INSTITUT NATIONAL DES SCIENCES APPLIQUÉES TOULOUSE

# Probabilités Avancées

## 4e Année Math Appliquées

Polycopié de ____________________

# Table des matières

# 1. Introduction

Le terme **martingale** désigne une stratégie assurant des gains à un jeux de hasard :
`https://fr.wikipedia.org/wiki/Martingale`

En mathématiques, cette stratégie se traduit par **suite de variables aléatoires**, partageant une relation de dépendance. Cet objet constitue le second cas (avec les chaines de Markov) où l'on peut traiter le cas de dépendance entre variable aléatoires.

## 1.1 Un Exemple Introductif

On joue à pile ou face avec une pièce équilibrée. On note $X$ le résultat du lancé, avec la convention $X = 1$ si on gagne notre pari, et $X = -1$ si on le perd :

$$\mathbb{P}(X = 1) = \frac{1}{2} = \mathbb{P}(X = -1).$$

Avant chaque lancé, on place sa mise,

- si on gagne, on double la mise
- si on perd, on perd la mise.

**Stratégie :** Doubler sa mise tant qu'on perd. On arrête de miser dès qu'on a gagné. On peut dresser un tableau de Gains :

## 1.2 Modélisation Mathématique

— Chaque lancé est modélisé par une variable aléatoire $X_n$ :

$$\mathbb{P}(X_n = 1) = \frac{1}{2} = \mathbb{P}(X_n = -1).$$

— On note $W_n$ la fortune après $n$ lancés, $W_0 = 0$ par convention.
On va montrer que la fortune $(W_n)_{n\geq 0}$ est une **martingale**.

**Proposition 1.2.1** Pour chaque $n \geq 0$, on a $W_n \in \{1, -2^n + 1\}$.

*Démonstration.*
— Si on gagne a $n^e$ lancé, la fortune totale sera :

$$\begin{aligned} W_n &= 2^n - (1 + 2 + \cdots + 2^{n-1}) \\ &= 2^n - \frac{1-2^n}{1-2} = 1 \end{aligned}$$

— Si on perd au $n^e$ lancé, la fortune totale sera :

$$\begin{aligned} W_n &= -(1 + 2 + \cdots + 2^{n-1}) \\ &= -\frac{1-2^n}{1-2} = -2^n + 1 \end{aligned}$$

■

On note $\mathscr{F}_n = \sigma(X_k, k \leq n)$, c'est à dire, tous les événements impliquant $X_1, X_2, \ldots, X_n$. C'est donc toute **l'information disponible** à l'instant $n$ :

■ **Exemple 1.1** L'événement

$$\{\text{on a perdu les 3 premiers lancés}\}$$

s'écrit

$$\{X_1 = -1, X_2 = -1, X_3 = -1\},$$

donc appartient à $\mathscr{F}_3$ ■

**Proposition 1.2.2** Pour chaque $n \geq 0$, on a $\mathbb{E}(W_{n+1}|\mathscr{F}_n) = W_n$.

*Démonstration.*
— Si $W_n = 1$, on a gagné à l'étape $n$, et donc le jeux s'est arrêté, le gain n'a pas changé :

$$W_{n+1} = W_n,$$

donc en particulier, $\mathbb{E}(W_{n+1}|W_n = 1) = 1 = W_n$.

— Si $W_n = -2^n + 1$, comme $W_{n+1} \in \{1, -2^{n+1} + 1\}$, son espérance s'écrit :

$$\begin{aligned} &\mathbb{E}(W_{n+1}|W_n = -2^n + 1) \\ = \ &1 \times \mathbb{P}(W_{n+1} = 1|W_n = -2^n + 1) \\ &+ (-2^{n+1} + 1) \times \mathbb{P}(W_{n+1} = -2^{n+1} + 1|W_n = -2^n + 1) \end{aligned}$$

Maintenant, on a

$$\begin{aligned} \mathbb{P}(W_{n+1} = 1|W_n = -2^n + 1) &= \mathbb{P}(\text{gagné}) = \frac{1}{2} \\ \mathbb{P}(W_{n+1} = -2^{n+1} + 1|W_n = -2^n + 1) &= \mathbb{P}(\text{perdu}) = \frac{1}{2} \end{aligned}$$

Donc le calcul d'espérance donne :

$$\mathbb{E}(W_{n+1}|W_n = -2^n + 1) = \frac{1}{2}(1 - 2^{n+1} + 1) = -2^n + 1$$

**Conclusion :**
Dans les deux cas, on a bien $\mathbb{E}(W_{n+1}|\mathscr{F}_n) = W_n$. ■

Une martingale est un objet aléatoire visant à encapsuler une certaine notion d'information au cours du temps. La **filtration** $\mathscr{F}_n$ représente l'information disponible au temps présent. Une **martingale** est une suite de variables aléatoires telles que, étant donné les information à l'instant $n$, en moyenne, ne fait pas pire à l'instant d'après. En d'autre termes, le présent est la **meilleure approximation** du futur disponible avec nos informations.

## 1.3 Rappels sur l'espérance conditionnelle

Une filtration est une collection croissante de tribus. Elle représente l'accumulation de l'information au cours du temps. Dans la définition d'une martingale, on prend une espérance conditionnelle par rapport à une tribu, et non plus par rapport à un événement. Le chapitre suivant est dédié à la construction formelle de l'espérance conditionnelle. Dans ce paragraphe, on va voir l'espérance conditionnelle par rapport à une tribu comme une généralisation de ce qu'on connait.

### 1.3.1 Le cas discret

Si $X$ est une variable aléatoire discrète, et qu'on note $\mathscr{X}$ l'ensemble des valeurs possibles pour $X$, **la loi** de $X$ est alors la collection de nombres positifs :

$$\{\mathbb{P}(X=x),\ x\in\mathscr{X}\},$$

qu'on pouvait par exemple la représenter dans un tableau :

| $X$ | $x_1$ | $x_2$ | $\cdots$ |
|---|---|---|---|
| $\mathbb{P}(X=x)$ | $p_1$ | $p_2$ | $\cdots$ |

On calcule alors l'espérance avec la formule :

$$\mathbb{E}[h(X)] = \sum_{x\in\mathscr{X}} h(x)\mathbb{P}(X=x).$$

Si on considère maintenant le couple $(X,Y)$ où $Y$ est une variable aléatoire discrète, la formule est similaire :

$$\mathbb{E}[h(X,Y)] = \sum_{x\in\mathscr{X}}\sum_{y\in\mathscr{Y}} h(x,y)\mathbb{P}(X=x,Y=y)$$

où on remarque qu'on somme contre les nombres $\mathbb{P}(X=x,Y=y)$, qui représentent **la loi** de $(X,Y)$ :

$$\{\mathbb{P}(X=x,Y=y),\ (x,y)\in\mathscr{X}\times\mathscr{Y}\}.$$

On peut également représenter ces nombres dans un tableau à double entrée (tableau de contingence), mais cette représentation se perd pour plus de 3 variables. Rappelons maintenant qu'on peut définir les **probabilités conditionnelles** de la façon suivante :

$$\mathbb{P}(X=x|Y=y) = \frac{\mathbb{P}(X=x,Y=y)}{\mathbb{P}(Y=y)},\ \forall y\in\mathscr{Y}.$$

Donc à $y\in\mathscr{Y}$ fixé, la **loi conditionnelle de** $X$ **sachant** $Y=y$ est la collection de nombres positifs

$$\{\mathbb{P}(X=x|Y=y)\ ;\ x\in\mathscr{X}\}.$$

Alors, **l'espérance conditionnelle** est simplement l'espérance par rapport à cette loi :

$$\mathbb{E}[h(X)|Y=y] = \sum_{x\in\mathscr{X}} h(x)\mathbb{P}(X=x|Y=y).$$

**Remarque** Remarquez que le nombre $\mathbb{E}[h(X)|Y=y]$ dépend de la valeur de $Y=y$. Ainsi, $\mathbb{E}[h(X)|Y=y]$ est une certaine fonction, disons $h$ de $y$. Pour chaque valeur de $Y=y$, on peut calculer le nombre $h(y)=\mathbb{E}[h(X)|Y=y]$, et on notera par abus

$$h(Y)=\mathbb{E}(X|Y).$$

■ **Exemple 1.2** En Chaines de Markov, la propriété de Markov disait :

$$\mathbb{P}(X_{n+1}=x_{n+1}|\, X_n=x_n,\dots X_0=x_0)=\mathbb{P}(X_{n+1}=x_{n+1}|\, X_n=x_n).$$

Donc pour une chaine de Markov, la loi conditionnellement à tout le passé est la loi conditionnelle à l'instant présent. ■

### 1.3.2 Le cas des variables continues

Dans le cas particulier où $(X,Y)$ sont à densité par rapport à la mesure de Lebesgue, tout ce qui est dit dans le paragraphe précédent peut être reformulé avec les densités.

- Si on note $f_{(X,Y)}$ la densité jointe du couple $(X,Y)$, et $f_X$ et $f_Y$ les densités marginales de $X$ et de $Y$.
- On définit la **densité conditionnelle** avec la formule :

$$f(x|Y=y)=\frac{f_{(X,Y)}(x,y)}{f_Y(y)},\ \forall y\in\mathscr{Y}.$$

- Alors, **l'espérance conditionnelle** est l'espérance par rapport à cette densité :

$$\mathbb{E}[h(X)|Y=y]=\int_{\mathbb{R}} h(x)f(x|Y=y)dx$$

On remarque toujours que $\mathbb{E}[h(X)|Y=y]$ est encore une fonction de $y$.

### 1.3.3 Le cas général

Quand les variables ne sont plus ni discrètes, ni continues, **la loi** de $(X,Y)$ est alors la collection de nombres positifs :

$$\{\mathbb{P}(X\in A,Y\in B),\ \ A,B\in\mathscr{B}(\mathbb{R})\}.$$

où $\mathscr{B}(\mathbb{R})$ désigne les Boréliens de $\mathbb{R}$. Alors, en général, pour calculer l'espérance conditionnelle de $h(X)$ sachant $Y$, il ne suffit plus de connaitre

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

il faudrait calculer pour tous $B\in\mathscr{B}(\mathbb{R})$ :

$$\mathbb{E}[h(X)|Y\in B]$$

**Remarque** La tribu Borélienne $\mathbb{R}$ qu'on note $\mathscr{B}(\mathbb{R})$ est la tribu engendré par les ouverts de $\mathbb{R}$. Comprenez donc qu'il faudrait calculer $\mathbb{E}[h(X)|Y\in B]$ pour tout ouvert $B$ de $\mathbb{R}$, ce qui fait quand même beaucoup de valeurs !

Mais on va voir par la suite que l'ensemble

$$\sigma(Y):=\{Y\in B\,;\,B\in\mathscr{B}(\mathbb{R})\}$$

est ce qu'on appelle **la tribu engendré** par $Y$. On a donc besoin de

$$\mathbb{E}[h(X)|\sigma(Y)].$$

Il faut donc qu'on apprenne à définir une espérance conditionnelle par rapport à une tribu. On donnera dans le chapitre suivant la construction formelle de l'espérance conditionnelle, ainsi qu'une liste bien utile de propriétés à connaitre et savoir manipuler pour calculer ces espérances conditionnelles.

■ **Exemple 1.3** Reprenons le cas des chaines de Markov. Partons de la **propriété de Markov** :

$$\mathbb{P}(X_{n+1}=x_{n+1}|\, X_n=x_n,\ldots X_0=x_0)=\mathbb{P}(X_{n+1}=x_{n+1}|\, X_n=x_n).$$

À partir des nombres $\mathbb{P}(X_{n+1}=x_{n+1}|\, X_n=x_n,\ldots X_0=x_0)$, on peut calculer l'espérance conditionnelle :

$$\mathbb{E}[\varphi(X_{n+1})|X_n=x_n,\ldots X_0=x_0]=\mathbb{E}[\varphi(X_{n+1})|X_n=x_n].$$

On voit encore que dans ce cas, le membre de droite de l'égalité définit une certaine fonction $g(x_n)$ de la variable $x_n$. Mais si on peut calculer cette quantité pour tous les $x_0,x_1,\ldots,x_n$ possibles, en fait on a calculé

$$\mathbb{E}[\varphi(X_{n+1})|\mathscr{F}_n]=g(X_n),$$

puisque $\mathscr{F}_n=\sigma(X_1,\ldots,X_n)$ est *engendrée* par les événements $\{X_n=x_n,\ldots X_0=x_0\}$. Cette dernière égalité est la définition de la propriété de Markov dans le cas général. En fait, on notera

$$\mathbb{E}[\varphi(X_{n+1})|\mathscr{F}_n]=P_{n+1,n}(\varphi)(X_n)$$

pour faire apparaitre la dépendance de la fonction $\varphi$, et $P_{n+1,n}$ est le semi-groupe associé à $(X_n)_{n\in\mathbb{N}}$.
■

## 1.4 Le teaser du cours

Les résultats principaux de ce cours sont les suivants :

— Théorèmes d'arrêt :
Qui permettent de calculer la valeur de $\mathbb{E}(W_T)$, où $T$ est un **temps d'arrêt**, un instant aléatoire.

— Théorèmes de convergence de martingales :
Une collection de théorèmes établissant la convergence de certaines suites de variables aléatoires. Ces convergences sont particulièrement utiles en **optimisation stochastique**, machine learning etc.

— Inégalité maximales
Qui sont des généralisation de **l'inégalité de Markov**.

— Loi des grand nombres et théorèmes central limite
Qui donnent une ouverture sur les **statistiques, la finance et l'assurance**.

# 2. Rappels sur l'espérance conditionnelle

## 2.1 Quelques éléments sur les tribus

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

**Définition 2.1.1** Soit $\Omega$ un ensemble et $\mathscr{A}$ une famille de parties de $\Omega$. On dit que $\mathscr{A}$ est une tribu (sur $\Omega$) si :

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

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

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

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

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

**Définition 2.1.2** Étant donnée une v.a. $X$ à valeurs dans un espace mesurable $(E, \mathscr{E})$, on appelle tribu engendrée par $X$, et on la note $\sigma(X)$, la sous-tribu de $\mathscr{A}$ engendrée par l'ensemble des images réciproques de $X$. Autrement dit,

$$\begin{aligned}
\sigma(X) &:= \sigma\left(\{X^{-1}(B) : B \in \mathscr{E}\}\right) \\
&= \sigma\left(\{\{\omega \in \Omega : X(\omega) \in B\} ; B \in \mathscr{E}\}\right) \\
&= \sigma\left(\{\{X \in B\} ; B \in \mathscr{E}\}\right),
\end{aligned}$$

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

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

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

**Lemme 2.1.1** (Doob)
Étant données des v.a.r. $X_1,\ldots,X_n$, une autre v.a.r. $Y$ est $\sigma(X_1,\ldots,X_n)$-mesurable si et seulement s'il exite une fonction borélienne $h:\mathbb{R}^n\to\mathbb{R}$ telle que $Y=h(X_1,\ldots,X_n)$.

## 2.2 Espérance conditionnelle

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

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

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

**Définition 2.2.1** Soit $X$ une v.a. appartenant à l'espace $L^2$ (respectivement à $L^1$) et soit $\mathscr{F}$ une sous-tribu quelconque de $\mathscr{A}$. Alors il existe une v.a. $Y$ dans $L^2(\mathscr{F})$ (resp. dans $L^1(\mathscr{F})$), unique à équivalence près, telle que pour toute v.a. $Z\in L^2(\mathscr{F})$ (resp. tout événement $A\in\mathscr{F}$),

$$\mathbb{E}[XZ] = \mathbb{E}[YZ] \quad (\text{resp.} \quad \mathbb{E}[X1_A] = \mathbb{E}[Y1_A]).$$

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

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

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

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

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

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

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

c'est-à-dire le résultat désiré. ■

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

**Exercice 2.1**

1. On lance deux pièces, et on définit les variables aléatoires :
$$\begin{aligned} X &= 1 \text{ si le 1er lancer est face} \\ Y &= \text{le nombre de face sur les deux lancers} \end{aligned}$$
Combien vaut $\mathbb{E}(X|Y)$ ?
2. Soit $Z$ une variable aléatoire de loi $\mathscr{E}(3)$.
À partir de cette variable aléatoire, on définit $W$ de loi $\mathscr{E}(Z)$.
Montrer que $\mathbb{E}(W|Z) = \frac{1}{Z}$.

■

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

**Proposition 2.2.1** L'espérance conditionnelle vérifie les propriétés suivantes.

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

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

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

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

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

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

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

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

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

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

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

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

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

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

**Proposition 2.2.2** Soit $X$ une v.a. intégrable. Alors l'espérance conditionnelle vérifie les propriétés suivantes.

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

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

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

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

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

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

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

$$\mathbb{E}[h(X,Y) \mid \mathscr{F}] = H(X), \qquad \text{avec} \qquad H(x) = \mathbb{E}[h(x,Y)].$$

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

# 3. Martingales

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

## 3.1 Définition et premières propriétés

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

**Définition 3.1.1** Un processus stochastique (ou aléatoire) à temps discret est une suite de v.a., non nécessairement indépendantes.

Dans la suite on parlera simplement de processus.

**Définition 3.1.2** Une suite $(\mathscr{F}_n)_{n\in\mathbb{N}}$ de sous-tribus de $\mathscr{A}$ est appelée une filtration de l'espace $(\Omega, \mathscr{A}, \mathbb{P})$ si

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

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

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

**Définition 3.1.3** Un processus $(M_n)_{n\in\mathbb{N}}$ est adapté à une filtration $(\mathscr{F}_n)_{n\in\mathbb{N}}$ si $M_n$ est $\mathscr{F}_n$-mesurable pour tout $n \in \mathbb{N}$.
Si $M_n$ est $\mathscr{F}_{n-1}$-mesurable pour tout $n \in \mathbb{N}^*$, alors $(M_n)_{n\in\mathbb{N}}$ est dit prévisible.

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

**Définition 3.1.4** Considérons un processus $(M_n)_{n\in\mathbb{N}}$ intégrable (tous ses éléments le sont) et adapté à une filtration $(\mathscr{F}_n)_{n\in\mathbb{N}}$. On dit que le processus $(M_n)_{n\in\mathbb{N}}$ est (pour $(\mathscr{F}_n)_{n\in\mathbb{N}}$) une

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

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

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

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

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

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

**Exercice 3.1** Soit $(X_n)_{n\in\mathbb{N}^*}$ une suite de v.a. indépendantes et intégrables telles que pour tout $n \in \mathbb{N}^*$, on a $\mathbb{E}[X_n] = 0$. Montrer que le processus $(S_n)_{n\in\mathbb{N}^*}$ défini par

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

est une martingale pour la filtration engendrée par la suite $(X_n)_{n\in\mathbb{N}^*}$. Notons que nous n'avons pas supposé que les v.a. étaient de même loi. ■

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

(i) Avec les notations de l'exercice 3.1, $(S_n)_{n\in\mathbb{N}^*}$ est une surmartingale (resp., sous-martingale) si $\mathbb{E}[X_n] \leq 0$ (resp. $\mathbb{E}[X_n] \geq 0$).

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

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

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

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

Le lecteur assidu s'essaiera à prouver les énoncés ci-dessus pour tester sa compréhension !

**Proposition 3.1.1** Soit $(M_n)_{n\in\mathbb{N}}$ un processus intégrable et adapté à une filtration $(\mathscr{F}_n)_{n\in\mathbb{N}}$. Alors c'est une martingale si et seulement si pour tous $n, p \in \mathbb{N}$,

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

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

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

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

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

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

**Proposition 3.1.2**

(i) L'opposé d'une surmartingale est une sous-martingale, et inversement.

(ii) La somme de deux surmartingales (resp. sous-martingales) est une surmartingale (resp. sous-martingale).

(iii) L'ensemble des martingales forme un espace vectoriel sur $\mathbb{R}$.

(iv) Si l'on compose une martingale avec une fonction convexe, on obtient une sous-martingale (sous l'hypothèse d'intégrabilité).

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

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

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

**Théorème 3.1.3** (Décomposition de Doob d'une sous-martingale)
Soit $(X_n)_{n\in\mathbb{N}}$ une sous-martingale. Alors, il existe une unique martingale $(M_n)_{n\in\mathbb{N}}$ et un unique processus croissant prévisible $(A_n)_{n\in\mathbb{N}}$, intégrable et issu de 0, tels que

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

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

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

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

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

on a que

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

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

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

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

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

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

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

**Remarque** On peut montrer à l'aide de la décomposition de Doob qu'une sous-martingale (ou surmartingale) d'espérance constante est en fait une martingale. Ce n'est pas la seule façon de prouver ce résultat.

**Exercice 3.2** On considère de $(Y_i)_{i\geq 0}$ une suite de variables *i.i.d.* de loi

$$\mathbb{P}(Y_i = -1) = \mathbb{P}(Y_i = 1) = \frac{1}{2}.$$

1. On pose $S_n = \sum_{k=1}^{n} Y_k$, rappeler pourquoi $(|S_n|)_{n\geq 0}$ est une sous-martingale.
2. Montrer que sa décomposition de Doob est $M_n + A_n$

$$M_n = \sum_{k=1}^{n} \operatorname{sgn}(S_{k-1})Y_k, \text{ et } A_n = \sum_{k=1}^{n} \mathbf{1}_{S_{k-1}=0}.$$

■

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

**Définition 3.1.5** (Transformation prévisible d'une martingale)
Soit $(M_n)_{n\in\mathbb{N}}$ une martingale et soit $(A_n)_{n\in\mathbb{N}^*}$ un processus prévisible. La transformation prévisible de la martingale $(M_n)_{n\in\mathbb{N}}$ par le processus $(A_n)_{n\in\mathbb{N}^*}$, notée $A\cdot M$, est définie par le processus donné par $(A\cdot M)_0 = M_0$ et

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

Si les v.a. $A_n$ sont bornées, alors $A\cdot M$ est une martingale.

**Remarque** Pour alléger les notations, il nous arrivera de poser $A\cdot M = (\widetilde{M}_n)_{n\in\mathbb{N}}$.

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

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

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

■ **Exemple 3.1** On se place dans le même cadre que l'exercice 3.2. Pour rappel, les $(Y_i)_{i\geq 0}$ sont une suite de variables *i.i.d.* telles que :

$$\mathbb{P}(Y_i = -1) = \mathbb{P}(Y_i = 1) = \frac{1}{2}.$$

D'après l'exercice 3.1, on sait que $S_n = \sum_{k=1}^{n} Y_k$, est une martingale. On définit le processus prévisible borné

$$A_n = \mathrm{sgn}(S_{n-1}).$$

On a alors l'identité suivante pour la transformation prévisible $A\cdot S$ :

$$(A\cdot S)_n = \sum_{k=1}^{n} \mathrm{sgn}(S_{k-1})\left(S_k - S_{k-1}\right) = \sum_{k=1}^{n} \mathrm{sgn}(S_{k-1}) Y_k.$$

Or, dans l'exercice 3.2, on a vu que $A\cdot S$ est en fait la partie martingale de la décomposition de Doob de la sous-martingale $(|S_n|)_{n\in\mathbb{N}}$.

■

## 3.2 Martingales de carré intégrable

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

**Définition 3.2.1** Un processus $(X_n)_{n\in\mathbb{N}}$ est dit de carré intégrable, ou simplement dans $L^2$, si $\mathbb{E}[X_n^2] < +\infty$ pour tout $n \in \mathbb{N}$.

**Remarque** Plus tard, nous utiliserons la notion de **borné dans** $L^2$ :

$$\sup_{n\geq 0} \mathbb{E}[|X_n|^2] < +\infty$$

Clairement, borné dans $L^2 \Rightarrow$ de carré intégrable.

**Exercice 3.3** Soient $X_i$ des variables aléatoires *i.i.d.* telles que

$$\mathbb{E}[|X_i|^2] = \sigma^2.$$

Montrer qu'alors, la marche aléatoire $S_n = \sum_{i=1}^n X_i$, est de **carré intégrable**. ■

Le résultat suivant est la version $L^2$ du théorème de transformation prévisible.

**Théorème 3.2.1** (Transformation prévisible d'une martingale, version $L^2$)
Soit $(A \cdot M_n)_{n\in\mathbb{N}}$ la transformation prévisible d'une martingale de carré intégrable $(M_n)_{n\in\mathbb{N}}$ par le processus prévisible $(A_n)_{n\in\mathbb{N}^*}$, supposé borné. Alors $(A \cdot M_n)_{n\in\mathbb{N}}$ est une martingale de carré intégrable.

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

Ce théorème permet de prouver facilement que certaines variables aléatoires sont des martingales.

**Exercice 3.4** Soit $(X_i)_{i\geq 0}$ une suite de variables aléatoires *i.i.d.*, telle que

$$\mathbb{P}(X_i = 1) = \mathbb{P}(X_i = -1) = \frac{1}{2}.$$

Montrer que :

$$M_n = \sum_{k=1}^n X_{k-1}X_k \text{ est une martingale de carré intégrable.}$$

■

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

**Définition 3.2.2** Soit $(M_n)_{n\in\mathbb{N}}$ une martingale de carré intégrable. Alors il existe un unique processus croissant et prévisible, intégrable et issu de 0, appelé la variation quadratique de $(M_n)_{n\in\mathbb{N}}$ et noté $([M,M]_n)_{n\in\mathbb{N}}$, tel que le processus $(M_n^2-[M,M]_n)_{n\in\mathbb{N}}$ soit une martingale.
La variation quadratique admet la représentation suivante : $[M,M]_0=0$ et pour tout $n\in\mathbb{N}^*$,

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

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

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

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

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

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

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

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

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

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

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

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

et en particulier

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

une identité nous permettant souvent de calculer la variance d'une martingale de carré intégrable lorsque l'on connaît explicitement sa variation quadratique.

**Exercice 3.5** Si $(X_i)_{i\geq 1}$ sont $i.i.d.$ telles que $\mathbb{E}(X_i) = 0$ et $\mathbb{E}(X_i^2) = \sigma^2$, quelle est la variation quadratique de

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

■

Enfin, le dernier résultat de cette partie est une formule exprimant la variation quadratique de la transformation prévisible en fonction de celle de la martingale originelle. Nous nous servirons de ce résultat dans le prochain chapitre pour démontrer la loi des grands nombres pour les martingales.

**Théorème 3.2.2** (Variation quadratique d'une transformation prévisible)
Soit $(A\cdot M_n)_{n\in\mathbb{N}}$ la transformation prévisible d'une martingale de carré intégrable $(M_n)_{n\in\mathbb{N}}$ par le processus prévisible $(A_n)_{n\in\mathbb{N}^*}$, supposé borné. Alors on a la formule suivante : pour tout $n \in \mathbb{N}^*$,

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

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

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

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

**Exercice 3.6** On considère de nouveaux $(Y_i)_{i\geq 0}$ une suite de variables *i.i.d.* de loi :

$$\mathbb{P}(Y_i = -1) = \mathbb{P}(Y_i = 1) = \frac{1}{2}.$$

On note $\mathscr{F}_n = \sigma(Y_k, k \leq n)$ et $S_n = Y_1 + \cdots + Y_n$ et on pose

$$M_n = \sum_{k=1}^{n} \operatorname{sgn}(S_{k-1}) Y_k, \text{ où}$$

1. Rappeler l'expression de $\left([S,S]_n\right)_{n\geq 0}$
2. Montrer que $(M_n)_{n\geq 0}$ est une martingale et calculer $\left([M,M]_n\right)_{n\geq 0}$.

**Remarque** Cet exercice à mettre en lien avec la décomposition de Doob de $(|S_n|)_{n\geq 0}$.

## 3.3 Théorème d'arrêt

Dans cette partie, nous allons établir le théorème d'arrêt pour les martingales, résultat très utile en pratique, par exemple lorsque l'on considère le problème de la ruine du joueur. L'objectif est de considérer notre martingale à un instant aléatoire. Par exemple, pour la stratégie au casino "doubler jusqu'à la victoire", $W_T = 1$ par construction. Bien évidemment, on ne pourra pas considérer n'importe quel temps aléatoire, ni n'importe quel processus stochastique. L'objectif de ce paragraphe est de définir les objects pour lesquels cette opération est possible. Commençons par un exemple introductif.

■ **Exemple 3.2** On lance un dé équilibré, et on somme les résultats jusqu'au 1er 6 obtenu. En moyenne, combien vaut la somme ?

On note $\tau$ le premier temps où le dé donne 6, et $S_n$ la somme des résultats jusqu'au temps $n$. On peut montrer que

$$\mathbb{E}\left(\sum_{k=1}^{\tau} X_k\right) = \mathbb{E}(\tau)\mathbb{E}(X_1).$$

■

L'exemple précédent, bien que très simple peut s'utiliser en assurance. C'est identité est l'égalité de Wald. Elle est utilisée en assurance pour calculer le prix des contrats :

- Les $X_k$ peuvent représenter la valeur des biens assurées à rembourser (claims en anglais)
- Le temps aléatoire $\tau$ est le temps où survient un sinistre (vol de voiture, cambriolage etc.).

Alors, chaque mois, les assurés paient une cotisation (premium en anglais), que fixe l'assureur pour que la variable aléatoire soit positive :

$$L_T = \ell + cT + \sum_{t=1}^{T} X_t$$

On peut alors choisir $\ell$ et $c$ (qui dépendront du nombre total d'assurés) pour qu'en moyenne, $L_T$ soit positive.

Avant tout, définissons la tribu correspondant à l'ensemble des informations disponibles. C'est la tribu engendrée par toutes les tribus $\mathscr{F}_n$, à savoir

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

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

**Définition 3.3.1** Une v.a. entière $\tau$, pouvant prendre la valeur $+\infty$, est un temps d'arrêt pour la filtration $(\mathscr{F}_n)_{n\in\mathbb{N}}$ si pour tout $n \in \mathbb{N}$,

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

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

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

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

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

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

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

**Proposition 3.3.1** Soient $\tau_1$ et $\tau_2$ deux temps d'arrêt. Alors $\tau_1 \wedge \tau_2$, $\tau_1 \vee \tau_2$ et $\tau_1 + \tau_2$ sont des temps d'arrêt.

Étant donné un temps d'arrêt $\tau$, on définit la famille d'événements suivante :

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

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

**Proposition 3.3.2** On a les propriétés suivantes :

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

*Démonstration.* $(i)$ Si $\tau$ est constant alors on a $\{\tau \leq n\} = \Omega$ ou $\emptyset$ selon que $n \geq k$ ou $n < k$. Dans les deux cas l'événement $\{\tau \leq n\}$ appartient à la tribu $\mathscr{F}_n$, d'où $\tau$ est un temps d'arrêt.
À présent, montrons par double inclusion que $\mathscr{F}_\tau = \mathscr{F}_k$. Si $A \in \mathscr{F}_k$ alors on a :

- si $n \geq k$ alors $A \cap \{\tau \leq n\} = A \in \mathscr{F}_k \subset \mathscr{F}_n$ ;
- si $n < k$ alors $A \cap \{\tau \leq n\} = \emptyset \in \mathscr{F}_n$.

D'où $A \in \mathscr{F}_\tau$. Quant à l'autre inclusion, on note que si $A \in \mathscr{F}_\tau$ alors $A \cap \{\tau \leq n\} \in \mathscr{F}_n$ pour tout $n \in \mathbb{N}$, en particulier pour $n \geq k$ on obtient que $A \in \mathscr{F}_k$.

$(ii)$ Le temps d'arrêt $\tau$ est $\mathscr{F}_\tau$-mesurable si pour tout $k \in \mathbb{N}$, on a $\{\tau = k\} \in \mathscr{F}_\tau$, c'est-à-dire si pour tous $k, n \in \mathbb{N}$, on a $\{\tau = k\} \cap \{\tau = n\} \in \mathscr{F}_n$, ce que l'on démontre sans difficulté après avoir considéré les cas $n \neq k$ et $n = k$.

$(iii)$ Soit $A \in \mathscr{F}_{\tau_1}$. Comme $\tau_1 \leq \tau_2$, on a pour tout $n \in \mathbb{N}$,

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

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

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

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

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

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

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

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

**Théorème 3.3.3** (Théorème d'arrêt pour les martingales)
Soit $(M_n)_{n\in\mathbb{N}}$ une martingale et soit $\tau$ un temps d'arrêt. Alors le processus arrêté $(M_{n\wedge\tau})_{n\in\mathbb{N}}$ est aussi une martingale. De plus, si $\tau_1\leq\tau_2$ sont deux temps d'arrêt bornés, alors on a

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

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

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

ce qui entraîne que la suite $(M_{n\wedge\tau})_{n\in\mathbb{N}}$ est la transformation prévisible de la martingale $(M_n)_{n\in\mathbb{N}}$ par rapport à $(A_n)_{n\in\mathbb{N}^*}$. Ainsi, c'est une martingale par le théorème 3.1.5 et la première affirmation du théorème est démontrée. Le lecteur averti aura noté que cette propriété de martingale peut être démontrée directement : montrer d'abord l'intégrabilité, ensuite le fait que $(M_{n\wedge\tau})_{n\in\mathbb{N}}$ est adapté à la filtration sous-jacente, puis la propriété de martingale en remarquant que

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

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

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

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

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

Enfin, si $\tau$ désigne $\tau_1$ ou $\tau_2$, comme p.s. $M_{n\wedge\tau} \underset{n\to+\infty}{\longrightarrow} M_\tau$ et que

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

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

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

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

L'exercise suivant est un exercice très classique d'application du théorème d'arrêt. Attention à bien vérifier les hypothèses du théorème !

**Exercice 3.7** Deux joueurs A et B jouent un nombre illimité de parties indépendantes, l'enjeu étant de 1 euro par partie : si A gagne la n-ième partie, il prend un euro à B et réciproquement. La probabilité que le joueur A gagne ou perde une partie donnée est $1/2$. Pour modéliser ce jeu, on considère la suite des gains de A, représentée par une suite $(Y_n)_{n\in\mathbb{N}^*}$ de v.a. indépendantes à valeurs dans $\{-1,1\}$, de loi commune de Rademacher de paramètre $1/2$ donnée par

$$\mathbb{P}(Y=1) = 1 - \mathbb{P}(Y=-1) = \frac{1}{2}.$$

On suppose que la fortune initiale de A est de $a$ euros et celle de $B$ de $b$ euros, où $a,b \in \mathbb{N}^*$. Considérons le processus $(M_n)_{n\in\mathbb{N}}$ défini par $M_0 = 0$ et

$$M_n = \sum_{k=1}^{n} Y_k, \quad n \in \mathbb{N},$$

modélisant le gain de A au cours des $n$ premières parties. Après la n-ième partie, la fortune de A est de $a + M_n$ euros, et celle de B de $b - M_n$ euros, tant que ces quantités restent positives. Le jeu s'arrête dès que l'un des deux joueurs est ruiné.
Soit $T$ la v.a. correspondant à l'instant où le jeu s'arrête, c'est-à-dire

$$T = \inf\{n \geq 1 : M_n = -a \text{ ou } M_n = b\} \in \mathbb{N}^* \cup \{+\infty\}.$$

On note également $\mathscr{F}_0 = \{\emptyset, \Omega\}$ et $\mathscr{F}_n = \sigma(Y_1,\ldots,Y_n)$ pour tout $n \in \mathbb{N}^*$. Tout au long de cet exercice, on admet que $T$ est une v.a. p.s. finie, c'est-à-dire que $\mathbb{P}(T<\infty) = 1$.

1. Montrez que $T$ est un temps d'arrêt.
2. Quelles sont les valeurs possibles de $M_T$ ?
3. Montrez que $\mathbb{E}[M_T] = 0$.
4. Déduisez-en la valeur des probabilités $\mathbb{P}(M_T = -a)$ et $\mathbb{P}(M_T = b)$. À quoi correspondent-elles ?
5. Montrez que le processus $(Z_n)_{n\in\mathbb{N}}$ défini par $Z_n = M_n^2 - n$ est une martingale.
6. Déduisez-en la valeur de l'espérance $\mathbb{E}[T]$.

■

## 3.4 Exercices du Chapitre 3

**Exercice 3.8** Soit $(X_n)_{n\in\mathbb{N}^*}$ une suite de v.a. indépendantes et intégrables et $(\mathscr{F}_n)_{n\in\mathbb{N}^*}$ sa filtration naturelle : $\mathscr{F}_n = \sigma(X_k, n \le n)$. Notons le processus $(S_n)_{n\in\mathbb{N}^*}$ défini par $S_n := \sum_{i=1}^n X_i$.

1. Montrez que $(S_n)_{n\in\mathbb{N}^*}$ est une martingale (resp. surmartingale, sous-martingale) pour la filtration $(\mathscr{F}_n)_{n\in\mathbb{N}^*}$ lorsque pour tout $n\in\mathbb{N}^*$, on a $\mathbb{E}[X_n] = 0$ (resp. $\mathbb{E}[X_n] \le 0$, $\mathbb{E}[X_n] \ge 0$).
2. Supposons que $\sigma^2 = \mathbb{E}(X_n^2)$, montrer que $n\in\mathbb{N}^*$ par $M_n := S_n^2 - n\sigma^2$, est une martingale.
3. Montrez que le processus donné pour tout $n\in\mathbb{N}$ par $M_n := \mathbb{E}[X \mid \mathscr{F}_n]$, où $X$ est une v.a. intégrable et $(\mathscr{F}_n)_{n\in\mathbb{N}}$ est une filtration quelconque, est une martingale.

**Exercice 3.9** Soit $(X_n)_{n\in\mathbb{N}^*}$ une suite de v.a. i.i.d., supposées intégrables, d'espérance commune $m$, et notons $(\mathscr{F}_n)_{n\in\mathbb{N}^*}$ sa filtration naturelle. Que peut-on dire du processus $(Y_n)_{n\in\mathbb{N}^*}$ défini par

$$Y_n = \sum_{i=1}^n iX_i - \frac{n(n+1)}{2} \times m, \quad n\in\mathbb{N}^* ?$$

**Exercice 3.10** Soit $(X_n)_{n\in\mathbb{N}^*}$ une suite de v.a. indépendantes et intégrables et $(\mathscr{F}_n)_{n\in\mathbb{N}^*}$ sa filtration naturelle : $\mathscr{F}_n = \sigma(X_k, n \le n)$. Notons le processus $(S_n)_{n\in\mathbb{N}^*}$ défini par $S_n := \sum_{i=1}^n X_i$.

1. Montrez que $(S_n)_{n\in\mathbb{N}^*}$ est une martingale (resp. surmartingale, sous-martingale) pour la filtration $(\mathscr{F}_n)_{n\in\mathbb{N}^*}$ lorsque pour tout $n\in\mathbb{N}^*$, on a $\mathbb{E}[X_n] = 0$ (resp. $\mathbb{E}[X_n] \le 0$, $\mathbb{E}[X_n] \ge 0$).
2. Supposons que $\sigma^2 = \mathbb{E}(X_n^2)$, montrer que $n\in\mathbb{N}^*$ par $M_n := S_n^2 - n\sigma^2$, est une martingale.
3. Montrez que le processus donné pour tout $n\in\mathbb{N}$ par $M_n := \mathbb{E}[X \mid \mathscr{F}_n]$, où $X$ est une v.a. intégrable et $(\mathscr{F}_n)_{n\in\mathbb{N}}$ est une filtration quelconque, est une martingale.

**Exercice 3.11** Soit $(S_n)_{n\in\mathbb{N}^*}$ la marche aléatoire symétrique simple sur $\mathbb{Z}$, i.e. : $S_n := \sum_{i=1}^n X_i$, où $X_i = \pm 1$ avec probabilité $\frac{1}{2}$. On note $(\mathscr{F}_n)_{n\in\mathbb{N}^*}$ sa filtration naturelle : $\mathscr{F}_n = \sigma(X_k, n \le n)$. Parmi les variables aléatoires suivantes, lesquelles sont des temps d'arrêts ?

1. Pour $M\in\mathbb{Z}$ fixé, $T_1 = \inf\{n \ge 0 | \ S_n = M\}$.
2. Pour $k\in\mathbb{N}$ fixé, $T_2 = \inf\{n \ge k | \ S_n = S_{n-k}\}$.
3. Pour $k\in\mathbb{N}$ fixé, $T_3 = \inf\{n \ge k | \ S_n = S_{n+k}\}$.
4. $T_4 = \inf\{n \ge T_1 | \ S_n = 0\}$.
5. Pour $N\in\mathbb{N}$ fixé, $T_5 = \sup\{n \in [\![0,N]\!] \ |S_n = 0\}$.
6. Pour $N\in\mathbb{N}$ fixé, $T_6 = \inf\{n \in [\![0,N]\!] \ |\forall m \in [\![0,N]\!], S_m \le S_n\}$.

**Exercice 3.12** Soit $T$ un temps d'arrêt pour une filtration $(\mathscr{F}_n)_{n\ge 0}$. On suppose qu'il existe $\varepsilon \in ]0,1[$ et $n_0\in\mathbb{N}^*$ tel que pour tous $n \ge 0$, on a presque surement :

$$\mathbb{P}(T < n + n_0 | \mathscr{F}_n) > \varepsilon.$$

Montrer que $T$ est fini presque surement et que $\mathbb{E}[T] < +\infty$.

**Exercice 3.13** Soit $(M_n)_{n\in\mathbb{N}}$ une martingale de carré intégrable, associée à une filtration $(\mathscr{F}_n)_{n\in\mathbb{N}}$. Montrez la formule suivante : pour tout $n\in\mathbb{N}^*$,

$$\mathrm{Var}(M_n) = \mathrm{Var}(M_0) + \sum_{i=1}^n \mathrm{Var}(M_i - M_{i-1}).$$

**Exercice 3.14** On considère $(Y_i)_{i\ge 0}$ une suite de variables *i.i.d.* de loi :

$$\mathbb{P}(Y_i = -1) = \mathbb{P}(Y_i = 1) = \frac{1}{2}.$$

On note $\mathscr{F}_n = \sigma(Y_k, k \le n)$ et $S_n = Y_1 + \cdots + Y_n$ et on pose

$$M_n = \sum_{i=1}^n \mathrm{sgn}(S_{k-1}) Y_k, \text{ où}$$

1. Rappeler l'expression de $\Big([S,S]_n\Big)_{n\geq 0}$
2. Montrer que $(M_n)_{n\geq 0}$ est une martingale et calculer $\Big([M,M]_n\Big)_{n\geq 0}$.
3. Quelle est la décomposition de Doob de $(|S_n|)_{n\geq 0}$ ?

**Exercice 3.15** On considère une suite de variables aléatoires $(Y_n)_{n\in\mathbb{N}}$ i.i.d. et positives, de densité commune

$$f(x) = \frac{\alpha}{x^{\alpha+1}} 1_{x\geq 1}, \quad \alpha > 1.$$

On note $(\mathscr{F}_n)_{n\in\mathbb{N}}$ sa filtration naturelle. Dans cet exercice, on s'intéresse à la suite $(X_n)_{n\in\mathbb{N}}$ définie par

$$X_n = \max\{Y_0, Y_1, \dots, Y_n\}, \quad n \in \mathbb{N}.$$

1. Si l'on note $x_+ := \max\{x, 0\}$ la partie positive d'un nombre $x$, montrez que
$$X_{n+1} = X_n + (Y_{n+1} - X_n)_+.$$
2. Montrez que $X_n$ est intégrable pour tout $n \in \mathbb{N}$.
3. Montrez que $(X_n)_{n\in\mathbb{N}}$ est une sous-martingale pour la filtration $(\mathscr{F}_n)_{n\in\mathbb{N}}$.
4. Calculez l'espérance conditionnelle $\mathbb{E}\left[(Y_{n+1} - X_n)_+ \mid \mathscr{F}_n\right]$.
5. Déduisez-en la décomposition de Doob de la sous-martingale $(X_n)_{n\in\mathbb{N}}$.

**Exercice 3.16** Soit $p \neq \frac{1}{2}$. On considère $(U_n)_{n\in\mathbb{N}}$ une suite de variables aléatoires indépendantes de loi de Rademacher :

$$\mathbb{P}(U_i = 1) = p, \quad \mathbb{P}(U_i = -1) = 1 - p = q.$$

On travaille dans $(\mathscr{F}_n)_{n\in\mathbb{N}}$ la filtration naturelle des $(U_n)_{n\in\mathbb{N}}$. On fixe $N > 0$ et $k \in \{0, \dots, N\}$, et on pose

$$S_n = k + \sum_{j=1}^{n} U_j, \quad Y_n = \left(\frac{q}{p}\right)^{S_n}.$$

1. Montrer que $(Y_n)_{n\in\mathbb{N}}$ est une martingale.
2. On pose $T = \inf\big\{n \in \mathbb{N};\ S_n \in \{0, N\}\big\}$. Montrer que $T$ est un temps d'arret.
3. On pose $p_k = \mathbb{P}(S_T = 0 | S_0 = k)$. Justifier soigneusement que
$$\mathbb{E}(Y_T) = p_k + \left(\frac{p}{q}\right)^N (1 - p_k).$$
4. En déduire que $p_k = \frac{(q/p)^k - (q/p)^N}{1 - (q/p)^N}$.

# 4. Inégalités maximales

## 4.1 Objectif de ces inégalités

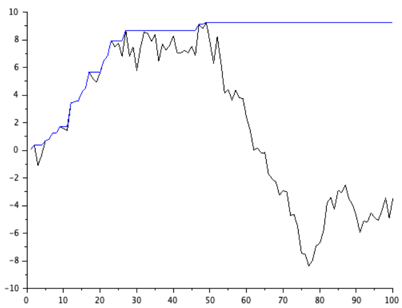
Lorsque l'on étudie un processus aléatoire, une question importante en pratique est de savoir contrôler son évolution. Dans le cas simple d'une seule variable aléatoire, on a l'inégalité de Markov $\mathbb{P}(X \geq t) \leq \frac{\mathbb{E}[X]}{t}$. Si maintenant, on cherche à appliquer cette inégalité à une **suite** de variables aléatoires $(X_n)_{n\geq 0}$, la quantité à estimer serait :

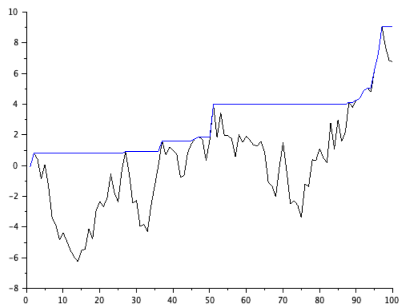
$$X_n^* = \sup_{k\leq n} X_k.$$

Par exemple, si $(X_n)_{n\geq 0}$ représente la **perte d'un joueur au casino**, on souhaiterait contrôler

$$\mathbb{P}\left(X_n^* \geq t\right) = \mathbb{P}\left(\sup_{k\leq n} X_k \geq t\right) \leq \frac{\mathbb{E}\left(\sup_{k\leq n} X_k\right)}{t},$$

la dernière inégalité découlant simplement de Markov. Alors, on voit qu'on est ramené à calculer l'espérance d'un sup : $\mathbb{E}\left(\sup_{k\leq n} X_k\right)$. Mais travailler avec $\sup_{k\leq n}$ dans l'espérance est difficile :




Supposons qu'il soit possible d'échanger $\sup_{k\leq n}$ et $\mathbb{E}$, c'est à dire qu'il soit possible d'écrire $\mathbb{E}\left(\sup_{k\leq n} X_k\right) = \sup_{k\leq n} \mathbb{E}\left(X_k\right)$. On peut alors estimer la probabilité :

$$\mathbb{P}\left(\sup_{k\leq n} X_k \geq t\right) \leq \frac{\sup_{k\leq n} \mathbb{E}\left(X_k\right)}{t}$$

Quand est-ce que l'interversion de $\sup_{k\leq n}$ et $\mathbb{E}$ est possible ? La réponse : pour les sous et surmartingales positives. On les appelle inégalités maximales de Doob.

## 4.2 Les inégalités de Doob

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

**Théorème 4.2.1 — Inégalités maximales de Doob.** Soit $(M_n)_{n\in\mathbb{N}}$ une sous-martingale positive. Alors pour tout $n \in \mathbb{N}$,

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

ce qui conclut la démonstration de la première inégalité.
Pour la seconde inégalité, on notera que si $(M_n)_{n\in\mathbb{N}}$ est une sous-martingale positive telle que $M_n \in L^p$ pour tout $n \in \mathbb{N}$, alors $(M_n^p)_{n\in\mathbb{N}}$ est une sous-martingale positive et l'inégalité précédente s'applique.

Enfin, pour la troisième inégalité, le processus $(M_n)_{n\in\mathbb{N}}$ étant une surmartingale positive, il suffit de reprendre le raisonnement précédent pour aboutir à l'inégalité

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

et le théorème de convergence monotone achève la démonstration. ■

Illustrons l'efficacité de ces inégalités par rapport à l'inégalité de Markov à l'aide de l'exercice suivant.

**Exercice 4.1** Soit $\xi_1, \xi_2, \ldots$ une suite de variables aléatoires indépendantes de loi $\mathcal{N}(0,\sigma^2)$.

1. Montrer qu'en utilisant l'inégalité de Markov, on a :

$$\mathbb{P}\left(\max_{1\le k\le n}\left(\xi_1+\cdots+\xi_k\right) \ge \varepsilon\right) \le n\exp\left(-\frac{\varepsilon^2}{2n\sigma^2}\right).$$

2. Montrer qu'en utilisant l'inégalité de Doob, on obtient :

$$\mathbb{P}\left(\max_{1\le k\le n}\left(\xi_1+\cdots+\xi_k\right) \ge \varepsilon\right) \le \exp\left(-\frac{\varepsilon^2}{2n\sigma^2}\right).$$

■

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

**Théorème 4.2.2 — Inégalités de Doob $L^p$.** Soit $p > 1$. Supposons que $(M_n)_{n\in\mathbb{N}}$ soit une sous-martingale positive telle que $M_n \in L^p$ pour tout $n \in \mathbb{N}$. Alors pour tout $n \in \mathbb{N}$, $M_n^* \in L^p$ et

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

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

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

En se ramenant à la première inégalité maximale du théorème 4.2.1, il vient

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

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

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

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

## 4.3 Exercices du Chapitre 4

**Exercice 4.2** Soit $(Y_n)_{n\in\mathbb{N}^*}$ une suite de v.a. indépendantes. On suppose que chaque $Y_n$ suit la loi de Rademacher de paramètre $p_n \in ]0,1/2[$. On considère le processus $(S_n)_{n\in\mathbb{N}}$ défini par $S_0 = 0$ et

$$S_{n+1} = S_n + (1+\cos(S_n)^2)\,Y_{n+1}, \quad n \in \mathbb{N}.$$

On note également $\mathscr{F}_0 = \{\emptyset, \Omega\}$ et $\mathscr{F}_n = \sigma(Y_1,\ldots,Y_n)$ pour tout $n \in \mathbb{N}^*$.

1. Montrez que le processus $(S_n)_{n\in\mathbb{N}}$ est une surmartingale pour la filtration $(\mathscr{F}_n)_{n\in\mathbb{N}}$.
2. Déterminez sa décomposition de Doob, i.e., les deux processus $(M_n)_{n\in\mathbb{N}}$ et $(A_n)_{n\in\mathbb{N}}$ où $(M_n)_{n\in\mathbb{N}}$ est une martingale et $(A_n)_{n\in\mathbb{N}}$ est un processus décroissant prévisible, intégrable et nul en 0.

3. Montrez qu'il existe une constante $C > 0$ telle que pour tout $n \in \mathbb{N}^*$,

$$\mathbb{E}\left[\sup_{0\leq k\leq n} |M_k|^2\right] \leq C \sum_{k=1}^{n} p_k(1-p_k).$$

4. Déduisez-en que la v.a. $\sup_{k\geq 0} |M_k|$ appartient à l'espace $L^2$ si et seulement si la série de terme général $p_k$ converge.

**Exercice 4.3 — Inégalités maximales pour la marche aléatoire asymétrique.** Considérons $(S_n)_{n\in\mathbb{N}}$ la marche aléatoire asymétrique sur $\mathbb{Z}$ et issue de 0, c'est-à-dire que $S_0 = 0$ et pour tout $n \in \mathbb{N}^*$,

$$S_n = \sum_{k=1}^{n} U_k,$$

où les $U_k$ sont i.i.d. de loi de Rademacher de paramètre $p \in ]0,1/2[$. On note également $\mathscr{F}_0 = \{\emptyset, \Omega\}$ et $\mathscr{F}_n = \sigma(U_1,\ldots,U_n)$ pour tout $n \in \mathbb{N}^*$.

1. Montrez que le processus $(Z_n)_{n\in\mathbb{N}}$ défini par $Z_n = (q/p)^{S_n}$ est une martingale positive.
2. Déduisez-en les inégalités suivantes : pour tout $k \in \mathbb{N}$,

$$\mathbb{P}\left(\sup_{n\geq 0} S_n \geq k\right) \leq \left(\frac{p}{q}\right)^k \quad \text{et} \quad \mathbb{E}\left[\sup_{n\geq 0} S_n\right] \leq \frac{q}{q-p}.$$

**Exercice 4.4 — Martingale Gaussienne.** Soit $(M_n)_{n\in\mathbb{N}}$ une martingale issue de 0 et de carré intégrable. On suppose que cette martingale est un processus gaussien, c'est-à-dire que tous les vecteurs extraits de ce processus (de taille finie) sont des vecteurs gaussiens. On note $(\mathscr{F}_n)_{n\in\mathbb{N}}$ la filtration naturelle de cette martingale.

1. Montrez que les accroissements $d_k = M_k - M_{k-1}$, $k \in \mathbb{N}^*$, avec $d_0 = 0$, sont indépendants.
2. Déduisez-en que $d_{n+1}$ est indépendante de la tribu $\mathscr{F}_n$ pour tout $n \in \mathbb{N}$.
3. On suppose dans la suite que $\mathrm{Var}(d_k) = \sigma^2 > 0$ pour tout $k \in \mathbb{N}^*$. Calculez la variation quadratique de la martingale $(M_n)_{n\in\mathbb{N}}$.
4. Soit $\theta \in \mathbb{R}$. Montrez que le processus $(X_n)_{n\in\mathbb{N}}$ défini pour tout $n \in \mathbb{N}$ par

$$X_n = \exp\left(\theta M_n - \frac{\theta^2}{2}[M,M]_n\right),$$

   est une martingale.
5. Montrez que l'on a l'inégalité maximale suivante : pour tout $n \in \mathbb{N}$ et tout $x \geq 0$,

$$\mathbb{P}\left(\sup_{0\leq k\leq n} M_k \geq x\right) \leq \exp\left(-\frac{x^2}{2n\sigma^2}\right).$$

# 5. Théorèmes de Convergence

Le joueur au casino qui décide de s'arrêter lorsqu'il atteint une somme $M$ ou lorsqu'il n'a plus d'argent se modélise par :

$$W_n = s_0 + \sum_{i=1}^{n\wedge T} X_i,$$

où $T = \inf\{n \geq 0;\ S_n = 0 \text{ ou } S_n = M\}$. Alors, $(W_n)_{n\geq 0}$ est une **martingale bornée :**

$$\forall n \geq 0,\ \ 0 \leq W_n \leq M$$

Intuitivement, $W_n$ finit par toucher un jour soit 0, soit $M$ et alors reste à cette position pour toujours :

$$\lim_{n\to+\infty} W_n = W_\infty,$$

où $W_\infty$ est une variable aléatoire valant 0 ou $M$.

Un autre exemple est la martingale

$$M_n = \mathbb{E}(Z|\mathscr{F}_n).$$

À chaque temps $n \geq 0$, $M_n$ est la meilleure approximation de $Z$ connaissant $\mathscr{F}_n$. $(\mathscr{F}_n)_{n\geq 0}$ forme une filtration ; **l'information grandit**, donc l'approximation se précise. Intuitivement, à la limite, on découvre $Z$ totalement, c'est à dire :

$$\lim_{n\to+\infty} M_n = Z.$$

Ces deux cas sont deux exemples de **convergence de martingales**. On remarque que contrairement à la loi des grand nombres par exemple, la limite est encore aléatoire : Pour le casino $W_\infty$ représente la fortune du joueur quand le jeu s'arrête : elle est aléatoire. Pour la $M_n = \mathbb{E}(Z|\mathscr{F}_n)$, par définition, $Z$ est une variable aléatoire. Le but de ce chapitre est de donner des conditions **nécessaires et suffisantes** pour garantir la convergence de martingales.

## 5.1 Les théorèmes de convergence

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

**Théorème 5.1.1** Soit $(M_n)_{n\in\mathbb{N}}$ une martingale bornée dans $L^2$, i.e.

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

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

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

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

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

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

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

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

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

$$\begin{aligned} \mathbb{E}[d_k d_l] &= \mathbb{E}[(M_k - M_{k-1})(M_l - M_{l-1})] \\ &= \mathbb{E}[M_k M_l] - \mathbb{E}[M_k M_{l-1}] - \mathbb{E}[M_{k-1} M_l] + \mathbb{E}[M_{k-1} M_{l-1}] \\ &= \mathbb{E}[\mathbb{E}[M_k M_l \mid \mathscr{F}_k]] - \mathbb{E}[\mathbb{E}[M_k M_{l-1} \mid \mathscr{F}_k]] - \mathbb{E}[\mathbb{E}[M_{k-1} M_l \mid \mathscr{F}_{k-1}]] \\ &\quad + \mathbb{E}[\mathbb{E}[M_{k-1} M_{l-1} \mid \mathscr{F}_{k-1}]] \\ &= \mathbb{E}[M_k \mathbb{E}[M_l \mid \mathscr{F}_k]] - \mathbb{E}[M_k \mathbb{E}[M_{l-1} \mid \mathscr{F}_k]] - \mathbb{E}[M_{k-1} \mathbb{E}[M_l \mid \mathscr{F}_{k-1}]] \\ &\quad + \mathbb{E}[M_{k-1} \mathbb{E}[M_{l-1} \mid \mathscr{F}_{k-1}]] \\ &= \mathbb{E}\left[M_k^2\right] - \mathbb{E}\left[M_k^2\right] - \mathbb{E}\left[M_{k-1}^2\right] + \mathbb{E}\left[M_{k-1}^2\right] \\ &= 0. \end{aligned}$$

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

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

qui est la somme partielle d'une série convergente par l'hypothèse de $L^2$-bornitude. Par ailleurs, le processus $(M_{k+m} - M_m)_{k\in\mathbb{N}}$ étant une martingale (pour la filtration translatée $(\mathscr{F}_{k+m})_{k\in\mathbb{N}}$), la suite $(M'_k)_{k\in\mathbb{N}}$ définie pour tout $k \in \mathbb{N}$ par $M'_k := (M_{k+m} - M_m)^2$ est une sous-martingale positive et l'inégalité de Doob du théorème 4.2.1, combinée au lemme de Fatou, s'applique de la manière suivante :

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

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

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

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

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

■

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

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

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

**Théorème 5.1.2** Une sous-martingale bornée, tout comme une surmartingale positive, est p.s. convergente.

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

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

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

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

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

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

c'est-à-dire que la martingale arrêtée est bornée, donc p.s. convergente par le théorème 5.1.1. Comme on a $X_n = X_{n\wedge\tau_p}$ lorsque $\tau_p = +\infty$, on en déduit que la sous-martingale $(X_n)_{n\in\mathbb{N}}$ converge p.s. sur l'événement $\{\tau_p = +\infty\}$. En utilisant la convergence monotone à deux reprises, on a pour tout $p \in \mathbb{N}$,

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

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

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

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

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

**Exercice 5.1** On considère le processus $(X_n)_{n\in\mathbb{N}}$ défini par $X_0 = p \in ]0,1[$ et pour tout $n \in \mathbb{N}$,

$$X_{n+1} = \theta X_n + (1-\theta)U_{n+1},$$

où $\theta \in ]0,1[$ est un paramètre donné et $(U_n)_{n\in\mathbb{N}^*}$ est une suite de v.a. de Bernoulli.

1. Montrez que $X_n \in ]0,1[$ pour tout $n \in \mathbb{N}$ .
2. Dans la suite, on note $\mathscr{F}_n = \sigma(X_0, X_1, \ldots, X_n)$ la filtration engendrée par le processus $(X_n)_{n\in\mathbb{N}}$ et l'on suppose que pour tout $n \in \mathbb{N}$,

$$\mathbb{P}(U_{n+1} = 1 \mid \mathscr{F}_n) = X_n \quad \text{et} \quad \mathbb{P}(U_{n+1} = 0 \mid \mathscr{F}_n) = 1 - X_n.$$

   Montrez que le processus $(X_n)_{n\in\mathbb{N}}$ a une espérance constante égale à $p$.
3. Montrez que $(X_n)_{n\in\mathbb{N}}$ est une martingale.
4. Justifiez la convergence p.s. et dans $L^2$ de cette martingale (vers une v.a. que l'on notera $L$).
5. Montrez que pour tout $n \in \mathbb{N}$, on a

$$\mathbb{E}\left[(X_{n+1} - X_n)^2\right] = (1-\theta)^2\, \mathbb{E}[X_n(1-X_n)].$$

6. Déduisez-en la valeur de l'espérance $\mathbb{E}[L(1-L)]$.
7. Déduisez-en la loi de $L$.
8. **Question bonus :** Montrez que pour tout $n \in \mathbb{N}$,

$$\mathrm{Var}(X_n) = p(1-p)\left(1 - (2\theta - \theta^2)^n\right).$$

■

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

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

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

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

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

**Théorème 5.1.3** Soit $(M_n)_{n\in\mathbb{N}}$ une martingale de carré intégrable. Si l'on a $[M,M]_\infty < +\infty$ p.s., la martingale est p.s. convergente.

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

**Théorème 5.1.4** Soit $(X_n)_{n\in\mathbb{N}}$ une sous-martingale bornée dans $L^1$, i.e.

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

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

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

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

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

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

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

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

**Théorème 5.1.5** Étant donné $p > 1$, soit $(M_n)_{n\in\mathbb{N}}$ une martingale bornée dans $L^p$, i.e.

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

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

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

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

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

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

■

## 5.2 Les Algorithmes de Robbins-Monro

Dans cette section, nous étudions une catégorie d'algorithmes stochastiques ayant une utilité particulière du point de vue des applications : les algorithmes de Robbins-Monro. Ces algorithmes stochastiques servent à approcher numériquement la solution d'une équation de la forme $h(\theta) = 0$, et sont définis de façon récursive et s'écrivent de la façon suivante :

$$X_{n+1} = X_n - \gamma_{n+1} h(X_n) + \gamma_{n+1}\Delta M_{n+1}, \quad X_0 \in \mathbb{R}^d,$$

où $(\Delta M_n)_{n\geq 1}$ est une suite d'incréments martingale,

$$\sum_{n\geq 1} \gamma_n = +\infty, \quad \sum_{n\geq 1} \gamma_n^2 < +\infty. \tag{5.1}$$

**Remarque** Il est possible d'inclure un reste dans la définition de $(X_n)_{n\geq 0}$ :

$$X_{n+1} = X_n - \gamma_{n+1} h(X_n) + \gamma_{n+1}\Delta M_{n+1} + \gamma_{n+1}^2 \rho_{n+1}, \quad X_0 \in \mathbb{R}^d.$$

où $\rho_{n+1}$ peut être aléatoire. Des hypothèses supplémentaires sont alors nécessaires. Pour simplifier la présentation, nous choisissons de ne pas inclure ce terme.

■ **Exemple 5.1** Un algorithme bien connu, celui du **Gradient Stochastique**, entre dans cette catégorie. Le but de cet algorithme est l'optimisation d'une fonction $f$. Dans ce cas on prend $h(x) = \nabla f(x)$, et on remarque la décomposition suivante :

$$X_{n+1} = \underbrace{X_n - \gamma_{n+1}\nabla f(X_n)}_{\text{algorithme du gradient}} + \underbrace{\gamma_{n+1}\Delta M_{n+1}}_{\text{terme de bruit}}.$$

Dans la suite, on donnera des hypothèses pour garantir la convergence de cet algorithme. ■

■ **Exemple 5.2** Supposons que pour l'équation $h(\theta) = 0$ que l'on cherche à résoudre, la fonction $h$ a la forme particulière :

$$h(\theta) = \mathbb{E}[H(\theta, \xi)] = 0.$$

Alors, on peut prendre $\Delta M_{n+1} = h(X_n) - H(X_n, \xi_{n+1})$, et l'algorithme se simplifie en

$$X_{n+1} = X_n - \gamma_{n+1} H(X_n, \xi_{n+1}), \tag{5.2}$$

où $(\xi_i)_{i\geq 1}$ sont des copies indépendantes de même loi que $\xi$. En particulier, si on prend comme fonction $H$ :

$$H(\theta, u) = \theta - \varphi(u),$$

où $\varphi$ est une fonction quelconque, résoudre $\mathbb{E}(H(\theta,\xi)) = 0$ revient à prendre :

$$\mathbb{E}(\theta - \varphi(\xi)) = 0 \Leftrightarrow \theta = \mathbb{E}[\varphi(\xi)]$$

Mais alors l'algorithme précédent devient en spécifiant $\gamma_n = 1/n$ :

$$\begin{aligned} \theta_{n+1} &= \theta_n - \gamma_{n+1} H(\theta_n, \xi_{n+1}) \\ &= \theta_n - \frac{1}{n+1}\Big(\theta_n - \varphi(\xi_{n+1})\Big) \end{aligned}$$

On peut arranger cette expression en :

$$\begin{aligned} \theta_{n+1} &= \theta_n - \frac{1}{n+1}\Big(\theta_n - \varphi(\xi_{n+1})\Big) \\ (n+1)\theta_{n+1} &= (n+1)\theta_n - \Big(\theta_n - \varphi(\xi_{n+1})\Big) \\ (n+1)\theta_{n+1} - n\theta_n &= \varphi(\xi_{n+1}) \end{aligned}$$

Ainsi, en sommant de $n = 0$ à $n = N-1$ :

$$N\theta_N - 0 \times \theta_0 = \sum_{n=0}^{N-1} \Big((n+1)\theta_{n+1} - n\theta_n\Big) = \sum_{n=0}^{N-1} \varphi(\xi_{n+1}).$$

Et donc :

$$\theta_N = \frac{1}{N}\sum_{n=0}^{N-1} \varphi(\xi_{n+1}).$$

Ainsi, les algorithmes de Robbins-Monro contiennent les algorithmes de Montecarlo.

■

**Théorème 5.2.1 — Théorème de Robbins Siegmund.** Soient $(V_n)_{n\geq 1}$, $(\alpha_n)_{n\geq 1}$, $(\beta_n)_{n\geq 1}$ et $(U_n)_{n\geq 1}$, des suites de variables aléatoires positives telles que :

1. $(\alpha_n)_{n\geq 1}$, $(\beta_n)_{n\geq 1}$ et $(U_n)_{n\geq 1}$ sont des suites de variables aléatoires prévisibles.
2. $\sup\limits_{\omega\in\Omega} \prod\limits_{n\geq 1} (1 + \alpha_n(\omega)) < +\infty$,
3. $\sum\limits_{n\geq 1} \mathbb{E}(\beta_n) < +\infty$,
4. $\forall n \geq 1$, on a la relation de récurrence :

$$E(V_{n+1}|\mathscr{F}_n) \leq V_n\big(1 + \alpha_{n+1}\big) + \beta_{n+1} - U_{n+1}.$$

Alors, on a les conclusions suivantes :

1. On a la convergence $V_n \underset{n\to+\infty}{\longrightarrow} V_\infty \in L^1$ presque sûrement et $\sup_{n\geq} E(V_n) < +\infty$,
2. $\sum\limits_{n\geq 1} E(U_n) < +\infty$ et $\sum\limits_{n\geq 1} U_n < +\infty$ presque sûrement.

*Démonstration.* La preuve consiste à trouver une sur-martingale positive, et conclure à l'aide des théorèmes de convergences. Considérons tout d'abord

$$V_n + \sum_{k=1}^{n} U_k.$$

En utilisant la relation de récurrence sur $(V_n)_{n\geq 0}$, on obtient :

$$\begin{aligned} \mathbb{E}\left(V_{n+1} + \sum_{k=1}^{n+1} U_k | \mathscr{F}_n\right) &\leq V_n(1 + \alpha_{n+1}) + \beta_{n+1} - U_{n+1} + \sum_{k=1}^{n+1} U_k \\ &= V_n(1 + \alpha_{n+1}) + \beta_{n+1} + \sum_{k=1}^{n} U_k \end{aligned}$$

Ce n'est pas encore une sur-martingale, mais on s'y rapproche. Pour nous débarrasser de $(1+\alpha_{n+1})$, posons :

$$S_n = \frac{V_n + \sum_{k=1}^n U_k}{\prod_{k=1}^n (1+\alpha_k)}.$$

Dans ce cas, on obtient la relation :

$$\begin{aligned}
\mathbb{E}(S_{n+1}|\mathscr{F}_n) &\leq \frac{V_n(1+\alpha_{n+1}) + \beta_{n+1} + \sum_{k=1}^n U_k}{\prod_{k=1}^{n+1}(1+\alpha_k)} \\
&\leq (1+\alpha_{n+1}) \times \frac{V_n + \sum_{k=1}^n U_k}{\prod_{k=1}^{n+1}(1+\alpha_k)} + \frac{\beta_{n+1}}{\prod_{k=1}^{n+1}(1+\alpha_k)} \\
&\leq S_n + \frac{\beta_{n+1}}{\prod_{k=1}^{n+1}(1+\alpha_k)}
\end{aligned}$$

Ce n'est toujours pas une sur-martingale, mais on en est pas loin ! Concentrons nous sur

$$B_n = \sum_{k=1}^n \frac{\beta_k}{\prod_{i=1}^k (1+\alpha_i)}.$$

La suite $(B_n)_{n\geq 0}$ est une somme de terme positifs, donc c'est une suite croissante. On a alors presque surement :

$$B_n \underset{n\to+\infty}{\longrightarrow} B_\infty,$$

quitte à prendre $B_\infty = +\infty$.

Mais comme par convergence monotone, on a :

$$\begin{aligned}
\mathbb{E}(B_\infty) &= \mathbb{E}(\lim_{n\to+\infty} B_n) = \lim_{n\to+\infty} \mathbb{E}(B_n) \\
&= \lim_{n\to+\infty} \mathbb{E}\left(\sum_{k=1}^n \frac{\beta_k}{\prod_{i=1}^k(1+\alpha_i)}\right) \\
&\leq \sum_{k=1}^{+\infty} \mathbb{E}\left(\frac{\beta_k}{\prod_{i=1}^k(1+\alpha_i)}\right) \leq \sum_{k=1}^{+\infty} \mathbb{E}(\beta_k) < +\infty,
\end{aligned}$$

on rappelle en effet que par hypothèse $\sum_{k=1}^{+\infty} \mathbb{E}(\beta_k) < +\infty$, on a donc forcément que $B_\infty$ est fini presque surement.

Notre sur-martingale sera alors :

$$\tilde{S}_n = S_n + \mathbb{E}(B_\infty|\mathscr{F}_n) - B_n$$

Vérifions qu'elle est positive. D'une part, on a bien :

$$S_n = \frac{V_n + \sum_{k=1}^n U_k}{\prod_{k=1}^n (1+\alpha_k)} \geq 0$$

D'autre part, puisque

$$B_n = \sum_{k=1}^n \frac{\beta_k}{\prod_{i=1}^k(1+\alpha_i)} \underset{n\to+\infty}{\nearrow} B_\infty$$

donc on a aussi $\mathbb{E}(B_\infty|\mathscr{F}_n) - B_n \geq 0$, et donc que $(\tilde{S}_n)_{n\geq 0}$ est bien positive. Vérifions la propriété de sur-martingale :

$$\begin{aligned}
\mathbb{E}(\tilde{S}_{n+1}|\mathscr{F}_n) &= \mathbb{E}(S_{n+1} + \mathbb{E}(B_\infty|\mathscr{F}_{n+1}) - B_{n+1}|\mathscr{F}_n) \\
&= S_n + \frac{\beta_{n+1}}{\prod_{k=1}^{n+1}(1+\alpha_k)} + \mathbb{E}(B_\infty|\mathscr{F}_n) - B_{n+1}
\end{aligned}$$

Mais comme $B_n = \sum_{k=1}^{n} \frac{\beta_k}{\prod_{i=1}^{k}(1+\alpha_i)}$, on a donc bien :

$$\mathbb{E}(\tilde{S}_{n+1}|\mathscr{F}_n) \leq S_n + \mathbb{E}(B_\infty|\mathscr{F}_n) - B_n = \tilde{S}_n.$$

Ainsi, $(\tilde{S}_n)_{n\geq 0}$ est une sur-martingale positive, elle converge presque surement. On récupère les conclusions du théorème grace à cette convergence. En effet : puisque $B_n \underset{n\to+\infty}{\nearrow} B_\infty$, on a que

$$\mathbb{E}(B_\infty|\mathscr{F}_n) - B_n \underset{n\to+\infty}{\longrightarrow} 0 \text{ presque surement, et dans } L^1.$$

donc, on en déduit que, $S_n = \frac{V_n + \sum_{k=1}^{n} U_k}{\prod_{k=1}^{n}(1+\alpha_k)}$ converge aussi vers $S_\infty \in L^1$. Mais puisque

$$V_n + \sum_{k=1}^{n} U_k = S_n \left( \prod_{k=1}^{n}(1+\alpha_k) \right)$$

on a :

$$\mathbb{E}(V_n) + \mathbb{E}\left( \sum_{k=1}^{n} U_k \right) = \mathbb{E}\left( S_n \left( \prod_{k=1}^{n}(1+\alpha_k) \right) \right).$$

Et comme $\sup_{n\in\mathbb{N}} \prod_{k=1}^{n}(1+\alpha_k) < +\infty$, on a :

$$\mathbb{E}\left( S_n \left( \prod_{k=1}^{n}(1+\alpha_k) \right) \right) \leq \mathbb{E}\left(S_n\right) \| \prod_{k=1}^{n}(1+\alpha_k) \|_\infty.$$

Remarquons que $\mathbb{E}\left(S_n\right) \leq \mathbb{E}\left(\tilde{S}_n\right) = \mathbb{E}\left(S_1\right) < +\infty$ puisque $\tilde{S}$ est une sur-martingale, on en déduit :

$$\sup_{n\in\mathbb{N}} \mathbb{E}(V_n) < +\infty \text{ and } \sup_{n\in\mathbb{N}} \mathbb{E}\left( \sum_{k=1}^{n} U_k \right) < +\infty.$$

Maintenant, on a clairement que

$$\sup_{n\in\mathbb{N}} \mathbb{E}\left( \sum_{k=1}^{n} U_k \right) = \mathbb{E}\left( \sum_{k=1}^{+\infty} U_k \right) < +\infty \Rightarrow \sum_{k=1}^{+\infty} U_k \ \ (p.s.)$$

On obtient finalement que :

$$V_n + \underbrace{\sum_{k=1}^{n} U_k}_{\text{converge}} = \underbrace{S_n}_{\text{converge}} \underbrace{\left( \prod_{k=1}^{n}(1+\alpha_k) \right)}_{\text{finite}}$$

On peut alors déduire que $V_n$ converge vers $V_\infty \in L^1$. Ainsi, toutes les conclusions du théorèmes ont été établies.

■

**Remarque** Il existe de nombreuses variantes de ce théorèmes, y comprit des versions déterministes, mais l'un des avantages de ce théorème est que toutes les suites utilisées peuvent être aléatoires.

Nous donnons maintenant une utilisation de ce théorème utile dans les applications : l'algorithme du bandit. Cet algorithme est un exemple d'algorithme de recommendation, et constitue la base de ceux utilisé par vos réseaux préférés. Nous étudions l'exemple simple du Bandit à 2 bras, des généralisations à $N$ bras sont possible avec d'autres techniques.

■ **Exemple 5.3** On considère une machine à sous avec deux leviers ; le levier $A$ et le levier $B$.
— Le levier $A$ gagne avec probabilité $p_A$, et le levier $B$ avec probabilité $p_B$.
— Mais on ne connait pas qui de $p_A$ ou $p_B$ est la plus grande.
Le but de cet algorithme est de déterminer lequel des deux bras il vaut mieux jouer.

On choisit un $x_0 \in (0,1)$, quelconque, et on se donne une suite $(U_n)_{n\geq 1}$ de variables aléatoires indépendantes de loi uniforme sur $(0,1)$. On construit récursivement

$$X_{n+1} = X_n + \gamma_{n+1} \begin{cases} \mathbf{1}_{A_{n+1}}(1-X_n) & \text{if } U_{n+1} \leq X_n \\ -\mathbf{1}_{B_{n+1}} X_n & \text{if } U_{n+1} > X_n \end{cases}$$

**Remarque** L'algorithme fonctionne de la façon suivante :
1. Choix de $x_0 \in [0,1]$
2. Simulation de $U$ uniforme sur [0,1]
   — Si : $x_0 \leq U$, alors, on joue $A$
     $\rightarrow$ on met à jour $x_0$ en $x_1$ si le bras $A$ gagne
   — Sinon, on joue $B$
     $\rightarrow$ on met à jour $x_0$ en $x_1$ si le bras $B$ gagne
3. Répète étape 2.

**Proposition 5.2.2** La suite $(X_n)_{n\geq 0}$ converge presque surement vers 0 ou 1.

On prend la convention que si $X_n = 1$, alors on joue toujours le bras $A$, si $X_n = 0$, on joue toujours le bras $B$. Ainsi, la procédure finit par se fixer sur l'un des deux bras.

*Démonstration.* On pose

$$H(X_n, U_{n+1}) = \begin{cases} \mathbf{1}_{A_{n+1}}(1-X_n) & \text{if } U_{n+1} \leq X_n \\ -\mathbf{1}_{B_{n+1}} X_n & \text{if } U_{n+1} > X_n \end{cases}$$

Ainsi, on remarque qu'on peut écrire

$$X_{n+1} = X_n + \gamma_{n+1} H(X_n, U_{n+1}), \text{ avec } h(x) = \mathbb{E}[H(x,U)] = x(1-x)\Big(p_A - p_B\Big).$$

On va utiliser le théorème de Robbins-Siegmund avec $V_n = 1 - X_n$. On remarque que :

$$\begin{aligned} \mathbb{E}(V_{n+1} | \mathscr{F}_n) &= 1 - \mathbb{E}(X_n + \gamma_{n+1} H(X_n, U_{n+1}) | \mathscr{F}_n) \\ &= V_n + \gamma_{n+1} \mathbb{E}(H(X_n, U_{n+1}) | \mathscr{F}_n) \\ &= V_n - \underbrace{\gamma_{n+1} X_n (1-X_n)\Big(p_A - p_B\Big)}_{\geq 0}. \end{aligned}$$

Avec les notations du théorème de Robbins-Siegmund, on peut poser que tout le terme sous l'accolade corresponde à la suite $U_{n+1}$, et $\alpha_n = \beta_n = 0$. Ainsi, les hypothèses de Robbins-Siegmund sont vérifiées, et on a les conclusions suivantes :
— $V_n$ converge, donc, $X_n$ converge aussi.
— $\sum_{k\geq 0} \mathbb{E}(U_k) < +\infty$
Il nous reste à prouver que les limites possible de $(X_n)_{n\geq 0}$ sont 0 ou 1. Pour cela, on se sert du fait que la série des $(U_n)_{n\geq 0}$ converge, donc que son terme général tend vers zéro :

$$\sum_{k\geq 0} \mathbb{E}(U_k) < +\infty \Leftrightarrow \sum_{k\geq 0} \gamma_{k+1} \mathbb{E}\Big(X_k(1-X_k)\Big) < +\infty$$

Mais comme $\sum \gamma_k = +\infty$, on a forcément que :

$$\lim_{k\to+\infty} \mathbb{E}\Big(X_k(1-X_k)\Big) = 0$$

Mais sachant que $X_n \underset{n\to+\infty}{\longrightarrow} X_\infty$ presque surement, et que $X_n \in [0,1]$ pour tous $n \in \mathbb{N}$, on a par le théorème de convergence dominée :

$$\mathbb{E}\Big(X_\infty(1-X_\infty)\Big) = 0.$$

Ainsi, $X_\infty(1-X_\infty) = 0$ presque surement (l'espérance d'une variable aléatoire positive presque surement est nulle entraine que la variable est nulle presque surement), ainsi :

$$X_\infty = 0 \text{ ou bien } X_\infty = 1.$$

■

**Remarque** Le résultat que l'on aurait aimé obtenir est que :

$$\lim_{n\to+\infty} X_n = \begin{cases} 1 & \text{if } p_A > p_B \\ 0 & \text{if } p_A < p_B \end{cases}$$

Pour l'instant, ce résultat est encore hors de portée. Pour l'établir, il faut étudier la stabilité de l'EDO :

$$\dot{x}(t) = x(t)(1-x(t))\Big(p_A - p_B\Big) = h(x(t))$$

On peut alors montrer que si $p_A > p_B$, alors, 1 est stable et 0 est instable, et réciproquement.

**Remarque** Cet algorithme présente des phénomènes de **fausse convergence**. En effet, nous avons remarqué que si $X_n = 0$ ou $X_n = 1$ pour un certain $n \geq 0$, alors, la suite reste fixée sur 0 ou 1. Cela signifie que si le premier pas est trop grand (nous faisant sortir de l'intervalle [0,1]), l'algorithme peut se fixer sur une mauvaise valeur. Ces phénomènes sont encore à l'étude à l'heure actuelle, avec un regain d'intérêt suite à l'avènement des algorithmes de recommendation et leur utilisation dans les plates-formes en ligne.

**Remarque** Il est possible de prouver la convergence de l'algorithme très facilement, en remarquant que $X_n$ est une sous-martingale bornée. Mais avec cet argument, il n'est pas possible d'établir que la convergence se fait certainement vers 0 ou vers 1.

■

L'exercice suivant utilise le même genre d'idée que l'exemple du Bandit. Il s'agit d'un autre exemple très utile dans les application, le cas d'observations bruité.

**Exercice 5.2** Dans cet exercice, on souhaite trouver $x^*$ telle que pour un certain $\alpha \in \mathbb{R}$ fixé, on satisfasse à l'équation suivante :

$$\phi(x^*) = \alpha$$

Le problème, c'est que l'observation directe de la fonction $\phi$ n'est pas possible, on a une observation bruité, pour tous $x \in \mathbb{R}$, on observe

$$Y = \phi(x) + \varepsilon,$$

où $\varepsilon$ est une variable aléatoire centrée. Pour cela, on pose l'algorithme :

$$X_{n+1} = X_n - \gamma_{n+1}(Y_{n+1} - \alpha), \text{ where } Y_{n+1} = \phi(X_n) + \varepsilon_{n+1}.$$

On note $(\mathscr{F}_n)_{n\geq 0}$ la filtration naturelle des $(\varepsilon_n)_{n\geq 0}$, qu'on suppose *i.i.d.*. On se donne les hypothèses suivantes :

— $\phi$ est continue, et il existe une unique solution à l'équation $\phi(x) = \alpha$.
— $(x - x^*)(\phi(x) - \alpha) > 0$
— $\mathbb{E}(Y_{n+1}|\mathscr{F}_n) = \phi(X_n)$, ou bien de façon équivalente $\mathbb{E}(\varepsilon_i) = 0$,
— $\mathbb{E}((Y_{n+1} - \alpha)^2|\mathscr{F}_n) \leq C\Big(1 + (X_n - x^*)^2\Big)$,
— $\sum \gamma_n = +\infty$, et $\sum \gamma_n^2 < +\infty$.

En utilisant Robbins Siegmund sur $V_n = (X_n - x^*)^2$ montrer que $X_n \underset{n\to+\infty}{\longrightarrow} x^*$.

Dans l'exercice précédent, le choix de la fonction $V(x) = (x - x^*)^2$ est assez crucial pour la preuve. C'est un exemple de **fonction de Lyapunov.** Le résultat suivant formule que si on trouve une telle fonction, alors, on est garanti de converger.

**Théorème 5.2.3 — Théorème de Convergence.** Soit l'algorithme stochastique :

$$\begin{aligned} X_{n+1} &= X_n - \gamma_{n+1} h(X_n) + \gamma_{n+1}\Delta M_{n+1}, \\ X_0 &= x_0 \in \mathbb{R}^d. \end{aligned}$$

On suppose que les conditions suivantes sont vérifiées :

1. Pas décroissant : $\sum_{k\geq 1}\gamma_k = +\infty$ et $\sum_{k\geq 1}\gamma_k^2 < +\infty$.
2. On suppose qu'il existe une fonction $V : \mathbb{R}^d \to \mathbb{R}$ positive, de classe $\mathscr{C}^2$ et telle que :
   (a) $\langle \nabla V(x) - \nabla V(y), x - y\rangle \leq L\|x-y\|^2$,
   (b) $\lim_{\|x\|\to+\infty} V(x) = +\infty$,
   (c) $\forall x \in \mathbb{R}^d$, $\langle \nabla V(x), h(x)\rangle \geq 0$,
   (d) $\forall x \in \mathbb{R}^d$, $\|h(x)\|^2 + \|\nabla V(x)\|^2 \leq C\big(1 + V(x)\big)$.
3. $(\Delta M_n)_{n\geq 1}$ est une suite d'accroissements martingales dans la filtration $(\mathscr{F}_n)_{n\geq 1}$ telle que $\forall n \geq 1$
$$\mathbb{E}\Big[\|\Delta M_n\|^2 | \mathscr{F}_{n-1}\Big] \leq C\big(1 + V(X_{n-1})\big).$$

Alors, on a les conclusions suivantes :

1. $\sup_{n\geq 1} \mathbb{E}[V(X_n)] < +\infty$,
2. $\sum_{k\geq 1} \gamma_{k+1}\langle \nabla V(X_n), h(X_n)\rangle < +\infty$ (p.s.),
3. $V(X_n) \underset{n\to+\infty}{\longrightarrow} V_\infty \in L^1$ presque sûrement,
4. $X_{n+1} - X_n \underset{n\to+\infty}{\longrightarrow} 0$ presque surement et dans $L^2$.

*Démonstration.* La preuve consiste à considérer la suite $(X_n)_{n\geq 0}$ le long des trajectoires de la fonction de Lyapunov $V$. La preuve est similaire à celle dans le cas déterministe (descente de gradient). On a :

$$\begin{aligned} V(X_{n+1}) &= V\Big(X_n - \gamma_{n+1}h(X_n) + \gamma_{n+1}\Delta M_{n+1} + \gamma_{n+1}^2\rho_{n+1}\Big) \\ &= V(X_n) - \gamma_{n+1}\langle \nabla V(X_n), h(X_n)\rangle \\ &\quad + \gamma_{n+1}\langle \nabla V(X_n), \Delta M_{n+1}\rangle + \gamma_{n+1}^2\langle \nabla V(X_n), \rho_{n+1}\rangle \\ &\quad + \frac{1}{2}D^2V(\xi_{n+1})\Delta X^{(2)} \end{aligned}$$

On utilise alors le théorème de Robbins-Siegmund sur $V_n = V(X_n)$. Puisque $\nabla V$ est Lipschitz, $D^2V$ est borné et on a directement :

$$\begin{aligned} |\frac{1}{2}D^2V(\xi_{n+1})\Delta X^{(2)}| &\leq C\|X_{n+1} - X_n\|^2 \\ &\leq C\gamma_{n+1}^2\Big(|h(X_n)|^2 + |\Delta M_{n+1}|^2 + |\rho_{n+1}|^2\Big) \end{aligned}$$

On a obtenu :

$$\mathbb{E}\Big(\Big|\frac{1}{2}D^2V(\xi_{n+1})\Delta X^{(2)}\Big| \;\Big|\; \mathscr{F}_n\Big) \leq C\gamma_{n+1}^2\Big(1 + V(X_n)\Big)$$

Comme $\Delta M_{n+1}$ est un incrément martingale, on a

$$\mathbb{E}(\langle \nabla V(X_n), \Delta M_{n+1}\rangle | \mathscr{F}_n) = 0$$

D'autre part :

$$\begin{aligned}
\mathbb{E}(\langle \nabla V(X_n), \rho_{n+1}\rangle | \mathscr{F}_n) &\leq |\nabla V(X_n)| \mathbb{E}(|\rho_{n+1}||\mathscr{F}_n) \\
&\leq |\nabla V(X_n)| \sqrt{\mathbb{E}(|\rho_{n+1}|^2 | \mathscr{F}_n)} \\
&\leq C|\nabla V(X_n)| \sqrt{\left(1+V(X_n)\right)} \\
&\leq C(1+V(X_n)).
\end{aligned}$$

Ainsi, on a la borne :

$$\begin{aligned}
& \mathbb{E}(V(X_{n+1})|\mathscr{F}_n) \\
\leq\ & V(X_n) - \gamma_{n+1}\langle \nabla V(X_n), h(X_n)\rangle + C\gamma_{n+1}^2(1+V(X_n)) \\
\leq\ & V(X_n)(1+C\gamma_{n+1}^2) - \gamma_{n+1}\langle \nabla V(X_n), h(X_n)\rangle + C\gamma_{n+1}^2
\end{aligned}$$

Posons alors avec les notations du Théorème 5.2.1 :

- $C\gamma_{n+1}^2 = \alpha_n$ qui est tel que le produit infini est fini ;
- $\gamma_{n+1}\langle \nabla V(X_n), h(X_n)\rangle = U_{n+1}$ est positif ;
- $C\gamma_{n+1}^2 = \beta_{n+1}$ est une série (d'espérance) convergente ;
- $V(X_n) = V_n$, qui satisfait à la récurrence du théorème 5.2.1.

Les conclusions du théorème de Robbins Siegmund sont :

1. $V(X_n) \underset{n\to+\infty}{\longrightarrow} V_\infty \in L^1$ presque surement,
2. $\sup_{n\in N} \mathbb{E}(V(X_n)) < +\infty$,
3. $\sum_{n\geq 1} \gamma_{n+1}\langle \nabla V(X_n), h(X_n)\rangle < +\infty$ presque surement.

Pour la dernière conclusion, on a de plus que :

$$\mathbb{E}(|X_{n+1}-X_n|^2) \leq \underbrace{C\gamma_{n+1}^2}_{\text{convergent series}} \left(1+\underbrace{\mathbb{E}(V(X_n))}_{\text{bounded}}\right).$$

Donc, $\sum_{n\geq 1} \mathbb{E}(|X_{n+1}-X_n|^2)$ est une série convergente et donc son terme général tend vers 0 (p.s.). ■

Notons que dans le théorème précédent, on ne garanti pas la **convergence des itérées**, seulement de la fonctionnelle : $V(X_n)$. Obtenir la convergence des $(X_n)_{n\geq 0}$, on a besoin d'une hypothèse supplémentaire.

**Corollaire 5.2.4** Convergence de Robbins Monro

On se place sous les hypothèse de la section 5.2.3, et on suppose de plus que

$$\{x \in \mathbb{R}^d \ ; \ \langle \nabla V(x), h(x)\rangle = 0\} = \{x^*\}.$$

Alors, on a :

1. $x^*$ est l'unique minimum de $V$ et l'unique zéro de $h$.
2. On a $X_n \underset{n\to+\infty}{\longrightarrow} x^*$ et $\langle \nabla V(X_n), h(X_n)\rangle \underset{n\to+\infty}{\longrightarrow} 0$ presque surement.
3. Soit $p > 0$ tel que $|x|^p \leq CV(x)^\rho$ pour un certain $\rho \in [0,1[$. Alors :

$$\mathbb{E}\left[\|X_n - x^*\|^p\right] \underset{n\to+\infty}{\longrightarrow} 0.$$

**Remarque** Sans cette hypothèse supplémentaire, on peut toujours garantir la convergence de $X_n \to X^*$, mais dans ce cas, $X^*$ reste aléatoire et son support est exactement l'ensemble :

$$\{x \in \mathbb{R}^d \ ; \ \langle \nabla V(x), h(x)\rangle = 0\}.$$

**Exercice 5.3** On note $sgn(x)$ la fonction qui à un réel $x$ associe son signe : $+1$ si $x$ est positif ou nul et $-1$ si $x$ est strictement négatif. On défini l'algorithme :

$$X_{n+1} = X_n - \gamma_{n+1} sgn(X_n) + \gamma_{n+1}\varepsilon_{n+1}, \quad X_0 \in \mathbb{R},$$

où $(\varepsilon_i)_{i\geq 1}$ est une suite de variables aléatoires *i.i.d.*, supposée centrée et de carré intégrable et la suite $(\gamma_k)_{k\geq 1}$ satisfait :

$$\sum_{k\geq 1} \gamma_k = +\infty, \quad \sum_{k\geq 1} \gamma_k^2 < +\infty.$$

1. Vérifiez que la fonction $V(x) = x^2 + 1$ satisfait aux hypothèses d'un théorème du cours.
2. En déduire que $(X_n)_{n\geq 1}$ converge (p.s.) et donner sa limite.

## 5.3 Exercices du Chapitre 5

**Exercice 5.4 — Le modèle de Black-Scholes discrétisé.** Le processus d'évolution temporelle d'une action boursière peut être modélisé par la donnée d'un processus défini par récurrence de la manière suivante : on pose $S_0 = s_0 > 0$ et pour tout $n \in \mathbb{N}$,

$$S_{n+1} = (1+\mu)S_n + \sigma S_n \varepsilon_{n+1},$$

où la suite $(\varepsilon_n)_{n\in\mathbb{N}^*}$ est une suite i.i.d. de loi de Rademacher de paramètre $1/2$ et où les paramètres réels $\mu$ et $\sigma$ (qui sont le taux d'actualisation et le coefficient de volatilité) satisfont l'inégalité $|\sigma| < 1+\mu$. On note $\lambda$ le réel strictement positif

$$\lambda = \sqrt{(1+\mu)^2 - \sigma^2},$$

et l'on considère la filtration $\mathscr{F}_0 = \{\emptyset, \Omega\}$ et $\mathscr{F}_n = \sigma(\varepsilon_1, \dots, \varepsilon_n)$ pour tout $n \in \mathbb{N}^*$.

1. Montrez que le processus $(S_n)_{n\in\mathbb{N}}$ est strictement positif, intégrable et adapté à la filtration $(\mathscr{F}_n)_{n\in\mathbb{N}}$.
2. Déterminez sa nature selon les valeurs de $\mu$.
3. Dans le cas où $\mu < 0$, montrez que le processus converge p.s. vers 0.
4. Calculez sa variance.
5. On définit le processus $(Z_n)_{n\in\mathbb{N}}$ par $Z_n = \ln(S_n)$ pour tout $n \in \mathbb{N}$. Déterminez la nature de $(Z_n)_{n\in\mathbb{N}}$ selon les valeurs du paramètre $\lambda$.
6. Déterminez la limite p.s. de $(Z_n)_{n\in\mathbb{N}}$ selon les valeurs de $\lambda$, et déduisez-en celle de $(S_n)_{n\in\mathbb{N}}$ (on se convaincra que $\mu < 0$ entraîne que $\lambda \in ]0,1[$).

**Exercice 5.5 — Récurrence nulle de la marche aléatoire symétrique.** Considérons $(S_n)_{n\in\mathbb{N}}$ la marche aléatoire symétrique sur $\mathbb{Z}$ et issue de 0, c'est-à-dire que $S_0 = 0$ et pour tout $n \in \mathbb{N}^*$,

$$S_n = \sum_{k=1}^{n} U_k,$$

où les $U_k$ sont i.i.d. de loi de Rademacher de paramètre $1/2$. On note également $\mathscr{F}_0 = \{\emptyset, \Omega\}$ et $\mathscr{F}_n = \sigma(U_1, \dots, U_n)$ pour tout $n \in \mathbb{N}^*$. Enfin on considère la v.a.

$$T = \inf\{n \geq 0 : S_n = a\},$$

correspondant au temps d'atteinte du niveau $a \in \mathbb{N}^*$. Cette chaîne de Markov étant récurrente, on admet dans cet exercice que $T$ est p.s. fini.

1. Étant donné $\theta$ un paramètre positif ou nul, montrez que le processus $(X_n^\theta)_{n\in\mathbb{N}}$ défini par
$$X_n^\theta = \frac{e^{\theta S_n}}{\cosh(\theta)^n}, \quad n \in \mathbb{N},$$
est une martingale.
2. Montrez que le processus arrêté $(X_{n\wedge T}^\theta)_{n\in\mathbb{N}}$ est borné et converge p.s.
3. Déduisez-en la valeur de l'espérance $\mathbb{E}[\cosh(\theta)^{-T}]$.
4. En dérivant cette relation en $\theta = 0$, montrez que $T$ n'est pas intégrable (traduisant la récurrence nulle de la chaîne).

**Exercice 5.6** Soit $\Phi$ une fonction de $\mathbb{R}$ dans $\mathbb{R}$, $C^2$ convexe telle que $\Phi''$ est bornée On fait l'hypothèse que $\Phi$ peut s'écrire sous la forme $\Phi(x) = \mathbb{E}[f(x,Z)]$ où $Z$ est une variable aléatoire à valeurs dans $\mathbb{R}^d$ et que $\Phi$ admet un unique minimum $x^*$. On considère l'algorithme défini par $X_0 = x_0 \in \mathbb{R}$ et

$$X_{n+1} = X_n - \gamma_{n+1} \frac{f(X_n + \rho_{n+1}, Z_{n+1}) - f(X_n - \rho_{n+1}, Z_{n+1})}{2\rho_{n+1}} \quad \forall n \geq 0,$$

où $\gamma_n = 1/n$, $\rho_n = 1/n^\alpha$ avec $\alpha > 1/2$ et $(Z_n)$ est une suite de v.a. i.i.d. de même loi que $Z$.

1. Montrer que pour tout $x \in \mathbb{R}$ et $\rho > 0$,
$$|\Phi(x+\rho) - \Phi(x-\rho) - 2\rho\Phi'(x)| \leq C\rho^2$$
où $C$ est une constante indépendante de $x$ et $\rho$.
2. On pose
$$A_{n+1} = \frac{f(X_n + \rho_{n+1}, Z_{n+1}) - f(X_n - \rho_{n+1}, Z_{n+1})}{2\rho_{n+1}}.$$
Montrer que
$$\mathbb{E}[A_{n+1}|\mathscr{F}_n] = \Phi'(X_n) + R_{n+1}$$
où $|R_{n+1}| \leq C\rho_{n+1}$.
3. On note $V_n = (X_n - x^*)^2$ et on admet dans la suite qu'il existe $C > 0$ tel que pour tout $z \in \mathbb{R}$,
$$|f(x,z) - f(x',z)| \leq C|x - x'|.$$
Montrer qu'il existe des constantes $C_1$ et $C_2$ positives telles que
$$\mathbb{E}[V_{n+1}|\mathscr{F}_n] \leq V_n(1 + C_1\gamma_{n+1}^2) + C_2\left(\gamma_{n+1}^2 + \rho_{n+1}^2\right) - 2\gamma_{n+1}(X_n - x^*)\Phi'(X_n).$$
4. En déduire que $(V_n)$ converge $p.s.$ vers $V_\infty \in L^1$ et que
$$\sum_{n\geq 1} \gamma_n (X_{n-1} - x^*)\Phi'(X_{n-1}) < +\infty \quad p.s.$$
5. En déduire que $X_n \to x^*$ $p.s.$

**Exercice 5.7 — Galton Watson.** Le processus de Galton-Watson est un exemple rudimentaire de processus de branchement, qui modélise la situation suivante. Soit $p = (p_k)_{k\in\mathbb{N}}$ une loi de probabilité sur $\mathbb{N}$ appelée loi de reproduction. Au temps $n = 0$, on part avec un seul individu. Puis naissent aléatoirement des enfants au temps $n = 1$, dont le nombre suit la loi $p$. Ensuite chaque enfant fait lui-même des enfants au temps $n = 2$, indépendamment de tout le reste, dont le nombre est supposé suivre la même loi $p$. Et ainsi de suite.

Le but de cet exercice est d'étudier la probabilité d'extinction de cette population.

Si $Z_n$ désigne le nombre d'individus à l'instant $n \in \mathbb{N}$, le processus $(Z_n)_{n\in\mathbb{N}}$ est alors défini par récurrence de la manière suivante : $Z_0 = 1$ et pour tout $n \in \mathbb{N}$,

$$Z_{n+1} = \sum_{k=1}^{Z_n} X_k^{(n+1)},$$

où la variable aléatoire $X_k^{(n+1)}$ représente le nombre d'enfants du k-ième individu à l'instant $n+1$ (on adopte la convention $\sum_\emptyset = 0$ de sorte que si $Z_n = 0$ alors $Z_{n+1} = 0$).

La famille $(X_k^{(n)})_{n,k}$ forment une suite i.i.d. de loi commune $p$, que l'on suppose de carré intégrable. De surcroît, pour éviter les cas particuliers, on fait l'hypothèse que $p_0 > 0$ et $p_0 + p_1 < 1$.

L'objectif est donc d'étudier la probabilité d'extinction de cette population, $\mathbb{P}(\tau_0 < +\infty)$, où

$$\tau_0 := \inf\{n \geq 0 : Z_n = 0\} \in \mathbb{N} \cup \{+\infty\}.$$

On note $m = \mathbb{E}[X_k^{(n)}]$, $\sigma^2 = \mathrm{Var}(X_k^{(n)})$ et enfin $(\mathscr{F}_n)_{n\in\mathbb{N}}$ désigne la filtration naturelle du processus $(Z_n)_{n\in\mathbb{N}}$.

1. Montrez que le processus $(Y_n)_{n\in\mathbb{N}}$ défini par $Y_n = Z_n/m^n$ est une martingale, et déduisez-en qu'elle converge p.s. vers une variable aléatoire $Y_\infty$ intégrable.
2. Déduisez-en que lorsque $m < 1$, on a $\mathbb{P}(\tau_0 < +\infty) = 1$.
3. On suppose dans les questions 3 à 7 que $m > 1$. Le but est de montrer que l'on a toujours $\mathbb{P}(\tau_0 < +\infty) < 1$. Montrez tout d'abord que $\mathbb{P}(\tau_0 < +\infty) > 0$.
4. Montrez que pour tout $n \in \mathbb{N}^*$,
$$\mathbb{E}[Z_n^2|\mathscr{F}_{n-1}] = \sigma^2 Z_{n-1} + m^2 Z_{n-1}^2.$$
5. Déduisez-en que pour tout $n \in \mathbb{N}^*$,
$$\mathbb{E}[Z_n^2] = \sigma^2 m^{n-1} \times \frac{m^n - 1}{m - 1} + m^{2n}.$$
6. Déduisez-en que $(Y_n)_{n\in\mathbb{N}}$ est une martingale bornée dans $L^2$ et déterminez $\mathbb{E}[Y_\infty^2]$.
7. Montrez que $\mathbb{P}(Y_\infty = 0) < 1$ puis que $\mathbb{P}(\tau_0 < +\infty) < 1$.
8. Dans la suite, on suppose que $m = 1$, i.e. que chaque individu a en moyenne un enfant. On va montrer que $\mathbb{P}(\tau_0 < +\infty) = 1$. On note $Z_\infty$ la limite p.s. de $Z_n$ (comme $m = 1$ on a $Z_n = Y_n$). Montrez que
$$\{Z_\infty = k\} = \bigcup_{n\geq 0} A_n \quad \text{où} \quad A_n = \bigcap_{p\in\mathbb{N}} \{Z_{n+p} = k\}.$$
9. Montrez que $(A_n)_{n\in\mathbb{N}}$ est une suite croissante pour l'inclusion puis que
$$\mathbb{P}(Z_\infty = k) = \lim_{n\to+\infty} \mathbb{P}(A_n).$$
10. Supposons que $k$ soit différent de 0 et considérons
$$A_n^p = \bigcap_{i=0}^{p} \{Z_{n+i} = k\}.$$
Montrez qu'il existe un nombre $\alpha_k \in ]0,1[$ tel que pour tous $n, p \in \mathbb{N}^*$, on ait
$$\mathbb{P}(A_n^p) = \alpha_k \mathbb{P}(A_n^{p-1}).$$
11. Déduisez-en que si $k \neq 0$, alors $\mathbb{P}(A_n) = 0$ pour tout $n \in \mathbb{N}$ puis que les seules valeurs possibles de $Z_\infty$ sont 0 et $+\infty$.
12. Concluez.

**Exercice 5.8** Soit $(X_n)_{n\in\mathbb{N}}$ une suite de variables aléatoires indépendantes et identiquement distribuées, et $(\mathscr{F}_n)_{n\in\mathbb{N}}$ la filtration naturelle des $(X_n)_{n\in\mathbb{N}}$. Soient $g_0$ et $g_1$ deux densités de probabilités strictement positives ($g_0 \neq g_1$). Le but de cet exercice est d'étudier le test d'hypothèses suivant :
$(H_0)$ Les $X_i$ sont à densité $g_0$,
$(H_1)$ Les $X_i$ sont à densité $g_1$.
On considère le rapport de vraisemblance :
$$L_n = \prod_{i=1}^{n} \frac{g_0(X_i)}{g_1(X_i)}.$$
On se place sous l'alternative $H_1$. Le but est de montrer que $L_n \underset{n\to+\infty}{\longrightarrow} 0$ (ps).

1. Montrer que $(L_n)_{n\in\mathbb{N}}$ est une martingale.
2. En déduire que $(L_n)_{n\in\mathbb{N}}$ converge vers $L_\infty$.
3. On considère la log-vraisemblance : $\log(L_n)$, montrer que $\log(L_n) = \sum_{i=1}^{n} \log\left(\frac{g_0(X_i)}{g_1(X_i)}\right)$.

4. En déduire que $\frac{\log(L_n)}{n} \underset{n\to+\infty}{\longrightarrow} m$, où $m = \mathbb{E}\left[\log\left(\frac{g_0(X_1)}{g_1(X_1)}\right)\right]$.
5. Montrer que $m \leq 0$, et en déduire que $L_\infty = 0$.
6. Expliquer la procédure de test.

**Exercice 5.9** Soit $(\xi_n)_{n\geq 0}$ une suite de variables aléatoires indépendantes de loi exponentielle de paramètre 1. On pose

$$M_n = \xi_1 + \cdots + \xi_n,$$

et on note

$$X_n = \frac{n!}{(1+M_n)^{n+1}} e^{M_n}.$$

1. Montrer que $(X_n)_{n\geq 0}$ est une martingale dans la filtration $\mathscr{F}_n = \sigma(\xi_k; k \leq n)$.
   *Indication :* On rappelle que la densité de la loi exponentielle de paramètre 1 est $f(x) = e^{-x}\mathbf{1}_{\{x\geq 0\}}$.
2. Que pouvez vous dire de $(X_n)$ lorsque $n$ tends vers l'infini ?

**Exercice 5.10** (Quantification Optimale)

**Problématique :** Pour une loi de probabilité donnée $\mu$ sur $\mathbb{R}^d$, l'idée de base de l'intégration numérique consiste à choisir des points $x_1, \ldots, x_N$ de $\mathbb{R}^d$ et une partition $C_1, \ldots, C_N$ de $\mathbb{R}^d$ telle que $x_i \in C_i$ et à approcher $I(f) = \int f(x)\mu(dx)$ par

$$I_N(f) = \sum_{i=1}^{N} \int_{C_i} f(x_i)\mu(dx) = \sum_{i=1}^{N} f(x_i)\mathbb{P}(X \in C_i)$$

où $X$ a pour loi $\mu$.
Pour un nombre $N$ donné, est-il possible d'optimiser (dans un sens à définir dans la suite), le choix des points $x_1, \ldots, x_N$ et de la partition $C_1, \ldots, C_N$, *i.e.* de minimiser l'erreur $I_N(f) - I(f)$ indépendemment de $f$ ?
**Distorsion, Mosaïque de Voronoï :** Soit $\mu$ une loi sur $\mathbb{R}^d$ et $N \in \mathbb{N}$. La fonction $D_N^\mu : (\mathbb{R}^d)^N \to \mathbb{R}$ définie par

$$D_N^\mu(x_1, \ldots, x_N) = \int \min_{1\leq i\leq N} |x_i - u|^2 \mu(du)$$

est appelée la *distorsion*. Dans la suite, on supposera pour simplifier $\mu$ absolument continue par rapport à la mesure de Lebesgue sur $\mathbb{R}^d$. **definition** On note $x = (x_1, \ldots, x_N)$. On appelle *mosaïque de Voronoï* associée à $x$ la *p.s.* partition $(C_i(x))_{i\in\{1,\ldots,N\}}$ de $\mathbb{R}^d$ définie pour tout $i \in \{1, \ldots, N\}$ :

$$C_i(x) = \{u \in \mathbb{R}^d, |u - x_i| < \min_{j\neq i} |u - x_j|\}.$$

Ainsi,

$$D_N^\mu(x_1, \ldots, x_N) = \sum_{k=1}^{N} \int_{C_i(x)} |x_i - u|^2 \mu(du).$$

1. On considère $x_1 = (0,0)$, $x_2 = (0,1)$, $x_3 = (1,0)$, $x_4 = (-1,0)$ et $x_5 = (0,-1)$. Dessiner la mosaïque de Voronoï associée à $x = (x_1, \ldots, x_5)$.
2. Soit $f : \mathbb{R}^d \to \mathbb{R}$ et $f$ Lipschitzienne. Alors,

$$\left| \int_{\mathbb{R}^d} f d\mu - \sum_{i=1}^{N} f(x_i)\mathbb{P}(X \in C_i(x)) \right| \leq [f]_1 \sqrt{D_N^\mu(x)}.$$

**Quantification Optimale :** En accord avec la question précédente, l'idée de la quantification optimale est de minimiser $x \to D_N^\mu(x)$, *i.e.* chercher les *quantifieurs optimaux* $x_1^*, \ldots, x_N^*$ associés à une loi donnée et à un nombre de points donnés. Ce problème étant complètement insoluble explicitement (sauf dans des cas très particuliers), il nécessite la mise en place d'algorithmes : l'approximation stochastique est une solution adaptée.

3. On suppose que $d = 1$. Montrer que $D_N^\mu$ est différentiable et que :

$$\frac{\partial D_N^\mu(x)}{\partial x_i} = 2\int_{C_i(x)} (x_i - u)\mu(du).$$

   *Indication :* On rappelle que $\mu$ est supposée absolument continue, elle ne charge donc pas le bord de $C_i(x)$.

4. Justifier qu'un algorithme stochastique permettant (a priori) d'approcher les quantifieurs optimaux est $X^k = (X_1^k, \ldots, X_N^k)$ avec la dynamique

$$X_i^{k+1} = X_i^k - 2\gamma_{k+1}(X_i^k - U^{k+1})1_{U^{k+1}\in C_i(X^k)}, \quad \forall i = 1, \ldots, N,$$

   avec $(U^k)$ suite de v.a. i.i.d. de loi $\mu$.

**Le cas de la loi uniforme :** Soit $U$ de loi uniforme sur $[0,1]$ et $N \in \mathbb{N}^*$. Sans perte de généralité, on peut supposer $0 \le x_1 < x_2 < \ldots < x_N$.

5. Calculer explicitement $\nabla D_N^\mu(x)$.
6. En déduire que $\nabla D_N^\mu(x)$ s'annule en un unique point $x^* = (x_1^*, \ldots, x_N^*)$ et que $x_i^* = (\frac{2i-1}{2N})_{1\le i\le N}$. Pourquoi ce résultat est-il logique ? N.B. On peut prouver que $x^*$ est l'unique minimum de $D_n^\mu$.

**Le cas de la loi Normale :**

Obtenir une bonne approximation des quantifieurs optimaux d'une loi prend du temps et n'a un intérêt que si l'on doit ensuite faire beaucoup de calculs d'espérances de fonctions sous cette loi. C'est bien sûr le cas de la loi normale (ex : calcul de prix en finance).

On admet dans cet exercice que $D_N^\mu$ admet un unique extremum $x^*$ et que cet unique extremum est un minimum global (En dimension 1, on peut montrer que c'est vrai dès que la densité est log-concave, ce qui est le cas pour la loi normale).

7. On note

$$\Pi_i^n = \frac{1}{n}\sum_{k=1}^{n} 1_{U^k\in C_i(X^{k-1})}.$$

   Expliquer rapidement pourquoi si $X_n \to x^*$ *p.s.*,

$$\Pi_i^n \xrightarrow{n\to+\infty} \mathbb{P}(X \in C_i(x^*)) \quad p.s.$$

   Les $\mathbb{P}(X \in C_i(x^*))$ sont appelés les paramètres compagnons.

8. Proposer un algorithme récursif pour approcher à la fois les quantifieurs optimaux ainsi que les paramètres compagnons.

   **Question TP :**

   On se propose de calculer le prix d'un call Européen d'échéance $T$ et de strike $K$ dans le modèle de Black-Scholes de paramètres $r$ et $\sigma$. On rappelle que le prix d'une telle option est alors :

$$C(S_0, r, \sigma, K, T) = e^{-rT}\mathbb{E}\left[\left(S_0 \exp((r - \frac{\sigma^2}{2})T + \sigma\sqrt{T}W) - K\right)_+\right]$$

où $W \sim \mathcal{N}(0,1)$. On rappelle également que ce prix possède une expression explicite donnée par la formule de Black-Scholes :

$$C(S_0, r, \sigma, K, T) = S_0 \Phi(d_1) - K e^{-rT} \Phi(d_2)$$

où $\Phi$ est la fonction de répartition de $W \sim \mathcal{N}(0,1)$, et

$$d_1 = \frac{1}{\sigma\sqrt{T}} \left( \log(\frac{S_0}{K}) + (r - \frac{\sigma^2}{2})T \right)$$

et

$$d_2 = d_1 + \sigma\sqrt{T}.$$

(a) Écrire une fonction permettant de calculer une approximation du prix du call européen dans le modèle de Black-Scholes en fonction de $S_0$, $K$, $T$, $r$, $\sigma$, des quantifieurs optimaux et des paramètres compagnons.

(b) Tester votre fonction avec les quantifieurs et les paramètres compagnons estimés puis avec les quantifieurs et paramètres références.

**Remarque :**
On peut généraliser la distorsion (et donc le critère que l'on souhaite minimiser) à :

$$D_N^\mu(x_1, \ldots, x_N) = \int \min_{1 \le i \le N} |x_i - u|^\alpha \mu(du).$$

Dans ce cas, on peut montrer que

$$\frac{\partial}{\partial x_i} D_N^\mu(x_1, \ldots, x_N) = \alpha \int_{C_i(x)} |x_i - u|^{\alpha-1} \frac{x_i - u}{||x_i - u||} \mu(du).$$

La preuve de la différentiabilité de $D_N^\mu(x_1, \ldots, x_N)$ dans ce cas est la Proposition 9 dans :

Pagès, G. (1998). A space quantization method for numerical integration, *Journal of computational and applied mathematics*, 89(1), 1-38.

# 6. Autres théorèmes de convergence

Dans ce chapitre, nous étudions la convergence des martingales dans $L^1$. Pour cela, nous avons besoin de généraliser le théorème de convergence dominée. Nous prouverons également la loi des grands nombres et le théorème central limite pour les martingales.

## 6.1 Le cas de la convergence $L^1$

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

**Définition 6.1.1** Une famille de v.a. $(X_i)_{i\in I}$ est dite uniformément intégrable si

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

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

**Proposition 6.1.1** Si une suite $(X_n)_{n\in\mathbb{N}}$ est uniformément intégrable et converge p.s. vers une v.a. $X_\infty$, alors $X_\infty \in L^1$ et la convergence a aussi lieu dans $L^1$.

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

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

Tout d'abord par le lemme de Fatou,

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

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

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

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

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

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

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

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

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

**Proposition 6.1.2** Les assertions suivantes sont vérifiées.

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

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

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

alors elle est uniformément intégrable.

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

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

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

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

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

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

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

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

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

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

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

**Théorème 6.1.3** Soit $(M_n)_{n\in\mathbb{N}}$ une martingale. Alors les assertions suivantes sont équivalentes :

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

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

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

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

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

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

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

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

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

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

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

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

$(ii) \Rightarrow (iii)$ : la martingale $(M_n)_{n\in\mathbb{N}}$ étant uniformément intégrable, elle est bornée dans $L^1$ par la proposition 6.1.2, et donc elle converge p.s. par le théorème 5.1.4. Une fois la convergence p.s. établie, la convergence dans $L^1$ est immédiate par la proposition 6.1.1.

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

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

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

On en déduit par le théorème 5.1.5 que toute martingale bornée dans un $L^p$, où $p > 1$, est fermée.

## 6.2 Loi des grands nombres et théorème central limite pour les martingales

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

**Lemme 6.2.1 — Kronecker.** Soit $(a_n)_{n\in\mathbb{N}}$ une suite de nombres strictement positifs tendant vers l'infini à l'infini et $(x_n)_{n\in\mathbb{N}}$ une suite de nombres réels. Si la série de terme général $x_n/a_n$ converge, alors

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

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

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

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

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

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

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

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

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

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

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

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

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

d'où la conclusion. ■

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

**Théorème 6.2.2 — LGN.** Soit $(M_n)_{n\in\mathbb{N}}$ une martingale de carré intégrable. Si p.s., $[M,M]_\infty = +\infty$, alors on a la convergence p.s. suivante :

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

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

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

On remarque que $(X_n)_{n\in\mathbb{N}}$ est la transformation prévisible de la martingale $(M_n)_{n\in\mathbb{N}}$ par le processus prévisible et borné $(1/(1+[M,M]_n))_{n\in\mathbb{N}}$. Ainsi, par le théorème 3.2.1, le processus $(X_n)_{n\in\mathbb{N}}$ est une martingale de carré intégrable. De plus, le théorème 3.2.2 nous donne la variation quadratique de la transformation prévisible : pour tout $m \in \mathbb{N}^*$,

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

la variation quadratique étant un processus croissant. Ainsi, par le théorème 5.1.3, la martingale $(X_n)_{n\in\mathbb{N}}$ est p.s. convergente. Enfin, comme on a supposé que $[M,M]_\infty = +\infty$ p.s., on peut appliquer le lemme de Kronecker aux suites aléatoires $a_n = 1+[M,M]_n$ et $x_n = M_n - M_{n-1}$ ($x_0 = M_0$) et l'on obtient que la suite

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

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

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

**Théorème 6.2.3** Soit $(X_n)_{n\in\mathbb{N}^*}$ une suite de v.a. indépendantes (non nécessairement de même loi) et de carré intégrable. Posons $m_n = \mathbb{E}[X_n]$ et $\sigma_n^2 = \mathrm{Var}(X_n)$. Si la série numérique de terme général $(\sigma_n)_{n\in\mathbb{N}^*}$ diverge, alors on a la convergence p.s. suivante :

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

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

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

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

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

Enfin, le théorème 6.2.2 achève la démonstration de cette LGN. ■

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

**Théorème 6.2.4 — TCL.** Soit $M$ une martingale dont les accroissements satisfont la propriété de convergence p.s. suivante :

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

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

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

Alors on a les convergences en loi

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

**Remarque** La borne uniforme sur les moments de $M_n - M_{n-1}$ est une condition très restrictive, mais rend la preuve plus abordable. Il est possible de remplacer cette condition par la condition plus faible suivante, $\forall \varepsilon > 0$ :

$$\lim_{n\to+\infty} \frac{1}{a_n} \sum_{k=1}^{n} \mathbb{E}\left(|M_k - M_{k-1}|^2 \mathbf{1}_{\{|M_k - M_{k-1}| \geq \varepsilon\sqrt{a_n}\}} \middle| \mathscr{F}_{k-1}\right) = 0 \text{ en probabilité.}$$

où $(a_n)_{n\in\mathbb{N}}$ est la même suite de l'énoncé du théorème précédent. Cet énoncé est chez Bercu et Chaffai [**bercu**], mais même eux font la preuve dans le cas borné ! Ils donnent toute fois une référence pour le cas général.

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

La démonstration se fait en plusieurs étapes.

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

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

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

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

d'où l'inégalité p.s.

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

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

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

Par la formule de Taylor-Lagrange, on a

$$\begin{aligned}\phi_{x,y}(t) &= \phi_{x,y}(0) + t\,\phi'_{x,y}(0) + \frac{t^2}{2}\,\phi''_{x,y}(0) + \frac{t^3}{3!}\,R_{x,y}(t)\\ &= 1 + itx + \frac{t^2}{2}\,(y - x^2) + \frac{t^3}{6}\,R_{x,y}(t),\end{aligned}$$

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

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

avec $C > 0$ une constante déterministe dont la valeur n'a aucune importance (dans la suite, on notera $C$ une telle constante, qui peut changer de ligne à ligne).

∘ Étape 3 : appliquons ceci à la suite de fonctions aléatoires $(Z_n)_{n\in\mathbb{N}^*}$ définie par

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

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

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

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

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

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

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

on a

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

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

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

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

∘ Étape 5 : par ailleurs on a
$$\begin{aligned}\mathbb{E}\left[\exp\left(it\frac{M_n}{\sqrt{n}}+\frac{t^2\sigma^2}{2}\right)\right]-1 &= \mathbb{E}\left[\exp\left(it\frac{M_n}{\sqrt{n}}\right)\left(\exp\left(\frac{t^2\sigma^2}{2}\right)-\exp\left(\frac{t^2[M,M]_n}{2n}\right)\right)\right]\\ &\quad +\mathbb{E}\left[Y_n\left(\frac{t}{\sqrt{n}}\right)\right]-1.\end{aligned}$$
Montrons que la première espérance dans le terme de droite tend vers 0 lorsque $n$ tend vers l'infini. On a
$$\left|\exp\left(it\frac{M_n}{\sqrt{n}}\right)\left(\exp\left(\frac{t^2\sigma^2}{2}\right)-\exp\left(\frac{t^2[M,M]_n}{2n}\right)\right)\right| = \left|\exp\left(\frac{t^2\sigma^2}{2}\right)-\exp\left(\frac{t^2[M,M]_n}{2n}\right)\right|,$$
qui tend p.s. vers 0 lorsque $n$ tend vers l'infini, la suite $[M,M]_n/n$ tendant p.s. vers $\sigma^2$ par hypothèse. De plus, comme la suite $[M,M]_n/n$ est bornée par 1 d'après l'étape 1, on a
$$\left|\exp\left(\frac{t^2\sigma^2}{2}\right)-\exp\left(\frac{t^2[M,M]_n}{2n}\right)\right| \leq \exp\left(\frac{t^2\sigma^2}{2}\right)+\exp\left(\frac{t^2}{2}\right),$$
cette dernière quantité étant bien intégrable car déterministe. Le théorème de convergence dominée et l'étape 4 nous donnent alors que
$$\lim_{n\to+\infty} \mathbb{E}\left[\exp\left(it\frac{M_n}{\sqrt{n}}+\frac{t^2\sigma^2}{2}\right)\right] = 1,$$
c'est-à-dire que
$$\lim_{n\to+\infty} \mathbb{E}\left[\exp\left(it\frac{M_n}{\sqrt{n}}\right)\right] = \exp\left(-\frac{t^2\sigma^2}{2}\right).$$
Enfin, le théorème de convergence de Lévy (la convergence des fonctions caractéristiques est équivalente à la convergence en loi) nous permet d'obtenir le résultat désiré, à savoir la convergence en loi de la suite $(M_n/\sqrt{n})_{n\in\mathbb{N}^*}$ vers une v.a. gaussienne centrée et de variance $\sigma^2$. ■

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

**Lemme 6.2.5 — Lemme de Slutsky.** Soient $(X_n)_{n\in\mathbb{N}_*}$ et $(Y_n)_{n\in\mathbb{N}_*}$ deux suites de v.a. convergeant en loi respectivement vers un nombre $c \in \mathbb{R}$ et une v.a. $Y$. Alors

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

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

L'hypothèse de moment pour le théorème central limite est très restrictive :

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

Dans les faits, on peut la remplacer par **la condition de Lindeberg** :

**Théorème 6.2.6 — Condition de Lindeberg.** Supposons que $\forall \varepsilon > 0$, on ait la convergence en probabilité :

$$\frac{1}{\mathbb{E}[M_n^2]} \sum_{k=1}^{n} \mathbb{E}\left[\left(M_k - M_{k-1}\right)^2 \mathbf{1}_{|M_k - M_{k-1}| > \varepsilon\sqrt{\mathbb{E}[M_n^2]}} \Big| \mathscr{F}_{k-1}\right] \underset{n\to+\infty}{\longrightarrow} 0.$$

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

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

Alors on a les convergences en loi

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

**Remarque**

La loi des Grands nombres et le théorème central limite martingale ont bien plus d'applications qu'il n'y parait. En effet, ce théorème permet, en particulier de se passer de l'hypothèse d'indépendance pour une marche aléatoire. Considérons l'exemple d'un sondage. Chaque personne interrogée répond par oui (qu'on notera $X_n = 1$) ou par zéro (qu'on notera par $X_n = 0$) à une question donnée. On compte alors les opinions favorables : $S_n = \sum_{k=1}^n X_k$.
Notons qu'alors, si l'on interroge les membres de différents milieu social, les réponses ne seront pas nécessairement de même loi.
Si $p_k$ est la vraie probabilité que la personne $k$ soit favorable alors

$$M_n = \sum_{k=1}^{n} \left(X_k - p_k\right) \text{ est une martingale.}$$

On remarquera qu'on **n'a pas besoin que les $X_i$ soient de même loi.**
Ainsi, il restera vrai que $\frac{M_n}{n}$ converge (p.s.) vers 0, et que $\frac{M_n}{\sqrt{n}}$ est asymptotiquement Gaussien, meme si les observations ne sont pas indépendantes.

Le théorème de Robbins-Monro 5.2.3 peut être vu comme la loi des grand nombre, pour ces algorithmes. On peut alors s'intéresser au théorème central limite. Comme dans le cas classique, une condition de moment supplémentaire est nécessaire. Notons que ce résultat est en fait plus général, comme on l'illustrera dans le paragraphe suivant : il est possible de montrer la convergence des trajectoires d'un algorithme de Robbins-Monro vers la solution d'une équation différentielle stochastique.

**Théorème 6.2.7 — Théorème central limite pour Robbins Monro.** Soit l'algorithme de Robbins Monro :

$$\theta_{n+1} = \theta_n - \gamma_{n+1} H(\theta_n, \xi_{n+1}), \quad \theta_0 \in \mathbb{R}^d.$$

On se place sous les hypothèses de la section 5.2.4 de sorte qu'on ait la convergence presque sûre $\theta_n \underset{n\to+\infty}{\longrightarrow} \theta^*$. On suppose de plus que :

1. $\exists \delta > 0$, tel que $\sup_{n\geq 0} \mathbb{E}[|H(\theta_n, \xi_{n+1})|^{2+\delta}] < C$.
2. La matrice suivante est définie positive :

$$\Sigma_H(\theta^*) = \mathbb{E}\left[\left(h(\theta^*) - H(\theta^*,\xi)\right)\left(h(\theta^*) - H(\theta^*,\xi)\right)^T\right]$$

3. Le pas est de la forme $\gamma_n = \frac{\alpha}{\beta+n}$, avec $\alpha, \beta > 0$, et $\alpha > \frac{1}{2Re(\lambda_{min})}$, où $Re(\lambda_{min})$ désigne la partie réelle de la plus petite valeur propre de $Dh(\theta^*)$.

Alors, on a la convergence en loi suivante :

$$\sqrt{n}(\theta_n - \theta^*) \Rightarrow \mathcal{N}(0, \alpha\Sigma),$$

où

$$\Sigma = \int_0^{+\infty} e^{-\left(Dh(\theta^*) - \frac{I_d}{2\alpha}\right)s} \Sigma_H(\theta^*) e^{-\left(Dh(\theta^*) - \frac{I_d}{2\alpha}\right)s} ds.$$

**Remarque** En dimension 1, $\lambda_{min} = h'(\theta^*)$, $\Sigma_H(\theta^*) = \mathrm{Var}[H(\theta^*,\xi)]$. Ainsi,

$$e^{-\left(Dh(\theta^*) - \frac{I_d}{2\alpha}\right)s} = e^{-(h'(\theta^*) - \frac{1}{2\alpha})s}.$$

On considère un algorithme de Robbins-Monro :

$$\theta_{n+1} = \theta_n - \gamma_{n+1} H(\theta, X_{n+1}), \quad \theta_0 \in \mathbb{R}^d,$$

avec $(\gamma_k)_{k\geq 1}$ une suite de réels positifs satisfaisant (5.1) et $(X_k)_{k\geq 1}$ des copies indépendantes de même loi. Cet algorithme approche le zéro $\theta^*$ d'une fonction $h : \mathbb{R} \to \mathbb{R}$, qui peut s'écrire sous la forme d'une espérance :

$$h(\theta) = \mathbb{E}[H(\theta, X)]$$

On note $(\bar{\theta}_t)_{t\geq 0}$ la solution de l'équation différentielle

$$\begin{cases} \frac{d}{dt}\bar{\theta}_t = -h(\bar{\theta}_t), \\ \bar{\theta}_0 \in \mathbb{R}^d. \end{cases}$$

Alors, l'algorithme $(\theta_n)_{n\geq 1}$ a tendance à suivre les solutions de l'EDO $(\bar{\theta}_t)_{t\ \geq 0}$. L'heuristique est la suivante. Prenons pour simplifier un pas constant $\gamma_k = \gamma > 0$ :

$$\begin{cases} \theta_{n+1} = \theta_n - \gamma H(\theta_n, \eta_{n+1}) \\ \theta_0 \in \mathbb{R}^d. \end{cases}$$

Soit alors $N > 0$ **très grand**, de sorte que $\theta_{n+N}$ soit proche de la stationnarité (les itérées ne bougent plus beaucoup). On a alors :

$$\begin{aligned} \theta_{n+N} &= \theta_n - \sum_{i=0}^{N-1} \gamma H(\theta_{i+n}, \eta_{n+i+1}) \approx \theta_n - \gamma \sum_{i=0}^{N-1} H(\theta_n, \eta_{n+i+1}) \\ &= \theta_n - \gamma N \frac{1}{N} \sum_{i=0}^{N-1} H(\theta_n, \eta_{n+i+1}) \approx \theta_n - N\gamma \cdot h(\theta_n). \end{aligned}$$

On pose alors $N\gamma = t_N$, on a :

$$\theta_{n+N} \approx \theta_n - t_N \cdot h(\theta_n).$$

Cette équation n'est autre que le schéma d'Euler pour l'EDO : $(\bar{\theta}_t)_{t\geq 0}$ :

$$\frac{d}{dt}\bar{\theta}_t = -h(\bar{\theta}_t),\ \bar{\theta}_0 = \theta_0,\ \text{with} : h(\theta) = E[H(\theta,\eta)].$$

Ainsi, les fluctuations de l'algorithme se considèrent autour des solutions de l'EDO.

**Théorème 6.2.8 — Comparaison Algorithme, Équation Différentielle..** On fixe un horizon de temps $T > 0$ et on note toujours $t_n = \gamma_1 + \cdots + \gamma_n$. Supposons que $h$ est Lipschitz, on pose

$$\gamma = \max_{n;\ t_n \leq T} \gamma_n.$$

Alors, $\forall \varepsilon > 0$, lorsque $\gamma$ est assez petit, on a :

$$\mathbb{P}\left(\sup_{n;\ t_n\leq T} \|\theta_n - \bar{\theta}_{t_n}\| > \varepsilon\right) \leq C(T,\gamma) \underset{\gamma\to 0}{\longrightarrow} 0.$$

■ **Exemple 6.1** Supposons qu'on souhaite minimiser

$$V(\theta) = \mathbb{E}[(X-\theta)^2],\ \text{où } X \sim \mathcal{N}(0,1).$$

On annule la dérivée : $V'(\theta) = -2\mathbb{E}(X-\theta)$ ce qui nous donne l'algorithme suivant :

$$\theta_{n+1} = \theta_n + 2\gamma_{n+1}(X_{n+1}-\theta_n) \underset{n\to+\infty}{\longrightarrow} 0.$$

L'EDO associée à cet algorithme est le suivant : :

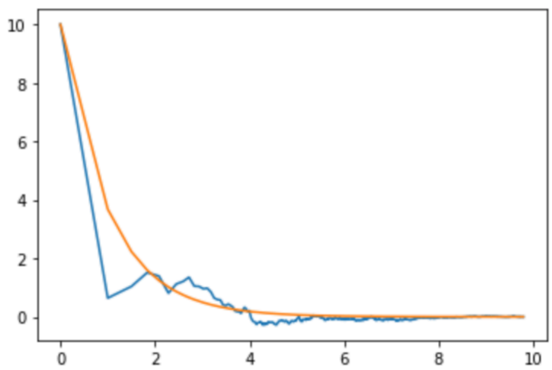
$$\begin{cases} \frac{d}{dt}\bar{\theta}_t = -h(\bar{\theta}_t) = -\theta_t, \\ \bar{\theta}_0 \in \mathbb{R}^d. \end{cases} \quad \Rightarrow \quad \bar{\theta}_t = \theta_0 e^{-t}.$$

La simulation de cet algorithme donne :

In [98]:
```
N=10000
theta=np.zeros(N)
T=np.zeros(N)
theta[0]=10
for i in range(N-1):
    x=np.random.normal(0,1)
    theta[i+1] = theta[i] + 1/(i+1)*(2*x - theta[i])
    T[i+1]=T[i]+1/(i+1)


plt.plot(T,theta)
plt.plot(T,theta[0]*np.exp(-T))
```

Out[98]: [<matplotlib.lines.Line2D at 0x7ffe456b51f0>]




■

## 6.3 Exercices du Chapitre 6

**Exercice 6.1** On peut se demander si les performances sportives seront toujours battues et si oui, à quel rythme. Afin de modéliser cette situation, on considère une suite de variables aléatoires $(R_n)_{n\in\mathbb{N}^*}$, chaque $R_k$ correspondant au rang de la $k$-ième performance parmi les $k$ premières. On les suppose indépendantes mais pas de même loi, chaque $R_k$ suivant la loi uniforme sur $\{1,2,\ldots,k\}$, c'est-à-dire

$$\mathbb{P}(R_k = i) = \frac{1}{k}, \quad i \in \{1,2,\ldots,k\}.$$

On dit qu'il se produit un record à l'instant $k$ si $R_k = 1$. On s'intéresse au comportement asymptotique des suites $(Z_n)_{n\in\mathbb{N}}$ et $(M_n)_{n\in\mathbb{N}}$ définies par $Z_0 = M_0 = 0$ et pour tout $n \in \mathbb{N}^*$,

$$Z_n = \sum_{k=1}^{n} 1_{\{R_k=1\}}, \quad \text{et} \quad M_n = Z_n - \sum_{k=1}^{n} \frac{1}{k}.$$

On note $\mathscr{F}_0 = \{\emptyset, \Omega\}$ et $\mathscr{F}_n = \sigma(R_1,\ldots,R_n)$ la filtration engendrée par le processus $(R_n)_{n\in\mathbb{N}^*}$.

1. Calculez pour tout $n \in \mathbb{N}^*$ la variance de $Z_n$.
2. Montrez que $(M_n)_{n\in\mathbb{N}}$ est une martingale de carré intégrable.
3. Calculez sa variation quadratique.
4. Montrez que $M_n/\ln(n)$ tend p.s. vers 0 lorsque $n$ tend vers l'infini.
   **Indication :** utilisez l'équivalent
   $$\sum_{k=1}^{n} \frac{1}{k} \underset{n\to+\infty}{\simeq} \ln(n).$$
5. Déduisez-en la convergence p.s. de $Z_n/\ln(n)$ vers une limite que l'on déterminera.

**Exercice 6.2** Soit $(X_n)_{n\in\mathbb{N}^*}$ une suite de v.a. indépendantes de loi de Rademacher de paramètre $1/2$ et considérons la somme définie par $M_0 = 0$ et pour tout $n \in \mathbb{N}^*$,

$$M_n = \sum_{k=1}^{n} k^{\alpha} X_k,$$

où $\alpha > 0$. On note également $(\mathscr{F}_n)_{n\in\mathbb{N}}$ la filtration naturelle de la suite $(X_n)_{n\in\mathbb{N}}$ avec de surcroît $\mathscr{F}_0 = \{\emptyset, \Omega\}$.

1. Montrez que le processus $(M_n)_{n\in\mathbb{N}}$ est une martingale de carré intégrable.
2. Calculez la variation quadratique de cette martingale et vérifiez que
   $$\lim_{n\to+\infty} \frac{[M,M]_n}{n^{2\alpha+1}} = \frac{1}{2\alpha+1}.$$
   **Indication :** on utilisera une comparaison somme - intégrale.
3. Déduisez-en que p.s. $\lim_{n\to+\infty} n^{-(2\alpha+1)} M_n = 0$.
4. On suppose que $\alpha \in ]0,1/6[$. Montrez que $(M_n)_{n\in\mathbb{N}}$ satisfait à un TCL dont on précisera la vitesse en fonction de $\alpha$.

# 7. Une ouverture vers les statistiques

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

**Définition 7.0.1** On appelle vraisemblance du modèle la fonction

$$L_n : (x_1, \ldots, x_n, \theta) \mapsto L_n(x_1, \ldots, x_n, \theta)$$

définie :

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

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

La v.a. obtenue en appliquant la fonction $(x_1, \ldots, x_n) \mapsto \operatorname{argmax}_{\theta \in \Theta} L_n(x_1, \ldots, x_n, \theta)$ au $n$-échantillon $(X_1, \ldots, X_n)$ s'appelle l'estimateur du maximum de vraisemblance (EMV) du paramètre $\theta$. On le note $\hat{\theta}_n$.

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

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

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

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

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

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

Ainsi, soit $(U_n)_{n\geq 1}$ une suite i.i.d. de loi commune $\mathcal{N}(0,1)$ et considérons la suite $(X_n)_{n\geq 0}$ définie par récurrence par $X_0 = x \in \mathbb{R}^*$ et

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

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

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

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

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

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

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

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

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

et le $\theta$ pour lequel cette quantité s'annule vaut

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

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

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

où $M$ est défini par la suite

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

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

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

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

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

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

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

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

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

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

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

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

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

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

# 8. Annales des années passées

**Exercice 8.1** On considère $(X_n)_{n\in\mathbb{N}^*}$ une suite i.i.d. de variables aléatoires de Rademacher de paramètre $p \in ]0,1/2[$. On définit le processus $(S_n)_{n\in\mathbb{N}}$ par $S_0 = 0$ et

$$S_n = \sum_{k=1}^{n} X_k, \quad n \in \mathbb{N}^*.$$

Enfin, on note $\mathscr{F}_0 = \{\emptyset, \Omega\}$ et $\mathscr{F}_n = \sigma(X_1, \ldots, X_n)$ la filtration naturelle de la suite $(X_n)_{n\in\mathbb{N}^*}$. Le but de cet exercice est de déterminer la loi de la variable aléatoire $W$ définie par $W = \sup_{n\in\mathbb{N}} S_n$.

En préambule, on rappelle la loi forte des grands nombres pour les suites de variables aléatoires i.i.d. intégrables : si $(Y_n)_{n\in\mathbb{N}^*}$ est une telle suite, alors $n^{-1}\sum_{k=1}^{n} Y_k$ converge p.s. vers $\mathbb{E}[Y_1]$ lorsque $n$ tend vers l'infini.

1. En utilisant la loi forte des grands nombres, montrez que $S_n$ tend p.s. vers $-\infty$ lorsque $n$ tend vers l'infini, et déduisez-en que $W < +\infty$ p.s.
2. On considère le processus $(M_n)_{n\in\mathbb{N}}$ défini par $M_0 = 1$ et pour tout $n \in \mathbb{N}^*$,
$$M_n = \left(\frac{q}{p}\right)^{S_n}.$$
Montrez que $\mathbb{E}[M_n] = 1$ pour tout $n \in \mathbb{N}^*$.
3. Montrez que $(M_n)_{n\in\mathbb{N}}$ est une martingale pour la filtration $(F_n)_{n\in\mathbb{N}}$.
4. Soit $k \in \mathbb{N}$ et $\tau_k$ le temps d'arrêt correspondant au premier temps de passage du processus $(S_n)_{n\in\mathbb{N}}$ par $k$, i.e.,
$$\tau_k = \inf\{n \geq 1 : S_n = k\}.$$
Montrez qu'au sens de la convergence p.s., on a
$$\lim_{n\to+\infty} M_{n\wedge\tau_k} = \left(\frac{q}{p}\right)^{k} 1_{\{\tau_k<+\infty\}}.$$
5. Déduisez-en la valeur de la probabilité $\mathbb{P}(\tau_k < +\infty)$.
6. Déduisez-en l'expression de la fonction de répartition de $W$ et concluez.

**Exercice 8.2** Soit $(Y_n)_{n\in\mathbb{N}^*}$ une suite de variables aléatoires i.i.d. suivant la loi normale centrée réduite. On considère également le processus $(X_n)_{n\in\mathbb{N}}$ défini par récurrence de la manière suivante : on pose $X_0 = x_0 \neq 0$ et pour tout $n \in \mathbb{N}$,

$$X_{n+1} = \theta X_n + Y_{n+1},$$

où $\theta$ est un paramètre réel inconnu. Le but de cet exercice est de construire un estimateur convergent du paramètre $\theta$ en supposant que le processus $(X_n)_{n\in\mathbb{N}}$ est observé. On note $\mathscr{F}_0 = \{\emptyset, \Omega\}$ et $\mathscr{F}_n = \sigma(Y_1, \ldots, Y_n)$ la filtration naturelle de la suite $(Y_n)_{n\in\mathbb{N}^*}$.

1. Montrez par récurrence sur $n \in \mathbb{N}^*$ que le processus $(X_n)_{n\in\mathbb{N}}$ est adapté et de carré intégrable.
2. On définit le processus $(M_n)_{n\in\mathbb{N}}$ par $M_0 = 0$ et pour tout $n \in \mathbb{N}^*$,
$$M_n = \sum_{k=1}^{n} X_{k-1} Y_k.$$
Montrez que $(M_n)_{n\in\mathbb{N}}$ est une martingale de carré intégrable.
3. Montrez que sa variation quadratique est donnée par $[M,M]_0 = 0$ et pour tout $n \in \mathbb{N}^*$ par
$$[M,M]_n = \sum_{k=1}^{n} X_{k-1}^2.$$
4. Étant donné $n \in \mathbb{N}^*$, on définit la log-vraisemblance associée au $n$-échantillon $(X_1, \ldots, X_n)$ par la formule
$$L_n(X_1, \ldots, X_n, \theta) = -\frac{1}{2} \sum_{k=1}^{n} (X_k - \theta X_{k-1})^2 - \frac{n}{2} \ln(2\pi).$$
Montrez qu'en tant que fonction (aléatoire) de $\theta$, la log-vraisemblance admet un unique maximum (aléatoire), noté $\theta_n$ et appelé estimateur du maximum de vraisemblance, dont l'expression explicite est donnée par l'égalité p.s. :
$$\theta_n = \theta + \frac{M_n}{[M,M]_n}, \quad n \in \mathbb{N}^*.$$
(justifiez la raison pour laquelle on peut diviser par $[M,M]_n$ dès que $n \in \mathbb{N}^*$).
5. On admet la propriété suivante : la suite $(Y_n)_{n\in\mathbb{N}}$ ne converge pas p.s. vers 0 lorsque $n$ tend vers l'infini. Déduisez-en que $\lim_{n\to+\infty}[M,M]_n = +\infty$ p.s.
6. Déduisez-en que $\theta_n$ converge p.s. vers $\theta$ lorsque $n$ tend vers l'infini.
7. Question bonus : montrez la propriété énoncée ci-dessus.

**Exercice 8.3** Soit $(Y_n)_{n\in\mathbb{N}}$ une suite i.i.d. de v.a. de Bernoulli de paramètre $p \in ]0,1[$. On suppose que la v.a. $Y_k$ correspond au k-ième lancer d'une même pièce de monnaie dont la probabilité de faire Pile vaut $p$. On note alors $(S_n)_{n\in\mathbb{N}}$ le processus à valeurs entières défini par $S_0 = 0$ et pour tout $n \in \mathbb{N}^*$ par
$$S_n = \sum_{k=1}^{n} Y_{k-1} Y_k.$$
On note également $(F_n)_{n\in\mathbb{N}}$ la filtration naturelle de la suite $(Y_n)_{n\in\mathbb{N}}$. L'objectif de cet exercice est de déterminer le comportement asymptotique du processus $(S_n)_{n\in\mathbb{N}}$.

1. À quoi correspond le processus $(S_n)_{n\in\mathbb{N}}$ du point de vue du jeu de Pile ou Face ?
2. On définit le processus $(M_n)_{n\in\mathbb{N}}$ par $M_0 = 0$ et pour tout $n \in \mathbb{N}^*$ par
$$M_n = \sum_{k=1}^{n} Y_{k-1} (Y_k - p).$$
Montrez que $(M_n)_{n\in\mathbb{N}}$ est une martingale pour la filtration $(\mathscr{F}_n)_{n\in\mathbb{N}}$, qui est de carré intégrable.

3. Quel résultat du cours vous permet d'obtenir directement la conclusion à la question précédente ?
4. Montrez que la variation quadratique de la martingale $(M_n)_{n\in\mathbb{N}}$ est donnée par $[M,M]_0 = 0$ et pour tout $n \in \mathbb{N}^*$ par
$$[M,M]_n = p(1-p)\sum_{k=1}^{n} Y_{k-1}^2.$$
5. Déduisez-en la limite p.s. suivante :
$$\lim_{n\to+\infty} \frac{M_n}{n} = 0.$$
6. Concluez sur la limite p.s. de $S_n/n$ lorsque $n$ tend vers l'infini. Quelle est votre interprétation de ce résultat en terme de jeu de Pile ou Face ?

**Exercice 8.4** On s'intéresse dans cet exercice à la version discrétisée en temps d'un objet fondamental dans le domaine du calcul stochastique : le mouvement brownien réel. Il s'agit tout simplement du processus $(B_n)_{n\in\mathbb{N}}$ à valeurs réelles défini par $B_0 = 0$ et pour tout $n \in \mathbb{N}^*$ par
$$B_n = \sum_{k=1}^{n} U_k,$$
où les $U_k$ sont i.i.d. de loi $\mathscr{N}(0,1)$. On note également $\mathscr{F}_0 = \{\emptyset, \Omega\}$ et $\mathscr{F}_n = \sigma(U_1,\ldots,U_n)$ pour tout $n \in \mathbb{N}^*$. Enfin on considère la v.a.
$$T = \inf\{n \geq 0 : B_n = a\},$$
correspondant au temps d'atteinte du niveau $a > 0$. On peut montrer que le mouvement brownien réel est une chaîne de Markov à valeurs réelles et l'on admet dans la suite que $T$ est un temps d'arrêt (pour la filtration $(\mathscr{F}_n)_{n\in\mathbb{N}}$) <u>fini</u> p.s., traduisant la propriété de récurrence de ce processus. Le but de cet exercice est de montrer par une méthode martingale que le mouvement brownien réel est récurrent nul, c'est-à-dire que $T$ n'est pas intégrable.
Dans la suite de l'exercice, $\theta$ désigne un paramètre positif ou nul quelconque.

1. Montrez par le calcul que l'on a
$$\mathbb{E}\left[\exp\left(\theta U - \frac{\theta^2}{2}\right)\right] = 1.$$
2. On définit le processus $(M_n)_{n\in\mathbb{N}}$ par
$$M_n = \exp\left(\theta B_n - \frac{\theta^2 n}{2}\right), \quad n \in \mathbb{N}.$$
Montrez que ce processus est une martingale pour la filtration $(\mathscr{F}_n)_{n\in\mathbb{N}}$.
3. En utilisant le théorème d'arrêt, montrez que
$$\mathbb{E}\left[\exp\left(-\frac{\theta^2 T}{2}\right)\right] = e^{-\theta a}.$$
4. Montrez que p.s.,
$$\sup_{\theta\in\mathbb{R}} \left|\frac{\partial}{\partial\theta}\exp\left(-\frac{\theta^2 T}{2}\right)\right| \leq e^{-1/2}\sqrt{T}.$$
5. Dans la suite, on suppose par l'absurde que $T$ est intégrable. Que peut-on dire de l'intégrabilité de $\sqrt{T}$ ?

6. Déduisez-en que la fonction
$$\theta \mapsto \mathbb{E}\left[\exp\left(-\frac{\theta^2 T}{2}\right)\right],$$
est alors de classe $\mathscr{C}^1$ sur $\mathbb{R}_+$.
7. En aboutissant à une contradiction, établissez la conclusion désirée de l'exercice.

**Exercice 8.5** Soit $(X_n)_{n\in\mathbb{N}^*}$ une suite de variables aléatoires i.i.d. à valeurs dans $\{-1,0,1\}$, de loi donnée par
$$\mathbb{P}(X_1 = 1) = \mathbb{P}(X_1 = -1) = \frac{p}{2} \quad \text{et} \quad \mathbb{P}(X_1 = 0) = 1-p,$$
où $p \in ]0,1[$. Considérons le processus $(M_n)_{n\in\mathbb{N}}$ défini par $M_0 = 0$ et pour tout $n \in \mathbb{N}^*$ par
$$M_n = \sum_{k=1}^n \frac{X_k}{\sqrt{k}}.$$
On note $(\mathscr{F}_n)_{n\in\mathbb{N}}$ la filtration naturelle de la suite $(X_n)_{n\in\mathbb{N}^*}$, avec $\mathscr{F}_0 = \{\emptyset, \Omega\}$.
1. Montrez que le processus $(M_n)_{n\in\mathbb{N}}$ est une martingale de carré intégrable.
2. Calculez sa variation quadratique.
3. Déduisez-en que $M_n/\ln n$ converge p.s. vers 0 lorsque $n \to +\infty$ (on rappelle que $\sum_{k=1}^n 1/k \underset{n\to+\infty}{\sim} \ln n$).
4. Montrez que la martingale vérifie un Théorème Central Limite (on précisera la vitesse ainsi que l'espérance et la variance limite).

**Exercice 8.6** Soit $(X_n)_{n\in\mathbb{N}^*}$ une suite de variables aléatoires i.i.d à valeurs dans $\{-1,0,1,2,3,\ldots\}$, supposées non constantes, intégrables et telles que $\mathbb{E}[X_1] \leq 0$. On définit le processus $(S_n)_{n\in\mathbb{N}}$ par $S_0 = 0$ et
$$S_n = \sum_{k=1}^n X_k, \quad n \in \mathbb{N}^*,$$
et l'on introduit $T$ son temps d'atteinte de $-1$, i.e.,
$$T = \inf\{n \in \mathbb{N}^* : S_n = -1\}.$$
On admet qu'il s'agit d'un temps d'arrêt pour la filtration $(\mathscr{F}_n)_{n\in\mathbb{N}}$ définie par $\mathscr{F}_n = \sigma(S_0,\ldots,S_n)$, $n \in \mathbb{N}$. Enfin, on note $g$ la fonction d'un paramètre réel $\beta$ définie par
$$g(\beta) = \ln \mathbb{E}\left[e^{-\beta X_1}\right].$$
Le but de cet exercice est de montrer que $\mathbb{P}(T < +\infty) = 1$, c'est-à-dire que p.s. le processus $(S_n)_{n\in\mathbb{N}}$ atteint la valeur $-1$.

**Partie I**

1. Montrez que la fonction $g$ est bien définie sur $[0,+\infty[$, et qu'elle est à valeurs dans $[0,+\infty[$ (pour ce dernier point, vous utiliserez l'inégalité de Jensen pour une fonction concave bien choisie).
2. Dans la suite de l'exercice, on suppose que $g$ est continue sur $[0,+\infty[$ et qu'elle est strictement positive sur $]0,+\infty[$. Étant donné $\beta > 0$, on considère le processus $(M_n)_{n\in\mathbb{N}}$ défini par
$$M_n = \exp\left(-\beta S_n - n g(\beta)\right), \quad n \in \mathbb{N}.$$
Montrez que sur l'événement $\{T = +\infty\}$, on a $\lim_{n\to+\infty} M_n = 0$ p.s.

3. Déduisez-en la convergence p.s. suivante :

$$\lim_{n\to+\infty} M_{n\wedge T} = M_T\, 1_{\{T<+\infty\}}.$$

4. Montrez que le processus $(M_{n\wedge T})_{n\in\mathbb{N}}$ est borné.
5. Déduisez-en que

$$\lim_{n\to+\infty} \mathbb{E}\left[M_{n\wedge T}\right] = e^{\beta} \times \mathbb{E}\left[e^{-Tg(\beta)}\, 1_{\{T<+\infty\}}\right].$$

**Partie II**

1. Montrez que $\mathbb{E}[M_n] = 1$ pour tout $n \in \mathbb{N}$.
2. Montrez que le processus $(M_n)_{n\in\mathbb{N}}$ est une martingale pour la filtration $(\mathscr{F}_n)_{n\in\mathbb{N}}$.
3. En utilisant le théorème d'arrêt et les résultats de la Partie I, montrez que

$$\mathbb{E}[e^{-Tg(\beta)}\, 1_{\{T<+\infty\}}] = e^{-\beta}.$$

4. Concluez en faisant tendre $\beta$ vers 0.

**Exercice 8.7** Cet exercice provient du problème de la ruine du joueur. Soit $(Y_n)_{n\in\mathbb{N}^*}$ une suite de variables aléatoires indépendantes de loi de Rademacher de paramètre $1/2$, i.e.,

$$\mathbb{P}(Y=1) = 1 - \mathbb{P}(Y=-1) = \frac{1}{2}.$$

Considérons le processus $(M_n)_{n\in\mathbb{N}}$ défini par $M_0 = 0$ et pour tout $n \in \mathbb{N}^*$, par

$$M_n = \sum_{k=1}^{n} Y_k.$$

Soit $T$ la variable aléatoire définie par

$$T = \inf\{n \geq 1 : M_n = -a \text{ ou } M_n = b\},$$

où $a$ et $b$ sont deux entiers strictement positifs. On note également $\mathscr{F}_0 = \{\emptyset, \Omega\}$ et $\mathscr{F}_n = \sigma(Y_1,\ldots,Y_n)$ pour tout $n \in \mathbb{N}^*$. Les résultats à votre disposition sont les suivants :

- la variable aléatoire $T$ est un temps d'arrêt p.s. fini.
- $\mathbb{P}(M_T = -a) = \frac{b}{a+b}$ et $\mathbb{P}(M_T = b) = \frac{a}{a+b}$.
- $\mathbb{E}[T] = ab$.

1. Montrez que le processus $(M_n^3 - 3nM_n)_{n\in\mathbb{N}}$ est intégrable.
2. Montrez que le processus $(M_n^3 - 3nM_n)_{n\in\mathbb{N}}$ est une martingale.
3. Déduisez-en l'égalité suivante : $\mathbb{E}[M_T^3] = 3\,\mathbb{E}[TM_T]$.
4. Déduisez-en la valeur de $\mathbb{E}[TM_T]$.
5. On suppose $a \neq b$. Déduisez-en que les variables aléatoires $T$ et $M_T$ ne sont pas indépendantes.

**Exercice 8.8** Soit $(Y_n)_{n\in\mathbb{N}}$ une suite de variables aléatoires indépendantes, non identiquement nulles, à valeurs dans $\mathbb{R}^+$ et telle que pour tout $n \in \mathbb{N}$,

$$\mathbb{E}[e^{\lambda Y_n}] < +\infty, \quad \forall \lambda \geq 0.$$

On considère également une suite $(U_n)_{n\in\mathbb{N}^*}$ de variables aléatoires indépendantes et de même loi uniforme sur $[0,1]$. On suppose que les deux suites $(Y_n)_{n\in\mathbb{N}}$ et $(U_n)_{n\in\mathbb{N}^*}$ sont indépendantes. On définit ensuite la suite $(Z_n)_{n\in\mathbb{N}^*}$ de la manière suivante : pour tout $n \in \mathbb{N}^*$,

$$Z_n = \begin{cases} e^{Y_{n-1}} - 1 & si \quad U_n \leq e^{-Y_{n-1}} \\ -1 & si \quad U_n > e^{-Y_{n-1}} \end{cases}$$

et l'on considère le processus $(M_n)_{n\in\mathbb{N}}$ défini par $M_0 = 0$ et pour tout $n \in \mathbb{N}^*$, par

$$M_n = \sum_{k=1}^{n} Z_k.$$

Enfin on note $\mathscr{F}_0 = \sigma(Y_0)$ et $\mathscr{F}_n = \sigma(Y_0, Y_1, \ldots, Y_n, U_1, U_2, \ldots, U_n)$ pour tout $n \in \mathbb{N}^*$.

1. Écrivez $Z_n$ sous la forme $Z_n = f(U_n, Y_{n-1})$ et déduisez-en qu'elle est $\mathscr{F}_n$-mesurable, et ce pour tout $n \in \mathbb{N}^*$.
2. Montrez que $Z_n \in L^2$ pour tout $n \in \mathbb{N}^*$.
3. Montrez que pour tout $n \in \mathbb{N}$,

$$\mathbb{P}(Z_{n+1} = e^{Y_n} - 1 \,|\, \mathscr{F}_n) = e^{-Y_n} \quad \text{et} \quad \mathbb{P}(Z_{n+1} = -1 \,|\, \mathscr{F}_n) = 1 - e^{-Y_n}.$$

4. Déduisez-en que $\mathbb{E}[Z_{n+1}|\mathscr{F}_n] = 0$ pour tout $n \in \mathbb{N}$.
5. Déduisez des questions précédentes que le processus $(M_n)_{n\in\mathbb{N}}$ est une martingale de carré intégrable.
6. Montrez que la variation quadratique de cette martingale est donnée par

$$[M,M]_n = \sum_{k=1}^{n} (e^{Y_{k-1}} - 1), \quad n \in \mathbb{N}^*.$$

7. À présent, les $Y_n$ sont des variables de Bernoulli de paramètre $p_n \in ]0,1[$.
   On suppose dans cette question que les $p_n$ ne dépendent pas de $n$. Montrez que $M_n/n$ tend p.s. vers 0 lorsque $n$ tend vers l'infini.
8. On suppose dans la suite que les $p_n$ vérifient $\sum_{n\in\mathbb{N}} p_n < +\infty$.
   (a) Calculez l'espérance de $[M,M]_n$ pour tout $n \in \mathbb{N}^*$.
   (b) Déduisez-en que le processus $(M_n)_{n\in\mathbb{N}}$ converge p.s. et dans $L^2$ (on ne demande pas de préciser la limite).

**Exercice 8.9 Optimisation d'une intégrale à paramètre**

Soit $Z$ une v.a.r. admettant une densité $f$ sur $\mathbb{R}$. On considère la fonction $V : \mathbb{R}^d \to \mathbb{R}$ définie par

$$V(x) = \int_{\mathbb{R}} \phi(x,z) f(z) dz,$$

où $\phi : \mathbb{R}^d \times \mathbb{R} \to \mathbb{R}$ est une fonction continue bornée, de classe $\mathscr{C}^2$ en la variable $x$ et telle que son gradient $\nabla_x$ et sa matrice hessienne $\nabla_x^2$ par rapport à $x$ soient bornés, uniformément en la variable $z$.

1. Montrez que $V$ est de classe $\mathscr{C}^2$ et que sa matrice hessienne est bornée.
2. On suppose de plus que $V$ tend vers l'infini à l'infini et n'a qu'un seul minimum $x^*$. Considérons $(Z_n)_{n\in\mathbb{N}^*}$ une suite de v.a.r. i.i.d. de même loi que $Z$ et $\alpha \in (1/2, 1]$. Montrez que l'algorithme défini par

$$X_{n+1} = X_n - \frac{1}{(n+1)^\alpha} \nabla_x \phi(X_n, Z_{n+1}), \quad n \in \mathbb{N}, \quad \text{avec} \quad X_0 = 0,$$

   converge p.s. vers $x^*$.

**Exercice 8.10** (Exam 2013)

La mise en oeuvre de méthodes de réduction de variance est souvent nécessaire pour rendre les algorithmes de Monte-Carlo efficaces. Parmi les méthodes de réduction de variance, il existe la méthode d'"Echantillonnage préférentiel" qui consiste à simuler des variables selon une loi adaptée à la fonction que l'on cherche à estimer. La question posée dans cet exercice est la suivante : comment bien choisir l'échantillonnage ? On se donne une loi $\mu$ sur $\mathbb{R}$ absolument continue par rapport à la mesure de Lebesgue et on note $p : \mathbb{R} \to \mathbb{R}$ sa densité. On considère $f : \mathbb{R} \to \mathbb{R}$ une fonction mesurable bornée et on s'intéresse au calcul de $\mathbb{E}[f(X)] = \int_{\mathbb{R}} f(x)p(x)dx$. On suppose pour simplifier que $\mathbb{P}(f(X) = 0) = 0$.

1. (a) Montrer que pour tout $\theta \in \mathbb{R}$,

$$\mathbb{E}[f(X)] = \mathbb{E}[f(X+\theta)\frac{p(X+\theta)}{p(X)}].$$

On notera

$$Y_\theta = f(X+\theta)\frac{p(X+\theta)}{p(X)}.$$

On a donc pour tout $\theta \in \mathbb{R}$, $\mathbb{E}[Y_\theta] = \mathbb{E}[f(X)]$. On peut donc envisager de choisir $Y_\theta$ ayant une variance minimale.

(b) Supposons qu'il existe $\theta^*$ tel que $\mathbb{E}[(Y_{\theta^*})^2] = \min_{\theta\in\mathbb{R}} \mathbb{E}[Y_\theta^2]$. A l'aide de la question précédente, montrer qu'alors

$$\mathrm{Var}(Y_{\theta^*}) = \min_{\theta\in\mathbb{R}} \mathrm{Var}(Y_\theta).$$

(c) On note $Q(\theta) = \mathbb{E}[(Y_\theta)^2]$. Montrer que

$$Q(\theta) = \mathbb{E}[f^2(X)\frac{p(X)}{p(X-\theta)}].$$

2. On suppose maintenant que $X \sim \mathscr{N}(0,1)$. Le but est de construire un algorithme permettant d'approcher $\theta^*$ (s'il existe).

(a) Montrer qu'alors

$$Q(\theta) = \mathbb{E}[f^2(X)e^{\frac{\theta^2}{2}-\theta X}].$$

(b) On admet que $Q$ est 2 fois dérivable (et satisfait les règles de dérivation des intégrales à paramètre). Calculer $Q'$ et $Q''$.

(c) En déduire que $Q'$ est strictement croissante. En admettant que $\lim_{\theta\to\pm\infty} Q(\theta) = +\infty$, en déduire que $Q$ admet un unique minimum en un point que l'on notera $\theta^*$.

3. Le but de la fin de l'exercice est de construire un algorithme permettant d'approcher $\theta^*$. La fonction $Q'$ ayant une croissance trop forte, on introduit une nouvelle fonction $h$ définie par

$$h(\theta) = \mathbb{E}[f^2(X-\theta)(2\theta - X)].$$

(a) En faisant un nouveau changement de variable sur la fonction $Q'$, montrer que $Q'(\theta) = h(\theta)e^{\theta^2}$. En déduire que $\theta^*$ est aussi l'unique point où la fonction $h$ s'annule (et que $h$ a le signe de $Q'$).

(b) Montrer que

$$h(\theta) \le C(1+|\theta|)$$

et que

$$\mathbb{E}[\big(f^2(X-\theta)(2\theta - X)\big)^2] \le C(1+\theta^2).$$

(c) Proposer un algorithme stochastique permettant d'approcher $\theta^*$ (où $\theta^*$ est vu comme l'unique zéro de la fonction $h$).

(d) Montrer à l'aide d'un théorème de Robbins-Monro (et de la question précédente) que cet algorithme converge vers $\theta^*$.

(e) L'idée naturelle pour approcher $\theta^*$ aurait été d'utiliser $Q'$ plutôt que $h$. Quel algorithme stochastique aurait-on alors proposé ? Pourquoi aurait-on eu des problèmes pour montrer sa convergence ?

**Exercice 8.11 Algorithme stochastique pour le quantile**

Soit $X$ une variable aléatoire intégrable quantifiant la perte de valeur d'un portefeuille entre un temps $t$ et un temps $t+\Delta_t$. On appelle alors "Value At Risk" et on note $V@R_\alpha(X)$ la quantité définie par :

$$V@R_\alpha(X) = \inf\{t \in \mathbb{R}, \mathbb{P}(X \le t) \ge \alpha\}.$$

Il s'agit donc du quantile d'ordre $\alpha$ associé à la variable $X$ ($\alpha$ est dans la pratique proche de 1 : 95 %, 99 %,...). On suppose dans la suite que $X$ est une variable aléatoire intégrable admettant une densité $f_X$ continue et strictement positive sur $\mathbb{R}$.
On définit alors la "Conditional Value At Risk" associée à $V@R_\alpha(X)$ par :

$$CV@R_\alpha(X) = \mathbb{E}[X|X \ge V@R_\alpha(X)].$$

$V@R_\alpha(X)$ et $CV@R_\alpha(X)$ sont des indicateurs pour évaluer le risque du marché. On les appelle mesures de risque. On désire dans cet exercice de construire un algorithme pour approcher simultanément $V@R$ et $CV@R$.

**Partie 1.**

1. Donner une interprétation (rapide) de la $V@R$ et de la $CV@R$.
2. On pose $f(\theta) = \mathbb{E}[(X-\theta)_+] = \mathbb{E}[\max(X-\theta, 0)]$. Montrer que pour tout $\theta \in \mathbb{R}$, $h \ge 0$,
$$f(\theta+h) - f(\theta) = -h\mathbb{P}(X \ge \theta+h) - \mathbb{E}[(X-\theta)1_{\{\theta \le X < \theta+h\}}].$$
En déduire que $f$ est dérivable sur $\mathbb{R}$ et que $f'(\theta) = -\mathbb{P}(X > \theta)$.
3. On pose
$$\psi(\theta) = \theta + \frac{1}{1-\alpha}\mathbb{E}[(X-\theta)_+].$$
Montrer que $\psi$ est de classe $\mathscr{C}^2$ sur $\mathbb{R}$ et donner $\psi'$ et $\psi''$.
4. Montrer que :
$$\frac{\psi(\theta)}{\theta} \to \begin{cases} -\frac{\alpha}{1-\alpha} & \text{lorsque } \theta \to -\infty \\ 1 & \text{lorsque } \theta \to +\infty. \end{cases}$$
5. Montrer que $\psi$ est strictement convexe et admet un unique minimum $\theta^* = V@R_\alpha(X)$.
6. Montrer que $\psi(\theta^*) = CV@R$.

**Partie 2.**

1. On considère l'algorithme stochastique $(\theta_n)_{n\ge 0}$ défini par
$$\theta_{n+1} = \theta_n - \frac{1}{n+1}\left(1 - \frac{1}{1-\alpha}\mathbf{1}_{\{X_{n+1} > \theta_n\}}\right)$$
où $(X_n)$ est une suite de variables aléatoires i.i.d. de même loi que $X$. Ecrire l'algorithme sous la forme $\theta_{n+1} = \theta_n - \gamma_{n+1}h(\theta_n) + \gamma_{n+1}\Delta M_{n+1}$ en précisant $\gamma_{n+1}$, $h$ et $\Delta M_{n+1}$ (On pourra vérifier que $h = \psi'$).
2. Montrer que $\theta_n \to \theta^*$ $p.s$ (On pourra utiliser la fonction $\psi$ comme fonction de Lyapounov).
3. On note $(X_k)_{k\ge 1}$ une suite de variables aléatoires i.i.d. de même loi que $X$ et on définit $C_n$ par
$$C_n = \frac{1}{n}\sum_{k=1}^{n} v(\theta_{k-1}, X_k) \quad \text{où } v(\theta, y) = \theta + \frac{1}{1-\alpha}(y-\theta)_+.$$
En utilisant la loi des grands nombres, montrer que $(C_n)$ converge $p.s.$ vers $CV@R_\alpha(X)$ (On pourra admettre que la fonction $x \mapsto \max(x, 0)$ est Lipschitzienne).

**Exercice 8.12** Une urne contient une boule blanche et une boule noire. On tire une boule au hasard et on la remet dans l'urne avec qu'une boule de la même couleur que celle tirée, et on répète l'expérience. Ainsi, avant le $n+1^e$ tirage, l'urne contient $n+2$ boules, dont $X_n$ noires.

1. Montrer que
$$\mathbb{P}(X_{n+1}=j|X_n=k)=\begin{cases}\dfrac{n+2-k}{n+2} & \text{si } j=k\\ \dfrac{k}{n+2} & \text{si } j=k+1\\ 0 & \text{sinon.}\end{cases}$$
2. En déduire par récurrence que $\forall n\geq 1, \forall j\in\{1,\dots,n+1\}, \mathbb{P}(X_n=j)=\dfrac{1}{n+1}$.
3. On pose $Y_n=\dfrac{X_n}{n+2}$. Calculez la fonction de répartition de $Y_n$, et en déduire que $Y_n$ converge en loi vers une limite à préciser.
4. Montrer que $\mathbb{E}(X_{n+1}|X_n=k)=\dfrac{n+3}{n+2}k$ et $\mathrm{Var}(X_{n+1}|X_n=k)=\dfrac{k}{n+2}\left(1-\dfrac{k}{n+2}\right)$.
   On rappelle que $\mathrm{Var}(Z|A)=\mathbb{E}\left((Z-E(Z|A))^2\middle|A\right)$.
5. On pose $\varepsilon_n=X_{n+1}-\frac{n+3}{n+2}X_n$. Montrer qu'on a la dynamique suivante pour $(Y_n)_{n\geq 1}$ :
$$Y_{n+1}=Y_n+\frac{1}{n+3}\varepsilon_{n+1}.$$
6. En déduire que $(Y_n)_{n\geq 1}$ est une martingale dans sa filtration naturelle.
7. À l'aide des questions précédentes, montrer que $(Y_n)_{n\geq 1}$ converge presque surement vers la limite identifiée en 3.

**Exercice 8.13** (Estimateurs de Monte-Carlo) Soit $f:\mathbb{R}\to\mathbb{R}$, mesurable bornée, et $X$ une variable aléatoire réelle. On étudie les estimateurs de Monte-Carlo :
$$\mathbb{E}[f(X)]=\lim_{n\to+\infty}\frac{1}{n}\sum_{i=1}^{n}f(X_i),$$
où les $(X_i)_{i\geq 1}$ sont des copies indépendantes de même loi que $X$.

1. On pose $\theta_0=0$ et $\forall n\geq 1$ :
$$\theta_n=\frac{1}{n}\sum_{i=1}^{n}f(X_i).$$
   Déterminer la fonction $H(\theta,x)$ de sorte que
$$\theta_{n+1}=\theta_n-\gamma_{n+1}H(\theta_n,X_{n+1}).$$
2. On propose la fonction de Lyapunov suivante
$$V(\theta)=\Big(\theta-\mathbb{E}[f(X)]\Big)^2+1.$$
   Prouver que la suite $(\theta_n)_{n\geq 1}$ converge vers $\theta^*=\mathbb{E}[f(X)]$.
3. Utiliser les résultats précédents pour fournir un estimateur de :
$$\int_0^1 f(x)dx,$$
   pour $f$ une fonction mesurable bornée.
4. Montrer que $\forall n\geq 1$, $H(\theta_n,X_{n+1})\leq C$ presque sûrement.

5. Montrer que $\mathrm{Var}[H(\theta^*, X)] = \mathrm{Var}[f(X)]$.
6. On précise que $f$ n'est pas constante. Déduire à l'aide du Théorème Central Limite que

$$\frac{\sqrt{n}}{\sqrt{\mathrm{Var}[f(X)]}}(\theta_n - \theta^*) \Rightarrow \mathcal{N}(0,1).$$

7. Soit $q_\alpha$ tel que $\mathbb{P}(|\mathcal{N}(0,1)| \leq q_\alpha) = 1 - \alpha$. Montrer que

$$\lim_{n\to+\infty} \mathbb{P}\left(|\theta_n - \theta^*| \leq q_\alpha \frac{K}{\sqrt{n}}\right) \geq 1 - \alpha.$$

   pour une constante $K$ qu'on déterminera.
   *Indication :* On peut chercher à montrer que quand $f$ est bornée $\mathrm{Var}[f(X)] \leq M$.
8. Déduire un intervalle de confiance asymptotique pour $\mathbb{E}[f(X)]$.

**Exercice 8.14** Soit $X$ une variable aléatoire à densité $f_X$, strictement positive et bornée. On rappelle que son quantile au niveau $\alpha$ est l'unique réel $q^*$ solution de l'équation

$$\mathbb{P}(X \leq q^*) = \int_{-\infty}^{q^*} f_X(x)dx = \alpha.$$

Pour résoudre numériquement cette équation, on propose le schéma suivant :

$$\hat{q}_{n+1} = \hat{q}_n - \frac{1}{n+1}\left(\mathbf{1}_{\{X_{n+1} \leq \hat{q}_n\}} - \alpha\right),\ \hat{q}_0 \in \mathbb{R}.$$

On considère une fonction $V$ telle que $V'(q) = h(q)$, où $h(q) = \mathbb{E}\left(\mathbf{1}_{\{X \leq q\}} - \alpha\right)$.

1. Justifiez que $(q - q^*)(V'(q) - V'(q^*)) \geq 0$.
2. En déduire le tableau de variation de $V$ suivant :

| $q$ | $-\infty$ | | $q^*$ | | $+\infty$ |
|---|---|---|---|---|---|
| $V(q)$ | | $\searrow$ | $V(q^*)$ | $\nearrow$ | |

3. Justifiez que sans perte de généralité, on peut prendre $V(q^*) > 0$.
4. Montrez que $V(q) \to +\infty$ quand $|q| \to +\infty$.
5. Justifiez de la convergence de l'algorithme.

**Exercice 8.15** On note $Y_n$ le capital d'une compagnie d'assurance à une certaine année $n$ . Chaque année, ses assurés lui versent une cotisation totale de $P$ euros (premium en anglais), supposé constant, mais doit débourser $C_n$ euros en remboursements. On suppose que les $C_i$ sont toutes indépendantes de loi $\mathcal{N}(\mu, \sigma^2)$, avec $P > \mu$. On a ainsi la relation suivante entre deux années successives :

$$Y_{n+1} = Y_n + P - C_{n+1}, \quad Y_0 > 0.$$

On va montrer que la probabilité de banqueroute satisfait :

$$\mathbb{P}(Y_n \leq 0 \text{ pour un certain } n) \leq \exp\left(-\frac{2(P-\mu)Y_0}{\sigma^2}\right).$$

1. Interprétez la condition $P > \mu$. Que devient la borne sur la probabilité dans le cas contraire ?
2. Montrer que la probabilité de banqueroute peut se ré-écrire :

$$\mathbb{P}(Y_n \leq 0 \text{ pour un certain } n) = \mathbb{P}\left(\sup_{n\geq 0} \sum_{k=1}^{n} (C_k - P) \geq Y_0\right).$$

3. On pose $X_n = \sum_{k=1}^{n}(C_k - P)$, avec $X_0 = 0$. Montrer qu'on peut choisir $\lambda_0 \neq 0$ pour que $(e^{\lambda_0 X_n})_{n\in\mathbb{N}}$ soit une martingale dans la filtration naturelle des $(C_n)_{n\in\mathbb{N}}$.
   *Indication :* On rappelle que si $Z \sim \mathcal{N}(\mu, \sigma^2)$, alors $\mathbb{E}(e^{tZ}) = e^{\mu t + \frac{\sigma^2}{2}t^2}$.

4. Dans la suite, on prendra $\lambda_0 = 2\frac{(P-\mu)}{\sigma^2} > 0$. En déduire l'inégalité demandée.

**Exercice 8.16** Soit $(X_n)_{n\in\mathbb{N}}$ une suite de variable aléatoires i.i.d., strictement positives et telles que pour tous $n \in \mathbb{N}$, $\mathbb{E}(X_n) = 1$. On note $\mathscr{F}_n$ la filtration naturelle de $(X_n)_{n\in\mathbb{N}}$. On pose $Y_n = \prod_{k=1}^n X_k$.

1. Montrer que $(Y_n)_{n\in\mathbb{N}}$ est une martingale qui converge presque surement. On note $Y_\infty$ sa limite.
2. On pose $a_n = \mathbb{E}(\sqrt{X_n})$. Montrer que $a_n \in ]0,1]$ et que $A_n = \prod_{k=1}^n a_i$ converge. On note $A_\infty$ sa limite.
3. On pose $Z_n = \prod_{k=1}^n \frac{\sqrt{X_k}}{a_k}$, montrer que $(Z_n)_{n\in\mathbb{N}}$ est une martingale qui converge presque surement. On note $Z_\infty$ sa limite.
4. Justifier que $Z_n^2 = \frac{Y_n}{A_n^2}$, $Y_n = Z_n^2 A_n^2$ et que $Y_n \leq Z_n^2$.
5. On suppose que $A_\infty = 0$, montrer qu'alors $Y_\infty = 0$.
6. On suppose maintenant que $A_\infty > 0$. On va montrer que la convergence a lieu dans $L^1$ et que $\mathbb{E}[Y_\infty] = 1$.
   (a) Soit $W_n = \max_{k\leq n} Z_k$ et $W_\infty = \max_{k\geq 1} Z_k$. Montrer que $\mathbb{E}[W_n^2] \leq 4\mathbb{E}[Z_n^2] \leq \frac{4}{A_n^2}$.
   (b) En déduire que $W_\infty^2$ est intégrable (*Indication :* On justifiera que $W_n^2$ est monotone).
   (c) En déduire que $(Y_n)_{n\in\mathbb{N}}$ est uniformément intégrable (*Indication :* On montrera que $Y_n \leq W_\infty^2$).
   (d) Conclure.

**Exercice 8.17** Soit $X$ une variable aléatoire réelle de densité $\sqrt{\frac{2}{\pi}} e^{-x^2/2}\mathbf{1}_{x\geq 0}$, par rapport à la mesure de Lebesgue. On note $\mathscr{L}$ sa transformée de Laplace :

$$\mathscr{L}(\theta) = \mathbb{E}[e^{-\theta X}].$$

Pour un $\alpha \in ]0,1[$ donné, on cherche à trouver $\theta^*$ tel que $\mathscr{L}(\theta^*) = \alpha$. Pour cela, on considère :

$$\theta_0 = 1,\ \theta_{n+1} = \theta_n - \gamma_{n+1} \times \begin{cases} (\alpha - 1) & \text{si } \theta_n < 0 \\ (\alpha - e^{-\theta_n X_{n+1}}) & \text{si } \theta_n \geq 0 \end{cases}$$

avec $(\gamma_k)_{k\in\mathbb{N}}$ telle que $\sum \gamma_k = +\infty$ et $\sum \gamma_k^2 < +\infty$.

1. Montrer que $\forall \theta \in \mathbb{R}_+$, $\mathscr{L}(\theta) \in [0,1]$, et est strictement décroissante.
2. En déduire que l'équation $\mathscr{L}(\theta^*) = \alpha$, a une unique solution pour $\alpha \in ]0,1[$.
3. On pose $H(\theta,x) = \alpha - \min(e^{-\theta x}, 1)$. Justifiez que l'algorithme peut s'écrire pour $n \geq 0$ :
$$\theta_0 = 1,\ \theta_{n+1} = \theta_n - \gamma_{n+1} H(\theta_n, X_{n+1}).$$
4. On définit $h(\theta) = \mathbb{E}[H(\theta,X)]$. Montrer que :
$$h(\theta) = \begin{cases} \alpha - 1 & \text{si } \theta < 0 \\ \alpha - \mathscr{L}(\theta) & \text{si } \theta \geq 0. \end{cases}$$
5. Résoudre $h(\theta) = 0$. Sans justifications, si l'algorithme converge, vers quelle valeur converge-t-il ?
6. Montrer que $h$ est bornée et croissante sur $\mathbb{R}$, et strictement croissante sur $\mathbb{R}_+$.
7. En utilisant le fonction de Lyapunov $V(\theta) = (\theta - \theta^*)^2 + 1$, justifier la convergence de l'algorithme.

**Exercice 8.18** On considère $(Y_n)_{n\geq 0}$ une suite de variables aléatoires i.i.d. telles que $\mathbb{P}(Y_0 = 1) = p$ et $\mathbb{P}(Y_0 = -1) = q = 1 - p$. On suppose que $0 < p < 1$. On pose enfin $S_0 = s_0 \in \mathbb{R}^*$ et pour $n \geq 1$,

$$S_n = S_0 + \sum_{j=1}^n Y_{j-1} Y_j.$$

On travaille dans la filtration $\mathscr{F}_n = \sigma(Y_0, \dots, Y_n)$.

1. Montrer que
$$\mathbb{P}(S_n > S_{n-1} | \mathscr{F}_{n-1}) = p\mathbf{1}_{Y_{n-1}>0} + q\mathbf{1}_{Y_{n-1}<0}.$$
*Indication :* On rappelle que $\mathbb{P}(S_n > S_{n-1} | \mathscr{F}_{n-1}) = \mathbb{E}[\mathbf{1}_{\{S_n > S_{n-1}\}} | \mathscr{F}_{n-1}]$.
2. En déduire que $\mathbb{P}(S_n > S_{n-1}) = p^2 + q^2 = 2p^2 - 2p + 1$.
3. Montrer que cette probabilité est strictement supérieure à $1/2$ lorsque $p \neq 1/2$.
On note $(X_n)_{n\geq 0}$ le processus défini par
$$X_0 = 0 \quad et \quad X_n := \sum_{k=1}^{n} \mathbb{E}[S_k - S_{k-1} | \mathscr{F}_{k-1}] \quad \forall n \geq 1,$$
et $(M_n)_{n\geq 0}$ le processus défini par $M_n = S_n - X_n$ pour tout $n \geq 0$.
4. Montrer que $(M_n)_{n\geq 0}$ est une $(\mathscr{F}_n)$-martingale de carré intégrable.
5. On rappelle que
$$[M,M]_n = \sum_{k=1}^{n} \mathbb{E}\left[(M_k - M_{k-1})^2 \middle| \mathscr{F}_{k-1}\right].$$
Montrer que $[M,M]_n = \sum_{k=1}^{n} \mathrm{Var}(Y_k)$.
6. En déduire que presque surement,
$$\frac{M_n}{n} \underset{n\to+\infty}{\longrightarrow} 0.$$
7. Calculer $\mathbb{E}[S_n - S_{n-1} | \mathscr{F}_{n-1}]$ pour tout $n \geq 1$.
8. En déduire que
$$X_n = (p-q) \sum_{k=1}^{n} Y_{k-1}.$$
9. En utilisant la loi des grand nombres standard, montrer que
$$\frac{S_n}{n} \underset{n\to+\infty}{\longrightarrow} (p-q)^2 \ \ (p.s.).$$

**Exercice 8.19** On a une population de taille $N$ fixée ($N \in \mathbb{N}^*$) qui se renouvelle entièrement à chaque génération et dont chaque individu est de type $A$ ou $a$. On note $X_n$ le nombre d'individus de type $A$ à la génération $n$ et on suppose que $X_0 = x_0 \in \{0, \dots, N\}$. On suppose également qu'à la génération $n+1$, chaque individu (de la $(n+1)$-ième génération) choisit (indépendemment des autres) un parent selon une loi uniforme sur $\{1, \dots, N\}$ et hérite du type de son parent (Un enfant a un unique parent mais un parent peut avoir plusieurs enfants). Le nombre d'individus de type $A$ à la génération $n+1$ peut alors se modéliser de la manière suivante :
$$X_{n+1} = \sum_{i=1}^{N} \mathbf{1}_{\{U_{n+1}^i \leq \frac{X_n}{N}\}}$$
où les $U_n^i$, $(n \in \mathbb{N}^*, i \in \{1, \dots, N\})$ sont des variables aléatoires $i.i.d.$ de loi uniforme sur $[0,1]$. On note enfin $\mathscr{F}_0 = \{\emptyset, \Omega\}$ et $\mathscr{F}_n = \sigma(U_k^i, k \in \{1, \dots, n\}, i \in \{1, \dots, N\})$ pour tout $n \in \mathbb{N}^*$.

1. Pour tout $k \in \{0, \dots, N\}$, montrer que
$$\mathbb{P}(X_{n+1} = k | \mathscr{F}_n) = \binom{N}{k} \left(\frac{X_n}{N}\right)^k \left(1 - \frac{X_n}{N}\right)^{N-k}.$$
2. Montrer que $(X_n)_{n\geq 0}$ est une $(\mathscr{F}_n)$-martingale.

3. Montrer que $(X_n)$ converge $p.s.$ vers $X_\infty$ intégrable. Calculer $\mathbb{E}[X_\infty]$.
4. Calculer $\mathbb{E}[X_{n+1}^2|\mathscr{F}_n]$. (On pensera à utiliser la variance d'une binômiale)
5. Pour tout $n \in \mathbb{N}$, on note
$$Z_n = \left(\frac{N}{N-1}\right)^n X_n(N-X_n).$$
Montrer que $(Z_n)_{n\geq 0}$ est une $(\mathscr{F}_n)$-martingale.
6. En déduire $\mathbb{E}[X_\infty(N-X_\infty)]$.
7. En déduire la loi de $X_\infty$.

**Exercice 8.20** On considère $(X_n)_{n\geq 1}$, une suite de variables aléatoires indépendantes de loi $\mathscr{N}(0,1)$. On note $(S_n)$, la suite définie par $S_0 = 0$ et $S_n = X_1 + \ldots + X_n$ pour $n \geq 1$. On note également $\mathscr{F}_n = \sigma(S_0, S_1, \ldots, S_n)$ et $\theta$ un nombre strictement positif.

1. Soit $n \geq 0$. Calculer
$$\mathbb{E}[e^{\theta S_{n+1}}|\mathscr{F}_n].$$
2. En déduire que $(W_n)_{n\geq 0}$ définie par :
$$W_n = e^{\theta S_n - n\frac{\theta^2}{2}}$$
est une $(\mathscr{F}_n)$-martingale.
3. Montrer que la martingale $(W_n)$ converge p.s.
4. Montrer que
$$\theta S_n - n\frac{\theta^2}{2} \underset{n\to+\infty}{\longrightarrow} -\infty.$$
5. En déduire la limite de $(W_n)$ lorsque $n \to +\infty$.
6. On pose $u_n = \mathbb{E}[W_n]$. Quelle est la limite de $(u_n)$ ?

**Le problème du Mariage de la princesse**

Il s'agit d'un problème d'arrêt optimal qui peut s'utiliser dans bien des applications. Un roi voudrait marier sa fille au meilleur prince. Pour cela, il fait venir les prétendant un par un, jusqu'à ce qu'il en trouve un qui lui convienne. Le problème est que dès qu'un candidat se présente, le roi doit décider tout de suite, il ne pourra pas rappeler le prince s'il a été refusé une fois.

Quelle stratégie employer ?

- On modélise les princes par une suite de variables aléatoires indépendantes de loi uniforme : $X_1, X_2, \dots X_n$. Alors, $100 \times X_i$ représente le score du $i^e$ prince : 0%, le prince ne convient pas du tout, 100%, le prince est parfait. On prendra pour convention $X_0 = 0$.
- La filtration considérée sera celle des princes : $\mathscr{F}_k = \sigma(X_1, \dots, X_k)$.

**Exercice 8.21 — calculs préliminaires.** On définit $(a_k)_{k\in\mathbb{N}}$ par $a_0 = 0$ et $a_{k+1} = \frac{1}{2}(1 + a_k^2)$.

1. Montrer que $\forall k \in \mathbb{N}$, $a_k < 1$.
2. Montrer que $a_k < a_{k+1}$.
3. Montrer que $a_k \underset{k\to+\infty}{\longrightarrow} 1$.
4. Montrer que si $X$ suit une loi uniforme sur [0,1], alors $\mathbb{E}[X \vee a_k] = a_{k+1}$.

**Exercice 8.22 — Pour un temps d'arrêt quelconque.** Soit $\rho$ un temps d'arrêt borné. On pose $Y_k = X_{\rho\wedge k} \vee a_{n-k}$. On va montrer que $(Y_k)_{k\geq 1}$ est une surmartingale, et lui appliquer le théorème d'arrêt optimal.

1. Justifiez que $\mathbf{1}_{\rho\leq k}$ et $X_\rho \mathbf{1}_{\rho\leq k}$ sont $\mathscr{F}_k$ mesurables.
2. En déduire que
$$\mathbb{E}\left(X_\rho \vee a_{n-(k+1)} \mathbf{1}_{\rho\leq k} \middle| \mathscr{F}_k\right) = X_\rho \vee a_{n-(k+1)} \mathbf{1}_{\rho\leq k}.$$
3. Justifiez que $X_\rho \vee a_{n-(k+1)} \leq X_\rho \vee a_{n-k}$.
4. En utilisant la question 4 de l'exercice 8.21, prouver que
$$\mathbb{E}\left(X_{k+1} \vee a_{n-(k+1)} \mathbf{1}_{\rho>k} \middle| \mathscr{F}_k\right) = a_{n-k}\mathbf{1}_{\rho>k} \leq X_{\rho\wedge k} \vee a_{n-k}\mathbf{1}_{\rho>k}$$
5. En déduire que $(Y_k)_{k\geq 0}$ est une surmartingale.
6. En utilisant la décomposition de Doob et le théorème d'arrêt, montrer que $\mathbb{E}(Y_\rho) \leq \mathbb{E}(Y_0) = a_n$.
7. Si le roi décide de s'arrêter au temps aléatoire $\rho$, donner une borne sur le score moyen du prince.

**Exercice 8.23 — Stratégie optimale.** Dans cette partie, on considère $Y_k = X_{\tau_n\wedge k} \vee a_{n-k}$, où on pose
$$\tau_n = \inf\{1 \leq k \leq n \,;\, X_k > a_{n-k}\}.$$

1. Justifier que $\tau_n$ est un temps d'arrêt borné.
2. En reprenant les arguments de la question 5 de l'exercice 8.22 et la définition de $\tau_n$, prouver que $(Y_k)_{k\leq n}$ est bien une martingale.
3. En déduire qu'alors, $\mathbb{E}(X_{\tau_n}) = \mathbb{E}(Y_\rho) = \mathbb{E}(Y_0) = a_n$
4. Expliquer pourquoi la stratégie $\tau_n$ est considérée comme optimale.
5. Il y a 5 prétendants dans le royaume. Voici les premières valeurs de la suite $(a_n)_{n\geq 0}$ :

| $n$ | 0 | 1 | 2 | 3 | 4 | 5 |
|---|---|---|---|---|---|---|
| $a_n$ | 0 | 0.5 | 0.625 | 0.695 | 0.742 | 0.775 |

Quelle sera la stratégie du roi et quel sera le score moyen du prince avec cette stratégie ?

**Solution :** Prenons sur un exemple $n = 5$, la princesse a 5 prétendants. On calcule les valeurs de la suite $(a_k)_{k\geq 1}$ et on donne la consigne suivante au roi :

— Si $X_1 > a_4 \approx 0.742$, il faut choisir le prince numéro 1 sinon
— Si $X_2 > a_3 \approx 0.695$ il faut choisir le prince numéro 2, sinon
— Si $X_3 > a_2 = 0.625$, il faut choisir le prince numéro 3, sinon
— Si $X_4 > a_1 = 0.5$, il faut choisir le prince numéro 4, sinon
— il faut choisir le prince numéro 5.

Avec cette stratégie, le score moyen du prince sera de $\mathbb{E}(X_{\tau_5}) = a_5 \approx 0.775$.

Auteur : Huang Lorick

LHUANG@INSA-TOULOUSE.FR

POLYCOPIÉ DE INSA DE TOULOUSE

*Version du 17 septembre 2025*