

Université de Paris-Sud 11

Mémoire de Magistère

Antoine MARCHINA

24 octobre 2014

Introduction

A la fin de ma licence de Mathématiques à l'Université de Versailles, mes encadrants m'ont conseillé et encouragé à intégrer le magistère de l'Université Paris-Sud 11. D'une part puisque Orsay possédait une préparation à l'Agrégation externe (qu'il n'y avait pas à Versailles) et d'autre part pour la qualité de l'enseignement qui y est dispensé. C'est ainsi que je suis arrivé en 2^{ème} année de magistère en 2011.

La première partie de ce mémoire est consacrée au déroulement de ma scolarité au sein du magistère et la seconde à la présentation de mon sujet de thèse, qui a débuté le 1 octobre 2014.

Table des matières

1	Parcours au sein du magistère	4
1.1	Master 1 de Mathématiques Fondamentales et Appliquées 2011-2012	4
1.2	Master 2 Probabilités et Statistiques 2013-2014	5
2	Présentation de mon sujet de thèse – Inégalités de concentration pour les fonctions de variables aléatoires indépendantes	6
	Références	11
A	Annexe : Mémoires rédigés dans le cadre du magistère	13
A.1	TER de M1 : Dynamique de quelques modèles de population .	13
A.2	Apprentissage hors-murs : Marches Aléatoires sur des Réseaux orientés	34
A.3	Mémoire de M2 : Inégalités de Concentration	61

1 Parcours au sein du magistère

1.1 Master 1 de Mathématiques Fondamentales et Appliquées 2011-2012

Souhaitant passer l'Agrégation l'année suivante j'ai choisi de suivre les cours de Mathématiques générales et Analyse qui commençaient en un sens la préparation à ce concours :

- « Mathématiques générales I » par M. Rumin
- « Analyse I » par S. Fischler
- « Mathématiques générales II » par F. Menous
- « Analyse II » par C. Frances

De plus, au premier semestre, j'ai suivi les deux cours fondamentaux :

- « Distributions » par T. Ramond
- « Probabilités » par J-F. Le Gall

Le cours de Probabilités m'a fasciné et c'est celui à l'origine de mon souhait de poursuivre mes études dans ce domaine. J'ai donc poursuivi au second semestre en suivant les cours :

- « Statistiques » par E. Gassiat
- « MAO Probabilités-Statistiques » par C. Lacour

Les deux cours spécifiques au magistère étaient

- « Compléments de théorie spectrale et d'analyse harmonique » par F. Paulin
- « Introduction aux systèmes dynamiques » par P. Pansu

Les deux cours ont été intéressants. Le premier a de plus été très utile pour l'Agrégation. Le deuxième m'a fait découvrir la théorie ergodique que j'ai parfois eu l'occasion de retrouver à divers moments.

Il fallait de plus pour le M1, réalisé un Travail Encadré de Recherche (TER) au second semestre. Désireux d'élargir mes connaissances en Probabilités, j'ai souhaité le faire dans ce domaine. J'ai donc réalisé un mémoire sous la direction de R. Cerf sur le sujet « Dynamique de quelques modèles de population », disponible dans l'Annexe A. Ce travail a été pour moi un premier pas dans le monde de la recherche et a de plus confirmé mon goût pour les Probabilités.

Venant juste d'intégrer le magistère en deuxième année, je devais faire l'apprentissage hors-mur que mes camarades avaient déjà fait lors de leur

première année. J'y ai vu l'occasion de découvrir vraiment le milieu de la recherche. En effet, lors du TER, la lecture d'article se déroulait chez moi ou en bibliothèque et je rencontrais R. Cerf à des intervalles réguliers. Mais je n'étais pas intégré à une équipe de recherche. J'ai alors réalisé cet apprentissage au sein du Laboratoire de Mathématiques de Versailles sous la direction de A. Devulder, en Probabilités sur le sujet « Marches Aléatoires sur des Réseaux Orientés », dont le mémoire est disponible dans l'Annexe A. Et donc pendant trois semaines, j'ai été accueilli dans le bureau des doctorants, intégré à l'équipe. Cet apprentissage a confirmé définitivement mon désir de poursuivre en recherche (dans le domaine des Probabilités) après avoir passé l'Agrégation.

1.2 Master 2 Probabilités et Statistiques 2013-2014

J'ai interrompu le magistère pendant un an afin de préparer l'Agrégation dans le M2 Formation des Professeurs Agrégés. J'ai été reçu à l'Agrégation session 2013. Puis j'ai intégré naturellement le M2 Recherche Probabilités-Statistiques. J'ai suivi les cours :

- « Calcul Stochastique et Processus de Markov » par J.F. Le Gall
- « Grandes Déviations » par T. Bodineau
- « Théorie des processus, Processus de Poisson et Processus de Levy » par C. Giraud
- « Simulation Aléatoire » par E. Gobet
- « Statistiques Asymptotiques » par E. Gassiat
- « Matrices Aléatoires » par E. Maurel-Ségala
- « Processus en Milieu Aléatoire » par A. Singh
- « Lacets et Entrelacs Markoviens » par Y. Le Jan
- « Modèles Stochastiques de Populations : comportement en temps long et limites d'échelles » par S. Méléard et V. Bansaye

Il fallait pour le magistère, suivre un cours supplémentaire par rapport au M2 « classique ». Je souhaitais poursuivre dans la recherche en Probabilités mais je ne savais pas vers quel sous domaine m'orienter. C'est pourquoi je n'ai pas choisi un cours supplémentaire en particulier mais un ensemble de cours (dont le nombre est supérieur à ce qui est nécessaire pour valider) qui me permettait de voir l'étendue de ce qui pouvait exister en Probabilités et un peu en Statistiques.

Pour le mémoire de M2, je souhaitais le réaliser dans l'optique de poursuivre en thèse le sujet. J'ai alors réalisé mon mémoire sur les Inégalités de

Concentration sous la direction de E. Rio dont le mémoire se trouve dans l'Annexe A. Et je vais continuer sur ce thème en thèse dont l'intitulé est « Inégalités de Concentration pour les fonctions de variables aléatoires indépendantes » sous la direction de E. Rio à l'Université de Versailles.

2 Présentation de mon sujet de thèse – Inégalités de concentration pour les fonctions de variables aléatoires indépendantes

Les inégalités de concentration sont des inégalités qui bornent la probabilité de déviation d'une variable aléatoire par rapport à sa moyenne d'une certaine valeur. La recherche sur ce sujet a été très intensive ces dernières décennies du fait des nombreuses applications dans des domaines tels que la théorie de l'apprentissage (machine learning), l'analyse probabiliste d'algorithmes, les graphes aléatoires,...

Les premiers travaux dans ce domaine ont été dans les années 60 sur la somme de variables aléatoires indépendantes avec les travaux de Hoeffding [8] et de Bennett [2] :

Théorème 2.1 ([8], Théorème 2). *Soient X_1, \dots, X_n des variables aléatoires réelles indépendantes telles que $a_i \leq X_i \leq b_i$ p.s. et $S_n = \sum_{i=1}^n X_i$. Alors pour tout $t > 0$, on a*

$$\mathbb{P}(S_n - \mathbb{E}[S_n] \geq t) \leq \exp\left(\frac{-2t^2}{\sum_{i=1}^n (b_i - a_i)^2}\right).$$

Cette inégalité est connue sous le nom d'inégalité de Hoeffding.

Théorème 2.2 (Inégalité de Bennett, [2]). *Soient X_1, \dots, X_n des variables aléatoires réelles indépendantes telles que $X_i \leq b$ p.s. où $b > 0$, $v = \sum_{i=1}^n \mathbb{E}[X_i^2]$ et $S_n = \sum_{i=1}^n X_i$. Alors pour tout $t > 0$,*

$$\mathbb{P}(S_n - \mathbb{E}[S_n] \geq t) \leq \exp\left(-\frac{v}{b^2} h\left(\frac{bt}{v}\right)\right),$$

où $h(u) = (1 + u) \log(1 + u) - u$ pour $u > 0$.

Ces deux inégalités reposent sur des majorations de la transformée de Laplace de S_n . Azuma [1] a étendu l'inégalité de Hoeffding aux martingales

dont les accroissements sont bornés. Ensuite dans les années 70 est apparue la méthode des martingales pour obtenir des inégalités de concentration avec les travaux de Yurinskii [21], Maurey [11] et Milman et Schechtman [14]. Ensuite McDiarmid [12] a pu obtenir à partir de cette méthode et de l'inégalité d'Azuma le résultat suivant :

Théorème 2.3 (Inégalité de McDiarmid, [12], p. 149). *Soient X_1, \dots, X_n des variables aléatoires indépendantes telles que pour tout k , X_k prend ses valeurs dans un espace mesurable (E_k, \mathcal{E}_k) . Soient c_1, \dots, c_n des réels positifs et $f : \prod_{k=1}^n E_k \rightarrow \mathbb{R}$ tels que*

$$|f(x_1, \dots, x_{i-1}, x_i, x_{i+1}, \dots, x_n) - f(x_1, \dots, x_{i-1}, x'_i, x_{i+1}, \dots, x_n)| \leq c_i$$

pour tout i et pour tous x_1, \dots, x_n, x'_i . Soit $Z = f(X_1, \dots, X_n)$. Alors pour tout $x > 0$,

$$\mathbb{P}(|Z - \mathbb{E}[Z]| \geq x) \leq 2 \exp\left(-\frac{2x^2}{\sum_{k=1}^n c_k^2}\right).$$

Par la suite, dans le milieu des années 90, Talagrand [18], [19], [20] introduit une nouvelle méthode appelée méthode d'induction puis Ledoux [9], [10] introduit les méthodes d'entropie reposant sur des inégalités log-Sobolev. Ces méthodes ont permis l'obtention de nouveaux résultats là où la méthode des martingales trouvait certaines limites.

Je vais m'intéresser dans le début de cette thèse à la méthode des martingales. Rappelons cette méthode. On considère X_1, X_2, \dots une suite finie de variables aléatoires indépendantes et de loi commune μ , définies sur un espace de probabilité $(\Omega, \mathcal{F}, \mathbb{P})$ à valeurs dans un espace mesurable (E, \mathcal{E}) . La variable d'intérêt est $Z = f(X_1, \dots, X_n)$ (que l'on peut supposer centrée sans perte de généralité) où f est une fonction à valeurs réelles. Soit $\mathcal{F}_k = \sigma(X_1, \dots, X_k)$ et $\mathcal{F}_0 = \{\emptyset, \Omega\}$. On définit $Z_k = \mathbb{E}[Z | \mathcal{F}_k]$. Maintenant on remarque que $(Z_k)_{0 \leq k \leq n}$ est une martingale (telle que $Z = Z_n$). Ainsi, obtenir des inégalités de concentration pour les martingales nous donnera des inégalités de concentration pour Z .

Comme exemple d'application, Delyon [7] utilise ces techniques et démontre notamment un contrôle sur le maxima de U-statistiques. Par des techniques assez proches, il obtient aussi une inégalité de concentration pour le nombre de triangles (cycles de longueur 3) dans un graphe aléatoire. Précisément, on considère le modèle d'Erdős-Rényi : n sommets et les arêtes sont représentées par des variables aléatoires indépendantes et identiquement distribuées de loi de Bernoulli de paramètre p . On note ces variables aléatoires

Y_{ab} , $1 \leq a < b \leq n$, avec les conventions $Y_{ab} = Y_{ba}$ et $Y_{aa} = 0$. Le nombre de triangle dans ce modèle est alors

$$Z = \sum_{\{a,b,c\}} Y_{ab}Y_{bc}Y_{ac} = \sum_{\tau \in \mathcal{T}} X_{\tau},$$

où \mathcal{T} est de cardinal $\binom{n}{3}$. Delyon obtient une inégalité à droite pour $\mathbb{P}(|Z - \mathbb{E}[Z]| \geq x)$ pour tout réel x . Ce type d'inégalité sur le nombre de triangles (ou autres motifs de sous-graphe) permet l'étude de propriétés du graphe.

Pendant mon mémoire de M2, j'ai étudié l'article de Bentkus [3], où il obtient des inégalités de concentration pour les martingales.

On considère une martingale $M_n = \sum_{i=1}^n X_i$, $M_0 = 0$ telle que $X_k \leq b$ p.s. où $b > 0$ et $\mathbb{E}[X_k^2 | \mathcal{F}_{k-1}] \leq \sigma_k^2$ p.s. où $\sigma_k^2 \geq 0$. La clé pour obtenir une inégalité de concentration repose sur la majoration de la quantité $\mathbb{E}[g(M_n)]$ où g est une fonction positive, croissante et convexe. En effet, par l'inégalité de Markov, on a

$$\mathbb{P}(M_n \geq x) \leq \inf_{g \in \mathcal{F}} \frac{\mathbb{E}[g(M_n)]}{g(x)}$$

pour toute classe \mathcal{F} de fonctions de \mathbb{R} dans \mathbb{R} , positives et croissantes. Puis l'hypothèse de convexité est utilisée pour majorer cette espérance.

Le choix $\mathcal{F} = \mathcal{E}$, où \mathcal{E} est la classe des fonctions exponentielles croissantes $x \in \mathbb{R} \mapsto e^{tx}$, $t > 0$, est classique (appelée méthode de Chernoff; les théorèmes 2.1 et 2.2 reposent dessus). Cependant, en prenant une variable aléatoire gaussienne Z et en comparant avec l'équivalent de la queue d'une gaussienne : $\mathbb{P}(Z \geq x) \sim (x\sqrt{2\pi})^{-1} \exp(-x^2/2)$ quand $x \rightarrow \infty$, il « manque » un facteur d'ordre $1/x$ pour des $x > 0$ grand. La raison de ce phénomène est due au fait que la classe \mathcal{E} est trop petite. On considère alors les classes \mathcal{H}^α (qui contiennent \mathcal{E}), où $\alpha > 0$, définie par : f est dans \mathcal{H}^α si et seulement si il existe une mesure borélienne positive μ_f telle que pour tout réel u ,

$$f(u) = \int (u - t)_+^\alpha \mu_f(dt).$$

Bentkus obtient la borne supérieure exacte pour l'espérance de $f(M_n)$ avec $f \in \mathcal{H}^2$ et prouve par suite une inégalité de concentration pour les martingales. Avant d'énoncer le résultat, introduisons des notations. Soit $a < 0 < b$, $\mathcal{B}(a, b)$ désigne la loi de Bernoulli prenant la valeur a avec probabilité $b/(b-a)$ et la valeur b avec probabilité $-a/(b-a)$. Pour une variable aléatoire réelle, la fonction $\mathbb{P}^\circ(X \geq \cdot)$ correspond à la fonction minimale log-concave majorant la fonction de survie $\mathbb{P}(X \geq \cdot)$

Théorème 2.4 ([3], Lemme 4.4). *Supposons que les accroissements X_k d'une martingale M_n vérifient $\mathbb{P}(X_k \leq b) = 1$ et $\mathbb{P}(s_k^2 \leq \sigma_k^2) = 1$ où $s_k^2 = \mathbb{E}[X_k^2 | \mathcal{F}_{k-1}]$, $b > 0$, $\sigma_k^2 \geq 0$. Alors pour tout $x \in \mathbb{R}$,*

$$\mathbb{P}(M_n \geq x) \leq \frac{e^2}{2} \mathbb{P}^\circ(S_n \geq x), \quad (2.1)$$

où $S_n = \sum_{i=1}^n \varepsilon_i$ où les ε_i sont des variables aléatoires indépendantes et identiquement distribuées de loi $\mathcal{B}(-\sigma^2/b, b)$ avec $\sigma^2 = \sum_{i=1}^n \sigma_i^2/n$.

De plus, on a

$$\mathbb{P}(M_n \geq x) \leq \frac{e^2}{2} \mathbb{P}^\circ\left(\eta \geq \lambda + \frac{x}{b}\right), \quad (2.2)$$

où $\eta \sim \mathcal{P}(\lambda)$ avec $\lambda = \sum_{i=1}^n \sigma_i^2/b^2$.

Bentkus obtient également deux autres résultats de concentration (du même ressort que ci-dessus) en changeant les hypothèses de majoration sur les accroissements. Il considère les cas suivants :

- $\mathbb{P}(-p_k \leq X_k \leq 1 - p_k) = 1$ où $0 \leq p_k \leq 1$ pour $k = 1, \dots, n$,
- $\mathbb{P}(X_k \leq b_k) = 1$ où $b_k \geq 0$ et $\mathbb{P}(s_k^2 \leq \sigma_k^2) = 1$ où $\sigma_k^2 \geq 0$ pour $k = 1, \dots, n$.

Pinelis ([17], Théorème 2.1 et Corollaire 2.2) considère la classe de fonctions \mathcal{H}^3 et ajoute une condition sur le troisième moment pour obtenir un résultat de comparaison pour $\mathbb{E}[f(M_n)]$ avec $f \in \mathcal{H}^3$, puis en déduit le théorème suivant.

Théorème 2.5. *Supposons que les accroissements X_k d'une martingale M_n vérifient $\mathbb{P}(X_k \leq b) = 1$, $\mathbb{P}(s_k^2 \leq \sigma_k^2) = 1$ où $s_k^2 = \mathbb{E}[X_k^2 | \mathcal{F}_{k-1}]$, $b > 0$, $\sigma_k^2 \geq 0$.*

Supposons de plus que $\mathbb{P}(t_k^3 \leq \beta_k) = 1$ où $t_k^3 = \mathbb{E}[(X_k)_+^3 | \mathcal{F}_{k-1}]$, $\beta_k > 0$.

Soit $\beta = \sum_{k=1}^n \beta_k/n$, $\sigma^2 = \sum_{k=1}^n \sigma_k^2/n$ et $\varepsilon = \beta/b\sigma^2$. Alors pour tout $x \in \mathbb{R}$,

$$\mathbb{P}(M_n \geq x) \leq \frac{2e^3}{9} \mathbb{P}^\circ\left(\Gamma_{(1-\varepsilon)n\sigma^2} + b\tilde{\Pi}_{\varepsilon\lambda} \geq x\right),$$

où $\lambda = \sum_{i=1}^n \sigma_i^2/b^2$, Γ_{a^2} désigne la loi normale centrée de variance a^2 et $\tilde{\Pi}_\theta$ désigne la loi de Poisson recentrée de paramètre θ .

Ce résultat est un raffinement du résultat de Bentkus. Cependant la constante $2e^3/9$ est moins bonne que $e^2/2$.

Pinelis [15], [16] et Bentkus [4] prouvent des résultats de comparaison pour $\mathbb{E}[M_n]$ où M_n est une martingale à accroissements bornés et obtiennent

par suite des comparaisons entre la probabilité de queue d'une martingale et celle d'une gaussienne :

$$\mathbb{P}(M_n \geq x) \leq c \mathbb{P}(sZ \geq x)$$

où c est une constante et s provient des hypothèses de majoration que l'on fait sur les accroissements de la martingale.

Pendant ma thèse, je vais alors chercher à appliquer ces résultats de concentration sur les martingales à des fonctions de variables aléatoires indépendantes pour lesquelles les accroissements de la martingales dans la décomposition martingale (rappelée ci-dessus) vérifient ces hypothèses de majoration afin d'obtenir de nouveaux résultats.

Enfin le cas de fonctions lipschitziennes de variables aléatoires indépendantes non bornées est encore mal connu. Citons par exemple les maxima de processus empiriques indexés par des classes de fonctions non bornées. Précisément, on note

$$S_n(f) = \sum_{i=1}^n f(X_i),$$

où $f \in \mathcal{F}$ et \mathcal{F} est une famille dénombrable de fonctions mesurables de E dans \mathbb{R} . La variable d'intérêt est

$$Z = \sup \{S_n(f) \mid f \in \mathcal{F}\}.$$

Les méthodes de Talagrand [19] et Ledoux [9] ont montré leur efficacité afin d'étudier la concentration de Z autour de sa moyenne quand les fonctions de \mathcal{F} sont bornées (possiblement que d'un seul côté). Quand les fonctions ne sont plus bornées, les résultats sont moins performants. Boucheron, Lugosi, Massart ([5], p. 443–446) donnent des estimations des moments $(Z - \mathbb{E}[Z])_+$, mais avec des constantes qui ne semblent pas optimales.

Par ailleurs, pour les inégalités de concentration avec vitesse polynomiale, des résultats récents de Chazottes et Gouëzel [6] montrent l'intérêt des méthodes de martingale dans le cas des systèmes dynamiques ou de chaînes de Markov.

Références

- [1] K. Azuma. Weighted sums of certain dependent random variables. *Tohoku Mathematical Journal*, 19(3) :357–367, 1967.
- [2] G. Bennett. Probability inequalities for the sum of independent random variables. *Journal of the American Statistical Association*, 57(297) :33–45, 1962.
- [3] V. Bentkus. On hoeffding’s inequalities. *The Annals of Probability*, 32(2) :1650–1673, 2004.
- [4] V. Bentkus. On measure concentration for separately lipschitz functions in product spaces. *Israel Journal of Mathematics*, 158(1) :1–17, 2007.
- [5] S. Boucheron, G. Lugosi, and P. Massart. *Concentration Inequalities : A Nonasymptotic Theory of Independance*. Oxford Univ. Press, 2013.
- [6] Jean-René Chazottes and Sébastien Gouëzel. Optimal concentration inequalities for dynamical systems. *Communications in Mathematical Physics*, 316(3) :843–889, 2012.
- [7] B. Delyon. Exponential inequalities for sums of weakly dependent variables. *Electronic Journal of Probability*, 14(28) :52–779, 2009.
- [8] W. Hoeffding. Probability inequalities for sums of bounded random variables. *Journal of the American Statistical Association*, 58(301) :13–30, March 1963.
- [9] M. Ledoux. Isoperimetry and gaussian analysis. In P. Bernard, editor, *Lectures on Probability Theory and Statistics*, pages 165–294. École d’Été de Probabilités de St Flour XXIV-1994, 1996.
- [10] M. Ledoux. On Talagrand’s deviation inequalities for product measures. *ESAIM : Probability and Statistics*, (1) :63–87, 1997.
- [11] B. Maurey. Construction de suites symétriques. *CRAS Paris Série I Mathématiques*, (288) :679–681, 1979.
- [12] C. McDiarmid. On the method of bounded differences. *Survey in Combinatorics*, pages 148–188, 1989.
- [13] C. McDiarmid. Concentration. In M. Habib, C. McDiarmid, and J. Ramirez-Alfonsin, editors, *Probabilistic Methods for Algorithmic Discrete Mathematics*, pages 195–248. Springer, 1998.
- [14] V.D. Milman and G. Schechtman. *Asymptotic Theory of Finite Dimensional Normed Spaces : Isoperimetric Inequalities in Riemannian Manifolds*, volume 1200 of *Lecture Notes in Mathematics*. Springer, 1986.

- [15] I. Pinelis. Optimal tail comparison based on comparison of moments. In Ernst Eberlein, Marjorie Hahn, and Michel Talagrand, editors, *High Dimensional Probability*, volume 43 of *Progress in Probability*, pages 297–314. Birkhauser, 1998.
- [16] I. Pinelis. On normal domination of (super)martingales. *Electronic Journal of Probability*, 11(39) :1049–1070, 2006.
- [17] I. Pinelis. On the bennett-hoeffding inequality. *Annales de l'Institut Henri Poincaré, Probabilités et Statistiques*, 50(1) :15–27, 2014.
- [18] M. Talagrand. Concentration of measure and isoperimetric inequalities in product spaces. *Publications Mathématiques de l'I.H.É.S.*, (81) :73–205, 1995.
- [19] M. Talagrand. New concentration inequalities in product spaces. *Inventiones Mathematicae*, (126) :505–563, 1996.
- [20] M. Talagrand. A new look at independance. *Annals of Probability*, (24) :1–34, 1996.
- [21] V.V. Yurinskii. Exponential inequalities for sums of random vector. *Journal of Multivariate Analysis*, (6) :472–499, 1976.

A Annexe : Mémoires rédigés dans le cadre du magistère

A.1 TER de M1 : Dynamique de quelques modèles de population

Université de Paris-Sud 11

TER de Mathématiques de 1^{ère} année de Master

Dynamique de quelques modèles de population

Antoine MARCHINA

8 juin 2012

Introduction

Ce TER est essentiellement basé sur l'article d'Amaury Lambert [3].

On étudie la dynamique de 3 différents modèles classiques de population, à savoir le modèle de Cannings, celui de Wright-Fisher (qui est un cas particulier de celui de Cannings) et enfin le modèle de Bienaymé-Galton-Watson (BGW). Ce sont des modèles stochastiques à temps discret, les 2 premiers sont à taille constante et celui de BGW à une population dont la taille varie aléatoirement (modèle de branchement).

La dernière partie consiste à voir des liens entre ces 2 différents types de modèles (taille fixe et modèle de branchement).

Je remercie Raphaël Cerf pour son encadrement durant ce projet.

Table des matières

1	Modèle de Cannings	4
2	Modèle de Wright-Fisher	8
3	Modèle de Bienaymé-Galton-Watson	9
3.1	Généralités sur les fonctions génératrices	10
3.2	Calcul de la probabilité d'extinction	11
3.3	Quelques remarques sur la convergence	14
4	Deux relations entre modèles de branchement et modèles à taille fixe	18
	Références	20

1 Modèle de Cannings

Le modèle de Cannings est un modèle stochastique de généalogie – en temps discret – d’une population de taille finie et fixée.

Fixons la taille de la population à $2N$, où N est un entier strictement positif. On peut interpréter ceci comme une population de diploïde de taille constante N dont on étudie la proportion de 2 allèles A et a dans un même locus. On traite les N individus comme $2N$ copies du locus sans se soucier de les rassembler par paire pour considérer des individus. Chaque individu n’a qu’un parent et à chaque étape, les $2N$ individus sont indexés aléatoirement $i = 1, \dots, 2N$. Pour tout i , l’individu i de la génération n engendre un nombre η_i d’enfants.

Définition 1.1. *La dynamique du modèle de Cannings est définie par les règles suivantes*

- La génération $n + 1$ provient des enfants de la génération n
- La taille de la population est constante, i.e $\sum_i \eta_i = 2N$
- La loi du $2N$ -uplet $(\eta_1, \dots, \eta_{2N})$ est échangeable

On rappelle la notion d’échangeabilité :

Définition 1.2. *La loi de (X_1, \dots, X_n) est échangeable si pour tout $\sigma \in \mathfrak{S}_n$,*

$$(X_1, \dots, X_n) \stackrel{\text{loi}}{=} (X_{\sigma(1)}, \dots, X_{\sigma(n)})$$

Considérons une sous-population de taille $Y_0 = y$ et on note Y_n le nombre de descendants de cette sous-population au temps n . Par exemple, on considère à partir d’un certain instant, la sous-population porteur de l’allèle a . On a ainsi directement la proposition suivante :

Proposition 1.3. *Le processus $(Y_n; n \geq 0)$ est une chaîne de Markov sur l’espace d’états $\{0, \dots, 2N\}$ avec 2 états absorbants : 0 et $2N$.*

Notons τ le temps d’absorption :

$$\tau = \inf \{n \geq 0 \mid Y_n \in \{0, 2N\}\}$$

Si on exclut le cas trivial où chaque individu engendre exactement 1 enfant (i.e. pour tout i , $\eta_i = 1$ p.s.), on a $\tau < \infty$ p.s. (car les 2 états absorbants sont accessibles).

Remarquons dès maintenant que τ est un temps d'arrêt. En effet, en notant $(\mathcal{F}_n)_{n \geq 0}$ la filtration canonique associée au processus $(Y_n)_{n \geq 0}$, pour tout entier positif n on a

$$\begin{aligned} \{\tau \leq n\} &= \{\inf \{k \mid Y_k \in \{0, 2N\}\} \leq n\} \\ &= \bigcup_{k=1}^n \{Y_k \in \{0, 2N\}\} \in \mathcal{F}_n \end{aligned}$$

Définition 1.4. *En toute généralité, l'événement $\{Y_\tau = 0\}$ est appelé Extinction et noté $\{Ext\}$, tandis que $\{Y_\tau = 2N\}$ est appelé Fixation et noté $\{Fix\}$.*

En reprenant notre interprétation avec les allèles, la fixation désigne le moment où toute la population n'est porteur que de l'allèle a (i.e. l'allèle A a totalement disparu de la population).

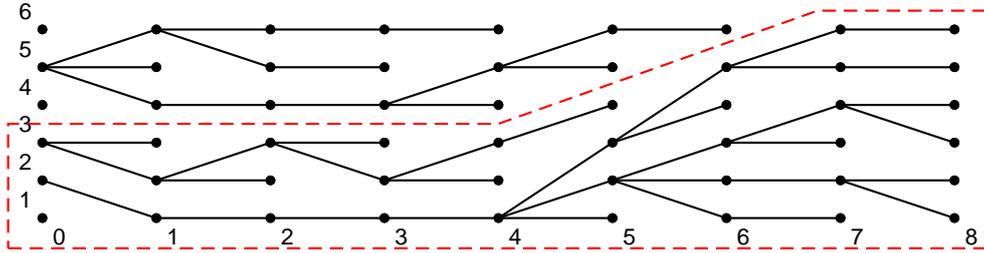


FIGURE 1 – Représentation du modèle de Cannings pour $2N = 6$ sur 8 étapes. Si on considère que les individus 1, 2 et 3 possèdent l'allèle a et les autres l'allèle A , on remarque qu'il y a fixation de l'allèle a dans la population à partir du temps 7.

Proposition 1.5. *La chaîne de Markov $(Y_n; n \geq 0)$ est une martingale (pour la filtration canonique associée à Y_n), et on a :*

$$\mathbb{P}_y(Fix) = \frac{y}{2N}$$

DÉMONSTRATION. Tout d'abord, l'échangeabilité entraîne que les η_i sont identiquement distribués et de plus comme $\sum_i \eta_i = 2N$, on doit avoir $\mathbb{E}(\eta_1) = 1$ car

$$\mathbb{E}\left(\sum_i \eta_i\right) = \sum_i \mathbb{E}(\eta_i) = 2N\mathbb{E}(\eta_1)$$

Ensuite on remarque que conditionnellement à Y_n , et toujours par échangeabilité, on a

$$Y_{n+1} \stackrel{\text{loi}}{=} \sum_{i=1}^{Y_n} \eta_i$$

On a donc $\mathbb{E}_y[Y_{n+1}|\mathcal{F}_n] = Y_n$. Ce qui prouve que $(Y_n)_{n \geq 0}$ est une martingale.

On peut alors appliquer le théorème d'arrêt à la martingale bornée (Y_n) et on obtient

$$\mathbb{E}_y(Y_0) = y = \mathbb{E}_y(Y_\tau) = \mathbb{E}_y(0 \cdot \mathbf{1}_{Y_\tau=0} + 2N \cdot \mathbf{1}_{Y_\tau=2N}) = 2N\mathbb{P}_y(\text{Fix}) \quad \square$$

Introduisons maintenant la matrice de transition de la chaîne de Markov Y_n que l'on note P . Donc P est une matrice carrée de taille $2N + 1$ et ses éléments sont :

$$P_{yz} = \mathbb{P}_y(Y_1 = z) \quad 0 \leq y, z \leq 2N$$

Avant d'énoncer le prochain théorème, rappelons-nous que la valeur propre dominante d'une matrice de transition est toujours 1.

En effet, en toute généralité considérons λ une valeur propre de Q , matrice de transition (stochastique), de vecteur propre associé $v = (v_i)_{1 \leq i \leq n}$. Comme $Qv = \lambda v$, on a alors pour tout $0 \leq i \leq n$:

$$|\lambda v_i| = \left| \sum_{j=1}^n q_{ij} v_j \right| \leq \sup_{0 \leq j \leq 2N} |v_j| \sum_{j=1}^n q_{ij} = \sup_{0 \leq j \leq 2N} |v_j|$$

Et donc $\lambda \leq 1$. On a égalité en considérant le vecteur $v = (1, \dots, 1)$.

Remarquons qu'ici e_0 et e_{2N} sont des vecteurs propres gauches de P (i.e. vecteurs propres de la transposée de P) associés à la valeur propre gauche 1. Et comme les valeurs propres gauches et les valeurs propres (droites) sont les mêmes (en considérant par exemple le polynôme caractéristique de P qui est égal à celui de tP). Ainsi 1 est valeur propre de P de multiplicité au moins 2.

Théorème 1.6 (Cannings). *Les valeurs propres de P rangées dans l'ordre décroissant sont $\lambda_0 = \lambda_1 = 1$ et*

$$\lambda_j = \mathbb{E}(\eta_1 \eta_2 \dots \eta_j) \quad 1 \leq j \leq 2N$$

DÉMONSTRATION. Considérons Z la matrice carrée de taille $2N + 1$ suivante :

$$Z = \begin{pmatrix} 1 & 0 & 0 & 0 & \dots & 0 \\ 1 & 1 & 1^2 & 1^3 & \dots & 1^{2N} \\ 1 & 2 & 2^2 & 2^3 & \dots & 2^{2N} \\ 1 & 3 & 3^2 & 3^3 & \dots & 3^{2N} \\ \vdots & \vdots & \vdots & \vdots & \dots & \vdots \\ 1 & 2N & (2N)^2 & (2N)^3 & \dots & (2N)^{2N} \end{pmatrix}$$

Z est inversible (matrice de Vandermonde et coefficients 2 à 2 distincts). On va montrer qu'il existe une matrice triangulaire supérieure A telle que $P = ZAZ^{-1}$ dont les éléments diagonaux sont $a_{00} = 1$ et

$$a_{jj} = \mathbb{E}(\eta_1 \eta_2 \dots \eta_j) \quad 1 \leq j \leq 2N$$

Ainsi les valeurs propres de Z sont les mêmes que celles de A (qui sont ses éléments diagonaux), ce qui finirait la preuve.

En réécrivant l'équation comme $PZ = ZA$, on cherche des coefficients $(a_{ij})_{i \leq j}$ tel que

$$(PZ)_{ij} = \sum_{k=0}^{2N} p_{ik} k^j = \sum_{k=0}^j i^k a_{kj} = (ZA)_{ij}$$

Comme P est la matrice de transition de Y , on a,

$$\begin{aligned} (PZ)_{ij} &= \mathbb{E}(Y_{n+1}^j \mid Y_n = i) \\ &= \mathbb{E}(Y_1^j \mid Y_0 = i) && \text{propriété de Markov} \\ &= \mathbb{E}\left((\eta_1 + \dots + \eta_i)^j\right) && \text{définition de } (Y_n)_{n \geq 0} \end{aligned}$$

De plus, la formule du multinôme nous donne grâce à la propriété d'échangeabilité du modèle

$$\mathbb{E}\left((\eta_1 + \dots + \eta_i)^j\right) = i \mathbb{E}(\eta_1) + \dots + i^{[j]} \mathbb{E}(\eta_1 \dots \eta_j)$$

avec $i^{[j]} = i(i-1)\dots(i-j+1)$. En remarquant que l'on peut écrire

$$\sum_{k=0}^j a_{kj} i^k = a_{jj} i^{[j]} + \text{termes en } i^{j-1}, \dots, i^0$$

On peut ainsi trouver les coefficients a_{ij} de la matrice A et telle que ses éléments diagonaux soient biens $a_{jj} = \mathbb{E}(\eta_1 \dots \eta_j)$, $j = 0, \dots, 2N$. \square

2 Modèle de Wright-Fisher

Le modèle de Wright-Fisher (WF) est un cas particulier du modèle de Cannings : il correspond au cas où $(\eta_1, \dots, \eta_{2N})$ suit une loi multinomiale de paramètres $(2N; 1/2N, \dots, 1/2N)$. C'est-à-dire que, si on a $\alpha_1 + \dots + \alpha_{2N} = 2N$, alors

$$\mathbb{P}(\eta_1 = \alpha_1, \dots, \eta_{2N} = \alpha_{2N}) = \frac{(2N)!}{\alpha_1! \dots \alpha_{2N}!} \frac{1}{(2N)^{2N}}$$

On peut interpréter une loi multinomiale comme le résultat d'un tirage avec remise de $2N$ boules (dans ce cas ci) dans une urne ayant $2N$ boules différentes. On peut donc également définir le modèle ainsi : Chaque individu de la génération $n+1$ choisit son père parmi la génération n de façon uniforme et indépendante.

Proposition 2.1. *La matrice de transition associée à la chaîne de Markov Y est donnée par*

$$P_{yz} = \binom{2N}{z} \left(\frac{y}{2N}\right)^z \left(1 - \frac{y}{2N}\right)^{2N-z}$$

C'est-à-dire que conditionnellement à $Y_n = y$, Y_{n+1} suit une loi binomiale :

$$Y_{n+1} \stackrel{\text{loi}}{=} \text{Bin}\left(2N, \frac{y}{2N}\right)$$

DÉMONSTRATION. Pour tout n dans \mathbb{N} , Y_n prend bien ses valeurs dans $0, \dots, 2N$. De plus on a facilement par définition que

$$\mathbb{P}(Y_{n+1} = 0 \mid Y_n = 0) = \mathbb{P}(Y_{n+1} = 2N \mid Y_n = 2N) = 1$$

Prenons maintenant $y \in \{1, \dots, 2N-1\}$. Il suffit de considérer la deuxième manière de définir le modèle (avec le tirage aléatoire), on voit que conditionnellement à $Y_n = y$, Y_{n+1} suit une loi binomiale de paramètres $(2N, y/2N)$; i.e.

$$\mathbb{P}(Y_{n+1} = z \mid Y_n = y) = P_{yz} = \binom{2N}{z} \left(\frac{y}{2N}\right)^z \left(1 - \frac{y}{2N}\right)^{2N-z} \quad \square$$

3 Modèle de Bienaymé-Galton-Watson

Donnons-nous une v.a. ξ à valeurs dans \mathbb{N} et une suite de réels positifs ou nuls $(p_k)_{k \in \mathbb{N}}$ telle que pour tous $k \in \mathbb{N}$, $p_k = \mathbb{P}(\xi = k)$. On suppose de plus que p_0 et p_1 sont différents de 0 et 1.

La taille de la population au temps n est notée Z_n . Et à chaque temps n , on note aléatoirement les individus de la population par $i = 1, \dots, Z_n$.

Définition 3.1. *La dynamique du modèle de Bienaymé-Galton-Watson (BGW) est définie par les règles suivantes*

- La génération $n + 1$ provient des enfants de la génération n
- Conditionnellement à Z_n , pour tout $i \in \llbracket 1, 2Z_n \rrbracket$, l'individu i de la génération n engendre un nombre ξ_i d'enfants
- Les ξ_i sont indépendants et de même loi que ξ

Remarquons que le deuxième point est équivalent à dire que conditionnellement à Z_n , on a

$$Z_{n+1} \stackrel{\text{loi}}{=} \sum_{i=1}^{Z_n} \xi_i \quad \text{et} \quad Z_0 = z$$

Proposition 3.2. *Le processus $(Z_n; n \geq 0)$ est une chaîne de Markov de matrice de transition Q donnée pour tous $x, y \in \mathbb{N}$ par*

$$Q(x, y) = \mu^{*x}(y)$$

où μ est la loi de ξ , $\mu^{*0} = \delta_0$ et $\mu^x = \underbrace{\mu * \dots * \mu}_{x \text{ fois}}$ si $x \geq 1$

DÉMONSTRATION. On a par indépendance des ξ_i

$$\begin{aligned} \mathbb{P}(Z_{n+1} = y \mid Z_0 = z, \dots, Z_n = z_n) &= \mathbb{P}\left(\sum_{i=1}^{Z_n} \xi_i = y \mid Z_0 = z, \dots, Z_n = z_n\right) \\ &= \mathbb{P}\left(\sum_{i=1}^{z_n} \xi_i = y \mid Z_0 = z, \dots, Z_n = z_n\right) \\ &= \mathbb{P}\left(\sum_{i=1}^{z_n} \xi_i = y\right) \\ &= Q(z_n, y) \end{aligned}$$

où $Q(x, \cdot)$ est la loi de $\sum_{i=1}^x \xi_i$ est qui est bien la formule donnée. \square

On peut interpréter ce modèle ainsi : On considère z particules au départ et chacune d'elles engendre k enfants avec probabilité p_k indépendamment les uns des autres. Puis au temps suivant, chacun des enfants engendre de même \tilde{k} enfants avec probabilité $p_{\tilde{k}}$ indépendamment les uns des autres et ainsi de suite.

On considère des processus qui débutent avec $Z_0 = z$ individus. Remarquons que l'on peut ramener l'étude d'un tel processus à l'étude d'un processus débutant avec un seul individu grâce à la propriété suivante :

Propriété 3.3 (des processus de branchement). *Si $Z(x)$ désigne un processus de BGW qui débute avec $Z_0 = x$ individus, on a*

$$Z(x + y) \stackrel{\text{loi}}{=} Z(x) + \tilde{Z}(y)$$

où \tilde{Z} est une copie indépendante de Z .

DÉMONSTRATION. En effet, supposons que $Z_0 = x + y$ où on désigne par x (resp. y) les x premiers (resp. y derniers) individus. On peut écrire $Z_n = Z_n^{(x)} + Z_n^{(y)}$ où $Z_n^{(x)}$ (resp. $Z_n^{(y)}$) est le nombre d'individu vivant à l'instant n et descendant des x (resp. y) individus initiaux. \square

3.1 Généralités sur les fonctions génératrices

Avant de poursuivre l'étude de ce modèle, rappelons quelques propriétés des fonctions génératrices de v.a. qui nous serviront par la suite. Les définitions, propriétés, théorèmes.. seront directement adaptés à notre modèle.

Définition 3.4. *La fonction génératrice de ξ est la fonction f définie sur $[0, 1]$ par*

$$f(s) = \mathbb{E} \left[s^\xi \right] = \sum_{k \geq 0} p_k s^k$$

Proposition 3.5. *Sur $[0, 1]$, la fonction f est indéfiniment dérivable. De plus f admet une dérivée à gauche en 1 si et seulement si $\mathbb{E}[\xi] < \infty$. Et dans ce cas $f'(1) = \mathbb{E}[\xi]$.*

DÉMONSTRATION. $z \mapsto f(z)$ est la somme d'une série entière de rayon de convergence supérieur ou égal à 1 (car $\sum_{k \geq 0} p_k = 1$). Donc cette fonction est holomorphe sur le disque unité ouvert et y est indéfiniment dérivable. Soit $s \in [0, 1[$, on a

$$\frac{f(1) - f(s)}{1 - s} = \frac{1}{1 - s} \left(\sum_{k \geq 0} p_k - \sum_{k \geq 0} p_k s^k \right) = \frac{1}{1 - s} \left(\sum_{k \geq 0} p_k (1 - s^k) \right)$$

Or pour tout $k \in \mathbb{N}$, $s \mapsto \frac{1 - s^k}{1 - s} = 1 + s + \dots + s^{k-1}$ est croissante et tend vers k quand s tend vers 1 (par valeurs inférieures). Donc le théorème de convergence monotone fournit :

$$\lim_{\substack{s \rightarrow 1 \\ s < 1}} \frac{f(1) - f(s)}{1 - s} = \sum_{k \geq 0} k p_k = \mathbb{E}[\xi] \quad \square$$

Proposition 3.6. *f est une fonction croissante et convexe telle que $f(1) = 1$.*

Et donc f possède au plus 2 points fixes dans $[0, 1]$.

DÉMONSTRATION. La croissance est directement donnée par la définition car f est une somme de fonctions de s strictement croissantes. Pour tout $s \in [0, 1[$, $f'(s) = \sum_{k \geq 1} k s^{k-1} p_k$ qui est également une somme de fonctions de s strictement croissantes donc f est strictement convexe. Maintenant comme elle est strictement convexe et que $f(1) = 1$ il existe au plus un autre point fixe selon que f coupe la diagonale du carré (i.e. $f'(1) > 1$) ou non (i.e. $f'(1) \leq 1$). \square

3.2 Calcul de la probabilité d'extinction

On note comme précédemment \mathbb{P}_z la probabilité conditionnelle $\mathbb{P}(\cdot | Z_0 = z)$. (\mathbb{P} désigne \mathbb{P}_1).

Proposition 3.7. *La fonction génératrice de Z_n est donnée par*

$$\mathbb{E}_z \left[s^{Z_n} \right] = f_n(s)^z \quad s \in [0, 1]$$

où f_n désigne la n -ième itération de f avec elle-même.

En particulier on a

$$\mathbb{E} [Z_n | Z_0 = z] = m^n z$$

DÉMONSTRATION. On a pour tout entier k non nul, en utilisant l'indépendance des v.a. ξ_i :

$$\begin{aligned}\mathbb{E} \left[s^{Z_{n+1}} \mid Z_n = k \right] &= \mathbb{E} \left[s^{\sum_{i=1}^k \xi_i} \right] \\ &= \prod_{i=1}^k s^{\xi_i} \\ &= f(s)^k\end{aligned}$$

Donc $\mathbb{E} \left[s^{Z_{n+1}} \mid Z_n \right] = f(s)^{Z_n}$. Et ainsi on a

$$\mathbb{E} \left[s^{Z_{n+1}} \right] = \mathbb{E} \left[f(s)^{Z_n} \right]$$

On finit de montrer le résultat par récurrence.

Au rang 0, on a bien $\mathbb{E}_z \left[s^{Z_0} \right] = \mathbb{E}_z \left[s^z \right] = s^z = f_0(s)^z$.

Supposons que l'on ait démontré la propriété jusqu'au rang $n \in \mathbb{N}$. On a alors en se rappelant que f est à valeur dans $[0, 1]$,

$$\mathbb{E}_z \left[s^{Z_{n+1}} \right] = \mathbb{E}_z \left[f(s)^{Z_n} \right] = \left(f_n(f(s)) \right)^z = f_{n+1}(s)^z$$

Dérivons maintenant l'expression obtenue sur $[0, 1[$. On a

$$\mathbb{E} \left[Z_n s^{Z_n-1} \right] = z \cdot f_n(s)^{z-1} \cdot \prod_{k=0}^{n-1} \left[f'(f_k(s)) \right]$$

Il suffit alors de faire tendre s vers 1 pour terminer la preuve. \square

Désignons par q le premier point fixe de f sur $[0, 1]$. C'est-à-dire

$$q = \inf \{ s \in [0, 1] \mid f(s) = s \}$$

Théorème 3.8. 1. Sur l'événement de non-extinction,

$$\lim_{n \rightarrow \infty} Z_n = +\infty \quad p.s.$$

2. De plus,

- si $m \leq 1$, $q = 1$ et $f_n(s)$ croît vers 1 pour tout $s \in [0, 1]$ quand n tend vers l'infini.
- si $m > 1$, $q \in [0, 1[$ et $f_n(s)$ croît (resp. décroît) vers q pour tout $s \in [0, q]$ (resp. $s \in]q, 1]$) quand n tend vers l'infini.

3. Enfin on a $\mathbb{P}_z(\text{Ext}) = q^z$.

DÉMONSTRATION. 1. Pour tout entier non nul k , la classe $\{1, 2, \dots, k\}$ est transiente. C'est-à-dire que la chaîne ne visite qu'un nombre fini de fois chacun des points. Donc presque sûrement, il existe un entier $n_0(\omega)$ tel que pour tout $n \geq n_0(\omega)$, $Z_n \geq k+1$. Ce qui montre exactement que

$$Z_n \xrightarrow[n \rightarrow +\infty]{} +\infty$$

2. On se rappelle de la proposition (3.6) et la croissance ou décroissance de $f_n(s)$ en découle rapidement (selon si la courbe est au-dessus ou en-dessous de la diagonale du carré)
3. On a par définition de l'extinction,

$$\{\text{Ext}\} = \bigcup_{n \geq 0} \uparrow \{Z_n = 0\}$$

et ainsi

$$\mathbb{P}(\text{Ext}) = \lim_{n \rightarrow \infty} \uparrow \mathbb{P}(Z_n = 0)$$

Et comme f_n est la fonction génératrice de Z_n , $\mathbb{P}_z(Z_n = 0) = f_n(0)^z$. On définit maintenant par récurrence la suite $(q_n)_{n \geq 0}$ par $q_0 = 0$ et $q_{n+1} = f(q_n)$.

On a donc $\mathbb{P}(\text{Ext}) = \lim_{n \rightarrow \infty} q_n$.

La continuité de f nous dit que cette limite est un point fixe de f donc appartient à $\{q, 1\}$. Or $0 = q_0 < q$ et par la croissance de f (proposition (3.6)) et en itérant, on a $q_n < q \leq 1$. \square

Définition 3.9. *Un modèle de BGW est dit*

- *surcritique quand $m > 1$ ($q < 1$)*
- *critique quand $m = 1$ ($q = 1$)*
- *sous-critique quand $m < 1$ ($q = 1$)*

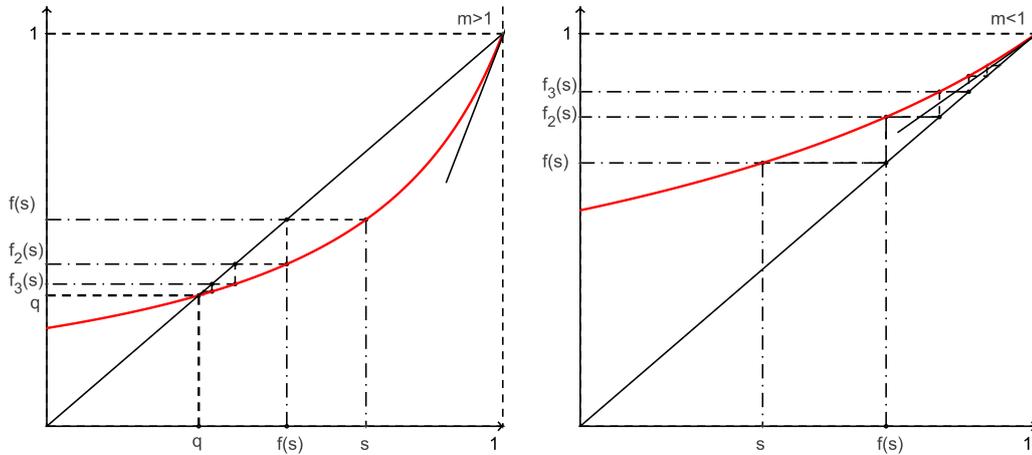


FIGURE 2 – Graphe de la fonction génératrice f (ici d'une loi géométrique) dans les cas surcritique ($m > 1$) à gauche et sous-critique ($m < 1$) à droite.

3.3 Quelques remarques sur la convergence

On a montré que $\mathbb{P}_z(\text{Ext}) = q^z$. Donc dans les cas sous-critiques et critiques, il y a extinction presque sûre. Dans le cas sur-critique, Z_n converge (presque sûrement) vers une v.a. limite Z_∞ qui ne prend que les valeurs 0 (c'est l'extinction) et $+\infty$ (c'est l'explosion) et on a

$$\mathbb{P}(Z_\infty = 0) = q^z \quad \mathbb{P}(Z_\infty = +\infty) = 1 - q^z$$

Analysons comment se produit l'extinction (ou la non extinction). Appelons τ le temps d'extinction.

Cas sous-critique

On a (comme τ est à valeurs strictement positives)

$$\mathbb{E}_z[\tau] = \sum_{n \geq 0} \mathbb{P}_z(\tau > n) = \sum_{n \geq 0} (1 - f_n(0))^z \quad (3.1)$$

car $\mathbb{P}_z(\tau \leq n) = \mathbb{P}_z(Z_n = 0)$. De plus $1 - f(s) \leq m(1 - s)$ car on a vu que f' est croissante sur $[0, 1]$, donc

$$\frac{f(1) - f(s)}{1 - s} = \frac{1 - f(s)}{1 - s} \leq f'(1) = m$$

On obtient par récurrence immédiate sur n , $1 - f_n(s) \leq m^n(1 - s)$. Ainsi

$$\mathbb{E}_1[\tau] \leq \sum_{n \leq 0} m^n < +\infty$$

Pour conclure on utilise la propriété des processus de branchements, en "décomposant" notre chaîne qui commence à $Z_0 = z$ en somme de chaînes qui débutent à 1. Cela permet de voir que la probabilité d'extinction de la population est égale au temps maximum d'extinction des sous-populations issues de chaque individu. (Ceci ramène exactement l'étude sous \mathbb{P}_z à l'étude sous \mathbb{P}_1). Ce qui prouve que

$$\mathbb{E}_z[\tau] < +\infty$$

L'espérance de l'extinction est donc finie. On peut également montrer que l'extinction a lieu très rapidement (décroissance à vitesse géométrique).

En effet

$$\mathbb{P}_z(Z_n \neq 0) = \mathbb{P}_z(Z_n \geq 1) \leq \mathbb{E}_z[Z_n] = m^n z$$

Cas critique

On reprend l'équation (3.1), cependant on ne peut pas conclure de la même manière (car $m = 1$), il faut faire une étude plus fine. On simplifie alors un peu le problème en posant $m_2 = \sum_{k \geq 0} k^2 p_k < +\infty$ (i.e. ξ admet un moment d'ordre 2). Dans ces nouvelles conditions on peut prouver une équivalence quand n tend vers l'infini de $1 - f_n(s)$ pour tout $s \in [0, 1[$.

Posons $a = (m_2 - 1)/2$ qui est strictement positif sous l'hypothèses $m = 1$. De plus

$$f''(1) = \sum_{k \geq 2} k(k-1)p_k = m_2 - m = 2a$$

Ainsi en faisant un développement limité à l'ordre 2 au voisinage de $s = 1$, on a pour tout $s \in [0, 1[$,

$$\begin{aligned} f(1) - f(s) &= (1-s)f'(1) - \frac{(1-s)^2}{2}f''(1) - (1-s)^2\varepsilon(s) \\ &= (1-s) - a(1-s)^2 - (1-s)^2\varepsilon(s) \\ &= (1-s)(1 - (1-s)(a + \varepsilon(s))) \end{aligned}$$

où $\lim_{s \rightarrow 1} \varepsilon(s) = 0$. En poursuivant on a

$$\begin{aligned} \frac{1-s}{1-f(s)} &= \frac{1}{1 - (1-s)(a + \varepsilon(s))} \\ &= 1 + (1-s)(a + \varepsilon(s)) + \lambda(s), \quad \text{où } \lim_{s \rightarrow 1} \lambda(s) = 0 \end{aligned}$$

Et ainsi pour tout entier n

$$\frac{1}{1 - f_{n+1}(s)} = \frac{1}{1 - f_n(s)} + a + \tilde{\varepsilon}(f_n(s)), \quad \text{où } \lim_{s \rightarrow 1} \tilde{\varepsilon}(s) = 0$$

On obtient donc facilement, en itérant, pour tout entier n

$$\frac{1}{1 - f_n(s)} = \frac{1}{1 - s} + na + \delta_n(s), \quad \text{où } \delta_n(s) = \sum_{k=0}^{n-1} \tilde{\varepsilon}(f_k(s))$$

Pour tout $0 \leq k \leq n$, pour $s \in [0, 1[$ fixé, $\tilde{\varepsilon}(f_k(s))$ tend vers 0 car $f_k(s)$ tend vers 1, ($m = 1$). Et donc $\delta_n(s)/n$ tend vers 0 (comme limite de Césaro de la suite $(\tilde{\varepsilon}(f_n(s)))_n$). Ainsi quand n tend vers l'infini, on a

$$1 - f_n(s) = \frac{1}{na} \frac{1}{1 + \frac{1}{na(1-s)} + \frac{\delta_n(s)}{na}} \sim \frac{1}{na} = \frac{2}{n(m_2 - 1)}$$

Revenons à notre problème.

Comme $\mathbb{E}_1[\tau] = \sum_{n \geq 0} (1 - f_n(0))$, on a $\mathbb{E}_1[\tau] = +\infty$ par l'équivalence que l'on vient de montrer. Cela montre que pour tout z ,

$$\mathbb{E}_z[\tau] = +\infty$$

Dans ce cas-ci la probabilité d'extinction tend plus lentement vers 0 que dans le cas $m < 1$. En effet, (toujours sous l'hypothèse $m_2 < +\infty$),

$$\mathbb{P}(Z_n \neq 0) = 1 - f_n(0) \sim \frac{2}{n(m_2 - 1)}, \quad \text{quand } n \rightarrow \infty$$

Remarque 3.1. *Le résultat $\mathbb{E}_z[\tau] = +\infty$ est toujours vrai sans l'hypothèse $m_2 < +\infty$ mais est beaucoup plus dur à démontrer.*

Cas surcritique Le théorème qui suit montre que Z_n croit vers $+\infty$ à une vitesse géométrique (en m^n).

Théorème 3.10. *On suppose que $m > 1$ et $m_2 < +\infty$. Alors $W_n = \frac{Z_n}{m^n}$ converge presque sûrement et dans L^2 vers une variable W strictement positive sur l'ensemble de non extinction et telle que $\mathbb{E}_z[W] = z$.*

DÉMONSTRATION. Considérons la filtration canonique $(\mathcal{F}_n)_n$ associée à $(Z_n)_n$. On a (en utilisant Fubini(positif))

$$\mathbb{E}_z[Z_{n+1} | \mathcal{F}_n] = \mathbb{E}_z \left[\sum_{k \geq 1} \mathbb{1}_{\{k \leq Z_n\}} \xi_k | \mathcal{F}_n \right] = \sum_{k \geq 1} \mathbb{E}_z \left[\mathbb{1}_{\{k \leq Z_n\}} \xi_k | \mathcal{F}_n \right]$$

Or $\mathbb{1}_{\{k \leq Z_n\}}$ est par définition \mathcal{F}_n -mesurable et pour tout k , ξ_k est indépendante de \mathcal{F}_n . Ainsi,

$$\mathbb{E}_z[Z_{n+1} | \mathcal{F}_n] = \sum_{k \geq 1} \mathbb{1}_{\{k \leq Z_n\}} \mathbb{E}_z[\xi_k] = Z_n m$$

Et donc

$$\mathbb{E}_z[W_{n+1} | \mathcal{F}_n] = W_n$$

Ce qui implique que W_n est une martingale. On a aussi directement $\mathbb{E}_z[W] = z$ car $\mathbb{E}_z[Z_n] = m^n z$ par la propriété (3.7)

De plus on a

$$\mathbb{E}_z[Z_{n+1}^2 | \mathcal{F}_n] = \mathbb{E}_z \left[\sum_{k \geq 1}^{Z_n} \xi_k^2 | \mathcal{F}_n \right] + 2 \mathbb{E}_z \left[\sum_{i \geq 1}^{Z_n-1} \sum_{j \geq i}^{Z_n} \xi_i \xi_j | \mathcal{F}_n \right]$$

En faisant le même raisonnement que précédemment, on trouve

$$\mathbb{E}_z[Z_{n+1}^2 | \mathcal{F}_n] = m_2 Z_n + m^2 Z_n (Z_n - 1)$$

Et donc $\mathbb{E}_z[W_{n+1}^2 | \mathcal{F}_n] \leq W_n^2 + \frac{m_2}{m^{n+2}} W_n$. On prend l'espérance de ce dernier résultat et on trouve $\mathbb{E}_z[W_{n+1}^2] \leq \mathbb{E}_z[W_n^2] + \frac{z m_2}{m^{n+2}}$. Une récurrence immédiate nous fournit alors

$$\begin{aligned} \mathbb{E}_z[W_{n+1}^2] &\leq \mathbb{E}[W_0^2] + z m_2 \sum_{k=2}^{n+2} \frac{1}{m^k} \\ &\leq z^2 + z m_2 \sum_{k \geq 0} \frac{1}{m^k} \\ &\leq z^2 + \frac{z m_2}{m(m-1)} \quad \square \end{aligned}$$

où on a utilisé que $W_0 = Z_0$ et que $m > 1$.

On a donc uniformément majoré les moments d'ordre 2 de la martingale $(W_n)_n$. On sait alors qu'elle converge presque sûrement et dans L^2 vers W .

4 Deux relations entre modèles de branchement et modèles à taille fixe

On commence par remarquer que la propriété de branchement (propriété (3.3)) nous dit que si le modèle de BGW est conditionné à avoir une taille fixe de $2N$ alors c'est un cas particulier du modèle de Cannings. La proposition qui suit nous donne un résultat plus précis.

Proposition 4.1. *Un modèle de BGW avec une loi de reproduction de Poisson, conditionné à avoir une taille fixe à la même loi qu'un modèle de Wright-Fisher.*

DÉMONSTRATION. Donnons d'abord du sens au "conditionnement à avoir une taille fixe".

Notons $\xi_i^{(n)}$ le nombre d'enfants de l'individu i de la génération $n - 1$ dans le modèle de BGW (sans contionnement). Ces v.a. sont i.i.d., donc pour tout entier $K' \geq K \geq 1$, la loi de $(\xi_i^{(n)}; 1 \leq i \leq 2N, 1 \leq n \leq K)$ conditionnée par $\{\sum_{i=1}^{2N} \xi_i^{(n)}; 1 \leq i \leq 2N, 1 \leq n \leq K'\}$ est indépendante de $K' \geq K$. Ceci donne donc du sens au fait de conditionner le modèle de BGW par $\{\sum_{i=1}^{2N} \xi_i^{(n)}; 1 \leq i \leq 2N, 1 \leq n \leq 1\}$. Ce qui se traduit par conditionner le modèle à avoir une taille fixe de $2N$.

Notons maintenant $\eta_i^{(n)}$ le nombre d'enfants de l'individu i de la génération $n - 1$ du modèle de BGW conditionné à avoir une taille fixe de $2N$. Les $2N$ -uplet $(\eta_i^{(n)}; 1 \leq i \leq 2N)$ pour $n \geq 1$ sont donc i.i.d. et pour tout $2N$ -uplet d'entiers (k_1, \dots, k_{2N}) de somme $2N$, la loi commune est donnée par

$$\begin{aligned} \mathbb{P}(\eta_1 = k_1, \dots, \eta_{2N} = k_{2N}) &= \mathbb{P}\left(\xi_1 = k_1, \dots, \xi_{2N} = k_{2N} \mid \sum_{i=1}^{2N} \xi_i = 2N\right) \\ &= \frac{\prod_{i=1}^{2N} \mathbb{P}(\xi_i = k_i)}{\mathbb{P}_{2N}(Z_1 = 2N)} \end{aligned}$$

La dernière égalité étant obtenue par indépendance des ξ_i et par définition de Z_n .

Si on note f la fonction génératrice d'une loi de Poisson de paramètre m , on a pour tout $s \in [0, 1]$, $f(s) = \exp(-m(1 - s))$.

De plus on a vu que la fonction génératrice de Z sous \mathbb{P}_{2N} est $f(s)^{2N}$, on a alors

$$\mathbb{E}_{2N}(s^{Z_1}) = \exp(-m(1 - s))^{2N} = \exp(-2Nm(1 - s))$$

On reconnaît alors la fonction génératrice d'une loi de Poisson de paramètre $2Nm$. Ainsi on a

$$\begin{aligned} \mathbb{P}(\eta_1 = k_1, \dots, \eta_{2N} = k_{2N}) &= \frac{\prod_{i=1}^{2N} (e^{-m} m^{k_i} / k_i!)}{(2Nm)^{2N} e^{-2Nm} / (2N)!} \\ &= e^{-2Nm} \frac{m^{2N}}{\prod_{i=1}^{2N} k_i!} \cdot e^{2Nm} \frac{(2N)!}{(2N)^{2N} m^{2N}} \\ &= \frac{(2N)!}{k_1! \dots k_{2N}!} \left(\frac{1}{2N} \right)^{2N} \end{aligned}$$

On reconnaît la distribution d'une loi multinomiale de paramètres $(2N; 1/2N, \dots, 1/2N)$ et donc la même loi que le modèle de Wright-Fisher. \square

On vient de voir qu'à partir d'un modèle de branchement, on peut obtenir un modèle de Wright-Fisher. Inversement, le théorème qui suit montre que l'on peut obtenir un modèle de branchement à partir du modèle de Wright-Fisher.

On note $Y^{(N)}$ la chaîne de Markov associée à un modèle de Wright-Fisher qui commence avec une sous-population de taille $Y_0^{(N)} = k$ indépendante de la taille $2N$ de la population.

Théorème 4.2. *La loi de la chaîne de Markov $(Y_n^{(N)}; n \geq 0)$ converge quand N tend vers l'infini, vers la loi d'un processus de BGW de loi de reproduction de Poisson de paramètre 1.*

Avant de démontrer ce théorème, rappelons un résultat qui sera utile dans la démonstration.

Théorème 4.3. *Soit $(X_n)_{n \geq 0}$ une suite de v.a. de loi binomiale de paramètres (n, p_n) et soit X une v.a. de loi de Poisson de paramètre $\lambda > 0$. On suppose que $\lim_{n \rightarrow \infty} np_n = \lambda$. Alors X_n converge en loi vers X quand n tend vers l'infini.*

DÉMONSTRATION (DE LA PROPOSITION (4.2)). On se rappelle que $Y_1^{(N)}$ suit une loi binomiale de paramètres $(2N, k/2N)$ (voir partie sur Wright-Fisher). Par le théorème (4.3), elle converge en loi vers une variable de loi de Poisson de paramètre k . Ce qui donne la convergence en loi des lois fini-dimensionnelles qui suffit à démontrer la convergence de la loi du processus dans ce cas-ci. \square

Références

- [1] R. Durrett. *Probability Models for DNA Sequence Evolution*. Probability and Its Applications. Springer, 2008.
- [2] W.J. Ewens. *Mathematical Population Genetics : Theoretical introduction*. Number vol. 1 in Interdisciplinary Applied Mathematics. Springer-Verlag, 2004.
- [3] A. Lambert. Population dynamics and random genealogies. *Stochastic Models*, 24(01) :45–163, 2008.
- [4] J.F. Le Gall. *Intégration, Probabilités et Processus Aléatoires*. FIMFA, 2006.
- [5] S. Méléard. *Modèles aléatoires en Ecologie et Evolution*. Ecole Polytechnique, 2009.

A.2 Apprentissage hors-murs : Marches Aléatoires sur des Réseaux orientés

Université de Paris-Sud 11

Stage de 2^{ème} année de Magistère de Mathématiques

Marches aléatoires sur des réseaux orientés

Antoine MARCHINA

21 septembre 2014

*Stage effectué au sein du laboratoire de Mathématiques de l'Université de
Versailles Saint-Quentin, sous la direction d'Alexis Devulder.*

Introduction et Remerciements

Ce rapport fait suite à l'apprentissage hors-mur pour le cursus du magistère. J'ai effectué ce stage de 3 semaines au sein du département de mathématiques de l'Université de Versailles Saint-Quentin en Yvelines sous la direction d'Alexis Devulder. L'objectif était l'étude d'un domaine actuel de recherche en Probabilités par le biais d'un article récent. J'ai également effectué des simulations (sur Matlab) des résultats théoriques démontrés dans l'article.

Ce rapport est essentiellement basé sur l'article de Massimo Campanino et Dimitri Petritis [1].

Je tiens à remercier en tout premier lieu Alexis Devulder qui a accepté de diriger ce stage et pour tous ses conseils pendant ces 3 semaines. Je remercie également tous les membres du département de maths qui m'ont fait découvrir la vie du laboratoire au quotidien, et particulièrement les thésards qui m'ont accueilli avec eux dans leur bureaux.

Table des matières

1	Présentation du cadre	4
1.1	Définitions	4
1.2	Résultats	6
2	Techniques préliminaires	6
2.1	Embarquement	6
2.2	Définitions	8
3	Preuves des théorèmes 1.4 and 1.5	10
3.1	Marche aléatoire sur le réseau \mathbb{L}	11
3.2	Marche aléatoire sur le réseau \mathbb{H}	14
4	Preuve du théorème 1.6	16
4.1	Techniques d'estimation	17
4.2	Preuve de la transience de la marche aléatoire \mathbb{O}_ε	18
	Références	26

1 Présentation du cadre

1.1 Définitions

Définition 1.1. Un graphe orienté simple \mathbb{G} est un couple (\mathbb{V}, \mathbb{A}) tel que

- \mathbb{V} est un ensemble dénombrable, appelé ensemble des sommets de \mathbb{G} .
- $\mathbb{A} \subset \{(x, y) \in \mathbb{V}^2 \mid x \neq y\}$ est un ensemble de couples d'éléments de \mathbb{V} , appelé ensemble des arcs orientés de \mathbb{G} .

Remarque 1.1. L'hypothèse $x \neq y$ dans la définition de \mathbb{A} se traduit par l'absence de boucles dans le graphe \mathbb{G} .

On définit maintenant les fonctions Range et Source notées respectivement r et s ainsi :

$$r: \begin{cases} \mathbb{A} \longrightarrow \mathbb{V} \\ a = (u, v) \longmapsto r(a) = v \end{cases} \quad s: \begin{cases} \mathbb{A} \longrightarrow \mathbb{V} \\ a = (u, v) \longmapsto s(a) = u \end{cases}$$

On définit de plus pour chaque sommet $v \in \mathbb{V}$, son degré entrant $d_v^+ = \#\{a \in \mathbb{A} \mid r(a) = v\}$ et son degré sortant $d_v^- = \#\{a \in \mathbb{A} \mid s(a) = v\}$.

Tous les graphes que l'on considère sont transitifs finis. C'est à dire que pour tout couple de sommets distincts de \mathbb{V} , il existe $k \in \mathbb{N}$ et une suite finie (w_0, \dots, w_k) avec pour tout $i = 0, \dots, k$, $w_i \in \mathbb{V}$, tel que $w_0 = u$, $w_k = v$ et $(w_i, w_{i+1}) \in \mathbb{A}$.

Remarque 1.2. Cette dernière propriété implique en particulier que tous les graphes considérés n'ont pas de "puits". C'est à dire pour tout $v \in \mathbb{V}$, $d_v^- \geq 1$.

Définition 1.2. Soit (\mathbb{V}, \mathbb{A}) un graphe orienté. Une marche aléatoire simple sur (\mathbb{V}, \mathbb{A}) est une chaîne de Markov $(M_n)_{n \in \mathbb{N}}$ sur \mathbb{V} de matrice de transition P définie par :

$$\mathbb{P}(u, v) = \mathbb{P}(M_{n+1} = v \mid M_n = u) = \begin{cases} \frac{1}{d_u^-} & \text{si } (u, v) \in \mathbb{A} \\ 0 & \text{sinon} \end{cases}$$

Maintenant nous considérerons comme graphe uniquement des réseaux à 2 dimensions, i.e. $\mathbb{V} = \mathbb{Z}^2$. On écrira souvent $\mathbb{V} = \mathbb{V}_1 \times \mathbb{V}_2 = \mathbb{Z}^2$, avec \mathbb{V}_1 et \mathbb{V}_2 isomorphes à \mathbb{Z} .

Définition 1.3. Soit $\varepsilon = (\varepsilon_y)_{y \in \mathbb{V}_2}$ une suite de variables à valeurs dans $\{-1, 1\}$ (suite déterministe ou aléatoire). On appelle réseau orienté ε -horizontalement $\mathbb{G} = \mathbb{G}(\mathbb{V}, \varepsilon)$, le graphe orienté simple tel que l'ensemble des sommets soit $\mathbb{V} = \mathbb{Z}$ et que $(u, v) \in \mathbb{A}$ si et seulement si $u \neq v$ et satisfont l'une des 2 conditions suivantes :

- Soit $v_1 = u_1$ et $v_2 = u_2 \pm 1$
- Soit $v_1 = u_1 + \varepsilon_{u_2}$ et $v_2 = u_2$

Remarque 1.3. Ici, le réseau est régulier dans le sens où pour tout $v \in \mathbb{V}$, $d_v^- = d_v^+ = d = 3$.

Dans la suite, nous étudierons 3 réseaux particuliers :

1. Réseau alterné \mathbb{L}

ε est une suite déterministe définie pour tout $y \in \mathbb{V}_2$ par $\varepsilon_y = (-1)^y$.

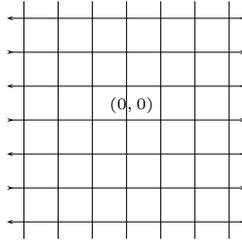


FIGURE 1 – Réseau alterné \mathbb{L}

2. Réseau bidirectionnel \mathbb{H}

ε est une suite déterministe définie pour tout $y \in \mathbb{V}_2$ par $\varepsilon_y = \begin{cases} 1 & \text{si } y \geq 0 \\ -1 & \text{si } y < 0 \end{cases}$

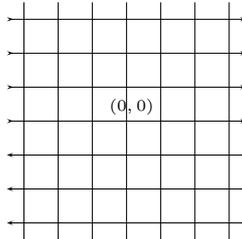


FIGURE 2 – Réseau bidirectionnel \mathbb{H}

3. Réseau aléatoirement orienté \mathbb{O}_ε
 ε est une suite de variables aléatoires de loi de Rademacher.

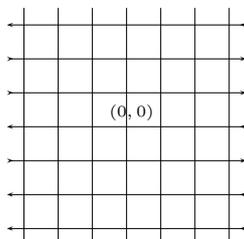


FIGURE 3 – Réseau aléatoirement orienté \mathbb{O}_ε

1.2 Résultats

Théorème 1.4. *La marche aléatoire simple sur le réseau alterné \mathbb{L} est récurrente.*

Théorème 1.5. *La marche aléatoire simple sur le réseau bidirectionnel \mathbb{H} est transiente.*

Théorème 1.6. *Pour presque toutes réalisations de l'environnement ε , la marche aléatoire sur le réseau aléatoirement orienté \mathbb{O}_ε est transiente.*

2 Techniques préliminaires

2.1 Embarquement

On suppose qu'il existe un espace probabilisé $(\Omega, \mathcal{A}, \mathbb{P})$ où sont définies toutes les v.a. que l'on va utiliser. En particulier, la chaîne de Markov $(M_n)_{n \in \mathbb{N}}$ y est définie et on note $(\tilde{\mathcal{F}}_n)_{n \in \mathbb{N}}$ sa filtration naturelle. On suppose de plus que la chaîne part de 0 (i.e. $M_0 = (0, 0)$). On note de plus \mathbf{e}_1 et \mathbf{e}_2 les vecteurs unités de \mathbb{Z}^2 . Enfin on désigne par $(\tilde{\psi}_n)_{n \in \mathbb{N}}$ la suite des projections verticales des accroissements de la chaîne de Markov. C'est à dire

$$\psi_{n+1} = \langle \mathbf{e}_2, M_{n+1} - M_n \rangle$$

où $\langle \cdot, \cdot \rangle$ désigne le produit scalaire euclidien sur \mathbb{R}^2 .

Les variables $\tilde{\psi}_n$ forment une suite de v.a. à valeurs dans $\{-1, 0, 1\}$ de loi

$$\mathbb{P}(\tilde{\psi}_i = -1) = \mathbb{P}(\tilde{\psi}_i = 1) = \mathbb{P}(\tilde{\psi}_i = 0) = \frac{1}{3}$$

Lemme 2.1. *Sur l'événement $\{M_n = u\}$, la chaîne de Markov (M_n) vérifie*

$$M_{n+1} = u + \varepsilon_{u_2} \mathbf{e}_1 \mathbf{1}_{\tilde{\psi}_{n+1}=0} + \tilde{\psi}_{n+1} \mathbf{e}_2 \mathbf{1}_{\tilde{\psi}_{n+1} \neq 0}$$

L'idée de l'embarquement est de décomposer la marche aléatoire (M_n) en une marche aléatoire simple à 1 dimension – la marche verticale (Y_n) –, une suite de temps d'attente (ξ_n) et une marche aléatoire à 1 dimension avec des sauts sans limites – la marche horizontale (X_n) .

On découpe alors en intervalle les moments où (M_n) réalise un mouvement vertical et ceux où elle réalise un mouvement horizontal. Pour ce faire, on introduit la suite de temps aléatoires suivante :

$$\begin{aligned} \tilde{\sigma}_1 &= 1 \\ \tilde{\tau}_1 &= \inf\{n \geq \tilde{\sigma}_1 \mid \tilde{\psi}_n \neq 0\} \geq \tilde{\sigma}_1 \\ \tilde{\sigma}_2 &= \inf\{n \geq \tilde{\tau}_1 \mid \tilde{\psi}_n \neq 0\} > \tilde{\tau}_1 \\ &\vdots \\ \tilde{\tau}_k &= \inf\{n \geq \tilde{\sigma}_k \mid \tilde{\psi}_n \neq 0\} > \tilde{\sigma}_k \\ \tilde{\sigma}_{k+1} &= \inf\{n \geq \tilde{\tau}_k \mid \tilde{\psi}_n \neq 0\} > \tilde{\tau}_k \\ &\vdots \end{aligned}$$

Et les intervalles suivants qui sont respectivement les moments où (M_n) réalise un mouvement vertical quand n est dans un I_k et un mouvement horizontal quand n est dans un J_k .

$$\begin{aligned} I_k &= [\tilde{\sigma}_k, \tilde{\tau}_k - 1] \cap \mathbb{N} \\ J_k &= [\tilde{\tau}_k, \tilde{\sigma}_{k+1} - 1] \cap \mathbb{N} \end{aligned}$$

On peut alors définir la suite $(\tilde{\xi}_j)_{j \in \mathbb{N}}$ par

$$\begin{aligned} \tilde{\xi}_{\sum_{i=1}^n |J_i| + \alpha + (n-1)} &= \tilde{\tau}_{n+1} - \tilde{\sigma}_n, \\ \tilde{\xi}_{\sum_{i=1}^n |J_i| + \alpha + (n-1) + k} &= 0, \quad \forall k \in \{1, \dots, |J_{n+1}|\} \end{aligned}$$

On définit également la suite $(\psi_j)_{j \geq 1}$ définie à partir de $(\tilde{\psi}_n)$ où on a enlevé tous les rangs k où $\tilde{\psi}_k = 0$ (c'est-à-dire que la suite (ψ_j) ne voit que les déplacements verticaux).

On a alors assez facilement les 3 lemmes suivants :

Lemme 2.2. *A une suite $(\tilde{\psi}_n)_{n \in \mathbb{N}}$ donnée, les suites $(\tilde{\xi}_j)_{j \geq 1}$ et $(\psi_j)_{j \geq 1}$ sont uniquement déterminées et inversement.*

Lemme 2.3. *La suite $(\psi_j)_{j \geq 1}$ est une suite de va iid de loi de Rademacher.*

Lemme 2.4. *La suite $(\xi_j)_{j \geq 1}$ est une suite de vaïid de loi géométrique sur \mathbb{N} de paramètres $p = 1/3$.*

Les constructions que l'on a effectuées permettent de donner le contexte de ce qu'on fera par la suite. On considérera alors directement que les suites (ψ_n) et (ξ_n) (sans plus parler de $(\tilde{\psi}_n)$).

2.2 Définitions

Définition 2.5. *Soit $(\psi_n)_{n \in \mathbb{N}}$ une suite de vaïid de loi de Rademacher et on pose $Y_0 = 0$ et*

$$Y_n = \sum_{k=1}^n \psi_k \quad \text{pour tout } n \geq 1$$

$(Y_n)_{n \in \mathbb{N}}$ est la marche aléatoire simple (à 1 dimension) à valeurs dans \mathbb{V}_2 . Ce processus est appelé le squelette vertical.

On note de plus, pour tout $n \in \mathbb{N}$ et $y \in \mathbb{V}_2$,

$$\eta_n(y) = \sum_{k=0}^n \mathbb{1}_{Y_k=y}$$

son temps d'occupation du niveau y .

Définition 2.6. *Supposons que le squelette vertical et l'environnement des orientations soient donnés. Soit $(\xi_n^{(y)})_{n \geq 1, y \in \mathbb{V}_2}$ une suite doublement infinie de vaïid de loi géométrique sur \mathbb{N} de paramètres p et $q = 1 - p$. Soit $(\eta_n(y))$ définie ci-dessus. On appelle marche aléatoire horizontale embarquée, le processus $(X_n)_{n \in \mathbb{N}}$ avec pour tout $n \in \mathbb{N}$*

$$X_n = \sum_{y \in \mathbb{V}_2} \varepsilon_y \sum_{i=1}^{\eta_{n-1}(y)} \xi_i^{(y)}$$

Remarque 2.1. *On prend comme convention que la somme en i est nulle dès que $\eta_{n-1}(y) = 0$.*

La v.a. X_n correspond au déplacement horizontal après $n-1$ déplacements verticaux.

On remarque que la marche (M_n) ne coïncide pas avec (X_n, Y_n) . Cependant ces 2 objets sont quand même proches comme le dis le lemme suivant qui découle directement des définitions des objets considérés :

Lemme 2.7. *Soit*

$$T_n = n + \sum_{y \in \mathbb{V}_2} \sum_{i=1}^{\eta_{n-1}(y)} \xi_i^{(y)}$$

l'instant juste après que la marche aléatoire (M_k) ait réalisé son n -ième mouvement vertical. On a alors

$$M_{T_n} = (X_n, Y_n)$$

On définit maintenant

$$\begin{aligned} \sigma_0 &= 0 \\ \sigma_n &= \inf\{k \geq \sigma_{n-1} \mid Y_k = 0\} > \sigma_{n-1}, \quad \forall n \geq 1 \end{aligned}$$

le n -ième retour à l'origine pour la marche squelette. Et donc, $M_{T_{\sigma_n}} = (X_{\sigma_n}, 0)$.

Remarquons que $M_k = (0, 0)$ ssi $X_k = 0$ et $Y_k = 0$. Et comme (Y_k) est une marche aléatoire simple, l'événement $\{Y_k = 0\}$ n'est réalisé qu'aux instants σ_n pour $n \geq 1$.

On va donner maintenant un critère de transience pour la marche (M_k) .

Lemme 2.8. *Soit $\mathcal{F} = \sigma(\psi_i, i \in \mathbb{N})$ et $\mathcal{G} = \sigma(\varepsilon_y, y \in \mathbb{V}_2)$. Et soit Z une va de même loi que ξ_1 .*

$$\sum_{l=0}^{\infty} \mathbb{P}(M_l = (0, 0) \mid \mathcal{F} \vee \mathcal{G}) = \sum_{n=0}^{\infty} \mathbb{P}(0 \in I(X_{\sigma_n}, \varepsilon_0 Z) \mid \mathcal{F} \vee \mathcal{G})$$

où, pour tout $x \in \mathbb{Z}$ et $z \in \mathbb{N}$ et $\varepsilon = \pm 1$, $I(x, \varepsilon z) = \{x, \dots, x + z\}$ si $\varepsilon = +1$ et $\{x - z, \dots, x\}$ si $\varepsilon = -1$.

DÉMONSTRATION. Pour que le processus (M_l) retourne à l'origine, il faut que les composantes verticales et horizontales soient 0 en même temps. On se rappelle $Y_{\sigma_n} = 0$ et alors seulement $M_{\sigma_n} = (X_{\sigma_n}, 0)$. Pour $k = T_{\sigma_n}, \dots, T_{\sigma_{n+1}-1}$, le processus M_k peut être à l'origine avant qu'il ne réalise un déplacement vertical. C'est-à-dire si l'un des $X_{\sigma_n} = 0$ ou si les points X_{σ_n} et $X_{\sigma_{n+1}}$ sont de part et d'autre du point 0. Or

$$\begin{aligned} X_{\sigma_{n+1}} - X_{\sigma_n} &= \sum_{y \in \mathbb{V}_2} \varepsilon_y \left(\sum_{i=1}^{\eta_{\sigma_n}(y)} \xi_i^{(y)} - \sum_{i=1}^{\eta_{\sigma_{n+1}-1}(y)} \xi_i^{(y)} \right) \\ &= \varepsilon_0 \xi_{\eta_{\sigma_n}(0)}^{(0)} \\ &\stackrel{\text{loi}}{=} \varepsilon_0 Z \end{aligned}$$

Et donc, le processus (M_l) peut retourner en 0 pour $l \in \{T_{\sigma_n}, \dots, T_{\sigma_{n+1}-1}\}$ ssi le point 0 appartient à l'ensemble $I(X_{\sigma_n}, \varepsilon_0 Z)$. \square

Lemme 2.9. *En gardant les mêmes notations que précédemment*

1. Si $\sum_{n=0}^{\infty} \mathbb{P}_0(X_{\sigma_n} = 0 \mid \mathcal{F} \vee \mathcal{G}) = \infty$ alors $\sum_{l=0}^{\infty} \mathbb{P}(M_l = (0, 0) \mid \mathcal{F} \vee \mathcal{G}) = \infty$
2. Si $(X_{\sigma_n})_{n \in \mathbb{N}}$ est transiente alors $(M_n)_{n \in \mathbb{N}}$ est aussi transiente.

DÉMONSTRATION.

$$\begin{aligned} \mathbb{P}(0 \in I(X_{\sigma_n}, \varepsilon_0 Z) \mid \mathcal{F} \vee \mathcal{G}) &= \mathbb{P}(X_{\sigma_n} = 0 \mid \mathcal{F} \vee \mathcal{G}) \\ &\quad + \mathbb{1}_{\{\varepsilon_0 = -1\}} \mathbb{P}(\cup_{x \in \mathbb{N}^*} \{X_{\sigma_n} = x; Z \geq x \mid \mathcal{F} \vee \mathcal{G}\}) \\ &\quad + \mathbb{1}_{\{\varepsilon_0 = 1\}} \mathbb{P}(\cup_{x \in \mathbb{N}^*} \{X_{\sigma_n} = -x; Z \geq x \mid \mathcal{F} \vee \mathcal{G}\}). \end{aligned}$$

Pour le cas 1, le lemme précédent donne directement le résultat. Pour le cas 2, comme le processus X_{σ_n} est transient, il existe une constante $C > 0$ telle que pour tout $x \in \mathbb{Z}$, on ait

$$\sum_{n \in \mathbb{N}} \mathbb{P}_0(X_{\sigma_n} = x \mid \mathcal{F} \vee \mathcal{G}) \leq C < \infty$$

et donc,

$$\sum_{n \in \mathbb{N}} \mathbb{P}_0(\cup_{x \in \mathbb{N}^*} \{X_{\sigma_n} = x; Z \geq x\} \mid \mathcal{F} \vee \mathcal{G}) = \sum_{n \in \mathbb{N}} \sum_{x \geq 1} q^x \mathbb{P}_0(X_{\sigma_n} = x \mid \mathcal{F} \vee \mathcal{G}) \leq \frac{q}{1-q} C,$$

en utilisant le théorème de Fubini positif pour la dernière inégalité. On a ainsi prouvé la transience de $(M_n)_{n \in \mathbb{N}}$. \square

3 Preuves des théorèmes 1.4 and 1.5

Soit ξ une va de même loi que $\xi_i^{(y)}$. Notons sa fonction caractéristique

$$\chi(\theta) = \mathbb{E} \exp(i\theta\xi) = \frac{p}{1 - q \exp(i\theta)} = r(\theta) \exp(i\alpha(\theta)), \quad \theta \in [-\pi, \pi]$$

où

$$r(\theta) = |\chi(\theta)| = \frac{p}{\sqrt{p^2 + 2q(1 - \cos \theta)}} = r(-\theta)$$

et

$$\alpha(\theta) = \arctan \frac{q \sin \theta}{1 - q \cos \theta} = -\alpha(-\theta)$$

On a

$$\begin{aligned}
\mathbb{E}[\exp(i\theta X_n)] &= \mathbb{E}[\mathbb{E}[\exp(i\theta X_n) \mid \mathcal{F} \vee \mathcal{G}]] \\
&= \mathbb{E}\left[\mathbb{E}\left[\exp\left(i\theta \sum_{y \in \mathbb{V}_2} \epsilon_y \sum_{i=1}^{\eta_{n-1}(y)} \xi_i^{(y)}\right) \mid \mathcal{F} \vee \mathcal{G}\right]\right] \\
&= \mathbb{E}\left[\prod_{y \in \mathbb{V}_2} \chi(\theta \epsilon_y)^{\eta_{n-1}(y)}\right]
\end{aligned}$$

3.1 Marche aléatoire sur le réseau \mathbb{L}

Lemme 3.1. *Pour tout $n \in \mathbb{N}^*$, le temps d'occupation de la marche aléatoire squelette vérifie*

$$\sum_{y \in \mathbb{V}_2} (-1)^y \eta_{\sigma_n-1}(y) = 0$$

DÉMONSTRATION. Montrons que

$$\sum_{y \in \mathbb{V}_2} (-1)^y \eta_{\sigma_n-1}(y) = \sum_{k=0}^{\sigma_n-1} (-1)^{Y_k}$$

Rappelons que

$$\eta_n(y) = \#\{0 \leq k \leq n \mid Y_k = y\}$$

Pour toute fonction mesurable f , on a alors,

$$\begin{aligned}
\sum_{y \in \mathbb{V}_2} f(y) \eta_n(y) &= \sum_{y \in \mathbb{V}_2} f(y) \sum_{k=0}^n \mathbf{1}_{Y_k=y} \\
&= \sum_{k=0}^n \sum_{y \in \mathbb{V}_2} f(y) \mathbf{1}_{Y_k=y} \\
&= \sum_{k=0}^n \sum_{y \in \mathbb{V}_2} f(Y_k) \mathbf{1}_{Y_k=y} \\
&= \sum_{k=0}^n f(Y_k)
\end{aligned}$$

Maintenant, comme σ_n est pair, $\sum_{k=0}^{\sigma_n-1} (-1)^{Y_k} = 0$ car pour tout $k \in \mathbb{N}$, $Y_{k+1} = Y_k \pm 1$. \square

DÉMONSTRATION DU THÉORÈME 1.4.

Reprenons le lemme 2.8, et prenons l'espérance des 2 cotés de l'égalité. On a ainsi,

$$\mathbb{E} \left[\sum_{l=0}^n \mathbb{P}(M_l = (0, 0) \mid \mathcal{F} \vee \mathcal{G}) \right] = \mathbb{E} \left[\sum_{n=0}^{+\infty} \mathbb{P}(0 \in I(X_{\sigma_n}, \varepsilon_0 Z) \mid \mathcal{F} \vee \mathcal{G}) \right]$$

On intervertit alors somme et espérance par Fubini positif et on a donc facilement

$$\sum_{n \in \mathbb{N}} \mathbb{P}(X_{\sigma_n} = 0) \leq \sum_{n \in \mathbb{N}} \mathbb{P}(M_n = (0, 0))$$

Notons f la fonction génératrice du temps du premier retour en 0 pour la marche aléatoire simple, i.e. pour $|s| \leq 1$, $f(s) = \mathbb{E}[s^{\sigma_1}]$.

Admettons le lemme suivant. L'idée de la démonstration est de transformer le temps du premier retour en 0 en temps d'atteinte de -1 ou 1 pour la marche aléatoire obtenue à partir de (Y_k) mais en supprimant la première étape Y_1 et de s'intéresser à $\mathbb{P}(\sigma_1 = 2k)$ et à la fonction génératrice du temps d'atteinte de 1 ou -1.

Lemme 3.2. *On a pour $|s| \geq 1$,*

$$f(s) = 1 - \sqrt{(1 - s^2)}$$

En utilisant de plus les propriétés du module et de l'argument de la fonction caractéristique, et le lemme 3.1.

$$\begin{aligned} \mathbb{E} [\exp(i\theta X_{\sigma_n})] &= \mathbb{E} \left[\prod_{y \in \mathbb{V}_2} \chi(\theta \varepsilon_y)^{\eta_{\sigma_n-1}(y)} \right] \\ &= \mathbb{E} \left[r(\theta)^{\sum_{y \in \mathbb{V}_2} \eta_{\sigma_n-1}(y)} \exp(i\alpha(\theta) \sum_{y \in \mathbb{V}_2} \varepsilon_y \eta_{\sigma_n-1}(y)) \right] \\ &= \mathbb{E} \left[r(\theta)^{\sum_{y \in \mathbb{V}_2} \eta_{\sigma_n-1}(y)} \right] \end{aligned}$$

Et on a facilement,

$$\sum_{y \in \mathbb{V}_2} \eta_{n-1}(y) = \sum_{y \in \mathbb{V}_2} \sum_{l=0}^{n-1} \mathbf{1}_{\{Y_l=y\}} = n$$

Donc

$$\mathbb{E} [\exp(i\theta X_{\sigma_n})] = \mathbb{E} [r(\theta)^{\sigma_n}]$$

Or pour tout $k \in \mathbb{N}^*$, $(\sigma_{k+1} - \sigma_k) \stackrel{loi}{=} \sigma_1$ (par la propriété de Markov) et les accroissements sont indépendants. Et donc, en notant $\sigma_n = (\sigma_n - \sigma_{n-1}) + (\sigma_{n-1} - \sigma_{n-2}) + \dots + (\sigma_2 - \sigma_1) + \sigma_1$,

$$\mathbb{E}[r(\theta)^{\sigma_n}] = \prod_{k=1}^n \mathbb{E}[r(\theta)^{(\sigma_k - \sigma_{k-1})}] = (\mathbb{E}[r(\theta)^{\sigma_1}])^n$$

Finalement, en utilisant le lemme 3.2,

$$\mathbb{E}[\exp(i\theta X_{\sigma_n})] = \left(1 - \sqrt{1 - r(\theta)^2}\right)^n$$

Par ailleurs,

$$\begin{aligned} \frac{1}{2\pi} \int_{-\pi}^{\pi} \mathbb{E}[\exp(i\theta X_{\sigma_n})] d\theta &= \frac{1}{2\pi} \int_{-\pi}^{\pi} \sum_{k \geq 0} \exp(i\theta 2k) \mathbb{P}(X_{\sigma_n} = 2k) d\theta \\ &= \sum_{k \geq 0} \frac{1}{2\pi} \int_{-\pi}^{\pi} \exp(i\theta 2k) d\theta \mathbb{P}(X_{\sigma_n} = 2k) \\ &= \mathbb{P}(X_{\sigma_n} = 0) \end{aligned}$$

où l'on a appliqué le théorème de Fubini dans la deuxième égalité.

Ainsi, (en appliquant le théorème de Fubini positif)

$$\begin{aligned} \sum_{n \geq 0} \mathbb{P}(X_{\sigma_n} = 0) &= \frac{1}{2\pi} \int_{-\pi}^{\pi} \left(1 - \sqrt{1 - r(\theta)^2}\right)^n d\theta \\ &= \frac{1}{\pi} \int_0^{\pi} \frac{1}{\sqrt{1 - r(\theta)^2}} d\theta \end{aligned}$$

Pour $\theta \rightarrow 0$, $\frac{1}{\sqrt{1 - r(\theta)^2}} = O\left(\frac{1}{|\theta|}\right)$ non intégrable en 0. Et donc, $\sum_{n \geq 0} \mathbb{P}(X_{\sigma_n} = 0) = \infty$, ce qui prouve la récurrence de la chaîne (M_n) . \square

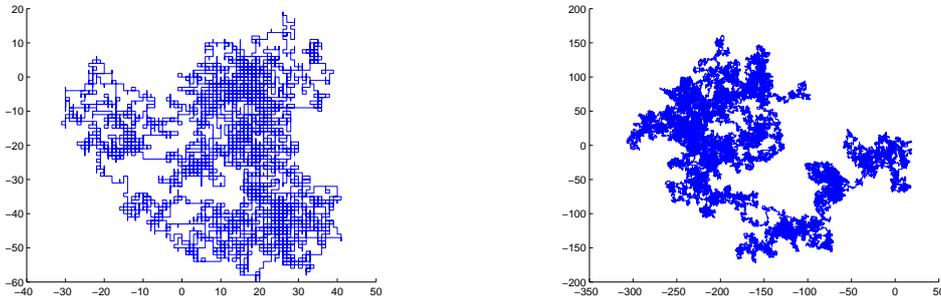


FIGURE 4 – Exemple de marche aléatoire \mathbb{L} sur 10 000 et 100 000 pas.

3.2 Marche aléatoire sur le réseau \mathbb{H}

Lemme 3.3. Soit $\varepsilon_y = 1$ si $y \geq 0$ et $\varepsilon_y = -1$ si $y < 0$. Notons $(\rho_k)_{k \in \mathbb{N}}$ une suite de va iid de loi de Rademacher et $(\tau_k)_{k \in \mathbb{N}}$ une suite de va iid, indépendante de la suite (ρ_k) , et telle que $\tau_1 = \sigma_1$. On a

$$\sum_{y \in \mathbb{V}_2} \varepsilon_y \eta_{\sigma_n-1}(y) \stackrel{\text{loi}}{=} \sum_{k=1}^n \rho_k (\tau_k - 1) + n$$

DÉMONSTRATION.

$$\begin{aligned} \sum_{y \in \mathbb{V}_2} \varepsilon_y \eta_{\sigma_n-1}(y) &= \sum_{y \in \mathbb{V}_2} \varepsilon_y \sum_{j=1}^n \sum_{k=\sigma_{j-1}}^{\sigma_j-1} \mathbb{1}_{\{Y_k=y\}} \\ &= \sum_{y \neq 0} \varepsilon_y \sum_{j=1}^n \sum_{k=\sigma_{j-1}+1}^{\sigma_j-1} \mathbb{1}_{\{Y_k=y\}} + \varepsilon_0 n \end{aligned}$$

(Y_k vaut 0 si et seulement si $k = \sigma_j$ pour un certain $j \in \mathbb{N}$, et comme $\eta_{\sigma_n-1}(y) = \sum_{k=0}^{\sigma_n-1} \mathbb{1}_{Y_k=y}$, cela justifie le terme " $\varepsilon_0 n$ ").

Par définition des σ_n , pour tout $j \in \{1, \dots, n\}$, le processus Y_k a le même signe pour tout $k \in \{\sigma_{j-1}+1, \dots, \sigma_j-1\}$. Ainsi, $\sum_{y \neq 0} \sum_{k=\sigma_{j-1}+1}^{\sigma_j-1} \mathbb{1}_{Y_k=y} = \sigma_j - \sigma_{j-1} \stackrel{\text{loi}}{=} \tau_j - 1$ en inversant les sommes et où on a utilisé comme précédemment que $\sigma_{k+1} - \sigma_k \stackrel{\text{loi}}{=} \sigma_1$.

Donc

$$\sum_{y \in \mathbb{V}_2} \varepsilon_y \eta_{\sigma_n-1}(y) = \tilde{\varepsilon}_j \sum_{k=\sigma_{j-1}+1}^{\sigma_j-1} \sum_{y \neq 0} \mathbb{1}_{Y_k=y}$$

où $\tilde{\varepsilon}_j$ vaut 1 si $Y_{\sigma_{j-1}+1}$ est positif et -1 si c'est négatif. Et par définition de Y_k , $\tilde{\varepsilon}_j = \rho_j$ ce qui conclut la preuve. \square

Proposition 3.4. Pour la marche aléatoire embarquée, on a

$$\mathbb{E}[\exp(i\theta X_{\sigma_n})] = g(\theta)^n$$

$$\text{où } g(\theta) = \frac{1}{2} \chi(\theta) \left[2 - \left(1 - \sqrt{1 - \chi(\theta)^2} \right) + \left(1 - \sqrt{1 - \bar{\chi}(\theta)^2} \right) \right]$$

DÉMONSTRATION. Notons $\mathcal{D} = \sigma(\rho_k, k \in \mathbb{N})$ la tribu engendrée par les

variables de Rademacher. Ensuite en utilisant le lemme 3.3 on a

$$\begin{aligned}
\mathbb{E}[\exp(i\theta X_{\sigma_n})] &= \mathbb{E}[\mathbb{E}[\exp(i\theta X_{\sigma_n})|\mathcal{D}]] \\
&= \mathbb{E}\left[\mathbb{E}\left[r(\theta)^{\sum_{j=1}^n(\tau_j-1)+n} \exp\left(i\alpha(\theta)\left(\sum_{j=1}^n \rho_j(\tau_j-1) + n\right)\right) \middle| \mathcal{D}\right]\right] \\
&= \chi(\theta)^n \mathbb{E}\left[\mathbb{E}\left[r(\theta)^{\sum_{j=1}^n \tau_j-1} \exp(i\alpha(\theta) \sum_{j=1}^n \rho_j(\tau_j-1)) \middle| \mathcal{D}\right]\right] \\
&= \chi(\theta)^n \mathbb{E}\left[\mathbb{E}\left[\prod_{j=1}^n (\chi(\theta \rho_j)^{\tau_j-1}) \middle| \mathcal{D}\right]\right] \\
&= \chi(\theta)^n \mathbb{E}\left[\mathbb{E}\left[\prod_{j=0}^{n-1} (\chi(\theta \rho_{j+1})^{\tau_j}) \middle| \mathcal{D}\right]\right] \\
&= \chi(\theta)^n \mathbb{E}\left[\mathbb{E}\left[\prod_{j=0}^{n-1} (\chi(\theta \rho_{j+1})^{\sigma_j}) \middle| \mathcal{D}\right]\right] \\
&= \chi(\theta)^n \prod_{j=0}^{n-1} \mathbb{E}\left[\left(1 - \sqrt{1 - \chi(\theta \rho_{j+1})^2}\right)\right] \\
&= g(\theta)^n. \quad \square
\end{aligned}$$

Proposition 3.5. *La marche aléatoire (M_n) sur le réseau \mathbb{H} vérifie*

$$\sum_{n=0}^{+\infty} \mathbb{P}(M_n = (0, 0)) < +\infty$$

DÉMONSTRATION. Comme $\varepsilon_0 = 1$ et en se rappelant du lemme 2.8, on a

$$\begin{aligned}
\sum_{n=0}^{\infty} \mathbb{P}(M_n = (0, 0)) &= \sum_{n=0}^{\infty} \mathbb{P}(0 \in I(X_{\sigma_n}, Z)) \\
&= \sum_{n=0}^{\infty} \sum_{x \geq 0} \mathbb{P}(X_{\sigma_n} = -x) \mathbb{P}(Z \geq x).
\end{aligned}$$

car $I(X_{\sigma_n}, Z) = \{X_{\sigma_n}, \dots, X_{\sigma_n} + Z\}$.

On a de plus en refaisant la même astuce que dans la démonstration sur le réseau \mathbb{L} ,

$$\mathbb{P}(X_{\sigma_n} = -x) = \frac{1}{2\pi} \int_{-\pi}^{\pi} \exp(i\theta x) g(\theta)^n d\theta$$

Et donc,

$$\begin{aligned} \sum_{x \geq 0} \mathbb{P}(X_{\sigma_n} = -x) \mathbb{P}(Z \geq x) &= \frac{1}{2\pi} \int_{-\pi}^{\pi} \left(\sum_{x \geq 0} \exp(i\theta x) q^x \right) g(\theta)^n d\theta \\ &= \frac{1}{2\pi} \int_{-\pi}^{\pi} \frac{1}{1 - q \exp(i\theta)} g(\theta)^n d\theta \end{aligned}$$

Ainsi,

$$\begin{aligned} \sum_{n=0}^{\infty} \mathbb{P}(M_n = (0, 0)) &= \sum_{n=0}^{\infty} \frac{1}{2\pi} \int_{-\pi}^{\pi} \frac{1}{1 - q \exp(i\theta)} g(\theta)^n d\theta \\ &= \frac{1}{2\pi} \int_0^{\pi} \frac{\operatorname{Re}(\chi(\theta))}{p(1 - g(\theta))} d\theta \end{aligned}$$

Pour pouvoir conclure sur la transience de la marche, il faut maintenant étudier l'intégrabilité de $(1 - g(\theta))^{-1}$. Or on a par calcul

$$\lim_{\theta \rightarrow 0^+} \frac{1 - g(\theta)}{\sqrt{\theta}} = \sqrt{2}$$

Le facteur $(1 - g(\theta))^{-1}$ est donc intégrable, ce qui conclut la preuve. \square

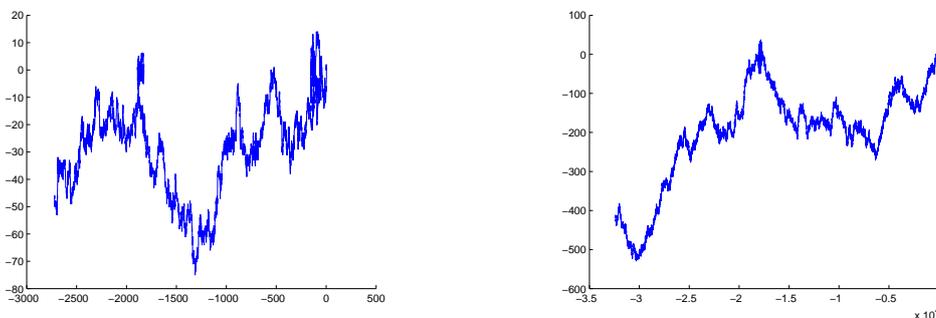


FIGURE 5 – Exemple de marche aléatoire \mathbb{H} sur 10 000 et 100 000 pas.

4 Preuve du théorème 1.6

Pour ce réseau, on ne peut plus faire les mêmes méthodes que précédemment car l'accroissement $X_{\sigma_{n+1}} - X_{\sigma_n}$ n'est plus indépendant de l'accroissement $X_{\sigma_n} - X_{\sigma_{n-1}}$ (ils peuvent avoir le même ε_y pour quelque y).

4.1 Techniques d'estimation

Introduisons d'abord les notations suivantes :

$$\begin{aligned}\mathcal{H} &= \sigma(\xi_i^{(y)}, i \in \mathbb{N}, y \in \mathbb{V}_2) \\ \mathcal{G} &= \sigma(\varepsilon_y, y \in \mathbb{V}_2) \\ \mathcal{F}_n &= \sigma(\psi_i, i = 1, \dots, n),\end{aligned}$$

avec $\mathcal{F} \equiv \mathcal{F}_\infty$. Ainsi que la suite d'événements $A_n = A_{n,1} \cap A_{n,2}$ et B_n avec

$$\begin{aligned}A_{n,1} &= \{\omega \in \Omega : \max_{0 \leq k \leq 2n} |Y_k| < n^{1/2+\delta_1}\} \text{ pour un } \delta_1 > 0, \\ A_{n,2} &= \{\omega \in \Omega : \max_{y \in \mathbb{V}_2} \eta_{2n-1}(y) < n^{1/2+\delta_2}\} \text{ pour un } \delta_2 > 0, \\ B_n &= \{\omega \in A_n : \left| \sum_{y \in \mathbb{V}_2} \varepsilon_y \eta_{2n-1}(y) \right| > n^{1/2+\delta_3}\} \text{ pour un } \delta_3 > 0.\end{aligned}$$

Les δ_i sont supposés petits. On notera $d_{n,i} = n^{1/2+\delta_i}$ pour $i = 1, 2, 3$.

Remarquons que $A_{n,1}$, $A_{n,2}$ et donc A_n appartiennent à \mathcal{F}_{2n} et $B_n \in \mathcal{F}_{2n} \vee \mathcal{G}$.

Comme $B_n \subseteq A_n$ et que les 2 événements sont $(\mathcal{F}_{2n} \vee \mathcal{G})$ -mesurables, on peut faire la décomposition :

$$1 = \mathbb{1}_{B_n} + \mathbb{1}_{A_n \setminus B_n} + \mathbb{1}_{A_n^c},$$

On a par (\mathcal{F}_{2n}) -mesurabilité de Y_{2n} ,

$$\begin{aligned}\mathbb{P}(X_{2n} = 0; Y_{2n} = 0 | \mathcal{F} \vee \mathcal{G}) &= \mathbb{1}_{B_n} \mathbb{1}_{\{Y_{2n}=0\}} \mathbb{P}(X_{2n} = 0 | \mathcal{F} \vee \mathcal{G}) \\ &\quad + \mathbb{1}_{A_n \setminus B_n} \mathbb{1}_{\{Y_{2n}=0\}} \mathbb{P}(X_{2n} = 0 | \mathcal{F} \vee \mathcal{G}) \\ &\quad + \mathbb{1}_{A_n^c} \mathbb{1}_{\{Y_{2n}=0\}} \mathbb{P}(X_{2n} = 0 | \mathcal{F} \vee \mathcal{G}),\end{aligned}$$

On prend alors les espérances de chaque coté de l'égalité pour obtenir l'équation :

$$p_n = p_{n,1} + p_{n,2} + p_{n,3},$$

où

$$\begin{aligned}p_n &= \mathbb{P}(X_{2n} = 0; Y_{2n} = 0) \\ p_{n,1} &= \mathbb{P}(X_{2n} = 0; Y_{2n} = 0; B_n) \\ p_{n,2} &= \mathbb{P}(X_{2n} = 0; Y_{2n} = 0; A_n \setminus B_n) \\ p_{n,3} &= \mathbb{P}(X_{2n} = 0; Y_{2n} = 0; A_n^c).\end{aligned}$$

4.2 Preuve de la transience de la marche aléatoire \mathbb{O}_ε

Proposition 4.1. *Pour n assez grand, il existe $\delta > 0$ et $c > 0$ tel que*

$$p_{n,3} = O(\exp(-cn^\delta))$$

DÉMONSTRATION. Ecrivons $A_n^c = A_{n,1}^c \cup A_{n,2}^c$. On a

$$\mathbb{P}(A_{n,1}^c | Y_{2n} = 0) = \mathbb{P}(\max_{0 \leq k \leq 2n} |Y_k| \geq d_{n,1} | Y_{2n} = 0)$$

Et notons $a_n = \lfloor d_{n,1} \rfloor$ et $R_n = a_n, \dots, n$.

$$\begin{aligned} \mathbb{P}(A_{n,1}^c | Y_{2n} = 0) &= \sum_{y \in R_n} \mathbb{P}(\max_{0 \leq k \leq 2n} |Y_k| = y | Y_{2n} = 0) \\ &\leq 2 \sum_{y \in R_n} \mathbb{P}_0(Y_{2n} = 2y) \\ &= 2\mathbb{P}_0(Y_{2n} \geq 2a_n), \end{aligned}$$

La deuxième inéquation étant obtenue d'abord en décomposant (sur $Y_{2n} = 0$), l'événement $\{\max_{0 \leq k \leq 2n} |Y_k| = y\}$ en $\{\max_{0 \leq k \leq 2n} Y_k = y\} \cup \{\min_{0 \leq k \leq 2n} Y_k = -y\}$, puis en appliquant le principe de réflexion ainsi : en se plaçant sur le point de coordonnées (m, y) (respectivement $(m, -y)$) où $0 \leq m \leq 2n$, le nombre de chemins reliant (m, y) à $(2n, 0)$ est égal au nombre de chemins reliant (m, y) à $(2n, 2y)$ (resp. $(2n, -2y)$).

Maintenant

$$\begin{aligned} \mathbb{P}(A_{n,1}^c | Y_{2n} = 0) &\leq 2\mathbb{P}_0(Y_{2n} \geq 2a_n) \\ &\leq 2 \inf_{t>0} \mathbb{P}_0(\exp(tY_{2n}) \geq \exp(2ta_n)) \\ &\leq 2 \inf_{t>0} \frac{(\cosh(t))^{2n}}{\exp(2ta_n)} \\ &= 2 \inf_{t>0} \left(\cosh(t) \exp\left(-t \frac{a_n}{n}\right) \right)^{2n} \end{aligned}$$

Dans la deuxième inégalité, nous avons utilisé l'inégalité de Markov. Puis

$$\begin{aligned} \mathbb{E}[\exp(tY_{2n})] &= \mathbb{E}\left[\exp\left(t \sum_{k=0}^{2n} \psi_k\right)\right] \\ &= (\mathbb{E}[\exp(t\psi_1)])^{2n} \\ &= \left(\frac{\exp(t)}{2} + \frac{\exp(-t)}{2}\right)^{2n} \end{aligned}$$

Une simple étude de fonctions permet de trouver que l'infimum est atteint en $t = \operatorname{arctanh}\left(\frac{a_n}{n}\right)$. Puis en appliquant en ce t et en prenant n assez grand, on trouve

$$\begin{aligned}\mathbb{P}(A_{n,1}^c | Y_{2n} = 0) &\leq \exp\left(-\frac{a_n^2}{n}\right) \exp\left(o\left(\frac{a_n^2}{n}\right)\right) \\ &\leq \exp\left(-n^{2\delta_1}\right) + o\left(n^{2\delta_1}\right)\end{aligned}$$

De plus,

$$\begin{aligned}\mathbb{P}(A_{n,2}^c | Y_{2n} = 0) &= \mathbb{P}\left(\max_{y \in \mathbb{V}_2}(\eta_{2n-1}(y) \geq d_{n,2} | Y_{2n} = 0)\right) \\ &\leq \sum_{y \in \mathbb{V}_2} \frac{\mathbb{P}(\eta_{2n-1}(y) \geq d_{n,2})}{\mathbb{P}(Y_{2n} = 0)}\end{aligned}$$

Or,

$$\mathbb{P}_0(\eta_{2n-1}(y) \geq d_{n,2}) \leq \mathbb{P}_y(\sigma_{y,[d_{n,2}]} \leq 2n)$$

où $\sigma_{y,k}$ désigne le temps du k -ième retour au point y pour la marche aléatoire squelette.

Ainsi pour tout $t > 0$,

$$\begin{aligned}\mathbb{P}_0(\eta_{2n-1}(y) \geq d_{n,2}) &\leq \mathbb{P}_y(\sigma_{y,[d_{n,2}]} \leq 2n) \\ &\leq \mathbb{P}_y(\exp(-t\sigma_{y,[d_{n,2}]}) \geq \exp(-2nt)) \\ &\leq \exp(2nt) \mathbb{E}_0(\exp(-t\sigma_{[d_{n,2}]})) \\ &= \exp(2nt) (1 - \sqrt{1 - \exp(-2t)})^{[d_{n,2}]}, \quad \forall t > 0 \\ &\leq \inf_{t>0} \exp(2nt) (1 - \sqrt{1 - \exp(-2t)})^{[d_{n,2}]}\end{aligned}$$

où l'on a utilisé l'inégalité de Markov puis la propriété de Markov dans la troisième inégalité.

L'infimum est atteint en $t = -\frac{1}{2} \ln\left(\frac{n^2 - n[d_{n,2}]}{([d_{n,2}]/2 - n)^2}\right)$. En appliquant alors cette valeur de t et en prenant n assez grand on trouve

$$\mathbb{P}_0(\eta_{2n-1}(y) \geq d_{n,2}) \leq \exp\left(-\frac{1}{4}n^{2\delta_2}\right) + o\left(n^{2\delta_2}\right)$$

Pour conclure la preuve, remarquons d'abord que $\mathbb{P}_0(Y_{2n}) = O\left(\frac{1}{\sqrt{n}}\right)$. En effet le nombre de chemins de $(0,0)$ à $(2n,0)$ est exactement $\binom{2n}{n}$, et il y a

2^{2n} trajectoires possibles de longueur $2n$. Ce qui implique que $\mathbb{P}_0(Y_{2n}) = \frac{\binom{2n}{n}}{2^{2n}}$. Il nous reste alors plus qu'à utiliser la formule de Stirling pour trouver le résultat.

Comme la somme $\sum_{y \in \mathbb{V}_2}$ est effectuée sur l'événement $\{Y_{2n} = 0\}$, qui contient au plus $2n + 1$ éléments, on a

$$\mathbb{P}(A_{n,2}^c | Y_{2n} = 0) \leq O(\sqrt{n})(2n + 1) \exp(-cn^{2\delta_2})$$

pour un certain $c > 0$.

On choisit alors $0 < \delta'_2 < \delta_2$ et $\delta = \min(2\delta_1, \delta'_2)$. Ainsi,

$$\mathbb{P}(A_n^c | Y_{2n} = 0) \leq 2 \exp(-n^{2\delta-1}) o(\exp(-cn^{\delta-1})) + C \exp(-cn^{\delta'_2}) = O(\exp(-cn^\delta))$$

Ceci conclut alors la preuve car,

$$\begin{aligned} p_{n,3} &\leq \mathbb{P}(A_n^c, Y_{2n} = 0) \\ &= \mathbb{P}(A_n^c | Y_{2n}=0) \mathbb{P}(Y_{2n}=0) \end{aligned}$$

et on utilise à nouveau l'estimation de $\mathbb{P}(Y_{2n} = 0)$. □

Corollaire 4.2.

$$\sum_{n \in \mathbb{N}} p_{n,3} < \infty$$

Rappelons nous maintenant que

$$X_{2n} = \sum_{y \in \mathbb{V}_2} \varepsilon_y \sum_{i=1}^{\eta_{2n-1}(y)} \xi_i^{(y)} = \sum_{k=1}^{2n} \varepsilon_{Y_k} \xi_k$$

et introduisons les variables aléatoires suivantes :

$$\begin{aligned} N_+ &= \sum_{k=1}^{2n} \mathbf{1}_{\{\varepsilon_{Y_k}=1\}} \\ N_- &= \sum_{k=1}^{2n} \mathbf{1}_{\{\varepsilon_{Y_k}=-1\}} \\ \Delta_n &= N_+ - N_- = \sum_{y \in \mathbb{V}_2} \varepsilon_y \eta_{2n-1}(y). \end{aligned}$$

Proposition 4.3. *Pour n assez grand, pour tout $\delta' \in]0, 2\delta_3[$, on a*

$$p_{n,1} = O(\exp(-n^{\delta'}))$$

DÉMONSTRATION. On doit montrer que pour n grand, $\mathbb{P}(X_{2n} = 0, Y_{2n} = 0, B_n) = O(\exp(-n^{\delta'}))$. Posons $m_1 = \mathbb{E}[\xi_1]$ et $m_2 = \mathbb{E}[\xi_1^2]$ et $s^2 = m_2 - m_1^2$. En remarquant que N_+ , N_- et Δ_n sont $\mathcal{F}_{2n} \vee \mathcal{G}$ -mesurables et que $N_+ + N_- = 2n$ on a par calcul

$$\begin{aligned}\mathbb{E}[X_{2n} | \mathcal{F}_{2n} \vee \mathcal{G}] &= m_1 \Delta_n \\ \mathbb{E}[X_{2n}^2 | \mathcal{F}_{2n} \vee \mathcal{G}] &= 2ns^2 + m_1^2 \Delta_n^2 \\ \mathbb{V}(X_{2n} | \mathcal{F}_{2n} \vee \mathcal{G}) &= 2ns^2.\end{aligned}$$

Pour $t \in]-\infty, -\ln(q)[$, on définit la fonction génératrice de la v.a. ξ_1 à savoir $\phi(t) = \mathbb{E}[\exp(t\xi_1)]$. Par définition de ξ_1 , pour des petites valeurs de $|t|$, la fonction génératrice se comporte comme $\exp(tm_1 + t^2 s^2/2 + O(t^3))$. Ainsi on obtient,

$$\mathbb{E}[\exp(tX_{2n}) | \mathcal{F}_{2n} \vee \mathcal{G}] = \phi(t)^{N_+} \phi(-t)^{N_-} = \exp(tm_1 \Delta_n + t^2 s^2 n + O(t^3 n))$$

Sur l'événement $\{\Delta_n > d_{n,3}\}$, on a pour $t < 0$,

$$\begin{aligned}\mathbb{P}(X_{2n} = 0 | \mathcal{F}_{2n} \vee \mathcal{G}) &\leq \mathbb{P}(X_{2n} \leq 0 | \mathcal{F}_{2n} \vee \mathcal{G}) \\ &= \mathbb{P}(\exp(tX_{2n}) \geq 1 | \mathcal{F}_{2n} \vee \mathcal{G}) \\ &\leq \mathbb{E}[\exp(tX_{2n}) | \mathcal{F}_{2n} \vee \mathcal{G}] \\ &= \exp(tm_1 \Delta_n + t^2 s^2 n + O(t^3 n))\end{aligned}$$

Maintenant, en prenant $t = -\frac{m_1 n^{\delta_3 - 1/2}}{2s^2}$ pour n suffisamment grand,

$$\mathbb{P}(X_{2n} = 0 | \mathcal{F}_{2n} \vee \mathcal{G}) \leq \exp\left(-\frac{m_1^2}{4s^2} n^{2\delta_3} + O(n^{3\delta_3 - 1/2})\right)$$

En raisonnant de manière similaire sur l'événement $\{\Delta_n < -d_{n,3}\}$, on a pour $t \in]0, -\ln q[$,

$$\mathbb{P}(X_{2n} = 0 | \mathcal{F}_{2n} \vee \mathcal{G}) \leq \exp(tm_1 \Delta_n + t^2 s^2 n + O(t^3 n))$$

puis on prend $t = \frac{m_1 n^{\delta_3 - 1/2}}{2s^2}$ pour n suffisamment grand.

Ceci conclut alors la preuve. En effet il suffit d'écrire

$$\begin{aligned}p_{n,1} &= \mathbb{P}(X_{2n} = 0, Y_{2n} = 0, B_n) \\ &\leq \mathbb{P}(X_{2n} = 0, B_n) \\ &= \mathbb{E}[\mathbb{1}_{B_n} \mathbb{P}(X_{2n} = 0 | \mathcal{F}_{2n} \vee \mathcal{G})] \\ &= \mathbb{E}[\mathbb{1}_{D_n > d_{n,3}} \mathbb{P}(X_{2n} = 0 | \mathcal{F}_{2n} \vee \mathcal{G})] + \mathbb{E}[\mathbb{1}_{D_n < -d_{n,3}} \mathbb{P}(X_{2n} = 0 | \mathcal{F}_{2n} \vee \mathcal{G})]\end{aligned}$$

où l'on a utilisé le fait que B_n est $\mathcal{F}_{2n} \vee \mathcal{G}$ -mesurable. Il ne reste plus qu'à utiliser les résultats trouvés plus haut pour finir. \square

Corollaire 4.4.

$$\sum_{n \in \mathbb{N}} p_{n,1} < \infty$$

On admet le lemme technique suivant.

Lemme 4.5. *Sur l'événement $A_n \setminus B_n$, on a*

$$\mathbb{P}(X_{2n} = 0 | \mathcal{F} \vee \mathcal{G}) = O\left(\sqrt{\frac{\ln n}{n}}\right)$$

Proposition 4.6. *Pour tout $\delta_5 > 0$, et pour n assez grand n*

$$\mathbb{P}(A_n \setminus B_n | \mathcal{F}) = O(n^{-1/4+\delta_5})$$

DÉMONSTRATION. Notons $\zeta_y = \varepsilon_y \eta_{2n-1}(y)$ et soit G une v.a. gaussienne centrée de variance $d_{n,3}^2$ (conditionnellement à \mathcal{F}) et indépendante des ζ_y . On a

$$\mathbb{P}\left(\left|\sum_{y \in \mathbb{V}_2} \zeta_y\right| \leq d_{n,3} \mid \mathcal{F}\right) \mathbb{P}(|G| \leq d_{n,3} \mid \mathcal{F}) \leq \mathbb{P}\left(\left|\sum_{y \in \mathbb{V}_2} \zeta_y + G\right| \leq 2d_{n,3} \mid \mathcal{F}\right)$$

donc

$$\mathbb{P}\left(\left|\sum_{y \in \mathbb{V}_2} \zeta_y\right| \leq d_{n,3} \mid \mathcal{F}\right) \leq \frac{1}{\mathbb{P}(|G| \leq d_{n,3} \mid \mathcal{F})} \mathbb{P}\left(\left|\sum_{y \in \mathbb{V}_2} \zeta_y + G\right| \leq 2d_{n,3} \mid \mathcal{F}\right)$$

Soient

$$\chi_2(t) = \mathbb{E}[\exp(it \sum_y \zeta_y) | \mathcal{F}] = \prod_y \cos(\eta_{2n-1}(y)t)$$

et

$$\chi_3(t) = \mathbb{E}[\exp(itG) | \mathcal{F}] = \exp(-t^2 d_{n,3}^2 / 2)$$

Ainsi par indépendance,

$$\mathbb{E}[\exp(it(\sum_y \zeta_y + G)) | \mathcal{F}] = \chi_2(t) \chi_3(t)$$

Soit $\delta > 0$, on a

$$\begin{aligned} \mathbb{P}\left(\left|\sum_y \zeta_y + G\right| \leq \delta \mid \mathcal{F}\right) &= \int_{-\delta}^{\delta} \frac{1}{2\pi} \int_{\mathbb{R}} \chi_2(t) \chi_3(t) \exp(-itu) dt du \\ &= \int_{\mathbb{R}} \int_{-\delta}^{\delta} \frac{1}{2\pi} \chi_2(t) \chi_3(t) \exp(-itu) dt du \\ &= \int_{\mathbb{R}} \chi_2(t) \chi_3(t) \frac{1}{\pi} \frac{\sin(t\delta)}{t} dt \end{aligned}$$

Ainsi on a

$$\mathbb{P}\left(\left|\sum_{y \in \mathbb{V}_2} \zeta_y + G\right| \leq d_{n,3} \mid \mathcal{F}\right) \leq C d_{n,3} I$$

où

$$I = \int \prod_y \cos(\eta_{2n-1}(y)t) \exp(-t^2 d_{n,3}^2/2) dt$$

Fixons maintenant $b_n = \frac{n^{\delta_4}}{d_{n,3}}$ pour un $\delta_4 > 0$ et séparons l'intégrale I en $I_1 + I_2$, où pour l'intégration se fait sur $|t| \leq b_n$ pour I_1 et sur $|t| > b_n$ pour I_2 .

On a

$$\begin{aligned} I_2 &\leq C \int_{|t| > b_n} \exp(-t^2 d_{n,3}^2/2) \frac{dt}{2\pi} \\ &= \frac{C}{d_{n,3}} \int_{|s| > n^{\delta_4}} \exp(-s^2/2) \frac{ds}{2\pi} \\ &\leq 2 \frac{C}{d_{n,3}} \frac{1}{n^{\delta_4}} \frac{\exp(-n^{2\delta_4}/2)}{2\pi}, \end{aligned}$$

où l'on a utilisé le fait que si Z suit une loi normale centrée réduite de densité notée ϕ , alors $P(Z > x) \leq \frac{\phi(x)}{x}$.

Pour I_1 on a

$$I_1 \leq \int_{|t| \leq b_n} \prod_y |\cos(\eta_{2n-1}(y)t)| dt$$

Comme $\sum_y \frac{\eta_{2n-1}(y)}{2n} = 1$, par l'inégalité de Hölder,

$$\begin{aligned} I_1 &\leq \prod_y \left[\left(\int_{|t| \leq b_n} |\cos(\eta_{2n-1}(y)t)|^{\frac{2n}{\eta_{2n-1}(y)}} dt \right)^{\frac{\eta_{2n-1}(y)}{2n}} \right] \\ &\leq \sup_{y: \eta_{2n-1}(y) \neq 0} \int_{|t| \leq b_n} |\cos(\eta_{2n-1}(y)t)|^{\frac{2n}{\eta_{2n-1}(y)}} dt, \end{aligned}$$

Sur l'événement A_n , pour tout y tel que $\eta_{2n-1}(y) \neq 0$, on a $|\eta_{2n-1}(y)| \leq b_n d_{n,2} = \frac{d_{n,2}}{d_{n,3}} n^{\delta_4}$ où l'on choisit $\delta_2, \delta_3, \delta_4$ tels que $\delta_2 + \delta_4 - \delta_3 = -\delta_5 < 0$. Pour

ces y ,

$$\begin{aligned} I_1 &\leq 2b_n \int_{|v| \leq \pi/2} \left| \cos\left(\frac{\eta_{2n-1}(y)}{d_{n,3}} n^{\delta_4} v\right) \right| \frac{2n}{\eta_{2n-1}(y)} dv \\ &\leq 2b_n \int_{|v| \leq \pi/2} \left| \cos v \right| \frac{2n}{\eta_{2n-1}(y)} dv. \end{aligned}$$

Pour $|v| < \pi/2$ on a $|\cos v| \leq \exp(-cv^2/2)$ et donc

$$\begin{aligned} I_1 &\leq 2b_n \int \exp\left(-c \frac{2nv^2}{\eta_{2n-1}(y)}\right) dv \\ &\leq cb_n \frac{\eta_{2n-1}(y)^{1/2}}{n^{1/2}} \int \exp(-v^2/2) \frac{dv}{\sqrt{2\pi}} \\ &\leq c \frac{d_{n,2}^{1/2}}{d_{n,3}} \frac{n^{\delta_4}}{n^{1/2}} \end{aligned}$$

Et finalement, la probabilité totale est majorée par $O(d_{n,3}(I_1 + I_2)) = O(n^{-1/4})$. \square

Corollaire 4.7.

$$\sum_{n \in \mathbb{N}} p_{n,2} < \infty$$

DÉMONSTRATION. On rappelle que pour la marche aléatoire simple, $\mathbb{P}(Y_{2n} = 0) = O(n^{-1/2})$ et grâce aux estimations obtenues précédemment

$$\begin{aligned} p_{n,2} &= \mathbb{P}(X_{2n} = 0; Y_{2n} = 0; A_n \setminus B_n) \\ &= \mathbb{E} \left[\mathbb{E} \left[\mathbf{1}_{Y_{2n}=0} \left(\mathbb{E}[\mathbf{1}_{A_n \setminus B_n} \mathbb{P}(X_{2n} = 0 \mid \mathcal{F} \vee \mathcal{G}) \mid \mathcal{F}] \right) \right] \right] \\ &= O(n^{-1/2} n^{-1/4} \sqrt{\frac{\ln n}{n}}) \\ &= O(n^{-5/4} \ln n) \end{aligned}$$

ce qui prouve la sommabilité de $p_{n,2}$. \square

Théorème 4.8. *Pour presque toutes réalisations de l'événement aléatoire ε , la marche aléatoire sur le réseau \mathbb{O}_ε est transiente.*

DÉMONSTRATION. La transience est une simple conséquence des propositions précédentes. On écrit, $p_n = p_{n,1} + p_{n,2} + p_{n,3}$ qui est sommable car on a montré que les $p_{n,i}$ sont toutes sommables. \square

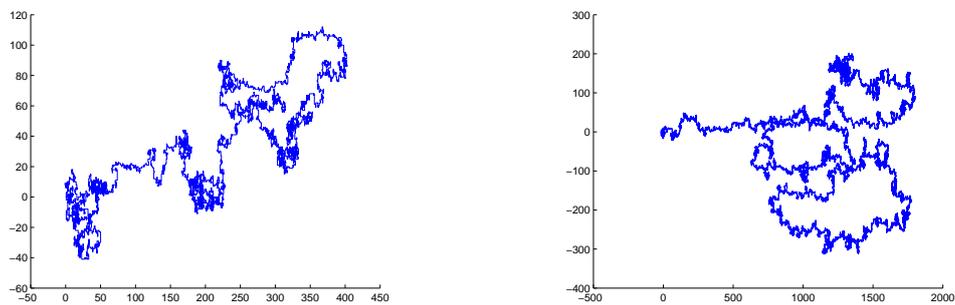


FIGURE 6 – Exemple de marche aléatoire \mathbb{O}_ε , pour ε une suite de v.a. de Rademacher, sur 10 000 et 100 000 pas.

Références

- [1] Massimo Campanino and Dimitri Petritis. Random walks on randomly oriented lattices. 2001.
- [2] R. Durrett. *Probability : Theory and Examples*. Cambridge Series in Statistical and Probabilistic Mathematics. Cambridge University Press, 2010.
- [3] B.E. Fristedt and L.F. Gray. *A Modern Approach to Probability Theory*. Probability and Its Applications. Birkhäuser Boston, 1996.
- [4] J.F. Le Gall. *Intégration, Probabilités et Processus Aléatoires*. FIMFA, 2006.

A.3 Mémoire de M2 : Inégalités de Concentration

Université de Paris-Sud 11

Mémoire de Master 2 Probabilités et Statistiques

Inégalités de Concentration

Antoine MARCHINA

15 septembre 2014

Sous la direction d'Emmanuel Rio

Notations

$\text{Int}(A)$	Intérieur de l'ensemble A
∂A	Frontière de l'ensemble A
$\nabla g(z)$	Gradient de la fonction g en z
x_+	Partie positive de x
v.a.	Variable Aléatoire
iid	Indépendante et Identiquement Distribuée
$\phi(x)$	Fonction de distribution de la loi normale standard
\mathbb{P}_X	Loi de la v.a. X
$\mathcal{B}(a, b), a \leq 0 \leq b$	Loi de Bernoulli prenant la valeur a avec probabilité $b/(b-a)$ et b avec $-a/(b-a)$
$\mathcal{B}(p), 0 \leq p \leq 1$	Loi de Bernoulli prenant la valeur 1 avec probabilité p et 0 avec $1-p$
(E, \mathcal{E})	Espace mesurable
$b\mathcal{E}$	Espace des fonctions bornées réelles mesurables sur (E, \mathcal{E})
$l_\infty(\mathcal{F}), \mathcal{F} \subset b\mathcal{E}$	Espace des fonctions réelles bornées sur \mathcal{F}
$\ \cdot\ _{\mathcal{F}}$	Norme sup sur $l_\infty(\mathcal{F})$ définie par $\ F\ _{\mathcal{F}} = \sup_{f \in \mathcal{F}} F(f) $
$M_1(E)$	Espace des mesures de probabilités sur (E, \mathcal{E})
τ	Topologie la plus fine sur $M_1(E)$ qui rend les fonctions $\nu \mapsto \int_E f d\nu$ continue pour tout $f : E \rightarrow \mathbb{R}$ bornée, \mathcal{E} -mesurable
$d_2(f, g)$	$(\int_E (f-g)^2 d\mu)^{1/2}$
PGD	Principe de Grandes Déviations
$\nu^{\mathcal{F}}, \nu \in M_1(E), \mathcal{F} \subset b\mathcal{E}$	Elément de $l_\infty(\mathcal{F})$ tel que pour tout $f \in \mathcal{F}$ $\nu^{\mathcal{F}}(f) = \nu(f) \stackrel{\text{def}}{=} \int f d\nu$
$H(\nu \mu)$	Entropie relative de ν par rapport à μ définie par $\int_E d \log(d) d\mu$ si $\nu \ll \mu$ où $d = d\nu/d\mu$ et $+\infty$ sinon

Table des matières

1	On Hoeffding's Inequalities	4
1.1	Enoncés	4
1.2	Concentration Convexe	5
1.3	Démonstrations	12
1.4	Commentaires	17
2	Large Deviations	20
2.1	Théorème	21
2.2	Complément	24
A	Enveloppe log-concave	27
B	Schur-concavité	28
C	Principes de Grandes Déviations	31
	Références	32

1 On Hoeffding's Inequalities

Cette partie repose sur l'étude de l'article de V. Bentkus [3].

On se place dans un espace mesurable (Ω, \mathcal{F}) et on considère une famille de tribus $\mathcal{F}_0 = \emptyset \subset \mathcal{F}_1 \subset \dots \subset \mathcal{F}_n \subset \mathcal{F}$. On considère une \mathcal{F}_n -martingale $M_n = \sum_{i=1}^n X_i$, $M_0 = 0$. Dans toute cette partie on note $s_k^2 = \mathbb{E}[X_k^2 | \mathcal{F}_{k-1}]$. On étudiera cette martingale sous différentes hypothèses de bornitude sur ses accroissements :

1. Pour tout k , il existe p_k tel que $\mathbb{P}(-p_k \leq X_k \leq 1 - p_k) = 1$.
2. Il existe $b > 0$ tel que pour tout k , $\mathbb{P}(X_k \leq b) = 1$ et pour tout k , il existe $\sigma_k^2 \geq 0$ tel que $\mathbb{P}(s_k^2 \leq \sigma_k^2) = 1$.
3. Pour tout k , il existe $b_k > 0$ et $\sigma_k^2 \geq 0$ tels que $\mathbb{P}(X_k \leq b_k) = 1$ et $\mathbb{P}(s_k^2 \leq \sigma_k^2) = 1$.

Notre intérêt se porte sur la probabilité de queue de la martingale i.e. $\mathbb{P}(M_n \geq x)$ pour tout $x \in \mathbb{R}$. Essentiellement, on montre (cf. Théorèmes 1.1, 1.2 et 1.3) que l'on peut majorer cette probabilité par la probabilité de queue d'une somme de n v.a. iid de Bernoulli.

Décrivons la démarche adoptée dans l'article. On commence par majorer de manière classique $\mathbb{P}(M_n \geq x)$: Soit f une fonction croissante positive, par l'inégalité de Markov on a

$$\mathbb{P}(M_n \geq x) \leq \frac{1}{f(x)} \mathbb{E}[f(M_n)]$$

Pour majorer $\mathbb{E}[f(M_n)]$ on cherche des résultats de concentration convexe :

$$\mathbb{E}[f(M_n)] \leq \mathbb{E}[f(S_n)]$$

où $f \in \mathcal{C}$ et \mathcal{C} est un ensemble de fonctions convexes et S_n est une somme de n v.a. iid de Bernoulli.

Commençons par énoncer les théorèmes.

1.1 Énoncés

Théorème 1.1. *Supposons que les accroissements X_k d'une martingale M_n vérifient $\mathbb{P}(-p_k \leq X_k \leq 1 - p_k) = 1$ où $0 \leq p_k \leq 1$. Alors pour tout $x \in \mathbb{R}$,*

$$\mathbb{P}(M_n \geq x) \leq e\mathbb{P}^\circ(S_n \geq x) \tag{1.1}$$

où $S_n = \sum_{i=1}^n \varepsilon_i$, où les ε_i sont des v.a. iid de loi $\mathcal{B}(-p, 1-p)$ avec $p = \sum_{i=1}^n p_i/n$. De plus, on a

$$\mathbb{P}(M_n \geq x) \leq \frac{e^3}{2} \mathbb{P}^\circ \left(\eta \geq \lambda + \frac{x}{1-p} \right) \quad (1.2)$$

où $\eta \sim \mathcal{P}(\lambda)$ avec $\lambda = pn/(1-p)$.

Théorème 1.2. *Supposons que les accroissements X_k d'une martingale M_n vérifient $\mathbb{P}(X_k \leq b) = 1$ et $\mathbb{P}(s_k^2 \leq \sigma_k^2) = 1$ où $b > 0, \sigma_k^2 \geq 0$. Alors pour tout $x \in \mathbb{R}$,*

$$\mathbb{P}(M_n \geq x) \leq \frac{e^2}{2} \mathbb{P}^\circ(S_n \geq x) \quad (1.3)$$

où $S_n = \sum_{i=1}^n \varepsilon_i$ où les ε_i sont des v.a. iid de loi $\mathcal{B}(-\sigma^2/b, b)$ avec $\sigma^2 = \sum_{i=1}^n \sigma_i^2/n$.

De plus, on a

$$\mathbb{P}(M_n \geq x) \leq \frac{e^2}{2} \mathbb{P}^\circ \left(\eta \geq \lambda + \frac{x}{b} \right) \quad (1.4)$$

où $\eta \sim \mathcal{P}(\lambda)$ avec $\lambda = \sum_{i=1}^n \sigma_i^2/b^2$.

Théorème 1.3. *Supposons que les accroissements X_k d'une martingale M_n vérifient $\mathbb{P}(X_k \leq b_k) = 1$ et $\mathbb{P}(s_k^2 \leq \sigma_k^2) = 1$ où $b_k > 0, \sigma_k^2 \geq 0$. Alors pour tout $x \in \mathbb{R}$,*

$$\mathbb{P}(M_n \geq x) \leq \frac{2e^3}{9} \mathbb{P}^\circ(S_n \geq x) \quad (1.5)$$

où $S_n = \sum_{i=1}^n \varepsilon_i$, où les ε_i sont des v.a. iid de loi $\mathcal{B}(-a, a)$, avec $a_k = \max\{b_k, \sigma_k\}$ et $a^2 = \sum_{i=1}^n a_i^2/n$.

On a de plus,

$$\mathbb{P}(M_n \geq x\sqrt{n}) \leq \frac{2e^3}{9} (1 - \phi(x/a)) \quad (1.6)$$

1.2 Concentration Convexe

Dans un premier temps, on va chercher à majorer l'espérance de $f(M_n)$ par l'espérance de $f(T_n)$ où T_n est la somme de n v.a. de Bernoulli indépendantes (non nécessairement de même loi). Les 3 premiers lemmes ci-dessous couvrent chacun une hypothèse faite sur les accroissements de la martingale M_n dans l'introduction.

Lemme 1.4. *i. Soit $f : \mathbb{R} \rightarrow \mathbb{R}$ une fonction convexe. Soit X une v.a. telle que $\mathbb{E}[X] = 0$ et $\mathbb{P}(a \leq X \leq b) = 1$ où $a \leq 0 \leq b$. Alors on a :*

$$\mathbb{E}[f(X)] \leq \mathbb{E}[f(\xi)]$$

où $\xi \sim \mathcal{B}(a, b)$.

ii. Soit $f : \mathbb{R}^n \rightarrow \mathbb{R}$ une fonction convexe de chaque variable x_1, \dots, x_n quand les autres $n - 1$ variables sont fixées. On suppose que les accroissements X_k d'une martingale M_n satisfont $\mathbb{P}(a_k \leq X_k \leq b_k) = 1$ où $a_k \leq 0 \leq b_k$. Soit $\xi_k \sim \mathcal{B}(a_k, b_k)$ des v.a. indépendantes. Alors :

$$\mathbb{E}[f(X_1, \dots, X_n)] \leq \mathbb{E}[f(\xi_1, \dots, \xi_n)]$$

DÉMONSTRATION. *i. Soit $u : [a, b] \rightarrow \mathbb{R}$ une fonction linéaire telle que $u(a) = f(a)$ et $u(b) = f(b)$. On a donc par convexité, $f(x) \leq u(x)$ sur $[a, b]$ et de plus, comme $\xi \in \{a, b\}$ p.s., $\mathbb{E}[u(\xi)] = \mathbb{E}[f(\xi)]$. En conclusion,*

$$\mathbb{E}[f(X)] \leq \mathbb{E}[u(X)] = \mathbb{E}[u(\xi)] = \mathbb{E}[f(\xi)]$$

ii. Pour $n = 1$ il s'agit du point i. On suppose ii. vérifié pour $1, \dots, n - 1$ (avec $n > 1$).

On définit $Z_0 = 0, Z_1 = X_2, \dots, Z_{n-1} = \sum_{k=2}^n X_k$. On remarque que pour X_1 donnée, $(Z_k)_{0 \leq k \leq n-1}$ définit une martingale dont les accroissements vérifient

$$\mathbb{P}(a_{k+1} \leq Z_k - Z_{k-1} \leq b_{k+1}) = 1$$

Ainsi,

$$\begin{aligned} \mathbb{E}[f(X_1, \dots, X_n)] &= \mathbb{E}[\mathbb{E}[f(X_1, \dots, X_n) \mid X_1]] \\ &\leq \mathbb{E}[\mathbb{E}[f(X_1, \xi_2, \dots, \xi_n) \mid X_1]] \\ &= \mathbb{E}[\mathbb{E}[f(X_1, \xi_2, \dots, \xi_n) \mid \xi_2, \dots, \xi_n]] \\ &\leq \mathbb{E}[\mathbb{E}[f(\xi_1, \dots, \xi_n) \mid \xi_2, \dots, \xi_n]] \\ &= \mathbb{E}[f(\xi_1, \dots, \xi_n)] \end{aligned}$$

où l'on a utilisé l'hypothèse de récurrence aux 2^{ème} et 4^{ème} lignes. \square

Lemme 1.5. *Soit $f_t : \mathbb{R} \rightarrow \mathbb{R}$ l'une des fonctions suivantes :*

$$\begin{aligned} f_t(x) &= (x - t)_+^2, & t \in \mathbb{R} \\ f_t(x) &= (x - t)_+^s, & s > 2, t \in \mathbb{R} \\ f_t(x) &= \exp(tx), & t > 0 \end{aligned}$$

i. Soit X une v.a. centrée telle que $\mathbb{P}(X \leq b) = 1$ et $\mathbb{E}[X^2] \leq \sigma^2$ où $b > 0, \sigma \geq 0$. Alors :

$$\mathbb{E}[f_t(X)] \leq \mathbb{E}[f_t(\theta)]$$

où $\theta \sim \mathcal{B}(-\sigma^2/b, b)$.

ii. On suppose que les accroissements X_k d'une martingale M_n satisfont $\mathbb{P}(X_k \leq b_k) = 1$ et $\mathbb{P}(s_k^2 \leq \sigma_k^2) = 1$ où $b_k > 0, \sigma_k \geq 0$. Soit $T_n = \sum_{k=1}^n \theta_k$ où les θ_k sont des v.a. indépendantes et $\theta_k \sim \mathcal{B}(-\sigma_k^2/b_k, b_k)$. Alors,

$$\mathbb{E}[f_t(M_n)] \leq \mathbb{E}[f_t(T_n)]$$

DÉMONSTRATION. Il suffit de prouver le théorème avec $f_t(x) = (x - t)_+^2$ où $t \in \mathbb{R}$.

En effet, soit g l'une des 2 autres fonctions, on peut écrire :

$$g(x) = \int_{\mathbb{R}} g(u) \delta_0(x - u) du \quad (1.7)$$

$$= \int_{\mathbb{R}} g'(u) H(x - u) du \quad (1.8)$$

$$= \int_{\mathbb{R}} g''(u) (x - u)_+ du \quad (1.9)$$

$$= \frac{1}{2} \int_{\mathbb{R}} g^{(3)}(u) (x - u)_+^2 du \quad (1.10)$$

Comme $g^{(3)}$ est positive, l'inégalité $\mathbb{E}[(M_n - u)_+^2] \leq \mathbb{E}[(T_n - u)_+^2]$ pour tout $u \in \mathbb{R}$ implique $\mathbb{E}[g(M_n)] \leq \mathbb{E}[g(T_n)]$.

Soit donc $f_t(x) = (x - t)_+^2$ où $t \in \mathbb{R}$.

i. Supposons que $t \leq -\sigma^2/b$. On a :

$$\mathbb{E}[f_t(X)] \leq \sigma^2 + t^2 = \mathbb{E}[f_t(\theta)]$$

Le cas $t \geq b$ est trivial. Il reste le cas $-\sigma^2/b < t < b$ à traiter.

On définit sur $[t, b]$

$$x \mapsto \frac{(x - t)_+}{(x + \sigma^2/b)^2}$$

Cette fonction est croissante et son maximum est atteint en b . Ainsi,

$$\frac{(x - t)_+}{(x + \sigma^2/b)^2} \leq \frac{b^2(b - t)^2}{(b^2 + \sigma^2)^2}$$

Et donc

$$(x - t)_+^2 \leq c \left(x + \frac{\sigma^2}{b} \right)^2$$

où $c = b^2(b-t)^2/(b^2 + \sigma^2)^2$. Par suite,

$$\mathbb{E}[f_t(X)] \leq c\mathbb{E}\left[\left(X + \sigma^2/b\right)^2\right] \leq c\left(\sigma^2 + \sigma^4/b^2\right) = \mathbb{E}[f_t(\theta)]$$

ii. Pour $n = 1$ il s'agit du point i. On suppose ii. vérifié pour $1, \dots, n-1$ (avec $n > 1$).

On considère la même suite $(Z_k)_{1 \leq k \leq n-1}$ que dans la démonstration du Lemme 1.5. Conditionnellement à X_1 , cette suite est une martingale dont les accroissements vérifient $\mathbb{P}(Z_k - Z_{k-1} \leq b_{k+1}) = 1$ et $\mathbb{E}\left[(Z_k - Z_{k-1})^2 \mid Z_1, \dots, Z_{k-1}\right] \leq \sigma_{k+1}^2$. Alors,

$$\begin{aligned} \mathbb{E}[f_t(M_n)] &= \mathbb{E}[\mathbb{E}[f_t(X_1 + \dots + X_n) \mid X_1]] \\ &\leq \mathbb{E}[\mathbb{E}[f_t(X_1 + \theta_2 + \dots + \theta_n) \mid X_1]] \\ &= \mathbb{E}[\mathbb{E}[f_t(X_1 + \theta_2 + \dots + \theta_n) \mid \theta_2, \dots, \theta_n]] \\ &\leq \mathbb{E}[\mathbb{E}[f_t(\theta_1 + \dots + \theta_n) \mid \theta_2, \dots, \theta_n]] \\ &= \mathbb{E}[f_t(\theta_1 + \dots + \theta_n)] \end{aligned}$$

où l'on a utilisé l'hypothèse de récurrence aux 2^{ème} et 4^{ème} lignes. \square

Lemme 1.6. Soit f_t l'une des fonctions du Lemme 1.5.

i. On suppose que X est une v.a. centrée telle que $\mathbb{P}(X \leq b) = 1$ et $\mathbb{E}[X^2] \leq \sigma^2$ où $b > 0, \sigma^2 \geq 0$. Alors on a

$$\mathbb{E}[f_t(X)] \leq \mathbb{E}[f_t(\theta)]$$

où $\theta \sim \mathcal{B}(-a, a)$ avec $a = \max\{\sigma, b\}$.

ii. On suppose que M_n est une martingale dont les accroissements X_k vérifient $\mathbb{P}(X_k \leq b_k) = 1$ et $\mathbb{P}(s_k^2 \leq \sigma_k^2) = 1$ où $b_k > 0, \sigma_k \geq 0$. Soit $T_n = \sum_{i=1}^n \theta_i$ où les θ_i sont des v.a. indépendantes et telles que $\theta_k \sim \mathcal{B}(-a_k, a_k)$ avec $a_k = \max\{\sigma_k, b_k\}$. Alors

$$\mathbb{E}[f_t(M_n)] \leq \mathbb{E}[f_t(T_n)]$$

DÉMONSTRATION. Comme dans le Lemme 1.5, il suffit de prouver i.

Supposons que $\sigma \leq b$. Le point i. du Lemme 1.5 implique $\mathbb{E}[f_t(X)] \leq \mathbb{E}[f_t(\theta_0)]$ où $\theta_0 \sim \mathcal{B}(-\sigma^2/b, b)$. De plus on a $\mathbb{P}(\theta_0 \leq b) = 1$ et $\mathbb{E}[\theta_0^2] \leq b^2$. Ainsi,

$$\mathbb{E}[f_t(\theta_0)] \leq \mathbb{E}[f_t(\theta)]$$

où $\theta \sim \mathcal{B}(-a, a)$ avec $a = \max\{\sigma, b\} = b$. Et on a donc bien,

$$\mathbb{E}[f_t(X)] \leq \mathbb{E}[f_t(\theta)]$$

Supposons que $\sigma > b$. Les conditions sur X implique $\mathbb{P}(X \leq \sigma) = 1$ et $\mathbb{E}[X^2] \leq \sigma^2$. On applique alors le point i. du Lemme 1.5 avec $a = \max\{\sigma, b\} = \sigma$. \square

Maintenant, par rapport à l'objectif, il faut majorer l'espérance de $f(T_n)$ où T_n est une somme de n v.a. de Bernoulli indépendantes par l'espérance de $f(S_n)$ où S_n est une somme de n v.a. iid de Bernoulli. De nouveau les 3 lemmes ci-dessous couvrent chacun une des hypothèses dans l'introduction. Ici, on aura besoin de la notion de Schur-concavité. Les résultats que l'on utilisera sont présentés dans l'Annexe B.

Lemme 1.7. *Soit f une fonction convexe à valeurs réelles. Soit $T_n = \sum_{i=1}^n \xi_i$ où les ξ_i sont des v.a. indépendantes telles que $\xi_i \sim \mathcal{B}(-p_i, 1 - p_i)$. Soit $S_n = \sum_{i=1}^n \varepsilon_i$ où les ε_i sont des v.a. iid de loi $\mathcal{B}(-p, 1 - p)$ où $p = \sum_{i=1}^n p_i/n$. Alors on a*

$$\mathbb{E}[f(T_n)] \leq \mathbb{E}[f(S_n)]$$

DÉMONSTRATION. Soit $g(p_1, \dots, p_n) = \mathbb{E}[f(T_n)]$ définie sur $[0, 1]^n$. Remarquons que $(p_1, \dots, p_n) \geq^* (p, \dots, p)$ (cf Lemme B.2) et $g(p, \dots, p) = \mathbb{E}[f(S_n)]$. Ainsi si on prouve que g est Schur-concave, on aura prouvé le lemme. Le fait que $g(p_1, \dots, p_n) = \mathbb{E}[\mathbb{E}[f(\xi_1 + \xi_2 + \sum_{i=3}^n \xi_i \mid \xi_3, \dots, \xi_n)]]$ et la Remarque B.8 nous permet de se restreindre au cas $n = 2$.

Soit donc $(p_1, p_2), (\tilde{p}_1, \tilde{p}_2) \in [0, 1]^2$ tels que $(p_1, p_2) \geq^* (\tilde{p}_1, \tilde{p}_2)$. On a

$$\begin{aligned} g(p_1, p_2) = & p_1 p_2 (f(-p_1 - p_2) - 2f(1 - p_1 - p_2) + f(2 - p_1 - p_2)) \\ & + (p_1 + p_2) (f(1 - p_1 - p_2) - f(-p_1 - p_2)) + f(-p_1 - p_2) \end{aligned}$$

Utilisons maintenant les faits suivants :

- La convexité de f entraîne $f(-p_1 - p_2) - 2f(1 - p_1 - p_2) + f(2 - p_1 - p_2) \geq 0$
- La fonction $(p_1, p_2) \mapsto p_1 p_2$ est Schur-concave.
- Comme $(p_1, p_2) \geq^* (\tilde{p}_1, \tilde{p}_2)$, on a $p_1 + p_2 = \tilde{p}_1 + \tilde{p}_2$

Ce qui nous donne :

$$g(p_1, p_2) \leq g(\tilde{p}_1, \tilde{p}_2)$$

Et donc g est Schur-concave. \square

Lemme 1.8. Soit x_1, \dots, x_n des réels positifs et $a = \sum_{i=1}^n x_i/n$ et f_t l'une des fonctions du Lemme 1.5. Soit $T_n = \sum_{i=1}^n \theta_i$ où les θ_i sont des v.a. indépendantes telles que $\theta_i \sim \mathcal{B}(-x_i, 1)$. Soit $S_n = \sum_{i=1}^n \varepsilon_i$ où les ε_i sont des v.a. iid de loi $\mathcal{B}(-a, 1)$. Alors pour tout $t \in \mathbb{R}$, on a

$$\mathbb{E}[f_t(T_n)] \leq \mathbb{E}[f_t(S_n)]$$

DÉMONSTRATION. Soit

$$\begin{aligned} q_k &= \mathbb{P}(\theta_k = -x_k) = \frac{1}{1+x_k} \\ p_k &= \mathbb{P}(\theta_k = 1) = \frac{x_k}{1+x_k} \end{aligned}$$

Soit $g_t(x_1, \dots, x_n) = \mathbb{E}[f_t(T_n)]$. Comme dans le lemme précédent, $(x_1, \dots, x_n) \geq^* (a, \dots, a)$ et $g(a, \dots, a) = \mathbb{E}[f_t(S_n)]$. On montre que g_t est une fonction Schur-concave. On peut prendre \mathcal{C} un cube ouvert assez grand tel que pour a donné, \mathcal{C} contient l'ensemble $\{x \in \mathbb{R}^n \mid \sum_{i=1}^n x_i = an, x_1, \dots, x_n \geq 0\}$. Comme le cube est ouvert, x_k peut prendre des valeurs négatives, on suppose que $x_k \geq -1/3$. Donc p_k peut-être négatif et on comprend alors $\mathbb{E}[h(\theta_k)]$ comme la valeur $h(-x_k)q_k + h(1)p_k$.

On peut se restreindre comme précédemment au cas $n = 2$.

Notons alors $x_1 = \alpha$ et $x_2 = \beta$. On va montrer en supposant que $\beta > \alpha$, $\partial_\alpha g_t - \partial_\beta g_t \stackrel{\text{def}}{=} I(t) \geq 0$. On a par indépendance,

$$g_t(x) = f_t(-\alpha - \beta)q_1q_2 + f_t(1 - \beta)p_1q_2 + f_t(1 - \alpha)q_1p_2 + f_t(2)p_1p_2$$

On étudie 5 cas :

- i. $t \leq -\alpha - \beta$
 - ii. $-\alpha - \beta \leq t \leq 1 - \beta$
 - iii. $1 - \beta \leq t \leq 1 - \alpha$
 - iv. $1 - \alpha \leq t \leq 2$
 - v. $t \geq 2$
- i. $f_t(x) = (x - t)^2$ sur le support de $\theta_1 + \theta_2$. Donc $g_t(x) = \alpha + \beta + t^2$, ce qui implique $I(t) = 0$.
 - ii. On écrit

$$g_t(x) = (1 - \beta - t)^2 p_1 q_2 + (1 - \alpha - t)^2 q_1 p_2 + (2 - t)^2 p_1 p_2 \quad (1.11)$$

$$\begin{aligned} &= (1 - \beta - t)^2 p_1 q_2 + (1 - \alpha - t)^2 q_1 p_2 + (2 - t)^2 p_1 p_2 \\ &\quad + (\alpha + \beta + t)^2 q_1 - q_2 - (\alpha + \beta + t)^2 q_1 - q_2 \end{aligned} \quad (1.12)$$

$$= \alpha + \beta + t^2 - (\alpha + \beta + t)^2 q_1 - q_2 \quad (1.13)$$

où l'on a utilisé le résultat du i. à la 3^{ème} ligne. Maintenant, il est facile de montrer que $I(t) = (\alpha + \beta + t)^2 q_1^2 q_2^2 (\beta - \alpha) \geq 0$

- iii. Ici $g_t(x) = (t + \alpha - 1)^2 q_1 p_2 + (t - 2)^2 p_1 p_2$. On peut écrire $I(t) = At^2 + Bt + C$ où $A = -q_2^2 < 0$. Ainsi $t \mapsto I(t)$ est concave. Or $I(1 - \beta) \geq 0$ par ii. et $I(1 - \alpha) \geq 0$ par iv. (voir plus loin)
- iv. $g_t(x) = (t - 2)^2 p_1 p_2$. On montre alors que $I(t) = (t - 2)^2 q_1 q_2 (\beta q_1 - \alpha q_2)$. Or $\beta q_1 - \alpha q_2 \geq 0$ est équivalent $(\beta - \alpha)(1 + \beta + \alpha) \geq 0$ qui est vraie par les hypothèses sur β et α . Ainsi $I(t) \geq 0$.
- v. C'est direct ($g_t(x) = 0$). □

Lemme 1.9. Soit a_1, \dots, a_n des réels positifs et $a^2 = \sum_{i=1}^n a_i^2 / n$ et f_t la fonction définie par $f_t(x) = (x - t)_+^3$ avec $t \in \mathbb{R}$. Soit $T_n = \sum_{i=1}^n \theta_i$ où les θ_i sont des v.a. indépendantes telles que $\theta_i \sim \mathcal{B}(-a_i, a_i)$. Soit $S_n = \sum_{i=1}^n \varepsilon_i$ où les ε_i sont des v.a. iid de loi $\mathcal{B}(-a, a)$. Alors pour tout $t \in \mathbb{R}$, on a

$$\mathbb{E}[f_t(T_n)] \leq \mathbb{E}[f_t(S_n)]$$

DÉMONSTRATION. Soit la fonction g_t définie sur $\{\eta = (\eta_1, \dots, \eta_n) \mid \forall i, \eta_i > 0\} \subset \mathbb{R}^n$ par

$$g_t(\eta) = \mathbb{E}[f_t(\sqrt{\eta_1} \tilde{\varepsilon}_1 + \dots + \sqrt{\eta_n} \tilde{\varepsilon}_n)]$$

où les $\tilde{\varepsilon}_i$ sont des v.a. iid de Rademacher. On a comme précédemment $(a_1^2, \dots, a_n^2) \geq^* (a^2, \dots, a^2)$ et $g_t(a^2, \dots, a^2) = \mathbb{E}[f_t(S_n)]$. Et on ne regarde que pour 2 variables et on suppose que $\eta_2 > \eta_1$. On a donc,

$$\begin{aligned} g_t(\eta) &= f_t(-\sqrt{\eta_1} - \sqrt{\eta_2}) \frac{1}{4} + f_t(-\sqrt{\eta_1} + \sqrt{\eta_2}) \frac{1}{4} \\ &\quad + f_t(\sqrt{\eta_1} - \sqrt{\eta_2}) \frac{1}{4} + f_t(\sqrt{\eta_1} + \sqrt{\eta_2}) \frac{1}{4} \end{aligned}$$

On étudie 5 cas :

- i. $t \leq -\sqrt{\eta_1} - \sqrt{\eta_2}$
- ii. $-\sqrt{\eta_1} - \sqrt{\eta_2} \leq t \leq \sqrt{\eta_1} - \sqrt{\eta_2}$
- iii. $\sqrt{\eta_1} - \sqrt{\eta_2} \leq t \leq \sqrt{\eta_2} - \sqrt{\eta_1}$
- iv. $\sqrt{\eta_2} - \sqrt{\eta_1} \leq t \leq \sqrt{\eta_2} + \sqrt{\eta_1}$
- v. $t \geq \sqrt{\eta_2} + \sqrt{\eta_1}$

La démonstration de ce point est similaire à la démonstration du Lemme 1.8. La démonstration est finie pour tout vecteur (a_1, \dots, a_n) sans coordonnées nulles. Pour le cas général, on conclut par continuité de f_t . □

1.3 Démonstrations

On aura besoin du lemme technique suivant pour conclure dans les preuves des théorèmes. En fait on s'en sert pour majorer $\mathbb{E}[f(S_n)]$. La notation B° désigne l'enveloppe log-concave de la fonction B (voir Annexe A).

Lemme 1.10. *Soit $s > 0$ et B une fonction de survie. Soit*

$$\begin{aligned}\alpha &= \sup\{y \mid B^\circ(y) = 1\} \\ \beta &= \inf\{y \mid B^\circ(y) = 0\}\end{aligned}$$

Alors pour tout x tel que $\alpha < x \leq \beta$, on a

$$\inf_{t < x} (x - t)^{-s} \int_t^{+\infty} s(z - t)^{s-1} B(z) dz \leq e^s s^{-s} \Gamma(s + 1) B^\circ(x)$$

DÉMONSTRATION. Par définition, la fonction, $z \mapsto -\log B^\circ(z)$ est convexe. De plus, cette fonction est strictement positive et strictement croissante sur $] \alpha, \beta]$. Donc pour chaque $x \in] \alpha, \beta]$, il existe une fonction linéaire $y(z) = a(x, B) + b(x, B)z$, pour un certain $b(x, B) > 0$ tel que $y(x) = -\log B^\circ(x)$ et $-\log B^\circ(z) \geq y(z)$ pour tout $z \in \mathbb{R}$.

En particulier, $B^\circ(x) = \exp(-a(x, B) - b(x, B)x)$ et $B^\circ(z) \leq \exp(-a(x, B) - b(x, B)z)$ pour tout $z \in \mathbb{R}$.

Et comme $B \leq B^\circ$,

$$\begin{aligned}\int_t^{+\infty} s(z - t)^{s-1} B(z) dz &\leq \exp(a(x, B)) \int_t^{+\infty} s(z - t)^{s-1} \exp(-b(x, B)z) dz \\ &= \Gamma(s + 1) b(x, B)^{-s} \exp(-a(x, B) - b(x, B)t) \\ &= \Gamma(s + 1) b(x, B)^{-s} \exp(b(x, B)(x - t)) B^\circ(x)\end{aligned}$$

On choisit t tel que $b(x, B)(x - t) = s$ et on obtient l'inégalité du lemme. \square

On a maintenant tous les ingrédients pour démontrer les Théorèmes 1.1, 1.2 et 1.3. On procède comme indiqué dans l'introduction.

DÉMONSTRATION (THÉORÈME 1.1). On suppose que $-pn < x < n - pn$ (dans les autres cas, l'inégalité recherchée est triviale). Soit $f_t(z) = (z - t)_+$ avec $t \in \mathbb{R}$. On a par l'inégalité de Markov, pour $t < x$,

$$\mathbb{P}(M_n \geq x) \leq \frac{1}{(x - t)} \mathbb{E}[f(M_n)]$$

Soit une suite de v.a. indépendantes $(\theta_i)_{1 \leq i \leq n}$ telle que $\theta_i \sim \mathcal{B}(-\sigma_i, 1)$. On peut alors appliquer le Lemme 1.4 avec la fonction f du lemme définie par $f(x_1, \dots, x_n) = f_t(\sum_{i=1}^n x_i)$ qui vérifie bien les conditions voulues,

$$\mathbb{E}[f_t(M_n)] \leq \mathbb{E}[f_t(T_n)]$$

où $T_n = \sum_{i=1}^n \xi_i$ où les ξ_i sont des v.a. indépendantes telles que $\xi_i \sim \mathcal{B}(-p_i, 1-p_i)$. Or on a par le Lemme 1.7

$$\mathbb{E}[f_t(T_n)] \leq \mathbb{E}[f_t(S_n)]$$

où S_n est comme dans l'énoncé du Théorème.

Maintenant,

$$\begin{aligned} E[f_t(S_n)] &= \mathbb{E}[(S_n - t)_+] \\ &= \int_t^{+\infty} (z - t)_+ \mathbb{P}_{S_n}(dz) \\ &= \int_t^{+\infty} \left(\int_t^z du \right) \mathbb{P}_{S_n}(dz) \\ &= \int_t^{+\infty} \left(\int_t^{+\infty} \mathbb{1}_{u \leq z} du \right) \mathbb{P}_{S_n}(dz) \\ &= \int_t^{+\infty} \left(\int_t^{+\infty} \mathbb{1}_{u \leq z} \mathbb{P}_{S_n}(dz) \right) du \\ &= \int_t^{+\infty} \mathbb{P}(S_n \geq u) du \end{aligned}$$

Et donc,

$$\mathbb{P}(M_n \geq x) \leq \inf_{t < x} \frac{1}{x - t} \int_t^{+\infty} \mathbb{P}(S_n \geq u) du$$

Ensuite une application du Lemme 1.10 avec $\alpha = -pn$, $\beta = n - pn$ et $B(z) = \mathbb{P}(S_n \geq z)$ donne

$$\mathbb{P}(M_n \geq x) \leq e \mathbb{P}^\circ(S_n \geq x)$$

Il reste à prouver la 2^{ème} inégalité du Théorème.

On applique la borne 1.4 du Théorème 1.2 (la démonstration de ce théorème suit),

$$\mathbb{P}(S_n \geq z) \leq \frac{e^2}{2} \mathbb{P}^\circ \left(\eta \geq \lambda + \frac{z}{1-p} \right)$$

où $\eta \sim \mathcal{P}(\lambda)$ avec $\lambda = pn/(1-p)$. Ainsi en combinant les 2 dernières inégalités, on a

$$\mathbb{P}(M_n \geq x) \leq \frac{e^2}{2} \inf_{t < x} \frac{1}{x - t} \int_t^{+\infty} \mathbb{P}^\circ \left(\eta \geq \lambda + \frac{z}{1-p} \right) dz$$

Et une nouvelle application du Lemme 1.10 permet de conclure. \square

DÉMONSTRATION (THÉORÈME 1.2). On procède de manière similaire à la démonstration du théorème précédent.

On peut supposer $b = 1$. On se place dans le cas $-n\sigma^2 < x \leq n$ (dans les autres cas, l'inégalité recherchée est triviale). Soit $f_t(z) = (z-t)_+^2$ avec $t \in \mathbb{R}$. On a par l'inégalité de Markov, pour $t < x$,

$$\mathbb{P}(M_n \geq x) \leq \frac{1}{(x-t)^2} \mathbb{E}[f_t(M_n)]$$

On peut appliquer le Lemme 1.5 :

$$\mathbb{E}[f_t(M_n)] \leq \mathbb{E}[f_t(T_n)]$$

où $T_n = \sum_{i=1}^n \theta_i$ où les θ_i sont des v.a. indépendantes telles que $\theta_i \sim \mathcal{B}(-\sigma_i^2, 1)$. Or on a par le Lemme 1.8

$$\mathbb{E}[f_t(T_n)] \leq \mathbb{E}[f_t(S_n)]$$

où S_n est comme dans l'énoncé du Théorème.

Maintenant, comme dans la démonstration précédente,

$$E[f_t(S_n)] = 2 \int_t^{+\infty} \mathbb{P}(S_n \geq u)(u-t)du$$

Et donc,

$$\mathbb{P}(M_n \geq x) \leq \inf_{t < x} \frac{2}{(x-t)^2} \int_t^{+\infty} (u-t)\mathbb{P}(S_n \geq u)du$$

Ensuite une application du Lemme 1.10 avec $\alpha = -n\sigma^2$, $\beta = n$ et $B(z) = \mathbb{P}(S_n \geq z)$ donne

$$\mathbb{P}(M_n \geq x) \leq \frac{e^2}{2} \mathbb{P}^\circ(S_n \geq x)$$

Il reste à prouver la 2^{ème} inégalité du théorème.

Soit $K_{n+m} = \sum_{i=1}^{n+m} Y_i$ la martingale définie par $Y_k = X_k$ pour $k = 1, \dots, n$ et $Y_k = 0$ sinon. On applique la 1^{ère} inégalité à K_{n+m} .

$$\begin{aligned} \mathbb{P}(M_n \geq x) &= \mathbb{P}(K_{n+m} \geq x) \\ &\leq \frac{e^2}{2} \mathbb{P}^\circ(S_{n+m} \geq x) \end{aligned}$$

où $S_{n+m} = \sum_{i=1}^{n+m} \varepsilon_i$ où les ε_i sont des v.a. iid de loi $\mathcal{B}(-\sigma^2/b, b)$ avec $\sigma^2 = \sum_{i=1}^n \sigma_i^2 / (n+m) = b^2 \lambda / (n+m)$.

Soit $Z_{n+m} = \sum_{i=1}^{n+m} \xi_i$ où les ξ_i sont des v.a. iid de loi $\mathcal{B}(p)$ avec $p = \lambda/(n+m+\lambda)$. On a

$$\mathbb{P}(S_{n+m} \geq x) = \mathbb{P}\left(Z_{n+m} \geq \frac{\lambda + x/b}{1 + \lambda/(n+m)}\right)$$

Comme $p(n+m) \xrightarrow{m \rightarrow \infty} \lambda$, on peut appliquer le théorème limite de Poisson, d'où

$$\lim_{m \rightarrow \infty} \mathbb{P}\left(Z_{n+m} \geq \frac{\lambda + x/b}{1 + \lambda/(n+m)}\right) = \mathbb{P}\left(\eta \geq \lambda + \frac{x}{b}\right) \quad \square$$

DÉMONSTRATION (THÉORÈME 1.3). On procède de manière similaire à la démonstration du théorème précédent.

Soit $f_t(z) = (z-t)_+^3$ avec $t \in \mathbb{R}$. On a par l'inégalité de Markov, pour $t < x$,

$$\mathbb{P}(M_n \geq x) \leq \frac{1}{(x-t)^3} \mathbb{E}[f_t(M_n)]$$

On peut appliquer le Lemme 1.6 :

$$\mathbb{E}[f_t(M_n)] \leq \mathbb{E}[f_t(T_n)]$$

où $T_n = \sum_{i=1}^n \theta_i$ où les θ_i sont des v.a. indépendantes telles que $\theta_i \sim \mathcal{B}(-a_i, a_i)$. Or on a par le Lemme 1.9

$$\mathbb{E}[f_t(T_n)] \leq \mathbb{E}[f_t(S_n)]$$

où S_n est comme dans l'énoncé du théorème. Maintenant, en procédant de manière similaire que précédemment,

$$\mathbb{E}[f_t(S_n)] = 3 \int_t^{+\infty} \mathbb{P}(S_n \geq u) (u-t)^2 du$$

Et donc,

$$\mathbb{P}(M_n \geq x) \leq \inf_{t < x} \frac{3}{(x-t)^3} \int_t^{+\infty} (u-t)^2 \mathbb{P}(S_n \geq u) du$$

On peut alors conclure par le Lemme 1.10,

$$\begin{aligned} \mathbb{P}(M_n \geq x) &\leq \frac{e^3}{27} \Gamma(4) \mathbb{P}^\circ(S_n \geq x) \\ &= \frac{2e^3}{9} \mathbb{P}^\circ(S_n \geq x) \end{aligned}$$

Il reste à montrer la 2^{ème} inégalité du théorème.

Soit $K_{n+m} = \sum_{i=1}^{n+m} Y_i$ la martingale définie par $Y_k = X_k$ pour $k = 1, \dots, n$ et $Y_k = 0$ sinon. On applique la 1^{ère} inégalité à K_{n+m} .

$$\begin{aligned} \mathbb{P}(M_n \geq x\sqrt{n}) &= \mathbb{P}(K_{n+m} \geq x\sqrt{n}) \\ &\leq \frac{2e^3}{9} \mathbb{P}^\circ(S_{n+m} \geq x\sqrt{n}) \end{aligned}$$

où $S_{n+m} = \sum_{i=1}^{n+m} \varepsilon_i$ où les ε_i sont des v.a. iid de loi $\mathcal{B}(-\tilde{a}, \tilde{a})$ avec $\tilde{a}^2 = \sum_{i=1}^n a_i^2 / (n+m)$. Ainsi,

$$S_{n+m} \stackrel{\text{loi}}{=} \frac{\sqrt{n}}{\sqrt{n+m}} a \sum_{i=1}^{n+m} \tilde{\varepsilon}_i$$

Donc,

$$\mathbb{P}(M_n \geq x\sqrt{n}) \leq \frac{2e^3}{9} \mathbb{P}^\circ \left(\sum_{i=1}^{n+m} \tilde{\varepsilon}_i \geq \frac{x\sqrt{n+m}}{a} \right)$$

On passe alors à la limite quand $m \rightarrow +\infty$ et le TCL permet de conclure :

$$\mathbb{P}(M_n \geq x\sqrt{n}) \leq \frac{2e^3}{9} (1 - \phi(x/a)) \quad \square$$

On peut donner des extensions au Théorème 1.2 ainsi :

Théorème 1.11. *Supposons que les accroissements d'une martingale M_n satisfont $\mathbb{P}(X_k \leq b_k) = 1$ et $\mathbb{P}(s_k^2 \leq \sigma_k^2) = 1$ où $b_i > 0, \sigma_i^2 \geq 0$. Soit $T_n = \sum_{i=1}^n \theta_i$ où les θ_i sont des v.a. indépendantes telles que $\theta_i \sim \mathcal{B}(-\sigma_i^2/b, b)$. Alors pour tout $x \in \mathbb{R}$ et $t > 0$,*

$$\mathbb{P}(M_n \geq x) \leq \exp(-tx) \mathbb{E}[\exp(hT_n)]$$

En particulier, si pour tout i , $b_i = b$,

$$\mathbb{P}(M_n \geq x) \leq \inf_{t>0} \exp(-tx) \mathbb{E}[\exp(hS_n)]$$

où $S_n = \sum_{i=1}^n \varepsilon_i$ où les ε_i sont des v.a. iid de loi $\mathcal{B}(-a/b, b)$ avec $a = \sum_{i=1}^n \sigma_i^2/n$.

DÉMONSTRATION. Une application directe de l'inégalité de Markov et du Lemme 1.5 donne la 1^{ère} inégalité. La 2^{ème} inégalité a été prouvée par Hoeffding [9]. \square

Théorème 1.12. *Supposons que les accroissements d'une martingale M_n satisfont $\mathbb{P}(X_k \leq b_k) = 1$ et $\mathbb{P}(s_k^2 \leq \sigma_k^2) = 1$ où $b_i > 0, \sigma_i^2 \geq 0$. Soit $f_t(x) = (x - t)_+^s$ avec $s \geq 2$. Soit $T_n = \sum_{i=1}^n \theta_i$ où les θ_i sont des v.a. indépendantes telles que $\theta_i \sim \mathcal{B}(-\sigma_i^2/b, b)$. Alors pour tout $t < x$ et tout $x \in \mathbb{R}$, on a*

$$\mathbb{P}(M_n \geq x) \leq \frac{1}{(x - t)^s} \mathbb{E}[f_t(T_n)]$$

En particulier, dans le cas où pour tout i , $b_i = b$, une application du Lemme 1.8 donne

$$\mathbb{P}(M_n \geq x) \leq \frac{1}{(x - t)^s} \mathbb{E}[f_t(S_n)]$$

où $S_n = \sum_{i=1}^n \varepsilon_i$ où les ε_i sont des v.a. iid de loi $\mathcal{B}(-a, b)$ avec $a = \sum_{i=1}^n \sigma_i^2/n$. Et

$$\mathbb{P}(M_n \geq x) \leq e^s s^{-s} \Gamma(s + 1) \mathbb{P}^\circ(T_n \geq x)$$

En particulier, dans le cas où pour tout i , $b_i = b$,

$$\mathbb{P}(M_n \geq x) \leq e^s s^{-s} \Gamma(s + 1) \mathbb{P}^\circ(S_n \geq x)$$

DÉMONSTRATION. La 1^{ère} inégalité découle directement de l'inégalité de Markov et du Lemme 1.5. Puis le Lemme 1.10 implique les 2 dernières inégalités. \square

1.4 Commentaires

- Pour obtenir des bornes pour $\mathbb{P}(M_n \leq x)$, il suffit d'appliquer les théorèmes à la martingale $-M_n$.
- **Constantes apparaissant dans les Théorèmes**

La constante apparaissant dans les Théorèmes 1.1, 1.2 et 1.3, n'est peut-être pas la constante optimale. Cependant on montre que cette constante est nécessairement plus grande que 2. En fait on montre

Lemme 1.13.

$$\sup_n \sup_{(M_k)_{k \in \mathbb{N}} \in \mathcal{M}} \mathbb{P}(M_n \geq x\sqrt{n}) \geq 2(1 - \phi(x))$$

où \mathcal{M} est une classe de martingales dont les accroissements sont des v.a. de Bernoulli centrées et bornées par 1.

DÉMONSTRATION. Soit $n \in \mathbb{N}$. Montrons que

$$\sup_{(M_k)_{k \in \mathbb{N}} \in \mathcal{M}} \mathbb{P}(M_n \geq y) = \sup_{(M_k)_{k \in \mathbb{N}} \in \mathcal{M}} \mathbb{P}\left(\max_{0 \leq k \leq n} M_k \geq y\right)$$

En effet, on a l'inégalité \leq car $M_n \leq \max_{0 \leq k \leq n} M_k$. Pour l'autre sens, en définissant $\tau_y = \inf\{n \geq 0 \mid M_n \geq y\}$, on a

$$\begin{aligned} \mathbb{P}\left(\max_{0 \leq k \leq n} M_k \geq y\right) &= \mathbb{P}\left(M_{n \wedge \tau_y} \geq y\right) \\ &\leq \sup_{(M_k)_{k \in \mathbb{N}} \in \mathcal{M}} \mathbb{P}(M_n \geq y) \end{aligned}$$

où dans la dernière inégalité on a utilisé le fait que la martingale $(M_{k \wedge \tau_y})_{k \in \mathbb{N}}$ est une martingale de \mathcal{M} .

On définit maintenant,

$$D_n(y) = \sup_{(S_k) \in \mathcal{M}} \mathbb{P}((S_k)_{k \leq n} \text{ visite l'intervalle } [y, \infty])$$

où S_n est une marche aléatoire simple basée sur une martingale (M_k) .

Par ce qu'on a fait avant, $D_n(y) = \sup_{(M_k) \in \mathcal{M}} \mathbb{P}(M_n \geq y)$.

De plus, comme $y \mapsto D_n(y)$ est décroissante, pour tout entier $k \geq y$, on a $D_n(y) \geq D_n(k)$.

Rappelons l'inégalité de Levy :

Théorème 1.14 (Inégalité de Levy). *Soit $(X_j)_{j \leq n}$ des v.a. indépendantes symétriques et $S_n = \sum_{j=1}^n X_j$. On a pour tout entier k ,*

$$\mathbb{P}\left(\max_{j \leq n} S_j \geq k\right) = 2\mathbb{P}(S_n \geq k) - \mathbb{P}(S_n = k)$$

Si on choisit un $k \geq y$ tel que $n - k$ soit impair, l'inégalité de Levy donne $D_n(k) = 2\mathbb{P}(S_n \geq k)$.

On pose maintenant $y = x\sqrt{n}$ et on considère l'entier $k_{x\sqrt{n}}$ qui est l'entier minimal vérifiant $k_{x\sqrt{n}} \geq x\sqrt{n}$ et $n - k_{x\sqrt{n}}$ est impair. Donc,

$$x\sqrt{n} \leq k_{x\sqrt{n}} \leq x\sqrt{n} + 2$$

On a donc,

$$\begin{aligned} D_n(x\sqrt{n}) &\geq 2\mathbb{P}(S_n \geq k_{x\sqrt{n}}) \\ &\geq 2\mathbb{P}\left(\frac{S_n}{\sqrt{n}} \geq x + \frac{2}{\sqrt{n}}\right) \\ &\stackrel{\text{def}}{=} 2\left(1 - F_n\left(x + \frac{x}{\sqrt{n}}\right)\right) \end{aligned}$$

Or par le TCL, $F_n(x + 2/\sqrt{n})$ tend vers $\phi(x)$ quand n tend vers l'infini. \square

- Dans les Théorèmes 1.1, 1.2 et 1.3, on ne peut pas remplacer $\mathbb{P}^\circ(S_n \geq x)$ par $\mathbb{P}(S_n \geq x)$.

En effet, on se place dans le cas $n = 1$. Soit X une v.a telle que $\mathbb{P}(X \leq 1) = 1$, $\mathbb{E}[X] = 0$ et $\mathbb{E}[X^2] \leq \sigma^2$. Soit $\varepsilon \sim \mathcal{B}(-\sigma^2, 1)$. Il suffit de montrer que

$$\sup_{\mathcal{L}(X)} \frac{\mathbb{P}(X \geq 0)}{\mathbb{P}(\varepsilon \geq 0)} = \infty$$

où $\mathcal{L}(X)$ désigne les v.a. de même loi que X .

On prend $X = 0$ alors $\mathbb{P}(X \geq 0) = 1$. Et comme $\mathbb{P}(\varepsilon \geq 0) = \sigma^2/(1+\sigma^2)$, la conclusion suit.

2 Large Deviations

Cette partie repose sur l'étude de l'article de L. Wu [13].

On cherche à obtenir des principes de grandes déviations en norme uniforme pour les processus empiriques.

Soit $(X_n)_{n \geq 1}$ une suite de v.a. iid définie sur un espace de probabilité $(\Omega, \mathcal{B}, \mathbb{P})$ à valeurs dans un espace mesurable (E, \mathcal{E}) de loi μ . On définit la mesure empirique :

$$L_n = \frac{1}{n} \sum_{i=1}^n \delta_{X_k}$$

qui est un élément (aléatoire) de $M_1(E)$.

Soit $\mathcal{F} \subset b\mathcal{E}$. Soit $(\varepsilon_n)_{n \geq 1}$ une suite de v.a. iid définie sur $(\Omega, \mathcal{B}, \mathbb{P})$ de loi $\mathcal{B}(1/2)$ et indépendante de $(X_n)_{n \geq 1}$. On note

$$H(n, \eta) = \mathbb{E} \left[\left\| \sum_{k=1}^n \varepsilon_k f(X_k) \right\|_{\mathcal{F}_\eta} \right]$$

où

$$\mathcal{F}_\eta = \{f - g \mid f, g \in \mathcal{F} \text{ et } d_2(f, g) \leq \eta\}$$

On aura besoin du lemme suivant de Talagrand [12]

Lemme 2.1 (Talagrand). *Soit $\mathcal{F} \subset b\mathcal{E}$ dénombrable et tel que pour tout $f \in \mathcal{F}$, $0 \leq f \leq 1$. Soit $\sigma \stackrel{\text{def}}{=} \sup_{f \in \mathcal{F}} (f(f - \mu(f))^2 d\mu)^{1/2}$ et $S \stackrel{\text{def}}{=} n\sigma^2 + H(n)$ où*

$$H(n) = \mathbb{E} \left[\left\| \sum_{k=1}^n \varepsilon_k f(X_k) \right\|_{\mathcal{F}} \right]$$

Alors pour une constante universelle $L > 0$, on a pour tout $t \geq LH(n)$,

$$\mathbb{P} \left(\left\| \sum_{k=1}^n f(X_k) - n\mu(f) \right\|_{\mathcal{F}} \geq t \right) \leq \exp(-\phi_{L,S}(t))$$

où la fonction $\phi_{L,S}(t)$ pour $t \geq 0$ est donnée par

$$\phi_{L,S}(t) = \frac{t^2}{L^2 S}, \quad \text{si } t \leq LS \tag{2.14}$$

$$\phi_{L,S}(t) = \frac{t}{L} \left(\log \left(\frac{et}{LS} \right) \right)^{1/2}, \quad \text{si } t \geq LS \tag{2.15}$$

2.1 Théorème

Théorème 2.2. *Soit $\mathcal{F} \subset b\mathcal{E}$ dénombrable et tel que pour tout $f \in \mathcal{F}$, $0 \leq f \leq 1$. Alors les propositions suivantes sont équivalentes,*

- i. $\mathbb{P}\left((L_n - \mu)^{\mathcal{F}} \in \cdot\right)$ satisfait un PGD sur $l_\infty(\mathcal{F})$ de vitesse n et de bonne fonction de taux donnée par*

$$h_{\mathcal{F}}(F) \stackrel{\text{def}}{=} \inf\{H(\nu | \mu) \mid \nu \in M_1(E) \text{ et } \nu^{\mathcal{F}} = F \text{ sur } \mathcal{F}\}$$

- ii. (\mathcal{F}, d_2) est précompact et*

$$\lim_{\eta \rightarrow 0} \limsup_{n \rightarrow \infty} \frac{H(n, \eta)}{n} = 0$$

- iii. (\mathcal{F}, d_2) est précompact et $(L_n - \mu)^{\mathcal{F}}$ tend vers 0 en probabilité dans $l_\infty(\mathcal{F})$*

DÉMONSTRATION. On s'intéresse dans ce mémoire à la preuve de *ii.* \Rightarrow *i.* On admettra ici que $h_{\mathcal{F}}$ est une bonne fonction de taux (c.f. Menneteau [11] pour une preuve).

Considérons tout d'abord le cas où \mathcal{F} est fini. C'est une conséquence du Théorème de Sanov (C.2) et du principe de contraction (C.3) puisque si \mathcal{F} est fini, $l_\infty(\mathcal{F})$ est séparable. On se place maintenant dans le cas où \mathcal{F} est infini. Montrons que l'on a pour tout $\delta > 0$,

$$\lim_{\eta \rightarrow 0} \limsup_{n \rightarrow \infty} \frac{1}{n} \log \mathbb{P}\left(\|L_n(\cdot) - \mu(\cdot)\|_{\mathcal{F}_\eta} \geq \delta\right) = -\infty \quad (2.16)$$

Soit $\mathcal{F}'_\eta = \{(f+1)/2, f \in \mathcal{F}_\eta\}$. On remarque que pour $g \in \mathcal{F}'_\eta$, on a $0 \leq g \leq 1$. En effet, pour $f \in \mathcal{F}_\eta$, on peut écrire, $f = h - \tilde{h}$ avec $h, \tilde{h} \in \mathcal{F}$ et avec l'hypothèse sur \mathcal{F} on a donc $-1 \leq f \leq 1$. On définit

$$H'(n, \eta) = \mathbb{E} \left[\left\| \sum_{k=1}^n \varepsilon_k f(X_k) \right\|_{\mathcal{F}'_\eta} \right]$$

On a

$$\begin{aligned}
|2H'(n, \eta) - H(n, \eta)| &= \left| \mathbb{E} \left[2 \left\| \sum_{k=1}^n \varepsilon_k \frac{f(X_k) + 1}{2} \right\|_{\mathcal{F}_\eta} - \left\| \sum_{k=1}^n \varepsilon_k f(X_k) \right\|_{\mathcal{F}_\eta} \right] \right| \\
&\leq \mathbb{E} \left[\left\| 2 \sum_{k=1}^n \varepsilon_k \frac{f(X_k) + 1}{2} - \sum_{k=1}^n \varepsilon_k f(X_k) \right\|_{\mathcal{F}_\eta} \right] \\
&= \mathbb{E} \left[\left| \sum_{k=1}^n \varepsilon_k \right| \right] \\
&= \mathbb{E} \left[\sqrt{\left(\sum_{k=1}^n \varepsilon_k \right)^2} \right] \\
&\leq \sqrt{n}
\end{aligned}$$

où l'on a utilisé l'inégalité de Jensen dans la dernière inégalité.

Ainsi on a

$$\frac{2}{n} H'(n, \eta) \leq \frac{1}{\sqrt{n}} + \frac{H(n, \eta)}{n} \quad (2.17)$$

Soit $\delta > 0$ et L la constante universelle dans le Lemme 2.1 de Talagrand. Comme on a supposé ii., on a pour η assez petit et n assez grand,

$$LH'(n, \eta) \leq \frac{n\delta}{2} \quad (2.18)$$

On remarque que,

$$\begin{aligned}
\mathbb{P} \left(\|L_n(\cdot) - \mu(\cdot)\|_{\mathcal{F}_\eta} \geq \delta \right) &= \mathbb{P} \left(\sup_{f \in \mathcal{F}_\eta} |L_n(f) - \mu(f)| \geq \delta \right) \\
&= \mathbb{P} \left(\sup_{f \in \mathcal{F}_\eta} \left| L_n \left(\frac{f+1}{2} \right) - \mu \left(\frac{f+1}{2} \right) \right| \geq \frac{\delta}{2} \right) \\
&= \mathbb{P} \left(\sup_{f \in \mathcal{F}'_\eta} |L_n(f) - \mu(f)| \geq \frac{\delta}{2} \right) \\
&= \mathbb{P} \left(\|L_n(\cdot) - \mu(\cdot)\|_{\mathcal{F}'_\eta} \geq \frac{\delta}{2} \right)
\end{aligned}$$

On applique alors le Lemme 2.1 de Talagrand à \mathcal{F}'_η : Pour tout $\delta > 0$,

$$\mathbb{P} \left(\|L_n(\cdot) - \mu(\cdot)\| \geq \frac{\delta}{2} \right) \leq \exp(-\phi_{L,S}(n\delta/2))$$

où $S = n\eta^2/4 + H'(n, \eta)$. En utilisant l'inégalité 2.17, on montre que

$$\frac{n}{2S} \geq \frac{1}{\eta^2/2 + H(n, \eta)/n + 1/\sqrt{n}}$$

Et donc $\phi_{L,S}(n\delta/2)$ est donnée par 2.15, puis

$$\lim_{\eta \rightarrow 0} \limsup_{n \rightarrow \infty} \frac{1}{n} \log \mathbb{P} \left(\|L_n(\cdot) - \mu(\cdot)\|_{\mathcal{F}_\eta} \geq \delta \right) = -\infty$$

Il nous reste maintenant à prouver que ce qu'on a fait implique le PGD recherché.

Soit C un fermé de $l_\infty(\mathcal{F})$ et soit $b = \inf_{F \in C} h_{\mathcal{F}}(F)$. On doit montrer que

$$\limsup_{n \rightarrow \infty} \frac{1}{n} \log \mathbb{P} \left((L_n - \mu)^{\mathcal{F}} \in C \right) \leq -b$$

Si $b = 0$ il n'y a rien à montrer. On suppose donc $b > 0$. Soit $0 < a < b$.

Commençons par quelques notations :

Soit $K \stackrel{\text{def}}{=} \{h_{\mathcal{F}} \leq a\}$. Soit $\eta > 0$ et on note \mathcal{F}^η un η -ensemble fini de \mathcal{F} : i.e. $\mathcal{F}^\eta \subset \mathcal{F}$ et pour tout $f \in \mathcal{F}$, il existe $g \in \mathcal{F}^\eta$ telle que $d_2(f, g) < \eta$. Et on définit h^η la fonction de taux correspondant à \mathcal{F}^η :

$$h^\eta(F) = \inf \{H(\nu | \mu) \mid \nu \in M_1(E) \text{ et } \nu^{\mathcal{F}} = F \text{ sur } \mathcal{F}^\eta\}$$

On définit $p_\eta(F)$ comme la restriction de F à \mathcal{F}^η . Et on note $K^\eta \stackrel{\text{def}}{=} \{h^\eta \leq a\}$. On remarque que $K^\eta = p_\eta(K)$.

Montrons que pour tout $\delta > 0$, il existe $\eta > 0$ tel que pour tout $F \in l_\infty(\mathcal{F})$, les conditions

$$p_\eta(F) \in K^\eta(\delta) \tag{2.19}$$

et

$$\sup\{|F(f) - F(g)|, f, g \in \mathcal{F}, d_2(f, g) < \eta\} < \delta \tag{2.20}$$

implique que $F \in K(3\delta)$, où l'on a noté $K(\varepsilon)$ un ε -voisinage de K dans $l_\infty(\mathcal{F})$ et $K^\eta(\varepsilon)$ un ε -voisinage de K^η dans $l_\infty(\mathcal{F}^\eta)$.

En effet, soit $f \in \mathcal{F}$ et $\tilde{F} \in K$. On considère $g \in \mathcal{F}^\eta$ telle que $d_2(f, g) < \eta$. On a

$$\begin{aligned} |F(f) - \tilde{F}(f)| &\leq |F(f) - F(g)| + |F(g) - \tilde{F}(g)| + |\tilde{F}(g) - \tilde{F}(f)| \\ &= \underbrace{|F(f) - F(g)|}_{=(1)} + \underbrace{|p_\eta(F)(g) - p_\eta(\tilde{F})(g)|}_{=(2)} + \underbrace{|\tilde{F}(g) - \tilde{F}(f)|}_{=(3)} \end{aligned}$$

- (1) $< \delta$ par la condition (2.20).
- La condition (2.19) implique qu'il existe $H \in K$ telle que $\sup_{f \in \mathcal{F}^\eta} |p_\eta(F)(f) - p_\eta(H)(f)| < \delta$. En prenant alors $\tilde{F} = H$ on a (2) $< \delta$.
- (3) $< \delta$ par la condition (2.20).

Et on a le résultat voulu.

Comme K est compact, il existe $\delta > 0$ telle que $C \cap K(3\delta) = \emptyset$ (puisque $b > a$). Ainsi pour η assez petit,

$$\begin{aligned} \mathbb{P}\left((L_n - \mu)^\mathcal{F} \in C\right) &\leq \mathbb{P}\left((L_n - \mu)^\mathcal{F} \notin K(3\delta)\right) \\ &\leq \mathbb{P}\left((L_n - \mu)^{\mathcal{F}^\eta} \notin K^\eta(\delta)\right) + \mathbb{P}\left(\|L_n - \mu\|_{\mathcal{F}^\eta} \geq \delta\right) \end{aligned}$$

Maintenant, l'égalité (2.16) nous permet d'écrire,

$$\mathbb{P}\left(\|L_n - \mu\|_{\mathcal{F}^\eta} \geq \delta\right) \leq \exp(-L(\eta).n)$$

où $L(\eta)$ tend vers l'infini quand η tend vers 0.

De plus, comme on a le résultat dans le cas fini, et par définition de K^η on a

$$\limsup_{n \rightarrow \infty} \frac{1}{n} \log \mathbb{P}\left((L_n - \mu)^{\mathcal{F}^\eta} \notin K^\eta(\delta)\right) \leq -a$$

Et finalement,

$$\mathbb{P}\left((L_n - \mu)^\mathcal{F} \in C\right) \leq \exp(-na + o(n)) + \exp(-L(\eta).n)$$

Comme $a < b$ est arbitraire, on a la borne sup voulue :

$$\limsup_{n \rightarrow \infty} \frac{1}{n} \log \mathbb{P}\left((L_n - \mu)^\mathcal{F} \in C\right) \leq -b = -\inf_{F \in C} h_\mathcal{F}(F) \quad \square$$

2.2 Complément

On voudrait maintenant à partir du PGD pour $(L_n - \mu)^\mathcal{F}$ dans $l_\infty(\mathcal{F})$ obtenir un PGD pour $\|(L_n - \mu)^\mathcal{F}\|_\mathcal{F}$ dans \mathbb{R} .

Comme déjà noté précédemment, dans le cas où \mathcal{F} est fini, $l_\infty(\mathcal{F})$ est séparable et on peut alors appliquer le principe de contraction C.3 : Ainsi, $\mathbb{P}\left(\|(L_n - \mu)^\mathcal{F}\|_\mathcal{F} \in \cdot\right)$ satisfait un PGD dans \mathbb{R} de vitesse n et de fonction de taux donnée par, pour $y \in [0, 1]$,

$$J(y) \stackrel{\text{def}}{=} \inf\{H(\nu | \mu) \mid \nu \in M_1(E), \|\nu^\mathcal{F}\|_\mathcal{F} = y\}$$

Identifions la fonction de taux :

Conjecture 2.3. On a pour tout $y \in \mathbb{R}_+$,

$$J(y) \geq \inf_{f \in \mathcal{F}} l_{\mathcal{F}}^*(y)$$

où

$$l_f^*(y) \stackrel{\text{def}}{=} \sup_{t > 0} \left\{ ty - \log \left(\int e^{tf} d\mu \right) \right\} \quad (2.21)$$

DÉMONSTRATION. On suppose $\nu \ll \mu$ et on note $d \stackrel{\text{def}}{=} d\nu/d\mu$. On a pour $t > 0$,

$$t \int df d\mu = \int d \underbrace{\left(tf - \log \left(\int e^{tf} d\mu \right) \right)}_{\stackrel{\text{def}}{=} g} d\mu + \log \left(\int e^{tf} d\mu \right)$$

Or par une application de l'inégalité de Fenchel-Young C.4,

$$\int dg d\mu \leq \int e^g d\mu + \int (d \log(d) - d) d\mu$$

Et comme $\int e^g d\mu = 1$, on a

$$\int dg d\mu \leq H(\nu | \mu)$$

Finalement en prenant \tilde{f} qui vérifie $y = \sup_{f \in \mathcal{F}} \nu(f) = \nu(\tilde{f})$ (\mathcal{F} est supposé fini), on a

$$H(\nu | \mu) \geq ty - \log \left(\int e^{t\tilde{f}} d\mu \right)$$

Comme cette inégalité est valable pour tout $t > 0$,

$$H(\nu | \mu) \geq \sup_{t > 0} \left\{ ty - \log \left(\int e^{t\tilde{f}} d\mu \right) \right\} = l_{\tilde{f}}^*(y)$$

Et donc,

$$J(y) \geq l_{\tilde{f}}^*(y) \geq \inf_{f \in \mathcal{F}} l_{\mathcal{F}}^*(y)$$

Dans ce mémoire, je n'ai pas abouti à l'inégalité contraire. Je donne seulement une piste possible pour la démonstration.

On définit pour $f \in \mathcal{F}$, $t > 0$,

$$\nu_{t,f} = \left(\int e^{tf} d\mu \right)^{-1} e^{tf} \mu$$

Un rapide calcul nous montre que

$$H(\nu_{t,f} | \mu) = t\nu_{t,f}(f) - \log \left(\int e^{tf} d\mu \right)$$

De plus, le sup dans (2.21) est atteint au t qui vérifie $y = \nu_{t,f}(f)$.
 Notons alors $\mathcal{F} = \{f_1, \dots, f_r\}$ et considérons $t_1, \dots, t_r > 0$ tels que pour tout i , $\nu_{t_i, f_i}(f_i) = y$. Ce qui implique

$$l_{f_i}^* = t_i y - \log \left(\int e^{t_i f_i} d\mu \right)$$

On se pose maintenant la question si il existe un i tel que $\sup_{f \in \mathcal{F}} \nu_{t_i, f_i}(f) = \nu_{t_i, f_i}(f_i)$.
 En effet, dans ce cas, on pourrait conclure, puisqu'on aurait

$$J(y) \leq H(\nu_{t_i, f_i} \mid \mu) = l_{f_i}^*(y) \quad \square$$

A Enveloppe log-concave

On explique ici la notation \mathbb{P}° et on en donne quelques propriétés.
On rappelle la définition suivante :

Définition A.1. Une fonction $f : A \subset \mathbb{R} \rightarrow \mathbb{R}_+$ est log-concave si la fonction $x \in A \mapsto -\log f(x)$ est convexe.

On définit maintenant l'enveloppe log-concave d'une fonction :

Définition A.2. Soit $f : \mathbb{R} \rightarrow \mathbb{R}_+$. Son enveloppe log-concave est la fonction $f^\circ : \mathbb{R} \rightarrow \mathbb{R}_+$ définie comme la fonction minimale log-concave telle que $f \leq f^\circ$.

On rappelle que la fonction de survie d'une v.a. X est la fonction $x \mapsto \mathbb{P}(X \geq x)$. On l'a dit discrète lorsque X est une v.a. discrète.
Soit B une fonction de survie d'une v.a., on définit alors $\mathbb{P}^\circ(S_n \geq x) = B^\circ(x)$.

Lemme A.3. *i. Soit $n \mapsto p_n$ et $n \mapsto q_n$, 2 fonctions log-concave telles que pour tout n , $p_n \geq 0$, $q_n \geq 0$. Alors,*

$$n \mapsto (p * q)_n \stackrel{\text{def}}{=} \sum_{k=-\infty}^{+\infty} p_{n-k} q_k$$

est une fonction log-concave.

ii. Soit $n \mapsto p_n$ une fonction log-concave telle que pour tout n , $p_n \geq 0$. Alors

$$n \mapsto t_n \stackrel{\text{def}}{=} \sum_{k \geq n} p_k$$

est une fonction log-concave.

iii. Les v.a. de Bernoulli ont une fonction de masse log-concave. Les fonctions de survie Binomiale sont log-concave.

iv. Soit (B_k) une suite de fonctions de survie log-concave qui ont une fonction de masse portée par \mathbb{Z} . Alors la limite (simple) $\lim_{k \rightarrow \infty} B_k$ est une fonction log-concave.

v. La fonction de survie d'une v.a. de Poisson est log-concave

DÉMONSTRATION. i. Soit $\delta = (p * q)_n^2 - (p * q)_{n-1}(p * q)_{n+1}$. On montre que $\delta \geq 0$.

Un calcul simple nous permet de voir que

$$2\delta = \sum_{k,r=-\infty}^{+\infty} \alpha_{k,r} \beta_{k,r}^{(n)}$$

où $\alpha_{k,r} = p_k p_r - p_{k+1} p_{r-1}$ et $\beta_{k,r}^{(n)} = q_{n-k} q_{n-r} - q_{n-k-1} q_{n-r+1}$. Maintenant, par hypothèse sur p_n et q_n , si $k \geq r$ on a $\alpha_{k,r} \geq 0$ et $\beta_{k,r}^{(n)} \geq 0$ et si $k < r$, $\alpha_{k,r} \leq 0$ et $\beta_{k,r}^{(n)} \leq 0$. Ce qui conclut.

- ii. On voit que $t_n = (p * q)_n$ où $q_n = \mathbb{1}_{n \leq 0}$ qui est une fonction log-concave. La conclusion s'en suit avec le i.
- iii. La propriété est claire pour la fonction de masse d'une v.a. de Bernoulli puisqu'elle ne prend que 2 valeurs et donc pour tout k , $p_{k-1} p_{k+1} = 0$. Ensuite, on voit une v.a. Binomiale comme une somme de v.a. de Bernoulli iid. Le i. donne alors la propriété recherchée. Soit $X \sim \text{Bin}(k, p)$, $n \mapsto \mathbb{P}(X \geq n) = \sum_{j \geq n} p_j$ où $j \mapsto p_j$ est la fonction de masse d'une v.a. Binomiale. Et donc par le point ii., la fonction de survie d'une v.a. Binomiale est log-concave.
- iv. C'est direct.
- v. La fonction de survie d'une v.a. de Poisson est une limite d'une suite de fonction de survie de v.a. Binomiale. On conclut en appliquant iii. et iv. \square

B Schur-concavité

Définition B.1. Soit $x, y \in \mathbb{R}^n$. On dit que x majorize y , noté $x \geq^* y$, si pour tout $k \in \{1 \dots n\}$,

$$\sum_{i=k}^n x_{(i)} \geq \sum_{i=k}^n y_{(i)}$$

et si

$$\sum_{i=1}^n x_{(i)} = \sum_{i=1}^n y_{(i)}$$

où $x_{(n)} \geq \dots \geq x_{(1)}$.

Lemme B.2. Soit $x \in \mathbb{R}^n$, $a_x = \sum_{i=1}^n x_i/n$ et $y(x) = (a_x, \dots, a_x) \in \mathbb{R}^n$. On a alors, $x \geq^* y(x)$.

DÉMONSTRATION. Soit $k \in \{1, \dots, n\}$. On a

$$\begin{aligned}
ka_x &= \frac{1}{n} \sum_{i=1}^n x_{(i)} + \dots + \frac{1}{n} \sum_{i=1}^n x_{(i)} \\
&= \frac{x_{(n)} + \dots + x_{(n)} + x_{(n-k)} + \dots + x_{(1)}}{n} \\
&\quad + \dots + \frac{x_{(n-k+1)} + \dots + x_{(n-k+1)} + x_{(n-k)} + \dots + x_{(1)}}{n} \quad \square \\
&\leq \frac{nx_{(n)}}{n} + \dots + \frac{nx_{(n-k+1)}}{n} \\
&\leq \sum_{i=n-k+1}^n x_{(i)}
\end{aligned}$$

Définition B.3. Une fonction g à valeurs réelles définie sur un ouvert $C \subset \mathbb{R}^n$ est dite Schur-concave sur C si $x \geq^* y$ implique $g(x) \leq g(y)$ (et Schur-convexe si $x \leq^* y$ implique $g(x) \leq g(y)$).

Théorème B.4 (Schur-Ostrowski). Soit C un ouvert convexe symétrique de \mathbb{R}^n d'intérieur non vide. On suppose que $g : C \mapsto \mathbb{R}$ est une fonction symétrique de ses arguments et continuellement différentiable. Alors g est Schur-concave si et seulement si pour tout $i \neq j$,

$$(x_i - x_j) (\partial_j g(x) - \partial_i g(x)) \geq 0$$

Remarque B.5. Par la symétrie de g , on peut remplacer la condition du Théorème B.4 par :

$$(x_1 - x_2) (\partial_2 g(x) - \partial_1 g(x)) \geq 0$$

Ce théorème découle des 2 lemmes suivants.

Lemme B.6. Soit $\mathcal{D} = \{x \in \mathbb{R}^n \mid x_1 \geq \dots \geq x_n\}$. Soit ϕ une fonction réelle définie sur \mathcal{D} . Alors on a équivalence entre :

- i. $x \leq^* y$ sur $\mathcal{D} \Rightarrow \phi(x) \leq \phi(y)$
- ii. Pour tout $z \in \mathcal{D}$ et $k \in \{1, \dots, n-1\}$,

$$\varepsilon \mapsto \phi(z_1, \dots, z_{k-1}, z_k + \varepsilon, z_{k+1} - \varepsilon, z_{k+2}, \dots, z_n)$$

est croissante sur le domaine $[0, \min\{z_{k-1} - z_k, z_{k+1} - z_{k+2}\}]$ pour $k \in \{1, \dots, n-2\}$ et sur $[0, z_{n-2} - z_{n-1}]$ pour $k = n-1$.

Lemme B.7. Soit ϕ une fonction à valeurs réelles, définie et continue sur \mathcal{D} et continuellement différentiable sur $\text{Int}(\mathcal{D})$. Alors on a équivalence entre

- i. $x \leq^* y \Rightarrow \phi(x) \leq \phi(y)$
- ii. Pour tout $z \in \text{Int}(\mathcal{D})$, $\nabla\phi(z) \in \mathcal{D}$

DÉMONSTRATION (LEMME B.6). Notons $\tilde{z}_k = \sum_{i=1}^k z_i$ pour $k \in \{1, \dots, n\}$. Sur \mathcal{D} , par définition, $x \leq^* y$ si et seulement si pour tout $k \in \{1, \dots, n-1\}$, $\tilde{x}_k \leq \tilde{y}_k$ et $\tilde{x}_n = \tilde{y}_n$.

On définit maintenant la relation \leq sur \mathbb{R}^n par $x \leq y$ si pour tout $i \in \{1, \dots, n\}$, $x_i \leq y_i$. On a alors équivalence entre :

- i. $u \leq v \Rightarrow \phi(u) \leq \phi(v)$
- ii. ϕ est croissante en chaque argument

Ces deux points donnent l'équivalence suivante :

- i. $x \leq^* y$ sur $\mathcal{D} \Rightarrow \phi(x) \leq \phi(y)$
- ii. Pour tout $k \in \{1, \dots, n-1\}$,

$$\tilde{z}_k \mapsto \phi(\tilde{z}_1, \tilde{z}_2 - \tilde{z}_1, \dots, \tilde{z}_k - \tilde{z}_{k-1}, \tilde{z}_{k+1} - \tilde{z}_k, \dots, \tilde{z}_n - \tilde{z}_{n-1}) = \phi(z)$$

est croissante sur le domaine où $z \in \mathcal{D}$.

ce qui est une réécriture du lemme. □

DÉMONSTRATION (LEMME B.7). Supposons i. Comme ϕ est continue sur $\partial\mathcal{D}$, on peut remplacer la condition du Lemme B.6 par : pour tout $z \in \text{Int}(\mathcal{D})$,

$$\varepsilon \mapsto \phi(z_1, \dots, z_{k-1}, z_k + \varepsilon, z_{k+1} - \varepsilon, z_{k+2}, \dots, z_n)$$

est croissante sur le domaine $]0, \min\{z_{k-1} - z_k, z_{k+1} - z_{k+2}\}[$ pour $k \in \{1, \dots, n-2\}$ et sur $]0, z_{n-2} - z_{n-1}[$ pour $k = n-1$. Comme de plus, ϕ est différentiable, cette condition est équivalente à

$$\frac{d}{d\varepsilon} \phi(z_1, \dots, z_{k-1}, z_k + \varepsilon, z_{k+1} - \varepsilon, z_{k+2}, \dots, z_n) \geq 0$$

sur le même domaine.

c'est-à-dire,

$$\begin{aligned} & \partial_k \phi(z_1, \dots, z_{k-1}, z_k + \varepsilon, z_{k+1} - \varepsilon, z_{k+2}, \dots, z_n) \\ & - \partial_{k+1} \phi(z_1, \dots, z_{k-1}, z_k + \varepsilon, z_{k+1} - \varepsilon, z_{k+2}, \dots, z_n) \geq 0 \end{aligned} \quad \square$$

DÉMONSTRATION (THÉORÈME B.4). Remarquons que la condition dans le théorème est équivalente à : pour tout $z \in \mathcal{D} \cap \mathcal{C}$, $\nabla\phi(z) \in \mathcal{D}$.

De plus, on a pour toute matrice de permutation Π , $x \leq^* x\Pi \leq^* x$. Ainsi, si ϕ est Schur-convexe (ou Schur-concave) sur \mathcal{C} symétrique (i.e. $x \in \mathcal{C}$ implique $x\Pi \in \mathcal{C}$ pour toute matrice de permutation Π) alors ϕ est symétrique sur \mathcal{C} (i.e. pour tout $x \in \mathcal{C}$ et pour toute matrice de permutation Π , $\phi(x) = \phi(x\Pi)$). Par conséquent, si ϕ est symétrique sur \mathcal{C} et Schur-convexe sur $\mathcal{D} \cap \mathcal{C}$, alors ϕ est Schur-convexe sur \mathcal{C} . Ce qui conclut. □

Remarque B.8. *Pour montrer qu'une fonction f est Schur-concave, on peut se restreindre au cas $n = 2$. En effet, il suffit de prouver $f(x) \leq f(y)$ pour $x \geq^* y$ lorsque y diffère de x de 2 composantes. Et comme f est symétrique, il suffit de prouver que $(z_1, z_2) \mapsto f(z_1, z_2, \dots, z_n)$ est Schur-concave.*

C Principes de Grandes Déviations

On rappelle ici quelques résultats de la théorie des Grandes Déviations que nous utilisons dans le mémoire.

Définition C.1. *Une fonction $I : E \rightarrow [0, \infty]$ est dite bonne fonction de taux si $I \not\equiv \infty$ et si pour tout $a \in \mathbb{R}$, $\{x \in E \mid I(x) \leq a\}$ est compact.*

Théorème C.2 (Sanov). *L_n satisfait un PGD sur $(M_1(E), \tau)$ de bonne fonction de taux $H(\cdot \mid \mu)$.*

Théorème C.3 (Principe de contraction). *Soit (P_n) une suite de probabilités sur X un espace Polonais satisfaisant un PGD de fonction de taux I et soit $f : X \rightarrow Y$ une fonction continue à valeurs dans un espace Polonais Y . Alors $Q_n \stackrel{\text{def}}{=} P_n \circ f^{-1}$ satisfait un PGD sur Y de fonction de taux J donnée par : pour $y \in Y$,*

$$J(y) = \inf_{\substack{x \in X \\ f(x)=y}} I(x)$$

Théorème C.4 (Inégalité de Fenchel-Young). *Soit $X \subset \mathbb{R}^n$ un ensemble convexe et $f : X \rightarrow \mathbb{R}$ une fonction convexe. Soit*

$$X^* \stackrel{\text{def}}{=} \left\{ x^* \mid \sup_{x \in X} (\langle x^*, x \rangle - f(x)) < \infty \right\}$$

Pour tout $x \in X$, pour tout $p \in X^$,*

$$\langle p, x \rangle \leq f(x) + f^*(p)$$

Références

- [1] V. Bentkus. An inequality for large deviation probabilities of sums of bounded i.i.d. random variables. *Lithuanian Mathematical Journal*, 41(2) :112–119, 2001.
- [2] V. Bentkus. An inequality for tail probabilities of martingales with bounded differences. *Lithuanian Mathematical Journal*, 42(3) :255–261, 2002.
- [3] V. Bentkus. On hoeffding’s inequalities. *The Annals of Probability*, 32(2) :1650–1673, 04 2004.
- [4] S. Boucheron, G. Lugosi, and P. Massart. *Concentration Inequalities : A Nonasymptotic Theory of Independence*. OUP Oxford, 2013.
- [5] A. De Acosta. On large deviations of empirical measures in the τ -topology. *Journal of Applied Probability*, 31 :41–47, 1994.
- [6] A. Dembo and O. Zeitouni. *Large Deviations Techniques and Applications*. Stochastic Modelling and Applied Probability. Springer, 2009.
- [7] D. Dzindzalieta. Random walks maximizing the probability to visit an interval. *ArXiv e-prints*, May 2013.
- [8] L. J. Gleser. On the distribution of the number of successes in independent trials. *The Annals of Probability*, 3(1) :182–188, 02 1975.
- [9] W. Hoeffding. Probability inequalities for sums of bounded random variables. *Journal of the American Statistical Association*, 58(301) :13–30, March 1963.
- [10] A.W. Marshall, I. Olkin, and B.C. Arnold. *Inequalities : Theory of Majorization and Its Applications : Theory of Majorization and Its Applications*. Springer Series in Statistics. Springer Science+Business Media, LLC, 2010.
- [11] L. Menneteau. *Large Deviations for Local Empirical Measures*, volume 47 of *Progress in Probability*. Birkhauser Boston, 2000.
- [12] M. Talagrand. Sharper bounds for gaussian and empirical processes. *The Annals of Probability*, 22(1) :28–76, 01 1994.
- [13] L. Wu. Large deviations, moderate deviations and lil for empirical processes. *The Annals of Probability*, 22(1) :17–27, 01 1994.