MACS203 : Martingales

1 Préliminaires de la théorie des probabilités

Variables aléatoires

Def. Soit $\mathbf{T}$ un ensemble et $\{X_\tau, \tau \in \mathbf{T}\}$ une famille quelconque de v.a. La σ-algèbre $\mathcal{X}$ engendrée par cette famille est la plus petite σ-algèbre sur Ω telle que X_τ est $\mathcal{X}$-mesurable pour tout $\tau \in \mathbf{T}$, i.e.

$$\mathcal{X} = \sigma(X_\tau, \tau \in \mathbf{T}) = \sigma(\{X_\tau^{-1}(A) \mid \tau \in \mathbf{T}, A \in \mathcal{B}(\mathbf{R})\}) .$$

Lem. *Soit X et Y deux v.a. sur $(\Omega, \mathcal{A}, \mathbf{P})$ à valeurs respectivement dans $\mathbf{R}$ et $\mathbf{R}^n$. Alors X est $\sigma(Y)$-mesurable ssi $\exists f\colon \mathbf{R}^n \to \mathbf{R}, X = f(Y)$.*

Espérance de variables aléatoires

Th (Inégalité de Jensen). *Soit $X \in \mathcal{L}^1(\mathcal{A}, \mathbf{P})$ et $g\colon \mathbf{R}^d \to \bar{\mathbf{R}}$ une fonction convexe telle que $\mathbf{E}(|g(X)|) < \infty$. Alors $\mathbf{E}(g(X)) \geqslant g(\mathbf{E}(X))$.*

Def. Soit X une v.a. à valeurs dans $\mathbf{R}^d$. Sa fonction caractéristique est $\Phi_X : \begin{array}{ccc} \mathbf{R}^d & \to & \mathbf{C} \\ u & \mapsto & \mathbf{E}\left[e^{i\langle u|X\rangle}\right] \end{array}$.

Lem. *$\Phi_X(0) = 1$ et Φ_X est continue bornée (par 1) sur $\mathbf{R}^d$.*

Prop. *Soit $X \sim \mathcal{N}(b, V)$. On a $\Phi_X(u) = e^{\langle u|b\rangle - \frac{1}{2}\langle u|Vu\rangle}$.*

Prop. *Soit X réelle avec $\mathbf{E}(|X|^p) < \infty$ pour un certain $p \in \mathbf{N}^*$. Alors Φ_X est p fois dérivable et $\forall k \in [\![1\,;p]\!], \Phi_X^{(k)}(0) = i^k\mathbf{E}(X^k)$.*

Espaces $\mathcal{L}^p$ et convergences fonctionnelles des v.a.

La corrélation entre deux v.a. X et Y est $\mathrm{Cor}(X,Y) = \frac{\mathrm{Cov}(X,Y)}{\|X\|_2\|Y\|_2}$.
Le théorème de Pythagore s'écrit

$$\mathbf{E}(XY) = 0 \implies \mathbf{E}[(X+Y)^2] = \mathbf{E}\left[X^2\right] + \mathbf{E}\left[Y^2\right] \qquad \text{ou} \qquad \mathrm{Cov}(X,Y) = 0 \implies \mathrm{Var}(X+Y) = \mathrm{Var}(X) + \mathrm{Var}(Y) .$$

et la loi du parallélogramme s'écrit $\|X+Y\|_2^2 + \|X-Y\|_2^2 = 2\,\|X\|_2^2 + 2\,\|Y\|_2^2$.

Lem. *La convergence p.s. ou la convergence en norme dans L^p impliquent la convergence en probabilité.*

Lem. *La convergence en probabilité est équivalente à la convergence au sens de la distance $D\colon (X,Y) \mapsto \mathbf{E}(|X-Y| \wedge 1)$*

Th. *(L^0, D) est un espace métrique complet.*

Def. Une famille C de v.a. est dite **uniformément intégrable** (U.I.) si $\lim_{c\to\infty} \sup_{X\in C} \mathbf{E}\left[|X|\,\mathbf{1}_{|X|\geqslant c}\right] = 0$.

Th. *Soit $(X_n)_n$ et X des v.a. dans $\mathcal{L}^1$. Alors $X_n \longrightarrow X$ dans L^1 si et seulement si $X_n \longrightarrow X$ en probabilité et $(X_n)_n$ est U.I.*

Convergence en loi

2 Vecteurs gaussiens

Def. X est un **vecteur gaussien** (ou variable gaussienne multivariée ou variable normale multivariée) si et seulement si $\forall a \in \mathbf{R}^d$, la loi de $\langle a \mid X\rangle$ est une loi gaussienne (éventuellement de variance nulle).

Th. *X est un vecteur gaussien d'espérance m et de matrice de covariance Γ si et seulement si sa fonction caractéristique est $t \mapsto \exp\left(i\,\langle t \mid m\rangle - \frac{1}{2}t^\intercal\Gamma t\right)$. On écrit $X \sim \mathcal{N}_d(m, \Gamma)$.*

Prop. *Soit (X,Y) un vecteur gaussien à valeurs dans $\mathbf{R}^n \times \mathbf{R}^m$, de moyenne $\mu = \begin{pmatrix}\mu_X \\ \mu_Y\end{pmatrix}$ et de matrice de variances-covariances $V = \begin{pmatrix} V_X & V_{XY}^\intercal \\ V_{XY} & V_Y \end{pmatrix}$. Supposons que $\mathrm{Var}(Y) = V_Y$ est inversible. Alors la loi conditionnelle de X sachant $Y = y$ est gaussienne de moyenne $\mathbf{E}(X \mid Y = y) = \mu_X + V_{XY}V_Y^{-1}(y-\mu_Y)$ et variance $\mathrm{Var}(X \mid Y = y) = V_X - V_{XY}V_Y^{-1}V_{XY}^\intercal$.*

3 Processus aléatoires et structure d'information

Def. Un **processus** est une suite $(X_n)_n$ de v.a. sur $(\Omega, \mathcal{A})$ à valeurs dans un ensemble mesuré $(E, \mathcal{E})$.

Def. Une **filtration** de $\mathcal{A}$ est une suite croissante $\mathbf{F} = (\mathcal{F}_n)_{n\geqslant 0}$ de sous-σ-algèbres de $\mathcal{A}$. On dit que $(\Omega, \mathcal{A}, \mathbf{F})$ est un espace probabilisable filtré et $(\Omega, \mathcal{A}, \mathbf{F}, \mathbf{P})$ un espace probabilisé filtré.

Ex. La suite $(\mathcal{F}_n^X)_{n\in\mathbf{N}} = (\sigma(X_i, i \leqslant n))_{n\in\mathbf{N}}$ est une filtration de $\mathcal{A}$ appelée **filtration naturelle** de X.

Def. Soit $X = (X_n)_n$ un processus aléatoire et $(\mathcal{F}_n)_n$ une filtration de $\mathcal{A}$. On dit que X est :

- **F-adapté** si $\forall n \in \mathbf{N}$, X_n est $\mathcal{F}_n$-mesurable,
- **F-prévisible** si $\forall n \in \mathbf{N}$, X_n est $\mathcal{F}_{n-1}$-mesurable, où $\mathcal{F}_{-1} := \{\varnothing, \Omega\}$.

Def. Un **temps d'arrêt** ν est une variable aléatoires à valeurs dans $\mathbf{N} \cup \{\infty\}$ telle que $\forall n \in \mathbf{N}, \{\nu = n\} \in \mathcal{F}_n$. On note $\mathcal{T}$ l'ensemble des temps d'arrêt.

Prop. *Soit $(X_n)_{n\in\mathbf{N}}$ un processus **F**-adapté à valeurs dans $(E, \mathcal{E})$. Pour tout $A \in \mathcal{E}$, on définit le **premier temps d'atteinte** $T_A := \inf\{n \in \mathbf{N} \mid X_n \in A\}$, avec la convention $\inf \varnothing = \infty$. Alors T_A est un temps d'arrêt.*

Prop. *Soit $\tau, \theta, (\tau_n)_{n\in\mathbf{N}}$ des temps d'arrêt.*

(i) $\tau \wedge \theta, \tau \vee \theta$ *et* $\tau + \theta$ *sont des temps d'arrêt,*

(ii) *soit* $c \geqslant 0$ *une constante, alors* $\tau + c$ *et* $(1 + c)\tau$ *sont des temps d'arrêt,*

(iii) $\liminf_n \tau_n$ *et* $\limsup_n \tau_n$ *sont des temps d'arrêt.*

Prop. *Soit $(X_n)_n$ un processus aléatoire à valeurs dans un espace mesuré $(E, \mathcal{E})$ et τ un temps d'arrêt. Alors $X_\tau : \omega \in \Omega \mapsto X_{\tau(\omega)}(\omega)$ est une v.a.*

Prop. *Pour tout temps d'arrêt $\tau \in \mathcal{T}$, $\mathcal{F}_\tau \subset \mathcal{A}$ est une sous-σ-algèbre de $\mathcal{A}$. Si X est un processus aléatoire **F**-adapté, X_τ est $\mathcal{F}_\tau$-mesurable.*

Def. L'**information disponible à un temps d'arrêt** est $\mathcal{F}_\tau := \{A \in \mathcal{A} \mid \forall n \in \mathbf{N}, A \cap \{\tau = n\} \in \mathcal{F}_n\}$.

Prop. *Pour tout temps d'arrêt $\tau \in \mathcal{T}$, $\mathcal{F}_\tau$ est une sous-σ-algèbre de $\mathcal{A}$. Si X est un processus aléatoire **F**-adapté, X_τ est $\mathcal{F}_\tau$-mesurable.*

Prop. *Soit τ et θ deux temps d'arrêt. Alors $\{\tau \leqslant \theta\}$, $\{\tau \geqslant \theta\}$ et $\{\tau = \theta\}$ appartiennent à $\mathcal{F}_\tau \cap \mathcal{F}_\theta$, et pour toute v.a. X intégrable, on a $\mathbf{E}(\mathbf{E}(X \mid \mathcal{F}_\tau) \mid \mathcal{F}_\theta) = \mathbf{E}(\mathbf{E}(X \mid \mathcal{F}_\theta) \mid \mathcal{F}_\tau) = \mathbf{E}(X \mid \mathcal{F}_{\tau\wedge\theta})$.*

4 Chaînes de Markov

Soit $X = (X_n)_{n\in\mathbf{N}}$ un processus stochastique défini sur un espace probabilisé $(\Omega, \mathcal{A}, \mathbf{P})$ et à valeurs dans un espace d'états discret E, fini ou dénombrable.

Not. π_n est la distribution marginale de X_n : $\forall x \in E, \pi_n(x) := \mathbf{P}(X_n = x)$.

Def. On dit que X est un **chaîne de Markov** si $\forall A \subset E, \forall n \in \mathbf{N}^*, \mathbf{P}[X_n \in A \mid \mathcal{F}^X_{n-1}] = \mathbf{P}[X_n \in A \mid X_{n-1}]$.

Les **probabilités de transition** sont représentées par les **matrices de transition** P_n définies par $\forall n \in \mathbf{N}, \forall x, y \in E, P_n(x, y) := \mathbf{P}[X_n = y \mid X_{n-1} = x]$. Ce sont des matrices stochastiques : leurs composantes sont positives et leurs lignes somment à l'unité.

Prop. *Soit $(P_n)_{n\in\mathbf{N}^*}$ une suite de matrices stochastiques sur E. Pour toute distribution initiale π_0 il existe une chaîne de Markov de loi initiale π_0 et de matrices de transition $(P_n)_{n\in\mathbf{N}^*}$.*

Les probabilités marginales π_n se déduisent par $\forall n \in \mathbf{N}^*, \pi_n = \pi_0 P_1 \dots P_n$, où π_0 est un vecteur ligne de taille $\mathrm{Card}(E)$.

Prop (Formule de Chapman-Kolmogorov). $\forall x, y \in E, \forall k \in [0\,;n], \mathbf{P}(X_n = y \mid X_0 = x) = \sum_{z\in E} \mathbf{P}(X_n = y \mid X_k = z)\mathbf{P}(X_k = z \mid X_0 = x)$.

Not. $\mathbf{P}_x$ est la probabilité conditionelle sachant $X_0 = x$, et $\mathbf{E}_x$ est l'espérance associée.

Th (Propriété de Markov forte). *Soit $(X_n)_{n\in\mathbf{N}}$ une chaîne de Markov et τ un temps d'arrêt à valeurs dans $\mathbf{N}$. Alors $\forall A \subset E, \forall n \in \mathbf{N}^*, \mathbf{P}(X_{\tau+n} \in A \mid \mathcal{F}^X_{\tau+n-1}) = \mathbf{P}(X_{\tau+n} \in A \mid X_{\tau+n-1})$.*

Def. Une chaîne de Markov est dite **homogène** si sa matrice de transition P_n est indépendante de n.

5 Lois invariantes et classification des états

Soit X une chaîne de Markov homogène de matrice de transition P.

Def. Une probabilité ν sur E est représentée par un vecteur ligne $(\nu(x))_{x\in E}$. On dit que ν est une probabilité invariante pour X si $\nu P = \nu$.

Th. *Soit E un espace d'état fini. Alors il existe au moins une probabilité invariante.*

Si $\forall x \in E, \forall n \in \mathbf{N} \pi_n(x) > 0$ on définit $Q_n(x, y) := \mathbf{P}(X_n = y \mid X_{n+1} = x) = \frac{P(y,x)\pi_n(y)}{\pi_{n+1}(x)}$.

Def. On dit que X (homogène) est **réversible** par rapport à une mesure de probabilité ν si $\forall x, y \in E, \nu(x)P(x, y) = \nu(y)P(y, x)$, i.e. si les lois marginales π_n sont données par ν, $Q_n = P$ pour tout n.

Prop. *Soit ν une mesure de probabilité par rapport à laquelle X est invariant. Alors ν est une probabilité invariante.*

Def. Soit $x \in E$. On définit le temps d'arrêt de premier retour à x : $R_x := R^x_1 = \inf\{n \in \mathbf{N}^* \mid X_n = x\}$. x est dit **récurrent** si $\mathbf{P}(R_x < \infty) = 1$, dont **récurrent positif** si $\mathbf{E}_x(R_x) < \infty$ et **récurrent nul** si $\mathbf{E}_x(R_x) = \infty$. Sinon on dit que x est **transitoire** ou **transient**.

On introduit les mesures à valeurs dans $[0\,;\infty]$ définies par $\forall x, y \in E, \mu_x(y) = \mathbf{E}_x\left[\sum_{n=0}^{R_x-1} \mathbf{1}_{\{X_n=y\}}\right] = \sum_{n\in\mathbf{N}} \mathbf{P}_x(R_x > n, X_n = y)$.

Prop. *Soit $x \in E$. Alors,*

(i) $\mu_x P = \mu_x$ *ssi* x *est un état récurrent,*

(ii) μ_x *est une mesure finie ssi* x *est récurrent positif, dans ce cas* $\nu_x = \frac{\mu_x}{\mathbf{E}_x(R_x)}$ *est une probabilité invariante.*

Def. Soit $x, y \in E$. On dit que :

- x **communique** avec y, noté $x \leftarrow y$ si $\exists n \in \mathbf{N}, x_1, \ldots, x_n, P(x, x_1) \cdots P(x_n, y) > 0$,
- x et y communiquent, noté $x \leftrightarrow y$, si $x \leftarrow y$ et $y \leftarrow x$.

Def. Une classe $E_0 \subset E$ est dite **irréductible** si $\forall x, y \in E_0, x \leftarrow y$. X est dite irréductible si E est irréductible. Une classe $E_0 \subset E$ est dite **fermée** si $\forall x, y \in E, (x \in E_0 \wedge x \leftarrow y) \implies y \in E_0$. Si $\{x_0\}$ est fermée, on dit que x_0 est absorbant.

On introduit le **nombre de visite d'un état** x : $N^x := \sum_{n \in \mathbf{N}} \mathbf{1}_{\{X_n = x\}}$.

Prop. *Soit* $x, y \in E$.

(i) Si $x \leftarrow y$ *et* x *est récurrent, alors* y *est récurrent et* $N^y = \infty$, $\mathbf{P}_x$*-p.s.*

(ii) Si $x \leftrightarrow y$*, alors* x *et* y *sont simultanément soit transitoires soit récurrents.*

Th. *Supposons* X *irréductible. Alors* X *est récurrente positive si et seulement si* X *admet une loi invariante* ν*. De plus,* ν *est unique, strictement positive, donnée par* $\forall x \in E, \nu(x) = \frac{1}{\mathbf{E}_x(R_x)}$.

Prop. *Soit* X *une chaîne de Markov sur un espace d'état dénombrable* E*, et* $x \in E$ *récurrent. Alors, pour toute mesure* ν *sur* E*,* $\nu \geqslant \nu P \implies \nu = \nu(x)\mu_x$.

6 Théorèmes ergodiques

Théorèmes ergodiques

Th. *Soit* X *une chaîne de Markov irréductible,* $\forall x, y \in E, \frac{1}{n} N_n^y := \frac{1}{n}\sum_{i=0}^n \mathbf{1}_{\{X_i = y\}} \longrightarrow \frac{1}{\mathbf{E}_y(R^y)}$, $\mathbf{P}_x$*-p.s.*

En particulier il vient $\forall x, y \in E, \frac{1}{n}\sum_{i=0}^n \mathbf{P}(X_i = x) = \frac{1}{n}\sum_{i=0}^n \pi_i(x) \longrightarrow \nu(x)$, $\mathbf{P}_y$-p.s. avec ν une loi invariante.

Th. *Soit* X *une chaîne de Markov irréductible et récurrente positive sur* E *dénombrable, de matrice de transition* P *et d'unique loi invariante* ν*. Alors, pour toute fonction* $g \colon E \times E \to \mathbf{R}$ *positive ou telle que* $\mathbf{E}_\nu\left[|g(X_0, X_1)|\right] < \infty$*, on a* $\forall \pi_0, \frac{1}{n}\sum_{i=1}^n g(X_{i-1}, X_i) \xrightarrow{p.s.} \mathbf{E}_\nu\left[g(X_0, X_1)\right] = \sum_{x \in E} \nu(x) \sum_{y \in E} P(x, y) g(x, y)$.

Th. *Soit* X *et* g *comme précedemment. Supposons qu'il existe* $x \in E$ *tel que*

$$s(x)^2 := \mathbf{E}_x\left[\sum_{i=1}^{R_x} \left(g(X_{i-1}, X_i) - \mathbf{E}_\nu(g(X_0, X_1))\right)^2\right] < \infty\,.$$

Alors $\sigma^2 := \nu(x) s(x)^2$ *est une constante (indépendante de* x*) et*

$$\sqrt{x}\left(\frac{1}{n}\sum_{i=1}^n g(X_{i-1}, X_i) - \mathbf{E}_\nu(g(X_0, X_1))\right) \longrightarrow \mathcal{N}(0, \sigma^2) \text{ en loi.}$$

Convergence des lois marginales et apériodicité

Not. Pour tout état $x \in E$ on définit $I(x) := \{n \in \mathbf{N}^* \mid P^n(x, x) > 0\}$ et $\mathbf{p}(x) := \operatorname{pgcd}(I(x))$.

Prop. *Soit* X *une chaîne de Markov irréductible. Alors la fonction* $\mathbf{p}(x) = \mathbf{p}_X$ *est constante.*

Def. Soit X une chaîne de Markov irréductible. On dit que X est **apériodique** si $\mathbf{p}_X = 1$.

Lem. *Pour* $x \in E$, $\mathbf{p}(x) = 1 \iff \exists \mathbf{n}(x) \in \mathbf{N}, \forall n \geqslant \mathbf{n}(x), P^n(x, x) > 0$.

Th. *Soit* X *une chaîne de Markov irréductible, apériodique et récurrente positive d'unique loi invariante* ν*. Alors* $\forall x \in E, \pi_n(x) \longrightarrow \nu(x)$.

Prop. *Soit* X^1 *et* X^2 *deux chaînes de Markov indépendantes de même matrice de transition* P *irréductible apériodique. Alors la chaîne produit* $Y := (X^1, X^2)$ *est irréductible apériodique. Si de plus* P *est récurrente positive, il en est de même pour* Y.

Prop. *Soit* X *une chaîne de Markov irréductible apériodique sur* E *fini. Alors sa matrice de transition* P *vérifie la* ***condition de Dobelin*** *: il existe* $k \in \mathbf{N}, \epsilon > 0$ *et une loi* δ *sur* E *tels que* $\forall x, y \in E, P^k(x, y) \geqslant \epsilon \cdot \delta(y)$.

Th. *Soit* P *une matrice de transition vérifiant la condition de Dobelin. Alors il existe une unique loi invariante* $\nu \geqslant \epsilon \cdot \delta$ *vérifiant*

$$\sup_{x \in E} \sum_{y \in E} |P^n(x, y) - \nu(y)| \leqslant 2(1 - \epsilon)^{\lfloor n/k \rfloor}\,.$$

7 Martingales en temps discret

Martingales et temps d'arrêt

Def. Soit $(X_n)_{n\geqslant 0}$ un processus aléatoire adapté sur l'espace probabilisé filtré $(\Omega, \mathcal{A}, \mathbf{F}, \mathbf{P})$. On dit que X est une **surmartingale** (resp. **sous-martingale**) si X_n est $\mathbf{P}$-intégrable pour tout n et $\forall n \in \mathbf{N}^*, \mathbf{E}[X_n \mid \mathcal{F}_{n-1}] \leqslant$ (resp. $\geqslant$)X_{n-1}. X est une **martingale** s'il est à la fois surmartingale et sous-martingale.

Def. Pour un processus aléatoire $X = (X_n)_{n\geqslant 0}$, on définit le **processus arrêté** au temps d'arrrêt ν par $\forall n \in \mathbf{N}, X_n^\nu := X_{n\wedge\nu}$.

Lem. *Soit X une surmartingale (resp. sous-martingale, martingale) et ν un temps d'arrêt. Alors le processus arrêté X^ν est une surmartingale (resp. sous-martingale, martingale).*

Th. *Soit X une martingale (resp. surmartingale) et $\underline{\nu}, \overline{\nu}$ deux temps d'arrêt bornés dans $\mathcal{T}$ vérifiant $\underline{\nu} \leqslant \overline{\nu}$ p.s. Alors $\mathbf{E}[X_{\overline{\nu}} \mid \mathcal{F}_{\underline{\nu}}] =$ (resp. $\leqslant$)$X_{\underline{\nu}}$.*

Prop. *Soit $X = (X_n)_n$ un processus aléatoire $\mathbf{F}$-adapté, $\forall n \in \mathbf{N}, \mathbf{E}(|X_n|) < \infty$. Alors X est une martingale ssi $\mathbf{E}[X_\nu] = \mathbf{E}[X_0]$ pour tout temps d'arrêt n borné.*

Martingales fermées

Def. Une martingale $(X_n)_n$ est **fermée** s'il existe une v.a. réelle intégrable Y telle que $\forall n \in \mathbf{N}, X_n = \mathbf{E}(Y \mid X_n)$.

Th. *Toute martingale fermée est uniformément intégrable.*

Inégalités et décomposition

Th (Inégalité maximale de Doob). *Soit $M = (M_n)_{n\geqslant 0}$ une sous-martingale, et $M_n^* := \sup_{k\leqslant n} M_k$ son processus de maximum courant.*

(i) $\forall c > 0, \forall n \in \mathbf{N}, c\mathbf{P}(M_n^* \geqslant c) \leqslant \mathbf{E}(M_n \mathbf{1}_{M_n^*\geqslant c})$

(ii) *Soit $p > 1$ et supposons que la sous-martingale M est positive et $\forall n \in \mathbf{N}, M_n \in \mathcal{L}^p$. Alors $M_n^* \in \mathcal{L}^p$ et $\|M_n^*\|_p \leqslant \frac{p}{p-1}\|M_n\|_p$.*

Prop (Décomposition de Doob). *Soit $(X_n)_n$ un processus aléatoire intégrable. Il existe une martingale $(M_n)_n$ et un processus $\mathbf{F}$-prévisible $(V_n)_n$ tels que $M_0 = V_0 = 0$ et $\forall n \geqslant 0, X_n = X_0 + M_n + V_n$. Cette décomposition est unique.*

Rem. On voit que : X est une surmartingale ssi V est décroissant, X est une sous-martingale ssi V est croissant et X est une martingale ssi $V = 0$.

Prop. *Soit $X = (X_n)_n$ une martingale de carré intégrable, et $\Delta X_n := X_n - X_{n-1}$. Alors $X_n^2 = X_0^2 + N_n + [X]_n$ où $N_n := 2\sum_{i=1}^n X_{i-1}\Delta X_i$, $[X]_n := \sum_{i=1}^n (\Delta X_i)^2$ et $N_0 = [X]_0 = 0$. Dans cette décomposition, $(N_n)_n$ est une martingale nulle en zéro, et $([X]_n)_n$ est un processus $\mathbf{F}$-adapté croissant intégrable appelé **variation quadratique** de la martingale X.*

Def. Un processus $X = (X_n)_{n\in\mathbf{N}}$ est une **martingale locale** s'il existe une suite de temps d'arrêt $(\tau_n)_n$ telle que $\tau_n \longrightarrow \infty$ $\mathbf{P}$-p.s. et le processus arrêté X^{τ_n} est une martingale pour tout n.

Lem. *Soit $X = \{X_n, n \in [\![0\,;N]\!]\}$ une martingale locale telle que $\mathbf{E}[X_N^-] < \infty$. Alors X est une martingale.*

8 Convergence des martingales

Rem. La suite $(\mathbf{E}[M_n^2])_{n\in\mathbf{N}}$ est croissante.

Th. *Soit $(M_n)_{n\in\mathbf{N}}$ une martingale bornée dans L^2, i.e. $\sup_n \mathbf{E}[M_n^2] < \infty$. Alors il existe une v.a. $M_\infty \in L^2$ telle que $M_n \xrightarrow{L^2} M_\infty$ et $M_n \xrightarrow{p.s.} M_\infty$.*

Th. *Soit $(M_n)_{n\in\mathbf{N}}$ une martingale de carré intégrable telle que $\sum_{n\geqslant 1} \frac{1}{n^2}\mathbf{E}[|\Delta M_n|^2] < \infty$. Alors $\frac{1}{n}M_n \longrightarrow 0$ p.s. et dans L^2.*

Th (Loi forte des grands nombres). *Soit $(X_n)_{n\geqslant 0}$ une suite iid de v.a. intégrables. Alors $\frac{1}{n}\sum_{i=1}^n X_i \xrightarrow{p.s.} \mathbf{E}[X_1]$.*

Lem. *Soit $(X_n)_{n\in\mathbf{N}}$ une sous-martingale, et $a < b$. Alors la moyenne du nombre de traversées montantes de l'intervalle $[a\,;b]$ vérifie $\mathbf{E}[U_n^{a,b}] \leqslant \frac{1}{b-a}\mathbf{E}[(X_n - a)^+]$.*

Th. *Soit $(X_n)_n$ une sous-martingale. Si $\sup_{n\geqslant 0}\mathbf{E}[X_n^+] < \infty$, alors $\exists X \in L^1, X_n \xrightarrow{p.s.} X$.*

Th. *Soit $(X_n)_n$ une sous-martingale. Si $\sup_{n\geqslant 0}\mathbf{E}[X_n^+] < \infty$ alors il existe une v.a. $X \in L^1$ telle que $X_n \xrightarrow{p.s.} X$.*

Rem. Une sous-martingale est bornée dans L^1 si et seulement si $\sup_{n\geqslant 0}\mathbf{E}[X_n^+] < \infty$.

Th. *Pour une martingale $M = (M_n)_n$, les deux assertions suivantes sont équivalentes :*

(i) *M est fermée et, en prenant $Y \in L^1$ tel que $\forall n \in \mathbf{N}, M_n = \mathbf{E}[Y \mid \mathcal{F}_n]$, on a $M_n \xrightarrow{L^1} Y$ et $M_n \xrightarrow{p.s.} Y$,*

(ii) *M est uniformément intégrable.*

Th (central limite martingale). *Soit $(M_n)_n$ une martingale telle que $\mathbf{E}[(\Delta M_n)^2 \mid \mathcal{F}_n] = 1$ et $K := \sup_{n\geqslant 1}\mathbf{E}(|\Delta M_n|^3) < \infty$. Alors $\frac{1}{\sqrt{n}}M_n \xrightarrow{\mathcal{L}} \mathcal{N}(0,1)$.*