_n)_{n \in \mathbb{N}_0}$.
2. Soient $(Z_n)_{n \in \mathbb{N}_0}$ une sous-martingale pour la filtration $(\mathcal{F}_n)_{n \in \mathbb{N}_0}$ et f une fonction convexe croissante Borel-mesurable. Si $\mathbb{E}[|f(Z_n)|] < +\infty$ pour tout $n \in \mathbb{N}_0$, alors $(f(Z_n))_{n \in \mathbb{N}_0}$ est une sous-martingale pour la filtration $(\mathcal{F}_n)_{n \in \mathbb{N}_0}$.

Démonstration : Remarquons que pour tout ensemble borélien $B \subseteq \mathbb{R}$, on a

$$(f(Z_n))^{-1}(B) = \{\omega \in \Omega : Z_n(\omega) \in f^{-1}(B)\} \in \mathcal{F}_n$$

puisque $f^{-1}(B)$ est un ensemble borélien de \mathbb{R} . De plus, dans le premier cas, on applique l'inégalité de Jensen pour les espérances conditionnelles et il vient

$$\mathbb{E}[f(Z_{n+1}) | \mathcal{F}_n] \geq f(\mathbb{E}[Z_{n+1} | \mathcal{F}_n]) = f(Z_n),$$

ce qui suffit. La dernière égalité devient une inégalité dans le deuxième cas. ■

En considérant différents choix de fonctions f , on obtient le corollaire suivant.

Corollaire 2.2.2

1. Si $(Z_n)_{n \in \mathbb{N}_0}$ est une martingale, alors $(|Z_n|)_{n \in \mathbb{N}_0}$ une sous-martingale.
2. Si $p > 1$ et si $(Z_n)_{n \in \mathbb{N}_0}$ est une martingale telle que $\mathbb{E}[|Z_n|^p] < +\infty$, alors $(|Z_n|^p)_{n \in \mathbb{N}_0}$ est une sous-martingale.
3. Si $(Z_n)_{n \in \mathbb{N}_0}$ est une sous-martingale, alors $((Z_n - a)^+)_{n \in \mathbb{N}_0}$ est une sous-martingale pour tout $a \in \mathbb{R}$.

Afin d'introduire la prochaine construction, nous devons considérer la notion de processus prévisible.

Définition 2.2.3 Une suite $(H_n)_{n \geq 2}$ de variables aléatoires est un *processus prévisible* pour la filtration $(\mathcal{F}_n)_{n \in \mathbb{N}_0}$ si H_n est \mathcal{F}_{n-1} -mesurable pour tout $n \geq 2$.

Proposition 2.2.4 Soit $(\mathcal{F}_n)_{n \in \mathbb{N}_0}$ une filtration. Soit $(Z_n)_{n \in \mathbb{N}_0}$ une sur-martingale et $(H_n)_{n \geq 2}$ un processus prévisible, non-négatif. Si H_n est borné pour tout $n \geq 2$, alors

$$(H \cdot Z)_n := \sum_{m=2}^n H_m(Z_m - Z_{m-1}), \quad n \geq 2$$

est une sur-martingale. De même, si $(Z_n)_{n \in \mathbb{N}_0}$ une martingale et $(H_n)_{n \geq 2}$ un processus prévisible tel que H_n est borné pour tout $n \geq 2$, alors $((H \cdot Z)_n)_{n \geq 2}$ est une martingale.

Démonstration : La vérification des deux premiers points est immédiate. Pour le troisième, on remarque que

$$\begin{aligned} \mathbb{E}[(H \cdot Z)_{n+1} | \mathcal{F}_n] &= (H \cdot Z)_n + \mathbb{E}[H_{n+1}(Z_{n+1} - Z_n) | \mathcal{F}_n] \\ &= (H \cdot Z)_n + H_{n+1} \mathbb{E}[(Z_{n+1} - Z_n) | \mathcal{F}_n] \\ &= (H \cdot Z)_n + H_{n+1} (\mathbb{E}[Z_{n+1} | \mathcal{F}_n] - Z_n) \\ &\leq (H \cdot Z)_n \end{aligned}$$

puisque H_{n+1} est positif et $\mathbb{E}[Z_{n+1} | \mathcal{F}_n] \leq Z_n$. On procède de même dans le cas d'une martingale. ■

Exemple 2.2.5 Supposons qu'un joueur mise de manière répétée sur les résultats du lancer d'une pièce. A chaque étape, il décide de miser une certaine quantité, en fonction des gains précédents. La stratégie est prévisible puisqu'elle ne dépend que des résultats passés. Supposons que Z_n est le résultat obtenu à l'étape n si le joueur partie un euro à chaque étape. Ainsi, $Z_n - Z_{n-1}$ donne le gain obtenu à l'étape n (1 ou -1). Si le joueur opte pour une stratégie H_n (qui dépend donc de Z_1, \dots, Z_{n-1}), il aura gagné au temps n la somme

$$(H \cdot Z)_n = \sum_{m=1}^n H_m(Z_m - Z_{m-1}).$$

Le résultat précédent affirme que cette suite est également une martingale. On en tire que

$$\mathbb{E}[(H \cdot Z)_n] = \mathbb{E}[(H \cdot Z)_2] = 0$$

pour tout n puisque $\mathbb{E}[(H \cdot Z)_2 | \mathcal{F}_2] = 0$. Cela signifie qu'aucune stratégie prévisible ne permet d'avoir des gains en moyenne strictement positifs.

Terminons cette section en montrant que toute sous-martingale peut être vue comme la somme d'une martingale et d'un processus prévisible croissant.

Proposition 2.2.6 *Toute sous-martingale $(Z_n)_{n \in \mathbb{N}_0}$ pour la filtration $(\mathcal{F}_n)_{n \in \mathbb{N}_0}$ peut s'écrire de manière presque sûrement unique sous la forme $Z_n = Y_n + H_n$, où $(Y_n)_{n \in \mathbb{N}_0}$ est une martingale pour la filtration $(\mathcal{F}_n)_{n \in \mathbb{N}_0}$ et $(H_n)_{n \in \mathbb{N}_0}$ est un processus prévisible pour la filtration $(\mathcal{F}_n)_{n \in \mathbb{N}_0}$ presque sûrement croissant tel que $H_1 = 0$.*

Démonstration : Supposons tout d'abord que la décomposition existe. Alors, on a

$$\mathbb{E}[Z_n | \mathcal{F}_{n-1}] = \mathbb{E}[Y_n | \mathcal{F}_{n-1}] + \mathbb{E}[H_n | \mathcal{F}_{n-1}] = Y_{n-1} + H_n = Z_{n-1} - H_{n-1} + H_n.$$

On en tire que

$$H_n = H_{n-1} + \mathbb{E}[Z_n | \mathcal{F}_{n-1}] - Z_{n-1}$$

et

$$Y_n = Z_n - H_n.$$

Par conséquent, si elle existe, la décomposition est unique et est donnée par récurrence par les deux formules ci-dessus (puisqu'on impose le cas de base $H_1 = 0$, et donc $Y_1 = Z_1$).

Il reste donc à montrer que les suites $(H_n)_{n \in \mathbb{N}_0}$ et $(Y_n)_{n \in \mathbb{N}_0}$ satisfont bien les propriétés de l'énoncé. Tout d'abord, comme $(Z_n)_{n \in \mathbb{N}_0}$ est une sous-martingale, on a $H_n \geq H_{n-1} \geq 0$. Par récurrence, il est également clair que H_n est \mathcal{F}_{n-1} mesurable. Enfin, on vérifie facilement que $\mathbb{E}[|Y_n|] < +\infty$ et que

$$\mathbb{E}[Y_n | \mathcal{F}_{n-1}] = \mathbb{E}[Z_n | \mathcal{F}_{n-1}] - H_n = Z_{n-1} - H_{n-1} = Y_{n-1},$$

ce qui suffit. ■

Remarque 2.2.7 Une martingale est en moyenne constante. Le processus $(H_n)_{n \in \mathbb{N}_0}$ doit donc cumuler les sauts de la sous-martingale $(Z_n)_{n \in \mathbb{N}_0}$. En effet, on a pour tout $n \geq 2$,

$$H_n = \sum_{k=2}^n \mathbb{E}[Z_k - Z_{k-1} | \mathcal{F}_{k-1}].$$

2.3 Temps d'arrêt et martingales arrêtées

Une martingale $(Z_n)_{n \in \mathbb{N}_0}$ sert principalement à décrire l'évolution d'un processus au cours du temps (Z_n est alors l'état du système à l'instant n). A tout processus évolutif, on associe la définition suivante.

Définition 2.3.1 Soit $(\mathcal{F}_n)_{n \in \mathbb{N}_0}$ une filtration. Une variable aléatoire T est un *temps d'arrêt* pour la filtration $(\mathcal{F}_n)_{n \in \mathbb{N}_0}$ si

1. T est à valeurs dans $\mathbb{N} \cup \{+\infty\}$,
2. $\{T \leq n\} \in \mathcal{F}_n$ pour tout $n \in \mathbb{N}_0$.

Un temps d'arrêt est donc un temps aléatoire qui suit l'évolution aléatoire sous-jacente. Dans notre exemple d'un jeu aléatoire répété, T peut être vu comme le temps auquel le joueur décide de quitter le jeu : cette décision de quitter le jeu directement après le temps n ne dépend que des résultats obtenus avant le temps n (inclus).

Remarque 2.3.2 La condition $\{T \leq n\} \in \mathcal{F}_n$ pour tout $n \in \mathbb{N}_0$ est équivalente à $\{T = n\} \in \mathcal{F}_n$. En effet, d'une part, on a

$$\{T = n\} = \{T \leq n\} \setminus \{T \leq n - 1\}$$

et d'autre part

$$\{T \leq n\} = \bigcup_{k=0}^n \{T = k\}.$$

Notons que cette définition ne se généralise pas au cas continu que nous aborderons par la suite.

Exemple 2.3.3 Soit $(X_n)_{n \in \mathbb{N}_0}$ une suite de variables aléatoires réelles adaptées à la filtration $(\mathcal{F}_n)_{n \in \mathbb{N}_0}$. Pour tout $a > 0$, le temps aléatoire

$$T_a = \inf\{n \in \mathbb{N}_0 : X_n > a\},$$

appelé *temps de premier passage au seuil a* , est un temps d'arrêt. En effet, on a

$$\{T_a \leq n\} = \{\exists m \leq n : X_m > a\} = \bigcup_{m=1}^n \{X_m > a\} \in \mathcal{F}_n.$$

D'une manière générale, le temps de première entrée dans un borélien B , défini par

$$T_B = \inf\{n \in \mathbb{N}_0 : X_n \in B\},$$

est un temps d'arrêt. Par contre, le temps de dernier passage au seuil a , défini par

$$\tilde{T}_a = \sup\{n \in \mathbb{N}_0 : X_n > a\},$$

n'est pas un temps d'arrêt.

Le théorème suivant sera la base de cette section.

Théorème 2.3.4 Soit $(Z_n)_{n \in \mathbb{N}_0}$ une martingale (resp. sous-martingale, sur-martingale) pour une filtration $(\mathcal{F}_n)_{n \in \mathbb{N}_0}$. Si T un temps d'arrêt pour $(\mathcal{F}_n)_{n \in \mathbb{N}_0}$, alors

$$Z_{T \wedge n}, \quad n \in \mathbb{N}_0$$

est une martingale (resp. sous-martingale, sur-martingale) pour $(\mathcal{F}_n)_{n \in \mathbb{N}_0}$ ^a. En particulier, pour tout $n \in \mathbb{N}_0$, on a

$$\mathbb{E}[Z_{T \wedge n}] = \mathbb{E}[Z_1]$$

(resp. $\mathbb{E}[Z_{T \wedge n}] \geq \mathbb{E}[Z_1]$, $\mathbb{E}[Z_{T \wedge n}] \leq \mathbb{E}[Z_1]$).

a. avec $x \wedge y = \min(x, y)$

Démonstration : Pour tout $n \geq 2$, on considère la variable aléatoire $H_n = \mathbf{1}_{\{T \geq n\}} = 1 - \mathbf{1}_{\{T \leq n-1\}}$. Comme T est un temps d'arrêt pour $(\mathcal{F}_n)_{n \in \mathbb{N}_0}$, le processus $(H_n)_{n \geq 2}$ est prévisible, non-négatif et borné. La Proposition 2.2.4 implique que $H \cdot Z$ est une martingale (resp. sous-martingale, sur-martingale). Or, on a

$$\begin{aligned} (H \cdot X)_n &= \sum_{m=2}^n \mathbf{1}_{\{T \geq m\}} (Z_m - Z_{m-1}) \\ &= \sum_{m=2}^{n \wedge T} (Z_m - Z_{m-1}) \\ &= Z_{n \wedge T} - Z_1 \end{aligned}$$

et on en déduit facilement que $(Z_{n \wedge T})_n$ est également une martingale (resp. sous-martingale, sur-martingale). ■

Reprenons notre exemple du jeu de hasard répété : on peut décider d'arrêter le jeu lorsqu'on aura gagné un certain montant A (donne le temps d'arrêt T , on montrera par la suite qu'il est fini presque sûrement) ou lorsqu'on aura joué 10 parties ($n = 10$) : Le gain espéré est alors

$$\mathbb{E}[Z_{T \wedge 10}] = \mathbb{E}[Z_1].$$

Par contre, si on joue jusqu'à avoir gagné un certain montant A , l'espérance de gain est donné par

$$\mathbb{E}[Z_T] = A.$$

Par conséquent, même si $\mathbb{E}[Z_{T \wedge n}] = \mathbb{E}[Z_1]$, il est possible de construire une martingale $(Z_n)_{n \in \mathbb{N}_0}$ et un temps d'arrêt T tels que

$$\mathbb{E}[Z_T] \neq \mathbb{E}[Z_1].$$

Le résultat suivant donne trois conditions suffisantes pour avoir l'égalité. Notons que pour que Z_T soit bien défini, on doit s'assurer que $\mathbb{P}[T < +\infty] = 1$.

Théorème 2.3.5 (Martingale Stopping Theorem) Soit $(Z_n)_{n \in \mathbb{N}_0}$ une martingale (resp. sous-martingale, sur-martingale) pour la filtration $(\mathcal{F}_n)_{n \in \mathbb{N}_0}$ et T un temps d'arrêt pour $(\mathcal{F}_n)_{n \in \mathbb{N}_0}$. Si l'une des trois conditions suivantes est satisfaite

- (1) T est bornée, c'est-à-dire il existe $M < +\infty$ tel que $T(\omega) < M$ pour tout $\omega \in \Omega$,
 - (2) $\mathbb{P}[T < +\infty] = 1$ et les $Z_{n \wedge T}$ sont uniformément bornés, c'est-à-dire il existe $M < +\infty$ tel que $|Z_{n \wedge T}(\omega)| \leq M$ pour tous $n \in \mathbb{N}_0$ et $\omega \in \Omega$,
 - (3) $\mathbb{E}[T] < +\infty$ et il existe $M < +\infty$ tel que $\mathbb{E}[|Z_{n+1} - Z_n| \mid \mathcal{F}_n] < M$ pour tout $n \in \mathbb{N}_0$,
- alors

$$\mathbb{E}[Z_T] = \mathbb{E}[Z_1]$$

(resp. $\mathbb{E}[Z_T] \geq \mathbb{E}[Z_1]$, $\mathbb{E}[Z_T] \leq \mathbb{E}[Z_1]$).

Démonstration : A nouveau, concentrons-nous sur le cas où $(Z_n)_{n \in \mathbb{N}_0}$ est une martingale. Dans le premier cas, remarquons que pour tout $n \geq M$, on a $Z_{T \wedge n} = Z_T$, d'où $\mathbb{E}[Z_T] = \mathbb{E}[Z_{T \wedge n}] = \mathbb{E}[Z_1]$ par le Théorème 2.3.4.

Pour les deux cas suivant, puisque $\mathbb{P}[T < +\infty] = 1$, on a presque sûrement, $Z_{n \wedge T} = Z_T$ pour tout n suffisamment grand. Il s'ensuit que presque sûrement, on a

$$\lim_{n \rightarrow \infty} Z_{n \wedge T} = Z_T,$$

d'où

$$\mathbb{E}\left[\lim_{n \rightarrow \infty} Z_{n \wedge T}\right] = \mathbb{E}[Z_T].$$

De plus, par le Théorème 2.3.4, on sait que

$$\mathbb{E}[Z_{n \wedge T}] = \mathbb{E}[Z_1]$$

et il reste donc à montrer qu'on peut bien intervertir l'espérance et la limite. On va appliquer à chaque fois le Théorème de la convergence dominée de Lebesgue. Pour (2), puisqu'on peut majorer $|Z_{n \wedge T}|$ par la constante M qui est intégrable (la mesure de probabilité est finie), on obtient que

$$\lim_{n \rightarrow \infty} \mathbb{E}[Z_{n \wedge T}] = \mathbb{E}\left[\lim_{n \rightarrow \infty} Z_{n \wedge T}\right] = \mathbb{E}[Z_T].$$

Donc $\mathbb{E}[Z_T] = \mathbb{E}[Z_1]$. Pour (3), puisque

$$Z_{n \wedge T} = Z_1 + \sum_{i=1}^{n \wedge T - 1} (Z_{i+1} - Z_i),$$

on a

$$|Z_{n \wedge T}| \leq |Z_1| + \sum_{i=1}^{n \wedge T - 1} |Z_{i+1} - Z_i| \leq |Z_1| + \sum_{i=1}^{+\infty} |Z_{i+1} - Z_i| \mathbf{1}_{\{T > i\}}.$$

De plus, puisque $\{T > i\} \in \mathcal{F}_i$, on a

$$\begin{aligned} \mathbb{E}[|Z_1|] + \sum_{i=1}^{+\infty} \mathbb{E}[\mathbb{E}[|Z_{i+1} - Z_i| \mathbf{1}_{\{T > i\}} | \mathcal{F}_i]] &= \mathbb{E}[|Z_1|] + \sum_{i=1}^{+\infty} \mathbb{E}[\mathbb{E}[|Z_{i+1} - Z_i| | \mathcal{F}_i] \mathbf{1}_{\{T > i\}}] \\ &\leq \mathbb{E}[|Z_1|] + M \sum_{i=1}^{+\infty} \mathbb{E}[\mathbf{1}_{\{T > i\}}] \\ &= \mathbb{E}[|Z_1|] + M \sum_{i=1}^{+\infty} \mathbb{P}[T > i] \\ &\leq \mathbb{E}[|Z_1|] + M \sum_{i=0}^{+\infty} \mathbb{P}[T > i] \\ &= \mathbb{E}[|Z_1|] + M \mathbb{E}[T] \end{aligned}$$

et on conclut à nouveau en utilisant le Théorème de la convergence dominée. ■

Présentons des applications directes de ce théorème.

Corollaire 2.3.6 (Identité de Wald) Soit $(X_n)_{n \in \mathbb{N}_0}$ une suite de variables aléatoires indépendantes de même espérance et T un temps d'arrêt pour la filtration naturelle $\mathcal{F}_n = \sigma(X_1, \dots, X_n)$. Supposons que T est borné. Alors

$$\mathbb{E} \left[\sum_{n=1}^T X_n \right] = \mathbb{E}[T] \mathbb{E}[X_1]$$

Démonstration : Notons $\mu = \mathbb{E}[X_1]$. La suite $Z_n = \sum_{i=1}^n (X_i - \mu)$ est une martingale de moyenne nulle et donc, par Théorème 2.3.5, on a

$$\begin{aligned} 0 &= \mathbb{E}[Z_T] \\ &= \mathbb{E} \left[\sum_{i=1}^T X_i - T\mu \right] \\ &= \mathbb{E} \left[\sum_{i=1}^T X_i \right] - \mathbb{E}[T]\mu, \end{aligned}$$

d'où la conclusion. ■

Corollaire 2.3.7 Si $(Z_n)_{n \in \mathbb{N}_0}$ est une sous-martingale pour la filtration $(\mathcal{F}_n)_{n \in \mathbb{N}_0}$ et T un temps d'arrêt pour $(\mathcal{F}_n)_{n \in \mathbb{N}_0}$ borné par $N \in \mathbb{N}_0$. Alors

$$\mathbb{E}[Z_1] \leq \mathbb{E}[Z_T] \leq \mathbb{E}[Z_N].$$

Démonstration : Soit $N \in \mathbb{N}_0$ tel que $T \leq N$. Puisque T est borné, on sait par le Théorème 2.3.5 que $\mathbb{E}[Z_T] \geq \mathbb{E}[Z_1]$. De plus, puisque T est un temps d'arrêt, on a $\{T = k\} \in \mathcal{F}_k$ et il vient

$$\mathbb{E}[Z_N \mathbf{1}_{\{T=k\}} | \mathcal{F}_k] = \mathbf{1}_{\{T=k\}} \mathbb{E}[Z_N | \mathcal{F}_k] \geq \mathbf{1}_{\{T=k\}} Z_k = \mathbf{1}_{\{T=k\}} Z_T$$

pour tout $k \leq N$. Par conséquent, on obtient

$$\mathbb{E}[Z_N \mathbf{1}_{\{T=k\}}] = \mathbb{E}[\mathbb{E}[Z_N \mathbf{1}_{\{T=k\}} | \mathcal{F}_k]] \geq \mathbb{E}[\mathbf{1}_{\{T=k\}} Z_T].$$

Puisque $1 = \sum_{k=1}^N \mathbf{1}_{\{T=k\}}$, on en tire que

$$\mathbb{E}[Z_N] = \sum_{k=1}^N \mathbb{E}[Z_N \mathbf{1}_{\{T=k\}}] \geq \sum_{k=1}^N \mathbb{E}[\mathbf{1}_{\{T=k\}} Z_T] = \mathbb{E}[Z_T],$$

d'où la conclusion. ■

2.4 Inégalités de Doob

Les inégalités de Doob présentées dans cette section ont de nombreuses applications, et permettent notamment d'étudier le supremum d'un processus stochastique.

Lemme 2.4.1 Soit $(Z_n)_{n \in \mathbb{N}_0}$ une sous-martingale. Alors,

$$\mathbb{E}[Z_m \mathbf{1}_A] \geq \mathbb{E}[Z_n \mathbf{1}_A]$$

pour tout $m \geq n$ et tout $A \in \mathcal{F}_n$.

Démonstration : On sait que $\mathbb{E}[Z_m | \mathcal{F}_n] \geq Z_n$ pour tout $m \geq n$. Par conséquent, pour tout $A \in \mathcal{F}_n$, on a $\mathbb{E}[Z_m \mathbf{1}_A | \mathcal{F}_n] = \mathbf{1}_A \mathbb{E}[Z_m | \mathcal{F}_n] \geq Z_n \mathbf{1}_A$. On en tire alors

$$\mathbb{E}[Z_m \mathbf{1}_A] = \mathbb{E}[\mathbb{E}[Z_m \mathbf{1}_A | \mathcal{F}_n]] \geq \mathbb{E}[Z_n \mathbf{1}_A].$$

■

Proposition 2.4.2 (Inégalités de Doob) Soit $(Z_n)_{n \in \mathbb{N}_0}$ une sous-martingale. Alors, pour tout $c > 0$, on a

$$c \mathbb{P}\left[\max_{k \in \{1, \dots, n\}} Z_k \geq c\right] \leq \mathbb{E}[Z_n \mathbf{1}_{\{\max_{k \in \{1, \dots, n\}} Z_k \geq c\}}]$$

pour tout $n \in \mathbb{N}_0$. Si de plus $(Z_n)_{n \in \mathbb{N}_0}$ est non-négatif, on en tire que

$$\mathbb{P}\left[\max_{k \in \{1, \dots, n\}} Z_k \geq c\right] \leq \frac{\mathbb{E}[Z_n]}{c}$$

pour tout $n \in \mathbb{N}_0$.

Démonstration : Considérons le temps d'arrêt

$$T = \min\{n \in \mathbb{N}_0 : Z_n \geq c\}.$$

Alors,

$$\mathbb{P}\left[\max_{k \in \{1, \dots, n\}} Z_k \geq c\right] = \mathbb{P}[T \leq n] = \sum_{k=1}^n \mathbb{P}[T = k].$$

De plus, on a

$$c \mathbb{P}[T = k] = c \mathbb{E}[\mathbf{1}_{\{T=k\}}] \leq \mathbb{E}[Z_k \mathbf{1}_{\{T=k\}}] \leq \mathbb{E}[Z_n \mathbf{1}_{\{T=k\}}]$$

par le Lemme 2.4.1. Il s'ensuit que

$$c \mathbb{P}\left[\max_{k \leq n} Z_k \geq c\right] = \sum_{k=1}^n c \mathbb{P}[T = k] \leq \sum_{k=1}^n \mathbb{E}[Z_n \mathbf{1}_{\{T=k\}}] = \mathbb{E}\left[Z_n \sum_{k=1}^n \mathbf{1}_{\{T=k\}}\right] = \mathbb{E}[Z_n \mathbf{1}_{\{T \leq n\}}].$$

La deuxième inégalité est évidente. ■

L'intérêt de cette inégalité est qu'elle permet de majorer une quantité faisant intervenir tout le processus jusqu'au temps n par une quantité ne dépendant que de Z_n .

Remarque 2.4.3 En particulier, si $(Z_n)_{n \in \mathbb{N}_0}$ est une martingale, alors pour tout $c > 0$, on a

$$\mathbb{P}\left[\sup_{k \leq n} |Z_k| \geq c\right] \leq \frac{\mathbb{E}[|Z_n|]}{c}.$$

Les inégalités de Doob permettent de retrouver les inégalités de Kolmogorov.

Proposition 2.4.4 (Inégalités de Kolmogorov) Soient X_1, \dots, X_n des variables aléatoires indépendantes centrées et de variance finie. Posons $S_j = \sum_{k=1}^j X_k$ pour tout $j \in \{1, \dots, n\}$. Alors pour tout $c > 0$,

$$\mathbb{P}\left[\max_{j \in \{1, \dots, n\}} |S_j| \geq c\right] \leq \frac{\mathbb{E}[S_n^2]}{c^2}.$$

Démonstration : On sait que la suite $(S_n)_n$ définit une martingale. Par conséquent, $(S_n^2)_n$ est une sous-martingale. En utilisant l'inégalité de Doob, on obtient

$$\mathbb{P}\left[\max_{j \in \{1, \dots, n\}} |S_j| \geq c\right] = \mathbb{P}\left[\max_{j \in \{1, \dots, n\}} S_j^2 \geq c^2\right] \leq \frac{1}{c^2} \mathbb{E}[S_n^2].$$

■

Les inégalités de Doob permettent également d'obtenir le résultat suivant, qui nous dit que la norme L^p du maximum d'une sous-martingale peut être contrôlée par la norme L^p de la sous-martingale. Cela nous permettra d'obtenir la convergence en norme L^p de martingales. Commençons par rappeler que si X est une variable aléatoire, alors

$$\|X\|_p = \mathbb{E}[|X|^p]^{1/p}.$$

Théorème 2.4.5 (Inégalité L^p de Doob) Considérons une sous-martingale $(Z_n)_{n \in \mathbb{N}_0}$ non-négative. Pour tout $p > 1$, on a

$$\left\| \max_{k \in \{1, \dots, n\}} Z_k \right\|_p \leq \frac{p}{p-1} \|Z_n\|_p, \quad \forall n \in \mathbb{N}_0.$$

Démonstration : Cela résulte immédiatement de la Proposition 2.4.2 et du lemme suivant avec $X = \max_{k \in \{1, \dots, n\}} Z_k$ et $Y = Z_n$. ■

Lemme 2.4.6 Soient X et Y deux variables aléatoires non-négatives et $p > 1$. Supposons que $\mathbb{E}[Y^p] < +\infty$ et que

$$c \mathbb{P}[X \geq c] \leq \mathbb{E}[Y \mathbf{1}_{\{X \geq c\}}]$$

pour tout $c > 0$. Alors

$$\|X\|_p \leq \frac{p}{p-1} \|Y\|_p.$$

Démonstration : Posons $X_n = \min(X, n)$. Alors pour tout n , X_n est une variable aléatoire bornée. De plus, on a

$$\mathbb{P}[X_n \geq c] = 0 \text{ si } n < c$$

et si $n \geq c$,

$$\mathbb{P}[X_n \geq c] = \mathbb{P}[X \geq c] \text{ et } \mathbf{1}_{\{X_n \geq c\}} = \mathbf{1}_{\{X \geq c\}}.$$

Ainsi, pour tout $c > 0$, la variable X_n satisfait

$$c \mathbb{P}[X_n \geq c] \leq \mathbb{E}[Y \mathbf{1}_{\{X_n \geq c\}}].$$

Pour tout $t \in [0, n]$, on a

$$t^p = p \int_0^t x^{p-1} dx = p \int_0^n x^{p-1} \mathbf{1}_{\{x \leq t\}} dx.$$

En évaluant cette relation en $t = X_n$ et en prenant l'espérance, le théorème de Fubini¹ donne

$$\begin{aligned} \mathbb{E}[X_n^p] &= p \int_0^n x^{p-1} \mathbb{E}[\mathbf{1}_{\{x \leq X_n\}}] dx \\ &= p \int_0^n x^{p-1} \mathbb{P}[x \leq X_n] dx \\ &\leq p \int_0^n x^{p-2} \mathbb{E}[Y \mathbf{1}_{\{X_n \geq x\}}] dx \\ &= p \mathbb{E} \left[Y \int_0^n x^{p-2} \mathbf{1}_{\{X_n \geq x\}} dx \right] \\ &= p \mathbb{E} \left[Y \int_0^{X_n} x^{p-2} dx \right] \\ &= \frac{p}{p-1} \mathbb{E}[Y X_n^{p-1}] \\ &\leq \frac{p}{p-1} \|Y\|_p \|X_n^{p-1}\|_q \end{aligned}$$

où la dernière ligne résulte de l'inégalité de Hölder, avec $\frac{1}{p} + \frac{1}{q} = 1$, et où toutes les espérances et intégrales sont finies puisque X_n est borné. Remarquons que

$$\|X_n^{p-1}\|_q = \mathbb{E}[X_n^{q(p-1)}]^{1/q} = \mathbb{E}[X_n^p]^{1-1/p}$$

puisque $q(p-1) = p$. On en tire donc que

$$\mathbb{E}[X_n^p] \leq \frac{p}{p-1} \|Y\|_p \mathbb{E}[X_n^p]^{1-1/p}.$$

Comme X_n est bornée, $\mathbb{E}[X_n^p]$ est fini et on obtient

$$\|X_n\|_p \leq \frac{p}{p-1} \|Y\|_p.$$

Pour conclure, puisque $X = \lim_{n \rightarrow +\infty} X_n$, le lemme de Fatou donne

$$\mathbb{E}[X^p]^{1/p} = \mathbb{E} \left[\liminf_{n \rightarrow +\infty} X_n^p \right]^{1/p} \leq \left(\liminf_{n \rightarrow +\infty} \mathbb{E}[X_n^p] \right)^{1/p} \leq \frac{p}{p-1} \|Y\|_p$$

d'où la conclusion. ■

1. tout est positif!

2.5 Convergence presque sûre des martingales

Les martingales ont l'élégance de ne pas exiger grand chose pour converger. Remarquons que la définition d'une sous-martingale est semblable à celle d'une suite qui, en tendance et conditionnellement au passé, est croissante. Nous savons qu'une suite de réels qui est croissante et majorée converge. Nous allons montrer un résultat analogue pour les sous-martingale, en remplaçant l'hypothèse de suite majorée par une majoration de l'espérance, i.e. en imposant $\sup_n \mathbb{E}[|Z_n|] < +\infty$. Nous allons dans cette section considérer la convergence presque sûre. Dans les sections suivantes, nous étudierons le cas de la convergence en norme L^p , le cas où $p = 1$ étant un peu plus délicat.

Soit $(Z_n)_{n \in \mathbb{N}_0}$ un processus adapté à la filtration $(\mathcal{F}_n)_{n \in \mathbb{N}_0}$. Soient $a < b$ deux réels. On note $N_n(a, b)$ le nombre de montées du processus entre a et b au temps n , c'est-à-dire le plus grand entier k pour lequel il existe deux suites de naturels entrelacées

$$1 \leq s_1 < t_1 < s_2 < t_2 < \dots < s_k < t_k \leq n$$

avec $Z_{s_i} \leq a$ et $Z_{t_i} \geq b$. Considérons les temps d'arrêt donnés par récurrence via les formules

$$\begin{cases} S_1 = \inf\{n \geq 1 : Z_n \leq a\} \\ T_k = \inf\{n > S_k : Z_n \geq b\} \\ S_k = \inf\{n > T_{k-1} : Z_n \leq a\}. \end{cases}$$

On vérifie facilement qu'il s'agit de temps d'arrêts (potentiellement infinis) en écrivant par exemple

$$\{T_k \leq n\} = \bigcup_{1 \leq s_1 < t_1 < s_2 < t_2 < \dots < s_k < t_k \leq n} \{Z_{s_1} \leq a, Z_{t_1} \geq b, \dots, Z_{s_k} \leq a, Z_{t_k} \geq b\}.$$

On a alors

$$N_n(a, b) = \sum_{k=1}^{+\infty} \mathbf{1}_{\{T_k \leq n\}}$$

puisque'on a effectué une montée entre a et b à chaque temps T_k . En particulier, $N_n(a, b)$ est \mathcal{F}_n -mesurable.

Lemme 2.5.1 (Nombre de montées) Soit $(Z_n)_{n \in \mathbb{N}}$ une sous-martingale pour la filtration $(\mathcal{F}_n)_{n \in \mathbb{N}_0}$. Pour tous réels $a < b$, on a

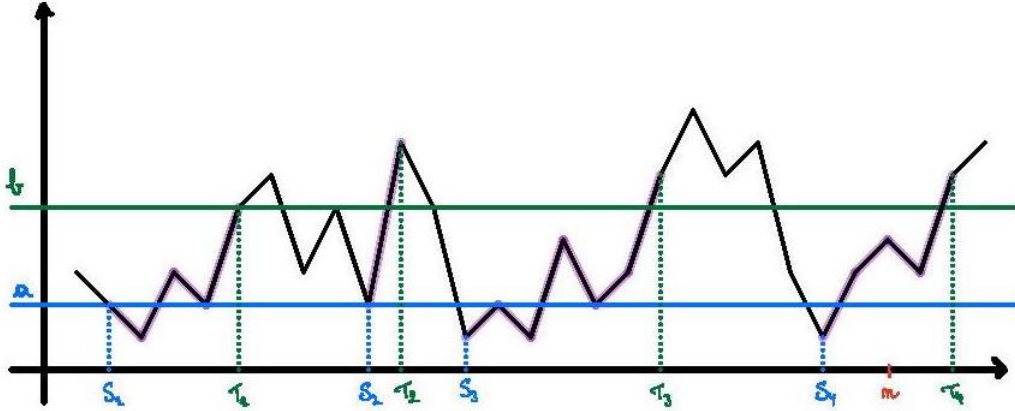
$$(b - a)\mathbb{E}[N_n(a, b)] \leq \mathbb{E}[(Z_n - a)^+ - (Z_1 - a)^+].$$

Démonstration : Posons $X_n = (Z_n - a)^+$. On sait par le Corollaire 2.2.2 que $(X_n)_{n \in \mathbb{N}_0}$ est une sous-martingale. Posons

$$H_m = \begin{cases} 1 & \text{si il existe } k \text{ tel que } S_k < m \leq T_k \\ 0 & \text{sinon,} \end{cases}$$

le processus qui vaut 1 si on est dans une "montée" (i.e. si on va de a à b), et 0 si on est dans une "descente". Le processus ainsi défini est prévisible puisque

$$H_m = \sum_{k=1}^{+\infty} \mathbf{1}_{\{S_k < m \leq T_k\}} = \sum_{k=1}^{+\infty} \mathbf{1}_{\{S_k \leq m-1\}} (1 - \mathbf{1}_{\{T_k \leq m-1\}}).$$


 FIGURE 2.1 – Sur cette réalisation, $N_n(a, b) = 3$.

Puisque H_j vaut 1 si $j \in \{S_k + 1, \dots, T_k\}$ et est nul si $j \in \{T_k + 1, \dots, S_{k+1}\}$ pour un k , et puisque $N_n(a, b)$ est le plus grand entier k tel que $T_k \leq n$, on a

$$\begin{aligned}
 (H \cdot X)_n &= \sum_{j=2}^n H_j (X_j - X_{j-1}) \\
 &= \sum_{k=1}^{N_n(a,b)} \sum_{j=S_k+1}^{T_k} (X_j - X_{j-1}) + \mathbf{1}_{\{S_{N_n(a,b)+1} < n\}} \sum_{j=S_{N_n(a,b)+1}+1}^n (X_j - X_{j-1}) \\
 &= \sum_{k=1}^{N_n(a,b)} (X_{T_k} - X_{S_k}) + \mathbf{1}_{\{S_{N_n(a,b)+1} < n\}} (X_n - X_{S_{N_n(a,b)+1}})
 \end{aligned}$$

où la première partie correspond à des sommes télescopiques sur chaque “montée”, et la deuxième à la dernière “montée” potentiellement incomplète. Remarquons que $X_{S_k} = 0$ et $X_{T_k} \geq b - a$, ce qui implique que

$$(H \cdot X)_n \geq N_n(a, b)(b - a) + \mathbf{1}_{\{S_{N_n(a,b)+1} < n\}} X_n \geq N_n(a, b)(b - a)$$

où la deuxième inégalité provient du fait que $X_n \geq 0$. Posons à présent $K_n = 1 - H_n$. La Proposition 2.2.4 implique que $(K \cdot X)_n$ est une sous-martingale et par conséquent, on sait que

$$\mathbb{E}[(K \cdot X)_n] \geq \mathbb{E}[(K \cdot X)_2] \geq 0$$

puisque $\mathbb{E}[(K \cdot X)_2 | \mathcal{F}_1] = K_2 \mathbb{E}[X_2 - X_1 | \mathcal{F}_1] \geq 0$. En utilisant les sommes télescopiques, on a

$$X_n - X_1 = \sum_{j=2}^n (X_j - X_{j-1}) = \sum_{j=2}^n (H_j + K_j)(X_j - X_{j-1}) = (H \cdot X)_n + (K \cdot X)_n,$$

d’où

$$(b - a)\mathbb{E}[N_n(a, b)] \leq \mathbb{E}[(H \cdot X)_n] \leq \mathbb{E}[(H \cdot X)_n] + \mathbb{E}[(K \cdot X)_n] = \mathbb{E}[X_n - X_1],$$

ce qui conclut la preuve. ■

Cette majoration du nombre de montées permet de démontrer le premier résultat de convergence suivant.

Théorème 2.5.2 (Convergence presque sûre) Soit $(Z_n)_{n \in \mathbb{N}_0}$ une sous-martingale. Supposons que

$$\sup_{n \in \mathbb{N}_0} \mathbb{E}[Z_n^+] < +\infty.$$

Alors $(Z_n)_{n \in \mathbb{N}_0}$ converge presque sûrement vers une variable aléatoire Z telle que $\mathbb{E}[|Z|] < +\infty$.

Démonstration : Soient $a, b \in \mathbb{Q}$ tels que $a < b$. Par le Lemme 2.5.1 du nombre de montées, on sait que pour tout $n \geq 1$, on a

$$(b - a)\mathbb{E}[N_n(a, b)] \leq \mathbb{E}[(Z_n - a)^+ - (Z_1 - a)^+] \leq \mathbb{E}[(Z_n - a)^+] \leq \mathbb{E}[Z_n^+] + |a|$$

puisque $(x - a)^+ \leq x^+ + |a|^2$. Si n tend vers l'infini, alors $N_n(a, b)$ tend vers le nombre total $N(a, b)$ de montée entre a et b . Comme $\sup_{n \in \mathbb{N}_0} \mathbb{E}[Z_n^+] < +\infty$, la relation précédente et le théorème de la convergence monotone impliquent que $\mathbb{E}[N(a, b)] < +\infty$. Par conséquent, $N(a, b)$ est fini presque sûrement. On en tire donc que l'événement

$$\bigcup_{a, b \in \mathbb{Q}} \left\{ \liminf_{n \rightarrow +\infty} Z_n \leq a < b \leq \limsup_{n \rightarrow +\infty} Z_n \right\}$$

a une probabilité nulle. Il s'ensuit que $\liminf_{n \rightarrow +\infty} Z_n = \limsup_{n \rightarrow +\infty} Z_n$ presque sûrement, et donc que $Z := \lim_{n \rightarrow +\infty} Z_n$ existe presque sûrement. Par le Lemme de Fatou, on a $\mathbb{E}[Z^+] \leq \liminf_{n \rightarrow +\infty} \mathbb{E}[Z_n^+] < +\infty$. Enfin, comme $(Z_n)_{n \in \mathbb{N}_0}$ est une sous-martingale, on a

$$\mathbb{E}[Z_n^-] = \mathbb{E}[Z_n^+] - \mathbb{E}[Z_n] \leq \mathbb{E}[Z_n^+] - \mathbb{E}[Z_1].$$

Une deuxième application du Lemme de Fatou permet de conclure. ■

Remarque 2.5.3 Remarquons que l'on vient de montrer que si $(Z_n)_{n \in \mathbb{N}_0}$ est une sous-martingale, alors

$$\sup_{n \in \mathbb{N}_0} \mathbb{E}[Z_n^+] < +\infty \iff \sup_{n \in \mathbb{N}_0} \mathbb{E}[|Z_n|] < +\infty.$$

Ainsi, on se ramène bien à l'hypothèse de majoration de la suite des espérances $(\mathbb{E}[|Z_n|])_n$ comme mentionné dans l'introduction.

Corollaire 2.5.4 Si $(Z_n)_{n \in \mathbb{N}_0}$ est une martingale non-négative, alors $(Z_n)_{n \in \mathbb{N}_0}$ converge presque sûrement vers une variable aléatoire finie.

Démonstration : Puisque Z_n est non-négative, on a

$$\sup_{n \in \mathbb{N}_0} \mathbb{E}[Z_n^+] = \sup_{n \in \mathbb{N}_0} \mathbb{E}[Z_n] = \mathbb{E}[Z_1] < +\infty,$$

ce qui suffit par le Théorème 2.5.2. ■

2. On a bien sûr $x - a \leq x^+ + |a|$ et comme le deuxième membre est toujours positif, on obtient la relation annoncée.

Exemple 2.5.5 Considérons le problème de la ruine du joueur participant à un jeu équitable : Z_n représente le gain obtenu après le $n^{\text{ième}}$ jeu. Supposons en outre que la banque ne fait pas crédit, donc à chaque partie, le joueur gagne ou perd une pièce, mais le jeu s'arrête si le joueur perd alors qu'il ne lui restait pas d'argent : Soit T le temps d'arrêt égal au nombre de parties jouées avant que le joueur ne perde tout son argent, i.e. $T = \{n \geq 1 : Z_n = 0\}$. Par le Théorème 2.3.4, $(Z_{n \wedge T})_{n \in \mathbb{N}_0}$ est une martingale et puisqu'elle est non-négative, la limite

$$\lim_{n \rightarrow +\infty} Z_{n \wedge T}$$

existe et est finie presque sûrement par le Corollaire 2.5.4. Or, il est clair que pour tout $n \in \mathbb{N}_0$, on a $|Z_{n+1} - Z_n| = 1$. Par conséquent, si $T = +\infty$ avec une probabilité non-nulle, alors la suite $(Z_{n \wedge T})_{n \in \mathbb{N}_0}$ ne converge pas vers une variable finie sur un événement de probabilité non-nulle, ce qui est impossible. On en tire donc que $\mathbb{P}[T < +\infty] = 1$ et le joueur finira par perdre tout son argent.

2.6 Convergence L^p , $p > 1$

Les inégalités de Doob dans L^p permettent facilement de déduire des résultats précédents la convergence des martingales en norme L^p pour tout $p > 1$.

Théorème 2.6.1 (Convergence dans L^p) Soient $(Z_n)_{n \in \mathbb{N}_0}$ une martingale pour la filtration $(\mathcal{F}_n)_{n \in \mathbb{N}_0}$ et $p > 1$. Supposons que

$$\sup_{n \in \mathbb{N}_0} \mathbb{E}[|Z_n|^p] < +\infty.$$

Alors, la suite $(Z_n)_{n \in \mathbb{N}_0}$ converge dans L^p .

Démonstration : En utilisant l'inégalité de Hölder, on a $\mathbb{E}[Z_n^+] \leq \mathbb{E}[|Z_n|] \leq (\mathbb{E}[|Z_n|^p])^{1/p}$. Le Théorème 2.5.2 implique que la suite $(Z_n)_{n \in \mathbb{N}_0}$ converge presque sûrement vers une variable aléatoire Z . Par le Théorème 2.4.5, on sait que

$$\mathbb{E}\left[\left|\max_{k \leq n} Z_k\right|^p\right] \leq \left(\frac{p}{p-1}\right)^p \mathbb{E}[|Z_n|^p] \leq \left(\frac{p}{p-1}\right)^p M$$

pour tout $n \in \mathbb{N}_0$ et pour une constante M finie. Le Théorème de la convergence monotone implique que $\mathbb{E}\left[\left|\sup_{n \in \mathbb{N}_0} Z_n\right|^p\right] < +\infty$. On peut alors appliquer le Théorème de la convergence dominée (en majorant $|Z_n|^p$ par $\left|\sup_{m \in \mathbb{N}_0} Z_m\right|^p$) pour obtenir que

$$\lim_{n \rightarrow +\infty} \mathbb{E}[|Z_n - Z|^p] = 0.$$

■

2.7 Convergence L^1

La convergence des martingales en norme L^1 est plus délicate à traiter et nécessite l'introduction de la notion d'intégrabilité uniforme.

Définition 2.7.1 Un processus stochastique $\{X(t) : t \in I\}$ est *uniformément intégrable* si

$$\lim_{M \rightarrow +\infty} \left(\sup_{t \in I} \mathbb{E}[\mathbf{1}_{\{|X(t)| > M\}} |X(t)|] \right) = 0.$$

Remarque 2.7.2 Pour une variable aléatoire, la notion d'intégrabilité uniforme est plus forte que la notion d'intégrabilité. En effet, supposons que X est une variable telle que $\mathbb{E}[|X|] < +\infty$. Puisque $\mathbf{1}_{\{|X| > M\}}|X| \leq |X|$, une simple application du Théorème de la convergence dominée implique que $\lim_{M \rightarrow +\infty} \mathbb{E}[\mathbf{1}_{\{|X| > M\}}|X|] = 0$.

Remarque 2.7.3 Si $\{X(t) : t \in I\}$ est uniformément intégrable, il existe M suffisamment grand tel que $\mathbb{E}[\mathbf{1}_{\{|X(t)| > M\}}|X(t)|] \leq 1$. Par conséquent, on a

$$\mathbb{E}[|X(t)|] \leq \mathbb{E}[\mathbf{1}_{\{|X(t)| \leq M\}}|X(t)|] + \mathbb{E}[\mathbf{1}_{\{|X(t)| > M\}}|X(t)|] \leq M + 1$$

d'où

$$\sup_{t \in I} \mathbb{E}[|X(t)|] < +\infty.$$

En particulier, la notion d'intégrabilité uniforme pour une suite $(Z_n)_{n \in \mathbb{N}_0}$ de variables aléatoires est plus forte que la condition $\sup_{n \in \mathbb{N}_0} \mathbb{E}[|X_n|] < +\infty$ considérée précédemment pour la convergence presque sûre des martingales.

Montrons à présent que les martingales de Doob introduites dans l'Exemple 2.1.6 sont uniformément intégrables. En fait, nous verrons que ces martingales caractérisent entièrement les martingales uniformément intégrables.

Proposition 2.7.4 Soit Y une variable aléatoire telle que $\mathbb{E}[|Y|] < \infty$. Soit $(\mathcal{F}_n)_{n \in \mathbb{N}_0}$ une filtration et pour tout $n \in \mathbb{N}_0$, posons

$$Z_n = \mathbb{E}[Y | \mathcal{F}_n].$$

Alors $(Z_n)_{n \in \mathbb{N}_0}$ est uniformément intégrable.

Démonstration : Fixons $\varepsilon > 0$. En utilisant l'inégalité de Jensen, on sait que pour tout n , $|\mathbb{E}[Y | \mathcal{F}_n]| \leq \mathbb{E}[|Y| | \mathcal{F}_n]$. Ainsi, pour tout M , il vient

$$\begin{aligned} \mathbb{E}[|\mathbb{E}[Y | \mathcal{F}_n]| \mathbf{1}_{\{|\mathbb{E}[Y | \mathcal{F}_n]| > M\}}] &\leq \mathbb{E}[\mathbb{E}[|Y| | \mathcal{F}_n] \mathbf{1}_{\{\mathbb{E}[|Y| | \mathcal{F}_n] > M\}}] \\ &= \mathbb{E}[\mathbb{E}[|Y| \mathbf{1}_{\{\mathbb{E}[|Y| | \mathcal{F}_n] > M\}} | \mathcal{F}_n]] \\ &= \mathbb{E}[|Y| \mathbf{1}_{\{\mathbb{E}[|Y| | \mathcal{F}_n] > M\}}] \end{aligned}$$

puisque $\{\mathbb{E}[|Y| | \mathcal{F}_n] > M\} \in \mathcal{F}_n$. De plus, comme $\mathbb{E}[|Y|] < \infty$, on sait par la Remarque 2.7.2 que Y est uniformément intégrable. Il existe donc $a > 0$ tel que

$$\mathbb{E}[|Y| \mathbf{1}_{\{|Y| > a\}}] < \frac{\varepsilon}{2}.$$

Par conséquent, si A est tel que $\mathbb{P}[A] < \varepsilon/2a$, on a

$$\mathbb{E}[|Y| \mathbf{1}_A] = \mathbb{E}[|Y| \mathbf{1}_{\{|Y| > a\} \cap A}] + \mathbb{E}[|Y| \mathbf{1}_{\{|Y| \leq a\} \cap A}] \leq \frac{\varepsilon}{2} + a\mathbb{P}[A] < \varepsilon.$$

En utilisant l'inégalité de Markov, on a

$$\mathbb{P}[\mathbb{E}[|Y| | \mathcal{F}_n] > M] \leq \frac{\mathbb{E}[\mathbb{E}[|Y| | \mathcal{F}_n]]}{M} = \frac{\mathbb{E}[|Y|]}{M}.$$

On en tire donc que

$$\mathbb{E}[|Y| \mathbf{1}_{\{\mathbb{E}[|Y| | \mathcal{F}_n] > M\}}] < \varepsilon$$

si $M > \frac{2a\mathbb{E}[|Y|]}{\varepsilon}$. ■

Lemme 2.7.5 *Supposons que $(X_n)_{n \in \mathbb{N}}$ est une suite uniformément intégrable de variables aléatoires qui converge vers en probabilité vers une variable aléatoire X telle que $\mathbb{E}[|X|] < +\infty$. Alors la suite $(X_n)_{n \in \mathbb{N}}$ converge vers X en norme L^1 .*

Démonstration : Pour tout $M > 0$, considérons la fonction

$$\varphi_M(x) = \begin{cases} M & \text{si } x \geq M \\ x & \text{si } |x| \leq M \\ -M & \text{si } x \leq -M. \end{cases}$$

Pour tout n , on a

$$\begin{aligned} \mathbb{E}[|X_n - X|] &\leq \mathbb{E}[|X_n - \varphi_M(X_n)|] + \mathbb{E}[|\varphi_M(X_n) - \varphi_M(X)|] + \mathbb{E}[|\varphi_M(X) - X|] \\ &\leq \mathbb{E}[|X_n| \mathbf{1}_{\{|X_n| \geq M\}}] + \mathbb{E}[|\varphi_M(X_n) - \varphi_M(X)|] + \mathbb{E}[|X| \mathbf{1}_{\{|X| \geq M\}}] \end{aligned}$$

puisque $|\varphi_M(x) - x| \leq |x| \mathbf{1}_{\{|x| \geq M\}}$. Il reste à montrer que chacun des 3 termes peut être rendu aussi petit que souhaité, en prenant n et M suffisamment grand. Soit donc $\varepsilon > 0$. Par l'hypothèse d'intégrabilité uniforme, on peut supposer que

$$\mathbb{E}[|X_n| \mathbf{1}_{\{|X_n| \geq M\}}] \leq \varepsilon$$

en prenant M suffisamment grand. De plus, par la Remarque 2.7.2, X est également uniformément intégrable. A nouveau, si M est suffisamment grand, on a

$$\mathbb{E}[|X| \mathbf{1}_{\{|X| \geq M\}}] \leq \varepsilon.$$

Enfin, on vérifie facilement que $|\varphi_M(x) - \varphi_M(y)| \leq |x - y|$ pour tous x, y . Ainsi, $(\varphi_M(X_n))_n$ converge également en probabilité vers $\varphi_M(X)$. Par conséquent,

$$\begin{aligned} \mathbb{E}[|\varphi_M(X_n) - \varphi_M(X)|] &= \mathbb{E}[|\varphi_M(X_n) - \varphi_M(X)| \mathbf{1}_{\{|\varphi_M(X_n) - \varphi_M(X)| > \frac{\varepsilon}{3}\}}] \\ &\quad + \mathbb{E}[|\varphi_M(X_n) - \varphi_M(X)| \mathbf{1}_{\{|\varphi_M(X_n) - \varphi_M(X)| \leq \frac{\varepsilon}{3}\}}] \\ &\leq 2M\mathbb{P}[|\varphi_M(X_n) - \varphi_M(X)| > \frac{\varepsilon}{3}] + \frac{\varepsilon}{3} \end{aligned}$$

puisque $|\varphi_M(x) - \varphi_M(y)| \leq 2M$ pour tous x, y . En utilisant la convergence en probabilité, il vient que

$$\mathbb{P}[|\varphi_M(X_n) - \varphi_M(X)| > \frac{\varepsilon}{3}] \leq \frac{\varepsilon}{3M}$$

pour tout n suffisamment grand, et donc

$$\mathbb{E}[|\varphi_M(X_n) - \varphi_M(X)|] \leq \varepsilon. \quad \blacksquare$$

Le lemme précédent permet d'obtenir facilement la caractérisation des martingales qui convergent en norme L^1 .

Théorème 2.7.6 Soient $(Z_n)_{n \in \mathbb{N}_0}$ une martingale pour la filtration $(\mathcal{F}_n)_{n \in \mathbb{N}_0}$. Les assertions suivantes sont équivalentes :

1. La martingale est uniformément intégrable.
2. La martingale converge presque sûrement et en norme L^1 .
3. Il existe une variable aléatoire Z telle que $Z_n = \mathbb{E}[Z | \mathcal{F}_n]$ pour tout $n \in \mathbb{N}_0$.

Dans ce cas, Z est la limite de $(Z_n)_{n \in \mathbb{N}_0}$ en norme L^1 .

Démonstration : $1 \Rightarrow 2$. En combinant les Remarques 2.5.3 et 2.7.3, on sait que $(Z_n)_{n \in \mathbb{N}_0}$ converge presque sûrement vers une variable aléatoire Z telle que $\mathbb{E}[|Z|] < +\infty$. Comme la convergence presque sûre implique la convergence en probabilité, le Lemme 2.7.5 implique la convergence en norme L^1 vers Z .

$2 \Rightarrow 3$. Soit Z la limite en norme L^1 de $(Z_n)_{n \in \mathbb{N}_0}$. Pour tout $A \in \mathcal{F}$, on a $\mathbb{E}[Z_n \mathbf{1}_A] \rightarrow \mathbb{E}[Z \mathbf{1}_A]$ puisque

$$|\mathbb{E}[Z_n \mathbf{1}_A - Z \mathbf{1}_A]| \leq \mathbb{E}[|Z_n \mathbf{1}_A - Z \mathbf{1}_A|] \leq \mathbb{E}[|Z_n - Z|] \rightarrow 0.$$

Si $A \in \mathcal{F}_n$, comme $(Z_n)_{n \in \mathbb{N}_0}$ est une martingale, on a pour tout $m > n$

$$\mathbb{E}[Z_n \mathbf{1}_A] = \mathbb{E}[\mathbb{E}[Z_m | \mathcal{F}_n] \mathbf{1}_A] = \mathbb{E}[\mathbb{E}[Z_m \mathbf{1}_A | \mathcal{F}_n]] = \mathbb{E}[Z_m \mathbf{1}_A].$$

La suite $(\mathbb{E}[Z_n \mathbf{1}_A])_n$ est donc constante et égale à sa limite $\mathbb{E}[Z \mathbf{1}_A]$. Autrement dit, on a

$$\mathbb{E}[Z_n \mathbf{1}_A] = \mathbb{E}[Z \mathbf{1}_A]$$

pour tout $A \in \mathcal{F}_n$, ce qui est exactement la définition de l'espérance conditionnelle, puisque Z_n est \mathcal{F}_n -mesurable. On obtient donc $Z_n = \mathbb{E}[Z | \mathcal{F}_n]$.

$3 \Rightarrow 1$. Il suffit d'appliquer la Proposition 2.7.4. ■

Chapitre 3

Processus stochastiques à temps continu

Dans cette section, nous nous intéressons à deux classes de processus stochastiques à temps continu. Dans un premier temps, nous étudions les processus admettant une modification dont les trajectoires sont presque sûrement continues. Cela nous permet de redéfinir la loi du processus sur $C([0, +\infty[, \mathbb{R})$. Nous nous intéressons ensuite à une classe très importante de processus, les processus gaussiens.

3.1 Régularité des trajectoires

A partir de maintenant, nous travaillons avec des processus à temps continu sur un espace probabilisé $(\Omega, \mathcal{F}, \mathbb{P})$ fixé et indexés par un intervalle $I \subseteq \mathbb{R}$. En général, on prendra $I = [0, +\infty[$. Nous avons vu que si $X = \{X(t) : t \in I\}$ est un processus stochastique sur $(\Omega, \mathcal{F}, \mathbb{P})$, alors l'application

$$X : (\Omega, \mathcal{F}) \rightarrow (\mathbb{R}^I, \mathbb{B}^I) : \omega \mapsto X(\cdot, \omega)$$

est une variable aléatoire, c'est-à-dire est mesurable. Cependant, si le processus présente de bonnes propriétés trajectoires, on peut améliorer l'espace d'arrivée, avec par exemple de meilleures propriétés topologiques pour cet espace. En particulier, si X est à trajectoires continues, alors X est à valeurs dans $\mathcal{C}(I, \mathbb{R})$.

Supposons que X est un processus stochastique indexé par $I = [0, +\infty[$ et à trajectoires continues. On munit $\mathcal{C}([0, +\infty[, \mathbb{R})$ de la topologie donnée par la convergence uniforme sur tous les compacts. Autrement dit, une base de topologie de $\mathcal{C}([0, +\infty[, \mathbb{R})$ est donnée par les semi-boules

$$b_n(f, \varepsilon) = \left\{ g \in \mathcal{C}([0, +\infty[, \mathbb{R}) : \sup_{t \in [0, n]} |f(t) - g(t)| < \varepsilon \right\}$$

pour $n \in \mathbb{N}_0$, $f \in \mathcal{C}([0, +\infty[, \mathbb{R})$ et $\varepsilon > 0$. On peut montrer que cette topologie est métrisable. Sur $\mathcal{C}([0, +\infty[, \mathbb{R})$, on peut considérer deux σ -algèbres : celle induite par $\mathbb{B}^{[0, +\infty[}$ et la σ -algèbre de Borel $\mathcal{B}(\mathcal{C}([0, +\infty[, \mathbb{R}))$. Le résultat suivant montre que ces deux σ -algèbres coïncident.

Lemme 3.1.1 On a

$$\mathbb{B}^{[0, +\infty[} \cap \mathcal{C}([0, +\infty[, \mathbb{R}) = \mathcal{B}(\mathcal{C}([0, +\infty[, \mathbb{R})).$$

Démonstration : Considérons les projections $\Pi_t : \mathcal{C}([0, +\infty[, \mathbb{R}) \rightarrow \mathbb{R} : f \mapsto f(t)$ ¹ et montrons qu'elles sont continues pour tout $t \in [0, +\infty[$. Soient $y \in \mathbb{R}$ et $\varepsilon > 0$. Fixons $f \in \Pi_t^{-1}(]y - \varepsilon, y + \varepsilon[)$. Alors il existe $r > 0$ tel que $|f(t) - y| < \varepsilon - r$. Si $n \geq t$, on a clairement $b_n(f, r) \subseteq \Pi_t^{-1}(]y - \varepsilon, y + \varepsilon[)$, ce qui suffit. Donc $\Pi_t : \mathcal{C}([0, +\infty[, \mathbb{R}) \rightarrow \mathbb{R}$ est continue, et par conséquent mesurable pour la σ -algèbre $\mathcal{B}(\mathcal{C}([0, +\infty[, \mathbb{R}))$. Soit C un cylindre de $\mathbb{B}^{[0, +\infty[}$. Alors il existe $t_1, \dots, t_n \geq 0$ et $B_1, \dots, B_n \in \mathbb{B}$ tels que

$$C \cap \mathcal{C}([0, +\infty[, \mathbb{R}) = \bigcap_{j=1}^n \Pi_{t_j}^{-1}(B_j).$$

Comme $\Pi_{t_j}^{-1}(B_j) \in \mathcal{B}(\mathcal{C}([0, +\infty[, \mathbb{R}))$ par mesurabilité de Π_{t_j} , on en tire que $C \cap \mathcal{C}([0, +\infty[, \mathbb{R}) \in \mathcal{B}(\mathcal{C}([0, +\infty[, \mathbb{R}))$. Comme les cylindres engendrent $\mathbb{B}^{[0, +\infty[}$, on obtient l'inclusion

$$\mathbb{B}^{[0, +\infty[} \cap \mathcal{C}([0, +\infty[, \mathbb{R}) \subseteq \mathcal{B}(\mathcal{C}([0, +\infty[, \mathbb{R})).$$

Pour la deuxième inclusion, montrons que toute semi-boule $b_n(f, \varepsilon)$ appartient à $\mathbb{B}^{[0, +\infty[}$. Remarquons que par continuité, on a

$$\begin{aligned} b_n(f, \varepsilon) &= \bigcup_{m \geq 1} \left\{ g \in \mathcal{C}([0, +\infty[, \mathbb{R}) : \sup_{t \in [0, n]} |f(t) - g(t)| \leq \varepsilon - \frac{1}{m} \right\} \\ &= \bigcup_{m \geq 1} \bigcap_{t \in [0, n] \cap \mathbb{Q}} \left\{ g \in \mathcal{C}([0, +\infty[, \mathbb{R}) : |f(t) - g(t)| \leq \varepsilon - \frac{1}{m} \right\} \\ &= \bigcup_{m \geq 1} \bigcap_{t \in [0, n] \cap \mathbb{Q}} \Pi_t^{-1}(]f(t) - \varepsilon + \frac{1}{m}, f(t) + \varepsilon - \frac{1}{m}[) \\ &= \bigcup_{m \geq 1} \bigcap_{t \in [0, n] \cap \mathbb{Q}} \pi_t^{-1}(]f(t) - \varepsilon + \frac{1}{m}, f(t) + \varepsilon - \frac{1}{m}[) \cap \mathcal{C}([0, +\infty[, \mathbb{R}). \end{aligned}$$

Ainsi, $b_n(f, \varepsilon) \in \mathbb{B}^{[0, +\infty[} \cap \mathcal{C}([0, +\infty[, \mathbb{R})$ et comme les boules engendrent la σ -algèbre de Borel², cela suffit. ■

Proposition 3.1.2 Si $X = \{X(t) : t \in [0, +\infty[\}$ est un processus stochastique à valeurs dans $\mathcal{C}([0, +\infty[, \mathbb{R})$, alors l'application

$$X : (\Omega, \mathcal{F}) \rightarrow (\mathcal{C}([0, +\infty[, \mathbb{R}), \mathcal{B}(\mathcal{C}([0, +\infty[, \mathbb{R}))) : \omega \mapsto X(\cdot, \omega)$$

est mesurable et \mathbb{P}_X définit une mesure sur $(\mathcal{C}([0, +\infty[, \mathbb{R}), \mathcal{B}(\mathcal{C}([0, +\infty[, \mathbb{R})))$.

Remarque 3.1.3 Soit $X = \{X(t) : t \in [0, +\infty[\}$ est un processus stochastique à valeurs dans $\mathcal{C}([0, +\infty[, \mathbb{R})$. Considérons une application continue $\varphi : \mathcal{C}([0, +\infty[, \mathbb{R}) \rightarrow \mathbb{R}$, donc mesurable par rapport aux σ -algèbres de Borel de $\mathcal{C}([0, +\infty[, \mathbb{R})$ et de \mathbb{R} . Alors

$$\varphi \circ X : (\Omega, \mathcal{F}) \mapsto (\mathbb{R}, \mathbb{B})$$

1. Nous avons utilisé la notation Π_t pour marquer la différence avec les projections π_t qui sont définies sur \mathbb{R}^I .
2. car $\mathcal{C}([0, +\infty[, \mathbb{R})$ est un espace métrique séparable.

est mesurable. On peut donc considérer les probabilités du type $\mathbb{P}[\varphi \circ X \in B]$ pour tout $B \in \mathbb{B}$. Par exemple, on peut prendre la fonction $\varphi(f) = \sup_{t \in [0,1]} |f(t)|$ et s'intéresser à la probabilité $\mathbb{P}[\sup_{x \in [0,1]} |X(t)| \leq a]$ pour $a > 0$.

Remarque 3.1.4 Une fonction $f : I \rightarrow \mathbb{R}$ est *càdlàg* (continu à droite, limite à gauche) si pour tout $t_0 \in I$, la limite à gauche $\lim_{t \rightarrow t_0, t \leq t_0} f(t)$ existe et la limite à droite satisfait $\lim_{t \rightarrow t_0, t \geq t_0} f(t) = f(t_0)$ (on a la continuité à droite). L'espace des fonctions càdlàg est appelé l'espace de *Skorokhod*. On impose classiquement aux trajectoires de processus stochastiques à temps continus d'être càdlàg afin d'assurer l'existence de probabilités "naturelles".

Il est donc intéressant de chercher à savoir si un processus a de bonnes propriétés de régularité. Notons qu'au vu de l'exemple 1.3.4, des versions d'un processus stochastiques n'ont pas toujours la même régularité. Le résultat suivant donne une condition pour l'existence d'une version continue d'un processus stochastique. En général, on travaillera alors avec ces versions.

Théorème 3.1.5 (Kolmogorov-Čentsov) *Soit X un processus stochastique indexé par un intervalle $I \subseteq \mathbb{R}$. S'il existe $\alpha > 0$, $\beta > 0$ et une constante $C > 0$ vérifiant*

$$\mathbb{E}[|X(t) - X(s)|^\alpha] < C|t - s|^{1+\beta}$$

pour tous $s, t \in I$, alors il existe une version \tilde{X} de X dont les trajectoires sont localement Höldériennes d'exposant γ pour tout $\gamma \in]0, \beta/\alpha[$, c'est-à-dire pour tout $\omega \in \Omega$ et pour tout compact $K \subseteq I$, il existe une constante $C_\gamma(\omega) > 0$ telle que

$$|\tilde{X}(t, \omega) - \tilde{X}(s, \omega)| \leq C_\gamma(\omega)|t - s|^\gamma, \quad \forall t, s \in K.$$

En particulier, \tilde{X} est une version continue de X à trajectoires continues.

Démonstration : Supposons par simplicité que $I = [0, 1]$; si ce n'est pas le cas, on écrit I comme une union dénombrable d'intervalles fermés et on utilise la même construction sur chaque intervalle.

Commençons par construire \tilde{X} sur l'ensemble D des dyadiques de $[0, 1]$, c'est-à-dire sur l'ensemble

$$D = \left\{ \frac{k}{2^j} : j \in \mathbb{N}_0, k \in \{0, \dots, 2^j - 1\} \right\}.$$

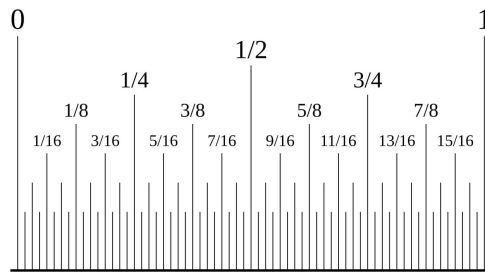


FIGURE 3.1 – Dyadiques de $[0, 1]$ pour $j \leq 4$.

Fixons $\gamma \in]0, \beta/\alpha[$. En utilisant l'inégalité de Markov et l'hypothèse, on a

$$\begin{aligned}
 & \mathbb{P} \left[\exists k \in \{0, \dots, 2^j - 1\} : \left| X \left(\frac{k+1}{2^j} \right) - X \left(\frac{k}{2^j} \right) \right| > \frac{1}{2^{j\gamma}} \right] \\
 &= \mathbb{P} \left[\bigcup_{k \in \{0, \dots, 2^j - 1\}} \left\{ \left| X \left(\frac{k+1}{2^j} \right) - X \left(\frac{k}{2^j} \right) \right| > \frac{1}{2^{j\gamma}} \right\} \right] \\
 &\leq \sum_{k=0}^{2^j-1} \mathbb{P} \left[\left| X \left(\frac{k+1}{2^j} \right) - X \left(\frac{k}{2^j} \right) \right|^\alpha > \frac{1}{2^{j\alpha\gamma}} \right] \\
 &\leq \sum_{k=0}^{2^j-1} 2^{j\alpha\gamma} C \left(\frac{1}{2^j} \right)^{1+\beta} \\
 &= C 2^{j(\alpha\gamma-\beta)}.
 \end{aligned}$$

Comme $\gamma < \frac{\beta}{\alpha}$, la série

$$\sum_{j=1}^{+\infty} C 2^{j(\alpha\gamma-\beta)}$$

converge et le lemme de Borel-Cantelli implique que

$$\mathbb{P} \left[\limsup_{j \rightarrow +\infty} \left\{ \exists k \in \{0, \dots, 2^j - 1\} : \left| X \left(\frac{k+1}{2^j} \right) - X \left(\frac{k}{2^j} \right) \right| > \frac{1}{2^{j\gamma}} \right\} \right] = 0.$$

Cela signifie qu'il existe un événement $\Omega_\gamma \subseteq \Omega$ de probabilité 1 tel que

$$\forall \omega \in \Omega_\gamma \exists J(\omega) \in \mathbb{N} \text{ tel que } \left| X \left(\frac{k+1}{2^j}, \omega \right) - X \left(\frac{k}{2^j}, \omega \right) \right| \leq \frac{1}{2^{j\gamma}}, \forall j \geq J(\omega), \forall k \in \{0, \dots, 2^j - 1\}.$$

Par conséquent, pour tout $\omega \in \Omega_\gamma$, puisqu'il y a un nombre fini de termes pour $j < J(\omega)$, il existe une constante $C_\gamma(\omega) > 0$ telle que

$$\left| X \left(\frac{k+1}{2^j}, \omega \right) - X \left(\frac{k}{2^j}, \omega \right) \right| \leq C_\gamma(\omega) \frac{1}{2^{j\gamma}}, \forall j \in \mathbb{N}_0, \forall k \in \{0, \dots, 2^j - 1\}.$$

Montrons à présent que pour tout $\omega \in \Omega_\gamma$, il existe une constante $C'_\gamma(\omega) > 0$ telle que pour tous $s, t \in D$, on a également

$$|X(t, \omega) - X(s, \omega)| \leq C'_\gamma(\omega) |t - s|^\gamma.$$

Supposons que $t > s$ et considérons $p \in \mathbb{N}$ tel que

$$\frac{1}{2^{p+1}} \leq t - s < \frac{1}{2^p}.$$

Puisque $t - s < \frac{1}{2^p}$, il existe $k \in \{0, \dots, 2^p - 1\}$ tel que $t, s \in [k/2^p, k+2/2^p[$ (ou $[k/2^p, k+1/2^p[$ si $k = 2^p - 1$). On peut alors écrire t et s sous la forme

$$t = \frac{k + \varepsilon_0}{2^p} + \frac{\varepsilon_1}{2^{p+1}} + \frac{\varepsilon_2}{2^{p+2}} + \dots + \frac{\varepsilon_n}{2^{p+n}}$$

et

$$s = \frac{k}{2^p} + \frac{\varepsilon'_1}{2^{p+1}} + \frac{\varepsilon'_2}{2^{p+2}} + \cdots + \frac{\varepsilon'_n}{2^{p+n}}$$

où $\varepsilon_0, \varepsilon_1, \varepsilon_2, \dots, \varepsilon_n, \varepsilon'_1, \varepsilon'_2, \dots, \varepsilon'_n \in \{0, 1\}$ et $n \in \mathbb{N}_0$. Pour tout $j \in \{0, \dots, n\}$, posons

$$t_j = \frac{k + \varepsilon_0}{2^p} + \frac{\varepsilon_1}{2^{p+1}} + \frac{\varepsilon_2}{2^{p+2}} s + \cdots + \frac{\varepsilon_j}{2^{p+j}}$$

et

$$s_j = \frac{k}{2^p} + \frac{\varepsilon'_1}{2^{p+1}} + \frac{\varepsilon'_2}{2^{p+2}} + \cdots + \frac{\varepsilon'_j}{2^{p+j}}.$$

Alors, sur Ω_γ , on a

$$\begin{aligned} |X(t) - X(s)| &= |X(t_n) - X(s_n)| \\ &\leq |X(t_0) - X(s_0)| + \sum_{j=0}^{n-1} |X(t_{j+1}) - X(t_j)| + \sum_{j=0}^{n-1} |X(s_{j+1}) - X(s_j)| \\ &\leq C_\gamma(\omega) 2^{-\gamma p} + 2C_\gamma(\omega) \sum_{j=0}^{n-1} 2^{-\gamma(p+j+1)} \\ &\leq 2^{-\gamma p+1} C_\gamma(\omega) \sum_{j=0}^{+\infty} 2^{-\gamma j} \\ &= 2^{-\gamma p+1} C_\gamma(\omega) \frac{2^\gamma}{2^\gamma - 1} \\ &\leq |t - s|^\gamma 2^{\gamma+1} C_\gamma(\omega) \frac{2^\gamma}{2^\gamma - 1} \end{aligned}$$

puisque $|t - s| \geq \frac{1}{2^{p+1}}$.

Considérons une suite croissante $(\gamma_n)_{n \in \mathbb{N}}$ de $]0, \beta/\alpha[$ qui converge vers β/α , et notons Ω_1 l'événement de probabilité 1

$$\Omega_1 = \bigcap_{n \in \mathbb{N}} \Omega_{\gamma_n}.$$

Par construction, sur Ω_1 , les trajectoires de X sont Höldériennes sur D d'exposant γ pour tout $\gamma \in]0, \beta/\alpha[$.

Il reste à présent à construire \tilde{X} . On procède de la manière suivante : Considérons tout d'abord $\omega \in \Omega_1$. Si $t \in D$, on pose $\tilde{X}(t, \omega) = X(t, \omega)$. Si $t \notin D$, on considère une suite $(t_n)_{n \in \mathbb{N}}$ de D qui converge vers t et on pose

$$\tilde{X}(t, \omega) = \lim_{n \rightarrow +\infty} X(t_n, \omega).$$

Sur l'ensemble $\Omega \setminus \Omega_1$ de probabilité 0, on pose $\tilde{X}(t, \omega) = 0$ pour tout t . Par construction, le processus \tilde{X} est bien à trajectoires Höldériennes d'exposant γ pour tout $\gamma \in]0, \beta/\alpha[$. Pour conclure, il suffit de montrer que \tilde{X} est une version de X . Si $t \in D$, on a $\tilde{X}(t) = X(t)$ sur l'événement Ω_1 de probabilité 1. Si $t \notin D$, il existe une suite $(t_n)_{n \in \mathbb{N}}$ de D qui converge vers t telle que $\tilde{X}(t, \omega) = \lim_{n \rightarrow +\infty} X(t_n, \omega)$. Pour tout $\varepsilon > 0$, on a

$$\mathbb{P}\left[|X(t_n) - X(t)| > \varepsilon\right] \leq \frac{\mathbb{E}\left[|X(t_n) - X(t)|^\alpha\right]}{\varepsilon^\alpha} \leq \frac{|t_n - t|^{1+\beta}}{\varepsilon^\alpha}$$

et donc $X(t_n)$ converge vers $X(t)$ en probabilité. Il s'ensuit qu'il existe une sous suite de $X(t_n)$ qui converge presque sûrement vers $X(t)$. Sur un événement de probabilité 1, par unicité de la limite, on a donc $\tilde{X}(t) = X(t)$. ■

Remarque 3.1.6 Les versions continues données par le Théorème 3.1.5 de Kolmogorov-Čentsov sont indistinguables par la Proposition 1.3.5.

3.2 Processus gaussiens

Dans cette section, nous présentons et étudions une classe importante de processus stochastiques.

Définition 3.2.1 Un processus stochastique $X = \{X(t) : t \in I\}$ est *gaussien* si toutes ses lois fini-dimensionnelles sont gaussiennes. Vu la caractérisation des vecteurs gaussiens, X est gaussien si et seulement si toute combinaison linéaire

$$a_1 X(t_1) + \dots + a_p X(t_p), \quad p \in \mathbb{N}_0, a_1, \dots, a_p \in \mathbb{R}, t_1, \dots, t_p \in I$$

de ses marginales suit une loi normale.

Puisque la loi d'un vecteur gaussien est déterminée par son vecteur moyenne et sa matrice de covariance, toutes les lois fini-dimensionnelles d'un processus gaussien X (et donc la loi du processus par la Proposition 1.2.9) sont connues dès que l'on se donne la *fonction de moyenne*

$$m(t) = \mathbb{E}[X(t)]$$

et l'*opérateur de covariance*

$$K(t, s) = \text{Cov}(X(t), X(s)).$$

En effet, la loi fini-dimensionnelle de $(X(t_1), \dots, X(t_p))$ est alors la loi gaussienne $\mathcal{N}(\mathbf{m}, \mathbf{K})$ de dimension p où $\mathbf{m} = (m(t_1), \dots, m(t_p))$ et $\mathbf{K} = (K(t_i, t_j))_{i,j \in \{1, \dots, p\}}$. En particulier, on a le résultat suivant.

Proposition 3.2.2 Soit X un processus gaussien d'opérateur de covariance K . Alors K est symétrique (i.e. $K(s, t) = K(t, s)$ pour tous $s, t \in I$) et K est de type positif, i.e. pour tous $p \in \mathbb{N}_0$ et tous $t_1, \dots, t_p \in I$, la matrice $(K(t_i, t_j))_{i,j \in \{1, \dots, p\}}$ est semi-définie positive.

La réciproque de ce résultat est donnée ci-dessous. Nous ne démontrerons pas ce résultat puisque la preuve est basée sur un résultat admis.

Proposition 3.2.3 Soient $K : I \times I \rightarrow \mathbb{R}$ un opérateur symétrique de type positif et $m : I \rightarrow \mathbb{R}$. Alors il existe un processus gaussien de fonction moyenne m et d'opérateur de covariance K .

Démonstration : Il s'agit d'une application du Théorème de Kolmogorov 1.4.2; nous renvoyons le lecteur à [2, Théorème 2.1] par exemple pour le détail. ■

Des bonnes conditions pour avoir une version assez régulière d'un processus gaussien sont données dans le résultat suivant, conséquence du Théorème de Kolmogorov-Čentsov. Notons qu'on dit qu'un processus gaussien est *centré* si sa fonction moyenne est identiquement nulle.

Proposition 3.2.4 (Kolmogorov-Čentsov gaussien) Soit X un processus gaussien centré d'opérateur de covariance K . S'il existe $\varepsilon > 0$ et $C > 0$ tels que

$$K(t, t) + K(s, s) - 2K(t, s) \leq C|t - s|^\varepsilon$$

pour tous $t, s \in I$, alors il existe une version \tilde{X} de X dont les trajectoires sont localement Höldériennes d'exposant γ pour tout $\gamma \in]0, \varepsilon/2[$. En particulier, \tilde{X} est une version continue de X à trajectoires continues.

Démonstration : Remarquons que

$$\mathbb{E}\left[|X(t) - X(s)|^2\right] = K(t, t) + K(s, s) - 2K(t, s) \leq C|t - s|^\varepsilon.$$

Malheureusement, on ne peut pas appliquer directement le Théorème de Kolmogorov-Čentsov puisque l'inégalité $\varepsilon > 1$ n'est pas garantie. Néanmoins, rappelons que si Z suit une loi normale centrée, on a³

$$\mathbb{E}[Z^{2m}] = \frac{(2m!)}{2^m m!} \text{var}[Z]^m$$

pour tout $m \in \mathbb{N}_0$. Par conséquent, il vient

$$\mathbb{E}\left[|X(t) - X(s)|^{2m}\right] \leq C^m \frac{(2m!)}{2^m m!} |t - s|^{m\varepsilon}$$

et en choisissant m tel que $m\varepsilon > 1$, le Théorème de Kolmogorov-Čentsov nous donne une version de X qui est localement Höldérienne d'exposant γ pour tout $\gamma < \frac{m\varepsilon - 1}{2m}$. En prenant la limite pour m tendant vers l'infini, on obtient des versions Höldériennes d'exposant γ pour tout $\gamma < \frac{\varepsilon}{2}$. ■

Exemple 3.2.5 Parmi les opérateurs de covariances K , on peut considérer l'opérateur défini par

$$K(t, s) = \frac{1}{2}(s^{2H} + t^{2H} - |s - t|^{2H}).$$

Le processus gaussien centré associé s'appelle le *mouvement Brownien fractionnaire*. On a

$$K(t, t) + K(s, s) - 2K(t, s) = |s - t|^{2H}$$

et donc le mouvement Brownien fractionnaire admet une version dont les trajectoires sont presque sûrement Höldériennes d'exposant γ pour tout $\gamma \in]0, H[$. Si $H = 1/2$, l'opérateur prend la forme

$$K(t, s) = \frac{1}{2}(s + t - |s - t|) = \min(t, s)$$

et on parle de *mouvement Brownien*. Nous étudierons ce processus dans le prochain chapitre.

3.3 Propriétés en loi de processus stochastiques

Dans cette dernière section, nous présentons d'autres propriétés en loi classiques et importantes que peuvent avoir des processus stochastiques.

3. voir (B.4).

Définition 3.3.1 Un processus stochastique $X = \{X(t) : t \geq 0\}$ est

— *stationnaire* (strict) si pour tout $h > 0$, on a

$$\{X(t+h) : t \geq 0\} \stackrel{\mathcal{L}}{=} \{X(t) : t \geq 0\},$$

c'est-à-dire pour tout h , tout $n \in \mathbb{N}_0$ et $t_1, \dots, t_n \geq 0$, on a

$$(X(t_1+h), \dots, X(t_n+h)) \stackrel{\mathcal{L}}{=} (X(t_1), \dots, X(t_n)).$$

— à *accroissements stationnaires* si pour tout $h > 0$, la loi des accroissements $X(t+h) - X(t)$ ne dépend pas de $t \geq 0$, c'est-à-dire si

$$X(t+h) - X(t) \stackrel{\mathcal{L}}{=} X(h) - X(0).$$

— à *accroissements indépendants* si pour tout $p \geq 2$ et tous $0 \leq t_1 < t_2 < \dots < t_p$, les variables aléatoires

$$X(t_2) - X(t_1), \dots, X(t_p) - X(t_{p-1})$$

sont indépendantes.

Exemple 3.3.2 Le processus le plus simple à accroissements indépendants et stationnaires est le processus de Poisson $\{N(t) : t \geq 0\}$ de paramètre $\lambda > 0$.

Dans le chapitre suivant, nous introduirons un autre processus à accroissements indépendants et stationnaires : le mouvement brownien.

La stationnarité d'un processus gaussien peut être facilement caractérisée.

Proposition 3.3.3 *Un processus gaussien X est stationnaire si et seulement si sa fonction moyenne $m(t) = \mathbb{E}[X(t)]$ est constante et son opérateur de covariance est de la forme $K(s-t)$.*

Démonstration : Les deux conditions sont clairement des conditions nécessaires de stationnarité (que le processus soit gaussien ou pas). En effet, par définition $\mathcal{L}(X(t)) = \mathcal{L}(X(s))$ pour tous s, t , donc $\mathbb{E}[X(t)] = \mathbb{E}[X(s)]$ et la fonction moyenne est constante. De même, $\mathcal{L}(X(t), X(s)) = \mathcal{L}(X(t+h), X(s+h))$ pour tous s, t, h , et on en tire donc que $\text{Cov}(X(t), X(s)) = \text{Cov}(X(t+h), X(s+h))$. Cela implique que la covariance ne peut dépendre que de la différence $t - s$.

Montrons que les deux conditions sont également suffisantes dans le cas gaussien. En effet, dans ce cas, la loi est caractérisée par sa fonction de moyenne et sa covariance. Ainsi, si $\mathbb{E}[X(t)]$ est constante et si $K(s, t) = K(s - t)$, alors la loi ne peut pas être modifiée par des translations. ■

Exemple 3.3.4 Un bruit blanc gaussien est un processus gaussien stationnaire dont l'opérateur de covariance est de la forme $K(t, s) = k(t)\delta(t - s)$. Ainsi, les valeurs prises par le processus à des instants différents sont décorrélés.

Chapitre 4

Définition et construction du mouvement Brownien

“La star des processus, le processus des stars”
S. N.

4.1 Introduction et motivation

Une théorie du *mouvement Brownien* a été postulée au début du 20^{ème} siècle par Albert Einstein (1879 - 1955), d’une part, et par Louis Bachelier (1870 - 1946), d’autre part, en réponse à deux questions de natures en apparence fort éloignées. Einstein cherchait à estimer le nombre d’Avogadro et, ce faisant, à apporter une réponse à une question du botaniste Robert Brown (1773 - 1858) concernant le mouvement irrégulier et aléatoire d’une particule de pollen en suspension dans de l’eau. Bachelier cherchait quant à lui à modéliser les fluctuations observées sur les marchés financiers et à proposer un modèle permettant d’évaluer correctement le prix d’options. Les deux constructions se sont révélées équivalentes, et ont toutes deux dans leurs disciplines été confirmées tant expérimentalement que mathématiquement. Les fondations rigoureuses de la théorie du mouvement Brownien sont dues au mathématicien Norbert Wiener (1894 - 1964). De nombreux autres grands mathématiciens se sont également penchés sur ces questions (Doob, Yor, Levy, Mandelbrot ...) et leurs découvertes ont donné naissance à une des branches les plus importantes des mathématiques contemporaines.

Bien que définissable en toute dimension et sur tout espace, nous allons nous concentrer dans ce cours sur le mouvement Brownien univarié. Les propriétés caractéristiques du mouvement brownien unidimensionnel sont données à la Définition 4.2.1. Comme nous le verrons, ce processus est à la famille des processus stochastiques ce que la loi normale est à la famille des lois de probabilités : un phare. Une bonne compréhension de son comportement s’est avérée cruciale dans presque toutes les branches des sciences tant physiques que sociales. Il s’agit également d’un objet mathématique d’une beauté peu commune. Sa perfection est le reflet d’une loi de la nature plutôt que d’une invention humaine.

Dans ce chapitre, nous donnons la définition du mouvement Brownien. Nous montrons ensuite que celui-ci existe bien. Une première méthode pour montrer l’existence du mouvement Brownien consiste à utiliser le Théorème 1.4.2 d’existence de Kolmogorov. Nous allons plutôt proposer une construction explicite du mouvement Brownien. Celle-ci se base sur une décomposition du processus dans une base particulière, appelée système de Faber-Schauder et lié à la base de Haar.

4.2 Définition du mouvement Brownien

Définition 4.2.1 Un processus stochastique $B = \{B(t) : t \geq 0\}$ à valeurs réelles est un *mouvement Brownien* si

1. $B(0) = 0$ presque sûrement,
2. les accroissements de B sont indépendants,
3. les accroissements de B sont stationnaires et $B(t) - B(s) \sim \mathcal{N}(0, t - s)$ pour tous $t > s \geq 0$,
4. les trajectoires de B sont presque sûrement continues.

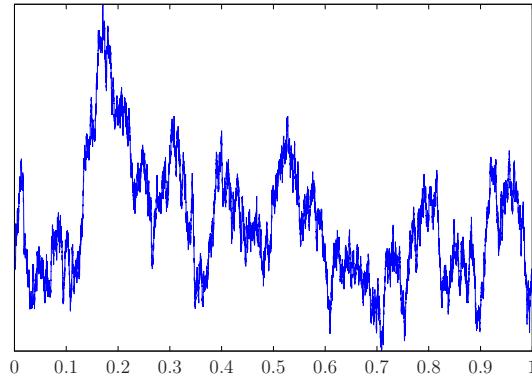


FIGURE 4.1 – Trajectoire d'un mouvement Brownien

Proposition 4.2.2 Soit B un mouvement Brownien.

1. Pour tout $t \geq 0$, $B(t) \sim \mathcal{N}(0, t)$.
2. Le processus B est gaussien et sa fonction de covariance est donnée par

$$K(t, s) = \min(t, s), \quad \forall t, s \geq 0.$$

Démonstration : Le premier point est immédiat puisque $B(t)$ est égal en loi à $B(t) - B(0)$. De plus, B est gaussien car si $n \in \mathbb{N}_0$, $a_1, \dots, a_n \in \mathbb{R}$ et $t_1, \dots, t_n \in [0, +\infty[$, on a

$$\begin{aligned} \sum_{i=1}^n a_i B(t_i) &= \sum_{i=1}^n a_i \left(\sum_{j=1}^i (B(t_j) - B(t_{j-1})) \right) \\ &= \sum_{j=1}^n \left(\sum_{i=j}^n a_i \right) (B(t_j) - B(t_{j-1})) \end{aligned}$$

où on a posé $t_0 = 0$, qui est une somme de gaussiennes indépendantes, donc suit bien une loi normale. Enfin, puisque B est centré, on a $K(t, s) = \mathbb{E}[B(t)B(s)]$. Si $t > s$, on en tire que

$$\begin{aligned} K(t, s) = \mathbb{E}[B(t)B(s)] &= \mathbb{E}[(B(t) - B(s) + B(s))B(s)] \\ &= \mathbb{E}[(B(t) - B(s))(B(s) - B(0))] + \mathbb{E}[B(s)^2] \\ &= 0 + s = \min(t, s) \end{aligned}$$

où on a utilisé l'indépendance des accroissements, le premier point et le fait que $B(0) = 0$ presque sûrement. ■

Proposition 4.2.3 Si $B = \{B(t) : t \geq 0\}$ est un processus gaussien centré d'opérateur de covariance $K(t, s) = \min(t, s)$ pour tous $t, s \geq 0$, alors ses accroissements sont indépendants et satisfont

$$B(t) - B(s) \sim \mathcal{N}(0, t - s), \quad \forall t > s \geq 0.$$

En particulier, si les trajectoires de B sont presque sûrement continues, c'est un mouvement Brownien.

Démonstration : Comme B est un processus gaussien, par la Proposition B.0.9, ses accroissements seront indépendants si et seulement si ils sont non-corrélés. Si $t > s \geq u > v \geq 0$, on a

$$\begin{aligned} \text{Cov}(B(t) - B(s), B(u) - B(v)) &= \mathbb{E}[(B(t) - B(s))(B(u) - B(v))] \\ &= \mathbb{E}[B(t)B(u)] - \mathbb{E}[B(t)B(v)] - \mathbb{E}[B(s)B(u)] + \mathbb{E}[B(s)B(v)] \\ &= u - v - u + v = 0 \end{aligned}$$

ce qui démontre l'indépendances des accroissements. De plus, pour tous $t > s \geq 0$, la variable aléatoire $B(t) - B(s)$ suit une loi normale centrée de variance

$$\begin{aligned} \text{Var}[B(t) - B(s)] &= \mathbb{E}[(B(t) - B(s))^2] \\ &= \mathbb{E}[B(t)^2] + \mathbb{E}[B(s)^2] - 2\mathbb{E}[B(t)B(s)] \\ &= \text{Cov}(B(t), B(t)) + \text{Cov}(B(s), B(s)) - 2\text{Cov}(B(t), B(s)) \\ &= t + s - 2\min(t, s) = t - s \end{aligned}$$

ce qui permet de conclure. ■

4.3 Construction du mouvement Brownien

L'existence du mouvement Brownien pourrait être démontrée en utilisant la Proposition 3.2.3. Ce résultat étant basé sur le Théorème de Kolmogorov 1.4.2, nous préférons dans ce cours proposer une approche constructive du mouvement Brownien. Commençons par introduire une base de $L^2([0, 1])$ qui va nous permettre de décomposer le processus en une série dont les coefficients sont des lois normales indépendantes.

4.3.1 Base de Haar et système de Faber-Schauder

Commençons par rappeler quelques résultats concernant les systèmes orthogonaux de l'espace $L^2([0, 1])$ muni de son produit scalaire usuel

$$\langle f, g \rangle = \int_0^1 f(x)\overline{g(x)}dx.$$

Définition 4.3.1 Soit $(f_n)_{n \in \mathbb{N}}$ une suite orthonormée de $L^2([0, 1])$ (i.e. $\langle f_n, f_m \rangle = \delta_{n,m}$ pour tous $n, m \in \mathbb{N}$). On dit que cette suite est *totale* si pour tout $f \in L^2([0, 1])$, on a

$$\langle f, f_n \rangle = 0 \quad \forall n \in \mathbb{N} \Rightarrow f = 0 \text{ p.p.}$$

Dans ce cas, pour tout $f \in L^2([0, 1])$, on peut écrire

$$f = \sum_{n=0}^{+\infty} \langle f, f_n \rangle f_n$$

où la convergence a lieu dans $L^2([0, 1])$.

Proposition 4.3.2 (Parseval) Si $(f_n)_{n \in \mathbb{N}}$ est une suite orthonormée totale de $L^2([0, 1])$, alors pour tous $g, f \in L^2([0, 1])$, on a

$$\langle f, g \rangle = \sum_{n=0}^{+\infty} \langle f, f_n \rangle \overline{\langle g, f_n \rangle}.$$

En particulier,

$$\|f\|_{L^2([0,1])}^2 = \sum_{n=0}^{+\infty} |\langle f, f_n \rangle|^2.$$

Présentons directement l'intérêt de l'égalité de Parseval dans notre cas. Considérons $f = \mathbf{1}_{[0,t]}$ et $g = \mathbf{1}_{[0,s]}$. Alors

$$\langle \mathbf{1}_{[0,t]}, \mathbf{1}_{[0,s]} \rangle = \int_0^1 \mathbf{1}_{[0,t]}(x) \mathbf{1}_{[0,s]}(x) dx = \min(t, s)$$

et l'égalité de Parseval implique que

$$\min(t, s) = \sum_{n=0}^{+\infty} \langle \mathbf{1}_{[0,t]}, f_n \rangle \overline{\langle \mathbf{1}_{[0,s]}, f_n \rangle} = \sum_{n=0}^{+\infty} \int_0^t f_n(x) dx \int_0^s \overline{f_n(x)} dx \quad (4.1)$$

ce qui donne une écriture alternative de l'opérateur de covariance d'un mouvement Brownien. La suite orthogonale qui va nous intéresser s'appelle la *base de Haar*. Elle consiste à utiliser les translatés et dilatés d'une fonction de base ψ pour remplir tout l'espace. Plus précisément, considérons la fonction

$$\psi(x) = \begin{cases} 1 & \text{si } 0 \leq x < \frac{1}{2} \\ -1 & \text{si } \frac{1}{2} \leq x < 1 \\ 0 & \text{sinon.} \end{cases}$$

Cette fonction est à support égal à $[0, 1]$. Considérons ensuite pour tout $j \in \mathbb{N}$ et tout $k \in \{0, \dots, 2^j - 1\}$, la fonction $\psi_{j,k}$ définie sur \mathbb{R} par

$$\psi_{j,k}(x) = 2^{j/2} \psi(2^j x - k).$$

Proposition 4.3.3 Avec la fonction constante égale à 1, les fonctions $\psi_{j,k}$, $j \in \mathbb{N}$ et $k \in \{0, \dots, 2^j - 1\}$, forment une suite orthonormée totale de $L^2([0, 1])$. On appelle cette suite la base de Haar.

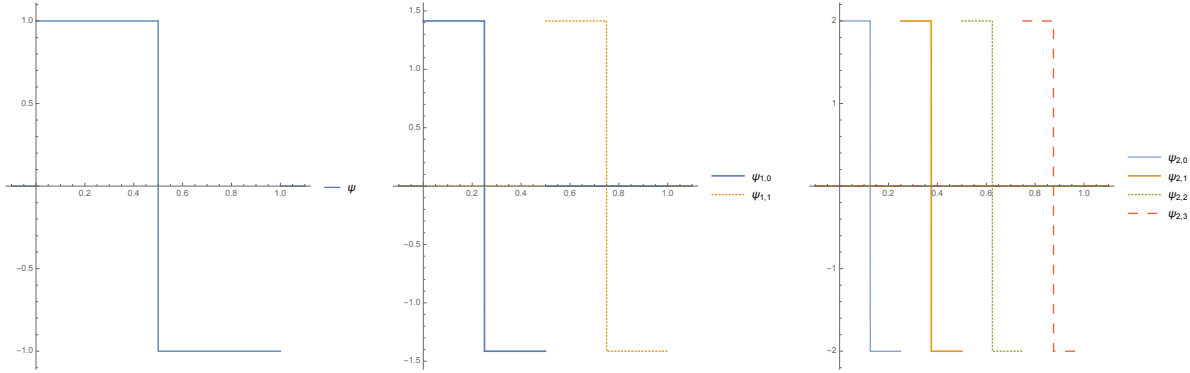


FIGURE 4.2 – Premières fonctions de la base de Haar

Démonstration : Nous esquissons juste les arguments de la preuve. Il n'est pas compliqué de vérifier que les fonctions 1 et $\psi_{j,k}$, $j \in \mathbb{N}$, $k \in \{0, \dots, 2^j - 1\}$ forment une suite orthogonale de $L^2([0, 1])$. Pour montrer que cette suite est totale, on considère $f \in L^2([0, 1])$ tel que $\langle f, 1 \rangle = 0$ et $\langle f, \psi_{j,k} \rangle = 0$ pour tous $j \in \mathbb{N}$ et $k \in \{0, \dots, 2^j - 1\}$. Pour montrer que f est nulle presque partout, l'idée est de considérer la fonction F définie sur $[0, 1]$ par

$$F(y) = \int_0^y f(x) dx.$$

Remarquons que

$$\begin{aligned} 0 &= 2^{-j/2} \int_0^1 f(x) \psi_{j,k}(x) dx = \int_{k2^{-j}}^{(2k+1)2^{-(j+1)}} f(x) dx - \int_{(2k+1)2^{-(j+1)}}^{(k+1)2^{-j}} f(x) dx \\ &= F\left(\frac{2k+1}{2^{j+1}}\right) - F\left(\frac{k}{2^j}\right) - F\left(\frac{k+1}{2^j}\right) + F\left(\frac{2k+1}{2^{j+1}}\right) \\ &= 2F\left(\frac{2k+1}{2^{j+1}}\right) - F\left(\frac{k}{2^j}\right) - F\left(\frac{k+1}{2^j}\right). \end{aligned}$$

Par induction sur j , cette relation donne alors que F s'annule en tous les points dyadiques. De plus, F est continue car par Cauchy-Schwarz, on a

$$|F(y') - F(y)| \leq \left| \int_y^{y'} f(x) dx \right| \leq \|f\|_{L^2([0,1])} \|\mathbf{1}_{[y,y']}\|_{L^2([0,1])} \leq C \sqrt{y' - y}$$

pour tous $0 \leq y \leq y' \leq 1$. Par densité des nombres dyadiques dans $[0, 1]$, on en tire que F est nulle partout. Puisque F est une primitive de f , cela implique que f est constante. Enfin, on a $\int_0^1 f(x) dx = 0$ ce qui permet d'obtenir la conclusion. ■

Au vu de l'égalité (4.1), les fonctions qui vont nous intéresser sont données par les primitives des fonctions $\psi_{j,k}$. On pose

$$\Lambda(x) = \int_0^x \psi(u) du = \begin{cases} x & \text{si } 0 \leq x < \frac{1}{2} \\ 1 - x & \text{si } \frac{1}{2} \leq x < 1 \\ 0 & \text{sinon.} \end{cases} \quad (4.2)$$

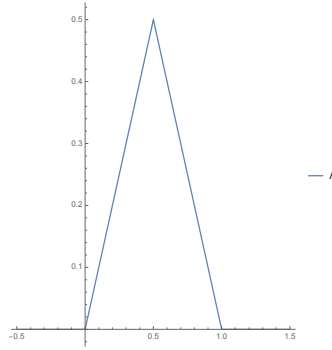


FIGURE 4.3 – Fonction chapeau

Cette fonction est parfois appelée la *fonction chapeau* (ou fonction triangle, fonction tente). Remarquons que pour tous $j \in \mathbb{N}$ et $k \in \{0, \dots, 2^j - 1\}$, on a

$$\int_0^x \psi_{j,k}(u) du = 2^{j/2} \int_0^x \psi(2^j u - k) du = 2^{-j/2} \int_0^{2^j x - k} \psi(v) dv = 2^{-j/2} \Lambda(2^j x - k). \quad (4.3)$$

L'ensemble des fonctions $\Lambda_{j,k}$, $j \in \mathbb{N}$ et $k \in \{0, \dots, 2^j - 1\}$, définies par $\Lambda_{j,k}(x) = 2^{-j/2} \Lambda(2^j x - k)$ forment le *système de Faber-Schauder* (aussi appelé *base de Schauder*). Ces fonctions sont continues et le support de $\Lambda_{j,k}$ est donné par l'intervalle dyadique $[\frac{k}{2^j}, \frac{k+1}{2^j}]$.

4.3.2 Construction du mouvement Brownien

On va construire tout d'abord le mouvement Brownien sur $[0, 1]$, on l'étendra ensuite à $[0, +\infty[$. On considère la fonction chapeau Λ introduite dans la sous-section précédente.

Théorème 4.3.4 Soient $\varepsilon, \varepsilon_{j,k}$, $j \in \mathbb{N}, k \in \{0, \dots, 2^j - 1\}$, des variables aléatoires indépendantes et identiquement distribuées selon une loi $\mathcal{N}(0, 1)$. Pour tout $t \in [0, 1]$, on pose

$$B(t) = \sum_{j=0}^{+\infty} \sum_{k=0}^{2^j-1} \varepsilon_{j,k} 2^{-j/2} \Lambda(2^j t - k) + \varepsilon t. \quad (4.4)$$

Alors $B = \{B(t) : t \in [0, 1]\}$ est un mouvement Brownien sur $[0, 1]$.

Nous allons décomposer la preuve de ce résultat en plusieurs lemmes intermédiaires. Le premier lemme va nous permettre de montrer que la série définie dans (4.4) converge presque sûrement et que la limite est alors continue.

Lemme 4.3.5 Presque sûrement, il existe $J \in \mathbb{N}$ tel que pour tout $j \geq J$,

$$\max_{k \in \{0, \dots, 2^j - 1\}} |\varepsilon_{j,k}| \leq \sqrt{2^j}.$$

Démonstration : Remarquons tout d'abord que si $Z \sim \mathcal{N}(0, 1)$, alors pour tout $a > 0$ on a

$$\begin{aligned} \mathbb{P}[|Z| > a] &= \frac{2}{\sqrt{2\pi}} \int_a^{+\infty} e^{-x^2/2} dx \\ &\leq \frac{2}{\sqrt{2\pi}} \int_a^{+\infty} \frac{x}{a} e^{-x^2/2} dx \\ &= \sqrt{\frac{2}{\pi}} \frac{e^{-a^2/2}}{a}. \end{aligned}$$

Pour tout $j \in \mathbb{N}$, notons A_j l'événement

$$\max_{k \in \{0, \dots, 2^j - 1\}} |\varepsilon_{j,k}| > \sqrt{2^j}.$$

On a

$$\begin{aligned} \mathbb{P}[A_j] &= \mathbb{P}[\exists k \in \{0, \dots, 2^j - 1\} : |\varepsilon_{j,k}| > \sqrt{2^j}] \\ &\leq \sum_{k=0}^{2^j-1} \mathbb{P}[|\varepsilon_{j,k}| > \sqrt{2^j}] \\ &\leq \sum_{k=0}^{2^j-1} \sqrt{\frac{2}{\pi}} \frac{e^{-j}}{\sqrt{2^j}} \\ &= \frac{1}{\sqrt{j\pi}} \left(\frac{2}{e}\right)^j. \end{aligned}$$

Puisque $2 < e$, on en tire que $\mathbb{P}[A_j]$ est le terme général d'une série qui converge. Le lemme de Borel-Cantelli implique alors que

$$\mathbb{P}\left[\limsup_{j \rightarrow +\infty} A_j\right] = 0,$$

d'où

$$\mathbb{P}\left[\liminf_{j \rightarrow +\infty} A_j^c\right] = 1.$$

Cette dernière égalité donne la conclusion. ■

Le résultat suivant nous donne la convergence presque sûre de la série. Cette convergence presque sûre est même uniforme, ce qui permet d'obtenir la continuité des trajectoires limites.

Lemme 4.3.6 *La série (4.4) converge presque sûrement uniformément sur $[0, 1]$. En particulier, le processus limite B est presque sûrement à trajectoires continues.*

Démonstration : Par le lemme 4.3.5, on sait que presque sûrement, il existe $J \in \mathbb{N}$ tel que pour tout $j \geq J$ et tout $k \in \{0, \dots, 2^j - 1\}$, $|\varepsilon_{j,k}| \leq \sqrt{2^j}$. On en tire que

$$\begin{aligned} \sup_{t \in [0,1]} \left| \sum_{k=0}^{2^j-1} \varepsilon_{j,k} 2^{-j/2} \Lambda(2^j t - k) \right| &\leq 2^{-j/2} \sqrt{2^j} \sup_{t \in [0,1]} \left| \sum_{k=0}^{2^j-1} \Lambda(2^j t - k) \right| \\ &\leq \frac{1}{2} 2^{-j/2} \sqrt{2^j} \end{aligned}$$

pour tout $j \geq J$ puisque pour tout $t \in [0, 1]$, il existe un unique $k \in \{0, \dots, 2^j - 1\}$ tel que $\Lambda(2^j t - k) \neq 0$ et pour ce k , $\Lambda(2^j t - k) \leq \frac{1}{2}$. Il s'ensuit que la série (4.4) converge presque sûrement uniformément sur $[0, 1]$ et que dans ce cas, la limite est continue (comme limite uniforme de fonctions continues). ■

Lemme 4.3.7 *Pour tout $t \in [0, 1]$, la série (4.4) converge en probabilité, en loi et en norme L^2 . Le processus limite B est un processus gaussien centré.*

Démonstration : Par le Lemme 4.3.6 et en utilisant le Théorème A.0.5 et la Proposition B.0.2, on obtient toutes les convergences annoncées de la série (4.4). Montrons que sa limite donne un processus gaussien. Soient $n \in \mathbb{N}_0$, $a_1, \dots, a_n \in \mathbb{R}$ et $t_1, \dots, t_n \in [0, 1]$. Montrons que la variable aléatoire

$$a_1 B(t_1) + \dots + a_n B(t_n)$$

est gaussienne. Pour tout $J \in \mathbb{N}$, considérons la variable aléatoire définie par

$$X_J = \sum_{l=1}^n a_l \left(\sum_{j=0}^J \sum_{k=0}^{2^j-1} \varepsilon_{j,k} 2^{-j/2} \Lambda(2^j t_l - k) \right).$$

Cette variable suit une loi normale centrée de variance

$$\sigma_J^2 = \left(\sum_{j=0}^J \sum_{k=0}^{2^j-1} \left(\sum_{l=1}^n a_l \Lambda(2^j t_l - k) \right)^2 2^{-j} \right) \leq \frac{n}{4} \left(\sum_{l=1}^n a_l \right)^2 \sum_{j=0}^J 2^{-j}$$

puisque'il y a au plus n termes non-nuls dans la somme sur k , quel que soit j . La suite $(\sigma_J^2)_{J \in \mathbb{N}}$ converge donc vers

$$\sigma^2 = \left(\sum_{j=0}^{+\infty} \sum_{k=0}^{2^j-1} 2^{-j} \left(\sum_{l=1}^n a_l \Lambda(2^j t_l - k) \right)^2 \right) < +\infty$$

et la Proposition B.0.2 implique que la suite des variables $(X_J)_{J \in \mathbb{N}}$ converge en loi vers une loi normale $\mathcal{N}(0, \sigma^2)$. On obtient la conclusion pour $a_1 B(t_1) + \dots + a_n B(t_n)$ puisque'il suffit alors d'ajouter à la série une loi gaussienne indépendante. On obtient également en cas particulier que le processus B est centré. ■

Remarque 4.3.8 Soient $(X_n)_{n \in \mathbb{N}}$ et $(Y_n)_{n \in \mathbb{N}}$ deux suites de variables aléatoires qui convergent en norme L^2 vers X et Y respectivement¹. Alors $(X_n + Y_n)_{n \in \mathbb{N}}$ converge en norme L^2 vers $X + Y$ et on en tire que

$$\mathbb{E}[X_n Y_n] \rightarrow \mathbb{E}[XY]$$

lorsque $n \rightarrow +\infty$. En effet, on a

$$\begin{aligned} \mathbb{E}[X_n Y_n] &= \frac{1}{2} \mathbb{E}[(X_n + Y_n)^2 - X_n^2 - Y_n^2] \\ &\rightarrow \frac{1}{2} \mathbb{E}[(X + Y)^2 - X^2 - Y^2] = \mathbb{E}[XY]. \end{aligned}$$

1. Dans notre cas, il s'agira des sommes partielles de la série définissant B .

Jusqu'à présent, nous avons juste utilisé le fait que la fonction Λ est à support dans $[0, 1]$ et est continue. Le lemme suivant justifie le choix particulier de Λ parmi les fonctions continues à support compact, motivé par (4.3).

Lemme 4.3.9 *L'opérateur de covariance du processus gaussien B défini par (4.4) est donné par*

$$K(t, s) = \min(t, s), \quad \forall t, s \geq 0.$$

Démonstration : Vu l'indépendance de la variable aléatoire ε par rapport aux variables $\varepsilon_{j,k}$, on a

$$\begin{aligned} \text{Cov}(B(t), B(s)) &= \mathbb{E}[(B(t) - \varepsilon t)(B(s) - \varepsilon s)] + \mathbb{E}[\varepsilon t(B(s) - \varepsilon s)] + \mathbb{E}[\varepsilon s(B(t) - \varepsilon t)] + \mathbb{E}[\varepsilon^2 st] \\ &= \mathbb{E}[(B(t) - \varepsilon t)(B(s) - \varepsilon s)] + \mathbb{E}[\varepsilon^2 st] \\ &= \mathbb{E}[(B(t) - \varepsilon t)(B(s) - \varepsilon s)] + st \end{aligned}$$

Puisque la série définissant B converge en norme L^2 , vu la remarque 4.3.8 et la linéarité de l'espérance, on obtient

$$\begin{aligned} \text{Cov}(B(t), B(s)) &= \mathbb{E} \left[\sum_{j=0}^{+\infty} \sum_{j'=0}^{+\infty} \sum_{k=0}^{2^j-1} \sum_{k'=0}^{2^{j'}-1} \varepsilon_{j,k} 2^{-j/2} \varepsilon_{j',k'} 2^{-j'/2} \Lambda(2^j t - k) \Lambda(2^{j'} s - k') \right] + st \\ &= \left(\sum_{j=0}^{+\infty} \sum_{j'=0}^{+\infty} \sum_{k=0}^{2^j-1} \sum_{k'=0}^{2^{j'}-1} \mathbb{E}[\varepsilon_{j,k} \varepsilon_{j',k'}] 2^{-j/2} 2^{-j'/2} \Lambda(2^j t - k) \Lambda(2^{j'} s - k') \right) + st \\ &= \left(\sum_{j=0}^{+\infty} \sum_{k=0}^{2^j-1} 2^{-j} \Lambda(2^j t - k) \Lambda(2^j s - k) \right) + st \\ &= \left(\sum_{j=0}^{+\infty} \sum_{k=0}^{2^j-1} \int_0^t \psi_{j,k}(x) dx \int_0^s \psi_{j,k}(x) dx \right) + \int_0^t 1 dx \int_0^s 1 dx \\ &= \left(\sum_{j=0}^{+\infty} \sum_{k=0}^{2^j-1} \langle \mathbf{1}_{[0,t[}, \psi_{j,k} \rangle \langle \mathbf{1}_{[0,s[}, \psi_{j,k} \rangle \right) + \langle \mathbf{1}_{[0,t[}, 1 \rangle \langle \mathbf{1}_{[0,s[}, 1 \rangle \\ &= \langle \mathbf{1}_{[0,t[}, \mathbf{1}_{[0,s[} \rangle \end{aligned}$$

où on a utilisé l'indépendance des variables $\varepsilon_{j,k}$, (4.3) et où la dernière égalité résulte de l'égalité de Parseval (4.1) pour la base donnée par la Proposition 4.3.3. Puisque

$$\langle \mathbf{1}_{[0,t[}, \mathbf{1}_{[0,s[} \rangle = \min(s, t),$$

on obtient la conclusion. ■

Démonstration du Théorème 4.3.4 : La proposition 4.2.3 et les lemmes 4.3.6, 4.3.7 et 4.3.9 prouvent le résultat. ■

Maintenant que nous avons construit le mouvement Brownien sur $[0, 1]$, nous pouvons l'étendre à $[0, +\infty[$ en considérant sur chaque intervalle $[n, n+1]$ une copie indépendante du mouvement

Brownien sur $[0, 1]$, en commençant chacun d'entre eux là où le précédent s'est terminé : on considère donc une famille $B^{(n)}$, $n \in \mathbb{N}$, de mouvements Browniens indépendants² sur $[0, 1]$. Pour tout $t \geq 0$, on pose ensuite

$$B(t) = \sum_{k=0}^{n-1} B^{(k)}(1) + B^{(n)}(t - n) \quad \text{si } t \in [n, n + 1[.$$

Autrement dit,

$$B(t) = \sum_{k=0}^{\lfloor t \rfloor - 1} B^{(k)}(1) + B^{(\lfloor t \rfloor)}(\{t\}),$$

où $\lfloor t \rfloor$ est la partie entière de t et $\{t\} = t - \lfloor t \rfloor$ sa partie fractionnaire. En utilisant la Proposition 4.2.3, il est facile de vérifier que B est bien un mouvement Brownien sur $[0, +\infty[$. Cette vérification est laissée en exercice.

2. Deux processus stochastiques X et Y sont indépendants si pour tout $n \in \mathbb{N}$ et tous t_1, \dots, t_n , les vecteurs aléatoires $(X(t_1), \dots, X(t_n))$ et $(Y(t_1), \dots, Y(t_n))$ sont indépendants.

Chapitre 5

Propriétés en loi du mouvement Brownien

5.1 Mesure de Wiener

Soit B un mouvement Brownien sur un espace probabilisé $(\Omega, \mathcal{F}, \mathbb{P})$. Comme les trajectoires du mouvement Brownien sont presque sûrement continues, on obtient une application mesurable

$$B : \Omega^* \rightarrow C([0, +\infty[, \mathbb{R}) : \omega \mapsto B(\cdot, \omega)$$

où $\Omega^* \subseteq \Omega$ est de probabilité 1 (donc $(\Omega^*, \mathcal{F} \cap \Omega^*, \mathbb{P})$ est un espace probabilisé).

Définition 5.1.1 La *mesure de Wiener* W est la loi du mouvement Brownien $B : \Omega^* \rightarrow C([0, +\infty[, \mathbb{R})$, c'est-à-dire

$$W(A) = \mathbb{P}[B \in A]$$

pour tout $A \in \mathcal{B}(C([0, +\infty[, \mathbb{R}))$.

On peut facilement décrire les valeurs prises par cette mesure sur les cylindres. Cela repose directement sur la loi des loi fini-dimensionnelles de B .

Proposition 5.1.2 Pour tous $0 < t_1 < \dots < t_n$, la loi du vecteur aléatoire $(B(t_1), \dots, B(t_n))$ a pour densité

$$(x_1, \dots, x_n) \mapsto \frac{1}{(2\pi)^{n/2} \sqrt{t_1(t_2 - t_1) \dots (t_n - t_{n-1})}} \exp \left(- \sum_{j=1}^n \frac{(x_j - x_{j-1})^2}{2(t_j - t_{j-1})} \right)$$

avec la convention que $x_0 = 0$ et $t_0 = 0$.

Démonstration : On sait que les variables aléatoires

$$Y_1 = B(t_1) - B(t_0), Y_2 = B(t_2) - B(t_1), \dots, Y_n = B(t_n) - B(t_{n-1})$$

sont indépendantes et pour tout $j \in \{1, \dots, n\}$, on a $Y_j \sim \mathcal{N}(0, t_j - t_{j-1})$. Il s'ensuit que la loi du vecteur aléatoire $Y = (Y_1, \dots, Y_n)$ a pour densité la fonction

$$f_Y(y_1, \dots, y_n) = \frac{1}{(2\pi)^{n/2} \sqrt{t_1(t_2 - t_1) \dots (t_n - t_{n-1})}} \exp \left(- \sum_{j=1}^n \frac{y_j^2}{2(t_j - t_{j-1})} \right).$$

Considérons la bijection linéaire

$$\varphi : \mathbb{R}^n \rightarrow \mathbb{R}^n : (y_1, \dots, y_n) \mapsto (y_1, y_1 + y_2, y_1 + y_2 + y_3, \dots, y_1 + y_2 + y_3 + \dots + y_n)$$

d'inverse

$$\varphi^{-1} : \mathbb{R}^n \rightarrow \mathbb{R}^n : (x_1, \dots, x_n) \mapsto (x_1, x_2 - x_1, x_3 - x_2, \dots, x_n - x_{n-1}).$$

Puisque $\varphi(Y_1, \dots, Y_n) = (B(t_1), \dots, B(t_n))$ et puisque le jacobien de φ^{-1} vaut 1, le théorème de changement de variable donne

$$\begin{aligned} f_{(B(t_1), \dots, B(t_n))}(x_1, \dots, x_n) &= 1 \cdot f_Y(\varphi^{-1}(x_1, \dots, x_n)) \\ &= f_Y(x_1, x_2 - x_1, x_3 - x_2, \dots, x_n - x_{n-1}) \\ &= \frac{1}{(2\pi)^{n/2} \sqrt{t_1(t_2 - t_1) \dots (t_n - t_{n-1})}} \exp\left(-\sum_{j=1}^n \frac{(x_j - x_{j-1})^2}{2(t_j - t_{j-1})}\right). \end{aligned}$$

■

Par conséquent, pour tout cylindre $C = \prod_{t \in I} A_t$ où $A_t \neq \mathbb{R}$ si $t \in \{t_1, \dots, t_n\}$ avec $0 < t_1 < \dots < t_n$, on a

$$\begin{aligned} W(C) &= \mathbb{P}[B \in C] \\ &= \frac{1}{(2\pi)^{n/2} \sqrt{t_1(t_2 - t_1) \dots (t_n - t_{n-1})}} \int_{A_{t_n}} \dots \int_{A_{t_1}} \exp\left(-\sum_{j=1}^n \frac{(x_j - x_{j-1})^2}{2(t_j - t_{j-1})}\right) dx_1 \dots dx_n \end{aligned}$$

avec $x_0 = 0$. De même, si $C = \prod_{t \in I} A_t$ où $A_t \neq \mathbb{R}$ si $t \in \{t_0, t_1, \dots, t_n\}$ avec $0 = t_0 < t_1 < \dots < t_n$, alors

$$W(C) = \mathbf{1}_{A_0}(0) \frac{1}{(2\pi)^{n/2} \sqrt{t_1(t_2 - t_1) \dots (t_n - t_{n-1})}} \int_{A_{t_n}} \dots \int_{A_{t_1}} \exp\left(-\sum_{j=1}^n \frac{(x_j - x_{j-1})^2}{2(t_j - t_{j-1})}\right) dx_1 \dots dx_n$$

Ces deux relations caractérisent entièrement la loi du mouvement Brownien par la Proposition 1.2.9. On en tire que la mesure de Wiener ne dépend pas du choix du mouvement Brownien choisi pour la définir. Ainsi, si \tilde{B} est un mouvement Brownien défini sur $(\tilde{\Omega}, \tilde{\mathcal{F}}, \tilde{\mathbb{P}})$, alors pour tout $A \in \mathcal{B}(C([0, +\infty[, \mathbb{R})))$, on a

$$\mathbb{P}[B \in A] = W(A) = \tilde{\mathbb{P}}[\tilde{B} \in A].$$

Remarque 5.1.3 Au vu de la Remarque 1.2.6, on pourrait définir le mouvement Brownien à partir de la mesure de Wiener. En effet, on prend l'espace probabilisé

$$(C([0, +\infty[, \mathbb{R})), \mathcal{B}(C([0, +\infty[, \mathbb{R})), W)$$

et on pose

$$B(\omega, t) = \omega(t).$$

Alors B est un mouvement Brownien sur $(C([0, +\infty[, \mathbb{R})), \mathcal{B}(C([0, +\infty[, \mathbb{R})), W)$. On appelle cette procédure la *construction canonique* du mouvement Brownien.

5.2 Propriétés immédiates

Dans cette section, nous fixons un mouvement Brownien B .

Proposition 5.2.1 *Le mouvement Brownien a les propriétés suivantes :*

1. *Symétrie : Le processus $\{-B(t) : t \geq 0\}$ est un mouvement Brownien.*
2. *Auto-similarité : Pour tout $c > 0$, le processus $\{B(ct)/\sqrt{c} : t \geq 0\}$ est un mouvement Brownien.*
3. *Inversion du temps : Le processus $\{X(t) : t \geq 0\}$ défini par*

$$X(t) = \begin{cases} tB\left(\frac{1}{t}\right) & \text{si } t > 0 \\ 0 & \text{si } t = 0 \end{cases}$$

est un mouvement Brownien.

4. *Propriété de Markov faible : Pour tout $t_0 > 0$, le processus $\{B(t_0 + t) - B(t_0) : t \geq 0\}$ est un mouvement Brownien indépendant de $\sigma(\{B(s) : s \leq t_0\})$.*

Démonstration : Dans tous les cas, il est clair que le processus considéré est un processus gaussien centré.

1. Les trajectoires de $\{-B(t) : t \geq 0\}$ sont presque sûrement continues puisque celles de B le sont. De plus, on a

$$\text{Cov}(-B(t), -B(s)) = \text{Cov}(B(t), B(s)) = \min(t, s),$$

ce qui suffit.

2. A nouveau, les trajectoires du processus $\{B(ct)/\sqrt{c} : t \geq 0\}$ sont presque sûrement continues. C'est un mouvement brownien car on a de plus

$$\text{Cov}\left(\frac{B(ct)}{\sqrt{c}}, \frac{B(cs)}{\sqrt{c}}\right) = \frac{1}{c} \text{Cov}(B(ct), B(cs)) = \frac{1}{c} \min(ct, cs) = \min(t, s).$$

3. Remarquons que

$$\begin{aligned} \text{Cov}(X(t), X(s)) &= \mathbb{E}\left[tB\left(\frac{1}{t}\right), sB\left(\frac{1}{s}\right)\right] = ts\mathbb{E}\left[B\left(\frac{1}{t}\right), B\left(\frac{1}{s}\right)\right] = \begin{cases} \frac{ts}{t} & \text{si } t > s \\ \frac{ts}{s} & \text{si } t \leq s \end{cases} \\ &= \min(t, s). \end{aligned}$$

En particulier, les lois fini-dimensionnelles de X sur $]0, +\infty[$ coïncident avec celles du mouvement Brownien. Sur $]0, +\infty[$, on a également la continuité presque sûre des trajectoires de X . Il reste à étudier la continuité des trajectoires en 0. On cherche donc à montrer que

$$\mathbb{P}\left[\lim_{t \rightarrow 0} X(t) = 0\right] = 1.$$

Par continuité presque sûre des trajectoires de X sur $]0, +\infty[$, on a

$$\begin{aligned} \mathbb{P} \left[\lim_{t \rightarrow 0} X(t) = 0 \right] &= \mathbb{P} \left[\bigcap_{n \geq 1} \bigcup_{p \geq 1} \bigcap_{t \in]0, 1/p[\cap \mathbb{Q}} \left\{ |X(t)| \leq \frac{1}{n} \right\} \right] \\ &= \lim_{n \rightarrow +\infty} \lim_{p \rightarrow +\infty} \mathbb{P} \left[\bigcap_{t \in]0, 1/p[\cap \mathbb{Q}} \left\{ |X(t)| \leq \frac{1}{n} \right\} \right] \end{aligned}$$

Notons $]0, 1/p[\cap \mathbb{Q} = \{q_m : m \in \mathbb{N}\}$. On a alors

$$\begin{aligned} \mathbb{P} \left[\bigcap_{t \in]0, 1/p[\cap \mathbb{Q}} \left\{ |X(t)| \leq \frac{1}{n} \right\} \right] &= \lim_{M \rightarrow +\infty} \mathbb{P} \left[\bigcap_{m \leq M} \left\{ |X(q_m)| \leq \frac{1}{n} \right\} \right] \\ &= \lim_{M \rightarrow +\infty} \mathbb{P} \left[\bigcap_{m \leq M} \left\{ |B(q_m)| \leq \frac{1}{n} \right\} \right] \end{aligned}$$

puisque par ce qui précède, on sait que le vecteur aléatoire $(X(q_1), \dots, X(q_M))$ a la même loi que le vecteur aléatoire $(B(q_1), \dots, B(q_M))$. Ainsi, il vient

$$\begin{aligned} \mathbb{P} \left[\lim_{t \rightarrow 0} X(t) = 0 \right] &= \lim_{n \rightarrow +\infty} \lim_{p \rightarrow +\infty} \lim_{M \rightarrow +\infty} \mathbb{P} \left[\bigcap_{m \leq M} \left\{ |B(q_m)| \leq \frac{1}{n} \right\} \right] \\ &= \mathbb{P} \left[\lim_{t \rightarrow 0} B(t) = 0 \right] \\ &= 1, \end{aligned}$$

ce qui permet de conclure.

4. Comme les trajectoires du processus $\{B(t_0 + t) - B(t_0) : t \geq 0\}$ sont presque sûrement continues, il suffit d'étudier sa fonction de covariance. On a

$$\begin{aligned} &\text{Cov}(B(t_0 + t) - B(t_0), B(t_0 + s) - B(t_0)) \\ &= \text{Cov}(B(t_0 + t), B(t_0 + s)) - \text{Cov}(B(t_0 + t), B(t_0)) - \text{Cov}(B(t_0), B(t_0 + s)) + \text{Cov}(B(t_0), B(t_0)) \\ &= \min(t_0 + t, t_0 + s) - t_0 - t_0 + t_0 \\ &= \min(t, s) \end{aligned}$$

et le processus est donc un mouvement brownien. Pour obtenir l'indépendance, remarquons que pour tous $n, m \geq 1$, $t_1 < \dots < t_n$ et $s_1 < \dots < s_m \leq t_0$, les vecteurs

$$(B(t_0 + t_1) - B(t_0), \dots, B(t_0 + t_n) - B(t_0 + t_{n-1})) \quad \text{et} \quad (B(s_1) - B(0), \dots, B(s_m) - B(s_{m-1}))$$

sont indépendants. On en tire directement l'indépendance des vecteurs

$$(B(t_0 + t_1) - B(t_0), \dots, B(t_0 + t_n) - B(t_0)) \quad \text{et} \quad (B(s_1), \dots, B(s_m))$$

ce qui suffit. ■

La propriété d'auto-similarité du mouvement Brownien reflète la nature fractale de celui-ci : en effet, il suffit de connaître le comportement du mouvement Brownien sur un intervalle de temps arbitrairement petit pour le connaître sur tout intervalle de temps arbitrairement grand. La propriété d'inversion du temps est également très forte : les premiers instants du mouvement Brownien sont suffisamment riches pour capturer le comportement du mouvement Brownien à l'infini. Enfin, rappelons que la propriété de Markov faible peut se réécrire comme l'indépendance du futur par rapport au passé, conditionnellement au présent.

5.3 Propriété de Markov forte et temps d'atteinte

Nous venons de voir que pour tout $t_0 \geq 0$, le processus $\{B(t + t_0) - B(t_0) : t \geq 0\}$ est un mouvement Brownien indépendant de $\sigma(\{B(s) : s \leq t_0\})$. La propriété forte de Markov établit un résultat similaire lorsque le temps auquel on commence à regarder notre processus est aléatoire. On impose à ce temps d'être un temps d'arrêt pour la filtration naturelle du mouvement Brownien. Nous en déduisons ensuite des informations sur le temps d'atteinte d'un niveau donné. Commençons par étendre les définitions vues dans le cas discret au cas continu.

Définition 5.3.1 Soit \mathcal{F} une σ -algèbre sur Ω . Une famille de sous- σ -algèbres $(\mathcal{F}_t)_{t \geq 0}$ est une *filtration* si $\mathcal{F}_s \subseteq \mathcal{F}_t$ pour tous $0 \leq s < t$. Un processus stochastique $\{X(t) : t \geq 0\}$ est *adapté à la filtration* $(\mathcal{F}_s)_{s \geq 0}$ si $X(t)$ est \mathcal{F}_t -mesurable pour tout $t \geq 0$.

Comme dans le cas discret, on peut associer à un processus stochastique une filtration naturelle pour laquelle le processus est adapté : il suffit de poser

$$\mathcal{F}_t^X = \sigma(\{X(s) : s \leq t\}).$$

Définition 5.3.2 Une variable aléatoire T est un *temps d'arrêt* pour la filtration $(\mathcal{F}_t)_{t \geq 0}$ si

1. T est à valeurs dans $[0, +\infty]$,
2. $\{T \leq t\} \in \mathcal{F}_t$ pour tout $t \geq 0$.

Remarque 5.3.3 Si T est un temps d'arrêt, on a également

$$\{T < t\} = \bigcup_{n \geq 1} \{T \leq t - \frac{1}{n}\} \in \mathcal{F}_t$$

et donc également

$$\{T = t\} \in \mathcal{F}_t.$$

Néanmoins, contrairement au cas discret, la réciproque de ce résultat est fautive en général. Ce problème est lié à la notion de continuité à droite de filtrations, que nous n'aborderons pas dans ce cours.

Le lemme suivant montre que tout temps d'arrêt peut être approché par une suite décroissante $(T_n)_{n \in \mathbb{N}}$ de temps d'arrêts tels que pour tout $n \in \mathbb{N}$, T_n est à valeurs dans un ensemble dénombrable.

Lemme 5.3.4 Soit T un temps d'arrêt pour la filtration $(\mathcal{F}_t)_{t \geq 0}$. Pour tout $n \in \mathbb{N}$, on pose

$$T_n = \begin{cases} \frac{k+1}{2^n} & \text{si } \frac{k}{2^n} < T \leq \frac{k+1}{2^n} \\ +\infty & \text{si } T = +\infty. \end{cases}$$

Alors $(T_n)_{n \in \mathbb{N}}$ est une suite de temps d'arrêts pour la filtration $(\mathcal{F}_t)_{t \geq 0}$ qui décroît vers T .

Démonstration : Pour tout $n \in \mathbb{N}$ et tout $t \geq 0$, on a

$$\{T_n \leq t\} = \bigcup_{k=0}^{[2^n t]-1} \left\{ T_n = \frac{k+1}{2^n} \right\} = \bigcup_{k=0}^{[2^n t]-1} \left\{ \frac{k}{2^n} < T \leq \frac{k+1}{2^n} \right\} = \left\{ T \leq \frac{[2^n t]}{2^n} \right\} \in \mathcal{F}_{\frac{[2^n t]}{2^n}} \subseteq \mathcal{F}_t.$$

De plus, il est évident que T_n décroît vers T . ■

Afin de pouvoir généraliser la propriété de Markov à un temps d'arrêt, il faut pouvoir donner un sens à la σ -algèbre engendrée par les variables aléatoires $B(s)$, $s \leq T$. Pour cela, on introduit la notion suivante.

Définition 5.3.5 La σ -algèbre \mathcal{F}_T des événements antérieurs à un temps d'arrêt T est définie par

$$\mathcal{F}_T = \{A \in \mathcal{F} : \forall t \geq 0, A \cap \{T \leq t\} \in \mathcal{F}_t\}.$$

Il est facile de vérifier qu'il s'agit d'une σ -algèbre. De plus, on a le résultat suivant.

Lemme 5.3.6 Les variables aléatoires T et $B(T)$ sont \mathcal{F}_T -mesurables.

Démonstration : Pour montrer que T est \mathcal{F}_T -mesurable, il faut montrer que pour tout $b \geq 0$, on a $\{T \leq b\} \in \mathcal{F}_T$. C'est immédiat car pour tout $t \geq 0$,

$$\{T \leq b\} \cap \{T \leq t\} = \{T \leq \min(t, b)\} \in \mathcal{F}_{\min(t, b)} \subseteq \mathcal{F}_t.$$

Pour $B(T)$, on commence par remarquer que

$$B(T) = \lim_{n \rightarrow +\infty} \sum_{k=0}^{+\infty} \mathbf{1}_{\{\frac{k}{2^n} < T \leq \frac{k+1}{2^n}\}} B\left(\frac{k}{2^n}\right)$$

et il suffit de montrer que chaque variable aléatoire $\mathbf{1}_{\{\frac{k}{2^n} < T \leq \frac{k+1}{2^n}\}} B\left(\frac{k}{2^n}\right)$ est \mathcal{F}_T -mesurable. Soit $b \in \mathbb{R}$ et $t \geq 0$. On veut montrer que

$$\left\{ \mathbf{1}_{\{\frac{k}{2^n} < T \leq \frac{k+1}{2^n}\}} B\left(\frac{k}{2^n}\right) \leq b \right\} \cap \{T \leq t\} \in \mathcal{F}_t.$$

Supposons tout d'abord que $t \leq \frac{k}{2^n}$. Alors $\mathbf{1}_{\{\frac{k}{2^n} < T \leq \frac{k+1}{2^n}\}} = 0$ et donc

$$\underbrace{\left\{ \mathbf{1}_{\{\frac{k}{2^n} < T \leq \frac{k+1}{2^n}\}} B\left(\frac{k}{2^n}\right) \leq b \right\}}_{\in \{\emptyset, \Omega\}} \cap \{T \leq t\} \in \mathcal{F}_t.$$

Supposons à présent que $\frac{k}{2^n} < t \leq \frac{k+1}{2^n}$. Alors

$$\begin{aligned} & \left\{ \mathbf{1}_{\left\{ \frac{k}{2^n} < T \leq \frac{k+1}{2^n} \right\}} B\left(\frac{k}{2^n}\right) \leq b \right\} \cap \{T \leq t\} \\ = & \left(\{0 \leq b\} \cap \left\{ T \leq \frac{k}{2^n} \right\} \cap \{T \leq t\} \right) \cup \left(\left\{ B\left(\frac{k}{2^n}\right) \leq b \right\} \cap \left\{ \frac{k}{2^n} < T \right\} \cap \{T \leq t\} \right) \end{aligned}$$

et on vérifie immédiatement que cet ensemble appartient à \mathcal{F}_t . Enfin, si $t > \frac{k+1}{2^n}$, la variable aléatoire $\mathbf{1}_{\left\{ \frac{k}{2^n} < T \leq \frac{k+1}{2^n} \right\}} B\left(\frac{k}{2^n}\right)$ est \mathcal{F}_t -mesurable, ce qui permet de conclure. ■

Théorème 5.3.7 (Propriété de Markov forte) *Soit T un temps d'arrêt pour la filtration naturelle du mouvement Brownien. Si $\mathbb{P}[T < +\infty] = 1$, alors $\{B(t+T) - B(T) : t \geq 0\}$ est un mouvement Brownien indépendant de \mathcal{F}_T .*

Démonstration : Commençons par démontrer le résultat pour la suite décroissante $(T_n)_{n \in \mathbb{N}}$ de temps d'arrêt donnée dans le Lemme 5.3.4. Par hypothèse, ces temps d'arrêt sont presque sûrement finis. Pour tous $k, n \geq 0$, on considère les processus stochastiques définis par

$$B_{k,n}(t) = B(t + k2^{-n}) - B(k2^{-n})$$

et

$$B_n(t) = B(t + T_n) - B(T_n) \quad \text{si } T_n < +\infty$$

pour tout $t \geq 0$. Pour tout $A \in \mathcal{F}_{T_n}$ et tout borélien C , on a

$$\begin{aligned} \mathbb{P}[\{B_n \in C\} \cap A] &= \sum_{k=1}^{+\infty} \mathbb{P} \left[\{B_{k,n} \in C\} \cap \underbrace{A \cap \{T_n = k2^{-n}\}}_{\in \mathcal{F}_{k2^{-n}}} \right] \\ &= \sum_{k=1}^{+\infty} \mathbb{P}[\{B_{k,n} \in C\}] \mathbb{P}[A \cap \{T_n = k2^{-n}\}] \\ &= \mathbb{P}[\{B \in C\}] \sum_{k=1}^{+\infty} \mathbb{P}[A \cap \{T_n = k2^{-n}\}] \\ &= \mathbb{P}[\{B \in C\}] \mathbb{P}[A] \end{aligned}$$

puisque par la propriété de Markov faible, on sait que $B_{k,n}$ est indépendant de $\mathcal{F}_{k2^{-n}}$ et que $\mathbb{P}[\{B_{k,n} \in C\}] = \mathbb{P}[\{B \in C\}]$. En prenant $A = \Omega$, on trouve

$$\mathbb{P}[B_n \in C] = \mathbb{P}[B \in C]$$

et donc B et B_n sont égaux en loi. Comme on a également la continuité presque sûre des trajectoires de B_n , c'est un mouvement brownien. Enfin, le calcul précédent montre qu'il est indépendant de \mathcal{F}_{T_n} .

Il reste à prouver que le résultat est également valide pour le temps d'arrêt T . Comme $(T_n)_{n \in \mathbb{N}}$ décroît vers T , on a clairement $\mathcal{F}_T \subseteq \mathcal{F}_{T_n}$. Par conséquent, $\{B(t + T_n) - B(T_n) : t \geq 0\}$ est un mouvement brownien indépendant de \mathcal{F}_{T_n} . Puisqu'on a presque sûrement

$$B(t + s + T) - B(t + T) = \lim_{n \rightarrow +\infty} B(t + s + T_n) - B(t + T_n)$$

par continuité des trajectoires de B , on trouve que $B(t + s + T) - B(t + T) \sim \mathcal{N}(0, s)$. De plus, la relation précédente implique que les accroissements de $\{B(t + T) - B(T) : t \geq 0\}$ sont indépendants. Enfin, il est clair que les trajectoires de $\{B(t + T) - B(T) : t \geq 0\}$ sont presque sûrement continues. On en tire qu'il s'agit bien d'un mouvement brownien. Il est indépendant de \mathcal{F}_T puisque $B(t + s + T_n) - B(t + T_n)$ est indépendant de $\mathcal{F}_{T_n} \supseteq \mathcal{F}_T$ pour tout n . ■

Dans ce qui suit, on va s'intéresser au temps d'arrêt T_a , définis pour tout $a \in \mathbb{R}_0$ par

$$T_a = \inf\{t \geq 0 : B(t) = a\}.$$

Remarquons que si $T_a < +\infty$, alors presque sûrement on a $B(T_a) = a$ par continuité des trajectoires du mouvement Brownien. Le principe de réflexion montre qu'après l'instant T_a , la droite $y = a$ est un axe de symétrie du mouvement brownien (comme l'est la droite $y = 0$ après l'instant $T_0 = 0$).

Proposition 5.3.8 (Principe de réflexion) Pour tout $a \in \mathbb{R}_0$, le processus Y défini par

$$Y(t) = \begin{cases} B(t) & \text{si } t \leq T_a \\ 2a - B(t) & \text{si } t > T_a \end{cases}$$

est un mouvement Brownien. De plus,

$$\mathbb{P}[T_a \leq t, B(t) \leq b] = \mathbb{P}[B(t) \geq 2a - b]$$

pour tout $b \leq a$.

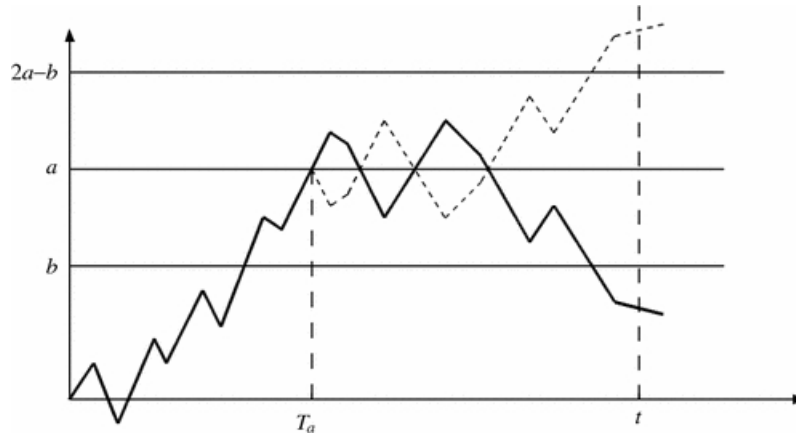


FIGURE 5.1 – Principe de réflexion.

Démonstration : Si T_a est infini¹, alors $Y(t) = B(t)$. Si T_a est fini, on sait par le Théorème 5.3.7 que le processus $B(t + T_a) - a$ est un mouvement brownien indépendant de \mathcal{F}_{T_a} . Par symétrie, on sait aussi que $a - B(t + T_a)$ est un mouvement brownien indépendant de \mathcal{F}_{T_a} . Par conséquent, le processus $\{B(t) : t > T_a\} = \{a + (B(t) - a) : t > T_a\}$ a la même loi que le processus $\{a + a - B(t) : t > T_a\} = \{2a - B(t) : t > T_a\}$, ce qui suffit pour la première partie.

1. on va voir dans le résultat suivant que presque sûrement, T_a est fini.

Pour la deuxième, on remarque que

$$\mathbb{P}[T_a \leq t, B(t) \leq b] = \mathbb{P}[B(t) \leq b | T_a \leq t] \mathbb{P}[T_a \leq t] = \mathbb{P}[2a - B(t) \leq b | T_a \leq t] \mathbb{P}[T_a \leq t]$$

puisque par la première partie, $2a - B(t)$ et $B(t)$ ont la même loi pour $t > T_a$. Ainsi,

$$\begin{aligned} \mathbb{P}[T_a \leq t, B(t) \leq b] &= \mathbb{P}[2a - B(t) \leq b | T_a \leq t] \mathbb{P}[T_a \leq t] \\ &= \mathbb{P}[2a - B(t) \leq b, T_a \leq t] \\ &= \mathbb{P}[B(t) \geq 2a - b] \end{aligned}$$

puisque $\{B(t) \geq 2a - b\} \subseteq \{B(t) \geq a\} \subseteq \{T_a \leq t\}$. ■

On tire du principe de réflexion le premier résultat suivant sur les trajectoires du processus.

Corollaire 5.3.9 *Pour tout $t > 0$ et tout $a > 0$, on a*

$$\mathbb{P} \left[\sup_{0 \leq s \leq t} B(s) \geq a \right] = 2 \mathbb{P}[B(t) \geq a].$$

En particulier, $\mathbb{P}[T_a < +\infty] = 1$.

Démonstration : Remarquons que $\mathbb{P}[\sup_{0 \leq s \leq t} B(s) \geq a] = \mathbb{P}[T_a \leq t]$ puisque les trajectoires du mouvement Brownien sont presque sûrement continues et puisque l'intervalle $[0, t]$ est compact. De plus, $B(t) \geq a$ implique que presque sûrement, $T_a \leq t$. Par conséquent, en utilisant le principe de réflexion, on a

$$\begin{aligned} \mathbb{P} \left[\sup_{0 \leq s \leq t} B(s) \geq a \right] &= \mathbb{P}[T_a \leq t] \\ &= \mathbb{P}[T_a \leq t, B(t) \leq a] + \mathbb{P}[T_a \leq t, B(t) \geq a] \\ &= \mathbb{P}[B(t) \geq 2a - a] + \mathbb{P}[B(t) \geq a] \\ &= 2 \mathbb{P}[B(t) \geq a] \end{aligned}$$

puisque $\{B(t) \geq a\} \subseteq \{T_a \leq t\}$.

Pour la deuxième partie, fixons $b > a$. En considérant la limite de la suite croissante d'événements $(\{\sup_{0 \leq s \leq n} B(s) \geq b\})_{n \in \mathbb{N}}$, on obtient que

$$\mathbb{P} \left[\sup_{s \geq 0} B(s) \geq b \right] = \lim_{n \rightarrow +\infty} 2 \mathbb{P}[B(n) \geq b] = \lim_{n \rightarrow +\infty} 2 \mathbb{P} \left[Z \geq \frac{b}{\sqrt{n}} \right] = 1$$

où $Z \sim \mathcal{N}(0, 1)$. En utilisant la continuité presque sûre des trajectoires, on trouve

$$\mathbb{P}[T_a < +\infty] \geq \mathbb{P} \left[\sup_{s \geq 0} B(s) \geq b \right] = 1,$$

d'où la conclusion. ■

Puisque le mouvement Brownien est symétrique, on obtient également que $\mathbb{P}[T_a < +\infty] = 1$ pour tout $a < 0$. Ainsi, le mouvement Brownien atteint presque sûrement tout niveau a . Nous verrons néanmoins que le temps moyen nécessaire pour atteindre ce niveau est infini !

5.4 Théorème de Donsker

Le Théorème de Donsker est l'équivalent du théorème central limite dans le cadre des marches aléatoires. Il donne l'une des justifications de la place du mouvement Brownien parmi les processus stochastiques les plus importants. Considérons une suite de variables aléatoires $(X_i)_{i \in \mathbb{N}_0}$ indépendantes et identiquement distribuées, telles que $\mathbb{E}[X_i] = 0$ et $\mathbb{E}[X_i^2] = 1$. Considérons ensuite la marche aléatoire

$$S_n = \sum_{i=1}^n X_i.$$

On peut l'étendre en un processus stochastique à temps continu en effectuant une interpolation linéaire : on définit le processus $\{X(t) : t \geq 0\}$ en posant

$$X(t) = \sum_{i=1}^{\lfloor t \rfloor} X_i + (t - \lfloor t \rfloor)X_{\lfloor t \rfloor + 1} = S_k + (t - k)X_{k+1} \quad \text{si } t \in [k, k+1[.$$

On obtient ainsi un processus dont les trajectoires sont des fonctions affines par morceaux. A présent, nous allons effectuer un changement d'échelle pour affiner notre approximation : on interpole les valeurs de la marche aléatoire sur les intervalles de longueur $1/n$: on considère le processus $\{Y_n(t) : t \geq 0\}$ défini par

$$Y_n(t) = \sum_{i=1}^{\lfloor nt \rfloor} X_i + (nt - \lfloor nt \rfloor)X_{\lfloor nt \rfloor + 1} = X(nt).$$

Remarquons que

$$Y_n\left(\frac{k}{n}\right) = \sum_{i=1}^k X_i$$

et par indépendance des variables X_i , $i \in \mathbb{N}_0$, on a

$$\mathbb{E}\left[\left(Y_n\left(\frac{k}{n}\right)\right)^2\right] = \sum_{i=1}^k \mathbb{E}[X_i^2] = k.$$

Notre objectif est de construire une suite de processus qui converge vers le mouvement Brownien. Or, pour un mouvement Brownien, on sait que $\mathbb{E}[B(t)^2] = t$ pour tout $t \geq 0$. Il semble donc naturel de considérer $\frac{1}{\sqrt{n}}Y_n(k/n)$ à la place de $Y_n(k/n)$. On est donc amenés à considérer la suite de processus $\{Z_n(t) : t \geq 0\}$ définis par

$$Z_n(t) = \frac{1}{\sqrt{n}} \sum_{i=1}^{\lfloor nt \rfloor} X_i + (nt - \lfloor nt \rfloor) \frac{1}{\sqrt{n}} X_{\lfloor nt \rfloor + 1}.$$

Proposition 5.4.1 *Pour tout $t \geq 0$ fixé, la suite $(Z_n(t))_{n \in \mathbb{N}_0}$ converge en loi vers $B(t)$.*

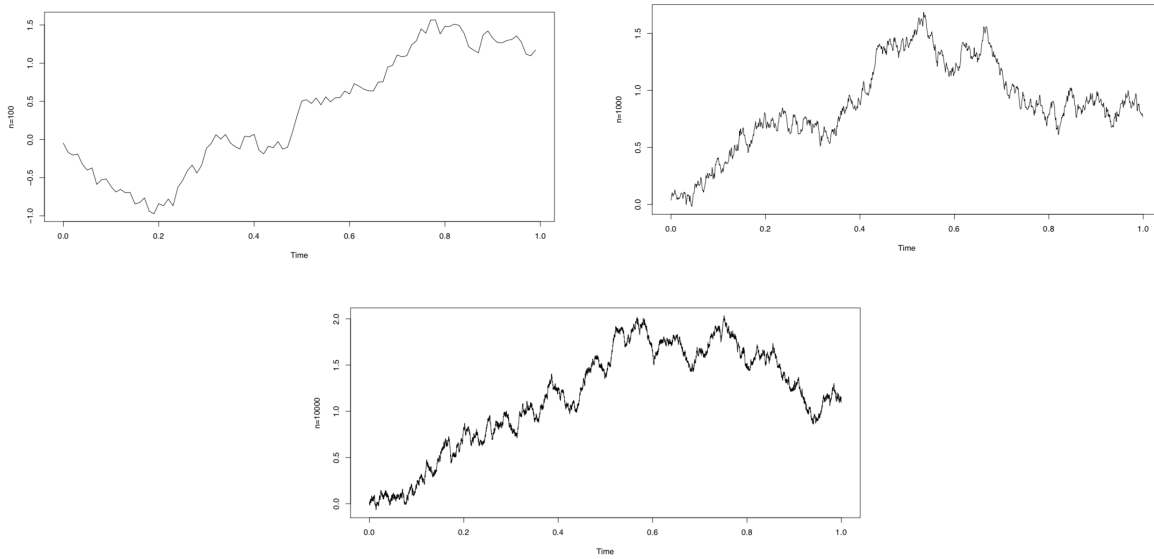


FIGURE 5.2 – Trajectoires de Z_n sur $[0, 1]$ pour $n = 100$, $n = 1000$ et $n = 10000$

Démonstration : Commençons par rappeler que $B(t) \sim \mathcal{N}(0, t)$. Remarquons que par application du Théorème central limite, on sait que la suite

$$\frac{1}{\sqrt{n}} \sum_{i=1}^{\lfloor nt \rfloor} X_i = \frac{S_{\lfloor nt \rfloor}}{\sqrt{n}}$$

converge en loi vers une loi normale $\mathcal{N}(0, t)$ si n tend vers l’infini. Nous allons à présent montrer que la suite $(nt - \lfloor nt \rfloor) \frac{1}{\sqrt{n}} X_{\lfloor nt \rfloor + 1}$ converge en probabilité vers 0 lorsque n tend vers l’infini. La conclusion s’ensuivra facilement puisque la convergence en probabilité implique la convergence en loi. Pour cela, fixons $\varepsilon > 0$ et appliquons l’inégalité de Bienaymé Tchebychev : il vient

$$\mathbb{P} \left[\left| (nt - \lfloor nt \rfloor) \frac{1}{\sqrt{n}} X_{\lfloor nt \rfloor + 1} \right| > \varepsilon \right] \leq \frac{(nt - \lfloor nt \rfloor)^2}{n\varepsilon^2} \leq \frac{1}{n\varepsilon^2} \rightarrow 0$$

lorsque $n \rightarrow +\infty$. ■

On peut généraliser ce résultat en étudiant la convergence des lois fini-dimensionnelles. Commençons par regarder les accroissements.

Proposition 5.4.2 *Pour tous $0 \leq t_1 < t_2 < \dots < t_d$, on a*

$$\lim_{n \rightarrow +\infty} \mathbb{P} [Z_n(t_1) - Z_n(t_0) \leq x_1, \dots, Z_n(t_d) - Z_n(t_{d-1}) \leq x_d] = \prod_{l=1}^d \mathbb{P} [B(t_l) - B(t_{l-1}) \leq x_l]$$

où B est un mouvement Brownien et où $x_1, \dots, x_d \in \mathbb{R}$.

Autrement dit, les accroissements de $\{Z_n(t) : t \geq 0\}$ sont asymptotiquement indépendants et distribués selon une loi normale de moyenne nulle et de variance donnée par la longueur de l'intervalle sur lequel on considère l'accroissement.

Démonstration : Tout accroissement $Z_n(t) - Z_n(s)$ avec $t > s \geq 0$ s'écrit

$$Z_n(t) - Z_n(s) = \frac{1}{\sqrt{n}} \sum_{i=[ns]+1}^{[nt]} X_i + (nt - [nt]) \frac{1}{\sqrt{n}} X_{[nt]+1} - (ns - [ns]) \frac{1}{\sqrt{n}} X_{[ns]+1}.$$

Comme les termes $(nt - [nt]) \frac{1}{\sqrt{n}} X_{[nt]+1}$ et $(ns - [ns]) \frac{1}{\sqrt{n}} X_{[ns]+1}$ convergent en probabilité vers 0, il suffit d'étudier la convergence en loi de la suite de vecteurs aléatoires

$$\left(\frac{1}{\sqrt{n}} \sum_{i=[nt_0]+1}^{[nt_1]} X_i, \dots, \frac{1}{\sqrt{n}} \sum_{i=[nt_{d-1}]+1}^{[nt_d]} X_i \right).$$

Remarquons que si n est suffisamment grand, les variables aléatoires intervenant les sommes ci-dessus sont différentes et donc indépendantes. Par conséquent, pour n suffisamment grand, les sommes considérées sont donc indépendantes. Il reste alors à montrer que toute somme du type $\frac{1}{\sqrt{n}} \sum_{i=[ns]+1}^{[nt]} X_i$ avec $t > s \geq 0$ converge en loi vers une loi normale $\mathcal{N}(0, t - s)$. Remarquons que si $(\tilde{X}_i)_{i \in \mathbb{N}_0}$ sont des variables iid selon la même loi que X_1 , alors la loi de $\frac{1}{\sqrt{n}} \sum_{i=[ns]+1}^{[nt]} X_i$ est la même que celle de

$$\frac{1}{\sqrt{n}} \sum_{i=1}^{[nt]-[ns]} \tilde{X}_i.$$

Par le Théorème central limite et puisque $[nt] - [ns] \approx n(t - s)$ lorsque $n \rightarrow +\infty$, cette suite converge en loi vers un loi normale $\mathcal{N}(0, t - s)$, d'où la conclusion. ■

Corollaire 5.4.3 Pour tout $t_1, \dots, t_d \geq 0$ fixés, on a

$$\lim_{n \rightarrow +\infty} \mathbb{P}[Z_n(t_1) \leq x_1, \dots, Z_n(t_d) \leq x_d] = \mathbb{P}[B(t_1) \leq x_1, \dots, B(t_d) \leq x_d]$$

où B est un mouvement Brownien et où $x_1, \dots, x_d \in \mathbb{R}$.

Démonstration : Il suffit de montrer qu'asymptotiquement, le vecteur $(Z_n(t_1), \dots, Z_n(t_d))$ est gaussien centré de covariance $K_{ij} = \min(t_i, t_j)$. ■

La convergence des lois fini-dimensionnelles n'est pas suffisante pour assurer la convergence d'une suite de processus. Néanmoins, de telles considérations nécessitent de développer de nouvelles notions. Nous admettrons donc le résultat suivant.

Théorème 5.4.4 (Donsker) Pour toute fonction continue $H : C([0, 1]) \rightarrow \mathbb{R}$, on a

$$\lim_{n \rightarrow +\infty} \mathbb{P}[H(Z_n) \leq x] = \mathbb{P}[H(B) \leq x]$$

pour tout $x \in \mathbb{R}$.

Ce Théorème est très puissant et permet de calculer de nombreuses probabilités pour des suites de variables indépendantes et identiquement distribuées, quelle que soit la loi de celles-ci (admettant des moments d'ordre 1 et 2). Par exemple, on a

$$\lim_{n \rightarrow +\infty} \mathbb{P} \left[\max_{0 \leq t \leq T} Z_n(t) \leq x \right] = \mathbb{P} \left[\max_{0 \leq t \leq T} B(t) \leq x \right].$$

Chapitre 6

Martingales en temps continu

Dans ce chapitre, nous étendons la définition de martingales à temps discret au cas continu. Nous présentons quelques propriétés sur les martingales arrêtées et sur la convergence des martingales, en nous basant sur les résultats obtenus dans le cas discret. Enfin, nous montrons que le mouvement Brownien est une martingale pour sa filtration naturelle.

6.1 Définitions et propriétés

La notion de martingale se généralise aisément au cas de processus à temps continu.

Définition 6.1.1 Un processus stochastique $\{Z(t) : t \geq 0\}$ est une *martingale* pour la filtration $(\mathcal{F}_t)_{t \geq 0}$ si

1. $\{Z(t) : t \geq 0\}$ est adapté à $(\mathcal{F}_t)_{t \geq 0}$,
2. $\mathbb{E}[|Z(t)|] < +\infty$ pour tout $t \geq 0$,
3. $\mathbb{E}[Z(t) | \mathcal{F}_s] = Z(s)$ pour tout $t \geq s$.

Comme dans le cas discret, on peut également parler de sous-martingales et de sur-martingales en changeant l'égalité du point 3 par l'inégalité \geq ou \leq respectivement. On trouve également que la suite des espérances $\mathbb{E}[Z(t)]$ est constante, croissante ou décroissante selon que $\{Z(t) : t \geq 0\}$ est une martingale, une sous-martingale ou une sur-martingale respectivement.

Comme dans le cas discret, en utilisant l'inégalité de Jensen pour les espérances conditionnelles, on obtient que si $\{Z(t) : t \geq 0\}$ est une martingale et si f est fonction convexe Borel-mesurable pour laquelle $\mathbb{E}[|f(Z(t))|] < +\infty$ pour tout $t \geq 0$, alors $\{f(Z(t)) : t \geq 0\}$ est une sous-martingale.

Evidemment, si $\{Z(t) : t \geq 0\}$ est une martingale pour la filtration $(\mathcal{F}_t)_{t \geq 0}$ etsi $D \subseteq [0, +\infty[$ est un ensemble dénombrable, alors la suite $(Z(n))_{n \in D}$ est une martingale à temps discret pour la filtration $(\mathcal{F}_n)_{n \in D}$. C'est grâce à cette constatation que nous allons obtenir des résultats pour les martingales à temps continu à partir des résultats obtenus dans le Chapitre 2. Notons que pour pouvoir passer du discret au continu, il est naturel d'imposer à notre processus des conditions de régularité sur ses trajectoires.

Proposition 6.1.2 (Inégalités de Doob) Soit $\{Z(t) : t \geq 0\}$ une sous-martingale non-négative dont les trajectoires sont presque sûrement continues. Alors pour tout $c > 0$, on a

$$c \mathbb{P} \left[\sup_{0 \leq s \leq t} Z(s) \geq c \right] \leq \mathbb{E}[Z(t)]$$

et pour tout $p > 1$,

$$\left\| \sup_{0 \leq s \leq t} Z(s) \right\|_p \leq \frac{p}{p-1} \|Z(t)\|_p.$$

Démonstration : Pour tout $n \in \mathbb{N}_0$, considérons l'ensemble

$$\left\{ \frac{k}{2^n} t : k \in \{0, \dots, 2^n\} \right\}.$$

Puisque les trajectoires de Z sont presque sûrement continues, on a

$$\lim_{n \rightarrow +\infty} \sup_{0 \leq k \leq 2^n} Z\left(\frac{k}{2^n} t\right) = \sup_{0 \leq s \leq t} Z(s).$$

presque sûrement. Puisque,

$$\mathbb{P} \left[\sup_{0 \leq k \leq 2^n} Z\left(\frac{k}{2^n} t\right) \geq c \right] = \mathbb{E} \left[\mathbf{1}_{\{\sup_{0 \leq k \leq 2^n} Z(\frac{k}{2^n} t) \geq c\}} \right]$$

le Lemme de Fatou nous donne

$$\mathbb{E} \left[\mathbf{1}_{\{\sup_{0 \leq s \leq t} Z(s) \geq c\}} \right] \leq \liminf_{n \rightarrow +\infty} \mathbb{E} \left[\mathbf{1}_{\{\sup_{0 \leq k \leq 2^n} Z(\frac{k}{2^n} t) \geq c\}} \right] = \liminf_{n \rightarrow +\infty} \mathbb{P} \left[\sup_{0 \leq k \leq 2^n} Z\left(\frac{k}{2^n} t\right) \geq c \right].$$

En appliquant la Proposition 2.4.2 à la sous-martingale discrète $M_k = Z\left(\frac{k}{2^n} t\right)$ pour la filtration $\mathcal{F}_{\frac{k}{2^n} t}$, on a

$$c \mathbb{P} \left[\sup_{0 \leq k \leq 2^n} Z\left(\frac{k}{2^n} t\right) \geq c \right] \leq \mathbb{E}[Z(t)].$$

Cela permet de conclure la première partie. Pour la deuxième partie, on procède de même en appliquant le lemme de Fatou à l'inégalité de Doob discrète

$$\left\| \sup_{0 \leq k \leq 2^n} Z\left(\frac{k}{2^n} t\right) \right\|_p \leq \frac{p}{p-1} \|Z(t)\|_p.$$

■

Le théorème suivant généralise le résultat sur les martingales discrètes arrêtées.

Théorème 6.1.3 Soit $\{Z(t) : t \geq 0\}$ une martingale pour la filtration $(\mathcal{F}_t)_{t \geq 0}$ dont les trajectoires sont presque sûrement continues. Si T est un temps d'arrêt pour $(\mathcal{F}_t)_{t \geq 0}$, alors

$$Y(t) = Z(T \wedge t), \quad t \geq 0$$

est une martingale pour $(\mathcal{F}_t)_{t \geq 0}$. En particulier,

$$\mathbb{E}[Z(T \wedge t)] = \mathbb{E}[Z(0)].$$

Démonstration : Soient $t > s \geq 0$. Considérons la suite $(T_n)_{n \in \mathbb{N}}$ de temps d'arrêts construite dans le Lemme 5.3.4. Pour tout n , désignons par D_n l'ensemble

$$D_n = \left\{ \frac{k}{2^n} : k \geq 0 \right\}$$

l'ensembles des dyadiques d'échelle n . Alors $(Z(m))_{m \in D_n}$ est une martingale discrète pour la filtration $(\mathcal{F}_m)_{m \in D_n}$. De plus, T_n est un temps d'arrêt pour $(\mathcal{F}_m)_{m \in D_n}$ car pour tout $m \in D_n$, on a $\{T_n \leq m\} \in \mathcal{F}_m$ (comme montré dans le Lemme 5.3.4). Par conséquent, $(Z(T_n \wedge m))_{m \in D_n}$ est une martingale et

$$\mathbb{E}[Z(T_n \wedge m) | \mathcal{F}_l] = Z(T_n \wedge l)$$

pour tout $m, l \in D$, $m \geq l$. Prenons $l = [2^n s]/2^n$ et $m = [2^n t]/2^n$. Par continuité des trajectoires, on a presque sûrement

$$\lim_{n \rightarrow +\infty} Z(T_n \wedge m) = Z(T \wedge t) \quad \text{et} \quad \lim_{n \rightarrow +\infty} Z(T_n \wedge l) = Z(T \wedge s)$$

et il reste à montrer que cette convergence a lieu également dans L^1 (admis). ■

Théorème 6.1.4 *Soit T un temps d'arrêt borné pour la filtration $(\mathcal{F}_t)_{t \geq 0}$. Si $\{Z(t) : t \geq 0\}$ est une martingale pour $(\mathcal{F}_t)_{t \geq 0}$ dont les trajectoires sont presque sûrement continues, alors*

$$\mathbb{E}[Z(T)] = \mathbb{E}[Z(0)].$$

Démonstration : Par le Théorème 6.1.3, on sait que $\mathbb{E}[Z(T \wedge t)] = \mathbb{E}[Z(0)]$ pour tout $t \geq 0$. En prenant t suffisamment grand, on a $T \wedge t = T$ puisque T est borné, d'où la conclusion. ■

Théorème 6.1.5 *Soit $\{Z(t) : t \geq 0\}$ une martingale pour $(\mathcal{F}_t)_{t \geq 0}$ dont les trajectoires sont presque sûrement continues. Supposons qu'il existe $p > 1$ et $M < +\infty$ tels que $\mathbb{E}[|Z(t)|^p] \leq M$ pour tout $t \geq 0$, alors il existe une variable aléatoire Z pour laquelle $\mathbb{E}[|Z|^p] < +\infty$ et telle $Z(t)$ converge presque sûrement et en norme L^p vers Z .*

Démonstration : Comme $(Z(n))_{n \in \mathbb{N}}$ est une martingale discrète, le Corollaire 2.6.1 donne l'existence d'une variable aléatoire Z satisfaisant $\mathbb{E}[|Z|^p] < +\infty$ et telle que la suite $(Z(n))_{n \geq 0}$ converge presque sûrement et en norme L^p vers Z . Pour tout $n \in \mathbb{N}$ et tout $t \geq n$, on a

$$|Z(t) - Z| \leq |Z(n) - Z| + \sup_{s \geq n} |Z(s) - Z(n)|$$

et donc

$$\limsup_{t \rightarrow +\infty} |Z(t) - Z| \leq \lim_{n \rightarrow +\infty} \sup_{s \geq n} |Z(s) - Z(n)|.$$

Afin d'obtenir la convergence presque sûre de $Z(t)$ vers Z , il suffit donc de montrer que presque sûrement, $\lim_{n \rightarrow +\infty} \sup_{s \geq n} |Z(s) - Z(n)| = 0$. Pour cela, remarquons que $\{Z(s) - Z(n) : s \geq n\}$ est une martingale pour la filtration $(\mathcal{F}_s)_{s \geq n}$. Par conséquent, $\{|Z(s) - Z(n)|^p : s \geq n\}$ est une sous-martingale non-négative et la Proposition 6.1.2 donne

$$\mathbb{P}\left[\sup_{n \leq s \leq m} |Z(s) - Z(n)| \geq c \right] \leq c^{-p} \mathbb{E}[|Z(m) - Z(n)|^p]$$

pour tout $n \leq m$. Si on fait tendre m vers l'infini, la continuité de la mesure de probabilité \mathbb{P} et la convergence de la suite $(Z(n))_{n \geq 0}$ vers Z en norme L^p impliquent que

$$\mathbb{P}\left[\sup_{n \leq s} |Z(s) - Z(n)| \geq c\right] \leq c^{-p} \mathbb{E}[|Z - Z(n)|^p].$$

En prenant à présent la limite sur n , il vient de même

$$\mathbb{P}\left[\lim_{n \rightarrow +\infty} \sup_{n \leq s} |Z(s) - Z(n)| \geq c\right] = 0.$$

Comme $c > 0$ est arbitraire, on trouve qu'avec une probabilité 1, $\lim_{n \rightarrow +\infty} \sup_{n \leq s} |Z(s) - Z(n)| = 0$, ce qui suffit pour la convergence presque sûre.

Il reste à montrer qu'on a également la convergence en norme L^p . Pour tout $n \in \mathbb{N}$ et tout $t \geq n$, on a

$$\|Z(t) - Z\|_p \leq \|Z(t) - Z(n)\|_p + \|Z(n) - Z\|_p.$$

Puisque $\{|Z(t) - Z(n)|^p : t \geq n\}$ est une sous-martingale, on sait que $\mathbb{E}[|Z(t) - Z(n)|^p] \leq \mathbb{E}[|Z(m) - Z(n)|^p]$ pour tout $m > t$, d'où

$$\|Z(t) - Z\|_p \leq \|Z(m) - Z(n)\|_p + \|Z(n) - Z\|_p.$$

La conclusion s'ensuit puisque les deux termes du membre de droite tendent vers 0 lorsque n et m tendent vers l'infini. ■

Remarque 6.1.6 La convergence presque sûre reste vraie sous l'hypothèse $\mathbb{E}[|Z(t)|] \leq M$ pour tout $t \geq 0$ [8]. La convergence en norme L^1 est à nouveau plus délicate à étudier et repose sur la notion d'intégrabilité uniforme comme dans le cas discret.

6.2 Martingales et Mouvement Brownien

Théorème 6.2.1 *Le mouvement Brownien B est une martingale pour sa filtration naturelle.*

Démonstration : Pour tous $t > s \geq 0$, on a

$$\mathbb{E}[B(t)|\mathcal{F}_s] = \mathbb{E}[B(t) - B(s)|\mathcal{F}_s] + \mathbb{E}[B(s)|\mathcal{F}_s] = \mathbb{E}[B(t) - B(s)] + B(s) = B(s)$$

vu la propriété de Markov faible. ■

En utilisant l'inégalité de Jensen, on a donc que $B^2(t)$ et $e^{\gamma B(t)}$ (pour $\gamma > 0$) sont des sous-martingales. Le résultat suivant montre qu'on peut les modifier de manière déterministe pour obtenir des martingales.

Proposition 6.2.2 1. $\{B^2(t) - t : t \geq 0\}$ est une martingale.
 2. Pour tout $\gamma \in \mathbb{R}_0$, $\{e^{\gamma B(t) - \gamma^2 t/2} : t \geq 0\}$ est une martingale.

Démonstration : 1. Pour tous $t > s \geq 0$, on a

$$\begin{aligned}
 \mathbb{E}[B^2(t) - t | \mathcal{F}_s] &= \mathbb{E}[B^2(s) + (B(s) - B(t))^2 + 2B(s)(B(t) - B(s)) - t | \mathcal{F}_s] \\
 &= B^2(s) + \mathbb{E}[(B(s) - B(t))^2] + 2\mathbb{E}[B(s)(B(t) - B(s))] - t \\
 &= B^2(s) + (s - t) + 0 - t \\
 &= B^2(s) - s.
 \end{aligned}$$

2. Remarquons tout d'abord que pour tout $t \geq 0$, on a

$$\mathbb{E}[e^{\gamma B(t) - \gamma^2 t/2}] = e^{-\gamma^2 t/2} \mathbb{E}[e^{\gamma B(t)}] = 1$$

puisque $B(t) \sim \mathcal{N}(0, t)$. De plus, pour tous $t > s \geq 0$, on a

$$\begin{aligned}
 \mathbb{E}[e^{\gamma B(t) - \gamma^2 t/2} | \mathcal{F}_s] &= e^{-\gamma^2 t/2} \mathbb{E}[e^{\gamma B(t)} | \mathcal{F}_s] \\
 &= e^{-\gamma^2 t/2} \mathbb{E}[e^{\gamma(B(t) - B(s))} e^{\gamma B(s)} | \mathcal{F}_s] \\
 &= e^{-\gamma^2 t/2} e^{\gamma B(s)} \mathbb{E}[e^{\gamma(B(t) - B(s))}] \\
 &= e^{-\gamma^2 t/2} e^{\gamma B(s)} e^{\gamma^2 (t-s)/2} \\
 &= e^{\gamma B(s) - \gamma^2 s/2}
 \end{aligned}$$

puisque $B(t) - B(s) \sim \mathcal{N}(0, t - s)$. ■

Chapitre 7

Propriétés trajectorielles du mouvement Brownien

Par le Théorème 3.2.4 de Kolmogorov-Čentsov gaussien, nous savons que les trajectoires du mouvement Brownien sont presque sûrement Höldériennes pour tout exposant $\gamma < \frac{1}{2}$. Le mouvement Brownien présente donc une certaine régularité. Dans ce chapitre, nous montrons que ces trajectoires ne sont néanmoins pas “trop régulières”.

7.1 Temps d’atteinte

L’objectif de cette section est d’obtenir des informations supplémentaires sur la loi du temps d’atteinte T_a d’un niveau a donné pour le mouvement Brownien. On a vu dans le Corollaire 5.3.9 que $\mathbb{P}[T_a < +\infty] = 1$. Néanmoins, nous allons à présent montrer que le temps moyen pris par le mouvement Brownien pour atteindre le niveau a est infini.

Proposition 7.1.1 *Soient $a < 0 < b$. Alors*

1. $\mathbb{P}[T_a < T_b] = \frac{b}{b-a}$.
2. $\mathbb{E}[T_a \wedge T_b] = -ab$.

Démonstration : 1. On sait que T_a et T_b sont finis presque sûrement. Ainsi, si $T = T_a \wedge T_b$, alors T est un temps d’arrêt qui est également fini presque sûrement. Puisque $B(T) = a$ si $T_a < T_b$ et $B(T) = b$ si $T_b < T_a$, il vient

$$\mathbb{E}[B(T)] = a\mathbb{P}[T_a < T_b] + b\mathbb{P}[T_b < T_a] = a\mathbb{P}[T_a < T_b] + b(1 - \mathbb{P}[T_a < T_b]).$$

Si on montre que $\mathbb{E}[B(T)] = 0$, on en tirera la conclusion en résolvant l’équation par rapport à $\mathbb{P}[T_a < T_b]$.

Comme T est un temps d’arrêt du mouvement brownien, on sait par le Théorème 6.1.3 que $B(T \wedge t)$ est une martingale. En particulier, on a

$$\mathbb{E}[B(T \wedge t)] = \mathbb{E}[B(0)] = 0.$$

Puisque $|B(T \wedge t)| \leq \max\{-a, b\}$ et puisque presque sûrement

$$\lim_{t \rightarrow +\infty} B(T \wedge t) = B(T),$$

le Théorème de la convergence dominée de Lebesgue donne

$$0 = \lim_{t \rightarrow +\infty} \mathbb{E}[B(T \wedge t)] = \mathbb{E}[B(T)],$$

et on en tire la conclusion.

2. A nouveau, considérons le temps d'arrêt $T = T_a \wedge T_b$. La Proposition 6.2.2 et le Théorème 6.1.3 impliquent que le processus $B^2(T \wedge t) - (T \wedge t)$ est une martingale. Par conséquent,

$$\mathbb{E}[B^2(T \wedge t) - (T \wedge t)] = \mathbb{E}[B^2(0) - 0] = 0$$

d'où

$$\mathbb{E}[B^2(T \wedge t)] = \mathbb{E}[T \wedge t].$$

Par le Théorème de la convergence monotone, puisque T est fini presque sûrement, on a

$$\lim_{t \rightarrow +\infty} \mathbb{E}[T \wedge t] = \mathbb{E}[T].$$

De plus, $|B^2(T \wedge t)| \leq \max\{a^2, b^2\}$. En utilisant le point 1 et le Théorème de la convergence dominée, on obtient

$$\lim_{t \rightarrow +\infty} \mathbb{E}[B^2(T \wedge t)] = \mathbb{E}[B^2(T)] = a^2 \frac{b}{b-a} + b^2 \left(1 - \frac{b}{b-a}\right) = -ab.$$

Au total, on a donc

$$\mathbb{E}[T] = \lim_{t \rightarrow +\infty} \mathbb{E}[T \wedge t] = \lim_{t \rightarrow +\infty} \mathbb{E}[B^2(T \wedge t)] = -ab$$

d'où le résultat annoncé. ■

Remarque 7.1.2 Notez la ressemblance entre la relation $\mathbb{P}[T_a < T_b] = \frac{b}{b-a}$ pour le mouvement Brownien, et la relation obtenue dans l'exercice 10 de la liste 1 pour la marche aléatoire non-biaisée.

Corollaire 7.1.3 Soient $a \in \mathbb{R}_0$ et $\lambda > 0$.

1. $\mathbb{E}[T_a] = +\infty$.
2. $\mathbb{E}[e^{-\lambda T_a}] = e^{-|a|\sqrt{2\lambda}}$.

Démonstration : 1. Par symétrie du mouvement Brownien, on peut supposer que $a < 0$. Pour tout $b > 0$, on a $T_a \geq T_a \wedge T_b$. Il s'ensuit que

$$\mathbb{E}[T_a] \geq \mathbb{E}[T_a \wedge T_b] = -ab.$$

On obtient le résultat en faisant tendre b vers l'infini.

2. Par symétrie du mouvement Brownien, on peut supposer que $a > 0$. Par la Proposition 6.2.2 et le Théorème 6.1.3, on sait que pour $\gamma > 0$ fixé, le processus arrêté $\{e^{\gamma B(T_a \wedge t) - \gamma^2(T_a \wedge t)/2} : t \geq 0\}$ est une martingale. On en tire donc que

$$\mathbb{E}[e^{\gamma B(T_a \wedge t) - \gamma^2(T_a \wedge t)/2}] = \mathbb{E}[e^{\gamma B(0) - 0}] = 1$$

pour tout $t \geq 0$. Puisque

$$|e^{\gamma B(T_a \wedge t) - \gamma^2(T_a \wedge t)/2}| \leq e^{\gamma a},$$

le Théorème de la convergence dominée et la continuité presque sûre des trajectoires du mouvement Brownien donnent

$$1 = \lim_{t \rightarrow +\infty} \mathbb{E}[e^{\gamma B(T_a \wedge t) - \gamma^2(T_a \wedge t)/2}] = \mathbb{E}[e^{\gamma B(T_a) - \gamma^2 T_a/2}].$$

Puisque $B(T_a) = a$, on trouve que

$$1 = e^{\gamma a} \mathbb{E}[e^{-\gamma^2 T_a/2}]$$

et le résultat s'obtient en prenant $\gamma = \sqrt{2\lambda}$. ■

Remarque 7.1.4 Par transformée de Laplace inverse, on peut déduire du résultat précédent la loi de T_a . Remarquons qu'on aurait aussi pu le déduire des propriétés en loi du mouvement Brownien. En effet, pour tout $a > 0$

$$\begin{aligned} \mathbb{P}[T_a \leq t] &= \mathbb{P}\left[\sup_{0 \leq s \leq t} B(s) \geq a\right] = 2\mathbb{P}[B(t) \geq a] \\ &= \mathbb{P}[|B(t)| \geq a] \\ &= \mathbb{P}[B^2(t) \geq a^2] \\ &= \mathbb{P}[tB^2(1) \geq a^2] \\ &= \mathbb{P}\left[\frac{a^2}{B^2(1)} \leq t\right] \end{aligned}$$

et donc T_a a la même loi que $\frac{a^2}{B^2(1)}$.

Des résultats précédents, on déduit également des informations sur le comportement à l'infini du mouvement Brownien.

Corollaire 7.1.5 *Presque sûrement, on a $T_a < +\infty$ pour tout $a \in \mathbb{R}$ et*

$$\limsup_{t \rightarrow +\infty} B(t) = +\infty \quad \text{et} \quad \liminf_{t \rightarrow +\infty} B(t) = -\infty.$$

En particulier, B visite presque sûrement chaque réel a une infinité de fois.

Démonstration : Pour tout $q \in \mathbb{Q}$, on sait que $\mathbb{P}[T_q < +\infty] = 1$. En prenant l'intersection dénombrable, on a donc également

$$\mathbb{P}[T_q < +\infty, \forall q \in \mathbb{Q}] = 1.$$

Puisque les trajectoires de B sont presque sûrement continues, on en tire que

$$\mathbb{P}[T_a < +\infty, \forall a \in \mathbb{R}] = 1.$$

Par conséquent, les trajectoires du mouvement Brownien visitent presque sûrement tous les réels. Montrons que cela implique que presque sûrement

$$\limsup_{t \rightarrow +\infty} B(t) = +\infty.$$

Sinon, pour presque toute trajectoire, il existerait $C > 0$ et $t_0 > 0$ tels que $B(t) < C$ pour tout $t \geq t_0$. Par conséquent, la trajectoire de B devrait visiter tous les réels de $[C, +\infty[$ sur l'intervalle compact $[0, t_0]$. C'est impossible par continuité presque sûre des trajectoires. On démontre de même que presque sûrement

$$\liminf_{t \rightarrow +\infty} B(t) = -\infty.$$

Pour montrer que B visite presque sûrement chaque réel une infinité de fois, il suffit encore d'utiliser la continuité presque sûre des trajectoires. ■

Remarquons que même si B prend des valeurs arbitrairement grandes à l'infini, son comportement reste sous-linéaire : en effet, par inversion du temps, on sait que le processus $\{X(t) : t \geq 0\}$ défini par $X(0) = 0$ et $X(t) = tB(1/t)$, $t > 0$ est un mouvement Brownien. En particulier,

$$0 = \lim_{t \rightarrow 0^+} X(t) = \lim_{t \rightarrow 0^+} tB\left(\frac{1}{t}\right) = \lim_{s \rightarrow +\infty} \frac{B(s)}{s}.$$

On peut en fait caractériser plus précisément la croissance asymptotique du mouvement Brownien : on a presque sûrement

$$\limsup_{s \rightarrow +\infty} \frac{B(s)}{\sqrt{2s \log \log s}} = 1$$

et

$$\liminf_{s \rightarrow +\infty} \frac{B(s)}{\sqrt{2s \log \log s}} = -1.$$

Ce résultat s'appelle la *loi du logarithme itéré*.

7.2 Zéros du mouvement Brownien

Dans cette section, nous nous intéressons aux zéros du mouvement Brownien et nous montrons qu'ils forment un ensemble parfait de $[0, +\infty[$. Commençons par démontrer un lemme qui nous donnera directement l'existence de zéros de B arbitrairement proches de 0.

Lemme 7.2.1 (Loi forte des grands nombres) *Presque sûrement,*

$$\limsup_{s \rightarrow 0^+} \frac{B(s)}{s} = +\infty \quad \text{et} \quad \liminf_{s \rightarrow 0^+} \frac{B(s)}{s} = -\infty.$$

Démonstration : Considérons le processus $\{X(t) : t \geq 0\}$ défini par $X(0) = 0$ et $X(t) = tB(1/t)$, $t > 0$. Ce processus est un mouvement Brownien et en utilisant le Corollaire 7.1.5, on obtient

$$+\infty = \limsup_{t \rightarrow +\infty} X(t) = \limsup_{t \rightarrow +\infty} tB\left(\frac{1}{t}\right) = \limsup_{s \rightarrow 0^+} \frac{B(s)}{s}.$$

En procédant de manière similaire, on a également

$$-\infty = \liminf_{t \rightarrow +\infty} X(t) = \liminf_{t \rightarrow +\infty} tB\left(\frac{1}{t}\right) = \liminf_{s \rightarrow 0^+} \frac{B(s)}{s}.$$

■

Proposition 7.2.2 *Presque sûrement, pour tout $\varepsilon > 0$, B prend des valeurs positives et négatives dans l'intervalle $[0, \varepsilon]$. En particulier,*

$$\mathbb{P}[\inf\{t > 0 : B(t) = 0\} = 0] = 1.$$

Démonstration : Par le Lemme 7.2.1, B doit prendre une valeur positive sur tout intervalle arbitrairement proche de 0, et B doit également y prendre une valeur négative. Cela suffit. ■

Nous allons étudier la structure de l'ensemble des zéros de B . Pour cela, considérons l'ensemble aléatoire

$$Z = \{t \geq 0 : B(t) = 0\}.$$

Par le Corollaire 7.1.5, on sait déjà que Z est presque sûrement non-borné.

Rappelons qu'un ensemble $P \subseteq \mathbb{R}$ est *parfait* s'il est fermé et s'il ne possède pas de point isolé, c'est-à-dire pour tout $x \in P$ et tout $\varepsilon > 0$, il existe $y \in P$ tel que $0 < |x - y| < \varepsilon$. Notons que tout ensemble parfait est non-dénombrable (exercice).

Proposition 7.2.3 *Presque sûrement, Z est un ensemble parfait de mesure de Lebesgue nulle.*

Démonstration : Comme les trajectoires du mouvement Brownien sont presque sûrement continues, l'ensemble Z est presque sûrement fermé. Montrons qu'il est presque sûrement parfait. Pour tout $q \in \mathbb{Q}$, on considère le temps d'arrêt

$$\tau(q) = \inf\{t \geq q : B(t) = 0\}.$$

Comme Z est un ensemble presque sûrement non-borné, on sait que presque sûrement, $\tau(q)$ est fini pour tout $q \in \mathbb{Q}$. De plus, par continuité presque sûre des trajectoires de B , $\tau(q) \in Z$ presque sûrement. Par la propriété forte de Markov, le processus B_q défini par

$$B_q(t) = B(\tau(q) + t) - B(\tau(q)) = B(\tau(q) + t)$$

est un mouvement Brownien. Par conséquent, la Proposition 7.2.2 implique que presque sûrement, B_q possède un zéro dans tout intervalle $]0, \varepsilon]$, et donc $\tau(q)$ n'est pas un zéro isolé.

Supposons à présent que t_0 est un zéro de B différent de $\tau(q)$ pour tout $q \in \mathbb{Q}$. Soit $(q_n)_{n \in \mathbb{N}}$ une suite croissante de rationnels qui tend vers t_0 . Puisque $t_0 \neq \tau(q_n)$ et $q_n < t_0$, on a également $q_n \leq \tau(q_n) < t_0$. On en tire que la suite $(\tau(q_n))_{n \in \mathbb{N}}$ de zéros de B converge vers t_0 , et donc t_0 n'est pas isolé.

Montrons pour conclure que Z est un ensemble de mesure de Lebesgue nulle. Considérons le processus $X(t) = \mathbf{1}_Z(t)$. En utilisant le théorème de Fubini, si λ désigne la mesure de Lebesgue, on a

$$\mathbb{E}[\lambda(Z)] = \mathbb{E}\left[\int_{\mathbb{R}} X(t) dt\right] = \int_{\mathbb{R}} \mathbb{E}[X(t)] dt = \int_{\mathbb{R}} \mathbb{P}[B(t) = 0] dt = 0.$$

Par conséquent $\lambda(Z) = 0$ presque sûrement. ■

Remarque 7.2.4 L'ensemble aléatoire Z est presque sûrement un ensemble Lebesgue-négligeable non-dénombrable. On peut le caractériser plus précisément grâce à sa dimension de Hausdorff : celle-ci vaut $1/2$.

7.3 Non-dérivabilité

Dans cette section, nous prouvons que les trajectoires du mouvement Brownien ne sont dérivables en aucun point. C'est un exemple classique de fonction "nulle part dérivable".

Proposition 7.3.1 *Pour tout $t \geq 0$, presque sûrement, les trajectoires du mouvement Brownien ne sont pas dérivables en t .*

Démonstration : Par la propriété de Markov, il suffit de montrer la non-dérivabilité à droite en 0, c'est-à-dire montrer que

$$\lim_{t \rightarrow 0^+} \frac{B(t) - B(0)}{t - 0} = \lim_{t \rightarrow 0^+} \frac{B(t)}{t}$$

n'existe pas. Cela découle immédiatement du Lemme 7.2.1. ■

La proposition suivante est plus forte que le résultat précédent puisqu'un nombre non-dénombrable de valeurs de t est pris en compte.

Proposition 7.3.2 *Presque sûrement, pour tout $t \geq 0$, les trajectoires du mouvement Brownien ne sont pas dérivables en t .*

Démonstration : Nous allons montrer que les trajectoires du mouvement Brownien ne sont nulle part dérivables sur $[0, 1[$; en recouvrant \mathbb{R} par une union dénombrable d'intervalles et en utilisant la propriété de Markov, on obtiendra la conclusion.

Pour tout $M \in \mathbb{N}_0$, $j \in \mathbb{N}$ et $k \in \{0, \dots, 2^j - 1\}$, soit $A_{j,k}^M$ l'événement défini par

$$\left| B\left(\frac{k+n+1}{2^j}\right) - B\left(\frac{k+n}{2^j}\right) \right| \leq \frac{5M}{2^j} \quad \forall n \in \{0, 1, 2\}.$$

En utilisant l'indépendance et la stationnarité des accroissements de B , on a

$$\begin{aligned} \mathbb{P}[A_{j,k}^M] &= \prod_{n=0}^2 \mathbb{P}\left[\left|B\left(\frac{k+n+1}{2^j}\right) - B\left(\frac{k+n}{2^j}\right)\right| \leq \frac{5M}{2^j}\right] \\ &= \left(\mathbb{P}\left[\left|B\left(\frac{1}{2^j}\right)\right| \leq \frac{5M}{2^j}\right]\right)^3 \\ &= \left(\mathbb{P}\left[|Z| \leq \frac{5M}{2^{j/2}}\right]\right)^3 \\ &= \left(\frac{2}{\sqrt{2\pi}} \int_0^{\frac{5M}{2^{j/2}}} e^{-x^2/2} dx\right)^3 \\ &\leq \left(\frac{10M}{\sqrt{2\pi}} \frac{1}{2^{j/2}}\right)^3 \\ &= C_0 2^{-3j/2} \end{aligned}$$

où $Z \sim \mathcal{N}(0, 1)$ et $C_0 = \left(\frac{10M}{\sqrt{2\pi}}\right)^3$. On en tire que

$$\mathbb{P}\left[\bigcup_{k=0}^{2^j-1} A_{j,k}^M\right] \leq 2^j C_0 2^{-3j/2} = C_0 2^{-j/2}$$

qui est le terme général d'une série convergente. Le Lemme de Borel-Cantelli implique alors que

$$\mathbb{P} \left[\bigcap_{J \in \mathbb{N}} \bigcup_{j \geq J} \bigcup_{k=0}^{2^j-1} A_{j,k}^M \right] = 0.$$

Il s'ensuit que

$$\mathbb{P} \left[\bigcup_{M \in \mathbb{N}_0} \bigcap_{J \in \mathbb{N}} \bigcup_{j \geq J} \bigcup_{k=0}^{2^j-1} A_{j,k}^M \right] = 0.$$

Pour conclure, il suffit de remarquer que

$$\{\exists t \in [0, 1[: B \text{ est dérivable en } t\} \subseteq \bigcup_{M \in \mathbb{N}_0} \bigcup_{J \in \mathbb{N}} \bigcap_{j \geq J} \bigcup_{k=0}^{2^j-1} A_{j,k}^M \subseteq \bigcup_{M \in \mathbb{N}_0} \bigcap_{J \in \mathbb{N}} \bigcup_{j \geq J} \bigcup_{k=0}^{2^j-1} A_{j,k}^M.$$

En effet, si B est dérivable en $t \in [0, 1[$, il existe deux constantes $M \in \mathbb{N}_0$ et $h_0 > 0$ telles que

$$\frac{|B(t+h) - B(t)|}{h} \leq M$$

pour tout $h \leq h_0$. Soit $J \in \mathbb{N}$ tel que $\frac{3}{2^J} \leq h_0$. Pour tout $j \geq J$, il existe $k \in \{0, \dots, 2^j - 1\}$ tel que $t \in [\frac{k}{2^j}, \frac{k+1}{2^j}[$. Alors pour tout $n \in \{0, 1, 2\}$, on a

$$\begin{aligned} \left| B\left(\frac{k+n+1}{2^j}\right) - B\left(\frac{k+n}{2^j}\right) \right| &\leq \left| B\left(\frac{k+n+1}{2^j}\right) - B(t) \right| + \left| B(t) - B\left(\frac{k+n}{2^j}\right) \right| \\ &\leq M \frac{3}{2^j} + M \frac{2}{2^j} \leq \frac{5M}{2^j}, \end{aligned}$$

ce qui suffit. ■

Mentionnons qu'il est possible de raffiner le résultat précédent pour montrer que les trajectoires du mouvement Brownien sont, presque sûrement, nulle part Höldérienne d'exposant γ pour tout $\gamma > 1/2$. Le cas critique $\gamma = 1/2$ a également été étudié, via le module de continuité ([8]) : on peut montrer que presque sûrement,

$$\limsup_{h \rightarrow 0^+} \sup_{t \in [0, 1[} \sup_{|t-s| \leq h} \frac{|B(s) - B(t)|}{\sqrt{2h \log(1/h)}} = 1.$$

7.4 Variation non-bornée

Nous avons obtenu plusieurs résultats prouvant le caractère irrégulier des trajectoires du mouvement Brownien. Nous allons à présent nous intéresser aux propriétés de variation de ces trajectoires. Il s'agit d'un outil important dans la théorie de l'intégration stochastique.

Définition 7.4.1 La *variation* d'une fonction $f : [a, b] \rightarrow \mathbb{R}$ est définie par

$$\mathcal{V}_{a,b}(f) = \sup_{\pi} \sum_{k=1}^p |f(t_k) - f(t_{k-1})|,$$

où le supremum est pris sur toutes partitions $\pi = \{a = t_0 < t_1 < \dots < t_p = b\}$ de $[a, b]$. On dit que f est à *variation bornée* (sur $[a, b]$) si $\mathcal{V}_{a,b}(f) < +\infty$.

Par exemple, toute fonction monotone sur $[a, b]$ est à variation bornée. De même, toute fonction Lipschitzienne sur $[a, b]$ est à variation bornée.

Nous allons montrer que presque sûrement, les trajectoires du mouvement Brownien ne sont pas à variation bornée. Ce résultat montre que même si les trajectoires browniennes sont continues, elles oscillent énormément. Ce phénomène constitue la difficulté majeure pour construire une intégrale de type Stieltjes par rapport au mouvement Brownien.

Afin de démontrer ce résultat, nous allons nous intéresser dans un premier temps à la variation quadratique du mouvement Brownien. Pour toute partition $\pi = \{a = t_0 < t_1 < \dots < t_p = b\}$ de $[a, b]$, on pose

$$|\pi| = \max_{1 \leq k \leq p} (t_k - t_{k-1}).$$

Proposition 7.4.2 Soit $t > 0$ et pour tout $n \in \mathbb{N}$, soit $\pi_n = \{0 = t_0^{(n)} < t_1^{(n)} < \dots < t_{p_n}^{(n)} = t\}$ une partition de $[0, t]$.

1. Si $\lim_{n \rightarrow +\infty} |\pi_n| = 0$, alors

$$\sum_{k=1}^{p_n} |B(t_k^{(n)}) - B(t_{k-1}^{(n)})|^2 \longrightarrow t.$$

en norme L^2 , et donc en probabilité.

2. Si $\sum_{n \in \mathbb{N}} |\pi_n| < +\infty$, alors la convergence a également lieu presque sûrement.

Démonstration : Remarquons tout d'abord que pour tout partition $\pi = \{0 = t_0 < t_1 < \dots < t_p = t\}$ de $[0, t]$, on a

$$\mathbb{E} \left[\sum_{k=1}^p |B(t_k) - B(t_{k-1})|^2 \right] = \sum_{k=1}^p t_k - t_{k-1} = t$$

et par indépendance des accroissements,

$$\begin{aligned} \text{Var} \left[\sum_{k=1}^p |B(t_k) - B(t_{k-1})|^2 \right] &= \sum_{k=1}^p \text{Var} \left[|B(t_k) - B(t_{k-1})|^2 \right] \\ &= \sum_{k=1}^p (t_k - t_{k-1})^2 \text{Var} [|Z|^2] \\ &= \sum_{k=1}^p (t_k - t_{k-1})^2 (\mathbb{E}[|Z|^4] - (\mathbb{E}[|Z|^2])^2) \\ &= 2 \sum_{k=1}^p (t_k - t_{k-1})^2 \\ &\leq 2|\pi| \sum_{k=1}^p (t_k - t_{k-1}) \\ &= 2|\pi|t \end{aligned}$$

où $Z \sim \mathcal{N}(0, 1)$ (le moment $\mathbb{E}[|Z|^4] = 3$ peut être calculé facilement à partir de la fonction génératrice des moments).

Pour le point 1, on a donc

$$\mathbb{E} \left[\left(\sum_{k=1}^{p_n} |B(t_k^{(n)}) - B(t_{k-1}^{(n)})|^2 - t \right)^2 \right] = \text{Var} \left[\sum_{k=1}^{p_n} |B(t_k^{(n)}) - B(t_{k-1}^{(n)})|^2 \right] \leq 2|\pi_n|t \rightarrow 0$$

lorsque n tend vers l'infini.

Pour le point 2, on utilise l'inégalité de Bienaymé-Tchebychev pour obtenir

$$\mathbb{P} \left[\left| \sum_{k=1}^{p_n} |B(t_k^{(n)}) - B(t_{k-1}^{(n)})|^2 - t \right| \geq \varepsilon \right] \leq \frac{2|\pi_n|t}{\varepsilon^2}.$$

En utilisant l'hypothèse, il vient

$$\sum_{n \in \mathbb{N}} \mathbb{P} \left[\left| \sum_{k=1}^{p_n} |B(t_k^{(n)}) - B(t_{k-1}^{(n)})|^2 - t \right| \geq \varepsilon \right] < +\infty$$

et le Lemme de Borel-Cantelli implique que

$$\mathbb{P} \left[\bigcup_{N \in \mathbb{N}} \bigcap_{n \geq N} \left\{ \left| \sum_{k=1}^{p_n} |B(t_k^{(n)}) - B(t_{k-1}^{(n)})|^2 - t \right| < \varepsilon \right\} \right] = 1.$$

Ainsi, presque sûrement,

$$\left| \sum_{k=1}^{p_n} |B(t_k^{(n)}) - B(t_{k-1}^{(n)})|^2 - t \right| < \varepsilon$$

pour tout n suffisamment grand. En prenant une suite ε_n qui converge vers 0, on obtient la convergence presque sûre de l'énoncé. \blacksquare

Proposition 7.4.3 *Presque sûrement, les trajectoires du mouvement Brownien sont à variation non-bornée sur tout intervalle non-trivial.*

Démonstration : En utilisant la propriété de Markov, il suffit de montrer que le mouvement Brownien est à variation non-bornée presque sûrement sur tout intervalle de la forme $[0, t]$. De plus, comme il suffit de regarder un nombre dénombrable d'intervalles pour avoir le comportement sur tout intervalle, on peut montrer que sur $[0, t]$, presque sûrement, les trajectoires du mouvement Brownien ne sont pas à variation bornée¹. Soit $\pi_n = \{0 = t_0^{(n)} < t_1^{(n)} < \dots < t_{p_n}^{(n)} = t\}$ une suite de partitions de $[0, t]$ telle que $\sum_{n=0}^{+\infty} |\pi_n| < +\infty$. Par la Proposition 7.4.2, on sait que

$$\sum_{k=1}^{p_n} |B(t_k^{(n)}) - B(t_{k-1}^{(n)})|^2$$

1. On peut également regarder ce qui se passe sur l'intervalle $[0, 1]$ puis utiliser la propriété d'autosimilarité du mouvement Brownien.

converge presque sûrement vers t . On a donc

$$t = \lim_{n \rightarrow +\infty} \sum_{k=1}^{p_n} |B(t_k^{(n)}) - B(t_{k-1}^{(n)})|^2 \leq \lim_{n \rightarrow +\infty} \left(\max_{1 \leq k \leq p_n} |B(t_k^{(n)}) - B(t_{k-1}^{(n)})| \right) \sum_{k=1}^{p_n} |B(t_k^{(n)}) - B(t_{k-1}^{(n)})|$$

presque sûrement. Puisque $\lim_{n \rightarrow +\infty} \left(\max_{1 \leq k \leq p_n} (t_k^{(n)} - t_{k-1}^{(n)}) \right) = 0$ et la continuité des trajectoires² implique que presque sûrement $\lim_{n \rightarrow +\infty} \left(\max_{1 \leq k \leq p_n} |B(t_k^{(n)}) - B(t_{k-1}^{(n)})| \right) = 0$. Par conséquent, il faut avoir

$$\lim_{n \rightarrow +\infty} \sum_{k=1}^{p_n} |B(t_k^{(n)}) - B(t_{k-1}^{(n)})| = +\infty,$$

ce qui montre que presque sûrement, $\mathcal{V}_{0,t}(B) = +\infty$. ■

2. On peut également utiliser la régularité Höldérienne des trajectoires pour $\gamma < 1/2$.

Chapitre 8

L'intégrale d'Itô

Le but de ce chapitre est de présenter le calcul stochastique d'Itô. Il s'agit de donner du sens à des équations différentielles stochastiques de la forme

$$dX(t) = g(t, X(t)) + f(t, X(t))dB(t).$$

Ces équations sont présentes dans de nombreuses applications (évolution d'un actif financier, croissance démographique,...) où il est apparu que les modèles étaient plus réalistes lorsque l'on autorise les coefficients des équations à être aléatoires. Ainsi, $X(t)$ représente un signal au temps t et $dX(t)$ son accroissement résultant d'un accroissement dt du temps infiniment petit. Les deux fonctions g, f sont des fonctions déterministes. Comme dans le cas des équations différentielles ordinaires l'équation différentielle stochastique s'écrit mathématiquement

$$X(t) = X(0) + \int_0^t g(s, X(s)) ds + \int_0^t f(s, X(s)) dB(s).$$

La première intégrale est une intégrale de Lebesgue classique calculée trajectoire par trajectoire. Il s'agit donc de donner du sens à la deuxième intégrale.

De manière générale, si $\{\Phi(t) : t \in [0, T]\}$ est un processus stochastique, on cherche à définir le processus

$$Y(t) = \int_0^t \Phi(s) dB(s).$$

Le problème est de donner du sens à l'élément différentiel $dB(s)$ puisque les trajectoires de B ne sont pas dérivables (ou même à variation bornée).

Rappelons en effet qu'une méthode d'intégration classique généralisant la méthode de Riemann est donnée par l'intégrale de Stieltjes (trajectoire par trajectoire) définie par

$$\int_a^b v(x) du(x) = \lim \sum_{i=1}^n v(\xi_i)(u(x_i) - u(x_{i-1}))$$

avec $a = x_0 < x_1 < \dots < x_n = b$, $\xi_i \in [x_{i-1}, x_i]$ et où le pas de la partition tend vers 0. Dans le cas où u est dérivable, alors si l'intégrale existe, on retrouve simplement l'intégrale

$$\int_a^b v(x) du(x) = \int_a^b v(x)u'(x) dx$$

et la formule d'intégration par partie. On peut considérer également à des intégrateurs plus irréguliers, mais une condition classique d'existence de l'intégrale de Stieltjes est l'hypothèse de variation bornée.

Un objet mathématique adéquat a été proposé par Itô, sous l'hypothèse cruciale d'adaptation du processus Φ à la filtration naturelle du mouvement Brownien.

8.1 Construction de l'intégrale d'Itô

Notre but est de donner du sens à des intégrales du type

$$\int_0^T \Phi(s) dB(s)$$

où $\{\Phi(s) : s \in [0, T]\}$ est un processus stochastique.

Commençons par définir la classe des intégrands aléatoires qui vont nous intéresser. Dans ce qui suit, nous désignerons par $(\mathcal{F}_t)_{t \geq 0}$ la filtration naturelle du mouvement Brownien.

Définition 8.1.1 Un processus stochastique $\{\Phi(t) : t \in [0, T]\}$ appartient à la classe $\mathcal{H}^2([0, T])$ si

1. l'application $(t, \omega) \in [0, T] \times \Omega \mapsto \Phi(t, \omega)$ est $\mathcal{B}([0, T]) \times \mathcal{F}$ -mesurable,
2. Φ est adapté à la filtration naturelle $(\mathcal{F}_t)_{t \in [0, T]}$ du mouvement Brownien,
3. Φ vérifie la condition d'intégrabilité

$$\mathbb{E} \left[\int_0^T \Phi^2(s) ds \right] < +\infty.$$

Remarquons que la condition d'intégrabilité est en fait une intégrale double et la première condition permet de donner du sens à cette double intégrale. Ces deux conditions impliquent donc $\mathcal{H}^2([0, T]) \subseteq L^2([0, T] \times \Omega)$. De plus, intuitivement parlant, dire que le processus est adapté à la filtration naturelle du mouvement Brownien signifie que $\Phi(t)$ ne dépend que de l'histoire du processus B jusqu'au temps t : en fait, presque sûrement, $\Phi(t)$ peut être obtenu à partir de $B(s)$, $s \in [0, t]$, au moyen d'un procédé déterministe.

Commençons par regarder ce que nous souhaiterions que l'intégrale d'Itô donne dans le cas le plus simple. Supposons donc que Φ est simplement une fonction déterministe donnée par l'indicatrice d'un intervalle $[a, b[$. Il semble alors naturel d'imposer que

$$\int_0^T \Phi(s) dB(s) = \int_a^b dB(s) = B(b) - B(a).$$

Comme on s'attend à ce que l'intégrale soit linéaire, on est amenés à considérer des sommes d'indicatrices.

Définition 8.1.2 La classe $\mathcal{H}_0^2([0, T])$ des processus *élémentaires* est l'ensemble des processus Φ pouvant s'écrire sous la forme

$$\Phi(t, \omega) = \sum_{i=0}^{n-1} a_i(\omega) \mathbf{1}_{[t_i, t_{i+1}[}(t)$$

où $0 \leq t_0 < t_1 < \dots < t_n \leq T$ et où pour tout $i \in \{0, \dots, n-1\}$, la variable aléatoire réelle a_i est \mathcal{F}_{t_i} -mesurable et satisfait $\mathbb{E}[a_i^2] < +\infty$ (de manière à avoir $\mathcal{H}_0^2([0, T]) \subseteq \mathcal{H}^2([0, T])$). L'intégrale d'Itô de Φ est alors définie par

$$\int_0^T \Phi(s) dB(s) := \sum_{i=0}^{n-1} a_i (B(t_{i+1}) - B(t_i)).$$

On note cette intégrale $I_T(\Phi)$.

Le lemme suivant donne une propriété d'isométrie de l'intégrale d'Itô dans le cas des processus élémentaires. Il s'agit d'un résultat clé pour l'extension de la définition de l'intégrale à la classe $\mathcal{H}^2([0, T])$.

Lemme 8.1.3 (Isométrie d'Itô pour les processus élémentaires) *Pour tout processus élémentaire $\Phi \in \mathcal{H}_0^2([0, T])$, on a*

$$\mathbb{E} \left[\left(\int_0^T \Phi(s) dB(s) \right)^2 \right] = \mathbb{E} \left[\int_0^T \Phi^2(s) ds \right].$$

Remarquons que cela revient à dire que $\|I_T(\phi)\|_{L^2(\Omega)} = \|\Phi\|_{L^2([0, T] \times \Omega)}$, ce qui montre que l'application $I_T : \Phi \in \mathcal{H}_0^2([0, T]) \mapsto I_T(\Phi) \in L^2(\Omega)$ préserve la norme (c'est une isométrie). En particulier, I_T transforme toute suite de Cauchy en une suite de Cauchy. Ce point sera crucial pour la construction générale de l'intégrale d'Itô.

Remarquons également que

$$\mathbb{E} \left[\int_0^T \Phi(s) dB(s) \right] = \sum_{i=0}^{n-1} \mathbb{E}[a_i] \mathbb{E}[B(t_{i+1}) - B(t_i)] = 0$$

puisque pour tout i , a_i est \mathcal{F}_{t_i} -mesurable et $B(t_{i+1}) - B(t_i)$ est indépendant de \mathcal{F}_{t_i} par la propriété de Markov. Ainsi,

$$\text{Var} \left[\int_0^T \Phi(s) dB(s) \right] = \mathbb{E} \left[\left(\int_0^T \Phi(s) dB(s) \right)^2 \right]$$

et l'isométrie d'Itô donne donc un calcul de la variance de l'intégrale stochastique.

Démonstration : Il s'agit d'une simple vérification. Supposons que

$$\Phi(t, \omega) = \sum_{i=0}^{n-1} a_i(\omega) \mathbf{1}_{[t_i, t_{i+1}[}(t).$$

Alors, puisque les intervalles intervenant dans la somme sont deux à deux disjoints, on a

$$\Phi^2(t) = \sum_{i=0}^{n-1} a_i^2 \mathbf{1}_{[t_i, t_{i+1}[}(t)$$

et donc

$$\mathbb{E} \left[\int_0^T \Phi^2(s) ds \right] = \sum_{i=0}^{n-1} \mathbb{E}[a_i^2] (t_{i+1} - t_i).$$

D'autre part, on a également

$$\mathbb{E} \left[\left(\int_0^T \Phi(s) dB(s) \right)^2 \right] = \mathbb{E} \left[\sum_{i=0}^{n-1} \sum_{j=0}^{n-1} a_i a_j (B(t_{i+1}) - B(t_i))(B(t_{j+1}) - B(t_j)) \right].$$

Si $i < j$, on sait que les variables a_i , a_j et $B(t_{i+1}) - B(t_i)$ sont \mathcal{F}_{t_j} -mesurables. De plus, par la propriété de Markov, $B(t_{j+1}) - B(t_j)$ est indépendant de \mathcal{F}_{t_j} . Par conséquent,

$$\begin{aligned} \mathbb{E} [a_i a_j (B(t_{i+1}) - B(t_i))(B(t_{j+1}) - B(t_j))] &= \mathbb{E} [a_i a_j (B(t_{i+1}) - B(t_i))] \mathbb{E} [(B(t_{j+1}) - B(t_j))] \\ &= 0. \end{aligned}$$

On procède de même si $j < i$. Ainsi, on obtient que

$$\begin{aligned} \mathbb{E} \left[\left(\int_0^T \Phi(s) dB(s) \right)^2 \right] &= \sum_{i=0}^{n-1} \mathbb{E} [a_i^2 (B(t_{i+1}) - B(t_i))^2] \\ &= \sum_{i=0}^{n-1} \mathbb{E} [a_i^2] \mathbb{E} [(B(t_{i+1}) - B(t_i))^2] \\ &= \sum_{i=0}^{n-1} \mathbb{E} [a_i^2] (t_{i+1} - t_i) \end{aligned}$$

puisque a_i est \mathcal{F}_{t_i} -mesurable et que $B(t_{i+1}) - B(t_i)$ est indépendant de \mathcal{F}_{t_i} par la propriété de Markov. Cela conclut la preuve. \blacksquare

L'idée d'Itô pour définir l'intégrale d'un processus Φ de $\mathcal{H}^2([0, T])$ est de trouver une suite de processus élémentaires $(\Phi_n)_{n \in \mathbb{N}}$ approchant Φ au sens de la norme de $L^2([0, T] \times \Omega)$, c'est-à-dire

$$\lim_{n \rightarrow +\infty} \mathbb{E} \left[\int_0^T (\Phi(s) - \Phi_n(s))^2 ds \right] = 0.$$

L'isométrie d'Itô permettra alors de définir

$$\int_0^T \Phi(s) dB(s) := \lim_{n \rightarrow +\infty} \int_0^T \Phi_n(s) dB(s)$$

où la limite prise dans $L^2(\Omega)$. En effet, si la suite $(\Phi_n)_{n \in \mathbb{N}}$ converge dans $L^2([0, T] \times \Omega)$, elle y est de Cauchy. Le Lemme 8.1.3 permet alors d'affirmer que la suite $(I_T(\Phi_n))_{n \in \mathbb{N}}$ est de Cauchy dans $L^2(\Omega)$ et y converge donc. La convergence étant assurée par l'isométrie d'Itô, il reste à montrer que

- la classe $\mathcal{H}_0^2([0, T])$ des processus élémentaires est dense dans la classe $\mathcal{H}^2([0, T])$ au sens décrit ci-dessus ;
- l'intégrale est bien définie, c'est-à-dire qu'elle ne dépend pas de la suite choisie pour la définir.

Le deuxième point se démontre de manière classique : soit $(\tilde{\Phi}_n)_{n \in \mathbb{N}}$ est une deuxième suite de $\mathcal{H}_0([0, T])$ qui converge vers Φ au sens de la norme de $L^2([0, T] \times \Omega)$. On forme la suite $(\Psi_n)_{n \in \mathbb{N}}$ de processus élémentaires en posant

$$\Psi_{2n} = \Phi_n \quad \text{et} \quad \Psi_{2n+1} = \tilde{\Phi}_n.$$

Bien sûr, la suite $(\Psi_n)_{n \in \mathbb{N}}$ converge vers Φ dans $L^2([0, T] \times \Omega)$. Par le raisonnement fait précédemment, la suite $(I_T(\Psi_n))_{n \in \mathbb{N}}$ converge dans $L^2(\Omega)$. Comme cette suite admet comme sous-suites les suites $(I_T(\Phi_n))_{n \in \mathbb{N}}$ et $(I_T(\tilde{\Phi}_n))_{n \in \mathbb{N}}$, elles doivent converger vers la même limite.

Pour le premier point (la densité de $\mathcal{H}_0^2([0, T])$), nous allons admettre le résultat suivant (voir [8] par exemple).

Proposition 8.1.4 *Pour tout $\Phi \in \mathcal{H}^2([0, T])$, il existe une suite $(\Phi_n)_{n \in \mathbb{N}}$ de $\mathcal{H}_0^2([0, T])$ satisfaisant*

$$\lim_{n \rightarrow +\infty} \mathbb{E} \left[\int_0^T (\Phi(s) - \Phi_n(s))^2 ds \right] = 0.$$

Nous pouvons à présente donner la définition propre de l'intégrale d'Itô.

Définition 8.1.5 Soit $\Phi \in \mathcal{H}^2([0, T])$ et soit $(\Phi_n)_{n \in \mathbb{N}}$ une suite de $\mathcal{H}_0^2([0, T])$ qui converge vers Φ dans $L^2(\Omega)$. L'intégrale d'Itô de Φ est la limite de la suite de variables aléatoires

$$\int_0^T \Phi_n(s) dB(s)$$

au sens de la norme $L^2(\Omega)$. Cette limite ne dépend pas de la suite $(\Phi_n)_{n \in \mathbb{N}}$ choisie. On la note $I_T(\Phi)$ ou

$$\int_0^T \Phi(s) dB(s).$$

Notons que l'isométrie d'Itô des processus élémentaires s'étend à la classe $\mathcal{H}^2([0, T])$, en utilisant un simple passage à la limite.

Proposition 8.1.6 (Isométrie d'Itô) *Pour tout $\Phi \in \mathcal{H}^2([0, T])$, on a*

$$\mathbb{E} \left[\left(\int_0^T \Phi(s) dB(s) \right)^2 \right] = \mathbb{E} \left[\int_0^T \Phi^2(s) ds \right].$$

Les propriétés suivantes sont immédiates à vérifier pour les processus élémentaires, et par passage à la limite pour les processus de $\mathcal{H}^2([0, T])$. Le calcul de la variance au point 2 est une conséquence de la propriété d'isométrie de l'intégrale d'Itô.

Proposition 8.1.7 *Soient Φ et Ψ deux processus de $\mathcal{H}^2([0, T])$.*

1. *Linéarité : on a*

$$\int_0^T (\Phi(s) + \Psi(s)) dB(s) = \int_0^T \Phi(s) dB(s) + \int_0^T \Psi(s) dB(s)$$

et

$$\int_0^T \alpha \Phi(s) dB(s) = \alpha \int_0^T \Phi(s) dB(s).$$

2. *Espérance et variance :*

$$\mathbb{E} \left[\int_0^T \Phi(s) dB(s) \right] = 0, \quad \text{Var} \left[\int_0^T \Phi(s) dB(s) \right] = \mathbb{E} \left[\int_0^T \Phi^2(s) ds \right].$$

3. *La variable aléatoire $\int_0^T \Phi(s) dB(s)$ est \mathcal{F}_T -mesurable.*

Enfin, pour tout $0 \leq u < t \leq T$, on définit

$$\int_u^t \Phi(s) dB(s) := \int_0^t \Phi(s) \mathbf{1}_{[u,t]} dB(s)$$

et on vérifie directement le résultat suivant.

Proposition 8.1.8 *Soient Φ et Ψ deux processus de $\mathcal{H}^2([0, T])$. Pour tous $0 \leq v < u < t \leq T$, on a*

$$\int_v^t \Phi(s) dB(s) = \int_v^u \Phi(s) dB(s) + \int_u^t \Phi(s) dB(s).$$

8.2 Cas déterministe

Supposons que $f : [0, T] \rightarrow \mathbb{R}$ est une fonction déterministe. On souhaite étudier

$$\int_0^T f(s) dB(s).$$

Il s'agit d'un cas particulier du cas étudié précédemment, avec $\Phi(t, \omega) = f(t)$ pour tout $t \in [0, T]$ et $\omega \in \Omega$. Afin d'appartenir à la classe $\mathcal{H}^2([0, T])$, on doit imposer à la fonction f d'appartenir à $L^2([0, T])$. En effet, dans ce cas, on a

$$\mathbb{E} \left[\int_0^T \Phi^2(s) ds \right] = \int_0^T f^2(s) ds < +\infty.$$

Si $f \in L^2([0, T])$, on sait qu'il existe une suite $(f_n)_{n \in \mathbb{N}}$ de fonctions de la forme

$$f_n(t) = \sum_{k=0}^{p_n-1} a_{n,k} \mathbf{1}_{[t_{n,k}, t_{n,k+1}[}(t)$$

où $p_n \in \mathbb{N}$, $a_{k,n} \in \mathbb{R}$ et $0 = t_{n,0} < \dots < t_{n,p_n} = T$, et qui converge en norme $L^2([0, T])$ vers f . Dans le cas des fonctions f continues, on a le résultat suivant.

Proposition 8.2.1 *Soit $f : [0, T] \rightarrow \mathbb{R}$ une fonction continue. Alors la suite $(f_n)_{n \in \mathbb{N}_0}$ définie par*

$$f_n(t) = \sum_{k=0}^{n-1} f\left(\frac{kT}{n}\right) \mathbf{1}_{\left[\frac{kT}{n}, \frac{(k+1)T}{n}\right[}(t)$$

converge dans $L^2([0, T])$ vers f .

Démonstration : Soit $\varepsilon > 0$. Comme f est continue sur l'intervalle compact $[0, T]$, elle y est uniformément continue. Par conséquent, il existe $\delta > 0$ tel que $|f(x) - f(y)| < \sqrt{\frac{\varepsilon}{T}}$ si $|x - y| < \delta$. Fixons $N \geq 1$ tel que $\frac{T}{N} < \delta$. Pour tout $n \geq N$, on a

$$\begin{aligned}
 \int_0^T |f_n(t) - f(t)|^2 dt &= \int_0^T \left| \sum_{k=0}^{n-1} f\left(\frac{kT}{n}\right) \mathbf{1}_{\left[\frac{kT}{n}, \frac{(k+1)T}{n}\right]}(t) - f(t) \sum_{k=0}^{n-1} \mathbf{1}_{\left[\frac{kT}{n}, \frac{(k+1)T}{n}\right]}(t) \right|^2 dt \\
 &= \int_0^T \sum_{k=0}^{n-1} \left(f\left(\frac{kT}{n}\right) - f(t) \right)^2 \mathbf{1}_{\left[\frac{kT}{n}, \frac{(k+1)T}{n}\right]}(t) dt \\
 &= \sum_{k=0}^{n-1} \int_{\frac{kT}{n}}^{\frac{(k+1)T}{n}} \left(f\left(\frac{kT}{n}\right) - f(t) \right)^2 dt \\
 &\leq \sum_{k=0}^{n-1} \int_{\frac{kT}{n}}^{\frac{(k+1)T}{n}} \frac{\varepsilon}{T} dt \\
 &= \varepsilon
 \end{aligned}$$

puisque les intervalles $\left[\frac{kT}{n}, \frac{(k+1)T}{n}\right]$ sont deux à deux disjoints et puisque $t \in \left[\frac{kT}{n}, \frac{(k+1)T}{n}\right]$ implique $\left|\frac{kT}{n} - t\right| \leq \frac{T}{n} < \delta$. ■

Corollaire 8.2.2 Soit $f : [0, T] \rightarrow \mathbb{R}$ une fonction continue. Alors, l'intégrale d'Itô

$$\int_0^T f(s) dB(s)$$

suit une loi normale centrée et de variance $\int_0^T f^2(s) ds$.

Démonstration : Par la Proposition 8.2.1, on sait que

$$\int_0^T f(s) dB(s) = \lim_{n \rightarrow +\infty} \sum_{k=0}^{n-1} f\left(\frac{kT}{n}\right) \left(B\left(\frac{(k+1)T}{n}\right) - B\left(\frac{kT}{n}\right) \right)$$

où la limite a lieu en norme $L^2(\Omega)$. Remarquons que chacun des variables aléatoires

$$\sum_{k=0}^{n-1} f\left(\frac{kT}{n}\right) \left(B\left(\frac{(k+1)T}{n}\right) - B\left(\frac{kT}{n}\right) \right)$$

suit une loi normale centrée par indépendance des accroissements du mouvement Brownien. En prenant la limite des variances (sommes de Riemann), on obtient la loi de l'intégrale d'Itô. On peut aussi remarquer que, par l'isométrie d'Itô, on a

$$\left[\int_0^T f(s) dB(s) \right] = \mathbb{E} \left[\left(\int_0^T f(s) dB(s) \right)^2 \right] = \mathbb{E} \left[\int_0^T f^2(s) ds \right] = \int_0^T f^2(s) ds.$$

Par le même raisonnement, ce résultat reste évidemment valide sous l'hypothèse moins forte $f \in L^2([0, T])$. ■

8.3 Un premier exemple

Dans cette section, nous allons calculer l'intégrale du mouvement Brownien par rapport à lui-même.

Proposition 8.3.1 *On a*

$$\int_0^T B(s) dB(s) = \frac{1}{2}(B^2(T) - T).$$

Ce premier calcul d'intégrale stochastique montre que la formule d'intégration par parties n'admet pas d'extension triviale au cas de l'intégrale d'Itô : on y voit apparaître un deuxième terme déterministe, qui est lié à la variation quadratique non-nulle du mouvement Brownien. Ce terme se retrouvera dans la formule d'Itô.

Démonstration : Commençons par remarquer que $B \in \mathcal{H}^2([0, T])$ puisque par Fubini, on a

$$\mathbb{E} \left[\int_0^T B^2(s) ds \right] = \int_0^T \mathbb{E}[B^2(s)] ds = \int_0^T s ds = \frac{T^2}{2} < +\infty.$$

Construisons à présent une suite de $\mathcal{H}_0([0, T])$ qui converge vers B . Pour tout $j \in \mathbb{N}$, on pose

$$B_j(s) = \sum_{k=0}^{2^j-1} B\left(\frac{kT}{2^j}\right) \mathbf{1}_{\left[\frac{kT}{2^j}, \frac{(k+1)T}{2^j}\right]}(s).$$

On a

$$\begin{aligned} \mathbb{E} \left[\int_0^T (B(s) - B_j(s))^2 ds \right] &= \mathbb{E} \left[\int_0^T \left(\sum_{k=0}^{2^j-1} \left(B(s) - B\left(\frac{kT}{2^j}\right) \right) \mathbf{1}_{\left[\frac{kT}{2^j}, \frac{(k+1)T}{2^j}\right]}(s) \right)^2 ds \right] \\ &= \mathbb{E} \left[\int_0^T \sum_{k=0}^{2^j-1} \left(B(s) - B\left(\frac{kT}{2^j}\right) \right)^2 \mathbf{1}_{\left[\frac{kT}{2^j}, \frac{(k+1)T}{2^j}\right]}(s) ds \right] \\ &= \mathbb{E} \left[\sum_{k=0}^{2^j-1} \int_{\frac{kT}{2^j}}^{\frac{(k+1)T}{2^j}} \left(B(s) - B\left(\frac{kT}{2^j}\right) \right)^2 ds \right] \\ &= \sum_{k=0}^{2^j-1} \int_{\frac{kT}{2^j}}^{\frac{(k+1)T}{2^j}} \mathbb{E} \left[\left(B(s) - B\left(\frac{kT}{2^j}\right) \right)^2 \right] ds \\ &= \sum_{k=0}^{2^j-1} \int_{\frac{kT}{2^j}}^{\frac{(k+1)T}{2^j}} \left(s - \frac{kT}{2^j} \right) ds \\ &= \sum_{k=0}^{2^j-1} \frac{1}{2} \left(\frac{T}{2^j} \right)^2 \\ &= \frac{T^2}{2} 2^{-j} \end{aligned}$$

et donc

$$\lim_{j \rightarrow +\infty} \mathbb{E} \left[\int_0^T (B(s) - B_j(s))^2 ds \right] = 0.$$

Ainsi, $(B_j)_{j \in \mathbb{N}}$ est bien une suite de processus élémentaires de $\mathcal{H}_0^2([0, T])$ qui converge dans $L^2([0, T] \times \Omega)$ vers B .

Il reste à montrer que

$$\int_0^T B_j(s) dB(s) = \sum_{k=0}^{2^j-1} B\left(\frac{kT}{2^j}\right) \left(B\left(\frac{(k+1)T}{2^j}\right) - B\left(\frac{kT}{2^j}\right) \right)$$

converge vers en norme L^2 vers $\frac{1}{2}(B^2(T) - T)$. Remarquons que

$$a^2 - b^2 = a^2 + b^2 - 2ab - 2b^2 + 2ab = (a - b)^2 + 2b(a - b)$$

et donc

$$b(a - b) = \frac{1}{2}(a^2 - b^2 - (a - b)^2).$$

Par conséquent, on a

$$\begin{aligned} & \sum_{k=0}^{2^j-1} B\left(\frac{kT}{2^j}\right) \left(B\left(\frac{(k+1)T}{2^j}\right) - B\left(\frac{kT}{2^j}\right) \right) \\ &= \frac{1}{2} \sum_{k=0}^{2^j-1} \left(B^2\left(\frac{(k+1)T}{2^j}\right) - B^2\left(\frac{kT}{2^j}\right) \right) - \frac{1}{2} \sum_{k=0}^{2^j-1} \left(B\left(\frac{(k+1)T}{2^j}\right) - B\left(\frac{kT}{2^j}\right) \right)^2 \\ &= \frac{1}{2} B^2(T) - \frac{1}{2} \sum_{k=0}^{2^j-1} \left(B\left(\frac{(k+1)T}{2^j}\right) - B\left(\frac{kT}{2^j}\right) \right)^2. \end{aligned}$$

On sait déjà par la Proposition 7.4.2 que

$$\sum_{k=0}^{2^j-1} \left(B\left(\frac{(k+1)T}{2^j}\right) - B\left(\frac{kT}{2^j}\right) \right)^2 \longrightarrow T$$

en norme L^2 lorsque j tend vers l'infini. Cela donne donc le résultat annoncé. \blacksquare

8.4 L'intégrale d'Itô comme processus stochastique

Afin d'associer un processus stochastique à l'intégrale stochastique d'un processus Φ , il semble naturel de considérer l'intégrale de $\Phi_t = \Phi \mathbf{1}_{[0,t]}$, c'est-à-dire considérer le processus défini pour tout $t \in [0, T]$ par

$$\int_0^t \Phi(s) dB(s) = \int_0^T \Phi_t(s) dB(s).$$

Il faut néanmoins être prudent avec cette construction. Pour chaque t fixé, l'intégrale est une limite en norme $L^2(\Omega)$ et est donc définie sur un événement Ω_t de probabilité 1 (en dehors de cet événement, la variable aléatoire peut prendre des valeurs arbitraires). En faisant varier t , puisque $[0, T]$ n'est pas dénombrable, on pourrait travailler sur un événement potentiellement vide! En fait, on peut montrer qu'il existe un processus X à trajectoires continues tel que

$$\mathbb{P} \left[X(t) = \int_0^T \Phi_t(s) dB(s) \right] = 1$$

pour tout $t \in [0, T]$ (voir [8] par exemple). Ainsi, la famille $(\int_0^t \Phi(s) dB(s))_{t \in [0, T]}$ admet une modification à trajectoires continues, et c'est avec cette modification que l'on travaille.

Proposition 8.4.1 Soit Φ un processus de $\mathcal{H}^2([0, T])$. Le processus stochastique

$$\left\{ \int_0^t \Phi(s) dB(s) : t \in [0, T] \right\}$$

est une martingale pour la filtration naturelle du mouvement Brownien.

Vu les propriétés de l'intégrale d'Itô, il suffit de montrer que pour tous $0 \leq s < t \leq T$, on a

$$\mathbb{E} \left[\int_s^t \Phi(u) dB(u) \mid \mathcal{F}_s \right] = 0$$

puisque $\int_0^s \Phi(u) dB(u)$ est \mathcal{F}_s -mesurable. On démontre cette relation pour un processus de $\mathcal{H}_0^2([0, T])$ et on considère ensuite les processus de $\mathcal{H}^2([0, T])$ par passage à la limite.

8.5 Extension de la classe des intégrants

La classe d'intégrants $\mathcal{H}^2([0, T])$, et plus particulièrement la condition d'intégrabilité, est parfois trop restrictive dans la pratique. En effet, si $f : \mathbb{R} \rightarrow \mathbb{R}$ est une fonction continue, on souhaiterait pouvoir définir

$$\int_0^T f(B(s)) dB(s).$$

La condition d'intégrabilité est trop forte pour de nombreuses fonctions, comme par exemple $f(x) = \exp(x^4)$. Il est néanmoins possible de contourner ce problème en utilisant une méthode de localisation basée sur les temps d'arrêt et les martingales locales ([8]), et de relaxer la condition par la contrainte moins restrictive

$$\mathbb{P} \left[\int_0^T \Phi^2(s) ds < +\infty \right] = 1.$$

La classe des intégrants est ainsi élargie et cela permet notamment de considérer des intégrants du type $f(B(s))$ où f est une fonction continue (puisque alors $f(B(s))$ est presque sûrement continu, et donc presque sûrement de carré intégrable sur l'intervalle compact $[0, T]$).

La définition de l'intégrale d'Itô dans ce cadre plus général nécessite un développement plus technique, mais il y a des cas importants où une représentation concrète est possible : en particulier, pour des fonctions continues du mouvement Brownien, l'intégrale d'Itô peut être interprétée via les sommes de Riemann (voir [8]).

Proposition 8.5.1 Pour toute fonction $f : \mathbb{R} \rightarrow \mathbb{R}$ continue, on a

$$\sum_{k=0}^{n-1} f \left(B \left(\frac{kT}{n} \right) \right) \left(B \left(\frac{(k+1)T}{n} \right) - B \left(\frac{kT}{n} \right) \right) \xrightarrow{\mathbb{P}} \int_0^T f(B(s)) dB(s)$$

lorsque n tend vers l'infini.

On retrouve donc le même genre de résultat que dans le cas déterministe. Notons que pour prouver que cette convergence a lieu dans $L^2(\Omega)$ (première étape avant de montrer la convergence en probabilité), on procède comme dans le cas de la preuve de la Proposition 8.2.1 puisque les trajectoires de $f \circ B$ sont presque sûrement continues.

8.6 La formule d'Itô

Nous disposons de nombreux outils pour le calcul d'intégrales classiques (non stochastiques), permettant de ne pas avoir recours à la définition de celles-ci. Citons-notamment le théorème fondamental du calcul intégral et la méthode d'intégration par parties. La situation avec l'intégrale d'Itô est similaire, ce qui a permis à cette intégrale de prendre une place importante dans la théorie du calcul stochastique.

Théorème 8.6.1 (Première formule d'Itô) *Si $f : \mathbb{R} \rightarrow \mathbb{R}$ est une fonction de classe C^2 , alors*

$$f(B(t)) = f(0) + \int_0^t f'(B(s)) dB(s) + \frac{1}{2} \int_0^t f''(B(s)) ds$$

pour tout $t \geq 0$.

Notons que les intégrales ont du sens au vu de la continuité de f' et f'' et des commentaires présentés dans la section précédente. Si $Y(t) = f(B(t))$, cette formule s'écrit souvent sous la forme condensée

$$dY(t) = f'(B(t)) dB(t) + \frac{1}{2} f''(B(t)) dt.$$

Démonstration : Nous présentons uniquement les idées générales de la preuve. Pour tous $n \in \mathbb{N}_0$ et $k \in \{0, \dots, n\}$, on note $t_{n,k} = \frac{kt}{n}$. En utilisant la formule de Taylor, on a

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

Ainsi,

$$\begin{aligned} f(B(t)) - f(0) &= \sum_{k=1}^{n-1} f(B(t_{n,k+1})) - f(B(t_{n,k})) \\ &= \sum_{k=1}^{n-1} f'(B(t_{n,k})) (B(t_{n,k+1}) - B(t_{n,k})) + \sum_{k=1}^{n-1} \frac{1}{2} f''(B(t_{n,k})) (B(t_{n,k+1}) - B(t_{n,k}))^2 \\ &\quad + \sum_{k=1}^{n-1} o((B(t_{n,k+1}) - B(t_{n,k}))^2). \end{aligned}$$

Le premier terme tend en probabilité vers l'intégrale $\int_0^t f'(B(s)) dB(s)$ par la Proposition 8.5.1. On montre que le troisième terme converge vers 0. Pour le deuxième, on montre que

$$\lim_{n \rightarrow +\infty} \mathbb{E} \left[\left| \sum_{k=1}^{n-1} f''(B(t_{n,k})) \left((B(t_{n,k+1}) - B(t_{n,k}))^2 - (t_{n,k+1} - t_{n,k}) \right) \right|^2 \right] = 0$$

en utilisant le fait que $\mathbb{E} \left[(B(t_{n,k+1}) - B(t_{n,k}))^2 \right] = t_{n,k+1} - t_{n,k}$ ainsi que les propriétés du mouvement Brownien. Comme la convergence en norme L^2 implique la convergence en probabilité, on trouve que la limite en probabilité de

$$\sum_{k=1}^{n-1} f''(B(t_{n,k})) (B(t_{n,k+1}) - B(t_{n,k}))^2$$

est donnée par

$$\lim_{n \rightarrow +\infty} \sum_{k=1}^{n-1} f''(B(t_{n,k})) (t_{n,k+1} - t_{n,k}) = \int_0^t f''(B(s)) ds$$

en utilisant les sommes de Riemann. ■

Grâce à cette formule, on retrouve l'intégrale d'Itô du mouvement Brownien calculée dans la Section 8.3 en ayant recours à la définition. En effet, en prenant $f(x) = x^2$, on retrouve

$$B^2(t) = \int_0^t 2B(s) dB(s) + t,$$

ce qui est à comparer avec la formule

$$x^2 = \int_0^x 2s ds.$$

La partie surprenante de cette formule vient de la présence de la deuxième intégrale, sans laquelle nous retrouverions le théorème fondamental du calcul intégral. En fait, cette différence est très importante car permet de donner une interprétation probabilistique des caractéristiques de $f(B(t))$. En particulier, la première intégrale est de moyenne nulle et capture l'essentiel de la variabilité de $f(B(t))$. La deuxième intégrale capture l'information de ce qu'on appelle le "drift" du processus $f(B(t))$. Une autre interprétation de la formule d'Itô est la décomposition de $f(B(t))$ en deux composantes : le signal et le bruit.

Une conséquence importante de la formule d'Itô est qu'elle permet de donner une définition plus facilement manipulable de l'intégrale d'Itô. Si $F : \mathbb{R} \rightarrow \mathbb{R}$ est une fonction de classe C^2 telle que $F' = f$ et $F(0) = 0$, alors la formule d'Itô se réécrit

$$\int_0^t f(B(s)) dB(s) = F(B(t)) - \frac{1}{2} \int_0^t f'(B(s)) ds.$$

Cela nous permet d'obtenir une définition *trajectorielle* de l'intégrale d'Itô.

La première formule d'Itô montre vite ses limites : Par exemple, on pourrait s'intéresser aux martingales $B^2(t) - t$ ou $\exp(\gamma B(t) - \gamma^2 t/2)$. Ces fonctions ne sont pas des fonctions de $B(t)$ seul, mais du couple $(t, B(t))$. La deuxième formule d'Itô permet de considérer des fonctions qui dépendent également de t .

Théorème 8.6.2 (Deuxième formule d'Itô) Soit $f : [0, +\infty[\times \mathbb{R} \rightarrow \mathbb{R}$ est une fonction de classe C^1 par rapport à sa première variable t et de classe C^2 par rapport à sa deuxième variable x . On a

$$f(t, B(t)) = f(0, 0) + \int_0^t \frac{\partial f}{\partial t}(s, B(s)) ds + \int_0^t \frac{\partial f}{\partial x}(s, B(s)) dB(s) + \frac{1}{2} \int_0^t \frac{\partial^2 f}{\partial x^2}(s, B(s)) ds.$$

Notons que si $X(t) = f(t, B(t))$, cette formule s'écrit couramment sous la forme condensée

$$dY(t) = \frac{\partial f}{\partial t}(t, B(t)) dt + \frac{\partial f}{\partial x}(t, B(t)) dB(t) + \frac{1}{2} \frac{\partial^2 f}{\partial x^2}(t, B(t)) dt.$$

Annexe A

Modes de convergence et le théorème central limite

Dans ce chapitre, nous rappelons les différentes notions de convergence de suites de variables aléatoires. Ces rappels seront utiles tout au long du cours.

Avant de commencer, nous faisons plusieurs remarques d'ordre général. Tous les vecteurs et variables aléatoires que nous considérons dans ce chapitre, et dans la suite du cours sauf mention explicite du contraire, sont supposés être définis sur un même espace probabilisé $(\Omega, \mathcal{F}, \mathbb{P})$. Par convention les vecteurs de \mathbb{R}^d , $d \geq 2$ sont des vecteurs colonnes. Le produit scalaire euclidien de $\mathbf{x} = (x_1, \dots, x_d)^T \in \mathbb{R}^d$ et $\mathbf{y} = (y_1, \dots, y_d)^T \in \mathbb{R}^d$ est $\langle \mathbf{x}, \mathbf{y} \rangle = \mathbf{x}^T \mathbf{y} = \sum_{i=1}^d x_i y_i$. La norme euclidienne s'écrit $|\mathbf{x}| = (x_1^2 + \dots + x_d^2)^{1/2}$. Le support d'une fonction $f : \mathbb{R}^d \rightarrow \mathbb{R}$ est l'adhérence de l'ensemble des points de son domaine en lesquels la fonction ne s'annule pas, i.e. $\text{supp}(f) = \overline{\{x \in \mathbb{R}^d \mid f(x) \neq 0\}}$. C'est une partie fermée de \mathbb{R}^d .

La *fonction de répartition* d'un vecteur aléatoire $\mathbf{X} = (X_1, \dots, X_d)^T \in \mathbb{R}^d$ est définie en chaque point $\mathbf{x} = (x_1, \dots, x_d)^T \in \mathbb{R}^d$ par $F_{\mathbf{X}}(\mathbf{x}) = \mathbb{P}(\mathbf{X} \leq \mathbf{x}) = \mathbb{P}(X_1 \leq x_1, \dots, X_d \leq x_d)$. Soient $\mathbf{X}, \mathbf{X}_1, \mathbf{X}_2, \dots$ des vecteurs aléatoires à valeurs dans \mathbb{R}^d .

Définition A.0.1 \mathbf{X}_n converge vers \mathbf{X} en loi (notation : $\mathbf{X}_n \xrightarrow{\mathcal{L}} \mathbf{X}$) si $F_{\mathbf{X}_n}(x) \rightarrow F_{\mathbf{X}}(x)$ lorsque $n \rightarrow \infty$ en tout point où $F_{\mathbf{X}}$ est continue.

La convergence en loi, aussi appelée convergence en distribution, est peut-être le mode de convergence le plus utilisé (notamment à cause du Théorème Central Limite, cf Théorème B.0.3). On appelle également ce mode de convergence la *convergence faible*, terminologie justifiée notamment par le fait qu'une suite de vecteurs aléatoires iid convergera nécessairement en loi vers la loi commune de la suite. La convergence en loi n'est pas, loin s'en faut, la seule façon dont une suite de vecteurs aléatoires peut converger.

Définition A.0.2 \mathbf{X}_n converge vers \mathbf{X} en probabilité (notation : $\mathbf{X}_n \xrightarrow{\mathbb{P}} \mathbf{X}$) si, pour tout $\varepsilon > 0$, $\mathbb{P}(|\mathbf{X}_n - \mathbf{X}| > \varepsilon) \rightarrow 0$ lorsque $n \rightarrow \infty$.

Définition A.0.3 \mathbf{X}_n converge vers \mathbf{X} en norme L^r , $r > 0$ (notation : $\mathbf{X}_n \xrightarrow{r} \mathbf{X}$) si $\mathbb{E}[|\mathbf{X}_n - \mathbf{X}|^r] \rightarrow 0$ lorsque $n \rightarrow \infty$. Dans le cas particulier $r = 2$ on parle de *convergence en moyenne quadratique*.

Définition A.0.4 \mathbf{X}_n converge vers \mathbf{X} presque sûrement (notation : $\mathbf{X}_n \xrightarrow{p.s.} \mathbf{X}$) si $\mathbb{P}(\mathbf{X}_n \rightarrow \mathbf{X}) = 1$, i.e. si $\mathbb{P}(\{\omega \mid \mathbf{X}_n(\omega) \rightarrow \mathbf{X}(\omega)\}) = 1$ lorsque $n \rightarrow \infty$.

Les modes de convergence présentés aux définitions A.0.1, A.0.2, A.0.3, et A.0.4 sont intimement liés les uns aux autres, comme en attestent les deux théorèmes suivants.

Théorème A.0.5

1. $\mathbf{X}_n \xrightarrow{p.s.} \mathbf{X} \Rightarrow \mathbf{X}_n \xrightarrow{\mathbb{P}} \mathbf{X}$.
2. $\mathbf{X}_n \xrightarrow{r} \mathbf{X}$ pour un $r > 0 \Rightarrow \mathbf{X}_n \xrightarrow{\mathbb{P}} \mathbf{X}$.
3. $\mathbf{X}_n \xrightarrow{\mathbb{P}} \mathbf{X} \Rightarrow \mathbf{X}_n \xrightarrow{\mathcal{L}} \mathbf{X}$.

Démonstration : Voir [4, Théorème 1] ; le point 2 est une conséquence immédiate de l'inégalité de Markov. ■

Aucun des items du Théorème A.0.5 n'est une équivalence (on peut construire des contre exemples pour chaque implication inverse). Toutefois, en imposant des hypothèses supplémentaires, on a les "inverses partiels" suivants.

Théorème A.0.6

1. Soit $c \in \mathbb{R}^d$ un point. $\mathbf{X}_n \xrightarrow{\mathcal{L}} c \Rightarrow \mathbf{X}_n \xrightarrow{\mathbb{P}} c$.
2. Si $\mathbf{X}_n \xrightarrow{p.s.} \mathbf{X}$ et si il existe $r > 0$ tel que $|\mathbf{X}_n|^r \leq Z$ avec $\mathbb{E}[Z] < \infty$ alors $\mathbf{X}_n \xrightarrow{r} \mathbf{X}$.
3. Si $\mathbf{X}_n \xrightarrow{p.s.} \mathbf{X}$, $\mathbf{X}_n \geq 0$ et si $\mathbb{E}[\mathbf{X}_n] \rightarrow \mathbb{E}[\mathbf{X}] < \infty$, alors $\mathbf{X}_n \xrightarrow{r} \mathbf{X}$ pour $r = 1$.
4. $\mathbf{X}_n \xrightarrow{\mathbb{P}} \mathbf{X}$ si et seulement si toute sous-suite n_1, n_2, \dots admet une sous-sous-suite m_1, m_2, \dots telle que $\mathbf{X}_{m_j} \xrightarrow{p.s.} \mathbf{X}$ lorsque $j \rightarrow \infty$.
5. Si $\mathbf{X}_n \xrightarrow{\mathcal{L}} \mathbf{X}$ alors il existe $((\mathbf{Y}_n)_{n \geq 1}, \mathbf{Y})$ tels que $\mathbf{X}_n \stackrel{\mathcal{L}}{=} \mathbf{Y}_n$, $\mathbf{X} \stackrel{\mathcal{L}}{=} \mathbf{Y}$ et $\mathbf{Y}_n \xrightarrow{p.s.} \mathbf{Y}$.

Démonstration : Les quatre premiers points sont prouvés dans [4, Théorème 2] ; le dernier provient de [1, page 70]. ■

La définition A.0.1 n'est pas nécessairement la forme la plus maniable de convergence en loi. Nous citons le théorème suivant, dont les différents items pourront également être utilisés en définition.

Théorème A.0.7 (Helly-Bray) Les assertions suivantes sont équivalentes lorsque $n \rightarrow \infty$.

1. $\mathbf{X}_n \xrightarrow{\mathcal{L}} \mathbf{X}$.
2. $\mathbb{E}[\varphi(\mathbf{X}_n)] \rightarrow \mathbb{E}[\varphi(\mathbf{X})]$ pour toute fonction réelle φ à support compact.
3. $\mathbb{E}[\varphi(\mathbf{X}_n)] \rightarrow \mathbb{E}[\varphi(\mathbf{X})]$ pour toute fonction réelle φ continue bornée.
4. $\mathbb{E}[\varphi(\mathbf{X}_n)] \rightarrow \mathbb{E}[\varphi(\mathbf{X})]$ pour toute fonction réelle φ mesurable bornée telle que $\mathbb{P}(\mathbf{X} \in C(\varphi)) = 1$, où $C(\varphi) = \{x : \varphi \text{ est continue en } x\}$ (i.e. pour toute fonction mesurable bornée continue $\mathbb{P}^{\mathbf{X}}$ -p.s.).

Démonstration : Voir [4, Théorème 3]. ■

Théorème A.0.8 (Théorème du portemanteau) Les assertions suivantes sont équivalentes lorsque $n \rightarrow \infty$.

1. $\mathbf{X}_n \xrightarrow{\mathcal{L}} \mathbf{X}$.
2. $\mathbb{E}[\varphi(\mathbf{X}_n)] \rightarrow \mathbb{E}[\varphi(\mathbf{X})]$ pour toute fonction réelle φ continue bornée.
3. $\mathbb{E}[\varphi(\mathbf{X}_n)] \rightarrow \mathbb{E}[\varphi(\mathbf{X})]$ pour toute fonction réelle φ uniformément continue et bornée.
4. pour tout fermé F , $\limsup_{n \rightarrow \infty} \mathbb{P}(\mathbf{X}_n \in F) \leq \mathbb{P}(\mathbf{X} \in F)$.
5. pour tout ouvert G , $\limsup_{n \rightarrow \infty} \mathbb{P}(\mathbf{X}_n \in G) \geq \mathbb{P}(\mathbf{X} \in G)$.
6. pour tout $A \in \mathcal{B}(\mathbb{R}^d)$ tel que $\mathbb{P}(\mathbf{X} \in \bar{A} \setminus \overset{\circ}{A}) = 0$, $\mathbb{P}(\mathbf{X}_n \in A) \rightarrow \mathbb{P}(\mathbf{X} \in A)$.

Démonstration : Voir [2, Théorème 1.3] pour une preuve dans un cadre plus général (vecteurs aléatoires à valeurs dans un espace métrique). ■

On déduit du théorème du porte-manteau :

Théorème A.0.9 (Continuous mapping theorem) Soit $f : \mathbb{R}^d \rightarrow \mathbb{R}^q$. Alors si $\mathbf{X}_n \xrightarrow{\mathcal{L}} \mathbf{X}$ et f est continue $\mathbb{P}^{\mathbf{X}}$ -p.s., on a aussi $f(\mathbf{X}_n) \xrightarrow{\mathcal{L}} f(\mathbf{X})$.

La convergence en loi de vecteurs aléatoires se ramène à un problème de convergence univariée grâce au théorème suivant.

Théorème A.0.10 (Cramér-Wold device) $\mathbf{X}_n \xrightarrow{\mathcal{L}} \mathbf{X}$ si et seulement si $\langle \mathbf{t}, \mathbf{X}_n \rangle \xrightarrow{\mathcal{L}} \langle \mathbf{t}, \mathbf{X} \rangle$ pour tout $\mathbf{t} \in \mathbb{R}^d$.

Démonstration : Voir [3, Page 168]. ■

Finalement nous mentionnons le

Théorème A.0.11 (Théorème de Paul Lévy) $\mathbf{X}_n \xrightarrow{\mathcal{L}} \mathbf{X}$ si et seulement si $\varphi_{\mathbf{X}_n}(t) \rightarrow \varphi_{\mathbf{X}}(t)$ pour tout $t \in \mathbb{R}^d$, où $\varphi_{\mathbf{X}}(t) = \mathbb{E}[e^{i\langle \mathbf{t}, \mathbf{X} \rangle}]$ est la fonction caractéristique de \mathbf{X} .

Démonstration : Voir [4, Théorème 3(e)]. ■

Remarquons que, dans le monde anglo-saxon, on parle plutôt de *Théorème de continuité*.

Annexe B

Vecteurs gaussiens

Une variable aléatoire réelle N est *gaussienne standard* (a.k.a. normale centrée réduite), et on écrit $N \sim \mathcal{N}(0, 1)$, si sa mesure engendrée est donnée par

$$\mathbb{P}^N(A) = \mathbb{P}(N \in A) = \frac{1}{\sqrt{2\pi}} \int_A e^{-x^2/2} dx \quad (\text{B.1})$$

pour tout borélien A . Une variable aléatoire X de moyenne $\mathbb{E}[X] = \mu \in \mathbb{R}$ et de variance $\text{Var}(X) = \sigma^2 \in [0, \infty)$ est de loi *gaussienne* (a.k.a. normale) si $X \stackrel{\mathcal{L}}{=} \sigma N + \mu$ (égalité en loi) pour $N \sim \mathcal{N}(0, 1)$; on écrit $X \sim \mathcal{N}(\mu, \sigma^2)$. La loi d'une variable aléatoire gaussienne est entièrement déterminée par sa moyenne et sa variance.

Remarque B.0.1 Une variable aléatoire gaussienne peut avoir une variance nulle; dans ce cas il s'agit d'une variable aléatoire dégénérée égale à sa moyenne, sa loi est une mesure de Dirac en μ : $\mathbb{P}^X = \delta_\mu$.

Si X est gaussien sa fonction génératrice des moments (FGM, MGF en anglais) et sa fonction caractéristique (FC, CF en anglais) sont

$$M_X(t) = \mathbb{E}[e^{tX}] = e^{\mu t + \frac{\sigma^2 t^2}{2}} \text{ et } \varphi_X(t) = \mathbb{E}[e^{itX}] = e^{i\mu t - \sigma^2 t^2/2} \quad (\text{B.2})$$

(toujours avec $\mu = \mathbb{E}[X]$ et $\sigma^2 = \text{Var}(X)$). Les moments de tous ordres d'une variable aléatoire gaussienne existent et s'expriment uniquement en termes de μ et σ^2 . La suite $(m_n)_{n \geq 0}$ des moments d'une variable aléatoire gaussienne centrée réduite satisfait la relation de récurrence

$$m_n = nm_{n-2}. \quad (\text{B.3})$$

Si $X \sim \mathcal{N}(0, \sigma^2)$ alors

$$\mathbb{E}[X^{2n}] = \frac{(2n)!}{2^n n!} \sigma^{2n} \text{ et } \mathbb{E}[X^{2n+1}] = 0 \quad (\text{B.4})$$

pour tout $n \in \mathbb{N}$. On peut montrer (cf [7, Lemme 3.1.1]) que la loi gaussienne est caractérisée par ses moments.

La loi gaussienne est *stable par addition*, i.e. si X_1, \dots, X_n sont indépendants gaussiens alors $\sum_{i=1}^n \alpha_i X_i$ est gaussien également, de moyenne $\sum_{i=1}^n \alpha_i \mu_i$ et de variance $\sum_{i=1}^n \alpha_i^2 \sigma_i^2$. La loi gaussienne est également *infinitement divisible*, i.e. pour tout $n \in \mathbb{N}$ il existe X_1, \dots, X_n iid tels que

$$X \stackrel{\mathcal{L}}{=} X_1 + \dots + X_n.$$

En effet, si $X \sim \mathcal{N}(\mu, \sigma^2)$ il suffit de prendre $X_i \sim \mathcal{N}(\mu/n, \sigma^2/n)$.

Proposition B.0.2 Soit $(X_n)_{n \geq 1}$ une suite de variables aléatoires $\mathcal{N}(\mu_n, \sigma_n^2)$.

1. La suite converge en loi si et seulement si $\mu_n \rightarrow \mu$ et $\sigma_n^2 \rightarrow \sigma^2 \in [0, \infty)$. La loi limite est alors $\mathcal{N}(\mu, \sigma^2)$.
2. $X_n \xrightarrow{\mathbb{P}} X$ alors $X_n \xrightarrow{r} X$ pour tout $r < \infty$.

Démonstration : Le point 1 sera démontré aux travaux pratiques (Liste 1, exercice 5). Pour le second nous renvoyons à [2, Proposition 0.5]. ■

Théorème B.0.3 (Théorème Central Limite) Soit $(X_n)_{n \geq 1}$ une suite de variables aléatoires iid, d'espérance μ et de variance finie $\sigma^2 > 0$. Soit $S_n = X_1 + \dots + X_n$ leur somme partielle et N une gaussienne standard. Alors

$$\frac{S_n - \mathbb{E}[S_n]}{\sqrt{\text{Var}(S_n)}} \xrightarrow{\mathcal{L}} N$$

lorsque $n \rightarrow \infty$. On écrira également également $S_n - \mathbb{E}[S_n] / \sqrt{\text{Var}(S_n)} \rightarrow \mathcal{N}(0, 1)$ ou encore $S_n \approx \mathcal{N}(\mathbb{E}[S_n], \text{Var}(S_n))$ ou encore $\bar{X}^{(n)} = \frac{1}{n} S_n \approx \mathcal{N}(\mu, \sigma^2/n)$.

Démonstration : Voir [9]. ■

L'extension de la loi gaussienne univariée au cas multivarié se fait typiquement de la façon suivante.

Définition B.0.4 Un vecteur aléatoire $\mathbf{X} = (X_1, \dots, X_d)^T \in \mathbb{R}^d$ est *gaussien* si toute combinaison linéaire de ses composantes est gaussienne.

Remarque B.0.5 Une autre définition plus transparente consiste à dire que \mathbf{X} est multivarié gaussien si il existe une matrice \mathbf{D} et un vecteur \mathbf{m} tels que $\mathbf{X} = \mathbf{D}\mathbf{N} + \mathbf{m}$ avec \mathbf{N} un vecteur dont les composantes sont $\mathcal{N}(0, 1)$ indépendantes. On peut montrer que les deux approches sont équivalentes.

Définition B.0.6 Soit $\mathbf{X} = (X_1, \dots, X_d)^T$ un vecteur gaussien. Son *vecteur moyen* (a.k.a. sa moyenne) est

$$\mathbf{m} = (\mathbb{E}[X_1], \dots, \mathbb{E}[X_d])^T$$

et sa *matrice de covariance* (a.k.a. sa variance) est

$$\begin{aligned} \mathbf{K} &= \text{Var}(\mathbf{X}) = \text{Cov}(\mathbf{X}) \\ &= \mathbb{E}[(\mathbf{X} - \mathbb{E}[\mathbf{X}])(\mathbf{X} - \mathbb{E}[\mathbf{X}])^T] = \left(\text{Cov}(X_i, X_j) \right)_{i,j \in \{1, \dots, d\}} \end{aligned}$$

Dans ce cas, on note $\mathbf{X} \sim \mathcal{N}(\mathbf{m}, \mathbf{K})$ (ou $\mathcal{N}_d(\mathbf{m}, \mathbf{K})$ s'il y a pas ambiguïté sur les dimensions).

Proposition B.0.7 La loi d'un vecteur gaussien \mathbf{X} est entièrement déterminée par sa moyenne et sa variance.

Démonstration : Ce point sera démontré aux travaux pratiques (Liste 1, exercice 1). ■

Puisque les marginales de \mathbf{X} suivent une loi gaussienne, les covariances $\text{Cov}(X_i, X_j)$ sont bien définies. Rappelons également que si le vecteur moyen est nul, on dit que \mathbf{X} est *centré*. Enfin, remarquons qu'une matrice de covariance est toujours symétrique – par construction – et semi-définie positive : en effet, pour tout $\mathbf{x} \in \mathbb{R}^d$, on a

$$\begin{aligned} \mathbf{x}^T \mathbf{K} \mathbf{x} &= \mathbf{x}^T \mathbb{E} [(\mathbf{X} - \mathbb{E}[\mathbf{X}])(\mathbf{X} - \mathbb{E}[\mathbf{X}])^T] \mathbf{x} = \mathbb{E} [\mathbf{x}^T (\mathbf{X} - \mathbb{E}[\mathbf{X}])(\mathbf{X} - \mathbb{E}[\mathbf{X}])^T \mathbf{x}] \\ &= \mathbb{E} \left[\left((\mathbf{X} - \mathbb{E}[\mathbf{X}])^T \mathbf{x} \right)^2 \right] \geq 0. \end{aligned}$$

Remarque B.0.8 1. Grâce au Théorème A.0.10, la proposition B.0.2 ainsi que le TCL (théorème B.0.3) s'étendent presque verbatim au cas de vecteurs aléatoires.

2. Si X_1, \dots, X_d sont des gaussiennes indépendantes, alors $(X_1, \dots, X_d)^T$ est gaussien multivarié.
3. La définition B.0.4 requiert beaucoup plus que la seule normalité des marginales. En effet, si $N \sim \mathcal{N}(0, 1)$ et

$$N' = \begin{cases} N & \text{si } |N| \leq 1; \\ -N & \text{si } |N| > 1. \end{cases}$$

Alors N' est également gaussien standard. Mais étant donné que $|N + N'| \leq 2$ et $N + N'$ n'est pas constante, $N + N'$ n'est pas gaussien et donc (N, N') n'est pas gaussien bivarié.

4. Un vecteur aléatoire gaussien de \mathbb{R}^d n'a pas nécessairement une densité par rapport à la mesure de Lebesgue λ sur \mathbb{R}^d ; par exemple (X, X) est gaussien bivarié si X est gaussien, mais $\mathbb{P}((X, X) \in \text{première bisectrice}) = 1$ alors que $\lambda(\text{première bisectrice}) = 0$.
5. Si \mathbf{X} est gaussien alors $\mathbf{A}\mathbf{X}$ est également gaussien pour toute matrice \mathbf{A} aux dimensions appropriées.

Proposition B.0.9 Soit $\mathbf{X} = (X_1, \dots, X_d)^T$ un vecteur aléatoire gaussien. Alors les composantes X_1, \dots, X_d de \mathbf{X} sont indépendantes si et seulement si elles sont non corrélées, i.e. pour tout $i, j \in \{1, \dots, p\}$ tels que $i \neq j$, X_i et X_j sont indépendantes si et seulement si $\text{Cov}(X_i, X_j) = 0$.

Démonstration : Voir travaux pratiques, Liste 1. ■

Bien entendu, en général une corrélation nulle n'implique en aucune manière l'indépendance. Même dans le cas de vecteurs à marginales gaussiennes, il faut bien veiller à considérer le comportement conjoint des composantes !

Exemple B.0.10 Considérons $N \sim \mathcal{N}(0, 1)$ et ε une variable aléatoire de loi de Rademacher (i.e. $\varepsilon \sim \text{Uni}(\{-1, 1\})$). Alors $X_1 = N$ et $X_2 = \varepsilon N$ sont deux variables aléatoires gaussiennes standard, de covariance nulle mais pas indépendantes. Le couple (X_1, X_2) n'est toutefois pas multivarié gaussien.

Définition B.0.11 Un vecteur gaussien \mathbf{X} est *non-dégénéré* si sa matrice de covariance est inversible.

Proposition B.0.12 $\det(\mathbf{K}) = 0$ si et seulement si au moins l'une des variables aléatoires X_1, \dots, X_d s'exprime comme une combinaison linéaire des autres variables.

Proposition B.0.13 Soit $\mathbf{X} \sim \mathcal{N}(\mathbf{m}, \mathbf{K})$. Le vecteur \mathbf{X} admet une densité (par rapport à la mesure de Lebesgue) si et seulement si \mathbf{X} est non-dégénérée. Dans ce cas, on a

$$f_{\mathbf{X}}(\mathbf{x}) = \frac{1}{(2\pi)^{d/2} \sqrt{\det(\mathbf{K})}} \exp\left(-\frac{1}{2}(\mathbf{x} - \mathbf{m})^T \mathbf{K}^{-1}(\mathbf{x} - \mathbf{m})\right)$$

pour tout $\mathbf{x} \in \mathbb{R}^d$.

Démonstration : Nous prouverons la condition nécessaire aux travaux pratiques, Liste 1. ■

Annexe C

Normes L^p et inégalités de Hölder

Définition C.0.1 Soit $p \in [1, +\infty[$ et soit (X, \mathcal{A}, μ) un espace mesuré. L'espace $\mathcal{L}^p(X, \mathcal{A}, \mu, \mathbb{R})$ est l'ensemble des applications mesurables $f : X \rightarrow \mathbb{R}$ telles que $|f|^p$ est intégrable.

Définition C.0.2 Soit (X, \mathcal{A}, μ) un espace mesuré. Si $p \in [1, +\infty[$, on pose pour tout $f \in \mathcal{L}^p(X, \mathcal{A}, \mu)$

$$\|f\|_p = \left(\int |f|^p d\mu \right)^{\frac{1}{p}}.$$

Nous allons à présent montrer que les applications $\|\cdot\|_p$ définissent des semi-normes. Commençons par introduire l'exposant conjugué de p .

Définition C.0.3 Soit $p \in]1, +\infty[$. Alors $\frac{1}{p} \in]0, 1[$ et il existe un nombre réel q tel que

$$\frac{1}{p} + \frac{1}{q} = 1.$$

Les nombres p et q sont appelés des *exposants conjugués*.

Remarquons que si p et q sont des exposants conjugués, alors

$$p + q = pq$$

et si p et q sont finis,

$$p = q(p - 1) \quad \text{et} \quad q = p(q - 1).$$

Lemme C.0.4 Soit $p \in]1, +\infty[$ et q son exposant conjugué. Alors

$$xy \leq \frac{x^p}{p} + \frac{y^q}{q}$$

pour tous $x, y \geq 0$.

Démonstration : Bien sûr, on peut supposer que x et y sont non-nuls. En posant $u = x^p$ et $v = y^q$, il suffit de prouver que

$$u^{1/p}v^{1/q} \leq \frac{u}{p} + \frac{v}{q}$$

pour tous $u, v > 0$. En posant $t = \frac{u}{v}$, il suffit de montrer que

$$t^{1/p} \leq \frac{t}{p} + \frac{1}{q}$$

pour tout $t > 0$. Il suffit alors de remarquer que la fonction

$$t > 0 \mapsto \frac{t}{p} + \frac{1}{q} - t^{1/p}$$

atteint son minimum en $t = 1$, et que ce minimum vaut 0. ■

Proposition C.0.5 (Inégalité de Hölder) Soient (X, \mathcal{A}, μ) un espace mesuré, $p \in]1, +\infty[$ et q son exposant conjugué. Si $f \in \mathcal{L}^p(X, \mathcal{A}, \mu)$ et $g \in \mathcal{L}^q(X, \mathcal{A}, \mu)$, alors $fg \in \mathcal{L}^1(X, \mathcal{A}, \mu)$ et

$$\int |fg| d\mu \leq \|f\|_p \|g\|_q.$$

Démonstration : Supposons à présent que $p, q \in]1, +\infty[$. Si $\|f\|_p = 0$ ou $\|g\|_q = 0$, alors $fg = 0$ presque partout et le résultat est démontré. Sinon, on pose $\tilde{f} = \frac{f}{\|f\|_p}$ et $\tilde{g} = \frac{g}{\|g\|_q}$. Par le Lemme C.0.4, on a

$$|\tilde{f}(x)\tilde{g}(x)| \leq \frac{|\tilde{f}(x)|^p}{p} + \frac{|\tilde{g}(x)|^q}{q}$$

pour tout $x \in X$. Par conséquent, $\tilde{f}\tilde{g} \in \mathcal{L}^1(X, \mathcal{A}, \mu)$ et

$$\int |\tilde{f}\tilde{g}| d\mu \leq \frac{1}{p} \int |\tilde{f}|^p d\mu + \frac{1}{q} \int |\tilde{g}|^q d\mu = \frac{1}{p} + \frac{1}{q} = 1.$$

On en tire que $fg \in \mathcal{L}(X, \mathcal{A}, \mu)$ et que

$$\frac{1}{\|f\|_p \|g\|_q} \int |fg| d\mu \leq 1.$$

Proposition C.0.6 (Inégalité de Minkowski) Soient (X, \mathcal{A}, μ) un espace mesuré, $p \in]1, +\infty[$. Si $f, g \in \mathcal{L}^p(X, \mathcal{A}, \mu)$, alors

$$\|f + g\|_p \leq \|f\|_p + \|g\|_p.$$

Démonstration : Si $\|f + g\|_p = 0$, le résultat est clair. Sinon, soit q l'exposant conjugué de p . Alors $p + q = pq$ et par conséquent,

$$|f + g|^p = (|f + g|^{p-1})^q.$$

On en tire que $|f + g|^{p-1} \in \mathcal{L}^q(X, \mathcal{A}, \mu)$. En utilisant l'inégalité de Hölder, on obtient

$$\begin{aligned}
 \|f + g\|_p^p &= \int |f + g|^p d\mu \leq \int |f + g|^{p-1} (|f| + |g|) d\mu \\
 &= \int |f + g|^{p-1} |f| d\mu + \int |f + g|^{p-1} |g| d\mu \\
 &\leq \| |f + g|^{p-1} \|_q \|f\|_p + \| |f + g|^{p-1} \|_q \|g\|_p \\
 &= (\|f\|_p + \|g\|_p) \| |f + g|^{p-1} \|_q \\
 &= (\|f\|_p + \|g\|_p) \left(\int |f + g|^p d\mu \right)^{1-1/p}
 \end{aligned}$$

puisque $q(p-1) = p$ et $1/q = 1 - 1/p$. En divisant par $(\int |f + g|^p d\mu)^{1-1/p}$, cette inégalité se réécrit

$$\|f + g\|_p \leq \|f\|_p + \|g\|_p.$$

■

Par conséquent, il est facile de vérifier que $\|\cdot\|_p$ définit bien une semi-norme sur l'espace $\mathcal{L}^p(X, \mathcal{A}, \mu)$.

Bibliographie

- [1] Billingsley, Patrick (1999) *Convergence of probability measures*. New York ; John Wiley & Sons.
- [2] Breton, Jean-Christophe (2018) *Processus stochastiques - M2 Mathématiques*. Université de Rennes 1.
- [3] Durrett, R. (2005) *Probability : Theory and Examples*. 3rd edition, Duxbury.
- [4] Ferguson, Thomas S. (2017). *A course in large sample theory*. Routledge.
- [5] Hörmann, Siegfried. (2017) Probabilités II, Slides Université Libre de Bruxelles.
- [6] Liggett, Thomas M. (2010) *Continuous time Markov processes*. Vol. 113. American Mathematical Society.
- [7] Nourdin, Ivan, and Giovanni Peccati. (2012) *Normal approximations with Malliavin calculus : from Stein's method to universality*. Vol. 192. Cambridge University Press.
- [8] Steele, J. Michael *Stochastic calculus and Financial Applications* Vol 45, Stochastic Modeling and Applied Probability, Springer.
- [9] Swan, Yvik (2018) *Un premier cours de probabilité* Université de Liège.
- [10] Williams, David (1991) *Probability with Martingales* Cambridge University Press.