Chapitre 3

Martingales discrètes

Résumé. Étant donné un espace de probabilité filtré par $\mathcal{F} = (\mathcal{F}_n)_{n\in\mathbb{N}}$, une *martingale* est un processus $(X_n)_{n\in\mathbb{N}}$ adapté à la filtration et tel que $\mathbb{E}[X_{n+1}|\mathcal{F}_n] = X_n$ pour tout n. Cette classe de processus est stable par *intégration stochastique contre des processus prévisibles*, et par coupure par des *temps d'arrêt*. Toute martingale bornée dans $\mathrm{L}^1(\Omega, \mathcal{B}, \mathbb{P})$ converge p.s., et dans L^1 si et seulement si elle est uniformément intégrable. Toute martingale bornée dans $\mathrm{L}^p(\Omega, \mathcal{B}, \mathbb{P})$ avec $p > 1$ converge p.s. et dans L^p. Cette théorie s'applique en particulier aux marches aléatoires et aux *processus de branchement*.

Dans ce chapitre, $(\Omega, \mathcal{B}, \mathbb{P})$ est un espace de probabilité, et $\mathcal{F} = (\mathcal{F}_n)_{n\in\mathbb{N}}$ est une filtration de l'espace, c'est-à-dire une suite croissante de sous-tribus de $\mathcal{B}$. Dans le cadre de processus à temps discret définis sur $(\Omega, \mathcal{B}, \mathbb{P})$, il faut comprendre $\mathcal{F}_n$ comme l'ensemble des événements mesurables au temps n. Ainsi, on dira qu'une suite de variables aléatoires $(X_n)_{n\in\mathbb{N}}$ définies sur $(\Omega, \mathcal{B}, \mathbb{P})$ forme un *processus adapté* relativement à la filtration $(\mathcal{F}_n)_{n\in\mathbb{N}}$ si X_n est $\mathcal{F}_n$-mesurable pour tout temps n. Tout processus est adapté à la filtration qu'il engendre, définie par $\mathcal{F}_n = \sigma(X_0, \ldots, X_n)$. On aura également besoin de la notion de *processus prévisible* : ainsi, $(X_n)_{n\geq 1}$ est dit prévisible si X_n est $\mathcal{F}_{n-1}$-mesurable pour tout temps $n \geq 1$.

1. Martingales et leurs transformations

1.1. Exemples élémentaires de martingales. Dans ce qui suit, un processus aléatoire $(X_n)_{n\in\mathbb{N}}$ sera dit intégrable si toutes les variables X_n le sont.

Définition 3.1 (Martingales). *Un processus aléatoire $(X_n)_{n\in\mathbb{N}}$ (à valeurs dans $\mathbb{R}$) est appelé martingale s'il est adapté, intégrable et si, pour tout n,*

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

On parle de sur-martingale si l'égalité est remplacée par l'inégalité $\mathbb{E}[X_{n+1}|\mathcal{F}_n] \leq X_n$, et de sous-martingale si l'égalité est remplacée par l'inégalité $\mathbb{E}[X_{n+1}|\mathcal{F}_n] \geq X_n$.

Étant donnée une martingale, on a plus généralement et par conditionnements successifs $\mathbb{E}[X_N|\mathcal{F}_n] = X_n$ pour $N \geq n$, et $\mathbb{E}[X_n] = \mathbb{E}[X_0]$ pour tout $n \in \mathbb{N}$.

Exemple. Soit $(Y_n)_{n\geq 1}$ une suite de variables aléatoires intégrables et indépendantes (mais pas forcément de même loi), $\mathcal{F} = (\mathcal{F}_n)_{n\in\mathbb{N}}$ la filtration engendrée par ces variables, et $X_n = \sum_{k=1}^n Y_k$. C'est un processus adapté, et si chaque Y_n vérifie $\mathbb{E}[Y_n] = 0$, alors c'est une martingale, car

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

En particulier, les marches aléatoires simples sur $\mathbb{Z}$ sont des martingales (et plus généralement, les marches aléatoires dont les incréments i.i.d. ont moyenne nulle). Si l'on remplace l'hypothèse $\mathbb{E}[Y_n] = 0$ par $\mathbb{E}[Y_n] \geq 0$ (respectivement, $\mathbb{E}[Y_n] \leq 0$), on obtient une sous-martingale (resp., une sur-martingale).

Exemple. Soit $(Y_n)_{n\geq 1}$ une suite de variables aléatoires positives, bornées et indépendantes, $\mathcal{F} = (\mathcal{F}_n)_{n\in\mathbb{N}}$ la filtration engendrée par ces variables, et $X_n = \prod_{k=1}^n Y_k$. C'est un processus adapté, et si chaque Y_n vérifie $\mathbb{E}[Y_n] = 1$, alors c'est une martingale, car

$$\mathbb{E}[X_{n+1}|\mathcal{F}_n] = \mathbb{E}[X_n Y_{n+1}|\mathcal{F}_n] = X_n\, \mathbb{E}[Y_{n+1}] = X_n.$$

Exemple. Soit $(X_n)_{n\in\mathbb{N}}$ une chaîne de Markov de matrice de transition P sur un espace d'états $\mathfrak{X}$. Une fonction sur $\mathfrak{X}$ est dite *harmonique* pour P si pour tout $x \in \mathfrak{X}$,

$$f(x) = (Pf)(x) = \sum_{y\in\mathfrak{X}} P(x,y)f(y).$$

Dans ce cas, le processus $(f(X_n))_{n\in\mathbb{N}}$ est une martingale pour la filtration engendrée par la chaîne de Markov, car $\mathbb{E}[f(X_{n+1})|\mathcal{F}_n] = Pf(X_n) = f(X_n)$. On obtient de même des sous- et des sur-martingales à partir de fonctions sous- ou sur-harmoniques.

Exemple. Soit X une variable aléatoire intégrable, et $X_n = \mathbb{E}[X|\mathcal{F}_n]$, avec $\mathcal{F}$ filtration fixée sur l'espace de probabilité $(\Omega, \mathcal{B}, \mathbb{P})$. Le processus $(X_n)_{n\in\mathbb{N}}$ est une martingale par la propriété des conditionnements successifs du chapitre 1.

D'un point de vue abstrait, la notion de martingale permet de mesurer la différence entre processus adaptés et processus prévisibles :

Proposition 3.2 (Décomposition de Doob). *Tout processus adapté intégrable $(X_n)_{n\in\mathbb{N}}$ s'écrit de manière unique comme somme d'une martingale $(M_n)_{n\in\mathbb{N}}$ et d'un processus prévisible $(P_n)_{n\in\mathbb{N}}$ issu de 0.*

L'existence est donnée par les formules

$$M_{n+1} - M_n = X_{n+1} - \mathbb{E}[X_{n+1}|\mathcal{F}_n]$$
$$P_{n+1} - P_n = \mathbb{E}[X_{n+1}|\mathcal{F}_n] - X_n$$

qui fournissent une définition récursive des deux composantes de X_n, avec $M_0 = X_0$ et $P_0 = 0$. Pour l'unicité, il suffit de noter qu'un processus qui est à la fois une martingale et prévisible est constant au cours du temps.

1.2. Transformations de martingales. À partir d'une martingale $(X_n)_{n\in\mathbb{N}}$, il existe plusieurs constructions qui permettent d'obtenir de nouvelles martingales ; dans le cadre des martingales à temps continu, ceci mène à la théorie du *calcul stochastique d'Itô*. Dans le cadre des martingales à temps discret, la théorie se réduit essentiellement aux résultats suivants. Notons que si $(X_n)_{n\in\mathbb{N}}$ est une sous-martingale, alors $(-X_n)_{n\in\mathbb{N}}$ est une sur-martingale ; les résultats seront donc énoncés seulement pour les sur-martingales.

Théorème 3.3 (Intégrale stochastique discrète). *Soit $(X_n)_{n\in\mathbb{N}}$ un processus adapté, et $(A_n)_{n\in\mathbb{N}}$ un processus prévisible. Notons*

$$(A\cdot X)_n = \sum_{k=1}^{n} A_k(X_k - X_{k-1}),$$

qui est un processus adapté issu de 0. Si $(X_n)_{n\in\mathbb{N}}$ est une martingale et $(A_n)_{n\in\mathbb{N}}$ est une suite de variables aléatoires bornées, alors $((A\cdot X)_n)_{n\in\mathbb{N}}$ est une martingale. De même, si $(X_n)_{n\in\mathbb{N}}$ est une sur-martingale et $(A_n)_{n\in\mathbb{N}}$ est une suite de variables aléatoires bornées positives, alors $((A\cdot X)_n)_{n\in\mathbb{N}}$ est une sur-martingale.

En effet, dans le cas des martingales,

$$\begin{aligned}\mathbb{E}[(A\cdot X)_{n+1}|\mathcal{F}_n] &= (A\cdot X)_n + \mathbb{E}[A_{n+1}(X_{n+1}-X_n)|\mathcal{F}_n]\\ &= (A\cdot X)_n + A_{n+1}(\mathbb{E}[X_{n+1}|\mathcal{F}_n] - X_n) = (A\cdot X)_n.\end{aligned}$$

Dans le cas des sur-martingales, la même preuve s'adapte, et l'on peut alternativement utiliser le résultat élémentaire suivant : un processus adapté intégrable est une sur-martingale si et seulement si sa partie prévisible au sens de la décomposition de Doob est un processus décroissant (appelé le *compensateur* de la sur-martingale).

Exemple. Soit $(X_n)_{n\in\mathbb{N}}$ une marche aléatoire sur $\mathbb{Z}$, d'incréments indépendants et identiquement distribués $Y_1,\ldots,Y_n,\ldots$, avec $\mathbb{E}[Y_1]=0$. On interprète chaque variable $Y_i = X_i - X_{i-1}$ comme le résultat d'un jeu aléatoire, et l'on décide à chaque étape $i-1$ d'une mise $A_i \geq 0$ pour le i-ième tirage du jeu. Cette mise est donc $\mathcal{F}_{i-1}$-mesurable, c'est-à-dire prévisible, et on peut imaginer que l'on décide de la mise A_i en fonction des résultats précédents du jeu $Y_1,\ldots,Y_{i-1}$. L'hypothèse $\mathbb{E}[Y_1]=0$ assure que pour tout temps n, $\mathbb{E}[X_n]=0$, c'est-à-dire qu'au temps n on a en moyenne gagné autant de fois que l'on a perdu. La proposition 3.3 assure que dans ce cas, quelque soit la stratégie des mises, le gain au temps n

$$G_n = (A\cdot X)_n = \sum_{i=1}^{n} A_i\,Y_i$$

est également une martingale, de sorte qu'on a aussi $\mathbb{E}[G_n]=0$ pour tout n. Autrement dit, si un jeu est équilibré, alors aucune stratégie prévisible ne permet d'avoir des gains en moyenne strictement positifs.

Soit $(X_n)_{n\in\mathbb{N}}$ une martingale. Comme toute fonction $(f(X_n))_{n\in\mathbb{N}}$ d'un processus adapté est de nouveau un processus adapté, sous des hypothèses d'intégrabilité, $f(X_n) = M_n^f + P_n^f$ est la somme d'une martingale et d'un processus prévisible. Ceci fournit de nouvelles techniques de constructions de martingales.

Exemple. Posons $f(x) = x^2$, et considérons une martingale $(X_n)_{n\in\mathbb{N}}$ dans $\mathrm{L}^2(\Omega,\mathcal{B},\mathbb{P})$, c'est-à-dire que chaque X_n est de carré intégrable. Alors, la partie prévisible de X_n^2 est donnée par

$$P_{n+1} - P_n = \mathbb{E}[(X_{n+1})^2|\mathcal{F}_n] - (X_n)^2 = \mathbb{E}[(X_{n+1} - X_n)^2|\mathcal{F}_n] \geq 0.$$

Ainsi, il existe un unique processus prévisible, positif et croissant, appelé variation quadratique de $(X_n)_{n\in\mathbb{N}}$ et noté $\langle X\rangle_n$, tel que

$$\langle X\rangle_n = \sum_{k=0}^{n-1} \mathbb{E}[(X_{k+1} - X_k)^2|\mathcal{F}_k] \qquad ; \qquad ((X_n)^2 - \langle X\rangle_n)_{n\in\mathbb{N}} \text{ est une martingale.}$$

Dans le cas de la marche aléatoire simple sur $\mathbb{Z}$, on a $(X_{k+1} - X_k)^2 = 1$ presque sûrement pour tout k, donc $\langle X\rangle_n = n$ presque sûrement, et $((X_n)^2 - n)_{n\in\mathbb{N}}$ est une martingale.

Exemple. Soit $(X_n)_{n\in\mathbb{N}}$ une marche aléatoire sur $\mathbb{Z}$, avec $X_n = Y_1 + \cdots + Y_n$ somme d'i.i.d. et $\mathbb{E}[\mathrm{e}^{tY_1}] < \infty$ pour un certain réel positif t. La transformée exponentielle $Z_n = \mathrm{e}^{tX_n - n\log\mathbb{E}[\mathrm{e}^{tY_1}]}$ est de nouveau une martingale, puisque

$$\mathbb{E}[Z_{n+1}|\mathcal{F}_n] = Z_n\,\mathbb{E}[\mathrm{e}^{tY_{n+1} - \log\mathbb{E}[\mathrm{e}^{tY_1}]}] = Z_n\,\frac{\mathbb{E}[\mathrm{e}^{tY_{n+1}}]}{\mathbb{E}[\mathrm{e}^{tY_1}]} = Z_n.$$

Dans le cas de la marche aléatoire simple sur $\mathbb{Z}$, $(\frac{\mathrm{e}^{tX_n}}{(\cosh t)^n})_{n\in\mathbb{N}}$ est donc une martingale pour tout paramètre $t \in \mathbb{R}$.

1.3. Théorèmes d'arrêt. Une autre transformation de martingales qui préserve cette propriété est la *coupure par un temps d'arrêt*. Soit $(X_n)_{n\in\mathbb{N}}$ un processus adapté. Rappelons qu'un temps d'arrêt pour la filtration $(\mathcal{F}_n)_{n\in\mathbb{N}}$ est une variable aléatoire T à valeurs dans $[\![0,+\infty]\!]$ et telle que $\{T = n\} \in \mathcal{F}_n$ pour tout $n \in \mathbb{N}$.

Théorème 3.4 (Coupure par un temps d'arrêt). *Soit $(X_n)_{n\in\mathbb{N}}$ une martingale, et T un temps d'arrêt. Le processus*

$$X_{\inf(n,T)} = \begin{cases} X_n & \text{si } n < T, \\ X_T & \text{si } n \geq T \end{cases}$$

est une martingale. On a le même résultat avec des sur-martingales à la place de martingales.

En effet, c'est une conséquence de la proposition 3.3 avec le processus prévisible $A_n = \mathbb{1}_{T\geq n} = 1 - \mathbb{1}_{T<n}$.

Exemple. Soit $(X_n)_{n\in\mathbb{N}}$ une marche aléatoire simple sur $\mathbb{Z}$, $T = T_{-1}$ le temps d'atteinte de -1. Comme $(X_{\inf(n,T)})_{n\in\mathbb{N}}$ est une martingale, pour tout n,

$$\begin{aligned} 0 = \mathbb{E}[X_{\inf(n,T)}] &= -\mathbb{P}[T \leq n] + \mathbb{E}[X_n\,\mathbb{1}_{T>n}] \\ &= -\mathbb{P}[T \leq n] + \mathbb{E}[X_n|T > n]\,(1 - \mathbb{P}[T \leq n]), \end{aligned}$$

donc l'espérance conditionnelle de X_n sachant $X_0, \dots, X_n \geq 0$ est égale à

$$\frac{\mathbb{P}[T \leq n]}{1 - \mathbb{P}[T \leq n]} \to_{n \to \infty} +\infty.$$

Exemple. Soient $a, b > 0$, et $T = \inf(T_a, T_{-b})$. Comme $(X_{\inf(n,T)})_{n \in \mathbb{N}}$ est une martingale, pour tout n,

$$0 = \mathbb{E}[X_{\inf(n,T)}] = a\,\mathbb{P}[T_a \leq \inf(n, T_{-b})] - b\,\mathbb{P}[T_{-b} \leq \inf(n, T_a)] + \mathbb{E}[X_n\, \mathbb{1}_{\inf(T_a, T_{-b}) > n}].$$

Comme la marche aléatoire simple est récurrente, T_a et T_{-b} sont finis presque sûrement, donc la limite de l'expression précédente est

$$0 = a\,\mathbb{P}[T_a \leq T_{-b}] - b\,\mathbb{P}[T_{-b} \leq T_a],$$

puisqu'on a la borne $|\mathbb{E}[X_n\, \mathbb{1}_{\inf(T_a, T_{-b}) > n}]| \leq \max(a, b)\,\mathbb{P}[\inf(T_a, T_{-b}) > n] \to 0$. On en déduit :

$$\mathbb{P}[T_a \leq T_{-b}] = \frac{b}{b+a} \quad ; \quad \mathbb{E}[X_T] = 0.$$

En règle générale, si l'on n'a pas de bonnes bornes sur T ou sur X_T, il est faux que $\mathbb{E}[X_T] = \mathbb{E}[X_0]$ pour un temps d'arrêt T. Le théorème suivant donne des conditions suffisantes pour que ce soit vrai :

Théorème 3.5 (Doob). *Soit T un temps d'arrêt fini presque sûrement, et $(X_n)_{n \in \mathbb{N}}$ une martingale. Si $\mathbb{E}[|X_T|] < \infty$ et $\lim_{n \to \infty} \mathbb{E}[|X_n|\, \mathbb{1}_{T > n}] = 0$, alors $\mathbb{E}[X_T] = \mathbb{E}[X_0]$. Ces hypothèses sont satisfaites dans les cas suivants :*

(1) T est borné presque sûrement.

(2) $(X_{\inf(n,T)})_{n \in \mathbb{N}}$ est uniformément borné presque sûrement.

(3) T est intégrable et les incréments $X_{n+1} - X_n$ sont uniformément bornés par une constante.

La troisième condition suffisante du théorème est en particulier satisfaite lorsque $(X_n)_{n \in \mathbb{N}}$ est l'intégrale stochastique d'un processus prévisible borné contre une martingale bornée.

Exemple. On considère la marche aléatoire biaisée sur $\mathbb{Z}$ avec $\mathbb{P}[X_{n+1} = X_n + 1] = 1 - \mathbb{P}[X_{n+1} = X_n - 1] = p = 1 - q$, issue de $k > 0$. On considère le temps $T = \inf\{T_0, T_L\}$ de sortie de l'intervalle $[\![0, T]\!]$. Le processus $Z_n = (\frac{q}{p})^{X_n}$ est une martingale :

$$\mathbb{E}[Z_{n+1}|\mathcal{F}_n] = p\,\frac{q}{p}\,Z_n + q\,\frac{p}{q}\,Z_n = (q+p)\,Z_n = Z_n.$$

D'autre part, $(Z_{\inf(n,T)})_{n \in \mathbb{N}}$ est clairement borné par $\max(1, (\frac{p}{q})^L)$. Par le théorème d'arrêt de Doob,

$$\mathbb{E}[Z_T] = \mathbb{E}[Z_0] = \left(\frac{q}{p}\right)^k = P_0 + P_L\,\left(\frac{q}{p}\right)^L = P_0 + (1 - P_0)\,\left(\frac{q}{p}\right)^L,$$

avec $P_0 = \mathbb{P}[X_T = 0]$ et $P_L = \mathbb{P}[X_T = L]$. Ainsi,

$$P_0 = \frac{(\frac{q}{p})^k - (\frac{q}{p})^L}{1 - (\frac{q}{p})^L}.$$

2. Convergence de martingales

L'intérêt principal de la théorie des martingales est l'existence de puissants théorèmes de convergence, semblables à la loi des grands nombres. Dans cette section, nous présentons ces résultats en étudiant successivement la convergence p.s., dans $\mathrm{L}^1(\Omega, \mathcal{B}, \mathbb{P})$ et dans les espaces $\mathrm{L}^p(\Omega, \mathcal{B}, \mathbb{P})$ avec $p > 1$.

2.1. Convergence presque sûre. Soit $(X_n)_{n\in\mathbb{N}}$ un processus adapté, $a < b$ deux réels. On note $M_n(a, b)$ le plus grand entier k tel qu'existe deux suites entrelacées

$$0 \leq s_1 \leq t_1 \leq s_2 \leq t_2 \leq \cdots \leq s_k \leq t_k \leq n$$

avec $X_{s_i} \leq a$ et $X_{t_i} \geq b$ pour tout i. C'est le *nombre de montées* entre a et b, et c'est une variable aléatoire $\mathcal{F}_n$-mesurable. En effet, on peut associer à cette variable les temps d'arrêt

$$S_k = \inf\{n > T_{k-1},\ X_n \leq a\}$$
$$T_k = \inf\{n > S_k,\ X_n \geq b\}$$

et $M_n(a, b) = \sum_{k=1}^{\infty} \mathbb{1}_{T_k \leq n}$ est bien une somme de variables $\mathcal{F}_n$-mesurables.

Proposition 3.6 (Nombre de montées). *Soit $(X_n)_{n\in\mathbb{N}}$ une sur-martingale. On a*

$$\mathbb{E}[M_n(a, b)] \leq \frac{\mathbb{E}[(X_n - a)_-]}{b - a}.$$

En effet, soit $P_n = \sum_{k=1}^{\infty} \mathbb{1}_{S_k < n \leq T_k}$, qui est un processus prévisible puisque $\mathbb{1}_{S_k < n \leq T_k} = \mathbb{1}_{S_k < n}(1 - \mathbb{1}_{T_k < n})$. Comme $(X_n)_{n\in\mathbb{N}}$ est une sur-martingale et $(P_n)_{n\in\mathbb{N}}$ est un processus prévisible positif borné, $((P \cdot X)_n)_{n\in\mathbb{N}}$ est une sur-martingale. Or,

$$\begin{aligned}(P \cdot X)_n &= \sum_{i=1}^{n} P_i(X_i - X_{i-1}) = \sum_{k=1}^{M_n(a,b)} (X_{T_k} - X_{S_k}) + (X_n - X_{S_{M_n(a,b)+1}})\, \mathbb{1}_{S_{M_n(a,b)+1} < n} \\ &\geq (b - a)\, M_n(a, b) - (X_n - a)_-.\end{aligned}$$

Comme $0 = \mathbb{E}[(P \cdot X)_0] \geq \mathbb{E}[(P \cdot X)_n]$, l'inégalité est démontrée. On en déduit que les nombres de montées sont finis p.s. sous des hypothèses assez faibles sur $(X_n)_{n\in\mathbb{N}}$; ceci entraîne la convergence presque sûre. Plus précisément :

Théorème 3.7 (Convergence presque sûre des martingales). *Soit $(X_n)_{n\in\mathbb{N}}$ une sur-martingale. Il existe une limite p.s. $X_\infty = \lim_{n\to\infty} X_n$ si*

(1) $(X_n)_{n\in\mathbb{N}}$ est positive,

(2) ou plus généralement, si $(X_n)_{n\in\mathbb{N}}$ est bornée dans $\mathrm{L}^1(\Omega,\mathcal{B},\mathbb{P})$ *:*

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

Le théorème s'applique en particulier aux martingales. Par le lemme de Fatou, la limite X_∞ est intégrable, mais elle n'est pas forcément la limite dans $\mathrm{L}^1(\Omega,\mathcal{B},\mathbb{P})$ des variables X_n. En particulier, on n'a pas forcément $\mathbb{E}[X_\infty]=\mathbb{E}[X_0]$.

Exemple. Soit $(X_n)_{n\in\mathbb{N}}$ une marche aléatoire de paramètres p et $q=1-p$ sur $\mathbb{Z}$, par exemple issue de 0, et avec $p\neq q$. On sait que $Z_n=(\frac{q}{p})^{X_n}$ est une martingale, et elle est positive ; elle admet donc une limite p.s. Z_∞, qui est intégrable. Les valeurs possibles de Z_∞ sont dans l'adhérence des valeurs possibles de $(X_n)_{n\in\mathbb{N}}$, c'est-à-dire dans

$$\left\{\left(\frac{q}{p}\right)^k,\ k\in\mathbb{Z}\right\}\sqcup\{0\}.$$

Si $Z_\infty=(q/p)^k$ a probabilité positive pour une valeur k finie, alors avec probabilité positive $(X_n)_{n\in\mathbb{N}}$ est stationnaire égale à k. Or, une marche aléatoire n'est jamais stationnaire, donc $Z_\infty=0$ presque sûrement. On en déduit que $X_n\to-\infty$ p.s. si $p<q$, et $X_n\to+\infty$ p.s. si $p>q$. Notons que dans ce cas, on n'a pas convergence dans L^1, car $\mathbb{E}[X_0]=1\neq 0$.

Exemple. On considère une urne contenant initialement B boules blanches et N boules noires. À chaque étape, on tire une boule de l'urne, et l'on ajoute $c\geq 1$ boules de cette couleur dans l'urne. Soit X_n la proportion de boules blanches dans l'urne au temps n ; le nombre de boules blanches au temps n est alors $(B+N+cn)X_n$, et le nombre de boules noires au temps n est $(B+N+cn)(1-X_n)$. On a $X_0=\frac{B}{B+N}$ et

$$\mathbb{E}[X_{n+1}|\mathcal{F}_n]=X_n\,\frac{(B+N+cn)X_n+c}{B+N+c(n+1)}+(1-X_n)\,\frac{(B+N+cn)X_n}{B+N+c(n+1)}=X_n,$$

donc $(X_n)_{n\in\mathbb{N}}$ est une martingale à valeurs dans $[0,1]$. Elle est donc uniformément bornée, et admet une limite presque sûre X_∞. La loi de cette limite peut être calculée comme suit. La probabilité pour obtenir au cours des n premiers tirages n_1 tirages blancs et n_2 tirages noirs est

$$\binom{n}{n_1}\frac{B(B+c)(B+2c)\cdots(B+(n_1-1)c)\ N(N+c)(N+2c)\cdots(N+(n_2-1)c)}{(B+N)(B+N+c)\cdots(B+N+(n-1)c)}.$$

Supposons $n_1/n\to x\in(0,1)$. Alors l'expression ci-dessus a pour asymptotique

$$\frac{1}{n}\,\frac{\Gamma((B+N)/c)}{\Gamma(B/c)\,\Gamma(N/c)}\,x^{\frac{B}{c}-1}\,(1-x)^{\frac{N}{c}-1},$$

et d'autre part, $n_1/n\to x$ équivaut à $X_n\to x$. On en déduit que X_∞ est la loi de densité

$$\frac{\Gamma((B+N)/c)}{\Gamma(B/c)\,\Gamma(N/c)}\,x^{\frac{B}{c}-1}\,(1-x)^{\frac{N}{c}-1}\,dx$$

c'est-à-dire la loi $\beta(B/c,N/c)$. On verra plus loin qu'on a aussi convergence dans $\mathrm{L}^1(\Omega,\mathcal{B},\mathbb{P})$, et en fait dans tous les espaces L^p, car la martingale est à valeurs bornées.

2.2. Convergence dans L^1. Pour obtenir la convergence dans $\mathrm{L}^1(\Omega, \mathcal{B}, \mathbb{P})$ d'une suite de martingales bornées au sens du théorème 3.7, on a besoin d'une borne "uniforme", c'est-à-dire de la notion d'uniforme intégrabilité :

DÉFINITION 3.8. *Un processus (adapté, intégrable) $(X_n)_{n\in\mathbb{N}}$ est dit uniformément intégrable si*

$$\lim_{K\to+\infty} \sup_{n\in\mathbb{N}} \left(\mathbb{E}[|X_n|\, \mathbb{1}_{|X_n|\geq K}]\right) = 0.$$

EXEMPLE. La martingale associée au modèle d'urnes de Pólya de la section précédente est uniformément intégrable, puisqu'elle est bornée dans $\mathrm{L}^\infty(\Omega, \mathcal{B}, \mathbb{P})$. En revanche, la marche aléatoire simple n'est pas uniformément intégrable, car

$$\sup_{n\in\mathbb{N}} \left(\mathbb{E}[|X_n|\, \mathbb{1}_{|X_n|\geq K}]\right) \geq K \sup_{n\in\mathbb{N}} \left(\mathbb{P}[|X_n| \geq K]\right) = K.$$

Il en va de même pour la marche aléatoire biaisée de paramètres $p \neq q$.

THÉORÈME 3.9 (Convergence dans $\mathrm{L}^1(\Omega, \mathcal{B}, \mathbb{P})$ des martingales). *Soit $(X_n)_{n\in\mathbb{N}}$ une martingale. Les assertions suivantes sont équivalentes :*

1. *La martingale est uniformément intégrable.*
2. *La martingale converge p.s. et dans $\mathrm{L}^1(\Omega, \mathcal{B}, \mathbb{P})$ vers sa limite X_∞.*
3. *La martingale est fermée, c'est-à-dire qu'il existe une variable intégrable X telle que $X_n = \mathbb{E}[X|\mathcal{F}_n]$ pour tout n. Dans ce cas, on peut supposer X mesurable pour $\mathcal{F}_\infty = \bigcup_{n\in\mathbb{N}} \mathcal{F}_n$ en remplaçant X par $\mathbb{E}[X|\mathcal{F}_\infty]$, et on a alors $X = X_\infty = \lim_{n\to\infty} X_n$ presque sûrement.*

Dans tous ces cas, on a $\mathbb{E}[X_\infty] = \mathbb{E}[X_0]$. De plus, si T est un temps d'arrêt, alors $\mathbb{E}[X_\infty|\mathcal{F}_T] = X_T$, et donc en particulier $\mathbb{E}[X_T] = \mathbb{E}[X_\infty] = \mathbb{E}[X_0]$.

EXEMPLE. La martingale associée au modèle d'urnes de Pólya converge dans L^1 vers une variable aléatoire de loi $\beta(B/c, N/c)$, puisqu'elle est à valeurs bornées dans $[0, 1]$, et donc uniformément intégrable.

EXEMPLE. On considère de nouveau la marche aléatoire biaisée $X_n = \sum_{k=1}^n Y_k$, avec $\mathbb{P}[Y_k = 1] = 1 - \mathbb{P}[Y_k = -1] = p = 1 - q$; et l'on note $T = \inf(T_a, T_{-b})$, où a et b sont deux entiers positifs. Si t est un réel fixé, alors la transformée exponentielle

$$Z_n = \frac{\mathrm{e}^{tX_n}}{\mathbb{E}[\mathrm{e}^{tY_1}]^n} = \frac{\mathrm{e}^{tX_n}}{(p\,\mathrm{e}^t + q\,\mathrm{e}^{-t})^n}$$

est une martingale. La martingale arrêtée $(Z_{\inf(n,T)})_{n\in\mathbb{N}}$ est donc une martingale bornée dans L^∞, puisque le numérateur de Z_n est borné par $\sup(\mathrm{e}^{ta}, \mathrm{e}^{-tb})$, et le dénominateur est plus grand que 1 par convexité de l'exponentielle. Elle est donc en particulier uniformément intégrable, et sa limite $Z_T = Z_{\inf(\infty,T)} = \frac{\mathrm{e}^{tX_T}}{(p\,\mathrm{e}^t+q\,\mathrm{e}^{-t})^T}$ satisfait

$$\begin{aligned} 1 &= \mathbb{E}[Z_0] = \mathbb{E}[Z_T] \\ &= \mathrm{e}^{ta}\,\mathbb{E}\left[\frac{1}{(p\,\mathrm{e}^t + q\,\mathrm{e}^{-t})^T}\,\mathbb{1}_{X_T=a}\right] + \mathrm{e}^{-tb}\,\mathbb{E}\left[\frac{1}{(p\,\mathrm{e}^t + q\,\mathrm{e}^{-t})^T}\,\mathbb{1}_{X_T=-b}\right]. \end{aligned}$$

Posons $s^{-1} = p\,\mathrm{e}^t + q\,\mathrm{e}^{-t}$, de sorte que $\mathrm{e}^t = \lambda_\pm(s) = \frac{1\pm\sqrt{1-4pqs^2}}{2ps}$. Ces deux solutions donnent deux équations linéaires pour les transformées de Laplace de T sur les événements $X_T = a$ et $X_T = -b$:

$$\begin{aligned} 1 &= (\lambda_+(s))^a\,\mathbb{E}[s^T\,\mathbb{1}_{X_T=a}] + (\lambda_+(s))^{-b}\,\mathbb{E}[s^T\,\mathbb{1}_{X_T=-b}] \\ &= (\lambda_-(s))^a\,\mathbb{E}[s^T\,\mathbb{1}_{X_T=a}] + (\lambda_-(s))^{-b}\,\mathbb{E}[s^T\,\mathbb{1}_{X_T=-b}]. \end{aligned}$$

On conclut que

$$\begin{aligned} \mathbb{E}[s^T\,\mathbb{1}_{X_T=a}] &= \frac{(\lambda_+(s))^b - (\lambda_-(s))^b}{(\lambda_+(s))^{a+b} - (\lambda_-(s))^{a+b}}; \\ \mathbb{E}[s^T\,\mathbb{1}_{X_T=-b}] &= \frac{(q/p)^b\,((\lambda_+(s))^a - (\lambda_-(s))^a)}{(\lambda_+(s))^{a+b} - (\lambda_-(s))^{a+b}}; \\ \mathbb{E}[s^T] &= \frac{((\lambda_+(s))^b - (\lambda_-(s))^b) + (q/p)^b\,((\lambda_+(s))^a - (\lambda_-(s))^a)}{(\lambda_+(s))^{a+b} - (\lambda_-(s))^{a+b}}, \end{aligned}$$

ce qui donne la loi de T.

Exemple. Avec le même exemple, on peut calculer la loi de $S = \sup_{n\in\mathbb{N}} X_n$ si $p < q$. Par la loi des grands nombres, $(X_n)_{n\in\mathbb{N}}$ converge p.s. vers $-\infty$, donc S est une variable aléatoire positive finie presque sûrement. De plus,

$$\{S \geq a\} = \{T_a < \infty\} = \bigcup_{b=0}^{\infty} \{T_a < T_{-b}\}$$

la réunion d'événements étant décroissante. Or, si $T = \inf(T_a, T_{-b})$, alors $(q/p)^{X_{\inf(n,T)}}$ est une martingale bornée dans L^∞, donc uniformément intégrable dans $\mathrm{L}^1(\Omega, \mathcal{B}, \mathbb{P})$. Il s'ensuit :

$$1 = \mathbb{E}\left[(q/p)^{X_0}\right] = \mathbb{E}\left[(q/p)^{X_T}\right] = \mathbb{P}[T_a < T_{-b}]\left(\frac{q}{p}\right)^a + (1 - \mathbb{P}[T_a < T_{-b}])\left(\frac{q}{p}\right)^{-b},$$

donc $\mathbb{P}[T_a < T_{-b}] = \frac{1-(q/p)^{-b}}{(q/p)^a - (q/p)^{-b}}$, et par passage à la limite

$$\mathbb{P}[S \geq a] = \mathbb{P}[T_a < \infty] = \left(\frac{p}{q}\right)^a.$$

La variable S suit donc une loi géométrique de paramètre $1 - \frac{p}{q}$.

2.3. Convergence dans L^p. À moins de couper la martingale par un temps d'arrêt pour assurer son caractère borné dans L^∞, il peut être difficile de vérifier l'uniforme intégrabilité de $(X_n)_{n\in\mathbb{N}}$, et donc d'appliquer les théorèmes précédents. Néanmoins, si l'on suppose que $(X_n)_{n\in\mathbb{N}}$ est un processus dans $\mathrm{L}^p(\Omega,\mathcal{B},\mathbb{P})$ avec $p>1$, c'est-à-dire que chaque variable a sa puissance p intégrable, alors les hypothèses deviennent beaucoup plus simples à vérifier. L'inégalité suivante est à la base de ces simplifications :

Proposition 3.10 (Inégalité maximale de Doob). *Soit $(X_n)_{n\in\mathbb{N}}$ une martingale. On a pour tout $x>0$ et tout $p\geq 1$:*

$$\mathbb{P}\left[\sup_{k\leq n}|X_k|\geq x\right]\leq \frac{\mathbb{E}[|X_n|^p]}{x^p}.$$

Des inégalités du même type existent pour le maximum des valeurs d'une sur- ou d'une sous-martingale. Elles conduisent aux inégalités suivantes :

$$\mathbb{E}\left[\left(\sup_{k\leq n}|X_k|\right)^p\right]\leq \left(\frac{p}{p-1}\right)^p\ \mathbb{E}[|X_n|^p]$$

pour $p>1$ et une martingale dans $\mathrm{L}^p(\Omega,\mathcal{B},\mathbb{P})$.

Théorème 3.11 (Convergence dans $\mathrm{L}^p(\Omega,\mathcal{B},\mathbb{P})$ des martingales). *Soit $(X_n)_{n\in\mathbb{N}}$ une martingale dans $\mathrm{L}^p(\Omega,\mathcal{B},\mathbb{P})$. Si*

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

alors $(|X_n|^p)_{n\in\mathbb{N}}$ est uniformément intégrable, et $(X_n)_{n\in\mathbb{N}}$ converge p.s. et dans L^p vers sa limite X_∞.

Ainsi, il suffit de vérifier le caractère borné dans L^p pour obtenir une convergence forte.

Exemple. Dans le cas $p=2$, le processus variation quadratique $(\langle X\rangle_n)_{n\in\mathbb{N}}$ fournit une preuve simple du théorème précédent. En effet, supposant pour simplifier $X_0=0$ p.s., on a

$$\mathbb{E}[(X_n)^2]=\mathbb{E}[\langle X\rangle_n]=\sum_{k=0}^{n-1}\mathbb{E}[(X_{k+1}-X_k)^2]$$

pour tout $n\in\mathbb{N}$. Il s'ensuit que $(X_n)_{n\in\mathbb{N}}$ est borné dans L^2 si et seulement si

$$\sum_{k=0}^{\infty}(X_{k+1}-X_k)^2$$

est intégrable. Dans ce cas, notons que $X_n=\sum_{k=0}^{n-1}(X_{k+1}-X_k)$ est une décomposition de X_n en variables orthogonales pour le produit scalaire associé à la norme L^2 : en effet, si $k>j$, alors

$$\begin{aligned}\mathbb{E}[(X_{k+1}-X_k)(X_{j+1}-X_j)]&=\mathbb{E}[\mathbb{E}[X_{k+1}-X_k|\mathcal{F}_{j+1}](X_{j+1}-X_j)]\\&=\mathbb{E}[0\,(X_{j+1}-X_j)]=0.\end{aligned}$$

L'intégrabilité de $\sum_{k=0}^{\infty}\mathbb{E}[(X_{k+1}-X_k)^2]$ assure l'uniforme intégrabilité des sommes partielles $\sum_{k=0}^{n-1}\mathbb{E}[(X_{k+1}-X_k)^2]$, et donc l'uniforme intégrabilité de $((X_n)^2)_{n\in\mathbb{N}}$. On a donc bien équivalence entre le caractère borné dans L^2 et l'uniforme intégrabilité des carrés. De plus, ces conditions impliquent que $\sum_{k=0}^{\infty}(X_{k+1}-X_k)$ converge dans L^2, donc $(X_n)_{n\in\mathbb{N}}$ a bien une limite dans L^2 sous la condition $\sup_{n\in\mathbb{N}}\mathbb{E}[(X_n)^2]<\infty$. Finalement, la borne dans L^2 implique une borne dans L^1, donc on a aussi convergence p.s. par le théorème 3.7.

3. Processus de branchement

Les *processus de branchement* fournissent des exemples de martingales auxquelles s'appliquent l'ensemble des résultats précédents. Soit μ une mesure de probabilité sur $\mathbb{N}$, telle que $\mu(1)\neq 1$ et $\sum_{k=0}^{\infty}k^2\,\mu(k)<\infty$. On note $m=\sum_{k=0}^{\infty}k\,\mu(k)$. On fixe une famille doublement indexée $(\zeta_{n,k})_{n,k\in\mathbb{N}}$ de variables indépendantes et identiquement distribuées de loi μ. Le *processus de branchement de Galton-Watson* de loi μ est la suite de variables aléatoires $(Y_n)_{n\in\mathbb{N}}$ définies récursivement par

$$Y_0=1\quad;\quad Y_{n+1}=\sum_{k=1}^{Y_n}\zeta_{n,k}.$$

La variable Y_n représente le nombre d'individus de la n-ième génération aléatoire d'une population dont chaque individu a un nombre d'enfants donné par la loi μ et indépendant des autres progénitures aléatoires. La suite $(Y_n)_{n\in\mathbb{N}}$ est une chaîne de Markov de matrice de transition

$$P(k,l)=\mu^{*k}(l),\quad \text{où } \mu^{*k} \text{ est la } k\text{-ième convolée de } \mu.$$

Pour étudier cette chaîne (récurrence, transience, *etc.*), on introduit la martingale $(M_n=\frac{Y_n}{m^n})_{n\in\mathbb{N}}$. C'est bien une martingale, car

$$\begin{aligned}\mathbb{E}[M_{n+1}|\mathcal{F}_n]&=\frac{1}{m^{n+1}}\;\mathbb{E}\left[\sum_{k=1}^{r}\zeta_{n,k}\right]_{|r=Y_n}\\&=\frac{m\,Y_n}{m^{n+1}}=\frac{Y_n}{m^n}=M_n.\end{aligned}$$

En tant que martingale positive, $(M_n)_{n\in\mathbb{N}}$ a une limite presque sûre M_∞, qui est intégrable. On peut alors distinguer 3 cas :

(1) Supposons $m<1$. Alors, comme Y_n reste un entier pour tout n, en écrivant $Y_n\simeq M_\infty\,m^n$, on en déduit que $Y_n\to 0$ p.s., c'est-à-dire qu'il y a extinction presque sûre (notons que $r=0$ est un état absorbant de la chaîne de Markov $(Y_n)_{n\in\mathbb{N}}$).

(2) Si $m=1$, alors $Y_n=M_n$ converge presque sûrement. Hors, aucun état de la chaîne sauf 0 n'est absorbant, et d'autre part, une suite à valeurs entières qui converge est stationnaire. Si $Y_n\to r\neq 0$ avec probabilité non nulle, alors Y_n reste stationnaire à r avec probabilité non nulle. Mais par la propriété de Markov et le lemme de Borel-Cantelli, Y_n quitte p.s. l'état r une fois qu'il l'a atteint, ce

qui mène à une contradiction. On conclut que $Y_n \to 0$ p.s., c'est-à-dire qu'il y a de nouveau extinction.

(3) Supposons finalement $m > 1$. L'hypothèse sur $\sum_{k=0}^{\infty} k^2\, \mu(k) = h < \infty$ montre que la convergence de $(M_n)_{n\in\mathbb{N}}$ a lieu dans L^2, puisque

$$\mathbb{E}[(M_{n+1})^2] = \frac{\mathbb{E}[(\sum_{k=1}^{Y_n} \zeta_{n,k})^2]}{m^{2n+2}} = \frac{\mathbb{E}[Y_n]\, h + \mathbb{E}[Y_n(Y_n-1)]\, m^2}{m^{2n+2}}$$
$$= \mathbb{E}[(M_n)^2] + \frac{\mathbb{E}[Y_n]\,(h-m^2)}{m^{2n+2}} = \mathbb{E}[(M_n)^2] + \frac{(h-m^2)}{m^{n+2}},$$

et donc

$$\sup_{n\in\mathbb{N}} \mathbb{E}[(M_n)^2] \leq 1 + \sum_{n=0}^{\infty} \frac{h-m^2}{m^{n+2}} = \frac{h-m}{m^2-m} < \infty.$$

Par le théorème 3.11, $M_n \to M_\infty$ dans L^2 et p.s., et en particulier, $\mathbb{E}[M_\infty] = \mathbb{E}[M_0] = 1$, et $Y_n \to \infty$ avec probabilité positive.

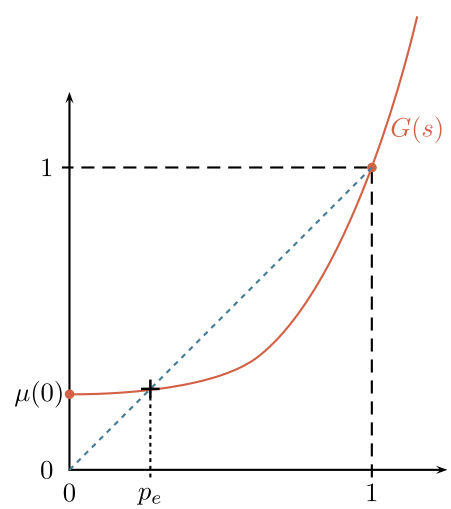
Dans le cas sur-critique $m > 1$, toujours sous l'hypothèse $\sum_{k=0}^{\infty} k^2 \mu(k) < \infty$, on peut montrer plus précisément :

Théorème 3.12 (Processus de branchement sur-critiques). *Soit* $G(s) = \sum_{k=0}^{\infty} \mu(k)\, s^k$ *la fonction génératrice de la loi* μ. *On suppose* $m = G'(1) > 1$, *et* $h = G''(1) + G'(1) < \infty$. *La probabilité d'extinction*

$$p_e = \mathbb{P}\left[\lim_{n\to\infty} Z_n = 0\right]$$

est l'unique point fixe dans $[0, 1)$ *de* G. *De plus, on a aussi* $p_e = \mathbb{P}[M_\infty = 0]$, *et donc, conditionnellement à la non-extinction,* $M_\infty > 0$, *c'est-à-dire que* Y_n *tend vers l'infini a vitesse exponentielle.*

Pour démontrer ceci, notons que G et toutes ses dérivées sont des fonctions croissantes sur $[0, 1]$, de sorte que G est convexe et a pour allure :



La valeur de G a l'origine est $\mu(0) \geq 0$, $G(1) = 1$, et il existe un unique autre point $t \in [0,1)$ tel que $G(t) = t$. La fonction génératrice G permet de calculer toutes les fonctions génératrices des populations Y_n, car

$$\mathbb{E}[s^{Y_{n+1}}] = \mathbb{E}[\mathbb{E}[s^{Y_{n+1}}|Y_n]] = \mathbb{E}[G(s)^{Y_n}],$$

de sorte que par récurrence,

$$\mathbb{E}[s^{Y_n}] = G \circ G \circ \cdots \circ G(s) = G^{\circ n}(s).$$

Soit $p_{e,n} = \mathbb{P}[Y_n = 0]$. C'est aussi $G^{\circ n}(0)$, et comme G est croissante $0 \leq G(0) = \mu(0)$, on en déduit la croissance de $Y_{e,n}$ vers sa limite p_e. Comme G est continu, $p_e = G(p_e)$, donc p_e est un point fixe de G. Comme $p_e \neq 1$, $p_e = t$ est l'unique point fixe de G dans $[0,1)$.

Considérons maintenant la probabilité $p_m = \mathbb{P}[M_\infty = 0]$. Notons que compte tenu de la convergence dans L^2, $\mathbb{E}[M_\infty] = 1$, donc $p_m < 1$. D'autre part, p_m est un point fixe de G. En effet, considérons les processus de Galton-Watson $(Y_n^{(i)})_{i \in [\![1,Y_1]\!]}$ correspondant aux Y_1 individus de la première génération. Conditionnellement à Y_1, il s'agit de Y_1 processus de Galton-Watson indépendants, de même loi que le processus originel, et donc tels qu'existent des limites

$$M_\infty^{(i)} = \lim_{n \to \infty} \frac{Y_n^{(i)}}{m^n}$$

indépendantes et toutes de même loi que M_∞. On a alors $M_\infty = \frac{1}{m} \sum_{i=1}^{Y_1} M_\infty^{(i)}$, et donc

$$p_m = \mathbb{P}[M_\infty = 0] = \mathbb{E}[\mathbb{P}[M_\infty = 0|Y_1]] = \mathbb{E}[(p_m)^{Y_1}] = G(p_m)$$

car M_∞ est nulle si et seulement si tous les $M_\infty^{(i)}$ le sont. Par conséquent, $p_m = p_e = t$ est de nouveau l'unique point fixe de G dans $[0,1)$.

Références. Pour la théorie générale des martingales discrètes, on renvoie à

(1) D. Williams, *Probability with Martingales*, Cambridge University Press, 1991 ; Part II.

(2) G. Grimmett and D. Stirzaker, *Probability and Random Processes*, 3rd edition, Oxford University Press, 2001 ; §12.

L'extension au cas des temps continus et l'application à l'étude des mouvements browniens est traité de façon exhaustive dans

(3) D. Revuz and M. Yor, *Continuous Martingales and Brownian Motion*, 3rd edition, Grundlehren der Mathematischen Wissenschaften 293, Springer-Verlag, 1999.

Enfin, pour la théorie des processus de branchement, et de nombreuses généralisations de l'étude du paragraphe 3, on conseille

(4) T. Harris, *The Theory of Branching Processes*, Dover, 1989.

Le chapitre 1 de cet ouvrage traite en particulier des processus de Galton-Watson simples étudiés ici.