MAITRISE de MATHEMATIQUES **Université Paul Sabatier**

PROCESSUS STOCHASTIQUES **U.F.R. M.I.G.**

36 h cours, 42 h TD.

1 Introduction

Un processus stochastique est une famille de variables aléatoires (X_t) indexées par le temps.Ce temps peut être discret ($t \in \mathbb{N}$) ce sera le cas dans l'étude des martingales et des chaînes de Markov, ou continu ($t \in \mathbb{R}^+$), ce sera le cas des processus de Poisson.

Ce cours comprendra quatre chapitres :

1. les martingales à temps discret ;
2. les chaînes de Markov à espace d'états fini;
3. les chaînes de Markov à espace d'états dénombrable;
4. les processus de Poisson et les processus de Markov à temps continu et espace d'état fini ou dénombrable.

Il s'agit dans tous les cas de modèles mathématiques rendant compte de phénomènes concrets, physiques ou économiques par exemple. Ces sujets figurent également au programme de l'agrégation, dans l'option orale de mathématiques appliquées.

La théorie des martingales est l'un des outils les plus puissants de la théorie moderne des probabilités. Les chaînes de Markov, quant à elles, sont les modèles les plus utilisés pour représenter les phénomènes aléatoires. Le processus de Poisson est le plus simple des processus de Markov à temps continu, et sert de "brique élémentaire" pour en construire de plus généraux sur les espaces finis ou dénombrables. Il y a bien sûr d'autres modèles, en particulier les processus de Markov à temps continu et à trajectoires continues (les diffusions), dont le modèle le plus simple est le mouvement brownien, mais nous n'en parlerons pas dans ce cours.

1.1 Plan

1. Quelques rappels : théorème des classes monotones, espérances conditionnelles, intégrabilité uniforme

2. Martingales à temps discret.

   Définitions, exemples, inégalités maximales, théorèmes de convergence, temps d'arrêt, théorème d'arrêt, problème de l'arrêt optimal, théorèmes de convergence, exemple de l'algorithme de Robbins Monro, loi des grands nombres et théorèmes de limite centrale.

3. Chaines de Markov à espace d'états fini:

   Définitions, exemples, classification des états, mesure invariante, convergence vers la mesure invariante, théorème ergodique, lois des temps d'atteinte; liens avec les martingales et certains problèmes de résolutions d'EDP, algorithmes MCMC.

4. Chaînes de Markov à espace d'états dénombrables : exemples, récurrence et transience, exemples avec les files d'attente, les marches aléatoires, les processus de vie et de mort.

5. Processus de Poisson :

   Définition par processus de comptage/temps de saut, propriété de Markov, propriété d'accroissements indépendants stationnaires. Construction de Processus de Markov à états finis ou dénombrable à partir du générateur; description du comportement des trajectoires.

2 Rappels

Dans ce chapitre, nous énonçons en vrac quelques définitions et propriétés dont nous nous servirons dans la suite. Tous ces points sont donnés sans démonstrations. La plupart d'entre elles se trouvent dans le livre [1].

Dans tout ce cours, nous utiliserons la notation $\sigma(E)$ pour désigner la plus petite tribu qui rend mesurable tous les éléments de E, que E soit formée de parties d'un ensemble, de fonctions définies sur cet ensemble, ou bien des deux. Ainsi, la tribu borélienne de $\mathbb{R}$ est elle $\sigma([a,b];\ a<b;\ a,b\in\mathbb{Q})$.

2.1 Théorème des classes monotones.

Les théorèmes de classes monotones sont l'un des outils essentiels de la théorie de la mesure. Ils permettent d'étendre facilementà toute une tribu des propriétés établies pour des classes de parties stables par intersection ou des familles de fonctions stables par multiplication. Traditionnellement, ils existent sous forme ensembliste et sous forme fonctionnelle. Nous ne livrons ici que la forme fonctionnelle, qui est de loin la plus souple à l'usage.

Soit Ω un ensemble, et $\mathcal{H}$ un sous-espace vectoriel de l'ensemble des fonctions **bornées** définies sur Ω et à valeurs réelles. On dit que $\mathcal{H}$ a la propriété de classe monotone (**PCM**) si, pour toute suite (f_n) d'éléments de $\mathcal{H}$, croissante, positive et bornée, la limite simple f de la suite (f_n) est encore dans $\mathcal{H}$. En d'autres termes,

$$(f_n\in\mathcal{H}\ ;\ \forall n,\ 0\leq f_n\leq f_{n+1}\leq 1)\Longrightarrow \lim_n f_n\in\mathcal{H}.$$

On notera $\mathcal{B}(\mathcal{A})$ l'espace vectoriel de toutes les fonctions mesurables bornées par rapport à la tribu $\mathcal{A}$, alors le théorème des classes monotones s'énonce

Théorème 2.1. *(TCM) Soit Ω un ensemble et $\mathcal{H}$ un espace vectoriel satisfaisant* **PCM**. *Si $\mathcal{C}$ est une partie de $\mathcal{H}$ stable par multiplication, alors*

$$\mathcal{B}(\sigma(\mathcal{C}))\subset\mathcal{H}.$$

2.2 Espérance conditionnelle.

Nous donnons dans cette section quelques définitions et propriétés qui nous seront utiles dans la suite de ce cours.

Définition 2.2. *Soit $(\Omega, \mathcal{A}, \mathbf{P})$ un espace de probabilité et $\mathcal{B}$ une sous-tribu de $\mathcal{A}$. Soit X une variable aléatoire réelle positive (resp. intégrable), $\mathcal{A}$ mesurable. Alors, il existe une variable aléatoire Y positive (resp. intégrable), $\mathcal{B}$-mesurable, unique (à l'égalité presque sûre près), telle que, pour tout ensemble B $\mathcal{B}$-mesurable, on ait*

$$\mathbf{E}(X\mathbf{1}_B) = \mathbf{E}(Y\mathbf{1}_B).$$

Cette variable aléatoire Y est appelée l'espérance conditionnelle de X sachant $\mathcal{B}$ et on note

$$Y = \mathbf{E}(X/\mathcal{B}).$$

Lorsque $\mathcal{B}$ est la triby $\sigma(Y)$, où Y est une variable aléatoire, alors on note plus simplement $\mathbf{E}(X/\mathcal{B}) = \mathbf{E}(X/Y)$.

L'espérance conditionnelle se comporte comme l'espérance, à ceci près qu'elle associe à une variable aléatoire non pas un nombre mais une nouvelle variable aléatoire.

En particulier, elle jouit des propriétés suivantes :

1. L'application $X \longrightarrow \mathbf{E}(X/\mathcal{B})$ est linéaire.
2. Si $X \leq Y$, $\mathbf{E}(X/\mathcal{B}) \leq \mathbf{E}(Y/\mathcal{B})$, (ps).
3. Si X est positive (resp intégrable), alors, pour toute variable aléatoire Z, $\mathcal{B}$-mesurable, positive (resp bornée), on a

$$\mathbf{E}[Z\mathbf{E}(X/\mathcal{B})] = \mathbf{E}[ZX].$$

   En particulier, et c'est sans doute la propriété la plus importante

$$\mathbf{E}[\mathbf{E}(X/\mathcal{B})] = \mathbf{E}[X].$$

4. (Conditionnement successifs) Si $\mathcal{C} \subset \mathcal{B}$ est une sous-tribu de $\mathcal{B}$, alors

$$\mathbf{E}(X/\mathcal{B}/\mathcal{C}) = \mathbf{E}(X/\mathcal{C}).$$

   Remarquons que c'est faux si $\mathcal{C}$ n'est pas incluse dans $\mathcal{B}$, et qu'en général pour deux sous-tribus quelconques, les espérances conditionnelles sachant $\mathcal{B}$ et sachant $\mathcal{C}$ ne commutent pas (Le fait de commuter est équivalent à la propriété d'indépendance conditionnelle de $\mathcal{B}$ et $\mathcal{C}$ par rapport à leur intersection.)

5. Si Z est $\mathcal{B}$-mesurable, alors $E(Z/\mathcal{B}) = Z$, et plus généralement

$$\mathbf{E}(XZ/\mathcal{B}) = Z\mathbf{E}(X/\mathcal{B}).$$

6. Si X est indépendante de $\mathcal{B}$, alors $E(X/\mathcal{B})$ est constante et vaut $\mathbf{E}(X)$.

7. Si (X, Y) est indépendante de la tribu $\mathcal{B}$, alors $\mathbf{E}(X/\sigma(Y, B) = \mathbf{E}(X/Y)$.

8. Si X est indépendante de $\mathcal{B}$ et si Y est $\mathcal{B}$-mesurable, alors $\mathbf{E}(f(X, Y)/\mathcal{B}) = k(Y)$, où $k(y) = \mathbf{E}(f(X, y))$.

9. Pour tout $p \in [1, \infty]$,
$$\|\mathbf{E}(X/\mathcal{B})\|_p \leq \|X\|_p.$$
L'espérance conditionnelle est donc une contraction dans tous les espaces L^p.

10. L'espérance conditionnelle est dans L^2 un projecteur sur le sous-espace fermé $L^2(\Omega, \mathcal{B}, \mathbf{P})$. En particulier, c'est un opérateur symétrique : pour tout couple (X_1, X_2) de variables de L^2, on a

$$\mathbf{E}[X_1 E(X_2/\mathcal{B})] = \mathbf{E}[E(X_1/\mathcal{B})X_2] = \mathbf{E}[E(X_1/\mathcal{B})E(X_2/\mathcal{B})].$$

11. (Inégalité de Jensen conditionnelle). Si ϕ est convexe, alors

$$\mathbf{E}(\phi(X)/\mathcal{B}) \geq \phi(\mathbf{E}(X/\mathcal{B})), (ps)$$

dès que les deux membres ont un sens. En particulier

$$|\mathbf{E}(X/\mathcal{B})| \leq \mathbf{E}(|X|/\mathcal{B}).(ps)$$

12. Si X est de carré intégrable, sa variance conditionnelle est la variable aléatoire (presque sûrement positive)

$$\sigma^2(X/\mathcal{B}) = \mathbf{E}(X^2/\mathcal{B}) - \mathbf{E}(X/\mathcal{B})^2 = \mathbf{E}[(X - \mathbf{E}(X/\mathcal{B})^2/\mathcal{B}].$$

13. (Inégalité de Hölder conditionnelle). Si $p \in [1, \infty]$ et si $\frac{1}{p} + \frac{1}{q} = 1$, alors

$$|\mathbf{E}[XY/\mathcal{B}]| \leq (ps)[\mathbf{E}(|X|^p/\mathcal{B})]^{1/p}[\mathbf{E}(|Y|^q/\mathcal{B})]^{1/q}.(ps)$$

14. (Théorème de convergence monotone conditionnelle). Si (X_n) est une suite de variables aléatoires positives qui converge en croissant vers X, alors la suite $(\mathbf{E}(X_n/\mathcal{B}))$ converge (en croissant), presque sûrement, vers $\mathbf{E}(X/\mathcal{B})$.

15. (Lemme de Fatou conditionnel)

    Si (X_n) est une suite de variables aléatoires positives, alors

$$\mathbf{E}(\liminf_n X_n/\mathcal{B}) \leq \liminf_n[\mathbf{E}(X_n/\mathcal{B})](ps).$$

16. (Convergence dominée conditionnelle) Si X_n est une suite de variables qui converge presque sûrement vers X, et s'il existe une variable Y positive (non nécessairement intégrable) telle que, pour tout n, $|X_n| \leq Y(ps)$, alors $\mathbf{E}(X_n/\mathcal{B})$ converge presque sûrement vers $\mathbf{E}(X/\mathcal{B})$ sur l'ensemble $\{\mathbf{E}(Y/\mathcal{B}) < \infty\}$.

Il faut faire attention à ce que, dans le cas positif, l'espérance conditionnelle d'une variable finie presque sûrement peut fort bien être infinie partout. Il suffit pour s'en convaincre de considérer, sur $[0,1]$, muni de la mesure de Lebesgue, la sous-tribu engendrée par les deux intervalles $[0,1/2[$ et $[1/2,1]$, ainsi que la variable $X(\omega) = \frac{1}{\omega} + \frac{1}{1-\omega}$.

2.3 Intégrabilité uniforme.

On dit qu'une famille de v.a. $(X_\alpha, \alpha \in I)$ est **uniformément intégrable** si

$$\lim_{n\to\infty} \sup_\alpha \int_{\{|X_\alpha|\geq n\}} |X_\alpha| d\mathbf{P} = 0.$$

Il est équivalent d'avoir :

$$\sup_\alpha E[|X_\alpha|] < +\infty \; ; \; \forall\varepsilon > 0, \exists\delta > 0, \forall A \in \mathcal{A} : \mathbf{P}(A) \leq \delta, \int_A |X_\alpha| d\mathbf{P} < \varepsilon.$$

Une partie finie de L^1 est uniformément intégrable, ainsi qu'une famille dominée dans L^1. Plus généralement, si $\mathcal{H}$ est uniformément intégrable, et si une famille $\mathcal{H}_1$ est telle que

$$\forall X \in \mathcal{H}_1, \; \exists Y \in \mathcal{H}, \; |X| \leq |Y|,$$

alors $\mathcal{H}_1$ est uniformément intégrable.

Le critère suivant est en pratique très utile pour s'assurer de l'intégrabilité uniforme d'une famille.

Proposition 2.3. *(Théorème de De La Vallée Poussin) Une famille $X_\alpha, \alpha \in I$ est uniformément intégrable si et seulement si il existe une fonction $\Phi(x) : \mathbb{R}_+ \to \mathbb{R}_+$, croissante, telle que $\Phi(x)/x$ converge vers $+\infty$ lorsque $x \to +\infty$, et telle que*

$$\sup_\alpha \mathbf{E}(\Phi(|X_\alpha|)) < \infty.$$

Ainsi, une partie bornée dans L^p est uniformément intégrable dès que $p > 1$.

Une autre classe de parties uniformément intégrables est donnée de la façon suivante.

Proposition 2.4. *Si X est dans L^1, l'ensemble des variables $\mathbf{E}(X/\mathcal{B})$, où $\mathcal{B}$ parcourt les sous-tribus de $\mathcal{A}$, est uniformément intégrable.*

La notion d'intégrabilité uniforme est principalement utilisée à cause de la propriété suivante:

Proposition 2.5. *Si (X_n) est une suite de variables aléatoires intégrables, qui converge vers X en probabilité lorsque n converge vers ∞, alors il est équivalent de dire*

1. *(X_n) est uniformément intégrable.*
2. *X_n converge vers X dans L^1.*

Dans ce cas, on a bien sûr $\mathbf{E}(X_n) \to \mathbf{E}(X)$.

Remarquons qu'on retrouve ainsi comme cas particulier le théorème de convergence dominée.

3 Filtrations, temps d'arrêt.

Dans ce chapitre, on introduit le vocabulaire élémentaire de la théorie des processus à temps discrets que nous utiliserons dans la suite du cours.

Un **processus** (à temps discret) sur $(\Omega, \mathcal{A}, \mathbf{P})$ est une suite de variables aléatoires $(X_n(\omega))_{n\geq 0}$, à valeurs dans un espace mesuré $(E, \mathcal{E})$. Lorsque E est l'ensemble des réels, on parlera de processus réel, ou scalaire. Il est souvent utile de le considérer comme une variable aléatoire $X(\omega, n)$ défini sur l'espace produit $\Omega \times \mathbb{N}$, muni de la tribu produit.

Une **filtration** $(\mathcal{F}_n)$ est la donnée d'une suite croissante de sous-tribus de $\mathcal{A}$. Sauf si c'est spécifié par ailleurs, on pose en général $\mathcal{F}_\infty = \sup_n \mathcal{F}_n$, c'est dire la plus petite tribu qui contienne toutes les $\mathcal{F}_n$.

Un exemple typique de filtration est obtenu de la façon suivante : on observe une suite (X_n) de variables aléatoires, X_n étant l'observation à l'instant n, et $\mathcal{F}_n = \sigma(X_0, \cdots, X_n)$ est la plus petite tribu qui rende mesurable les variables $(X_0, \cdots, X_n)$. On l'appelle la **filtration naturelle** du processus X_n. On la notera $\mathcal{F}^X$, ou $(\mathcal{F}_n^X)$.

Il est souvent commode (mais pas indispensable) de supposer que $\mathcal{F}_0$ contient tous les négligeables de $\mathcal{F}_\infty$, c'est à dire tous les ensembles A qui sont inclus dans un ensemble de $\mathcal{F}_\infty$ qui est de probabilité nulle. On peut toujours augmenter toutes les tribus de ces négligeables (on dit qu'on les complète). Cela permet de supposer qu'un ensemble de mesure nulle est toujours dans $\mathcal{F}_0$, et d'avoir la même notion d'égalité ps dans toutes les tribus.

On dit qu'un processus (X_n) est **adapté** à $(\mathcal{F}_n)$ si, pour tout n, X_n est $\mathcal{F}_n$-mesurable.

Ainsi, la filtration naturelle du processus (X_n) est la plus petite filtration par rapport à laquelle le processus (X_n) soit adapté.

On dit qu'un processus (X_n) est prévisible si X_0 est $\mathcal{F}_0$-mesurable, et que, pour tout $n \geq 1$, X_n est $\mathcal{F}_{n-1}$-mesurable. Ainsi, $Y_n = X_{n-1}$ est prévisible dans la filtration naturelle de (X_n).

Un **Temps d'Arrêt** est une application de Ω dans $\bar{\mathbb{N}} = \mathbb{N} \cup \{\infty\}$ qui a la propriété suivante :

$$\forall n \in \mathbb{N}, \ \{T \leq n\} \in \mathcal{F}_n.$$

Remarques

1. Il est important d'autoriser aux *temps d'arrêt* de prendre la valeur $+\infty$. Noter que l'événement $\{T = \infty\}$ est automatiquement $\mathcal{F}_\infty$ mesurable.

2. La variable aléatoire $T : \Omega \longrightarrow \bar{\mathbb{N}}$ est un *temps d'arrêt* si et seulement si, pour tout n, l'événement $\{T = n\}$ est $\mathcal{F}_n$ mesurable. La démonstration est laissée à titre d'exercice.

3. Si T est une variable à valeurs dans $\bar{\mathbb{N}}$, on peut considérer le processus $\mathbf{1}_{[T,\infty]}(\omega, n) = \mathbf{1}_{\{T \leq n\}}(\omega)$, ou le processus $\mathbf{1}_{[0,T]}(\omega, n) = \mathbf{1}_{\{n \leq T\}}(\omega)$. Alors, T est un *temps d'arrêt* si et seulement si $\mathbf{1}_{[T,\infty]}$ est adapté, ou encore si et seulement si $\mathbf{1}_{[0,T]}$ est prévisible.

L'exemple fondamental de *temps d'arrêt* est le suivant : soit (X_n) un processus adapté et A_n une famille d'ensembles mesurables dans l'espace où (X_n) prend ses valeurs. Alors, la variable

$$T(\omega) = \inf\{n |\ X_n \in A_n\}$$

est un *temps d'arrêt*. (Par convention, on pose $\inf\{\emptyset\} = +\infty$.)

Pour le voir, il suffit d'écrire

$$\{T = n\} = \cap_{k=0}^{n-1}\{X_k \notin A_k\} \cap \{X_n \in A_n\},$$

dont on voit immédiatement qu'il est dans $\mathcal{F}_n$.

Réciproquement, tout temps d'arrêt peut s'écrire comme cela, en considérant le processus $\mathbf{1}_{[T,\infty]}$ et $A_n = \{1\}$.

Quelques propriétés élémentaires des *temps d'arrêt* : si S et T sont des *temps d'arrêt*, $\sup(S,T)$, $\inf(S,T)$sont des *temps d'arrêt*, $S+T$ est un *temps d'arrêt*, mais en général $T-1$ n'est pas un *temps d'arrêt*. Les vérifications de ces assertions sont immédiates.

Notation Dans toute la suite, nous désignerons le $\sup$ par le symbole $\vee$ et l'$\inf$ par le symbole $\wedge$. Ainsi

$$\sup(S,T) = S \vee T,\ \ \inf(S,T) = S \wedge T.$$

Tribus associées aux temps d'arrêt : si T est un *temps d'arrêt*, on définit la tribu des événements antérieurs à T, et on note $\mathcal{F}_T$, l'ensemble des événements A de $\mathcal{F}_\infty$ qui ont la propriété suivante : $\forall n \in \mathbb{N},\ A \cap \{T \leq n\} \in \mathcal{F}_n.$

Pour vérifier qu'une variable est $\mathcal{F}_T$-mesurable, on a la proposition suivante:

Proposition 3.1. *Une variable réelle X est $\mathcal{F}_T$ mesurable si et seulement si, pour tout n,*

$X\mathbf{1}_{T \leq n}$ est $\mathcal{F}_n$ mesurable.

C'est immédiat, si on se rappelle qu'une variable $\mathcal{F}_T$-mesurable est limite simple de combinaisons linéaires d'indicatrices d'ensembles $\mathcal{F}_n$-mesurables.

Remarque. — Il ne faut pas confondre la tribu $\mathcal{F}_T$ et la tribu $\sigma(T)$, la plus petite tribu rendant T mesurable : pour tout k, on a $\{T = k\} \in \mathcal{F}_T$, et donc $\sigma(T) \subset \mathcal{F}_T$, mais l'inverse est faux en général.

On vérifie immédiatement les propriétés suivantes :

1. Si S et T sont des *temps d'arrêt*, et si $A \in \mathcal{F}_S$, alors $A \cap \{S \leq T\} \in \mathcal{F}_T$.

2. Les événements $\{S \leq T\}$, $\{S < T\}$, $\{S = T\}$ appartiennent à $\mathcal{F}_S \cap \mathcal{F}_T$.

3. Si $S \leq T$, alors $\mathcal{F}_S \subset \mathcal{F}_T$.

4. Si S et T sont des *temps d'arrêt*, alors $\mathcal{F}_S \cap \mathcal{F}_T = \mathcal{F}_{S \wedge T}$.

5. Si T est un *temps d'arrêt* fini partout (c'est à dire qui ne prend pas la valeur $+\infty$), alors F_T est engendrée par les éléments de la forme $A \cap \{n \leq T\}$, avec $A \in \mathcal{F}_n$. Pour le voir, on remarque d'un côté que les générateurs proposés sont bien dans $\mathcal{F}_T$, et d'autre part que, si A est dans $\mathcal{F}_T$, alors

$$A = \bigcup_n A \cap \{T = n\} = \bigcup_n [(A \cap \{T \leq n\}) \cap \{n \leq T\}],$$

   et, puisque $(A \cap \{T \leq n\}) \in \mathcal{F}_n$, on a donc bien écrit A à l'aide des générateurs proposés.

   En fait, cette formule reste vraie même si T n'est pas fini, mais ceci apparaîtra comme une conséquence du théorème de convergence des martingales, et ne sera vrai qu'aux ensembles de mesure nulle près.

6. Si T et S sont des *temps d'arrêt* finis, alors $\mathcal{F}_S \vee \mathcal{F}_T = \mathcal{F}_{S \vee T}$, où $\mathcal{F}_S \vee \mathcal{F}_T$ désigne la plus petite tribu qui contienne à la fois F_T et $\mathcal{F}_S$. On peut le voir en utilisant les générateurs décrits plus hauts.

7. Si (X_n) est adapté et que T est un *temps d'arrêt*, alors la variable X_T définie par

$$X_T(\omega) = X_{T(\omega)}(\omega) \mathbf{1}_{T < \infty} = \sum_n X_n \mathbf{1}_{T=n}$$

   est $\mathcal{F}_T$ mesurable. Cette propriété est immédiate si on utilise la proposition précédente qui caractérise les variables $\mathcal{F}_T$-mesurables. Remarquons que nous avons pris la précaution de poser $X_T = 0$ sur $\{T = \infty\}$. Si une variable X_∞ $\mathcal{F}_\infty$-mesurable est définie, alors on a le même résultat en posant $X_T = X_\infty$ sur $\{T = \infty\}$.

8. Si T est un *temps d'arrêt*, et si A est un ensemble mesurable, la variable $T^A = T\mathbf{1}_A + \infty \mathbf{1}_{A^c}$ est un temps d'arrêt si et seulement si $A \in \mathcal{F}_T$, puisque $\{T^A \leq n\} = A \cap \{T \leq n\}$.

4 Martingales à temps discret.

Voir le livre de J. NEVEU [8], le livre de P. BARBE et M. LEDOUX. [1], ou celui de P.S. TOULOUSE [11].

La théorie des martingales a son origine dans l'étude des jeux : elle modélise d'une part le caractère aléatoire d'un phénomène mais aussi son évolution dans le temps. On étudie ici le temps discret.

4.1 Définitions et premiers exemples

Définition 4.1. *Soit un processus adapté X sur l'espace de probabilité filtré $(\Omega, \mathcal{A}, (\mathcal{F}_n, n \in \mathbb{N}), \mathbf{P})$ tel que pour tout entier n, X_n est intégrable. On dit que X est une* **martingale** *si pour tout entier n,*

$$E[X_{n+1}/\mathcal{F}_n] = X_n, \ \text{presque sûrement.}$$

On dit que X est une **sous-martingale** *si pour tout entier n,*

$$E[X_{n+1}/\mathcal{F}_n] \geq X_n, \ \text{presque sûrement.}$$

On dit que X est une **sur-martingale** *si pour tout entier n,*

$$E[X_{n+1}/\mathcal{F}_n] \leq X_n, \ \text{presque sûrement.}$$

4.1.1 Exemples

(Les démonstrations des propriétés qui suivent sont laissés au lecteur à titre d'exercice.)

(a) Si $H \in L^1(\Omega, \mathcal{A})$, $X_n = E[H/\mathcal{F}_n]$ est une martingale. D'après le critère 2.4, cette martingale est uniformément intégrable.

(b) Soit $(Z_n, n \in \mathbb{N})$ une suite de variables aléatoires indépendantes et intégrables et $X_n = \sum_{i=1}^n Z_i$. Alors les filtrations $\mathcal{F}^X$ et $\mathcal{F}^Z$ sont égales et, pour cette filtration

1. Si pour tout entier n, $E(Z_n) = 0$, $X = (X_n)$ est une martingale;
2. Si pour tout entier n, $E(Z_n) \geq 0$, $X = (X_n)$ est une sous- martingale;
3. Si pour tout entier n, $E(Z_n) \leq 0$, $X = (X_n)$ est une sur-martingale;

4. Dans tous les cas, si toutes les variables Z_i ont même espérance m, $X_n - nm$ est une martingale.

(c) L'exemple le plus simple de théorie des jeux est celui où les variables aléatoires Z_n sont des variables aléatoires de loi de Bernoulli de paramètre p avec valeurs $+1$ et -1. Alors, X_n représente la fortune d'un joueur après n paris. Alors $M_n = X_n - n(2p-1)$ est une martingale pour la filtration naturelle $\mathcal{F}^X$.

(d) Soit $(Z_n, n \in \mathbb{N})$ une suite de variables aléatoires indépendantes et intégrables, et on pose $X_n = \sum_{i=1}^n Z_i$. On suppose que $\mathbf{E}[\exp(aZ_n)] := \exp(r_n)$ existe et est finie. On pose $R_n = r_1 + \cdots + r_n$. (Ici, par convention, $X_0 = R_0 = 0$.)

Alors, $M_n = \exp(aX_n - R_n)$ est une martingale pour la filtration naturelle $\mathcal{F}^X$.

Remarquons enfin qu'on peut avoir le choix de la filtration par rapport à laquelle X est une martingale, (resp une sous-m., une sur-m.). En effet

Proposition 4.2. *Si X est une martingale (resp. une sur-m., une sous-m.) par rapport à une filtration $(\mathcal{F}_n)$ et qu'elle est adaptée à une filtration $(\mathcal{G}_n)$ plus petite que $(\mathcal{F}_n)$ (au sens où, pour tout n, $\mathcal{G}_n \subset \mathcal{F}_n$), alors X est une martingale (resp. une sur.m., une sous-m.) par rapport à la filtration $\mathcal{G}_n$. En particulier, X est une martingale (resp. une sur-m., une sous-m.) par rapport à sa filtration naturelle.*

Démonstration. — Il suffit d'appliquer le principe des conditionnements successifs. ∎

Proposition 4.3. *Soit X une $\mathcal{F}$-martingale. Alors*

1. *$\forall n \geq 0, \forall k \geq 0$, $E[X_{n+k}/\mathcal{F}_n] = X_n$; $E[X_n] = E[X_0]$.*
2. *Si la martingale est de carré intégrable, les accroissements $(\Delta X)_n := X_n - X_{n-1}$ de X sont orthogonaux :*

$$n \neq m \implies \mathbf{E}[(\Delta X)_n (\Delta X)_m] = 0.$$

3. *Si X est une sur-martingale, $-X$ est une sous-martingale.*
4. *L'espace des martingales est un espace vectoriel réel.*
5. *Si X est une martingale et ϕ une application convexe telle que $\phi(X_n)$ est intégrable, alors le processus Y défini par $Y_n = \phi(X_n)$ pour tout n est une sous-martingale si pour tout $n, Y_n \in L^1$.*

6. *De même, si X est une sous-martingale, et ϕ croissante et convexe, $\phi(X)$ est une sous-martingale.*

Démonstration. — (Exercices) ∎

Pour les martingales de carré intégrable, nous avons en outre

Proposition 4.4. *Si M_n est une martingale de carré intégrable et si $(\Delta M)_n$ désigne l'accroissement $(M_n - M_{n-1})$, alors*

$$\forall n \leq p,\ \mathbf{E}[(M_p - M_n)^2] = \sum_{n+1}^{p} \mathbf{E}[(\Delta M)_k^2].$$

Démonstration. — Il suffit d'appliquer le propriété d'accroissements orthogonaux de la proposition précédente. ∎

Corollaire 4.5. *Une martingale bornée dans L^2 converge dans L^2.*

Démonstration. — Si la martingale est bornée dans L^2, alors la série

$$\sum_{0}^{p} \mathbf{E}[(\Delta M)_k^2]$$

est bornée, et donc converge. Par conséquent,

$$\lim_{n\to\infty} \sup_{p\geq n} \sum_{n}^{p} \mathbf{E}[(\Delta M)_k^2] = 0.$$

D'après la proposition précédente, la suite est donc de Cauchy dans L^2, et y converge.

∎

Nous verrons un peu plus bas un théorème beaucoup plus fort.

4.2 Décompositions

La première proposition décompose une sous-martingale en la somme d'une martingale et d'une suite croissante de variables aléatoires :

Théorème 4.6. **Décomposition de Doob** *: soit X une sous-martingale ; il existe une martingale M et un processus croissant prévisible A, nul en 0, uniques, tels que pour tout entier n, $X_n = M_n + A_n$.*

Le processus A est appelé “compensateur” de X.
Démonstration. — On part de $A_0 = 0$ et $M_0 = X_0$. Puis pour $n \geq 1$, on définit A_n de la façon suivante : on pose $\Delta_n = \mathbf{E}(X_n/\mathcal{F}_{n-1}) - X_{n-1}$, et

$$A_n = \Delta_1 + \cdots + \Delta_n.$$

On pose enfin $M_n = X_n - A_n$. Par construction A_n est prévisible, et puisque X_n est une sous-martingale, $\Delta_n \geq 0$, et donc A_n est croissante. Ensuite,

$$\mathbf{E}(M_{n+1}/\mathcal{F}_n) = \mathbf{E}(X_{n+1}/\mathcal{F}_n) - A_{n+1} = X_n + \Delta_n - A_{n+1} = M_n.$$

et donc M_n est une martingale.

L’unicité de la décompositon s’écrit de la même façon, et remarquant que, si une telle décomposition a lieu, on doit avoir

$$\mathbf{E}(X_{n+1} - X_n/\mathcal{F}_n) = A_{n+1} - A_n,$$

ce qui caractérise A_n si l’on sait que $A_0 = 0$. ∎

Explicitons cette décomposition dans le cas particulier des martingales de carré intégrable.

Proposition 4.7. *Soit M_n une martingale de carré intégrable. Posons*

$$(\Delta M)_n = M_n - M_{n-1} \textit{ et } U_n = \mathbf{E}[(\Delta M)_n^2/\mathcal{F}_{n-1}).$$

Alors $M_n^2 - \sum_{k=1}^n U_k$ est une martingale.

Démonstration. — C’est la décomposition précédente dans le cas de la sous-martingale M_n^2, en remarquant que

$$\mathbf{E}[(\Delta M)_n^2/\mathcal{F}_{n-1}) = \mathbf{E}(M_n^2/\mathcal{F}_{n-1}] - M_{n-1}^2.$$

Il suffit de d’evelopper le double produit et d’utiliser la propriété de martingale.

∎

La seconde décomposition est beaucoup moins utilisée que la première, mais rend quelques services:

Théorème 4.8. Décomposition de Krickeberg *Soit X une martingale bornée dans L^1 ; il existe deux martingales positives M et N telles que $X_n = M_n - N_n$ pour tout entier n.*

Démonstration. — (On en verra une démonstration élémentaire plus bas dans le cas des martingales uniformément intégrables.)

On peut tout d'abord supposer que $X_0 = 0$.

L'hypothèse est qu'il existe une constante M telle que pour tout n, $E[|X_n|] \leq M$. Or, la suite $(|X_n|, n \in \mathbb{N})$ est une sous martingale. Par la décomposition de Doob : il existe une martingale M et un processus croissant prévisible A tels que pour tout entier $n, |X_n| = M_n + A_n$. La suite croissante A_n est bornée dans L^1, et donc converge vers une limite presque sûre notée A_∞ intégrable. On a $|X_n| \leq M_n + A_\infty$, et on peut donc en prendre l'espérance conditionnelle :

$$\forall n, |X_n| \leq M_n + E[A_\infty / \mathcal{F}_n] = Y_n.$$

La suite Y est une martingale positive, $Y - X$ aussi puisque $Y \geq |X| \geq X$. ∎

4.3 Transformation des martingales

La propriété de martingale (et également de sur- et sous-martingale) a l'énorme avantage d'être stable par un certain nombre de transformations élémentaires, ne seraient pas possibles dans le cadre de l'étude des sommes de variables aléatoires indépendantes.

La principale transformation est celle-ci :

Proposition 4.9. *Soit (X_n) un processus adapté et (V_n) un processus prévisible tel que, pour tout n, la variable $V_n(X_n - X_{n-1})$ soit intégrable. Alors, on définit le processus $(V.X)$ de la façon suivante*

$$(V.X)_n = V_0 X_0 + \sum_{k=1}^{n} V_n (X_n - X_{n-1}).$$

Alors, si X est une martingale, il en va de même de $(V.X)$. Si X est une sur- (resp sous-) martingale, et si V est de plus positif, alors $(V.X)$ est une sur- (resp. sous-) martingale.

Démonstration. — En reprenant les notations déjà utlisées, le processus $(V.X)$ vérifie

$$(\Delta(V.X))_n = V_n(\Delta X)_n.$$

La démonstration est alors immédiate et laissée au lecteur. ∎

Dans un casino par exemple, un processus V correspond à une stratégie du joueur : en fonction de toutes les observations dont il dispose à l'instant n, il va miser à l'instant $n+1$ une somme V_{n+1}, pour obtenir un gain $V_{n+1}(X_{n+1} - X_n)$.

Une conséquence importante est la suivante :

Proposition 4.10. *Soit (X_n) une martingale (resp. une sous-, une sur-martingale), et soit T un temps d'arrêt. Alors le processus X^T défini par $X^T_n = X_{T\wedge n}$ est une martingale (resp. une sous-, une sur-martingale).*

Démonstration. — Il suffit de considérer le processus $V = \mathbf{1}_{[0,T]}$, dont on a déjà vu qu'il est prévisible. Dans ce cas, le processus $(V.X)$ n'est rien d'autre que X^T :

$$(V.X)_n = X_0 + \sum_{n\leq T}(X_{n+1} - X_n).$$

∎

4.4 Théorème d'arrêt

Dans cette section, nous énonçons le théorème d'arrêt pour les *temps d'arrêt* bornés : il sera étendu plus tard aux *temps d'arrêt* quelconques, mais uniquement pour les martingales uniformément intégrables.

Théorème 4.11. *(***Théorème d'arrêt***.)*

Soit $(X_n, n\in\mathbb{N})$ une martingale et S et T deux temps d'arrêt bornés (c'est à dire qu'il existe un entier N tel que $S\vee T\leq N$, presque sûrement). Alors

$$\mathbf{E}(X_T/F_S) = X_{S\wedge T}.$$

De même, si X est une sous- (resp une sur-)martingale, on a

$$X_{S\wedge T} \leq \text{ (resp } \geq\text{) } \mathbf{E}(X_T/\mathcal{F}_S).$$

Démonstration. — Nous n'allons démontrer le résultat que pour les martingales, le cas des sur et sous- martingales se traitant de la même façon.

Nous allons d'abord montrer le théorème lorsque $T = N$ et $S \le N$. Nous avons

$$X_S = \sum_{k=0}^{N} X_k \mathbf{1}_{S=k}.$$

Nous savons déjà que X_S est $\mathcal{F}_S$ mesurable; elle est de plus intégrable comme combinaison linéaire de finie de variables intégrables.

Il nous suffit donc de démontrer que, pour tout $A \in \mathcal{F}_S$,

$$\mathbf{E}(X_S \mathbf{1}_A) = \mathbf{E}(X_N \mathbf{1}_A).$$

Ceci s'écrit

$$\sum_{k=0}^{N} \mathbf{E}(X_k \mathbf{1}_{A\cap\{S=k\}}) = \sum_{k=0}^{N} \mathbf{E}(X_N \mathbf{1}_{A\cap\{S=k\}}).$$

Mais, puisque $A \in \mathcal{F}_S$, alors $A \cap \{S = k\} \in \mathcal{F}_k$, et la propriété de martingale s'applique pour donner, pour tout $k \le N$,

$$\mathbf{E}(X_k \mathbf{1}_{A\cap\{S=k\}}) = \mathbf{E}(X_N \mathbf{1}_{A\cap\{S=k\}}).$$

Pour passer au cas général, il suffit d'appliquer ce résultat à la martingale arrêtée X^T, qui vaut X_T en N puisque $T \le N$. Alors $X_S^T = X_{S\wedge T}$.

∎

Corollaire 4.12. *Soit (T_n) une suite croissante de temps d'arrêt bornés, et X une (sous, sur) martingale ; alors $(X_{T_n}, n \in \mathbb{N})$ est une (sous, sur) martingale pour la filtration $(\mathcal{F}_{T_n}, n \in \mathbb{N})$.*

Démonstration. — (en exercice) ∎

Corollaire 4.13. *Soit X est une variable intégrable, et soit (X_n) la martingale $\mathbf{E}(X/\mathcal{F}_n)$. Si T est un temps d'arrêt borné, alors*

$$\mathbf{E}(X/\mathcal{F}_T) = X_T.$$

Si S et T sont deux temps d'arrêt bornés,

$$\mathbf{E}(X/\mathcal{F}_S/\mathcal{F}_T) = \mathbf{E}(X/\mathcal{F}_T/\mathcal{F}_S) = \mathbf{E}(X/\mathcal{F}_{S\wedge T}) = X_{S\wedge T}.$$

En d'autres termes, $\mathcal{F}_S$ et $\mathcal{F}_T$ sont conditionnellement indépendantes sachant $\mathcal{F}_{S\wedge T} = \mathcal{F}_S \cap \mathcal{F}_T$.

Démonstration. — Soit N un majorant commun à T et S. On considère la martingale $X_n = \mathbf{E}(X/\mathcal{F}_n)$. Par le principe des conditionnements successifs, $\mathbf{E}(X/\mathcal{F}_T) = \mathbf{E}(X_N/\mathcal{F}_T)$. Puis le théorème d'arrêt nous donne $\mathbf{E}(X_N/\mathcal{F}_T) = X_T$.

Si S et T sont deux temps d'arrêt bornés,

$$\mathbf{E}(X_T/\mathcal{F}_S) = X_{S\wedge T} = \mathbf{E}(X/\mathcal{F}_{S\wedge T}).$$

Ensuite, on retrouve la caractérisation de l'indépendance conditionnelle : deux tribus sont conditionnellement indépendantes par rapport à leur intersection si et seulement si les espérances conditionnelles par rapport à ces tribus commutent.

∎

Corollaire 4.14. *Soit $(X_n, n \in \mathbb{N})$ une $\mathcal{F}$-(sous)martingale et T un $\mathcal{F}$-temps d'arrêt presque sûrement fini. Alors, si l'une des deux hypothèses suivantes est réalisée :*

1. *T est un temps d'arrêt borné.*
2. *La suite $(X_{T\wedge n})$ est uniformément intégrable. (En particulier s'il existe $Y \in L^1$ telle que pour tout n, $|X_{T\wedge n}| \leq Y$.)*

alors $E[X_0](\leq) = E[X_T]$.

Démonstration. — Exercice : dans le premier cas, puisque T et 0 sont deux temps d'arrêt bornés et $T \geq 0$, le théorème d'arrêt montre que $E(X_T/\mathcal{F}_0)(\geq) = X_0$, et on prend l'espérance.
Dans le deuxième cas, on prend d'abord comme temps d'arrêt $T \wedge n$ qui est borné, et vérifie donc $E[X_{T\wedge n}](\geq) = E[X_0]$. L'hypothèse (ii) permet alors d'utiliser le théorème de Lebesgue de convergence dominée, puisque le fait que T est presque sûrement fini implique que $X_{T\wedge n}$ converge presque sûrement vers X_T. ∎

Applications. —

1. **Exemple : la ruine du joueur.** (Exercice) Dans l'exemple ci-dessus, on considère que Y_n est le résultat d'un jeu au temps n, variable aléatoire de Bernoulli valant $+1$ ou -1 avec probabilité $p, 1-p$. La fortune initiale du joueur est a et il espère atteindre la quantité $b > a$. Le jeu est terminé lorsque la fortune du joueur est soit nulle, auquel cas il a perdu, soit égale à b, auquel cas il a gagné. On note T l'instant de la fin du jeu.

   On demande

- la probabilité $\mathbf{P}(G)$ que le joueur gagne,

- l'espérance de T.

Traitons d'abord le cas où $m = 2p - 1$ est non nul.

Tout d'abord, il faut montrer que le temps d'atteinte T est fini presque sûrement. Si S_n est la fortune du joueur à l'instant n, $S_n - mn$ est une martingale. En considérant le temps $T \wedge n$, et la martingale précédente, on obtient, lorsque $m > 0$, que $\mathbf{E}(T) \leq (b - a)/m$. On en déduit que le temps T est fini (ps) lorsque $m > 0$. Par symétrie, on déduit de même que T est fini (ps) lorsque $m < 0$ (Donner une majoration de $\mathbf{E}(T)$ dans ce cas).

Chercher un nombre positif r tel que r^{S_n} soit aussi une martingale; appliquer le théorème d'arrêt à ces deux martingales pour obtenir deux relations qui lient $\mathbf{E}(T)$ et la probabilité de gain.

Lorsque m est nul, S_n est elle même une martingale, et $S_n^2 - n$ est encore une martingale. Utiliser alors un raisonnement analogue.

(On trouve $\mathbf{P}(G) = a/b$ et $\mathbf{E}(T) = a(b - a)$ dans le cas $m = 0$, et sinon, $\mathbf{P}(G) = (r^a - 1)/(r^b - 1)$, $\mathbf{E}(T) = (b\mathbf{P}(G) - a)/m$, avec $r = (1 - p)/p$.)

On peut déduire de ce qui précède d'autres informations. Par exemple, dans le cas où $m = 0$, que le temps d'arrêt $T_b = \inf\{n \mid X_n = b\}$ est fini presque sûrement. En effet, supposons pour simplifier que la martingale X_n parte de 0 à l'instant n, et appelons $T_{b,k} = T_b \wedge T_{-k} = \inf\{n \mid X_n \in \{-k, b\}\}$. En décalant de k le résultat précédent, nous trouvons que $\mathbf{P}(T_b < T_{-k}) = \frac{k}{b+k}$. Puisque T_{-k} croît vers l'infini avec k, nous obtenons en passant à la limite que $\mathbf{P}(T_b < \infty) = 1$. Mais observer qu'on ne peut pas appliquer le théorème d'arrêt au temps T_b, ce qui mènerait à une contradiction. Á titre d'exercice, le lecteur pourra calculer $\mathbf{P}(T_b < \infty)$, pour $b > 0$, dans le cas où $m < 0$.

Un autre exemple d'application important est celui de l'arrêt optimal. Ce problème se rencontre dans la pratique lorsque, connaissant la loi d'une suite de variables aléatoires X_n, on cherche à maximiser un profit, ou à minimiser un coût.

2. **Arrêt optimal.** Dans une filtration$(\mathcal{F}_n)$, on se donne un entier N, une suite de variable aléatoires $(V_n)_{n \leq N}$ $\mathcal{F}_n$ mesurables, intégrables, et on cherche à trouver le *temps d'arrêt* T inférieur à N qui maximise $\mathbf{E}(V_T)$.

   La réponse est donnée de la façon suivante : on définit la suite $(\hat{V}_n)_{n \leq N}$ par

$$\hat{V}_N = V_N;\ \hat{V}_k = \max(V_k, \mathbf{E}(\hat{V}_{k+1}/\mathcal{F}_k))\ (k = 0, \cdots, N - 1).$$

Alors $T = \inf(n |\ V_n = \hat{V}_n\}$.

Démonstration. — Remarquons d'abord que $\hat{V}_n$ est une sur-martingale. En fait, c'est la plus petite sur-martingale qui majore (V_n). (Elle n'est définie que pour $n \leq N$, mais le lecteur formaliste pourra la prolonger à tous les entiers en la supposant arrêtée à N.) Elle s'appelle **l'enveloppe de Snell** de la suite (V_n).

Remarquons que le processus arrêté $\hat{V}^T$ est une martingale: en effet, si k est un entier ($\leq N-1$), on a

$$\hat{V}_{k\wedge T} = \mathbf{E}(\hat{V}_{(k+1)\wedge T}/\mathcal{F}_k).$$

Pour le voir, nous décomposons

$$\mathbf{E}(\hat{V}_{(k+1)\wedge T}/\mathcal{F}_k) = \mathbf{E}(\hat{V}_T \mathbf{1}_{T\leq k}/\mathcal{F}_k) + \mathbf{E}(\hat{V}_{k+1}\mathbf{1}_{T\geq k+1}/\mathcal{F}_k).$$

Les événements $\{T \leq k\}$ et $\{T \geq k+1\}$ sont dans $\mathcal{F}_k$, et la variable $\hat{V}_T\mathbf{1}_{T\leq k}$ est donc $\mathcal{F}_k$-mesurable. Donc

$$\mathbf{E}(\hat{V}_T\mathbf{1}_{T\leq k}/\mathcal{F}_k) = \hat{V}_T\mathbf{1}_{T\leq k}.$$

Pour le deuxième terme, nous observons que sur $\{k+1 \leq T\}$, $\hat{V}_k = \mathbf{E}(\hat{V}_{k+1}/\mathcal{F}_k)$, par définition de $\hat{V}$ et de T, et donc

$$\mathbf{E}(\hat{V}_{k+1}\mathbf{1}_{T\geq k+1}/\mathcal{F}_k) = \hat{V}_k\mathbf{1}_{T\geq k+1}.$$

D'où

$$\mathbf{E}(\hat{V}_{(k+1)\wedge T}/\mathcal{F}_k) = \hat{V}_T\mathbf{1}_{T\leq k} + \hat{V}_k\mathbf{1}_{T\geq k+1} = \hat{V}_k^T.$$

Choisissons un *temps d'arrêt* $S \leq N$, et montrons que $\mathbf{E}(V_S) \leq \mathbf{E}(V_T)$. On écrit

$$\mathbf{E}(V_S) \leq \mathbf{E}(\hat{V}_S) = \mathbf{E}(\hat{V}_S\mathbf{1}_{S\geq T}) + \hat{\mathbf{E}}(V_S\mathbf{1}_{S<T}).$$

Pour le premier terme, on utilise la propriété de sur-martingale :

$$\mathbf{E}(\hat{V}_S\mathbf{1}_{S\geq T}) \leq \mathbf{E}(\hat{V}_T\mathbf{1}_{S\geq T}) = \mathbf{E}(\ V_T\mathbf{1}_{S\geq T}).$$

Le second terme se majore de la façon suivante. On écrit

$$\mathbf{E}(\hat{V}_S\mathbf{1}_{S<T}) = \mathbf{E}(\hat{V}^T_{S\wedge T}\mathbf{1}_{S<T}).$$

Nous savons que

$$\{S < T\} \in \mathcal{F}_S \cap \mathcal{F}_T = \mathcal{F}_{S\wedge T}.$$

Utilisant la propriété de martingale de $\hat{V}^T$, et le théorème d'arrêt, nous avons

$$\mathbf{E}(\hat{V}^T_{S\wedge T}\mathbf{1}_{S<T}) = \mathbf{E}(\hat{V}^T_T\mathbf{1}_{S<T}) = \mathbf{E}(\hat{V}_T\mathbf{1}_{S<T}) = \mathbf{E}(V_T\mathbf{1}_{S<T}).$$

D'où la majoration de $\mathbf{E}(V_S)$ par $\mathbf{E}(V_T)$.

∎

Á titre d'exercice, on pourra considérer le problème suivant. On cherche à quel instant optimal vendre sa voiture, sachant que chaque jour, elle a une décote constante δ, que chaque jour d'utilisation, on en tire un bénéfice d'usage $b > \delta$, et que la probabilité qu'elle tombe définitivement en panne d'ici demain est p. Dans ce cas, si X_n est sa valeur d'usage au temps n, sa valeur au temps $n+1$ sera $(X_n + b - \delta)(1 - \varepsilon_{n+1})$, où ε_n est une suite de variables aléatoires indépendantes, prenant les valeurs 1 ou 0 avec les probabilités p et $1-p$ ($0 < p < 1$). Sur cet exemple, on pourra vérifier que $\hat{V}_k = F_{n-k}(X_k)$, où la fonction F_k est une fonction affine par morceaux. On pourra aussi vérifier qu'elle vonverge vers l'infini lorsque k converge vers l'infini, ce qui montre que sur ce modèle, à horizon infini, on n'a jamais intérêt à vendre sa voiture.

Enfin, nous énonçons une dernière conséquence du théorème d'arrêt.

Corollaire 4.15. *(Identité de Wald) Soit* $(Y_n, n \in \mathbb{N})$ *une suite de variables aléatoires réelles adaptées, de même loi,d'espérance* m, *et pour tout* n, Y_n *est indépendante de* $\mathcal{F}_{n-1}$. *On pose* $S_n = Y_0 + \cdots + Y_n$. *Alors, si* T *un* $\mathcal{F}$*-temps d'arrêt intégrable,* $E[S_T] = E[Y_1]E[T]$.

Démonstration. — On montre que le processus $(X_n = S_n - nm, n \geq 0)$ est une martingale. Alors, le théorème est immédiat pour des temps d'arrêt bornés.

Puis on traite d'abord le cas où les variables Y_i sont positives, en appliquant le théorème à $T \wedge n$ et en passant à la limite par convergence monotone. Enfin, pour passer au cas général, on observe que $|S_{T\wedge n}|$ est majorée par $\hat{S}_T$, où $\hat{S}_n = \sum_0^n |Y_i|$, variable intégrable d'après ce qui précède. On peut donc appliquer le théorème de convergence dominée. (On aurait pu tout aussi bien appliquer la décomposition de Krickeberg.) ∎

Le théorème suivant permet de s'affranchir de l'hypothèse que les temps d'arrêt sont presque sûrement bornés.

Théorème 4.16. *(DOOB) : Soit* $(X_n, n \in \mathbb{N})$ *une* $\mathcal{F}$*-martingale et* S *et* T *deux* $\mathcal{F}$*-temps d'arrêt tels que :*

(i) $E[|X_T|]$ *et* $E[|X_{T\wedge S}|]$ *sont finies,*

(ii) $\lim_{n\to+\infty}\int_{\{T>n\}}|X_n|d\mathbf{P} = \lim_{n\to+\infty}\int_{\{T\wedge S>n\}}|X_n|d\mathbf{P} = 0,$

Alors $E[X_T/\mathcal{F}_{T\wedge S}] = X_S.$

Démonstration. — Soit $A \in \mathcal{F}_S$ et n entier. Appliquons le théorème d'arrêt à $T\wedge n$

$$E[X_{T\wedge n}\mathbf{1}_A] = E[X_{T\wedge n\wedge S}\mathbf{1}_A].$$

Traitons le premier membre de l'égalité précédente :

$$E[X_{T\wedge n}\mathbf{1}_A] = \mathbf{E}[X_T\mathbf{1}_A\mathbf{1}_{T\leq n}) + \mathbf{E}(X_n\mathbf{1}_{T>n}\mathbf{1}_A.$$

En faisant converger n vers $+\infty$, le premier terme converge vers $\mathbf{E}(X_T\mathbf{1}_A)$ par convergence dominée, et le second vers 0 par l'hypothèse.

On fait de même avec le second membre et on obtient l'identité cherchée.

∎

Exemple : Si la famille de v.a. $(X_n, n\in\mathbb{N})$ est uniformément intégrable, alors les conditions ci-dessus (i) et (ii) sont satisfaites. En effet, on a alors par définition, $\sup_n E[|X_n|] = K < +\infty$ et $\sup_n\int_A|X_n|d\mathbf{P}\to 0$ quand $\mathbf{P}(A)\to 0$. On en déduit

$$\int_{T>n}|X_n|d\mathbf{P} \leq \sup_m\int_{T>n}|X_m|d\mathbf{P}\to 0$$

puisque T est presque sûrement fini ; par ailleurs, puisque $(|X_n|)$ est une sous-martingale, et que $T\wedge n$ est fini, on a $\mathbf{E}[|X_{T\wedge n}|] \leq \mathbf{E}(|X_n|) \leq K$, et donc, par le lemme de Fatou, $\mathbf{E}(|X_T|)\leq K$.

Donc, en particulier, une martingale uniformément intégrable vérifie le théorème de Doob pour tout couple de *temps d'arrêt* S et T presque sûrement finis.

On va voir plus bas qu'on pourra s'affranchir dans ce cas de l'hypothèse de finitude sur les *temps d'arrêt*.

4.5 Inégalités.

Théorème 4.17. *(Inégalité Maximale de Doob.) Soit* $(X_n, n\in\mathbb{N})$ *une sous martingale positive et* $\lambda\geq 0$. *Apelons* $X^*_n = \sup_{k=0}^n X_k$. *Alors*

$$\forall n\in\mathbb{N},\ \lambda\mathbf{P}\{X^*_n\geq\lambda\} \leq \mathbf{E}[X_n\mathbf{1}_{\{X^*_n\geq\lambda\}}] \leq E[X_n].$$

Démonstration. — Considérons le *temps d'arrêt* $T = \inf\{k \in \mathbb{N}, X_k \geq \lambda\}$. Alors,

$$\{T \leq n\} = \{X_n^* \geq \lambda\}.$$

Posons $S = T \wedge (n+1)$, qui est un *temps d'arrêt* borné. On a

$$A = \{S \leq n\} = \{T \leq n\} \in \mathcal{F}_S.$$

Ecrivons le théorème d'arrêt pour cette sous-martingale, entre les temps n et S.

$$\mathbf{E}(X_{S\wedge n}\mathbf{1}_A) \leq \mathbf{E}(X_n \mathbf{1}_A).$$

Cela s'écrit encore

$$\mathbf{E}(X_T \mathbf{1}_{T\leq n}) \leq \mathbf{E}(X_n \mathbf{1}_{T\leq n}).$$

Or, sur $\{T \leq n\}$, $X_T \geq \lambda$, et on obtient l'inégalité cherchée en minorant le membre de gauche de l'inégalité précédente. ∎

Corollaire 4.18. *Si* $(X_n, n \in \mathbb{N})$ *est une martingale,* $(|X_n|, n \in \mathbb{N})$ *est une sous martingale positive et on obtient donc*

$$\forall n \in \mathbb{N},\ \lambda \mathbf{P}\{\max_{k\leq n} |X_k| \geq \lambda\} \leq E[|X_n|\mathbf{1}_{\max_{k\leq n}|X_k|\geq\lambda}] \leq E[|X_n|].$$

Théorème 4.19. *Soit* $(X_n, n \in \mathbb{N})$ *une sous martingale positive et* $p > 1$. *Alors*

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

Démonstration. — Pour toute variable aléatoire positive U de L^p, on a

$$E(U^p) = p\int_0^\infty t^{p-1}\mathbf{P}(U > t)dt.$$

On a donc

$$E[(X_n^*)^p] = r\int_0^\infty t^{p-1}\mathbf{P}(X_n^* \geq t)dt,$$

et on applique le théorème précédent qui permet la majoration

$$p\int_0^\infty t^{p-2}E[X_n\mathbf{1}_{\{X_n^*\geq t\}}]dt = pE[X_n\int_0^\infty t^{p-2}\mathbf{1}_{\{X_n^*\geq t\}}dt] = pE[X_n(X_n^*)^{p-1}/(p-1)]$$

que l'on majore par l'inégalité de Hölder

$$E[X_n(X_n^*)^{p-1}] \leq \|X_n\|_p\|X_n^*\|_p^{p-1}.$$

Puisque X_n^* est majorée par $\sum_0^n X_k$, c'est une variable de L^p, et on peut donc simplifier pour obtenir le résultat. ∎

En particulier, on en déduit

Corollaire 4.20. *Soit (X_n) une sous-martingale positive bornée dans L^1. Alors la variable $X^* = \sup_n X_n$ est finie presque sûrement. Si (X_n) est bornée dans L^p ($p > 1$), alors X^* est dans L^p.*

(Ce dernier résultat est faux pour $p = 1$.)

Les mêmes conclusions sont vraies pour les martingales (non nécessairement positives).

Démonstration. — La suite X_n^* converge en croissant vers X^*. Il suffit donc d'appliquer l'inégalité de Doob pour obtenir

$$\lambda \mathbf{P}(X_n^* > \lambda) \leq \sup_n \mathbf{E}(|X_n|) = K < \infty.$$

En passant à la limite

$$\lambda \mathbf{P}(X^* > \lambda) \leq K,$$

et donc $\mathbf{P}(X^* > \lambda) \to 0$ ($\lambda \to \infty$). Ceci montre la finitude de X^*.

La seconde assertion se traîte de même en utlisant le théorème 4.19.

∎

Remarque. — La démonstration du théorème 4.19, on on aurait remplacé la fonction t^p par la fonction $t \log_+ t$, nous dirait que, si $\sup_n \mathbf{E}(X_n \log_+ X_n) < \infty$, alors X^* est dans L^1. C'est la meilleure condition possible sur l'intégrabilité des variables X_n qui assure l'intégrabilité de X^*.

4.6 Convergence des martingales

Dans cette section, nous allons démontrer le théorème fondamental de convergence presque sûre des martingales.

Commençons par établir une inégalité qui contrôle le nombre d'oscillations d'une sous-martingale entre deux valeurs, et plus exactement son nombre de descentes à travers un intervalle.

Soit $x = (x_n)$ une suite de réels. Pour tout $n \geq 0$, pour tout couple de réels (a, b) tels que $a < b$, on définit le **nombre de descentes** $D_n^{a,b}(x)$ de la suite x sur $\{0, \cdots, n\}$ à travers l'intervalle (a, b) en comptant les passages au dessus de b et en dessous de a de la façon suivante : on pose $A_0 = B_0 = 0$, puis, par récurrence

$$B_n = \inf\{p \geq A_{n-1} \mid x_p > b\}; \; A_n = \inf\{p \geq B_n \mid x_p < a\},$$

avec comme d'habitude $\inf\{\emptyset\} = +\infty$, par convention. Alors,

$$D_n^{a,b}(x) = \sum_{p=1}^{\infty} 1_{\{A_p < n\}}.$$

Ce nombre compte exactement le nombre de passages au dessus de b de la suite qui sont suivis strictement avant n d'un passage en dessous de a.

Le nombre total de passages sera défini comme la limite croissante $D^{a,b}(x)$ des nombres $D_n^{a,b}(x)$.

Remarquons que la somme qui apparaît dans la définition de D_n est toujours finie, car $A_n > B_{n-1} > A_{n-1}$, si bien que $A_n \geq A_{n-1} + 2$, et donc $A_n \geq 2n$.

D'autre part, si $X = (X_n)$ est une suite de variables aléatoires, alors $D_n^{a,b}(X)$ est une variable aléatoire; lorsque n croît, ces variables forment une suite croissante. De plus, si la suite (X_n) est adaptée à une filtration $\mathcal{F}_n$, alors les variables aléatoires A_n et B_n intervenant dans la définition de $D_n^{a,b}(X)$ sont des *temps d'arrêt*.

L'intérêt de ces nombres de descentes est le suivant.

Si $x = (x_n)$ une suite de nombres réels, posons $x^* = \sup_n |x_n|$. Alors, on a le

Lemme 4.21. *Soit (x_n) une suite de nombres réels. Une CNS pour que la suite converge est que x^* soit fini et que, pour tout couple de rationnels (a, b) tel que $a < b$, la quantité $D^{a,b}(x)$ soit finie.*

Remarque. — Nous nous sommes restreints aux couples de rationnels dans l'énoncé précédent pour des raisons de dénombrabilité. N'importe quel ensemble dense dans $\mathbb{R}$ aurait aussi bien fait l'affaire.

Démonstration. — Montrons que la condition est suffisante. Si x^* est fini, la suite est bornée : appelons M et m les limites supérieures et inférieures de x, qui sont finies. Si $m < M$, choisissons deux rationnels a et b tels que $m < a < b < M$. Alors, on voit immédiatement que $D^{a,b}(x) = \infty$. D'où une contradiction et donc $m = M$, et la suite converge.

Réciproquement, une suite convergente est bornée, et, si un couple (a, b) avec $a < b$ est tel que $D^{a,b}(x) = \infty$, alors $\liminf_n x_n \leq a$ et $\limsup_n x_n \geq b$, ce qui est impossible. ∎

Nous pouvons maintenant énoncer le résultat sur les descentes

Proposition 4.22. *Soit (X_n) une sous martingale positive. Considérons un couple (a, b) avec $a < b$ de réels et appelons D_n le nombre de descentes de X à travers* *l'intervalle* (a, b) *sur* $\{0, \cdots, n\}$. *Alors,*

$$\mathbf{E}(D_n) \leq \frac{1}{b-a} \mathbf{E}[(X_n - b)_+].$$

Ici, comme d'habitude, $(X_n - b)_+$ désigne la partie positive de $X_n - b$.

Démonstration. — Fixons n, et commençons par arrêter la sous-martingale à n (c'est à dire remplaçons X par X^n.) Considérons les *temps d'arrêt* intervenant dans la définition des montées, que nous notons A_p et B_p pour simplifier. Rappelons que $B_p < A_p$ (on compte par convention les passages au dessus de b avant les passages en dessous de a). Posons $A_p^n = A_p \wedge n$ et $B_p^n = B_p \wedge n$, qui sont des *temps d'arrêt* bornés.

D'après le théorème d'arrêt, nous avons

$$\mathbf{E}(X_{A_p^n} - X_{B_p^n}) \geq 0.$$

Décomposons cette espérance sur les ensembles $\{A_p < n\}$ et $\{A_p \geq n\}$. Nous obtenons

$$\mathbf{E}[(X_{B_p^n} - X_{A_p^n})\mathbf{1}_{\{A_p<n\}}] \leq \mathbf{E}[(X_{A_p^n} - X_{B_p^n})\mathbf{1}_{\{A_p \geq n\}}].$$

Or, sur $\{A_p < n\}$,

$$X_{B_p^n} - X_{A_p^n} = X_{B_p} - X_{A_p} \geq b - a.$$

D'autre part, sur $\{A_p \geq n\}$, $X_{A_p^n} - X_{B_p^n} = X_n - X_{B_p^n}$. Cette dernière quantité est nulle sur $\{B_p \geq n\}$. Sur $\{B_p < n\}$, elle est inférieure à $X_n - b$ donc à $(X_n - b)_+$.

On obtient donc

$$(b-a)\mathbf{E}(\mathbf{1}_{\{A_p<n\}}) \leq \mathbf{E}[(X_n - b)_+ \mathbf{1}_{\{B_p<n\leq A_p\}}].$$

Remarquons encore que les ensembles $\{B_p < n \leq A_p\}$ sont disjoints lorsque p varie, puis sommons en p l'inégalité précédente. Le membre de gauche devient $(b - a)\mathbf{E}(D_n)$ par définition du nombre de descentes, tandis que le second est inférieur à $\mathbf{E}[(X_n - b)_+]$. ∎

Corollaire 4.23. *Soit* (X_n) *une sous-martingale positive, d'espérance bornée. Alors, pour tout couple* (a, b), $a < b$, *le nombre de descentes de la suite* (X_n) *à travers* (a, b) *est presque sûrement fini, et la suite* (X_n) *est presque sûrement convergente.*

Démonstration. — Montrons tout d'abord le premier point. Soit M un majorant des $\mathbf{E}(X_n)$, et un couple de réels (a, b) tel que $a < b$. En appelant D_n le nombre de descentes à travers (a, b) sur l'intervalle $\{0, n\}$, et D le nombre total de descentes, on voit que

$$\mathbf{E}(D_n) \leq \frac{1}{b-a}(M + |b|).$$

En passant à la limite en n, nous voyons qu'alors D est intégrable donc finie presque sûrement.

Appelons $N_{a,b}$ l'ensemble sur lequel le nombres total de descentes à travers (a, b) est infini, et N l'union de tous les ensembles $N_{a,b}$ pour tous les couples (a, b) de rationnels tels que $a < b$: c'est une union dénombrable de négligeables et c'est donc un ensemble négligeable.

D'autre part, nous savons d'après le corollaire 4.20 que la variable $X^* = \sup_n X_n$ est finie presque sûrement. Soit N_1 l'ensemble où elle est infinie.

L'ensemble $N \cup N_1$ est négligeable, et, sur le complémentaire de cet ensemble, la suite $(X_n(\omega))$ satisfait aux hypothèse du lemme (rappelons que (X_n) est positive par hypothèse): elle y est donc convergente. ∎

Le résultat précédent se prolonge immédiatement aux martingales bornées dans L^1, aux sous-et sur-martingales bornées dans L^1 non nécessairement positives :

Corollaire 4.24. *Soit X_n une martingale, ou une sous-martingale, ou une sur-martingale, bornée dans L^1. Alors, X_n converge presque sûrement vers une variable X_∞.*

Démonstration. — Commençons par le cas des martingales : les fonctions $x_+ = \sup(x, 0)$ et $x_- = \sup(-x, 0)$ sont convexes: donc, si $M_,$ est une martingale bornée dans L^1, les processus M_{n+} et M_{n-} sont des sous-martingales positives bornées dans L^1. Elles convergent donc presque sûrement et il en va de même de la martingale $M_n = M_{n+} - M_{n-}$.

Traitons ensuite le cas d'une sous-martingale (X_n) bornée dans L^1 : nous avons vu plus haut qu'elle peut s'écrire $X_n = M_n + A_n$, où (M_n) est une martingale et (A_n) est une suite croissante nulle en 0.

Si nous posons $K = \sup_n \mathbf{E}(|X_n|)$, on a $E(|X_0|) \leq K$, et donc

$$|\mathbf{E}(M_n)| = |\mathbf{E}(M_0)| = |\mathbf{E}(X_0)| \leq K,$$

et donc

$$\mathbf{E}(|A_n|) = \mathbf{E}(A_n) = \mathbf{E}(X_n - M_n) \leq 2K,$$

et A_n est bornée dans L^1, et donc également M_n par différence.

Il s'ensuit que M_n converge presque sûrement, ainsi que A_n (une suite croissante positive bornée dans L^1 est presque sûrement convergente vers une limite finie et même intégrable, par le théorème de convergence monotone). Par différence, X_n est presque sûrement convergente.

Le cas des sur-martingales est une conséquence du cas des sous-martingales, par changement de signe. ∎

Remarquons qu'une conséquence immédiate est qu'une sur-martingale positive est toujours presque sûrement convergente, (car elle est automatiquement bornée dans L^1), ainsi qu'une sur-martingale minorée (par une constante), ou qu'une sous-martingale majorée.

Par le lemme de Fatou, il est clair que la limite d'une martingale bornée dans L^1 est une variable intégrable M_∞. Il n'est pas sûr par contre que $M_n = \mathbf{E}(M_\infty/\mathcal{F}_n)$. C'est le cas des martingales uniformément intégrables, comme le montre la proposition suivante. (Nous renvoyons le lecteur à la section 2.3 pour les propriétés élémentaires des familles uniformément intégrables.)

Proposition 4.25. *Soit M_n une martingale bornée dans L^1, et soit M_∞ la limite de M_n lorsque $n \to \infty$. Les propositions suivantes sont équivalentes*

1. *M_n converge dans L^1 vers M_∞.*
2. *M_n est uniformément intégrable.*
3. *$M_n = \mathbf{E}[M_\infty/\mathcal{F}_n]$.*
4. *Il existe une variable intégrable M telle que $M_n = \mathbf{E}[M/\mathcal{F}_n]$. De plus, dans ce cas, $M_\infty = \mathbf{E}[M/\mathcal{F}_\infty]$.*

(Dans ce qui précède, nous avons supposé que $\mathcal{F}_\infty = \vee_n \mathcal{F}_n$.)

Démonstration. — Nous savons que, pour une suite de variables intégrables qui converge presque sûrement, il est équivalent de converger dans L^1 ou d'être uniformément intégrable. De plus, pour toute variable aléatoire M intégrable, l'ensemble des $\mathbf{E}(M/\mathcal{B})$, où $\mathcal{B}$ parcourt les sous tribus de $\mathcal{A}$ forme une famille uniformément intégrable. Il ne reste donc à montrer que les points suivants :

1. Si (M_n) est uniformément intégrable, alors $M_n = \mathbf{E}(M_\infty/\mathcal{F}_n)$;
2. Si M est intégrable, alors la martingale $M_n = \mathbf{E}(M/\mathcal{F}_n)$ converge vers $\mathbf{E}(M/\mathcal{F}_\infty)$.

Commençons par le premier point: nous écrivons, pour $p \geq n$ $M_n = \mathbf{E}(M_p/\mathcal{F}_n)$, puis nous passons à la limite en p: l'espérance conditionnelle étant continue dans L^1, et puisque M_p converge vers M_∞ dans L^1 par hypothèse, nous obtenons le résultat.

Pour le second point, remarquons d'abord que M_∞ est $\mathcal{F}_\infty$ mesurable par construction. Il suffit donc de montrer que, pour tout $A \in \mathcal{F}_\infty$, nous avons $\mathbf{E}(M_\infty \mathbf{1}_A) = \mathbf{E}(M\mathbf{1}_A)$. Or, ceci est vrai lorsque A est dans l'une des sous-tribus $\mathcal{F}_n$, puisqu'alors

$$\mathbf{E}(M\mathbf{1}_A) = \mathbf{E}(M_n\mathbf{1}_A) = \mathbf{E}(M_\infty\mathbf{1}_A).$$

Alors, l'identité cherchée, qui est vraie pour tous les éléments de l'algèbre $\cup_n \mathcal{F}_n$, reste vraie pour tous les éléments de la σ -algèbre engendrée par cette algèbre, qui est $\mathcal{F}_\infty$. (Argument de classes monotones.) ∎

Remarques

1. Un énoncé analogue à la proposition 4.25 est vrai pour les sous- et sur-martingales; nous en laissons le soin au lecteur.
2. Si la martingale est bornée dans L^p pour un $p > 1$, alors elle est dominée par une variable de L^p et elle converge alors dans L^p.

Un des intérêts principaux des martingales uniformément intégrables est que les résultats énoncés auparavant pour les temps d'arrêt bornés ou finis presque sûrement s'étendent aux temps d'arrêt quelconques. C'est en particulier vrai pour le plus important d'entre eux, le théorème d'arrêt :

Théorème 4.26. *(***Théorème d'arrêt, deuxième version.***) Soit M_n une martingale uniformément intégrable et soit T un temps d'arrêt quelquonque; alors, en définissant $M_T = M_\infty$ sur $\{T = \infty\}$, on a*

1. *$M_T = \mathbf{E}(M_\infty/\mathcal{F}_T)$.*
2. *L'ensemble des variables (M_T), où T est un temps d'arrêt quelconque, est uniformément intégrable.*
3. *Si S et T sont des temps d'arrêt quelconques, on a*
$$\mathbf{E}(M_T/\mathcal{F}_S) = M_{S\wedge T}.$$
4. *Pour une variable $\mathcal{A}$-mesurable M intégrable, en posant $M_n = \mathbf{E}(M/\mathcal{F}_n)$, on a $M_T = \mathbf{E}(M/\mathcal{F}_T)$.*

Démonstration. — Pour le premier point, il suffit de réécrire la démonstration du théorème d'arrêt dans ce cas. Si $\mathcal{A}$ est dans $\mathcal{F}_T$, alors

$$\begin{aligned}\mathbf{E}(M_T\mathbf{1}_A) &= \sum_{k\in\mathbb{N}\bigcup\infty}\mathbf{E}(M_k\mathbf{1}_{A\bigcap\{T=k\}})\\ &= \sum_{k\in\mathbb{N}\bigcup\infty}\mathbf{E}(M_\infty\mathbf{1}_{A\bigcap\{T=k\}}) = \mathbf{E}(M_\infty\mathbf{1}_A).\end{aligned}$$

Ensuite, la famille des M_T est une sous-famille de la famille uniformément intégrable des $\mathbf{E}(M_\infty/\mathcal{B})$, où$\mathcal{B}$ parcourt les sous-tribus de $\mathcal{A}$.

La martingale arrêtée M^T est donc uniformément intégrable. En lui appliquant le théorème d'arrêt au temps S, on obtient

$$\mathbf{E}(M_T/\mathcal{F}_S) = M_{S\wedge T}.$$

Enfin, il suffit d'écrire

$$\mathbf{E}(M/\mathcal{F}_T) = \mathbf{E}(\mathbf{E}(M/\mathcal{F}_\infty)/\mathcal{F}_T) = \mathbf{E}(M_\infty/\mathcal{F}_T) = M_T.$$

∎

Enfin, la convergence dans L^2 des martingales bornées dans L^2 est beaucoup plus simple à obtenir, comme nous l'avons déjà remarqué, et n'exige pas l'inégalité de Doob. Bien que ce résultat soit inclus dans ceux qui précède, cela vaut sans doute la peine de l'énoncer séparément.

Théorème 4.27. *Soit $(X_n, n\in\mathbb{N})$ une sous martingale positive bornée dans L^2. Alors la suite $(X_n, n\in\mathbb{N})$ converge dans L^2.*

Remarques

1. Il est important de remarquer que la condition d'intégrabilité uniforme est bien nécessaire pour obtenir le théorème d'arrêt. En effet, reprenons l'exemple de la ruine du joueur $M_n = 1+\sum_1^n X_p$, avec des variables X_p de Bernoulli, indépendantes et centrées. Appelons $T=\inf\{n|\ X_n=0\}$. Il n'est pas difficile de voir que T est fini presque sûrement (exercice!). Alors, la martingale arrêtée M^T converge presque sûrement vers 0, et donc $M_\infty=0$, alors que $M_0=1$.

   Or, elle est bien bornée dans L^1 puisque $\mathbf{E}(|M_n^T|) = \mathbf{E}(|M_{T\wedge n}|) = \mathbf{E}(M_{T\wedge n})$ (puisque $M_{T\wedge n}\geq 0$) $= \mathbf{E}(M_0) = 1$.

2. La décomposition de Krickeberg des martingales uniformément intégrables est immédiate : on décompose M_∞ en ses parties positives et négatives M_∞^+ et M_∞^-, et on écrit

$$M_n = \mathbf{E}(M_\infty^+/\mathcal{F}_n) - \mathbf{E}(M_\infty^-/\mathcal{F}_n).$$

4.7 Martingales inverses

(Neveu pages 115-119)

Les martingales inverses jouent un rôle important en théorie des martingales à temps continu. Mais elles ont aussi une utilité directe (voir plus bas la preuve de la loi forte des grands nombres). De plus, elles sont baucoup plus simples à étudier que les martingales ordinaires.

Sur l'espace de probabilité $(\Omega, \mathcal{A}, \mathbf{P})$, on considère une suite décroissante de tribus : $(\mathcal{F}_n, n \in \mathbb{N})$. On pose alors $\mathcal{F}_\infty = \cap_n \mathcal{F}_n$.

Définition 4.28. *Une suite de variables aléatoires $(X_n, n \in \mathbb{N})$ sur $(\Omega, \mathcal{A}, \mathbf{P})$ est une* **(sous)martingale inverse** *si pour tout n, X_n est $\mathcal{F}_n$-mesurable, intégrable, et $E(X_n/\mathcal{F}_{n+1})(\geq) = X_{n+1}$.*

Ces processus ont des propriétés analogues à celles des martingales comme par exemple : une fonction convexe d'une martingale est une sous-martingale, l'espérance d'une martingale est constante, l'espérance d'une sous martingale est décroissante. Il existe également une décomposition de Doob.

En fait, il suffit de changer n en $-n$ pour retrouver les propriétés des martingales ordinaires, mais cette fois-ci l'ensemble des temps est $\mathbb{Z}^-$, l'ensemble des entiers négatifs.

Proposition 4.29. *Soit (X_n) une sous-martingale inverse; il existe un unique processus croissant (A_n), tel que pour tout n A_n soit $\mathcal{F}_{n+1}$ mesurable, et tel que*

$$X_n = \mathbf{E}(X_0/\mathcal{F}_n) - A_n.$$

Démonstration. — On fait comme dans le cas des sous-martingales ordinaires. On pose $A_0 = 0$ et on pose

$$A_{n+1} - A_n = \mathbf{E}(X_n/\mathcal{F}_{n+1}) - X_{n+1}.$$

C'est une suite croissante par la propriété de sous-martingale, et on voit que $X_n = M_n - A_n$, où M_n est une martingale inverse. Donc, $M_n = \mathbf{E}(M_0/\mathcal{F}_n)$, et on obtient la décomposition annoncée.

L'unicité est évidente. ∎

Si M est une martingale inverse, alors $M_n = \mathbf{E}(M_0/\mathcal{F}_n)$. Une martingale inverse est donc toujours uniformément intégrable. C'est également le cas d'une sous-martingale inverse .

Proposition 4.30. *Une sous-martingale inverse bornée dans L^1 est uniformément intégrable.*

Démonstration. — On écrit la décomposition de Doob $X_n = \mathbf{E}(X_0/\mathcal{F}_n) - A_n$. Alors, A_n est bornée dans L_1, et croissante donc convergente presque partout vers $A_\infty \in L^1$. On a $0 \leq A_n \leq A_\infty$, et la suite A_n est donc uniformément intégrable.

La martingale inverse $\mathbf{E}(X_0/\mathcal{F}_n)$ est également uniformément intégrable, d'où le résultat. ∎

Tous les résultats obtenus sur les martingales et sous-martingales ordinaires vont s'appliquer (presque) immédiatement aux martingales et sous-martingales inverses, en retournant le temps. Plus précisément, considérons une (sous-)martingale inverse (X_n) et fixons un entier N: Notons $\mathcal{G}_n = \mathcal{F}_{(N-n)\vee 0}$ et $Y_n = X_{(N-n)\vee 0}$. Alors, $\mathcal{G}_n$ est une filtration ordinaire par rapport à laquelle Y_n est une (sous-)martingale ordinaire.

On peut lui appliquer alors les inégalités maximales et les inégalités sur les descentes (qui deviennent des montées dans ce cas) : on obtient les mêmes inégalités qu'avant, mais c'est la variable X_0 qui joue le rôle qui était alors dévolu à X_n.

Théorème 4.31. *Soit (X_n) une sous-martingale inverse positive ou une martingale inverse. Soit $X_N^* = \sup_{k=0}^N |X_k|$.*

Alors, pour tout $\lambda > 0$,

$$\lambda \mathbf{P}(X_N^* \geq \lambda) \leq \mathbf{E}[|X_0| \mathbf{1}_{X_N^* \geq \lambda}].$$

Théorème 4.32. *Soit (X_n) une sous-martingale positive. Considérons deux réels $a < b$ et appelons $U_n^{(a,b)}$ le nombres de montées de la suite X_n à travers l'intervalle $[a, b]$. Alors*

$$(b-a)\mathbf{E}(U_n^{(a,b)}) \leq \mathbf{E}[(X_0 - b)_+].$$

Les théorèmes de convergence s'en déduisent alors immédiatement.

Théorème 4.33. *Soit X_n une martingale inverse, ou une sous-martingale inverse bornée dans L^1.*

Alors X_n converge presque sûrement et dans L^1 vers $X_\infty = (\leq)\mathbf{E}(X_0/\mathcal{F}_\infty)$.

Démonstration. — Il suffit de suivre le shéma de démonstration donné pour les martingales ordinaires. La seule chose à démontrer est que

$$X_\infty = (\leq)\mathbf{E}(X_0/\mathcal{F}_\infty).$$

C'est immédiat en utilisant l'intégrabilité uniforme. ∎

Remarque. — Si nous mettons ensemble les deux résultats de convergence pour les martingales ordinaires et inverses uniformément intégrables, nous obtenons la régularitè de l'espérance conditionnelle par rapport à la "variable tribu" : si $\mathcal{F}_n$ est une suite croissante ou décroissante de tribus qui converge vers $\mathcal{F}_\infty$ (au sens où $\mathcal{F}_\infty = \vee_n \mathcal{F}_n$ dans le cas croissant et $\mathcal{F}_\infty = \cap_n \mathcal{F}_n$ dans le cas décroissant), alors, pour toute variable dans X de L^1, on a

$$\mathbf{E}(X/\mathcal{F}_n) \to \mathbf{E}(X/\mathcal{F}_\infty), \text{ ps et dans } L^1.$$

Exemple : L'application la plus spectaculaire de ce résultat est la loi forte des grands nombres: nous l'énonçons comme exercice :

Soit une suite (Z_n) de variables aléatoires indépendantes intégrables et de même loi. On pose $S_n = \frac{Z_1+\cdots+Z_n}{n}$ et $\mathcal{F}_n = \sigma(S_n, S_{n+1}, \ldots.)$. Alors S est une $\mathcal{F}$-martingale inverse, qui converge vers $\mathbf{E}(Z_1)$, presque sûrement et dans L^1.

Indication : Montrer d'abord que $\mathcal{F}_n = \sigma(S_n) \vee \sigma(Z_{n+1}, Z_{n+2}, \ldots)$. En déduire que $\mathbf{E}(S_n/\mathcal{F}_{n+1}) = \mathbf{E}(S_n/S_{n+1})$, puis la propriété de martingale inverse en remarquant calculant pour tout n $\mathbf{E}(X_1/S_n)$.

En déduire que la suite S converge presque sûrement et dans $L^1(\Omega, \mathcal{A}, \mathbf{P})$ vers $E(S_\infty)$. Utiliser la loi du $\{0,1\}$ pour montrer que $Z_\infty = \mathbf{E}(Z_1)$.

Enfin, le théorème de convergence de martingales montre que la tribu $\mathcal{F}_{T\wedge n}$ converge vers $\mathcal{F}_T$, aux ensembles de mesure nulle près :

Proposition 4.34. *1. Soit T un temps d'arrêt, et soit $\mathcal{F}_{T_-} = \vee_n \mathcal{F}_{T\wedge n}$. Alors, aux ensembles de mesure nulle près, $\mathcal{F}_{T_-} = \mathcal{F}_T$.*

2. Si S et T sont deux temps d'arrêt quelconques, alors $\mathcal{F}_{S\wedge T} = \mathcal{F}_S \wedge \mathcal{F}_T$, aux ensembles de mesure nulle près.

Démonstration. — Commenons par le premier point : il suffit de démontrer que, pour toute variable aléatoire X intégrable (bornée suffirait puisqu'il s'agit d'appliquer ce qui suit à $\mathbf{1}_A$), $\mathbf{E}(X/\mathcal{F}_{T_-}) = \mathbf{E}(X/\mathcal{F}_T)$.

Or, considérons la martingale $M_n = \mathbf{E}(X/\mathcal{F}_n)$. On sait que $\mathbf{E}(X/\mathcal{F}_T) = M_T$, et que $\mathbf{E}(X/\mathcal{F}_{T\wedge n}) = M_{T\wedge n}$. Si $\mathcal{G}_n$ désigne la tribu $\mathcal{F}_{T\wedge n}$, nous avons $\mathcal{G}_\infty = \mathcal{F}_{T_-}$, et, par le théorème de convergence,

$$M_{T\wedge n} = \mathbf{E}(X/\mathcal{G}_n) \to \mathbf{E}(X/\mathcal{G}_\infty) = \mathbf{E}(X/\mathcal{F}_{T_-}).$$

Par ailleurs,

$$M_{T\wedge n} \to M_T = \mathbf{E}(X/\mathcal{F}_T),$$

ces deux convergence ayant lieu presque sûrement. D'où l'égalite.

Pour le second point, nous savions que c'était vrai pour les *temps d'arrêt* finis. Il suffit de tronquer S et T à n et de passer à la limite, en utilisant ce qui précède. ∎

4.8 Algorithme de Robbins-Monro

Exemple : le dosage.

Il s'agit de déterminer le dosage optimal d'un produit chimique pour obtenir l'effet voulu ; on sait, à cet effet, effectuer des tests.

Au dosage $z \in \mathbb{R}$, on associe l'effet $F(z, \eta)$, où η est une variable aléatoire dont la réalisation dépend du test effectué (pas du dosage). On connaît la loi de η et la fonction F.

L'effet moyen du dosage z est donné par $f(z) = E[F(z, \eta)]$; on suppose que f est strictement croissante et continue.

Soit $a \in \mathbb{R}$ l'effet désiré. On cherche à déterminer l'unique z^* tel que $f(z^*) = a$, et ceci sans avoir à calculer la fonction f, calcul qui peut s'avérer impossible dans la pratique.

Pour cela, on se donne une suite de variables aléatoires (η_n) indépendantes et de même loi que η, et une suite γ_n, dont on précisera les propriétés plus bas, mais la suite $\gamma_n = 1/n$ par exemple convient), et on construit une suite de variables aléatoires $(Z_n, n \in \mathbb{N})$ définie par récurrence :

$$Z_{n+1} = Z_n - \gamma_n[F(Z_n, \eta_{n+1}) - a].$$

Alors, si

$$E(F(z,\eta)^2) \leq c(1+z^2),$$

la suite Z_n converge presque sûrement vers la valeur z^*.

En fait, on va démontrer le résultat un peu plus général suivant :

Théorème 4.35. *de ROBBINS-MONRO : Soit $h \in C(\mathbb{R}, \mathbb{R})$, (γ_n) une suite de réels positifs et (ε_n) une suite de variables aléatoires de carré intégrable. On pose $\mathcal{F}_n = \sigma(\varepsilon_1, \cdots, \varepsilon_n)$, et on suppose que*

1. $E(\varepsilon_{n+1}/\mathcal{F}_n) = 0$;
2. $\sum_n \gamma_n = +\infty$; $\sum_n \gamma_n^2 < +\infty$;
3. $E(\varepsilon_{n+1}^2/\mathcal{F}_n) \leq c(1+Z_n^2)$;
4. $|h(z)|^2 \leq c(1+z^2)$;
5. *Il existe un unique $z^* \in \mathbb{R} : h(z^*) = 0, (z-z^*)h(z) < 0$, $\forall z \neq z^*$.*

Alors, si on pose

$$Z_{n+1} = Z_n + \gamma_n[h(Z_n) + \varepsilon_{n+1}], Z_0 \text{ donnée constante,}$$

la suite (Z_n) converge presque sûrement vers z^*.

Pour le problème du dosage, on pose $\varepsilon_{n+1} = f(Z_n) - F(Z_n, \eta_{n+1})$, et $h(z) = a - f(z)$. Les conditions sont alors vérifiées.

Démonstration. — Nous ne donnons que des indications sur les grandes lignes, les détails sont laissés à titre d'exercice.

Première étape : Il existe deux constantes α et β telles que, pour tout n, on ait la majoration

$$E[(Z_{n+1} - z^*)^2/\mathcal{F}_n] \leq (Z_n - z^*)^2(1 + \alpha\gamma_n^2) + \gamma_n^2\beta.$$

En effet, d'une part les hypothèses montrent par récurrence que Z_n est de carré intégrable pour tout n, et d'autre part :

(4.1)
$$(Z_{n+1} - z^*)^2 = (Z_n - z^*)^2 + 2\gamma_n(Z_n - z^*)(h(Z_n) + \varepsilon_{n+1}) + \gamma_n^2(h(Z_n) + \varepsilon_{n+1})^2$$

que l'on conditionne par $\mathcal{F}_n$:

$$\begin{aligned} E[(Z_{n+1}-z^*)^2/\mathcal{F}_n] &= \\ &= (Z_n-z^*)^2+2\gamma_n(Z_n-z^*)h(Z_n)+ \\ &\quad \gamma_n^2(h(Z_n)^2+E[\varepsilon_{n+1}^2/\mathcal{F}_n]. \end{aligned}$$

Le deuxième terme est négatif par (iii) et le dernier se majore par (ii) et (iii) :

$$E[(Z_{n+1}-z^*)^2/\mathcal{F}_n] \leq (Z_n-z^*)^2+\gamma_n^2 2c(1+Z_n^2).$$

Soit une majoration de la forme $(Z_n-z^*)^2(1+\alpha\gamma_n^2)+\gamma_n^2\beta$.

Deuxième étape : on pose $u_n=\Pi_{k=1}^{n-1}(1+\alpha\gamma_k^2)$, suite convergente puisque la série $\sum_n \gamma_n^2$ converge ; et $v_n=\sum_{k=1}^{n-1}\frac{\beta\gamma_k^2}{u_{k+1}}$ qui converge aussi puisque $u_{k+1}\geq 1$ et la série $\sum_n\gamma_n^2$ converge.

On définit la suite de variables aléatoires $X_n=\frac{(Z_n-z^*)^2}{u_n}-v_n$, soit

$$(Z_n-z^*)^2=u_n(X_n+v_n).$$

On vérifie que X est une surmartingale : X_n est $\mathcal{F}_n$-mesurable par définition et intégrable puisque Z_n est de carré intégrable ; on a

$$\begin{aligned} E[X_{n+1}/\mathcal{F}_n] &= \frac{1}{u_{n+1}}(E[(Z_{n+1}-z^*)^2/\mathcal{F}_n])-v_{n+1} \\ &\leq \frac{1}{u_{n+1}}[(Z_n-z^*)^2(1+\alpha\gamma_n^2)+\gamma_n^2\beta]-v_{n+1} \\ &= X_n. \end{aligned}$$

De plus, elle est minorée par la suite convergente $(-v_n, n\in\mathbb{N})$. C'est donc une sur-martingale bornée dans L^1 et elle converge presque sûrement.

Troisième étape : La suite $((Z_n-z^*)^2, n\in\mathbb{N})$ converge vers $(Z_\infty-z^*)^2$. Il faut montrer que cette limite est nulle. On reprend l'équation (4.1) en gardant le deuxième terme négatif que l'on passe à gauche et il vient :

$$0\leq 2\gamma_n(z^*-Z_n)h(Z_n)\leq (Z_n-z^*)^2(1+\alpha\gamma_n^2)+\gamma_n^2\beta-E[(Z_{n+1}-z^*)^2/\mathcal{F}_n]$$

que l'on réexprime avec les suites (u_n) et (v_n) :

$$0\leq 2\gamma_n(z^*-Z_n)h(Z_n)\leq u_n(X_n+v_n)\frac{u_{n+1}}{u_n}+\beta\gamma_n^2-u_{n+1}(E[X_{n+1}/\mathcal{F}_n]+v_{n+1}),$$

ce qui donne

$$0 \leq 2\gamma_n(z^* - Z_n)h(Z_n) \leq u_{n+1}(X_n - E[X_{n+1}/\mathcal{F}_n]).$$

Ces termes sont tous positifs. La série des espérances $u_{n+1}(E[X_n] - E[X_{n+1}])$ est convergente, car le terme général de la série des espérances est qui est convergente puisque $E(X_n)$ est une suite décroissante et u_n est une suite convergente.

Donc la série de terme général $\gamma_n(z^* - Z_n)h(Z_n)$ est positive, la série de ses espérances converge, et elle converge donc ps et dans L^1.

La divergence de la série de terme général γ_n et l'unicité du zéro de h permet alors de conclure. ∎

4.9 Compléments sur les martingales de carré intégrable.

Rappelons que, si M_n est une martingale de carré intégrable, il existe un unique processus croissant prévisible, nul en 0, tel que $M_n - A_n$ soit une martingale. Nous le noterons $\langle M \rangle_n$. Rappelons sa valeur

$$\langle M \rangle_n = \sum_{k=1}^{n} \mathbf{E}[(M_k - M_{k-1})^2/\mathcal{F}_{k-1}].$$

Nous l'appelerons "la variation quadratique" de M.

Il est immédiat que, si T est un *temps d'arrêt*, la variation quadratique de la martingale arrêtée M^T est le processus arrêté $\langle M \rangle^T$, et, comme nous l'avons vu

$$\mathbf{E}(M_n - M_0)^2 = \mathbf{E}(\langle M \rangle_n).$$

Remarquons que cette identité est encore vraie pour $n = \infty$, par un passage à la limite (avec la convention que $\mathbf{E}[(M_\infty - M_0)^2] = \infty$ si M n'est pas bornée dans L^2.)

Cette identité s'étend immédiatement aux *temps d'arrêt*, bornés ou non.

Proposition 4.36. *Si T est un temps d'arrêt, alors*

$$\mathbf{E}[(M_T - M_0)^2] = \mathbf{E}(\langle M \rangle_T].$$

Démonstration. — C'est immédiat pour un temps borné, car on applique l'identité précédente à la martingale M^T, et à un temps n qui majore T.

Si T n'est pas borné, nous appliquons ce qui précède au *temps d'arrêt* $T \wedge n$, et on passe à la limite.

∎

On voit donc qu'il suffit de savoir que $\langle M\rangle_T$ est intégrable pour en déduire que la martingale est bornée dans L^2.

Dans la pratique, il est souvent possible que le processus $\langle M\rangle_n$ soit très facile à calculer : par exemple, si M_n est une somme de variables aléatoires indépendantes et centrées $\sum_0^n \varepsilon_k$, alors $\langle M\rangle_n = \sum_0^n \mathbf{E}(\varepsilon_k^2)$.

On peut améliorer ce résultat pour décider si la martingale arrêtée M^T est uniformément intégrable. On a

Théorème 4.37. *(cf [8], p.148)*

1. *M_n converge presque sûrement vers une limite M_∞ sur l'événement $\{\langle M\rangle_\infty < \infty\}$.*
2. *Si M est une martingale de carré intégrable nulle en 0, et si $M^* = \sup_n |M_n|$, alors*

$$\mathbf{E}(M^*) \leq 3\mathbf{E}(\sqrt{\langle M\rangle_\infty}).$$

3. *En particulier, M^T est uniformément intégrable dès que $\sqrt{\langle M\rangle_T}$ est intégrable.*

Démonstration. — Commencons par traiter le cas des martingales bornées dans L^2. Il suffira ensuite d'arrêter la martingale en n, puis de passer à la limite pour obrtenir le résultat. Fixons $a > 0$ et posons

$$T_a = \inf\{n|\ \langle M\rangle_{n+1} > a^2\},$$

avec comme d'habitude $\inf\{\emptyset\} = +\infty$. Nous avons donc

$$\{T_a = \infty\} = \{\langle M\rangle_\infty \leq a^2\}.$$

D'autre part, $\langle M\rangle_{T_a} \leq a^2$ sur $\{T_a < \infty\}$: on voit donc que $\langle M\rangle_{T_a} \leq \langle M_\infty \wedge a^2$, et que M^{T_a} est bornée dans L^2. Elle converge donc vers une limite. En faisant converger a vers l'infini, nous voyons déjà la convergence de M_n sur $\{\langle M\rangle_\infty < \infty\}$.

Ensuite, pour démontrer l'inégalité

$$\mathbf{E}(M^*) \leq 3\mathbf{E}(\sqrt{\langle M\rangle_\infty}),$$

nous écrivons

$$\begin{aligned} \mathbf{P}(M^* > a) &\leq \mathbf{P}(T_a < \infty) + \mathbf{P}(T_a = \infty; M^* > a) \\ &\leq \mathbf{P}(T_a < \infty) + \mathbf{P}((M^{T_a})^* > a). \end{aligned}$$

On applique ensuite l'inégalité de Doob à la sous-martingale M^{2T_a}, pour obtenir

$$\mathbf{P}((M^{T_a})^* > a) \leq \frac{1}{a^2}\mathbf{E}(M^2_{T_a}) = \frac{1}{a^2}\mathbf{E}(\langle M\rangle_{T_a}).$$

Mais $\mathbf{E}(\langle M\rangle_{T_a}) \leq \mathbf{E}(\langle M\rangle_\infty \wedge a^2)$;

En reportant cette majoration dans ce qui précède, on obtient

$$\mathbf{P}(M^* > a) \leq \mathbf{P}(\langle M\rangle_\infty > a^2) + \frac{1}{a^2}\mathbf{E}(\langle M\rangle_\infty \wedge a^2).$$

En intégrant cette inégalité par rapport à la mesure de Lebesgue en a, sur $[0,\infty[$, on obtient

$$\mathbf{E}(M^*_\infty) \leq 3\mathbf{E}(\sqrt{\langle M\rangle_\infty}).$$

Pour obtenir le dernier point, on applique le résultat à la martingale arrêtée M^T.

∎

Nous avons vu que M converge vers une valeur finie sur $\{\langle M\rangle_\infty < \infty\}$. Que se passe-t-il sur $\{\langle M\rangle_\infty = \infty\}$? La réponse est donnée par la proposition suivante

Proposition 4.38. *(***Loi des grands nombres***) Si $f : \mathbb{R}_+ \to \mathbb{R}_+$ est une fonction croissante, telle que*

$$\int_0^\infty \frac{dt}{[1+f(t)]^2} < \infty,$$

alors $M_n/f(\langle M\rangle_n)$ converge presque sûrement vers 0 lorsque $n \to \infty$ sur $\{\langle M\rangle_\infty = \infty\}$.

Remarquons que ce résultat s'applique en particulier avec $f(t) = t$, et aussi avec $f(t) = t^\alpha$ si $\alpha > 1/2$.

Démonstration. — Posons $A_n = (1 + f(\langle M\rangle_n))^{-1}$. Considérons la martingale $N_n = (A.M)_n$. Alors

$$\begin{aligned}\langle N\rangle_n &= \sum_{k=1}^n A_k^2(\langle M\rangle_k - \langle M\rangle_{k-1})\\ &= \sum_{k=1}^n \frac{\langle M\rangle_k - \langle M\rangle_{k-1}}{(1+f(\langle M\rangle_k))^2}\\ &\leq \sum_{k\geq 1}\int_{\langle M\rangle_{k-1}}^{\langle M\rangle_k}\frac{dt}{[1+f(t)]^2}.\end{aligned}$$

La martingale N est donc de carré intégrable. Elle est donc presque sûrement convergente.

La démonstration sera donc achevée dès qu'on aura démontré le lemme suivant.
∎

Lemme 4.39. *(***Lemme de Kronecker***) Soit u_n une suite croissante de réels positifs qui converge vers l'infini, et δ_n une suite de réels. Alors, si la série $\sum_n \delta_n/u_n$ converge, la suite*

$$\frac{1}{u_n}\sum_0^n \delta_k$$

converge vers 0.

Démonstration. — (Exercice.)

∎

Corollaire 4.40. *Si la martingale satisfait à $\mathbf{E}[(M_{n+1} - M_n)^2/\mathcal{F}_n] \leq C$, alors $M_n/n^\alpha \to 0, \forall\alpha > 1/2$.*

Démonstration. — Dans ce cas, nous avons $\langle M\rangle_n \leq Cn$, et on applique le résultat précédent. ∎

Nous énonçons sans démonstration les résultats plus fins sur le comportement asymptotique de la suite M_n, et nous renvoyons à [8], p. 153, pour plus de détails.

Proposition 4.41. *(***Loi du logarithme itéré***) Si M_n est une martingale de carré intégrable telle que $\|\sup_n |M_{n+1} - M_n|\|_\infty < \infty$, alors, presque sûrement sur* *l'événement* $\{\langle M\rangle_n = \infty\}$,

$$\begin{aligned}\limsup_n \frac{M_n}{\sqrt{2\langle M\rangle_n \log\log\langle M\rangle_n}} &= 1\\ \liminf_n \frac{M_n}{\sqrt{2\langle M\rangle_n \log\log\langle M\rangle_n}} &= -1\end{aligned}$$

Enfin, la proposition 4.38 sert de loi des grands nombres pour les martingales. Le résultat suivant sert quand à lui de théorème de la limite centrale. Nous ne l'énonçons pas dans toute sa généralité mais seulement dans la forme qui nous servira plus tard pour les chaînes de Markov.

Théorème 4.42. *(***Théorème de la limite centrale***) Soit M_n une martingale de carré intégrable. Pour $n \geq 0$, posons $\Delta_n = M_{n+1} - M_n$, et $\Sigma_n = \mathbf{E}\big(\Delta_n^2/\mathcal{F}_n\big)$. Nous supposons qu'il existe une constante K telle que, pour tout n, $\mathbf{E}(|\Delta_n|^3/\mathcal{F}_n) \leq K$.*

Si la suite $(\langle M\rangle_n/n)$ converge en probabilité vers une constante σ^2, alors $M_n/\sqrt{n}$ converge en loi vers une variable gaussienne $N(0,\sigma^2)$.

Remarquons que la même propriété s'en déduit immédiatement s'il y a convergence en loi de $\langle M\rangle_n/n$ vers σ^2, puisque la convergence en loi vers une constante implique la convergence en proba.

Démonstration. — Si U_n converge en loi vers une mesure de probabilité μ, et si V_n converge en probabilité vers 0, alors $U_n + V_n$ converge en loi vers la même limite. On peut donc se ramener au cas où $M_0 = 0$.

Ensuite, quitte à remplacer M_n par $M_n/K^{3/2}$, on peut se ramener au cas où $K = 1$, et dans ce cas $\Sigma_n \leq 1$, par l'inégalité de Jensen conditionnelle.

En appliquant le théorème de P. Lévy, il nous suffit alors de démontrer que, pour tout réel t,

$$\lim_{n\to\infty} \mathbf{E}\big(\exp(itM_n/\sqrt{n} + t^2\sigma^2/2)\big) = 1.$$

Pour cela, nous allons montrer qu'il existe une constante universelle K_1 telle que, pour tout $t \in [-1,1]$, on ait

$$|\mathbf{E}\big(\exp(itM_n + t^2\langle M\rangle_n/2)\big) - 1| \leq nK_1t^3\exp(nt^2). \tag{4.2}$$

Pour le voir, posons

$$\mathbf{E}\big(\exp(itM_n + t^2\langle M\rangle_n/2)\big) = 1 + v_n,$$

et remarquons tout d'abord qu'il existe une constante universelle K_2 telle que, pour tout réel $y \in [0,1]$, tout réel $t \in [-1,1]$,

$$\exp(itx + t^2 y) = 1 + itx - t^2x^2/2 + t^2 y + R(t,x,y),$$

avec $|R(t,x,y)| \leq K_2(1+|x|^3)t^3\exp(t^2)$: écrire la formule de Taylor avec reste intégral pour la fonction $f(t) = \exp(itx + t^2y)$:

$$f(t) = 1 + tx + t^2/2(y - x^2) + \int_0^t f^{(3)}(s)(t-s)^2\,ds,$$

et de remarquer que la dérivée troisième de f, qui est égale à

$$\big((2ty + ix)^2 + 2y)\big)(ix + 2ty)f$$

est majorée par $K_2(1+|x|^3)$ sur le domaine considéré.

Ensuite, pour démontrer la majoration 4.2, nous posons

$$Y_n = \exp(itM_n + t^2\langle M\rangle_n/2),$$

et écrivons $Y_{n+1} = Y_n Z_n$, avec

$$Z_n = \exp(it\Delta_n + t^2\Sigma_n/2).$$

Commençons par calculer $\mathbf{E}(Y_{n+1}/\mathcal{F}_n)$. La variable Y_n étant mesurable par rapport à $\mathcal{F}_n$, cette dernière quantité est égale à $Y_n\mathbf{E}(Z_n/\mathcal{F}_n)$. Utilisons alors le développement limité précédent, avec $x = \Delta_n$, $y = \Sigma_n = \mathbf{E}(\Delta_n^2/\mathcal{F}_n)$, développement justifié puisque $|\Sigma_n| \leq 1$. On obtient ainsi

$$\begin{aligned}\mathbf{E}(Z_n/\mathcal{F}_n) = \quad & 1 + it\mathbf{E}(M_n/\mathcal{F}_n) - t^2/2\mathbf{E}(\Delta_n^2/\mathcal{F}_n) + t^2/2\mathbf{E}(\Sigma_n/\mathcal{F}_n) \\ & + \mathbf{E}\big(R(t,\Delta_n,\Sigma_n)/\mathcal{F}_n\big).\end{aligned}$$

Tout est fait pour que ce dévelopement se résume à

$$\mathbf{E}(Z_n/\mathcal{F}_n) = 1 + \mathbf{E}\big(R(t,\Delta_n,\Sigma_n) \mid \mathcal{F}_n\big) = 1 + V_n,$$

où la variable aléatoire $|V_n|$ est majorée par

$$K_2 t^3 \exp(t^2)\mathbf{E}(1 + |\Delta_n|^3/\mathcal{F}_n) \leq 2K_2 t^3\exp(t^2).$$

Posons alors $K_1 = 2K_2$, et établissons la majoration cherchée sur v_n par récurrence sur n. En prenant les espérances des espérances conditionnelles, nous obtenons

$$1 + v_{n+1} = \mathbf{E}\big(Y_n(1+V_n)\big) = 1 + v_n + \mathbf{E}(Y_nV_n).$$

Nous voyons d'autre part que $|Y_n| \leq \exp(nt^2/2)$, et donc, si par hypothèse $|v_n| \leq nK_1 t^3 \exp(nt^2)$, nous obtenons

$$|v_{n+1}| \leq K_1 t^3 \exp(nt^2) + K_1 t^3 \exp(t^2) \exp(nt^2/2) \leq (n+1) K_1 t^3 \exp\big(((n+1)t^2\big).$$

Cette majoration étant établie, il ne reste plus qu'à changer t et $t/\sqrt{n}$ pour obtenir

$$\lim_{n\to\infty} \mathbf{E}\Big(\exp\big(itM_n/\sqrt{n} + t^2 \langle M\rangle_n/(2n)\big)\Big) = 1.$$

Ensuite, nous écrivons

$$\begin{aligned} & \mathbf{E}\big(|\exp(itM_n - \sqrt{n} + t^2\sigma^2/2) - \exp(itM_n/\sqrt{n} + t^2/(2n)\langle M\rangle_n)|\big) \\ = \; & \mathbf{E}\big(|\exp(t^2\sigma^2/2) - \exp(t^2/(2n)\langle M\rangle_n)|\big), \end{aligned}$$

et cette dernière quantité converge vers 0, car les variables $\exp\big(t^2\langle M\rangle_n/(2n)\big)$ sont bornées, convergent en probabilité vers $\exp(t^2\sigma^2/2)$. Elles y convergent donc dans L^1.

∎

4.10 Le théorème de Radon-Nikodym.

Nous terminons ce chapitre par une application classique du théorème de convergence des martingales au théorème de Radon-Nikodym. Nous l'énonçons dans le cadre des tribus séparables (engendrées par un nombre dénombrable d'événement), mais la démonstration présentée ici reste vraie pour des tribus plus générales, par exemple les tribus engendrées aux ensembles de mesure nulle près par un nombre dénombrable d'événements, ce qui englobe une classe infiniment plus vaste.

Rappelons qu'on dit qu'une probabilité $\mathbf{Q}$ admet une densité par rapport à $\mathbf{P}$ s'il existe une variable aléatoire X, intégrable par rapport à $\mathbf{P}$, d'intégrale 1, positive, telle que, pour tout élément A de la tribu $\mathcal{A}$, on ait

$$\mathbf{Q}(A) = \int_A X \, d\mathbf{P}.$$

On la notera alors $\mathbf{Q} = X.\mathbf{P}$. Une condition nécessaire et suffisante pour que Q ait une densité par rapport à $\mathbf{P}$ est la suivante :

Théorème 4.43. *(***De Radon-Nikodym***) Soit $(\Omega, \mathcal{A}, \mathbf{P})$ un espace de probabilité, et on suppose la tribu $\mathcal{A}$ séparable. Soit $\mathbf{Q}$ une autre probabilité sur $(\Omega, \mathcal{A})$. Une CNS pour que $\mathbf{Q}$ admette une densité par rapport à $\mathbf{P}$ est que,*

$$\forall A \in \mathcal{A}, \; \mathbf{P}(A) = 0 \Longrightarrow \mathbf{Q}(A) = 0.$$

Démonstration. — La condition est évidemment nécessaire, et tout le problème est de montrer qu'elle est suffisante. L'idée de la preuve est de construire la variable aléatoire X grâce au théorème de convergence des martingales.

Dans ce qui suit, nous appellerons (A_n) une suite d'éléments qui engendrent la tribu $\mathcal{A}$, et la tribu $\sigma(A_0, \cdots, A_n)$ sera dénotée $\mathcal{F}_n$. La suite $(\mathcal{F}_{n)}$ est évidement croissante, et la tribu $\mathcal{F}_n$, engendrée par une famille finie de variables aléatoires, est en fait engendrée par une partition finie $\{B_1^n, B_2^n, \cdots, B_{p_n}^n\}$. Il est facile d'exhiber une densité de $\mathbf{Q}$ par rapport à $\mathbf{P}$ sur la tribu $\mathcal{F}_n$, c'est à dire une variable aléatoire X_n positive telle que, pour tous les B de $\mathcal{F}_n$, on ait $\mathbf{Q}(B) = \int_B X_n \, d\mathbf{P}$. Il suffit de poser

$$X_n = \sum_{i=1}^{p_n} \frac{\mathbf{Q}(B_i^n)}{\mathbf{P}(B_i^n)} \mathbf{1}_{B_i^n}.$$

A cause de l'hypothèse, si dans la somme précédente l'un des numérateurs est nul, il en va de même du dénominateur, et on n'a qu'à poser alors $0/0 = 1$. On voit que X_n joue bien le rôle d'une densité car tout élément de $\mathcal{F}_n$ est une réunion finie de B_i^n.

Bien sûr, la suite (X_n) est adaptée, positive et d'espérance 1 (sous $\mathbf{P}$) puisque c'est une densité de lois de probabilité. La remarque fondamentale dans ce qui suit est la suivante : (X_n) est une martingale sous la probabilité $\mathbf{P}$. Ceci provient d'une remarque plus générale : si $\mathcal{B}_1 \subset \mathcal{B}_2$ sont deux sous-tribus de $\mathcal{A}$, et si X_1 et X_2 sont deux variables qui jouent le rôle des densités de $\mathbf{Q}$ par rapport à $\mathbf{P}$ respectivement sur $\mathcal{B}_1$ et $\mathcal{B}_2$, alors

$$\mathbf{E_P}(X_2 \mid \mathcal{B}_1) = X_1,$$

où $\mathbf{E_P}$ désigne ici l'espérance (conditionnelle) par rapport à $\mathbf{P}$. En effet, on voit que X_1 est $\mathcal{B}_1$ mesurable par construction, et, pour tout B de $\mathcal{B}_1$, on a

$$\mathbf{E_P}(\mathbf{1}_B X_1) = \mathbf{Q}(B) = \mathbf{E_P}(\mathbf{1}_B X_2),$$

car $B \in \mathcal{B}_2$.

En appliquant ceci à la tribu $\mathcal{F}_n \subset \mathcal{F}_{n+1}$, on voit que (X_n) est une martingale sous $\mathbf{P}$. Elle est positive d'espérance 1, donc converge presque sûrement pour P. Nous allons nous attacher à montrer qu'elle est uniformément intégrable, et ceci grâce à l'hypothèse. En effet, supposons qu'elle ne le soit pas; alors, il existe $\varepsilon > 0$ et une sous-suite n_k tendant vers l'infini telle que

$$\mathbf{E_P}(X_{n_k} \mathbf{1}_{\{X_{n_k} \geq k\}}) \geq \varepsilon.$$

Ceci s'écrit encore $\mathbf{Q}(X_{n_k} \geq k) \geq \varepsilon$, d'où encore

$$\mathbf{Q}(\limsup_k(\{X_{n_k} \geq k\}) \geq \limsup \mathbf{Q}(X_{n_k} \geq k) \geq \varepsilon.$$

Si l'on appelle B l'événement $\{\limsup_n X_n = \infty\}$, nous voyons donc que $\mathbf{Q}(B) \geq \varepsilon$, alors que $\mathbf{P}(B) = 0$, d'où la contradiction.

La suite (X_n) étant sous $\mathbf{P}$ uniformément intégrable, elle converge sous $\mathbf{P}$ et dans L^1 vers une variable aléatoire X positive intégrable et d'intégrale 1. Il nous reste à monter que X est la densité cherchée de $\mathbf{Q}$ par rapport à $\mathbf{P}$. Pour ce faire, nous commençons à remarquer que, si B est dans l'une des tribus $\mathcal{F}_n$, alors

$$\mathbf{Q}(B) = \int_B X_n \, d\mathbf{P} = \int_B X \, d\mathbf{P},$$

car $X_n = \mathbf{E}_{\mathbf{P}}(X/\mathcal{F}_n)$. Les probabilités $\mathbf{Q}$ et $X.\mathbf{P}$ coïncident alors sur l'algèbre $\cup_n \mathcal{F}_n$, et donc sur la σ-algèbre $\mathcal{A}$ engendrée. C'est ce que nous voulions démontrer. ∎

References

[1] P. BARBE et M. LEDOUX, “Probabilité”, De la licence à l’agrégation, Ed. Espace 34, Belin, Montpellier, 1998.

[2] R. DURETT : “Probability : Theory and Examples”, Wadsworth and Brooks, Pacific Grove, california, 1991.

[3] W. FELLER : “An Introduction to Probability Theory and its Applications” I and II, John Wiley, New York, 1966.

[4] G. GRIMMETT, D. STIRZAKER, “Probability and Random Processes”, third edition, Oxford University Press, Oxford, 2001.

[5] P. JAFFARD : “Initiation aux méthodes de la statistique et du calcul des probabilités”, Masson, Paris, 1978 (réédité en 1990).

[6] I. KARATZAS, S.E.SHREVE : “Brownian motion and Stochastic Calculus”,Springer, Berlin, 1999.

[7] L. MAZLIAK, P. PRIOURET et P. BALDI : “Martingales et chaînes de Markov”, Hermann, Paris, 1998.

[8] J. NEVEU : “Martingales à temps discret”, Masson, Paris, 1972.

[9] J. Norris : “Markov Chains”, Cambridge University Press, Cambridge, 1997.

[10] P. PROTTER : “Stochastic Integration and Differential Equations”, Springer, Berlin, 1990.

[11] P. S. TOULOUSE, “Thèmes de Probabilités et Statistique, Agrégation de mathématiques, Dunod, Paris, 1999.