Département GMM
4ème année Spécialité MMS

Martingales et Applications

Aldéric Joulin

A. Joulin
Bureau 115 - GMM
ajoulin@insa-toulouse.fr

Table des matières

Chapitre 0

Rappels sur l'espérance conditionnelle

0.1 Quelques éléments sur les tribus

Avant de rentrer dans le coeur du sujet, faisons quelques rappels sur la notion de tribu.

Définition 0.1.1. *Soit Ω un ensemble et $\mathcal{A}$ une famille de parties de Ω. On dit que $\mathcal{A}$ est une tribu (sur Ω) si :*

(i) $\Omega \in \mathcal{A}$.

(ii) pour tout ensemble $A \in \mathcal{A}$, on a $A^c \in \mathcal{A}$, où A^c désigne le complémentaire de A dans Ω (stabilité par passage au complémentaire).

(iii) pour toute famille $(A_n)_{n\in\mathbb{N}}$ de Ω satisfaisant $A_n \in \mathcal{A}$ pour tout $n \in \mathbb{N}$, alors $\cup_{n\in\mathbb{N}} A_n \in \mathcal{A}$ (stabilité par union dénombrable).

En particulier, on appelle tribu engendrée par une classe de parties $\mathcal{C}$ de Ω la plus petite tribu sur Ω contenant $\mathcal{C}$, c'est-à-dire l'intersection de toutes les tribus contenant $\mathcal{C}$. On la note $\sigma(\mathcal{C})$. Un exemple classique est la tribu borélienne sur $\Omega = \mathbb{R}^d$, où $\mathcal{C}$ désigne l'ensemble des ouverts (ou fermés) de $\mathbb{R}^d$.

Dans la suite de ce cours, l'espace de probabilité sur lequel les variables aléatoires (v.a.) sont définies est noté $(\Omega, \mathcal{A}, \mathbb{P})$.

Définition 0.1.2. *Étant donnée une v.a. X à valeurs dans un espace mesurable $(E, \mathcal{E})$, on appelle tribu engendrée par X, et on la note $\sigma(X)$, la sous-tribu de $\mathcal{A}$ engendrée par l'ensemble des images réciproques de X. Autrement dit,*

$$\sigma(X) \;:=\; \sigma\left(\{X^{-1}(B) : B \in \mathcal{E}\}\right)$$

$$\begin{aligned} &= \sigma\left(\{\{\omega \in \Omega : X(\omega) \in B\}\,;B \in \mathcal{E}\}\right) \\ &= \sigma\left(\{\{X \in B\}\,;B \in \mathcal{E}\}\right), \end{aligned}$$

où la dernière égalité est simplement une notation. Il s'agit donc de la plus petite tribu sur Ω rendant X mesurable. De même, si $X_1, \ldots, X_n$ sont des v.a. à valeurs dans $(E, \mathcal{E})$, alors on définit la tribu engendrée par ces v.a. comme

$$\sigma\left(X_1, \ldots, X_n\right) := \sigma\left(\{\cap_{i=1}^n \{X_i \in B_i\}\,; B_1, \ldots, B_n \in \mathcal{E}\}\right).$$

À présent, nous allons énoncer un résultat très utile en pratique, le lemme de Doob, dû à un célèbre probabiliste américain du milieu de 20ème siècle. En particulier, ce lemme nous donne un critère simple pour établir la mesurabilité d'une v.a. réelle (v.a.r.) Y par rapport à la tribu engendrée par d'autres v.a.r. $X_1, \ldots, X_n$.

Lemme 0.1.3. *(Doob)*
Étant données des v.a.r. $X_1, \ldots, X_n$, une autre v.a.r. Y est $\sigma(X_1, \ldots, X_n)$-mesurable si et seulement s'il exite une fonction borélienne $h : \mathbb{R}^n \to \mathbb{R}$ telle que $Y = h(X_1, \ldots, X_n)$.

0.2 Espérance conditionnelle

Maintenant, nous sommes en mesure d'introduire la notion d'espérance conditionnelle. Dans la suite et sauf mention du contraire, toutes nos v.a. seront à valeurs réelles (donc des v.a.r.). Par ailleurs, étant donné un nombre $p \in [1, +\infty]$, nous noterons pour alléger la notation $L^p := L^p(\Omega, \mathcal{A}, \mathbb{P})$ et $L^p(\mathcal{F}) = L^p(\Omega, \mathcal{F}, \mathbb{P})$, où

$$L^p(\Omega, \mathcal{A}, \mathbb{P}) := \left\{ \text{ v.a. } X : \|X\|_{L^p} := \mathbb{E}\left[|X|^p\right]^{1/p} < +\infty \right\},$$

et $L^p(\Omega, \mathcal{F}, \mathbb{P})$ est le sous-ensemble de $L^p(\Omega, \mathcal{A}, \mathbb{P})$ constitué des v.a. $\mathcal{F}$-mesurables. Enfin, on rappelle qu'une propriété faisant intervenir des v.a. est vérifiée presque sûrement (on note p.s.) si elle est vérifiée pour tout ω en dehors d'un ensemble négligeable, i.e. de probabilité 0. Il s'agit simplement de la version probabiliste du "presque partout", une probabilité étant une mesure de masse totale égale à 1. Par exemple, on rencontre souvent cette notion lorsque l'on a une égalité entre deux v.a.r. (on parle de v.a. équivalentes), ou encore une inégalité, un résultat de convergence, etc...

Définition et théorème 0.2.1. *Soit X une v.a. appartenant à l'espace L^2 (respectivement à L^1) et soit $\mathcal{F}$ une sous-tribu quelconque de $\mathcal{A}$. Alors il existe une v.a. Y dans $L^2(\mathcal{F})$ (resp. dans $L^1(\mathcal{F})$), unique à équivalence près, telle que pour toute v.a. $Z \in L^2(\mathcal{F})$ (resp. tout événement $A \in \mathcal{F}$),*

$$\mathbb{E}\left[XZ\right] = \mathbb{E}\left[YZ\right] \quad (\textit{resp.} \quad \mathbb{E}\left[X 1_A\right] = \mathbb{E}\left[Y 1_A\right]).$$

On note $Y = \mathbb{E}\left[X \mid \mathcal{F}\right]$: c'est l'espérance conditionnelle de X sachant la tribu $\mathcal{F}$.

Démonstration. Nous n'allons faire que la démonstration dans $L^2(\mathcal{F})$, le cas $L^1(\mathcal{F})$ se traitant ensuite par densité et un passage à la limite. Posons $F = L^2(\mathcal{F})$, qui est un espace de Hilbert donc complet (i.e. toute suite de Cauchy converge pour la norme associée). De plus, F étant inclus dans L^2, il est fermé (tout sous-espace complet d'un espace métrique, non nécessairement complet, est fermé). Ainsi, par le théorème du supplémentaire orthogonal d'un fermé dans un espace de Hilbert, on a

$$L^2 = F \oplus F^\perp,$$

c'est-à-dire que l'espace L^2 se décompose comme la somme directe de F et de son orthogonal $F^\perp$ défini par

$$F^\perp := \{U \in L^2 : \mathbb{E}[U\,Z] = 0 \text{ pour tout } Z \in F\}.$$

En d'autres termes, toute v.a. de L^2 se décompose de manière unique comme la somme de 2 v.a., l'une dans F et l'autre dans $F^\perp$. Lorsque l'on applique ce résultat à X on obtient l'existence et l'unicité d'une v.a. $Y \in F$ telle que $X = Y + (X - Y)$ où $X - Y$ est dans $F^\perp$. Ainsi, on en déduit que pour tout $Z \in F$,

$$\mathbb{E}[(X - Y)\,Z] = 0,$$

c'est-à-dire le résultat désiré. ■

Cette définition généralise le cas de l'espérance conditionnelle définie par rapport à une tribu engendrée par une partition, que vous avez pu voir dans le cours de Compléments de Probabilités. Notons par ailleurs que nous imposons de l'intégrabilité dans la définition de l'espérance conditionnelle afin de s'assurer que cette dernière ait un sens, mais cette hypothèse peut bien évidemment être relaxée en supposant que X est seulement positive, auquel cas $\mathbb{E}[X \mid \mathcal{F}]$ peut prendre la valeur $+\infty$. De surcroît, on peut montrer que l'on ne change pas la notion d'espérance conditionnelle dans le cas intégrable en remplaçant les indicatrices 1_A par des v.a. $Z \in L^\infty(\mathcal{F})$, c'est-à-dire des v.a. $\mathcal{F}$-mesurables et bornées (cf. le cours de théorie de la mesure et l'approximation des fonctions mesurables bornées par des suites de fonctions étagées). On rappelle qu'une v.a. Z est dite bornée s'il existe un nombre (déterministe) $c \geq 0$ tel que $\mathbb{P}(|Z| \leq c) = 1$ (c'est le cas des v.a. de Bernoulli, binomiale, uniforme sur un intervalle, tandis que les v.a. de Poisson, géométrique, exponentielle, normale ne le sont pas).

Tout comme l'espérance et comme conséquence directe des résultats classiques issus de la théorie de la mesure, l'espérance conditionnelle satisfait les propriétés de linéarité, de positivité, de convergences monotone et dominée, mais aussi les inégalités de Jensen, de Cauchy-Schwarz et de Hölder.

Proposition 0.2.2. *L'espérance conditionnelle vérifie les propriétés suivantes.*

(i) *Si* $X \geq 0$ *p.s. alors* $\mathbb{E}[X \mid \mathcal{F}] \geq 0$ *p.s.*

(ii) *L'application* $X \to \mathbb{E}[X \mid \mathcal{F}]$ *est linéaire.*

(iii) *(Convergence monotone) Soit* $(X_n)_{n\in\mathbb{N}}$ *une suite croissante de v.a. positives, tendant p.s. vers une v.a.* X*. Alors on a la convergence p.s*

$$\lim_{n\to+\infty} \mathbb{E}[X_n \mid \mathcal{F}] = \mathbb{E}[X \mid \mathcal{F}].$$

(iv) *(Lemme de Fatou) Soit* $(X_n)_{n\in\mathbb{N}}$ *une suite de v.a. positives. Alors p.s.,*

$$\mathbb{E}[\liminf_{n\to+\infty} X_n \mid \mathcal{F}] \leq \liminf_{n\to+\infty} \mathbb{E}[X_n \mid \mathcal{F}].$$

(v) *(Convergence dominée) Soit* $(X_n)_{n\in\mathbb{N}}$ *une suite de v.a. tendant p.s. vers une v.a.* X *et telles que* $\sup_{n\in\mathbb{N}} |X_n| \leq V \in L^1$*. Alors p.s.,*

$$\lim_{n\to+\infty} \mathbb{E}[X_n \mid \mathcal{F}] = \mathbb{E}[X \mid \mathcal{F}].$$

(vi) *(Inégalité de Jensen) Si* $\varphi : \mathbb{R} \to \mathbb{R}$ *est une fonction convexe telle que* $\mathbb{E}[|\varphi(X)|] < +\infty$*, alors on a l'inégalité p.s.,*

$$\varphi\left(\mathbb{E}[X \mid \mathcal{F}]\right) \leq \mathbb{E}[\varphi(X) \mid \mathcal{F}].$$

(vii) *(Inégalité de Hölder) Soient* X, Y *deux v.a. respectivement dans* L^p *et* L^q*, où* p *et* q *vérifient* $1/p + 1/q = 1$ *avec* $p, q > 1$*. Alors on a l'inégalité p.s.,*

$$|\mathbb{E}[XY \mid \mathcal{F}]| \leq \left(\mathbb{E}[|X|^p \mid \mathcal{F}]\right)^{1/p} \left(\mathbb{E}[|Y|^q \mid \mathcal{F}]\right)^{1/q}.$$

Pour $p = q = 2$ *c'est la célèbre inégalité de Cauchy-Schwarz.*

En pratique, les propriétés que l'on utilise le plus souvent sont les suivantes.

Proposition 0.2.3. *Soit* X *une v.a. intégrable. Alors l'espérance conditionnelle vérifie les propriétés suivantes.*

(i) $\mathbb{E}[\mathbb{E}[X \mid \mathcal{F}]] = \mathbb{E}[X]$.

(ii) *Si les tribus* $\sigma(X)$ *et* $\mathcal{F}$ *sont indépendantes, alors* $\mathbb{E}[X \mid \mathcal{F}] = \mathbb{E}[X]$.

(iii) *Si* X *est* $\mathcal{F}$*-mesurable, alors* $\mathbb{E}[X \mid \mathcal{F}] = X$.

(iv) *Si* $\mathcal{G} \subset \mathcal{F}$ *alors* $\mathbb{E}[\mathbb{E}[X \mid \mathcal{F}] \mid \mathcal{G}] = \mathbb{E}[X \mid \mathcal{G}]$.

(v) *Si* Y *est une v.a.* $\mathcal{F}$*-mesurable avec de surcroît le produit* XY *qui est intégrable, alors*

$$\mathbb{E}[XY \mid \mathcal{F}] = Y\mathbb{E}[X \mid \mathcal{F}].$$

(vi) Supposons que X est $\mathcal{F}$-mesurable et soit Y une v.a. indépendante de $\mathcal{F}$. Soit $h : \mathbb{R}^2 \to \mathbb{R}$ une application borélienne telle que $\mathbb{E}[|h(X,Y)|] < \infty$. Alors on a

$$\mathbb{E}[h(X,Y) \mid \mathcal{F}] = H(X), \qquad \textit{avec} \qquad H(x) = \mathbb{E}\left[h(x,Y)\right].$$

(vii) Si X est de carré intégrable, alors $\mathbb{E}\left[X \mid \mathcal{F}\right]$ est le projeté orthogonal de X sur le sous-espace fermé $L^2(\mathcal{F})$.

Démonstration. (i) : Prendre $Z = 1$ qui est bien $\mathcal{F}$-mesurable et bornée.
(ii) : Supposons les tribus $\sigma(X)$ et $\mathcal{F}$ indépendantes. Alors pour toute v.a. $\mathcal{F}$-mesurable et bornée Z,

$$\mathbb{E}\left[X\, Z\right] \;=\; \mathbb{E}\left[X\right]\, \mathbb{E}\left[Z\right].$$

Or $\mathbb{E}\left[X\right]$ étant constante, elle est bien $\mathcal{F}$-mesurable et donc on obtient le résultat par unicité de l'espérance conditionnelle.
(iii) : Même raisonnement que précédemment.
(iv) et (v) : Il suffit de l'écrire (un peu pénible tout de même).
(vi) : Démonstration admise.
(vii) : La v.a. $Y := \mathbb{E}[X \mid \mathcal{F}]$ est le projeté orthogonal de X sur le fermé $F := L^2(\mathcal{F})$ si et seulement si

$$Y \in F \quad \text{and} \quad \|X - Y\|_{L^2} \;=\; \inf_{Z \in F} \|X - Z\|_{L^2}.$$

On sait déjà que $Y \in F$. Ainsi, soit $Z \in F$ et posons $U := X - Y \in F^{\perp}$ et $V := Y - Z \in F$, comme différence de deux éléments de F. Alors $\mathbb{E}\left[UV\right] = 0$ et l'on obtient

$$\begin{aligned}
\mathbb{E}\left[(X-Z)^2\right] &= \mathbb{E}\left[(U+V)^2\right] \\
&= \mathbb{E}\left[U^2\right] + \mathbb{E}\left[V^2\right] + 2\,\mathbb{E}\left[UV\right] \\
&= \mathbb{E}\left[U^2\right] + \mathbb{E}\left[V^2\right],
\end{aligned}$$

quantité qui est minimale pour le choix de $V = 0$, i.e. $Z = Y$. ∎

Chapitre 1

Martingales

Les martingales proviennent en partie de la théorie des jeux. Si M_n désigne la fortune d'un joueur à l'instant n et $\mathcal{F}_n$ l'information disponible sur le jeu jusqu'à ce même instant n, la suite $(M_n)_{n\in\mathbb{N}}$ est une martingale si en moyenne et connaissant $\mathcal{F}_n$, la fortune du joueur à l'instant suivant vaut M_n. Ce concept traduit le fait que le jeu est équitable. L'exemple le plus simple est le suivant : à chaque instant n, vous jouez à pile ou face avec la même pièce de monnaie, chacune des deux faces étant obtenue avec probabilité 1/2. Vous gagnez (resp. perdez) un Euro si vous faites pile (resp. face). La suite de v.a. $(M_n)_{n\in\mathbb{N}}$, où M_n désigne votre fortune à l'instant n, est alors une martingale. En revanche, si la probabilité d'obtenir pile (resp. face) est p (resp. $1-p$), où le paramètre p est différent de 1/2, alors on a affaire à une sous ou sur-martingale selon la valeur de p, la symétrie du jeu étant brisée et le jeu étant biaisé. Du point de vue mathématique, les martingales sont des suites de v.a. ayant un certain nombre de propriétés agréables, en particulier en terme de convergence. L'objectif de ce cours est donc à la fois de présenter les propriétés générales de ces processus et de voir à travers différents exemples comment elles peuvent être utilisées en modélisation mathématique.

1.1 Définition et premières propriétés

Tout d'abord, commençons par un élément de sémantique issue de la théorie du calcul stochastique, c'est-à-dire de l'étude des phénomènes aléatoires indexés par le temps continu.

Définition 1.1.1. *Un processus stochastique (ou aléatoire) à temps discret est une suite de v.a., non nécessairement indépendantes.*

Dans la suite on parlera simplement de processus.

Définition 1.1.2. *Une suite $(\mathcal{F}_n)_{n\in\mathbb{N}}$ de sous-tribus de $\mathcal{A}$ est appelée une filtration de l'espace $(\Omega, \mathcal{A}, \mathbb{P})$ si*

$$\mathcal{F}_0 \subset \mathcal{F}_1 \subset \cdots \subset \mathcal{F}_n \subset \cdots \subset \mathcal{A}.$$

L'espace $(\Omega, \mathcal{A}, (\mathcal{F}_n)_{n\in\mathbb{N}}, \mathbb{P})$ est alors appelé un espace de probabilité filtré.

La notion de tribu est liée à l'information dont nous disposons. Ainsi, supposer cette suite de tribus croissante traduit simplement le fait que plus on avance dans le temps, plus on a d'informations.

Définition 1.1.3. *Un processus $(M_n)_{n\in\mathbb{N}}$ est adapté à une filtration $(\mathcal{F}_n)_{n\in\mathbb{N}}$ si M_n est $\mathcal{F}_n$-mesurable pour tout $n \in \mathbb{N}$.*
Si M_n est $\mathcal{F}_{n-1}$-mesurable pour tout $n \in \mathbb{N}^*$, alors $(M_n)_{n\in\mathbb{N}}$ est dit prévisible.

Notons qu'un processus prévisible est forcément adapté. Par ailleurs, le processus $(M_n)_{n\in\mathbb{N}}$ est évidemment adapté à sa filtration naturelle, définie par $\mathcal{F}_n := \sigma(M_0, M_1, \ldots, M_n)$ pour tout $n \in \mathbb{N}$.

Définition 1.1.4. *Considérons un processus $(M_n)_{n\in\mathbb{N}}$ intégrable (tous ses éléments le sont) et adapté à une filtration $(\mathcal{F}_n)_{n\in\mathbb{N}}$. On dit que le processus $(M_n)_{n\in\mathbb{N}}$ est (pour $(\mathcal{F}_n)_{n\in\mathbb{N}}$) une*

(i) martingale si $\mathbb{E}[M_{n+1} \mid \mathcal{F}_n] = M_n$ pour tout $n \in \mathbb{N}$.

(ii) surmartingale si $\mathbb{E}[M_{n+1} \mid \mathcal{F}_n] \leq M_n$ pour tout $n \in \mathbb{N}$.

(iii) sous-martingale si $\mathbb{E}[M_{n+1} \mid \mathcal{F}_n] \geq M_n$ pour tout $n \in \mathbb{N}$.

Ainsi, le processus intégrable $(M_n)_{n\in\mathbb{N}}$ qui est adapté à la filtration $(\mathcal{F}_n)_{n\in\mathbb{N}}$ est une martingale si et seulement si pour tout événement $A \in \mathcal{F}_n$, on a l'égalité

$$\mathbb{E}\left[M_{n+1} 1_A\right] = \mathbb{E}\left[M_n 1_A\right].$$

Dans le cas d'une sous ou surmartingale, il faut remplacer l'égalité par l'inégalité qui convient. Observons par ailleurs qu'une martingale est à la fois une surmartingale et une sous-martingale. Enfin, elle est forcément d'espérance constante, tandis que celle d'une surmartingale décroît et celle d'une sous-martingale croît.

Donnons-nous quelques exemples bien sympathiques que l'on rencontre souvent en pratique.

(i) Si $(X_n)_{n\in\mathbb{N}^*}$ désigne une suite de v.a. indépendantes et intégrables, alors le processus $(S_n)_{n\in\mathbb{N}^*}$ défini par $S_n := \sum_{i=1}^n X_i$ est une martingale (resp. surmartingale, sous-martingale) pour la filtration engendrée par la suite $(X_n)_{n\in\mathbb{N}^*}$ lorsque pour tout $n \in \mathbb{N}^*$, on a $\mathbb{E}[X_n] = 0$ (resp. $\mathbb{E}[X_n] \leq 0$, $\mathbb{E}[X_n] \geq 0$). Notons que nous n'avons pas supposé que les v.a. étaient de même loi.

(ii) Même conclusion pour le processus donné pour tout $n \in \mathbb{N}^*$ par $M_n := S_n^2 - n\sigma^2$, sous réserve que les X_i ont même variance σ^2.

(iii) Si $(X_n)_{n\in\mathbb{N}^*}$ désigne une suite de v.a. indépendantes, intégrables et d'espérance commune égale à 1, alors le processus $M_n := \prod_{i=1}^n X_i$ est une martingale par rapport à la filtration naturelle de $(X_n)_{n\in\mathbb{N}^*}$.

(iv) Si $(M_n)_{n\in\mathbb{N}}$ est un processus intégrable, centré (chaque élément l'est), et à accroissements indépendants (c'est-à-dire que la suite $(M_{n+1} - M_n)_{n\in\mathbb{N}}$ forme une suite de v.a. indépendantes), alors c'est une martingale par rapport à sa filtration naturelle.

(v) Si $(\mathcal{F}_n)_{n\in\mathbb{N}}$ est une filtration quelconque et X une v.a. intégrable, alors $M_n := \mathbb{E}[X \mid \mathcal{F}_n]$ est une martingale pour $(\mathcal{F}_n)_{n\in\mathbb{N}}$.

Proposition 1.1.5. *Soit $(M_n)_{n\in\mathbb{N}}$ un processus intégrable et adapté à une filtration $(\mathcal{F}_n)_{n\in\mathbb{N}}$. Alors c'est une martingale si et seulement si pour tous $n, p \in \mathbb{N}$,*

$$\mathbb{E}[M_{n+p} \mid \mathcal{F}_n] = M_n.$$

On a le même résultat lorsque $(M_n)_{n\in\mathbb{N}}$ est une sous ou surmartingale (remplacer l'égalité par l'inégalité qui convient).

Démonstration. La condition suffisante est triviale (prendre $p = 1$). Pour la condition nécessaire, notons que l'on a pour $p \in \mathbb{N}^*$,

$$\begin{aligned}
\mathbb{E}[M_{n+p} \mid \mathcal{F}_n] &= \mathbb{E}\left[\sum_{i=1}^p (M_{n+i} - M_{n+i-1}) + M_n \mid \mathcal{F}_n\right] \\
&= \sum_{i=1}^p \mathbb{E}\left[\mathbb{E}\left[M_{n+i} - M_{n+i-1} \mid \mathcal{F}_{n+i-1}\right] \mid \mathcal{F}_n\right] + M_n \\
&= M_n,
\end{aligned}$$

car $(\mathcal{F}_n)_{n\in\mathbb{N}}$ étant une filtration, on a $\mathcal{F}_n \subset \mathcal{F}_{n+i-1}$ pour tout $i \in \{1, \dots, p\}$. ■

Dans la suite, nous supposerons l'espace de probabilité filtré par une filtration générique $(\mathcal{F}_n)_{n\in\mathbb{N}}$. Dans la proposition suivante, on présente quelques propriétés immédiates.

Proposition 1.1.6. (i) *L'opposé d'une surmartingale est une sous-martingale, et inversement.*
(ii) *La somme de deux surmartingales (resp. sous-martingales) est une surmartingale (resp. sous-martingale).*
(iii) *L'ensemble des martingales forme un espace vectoriel sur $\mathbb{R}$.*
(iv) *Si l'on compose une martingale avec une fonction convexe, on obtient une sous-martingale (sous l'hypothèse d'intégrabilité).*

Démonstration. Les trois premières propriétés sont immédiates. Pour la dernière, il suffit d'utiliser l'inégalité de Jensen pour l'espérance conditionnelle. ■

Par exemple, si $(M_n)_{n\in\mathbb{N}}$ est une martingale alors $(|M_n|^p)_{n\in\mathbb{N}}$ avec $p \geq 1$ (sous réserve que chaque M_n est dans L^p) et $(M_n^+)_{n\in\mathbb{N}}$ sont des sous-martingales. Par ailleurs, on déduit de la proposition précédente que la plupart des résultats valables pour une sous-martingale le sera aussi pour une surmartingale, quitte à modifier légèrement les hypothèses (transformer un + en −, une suite croissante en suite décroissante, etc...).

À présent, introduisons deux notions importantes à propos des martingales : la décomposition de Doob et la transformation prévisible.

Théorème 1.1.7. *(Décomposition de Doob d'une sous-martingale)*
Soit $(X_n)_{n\in\mathbb{N}}$ une sous-martingale. Alors, il existe une unique martingale $(M_n)_{n\in\mathbb{N}}$ et un unique processus croissant prévisible $(A_n)_{n\in\mathbb{N}}$, intégrable et issu de 0, tels que

$$X_n = M_n + A_n, \quad n \in \mathbb{N}.$$

Démonstration. Pour l'existence, on définit récursivement les processus $(M_n)_{n\in\mathbb{N}}$ et $(A_n)_{n\in\mathbb{N}}$ de la manière suivante : $A_0 = 0$ et $M_0 = X_0$ puis pour tout $n \geq 1$,

$$\begin{aligned} M_n &= M_{n-1} + X_n - X_{n-1} - \mathbb{E}[X_n - X_{n-1} \mid \mathcal{F}_{n-1}] \\ A_n &= A_{n-1} + \mathbb{E}[X_n - X_{n-1} \mid \mathcal{F}_{n-1}]. \end{aligned}$$

On vérifie aisément que $(M_n)_{n\in\mathbb{N}}$ est une martingale par construction, que $(A_n)_{n\in\mathbb{N}}$ est un processus croissant par la propriété de sous-martingale de la suite $(X_n)_{n\in\mathbb{N}}$, prévisible et intégrable par construction. Enfin, ils satisfont $X_n = M_n + A_n$ pour tout $n \in \mathbb{N}$.
Pour l'unicité, supposons qu'il existe deux couples de processus $(M_n)_{n\in\mathbb{N}}$, $(A_n)_{n\in\mathbb{N}}$ et $(N_n)_{n\in\mathbb{N}}$, $(B_n)_{n\in\mathbb{N}}$ vérifiant la conclusion, et montrons qu'ils sont égaux. Comme pour tout $n \in \mathbb{N}$,

$$X_n = M_n + A_n = N_n + B_n,$$

on a que

$$M_n - N_n = B_n - A_n, \quad n \in \mathbb{N}.$$

Ainsi, le processus $(B_n - A_n)_{n\in\mathbb{N}}$ est une martingale et l'on obtient pour tout $n \in \mathbb{N}$,

$$\mathbb{E}[B_{n+1} - A_{n+1} \mid \mathcal{F}_n] = B_n - A_n.$$

Or les processus $(A_n)_{n\in\mathbb{N}}$ et $(B_n)_{n\in\mathbb{N}}$ étant supposés prévisibles, on a

$$\mathbb{E}[B_{n+1} \mid \mathcal{F}_n] = B_{n+1} \quad \text{et} \quad \mathbb{E}[A_{n+1} \mid \mathcal{F}_n] = A_{n+1}.$$

Autrement dit, le processus $(B_n - A_n)_{n\in\mathbb{N}}$ est constant et vaut donc $B_0 - A_0 = 0$. Ainsi, les processus $(A_n)_{n\in\mathbb{N}}$ et $(B_n)_{n\in\mathbb{N}}$ coïncident et donc les processus $(M_n)_{n\in\mathbb{N}}$ et $(N_n)_{n\in\mathbb{N}}$ aussi. ∎

L'intérêt de la décomposition de Doob réside dans le fait que l'étude d'une sous-martingale peut souvent se ramener à celle d'une martingale, en général plus facile à appréhender (modulo la présence d'un processus croissant prévisible). Cette remarque sera illustrée dans la partie convergence des martingales.

La transformation prévisible d'une martingale permet de construire une nouvelle martingale à partir d'une martingale originelle, sous certaines conditions sur la suite prévisible.

Définition et théorème 1.1.8. *(Transformation prévisible d'une martingale)* *Soit $(M_n)_{n\in\mathbb{N}}$ une martingale et soit $(A_n)_{n\in\mathbb{N}^*}$ un processus prévisible. La transformation prévisible de la martingale $(M_n)_{n\in\mathbb{N}}$ par le processus $(A_n)_{n\in\mathbb{N}^*}$ est définie par le processus donné par $\widetilde{M}_0 = M_0$ et*

$$\widetilde{M}_n := \widetilde{M}_0 + \sum_{i=1}^{n} A_i\,(M_i - M_{i-1}), \quad n \in \mathbb{N}^*.$$

Si les v.a. A_n sont bornées, alors $(\widetilde{M}_n)_{n\in\mathbb{N}}$ est une martingale.

Démonstration. Tout d'abord, remarquons que $(\widetilde{M}_n)_{n\in\mathbb{N}}$ est clairement adapté par rapport à la filtration $(\mathcal{F}_n)_{n\in\mathbb{N}}$, et que la bornitude des A_n assure simplement l'intégrabilité des éléments $\widetilde{M}_n$. Étant donné un entier n, on a par la prévisibilité de A_{n+1},

$$\begin{aligned} \mathbb{E}\left[\widetilde{M}_{n+1} - \widetilde{M}_n \mid \mathcal{F}_n\right] &= \mathbb{E}\left[A_{n+1}(M_{n+1} - M_n) \mid \mathcal{F}_n\right] \\ &= A_{n+1}\,\mathbb{E}\left[M_{n+1} - M_n \mid \mathcal{F}_n\right] \\ &= 0, \end{aligned}$$

car $(M_n)_{n\in\mathbb{N}}$ est une martingale. ■

1.2 Martingales de carré intégrable

Dans cette partie, nous allons nous intéresser à une sous-classe très importante des martingales : les martingales de carré intégrable. En particulier, cette étude va donner lieu à une version à temps discret d'un objet fondamental de la théorie du calcul stochastique : l'intégrale stochastique d'Itô.

Définition 1.2.1. *Un processus $(X_n)_{n\in\mathbb{N}}$ est dit de carré intégrable, ou simplement dans L^2, si $\mathbb{E}[X_n^2] < +\infty$ pour tout $n \in \mathbb{N}$.*

Le résultat suivant est la version L^2 du théorème de transformation prévisible. La démonstration, laissée en exercice, est la même que celle effectuée dans le cas classique, avec en plus le caractère L^2 (il faut montrer que le processus $(\widetilde{M}_n)_{n\in\mathbb{N}}$ est de carré intégrable). Notons que la transformation prévisible peut être vue comme une discrétisation temporelle de la célèbre intégrale stochastique d'Itô.

Théorème 1.2.2. *(Transformation prévisible d'une martingale, version L^2)*
Soit $(\widetilde{M}_n)_{n\in\mathbb{N}}$ la transformation prévisible d'une martingale de carré intégrable $(M_n)_{n\in\mathbb{N}}$ par le processus prévisible $(A_n)_{n\in\mathbb{N}^*}$, supposé borné. Alors $(\widetilde{M}_n)_{n\in\mathbb{N}}$ est une martingale de carré intégrable.

À présent, nous allons introduire la variation quadratique d'une martingale de carré intégrable.

Définition et théorème 1.2.3. *Soit $(M_n)_{n\in\mathbb{N}}$ une martingale de carré intégrable. Alors il existe un unique processus croissant et prévisible, intégrable et issu de 0, appelé la variation quadratique de $(M_n)_{n\in\mathbb{N}}$ et noté $([M,M]_n)_{n\in\mathbb{N}}$, tel que le processus $(M_n^2 - [M,M]_n)_{n\in\mathbb{N}}$ soit une martingale.*
La variation quadratique admet la représentation suivante : $[M,M]_0 = 0$ et pour tout $n \in \mathbb{N}^*$,

$$[M,M]_n = \sum_{i=1}^{n} \mathbb{E}\left[(M_i - M_{i-1})^2 \mid \mathcal{F}_{i-1}\right].$$

Démonstration. La martingale $(M_n)_{n\in\mathbb{N}}$ étant de carré intégrable, le processus $(M_n^2)_{n\in\mathbb{N}}$ est une sous-martingale et par la décomposition de Doob, il existe une unique martingale $(N_n)_{n\in\mathbb{N}}$ et un unique processus croissant $(A_n)_{n\in\mathbb{N}}$, prévisible, intégrable et issu de 0, tels que

$$M_n^2 = N_n + A_n, \quad n \in \mathbb{N}.$$

Le processus $(A_n)_{n\in\mathbb{N}}$ est la variation quadratique de la martingale de carré intégrable $(M_n)_{n\in\mathbb{N}}$, et est noté dans la suite $([M,M]_n)_{n\in\mathbb{N}}$.
À présent, établissons l'expression explicite de la variation quadratique. Pour ce faire il suffit de montrer, par unicité de la décomposition de Doob, que le processus suivant est une martingale : $N_0 := M_0^2$ et

$$N_n := M_n^2 - \sum_{i=1}^{n} \mathbb{E}\left[(M_i - M_{i-1})^2 \mid \mathcal{F}_{i-1}\right], \quad n \in \mathbb{N}^*.$$

On a pour tout $n \in \mathbb{N}$,

$$\mathbb{E}[N_{n+1} - N_n \mid \mathcal{F}_n] = \mathbb{E}[M_{n+1}^2 \mid \mathcal{F}_n] - M_n^2 - \mathbb{E}\left[(M_{n+1} - M_n)^2 \mid \mathcal{F}_n\right]$$

$$
\begin{aligned}
&= \mathbb{E}[M_{n+1}^2 \mid \mathcal{F}_n] - M_n^2 - \mathbb{E}\left[M_{n+1}^2 \mid \mathcal{F}_n\right] - M_n^2 + 2M_n\mathbb{E}\left[M_{n+1} \mid \mathcal{F}_n\right] \\
&= \mathbb{E}[M_{n+1}^2 \mid \mathcal{F}_n] - M_n^2 - \mathbb{E}\left[M_{n+1}^2 - M_n^2 \mid \mathcal{F}_n\right] \\
&= 0,
\end{aligned}
$$

car le processus $(M_n)_{n\in\mathbb{N}}$ est une martingale. ∎

Notons que nous avons démontré le résultat suivant, que l'on utilise souvent en pratique : si $(M_n)_{n\in\mathbb{N}}$ est une martingale de carré intégrable, alors on a pour tout $n \in \mathbb{N}$,

$$\mathbb{E}\left[(M_{n+1} - M_n)^2 \mid \mathcal{F}_n\right] = \mathbb{E}\left[M_{n+1}^2 \mid \mathcal{F}_n\right] - M_n^2.$$

Par ailleurs, en prenant l'espérance de la martingale $(M_n^2 - [M,M]_n)_{n\in\mathbb{N}}$, on obtient

$$\mathbb{E}\left[M_n^2\right] = \mathbb{E}\left[M_0^2\right] + \mathbb{E}\left[[M,M]_n\right], \quad n \in \mathbb{N},$$

et en particulier

$$\mathrm{Var}(M_n) = \mathrm{Var}(M_0) + \mathbb{E}\left[[M,M]_n\right], \quad n \in \mathbb{N},$$

une identité nous permettant souvent de calculer la variance d'une martingale de carré intégrable lorsque l'on connaît explicitement sa variation quadratique. Enfin, le dernier résultat de cette partie est une formule exprimant la variation quadratique de la transformation prévisible en fonction de celle de la martingale originelle. Nous nous servirons de ce résultat dans le prochain chapitre pour démontrer la loi des grands nombres pour les martingales.

Théorème 1.2.4. *(Variation quadratique d'une transformation prévisible)*
Soit $(\widetilde{M}_n)_{n\in\mathbb{N}}$ la transformation prévisible d'une martingale de carré intégrable $(M_n)_{n\in\mathbb{N}}$ par le processus prévisible $(A_n)_{n\in\mathbb{N}^*}$, supposé borné. Alors on a la formule suivante : pour tout $n \in \mathbb{N}^*$,

$$[\widetilde{M}, \widetilde{M}]_n = \sum_{i=1}^{n} A_i^2 \left([M,M]_i - [M,M]_{i-1}\right).$$

Démonstration. On a $[\widetilde{M}, \widetilde{M}]_0 = 0$ par définition et pour tout $n \in \mathbb{N}$,

$$
\begin{aligned}
[\widetilde{M}, \widetilde{M}]_n &= \sum_{i=1}^{n} \mathbb{E}\left[(\widetilde{M}_i - \widetilde{M}_{i-1})^2 \mid \mathcal{F}_{i-1}\right] \\
&= \sum_{i=1}^{n} \mathbb{E}\left[A_i^2 (M_i - M_{i-1})^2 \mid \mathcal{F}_{i-1}\right] \\
&= \sum_{i=1}^{n} A_i^2 \mathbb{E}\left[(M_i - M_{i-1})^2 \mid \mathcal{F}_{i-1}\right]
\end{aligned}
$$

$$= \sum_{i=1}^{n} A_i^2 \left([M,M]_i - [M,M]_{i-1}\right),$$

car le processus $(A_n)_{n\in\mathbb{N}^*}$ est prévisible, d'où le résultat désiré. ■

1.3 Théorème d'arrêt

Dans cette partie, nous allons établir le théorème d'arrêt pour les martingales, résultat très utile en pratique, par exemple lorsque l'on considère le problème de la ruine du joueur. Avant tout, définissons la tribu correspondant à l'ensemble des informations disponibles. C'est la tribu engendrée par toutes les tribus $\mathcal{F}_n$, à savoir

$$\mathcal{F}_\infty := \sigma\left(\cup_{n\geq 0}\mathcal{F}_n\right).$$

En d'autres termes, c'est la plus petite tribu contenant toutes les tribus $\mathcal{F}_n$. On rappelle que l'union de tribus n'étant pas forcément une tribu, il nous faut prendre dans la définition la tribu engendrée par l'union des tribus. À présent, introduisons la notion de temps d'arrêt relativement à la filtration $(\mathcal{F}_n)_{n\in\mathbb{N}}$.

Définition 1.3.1. *Une v.a. entière τ, pouvant prendre la valeur $+\infty$, est un temps d'arrêt pour la filtration $(\mathcal{F}_n)_{n\in\mathbb{N}}$ si pour tout $n\in\mathbb{N}$,*

$$\{\tau\leq n\}\in\mathcal{F}_n.$$

Notons que dans cette définition on peut remplacer l'événement $\{\tau\leq n\}$ par $\{\tau = n\}$. En effet, la propriété "$\{\tau\leq n\}\in\mathcal{F}_n$ pour tout $n\in\mathbb{N}$" est équivalent à celle "$\{\tau = n\}\in\mathcal{F}_n$ pour tout $n\in\mathbb{N}$".

Pour illuster cette notion de temps d'arrêt, prenons l'exemple d'un joueur rentrant avec une somme S_0 dans un casino. On note S_n sa fortune à l'instant n. Ce joueur décide de jouer tant qu'il a encore de l'argent dans son portefeuille (on suppose le casino ouvert 24h sur 24). Cela signifie qu'il joue jusqu'à l'instant

$$\tau := \inf\{n\in\mathbb{N} : S_n = 0\},$$

qui est un temps d'arrêt pour la filtration naturelle de la suite $(S_n)_{n\in\mathbb{N}}$. En effet, on a pour tout $n\in\mathbb{N}$,

$$\{\tau\leq n\} = \bigcup_{i=0}^{n}\{S_i = 0\}\in\mathcal{F}_n.$$

Dans la proposition suivante, on énonce quelques propriétés des temps d'arrêt, dont les démonstrations sont immédiates. On note dans la suite $a\wedge b = \min\{a,b\}$ et $a\vee b = \max\{a,b\}$.

Proposition 1.3.2. *Soient τ_1 et τ_2 deux temps d'arrêt. Alors $\tau_1 \wedge \tau_2$, $\tau_1 \vee \tau_2$ et $\tau_1 + \tau_2$ sont des temps d'arrêt.*

Étant donné un temps d'arrêt τ, on définit la famille d'événements suivante :

$$\mathcal{F}_\tau := \{A \in \mathcal{F}_\infty : A \cap \{\tau \leq n\} \in \mathcal{F}_n \quad \forall n \in \mathbb{N}\}.$$

On peut démontrer que cette famille est une tribu, appelée tribu des événements antérieurs à τ. Là aussi, remplacer l'inégalité par une égalité ne modifie en rien la tribu (on pourra utiliser cette définition alternative par la suite). Notons par ailleurs que si τ n'est pas un temps d'arrêt, l'univers Ω n'est plus dans la famille $\mathcal{F}_\tau$ qui n'est plus une tribu.
De cette définition relativement difficile, il faut retenir que $\mathcal{F}_\tau$ représente l'ensemble des informations dont on dispose à l'instant aléatoire τ. La proposition suivante justifie les résultats auxquels on s'attend en considérant la tribu des événements antérieurs à un temps d'arrêt.

Proposition 1.3.3. *On a les propriétés suivantes :*
(i) Si $\tau(\omega) = k$ pour tout $\omega \in \Omega$, alors τ est un temps d'arrêt et $\mathcal{F}_\tau = \mathcal{F}_k$ (il n'y a donc pas d'ambiguïté de notation).
(ii) Le temps d'arrêt τ est $\mathcal{F}_\tau$-mesurable.
(iii) Si $\tau_1 \leq \tau_2$, alors $\mathcal{F}_{\tau_1} \subset \mathcal{F}_{\tau_2}$.
(iv) On a l'égalité $\mathcal{F}_{\tau_1 \wedge \tau_2} = \mathcal{F}_{\tau_1} \cap \mathcal{F}_{\tau_2}$.

Démonstration. (i) Si τ est constant alors on a $\{\tau \leq n\} = \Omega$ ou $\emptyset$ selon que $n \geq k$ ou $n < k$. Dans les deux cas l'événement $\{\tau \leq n\}$ appartient à la tribu $\mathcal{F}_n$, d'où τ est un temps d'arrêt.
À présent, montrons par double inclusion que $\mathcal{F}_\tau = \mathcal{F}_k$. Si $A \in \mathcal{F}_k$ alors on a :
∘ si $n \geq k$ alors $A \cap \{\tau \leq n\} = A \in \mathcal{F}_k \subset \mathcal{F}_n$;
∘ si $n < k$ alors $A \cap \{\tau \leq n\} = \emptyset \in \mathcal{F}_n$.
D'où $A \in \mathcal{F}_\tau$. Quant à l'autre inclusion, on note que si $A \in \mathcal{F}_\tau$ alors $A \cap \{\tau \leq n\} \in \mathcal{F}_n$ pour tout $n \in \mathbb{N}$, en particulier pour $n \geq k$ on obtient que $A \in \mathcal{F}_k$.
(ii) Le temps d'arrêt τ est $\mathcal{F}_\tau$-mesurable si pour tout $k \in \mathbb{N}$, on a $\{\tau = k\} \in \mathcal{F}_\tau$, c'est-à-dire si pour tous $k, n \in \mathbb{N}$, on a $\{\tau = k\} \cap \{\tau = n\} \in \mathcal{F}_n$, ce que l'on démontre sans difficulté après avoir considéré les cas $n \neq k$ et $n = k$.
(iii) Soit $A \in \mathcal{F}_{\tau_1}$. Comme $\tau_1 \leq \tau_2$, on a pour tout $n \in \mathbb{N}$,

$$A \cap \{\tau_2 \leq n\} = A \cap \{\tau_1 \leq n\} \cap \{\tau_2 \leq n\} \in \mathcal{F}_n,$$

comme intersection de deux événements de la tribu $\mathcal{F}_n$ (τ_2 est un temps d'arrêt).
(iv) La première inclusion étant immédiate par (iii), démontrons seulement l'inclusion $\mathcal{F}_{\tau_1} \cap \mathcal{F}_{\tau_2} \subset \mathcal{F}_{\tau_1 \wedge \tau_2}$. Soit $A \in \mathcal{F}_{\tau_1} \cap \mathcal{F}_{\tau_2}$. On a pour tout $n \in \mathbb{N}$,

$$A \cap \{\tau_1 \wedge \tau_2 \leq n\} = (A \cap \{\tau_1 \leq n\}) \cup (A \cap \{\tau_2 \leq n\}) \in \mathcal{F}_n,$$

comme union de deux événements de la tribu $\mathcal{F}_n$. Ainsi, $A \in \mathcal{F}_{\tau_1 \wedge \tau_2}$, d'où l'inclusion désirée. ■

Dans la pratique, comme par exemple pour le problème de la ruine du joueur, ce qui nous intéresse est le comportement d'un processus $(X_n)_{n\in\mathbb{N}}$ évalué au temps d'arrêt τ. S'il est supposé fini p.s., alors on peut définir la v.a. X_τ comme

$$X_\tau := \sum_{n\in\mathbb{N}} 1_{\{\tau=n\}} X_n.$$

On montre facilement que X_τ est $\mathcal{F}_\tau$-mesurable. Lorsque que le temps d'arrêt prend la valeur $+\infty$, c'est-à-dire que $\mathbb{P}(\tau = +\infty) > 0$, nous pouvons toujours le tronquer en considérant plutôt pour un entier n donné le temps d'arrêt $n \wedge \tau$, qui est fini et même borné. Dans ce cas, la v.a. $X_{n\wedge\tau}$ est bien définie.

Énonçons à présent le théorème d'arrêt pour les martingales (valable aussi pour les sous et surmartingales sous les mêmes hypothèses, sous réserve de modifier les égalités par les inégalités qui conviennent).

Théorème 1.3.4. *(Théorème d'arrêt pour les martingales)*
Soit $(M_n)_{n\in\mathbb{N}}$ une martingale et soit τ un temps d'arrêt. Alors le processus arrêté $(M_{n\wedge\tau})_{n\in\mathbb{N}}$ est aussi une martingale. De plus, si $\tau_1 \leq \tau_2$ sont deux temps d'arrêt bornés, alors on a

$$\mathbb{E}\left[M_{\tau_2} \mid \mathcal{F}_{\tau_1}\right] = M_{\tau_1} \quad \text{et} \quad \mathbb{E}[M_{\tau_1}] = \mathbb{E}[M_{\tau_2}] = \mathbb{E}[M_0].$$

Démonstration. Quitte à remplacer $(M_n)_{n\in\mathbb{N}}$ par la martingale $(M_n - M_0)_{n\in\mathbb{N}}$, supposons sans perte de généralité que $M_0 = 0$ (ce que l'on fera de temps en temps dans la suite de ce cours). La v.a. τ étant un temps d'arrêt, les v.a. définies pour tout $n \in \mathbb{N}^*$ par $A_n := 1_{\{\tau\geq n\}}$ forment un processus prévisible. De plus, on a

$$\sum_{i=1}^{n} A_i \,(M_i - M_{i-1}) = \sum_{i=1}^{n\wedge\tau} (M_i - M_{i-1}) = M_{n\wedge\tau},$$

ce qui entraîne que la suite $(M_{n\wedge\tau})_{n\in\mathbb{N}}$ est la transformation prévisible de la martingale $(M_n)_{n\in\mathbb{N}}$ par rapport à $(A_n)_{n\in\mathbb{N}^*}$. Ainsi, c'est une martingale par le théorème 1.1.8 et la première affirmation du théorème est démontrée. Le lecteur averti aura noté que cette propriété de martingale peut être démontrée directement : montrer d'abord l'intégrabilité, ensuite le fait que $(M_{n\wedge\tau})_{n\in\mathbb{N}}$ est adapté à la filtration sous-jacente, puis la propriété de martingale en remarquant que

$$\begin{aligned} M_{n\wedge\tau} &= M_n 1_{\{\tau\geq n\}} + M_\tau 1_{\{\tau<n\}} \\ &= M_n 1_{\{\tau\geq n\}} + \sum_{i=1}^{n-1} M_i 1_{\{\tau=i\}}. \end{aligned}$$

Démontrons maintenant les deux égalités. Supposons τ_2 borné par une constante positive κ. Comme $(M_{n\wedge\tau_2})_{n\in\mathbb{N}}$ est une martingale, on a pour $n_0 := \lfloor\kappa\rfloor + 1$ (la notation $\lfloor\cdot\rfloor$ désignant la partie entière) et tout événement $A \in \mathcal{F}_{\tau_1}$ que

$$\begin{aligned}
\mathbb{E}\left[M_{\tau_2} 1_A\right] &= \mathbb{E}\left[M_{n_0\wedge\tau_2} 1_A\right] \\
&= \sum_{p=0}^{n_0-1} \mathbb{E}\left[M_{n_0\wedge\tau_2} 1_{\{\tau_1=p\}\cap A}\right] \qquad \text{car } \tau_1 < n_0 \\
&= \sum_{p=0}^{n_0-1} \mathbb{E}\left[M_{p\wedge\tau_2} 1_{\{\tau_1=p\}\cap A}\right] \qquad \text{car } \{\tau_1 = p\} \cap A \in \mathcal{F}_p \\
&= \mathbb{E}\left[M_{\tau_1\wedge\tau_2} 1_A\right] \\
&= \mathbb{E}\left[M_{\tau_1} 1_A\right],
\end{aligned}$$

c'est-à-dire, en terme d'espérance conditionnelle, que

$$\mathbb{E}\left[M_{\tau_2} \mid \mathcal{F}_{\tau_1}\right] = M_{\tau_1}.$$

Enfin, si τ désigne τ_1 ou τ_2, comme p.s. $M_{n\wedge\tau} \underset{n\to+\infty}{\longrightarrow} M_\tau$ et que

$$\begin{aligned}
|M_{n\wedge\tau}| &\leq \sum_{i=1}^{n\wedge\tau} |M_i - M_{i-1}| \\
&\leq \sum_{i=1}^{\kappa} (|M_i| + |M_{i-1}|),
\end{aligned}$$

qui est intégrable sans dépendre de n, le théorème de convergence dominée entraîne que

$$\mathbb{E}[M_\tau] = \lim_{n\to+\infty} \mathbb{E}[M_{n\wedge\tau}] = 0.$$

Ainsi, le théorème d'arrêt est démontré dans sa totalité. ■

Chapitre 2

Inégalités maximales et théorèmes de convergence

2.1 Inégalités maximales de Doob

Lorsque l'on étudie un processus aléatoire, une question importante en pratique est de savoir contrôler son évolution. En particulier, il s'avère que pour les sous et sur-martingales positives, nous pouvons obtenir des inégalités sur son processus supremum, à savoir

$$M_n^* := \sup_{0 \le k \le n} M_k, \quad n \in \mathbb{N}.$$

On les appelle inégalités maximales de Doob. Dans cette partie, nous établissons des inégalités maximales pour les sous et surmartingales positives qui nous serviront comme ingrédient de base dans la démonstration des théorèmes de convergence. En particulier, une fois que tous les résultats ci-dessous seront démontrés pour les sous-martingales positives, ils seront immédiatement valables pour des martingales, en remplaçant M_n^* par $\sup_{0 \le k \le n} |M_k|$.

Théorème 2.1.1 (Inégalités maximales de Doob)**.** *Soit $(M_n)_{n\in\mathbb{N}}$ une sous-martingale positive. Alors pour tout $n \in \mathbb{N}$,*

$$\mathbb{P}(M_n^* \ge \lambda) \le \frac{\mathbb{E}[M_n \, 1_{\{M_n^* \ge \lambda\}}]}{\lambda} \le \frac{\mathbb{E}[M_n]}{\lambda}, \quad \lambda > 0.$$

Par ailleurs, si $(M_n)_{n\in\mathbb{N}}$ a tous ses éléments dans L^p, où $p \ge 1$, alors pour tout $n \in \mathbb{N}$,

$$\mathbb{P}(M_n^* \ge \lambda) \le \frac{\mathbb{E}[M_n^p]}{\lambda^p}, \quad \lambda > 0.$$

Enfin, si $(M_n)_{n\in\mathbb{N}}$ *une surmartingale positive, on a*

$$\mathbb{P}(M_\infty^* \geq \lambda) \leq \frac{\mathbb{E}[M_0]}{\lambda}, \quad \lambda > 0.$$

Démonstration. Établissons tout d'abord la première inégalité. Notons que pour une sous-martingale, on a

$$M_k \leq \mathbb{E}[M_n \mid \mathcal{F}_k], \quad 0 \leq k \leq n,$$

ce qui entraîne que pour tout $A \in \mathcal{F}_k$, en multipliant par 1_A et en prenant l'espérance,

$$\mathbb{E}[M_k\, 1_A] \leq \mathbb{E}[M_n\, 1_A].$$

Par ailleurs, on peut écrire l'identité suivante : pour tout $\lambda > 0$,

$$\{M_n^* \geq \lambda\} = \{\tau_\lambda \leq n\}, \quad n \in \mathbb{N},$$

où τ_λ est le temps d'arrêt $\tau_\lambda := \inf\{n \in \mathbb{N} : M_n \geq \lambda\}$. Si $\tau_\lambda \leq n$, alors on a $M_{\tau_\lambda} \geq \lambda$ et l'on peut écrire

$$\lambda\, 1_{\{\tau_\lambda \leq n\}} \leq M_{\tau_\lambda}\, 1_{\{\tau_\lambda \leq n\}}.$$

En passant à l'espérance et en utilisant l'inégalité entre espérances précédente, on obtient

$$\begin{aligned}
\lambda\, \mathbb{P}(\tau_\lambda \leq n) &= \lambda\, \mathbb{E}[1_{\{\tau_\lambda \leq n\}}] \\
&\leq \mathbb{E}[M_{\tau_\lambda}\, 1_{\{\tau_\lambda \leq n\}}] \\
&= \sum_{k=0}^n \mathbb{E}[M_k\, 1_{\{\tau_\lambda = k\}}] \\
&\leq \sum_{k=0}^n \mathbb{E}[M_n\, 1_{\{\tau_\lambda = k\}}] \quad \text{car } \{\tau_\lambda = k\} \in \mathcal{F}_k \\
&= \mathbb{E}[M_n\, 1_{\{\tau_\lambda \leq n\}}],
\end{aligned}$$

ce qui conclut la démonstration de la première inégalité.
Pour la seconde inégalité, on notera que si $(M_n)_{n\in\mathbb{N}}$ est une sous-martingale positive telle que $M_n \in L^p$ pour tout $n \in \mathbb{N}$, alors $(M_n^p)_{n\in\mathbb{N}}$ est une sous-martingale positive et l'inégalité précédente s'applique.
Enfin, pour la troisième inégalité, le processus $(M_n)_{n\in\mathbb{N}}$ étant une surmartingale positive, il suffit de reprendre le raisonnement précédent pour aboutir à l'inégalité

$$\lambda\, \mathbb{P}(\tau_\lambda \leq n) \leq \mathbb{E}[M_0\, 1_{\{\tau_\lambda \leq n\}}] \leq \mathbb{E}[M_0],$$

et le théorème de convergence monotone achève la démonstration. ∎

Ce qui est remarquable dans les inégalités maximales, c'est que l'on réussit à faire passer le "sup" à l'extérieur de la probabilité. De plus, on voit que l'on contrôle ici l'évolution du processus par l'espérance de la valeur initiale ou de la valeur terminale du processus. À présent, donnons une seconde inégalité maximale, comparant l'espérance du processus supremum avec celle du processus originel.

Théorème 2.1.2 (Inégalités de Doob L^p)**.** *Soit $p > 1$. Supposons que $(M_n)_{n\in\mathbb{N}}$ soit une sous-martingale positive telle que $M_n \in L^p$ pour tout $n \in \mathbb{N}$. Alors pour tout $n \in \mathbb{N}$, $M_n^* \in L^p$ et*

$$\mathbb{E}[(M_n^*)^p] \leq \left(\frac{p}{p-1}\right)^p \mathbb{E}[M_n^p].$$

Démonstration. Tout d'abord, le fait que $M_n^* \in L^p$ est une conséquence des inégalités

$$(M_n^*)^p \leq \left(\sum_{k=0}^n M_k\right)^p \leq (n+1)^{p-1} \sum_{k=0}^n M_k^p.$$

En se ramenant à la première inégalité maximale du théorème 2.1.1, il vient

$$\begin{aligned}
\mathbb{E}[(M_n^*)^p] &= p \int_0^\infty x^{p-1}\, \mathbb{P}(M_n^* \geq x)\, dx \\
&\leq p \int_0^\infty x^{p-2}\, \mathbb{E}[M_n\, 1_{\{M_n^* \geq x\}}]\, dx \\
&= p\, \mathbb{E}[M_n \int_0^{M_n^*} x^{p-2}\, dx] \\
&= \frac{p}{p-1}\, \mathbb{E}[M_n\, (M_n^*)^{p-1}] \\
&\leq \frac{p}{p-1}\, \mathbb{E}[M_n^p]^{1/p}\, \mathbb{E}[(M_n^*)^p]^{(p-1)/p},
\end{aligned}$$

où dans la dernière inégalité nous avons utilisé l'inégalité de Hölder avec les exposants p et $q = p/(p-1)$. Enfin, on regroupe les termes comme il faut dans l'inégalité ci-dessus et ceci achève la démonstration. ■

L'inégalité de Doob la plus utilisée est celle pour l'espace L^2, c'est-à-dire

$$\mathbb{E}\left[\sup_{0\leq k\leq n} |M_k|^2\right] \leq 4\, \mathbb{E}\left[M_n^2\right], \quad n \in \mathbb{N}.$$

2.2 Convergence des martingales

L'une des raisons qui expliquent l'importance des martingales sont les théorèmes de convergence, qui garantissent qu'une martingale converge p.s. sous des hypothèses souvent faciles à vérifier. Commençons par énoncer un premier résultat dans le cas des martingales bornées dans L^2, dont la preuve repose essentiellement sur une inégalité maximale de Doob.

Théorème 2.2.1. *Soit $(M_n)_{n\in\mathbb{N}}$ une martingale bornée dans L^2, i.e.*

$$\sup_{n\in\mathbb{N}} \mathbb{E}\left[M_n^2\right] < +\infty.$$

Alors la suite $(M_n)_{n\in\mathbb{N}}$ converge p.s. et dans L^2 vers une v.a. $M_\infty \in L^2$.
En particulier, une martingale bornée (i.e. telle que la v.a. $\sup_{n\in\mathbb{N}} |M_n|$ est bornée) est p.s. convergente.

Démonstration. Pour toute paire arbitraire de rationnels $a < b$, notons l'ensemble

$$A_{a,b} := \left\{\omega \in \Omega : \liminf_{n\to+\infty} M_n(\omega) \leq a < b \leq \limsup_{n\to+\infty} M_n(\omega)\right\}.$$

Ainsi, la suite $(M_n)_{n\in\mathbb{N}}$ va converger p.s. (éventuellement vers l'infini) à partir du moment où l'on montre que $\mathbb{P}(\cup_{a,b\in\mathbb{Q}} A_{a,b}) = 0$. Afin de faire le rapprochement avec les inégalités maximales, observons que l'on a

$$A_{a,b} \subset \left\{\omega \in \Omega : \sup_{k\geq m} |M_k(\omega) - M_m(\omega)| \geq \frac{b-a}{2}\right\}, \quad m \in \mathbb{N}.$$

En effet, si $\omega \in A_{a,b}$, alors pour tout $m \in \mathbb{N}$,

$$\begin{aligned} b-a &\leq \sup_{k\geq m} M_k(\omega) - \inf_{k\geq m} M_k(\omega) \\ &= \sup_{k,l\geq m} |M_k(\omega) - M_l(\omega)| \\ &\leq 2 \sup_{k\geq m} |M_k(\omega) - M_m(\omega)|. \end{aligned}$$

Considérons à présent les accroissements définis par $d_k := M_k - M_{k-1}$, $k \in \mathbb{N}^*$. Comme la suite $(M_n)_{n\in\mathbb{N}}$ est de carré intégrable, on en déduit que les $(d_k)_{k\in\mathbb{N}^*}$ sont orthogonaux entre eux. En effet, on a pour tous $1 \leq k < l$,

$$\begin{aligned} \mathbb{E}[d_k d_l] &= \mathbb{E}[(M_k - M_{k-1})(M_l - M_{l-1})] \\ &= \mathbb{E}[M_k M_l] - \mathbb{E}[M_k M_{l-1}] - \mathbb{E}[M_{k-1} M_l] + \mathbb{E}[M_{k-1} M_{l-1}] \\ &= \mathbb{E}[\mathbb{E}[M_k M_l \mid \mathcal{F}_k]] - \mathbb{E}[\mathbb{E}[M_k M_{l-1} \mid \mathcal{F}_k]] - \mathbb{E}[\mathbb{E}[M_{k-1} M_l \mid \mathcal{F}_{k-1}]] \\ &\quad + \mathbb{E}[\mathbb{E}[M_{k-1} M_{l-1} \mid \mathcal{F}_{k-1}]] \\ &= \mathbb{E}[M_k \mathbb{E}[M_l \mid \mathcal{F}_k]] - \mathbb{E}[M_k \mathbb{E}[M_{l-1} \mid \mathcal{F}_k]] - \mathbb{E}[M_{k-1} \mathbb{E}[M_l \mid \mathcal{F}_{k-1}]] \end{aligned}$$

$$
\begin{aligned}
&+\mathbb{E}\left[M_{k-1}\mathbb{E}[M_{l-1} \mid \mathcal{F}_{k-1}]\right] \\
&= \mathbb{E}\left[M_k^2\right] - \mathbb{E}\left[M_k^2\right] - \mathbb{E}\left[M_{k-1}^2\right] + \mathbb{E}\left[M_{k-1}^2\right] \\
&= 0.
\end{aligned}
$$

Comme précédemment nous pouvons supposer sans perte de généralité que $M_0 = 0$, et l'on obtient alors

$$
\mathbb{E}[M_n^2] = \mathbb{E}\left[\left(\sum_{k=1}^n d_k\right)^2\right] = \sum_{k=1}^n \mathbb{E}\left[d_k^2\right],
$$

qui est la somme partielle d'une série convergente par l'hypothèse de L^2-bornitude. Par ailleurs, le processus $(M_{k+m} - M_m)_{k\in\mathbb{N}}$ étant une martingale (pour la filtration translatée $(\mathcal{F}_{k+m})_{k\in\mathbb{N}}$), la suite $(M'_k)_{k\in\mathbb{N}}$ définie pour tout $k \in \mathbb{N}$ par $M'_k := (M_{k+m} - M_m)^2$ est une sous-martingale positive et l'inégalité de Doob du théorème 2.1.1, combinée au lemme de Fatou, s'applique de la manière suivante :

$$
\begin{aligned}
\mathbb{P}\left(\sup_{k\geq m} |M_k - M_m| \geq \frac{b-a}{2}\right) &= \mathbb{P}\left(\sup_{k\geq 0} M'_k \geq \frac{(b-a)^2}{4}\right) \\
&\leq \frac{4}{(b-a)^2} \liminf_{k\to\infty} \mathbb{E}\left[M'_k\right] \\
&= \frac{4}{(b-a)^2} \sup_{k\geq 0} \mathbb{E}\left[M'_k\right] \\
&= \frac{4}{(b-a)^2} \sup_{k\geq 0} \sum_{i=m}^{k+m} \mathbb{E}\left[d_i^2\right] \\
&= \frac{4}{(b-a)^2} \sum_{i\geq m} \mathbb{E}\left[d_i^2\right].
\end{aligned}
$$

Enfin, le terme de droite tend vers 0 lorsque $m \to +\infty$, ce qui démontre que $\mathbb{P}(A_{a,b}) = 0$, et donc que $\mathbb{P}(\cup_{a,b\in\mathbb{Q}} A_{a,b}) = 0$, toute union dénombrable d'ensembles négligeables restant négligeable.
Pour démontrer la seconde assertion, notons M_∞ la limite p.s. de la martingale $(M_n)_{n\in\mathbb{N}}$. Par le lemme de Fatou, on a

$$
\mathbb{E}\left[M_\infty^2\right] \leq \liminf_{n\to+\infty} \mathbb{E}\left[M_n^2\right] \leq \sup_{n\in\mathbb{N}} \mathbb{E}\left[M_n^2\right] < +\infty,
$$

d'où M_∞ est non seulement finie mais aussi dans L^2. Enfin, la convergence dans L^2 est immédiate d'après ce qui précède :

$$
\mathbb{E}\left[(M_\infty - M_n)^2\right] = \sum_{k\geq n+1} \mathbb{E}\left[d_k^2\right] \underset{n\to+\infty}{\longrightarrow} 0.
$$

■

Notons que d'après la formule que nous avons déjà vue,

$$\mathbb{E}\left[M_n^2\right] = \mathbb{E}\left[M_0^2\right] + \mathbb{E}\left[[M,M]_n\right], \quad n \in \mathbb{N},$$

une martingale de carré intégrable est bornée dans L^2 si et seulement si $[M,M]_\infty \in L^1$. À présent, donnons un résultat de convergence pour les sous et surmartingales. En particulier, une martingale positive étant une surmartingale positive, elle est p.s. convergente.

Théorème 2.2.2. *Une sous-martingale bornée, tout comme une surmartingale positive, est p.s. convergente.*

Démonstration. Commençons par démontrer la première assertion. Soit $(X_n)_{n\in\mathbb{N}}$ une sous-martingale bornée en valeur absolue par K (une constante déterministe indépendante de n) et notons $X_n = M_n + A_n$ sa décomposition de Doob. On a

$$\mathbb{E}[A_n] = \mathbb{E}[X_n] - \mathbb{E}[M_0] \leq 2K,$$

et donc par convergence monotone, $\mathbb{E}[A_\infty] \leq 2K$, d'où $A_\infty < +\infty$ p.s. On pose

$$\tau_p = \inf\{n \geq 0 : A_{n+1} > p\}, \quad p \in \mathbb{N}.$$

On montre facilement que τ_p est un temps d'arrêt et, par le théorème d'arrêt, le processus arrêté $(M_{n\wedge\tau_p})_{n\in\mathbb{N}}$ est une martingale. Par ailleurs on a $A_{n\wedge\tau_p} \leq p$ et donc

$$|M_{n\wedge\tau_p}| = |X_{n\wedge\tau_p} - A_{n\wedge\tau_p}| \leq K + p,$$

c'est-à-dire que la martingale arrêtée est bornée, donc p.s. convergente par le théorème 2.2.1. Comme on a $X_n = X_{n\wedge\tau_p}$ lorsque $\tau_p = +\infty$, on en déduit que la sous-martingale $(X_n)_{n\in\mathbb{N}}$ converge p.s. sur l'événement $\{\tau_p = +\infty\}$. En utilisant la convergence monotone à deux reprises, on a pour tout $p \in \mathbb{N}$,

$$\begin{aligned}
\mathbb{P}(\tau_p < +\infty) &= \lim_{n\to+\infty} \mathbb{P}(\tau_p \leq n) \\
&= \lim_{n\to+\infty} \mathbb{P}(A_{n+1} > p) \\
&= \mathbb{P}(A_\infty > p),
\end{aligned}$$

probabilité qui tend vers 0 lorsque p tend vers l'infini, la v.a. A_∞ étant p.s. finie. D'où la convergence monotone entraîne que $\mathbb{P}(\cap_{p\in\mathbb{N}}\{\tau_p < +\infty\}) = 0$, c'est-à-dire que $\cup_{p\in\mathbb{N}}\{\tau_p = +\infty\}$ est un événement de probabilité 1. Finalement, la sous-martingale $(X_n)_{n\in\mathbb{N}}$ convergeant p.s. sur l'événement $\cup_{p\in\mathbb{N}}\{\tau_p = +\infty\}$, on en déduit qu'elle converge p.s., ce qui achève la démonstration de la première assertion.

À présent si $(X_n)_{n\in\mathbb{N}}$ est une surmartingale positive, la fonction exponentielle étant convexe et croissante, le processus $(e^{-X_n})_{n\in\mathbb{N}}$ est une sous-martingale à valeurs dans l'intervalle $[0,1]$. Par ce qui précède on obtient la convergence p.s. de cette sous-martingale, et donc celle de $(X_n)_{n\in\mathbb{N}}$, qui peut néanmoins tendre vers $+\infty$. Si l'on note X_∞ cette limite, le lemme de Fatou nous donne

$$\mathbb{E}[X_\infty] \leq \liminf_{n\to+\infty} \mathbb{E}[X_n] \leq \mathbb{E}[X_0],$$

la dernière inégalité étant due à la propriété de surmartingale. On en déduit que $X_\infty < +\infty$ p.s., ce qui termine la preuve. ■

Si l'on reprend la démonstration précédente en remplaçant la décomposition de Doob de la sous-martingale positive $(X_n)_{n\in\mathbb{N}}$ par celle de $(M_n^2)_{n\in\mathbb{N}}$, où $(M_n)_{n\in\mathbb{N}}$ est une martingale de carré intégrable, alors en notant le temps d'arrêt

$$\tau_p = \inf\{n \geq 0 : [M,M]_{n+1} > p\},$$

on montre que la martingale arrêtée $(M_{n\wedge\tau_p})_{n\in\mathbb{N}}$ est bornée dans L^2, donc est p.s. convergente. Ainsi, $(M_n)_{n\in\mathbb{N}}$ converge p.s. sur l'événement $\{\tau_p = +\infty\}$ et comme on a

$$\bigcup_{p\in\mathbb{N}} \{\tau_p = +\infty\} = \{[M,M]_\infty < +\infty\},$$

on en déduit que la martingale converge p.s. si $[M,M]_\infty < +\infty$ p.s. On a démontré le résultat suivant.

Théorème 2.2.3. *Soit $(M_n)_{n\in\mathbb{N}}$ une martingale de carré intégrable. Si l'on a $[M,M]_\infty < +\infty$ p.s., la martingale est p.s. convergente.*

Un autre critère intéressant pour assurer la convergence p.s. des sous-martingales est la bornitude dans L^1. Notons que contrairement au cas L^2 (ou même L^p comme on le verra ci-dessous), la bornitude dans L^1 n'entraîne pas en général la convergence dans L^1.

Théorème 2.2.4. *Soit $(X_n)_{n\in\mathbb{N}}$ une sous-martingale bornée dans L^1, i.e.*

$$\sup_{n\in\mathbb{N}} \mathbb{E}[|X_n|] < +\infty.$$

Alors le processus $(X_n)_{n\in\mathbb{N}}$ est p.s. convergent.

Démonstration. Pour démontrer ce résultat, il suffit de montrer que $(X_n)_{n\in\mathbb{N}}$ s'écrit comme la différence d'une martingale positive et d'une surmartingale positive, puis d'appliquer le théorème précédent. Soit $X_n^+ = M_n + A_n$ la décomposition de Doob de la sous-martingale $(X_n^+)_{n\in\mathbb{N}}$. On a

$$\mathbb{E}[A_\infty] = \sup_{n\in\mathbb{N}} \mathbb{E}[A_n] \leq \mathbb{E}[|M_0|] + \sup_{n\in\mathbb{N}} \mathbb{E}[|X_n|] < +\infty.$$

Ainsi, le processus $(Y_n)_{n\in\mathbb{N}}$ défini par $Y_n = M_n + \mathbb{E}\left[A_\infty \mid \mathcal{F}_n\right]$ est une martingale comme somme de deux martingales. Par ailleurs, comme le processus $(A_n)_{n\in\mathbb{N}}$ est croissant et adapté (car prévisible), on a

$$Y_n = M_n + \mathbb{E}\left[A_\infty \mid \mathcal{F}_n\right] \geq M_n + A_n = X_n^+ \geq 0,$$

c'est-à-dire que $(Y_n)_{n\in\mathbb{N}}$ est une martingale positive. Enfin, le processus $(Y_n - X_n)_{n\in\mathbb{N}}$ est une surmartingale comme somme de deux surmartingales, qui est positive car $Y_n \geq X_n^+ \geq X_n$. La démonstration est établie. ∎

À présent, nous sommes en mesure d'énoncer la version L^p des théorèmes de convergence pour les martingales.

Théorème 2.2.5. *Étant donné $p > 1$, soit $(M_n)_{n\in\mathbb{N}}$ une martingale bornée dans L^p, i.e.*

$$\sup_{n\in\mathbb{N}} \mathbb{E}\left[|M_n|^p\right] < +\infty.$$

Alors le processus $(M_n)_{n\in\mathbb{N}}$ converge p.s. et dans L^p vers une v.a. $M_\infty \in L^p$.

Démonstration. La martingale $(M_n)_{n\in\mathbb{N}}$ étant bornée dans L^p, elle est donc bornée dans L^1 par l'inégalité de Hölder, donc p.s. convergente par le théorème précédent. Ainsi, notons M_∞ sa limite p.s. Par ailleurs, d'après l'inégalité de Doob L^p, la v.a. $M_\infty^* = \sup_{n\in\mathbb{N}} |M_n|$ appartient à l'espace L^p. Grâce à l'inégalité suivante :

$$|M_n - M_\infty| \leq 2M_\infty^*,$$

on peut appliquer le théorème de convergence dominée pour obtenir la convergence dans L^p, à savoir

$$\lim_{n\to+\infty} \mathbb{E}\left[|M_n - M_\infty|^p\right] = 0.$$

∎

2.3 Le cas de la convergence L^1

Étudions en détail le cas de la convergence L^1 qui est plus délicat et pour lequel nous avons une complète caractérisation. Pour ce faire, introduisons tout d'abord le concept d'intégrabilité uniforme.

Définition 2.3.1. *Une famille de v.a. $(X_i)_{i\in I}$ est dite uniformément intégrable si*

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

L'intérêt d'introduire cette notion est qu'elle va nous permettre d'obtenir la convergence dans L^1 à partir de la convergence p.s., lorsque le théorème de convergence dominée ne s'applique pas.

Proposition 2.3.2. *Si une suite $(X_n)_{n\in\mathbb{N}}$ est uniformément intégrable et converge p.s. vers une v.a. X_∞, alors $X_\infty \in L^1$ et la convergence a aussi lieu dans L^1.*

Démonstration. Notons pour tout $a > 0$ la quantité

$$\rho(a) := \sup_{n\in\mathbb{N}} \mathbb{E}\left[|X_n|\, 1_{\{|X_n|>a\}}\right].$$

Tout d'abord par le lemme de Fatou,

$$\begin{aligned} \mathbb{E}\left[|X_\infty|\, 1_{\{|X_\infty|>a\}}\right] &\leq \liminf_{n\to+\infty} \mathbb{E}\left[|X_n|\, 1_{\{|X_n|>a\}}\right] \\ &\leq \rho(a), \end{aligned}$$

ce qui entraîne que $X_\infty \in L^1$ et même $\mathbb{E}[|X_\infty|] \leq \rho(a) + a$. Maintenant, on majore $|X_n - X_\infty|$ de la manière suivante :

$$|X_n - X_\infty| \leq |X_n - X_\infty|\, 1_{\{|X_n|\leq a\}} + |X_\infty|\, 1_{\{|X_n|>a\}} + |X_n|\, 1_{\{|X_n|>a\}}.$$

Les deux premiers termes sont majorés respectivement par $a + |X_\infty|$ et $|X_\infty|$, tous deux dans L^1, donc le théorème de convergence dominée entraîne que

$$\mathbb{E}\left[|X_n - X_\infty|\, 1_{\{|X_n|\leq a\}}\right] \underset{n\to+\infty}{\longrightarrow} 0 \quad \text{et} \quad \mathbb{E}\left[|X_\infty|\, 1_{\{|X_n|>a\}}\right] \underset{n\to+\infty}{\longrightarrow} \mathbb{E}\left[|X_\infty|\, 1_{\{|X_\infty|>a\}}\right] \leq \rho(a),$$

tandis que le troisième et dernier terme est majoré en espérance par $\rho(a)$. Ainsi, on obtient que

$$\limsup_{n\to+\infty} \mathbb{E}\left[|X_n - X_\infty|\right] \leq 2\rho(a),$$

et nous concluons la preuve en observant que $\rho(a) \to 0$ lorsque $a \to +\infty$. ∎

La proposition suivante nous donne des critères intéressants en pratique pour établir la propriété d'intégrabilité uniforme.

Proposition 2.3.3. *Les assertions suivantes sont vérifiées.*

(i) Si une famille de v.a. $(X_i)_{i\in I}$ est bornée par une v.a. intégrable, alors elle est uniformément intégrable.

(ii) S'il existe $p > 1$ tel que la famille de v.a. $(X_i)_{i\in I}$ soit bornée dans L^p, i.e.

$$\sup_{i\in I} \mathbb{E}\left[|X_i|^p\right] < +\infty,$$

alors elle est uniformément intégrable.

(iii) *Une famille de v.a.* $(X_i)_{i\in I}$ *est uniformément intégrable si et seulement si elle est bornée dans* L^1 *et équicontinue, i.e. pour tout* $\varepsilon > 0$, *il existe* $\eta > 0$ *tel que*

$$\mathbb{P}(A) < \eta \Longrightarrow \sup_{i\in I} \mathbb{E}\left[|X_i|\, 1_A\right] < \varepsilon.$$

Démonstration. Le point (i) est immédiat en utilisant le théorème de convergence dominée, tandis que pour le point (ii), on utilise les inégalités de Hölder puis de Chebyshev. Établissons le point (iii). On a pour tout $A \in \mathcal{A}$ et tout $a > 0$,

$$\mathbb{E}[|X_i|\, 1_A] \leq a\, \mathbb{P}(A) + \mathbb{E}[|X_i|\, 1_{\{|X_i|>a\}}].$$

Ainsi, en prenant $A = \Omega$, ceci entraîne par intégrabilité uniforme que $\sup_{i\in I} \mathbb{E}[|X_i|] < +\infty$. Montrons maintenant l'équicontinuité. Soit $\varepsilon > 0$. On choisit alors a de telle sorte que $\sup_{i\in I} \mathbb{E}[|X_i|\, 1_{\{|X_i|>a\}}] < \varepsilon/2$. De l'inégalité ci-dessus, il s'ensuit

$$\sup_{i\in I} \mathbb{E}[|X_i|\, 1_A] \leq a\, \mathbb{P}(A) + \frac{\varepsilon}{2}.$$

Enfin, le choix de $\eta := \varepsilon/(2a)$ nous permet d'établir l'équicontinuité.
Réciproquement, posons $M := \sup_{i\in I} \mathbb{E}\left[|X_i|\right] < +\infty$. Alors pour tout $i \in I$, l'inégalité de Chebyshev nous dit que

$$\sup_{i\in I} \mathbb{P}(|X_i| > a) \leq \frac{M}{a},$$

et par suite, le choix de a assez grand et l'équicontinuité achèvent la démonstration de l'intégrabilité uniforme. ∎

En particulier, si une v.a. est intégrable, alors on peut adapter la preuve précédente pour montrer, en utilisant le théorème de convergence dominée, qu'elle est équicontinue. Ceci se généralise à une famille finie de variables aléatoires.

À présent, nous sommes en mesure d'établir la caractérisation de la convergence dans L^1 des martingales.

Théorème 2.3.4. *Soit* $(M_n)_{n\in\mathbb{N}}$ *une martingale. Alors les assertions suivantes sont équivalentes :*

(i) *il existe une v.a.* $Z \in L^1$ *telle que*

$$M_n = \mathbb{E}[Z \mid \mathcal{F}_n], \quad n \in \mathbb{N}.$$

(ii) *la martingale* $(M_n)_{n\in\mathbb{N}}$ *est uniformément intégrable.*

(iii) *la martingale* $(M_n)_{n\in\mathbb{N}}$ *converge p.s. et dans* L^1 *vers une variable* $M_\infty \in L^1$.

Dans ce cas, on a $Z = M_\infty$ *et la martingale* $(M_n)_{n\in\mathbb{N}}$ *est dite fermée (par* M_∞*).*

Démonstration. $(i) \Rightarrow (ii)$: étant donné un entier n, considérons l'ensemble $A_n := \{|M_n| > a\}$. D'après les inégalités de Chebyshev puis de Jensen, on a

$$\sup_{n\in\mathbb{N}} \mathbb{P}(A_n) \leq \frac{\sup_{n\in\mathbb{N}} \mathbb{E}[|M_n|]}{a} \leq \frac{\mathbb{E}[|Z|]}{a}.$$

Par ailleurs, A_n étant $\mathcal{F}_n$-mesurable, l'inégalité de Jensen entraîne aussi

$$\mathbb{E}[|M_n|\, 1_{A_n}] \leq \mathbb{E}[|Z|\, 1_{A_n}].$$

La v.a. Z étant équicontinue car intégrable, pour tout $\varepsilon > 0$, il existe $\eta > 0$ tel que pour tout $n \in \mathbb{N}$,

$$\mathbb{P}(A_n) < \eta \Longrightarrow \mathbb{E}[|Z|\, 1_{A_n}] < \varepsilon.$$

Il suffit enfin de choisir a assez grand de sorte que $\sup_{n\in\mathbb{N}} \mathbb{P}(A_n) < \eta$, et on aura alors $\mathbb{E}[|M_n|\, 1_{A_n}] < \epsilon$ pour tout $n \in \mathbb{N}$, autrement dit l'intégrabilité uniforme.

$(ii) \Rightarrow (iii)$: la martingale $(M_n)_{n\in\mathbb{N}}$ étant uniformément intégrable, elle est bornée dans L^1 par la proposition 2.3.3, et donc elle converge p.s. par le théorème 2.2.4. Une fois la convergence p.s. établie, la convergence dans L^1 est immédiate par la proposition 2.3.2.

$(iii) \Rightarrow (i)$: notons que pour tous $n, p \in \mathbb{N}$ et tout $A \in \mathcal{F}_n$, la propriété de martingale entraîne que

$$\begin{aligned} |\mathbb{E}[(M_\infty - M_n)\, 1_A]| &= |\mathbb{E}[(M_\infty - M_{n+p})\, 1_A]| \\ &\leq \mathbb{E}[|M_\infty - M_{n+p})|] \\ &\underset{p\to+\infty}{\longrightarrow} 0. \end{aligned}$$

Finalement, on obtient que $\mathbb{E}[(M_\infty - M_n)\, 1_A] = 0$ pour tout $A \in \mathcal{F}_n$, donc que $M_n = \mathbb{E}[M_\infty \mid \mathcal{F}_n]$. ∎

On en déduit par le théorème 2.2.5 que toute martingale bornée dans un L^p, où $p > 1$, est fermée.

2.4 Loi des grands nombres et théorème central limite pour les martingales

On a vu précédemment qu'une martingale de carré intégrable dont la variation quadratique admet une limite p.s. est p.s. convergente. En revanche, quel résultat de convergence peut-on établir lorsque la variation quadratique tend vers l'infini ? Pour obtenir un résultat de convergence, il faut renormaliser la martingale de carré intégrable par sa variation quadratique : nous rentrons dans le cadre de la loi des grands nombres. Avant de donner le résultat, rappelons le lemme de Kronecker qui est un résultat de convergence pour les séries numériques.

Lemme 2.4.1 (Kronecker)**.** *Soit $(a_n)_{n\in\mathbb{N}}$ une suite de nombres strictement positifs tendant vers l'infini à l'infini et $(x_n)_{n\in\mathbb{N}}$ une suite de nombres réels. Si la série de terme général x_n/a_n converge, alors*

$$\lim_{n\to+\infty} \frac{1}{a_n} \sum_{k=1}^{n} x_k = 0.$$

Démonstration. Soit $(u_n)_{n\in\mathbb{N}}$ la suite définie par

$$u_n = \sum_{k=1}^{n} \frac{x_k}{a_k}, \quad u_0 = 0,$$

qui est convergente par hypothèse. Notons u la somme de la série. On a

$$\frac{1}{a_n} \sum_{k=1}^{n} x_k = \frac{1}{a_n} \sum_{k=1}^{n} a_k(u_k - u_{k-1}) = u_n - \frac{1}{a_n} \sum_{k=1}^{n} u_{k-1}(a_k - a_{k-1}).$$

Pour montrer la convergence désirée, il nous faut montrer que la suite $(\rho_n)_{n\in\mathbb{N}}$ donnée par

$$\rho_n = \frac{1}{a_n} \sum_{k=1}^{n} u_{k-1}(a_k - a_{k-1}) - u,$$

tend vers 0 lorsque n tend vers l'infini. On a pour tout $p \in \{0, \cdots, n-1\}$,

$$\begin{aligned} |\rho_n| &\leq \frac{1}{a_n} \left| \sum_{k=1}^{p} u_{k-1}(a_k - a_{k-1}) \right| + \frac{1}{a_n} \left| \sum_{k=p+1}^{n} (u_{k-1} - u)(a_k - a_{k-1}) \right| + \frac{a_p}{a_n}|u| \\ &\leq \frac{1}{a_n} \left| \sum_{k=1}^{p} u_{k-1}(a_k - a_{k-1}) \right| + \frac{|a_n - a_p|}{a_n} \sup_{k\geq p+1} |u_{k-1} - u| + \frac{a_p}{a_n}|u|. \end{aligned}$$

En passant à la limite en n, on a pour tout $p \in \mathbb{N}$,

$$\limsup_{n\to+\infty} |\rho_n| \leq 0 + \sup_{k\geq p+1} |u_{k-1} - u| + 0,$$

la suite $(a_n)_{n\in\mathbb{N}}$ tendant vers l'infini à l'infini. Enfin en passant à la limite en p dans l'inégalité, on obtient

$$\limsup_{n\to+\infty} |\rho_n| \leq \lim_{p\to+\infty} \sup_{k\geq p+1} |u_{k-1} - u| = \limsup_{p\to+\infty} |u_p - u| = 0,$$

d'où la conclusion. ∎

À présent, nous allons énoncer la loi des grands nombres (LGN) pour les martingales. La dénomination est due au fait que ce résultat est une généralisation de la LGN pour les v.a. i.i.d., sous l'hypothèse supplémentaire d'un moment d'ordre deux.

Théorème 2.4.2 (LGN)**.** *Soit $(M_n)_{n\in\mathbb{N}}$ une martingale de carré intégrable. Si p.s., $[M,M]_\infty = +\infty$, alors on a la convergence p.s. suivante :*

$$\lim_{n\to+\infty} \frac{M_n}{[M,M]_n} = 0.$$

Démonstration. Considérons le processus $(X_n)_{n\in\mathbb{N}}$ défini par $X_0 = 0$ et

$$X_n = \sum_{k=1}^{n} \frac{M_k - M_{k-1}}{1 + [M,M]_k}, \quad n \geq 1.$$

On remarque que $(X_n)_{n\in\mathbb{N}}$ est la transformation prévisible de la martingale $(M_n)_{n\in\mathbb{N}}$ par le processus prévisible et borné $(1/(1+[M,M]_n))_{n\in\mathbb{N}}$. Ainsi, par le théorème 1.2.2, le processus $(X_n)_{n\in\mathbb{N}}$ est une martingale de carré intégrable. De plus, le théorème 1.2.4 nous donne la variation quadratique de la transformation prévisible : pour tout $m \in \mathbb{N}^*$,

$$\begin{aligned}
[X,X]_m &= \sum_{i=1}^{m} \frac{[M,M]_i - [M,M]_{i-1}}{(1+[M,M]_i)^2} \\
&\leq \sum_{i=1}^{m} \frac{[M,M]_i - [M,M]_{i-1}}{(1+[M,M]_i)(1+[M,M]_{i-1})} \\
&= \sum_{i=1}^{m} \left(\frac{1}{1+[M,M]_{i-1}} - \frac{1}{1+[M,M]_i} \right) \\
&= \frac{1}{1+[M,M]_0} - \frac{1}{1+[M,M]_m} \\
&\leq 1,
\end{aligned}$$

la variation quadratique étant un processus croissant. Ainsi, par le théorème 2.2.3, la martingale $(X_n)_{n\in\mathbb{N}}$ est p.s. convergente. Enfin, comme on a supposé que $[M,M]_\infty = +\infty$ p.s., on peut appliquer le lemme de Kronecker aux suites aléatoires $a_n = 1 + [M,M]_n$ et $x_n = M_n - M_{n-1}$ ($x_0 = M_0$) et l'on obtient que la suite

$$\frac{M_n}{[M,M]_n} = \left(\frac{M_n - M_0}{1+[M,M]_n} + \frac{M_0}{1+[M,M]_n} \right) \times \frac{1+[M,M]_n}{[M,M]_n},$$

converge p.s. vers 0, ce qui achève la démonstration de la LGN pour les martingales. ■

Comme nous l'avons mentionné ci-dessus, le résultat précédent généralise la LGN pour les v.a. i.i.d., au sens où elles ne sont plus nécessairement de même loi. Le prix à payer pour cette amélioration est de supposer qu'elles sont de carré intégrable et non seulement intégrables.

Théorème 2.4.3. *Soit $(X_n)_{n\in\mathbb{N}^*}$ une suite de v.a. indépendantes (non nécessairement de même loi) et de carré intégrable. Posons $m_n = \mathbb{E}[X_n]$ et $\sigma_n^2 = \mathrm{Var}(X_n)$. Si la série numérique de terme général $(\sigma_n)_{n\in\mathbb{N}^*}$ diverge, alors on a la convergence p.s. suivante :*

$$\lim_{n\to+\infty} \frac{1}{\sum_{k=1}^n \sigma_k^2} \times \sum_{k=1}^n (X_k - m_k) = 0.$$

Démonstration. Soit $(M_n)_{n\in\mathbb{N}^*}$ le processus défini par

$$M_n = \sum_{k=1}^n (X_k - m_k), \quad n \in \mathbb{N}^*.$$

Le processus $(M_n)_{n\in\mathbb{N}^*}$ étant la somme de v.a. indépendantes, intégrables et centrées, nous avons vu que $(M_n)_{n\in\mathbb{N}^*}$ était une martingale pour la filtration $\mathcal{F}_n = \sigma(X_1, \cdots, X_n)$, $n \in \mathbb{N}^*$, qui de surcroît est de carré intégrable. Pour tout $n \in \mathbb{N}^*$, on a

$$\begin{aligned}
[M, M]_n &= \sum_{k=1}^n \mathbb{E}\left[(M_k - M_{k-1})^2 \mid \mathcal{F}_{k-1}\right] \\
&= \sum_{k=1}^n \mathbb{E}\left[(X_k - m_k)^2 \mid \mathcal{F}_{k-1}\right] \\
&= \sum_{k=1}^n \mathbb{E}\left[(X_k - m_k)^2\right] \qquad \text{car } \sigma(X_k) \text{ et } \mathcal{F}_{k-1} \text{ sont indépendantes} \\
&= \sum_{k=1}^n \mathrm{Var}(X_k) \\
&= \sum_{k=1}^n \sigma_k^2.
\end{aligned}$$

Enfin, le théorème 2.4.2 achève la démonstration de cette LGN. ■

Une fois la LGN pour les martingales établie, on peut se demander si un théorème central limite (TCL) a lieu. Il s'avère que la réponse à cette question est positive et l'on a même plusieurs résultats à notre disposition selon les hypothèses que l'on formule sur les martingales. Cependant, nous n'allons énoncer que celui dont la démonstration n'est pas trop pénible. Bien évidemment, on retrouve le TCL pour les variables i.i.d. centrées avec le choix de la suite $a_n = n$, sous une hypothèse supplémentaire sur les accroissements.

Théorème 2.4.4 (TCL)**.** *Soit M une martingale dont les accroissements satisfont la propriété de convergence p.s. suivante :*

$$\sup_{n\in\mathbb{N}^*} \mathbb{E}\left[|M_n - M_{n-1}|^3 \mid \mathcal{F}_{n-1}\right] < +\infty, \quad p.s.,$$

et soit $(a_n)_{n\in\mathbb{N}}$ une suite positive tendant vers l'infini lorsque n tend vers l'infini. On suppose qu'il existe une constante $\sigma^2 > 0$ telle que l'on ait la convergence p.s. suivante :

$$\lim_{n\to+\infty} \frac{[M,M]_n}{a_n} = \sigma^2.$$

Alors on a les convergences en loi

$$\frac{M_n}{\sqrt{a_n}} \underset{n\to+\infty}{\Longrightarrow} \mathcal{N}(0,\sigma^2) \quad \text{et} \quad \sqrt{a_n}\,\frac{M_n}{[M,M]_n} \underset{n\to+\infty}{\Longrightarrow} \mathcal{N}(0,\sigma^{-2}).$$

Démonstration. La première convergence en loi est le TCL proprement dit tandis que la seconde en est une conséquence grâce au point (iii) du lemme de Slutsky énoncé ci-dessous après la fin de la démonstration. Ainsi, démontrons seulement la première convergence en loi et considérons pour simplifier la suite $a_n = n$ (le cas général est identique).

La démonstration se fait en plusieurs étapes.

∘ Étape 1 : tout d'abord notons K le supremum de l'énoncé et pour tout $n \in \mathbb{N}^*$ la quantité

$$\sigma_n^2 := \mathbb{E}\left[(M_n - M_{n-1})^2 \mid \mathcal{F}_{n-1}\right].$$

L'inégalité de Hölder pour les espérances conditionnelles appliquée avec $p = 3/2$ et $q = 3$ indique que l'on a p.s.,

$$\sup_{k\in\mathbb{N}^*} \sigma_k^2 \leq K^{2/3},$$

d'où l'inégalité p.s.

$$[M,M]_n = \sum_{k=1}^n \sigma_k^2 \leq nK^{2/3}.$$

∘ Étape 2 : dans la suite, on supposera sans perte de généralité que $K = 1$. Soient $x \in \mathbb{R}$, $y \in [0,1]$ des paramètres fixés et posons

$$\phi_{x,y}(t) := \exp\left(itx + \frac{t^2 y}{2}\right), \quad t \in [-1,1].$$

Par la formule de Taylor-Lagrange, on a

$$\phi_{x,y}(t) \;=\; \phi_{x,y}(0) + t\,\phi'_{x,y}(0) + \frac{t^2}{2}\,\phi''_{x,y}(0) + \frac{t^3}{3!}\,R_{x,y}(t)$$

$$= \quad 1 + itx + \frac{t^2}{2}\,(y - x^2) + \frac{t^3}{6}\,R_{x,y}(t),$$

où $R_{x,y}$ est une fonction de t, dépendant de x et de y et telle que l'on ait

$$|R_{x,y}(t)| \;\leq\; \sup_{s\in[0,t]} |\phi'''_{x,y}(s)| \;\leq\; C\,\left(1 + |x|^3\right)\, e^{\frac{t^2}{2}},$$

avec $C > 0$ une constante déterministe dont la valeur n'a aucune importance (dans la suite, on notera C une telle constante, qui peut changer de ligne à ligne).
∘ Étape 3 : appliquons ceci à la suite de fonctions aléatoires $(Z_n)_{n\in\mathbb{N}^*}$ définie par

$$Z_n(t) := \phi_{M_n - M_{n-1},\sigma_n^2}(t) = \exp\left(it(M_n - M_{n-1}) + \frac{t^2\sigma_n^2}{2}\right), \quad t \in [-1,1].$$

Notons que l'étape 1 nous le permet car l'on a bien $\sigma_n^2 \in [0,1]$ pour tout $n \in \mathbb{N}^*$. On a alors

$$\begin{aligned}
\mathbb{E}\left[Z_n(t) \mid \mathcal{F}_{n-1}\right] \;&=\; 1 + it\,\underbrace{\mathbb{E}\left[M_n - M_{n-1} \mid \mathcal{F}_{n-1}\right]}_{=0 \text{ car } M \text{ martingale}} + \frac{t^2}{2}\,\underbrace{\mathbb{E}\left[\sigma_n^2 - (M_n - M_{n-1})^2 \mid \mathcal{F}_{n-1}\right]}_{=0 \text{ par définition}} \\
&\quad + \frac{t^3}{6}\,\mathbb{E}\left[R_{M_n - M_{n-1},\sigma_n^2}(t) \mid \mathcal{F}_{n-1}\right] \\
&=\; 1 + V_{n-1}(t),
\end{aligned}$$

où $V_{n-1}(t)$ est une v.a. $\mathcal{F}_{n-1}$ mesurable vérifiant l'inégalité p.s. :

$$|V_{n-1}(t)| \leq C\,|t|^3\, e^{\frac{t^2}{2}}.$$

∘ Étape 4 : à présent si l'on pose pour tout $n \in \mathbb{N}^*$ la fonction aléatoire

$$Y_n(t) := \phi_{M_n,[M,M]_n}(t) = \exp\left(itM_n + \frac{t^2[M,M]_n}{2}\right), \quad t \in [-1,1],$$

on a

$$\begin{aligned}
\mathbb{E}[Y_n(t)] \;&=\; \mathbb{E}\left[Y_{n-1}(t)\, Z_n(t)\right] \\
&=\; \mathbb{E}\left[Y_{n-1}(t)\, \mathbb{E}\left[Z_n(t) \mid \mathcal{F}_{n-1}\right]\right] \qquad \text{car } Y_{n-1}(t) \text{ est } \mathcal{F}_{n-1} \text{ mesurable} \\
&=\; \mathbb{E}\left[Y_{n-1}(t)\,(1 + V_n(t))\right] \\
&=\; \mathbb{E}\left[Y_{n-1}(t)\right] + \mathbb{E}\left[Y_{n-1}(t)\, V_n(t)\right],
\end{aligned}$$

et l'on démontre facilement par récurrence sur $n \in \mathbb{N}^*$ que

$$\mathbb{E}[Y_n(t)] = 1 + v_n(t),$$

où $v_n(t)$ satisfait l'inégalité
$$|v_n(t)| \leq C\, n|t|^3\, e^{\frac{nt^2}{2}}.$$
Ainsi on en déduit en remplaçant t par $t/\sqrt{n}$ (t est à présent fixé dans $\mathbb{R}$ et non plus dans $[-1,1]$ mais comme n a vocation à tendre vers l'infini on peut supposer que $t/\sqrt{n} \in [-1,1]$) et en passant à la limite dans l'égalité précédente que

$$\lim_{n\to+\infty} \mathbb{E}\left[Y_n\left(\frac{t}{\sqrt{n}}\right)\right] = 1.$$

∘ Étape 5 : par ailleurs on a

$$\begin{aligned} \mathbb{E}\left[\exp\left(it\frac{M_n}{\sqrt{n}} + \frac{t^2\sigma^2}{2}\right)\right] - 1 \; &= \; \mathbb{E}\left[\exp\left(it\frac{M_n}{\sqrt{n}}\right)\left(\exp\left(\frac{t^2\sigma^2}{2}\right) - \exp\left(\frac{t^2[M,M]_n}{2n}\right)\right)\right] \\ &\quad + \mathbb{E}\left[Y_n\left(\frac{t}{\sqrt{n}}\right)\right] - 1. \end{aligned}$$

Montrons que la première espérance dans le terme de droite tend vers 0 lorsque n tend vers l'infini. On a

$$\left|\exp\left(it\frac{M_n}{\sqrt{n}}\right)\left(\exp\left(\frac{t^2\sigma^2}{2}\right) - \exp\left(\frac{t^2[M,M]_n}{2n}\right)\right)\right| = \left|\exp\left(\frac{t^2\sigma^2}{2}\right) - \exp\left(\frac{t^2[M,M]_n}{2n}\right)\right|,$$

qui tend p.s. vers 0 lorsque n tend vers l'infini, la suite $[M,M]_n/n$ tendant p.s. vers σ^2 par hypothèse. De plus, comme la suite $[M,M]_n/n$ est bornée par 1 d'après l'étape 1, on a
$$\left|\exp\left(\frac{t^2\sigma^2}{2}\right) - \exp\left(\frac{t^2[M,M]_n}{2n}\right)\right| \leq \exp\left(\frac{t^2\sigma^2}{2}\right) + \exp\left(\frac{t^2}{2}\right),$$
cette dernière quantité étant bien intégrable car déterministe. Le théorème de convergence dominée et l'étape 4 nous donnent alors que

$$\lim_{n\to+\infty} \mathbb{E}\left[\exp\left(it\frac{M_n}{\sqrt{n}} + \frac{t^2\sigma^2}{2}\right)\right] = 1,$$

c'est-à-dire que
$$\lim_{n\to+\infty} \mathbb{E}\left[\exp\left(it\frac{M_n}{\sqrt{n}}\right)\right] = \exp\left(-\frac{t^2\sigma^2}{2}\right).$$
Enfin, le théorème de convergence de Lévy (la convergence des fonctions caractéristiques est équivalente à la convergence en loi) nous permet d'obtenir le résultat désiré, à savoir la convergence en loi de la suite $(M_n/\sqrt{n})_{n\in\mathbb{N}^*}$ vers une v.a. gaussienne centrée et de variance σ^2. ∎

Pour terminer cette partie, on rappelle sans démonstration le lemme de Slutsky.

Lemme 2.4.5 (Lemme de Slutsky)**.** *Soient* $(X_n)_{n\in\mathbb{N}_*}$ *et* $(Y_n)_{n\in\mathbb{N}_*}$ *deux suites de v.a. convergeant en loi respectivement vers un nombre* $c \in \mathbb{R}$ *et une v.a.* Y*. Alors*

(*i*) *la somme* $X_n + Y_n$ *converge en loi vers* $c + Y$*.*
(*ii*) *le produit* $X_n Y_n$ *converge en loi vers* cY*.*
(*iii*) *le ratio* Y_n/X_n *converge en loi vers* Y/c *dès que* $c \neq 0$*.*

Dans l'énoncé, l'hypothèse selon laquelle X_n converge vers une constante est cruciale. En effet, si la limite était une v.a., le résultat ne serait plus valide et il faudrait une hypothèse plus forte comme la convergence en loi du couple (X_n, Y_n) pour que le résultat reste vrai. Par ailleurs, le lemme reste valide lorsque l'on remplace toutes les convergences en loi par des convergences en probabilité.

2.5 Une ouverture vers les statistiques

Terminons ce cours sur les martingales par une application des LGN et TCL vus ci-dessus à l'estimation paramétrique par maximum de vraisemblance dans un modèle selon lequel les v.a. considérées ne sont plus i.i.d. mais forment une chaîne de Markov. Avant de rentrer dans les détails, rappelons brièvement le contexte de l'estimation statistique d'un paramètre inconnu par maximum de vraisemblance. On se donne un n-échantillon $(X_1, \ldots, X_n)$ d'observations dépendant d'un paramètre inconnu $\theta \in \Theta$, c'est-à-dire une suite de v.a. (i.i.d. ou non) dont la loi jointe P_θ dépend de θ. L'ensemble Θ, quant à lui, est un sous-ensemble de $\mathbb{R}$ (ou de $\mathbb{R}^d$, mais dans ce qui va suivre nous ne considérerons que la dimension 1) supposé implicitement compact ou borné pour assurer des résultats de régularité suffisants (sur lesquels nous n'insisterons pas). Rappelons que $\mathbb{P}$ est une probabilité sur $(\Omega, \mathcal{A})$ tandis que P_θ est une probabilité sur l'ensemble des valeurs prises par le vecteur $(X_1, \ldots, X_n)$ muni de sa tribu naturelle (la tribu borélienne produit dans le cas d'un espace continu et l'ensemble des parties produit dans le cas discret d'un ensemble fini ou dénombrable). On suppose aussi que la famille $(P_\theta)_{\theta\in\Theta}$ vérifie la condition d'identifiabilité, c'est-à-dire que l'application $\theta \mapsto P_\theta$ est injective : si deux lois de probabilités P_{θ_1} et P_{θ_2} sont égales alors $\theta_1 = \theta_2$.

Définition 2.5.1. *On appelle vraisemblance du modèle la fonction* $L_n : (x_1, \ldots, x_n, \theta) \mapsto L_n(x_1, \ldots, x_n, \theta)$ *définie :*

- dans le cas de v.a. discrètes par $L_n(x_1, \ldots, x_n, \theta) = P_\theta(\{x_1, \ldots, x_n\})$ *;*

- dans le cas continu par la densité de la loi jointe du n*-échantillon par rapport à la mesure de Lebesgue, c'est-à-dire* $L_n(x_1, \ldots, x_n, \theta) = f_\theta(x_1, \ldots, x_n)$*.*

La v.a. obtenue en appliquant la fonction $(x_1, \ldots, x_n) \mapsto$ *argmax*$_{\theta\in\Theta}$ $L_n(x_1, \ldots, x_n, \theta)$ *au*

n-échantillon $(X_1, \ldots, X_n)$ s'appelle l'estimateur du maximum de vraisemblance (EMV) du paramètre θ. On le note $\hat{\theta}_n$.

Dans le cas où les v.a. sont indépendantes, la loi jointe étant le produit des lois marginales, la vraisemblance devient simplement le produit des lois marginales. En revanche ce n'est plus le cas si les v.a. ne sont plus indépendantes comme lorsque le n-échantillon représente les termes successifs d'une chaîne de Markov. Notons par ailleurs que l'EMV peut être unique, ne pas être unique, ou même ne pas exister. Lorsque la vraisemblance est strictement positive et grâce à la croissance stricte du logarithme népérien (noté log dans la suite), il est équivalent (et souvent plus simple en pratique) de maximiser la log-vraisemblance, c'est-à-dire le logarithme népérien de la vraisemblance (dans le cas indépendant le produit se transforme en somme, ce qui est plus simple à dériver). La recherche de l'EMV est alors un problème d'optimisation de la log-vraisemblance : si c'est une fonction assez régulière de θ, alors il s'agit :

- de trouver un point critique, i.e. un point pour lequel la dérivée de la log-vraisemblance s'annule, puis

- de vérifier que la dérivée seconde en ce point est négative (il s'agit donc d'un maximum local) puis

- de montrer que ce maximum local est en fait global. Ce dernier point est en général assuré par l'éventuelle concavité de la log-vraisemblance.

Dans le cas des v.a. i.i.d. et sous des hypothèses de régularité raisonnables, l'EMV est un estimateur consistant de θ, c'est-à-dire qu'il converge en probabilité vers θ lorsque la taille n de l'échantillon tend vers l'infini. De plus, il est asymptotiquement normal au sens de la convergence en loi vers une v.a. gaussienne centrée dont la variance est l'inverse d'une quantité strictement positive appelée l'information de Fisher. En particulier l'EMV atteint à la limite la borne de Cramer-Rao : il est asymptotiquement sans biais et de variance minimale : on dit qu'il est asymptotiquement efficace.

Dans le cas des chaînes de Markov, il existe des résultats analogues sous de bonnes hypothèses de convergence en temps long vers une mesure invariante. Plutôt que d'énoncer un résultat général, regardons ce que l'on obtient sur un exemple particulier et voyons comment la théorie des martingales entre en jeu. L'exemple que nous allons étudier fait partie de la classe importante des processus auto-régressifs d'ordre 1, notés processus AR(1), intervenant comme modèle de régression pour des séries temporelles (dans lequel la série est expliquée par ses valeurs passées plutôt que par d'autres variables). Dans la suite on supposera la condition initiale X_0 déterministe par simplicité, mais ce qui va suivre reste valide dans le cas d'une v.a. quelconque, l'important étant que l'on connaisse sa loi. En effet, la valeur de la chaîne au temps 0 est observée en pratique et peut donc être considérée comme connue (sa loi ne dépendra pas du paramètre inconnu θ).

Ainsi, soit $(U_n)_{n\geq 1}$ une suite i.i.d. de loi commune $\mathcal{N}(0,1)$ et considérons la suite

$(X_n)_{n\geq 0}$ définie par récurrence par $X_0 = x \in \mathbb{R}^*$ et

$$X_{n+1} \;=\; \theta\, X_n + U_{n+1}, \quad n \in \mathbb{N},$$

où $\theta \in]-1,1[$ est un paramètre inconnu que l'on souhaite estimer. La suite $(X_n)_{n\geq 0}$ s'écrivant sous la forme

$$X_{n+1} = f(X_n, U_{n+1}),$$

où $f : \mathbb{R}\times\mathbb{R} \to \mathbb{R}$ est la fonction borélienne $f(x,u) = \theta x + u$, et les $(U_n)_{n\geq 1}$ étant i.i.d. et (évidemment) indépendantes de X_0, c'est une chaîne de Markov homogène qui est la version à temps discret d'un processus stochastique à temps continu appelé processus d'Ornstein-Uhlenbeck. Il n'est pas difficile de montrer que le vecteur $(X_1,\ldots,X_n)$ des observations est gaussien (la condition d'identifiabilité est alors immédiatement vérifiée). Plus précisément, comme la densité jointe est le produit des densités conditionnelles (rappelez-vous votre cours sur les chaînes de Markov), la densité jointe est

$$f_\theta(x_1,\ldots,x_n) \;=\; \frac{1}{(2\pi)^{n/2}}\exp\left(-\frac{1}{2}\sum_{i=1}^n (x_i - \theta\, x_{i-1})^2\right), \quad (x_1,\ldots,x_n)\in\mathbb{R}^n,$$

où par convention $x_0 = x$. Ainsi, la log-vraisemblance est donnée par

$$\log L_n(x_1,\ldots,x_n,\theta) = -\frac{n}{2}\log(2\pi) - \frac{1}{2}\sum_{i=1}^n (x_i - \theta\, x_{i-1})^2, \quad (x_1,\ldots,x_n)\in\mathbb{R}^n.$$

On voit directement que cette fonction est une parabole en θ dont les branches pointent vers le bas : il s'agit d'une fonction concave et régulière de θ, qui admet donc un unique maximum. Sa dérivée est

$$\frac{\partial}{\partial\theta}\log L_n(x_1,\ldots,x_n,\theta) = -\sum_{i=1}^n x_{i-1}\,(\theta\, x_{i-1} - x_i),$$

et le θ pour lequel cette quantité s'annule vaut

$$\frac{\sum_{i=1}^n x_{i-1}\, x_i}{\sum_{i=1}^n x_{i-1}^2}.$$

Notons que ce ratio est bien défini, le dénominateur étant strictement positif ($x_0 = x \neq 0$). Ainsi l'EMV est égal à

$$\hat\theta_n \;=\; \frac{\sum_{i=1}^n X_{i-1}\, X_i}{\sum_{i=1}^n X_{i-1}^2} \;=\; \theta + \frac{\sum_{i=1}^n X_{i-1}\, U_i}{\sum_{i=1}^n X_{i-1}^2} \;=\; \theta + \frac{M_n}{[M,M]_n},$$

où M est défini par la suite

$$M_n := \sum_{i=1}^{n} X_{i-1}\, U_i, \quad n \in \mathbb{N}^*, \quad M_0 := 0.$$

La notation M n'est pas innocente : on montre que M est une martingale de carré intégrable. Nous y voilà : la consistance et la potentielle normalité asymptotique de l'EMV dépendent du comportement du ratio martingale/variation quadratique. Pour la consistance, on souhaite utiliser la LGN et pour ce faire, il faut que la variation quadratique tende p.s. vers l'infini lorsque n tend vers l'infini. Remarquons que cette dernière est, lorsque l'on divise par n, de la forme

$$\frac{[M,M]_n}{n} = \frac{1}{n} \sum_{i=1}^{n} g(X_{i-1}), \qquad \text{avec} \qquad g(u) = u^2.$$

Ainsi, si la chaîne de Markov est irréductible et récurrente positive, alors la chaîne possède une unique probabilité invariante π et dans le cas où $g \in L^1(\pi)$, la somme renormalisée ci-dessus converge p.s. vers l'intégrale de la fonction g par rapport à π : c'est le théorème ergodique. L'hypothèse de la LGN est alors satisfaite et l'on obtient ainsi la consistance de l'EMV, la convergence p.s. entraînant celle en probabilité. On peut montrer facilement que la chaîne ci-dessus est bien irréductible (on a $P(z,A) > 0$ pour tout $z \in \mathbb{R}$ et tout borélien A de mesure de Lebesgue strictement positive, les probabilités de transition étant des lois gaussiennes, donc de support $\mathbb{R}$ tout entier) et qu'elle admet pour unique probabilité invariante la loi gaussienne centrée et de variance $1/(1-\theta^2)$ (rappelons que θ est supposé être dans l'intervalle $]-1,1[$).
De même, pour la normalité asymptotique, on va utiliser le TCL. Tout d'abord on a bien p.s.

$$\lim_{n\to+\infty} \frac{[M,M]_n}{n} = \sqrt{\frac{1-\theta^2}{2\pi}} \int_{\mathbb{R}} u^2\, e^{-\frac{u^2(1-\theta^2)}{2}}\, du = \frac{1}{1-\theta^2} =: \sigma^2.$$

Néanmoins l'autre condition du TCL n'est pas vérifiée. En effet, comme X_{n-1} est $\mathcal{F}_{n-1}$-mesurable et U_n est indépendante de $\mathcal{F}_{n-1}$, on a

$$\mathbb{E}\left[|M_n - M_{n-1}|^3 \mid \mathcal{F}_{n-1} \right] \,=\, \mathbb{E}\left[|X_{n-1}\, U_n|^3 \mid \mathcal{F}_{n-1} \right] \,=\, |X_{n-1}|^3\, \mathbb{E}\left[|U_n|^3 \right],$$

quantité qui n'est pas bornée car X_{n-1} ne l'est pas (son support est $\mathbb{R}$ tout entier). Cependant le TCL a tout de même lieu lorsque cette condition est remplacée par la condition plus faible de la convergence p.s. suivante :

$$\lim_{n\to+\infty} \frac{1}{n^{3/2}} \sum_{k=1}^{n} \mathbb{E}\left[|M_k - M_{k-1}|^3 \mid \mathcal{F}_{k-1} \right] = 0,$$

condition qui est satisfaite par notre modèle grâce au théorème ergodique. Ainsi, on a la convergence en loi suivante :

$$\sqrt{n}\,\left(\hat{\theta}_n - \theta\right) \underset{n\to+\infty}{\Longrightarrow} \mathcal{N}(0, 1-\theta^2).$$

Enfin, si l'on désire une variance ne dépendant pas du paramètre inconnu θ, il suffit d'utiliser la consistance de l'EMV puis le lemme de Slutsky pour obtenir la convergence en loi suivante :

$$\sqrt{n} \times \frac{\hat{\theta}_n - \theta}{\sqrt{1-\hat{\theta}_n^2}} \underset{n\to+\infty}{\Longrightarrow} \mathcal{N}(0, 1).$$

Notons que comme $\theta \in]-1, 1[$ et que l'on a la convergence p.s. de $\hat{\theta}_n$ vers θ, on aura que p.s. $\hat{\theta}_n \in]-1, 1[$ à partir d'un certain rang (aléatoire), ce qui évite le problème de définition au dénominateur ci-dessus.

Bibliographie

[1] P. Baldi, L. Mazliak et P. Priouret. Martingales et chaînes de Markov. Théorie élémentaire et exercices corrigés. Hermann, 2001.

[2] M. Cottrel, C. Duhamel et V. Genon-Catalot. Exercices de Probabilités. Cassini, 2011.