

ISFA 2
Université Claude Bernard Lyon 1
Année universitaire 2006-2007

Processus stochastiques

F. Bienvenüe-Duheille

Jusqu'à maintenant, les cours de probabilités que vous avez suivis se sont intéressés essentiellement à l'étude d'une variable aléatoire, notamment sa loi, son espérance, sa variance, ou plus généralement à l'étude d'un couple ou vecteur aléatoire. Vous avez également étudié le comportement de suites formées à partir de variables aléatoires indépendantes et identiquement distribuées avec en particulier les lois des grands nombres et le théorème central limite mais aussi tous les théorèmes statistiques.

Ce cours s'intéresse aux processus stochastiques, c'est-à-dire aux familles de variables aléatoires indexées par un paramètre discret ou continu, représentant le temps. La loi d'un processus $(X_t)_{t \in I}$ sera donnée par la loi conjointe de $(X_{t_1}, \dots, X_{t_n})$, pour tout choix de $n \in \mathbb{N}^*$ et de $t_i \in I$, et non simplement par la loi de chacune des variables aléatoires X_t .

Avant de vous pencher sur ce cours, prenez le temps de revisiter vos cours de calcul intégral et probabilité, en ouvrant un oeil éveillé sur les notions de mesurabilité, lois usuelles discrètes et à densités, lois conjointes, espérance conditionnelle, lois fortes de grands nombres, théorème central limite... sans oublier les différents modes de convergence (p.s., en loi, dans $\mathcal{L}^p$) et sans que cela soit limitatif!

Chapitre 1

Chaînes de Markov

1 Définitions, propriétés

1.1 Le modèle

On dispose :

- d'un **espace d'états**, c'est-à-dire d'un ensemble E fini ou dénombrable,
- d'une loi de probabilité μ_0 sur E qui jouera le rôle de **loi initiale**,
- des **probabilités de transition (ou de passage)** de x vers y , c'est-à-dire d'une famille $(p(x, y))_{(x, y) \in E^2}$ de nombres réels positifs vérifiant

$$\sum_{y \in E} p(x, y) = 1.$$

Remarque : Lorsque l'espace d'états E est fini, les $(p(x, y))$ peuvent être écrits sous la forme d'une matrice, appelée **matrice de transition**, et dans ce cas, les « sommes en ligne » de cette matrice doivent toutes être égales à 1. On dit que cette matrice est **stochastique**.

Définition 1.1 Une chaîne de Markov $(X_n)_{n \geq 0}$ sur E , de loi initiale μ_0 , de probabilités de transition $(p(x, y))_{x, y}$ est une suite de variables aléatoires à valeurs dans E telle que

1. pour tout $x \in E$, $\mathbf{P}(X_0 = x) = \mu_0(x)$,
2. pour tout $n \geq 1$ et pour tout $(n + 1)$ -uplet $(x_0, x_1, \dots, x_n) \in E^{n+1}$, on a

$$\mathbf{P}(X_0 = x_0, X_1 = x_1, \dots, X_n = x_n) = \mu_0(x_0)p(x_0, x_1)p(x_1, x_2) \dots p(x_{n-1}, x_n).$$

En particulier, on aura $p(x, y) = \mathbf{P}(X_1 = y | X_0 = x) = \mathbf{P}(X_{n+1} = y | X_n = x)$.

On notera par la suite $p^{(n)}(x, y) = \mathbf{P}(X_n = y | X_0 = x) = \mathbf{P}(X_{n+k} = y | X_k = x)$, c'est-à-dire la probabilité de passer de l'état x à l'état y en n étapes.

Exemples :

- Une suite de variables aléatoires indépendantes et identiquement distribuées constitue une chaîne de Markov (le vérifier!), mais ce n'est évidemment pas un cas générique.
- Une chaîne de Markov peut servir à modéliser la cotisation d'assurance automobile versée par un assuré : la cotisation versée pour l'année $n + 1$ dépend de sa cotisation pour l'année n , et du nombre d'accidents qu'il a occasionnés au cours de l'année n .
- On peut modéliser le niveau de stock de produits (avec par exemple un réapprovisionnement dès que le stock est inférieur à un seuil), l'état d'une machine (fonctionnement/panne/réparation), la santé d'un individu (sain/malade/mort, avec éventuellement différents niveaux de gravité de la maladie).

- On peut également utiliser une chaîne de Markov pour modéliser le temps qu'il fait. Les états seront alors par exemple : beau, nuageux, pluvieux. Dans le modèle le plus simple, le temps d'un jour donné ne dépend que du temps de la veille. On peut également envisager une modélisation faisant intervenir le temps de deux jours successifs (ou plus). Les états seront alors les neuf couples (beau, beau), (beau, nuageux), (beau, pluvieux), (nuageux, beau), ..., la première coordonnée désignant le temps de l'avant-veille et la deuxième le temps de la veille. Un certain nombre de transitions est alors impossible (par exemple de (beau, beau) vers (pluvieux, pluvieux)). Pourquoi ? Lister les différentes transitions possibles.

La proposition suivante permet de caractériser les chaînes de Markov :

Proposition 1.2 *Une suite de variables aléatoires $(X_n)_{n \geq 0}$ est une chaîne de Markov de probabilités de transition $(p(x, y))$ si et seulement si, pour tout $n \geq 0$ et tout $(n+1)$ -uplet $(x_0, x_1, \dots, x_n)$ de E^{n+1} , on a*

$$\mathbf{P}(X_{n+1} = x_{n+1} | X_n = x_n, \dots, X_1 = x_1, X_0 = x_0) = \mathbf{P}(X_{n+1} = x_{n+1} | X_n = x_n) = p(x_n, x_{n+1}). \quad (1.1)$$

Preuve : Supposons qu'une suite $(X_n)_{n \geq 0}$ de variables aléatoires vérifie l'égalité 1.1 ci-dessus, pour tout $(n+1)$ -uplet ; notons μ la loi de la variable aléatoire X_0 et montrons, par récurrence sur n , que $(X_n)_{n \geq 0}$ est une chaîne de Markov.

L'hypothèse de récurrence $\mathcal{H}_n$ est :

$$(\mathcal{H}_n) \quad \begin{cases} \text{Pour tout } (n+1)\text{-uplet } (x_0, \dots, x_n) \in E^{n+1}, \\ \mathbf{P}(X_0 = x_0, X_1 = x_1, \dots, X_n = x_n) = \mu(x_0)p(x_0, x_1) \dots p(x_{n-1}, x_n) \end{cases}$$

L'hypothèse $\mathcal{H}_0$ est bien vérifiée puisque, pour tous $(x_0, x_1) \in E^2$ on a

$$\mathbf{P}(X_0 = x_0, X_1 = x_1) = \mathbf{P}(X_1 = x_1 | X_0 = x_0)\mathbf{P}(X_0 = x_0) = p(x_0, x_1)\mu(x_0).$$

Supposons maintenant que $\mathcal{H}_n$ est vérifiée et prouvons que $\mathcal{H}_{n+1}$ l'est également. Fixons un $(n+2)$ -uplet $(x_0, \dots, x_{n+1}) \in E^{n+2}$. On a

$$\begin{aligned} \mathbf{P}(X_0 = x_0, X_1 = x_1, \dots, X_n = x_n, X_{n+1} = x_{n+1}) &= \mathbf{P}(X_{n+1} = x_{n+1} | X_0 = x_0, X_1 = x_1, \dots, X_n = x_n) \\ &\quad \times \mathbf{P}(X_0 = x_0, X_1 = x_1, \dots, X_n = x_n). \end{aligned}$$

L'égalité 1.1 et l'hypothèse de récurrence au rang n permettent maintenant d'écrire

$$\mathbf{P}(X_0 = x_0, X_1 = x_1, \dots, X_n = x_n, X_{n+1} = x_{n+1}) = \mu(x_0)p(x_0, x_1) \dots p(x_{n-1}, x_n)p(x_n, x_{n+1}).$$

La preuve de la réciproque est laissée en exercice.

Exercice : On se donne une chaîne de Markov $(X_n)_{n \geq 0}$ de mesure initiale μ_0 et de probabilités de transition $p(x, y)$. Écrivez la loi conjointe de (X_0, X_1) , puis la loi de X_1 , puis celle de X_2 .

Dans la suite, c'est les probabilités de transition qui seront importantes, beaucoup plus que la loi initiale. Le plus souvent, on parlera ainsi d'une chaîne de Markov de probabilités de transition données, sans préciser la loi initiale. Néanmoins, on utilisera les notations suivantes pour spécifier la loi initiale :

- Si $e \in E$ est un état, $\mathbf{P}_e$ désignera la mesure de probabilité se rapportant à la chaîne vérifiant $X_0 = e$ presque sûrement. On a ainsi $\mathbf{P}_e(X_n = x) = \mathbf{P}(X_n = x | X_0 = e)$, et en particulier $\mathbf{P}_e(X_0 = x) = 1$ si $x = e$ et 0 sinon et $\mathbf{P}_e(X_1 = x) = p(e, x)$.
- Si μ est une mesure sur E , $\mathbf{P}_\mu$ désignera la loi de la chaîne de mesure initiale μ . On aura donc $\mathbf{P}_\mu(X_0 = x) = \mu(x)$ et $\mathbf{P}_\mu(X_1 = x) = \sum_{y \in E} \mu(y)p(y, x)$.

De façon similaire, lorsque l'on calcule des espérances, on écrit $\mathbf{E}_e$ ou $\mathbf{E}_\mu$.

Le fait d'être une chaîne de Markov n'est pas stable par les opérations habituelles (somme, produit...). L'une des opérations que l'on peut faire pour conserver le caractère markovien est décrite dans la proposition suivante :

Proposition 1.3 *Soit $(X_n)_{n \geq 0}$ une chaîne de Markov dont les probabilités de transition sont données par la matrice (éventuellement infinie) $\Pi = (p(x, y))$. Alors la suite $(X_{2n})_{n \geq 0}$ est une chaîne de Markov de probabilités de transition, notées $\Pi^{(2)} = (p^{(2)}(x, y))$, données par*

$$p^{(2)}(x, y) = \sum_{z \in E} p(x, z)p(z, y).$$

1.2 Les contraintes du modèle

Étudions pour une chaîne de Markov $(X_n)_{n \geq 0}$ sur un espace d'état E le temps passé dans un état donné avant de le quitter. Supposons pour fixer les idées $X_0 = x$ p.s. et notons $T = \inf\{n \geq 1, X_n \neq x\}$. $T - 1$ est alors le temps passé en l'état x avant de le quitter. On a bien sûr $\mathbf{P}_x(T = 1) = \mathbf{P}_x(X_1 \neq x) = 1 - p(x, x)$ et plus généralement, pour $n \geq 2$:

$$\begin{aligned} \mathbf{P}_x(T = n) &= \mathbf{P}_x(X_1 = x, \dots, X_{n-1} = x, X_n \neq x) \\ &= \mathbf{P}_x(X_1 = x, \dots, X_{n-1} = x) - \mathbf{P}_x(X_1 = x, \dots, X_{n-1} = x, X_n = x) \\ &= (p(x, x))^{n-1} - (p(x, x))^n = (p(x, x))^{n-1}(1 - p(x, x)) \end{aligned}$$

La loi de T est donc une loi géométrique. C'est une loi sans mémoire, qui vérifie

$$\mathbf{P}(T \geq n + k | T \geq n) = \mathbf{P}(T \geq k).$$

Une possibilité pour éviter cela est de permettre que les probabilités de transitions dépendent du temps n : la loi de $(X_0, \dots, X_n)$ est alors donnée par

$$\mathbf{P}(X_0 = x_0, X_1 = x_1, \dots, X_n = x_n) = \mu_0(x_0)p(x_0, x_1, 1)p(x_1, x_2, 2) \dots p(x_{n-1}, x_n, n)$$

et on a $p(x, y, n) = \mathbf{P}(X_n = y | X_{n-1} = x)$. On parle alors de chaîne de Markov **inhomogène en temps**, par opposition aux chaînes de Markov décrites au paragraphe 1.1 qui sont dites homogènes en temps.

Le problème est alors que les résultats décrits dans la suite de ce chapitre ne s'appliquent plus. La seule méthode pour étudier de telles chaînes de Markov consiste à effectuer des simulations informatiques.

1.3 Irréductibilité, stationnarité

On se donne une chaîne de Markov $(X_n)_{n \geq 0}$ sur un espace d'états E fini ou dénombrable.

Définition 1.4 *On dira que deux états x et y **communiquent** s'il existe deux entiers $n, m > 0$ tels que $\mathbf{P}(X_n = y | X_0 = x) > 0$ et $\mathbf{P}(X_m = x | X_0 = y) > 0$. La chaîne sera dite **irréductible** si tous les états communiquent.*

L'irréductibilité traduit donc la possibilité de passer d'un état à un autre, même avec une très faible probabilité.

Pour vérifier qu'une chaîne de Markov est irréductible, il est souvent utile de dessiner un graphe orienté dont les sommets sont les états de la chaîne et où une arête représente une transition possible (l'arête (x, y) existe uniquement si $p(x, y) > 0$). La chaîne est alors irréductible si et seulement s'il

existe un chemin fermé passant au moins une fois par tous les états de la chaîne (on dit que le graphe est connexe). La notion d'irréductibilité dépend uniquement des probabilités de transition, et non de la loi initiale.

Définition 1.5 Une mesure de probabilité μ est dite **stationnaire** pour les probabilités de transition $(p(x, y))$ si, pour tout $x \in E$, $\mathbf{P}_\mu(X_1 = x) = \mu(x)$, c'est-à-dire que si μ est la loi de X_0 , alors c'est aussi celle de X_1 , et de X_n pour tout $n \geq 1$. De façon pratique, une mesure μ , donnée par les $\mu(x)$, $x \in E$, est stationnaire pour $(p(x, y))$ si elle vérifie

$$\sum_{x \in E} \mu(x) = 1$$

$$\forall x \in E, \sum_{y \in E} \mu(y)p(y, x) = \mu(x).$$

Attention : si les $p(x, y)$ sont disposés sous la forme d'une matrice notée Π (les x représentant les lignes et les y les colonnes), μ est solution d'un système donné par les **colonnes** de cette matrice. En fait, μ est très exactement un vecteur propre de la transposée de Π , associé à la valeur propre 1. Lorsque E est de cardinal fini, pourquoi 1 est-il toujours valeur propre de Π ?

Remarque : Si la loi de X_0 est la mesure stationnaire (ou une mesure stationnaire s'il en existe plusieurs), les (X_n) forment une suite de variables aléatoires identiquement distribuées (mais pas indépendantes en général!).

Proposition 1.6

1. Si l'espace d'états est fini, il existe toujours au moins une mesure stationnaire.
2. Si la chaîne est irréductible, il existe au plus une mesure stationnaire.

Corollaire 1.7 Si la chaîne est irréductible sur un espace d'états fini, il existe une unique mesure stationnaire.

Exercice : Vérifiez que la chaîne de Markov formée par une suite de variables aléatoires indépendantes et identiquement distribuées est irréductible, apériodique et que tous les états sont récurrents positifs. Quelle est la mesure stationnaire ?

1.4 Transience, récurrence, période

Une des questions importantes qui se posent face à une chaîne de Markov est de savoir si, partant d'un point, on y revient avec probabilité 1. Autrement dit, si on se promène en suivant une chaîne de Markov, est-on sûr de repasser par son point de départ ? Pour étudier ces questions, on fixe donc un état $x \in E$ et on travaille avec la mesure $\mathbf{P}_x$, c'est-à-dire que l'on a $X_0 = x$ p.s..

On notera τ_x l'**instant du premier retour** en x : $\tau_x = \inf\{n \geq 1, X_n = x\}$.

Définition 1.8

- L'état x est **transient** ou **transitoire** si $\mathbf{P}_x(\tau_x < \infty) < 1$.
- L'état x est **récurrent** si $\mathbf{P}_x(\tau_x < \infty) = 1$.
- L'état x est **récurrent positif** si $\mathbf{E}_x(\tau_x) < \infty$. Il est **récurrent nul** si $\mathbf{P}_x(\tau_x < \infty) = 1$ et $\mathbf{E}_x(\tau_x) = \infty$.
- La **période** de l'état x est le pgcd (plus grand commun diviseur) de toutes les longueurs de chemins reliant x à lui-même, et parcourus avec une probabilité strictement positive. Lorsque le pgcd obtenu est égale à 1, on dit que l'état x est **apériodique**.

- Lorsque tous les états d'une chaîne de Markov sont de la même nature, on dit que la chaîne est, selon le cas, transiente, récurrente positive ou récurrente nulle. On parle également de période de la chaîne si tous les états ont la même période.

Attention : La récurrence et la transience ne sont pas directement liées à l'irréductibilité de la chaîne de Markov : la chaîne peut tout à fait être irréductible sans qu'aucun de ses états ne soit récurrent.

Rappelons que l'on note $p^{(n)}(x, y) = \mathbf{P}_x(X_n = y) = \mathbf{P}(X_n = y | X_0 = x)$ et que $p^{(n)}(x, y)$ est le terme (x, y) de la matrice Π^n . La proposition suivante permet de caractériser la récurrence/transience d'un état :

Théorème 1.9 (Caractérisation de la récurrente/transience) .

1. x est récurrent $\iff \mathbf{P}_x(\exists \text{ une infinité d'indices } n \text{ tels que } X_n = x) = 1$.
2. x est transient $\iff \sum_{n \geq 1} p^{(n)}(x, x) < \infty$.
3. x est récurrent nul $\iff \sum_{n \geq 1} p^{(n)}(x, x) = \infty$ et $\lim_n p^{(n)}(x, x) = 0$.
4. Si x est récurrent positif et apériodique, $p^{(n)}(x, x)$ admet une limite strictement positive.
5. Si x est récurrent positif et périodique de période t , $p^{(nt)}(x, x)$ admet une limite strictement positive.

Preuve du point 1. Fixons un état x .

$$\begin{aligned}
\mathbf{P}_x(\text{On revient au moins 2 fois en } x) &= \mathbf{P}_x(\cup_n \{\tau_x = n, X_n = x, \exists m > n, X_m = x\}) \\
&= \sum_n \mathbf{P}_x(\tau_x = n, X_n = x, \exists m > n, X_m = x) \\
&= \sum_n \mathbf{P}_x(\exists m > n, X_m = x | \tau_x = n, X_n = x) \times \\
&\quad \mathbf{P}_x(\tau_x = n, X_n = x) \\
&= \sum_n \mathbf{P}_x(\exists m > n, X_m = x | X_n = x) \mathbf{P}_x(\tau_x = n) \\
&= \sum_n \mathbf{P}_x(\tau_x < \infty) \mathbf{P}_x(\tau_x = n) \\
&= (\mathbf{P}_x(\tau_x < \infty))^2
\end{aligned}$$

Où utilise-t-on le caractère markovien ?

Plus généralement, $\mathbf{P}_x(\text{On revient au moins } n \text{ fois en } x) = (\mathbf{P}_x(\tau_x < \infty))^n$.

Comme les événements $\{\text{On revient au moins } n \text{ fois en } x\}$ sont décroissants, on a

$$\mathbf{P}_x(\text{On revient une infinité de fois en } x) = \lim_n (\mathbf{P}_x(\tau_x < \infty))^n$$

Si x est récurrent, $\mathbf{P}_x(\tau_x < \infty) = 1$, on en déduit que la chaîne issue de x repasse presque sûrement en x une infinité de fois.

Si x est transient, on obtient que la probabilité que la chaîne repasse une infinité en x est nulle.

Preuve du point 2.

- Supposons que $\sum_n p^{(n)}(x, x) < \infty$. Le lemme de Borel Cantelli (partie « triviale ») implique que

$$\mathbf{P}_x \left(\bigcap_{N} \bigcup_{n \geq N} \{X_n = x\} \right) = 0.$$

L'événement précédent est l'ensemble :

$$\{\omega \in \Omega; \text{il existe une sous-suite infinie } n_k(\omega) \text{ telle que } X_{n_k}(\omega) = x\},$$

ou, autrement dit, c'est l'ensemble des ω pour lesquels la suite $(X_n(\omega))_n$ visite l'état x une infinité de fois.

Puisque cet événement est de probabilité nulle, le point 1 implique que x est transient.

• Supposons maintenant que x est transient et notons N_x le nombre de passages de $(X_n)_{n \geq 0}$ par x . En reprenant la preuve du point 1, on voit que $\mathbf{P}_x\{N_x \geq l\} = (\mathbf{P}_x(\tau_x < \infty))^l$, c'est-à-dire que N_x suit une loi géométrique sur $\mathbb{N}$ de paramètre $1 - \mathbf{P}_x(\tau_x < \infty)$. En particulier, $\mathbf{E}_x(N_x) < \infty$.

Or

$$\begin{aligned} \mathbf{E}_x(N_x) &= \mathbf{E}_x\left(\sum_{n \geq 1} \mathbf{1}_{X_n=x}\right) \\ &= \sum_{n \geq 1} \mathbf{E}_x(\mathbf{1}_{X_n=x}) \\ &= \sum_{n \geq 1} p^{(n)}(x, x). \end{aligned}$$

On a donc équivalence entre transience d'un état et convergence de la série $\sum_n p^{(n)}(x, x)$.

Proposition 1.10 *Si x et y sont deux états qui communiquent, ils sont de même nature (récurrents positifs, récurrents nuls ou transients) et sont de même période.*

Preuve : Soient x et y deux états qui communiquent. Choisissons deux entiers N et M tels que $\alpha = p^{(N)}(x, y) > 0$ et $\beta = p^{(M)}(y, x) > 0$.

Pour tout n positif, on a :

$$p^{(N+M+n)}(x, x) \geq p^{(N)}(x, y)p^{(n)}(y, y)p^{(M)}(y, x) = \alpha\beta p^{(n)}(y, y) \quad (1.2)$$

et

$$p^{(N+M+n)}(y, y) \geq p^{(M)}(y, x)p^{(n)}(x, x)p^{(N)}(x, y) = \alpha\beta p^{(n)}(x, x) \quad (1.3)$$

donc les séries $\sum_n p^{(n)}(x, x)$ et $\sum_n p^{(n)}(y, y)$ convergent ou divergent simultanément et la limite de $p^{(n)}(x, x)$ (ou de $p^{(nt)}(x, x)$) lorsque n tend vers $+\infty$ est non nulle si et seulement si la limite de $p^{(n)}(y, y)$ (ou de $p^{(nt)}(y, y)$) est non nulle : x est récurrent nul si et seulement si y l'est.

Passons maintenant à l'étude de la période.

Supposons que x soit de période t . Comme $p^{(N+M)}(x, x) \geq p^{(N)}(x, y)p^{(M)}(y, x) > 0$, $N + M$ est nécessairement divisible par t . On en déduit que si n n'est pas un multiple de t , $p^{(N+M+n)}(x, x) = 0$; puis en utilisant la relation 1.2, il vient que, toujours si n n'est pas un multiple de t , $p^{(n)}(y, y) = 0$.

Donc la période de y est un multiple de celle de x : nécessairement, les deux périodes sont égales.

On déduit immédiatement de la proposition précédente le corollaire suivant :

Corollaire 1.11

- Si la chaîne est irréductible, tous les états de la chaîne sont de même nature et ont la même période.
- Si la chaîne est irréductible, il existe une (nécessairement unique) mesure stationnaire μ si et seulement tous les états sont récurrents positifs et on a alors $\mathbf{E}_x(\tau_x) = 1/\mu(x)$, pour tout $x \in E$. Si la chaîne est de plus apériodique, on a $p^{(n)}(x, x) \rightarrow \mu(x)$, et si elle est de période t , on a $p^{(nt)}(x, x) \rightarrow t\mu(x)$
- En particulier si l'espace d'états est fini et si la chaîne est irréductible, tous les états sont positivement récurrents et de même période.

1.5 Convergence

Le principal résultat est le suivant :

Proposition 1.12 *Si (X_n) est une chaîne de Markov irréductible, apériodique et s'il existe une unique mesure stationnaire μ , alors, pour toute mesure initiale μ_0 , la suite $(X_n)_n$ converge en loi vers μ .*

Autrement dit, pour tout $x \in E$, $\mathbf{P}_{\mu_0}(X_n = x)$ tend vers $\mu(x)$ lorsque n tend vers $+\infty$.

On connaît de plus un majorant de la vitesse de convergence :

$$|\mathbf{P}_{\mu_0}(X_n = x) - \mu(x)| \leq c\lambda^n,$$

où λ est la deuxième plus grande valeur propre de la « matrice » de transition (la plus grande valeur propre est toujours 1 : pourquoi?).

La convergence que l'on observe est donc en loi, et il y a très peu de chance d'obtenir une convergence presque sûre : cela signifierait que la chaîne de Markov est constante à partir d'un certain rang.

La convergence est très rapide : λ est strictement inférieur à 1. On dit dans ce cas que la vitesse est exponentielle.

1.6 Fonction de coût

Les transitions d'une chaîne de Markov ont fréquemment un coût (ou un gain) associé : le coût de la transition de l'état x vers l'état y . On peut alors calculer le coût d'un chemin, comme étant la somme des coûts des différentes transitions.

On se donne donc une fonction $c : E \times E \rightarrow \mathbb{R}$, qui décrit le coût des transitions. Le coût du chemin $\gamma = (x_0 \rightarrow x_1 \rightarrow \dots \rightarrow x_n)$ est ainsi $c(\gamma) = c(x_0, x_1) + c(x_1, x_2) + \dots + c(x_{n-1}, x_n)$, et le coût moyen calculé le long de ce chemin est $c(\gamma)/n$.

Théorème 1.13 *Soit $(X_n)_{n \geq 0}$ une chaîne irréductible, de mesure initiale μ_0 et de mesure stationnaire μ . Soit $c : E \times E \rightarrow \mathbb{R}$ une fonction de coût telle que*

$$\mathbf{E}_{\mu}|c(X_0, X_1)| = \sum_{x, y \in E} \mu(x)p(x, y)|c(x, y)| < \infty.$$

On note C_n le coût du chemin $(X_0 \rightarrow X_1 \rightarrow \dots \rightarrow X_n)$. Alors, presque sûrement,

$$\frac{C_n}{n} \rightarrow \mathbf{E}_{\mu}(c(X_0, X_1)) = \sum_{x, y \in E} \mu(x)p(x, y)c(x, y).$$

Ce résultat est une extension de la loi forte des grands nombres : en prenant pour $(X_n)_{n \geq 0}$ une suite de variables aléatoires indépendantes et identiquement distribuées, qui est, rappelons-le, une chaîne de Markov, et en posant $c(x, y) = y$, on retrouve exactement l'énoncé de la loi forte des grands nombres.

On en déduit le résultat suivant :

Proposition 1.14 *Soit $(X_n)_{n \geq 0}$ une chaîne de Markov irréductible admettant une mesure stationnaire μ . Fixons un état $e \in E$ et notons $\rho_e(n)$ le nombre de visites de l'état e par la chaîne de Markov avant l'instant n :*

$$\rho_e(n) = \text{card}\{k \leq n, X_k = e\} = \sum_{k=1}^n \mathbf{1}_{X_k=e}.$$

Alors $\rho_e(n)/n$ tend presque sûrement vers $\mu(e)$.

Preuve : Il suffit d'appliquer la proposition 1.13 à la fonction $f(x, y) = \mathbf{1}_{y=e}$.

1.7 Simulation

Pourquoi faire

Lorsque l'on simule un échantillon aléatoire, on répète généralement un certain nombre k de fois la même opération, et on obtient un échantillon de taille k , c'est-à-dire k nombres répartis suivant une loi fixée.

Pour simuler une chaîne de Markov, l'objectif est différent : il faut simuler les chemins $X_0 \rightarrow X_1 \rightarrow \dots \rightarrow X_n$. Un échantillon de taille 1 est donc un chemin, dont on fixe a priori la longueur n .

La simulation peut avoir plusieurs buts :

- Obtenir une estimation du coût moyen d'un chemin. Dans ce cas, un échantillon de taille 1 suffit, puisque C_n/n tend presque sûrement vers $\mathbf{E}_\mu(c(X_0, X_1))$.
- Simuler la mesure stationnaire, en obtenir un histogramme ou faire des tests sur cette loi, sans la calculer explicitement. Pour n assez grand, X_n suit « presque » la loi stationnaire, lorsque la chaîne est **irréductible** et **apériodique**, et si cette loi existe. Pour obtenir une simulation de la mesure stationnaire, on peut donc simuler k chemins $x_0 \rightarrow x_1 \rightarrow \dots \rightarrow x_n$, en ne gardant que x_n .

Les ingrédients

Il faut connaître la loi initiale et chacune des probabilités de transition.

La loi initiale n'est pas très importante dans le cas des chaînes irréductibles et apériodiques admettant une mesure stationnaire, puisque très rapidement, la loi de X_n est très proche de la loi stationnaire.

Il faut évidemment également disposer d'une liste suffisamment longue de nombres uniformément répartis sur $[0, 1]$.

La recette

Pour obtenir un chemin simulé, on commence par simuler l'état initial x_0 suivant la loi initiale. La loi de X_1 sachant $\{X_0 = x_0\}$ est ensuite donnée par $p(x_0, x) : \mathbf{P}(X_1 = x | X_0 = x_0) = p(x_0, x)$; on simule donc une valeur suivant la loi donnée par la ligne x_0 de la matrice de transition et on obtient une valeur x_1 . Pour simuler X_2 , on recommence de la même façon, mais avec la loi donnée par la ligne x_1 de la matrice de transition et ainsi de suite, jusqu'à obtenir le chemin de longueur désirée.

Attention : Même si la loi initiale est la mesure stationnaire μ , cela ne revient pas du tout au même de simuler n fois μ (comme on le ferait pour obtenir un échantillon « standard »), et de simuler le chemin de la chaîne de Markov, pour lequel il peut y avoir par exemple des transitions interdites.

Pour simuler des lois discrètes : Nous décrivons ici comment simuler une loi discrète ν sur $\{1, \dots, m\}$, donnée par $p_1 = \nu(1), \dots, p_m = \nu(m)$.

Pour obtenir un échantillon de taille 1, on découpe l'intervalle $[0, 1]$ en m sous-intervalles de longueur $p_1, p_2, \dots, p_m$. On a ainsi par exemple $I_1 = [0, p_1[$, $I_2 = [p_1, p_1 + p_2[$, $\dots$, $I_{m-1} = [p_1 + \dots + p_{m-2}, p_1 + \dots + p_{m-1}[$ et $I_m = [1 - p_m, 1]$. On prend une valeur u (la première valeur) de l'échantillon uniforme, et si u appartient à l'intervalle I_j , on pose $x_1 = j$.

Pour obtenir un échantillon de taille k , on recommence k fois l'opération précédente, en renouvelant le tirage de la valeur uniforme : on utilise successivement $u_1, u_2, \dots, u_k$.

En résumé, le plus simple est de commencer par décrire la méthode de simulation choisie, pour chacune des lois dont on va avoir besoin : la loi initiale, puis la loi donnée par chacune des lignes de la matrice. On procède alors à la simulation en utilisant l'échantillon uniforme, sans reprendre deux fois le même u_i . Si on a besoin de plusieurs chemins, on applique autant de fois qu'il le faut la méthode,

en utilisant à chaque fois des échantillon uniformes disjoints : par exemple les $n + 1$ premières valeurs de l'échantillon uniforme permettront de construire le premier chemin de longueur n , les $n + 1$ valeurs suivantes le deuxième chemin et ainsi de suite.

2 Marches aléatoires au plus proche voisin

Nous étudions dans ce paragraphe une chaîne de Markov particulière : la marche aléatoire au plus proche voisin sur $\mathbb{Z}$: le « marcheur » se déplace sur $\mathbb{Z}$ vers la gauche ou vers la droite en faisant un pas par unité de temps, indépendamment du chemin déjà parcouru.

Cette chaîne peut s'écrire de la façon suivante : $X_0 = 0$ et $X_n = X_{n-1} + Y_n$, où les Y_n sont des variables aléatoires indépendantes et identiquement distribuées telles que $\mathbf{P}(Y_n = 1) = p$, $\mathbf{P}(Y_n = -1) = q = 1 - p$.

C'est une chaîne irréductible et de période 2. On dira qu'elle est symétrique lorsque $p = q = 1/2$ et asymétrique pour $p \neq q$.

2.1 Récurrence et transience

Lorsque $p \neq q$, il est facile de voir que la marche est transiente. Supposons en effet $p > q$. On a alors $\mathbf{E}(Y_1) = p - q > 0$ et, par la loi forte des grands nombres, $(Y_1 + \dots + Y_n)/n$ tend presque sûrement vers $\mathbf{E}(Y_1)$. Donc X_n tend presque sûrement vers $+\infty$: il n'existe pas une infinité d'indices n tels que $X_n = 0$, c'est-à-dire que la chaîne de Markov $(X_n)_{n \geq 0}$ est transiente (cf p.6). Ce raisonnement ne permet pas de résoudre le cas $p = q = 1/2$.

Proposition 1.15 *La marche au plus proche voisin sur $\mathbb{Z}$ est récurrente nulle si $p = 1/2$ et transiente si $p \neq 1/2$.*

Preuve : Comme la chaîne est irréductible, on s'intéresse uniquement à la nature de l'état 0. Nous allons utiliser le critère du théorème 1.9 pour vérifier que cela.

$$\begin{aligned} p^{(2n)}(0,0) &= \mathbf{P}(n \text{ des variables } (Y_k)_{k \leq 2n} \text{ sont égales à } 1 \text{ et les autres sont égales à } -1) \\ &= C_{2n}^n p^n q^n \\ &\sim \frac{(2n)^{2n} \sqrt{4\pi n} e^{-2n}}{(n^n \sqrt{2\pi n} e^{-n})^2} (pq)^n = \sqrt{\frac{1}{n\pi}} (4pq)^n \end{aligned}$$

Si $p = 1/2$, la série de terme général $p^{(2n)}(0,0)$ est donc divergente, ce qui implique que 0 est un état récurrent et $p^{(2n)}$ tend vers 0, ce qui implique que 0 est nécessairement récurrent nul.

Si $p \neq 1/2$, $4pq = 4p(1-p)$ est strictement inférieur à 1 donc la série de terme général $p^{(2n)}(0,0)$ est convergente : on retrouve que 0 est transient.

2.2 Temps de retour

À titre « culturel », nous allons également nous intéresser à la loi de $\tau_1 = \inf\{n \geq 1, X_n = 1\}$ et de $\tau_{-1} = \inf\{n \geq 1, X_n = -1\}$, ainsi qu'à celle de $\tau_0 = \inf\{n, X_n = 0\}$, pour une chaîne issue de $X_0 = 0$. Notons Φ la fonction génératrice de la loi de τ_1 définie pour $s \in [0, 1]$ par

$$\Phi(s) = \sum_{n \geq 1} \mathbf{P}_0(\tau_1 = n) s^n.$$

Remarquons pour la suite que $\Phi(1) = \mathbf{P}_0(\tau_1 < \infty)$.

Calculons les premières valeurs de $\mathbf{P}_0(\tau_1 = n)$: $\mathbf{P}_0(\tau_1 = 0) = 0$, $\mathbf{P}_0(\tau_1 = 1) = p$, $\mathbf{P}_0(\tau_1 = 2) = 0$ et $\mathbf{P}_0(\tau_1 = 3) = \mathbf{P}_0(X_1 = -1, X_2 = 0, X_3 = 1) = qp^2$.

Nous allons maintenant prouver une relation de récurrence entre les $\mathbf{P}_0(\tau_1 = n)$ en décomposant les trajectoires. Pour $n \geq 3$, pour avoir $\tau_1(\omega) = n$, il faut que $X_1(\omega) = -1$, puis il faut mettre un certain nombre m de pas pour rejoindre 0, puis $n - m - 1$ pas pour rejoindre 0. On a ainsi

$$\begin{aligned}
\mathbf{P}_0(\tau_1 = n) &= \sum_{m=1}^{n-2} \mathbf{P}(X_1 = -1, \\
&\quad \text{et la suite } (-1 + Y_2 + \dots + Y_k) \text{ atteint 0 pour la 1}^{\text{ère}} \text{ fois en } k = m + 1, \\
&\quad \text{et la suite } (0 + Y_{m+2} + \dots + Y_l) \text{ atteint 1 pour la 1}^{\text{ère}} \text{ fois en } l = n) \\
&= q \sum_{m=1}^{n-2} \mathbf{P}_0(\text{ la suite } (Y_2 + \dots + Y_k) \text{ atteint 1 pour la 1}^{\text{ère}} \text{ fois en } k = m + 1, \\
&\quad \text{et la suite } (Y_{m+2} + \dots + Y_l) \text{ atteint 1 pour la 1}^{\text{ère}} \text{ fois en } l = n) \\
&= q \sum_{m=1}^{n-2} \mathbf{P}(\text{ la suite } (Y_2 + \dots + Y_k) \text{ atteint 1 pour la 1}^{\text{ère}} \text{ fois en } k = m + 1) \times \\
&\quad \mathbf{P}(\text{ la suite } (Y_{m+2} + \dots + Y_l) \text{ atteint 1 pour la 1}^{\text{ère}} \text{ fois en } l = n) \\
&= q \sum_{m=1}^{n-2} \mathbf{P}_0(\tau_1 = m) \mathbf{P}_0(\tau_1 = n - m - 1)
\end{aligned}$$

On peut vérifier alors, en calculant Φ^2 en tant que carré d'une série entière, que Φ est solution de l'équation suivante : $\Phi(s) = ps + qs\Phi^2(s)$

Φ est donc en fait l'une des racines d'un polynôme de degré 2. On obtient

$$\Phi(s) = \frac{1 \pm \sqrt{1 - 4pqs^2}}{2qs}.$$

En remarquant que $\Phi(0) = 0$, on conclut finalement que

$$\Phi(s) = \frac{1 - \sqrt{1 - 4pqs^2}}{2qs}.$$

Pour obtenir $\mathbf{P}_0(\tau_1 = n)$, il suffit alors de décomposer Φ en série entière. On retrouve notamment que lorsque n est pair, $\mathbf{P}_0(\tau_1 = n) = 0$.

En utilisant cette expression de Φ , on peut maintenant calculer $\mathbf{P}_0(\tau_1 < \infty)$: $\mathbf{P}_0(\tau_1 < \infty) = \Phi(1)$. En remarquant que $1 - 4pq = (1 - 2p)^2$, on en déduit que lorsque $p \leq 1/2 \leq q$, on a $\mathbf{P}_0(\tau_1 < \infty) = p/q$ et lorsque $p \geq 1/2 \geq q$, on a $\mathbf{P}_0(\tau_1 < \infty) = 1$.

La fonction génératrice de τ_{-1} et sa loi s'obtiennent à partir de celles de τ_1 en échangeant les rôles de p et q .

Passons maintenant à la loi de τ_0 . On décompose la trajectoire suivant le premier pas de la marche. Pour que τ_0 soit égal à n , il faut commencer par faire un pas soit vers le bas, soit vers le haut. Ensuite, on met $n - 1$ pas pour rejoindre 0, en venant soit du bas, soit du haut, ce qui correspond soit à un $\tilde{\tau}_1$, soit à un $\tilde{\tau}_{-1}$, qui est indépendant du premier pas effectué. On a donc, pour tout $n \geq 1$,

$$\begin{aligned}
\mathbf{P}_0(\tau_0 = n) &= \mathbf{P}_0(X_1 = -1) \mathbf{P}_{-1}(\tau_0 = n - 1) + \mathbf{P}_0(X_1 = 1) \mathbf{P}_1(\tau_0 = n - 1). \\
&= q \mathbf{P}_0(\tau_1 = n - 1) + p \mathbf{P}_0(\tau_{-1} = n - 1)
\end{aligned}$$

En particulier, en sommant ces égalités, il vient :

$$\mathbf{P}_0(\tau_0 < \infty) = q \mathbf{P}_0(\tau_1 < \infty) + p \mathbf{P}_0(\tau_{-1} < \infty).$$

Si $p = q = 1/2$, $\mathbf{P}_0(\tau_1 < \infty) = \mathbf{P}_0(\tau_{-1} < \infty) = 1$, donc $\mathbf{P}_0(\tau_0 < \infty) = 1$. On retrouve donc que 0 est récurrent. Φ n'étant pas dérivable en 1, $\mathbf{E}_0(\tau_1) = +\infty$. Nécessairement, on a également $\mathbf{E}_0(\tau_0) = +\infty$, c'est-à-dire que 0 est récurrent nul.

Si $p < q$, $\mathbf{P}_0(\tau_1 < \infty) = p/q < 1$ et $\mathbf{P}_0(\tau_{-1} < \infty) = 1$ donc $\mathbf{P}_0(\tau_0 < \infty) = p + p < 1$. L'état 0 est transient.

Remarque : Obtenir la loi de τ_0 est un moyen de vérifier la transience ou la récurrence d'un point, mais ce n'est pas le plus simple ! Le critère 1.9 est généralement bien plus simple à mettre en œuvre.

Chapitre 2

Espérance conditionnelle et martingales

1 Construction et propriétés de l'espérance conditionnelle

Il s'agit uniquement de rappels ! Consultez votre cours de 1^{ère} année pour les preuves et compléments. Prenez garde à bien différencier les signes $\in$ et $\subset$.

1.1 Définitions

Rappelons avant tout la définition d'une tribu sur un ensemble Ω :

$\mathcal{F} \subset \mathcal{P}(\Omega)$ est une **tribu** si,

- $\emptyset \in \mathcal{F}$
- Si $A \in \mathcal{F}$ alors $\Omega \setminus A \in \mathcal{F}$
- si les $(A_n)_{n \geq 1}$ sont dans $\mathcal{F}$, alors $\cup_n A_n \in \mathcal{F}$.

Une **sous-tribu** $\mathcal{F}$ d'une tribu Σ est telle que $\mathcal{F} \subset \Sigma$, c'est-à-dire que si $A \in \mathcal{F}$ alors $A \in \Sigma$.

La **tribu engendrée** par une variable aléatoire Y est la tribu $\sigma(Y)$ définie par

$$\sigma(Y) = \{Y \in B, B \in \mathcal{B}(\mathbb{R})\}.$$

Une variable aléatoire est **indépendante** d'une tribu $\mathcal{F}$ si $\forall B \in \mathcal{B}(\mathbb{R})$ et $\forall A \in \mathcal{F}$,

$$\mathbf{P}(\{X \in B\} \cup A) = \mathbf{P}(X \in B)\mathbf{P}(A).$$

Soient maintenant une variable aléatoire $X : (\Omega, \Sigma, \mathbf{P}) \rightarrow (\mathbb{R}, \mathcal{B}(\mathbb{R}))$ et une sous-tribu $\mathcal{F}$ de Σ . On suppose que X est **intégrable** (c'est-à-dire $\mathbf{E}|X| < \infty$). L'**espérance conditionnelle** de X sachant $\mathcal{F}$ est alors la **variable aléatoire** Y $\mathcal{F}$ -mesurable telle que, pour tout $A \in \mathcal{F}$,

$$\int_A X \, d\mathbf{P} = \int_A Y \, d\mathbf{P}.$$

Cette variable aléatoire est unique, moyennant les « presque sûrement » usuels. Si $\mathcal{F}$ est la tribu engendrée par une variable aléatoire Z , on parle d'espérance conditionnelle sachant Z et on note $\mathbf{E}(X|Z) = \mathbf{E}(X|\sigma(Z))$.

Justification de l'existence : voir le théorème de Radon-Nicodyn.

Exemples :

- Si X est intégrable et $\mathcal{F}$ -mesurable, alors $\mathbf{E}(X|\mathcal{F}) = X$.
- Si X est $\mathcal{F}$ -mesurable et Y est une variable aléatoire intégrable telle que XY est intégrable, alors $\mathbf{E}(XY|\mathcal{F}) = X\mathbf{E}(Y|\mathcal{F})$.
- Si X est une variable aléatoire intégrable et indépendante de $\mathcal{F}$ alors $\mathbf{E}(X|\mathcal{F}) = \mathbf{E}(X)$.

- Si X est une variable aléatoire intégrable et si $\mathcal{F}$ est la tribu engendrée par une variable aléatoire N à valeurs dans $\mathbb{N}$, alors

$$\mathbf{E}(X|\mathcal{F}) = \sum_{n \geq 0} \mathbf{E}(X|\{N = n\}) \mathbf{1}_{N=n} = \sum_{n \geq 0} \frac{\mathbf{E}(X \mathbf{1}_{N=n})}{\mathbf{P}(N = n)} \mathbf{1}_{N=n}.$$

1.2 Propriétés

Soient X et Y deux variables aléatoires intégrables et $\mathcal{F}$ et $\mathcal{G}$ deux tribus.

L'espérance conditionnelle vérifie les propriétés suivantes :

Proposition 2.1

1. $\mathbf{E}(\mathbf{E}(X|\mathcal{F})) = \mathbf{E}(X)$.
2. Pour tous réels a et b , $\mathbf{E}(aX + bY|\mathcal{F}) = a\mathbf{E}(X|\mathcal{F}) + b\mathbf{E}(Y|\mathcal{F})$.
3. Si $\mathcal{F} \subset \mathcal{G}$, $\mathbf{E}(\mathbf{E}(X|\mathcal{G})|\mathcal{F}) = \mathbf{E}(X|\mathcal{F})$ et $\mathbf{E}(\mathbf{E}(X|\mathcal{F})|\mathcal{G}) = \mathbf{E}(X|\mathcal{F})$.
4. Si $X \leq Y$ p.s., alors $\mathbf{E}(X|\mathcal{F}) \leq \mathbf{E}(Y|\mathcal{F})$.
5. Si Φ est une fonction convexe positive (ou telle que $\Phi(X)$ soit intégrable), alors

$$\mathbf{E}(\Phi(X)|\mathcal{F}) \geq \Phi(\mathbf{E}(X|\mathcal{F})).$$

Commentaires :

- Le point 1 est évident (mais très utile). Pour le prouver, il suffit d'écrire la définition de l'espérance conditionnelle en remplaçant A par Ω .
- Le dernier point est l'inégalité de Jensen. Pour mémoriser le sens de l'inégalité, pensez au cas $\Phi(x) = x^2$.

Exercice : Faites la preuve de la proposition précédente !

2 Martingales en temps discret

2.1 Définition, exemples

Définition 2.2

- Une suite croissante de tribus $(\mathcal{F}_n)_{n \geq 0}$, i.e. telle que $\mathcal{F}_n \subset \mathcal{F}_{n+1}$ est appelée une **filtration**.
- Un processus $(M_n)_n$ est **adapté** à la filtration $(\mathcal{F}_n)$ si tout n , la variable aléatoire M_n est $\mathcal{F}_n$ -mesurable.
- Soit $(\mathcal{F}_n)_{n \geq 0}$ une filtration et $(M_n)_{n \geq 0}$ une suite de variables aléatoires. $(M_n, \mathcal{F}_n)_{n \geq 0}$ est une **martingale** si
 - pour tout $n \geq 0$, M_n est une variable aléatoire intégrable : $\mathbf{E}|M_n| < \infty$,
 - M_n est $\mathcal{F}_n$ -mesurable (ou (M_n) est $\mathcal{F}_n$ -adapté),
 - $\mathbf{E}(M_{n+1}|\mathcal{F}_n) = M_n$ p.s.

Exemples :

- La marche symétrique : $S_0 = 0$, $S_n = X_1 + \dots + X_n$, où les X_i sont des variables aléatoires indépendantes et identiquement distribuées, $\mathbf{P}(X_1 = 1) = \mathbf{P}(X_1 = -1) = 1/2$, est une martingale. Une filtration adéquate est donnée par $\mathcal{F}_0 = \{\emptyset, \Omega\}$ et $\mathcal{F}_n = \sigma\{X_1, \dots, X_n\} = \sigma\{S_1, \dots, S_n\}$.
- Plus généralement : on se donne des variables aléatoires (X_n) indépendantes, intégrables et centrées et on pose $S_0 = 0$, $S_n = X_1 + \dots + X_n$. Vérifiez que le processus obtenu est une martingale (même filtration que dans l'exemple précédent).

- Toutes les martingales ne s'écrivent pas comme des sommes de variables aléatoires indépendantes : en prenant les (X_i) de carré intégrable dans l'exemple ci-dessus, et de même variance σ^2 , et en posant $M_n = S_n^2 - n\sigma^2$, on obtient à nouveau une martingale. $M_{n+1} - M_n = X_{n+1}^2 + 2X_{n+1}S_n - \sigma^2$ n'est en général pas indépendante de M_n .

Une propriété importante des martingales est la suivante :

Proposition 2.3 Soit $(M_n, \mathcal{F}_n)_{n \geq 0}$ une martingale.

- Pour tout $k \geq n \geq 0$, on a $\mathbf{E}(M_k | \mathcal{F}_n) = M_n$ p.s.
- $\mathbf{E}(M_n) = \mathbf{E}(M_0)$.

Preuve : en exercice.

Définition 2.4 Soit $(\mathcal{F}_n)_n$ une filtration et $(M_n)_n$ une suite de variables aléatoires **intégrables** et $\mathcal{F}_n$ -**mesurables**.

- On dit que $(M_n, \mathcal{F}_n)$ est une **sous-martingale** si $\mathbf{E}(M_{n+1} | \mathcal{F}_n) \geq M_n$.
- On dit que $(M_n, \mathcal{F}_n)$ est une **sur-martingale** si $\mathbf{E}(M_{n+1} | \mathcal{F}_n) \leq M_n$.

Attention : Ne pas oublier de vérifier l'intégrabilité et la mesurabilité !

Exercice : Montrez que si (M_n) est une sous-martingale, la suite $(\mathbf{E}(M_n))_n$ est croissante.

Exemple : Si $(M_n, \mathcal{F}_n)$ est une martingale telle que pour tout n , $\mathbf{E}(|M_n|^p) < \infty$, où $p > 1$ est un réel fixé, alors $(|M_n|^p, \mathcal{F}_n)_n$ est une sous-martingale.

2.2 Temps d'arrêt

Définition 2.5 Une variable aléatoire τ , à valeurs dans $\mathbb{N} \cup \{+\infty\}$ est un temps d'arrêt pour la filtration $(\mathcal{F}_n)_n$ si pour tout $n \in \mathbb{N}$, on a $\{\tau = n\} \in \mathcal{F}_n$.

Une autre caractérisation est la suivante :

Proposition 2.6 Une variable aléatoire τ à valeurs dans $\mathbb{N} \cup \{+\infty\}$ est un temps d'arrêt pour la filtration $(\mathcal{F}_n)_n$ si et seulement si, pour tout $n \in \mathbb{N}$, on a $\{\tau \leq n\} \in \mathcal{F}_n$.

Convention : Dans la définition d'un temps d'arrêt, on utilisera fréquemment la convention suivante : $\inf\{\emptyset\} = +\infty$ (et sa contrepartie : $\sup\{\emptyset\} = -\infty$ ou 0).

Exemples : Soit $(M_n, \mathcal{F}_n)$ une martingale et λ un réel positif.

- $\tau = \inf\{n \geq 0, M_n \leq \lambda\}$ est un temps d'arrêt.
- $\sigma = \inf\{n \geq 0, |M_n| \leq \lambda\}$ est un temps d'arrêt.
- $\rho = \inf\{n \geq 40, M_n \geq \lambda\}$ est un temps d'arrêt.
- $\phi = \sup\{n \geq 0, M_n \leq \lambda\}$ n'est pas un temps d'arrêt.
- $\tilde{\rho} = \inf\{n \geq \sigma, M_n \geq \lambda\}$ est un temps d'arrêt si σ en est un.

Exercice : Les constantes sont des temps d'arrêt. Si τ et σ sont des temps d'arrêt, les variables aléatoires $\min(\tau, \sigma)$ et $\max(\tau, \sigma)$ sont des temps d'arrêt.

Définition 2.7 On définit la tribu $\mathcal{F}_\tau$ par

$$\mathcal{F}_\tau = \{A \in \Sigma; \forall n \in \mathbb{N}, A \cap \{\tau = n\} \in \mathcal{F}_n\}.$$

Exercice :

1. Vérifiez que le membre de droite de l'égalité ci-dessus définit bien une tribu.
2. Montrez que si τ est un temps d'arrêt constant égal à n , alors $\mathcal{F}_\tau$ (comme défini ci-dessus) est la tribu $\mathcal{F}_n$ (de la filtration).
3. Si τ et σ sont deux temps d'arrêt, l'événement $\{\tau \leq \sigma\}$ appartient aux tribus $\mathcal{F}_\tau$ et $\mathcal{F}_\sigma$.
4. τ est une variable aléatoire $\mathcal{F}_\tau$ -mesurable.
5. Montrez que si $\tau \leq \sigma$ **presque sûrement**, alors $\mathcal{F}_\tau \subset \mathcal{F}_\sigma$.
6. Montrez que si τ et σ sont deux temps d'arrêt et si $\rho = \min(\tau, \sigma)$ alors $\mathcal{F}_\rho = \mathcal{F}_\sigma \cap \mathcal{F}_\tau$.
7. Montrez que si $(M_n)_n$ est une $\mathcal{F}_n$ -martingale et si τ est un temps d'arrêt par rapport à $(\mathcal{F}_n)_n$ vérifiant $\mathbf{P}(\tau < \infty) = 1$ alors M_τ est $\mathcal{F}_\tau$ -mesurable.

Le théorème suivant, appelé **théorème d'arrêt** ou **théorème optionnel de Doob** est le plus important de ce chapitre sur les martingales. Moralement, il signifie que, moyennant quelques conditions, il est possible de conditionner par rapport à des tribus $\mathcal{F}_\sigma$ (avec un temps d'arrêt σ) de la même façon que par rapport à une tribu $\mathcal{F}_n$.

Théorème 2.8 Soit $(M_n, \mathcal{F}_n)$ une martingale et τ et σ deux temps d'arrêt à valeurs dans $\mathbb{N}$ (et non $\mathbb{N} \cup \{\infty\}$) tels que :

- (i) $\mathbf{E}|M_\sigma| < \infty$, (ii) $\mathbf{E}|M_\tau| < \infty$,
- (iii) $\liminf_n \mathbf{E}(|M_n| \mathbf{1}_{\tau \geq n}) = 0$.

Alors on a

$$\mathbf{E}(M_\tau \mathbf{1}_{\tau \geq \sigma} | \mathcal{F}_\sigma) = M_\sigma \mathbf{1}_{\tau \geq \sigma} \quad p.s..$$

Ce théorème est également vrai pour les sous-martingales et les sur-martingales :

Théorème 2.9 Soient $(\mathcal{F}_n)_n$ une filtration, τ et σ deux temps d'arrêt par rapport à cette filtration, à valeurs dans $\mathbb{N}$, et $(M_n)_n$ un processus $\mathcal{F}_n$ -intégrable vérifiant les trois hypothèses (i), (ii) et (iii) du théorème précédent. Alors

– Si $(M_n)_n$ est une sous-martingale, on a

$$\mathbf{E}(M_\tau \mathbf{1}_{\tau \geq \sigma} | \mathcal{F}_\sigma) \geq M_\sigma \mathbf{1}_{\tau \geq \sigma} \quad p.s..$$

– Si $(M_n)_n$ est une sur-martingale, on a

$$\mathbf{E}(M_\tau \mathbf{1}_{\tau \geq \sigma} | \mathcal{F}_\sigma) \leq M_\sigma \mathbf{1}_{\tau \geq \sigma} \quad p.s..$$

Les hypothèses des deux théorèmes ci-dessus sont fréquemment difficiles à prouver. Très souvent, on utilisera ces théorèmes sous l'une des conditions suivantes :

1. Si M_n est **uniformément intégrable**, c'est-à-dire si $\lim_k \sup_n \mathbf{E}(|M_n| \mathbf{1}_{|M_n| \geq k}) = 0$, alors (iii) est vérifiée.
2. Si τ est borné par une constante déterministe K , les conditions (ii) et (iii) sont remplies.
3. On appliquera souvent le théorème d'arrêt avec $\sigma = 0$ ou 1 . La tribu $\mathcal{F}_\sigma$ est alors particulièrement simple, et on s'intéresse en fait à $\mathbf{E}(M_\tau)$.
4. Toujours avec $\sigma = 0$ ou 1 et τ un temps d'arrêt quelconque. Si l'on ne parvient pas à prouver les hypothèses (ii) et (iii), on remplace τ par $\tau \wedge N$ (N déterministe) et on fait tendre N vers l'infini. Pour le temps d'arrêt $\tau \wedge N$, les hypothèses (ii) et (iii) sont vérifiées et il reste à justifier un passage à la limite sous l'espérance, ce qui se fait classiquement en utilisant les théorèmes de convergence dominée de Lebesgue ou de convergence monotone (Beppo-Levi).

Preuve : pour une sous-martingale.

Il faut montrer que $\mathbf{E}(M_\tau \mathbf{1}_{\tau \geq \sigma} | \mathcal{F}_\sigma) \geq M_\sigma \mathbf{1}_{\tau \geq \sigma}$, ce qui est équivalent à

$$\forall A \in \mathcal{F}_\sigma, \quad \mathbf{E}(\mathbf{1}_A M_\tau \mathbf{1}_{\tau \geq \sigma}) \geq \mathbf{E}(\mathbf{1}_A M_\sigma \mathbf{1}_{\tau \geq \sigma})$$

ou encore

$$\forall A \in \mathcal{F}_\sigma, \quad \forall n \geq 0, \quad \mathbf{E}(\mathbf{1}_{A \cap \{\sigma=n\} \cap \{\tau \geq \sigma\}} M_\tau) \geq \mathbf{E}(\mathbf{1}_{A \cap \{\sigma=n\} \cap \{\tau \geq \sigma\}} M_\sigma).$$

Fixons donc $A \in \mathcal{F}_\sigma$ et $n \geq 0$. L'événement $A \cap \{\sigma = n\}$ appartient à la tribu $\mathcal{F}_n$.

Remarquons que l'on peut décomposer l'événement $A \cap \{\sigma = n\} \cap \{\tau \geq \sigma\}$ en

$$A \cap \{\sigma = n\} \cap \{\tau \geq \sigma\} = (A \cap \{\tau = \sigma = n\}) \cup (A \cap \{\sigma = n\} \cap \{\tau \geq n+1\})$$

et l'événement $A \cap \{\sigma = n\} \cap \{\tau \geq n+1\}$ appartient à la tribu $\mathcal{F}_n$.

On a donc, en utilisant le fait que (M_n) est une sous-martingale,

$$\begin{aligned} \mathbf{E}(\mathbf{1}_{A \cap \{\sigma=n\} \cap \{\tau \geq \sigma\}} M_\sigma) &= \mathbf{E}(\mathbf{1}_{A \cap \{\sigma=n\} \cap \{\tau \geq \sigma\}} M_n) \\ &= \mathbf{E}(\mathbf{1}_{A \cap \{\tau=\sigma=n\}} M_n) + \mathbf{E}(\mathbf{1}_{A \cap \{\sigma=n\} \cap \{\tau \geq n+1\}} M_n) \\ &\leq \mathbf{E}(\mathbf{1}_{A \cap \{\tau=\sigma=n\}} M_\tau) + \mathbf{E}(\mathbf{1}_{A \cap \{\sigma=n\} \cap \{\tau \geq n+1\}} M_{n+1}) \end{aligned}$$

Recommençons les mêmes opérations avec $\mathbf{E}(\mathbf{1}_{A \cap \{\sigma=n\} \cap \{\tau \geq n+1\}} M_{n+1})$.

$$A \cap \{\sigma = n\} \cap \{\tau \geq n+1\} = (A \cap \{\sigma = n\} \cap \{\tau = n+1\}) \cup (A \cap \{\sigma = n\} \cap \{\tau \geq n+2\})$$

et l'événement $A \cap \{\sigma = n\} \cap \{\tau \geq n+2\}$ appartient à $\mathcal{F}_{n+1}$ donc

$$\begin{aligned} \mathbf{E}(\mathbf{1}_{A \cap \{\sigma=n\} \cap \{\tau \geq n+1\}} M_{n+1}) &= \mathbf{E}(\mathbf{1}_{A \cap \{\sigma=n\} \cap \{\tau=n+1\}} M_\tau) + \mathbf{E}(\mathbf{1}_{A \cap \{\sigma=n\} \cap \{\tau \geq n+2\}} M_{n+1}) \\ &\leq \mathbf{E}(\mathbf{1}_{A \cap \{\sigma=n\} \cap \{\tau=n+1\}} M_\tau) + \mathbf{E}(\mathbf{1}_{A \cap \{\sigma=n\} \cap \{\tau \geq n+2\}} M_{n+2}) \end{aligned}$$

On obtient :

$$\mathbf{E}(\mathbf{1}_{A \cap \{\sigma=n\} \cap \{\tau \geq \sigma\}} M_\sigma) \leq \mathbf{E}(\mathbf{1}_{A \cap \{n=\sigma \leq \tau \leq n+1\}} M_\tau) + \mathbf{E}(\mathbf{1}_{A \cap \{\sigma=n\} \cap \{\tau \geq n+2\}} M_{n+2}).$$

En itérant le procédé, on montre que, pour tout $p \geq 1$,

$$\mathbf{E}(\mathbf{1}_{A \cap \{\sigma=n\} \cap \{\tau \geq \sigma\}} M_\sigma) \leq \mathbf{E}(\mathbf{1}_{A \cap \{n=\sigma \leq \tau \leq n+p\}} M_\tau) + \mathbf{E}(\mathbf{1}_{A \cap \{\sigma=n\} \cap \{\tau \geq n+p+1\}} M_{n+p+1}).$$

(Exercice : Écrire proprement la récurrence !)

Lorsque p tend vers $+\infty$, $\mathbf{E}(\mathbf{1}_{A \cap \{n=\sigma \leq \tau \leq n+p\}} M_\tau)$ converge vers $\mathbf{E}(\mathbf{1}_{A \cap \{n=\sigma \leq \tau\}} M_\tau)$ car M_τ est intégrable. De plus, l'hypothèse (iii) assure l'existence d'une sous-suite p_k telle que

$$\mathbf{E}(\mathbf{1}_{\{\tau \geq n+p_k+1\}} |M_{n+p_k+1}|) \text{ tend vers } 0.$$

On aura alors, pour tout $A \in \Sigma$,

$$\begin{aligned} |\mathbf{E}(\mathbf{1}_{A \cap \{\sigma=n\} \cap \{\tau \geq n+p_k+1\}} M_{n+p_k+1})| &\leq \mathbf{E}|\mathbf{1}_{A \cap \{\sigma=n\} \cap \{\tau \geq n+p_k+1\}} M_{n+p_k+1}| \\ &\leq \mathbf{E}|\mathbf{1}_{\tau \geq n+p_k+1} M_{n+p_k+1}| \rightarrow 0. \end{aligned}$$

On obtient finalement en passant à la limite suivant la sous-suite p_k :

$$\mathbf{E}(\mathbf{1}_{A \cap \{\sigma=n\} \cap \{\tau \geq \sigma\}} M_\sigma) \leq \mathbf{E}(\mathbf{1}_{A \cap \{n=\sigma \leq \tau\}} M_\tau),$$

ce qui est la conclusion du théorème.

2.3 Convergence

Définition 2.10 On dit qu'une suite (X_n) de variables aléatoires est uniformément bornée dans $\mathcal{L}^p$ si $\sup_n \mathbf{E}(|X_n|^p) < \infty$.

Théorème 2.11

- Si (M_n) est une sous-martingale telle que

$$\sup_n \mathbf{E}(M_n^+) < \infty$$

alors M_n converge presque sûrement vers une variable aléatoire M_∞ intégrable.

- Si (M_n) est une sous-martingale uniformément bornée dans $\mathcal{L}^p$, $p > 1$, alors M_n converge presque sûrement et dans $\mathcal{L}^p$ vers une variable aléatoire notée M_∞ .
- Si (M_n) est une martingale uniformément intégrable (c.f. p.17), elle converge presque sûrement et dans $\mathcal{L}^1$ et on a $\mathbf{E}(M_\infty|\mathcal{F}_n) = M_n$.

Les deux premiers points du théorème précédent sont dûs à J.L. Doob. Le dernier point est un résultat de P. Lévy.

Un corollaire du point 1) est le suivant :

Corollaire 2.12 Si (M_n) est une sous martingale majorée par une constante alors elle converge presque sûrement.

3 Martingales en temps continu

3.1 Tribus, filtrations

Les filtrations que l'on utilisera pour étudier les processus en temps continu (et en particulier les martingales) vérifieront toujours les conditions suivantes :

- $\mathcal{F}_0$ est **complète** : si N est un ensemble négligeable, $N \subset \mathcal{F}_0$.
- $(\mathcal{F}_t)_{t \geq 0}$ est une **filtration** : $\forall t \geq s \geq 0, \mathcal{F}_s \subset \mathcal{F}_t$.
- La filtration est **continue à droite** : $\mathcal{F}_t = \bigcap_{s > t} \mathcal{F}_s$.

3.2 Les processus càd-làg

Un processus $(X_t)_{t \geq 0}$ est dit **càd-làg** (continu à droite, limite à gauche) s'il existe un ensemble négligeable $N \in \Omega$ tel que, pour tout $\omega \notin N$, la trajectoire $t \rightarrow X_t(\omega)$ est continue à droite en tout $t \geq 0$ et admet une limite à gauche en tout $t > 0$.

3.3 Martingales

Définition 2.13 Un processus $(M_t)_t$ est **adapté** à la filtration $(\mathcal{F}_t)$ si pour tout t , la variable aléatoire M_t est $\mathcal{F}_t$ -mesurable.

Une **martingale en temps continu** par rapport à une filtration $(\mathcal{F}_t)$ complète et continue à droite sera pour nous un processus $(M_t)_{t \geq 0}$ càd-làg et vérifiant les conditions suivantes :

1. Pour tout $t \geq 0$, M_t est intégrable,
2. Pour tout $t \geq 0$, M_t est $\mathcal{F}_t$ -mesurable (ou (M_t) est $\mathcal{F}_t$ -adapté),
3. Pour tous $t \geq s \geq 0$, $\mathbf{E}(M_t|\mathcal{F}_s) = M_s$.

Un **temps d'arrêt** est une variable T à valeurs réelles vérifiant $\{T \leq t\} \in \mathcal{F}_t$, pour tout $t \geq 0$.

Proposition 2.14 *T est un temps d'arrêt si et seulement si $\{T < t\} \in \mathcal{F}_t$ pour tout $t \geq 0$.*

Preuve : Soit $t \geq 0$ fixé et supposons que pour tout $s \geq 0$, $\{T < s\} \in \mathcal{F}_s$. On a :

$$\begin{aligned}\{T \leq t\} &= \bigcap_{k \in \mathbb{N}^*} \left\{ T < t + \frac{1}{k} \right\} \\ &= \bigcap_{k \geq K} \left\{ T < t + \frac{1}{k} \right\}\end{aligned}$$

L'événement $\{T \leq t\}$ appartient donc à la tribu $\mathcal{F}_{t+1/K}$, pour tout K , donc il appartient à l'intersection de ces tribus, qui est la tribu $\mathcal{F}_t$.

Supposons maintenant que pour tout $s \geq 0$, $\{T \leq s\}$ appartienne à $\mathcal{F}_s$. On a :

$$\{T < t\} = \bigcup \left\{ T \leq t - \frac{1}{k} \right\}$$

donc $\{T < t\}$ appartient à $\mathcal{F}_t$.

Proposition 2.15 *Soit $(M_n, \mathcal{F}_n)_{n \geq 0}$ une martingale en temps discret. On suppose que $\mathcal{F}_0$ est complète. Pour tout $t \geq 0$, on pose $N_t = M_{[t]}$ et $\mathcal{G}_t = \mathcal{F}_{[t]}$, où $[t]$ désigne la partie entière de t .*

Alors $(\mathcal{G}_t)_t$ est une filtration continue à droite, $(N_t)_t$ est un processus càd-làg et c'est une martingale.

Les propriétés des temps d'arrêt en temps continu sont les mêmes qu'en temps discret. En particulier, le min, le max de deux temps d'arrêt sont des temps d'arrêt.

On définit également la tribu $\mathcal{F}_T$ par $\mathcal{F}_T = \{A \in \Sigma; \forall t \geq 0, A \cap \{T \leq t\} \in \mathcal{F}_t\}$. Comme ci-dessus, la condition $A \cap \{T \leq t\} \in \mathcal{F}_t$ peut être remplacée par $A \cap \{T < t\} \in \mathcal{F}_t$.

On peut vérifier les propriétés suivantes :

Proposition 2.16 *Soit T un temps d'arrêt par rapport à une filtration $(\mathcal{F}_t)$.*

- Si T est une constante t , $\mathcal{F}_T = \mathcal{F}_t$.
- Si T et S sont des temps d'arrêt vérifiant $T \leq S$ **presque sûrement**, alors $\mathcal{F}_T \subset \mathcal{F}_S$.
- T est $\mathcal{F}_T$ -mesurable.
- Si T et S sont deux temps d'arrêt, les événements $\{T = S\}$, $\{T \leq S\}$ et $\{T < S\}$ appartiennent aux tribus $\mathcal{F}_T$ et $\mathcal{F}_S$.

Le théorème d'arrêt de Doob reste lui aussi valable :

Théorème 2.17 *Soient $(M_t, \mathcal{F}_t)$ une martingale càd-làg et τ et σ deux temps d'arrêt à valeurs dans $\mathbb{R}^+$ (et non $\mathbb{R}^+ \cup \{+\infty\}$). Si les variables aléatoires M_τ et M_σ sont intégrables et si $\liminf_t \mathbf{E}(|M_t| \mathbf{1}_{\tau \geq t}) = 0$ alors on a*

$$\mathbf{E}(M_\tau \mathbf{1}_{\tau \geq \sigma} | \mathcal{F}_\sigma) = M_\sigma \mathbf{1}_{\tau \geq \sigma} \quad p.s.$$

Idée de la preuve : Essentiellement, on discrétise le temps en considérant $(M_t^{(n)}) = (M_{[nt]/n})$ et on pose $\tau^{(n)} = [n\tau]/n$ et $\sigma^{(n)} = [n\sigma]/n$. $(M_t^{(n)})_t$ est une martingale en temps discret (en tout cas, elle y ressemble beaucoup) par rapport à la filtration $(\mathcal{F}_{[nt]/n})_t$, et $\sigma^{(n)}$ et $\tau^{(n)}$ sont des temps d'arrêt par rapport à cette même filtration. On peut leur appliquer le théorème optionnel de Doob en temps discret. Pour conclure, il reste encore à effectuer un passage à la limite à la fois au niveau de la tribu $\mathcal{F}_\sigma$ et au niveau de l'espérance conditionnelle.

3.4 Martingales locales

L'intégrabilité est une hypothèse forte dans la construction des martingales : certains processus ressemblent fortement à des martingales du point de vue du conditionnement mais ne sont pas intégrables et on ne peut donc pas calculer leur espérance conditionnelle. Les martingales locales permettent de contourner ce problème :

Définition 2.18 *Un processus càd-làg $(M_t)_{t \geq 0}$ est une **martingale locale** pour une filtration $(\mathcal{F}_t)_{t \geq 0}$ s'il existe une suite croissante de temps d'arrêt $(\tau_n)_{n \geq 1}$ tendant vers $+\infty$ et telle que le processus $(M_{t \wedge \tau_n})_{t \geq 0}$ soit une martingale. On définit de façon similaire les sous-martingales locales et les sur-martingales locales.*

3.5 Variation quadratique

On se donne une martingale $(M_t, \mathcal{F}_t)$ de carré intégrable, c'est-à-dire que pour tout $t \geq 0$, $\mathbf{E}(M_t^2)$ est fini. On peut montrer qu'il existe un unique processus croissant noté $\langle M \rangle$ et appelé **variation quadratique** de M tel que $\langle M \rangle_0 = 0$ et tel que $(M_t^2 - \langle M \rangle_t)$ soit une martingale.

Lorsque (M_n) est une martingale discrète de carré intégrable, l'existence de $\langle M \rangle_n$ est claire : en effet, posons $X_n = M_n - M_{n-1}$, $X_0 = M_0$ et pour tout $n \geq 1$, $V_n = \sum_{i=1}^n X_i^2$. Les V_n sont des variables aléatoires intégrables et $\mathcal{F}_n$ -mesurables. On a

$$\begin{aligned} \mathbf{E}(M_{n+1}^2 - V_{n+1} | \mathcal{F}_n) &= \mathbf{E}(M_n^2 + 2M_n X_{n+1} + X_{n+1}^2 - V_n - X_{n+1}^2 | \mathcal{F}_n) \\ &= M_n^2 - V_n^2 + 2M_n \mathbf{E}(X_{n+1} | \mathcal{F}_n) \\ &= M_n^2 - V_n^2 \end{aligned}$$

car $\mathbf{E}(X_{n+1} | \mathcal{F}_n) = \mathbf{E}(M_{n+1} - M_n | \mathcal{F}_n) = \mathbf{E}(M_{n+1} | \mathcal{F}_n) - M_n = 0$.

V_n est donc un processus vérifiant $V_0 = 0$ et $(M_n^2 - V_n)$ est une $\mathcal{F}_n$ -martingale : c'est la variation quadratique de (M_n) . On remarque au passage que le processus obtenu est bien croissant, puisque c'est une somme de carrés.

L'existence de la variation quadratique d'une martingale en temps continu se base sur la discrétisation du temps (voir les idées de la preuve du théorème 2.17 d'arrêt de Doob).

Proposition 2.19 *La variation quadratique $(\langle M \rangle_t)$ est un processus $\mathcal{F}_t$ -mesurable et croissant.*

Remarque : La variation quadratique peut tout à fait être définie pour une martingale locale, ou pour une martingale qui ne serait pas de carré intégrable en ajoutant une suite de temps d'arrêt où il faut.

3.6 Convergence

Les résultats de convergence pour les martingales en temps continu sont les mêmes que pour les martingales en temps discret (cf p. 19) :

Théorème 2.20 *Soit $(M_t)_t$ une sous-martingale.*

1. *Si $\sup_t \mathbf{E}(M_t^+) < \infty$ alors M_t converge presque sûrement vers une variable aléatoire M_∞ intégrable.*
2. *Si, pour un certain $p > 1$, (M_t) est uniformément bornée dans $\mathcal{L}^p$, c'est-à-dire si*

$$\sup_t \mathbf{E}(|M_t|^p) < \infty,$$

alors (M_t) converge presque sûrement et dans $\mathcal{L}^p$.

3. *Si (M_t) est une martingale uniformément intégrable (c.f. p.17), elle converge presque sûrement et dans $\mathcal{L}^1$ et on a $\mathbf{E}(M_\infty | \mathcal{F}_t) = M_t$.*

Chapitre 3

Mouvement brownien, intégration stochastique

1 Processus gaussiens

1.1 Vecteurs gaussiens

Définition 3.1 $\bar{X} = (X_1, \dots, X_n)$ est un **vecteur aléatoire gaussien** si, pour tout vecteur $\bar{\alpha} = (\alpha_1, \dots, \alpha_n) \in \mathbb{R}^n$, $\langle \bar{\alpha}, \bar{X} \rangle = \sum_k \alpha_k X_k$ suit une loi de Gauss sur $\mathbb{R}$.

Il y a une grande différence entre un vecteur formé de variables aléatoires gaussiennes et un vecteur gaussien : en effet, si un vecteur est gaussien, chaque composante suit une loi de Gauss (prendre les α_i tous nuls sauf un dans la définition ci-dessus), mais la réciproque est fausse.

Par exemple : si X suit une loi de Gauss centrée et de variance 1 et si ε est une variable aléatoire indépendante de X et de loi donnée par $\mathbf{P}(\varepsilon = 1) = \mathbf{P}(\varepsilon = -1) = 1/2$, $Y = \varepsilon X$ suit également une loi de Gauss centrée et de variance 1 (à faire en exercice), mais $X + Y$ ne suit pas une loi de Gauss donc le couple (X, Y) n'est pas gaussien. Par contre si X_1 et X_2 sont des variables aléatoires gaussiennes indépendantes, le couple (X_1, X_2) est gaussien, tout comme le couple $(aX_1 + bX_2, cX_1 + dX_2)$ pour tout choix de réels a, b, c et d .

Remarque : Les variables aléatoires **constantes** sont considérées comme gaussiennes. En fait, la bonne caractérisation du caractère gaussien est la fonction caractéristique : une variable aléatoire X est gaussienne si sa fonction caractéristique est de la forme

$$\mathbf{E}(\exp(i\alpha X)) = \exp(i\alpha m) \exp(-\alpha^2 \sigma^2/2),$$

où m est la moyenne et σ^2 est la variance de la variable aléatoire. Une variable aléatoire constante a bien une telle fonction caractéristique.

Cette caractérisation des variables aléatoires gaussiennes par leur fonction caractéristique est bien plus générale et s'étend aux vecteurs gaussiens.

Proposition 3.2 Un vecteur $\bar{X} = (X_1, \dots, X_n)$ est un vecteur aléatoire gaussien si et seulement s'il existe une matrice symétrique positive $\Gamma \in \mathcal{M}_n(\mathbb{R})$ et un vecteur $\bar{m} = (m_1, \dots, m_n)$ tels que pour tout $\bar{\alpha} = (\alpha_1, \dots, \alpha_n) \in \mathbb{R}^n$,

$$\mathbf{E}(\exp(i\langle \bar{\alpha}, \bar{X} \rangle)) = \mathbf{E}\left(\exp\left(i\sum_k \alpha_k X_k\right)\right) = \exp\left(i\langle \bar{\alpha}, \bar{m} \rangle - \frac{1}{2}\langle \bar{\alpha}, \Gamma \bar{\alpha} \rangle\right).$$

On a dans ce cas $m_k = \mathbf{E}(X_k)$ et $\Gamma_{jk} = \Gamma_{kj} = \mathbf{cov}(X_j, X_k) = \mathbf{E}(X_j X_k) - \mathbf{E}(X_j)\mathbf{E}(X_k)$. Γ est la **matrice de covariance** du vecteur $\bar{X}$.

Remarque : Rappelons qu'une matrice M est positive si pour tout vecteur v , le produit scalaire $\langle v, Mv \rangle$ est positif. Elle est définie positive si ce même produit scalaire est strictement positif pour tout vecteur v non nul. Cela n'a pas de rapport direct avec la positivité des coefficients de la matrice !

La matrice de covariance et le vecteur des moyennes permettent donc de caractériser complètement la loi d'un vecteur gaussien. En particulier, si un couple (X, Y) est gaussien, ses composantes sont indépendantes si et seulement si la covariance $\text{cov}(X, Y)$ est nulle (c'est-à-dire si sa matrice de covariance est diagonale). C'est bien le seul cas que vous rencontrerez où vous pourrez conclure que deux variables aléatoires sont indépendantes en regardant leur covariance ! Évidemment, c'est un artefact, puisqu'il faut également savoir que le couple est gaussien.

Exercice : Montrez qu'une matrice de covariance est toujours une matrice symétrique et positive (au sens où $\langle v, Mv \rangle \geq 0$ pour tout vecteur v).

Du côté des densités : rappelons que la densité d'une loi de Gauss sur $\mathbb{R}$ de moyenne m et de variance $\sigma^2 > 0$ est $f(x) = \exp(-(x - m)^2/(2\sigma^2))/\sqrt{2\pi\sigma^2}$, et que la primitive de cette fonction ne s'exprime pas de façon simple.

Pour un vecteur aléatoire gaussien, la densité, si elle existe, s'exprime à l'aide de la matrice de covariance et du vecteur des moyennes :

Proposition 3.3 *On considère un vecteur gaussien $\bar{X} = (X_1, \dots, X_n)$ sur $\mathbb{R}^n$ de moyenne $\bar{m} = (m_1, \dots, m_n)$ et de matrice de covariance $\Gamma = (\Gamma_{jk})_{1 \leq j, k \leq n}$.*

$\bar{X}$ admet une densité par rapport à la mesure de Lebesgue sur $\mathbb{R}^n$ si et seulement si Γ est une matrice définie positive.

La densité est dans ce cas la fonction de n variables f donnée par

$$f(\bar{x}) = \frac{\exp(-\langle \bar{x} - \bar{m}, \Gamma^{-1}(\bar{x} - \bar{m}) \rangle / 2)}{(2\pi)^{n/2} \sqrt{\det(\Gamma)}}.$$

Juste pour vérifier la formule de la densité, supposons que $\bar{X}$ est un vecteur gaussien formé de variables aléatoires indépendantes centrées (pour alléger l'écriture !) et de variances non nulles σ_k^2 . Dans ce cas la densité de chaque X_k est $\exp(-x_k^2/(2\sigma_k^2))/\sqrt{2\pi\sigma_k^2}$ et, puisque les variables aléatoires sont indépendantes, la densité du vecteur $\bar{X}$ est le produit des densités :

$$f(\bar{x}) = \frac{\exp(-\sum_k x_k^2/(2\sigma_k^2))}{(2\pi)^{n/2} \prod_k \sigma_k}.$$

Les composantes de $\bar{X}$ étant indépendantes, la matrice Γ des covariances de $\bar{X}$ est diagonale, et on retrouve bien l'expression de la proposition 3.3.

Définition 3.4 *Une famille $(X_i)_{i \in I}$ de variables aléatoires est dite gaussienne si pour tous $n \in \mathbb{N}^*$ et $(i_1, \dots, i_n) \in I^n$, le vecteur $(X_{i_1}, \dots, X_{i_n})$ est gaussien.*

Cette définition s'applique bien entendu aux processus.

Proposition 3.5 *Soit $\bar{X} = (X_n)_{n \in \mathbb{N}}$ une famille gaussienne de moyenne $\bar{m}$ et de covariance Γ . On note*

$$H = \{ \langle \bar{\alpha}, \bar{X} \rangle, \alpha_k \in \mathbb{R} \text{ tels que } \langle \bar{\alpha}, \bar{m} \rangle < \infty \text{ et } \langle \bar{\alpha}, \Gamma \bar{\alpha} \rangle < \infty \}$$

l'espace hilbertien engendré par les $(X_n)_{n \geq 0}$.

Alors toute variable $Y \in H$ est gaussienne et plus généralement, toute suite (finie ou non) $(Y_n)_n$ de H est une famille gaussienne.

1.2 Conditionnement

Calculer une espérance conditionnelle est rarement une opération simple. Le cas des vecteurs gaussiens est ici particulier : conditionner revient à projeter orthogonalement dans $\mathcal{L}^2$.

Proposition 3.6 *Soit $(X, Y_1, \dots, Y_n)$ un vecteur gaussien centré, alors l'espérance conditionnelle $\mathbf{E}(X|Y_1, \dots, Y_n)$ est presque sûrement égale à la projection orthogonale dans $\mathcal{L}^2(\mathbf{P})$ de X sur l'espace vectoriel engendré par $\{Y_1, \dots, Y_n\}$.*

Remarque : Vérifier que, sans aucune hypothèse de normalité, si $Z = \mathbf{E}(X|Y)$ alors $\mathbf{E}((X-Z)Y) = 0$, c'est-à-dire que $X - Z$ est orthogonal à Y .

En pratique, pour calculer l'espérance conditionnelle d'une variable gaussienne X , on écrit $Z = \mathbf{E}(X|Y_1, \dots, Y_n)$ sous la forme d'une combinaison linéaire des (Y_i) , dont les coefficients sont déterminés les conditions d'orthogonalité :

$$Z = \sum_{i=1}^n \alpha_i Y_i$$

$$\forall j, \mathbf{E}((X - Z)Y_j) = 0.$$

Les valeurs des espérances $\mathbf{E}(XY_i)$ et $\mathbf{E}(Y_i Y_j)$ sont fournies par les termes de la matrice de covariance, puisque les variables aléatoires sont centrées.

Remarque : Si le vecteur n'est pas centré, on commence par le centrer en posant : $X^0 = X - \mathbf{E}(X)$ et $Y_i^0 = Y_i - \mathbf{E}(Y_i)$. Si le vecteur $(X, Y_1, \dots, Y_n)$ est gaussien, le vecteur $(X^0, Y_1^0, \dots, Y_n^0)$ l'est aussi et la matrice de covariance est conservée. De plus, $\mathbf{E}(X|Y_1, \dots, Y_n) = \mathbf{E}(X) + \mathbf{E}(X^0|Y_1^0, \dots, Y_n^0)$.

On peut également parler de conditionnement d'un vecteur par rapport à un autre vecteur : dans ce cas on a simplement $\mathbf{E}((X_1, \dots, X_m)|\bar{Y}) = (\mathbf{E}(X_1|\bar{Y}), \dots, \mathbf{E}(X_m|\bar{Y}))$.

Définition 3.7 *La loi conditionnelle de $(X_1, \dots, X_m)$ sachant l'événement $\{(Y_1, \dots, Y_n) = (y_1, \dots, y_n)\}$ est la famille de mesures de probabilité $\mu_{y_1, \dots, y_n}$ sur $\mathbb{R}^m$ telles que, pour tout $A \in \mathcal{B}(\mathbb{R}^m)$ et tout $B \in \mathcal{B}(\mathbb{R}^n)$*

$$\mathbf{P}((X_1, \dots, X_m) \in A, (Y_1, \dots, Y_n) \in B) = \int_B \mu_{y_1, \dots, y_n}(A) d(Y_1, \dots, Y_n)(\mathbf{P}).$$

Proposition 3.8 *Soit $(X_1, \dots, X_m, Y_1, \dots, Y_n)$ un vecteur gaussien centré. On suppose que, pour tout i entre 1 et m ,*

$$\mathbf{E}(X_i|Y_1, \dots, Y_m) = \sum_{j=1}^n \alpha_{ij} Y_j.$$

La loi conditionnelle de $(X_1, \dots, X_m)$ sachant $(Y_1, \dots, Y_n) = (y_1, \dots, y_n)$ coïncide alors avec la loi du vecteur gaussien dont les m composantes sont

$$X_i - \sum_{j=1}^n \alpha_{ij} Y_j + \sum_{j=1}^n \alpha_{ij} y_j.$$

Exercice : Lorsque le vecteur $(X_1, \dots, X_m, Y_1, \dots, Y_n)$, non nécessairement gaussien, admet une densité que l'on notera $f_{\bar{X}, \bar{Y}}$ sur $\mathbb{R}^{m+n}$, vérifiez que la loi conditionnelle de $\bar{X}$ sachant $(Y_1, \dots, Y_n) = (y_1, \dots, y_n)$ admet pour densité

$$(x_1, \dots, x_m) \rightarrow \frac{f_{\bar{X}, \bar{Y}}(x_1, \dots, x_m, y_1, \dots, y_n)}{f_{\bar{Y}}(y_1, \dots, y_n)},$$

où $f_{\bar{Y}}$ est la densité de $\bar{Y}$. Vérifiez (en commençant par la cas $n = m = 1$) que, lorsque $(\bar{X}, \bar{Y})$ est gaussien centré, la loi obtenue dans la proposition 3.8 est bien la loi conditionnelle cherchée.

1.3 Convergence des séries gaussiennes

Proposition 3.9 Soit (X_n) une suite de variables aléatoires indépendantes, centrées et de loi de Gauss. Les propriétés suivantes sont équivalentes :

1. $\sum X_n$ converge p.s.
2. $\sum X_n$ converge en probabilité.
3. $\sum X_n$ converge dans $\mathcal{L}^2$.
4. $\sum X_n$ converge en loi.

La proposition précédente intervient en particulier dans la construction du mouvement brownien.

Preuve : Les implications $1 \Rightarrow 2$, $3 \Rightarrow 2$ et $1 \Rightarrow 4$ sont toujours vérifiées (indépendamment du caractère gaussien de la suite).

Montrons $2 \Rightarrow 1$. On note $S_{N,p} = \sum_N^p X_i$.

On fixe deux entiers $N \leq M$ et un réel $\varepsilon > 0$ et on note

$$\tau = \inf\{p \leq M, |S_{N,p}| > \varepsilon\} \text{ où } \inf\{\emptyset\} = +\infty.$$

On a

$$\mathbf{P}\left(\sup_{N \leq p \leq M} |S_{N,p}| > \varepsilon\right) = \sum_{p=N}^M \mathbf{P}(\tau = p, |S_{N,p}| > \varepsilon).$$

Or

$$S_{N,p} = \frac{1}{2}(S_{N,M} + (X_n + \cdots + X_p) - (X_{p+1} + \cdots + X_M))$$

donc

$$\begin{aligned} \mathbf{P}(\tau = p, |S_{N,p}| > \varepsilon) &\leq \mathbf{P}(\tau = p, |S_{N,M}| > \varepsilon) \\ &\quad + \mathbf{P}(\tau = p, |(X_n + \cdots + X_p) - (X_{p+1} + \cdots + X_M)| > \varepsilon) \\ &= 2\mathbf{P}(\tau = p, |S_{N,M}| > \varepsilon). \end{aligned}$$

D'où $\mathbf{P}(\sup_{N \leq p \leq M} |S_{N,p}| > \varepsilon) \leq 2\mathbf{P}(|S_{N,M}| > \varepsilon)$.

et

$$\mathbf{P}\left(\sup_{N \leq p, q \leq M} |S_{N,p} - S_{N,q}| > \varepsilon\right) \leq 2\mathbf{P}(|S_{N,M}| > \varepsilon/2).$$

Notons S la limite (en probabilité) de $S_{1,N}$. On a

$$\mathbf{P}\left(\sup_{N \leq p, q \leq M} |S_{N,p} - S_{N,q}| > \varepsilon\right) \leq 2\mathbf{P}(|S_{1,M} - S| > \varepsilon/4) + 2\mathbf{P}(|S_{1,N-1} - S| > \varepsilon/4).$$

Lorsque M tend vers $+\infty$, il vient

$$\mathbf{P}\left(\sup_{p, q \geq N} |S_{N,p} - S_{N,q}| > \varepsilon\right) \leq 2\mathbf{P}(|S_{1,N-1} - S| > \varepsilon/4)$$

ce qui tend vers 0 lorsque N tend vers $+\infty$. La série $S_{1,N}$ est donc de Cauchy uniformément en probabilité : c'est alors un exercice classique de montrer qu'elle converge presque sûrement.

$4 \Rightarrow 3$: Par hypothèse il existe une mesure de probabilité μ telle que, pour tout α ,

$$\int_{\mathbb{R}} e^{i\alpha x} d\mu(x) = \lim_{n \rightarrow \infty} e^{-\alpha \sigma_n^2/2}$$

où $\sigma_n^2 = \sum_{i=1}^n \mathbf{E}(X_i^2)$. Donc σ_n^2 converge, vers une limite $\sigma^2 \in \mathbb{R}^+$ et μ est la loi de Gauss centrée et de variance σ^2 .

De plus, pour $n < m$, $\mathbf{E}((S_{1,n} - S_{1,m})^2) = \sum_{i=n+1}^m \mathbf{E}(X_i^2)$. Puisque la série de terme général $\mathbf{E}(X_i^2)$ est convergente, $(S_{1,n})$ est une suite de Cauchy dans $\mathcal{L}^2$ qui est un espace de Hilbert. On conclut donc que $(S_{1,n})$ converge dans $\mathcal{L}^2$.

2 Mouvement brownien

2.1 Introduction

Rappelons que $[x]$ désigne la partie entière d'un nombre réel x (ou plus grand nombre entier inférieur ou égal à x).

Considérons une suite $(X_i)_i$ de variables aléatoires indépendantes et identiquement distribuées de loi donnée par $\mathbf{P}(X_i = 1) = \mathbf{P}(X_i = -1) = 1/2$. Remarquons que les X_i sont centrées et de variance 1 et notons $S_0 = 0$ et $S_n = S_{n-1} + X_n$, la suite des sommes partielles. Il est bien connu que S_n/n tend presque sûrement vers 0 (cela découle de la loi forte des grands nombres) et que $S_n/\sqrt{n}$ tend en loi vers une loi de Gauss centrée et de variance 1 (c'est le théorème central limite). Toujours avec le théorème central limite, on peut voir que pour tout $t > 0$, $S_{[nt]}/\sqrt{n}$ tend en loi vers une loi de Gauss centrée et de variance t . En utilisant une généralisation du théorème central limite, on peut également vérifier que si $0 < t_1 < \dots < t_k$ sont des instants fixés, le vecteur $(S_{[nt_1]}/\sqrt{n}, \dots, S_{[nt_k]}/\sqrt{n})$ tend en loi, lorsque $n \rightarrow \infty$, vers un vecteur gaussien $(V_1, \dots, V_k)$ centré dont les covariances sont $\text{cov}(V_j, V_k) = \min(t_j, t_k)$.

Théorème 3.10 (Donsker) *Soit (X_n) une suite de variables aléatoires de carrés intégrables, centrées, indépendantes et identiquement distribuées et S_n la suite de ses sommes partielles. On note $\sigma^2 = \text{var}(X_n)$. Alors, pour tout entier $k \geq 1$ et tous réels $0 < t_1 < \dots < t_k$, le vecteur $(S_{[nt_1]}/\sqrt{n}, \dots, S_{[nt_k]}/\sqrt{n})$ tend en loi vers un vecteur gaussien centré $(V_1, \dots, V_k)$ de covariances $\text{cov}(V_j, V_k) = \sigma^2 \min(t_j, t_k)$.*

Définition 3.11 *Le mouvement brownien réel standard (ou mouvement brownien tout court) est un processus $(B_t)_{t \geq 0}$ tel que*

1. $B_0 = 0$ (ou éventuellement $B_0 = x$ fixé)
2. Pour tous réels positifs t et s , $B_t - B_s$ suit une loi de Gauss centrée et de variance $|t - s|$.
3. Pour tous $0 \leq t_0 \leq t_1 \leq \dots \leq t_n$, les variables aléatoires $B_{t_1} - B_{t_0}, B_{t_2} - B_{t_1}, \dots, B_{t_n} - B_{t_{n-1}}$ sont indépendantes.

Remarque : Cette définition implique que le mouvement brownien est un processus gaussien.

On notera de façon indifférenciée $B_t, B_t(\omega), B(t), B(t, \omega)$.

Définition 3.12

- Lorsqu'un processus vérifie le point 2 ci-dessus, on dit qu'il est à accroissements indépendants.
- Si (X_t) est un processus tel que, pour tous $t \geq s$, la loi de $X_t - X_s$ est la même que celle de $X_{t-s} - X_0$, on dit qu'il est stationnaire.

2.2 Construction

Le but de la construction du mouvement brownien est de vérifier qu'il existe un processus (B_t) satisfaisant la définition 3.11.

Fixons des réels positifs $t_1, \dots, t_k$ et notons Γ la matrice $n \times n$ définie par $\Gamma_{ij} = \min(t_i, t_j)$. Nous allons commencer par vérifier que Γ est une matrice positive. Remarquons tout d'abord que, si s et t sont des réels positifs,

$$\min(s, t) = \int_0^{+\infty} \mathbf{1}_{[0,t]}(u) \mathbf{1}_{[0,s]}(u) du.$$

Prenons un vecteur $\bar{\alpha} = (\alpha_1, \dots, \alpha_k)$ de $\mathbb{R}^k$.

$$\begin{aligned} \langle \bar{\alpha}, \Gamma \bar{\alpha} \rangle &= \sum_{i,j} \alpha_i \alpha_j \min(t_i, t_j) \\ &= \int_0^{+\infty} \sum_{i,j} \alpha_i \alpha_j \mathbf{1}_{[0,t_i]}(u) \mathbf{1}_{[0,t_j]}(u) du \\ &= \int_0^{+\infty} \left(\sum_i \alpha_i \mathbf{1}_{[0,t_i]}(u) \right)^2 du \end{aligned}$$

La matrice Γ est donc bien positive.

Soit maintenant (e_n) une base hilbertienne de $\mathcal{L}^2([0, 1], du)$ et notons

$$\mathbf{1}_{[0,t]}(u) = \sum_n a_n(t) e_n(u).$$

On a alors

$$\min(s, t) = \langle \mathbf{1}_{[0,s]}, \mathbf{1}_{[0,t]} \rangle = \sum_n a_n(s) a_n(t).$$

Considérons maintenant une suite (X_n) de variables aléatoires indépendantes et identiquement distribuées de loi $\mathcal{N}(0, 1)$ et notons, pour tout $t \in [0, 1]$,

$$B_t = \sum_n a_n(t) X_n.$$

Pour tout $t \in [0, 1]$, la suite définissant B_t converge dans $\mathcal{L}^2(\mathbf{P})$ puisque

$$\sum_n a_n^2(t) = \langle \mathbf{1}_{[0,t]}, \mathbf{1}_{[0,t]} \rangle = t < \infty.$$

Le processus $(B_t)_t$ obtenu est gaussien centré, vérifie $B_0 = 0$ et on a

$$\begin{aligned} \mathbf{E}(B_t B_s) &= \sum_n a_n(t) a_n(s) \mathbf{E}(X_n^2) + 2 \sum_{n < m} a_n(t) a_m(s) \mathbf{E}(X_n X_m) \\ &= \sum_n a_n(t) a_n(s) = \min(t, s). \end{aligned}$$

De plus, pour $0 \leq s \leq t \leq 1$,

$$B_t - B_s = \sum (a_n(t) - a_n(s)) X_n$$

suit une loi de Gauss centrée et de variance

$$\begin{aligned} \sum (a_n(t) - a_n(s))^2 &= \sum a_n^2(t) + \sum a_n^2(s) - 2 \sum a_n(t) a_n(s) \\ &= t + s - 2s = t - s. \end{aligned}$$

Il reste à vérifier que le processus est à accroissements indépendants. Comme c'est un processus gaussien, il suffit de vérifier que la covariance de $B_t - B_s$ et de B_u est nulle pour tous $0 \leq u \leq s < t \leq 1$:

$$\begin{aligned} \text{cov}(B_t - B_s, B_u) &= \mathbf{E}((B_t - B_s)B_u) \\ &= \mathbf{E}(B_t B_u) - \mathbf{E}(B_s B_u) \\ &= \min(t, u) - \min(s, u) = 0 \end{aligned}$$

Le processus $(B_t)_{t \in [0,1]}$ ainsi construit vérifie bien la définition 3.11.

Construisons maintenant le mouvement brownien indexé par $\mathbb{R}^+$: posons, pour tout $t \geq 0$, $Z_t = (1+t)B_{(1+t)^{-1}} - B_1$. Le processus $(Z_t)_{t \geq 0}$ est gaussien centré, $Z_0 = 0$ et pour $t \geq s \geq 0$,

$$\begin{aligned} \mathbf{E}(Z_t Z_s) &= (1+t)(1+s)\mathbf{E}(B_{(1+t)^{-1}} B_{(1+s)^{-1}}) + \mathbf{E}(B_1^2) - (1+t)\mathbf{E}(B_1 B_{(1+t)^{-1}}) \\ &\quad - (1+s)\mathbf{E}(B_1 B_{(1+s)^{-1}}) \\ &= (1+t)(1+s)(\max(1+t, 1+s))^{-1} + 1 - 1 - 1 \\ &= 1 + s - 1 = s \end{aligned}$$

Il est ensuite très facile de vérifier que $Z_t - Z_s$ suit une loi de Gauss centrée et de variance $|t - s|$ et que les accroissements de (Z_t) sont indépendants.

2.3 Propriétés

Proposition 3.13 *(B_t) est un mouvement brownien si et seulement si c'est un processus gaussien centré de covariance $\text{cov}(B_t, B_s) = \min(t, s)$.*

Preuve : en exercice (c'est facile et c'est en fait déjà fait dans la section précédente !)

Proposition 3.14 (scaling) *Soit (B_t) un mouvement brownien réel standard et $c > 0$ un réel. On pose $X_t = cB_{t/c^2}$. Alors $(X_t)_t$ est un mouvement brownien réel standard.*

Preuve : le processus (X_t) est évidemment gaussien et il suffit de calculer les covariances pour montrer que c'est un mouvement brownien.

Proposition 3.15 (inversion temporelle) *Soit (B_t) un mouvement brownien réel standard. On pose pour tout $t > 0$, $Z_t = tB_{1/t}$. Alors (Z_t) est un mouvement brownien réel standard.*

Preuve : Le processus (Z_t) est gaussien, et comme ci-dessus, il suffit de calculer les covariances.

Proposition 3.16 1. *Un mouvement brownien (B_t) est un processus presque sûrement continu, c'est-à-dire qu'il existe un ensemble négligeable N tel que, pour tout $\omega \notin N$, la trajectoire $t \rightarrow B_t(\omega)$ est continue.*

2. *Les trajectoires du mouvement brownien sont nulle part dérivables : il existe un ensemble négligeable N' tel que, pour tout $\omega \notin N'$, la fonction $t \rightarrow B_t(\omega)$ est nulle part dérivable.*

La continuité n'est pas triviale : en effet, soit (X_t) un processus à accroissement indépendant et stationnaire et notons $\phi_t(\alpha) = \mathbf{E}(\exp(i\alpha X_t))$. Pour tous réels positifs t et s , on a

$$\mathbf{E}(e^{i\alpha X_{t+s}}) = \mathbf{E}\left((e^{i\alpha(X_{t+s}-X_s)} e^{i\alpha(X_s)})\right) = \mathbf{E}\left(e^{i\alpha(X_{t+s}-X_s)}\right) \mathbf{E}\left(e^{i\alpha(X_s)}\right)$$

c'est-à-dire, avec la stationnarité,

$$\phi_{t+s}(\alpha) = \phi_t(\alpha)\phi_s(\alpha)$$

C'est alors un exercice classique de montrer (moyennant une hypothèse de continuité ou même simplement de mesurabilité) que ϕ est de la forme $\phi_t(\alpha) = \exp(-c(\alpha)t)$. Pour $c(\alpha) = \alpha^2/2$, on retrouve le mouvement brownien, et pour $c(\alpha) = |\alpha|$, le processus obtenu est le *processus de Cauchy*. Le mouvement brownien est continu, alors que l'ensemble des discontinuités du processus de Cauchy est presque sûrement dense dans $\mathbb{R}^+$.

Preuve : de la continuité. Il est facile de voir que $t \rightarrow B_t$ est continu dans $\mathcal{L}^2(\mathbf{P})$ pour tout t fixé, ce qui assure la continuité presque sûre de $t \rightarrow B_t$ en tout t_0 fixé, c'est-à-dire

$$\mathbf{P}(B_t \text{ est continu en } t_0) = 1.$$

Il reste à voir qu'il existe un négligeable n tel que pour tout $\omega \notin N$, $t \rightarrow B_t$ est continu sur $\mathbb{R}^+$.

Le lemme technique suivant, base de la continuité, est admis,

Lemme 3.17 *Soit (X_t) , $t \in [0, 1]$, un processus continu en probabilité. Si D est une partie dénombrable dense de $[0, 1]$, on pose*

$$U_\varepsilon = \sup\{|X_t - X_s|; s, t \in D, |t - s| \leq \varepsilon\}.$$

On suppose que U_ε tend presque sûrement vers 0 lorsque ε tend vers 0. Alors il existe une version $\tilde{X}$ de X dont toutes les trajectoires sont continues.

Définition 3.18 $\tilde{X}$ est une **version** de X si, pour tout t fixé, $\mathbf{P}(\tilde{X}_t = X_t) = 1$.

Montrons que B_t est continu en probabilité : soit $\varepsilon > 0$.

$$\mathbf{P}(|B_t - B_{t_0}| \geq \varepsilon) = \mathbf{P}(|B_{|t-t_0|}| \geq \varepsilon) \leq \frac{1}{\varepsilon^2} \mathbf{E}(B_{|t-t_0|}^2) = \frac{|t - t_0|}{\varepsilon^2}$$

donc B_t est continu en probabilité en tout $t_0 \geq 0$ fixé.

Notons

$$D = \{k2^{-n}, k = 0, \dots, 2^n, n \in \mathbb{N}^*\}$$

et posons

$$U_n = \sup_{s, t \in D, |t-s| \leq 2^{-n}} |B_t - B_s|.$$

La suite U_n étant décroissante, sa convergence en probabilité vers 0 équivaut à sa convergence presque sûre vers 0. Or

$$\{U_n > \varepsilon\} \subset \bigcup_{k=0}^{2^n-1} \left\{ \sup_{t \in D \cap [k2^{-n}, (k+1)2^{-n}]} |B_t - B_{k2^{-n}}| > \varepsilon/3 \right\}$$

donc

$$\mathbf{P}(U_n > \varepsilon) \leq \sum_{k=0}^{2^n-1} \mathbf{P} \left(\sup_{t \in D \cap [k2^{-n}, (k+1)2^{-n}]} |B_t - B_{k2^{-n}}| > \varepsilon/3 \right)$$

ce qui donne en utilisant la propriété de scaling 3.14

$$\mathbf{P}(U_n > \varepsilon) \leq 2^n \mathbf{P} \left(\sup_{t \in D \cap [0, 2^{-n}]} |B_t| > \varepsilon/3 \right) = 2^n \mathbf{P} \left(\sup_{s \in D \cap [0, 1]} |B_s| \geq 2^{n/2} \varepsilon/3 \right)$$

Le lemme suivant permet alors de conclure :

Lemme 3.19 *Soient $0 \leq t_1 \dots \leq t_n \leq 1$ et $\alpha \geq 0$. Alors*

$$(i) \quad \mathbf{P} \left(\sup_i B_{t_i} \geq \alpha \right) \leq 2\mathbf{P}(B_1 \geq \alpha)$$

$$\text{et } (ii) \quad \mathbf{P} \left(\sup_i |B_{t_i}| \geq \alpha \right) \leq 2\mathbf{P}(|B_1| \geq \alpha)$$

Terminons la preuve de la continuité en utilisant une version de l'inégalité de Markov :

$$\mathbf{P}(U_n > \varepsilon) \leq 2^{n+1} \mathbf{P}(|B_1| \geq 2^{n/2} \varepsilon / 3) \leq 2^{n+1} \frac{\mathbf{E}(B_1^4)}{(2^{n/2} \varepsilon / 3)^4} = \frac{3^4}{\varepsilon^4} 2^{-n+1} 3$$

qui tend bien vers 0 lorsque n tend vers $+\infty$, ce qui termine la preuve de la continuité.

Terminons par la preuve du lemme 3.19 :

(i) $\Rightarrow$ (ii) :

$$\begin{aligned} \mathbf{P}\left(\sup_i |B_{t_i}| \geq \alpha\right) &= \mathbf{P}\left(\left\{\sup_i B_{t_i} \geq \alpha\right\} \cup \left\{\sup_i (-B_{t_i}) \geq \alpha\right\}\right) \\ &\leq 2\mathbf{P}\left(\sup_i B_{t_i} \geq \alpha\right) \\ &\leq 4\mathbf{P}(B_1 \geq \alpha) = 2\mathbf{P}(|B_1| \geq \alpha) \end{aligned}$$

(i) : Soit $\tau = \inf\{i, B_{t_i} \geq \alpha\}$.

$$\begin{aligned} \mathbf{P}\left(\sup_i B_{t_i} \geq \alpha\right) &= \mathbf{P}\left(\bigcup_{i=1}^n \{\tau = i\}\right) \\ &= \sum_{i=1}^n \mathbf{P}(\tau = i) \\ &\leq 2 \sum_{i=1}^n \mathbf{P}(\tau = i) \mathbf{P}(B_1 - B_{t_i} \geq 0) \\ &\leq 2 \sum_{i=1}^n \mathbf{P}(\tau = i \text{ et } B_1 - B_{t_i} \geq 0) \\ &\leq 2 \sum_{i=1}^n \mathbf{P}(\tau = i \text{ et } B_1 \geq \alpha) \\ &\leq 2\mathbf{P}(B_1 \geq \alpha). \end{aligned}$$

2.4 Martingale et temps d'arrêt

Proposition 3.20 *Le mouvement brownien $(B_t)_{t \geq 0}$ est une martingale continue, par rapport à la filtration $(\mathcal{F}_t)$ continue à droite et complétée par les négligeables définie par*

$$\mathcal{F}_t = \bigcap_{\varepsilon \downarrow 0} \overline{\{B_s, s \leq t + \varepsilon\}}.$$

Preuve : Il est clair que les $(\mathcal{F}_t)$ forment une filtration complète et continue à droite, et que B_t est $\mathcal{F}_t$ -mesurable. Il faut donc montrer que, pour tous $t \geq s \geq 0$, $\mathbf{E}(B_t | \mathcal{F}_s) = B_s$.

Écrivons $B_t = B_s + (B_t - B_s)$ et notons, pour tout $u \geq 0$, $Z_u = B_{s+u} - B_s$. Le processus $(Z_u)_{u \geq 0}$ est un processus gaussien centré, et la famille formée par les $(B_r)_{r \geq 0}$ et les $(Z_u)_{u \geq 0}$ est gaussienne. Les covariances $\mathbf{cov}(B_r, Z_u)$ sont faciles à calculer et sont toutes nulles pour $r \in [0, s]$ et $u \geq 0$, ce qui prouve que $(Z_u)_{u \geq 0}$ est indépendant de la tribu $\mathcal{F}_s$. On a donc

$$\mathbf{E}(B_t | \mathcal{F}_s) = \mathbf{E}(B_s | \mathcal{F}_s) + \mathbf{E}(Z_{t-s} | \mathcal{F}_s) = B_s.$$

Proposition 3.21 *La variation quadratique $(B_t)_{t \geq 0}$ est égale à t .*

Preuve : Pour prouver cela, il faut montrer que $M_t = B_t^2 - t$ est une martingale. Calculons donc $\mathbf{E}(B_t^2 - t | \mathcal{F}_s)$, pour $t \geq s \geq 0$, en utilisant les mêmes notations que ci-dessus.

$$\begin{aligned} \mathbf{E}(B_t^2 - t | \mathcal{F}_s) &= \mathbf{E}\left(\left((B_t - B_s) + B_s\right)^2 | \mathcal{F}_s\right) - t \\ &= \mathbf{E}\left(Z_{t-s}^2 | \mathcal{F}_s\right) + (B_s^2 | \mathcal{F}_s) + 2\mathbf{E}(B_s Z_{t-s} | \mathcal{F}_s) - t \end{aligned}$$

On a vu ci-dessus que le processus $(Z_u)_u$ est indépendant de la tribu $\mathcal{F}_s$ donc

$$\mathbf{E}\left(Z_{t-s}^2 | \mathcal{F}_s\right) = \mathbf{E}\left(Z_{t-s}^2\right).$$

Par définition du mouvement brownien, $Z_{t-s} = B_t - B_s$ suit une loi de Gauss centrée et de variance $t - s$. On a donc

$$\mathbf{E}(Z_{t-s}^2) = t - s.$$

De plus, B_s est une variable aléatoire $\mathcal{F}_s$ -mesurable donc

$$\mathbf{E}(B_s Z_{t-s} | \mathcal{F}_s) = B_s \mathbf{E}(Z_{t-s} | \mathcal{F}_s),$$

et en utilisant à nouveau l'indépendance de Z_{t-s} et de $\mathcal{F}_s$, il vient

$$\mathbf{E}(Z_{t-s} | \mathcal{F}_s) = \mathbf{E}(Z_{t-s}) = \mathbf{E}(B_t) - \mathbf{E}(B_s) = 0.$$

Finalement, on a bien

$$\mathbf{E}(B_t^2 - t | \mathcal{F}_s) = B_s^2 - s.$$

Il n'est pas très compliqué de vérifier que le processus $(Z_u)_{u \geq 0}$ qui apparaît ci-dessus est un mouvement brownien indépendant de $\mathcal{F}_s$: la preuve est essentiellement basée sur le caractère gaussien de ce processus et sur le calcul des covariances nécessaires (voir la preuve de la proposition 3.20). Ce résultat s'appelle la propriété de Markov faible. Elle est encore vraie lorsque l'instant s est remplacé par un temps d'arrêt. Elle ne l'est évidemment plus si on remplace s par une variable aléatoire quelconque. Avant d'énoncer ces deux propriétés, quelques notations :

- On écrira $\mathbf{E}_x(f(B_t))$ ou $\mathbf{P}_x(B_t \in A)$ pour signifier que le mouvement brownien $(B_t)_t$ est issu de x ; autrement dit,

$$\mathbf{E}_x(f(B_t)) = \mathbf{E}_0(f(B_t + x)) \quad \text{et} \quad \mathbf{P}_x(B_t \in A) = \mathbf{P}_0(x + B_t \in A).$$

Par exemple, $\mathbf{E}_x(B_t) = x$ et $\mathbf{E}_x(B_t^2) = \mathbf{E}_0((x + B_t)^2) = x^2 + t$.

- Les expressions $\mathbf{E}_{B_s}(f(B_t))$ ou $\mathbf{P}_{B_s}(B_t \in A)$ signifient que l'on calcule les expressions comme ci-dessus, et que l'on remplace à la fin x par B_s . Autrement dit, on se donne un mouvement brownien $(\widetilde{B}_u)_{u \geq 0}$ indépendant de B_s (et de son passé) et on intègre uniquement par rapport à l'aléatoire de $(\widetilde{B}_u)_{u \geq 0}$, et pas par rapport à l'aléatoire de B_s . Par exemple, $\mathbf{E}_{B_s}(B_t) = B_s$ et $\mathbf{E}_{B_s}(B_t^2) = B_s^2 + t$.
- Pour finir, si μ est une mesure de probabilité, on définit $\mathbf{E}_\mu$ (et similairement $\mathbf{P}_\mu$) de la façon suivante :

$$\mathbf{E}_\mu(f(B_t)) = \int_{\mathbb{R}} \mathbf{E}_x(f(B_t)) \, d\mu(x).$$

Proposition 3.22 (Propriétés de Markov)

1. **Markov faible.** Soit $s \geq 0$ fixé. Alors le processus $(B_{s+u} - B_s)_{u \geq 0}$ est un mouvement brownien indépendant de $(B_v)_{0 \leq v \leq s}$, et donc de la tribu $\mathcal{F}_s$. Plus concrètement, pour toute fonction $f : \mathbb{R} \rightarrow \mathbb{R}$ borélienne bornée,

$$\mathbf{E}_x(f(B_{s+u}) | \mathcal{F}_s) = \mathbf{E}_{B_s}(f(B_u))$$

et pour toute fonction $g : \mathbb{R}^2 \rightarrow \mathbb{R}$ borélienne bornée,

$$\mathbf{E}_x(g(B_{s+u} - B_s, B_s) | \mathcal{F}_s) = \int_{\mathbb{R}} g(x, B_s) e^{-x^2/(2u)} \frac{dx}{\sqrt{2\pi u}}.$$

2. **Markov fort.** Soit τ un temps d'arrêt presque sûrement fini. Alors le processus $(B_{\tau+u} - B_\tau)_{u \geq 0}$ est un mouvement brownien indépendant de la tribu $\mathcal{F}_\tau$. Plus concrètement, pour toute fonction $f : \mathbb{R} \rightarrow \mathbb{R}$ borélienne bornée,

$$\mathbf{E}_x (f(B_{\tau+u}) | \mathcal{F}_\tau) = \mathbf{E}_{B_\tau} (f(B_u))$$

et pour toute fonction $g : \mathbb{R}^2 \rightarrow \mathbb{R}$ borélienne bornée,

$$\mathbf{E}_x (g(B_{\tau+u} - B_\tau, B_\tau) | \mathcal{F}_\tau) = \int_{\mathbb{R}} g(x, B_\tau) e^{-x^2/(2u)} \frac{dx}{\sqrt{2\pi u}}.$$

2.5 Simulation

Tout dépend de ce dont on a besoin :

- soit de (B_t) pour t grand, pour étudier le comportement asymptotique d'un phénomène. On posera alors $B_n = X_1 + \dots + X_n$ où les (X_i) sont des variables aléatoires indépendantes et identiquement distribuées de loi normale centrée réduite.
- soit de (B_t) , $t \in [0, N]$, en des instants de plus en plus proches. On commence par simuler (B_k) , $k = 1 \dots, N$ comme ci-dessus. Puis on simule $B(k + 1/2)$ en remarquant que

$$B\left(k + \frac{1}{2}\right) = \frac{B(k) + B(k+1)}{2} + \frac{Y_k^{(1)}}{2}$$

où les $Y_k^{(1)}$ sont indépendantes et de loi de Gauss centrée réduite.

On remarque ensuite que

$$B\left(\frac{n+1}{2} + \frac{1}{4}\right) = \frac{1}{2} \left(B\left(\frac{n+1}{2}\right) + B\left(\frac{n+2}{2}\right) \right) + \frac{1}{2\sqrt{2}} Y_n^{(2)}$$

et plus généralement,

$$B\left(t + \frac{\tau}{2}\right) = \frac{1}{2} (B(t) + B(t + \tau)) + \frac{\sqrt{\tau}}{2} Y$$

où Y suit une loi de Gauss centrée réduite, indépendante de toutes les variables précédemment utilisées.

La vérification de ces méthodes de simulation est basée sur le caractère gaussien du processus et sur le calcul des covariances.

3 Construction de l'intégrale stochastique, formule de Itô

Le but est ici de donner un sens à l'expression

$$\int_0^t H_s dB_s$$

ou

$$\int_0^t H_s dM_s,$$

pour certains processus H et une martingale continue M . Les intégrales de ce type sont appelées "intégrales stochastiques", c'est-à-dire, intégrales par rapport à des processus stochastiques. Le problème essentiel est que ni le mouvement brownien, ni les martingales continues ne sont des processus à variations bornées. Ces quantités ne peuvent donc pas se définir au sens classique de l'intégrale de Lebesgue-Stieljes.

Rappels :

1. Si f est une fonction continue et dérivable, on n'a pas nécessairement

$$f(b) - f(a) = \int_a^b f'(s) \, ds.$$

Il suffit pour que cette égalité soit vérifiée que f' soit Riemann-intégrable sur $[a, b]$. Si f' est intégrable sur $\mathbb{R}$, on définit une mesure (signée) en posant $\mu(A) = \int_{\mathbb{R}} \mathbf{1}_A(s) f'(s) \, ds$.

2. Les trajectoires du mouvement brownien ne sont pas des fonctions dérivables. Elles ne sont pas non plus à variations bornées, ce qui signifie que

$$\mu(A) = \int_{\mathbb{R}^+} \mathbf{1}_A(s) \, dB_s$$

ne définit pas une mesure au sens habituel du terme.

Néanmoins, il est possible de définir une intégrale par rapport au mouvement brownien ou par rapport à une martingale continue, mais dans $\mathcal{L}^2(\mathbf{P})$. Nous verrons que l'on construit ainsi un processus continu, qui est une martingale.

Notation : Pour tout couple de réels x et y , on note $x \wedge y = \min(x, y)$.

3.1 Intégration par rapport au mouvement brownien

L'intégrale stochastique se définit pour les processus $H_\cdot$ qui sont $\mathcal{F}_\cdot$ -adaptés (cf définition 2.13), où $\mathcal{F}_t$ est la filtration continue à droite associée au mouvement brownien. Si on veut définir une fonctionnelle

$$H_\cdot \rightarrow \int_0^t H_s \, dB_s$$

qui ressemble à une intégrale, l'une des propriétés "inévitables" est la suivante : si H est de la forme

$$H_s(\omega) = K(\omega) \mathbf{1}_{[a, b[}(s),$$

où a et b sont des constantes fixées et K une variable aléatoire, on doit avoir

$$\int_0^t H_s \, dB_s = K(B_{t \wedge b} - B_{t \wedge a}).$$

On remarque facilement que H est adapté si et seulement si K est $\mathcal{F}_a$ -mesurable.

Suivant le même raisonnement, la fonctionnelle "intégration stochastique" doit être additive, c'est-à-dire que si H est de la forme

$$H_s(\omega) = \sum_{i=1}^n K_i(\omega) \mathbf{1}_{[a_i, b_i[}(s) \tag{3.1}$$

où les K_i sont $\mathcal{F}_{a_i}$ -mesurables et où $0 \leq a_1 \leq b_1 \leq a_2 \leq \dots$ sont des constantes, on doit avoir

$$\int_0^t H_s \, dB_s = \sum_{i=1}^n K_i(B_{t \wedge b_i} - B_{t \wedge a_i}).$$

Les processus de la forme (3.1) sont appelés des **processus simples**, et ils engendrent les processus $\mathcal{F}_\cdot$ -adaptés. Il nous reste donc à voir que la limite au sens $\mathcal{L}^2$ existe, moyennant une condition d'intégrabilité sur H , pour terminer la construction de l'intégrale stochastique.

Commençons par fixer un processus simple $H_s(\omega) = \sum_i K_i(\omega) \mathbf{1}_{[a_i, b_i[}(s)$ vérifiant

$$\mathbf{E} \left(\int_0^\infty H_s^2 ds \right) < \infty$$

et calculons $\mathbf{E}(\int_0^\infty H_s dB_s)^2$.

En utilisant l'expression de l'intégrale stochastique d'un processus simple, on obtient :

$$\mathbf{E} \left(\int_0^\infty H_s dB_s \right)^2 = \sum_{i=1}^n \mathbf{E}(K_i^2 (B_{b_i} - B_{a_i})^2) + 2 \sum_{j < i} \mathbf{E}(K_i K_j (B_{b_i} - B_{a_i})(B_{b_j} - B_{a_j})).$$

Or, en conditionnant par rapport à $\mathcal{F}_{a_i}$, on a :

$$\begin{aligned} \mathbf{E}(K_i^2 (B_{b_i} - B_{a_i})^2) &= \mathbf{E} \left(K_i^2 \mathbf{E}((B_{b_i} - B_{a_i})^2 | \mathcal{F}_{a_i}) \right) \\ &= \mathbf{E}(K_i^2 (b_i - a_i)) \\ &= \mathbf{E} \left(\int_0^\infty K_i^2 \mathbf{1}_{[a_i, b_i[}(s) ds \right) \end{aligned}$$

et

$$\begin{aligned} \mathbf{E}[K_i K_j (B_{b_i} - B_{a_i})(B_{b_j} - B_{a_j})] &= \mathbf{E} [K_i K_j (B_{b_j} - B_{a_j}) \mathbf{E}((B_{b_i} - B_{a_i}) | \mathcal{F}_{a_i})] \\ &= 0. \end{aligned}$$

Finalement, on obtient la proposition suivante :

Proposition 3.23 *Pour tout processus simple H vérifiant $\mathbf{E}(\int_0^\infty H_s^2 ds) < \infty$,*

$$\mathbf{E} \left(\int_0^\infty H_s dB_s \right)^2 = \mathbf{E} \left(\int_0^\infty H_s^2 ds \right).$$

On obtient donc une isométrie entre l'espace $\mathcal{H}$ des processus H , $\mathcal{F}_t$ -adaptés, muni de la norme

$$\|H\|_{\mathcal{H}} = \left(\mathbf{E} \int_0^\infty H_s^2 ds \right)^{1/2}$$

et l'espace des variables aléatoires de la forme $\int_0^\infty H_s dB_s$ muni de la norme $\|X\|_2 = \sqrt{\mathbf{E}(X^2)}$. Cela suffit pour prouver l'existence de l'intégrale stochastique d'un processus H , $\mathcal{F}$ -adapté et vérifiant

$$(*) \quad \mathbf{E} \left(\int_0^\infty H_s^2 ds \right) < \infty,$$

comme limite dans $\mathcal{L}^2(\mathbf{P})$ de l'intégrale stochastique de processus simples qui tendent vers H au sens de la norme $\|\cdot\|_{\mathcal{H}}$. La limite existe et ne dépend pas de la suite de processus simples choisis, en vertu de la complétude de $\mathcal{L}^2(\mathbf{P})$.

Proposition 3.24 *Soit H un processus $\mathcal{F}_t$ -adapté vérifiant $\int_0^t \mathbf{E}(H_s^2) ds < \infty$. On définit un processus N par*

$$N_t = \int_0^t H_s dB_s = \int_{\mathbb{R}^+} H_s \mathbf{1}_{[0, t[}(s) dB_s.$$

N_t est une martingale continue dans $\mathcal{L}^2(\mathbf{P})$, de variation quadratique donnée par

$$\langle N \rangle_t = \int_0^t H_s^2 ds.$$

Preuve : Il est clair que N est bien défini puisque pour tout t fixé, $\mathbf{E} \left(\int_0^t H_s^2 ds \right)$ est finie. De plus, la relation de Challes est vérifiée, c'est-à-dire que pour tous réels $t > u \geq 0$, on a bien

$$N_t - N_u = \int_u^t H_s dB_s.$$

Pour tout $t > u$, on a

$$\mathbf{E}((N_t - N_u)^2) = \mathbf{E} \left(\int_u^t H_s^2 ds \right),$$

et cette quantité tend vers 0 lorsque t tend vers u . De même si t tend vers u par valeurs inférieures, ce qui prouve la continuité dans $\mathcal{L}^2(\mathbf{P})$.

Si H est un processus simple, on vérifie que N est une martingale continue, de variation quadratique

$$\langle N \rangle_t = \int_0^t H_s^2 ds.$$

En effet, écrivons H sous la forme $H_t = \sum K_i \mathbf{1}_{[a_i, b_i[}(t)$, où $a_1 \leq b_1 \leq a_2 \leq b_2 \leq \dots$ et où les K_i sont $\mathcal{F}_{a_i}$ -mesurables. On a, en notant $i(t) = \max\{i, a_i \leq t\}$,

$$N_t = \sum_{i \leq i(t)} K_i (B_{b_i \wedge t} - B_{a_i}).$$

La continuité de N_t est évidente, ainsi que le fait que N_t est $\mathcal{F}_t$ -mesurable. Puis, pour tout i vérifiant $a_i \geq s$, et pour tout $u \geq a_i$, on obtient en utilisant le fait que le mouvement brownien est une martingale :

$$\begin{aligned} \mathbf{E}(K_i(B_u - B_{a_i})|\mathcal{F}_s) &= \mathbf{E}(\mathbf{E}(K_i(B_u - B_{a_i})|\mathcal{F}_{a_i})|\mathcal{F}_s) \\ &= \mathbf{E}(K_i(\mathbf{E}(B_u|\mathcal{F}_{a_i}) - B_{a_i})|\mathcal{F}_s) \\ &= 0. \end{aligned}$$

On en déduit ainsi que (N_t) est une martingale.

En ce qui concerne sa variation quadratique, on procède de façon similaire : on note

$$V_t = \sum_{i \leq i(t)} K_i^2 ((b_i \wedge t) - a_i).$$

On a

$$N_t^2 - V_t = \sum_{i \leq i(t)} K_i^2 ((B_{b_i \wedge t} - B_{a_i})^2 - (b_i \wedge t) + a_i) + 2 \sum_{i < j \leq i(t)} K_i K_j (B_{b_i \wedge t} - B_{a_i})(B_{b_j \wedge t} - B_{a_j}).$$

En utilisant le fait que $(B_t^2 - t)$ est une martingale, on démontre alors facilement que $(N_t^2 - V_t)$ est également une martingale.

On passe alors à la limite pour obtenir ces mêmes résultats pour les processus H adaptés.

Nous allons maintenant définir l'intégrale stochastique par rapport à une martingale continue et de carré intégrable. Nous montrerons notamment que le processus obtenu est une martingale continue.

3.2 Intégrale stochastique par rapport à une martingale continue

On se donne une $\mathcal{F}_t$ -martingale continue (M_t) , de carré intégrable et on suppose que le supréum pour $t \geq 0$ de $\mathbf{E}M_t^2$ est fini.

Nous allons voir que, si H est un processus adapté vérifiant

$$\mathbf{E} \left(\int_0^\infty H_s^2 d\langle M \rangle_s \right) < \infty,$$

il est possible de définir l'intégrale stochastique de H par rapport à M .

Remarques :

1. La variation quadratique $\langle M \rangle_t$ d'une martingale continue M est un processus continu et croissant. La fonction $t \rightarrow \langle M \rangle_t$ est alors la fonction de répartition d'une mesure positive, portée par $\mathbb{R}^+$, ce qui définit de façon non ambiguë l'intégration par rapport à $\langle M \rangle_t$.
2. La condition sur le supréum de $\mathbf{E}M_t^2$ n'est pas restrictive car, en général, on s'intéressera à $\int_0^t H_s dM_s$, et la condition devient alors $\sup\{\mathbf{E}M_s^2; s \leq t\} < \infty$.

Commençons tout d'abord, comme dans le premier paragraphe, par étudier l'intégrale stochastique d'un processus simple. Soit $a \geq 0$ fixé, et K une v.a. bornée $\mathcal{F}_a$ -mesurable et $H(s) = K\mathbf{1}_{[a,b[}(s)$. On définit naturellement

$$N_t = \int_0^t H_s dM_s = K(M_{t \wedge b} - M_{t \wedge a}).$$

Nous allons calculer, comme dans le cas du mouvement brownien, $\mathbf{E}(N_t^2)$. On a le lemme suivant :

Lemme 3.25 $(N_t)_{t \geq 0}$ est une martingale continue, $\mathbf{E}N_\infty^2 = \mathbf{E}(K^2(\langle M \rangle_b - \langle M \rangle_a))$ et

$$\langle N \rangle_t = \int_0^t K^2 \mathbf{1}_{[a,b[}(s) d\langle M \rangle_s.$$

Preuve : du lemme 3.25. La continuité et l'intégrabilité de N_t sont évidentes. Pour montrer que $(N_t)_{t \geq 0}$ est une martingale, fixons $s < t$ et montrons que $\mathbf{E}(N_t | \mathcal{F}_s) = N_s$. Il y a six cas à faire suivant les valeurs respectives de s et t par rapport à a et b .

Par exemple, si $s \leq a \leq t \leq b$, on a $N_t = K(M_t - M_a)$, d'où, en conditionnant par rapport à $\mathcal{F}_a$, $\mathbf{E}(N_t | \mathcal{F}_s) = \mathbf{E}(K\mathbf{E}(M_t - M_a | \mathcal{F}_a) | \mathcal{F}_s) = 0 = N_s$. Les autres cas sont similaires.

Pour tout $t \geq b$, N_t est égal à $K(M_b - M_a)$. On a donc

$$\mathbf{E}(N_\infty^2) = \mathbf{E}(K^2(M_b - M_a)^2) = \mathbf{E}(K^2\mathbf{E}(M_b^2 + M_a^2 - 2M_bM_a | \mathcal{F}_a)).$$

En utilisant le fait que $M^2 - \langle M \rangle$ est une martingale, on obtient

$$\begin{aligned} \mathbf{E}(M_b^2 + M_a^2 - 2M_bM_a | \mathcal{F}_a) &= \mathbf{E}(M_b^2 - \langle M \rangle_b | \mathcal{F}_a) + \mathbf{E}(\langle M \rangle_b | \mathcal{F}_a) + M_a^2 - 2M_a\mathbf{E}(M_b | \mathcal{F}_a) \\ &= M_a^2 - \langle M \rangle_a + \mathbf{E}(\langle M \rangle_b | \mathcal{F}_a) - M_a^2 \\ &= \mathbf{E}(\langle M \rangle_b - \langle M \rangle_a | \mathcal{F}_a). \end{aligned}$$

On conclut que $\mathbf{E}(N_\infty^2) = \mathbf{E}(K^2(\langle M \rangle_b - \langle M \rangle_a))$.

La dernière assertion consiste à vérifier que $N_t^2 - \int_0^t K^2 \mathbf{1}_{[a,b[}(s) d\langle M \rangle_s$ est une martingale, ce qui se fait en distinguant les positions relatives de t , s , a et b .

Lemme 3.26 Si H^n est une suite de processus simples tels que

$$\lim_{n,m \rightarrow \infty} \mathbf{E} \left(\int_0^\infty (H_s^n - H_s^m)^2 d\langle M \rangle_s \right) = 0$$

alors

$$\lim_{n,m \rightarrow \infty} \mathbf{E} \sup_{s < \infty} (N_s^m - N_s^n)^2 = 0.$$

Preuve : du lemme 3.26. Soit $H = \sum K_i \mathbf{1}_{[a_i, b_i[}$ un processus simple. On définit $N_t = \int_0^t H_s dM_s$ par linéarité. D'après le lemme 3.25, N_t est une somme de martingales de carré intégrable, donc une martingale de carré intégrable. On peut lui appliquer l'inégalité de Doob et on obtient :

$$\mathbf{E} \left(\sup_{s \geq 0} N_s^2 \right) \leq c \mathbf{E}(N_\infty^2),$$

où c est une constante positive. Un calcul simple montre que

$$\mathbf{E}(N_\infty^2) = \sum \mathbf{E} (K_i^2 (\langle M \rangle_{b_i} - \langle M \rangle_{a_i})) = \mathbf{E} \left(\int_0^\infty H_s^2 d\langle M \rangle_s \right).$$

Si on applique ce résultat au processus simple $H^n - H^m$, l'intégrale stochastique obtenue est $N^n - N^m$ et on obtient le résultat souhaité.

Il nous reste maintenant à passer à la limite pour définir l'intégrale stochastique d'un processus adapté H .

Théorème 3.27 *Soit H un processus adapté vérifiant*

$$\mathbf{E} \left(\int_0^\infty H_s^2 d\langle M \rangle_s \right) < \infty$$

et H^n une suite de processus simples vérifiant

$$\lim_n \mathbf{E} \left(\int_0^\infty (H_s^n - H_s)^2 d\langle M \rangle_s \right) = 0.$$

Alors N_s^n converge uniformément pour $s \geq 0$ dans $\mathcal{L}^2$ vers une limite notée N_s . Le processus N_s est une martingale continue et ne dépend pas de la suite H_n . On a

$$\langle N \rangle_t = \int_0^t H_s^2 d\langle M \rangle_s \text{ p.s..}$$

Preuve : D'après le lemme 3.26, N_s^n est une suite de Cauchy pour la norme $\mathcal{L}^2$ du suprémum, elle converge donc dans $\mathcal{L}^2$ et on peut extraire une sous-suite qui converge presque-sûrement uniformément pour $s \geq 0$. Les martingales N_s^n étant continues, on en déduit alors que N_s est continu sur $\mathbb{R}^+$.

Fixons $s \leq t$ et étudions $\mathbf{E}((\mathbf{E}N_t^n | \mathcal{F}_s - \mathbf{E}N_t | \mathcal{F}_s)^2)$. On a

$$\mathbf{E}((\mathbf{E}N_t^n | \mathcal{F}_s - \mathbf{E}N_t | \mathcal{F}_s)^2) \leq \mathbf{E}(\mathbf{E}(N_t^n - N_t)^2 | \mathcal{F}_s) = \mathbf{E}(N_t^n - N_t)^2.$$

Le membre de droite tend vers 0 lorsque N tend vers $+\infty$ et le membre de droite est égal à $\mathbf{E}((N_s^n - \mathbf{E}N_t | \mathcal{F}_s)^2)$. On a donc, en passant à la limite : $\mathbf{E}(N_t | \mathcal{F}_s) = N_s$ p.s..

Il est trivial que la limite ne dépend pas de la suite de processus simples choisis, et la dernière assertion s'obtient par passage à la limite.

3.3 Généralisations

Processus adaptés quelconques

Le premier pas est de généraliser l'intégrale stochastique à des processus adaptés vérifiant

$$\int_0^\infty H_s^2 d\langle M \rangle_s < \infty \text{ p.s.}$$

mais pour lesquels la quantité ci-dessus n'est pas intégrable. Dans ce but, considérons les temps d'arrêt

$$T_N = \inf\{t > 0 : \int_0^t H_s^2 d\langle M \rangle_s > N\},$$

et définissons

$$\int_0^t H_s dM_s = \int_0^t H_s dM_{s \wedge T_N} \text{ si } t \leq T_N.$$

Comme $d\langle M \rangle_{s \wedge T_N} = 0$ pour $s > T_N$, la martingale $M_{s \wedge T_N}$ et le processus H vérifient la condition d'intégrabilité nécessaire. Lorsque $N \rightarrow \infty$, $T_N \rightarrow \infty$: on définit ainsi $\int_0^t H_s dM_s$ pour tout t positif, et la définition obtenue est consistante.

Martingales locales ou de carré non intégrable

On se donne une martingale locale continue (M_t) , sans condition d'intégrabilité. On considère ici les temps d'arrêt $S_N = \inf\{t, |M_t| > N\}$ et on définit comme précédemment

$$\int_0^t H_s dM_s = \int_0^t H_s dM_{s \wedge S_N} \text{ si } t \leq S_N.$$

Intégration par rapport à des semi-martingales continues

Définition 3.28 Une semi-martingale $(X_t)_{t \geq 0}$ est un processus de la forme $X_t = M_t + A_t$ où M_t est une martingale et A_t est un processus adapté à variations bornées.

Considérons une semi-martingale continue $(X_t)_t = (M_t + A_t)_t$. $(A_s)_s$ étant à variations bornées, on peut définir "proprement" (au sens de Lebesgue–Stieljes) l'intégration par rapport à dA_s . On pose alors, pour tout processus adapté H tel que $\int H_s^2 d\langle M \rangle_s + \int_0^t |H_s| |dA_s| < \infty$,

$$\int_0^t H_s dX_s = \int_0^t H_s dM_s + \int_0^t H_s dA_s.$$

Remarque : La variation quadratique de X est, par définition, égale à celle de M .

Une autre façon d'écrire une semi-martingale est d'écrire une équation du type

$$dX_s = dM_s + dA_s$$

ou

$$dX_s = f(B_s)dB_s + g(B_s)ds.$$

Une telle équation s'appelle une **équation différentielle stochastique** (ou EDS). Nous ne nous intéressons pas ici aux conditions d'existence des solutions d'une EDS.

3.4 Formule de Itô

Énoncé

La formule de Itô est une formule de changement de variable dans une intégrale stochastique.

Théorème 3.29 Soit (M_t) une martingale continue et f une fonction de classe $\mathcal{C}^2$ sur $\mathbb{R}$. On a alors presque sûrement :

$$f(M_t) - f(M_0) = \int_0^t f'(M_s) dM_s + \frac{1}{2} \int_0^t f''(M_s) d\langle M \rangle_s.$$

Un exemple important est le cas où M est le mouvement brownien réel issu d'un point $x \in \mathbb{R}$. On obtient le corollaire suivant.

Corollaire 3.30 Soit $(B_t)_{t \geq 0}$ un mouvement brownien réel issu d'un point $x \in \mathbb{R}$. On a, pour toute fonction f de classe $\mathcal{C}^2$,

$$f(B_t) = f(x) + \int_0^t f'(B_s) dB_s + \frac{1}{2} \int_0^t f''(B_s) ds.$$

Exemple : En prenant pour f la fonction $f(x) = x^2$, on obtient :

$$M_t^2 = M_0^2 + 2 \int_0^t M_s dM_s + \langle M \rangle_t.$$

Exercice : Soit f une fonction de classe $\mathcal{C}^2$ et M_t la martingale définie par l'intégrale stochastique suivante :

$$M_t = \int_0^t H_s dB_s.$$

Écrivez la formule de Itô pour $f(M_t)$.

La formule de Itô pour les semi-martingales

Rappelons que si X est une semi-martingale continue, $X = M + A$ où M est une martingale continue et A un processus continu à variations bornées, et $\langle X \rangle$ est par définition égal à $\langle M \rangle$.

Théorème 3.31 Soit X une semi-martingale continue et f une fonction de classe $\mathcal{C}^2$. On a

$$f(X_t) - f(X_0) = \int_0^t f'(X_s) dX_s + \frac{1}{2} \int_0^t f''(X_s) d\langle X \rangle_s.$$

Attention : Ne pas oublier le terme de variation quadratique $d\langle X \rangle_s$, en particulier lorsque la semi-martingale est donnée sous la forme d'une équation différentielle stochastique !

Exercice : Soit g une fonction de classe $\mathcal{C}^2$ et convexe. Écrivez la formule de Itô pour $g(B_t)$, où B_t est un mouvement brownien, et déduisez-en la décomposition de Doob-Meyer de $g(B_t)$ (i.e., écrivez $g(B_t)$ sous la forme d'une martingale additionnée d'un processus à variations bornées). Écrivez les deux formules de Itô possibles pour $f(g(B_t))$, pour une fonction f de classe $\mathcal{C}^2$.

La formule de Itô dans $\mathbb{R}^d$

Notation : Si M et N sont deux martingales de carré intégrable, on définit le « **crochet croisé** » de M et N par $\langle M, N \rangle_t = (\langle M + N \rangle_t - \langle M \rangle_t - \langle N \rangle_t)/2$. On vérifie alors que $(M_t N_t - \langle M, N \rangle_t)_t$ est une martingale.

Exercice : Calculez les crochets $\langle M, M \rangle$, et $\langle B_1, B_2 \rangle$ où B_1 et B_2 sont deux mouvements browniens indépendants.

Exercice : Vérifiez que si M et N sont deux martingales indépendantes et suffisamment intégrables, par rapport à la filtration $\mathcal{F}_t = \sigma\{M_s, N_s, s \leq t\}$, $(M_t N_t)$ est une martingale et en déduire que $\langle M, N \rangle_t = 0$ p.s.

Exercice : Soient $(M_t)_t$ et $(N_t)_t$ deux martingales continues par rapport à la même filtration et $(K_t)_t$ et $(H_t)_t$ deux processus adaptés. On note $X_t = \int_0^t H_s dM_s$ et $Y_t = \int_0^t K_s dN_s$. Montrez, en commençant par le cas où H et K sont des processus simples, que

$$\langle X, Y \rangle_t = \int_0^t H_s K_s d\langle M, N \rangle_s.$$

Théorème 3.32

1. Soient f une fonction de classe $\mathcal{C}^2$ sur $\mathbb{R}^d$ et $M = (M^1, \dots, M^d)$ un d -uplet de semi-martingales continues. On a alors

$$f(M_t) - f(M_0) = \sum_{i=1}^d \int_0^t \frac{\partial f}{\partial x_i}(M_s) dM_s^i + \frac{1}{2} \sum_{i,j=1}^d \int_0^t \frac{\partial^2 f}{\partial x_i \partial x_j}(M_s) d\langle M^i, M^j \rangle_s.$$

2. Soient f une fonction de classe $\mathcal{C}^2$ sur $\mathbb{R}^d$ et $(B_t)_{t \geq 0}$ un mouvement brownien sur $\mathbb{R}^d$. On a alors

$$f(B_t) - f(B_0) = \int_0^t \nabla f(B_s) dB_s + \frac{1}{2} \int_0^t \Delta f(B_s) ds,$$

où

$$\begin{aligned} \nabla f &= \left(\frac{\partial f}{\partial x_1}, \dots, \frac{\partial f}{\partial x_d} \right), \\ \nabla f(B_s) dB_s &= \sum_{i=1}^d \frac{\partial f}{\partial x_i}(B_s) dB_s^i \\ \Delta f &= \sum_{i=1}^d \frac{\partial^2 f}{\partial x_i^2}. \end{aligned}$$

Exemple : Soient X et Y deux martingales continues. On a

$$X_t Y_t = \int_0^t Y_s dX_s + \int_0^t X_s dY_s + \int_0^t d\langle X, Y \rangle_s,$$

c'est-à-dire que $X_t Y_t - \langle X, Y \rangle_t$ est une martingale.

Exercice :

1. Montrez directement puis en utilisant la formule de Itô que $(\exp(\lambda B_t - \lambda^2 t/2))_t$ est une martingale, pour tout réel λ fixé.
2. Soit $T_a = \inf\{t, B_t = a\}$. Déduisez de 1) et du théorème d'arrêt de Doob que $\mathbf{E}_0 \exp(-\alpha T_a) = \exp(-a\sqrt{2\alpha})$.

3.5 Preuve de la formule de Itô.

Soit M une martingale continue. En considérant les temps d'arrêt

$$T_N = \inf\{t \geq 0, |M_t| > N \text{ ou } \langle M \rangle_t > N\},$$

on voit que l'on peut supposer M et $\langle M \rangle$ bornés.

Écrivons la formule de Taylor pour la fonction f de classe $\mathcal{C}^2$: pour tous réels $a < b$, il existe $c(a, b) \in]a, b[$ tel que

$$f(b) - f(a) = f'(a)(b - a) + \frac{1}{2} f''(c(a, b))(b - a)^2.$$

Intuitivement, on va l'utiliser en discrétisant la martingale M . Le terme en $(b - a)$ donnera l'intégrale stochastique, et le terme en $(b - a)^2$, l'intégrale par rapport à la variation quadratique.

Soient t et ε deux réels positifs fixés. Nous construisons une suite croissante de temps d'arrêt t_i , dont la limite est t , et tels que la martingale bouge peu sur les intervalles $[t_i, t_{i+1}]$:

$$t_{i+1} = t_{i+1}^\varepsilon = t \wedge \inf\{s > t_i, s - t_i > \varepsilon \text{ ou } |M_s - M_{t_i}| > \varepsilon \text{ ou } \langle M \rangle_s - \langle M \rangle_{t_i} > \varepsilon\}.$$

On a donc, en posant $\alpha_i = f''(c(M_{t_i}, M_{t_{i+1}}))$:

$$f(M_t) - f(M_0) = \sum_i f'(M_{t_i})(M_{t_{i+1}} - M_{t_i}) + \frac{1}{2} \sum \alpha_i (M_{t_{i+1}} - M_{t_i})^2.$$

Montrons que $S_1^\varepsilon = \sum_i f'(M_{t_i})(M_{t_{i+1}} - M_{t_i}) \rightarrow \int_0^t H_s dM_s$ lorsque ε tend vers 0. Considérons le processus simple H^ε défini par

$$H_s^\varepsilon = \sum f'(M_{t_i}) \mathbf{1}_{[t_i, t_{i+1}[}.$$

On a $S_1^\varepsilon = \int_0^t H_s^\varepsilon dM_s$. Comme M est une martingale bornée, f' est uniformément continue sur l'ensemble des valeurs prises par M . En notant $H_s = f'(M_s) \mathbf{1}_{[0, t]}$, on en déduit que $H_s^\varepsilon - H_s$ tend uniformément vers 0. Comme $\langle M \rangle_t$ est bornée, on a obtenu que H^ε tend vers H au sens de la norme $\|\cdot\|_{\mathcal{H}}$, ce qui implique la convergence des intégrales stochastiques dans $\mathcal{L}^2$, et leur convergence presque sûre si on extrait une sous-suite.

Il nous reste à montrer que

$$\sum \alpha_i (M_{t_{i+1}} - M_{t_i})^2 \rightarrow \int_0^t f''(M_s) d\langle M \rangle_s.$$

Les sommes ci-dessus comportent un nombre fini de termes non nuls.

Étudions d'abord le cas où $f'' \equiv 1$.

Lemme 3.33 *On a*

$$\sum_i (M_{t_{i+1}} - M_{t_i})^2 \rightarrow \langle M \rangle_t, \text{ dans } \mathcal{L}^2.$$

Preuve : du lemme 3.33. Notons $\Delta_i = (M_{t_{i+1}} - M_{t_i})^2 - (\langle M \rangle_{t_{i+1}} - \langle M \rangle_{t_i})$ et montrons que $\sum \Delta_i$ tend vers 0 dans $\mathcal{L}^2$. Remarquons que $\mathbf{E}\Delta_i | \mathcal{F}_{t_i} = 0$ et, si $i < j$,

$$\mathbf{E}(\Delta_i \Delta_j) = \mathbf{E}(\Delta_i \mathbf{E}(\Delta_j | \mathcal{F}_{t_j})) = 0.$$

On en déduit que

$$\mathbf{E}(\sum \Delta_i)^2 = \mathbf{E} \sum (\Delta_i^2).$$

En utilisant le fait que $(a - b)^2 \leq 2a^2 + 2b^2$, on majore Δ_i^2 :

$$\Delta_i^2 \leq 2(M_{t_{i+1}} - M_{t_i})^4 + 2(\langle M \rangle_{t_{i+1}} - \langle M \rangle_{t_i})^2.$$

Par définition de t_i , on a

$$\begin{aligned} \mathbf{E} \left(\sum (M_{t_{i+1}} - M_{t_i})^4 \right) &\leq \varepsilon^2 \mathbf{E} \left(\sum (M_{t_{i+1}} - M_{t_i})^2 \right) \\ &= \varepsilon^2 \mathbf{E} \left(\sum (M_{t_{i+1}} - M_{t_i}) \right)^2 \\ &= \varepsilon^2 \mathbf{E} (M_t - M_0)^2, \end{aligned}$$

et

$$\mathbf{E} \left(\sum (\langle M \rangle_{t_{i+1}} - \langle M \rangle_{t_i})^2 \right) \leq \varepsilon \mathbf{E} \left(\sum (\langle M \rangle_{t_{i+1}} - \langle M \rangle_{t_i}) \right) = \varepsilon \mathbf{E} \langle M \rangle_t.$$

Lorsque ε tend vers 0, on obtient le résultat souhaité, c'est-à-dire

$$\sum (M_{t_{i+1}} - M_{t_i})^2 \rightarrow \langle M \rangle_t, \quad \text{dans } \mathcal{L}^2.$$

Fin de la preuve de la formule d'Itô. Reprenons notre fonction f , $\alpha_i = f''(c(M_{t_i}, M_{t_{i+1}}))$, et notons

$$G_s^\varepsilon = \sum \alpha_i \mathbf{1}_{[t_i, t_{i+1}[}(s) + f''(M_t) \mathbf{1}_{[s, \infty[}$$

et

$$A_s^\varepsilon = \sum_{t_i \leq s} (M_{t_{i+1}} - M_{t_i})^2.$$

Avec ces notations, on a

$$\sum \alpha_i (M_{t_{i+1}} - M_{t_i})^2 = \int_0^t G_s^\varepsilon dA_s^\varepsilon.$$

Le lemme 3.33 implique que A_s^ε converge en probabilité vers $\langle M \rangle_s$ et G_s^ε converge uniformément en s vers $f''(M_{s \wedge T})$ en vertu de la continuité uniforme de f'' sur l'ensemble des valeurs prises par M . On peut extraire une sous-suite ε_k telle que $A_s^{\varepsilon_k}$ converge p.s. vers $\langle M \rangle_s$ et $\sup_{\varepsilon_k} A_s^{\varepsilon_k} < \infty$. Pour conclure, il nous reste donc à justifier le passage à la limite sous l'intégrale. On a :

$$\int_0^t G_s^{\varepsilon_k} dA_s^{\varepsilon_k} - \int_0^t f''(M_s) d\langle M \rangle_s \leq \int_0^t |G_s^{\varepsilon_k} - f''(M_s)| dA_s^{\varepsilon_k} + \left| \int_0^t f''(M_s) d(A_s^{\varepsilon_k} - \langle M \rangle_s) \right|.$$

La fonction f'' est uniformément continue sur l'ensemble des valeurs prises par M . Pour tout $\delta > 0$ fixé, on peut donc trouver un $\varepsilon_0 > 0$ tel que, pour tout $|x - y| < \varepsilon_0$ implique $|f''(x) - f''(y)| \leq \delta$. Pour tout k assez grand et tout $s \in \mathbb{R}^+$, on aura $|G_s^{\varepsilon_k} - f''(M_s)| \leq \delta$ d'où

$$\int_0^t |G_s^{\varepsilon_k} - f''(M_s)| dA_s^{\varepsilon_k} \leq \delta A_t^{\varepsilon_k}.$$

Comme $A_t^{\varepsilon_k}$ est borné sur $k \in \mathbb{N}$, ce terme tend vers 0.

Fixons ω , notons $\phi(s) = f''(M_s(\omega))$ et montrons pour terminer que $\int_0^t \phi(s) dA_s^{\varepsilon_k}(\omega)$ converge vers $\int_0^t \phi(s) d\langle M \rangle_s(\omega)$. On sait que $A_s^{\varepsilon_k}(\omega) \rightarrow \langle M \rangle_s(\omega)$, pour tout $s \geq 0$, ce qui signifie que la mesure dA_s^ε converge vers la mesure $d\langle M \rangle_s$, et ces mesures sont bornées. La fonction ϕ étant bornée, c'est exactement ce qu'il nous faut pour conclure.

3.6 Formule de Girsanov

Le but de la formule de Girsanov est, par un changement de mesure de probabilité, de transformer un processus en une martingale.

Avant tout, quelques notations : on se donne deux mesures de probabilité $\mathbf{P}$ et $\mathbf{Q}$ sur (Ω, Σ) , que l'on suppose équivalentes, c'est-à-dire qu'il existe une densité « dans les deux sens ». On notera α la densité de $\mathbf{Q}$ par rapport à $\mathbf{P}$: $d\mathbf{Q}(\omega) = \alpha(\omega) d\mathbf{P}(\omega)$. On se donne également une filtration continue à droite $(\mathcal{F}_t)_t$. On notera $\mathbf{P}_t$ et $\mathbf{Q}_t$ les mesures $\mathbf{P}$ et $\mathbf{Q}$ restreintes $\mathcal{F}_t$ et α_t l'espérance conditionnelle (sous $\mathbf{P}$) de α :

$$\alpha_t = \mathbf{E}_{\mathbf{P}}(\alpha | \mathcal{F}_t).$$

Proposition 3.34

1. α_t est la densité de $\mathbf{Q}_t$ par rapport à $\mathbf{P}_t$:

$$\alpha_t = \frac{d\mathbf{Q}_t}{d\mathbf{P}_t}.$$

2. Un processus (Y_t) est une martingale pour $\mathbf{Q}$ si et seulement si $(\alpha_t Y_t)$ est une martingale par rapport à $\mathbf{P}$.

Preuve du point 1. Soit $A \in \mathcal{F}_t$. On a

$$\begin{aligned} \int \mathbf{1}_A \alpha_t d\mathbf{P}_t &= \int \mathbf{1}_A \alpha_t d\mathbf{P} \quad [\text{mesure restreinte}] \\ &= \int \mathbf{1}_A \alpha d\mathbf{P} \quad [\text{espérance conditionnelle}] \\ &= \int \mathbf{1}_A d\mathbf{Q} \quad [\text{densité}] \\ &= \int \mathbf{1}_A d\mathbf{Q}_t \quad [\text{mesure restreinte}] \end{aligned}$$

d'où le résultat.

Preuve du point 2. Soit $(Y_t)_t$ une $\mathbf{Q}$ -martingale. Fixons $A \in \mathcal{F}_s$, pour $s \leq t$. Il faut montrer que

$$\mathbf{E}_{\mathbf{P}}(\mathbf{1}_A \alpha_t Y_t) = \mathbf{E}_{\mathbf{P}}(\mathbf{1}_A \alpha_s Y_s).$$

On a

$$\begin{aligned} \mathbf{E}_{\mathbf{P}}(\mathbf{1}_A \alpha_t Y_t) &= \int \mathbf{1}_A Y_t \alpha_t d\mathbf{P} \\ &= \int \mathbf{1}_A Y_t \alpha_t d\mathbf{P}_t \quad [\text{restriction}] \\ &= \int \mathbf{1}_A Y_t d\mathbf{Q}_t \quad [\text{densité}] \\ &= \int \mathbf{1}_A Y_t d\mathbf{Q} \quad [\text{restriction}] \\ &= \int \mathbf{1}_A Y_s d\mathbf{Q} \quad [Y \text{ est une } \mathbf{Q}\text{-martingale}] \\ &= \int \mathbf{1}_A Y_s d\mathbf{Q}_s \quad [\text{restriction}] \\ &= \int \mathbf{1}_A Y_s \alpha_s d\mathbf{P}_s \quad [\text{densité}] \\ &= \int \mathbf{1}_A Y_s \alpha_s d\mathbf{P} \quad [\text{restriction}] \end{aligned}$$

Théorème 3.35 (Formule de Girsanov) Soit $(X_t)_t$ une $\mathbf{P}$ -martingale. On note A_t le processus à variations bornées suivant :

$$A_t = \int_0^t \frac{1}{\alpha_s} d\langle \alpha, X \rangle_s.$$

Alors $(X_t - A_t)$ est une $\mathbf{Q}$ martingale.

Preuve : Appliquer la formule de Itô à la fonction $f(x, a, \alpha) = \alpha(x - a)$.

Application : Soit (X_t) une $\mathbf{P}$ -martingale. On se donne une fonction bornée b . On cherche une mesure $\mathbf{Q}$ telle que $(X_t - \int_0^t b(X_s) d\langle X \rangle_s)$ soit une $\mathbf{Q}$ -martingale.

En pratique, on cherche une densité α telle que

$$\int_0^t b(X_s) d\langle X \rangle_s = \int_0^t \frac{d\langle \alpha, X \rangle_s}{\alpha_s}.$$

On écrit α_t sous la forme

$$\alpha_t = 1 + \int_0^t H_s dX_s.$$

On a alors

$$\langle \alpha, X \rangle_t = \int_0^t H_s d\langle X \rangle_s$$

ce qui induit

$$\frac{H_s}{\alpha_s} = b_s.$$

On obtient donc une équation intégrale sur α :

$$\alpha_t - 1 = \int_0^t b(X_s) \alpha_s dX_s.$$

On remarque alors que si $Y_t = \int_0^t b(X_s) dX_s$, $\alpha_t = \exp(Y_t - \langle Y_t \rangle/2)$ est solution de cette équation intégrale.

Chapitre 4

Processus de renouvellement et de Poisson

1 Processus de renouvellement

1.1 Définition

Soit $(X_n)_{n \geq 0}$ une suite de variables aléatoires positives. On note S_n la suite des sommes partielles : $S_0 = 0$ et $S_n = X_n + S_{n-1}$, pour tout $n \geq 1$. On considère alors le processus R_t défini comme suit :

$$R_t = \text{card}\{n \geq 1, S_n \leq t\} = \sum_{n \geq 1} \mathbf{1}_{\{S_n \leq t\}}.$$

Par exemple, si les X_n modélisent les durées de vie d'une ampoule, R_t représente le nombre d'ampoules changées avant l'instant t ; les X_n peuvent également représenter le temps séparant deux ventes successives, ou deux sinistres successifs pour une compagnie d'assurance. R_t désignera alors, suivant le cas, le nombre d'articles vendus ou le nombre de sinistres survenus au cours de l'intervalle de temps $[0, t]$. La suite (S_n) est appelée **processus de renouvellement associé** aux $(X_n)_{n \geq 0}$ et le processus (R_t) est le **processus de comptage**.

Par abus de langage, on appelle également R_t processus de renouvellement.

On se placera dans le cas où les $(X_n)_{n \geq 0}$ sont des variables aléatoires indépendantes et identiquement distribuées et on étudiera plus particulièrement le comportement asymptotique lorsque t tend vers $+\infty$ de $\mathbf{E}(R_t)$, $\mathbf{E}(R_{t+a} - R_t)$, $S_{R_t+1} - t$, $t - S_{R_t}$.

1.2 Quelques résultats élémentaires

On se place dans ce paragraphe et les suivants dans le cas où les (X_n) sont des variables aléatoires positives indépendantes et identiquement distribuées telles que $\mathbf{P}(X_1) = 0 < 1$.

Proposition 4.1 *Pour tout t fixé, R_t est intégrable et on a*

$$\mathbf{E}(R_t) = \sum_{n \geq 1} \mathbf{P}(S_n \leq t).$$

Preuve : Fixons $t \in \mathbb{R}^+$. Par définition de R_t ,

$$R_t = \sum_{n \geq 1} \mathbf{1}_{\{S_n \leq t\}}.$$

d'où

$$\mathbf{E}(R_t) = \sum_{n \geq 1} \mathbf{E}(\mathbf{1}_{\{S_n \leq t\}}) = \sum_{n \geq 1} \mathbf{P}\{S_n \leq t\}.$$

De plus, on a

$$\{X_1 + \cdots + X_{n+p} \leq t\} \subset \{X_1 + \cdots + X_n \leq t\} \cap \{X_{n+1} + \cdots + X_{n+p} \leq t\}$$

donc

$$\mathbf{P}\{X_1 + \cdots + X_{n+p} \leq t\} \leq \mathbf{P}\{X_1 + \cdots + X_n \leq t\} \mathbf{P}\{X_{n+1} + \cdots + X_{n+p} \leq t\}$$

c'est-à-dire

$$\mathbf{P}\{S_{n+p} \leq t\} \leq \mathbf{P}\{S_n \leq t\} \mathbf{P}\{S_p \leq t\}.$$

On dit que la suite $u_n = \mathbf{P}\{S_n \leq t\}$ est sous-multiplicative. Il est facile de voir que $\sum_n \mathbf{P}\{S_n \leq t\}$ converge.

En effet, puisque $S_n \rightarrow \infty$ p.s., il existe $p \geq 1$ tel que $\mathbf{P}(S_p \leq t) < 1$. Fixons un tel p et écrivons, à l'aide de la division euclidienne, les entiers n sous forme $n = kp + r$, où $k \geq 0$, $0 \leq r \leq p - 1$, on a

$$\sum_n \mathbf{P}\{S_n \leq t\} = \sum_{r=0}^{p-1} \sum_{k=0}^{\infty} \mathbf{P}\{S_{kp+r} \leq t\}$$

puis

$$\begin{aligned} \mathbf{P}\{S_{kp+r} \leq t\} &\leq \mathbf{P}\{S_{kp} \leq t\} \mathbf{P}\{S_r \leq t\} \\ &\leq (\mathbf{P}\{S_p \leq t\})^k \mathbf{P}\{S_r \leq t\} \end{aligned}$$

On obtient donc

$$\sum_n \mathbf{P}\{S_n \leq t\} \leq \sum_{r=0}^{p-1} \sum_{k=0}^{\infty} (\mathbf{P}\{S_p \leq t\})^k \mathbf{P}\{S_r \leq t\} \leq \sum_{r=0}^{p-1} \mathbf{P}\{S_r \leq t\} \frac{1}{1 - \mathbf{P}\{S_p \leq t\}} < \infty.$$

Proposition 4.2 *Toujours sous les mêmes conditions, et en supposant de plus que les X_n sont intégrables, on a*

$$\lim \frac{R_t}{t} = \frac{1}{\mathbf{E}(X_1)} \text{ p.s.}$$

Preuve : Il est clair que (R_t) est un processus croissant et que $\lim R_t = +\infty$ p.s.

La loi forte des grands nombres permet de dire que

$$\frac{S_n}{n} \rightarrow \mathbf{E}(X_1) \text{ p.s.}$$

Par le principe d'extraction de sous-suite, on en déduit que

$$\frac{S_{R_t}}{R_t} \rightarrow \mathbf{E}(X_1) \text{ p.s.}$$

puis on remarque que

$$\frac{S_{R_t}}{R_t} \leq \frac{t}{R_t} \leq \frac{S_{R_t+1}}{R_t+1} \frac{R_t+1}{R_t}.$$

Il reste à faire tendre t vers l'infini pour conclure que R_t/t tend vers $1/\mathbf{E}(X_1)$ p.s.

1.3 Le théorème du renouvellement

Théorème 4.3 *Si les (X_n) sont des variables aléatoires positives, indépendantes, identiquement distribuées, intégrables et telles que $\mathbf{P}(X_1 = 0) < 1$ et telles que, pour tout $\lambda > 0$, $\mathbf{P}(X_1 \in \lambda\mathbb{N}) < 1$, les limites suivantes ont lieu presque sûrement lorsque t tend vers $+\infty$:*

$$\begin{aligned}\lim_t \frac{\mathbf{E}(R_t)}{t} &= \frac{1}{\mathbf{E}(X_1)} \\ \lim_t \mathbf{E}R_{t+a} - \mathbf{E}(R_t) &= \frac{a}{\mathbf{E}(X_1)} \\ \lim_t \mathbf{P}(S_{R_t+1} - t \leq x) &= \lim_t \mathbf{P}(t - S_{R_t} \leq x) = \int_0^x \frac{\mathbf{P}(X_1 \geq y)}{\mathbf{E}(X_1)} dy \\ \lim_t \mathbf{P}(S_{R_t+1} - t \geq x \text{ et } t - S_{R_t} \geq y) &= \int_{x+y}^{+\infty} \frac{\mathbf{P}(X_1 \geq u)}{\mathbf{E}(X_1)} du\end{aligned}$$

1.4 Le théorème central limite pour les processus de renouvellement

Théorème 4.4 *On suppose que les X_n sont des variables aléatoires positives, indépendantes, identiquement distribuées et de carré intégrable. On note $\mu = \mathbf{E}(X_1) > 0$ et $\sigma^2 = \text{var}(X_1)$. Alors, lorsque t tend vers $+\infty$, on a la convergence en loi suivante :*

$$\mathbf{P}\left(\frac{R_t - t/\mu}{\sigma\sqrt{t/\mu^3}} \leq x\right) \longrightarrow \frac{1}{\sqrt{2\pi}} \int_{-\infty}^x e^{-u^2/2} du.$$

2 Processus de Poisson

2.1 Définition

On se donne une suite de variables aléatoires **indépendantes et identiquement distribuées** $(X_n)_{n \geq 0}$ de loi **exponentielle** (c'est-à-dire de densité $x \rightarrow \lambda e^{-\lambda x} \mathbf{1}_{\mathbb{R}^+}$) et on considère le processus de renouvellement, défini par $S_0 = 0$ et $S_n = S_{n-1} + X_n$.

On appelle **processus de Poisson homogène** d'intensité $\mathbf{E}(R_t) = \lambda t$ le processus de comptage associé à S_n :

$$R_t = \text{card}\{n \geq 1, S_n \leq t\}.$$

La proposition suivante permet de caractériser les processus de Poisson homogènes par leur loi :

Proposition 4.5 *$(R_t)_{t \geq 0}$ est un processus de Poisson si et seulement si les quatre propriétés suivantes sont vérifiées :*

1. $R_0 = 0$ p.s.
2. Pour tout $t > 0$, R_t suit une loi de Poisson d'espérance λt .
3. Pour tous $t > s \geq 0$, $R_t - R_s$ suit une loi de Poisson d'espérance $\lambda(t - s)$.
4. Pour tous $0 \leq t_1 < t_2 < \dots < t_n$, les variables aléatoires $R_{t_1}, R_{t_2} - R_{t_1}, \dots, R_{t_n} - R_{t_{n-1}}$ sont indépendantes.

2.2 Quelques propriétés

Les propriétés asymptotiques vérifiées par les processus de renouvellement deviennent des propriétés exactes pour le processus de Poisson.

Proposition 4.6 Soit (R_t) un processus de Poisson homogène d'intensité $\mathbf{E}(R_t) = \lambda t$. On a pour tout $t > 0$ et tout $a > 0$

$$\begin{aligned}\frac{\mathbf{E}(R_t)}{t} &= \frac{1}{\mathbf{E}(X_1)} \\ \mathbf{E}(R_{t+a}) - \mathbf{E}(R_t) &= \frac{a}{\mathbf{E}(X_1)} \\ \mathbf{P}(S_{R_t+1} - t \leq x) &= \mathbf{P}(t - S_{R_t} \leq x) = \int_0^x \lambda e^{-\lambda u} du \\ \mathbf{P}(S_{R_t+1} - t \geq x \text{ et } t - S_{R_t} \geq y) &= \int_{x+y}^{+\infty} \lambda e^{-\lambda u} du.\end{aligned}$$

En particulier, les variables aléatoires $S_{R_t+1} - t$ et $t - S_{R_t}$ sont indépendantes et de loi exponentielle : ces résultats qui ne sont vrais qu'asymptotiquement pour un processus de renouvellement quelconque, deviennent vrais pour tout $t > 0$ pour un processus de Poisson homogène.

Le preuve de cette proposition est basée sur le calcul explicite des lois de R_t et S_{R_t} . Voir la fiche d'exercice pour les détails.

Un paradoxe. On pourrait penser que X_{R_t+1} (c'est-à-dire la hauteur du pas qui permet de dépasser t) est de moyenne $1/\lambda$. Il n'en est rien, puisqu'il est de moyenne $2/\lambda$. En effet : la proposition 4.6 montre que $S_{R_t+1} - t$ et $t - S_{R_t}$ sont toutes deux de loi exponentielle d'espérance $1/\lambda$ donc la somme de ces deux variables aléatoires, qui est égale à X_{R_t+1} est de moyenne $2/\lambda$. Plus précisément, X_{R_t+1} suit une loi $\Gamma(2, \lambda)$. Donc, si on s'est fixé un instant t , le premier saut de (R_u) après t a lieu en moyenne $1/\lambda$ unité de temps après t , de même que le dernier saut avant t .

2.3 Processus de Poisson inhomogène

Soit $m : \mathbb{R}^+ \rightarrow \mathbb{R}^+$ une fonction croissante. Un processus de Poisson inhomogène de moyenne $m(t)$ est un processus $(R_t)_{t \geq 0}$ càd-làg, à valeurs dans $\mathbb{N}$ vérifiant :

- $R_0 = 0$,
- Pour tout $t > 0$, $R_t - R_{t-} = 0$ ou 1 ,
- Pour tout $t \geq 0$, la variable aléatoire R_t suit une loi de Poisson de moyenne $m(t)$,
- Pour tous $t \geq s \geq 0$, $R_t - R_s$ suit une loi de Poisson de moyenne $m(t) - m(s)$,
- Pour tous $0 \leq t_1 \leq t_2 \leq \dots \leq t_n$, les variables aléatoires $R_{t_1}, R_{t_2} - R_{t_1}, \dots, R_{t_n} - R_{t_{n-1}}$ sont indépendantes.

2.4 Simulation des processus de Poisson

Le processus de Poisson homogène

Il existe deux méthodes standard pour simuler un processus de Poisson (N_t) homogène de moyenne $\mathbf{E}N_t = \lambda t$ sur un intervalle de temps $[0, T]$. Le principe commun de ces deux méthodes est de simuler les instants de saut du processus.

Première méthode : On part d'un échantillon aléatoire (u_i) de nombres uniformément distribués sur l'intervalle $[0, 1]$. Puisque les intervalles de temps entre deux sauts consécutifs forment une suite de variables aléatoires de loi exponentielle, on peut poser $x_i = -\lambda^{-1} \ln u_i$. Les instants de saut sont alors donnés par $t_1 = x_1$ et, pour tout $n \geq 2$, $t_n = t_{n-1} + x_n$. On arrête la simulation lorsque $t_n > T$. Le processus simulé comptera alors $n - 1$ sauts sur l'intervalle $[0, T]$.

Deuxième méthode : On commence par simuler le nombre de sauts du processus entre 0 et T , qui suit une loi de Poisson de moyenne λT . Conditionnellement au nombre de sauts sur $[0, T]$, les instants de saut sont uniformément distribués sur $[0, T]$. Si la simulation du nombre de sauts fournit comme

résultat n , les instants de saut seront donc $T.u_1, T.u_2, \dots, T.u_n$. Cette méthode, utile d'un point de vue théorique, présente deux inconvénients pratiques. D'une part, la méthode la plus utilisée pour simuler une variable de Poisson est basée sur les sommes d'exponentielles ; autrement dit, en simulant une variable de Poisson, on obtient directement les instants de saut du processus de Poisson. D'autre part, les instants de sauts obtenus par cette méthode ne sont pas classés par ordre croissant. Il est donc nécessaire de trier les valeurs obtenues pour reconstituer le processus, et un tri est une opération lourde du point de vue informatique dès que l'échantillon est grand.

Le processus de Poisson général

On souhaite simuler un processus de Poisson de moyenne $\mathbf{E}(R_t) = m(t)$ où m est une fonction croissante et dérivable sur $[0, T]$. On notera $\lambda(t) = m'(t)$ et on supposera que

$$\sup_{[0, T]} \lambda(t) = \lambda^* < \infty.$$

On commence par simuler un processus de Poisson homogène de moyenne λ^*t sur $[0, T]$, avec pour instants de saut simulés $s_1, \dots, s_n$. On va maintenant supprimer certains des s_i par une méthode du rejet : on prend un échantillon uniforme $(u_i)_{i \leq n}$ indépendant de celui utilisé pour construire les instants $(s_i)_{i \leq n}$. On conserve s_i comme instant de saut du processus inhomogène si $u_i \leq \lambda(t)/\lambda^*$ et on le supprime sinon. Les instants de saut du processus non-homogène simulé sont donc les instants non supprimés.

2.5 Files d'attente

Introduction

Un système de file d'attente cherche à modéliser tout processus que l'on peut découper suivant les étapes suivantes : des « clients » se présentent à l'entrée d'un service, ils se font servir, puis s'en vont. Les « clients » peuvent être des requêtes informatiques, ou des personnes se présentant aux caisses d'un magasin. Évidemment, ces clients arrivent en des temps aléatoires, les temps de service sont eux aussi aléatoires, il peut y avoir un seul guichet, ou un nombre déterminé k de guichets, on peut également limiter la longueur de la file d'attente... Les possibilités sont à peu près infinies !

Les modèles les plus simples à étudier théoriquement sont ceux pour lesquels les clients arrivent suivant un processus de Poisson homogène (les temps inter-arrivées sont exponentiels), et où le temps de service est lui aussi exponentiel, avec toutes les indépendances possibles : on parle de file d'attente M/M/1 ou M/M/ k s'il y a k guichets (M=memoryless). Si la capacité totale du service (attente+guichets) est égale à N , on parle de système M/M/1/ N ou M/M/ k / N . Lorsque les temps inter-arrivées et/ou de service ne sont plus exponentiels, il s'agira de système G/M ou M/G ou G/G (G=general), avec là aussi éventuellement plusieurs guichets et une longueur bornée de file d'attente. En pratique dès que le système n'est plus de type M/M, on se contentera d'une simulation informatique pour le traiter.

2.6 Modélisation

Les clients arrivent dans le système à des instants $T_1, T_1 + T_2, \dots, T_1 + \dots + T_n, \dots$, les (T_i) formant une suite de variables aléatoires positives, indépendantes et identiquement distribuées de moyenne $1/\lambda$. Les arrivées forment donc un processus de renouvellement (R_t) .

La durée du service du n -ième client est une variable aléatoire S_n , les S_n formant une suite de variables aléatoires positives indépendantes et identiquement distribuées, de moyenne $1/\mu$. Les (S_n) sont indépendantes des (T_n) , et, dans le cas de plusieurs serveurs en parallèle, leur loi est la même quel que soit le serveur.

Lorsqu'un client se présente à l'entrée du système, si un serveur est libre, son service commence immédiatement. Sinon, il rentre dans la file d'attente sauf si la capacité maximale du système est déjà atteinte auquel cas il est définitivement rejeté.

On s'intéresse au comportement du système en régime stationnaire, c'est-à-dire que l'on suppose que le nombre (X_t) de clients présents dans le service converge en loi : $\lim_t \mathbf{P}(X_t = n) = p_n$ et $\sum p_n = 1$.

On notera alors

- $L = \sum np_n$ le nombre moyen de clients dans le service en régime stationnaire,
- L_Q le nombre moyen de clients qui attendent,
- W le temps moyen passé dans le service par un client en régime stationnaire,
- W_Q le temps moyen passé à attendre.

Ces quantités sont liées par les **formules de J.D. Little** :

$$L = \lambda_a W, \quad L_Q = \lambda_a W_Q, \quad W_Q = W - 1/\mu,$$

où λ_a est le nombre moyen de clients qui rentrent dans le service par unité de temps en régime stationnaire. Si la capacité du système est infinie, c'est simplement $\lim \mathbf{E}R_t/t$ (c'est donc λ dans le cas d'une entrée poissonnienne de moyenne λt). Si la capacité est finie, il faudra tenir compte des clients rejetés et on aura $\lambda_a < \lim \mathbf{E}R_t/t$.

Le système M/M/1 en régime stationnaire

On suppose donc que les clients arrivent suivant un processus de Poisson R_t d'intensité $\mathbf{E}(R_t) = \lambda t$ et que le temps de service suit une loi exponentielle de moyenne $1/\mu$. Pour fixer les idées, on suppose qu'au temps $t = 0$, il n'y a personne dans la file d'attente et personne en train de se faire servir. On note alors X_t le nombre de personnes dans le système complet à l'instant t et L_t le nombre de personnes en train d'attendre.

Notons $p_n(t) = \mathbf{P}(X_t = n)$ et examinons ce qui se passe au cours d'un intervalle de temps δt . Pour reprendre un langage cohérent avec celui des chaînes de Markov, on dira que le système est dans l'état n à l'instant t si $X_t = n$.

La probabilité qu'il y ait une arrivée dans le système au cours d'un intervalle de temps de longueur δt est $\lambda \delta t$, et la probabilité qu'il y ait un départ est $\mu \delta t$, sauf si on est dans l'état 0. La variation de p_n au cours de cet intervalle de temps provient

- soit de la sortie du système de l'état n : un client arrive et on passe à l'état $n + 1$, ou un client part, et on passe à l'état $n - 1$;
- soit de l'entrée du système dans l'état n : le système était dans l'état $n + 1$ et un client part, ou bien le système était dans l'état $n - 1$ et un client arrive.

On a ainsi, pour $n = 0$,

$$p_0(t + \delta t) - p_0(t) = -\lambda \delta t p_0(t) + \mu \delta t p_1(t)$$

et, pour tout $n \geq 1$,

$$p_n(t + \delta t) - p_n(t) = -(\lambda + \mu) \delta t p_n(t) + \lambda \delta t p_{n-1}(t) + \mu \delta t p_{n+1}(t)$$

En divisant par δt et en faisant tendre cette quantité vers 0, il vient :

$$p'_0(t) = -\lambda p_0(t) + \mu p_1(t)$$

et, pour tout $n \geq 1$,

$$p'_n(t) = -(\lambda + \mu) p_n(t) + \lambda p_{n-1}(t) + \mu p_{n+1}(t).$$

Remarque : Le fait que deux événements (deux arrivées, une arrivée et un départ ou deux départs) ou plus interviennent au cours du laps de temps δt est d'ordre δt^2 , qui est donc négligeable devant δt .

Supposons que l'on est dans un état stationnaire, c'est-à-dire $p_n(t)$ ne dépend pas de t . On aura bien sûr $p'_n(t) = 0$ et on notera pour simplifier $p_n(t) = p_n$.

On obtient donc le système :

$$\lambda p_0 = \mu p_1$$

et, pour tout $n \geq 1$,

$$(\lambda + \mu) p_n = \lambda p_{n-1} + \mu p_{n+1}$$

ainsi que, bien évidemment

$$\sum_{n \geq 0} p_n = 1.$$

En posant $\rho = \lambda/\mu$, il vient : $p_1 = \rho p_0$, puis, par récurrence, pour tout $n \geq 1$, $p_{n+1} = \rho p_n$.

La solution (stationnaire) du système est donc de la forme $p_n = \rho^n p_0$, et on détermine p_0 à l'aide de la dernière équation : p_0 (et les autres p_n non plus) n'existe pas si $\rho \geq 1$, puisque la série $\sum \rho^n$ diverge. Par contre, si $0 < \rho < 1$, la série converge et on obtient $p_0 = 1 - \rho$, et $p_n = \rho^n (1 - \rho)$.

La condition $\rho < 1$ se traduit par

$$\frac{1}{\mu} < \frac{1}{\lambda}$$

c'est-à-dire que le temps de service moyen est strictement plus court que le temps interarrivées moyen. Lorsque ce n'est pas le cas, il y a engorgement du système et la longueur de la file d'attente explose.

Exercice : Vérifiez que le nombre moyen de personnes dans le système est $\lambda/(\mu - \lambda)$. Calculez également le nombre moyen de personnes en attente. Vérifiez la concordance de vos résultats avec les formules de Little. Vérifiez que le temps moyen passé dans le service est égal à $1/(\mu - \lambda)$.

Application numérique : Un client arrive toutes les 15 minutes et la durée moyenne du service est de 10 minutes. Donnez les valeurs de L , L_Q , W , W_Q .

Le système M/M/k

On se place à nouveau en régime stationnaire et on note toujours $p_n = \mathbf{P}(X_t = n)$, la probabilité qu'il y ait n personnes dans le service à l'instant n . Il y a k serveurs et non plus un seul donc les équation en régime stationnaire deviennent :

$$\begin{aligned} \lambda p_0 &= \mu p_1 \\ (\lambda + \mu) p_1 &= \lambda p_0 + 2\mu p_2 \\ (\lambda + 2\mu) p_2 &= \lambda p_1 + 3\mu p_3 \\ &\vdots = \vdots \\ (\lambda + (k-1)\mu) p_{k-1} &= \lambda p_{k-2} + k\mu p_k \\ (\lambda + k\mu) p_k &= \lambda p_{k-1} + k\mu p_{k+1} \\ &\vdots = \vdots \end{aligned}$$

Autrement dit : les arrivées se passent toujours suivant un taux λ . Mais les taux de départ dépendent de l'état du système : si le système est dans l'état n , avec $n \leq k$ le taux de départ est $n\mu$, alors que si $n \geq k$, ce taux vaut $k\mu$. En effet, si n serveurs sont en service simultanément, et si les temps de service sont indépendants et de même loi exponentielle de moyenne $1/\mu$, la loi du premier départ est encore une loi exponentielle, de moyenne $1/(n\mu)$. C'est la loi du minimum des n exponentielles en question.

La solution du système existe si $\lambda < k\mu$ et on a alors $p_n = (\lambda/\mu)^n/n!$ si $n \leq k$ et $p_n = (\lambda/(k\mu))^n k^k p_0/k!$ où la condition $\sum_{n \geq 0} p_n = 1$ détermine p_0 .

La probabilité de trouver tous les serveurs occupés lorsque l'on arrive dans le service est

$$\sum_{n \geq k} p_n = \left(\frac{\lambda}{\mu}\right)^k \frac{k\mu}{k\mu - \lambda} p_0$$

et est appelée formule C d'Erlang.

Exercice : Écrivez les équations régissant les modèles M/M/k et M/M/1/N en régime stationnaire, et les résolvez-les.

Exercice : Le processus de sortie d'un système de file d'attente M/M/1 est un processus de Poisson. Quelle est son intensité?

3 Le processus de Poisson composé

Soit (Y_i) une suite de variables aléatoires indépendantes et identiquement distribuées et (T_n) la suite de ses sommes partielles : $T_0 = 0$ et pour tout $n \geq 1$, $T_{n+1} = T_n + Y_{n+1}$. On se donne également un processus de Poisson homogène $(R_t)_{t \geq 0}$, indépendant des (Y_i) . On appelle processus de Poisson composé le processus

$$N_t = T_{R_t} = \sum_{n \geq 1} T_n \mathbf{1}_{R_t = n} = \sum_{n \geq 1} Y_n \mathbf{1}_{R_t \geq n}.$$

Ce type de processus sert par exemple à modéliser le nombre d'animaux vivant dans une région donnée : le processus de Poisson décrit le nombre de terriers (par unité d'aire) et les Y_i représentent le nombre d'animaux par terrier.

Notons $G(s) = \mathbf{E}(s^{Y_1})$. La loi d'un processus de Poisson composé est caractérisée par sa fonction génératrice définie par

$$h(s, t) = \mathbf{E}(s^{N_t}) = \exp(-\lambda t(1 - G(s))).$$

En effet, on a

$$\begin{aligned} h(s, t) &= \mathbf{E}(s^{N_t}) \\ &= \sum_k \mathbf{E}(s^{T_k} \mathbf{1}_{R_t = k}) \\ &= \sum_k \mathbf{E}(s^{T_k}) \mathbf{P}(R_t = k) \\ &= \sum_k (\mathbf{E}(s^{Y_1}))^k \mathbf{P}(R_t = k) \\ &= \sum_k G(s)^k e^{-\lambda t} \frac{(\lambda t)^k}{k!} \\ &= \exp(-\lambda t + \lambda t G(s)) \end{aligned}$$

La connaissance de h permet naturellement d'obtenir facilement les moments de N_t . On a ainsi :

$$\mathbf{E}(N_t) = \frac{\partial}{\partial s} h(1, t) = \lambda t G'(1) = \lambda t \mathbf{E}(Y_1)$$

et

$$\text{var}(N_t) = \frac{\partial^2}{\partial s^2} h(1, t) + \mathbf{E}(N_t) - (\mathbf{E}(N_t))^2 = \lambda t (G''(1) + G'(1)) = \lambda t \mathbf{E}(Y_1^2).$$

Remarquons que h vérifie :

Proposition 4.7 *Pour tous nombres réels positifs t_1 et t_2 , on a $h(s, t_1 + t_2) = h(s, t_1)h(s, t_2)$.*

Réciproquement, dès que la fonction génératrice d'un processus vérifie cette relation, c'est la fonction génératrice d'un processus de Poisson composé.

Exemple :

- La loi binomiale négative, qui est, rappelons-le, la loi du r -ième succès dans un tirage aléatoire, a pour fonction génératrice

$$h(s, r) = \left(\frac{p}{1 - qs} \right)^r.$$

C'est une loi de Poisson composé avec

$$\lambda = -\log p, \quad G_{Y_1} = \frac{1}{\lambda} \log \frac{1}{1 - qs}.$$

En décomposant la fonction ci-dessus en série entière, on obtient la loi de Y_1 : elle est donnée par $\mathbf{P}(Y_1 = n) = \lambda q^n / n$. Il s'agit de la loi logarithmique.

- On classe les accidents de voitures suivant le nombre de véhicules impliqués : on suppose que le nombre d'accidents survenus au cours de l'intervalle de temps $[0, t]$ et impliquant k véhicules suit une loi de Poisson de paramètre $\lambda_k t$, et que ces différentes variables aléatoires sont indépendantes. Le nombre total de véhicules impliqués a pour fonction génératrice

$$h(s, t) = e^{-\lambda_1 t(1-s)} e^{-\lambda_2 t(1-s^2)} e^{-\lambda_3 t(1-s^3)} \dots$$

En posant $\lambda = \sum \lambda_i$ et Y_1 de loi $\mathbf{P}(Y = n) = \lambda_n / \lambda$, on remarque que le processus de Poisson composé associé à λ et aux (Y_i) de même loi que Y admet cette même fonction génératrice. Réciproquement, on peut décomposer toute fonction génératrice d'un processus de Poisson composé sous la forme :

$$h(s, t) = \prod_{n \geq 1} e^{-\lambda t(1-s^n)}$$

c'est-à-dire que tout processus de Poisson composé peut être vu comme un effet cumulatif.

Table des matières

1	Chaînes de Markov	2
1	Définitions, propriétés	2
1.1	Le modèle	2
1.2	Les contraintes du modèle	4
1.3	Irréductibilité, stationnarité	4
1.4	Transience, récurrence, période	5
1.5	Convergence	8
1.6	Fonction de coût	8
1.7	Simulation	9
2	Marches aléatoires au plus proche voisin	10
2.1	Récurrence et transience	10
2.2	Temps de retour	10
2	Espérance conditionnelle et martingales	14
1	Construction et propriétés de l'espérance conditionnelle	14
1.1	Définitions	14
1.2	Propriétés	15
2	Martingales en temps discret	15
2.1	Définition, exemples	15
2.2	Temps d'arrêt	16
2.3	Convergence	19
3	Martingales en temps continu	19
3.1	Tribus, filtrations	19
3.2	Les processus càd-làg	19
3.3	Martingales	19
3.4	Martingales locales	21
3.5	Variation quadratique	21
3.6	Convergence	21
3	Mouvement brownien, intégration stochastique	22
1	Processus gaussiens	22
1.1	Vecteurs gaussiens	22
1.2	Conditionnement	24
1.3	Convergence des séries gaussiennes	25
2	Mouvement brownien	26
2.1	Introduction	26
2.2	Construction	26
2.3	Propriétés	28
2.4	Martingale et temps d'arrêt	30

2.5	Simulation	32
3	Construction de l'intégrale stochastique, formule de Itô	32
3.1	Intégration par rapport au mouvement brownien	33
3.2	Intégrale stochastique par rapport à une martingale continue	36
3.3	Généralisations	37
3.4	Formule de Itô	38
3.5	Preuve de la formule de Itô.	40
3.6	Formule de Girsanov	42
4	Processus de renouvellement et de Poisson	45
1	Processus de renouvellement	45
1.1	Définition	45
1.2	Quelques résultats élémentaires	45
1.3	Le théorème du renouvellement	47
1.4	Le théorème central limite pour les processus de renouvellement	47
2	Processus de Poisson	47
2.1	Définition	47
2.2	Quelques propriétés	47
2.3	Processus de Poisson inhomogène	48
2.4	Simulation des processus de Poisson	48
2.5	Files d'attente	49
2.6	Modélisation	49
3	Le processus de Poisson composé	52