Calcul stochastique et modèles de diffusion

Cyril Labbé[1]

Année universitaire 2024/2025
M2MO : Modélisation Aléatoire, Finance et Data Science
Université Paris Cité

1. Université Paris Cité - LPSM - clabbe@lpsm.paris

Table des matières

La théorie des processus stochastiques à temps continu contient de nombreuses subtilités et certains arguments sont très techniques. Dans ces notes de cours, on utilise le symbole ♠ pour indiquer qu'un passage (preuve, remarque, exercice) est "plus avancé" et ne sera pas présenté en détail en cours.

Chapitre I

Mouvement Brownien

Approx. 2 séances de cours

1 Processus à temps continu

Dans cette section, nous introduisons des concepts généraux sur les processus à temps continu. Certains passages sont assez abstraits, et l'on conseille de ne pas s'y attarder en première lecture.

Définition I.1

Soit $(E, \mathcal{E})$ un ensemble mesurable, $\mathbb{T}$ un ensemble quelconque et $(\Omega, \mathcal{F}, \mathbb{P})$ un espace de probabilité. On dit que $(X_t, t \in \mathbb{T})$ est un processus stochastique sur $(\Omega, \mathcal{F}, \mathbb{P})$ à valeurs dans $(E, \mathcal{E})$ si pour tout $t \in \mathbb{T}$, X_t est une variable aléatoire de $(\Omega, \mathcal{F})$ dans $(E, \mathcal{E})$.

Lorsque $\mathbb{T} = \mathbb{N}$, on parle de processus à temps discret. Lorsque $\mathbb{T}$ est égal à $\mathbb{R}_+$ ou à un intervalle quelconque de $\mathbb{R}$, on parle de processus à temps continu. Lorsque $(E, \mathcal{E}) = (\mathbb{R}, \mathcal{B}(\mathbb{R}))$ on parle de processus à valeurs réelles.
L'essentiel de ce cours porte sur les processus stochastiques à temps continu et à valeurs réelles. De façon générique, on supposera dans la suite que $\mathbb{T} = \mathbb{R}_+$, et que $E = \mathbb{R}$ muni de la tribu Borélienne $\mathcal{E} = \mathcal{B}(\mathbb{R})$ (mais toutes les définitions présentées font encore sens si $\mathbb{T}$ est un intervalle borné, etc.).

Etant donné un processus stochastique $(X_t, t \in \mathbb{R}_+)$ et un n-uplet $t_1, \ldots, t_n \in \mathbb{R}_+$, le vecteur $(X_{t_1}, \ldots, X_{t_n})$ forme une variable aléatoire à valeurs dans $\mathbb{R}^n$. Notons sa loi $\mu_{(t_1,\ldots,t_n)}$: il s'agit d'une mesure de probabilité sur $\mathbb{R}^n$.
L'ensemble des variables aléatoires $(X_{t_1}, \ldots, X_{t_n})$ indicé par n et $t_1, \ldots, t_n$ est l'ensemble des *marginales finies-dimensionnelles* du processus stochastique. Les lois associées vérifient la condition de compatibilité suivante : pour tout $n \geq 1$, tout $t_1, \ldots, t_n \in \mathbb{R}_+$ et tout $t \notin \{t_1, \ldots, t_n\}$

$$\mu_{(t_1,\ldots,t_n,t)}(A \times \mathbb{R}) = \mu_{(t_1,\ldots,t_n)}(A) \ , \quad \forall A \in \mathcal{B}(\mathbb{R}^n) \ .$$

Cette identité peut s'exprimer de façon très intuitive : la restriction de la loi de $(X_{t_1}, \ldots, X_{t_n}, X_t)$

à ses n premières coordonnées est la loi de $(X_{t_1}, \dots, X_{t_n})$.

Les marginales finies-dimensionnelles donnent accès à des informations partielles sur le processus. Elles peuvent être vues comme des projections d'un objet plus gros : l'application

$$\begin{aligned} X : \Omega &\to \mathbb{R}^{\mathbb{R}_+} \\ \omega &\mapsto (X_t(\omega), t \in \mathbb{R}_+) \end{aligned}$$

qui prend ses valeurs dans un espace de *trajectoires*.

Nous souhaiterions voir X comme une variable aléatoire. Pour cela, il nous faut répondre à la question suivante : de quelle tribu peut-on munir l'espace de trajectoires ?

Lemme I.2

On définit $\mathcal{B}(\mathbb{R})^{\otimes \mathbb{R}_+}$ comme étant la plus petite tribu sur l'espace des trajectoires $\mathbb{R}^{\mathbb{R}_+}$ rendant mesurables les applications coordonnées $\pi_t : x \mapsto x_t$ pour tout $t \in \mathbb{R}_+$. Cette tribu coïncide avec la plus petite tribu contenant la collection $\mathcal{C}$ de tous les cylindres de dimension finie, c'est-à-dire, de tous les sous-ensembles de $\mathbb{R}^{\mathbb{R}_+}$ de la forme

$$\prod_{t \in \mathbb{R}_+} A_t \ , \quad \text{avec } A_t \in \mathcal{B}(\mathbb{R}_+) \ , \text{ et } \quad \#\{t \in \mathbb{R}_+ : A_t \neq \mathbb{R}\} < \infty \ .$$

Si $(X_t, t \in \mathbb{R}_+)$ est un processus stochastique, alors l'application X définie ci-dessus est mesurable de $(\Omega, \mathcal{F})$ dans $(\mathbb{R}^{\mathbb{R}_+}, \mathcal{B}(\mathbb{R})^{\otimes \mathbb{R}_+})$.

Démonstration [♠] On souhaite montrer que $\mathcal{B}(\mathbb{R})^{\otimes \mathbb{R}_+} = \sigma(\mathcal{C})$. Soit A un cylindre de dimension finie. Il existe n, $t_1 < \ldots < t_n$ et $A_t \in \mathcal{B}(\mathbb{R})$ tels que

$$A = \prod_{t \in \mathbb{R}_+} A_t \ , \quad A_t = \mathbb{R} \forall t \notin \{t_1, \dots, t_N\} \ .$$

Ainsi $A = \cap_{i=1}^n \pi_{t_i}^{-1}(A_{t_i})$ et ainsi $A \in \mathcal{B}(\mathbb{R})^{\otimes \mathbb{R}_+}$. Réciproquement la tribu $\mathcal{B}(\mathbb{R})^{\otimes \mathbb{R}_+}$ rend mesurable chaque application π_t. Il est n'est pas difficile de vérifier qu'elle rend alors mesurable les applications

$$\pi_{t_1, \dots, t_n} : x \mapsto (x_{t_1}, \dots, x_{t_n})$$

et ainsi contient la tribu engendrée par les cylindres de dimension finie.
Si l'on note $\mathcal{C}$ la classe des cylindres de dimension finie, alors $X^{-1}(\mathcal{C}) \subset \mathcal{F}$ par définition d'un processus stochastique ; or $\sigma(\mathcal{C})$ n'est rien d'autre que la tribu produit et une propriété générale de théorie de la mesure assure que $X^{-1}(\sigma(\mathcal{C})) = \sigma(X^{-1}(\mathcal{C}))$ et ainsi X est mesurable. □

Ainsi un processus stochastique $(X_t, t \in \mathbb{R}_+)$ peut être vu comme *une variable aléatoire* à valeurs dans un espace de trajectoires. La loi μ d'un tel processus stochastique est la mesure image de $\mathbb{P}$ par l'application X. Le résultat suivant est important.

Lemme I.3

La loi μ d'un processus stochastique est entièrement caractérisée par l'ensemble des lois finies-dimensionnelles $\mu_{(t_1,\ldots,t_n)}$ pour $n \geq 1$ et $t_1, \ldots, t_n \in \mathbb{R}_+$. En d'autres termes, si Y est un autre processus stochastique et si pour tout n et tous $t_1, \ldots, t_n \in \mathbb{R}_+$ les vecteurs $(X_{t_1}, \ldots, X_{t_n})$ et $(Y_{t_1}, \ldots, Y_{t_n})$ ont même loi, alors X et Y ont même loi.

Démonstration [♠] Comme la classe des cylindres est stable par intersection finie, un corollaire du théorème de classe monotone assure que la loi d'un processus stochastique est entièrement caractérisée par la donnée de $\mu(A)$ pour tout $A \in \mathcal{C}$. Or si l'on note $t_1, \ldots, t_n$ les indices pour lesquels $A_t \neq \mathbb{R}$ on observe que

$$\mu(A) = \mathbb{P}((X_{t_1}, \ldots, X_{t_n}) \in A_{t_1} \times \cdots \times A_{t_n}) = \mu_{(t_1,\ldots,t_n)}(A_{t_1} \times \ldots \times A_{t_n}) \ .$$

□

Remarque I.4 (♠)

Si $(X_t, t \in \mathbb{R}_+)$ est un processus stochastique, rien n'assure que les ensembles

$$\{\omega : \forall t \in [0,1] : X_t(\omega) \leq 10\} \ , \quad \{\omega : t \mapsto X_t(\omega) \text{ est continu}\} \ ,$$

soient mesurables en général. En effet, ces ensembles dépendent d'une collection non-dénombrables *d'indices* $t \in \mathbb{R}_+$, *et ne peuvent pas s'écrire a priori comme union/intersection dénombrables d'événements ne faisant intervenir que les marginales finies-dimensionnelles. Autrement dit, les ensembles*

$$\{\forall t \in [0,1] : x_t \leq 10\} \ , \quad \{t \mapsto x_t \text{ est continu}\} \ ,$$

n'appartiennent pas à la tribu produit et rien ne dit que X reste mesurable si l'on munit $\mathbb{R}^{\mathbb{R}_+}$ d'une tribu contenant ces ensembles.
Nous reviendrons sur la continuité des trajectoires d'un processus stochastique un peu plus loin dans ce chapitre.

Examinons le cas particulier des processus gaussiens.

Définition I.5

On dit qu'un processus stochastique $(X_t, t \in \mathbb{R}_+)$ est gaussien si toutes ses marginales finies-dimensionnelles forment des vecteurs gaussiens.

Autrement dit, un processus stochastique $(X_t, t \in \mathbb{R}_+)$ est gaussien si pour tous $t_1, \ldots, t_n \in \mathbb{R}_+$ et tous $\lambda_1, \cdots, \lambda_n \in \mathbb{R}$, la variable aléatoire $\sum_{i=1}^n \lambda_i X_{t_i}$ est gaussienne.

Proposition I.6

La loi d'un processus gaussien $(X_t, t \in \mathbb{R}_+)$ est complétement caractérisée par la donnée de sa fonction moyenne et de sa fonction de covariance :

$$t \mapsto \mathbb{E}[X_t] \ , \quad (s,t) \mapsto \mathrm{Cov}(X_s, X_t) \ .$$

Démonstration Il est bien connu que la loi d'un vecteur gaussien $(X_{t_1}, \dots, X_{t_n})$ est caractérisée par son vecteur moyenne et sa matrice de covariance. Or ces deux quantités sont complétement décrites par les fonctions introduites dans l'énoncé. Pour conclure, on utilise le fait prouvé précédemment que la loi d'un processus stochastique est complétement caractérisée par les lois de ses marginales finies-dimensionnelles. □

Attention : il n'est pas vrai en général que la loi d'un processus stochastique est caractérisée par sa fonction moyenne et sa fonction de covariance ; le caractère gaussien joue un rôle crucial ici.

On notera que si X est un processus gaussien de moyenne m et de fonction de covariance Γ, alors $(X_t - m_t, t \in \mathbb{R}_+)$ est un processus gaussien centré de même fonction de covariance.

Une question naturelle est de se demander s'il existe des processus gaussiens. Plus précisément, on se donne une fonction $t \mapsto m_t$ de $\mathbb{R}_+$ dans $\mathbb{R}$ ainsi qu'une fonction $\Gamma : \mathbb{R}_+ \times \mathbb{R}_+ \to \mathbb{R}$ vérifiant :

$$\Gamma(s,t) = \Gamma(t,s) \ , \quad \sum_{i,j=1}^{n} \lambda_i \lambda_j \Gamma(t_i, t_j) \geq 0 \ , \quad \forall n \geq 1 \ , \lambda_i \in \mathbb{R} \ , t_i \in \mathbb{R}_+ \ .$$

Existe-t-il un processus gaussien de fonction de moyenne m et de fonction de covariance Γ ?

Pour répondre à cette question, nous allons faire appel à un résultat général de la théorie des processus stochastique (qui ne s'applique pas seulement aux processus gaussiens). Sa preuve repose sur des arguments délicats de construction de mesures et est omise.

Théorème I.7 (*Théorème d'extension de Kolmogorov*)

Pour tout entier $n \geq 1$ et toute collection finie $t_1, \dots, t_n$ de $\mathbb{R}_+$, soit $\mu_{(t_1,\dots,t_n)}$ une mesure de probabilité sur $\mathbb{R}^n$. On suppose que ces mesures sont compatibles au sens suivant : pour tout $n \geq 1$, tout $(t_1, \dots, t_n) \in \mathbb{R}_+^n$ et tout $t \in \mathbb{R}_+ \backslash \{t_1, \dots, t_n\}$

$$\mu_{(t_1,\dots,t_n,t)}(A \times \mathbb{R}) = \mu_{(t_1,\dots,t_n)}(A) \ , \quad \forall A \in \mathcal{B}(\mathbb{R}^n) \ .$$

Alors il existe un processus stochastique $(X_t, t \in \mathbb{R}_+)$ tel que pour tout $n \geq 1$ et tout $t_1, \dots, t_n$, $(X_{t_1}, \dots, X_{t_n})$ a pour loi $\mu_{(t_1,\dots,t_n)}$.

Appliquons ce résultat à la question sur l'existence de processus gaussiens formulée plus haut. Pour cela, il nous suffit de vérifier la condition de compatibilité, qui peut se ré-exprimer de la façon suivante. Si $(Z_1, \dots, Z_{n+1})$ est un vecteur gaussien de moyenne $(m_1, \dots, m_{n+1})$ et de matrice de covariance $(C(i,j))_{1 \leq i,j \leq n+1}$ alors $(Z_1, \dots, Z_n)$ est un vecteur gaussien de moyenne $(m_1, \dots, m_n)$ et de matrice de covariance $(C(i,j))_{1 \leq i,j \leq n}$. Cette propriété est facile à vérifier et est laissée en exercice.

La tribu engendrée par un processus stochastique X, parfois notée $\sigma(X)$, est la plus petite tribu sur Ω rendant mesurables les v.a. X_t, $t \in \mathbb{R}_+$. On dira qu'un processus stochastique

X est indépendant d'une tribu $\mathcal{G}$ (sur Ω) si les tribus $\sigma(X)$ et $\mathcal{G}$ sont indépendantes. On peut vérifier que cela est vrai si et seulement si toutes les marginales finies dimensionnelles de X sont indépendantes de la tribu $\mathcal{G}$. De même, on dira que deux processus stochastiques X et Y, définis sur un même espace de probabilité, sont indépendants si $\sigma(X)$ et $\sigma(Y)$ sont indépendants.

Lemme I.8

Soient X et Y deux processus stochastiques sur $(\Omega, \mathcal{F})$. Si pour tous $n, m \geq 1$ et tout $s_1 < \ldots < s_m$ et $t_1 < \ldots < t_n$, les variables aléatoires $(X_{t_1}, \ldots, X_{t_n})$ et $(Y_{s_1}, \ldots, Y_{s_m})$ sont indépendantes, alors les processus X et Y sont indépendants.

(En fait, il suffit de prendre $n = m$ et $s_i = t_i$.)

Démonstration [♠] On note $\mathcal{C}_X$ la collection des événements

$$\{(X_{t_1}, \ldots, X_{t_n}) \in A_1 \times \ldots \times A_n\} ,$$

pour tout $A_1, \ldots, A_n \in \mathcal{B}(\mathbb{R})$, tout $n \geq 1$ et tout $t_1 < \ldots < t_n$. Il s'agit d'une classe stable par intersection finie, qui engendre $\sigma(X)$. On introduit $\mathcal{C}_Y$ de manière analogue. L'indépendance des v.a. de l'énoncé assure que tout événement de $\mathcal{C}_X$ est indépendant de tout événement de $\mathcal{C}_Y$. Un résultat classique assure alors que $\sigma(X)$ et $\sigma(Y)$ sont indépendantes. □

2 Définition du mouvement Brownien

Définition I.9

Un processus stochastique $(B_t, t \in \mathbb{R}_+)$ est appelé mouvement Brownien (issu de 0) si :

1. *B est un processus gaussien centré de fonction de covariance*

$$\mathrm{Cov}(B_s, B_t) = \mathbb{E}[B_s B_t] = \min(s, t) , \quad \forall s, t \in \mathbb{R}_+ ,$$

2. *les trajectoires de B sont continues.*

Quelques remarques s'imposent. Tout d'abord, un tel processus vérifie $B_0 = 0$ p.s. car toute variable gaussienne de variance nulle est égale p.s. à sa moyenne. Par ailleurs, par la Proposition I.6 la loi d'un processus vérifiant cette définition est unique. On pourra parler de la loi du mouvement Brownien sans ambiguité.

L'existence d'un processus vérifiant la première condition a déjà été démontrée : en effet, il n'est pas difficile de vérifier que la fonction de covariance de la définition est bien symétrique positive et l'on peut alors appliquer le théorème d'extension de Kolmogorov comme expliqué à la section précédente. En revanche la continuité d'un tel processus n'est pas acquise : ce faisant, l'existence du mouvement Brownien reste à établir. Ce sera l'objet des deux sections suivantes.

Notons enfin qu'il existe des processus gaussiens qui ne sont pas continus. Par exemple, si X est un processus gaussien centré de fonction de covariance $\Gamma(s, t) = \min(s, t)\mathbf{1}_{\{s,t \in \mathbb{Q}_+\}}$, alors

X est constant égal à 0 aux temps irrationnels et a une variance non-triviale aux temps rationnels : s'il était continu, il serait constant égal à 0 partout.
Commençons par manipuler les propriétés présentées dans la définition.

Proposition I.10

Soit $(B_t, t \in \mathbb{R}_+)$ un processus stochastique. Il y a équivalence entre :

(i) B est un processus gaussien centré de fonction de covariance $\mathrm{Cov}(B_s, B_t) = \min(s,t)$,

(ii) B vérifie :

(a) $B_0 = 0$ p.s.,

(b) pour tout $n \geq 2$, pour tous $0 \leq t_1 < t_2 < \ldots < t_n$, les variables aléatoires B_{t_1}, $B_{t_2} - B_{t_1}$, …, $B_{t_n} - B_{t_{n-1}}$ sont indépendantes,

(c) pour tous $0 \leq s \leq t$, la variable $B_t - B_s$ suit la loi gaussienne $\mathcal{N}(0, t-s)$

Ainsi tout processus B dont les trajectoires sont continues et qui satisfait les propriétés (a), (b) et (c) est un mouvement Brownien.

Démonstration Supposons que B vérifie (i). On a déjà vu plus haut que $B_0 = 0$ p.s, d'où le point (a). Par ailleurs la variable $B_t - B_s$ est gaussienne centrée de variance

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

ce qui prouve le point (c). En ce qui concerne les point (b), le vecteur $(B_{t_1}, B_{t_2} - B_{t_1}, \cdots, B_{t_n} - B_{t_{n-1}})$ étant gaussien, il suffit de vérifier que sa matrice de covariance est diagonale (exercice).
Supposons que B vérifie (ii). Les points (b) et (c) assurent que pour tout $n \geq 2$ et pour tous $0 \leq t_1 < t_2 < \ldots < t_n$, les variables B_{t_1}, $B_{t_2} - B_{t_1}$, …, $B_{t_n} - B_{t_{n-1}}$ sont indépendantes et gaussiennes centrées de variances t_1, $t_2 - t_1$, $\cdots$, $t_n - t_{n-1}$; ces variables forment donc un vecteur gaussien de matrice de covariance diagonale avec pour entrées sur la diagonale les variances précédentes. La stabilité sous transformation linéaire des vecteurs gaussiens assure que $(B_{t_1}, \cdots, B_{t_n})$ forme un vecteur gaussien centré tel que pour tous $1 \leq i \leq j \leq n$

$$\begin{aligned}
\mathrm{Cov}(B_{t_i}, B_{t_j}) &= \sum_{k=1}^{i} \sum_{\ell=1}^{j} \mathrm{Cov}(B_{t_k} - B_{t_{k-1}}, B_{t_\ell} - B_{t_{\ell-1}}) \\
&= \sum_{k=1}^{i} \mathrm{Var}(B_{t_k} - B_{t_{k-1}}) \\
&= t_i = \min(t_i, t_j)\ .
\end{aligned}$$

On a donc prouvé que B est un processus gaussien centré de fonction de covariance $\mathrm{Cov}(B_s, B_t) = \min(s,t)$. □

Corollaire I.11

Soit B un mouvement Brownien. Pour tout $n \geq 1$ et tous $0 \leq t_1 < \cdots < t_n$, la loi du *vecteur* $(B_{t_1}, B_{t_2}, \dots, B_{t_n})$ *admet une densité sur* $\mathbb{R}^n$ *donnée par*

$$p(x_1, \dots, x_n) = \frac{1}{(2\pi)^{n/2}\sqrt{t_1(t_2-t_1)\cdots(t_n-t_{n-1})}} \exp\left(-\sum_{i=1}^{n} \frac{(x_i - x_{i-1})^2}{2(t_i - t_{i-1})}\right) ,$$

avec $x_0 = 0$ *et* $(x_1, \dots, x_n) \in \mathbb{R}^n$.

Démonstration Comme on l'a vu dans la preuve précédente, le vecteur $(B_{t_1}, B_{t_2} - B_{t_1}, \cdots, B_{t_n} - B_{t_{n-1}})$ est composé de v.a. gaussiennes centrées indépendantes, il est donc gaussien et admet pour densité (en posant $t_0 := 0$)

$$p(y_1, \dots, y_n) = \frac{1}{(2\pi)^{n/2}\sqrt{t_1(t_2-t_1)\cdots(t_n-t_{n-1})}} \exp\left(-\sum_{i=1}^{n} \frac{y_i^2}{2(t_i - t_{i-1})}\right) .$$

En appliquant le changement de variables $x_i = y_1 + \cdots + y_i$ (de Jacobien égal à 1) on obtient le résultat voulu. □

Pour l'instant, nous n'avons pas établi l'existence du mouvement Brownien : la difficulté réside dans la continuité des trajectoires, que nous examinons à présent.

3 Continuité des trajectoires

Afin de prouver l'existence du mouvement Brownien, nous allons nous appuyer sur un résultat général dû à Kolmogorov qui affirme que, sous une hypothèse sur les moments d'un processus stochastique, on peut "modifier" le processus pour le rendre continu. Commençons par introduire cette notion de modification :

Définition I.12

Soient deux processus stochastiques $(X_t, t \in I)$ *et* $(\tilde{X}_t, t \in I)$ *indexés par un ensemble* $I \subset \mathbb{R}$*, et définis sur un même espace de probabilité. On dit que* $\tilde{X}$ *est une modification de* X *si*

$$\forall t \in I, \quad \mathbb{P}(X_t = \tilde{X}_t) = 1 .$$

On notera que si $\tilde{X}$ est une modification de X, alors les marginales de dimension finie des deux processus coïncident et ainsi les deux processus ont même loi. En revanche, il se peut que les trajectoires des deux processus soient très différentes ! En effet, l'égalité de la définition ne fournit aucun contrôle sur des collections *non-dénombrables* d'indices $t \in I$. Pour cela, il faut imposer une condition plus forte, ce qui donne lieu à la notion suivante.

Définition I.13

Soient deux processus stochastiques $(X_t, t \in I)$ *et* $(\tilde{X}_t, t \in I)$ *indexés par un ensemble* $I \subset \mathbb{R}$*, et définis sur un même espace de probabilité. On dit que* X *et* $\tilde{X}$ *sont indistinguables si*

$$\mathbb{P}(\forall t \in I : X_t = \tilde{X}_t) = 1 .$$

Remarque I.14 (♠)

Une formulation plus rigoureuse est : On dit que X et $\tilde{X}$ sont indistinguables s'il existe un sous-ensemble $N \subset \Omega$ qui est négligable, c'est-à-dire, N est inclus dans un événement de mesure nulle et tel que

$$\forall \omega \in \Omega \backslash N, \forall t \in I, \quad X_t(\omega) = \tilde{X}_t(\omega) \ .$$

En supposant N négligeable, on relâche la contrainte de mesurabilité de l'ensemble des réalisations où les deux processus diffèrent. Il s'agit là d'un point de détail.

Il est facile de vérifier que si deux processus sont indistinguables, alors l'un est une modification de l'autre.

Théorème I.15 (*Théorème de continuité de Kolmogorov*)

Soit $(X_t, t \in I)$ un processus stochastique à valeurs réelles indicé par un intervalle borné *$I \subset \mathbb{R}$. Supposons qu'il existe des réels $q, \epsilon, C > 0$ tels que pour tous $s, t \in I$*

$$\mathbb{E}[|X_s - X_t|^q] \leq C|t-s|^{1+\epsilon} \ .$$

Alors il existe une modification $\tilde{X}$ de X dont les trajectoires sont Hölderiennes d'exposant α pour tout $\omega \in \Omega$ et tout $\alpha \in]0, \epsilon/q[$, c'est-à-dire, pour tout $\alpha \in]0, \epsilon/q[$ il existe une variable aléatoire positive C_α telle que pour tous $s, t \in I$ et tout $\omega \in \Omega$

$$|\tilde{X}_s(\omega) - \tilde{X}_t(\omega)| \leq C_\alpha(\omega)|t-s|^\alpha \ .$$

Dans la preuve, nous aurons besoin du lemme technique suivant. Soit D l'ensemble dénombrable des nombres dyadiques de l'intervalle $[0, 1]$.

Lemme I.16

Soit $f : [0, 1] \to \mathbb{R}$. On suppose qu'il existe $\alpha, K > 0$ tels que pour tout entier $n \geq 1$ et tout $i \in \{1, \ldots, 2^n\}$

$$|f((i-1)2^{-n}) - f(i2^{-n})| \leq K2^{-n\alpha} \ .$$

Alors pour tous $s, t \in D$

$$|f(s) - f(t)| \leq \frac{2K}{1-2^{-\alpha}}|t-s|^\alpha \ .$$

Démonstration Soient $s, t \in D$ tels que $s < t$. On note $p \geq 0$ l'unique entier tel que $2^{-p} \leq t - s < 2^{-(p-1)}$. Nécessairement il existe k, ℓ, m entiers tels que

$$\begin{aligned} s &= k2^{-p} - \epsilon_{p+1}2^{-p-1} - \cdots - \epsilon_{p+\ell}2^{-p-\ell} \ , \\ t &= k2^{-p} + \epsilon'_p 2^{-p} + \epsilon'_{p+1}2^{-p-1} + \cdots + \epsilon'_{p+m}2^{-p-m} \ . \end{aligned}$$

où les ϵ_i, ϵ'_j valent 0 ou 1. On pose

$$\begin{aligned} s_i &= k2^{-p} - \epsilon_{p+1}2^{-p-1} - \cdots - \epsilon_{p+i}2^{-p-i} \;, \quad i \in \{0,\dots,\ell\} \\ t_j &= k2^{-p} + \epsilon'_p 2^{-p} + \epsilon'_{p+1}2^{-p-1} + \cdots + \epsilon'_{p+j}2^{-p-j} \;, \quad j \in \{0,\dots,m\} \;. \end{aligned}$$

On observe que $s_\ell = s$ et $t_m = t$. Ainsi l'hypothèse du lemme assure que

$$\begin{aligned} |f(s)-f(t)| &= |f(s_\ell)-f(t_m)| \\ &\le |f(s_0)-f(t_0)| + \sum_{i=1}^{\ell} |f(s_{i-1})-f(s_i)| + \sum_{j=1}^{m} |f(t_{j-1})-f(t_j)| \\ &= K\Big(2^{-p\alpha} + \sum_{i=1}^{\ell} 2^{-(p+i)\alpha} + \sum_{j=1}^{m} 2^{-(p+j)\alpha}\Big) \\ &\le 2K2^{-p\alpha}(1-2^{-\alpha})^{-1} \le 2K(t-s)^\alpha(1-2^{-\alpha})^{-1} \;. \end{aligned}$$

□

Présentons à présent la preuve du théorème. On va se restreindre aux indices de temps dyadiques et montrer que le processus X, restreint aux dyadiques, est continu p.s. (cela fait sens car on ne manipule qu'un nombre dénombrable de v.a.). Puis on étendra cette définition à tous les indices de temps par densité des dyadiques et continuité. Il ne restera plus qu'à vérifier que l'objet ainsi construit est bien une modification du processus de départ.

Démonstration [du Théorème de continuité de Kolmogorov] Sans perte de généralité, on peut supposer que $I = [0,1]$. On fixe $\alpha \in]0, \epsilon/q[$. L'hypothèse de l'énoncé combinée à l'inégalité de Markov assure que pour tout $n \ge 1$ et pour tout $i \in \{1,\dots,2^n\}$

$$\mathbb{P}(|X_{(i-1)2^{-n}} - X_{i2^{-n}}| > 2^{-n\alpha}) \le 2^{n\alpha q}\mathbb{E}[|X_{(i-1)2^{-n}} - X_{i2^{-n}}|^q] \le C2^{-n(1+\epsilon)}2^{n\alpha q} \;.$$

Ainsi

$$\mathbb{P}\left(\bigcup_{i=1}^{2^n}\{|X_{(i-1)2^{-n}} - X_{i2^{-n}}| > 2^{-n\alpha}\}\right) \le C2^{-n\epsilon}2^{n\alpha q} \;.$$

Comme $\epsilon - \alpha q > 0$ on obtient

$$\sum_{n\ge 1}\mathbb{P}\left(\bigcup_{i=1}^{2^n}\{|X_{(i-1)2^{-n}} - X_{i2^{-n}}| > 2^{-n\alpha}\}\right) < \infty \;.$$

Le lemme de Borel-Cantelli assure donc que pour $\mathbb{P}$-presque tout $\omega \in \Omega$, il existe un entier $n_0 = n_0(\omega)$ tel que

$$\forall n \ge n_0, \forall i \in \{1,\dots,2^n\}, \quad |X_{(i-1)2^{-n}} - X_{i2^{-n}}| \le 2^{-n\alpha} \;.$$

Ceci assure que la variable aléatoire

$$K_\alpha := \sup_{n\ge 1}\sup_{i\in\{1,\dots,2^n\}} \frac{|X_{(i-1)2^{-n}} - X_{i2^{-n}}|}{2^{-n\alpha}} \;,$$

est finie presque sûrement. (En effet pour $n \geq n_0(\omega)$ cette quantité est majorée par 1, tandis que le supremum sur $n < n_0$ porte sur un nombre fini de termes.)
Sur l'événement $\{K_\alpha < \infty\}$, qui est de probabilité 1, on déduit du lemme précédent que pour tous $s, t \in D$

$$|X_s(\omega) - X_t(\omega)| \leq C_\alpha(\omega)|t-s|^\alpha ,$$

avec $C_\alpha(\omega) = 2K_\alpha(\omega)/(1-2^{-\alpha})$. On a donc montré que sur l'événement $\{K_\alpha < \infty\}$, les trajectoires de X restreintes à l'ensemble D sont α-Höldériennes. On peut alors poser sans ambiguité

$$\tilde{X}_t(\omega) := \begin{cases} \lim_{s\to t, s\in D} X_s(\omega) & \text{si } K_\alpha(\omega) < \infty \\ 0 & \text{si } K_\alpha(\omega) = \infty \ . \end{cases}$$

Il reste à montrer que $\tilde{X}$ est une modification de X. Si $t \in D$ alors

$$\mathbb{P}(\tilde{X}_t \neq X_t) \leq \mathbb{P}(K_\alpha = \infty) = 0 \ .$$

Supposons à présent que $t \notin D$. L'hypothèse du théorème assure que pour tout $s \in [0,1]$

$$\mathbb{E}[1 \wedge |X_t - X_s|^q] \leq \mathbb{E}[|X_t - X_s|^q] \leq C|t-s|^{1+\epsilon} \ .$$

Pour $s \in D$, on sait déjà que $X_s = \tilde{X}_s$ p.s. de sorte que

$$\mathbb{E}[1 \wedge |X_t - \tilde{X}_s|^q] \leq C|t-s|^{1+\epsilon} \ .$$

En faisant tendre $s \to t$ avec $s \in D$, le théorème de convergence dominée assure que

$$\mathbb{E}[1 \wedge |X_t - \tilde{X}_t|^q] = 0 \ ,$$

et ainsi $X_t = \tilde{X}_t$ p.s. □

Corollaire I.17

Le mouvement Brownien existe !

Démonstration Soit B un processus gaussien centré de fonction de covariance $\mathrm{Cov}(B_s, B_t) = \min(s,t)$. (L'existence d'un tel processus a déjà été discutée plus haut). Afin d'appliquer le théorème de continuité de Kolmogorov, il nous faut évaluer les moments des incréments de ce processus. On remarque que $B_t - B_s$ a même loi que $\sqrt{t-s}\,Y$ où Y suit une loi normale centrée réduite. Ainsi en posant $C_q := \mathbb{E}[|Y|^q]$ qui est fini (les gaussiennes admettent des moments de tous ordres) on obtient

$$\mathbb{E}[|B_t - B_s|^q] = C_q(t-s)^{q/2} \ .$$

En choisissant $q > 2$ et un intervalle borné $I \subset \mathbb{R}_+$ arbitraire, on déduit du théorème l'existence d'une modification $(\tilde{B}_t, t \in I)$ dont les trajectoires sont Höldériennes donc continues. En appliquant cette construction sur chacun des intervalles $[0,1]$, $[1,2]$, $\cdots$, $[n, n+1]$, $\cdots$ on obtient alors une collection de modifications que l'on peut concaténer afin de former un processus $(\tilde{B}_t, t \in \mathbb{R}_+)$ à trajectoires continues et qui est une modification de $(B_t, t \in \mathbb{R}_+)$. □

Construction du mouvement Brownien par Paul Lévy. Nous venons d'appliquer un théorème général de la théorie des processus afin de construire le mouvement Brownien. Il se trouve que l'on peut donner une construction beaucoup plus géométrique (et élégante!) du mouvement Brownien. Celle-ci est due à Paul Lévy et s'appuie sur l'observation suivante

Exercice I.18

Pour tout $0 \le s < t$, le vecteur $(B_s, B_{(s+t)/2}, B_t)$ a même loi que le vecteur

$$(B_s, \frac{B_s + B_t}{2} + Y, B_t) ,$$

où Y est une variable aléatoire $\mathcal{N}(0,(t-s)/4)$ indépendante de B_s et B_t.

On se restreint à l'intervalle $[0,1]$, on note $D_n := \{k2^{-n} : k \in \{0,\ldots,2^n\}\}$ et l'on construit de façon récursive une suite de processus $(B_t^{(n)}, t \in [0,1])$ tels que :

1. Les marginales de dimensions finies de $B^{(n)}$ restreint à D_n coïncident avec celles du Brownien. Plus précisément : les vecteurs $(B_t^{(n)}, t \in D_n)$ et $(B_t, t \in D_n)$ ont même loi.
2. La procédure est consistente sur les dyadiques. Plus précisément : $B_t^{(n)} = B_t^{(m)}$ pour tout $t \in D_n \cap D_m$.
3. Chaque processus $B^{(n)}$ est affine entre deux points consécutifs de D_n. Plus précisément :
$$B_t^{(n)} = 2^n(t - k2^{-n})B^{(n)}_{(k+1)2^{-n}} + 2^n((k+1)2^{-n} - t)B^{(n)}_{k2^{-n}} , \quad t \in [k2^{-n}, (k+1)2^{-n}] .$$

On commence la procédure en se donnant une variable aléatoire X de loi $\mathcal{N}(0,1)$ et l'on pose

$$B_t^{(0)} := tX .$$

On suppose avoir construit le processus $B^{(n)}$ pour un certain entier $n \ge 0$. Pour construire $B^{(n+1)}$ il suffit de définir correctement les marginales aux temps $t \in D_{n+1} \backslash D_n$ puis d'interpoler linéairement entre les points de D_{n+1}. Pour ce faire, on se donne une collection de v.a. $X_{k,n}$, $k \in \{0,\ldots,2^n - 1\}$ i.i.d. $\mathcal{N}(0, 2^{-n}/4)$, indépendantes de toutes les v.a. utilisées jusqu'alors dans la construction. Pour chaque $k \in \{0,\ldots,2^n - 1\}$, on pose

$$B^{(n+1)}_{(2k+1)2^{-(n+1)}} := B^{(n)}_{(2k+1)2^{-(n+1)}} + X_{k,n} ,$$

tandis que pour tout $t \in D_n$, on pose

$$B_t^{(n+1)} = B_t^{(n)} .$$

Il est clair que $(B_t^{(n+1)}, t \in D_{n+1})$ est un vecteur gaussien centré. Pour s'assurer qu'il a la bonne loi, il suffit de calculer les covariances entre les entrées de ce vecteur (exercice). On peut ensuite achever la construction de $B^{(n+1)}$ en interpolant de façon affine entre deux points consécutifs de D_{n+1}.

On peut alors montrer qu'avec probabilité 1, la suite de fonctions continues $(B_t^{(n)}, t \in [0,1])$ converge uniformément vers un objet limite que l'on note $(B_t, t \in [0,1])$.

Nécessairement B est continu. Par construction, ses marginales de dimensions finies restreintes à D coïncident en loi avec celle du mouvement Brownien. Pour conclure, on utilise le lemme suivant.

Lemme I.19

♠ *Soient* $(X_t, t \in [0,1])$ *et* $(Y_t, t \in [0,1])$ *deux processus stochastiques dont les trajectoires sont continues. Supposons qu'il existe un ensemble* $J \subset [0,1]$ *dense dans* $[0,1]$ *tel que les marginales de dimensions finies de* X *et* Y*, restreintes à* J*, coïncident en loi. Alors les deux processus* $(X_t, t \in [0,1])$ *et* $(Y_t, t \in [0,1])$ *ont même loi.*

Démonstration Il suffit de montrer que pour tout $n \geq 1$ et tous $0 \leq t_1 \leq t_2 \leq \ldots \leq t_n$, les lois des vecteurs $(X_{t_1}, \ldots, X_{t_n})$ et $(Y_{t_1}, \ldots, Y_{t_n})$ coïncident.
On sait que cela est vrai dès que tous les indices de temps appartiennent à J. Pour conclure, il suffit de montrer que le vecteur aléatoire $(X_{s_1}, \ldots, X_{s_n})$ (resp. $(Y_{s_1}, \ldots, Y_{s_n})$) converge en loi vers $(X_{t_1}, \ldots, X_{t_n})$ (resp. $(Y_{t_1}, \ldots, Y_{t_n})$) dès que $s_i \to t_i$ pour tout $i \in \{1, \ldots, n\}$ avec les $s_i \in J$. Or, sous cette hypothèse, la continuité des trajectoires assure que $(X_{s_1}, \ldots, X_{s_n})$ converge presque sûrement vers $(X_{t_1}, \ldots, X_{t_n})$, et ainsi la convergence a également lieu en loi. □

4 Quelques propriétés du mouvement Brownien

Dans cette partie, $B = (B_t, t \in \mathbb{R}_+)$ désignera toujours un mouvement Brownien. Pour tout $t \geq 0$, on note $\mathcal{F}_t := \sigma(B_s, s \in [0,t])$: intuitivement, il s'agit de la plus petite tribu sur Ω contenant toute l'information encodée par la trajectoire Brownienne jusqu'au temps t. On introduit également $\mathcal{F}_\infty := \sigma(B_s, s \geq 0)$.

Proposition I.20 (*Propriétés d'invariance*)

1. *$-B$ est un mouvement Brownien [invariance par changement de signe],*
2. *pour tout $\lambda > 0$, le processus B^λ, défini par $B_t^\lambda := \lambda^{-1} B_{t\lambda^2}$, est un mouvement Brownien [invariance par changement d'échelle],*
3. *pour tout $T \geq 0$, le processus $B^{(T)}$, défini par $B_t^{(T)} := B_{T+t} - B_T$, est un mouvement Brownien [invariance par translation]*

Démonstration Dans les trois cas, il est immédiat que les processus sont gaussiens centrés : en effet, toute combinaison linéaire de marginales de dimension finie du processus $-B$, resp. B^λ, resp. $B^{(T)}$, est combinaison linéaire de marginales de dimension finie du processus B, et ainsi, est une gaussienne centrée. Par ailleurs, la continuité des trajectoires est préservée dans chaque cas. Il reste à calculer la fonction de covariance dans chaque cas.

$$\begin{aligned}
\mathrm{Cov}(-B_t, -B_s) &= \mathrm{Cov}(B_t, B_s) = \min(s,t) \ , \\
\mathrm{Cov}(B_t^\lambda, B_s^\lambda) &= \lambda^{-2}\,\mathrm{Cov}(B_{t\lambda^2}, B_{s\lambda^2}) = \lambda^{-2}\min(t\lambda^2, s\lambda^2) = \min(t,s) \ , \\
\mathrm{Cov}(B_t^{(T)}, B_s^{(T)}) &= \mathrm{Cov}(B_{T+t} - B_T, B_{T+s} - B_T) \\
&= \mathrm{Cov}(B_{T+t}, B_{T+s}) - \mathrm{Cov}(B_T, B_{T+s}) - \mathrm{Cov}(B_T, B_{T+t}) + \mathrm{Cov}(B_T, B_T) \\
&= \min(T+s, T+t) + T - 2T = \min(s,t) \ .
\end{aligned}$$

□

La propriété d'invariance par translation s'agrémente d'une propriété d'indépendance.

Proposition I.21 (*Propriété de Markov simple*)

Pour tout $T \geq 0$, le processus $B^{(T)}$, défini par $B_t^{(T)} := B_{T+t} - B_T$, est un mouvement Brownien indépendant de la tribu $\mathcal{F}_T$.

Démonstration Nous avons déjà montré que $B^{(T)}$ était un mouvement Brownien. Il nous faut prouver qu'il est indépendant de $\mathcal{F}_T$. Rappelons que cela signifie que les tribus engendrées par $(B_s, s \in [0,T])$ d'une part et par $(B_{T+t} - B_T, t \in \mathbb{R}_+)$ d'autre part sont indépendantes. Pour prouver cette assertion, il suffit de montrer que pour tous $n, m \geq 1$ et tous $0 \leq s_1 \leq s_2 \leq \ldots \leq s_n \leq T$, $0 \leq t_1 \leq t_2 \leq \ldots \leq t_m$ les vecteurs

$$(B_{s_1}, \ldots, B_{s_n}) \ , \quad (B_{T+t_1} - B_T, \ldots, B_{T+t_m} - B_T) \ ,$$

sont indépendants. Comme ces deux vecteurs forment conjointement un vecteur gaussien dans $\mathbb{R}^{n+m}$, l'indépendance des deux vecteurs est équivalente aux identités

$$\mathrm{Cov}(B_{s_i}, B_{T+t_j} - B_T) = 0 \ , \quad \forall i \in \{1, \ldots, n\} \ , \forall j \in \{1, \ldots, m\} \ .$$

On vérifie aisément que ces identités sont vraies. □

Le résultat suivant montre que le comportement "juste après" le temps 0 suit la loi du tout ou rien.

Proposition I.22

La tribu

$$\mathcal{F}_{0+} := \bigcap_{s>0} \mathcal{F}_s \ ,$$

est grossière au sens suivant : pour tout $A \in \mathcal{F}_{0+}$, $\mathbb{P}(A) = 0$ ou 1.

Démonstration Il suffit de montrer que $\mathcal{F}_{0+}$ est indépendante de $\mathcal{F}_\infty$. En effet, comme $\mathcal{F}_{0+} \subset \mathcal{F}_\infty$, cela permet de déduire que $\mathcal{F}_{0+}$ est indépendante d'elle-même, et est donc grossière.

Soient $n \geq 1$, $0 < t_1 < \ldots < t_n$ et $f : \mathbb{R}^n \to \mathbb{R}$ une fonction continue bornée. La continuité des trajectoires combinée au Théorème de Convergence Dominée assure que pour tout $A \in \mathcal{F}_{0+}$

$$\mathbb{E}[\mathbf{1}_A f(B_{t_1}, \ldots, B_{t_n})] = \lim_{\epsilon \to 0} \mathbb{E}[\mathbf{1}_A f(B_{t_1} - B_\epsilon, \ldots, B_{t_n} - B_\epsilon)] \ .$$

Or $A \in \mathcal{F}_\epsilon$ et la propriété de Markov simple assure que le vecteur $(B_{t_1} - B_\epsilon, \dots, B_{t_n} - B_\epsilon)$ est indépendant de $\mathcal{F}_\epsilon$. Ainsi

$$\mathbb{E}[\mathbf{1}_A f(B_{t_1}, \dots, B_{t_n})] = \lim_{\epsilon \to 0} \mathbb{P}(A)\mathbb{E}[f(B_{t_1} - B_\epsilon, \dots, B_{t_n} - B_\epsilon)] = \mathbb{P}(A)\mathbb{E}[f(B_{t_1}, \dots, B_{t_n})] \ .$$

On a donc montré que $\mathcal{F}_{0+}$ est indépendante de $\sigma(B_t, t > 0)$. Par continuité des trajectoires, $B_0 = \lim_{t \downarrow 0} B_t$ et ainsi B_0 est $\sigma(B_t, t > 0)$-mesurable, ce qui assure que $\sigma(B_t, t > 0) = \sigma(B_t, t \geq 0) = \mathcal{F}_\infty$, et achève la preuve. □

On déduit de cette dernière proposition des propriétés trajectorielles remarquables.

Corollaire I.23

On a presque sûrement

$$\forall \epsilon > 0, \quad \sup_{0 \leq s \leq \epsilon} B_s > 0 \ , \quad \inf_{0 \leq s \leq \epsilon} B_s < 0 \ .$$

Par ailleurs pour tout $a \in \mathbb{R}$, si l'on note $T_a := \inf\{t \geq 0 : B_t = a\}$ alors presque sûrement

$$\forall a \in \mathbb{R}, \quad T_a < \infty \ .$$

En conséquence presque sûrement

$$\limsup_{t \to \infty} B_t = +\infty \ , \quad \liminf_{t \to \infty} B_t = -\infty \ .$$

Démonstration Soit $(\epsilon_n)_n$ une suite de réels strictement positifs décroissant vers 0. On pose

$$A := \bigcap_n \{ \sup_{0 \leq s \leq \epsilon_n} B_s > 0 \} \ .$$

L'événement A est $\mathcal{F}_{0+}$-mesurable et est une intersection d'événements décroissants, de sorte que

$$\mathbb{P}(A) = \lim_{n \to \infty} \mathbb{P}(\sup_{0 \leq s \leq \epsilon_n} B_s > 0) \ .$$

Or

$$\mathbb{P}(\sup_{0 \leq s \leq \epsilon_n} B_s > 0) \geq \mathbb{P}(B_{\epsilon_n} > 0) = 1/2 \ ,$$

et ainsi la probabilité de A est supérieure ou égale à $1/2$. Cet événement étant trivial, sa probabilité est donc égale à 1. On a donc montré que presque sûrement $\forall \epsilon > 0$, $\sup_{0 \leq s \leq \epsilon} B_s > 0$. L'assertion concernant l'infimum est obtenue en remplaçant B par $-B$.
On prouve à présent la finitude p.s. de tous les T_a. On commence par écrire

$$1 = \mathbb{P}(\sup_{0 \leq s \leq 1} B_s > 0) = \lim_{\delta \downarrow 0} \mathbb{P}(\sup_{0 \leq s \leq 1} B_s > \delta) \ .$$

On fixe $a > 0$. En posant $\lambda = \delta / a$ et en utilisant l'invariance d'échelle du mouvement Brownien on obtient

$$\mathbb{P}(\sup_{0 \leq s \leq 1} B_s > \delta) = \mathbb{P}(\sup_{0 \leq s \leq 1} \lambda^{-1} B_s > \lambda^{-1} \delta) = \mathbb{P}(\sup_{0 \leq s \leq \lambda^{-2}} \lambda^{-1} B_{s\lambda^2} > a) = \mathbb{P}(\sup_{0 \leq s \leq \lambda^{-2}} B_s > a) \ .$$

En prenant la limite $\delta \downarrow 0$, c'est-à-dire, $\lambda^{-2} \uparrow \infty$ on obtient que

$$\mathbb{P}(\sup_{s\geq 0} B_s > a) = 1 \ .$$

On a donc montré que $T_a < \infty$ p.s. On en déduit que presque sûrement, pour tout $a \in (0,\infty) \cap \mathbb{Q}$, $T_a < \infty$. Or la continuité des trajectoires assure que pour tout $0 < a < b$, $T_b < \infty \Rightarrow T_a < \infty$. On a donc montré que presque sûrement pour tout $a > 0$, $T_a < \infty$.
En remplaçant B par $-B$, on en déduit que cela reste vrai pour $a < 0$.
On a donc montré que presque sûrement la trajectoire $t \mapsto B_t$ visite toutes les valeurs $a \in \mathbb{R}$. Nécessairement $\limsup_{t\to\infty} B_t = +\infty$ et $\liminf_{t\to\infty} B_t = -\infty$. □

5 Propriété de Markov forte du mouvement Brownien

Nous allons montrer que la propriété de Markov simple reste vraie lorsque le temps T est aléatoire mais n'anticipe pas le futur, c'est-à-dire, dès que T est un temps d'arrêt.

Définition I.24

On dit qu'une variable aléatoire T à valeurs dans $[0,\infty]$ est un temps d'arrêt si pour tout $t \geq 0$, l'événement $\{T \leq t\} \in \mathcal{F}_t$.

En termes intuitifs, cette définition impose que l'information encodée dans la tribu $\mathcal{F}_t$ suffit à déterminer si T s'est réalisé avant le temps t ou pas.

Donnons quelques exemples et contre-exemples de temps d'arrêt. La v.a. constante $T = t_0$ est un temps d'arrêt, ainsi que le premier temps d'atteinte de a par le Brownien

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

En revanche, le dernier temps d'atteinte de 0 avant le temps 1 par le Brownien

$$\sup\{t \leq 1 : B_t = 0\} \ ,$$

ainsi que le premier temps où la trajectoire Brownienne atteint son maximum sur $[0,1]$

$$\inf\{t \in [0,1] : B_t = \sup_{s\in[0,1]} B_s\} \ ,$$

ne sont pas des temps d'arrêt.

Enonçons quelques propriétés élémentaires vérifiées par tout temps d'arrêt.

Lemme I.25

Soit T un temps d'arrêt. Pour tout $t \geq 0$, les événements $\{T < t\}$ et $\{T = t\}$ sont dans $\mathcal{F}_t$. Par ailleurs, pour tout $n \geq 1$, $T_n := 2^{-n}(\lfloor 2^n T \rfloor + 1)$ est un temps d'arrêt.

Démonstration On note que

$$\{T < t\} = \bigcup_{q \in [0,t[\cap \mathbb{Q}} \{T \leq q\} ,$$

de sorte que cet événement est bien dans $\mathcal{F}_t$. En écrivant $\{T = t\} = \{T \leq t\} \backslash \{T < t\}$ la $\mathcal{F}_t$ mesurabilité de $\{T = t\}$ s'en suit.
On observe que T_n prend ses valeurs dans l'ensemble $k2^{-n}$, $k \geq 0$. Pour tout $t \geq 0$, on introduit $k \geq 0$ tel que $k2^{-n} \leq t < (k+1)2^{-n}$ et l'on écrit

$$\{T_n \leq t\} = \{T_n \leq k2^{-n}\} = \{\lfloor 2^n T \rfloor + 1 \leq k\} = \{2^n T < k\} = \{T < 2^{-n} k\} .$$

Comme $2^{-n} k \leq t$, ce dernier événement est $\mathcal{F}_t$-mesurable. □

Etant donné un temps d'arrêt T, on introduit la tribu du passé avant T en posant

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

(Nous reviendrons sur cette définition dans le chapitre suivant).
Pour se familiariser avec la tribu du passé, on pourra résoudre les deux exercices suivants.

Exercice I.26

Vérifier que $\mathcal{F}_T$ est bien une tribu.

Exercice I.27

Montrer que T est $\mathcal{F}_T$-mesurable.

L'exercice ci-dessous est plus difficile (sa solution est donnée dans un cadre plus général au chapitre suivant).

Exercice I.28 (♠)

Montrer que $\mathbf{1}_{\{T<\infty\}} B_T$ est $\mathcal{F}_T$-mesurable.

Théorème I.29 (*Propriété de Markov forte*)

Soit T un temps d'arrêt fini presque sûrement. On pose pour tout $t \geq 0$

$$B_t^{(T)} := \mathbf{1}_{\{T<\infty\}} (B_{T+t} - B_T) .$$

Le processus $B^{(T)}$ est un mouvement Brownien indépendant de $\mathcal{F}_T$.

Remarque I.30

Plus généralement, si T n'est pas fini presque sûrement mais que $T < \infty$ a une probabilité strictement positive, alors conditionnellement à l'événement $\{T < \infty\}$, le processus $B^{(T)}$ est un mouvement Brownien indépendant de $\mathcal{F}_T$.

La propriété de Markov forte a de très nombreuses applications dans l'étude du mouvement Brownien, nous en verrons certaines un peu plus loin. Notons qu'elle est fondamentalement due à l'indépendance et à la stationarité des incréments du mouvement Brownien.

Démonstration Nous allons montrer que pour tout $A \in \mathcal{F}_T$, tout $m \geq 1$, tous $0 \leq t_1 < t_2 < \ldots < t_m$, et toute fonction continue bornée $f : \mathbb{R}^m \to \mathbb{R}$

$$\mathbb{E}[\mathbf{1}_A f(B^{(T)}_{t_1}, \ldots, B^{(T)}_{t_m})] = \mathbb{P}(A)\mathbb{E}[f(B_{t_1}, \ldots, B_{t_m})] \;. \tag{I.1}$$

Cela suffit pour établir les assertions du théorème. En effet, en prenant $A = \Omega$, on déduit que $B^{(T)}$ a même loi qu'un mouvement Brownien : comme ses trajectoires sont également continues, on en déduit alors que c'est un mouvement Brownien. Par ailleurs cette identité assure que le vecteur $(B^{(T)}_{t_1}, \ldots, B^{(T)}_{t_m})$ est indépendant de la tribu $\mathcal{F}_T$, et ce pour tout choix de m et de valeurs $t_1, \ldots, t_m$. L'indépendance entre la tribu engendrée par $B^{(T)}$ et $\mathcal{F}_T$ s'en suit.

Pour établir (I.1), commençons par observer que le cas où T est déterministe est déjà couvert par la propriété de Markov simple. Supposons à présent que T prend ses valeurs dans un ensemble dénombrable $\{r_k, k \geq 1\}$ donné. On calcule alors

$$\begin{aligned}\mathbb{E}[\mathbf{1}_A f(B^{(T)}_{t_1}, \ldots, B^{(T)}_{t_m})] &= \sum_{k\geq 1} \mathbb{E}[\mathbf{1}_A f(B^{(T)}_{t_1}, \ldots, B^{(T)}_{t_m})\mathbf{1}_{\{T=r_k\}}] \\ &= \sum_{k\geq 1} \mathbb{E}[f(B_{r_k+t_1} - B_{r_k}, \ldots, B_{r_k+t_m} - B_{r_k})\mathbf{1}_{A\cap\{T=r_k\}}]\end{aligned}$$

Or $A \cap \{T = r_k\} \in \mathcal{F}_{r_k}$. Ainsi la propriété de Markov simple permet de poursuivre le calcul et d'obtenir

$$\begin{aligned}&= \sum_{k\geq 1} \mathbb{E}[f(B_{t_1}, \ldots, B_{t_m})]\mathbb{P}(A \cap \{T = r_k\}) \\ &= \mathbb{E}[f(B_{t_1}, \ldots, B_{t_m})]\mathbb{P}(A) \;.\end{aligned}$$

Dans le cas général où T prend ses valeurs dans $[0, \infty]$ (et est fini p.s.), on introduit la suite d'approximations $T_n := 2^{-n}(\lfloor 2^n T\rfloor + 1)$, $n \geq 1$. On note que $T_n \to T$ presque sûrement quand $n \to \infty$. Ainsi la continuité des trajectoires combinée au Théorème de Convergence Dominée assure que

$$\mathbb{E}[\mathbf{1}_A f(B^{(T)}_{t_1}, \ldots, B^{(T)}_{t_m})] = \lim_{n\to\infty} \mathbb{E}[\mathbf{1}_A f(B^{(T_n)}_{t_1}, \ldots, B^{(T_n)}_{t_m})] \;.$$

Le cas dénombrable déjà établi s'applique à T_n et $B^{(T_n)}$, et l'on obtient (en notant que si $A \in \mathcal{F}_T$ alors $A \in \mathcal{F}_{T_n}$)

$$\mathbb{E}[\mathbf{1}_A f(B^{(T_n)}_{t_1}, \ldots, B^{(T_n)}_{t_m})] = \mathbb{P}(A)\mathbb{E}[f(B_{t_1}, \ldots, B_{t_m})] \;,$$

de sorte que

$$\mathbb{E}[\mathbf{1}_A f(B^{(T)}_{t_1}, \ldots, B^{(T)}_{t_m})] = \mathbb{P}(A)\mathbb{E}[f(B_{t_1}, \ldots, B_{t_m})] \;.$$

□

Nous présentons une très jolie identité en loi portant sur le supremum courant du mouvement Brownien

$$S_t := \sup_{0\leq s\leq t} B_s \;, \quad t \geq 0 \;.$$

La preuve de cette identité repose sur la propriété de Markov forte du mouvement Brownien.

Théorème I.31 (*Principe de réflexion*)
Pour tout $t > 0$, tout $a \geq 0$ et tout $b \leq a$ on a

$$\mathbb{P}(S_t \geq a, B_t \leq b) = \mathbb{P}(B_t \geq 2a - b) \ .$$

En particulier, S_t a même loi que $|B_t|$.

La preuve de ce théorème sera vue en TD.

Démonstration On a déjà vu que $T_a := \inf\{t \geq 0 : B_t = a\} < \infty$ p.s. On écrit alors

$$\mathbb{P}(S_t \geq a, B_t \leq b) = \mathbb{P}(T_a \leq t, B_t \leq b) = \mathbb{P}(T_a \leq t, B^{(T_a)}_{t-T_a} \leq b - a) \ .$$

Le Théorème I.29 assure que $B^{(T_a)}$ est indépendant de T_a. Comme $B^{(T_a)}$ a même loi que $-B^{(T_a)}$ on en déduit que $(T_a, B^{(T_a)})$ a même loi que $(T_a, -B^{(T_a)})$ et ainsi

$$\mathbb{P}(T_a \leq t, B^{(T_a)}_{t-T_a} \leq b-a) = \mathbb{P}(T_a \leq t, -B^{(T_a)}_{t-T_a} \leq b-a) = \mathbb{P}(T_a \leq t, B_{T_a} - B_t \leq b-a) = \mathbb{P}(T_a \leq t, B_t \geq 2a-b) \ .$$

Comme $2a - b \geq a$, si $B_t \geq 2a - b$ alors nécessairement $T_a \leq t$ et ainsi

$$\mathbb{P}(S_t \geq a, B_t \leq b) = \mathbb{P}(B_t \geq 2a - b) \ .$$

Pour prouver la deuxième assertion, on écrit

$$\mathbb{P}(S_t \geq a) = \mathbb{P}(S_t \geq a, B_t \leq a) + \mathbb{P}(S_t \geq a, B_t > a) \ .$$

On note que si $B_t > a$ alors nécessairement $S_t \geq a$. En utilisant le principe de réflexion ainsi que cette dernière observation on obtient

$$\mathbb{P}(S_t \geq a) = \mathbb{P}(B_t \geq a) + \mathbb{P}(B_t > a) = 2\mathbb{P}(B_t \geq a) = \mathbb{P}(|B_t| \geq a) \ .$$

□

6 Mesure de Wiener, mouvement Brownien multi-dimensionnel

Soit $(B_t, t \geq 0)$ un mouvement Brownien. L'application

$$\begin{aligned} B : \Omega &\to \mathbb{R}^{\mathbb{R}_+} \\ \omega &\mapsto (B_t(\omega), t \in \mathbb{R}_+) \end{aligned}$$

est mesurable par rapport aux tribus $\mathcal{F}$ et $\mathcal{B}(\mathbb{R})^{\otimes \mathbb{R}_+}$. Par définition du mouvement Brownien, nous savons que les trajectoires de B sont continues et ainsi l'application prend ses valeurs dans $C(\mathbb{R}_+, \mathbb{R})$. Ainsi B est une application mesurable de $(\Omega, \mathcal{F})$ dans $C(\mathbb{R}_+, \mathbb{R})$ muni de la tribu produit, c'est-à-dire, de la plus petite tribu rendant mesurables les applications $x \mapsto x_t$ pour tous $t \in \mathbb{R}_+$. La loi du mouvement Brownien sur $C(\mathbb{R}_+, \mathbb{R})$ muni de la tribu produit est appelée *mesure de Wiener*.

Terminons ce chapitre en introduisant la version multi-dimensionnelle du mouvement Brownien. Un processus stochastique $(B_t, t \in \mathbb{R}_+) = ((B^{(1)}_t, \ldots, B^{(d)}_t), t \in \mathbb{R}_+)$ à valeurs dans $\mathbb{R}^d$ est un mouvement Brownien en dimension d si les processus $(B^{(i)}_t, t \in \mathbb{R}_+)$, $i \in \{1, \ldots, d\}$ sont des mouvements Browniens indépendants.

Annexe A

Généralités

1 Variables aléatoires

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

Définition A.1

Une application $X : \Omega \to E$ mesurable par rapport aux tribus $\mathcal{F}$ et $\mathcal{E}$ est appelée variable aléatoire.

On rappelle que X est mesurable par rapport aux tribus $\mathcal{F}$ et $\mathcal{E}$ si pour tout $B \in \mathcal{E}$, $X^{-1}(B) \in \mathcal{F}$.

Définition A.2

Soit X un variable aléatoire de $(\Omega, \mathcal{F}, \mathbb{P})$ dans $(E, \mathcal{E})$. L'application de $\mu_X : \mathcal{E} \to [0,1]$ définie par

$$\mu_X(B) := \mathbb{P}(X^{-1}(B)) \ , \quad B \in \mathcal{E} \ ,$$

est une mesure de probabilité sur $(E, \mathcal{E})$ que l'on appelle loi de X.

On notera que la loi de X n'est rien d'autre que l'image par X de la mesure de probabilité $\mathbb{P}$.

Quelques exemples :

1. Jet de dés : $\Omega = \{1, \ldots, 6\}^n$ muni de la tribu $\mathcal{F}$ de toutes les parties de Ω et de la mesure uniforme $\mathbb{P}$. L'application $X_k : \omega \mapsto \omega_k$ qui retourne la k-ème coordonnée est une variable aléatoire à valeurs dans $\{1, \ldots, 6\}$ muni de la tribu de toutes ses parties. Sa loi est uniforme. L'application $S : \Omega \to \sum_{k=1}^n X_k$ est une variable aléatoire à valeurs dans $\{n, \ldots, 6n\}$. Sa loi peut être explicitée.
2. Variable gaussienne : $\Omega = E = \mathbb{R}$, $\mathcal{F} = \mathcal{E} = \mathcal{B}(\mathbb{R})$ (la tribu Borélienne) et $\mathbb{P}$ la mesure de probabilté
$$\mathbb{P}(dx) = \frac{1}{\sqrt{2\pi}} e^{-x^2/2} dx \ .$$
3. Variable constante égale à $a \in E$: $\mu_X = \delta_a$ où l'on définit la masse de Dirac $\delta_a(B) = \mathbf{1}_B(a)$ pour tout $B \in \mathcal{E}$.

Proposition A.3 (*Formule de transfert*)

Soit X une variable aléatoire à valeurs dans un espace $(E, \mathcal{E})$ et $h : E \to \mathbb{R}$ une fonction $(\mathcal{E}, \mathcal{B}(\mathbb{R}))$-mesurable. Si $h(X)$ est intégrable alors

$$\mathbb{E}[h(X)] = \int_E h \, d\mu_X \, .$$

2 Variables gaussiennes

La loi gaussienne uni-dimensionnelle de moyenne $m \in \mathbb{R}$ et de variance $\sigma^2 > 0$ est la mesure de probabilité sur $\mathbb{R}$ de densité

$$\frac{1}{\sqrt{2\pi\sigma^2}} e^{-\frac{|x-m|^2}{2\sigma^2}} \, , \quad x \in \mathbb{R} \, .$$

Dans le cas $\sigma^2 = 0$, il s'agit de la mesure de probabilité sur $\mathbb{R}$ donnée par la masse de Dirac en m :

$$\delta_m(B) = \mathbf{1}_B(m) \, , \quad B \in \mathcal{B}(\mathbb{R}) \, .$$

Définition A.4

On dit qu'une variable aléatoire à valeurs dans $\mathbb{R}^d$ pour $d \geq 1$ est gaussienne si toute combinaison linéaire de ses coordonnées est une variable aléatoire gaussienne (uni-dimensionnelle).

Comme pour tout vecteur aléatoire à valeurs dans $\mathbb{R}^d$ (dont la norme euclidienne admet un moment d'ordre 2) on peut définir le vecteur moyenne m et la matrice de covariance Σ d'une v.a. gaussienne X

$$m_k := \mathbb{E}[X_k] \, , \quad \Sigma(k, \ell) = \mathrm{Cov}(X_k, X_\ell) \, .$$

On notera bien que m est un élément de $\mathbb{R}^d$ alors que Σ est une matrice symétrique positive de taille $d \times d$.

Proposition A.5

La fonction caractéristique d'une variable aléatoire gaussienne X est donnée par

$$\Phi_X(t) = \exp(i\langle t, m\rangle - \frac{\langle t, \Sigma t\rangle}{2}) \, , \quad t \in \mathbb{R}^d \, .$$

En conséquence, la loi d'une variable aléatoire gaussienne est entièrement déterminée par son vecteur moyenne et sa matrice de covariance.

Soit $Y = (Y_1, \ldots, Y_d)$ un vecteur aléatoire dont les coordonnées sont des v.a. gaussiennes uni-dimensionnelles, indépendantes centrées réduites. Il est facile de vérifier que Y est une variable aléatoire gaussienne de moyenne le vecteur nul et de matrice de covariance la matrice identité.

Etant donné un vecteur $m \in \mathbb{R}^d$ et une matrice $d \times d$ symétrique positive Σ, la variable aléatoire $X := m + \Sigma^{1/2} Y$ est gaussienne de moyenne m et de matrice de covariance Σ. (La matrice $\Sigma^{1/2}$ peut être définie comme suit. Soit P une matrice orthogonale telle que $D = P\Sigma P^t$ est diagonale, on définit alors $D^{1/2}$ comme la matrice diagonale dont les entrées sont les racines carrés des entrées de D (celles-ci sont positives car Σ est positive). On pose alors $\Sigma^{1/2} := P^t D^{1/2} P$.) On note $\mathcal{N}(m, \Sigma)$ la loi correspondante.

Proposition A.6

Soit $X \sim \mathcal{N}(m, \Sigma)$. X admet une densité sur $\mathbb{R}^d$ si et seulement si Σ est définie positive, et dans ce cas celle-ci est donnée par

$$\frac{1}{(2\pi)^{d/2}\sqrt{\det(\Sigma)}} \exp\Big(-\frac{\langle \Sigma^{-1}(x-m), x-m\rangle}{2}\Big), \quad x \in \mathbb{R}^d .$$

3 Espérance conditionnelle

Dans cette partie, on manipule des variables aléatoires réelles.

Définition A.7 (*Espérance conditionnelle dans L^1*)

Soit $(\Omega, \mathcal{F}, \mathbb{P})$ un espace de probabilité, $\mathcal{B} \subseteq \mathcal{F}$ une sous-tribu et $X \in L^1(\Omega, \mathcal{F}, \mathbb{P})$. On définit l'espérance conditionnelle de X sachant $\mathcal{B}$, que l'on note $\mathbb{E}[X \mid \mathcal{B}]$, comme l'unique variable aléatoire $U \in L^1(\Omega, \mathcal{B}, \mathbb{P})$ qui satisfait

$$\mathbb{E}[XZ] = \mathbb{E}[UZ],$$

pour toute variable aléatoire bornée Z qui est $\mathcal{B}$-mesurable.

On définit alors l'espérance conditionnelle d'une variable aléatoire X sachant une autre variable aléatoire Y comme l'espérance conditionnelle de X sachant $\sigma(Y)$, que l'on notera souvent $\mathbb{E}[X \mid Y]$.

Citons un cas particulier. Supposons que Y prenne ses valeurs dans un ensemble dénombrable, c'est-à-dire, on suppose qu'il existe un ensemble dénombrble $\mathcal{Y}$ tel que $\mathbb{P}(Y = y) > 0$ pour tout $y \in \mathcal{Y}$ et $\sum_{y\in\mathcal{Y}} \mathbb{P}(Y = y) = 1$. Alors

$$\mathbb{E}[X \mid Y] = \sum_{y\in\mathcal{Y}} \mathbb{E}[X \mid \{Y = y\}] \mathbf{1}_{\{Y=y\}} .$$

Proposition A.8 (*Propriétés de l'espérance conditionnelle*)

Soit $(\Omega, \mathcal{F}, \mathbb{P})$ un espace de probabilité, $\mathcal{B} \subseteq \mathcal{F}$ une sous-tribu et $X \in L^1(\Omega, \mathcal{F}, \mathbb{P})$. Les propriétés suivantes sont vérifiées :

(a) $\forall a, b \in \mathbb{R}$ et pour toute v.a. $Y \in L^1(\Omega, \mathcal{F}, \mathbb{P})$, on a $\mathbb{E}[aX + bY \mid \mathcal{B}] = a\mathbb{E}[X \mid \mathcal{B}] + b\mathbb{E}[Y \mid \mathcal{B}]$.

(b) *Si* $X \geq 0$*, alors* $\mathbb{E}[X \mid \mathcal{B}] \geq 0$.

(c) *Si* a *est une constante, alors* $\mathbb{E}[a \mid \mathcal{B}] = a$.

(d) *Si* $\mathcal{B} = \{\emptyset, \Omega\}$*, alors* $\mathbb{E}[X \mid \mathcal{B}] = \mathbb{E}[X]$ *et* $\mathbb{E}[X \mid \mathcal{F}] = X$.

(e) *Soit* Z *une variable* $\mathcal{B}$*-mesurable telle que* $XZ \in L^1$*. Alors* $\mathbb{E}[XZ \mid \mathcal{B}] = Z\mathbb{E}[X \mid \mathcal{B}]$*. En particulier si* X *est* $\mathcal{B}$*-mesurable,* $\mathbb{E}[X \mid \mathcal{B}] = X$.

(f) *Si* $\mathcal{C} \subseteq \mathcal{B}$*, alors* $\mathbb{E}\left[\mathbb{E}[X \mid \mathcal{B}] \mid \mathcal{C}\right] = \mathbb{E}[X \mid \mathcal{C}]$*. En particulier,* $\mathbb{E}\left[\mathbb{E}[X \mid \mathcal{B}]\right] = \mathbb{E}[X]$.

(g) *Si* $\sigma(X)$ *et* $\mathcal{B}$ *sont indépendantes, alors* $\mathbb{E}[X \mid \mathcal{B}] = \mathbb{E}[X]$*, et plus généralement, pour toute* Y $\mathcal{B}$*-mesurable*

$$\mathbb{E}[g(X,Y) \mid \mathcal{B}] = \int g(x,Y)\,\mathcal{L}_X(dx),$$

pour toute fonction mesurable $g : \mathbb{R}^2 \to \mathbb{R}$ *telle que* $g(X,Y)$ *est intégrable.*

(h) *(Jensen) Si* $\phi : \mathbb{R} \to \mathbb{R}_+$ *est convexe et* $X \in L^1(\Omega, \mathcal{F}, \mathbb{P})$*, alors*

$$\phi(\mathbb{E}[X \mid \mathcal{B}]) \leq \mathbb{E}[\phi(X) \mid \mathcal{B}] \,.$$

Annexe B

Uniforme intégrabilité

Commençons par rappeler que toute variable aléatoire intégrable Y sur $(\Omega, \mathcal{F}, \mathbb{P})$ satisfait (de par le théorème de convergence dominée)

$$\lim_{a\to\infty} \mathbb{E}[|Y|\mathbf{1}_{|Y|>a}] = 0 \ . \tag{B.1}$$

Plus généralement :

Lemme B.1 (*Uniforme continuité de l'intégrale*)

Pour tout $\varepsilon > 0$, il existe $\delta > 0$ tel que pour tout $B \in \mathcal{F}$

$$\mathbb{P}(B) < \delta \Rightarrow \mathbb{E}[|Y|\mathbf{1}_B] < \varepsilon \ .$$

Démonstration Il suffit d'écrire

$$\begin{aligned}\mathbb{E}[|Y|\mathbf{1}_B] &= \mathbb{E}[|Y|\mathbf{1}_B\mathbf{1}_{|Y|\leq a}] + \mathbb{E}[|Y|\mathbf{1}_B\mathbf{1}_{|Y|>a}] \\ &\leq a\mathbb{P}(B) + \mathbb{E}[|Y|\mathbf{1}_{|Y|>a}] \ .\end{aligned}$$

Il suffit alors de choisir a tel que $\mathbb{E}[|Y|\mathbf{1}_{|Y|>a}] < \varepsilon/2$ puis $\delta = \varepsilon/2a$. □

Nous allons généraliser ces idées à des collections de v.a.

Soit $(X_i, i \in I)$ une collection (pas forcément dénombrable) de variables aléatoires à valeurs dans $\mathbb{R}$ (ou $\mathbb{R}^d$).

Définition B.2

On dit que la famille $(X_i, i \in I)$ est uniformément intégrable (u.i.) si

$$\lim_{a\to\infty} \sup_{i\in I} \mathbb{E}[|X_i|\mathbf{1}_{|X_i|>a}] = 0 \ .$$

La convergence (B.1) montre que toute variable aléatoire intégrable Y forme une famille (à un élément) qui est u.i. Plus généralement, il est facile de vérifier que toute collection finie de v.a. intégrables est uniformément intégrable.

Donnons à présent une caractérisation de l'uniforme intégrabilité, qui se rapproche de la continuité uniforme de l'intégrale.

Proposition B.3

La famille $(X_i, i \in I)$ est uniformément intégrable si et seulement si

1. $\sup_{i\in I} \mathbb{E}[|X_i|] < \infty$
2. *pour tout* $\varepsilon > 0$, *il existe* $\delta > 0$ *tel que pour tout* $B \in \mathcal{F}$

$$\mathbb{P}(B) < \delta \Rightarrow \sup_{i\in I} \mathbb{E}[|X_i|\mathbf{1}_B] < \varepsilon \ .$$

Démonstration Supposons que la famille est u.i. On a pour tout $B \in \mathcal{F}$

$$\mathbb{E}[|X_i|\mathbf{1}_B] = \mathbb{E}[|X_i|\mathbf{1}_B\mathbf{1}_{|X_i|\le a}] + \mathbb{E}[|X_i|\mathbf{1}_B\mathbf{1}_{|X_i|>a}] \le a\mathbb{P}(B) + \mathbb{E}[|X_i|\mathbf{1}_{|X_i|>a}] \ .$$

Si l'on prend $B = \Omega$, on trouve 1. Si l'on choisit a tel que

$$\sup_{i\in I} \mathbb{E}[|X_i|\mathbf{1}_{|X_i|>a}] < \varepsilon/2 \ ,$$

ainsi que $\delta = \varepsilon/2a$ on trouve 2.
Réciproquement on suppose que 1. et 2. sont vérifiées. Soit $\varepsilon > 0$ fixé, on souhaite montrer que pour tout a assez grand on a

$$\sup_{i\in I} \mathbb{E}[|X_i|\mathbf{1}_{|X_i|>a}] < \varepsilon \ .$$

On pose $M := \sup_{i\in I} \mathbb{E}[|X_i|]$, et $a = M/\delta$. Ainsi pour tout $i \in I$

$$\mathbb{P}(|X_i| > a) \le M/a \le \delta \ ,$$

et 2. assure que

$$\mathbb{E}[|X_i|\mathbf{1}_{|X_i|>a}] \le \varepsilon \ .$$

□

Énonçons à présent deux critères assurant l'uniforme intégrabilité.

Proposition B.4

S'il existe Y dans L^1 telle que pour tout $i \in I$, p.s. $|X_i| \le |Y|$ alors la famille $(X_i, i \in I)$ est uniformément intégrable.

Démonstration Il suffit de vérifier les points 1. et 2. de la proposition précédente. Le point 1. est clair. Concernant le point 2., on sait que

$$\sup_{i\in I} \mathbb{E}[|X_i|\mathbf{1}_B] \le \mathbb{E}[|Y|\mathbf{1}_B] \ ,$$

et l'on peut conclure en utilisant l'uniforme continuité de l'intégrale (de Y). □

Proposition B.5

Soit $g : \mathbb{R}_+ \to \mathbb{R}_+$ une fonction telle que

$$\lim_{t\to\infty} \frac{g(t)}{t} = +\infty .$$

Si $\sup_{i\in I} \mathbb{E}[g(|X_i|)] < \infty$ alors la famille $(X_i, i \in I)$ est uniformément intégrable. En particulier, si $(X_i, i \in I)$ est bornée dans L^p pour un certain $p > 1$ alors elle est uniformément intégrable.

Démonstration On note $C_a := \inf_{t\geq a} g(t)/t$. Par hypothèse $C_a \to +\infty$ quand $a \to \infty$. On pose $M := \sup_{i\in I} \mathbb{E}[g(|X_i|)]$ et l'on obtient pour tout $a > 0$

$$\sup_{i\in I} \mathbb{E}[|X_i|\mathbf{1}_{|X_i|>a}] \leq \sup_{i\in I} \frac{1}{C_a}\mathbb{E}[g(|X_i|)\mathbf{1}_{|X_i|>a}] \leq \frac{M}{C_a} .$$

Il suffit alors de passer à la limite $a \to \infty$. □

Théorème B.6 (*Théorème de convergence dominée optimal*)

Soit $(X_n)_n$ une suite de v.a. intégrables qui converge en probabilité vers une variable aléatoire X. Il y a équivalence entre :

1. *$(X_n, n \geq 1)$ est uniformément intégrable,*
2. *X_n converge vers X dans L^1.*

Démonstration On commence par l'implication 1. ⇒ 2. La convergence en probabilité assure qu'il existe une sous-suite n_k le long de laquelle X_{n_k} converge p.s. vers X. Par Fatou, on trouve

$$\mathbb{E}[|X|] \leq \liminf_{k\to\infty} \mathbb{E}[|X_{n_k}|] \leq \sup_n \mathbb{E}[|X_n|] .$$

Par la Proposition B.3, on sait que l'u.i. assure que $(X_n)_n$ est bornée dans L^1, et ainsi $X \in L^1$. On fixe $\varepsilon > 0$. On écrit

$$\begin{aligned}\mathbb{E}[|X_n - X|] &= \mathbb{E}[|X_n - X|\mathbf{1}_{\{|X_n-X|\leq\varepsilon\}}] + \mathbb{E}[|X_n - X|\mathbf{1}_{\{|X_n-X|>\varepsilon\}}] \\ &\leq \varepsilon + \mathbb{E}[|X_n|\mathbf{1}_{\{|X_n-X|>\varepsilon\}}] + \mathbb{E}[|X|\mathbf{1}_{\{|X_n-X|>\varepsilon\}}]\end{aligned}$$

La convergence en probabilité assure que $\mathbb{P}(|X_n - X| > \varepsilon) \to 0$. Par la Proposition B.3, on sait alors que pour tout n assez grand

$$\mathbb{E}[|X_n|\mathbf{1}_{\{|X_n-X|>\varepsilon\}}] \leq \varepsilon .$$

Par ailleurs, l'uniforme continuité de l'intégrale assure que pour tout n assez grand

$$\mathbb{E}[|X|\mathbf{1}_{\{|X_n-X|>\varepsilon\}}] \leq \varepsilon .$$

Montrons à présent le sens réciproque, en utilisant la définition équivalente de l'u.i. fournie par la Proposition B.3. Tout d'abord la convergence dans L^1 assure que

$$\sup_{n\geq 1} \mathbb{E}[|X_n|] < \infty .$$

On fixe $\epsilon > 0$. Il existe $n_0 \geq 1$ tel que pour tout $n \geq n_0$ $\mathbb{E}[|X_n - X|] < \epsilon/2$. On écrit alors pour tout événement $B \in \mathcal{F}$

$$\mathbb{E}[|X_n|\mathbf{1}_B] \leq \mathbb{E}[\big(|X_n - X| + |X|\big)\mathbf{1}_B] \leq \mathbb{E}[|X_n - X|\mathbf{1}_B] + \mathbb{E}[|X|\mathbf{1}_B] \; .$$

Par l'uniforme continuité de l'intégrale, il existe $\delta_1 > 0$ tel que pour tout B tel que si $\mathbb{P}(B) < \delta_1$ alors $\mathbb{E}[|X|\mathbf{1}_B] < \epsilon/2$. Par ailleurs, la collection finie de v.a. $X_n - X$, $n \leq n_0$ est u.i. Ainsi il existe $\delta_2 > 0$ tel que si $\mathbb{P}(B) < \delta_1$ alors $\sup_{n \leq n_0} \mathbb{E}[|X_n - X|\mathbf{1}_B] < \epsilon/2$. Enfin pour tout $n \geq n_0$, $\mathbb{E}[|X_n - X|\mathbf{1}_B] \leq \mathbb{E}[|X_n - X|] \leq \epsilon/2$. Ainsi si $\mathbb{P}(B) \leq \delta_1 \wedge \delta_2$ on obtient

$$\sup_{n \geq 1} \mathbb{E}[|X_n|\mathbf{1}_B] \leq \epsilon \; .$$

□

On conclut avec un résultat d'uniforme intégrabilité qui sera utilisé à plusieurs reprises dans le cours.

Proposition B.7

Soit X une variable aléatoire intégrable. Soit $(\mathcal{F}_i)_{i \in I}$ une collection de sous-tribus de $\mathcal{F}$ et soit $X_i := \mathbb{E}[X \mid \mathcal{F}_i]$ pour tout $i \in I$. Alors la famille $(X_i, i \in I)$ est uniformément intégrable.

Démonstration Par l'inégalité de Jensen pour l'espérance conditionnelle

$$|X_i| \leq \mathbb{E}[|X| \mid \mathcal{F}_i] \; ,$$

et ainsi

$$\sup_i \mathbb{E}[|X_i|] \leq \sup_i \mathbb{E}[\mathbb{E}[|X| \mid \mathcal{F}_i]] = \mathbb{E}[|X|] < \infty \; .$$

Ainsi $(X_i, i \in I)$ est bornée dans L^1. Par ailleurs l'événement $\{|X_i| > a\}$ est $\mathcal{F}_i$ mesurable et ainsi

$$\begin{aligned}
\mathbb{E}[|X_i|\mathbf{1}_{\{|X_i|>a\}}] &= \mathbb{E}[|\mathbb{E}[X \mid \mathcal{F}_i]|\mathbf{1}_{\{|X_i|>a\}}] \\
&= \mathbb{E}[|\mathbb{E}[X\mathbf{1}_{\{|X_i|>a\}} \mid \mathcal{F}_i]|] \\
&\leq \mathbb{E}[\mathbb{E}[|X|\mathbf{1}_{\{|X_i|>a\}} \mid \mathcal{F}_i]] \\
&\leq \mathbb{E}[|X|\mathbf{1}_{\{|X_i|>a\}}] \; .
\end{aligned}$$

Fixons $\varepsilon > 0$. L'uniforme continuité de l'intégrale assure qu'il existe $\delta > 0$ tel que

$$\mathbb{P}(A) < \delta \Rightarrow \mathbb{E}[|X|\mathbf{1}_A] < \varepsilon \; .$$

Comme

$$\sup_i \mathbb{P}(|X_i| > a) \leq \frac{1}{a} \sup_i \mathbb{E}[|X_i|] \leq \frac{1}{a}\mathbb{E}[|X|] \; .$$

Il existe $a > 0$ tel que

$$\sup_i \mathbb{P}(|X_i| > a) < \delta \; ,$$

et ainsi

$$\sup_i \mathbb{E}[|X_i|\mathbf{1}_{\{|X_i|>a\}}] \leq \varepsilon \ .$$

On en déduit l'u.i. de $(X_i, i \in I)$.

$\square$

Par exemple, soit X une variable aléatoire intégrable et $(\mathcal{F}_i)$ une filtration. Alors, la famille $\{\mathbb{E}[X \mid \mathcal{F}_T] : T \text{ est un temps d'arrêt}\}$ est uniformément intégrable.

Annexe C

Martingales à temps discret

Soit $(\Omega, \mathcal{F}, (\mathcal{F}_n)_{n\in\mathbb{N}}, \mathbb{P})$ un espace de probabilité filtré, c'est-à-dire, un espace de probabilité $(\Omega, \mathcal{F}, \mathbb{P})$ muni d'une collection $(\mathcal{F}_n)_{n\in\mathbb{N}}$ de sous-tribus de $\mathcal{F}$ satisfaisant

$$\mathcal{F}_0 \subset \mathcal{F}_1 \subset \ldots .$$

Une telle collection est appelée *filtration*.

Définition C.1

Soit $M = (M_n)_{n\in\mathbb{N}}$ une collection de variables aléatoires définies sur $(\Omega, \mathcal{F}, \mathbb{P})$ et à valeurs réelles. On dit que M est une martingale (resp. sur-martingale, resp. sous-martingale) si :

1. *M est adapté à la filtration $(\mathcal{F}_n)_n$, c'est-à-dire, pour tout $n \in \mathbb{N}$, M_n est $\mathcal{F}_n$-mesurable,*
2. *pour tout $n \in \mathbb{N}$, M_n est intégrable,*
3. *pour tout $n \in \mathbb{N}$, $\mathbb{E}[M_{n+1} \mid \mathcal{F}_n] = M_n$ p.s. (resp. $\mathbb{E}[M_{n+1} \mid \mathcal{F}_n] \leq M_n$ p.s., resp. $\mathbb{E}[M_{n+1} \mid \mathcal{F}_n] \geq M_n$ p.s.).*

Quelques exemples : la marche aléatoire simple est une martingale ; la marche aléatoire biaisée (proba p d'augmenter de 1 et $1-p$ de diminuer de 1) est une sous-martingale lorsque $p \geq 1/2$ et une sur-martingale lorsque $p \leq 1/2$; pour toute collection de variables aléatoires indépendantes intégrables $(X_k)_{k\geq 0}$ le processus $M_n = X_0 + \ldots + X_n$ est une martingale dans la filtration $\mathcal{F}_n := \sigma(X_0, \ldots, X_n)$ pour peu que les $(X_k)_{k\geq 0}$ soient d'espérance nulle.

Exercice C.2

Soit X un processus adapté à $(\mathcal{F}_n)_n$, et soit $(H_n)_{n\geq 1}$ un processus prévisible, c'est-à-dire, pour tout $n \geq 1$, H_n est $\mathcal{F}_{n-1}$-mesurable. On suppose également que chaque variable aléatoire H_n est bornée. On pose

$$Y_0 := 0 , \quad Y_n := \sum_{k=1}^{n} H_k (X_k - X_{k-1}) .$$

— *Montrer que si X est une martingale alors H est également une martingale.*

— *Montrer que si X est une sous-martingale (resp. sur-martingale) et si $H_n \geq 0$ pour tout $n \geq 1$, alors Y est une sous-martingale (resp. sur-martingale).*

Exercice C.3 (♠)

Soit X une sous-martingale. Montrer qu'il existe une martingale M et un processus A croissant, prévisible et vérifiant $A_0 = 0$ tel que $X = M + A$. Montrer que cette décomposition est unique. (Indice : considérer $\mathbb{E}[X_{n+1} - X_n \mid \mathcal{F}_n]$).

Commençons par rappeler quelques propriétés simples.

Lemme C.4

Soient M une martingale (resp. une sous-martingale) et $f : \mathbb{R} \to \mathbb{R}$ convexe (resp. convexe croissante). On suppose que pour tout $n \in \mathbb{N}$, $f(M_n)$ est intégrable. Alors $f(M)$ est une sous-martingale.

Démonstration En appliquant l'inégalité de Jensen pour l'espérance conditionnelle, on obtient que pour tout $n \geq 0$ presque sûrement

$$f(\mathbb{E}[M_{n+1} \mid \mathcal{F}_n]) \leq \mathbb{E}[f(M_{n+1}) \mid \mathcal{F}_n] \;.$$

Si M est une martingale alors le terme de gauche coïncide avec $f(M_n)$ et l'on conclut. Si M est une sous-martingale, la croissance de f assure que le terme de gauche est supérieur ou égal à $f(M_n)$ et l'on conclut. □

On rappelle qu'une variable aléatoire $T : \Omega \to \mathbb{N} \cup \{+\infty\}$ est un temps d'arrêt si pour tout $n \in \mathbb{N}$, $\{T \leq n\} \in \mathcal{F}_n$. On notera que si T est un temps d'arrêt, alors $\{T > n\} = \{T \leq n\}^\complement \in \mathcal{F}_n$ et $\{T = n\} = \{T \leq n\} \cap \{T \leq n-1\}^\complement \in \mathcal{F}_n$.

Exercice C.5

Montrer que T est un temps d'arrêt si et seulement si pour tout $n \in \mathbb{N}$, $\{T = n\} \in \mathcal{F}_n$.

Exercice C.6

Soit $X = (X_n)_{n\in\mathbb{N}}$ un processus adapté et A un Borélien de $\mathbb{R}$. Montrer que le premier temps d'entrée dans A

$$T_A := \inf\{n \in \mathbb{N} : X_n \in A\} \;,$$

est un temps d'arrêt.

Exercice C.7

Soit $X = (X_n)_{n\in\mathbb{N}}$ un processus adapté, A un Borélien de $\mathbb{R}$ et $N \geq 1$ un entier. Montrer que le dernier temps avant N où X est dans A

$$L_A := \sup\{n \leq N : X_n \in A\} \;,$$

n'est pas un temps d'arrêt en général.

Lemme C.8

Soient M une martingale (resp. sous-martingale, resp. sur-martingale) et T un temps d'arrêt. On pose $M_n^T := M_{T\wedge n}$ pour tout $n \in \mathbb{N}$. Alors M^T est une martingale (resp. sous-martingale, resp. sur-martingale).

Démonstration Tout d'abord $|M_n^T| \le |M_0| + \ldots + |M_n|$ et ainsi M_n^T est intégrable. Par ailleurs, chaque variable $M_k \mathbf{1}_{T=k}$ est $\mathcal{F}_k$-mesurable et ainsi la somme

$$M_n^T = \sum_{k=0}^{n} M_k \mathbf{1}_{T=k} + M_n \mathbf{1}_{T>n} \ ,$$

est $\mathcal{F}_n$-mesurable. Enfin on observe que

$$M_{n+1}^T = \sum_{k=0}^{n} M_k \mathbf{1}_{T=k} + M_{n+1} \mathbf{1}_{T>n} \ ,$$

et l'on calcule

$$\begin{aligned}
\mathbb{E}[M_{n+1}^T \mid \mathcal{F}_n] &= \sum_{k=0}^{n} \mathbb{E}[M_k \mathbf{1}_{T=k} \mid \mathcal{F}_n] + \mathbb{E}[M_{n+1} \mathbf{1}_{T>n} \mid \mathcal{F}_n] \\
&= \sum_{k=0}^{n} M_k \mathbf{1}_{T=k} + \mathbf{1}_{T>n} \mathbb{E}[M_{n+1} \mid \mathcal{F}_n] \\
&= \sum_{k=0}^{n} M_k \mathbf{1}_{T=k} + M_n \mathbf{1}_{T>n} \\
&= M_n^T \ .
\end{aligned}$$

□

Les martingales apparaissent dans de nombreux contextes, il est donc utile d'établir des résultats généraux les concernant. Nous allons aborder trois thèmes :

1. les théorèmes d'arrêt, qui permettent d'évaluer $\mathbb{E}[M_T]$ où T est un temps d'arrêt,
2. les inégalités maximales, qui fournissent des bornes sur la queue de probabilité du supremum $\sup_{0\le k\le n} M_n$ d'une martingale (resp. sur-martingale, resp. sous-martingale),
3. les théorèmes de convergence, qui fournissent des informations sur le comportement asymptotique d'une martingale (resp. sur-martingale, resp. sous-martingale).

1 Théorème d'arrêt borné

Etant donné un temps d'arrêt T, on introduit la tribu du passé avant T :

$$\mathcal{F}_T := \{A \in \mathcal{F} : A \cap \{T \le n\} \in \mathcal{F}_n \ , \quad \forall n \in \mathbb{N}\} \ .$$

On peut vérifier que $A \in \mathcal{F}_T$ si et seulement si A s'écrit

$$A = \bigcup_n \Big(B_n \cap \{T = n\} \Big) ,$$

où $B_n \in \mathcal{F}_n$. Cela justifie l'appellation "tribu du passé" car les événements ne peuvent pas s'appuyer sur une information postérieure à T.

Exercice C.9

Soit X un processus adapté et T un temps d'arrêt. On définit $\mathbf{1}_{\{T<\infty\}} X_T$ comme la variable aléatoire égale à $X_{T(\omega)}$ lorsque $T(\omega) < \infty$ et égale à 0 lorsque $T(\omega) = +\infty$. Montrer que $\mathbf{1}_{\{T<\infty\}} X_T$ est $\mathcal{F}_T$-mesurable.

Théorème C.10 (*Théorème d'arrêt borné*)

Soit M une martingale (resp. une sous-martingale, resp. une sur-martingale). Soient $S \leq T$ deux temps d'arrêt bornés, *c'est-à-dire, tels qu'il existe $m \geq 0$ tel que $0 \leq S \leq T \leq m$ p.s. Alors M_S, M_T sont intégrables et presque sûrement*

$$\mathbb{E}[M_T \mid \mathcal{F}_S] = M_S , \quad \text{resp. } \geq M_S , \quad \text{resp. } \leq M_S .$$

Démonstration On a $|M_T| \leq |M_0| + \ldots + |M_m|$ p.s., et ainsi M_T est intégrable. De même pour M_S. Par ailleurs pour tout $A \in \mathcal{F}_S$

$$\mathbb{E}[M_T \mathbf{1}_A] = \mathbb{E}[M_{T\wedge m} \mathbf{1}_A] = \sum_{k=0}^{m} \mathbb{E}[M_{T\wedge m} \mathbf{1}_{A\cap\{S=k\}}] .$$

Or $(M_{T\wedge n}, n \in \mathbb{N})$ est une martingale et $A \cap \{S = k\}$ est $\mathcal{F}_k$ mesurable de sorte que

$$\mathbb{E}[M_{T\wedge m} \mathbf{1}_{A\cap\{S=k\}}] = \mathbb{E}[M_{T\wedge k} \mathbf{1}_{A\cap\{S=k\}}] .$$

Ainsi

$$\mathbb{E}[M_T \mathbf{1}_A] = \sum_{k=0}^{m} \mathbb{E}[M_{T\wedge k} \mathbf{1}_{A\cap\{S=k\}}] = \sum_{k=0}^{m} \mathbb{E}[M_S \mathbf{1}_{A\cap\{S=k\}}] = \mathbb{E}[M_S \mathbf{1}_A] .$$

□

Nous verrons plus loin des conditions sur la martingale M assurant que cette égalité reste vraie pour tous temps d'arrêt $S \leq T$.

2 Inégalités maximales

Proposition C.11

— *Si M est une sous-martingale alors pour tout $a > 0$*

$$a\mathbb{P}(\sup_{0\le k\le n} M_k > a) \le \mathbb{E}[M_n \mathbf{1}_{\sup_{0\le k\le n} M_k > a}] \le \mathbb{E}[M_n^+] ,$$

— *Si M est une sur-martingale alors pour tout $a > 0$*

$$a\mathbb{P}(\sup_{0\le k\le n} M_k > a) \le \mathbb{E}[M_0] + \mathbb{E}[M_n^-] .$$

Démonstration On fixe $n \in \mathbb{N}$ et l'on note $T := \inf\{k \in \mathbb{N} : M_k > a\} \wedge n$. On note que l'on a $\{\sup_{0\le k\le n} M_k \le a\} \subset \{T = n\}$ et $\{\sup_{0\le k\le n} M_k > a\} = \{M_T > a\}$. On calcule alors

$$\mathbb{E}[M_T] = \mathbb{E}[M_T \mathbf{1}_{\sup_{0\le k\le n} M_k > a}] + \mathbb{E}[M_T \mathbf{1}_{\sup_{0\le k\le n} M_k \le a}] \ge a\mathbb{P}(\sup_{0\le k\le n} M_k > a) + \mathbb{E}[M_n \mathbf{1}_{\sup_{0\le k\le n} M_k \le a}] .$$

— Si M est une sous-martingale alors le théorème d'arrêt borné appliqué aux temps d'arrêt $T \le n$ assure que

$$\mathbb{E}[M_n] \ge \mathbb{E}[M_T] \ge a\mathbb{P}(\sup_{0\le k\le n} M_k > a) + \mathbb{E}[M_n \mathbf{1}_{\sup_{0\le k\le n} M_k \le a}] ,$$

et ainsi

$$a\mathbb{P}(\sup_{0\le k\le n} M_k > a) \le \mathbb{E}[M_n \mathbf{1}_{\sup_{0\le k\le n} M_k > a}] \le \mathbb{E}[M_n^+] .$$

— Si M est une sur-martingale alors le théorème d'arrêt borné appliqué aux temps d'arrêt $0 \le T$ assure que

$$\mathbb{E}[M_0] \ge \mathbb{E}[M_T] \ge a\mathbb{P}(\sup_{0\le k\le n} M_k > a) + \mathbb{E}[M_n \mathbf{1}_{\sup_{0\le k\le n} M_k \le a}] ,$$

et ainsi

$$a\mathbb{P}(\sup_{0\le k\le n} M_k > a) \le \mathbb{E}[M_0] - \mathbb{E}[M_n \mathbf{1}_{\sup_{0\le k\le n} M_k \le a}] \le \mathbb{E}[M_0] + \mathbb{E}[M_n^-] .$$

□

Proposition C.12 (*Inégalité maximale de Doob*)

Soit M une martingale. Posons $M_n^* := \sup_{0\le k\le n} |M_k|$ pour tout $n \in \mathbb{N}$. Pour tout $p > 1$ et tout $n \in \mathbb{N}$

$$\mathbb{E}[(M_n^*)^p] \le \Big(\frac{p}{p-1}\Big)^p \mathbb{E}[|M_n|^p] .$$

Démonstration On va prouver une propriété plus générale. Soit X une sous-martingale positive et soit $\tilde{X}_n := \sup_{0\le k\le n} X_k$. Pour tout $p > 1$ et tout $n \in \mathbb{N}$

$$\mathbb{E}[(\tilde{X}_n)^p] \le \Big(\frac{p}{p-1}\Big)^p \mathbb{E}[X_n^p] .$$

Une fois cette propriété établie, il suffit de prendre $X_n = |M_n|$ pour déduire le résultat.
On peut supposer que $\mathbb{E}[X_n^p] < \infty$ car sinon il n'y a rien à montrer. Comme $x \mapsto x^p$ est convexe croissante sur $\mathbb{R}_+$, X^p est toujours une sous-martingale positive et on en déduit que $\mathbb{E}[X_k^p] < \infty$ pour tout $0 \le k \le n$, et ainsi

$$\mathbb{E}[(\tilde{X}_n)^p] \le \mathbb{E}[X_0^p + X_1^p + \ldots + X_n^p] \le \mathbb{E}[X_0^p] + \ldots + \mathbb{E}[X_n^p] < \infty \; .$$

L'inégalité maximale précédemment prouvée montre que

$$a\mathbb{P}(\sup_{0\le k\le n} X_k > a) \le \mathbb{E}[X_n \mathbf{1}_{\sup_{0\le k\le n} X_k > a}] \; .$$

En multipliant chaque membre de cette inégalité par a^{p-2} et en intégrant par rapport à la mesure de Lebesgue sur $[0,\infty[$ on obtient

$$\int_0^\infty a^{p-1}\mathbb{P}(\sup_{0\le k\le n} X_k > a)da = \frac{1}{p}\mathbb{E}[(\tilde{X}_n)^p] \; ,$$

ainsi que

$$\int_0^\infty a^{p-2}\mathbb{E}[X_n \mathbf{1}_{\sup_{0\le k\le n} X_k > a}]da = \frac{1}{p-1}\mathbb{E}[X_n(\tilde{X}_n)^{p-1}] \le \frac{1}{p-1}\mathbb{E}[(X_n)^p]^{1/p}\mathbb{E}[(\tilde{X}_n)^p]^{\frac{p-1}{p}} \; .$$

Ainsi

$$\frac{1}{p}\mathbb{E}[(\tilde{X}_n)^p] \le \frac{1}{p-1}\mathbb{E}[(X_n)^p]^{1/p}\mathbb{E}[(\tilde{X}_n)^p]^{\frac{p-1}{p}} \; ,$$

et l'inégalité voulue en découle. □

3 Convergence de martingales

La preuve du résultat suivant est assez longue, nous ne la rappellerons pas et nous renvoyons vers [LG06, Th 12.3.3].

Théorème C.13

Soit M une martingale (resp. sous-martingale, resp. sur-martingale) bornée dans L^1, c'est-à-dire, telle que $\sup_{n\in\mathbb{N}} \mathbb{E}[|M_n|] < \infty$. Alors M_n converge presque sûrement vers une variable aléatoire intégrable M_∞.

On notera qu'il suffit de prouver le résultat pour une sous-martingale : en effet, si M est une sur-martingale bornée dans L^1 alors $-M$ est une sous-martingale bornée dans L^1.

Il se trouve qu'en général cette convergence n'a pas lieu dans L^1. Considérons l'exemple suivant. Soit S la marche aléatoire simple issue de 1 et $T := \inf\{n \in \mathbb{N} : S_n = 0\}$ son premier temps d'atteinte de 0. On notera que $T < \infty$ presque sûrement mais que T n'est pas borné. Par un résultat précédent, le processus $M_n = S_{T\wedge n}$ est une martingale et l'on a

$$\mathbb{E}[M_n] = \mathbb{E}[M_0] = 1 \; .$$

Etant positive, la martingale arrêtée M est alors bornée dans L^1. On en déduit qu'elle converge p.s. vers une variable aléatoire intégrable M_∞. Or pour tout $k \in \mathbb{N}$ et pour tout $n \geq k$

$$M_n \mathbf{1}_{T=k} = S_{T \wedge n} \mathbf{1}_{T=k} = 0 \; .$$

Ainsi $M_\infty = 0$ sur l'événement $\{T = k\}$, et comme $T < \infty$ p.s., on en déduit que $M_\infty = 0$ p.s.

Si M_n convergeait vers M_∞ dans L^1 alors $1 = \mathbb{E}[M_n] \to \mathbb{E}[M_\infty] = 0$.

4 Martingales fermées

L'objectif de cette section est double. Tout d'abord, on souhaiterait trouver des conditions sur une martingale sous lesquelles le théorème d'arrêt est encore vrai pour des temps d'arrêt quelconques (pas forcément bornés). Par ailleurs, on souhaiterait trouver des conditiones sous lesquelles la convergence p.s. du théorème précédent a également lieu dans L^1. Il se trouve que dans les deux cas, la notion de martingale fermée joue un rôle prépondérant.

Définition C.14

On dit qu'une martingale M est fermée s'il existe une v.a. $X \in L^1$ telle que pour tout $n \in \mathbb{N}$

$$M_n = \mathbb{E}[X \mid \mathcal{F}_n] \; .$$

On notera qu'une martingale M peut être fermée par deux v.a. X et X' distinctes.

Théorème C.15

Il y a équivalence entre :

(i) M est fermée

(ii) M est uniformément intégrable.

(iii) M converge p.s. et dans L^1 vers une variable aléatoire M_∞ intégrable.

Si l'une des trois conditions est vérifiée, alors nécessairement M est fermée par sa limite M_∞.

On notera que la condition de martingale fermée est nécessaire et suffisante pour que la convergence p.s. du théorème vu précédemment soit renforcée en une convergence L^1 !

Démonstration (i) $\Rightarrow$ (ii). C'est une conséquence de la Proposition B.7.

(ii) $\Rightarrow$ (iii). Etant u.i., la suite $(M_n, n \in \mathbb{N})$ est bornée dans L^1. Comme c'est en outre une martingale, elle converge p.s. vers une variable aléatoire intégrable M_∞. Le théorème de convergence dominée optimal, Théorème B.6, montre alors que M_n converge aussi dans L^1 vers M_∞.

(iii) $\Rightarrow$ (i). Pour tout $A \in \mathcal{F}_n$, on a

$$\mathbb{E}[M_n \mathbf{1}_A] = \mathbb{E}[M_{n+k} \mathbf{1}_A] \; .$$

Or $M_{n+k} \to M_\infty$ dans L^1 quand $k \to \infty$. Ainsi

$$\mathbb{E}[M_{n+k} \mathbf{1}_A] \to \mathbb{E}[M_\infty \mathbf{1}_A] \; ,$$

et l'on en déduit que

$$\mathbb{E}[M_n \mathbf{1}_A] = \mathbb{E}[M_\infty \mathbf{1}_A] .$$

On a donc montré que $\mathbb{E}[M_\infty \mid \mathcal{F}_n] = M_n$ p.s., pour tout $n \in \mathbb{N}$. □

Soit M une martingale fermée. On vient de voir qu'il existe alors une v.a. M_∞ telle que p.s. $M_n \to M_\infty$. Ainsi pour tout temps T, on peut définir

$$M_T(\omega) = M_{T(\omega)}(\omega) .$$

(En particulier, sur l'événement $\{T = \infty\}$ cette définition fait bien sens).

Théorème C.16

Soit M une martingale fermée. Pour tout temps d'arrêt T on a presque sûrement

$$M_T = \mathbb{E}[M_\infty \mid \mathcal{F}_T] .$$

Par ailleurs pour tous temps d'arrêt $S \leq T$ on a presque sûrement

$$M_S = \mathbb{E}[M_T \mid \mathcal{F}_S] .$$

Démonstration La deuxième propriété est une conséquence de la première. En effet

$$\mathbb{E}[M_T \mid \mathcal{F}_S] = \mathbb{E}[\mathbb{E}[M_\infty \mid \mathcal{F}_T] \mid \mathcal{F}_S] = \mathbb{E}[M_\infty \mid \mathcal{F}_S] = M_S .$$

La première propriété se prouve comme suit. Supposons que pour tout $n \in \mathbb{N} \cup \{\infty\}$

$$\mathbb{E}[M_\infty \mid \mathcal{F}_T]\mathbf{1}_{\{T=n\}} = \mathbb{E}[M_\infty \mid \mathcal{F}_n]\mathbf{1}_{\{T=n\}} . \tag{C.1}$$

Alors p.s.

$$\mathbb{E}[M_\infty \mid \mathcal{F}_T]\mathbf{1}_{\{T=n\}} = \mathbb{E}[M_\infty \mid \mathcal{F}_n]\mathbf{1}_{\{T=n\}} = M_n \mathbf{1}_{\{T=n\}} = M_T \mathbf{1}_{\{T=n\}} ,$$

et comme cela est vrai pour tout $n \in \mathbb{N} \cup \{\infty\}$, on en déduit que $M_T = \mathbb{E}[M_\infty \mid \mathcal{F}_T]$.
Il nous faut donc prouver l'identité ci-dessus. Pour cela, on commence par prouver que le produit d'une v.a. X $\mathcal{F}_T$ mesurable par $\mathbf{1}_{\{T=n\}}$ est $\mathcal{F}_n$ mesurable. En effet pour tout Borélien B ne contenant pas 0, du fait que $\{X \in B\} \in \mathcal{F}_T$ on a

$$\{X\mathbf{1}_{\{T=n\}} \in B\} = \{X \in B\} \cap \{T = n\} \in \mathcal{F}_n .$$

Par ailleurs pour tout Borélien B contenant 0 on a

$$\{X\mathbf{1}_{\{T=n\}} \in B\} = \{T \neq n\} \cup \Big(\{X \in B\} \cap \{T = n\}\Big) ,$$

qui est également $\mathcal{F}_n$ mesurable.
Les deux membres de (C.1) sont donc $\mathcal{F}_n$ mesurables. Nous allons montrer que leurs espérances contre $\mathbf{1}_A$ coïncident, pour tout $A \in \mathcal{F}_n$. Commençons par observer que $A \cap \{T = n\}$

est $\mathcal{F}_T$ mesurable : en effet, si $k < n$ alors $A \cap \{T = n\} \cap \{T \leq k\} = \emptyset$ et si $k \geq n$, $A \cap \{T = n\} \cap \{T \leq k\} = A \cap \{T = n\}$. Ainsi

$$\begin{aligned}\mathbb{E}[\mathbb{E}[M_\infty \mid \mathcal{F}_T]\mathbf{1}_{A\cap\{T=n\}}] &= \mathbb{E}[\mathbb{E}[M_\infty \mathbf{1}_{A\cap\{T=n\}} \mid \mathcal{F}_T]] \\ &= \mathbb{E}[M_\infty \mathbf{1}_{A\cap\{T=n\}}] \\ &= \mathbb{E}[\mathbb{E}[M_\infty \mathbf{1}_{A\cap\{T=n\}} \mid \mathcal{F}_n]] \\ &= \mathbb{E}[\mathbb{E}[M_\infty \mid \mathcal{F}_n]\mathbf{1}_{A\cap\{T=n\}}] \,.\end{aligned}$$

□

Bibliographie

[cn11] E. Çinlar. *Probability and stochastics*, vol. 261 of *Graduate Texts in Mathematics*. Springer, New York, 2011, xiv+557. doi:10.1007/978-0-387-87859-1.

[ELVE19] W. E, T. Li, and E. Vanden-Eijnden. *Applied stochastic analysis*, vol. 199 of *Graduate Studies in Mathematics*. American Mathematical Society, Providence, RI, 2019, xxi+305. doi:10.1090/gsm/199.

[Fou24] N. Fournier. Calcul stochastique et processu de diffusion, 2024. URL https://perso.lpsm.paris/~nfournier/PolyCS.pdf.

[KS91] I. Karatzas and S. E. Shreve. *Brownian motion and stochastic calculus*, vol. 113 of *Graduate Texts in Mathematics*. Springer-Verlag, New York, second ed., 1991, xxiv+470. doi:10.1007/978-1-4612-0949-2.

[LG06] J.-F. Le Gall. Intégration, probabilités et processus aléatoires, 2006. URL https://www.imo.universite-paris-saclay.fr/~jean-francois.le-gall/IPPA2.pdf.

[LG13] J.-F. Le Gall. *Mouvement brownien, martingales et calcul stochastique*, vol. 71 of *Mathématiques & Applications (Berlin) [Mathematics & Applications]*. Springer, Heidelberg, 2013, viii+176. doi:10.1007/978-3-642-31898-6.

[RY99] D. Revuz and M. Yor. *Continuous martingales and Brownian motion*, vol. 293 of *Grundlehren der mathematischen Wissenschaften [Fundamental Principles of Mathematical Sciences]*. Springer-Verlag, Berlin, third ed., 1999, xiv+602. doi:10.1007/978-3-662-06400-9.

[Sal22] J. Salez. Introduction to stochastic calculus, 2022. URL https://www.ceremade.dauphine.fr/~salez/stoc.pdf.