

Processus Aléatoires

JEAN FRANÇOIS DELMAS BENJAMIN JOURDAIN
BERNARD LAPEYRE

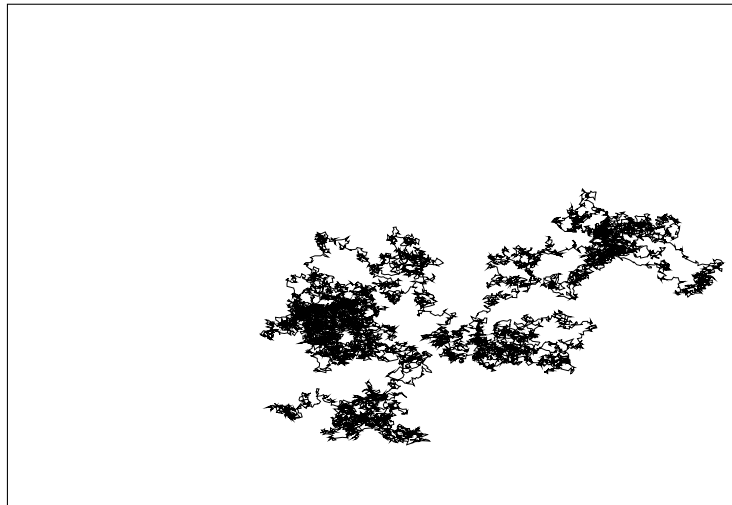


Table des matières

1	Rappels de Probabilités	7
1.1	Notion de tribu et de variables aléatoires	7
1.2	Espérance d'une variable aléatoire	8
1.3	Théorèmes de convergences pour les espérances	10
1.4	Loi d'une variable aléatoire réelle	12
1.5	Loi d'un vecteur aléatoire	13
1.6	Convergence presque sûre et théorèmes liés	16
1.7	Convergence en loi d'une famille de variables aléatoires	18
1.8	Autres type de convergence	20
1.9	Problème corrigé	22
1.10	Exercices	29
2	Vecteurs gaussiens - Mouvement Brownien	33
2.1	Vecteurs gaussiens	33
2.2	Mouvement Brownien	37
2.3	Vers une construction du mouvement brownien	39
2.4	Régularité des trajectoires	41
2.5	Caractère gaussien du mouvement brownien	42
2.6	Exercices	44
2.7	Travail dirigé : Test d'adéquation à une loi	46
3	Espérance conditionnelle	49
3.1	Tribus construites à partir de variables aléatoires	49
3.2	Notion d'espérance conditionnelle	50
3.3	Cas d'un couple gaussien	54
3.4	Travail dirigé : Estimation bayésienne	56
3.5	Exercices	58
4	Martingales à temps discrets	61
4.1	Introduction à la notion de martingale	61
4.2	Exemples	62
4.3	Un exemple générique de martingale	64
4.4	Notion de temps d'arrêt	65
4.5	Théorème de convergence des martingales	67
4.6	Convergence de Martingales et algorithmes stochastiques	69
4.7	Travail dirigé : Temps de sortie d'une marche aléatoire	75
4.8	Exercices	77

5	Chaînes de Markov à temps discret	79
5.1	Chaîne de Markov	79
5.2	Calcul de lois	82
5.3	Exemple d'utilisation des chaînes de Markov	84
5.4	Chaînes de Markov et espérance conditionnelle	86
5.4.1	Solution d'un problème d'arrêt optimal	88
5.5	Exercices	92
5.6	Travail dirigé : Modélisation de l'évolution d'une population	95
5.7	Travail dirigé : Algorithme de Hastings-Metropolis	97
5.8	Travail dirigé : Récurrence de la marche aléatoire simple	98
6	Propriété de Markov forte	101
6.1	Propriété de Markov générale	101
6.2	Introduction à la propriété de Markov forte	102
6.3	Tribu des événements antérieurs à un temps d'arrêt	102
6.4	Propriété de Markov forte : première approche	105
6.5	Exercices	107
6.6	Contrôle : Loi du supremum d'une marche aléatoire	109
7	Options européennes dans le modèle de Cox-Ross-Rubinstein	111
7.1	Le modèle	112
7.1.1	description	112
7.1.2	liens entre les paramètres r , a et b	113
7.1.3	le cas d'une seule période de temps : $N = 1$	114
7.2	Portefeuilles, arbitrages	115
7.2.1	portefeuille	115
7.2.2	Absence d'Opportunités d'Arbitrage	116
7.3	Pricing des options européennes	118
8	Mouvement brownien et martingales	121
8.1	Généralités sur les processus à temps continu	121
8.2	Extensions de la définition du mouvement brownien	122
8.3	Exemples de martingales browniennes	122
8.4	Exemples d'utilisation de la propriété de martingale	123
8.5	Propriété de Markov forte du mouvement brownien	125
8.6	Exercices	127
8.7	Travail dirigé : Principe de réflexion	129

Table des figures

1.1	Densité $e^{-\frac{1}{2}x^2}/\sqrt{2\pi}$ de $\mathcal{N}(0, 1)$	13
1.2	Fonction de répartition	13
1.3	Loi de Y_n	25
1.4	Loi de $Y_{2n} - Y_n$	25
1.5	Loi de Z_n	26
3.1	Projection orthogonale sur \mathcal{H}_B	50
4.1	Trajectoire de la marche aléatoire symétrique	62
4.2	Temps de sortie de $[a, b]$ par une marche aléatoire	75
5.1	Histogramme de la loi du nombre maximum de F consécutifs après 100 tirages	85
5.2	Probabilité d'avoir au moins n F consécutifs après 100 tirages	85
8.1	Principe de réflexion	129

Chapitre 1

Rappels de Probabilités

1.1 Notion de tribu et de variables aléatoires

Définition 1.1.1 Soit Ω un ensemble et \mathcal{A} un sous ensemble de l'ensemble $\mathcal{P}(\Omega)$ des parties de Ω . On dit que \mathcal{A} est une *tribu* si cet ensemble est stable par les opérations ensemblistes naturelles, plus précisément :

- stabilité par \cap , \cup et passage au complémentaire : si A et B appartiennent à \mathcal{A} , alors $A \cap B$, $A \cup B$ et A^c appartiennent à \mathcal{A} .
- stabilité par réunion et intersection *dénombrables* : si pour tout $i \in \mathbf{N}^*$, $A_i \in \mathcal{A}$ alors $\bigcup_{i \geq 1} A_i$ et $\bigcap_{i \geq 1} A_i$ sont dans \mathcal{A} .
- $\emptyset, \Omega \in \mathcal{A}$.

Remarque 1.1.2 \mathcal{A} représente une information disponible.

Exemples

- $\mathcal{A} = \{\emptyset, \Omega\}$ est la plus petite tribu. On l'appelle la tribu triviale. Elle représente l'absence totale d'information.
- $\mathcal{A} = \mathcal{P}(\Omega)$ est une tribu appelée tribu discrète qui représente l'information totale.
- Soit (B_1, \dots, B_n) une partition de Ω , alors :

$$\mathcal{A} = \left\{ \bigcup_{\text{réunion finie}} B_{i_k} \right\},$$

est une tribu.

- Soit $\Omega = \mathbf{R}$, on appelle tribu borélienne la plus petite tribu contenant les intervalles de \mathbf{R} . On la note $\mathcal{B}(\mathbf{R})$.

Remarque 1.1.3 Il est facile de vérifier que l'intersection d'une famille quelconque de tribus reste une tribu. La tribu borélienne de \mathbf{R} est donc définie comme l'intersection de toutes les tribus qui contiennent les intervalles de \mathbf{R} . Comme tout ouvert de \mathbf{R} s'exprime comme union dénombrable d'intervalles ouverts disjoints, la tribu borélienne contient tous les ouverts de \mathbf{R} et tous les fermés par passage au complémentaire.

En fait la tribu borélienne $\mathcal{B}(\mathbf{R})$ est strictement plus petite que l'ensemble de toutes les parties de \mathbf{R} , mais la preuve de ce résultat délicat repose sur l'utilisation de l'axiome du choix.

La notion de tribu permet de préciser la définition d'une variable aléatoire.

Définition 1.1.4 Une application X de Ω dans \mathbf{R} est une variable aléatoire *mesurable par rapport à la tribu \mathcal{A}* , si pour tout $B \in \mathcal{B}(\mathbf{R})$:

$$\{\omega \in \Omega, X(\omega) \in B\} = \{X \in B\} \in \mathcal{A}.$$

On dit alors que X est une variable aléatoire \mathcal{A} -mesurable.

Définition 1.1.5 Une probabilité \mathbf{P} sur (Ω, \mathcal{A}) est une mesure positive de masse totale 1 définie sur \mathcal{A} . Cela signifie que :

- pour tout $A \in \mathcal{A}$, $\mathbf{P}(A) \in [0, 1]$ est défini,
- $\mathbf{P}(\Omega) = 1$,
- si pour tout entier $i \geq 1$, $A_i \in \mathcal{A}$ et la famille des A_i est disjointe, alors :

$$\mathbf{P}(\cup_{i \geq 1} A_i) = \sum_{i \geq 1} \mathbf{P}(A_i).$$

Le triplet $(\Omega, \mathcal{A}, \mathbf{P})$ s'appelle espace de probabilité.

On voit que pour une variable aléatoire \mathcal{A} -mesurable on peut définir $\mathbf{P}(X \in A)$ pour tout A , borélien de \mathbf{R} .

Remarque 1.1.6 En choisissant $A_i = \emptyset$ pour $i \geq n + 1$, on obtient que si A_1, \dots, A_n sont disjoints $\mathbf{P}(\cup_{i=1}^n A_i) = \sum_{i=1}^n \mathbf{P}(A_i)$.

Définition 1.1.7 On dit qu'une propriété est vraie presque sûrement, si cette propriété est vérifiée avec probabilité 1. On dit ainsi qu'une suite $(X_n, n \geq 1)$ converge presque sûrement si :

$$\mathbf{P}\left(\lim_{n \rightarrow +\infty} X_n \text{ existe}\right) = 1.$$

1.2 Espérance d'une variable aléatoire

La notion d'espérance est la formalisation du concept de moyenne.

Cas des variables aléatoires positives Soit $(\Omega, \mathcal{A}, \mathbf{P})$ un espace de probabilité. Soit X une variable aléatoire réelle à valeurs $\mathbf{R}_+ \cup \{+\infty\}$ mesurable par rapport à la tribu \mathcal{A} (en plus de la définition 1.1.4, on suppose simplement que $\{X = +\infty\} \in \mathcal{A}$). Notez bien que l'on accepte que X puisse prendre la valeur $+\infty$ avec probabilité non nulle

Si X prend un nombre fini de valeurs de \mathbf{R} , $\{x_1, \dots, x_n\}$, on définit l'espérance naturellement par

$$\mathbf{E}(X) := \sum_{i=1}^n x_i \mathbf{P}(X = x_i).$$

On étend l'espérance à toute variable aléatoire X positive en "approchant" X par la suite croissante $(X_n, n \geq 1)$ donnée par

$$X_n = \sum_{k=0}^{n2^n-1} \frac{k}{2^n} \mathbf{1}_{\left\{\frac{k}{2^n} \leq X < \frac{k+1}{2^n}\right\}} + n \mathbf{1}_{\{X \geq n\}}.$$

Notez que pour un n fixé on sait donner une valeur à $\mathbf{E}(X_n)$ (car X_n prend un nombre fini de valeurs)

$$\mathbf{E}(X_n) = \sum_{k=0}^{n2^n} \frac{k}{2^n} \mathbf{P}\left(\frac{k}{2^n} \leq X < \frac{k+1}{2^n}\right) + n \mathbf{P}(X \geq n).$$

On définit alors l'espérance $\mathbf{E}(X)$ par

$$\mathbf{E}(X) = \lim_{n \rightarrow \infty} \mathbf{E}(X_n).$$

La limite existe forcément comme limite croissante de nombres réels positifs. Elle peut être égale à $+\infty$. Dans tous les cas elle définit l'espérance de la variable aléatoire positive X .

Notez que si $\mathbf{P}(X = +\infty) > 0$, on a par construction pour tout n

$$\mathbf{E}(X) \geq n\mathbf{P}(X \geq n) \geq n\mathbf{P}(X = +\infty).$$

On a donc $\mathbf{E}(X) = +\infty$. Par contraposée, on voit qu'une variable aléatoire positive d'espérance finie est finie avec probabilité 1. Autrement dit

$$\mathbf{E}(X) < +\infty \text{ implique } \mathbf{P}(X < +\infty) = 1. \quad (1.1)$$

Exercice 1 Vérifier à partir de la construction de l'espérance que, si $0 \leq X \leq Y$ on a $\mathbf{E}(X) \leq \mathbf{E}(Y)$.

Cas des variables aléatoire de signe quelconque Si X est une v.a. réelle de signe quelconque, on dit qu'elle est intégrable si $\mathbf{E}(|X|) < \infty$ ($\mathbf{E}(|X|)$ est toujours définie d'après ce qui précède). Notez que 1.1 implique que $|X|$ est finie presque sûrement. Comme

$$X\mathbf{1}_{X \geq 0} \leq |X| \text{ et } (-X)\mathbf{1}_{X < 0} \leq |X|$$

On a, par croissance de l'espérance, $\mathbf{E}(X\mathbf{1}_{X \geq 0}) \leq \mathbf{E}(|X|) < +\infty$ et $\mathbf{E}((-X)\mathbf{1}_{X < 0}) \leq \mathbf{E}(|X|) < +\infty$. On peut donc définir l'espérance par la différence de deux nombres réels (finis !)

$$\mathbf{E}(X) = \mathbf{E}(X\mathbf{1}_{X \geq 0}) - \mathbf{E}((-X)\mathbf{1}_{X < 0}).$$

On admet les propriétés suivantes de l'espérance.

1. Linéarité. Soient X et Y deux v.a. intégrables, soient $\alpha, \beta \in \mathbf{R}$. Alors la v.a. $\alpha X + \beta Y$ est intégrable, et on a

$$\mathbf{E}[\alpha X + \beta Y] = \alpha \mathbf{E}[X] + \beta \mathbf{E}[Y].$$

2. $\mathbf{E}(\mathbf{1}_A) = \mathbf{P}(A)$.
3. Positivité. Soit X une v.a. réelle positive p.s., c'est-à-dire telle que $\mathbf{P}(X \geq 0) = 1$, alors on a $\mathbf{E}(X) \in [0, \infty]$. En particulier $\mathbf{E}(X) \geq 0$.
4. Soient X et Y deux v.a. réelles intégrables telles que $X \leq Y$ p.s., c'est-à-dire telles que $\mathbf{P}(X \leq Y) = 1$, alors on a

$$\mathbf{E}(X) \leq \mathbf{E}(Y).$$

1.3 Théorèmes de convergences pour les espérances

Les théorèmes suivants sont fondamentaux lorsque l'on cherche à justifier proprement que :

$$\lim_{n \rightarrow \infty} \mathbf{E}(X_n) = \mathbf{E} \left(\lim_{n \rightarrow \infty} X_n \right).$$

Cette égalité bien que naturelle ne va pas de soi et est en général fautive, même si on peut la justifier dans de nombreux cas. L'espérance d'une variable aléatoire positive X a été construite comme la limite des espérances de variables aléatoires X_n qui convergent vers X en croissant. C'est pourquoi le théorème suivant que nous admettrons est naturel.

Théorème 1.3.1 (Théorème de convergence monotone) Si $(X_n, n \geq 1)$ est une suite de variables aléatoires croissantes positives, i.e.

$$0 \leq X_n \leq X_{n+1},$$

on peut affirmer que :

$$\mathbf{E} \left(\lim_{n \rightarrow +\infty} X_n \right) = \lim_{n \rightarrow +\infty} \mathbf{E}(X_n).$$

Noter que l'on n'impose ni l'intégrabilité des X_n ni celle de $\lim_{n \rightarrow +\infty} X_n$, ce qui rend ce théorème très pratique : rien n'interdit aux limites à droite et à gauche de l'égalité de prendre la valeur $+\infty$. Si l'une vaut $+\infty$ cela implique que l'autre vaut aussi $+\infty$.

Remarque 1.3.1 Une application directe de ce dernier résultat montre que l'on a *toujours* :

$$\sum_{n \geq 0} \mathbf{E}(|X_n|) = \mathbf{E} \left(\sum_{n \geq 0} |X_n| \right).$$

Lorsque l'on peut prouver que $\sum_{n \geq 0} \mathbf{E}(|X_n|) < +\infty$, on est donc sûr que

$$\mathbf{E} \left(\sum_{n \geq 0} |X_n| \right) < +\infty.$$

La variable aléatoire $\sum_{n \geq 0} |X_n|$ est donc finie presque sûrement et la série $\sum_{n \geq 0} X_n$ est donc forcément (absolument) convergente. En particulier, X_n tend presque sûrement vers 0.

On démontrerait de façon identique que si, pour un entier p plus grand que 1

$$\sum_{n \geq 0} \mathbf{E}(|X_n|^p) < +\infty,$$

alors $\lim_{n \rightarrow +\infty} X_n = 0$. Ce résultat (très facile à obtenir !) est l'un des rares moyens qui permet de prouver la convergence presque sûre.

Le résultat le plus important pour passer à la limite dans les espérances est le

Théorème 1.3.2 (Théorème de convergence dominée (ou de Lebesgue)) On suppose que la suite $(X_n, n \geq 1)$ vérifie

– **convergence presque sûre :**

$$\mathbf{P} \left(\lim_{n \rightarrow +\infty} X_n \text{ existe} \right) = 1.$$

– **domination :** pour tout n ,

$$|X_n| \leq \hat{X}, \text{ avec } \mathbf{E}(\hat{X}) < +\infty.$$

On peut alors affirmer que $\lim_{n \rightarrow +\infty} X_n$ est intégrable et que :

$$\mathbf{E} \left(\lim_{n \rightarrow +\infty} X_n \right) = \lim_{n \rightarrow +\infty} \mathbf{E}(X_n).$$

Nous allons montrer que le théorème de convergence dominée découle du théorème de convergence monotone. Pour cela nous allons commencer par établir le

Théorème 1.3.3 (Lemme de Fatou) Soit $(X_n)_{n \in \mathbb{N}^*}$ une suite de v.a. positives p.s. On a :

$$\mathbf{E} \left(\lim_{n \rightarrow \infty} \inf_{k \geq n} X_k \right) \leq \lim_{n \rightarrow \infty} \inf_{k \geq n} \mathbf{E}(X_k).$$

Démonstration : On pose $Y_n = \inf_{k \geq n} X_k$. La suite (Y_n) est une suite croissante de variables positives. Par le théorème de convergence monotone,

$$\mathbf{E}(\lim_n Y_n) = \lim_n \mathbf{E}(Y_n).$$

Pour $k \geq n$, $Y_n \leq X_k$ et par croissance de l'espérance, $\mathbf{E}(X_k) \geq \mathbf{E}(Y_n)$. Donc $\inf_{k \geq n} \mathbf{E}(X_k) \geq \mathbf{E}(Y_n)$. En passant à la limite $n \rightarrow +\infty$ dans cette inégalité, on obtient

$$\lim_{n \rightarrow \infty} \inf_{k \geq n} \mathbf{E}(X_k) \geq \lim_n \mathbf{E}(Y_n) = \mathbf{E}(\lim_n Y_n).$$

■

On peut alors démontrer le théorème de convergence dominée :

Démonstration : On note $X = \lim_n X_n$. Par passage à la limite dans l'inégalité de domination $|X_n| \leq \hat{X}$, on obtient que $|X| \leq \hat{X}$, ce qui assure l'intégrabilité de X . On a également $|X - X_n| \leq |X| + |X_n| \leq 2\hat{X}$. Ainsi les variables $Y_n = 2\hat{X} - |X - X_n|$ sont positives. Le lemme de Fatou implique alors

$$\mathbf{E} \left(\lim_{n \rightarrow \infty} \inf_{k \geq n} (2\hat{X} - |X - X_k|) \right) \leq \lim_{n \rightarrow \infty} \inf_{k \geq n} \mathbf{E}(2\hat{X} - |X - X_k|),$$

inégalité qui se réécrit

$$2\mathbf{E}(\hat{X}) - \mathbf{E} \left(\lim_{n \rightarrow \infty} \sup_{k \geq n} |X - X_k| \right) \leq 2\mathbf{E}(\hat{X}) - \lim_{n \rightarrow \infty} \sup_{k \geq n} \mathbf{E}(|X - X_k|).$$

Comme par convergence de X_n vers X , $\sup_{k \geq n} |X - X_k|$ converge presque sûrement vers 0 lorsque $n \rightarrow +\infty$, on en déduit que

$$\lim_{n \rightarrow \infty} \sup_{k \geq n} \mathbf{E}(|X - X_k|) \leq 0$$

et donc que $\lim_n \mathbf{E}(|X - X_n|) = 0$. On conclut en utilisant l'inégalité

$$|\mathbf{E}(X) - \mathbf{E}(X_n)| \leq \mathbf{E}(|X - X_n|)$$

qui s'obtient par linéarité et croissance de l'espérance. ■

1.4 Loi d'une variable aléatoire réelle

Définition 1.4.1 Soit X une variable aléatoire à valeurs dans \mathbf{R} . On dit que cette variable aléatoire admet une densité $p(x)$ ($p(x)$ étant une fonction positive d'intégrale sur \mathbf{R} égale à 1) si pour toute fonction $f : \mathbf{R} \rightarrow \mathbf{R}$ bornée ou positive, on a :

$$\mathbf{E}(f(X)) = \int_{\mathbf{R}} f(x)p(x)dx.$$

Remarque 1.4.2 Notons que la loi d'une variable aléatoire est caractérisée par :

- sa fonction de répartition :

$$u \in \mathbf{R} \rightarrow F_X(u) = \mathbf{P}(X \leq u),$$

- ou sa fonction caractéristique :

$$u \in \mathbf{R} \rightarrow \Phi_X(u) = \mathbf{E}(e^{iuX}).$$

Si X admet la densité p , $F_X(u) = \int_{-\infty}^u p(x)dx$ et $\Phi_X(u) = \int_{\mathbf{R}} e^{iux}p(x)dx$.

Exemple de loi : la loi gaussienne Un exemple de loi à densité sur \mathbf{R} très utile est la loi gaussienne. On dit que G suit une loi gaussienne centrée réduite si pour toute fonction f bornée :

$$\mathbf{E}(f(G)) = \int_{\mathbf{R}} f(x)e^{-\frac{x^2}{2}} \frac{dx}{\sqrt{2\pi}}.$$

La fonction caractéristique de cette loi vaut :

$$\mathbf{E}(e^{iuG}) = e^{-\frac{u^2}{2}}.$$

la fonction de répartition n'admet pas de formule explicite et on la note souvent N :

$$N(x) = \int_{-\infty}^x e^{-\frac{u^2}{2}} \frac{du}{\sqrt{2\pi}}.$$

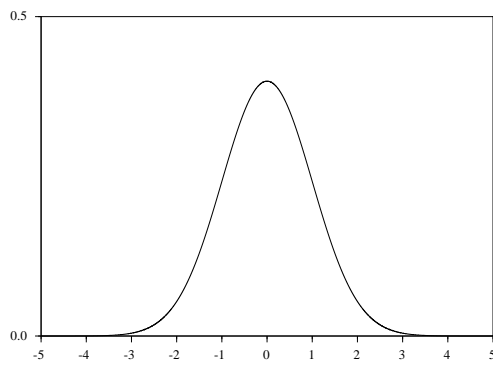
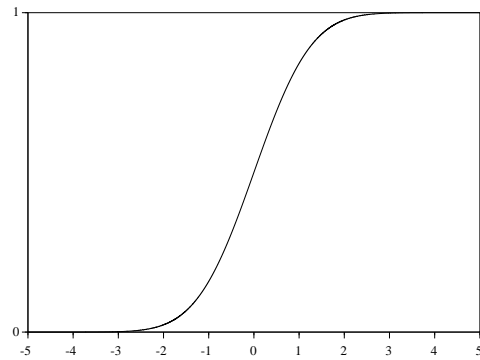
FIG. 1.1 – Densité $e^{-\frac{1}{2}x^2}/\sqrt{2\pi}$ de $\mathcal{N}(0, 1)$ 

FIG. 1.2 – Fonction de répartition

On dit que X suit une loi gaussienne de moyenne m et d'écart-type σ (ou de variance¹ σ^2) si, pour toute fonction positive :

$$\mathbf{E}(f(X)) = \int_{-\infty}^{+\infty} f(x) e^{-\frac{(x-m)^2}{2\sigma^2}} \frac{dx}{\sigma\sqrt{2\pi}}.$$

On vérifie que $\mathbf{E}(X) = m$ et $\text{Var}(X) = \sigma^2$ et l'on dit que X suit la loi $\mathcal{N}(m, \sigma^2)$.

Notons que si X suit une loi gaussienne de moyenne m et de variance σ^2 , alors $G = (X - m)/\sigma$ est une gaussienne de moyenne 0 et de variance 1 (on dit aussi centrée réduite). On voit donc que l'on peut toujours écrire une gaussienne (m, σ^2) sous la forme $X = m + \sigma G$, G étant une gaussienne centrée réduite. Cette remarque est souvent utile pour simplifier certains calculs.

Il est facile d'expliciter la fonction caractéristique d'une gaussienne (m, σ^2) en effet :

$$\mathbf{E}(e^{iuX}) = \mathbf{E}(e^{iu(m+\sigma G)}) = e^{ium} \mathbf{E}(e^{iu\sigma G}) = e^{ium} e^{-\frac{(u\sigma)^2}{2}} = e^{iu\mathbf{E}(X)} e^{-\frac{u^2}{2} \text{Var}(X)}.$$

1.5 Loi d'un vecteur aléatoire

Définition 1.5.1 Soit $X = (X_1, \dots, X_d)$ une variable aléatoire prenant ses valeurs dans \mathbf{R}^d . On dit que X admet une densité $p(x_1, \dots, x_d)$, si pour toute fonction positive ou bornée f de \mathbf{R}^d dans \mathbf{R} on a :

$$\mathbf{E}(f(X_1, \dots, X_d)) = \int_{\mathbf{R}^d} f(x_1, \dots, x_d) p(x_1, \dots, x_d) dx_1 \dots dx_d.$$

La densité d'un vecteur aléatoire permet de calculer la densité de chacune de ses composantes grâce à la formule des densités marginales :

Proposition 1.5.2 Si le vecteur aléatoire $X = (X_1, \dots, X_d)$ possède la densité $p(x_1, \dots, x_d)$ alors pour tout $1 \leq i \leq d$, X_i possède la densité $p_i(x_i)$ obtenue à partir de p en intégrant sur les autres variables :

$$p_i(x_i) = \int_{\mathbf{R}^{d-1}} p(x_1, \dots, x_d) dx_1 \dots dx_{i-1} dx_{i+1} \dots dx_d.$$

¹La variance d'une variable aléatoire est donnée par :

$$\text{Var}(X) = \mathbf{E}(X^2) - \mathbf{E}(X)^2 = \mathbf{E}\{(X - \mathbf{E}(X))^2\}.$$

La fonction de répartition d'un vecteur est rarement utilisée, mais on peut définir la *fonction caractéristique* d'un vecteur par, pour tout $\mathbf{u} = (u_1, \dots, u_d)$ vecteur de \mathbf{R}^d :

$$\Phi_X(\mathbf{u}) = \mathbf{E} (e^{i\mathbf{u} \cdot X}) = \int_{\mathbf{R}^n} e^{i\mathbf{u} \cdot x} p(x) dx,$$

avec $\mathbf{u} \cdot x = u_1 x_1 + \dots + u_d x_d$ et $dx = dx_1 \dots dx_d$.

La fonction caractéristique caractérise la loi du vecteur X : on peut montrer que si X et Y sont deux vecteurs à valeurs dans \mathbf{R}^d tels que pour tout $\mathbf{u} \in \mathbf{R}^d$:

$$\mathbf{E} (e^{i\mathbf{u} \cdot X}) = \mathbf{E} (e^{i\mathbf{u} \cdot Y}),$$

alors la loi de X est identique à celle de Y et donc pour toute fonction réelle positive ou bornée f sur \mathbf{R}^d :

$$\mathbf{E} (f(X)) = \mathbf{E} (f(Y)).$$

La notion d'indépendance est centrale en probabilité. C'est la façon naturelle de spécifier un modèle d'évolution.

Commençons par définir l'indépendance de deux variables aléatoires.

Définition 1.5.3 On dit que deux variables aléatoires réelles X et Y sont indépendantes, si l'on a l'une des deux propriétés équivalentes suivantes

- pour tout A et B ensemble (de la tribu borélienne pour les amateurs !)

$$\mathbf{P}(X \in A, Y \in B) = \mathbf{P}(X \in A) \mathbf{P}(Y \in B),$$

- Pour tout f et g fonctions bornées (et mesurable pour les mêmes amateurs) de \mathbf{R} dans \mathbf{R}

$$\mathbf{E} (f(X)g(Y)) = \mathbf{E} (f(X)) \mathbf{E} (g(Y)).$$

On peut généraliser cette définition à n variables aléatoires.

Définition 1.5.4 On dit que n variables aléatoires réelles (X_1, \dots, X_n) sont indépendantes, si l'on a l'une des deux propriétés équivalentes suivantes

- pour tout ensemble A_1, \dots, A_n

$$\mathbf{P}(X_1 \in A_1, \dots, X_n \in A_n) = \mathbf{P}(X_1 \in A_1) \dots \mathbf{P}(X_n \in A_n),$$

- Pour tout $(f_k, 1 \leq k \leq n)$ fonctions bornées de \mathbf{R} dans \mathbf{R}

$$\mathbf{E} (f_1(X_1) \dots f_n(X_n)) = \mathbf{E} (f_1(X_1)) \dots \mathbf{E} (f_n(X_n)).$$

Définition 1.5.5 On dit que $(X_n, n \geq 0)$ est une suite de variables aléatoires indépendantes si, tout vecteur de longueur fini est constitué de variables aléatoires indépendantes.

On peut caractériser l'indépendance d'un vecteur aléatoire sur la densité de la loi. Le résultat est exprimé pour un couple de variables aléatoires mais s'étend sans difficulté au cas vectoriel.

Proposition 1.5.6 Un couple (X, Y) de variables aléatoires réelles admettant une loi de densité $p(x, y)$, est un couple de variables aléatoires indépendantes si et seulement si, on peut mettre p sous la forme

$$p(x, y) = p_1(x)p_2(y),$$

p_1 et p_2 étant deux densités de loi à valeurs réelles. En outre, X admet alors la densité p_1 et Y la densité p_2 .

Démonstration : Supposons d'abord que $p(x, y) = p_1(x)p_2(y)$. Soit f et g deux fonctions bornées

$$\mathbf{E}(f(X)g(Y)) = \int_{\mathbf{R}^2} f(x)g(y)p(x, y)dx dy = \int_{\mathbf{R}^2} f(x)g(y)p_1(x)p_2(y)dx dy.$$

En utilisant le théorème de Fubini pour des intégrales réelles on obtient

$$\mathbf{E}(f(X)g(Y)) = \int_{\mathbf{R}} f(x)p_1(x)dx \int_{\mathbf{R}} g(y)p_2(y)dy.$$

En faisant $g = 1$ on obtient

$$\mathbf{E}(f(X)) = \int_{\mathbf{R}} f(x)p_1(x)dx,$$

et de même

$$\mathbf{E}(g(Y)) = \int_{\mathbf{R}} g(y)p_2(y)dy.$$

On a finalement

$$\mathbf{E}(f(X)g(Y)) = \mathbf{E}(f(X))\mathbf{E}(g(Y)).$$

X et Y sont donc indépendantes.

Réciproquement supposons que X et Y soient indépendantes. Commençons par construire p_1 et p_2 . Pour cela, notons que, en toute généralité, on a

$$\mathbf{E}(f(X)) = \int_{\mathbf{R}^2} f(x)p(x, y)dx dy = \int_{\mathbf{R}} f(x) \left(\int_{\mathbf{R}} p(x, y)dy \right) dx.$$

Ce qui prouve que la loi de X admet une densité donnée par

$$p_1(x) = \int_{\mathbf{R}} p(x, y)dy.$$

De même la densité de la loi de Y est donnée par

$$p_2(y) = \int_{\mathbf{R}} p(x, y)dx.$$

Pour simplifier la preuve, on supposera que p , p_1 et p_2 sont des fonctions continues. On a

$$\mathbf{P}(X \in [x, x + \epsilon], Y \in [y, y + \epsilon]) = \mathbf{P}(X \in [x, x + \epsilon]) \mathbf{P}(Y \in [y, y + \epsilon]).$$

Donc :

$$\int_x^{x+\epsilon} \int_y^{y+\epsilon} p(u, v)du dv = \int_x^{x+\epsilon} p_1(u)du \int_y^{y+\epsilon} p_2(v)dv.$$

On peut alors diviser par ϵ^2 et passer à la limite lorsque ϵ tend vers 0, pour obtenir, pour tout x, y :

$$p(x, y) = p_1(x)p_2(y).$$

■

Simulation d'une gaussienne centrée réduite Nous allons voir comment l'on peut simuler des variables aléatoires suivant la loi gaussienne centrée réduite. Pour cela, appelons G_1, G_2 deux gaussiennes centrées réduites indépendantes et notons R le rayon polaire :

$$R = \bar{r}(G_1, G_2) = \sqrt{G_1^2 + G_2^2},$$

et $\Theta = \bar{\theta}(G_1, G_2)$, l'angle polaire (il est plus prudent de ne pas expliciter la formule de l'angle polaire !). Soit f une fonction positive, on a :

$$\mathbf{E}(f(R, \Theta)) = \int_{\mathbf{R}^2} f(\bar{r}(x, y), \bar{\theta}(x, y)) e^{-\frac{x^2+y^2}{2}} \frac{1}{2\pi} dx dy.$$

Un changement de variable polaire permet alors d'affirmer que :

$$\mathbf{E}(f(R, \Theta)) = \int_{r>0, \theta \in [0, 2\pi[} f(r, \theta) e^{-\frac{r^2}{2}} \frac{1}{2\pi} r dr d\theta.$$

Ceci permet d'identifier la loi du couple (R, Θ) comme étant la loi de densité :

$$\mathbf{1}_{\{r > 0\}} r e^{-r^2/2} \times \frac{\mathbf{1}_{\{\theta \in [0, 2\pi[\}}}{2\pi}.$$

On en déduit que les deux variables aléatoires R et Θ sont indépendantes. Θ suit une loi uniforme sur $[0, 2\pi[$ et un changement de variable simple (exercice) montre que R^2 suit une loi exponentielle de paramètre $1/2$.

On en déduit que :

$$U_1 = e^{-R^2/2} \text{ et } U_2 = \frac{\Theta}{2\pi},$$

sont deux variables aléatoires indépendantes suivant des lois uniformes sur $[0, 1]$. Il est facile d'exprimer G_1 et G_2 à l'aide de U_1 et U_2 :

$$\begin{aligned} G_1 &= \sqrt{-2 \log(U_1)} \cos(2\pi U_2) \\ G_2 &= \sqrt{-2 \log(U_1)} \sin(2\pi U_2). \end{aligned}$$

Ceci conduit à la méthode classique de simulation d'une variable aléatoire gaussienne :

- tirer U_1 et U_2 indépendantes selon une loi uniforme sur $[0, 1]$,
- renvoyer $\sqrt{-2 \log(U_1)} \sin(2\pi U_2)$.

1.6 Convergence presque sûre et théorèmes liés

Rappelons l'un des théorèmes fondamentaux du calcul des probabilités : la Loi Forte des Grand Nombres.

Théorème 1.6.1 Soit $(X_i, i \geq 1)$ une suite de variables aléatoires indépendantes suivant toutes la même loi qu'une variable aléatoire X . On suppose que $\mathbf{E}(|X|) < +\infty$. Alors, pour presque tout ω :

$$\mathbf{E}(X) = \lim_{n \rightarrow +\infty} \frac{1}{n} (X_1(\omega) + \dots + X_n(\omega)).$$

Ce résultat est d'une importance centrale en statistique et les méthodes de Monte-Carlo en sont une application directe.

Remarque 1.6.1 La condition que $\mathbf{E}(|X|) < +\infty$ est évidemment essentielle. On vérifie, par exemple, que si $(X_i, i \geq 1)$ est une suite de variables aléatoires indépendantes suivant une loi de Cauchy standard, c'est à dire de densité donnée par :

$$\frac{1}{\pi(1+x^2)},$$

alors la loi de $(X_1 + \dots + X_n)/n$ est égale à celle de X_1 (on pourra démontrer ce résultat en exercice, en admettant que la fonction caractéristique de la loi de Cauchy est donnée par $\mathbf{E}(e^{iuX_1}) = e^{-|u|}$). On peut alors en déduire que $(X_1 + \dots + X_n)/n$ ne peut converger vers une constante.

Application On considère un modèle d'évolution du cours d'une action en temps discret. Entre l'instant n et l'instant $n+1$, on suppose le cours peut soit monter en étant multiplié par un facteur $\beta > 1$ soit baisser en étant multiplié par un facteur $\alpha \in]0, 1[$ mais que sa moyenne reste constante.

Pour formaliser cette idée, on se donne une suite $(T_i, i \geq 1)$ de variables aléatoires i.i.d. à valeurs dans $\{\alpha, \beta\}$ d'espérance $\mathbf{E}(T_i) = 1$. Un rapide calcul montre que

$$\mathbf{P}(T_i = \alpha) = p_\alpha = \frac{\beta - 1}{\beta - \alpha} \quad \text{et} \quad \mathbf{P}(T_i = \beta) = p_\beta = 1 - p_\alpha.$$

Le cours de l'action à l'instant n est

$$S_n = T_1 \times T_2 \times \dots \times T_n = \prod_{i=1}^n T_i.$$

Les variables $\ln(T_i)$ à valeurs dans $\{\ln(\alpha), \ln(\beta)\}$ sont intégrables. Comme pour tout $x > 0$ différent de 1, $\ln(x) < 1 - x$,

$$\mathbf{E}(\ln(T_i)) = p_\alpha \ln(\alpha) + p_\beta \ln(\beta) < p_\alpha(1 - \alpha) + p_\beta(1 - \beta) = 1 - \mathbf{E}(T_i) = 0.$$

En appliquant la loi forte des grands nombres, on obtient que $\frac{1}{n} \sum_{i=1}^n \ln(T_i)$ converge presque sûrement vers $\mathbf{E}(\ln(T_1)) < 0$ lorsque $n \rightarrow +\infty$. Donc $\sum_{i=1}^n \ln(T_i) = \ln(S_n)$ tend presque sûrement vers $-\infty$. Par composition avec la fonction exponentielle, on conclut que S_n converge presque sûrement vers 0.

Mais par indépendance des T_i , $\mathbf{E}(S_n) = \prod_{i=1}^n \mathbf{E}(T_i) = 1$. On a donc

$$1 = \lim_n \mathbf{E}(S_n) > \mathbf{E}(\lim_n S_n) = 0.$$

Cet exemple naturel montre que l'espérance de la limite n'est pas toujours égale à la limite des espérances. Il ne faut toutefois pas se décourager lorsque l'on doit calculer l'espérance d'une limite. Les théorèmes de convergence monotone et surtout de convergence dominée permettent de conclure dans bien des cas. L'exemple illustre simplement la nécessité de vérifier l'hypothèse de positivité et de croissance pour appliquer le théorème de convergence monotone et l'hypothèse de domination pour appliquer celui de convergence dominée.

Ici la suite S_n n'est pas croissante, ce qui explique que le théorème de convergence monotone ne s'applique pas. Et comme la conclusion du théorème de convergence dominée est infirmée, on en déduit que l'hypothèse de domination n'est pas vérifiée : $\mathbf{E}(\sup_n S_n) = +\infty$. Enfin on peut noter que l'inégalité donnée par le Lemme de Fatou est stricte.

1.7 Convergence en loi d'une famille de variables aléatoires

Définition 1.7.1 Soit $(X_n, n \geq 1)$ une suite de variables aléatoires réelles, on dit que cette suite converge en loi vers une variable aléatoire X si, pour toute fonction continue et bornée $f : \mathbf{R} \rightarrow \mathbf{R}$,

$$\lim_{n \rightarrow +\infty} \mathbf{E}(f(X_n)) = \mathbf{E}(f(X)).$$

On peut montrer le résultat, important mais délicat à prouver, suivant.

Proposition 1.7.2 $(X_n, n \geq 1)$ converge en loi vers une variable aléatoire X si et seulement si, pour tout réel u :

$$\lim_{n \rightarrow +\infty} \mathbf{E}(e^{iuX_n}) = \mathbf{E}(e^{iuX}).$$

Nous pouvons maintenant rappeler l'un des résultats fondamentaux du cours de probabilité de 1ère année : le théorème de la limite centrale.

Théorème 1.7.1 Soit $(X_n, n \geq 1)$ une suite de variables aléatoires indépendantes suivant toutes la même loi qu'une variable aléatoire X . On suppose que $\mathbf{E}(X^2) < +\infty$. On note σ^2 la variance de X , alors :

$$\sqrt{n} \left(\frac{1}{n} (X_1 + \dots + X_n) - \mathbf{E}(X) \right) \text{ converge en loi vers } \sigma G,$$

G étant une variable aléatoire suivant une loi gaussienne centrée réduite.

Démonstration : Quitte à retrancher $\mathbf{E}(X)$ à X , on se ramène au cas où X est une variable aléatoire de moyenne nulle. Notons alors :

$$\Phi_n(u) = \mathbf{E} \left(e^{iu \frac{X_1 + \dots + X_n}{\sqrt{n}}} \right).$$

En utilisant le fait que les X_n sont indépendantes et de même loi, on obtient :

$$\Phi_n(u) = \mathbf{E} \left(e^{iu \frac{X}{\sqrt{n}}} \right)^n.$$

Puis, en faisant un développement de Taylor à l'ordre 2 en 0 de $u \rightarrow \mathbf{E} \left(e^{iu \frac{X}{\sqrt{n}}} \right)$, on obtient (voir la remarque 1.7.3 pour un argument précis) :

$$\mathbf{E} \left(e^{iu \frac{X}{\sqrt{n}}} \right) = 1 + \mathbf{E}(X) \frac{u}{\sqrt{n}} - \frac{1}{2} \mathbf{E}(X^2) \frac{u^2}{n} + \frac{1}{n} \epsilon_n.$$

avec $\lim_{n \rightarrow +\infty} \epsilon_n = 0$. Finalement, comme $\mathbf{E}(X) = 0$, $\mathbf{E}(X^2) = \sigma^2$ et

$$\mathbf{E} \left(e^{iu \frac{X}{\sqrt{n}}} \right) = 1 - \frac{1}{2n} \sigma^2 u^2 + \frac{1}{n} \epsilon_n.$$

Donc :

$$\lim_{n \rightarrow +\infty} \Phi_n(u) = \lim_{n \rightarrow +\infty} \left(1 - \frac{1}{2n} \sigma^2 u^2 + \frac{1}{n} \epsilon_n \right)^n = e^{-\frac{\sigma^2 u^2}{2}}.$$

Ceci prouve le résultat annoncé grâce à la proposition 1.7.2.

Remarque 1.7.3 Pour justifier en détail le développement limité, il faut vérifier que la fonction $\Phi(u)$ définie par :

$$\Phi(u) = \mathbf{E} \left(e^{iuX} \right),$$

est 2 fois différentiable, les dérivées secondes partielles étant continues. Pour ceci, on commence par prouver que, comme $\mathbf{E}(|X|) < +\infty$, on a :

$$\Phi'(u) = \mathbf{E}\left(iXe^{iuX}\right).$$

Puis, comme $\mathbf{E}(|X|^2) < +\infty$, on a déduit que :

$$\Phi''(u) = -\mathbf{E}\left(X^2 e^{iuX}\right),$$

et l'on peut aussi prouver que $\Phi''(u)$ est continue en utilisant le théorème de convergence dominée. ■

Notons $\epsilon_n = \frac{1}{n}(X_1 + \dots + X_n) - \mathbf{E}(X)$. Considérons la fonction $f(x) = \mathbf{1}_{\{x \in [a, b]\}}$. Cette fonction n'est pas continue aux points a et b . Mais comme $\mathbf{P}(G = a) = \mathbf{P}(G = b) = 0$, on peut étendre (exercice difficile) le résultat du théorème précédent pour obtenir, que pour tout $a < b$:

$$\lim_{n \rightarrow +\infty} \mathbf{P}\left(\frac{\sigma}{\sqrt{n}}a \leq \epsilon_n \leq \frac{\sigma}{\sqrt{n}}b\right) = \int_a^b e^{-\frac{x^2}{2}} \frac{dx}{\sqrt{2\pi}}.$$

On prend, alors, $b = -a$ et on choisit a de façon à avoir :

$$\int_{-a}^a e^{-\frac{x^2}{2}} \frac{dx}{\sqrt{2\pi}} = p \text{ proche de } 1.$$

Les valeurs classiques de p sont $p = 0.95$ (qui correspond à $a = 1.96$) et $p = 0.99$ ($a = 2.6$).

Dans les applications pratiques on "oublie le passage à la limite" pour affirmer que $\mathbf{P}\left(|\epsilon_n| \leq 1.96 \frac{\sigma}{\sqrt{n}}\right)$ est de l'ordre de 0.95. On a donc avec une probabilité proche de 0.95 que :

$$\mathbf{E}(X) \in \left[\frac{1}{n}(X_1 + \dots + X_n) - 1.96 \frac{\sigma}{\sqrt{n}}, \frac{1}{n}(X_1 + \dots + X_n) + 1.96 \frac{\sigma}{\sqrt{n}} \right]$$

Application à la méthode de Monte-Carlo On cherche à calculer $\mathbf{E}(X)$. On suppose que $\mathbf{E}(X^2) < +\infty$ et que, de plus, l'on sait *simuler* une suite de variables aléatoires $(X_n, n \geq 1)$ (cela signifie que l'on peut considérer les variables aléatoires comme indépendantes et tirées selon la loi de X).

À l'issue de n tirages on va estimer la moyenne m par la moyenne empirique \bar{m} définie par :

$$\bar{m} = \frac{1}{n}(X_1 + \dots + X_n).$$

Quelle est la précision de cette estimation ? Le théorème de la limite centrale nous dit que :

$$\mathbf{P}\left(m \in \left[\bar{m} - 1.96 \frac{\sigma}{\sqrt{n}}, \bar{m} + 1.96 \frac{\sigma}{\sqrt{n}}\right]\right) \approx 0.95.$$

L'écart-type σ n'est évidemment pas explicitement connu lorsque l'on utilise une méthode de Monte-Carlo et l'on doit donc trouver une méthode d'estimation pour σ . La méthode habituellement utilisée est la suivante, on pose :

$$\bar{\sigma}_n^2 = \frac{1}{n-1} \sum_{i=1}^n (X_i - \bar{m})^2 = \frac{1}{n-1} \sum_{i=1}^n X_i^2 - \frac{n}{n-1} \bar{m}^2.$$

On peut alors montrer (exercice facile) que $\mathbf{E}(\bar{\sigma}_n^2)\sigma^2$ (on dit que l'on a affaire à un estimateur sans biais de σ^2) et que, presque sûrement, $\lim_{n \rightarrow +\infty} \bar{\sigma}_n^2 = \sigma^2$ (on dit que l'estimateur est convergent).

Dans la pratique, au cours de la simulation, en plus de stocker la somme des X_i on stocke également la somme $\sum_i X_i^2$ de leurs carrés (cela ne coûte pratiquement rien de plus) pour calculer $\bar{\sigma}_n^2$ à l'aide de sa seconde expression ci-dessus. On remplace σ par $\bar{\sigma}_n$ dans les extrémités de l'intervalle et l'on affirme que (bien que l'on ait rajouté une autre source d'erreur) :

$$\mathbf{P}\left(m \in \left[\bar{m} - 1.96 \frac{\bar{\sigma}_n}{\sqrt{n}}, \bar{m} + 1.96 \frac{\bar{\sigma}_n}{\sqrt{n}}\right]\right) \approx 0.95.$$

Exercice 2 Implementer la méthode de Monte-Carlo qui vous est suggérée avec G une gaussienne centrée réduite et $X = \mathbf{1}_{\{G \geq \lambda\}}$ pour $\lambda = 0, 1, 2, \dots, 5$. Comment évolue l'erreur relative de la méthode ?

Exercice 3 Calculer par simulation (ou au moins essayer !) $\mathbf{E}(e^{\sigma G})$ pour $\sigma = 1, 2, 5$. Préciser à chaque fois l'intervalle de confiance. Que constatez vous ? Comment interpréter ce résultat.

1.8 Autres type de convergence

Définition 1.8.1 On dit qu'une suite de v.a. $(X_n)_{n \in \mathbf{N}^*}$ **converge en probabilité** vers une v.a. X si et seulement si pour tout $\varepsilon > 0$,

$$\lim_{n \rightarrow \infty} \mathbf{P}(|X_n - X| > \varepsilon) = 0.$$

Proposition 1.8.2 *La convergence p.s. entraîne la convergence en probabilité.*

Démonstration : Soit $(X_n, n \in \mathbf{N}^*)$ une suite de v.a qui converge p.s. vers X . La suite de v.a. discrètes positives $(\mathbf{1}_{|X_n - X| > \varepsilon})_{n \in \mathbf{N}^*}$ converge p.s. vers 0. De plus elle est uniformément bornée par 1. Par le théorème de convergence dominée, on en déduit que :

$$\lim_{n \rightarrow \infty} \mathbf{E}[\mathbf{1}_{|X_n - X| > \varepsilon}] = \lim_{n \rightarrow \infty} \mathbf{P}(|X_n - X| > \varepsilon) = \mathbf{E}\left[\lim_{n \rightarrow \infty} \mathbf{1}_{|X_n - X| > \varepsilon}\right] = 0.$$

■

Proposition 1.8.3 *De toute suite qui converge en probabilité, on peut extraire une sous-suite qui converge presque sûrement.*

Démonstration : Soit $(X_n)_{n \in \mathbf{N}^*}$ une suite de v.a. qui converge en probabilité vers X . On définit la sous-suite de manière suivante : $\sigma(1) = 1$ et

$$\sigma(n+1) = \inf \left\{ p > \sigma(n) \text{ tel que } \mathbf{P}\left(|X_p - X| > \frac{1}{n}\right) \leq \frac{1}{n^2} \right\}.$$

La suite (X_n) converge en probabilité, cela assure que la sous-suite $\sigma(n)$ est bien définie. On en déduit, par convergence monotone, que

$$\mathbf{E}\left[\sum_{n \geq 1} \mathbf{1}_{|X_{\sigma(n)} - X| > \frac{1}{n}}\right] \leq \sum_{n \geq 1} \frac{1}{n^2} < \infty.$$

Cela implique que p.s. $\sum_{n \geq 1} \mathbf{1}_{|X_{\sigma(n)} - X| > \frac{1}{n}} < \infty$. Les termes d'une série convergente tendent vers 0. Donc p.s. $\lim_{n \rightarrow \infty} \mathbf{1}_{|X_{\sigma(n)} - X| > \frac{1}{n}} = 0$. Comme la fonction indicatrice ne prend que deux valeurs 0 ou 1, cela entraîne que p.s. $\mathbf{1}_{|X_{\sigma(n)} - X| > \frac{1}{n}}(\omega)$ est nul à partir d'un certain rang n_0 (qui dépend de ω). Donc, à partir d'un certain rang, $|X_{\sigma(n)} - X| < \frac{1}{n}$. En particulier, cela implique que p.s. $\lim_{n \rightarrow \infty} X_{\sigma(n)} = X$. Donc la sous-suite $(X_{\sigma(n)})$ converge p.s. vers X . ■

Proposition 1.8.4 *La convergence en probabilité implique la convergence en loi.*

Remarque 1.8.5 En fait, on a le résultat plus fort dû à Paul Lévy : soit $(X_n, n \in \mathbf{N}^*)$ une suite de v.a. réelle (vectorielle à valeurs dans \mathbf{R}^d). Si ψ_{X_n} converge vers une fonction ψ continue en 0, alors ψ est la fonction caractéristique d'une v.a. vectorielle X et la suite $(X_n, n \in \mathbf{N}^*)$ converge en loi vers X .

Remarque 1.8.6 Il existe un résultat similaire pour les transformées de Laplace.

Proposition 1.8.7 *Soit $(X_n, n \in \mathbf{N}^*)$ une suite de v.a. qui converge en loi vers la loi d'une v.a. X . Soit f une fonction à valeurs réelles, mesurable bornée. Soit A l'ensemble des points de discontinuité de f ($A \in \mathcal{B}(\mathbf{R}^d)$).*

$$\boxed{\text{Si } \mathbf{P}(X \in A) = 0, \text{ alors } \lim_{n \rightarrow \infty} \mathbf{E}[f(X_n)] = \mathbf{E}[f(X)].}$$

Remarque 1.8.8 On peut également montrer que la suite $(X_n, n \in \mathbf{N}^*)$ converge en loi vers X si et seulement si la suite de fonctions de répartition $F_{X_n}(x)$ converge presque partout vers la fonction de répartition de X .

1.9 Problème corrigé

1. Soit X une variable aléatoire suivant une loi de Cauchy réduite, c'est à dire suivant la loi

$$\frac{dx}{\pi(1+x^2)}.$$

Calculer la fonction de répartition de X , notée F .

◁ La fonction de répartition de X est donné, pour tout x réel, par

$$F(x) = \int_{-\infty}^x \frac{du}{\pi(1+u^2)}$$

d'où :

$$F(x) = \frac{1}{\pi} \text{Arc tan}(x) + \frac{1}{2}$$

▷

2. Vérifier que F est une bijection de \mathbf{R} dans $]0, 1[$. On note F^{-1} son inverse et l'on considère une variable aléatoire U de loi uniforme sur $[0, 1]$. Quelle est la probabilité que U vaille 0 ou 1 ? Montrer que $F^{-1}(U)$ suit la même loi que X . En déduire une méthode de simulation selon la loi de Cauchy.

◁ La fonction F est définie sur \mathbf{R} et prend ses valeurs dans $]0, 1[$, de plus elle est strictement croissante puisque sa dérivée est la densité de probabilité de X qui est strictement positive.

F est donc bien une bijection de \mathbf{R} dans $]0, 1[$.

L'ensemble $\{0, 1\}$ est un ensemble de mesure nulle pour la mesure de Lebesgue, et donc :

$$\begin{cases} \mathbf{P}(U = 0) & = & 0 \\ \mathbf{P}(U = 1) & = & 0 \end{cases}$$

De ce qui précède, $F^{-1}(U)$ est bien définie. On va montrer que X et $F^{-1}(U)$ ont même fonction de répartition (elles auront donc même loi). La fonction de répartition de $F^{-1}(U)$ est pour tout x réel :

$$\mathbf{P}(F^{-1}(U) \leq x) = \mathbf{P}(U \leq F(x)),$$

puisque F est une bijection croissante de \mathbf{R} dans $]0, 1[$. On a donc

$$\mathbf{P}(U \leq F(x)) = \int_0^{F(x)} du = F(x).$$

$F^{-1}(U)$ et X ont donc même loi.

Pour simuler la loi de Cauchy il suffit de simuler une variable aléatoire de loi uniforme sur $[0, 1]$ et d'appliquer la fonction $F^{-1}(x) = \tan(\pi(x - \frac{1}{2}))$. Les probabilités de tirer 0 ou 1 étant nulle, il n'y a pas de problème. ▷

3. Soit V une variable aléatoire qui vaut 1 ou -1 avec probabilité $1/2$ et Z une variable aléatoire qui suit une loi exponentielle de paramètre 1. Quelle est la loi de VZ ? Calculer sa fonction caractéristique. En déduire, en utilisant la formule d'inversion de la transformation de Fourier, que :

$$\int_{-\infty}^{+\infty} e^{iux} \frac{dx}{\pi(1+x^2)} = e^{-|u|}.$$

◁ Calculons la loi de VZ. On a pour toute f fonction bornée

$$\begin{aligned} \mathbf{E}(f(\text{VZ})) &= \mathbf{E}(f(Z)\mathbf{1}_{\{V=1\}}) + \mathbf{E}(f(-Z)\mathbf{1}_{\{V=-1\}}) \\ &= \frac{1}{2} \int_0^{+\infty} f(u)e^{-u} du + \frac{1}{2} \int_0^{+\infty} f(-u)e^{-u} du \\ &= \frac{1}{2} \int_{-\infty}^{+\infty} f(u)(\mathbf{1}_{\{u<0\}}e^u + \mathbf{1}_{\{u>0\}}e^{-u}) du \end{aligned}$$

D'où VZ suit la loi

$$\frac{1}{2}e^{-|u|} du.$$

La fonction caractéristique de VZ est alors pour tout u réel :

$$\begin{aligned} \int_{-\infty}^{+\infty} \frac{1}{2} e^{iux-|x|} dx &= \int_0^{+\infty} \frac{1}{2} e^{iux-x} dx + \int_{-\infty}^0 \frac{1}{2} e^{iux+x} dx \\ &= \frac{1}{2} \left(\frac{1+i u}{1+u^2} + \frac{1-i u}{1+u^2} \right) \\ &= \frac{1}{1+u^2} \end{aligned}$$

On en déduit donc en utilisant l'inversion de la transformation de Fourier que $\int_{-\infty}^{+\infty} \frac{e^{iux}}{1+x^2} dx = 2\pi \frac{1}{2} e^{-|u|}$ d'où la fonction caractéristique de la loi de Cauchy réduite :

$$\boxed{\int_{-\infty}^{+\infty} \frac{e^{iux}}{\pi(1+x^2)} dx = e^{-|u|}}$$

▷

4. Soient X et Y deux variables aléatoires indépendantes suivant des lois de Cauchy de paramètres respectifs a et b . Calculer la loi de $X + Y$.

◁ Si les variables aléatoires X et Y sont indépendantes, on peut calculer la fonction caractéristique ϕ_{X+Y} de $X + Y$ pour tout u réel, par :

$$\begin{aligned} \phi_{X+Y}(u) &= \mathbf{E}(e^{iu(X+Y)}) \\ &= \mathbf{E}(e^{iuX} e^{iuY}) \\ &= \mathbf{E}(e^{iuX}) \mathbf{E}(e^{iuY}) \\ &= \phi_X(u) \phi_Y(u) \end{aligned}$$

Or, pour tout réel u la fonction caractéristique d'une variable aléatoire de loi de Cauchy de paramètre λ est $e^{-|\lambda||u|}$ avec la question précédente. D'où, $\phi_{X+Y}(u) = e^{-(|a|+|b|)|u|}$.

$X + Y$ suit une loi de Cauchy de paramètre $|a| + |b|$.

▷

5. Soit $(Y_n, n \geq 1)$, une suite de variables aléatoires réelles convergeant presque sûrement vers une variable aléatoire Z . Montrer, en utilisant le théorème de Lebesgue, que l'on a, pour tout $\epsilon > 0$

$$\lim_{n \rightarrow +\infty} \mathbf{P}(|Y_{2n} - Y_n| \geq \epsilon) = 0.$$

◁ Soit ϵ un réel strictement positif. On considère la variable aléatoire $U_n = \mathbf{1}_{|Y_{2n} - Y_n| \geq \epsilon}$. On peut écrire pour tout entier n ,

$$|Y_{2n} - Y_n| = |Y_{2n} - Z + Z - Y_n| \leq |Y_{2n} - Z| + |Z - Y_n|.$$

Or, on sait que la suite $(Y_n, n \geq 1)$ converge vers Z presque sûrement. On en déduit que la suite $(|Y_{2n} - Y_n|, n \geq 1)$ converge presque sûrement vers 0 et donc qu'il existe un rang N avec N entier naturel dépendant de ϵ tel que pour tout $n > N$ on a $U_n = 0$ presque sûrement.

Ainsi, $(U_n, n \geq 1)$ converge presque sûrement vers 0. De plus, pour tout entier naturel n on a $U_n \leq 1$ avec 1 qui est une variable aléatoire d'espérance finie égale à 1. On peut donc appliquer le théorème de convergence de Lebesgue :

$$\lim_{n \rightarrow +\infty} \mathbf{E}(U_n) = \mathbf{E}(\lim_{n \rightarrow +\infty} U_n).$$

D'où :

$$\lim_{n \rightarrow +\infty} \mathbf{P}(|Y_{2n} - Y_n| \geq \varepsilon) = 0.$$

▷

6. Soit $(X_n, n \geq 1)$ une suite de variables aléatoires indépendantes suivant une loi de Cauchy de paramètre 1. On considère la suite

$$Y_n = \frac{X_1 + X_2 + \dots + X_n}{n}.$$

Calculer la loi de $Y_{2n} - Y_n$. La suite des Y_n converge t'elle en loi ? presque sûrement ?

◁ Pour tout entier $n > 0$,

$$Y_{2n} - Y_n = \frac{-(X_1 + \dots + X_n) + (X_{n+1} + \dots + X_{2n})}{2n}.$$

Comme $-X_1, -X_2, \dots, -X_n, X_{n+1}, \dots, X_{2n}$ sont des variables aléatoires indépendantes suivant des lois de Cauchy, on peut faire comme dans la question 3 et montrer que pour tout entier n :

la loi de $Y_{2n} - Y_n$ qui est une loi de Cauchy de paramètre 1.

De la même manière on peut dire que pour tout n , Y_n suit une loi de Cauchy de paramètre 1.

Y_n converge donc en loi vers une loi de Cauchy de paramètre 1.

Si Y_n convergerait presque sûrement vers une variable aléatoire Z , alors avec la question 5) on aurait

$$\lim_{n \rightarrow +\infty} \mathbf{P}(|Y_{2n} - Y_n| \geq 1) = 0,$$

or $Y_{2n} - Y_n$ suit une loi de Cauchy de paramètre 1 et donc pour tout entier naturel n ,

$$\mathbf{P}(|Y_{2n} - Y_n| \geq 1) = \int_{|x| \geq 1} \frac{dx}{\pi(1+x^2)},$$

est indépendant de n et est strictement positif. Il y a alors une contradiction et donc :

Y_n ne converge pas presque sûrement.

▷

7. Montrer que la suite $(Z_n, n \geq 1)$ définie par

$$Z_n = \left(\sqrt{|X_1|} + \sqrt{|X_2|} + \dots + \sqrt{|X_n|} \right) / n$$

converge presque sûrement et écrire sa limite sous forme d'une intégrale.

◁ On va vérifier que l'on peut utiliser la loi des grands nombres pour la variable aléatoire $\sqrt{|X_i|}$. Pour cela notons que, comme $\sqrt{|x|}$ est une fonction positive on a :

$$\mathbf{E}(\sqrt{|X_i|}) = \int_{\mathbf{R}} \frac{\sqrt{|x|} dx}{\pi(1+x^2)}.$$

Comme, pour x grand :

$$\frac{\sqrt{|x|}}{\pi(1+x^2)} \approx \frac{1}{|x|^{3/2}},$$

$\mathbf{E}(\sqrt{|X_i|}) < +\infty$. La loi des grands nombres permet alors d'affirmer que

$$\underline{(Z_n, n \geq 1) \text{ converge presque sûrement vers } \int_{\mathbf{R}} \frac{\sqrt{|x|} dx}{\pi(1+x^2)}}.$$

▷

8. Vérifier par simulation que Y_n diverge (p.s.) et que Z_n converge (p.s.).

◁ On commence à voir pour 3000 tirages de Y_{40000} et de $Y_{40000} - Y_{20000}$ que Y_n (figure 1.3) et $Y_{2n} - Y_n$ (figure 1.4) suivent une loi de Cauchy de paramètre 1 :

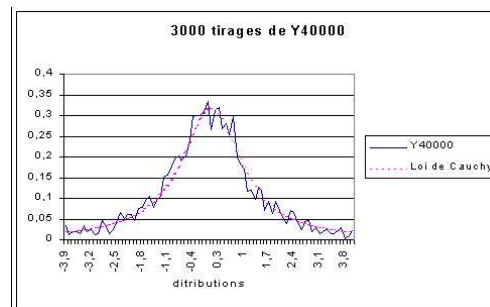


FIG. 1.3 – Loi de Y_n

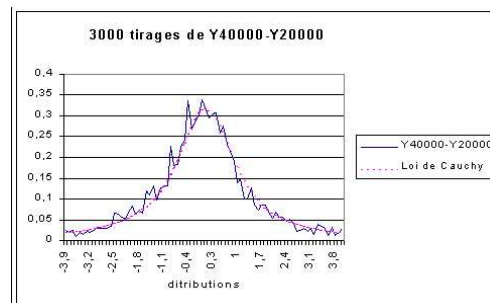
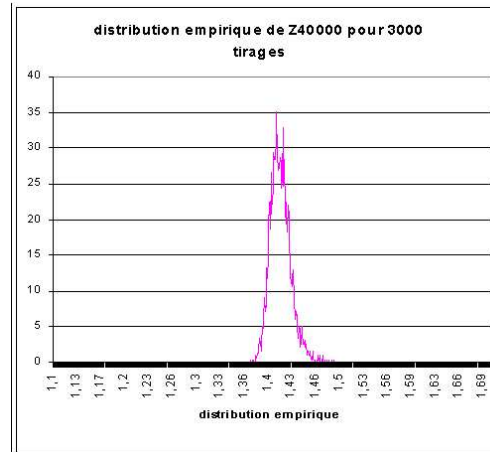


FIG. 1.4 – Loi de $Y_{2n} - Y_n$

Pour 3000 tirages de Z_{40000} , on voit un pic (figure 1.5) se former :

FIG. 1.5 – Loi de Z_n

▽

Exercice 2

Soient f et g deux fonctions de \mathbf{R} dans \mathbf{R}^+ telles que $f(x)$ et $g(x)$ soient les densités de lois de variables aléatoires à valeurs dans \mathbf{R} . On suppose de plus que, pour tout $x \in \mathbf{R}$

$$f(x) \leq kg(x).$$

Soient $(Y_1, Y_2, \dots, Y_n, \dots)$ une suite de variables aléatoires indépendantes suivant la loi de densité $g(x)$ et $(U_1, U_2, \dots, U_n, \dots)$ une suite de variables aléatoires indépendantes suivant une loi uniforme sur $[0, 1]$ indépendante de la suite des Y_i . On pose $N = \inf\{n \geq 1, kU_n g(Y_n) < f(Y_n)\}$.

1. Démontrer que N est fini presque sûrement et suit une loi géométrique. Quelle est sa moyenne ?

◁ Commençons par remarquer que $p = P(kU_n g(Y_n) < f(Y_n))$ ne dépend pas de n (cette probabilité s'exprime en fonction de la loi du couple (Y_n, U_n) qui ne dépend pas de n). De plus en utilisant l'indépendance de Y_n et U_n , on peut aussi écrire p sous la forme

$$p = \int_0^1 \int_{\mathbf{R}} \mathbf{1}_{\{kug(x) < f(x)\}} g(x) du dx.$$

Commençons par montrer que p est non nul. Pour cela, remarquons que si $A = \{x \in \mathbf{R}, g(x) = 0\}$, on a

$$p = \int_0^1 \int_{\mathbf{R}/A} \mathbf{1}_{\{kug(x) < f(x)\}} g(x) du dx.$$

D'où

$$p = \int_{\mathbf{R}/A} \left(\int_0^{\frac{f(x)}{kg(x)}} du \right) g(x) dx = \int_{\mathbf{R}/A} \frac{f(x)}{k} dx.$$

En tenant compte de l'inégalité $f(x) \leq kg(x)$, on obtient pour tout x de A , $f(x) = 0$. D'où $p = \int_{\mathbf{R}} \frac{f(x)}{k} dx$. On en déduit, comme f est une densité, que $p = \frac{1}{k}$. p est en particulier strictement positif.

Pour tout entier non nul n , on a (en utilisant les propriétés d'indépendance)

$$\begin{aligned} \mathbf{P}(N \geq n) &= \mathbf{P}(\forall k, 1 \leq k < n, kU_k g(Y_k) \geq f(Y_k)) \\ &= \prod_{k=1}^n \mathbf{P}(kU_k g(Y_k) \geq f(Y_k)) = (1 - p)^n. \end{aligned}$$

Comme $p > 0$ on obtient (par σ -additivité de la probabilité \mathbf{P} ... ou par convergence monotone, ... ou par convergence dominée) :

$$\mathbf{P}(N = +\infty) = \lim_{n \rightarrow +\infty} \mathbf{P}(N \geq n) = 0.$$

N est donc fini presque sûrement.

Pour tout n entier naturel non nul, on a (par indépendance)

$$\begin{aligned} \mathbf{P}(N = n) &= \mathbf{P}(\forall l, 1 \leq l < n, kU_l g(Y_l) \geq f(Y_l), kU_n g(Y_n) < f(Y_n)) \\ &= \prod_{l=1}^n \mathbf{P}(kU_l g(Y_l) < f(Y_l)) \times \mathbf{P}(kU_n g(Y_n) < f(Y_n)). \end{aligned}$$

D'où

$$\mathbf{P}(N = n) = p(1-p)^{n-1}$$

N suit bien une loi géométrique de paramètre $p = \frac{1}{k}$.

La moyenne de N est donnée par

$$\mathbf{E}(N) = \sum_{n \geq 1} np(1-p)^{n-1},$$

grâce au résultat connu sur les séries :

$$\mathbf{E}(N) = \frac{1}{p} = k$$

▷

2. On définit alors la variable aléatoire X en posant :

$$X = Y_N = \sum_{i \geq 1} Y_i \mathbf{1}_{\{N = i\}}.$$

Calculer pour n fixé et f bornée

$$\mathbf{E} \left(\mathbf{1}_{\{N = n\}} f(X) \right).$$

En déduire la loi de X . Quelle est la loi du couple (N, X) ?

◁ Soit n un entier naturel et h une fonction réelle bornée. On a

$$\mathbf{E}(\mathbf{1}_{\{N=n\}} h(X)) = \mathbf{E}(\mathbf{1}_{\{\forall l \leq l < n, kU_l g(Y_l) \geq f(Y_l)\}} \times \mathbf{1}_{\{kU_n g(Y_n) < f(Y_n)\}} h(Y_n)).$$

D'où par indépendance entre les tirages, on a

$$\mathbf{E}(\mathbf{1}_{\{N=n\}} h(X)) = (1-p)^{n-1} \mathbf{E}(\mathbf{1}_{\{kU_n g(Y_n) < f(Y_n)\}} h(Y_n)).$$

Et par le même raisonnement que dans la question 1) on obtient :

$$\mathbf{E}(\mathbf{1}_{\{kU_n g(Y_n) < f(Y_n)\}} h(Y_n)) = p \int_{\mathbf{R}} h(y) f(y) dy.$$

Donc :

$$\mathbf{E}(\mathbf{1}_{\{N=n\}} h(X)) = (1-p)^{n-1} p \int_{\mathbf{R}} h(y) f(y) dy.$$

Cette relation permet d'identifier la loi du couple X, N et prouve en particulier que X et N sont des variables aléatoires indépendantes.

Pour identifier la loi de X , il suffit de remarquer que toute fonction réelle bornée h ,

$$\mathbf{E}(h(X)) = \sum_{n \geq 1} \mathbf{E}(\mathbf{1}_{N=n} h(X)) = \int_{\mathbf{R}} h(y) f(y) dy \times \left(\sum_{n \geq 1} p(1-p)^{n-1} \right),$$

d'où $\mathbf{E}(h(X)) = \int_{\mathbf{R}} h(y) f(y) dy$.

X suit donc la loi de densité f .

▷

3. En déduire comment on peut simuler une variable aléatoire de loi $f(x) dx$ si on sait simuler une variable aléatoire de loi $g(x) dx$.

◁ On a vu que X suit une loi de densité f . Il suffit donc de simuler les suites $(U_n, n \geq 1)$ et $(Y_n, n \geq 1)$ de manière indépendante jusqu'à N (qui est une variable aléatoire) : Y_N donne alors un tirage d'une variable aléatoire de loi $f(x)dx$.

On peut remarquer qu'il faut trouver une densité g telle qu'il existe un k permettant d'assurer que $f(x) \leq kg(x)$ pour tout x . C'est souvent possible, quitte à prendre k grand. Cependant plus k est grand, plus l'espérance de N , $E(N) = k$ est grande et plus la simulation sera inefficace. Il est donc conseillé de choisir g de façon à avoir k aussi proche de 1 que possible. ▷

1.10 Exercices

Exercice 4 Soit $(X_i)_{i \in \mathbb{N}}$ une suite de variables aléatoires indépendantes, les X_i suivant une loi de moyenne m et de variance σ^2 .

1. Calculez l'estimation de $\mathbf{P}\left(\left|\frac{X_1 + \dots + X_n}{n} - m\right| \geq \alpha\right)$ en fonction de la variance que l'on obtient en appliquant l'inégalité de Bienaymé-Tchebychev. En déduire un intervalle de confiance à 5% près, par excès, de la moyenne à l'issue de n tirages.
2. Appliquer le théorème de la limite centrale et en déduire un intervalle de confiance pour l'estimation de la moyenne à 5% près.
3. On suppose que $\phi(\lambda) = \mathbf{E}(e^{\lambda X}) < +\infty$ pour un $\lambda \in \mathbf{R}$. Calculer $\phi(\lambda)$ si X suit une loi gaussienne centrée de variance σ^2 (supposée connue).
4. On suppose que la loi de X est symétrique, ce qui signifie que :

$$\mathbf{E}(f(X)) = \mathbf{E}(f(-X)),$$

pour toute fonction borélienne positive (ou bornée) f . Démontrer que dans ce cas pour tout $\beta > 0$:

$$\mathbf{P}(|X| \geq \lambda) \leq \frac{2}{e^{\beta\lambda}} \mathbf{E}(e^{\beta X})$$

5. Démontrer que pour tout $\beta > 0$:

$$\mathbf{P}\left(\left|\frac{X_1 + \dots + X_n}{n} - m\right| \geq \alpha\right) \leq \frac{2}{e^{\beta\alpha}} \mathbf{E}\left(e^{\frac{\beta}{n}(X_1 - m)}\right)^n$$

6. En déduire que $\mathbf{P}\left(\left|\frac{X_1 + \dots + X_n}{n} - m\right| \geq \alpha\right) \leq 2 e^{-n\psi(\alpha)}$ où :

$$\psi(u) = \sup_{\beta \in \mathbf{R}^+} \{\beta u - \log(\mathbf{E}(e^{\beta(X_1 - m)}))\}.$$

7. Calculer $\psi(u)$ pour une gaussienne de moyenne m et de variance σ^2 . Déduire pour cette loi un intervalle de confiance à 5% de la moyenne. Comparer avec les résultats obtenus en 1 et en 2.

Exercice 5 Soit X et Y deux variables aléatoires gaussiennes centrées de variance 1 indépendantes. Trouver la loi du couple $\left(\frac{X-Y}{\sqrt{2}}, \frac{X+Y}{\sqrt{2}}\right)$. En déduire la fonction caractéristique de $\frac{X^2 - Y^2}{2}$.

Exercice 6 Soit $(U_n, n \geq 1)$ une suite de variables aléatoires réelles indépendantes identiquement distribuées. On pose :

$$Y_n = \sup_{1 \leq k \leq n} U_k \text{ et } Z_n = \inf_{1 \leq k \leq n} U_k.$$

1. On suppose la loi des U_i uniforme sur $[0, 1]$. Calculer la loi de Y_n et de Z_n . Quel est le comportement presque sûr de Y_n et Z_n ?
2. Déterminer la loi du couple (Y_n, Z_n) . On commencera par calculer

$$\mathbf{P}(Y_n \leq x, Z_n \geq y).$$

3. Quelle est la limite en loi du couple $((n-2)(1-Y_n), (n-2)Z_n)$?
4. Si les U_i suivent des lois exponentielles de paramètre 1, quelle est la loi de Y_n et celle Z_n . Quel est le comportement presque sûr de ces variables lorsque n tend vers $+\infty$.
5. On suppose que les U_i suivent des lois uniformes sur $[0, \theta]$, θ étant un paramètre que l'on cherche à estimer. On va estimer θ par $\theta_n = Y_n$. Calculer la moyenne de θ_n et en déduire un estimateur sans biais de θ . Calculer la variance de l'estimateur.

Exercice 7 Soit $(X_n, n \geq 0)$ une suite de variables aléatoires indépendantes qui ont toutes pour loi une loi exponentielle de paramètre λ .

1. On pose $S_n = X_0 + \dots + X_n$. Calculer la loi de S_n .
2. Soit $T = \inf\{n \geq 0, S_n > 1\}$. Démontrer que presque sûrement $T < +\infty$.
3. Calculer la loi de T et l'identifier comme une loi de Poisson.
4. Proposer une méthode de simulation de la loi de Poisson, en prenant soin d'utiliser au maximum des fonctions élémentaires (addition, multiplication).

Exercice 8 Soit $X = (X_1, \dots, X_n)$ un vecteur centré, de matrice de covariance K . Calculer la variance de $\sum_{i=1}^n u_i X_i$. Démontrer que si il existe un vecteur $u \neq 0$ tel que $Ku = 0$, alors la loi de X est portée par un sous espace vectoriel strict de \mathbf{R}^n .

Exercice 9 Soit X une variable aléatoire normale centrée réduite.

1. Calculer la loi de X^3 .
2. Soit Z la variable aléatoire définie par, si $\alpha > 0$:

$$Z = \begin{cases} -X & \text{si } |X| < \alpha \\ X & \text{si } |X| \geq \alpha \end{cases}$$

Calculer la loi de Z . Le couple (X, Z) est-il gaussien ?

3. Donner un exemple de couple de variables aléatoires (X, Z) centrées, telles que $\mathbf{E}(XZ) = 0$, X et Z suivent des lois gaussiennes et telles que X et Z ne sont pas indépendantes.

Exercice 10 Soit $((X_i, Y_i), i \geq 1)$ une suite de variables aléatoires indépendantes à valeurs dans \mathbf{R}^2 suivant toutes la même loi. On suppose de plus que $\mathbf{E}(X_1) = \mathbf{E}(Y_1) = 0$, que $\mathbf{E}(X_1^2) < +\infty$ et $\mathbf{E}(Y_1^2) < +\infty$.

1. Quelle est la loi limite de $(X_1 + \dots + X_n)/\sqrt{n}$, de $(Y_1 + \dots + Y_n)/\sqrt{n}$?
2. Quelle est la loi limite du couple $((X_1 + \dots + X_n)/\sqrt{n}, (Y_1 + \dots + Y_n)/\sqrt{n})$?
3. A quelle condition la loi limite est elle celle d'un couple de variables aléatoires indépendantes ?
4. On suppose que le couple des (X_i, Y_i) suit la loi suivante : avec probabilité $1/4$ le couple vaut soit $(\cos(\theta), \sin(\theta))$, soit $(-\sin(\theta), \cos(\theta))$, soit $(-\cos(\theta), -\sin(\theta))$, soit $(\sin(\theta), -\cos(\theta))$. A quelle condition sur θ , les X_i et les Y_i sont elles indépendantes ? Montrer que, cependant, quelle que soit la valeur de θ , la loi limite est celle d'un couple de variables aléatoires indépendantes dont on donnera la loi.

Exercice 11 On se donne $2n$ ($n \geq 2$) variables aléatoires réelles gaussiennes indépendantes $X_1, \dots, X_n, Y_1, \dots, Y_n$ telles que, pour $1 \leq i \leq n$:

$$\mathbf{E}(X_i) = \lambda, \mathbf{E}(Y_i) = \mu,$$

et :

$$\text{Var}(X_i) = \text{Var}(Y_i) = \sigma^2,$$

λ, μ, σ étant trois paramètres réels. On pose :

$$L_n = \frac{X_1 + \dots + X_n}{n} \quad M_n = \frac{Y_1 + \dots + Y_n}{n}.$$

1. Trouver la loi du vecteur aléatoire (L_n, M_n) .
2. Exprimer lorsque $\lambda = \mu$, la fonction définie, pour $\alpha \geq 0$ par :

$$f(\alpha) = \mathbf{P}(|L_n - M_n| > \alpha\sigma/\sqrt{n}),$$

à l'aide de $N(x) = \int_{-\infty}^x e^{-u^2/2} du / \sqrt{2\pi}$.

3. On pose :

$$S_n^2 = \frac{1}{n-1} \sum_{i=1}^n (X_i - L_n)^2 \quad T_n^2 = \frac{1}{n-1} \sum_{i=1}^n (Y_i - M_n)^2$$

Montrer que S_n^2, T_n^2 et $(S_n^2 + T_n^2)/2$ sont des estimateurs sans biais de σ^2 . Lequel vous paraît-il préférable ?

4. En déduire que lorsque $\lambda = \mu$, la variable aléatoire :

$$\sqrt{n} \frac{L_n - M_n}{\sqrt{S_n^2 + T_n^2}},$$

suit une loi de Student de paramètre $n - 1$.

5. Que se passe-t-il lorsque $\lambda \neq \mu$ et que n tend vers $+\infty$. Proposer un test de l'hypothèse $\lambda = \mu$ contre l'alternative $\lambda \neq \mu$.
6. Soit deux groupes de 21 élèves. On constate que la moyenne empirique est de 13.8 et la variance empirique de 2.4 pour le premier groupe et de 13 et 3.1 pour le deuxième groupe. L'écart entre les deux groupes vous paraît-il significatif ? (Student = 0.93)

Exercice 12 On considère le modèle de transmission suivant : un signal déterministe s_n est bruité par une variable aléatoire B_n . On pose $X_n = \alpha s_n + \sigma B_n$ où X_n est l'observation effective (bruitée). α et σ sont des paramètres fixés mais inconnus. On cherche en particulier une méthode d'estimation de α . On suppose que les $(B_k)_{1 \leq k \leq n}$ suivent des lois gaussiennes centrées réduites et sont indépendantes.

1. Trouver la loi du vecteur (X_1, \dots, X_n) .
2. On note $x = (x_1, \dots, x_n)$ un point de \mathbf{R}^n . Trouver la valeur $\hat{\alpha}(x)$ qui maximise cette densité. Calculer :

$$\mathbf{E}(\hat{\alpha}(X_1, \dots, X_n)) \text{ et } \text{Var}(\hat{\alpha}(X_1, \dots, X_n)).$$

A quelle condition $\hat{\alpha}$ est-il un estimateur convergent en moyenne quadratique.

Chapitre 2

Vecteurs gaussiens - Mouvement Brownien

2.1 Vecteurs gaussiens

Définition 2.1.1 On dit qu'un vecteur aléatoire $X = (X_1, \dots, X_d)$ est un vecteur gaussien si, pour tout $u \in \mathbf{R}^d$, la variable aléatoire réelle $u \cdot X$ est une variable aléatoire gaussienne.

Notons que par linéarité de l'espérance et par bilinéarité de la variance on a :

$$\mathbf{E}(u \cdot X) = u_1 \mathbf{E}(X_1) + \dots + u_d \mathbf{E}(X_d) = u \cdot \mathbf{E}(X),$$

et :

$$\text{Var}(u \cdot X) = \sum_{i=1}^d \sum_{j=1}^d u_i u_j \text{Cov}(X_i, X_j),$$

avec $\text{Cov}(Y, Z) = \mathbf{E}(YZ) - \mathbf{E}(Y)\mathbf{E}(Z) = \mathbf{E}\{(Y - \mathbf{E}(Y))(Z - \mathbf{E}(Z))\}$.

Notons Γ , avec $\Gamma_{ij} = \text{Cov}(X_i, X_j)$, la matrice de variance covariance du vecteur X et m , avec $m_i = \mathbf{E}(X_i)$, son vecteur moyenne. On voit donc que :

$$\mathbf{E}(e^{iu \cdot X}) = e^{i\mathbf{E}(u \cdot X) - \frac{1}{2} \text{Var}(u \cdot X)} = e^{iu \cdot m - \frac{1}{2} u \cdot \Gamma u}.$$

On constate que la fonction caractéristique (et donc la loi) ne dépend que du couple (m, Γ) . On dit dans ce cas que X suit la loi $\mathcal{N}(m, \Gamma)$.

Exemples : Si G_1, \dots, G_d sont des variables gaussiennes centrées réduites indépendantes alors $G = (G_1, \dots, G_d)$ est un vecteur gaussien centré de matrice de variance covariance égale à la matrice identité de \mathbf{R}^d . En effet pour $u \in \mathbf{R}^d$, par indépendance, la fonction caractéristique de $u \cdot G$ est

$$\Phi_{u \cdot G}(v) = \mathbf{E}(e^{iv(u \cdot G)}) = \prod_{j=1}^d \mathbf{E}(e^{iv u_j G_j}) = \prod_{j=1}^d e^{-\frac{(v u_j)^2}{2}} = e^{-\frac{v^2}{2} u \cdot u},$$

ce qui implique que cette variable aléatoire suit la loi gaussienne centrée de variance $u \cdot u$.

Remarque 2.1.2 Il ne suffit pas que X_1 et X_2 soient des variables gaussiennes réelles pour que (X_1, X_2) soit un vecteur gaussien. Ce résultat est vrai dans le cas particulier où X_1 et X_2 sont indépendantes mais pas en toute généralité.

Notons que la matrice de variance covariance Γ de n'importe quel vecteur X est une matrice symétrique positive (i.e. $u \cdot \Gamma u \geq 0$ pour tout vecteur u de \mathbf{R}^d) : en effet, $u \cdot \Gamma u = \text{Var}(u \cdot X) \geq 0$. D'autre part, on peut construire (ou simuler) un vecteur gaussien de vecteur moyenne m et de

matrice de covariance Γ données *si et seulement si* la matrice Γ est symétrique positive. Nous prouvons ce résultat dans le paragraphe consacré à la simulation d'un vecteur gaussien. Notons, de plus, que Γ peut être dégénérée (i.e. il existe $u_0 \neq 0$ tel que $\Gamma u_0 = 0$) et dans ce cas X prend ses valeurs dans un sous espace affine strict de \mathbf{R}^d .

Stabilité du caractère gaussien par transformation linéaire Si X est un vecteur aléatoire à valeurs \mathbf{R}^d et $Y = \alpha + MX$ pour un vecteur (constant) $\alpha \in \mathbf{R}^n$ et une matrice M de taille $n \times d$, alors toute combinaison linéaire des coordonnées de Y est combinaison linéaire des coordonnées de X à une constante près : pour $v \in \mathbf{R}^n$, $v \cdot Y = v \cdot \alpha + (M^t v) \cdot X$. Donc si X est gaussien, Y l'est aussi.

Exercice 13 Vérifier que $\mathbf{E}(Y) = \alpha + M \cdot \mathbf{E}(X)$ et que la matrice de variance covariance Λ de Y s'exprime en fonction de celle Γ de X par la relation $\Lambda = M \Gamma M^t$.

Exemples : Si $X = (X_1, \dots, X_d)$ est un vecteur gaussien à valeurs \mathbf{R}^d alors pour $k \leq d$, le vecteur (X_1, \dots, X_k) est gaussien (de même que tout vecteur obtenu à partir de X en enlevant certaines coordonnées). La moyenne empirique $\frac{1}{d} \sum_{i=1}^d X_i$ est une gaussienne réelle.

Simulation d'un vecteur gaussien Supposons que l'on cherche simuler un vecteur gaussien dont on connaît le vecteur moyenne m et la matrice de variance-covariance Γ .

La matrice Γ est symétrique positive. Il existe donc une racine carrée A telle que :

$$AA^t = \Gamma.$$

Voir l'exercice 14 pour une preuve.

D'après ce qui précède, si $G = (G_1, \dots, G_d)$ est un vecteur de variables aléatoires gaussiennes centrées réduites indépendantes alors $m + AG$ suit la loi $\mathcal{N}(m, \Gamma)$.

On débouche donc sur l'algorithme suivant pour simuler un vecteur gaussien m, Γ :

Simulation d'un vecteur gaussien :

- Calculer une racine carrée A de Γ ,
- tirer un vecteur de gaussiennes centrées réduites indépendantes $G = (G_1, \dots, G_d)$,
- retourner le vecteur $m + AG$.

Exercice 14 Algorithme de Cholevsky En supposant que est A est une matrice triangulaire supérieure, résoudre itérativement l'équation $AA^t = \Gamma$. Programmer la méthode ainsi identifiée.

Vecteurs gaussiens et indépendance Si les coordonnées du vecteur aléatoire $X = (X_1, \dots, X_d)$ sont indépendantes et de carré intégrable, alors sa matrice de variance covariance est diagonale. En effet pour $i \neq j$, $\mathbf{E}(X_i X_j) = \mathbf{E}(X_i) \mathbf{E}(X_j)$ i.e. $\text{Cov}(X_i, X_j) = 0$.

Dans le cas où X est un vecteur gaussien, le caractère diagonal de la matrice de covariance s'avère une condition suffisante d'indépendance :

Proposition 2.1.3 Les coordonnées d'un vecteur gaussien $X = (X_1, \dots, X_d)$ sont indépendantes si et seulement si sa matrice de variance covariance Γ est diagonale.

Démonstration : Nous avons déjà démontré la condition nécessaire. Supposons que la matrice de variance covariance Γ est diagonale et notons m le vecteur espérance de X . Si Y_1, \dots, Y_d sont des variables gaussiennes indépendantes d'espérance et variance respectives $m_1, \Gamma_{11}, \dots, m_d, \Gamma_{dd}$ alors le vecteur

$Y = (Y_1, \dots, Y_d)$ suit la loi $\mathcal{N}(m, \Gamma)$. Donc X et Y ont même loi et les coordonnées X_i sont des variables indépendantes. ■

Exercice 15 Soit Y_1, \dots, Y_n des vecteurs aléatoires à valeurs respectivement dans $\mathbf{R}^{d_1}, \dots, \mathbf{R}^{d_n}$ t.q. le vecteur $Y = (Y_1, \dots, Y_n)$ est gaussien (cela implique en particulier que les Y_i sont des vecteurs gaussiens). Montrer que les vecteurs Y_1, \dots, Y_n sont indépendants si et seulement si la matrice de variance covariance Γ de Y est diagonale par blocs au sens où :

$$\forall (j, k) \notin \bigcup_{i=1}^n [d_1 + \dots + d_{i-1} + 1, d_1 + \dots + d_i]^2, \Gamma_{jk} = 0.$$

Convergence en loi de vecteurs aléatoires La définition de la convergence en loi est identique au cas unidimensionnel.

Définition 2.1.4 On dit que $(X_n, n \geq 1)$ suite de variables aléatoires à valeurs dans \mathbf{R}^d converge en loi vers X variable aléatoire à valeurs dans \mathbf{R}^d si, pour toute fonction continue et bornée $f : \mathbf{R}^d \rightarrow \mathbf{R}$,

$$\lim_{n \rightarrow +\infty} \mathbf{E}(f(X_n)) = \mathbf{E}(f(X)).$$

On peut étendre au cas vectoriel le critère de convergence utilisant la fonction caractéristique.

Théorème 2.1.1 Soit $(X_n, n \geq 1)$ une suite de variables aléatoires à valeurs dans \mathbf{R}^d . Cette suite converge en loi vers X si et seulement si, pour tout $u = (u_1, \dots, u_d)$, on a :

$$\lim_{n \rightarrow +\infty} \mathbf{E}(e^{iu \cdot X_n}) = \mathbf{E}(e^{iu \cdot X}).$$

Ce résultat est délicat à obtenir et nous l'admettrons. Nous allons voir qu'il permet de démontrer simplement une généralisation du théorème de la limite centrale.

Théorème 2.1.2 (de la limite centrale multidimensionnel) Soit $(X_n, n \geq 1)$ une suite de variables aléatoires à valeurs dans \mathbf{R}^d indépendantes suivant toutes la même loi que X . On suppose que $\mathbf{E}(|X|^2) < +\infty$ (où $|x|$ désigne la norme euclidienne d'un vecteur de \mathbf{R}^d). On note m le vecteur des espérances de X , $m_i = \mathbf{E}(X_i)$ pour $i = 1, \dots, d$. et Γ la matrice de variance-covariance du vecteur X , $\Gamma_{ij} = \text{Cov}(X_i, X_j)$. Alors :

$$\sqrt{n} \left(\frac{1}{n} (X_1 + \dots + X_n) - m \right)$$

converge en loi vers un vecteur gaussien centré de matrice de variance-covariance Γ .

Démonstration : Soit $u \in \mathbf{R}^d$ et $Y_n = u \cdot X_n$. Les variables Y_n sont des variables aléatoires réelles indépendantes et identiquement distribuées. Elles sont de carré intégrable. Leur espérance est $m_u = u \cdot m$ et leur variance $\sigma_u^2 = u \cdot \Gamma u$.

Donc, par application du théorème de la limite centrale dans \mathbf{R} (théorème 1.7.1), la suite $\sqrt{n} \left(\frac{1}{n} (Y_1 + \dots + Y_n) - m_u \right)$ converge en loi vers $\sigma_u G$. En utilisant la caractérisation de la convergence en loi des variables aléatoires réelles par les fonctions caractéristiques, on en déduit :

$$\lim_{n \rightarrow +\infty} \mathbf{E} \left(e^{i\sqrt{n} \left(\frac{1}{n} (Y_1 + \dots + Y_n) - m_u \right)} \right) = \mathbf{E} \left(e^{i\sigma_u G} \right) = e^{-\frac{1}{2} u \cdot \Gamma u}.$$

En utilisant la définition de Y_n , on obtient

$$\lim_{n \rightarrow +\infty} \mathbf{E} \left(e^{iu \cdot \sqrt{n} \left(\frac{1}{n} (X_1 + \dots + X_n) - m \right)} \right) = e^{-\frac{1}{2} u \cdot \Gamma u}.$$

Comme $u \in \mathbf{R}^d$ est arbitraire, on conclut en utilisant la caractérisation de la convergence en loi des vecteurs aléatoires dans \mathbf{R}^d par les fonctions caractéristiques. ■

Stabilité des vecteurs gaussiens par convergence en loi Le résultat suivant qui assure que toute limite en loi (à fortiori presque sûre, en probabilité, dans L^1 ou dans L^2 puisque ces convergences impliquent la convergence en loi) de variables gaussiennes réelles est gaussienne, est très utile pour obtenir le caractère gaussien de variables obtenues par un procédé de passage à la limite.

Proposition 2.1.5 Soit $(X_n)_n$ une suite de variables gaussiennes réelles qui converge en loi vers X . Alors X est gaussienne.

Démonstration : Supposons dans un premier temps que X_n converge vers X dans L^2 i.e. que $\mathbf{E}((X_n - X)^2) \rightarrow 0$. Comme $(\mathbf{E}(X_n) - \mathbf{E}(X))^2 \leq \mathbf{E}((X_n - X)^2)$, $\mathbf{E}(X_n) \rightarrow \mathbf{E}(X)$.

Par l'inégalité de Cauchy-Schwarz, $\mathbf{E}(X_n X) \leq \sqrt{\mathbf{E}(X_n^2)} \sqrt{\mathbf{E}(X^2)}$.

Donc

$$\mathbf{E}((X_n - X)^2) = \mathbf{E}(X_n^2) - 2\mathbf{E}(X_n X) + \mathbf{E}(X^2) \geq (\sqrt{\mathbf{E}(X_n^2)} - \sqrt{\mathbf{E}(X^2)})^2.$$

Ainsi $\mathbf{E}(X_n^2) \rightarrow \mathbf{E}(X^2)$ et $\text{Var}(X_n) \rightarrow \text{Var}(X)$.

On en déduit que pour $u \in \mathbf{R}$, $\mathbf{E}(e^{iuX_n}) = e^{iu\mathbf{E}(X_n) - \frac{u^2}{2}\text{Var}(X_n)} \rightarrow e^{iu\mathbf{E}(X) - \frac{u^2}{2}\text{Var}(X)}$. La convergence dans L^2 de X_n vers X entraîne la convergence en loi de X_n vers X et donc d'après les théorèmes 1.7.2 et 2.1.1, la convergence de $\mathbf{E}(e^{iuX_n})$ vers $\mathbf{E}(e^{iuX})$. On conclut donc que

$$\forall u \in \mathbf{R}, \mathbf{E}(e^{iuX}) = e^{iu\mathbf{E}(X) - \frac{u^2}{2}\text{Var}(X)},$$

ce qui implique que X est gaussienne.

Traisons maintenant le cas plus délicat où l'on suppose seulement la convergence en loi de la suite X_n vers X . Comme la fonction caractéristique $\Phi_X(u) = \mathbf{E}(e^{iuX})$ est continue (conséquence du théorème de convergence dominée) et vérifie $\Phi_X(0) = 1$, il existe $v > 0$ t.q. $|\Phi_X(v)| \neq 0$. La convergence en loi de X_n vers X entraîne celle de $\mathbf{E}(e^{ivX_n})\mathbf{E}(e^{-ivX_n}) = e^{-v^2\text{Var}(X_n)}$ vers $\mathbf{E}(e^{ivX})\mathbf{E}(e^{-ivX}) = |\Phi_X(v)|^2$. On en déduit que la suite $(\text{Var}(X_n))_n$ converge vers une limite $\sigma^2 \in \mathbf{R}_+$.

Comme pour tout $u \in \mathbf{R}$, $\mathbf{E}(e^{iuX_n}) = e^{iu\mathbf{E}(X_n) - \frac{u^2}{2}\text{Var}(X_n)}$ converge vers $\mathbf{E}(e^{iuX})$, pour tout $u \in \mathbf{R}$ la suite $e^{iu\mathbf{E}(X_n)}$ converge vers une limite $\psi(u)$ avec $|\psi(u)| = 1$. Montrons par l'absurde que l'on peut extraire de la suite $(\mathbf{E}(X_n))_n$ une sous-suite $(\mathbf{E}(X_{n_k}))_k$ qui converge vers une limite finie $m \in \mathbf{R}$. Supposons qu'au contraire $|\mathbf{E}(X_n)| \rightarrow +\infty$. Soit φ une fonction C^∞ à support compact sur \mathbf{R} . Par convergence dominée

$$\int_{\mathbf{R}} \psi(u)\varphi(u)du = \lim_{n \rightarrow +\infty} \int_{\mathbf{R}} e^{iu\mathbf{E}(X_n)}\varphi(u)du.$$

Or par intégration par parties,

$$\int_{\mathbf{R}} e^{iu\mathbf{E}(X_n)}\varphi(u)du = \frac{i}{\mathbf{E}(X_n)} \int_{\mathbf{R}} e^{iu\mathbf{E}(X_n)}\varphi'(u)du.$$

ce qui entraîne que $|\int_{\mathbf{R}} e^{iu\mathbf{E}(X_n)}\varphi(u)du| \leq \frac{1}{|\mathbf{E}(X_n)|} \int_{\mathbf{R}} |\varphi'(u)|du$. On conclut que pour toute fonction C^∞ à support compact φ ,

$$\int_{\mathbf{R}} \psi(u)\varphi(u)du = 0.$$

Ainsi la distribution associée à la fonction localement intégrable ψ est nulle i.e. $\psi \equiv 0$, ce qui contredit $|\psi| \equiv 1$. Donc il existe une sous-suite $(\mathbf{E}(X_{n_k}))_k$ qui converge vers une limite $m \in \mathbf{R}$. On conclut que

$$\forall u \in \mathbf{R}, \mathbf{E}(e^{iuX}) = \lim_k \mathbf{E}(e^{iuX_{n_k}}) = \lim_k e^{iu\mathbf{E}(X_{n_k}) - \frac{u^2}{2} \text{Var}(X_{n_k})} = e^{i um - \frac{u^2}{2} \sigma^2}.$$

■

Corollaire 2.1.6 Soit $(X_n)_n$ une suite de vecteurs gaussiens à valeurs \mathbf{R}^d qui converge en loi vers un vecteur X . Alors X est gaussien.

Démonstration : Comme pour $u \in \mathbf{R}^d$, l'application $x \in \mathbf{R}^d \rightarrow u \cdot x \in \mathbf{R}$ est continue, la convergence en loi de X_n vers X entraîne celle de $u \cdot X_n$ vers $u \cdot X$. Les variables $u \cdot X_n$ étant gaussiennes, on en déduit avec la proposition 2.1.5 que $u \cdot X$ est une variable gaussienne. Comme $u \in \mathbf{R}^d$ est arbitraire, X est un vecteur gaussien. ■

2.2 Mouvement Brownien

Définition 2.2.1 On appelle processus stochastique à temps continu à valeurs dans un espace E muni d'une tribu \mathcal{E} , une famille $(X_t, t \geq 0)$ de variables aléatoires à valeurs dans E définies sur un espace de probabilité $(\Omega, \mathcal{A}, \mathbf{P})$.

Remarque 2.2.2 L'indice $t \in [0, +\infty[$ représente le temps. Notons que l'on peut associer, à chaque $\omega \in \Omega$, une trajectoire :

$$t \rightarrow X_t(\omega).$$

Définition 2.2.3 Un processus stochastique $(B_t, t \geq 0)$ à valeurs réelles est appelé mouvement brownien (standard) s'il vérifie les quatre propriétés suivantes :

- (i) $B_0 = 0$
- (ii) Pour tout $s \leq t$, l'accroissement $B_t - B_s$ suit la loi gaussienne centrée de variance $t - s$.
- (iii) si $0 \leq t_1 \leq t_2 \leq \dots \leq t_n$, les accroissements $B_{t_1}, B_{t_2} - B_{t_1}, \dots, B_{t_n} - B_{t_{n-1}}$ sont indépendants
- (iv) En dehors d'un ensemble de probabilité nulle, les trajectoires $t \rightarrow B_t(\omega)$ sont continues.

Notons que (i) et (ii) implique que $B_t = B_t - B_0$ suit la loi gaussienne centrée de variance t dont la densité est

$$\frac{1}{\sqrt{2\pi t}} e^{-\frac{x^2}{2t}}.$$

Définition 2.2.4 Un processus $(X_t, t \geq 0)$ à valeurs réelles est dit :

- continu si les trajectoires $t \rightarrow X_t(\omega)$ le sont.
- à accroissements indépendants si $\forall n \in \mathbf{N}^*$ et $0 \leq t_1 \leq t_2 \leq \dots \leq t_n$, les accroissements $X_{t_1} - X_0, X_{t_2} - X_{t_1}, \dots, X_{t_n} - X_{t_{n-1}}$ sont indépendants.
- à accroissements stationnaires si $\forall s, t \geq 0, X_{t+s} - X_t$ a même loi que $X_s - X_0$.

Un mouvement brownien standard est donc un processus continu (iv) à accroissements indépendants (iii) et stationnaires (ii). En fait, on peut caractériser tout les processus qui vérifient ces trois propriétés

Théorème 2.2.1 Si $(X_t, t \geq 0)$ est un processus continu à accroissements indépendants et stationnaires alors il existe deux constantes réelles r et σ t.q. $\forall t \geq 0, X_t - X_0 = rt + \sigma B_t$ avec $(B_t, t \geq 0)$ un mouvement brownien.

Démonstration : La démonstration du caractère gaussien de $X_t - X_0$ est délicate. Elle repose sur des variantes du théorème de la limite centrale. Nous admettrons ce résultat.

Par contre, il est très facile de calculer la moyenne et la variance de $X_t - X_0$. En effet :

$$X_{t+s} - X_0 = (X_{t+s} - X_t) + (X_t - X_0).$$

La loi de $(X_{t+s} - X_t)$ est identique à celle de $(X_s - X_0)$ par hypothèse. Comme nous avons admis qu'il s'agissait d'une variable aléatoire gaussienne, cette variable aléatoire a une espérance et une variance finie ainsi que $(X_t - X_0)$. Donc :

$$\mathbf{E}(X_{t+s} - X_0) = \mathbf{E}(X_{t+s} - X_t) + \mathbf{E}(X_t - X_0) = \mathbf{E}(X_s - X_0) + \mathbf{E}(X_t - X_0).$$

Si l'on note $\phi(t) = \mathbf{E}(X_t - X_0)$, on a donc l'équation fonctionnelle :

$$\phi(t+s) = \phi(t) + \phi(s).$$

On en déduit que pour $k, n \in \mathbf{N}^*$, $\phi(k/n) = k\phi(1/n)$. Si on pose $r = \phi(1)$, le choix $k = n$ permet de vérifier que $\forall k, n \in \mathbf{N}^*$, $\phi(k/n) = rk/n$. On conclut alors, sans presque aucune hypothèse de régularité sur ϕ , que $\phi(t) = rt$.

De même, pour la variance, comme $X_{t+s} - X_t$ et $X_t - X_0$ sont deux variables aléatoires indépendantes, on a :

$$\text{Var}(X_{t+s} - X_0) = \text{Var}(X_{t+s} - X_t) + \text{Var}(X_t - X_0) = \text{Var}(X_s - X_0) + \text{Var}(X_t - X_0).$$

Le même argument permet alors d'affirmer que $\text{Var}(X_t - X_0) = \sigma^2 t$. On vérifie ensuite facilement que $B_t = \frac{1}{\sigma}(X_t - X_0 - rt)$ est un mouvement brownien. ■

Remarque 2.2.5 Si on souhaite modéliser le cours $(S_t)_{t \geq 0}$ d'une action par un processus continu strictement positif à accroissements relatifs

$$\frac{S_t - S_u}{S_u} = \frac{S_t}{S_u} - 1, \quad u \leq t$$

indépendants et stationnaires alors $X_t = \ln(S_t)$ est un processus continu à accroissements indépendants et stationnaires. Dans ces conditions, d'après le théorème 2.2.1, il existe deux constantes réelles r et σ et un mouvement brownien $(B_t, t \geq 0)$ t.q.

$$\forall t \geq 0, S_t = S_0 \exp(\sigma B_t + rt).$$

Ce modèle d'actif est appelé modèle de Black-Scholes.

Exercice 16 Montrer que pour tout $\lambda \in \mathbf{R}$, $\mathbf{E}(e^{\lambda B_t}) = e^{\frac{\lambda^2 t}{2}}$. En déduire que

$$\mathbb{E}(B_t^n) = \begin{cases} 0 & \text{si } n \text{ impair} \\ \frac{t^{n/2} n!}{2^{n/2} (n/2)!} & \text{si } n \text{ pair.} \end{cases}$$

2.3 Vers une construction du mouvement brownien

Nous allons donner une réponse (partielle) à la question : Le mouvement brownien (standard) existe t'il ? Le but de cette partie sera d'écrire un programme informatique permettant de simuler une trajectoire d'un mouvement brownien.

Pour mener à bien cette construction, nous allons chercher à identifier la "loi conditionnelle de $B_{\frac{t+s}{2}}$ lorsque l'on connaît B_t et B_s ".

Pour cela, notons que si $s \leq t$ alors le vecteur :

$$(B_s, B_{\frac{t+s}{2}}, B_t)$$

est un vecteur gaussien puisque il s'écrit linéairement à partir du vecteur de gaussiennes *indépendantes* :

$$(B_s, B_{\frac{t+s}{2}} - B_s, B_t - B_{\frac{t+s}{2}}).$$

L'indépendance de ces trois variables aléatoires est une conséquence de la définition : les accroissements du brownien sont indépendants.

On cherche à déterminer α et β de façon à ce que la variable aléatoire $Z_{\alpha,\beta}$:

$$Z_{\alpha,\beta} = B_{\frac{t+s}{2}} - \alpha B_s - \beta B_t,$$

soit indépendante du couple (B_s, B_t) .

Pour cela, il convient de noter que le vecteur $(Z_{\alpha,\beta}, B_s, B_t)$ est un vecteur gaussien (comme combinaison linéaire de variables aléatoires gaussiennes indépendantes) centré (puisque $\mathbf{E}(B_t) = \mathbf{E}(B_{\frac{t+s}{2}}) = \mathbf{E}(B_s) = 0$ et donc $\mathbf{E}(Z_{\alpha,\beta}) = 0$).

Ce vecteur étant gaussien, d'après l'exercice 15, $Z_{\alpha,\beta}$ sera indépendante du couple (B_s, B_t) si et seulement si :

$$\text{Cov}(Z_{\alpha,\beta}, B_s) = 0 \text{ et } \text{Cov}(Z_{\alpha,\beta}, B_t) = 0.$$

On obtient, sans difficultés, que :

$$\text{Cov}(Z_{\alpha,\beta}, B_s) = \mathbf{E}\left(B_{\frac{t+s}{2}} B_s\right) - \alpha \mathbf{E}(B_s^2) - \beta \mathbf{E}(B_t B_s).$$

On voit donc qu'il faut savoir calculer, pour un mouvement brownien standard $\mathbf{E}(B_u B_v)$ si $u \leq v$. Pour cela remarquons que :

$$\mathbf{E}(B_u B_v) = \mathbf{E}(B_u^2 + B_u(B_v - B_u)) = \mathbf{E}(B_u^2) + \mathbf{E}(B_u(B_v - B_u)).$$

Mais, B_u et $B_v - B_u$ sont indépendantes et $\mathbf{E}(B_u) = 0$ donc :

$$\mathbf{E}(B_u(B_v - B_u)) = 0.$$

De plus, B_u suit une loi gaussienne centrée de variance u donc $\mathbf{E}(B_u^2) = u$. On voit donc que, si $u \leq v$:

$$\mathbf{E}(B_u B_v) = u.$$

Plus généralement, cela implique que pour tout t et pour tout s :

$$\mathbf{E}(B_t B_s) = t \wedge s = \inf(s, t).$$

Ce résultat intermédiaire étant établi, on obtient :

$$\text{Cov}(Z_{\alpha,\beta}, B_s) = \mathbf{E}\left(B_{\frac{t+s}{2}} B_s\right) - \alpha \mathbf{E}(B_s^2) - \beta \mathbf{E}(B_t B_s) = s - \alpha s - \beta s,$$

et :

$$\text{Cov}(Z_{\alpha,\beta}, B_t) = \mathbf{E}\left(B_{\frac{t+s}{2}} B_t\right) - \alpha \mathbf{E}(B_t B_s) - \beta \mathbf{E}(B_t^2) = \frac{t+s}{2} - \alpha s - \beta t.$$

On a donc $\text{Cov}(Z_{\alpha,\beta}, B_s) = \text{Cov}(Z_{\alpha,\beta}, B_t) = 0$ si et seulement si $\alpha = \beta = 1/2$. Posons $Z = Z_{1/2,1/2}$, soit :

$$Z = B_{\frac{t+s}{2}} - \frac{1}{2}(B_s + B_t).$$

Z est par construction une variable aléatoire indépendante du couple (B_s, B_t) . De plus elle suit une loi gaussienne puisque elle s'exprime linéairement en fonction du vecteur gaussien $(B_s, B_{\frac{t+s}{2}}, B_t)$. Pour identifier la loi de Z il suffit de calculer sa moyenne et sa variance. On obtient facilement que $\mathbf{E}(Z) = 0$ puis que :

$$\text{Var}(Z) = \mathbf{E}(Z^2) = \mathbf{E}\left\{\left(B_{\frac{t+s}{2}} - \frac{1}{2}(B_s + B_t)\right)^2\right\}.$$

En développant le carré, on obtient :

$$\text{Var}(Z) = \frac{t+s}{2} + \frac{t}{4} + \frac{s}{4} - \frac{t+s}{2} - s + \frac{s}{2} = \frac{1}{4}(t-s).$$

On voit donc que Z s'écrit sous la forme :

$$Z = \frac{1}{2}\sqrt{t-s} G_{s,t},$$

$G_{s,t}$ étant une gaussienne centrée réduite indépendante du couple (B_s, B_t) .

Remarque 2.3.1 On peut montrer que $G_{s,t}$ est en fait indépendante de B_u pour tout $u \leq s$ ou $u \geq t$.

On peut résumer le résultat qui vient d'être établi de la façon suivante :

$$\begin{cases} B_{\frac{t+s}{2}} = \frac{1}{2}(B_s + B_t) + \frac{1}{2}\sqrt{t-s} G_{s,t} \\ G_{s,t} \text{ étant indépendante de } B_u \text{ pour tout } u \leq s \text{ ou } u \geq t. \end{cases} \quad (2.1)$$

Cela suggère la procédure de simulation (ou de construction) suivante de la trajectoire $(B_s, 0 \leq s \leq 1)$ à partir d'une suite $(G_j)_{j \geq 1}$ de variables gaussiennes centrées réduites indépendantes :

1. Poser $B_1 = G_1$,
2. utiliser (2.1), avec $s = 0$ et $t = 1$ pour poser $B_{1/2} = \frac{1}{2}(B_1 + G_2)$,
3. utiliser (2.1), avec $s = 0$ et $t = 1/2$ pour poser $B_{1/4} = \frac{1}{2}(B_{1/2} + \frac{1}{\sqrt{2}}G_2)$,
4. utiliser (2.1), avec $s = 1/2$ et $t = 1$ pour poser $B_{3/4} = \frac{1}{2}(B_{1/2} + B_1 + \frac{1}{\sqrt{2}}G_4)$,
5. etc ...

Il est facile d'écrire un programme C qui fait ce travail.

```
double gaussienne()
// renvoie une gaussienne centrée réduite
{
    return ...
}

void Pont(s,x,t,y)
```



```

{
  double z;

  if( |t-s| < epsilon)
  {
    tracer la droite de (s,x) à (t,y);
    return;
  }
  z = (x+y)/2 + sqrt(t-s)/2 * gaussienne();
  Pont(s,x,(s+t)/2,z);
  Pont((s+t)/2,z,t,y);
}

void main(void)
{
  Pont(0,0,1,gaussienne());
}

```

2.4 Régularité des trajectoires

Une propriété importante (bien que parfois gênante !) du mouvement brownien est le manque de régularité des trajectoires. Nous admettrons le théorème suivant.

Théorème 2.4.1 Soit $(B_t, t \geq 0)$ un mouvement brownien, alors, en dehors d'un ensemble de probabilité nulle, il n'existe aucun point où la trajectoire est différentiable.

Nous allons prouver un résultat lié à ce théorème bien que moins fort.

Proposition 2.4.1 Soit $(B_t, t \geq 0)$ un mouvement brownien, soit T un réel positif, n un entier. On pose :

$$t_i^n = \frac{iT}{n}, \text{ pour } 0 \leq i \leq n.$$

Alors, au sens de la convergence L^2 :

$$\lim_{n \rightarrow +\infty} \sum_{i=0}^{n-1} (B_{t_{i+1}^n} - B_{t_i^n})^2 = T.$$

Démonstration : En utilisant l'indépendance des accroissements de $(B_t, t \geq 0)$ on obtient :

$$\text{Var} \left(\sum_{i=0}^{n-1} (B_{t_{i+1}^n} - B_{t_i^n})^2 \right) = \sum_{i=0}^{n-1} \text{Var} \left((B_{t_{i+1}^n} - B_{t_i^n})^2 \right).$$

Mais :

$$\text{Var} \left((B_{t_{i+1}^n} - B_{t_i^n})^2 \right) = \mathbf{E} \left((B_{t_{i+1}^n} - B_{t_i^n})^4 \right) - \mathbf{E} \left((B_{t_{i+1}^n} - B_{t_i^n})^2 \right)^2.$$

La loi de $B_{t_{i+1}^n} - B_{t_i^n}$ est celle d'une gaussienne centrée de variance $t_{i+1}^n - t_i^n = T/n$, donc si G est une gaussienne centrée réduite, on a :

$$\text{Var} \left((B_{t_{i+1}^n} - B_{t_i^n})^2 \right) = \frac{T^2}{n^2} \left(\mathbf{E}(G^4) - \mathbf{E}(G^2)^2 \right) = \frac{T^2}{n^2} (\mathbf{E}(G^4) - 1).$$

De plus :

$$\mathbf{E} \left(\sum_{i=0}^{n-1} \left(B_{t_{i+1}^n} - B_{t_i^n} \right)^2 \right) = \sum_{i=0}^{n-1} \mathbf{E} \left(\left(B_{t_{i+1}^n} - B_{t_i^n} \right)^2 \right) = n \frac{T}{n} = T.$$

Finalement on obtient :

$$\mathbf{E} \left(\left[\sum_{i=0}^{n-1} \left(B_{t_{i+1}^n} - B_{t_i^n} \right)^2 - T \right]^2 \right) = \frac{T^2}{n} (\mathbf{E}(G^4) - 1) = \frac{2T^2}{n}.$$

Ce qui prouve le résultat. ■

Remarque 2.4.2 Ceci implique que $(B_t, t \geq 0)$ ne peut être lipschitzienne sur l'intervalle $[0, T]$. En effet la convergence L^2 implique la convergence presque sûre pour une sous suite et, de plus, si $|B_t - B_s| \leq K|t - s|$, on a :

$$\sum_{i=0}^{n-1} \left(B_{t_{i+1}^n} - B_{t_i^n} \right)^2 \leq \sum_{i=0}^{n-1} \frac{K^2 T^2}{n^2} \leq \frac{K^2 T}{n} \rightarrow_{n \rightarrow +\infty} 0.$$

Ceci prouve que la trajectoire ne peut être lipschitzienne sur l'intervalle $[0, T]$.

2.5 Caractère gaussien du mouvement brownien

Nous avons vu que si $(B_t, t \geq 0)$ est un mouvement brownien alors B_t suit une loi gaussienne. Une propriété plus forte est vérifiée par le processus $(B_t, t \geq 0)$: c'est un processus gaussien.

Définition 2.5.1 On dit qu'un processus $(X_t, t \geq 0)$ est un *processus gaussien*, si pour tout entier n et pour tout n -uplet, $0 \leq t_1 < t_2 < \dots < t_n < +\infty$, le vecteur $(X_{t_1}, \dots, X_{t_n})$ est un *vecteur gaussien*.

Il est alors facile de vérifier que :

Théorème 2.5.1 *Un mouvement brownien est un processus gaussien.*

Démonstration : Appelons $(B_t, t \geq 0)$ ce mouvement brownien. Alors

$$(B_{t_1}, B_{t_2} - B_{t_1}, \dots, B_{t_n} - B_{t_{n-1}})$$

est un vecteur formé de gaussiennes indépendantes, donc un vecteur gaussien. Le vecteur $(B_{t_1}, \dots, B_{t_n})$ qui s'obtient par transformation linéaire du précédent est également gaussien. ■

On peut de plus prouver le résultat suivant.

Théorème 2.5.2 *Soit $(B_t, t \geq 0)$ un processus gaussien centré continu t.q.*

$$\forall s, t \geq 0, \text{Cov}(B_s, B_t) = \min(s, t).$$

Alors $(B_t, t \geq 0)$ est un mouvement brownien.

Démonstration : Le point (iv) de la définition 2.2.3 est vérifié. Comme $\mathbf{E}(B_0) = 0$ (processus centré) et $\text{Var}(B_0) = \min(0, 0) = 0$, B_0 est nul avec probabilité 1 et le point (i) est vérifié.

Pour les points (ii) et (iii) on se donne $0 = t_0 \leq t_1 \leq \dots \leq t_n$. Pour $1 \leq i \leq n$,

$$\text{Var}(B_{t_i} - B_{t_{i-1}}) = \text{Var}(B_{t_i}) - 2\text{Cov}(B_{t_i}, B_{t_{i-1}}) + \text{Var}(B_{t_{i-1}}) = t_i - 2\min(t_i, t_{i-1}) + t_{i-1} = t_i - t_{i-1}.$$

Et pour $1 \leq i < j \leq n$,

$$\begin{aligned} \text{Cov}(B_{t_i} - B_{t_{i-1}}, B_{t_j} - B_{t_{j-1}}) &= \text{Cov}(B_{t_i}, B_{t_j}) - \text{Cov}(B_{t_i}, B_{t_{j-1}}) - \text{Cov}(B_{t_{i-1}}, B_{t_j}) + \text{Cov}(B_{t_{i-1}}, B_{t_{j-1}}) \\ &= \min(t_i, t_j) - \min(t_i, t_{j-1}) - \min(t_{i-1}, t_j) + \min(t_{i-1}, t_{j-1}) \\ &= t_i - t_i - t_{i-1} + t_{i-1} = 0. \end{aligned}$$

Comme (B_t) est un processus gaussien, on conclut que les variables $B_{t_1}, B_{t_2} - B_{t_1}, \dots, B_{t_n} - B_{t_{n-1}}$ sont des gaussiennes centrées indépendantes de variances respectives $t_1, t_2 - t_1, \dots, t_n - t_{n-1}$. ■

Exercice 17 Montrer que si $(B_t, t \geq 0)$ est un mouvement brownien alors :

- $(-B_t, t \geq 0)$
- $(\frac{1}{c}B_{c^2t}, t \geq 0)$ où $c \neq 0$,
- $(B_{t_0+t} - B_{t_0}, t \geq 0)$ où $t_0 \geq 0$

sont des mouvements browniens. En déduire que $\text{Loi}(\int_0^T B_s ds) = \text{Loi}(T^{3/2} \int_0^1 B_s ds)$.

Une application du caractère gaussien du mouvement brownien Quelle est la loi de $\int_0^T B_s ds$ si $(B_t, t \geq 0)$ est un mouvement brownien ?

Comme $(B_t, t \geq 0)$ est un processus continu, on a, au sens d'une limite presque sûre :

$$\lim_{n \rightarrow +\infty} \frac{T}{n} \sum_{i=0}^{n-1} B_{t_i^n} = \int_0^T B_s ds,$$

si $t_i^n = \frac{iT}{n}$ (approximation d'une intégrale par une somme de Riemann).

Comme le processus $(B_t, t \geq 0)$ est un processus gaussien les variables aléatoires $\frac{T}{n} \sum_{i=0}^{n-1} B_{t_i^n}$ sont des gaussiennes centrées. Avec la proposition 2.1.5, on en déduit que $\int_0^T B_s ds$ est une gaussienne. Comme $\mathbf{E}(\int_0^T B_s ds) = \int_0^T \mathbf{E}(B_s) ds = 0$, pour identifier sa loi il nous reste à calculer sa variance :

$$\begin{aligned} \text{Var}\left(\int_0^T B_s ds\right) &= \mathbf{E}\left(\left[\int_0^T B_s ds\right]^2\right) \\ &= \mathbf{E}\left(\int_0^T \int_0^T B_s B_{s'} ds ds'\right) \\ &= \int_0^T \int_0^T \mathbf{E}(B_s B_{s'}) ds ds'. \end{aligned}$$

Comme $\mathbf{E}(B_s B_t) = \min(s, t)$, un calcul simple montre que :

$$\text{Var}\left(\int_0^T B_s ds\right) = \frac{T^3}{3}.$$

Exercice 18 Montrer que le couple $(B_T, \int_0^T B_s ds)$ est un vecteur gaussien et identifier sa loi.

2.6 Exercices

Exercice 19 Soit X et Y des variables gaussiennes réelles indépendantes. Donner une condition nécessaire et suffisante pour que les variables $X + Y$ et $X - Y$ soient indépendantes.

Exercice 20 Soit X une variable gaussienne centrée réduite et ε une variable vérifiant $\mathbf{P}(\varepsilon = 1) = \mathbf{P}(\varepsilon = -1) = 1/2$ indépendantes.

1. Quelle est la loi de $Y = \varepsilon X$?
2. Donner la matrice de covariance du vecteur (X, Y) .
3. Comparer $\mathbf{E}(X^2 Y^2)$ à $\mathbf{E}(X^2)\mathbf{E}(Y^2)$. Les variables X et Y sont-elles indépendantes ?
4. Que peut-on conclure sur le couple (X, Y) ?

Exercice 21 Soit X_1, \dots, X_n des variables gaussiennes centrées réduites indépendantes. On note $M_n = \frac{1}{n} \sum_{i=1}^n X_i$ et $V_n = \frac{1}{n-1} \sum_{i=1}^n (X_i - M_n)^2$ la moyenne et la variance empirique.

1. Montrer que M_n est indépendante du vecteur $(X_1 - M_n, X_2 - M_n, \dots, X_n - M_n)$.
2. Conclure que M_n et V_n sont indépendantes.

Exercice 22 Dans toute la suite on suppose que $(B_t, t \geq 0)$ est un mouvement brownien standard.

1. Calculer $\mathbf{E}(|B_t|)$.
2. Calculer $\mathbf{E}(B_t^2 B_s^2)$.

Exercice 23 Soient f_1, f_2 deux fonctions continues sur $[0, 1]$. On note $X_j = \int_0^1 B_s f_j(s) ds$.

1. Montrer que presque sûrement

$$X_j = \lim_{n \rightarrow \infty} \frac{1}{n} \sum_{k=0}^{n-1} B_{k/n} f_j(k/n).$$

2. En déduire que (X_1, X_2) est un vecteur gaussien.
3. Calculer la moyenne et la matrice de variance covariance de ce vecteur.

Exercice 24 Lemme de Borel-Cantelli

1. Soit $(X_k, k \geq 1)$, une suite de variables aléatoires réelles, montrer que l'on a toujours :

$$\mathbf{E} \left(\sum_{k \geq 1} |X_k| \right) = \sum_{k \geq 1} \mathbf{E} (|X_k|).$$

2. En déduire que si $\mathbf{E} \left(\sum_{k \geq 1} |X_k| \right) < +\infty$, alors, presque sûrement $\lim_{k \rightarrow +\infty} X_k = 0$.
3. On considère une suite d'évènements $(A_k, k \geq 1)$. Montrez, en posant $X_k = \mathbf{1}_{A_k}$ que si $\sum_{k \geq 1} \mathbf{P}(A_k) < +\infty$ alors, presque sûrement, il existe un rang $N(\omega)$ (dépendant du hasard) tel que si $k \geq N(\omega)$ alors $\omega \notin A_k$.

Ce résultat, très utile pour prouver des convergences, est connu sous le nom de Lemme de Borel-Cantelli. Il sera utilisé dans l'exercice suivant.

Exercice 25 1. On note $Z_t = tB_{1/t}$. Montrer que le vecteur $(Z_{t_1}, Z_{t_2} - Z_{t_1}, \dots, Z_{t_n} - Z_{t_{n-1}})$, où $0 < t_1 < \dots < t_n$, est un vecteur gaussien dont on déterminera les caractéristiques. En admettant que $\lim_{t \rightarrow 0} Z_t = 0$, montrer que $(Z_t, t \geq 0)$ est un mouvement brownien. Le reste de cet exercice consiste à prouver que $\lim_{t \rightarrow 0} Z_t = 0$.

2. Montrer que presque sûrement $\lim_{n \rightarrow \infty} B_n/n = 0$.

3. On considère $M_n = \sup\{|B_t - B_n|; n \leq t \leq n+1\}$. Montrer que $(M_n, n \geq 0)$ est une suite de variables aléatoires indépendantes de même loi. **On admettra dans la suite de cet exercice que $\mathbf{E}(|M_1|) < +\infty$.**

4. Montrer que, pour tout $\epsilon > 0$

$$\sum_{n \geq 1} \mathbf{P}(M_n \geq \epsilon n) = \sum_{n \geq 1} \mathbf{P}(M_1 \geq \epsilon n) \leq \frac{1}{\epsilon} \mathbf{E}(|M_1|).$$

En utilisant le fait que $\mathbf{E}(|M_1|) < +\infty$ et le lemme de Borel-Cantelli en déduire que, presque sûrement, $\lim_{n \rightarrow +\infty} M_n/n = 0$.

5. Montrez que pour $n \leq t < n+1$, on a

$$\left| \frac{B_t}{t} - \frac{B_n}{n} \right| \leq \frac{|B_n|}{n(n+1)} + \frac{M_n}{n}.$$

En déduire que presque sûrement

$$\lim_{t \rightarrow \infty} B_t/t = 0.$$

Exercice 26 Soit $0 \leq a < b < \infty$ fixés. Soit τ une subdivision finie de $[a, b]$: $\tau = (t_0, \dots, t_n)$ où $a = t_0 < \dots < t_n = b$. Le pas $\Delta(\tau)$ de la subdivision τ est

$$\Delta(\tau) = \sup\{t_{i+1} - t_i, 0 \leq i < n\}.$$

Enfin on note $X_\tau = \sum_{i=1}^n [B_{t_{i+1}} - B_{t_i}]^2$. Soit $(\tau_k, k \geq 1)$ une suite de subdivisions dont le pas tend vers zéro.

1. Montrer que la suite $(X_{\tau_k}, k \geq 1)$ converge dans $L^2(\mathbf{P})$ vers $b - a$.

2. Montrer qu'il existe une sous-suite $(k_l, l \geq 1)$ telle que la suite $(X_{\tau_{k_l}}, l \geq 1)$ converge presque sûrement vers $b - a$ (On utilisera le lemme de Borel-Cantelli).

3. Soit $\epsilon > 0$. Déduire de la question précédente que presque sûrement

$$\sup_{(s,t) \in [a,b]^2, s \neq t} \frac{|B_s - B_t|}{(s-t)^{\frac{1}{2} + \epsilon}} = \infty.$$

Donner une interprétation intuitive en terme de régularité d'une trajectoire d'un mouvement brownien ?

2.7 Travail dirigé : Test d'adéquation à une loi

On vous propose de jouer avec un dé et cinq partenaires. Tout le monde mise 1F, on lance le dé. La face exposée détermine le gagnant qui empêche toutes les mises. Calculer votre gain moyen au bout de n coups. Pour accepter de jouer, il vous faut savoir si le dé est biaisé ou non. Notre but est d'établir une procédure simple pour déterminer si le dé est non-biaisé. On parle de test d'adéquation à une loi.

On note $(X_i, i \geq 1)$ des variables aléatoires indépendantes et de même loi à valeurs dans $\{x_1, \dots, x_k\}$, dont la loi est caractérisée par $p_j = \mathbf{P}[X_1 = x_j] > 0$ et $\sum_{j=1}^k p_j = 1$. (Dans l'exemple ci-dessus $k = 6$.) On note

$$N_j = \sum_{i=1}^n \mathbf{1}_{X_i=x_j} \quad \text{et} \quad T_n = \sum_{j=1}^k \frac{(N_j - np_j)^2}{np_j}.$$

1. Déterminer la loi de N_j .

Le cas d'une pièce : $k = 2$

1. Déterminer la loi du couple (N_1, N_2) .
2. Déterminer le comportement de la suite $(N_1/n, n \geq 1)$.
3. Déterminer le comportement de la suite $(\frac{N_1 - np_1}{\sqrt{n}}, n \geq 1)$.

La loi du χ^2

1. Soit $(X_n, n \geq 1)$ une suite de variables aléatoires vectorielles qui converge en loi vers X . Soit f une fonction continue. Déterminer le comportement de la suite $(f(X_n), n \geq 1)$.
2. Calculer T_n . Comportement de $(T_n, n \geq 1)$.
3. Donner la densité de la loi limite. Identifier la loi limite comme une loi Gamma : $\Gamma(1/2, d/2)$ avec $d = 1$. Rappeler sa fonction caractéristique. On parle également de loi du χ^2 à d degrés de liberté. Ces lois sont tabulées.
4. En déduire un test pour déterminer si une pièce est non-biaisée.

Le cas général

1. Donner la loi de la somme de deux χ^2 indépendants à k_1 et k_2 degrés de liberté. En déduire la loi de $\sum_{j=1}^k X_j^2$, où X_1, \dots, X_k sont k gaussiennes indépendantes $\mathcal{N}(0, 1)$. On pose

$$Z_n = \left(\frac{N_1 - np_1}{\sqrt{np_1}}, \dots, \frac{N_k - np_k}{\sqrt{np_k}} \right).$$

2. Donner la limite en loi, Z , de $(Z_n, n \geq 1)$ quand $n \rightarrow \infty$. Déterminer la matrice de variance-covariance Γ .
3. Calculer $\sum_{j=1}^k \sqrt{p_j} Z_n(j)$. En déduire que Γ est de rang strictement inférieur à k .
4. Montrer que $\Gamma^2 = \Gamma$. Montrer que Γ est de rang $k - 1$. On pourra considérer les vecteurs e_k et Γe_k où $e_k(1) = \sqrt{p_k}$, $e_k(k) = -\sqrt{p_1}$ et $e_k(i) = 0$ si $i \notin \{1; k\}$ pour $k \geq 2$. En déduire qu'il existe une matrice unitaire U telle que $U^t \Gamma U$ est diagonale.
5. Donner la loi de $U^t Z$. En déduire la loi de $\sum_{j=1}^k Z(j)^2$.
6. Donner la limite de $(T_n, n > 1)$. En déduire un test pour savoir si un dé est non-biaisé.

Exemples numériques

1. On considère les naissances aux U.S.A. On souhaite savoir si les naissances sont équi-réparties sur les jours de la semaine. Les nombres moyens de naissances par jours de semaine pour 1997 sont les suivants (source NVSR 1999) : $N_{\text{Lundi}} = 10861$, $N_{\text{Mardi}} = 12104$, $N_{\text{Mercredi}} = 11723$, $N_{\text{Jeudi}} = 11631$, $N_{\text{Vendredi}} = 11640$, $N_{\text{Samedi}} = 8670$ et $N_{\text{Dimanche}} = 7778$. On obtient $T_n = 1639$. On lit dans la table du χ^2 à 6 degrés de liberté que $\mathbf{P}(\chi^2(6) > 23) = 0,1\%$. On rejette donc l'hypothèse des naissances uniformes sur les jours de la semaine au niveau 0,1%.
2. Weldon a effectué $n = 26306$ lancers de 12 dés à six faces. On note X_i le nombre de faces indiquant 5 ou 6 lors du i -ième lancer. Les fréquences empiriques observées sont notées $\hat{f}_j = \frac{N_j}{n}$, où N_j est le nombre de fois où l'on a observé j faces 5 et 6 sur les 12 dés lancés : $N_j = \sum_{i=1}^n \mathbf{1}_{X_i=j}$. Les résultats obtenus sont : Si les dés sont non biaisés,

$\hat{f}_0 = 0,007033$	$\hat{f}_1 = 0,043678$	$\hat{f}_2 = 0,124116$	$\hat{f}_3 = 0,208127$
$\hat{f}_4 = 0,232418$	$\hat{f}_5 = 0,197445$	$\hat{f}_6 = 0,116589$	$\hat{f}_7 = 0,050597$
$\hat{f}_8 = 0,015320$	$\hat{f}_9 = 0,003991$	$\hat{f}_{10} = 0,000532$	$\hat{f}_{11} = 0,000152$
$\hat{f}_{12} = 0,000000$			

la probabilité d'observer les faces 5 ou 6 dans un lancer de dé est $1/3$. Les variables aléatoires ($X_i, 1 \leq i \leq n$) suivent donc une loi binomiale de paramètre $(12, 1/3)$. Les fréquences théoriques sont :

$f_0 = 0,007707$	$f_1 = 0,046244$	$f_2 = 0,127171$	$f_3 = 0,211952$
$f_4 = 0,238446$	$f_5 = 0,190757$	$f_6 = 0,111275$	$f_7 = 0,047689$
$f_8 = 0,014903$	$f_9 = 0,003312$	$f_{10} = 0,000497$	$f_{11} = 0,000045$
$f_{12} = 0,000002$			

Calculer T_n , et en déduire que l'on rejette l'hypothèse des dés non biaisés. Les données sont issues du livre de Feller Tome 1 : "An introduction to probability theory and its applications", page 149.

Chapitre 3

Espérance conditionnelle

3.1 Tribus construites à partir de variables aléatoires

Soit Ω un ensemble muni d'une tribu \mathcal{A} c'est-à-dire un sous-ensemble de l'ensemble $\mathcal{P}(\Omega)$ des parties de Ω

- contenant \emptyset et Ω
- stable par passage au complémentaire
- stable par union ou intersection dénombrable.

Rappelons la définition d'une variable aléatoire réelle sur (Ω, \mathcal{A}) :

Définition 3.1.1 Une application X de Ω muni de la tribu \mathcal{A} dans \mathbf{R} est une variable aléatoire si pour tout B dans la tribu borélienne de \mathbf{R} (plus petite tribu qui contient tous les intervalles) :

$$\{\omega \in \Omega, X(\omega) \in B\} = \{X \in B\} \in \mathcal{A}.$$

Cette propriété est la \mathcal{A} -mesurabilité de X .

Une sous-tribu \mathcal{B} de \mathcal{A} est une tribu sur Ω t.q. $\mathcal{B} \subset \mathcal{A}$.

- Soit X une variable aléatoire sur (Ω, \mathcal{A}) à valeurs dans \mathbf{R} . On note : $\sigma(X)$ la plus petite sous tribu \mathcal{B} de \mathcal{A} , rendant l'application X \mathcal{B} -mesurable. Cette tribu représente l'information donnée par la connaissance de X . Notez, par exemple, que si $A \in \mathcal{A}$, alors la tribu engendrée par la variable aléatoire $\mathbf{1}_A$ est égale à :

$$\{\emptyset, A, A^c, \Omega\}.$$

- Soit X_1, \dots, X_n , n variables aléatoires sur (Ω, \mathcal{A}) à valeurs dans \mathbf{R} , on note :

$$\sigma(X_1, \dots, X_n)$$

la plus petite sous tribu \mathcal{B} de \mathcal{A} rendant les applications X_1, \dots, X_n \mathcal{B} -mesurables. Cette tribu représente l'information donnée par la connaissance des variables aléatoires X_1, \dots, X_n .

- Soit $(X_n, n \geq 1)$ une suite de variables aléatoires réelles \mathcal{A} -mesurables. On note, pour tout $n \geq 1$:

$$\mathcal{F}_n = \sigma(X_k, 1 \leq k \leq n).$$

Alors $(\mathcal{F}_n, n \geq 1)$ est une suite croissante de tribus qui porte le nom de *filtration naturelle*. Pour un n fixé, \mathcal{F}_n représente l'information disponible à l'instant n .

Le résultat suivant dont la preuve fait l'objet de l'exercice 35 clarifie l'interprétation d'une tribu comme information.

Proposition 3.1.2 Soit X et Y deux variables aléatoires réelles. Alors Y est $\sigma(X)$ -mesurable si et seulement si il existe une fonction¹ f de \mathbf{R} dans \mathbf{R} , telle que :

$$\mathbf{P}(Y = f(X)) = 1.$$

Remarque 3.1.3 Si Y est $\sigma(X)$ -mesurable et, si l'on connaît la valeur de X , on peut en déduire la valeur de Y .

3.2 Notion d'espérance conditionnelle

Soit X une variable aléatoire réelle définie sur (Ω, \mathcal{A}) et soit $\mathcal{B} \subset \mathcal{A}$ une sous tribu de \mathcal{A} . L'espérance conditionnelle vise à définir la notion de *meilleure approximation* de X par une variable aléatoire \mathcal{B} -mesurable.

Cette notion est, évidemment, relative au choix d'une probabilité \mathbf{P} sur (Ω, \mathcal{A}) . Nous allons commencer par traiter le cas où X est une variable aléatoire de carré intégrable, c'est à dire telle que $\mathbf{E}(X^2) < +\infty$ (on le note aussi $X \in L^2(\mathbf{P})$).

La construction se fait en utilisant le théorème de projection sur un sous espace fermé d'un espace de Hilbert.

Pour cela posons :

- $\mathcal{H}_{\mathcal{A}} = \{X \text{ variable aléatoire } \mathcal{A}\text{-mesurable telle que } \mathbf{E}(X^2) < +\infty\}$,
- $\mathcal{H}_{\mathcal{B}} = \{X \text{ variable aléatoire } \mathcal{B}\text{-mesurable telle que } \mathbf{E}(X^2) < +\infty\}$.

Alors $\mathcal{H}_{\mathcal{A}}$ est un espace de Hilbert : c'est un espace vectoriel complet muni d'une norme dérivant du produit scalaire $\langle X, Y \rangle = \mathbf{E}(XY)$ et $\|X\|^2 = \langle X, X \rangle$. $\mathcal{H}_{\mathcal{B}}$ est un sous espace vectoriel fermé inclus dans cet espace de Hilbert. On peut alors montrer, en utilisant le théorème de projection sur un sous espace fermé dans un espace de Hilbert, le résultat suivant.

Proposition 3.2.1 Si X est une variable aléatoire telle que $\mathbf{E}(X^2) < +\infty$, il existe une unique variable aléatoire Z , telle que $\mathbf{E}(Z^2) < +\infty$, \mathcal{B} -mesurable et vérifiant :

$$\mathbf{E}((X - Z)^2) = \inf_{Y \in \mathcal{H}_{\mathcal{B}}} \mathbf{E}((X - Y)^2).$$

Z est, ainsi, la variable aléatoire la "plus proche" de X au sens de la norme de l'espace de Hilbert $\mathcal{H}_{\mathcal{A}}$, c'est à dire de la norme $\sqrt{\mathbf{E}(X^2)}$. Z est l'**espérance conditionnelle** de X sachant \mathcal{B} que l'on note :

$$\mathbf{E}(X|\mathcal{B}).$$

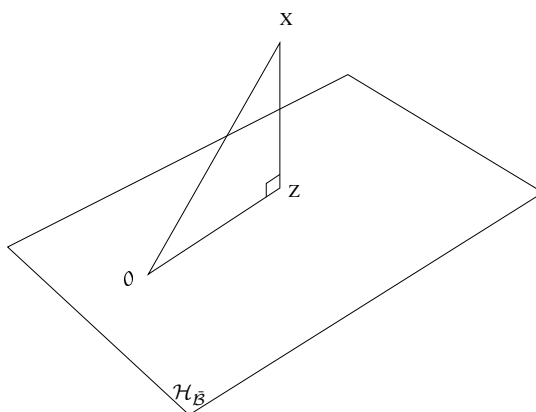


FIG. 3.1 – Projection orthogonale sur $\mathcal{H}_{\mathcal{B}}$

¹mesurable au sens où l'image réciproque par f de tout élément de la tribu borélienne de \mathbf{R} est dans la tribu borélienne de \mathbf{R}

La définition implique que pour tout réel λ , et pour toute variable aléatoire Y de $\mathcal{H}_{\mathcal{B}}$ on a :

$$\mathbf{E}((X - Z)^2) \leq \mathbf{E}((X - Z - \lambda Y)^2).$$

On voit donc que $\lambda = 0$ réalise le minimum de la fonction :

$$f(\lambda) = \lambda^2 \mathbf{E}(Y^2) + \mathbf{E}((X - Z)^2) - 2\lambda \mathbf{E}((X - Z)Y).$$

On en déduit en écrivant $f'(0) = 0$ que :

$$\mathbf{E}((X - Z)Y) = 0.$$

Réciproquement, si Z , tel que $\mathbf{E}(Z^2) < +\infty$ vérifie

$$\mathbf{E}(ZY) = \mathbf{E}(XY),$$

pour tout Y de $\mathcal{H}_{\mathcal{B}}$, alors on peut prouver, en utilisant le théorème de Pythagore que Z minimise la distance à X dans $\mathcal{H}_{\mathcal{B}}$. On obtient donc le résultat suivant :

Proposition 3.2.2 *Il existe une unique² variable aléatoire \mathcal{B} -mesurable telle que $\mathbf{E}(Z^2) < +\infty$, vérifiant pour toute variable aléatoire Y , \mathcal{B} -mesurable et telle que $\mathbf{E}(Y^2) < +\infty$:*

$$\mathbf{E}(XY) = \mathbf{E}(ZY).$$

Cette variable aléatoire Z est la meilleure approximation, au sens de la norme $\sqrt{\mathbf{E}(X^2)}$, de X parmi les variables aléatoires \mathcal{B} -mesurables de carré intégrable.

Cette définition ne permet pas de traiter le cas de variables aléatoires seulement intégrables (i.e. telle que $\mathbf{E}(|X|) < +\infty$). L'extension au cas intégrable, pour l'essentiel identique au cas de carré intégrable, est plus technique et nous admettrons donc le résultat suivant.

Théorème 3.2.1 *Soit X une variable aléatoire réelle telle que $\mathbf{E}(|X|) < +\infty$, alors il existe une unique variable aléatoire \mathcal{B} -mesurable Z telle que $\mathbf{E}(|Z|) < +\infty$ vérifiant pour toute variable aléatoire Y \mathcal{B} -mesurable et bornée :*

$$\mathbf{E}(XY) = \mathbf{E}(ZY).$$

On note la variable aléatoire Z sous la forme :

$$\mathbf{E}(X|\mathcal{B}).$$

C'est l'espérance conditionnelle de X sachant \mathcal{B} .

Remarque 3.2.3 Evidemment pour une variable aléatoire de carré intégrable, les deux définitions coïncident.

Remarque 3.2.4 Lorsque la tribu \mathcal{B} est celle engendrée par une variable aléatoire X ($\mathcal{B} = \sigma(X)$), on note :

$$\mathbf{E}(Y|X) := \mathbf{E}(Y|\sigma(X)).$$

Noter que comme toute variable aléatoire $\sigma(X)$ -mesurable est de la forme $f(X)$, f étant une fonction (voir la proposition 3.1.2), on a :

$$\mathbf{E}(Y|X) = f(X).$$

$\mathbf{E}(Y|X)$ est ainsi la meilleure approximation de Y par une fonction de X .

²unique au sens des variables aléatoires : deux variables aléatoires sont considérées comme égales lorsque elles sont égales avec probabilité 1

Exercice 27 Soit B_1, \dots, B_n une partition de Ω et

$$\mathcal{B} = \left\{ \bigcup_{\text{réunion finie}} B_{i_k} \right\}.$$

Soit Y une variable \mathcal{B} -mesurable. Pour $1 \leq i \leq n$, on note ω_i un élément de B_i et $y_i = Y(\omega_i)$. L'ensemble $\{Y = y_i\}$ contient ω_i et appartient à \mathcal{B} d'après l'hypothèse sur Y . En utilisant la définition de \mathcal{B} , on en déduit que $\{Y = y_i\}$ contient B_i . Ainsi $Y = \sum_{i=1}^n \mathbf{1}_{B_i} y_i$. Inversement si Y est de cette forme, alors il est facile de vérifier que Y est \mathcal{B} -mesurable.

En utilisant cette caractérisation des variables aléatoires \mathcal{B} -mesurables, vérifier que pour toute variable aléatoire réelle X intégrable sur (Ω, \mathcal{A}) ,

$$\mathbf{E}(X|\mathcal{B}) = \sum_{i: \mathbf{P}(B_i) > 0} \mathbf{1}_{B_i} \frac{\mathbf{E}(X \mathbf{1}_{B_i})}{\mathbf{P}(B_i)}.$$

Propriétés de l'espérance conditionnelle On peut démontrer (exercice sans difficultés) les propriétés suivantes de l'espérance conditionnelle qu'il faut impérativement connaître. Dans ce qui suit toutes les variables aléatoires considérées sont supposées intégrables.

1. Linéarité :

$$\mathbf{E}(\lambda X + \mu Y|\mathcal{B}) = \lambda \mathbf{E}(X|\mathcal{B}) + \mu \mathbf{E}(Y|\mathcal{B}).$$

2. Positivité :

$$\text{Si } X \geq 0, \text{ alors } \mathbf{E}(X|\mathcal{B}) \geq 0.$$

$$\text{Si } X \geq Y, \text{ alors } \mathbf{E}(X|\mathcal{B}) \geq \mathbf{E}(Y|\mathcal{B}).$$

3. Si X est intégrable et \mathcal{B} -mesurable alors :

$$\mathbf{E}(X|\mathcal{B}) = X.$$

4. Si Z est une variable aléatoire \mathcal{B} -mesurable et bornée :

$$\mathbf{E}(ZX|\mathcal{B}) = Z\mathbf{E}(X|\mathcal{B}).$$

5.

$$\mathbf{E}(\mathbf{E}(X|\mathcal{B})) = \mathbf{E}(X).$$

Soit \mathcal{C} une sous tribu de \mathcal{B} alors :

$$\mathbf{E}(\mathbf{E}(X|\mathcal{B})|\mathcal{C}) = \mathbf{E}(X|\mathcal{C}).$$

6. Si X est indépendante de la tribu \mathcal{B} (c'est à dire indépendante de toutes les variables aléatoires \mathcal{B} -mesurables) :

$$\mathbf{E}(X|\mathcal{B}) = \mathbf{E}(X).$$

7. Contraction pour la norme L^2 :

$$\mathbf{E}\left(|\mathbf{E}(X|\mathcal{B})|^2\right) \leq \mathbf{E}(|X|^2).$$

Remarque 3.2.5 On déduit de la propriété 2 :

$$|\mathbf{E}(X|\mathcal{B})| \leq \mathbf{E}(|X||\mathcal{B}),$$

et (contraction pour la norme L^1) :

$$\mathbf{E}(|\mathbf{E}(X|\mathcal{B})|) \leq \mathbf{E}(|X|).$$

Le résultat suivant est très utilisé dans les calculs.

Lemme 3.2.6 Soit \mathcal{B} une sous-tribu de \mathcal{A} et X, Y deux variables aléatoires réelles t.q. X est indépendante de \mathcal{B} (i.e. de toutes les variables \mathcal{B} -mesurables) et Y est \mathcal{B} -mesurable. Soit, en outre, f une fonction mesurable bornée de \mathbf{R}^2 dans \mathbf{R} , alors :

$$\mathbf{E}(f(X, Y)|\mathcal{B}) = \psi(Y),$$

avec :

$$\psi(y) = \mathbf{E}(f(X, y)).$$

En particulier, lorsque X et Y sont indépendantes, X est indépendante de $\sigma(Y)$ et $\mathbf{E}(f(X, Y)|Y) = \psi(Y)$ où ψ est définie dans l'énoncé du lemme.

Démonstration : Nous donnerons une démonstration dans le cas particulier où $\mathcal{B} = \sigma(Y)$ et où X et Y ont des lois admettant des densités notées respectivement $p_X(x)$ et $p_Y(y)$. Cette hypothèse technique simplifie la démonstration mais n'est pas nécessaire.

$\psi(Y)$ étant une variable aléatoire $\sigma(Y)$ -mesurable, il suffit de vérifier que pour toute variable aléatoire Z bornée $\sigma(Y)$ -mesurable bornée, donc de la forme $h(Y)$, on a :

$$\mathbf{E}(\psi(Y)h(Y)) = \mathbf{E}(f(X, Y)h(Y)).$$

Pour cela on va expliciter $\mathbf{E}(\psi(Y)h(Y))$. On a :

$$\mathbf{E}(\psi(Y)h(Y)) = \int_{\mathbf{R}} \psi(y)h(y)p_Y(y)dy,$$

mais, $\psi(y) = \mathbf{E}(f(X, y)) = \int_{\mathbf{R}} f(x, y)p_X(x)dx$, donc :

$$\mathbf{E}(\psi(Y)h(Y)) = \int_{\mathbf{R}} \left(\int_{\mathbf{R}} f(x, y)p_X(x)dx \right) h(y)p_Y(y)dy.$$

En utilisant le théorème de Fubini, on obtient :

$$\mathbf{E}(\psi(Y)h(Y)) = \int_{\mathbf{R}^2} f(x, y)h(y)p_X(x)p_Y(y)dy = \mathbf{E}(f(X, Y)h(Y)).$$

Ceci permet d'en déduire le résultat annoncé. ■

3.3 Cas d'un couple gaussien

Soit (X, Y) un couple de variables aléatoires centrées ($\mathbf{E}(X) = 0$ et $\mathbf{E}(Y) = 0$) qui est un vecteur gaussien. Nous allons chercher à calculer :

$$\mathbf{E}(X|Y).$$

Remarque 3.3.1 Rappelons que d'après la proposition 2.1.3, les coordonnées d'un vecteur gaussien sont indépendantes si et seulement si sa matrice de variance-covariance est diagonale.

Proposition 3.3.2 Soit (X, Y) un couple qui forme un vecteur gaussien centré. On notera σ_1^2 la variance de X , σ_2^2 la variance de Y (que l'on supposera strictement positive) et ρ le coefficient de corrélation³ de X et de Y :

$$\rho = \frac{\text{Cov}(X, Y)}{\sqrt{\text{Var}(X)\text{Var}(Y)}}.$$

Alors :

$$\mathbf{E}(X|Y) = \frac{\text{Cov}(X, Y)}{\text{Var}(Y)}Y = \rho \frac{\sigma_1}{\sigma_2}Y,$$

de plus :

$$X = \mathbf{E}(X|Y) + Z,$$

Z étant une variable aléatoire gaussienne centrée indépendante de Y et de variance :

$$\text{Var}(Z) = (1 - \rho^2)\sigma_1^2.$$

Remarque 3.3.3 Dans ce cas, très particulier, la meilleure approximation de X par une fonction linéaire de Y donne la meilleure approximation de X parmi toutes les fonctions de Y .

Démonstration : Cherchons $\lambda \in \mathbf{R}$ t.q. la variable $Z = X - \lambda Y$ soit indépendante de Y . Comme le couple (Z, Y) est obtenu par transformation linéaire à partir du couple gaussien centré (X, Y) , il est gaussien centré. Donc Z est indépendant de Y ssi $\mathbf{E}(ZY) = 0$ soit

$$\mathbf{E}(XY) = \lambda \mathbf{E}(Y^2).$$

Posons :

$$\lambda_0 = \frac{\text{Cov}(X, Y)}{\text{Var}(Y)} = \rho \frac{\sigma_1}{\sigma_2}$$

$$Z = X - \lambda_0 Y,$$

Z est donc indépendant de Y . Calculons maintenant $\mathbf{E}(X|Y)$, on a (par linéarité) :

$$\mathbf{E}(X|Y) = \mathbf{E}(Z|Y) + \lambda_0 \mathbf{E}(Y|Y).$$

Mais $\mathbf{E}(Y|Y) = Y$ (Y est bien sûr $\sigma(Y)$ -mesurable) et $\mathbf{E}(Z|Y) = \mathbf{E}(Z)$ (puisque Z est indépendant de Y par construction). De plus $\mathbf{E}(Z) = \mathbf{E}(X) - \lambda_0 \mathbf{E}(Y) = 0$, donc :

$$\mathbf{E}(X|Y) = \lambda_0 Y = \rho \frac{\sigma_1}{\sigma_2} Y.$$

De plus, le couple (X, Y) formant un vecteur gaussien, comme $Z = X - \lambda_0 Y$, Z suit une loi gaussienne centrée de variance donnée par :

$$\text{Var}(Z) = \text{Var}(X) + \lambda_0^2 \text{Var}(Y) - 2\lambda_0 \text{Cov}(X, Y) = \text{Var}(X) (1 - \rho^2).$$

■

³Noter que ρ est un nombre réel entre -1 et 1

3.4 Travail dirigé : Estimation bayésienne

Mesure de la qualité. On considère une entreprise qui fabrique des résistances électroniques. Les imperfections de la production entraînent que la résistance d'un composant ne correspond pas exactement à la valeur nominale m_θ . On modélise la valeur de la résistance par une variable aléatoire Θ gaussienne $\mathcal{N}(m_\theta, \sigma_\theta^2)$. Cette modélisation est issue d'une étude antérieure. Pour vérifier la qualité de production, l'ingénieur qualité prélève un composant et mesure la valeur de sa résistance. L'appareil de mesure quoique de bonne qualité induit une erreur aléatoire E sur la mesure. On suppose que l'appareil est bien calibré et que l'erreur est une variable aléatoire gaussienne $\mathcal{N}(0, \sigma_e^2)$ indépendante de la quantité mesurée. Le résultat de la mesure est donc $X = \Theta + E$. Le problème qui suit a pour but de donner une estimation de la résistance, alors que l'on connaît le résultat X de la mesure.

1. Quelle est la loi du couple (Θ, E) . En déduire la loi de X .
2. Trouver $\lambda \in \mathbf{R}$ t.q. $Z = \Theta - \lambda X$ soit indépendant de X . Quelles sont l'espérance et la variance de Z ?
3. Calculer $\hat{\Theta} = \mathbf{E}[\Theta | X]$. Interpréter les cas limites $\sigma_\theta \rightarrow 0$ et $\sigma_e \rightarrow 0$.
4. Calculer $\mathbf{E}[(\Theta - \hat{\Theta})^2]$. Cette quantité caractérise l'erreur (quadratique) de l'estimation. Pourquoi dit-on que $\hat{\Theta}$ est le meilleur estimateur de Θ ?

Où l'on dispose de plusieurs mesures. On répète n fois la mesure de la **même** résistance. À chaque mesure i l'appareil de mesure produit une erreur E_i . On suppose que les variables aléatoires $E_i, 1 \leq i \leq n$ sont indépendantes et de même loi $\mathcal{N}(0, \sigma_e^2)$. On suppose de plus que les erreurs sont indépendantes de la valeur Θ de la résistance mesurée.

1. Quelle est la loi du vecteur $(\Theta, E_1, \dots, E_n)$?
2. On note $X_i = \Theta + E_i$ le résultat de la mesure i . Pour calculer la loi de Θ sachant $X = (X_1, \dots, X_n)^t$, on applique la même méthode que précédemment. On cherche $\lambda \in \mathbf{R}^n$ tel que la variable Z définie par

$$\Theta = \lambda^t X + Z,$$

soit indépendante de X . Montrer que cette condition se traduit par

$$\lambda^t = \text{Cov}(\Theta, X^t) (\text{Var}X)^{-1}.$$

où $\text{Var}X$ désigne la matrice de variance covariance de X .

3. Montrer que $\text{Var}X = \sigma_\theta^2 \mathbf{1}_n + \sigma_e^2 \mathbf{I}_n$, où $\mathbf{1}_n$ est la matrice de taille $n \times n$ formée de 1 et \mathbf{I}_n est la matrice identité de \mathbf{R}^n .
4. Montrer que $\frac{1}{\sigma_e^2} \mathbf{I}_n - \frac{\sigma_\theta^2}{\sigma_e^2(n\sigma_\theta^2 + \sigma_e^2)} \mathbf{1}_n$ est l'inverse de $\text{Var}X$.
5. Montrer que $\lambda^t = \frac{\sigma_\theta^2}{n\sigma_\theta^2 + \sigma_e^2} (1, \dots, 1)$.
6. Montrer que

$$\mathbf{E}[Z] = \frac{\sigma_\theta^2 m_\theta}{n\sigma_\theta^2 + \sigma_e^2} \quad \text{et} \quad \text{Var}Z = \frac{\sigma_\theta^2 \sigma_e^2}{n\sigma_\theta^2 + \sigma_e^2}.$$

7. Exprimer $\hat{\Theta}_n = \mathbf{E}[\Theta | X]$ en fonction de la moyenne empirique $\bar{X}_n = \frac{1}{n} \sum_{i=1}^n X_i$ des mesures.

8. Montrer que l'erreur quadratique moyenne est

$$\mathbf{E} \left[(\Theta - \hat{\Theta}_n)^2 \right] = \frac{\sigma_\theta^2 \sigma_e^2}{n\sigma_\theta^2 + \sigma_e^2}.$$

Quel est l'intérêt de faire plusieurs mesures ?

9. Montrer que \bar{X}_n et $\hat{\Theta}_n$ convergent presque sûrement quand $n \rightarrow \infty$. Quelle est la limite. On parle d'estimateurs convergents.
10. Calculer la loi de $\sqrt{n}(\bar{X}_n - \Theta)$ et la loi de $\sqrt{n}(\hat{\Theta}_n - \Theta)$. Comparer les erreurs quadratiques moyennes de ces deux estimateurs. Conclusion.
11. Donner la loi asymptotique de $\sqrt{n}(\bar{X}_n - \Theta)$ et de $\sqrt{n}(\hat{\Theta}_n - \Theta)$. Que se passe-t-il si l'on dispose d'un grand nombre de mesures ?

3.5 Exercices

Exercice 28 Soit $(B_t)_{t \geq 0}$ un mouvement brownien et $0 < s \leq t$.

1. Calculer $\mathbf{E}(B_t | B_s)$ et $\mathbf{E}(B_s | B_t)$.
2. Déterminer $\mathbf{E}(B_t^2 | B_s)$ et $\mathbf{E}(e^{\lambda B_t} | B_s)$ où $\lambda \in \mathbf{R}$

Exercice 29 Soit A, B deux sous-ensembles mesurables de $(\Omega, \mathcal{A}, \mathbf{P})$. Calculer :

$$\mathbf{E}[\mathbf{1}_A | \mathbf{1}_B].$$

Exercice 30 Soit (Z, Z_1, \dots, Z_n) un vecteur centré de carré intégrable.

1. Montrer que si $\bar{Z} = \sum_{i=1}^n \lambda_i^0 Z_i$ est tel que, pour tout i :

$$\mathbf{E}((Z - \bar{Z})Z_i) = 0,$$

alors \bar{Z} réalise le minimum de $\mathbf{E}((Z - \sum_{i=1}^n \lambda_i Z_i)^2)$.

2. Si on suppose de plus que (Z, Z_1, \dots, Z_n) est un vecteur gaussien centré, démontrer que $Z - \bar{Z}$ est une gaussienne indépendante de (Z_1, \dots, Z_n) .

Exercice 31 Soit $(X_n, n \geq 1)$ une suite de v.a.r. i.i.d et de carré intégrable. Soit τ une variable aléatoire à valeurs dans \mathbf{N}^* , de carré intégrable et indépendante des v.a.r. précédentes. On note $S_\tau = X_1 + \dots + X_\tau$.

1. Calculer la moyenne conditionnelle $\mathbf{E}[S_\tau | \tau]$ et $\mathbf{E}[(S_\tau - \mathbf{E}[S_\tau | \tau])^2 | \tau]$, la variance conditionnelle.

2. En déduire la valeur de $\mathbf{E}[S_\tau]$ et de $\mathbf{E}[(S_\tau - \mathbf{E}[S_\tau])^2]$.

Exercice 32 1. Soit $(X_n, n \geq 1)$ une suite de v.a.r. i.i.d. telle que $\mathbf{E}[|X_1|] < \infty$. Montrer que p.s.

$$\mathbf{E}[X_1 | X_1 + X_2] = \mathbf{E}[X_2 | X_1 + X_2] = [X_1 + X_2]/2.$$

2. On note $S_n = \sum_{i=1}^n X_i$. En déduire que p.s.

$$\mathbf{E}[X_1 | \sigma(S_n, S_{n+1}, \dots)] = S_n/n.$$

Exercice 33 Soit X_0 une variable aléatoire gaussienne centrée de variance σ_0^2 et $(Y_j, j \geq 0)$ une suite infinie, indépendante de X_0 , de variables aléatoires réelles gaussiennes centrées indépendantes de variances égales à :

$$\text{Var}(Y_j) = \sigma^2, \text{ si } j \geq 0,$$

pour des réels σ_0 et σ positifs donnés.

Soit $(X_j, j \geq 0)$ la suite des variables aléatoires réelles définies par récurrence sur j à partir de X_0 par la formule, si $j \geq 0$:

$$X_{j+1} = \alpha X_j + Y_j,$$

pour un réel fixé α tel que $|\alpha| < 1$.

1. Établir que pour tout $n \geq 1$ la variable aléatoire réelle X_n est une variable aléatoire gaussienne centrée. Montrer qu'elle est indépendante de (Y_n, Y_{n+1}, \dots) .
2. Établir une relation de récurrence entre les variances $\sigma_j^2 = \text{Var}(X_j)$. À quelle condition nécessaire et suffisante sur σ_0^2, σ^2 et α les variables aléatoires ont elles la même variance. On suppose dans la suite que cette condition est vérifiée.

3. Exprimer X_{n+k} à l'aide de X_n et de Y_n, Y_{n+1}, \dots , pour $n \geq 0$ et $k > 0$. En déduire les valeurs des covariances $\text{Cov}(X_n, X_{n+k})$ en fonction des paramètres initiaux.
4. Trouver $\lambda_0, \dots, \lambda_n$ t.q. $X_{n+k} - \sum_{i=0}^n \lambda_i X_i$ soit indépendante de (X_0, \dots, X_n) .
5. Quelle est la meilleure approximation linéaire dans L^2 de X_{n+k} en fonction de X_0, \dots, X_n , $n \geq 0, k > 0$?

Exercice 34 Soit X une variable aléatoire à valeurs dans \mathbf{R}^d définie sur un espace de probabilité $(\Omega, \mathcal{A}, \mathbf{P})$. On suppose que X est bornée. Soit \mathcal{B} une sous tribu de \mathcal{A} et φ une fonction convexe de \mathbf{R}^d dans \mathbf{R} .

1. On suppose que φ est C^1 . Montrer que

$$\varphi(\mathbf{E}(X|\mathcal{B})) + \nabla \varphi(\mathbf{E}(X|\mathcal{B}))(X - \mathbf{E}(X|\mathcal{B})) \leq \varphi(X).$$

En déduire l'inégalité de Jensen,

$$\varphi(\mathbf{E}(X|\mathcal{B})) \leq \mathbf{E}(\varphi(X)|\mathcal{B}).$$

2. Retrouver l'inégalité de Jensen sans faire d'hypothèse de régularité sur la fonction φ mais en utilisant le fait que toute fonction convexe est l'enveloppe supérieure d'une famille dénombrable de fonctions affines (i.e. il existe une famille dénombrable $(A_n, n \geq 1)$ de fonctions affines telles que pour tout $x \in \mathbf{R}^d$, $\varphi(x) = \sup_{n \geq 1} A_n(x)$).

Exercice 35 Soit X et Y deux variables aléatoires réelles t.q. Y est $\sigma(X)$ -mesurable. On note $\mathcal{B}(\mathbf{R})$ la tribu borélienne de \mathbf{R} .

1. Montrer que $\mathcal{C} = \{X^{-1}(B), B \in \mathcal{B}(\mathbf{R})\}$ est une tribu. En déduire que $\mathcal{C} = \sigma(X)$.
2. Montrer que pour tout $a < b \in \mathbf{R}$, il existe $B_{a,b} \in \mathcal{B}(\mathbf{R})$ t.q. $\{a < Y \leq b\} = X^{-1}(B_{a,b}) = \{X \in B_{a,b}\}$.
3. Soit $Y_n = \sum_{p \in \mathbf{Z}} \frac{p}{2^n} \mathbf{1}_{\{\frac{p}{2^n} < Y \leq \frac{p+1}{2^n}\}}$. Montrer que $Y_n = f_n(X)$ pour

$$f_n(x) = \sum_{p \in \mathbf{Z}} \frac{p}{2^n} \mathbf{1}_{B_{\frac{p}{2^n}, \frac{p+1}{2^n}}}(x).$$

4. Montrer que $\forall n \geq n_0, \sup_{x \in \mathbf{R}} |f_n(x) - f_{n_0}(x)| \leq \frac{1}{2^{n_0}}$. En déduire que la fonction f_n converge uniformément vers une fonction f .
5. Quelle est la limite de $(Y_n)_n$? Conclure que $Y = f(X)$.

Il est facile de vérifier que les fonctions f_n sont mesurables au sens où $\forall B \in \mathcal{B}(\mathbf{R}), f_n^{-1}(B) \in \mathcal{B}(\mathbf{R})$. Cela implique que leur limite f est également mesurable.

Chapitre 4

Martingales à temps discrets

4.1 Introduction à la notion de martingale

Soit $(\Omega, \mathcal{A}, \mathbf{P})$ un espace muni d'une tribu et d'une probabilité. On se donne une suite croissante de sous-tribus de \mathcal{A} , $(\mathcal{F}_n, n \geq 0)$. On dit alors dans cette situation, que $(\mathcal{F}_n, n \geq 0)$ est une *filtration* sur $(\Omega, \mathcal{A}, \mathbf{P})$.

Soit $(X_n, n \geq 0)$ une suite de variables aléatoires. On dit que $(X_n, n \geq 0)$ est adaptée à la filtration $(\mathcal{F}_n, n \geq 0)$ si pour tout n , la variable aléatoire X_n est \mathcal{F}_n -mesurable.

Définition 4.1.1 Soit $(\Omega, \mathcal{A}, \mathbf{P})$ un espace et $(\mathcal{F}_n, n \geq 0)$ une filtration sur cet espace.

Une suite $(M_n, n \geq 0)$ de variables aléatoires réelles est une \mathcal{F}_n -martingale si :

- $(M_n, n \geq 0)$ est \mathcal{F}_n -adapté, et, pour tout $n \geq 0$, $\mathbf{E}(|M_n|) < +\infty$,
- pour tout $n \geq 0$:

$$\mathbf{E}(M_{n+1}|\mathcal{F}_n) = M_n. \quad (4.1)$$

Remarque 4.1.2 Si $(M_n, n \geq 0)$ est une \mathcal{F}_n -martingale, alors pour tout $n \geq 0$, M_n est \mathcal{F}_n mesurable et donc $\mathcal{F}_n^0 = \sigma(M_0, \dots, M_n) \subset \mathcal{F}_n$. On en déduit, grâce aux propriétés 6 et 3 de l'espérance conditionnelle données page 53, que :

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

On voit donc que $(M_n, n \geq 0)$ est, aussi, une \mathcal{F}_n^0 -martingale.

Remarque 4.1.3 Cette définition signifie que la meilleure prévision de M_{n+1} compte tenu de l'information disponible à l'instant n (i.e. \mathcal{F}_n) est donnée par M_n .

Remarque 4.1.4 Notez que (4.1) est équivalent à dire que pour tout $p \geq 0$:

$$\mathbf{E}(M_{n+p}|\mathcal{F}_n) = M_n.$$

En effet, en utilisant le fait que la suite des tribus \mathcal{F}_n est croissante on a :

$$\mathbf{E}(M_{n+2}|\mathcal{F}_n) = \mathbf{E}(\mathbf{E}(M_{n+2}|\mathcal{F}_{n+1})|\mathcal{F}_n) = \mathbf{E}(M_{n+1}|\mathcal{F}_n) = M_n.$$

On généralise ce résultat, sans difficultés, au cas $p > 2$ par récurrence.

Remarque 4.1.5 Si $(M_n, n \geq 0)$ est une martingale alors :

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

L'espérance d'une martingale est donc constante au cours du temps.

4.2 Exemples

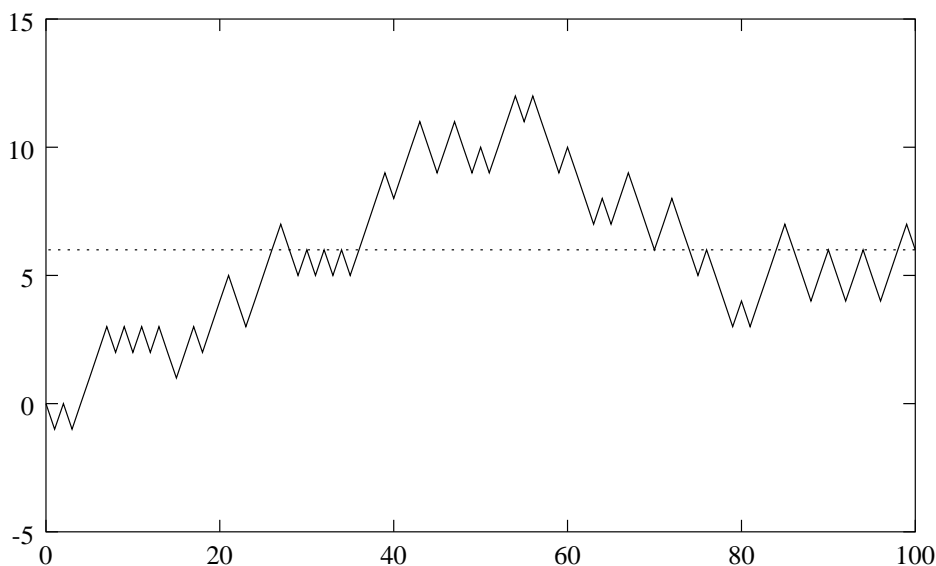


FIG. 4.1 – Trajectoire de la marche aléatoire symétrique

Soit $(X_n, n \geq 1)$ une suite de variables aléatoires indépendantes de même loi. On pose

$$\begin{cases} S_0 = 0 \\ S_n = X_1 + \dots + X_n \text{ si } n \geq 1 \end{cases}$$

On dit que $(S_n, n \geq 0)$ est une marche aléatoire. On note $\mathcal{F}_0 = \{\emptyset, \Omega\}$ et $\mathcal{F}_n = \sigma(X_1, \dots, X_n)$ pour $n \geq 1$. On peut remarquer que $\mathcal{F}_n = \sigma(S_0, \dots, S_n)$. On a alors :

Proposition 4.2.1 Si $\mathbf{E}(|X_1|) < +\infty$ et $\mathbf{E}(X_1) = 0$, $(S_n, n \geq 0)$ est une \mathcal{F}_n -martingale.

Démonstration : S_n est bornée puisque elle prend un nombre fini de valeurs, donc $\mathbf{E}(|S_n|) < +\infty$. De plus, on a :

$$\mathbf{E}(S_{n+1}|\mathcal{F}_n) = \mathbf{E}(S_n + X_{n+1}|\mathcal{F}_n) = \mathbf{E}(S_n|\mathcal{F}_n) + \mathbf{E}(X_{n+1}|\mathcal{F}_n).$$

Mais X_{n+1} est par construction indépendante de \mathcal{F}_n , donc $\mathbf{E}(X_{n+1}|\mathcal{F}_n) = \mathbf{E}(X_{n+1}) = 0$. De plus S_n est \mathcal{F}_n -mesurable (puisque c'est une fonction de (X_1, \dots, X_n)), donc $\mathbf{E}(S_n|\mathcal{F}_n) = S_n$. On obtient donc :

$$\mathbf{E}(S_{n+1}|\mathcal{F}_n) = S_n.$$

■

Proposition 4.2.2 Si $\mathbf{E}(X_1^2) < +\infty$ et $\mathbf{E}(X_1) = 0$, alors :

$$(V_n = S_n^2 - n\mathbf{E}(X_1^2), n \geq 0),$$

est une \mathcal{F}_n -martingale.

Démonstration : V_n prenant un nombre fini de valeurs, $\mathbf{E}(|V_n|) < +\infty$. Commençons par calculer $\mathbf{E}((S_{n+1} - S_n)^2 | \mathcal{F}_n)$, en utilisant le fait que S_n est \mathcal{F}_n mesurable on obtient :

$$\begin{aligned} \mathbf{E}((S_{n+1} - S_n)^2 | \mathcal{F}_n) &= \mathbf{E}(S_{n+1}^2 - 2S_n S_{n+1} + S_n^2 | \mathcal{F}_n) \\ &= \mathbf{E}(S_{n+1}^2 | \mathcal{F}_n) - 2S_n \mathbf{E}(S_{n+1} | \mathcal{F}_n) + S_n^2, \end{aligned}$$

mais comme $\mathbf{E}(S_{n+1} | \mathcal{F}_n) = S_n$ (S_n est une martingale), on obtient :

$$\mathbf{E}((S_{n+1} - S_n)^2 | \mathcal{F}_n) = \mathbf{E}(S_{n+1}^2 | \mathcal{F}_n) - S_n^2.$$

De plus :

$$\mathbf{E}((S_{n+1} - S_n)^2 | \mathcal{F}_n) = \mathbf{E}(X_{n+1}^2 | \mathcal{F}_n),$$

et X_{n+1} est indépendante de \mathcal{F}_n donc :

$$\mathbf{E}(S_{n+1}^2 | \mathcal{F}_n) - S_n^2 = \mathbf{E}((S_{n+1} - S_n)^2 | \mathcal{F}_n) = \mathbf{E}(X_{n+1}^2) = \mathbf{E}(X_1^2).$$

Soit encore :

$$\mathbf{E}(S_{n+1}^2 - (n+1)\mathbf{E}(X_1^2) | \mathcal{F}_n) = S_n^2 - n\mathbf{E}(X_1^2).$$

■

Proposition 4.2.3 Soit λ un réel tel que :

$$\phi(\lambda) = \log \mathbf{E}(e^{\lambda X_1}) < +\infty.$$

Alors :

$$(Z_n^\lambda = e^{\lambda S_n - n\phi(\lambda)}, n \geq 0),$$

est une \mathcal{F}_n -martingale.

Démonstration : Comme S_n est \mathcal{F}_n -mesurable, on a :

$$\mathbf{E}(e^{\lambda S_{n+1}} | \mathcal{F}_n) = \mathbf{E}(e^{\lambda S_n} e^{\lambda X_{n+1}} | \mathcal{F}_n) = e^{\lambda S_n} \mathbf{E}(e^{\lambda X_{n+1}} | \mathcal{F}_n).$$

Mais X_{n+1} est indépendant de \mathcal{F}_n donc :

$$\mathbf{E}(e^{\lambda X_{n+1}} | \mathcal{F}_n) = \mathbf{E}(e^{\lambda X_{n+1}}) = \mathbf{E}(e^{\lambda X_1}) = e^{\phi(\lambda)}.$$

D'où :

$$\mathbf{E}(e^{\lambda S_{n+1}} | \mathcal{F}_n) = e^{\lambda S_n} e^{\phi(\lambda)},$$

puis le résultat.

■

4.3 Un exemple générique de martingale

Proposition 4.3.1 Soit $(M_n, n \geq 0)$ une \mathcal{F}_n -martingale. Soit $(H_n, n \geq 0)$ un processus tel que H_n est \mathcal{F}_n -mesurable et borné. On se donne une variable aléatoire G_0 intégrable et \mathcal{F}_0 -mesurable. Et pour $n \geq 1$, on pose :

$$G_n = G_0 + \sum_{k=0}^{n-1} H_k (M_{k+1} - M_k),$$

ou, de façon équivalente, pour $n \geq 0$,

$$G_{n+1} - G_n = H_n (M_{n+1} - M_n).$$

Alors $(G_n, n \geq 0)$ est une \mathcal{F}_n -martingale.

Démonstration : Il suffit de remarquer que G_n est \mathcal{F}_n -mesurable que $\mathbf{E}(|G_n|) < +\infty$ et que :

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

■

Exemple : gain d'une stratégie dans un jeu de hasard équilibré On note $(H_n, n \geq 0)$ une suite de quantité que l'on va parier sur un tirage qui rapporte X_{n+1} . On suppose que les H_n s'écrivent sous la forme $H_n = f(n, X_1, \dots, X_n)$ (en particulier H_0 est déterministe) et que les H_n sont des variables aléatoires bornées. Si l'on note G_n le gain cumulé à l'instant n , on a :

$$G_{n+1} - G_n = H_n X_{n+1} = H_n (S_{n+1} - S_n).$$

On voit donc que si la suite des $(X_n, n \geq 1)$ est une suite I.I.D. de variables aléatoires intégrables avec $\mathbf{E}(X_1) = 0$, $(G_n, n \geq 0)$ est une martingale, d'après le résultat précédent appliqué avec $M_n = S_n = X_1 + \dots + X_n$. L'espérance d'une martingale étant constante on en déduit que :

Le gain moyen au bout d'un nombre fini de tirages dans un jeu équilibré est nul.

$$\mathbf{E}(G_n) = \mathbf{E}(G_0).$$

Il est, en particulier, exclu de gagner en moyenne dans un jeu de pile ou face équilibré.

Il est aussi impossible (encore plus !) de gagner à coup sûr sans prendre de risque (i.e. $\mathbf{P}(G_n \geq G_0) = 1$ et $\mathbf{P}(G_n > G_0) > 0$) dans les mêmes conditions.

Il existe cependant des temps aléatoires τ "raisonnables" finis tels que $\mathbf{P}(S_\tau > 0) = 1$. Nous verrons, cependant, qu'il faut accepter de ne pas borner ses pertes pour arriver à un tel résultat.

4.4 Notion de temps d'arrêt

Définition 4.4.1 Soit $(\mathcal{F}_n, n \geq 0)$ une filtration (i.e. une suite croissante de tribus). On dit que τ , une variable aléatoire prenant ses valeurs dans $\mathbf{N} \cup \{+\infty\}$, est un \mathcal{F}_n -temps d'arrêt si, pour tout $n \in \mathbf{N}$:

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

Remarque 4.4.2 Cela signifie que à l'instant n , on *sait* si $\tau \leq n$ ou pas.

Soit $(X_n, n \geq 0)$ un processus adapté à la filtration $(\mathcal{F}_n, n \geq 0)$ (i.e. pour tout n , X_n est \mathcal{F}_n mesurable), posons :

$$\tau = \inf\{n \geq 0, X_n \in [a, b]\},$$

alors τ est un \mathcal{F}_n -temps d'arrêt. On dit que τ est le temps d'entrée du processus $(X_n, n \geq 0)$ dans l'intervalle $[a, b]$.

Par contre τ , le temps réalisant le maximum d'un processus :

$$\tau = \text{ArgMax}\{\phi(X_n), n \geq 0\},$$

et τ' , le dernier temps où un processus est dans un ensemble :

$$\tau' = \sup\{n \geq 0, X_n \in [a, b]\},$$

ne sont généralement *pas des temps d'arrêt*.

Nous allons voir que la notion de martingale se marie bien avec celle de temps d'arrêt.

Définition 4.4.3 On note $x \wedge y = \inf\{x, y\}$. Soit $(M_n, n \geq 0)$ un processus adapté à la filtration $(\mathcal{F}_n, n \geq 0)$ et τ un temps d'arrêt par rapport à la même filtration. On définit alors un nouveau processus par :

$$N_n = M_{\tau \wedge n},$$

Ceci signifie que $N_n(\omega) = M_n(\omega)$ si $\tau(\omega) \geq n$ et $N_n = M_\tau$ si $\tau \leq n$ ¹. On dit que N est le processus M *arrêté* au temps τ .

Proposition 4.4.4 Si $(M_n, n \geq 0)$ est une martingale par rapport à $(\mathcal{F}_n, n \geq 0)$ et si τ est un temps d'arrêt par rapport à la même filtration, alors $(N_n, n \geq 0)$ est encore une \mathcal{F}_n -martingale.

Démonstration : On a :

$$N_{n+1} - N_n = \mathbf{1}_{\{\tau > n\}} (M_{n+1} - M_n).$$

De plus $H_n = \mathbf{1}_{\{\tau > n\}} = 1 - \mathbf{1}_{\{\tau \leq n\}}$ est \mathcal{F}_n -mesurable. $(N_n, n \geq 0)$ apparaît donc comme une transformée de la martingale M au sens de la proposition 4.3.1. Ce qui permet de conclure. ■

On obtient alors, sans difficultés, le résultat important suivant connu sous le nom de théorème d'arrêt.

Proposition 4.4.5 Si τ est un temps d'arrêt borné ($\tau \leq C$, presque sûrement avec C réel positif). Cette propriété est essentielle par rapport à une filtration $(\mathcal{F}_n, n \geq 0)$ et si $(M_n, n \geq 0)$ est une martingale par rapport à la même filtration alors :

$$\mathbf{E}(M_\tau) = \mathbf{E}(M_0).$$

¹On notera que le fait que τ puisse être infini ne pose pas de problèmes dans cette définition.

Démonstration : Il suffit de remarquer qu'il existe un entier K tel que, presque sûrement, $\tau \leq K$ et donc :

$$N_K = M_{\tau \wedge K} = M_\tau.$$

Mais comme $(N_n, n \geq 1)$ est une martingale on a :

$$\mathbf{E}(N_K) = \mathbf{E}(N_0),$$

ce qui se réécrit $\mathbf{E}(M_\tau) = \mathbf{E}(M_0)$. ■

On voit donc que l'égalité $\mathbf{E}(M_n) = \mathbf{E}(M_0)$ se généralise à certains temps aléatoires (les temps d'arrêts bornés). Les deux hypothèses :

1. τ temps d'arrêt,
2. τ borné,

sont cruciales. Évidemment, si $\tau = \text{ArgMax}\{M_n, n \leq N\}$, $\mathbf{E}(M_\tau) > \mathbf{E}(M_0)$ sauf dans des cas triviaux. On voit donc que **1** doit figurer dans les hypothèse du théorème. Pour justifier **2** nous allons prouver le résultat suivant.

Proposition 4.4.6 Soit $(X_n, n \geq 1)$ une suite de variables aléatoires indépendantes valant ± 1 avec probabilité $1/2$. On pose $S_0 = 0$ et pour $n \geq 1$:

$$S_n = X_1 + \dots + X_n.$$

Pour un entier strictement positif non nul a , on introduit :

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

(avec la convention que $\tau = +\infty$, si l'ensemble précédent est vide).

Alors τ est un temps d'arrêt vérifiant $\mathbf{P}(\tau < +\infty) = 1$ (on dit que τ est fini presque sûrement) et $\mathbf{E}(S_\tau) = a > 0$.

On voit donc que l'on ne peut pas appliquer le théorème d'arrêt à la martingale $(S_n, n \geq 1)$ et au temps d'arrêt τ .

Démonstration : Remarquons que :

$$^c\{\tau \leq n\} = \{S_1 < a, \dots, S_n < a\} \in \mathcal{F}_n.$$

$\{\tau \leq n\}$ est donc dans \mathcal{F}_n et τ est un temps d'arrêt.

Montrer que $\mathbf{P}(\tau < +\infty) = 1$ est plus délicat. Nous allons utiliser le théorème d'arrêt pour la martingale :

$$(e^{\lambda S_n - n\phi(\lambda)}, n \geq 0),$$

avec :

$$\phi(\lambda) = \log \mathbf{E}(e^{\lambda X_1}).$$

Lorsque $\lambda \neq 0$, on a $\mathbf{E}(e^{\lambda X_1}) = (e^\lambda + e^{-\lambda})/2 = \cosh(\lambda) > 1$ et donc $\phi(\lambda) > 0$.

τ n'est pas un temps d'arrêt borné mais $\tau \wedge N$, pour N un entier fixé, est borné et reste un temps d'arrêt. En effet, $\{\tau \wedge N \leq n\}$ est égal à $\{\tau \leq n\}$ si $N > n$ et à Ω si $N \leq n$ et il s'agit dans les deux cas d'ensembles de \mathcal{F}_n .

En utilisant le théorème d'arrêt on obtient :

$$\mathbf{E}(e^{\lambda S_{\tau \wedge N} - (\tau \wedge N)\phi(\lambda)}) = \mathbf{E}(e^0) = 1.$$

Nous allons maintenant chercher à faire tendre N vers $+\infty$. Que se passe t'il ?

- Nous avons vu que $\phi(\lambda) \geq 0$ et si l'on suppose de plus que $\lambda > 0$, comme $S_{\tau \wedge N} \leq a$, on voit que l'on a :

$$e^{\lambda S_{\tau \wedge N} - (\tau \wedge N)\phi(\lambda)} \leq e^{\lambda S_{\tau \wedge N}} \leq e^{\lambda a}.$$

- De plus lorsque N tends vers $+\infty$, on a $\tau \wedge N$ qui tends vers τ (c'est clair si l'on sépare les cas $\tau < +\infty$ et $\tau = +\infty$). On a donc, presque sûrement :

$$\lim_{N \rightarrow +\infty} e^{\lambda S_{\tau \wedge N} - (\tau \wedge N)\phi(\lambda)} = \mathbf{1}_{\{\tau < +\infty\}} e^{\lambda S_{\tau} - \tau\phi(\lambda)} + \mathbf{1}_{\{\tau = +\infty\}} \times 0.$$

Le théorème de Lebesgue (voir 1.3.2 page 11) s'applique et prouve que :

$$\mathbf{E} \left(\mathbf{1}_{\{\tau < +\infty\}} e^{\lambda S_{\tau} - \tau\phi(\lambda)} \right) = 1.$$

Mais comme, sur l'événement $\{\tau < +\infty\}$, $S_{\tau} = a$, on en déduit que, pour tout $\lambda > 0$:

$$\mathbf{E} \left(\mathbf{1}_{\{\tau < +\infty\}} e^{-\tau\phi(\lambda)} \right) = e^{-\lambda a}.$$

Il ne reste plus qu'à faire tendre λ vers 0 en décroissant par valeurs positives et à utiliser le théorème de convergence monotone (théorème 1.3.1 page 10) pour, en déduire, que :

$$\mathbf{P}(\tau < +\infty) = 1.$$

■

Remarque 4.4.7 On déduit de ce qui précède que, pour tout $\lambda > 0$:

$$\mathbf{E} \left(e^{-\tau\phi(\lambda)} \right) = e^{-\lambda a}.$$

On peut montrer que cette égalité suffit à identifier la loi de τ .

Remarque 4.4.8 On a vu que, dans le jeu de pile ou face équilibré, il n'existe pas de stratégie en nombre fini de coups permettant de gagner en moyenne. La proposition précédente permet d'affirmer que si on joue suffisamment longtemps 1F à pile ou face on finit par gagner à coup sûr 1F. Quelle est la limitation de cette stratégie ?

Comme $\tau \wedge n$ est un temps d'arrêt borné et $(S_n, n \geq 0)$ une martingale on a :

$$\mathbf{E}(S_{n \wedge \tau}) = \mathbf{E}(S_0) = 0.$$

Sous l'hypothèse $\mathbf{E}(\sup_{n \leq \tau} |S_n|) < +\infty$, on peut utiliser le théorème de Lebesgue et en déduire que $\mathbf{E}(S_{\tau}) = 0$. Or ceci est clairement impossible, puisque par construction $\mathbf{E}(S_{\tau}) = a > 0$. Donc $\mathbf{E}(\sup_{n \leq \tau} |S_n|) = +\infty$. Mais, comme $S_n \leq a$ pour $n \leq \tau$, on peut aussi affirmer que $\mathbf{E}(\sup_{n \leq \tau} (S_n)^-) = +\infty$ (x^- désignant la partie négative de x).

On voit que dans cette stratégie, avant de gagner 1F, on doit être prêt à assumer une perte dont l'espérance est infinie. En particulier on doit être prêt à assumer une perte non bornée.

4.5 Théorème de convergence des martingales

On cherche à voir sous quelles conditions une suite $(M_n, n \geq 1)$ converge lorsque n tends vers $+\infty$.

Théorème 4.5.1 *Supposons que $(M_n, n \geq 0)$ soit une martingale par rapport à $(\mathcal{F}_n, n \geq 0)$ et que de plus :*

$$\sup_{n \geq 0} \mathbf{E} (M_n^2) < +\infty.$$

On dit que $(M_n, n \geq 0)$ est bornée en moyenne quadratique (ou dans L^2).

Sous cette hypothèse il existe une variable aléatoire M_∞ de L^2 telle que :

$$\lim_{n \rightarrow +\infty} M_n = M_\infty,$$

au sens de la convergence en moyenne quadratique et presque sûre. Cela signifie que :

- convergence en moyenne quadratique : $\lim_{n \rightarrow +\infty} \mathbf{E} (|M_n - M_\infty|^2) = 0.$*
- convergence presque sûre : $\mathbf{P} (\lim_{n \rightarrow +\infty} M_n = M_\infty) = 1.$*

Démonstration : Rappel : Nous avons déjà vu que :

$$L^2 = \left\{ X \text{ v.a. telle que } \|X\|_{L^2}^2 = \mathbf{E}(X^2) < +\infty \right\},$$

est un espace de Hilbert. En particulier, il est complet. Ceci signifie que si :

$$\|X_{n+p} - X_n\|_{L^2},$$

peut être rendu petit pour n grand indépendamment de $p \geq 0$ (on dit alors que $(X_n, n \geq 1)$ est une suite de Cauchy), alors il existe X_∞ variable aléatoire de L^2 telle que :

$$\lim_{n \rightarrow +\infty} X_n = X_\infty \text{ dans } L^2.$$

Nous allons commencer par montrer que la suite $(M_n, n \geq 1)$ est une suite de Cauchy. Remarquons que :

$$\begin{aligned} \mathbf{E} \left((M_{n+p} - M_n)^2 \right) &= \mathbf{E} (M_{n+p}^2 - 2M_n M_{n+p} + M_n^2) \\ &= \mathbf{E} (M_{n+p}^2) + \mathbf{E} (M_n^2) - 2\mathbf{E} (\mathbf{E} (M_n M_{n+p} | \mathcal{F}_n)), \end{aligned}$$

mais :

$$\mathbf{E} (\mathbf{E} (M_n M_{n+p} | \mathcal{F}_n)) = \mathbf{E} (M_n \mathbf{E} (M_{n+p} | \mathcal{F}_n)) = \mathbf{E} (M_n^2).$$

Donc :

$$\mathbf{E} \left((M_{n+p} - M_n)^2 \right) = \mathbf{E} (M_{n+p}^2) - \mathbf{E} (M_n^2). \quad (4.2)$$

On déduit de cette dernière égalité que :

1. $\mathbf{E} (M_{n+p}^2) - \mathbf{E} (M_n^2) \geq 0$, donc $\mathbf{E} (M_n^2)$ est croissante. Comme elle est bornée, par hypothèses, $\mathbf{E} (M_n^2)$ converge vers une valeur α .
2. on a :

$$\sup_{p \geq 0} \mathbf{E} \left((M_{n+p} - M_n)^2 \right) = \sup_{p \geq 0} \mathbf{E} (M_{n+p}^2) - \mathbf{E} (M_n^2).$$

La suite $\mathbf{E} (M_n^2)$ étant convergente on en déduit que $(M_n, n \geq 1)$ est une suite de Cauchy dans L^2 . Elle converge donc vers une variable aléatoire M_∞ dans L^2 .

La convergence presque sûre est plus délicate à prouver et nous l'admettrons. ■

Exemple Soit $(U_n, n \geq 1)$ une suite de variables aléatoires indépendantes suivant une loi uniforme sur $[0, 1]$. Soit f une fonction bornée telle que $\mathbf{E}(f(U_1)) = 0$ et $(\gamma_n, n \geq 1)$ une suite de réels telle que $\sum_{n \geq 1} \gamma_n^2 < +\infty$ alors la série suivante :

$$\sum_{n \geq 1} \gamma_n f(U_n),$$

existe au sens de la convergence dans L^2 et presque sûre.

4.6 Convergence de Martingales et algorithmes stochastiques

Le but de cette section est de montrer, dans un cas simple, comment l'on peut utiliser le résultat de convergence des martingales cité plus haut pour construire des procédures d'approximation stochastique.

Cadre d'utilisation Nous allons supposer que l'on cherche à trouver un zéro d'une fonction $f(x)$ de \mathbf{R}^n dans \mathbf{R} . On supposera de plus que la fonction f est difficile à calculer, mais qu'elle se représente sous la forme :

$$f(x) = \mathbf{E}(F(x, U)),$$

F étant une fonction aisément calculable et U une variable aléatoire de loi donnée, facile à simuler.

Un exemple typique où ce genre d'hypothèse est pertinente est lorsque l'on cherche à évaluer un quantile d'une loi connue. Plus précisément, on note $f(x)$ la fonction de répartition d'une loi U donnée par :

$$f(x) = \mathbf{P}(U \leq x),$$

et, pour une probabilité α proche de 1, on cherche à identifier x_α tel que :

$$\mathbf{P}(U \leq x_\alpha) = \alpha.$$

On dit, dans ce cas, que x_α est le quantile d'ordre α de U . On voit que dans ce cas, on cherche à trouver un zéro de $f(x) = \mathbf{P}(U \leq x) - \alpha$.

Pour fixer les idées supposons que U suive une loi gaussienne. La fonction de répartition de U n'est pas connue explicitement, mais on sait simuler facilement une variable aléatoire selon la loi de U . De plus

$$\mathbf{P}(U \leq x) = \mathbf{E}\left(\mathbf{1}_{\{U \leq x\}}\right),$$

et l'on peut donc choisir $F(x, u) = \mathbf{1}_{\{u \leq x\}}$.

L'algorithme de Robbins et Monro Lorsque la fonction f est croissante, nous allons voir que l'on peut construire un algorithme n'utilisant que la fonction F convergent vers l'unique zéros de f . Cet algorithme porte le nom d'algorithme de Robbins et Monro et se présente de la façon suivante. On choisit une valeur initiale X_0 quelconque, puis l'on définit itérativement X_n en posant

$$X_{n+1} = X_n - \gamma_n F(X_n, U_{n+1}).$$

Nous allons prouver que sous certaines hypothèses sur F et γ cet algorithme est convergent.

Les hypothèses que nous ferons sur les fonctions F et f seront les suivantes :

- F est uniformément bornée en x, u .
- il existe $x^* \in \mathbf{R}^d$ tel que $f(x) \cdot (x - x^*) \geq |x - x^*|^2$.

Remarque 4.6.1 Notez que la première hypothèse implique que f est uniformément bornée en x . Lorsque f est supposée de plus continue, la deuxième hypothèse implique que x^* est l'unique zéro de f . Pour se convaincre de cette deuxième propriété, considérons un vecteur arbitraire de \mathbf{R}^n et ϵ un réel positif. On a alors :

$$f(x^* + \epsilon v) \cdot \epsilon v \geq \epsilon^2 |v|^2.$$

On en déduit que $f(x^* + \epsilon v) \cdot v \geq \epsilon |v|^2$ et en passant à la limite lorsque ϵ tend vers 0 que $f(x^*) \cdot v \geq 0$. Cette relation étant vérifiée pour tout vecteur v , on en déduit en changeant le signe de v que $f(x^*) \cdot v = 0$. Cette dernière égalité étant vérifiée pour tout v , implique que $f(x^*) = 0$.

L'unicité du zéro de f est claire puisque, si $x \neq x^*$ on a

$$f(x) \cdot (x - x^*) \geq c |x - x^*|^2 > 0.$$

Ce qui interdit à $f(x)$ d'être nul.

Notez que en dimension 1, la deuxième condition signifie que $f(x) \geq c(x - x^*)$ lorsque $x > x^*$ et que $f(x) \leq c(x - x^*)$ lorsque $x < x^*$.

Nous allons maintenant énoncer un résultat de convergence pour l'algorithme suggéré au début de ce paragraphe.

Théorème 4.6.1 Soit f une fonction de \mathbf{R}^n dans \mathbf{R} qui peut se mettre sous la forme

$$f(x) = \mathbf{E}(F(x, U)),$$

U étant une variable aléatoire de loi connue prenant ses valeurs dans \mathbf{R}^p et F une fonction de $\mathbf{R}^n \times \mathbf{R}^p$ dans \mathbf{R}^n .

On suppose que :

- F est une fonction bornée ($|F(x, u)| \leq K$ pour tout x et u),
- $f(x) \cdot (x - x^*) \geq c|x - x^*|^2$, c étant une constante strictement positive,
- $(\gamma_n, n \geq 0)$ est une suite de réels positif qui vérifient $\sum_{n \geq 0} \gamma_n = +\infty$ et $\sum_{n \geq 0} \gamma_n^2 < +\infty$.

On définit une suite de variables aléatoires $(X_n, n \geq 0)$ en posant pour $n \geq 0$:

$$X_{n+1} = X_n - \gamma_n F(X_n, U_{n+1}), X_0 = x,$$

x étant une valeur déterministe arbitraire et $(U_n, n \geq 1)$ une suite de variables aléatoires indépendantes tirées selon la loi de U .

Alors, $\lim_{n \rightarrow +\infty} X_n = x^*$ presque sûrement et dans L^2 .

Avant d'aborder la démonstration de ce théorème nous rappelons un résultat classique sur les suites de nombres réels connue sous le nom de lemme de Kronecker.

Lemme 4.6.2 Soit $(A_n, n \geq 1)$ une suite de nombres réels qui croissent vers $+\infty$.

Soit $S_n = \sum_{k=1}^n \epsilon_k$ la somme partielle d'une série que l'on suppose (simplement) convergente.

Alors

$$\lim_{n \rightarrow +\infty} \frac{1}{A_n} \sum_{k=1}^n A_k \epsilon_k = 0.$$

Démonstration : Par convention on pose $A_0 = S_0 = 0$. La série S_n étant convergente, on peut poser $S_n = S_\infty + \eta_n$ avec $\lim_{n \rightarrow +\infty} \eta_n = 0$. Avec ces notations, on a

$$\sum_{k=1}^n A_k \epsilon_k = \sum_{k=1}^n A_k (S_k - S_{k-1}) = \sum_{k=1}^n A_k (\eta_k - \eta_{k-1}) = A_n \eta_n - \sum_{k=1}^n (A_k - A_{k-1}) \eta_{k-1}.$$

D'où :

$$\frac{1}{A_n} \sum_{k=1}^n A_k \epsilon_k = \eta_n - \frac{1}{A_n} \sum_{k=1}^n (A_k - A_{k-1}) \eta_{k-1}.$$

Le premier terme tend clairement vers 0. Pour le deuxième, on se donne un $\epsilon > 0$, et l'on choisit N tel que pour $k \geq N$, $|\eta_k| \leq \epsilon$. On a alors :

$$\left| \frac{1}{A_n} \sum_{k=1}^n (A_k - A_{k-1}) \eta_{k-1} \right| \leq \frac{1}{A_n} \sum_{k=1}^N (A_k - A_{k-1}) |\eta_{k-1}| + \frac{\epsilon}{A_n} \sum_{k=N+1}^n (A_k - A_{k-1}).$$

On a donc, pour $n \geq N$

$$\left| \frac{1}{A_n} \sum_{k=1}^n (A_k - A_{k-1}) \eta_{k-1} \right| \leq \frac{C_N}{A_n} + \epsilon.$$

Ce qui permet de conclure en remarquant que $\lim_{n \rightarrow +\infty} A_n = +\infty$. ■

Démonstration : On peut se ramener sans perte de généralité au cas $x^* = 0$ par une translation. Nous ne traiterons donc que ce cas. D'autre part, vu les hypothèses sur γ_n , il est clair que γ_n tend vers 0, lorsque n tend vers $+\infty$. Nous supposons donc, dans la suite, toujours sans perte de généralité que, pour tout $n \geq 0$, $\gamma_n \leq 1/(2c)$ (il suffit d'étudier l'algorithme à partir d'un rang pour lequel l'inégalité est vérifiée).

Nous allons commencer par montrer que X_n tend vers 0 dans L^2 , c'est à dire que $\lim_{n \rightarrow +\infty} \mathbf{E}(|X_n|^2) = 0$. Pour cela notons $\tilde{F}(x, u) = F(x, u) - f(x)$. On a alors $\mathbf{E}(\tilde{F}(x, U)) = 0$. De plus :

$$\begin{aligned} |X_{n+1}|^2 &= |X_n - \gamma_n F(X_n, U_{n+1})|^2 \\ &= |X_n|^2 + \gamma_n^2 |F(X_n, U_{n+1})|^2 - 2\gamma_n X_n \cdot f(X_n) - 2\gamma_n X_n \cdot \tilde{F}(X_n, U_{n+1}) \\ &\leq |X_n|^2 + \gamma_n^2 K^2 - 2c\gamma_n |X_n|^2 - 2\gamma_n X_n \cdot \tilde{F}(X_n, U_{n+1}). \end{aligned} \quad (4.3)$$

K étant un majorant de F . On prend alors l'espérance des 2 membres de la précédente inégalité pour obtenir :

$$\mathbf{E}(|X_{n+1}|^2) \leq (1 - 2c\gamma_n) \mathbf{E}(|X_n|^2) + \gamma_n^2 K^2 - 2\gamma_n \mathbf{E}(X_n \cdot \tilde{F}(X_n, U_{n+1})).$$

Mais X_n est $\mathcal{F}_n = \sigma(U_1, \dots, U_n)$ mesurable et U_{n+1} est indépendant de \mathcal{F}_n par construction de la suite $(U_n, n \geq 1)$, on a donc (voir le lemme 3.2.6) :

$$\mathbf{E}(X_n \cdot \tilde{F}(X_n, U_{n+1}) | \mathcal{F}_n) = \psi(X_n),$$

avec $\psi(x) = \mathbf{E}(x \cdot \tilde{F}(x, U_{n+1})) = x \cdot \mathbf{E}(\tilde{F}(x, U_{n+1})) = 0$. On a donc :

$$\mathbf{E}(X_n \cdot \tilde{F}(X_n, U_{n+1})) = \mathbf{E}(\mathbf{E}(X_n \cdot \tilde{F}(X_n, U_{n+1}) | \mathcal{F}_n)) = 0. \quad (4.4)$$

En notant $\alpha_n = \mathbf{E}(|X_n|^2)$, on vient d'obtenir

$$\alpha_{n+1} \leq \alpha_n(1 - 2c\gamma_n) + K^2\gamma_n^2.$$

En posant, $A_1 = 1$ et :

$$A_n = \prod_{i=0}^{n-1} \frac{1}{1 - 2c\gamma_i},$$

on obtient :

$$A_{n+1}\alpha_{n+1} \leq A_n\alpha_n + K^2\gamma_n^2 A_{n+1}.$$

On en déduit en sommant entre 0 et $n - 1$ que :

$$\alpha_n \leq \frac{\mathbf{E}(|X_0|^2)}{A_n} + \frac{1}{A_n} \sum_{k=1}^n K^2\gamma_{k-1}^2 A_k.$$

Pour montrer que α_n tends vers 0, on applique alors le lemme de Kronecker. D'une part $\sum_{k=1}^n \gamma_{k-1}^2 < +\infty$. D'autre part, on note que $\gamma_n \leq 1/(2c)$ (à partir d'un certain rang) ce qui implique que A_n est croissante et l'on a $\log(1+x) \leq x$ pour tout x réel, d'où :

$$A_n = e^{-\sum_{i=1}^n \log(1-2c\gamma_i)} \geq e^{2c \sum_{i=1}^n \gamma_i}.$$

Comme, $\sum_{n \geq 1} \gamma_n = +\infty$, on obtient $\lim_{n \rightarrow +\infty} A_n = +\infty$, et l'on est en mesure d'appliquer le lemme de Kronecker pour prouver que $\lim_{n \rightarrow +\infty} \mathbf{E}(|X_n|^2) = 0$.

Pour montrer la convergence presque sûre, nous allons utiliser le théorème 4.5.1 qui permet d'affirmer qu'une martingale bornée dans L^2 est convergente. Pour cela on commence par réécrire l'inéquation (4.3) sous la forme :

$$|X_{n+1}|^2 \leq |X_n|^2 + \gamma_n^2 K^2 - 2c\gamma_n |X_n|^2 + M_{n+1} - M_n, \quad (4.5)$$

où l'on a défini la suite $(M_n, n \geq 0)$ par $M_0 = 0$ et

$$M_{n+1} - M_n = -2\gamma_n X_n \tilde{F}(X_n, \mathbf{U}_{n+1}).$$

Il est facile de vérifier que $(M_n, n \geq 0)$ est une martingale puisque

$$\mathbf{E}(M_{n+1} - M_n | \mathcal{F}_n) = -2\gamma_n \mathbf{E}(X_n \tilde{F}(X_n, \mathbf{U}_{n+1})) = 0,$$

vue l'égalité (4.4). D'autre part, en sommant l'inégalité (4.5) entre 0 et $n - 1$, on obtient :

$$|X_n|^2 \leq \frac{|X_0|^2}{A_n} + \frac{1}{A_n} \sum_{k=1}^n A_k (K^2\gamma_{k-1}^2 + M_k - M_{k-1}).$$

Si l'on est en mesure de prouver que $(M_n, n \geq 0)$ est une suite convergente, la série de terme général $K^2\gamma_{k-1}^2 + M_k - M_{k-1}$ sera elle aussi convergente et le lemme de Kronecker permettra de conclure. Pour terminer la démonstration, il nous reste donc à démontrer que M_n converge lorsque n tends vers l'infini. Pour montrer ceci, nous utilisons le théorème 4.5.1. On sait déjà que $(M_n, n \geq 0)$ est une martingale, il suffit donc de vérifier que

$$\sup_{n \geq 0} \mathbf{E}(M_n^2) < +\infty.$$

En utilisant l'équation (4.2), on obtient :

$$\begin{aligned} \mathbf{E}(M_{n+1}^2) - \mathbf{E}(M_n^2) &= \mathbf{E}\left((M_{n+1} - M_n)^2\right) \\ &= 4\gamma_n^2 \mathbf{E}\left(|X_n \cdot \tilde{F}(X_n, U_{n+1})|^2\right) \\ &\leq 4\gamma_n^2 \sqrt{\mathbf{E}(|X_n|^2)} \sqrt{\mathbf{E}(\tilde{F}(X_n, U_{n+1})^2)}. \end{aligned}$$

Mais F est bornée donc f et \tilde{F} aussi et nous venons de voir que $\mathbf{E}(|X_n|^2)$ tend vers 0 (et reste donc bornée). On a donc :

$$\mathbf{E}(M_{n+1}^2) - \mathbf{E}(M_n^2) \leq C\gamma_n^2.$$

En sommant cette inégalité entre 0 et $n - 1$, on obtient :

$$\mathbf{E}(M_n^2) - \mathbf{E}(M_0^2) \leq \sum_{k=0}^{n-1} \gamma_k^2.$$

La série $\sum_{k \geq 0} \gamma_k^2$ étant convergente, on en déduit que $\mathbf{E}(M_n^2)$ reste borné. Le théorème 4.5.1 permet alors de conclure à la convergence presque sûre de M_n et le lemme de Kronecker conduit à la convergence presque sûre de X_n vers 0. ■

Remarque 4.6.3 L'efficacité de l'algorithme est liée au choix de la suite $(\gamma_n, n \geq 0)$. Ce choix est souvent délicat. On s'en convaincra en utilisant pour des diverses valeurs de b et c , le programme suivant.

Un programme Scilab pour le calcul d'un fractile

```
// fractile a calculer
alpha=0.2; // 20 %

// tirages aleatoires communs aux trois algorithmes
n=1000;
u=rand(1,n);

// algorithme de Robbins et Monro
// choix de gamma_n
c=1.0;b=1.0;
Gamma=c*1./(1+b*[1:n]);
// algorithme
y=1/2;
x_rm=zeros(1,n);
for i=1:n
    y=y-Gamma(i)*((u(i)<=y)-alpha);
    x_rm(i)=y;
end

// moyenne des tirages Robbins Monro
z=cumsum(x_rm)./ [1:n];

// methode empirique classique
w=zeros(1,n);
ww=[];
for i=1:n
```

```
ww=[ww,u(i)];
xx=sort(ww);
h=xx(ceil(i*(1-alpha)));
w(i)=h;
end;

xbasc();
// algorithme de Robbins et Monro
plot2d([1:n],x_rm,1,'011','',[0,-.2,n,1],[0,4,0,6]);
// moyenne des tirages Robbins Monro
plot2d([1:n],z,2,'010','',[0,-.2,n,1],[0,4,0,3]);
// methode empirique
plot2d([1:n],w,6,'010','',[0,-.2,n,1],[0,4,0,3]);
// resultat exact
plot2d([1:n],alpha*ones(1:n),3,'010','',[0,-.2,n,1],[0,4,0,3]);
```

4.7 Travail dirigé : Temps de sortie d'une marche aléatoire

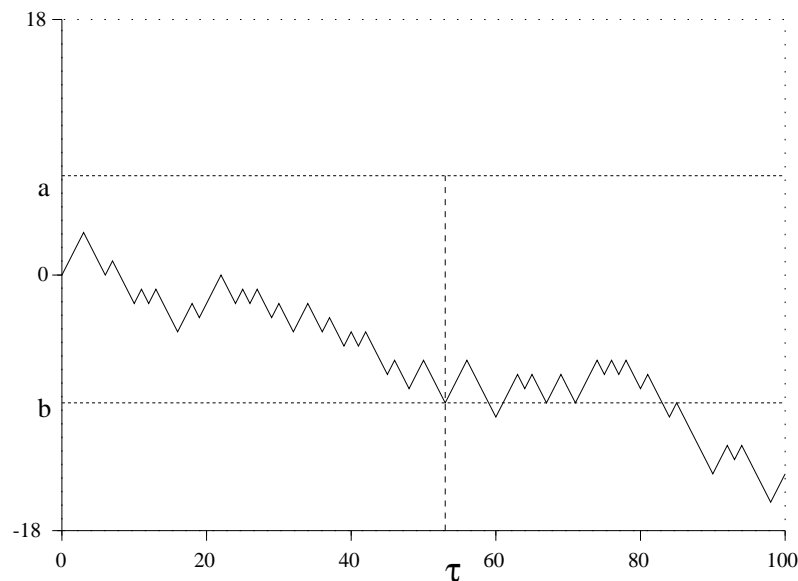


FIG. 4.2 – Temps de sortie de $[a, b]$ par une marche aléatoire

Soient $(X_i, i \geq 1)$ des variables aléatoires indépendantes de même loi, telles que $\mathbf{P}[X_i = 1] = p$ et $\mathbf{P}[X_i = -1] = q = 1 - p$. On suppose $p \in]0, 1[$. On note $S_n = S_0 + X_1 + \dots + X_n$. La suite S_n modélise la richesse d'un joueur à un jeu de pile ou face avec une pièce biaisée. $S_0 = x \in \mathbf{Z}$ est la richesse initiale du joueur. Dans les calculs on notera \mathbf{P}_x pour \mathbf{P} afin de rappeler que l'on part de la condition initiale $S_0 = x$.

Le cas $p = q = 1/2$.

1. Montrer que $(S_n, n \geq 0)$ est une \mathcal{F}_n -martingale où $\mathcal{F}_n = \sigma(X_1, \dots, X_n)$.
2. Pour $a < b \in \mathbf{Z}$, on note

$$\tau_{a,b} = \inf\{n \geq 0; S_n \geq b \text{ ou } S_n \leq a\}.$$

Que représente $\tau_{a,b}$? Montrer qu'il s'agit d'un temps d'arrêt.

3. Montrer en utilisant le théorème de la limite centrale que $\tau_{a,b}$ est fini p.s. Montrer qu'il n'est pas borné si $a < x < b - 1$ ou si $a + 1 < x < b$.
4. En utilisant le théorème d'arrêt, montrer que, pour $x \in [a, b]$:

$$\mathbf{E}_x[S_{\tau_{a,b}}] = x.$$

5. On suppose $x \in]a, b[$. Calculer $\mathbf{P}_x[S_{\tau_{a,b}} = a]$.
6. Trouver la suite $(b_n, n \geq 0)$ pour que $((S_n - x)^2 - b_n, n \geq 0)$ soit une \mathcal{F}_n -martingale.
7. Calculer $\mathbf{E}_x[\tau_{a,b}]$.

8. Utiliser les martingales exponentielles pour calculer

$$e^{at} \mathbf{E}_x \left[s^{-\tau_{a,b}} \mathbf{1}_{\{S_{\tau_{a,b}} = a\}} \right] + e^{bt} \mathbf{E}_x \left[s^{-\tau_{a,b}} \mathbf{1}_{\{S_{\tau_{a,b}} = b\}} \right],$$

où $s = (e^t + e^{-t})/2$ et $t \in \mathbf{R}$.

9. Exprimer e^t en fonction de s . En déduire $\mathbf{E}_x \left[s^{-\tau_{a,b}} \mathbf{1}_{\{S_{\tau_{a,b}} = a\}} \right]$.

10. En déduire la fonction génératrice de $\tau_{a,b}$.

11. On note $\tau_b = \inf\{n \geq 0; S_n \geq b\}$. Vérifier que $(\tau_{-n,b}, n \geq 0)$ converge en croissant vers τ_b . En déduire que

$$\mathbf{P}[\tau_b < +\infty] = \lim_{n \rightarrow \infty} \mathbf{P}[S_{\tau_{-n,b}} = b] = 1.$$

12. Calculer $\mathbf{E}[\tau_b]$. Calculer la fonction génératrice de τ_b . Peut-on appliquer le théorème d'arrêt à S_{τ_b} ? En déduire que $\mathbf{E}[\sup_{n \leq \tau_b} |S_n|] = +\infty$.

Le cas $p \neq q$.

On reprend les mêmes notations que ci-dessus.

1. Montrer que $((q/p)^{S_n}, n \geq 0)$ est une \mathcal{F}_n -martingale.
2. Montrer en utilisant la loi forte des grands nombres que $\tau_{a,b}$ est fini p.s. Montrer qu'il n'est pas borné si $a < x < b - 1$ ou si $a + 1 < x < b$.
3. En utilisant le théorème d'arrêt, montrer que, pour $x \in [a, b]$:

$$\mathbf{E}_x[(q/p)^{S_{\tau_{a,b}}}] = (q/p)^x.$$

4. On suppose $x \in]a, b[$. Calculer $\mathbf{P}_x[S_{\tau_{a,b}} = a]$. Que se passe-t-il quand $p \rightarrow 1/2$ (on cherchera une convergence en loi) ?
5. Trouver une suite déterministe $(c_n, n \geq 0)$ telle que $(S_n - c_n, n \geq 0)$ est une \mathcal{F}_n -martingale. En déduire $\mathbf{E}_x[\tau_{a,b}]$.
6. Utiliser les martingales exponentielles pour calculer

$$e^{at} \mathbf{E}_x \left[s^{-\tau_{a,b}} \mathbf{1}_{\{S_{\tau_{a,b}} = a\}} \right] + e^{bt} \mathbf{E}_x \left[s^{-\tau_{a,b}} \mathbf{1}_{\{S_{\tau_{a,b}} = b\}} \right],$$

où $s = p e^t + q e^{-t}$.

7. Exprimer e^t en fonction de s . En déduire $\mathbf{E}_x \left[s^{-\tau_{a,b}} \mathbf{1}_{\{S_{\tau_{a,b}} = a\}} \right]$.
8. En déduire la fonction génératrice de $\tau_{a,b}$. Que se passe-t-il quand $p \rightarrow 1/2$?
9. Montrer que τ_b n'est pas toujours fini p.s. Calculer $\mathbf{P}_x[\tau_b < +\infty]$.
10. Calculer la fonction génératrice de τ_b . Que se passe-t-il quand $p \rightarrow 1/2$? Peut-on appliquer le théorème d'arrêt à S_{τ_b} ?

4.8 Exercices

Exercice 36 Soit ν un temps d'arrêt par rapport à une filtration $(\mathcal{F}_n)_{0 \leq n \leq N}$. On note \mathcal{F}_ν l'ensemble des événements A tels que $A \cap \{\nu = n\} \in \mathcal{F}_n$, pour tout $n \in \{0, \dots, N\}$.

1. Montrer que \mathcal{F}_ν est une sous-tribu de \mathcal{F}_N . \mathcal{F}_ν est souvent appelée “tribu des événements antérieurs à ν ”.
2. Montrer que la variable aléatoire ν est \mathcal{F}_ν -mesurable.
3. Soit X une variable aléatoire réelle. Montrer l'égalité :

$$\mathbf{E}(X|\mathcal{F}_\nu) = \sum_{j=0}^N \mathbf{1}_{\{\nu = j\}} \mathbf{E}(X|\mathcal{F}_j)$$

4. Soit τ un temps d'arrêt tel que $\tau \geq \nu$. Montrer que $\mathcal{F}_\nu \subset \mathcal{F}_\tau$.
5. Sous les mêmes hypothèses, montrer que si (M_n) est une martingale, on a

$$M_\nu = \mathbf{E}(M_\tau|\mathcal{F}_\nu).$$

(On pourra traiter le cas $\tau = N$ d'abord.)

Chapitre 5

Chaînes de Markov à temps discret

5.1 Chaîne de Markov

Définition 5.1.1 Soit E un espace fini ou dénombrable. On dit que :

$$(P(x, y), x \in E, y \in E),$$

est une matrice de transition d'une chaîne de Markov si :

- $P(x, y) \geq 0$ pour tout x et y dans E ,
- $\sum_{y \in E} P(x, y) = 1$.

Pour tout x , on peut définir une probabilité sur E , en posant :

$$P(x, A) = \sum_{y \in A} P(x, y).$$

Définition 5.1.2 On dit qu'un processus $(X_n, n \geq 0)$ est une chaîne de Markov sur un espace E de probabilité de transition P si l'on a, pour tout x_0, \dots, x_n, x_{n+1} :

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

Un exemple générique de chaîne de Markov Soit $(U_n, n \geq 1)$ une suite de variables aléatoires indépendantes sur un espace F , ayant toutes la même loi. Soit H une application de $E \times F$ dans E , alors si l'on pose $X_0 = x_0$ un point fixé et si l'on définit X_n par la formule de récurrence :

$$X_{n+1} = H(X_n, U_{n+1}),$$

alors, la suite $(X_n, n \geq 0)$ est une chaîne de Markov dont la matrice de transition est donnée par :

$$P(x, y) = \mathbf{P}(H(x, U_1) = y).$$

Démonstration :

$$\begin{aligned} \mathbf{P}(X_{n+1} = x_{n+1}, X_n = x_n, \dots, X_0 = x_0) \\ &= \mathbf{P}(H(X_n, U_{n+1}) = x_{n+1}, X_n = x_n, \dots, X_0 = x_0) \\ &= \mathbf{P}(H(x_n, U_{n+1}) = x_{n+1}, X_n = x_n, \dots, X_0 = x_0). \end{aligned}$$

Mais $H(x_n, U_{n+1})$ est indépendante du vecteur (U_1, \dots, U_n) et X_0, \dots, X_n s'écrivent comme des fonctions de (U_1, \dots, U_n) . Donc $H(x_n, U_{n+1})$ est indépendante du vecteur X_0, \dots, X_n . On en déduit que :

$$\begin{aligned} \mathbf{P}(X_{n+1} = x_{n+1}, X_n = x_n, \dots, X_0 = x_0) \\ &= \mathbf{P}(H(x_n, U_{n+1}) = x_{n+1}) \mathbf{P}(X_n = x_n, \dots, X_0 = x_0). \end{aligned}$$

Par définition du conditionnement par un ensemble, on a :

$$\mathbf{P}(X_{n+1} = x_{n+1} | X_n = x_n, \dots, X_0 = x_0) = \mathbf{P}(H(x_n, \mathcal{U}_{n+1}) = x_{n+1}) = P(x_n, x_{n+1}).$$

■

Voici quelques exemples construits à partir de ce résultat.

Nombre de piles consécutifs dans un jeu de pile ou face Soit $(U_n, n \geq 1)$ une suite de variables aléatoires valant 1 avec probabilité p et -1 avec probabilité $1 - p$, p étant un nombre réel de l'intervalle $[0, 1]$. Appelons N_n le nombre de 1 consécutifs avant le n -ième tirage. Par convention $N_0 = 0$ et $N_n = 0$ si le n -ième tirage est 0. Il est, alors, facile de vérifier que pour $n \geq 0$:

$$N_{n+1} = (N_n + 1) \mathbf{1}_{\{U_{n+1} = 1\}}.$$

$(N_n, n \geq 0)$ est donc une chaîne de Markov de probabilité de transition :

$$\begin{cases} P(n, n+1) = p, & n \geq 0, \\ P(n, 0) = 1 - p, & n \geq 0, \\ P(x, y) = 0, & \text{sinon.} \end{cases}$$

Modèle de Cox-Ross-Rubinstein Soit $(V_n, n \geq 1)$ une suite de variables aléatoires indépendantes telle que $\mathbf{P}(V_n = u) = p$ et $\mathbf{P}(V_n = d) = 1 - p$, p étant un réel compris entre 0 et 1, u et v deux réels positifs. Posons $X_0 = x$ et définissons par récurrence X_n par :

$$X_{n+1} = X_n V_{n+1}.$$

$(X_n, n \geq 0)$ est une chaîne de Markov de matrice de transition P donnée par :

$$\begin{cases} P(x, xu) = p, & x \in \mathbf{R}, \\ P(x, xd) = 1 - p, & x \in \mathbf{R}, \\ P(x, y) = 0, & \text{sinon.} \end{cases}$$

C'est le modèle discret le plus couramment utilisé en finance. Il porte le nom de modèle de Cox-Ross-Rubinstein.

Marches aléatoires Soit $(X_n, n \geq 1)$ une suite de variables aléatoires réelles indépendantes suivant toutes la même loi, alors si $S_0 = 0$ et :

$$S_n = X_1 + \dots + X_n,$$

$(S_n, n \geq 0)$ est une chaîne de Markov de matrice de transition donnée par :

$$P(x, y) = \mathbf{P}(X_1 = y - x).$$

Les calculs des lois liées aux chaînes de Markov sont, souvent, faciles à expliciter.

Gestion de stock On décrit la demande des clients par une suite de variables aléatoires indépendantes de même loi $(D_n, n \geq 1)$. On se donne 2 nombres s et S tels que $0 \leq s \leq S$ qui représentent respectivement un niveau de stock maximum et minimum. On procède alors de la façon suivante :

- A l'instant n , si le stock X_n est inférieur à s on la porte à S (instantanément ! mais on raffiner le modèle pour améliorer ce point),
- On retranche la demande à l'instant $n + 1$, D_{n+1} au nouveau stock (si c'est possible).

Cela donne :

$$X_n = \begin{cases} (X_n - D_{n+1})_+ & \text{si } s < X_n \leq S \\ (S - D_{n+1})_+ & \text{si } X_n \leq s. \end{cases}$$

On suppose que la suite décrivant la demande $(D_n, n \geq 1)$ est une suite de variables aléatoires indépendantes suivant toutes la même loi. La suite $(X_n, n \geq 0)$ est, sous cette hypothèse, une chaîne de Markov. Voici quelques questions auxquelles on souhaiterait pouvoir répondre :

- que vaut $\frac{1}{n} \sum_{i=0}^n X_i$ la moyenne empirique du stock,
- que vaut $\mathbf{P}(X_j = a)$, la loi du stock,
- comment évaluer les demandes non satisfaites.

Processus de vie et de mort On considère $(\xi_i^{(n)}, i \geq 1, n \geq 0)$ une suite doublement indexée de variables aléatoires indépendantes de même loi, à valeurs dans les entiers positifs. On note $\phi(z)$ sa fonction génératrice de Laplace donnée par :

$$\phi(z) = \mathbf{E}(z^\xi), \text{ pour } 0 \leq z \leq 1.$$

On construit alors le processus $(X_n, n \geq 0)$ de la façon suivante, on pose $X_0 = 1$ puis :

$$X_{n+1} = \sum_{i=1}^{X_n} \xi_i^{(n+1)},$$

avec la convention que, si $X_n = 0$ la somme vaut 0.

L'interprétation intuitive, qui justifie le nom du processus, est que chaque individu donne naissance à un nombre d'enfants qui suit la loi de ξ puis disparaît.

Si l'on note $\xi^{(n)} = (\xi_i^{(n)}, i \geq 1)$, X_{n+1} peut s'écrire sous la forme :

$$X_{n+1} = \phi(X_n, \xi^{(n+1)}),$$

où $\phi(x, \xi) = \sum_{i=1}^x \xi_i$. La suite $(\xi^{(n)}, n \geq 1)$ étant une suite de variables aléatoires (à valeurs dans un espace de suite !) de même loi. On en déduit que $(X_n, n \geq 0)$ est une chaîne de Markov et que sa matrice de transition P est donnée par :

$$P(x, y) = \mathbf{P} \left(\sum_{i=1}^x \xi_i = y \right).$$

Nous allons voir que l'on est capable d'évaluer la probabilité d'extinction de la population que l'on notera p . Pour cela, notons que :

$$\mathbf{E} (z^{X_{n+1}} | \mathcal{F}_n) = \mathbf{E} \left(z^{\sum_{i=1}^{X_n} \xi_i^{(n+1)}} | \mathcal{F}_n \right)$$

Comme X_n est \mathcal{F}_n -mesurable et $\xi^{(n+1)}$ est indépendante de \mathcal{F}_n , en utilisant le lemme 3.2.6 page 54 on obtient :

$$\mathbf{E} (z^{X_{n+1}} | \mathcal{F}_n) = \mathbf{E}(z^\xi)^{X_n} = \phi(z)^{X_n}.$$

En prenant l'espérance des deux membres de cette égalité on obtient :

$$\mathbf{E}(z^{X_{n+1}}) = \mathbf{E}(\phi(z)^{X_n}).$$

Appellons $\phi_n(z)$ la fonction génératrice de Laplace de Z_n :

$$\phi_n(z) = \mathbf{E}(z^{X_n}),$$

on peut réécrire la dernière équation sous la forme :

$$\phi_{n+1}(z) = \phi_n(\phi(z)),$$

Comme $\phi_0(z) = z$, on en déduit que :

$$\phi_n(z) = \phi(\phi(\dots\phi(z))) = \phi^{\circ n}(z),$$

puis que :

$$\phi_{n+1}(z) = \phi(\phi_n(z)). \quad (5.1)$$

D'autre part la probabilité d'extinction p est donnée par :

$$p = \mathbf{P}(\text{il existe un instant } n \text{ tel que } X_n = 0),$$

Comme la famille d'événements $\{X_n = 0\}$ est croissante on en déduit que :

$$p = \lim_{n \rightarrow +\infty} \mathbf{P}(X_n = 0).$$

Remarquons, d'autre part, que :

$$\mathbf{P}(X_n = 0) = \mathbf{E}(0^{X_n}) = \phi_n(0).$$

On peut alors utiliser l'équation (5.1) et la continuité de ϕ pour prouver que :

$$p = \phi(p).$$

Le problème probabiliste de calcul de la probabilité d'extinction de la population se ramène simplement au problème de l'étude de la fonction ϕ . On peut déduire des propriétés de ϕ (exercice) que :

- si $\mathbf{E}(\xi) \leq 1$ alors $p = 1$,
- si $\mathbf{E}(\xi) > 1$, il existe une unique solution σ à l'équation $\phi(\sigma) = \sigma$ avec $\sigma < 1$ et l'on a $p = \sigma$.

On voit donc que :

- si $\mathbf{E}(\xi) \leq 1$, la population disparaît forcément au bout d'un temps fini,
- si $\mathbf{E}(\xi) > 1$, la population peut survivre indéfiniment.

5.2 Calcul de lois

Proposition 5.2.1 Soit $(X_n, n \geq 0)$ une chaîne de Markov sur un espace E de matrice de transition P et de loi initiale μ (la loi initiale est donnée par la loi de X_0 , i.e. par la famille de nombres $\mu(x) = \mathbf{P}(X_0 = x)$). La loi du vecteur (X_0, \dots, X_n) se calcule par la formule :

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

Démonstration : Par définition de la probabilité conditionnelle, on a :

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

Comme $(X_n, n \geq 0)$ est une chaîne de Markov on en déduit que :

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

puis par récurrence que :

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

■

Remarque 5.2.2 Notez que la loi du vecteur (X_0, \dots, X_n) ne dépend que de P , la matrice de transition, et μ , la loi initiale.

On déduit de ce résultat une expression pour la loi de X_n . Introduisons les notations suivantes :

- $(Pf)(x) = \sum_{y \in E} P(x, y) f(y)$, pour f une fonction de E dans \mathbf{R} ,
- $(\mu P)(y) = \sum_{x \in E} \mu(x) P(x, y)$, pour μ une loi de probabilité sur E ,
- $(PQ)(x, y) = \sum_{z \in E} P(x, z) Q(z, y)$, pour P et Q deux matrices de transition.
- P^n est définie par récurrence par :

$$P^{n+1}(x, y) = \sum_{z \in E} P(x, z) P^n(z, y).$$

Noter que Pf est une fonction sur E , PQ une matrice de transition et μP une loi de probabilité sur E .

Proposition 5.2.3 Soit $(X_n, n \geq 0)$ un chaîne de Markov de matrice de transition P et de loi initiale μ . Alors :

$$\mathbf{P}(X_n = x_n) = \sum_{x_0 \in E, x_1 \in E, \dots, x_{n-1} \in E} \mu(x_0) P(x_0, x_1) \times \dots \times P(x_{n-1}, x_n).$$

soit, avec les notations précédentes :

$$\mathbf{P}(X_n = x_n) = (\mu P^n)(x_n),$$

Remarque 5.2.4 On voit donc que le calcul de la loi de X_n se ramène à un calcul matriciel.

Démonstration : La démonstration est une application directe de la formule donnant la loi marginale d'une des coordonnées d'un vecteur. ■

5.3 Exemple d'utilisation des chaînes de Markov

Combien de piles consécutifs voit-on dans 100 tirages à pile ou face Soit $(X_n, n \geq 1)$ une suite de variables aléatoires indépendantes valant 1 (ou pile) avec probabilité p , 0 (ou face) avec probabilité $1-p$ ($0 \leq p \leq 1$). Nous avons vu que le nombre de piles consécutifs avant l'instant n , définit par $N_0 = 0$ et :

$$N_{n+1} = (N_n + 1) \mathbf{1}_{\{X_{n+1} = 1\}}.$$

est une chaîne de Markov. Définissons maintenant τ_l , pour un entier l fixé, par :

$$\tau_l = \inf\{n \geq 1, N_n = l\}.$$

τ_l est ainsi le premier instant où l'on voit une suite de l -piles consécutifs. Introduisons N_n^l (le processus arrêté à l'instant τ_l) par :

$$N_n^l = N_{n \wedge \tau_l}.$$

Notons que l'événement $\{N_n^l = l\}$ correspond à :

“J'ai vu *au moins* l consécutifs dans les n premiers tirages”.

D'autre part $(N_n^l, n \geq 0)$ reste une chaîne de Markov issue de 0 puisque :

$$N_{n+1}^l = (N_n^l + 1) \mathbf{1}_{\{X_{n+1} = 1, N_n^l < l\}} + l \mathbf{1}_{\{N_n^l = l\}}.$$

Notons $(P(x, y), x, y \in \{0, \dots, l\})$ sa matrice de transition, alors :

$$\mathbf{P}(N_n^l = l) = P^n(0, l).$$

Pour évaluer cette probabilité il nous reste à expliciter P et à calculer sa puissance pour diverses valeurs de l (mieux vaut alors utiliser un logiciel du genre MathLab (payant) ou Scilab (gratuit <http://www-rocq.inria.fr/scilab/> !).

Les seuls éléments non nul de la matrice P sont donnés par :

$$\begin{cases} P(x, x+1) = p & \text{si } x < l, \\ P(x, 0) = 1-p & \text{si } x < l, \\ P(l, l) = 1. \end{cases}$$

P s'exprime donc sous forme matricielle par :

$$P = \begin{pmatrix} 1-p & p & 0 & \dots & 0 & 0 \\ 1-p & 0 & p & \dots & 0 & 0 \\ \dots & \dots & \dots & \dots & \dots & \dots \\ 1-p & 0 & 0 & \dots & p & 0 \\ 1-p & 0 & 0 & \dots & 0 & p \\ 0 & 0 & 0 & \dots & 0 & 1 \end{pmatrix}.$$

Voici une procédure Scilab qui calcule P , sa puissance 100 et enfin $P^{100}(0, l)$.

```

function [Pro]= proba(l,p)
    P=[p*diag(ones(l,1));zeros(1,l-1),1];
    P=[(1-p)*ones(l,1);0],P];
    P100=P^100;
    Pro = P100(1,l+1)

```

On peut alors obtenir sans difficultés la fonction de répartition ou la loi du nombre maximal de 1 consécutifs dans un séquence aléatoire de 100 0 ou 1 indépendants avec $p = 1/2$. Pour la loi, après avoir tapé :

```
for l=2:10;
    write(%io(2),string(l)+' '+string(proba(l,1/2)-
    proba(l+1,1/2)));
end;
```

on obtient :

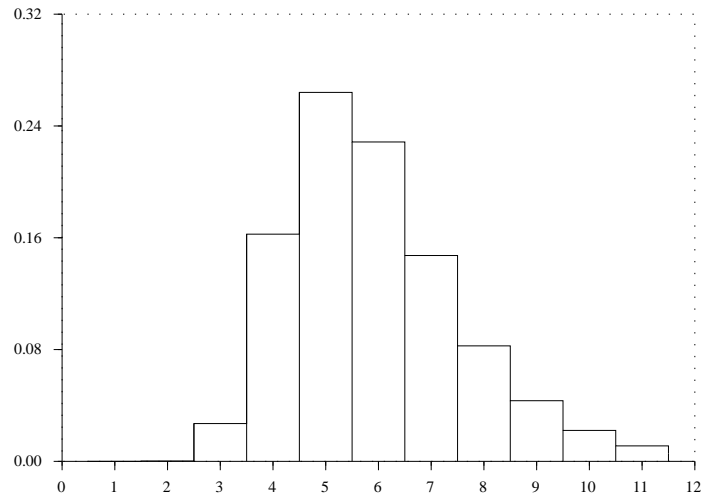


FIG. 5.1 – Histogramme de la loi du nombre maximum de F consécutifs après 100 tirages

et pour la fonction de répartition :

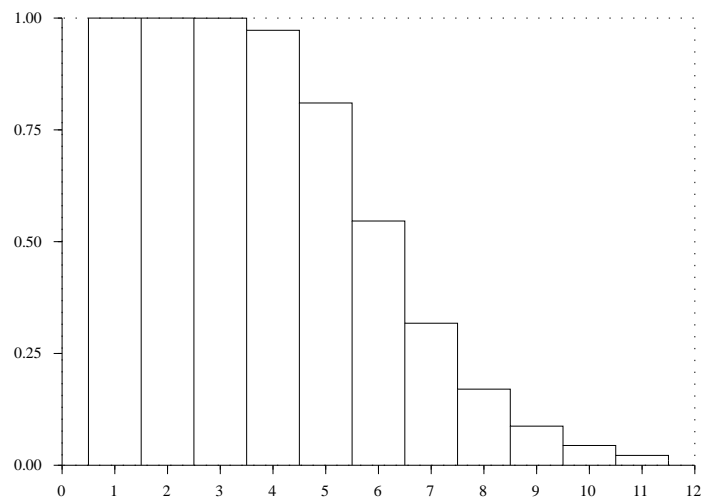


FIG. 5.2 – Probabilité d'avoir au moins n F consécutifs après 100 tirages

On est donc “à peu près” sûr (on se trompe 2 fois sur 10, quand même !) de voir 5 piles consécutifs dans une séquence 100 tirages à pile ou face, pour une pièce équilibrée. Le fait de ne pas voir 5 1 consécutifs doit être considéré comme louche et peut servir de test pour savoir si une suite a vraiment été tirée au hasard. Il est en effet très rare qu’une personne tentant d’imiter une suite au hasard ait l’idée de mettre 5 1 à la suite. Ça marche plutôt bien (exercice essayer avec l’un de vos camarades, qui ne suit pas le cours de processus bien sûr !).

5.4 Chaînes de Markov et espérance conditionnelle

Définition 5.4.1 Soit μ une probabilité sur E , un ensemble fini et P une matrice de transition sur E on dit que $(X_n, n \geq 0)$ est une chaîne de Markov par rapport à une filtration $(\mathcal{F}_n, n \geq 0)$, si c’est un processus adapté à cette filtration tel que pour toute fonction f de E dans \mathbf{R} , on a :

$$\mathbf{E}(f(X_{n+1})|\mathcal{F}_n) = (Pf)(X_n),$$

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

Remarque 5.4.2 – Notons que cette définition est équivalente à :

$$\mathbf{P}(X_{n+1} \in A|\mathcal{F}_n) = P(X_n, A).$$

avec $P(x, A) = \sum_{y \in A} P(x, y)$. Remarquons que si A est réduit à un point, on a alors :

$$\mathbf{P}(X_{n+1} = y|\mathcal{F}_n) = P(X_n, y).$$

– Notons aussi, qu’un processus de Markov par rapport à une filtration $(\mathcal{F}_n, n \geq 0)$ donnée l’est aussi par rapport à la filtration naturelle définie par :

$$\mathcal{F}_n^0 = \sigma(X_0, \dots, X_n).$$

En effet, comme $(X_n, n \geq 0)$ est adapté à $(\mathcal{F}_n, n \geq 0)$ on a pour tout n , $\mathcal{F}_n^0 \subset \mathcal{F}_n$ et donc :

$$\mathbf{E}(f(X_{n+1})|\mathcal{F}_n^0) = \mathbf{E}(\mathbf{E}(f(X_{n+1})|\mathcal{F}_n)|\mathcal{F}_n^0) = \mathbf{E}((Pf)(X_n)|\mathcal{F}_n^0) = (Pf)(X_n).$$

Bien sûr si $(X_n, n \geq 0)$ est une chaîne de Markov par rapport à une filtration donnée c’est aussi un processus de Markov au sens élémentaire.

Proposition 5.4.3 Soit $(X_n, n \geq 0)$ est une chaîne de Markov par rapport à la filtration $(\mathcal{F}_n, n \geq 0)$, alors :

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

Démonstration :

$$\begin{aligned} & \mathbf{P}(X_{n+1} = x_{n+1}, X_n = x_n, \dots, X_0 = x_0) \\ &= \mathbf{E} \left(\mathbf{E} \left(\mathbf{1}_{\{X_{n+1} = x_{n+1}\}} \mathbf{1}_{\{X_n = x_n, \dots, X_0 = x_0\}} | \mathcal{F}_n \right) \right) \\ &= \mathbf{E} \left(\mathbf{1}_{\{X_n = x_n, \dots, X_0 = x_0\}} \mathbf{E} \left(\mathbf{1}_{\{X_{n+1} = x_{n+1}\}} | \mathcal{F}_n \right) \right) \\ &= \mathbf{E} \left(\mathbf{1}_{\{X_n = x_n, \dots, X_0 = x_0\}} P(X_n, x_{n+1}) \right) \\ &= \mathbf{E} \left(\mathbf{1}_{\{X_n = x_n, \dots, X_0 = x_0\}} P(x_n, x_{n+1}) \right) \\ &= P(x_n, x_{n+1}) \mathbf{P}(X_n = x_n, \dots, X_0 = x_0). \end{aligned}$$

La dernière égalité donne le résultat souhaité par définition de la probabilité conditionnelle élémentaire.

■

Remarque 5.4.4 On peut montrer qu'un processus de Markov au sens élémentaire est un processus de Markov par rapport à sa filtration naturelle. Nous démontrons pas ce résultat qui pourra être traité en exercice par les élèves motivés par les aspects mathématiques.

Ces deux résultats montrent que l'équivalence des deux définitions de processus de Markov lorsque l'on se restreint à la filtration naturelle, i.e. $(\mathcal{F}_n^0, n \geq 0)$ avec $\mathcal{F}_n^0 = \sigma(X_0, \dots, X_n)$. L'intérêt de la définition abstraite est de permettre de prendre en compte une filtration $(\mathcal{F}_n, n \geq 0)$ contenant plus d'information que la filtration naturelle (donc telle que $\mathcal{F}_n^0 \subset \mathcal{F}_n$).

Calcul algorithmique de $\mathbf{E}(f(X_N))$ Nous allons donner une illustration de l'intérêt de nouveau formalisme pour les chaînes de Markov.

Soit $(X_n, n \geq 0)$ est une chaîne de Markov de matrice de transition P , partant de x_0 en 0. On a :

$$\mathbf{E}(f(X_N)) = (P^N f)(x_0).$$

On peut alors calculer $\mathbf{E}(f(X_N))$ itérativement en posant, pour n décroissant à partir de N :

$$\begin{cases} u(n, x) = [P u(n+1, \cdot)](x) = \sum_{y \in E} P(x, y) u(n+1, y), \\ u(N, x) = f(x), \end{cases} \quad (5.2)$$

puis :

$$\mathbf{E}(f(X_N)) = (P^N f)(x_0) = u(0, x_0).$$

Notez que l'équation (5.2) décrit un algorithme (récursif) permettant d'évaluer $\mathbf{E}(f(X_N))$ qui explicite le fait que la loi de X_n est donnée par $(P^n(x_0, y), y \in E)$.

Pour ce convaincre de ce résultat nous allons montrer que la processus $(u(n, X_n), n \geq 0)$ est une martingale. Pour cela, calculons $\mathbf{E}(u(n+1, X_{n+1}) | \mathcal{F}_n)$:

$$\mathbf{E}(u(n+1, X_{n+1}) | \mathcal{F}_n) = \mathbf{E}(f(X_{n+1}) | \mathcal{F}_n),$$

si $\phi(x) = u(n+1, x)$, on a donc :

$$\mathbf{E}(u(n+1, X_{n+1}) | \mathcal{F}_n) = (P\phi)(X_n) = \sum_{y \in E} P(x, y) u(n+1, y) \Big|_{x=X_n}.$$

Mais on a, par définition de u :

$$\sum_{y \in E} P(x, y) u(n+1, y) = u(n, x),$$

donc :

$$\mathbf{E}(u(n+1, X_{n+1}) | \mathcal{F}_n) = u(n, X_n).$$

Ce qui prouve que $(u(n, X_n), n \geq 0)$ est une martingale. On en déduit que son espérance est constante, soit :

$$\mathbf{E}(u(0, X_0)) = \mathbf{E}(u(n, X_n)).$$

Mais $u(n, x) = f(x)$ et $X_0 = x_0$ et la précédente égalité se réécrit :

$$u(0, x_0) = \mathbf{E}(f(X_N)).$$

Exemples : Soit $(S_n, n \geq 1)$ la marche aléatoire symétrique et supposons que l'on cherche à calculer, pour f une fonction donnée :

$$\mathbf{E}(f(S_N)).$$

Le théorème précédent, nous dit que, cette quantité est égale à $u(0, 0)$, et que u se calcule à l'aide de l'équation :

$$\begin{cases} u(N, x) = f(x), \\ u(n, x) = \frac{1}{2}u(n+1, x-1) + \frac{1}{2}u(n+1, x+1). \end{cases}$$

Notez que ces deux équations permettent d'écrire un programme (inefficace car de complexité exponentielle) calculant $v(n, x)$.

```
double fonction u(int n, double x)
{
    if(n==N) then return f(x);
    return 0.5 * (u(n+1, x-1)+u(n+1, x+1));
}
```

Il est cependant possible d'écrire un version efficace de ce programme en utilisant le fait que à l'instant n il n'est nécessaire de calculer u qu'en $x-n, x-n+1, \dots, x+n-1, x+n$.

Exemples : Imaginons que l'on cherche à calculer le prix d'un put américain dans le modèle de Cox Ross. Cela signifie que, a, b sont des réels positifs, p est un nombre réel entre 0 et 1, que la matrice de transition de la chaîne S_n s'écrit $P(x, x(1+a)) = p, P(x, x(1+b)) = 1-p$, et que l'on cherche à calculer :

$$\mathbf{E}((K - S_N)_+).$$

Le théorème précédent nous dit que, $\mathbf{E}((K - S_N)_+) = u(0, S_0)$ si v est l'unique solution de :

$$\begin{cases} v(N, x) = (K - x)_+, \\ v(n, x) = pu(n+1, x(1+a)) + (1-p)u(n+1, x(1+b)). \end{cases}$$

5.4.1 Solution d'un problème d'arrêt optimal

On se propose de calculer :

$$\bar{V}_n = \sup_{\tau \text{ } \mathcal{F}_n\text{-temps d'arrêt}, n \leq \tau \leq N} \mathbf{E}(\psi(X_\tau) | \mathcal{F}_n),$$

Lorsque $(X_n, n \geq 0)$ est une chaîne de Markov. Ce résultat est directement relié à des algorithmes de calcul de prix d'option américaine en finance et d'autres problèmes appliqués. Commençons par énoncer le résultat essentiel de cette partie.

Théorème 5.4.1 Soit $(X_n, n \geq 0)$ une chaîne de Markov issue du point x_0 en 0, de matrice de transition $P(x, y)$ sur un espace d'état fini E , par rapport à une filtration $(\mathcal{F}_n, n \geq 0)$. Soit $\psi(x)$ une fonction définie pour $x \in E$. On cherche à calculer :

$$\sup_{\tau \text{ } \mathcal{F}_n\text{-temps d'arrêt}, \tau \leq N} \mathbf{E}(\psi(X_\tau)).$$

Soit u la solution unique de :

$$\begin{cases} u(N, x) = \psi(N, x), & x \in E \\ u(n, x) = \max(\{Pu(n+1, \cdot)\}(x), \psi(x)) \\ = \max\left(\sum_{y \in E} P(x, y)u(n+1, y), \psi(x)\right) & n < N, x \in E. \end{cases} \quad (5.3)$$

Alors

$$\sup_{\tau \mathcal{F}_n\text{-t.a.}, 0 \leq \tau \leq N} \mathbf{E}(\psi(X_\tau)) = u(0, x_0).$$

De plus le temps d'arrêt :

$$\tau_0 = \inf\{k \geq 0, u(k, X_k) = \psi(X_k)\},$$

est un temps d'arrêt optimal i.e. qui vérifie

$$\sup_{\tau \mathcal{F}_n\text{-t.a.}, 0 \leq \tau \leq N} \mathbf{E}(\psi(X_\tau)) = \mathbf{E}(\psi(X_{\tau_0})).$$

Démonstration : On a, bien sûr, vu la définition de u :

$$u(n, x) \geq \{Pu(n+1, \cdot)\}(x).$$

De plus, par définition d'une chaîne de Markov :

$$\{Pu(n+1, \cdot)\}(X_n) = \mathbf{E}(u(n+1, X_{n+1}) | \mathcal{F}_n).$$

On a donc :

$$u(n+1, X_{n+1}) - u(n, X_n) \leq u(n+1, X_{n+1}) - \mathbf{E}(u(n+1, X_{n+1}) | \mathcal{F}_n).$$

Posons :

$$Y_n = u(n, X_n) - \mathbf{E}(u(n, X_n) | \mathcal{F}_{n-1})$$

et $M_n = u(0, x_0) + Y_1 + \dots + Y_n$, pour $n \geq 0$. M_n est une \mathcal{F}_n martingale, en effet :

$$\begin{aligned} \mathbf{E}(M_{n+1} - M_n | \mathcal{F}_n) &= \mathbf{E}(Y_{n+1} | \mathcal{F}_n) \\ &= \mathbf{E}(u(n+1, X_{n+1}) - \mathbf{E}(u(n+1, X_{n+1}) | \mathcal{F}_n) | \mathcal{F}_n) \\ &= 0. \end{aligned}$$

De plus, comme $u(n+1, X_{n+1}) - u(n, X_n) \leq Y_{n+1}$ et $u(0, x_0) = M_0$, il est immédiat d'en déduire par récurrence que, pour tout $n \geq 0$:

$$u(n, X_n) \leq M_n.$$

Maintenant si τ est un temps d'arrêt tel que $\tau \leq N$, on a :

$$u(\tau, X_\tau) \leq M_\tau.$$

Alors, en utilisant le fait que $(M_n, n \geq 0)$ est une martingale, et en utilisant le théorème d'arrêt on obtient :

$$\mathbf{E}(M_\tau) = \mathbf{E}(M_0) = u(0, x_0).$$

On en déduit donc que :

$$\mathbf{E}(u(\tau, X_\tau)) \leq \mathbf{E}(M_\tau) = u(0, x_0).$$

On vient donc de prouver que $\mathbf{E}(u(\tau, X_\tau)) \leq u(0, x_0)$. Comme, de plus, $u(n, x) \geq \psi(x)$, on a pour tout temps d'arrêt inférieur à N :

$$\mathbf{E}(\psi(X_\tau)) \leq u(0, x_0).$$

Donc :

$$\sup_{n \leq \tau \leq N, \tau \text{ t.a.}} \mathbf{E}(\psi(X_\tau)) \leq u(0, x_0).$$

Pour finir la démonstration du résultat annoncé, il suffit de trouver un temps d'arrêt τ_0 inférieur à N tel que :

$$\mathbf{E}(\psi(X_{\tau_0})) = u(0, x_0).$$

Pour cela posons :

$$\tau_0 = \inf\{p \geq 0, u(p, X_p) = \psi(X_p)\}.$$

On peut démontrer (exercice) que τ_0 est un temps d'arrêt. D'autre part on a $\tau_0 \leq N$ puisque $u(N, x) = \psi(N, x)$.

Sur l'événement $\{p < \tau_0\}$ on a $u(p, X_p) \neq \psi(X_p)$, et donc :

$$u(p, X_p) = \{P u(p+1, \cdot)\}(X_p) = \mathbf{E}(u(p+1, X_{p+1}) | \mathcal{F}_p).$$

Ceci implique que, sur l'événement $\{p < \tau_0\}$:

$$u(p+1, X_{p+1}) - u(p, X_p) = Y_{p+1}.$$

En sommant ces égalités entre $p = 0$ et $\tau_0 - 1$ on obtient :

$$u(\tau_0, X_{\tau_0}) = M_{\tau_0}.$$

On voit donc, en utilisant le théorème d'arrêt, que :

$$\mathbf{E}(u(\tau_0, X_{\tau_0})) = \mathbf{E}(M_{\tau_0}) = \mathbf{E}(M_0) = u(0, x_0).$$

Comme, par définition de τ_0 , l'on a $u(\tau_0, X_{\tau_0}) = \psi(X_{\tau_0})$, on en déduit que :

$$\mathbf{E}(\psi(X_{\tau_0})) = u(0, x_0).$$

Ceci prouve donc que :

$$\sup_{\tau \leq N, \tau \text{ t.a.}} \mathbf{E}(\psi(X_\tau)) = u(0, x_0).$$

■

Exemples : Soit $(S_n, n \geq 1)$ la marche aléatoire symétrique et supposons que l'on cherche à calculer, pour f une fonction donnée :

$$\sup_{\tau \leq N, \tau \text{ t.a.}} f(S_\tau).$$

Le théorème précédent, nous dit que, cette quantité est égale à $u(0, 0)$, et que u se calcule à l'aide de l'équation :

$$\begin{cases} u(N, x) = f(x), \\ u(n, x) = \max\left(\frac{1}{2}u(n+1, x-1) + \frac{1}{2}u(n+1, x+1), f(x)\right). \end{cases}$$

Une fois que u a été calculée on peut expliciter le temps d'arrêt optimal par

$$\tau_0 = \inf\{n \geq 0, u(n, S_n) = f(S_n)\}.$$

Exemples : Imaginons que l'on cherche à calculer le prix d'un put américain dans le modèle de Cox Ross. Cela signifie que, α, b, r sont des réels positifs, p est un nombre réel entre 0 et 1, que la matrice de transition de la chaîne S_n s'écrit $P(x, x(1 + \alpha)) = p, P(x, x(1 + b)) = 1 - p$, et que l'on cherche à calculer :

$$V_0 = \sup_{\tau \text{ t.a.}, \tau \leq N} \mathbf{E}((K - S_\tau)_+).$$

Le théorème précédent nous dit que, $V_0 = u(0, S_0)$ avec v solution de :

$$\begin{cases} v(N, x) = (K - x)_+, \\ v(n, x) = \max(pu(n + 1, x(1 + \alpha)) + (1 - p)u(n + 1, x(1 + b)), f(x)). \end{cases}$$

5.5 Exercices

Exercice 37 On considère l'espace à deux états $E = \{a, b\}$. La matrice markovienne la plus générale s'écrit

$$P = \begin{pmatrix} 1 - \alpha & \alpha \\ \beta & 1 - \beta \end{pmatrix}.$$

1. Condition pour que la chaîne de Markov ne soit pas périodique.
2. Condition pour que la chaîne de Markov soit irréductible.
3. On suppose que la chaîne est irréductible. Calculer la probabilité invariante.
4. On suppose que la chaîne de Markov associée à P est irréductible. Montrer que les puissances de P s'écrivent

$$P^n = \begin{pmatrix} 1 - p & p \\ 1 - p & p \end{pmatrix} + \gamma^n \begin{pmatrix} p & -p \\ -1 + p & 1 - p \end{pmatrix}.$$

où on déterminera p et γ .

5. On suppose que la chaîne de Markov est aperiodique et irréductible. En déduire que $\lim_{n \rightarrow \infty} \pi_0 P^n$ est indépendant de la loi initiale π_0 .

Exercice 38 Soit $(X_{2n+1}, n \geq 0)$ une suite de v.a.r. i.i.d telles que

$$\mathbf{P}[X_1 = 1] = \mathbf{P}[X_1 = -1] = 1/2.$$

On définit pour tout $n \geq 1$, $X_{2n} = X_{2n-1}X_{2n+1}$.

1. Vérifier que les v.a.r. $X_{2n}, n \geq 1$ sont i.i.d. Donner leur loi.
2. Montrer que X_n et X_{n+1} sont indépendantes.
3. Déduire de 1 et 2 que les v.a.r. $X_n, n \geq 1$ sont indépendantes deux à deux.
4. Calculer $\mathbf{P}[X_{m+n} = j \mid X_m = i]$ pour tout $n \geq 1, i, j = \pm 1$. En déduire que les équations de Chapman-Kolmogorov sont satisfaites.
5. Calculer $\mathbf{P}[X_{2n+1} = 1 \mid X_{2n} = -1, X_{2n-1} = 1]$. En déduire que X n'est pas une chaîne de Markov.
6. Montrer en revanche que $(Z_n = (X_n, X_{n+1}), n \geq 1)$ est une chaîne de Markov non homogène à valeurs dans $\{-1, 1\}^2$. On déterminera la matrice de transition en distinguant suivant que n est pair ou impair.

Exercice 39 Soit X une chaîne de Markov à espace d'état dénombrable. Soit $n \geq 1$. On note $p_{i,j}(n) = \mathbf{P}[X_n = j \mid X_0 = i]$, la probabilité de transition de l'état i à l'état j en n étapes. Par convention on pose $p_{i,i}(0) = 1$ et $p_{i,j}(0) = 0$. On note

$$l_{i,j}(n) = \mathbf{P}[X_n = j; X_k \neq i, 1 \leq k < n \mid X_0 = i],$$

la probabilité partant de l'état i d'atteindre l'état j en n étapes, sans repasser par l'état i . On note également pour $i \neq j$,

$$f_{i,j}(n) = \mathbf{P}[X_n = j; X_k \neq j, 1 \leq k < n \mid X_0 = i],$$

la probabilité partant de l'état i d'atteindre pour la première fois l'état j en n étapes. On introduit les fonctions suivantes définies sur $[-1, 1]$:

$$\begin{aligned} P_{i,j}(s) &= \sum_{n \geq 0} p_{i,j}(n) s^n \\ L_{i,j}(s) &= \sum_{n \geq 1} l_{i,j}(n) s^n \\ F_{i,j}(s) &= \sum_{n \geq 1} f_{i,j}(n) s^n. \end{aligned}$$

1. Montrer que pour tout $i \neq j$,

$$P_{i,j}(s) = P_{i,i}(s)L_{i,j}(s)$$

$$P_{i,j}(s) = P_{j,j}(s)F_{i,j}(s)$$

2. On définit le premier temps d'atteinte de j par $\tau = \inf\{k; X_k = j\}$ et le dernier temps de passage par $\tau'_n = [n - \sup_{k \leq n}\{X_k = i\}]\mathbf{1}\{X_n = j\}$. Montrer que les variables aléatoires τ'_n ne sont pas des temps d'arrêt. Dédurre du 1 que si $P_{i,i} = P_{j,j}$, alors on a $\mathbf{P}[\tau = k] = \mathbf{P}[\tau'_n = k]$ pour tout $n \geq k$. Donc de manière imagée le premier temps de passage et le dernier temps de passage "ont même loi".

3. Donner un exemple de chaîne de Markov à espace d'état dénombrable vérifiant $P_{j,j} = P_{i,i}$ pour tout i et j .

Exercice 40 On considère une marche aléatoire symétrique $S_0 = 0$ et $S_n = X_1 + \dots + X_n$, où les v.a.r. $(X_n, n \geq 1)$ sont i.i.d. et

$$\mathbf{P}[X_1 = 1] = \mathbf{P}[X_1 = -1] = 1/2.$$

On considère la famille de variables aléatoires

$$Y_n = \sup\{k \leq n; S_{2k} = 0\}, \quad n \geq 1.$$

On désire évaluer la valeur $\mathbf{P}[Y_n = k]$ pour $k \leq n$. Dans le jeu de pile ou face, la v.a.r. Y_n est le dernier instant où les deux joueurs sont à égalité.

1. Calculer $\mathbf{P}[S_m = b]$.

2. On admet la formule suivante :

$$\mathbf{P}[S_1 \neq 0; \dots; S_m \neq 0; S_m = b] = \frac{|b|}{m} \mathbf{P}[S_m = b].$$

Calculer $\mathbf{P}[S_1 \neq 0; \dots; S_{2n} \neq 0]$. On pourra utiliser le fait que

$$\frac{2k}{(n+k)(n-k)} = \frac{1}{n-k-1} - \frac{1}{n+k-1}.$$

3. En déduire que $\mathbf{P}[S_1 \neq 0; \dots; S_{2n} \neq 0] = \mathbf{P}[S_{2n} = 0]$.

4. En utilisant le résultat précédent montrer que

$$\mathbf{P}[Y_n = k] = \mathbf{P}[S_{2k} = 0] \mathbf{P}[S_{2n-2k} = 0].$$

Remarquer que contrairement à l'intuition la loi du dernier zéro avant $2n$ est symétrique par rapport à n .

5. Soit $0 < y < x < 1$. En utilisant la formule de Stirling ($n! \sim n^{n+1/2} e^{-n} \sqrt{2\pi}$ pour n grand), montrer que pour n grand

$$\mathbf{P}[ny \leq Y_n \leq nx] \sim \frac{1}{\pi} \int_y^x \frac{1}{\sqrt{u(1-u)}} du.$$

6. En déduire que la suite $(Y_n/n, n \geq 1)$ converge en loi.

7. Montrer que asymptotiquement, avec probabilité $1/2$,

$$Y_n/n \in \left[0, \frac{2-\sqrt{2}}{4}\right] \cup \left[\frac{2+\sqrt{2}}{4}, 1\right].$$

Exercice 41 1. Soit Y une chaîne de Markov à espace d'état discret. On suppose que

$$\mathbf{P}_j[Y_n = j \text{ infiniment souvent}] = 0.$$

On note $\tau = \sup\{n; Y_n = j\}$. Montrer que $\sum_{n \geq 0} \mathbf{P}_j[\tau = n] = 1$. En déduire que

$$\mathbf{P}_j[Y_n \neq j \quad \forall n \geq 1] = \frac{1}{1 + \sum_{n \geq 1} \mathbf{P}_j[Y_n = j]}.$$

On considère maintenant $(X_n, n \geq 1)$ une suite de v.a.r. i.i.d. telles que

$$\mathbf{P}[X_n = 1] = p, \quad \mathbf{P}[X_n = -1] = q = 1 - p.$$

On définit $S_0 = 0, S_n = X_1 + \dots + X_n$. Vérifier que $(S_n, n \geq 0)$ est une chaîne de Markov homogène à valeurs dans \mathbf{Z} . On suppose $p \neq 1/2$.

2. Évaluer $\mathbf{P}[S_n = 0]$.

3. Écrire la fonction $f(x) = [1 - x]^{-1/2}$ sous forme d'une série entière.

4. En déduire que $\mathbf{P}[S_n \neq 0 \quad \forall n \geq 1] = |p - q|$.

Exercice 42 Soit $(X_n, n \geq 0)$ une chaîne de Markov homogène à espace d'état dénombrable E . On note $T_y = \inf\{n \geq 1; X_n = y\}$. Si x et y sont deux états distincts, on note

$$\rho_{xy} = \mathbf{P}[T_y < \infty \mid X_0 = x].$$

1. Montrer que

$$\rho_{xy} > 0 \iff \exists n \geq 1 \text{ tel que } \mathbf{P}[X_n = y \mid X_0 = x] > 0.$$

On suppose $\rho_{xy} > 0$. Soit $n_0 = \inf\{n; \mathbf{P}[X_n = y \mid X_0 = x] > 0\}$.

2. Montrer qu'il existe des états $x_1, x_2, \dots, x_{n_0-1}$ tels que

$$p(x, x_1)p(x_1, x_2) \cdots p(x_{n_0-1}, y) > 0, \text{ où } p(x, z) = \mathbf{P}[X_1 = z \mid X_0 = x].$$

Montrer que ces états sont distincts.

3. Si E est fini et de cardinal d . Montrer alors que $n_0 \leq d - 1$ et $\mathbf{P}[T_y \leq d - 1 \mid X_0 = x] > 0$.

5.6 Travail dirigé : Modélisation de l'évolution d'une population

On désire étudier le modèle de population suivant. À l'instant n , la population comporte Z_n individus. Chaque individu i donne naissance à $\zeta_i^{(n)}$ enfant(s) à la génération suivante et meurt. Ainsi la population à l'instant $n + 1$ comporte

$$Z_{n+1} = \sum_{i=1}^{Z_n} \zeta_i^{(n)}$$

individus. Si $Z_n = 0$, alors la population est définitivement éteinte et $Z_{n+1} = 0$. On fait les hypothèses suivantes. La suite $(\zeta_i^{(n)}, i \geq 1, n \geq 1)$ est une suite de variables aléatoires discrètes indépendantes et de même loi, caractérisée par la fonction génératrice : $z \in [0, 1]$,

$$G(z) = \sum_{k=0}^{\infty} p_k z^k \quad \text{où} \quad p_k = \mathbf{P}[\zeta_i^{(n)} = k], \quad k \in \mathbf{N}.$$

On suppose que $0 < p_0 < 1$, $\text{Var}\zeta_i^{(n)} = \sigma^2 < \infty$ et on pose $m = \mathbf{E}[\zeta_i^{(n)}]$. On rappelle que la fonction G est sous ces hypothèses de classe C^2 sur $[0, 1]$.

On suppose également que la population comporte un seul individu à l'instant $n = 0$ (i.e. $Z_0 = 1$).

1. Vérifier que 0 est un état absorbant pour la suite $(Z_n, n \geq 0)$, c'est-à-dire montrer que $\{Z_n = 0\} \subset \{Z_{n+k} = 0\}$ pour tout $k \geq 0$.
2. Montrer que $Z = (Z_n, n \geq 0)$ est une chaîne de Markov sur un espace E que l'on précisera. On ne calculera pas la matrice de transition.
3. On note $T = \inf\{n \geq 0, Z_n = 0\}$ avec la convention $\inf \emptyset = +\infty$. T est le temps d'extinction de la population. Montrer que T est un temps d'arrêt.
4. Montrer que la suite $(W_n = m^{-n} Z_n, n \geq 0)$ est une martingale.
5. Calculer $\mathbf{E}[Z_n]$.
6. Montrer que $\mathbf{E}[Z_n] \geq \mathbf{P}[Z_n \neq 0]$. En déduire que si $m < 1$, alors presque sûrement $Z_n = 0$ pour n assez grand (cela revient à montrer que $\mathbf{P}[T < +\infty] = 1$).
7. Vérifier que $G'(1) = m$.
8. On définit $G_1 = G$ et pour $n \geq 1$, $G_{n+1}(z) = G_n(G(z))$ pour $z \in [0, 1]$. Montrer que pour $z \in [0, 1]$, $\mathbf{E}[z^{Z_n}] = G_n(z)$.
9. Soit $q \in [0, 1]$ tel que $G(q) = q$. Montrer que $(q^{Z_n}, n \geq 0)$ est une martingale.
On suppose dans toutes les questions suivantes que $m > 1$.
10. Montrer que G est convexe sur $[0, 1]$. Tracer l'allure de la fonction $G(z)$ pour $z \in [0, 1]$. Vérifier qu'il existe un unique réel $\eta \in]0, 1[$, tel que $\eta = G(\eta)$.

La martingale $(\eta^{Z_n}, n \geq 0)$ On rappelle le théorème suivant du cours : Si $(Y_n, n \geq 0)$ est une martingale telle que $\sup_{n \geq 0} \mathbf{E}[Y_n^2]$ est fini, alors il existe une variable aléatoire Y telle que $\mathbf{E}[Y^2] < \infty$, et $\lim_{n \rightarrow \infty} \mathbf{E}[(Y_n - Y)^2] = 0$, $\lim_{n \rightarrow \infty} \mathbf{E}[|Y_n - Y|] = 0$ et $(Y_n, n \geq 0)$ converge presque sûrement vers Y .

11. Vérifier que $(X_n = \eta^{Z_n}, n \geq 0)$ est une martingale bornée par 1 qui converge dans L^2 , dans L^1 et presque sûrement vers une variable aléatoire X à valeurs dans $[0, 1]$.
12. Montrer que si H est une variable aléatoire réelle à valeurs dans $[0, 1]$ alors,

$$\mathbf{P}[H \in]0, 1[] > 0 \Rightarrow \mathbf{E}[H^2] < \mathbf{E}[H].$$

13. Calculer $\mathbf{E}[X]$.
14. Vérifier que pour $z \in [0, \eta]$, $|G(z) - \eta| \leq \theta |z - \eta|$ où $\theta < 1$. En déduire $\mathbf{E}[X^2]$.
15. En déduire la valeur de $\mathbf{P}[X \in]0, 1[]$.
16. Déduire de la question précédente la valeur de $\mathbf{P}[T < +\infty]$.
17. Calculer $\mathbf{E}[Z_{n+1}^2 | Z_n]$. En déduire que pour $m \neq 1$,

$$\text{Var}Z_n = \sigma^2 m^{n-1} \frac{m^n - 1}{m - 1}.$$

18. Vérifier que $(W_n, n \geq 0)$ converge presque sûrement vers une variable aléatoire positive W .
19. Montrer que $G_{n+1}(z) = G(G_n(z))$ pour $z \in [0, 1]$. En déduire que $\mathbf{E}[e^{-tZ_{n+1}}] = G(\mathbf{E}[e^{-tZ_n}])$.
20. Montrer que la transformée de Laplace de W définie pour $t \geq 0$ par $\varphi(t) = \mathbf{E}[e^{-tW}]$, est solution de

$$\varphi(mt) = G(\varphi(t)) \quad t \geq 0.$$

21. Montrer que φ est décroissante, continue et que $\lim_{t \rightarrow \infty} \varphi(t) = \mathbf{P}[W = 0]$. Calculer $\mathbf{E}[W]$ et montrer que $\mathbf{P}[W = 0] = \eta$.
22. En utilisant les résultats précédents, montrer que presque sûrement $Z_n = 0$ pour n assez grand ou bien $\lim_{n \rightarrow \infty} m^{-n} Z_n$ existe et est non triviale c'est-à-dire qu'elle est différente de 0 ou $+\infty$.

5.7 Travail dirigé : Algorithme de Hastings-Metropolis

Le but est de simuler des lois compliquées à l'aide de chaînes de Markov simples.

Soit $(X_n, n \geq 0)$ une chaîne de Markov de noyau de transition P sur un espace d'état au plus dénombrable E . On suppose dans un premier temps qu'il existe $\alpha > 0$ et c une mesure de probabilité sur E telle que pour tout $(x, y) \in E^2$,

$$P(x, y) \geq \alpha c(y). \quad (5.4)$$

1) Montrer que si μ, μ' sont deux mesures de probabilités sur E , alors

$$|\mu P - \mu' P| \leq (1 - \alpha) |\mu - \mu'|,$$

où $|\nu| = \sum_{x \in E} |\nu(x)|$ est la norme en variation de la mesure ν . On rappelle que $\nu P(z) = \sum_{x \in E} \nu(x) P(x, z)$.

2) En déduire qu'il existe une unique mesure de probabilité π sur E telle que $\pi P = \pi$. On dit que π est la mesure invariante de la chaîne de Markov.

3) Que se passe-t-il si X_0 a pour loi π ?

4) Montrer que quelle que soit la loi initiale de X_0 , $(X_n, n \geq 0)$ converge en loi vers la mesure invariante. (On démontre en fait que la convergence pour la norme en variation implique la convergence en loi.) Donner une majoration de la vitesse de convergence.

5) Montrer que le résultat reste inchangé si l'on suppose seulement que pour tout $(x, y) \in E^2$,

$$P^l(x, y) \geq \alpha c(y), \quad \text{où } l \geq 1. \quad (5.5)$$

On cherche à simuler une loi ν_β , $\beta > 0$, sur E **fini** de la forme

$$\nu_\beta(x) = C e^{-\beta H(x)}, \quad x \in E,$$

où H est une fonction positive connue et C est la constante de normalisation. Dans le cadre de la mécanique statistique H représente l'énergie du système dans la configuration x . En général la constante de normalisation (appelée fonction de partition) est mal connue. On ne peut donc pas simuler directement une variable aléatoire de loi ν_β .

On suppose donné P un noyau de transition d'une chaîne de Markov sur E vérifiant 5.4 tel que

$$P(x, y) = P(y, x), \quad \text{pour tout } (x, y) \in E^2.$$

On choisit une suite $\alpha(x, y) \in]0, 1]$ où $(x, y) \in E^2$. Soit $(Y_{n,x}, n \geq 1, x \in E)$ et $(Z_{n,x,y}, n \geq 1, (x, y) \in E^2)$ des suites de variables aléatoires indépendantes. Leurs lois sont caractérisées par $\mathbf{P}[Y_{n,x} = y] = P(x, y)$ et $\mathbf{P}[Z_{n,x,y} = 1] = 1 - \mathbf{P}[Z_{n,x,y} = 0] = \alpha(x, y)$. On construit la suite de variable aléatoire $(X_n, n \geq 0)$ de la manière suivante. $X_0 = x_0$, et pour $n \geq 1$, on pose $X_n = x$ si $Y_{n, X_{n-1}} = x$ et si $Z_{n, X_{n-1}, x} = 1$; sinon $X_n = X_{n-1}$.

6) Montrer que $(X_n, n \geq 0)$ est une chaîne de Markov dont on donnera le noyau de transition Q .

7) Montrer que Q vérifie 5.4.

8) Montrer que s'il existe une mesure de probabilité ν sur E telle que pour tout $(x, y) \in E^2$, $\nu(x)Q(x, y) = \nu(y)Q(y, x)$, alors ν est la mesure invariante.

9) Quelle valeur peut-on donner à $(\alpha(x, y), (x, y) \in E^2)$ pour que la mesure invariante soit ν_β . Donner une méthode pour simuler une variable aléatoire de loi ν_β .

10) Que se passe-t-il quand $\beta \rightarrow \infty$?

11) Proposer intuitivement une méthode pour obtenir un minimum de H en utilisant un algorithme de Hastings-Métropolis. On parle alors d'algorithme de recuit-simulé.

5.8 Travail dirigé : Récurrence de la marche aléatoire simple

On considère un processus de Markov homogène en temps X à espace d'état dénombrable, issu de $X_0 = j$ sous \mathbf{P}_j . On définit les v.a.r.

$$U_n = \begin{cases} 1 & \text{si } X_n = j \\ 0 & \text{sinon.} \end{cases}$$

Soit $B \in \mathcal{B}(\{0, 1\}^{\mathbb{N}})$ et $(s_1, \dots, s_{n-1}) \in \{0, 1\}^{n-1}$.

1) Montrer que

$$\mathbf{P}_j[(U_{n+1}, \dots) \in B \mid U_n = 1, U_{n-1} = s_{n-1}, \dots, U_1 = s_1] = \mathbf{P}_j[(U_1, \dots) \in B].$$

On dit que U est un processus de renouvellement.

2) On note $F_n = \{U_n = 1, U_{n+k} = 0, \forall k \geq 1\}$ et $F_0 = \{U_k = 0, \forall k \geq 1\}$. Montrer que

$$1 - \mathbf{P}_j[X_n = j \text{ infiniment souvent}] = \sum_{n \geq 0} \mathbf{P}_j[F_n].$$

3) Montrer que $\mathbf{P}_j[F_n] = \mathbf{P}_j[F_0] \mathbf{P}_j[U_n = 1]$.

4) On note $G = \cup_{n \geq 1} \{X_n = j\}$. Dédurre de 2) et 3) que

$$\mathbf{P}_j[G] < 1 \Rightarrow \begin{cases} \mathbf{P}_j[X_n = j \text{ infiniment souvent}] & = 0 \\ \sum_{n \geq 1} \mathbf{P}_j[X_n = j] & < \infty. \end{cases}$$

5) Dédurre de 2) et 3) que

$$\mathbf{P}_j[G] = 1 \Rightarrow \begin{cases} \mathbf{P}_j[X_n = j \text{ infiniment souvent}] & = 1 \\ \sum_{n \geq 1} \mathbf{P}_j[X_n = j] & = \infty. \end{cases}$$

En déduire que $\mathbf{P}_j[X_n = j \text{ infiniment souvent}] = 1 \Leftrightarrow \sum_{n \geq 1} \mathbf{P}_j[X_n = j] = +\infty$ (on dit que l'état j est récurrent) et que $\mathbf{P}_j[X_n = j \text{ infiniment souvent}] = 0 \Leftrightarrow \sum_{n \geq 1} \mathbf{P}_j[X_n = j] < +\infty$ (on dit que l'état j est transient).

6) Soit la marche aléatoire symétrique sur \mathbf{Z} : $S_0 = 0$, $S_n = Y_1 + \dots + Y_n$, où les v.a.r. $(Y_n, n \geq 1)$ sont i.i.d. et

$$\mathbf{P}[Y_1 = 1] = \mathbf{P}[Y_1 = -1] = 1/2.$$

Calculer $\mathbf{P}[S_{2n} = 0]$. En utilisant la formule de Stirling ($n! \sim \sqrt{2\pi n} n^{n+1/2} e^{-n}$ pour n grand) les questions 5) et 4), montrer que $\mathbf{P}[S_n = 0 \text{ infiniment souvent}] = 1$.

7) On considère la marche aléatoire symétrique sur \mathbf{Z}^2 : $S_0 = 0$, $S_n = X_1 + \dots + X_n$, où les v.a.r. $(X_n, n \geq 1)$ sont i.i.d. et

$$\mathbf{P}[X_1 = (0, 1)] = \mathbf{P}[X_1 = (0, -1)] = \mathbf{P}[X_1 = (1, 0)] = \mathbf{P}[X_1 = (-1, 0)] = 1/4.$$

Calculer $\mathbf{P}[S_{2n} = 0]$. En utilisant la même méthode que pour la question 1, montrer que 0 est récurrent pour la chaîne de Markov S . On pourra utiliser que $(1+x)^n(1+x)^n = (1+x)^{2n}$ pour calculer $\sum_{k=0}^n \frac{(n!)^2}{[k!(n-k)!]^2}$.

8) On considère la marche aléatoire symétrique sur \mathbf{Z}^3 . Montrer que 0 est transient. On utilisera que pour $n \geq 4$:

$$\sum_{i+j+k=n} \left(\frac{n!}{3^n i! j! k!} \right)^2 \leq \left(\sum_{i+j+k=n} \frac{n!}{3^n i! j! k!} \right) \max_{i+j+k=n} \left(\frac{n!}{3^n i! j! k!} \right) \leq \frac{n!}{3^n ([n/3]!)^3}.$$

9) En déduire que la marche aléatoire symétrique en dimension $d \geq 3$ est transiente.

Remarque

Plus généralement, il existe un critère du à Chung et Fuchs (cf Feller tome II p.614) sur la fonction génératrice qui détermine si la marche aléatoire est récurrente ou transiente.

Chapitre 6

Propriété de Markov forte

6.1 Propriété de Markov générale

Soit P la matrice d'une chaîne de Markov sur un espace fini E , nous noterons :

$$\mathbf{E}_x (f(X_0, X_1, \dots, X_n)) = \mathbf{E} (f(X_0^x, X_1^x, \dots, X_n^x)),$$

où $(X_n^x, n \geq 0)$ est une chaîne de Markov de matrice de transition P partant de x à l'instant 0 (i.e. $\mathbf{P}(X_0^x = x) = 1$)¹.

Plus généralement, on peut définir \mathbf{E}_x pour toute fonction (positive ou bornée) de la trajectoire par :

$$\mathbf{E}_x (f(X_k, k \geq 0)) = \mathbf{E} (f(X_k^x, k \geq 0)).$$

On peut alors étendre la propriété de Markov de la façon suivante.

Proposition 6.1.1 *Soit $(X_n, n \geq 0)$ une chaîne de Markov de matrice de transition P par rapport à une filtration $(\mathcal{F}_n, n \geq 0)$, alors :*

$$\mathbf{E} (f(X_{k+n}, k \geq 0) | \mathcal{F}_n) = \phi(X_n)$$

avec :

$$\phi(x) = \mathbf{E}_x (f(X_k, k \geq 0)).$$

On écrit souvent cette relation sous la forme :

$$\mathbf{E} (f(X_{k+n}, k \geq 0) | \mathcal{F}_n) = \mathbf{E}_{X_n} (f(X_k, k \geq 0)).$$

Démonstration : Nous ne donnerons la démonstration de ce résultat que dans un cas très particulier. La généralisation n'est pas très difficile mais repose sur des arguments techniques particulièrement indigestes.

Nous supposerons f de la forme :

$$f(X_p, p \geq 0) = f(X_1),$$

avec un abus évident de langage. On a, alors, grâce à la définition de la propriété de Markov :

$$\mathbf{E} (f(X_{n+1}) | \mathcal{F}_n) = (Pf)(X_n).$$

¹Notez que cette définition ne dépend pas du choix de la suite X^x mais uniquement de sa loi qui est déterminée par P et le point initial x .

De plus :

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

puisque la loi de X_1 , pour une chaîne de Markov de probabilité de transition P issue de x en 0 est donnée par $(P(x, y), y \in E)$.

Le résultat est donc clair pour ce type de fonction. ■

Nous allons maintenant aborder la notion de propriété de Markov forte. Il s'agit d'étendre la propriété que nous venons d'énoncer non plus à des temps fixe n mais à des *temps d'arrêt finis*.

6.2 Introduction à la propriété de Markov forte

Nous allons commencer par illustrer cette propriété importante dans le cas d'une suite de variables aléatoires indépendantes de même loi $(U_n, n \geq 1)$.

Nous noterons $(\mathcal{F}_n, n \geq 1)$ la filtration définit par :

$$\mathcal{F}_n = \sigma(U_1, \dots, U_n),$$

i.e. la filtration naturelle de la suite. Soit S un temps d'arrêt par rapport à la filtration $(\mathcal{F}_n, n \geq 1)$. Cela signifie que pour tout entier n , $\{S \leq n\}$ est un ensemble de \mathcal{F}_n ou de façon équivalente que $\{S = n\}$ est un ensemble de \mathcal{F}_n . Nous supposons, de plus, que $\mathbf{P}(S < +\infty) = 1$.

Notre but est d'identifier la loi de la suite $(U_{S+1}, \dots, U_{S+n})$. Pour cela nous devons calculer, pour toute fonction positive :

$$\mathbf{E}(f(U_{S+1}, \dots, U_{S+n})).$$

Or, nous avons :

$$\begin{aligned} \mathbf{E}(f(U_{S+1}, \dots, U_{S+n})) &= \mathbf{E}\left(\sum_{k \geq 1} \mathbf{1}_{\{S = k\}} f(U_{S+1}, \dots, U_{S+n})\right) \\ &= \sum_{k \geq 1} \mathbf{E}\left(\mathbf{1}_{\{S = k\}} f(U_{k+1}, \dots, U_{k+n})\right). \end{aligned}$$

On note alors que $\{S = k\}$ est dans \mathcal{F}_k et que $f(U_{k+1}, \dots, U_{k+n})$ est une variable aléatoire indépendante de \mathcal{F}_k et l'on en déduit que :

$$\begin{aligned} \mathbf{E}(f(U_{S+1}, \dots, U_{S+n})) &= \sum_{k \geq 1} \mathbf{P}(S = k) \mathbf{E}(f(U_{k+1}, \dots, U_{k+n})), \\ &= \left[\sum_{k \geq 1} \mathbf{P}(S = k)\right] \mathbf{E}(f(U_1, \dots, U_n)), \\ &= \mathbf{E}(f(U_1, \dots, U_n)). \end{aligned}$$

Ceci prouve que $(U_{S+1}, \dots, U_{S+n})$ a la même loi que (U_1, \dots, U_n) et l'on peut en déduire que la suite $(U_{S+k}, k \geq 1)$ a la même loi que $(U_k, k \geq 1)$, c'est à dire est une suite de variables aléatoires indépendantes de même loi.

6.3 Tribu des événements antérieurs à un temps d'arrêt

Définition 6.3.1 Soit $(\mathcal{F}_n, n \geq 0)$ une filtration, S un temps d'arrêt par rapport à cette filtration on note \mathcal{F}_S la tribu des événements antérieurs à S définie par :

$$\mathcal{F}_S = \{A \in \mathcal{A}, \text{ pour tout } n, A \cap \{S = n\} \in \mathcal{F}_n\}.$$

Remarque 6.3.2 On démontre sans difficulté que S et X_S sont des variables aléatoires \mathcal{F}_S mesurable, si $(X_n, n \geq 0)$ est un processus adapté.

De même $\{T \leq S\}, \{X_S \in A\}$ sont dans \mathcal{F}_S , mais $\{X_{S+1} \in A\}$ n'est pas dans \mathcal{F}_S et plus généralement $\{X_t \in A\}$, sauf si $S \geq t$?

Remarque 6.3.3 Soit S un temps d'arrêt fini alors :

$$\mathbf{E}(X|\mathcal{F}_S) = \sum_{n \geq 0} \mathbf{1}_{\{S = n\}} \mathbf{E}(X|\mathcal{F}_n). \quad (6.1)$$

Ceci se prouve sans difficulté en posant :

$$\bar{X} = \sum_{n \geq 0} \mathbf{1}_{\{S = n\}} \mathbf{E}(X|\mathcal{F}_n),$$

et en remarquant que si A est un élément de \mathcal{F}_S , on a :

$$\mathbf{E}(\bar{X}\mathbf{1}_A) = \sum_{n \geq 0} \mathbf{E}(\mathbf{1}_{\{S=n\} \cap A} \mathbf{E}(X|\mathcal{F}_n)).$$

Comme $\{S = n\} \cap A$ est un élément de \mathcal{F}_n , on en déduit que :

$$\mathbf{E}(\bar{X}\mathbf{1}_A) = \sum_{n \geq 0} \mathbf{E}(\mathbf{E}(X\mathbf{1}_{\{S=n\} \cap A}|\mathcal{F}_n)) = \sum_{n \geq 0} \mathbf{E}(X\mathbf{1}_{\{S=n\} \cap A}) = \mathbf{E}(X\mathbf{1}_A).$$

Le théorème 3.2.1, définissant l'espérance conditionnelle permet alors d'affirmer que :

$$\bar{X} = \mathbf{E}(X|\mathcal{F}_S).$$

Retour au cas d'une suite indépendante Revenons au cas d'une suite $(U_n, n \geq 1)$ de variables aléatoires indépendantes de même loi. Nous notons dans ce cas $\mathcal{F}_n = \sigma(U_1, \dots, U_n)$ et nous supposons que S est un temps d'arrêt, fini presque sûrement, par rapport à cette tribu.

Nous cherchons, tout d'abord, à calculer :

$$\mathbf{E}(f(U_{S+1}, \dots, U_{S+n})|\mathcal{F}_S).$$

Pour cela nous allons utiliser l'équation 6.1 :

$$\begin{aligned} \mathbf{E}(f(U_{S+1}, \dots, U_{S+p})|\mathcal{F}_S) &= \sum_{n \geq 0} \mathbf{E}(\mathbf{1}_{\{S = n\}} f(U_{S+1}, \dots, U_{S+p})|\mathcal{F}_n) \\ &= \sum_{n \geq 0} \mathbf{E}(\mathbf{1}_{\{S = n\}} f(U_{n+1}, \dots, U_{n+p})|\mathcal{F}_n). \end{aligned}$$

Comme $(U_{n+1}, \dots, U_{n+p})$ est indépendante de \mathcal{F}_n et $\{S = n\}$ est dans \mathcal{F}_n on en déduit que :

$$\begin{aligned} \mathbf{E}(f(U_{S+1}, \dots, U_{S+p})|\mathcal{F}_S) &= \sum_{n \geq 0} \mathbf{1}_{\{S = n\}} \mathbf{E}(f(U_{n+1}, \dots, U_{n+p})) \\ &= \sum_{n \geq 0} \mathbf{1}_{\{S = n\}} \mathbf{E}(f(U_1, \dots, U_p)) \\ &= \mathbf{E}(f(U_1, \dots, U_p)). \end{aligned}$$

En prenant l'espérance des deux membres on retrouve le fait déjà établi que la loi de $(U_{S+k}, k \geq 1)$ est identique à celle de $(U_k, k \geq 1)$. Nous allons voir que ce résultat implique de plus que $(U_{S+k}, k \geq 1)$ est une suite indépendante de la tribu \mathcal{F}_S . Pour se convaincre de ceci, il suffit de noter le résultat suivant.

Proposition 6.3.4 Soit X une variable aléatoire et \mathcal{B} une tribu. Supposons que pour toute fonction positive ou bornée f on a :

$$\mathbf{E}(f(X)|\mathcal{B}) = \text{une constante} = \mathbf{E}(f(X)).$$

Alors X est indépendante de \mathcal{B} .

Démonstration : Soit Z une variable aléatoire \mathcal{B} -mesurable et g une fonction bornée, alors :

$$\mathbf{E}(f(X)g(Z)) = \mathbf{E}(\mathbf{E}(f(X)g(Z)|\mathcal{B})) = \mathbf{E}(g(Z)\mathbf{E}(f(X)|\mathcal{B})).$$

Mais $\mathbf{E}(f(X)|\mathcal{B}) = C$ et en prenant l'espérance des deux membres on a forcément $C = \mathbf{E}(f(X))$. On obtient donc :

$$\mathbf{E}(f(X)g(Z)) = \mathbf{E}(g(Z))\mathbf{E}(f(X)).$$

X est donc indépendante de toute variable aléatoire Z \mathcal{B} -mesurable. Elle est donc, par définition, indépendante de \mathcal{B} . ■

Or, nous avons vu que, pour toute fonction f et pour tout p :

$$\mathbf{E}(f(\mathbf{U}_{S+1}, \dots, \mathbf{U}_{S+p})|\mathcal{F}_S) = \mathbf{E}(f(\mathbf{U}_1, \dots, \mathbf{U}_p)),$$

On déduit, du résultat précédent que, pour tout p , $(\mathbf{U}_{S+1}, \dots, \mathbf{U}_{S+p})$ est indépendant de \mathcal{F}_S , puis par un argument technique que $(\mathbf{U}_{S+k}, k \geq 1)$ est indépendante de \mathcal{F}_S .

Un exemple d'utilisation : jeu de pile ou face. Soit $(\mathbf{U}_n, n \geq 1)$ une suite de variables aléatoires indépendantes telles que $\mathbf{P}(\mathbf{U}_1 = 1) = p$ et $\mathbf{P}(\mathbf{U}_1 = 0) = 1 - p$. Soit τ_1 le temps de première apparition de pile :

$$\tau_1 = \inf\{n \geq 1, \mathbf{U}_n = 1\}.$$

Il est bien connu que la loi de τ_1 est une loi géométrique de paramètre p , i.e. :

$$\mathbf{P}(\tau_1 = k) = p(1 - p)^{k-1}, k \geq 1.$$

De même τ_2 le temps de première apparition de face :

$$\tau_2 = \inf\{n \geq 1, \mathbf{U}_n = 0\},$$

suit une loi géométrique de paramètre $1 - p$:

$$\mathbf{P}(\tau_2 = k) = (1 - p)p^{k-1}, k \geq 1.$$

Quelle est la loi de l'instant de première apparition de la séquence 10, τ_{10} ? Pour répondre à cette question il suffit de remarquer que τ_{01} est la somme du temps d'atteinte de 1, τ_1 et du temps d'atteinte de 0 pour la suite $(\mathbf{U}_{\tau_1+k}, k \geq 1)$, τ'_2 . Cette suite est indépendante de \mathcal{F}_{τ_1} donc de τ_1 et a même loi que $(\mathbf{U}_k, k \geq 1)$. On obtient donc, sans aucun calcul, que :

$$\tau_{10} = \tau_1 + \tau'_2,$$

τ_1 suivant une loi géométrique de paramètre p et τ'_2 étant indépendante de τ_1 suivant une loi géométrique de paramètre $1 - p$.

6.4 Propriété de Markov forte : première approche

Nous allons généraliser les idées précédentes au cas des chaînes de Markov.

Proposition 6.4.1 Soit $(X_n, n \geq 1)$ une chaîne de Markov par rapport à une filtration $(\mathcal{F}_n, n \geq 0)$ de matrice de transition P , soit S un temps d'arrêt par rapport à la même filtration, fini presque sûrement. Alors pour toute fonction f bornée :

$$\mathbf{E}(f(X_{S+k}, k \geq 0) | \mathcal{F}_S) = \mathbf{E}_{X_S}(f(X_k, k \geq 0)).$$

et en particulier :

- $(X_{S+n}, n \geq 0)$ est une chaîne de Markov de matrice de transition P .
- Si X_S est une variable aléatoire constante, $(X_{S+n}, n \geq 0)$ est indépendante de la tribu \mathcal{F}_S .

Démonstration : Nous allons procéder comme dans le cas précédent et évaluer l'espérance conditionnelle $\mathbf{E}(f(X_S, X_{S+1}, \dots, X_{S+p}) | \mathcal{F}_S)$ pour cela notons que :

$$\begin{aligned} \mathbf{E}(f(X_S, X_{S+1}, \dots, X_{S+p}) | \mathcal{F}_S) &= \sum_{n \geq 0} \mathbf{1}_{\{S = n\}} \mathbf{E}(f(X_S, X_{S+1}, \dots, X_{S+p}) | \mathcal{F}_n), \\ &= \sum_{n \geq 0} \mathbf{E}\left(\mathbf{1}_{\{S = n\}} f(X_n, X_{n+1}, \dots, X_{n+p}) | \mathcal{F}_n\right), \\ &= \sum_{n \geq 0} \mathbf{1}_{\{S = n\}} \mathbf{E}(f(X_n, X_{n+1}, \dots, X_{n+p}) | \mathcal{F}_n), \\ &= \sum_{n \geq 0} \mathbf{1}_{\{S = n\}} \mathbf{E}_{X_n}(f(X_0, X_1, \dots, X_p)). \end{aligned}$$

La dernière égalité s'obtient en utilisant la propriété de Markov étendue à des fonctions générales (voir proposition 6.1.1). On peut réécrire la dernière égalité sous la forme :

$$\mathbf{E}(f(X_S, X_{S+1}, \dots, X_{S+p}) | \mathcal{F}_S) = \mathbf{E}_{X_S}(f(X_0, X_1, \dots, X_p)).$$

Par un argument technique on étend le résultat aux fonctions f de toute la trajectoire du processus :

$$\mathbf{E}(f(X_{S+k}, k \geq 0) | \mathcal{F}_S) = \mathbf{E}_{X_S}(f(X_k, k \geq 0)).$$

En prenant l'espérance des deux membres, on en déduit que la loi de $(X_{S+k}, k \geq 0)$ est celle d'une chaîne de Markov de probabilité de transition P issue de X_S à l'instant 0. Dans le cas où X_S est une constante, l'espérance conditionnelle se réduit à une constante et le lemme 6.3.4 permet alors d'affirmer que $(X_{S+k}, k \geq 0)$ est indépendante de \mathcal{F}_S . ■

Application : temps d'atteinte d'une marche aléatoire Soit $(U_k, k \geq 1)$ une suite de variables aléatoires indépendantes de même loi, telle que $\mathbf{P}(U_1 = \pm 1) = 1/2$. Soit τ_a le temps d'atteinte du niveau a :

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

pour a un entier strictement positif.

Comme $S_{\tau_1} = 1$, la suite $(S_{\tau_1+k}, k \geq 0)$ est indépendante de \mathcal{F}_{τ_1} et la loi de cette suite est celle d'une marche aléatoire partant de 1 à l'instant 0. De plus τ_2 est la somme de τ_1 et de τ'_1 le temps d'atteinte de 1 pour la marche aléatoire issue de 0 en 0, $(S_{\tau_1+k} - 1, k \geq 0)$. On en déduit donc, sans aucun calcul que :

$$\tau_2 = \tau_1 + \tau'_1,$$

τ_1 et τ'_1 étant deux variables aléatoires indépendantes, la loi de τ'_1 étant identique à celle de τ_1 . On généralise sans difficulté cette relation pour a entier arbitraire.

$$\tau_a = \tau_1^1 + \tau_1^2 + \dots + \tau_1^a,$$

les $(\tau_1^1, \dots, \tau_1^a)$ étant indépendants, suivant tous la même loi que τ_1 . On pourra montrer, en exercice, que $(\tau_a, a \geq 0)$ est une marche aléatoire.

6.5 Exercices

Exercice 43 Un petit restaurateur qui ne possède que trois tables, remarque que si son restaurant est vide alors la loi d'arrivée du premier client est une loi exponentielle de paramètre 1. Le temps de service d'un client est une loi exponentielle de paramètre 1. Mais dès qu'un client est installé, sa présence attire d'autres clients, et une voire deux tables en même temps, s'il en reste deux, peuvent se remplir. Les temps d'arrivée suivent toujours des lois exponentielles de paramètre 1. Enfin si toutes les tables sont occupées, le restaurateur presse le service et une ou deux tables en même temps peuvent se libérer suivant des temps exponentiels de paramètre 1.

On modélise l'évolution du nombre de tables occupées à l'aide d'une chaîne de Markov X à temps continu sur l'espace d'état $E = \{0, 1, 2, 3\}$. On note P_t le semi-groupe de transition du processus X et G le générateur infinitésimal ($P_t = e^{tG}$).

1. Montrer que G s'écrit

$$G = \begin{pmatrix} -1 & 1 & 0 & 0 \\ 1 & -3 & 1 & 1 \\ 0 & 1 & -2 & 1 \\ 0 & 1 & 1 & -2 \end{pmatrix}.$$

2. On note $T = \inf\{t \geq 0; X_t \neq X_0\}$, le premier instant où la chaîne change d'état. Calculer $\mathbf{E}[T \mid X_0 = i]$ pour $i \in \{0, 1, 2, 3\}$.

3. On note $R = \inf\{t \geq 0; X_t = 3\}$, le premier instant où les trois tables sont occupées. On désire estimer $\mathbf{E}[R \mid X_0 = 0]$, le temps moyen pour que le restaurant initialement vide ait ses trois tables occupées. On note $b_i = \mathbf{E}[R \mid X_0 = i]$. Montrer que $b_3 = 0$ et que pour tout $i \in \{0, 1, 2\}$

$$b_i = \mathbf{E}[T \mid X_0 = i] + \sum_{j \neq i} b_j \mathbf{P}[X_T = j \mid X_0 = i].$$

4. En déduire la valeur de b_0 .

5. Montrer que $G = Q^{-1}DQ$, où D est une matrice diagonale.

6. Calculer le semi-groupe de transition $P_t = e^{tG}$.

7. Calculer $\lim_{t \rightarrow \infty} P_t$.

8. En déduire la probabilité invariante π_0 .

9. Quel est la limite quand $t \rightarrow \infty$ du taux moyen de remplissage du restaurant ?

Exercice 44 On note G le générateur infinitésimal d'une chaîne de Markov X à temps continu sur un espace d'état dénombrable E . On peut considérer G comme une matrice $G = (g(x, y), (x, y) \in E^2)$. On note P_t le semi-groupe de transition du processus X ($P_t = e^{tG}$). On rappelle que si f est une fonction mesurable positive définie sur E , $P_t f(x) = \sum_{y \in E} f(y) p_t(x, y)$ où

$$p_t(x, y) = \mathbf{P}[X_t = y \mid X_0 = x].$$

On suppose que $\|G\| = \sup_{x \in E} \sum_{y \in E} |g(x, y)| < \infty$.

1. Montrer que $\|G^n\| \leq \|G\|^n$ et que $\|e^{tG}\| \leq e^{t\|G\|}$. En déduire que

$$P_t = I + tG + t\varepsilon_t,$$

où $\lim_{t \rightarrow 0} \|\varepsilon_t\| = 0$.

2. Faire un développement de $p_t(x, y)$ à l'ordre $o(t)$.

3. Montrer que pour tout $(x, y) \in E^2$ tel que $x \neq y$, alors $g(x, y) \geq 0$. Montrer que pour tout $x \in E$,

$$\sum_{y \in E} g(x, y) = 0.$$

4. Soit π une probabilité sur E . L'expression $\pi P_t = \pi$ signifie que $\sum_{x \in E} \pi(x) p_t(x, y) = \pi(y)$ pour tout $y \in E$. De même $\pi G = 0$ signifie que $\sum_{x \in E} \pi(x) g(x, y) = 0$. Montrer que

$$\pi P_t = \pi, \quad \forall t \geq 0 \iff \pi G = 0.$$

Si l'une des conditions ci-dessus est vérifiée, on dit que π est la probabilité invariante du processus.

Exercice 45 On modélise l'évolution d'une population, en fait du cardinal de la population, à l'aide d'une chaîne de Markov X à temps continu sur \mathbf{N} de la manière suivante.

$$\mathbf{P}[X_{t+h} = n + m \mid X_t = n] = \begin{cases} \lambda_n h + o(h) & \text{si } m = 1 \\ \mu_n h + o(h) & \text{si } m = -1 \\ o(h) & \text{si } |m| > 1 \end{cases}$$

On suppose $\lambda_0 \geq 0$, $\mu_0 = 0$ et pour tout $n \geq 1$, $\lambda_n > 0$, $\mu_n > 0$.

1. Interpréter les valeurs λ_n et μ_n .
2. Déterminer le générateur infinitésimal G de la chaîne de Markov ($P_t = e^{tG}$).
3. Donner la loi de $T(t) = \inf\{s \geq t; X_s \neq X_t\}$ conditionnellement à $X(t) = n$.
4. Calculer $\mathbf{P}[X_{T(t)} = n + m \mid X_t = n]$.
5. En utilisant la question 4. de l'exercice 2, donner une condition suffisante pour l'existence d'une probabilité invariante π . Calculer alors cette probabilité.
6. On considère le modèle suivant pour l'évolution de la population. À t fixé, deux phénomènes sont en concurrence. Soit un immigrant arrive (à un temps exponentiel de paramètre λ), soit un membre de la population meurt (à un temps exponentiel de paramètre μ). Calculer $\mathbf{P}[X_{T(t)} = n + m \mid X_t = n]$. En déduire les valeurs de λ_n et μ_n .
7. Montrer que la probabilité invariante est une loi de Poisson de paramètre $\rho = \lambda/\mu$.
8. On suppose que X_0 est distribué suivant la probabilité invariante π . Calculer le nombre moyen de personnes dans la population à l'instant $t > 0$.

Exercice 46 On reprend les résultats et les notations de l'exercice 2. On note

$$G = \begin{pmatrix} -\mu & \mu \\ \lambda & -\lambda \end{pmatrix},$$

où $\mu \geq 0$, $\lambda \geq 0$, le générateur infinitésimal le plus général sur l'espace à deux états $E = \{a, b\}$.

1. Sous quelles conditions la chaîne de Markov à temps continu associé à G est-elle constante (et donc déterministe) ?
2. On suppose $\mu + \lambda > 0$. Écrire $G = Q^{-1}DQ$ où D est une matrice diagonale. Montrer que

$$P_t = \frac{1}{\mu + \lambda} \begin{pmatrix} \lambda + \mu h(t) & \mu(1 - h(t)) \\ \lambda(1 - h(t)) & \mu + \lambda h(t) \end{pmatrix},$$

où h est une fonction que l'on calculera explicitement.

3. Calculer $\mathbf{P}[X_t = b \mid X_{2t} = a, X_0 = a]$.
4. Calculer la probabilité invariante π ($\pi P_t = \pi$).
5. Calculer la limite de P_t quand t tend vers ∞ .
6. En déduire que X_t converge en loi quand t tend vers ∞ . Donner la loi limite.
7. On note

$$P = \begin{pmatrix} \alpha & 1 - \alpha \\ 1 - \alpha & \alpha \end{pmatrix},$$

où $\alpha \in [0, 1]$. Montrer qu'il existe un semi-groupe de transition continu en temps $(P_t, t \geq 0)$ tel que $P_1 = P$ si et seulement si $1 \geq \alpha > 1/2$. On distinguera le cas $\alpha = 1$.

8. Montrer qu'alors le semi groupe $(P_t, t \geq 0)$ est unique et le calculer à l'aide de α .

6.6 Contrôle : Loi du supremum d'une marche aléatoire

Le but de ce problème est d'étudier la loi du maximum global de certaines marches aléatoires sur \mathbf{Z} . On montrera que sous certaines hypothèses, la loi du maximum est (presque) une loi géométrique.

Soit $(X_i, i \geq 1)$, une suite de variables aléatoires indépendantes, de même loi à valeurs dans \mathbf{Z} telles que $\mathbf{P}(X_i = 1) = p > 0$ et $\mathbf{P}(X_i \in \{1, 0, -1, -2, \dots\}) = 1$ (donc on a $\mathbf{P}(X_i \geq 2) = 0$). On suppose de plus que les variables aléatoires X_i sont intégrables ($\mathbf{E}[|X_i|] < \infty$) et que $\mathbf{E}[X_i] = m < 0$. On pose $S_0 = 0$ et pour $n \geq 1$, $S_n = \sum_{i=1}^n X_i$. Le maximum global de la marche aléatoire $(S_n, n \geq 0)$ est $S = \sup_{n \geq 0} S_n$.

Remarques préliminaires

1. Rappeler en quel sens la suite $(\frac{S_n}{n}, n \geq 1)$ converge et donner sa limite. En déduire $\mathbf{P}(S < +\infty)$.
2. On pose $\varphi(\lambda) = \mathbf{E}[e^{\lambda X_i}]$ pour $\lambda \geq 0$. Expliquer pourquoi $\varphi(\lambda)$ est bien définie et est une fonction continue.
3. Montrer que $(M_n = \varphi(\lambda)^{-n} e^{\lambda S_n}, n \geq 1)$ est une martingale.
4. On admet que φ est de classe C^∞ sur $]0, +\infty[$, et que les dérivées d'ordre k de φ sont données par $\frac{d^k}{d\lambda^k} \varphi(\lambda) = \mathbf{E}[X_i^k e^{\lambda X_i}]$. En déduire que φ est convexe. Tracer l'allure du graphe de φ . En déduire qu'il existe un unique $\lambda_0 \in]0, \infty[$ tel que $\varphi(\lambda_0) = 1$.
5. Vérifier que $\varphi'(\lambda_0) > 0$.

Changement de probabilité

1. On pose $p_k = e^{\lambda_0 k} \mathbf{P}(X_i = k)$. Vérifier que $\sum_{k \in \mathbf{Z}} p_k = 1$.
2. Soit $(Y_i, i \geq 1)$ une suite de variables aléatoires indépendantes et de même loi à valeurs dans \mathbf{Z} telles que pour $k \in \mathbf{Z}$, $\mathbf{P}(Y_i = k) = p_k$. On pose $Z_0 = 0$ et $Z_n = \sum_{i=1}^n Y_i$ pour $n \geq 1$. Montrer que pour toute fonction f bornée ou positive (mesurable), on a $\mathbf{E}[f(Y_i) e^{-\lambda_0 Y_i}] = \mathbf{E}[f(X_i)]$. En déduire que pour tout $u_1, \dots, u_n \in \mathbf{R}$, on a

$$\mathbf{E}\left[e^{i \sum_{k=1}^n u_k Y_k - \lambda_0 Z_n}\right] = \mathbf{E}\left[e^{i \sum_{k=1}^n u_k X_k}\right].$$

On admet que cela implique que pour toute fonction bornée ou positive (mesurable) f , on a $\mathbf{E}[f(Y_1, \dots, Y_n) e^{-\lambda_0 Z_n}] = \mathbf{E}[f(X_1, \dots, X_n)]$.

3. Vérifier que pour toute fonction f bornée ou positive, on a $\mathbf{E}[f(Y_i)] = \mathbf{E}[f(X_i) e^{\lambda_0 X_i}]$. En déduire que Y_i est intégrable et, en utilisant la question 5, que $\mathbf{E}[Y_i] > 0$.

Temps d'arrêt et martingale

1. Montrer que $(N_n = e^{-\lambda_0 Z_n}, n \geq 1)$ est une martingale. On précisera la filtration.
2. On introduit les premiers temps d'atteinte du niveau $k \geq 0$ pour les marches $(S_n, n \geq 0)$ et $(Z_n, n \geq 0)$:

$$\tau_k = \inf\{n \geq 0; S_n \geq k\} \quad \text{et} \quad \rho_k = \inf\{n \geq 0; Z_n \geq k\}.$$

Par convention, on pose $\inf \emptyset = \infty$. Montrer que ρ_k et τ_k sont des temps d'arrêt par rapport à des filtrations que l'on précisera.

3. Quel est le comportement asymptotique de la suite $(Z_n, n \geq 0)$ quand $n \rightarrow \infty$? En déduire $\mathbf{P}(\rho_k < \infty)$.
4. Si $\rho_k(\omega) = i$, que vaut $Z_i(\omega)$?
5. En utilisant la remarque 5.3.3. du polycopié, calculer

$$\mathbf{E} [\mathbf{1}_{\rho_k \leq n} e^{-\lambda_0 Z_n} \mid \mathcal{F}_{\rho_k}].$$

6. Déduire des questions précédentes que $\mathbf{P}(\tau_k \leq n) = \mathbf{E} [\mathbf{1}_{\rho_k \leq n} e^{-\lambda_0 Z_n}]$.
7. Montrer que $\mathbf{P}(\tau_k \leq n) = e^{-\lambda_0 k} \mathbf{P}(\rho_k \leq n)$. En déduire $\mathbf{P}(\tau_k < \infty)$.
8. Montrer que $S + 1$ est une variable aléatoire géométrique : $\mathbf{P}(S + 1 = k) = \alpha(1 - \alpha)^{k-1}$ pour $k \in \mathbf{N}^*$. On explicitera α .

Le cas particulier de la marche aléatoire simple

On suppose dorénavant que $\mathbf{P}(X_i = 1) = p$ et $\mathbf{P}(X_i = -1) = q = 1 - p$ avec $1 > q > p > 0$. On désire calculer la probabilité pour que la marche aléatoire $(S_n, n \geq 0)$ ne revienne jamais en 0.

1. Calculer $e^{-\lambda_0}$, et vérifier que $e^{-\lambda_0} = \frac{p}{q}$.
2. Montrer, en utilisant la propriété de Markov, que $\mathbf{P}(S_n < 0; \forall n \geq 1) = q\mathbf{P}(S = 0)$.
3. Déduire des questions précédentes la valeur de $\mathbf{P}(S_n < 0; \forall n \geq 1)$.
4. Rappeler la valeur de $\mathbf{P}(S_n < 0; \forall n \geq 1)$ quand $p = q = 1/2$ (vu en cours ou en travail dirigé). En déduire la valeur de $\mathbf{P}(S_n < 0; \forall n \geq 1)$ pour $p = 1 - q \in]0, 1[$.

Chapitre 7

Options européennes dans le modèle de Cox-Ross-Rubinstein

Une option est un produit financier qui donne le droit à son acheteur d'effectuer une opération portant sur un sous-jacent (qui peut être une action, une obligation, une cargaison de pétrole) avant une date appelée date d'échéance.

Exemples :

- Le détenteur d'une option européenne d'achat (Call en anglais) d'échéance $T = 6$ mois qui porte sur une action France Telecom pour un prix d'exercice (Strike en anglais) de $K = 50$ euros pourra acheter dans 6 mois une action France Telecom au prix de 50 euros. Bien entendu, il n'effectuera cette opération que si elle lui est profitable i.e. si $S_T \geq K$ où $(S_t)_t$ désigne le cours de l'action France Telecom à l'instant t ; en revendant l'action sur le marché, il réalisera alors le profit $(S_T - K)$.

Dans le cas où $S_T < K$, il n'exercera pas l'option puisqu'il peut acheter l'action sur le marché au cours S_T .

On fait la synthèse des deux cas en remarquant que le profit retiré de l'option par l'acheteur est $h = (S_T - K)^+$.

C'est le vendeur de l'option qui lui fournit la richesse h en T . Bien entendu, en contrepartie, il va demander à l'acheteur de verser une prime P à l'instant initial $t = 0$. La question du pricing de l'option c'est-à-dire du calcul de la prime P est intimement liée à celle de la couverture c'est-à-dire de la gestion du risque induit par la nécessité de fournir la richesse h à l'acheteur en T . Comme nous allons le voir il s'agit pour le vendeur d'investir la richesse P à l'instant initial et de la gérer "au mieux" jusqu'à l'échéance pour être en mesure de fournir h .

- De façon symétrique, une option européenne de vente (Put en anglais) d'échéance T , de prix d'exercice K portant sur le sous-jacent $(S_t)_t$ donne le droit à son détenteur de vendre une unité du sous-jacent au prix K à l'échéance T . Elle lui rapporte donc $h = (K - S_T)^+$.

Bien sûr, à l'instant initial il y a beaucoup de scénarii d'évolution possibles pour le cours de l'actif sous-jacent. Il est donc naturel de modéliser $(S_t)_{t \in [0, T]}$ comme un processus aléatoire à valeurs dans \mathbf{R}_+^* i.e. comme une famille de variables aléatoires positives. Le modèle le plus classique est celui de Black-Scholes

$$S_t = S_0 \exp(\sigma B_t + \mu t) \text{ où } B_t \text{ est un mouvement brownien.}$$

Ici, nous nous placerons dans le modèle plus simple de Cox-Ross-Rubinstein qui peut être vu comme une approximation en temps discret du modèle de Black-Scholes pour traiter le problème du pricing et de la couverture (hedging en anglais).

Enfin nous nous intéresserons seulement aux options européennes qui ne peuvent être exercées qu'à l'échéance. D'autres options, appelées américaines, peuvent être exercées par l'acheteur à tout instant jusqu'à l'échéance. Elles offrent plus de droits et déterminer leur prix et leur couverture est plus compliqué que dans le cas européen.

7.1 Le modèle

7.1.1 description

Le modèle comporte N périodes de temps.

On considère qu'il est possible de placer ou d'emprunter de l'argent au taux d'intérêt discret r qu'il est naturel de supposer positif d'un point de vue financier. Le montant obtenu (resp. à rembourser) à l'instant n en plaçant (resp. empruntant) un euro à l'instant initial est

$$S_n^0 = (1 + r)^n.$$

On se donne également un sous-jacent risqué (celui sur lequel vont porter les options) et pour $0 \leq n \leq N$, on note S_n la valeur de cet actif risqué à l'instant n . L'évolution de cet actif dépend de deux paramètres $b > a > -1$: entre l'instant n et l'instant $n + 1$ l'augmentation relative du cours peut prendre la valeur b (hypothèse haute : marché haussier) $S_{n+1} - S_n = bS_n$ ou bien la valeur a (hypothèse basse : marché baissier) $S_{n+1} - S_n = aS_n$. On suppose que tous les scénarii d'évolution possibles entre l'instant initial et l'instant terminal se produisent avec probabilité strictement positive.

On formalise cette dynamique de la manière suivante. On se donne sur un espace de probabilité $(\Omega, \mathcal{A}, \mathbf{P})$ des variables aléatoires T_1, \dots, T_N à valeurs dans $\{1 + a, 1 + b\}$ t.q.

$$\forall (x_1, \dots, x_N) \in \{1 + a, 1 + b\}^N, \mathbf{P}(T_1 = x_1, \dots, T_N = x_N) > 0. \quad (7.1)$$

La valeur initiale $S_0 > 0$ de l'actif risqué est supposée déterministe (c'est la valeur connue sur le marché du sous-jacent que l'on modélise). Ensuite S_n s'obtient par la relation de récurrence

$$\forall 0 \leq n \leq N - 1, S_{n+1} = S_n T_{n+1}.$$

On en déduit facilement que

$$\forall 0 \leq n \leq N, S_n = S_0 T_1 \dots T_n = S_0 \prod_{m=1}^n T_m.$$

On introduit la filtration $(\mathcal{F}_n)_{0 \leq n \leq N}$ définie par

$$\mathcal{F}_n = \begin{cases} \{\emptyset, \Omega\} & \text{si } n = 0 \\ \sigma(T_1, \dots, T_n) & \text{sinon.} \end{cases}$$

Notons que $\forall 0 \leq n \leq N, \mathcal{F}_n = \sigma(S_0, \dots, S_n)$.

Une option européenne est modélisée par une variable aléatoire h \mathcal{F}_N -mesurable positive qui représente le profit obtenu à l'instant N par son détenteur en l'exerçant.

Exemples :

- Pour Put européen standard (vanilla) de prix d'exercice K , $h = (K - S_N)^+$.
- Pour un Call européen sur-moyenne, $h = (\frac{1}{N+1} \sum_{n=0}^N S_n - K)^+$.
- Pour un Call sur maximum avec Strike flottant, $h = \max_{0 \leq n \leq N} S_n - S_N$.

Comme h est $\sigma(T_1, \dots, T_N)$ mesurable, dans l'esprit de la Proposition 3.1.2,

$$\exists f : \{1 + a, 1 + b\}^N \rightarrow \mathbf{R}_+ \text{ t.q. } h = f(T_1, \dots, T_N). \quad (7.2)$$

Dans l'exemple du Put européen standard $f(x_1, \dots, x_N) = (K - S_0 \prod_{i=1}^N x_i)^+$.

7.1.2 liens entre les paramètres r , a et b

Une hypothèse de modélisation très naturelle consiste à supposer l'Absence d'Opportunités d'Arbitrage. D'un point de vue intuitif, cette hypothèse signifie qu'il n'est pas possible de faire du profit sans prendre de risque.

Elle se traduit de la manière suivante d'un point de vue mathématique : partant d'une richesse nulle $V_0 = 0$ à l'instant initial, il n'est pas possible de faire des investissements en actif sans risque (placement à la banque au taux r) et en actif risqué qui permettent sans jamais injecter d'argent d'obtenir une richesse V_N à l'instant terminal qui vérifie

$$\mathbf{P}(V_N \geq 0) = 1 \text{ et } \mathbf{P}(V_N > 0) > 0.$$

Une conséquence facile de cette propriété d'Absence d'Opportunités d'Arbitrage est que deux actifs financiers qui ont la même valeur en N doivent avoir la même valeur à l'instant initial. Sinon acheter le moins cher en vendant le plus cher et placer la différence en actif sans risque permet de réaliser un arbitrage. En effet, à l'instant final N , le profit tiré de l'actif acheté compense exactement la somme à verser à l'acheteur de l'actif vendu. Et partant d'une richesse initiale nulle, on dispose de la somme placée en actif sans risque.

Même si elle se produit parfois sur le marché, la situation où les deux actifs ont des valeurs initiales différentes est très instable. En effet, les agents qui décèlent cet arbitrage vont essayer d'en profiter en achetant l'actif le moins cher et en vendant le plus cher, ce qui par la loi de l'offre et de la demande, va entraîner le retour à la situation stable où les deux prix sont égaux.

Exemple : A la différence d'une option où la prime est versée par l'acheteur à l'instant initial en échange de la richesse h en N , dans un contrat à terme on convient à l'instant initial du montant qui sera versé en N en échange d'un actif (l'actif est remis en N).

On peut déterminer par A.O.A., le montant M à verser à terme en échange d'une unité d'actif risqué. En effet, le montant à verser à l'instant initial en échange d'une unité d'actif risqué en N est S_0 . Et la valeur à l'instant initial de M euros à l'instant N est $M/S_N^0 = M/(1+r)^N$. On doit donc avoir $M/(1+r)^N = S_0$ soit $M = S_0(1+r)^N$.

Dans le modèle de Cox-Ross-Rubinstein, il est facile de vérifier que

$$\boxed{a < r < b} \tag{7.3}$$

est une condition nécessaire pour qu'il y ait Absence d'Opportunités d'Arbitrage. Rappelons que l'on a supposé $-1 < a < b$.

- Si $r \leq a$ alors il suffit d'acheter une unité d'actif risqué à l'instant initial en empruntant le montant correspondant à la banque pour réaliser un arbitrage. En effet, partant de $V_0 = 0$, on disposera en N , après remboursement de la somme empruntée, de la richesse

$$V_N = S_N - S_0(1+r)^N = S_0 \left(\prod_{n=1}^N T_n - (1+r)^N \right) \geq S_0((1+a)^N - (1+r)^N) \geq 0.$$

Et l'hypothèse (7.1) implique que $\mathbf{P}(V_N = S_0((1+b)^N - (1+r)^N)) > 0$.

- Si $r \geq b$, alors on va emprunter une unité d'actif risqué à l'instant initial, la vendre et placer la somme reçue S_0 à la banque. La possibilité de vendre une unité d'actif qu'on ne possède pas peut sembler étonnante mais elle existe sur les marchés financiers : cela s'appelle de la vente à découvert. Bien sûr en N , il faut rembourser S_N et on dispose encore de

$$V_N = S_0(1+r)^N - S_N \geq S_0((1+r)^N - (1+b)^N) \geq 0.$$

Le fait que $\mathbf{P}(V_N > 0) > 0$ découle à nouveau de (7.1).

Nous supposons donc désormais que l'hypothèse (7.3) est vérifiée.

L'hypothèse d'Absence d'Opportunités d'Arbitrage a une conséquence très importante pour traiter le problème du pricing et de la couverture de l'option européenne définie par h :

Supposons que le vendeur de l'option h sache constituer un portefeuille qui comporte de l'actif sans risque (placement à la banque) et de l'actif risqué sans ajouter ni enlever d'argent lors des éventuelles modifications de la composition du portefeuille après l'instant initial et dont la valeur terminale est $V_N = h$. Alors, comme le portefeuille et l'option ont même valeur à l'instant terminal, par Absence d'Opportunités d'Arbitrage, la prime de l'option est égale à la valeur initiale du portefeuille. En outre, le vendeur est en mesure de couvrir le risque lié à l'option : en investissant la prime pour gérer le portefeuille, il sera en mesure de fournir la richesse h à l'acheteur à l'instant N .

Dans le paragraphe suivant, nous verrons comment constituer un tel portefeuille dans le cas très simplifié où il y a une seule période de temps : $N = 1$. Nous aborderons ensuite le cas général.

7.1.3 le cas d'une seule période de temps : $N = 1$

D'après (7.2), il existe $f : \{1 + a, 1 + b\} \rightarrow \mathbf{R}_+$ t.q. le profit de l'acheteur de l'option est $h = f(T_1)$.

On souhaite trouver des quantités ϕ^0 et ϕ à investir respectivement en actif sans risque et en actif risqué à l'instant initial de façon à ce que la valeur $V_1 = \phi^0 S_1^0 + \phi S_1$ du portefeuille correspondant soit égale à $h = f(T_1)$ l'instant 1. Compte tenu de (7.1), on obtient

$$\begin{cases} \phi^0(1+r) + \phi S_0(1+a) = f(1+a) \\ \phi^0(1+r) + \phi S_0(1+b) = f(1+b). \end{cases}$$

Ce système admet pour unique solution

$$\begin{cases} \phi = \frac{f(1+b)-f(1+a)}{S_0(b-a)} \\ \phi^0 = \frac{1}{1+r} \left(f(1+a) - \frac{f(1+b)-f(1+a)}{(b-a)}(1+a) \right) = \frac{(1+b)f(1+a)-(1+a)f(1+b)}{(1+r)(b-a)}. \end{cases}$$

Suivant l'hypothèse d'Absence d'Opportunités d'Arbitrage, la valeur de l'option à l'instant initial doit être égale à la valeur initiale du portefeuille

$$\begin{aligned} V_0 = \phi^0 + \phi S_0 &= \frac{(1+b)f(1+a) - (1+a)f(1+b)}{(1+r)(b-a)} + \frac{f(1+b) - f(1+a)}{(b-a)} \\ &= \frac{1}{1+r} \left(\frac{b-r}{b-a} f(1+a) + \frac{r-a}{b-a} f(1+b) \right). \end{aligned}$$

Du fait de l'inégalité $a < r < b$ qui est une conséquence de l'Absence d'Opportunités d'Arbitrage, on peut définir \mathbf{P}^* probabilité t.q. $\mathbf{P}^*(T_1 = 1+a) = (b-r)/(b-a)$ et $\mathbf{P}^*(T_1 = 1+b) = (r-a)/(b-a)$. Alors si on note \mathbf{E}^* l'espérance associée, on a

$$V_0 = \frac{1}{1+r} \mathbf{E}^*(f(T_1)) = \mathbf{E}^* \left(\frac{h}{1+r} \right).$$

Ainsi la valeur de l'option est l'espérance sous la probabilité \mathbf{P}^* du profit h de l'acheteur actualisé (présence du facteur $1/(1+r)$).

Remarque 7.1.1 En général le prix n'est pas égal à l'espérance du profit (ou payoff) actualisé sous la probabilité de départ.

L'approche n'est pas une approche de type assurance. On peut supposer (en première approximation) que les remboursements $(X_i)_{1 \leq i \leq n}$ effectués par un assureur automobile à ses n (n grand) clients sont des variables indépendantes et identiquement distribuées. Pour déterminer la prime qu'il demande à ses clients, l'assureur tient compte de la loi forte des grands nombres qui indique que $\frac{1}{n} \sum_{i=1}^n X_i$ est proche de $\mathbf{E}(X_1)$ (espérance sous la probabilité de départ cette fois) : il demande donc une prime égale à $\mathbf{E}(X_1)$ plus une marge qui sert à la fois à le rémunérer et à le prémunir contre les fluctuations de la somme totale à rembourser $\sum_{i=1}^n X_i$ autour de $n\mathbf{E}(X_1)$ (le théorème de la limite centrale donne le comportement de ces fluctuations).

En revanche, dans le cas des marchés financiers, seul un des scénarii d'évolution possibles se réalise (il y a deux alternatives dans le cas particulier du modèle de Cox-Ross-Rubinstein à une période de temps : $T_1 = 1 + a$ et $T_1 = 1 + b$) et le vendeur de l'option ne peut pas jouer sur la loi forte des grands nombres. C'est l'existence du portefeuille de réplcation qui permet de donner un prix à l'option. Et le fait que la valeur initiale de ce portefeuille et donc de l'option s'écrit comme espérance sous \mathbf{P}^* du profit h actualisé est une conséquence des calculs menés pour déterminer le portefeuille et non de la loi forte des grands nombres.

Remarque 7.1.2 On a

$$\mathbf{E}^*(S_1) = S_0 \mathbf{E}^*(T_1) = S_0 \left((1+a) \frac{b-r}{b-a} + (1+b) \frac{r-a}{b-a} \right) = S_0(1+r).$$

Ainsi, sous la probabilité \mathbf{P}^* , l'actif risqué rapporte autant en moyenne que l'actif sans risque. En ce sens, il devient indifférent (ou neutre) d'investir en actif sans risque ou en actif risqué. C'est pourquoi \mathbf{P}^* est appelée probabilité risque-neutre.

7.2 Portefeuilles, arbitrages

Pour traiter le cas général $N \geq 1$, il faut introduire la notion de stratégie autofinancée de gestion de portefeuille.

7.2.1 portefeuille

Définition 7.2.1 Une stratégie de gestion de portefeuille Φ est la donnée d'une suite $((\phi_n^0, \phi_n))_{1 \leq n \leq N}$ \mathcal{F}_{n-1} -adaptée où ϕ_n^0 et ϕ_n représentent respectivement la quantité d'actif sans risque et la quantité d'actif risqué détenues sur la période $[n-1, n]$.

Il est naturel de supposer que (ϕ_n^0, ϕ_n) est \mathcal{F}_{n-1} -mesurable car c'est au vu de l'information \mathcal{F}_{n-1} disponible à l'instant $n-1$ que l'investisseur choisit la composition de son portefeuille sur la période $[n-1, n]$.

La valeur du portefeuille à l'instant n est

$$V_n^\Phi = \begin{cases} \phi_1^0 + \phi_1 S_0 & \text{si } n = 0 \\ \phi_n^0 (1+r)^n + \phi_n S_n & \text{si } 1 \leq n \leq N \end{cases}$$

On définit également la valeur actualisée de l'actif risqué

$$\tilde{S}_n = \frac{S_n}{S_n^0} = \frac{S_n}{(1+r)^n}$$

et du portefeuille

$$\tilde{V}_n^\Phi = \begin{cases} V_0^\Phi = \phi_0^0 + \phi_1 S_0 & \text{si } n = 0 \\ (1+r)^{-n} V_n^\Phi = \phi_n^0 + \phi_n \tilde{S}_n & \text{si } 1 \leq n \leq N \end{cases}$$

Une notion très importante en vue de définir et de caractériser les arbitrages est celle de stratégie autofinancée.

Définition 7.2.2 Une stratégie est dite autofinancée si on n'ajoute pas et on n'enlève pas d'argent du portefeuille après l'instant initial. Autrement dit, aux instants $1 \leq n \leq N-1$, les modifications de la composition du portefeuille se font à valeur constante :

$$\forall 1 \leq n \leq N-1, \phi_n^0(1+r)^n + \phi_n S_n = \phi_{n+1}^0(1+r)^n + \phi_{n+1} S_n.$$

La stratégie est autofinancée si et seulement si l'évolution de la valeur du portefeuille provient uniquement de l'évolution de l'actif sans risque et de l'actif risqué entre les instants de discrétisation successifs :

Proposition 7.2.3 Il y a équivalence entre

i) La stratégie Φ est autofinancée.

ii) $\forall 1 \leq n \leq N, V_n^\Phi = V_0^\Phi + \sum_{j=1}^n (\phi_j^0 \Delta S_j^0 + \phi_j \Delta S_j)$,
où pour $1 \leq j \leq N$, on note $(\Delta S_j^0, \Delta S_j) = (S_j^0 - S_{j-1}^0, S_j - S_{j-1})$.

iii) $\forall 1 \leq n \leq N, \tilde{V}_n^\Phi = V_0^\Phi + \sum_{j=1}^n \phi_j \Delta \tilde{S}_j$,
où pour $1 \leq j \leq N, \Delta \tilde{S}_j = \tilde{S}_j - \tilde{S}_{j-1}$.

Démonstration : Nous allons montrer l'équivalence de i) et iii). L'équivalence de i) et ii) découle des mêmes idées.

Pour $1 \leq j \leq N$,

$$\tilde{V}_j^\Phi = \phi_j^0 + \phi_j \tilde{S}_j = \phi_j^0 + \phi_j \tilde{S}_{j-1} + \phi_j \Delta \tilde{S}_j. \quad (7.4)$$

Lorsque $2 \leq j \leq N$, la condition d'autofinancement à l'instant $j-1$ s'écrit après division par $(1+r)^{j-1}$,

$$\phi_j^0 + \phi_j \tilde{S}_{j-1} = \phi_{j-1}^0 + \phi_{j-1} \tilde{S}_{j-1} = \tilde{V}_{j-1}^\Phi.$$

Avec (7.4), on en déduit que l'autofinancement est équivalent à

$$\forall 1 \leq j \leq N, \tilde{V}_j^\Phi - \tilde{V}_{j-1}^\Phi = \phi_j \Delta \tilde{S}_j.$$

Et cette famille d'égalités implique iii) par sommation. Elle s'en déduit en écrivant l'égalité de iii) aux instants n et $n-1$ et en effectuant la différence. ■

7.2.2 Absence d'Opportunités d'Arbitrage

Nous allons donner une définition mathématique de l'Absence d'Opportunités d'Arbitrage. Pour cela, il faut commencer par définir une stratégie d'arbitrage

Définition 7.2.4 – On appelle stratégie d'arbitrage une stratégie Φ autofinancée de valeur initiale nulle : $V_0^\Phi = 0$ et de valeur finale positive non nulle : $V_N^\Phi \geq 0$ p.s. et $\mathbf{P}(V_N^\Phi > 0) > 0$.

– On dit qu'il y a Absence d'Opportunités d'Arbitrage s'il n'existe pas de stratégie d'arbitrage. Dans ce cas, le marché est dit viable.

Nous avons déjà vu que $a < r < b$ est une condition nécessaire pour qu'il y ait Absence d'Opportunités d'Arbitrage. C'est en fait également une condition suffisante :

Proposition 7.2.5 *Il y a Absence d'Opportunités d'Arbitrage ssi $a < r < b$.*

Pour démontrer la condition suffisante, nous supposons $a < r < b$ et nous introduisons \mathbf{P}^* probabilité t.q sous \mathbf{P}^* , les variables T_1, \dots, T_N sont i.i.d. avec

$$\mathbf{P}^*(T_j = 1 + a) = \frac{b - r}{b - a} \quad \text{et} \quad \mathbf{P}^*(T_j = 1 + b) = \frac{r - a}{b - a}.$$

Notons que cette définition généralise celle donné dans le cas d'une seule période de temps.

Lemme 7.2.6 *Sous \mathbf{P}^* ,*

- *la valeur actualisée $(\tilde{S}_n)_{0 \leq n \leq N}$ de l'actif risqué est une \mathcal{F}_n -martingale.*
- *pour toute stratégie autofinancée Φ , la valeur actualisée $(\tilde{V}_n^\Phi)_{0 \leq n \leq N}$ du portefeuille associé est une \mathcal{F}_n -martingale.*

En admettant (provisoirement) le lemme, il est facile de démontrer la proposition.

Soit Φ une stratégie autofinancée de valeur initiale nulle : $V_0^\Phi = 0$ et de valeur terminale positive : $\mathbf{P}(V_N^\Phi \geq 0) = 1$. Comme V_N^Φ est $\mathcal{F}_N = \sigma(T_1, \dots, T_N)$ -mesurable,

$$\exists v_N : \{1 + a, 1 + b\}^N \rightarrow \mathbf{R} \text{ t.q. } V_N^\Phi = v_N(T_1, \dots, T_N).$$

D'après l'hypothèse (7.1), tous les scenarii d'évolution de l'actif risqué ont une probabilité strictement positive : $\forall (x_1, \dots, x_N) \in \{1 + a, 1 + b\}^N, \mathbf{P}(T_1 = x_1, \dots, T_N = x_N) > 0$. Avec $\mathbf{P}(V_N^\Phi \geq 0) = 1$, on en déduit que la fonction v_N est à valeurs dans \mathbf{R}_+ .

Le fait que $(\tilde{V}_n^\Phi)_n$ soit une martingale sous la probabilité \mathbf{P}^* entraîne que

$$\mathbf{E}^*(\tilde{V}_N^\Phi) = V_0^\Phi = 0.$$

En multipliant par $(1 + r)^N$, on en déduit que

$$0 = \mathbf{E}^*(v_N(T_1, \dots, T_N)) = \sum_{(x_1, \dots, x_N) \in \{1+a, 1+b\}^N} v_N(x_1, \dots, x_N) \mathbf{P}^*(T_1 = x_1) \dots \mathbf{P}^*(T_N = x_N).$$

Cette dernière somme est constituée de termes positifs $v_N(x_1, \dots, x_N)$ multipliés par des termes strictement positifs $\mathbf{P}^*(T_1 = x_1) \dots \mathbf{P}^*(T_N = x_N)$. Comme elle est nulle, la fonction v_N est nécessairement identiquement nulle. D'où $\mathbf{P}(V_N^\Phi = 0) = 1$ et il n'existe pas de stratégie d'arbitrage.

Il reste à démontrer le lemme pour achever la preuve de la proposition :

Démonstration : La suite $(\tilde{S}_n)_{0 \leq n \leq N}$ est \mathcal{F}_n -adapté. En outre chaque variable \tilde{S}_n est intégrable sous \mathbf{P}^* comme produit de variables aléatoires indépendantes intégrables.

Soit $0 \leq n \leq N - 1$. On a $\tilde{S}_{n+1} = \frac{\tilde{S}_n}{1+r} T_{n+1}$. La variable \tilde{S}_n est \mathcal{F}_n -mesurable. Sous \mathbf{P}^* , T_{n+1} est indépendante de (T_1, \dots, T_n) et donc de $\mathcal{F}_n = \sigma(T_1, \dots, T_n)$. En utilisant les propriétés 4 et 6 de l'espérance conditionnelle, on en déduit que

$$\begin{aligned} \mathbf{E}^*(\tilde{S}_{n+1} | \mathcal{F}_n) &= \frac{\tilde{S}_n}{1+r} \mathbf{E}^*(T_{n+1}) \\ &= \frac{\tilde{S}_n}{1+r} \left((1+a) \frac{b-r}{b-a} + (1+b) \frac{r-a}{b-a} \right) \\ &= \frac{\tilde{S}_n}{1+r} (1+r) = \tilde{S}_n. \end{aligned}$$

Ainsi $(\tilde{S}_n)_{0 \leq n \leq N}$ est une \mathcal{F}_n -martingale sous \mathbf{P}^* .

Soit maintenant Φ une stratégie autofinancée. Il est facile de vérifier que la suite $(\tilde{V}_n)_{0 \leq n \leq N}$ est \mathcal{F}_n -adaptée.

Soit $0 \leq n \leq N - 1$. Comme la stratégie est autofinancée, d'après la proposition 7.2.3,

$$\tilde{V}_{n+1}^\Phi = \tilde{V}_n^\Phi + \phi_{n+1} \Delta \tilde{S}_{n+1}. \quad (7.5)$$

La quantité ϕ_{n+1} d'actif risqué détenu dans le portefeuille sur la période $[n, n + 1]$ est $\mathcal{F}_n = \sigma(T_1, \dots, T_n)$ -mesurable. Donc dans l'esprit de la proposition 3.1.2, il existe une fonction $f_n : \{1 + a, 1 + b\}^n \rightarrow \mathbf{R}$ t.q. $\phi_{n+1} = f_n(T_1, \dots, T_n)$. Cela implique en particulier que la variable aléatoire ϕ_{n+1} est bornée. En prenant l'espérance conditionnelle sachant \mathcal{F}_n dans (7.5) et en utilisant notamment la propriété 4 de l'espérance conditionnelle on en déduit

$$\mathbf{E}(\tilde{V}_{n+1}^\Phi | \mathcal{F}_n) = \tilde{V}_n^\Phi + \phi_n (\mathbf{E}(\tilde{S}_{n+1} | \mathcal{F}_n) - \tilde{S}_n) = \tilde{V}_n^\Phi.$$

En fait $(\tilde{V}_n^\Phi)_{0 \leq n \leq N}$ est une transformée de martingales au sens de la proposition 4.3.1 et nous venons de redémontrer cette proposition dans un cas particulier. ■

Remarque 7.2.7 D'après le lemme 7.2.6, sous \mathbf{P}^* , la valeur actualisée de l'actif risqué est une martingale. C'est pourquoi on dit que \mathbf{P}^* est une probabilité martingale.

La constance de l'espérance de cette martingale s'écrit

$$\forall 1 \leq n \leq N, \mathbf{E}^*(S_n) = (1 + r)^n S_0,$$

c'est à dire que sous \mathbf{P}^* , l'actif risqué rapporte en moyenne autant que l'actif sans risque. En ce sens, il est indifférent (ou neutre) d'investir en actif sans risque ou en actif risqué. C'est pourquoi on dit également que \mathbf{P}^* est une probabilité risque-neutre.

Remarque 7.2.8 On peut légitimement se demander comment construire \mathbf{P}^* . Il suffit en fait de se placer sur l'espace $\Omega = \{1 + a, 1 + b\}^N$ muni de la tribu discrète $\mathcal{F} = \mathcal{P}(\Omega)$. Pour $\omega = (x_1, \dots, x_N) \in \Omega$, on pose

$$\begin{cases} T_n(\omega) = x_n \text{ pour } 1 \leq n \leq N \\ \mathbf{P}(\omega) = \left(\frac{b-r}{b-a}\right)^{N_a} \left(\frac{r-a}{b-a}\right)^{N-N_a} \text{ où } N_a = \#\{i : x_i = 1 + a\} \end{cases}$$

Alors il est facile (exercice) de vérifier que sous \mathbf{P}^* , les variables T_1, \dots, T_N sont i.i.d. de loi donnée plus haut.

7.3 Pricing des options européennes

On rappelle que l'on se place du point de vue du vendeur d'une option européenne définie par une variable aléatoire h positive \mathcal{F}_N -mesurable : on souhaite déterminer le prix et la couverture de cette option sous l'hypothèse d'Absence d'Opportunités d'Arbitrage : $a < r < b$.

Supposons qu'il existe une stratégie de portefeuille autofinancée qui réplique l'option au sens où la valeur terminale du portefeuille est égale au profit obtenu par l'acheteur de l'option : $V_N^\Phi = h$.

Alors comme l'option et le portefeuille ont même valeur à l'instant terminal et comme on ne fait ni ajout ni retrait d'argent au cours de la gestion du portefeuille entre l'instant initial et l'instant final, il est naturel de définir la valeur de l'option à un instant n antérieur comme la valeur V_n^Φ du portefeuille à cet instant.

D'après le Lemme 7.2.6, sous la probabilité \mathbf{P}^* introduite dans le paragraphe qui précède, la valeur actualisée du portefeuille $(\tilde{V}_n^\Phi)_{0 \leq n \leq N}$ est une \mathcal{F}_n -martingale. Donc la valeur de l'option à l'instant n est alors

$$V_n^\Phi = (1+r)^n \tilde{V}_n^\Phi = (1+r)^n \mathbf{E}^*(\tilde{V}_N^\Phi | \mathcal{F}_n) = (1+r)^{n-N} \mathbf{E}^*(h | \mathcal{F}_n).$$

Nous allons montrer que l'on peut construire une telle stratégie :

Théorème 7.3.1 *On suppose $a < r < b$. Alors il existe une stratégie Φ autofinancée qui permet de répliquer l'option i.e. t.q. $V_N^\Phi = h$. Et pour $0 \leq n \leq N$, la valeur de l'option à l'instant n est*

$$V_n^\Phi = (1+r)^{n-N} \mathbf{E}^*(h | \mathcal{F}_n).$$

Démonstration : On note $\tilde{V}_n = \mathbf{E}^*((1+r)^{-N} h | \mathcal{F}_n)$ la valeur actualisée présumée de l'option à l'instant n . La suite \tilde{V}_n est une \mathcal{F}_n -martingale sous la probabilité \mathbf{P}^* .

On cherche pour $1 \leq n \leq N$ une quantité ϕ_n \mathcal{F}_{n-1} -mesurable t.q.

$$\tilde{V}_n = \tilde{V}_{n-1} + \phi_n \Delta \tilde{S}_n. \quad (7.6)$$

Comme la variable aléatoire $\tilde{V}_n = \mathbf{E}^*(h | (1+r)^N | \mathcal{F}_n)$ est $\mathcal{F}_n = \sigma(T_1, \dots, T_n)$ -mesurable,

$$\exists \tilde{v}_n : \{1+a, 1+b\}^n \rightarrow \mathbf{R}_+ \text{ t.q. } \tilde{V}_n = \tilde{v}_n(T_1, \dots, T_n).$$

Pour $n = 0$, cela signifie que \tilde{v}_0 est une constante. Pour $1 \leq n \leq N$, l'égalité $\tilde{V}_{n-1} = \mathbf{E}^*(\tilde{V}_n | \mathcal{F}_{n-1})$ se réécrit

$$\tilde{v}_{n-1}(T_1, \dots, T_{n-1}) = \mathbf{E}^*(\tilde{v}_n(T_1, \dots, T_n) | \mathcal{F}_{n-1}).$$

La variable (T_1, \dots, T_{n-1}) est \mathcal{F}_{n-1} mesurable. En outre sous \mathbf{P}^* , T_n est indépendante de $\mathcal{F}_{n-1} = \sigma(T_1, \dots, T_{n-1})$. Donc d'après le Lemme 3.2.6,

$$\tilde{v}_{n-1}(T_1, \dots, T_{n-1}) = \frac{b-r}{b-a} \tilde{v}_n(T_1, \dots, T_{n-1}, 1+a) + \frac{r-a}{b-a} \tilde{v}_n(T_1, \dots, T_{n-1}, 1+b). \quad (7.7)$$

Compte tenu des deux alternatives $T_n = 1+a$ et $T_n = 1+b$, l'équation (7.6) est équivalente au système de deux équations à une seule inconnue ϕ_n :

$$\begin{cases} \tilde{v}_n(T_1, \dots, T_{n-1}, 1+a) = \tilde{v}_{n-1}(T_1, \dots, T_{n-1}) + \phi_n \left(\frac{1+a}{1+r} - 1 \right) \tilde{S}_{n-1} \\ \tilde{v}_n(T_1, \dots, T_{n-1}, 1+b) = \tilde{v}_{n-1}(T_1, \dots, T_{n-1}) + \phi_n \left(\frac{1+b}{1+r} - 1 \right) \tilde{S}_{n-1}. \end{cases}$$

Si ce système admet une solution, alors en faisant la différence entre la seconde et la première équation, on obtient qu'elle est égale à

$$\phi_n = (1+r) \frac{\tilde{v}_n(T_1, \dots, T_{n-1}, 1+b) - \tilde{v}_n(T_1, \dots, T_{n-1}, 1+a)}{(b-a) \tilde{S}_{n-1}}.$$

La relation (7.7) est précisément la relation de compatibilité qui assure que cette valeur trouvée pour ϕ_n est bien solution des deux équations.

On a donc trouvé une suite $(\phi_n)_{1 \leq n \leq N}$ \mathcal{F}_{n-1} -adaptée t.q. pour $1 \leq n \leq N$, l'égalité (7.6) est vérifiée. On pose $\phi_n^0 = \tilde{V}_{n-1} - \phi_n \tilde{S}_{n-1}$ pour $1 \leq n \leq N$. La stratégie de portefeuille $\Phi = ((\phi_n^0, \phi_n))_{1 \leq n \leq N}$ est telle que

$$\begin{cases} V_0^\Phi = \phi_1^0 + \phi_1 S_0 = \tilde{V}_0 = V_0 \\ \forall 1 \leq n \leq N, \tilde{V}_n^\Phi = \phi_n^0 + \phi_n \tilde{S}_n = \phi_n^0 + \phi_n \tilde{S}_{n-1} + \phi_n \Delta \tilde{S}_n = \tilde{V}_{n-1} + \phi_n \Delta \tilde{S}_n = \tilde{V}_n. \end{cases}$$

En particulier $V_N^\Phi = (1+r)^N \tilde{V}_N = h$ et (7.6) se réécrit $\forall 1 \leq n \leq N$, $\tilde{V}_n^\Phi = \tilde{V}_{n-1}^\Phi + \phi_n \Delta \tilde{S}_n$. Ainsi la stratégie Φ réplique l'option et elle est autofinancée d'après la proposition 7.2.3. ■

Exercice 47 On se place dans le modèle de Cox-Ross-Rubinstein à N périodes avec $-1 < a < r < b$. On s'intéresse à une option européenne standard d'échéance N et de fonction de payoff $g : \mathbf{R} \rightarrow \mathbf{R}_+$ i.e. t.q. le profit réalisé par l'acheteur à l'instant n est

$$h = g(S_n).$$

1. Pour quelle probabilité \mathbb{P}^* , la valeur de l'option à l'instant n est-elle

$$V_n = \mathbb{E}^*((1+r)^{n-N} h(S_N) | \mathcal{F}_n)?$$

2. Montrer que $V_n = C(n, S_n)$ pour une fonction $C(n, x)$ que l'on précisera. Que peut-on dire de $x \rightarrow C(n, x)$ si g est croissante ?
3. Montrer que $C(n, S_n) = \frac{1}{1+r} \mathbb{E}^*(C(n+1, S_{n+1}) | \mathcal{F}_n)$. En déduire une relation de récurrence descendante qui permet de calculer $C(n, x)$ en fonction de $C(n+1, \cdot)$.
4. Montrer que la quantité d'actif risqué à détenir sur la période $[n-1, n]$ pour couvrir l'option est donnée par

$$\phi_n = \frac{C(n, S_{n-1}(1+b)) - C(n, S_{n-1}(1+a))}{S_{n-1}(b-a)}.$$

Que peut-on dire de cette quantité si g est croissante ?

5. On suppose dans cette question que pour un certain $c > 0$, la fonction $x \rightarrow cx - g(x)$ est positive et croissante. Montrer qu'alors $\phi_n \leq c$.
6. En déduire des bornes pour le ratio de couverture ϕ_n dans le cas du Call.

Chapitre 8

Mouvement brownien et martingales

8.1 Généralités sur les processus à temps continu

Comme dans le cas discret, on introduit la notion de filtration qui représente l'information disponible.

Définition 8.1.1 Soit (Ω, \mathcal{A}) un espace muni d'une tribu, on appelle filtration une suite croissante de sous tribus de \mathcal{A} , $(\mathcal{F}_t, t \geq 0)$. Un processus $(X_t, t \geq 0)$ est un processus adapté à une filtration $(\mathcal{F}_t, t \geq 0)$, si pour tout $t \geq 0$, X_t est \mathcal{F}_t -mesurable.

Remarque 8.1.2 La filtration naturelle associée à un processus $(X_t \geq 0)$ est par définition la famille de sous tribus :

$$\mathcal{F}_t = \sigma(X_s, s \leq t),$$

où \mathcal{F}_t est la plus petite tribu rendant mesurable les applications $\omega \rightarrow X_s(\omega)$, pour $s \leq t$.

Comme dans le cas discret, on définit la notion de temps d'arrêt.

Définition 8.1.3 Soit $(\mathcal{F}_t, t \geq 0)$ une filtration, soit τ une variable aléatoire positive à valeurs dans $\mathbf{R}^+ \cup \{+\infty\}$. On dit que τ est un temps d'arrêt si pour tout t :

$$\{\tau \leq t\} \in \mathcal{F}_t.$$

Remarque 8.1.4 Il est facile de vérifier que si S et T sont deux temps d'arrêt, alors $S \wedge T = \inf(S, T)$ est encore un temps d'arrêt. En particulier, si $t \in \mathbf{R}^+$, $S \wedge t$ est un temps d'arrêt.

Martingales à temps continu On peut étendre la notion de martingale au cas des processus à temps continu.

Définition 8.1.5 Soit $(\mathcal{F}_t, t \geq 0)$ une filtration sur un espace $(\Omega, \mathcal{A}, \mathbf{P})$ et $(M_t, t \geq 0)$ un processus adapté à cette filtration. On dit que $(M_t, t \geq 0)$ est une martingale par rapport à la filtration $(\mathcal{F}_t, t \geq 0)$, si pour tout $t \in \mathbf{R}^+$, $\mathbf{E}(|M_t|) < +\infty$ et si pour tout $s \leq t$:

$$\mathbf{E}(M_t | \mathcal{F}_s) = M_s.$$

Remarque 8.1.6 Dans le cas discret la définition affirme que :

$$\mathbf{E}(M_{n+1} | \mathcal{F}_n) = M_n,$$

et nous avons vu que l'on peut en déduire que si $p \geq n$:

$$\mathbf{E}(M_p | \mathcal{F}_n) = M_n.$$

La définition précédente généralise donc celle du cas discret.

Théorème d'arrêt pour les martingales continues Si $(M_t, t \geq 0)$ est une martingale *continue* (i.e. les trajectoires $t \rightarrow M_t(\omega)$ sont continues, sauf pour un ensemble de ω de probabilité nulle pour \mathbf{P}), on peut étendre l'égalité :

$$\mathbf{E}(M_t) = \mathbf{E}(M_0),$$

qui découle de la définition en prenant l'espérance des deux membres a priori vraie pour un temps fixe t à un *temps d'arrêt borné*. Ce résultat porte le nom de théorème d'arrêt.

Théorème 8.1.1 Soit T un temps d'arrêt par rapport à une filtration $(\mathcal{F}_t, t \geq 0)$. Soit $(M_t, t \geq 0)$ une martingale continue par rapport à la même filtration. Si $T \leq k < +\infty$, k étant une constante réelle, alors on a :

$$\mathbf{E}(M_T) = \mathbf{E}(M_0).$$

8.2 Extensions de la définition du mouvement brownien

Nous commençons par définir le mouvement brownien par rapport à une filtration donnée.

Définition 8.2.1 On dit que $(X_t, t \geq 0)$ est un mouvement brownien (standard) par rapport à une filtration $(\mathcal{F}_t, t \geq 0)$, si $(X_t, t \geq 0)$ est un processus *continu* adapté à cette filtration, nul en 0 et si les *accroissements* $X_{t+h} - X_t$ sont *indépendants* de \mathcal{F}_t et de loi gaussienne centrée de variance h .

Remarque 8.2.2 Noter que cette définition implique que $(X_t, t \geq 0)$ est un mouvement brownien au sens classique (puisque $\sigma(X_s, s \leq t) \subset \mathcal{F}_t$).

De cette définition on peut déduire les propriétés suivantes.

Proposition 8.2.3 Si $(X_t, t \geq 0)$ est un mouvement brownien par rapport à $(\mathcal{F}_t, t \geq 0)$ alors :

- $(-X_t, t \geq 0)$ est un mouvement brownien par rapport à $(\mathcal{F}_t, t \geq 0)$,
- si $c > 0$, $(\frac{1}{c}X_{c^2t}, t \geq 0)$ est un mouvement brownien par rapport à $(\mathcal{F}_{c^2t}, t \geq 0)$,
- si $s \geq 0$, $(X_{t+s} - X_s, t \geq 0)$ est un mouvement brownien par rapport à $(\mathcal{F}_{t+s}, t \geq 0)$, de plus ce processus est indépendant de \mathcal{F}_s .

8.3 Exemples de martingales browniennes

Proposition 8.3.1 Soit $(X_t, t \geq 0)$ un mouvement brownien standard, alors si $\mathcal{F}_t = \sigma(X_s, s \leq t)$:

- $(X_t, t \geq 0)$ est une martingale par rapport à $(\mathcal{F}_t, t \geq 0)$,
- $(X_t^2 - t, t \geq 0)$ est une martingale par rapport à $(\mathcal{F}_t, t \geq 0)$,
- $(e^{\sigma X_t - \frac{\sigma^2}{2}t}, t \geq 0)$ est une martingale par rapport à $(\mathcal{F}_t, t \geq 0)$.

Démonstration : Notons que si $s \leq t$, on a, comme $X_t - X_s$ est indépendante de la tribu \mathcal{F}_s :

$$\mathbf{E}(X_t - X_s | \mathcal{F}_s) = \mathbf{E}(X_t - X_s).$$

Mais $\mathbf{E}(X_t - X_s) = \mathbf{E}(X_t) - \mathbf{E}(X_s) = 0 - 0 = 0$, donc :

$$\mathbf{E}(X_t - X_s | \mathcal{F}_s) = 0.$$

X_s étant \mathcal{F}_s -mesurable on en déduit que $\mathbf{E}(X_s | \mathcal{F}_s) = X_s$ puis que :

$$\mathbf{E}(X_t | \mathcal{F}_s) = X_s.$$

Ceci prouve le premier point.

Pour démontrer la deuxième propriété remarquons que :

$$\mathbf{E} \left((X_t - X_s)^2 | \mathcal{F}_s \right) = \mathbf{E}(X_t^2 | \mathcal{F}_s) - 2X_s \mathbf{E}(X_t | \mathcal{F}_s) + X_s^2.$$

Comme X_s est une martingale, on a $\mathbf{E}(X_t | \mathcal{F}_s) = X_s$ et donc :

$$\mathbf{E} \left((X_t - X_s)^2 | \mathcal{F}_s \right) = \mathbf{E}(X_t^2 | \mathcal{F}_s) - X_s^2.$$

Notons, de plus, que comme $X_t - X_s$ est indépendant de \mathcal{F}_s , on a :

$$\mathbf{E} \left((X_t - X_s)^2 | \mathcal{F}_s \right) = \mathbf{E} \left((X_t - X_s)^2 \right).$$

Mais $X_t - X_s$ suit la même loi que $X_{t-s} - X_0 = X_{t-s}$, c'est à dire une loi gaussienne centrée de variance $t - s$, on a donc :

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

Donc il vient :

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

ou encore :

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

Ce qui prouve la deuxième propriété.

Pour prouver la troisième propriété notons, en utilisant l'indépendance des accroissements, que :

$$\mathbf{E} \left(e^{\sigma(X_t - X_s)} | \mathcal{F}_s \right) = \mathbf{E} \left(e^{\sigma(X_t - X_s)} \right).$$

Mais la loi de $X_t - X_s$ est celle d'une gaussienne centrée de variance $t - s$ (dont on peut calculer la transformée de Laplace) donc :

$$\mathbf{E} \left(e^{\sigma(X_t - X_s)} \right) = e^{\frac{\sigma^2}{2}(t-s)}.$$

On a donc :

$$\mathbf{E} \left(e^{\sigma(X_t - X_s)} | \mathcal{F}_s \right) = e^{\frac{\sigma^2}{2}(t-s)}.$$

La variable aléatoire X_s étant \mathcal{F}_s -mesurable on obtient donc :

$$\mathbf{E} \left(e^{\sigma X_t - \frac{\sigma^2}{2} t} | \mathcal{F}_s \right) = e^{\sigma X_s - \frac{\sigma^2}{2} s}.$$

■

8.4 Exemples d'utilisation de la propriété de martingale

Temps de sortie d'une bande Soit $(X_t, t \geq 0)$ un mouvement brownien. On pose :

$$T = \inf \{ t \geq 0, X_t \notin [a, b] \},$$

avec $a < 0$ et $b > 0$. On peut montrer (par exemple, en utilisant une technique de martingale quasi identique à celle utilisée pour la marche aléatoire symétrique, voir plus bas) que $T < +\infty$ presque sûrement. De plus la continuité des trajectoires de $(X_t, t \geq 0)$ implique que soit $X_T = a$, soit $X_T = b$.

Comme $T \wedge n$ est encore un temps d'arrêt et qu'il est borné, on peut appliquer le théorème d'arrêt pour affirmer que :

$$\mathbf{E}(X_{T \wedge n}) = 0.$$

On peut alors faire tendre n vers l'infini et appliquer le théorème de Lebesgue, puisque $\sup_{n \geq 1} |X_{T \wedge n}| \leq (|a| + |b|)$. On obtient alors :

$$\mathbf{E}(X_T) = 0.$$

En tenant compte du fait que soit $X_T = a$, soit $X_T = b$, on obtient :

$$a\mathbf{P}(X_T = a) + b\mathbf{P}(X_T = b) = 0.$$

Mais comme $\mathbf{P}(X_T = a) + \mathbf{P}(X_T = b) = 1$, on obtient :

$$\mathbf{P}(X_T = a) = \frac{|b|}{|a| + |b|}.$$

Espérance du temps de sortie On a, de plus, pour tout n , en utilisant la propriété de martingale de $(X_t^2 - t, t \geq 0)$:

$$\mathbf{E}(X_{T \wedge n}^2 - T \wedge n) = 0.$$

Soit encore :

$$\mathbf{E}(X_{T \wedge n}^2) = \mathbf{E}(T \wedge n).$$

On peut passer à la limite lorsque n tend vers $+\infty$, en utilisant le théorème de convergence dominée à gauche et le théorème de convergence monotone à droite. On obtient alors :

$$\mathbf{E}(X_T^2) = \mathbf{E}(T).$$

Ceci permet d'évaluer la valeur de $\mathbf{E}(T)$:

$$\mathbf{E}(T) = \frac{|a|^2|b|}{|a| + |b|} + \frac{|b|^2|a|}{|a| + |b|} = |a||b|.$$

Le mouvement brownien atteint tout niveau Soit σ un réel positif, considérons la martingale $(M_t, t \geq 0)$ avec :

$$M_t = e^{\sigma X_t - \frac{\sigma^2}{2}t}.$$

Posons, pour a un réel strictement positif :

$$T_a = \inf\{t > 0, X_t = a\}.$$

On montre (exercice) que T_a est un temps d'arrêt. Noter que T_a n'est pas, a priori, presque sûrement fini. Nous allons appliquer le théorème d'arrêt au temps d'arrêt $T_a \wedge n$ qui est *borné* par n . On obtient :

$$\mathbf{E}\left(e^{\sigma X_{T_a \wedge n} - \frac{\sigma^2}{2}T_a \wedge n}\right) = 1.$$

Faisons tendre n vers $+\infty$. En utilisant le fait que σ et a sont positifs, on obtient que presque sûrement :

$$\lim_{n \rightarrow +\infty} e^{\sigma X_{T_a \wedge n} - \frac{\sigma^2}{2}T_a \wedge n} = \mathbf{1}_{\{T_a < +\infty\}} e^{\sigma X_{T_a} - \frac{\sigma^2}{2}T_a}.$$

Comme, de plus :

$$\left| e^{\sigma X_{T_a \wedge n} - \frac{\sigma^2}{2} T_a \wedge n} \right| \leq e^{\sigma a},$$

on peut appliquer le théorème de Lebesgue pour obtenir :

$$\mathbf{E} \left(\mathbf{1}_{\{T_a < +\infty\}} e^{\sigma X_{T_a} - \frac{\sigma^2}{2} T_a} \right) = 1.$$

Comme $X_{T_a} = a$ lorsque $T_a < +\infty$, on obtient :

$$\mathbf{E} \left(\mathbf{1}_{\{T_a < +\infty\}} e^{\sigma a - \frac{\sigma^2}{2} T_a} \right) = 1.$$

Faisons, maintenant, tendre σ vers 0 en décroissant par valeurs positives, le théorème de convergence monotone permet d'affirmer que :

$$\mathbf{P}(T_a < +\infty) = 1.$$

T_a est donc un temps d'arrêt presque sûrement fini. On peut, de plus, calculer explicitement sa loi. Pour cela il suffit de noter, que comme nous savons que $\mathbf{P}(T_a < +\infty) = 1$, on a :

$$\mathbf{E} \left(e^{-\frac{\sigma^2}{2} T_a} \right) = e^{-\sigma a}.$$

Cette égalité caractérise la loi de T_a et l'on peut en déduire, par transformation de Laplace inverse, que la loi de T_a admet une densité donnée par :

$$\frac{\alpha}{\sqrt{2\pi t^3}} e^{-\frac{\alpha^2}{2t}}.$$

Exercice 48 Donner (sans calcul) la loi de T_{-a} , pour $\alpha > 0$.

8.5 Propriété de Markov forte du mouvement brownien

Nous avons vu que, si s est un réel positif et si $(X_t, t \geq 0)$ est un mouvement brownien, alors $(X_{t+s} - X_s, t \geq 0)$ est un nouveau mouvement brownien indépendant de \mathcal{F}_s . La propriété de Markov forte du mouvement brownien affirme que ce résultat peut se généraliser en remplaçant s par un temps d'arrêt fini arbitraire.

Théorème 8.5.1 Soit $(X_t, t \geq 0)$ un mouvement brownien par rapport à une filtration $(\mathcal{F}_t, t \geq 0)$. Soit S un temps d'arrêt, fini presque sûrement, par rapport à cette filtration ; Alors, le processus :

$$(X_{S+t} - X_S, t \geq 0),$$

est un mouvement brownien par rapport à la filtration $(\mathcal{F}_{S+t}, t \geq 0)$. De plus, ce mouvement brownien est indépendant de la tribu \mathcal{F}_S .

Nous admettrons ce résultat.

Application de ce résultat Soit $(X_t, t \geq 0)$ un mouvement brownien et $T_a = \inf\{t > 0, X_t = a\}$ pour un $a > 0$. Nous avons vu que $T_a < +\infty$ presque sûrement. On peut donc affirmer que $(\tilde{X}_t, t \geq 0)$, avec :

$$\tilde{X}_t = X_{T_a+t} - X_{T_a} = X_{T_a+t} - a,$$

est un nouveau mouvement brownien indépendant de \mathcal{F}_{T_a} .

Considérons $b > a > 0$, on a :

$$T_b = T_a + \inf\{t > 0, X_{T_a+t} = b\}.$$

Ou encore :

$$T_b = T_a + \inf\{t > 0, \tilde{X}_t = b - a\}.$$

On en déduit que :

- la loi de $T_b - T_a$ est identique à celle de T_{b-a} ,
- $T_b - T_a$ est indépendant de la tribu \mathcal{F}_{T_a} .

$(T_a, a \geq 0)$ est donc un processus à accroissements indépendants et stationnaires ! Il ne peut être continu, sinon ce serait un mouvement brownien et ceci lui interdirait d'être croissant ! Nous venons donc d'identifier un exemple de processus à accroissement indépendant et stationnaire non continu.

Loi du maximum du mouvement brownien Soit $(B_t, t \geq 0)$ un mouvement brownien, on note S_t pour :

$$S_t = \sup\{s \leq t, B_s\}.$$

Calculons :

$$\mathbf{P}(S_t \geq a, B_t \leq a),$$

On a :

$$\mathbf{P}(S_t \geq a, B_t \leq a) = \mathbf{P}(T_a \leq t, B_t \leq a).$$

Mais si $\tilde{B}_t = B_{T_a+t} - a$, on obtient :

$$\mathbf{P}(T_a \leq t, B_t \leq a) = \mathbf{P}(T_a \leq t, \tilde{B}_{t-T_a} \leq 0) = \mathbf{E}\left(\mathbf{1}_{\{T_a \leq t\}} \mathbf{P}(\tilde{B}_{t-T_a} \leq 0 | \mathcal{F}_{T_a})\right).$$

Comme T_a est \mathcal{F}_{T_a} -mesurable et $(\tilde{B}_t, t \geq 0)$ est indépendant de \mathcal{F}_{T_a} , on a :

$$\mathbf{P}(\tilde{B}_{t-T_a} \leq 0 | \mathcal{F}_{T_a}) = \mathbf{P}(\tilde{B}_{t-s} \leq 0) \Big|_{s=T_a} = 1/2.$$

De même, on prouve que :

$$\mathbf{P}(-\tilde{B}_{t-T_a} \leq 0 | \mathcal{F}_{T_a}) = \mathbf{P}(-\tilde{B}_{t-s} \leq 0) \Big|_{s=T_a} = 1/2 = \mathbf{P}(\tilde{B}_{t-T_a} \leq 0 | \mathcal{F}_{T_a})$$

D'où :

$$\mathbf{P}(T_a \leq t, B_t \leq a) = \mathbf{P}(T_a \leq t, -\tilde{B}_{t-T_a} \leq 0) = \mathbf{P}(T_a \leq t, B_t \geq a).$$

Que l'on peut réécrire sous la forme :

$$\mathbf{P}(S_t \geq a, B_t \leq a) = \mathbf{P}(S_t \geq a, B_t \geq a).$$

Mais :

$$\mathbf{P}(S_t \geq a, B_t \geq a) = \mathbf{P}(B_t \geq a),$$

Donc, en remarquant que $\mathbf{P}(S_t \geq a, B_t = a) = 0$, on :

$$\mathbf{P}(S_t \geq a) = \mathbf{P}(S_t \geq a, B_t \leq a) + \mathbf{P}(S_t \geq a, B_t \geq a) = 2\mathbf{P}(B_t \geq a) = \mathbf{P}(|B_t| \geq a).$$

S_t et $|B_t|$ ont donc même loi.

8.6 Exercices

Soit $B = (B_t, t \geq 0)$ un mouvement brownien dans \mathbf{R} . On note $(\mathcal{F}_s, s \geq 0)$ où $\mathcal{F}_s = \sigma(B_u, u \leq s)$, la filtration engendrée par le mouvement brownien.

Exercice 49 – Calculer :

$$\text{a) } \mathbf{E}[B_s B_t] \quad \text{b) } \mathbf{E}[e^{\lambda B_t} | \mathcal{F}_s] \quad \text{c) } \mathbf{E}[B_t^n], \quad n \geq 1.$$

On pourra utiliser un développement en série entière de l'exponentielle pour calculer c) après avoir démontré que $\mathbf{E}[e^{\lambda |B_t|}] < \infty$.

– Calculer, en utilisant le fait que $(B_t, B_{t+h}, B_{t+h+s})$ est un vecteur gaussien, la loi conditionnelle de B_{t+h} sachant (B_t, B_{t+h+s}) . Comparer pour f une fonction positive mesurable

$$\mathbf{E}[f(B_{t+h}) | B_t, B_{t+h+s}] \quad \text{et} \quad \int dz f(z) p_h(B_t, z) p_s(z, B_{t+h+s}) / p_{h+s}(B_t, B_{t+h+s}).$$

où $p_u(x, y) = \frac{1}{\sqrt{2\pi u}} e^{-(x-y)^2/2u}$. Interpréter.

– Calculer $\mathbf{P}[B_2 > 0, B_1 > 0]$. Les événements $\{B_2 > 0\}$ et $\{B_1 > 0\}$ sont-ils indépendants ?

Exercice 50 Soit $a > 0$. On note le temps d'arrêt $T = \inf\{t \geq 0; B_t = a\}$ le premier instant où le mouvement brownien B atteint a . On rappelle que $(B_{T+s} - B_T, s \geq 0)$ est un mouvement brownien issu de 0 indépendant de T .

1. Montrer, en utilisant la propriété forte de Markov à l'instant T que

$$\mathbf{P}[T < t, B_t - B_T \leq 0] = \frac{1}{2} \mathbf{P}[T < t].$$

2. On note $S_t = \sup\{B_u; u \leq t\}$. Calculer $\mathbf{P}[S_t > a, B_t \leq a]$. De la même manière calculer $\mathbf{P}[B_t > a] = \mathbf{P}[S_t > a, B_t > a]$.

3. Calculer $\mathbf{P}[S_t > a]$. En déduire que S_t et $|B_t|$ ont même loi sous \mathbf{P} . Les deux processus $(S_t, t \geq 0)$ et $(|B_t|, t \geq 0)$ ont-ils même loi ?

4. Calculer à l'aide d'une intégration par partie la densité de la loi de T sous \mathbf{P} . Calculer $\mathbf{E}[T]$. Conclusion.

Exercice 51 On note $X_t = \frac{1}{\sqrt{1-t}} e^{-B_t^2/2(1-t)}$ pour $t \in [0, 1)$.

1. Soit $0 < t + s < 1$. Calculer $\mathbf{E}[X_{t+s} | \mathcal{F}_t]$. En déduire que $(X_t, t < 1)$ est une martingale.

2. En déduire $\mathbf{E}[X_t]$ pour tout $t < 1$.

3. Montrer que p.s. $\lim_{t \rightarrow 1} X_t = 0$.

4. Pourquoi ne peut-on pas appliquer le théorème d'arrêt ?

Exercice 52 1. Montrer que \mathbf{P} -p.s.

$$\lim_{n \rightarrow \infty} B_n/n = 0.$$

2. On rappelle que, à t fixé, sous \mathbf{P} , $S_t = \sup_{s \leq t} B_s$ a même loi que $|B_t|$. On considère $M_n = \sup\{|B_t - B_n|; n \leq t \leq n+1\}$. Majorer

$$\mathbf{P}[M_n \geq 2\sqrt{\ln n}].$$

En déduire que \mathbf{P} -p.s. $\lim_{n \rightarrow \infty} M_n/n = 0$.

3. Montrer que \mathbf{P} -p.s.

$$\lim_{t \rightarrow \infty} B_t/t = 0.$$

4. On note $Z_0 = 0$ et $Z_t = tB_{1/t}$. Montrer que le vecteur $(Z_{t_1}, Z_{t_2} - Z_{t_1}, \dots, Z_{t_n} - Z_{t_{n-1}})$, où $0 \leq t_1 < \dots < t_n$, est un vecteur gaussien dont on déterminera les caractéristiques. En déduire que $(Z_t, t \geq 0)$ est un mouvement brownien.

Exercice 53 On désire évaluer $\mathbf{P}[\forall s \in [1, t], B_s \neq 0]$.

1. On rappelle que, à t fixé, sous \mathbf{P} , $S_t = \sup_{s \leq t} B_s$ a même loi que $|B_t|$. Soit $t > 1$. En déduire en utilisant la propriété de Markov à l'instant 1, que

$$\mathbf{1}\{B_1 \leq 0\} \mathbf{P}[\exists s \in [1, t], \text{ tel que } B_s = 0 \mid B_1] = \mathbf{1}\{B_1 \geq 0\} \mathbf{P}[|B'_{t-1}| \leq B_1 \mid B_1],$$

où $(B'_t, t \geq 0)$ est un mouvement brownien indépendant de $(B_t, t \geq 0)$.

2. Montrer que

$$\mathbf{P}[\exists s \in [1, t], \text{ tel que } B_s = 0] = \frac{2}{\pi} \int_0^\infty dz \int_{z/\sqrt{t-1}}^\infty e^{-(z^2+y^2)/2} dy.$$

3. Utiliser les coordonnées polaires pour calculer

$$\mathbf{P}[\forall s \in [1, t], B_s \neq 0]$$

Exercice 54 1. Montrer que $\lim_{n \rightarrow \infty} \frac{B_n}{n} = 0$.

2. Montrer, en utilisant le théorème de la limite centrale, que pour tout $a > 0$,

$$\mathbf{P}(\forall n \geq 1, B_n \in [-a, a]) = 0.$$

3. On définit τ_a comme le premier temps de sortie de $[-a, a]$ pour le mouvement brownien :

$$\tau_a = \inf\{t > 0; |B_t| > a\}.$$

Montrer que τ_a est un temps d'arrêt.

4. Déduire des questions précédentes que τ_a est fini p.s. Quelles valeurs peut prendre B_{τ_a} ?

5. Montrer en utilisant la martingale $(e^{\lambda B_t - \frac{\lambda^2}{2}t}, t \geq 0)$ que

$$\begin{aligned} e^{\lambda a} \mathbf{E} \left[\mathbf{1}_{B_{\tau_a} = a} e^{-\frac{\lambda^2}{2} \tau_a} \right] + e^{-\lambda a} \mathbf{E} \left[\mathbf{1}_{B_{\tau_a} = -a} e^{-\frac{\lambda^2}{2} \tau_a} \right] &= 1 \\ e^{-\lambda a} \mathbf{E} \left[\mathbf{1}_{B_{\tau_a} = a} e^{-\frac{\lambda^2}{2} \tau_a} \right] + e^{\lambda a} \mathbf{E} \left[\mathbf{1}_{B_{\tau_a} = -a} e^{-\frac{\lambda^2}{2} \tau_a} \right] &= 1 \end{aligned}$$

6. En déduire que pour tout $\lambda > 0$,

$$\mathbf{E} \left[e^{-\frac{\lambda^2}{2} \tau_a} \right] = \frac{2}{e^{\lambda a} + e^{-\lambda a}}.$$

7. Montrer que

$$\mathbf{P}(\tau_a < 1) \leq \mathbf{E} \left[e^{-\frac{\lambda^2}{2} \tau_a} \right] e^{\frac{\lambda^2}{2}}.$$

En déduire que $\mathbf{P}(\tau_a < 1) \leq 2e^{-\frac{a^2}{2}}$.

8. Soit $(X_k, k \geq 1)$ une suite de variables aléatoires indépendantes et de même loi : celle de $\sup_{t \in [0, 1]} |B_t|$. En utilisant le paragraphe précédent majorer $\mathbf{P}(X_k \geq \sqrt{4 \ln k})$.

9. En déduire que $\mathbf{E} \left[\sum_{k \geq 1} \mathbf{1}_{X_k \geq \sqrt{4 \ln k}} \right]$ est finie. Montrer que p.s. il existe k_0 (aléatoire) tel que pour tout $k \geq k_0$, $X_k \leq \sqrt{4 \ln k}$.

10. Déduire des questions précédentes que $\lim_{t \rightarrow \infty} \frac{B_t}{t} = 0$.

8.7 Travail dirigé : Principe de réflexion

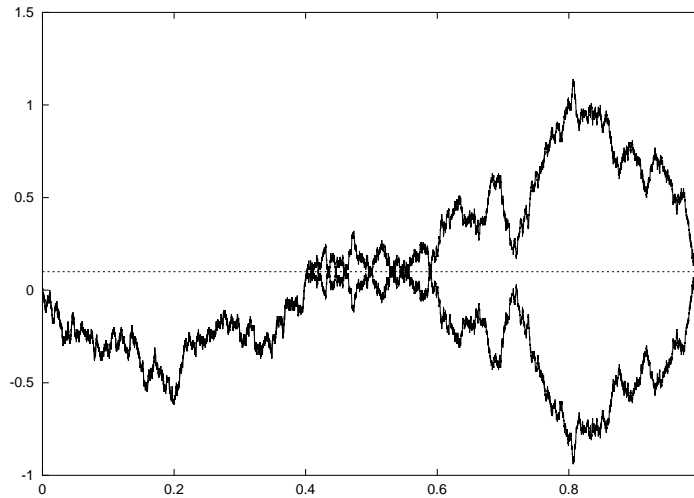


FIG. 8.1 – Principe de réflexion

Soit $B = (B_t, t \geq 0)$ un mouvement brownien standard à valeurs dans \mathbf{R} issu de 0 sous la probabilité \mathbf{P} . On note $S_t = \sup\{B_u; u \in [0, t]\}$.

La loi de S_t On a

$$\begin{aligned} \mathbf{P}[S_t \geq a] &= \mathbf{P}[S_t \geq a, B_t \geq a] + \mathbf{P}[S_t > a, B_t < a] \\ &= \mathbf{P}[B_t \geq a] + \mathbf{P}[S_t > a, B_t < a]. \end{aligned}$$

On note $T_a = \inf\{t \geq 0; B_t = a\}$.

1. Montrer, en utilisant la propriété de Markov forte que $\mathbf{P}[S_t \geq a, B_t < a] = \frac{1}{2} \mathbf{P}[T_a \leq t]$.
2. Vérifier que $\mathbf{P}[S_t \geq a] = \mathbf{P}[T_a \leq t]$.
3. En déduire que à t fixé S_t et $|B_t|$ ont même loi. Les deux processus $|B|$ et $S = (S_t, t \geq 0)$ ont-ils même loi ?
4. Donner un équivalent de $\mathbf{P}[B_s \leq 1; \forall s \leq t]$ quand $t \rightarrow \infty$.
5. Montrer que le processus

$$\begin{aligned} X_t &= B_t \quad \text{si } t \leq T_a \\ X_t &= 2a - B_t \quad \text{si } t > T_a, \end{aligned}$$

est un mouvement brownien.

6. Pour $a \leq b$ et $b > 0$, montrer que

$$\mathbf{P}[S_t > b, B_t < a] = \mathbf{P}[B_t < a - 2b].$$

7. En déduire la loi du couple (S_t, B_t) .
8. Calculer et reconnaître la loi de $S_t - B_t$ à t fixé.
9. Calculer et reconnaître la loi de $2S_t - B_t$ à t fixé.

10. Montrer que T_a et $a^2 T_1$ ont même loi.
11. Montrer que T_1 a même loi que $\frac{1}{|B_1|^2}$.
12. Calculer la densité de la loi de T_a .
13. Calculer la transformée de Laplace de T_a ($\mathbf{E}[e^{-\lambda T_a}]$ pour $\lambda \geq 0$) en utilisant la densité de T_a ou les martingales browniennes exponentielles.

Bibliographie

- [AS70] M. Abramowitz et I.A. Stegun, éditeurs. *Handbook of Mathematical Functions*. Dover, 9th edition, 1970.
- [Bou86] N. Bouleau. *Probabilités de l'Ingénieur*. Hermann, 1986.
- [Bou88] N. Bouleau. *Processus Stochastiques et Applications*. Hermann, 1988.
- [Bre68] Breiman, L. *Probability*. Addison-Wesley, 1968.
- [DCD82] D. Dacunha-Castelle et M. Duflo. *Probabilités et statistiques, tome 1, Problèmes à temps fixe*. Masson, 1982.
- [DCD83] D. Dacunha-Castelle et M. Duflo. *Probabilités et statistiques, tome 2, Problèmes à temps mobile*. Masson, 1983.
- [Fel66] Feller, W. *An introduction to probability theory and its applications. Vol. II*. New York-London-Sydney, 1966.
- [Fel68] Feller, W. *An introduction to probability theory and its applications. I*. New York-London-Sydney, 1968.
- [Gar88] T. Gard. *Introduction to Stochastic Differential Equation*. Marcel Dekker, 1988.
- [GS80] I.I. Gihman et A.V. Skorohod. *Introduction à la Théorie des Processus Aléatoires*. Mir, 1980.
- [Knu81] D.E. Knuth. *The Art of Computer programming, Vol. 2, Seminumerical Algorithms*. Addison-Wesley, 1981.
- [KS88] I. Karatzas et S.E. Shreve. *Brownian Motion and Stochastic Calculus*. Springer-Verlag, New-York, 1988.
- [L'E90] P. L'Ecuyer. Random numbers for simulation. *Communications of the ACM*, 33, 10 1990.
- [Nev72] J. Neveu. *Martingales à temps discret*. Masson, 1972.
- [RW87] L.C.G. Rogers et D. Williams. *Diffusions, Markov Processes and Martingales, Tome 2, Itô Calculus*. John Wiley and Sons, New York, 1987.
- [RY90] A. Revuz et M. Yor. *Continuous Martingale Calculus*. Springer-Verlag, 1990.
- [Wil91] D. Williams. *Probability with martingale*. Cambridge Mathematical textbooks. Cambridge University Press, 1991.

Index

- Convergence
 - en loi, 20
- convergence dominée
 - espérance, 131
 - intégration, 142
 - série, 140
- Espérance conditionnelle, 51
 - Propriétés, 53
- Fonction caractéristique, 14
- Fonction de répartition, 14
- Gaussienne, 14
 - simulation, 18
- Loi forte des grands nombres, 18
- Théorème
 - de convergence monotone, 12
 - de Lebesgue, 13
- Tribu, 9
- Variables aléatoires, 9
- Variance, 15