L'essentiel sur les martingales

Le minimum, en bref. Une *martingale* est une suite $(M_n)_{n\geq 0}$ de variables aléatoires qui vérifie

$$\mathbb{E}[M_{n+1}|M_0,\ldots,M_n] = M_n.$$

On parle de *sous-martingale* si l'égalité est remplacée par $\geq$ et de *surmartingale* avec $\leq$. Une sous-martingale a donc tendance à croître, en moyenne, et inversement pour une surmartingale, tandis qu'une martingale reste centrée autour de la valeur précédente. On note souvent $\mathcal{F}_n = \sigma(M_0,\ldots,M_n)$ (tribu engendrée par $M_0,\ldots,M_n$) de sorte que la définition s'écrit $\mathbb{E}[M_{n+1}|\mathcal{F}_n] = M_n$. La suite $(\mathcal{F}_n)_n$ est une suite croissante de tribus, ce qu'on appelle une *filtration*.
On a par récurrence, pour tout n, $\mathbb{E}[M_n] = \mathbb{E}[M_0]$ (et $\mathbb{E}[M_n] \geq \mathbb{E}[M_0]$ pour une sous-martingale, $\mathbb{E}[M_n] \leq \mathbb{E}[M_0]$ pour une surmartingale).
Si τ est un temps d'arrêt (par exemple un temps d'atteinte), c'est-à-dire que l'événement $\{\tau = n\}$ ne dépend que de $M_0,\ldots,M_n$ pour tout n, alors la *martingale arrêtée au temps* τ, $M^\tau = (M_{n\wedge\tau})_{n\geq 0}$ (qui vaut M_n pour $n \leq \tau$ et M_τ ensuite), est encore une martingale. De même pour sur/sous-martingale. Notamment, si τ est un temps d'arrêt *borné*, on en déduit $\mathbb{E}[M_\tau] = \mathbb{E}[M_0]$: ceci est appelé un *théorème d'arrêt*.
Les principaux théorèmes sont les théorèmes de convergence : une sous-martingale majorée converge p.s., de même par symétrie pour une surmartingale minorée (comme pour les suites croissantes et décroissantes), et donc **(essentiel)** une martingale majorée ou minorée converge p.s.. En particulier, une martingale positive converge p.s.. Il y a un autre théorème de convergence important : si une martingale est bornée dans L^p pour un $p \geq 1$, c'est-à-dire que $E[|M_n|^p] \leq C$ pour tout n, alors elle converge p.s. et aussi, si $p > 1$, dans L^p.
Enfin, les (sur,sous)-martingales vérifient diverses inégalités, notamment l'*inégalité maximale de Doob* : si $(M_n)_{n\geq 0}$ est une sous-martingale positive, alors pour tout $a > 0$ et tout $n \geq 0$,

$$\mathbb{P}\Big(\max_{0\leq k\leq n} M_k \geq a\Big) \leq \frac{\mathbb{E}[M_n]}{a}.$$

(Pour s'en rappeler : l'inégalité de Markov donne $\mathbb{P}(M_n \geq a) \leq \frac{\mathbb{E}[M_n]}{a}$ pour toute v.a. positive M_n, et ici intuitivement on peut, par « croissance », remplacer M_n par $\max(M_0,\ldots,M_n)$)

Soit $(\Omega,\mathcal{F},\mathbb{P})$ un espace de probabilité.

Définition 1. Une *filtration* sur $(\Omega,\mathcal{F})$ est une suite croissante $(\mathcal{F}_n)_{n\geq 0}$ de tribus incluses dans $\mathcal{F}$:

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

Note sur les tribus. Une tribu sur Ω est un ensemble de parties qui contient $\emptyset$, est stable par complémentaire et par union dénombrable. Une tribu contient donc toujours Ω et est stable aussi par intersection dénombrable. Les éléments d'une tribu $\mathcal{F}$ sont dits $\mathcal{F}$-mesurables (ou mesurables par rapport à $\mathcal{F}$). Une mesure se définit sur une tribu (c'est une fonction $\mu : \mathcal{F} \to \mathbb{R}_+$ qui est additive sur les réunions disjointes dénombrables), et on introduit d'ailleurs en général en cours les tribus en les motivant par des raisons techniques : on a besoin de parler des boréliens (et des lebesguiens) de $\mathbb{R}$ car on ne peut définir la mesure de Lebesgue de n'importe quelle partie de $\mathbb{R}$. En probabilités, elles ont également un sens intuitif d'« information », le cas le plus clair étant celui de la tribu engendrée par une (ou plusieurs) variable aléatoire : pour une v.a. réelle X, $\sigma(X) = \{X^{-1}(B)|B \in \mathcal{B}(\mathbb{R})\}$ est la tribu qui contient tous les événements relatifs à la seule variable aléatoire X. Connaître $\sigma(X)$, c'est-à-dire savoir si les événements de cette tribu sont réalisés ou non, c'est connaître la valeur de X. Plus généralement, une sous-tribu de $\mathcal{F}$ peut être vue comme un ensemble d'événements dont on peut avoir connaissance de la réalisation ou non [1]. C'est une intuition à garder à l'esprit quand on parle d'espérance conditionnelle ou de probabilité conditionnelle *sachant une tribu* $\mathcal{F}$, ce qui de prime abord paraît très abstrait. Dans le cas d'une filtration $(\mathcal{F}_n)_{n\geq 0}$, on considère en général n comme un temps, et $\mathcal{F}_n$ comme l'information apportée par le passé jusqu'au temps n : plus le temps passe, plus on accumule d'information.

1. c'est cohérent avec la définition de tribu : si on sait si des événements A et B sont satisfaits, alors on le sait logiquement pour A^c, et pour $A \cup B$ aussi, ou d'ailleurs pour une réunion quelconque d'événements que l'on connaît. Au-delà des raisons techniques, on peut éventuellement justifier la restriction aux réunions dénombrables en disant que, pour savoir si $\bigcup_{i\in I} A_i$ est vrai, il faut savoir si l'un des A_i est vrai, ce qui demande de considérer les A_i un à un, opération qui n'est possible que si I est dénombrable. Ainsi, les événements dont on connaît le statut (réalisé ou non) forment une tribu.

Définition 2. Soit $(\mathcal{F})_n$ une filtration. Une suite $(M_n)_{n\geq 0}$ de variables aléatoires intégrables est une *martingale (resp. surmartingale, resp. sous-martingale) par rapport à la filtration* $(\mathcal{F}_n)_{n\geq 0}$ si, pour tout $n \geq 0$,
– M_n est $\mathcal{F}_n$-mesurable (on dit que $(M_n)_n$ est *adaptée à la filtration*) ;
– $\mathbb{E}[M_{n+1}|\mathcal{F}_n] = M_n$ (resp. $\leq M_n$, resp. $\geq M_n$).

La plupart du temps, la filtration est en fait définie à partir de la suite $(M_n)_{n\geq 0}$ par

$$\mathcal{F}_n = \sigma(M_0, \ldots, M_n)$$

(tribu du passé avant l'instant n). On dit que c'est la filtration associée au processus $(M_n)_{n\geq 0}$, ou simplement la filtration naturelle. Si on ne précise pas la filtration quand on parle d'une martingale, c'est que l'on considère implicitement celle-ci. Le premier point de la définition est alors automatique. Et le second point de la définition peut s'écrire

$$\mathbb{E}[M_{n+1}|M_0, \ldots, M_n] = M_n \quad (\text{resp. } \leq M_n, \text{ resp. } \geq M_n).$$

Une martingale est donc une suite de variables aléatoires qui à chaque pas est centrée par rapport à la valeur précédente, même si on connaît son passé jusque-là (on parle de « passé » en voyant n comme un « temps »). Une surmartingale a quant à elle une tendance à décroître à chaque pas, tandis qu'une sousmartingale a une tendance à croître à chaque pas.

Les (sous,sur)martingales jouent un rôle d'outil en probabilités que l'on peut rapprocher de celui des suites croissantes et décroissantes en analyse réelle : on verra qu'une sous-martingale majorée converge et qu'une surmartingale minorée converge (et donc qu'une martingale minorée ou majorée converge !), et que l'on peut obtenir diverses inégalités sur celles-ci, de sorte qu'il est souvent pratique d'introduire une (sous,sur)martingale liée au processus que l'on étudie.

Exemples de base, à avoir en tête (on utilise la filtration naturelle dans les deux premiers exemples)

1. Si $(\xi_n)_{n\geq 0}$ est une suite de v.a. indépendantes, et
$$S_n = \xi_1 + \cdots + \xi_n,$$
alors on a $\mathbb{E}[S_{n+1}|\mathcal{F}_n] = \mathbb{E}[S_n + \xi_{n+1}|\mathcal{F}_n] = \mathbb{E}[S_n|\mathcal{F}_n] + \mathbb{E}[\xi_{n+1}|\mathcal{F}_n] = S_n + \mathbb{E}[\xi_{n+1}]$ (car S_n est $\mathcal{F}_n$-mesurable, et ξ_{n+1} est indépendante de $\mathcal{F}_n$), donc $(S_n)_{n\geq 0}$ est une martingale si $\mathbb{E}[\xi_n] = 0$ pour tout n (sous-martingale si ≥ 0, surmartingale si ≤ 0)
2. Si $(\xi_n)_{n\geq 0}$ est une suite de v.a. indépendantes, et
$$U_n = \xi_1 \cdots \xi_n,$$
alors on a $\mathbb{E}[U_{n+1}|\mathcal{F}_n] = \mathbb{E}[U_n\xi_{n+1}|\mathcal{F}_n] = U_n\mathbb{E}[\xi_{n+1}|\mathcal{F}_n] = U_n\mathbb{E}[\xi_{n+1}]$ (pour les mêmes raisons que plus haut), donc $(U_n)_n$ est une martingale si $\mathbb{E}[\xi_n] = 1$ pour tout n (sous-martingale si ≥ 1, surmartingale si ≤ 1).
3. Si X est une variable aléatoire intégrable, et $(\mathcal{F}_n)_{n\geq 0}$ est une filtration, alors la suite de terme général
$$M_n = \mathbb{E}[X|\mathcal{F}_n]$$
est une martingale. En effet, $\mathbb{E}[M_{n+1}|\mathcal{F}_n] = \mathbb{E}[\mathbb{E}[X|\mathcal{F}_n]|\mathcal{F}_{n+1}]$ et $\mathcal{F}_n \subset \mathcal{F}_{n+1}$ donc $\mathbb{E}[M_{n+1}|\mathcal{F}_n] = \mathbb{E}[X|\mathcal{F}_n] = M_n$ (propriété du double conditionnement : seul reste le conditionnement par la tribu la plus petite).

Le premier exemple peut se voir comme la suite des gains cumulés à un jeu de hasard en répétant la même mise. La martingale correspond à un jeu équitable, la surmartingale à un jeu défavorable (au joueur) et la sous-martingale à une jeu favorable. *Dans le langage courant, une martingale est une stratégie de jeu qui mène à un succès assuré, ce qui serait plutôt une sous-martingale au sens mathématique. Et, en pratique, tous les jeux de casinos fournissent plutôt des surmartingales... Une conséquence de la théorie que l'on va voir est qu'il n'est pas possible, via une stratégie de mises, de gagner à coup sûr à partir d'une martingale ou d'une surmartingale, en tout cas si les mises sont par exemple bornées.*

Le premier exemple (et le deuxième) est assez particulier puisque les incréments sont indépendants, alors que pour avoir une martingale on pourrait simplement supposer que $\mathbb{E}[\xi_{n+1}|\xi_0, \ldots, \xi_n] = 0$ pour tout n, ce qui autoriserait ξ_{n+1} à dépendre de $\xi_0, \ldots, \xi_n$, du moment que la variable reste centrée (d'espérance nulle) quels que soient $\xi_0, \ldots, \xi_n$. Par exemple, ceci est vérifié en prenant $\xi_{n+1} = \xi_n Y_{n+1}$ où Y_{n+1} est indépendante de $\xi_0, \ldots, \xi_n$ et $\mathbb{E}[Y_{n+1}] = 0$, ce qui montre que

$$M_n = Y_0 + Y_0Y_1 + \cdots + Y_0 \cdots Y_n$$

est une martingale si les Y_i sont indépendantes et centrées.

Cela dit, le premier exemple reste essentiel, en soi et comme point de départ pour construire d'autres sur- ou sous-martingales. Avant cela, donnons juste quelques conséquences directes de la définition :

Proposition 1. Soit $(M_n)_{n\geq 0}$ une martingale.
a) Pour tout $n \geq 0$, $\mathbb{E}[M_n] = \mathbb{E}[M_0]$.
b) Pour tous $0 \leq m \leq n$, $\mathbb{E}[M_n|\mathcal{F}_m] = M_m$.

La même chose vaut pour les sur- et sous-martingales, avec les inégalités correspondantes.

1 Transformées de martingales

Définition 3. Soit $M = (M_n)_{n\geq 0}$ une (sur,sous)martingale par rapport à la filtration $(\mathcal{F}_n)_n$, et $\phi = (\phi_n)_{n\geq 1}$ un processus prévisible, c'est-à-dire que ϕ_n est bornée et $\mathcal{F}_{n-1}$ mesurable pour tout $n \geq 1$. La *transformée de martingale* définie par M et ϕ est le processus $(\langle M, \phi\rangle_n)_{n\geq 0}$ défini par, pour tout $n \geq 0$,

$$\langle M, \phi\rangle_n = M_0 + \sum_{k=1}^{n}(M_k - M_{k-1})\phi_k.$$

L'intuition est la suivante : dans un jeu de hasard où une mise constante égale à 1 donnerait des gains cumulés $(M_n)_{n\geq 0}$ (comme dans le premier exemple, disons), ϕ_n correspond à la mise choisie au temps n, qui peut dépendre des résultats déjà observés $M_0, \ldots, M_{n-1}$, et qui mène à un gain $(M_n - M_{n-1})\phi_n$ vu que $M_n - M_{n-1}$ est le résultat du jeu au temps n pour une mise égale à 1. Donc $\langle M, \phi\rangle_n$ est le capital au temps n si on applique la stratégie ϕ.

Proposition 2. Si $(M_n)_{n\geq 0}$ est une martingale, alors $\langle M, \phi\rangle$ est une martingale.
Si $(M_n)_{n\geq 0}$ est une sous-martingale (resp. surmartingale) et $\phi_n \geq 0$ pour tout n, alors $\langle M, \phi\rangle$ est une sous-martingale (resp. une surmartingale).

(La preuve est directe)
Autrement dit, une stratégie, aussi futée soit-elle, ne permettra pas d'obtenir autre chose qu'une surmartingale si le jeu est défavorable.
Une des stratégies les plus simples consiste à s'arrêter de miser à partir d'un certain temps T. Ce temps doit être décidé en fonction des valeurs déjà observées : le fait que $T = n$ ou non ne doit dépendre que de $M_0, \ldots, M_n$. Ce doit donc être un temps d'arrêt :

Définition 4. Une variable aléatoire T à valeurs dans $\mathbb{N}$ est un *temps d'arrêt* par rapport à la filtration $(\mathcal{F}_n)_n$ si, pour tout $n \in \mathbb{N}$,

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

Si M est une (sur,sous)martingale, et T un temps d'arrêt, la *(sur,sous)martingale arrêtée* au temps T est donnée par

$$M_n^T = M_{T\wedge n} = \begin{cases} M_n & \text{si } n \leq T \\ M_T & \text{si } n \geq T. \end{cases}$$

(où on note $a \wedge b = \min(a, b)$). C'est la transformée $\langle M, \phi\rangle$ avec $\phi_n = \mathbf{1}_{\{n\leq T\}}$ (on mise 1 avant T et 0 après), donc par la proposition c'est une (sur,sous)martingale.
En particulier, on a $\mathbb{E}[M_{n\wedge T}] = \mathbb{E}[M_0]$ pour tout n dans le cas martingale. Si T est borné ($T \leq t$ p.s.), en choisissant un $n \geq t$ on a donc $\mathbb{E}[M_T] = \mathbb{E}[M_0]$. La conclusion vaut aussi si T est fini p.s. et $|M_{n\wedge T}| \leq Z$ avec Z intégrable (par exemple Z constante), car $M_{n\wedge T} \underset{n}{\longrightarrow} M_T$ (c'est une suite stationnaire) et on peut alors appliquer le théorème de convergence dominée. On verra un exemple à la fin.
Donnons une inégalité importante qui se prouve avec les martingales arrêtées :

Théorème 1 – *Inégalité maximale de Doob.* Soit $(M_n)_{n\geq 0}$ une sous-martingale positive. Pour tout $a > 0$ et $n \in \mathbb{N}$,

$$\mathbb{P}\Big(\max_{0\leq k\leq n} M_k \geq a\Big) \leq \frac{\mathbb{E}[M_n]}{a}.$$

On peut signaler une autre transformation usuelle :

Proposition 3. Soit $\varphi : \mathbb{R} \to \mathbb{R}_+$ une fonction convexe, et $(X_n)_n$ une suite de variables aléatoires telle que $E[\varphi(X_n)] < \infty$ pour tout $n \in \mathbb{N}$.

a) Si $(X_n)_{n\geq 0}$ est une martingale, alors $(\varphi(X_n))_{n\geq 0}$ est une sous-martingale.

b) Si $(X_n)_{n\geq 0}$ est une sous-martingale et si φ est croissante, alors $(\varphi(X_n))_{n\geq 0}$ est une sous-martingale.

Par exemple, si $(M_n)_n$ est une martingale, alors $(|M_n|)_n$ est une sous-martingale, de même pour $(M_n^2)_n$.

2 Théorèmes de convergence

Le théorème **fondamental** est le suivant :

Théorème 2. Soit $(M_n)_{n\geq 0}$ une surmartingale positive. Il existe une variable aléatoire M_∞ telle que M_n converge vers M_∞ presque sûrement.

On pourrait étendre le théorème (par translation) en disant qu'une surmartingale minorée par une constante converge presque sûrement. Il en va de même, par symétrie, d'une sous-martingale majorée. Et donc puisqu'une martingale est à la fois une sous-martingale et une surmartingale, elle converge dès qu'elle est majorée ou minorée.
On dispose d'un second théorème de convergence :

Théorème 3. Soit $(M_n)_{n\geq 0}$ une martingale, et $p > 1$.
Si $(M_n)_{n\geq 0}$ est bornée dans L^p (c'est-à-dire qu'il existe une constante C telle que $\mathbb{E}[|M_n|^p] < C$ pour tout $n \geq 0$), alors il existe une variable aléatoire $X_\infty \in L^p$ telle que X_n converge p.s. et dans L^p vers X_∞.

Attention, si $p = 1$, on peut montrer qu'il y a convergence p.s. vers $X_\infty \in L^1$ mais pas nécessairement convergence dans L^1.

3 Quelques exemples d'utilisation

3.1 Ruine du joueur

Soit $(\xi_n)_{n\geq 1}$ une suite de v.a. indépendantes de même loi donnée par $\mathbb{P}(\xi_1 = 1) = p$ et $\mathbb{P}(\xi_1 = -1) = q = 1 - p$.
On considère la suite

$$S_n = \xi_1 + \cdots + \xi_n.$$

Alors $\mathbb{E}[\xi_1] = p - q = 2p - 1$, donc $(S_n - (p-q)n)_{n\geq 0}$ est une martingale.
Soit $a, b \in \mathbb{N}^*$. On s'intéresse au temps d'arrêt $\tau = \min\{n \in \mathbb{N} \mid S_n = -a \text{ ou } S_n = b\}$ (si S_n atteint d'abord a, c'est la ruine, sinon la réussite). On recherche la probabilité de ruine $\mathbb{P}(S_\tau = -a)$ et le temps moyen de ruine ou succès $\mathbb{E}[\tau]$.
Comme τ est un temps d'arrêt, $(S_{n\wedge\tau} - (p-q)(n\wedge\tau))_{n\geq 0}$ est une martingale.
Par la théorie des chaînes de Markov (par exemple), on sait que $\tau < \infty$ p.s. (une chaîne de Markov irréductible finie est récurrente). On peut aussi le justifier par un argument de martingale : la martingale ci-dessus est majorée si $p \geq q$ (par b) et minorée si $p \leq q$ (par $-a$), donc dans les deux cas elle converge p.s., ce qui implique $\tau < \infty$ p.s. (en effet, à supposer que $\tau = \infty$, si $p \neq q$, la martingale divergerait vers $\pm\infty$ vu que $-a \leq S_{n\wedge\tau} \leq b$, et si $p = q$, les accroissements seraient toujours de ± 1 ce qui rendrait la convergence impossible).
Donc $S_{n\wedge\tau} \underset{n}{\longrightarrow} S_\tau$, et par ailleurs $\mathbb{E}[S_\tau] = -a\mathbb{P}(S_\tau = -a) + b\mathbb{P}(S_\tau = b)$.
Par la propriété de martingale, $\mathbb{E}[S_{n\wedge\tau} - (p-q)(n\wedge\tau)] = \mathbb{E}[S_0 - 0] = 0$.
Enfin, $\mathbb{E}[S_{n\wedge\tau} - (p-q)(n\wedge\tau)] = \mathbb{E}[S_{n\wedge\tau}] - (p-q)\mathbb{E}[n\wedge\tau]$ et, d'une part la suite $n\wedge\tau$ est croissante, positive, et stationne en τ, donc $\mathbb{E}[n\wedge\tau] \underset{n}{\longrightarrow} \mathbb{E}[\tau]$ par théorème de convergence monotone ; et d'autre part, pour $0 \leq n \leq \tau$, $-a \leq S_n \leq b$, donc pour tout n on a $-a \leq S_{n\wedge\tau} \leq b$, et $S_{n\wedge\tau} = S_\tau$ pour n grand donc par théorème de convergence dominée $\mathbb{E}[S_{n\wedge\tau}] \underset{n}{\longrightarrow} \mathbb{E}[S_\tau]$.
Au final, on a donc

$$0 = \mathbb{E}[S_{n\wedge\tau} - (p-q)(n\wedge\tau)] \underset{n}{\longrightarrow} \mathbb{E}[S_\tau] - (p-q)\mathbb{E}[\tau].$$

Dans le cas $p = q = \frac{1}{2}$, on a donc

$$0 = \mathbb{E}[S_\tau] = -a\mathbb{P}(S_\tau = -a) + b\mathbb{P}(S_\tau = b) = -a\mathbb{P}(S_\tau = -a) + b\big(1 - \mathbb{P}(S_\tau = -a)\big)$$

d'où

$$\mathbb{P}(S_\tau = -a) = \frac{b}{b+a}.$$

Dans le cas $p \neq q$, on ne peut conclure tout de suite. Par contre on remarque que la suite $U_n = \left(\dfrac{p}{q}\right)^{S_n}$ est une martingale. Notamment, $\mathbb{E}[U_{n\wedge\tau}] = \mathbb{E}[U_0]$ puisque $U_{n\wedge\tau}$ est une martingale.
On en déduit de la même manière que plus haut (par théorème de convergence dominée) que

$$1 = \mathbb{E}[U_0] \underset{n}{=\!\!\longrightarrow} \mathbb{E}[U_{n\wedge\tau}] = \mathbb{E}[U_\tau] = \left(\frac{p}{q}\right)^{-a}\mathbb{P}(S_\tau = -a) + \left(\frac{p}{q}\right)^{b}\mathbb{P}(S_\tau = b)$$

et $\mathbb{P}(S_\tau = b) = 1 - \mathbb{P}(S_\tau = -a)$ d'où

$$\mathbb{P}(S_\tau = -a) = \frac{1 - \left(\frac{p}{q}\right)^b}{\left(\frac{p}{q}\right)^{-a} - \left(\frac{p}{q}\right)^b}$$

et par l'identité obtenue plus haut

$$\mathbb{E}[\tau] = \frac{1}{p-q}\mathbb{E}[S_\tau] = \frac{1}{p-q}\left(-a\frac{1-\left(\frac{p}{q}\right)^{b}}{\left(\frac{p}{q}\right)^{-a}-\left(\frac{p}{q}\right)^{b}} + b\frac{\left(\frac{p}{q}\right)^{-a}-1}{\left(\frac{p}{q}\right)^{-a}-\left(\frac{p}{q}\right)^{b}}\right)$$

On peut aussi obtenir $\mathbb{E}[\tau]$ quand $p = q$ par une méthode similaire, mais cette fois-ci en remarquant le fait qu'alors $S_n^2 - n$ est une martingale. C'est moins évident qu'avant :

$$(S_{n+1})^2 = (S_n + \xi_{n+1})^2 = (S_n)^2 + 2S_n\xi_{n+1} + \xi_{n+1}^2$$

d'où

$$\mathbb{E}[S_{n+1}^2|\mathcal{F}_n] = S_n^2 + 2S_n\mathbb{E}[\xi_{n+1}] + \mathbb{E}[\xi_{n+1}^2] = S_n^2 + 0 + 1 = S_n^2 - n + (n+1)$$

ce qui donne $\mathbb{E}[S_{n+1}^2 - (n+1)|\mathcal{F}_n] = S_n^2 - n^2$ et ainsi le fait annoncé : $S_n^2 - n$ est une martingale. Notamment, comme avant,

$$0 = \mathbb{E}[S_{n\wedge\tau}^2 - (n\wedge\tau)] = \mathbb{E}[S_{n\wedge\tau}^2] - \mathbb{E}[n\wedge\tau] \underset{n}{\longrightarrow} \mathbb{E}[S_\tau^2] - \mathbb{E}[\tau] = a^2\mathbb{P}(S_\tau = -a) + b^2\mathbb{P}(S_\tau = b) - \mathbb{E}[\tau]$$

et donc

$$\mathbb{E}[\tau] = a^2\frac{b}{a+b} + b^2\frac{a}{a+b} = ab.$$

3.2 Série aléatoire

Supposons que les ξ_n soient indépendantes et de même loi dans L^2, centrés ($\mathbb{E}[\xi_k] = 0$) et que la suite réelle $(a_n)_n$ est telle que $\sum_n (a_n)^2 < \infty$.
Alors $M_n = \sum_{k=1}^n a_k\xi_k$ est une martingale bornée dans L^2 : en effet c'est le premier exemple de martingale, et vu que $\mathbb{E}[M_n] = 0$,

$$\mathbb{E}[(M_n)^2] = \mathrm{Var}(M_n) = \sum_{k=1}^n a_k^2\,\mathrm{Var}(\xi_k) = \mathrm{Var}(\xi_1)\sum_{k=1}^n a_k^2 \le \mathrm{Var}(\xi_1)\sum_{k=1}^\infty a_k^2.$$

Par suite, $(M_n)_{n\ge 0}$ converge p.s. et dans L^2. Autrement dit, la série $\sum_n a_n\xi_n$ converge p.s. et dans L^2.
On peut noter que par exemple si de plus $\sum_n |a_n| = \infty$ (disons $a_n = \frac{1}{n}$) et que ξ_n suit la loi uniforme sur $\{-1, 1\}$, on a $\sum_n |a_n\xi_n| = \sum_n |a_n| = \infty$ p.s. donc la convergence n'est p.s. pas absolue.

3.3 Modèle de Wright-Fischer

Soit $p \in]0, 1[$. On considère une chaîne de Markov d'espace d'états $\{0, 1, \dots, N\}$ de transition donnée par :

$$p(i,j) = \binom{N}{j}\left(\frac{i}{N}\right)^j\left(1-\frac{i}{N}\right)^{N-j}$$

(où $0 \le i, j \le N$). Autrement dit, sachant X_n, la variable aléatoire X_{n+1} suit la loi binomiale de paramètres N et $\frac{X_n}{N}$. Supposons que $X_0 = x_0$ est fixé.
Cette chaîne de Markov intervient dans la modélisation de la répartition des allèles d'un gène dans une population de taille constante, au fil des générations.
On note que $p(0, 0) = 1$ et $p(N, N) = 1$: 0 et N sont absorbants, au contraire de $1, 2, \dots, N-1$, qui communiquent entre eux et vérifient $p(1, 0) > 0$ et $p(N-1, N) > 0$. Les classes sont $\{0\}$, $\{N\}$ et $\{1, 2, \dots, N-1\}$. Les deux premières sont fermées (et finies) donc récurrentes, la dernière est ouverte donc transiente. Il n'y a qu'un nombre fini d'états transients, donc pour n assez grand X_n est dans l'une des classes récurrentes : $X_n = 0$ ou $X_n = N$ (pour le modèle biologique, ceci revient à l'extinction d'un allèle du gène étudié). On note τ le temps d'atteinte de 0 ou N.
Pour déterminer $\mathbb{P}(X_\tau = N)$, on note que X_n est elle-même une martingale. En effet, comme l'espérance de la loi $\mathcal{B}(n, p)$ est np, on a

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

Ainsi, $\mathbb{E}[X_n] = \mathbb{E}[X_0] = x_0$. Or $X_n \underset{n}{\longrightarrow} X_\tau$ et $(X_n)_n$ est bornée donc par théorème de convergence dominée

$$x_0 = \mathbb{E}[X_n] = \mathbb{E}[X_\tau] = 0\cdot\mathbb{P}(X_\tau = 0) + N\cdot\mathbb{P}(X_\tau = N)$$

et finalement

$$\mathbb{P}(X_\tau = N) = \frac{x_0}{N}.$$

3.4 Inégalité d'Azuma

Si $(M_n)_n$ est une martingale telle que $|M_{n+1} - M_n| \leq 1$ pour tout n, alors, pour tout $m \geq 1$,

$$\mathbb{P}\big(\big|M_m - \mathbb{E}[M_m]\big| \geq \lambda\big) \leq 2\exp\left(-\frac{\lambda^2}{2m}\right).$$

(voir wikipedia)

3.5 Urne de Pólya

On considère une urne contenant des boules rouges et bleues, où on effectue des tirages selon la règle suivante : après chaque tirage, on remet dans l'urne la boule tirée ainsi qu'une nouvelle boule de la même couleur (cette couleur devient donc plus probable pour le prochain tirage). Alors on vérifie que la proportion de boules de couleur rouge est une martingale. Elle est bornée (car dans $[0,1]$), donc converge p.s.. On peut montrer que sa limite est aléatoire, de loi Beta de paramètre (a,b) où a,b sont les nombres initiaux de boules rouges et bleues.

3.6 Lien avec les chaînes de Markov

Certains critères de récurrence ou transience (critères de Foster) utilisent des martingales.
• Par exemple, si $(X_n)_n$ est une chaîne de Markov irréductible d'espace d'états E infini et s'il existe une fonction $f : E \to \mathbb{R}_+$ telle que f diverge à l'infini (c'est-à-dire que, pour tout $A > 0$, il existe $F \subset E$ fini tel que $f(x) \geq A$ pour tout $x \in E \setminus F$) et $Pf(x) \leq f(x)$ pour tout x sauf un nombre fini (notons $G = \{x | Pf(x) > f(x)\}$, fini), où

$$Pf(x) = \sum_{y \in E} P(x,y)f(y) = \mathbb{E}_x[f(X_1)],$$

alors $(X_n)_n$ est récurrente.
En effet, la condition $Pf(x) \leq f(x)$ s'écrit aussi $\mathbb{E}_x[f(X_1)] \leq f(x)$, d'où on déduit $\mathbb{E}[f(X_{n+1})|X_n] \leq f(X_n)$ si $X_n \notin G$, et donc

$$\mathbb{E}\big[f(X_{n+1})\big|X_0, \ldots, X_n\big] \leq f(X_n) \quad \text{si } X_n \notin G,$$

ce qui implique que $(f(X_{n\wedge\tau}))_{n\geq 0}$, où $\tau = \inf\{n \geq 0 | X_n \in G\}$, est une surmartingale positive sous $\mathbb{P}_x$ pour tout état initial $x \in E$. Donc elle converge $\mathbb{P}_x$-p.s.. Vu que f diverge à l'infini, $(X_{n\wedge\tau})_n$ reste donc presque sûrement dans un ensemble fini. Comme $(X_n)_n$ ne peut pas rester indéfiniment dans un ensemble fini (car elle est irréductible et E est infini), on en déduit que $\tau < \infty$ $\mathbb{P}_x$-p.s. Ainsi, G est atteint p.s. depuis tout point de départ x. On en déduit facilement la récurrence de la chaîne de Markov ($(X_n)_n$ visite infiniment souvent G par ce qui précède et par la propriété de Markov forte, donc un sommet de G est visité infiniment souvent, d'où la récurrence).
• Et s'il existe même $\varepsilon > 0$ tel que $Pf(x) \leq f(x) - \varepsilon$ pour tout $x \notin G$ (avec G fini), alors on peut même conclure que $(X_n)_n$ est récurrente positive.
En effet, on constate que $\mathbb{E}_x[f(X_1)] \leq f(x) - \varepsilon$ si $x \notin G$, ce qui donne en général

$$\mathbb{E}[f(X_{n+1})|X_0, \ldots, X_n] \leq f(X_n) - \varepsilon$$

tant que $X_n \notin G$, ce qui peut aussi se réécrire

$$\mathbb{E}[f(X_{n+1}) + (n+1)\varepsilon | X_0, \ldots, X_n] \leq f(X_n) + n\varepsilon,$$

et montre que la suite $(f(X_{n\wedge\tau}) + (n \wedge \tau)\varepsilon)_{n\geq 0}$ est une surmartingale. Par suite,

$$f(x) \geq \mathbb{E}_x[f(X_{n\wedge\tau}) + (n\wedge\tau)\varepsilon] = \mathbb{E}_x[f(X_{n\wedge\tau})] + \mathbb{E}_x[n\wedge\tau]\varepsilon \geq \mathbb{E}_x[n\wedge\tau]\varepsilon$$

et par théorème de convergence monotone $\mathbb{E}_x[n\wedge\tau] \uparrow_n \mathbb{E}_x[\tau]$ donc $\mathbb{E}_x[\tau] \leq \varepsilon^{-1}f(x) < \infty$. De là, on déduit la récurrence positive (on a montré que le temps de retour dans G est intégrable, donc par un théorème ergodique, $(X_n)_n$ passe une proportion strictement positive de son temps dans G, et donc dans au moins un état de G, qui est par conséquent récurrent positif...).
• Exemple d'application : file d'attente $X_{n+1} = X_n - b(X_n) + \xi_{n+1}$ (espace d'états $\mathbb{N}$), où $b : \mathbb{N} \to \mathbb{N}$ vérifie $0 \leq b(x) \leq x$ (la quantité $b(x)$ est le nombre de clients servis s'il y en a x dans la queue), et $(\xi_n)_n$ sont des variables aléatoires positives indépendantes et de même loi (nombre de nouveaux clients). Dès que $b(x) \geq \mathbb{E}[\xi_1]$ pour tous les x sauf un nombre fini, la fonction $f(x) = x$ convient donc la chaîne est récurrente. On peut par exemple prendre $b(x) = 0$ si $x \leq \mathbb{E}[\xi_1]$ (on ne sert personne s'il y a peu de monde) et $b(x) = m > \mathbb{E}[\xi_1]$ (où $m \in \mathbb{N}$) dès que $x > \mathbb{E}[\xi_1]$. Il y a récurrence positive si $b(x) \geq \mathbb{E}[\xi_1] + \varepsilon$ pour x assez grand.

3.7 Processus de branchement de Galton-Watson

On vérifie facilement que $M_n = \frac{Z_n}{m^n}$ est une martingale, où m est le nombre moyen d'enfants. Si la limite M_∞ est strictement positive, alors $Z_n \sim M_\infty m^n$ diverge exponentiellement vite. Il resterait à étudier $\mathbb{P}(M_\infty = 0)$.