

ESPÉRANCE CONDITIONNELLE MARTINGALES

Préparation à l'Agrégation Bordeaux 1

Année 2013 - 2014

Jean-Jacques Ruch et Marie-Line Chabanol

Table des Matières

Chapitre I. Espérance conditionnelle	5
1. Introduction	5
2. Définition de l'espérance conditionnelle	7
2.A. Cas particulier des variables aléatoires de carré intégrable	7
2.B. Cas général	7
3. Propriétés de l'espérance conditionnelle analogues à celles de l'espérance	9
4. Propriétés spécifiques à l'espérance conditionnelle	10
5. Calculs d'espérance conditionnelle, loi conditionnelle	10
Chapitre II. Martingales	13
1. Introduction	13
2. Définition des martingales	14
3. Surmartingale et sous-martingale	15
4. Temps d'arrêt	16
5. Propriétés des martingales par rapport aux temps d'arrêts	17
6. Théorèmes d'arrêt	18
7. Inégalités maximales	19
8. Convergence des martingales	21
9. Convergence des martingales L^2	24
10. Complément : convergence dans L^1	25

CHAPITRE I

Espérance conditionnelle

1. Introduction

Pour de nombreux problèmes concrets (prédiction, observation incomplète, etc.) il est important de pouvoir estimer une variable aléatoire sur laquelle on n'a qu'une information partielle. Dès lors, on comprend l'importance de la notion d'espérance conditionnelle. La définition axiomatique de cette notion est motivée par le cas discret traité dans le premier paragraphe. Le calcul explicite des espérances conditionnelles, qui est en général un problème difficile, est illustré sur plusieurs cas, dont le cas gaussien particulièrement important pour les applications.

On note $(\Omega, \mathcal{A}, \mathbb{P})$ un espace probabilisé. Soit $B \in \mathcal{A}$ un événement tel que $\mathbb{P}(B) > 0$. On peut définir une nouvelle probabilité sur (Ω, \mathcal{A}) , appelée *probabilité conditionnelle* sachant B , en posant pour tout $A \in \mathcal{A}$,

$$\mathbb{P}(A|B) = \frac{\mathbb{P}(A \cap B)}{\mathbb{P}(B)} = \frac{\mathbb{E}[\mathbf{1}_A \mathbf{1}_B]}{\mathbb{P}(B)}$$

De même, pour toute variable aléatoire X positive ou dans $L^1(\Omega, \mathcal{A}, \mathbb{P})$, l'*espérance conditionnelle* de X sachant B est définie par

$$\mathbb{E}[X|B] = \frac{\mathbb{E}[X \mathbf{1}_B]}{\mathbb{P}(B)}.$$

Cette quantité est aussi l'espérance de X sous la probabilité $\mathbb{P}(\cdot|B)$, et elle s'interprète comme la valeur moyenne de X quand B est réalisé.

En particulier, si X est discrète, on obtient (de façon assez conforme à l'intuition ?)

$$\mathbb{E}[X|B] = \sum_k x_k \mathbb{P}(X = x_k|B)$$

Considérons maintenant une variable aléatoire Y à valeurs dans un espace E dénombrable et soit $y \in E$ tel que $\mathbb{P}(Y = y) > 0$. Pour toute variable aléatoire $X \in L^1(\Omega, \mathcal{A}, \mathbb{P})$ on peut définir, comme un cas particulier de ce qui précède,

$$\mathbb{E}[X|Y = y] = \frac{\mathbb{E}[X \mathbf{1}_{Y=y}]}{\mathbb{P}(Y = y)}.$$

Définition 1. Soit $X \in L^1(\Omega, \mathcal{A}, \mathbb{P})$ et Y une variable aléatoire discrète sur Ω . L'espérance conditionnelle de X sachant Y est la variable aléatoire réelle définie par

$$\mathbb{E}[X|Y] = \varphi(Y),$$

où la fonction $\varphi : E \rightarrow \mathbb{R}$ est donnée par

$$\varphi(y) = \begin{cases} \mathbb{E}[X|Y = y] & \text{si } y \text{ est tel que } \mathbb{P}(Y = y) > 0 \\ 0 & \text{sinon} \end{cases}$$

En particulier si X est également discrète, $\mathbb{E}[X|Y = y] = \sum_k x_k \mathbb{P}(X = x_k|Y = y)$.

Le choix de la valeur de φ lorsque $\mathbb{P}(Y = y) = 0$ n'a pas d'importance, puisque c'est un ensemble de probabilité nulle. En effet si on note $E' = \{y \in E \mid \mathbb{P}(Y = y) = 0\}$ alors

$$\mathbb{P}(Y \in E') = \sum_{y \in E'} \mathbb{P}(Y = y) = 0.$$

Donc, si on changeait la définition de φ sur E' cela donnerait la même variable aléatoire $\mathbb{E}[X|Y]$ à un ensemble de mesure nulle près.

Dans le cas général, l'espérance conditionnelle sera toujours définie à un ensemble de probabilité nulle près. En comparant avec le conditionnement par rapport à un événement, on observe que l'espérance conditionnelle $\mathbb{E}[X|Y]$ est maintenant une variable aléatoire : c'est la variable aléatoire qui donne la valeur moyenne de X quand on connaît Y : on a presque sûrement

$$\mathbb{E}[X|Y](\omega) = \mathbb{E}[X|Y = y], \text{ si } Y(\omega) = y.$$

On a donc aussi les écritures suivantes, avec plus ou moins d'abus de langage... :

Lorsque Y est une variable discrète à valeurs dans E

$$\mathbb{E}[X|Y](\omega) = \sum_{y \in E} 1_{Y^{-1}(\{y\})}(\omega) \mathbb{E}[X|Y = y] = \sum_{y \in E} 1_{Y(\omega)=y} \mathbb{E}[X|Y = y] = \mathbb{E}[X|Y = Y(\omega)].$$

Remarquons que $\mathbb{E}[X|Y]$ est une fonction de Y (puisque'elle ne dépend de ω que par la valeur $Y(\omega)$) donc une variable aléatoire $\sigma(Y)$ -mesurable. Dans un sens qui sera précisé plus loin, c'est la meilleure approximation de X par une fonction de Y .

Exemple : Lancer d'un dé. On prend $\Omega = \{1, 2, \dots, 6\}$ et $\mathbb{P}(\{\omega\}) = 1/6$ pour tout ω . Soient

$$Y(\omega) = \begin{cases} 1 & \text{si } \omega \text{ est impair} \\ 0 & \text{si } \omega \text{ est pair} \end{cases}$$

et $X(\omega) = \omega$. Alors,

$$\mathbb{E}[X|Y](\omega) = \begin{cases} 3 & \text{si } \omega \in \{1, 3, 5\} \\ 4 & \text{si } \omega \in \{2, 4, 6\} \end{cases}.$$

Ou encore

$$\mathbb{E}[X|Y] = 3 1_{Y=1} + 4 1_{Y=0}$$

Proposition 2. Soit $X \in L^1(\Omega, \mathcal{A}, \mathbb{P})$. On a

$$\mathbb{E}[|\mathbb{E}[X|Y]|] \leq \mathbb{E}[|X|];$$

en particulier $\mathbb{E}[X|Y] \in L^1(\Omega, \mathcal{A}, \mathbb{P})$.

De plus pour toute variable aléatoire Z bornée et $\sigma(Y)$ -mesurable

$$\mathbb{E}[Z\mathbb{E}[X|Y]] = \mathbb{E}[ZX].$$

en particulier $\mathbb{E}[\mathbb{E}[X|Y]] = \mathbb{E}[X]$.

DÉMONSTRATION. D'après la définition de l'espérance conditionnelle $\mathbb{E}[X|Y]$, on a

$$\begin{aligned} \mathbb{E}[|\mathbb{E}[X|Y]|] &= \sum_{y \in E \setminus E'} \mathbb{P}(Y = y) |\mathbb{E}[X|Y = y]| \\ &= \sum_{y \in E \setminus E'} \mathbb{P}(Y = y) \left| \frac{\mathbb{E}[X 1_{Y=y}]}{\mathbb{P}(Y = y)} \right| \leq \sum_{y \in E} \mathbb{E}[|X| 1_{Y=y}] = \mathbb{E}[|X|]. \end{aligned}$$

Pour la deuxième assertion, on utilise le fait qu'on peut écrire $Z = \psi(Y)$, avec ψ une fonction bornée. Alors,

$$\begin{aligned} \mathbb{E}[\psi(Y)\mathbb{E}[X|Y]] &= \sum_{y \in E \setminus E'} \mathbb{P}(Y = y) \psi(y) \mathbb{E}[X|Y = y] = \sum_{y \in E \setminus E'} \mathbb{P}(Y = y) \psi(y) \frac{\mathbb{E}[X 1_{Y=y}]}{\mathbb{P}(Y = y)} \\ &= \sum_{y \in E \setminus E'} \psi(y) \mathbb{E}[X 1_{Y=y}] = \sum_{y \in E \setminus E'} \mathbb{E}[\psi(y) X 1_{Y=y}] \\ &= \sum_{y \in E \setminus E'} \mathbb{E}[\psi(Y) X 1_{Y=y}] = \mathbb{E} \left[\sum_{y \in E \setminus E'} \psi(Y) 1_{Y=y} X \right] = \mathbb{E}[\psi(Y) X] \end{aligned}$$

□

Enfin, on peut vérifier que si Y' est une autre variable aléatoire discrète telle que $\sigma(Y) = \sigma(Y')$, on a $\mathbb{E}[X|Y] = \mathbb{E}[X|Y']$ p.s.

Ceci suggère que la bonne notion de conditionnement est la notion de conditionnement par rapport à une tribu. C'est cette notion que nous allons développer dans la suite.

2. Définition de l'espérance conditionnelle

2.A. Cas particulier des variables aléatoires de carré intégrable. Si Y est discrète, on peut vérifier que la définition précédente entraîne que si X est de carré intégrable, alors $E[X|Y]$ aussi ; de plus la proposition précédente entraîne alors que $(E[X|Y] - X)$ est orthogonal (au sens L^2) à toute variable aléatoire Z bornée $\sigma(Y)$ mesurable.

Cela suggère une généralisation de la définition lorsque Y n'est pas forcément discrète en terme de projection orthogonale.

Avant d'énoncer le résultat, rappelons que si \mathcal{B} est une sous-tribu de \mathcal{A} alors $L^2(\Omega, \mathcal{B}, \mathbb{P})$ s'identifie à un sous-espace fermé de $L^2(\Omega, \mathcal{A}, \mathbb{P})$, à savoir l'espace des éléments de $L^2(\Omega, \mathcal{A}, \mathbb{P})$ dont un représentant au moins est \mathcal{B} -mesurable.

Définition 3. Si $X \in L^2(\Omega, \mathcal{A}, \mathbb{P})$ et si \mathcal{B} est une sous tribu de \mathcal{A} alors $\mathbb{E}[X|\mathcal{B}]$ est la projection orthogonale de X sur $L^2(\Omega, \mathcal{B}, \mathbb{P})$. En particulier $\mathbb{E}[X|\mathcal{B}] \in L^2(\Omega, \mathcal{B}, \mathbb{P})$.
Si Y est une variable aléatoire, on note $\mathbb{E}[X|Y] = \mathbb{E}[X|\sigma(Y)]$: c'est la variable aléatoire fonction de Y qui approche le mieux X au sens L^2 .

Remarque : si $\mathcal{B} = \{\emptyset, \Omega\}$, $L^2(\Omega, \mathcal{B}, \mathbb{P}) = \{v.a.\text{ constantes}\}$, et on a alors $\mathbb{E}[X|\mathcal{B}] = \mathbb{E}[X]$.

Si On en déduit immédiatement la proposition suivante :

Proposition 4.

- Si $X \in L^2(\Omega, \mathcal{A}, \mathbb{P})$ alors

$$\mathbb{E}[(X - \mathbb{E}[X|\mathcal{B}])^2] = \inf_{Z \in L^2(\Omega, \mathcal{B}, \mathbb{P})} \mathbb{E}[(X - Z)^2].$$

- On a pour toute variable aléatoire $Z \in L^2(\Omega, \mathcal{B}, \mathbb{P})$

$$\mathbb{E}[Z\mathbb{E}[X|\mathcal{B}]] = \mathbb{E}[ZX]$$

En particulier, pour toute fonction mesurable ψ telle que $\psi(Y)$ est de carré intégrable,

$$\mathbb{E}[\psi(Y)\mathbb{E}[X|Y]] = \mathbb{E}[\psi(Y)X]$$

- L'espérance conditionnelle $\mathbb{E}[X|\mathcal{B}]$ est caractérisée parmi les variables de $L^2(\Omega, \mathcal{B}, \mathbb{P})$ par $\forall B \in \mathcal{B}, \mathbb{E}[1_B \mathbb{E}[X|\mathcal{B}]] = \mathbb{E}[1_B X]$

Ces propriétés suggèrent la définition dans le cas L^1 .

2.B. Cas général.

Théorème 5.

Soit \mathcal{B} une sous-tribu de \mathcal{A} , et soit une variable aléatoire $X \in L^1(\Omega, \mathcal{A}, \mathbb{P})$. Il existe alors une unique variable aléatoire dans $L^1(\Omega, \mathcal{B}, \mathbb{P})$, \mathcal{B} mesurable, notée $\mathbb{E}[X|\mathcal{B}]$, telle que

$$\forall B \in \mathcal{B}, \quad \mathbb{E}[X1_B] = \mathbb{E}[\mathbb{E}[X|\mathcal{B}]1_B].$$

De manière équivalente on a, pour toute variable aléatoire, Z , \mathcal{B} -mesurable et bornée

$$\mathbb{E}[XZ] = \mathbb{E}[\mathbb{E}[X|\mathcal{B}]Z].$$

En particulier, si $Z = 1$ on a

$$\mathbb{E}[\mathbb{E}[X|\mathcal{B}]] = \mathbb{E}[X].$$

L'espérance conditionnelle par rapport une tribu est caractérisée par l'une des deux propriétés ci-dessus. L'équivalence entre les deux points est assez facile à voir. La premier point nous dit que l'on a le résultat pour toutes les indicatrices \mathcal{B} -mesurables. Donc par somme et passage à la limite, il est encore vrai pour les fonctions étagées puis pour les fonctions bornées et \mathcal{B} -mesurables. Pour la réciproque il suffit de poser $Z = 1_B$.

Ainsi, par exemple, pour démontrer un résultat du style $\mathbb{E}[X|Y] = U$, une méthode est :

- On commence par vérifier que U est L^1 et est fonction de Y
- Puis on montre $\forall \psi$ mesurable bornée, $\mathbb{E}[X\psi(Y)] = \mathbb{E}[XU]$

DÉMONSTRATION. 1. Existence

On a vu plus haut le cas où $X \in L^2(\Omega, \mathcal{A}, \mathbb{P})$: dans ce cas $\mathbb{E}[X|\mathcal{B}]$ est la projection orthogonale de X sur $L^2(\Omega, \mathcal{B}, \mathbb{P})$, et on a bien l'existence.

Passons maintenant au cas général, c'est-à-dire $X \in L^1(\Omega, \mathcal{A}, \mathbb{P})$. En posant classiquement $X = X^+ - X^-$, il est clair que l'on peut se ramener au cas où $X \geq 0$. Soit pour $n \in \mathbb{N}$, $X_n = X \wedge n$. D'après ce qui précède on peut prendre son espérance conditionnelle par rapport à \mathcal{B} , $Y_n = \mathbb{E}[X_n|\mathcal{B}]$. D'autre part, X_n tend simplement en croissant vers X . Pour les Y_n remarquons que $0 \leq Y_n \leq Y_{n+1}$ presque sûrement. Il suffit pour cela de vérifier que si $U \geq 0$ alors son espérance conditionnelle vérifie $V = \mathbb{E}[U|\mathcal{B}] \geq 0$. En effet, par l'absurde : si $\mathbb{P}(V < 0) > 0$ alors il existe $\varepsilon > 0$ tel que $\mathbb{P}(V < -\varepsilon) > 0$. Or comme $\{V < -\varepsilon\} \in \mathcal{B}$ on a

$$0 \leq \mathbb{E}[U1_{V < -\varepsilon}] = \mathbb{E}[V1_{V < -\varepsilon}] \leq -\varepsilon$$

ce qui est impossible.

Posons $Y = \limsup Y_n$ qui est \mathcal{B} mesurable. Pour tout $B \in \mathcal{B}$ on a :

$$\begin{aligned} \mathbb{E}[Y1_B] &= \lim_{n \rightarrow +\infty} \mathbb{E}[Y_n1_B] \quad \text{par convergence monotone} \\ &= \lim_{n \rightarrow +\infty} \mathbb{E}[X_n1_B] \\ &= \mathbb{E}[X1_B] \quad \text{par convergence monotone.} \end{aligned}$$

2. Unicité. Soient Y et Y' deux variables aléatoires dans $L^1(\Omega, \mathcal{A}, \mathbb{P})$, \mathcal{B} -mesurables, et telles que

$$\forall B \in \mathcal{B}, \quad \mathbb{E}[Y1_B] = \mathbb{E}[X1_B] = \mathbb{E}[Y'1_B].$$

Comme Y et Y' sont \mathcal{B} -mesurables, $B_1 = \{Y > Y'\}$ et $B_2 = \{Y' > Y\}$ le sont aussi. D'où on obtient

$$\mathbb{E}[Y1_{B_1}] = \mathbb{E}[Y'1_{B_1}] \quad \text{et} \quad \mathbb{E}[Y1_{B_2}] = \mathbb{E}[Y'1_{B_2}]$$

Donc, on en déduit $(Y - Y')1_{B_1} = 0$ p.s. et $(Y' - Y)1_{B_2} = 0$ p.s. ce qui entraîne que $Y = Y'$ p.s. \square

3. Propriétés de l'espérance conditionnelle analogues à celles de l'espérance

Soit X une variable aléatoire dans $L^1(\Omega, \mathcal{A}, \mathbb{P})$. On note \mathcal{B} une sous-tribu de \mathcal{A} .

– **a)** Pour tous réels a et b et toute variable aléatoire réelle X intégrable,

$$\mathbb{E}[aX + b|\mathcal{B}] = a\mathbb{E}[X|\mathcal{B}] + b;$$

et pour toutes variables aléatoires réelles X_1, X_2 intégrables

$$\mathbb{E}[X_1 + X_2|\mathcal{B}] = \mathbb{E}[X_1|\mathcal{B}] + \mathbb{E}[X_2|\mathcal{B}]$$

– **b)** Si $X_1 \leq X_2$ p.s. alors $\mathbb{E}[X_1|\mathcal{B}] \leq \mathbb{E}[X_2|\mathcal{B}]$.

DÉMONSTRATION. Le point **a)** est une conséquence de l'unicité de l'espérance conditionnelle.

Pour le dernier point, on commence par montrer que si $X \geq 0$ alors $\mathbb{E}[X|\mathcal{B}] \geq 0$. C'est un corollaire de la preuve précédente : si $Y = \mathbb{E}[X|\mathcal{B}]$ vérifie $P(Y < 0) > 0$, alors $\mathbb{E}[Y1_{Y<0}] < 0$; mais $1_{Y<0}$ est \mathcal{B} mesurable, donc cette quantité est aussi égale à $\mathbb{E}[X1_{Y<0}]$ qui est positive, d'où la contradiction. \square

– **c)** Si X et X_n sont des variables aléatoires positives dans $L^1(\Omega, \mathcal{A}, \mathbb{P})$ alors

$$X_n \uparrow X \Rightarrow \mathbb{E}[X_n|\mathcal{B}] \uparrow \mathbb{E}[X|\mathcal{B}].$$

– **d)** Si X_n sont des variables aléatoires positives, alors

$$\mathbb{E}[\liminf X_n|\mathcal{B}] \leq \liminf \mathbb{E}[X_n|\mathcal{B}]$$

– **e)** Si $X_n \rightarrow X$ p.s. avec pour tout n , $|X_n| \leq Z \in L^1(\Omega, \mathcal{A}, \mathbb{P})$, alors

$$\lim \mathbb{E}[X_n|\mathcal{B}] = \mathbb{E}[X|\mathcal{B}].$$

– **f)** Soit f une fonction continue et convexe et X une variable aléatoire réelle telle que X et $f(X)$ sont intégrables, alors

$$f(\mathbb{E}[X|\mathcal{B}]) \leq \mathbb{E}[f(X)|\mathcal{B}];$$

– **g)** En particulier $|\mathbb{E}[X|\mathcal{B}]| \leq \mathbb{E}[|X||\mathcal{B}]$, et par conséquent $\mathbb{E}[|\mathbb{E}[X|\mathcal{B}]|] \leq \mathbb{E}[|X|]$.

DÉMONSTRATION. Pour le point **c)** : on pose $Y = \lim \mathbb{E}[X_n|\mathcal{B}] = \limsup \mathbb{E}[X_n|\mathcal{B}]$ (d'après la croissance) qui est \mathcal{B} -mesurable. On a pour tout $B \in \mathcal{B}$

$$\begin{aligned} \mathbb{E}[Y1_B] &= \lim \mathbb{E}[\mathbb{E}[X_n|\mathcal{B}]1_B] \quad \text{par convergence monotone} \\ &= \lim \mathbb{E}[X_n1_B] \\ &= \mathbb{E}[X1_B] \quad \text{par convergence monotone} \end{aligned}$$

Pour le point **d)** : On a d'après le résultat précédent

$$\begin{aligned} \mathbb{E}[\liminf X_n|\mathcal{B}] &= \lim_{n \rightarrow +\infty} \mathbb{E}[\inf_{k \geq n} X_k|\mathcal{B}] \\ &\leq \lim_{n \rightarrow +\infty} \inf_{k \geq n} \mathbb{E}[X_k|\mathcal{B}] = \liminf \mathbb{E}[X_n|\mathcal{B}]. \end{aligned}$$

Pour les derniers points (convergence dominée conditionnelle et inégalité de Jensen conditionnelle) il suffit de reprendre les démonstrations faites dans le cas de l'espérance classique. \square

4. Propriétés spécifiques à l'espérance conditionnelle

- a) Si X est intégrable alors $\mathbb{E}[X|\mathcal{B}]$ l'est aussi et $\mathbb{E}[\mathbb{E}[X|\mathcal{B}]] = \mathbb{E}[X]$.
- b) Si X est une variable aléatoire réelle \mathcal{B} -mesurable alors

$$\mathbb{E}[X|\mathcal{B}] = X \text{ p.s.}$$

en particulier $\mathbb{E}[1|\mathcal{B}] = 1$.

Donc si ψ est une fonction mesurable telle que $\psi(Y)$ est intégrable,

$$\mathbb{E}[\psi(Y)|Y] = \psi(Y).$$

- c) Soient X et Z deux variables aléatoires réelles intégrables telles que XZ soit aussi intégrable. Supposons Z \mathcal{B} -mesurable alors

$$\mathbb{E}[XZ|\mathcal{B}] = Z\mathbb{E}[X|\mathcal{B}] \text{ p.s.};$$

En particulier si ψ est une fonction mesurable telle que $\psi(Y)$ et $X\psi(Y)$ soient intégrables,

$$\mathbb{E}[X\psi(Y)|Y] = \psi(Y)\mathbb{E}[X|Y]$$

- d) Si \mathcal{B}_1 et \mathcal{B}_2 sont deux sous-tribus de \mathcal{A} telles que $\mathcal{B}_1 \subset \mathcal{B}_2$ alors

$$\mathbb{E}[\mathbb{E}[X|\mathcal{B}_1]|\mathcal{B}_2] = \mathbb{E}[X|\mathcal{B}_1] \quad \text{et} \quad \mathbb{E}[\mathbb{E}[X|\mathcal{B}_2]|\mathcal{B}_1] = \mathbb{E}[X|\mathcal{B}_1].$$

- e) Soit \mathcal{B} une sous-tribus de \mathcal{A} . Si X est une variable aléatoire dans $L^1(\Omega, \mathcal{A}, \mathbb{P})$, telle que $\sigma(X)$ et \mathcal{B} sont indépendantes, alors

$$\mathbb{E}[X|\mathcal{B}] = \mathbb{E}[X];$$

en particulier si X et Y sont deux variables aléatoires réelles indépendantes telles que $X \in L^1(\Omega, \mathcal{A}, \mathbb{P})$ alors

$$\mathbb{E}[X|Y] = \mathbb{E}[X]$$

la réciproque de ce dernier point étant fausse.

DÉMONSTRATION. Les points a) et b) se déduisent de la définition.

Lorsque Z est bornée, le point c) se déduit de la définition de l'espérance conditionnelle et de son unicité. Puis on applique la machine standard, qui montre que la propriété est vérifiée si Z est une indicatrice, ensuite une fonction en escalier positive, puis une fonction positive \mathcal{B} -mesurable et enfin une fonction, Z , \mathcal{B} -mesurable telle que XZ soit intégrable.

Le point suivant est laissé en exercice. Enfin pour le dernier point on a pour tout $B \in \mathcal{B}$:

$$\mathbb{E}[\mathbb{E}[X|\mathcal{B}]1_B] = \mathbb{E}[X1_B] = \mathbb{E}[X]\mathbb{E}[1_B] = \mathbb{E}[\mathbb{E}[X]1_B]$$

d'où par unicité de l'espérance conditionnelle on obtient le résultat. \square

5. Calculs d'espérance conditionnelle, loi conditionnelle

- $\mathcal{B} = \{\emptyset, \Omega\}$ est la tribu triviale.

Une variable aléatoire réelle est \mathcal{B} -mesurable si elle est constante sur Ω . Dans ce cas, $\mathbb{E}[X|\mathcal{B}]$ est l'unique variable aléatoire réelle constante sur Ω et égale à $\mathbb{E}[X]$.

- \mathcal{B} est engendré par une partition finie ou dénombrable de Ω (cela correspond au cas où $\mathcal{B} = \sigma(Y)$ lorsque Y est une variable aléatoire discrète)

On note $\{A_i, i \in J\}$, avec J dénombrable, une partition de Ω qui engendre \mathcal{B} . On définit L comme étant l'ensemble des indices dans J tels que $\mathbb{P}(A_i) > 0$. Alors on a

$$\mathbb{E}[X|\mathcal{B}] = \sum_{i \in L} \frac{\mathbb{E}[1_{A_i} X]}{\mathbb{P}(A_i)} 1_{A_i}.$$

On retrouve bien la formule donnée au début du chapitre.

- En particulier si X et Y sont toutes les deux discrètes, avec $\mathbb{P}(X = x_k, Y = y_l) = p_{kl}$ et $\mathbb{P}(Y = y_l) = q_l$, on a

$$\mathbb{P}(X = x_k | Y = y_l) = \frac{p_{kl}}{q_l}$$

Si l est fixée, la loi de probabilité sur les (x_k) donnée par $\mu(x_k) = \frac{p_{kl}}{q_l}$ est appelée “loi conditionnelle de X sachant $Y = y_l$ ”. On a ainsi une famille de lois de probabilité indexées par les y_l .

En particulier, si ψ est une fonction bornée,

$$\mathbb{E}[\psi(X)|Y = y_l] = \sum_k \psi(x_k) \frac{p_{kl}}{q_l}$$

et

$$\mathbb{E}[\psi(X)|Y] = \sum_l 1_{Y=y_l} \mathbb{E}[\psi(X)|Y = y_l]$$

On a alors

$$\mathbb{E}[\psi(X)] = \sum_l P(Y = y_l) \mathbb{E}[\psi(X)|Y = y_l]$$

- Cas particulier d’une somme aléatoire de variables aléatoires.

Soit $(X_i)_{i \geq 1}$ une suite de variables aléatoires indépendantes intégrables identiquement distribuées.

Soit N une variable aléatoire à valeurs dans \mathbb{N}^* indépendante des X_i . On pose $S_N = \sum_{i=1}^N X_i$. Alors

$E[S_N|N] = NE[X_1]$. En effet on a pour tout entier k

$$\mathbb{E}[S_N|N = k] = \frac{\mathbb{E}[S_N 1_{N=k}]}{\mathbb{P}(N=k)} = \frac{\mathbb{E}[(\sum_{i=1}^k X_i) 1_{N=k}]}{\mathbb{P}(N=k)} = k\mathbb{E}[X_1].$$

On en déduit entre autres $E[S_N] = \mathbb{E}[N]\mathbb{E}[X_1]$. On peut calculer de manière analogue la variance de S_N .

En particulier, si les X_i sont des variables de Bernoulli de paramètre p , on dira que la loi conditionnelle de S_N sachant N est une loi binomiale de paramètres (N, p) .

- \mathcal{B} est la tribu engendré par une variable aléatoire à densité à valeurs dans \mathbb{R}^d .

Pour simplifier les écritures on suppose que X et Y sont des variables aléatoires à densité à valeurs dans \mathbb{R} , de couple à densité. On note $f_{(X,Y)}(x, y)$ la densité jointe et les densités marginales $f_X(x) = \int f_{(X,Y)}(x, y) dy$ et $f_Y(y) = \int f_{(X,Y)}(x, y) dx$. On pose alors

$$f_{X|Y}(x|y) = \frac{f_{(X,Y)}(x, y)}{f_Y(y)} 1_{f_Y(y) \neq 0}.$$

On a alors :

$$\mathbb{E}[X|Y](\omega) = \varphi(Y)(\omega) = \int_{\mathbb{R}} x f_{X|Y}(x|Y(\omega)) dx.$$

$f_{X|Y}(x|y)$ est appelée “densité conditionnelle de X sachant $Y = y$ ” : si $f_Y(y) \neq 0$, $x \mapsto f_{X|Y}(x|y)$ est bien une densité. Ainsi on a une famille de densités (indexées par y). Cette famille de densités permet en fait de définir la “loi conditionnelle de X sachant Y ” : en effet on a bien, si ψ est une fonction borélienne bornée, l’égalité presque sûre :

$$\mathbb{E}[\psi(X)|Y] = \int_{\mathbb{R}} \psi(x) f_{X|Y}(x|Y) dx.$$

Réciproquement, si on connaît la densité de y et la loi conditionnelle de X sachant $Y = y$, on retrouve la densité du couple

$$f_{(X,Y)}(x, y) = f_{X|Y}(x|y) f_Y(y)$$

Remarque : on peut vérifier que $P(\{\omega, f_Y(Y(\omega)) = 0\}) = \int_{f_Y^{-1}(0)} f_Y(y) dy = 0 \dots$

Exemple 1

Si (X, Y) est uniforme sur le disque $D(0, 1)$, la loi de X sachant Y est uniforme sur $[-\sqrt{1 - Y^2}, \sqrt{1 - Y^2}]$.

En effet, si $y \in [-1, 1]$,

$$f_{X|Y}(x|y) = \frac{1_D(x, y)}{f_Y(y)} = \frac{1_{[-\sqrt{1-y^2}, \sqrt{1-y^2}]}(x)}{f_Y(y)}.$$

Exemple 2

Si (X, Y) a pour densité $2 \exp(-(x+y)) 1_{0 \leq x \leq y}$, on a $f_Y(y) = 2 \exp(-y)(1 - \exp(-y)) 1_{y > 0}$, donc $\forall y > 0$, la loi de X sachant $Y = y$ a pour densité $2 \frac{\exp(-x)}{1 - \exp(-y)} 1_{[0, y]}(x)$.

- Exemple important des vecteurs gaussiens

Soit (X_1, X_2) un vecteur gaussien de loi $N(M, \Gamma)$. Alors il existe des constantes réelles a et b telles que $X_1 - bX_2$ est indépendant de X_2 et

$$\mathbb{E}[X_1|X_2] = a + bX_2.$$

En effet il suffit de choisir b tels que $\text{Cov}(X_1 - bX_2, X_2) = 0$, c'est-à-dire $\text{Cov}(X_1, X_2) - b\text{Var}(X_2) = 0$. Puis $\mathbb{E}[X_1|X_2] = \mathbb{E}[X_1 - bX_2 + bX_2|X_2] = \mathbb{E}[X_1 - bX_2] + bX_2$.

Ainsi l'espérance conditionnelle d'une composante d'un vecteur gaussien par rapport à une autre composante est une fonction linéaire de cette composante. Elle suit donc à nouveau une loi normale...

CHAPITRE II

Martingales

Biblio : Cottrell pour des rappels de cours rapides ; Foata-Fuchs (Processus stochastiques) et Bercu

1. Introduction

Définition 1. *Sous le nom de processus aléatoire ou processus stochastique on entend un modèle permettant d'étudier un phénomène aléatoire évoluant au cours du temps. Pour le décrire, on se donne :*

- 1) *un espace probabilisé $(\Omega, \mathcal{A}, \mathbb{P})$;*
- 2) *un espace mesurable (E, \mathcal{B}) , où E est appelé espace des états du processus ;*
- 3) *une famille $(X_t)_{t \in T}$ de variables aléatoires définies sur $(\Omega, \mathcal{A}, \mathbb{P})$ et à valeurs dans E .*

L'ensemble T des indices t est l'espace des temps.

Etant donné $\omega \in \Omega$, on appelle trajectoire du processus, l'application $t \mapsto X_t(\omega)$.

Pour $\omega \in \Omega$ et $t \in T$, la quantité $X_t(\omega)$ est appelée état du processus à l'instant t .

Lorsque T est discret on représente une trajectoire par une suite de points dans le plan.

On distingue plusieurs type de processus :

- les *processus à temps discret* et les *processus à temps continu* si $T \subset \mathbb{N}$ et respectivement si $T = [0, 1]$ ou \mathbb{R} ;
- les *processus à espace d'états fini*, à *espace d'états dénombrable* ou à *espace d'états continu*, si E est fini, dénombrable ou respectivement continu.

Dans la suite sauf exception on ne considérera que des processus à temps discret c'est-à-dire une suite X_0, X_1, X_2, \dots de variables aléatoires sur $(\Omega, \mathcal{A}, \mathbb{P})$ à valeurs dans (E, \mathcal{B}) . Si E est fini ou dénombrable ces variables aléatoires sont forcément discrètes.

Définition 2. *Une filtration de $(\Omega, \mathcal{A}, \mathbb{P})$ est une suite croissante $(\mathcal{F}_n)_{n \in \mathbb{N}}$ de sous-tribus de \mathcal{A} . On dit alors que $(\Omega, \mathcal{A}, (\mathcal{F}_n)_{n \in \mathbb{N}}, \mathbb{P})$ est un espace probabilisé filtré.*

On définit souvent pour un filtration la tribu \mathcal{F}_{-1} par $\mathcal{F}_{-1} = \{\emptyset, \Omega\}$.

Définition 3. *Un processus $(X_n)_{n \in \mathbb{N}}$ est adapté à la filtration $(\mathcal{F}_n)_{n \in \mathbb{N}}$ si pour tout $n \in \mathbb{N}$, X_n est mesurable par rapport à la tribu \mathcal{F}_n .*

Soit $(X_n)_{n \in \mathbb{N}}$ un processus sur $(\Omega, \mathcal{A}, \mathbb{P})$. On définit \mathcal{F}_n^X comme étant la plus petite tribu rendant les variables aléatoires X_0, \dots, X_n mesurables :

$$\mathcal{F}_n^X = \sigma(X_0, \dots, X_n).$$

Alors $(\mathcal{F}_n^X)_{n \in \mathbb{N}}$ est une filtration appelée filtration canonique du processus aléatoire $(X_n)_{n \in \mathbb{N}}$.

D'après la définition, si $(X_n)_{n \in \mathbb{N}}$ est adapté à la filtration $(\mathcal{F}_n)_{n \in \mathbb{N}}$, alors X_n mesurable par rapport à la tribu \mathcal{F}_m pour $m \geq n$.

La filtration canonique est par construction la plus petite filtration qui rende le processus adapté.

Donnons deux exemples de processus, que nous étudierons plus en détails.

Un processus aléatoire $(X_n)_{n \in \mathbb{N}}$ adapté à la filtration $(\mathcal{F}_n)_{n \in \mathbb{N}}$ tel que chaque X_n est intégrable et

$$\mathbb{E}(X_{n+1}|\mathcal{F}_n) = X_n \text{ p.s.}$$

est appelé une *martingale*.

Un processus aléatoire $(X_n)_{n \in \mathbb{N}}$ à valeurs dans (E, \mathcal{B}) est une *chaîne de Markov* si pour tout $n \in \mathbb{N}$ et tout $B \in \mathcal{B}$ on a

$$\mathbb{P}(X_{n+1} \in B | X_0, \dots, X_n) = \mathbb{P}(X_{n+1} \in B | X_n).$$

Cela signifie que si l'on connaît la position X_n du processus à l'instant n et si on veut prédire sa position X_{n+1} , la connaissance de ce qui c'est passé avant l'instant n n'apporte aucun renseignement utile.

2. Définition des martingales

On se fixe, pour toute la suite, un espace probabilisé filtré $(\Omega, \mathcal{A}, (\mathcal{F}_n)_{n \geq 0}, \mathbb{P})$.

Définition 4. Un processus stochastique $(X_n)_{n \geq 0}$ est une martingale par rapport à la filtration $(\mathcal{F}_n)_{n \geq 0}$ si :

- (1) $\mathbb{E}[|X_n|] < +\infty$ (ie X_n est intégrable);
- (2) $(X_n)_{n \geq 0}$ est adapté à la filtration $(\mathcal{F}_n)_{n \geq 0}$;
- (3) $\mathbb{E}[X_{n+1}|\mathcal{F}_n] = X_n$ p.s..

Lorsqu'on ne précise pas la filtration, on suppose que l'on a pris la filtration canonique ou naturelle. On dira que $(X_n)_{n \geq 0}$ est une martingale par rapport au processus $(Y_n)_{n \geq 0}$, si on a choisi $\mathcal{F}_n = \sigma(Y_0, \dots, Y_n)$. On peut remarquer que, par définition de l'espérance conditionnelle la dernière propriété est équivalente à

$$\forall A \in \mathcal{F}_n, \quad \mathbb{E}[1_A X_{n+1}] = \mathbb{E}[1_A X_n],$$

ou encore à

$$\mathbb{E}[(X_{n+1} - X_n)|\mathcal{F}_n] = 0.$$

Historiquement, les martingales ont été introduites en lien avec les jeux de hasard ; X_n représentait alors la fortune du joueur à la n -ième partie, et la condition de martingale exprimait le fait que le jeu était équilibré. **Exemples**

- Si $(\mathcal{F}_n)_{n \in \mathbb{N}}$ est une filtration et si X est une variable aléatoire intégrable, alors $X_n = \mathbb{E}(X|\mathcal{F}_n)$ définit une martingale. C'est la *martingale de Doob*.
- Si $(X_n)_{n \geq 0}$ est un processus adapté intégrable, alors $S_n = X_0 + \dots + X_n$ définit une martingale si et seulement si $\mathbb{E}[X_{n+1}|\mathcal{F}_n] = 0$. En particulier si $(X_n)_{n \geq 0}$ est une suite de variables aléatoires indépendantes centrées telles que $X_0 = 0$ alors $S_n = X_0 + \dots + X_n$ est une martingale par rapport à $\mathcal{F}_n = \sigma(X_0, \dots, X_n)$. Donnons deux premières propriétés simples mais importantes des martingales.

Proposition 5. Si $(X_n)_{n \geq 0}$ est une martingale alors $\mathbb{E}(X_n) = \mathbb{E}(X_0)$. On dit qu'une martingale est à espérance mathématique constante.

DÉMONSTRATION. D'après la propriété des martingales on a

$$\mathbb{E}(X_n) = \mathbb{E}[\mathbb{E}[X_{n+1}|\mathcal{F}_n]] = \mathbb{E}(X_{n+1})$$

d'où par récurrence on obtient le résultat. □

Proposition 6. Si $(X_n)_{n \geq 0}$ est une martingale alors pour tout $n \geq 0$ et tout $k \geq 0$ on a

$$\mathbb{E}[X_{n+k}|\mathcal{F}_n] = X_n$$

Une manière équivalente de donner ce résultat est de dire que pour $m < n$, $\mathbb{E}[X_n|\mathcal{F}_m] = X_m$.

DÉMONSTRATION. On sait déjà que le propriété est vraie pour $k = 0$ et $k = 1$. Procédons par récurrence sur $k \geq 1$. D'après la propriété des martingales, et les propriétés de l'espérance conditionnelle on a

$$\mathbb{E}[X_{n+k+1}|\mathcal{F}_n] = \mathbb{E}[\mathbb{E}[X_{n+k+1}|\mathcal{F}_{n+k}]|\mathcal{F}_n] = \mathbb{E}[X_{n+k}|\mathcal{F}_n] = X_n$$

la dernière égalité s'obtenant avec l'hypothèse de récurrence. D'où, on obtient le résultat. \square

3. Surmartingale et sous-martingale

Définition 7. Un processus stochastique $(X_n)_{n \geq 0}$ est une surmartingale (respectivement une sous-martingale) par rapport à la filtration $(\mathcal{F}_n)_{n \geq 0}$ si :

- (1) si $\mathbb{E}[|X_n|] < +\infty$ (ie X_n est intégrable);
- (2) si $(X_n)_{n \geq 0}$ est adapté à la filtration $(\mathcal{F}_n)_{n \geq 0}$;
- (3) si $\mathbb{E}[X_{n+1}|\mathcal{F}_n] \leq X_n$ p.s. (resp. $\mathbb{E}[X_{n+1}|\mathcal{F}_n] \geq X_n$ p.s.).

Il est évident que (X_n) est une surmartingale si et seulement si $(-X_n)$ est une sous-martingale. De plus (X_n) est une martingale si et seulement si c'est à la fois une surmartingale et une sous-martingale.

Les surmartingales et sousmartingales vont permettre d'avoir des résultats analogues au résultat "Une suite croissante majorée converge" ou "Une suite décroissante minorée converge"... mais dans un cadre probabiliste où les suites ne sont croissantes "qu'en espérance conditionnelle".

Exemples

Considérons la marche aléatoire dans \mathbb{Z} définie par : $S_0 = X_0 = a$ et pour $n \geq 1$,

$$S_n = \sum_{j=0}^n X_j$$

où les X_j ($j \geq 1$) sont des variables aléatoires indépendantes de même loi $p\delta_1 + (1-p)\delta_{-1}$ avec $0 < p < 1$. Il est facile de voir que l'on définit ainsi une martingale, respectivement une surmartingale ou une sous-martingale, si $p = 1/2$, respectivement si $p < 1/2$ ou si $p > 1/2$. Cette marche peut modéliser la fortune d'un joueur qui joue à pile ou face et qui, à chaque lancer gagne ou perd un euro.

Donnons maintenant des exemples de constructions de surmartingales et sous-martingales.

Proposition 8. Soit $(X_n)_{n \geq 0}$ une martingale (resp. une sous-martingale). Si φ est une fonction convexe (resp. convexe croissante) telle que $\varphi(X_n)$ soit intégrable alors $(\varphi(X_n))_{n \geq 0}$ est une sous-martingale.

Soit $(X_n)_{n \geq 0}$ une martingale (resp. une surmartingale). Si φ est une fonction concave (resp. concave croissante) telle que $\varphi(X_n)$ soit intégrable alors $(\varphi(X_n))_{n \geq 0}$ est une surmartingale.

DÉMONSTRATION. Pour la première partie de la proposition, d'après l'inégalité de Jensen on a :

$$\mathbb{E}[\varphi(X_{n+1})|\mathcal{F}_n] \geq \varphi(\mathbb{E}[X_{n+1}|\mathcal{F}_n])$$

ce dernier terme étant égal à $\varphi(X_n)$ si au départ on a une martingale, sinon il est supérieur ou égal à $\varphi(X_n)$ si on a une sous-martingale et une fonction croissante. La deuxième partie s'obtient à partir de la première par passage à l'opposé. \square

Corollaire 9.

- (1) Soit $(X_n)_{n \geq 0}$ une martingale. Alors $(|X_n|)_{n \geq 0}$ et $(X_n^2)_{n \geq 0}$ (si X_n^2 est intégrable) sont des sous-martingales.
- (2) Soit $(X_n)_{n \geq 0}$ une sous-martingale alors $(X_n^+ = \sup(X_n, 0))_{n \geq 0}$ est une sous-martingale.

Nous allons maintenant énoncer quelques propriétés qui sont identiques à celles des martingales. Nous ne démontrerons pas ces résultats, il suffira d'adapter les démonstrations données dans le paragraphe des martingales.

Proposition 10. *Si $(X_n)_{n \geq 0}$ est une surmartingale (resp. une sous-martingale) alors :*

$$(\mathbb{E}(X_n))_{n \geq 0} \text{ est une suite décroissante (resp. croissante);}$$

$$\forall n \geq 0, \forall k \geq 0, \mathbb{E}[X_{n+k} | \mathcal{F}_n] \leq X_n \text{ (resp. } \geq) \quad \text{ou encore} \quad \forall m \geq n, \mathbb{E}[X_n | \mathcal{F}_m] \leq X_m \text{ (resp. } \geq)$$

4. Temps d'arrêt

Soit $(\Omega, \mathcal{A}, (\mathcal{F}_n)_{n \in \mathbb{N}}, \mathbb{P})$ est un espace probabilisé filtré. On pose

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

Définition 11. *Une variable aléatoire $T : \Omega \rightarrow \bar{\mathbb{N}} = \mathbb{N} \cup \{\infty\}$ est appelée temps d'arrêt si pour tout entier $n \in \mathbb{N}$, on a*

$$\{T = n\} \in \mathcal{F}_n.$$

Il est facile de voir que cela équivaut à $\forall n, \{T \leq n\} \in \mathcal{F}_n$ ou encore $\forall n, \{T > n\} \in \mathcal{F}_n$. De plus cela entraîne facilement que $\{T < n\} \in \mathcal{F}_{n-1}$ et $\{T \geq n\} \in \mathcal{F}_{n-1}$.

Il est important de remarquer que la valeur $+\infty$ est autorisée. En écrivant

$$\{T = +\infty\} = \Omega \setminus \cup_{n \geq 0} \{T = n\}$$

on voit que $\{T = +\infty\} \in \mathcal{F}_\infty$.

Exemples :

- (i) Si $k \in \mathbb{N}$, le temps constant $T = k$ est évidemment un temps d'arrêt.
- (ii) Soit $(X_n)_{n \in \mathbb{N}}$ un processus adapté à valeurs dans (E, \mathcal{B}) . Pour $B \in \mathcal{B}$ on pose

$$T_B = \begin{cases} \inf\{n \in \mathbb{N} : X_n \in B\} & \text{si } \{n \in \mathbb{N} : X_n \in B\} \neq \emptyset \\ +\infty & \text{sinon} \end{cases}$$

est un temps d'arrêt, appelé *temps d'entrée* dans B . En général on omet la deuxième partie de la définition et on prend juste la convention $\inf \emptyset = +\infty$. Il suffit de remarquer :

$$\{T_B = n\} = \{X_0 \notin B, X_1 \notin B, \dots, X_{n-1} \notin B, X_n \in B\} \in \mathcal{F}_n.$$

Proposition 12. *Soit S et T deux temps d'arrêt. Alors, $S+T$, $S \wedge T$ et $S \vee T$ sont des temps d'arrêts. En particulier, pour $k \in \mathbb{N}$, $T \wedge k$ est un temps d'arrêt borné.*

Par généralisation on a, si $s = (T_k)_{k \geq 0}$ est une suite de temps d'arrêt, alors $\inf T_k$, $\sup T_k$, $\liminf T_k$ et $\limsup T_k$ sont aussi des temps d'arrêt.

DÉMONSTRATION. Il suffit de remarquer que

$$\begin{aligned} \{S + T = n\} &= \cup_{0 \leq k \leq n} \{S = k\} \cap \{T = n - k\} \\ \{S \wedge T \leq n\} &= \{S \leq n\} \cup \{T \leq n\} \\ \{S \vee T \leq n\} &= \{S \leq n\} \cap \{T \leq n\} \end{aligned}$$

et par exemple $\{\inf T_k \leq n\} = \cup \{T_k \leq n\}$ et $\{\liminf T_k \leq n\} = \cup_{m \in \mathbb{N}} \cap_{k \geq m} \{T_k \leq n\}$. □

Définition 13. Si T est un temps d'arrêt, on appelle tribu des événements antérieurs à T la tribu suivante

$$\mathcal{F}_T = \{A \in \mathcal{F}_\infty : \forall n \in \mathbb{N}, A \cap \{T = n\} \in \mathcal{F}_n\}$$

On peut vérifier que T est une variable aléatoire \mathcal{F}_T -mesurable. En effet, pour tout entier $k \geq 0$, on a pour tout $n \geq 0$, $\{T = k\} \cap \{T = n\}$ est égal à $\{T = n\} \in \mathcal{F}_n$ si $k = n$, où est égal à l'ensemble vide; d'où $\{T = k\} \in \mathcal{F}_T$.

Proposition 14. Soit S et T deux temps d'arrêt. Alors :

$$S \leq T \Rightarrow \mathcal{F}_S \subset \mathcal{F}_T$$

DÉMONSTRATION. Soit $A \in \mathcal{F}_S$. Alors on a

$$A \cap \{T = n\} = \bigcup_{k=0}^n [A \cap \{S = k\} \cap \{T = n\}]$$

Or $A \cap \{S = k\} \in \mathcal{F}_k \subset \mathcal{F}_n$, d'où par passage à la réunion $A \cap \{T = n\} \in \mathcal{F}_n$. \square

Proposition 15. Soit $(X_n)_{n \geq 0}$ un processus adapté, et T un temps d'arrêt. Alors la variable aléatoire $1_{T < +\infty} X_T$ définie par

$$1_{T < +\infty} X_T(\omega) = \begin{cases} X_n(\omega) & \text{si } T(\omega) = n \in \mathbb{N} \\ 0 & \text{si } T = +\infty \end{cases}$$

est \mathcal{F}_T -mesurable.

DÉMONSTRATION. Si les $(X_n)_{n \geq 0}$ sont à valeurs dans (E, \mathcal{B}) et si $B \in \mathcal{B}$, alors $\{X_T \in B\} \in \mathcal{F}_T$ si et seulement si pour tout entier $n \geq 0$, $\{X_T \in B\} \cap \{T = n\} \in \mathcal{F}_n$. Or on a $\{X_T \in B\} \cap \{T = n\} = \{X_n \in B\} \cap \{T = n\}$ qui est bien dans \mathcal{F}_n . \square

5. Propriétés des martingales par rapport aux temps d'arrêts

Soit $(X_n)_{n \geq 0}$ un processus adapté à une filtration $(\mathcal{F}_n)_{n \geq 0}$ et T un temps d'arrêt adapté à la même filtration. On définit un nouveau processus, appelé *processus arrêté* et noté $X^{|T} = (X_{n \wedge T})_{n \geq 0}$, en posant

$$X_{n \wedge T}(\omega) = \begin{cases} X_n(\omega) & \text{si } n < T(\omega) \\ X_{T(\omega)}(\omega) & \text{si } n \geq T(\omega) \end{cases}$$

Par conséquent on a

$$X_{n \wedge T}(\omega) = X_n(\omega) 1_{\{n < T(\omega)\}} + X_{T(\omega)}(\omega) 1_{\{T(\omega) \leq n\}}$$

ce qui montre bien que ce processus est encore adapté à la filtration $(\mathcal{F}_n)_{n \geq 0}$. Une autre manière d'exprimer ce processus est la suivante

$$\begin{aligned} X_{n \wedge T} &= X_0 1_{\{T=0\}} + X_1 1_{\{T=1\}} + \cdots + X_n 1_{\{T=n\}} + X_n 1_{\{T > n\}} \\ &= X_0 + \sum_{k=0}^{n-1} (X_{k+1} - X_k) 1_{\{T > k\}}. \end{aligned}$$

Théorème 16.

Si $(X_n)_{n \geq 0}$ est une martingale et T un temps d'arrêt par rapport à la filtration $(\mathcal{F}_n)_{n \geq 0}$ alors le processus arrêté, $(X_{n \wedge T})_{n \geq 0}$, est encore une martingale.

DÉMONSTRATION. Si $\{T \leq n\}$ alors $n \wedge T = (n+1) \wedge T$ et par conséquent $(X_{(n+1) \wedge T} - X_{n \wedge T})1_{\{T \leq n\}} = 0$. D'où, comme $\{T \geq n+1\} = \{T \leq n\}^c \in \mathcal{F}_n$, on a

$$\mathbb{E}[X_{(n+1) \wedge T} - X_{n \wedge T} | \mathcal{F}_n] = \mathbb{E}[(X_{(n+1) \wedge T} - X_{n \wedge T})1_{\{T \geq n+1\}} | \mathcal{F}_n] = 1_{\{T \geq n+1\}} \mathbb{E}[X_{n+1} - X_n | \mathcal{F}_n] = 0.$$

□

Donnons un résultat similaire pour les surmartingales et les sous-martingales.

Proposition 17. *Si $(X_n)_{n \geq 0}$ est une surmartingale (resp. une sous-martingale) et si T est un temps d'arrêt par rapport à la filtration $(\mathcal{F}_n)_{n \geq 0}$ alors le processus arrêté, $(X_{n \wedge T})_{n \geq 0}$, est encore une surmartingale (resp. une sous-martingale).*

On souhaite maintenant étudier le comportement de la variable aléatoire X_T , obtenue lorsqu'on remplace le temps linéaire n dans une martingale par un temps d'arrêt T . On se placera dans le cas particulier où T est un temps d'arrêt borné.

La variable aléatoire X_T est appelée *variable aléatoire terminale* du processus arrêté $(X_{n \wedge T})_{n \geq 0}$. Elle est définie si le temps d'arrêt est presque sûrement fini.

Proposition 18. *Si T est un temps d'arrêt presque sûrement fini, i.e. $\mathbb{P}(T < +\infty) = 1$ alors*

$$X_{n \wedge T} \xrightarrow[n \rightarrow \infty]{p.s.} X_T$$

DÉMONSTRATION. Pour presque tout ω on a $T(\omega) < +\infty$. Soit alors un tel ω . Pour tout $n \geq T(\omega)$, $X_{n \wedge T}(\omega) = X_{n \wedge T(\omega)}(\omega) = X_{T(\omega)}(\omega)$. □

On peut remarquer que si $T(\omega) = +\infty$ alors $X_{n \wedge T}(\omega) = X_n(\omega)$. Si par exemple la suite $(X_n(\omega))$ est convergente alors, on pose $X_T(\omega) = \lim_{n \rightarrow +\infty} X_n(\omega)$. On peut donc prolonger la définition de la variable aléatoire X_T sur $\{T = +\infty\}$ et donc pour tout ω on a $X_T(\omega) = \lim_{n \rightarrow +\infty} X_{n \wedge T}(\omega)$.

Si $X_{n \wedge T}$ est une martingale, on a $\mathbb{E}[X_{n \wedge T}] = \mathbb{E}[X_0]$. Si de plus on a, comme dans la proposition précédente, $X_{n \wedge T} \xrightarrow[n \rightarrow \infty]{p.s.} X_T$, on a envie d'en déduire $\mathbb{E}[X_T] = \mathbb{E}[X_0]$. Les conditions sous lesquelles cette égalité est valide forment ce qu'on appelle des "théorèmes d'arrêt".

6. Théorèmes d'arrêt

Le théorème le plus simple est le suivant

Théorème 19.

Théorème d'arrêt borné. *Soient $(X_n)_{n \geq 0}$ une martingale (resp. une surmartingale, resp. une sous-martingale) et T un temps d'arrêt (pour la même filtration) borné par un entier N . Alors on a*

$$\mathbb{E}(X_0) = \mathbb{E}(X_T) = \mathbb{E}(X_N) \quad (\text{resp. } \mathbb{E}[X_0] \geq \mathbb{E}(X_T) \geq \mathbb{E}(X_N), \text{ resp. } \mathbb{E}(X_0) \leq \mathbb{E}(X_T) \leq \mathbb{E}(X_N))$$

DÉMONSTRATION. Comme T est borné par N , $X_{N \wedge T} = X_T$. Par conséquent, en utilisant le théorème 16 et la proposition 17 on obtient facilement le résultat concernant X_0 et X_T .

De plus, on peut remarquer que

$$X_N - X_T = \sum_{k=T+1}^N (X_k - X_{k-1}) = \sum_{k=1}^N (X_k - X_{k-1})1_{k-1 \geq T}$$

Donc

$$\mathbb{E}(X_N - X_T) = \sum_{k=1}^N \mathbb{E}((X_k - X_{k-1})1_{k-1 \geq T}) = \sum_{k=1}^N \mathbb{E}(\mathbb{E}[(X_k - X_{k-1})1_{k-1 \geq T} | \mathcal{F}_{k-1}]) = \sum_{k=1}^N \mathbb{E}(1_{k-1 \geq T} \mathbb{E}[(X_k - X_{k-1}) | \mathcal{F}_{k-1}])$$

où on a utilisé le fait que l'événement $T \leq (k-1)$ est \mathcal{F}_{k-1} mesurable. Or, par définition, tous les $\mathbb{E}[(X_k - X_{k-1}) | \mathcal{F}_{k-1}]$ sont de même signes (positif ou négatif selon l'on a une sur- ou une sous-martingale), on en déduit le signe de $\mathbb{E}(X_N - X_T)$. \square

Remarque : on peut étendre le théorème précédent au cas où l'on a deux temps d'arrêt T et S bornés, tels que $T \leq S$ p.s..

Un théorème d'arrêt n'est rien d'autre qu'un théorème d'interversion limite et intégrale. On a donc tout aussi facilement :

Théorème 20.

Soient $(X_n)_{n \geq 0}$ une martingale (resp. une surmartingale, resp. une sous-martingale) et T un temps d'arrêt presque sûrement fini tous les deux adaptés à la filtration $(\mathcal{F}_n)_{n \geq 0}$. Alors s'il existe une variable aléatoire Y intégrable telle que $\forall n, |X_{T \wedge n}| \leq Y$, on a

$$\mathbb{E}(X_T) = \mathbb{E}(X_0) \quad (\text{resp. } \mathbb{E}(X_T) \leq \mathbb{E}(X_0), \text{ resp. } \mathbb{E}(X_T) \geq \mathbb{E}(X_0)).$$

Ce théorème s'applique en particulier lorsque $X_{T \wedge n}$ est bornée.

DÉMONSTRATION. (dans le cas des martingales) : On sait que $\mathbb{E}[X_{n \wedge T}] = \mathbb{E}[X_0]$. Si de plus on a, comme dans la proposition précédente, $X_{n \wedge T} \xrightarrow[n \rightarrow \infty]{\text{p.s.}} X_T$ Il suffit ensuite d'appliquer le théorème de convergence dominée. \square

On peut faire un peu plus sophistiqué :

Théorème 21.

Théorème d'arrêt non borné Soient $(X_n)_{n \geq 0}$ une martingale (resp. une surmartingale, resp. une sous-martingale) et T un temps d'arrêt presque sûrement fini tous les deux adaptés à la filtration $(\mathcal{F}_n)_{n \geq 0}$. On suppose que $\mathbb{E}[|X_T|] < \infty$ et que $\lim_{n \rightarrow \infty} \mathbb{E}[1_{T > n} X_n] = 0$. Alors on a

$$\mathbb{E}(X_T) = \mathbb{E}(X_0) \quad (\text{resp. } \mathbb{E}(X_T) \leq \mathbb{E}(X_0), \text{ resp. } \mathbb{E}(X_T) \geq \mathbb{E}(X_0)).$$

DÉMONSTRATION. On a $X_T - X_{T \wedge n} = (X_T - X_{T \wedge n})1_{T \leq n} + (X_T - X_{T \wedge n})1_{T > n} = X_T 1_{T > n} - X_n 1_{T > n}$
 En prenant l'espérance, on obtient $\mathbb{E}[X_T] - \mathbb{E}[X_{T \wedge n}] = \mathbb{E}[X_T 1_{T > n}] - \mathbb{E}[X_n 1_{T > n}]$
 Le deuxième terme du membre de gauche tend vers 0 par hypothèse. Le premier également : en effet $\lim X_T 1_{T > n} = 0$ car T est p.s. fini ; de plus $|X_T 1_{T > n}| \leq |X_T|$ qui est intégrable. On peut donc utiliser le théorème de convergence dominée. \square

7. Inégalités maximales

Ce sont des inégalités plus fortes que les inégalités de Markov.

Théorème 22.**Inégalité maximale de Doob**

Soit $(X_n)_{n \geq 0}$ une sous-martingale positive. Alors, pour tout $\lambda > 0$, on a :

$$\mathbb{P} \left(\max_{0 \leq k \leq n} X_k > \lambda \right) \leq \frac{\mathbb{E}(X_n)}{\lambda}$$

On a alors pour $\lambda > 0$

$$\mathbb{P} \left(\sup_{k \in \mathbb{N}} X_k > \lambda \right) \leq \frac{1}{\lambda} \sup_{k \in \mathbb{N}} \mathbb{E}[|X_k|].$$

DÉMONSTRATION. Posons $M_n = \{\max_{0 \leq k \leq n} X_k > \lambda\} = \cup_{0 \leq k \leq n} \{X_k > \lambda\}$. On considère le temps d'arrêt T défini par $T = \inf\{0 \leq k, X_k > \lambda\}$. $T \wedge n$ est alors un temps d'arrêt borné par n , auquel on peut appliquer le théorème d'arrêt ; on a alors :

$$\begin{aligned} \mathbb{E}(X_n) &\geq \mathbb{E}(X_{n \wedge T}) \geq \mathbb{E}(X_T 1_{M_n}) + \mathbb{E}(X_n 1_{M_n^c}) \\ &\geq \mathbb{E}(X_T 1_{M_n}) \quad \text{car la sous-martingale est positive} \\ &\geq \lambda \mathbb{E}(1_{M_n}) = \lambda \mathbb{P}(M_n) \end{aligned}$$

d'où le résultat. □

Corollaire 23. Soit $(X_n)_{n \geq 0}$ une martingale. Alors, pour tout $\lambda > 0$, on a :

$$\mathbb{P} \left(\max_{0 \leq k \leq n} |X_k| > \lambda \right) \leq \frac{\mathbb{E}(|X_n|)}{\lambda}$$

DÉMONSTRATION. C'est une conséquence du résultat précédent et du fait que si $(X_n)_{n \geq 0}$ est une martingale alors $(|X_n|)_{n \geq 0}$ est une sous-martingale positive. □

Corollaire 24. Soit $(X_n)_{n \geq 0}$ une martingale telle que pour tout $n \geq 0$ on ait $\mathbb{E}(X_n^2) < +\infty$. Alors, pour tout $\lambda > 0$, on a :

$$\mathbb{P} \left(\max_{0 \leq k \leq n} |X_k| > \lambda \right) \leq \frac{\mathbb{E}(X_n^2)}{\lambda^2}$$

DÉMONSTRATION. Avec les hypothèses du corollaire, on sait que $(X_n^2)_{n \geq 0}$ est une sous-martingale positive. Ensuite il suffit d'appliquer la proposition ci-dessus. □

Le corollaire ci-dessus généralise l'inégalité de Kolmogorov (voir la loi forte des grands nombres) : soit $(Y_n)_{n \geq 1}$ une suite de variables aléatoires indépendantes identiquement distribuées d'espérance nulle et de variance σ^2 . On pose $Y_0 = 0$ et $S_n = Y_0 + Y_1 + \dots + Y_n$; alors pour tout $\varepsilon > 0$

$$\mathbb{P} \left(\sup_{0 \leq k \leq n} |S_k| > \varepsilon \right) \leq \frac{\mathbb{E}(S_n^2)}{\varepsilon^2}$$

En effet, sous les hypothèses, $(S_n)_{n \geq 0}$ est une martingale telle que $\mathbb{E}(S_n^2) < +\infty$, on peut donc appliquer le corollaire.

Pour terminer ce paragraphe, nous donnons un résultat pour les surmartingales positives.

Théorème 25.

Soit $(X_n)_{n \geq 0}$ une surmartingale positive. Alors, pour tout $\lambda > 0$, on a :

$$\mathbb{P} \left(\max_{0 \leq k \leq n} X_k > \lambda \right) \leq \frac{\mathbb{E}(X_0)}{\lambda}$$

DÉMONSTRATION. Posons $M_n = \{\max_{0 \leq k \leq n} X_k > \lambda\}$, et soit T le temps d'arrêt défini par $T = \inf \{k : 0 \leq k \leq n, X_k > \lambda\}$. Si on applique le théorème d'arrêt à la surmartingale $(X_n)_{n \geq 0}$ pour le temps d'arrêt borné $T \wedge n$, on a par positivité,

$$\mathbb{E}[X_0] \geq \mathbb{E}[X_{n \wedge T}] = \mathbb{E}[X_T 1_{M_n}] + \mathbb{E}[X_n 1_{M_n^c}] \geq \mathbb{E}[X_T 1_{M_n}] \geq \lambda \mathbb{P}(M_n).$$

□

8. Convergence des martingales

Il est naturel de se poser la question suivante :

Si $(X_n)_{n \geq 0}$ est une martingale alors existe-t-il une variable aléatoire X telle que $\lim_{n \rightarrow +\infty} X_n = X$? Si c'est le cas, en quel sens a lieu cette convergence (presque sûre, en probabilité, en loi, dans L_1 ou L_2) ? D'autre part a-t-on, alors $\mathbb{E}[X] = \mathbb{E}[X_0]$ ou encore $\mathbb{E}[X | \mathcal{F}_n] = X_n$?

Nous allons essayer de répondre à ces questions.

Théorème 26.

Convergence p.s. pour les martingales bornées dans L_1

Soit $(X_n)_{n \geq 0}$ une martingale bornée dans L^1 , i.e. $\sup_{n \geq 0} \mathbb{E}(|X_n|) < +\infty$. Alors $(X_n)_{n \geq 0}$ converge presque sûrement vers une variable aléatoire X_∞ intégrable.

Pour démontrer le résultat ci-dessus on va démontrer

$$\mathbb{P}(\liminf_{n \rightarrow +\infty} X_n < \limsup_{n \rightarrow +\infty} X_n) = 0$$

Pour cela on remarque que

$$\{\liminf_{n \rightarrow +\infty} X_n < \limsup_{n \rightarrow +\infty} X_n\} = \cup_{a,b \in \mathbb{Q}, a < b} \{\liminf_{n \rightarrow +\infty} X_n < a < b < \limsup_{n \rightarrow +\infty} X_n\}$$

Or l'événement $\{\liminf_{n \rightarrow +\infty} X_n < a < b < \limsup_{n \rightarrow +\infty} X_n\}$ entraîne qu'il y a une infinité d'indices n tels que $\{X_n < a\}$ se réalise et une infinité d'indices n tels que $\{X_n > b\}$ se réalise. Donc, si $M_{a,b}$ compte le nombre de traversées de la bande $[a, b]$ par la suite X_n ,

$$\{\liminf_{n \rightarrow +\infty} X_n < \limsup_{n \rightarrow +\infty} X_n\} \subset \cup_{a,b \in \mathbb{Q}, a < b} \{M_{a,b} = +\infty\}$$

Il suffit donc de montrer que pour tout $a < b$, $\mathbb{P}(M_{a,b} = +\infty) = 0$. On va en fait montrer à l'aide de deux lemmes que $\mathbb{E}[M_{a,b}] < \infty$, ce qui est un résultat plus fort.

De façon plus précise, on se donne deux nombres a et b tels que $a < b$ et on définit une double suite de temps d'arrêt $S_1 < T_1 < S_2 < T_2 < \dots$ par

$$\begin{aligned} S_1 &= \inf_{n \geq 0} \{X_n \leq a\}, & T_1 &= \inf_{n > S_1} \{X_n \geq b\} \\ S_2 &= \inf_{n > T_1} \{X_n \leq a\}, & T_2 &= \inf_{n > S_2} \{X_n \geq b\} \end{aligned}$$

et ainsi de suite, avec la convention que si l'une des bornes inférieures n'existe pas on donne la valeur infini. On note

$$M_{a,b} = \sum_{k \geq 1} 1_{\{T_k < +\infty\}}$$

qui représente le nombre total de traversées de $[a, b]$, en montant effectuées par la trajectoire $n \mapsto X_n$.

Lemme 27. Inégalité de Dubins

Pour tout $k \geq 1$ et tout $n \geq 1$, on a :

$$(b - a)\mathbb{P}(T_k < n) \leq \mathbb{E}[(a - X_n)1_{\{S_k \leq n < T_k\}}]$$

DÉMONSTRATION. L'entier n étant fixé, posons

$$D_k = X_{T_k \wedge n} - X_{S_k \wedge n}$$

En appliquant le théorème d'arrêt aux temps d'arrêts bornés $T_k \wedge n$ et $S_k \wedge n$ on remarque que

$$\mathbb{E}[D_k] = \mathbb{E}[X_1] - \mathbb{E}[X_1] = 0.$$

D'autre part, par définition des S_k et T_k on a

$$\begin{aligned} \{n < S_k\} &\subset \{D_k = 0\} & \text{et} & & \{T_k \leq n\} &\subset \{D_k \geq b - a\}; \\ \{S_k \leq n < T_k\} &\subset \{D_k = X_n - X_{S_k}\} &\subset \{D_k \geq X_n - a\}; \end{aligned}$$

et par conséquent

$$(b - a)1_{\{T_k \leq n\}} + (X_n - a)1_{\{S_k \leq n < T_k\}} \leq D_k$$

et en prenant l'espérance mathématique :

$$(b - a)\mathbb{P}(T_k \leq n) + \mathbb{E}[(X_n - a)1_{\{S_k \leq n < T_k\}}] \leq \mathbb{E}[D_k] = 0$$

□

Lemme 28. Lemme de Doob

Avec les mêmes notations on a :

$$\mathbb{E}[M_{a,b}] \leq \frac{|a|}{b - a} + \frac{1}{b - a} \sup_{n \geq 0} \mathbb{E}(|X_n|)$$

DÉMONSTRATION. D'après le lemme précédent on a

$$(b - a) \sum_{k \geq 1} \mathbb{P}(T_k \leq n) \leq \sum_{k \geq 1} \mathbb{E}[(a - X_n)1_{\{S_k \leq n < T_k\}}].$$

Or les événements $\{S_k \leq n < T_k\}$ pour $k \geq 1$ sont disjoints. On note A leur réunion. Comme $\{S_k \leq n < T_k\} = \emptyset$ si $k \geq n$, on a

$$\begin{aligned} \sum_{k \geq 1} \mathbb{E}[(a - X_n)1_{\{S_k \leq n < T_k\}}] &= \sum_{1 \leq k \leq n} \mathbb{E}[(a - X_n)1_{\{S_k \leq n < T_k\}}] = \mathbb{E} \left[(a - X_n) \sum_{1 \leq k \leq n} 1_{\{S_k \leq n < T_k\}} \right] \\ &= \mathbb{E} \left[(a - X_n) \sum_{k \geq 1} 1_{\{S_k \leq n < T_k\}} \right] = \mathbb{E}[(a - X_n)1_A] \end{aligned}$$

d'où

$$(b - a) \sum_{k \geq 1} \mathbb{P}(T_k \leq n) \leq \mathbb{E}[(a - X_n)1_A]$$

Maintenant, on a la majoration : $\mathbb{E}[(a - X_n)1_A] \leq |a| + \sup_{n \geq 0} \mathbb{E}(|X_n|)$; d'où la majoration pour tout n

$$\Sigma_n = \sum_{k \geq 1} \mathbb{P}(T_k \leq n) \leq \frac{|a| + \sup_{n \geq 0} \mathbb{E}(|X_n|)}{b - a} = \alpha.$$

La suite $(\Sigma_n)_{n \geq 0}$ étant croissante, on en déduit qu'elle est convergente. D'où, comme $\mathbb{P}(T_k < +\infty) = \lim_{n \rightarrow +\infty} \mathbb{P}(T_k < n)$, on obtient

$$\mathbb{E}[M_{a,b}] = \sum_{k \geq 1} \mathbb{P}(T_k < +\infty) \leq \alpha.$$

□

Ceci nous permet de terminer la preuve de la convergence presque sûre des martingales L^1 .

Lemme 29. *La suite $(X_n)_{n \geq 0}$ est presque sûrement convergente*

DÉMONSTRATION. La variable aléatoire $M_{a,b}$ est positive et intégrable. Elle est donc presque sûrement finie, ou encore $\mathbb{P}(M_{a,b} = +\infty) = 0$. La réunion dénombrable

$$\cup_{a,b \in \mathbb{Q}, a < b} \{M_{a,b} = +\infty\}$$

est donc aussi négligeable.

Puisque

$$\{\liminf_{n \rightarrow +\infty} X_n < \limsup_{n \rightarrow +\infty} X_n\} = \cup_{a,b \in \mathbb{Q}, a < b} \{\liminf_{n \rightarrow +\infty} X_n < a < b < \limsup_{n \rightarrow +\infty} X_n\} \subset \cup_{a,b \in \mathbb{Q}, a < b} \{M_{a,b} = +\infty\}$$

on a bien

$$\mathbb{P}(\liminf_{n \rightarrow +\infty} X_n < \limsup_{n \rightarrow +\infty} X_n) = 0$$

par conséquent on obtient le résultat. □

Pour terminer la démonstration du théorème, il reste à montrer que la limite presque sûre des X_n est une variable aléatoire intégrable. On a $|X_n| \rightarrow |X_\infty|$ presque sûrement et pour tout $n \geq 0$, $\mathbb{E}[|X_n|] \leq \sup_{n \geq 0} \mathbb{E}(|X_n|) < +\infty$, d'où par le lemme de Fatou on a

$$\mathbb{E}[|X_\infty|] = \mathbb{E}[\liminf_{n \rightarrow +\infty} |X_n|] = \mathbb{E}[\liminf_{n \rightarrow +\infty} |X_n|] \leq \liminf_{n \rightarrow +\infty} \mathbb{E}[|X_n|] \leq \sup_{n \geq 0} \mathbb{E}(|X_n|) < +\infty.$$

Corollaire 30. *Le théorème est encore vrai si on prend une surmartingale ou une sous-martingale.*

DÉMONSTRATION. Si $(X_n)_{n \geq 0}$ est une surmartingale, on a $\mathbb{E}[D_n] \leq 0$, car une surmartingale est décroissante ; ensuite les inégalités restent toutes dans le même sens.

Si on part maintenant d'une sous-martingale $(X_n)_{n \geq 0}$, alors $(-X_n)_{n \geq 0}$ est une surmartingale. A toute traversée en montant de $[a, b]$ correspond une traversée en descendant de $[-b, -a]$. Donc le théorème appliqué pour les surmartingales entraîne le résultat pour les sous-martingales. □

Corollaire 31. *Une martingale de signe constant (resp. une surmartingale positive, resp. une sous-martingale négative) est presque sûrement convergente.*

DÉMONSTRATION. Si $(X_n)_{n \geq 0}$ est une martingale positive, alors on a pour tout $n \geq 0$ les relations $\mathbb{E}[|X_n|] = \mathbb{E}[X_n] = \mathbb{E}[X_0] < +\infty$ et le théorème s'applique (on procède de même si on a une martingale négative). Pour une surmartingale positive, on a $\mathbb{E}[|X_n|] = \mathbb{E}[X_n] \leq \mathbb{E}[X_0] < +\infty$. Et enfin pour une sous-martingale négative $\mathbb{E}[|X_n|] = -\mathbb{E}[X_n] \leq -\mathbb{E}[X_0] < +\infty$. \square

9. Convergence des martingales L^2

Remarquons tout d'abord que si $(X_n)_{n \geq 0}$ est une martingale telle que pour $n \geq 0$, $\mathbb{E}(X_n^2) < +\infty$, alors pour tout $n \geq 0$, et tout $p \geq 1$ on a

$$\mathbb{E}[(X_{n+p} - X_n)^2] = \mathbb{E}[X_{n+p}^2] - \mathbb{E}[X_n^2]$$

Ce n'est rien d'autre que la formule de Pythagore appliquée à X_n qui est la projection orthogonale de X_{n+p} sur \mathcal{F}_n .

Théorème 32.

Convergence p.s. et L^2 des martingales bornées dans L^2 Soit $(X_n)_{n \geq 0}$ une martingale bornée dans L^2 , i.e. $\sup_{n \geq 0} \mathbb{E}(X_n^2) < +\infty$. Alors $(X_n)_{n \geq 0}$ converge dans L^2 et presque sûrement vers une variable aléatoire X_∞ telle que $X_n = \mathbb{E}[X_\infty | \mathcal{F}_n]$.
En particulier $\mathbb{E}[X_\infty] = \mathbb{E}[X_0]$.

DÉMONSTRATION. On pose $m = \sup_{n \geq 0} \mathbb{E}(X_n^2) < +\infty$. Comme $(X_n)_{n \geq 0}$ est une martingale, on sait que $(X_n^2)_{n \geq 0}$ est une sous-martingale et par conséquent $(\mathbb{E}(X_n^2))_{n \geq 0}$ est une suite réelle croissante. Cette suite est majorée par m , elle est donc convergente. Comme

$$\mathbb{E}[(X_{n+p} - X_n)^2] = \mathbb{E}[X_{n+p}^2] - \mathbb{E}[X_n^2]$$

on en déduit alors que la suite $(X_n)_{n \geq 0}$ est de Cauchy dans L^2 , et donc qu'elle converge dans L^2 , L^2 étant complet.

Montrons maintenant qu'on a convergence presque sûre. On vérifie facilement que pour $m \geq 1$ fixé, $((X_{m+k} - X_m)^2)_{k \geq 0}$ est une sous-martingale positive, donc d'après l'inégalité maximale de Doob, on a pour tout $a > 0$

$$\begin{aligned} \mathbb{P}\left(\max_{0 \leq k \leq n} |X_{m+k} - X_m| > a\right) &= \mathbb{P}\left(\max_{0 \leq k \leq n} (X_{m+k} - X_m)^2 > a^2\right) \\ &\leq \frac{\mathbb{E}[(X_{m+n} - X_m)^2]}{a^2} \leq \frac{\sup_{j \geq 0} \mathbb{E}[(X_{m+j} - X_m)^2]}{a^2} \end{aligned}$$

Or $(\{\max_{0 \leq k \leq n} |X_{m+k} - X_m| > a\})_{n \geq 0}$ est une suite croissante d'événements dont la réunion est égale à $\{\sup_{k \geq 0} |X_{m+k} - X_m| > a\}$; d'où

$$\mathbb{P}\left(\sup_{k \geq 0} |X_{m+k} - X_m| > a\right) = \lim_{n \rightarrow +\infty} \mathbb{P}\left(\max_{0 \leq k \leq n} |X_{m+k} - X_m| > a\right) \leq \frac{\sup_{j \geq 0} \mathbb{E}[(X_{m+j} - X_m)^2]}{a^2}$$

Par conséquent, $\lim_{m \rightarrow +\infty} \mathbb{P}(\sup_{k \geq 0} |X_{m+k} - X_m| > a) = 0$ pour tout $a > 0$ et donc en particulier si $a = 1/l$ pour $l \in \mathbb{N}^*$. Alors, si on pose $A_{m,l} = \{\sup_{k \geq 0} |X_{m+k} - X_m| > 1/l\}$, comme pour tout $m \geq 1$, $\mathbb{P}(\cap_{m \geq 1} A_{m,l}) \leq \mathbb{P}(A_{m,l})$, on a

$$\mathbb{P}(\cap_{m \geq 1} A_{m,l}) = 0 \quad \text{et donc} \quad \mathbb{P}(\cup_{l \geq 1} \cap_{m \geq 1} A_{m,l}) = 0$$

Si $\omega \notin \cup_{l \geq 1} \cap_{m \geq 1} A_{m,l}$, alors pour tout $l \geq 1$, il existe $m \geq 1$ tel que pour tout $k \geq 1$

$$|X_{m+k}(\omega) - X_m(\omega)| < 1/l.$$

Ceci implique que la suite $(X_n(\omega))_{n \geq 0}$ est de Cauchy dans \mathbb{R} et donc est convergente. On en déduit que $(X_n)_{n \geq 0}$ est presque sûrement convergente.

Pour le dernier point remarquons que pour tout $A \in \mathcal{F}_n$ on a

$$\begin{aligned} \mathbb{E}[1_A X_n] - \mathbb{E}[1_A \mathbb{E}[X_\infty | \mathcal{F}_n]] &= \mathbb{E}[1_A \mathbb{E}[X_{n+k} | \mathcal{F}_n] - 1_A \mathbb{E}[X_\infty | \mathcal{F}_n]] = \mathbb{E}[\mathbb{E}[1_A (X_{n+k} - X_\infty) | \mathcal{F}_n]] \\ &\leq \mathbb{E}[1_A |X_{n+k} - X_\infty|] \leq \mathbb{E}[|X_{n+k} - X_\infty|] \leq \sqrt{\mathbb{E}[|X_{n+k} - X_\infty|^2]} \end{aligned}$$

ce dernier terme tend vers 0 lorsque k tend vers l'infini. D'où on obtient que $X_n = \mathbb{E}[X_\infty | \mathcal{F}_n]$. \square

10. Complément : convergence dans L^1

Théorème 33.

Soit $(X_n)_{n \geq 0}$ une martingale. Les deux conditions suivantes sont équivalentes.

(i) La suite X_n converge vers X_∞ p.s. et dans L^1 .

(ii) Il existe une variable aléatoire Y intégrable telle que $X_n = \mathbb{E}[Y | \mathcal{F}_n]$ pour tout $n \in \mathbb{N}$.

De plus, si ces conditions sont satisfaites, on peut prendre $Y = X_\infty$ dans (ii). On dit alors que la martingale est fermée.

DÉMONSTRATION. Supposons d'abord (i). On a pour $m \geq n$, $X_n = \mathbb{E}[X_m | \mathcal{F}_n]$. D'autre part, $X \mapsto \mathbb{E}[X | \mathcal{F}_n]$ est une contraction de L^1 , i.e. $\mathbb{E}[|\mathbb{E}[X | \mathcal{F}_n]|] \leq \mathbb{E}[|X|]$. On en déduit que

$$X_n = \lim_{m \rightarrow +\infty} \mathbb{E}[X_m | \mathcal{F}_n] = \mathbb{E}[X_\infty | \mathcal{F}_n].$$

Pour la réciproque on a une martingale $(X_n)_{n \in \mathbb{N}}$ bornée dans L^1 et donc on sait qu'elle converge presque sûrement. Pour la convergence L^1 supposons d'abord que Y est bornée par une constante $K < \infty$. Il en est donc de même pour les variables aléatoires X_n , et le théorème de convergence dominée donne alors le résultat voulu.

Dans le cas général soit $\varepsilon > 0$, fixons M assez grand pour que

$$\mathbb{E}[|Y - Y1_{|Y| \leq M}|] < \varepsilon.$$

Alors, pour tout n ,

$$\mathbb{E}[|X_n - \mathbb{E}[Y1_{|Y| \leq M} | \mathcal{F}_n]|] \leq \mathbb{E}[|\mathbb{E}[Y - Y1_{|Y| \leq M} | \mathcal{F}_n]|] < \varepsilon.$$

D'après le cas borné, la martingale $\mathbb{E}[Y1_{|Y| \leq M} | \mathcal{F}_n]$ converge dans L^1 . Donc on peut choisir n_0 assez grand pour que, pour tous $m, n \geq n_0$,

$$\mathbb{E}[|\mathbb{E}[Y1_{|Y| \leq M} | \mathcal{F}_m] - \mathbb{E}[Y1_{|Y| \leq M} | \mathcal{F}_n]|] \leq \varepsilon.$$

En combinant les majorations on obtient pour, $m, n \geq n_0$,

$$\mathbb{E}[|X_m - X_n|] < 3\varepsilon$$

ce qui entraîne que la martingale (X_n) est de Cauchy dans L^1 et donc le résultat. \square

Corollaire 34. Soit Y une variable aléatoire intégrable sur un espace filtré $(\Omega, \mathcal{A}, \mathcal{F}, \mathbb{P})$. Alors la martingale définie par $X_n = \mathbb{E}[Y | \mathcal{F}_n]$ converge presque sûrement et dans L^1 vers $X_\infty = \mathbb{E}[Y | \mathcal{F}_\infty]$ où $\mathcal{F}_\infty = \sigma(\cup_{n \geq 1} \mathcal{F}_n)$.

DÉMONSTRATION. D'après le théorème précédent, il reste à montrer que $X_\infty = \mathbb{E}[Y | \mathcal{F}_\infty]$. Remarquons d'abord que X_∞ est \mathcal{F}_∞ -mesurable car les X_n le sont. Ensuite pour tout $n \in \mathbb{N}$ et $A \in \mathcal{F}_n$ on a :

$$\begin{aligned} \mathbb{E}[Y1_A] &= \mathbb{E}[\mathbb{E}[Y1_A | \mathcal{F}_n]] = \mathbb{E}[\mathbb{E}[Y | \mathcal{F}_n]1_A] \\ &= \mathbb{E}[X_n 1_A] = \mathbb{E}[\mathbb{E}[X_\infty | \mathcal{F}_n]1_A] = \mathbb{E}[\mathbb{E}[X_\infty 1_A | \mathcal{F}_n]] \\ &= \mathbb{E}[X_\infty 1_A]. \end{aligned}$$

Comme le résultat est vrai pour tout $A \in \sigma(\cup_{n \geq 1} \mathcal{F}_n)$ on en déduit qu'il est encore vrai pour tout $A \in \mathcal{F}_\infty$ (argument de classes monotones). D'où on obtient le corollaire. \square

Pour aller un peu plus loin nous avons besoin de parler de variables uniformément intégrables.

Définition 35. La famille $(X_i)_{i \in I}$ est uniformément intégrable (on parle aussi d'équi-intégrabilité) si

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

Remarquons qu'une famille uniformément intégrable est bornée dans L^1 . En effet on peut choisir a assez grand pour que

$$\sup_{i \in I} \mathbb{E}[|X_i| 1_{\{|X_i| > a\}}] \leq 1$$

et écrire ensuite

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

Attention la réciproque est fautive.

En utilisant le théorème de convergence dominée, on montre qu'une famille réduite à un singleton et plus généralement une famille finie est uniformément intégrable.

Rappelons le résultat suivant, liant la convergence en probabilité et la convergence L^1 .

Proposition 36. Soit $(X_n)_{n \in \mathbb{N}}$ une suite de variables aléatoires intégrables qui converge en probabilité vers X . Alors il y a équivalence entre

- (i) La suite $(X_n)_{n \in \mathbb{N}}$ converge vers X dans L^1 .
- (ii) La suite $(X_n)_{n \in \mathbb{N}}$ est uniformément intégrable.

Cette proposition nous donne une hypothèse sous laquelle la convergence en probabilité entraîne la convergence L^1 .

Appliquons ceci aux martingales.

Théorème 37.

Soit $(X_n)_{n \geq 0}$ une martingale. Les trois conditions suivantes sont équivalentes.

- (i) La suite X_n converge vers X_∞ p.s. et dans L^1 .
- (ii) Il existe une variable aléatoire Y intégrable telle que $X_n = \mathbb{E}[Y | \mathcal{F}_n]$ pour tout $n \in \mathbb{N}$.
- (iii) La suite $(X_n)_{n \in \mathbb{N}}$ est uniformément intégrable.

Rappelons qu'alors on a $X_n = \mathbb{E}[X_\infty | \mathcal{F}_n]$ et donc aussi $\mathbb{E}[X_\infty] = \mathbb{E}[X_0]$.

Le résultat reste vrai pour les surmartingales et les sous-martingales.