École Normale Supérieure de Rennes

RAPPORT DE STAGE

Marches Aléatoires à Incréments à Queues Lourdes

MIANNAY Matteo
Promo 2022

TUTEUR : PETRELIS Nicolas

https://perso.eleves.ens-rennes.fr/people/matteo.miannay (lien cliquable)

Mai 2023

Table des matières

Chapitre 1

Introduction

Sommaire

1.1 Motivations

Dans ce rapport, nous étudions l'article *An Introduction to Heavy-tailed and Subexponential Distributions* [1] de Foss, Korshunov et Zachary. De cet article nous allons nous intéresser à la dernière partie, sur les maximum des marches aléatoires à incréments à queue lourde avec une déviation négative. L'objet de ce rapport est donc d'extraire de l'article, long et très complet, les propositions utiles à cette partie, et de les réagencer de sorte à rendre plus progressive et accessible la preuve.

1.2 Définitions et boîte à outils

Définition 1. *Si F est une loi (que l'on nommera souvent ainsi afin de les confondre avec leur fonction de survie/ de répartition), on note $\overline{F}$ sa fonction de survie définie par $\overline{F}(x) = P(X \geq x)$.*

Définition 2 (Loi à queue lourde). *Une loi de probabilité F sur $\mathbb{R}$ est dite à queue lourde si*

$$\int_{-\infty}^{+\infty} e^{\lambda x} F(dx) = +\infty$$

pour tout $\lambda > 0$.

Les lois à queues lourdes sont donc les lois qui n'ont aucun moment exponentiel positif. Autrement, on dit que F est *à queue légère.*

Définition 3 (Fonction à queue lourde). *Une fonction $f \geq 0$ est à queue lourde si et seulement si, pour tout $\lambda > 0$,*

$$\limsup_{x \to +\infty} f(x)e^{\lambda x} = +\infty.$$

On s'en doute, il y a un lien entre lois et fonctions à queues lourdes..

Définition 4 (Fonction de risque et taux de risque). *Soit F une loi de probabilité dont le support n'est pas borné à droite, on pose $R(x) = -\ln(\overline{F}(x))$ la fonction de risque. Si cette fonction est dérivable, on note r sa dérivée et on l'appelle le taux de risque. En particulier, si F est à densité $f : r(x) = \dfrac{f(x)}{\overline{F}(x)}$.*

On remarque en particulier que $\overline{F}(x) = \exp\Big(-\displaystyle\int_{-\infty}^{x} r(y)dy\Big)$.

Propriété 1 (Taux de risque et queue lourde). *Soit F une loi à densité, r son taux de risque. Si $\lim\limits_{x \to +\infty} r(x) = 0$, alors F est à queue lourde, tandis que si $\liminf\limits_{x \to +\infty} r(x) > 0$ alors F est à queue légère.*

Démonstration. Dans le cas où $\liminf\limits_{x \to +\infty} r(x) > 0$, soit $0 < \lambda < \liminf\limits_{x \to +\infty} r(x)$, alors :

$$\int_a^b e^{\lambda x} F(dx) = \int_a^b e^{\lambda x} f(x)dx = \big[-F(x)e^{\lambda x}\big]_a^b + \frac{1}{\lambda}\int_a^b \overline{F}(x)e^{\lambda x}dx.$$

Mais alors : $e^{\lambda x}\overline{F}(x) = e^{\lambda x \int_{\infty}^{x} -r(y)dy}$

Or, $\limsup\limits_{y \to +\infty} -r(y) = -\liminf\limits_{y \to +\infty} r(y) < -\lambda$, d'où la convergence des crochets et de l'intégrale, et le résultat.

Maintenant, dans le cas où $\lim r(x) = 0$, alors si $\lambda > 0$ il existe un rang x_0 tel que pour tout $x > x_0$, $r(x) \leq \dfrac{\lambda}{2}$, d'où $-r(x) \geq -\dfrac{\lambda}{2}$ et donc $e^{\lambda x}\overline{F}(x) = e^{\lambda x \int_{\infty}^{x} -r(y)dy} \geq e^{\frac{\lambda}{2}x}$ (à une constante multiplicative non nulle près), qui tend vers l'infini et montre bien qu'on a pas de moment d'ordre λ.

□

Remarque 1. *Dans le cas où $\liminf\limits_{x \to +\infty} r(x) = 0$ mais qu'il n'y a pas convergence ; tout peut arriver.*

Théorème 1. *Soit F une loi. Alors les assertions suivantes sont équivalentes :*

1. *F est une loi à queue lourde ;*
2. *$\overline{F}$ est une fonction à queue lourde ;*
3. *La fonction de risque vérifie $\liminf\limits_{x \to +\infty} \dfrac{R(x)}{x} = 0$;*

4. *Pour tout $T > 0$, $F(x, x+T]$ est une fonction à queue lourde ;*
5. *Il existe $T > 0$, tel que $F(x, x+T)$ soit une fonction à queue lourde ;*

Ce théorème établit donc un premier lien entre fonction à queue lourde et distribution à queue lourde.

Démonstration. Nous allons montrer les implications suivantes : $1 \implies 4 \implies 5 \implies 2 \implies 3 \implies 1$, et la boucle sera bouclée.
Montrons d'abord que $1 \implies 4$:
Si $F(x, x+T]$ n'est pas à queue lourde, alors soit $c = \sup F(x, x+T)e^{\lambda' x}$ où λ' est choisi de sorte à ce que cette quantité soit finie (ceci existe car $F(x, x+T]$ n'est pas à queue lourde). Alors si $\lambda < \lambda'$, on a :

$$\int_0^{+\infty} e^{\lambda x} F(dx) \leq \sum_{n=0}^{\infty} e^{\lambda(n+1)T} F(nt, nt+T] \leq ce^{\lambda T} \sum_{n=0}^{+\infty} e^{(\lambda-\lambda')nT} < \infty.$$

On a évidemment $4 \implies 5$.

Montrons maintenant que $5 \implies 2$:
On a : $\overline{F}(x) \geq F(x, x+T)$. On trouve le résultat en multipliant par $e^{\lambda x}$ avec $\lambda > 0$, et en intégrant.

Maintenant, $2 \implies 3$:
Supposons que $\liminf \frac{R(x)}{x} > 0$. Alors, on a $\epsilon > 0, x_0 > 0$ tels que si $x \geq x_0, R(x) \geq \epsilon x$. On rappelle que $R(x) = -\ln(\overline{F}(x))$. On a donc pour $x \geq x_0$ l'inégalité $\overline{F}(x) \leq e^{-\epsilon x}$, ce qui contredit le fait que $\overline{F}$ est à queue lourde.

Enfin, $3 \implies 1$:
On suppose que F est à queue légère. Alors, il existe $\lambda > 0$ et $c > 0$ tels que $\overline{F}(x) \leq ce^{-\lambda x}$. En effet, c'est l'inégalité de Chernoff (on peut bien l'appliquer car la fonction génératrice des moments est bien définie en un λ car F est à queue légère). □

Propriété 2 (Loi sur un réseau). *Soit F une loi de probabilité sur un réseau $\{a + hn, n \in \mathbb{N}\}$ où $a \in \mathbb{R}$, $h > 0$. Alors si on note $p_n = F(\{a+hn\})$, F est à queue lourde si et seulement si $(p_n)_{n\in\mathbb{N}}$ est à queue lourde, c'est à dire :*

$$\limsup_{n\to+\infty} p_n e^{\lambda n} = +\infty, \quad \forall \lambda > 0.$$

Démonstration. On applique le théorème précédent avec $T = \frac{h}{2}$. □

Définition 5 (Loi à queue longue). *Une loi F est à queue longue si le support de F n'est pas borné à droite et que, pour tout $y > 0$,*

$$\lim_{x\to+\infty} \frac{\overline{F}(x+y)}{\overline{F}(x)} = 1$$

Définition 6 (Fonction à queue longue). *Une fonction positive en l'infini est dite à queue longue si :*

$$\forall y > 0, \quad \lim_{x\to+\infty} \frac{f(x+y)}{f(x)} = 1.$$

1.3 Exemples

Nous allons maintenant voir quelques exemples de lois à queues lourdes.

Exemple 1 (La loi de Pareto). *C'est l'exemple typique de loi à queue lourde. Sa fonction de survie est donnée par*

$$\overline{F}(x) = \left(\frac{\kappa}{x+\kappa}\right)^{\alpha}$$

On appelle $\kappa > 0$ l'échelle, et $\alpha > 0$ la forme. Les moments d'ordre strictements inférieurs à α existent, les autres sont infinis. Il est facile de les simuler en inversant la fonction de répartition.

Démonstration. Soit $\kappa > 0, \alpha > 0, \lambda > 0$. On a bien $\overline{F}(x)e^{\lambda x} = \left(\frac{\kappa}{x+\kappa}\right)^{\alpha} e^{\lambda x} \rightarrow_{x\to+\infty} = +\infty$. Donc $\overline{F}$ est à queue lourde, donc F l'est. La loi est donc bien à queue lourde. □

Exemple 2 (La loi de Burr). *Sa fonction de survie est donnée par $\overline{F}(x) = \left(\frac{\kappa}{x^{\tau}+\kappa}\right)^{\alpha}$ où les paramètres sont des réels strictement positifs. Ses moments d'ordre strictements plus petits que $\alpha\tau$ existent, les autres sont infinis.*

Démonstration. Idem qu'au dessus. □

Exemple 3 (La loi de Cauchy). *Sa fonction de densité est donnée par $f(x) = \frac{1}{\pi((x-a)^2+1}$. Les moments d'ordre strictement plus petits que 1 existent, les autres non.*

Démonstration. Soit $\lambda > 0$. $e^{\lambda x}f(x) = e^{\lambda x}\frac{1}{\pi((x-a)^2+1}$, d'où :

$$\lim_{x\to+\infty} e^{\lambda x}f(x) = +\infty$$

donc n'est pas intégrable.

□

Exemple 4 (La loi lognormale). *Sa fonction de densité est $f(x) = \frac{1}{\sqrt{2\pi}\sigma x}\exp(-\frac{(\ln(x)-\mu)^2}{2\sigma^2})$, où $\mu \in \mathbb{R}$, et $\sigma > 0$. Tout ses moments sont finis. Une variable aléatoire ζ suit une loi lognormale de paramètres μ et σ si $\ln(\zeta)$ suit une loi normale (μ, σ^2).*

Démonstration. Soit $\lambda > 0$, alors $e^{\lambda x}f(x) = \frac{1}{\sqrt{2\pi}\sigma x}\exp(-\frac{(\ln(x)-\mu)^2}{2\sigma^2} + \lambda x)$ donc $\lim_{x\to+\infty} e^{\lambda x}f(x) = +\infty$ par croissances comparées. Donc la fonction n'a aucun moment exponentiel, et on conclut. □

Exemple 5 (La loi de Weibull). *Sa fonction de survie est $\overline{F}(x) = e^{-(\frac{x}{\lambda})^{\alpha}}$ où $\lambda, \alpha > 0$. C'est une loi à queue lourde si et seulement si $\alpha < 1$. On retrouve au passage pour $\alpha = 1$ la loi exponentielle qui n'est donc pas à queue lourde. Tout ses moments sont finis.*

Démonstration. Soit $\lambda > 0$, alors $\overline{F}(x)e^{\mu x} = e^{-(\frac{x}{\lambda})^{\alpha}+\mu x}$. Si $\alpha < 1$, on a bien que $\lim_{x\to+\infty} \overline{F}(x)e^{\lambda x} = +\infty$, et donc que la loi est à queue lourde, sinon $x \to \overline{F}(x)e^{\mu x}$ est bornée et alors la loi est à queue légère. □

1.4 Lois sous-exponentielles

En pratique, les lois à queues lourdes rencontrées sont régulières et appartiennent à une classe plus grande de fonctions, les lois sous-exponentielles.

Introduisons ces dernières. On notera $F * G$ le produit de convolution. En particulier, on a, si ζ et ν sont des variables aléatoires indépendantes de distributions F et G, alors pour tout x, $F * G(x) = P(\zeta + \nu > x)$.

Nous allons nous intéresser à l'estimation de

$$\liminf_{x \to +\infty} \frac{\overline{F * F}(x)}{\overline{F}(x)}$$

.

Propriété 3. *Soit $F_1, ..., F_n$ des lois sur $\mathbb{R}_+$ à support non borné à droite. On a alors :*

$$\liminf \frac{\overline{F_1 * .... * F_n}(x)}{\overline{F}_1(x) + ... + \overline{F}_n(x)} \geq 1.$$

Démonstration. Soit $\zeta_1, ..., \zeta_n$ des variables aléatoires indépendantes de distribution $F_1, ..., F_n$. Alors les événements $\{\zeta_k > x,\ \zeta_j \in [0; x]\ \forall j \neq k\}$ sont disjoints. Alors :

$$\overline{F_1 * ..... * F_n(x)} \geq \sum_{k=1}^{n} P(\zeta_k \geq x,\ \zeta_j \in [0; x],\ \forall j \neq k) = \sum_{k=1}^{n} \overline{F}_k(x) \prod_{k \neq j} F_j(x) \underset{x \to +\infty}{\sim} \sum_{k=1}^{n} \overline{F}_k(x).$$

□

Remarque 2. *L'hypothèse de positivité est essentielle : sinon, la première inégalité n'est plus vraie, l'inclusion n'étant pas vérifiée (les potentiels ζ_i négatifs pouvant "tirer en arrière" la somme).*

Propriété 4. *En particulier, si F est une distribution sur $\mathbb{R}_+$ à support non borné, on a, pour $n \geq 2$,*

$$\liminf_{x \to +\infty} \frac{\overline{F^{*n}(x)}}{\overline{F}(x)} \geq n$$

Démonstration. On applique la propriété précédente. □

Les distributions exponentielles affinent ce résultat. En effet, le cas d'égalité est rarement atteint (dans le cas d'une loi exponentielle, on trouve d'ailleurs l'infini). Par contre, nous avons le résultat suivant :

Théorème 2. *Si F est une loi à queue lourde sur $\mathbb{R}_+$, on a :*

$$\liminf_{x \to +\infty} \frac{\overline{F * F}(x)}{\overline{F}(x)} = 2.$$

Pour montrer ce théorème, nous avons besoin de ce résultat assez fort :

Théorème 3. *Soit $\zeta \geq 0$ une variable aléatoire à queue lourde. Soit g une fonction telle que $\lim_{x\to+\infty} g(x) = +\infty$. Alors il existe une fonction h de $\mathbb{R}^+$ dans lui-même vérifiant $h(x) = o(x)$ en l'infini, $\mathbb{E}(e^{h(\zeta)}) < +\infty$ et $\mathbb{E}(e^{h(\zeta)+g(\zeta)}) = +\infty$.*

Exemple 6 (Loi de Weibull)**.** *Par exemple, si ζ suit une loi de Weibull avec une fonction de survie $\overline{F}(x) = e^{-x^\alpha}$ où $0 < \alpha < 1$, que l'on prend $g(x) = \ln(x)$ pour $x > 0$, alors $h(x) = (x+c)^\alpha - ln(x+c)$ convient, avec c assez grand.*

Remarque 3. *En réalité, ceci est une condition nécessaire et suffisante. En effet, si ζ est à queue légère, alors il existe $\lambda > 0$ tel que $\mathbb{E}(e^{\lambda\zeta}) < +\infty$. On choisit alors $g = \ln$ et donc, si $h(x) = o(x)$ en l'infini, il existe c tel que $h(x) < c + \lambda\frac{x}{2}$ et donc :*

$$\mathbb{E}(e^{h(\zeta)+g(\zeta)}) \leq \mathbb{E}(\zeta e^{c+\zeta\frac{\lambda}{2}}) < +\infty.$$

Cette caractérisation ne semble pas être très utile.

Montrons ce résultat.

Démonstration. Nous allons construire h comme une fonction affine par morceaux. A cet égard, nous construirons deux suites positives monotones (x_n) et (ϵ_n) de limites respectives $+\infty$ et 0, et on posera :

$$h(x) = h(x_{n-1}) + \epsilon_n(x - x_{n-1}) \quad \text{si} \quad x \in (x_{n-1}, x_n].$$

Cette fonction est clairement monotone par monotonie et positivité des deux suites. De plus, elle est concave.

On prend d'abord $x_0 = h(0) = 0$. ζ étant à queue lourde et $\lim_{x\to+\infty} g(x) = +\infty$, on peut se donner x_1 tel que $e^{g(x)} \geq 2^1$ pour $x > x_1$ et

$$\mathbb{E}(e^{\zeta}\mathbb{1}_{\zeta\in(x_0,x_1]}) + e^{x_1}\overline{F}(x_1) > \overline{F}(x_0) + 1.$$

D'après le théorème des valeurs intermédiaires, on peut choisir $\epsilon_1 > 0$ tel que

$$\mathbb{E}(e^{\epsilon_1\zeta}\mathbb{1}_{\zeta\in(x_0,x_1]}) + e^{\epsilon_1 x_1}\overline{F}(x_1) = \overline{F}(0) + \frac{1}{2}$$

Ceci se réecrit donc :

$$\mathbb{E}(e^{h(\zeta)}\mathbb{1}_{\zeta\in(x_0,x_1]}) + e^{h(x_1)}\overline{F}(x_1) = e^{h(x_0)}\overline{F}(0) + \frac{1}{2}.$$

Pour s'en convaincre, c'est l'indicatrice dans l'espérance qui fait apparaître les bons termes, il n'y a qu'à l'écrire en remplaçant avec la définition de h sur $(x_0, x_1]$.

Pour construire la fonction sur $(x_n, x_{n+1}]$ pour $n \geq 1$, on procède par récurrence en construisant une suite croissante (x_n), une suite décroissante (ϵ_n) telle que $e^{g(x)} \geq 2^n$ pour $x > x_n$ et

$$\mathbb{E}(e^{h(\zeta)}\mathbb{1}_{\zeta\in(x_{n-1},x_n]}) + e^{h(x_n)}\overline{F}(x_n) = e^{h(x_{n-1})}\overline{F}(x_{n-1}) + \frac{1}{2^n}.$$

On a déjà effectué l'initialisation, supposons la propriété vraie pour un certain $n \geq 2$. Comme g tend vers l'infini en l'infini, on peut trouver x_{n+1} de telle sorte que $e^{g(x)} \geq 2^{n+1}$ pour $x > x_{n+1}$ et

$$\mathbb{E}(e^{\epsilon_n(\zeta - x_n)}\mathbb{1}_{\zeta \in (x_n; x_{n+1}]} + e^{\epsilon_n(x_{n+1}-x_n)}\overline{F}(x_{n+1}) > 2.$$

Or, le terme de gauche est continu en ϵ_n et tend vers $\overline{F}(x_n)$ en décroissant quand ϵ_n tend vers 0. On peut donc choisir $0 < \epsilon_{n+1} < \epsilon_n$ tel que :

$$\mathbb{E}(e^{\epsilon_n(\zeta - x_n)}\mathbb{1}_{\zeta \in (x_n; x_{n+1}]}) + e^{\epsilon_n(x_{n+1}-x_n)}\overline{F}(x_{n+1}) = \overline{F}(x_n) + \frac{1}{2^{n+1}e^{h(x_n)}}.$$

Or, d'après la définition de h, on retrouve comme lors de l'initialisation l'égalité suivante :

$$\mathbb{E}(e^{h(\zeta)}\mathbb{1}_{\zeta \in (x_n, x_{n+1}]}) + e^{h(x_{n+1})}\overline{F}(x_{n+1}) = e^{h(x_n)}\overline{F}(x_n) + \frac{1}{2^{n+1}}.$$

Nous avons donc bien montré l'hérédité.

Montrons maintenant que cette fonction convient.

Soit $N \in \mathbb{N}$. Alors on a :

$$\mathbb{E}(e^{h(\zeta)}\mathbb{1}_{\zeta \leq x_{N+1}}) = \sum_{n=0}^{N} \mathbb{E}(e^{h(\zeta)}\mathbb{1}_{\zeta \in (x_n, x_{n+1}]})$$

Or,

$$\sum_{n=0}^{N} \mathbb{E}(e^{h(\zeta)}\mathbb{1}_{\zeta \in (x_n, x_{n+1}]}) = \sum_{n=0}^{N} \left(e^{h(x_n)}\overline{F}(x_n) - e^{h(x_{n+1})}\overline{F}(x_{n+1}) + \frac{1}{2^{n+1}}\right)$$

et

$$\sum_{n=0}^{N} \left(e^{h(x_n)}\overline{F}(x_n) - e^{h(x_{n+1})}\overline{F}(x_{n+1}) + \frac{1}{2^{n+1}}\right) \leq e^{h(x_0)}\overline{F}(x_0) + 1$$

Donc on a

$$\mathbb{E}(e^{h(\zeta)}\mathbb{1}_{\zeta \leq x_{N+1}}) \leq e^{h(x_0)}\overline{F}(x_0) + 1$$

On a donc bien montré, d'après le théorème de convergence monotone, que $\mathbb{E}(e^{h(\zeta)})$ est finie. Maintenant, comme $e^{g(x)} \geq 2^n$ si $x > x_n$, alors :

$$\mathbb{E}(e^{h(\zeta)+g(\zeta)}\mathbb{1}_{\zeta > x_n}) \geq 2^n\mathbb{E}(e^h(\zeta)\mathbb{1}_{\zeta > x_n}) \geq 2^n\big(\mathbb{E}(e^{h(\zeta)}\mathbb{1}_{\zeta \in (x_n; x_{n+1}]}) + e^{h(x_{n+1})}\overline{F}(x_{n+1})\big)$$

Or,

$$2^n\big(\mathbb{E}(e^{h(\zeta)}\mathbb{1}_{\zeta \in (x_n; x_{n+1}]}) + e^{h(x_{n+1})}\overline{F}(x_{n+1})\big) = 2^n(e^{h(x_n)}\overline{F}(x_n) + \frac{1}{2^{n+1}}),$$

où l'on a remplacé h par sa définition en s'aidant de l'indicatrice.

Donc,

$$\mathbb{E}(e^{h(\zeta)+g(\zeta)}\mathbb{1}_{\zeta>x_n}) \geq 2^n(e^{h(x_n)}\overline{F}(x_n) + \frac{1}{2^{n+1}}).$$

On a donc $\mathbb{E}(e^{h(\zeta)+g(\zeta)}\mathbb{1}_{\zeta>x_n}) \geq \frac{1}{2}$ pour tout $n \in \mathbb{N}$. Donc on a bien $\mathbb{E}(e^{h(\zeta)+g(\zeta)}) = \infty$, ce qu'il nous manquait pour conclure. □

Nous pouvons donc maintenant prouver le théorème.

Démonstration. Il nous reste à prouver :

$$\liminf_{x\to+\infty} \frac{\overline{F*F}(x)}{\overline{F}(x)} \leq 2,$$

l'autre inégalité ayant été montrée plus haut.

Nous allons procéder par l'absurde. On suppose alors qu'il existe $\delta > 0$, un réel x_0 tels que si $x \geq x_0$, $\overline{F*F}(x) \geq (2+\delta)\overline{F}(x)$. On applique alors le résultat précédent avec $g = \ln$. On obtient alors $h : \mathbb{R}^+ \longrightarrow \mathbb{R}^+$ tel que $\mathbb{E}(e^{h(\zeta)}) < \infty$ et $\mathbb{E}(\zeta e^{h(\zeta)}) = \infty$. On pose alors la fonction h_b pour $b > 0$ suivante :

$$h_b(x) = min\big(h(x), bx\big)$$

Or, $h(x) = o(x)$ en l'infini, donc pour tout b il existe un x_b tel que $h_b(x) = h(x)$ pour $x > x_b$. On a donc $\mathbb{E}(e^{h_b(x)}) < \infty$ et $\mathbb{E}(\zeta e^{h_b(x)}) = \infty$.

De plus, pour tout $x \geq 0$, $\lim_{b\longrightarrow 0} h_b(x) = 0$ donc $\lim_{b\longrightarrow 0} \mathbb{E}(e^{h_b(\zeta_1)}) = 1$. On se donne alors b tel que $\mathbb{E}(e^{h_b(\zeta_1)}) \leq 1 + \frac{\delta}{4}$.

Maintenant, pour a et t réels, on pose $a^{[t]} = \min(a,t)$.

Alors :

$$\mathbb{E}((\zeta_1^{[t]} + \zeta_2^{[t]})e^{h_b(\zeta_1+\zeta_2)}) = 2\mathbb{E}(\zeta_1^{[t]}e^{h_b(\zeta_1+\zeta_2)}) \leq 2\mathbb{E}(\zeta_1^{[t]}e^{h_b(\zeta_1+\zeta_2)})$$

par concavité de h_b (comme minimum de fonctions concaves).

Mais alors :

$$\frac{\mathbb{E}((\zeta_1^{[t]} + \zeta_2^{[t]})e^{h_b(\zeta_1+\zeta_2)})}{\mathbb{E}(\zeta_1^{[t]}e^{h_b(\zeta_1)})} \leq 2\frac{\mathbb{E}(\zeta_1^{[t]}e^{h_b(\zeta_1)})\mathbb{E}(e^{h_b(\zeta_2)})}{\mathbb{E}(\zeta_1^{[t]}e^{h_b(\zeta_1)})} = 2\mathbb{E}(e^{h_b(\zeta_2)}) \leq 2 + \frac{\delta}{2}.$$

De plus, on remarque que $(\zeta_1+\zeta_2)^{[t]} \leq \zeta_1^{[t]} + \zeta_2^{[t]}$, on en déduit que :

$$\frac{\mathbb{E}((\zeta_1^{[t]} + \zeta_2^{[t]})e^{h_b(\zeta_1+\zeta_2)})}{\mathbb{E}(\zeta_1^{[t]}e^{h_b(\zeta_1)})} \geq \frac{\mathbb{E}((\zeta_1+\zeta_2)^{[t]}e^{h_b(\zeta_1+\zeta_2)})}{\mathbb{E}(\zeta_1^{[t]}e^{h_b(\zeta_1)})} = \frac{\int_0^\infty x^{[t]}e^{h_b(x)}F*F(dx)}{\int_0^\infty x^{[t]}e^{h_b(x)}F(dx)}.$$

De fastidieuses intégrations par parties montrent que le dernier morceau est égal à

$$\frac{\int_0^\infty \overline{F*F}(x)d(x^{[t]}e^{h_b(x)})}{\int_0^\infty \overline{F}(x)d(x^{[t]}e^{h_b(x)})}$$

Or, on sait que $\mathbb{E}(\zeta_1 e^{h_b(\zeta_1)}) = \infty$, ce qui nous montre que les deux intégrales tendent vers l'infini. Or, de $\overline{F*F}(x) \geq (2+\delta)\overline{F}(x)$ pour $x > x_0$ on déduit que

$$\liminf_{x\to+\infty} \frac{\int_0^\infty \overline{F*F}(x)d(x^{[t]}e^{h_b(x)})}{\int_0^\infty \overline{F}(x)d(x^{[t]}e^{h_b(x)})} \geq 2+\delta$$

Donc

$$\liminf_{x\to+\infty} \frac{\mathbb{E}((\zeta_1^{[t]}+\zeta_2^{[t]})e^{h_b(\zeta_1+\zeta_2)}}{\mathbb{E}(\zeta_1^{[t]}e^{h_b(\zeta_1)}} \geq 2+\delta.$$

Or, on a montré plus haut que :

$$\frac{\mathbb{E}((\zeta_1^{[t]}+\zeta_2^{[t]})e^{h_b(\zeta_1+\zeta_2)})}{\mathbb{E}(\zeta_1^{[t]}e^{h_b(\zeta_1)})} \leq 2+\frac{\delta}{2}$$

ce qui établit une contradiction. □

Ainsi, si une loi à queue lourde est suffisamment régulière, la limite inférieure $\liminf\limits_{x\to+\infty} \frac{\overline{F*F}(x)}{\overline{F}(x)}$ est une limite, et vaut 2. On dit qu'une telle loi est une loi sous-exponentielle. Ces lois sont très commodes, et la plupart des lois à queues lourdes en font partie.

Définition 7 (Loi sous-exponentielle positive). *Si F est une loi sur $\mathbb{R}_+$ à support non borné, on dit que F est sous-exponentielle, que l'on note $F \in \mathcal{S}$ si $\overline{F*F}(x) \underset{x\to+\infty}{\sim} 2\overline{F}(x)$.*

Définition 8 (Loi sous-exponentielle). *Une loi sur $\mathbb{R}$ est dite sous-exponentielle si F^+ est sous-exponentielle, où $F^+ = max(0, F)$.*

Propriété 5 (Principe du simple saut). *Soit ζ_1 et ζ_2 des variables aléatoires indépendantes sur $\mathbb{R}_+$ de distribution F. Alors F est sous-exponentielle si et seulement si*

$$P(\zeta_1+\zeta_2 > x) \underset{x\to+\infty}{\sim} 2P(\zeta_1 > x).$$

Ceci se réécrit

$$\lim_{x\to+\infty} P(\zeta_1 > x | \zeta_1+\zeta_2 > x) = \frac{1}{2},$$

ou encore

$$P(\max(\zeta_1,\zeta_2) > x) = 1-(1-P(\zeta_1 > x))^2 \underset{x\to+\infty}{\sim} 2P(\zeta_1 > x)$$

Remarque 4. *L'idée est la suivante : si la somme de deux variables aléatoires sous-exponentielles indépendantes est plus grande que x très grand, alors c'est que l'une d'entre elle excède x. En gros, la somme est grande si quelqu'un est grand ; on fait de grandes choses si quelqu'un se motive et fait tout pour tout le monde !*

Définition 9 (Equivalence faible des queues)**.** *Deux lois F et G dont le support n'est pas borné à droite sont dites faiblement équivalentes si il existe $c_1 > 0$, $c_2 < +\infty$ tels que, pour tout $x \in \mathbb{R}$,*

$$c_1 \leq \frac{\overline{F}(x)}{\overline{G}(x)} \leq c_2.$$

Cette notion d'équivalence faible permettra de montrer à partir d'une première loi sous-exponentielle qu'une autre l'est. C'est ce que nous allons voir maintenant :

Définition 10. *Une loi F est dite insensible à une fonction h si sa fonction de survie $\overline{F}$ vérifie $\overline{F}(x \pm h(x)) \underset{x\to+\infty}{\sim} \overline{F}(x)$.*

Remarque 5. *Ce n'est pas la vraie définition, la vraie définition utilise la notion de fonction insensible à h ; mais cela ne nous servirait pas ici.*

Propriété 6. *Soit G à queue longue, h une fonction croissante telle que $\lim\limits_{x\to+\infty} h(x) = +\infty$, et G est insensible à h. Alors pour toute loi F sur $\mathbb{R}$, on a :*

$$\int_{-\infty}^{h(x)} \overline{G}(x-y)F(dy) \underset{x\to+\infty}{\sim} \overline{G}(x),$$

et :

$$\int_{x-h(x)}^{+\infty} \overline{F}(x-y)G(dy) \underset{x\to+\infty}{\sim} \overline{G}(x).$$

Démonstration. D'après les propriétés évidentes sur les fonctions de survie, on a :

$$\int_{-\infty}^{h(x)} \overline{G}(x-y)F(dy) \leq \overline{G}(x-h(x))$$

. De plus,

$$\int_{-\infty}^{h(x)} \overline{G}(x-y)F(dy) \geq \int_{-h(x)}^{h(x)} \overline{G}(x-y)F(dy) \geq F((-h(x),h(x)])\overline{G}(x+h(x))$$

$$\underset{x\to+\infty}{\sim} \overline{G}(x+h(x)) \underset{x\to+\infty}{\sim} \overline{G}(x).$$

Où la dernière équivalence provient de l'insensibilité de G à h, et l'avant dernière du fait que $F(\mathbb{R}) = 1$.

La seconde partie se montre de manière absolument similaire. □

Propriété 7. *Soit F une loi à queue longue, ζ_1 et ζ_2 des variables aléatoires indépendantes de distribution F. Soit h vérifiant $\lim\limits_{x\to+\infty} h(x) = +\infty$, et telle que F est insensible à h. Alors F est sous-exponentielle si et seulement si :*

$$P(\zeta_1+\zeta_2 > x, \zeta_1 > h(x), \zeta_2 > h(x)) = o(\overline{F}(x))$$

en $+\infty$.

Démonstration. On suppose dans un premier temps que $h(x) \leq \frac{x}{2}$ pour tout $x \in \mathbb{R}$. Alors, à l'aide d'une union disjointe :

$$P(\zeta_1 + \zeta_2 > x) =$$
$$P(\zeta_1+\zeta_2 > x, \zeta_1 \leq h(x))+P(\zeta_1+\zeta_2 > X, \zeta_2 \leq h(x))+P(\zeta_1+\zeta_2 > x, \zeta_1 > h(x), \zeta_2 > h(x)).$$

Mais, F est à queue longue, donc d'après la 6, on a que :

$$P(\zeta_1 + \zeta_2 > x, \zeta_i \leq h(x)) \underset{x\to+\infty}{\sim} \overline{F}(x).$$

On montre alors que $P(\zeta_1+\zeta_2 > x) \underset{x\to+\infty}{\sim} 2\overline{F}(x)$, ce qui montre que F est sous-exponentielle.

Le fait d'avoir supposé que $h(x) \leq \frac{x}{2}$ permet d'obtenir une "preuve de poche". Si on ne le suppose plus, il faut refaire les calculs avec $\min(h(x), \frac{x}{2})$, ce qui change certaines égalités en des inégalités, et permet d'obtenir le même résultat. □

Un dernier lemme un peu technique, avant de pouvoir énoncer les résultats...

Propriété 8. *Soit h croissante telle que $\lim_{x\to+\infty} h(x) = +\infty$. Alors si F_1, F_2, G_1 et G_2 sont des lois sur $\mathbb{R}$, on a :*

$$\limsup_{x\to+\infty} \frac{P(\zeta_1 + \nu_1 > x, \zeta_1 > h(x), \nu_1 > h(x))}{P(\zeta_2 + \nu_2 > x, \zeta_2 > h(x), \nu_2 > h(x))} \leq \limsup_{x\to+\infty} \frac{\overline{F_1}(x)}{F_2(x)} \limsup_{x\to+\infty} \frac{\overline{G_1}(x)}{\overline{G_2}(x)}$$

Démonstration. D'après 6, on sait que :

$$P(\zeta_1 + \nu_1 > x, \zeta_1 > h(x), \nu_1 > h(x)) \leq \sup_{z>h(x)} \frac{\overline{F_1}(z)}{\overline{F_2}(z)} \int_{h(x)}^{+\infty} \overline{F_2}(\max(h(x), x-y))G_1(dy)$$
$$= \sup_{z>h(x)} \frac{\overline{F_1}(z)}{\overline{F_2}(z)} \int_{h(x)}^{+\infty} \overline{G_1}(\max(h(x), x-y))F_2(dy).$$

De même,

$$\int_{h(x)}^{+\infty} \overline{G_1}(\max(h(x), x-y))F_2(dy) \leq \sup_{z>h(x)} \frac{\overline{G_1}(z)}{\overline{G_2}(z)} \int_{h(x)}^{+\infty} \overline{G_2}(\max(h(x), x-y))F_2(dy)$$
$$= \sup_{z>h(x)} \frac{\overline{G_1}(z)}{\overline{G_2}(z)} P(\zeta_2 + \nu_2 > x, \zeta_2 > h(x), \nu_2 > h(x)).$$

On en déduit de la deuxième inégalité que :

$$P(\zeta_2 + \nu_2 > x, \zeta_2 > h(x), \nu_2 > h(x)) \geq \int_{h(x)}^{+\infty} \overline{G_1}(\max(h(x), x-y))F_2(dy) \cdot \frac{1}{\sup_{z>h(x)} \frac{\overline{G_1}(z)}{\overline{G_2}(z)}}.$$

En faisant le quotient de la première inégalité et de celle-ci, on obtient le résultat. (En particulier, on retrouve bien la $\limsup$ car $\lim_{x\to+\infty} h(x) = +\infty$). □

Propriété 9. *Soit F sous-exponentielle, et h telle que $\lim_{x\to+\infty} h(x) = +\infty$. Soit G_1 et G_2 telles que $\overline{G_i}(x) = O(\overline{F}(x))$ en l'infini. Alors si ν_1 et ν_2 sont des variables aléatoires indépendantes de distributions G_1 et G_2, on a que*

$$P(\nu_1 + \nu_2 > x, \nu_1 > h(x), \nu_2 > h(x)) = o(\overline{F}(x))$$

en $+\infty$.

Démonstration. Soit ζ_1 et ζ_2 des variables aléatoires indépendantes de loi F, comme $\overline{G_i}(x) = O(\overline{F}(x))$ la 8 montre que pour un certain c, on a :

$$P(\nu_1 + \nu_2 > x, \nu_1 > h(x), \nu_2 > h(x)) \leq cP(\zeta_1 + \zeta_2 > x, \zeta_1 > h(x), \zeta_2 > h(x)).$$

Or, F est sous-exponentielle, donc d'après la 7, on peut conclure. □

Propriété 10. *Si F est sous-exponentielle, que G est à queue longue, et que F et G sont faiblement équivalentes. Alors G est sous-exponentielle.*

Démonstration. Soit h telle que $\lim_{x\to+\infty} h(x) = +\infty$, et telle que G soit insensible à h. Soit ζ_1 et ζ_2 des variables aléatoires indépendantes de distribution G. Alors, d'après la 9 et l'équivalence faible,

$$P(\zeta_1 + \zeta_2 > x, \zeta_1 > h(x), \zeta_2 > h(x)) = o(\overline{F}(x)) = o(\overline{G}(x)).$$

Donc, d'après, on a bien le fait que G est sous-exponentielle. □

Définition 11 (Equivalence proportionnelle)**.** *F et G sont dites proportionnellement équivalentes si il existe $c > 0$ tel que $\overline{F}(x) \underset{x\to+\infty}{\sim} c\overline{G}(x)$.*

Propriété 11. *Si F et G sont proportionnellement équivalentes, que F est sous-exponentielle, alors G l'est.*

Démonstration. Conséquence directe de la propriété précédente. □

1.5 Temps d'arrêt et lois sous-exponentielles

Dans cette partie, nous allons chercher à montrer le résultat suivant, qui nous sera utile plus tard :

Propriété 12. *Soit τ un temps d'arrêt intégrable vérifiant $\mathbb{E}((1+\delta)^\tau) < \infty$ pour un certain $\delta > 0$, F une loi sous-exponentielle. Alors :*

$$\lim_{x\to+\infty} \frac{P(S_\tau > x)}{\overline{F}(x)} = \mathbb{E}(\tau).$$

Cette partie est dédiée à la preuve de ce résultat. La loi que nous considérons est sous-exponentielle, donc nous savons que $\lim_{x\to+\infty} \frac{\overline{F*F}(x)}{\overline{F}(x)} = 2$. Néanmoins, nous allons avoir besoin d'appliquer le théorème de convergence dominée ; et pour ce faire, il nous faut une majoration. Cette majoration est nommée "l'estimation de Kesten".

Propriété 13 (Estimation de Kesten). *Soit $F \in \mathcal{S}$. Alors pour tout $\epsilon > 0$, il existe $c > 0$ tel que pour tout $x \geq 0$, $n \geq 1$, on ait :*

$$\overline{F^{*n}}(x) \leq c(1+\epsilon)^n \overline{F}(x).$$

Montrons ce résultat.

Démonstration. Soit (ζ_n) une suite de variables aléatoires indépendantes de distribution F, et $S_n = \sum_{i=1}^{n} \zeta_i$. Si $x_0 > 0$ et $k \geq 1$, on pose :

$$A_k = \sup_{x > x_0} \frac{\overline{F^{*k}}(x)}{\overline{F}(x)}.$$

On peut supposer que les ζ_i sont positives car $\zeta \leq \zeta^+$ donc l'inégalité n'en serait pas perturbée. On a donc l'existence de x_0 tel que, pour $x > x_0$, on ait :

$$P(\zeta_1 + \zeta_2, \zeta_2 \leq x) = P(\zeta_1 + \zeta_2 > x) - P(\zeta_2 > x) \leq (1 + \frac{\epsilon}{2}\overline{F}(x)).$$

On a, dans la première égalité, utilisé la positivité des lois, et dans la seconde utilisé la sous exponentialité pour s'occuper du premier terme.

Maintenant, on écrit :

$$P(S_n > x) = P(S_n > x, \zeta_2 \leq x - x_0) + P(S_n > x, \zeta_n > x - x_0) = P_1(x) + P_2(x).$$

On remarque alors :

$$P_1(x) = \int_0^{x-x_0} P(\S_{n-1} > x - y) dP(\zeta_n \in dy) \leq A_{n-1} \int_0^{x-x_0} \overline{F}(x-y) dP(\zeta_n \in dy)$$

$$= A_{n-1} P(\zeta_1 + \zeta_n > x, \zeta_n \leq x - x_0) \leq A_{n-1}(1 + \frac{\epsilon}{2}).$$

Aucune de ces égalités ou inégalités ne pose de problème : la première n'est qu'une définition, la seconde est immédiate avec la définition de de A_{n-1}, la troisième n'est qu'une réécriture de la seconde, et la quatrième provient de la définition de x_0.

Maintenant, par une simple inclusion d'événénements :

$$P_2(x) \leq P(\zeta_n > x - x_0) \leq L\overline{F}(x),$$

où

$$L = \sup_y \frac{\overline{F}(y - x_0)}{\overline{F}(y)}.$$

De plus, L est finie car $\overline{F}$ est à queue longue.

Mais alors, par récurrence, on obtient :

$$A_n \leq A_1(1+\frac{\epsilon}{2})^{n-1} + L\sum_{k=0}^{n-2}(1+\frac{\epsilon}{2})^k \leq Ln(1+\frac{\epsilon}{2})^{n-1},$$

d'où, pour tout $x > x_0$:

$$\frac{\overline{F^{*n}}(x)}{\overline{F}(x)} \leq A_n = \sup_{x>x_0} \frac{\overline{F^{*n}}(x)}{\overline{F}(x)} \leq Ln(1+\frac{\epsilon}{2})^{n-1},$$

c'est bien le résultat attendu. □

Enfin, il nous reste deux propriétés à montrer sur les convolutions de loi sous-exponentielles.

Propriété 14. *Soit F une loi sous-exponentielle. Soit $G_1, ..., G_n$ des lois sous-exponentielles telles que $\overline{F} + \overline{G_i}$ soit à queue longue et que $\overline{G_i} = O(F)$ en l'infini. Alors, en $+\infty$:*

$$\overline{G_1 * ..... * G_n}(x) = \overline{G_1}(x) + ...\overline{G_n}(x) + o(\overline{F}(x)).$$

Démonstration. C'est du calcul, avec les résultats vus précédemment sur les fonctions h insensible. Ce n'est pas très intéressant. □

Propriété 15. *Dans les mêmes conditions que la 14, en supposant également que G_1 est à queue longue et faiblement équivalente à F. Alors $G_1 * .... * G_n$ est à queue longue, faiblement équivalente à G_1. En particulier, elle est faiblement équivalente à F et est donc sous-exponentielle.*

Démonstration. La preuve est simple : la 10 montre que G_1 est sous-exponentielle. La faible équivalence montre que $\overline{G_k}(x) = O(\overline{G_1}(x))$ en $+\infty$ pour tout $1 \leq k \leq n$. Donc en appliquant 14 avec $F = G_1$, on obtient directement que $G_1 * ..... * G_n$ est à queue longue, et faiblement équivalent à G_1, donc à F, et donc d'après 10 encore, $G_1 * .... * G_n$ est sous-exponentielle. □

Finissons par une conséquences des résultats précédents qui est la propriété dont nous aurons besoin :

Propriété 16. *Soit F sous-exponentielle, G_1,...,G_n des distributions telles que $\lim_{x\to+\infty} \frac{\overline{G_i}(x)}{\overline{F}(x)} = c_i$, où $c_i \geq 0$. Alors :*

$$\lim_{x\to+\infty} \frac{\overline{G_1 * ... * G_n}(x)}{\overline{F}(x)} = c_1 + ... + c_n.$$

De plus, si $c_1 + ... + c_n > 0$, $G_1 * .... * G_n$ est sous-exponentielle.

Démonstration. La première partie est une conséquence du 14 ; les hypothèses sont immédiatement vérifiées et le résultat tombe. Quant à la seconde partie, si la somme est strictement positive alors l'un des c_i est strictement positif, et on applique 15 avec $G_1 = G_i$. □

On en déduit la propriété suivante, peut être un peu attendue..

Propriété 17. *Soit F sous-exponentielle. Alors, si $n \geq 2$,*

$$\lim_{x \to +\infty} \frac{\overline{F^{*n}}(x)}{\overline{F}(x)} = n.$$

En particulier, $\overline{F^{*n}}$ est sous-exponentielle.

Démonstration. Cas particulier de 16. □

Revenons à nos moutons. Le résultat que nous voulons prouver, énoncé précédemment est :

Propriété 18. *Soit τ un temps d'arrêt intégrable vérifiant $\mathbb{E}((1+\delta)^{\tau}) < \infty$ pour un certain $\delta > 0$, F une loi sous-exponentielle. Alors :*

$$\lim_{x \to +\infty} \frac{P(S_\tau > x)}{\overline{F}(x)} = \mathbb{E}(\tau).$$

Avec le travail effectué précédemment, ça va très vite :

Démonstration. Nous allont utiliser 17 et 13, ainsi que le théorème de convergence dominée. D'après 17, $\dfrac{P(S_{t\wedge n} > x)}{\overline{F}(x)} = \mathbb{E}(\tau \wedge n)$. D'après 13, on peut appliquer le théorème de convergence dominée ; et on en déduit le résultat.

□

1.6 Approximation pour les fonctions et distributions à queues lourdes

Si f est intégrable au voisinage de $+\infty$, on note $f_I = \displaystyle\int_x^{+\infty} f(y)dy$.

Propriété 19. *Soit f positive intégrable en $+\infty$. Si l'une des deux propriétés suivantes est vérifiée :*

1. *f est décroissante et f_I à queue longue*
2. *f est à queue longue*

Alors, pour $a > 0$,

$$\sum_{n=0}^{\infty} f(x+na) \underset{x \to +\infty}{\sim} \frac{1}{a} f_I(x)$$

Démonstration. On suppose d'abord (1).

Alors, par décroissance de f,

$$\frac{1}{a}\int_{x+na}^{x+(n+1)a} f(t)dt \leq f(x+na) \leq \frac{1}{a}\int_{x+(n-1)a}^{x+na} f(t)dt$$

En sommant, il vient :

$$\frac{1}{a}f_I(x) \leq \sum_{n=0}^{+\infty} f(x+na) \leq \frac{1}{a}f_I(x-a)$$

On conclut par définition d'une fonction à queue longue.

Maintenant, on suppose (2) :

Soit $\epsilon > 0$. f étant à queue longue, pour x assez grand :

$$(1-\epsilon)f(t) \leq f(x+na) \leq (1+\epsilon)f(t), \quad \text{où} \quad t \in [x+na, x+(n+1)a]$$

On intègre alors et;

$$\frac{1-\epsilon}{a}\int_{x+na}^{x+(n+1)a} f(t)dt \leq \sum_{n=0}^{\infty} f(x+na) \leq \frac{1+\epsilon}{a}\int_{x+na}^{x+(n+1)a} f(t)dt$$

Donc, en sommant, pour x grand on a :

$$\frac{1-\epsilon}{a}f_I(x) \leq \sum_{n=0}^{\infty} f(x+na) \leq \frac{1+\epsilon}{a}f_I(x)$$

C'est bien ce qu'on voulait. □

Chapitre 2

Maximum de marches aléatoires

Sommaire

2.1 Introduction

On considérera donc ici une marche aléatoire, avec une déviation négative, c'est à dire que l'espérance commune des incréments est strictement négative. Nous allons montrer que le maximum d'une telle marche aléatoire est grand à condition qu'un "grand saut" ait été effectué, c'est à dire qu'un incrément soit lui même grand.

Ainsi, nous considérerons $(\zeta_n)_{n\in\mathbb{N}}$ une suite de variables aléatoires indépendantes et identiquement distribuées, F leur distribution commune, et on note $\mathbb{E}(\zeta_1) = -a < 0$. On pose $M_n = \max\{S_i, 0 \leq i \leq n\}$ et $M = \sup\{S_n, n \in \mathbb{N}\}$.

Propriété 20. *Presque sûrement,* $M < \infty$.

Démonstration. En effet, d'après la loi forte des grands nombres, on a $S_n \underset{n\to+\infty}{\sim} -na$ presque sûrement, ce qui donne le résultat. $\square$

2.2 Asymptotique du sup d'une marche aléatoire avec une déviation négative

Propriété 21. *Pour tout* $x \geq 0$,

$$P(M > x) \geq \frac{\int_x^{+\infty} \overline{F}(y)dy}{a + \int_x^{+\infty} \overline{F}(y)dy}$$

En particulier,

$$\liminf_{x\to+\infty} \frac{P(M>x)}{\overline{F}_I(x)} \geq \frac{1}{a}.$$

Remarque 6. *On observe que :*

$$\int_x^{\infty} \overline{F}(y)dy = \int_x^{\infty}\int_y^{\infty} dFdy = \int_x^{\infty}\int_x^{t} dydF = \int_x^{\infty}(t-x)dF =$$

$$\int_0^{\infty} \mathbb{1}_{t>x} tdF - x\int_x^{\infty} dF = \mathbb{E}(\zeta \mathbb{1}_{\zeta-x>0}) - xE(\mathbb{1}_{\zeta-x>0}) = \mathbb{E}((\zeta-x)\mathbb{1}_{\zeta-x>0})$$

où on a applique le théorème de Fubini positif pour la deuxième inégalité.

Démonstration. On rappelle que les notations M et M_n sont celles définies dans l'introduction du chapitre 2, et que si f est intégrable au voisinage de $+\infty$, alors pour tout x réel,

$$f_I(x) = \int_x^{+\infty} f(y)dy.$$

Soit ζ une variable aléatoire de distribution F, indépendante de M. Alors M et $(M+\zeta)^+$ ont même distribution. En effet, $(M+\zeta)$ a même loi que $M' = \sup\{M_n, n \geq 1\}$, et M'^+ a même loi que M (car notre marche aléatoire commence à 0).

Soit x et y des réels positifs. On pose alors la fonction affine par morceaux suivante :

$$L_z^x(y) = \begin{cases} x \text{ si } y \leq x \\ y \text{ si } y \in [x, x+z] \\ x+z \text{ sinon} \end{cases}$$

Cette fonction va nous servir à nous débarasser des termes grands, on les récupérera avec une convergence dominée. En effet, cette fonction est bornée donc $\mathbb{E}(L_z^x(M))$ existe et est égale à $\mathbb{E}(L_z^x(M+\zeta))$ car $x \geq 0$.

Donc :

$$\mathbb{E}(L_z^x(M+\zeta) - L_z^x(M)) = 0$$

Or, $|L_z^x(M+\zeta) - L_z^x(M)| \leq |\zeta|$ (car L_y^x est 1-liepschitzienne) et $\lim\limits_{z\to+\infty} L_z^x(M+\zeta) - L_z^x(M) = L^x(M+\zeta) - L^x(M)$ où :

$$L^x(y) = \begin{cases} x \text{ si } y \leq x \\ y \text{ si } y \geq x \end{cases}$$

On peut donc appliquer le théorème de convergence dominée car $\zeta \in \mathrm{L}_1$ et donc :

$$\mathbb{E}(L^x(M+\zeta) - L^x(M)) = 0$$

Maintenant, on remarque que si $y \in [0; x]$,

$$L^x(y+\zeta) - L^x(y) = (y+\zeta-x)\mathbb{1}_{\{y+\zeta>x\}} \geq (\zeta-x)\mathbb{1}_{\{\zeta>x\}},$$

ce qui donne, par indépendance de M et ζ,

$$\mathbb{E}(L^x(M+\zeta) - L^x(M)\mathbb{1}_{M\leq x}) \geq \mathbb{E}(\zeta - x\mathbb{1}_{\zeta > x})P(M \leq x) \tag{2.1}$$

d'autre part si $y > x$, on observe que $L^x(y+\zeta) - L^x(y) \geq \zeta$, d'où;

$$\mathbb{E}(L^x(M+\zeta) - L^x(M)\mathbb{1}_{M>x}) \geq \mathbb{E}(\zeta)P(M > x) \tag{2.2}$$

Ce qui donne, en sommant $(2,1)$ et $(2,2)$:

$$\mathbb{E}(\zeta - x; \zeta > x)P(M \leq x) \leq -\mathbb{E}(\zeta)P(M > x).$$

Qui se réécrit :

$$P(M > x) \geq \frac{\mathbb{E}(\zeta - x; \zeta > x)P(M \leq x)}{a} = \frac{\mathbb{E}(\zeta - x; \zeta > x)(1 - P(M > x))}{a}$$

D'où, en mettant tout le monde du bon côté :

$$P(M > x) \geq \frac{\mathbb{E}(\zeta - x; \zeta > x)}{a + \mathbb{E}(\zeta - x; \zeta > x)}$$

On obtient donc bien :

$$P(M > x) \geq \frac{\int_x^\infty \overline{F}(y)dy}{a + \int_x^\infty \overline{F}(y)dy} = \frac{\overline{F}_I(x)}{a + \overline{F}_I(x)}$$

□

On peut affiner ce résultat en ajoutant une hypothèse de sous exponentiennalité :

Théorème 4. *Si en plus d'avoir une déviation négative F_I est sous-exponentielle, alors :*

$$P(M > x) \underset{x\to+\infty}{\sim} \frac{1}{a}\overline{F}_I(x)$$

Démonstration. Grâce au résultat précédent, on a juste à montrer la majoration.
Soit $\epsilon > 0$ et $A > a$. On définit $\tau_1, \tau_2 ...$ par : $\tau_1 = \min_{j\geq 1}\{j, S_j \geq A - j(a-\epsilon)\}$ et $\tau_k = +\infty$ si $\tau_{k-1} = +\infty,\ \tau_k + \min_{j\geq 1}\{j,\ S_{\tau_{k-1}+j} - S_{\tau_{k-1}} > A - j(a-\epsilon)\}$ sinon. Ce sont des temps de renouvellement.
Par propriété de Markov, les τ_k étant des temps d'arrêt, les vecteurs aléatoires

$$(\tau_1, S_{\tau_1}), (\tau_2 - \tau_1, S_{\tau_2} - S_{\tau_1}), ..., (\tau_k - \tau_{k-1}, S_{\tau_k} - S_{\tau_{k-1}})$$

sont iid sous la condition τ_k est fini.
De plus, $\lim_{A\to+\infty} \gamma^A = P(\tau_1 < \infty) = 0$.

En effet, il existe presque sûrement $N \in \mathbb{N}$ tel que pour tout $n \geq N$, $\dfrac{S_n}{n} \leq -a + \dfrac{\epsilon}{2}$. Or, $[\tau_1 < +\infty] = [\exists n \geq 1, S_n > -an + \epsilon n + A]$. Or, pour $n \geq \mathbb{N}$, $S_n \leq (-an + \epsilon n) - \dfrac{\epsilon}{2} n < -an + \epsilon n + A$ pour A assez grand, ce qu'on voulait.

On remarque par ailleurs que si $\tau_1 = n$, alors $S_{n-1} \leq A - (n-1)(a - \epsilon)$. C'est pourquoi on peut écrire, en écrivant que $S_\infty = -\infty$ pour ne pas avoir à conditionner :

$$P(S_{\tau_1} > x) = \sum_{n=1}^{\infty} P(\tau_1 = n, S_n > x)$$

$$\leq \sum_{n=1}^{\infty} P(S_{n-1} \leq A - (n-1)(a-\epsilon), S_n > x)$$

$$\leq \sum_{n=1}^{\infty} P(\zeta_n > x - A + (n-1)(a-\epsilon)),$$

où la première égalité est la formule des probabilités totales, la seconde une inégalité dûe à l'inclusion des événements susmentionnée, et la troisième est également une inclusion d'événements évidente (c'est le saut que ζ_n doit effectuer pour être dans l'intersection..).

Ainsi, $P(S_{\tau_1} > x) \leq \sum_{n=1}^{\infty} \overline{F}(x - A + n(a - \epsilon)) \leq \dfrac{1}{a - \epsilon} \overline{F}_I(x - A - a + \epsilon)$, où la deuxième inégalité correspond à la comparaison série intégrale effectuée pour prouver la propriété 6 sous l'hypothèse 1.

Maintenant, ça devient délicat :

Soit $\varphi_1, \ldots$. des variables aléatoires i.i.d telles que $P(\varphi_1 > x) = P(S_{\tau_1} > x | \tau_1 < \infty)$. L'existence provient du fait qu'on a fixé les fonctions de répartition, donc de telles variables existent, mais on peut les choisir indépendantes sur l'espace produit ...

D'après l'inégalité obtenue précédemment, On a, pour tout $x \in \mathbb{R}$, $P(\varphi_1 > x) \leq \overline{G}(x)$, où G est une loi sur $\mathbb{R}$ vérifiant $\dfrac{\overline{G}(x)}{\overline{F}_I(x)} = \dfrac{1}{\gamma(a - \epsilon)}$.

En effet, c'est l'inégalité vue plus haut, $P(S_{\tau_1} > x) \leq \dfrac{1}{a - \epsilon} \overline{F}_I(x - A - a + \epsilon)$. On utilise alors que F_I étant à queue longue, $\overline{F}_I(x - A - a + \epsilon) \underset{x \to +\infty}{\sim} \overline{F}_I(x)$.

D'après les résultats vu dans la partie sur les lois sous-exponentielles, en particulier 10, on déduit du fait que F_I soit sous-exponentielle et du fait que les deux lois soient proportionnellement équivalentes que G est sous-exponentielle.

Ainsi, on est dans le cadre d'application du théorème sur les lois sous-exponentielles et les temps d'arrêt (12). On peut alors se donner un temps d'arrêt suivant une loi géométrique de paramètre $1 - \gamma$, et on obtient que

$$\lim_{x \longrightarrow +\infty} \frac{P(S'_\tau > x)}{\overline{G}(x)} = \mathbb{E}(\tau).$$

où S' désigne la marche aléatoire associée à G. Soit :

$$P(S'_\tau > x) \underset{x\to+\infty}{\sim} \overline{G}(x)\mathbb{E}(\tau) = \frac{\overline{G}(x)}{1-\gamma}$$

Et, la formule des probabilités totales donne :

$$\mathbb{P}(S'_\tau > x) = \sum_{n=1}^{\infty} P(\tau = n)P(S'_n > x) = (1-\gamma)\sum_{n=1}^{\infty} \gamma^{k-1}\overline{G^{*k}}(x)$$

On trouve donc :

$$(1-\gamma)\sum_{k=0}^{\infty} \gamma^k \overline{G^{*k}}(x) \underset{x\to+\infty}{\sim} \frac{\gamma}{1-\gamma}\overline{G}(x).$$

Autrement dit, on a montré :

$$\sum_{k=1}^{\infty} \gamma^k P(\varphi_1 + ... + \varphi_k > x) \leq \frac{\gamma + o(1)}{(1-\gamma)^2}\overline{G}(x) \leq \frac{\gamma + o(1)}{(1-\gamma)^2}\frac{\overline{F}_I(x)}{\gamma(a-\epsilon)} = \frac{1+o(1)}{(1-\gamma)^2(a-\epsilon)}\overline{F}_I(x)$$

Maintenant ; si $M > x$, il existe $k > 0$ et $\tau_k \leq j < \tau_{k+1}$ tel que $S_j > x$. (l'existence provient du fait qu'on peut avoir $\tau_{k+1} = \infty$). Alors on a $S_{\tau_k} > x - A + a - \epsilon$. En effet, si ce n'est pas le cas, $S_{\tau_k} < x$ (on rappelle que $A >> a$), dans ce cas $\tau_k < j < \tau_{k+1}$ et $S_j - S_{\tau_k} > x - (x - A + a - \epsilon) = A - a + \epsilon$ et donc $\tau_{k+1} \leq j$, contradiction.

Donc :

$$[M \geq x] \subset \bigcup_{k=1}^{\infty}[S_{\tau_k} > x - A + \epsilon]$$

Donc, pour x assez grand,

$$P[M \geq x] \leq \sum_{k=1}^{\infty} P(S_{\tau_k} > x - A + \epsilon)$$

$$= \sum_{k=1}^{\infty} P(\tau_k < \infty)P(S_{\tau_k} > x - A + \epsilon | \tau_k < \infty)$$

$$= \sum_{k=1}^{\infty} P(\tau_k < \infty)P(S_{\tau_k} - S_{\tau_{k-1}} + S_{\tau_{k-1}} - S_{\tau_{k-2}} ... + S_{\tau_1} > x - A + \epsilon | \tau_k < \infty)$$

$$= \sum_{k=1}^{\infty} \gamma^k P(\varphi_1 + ... + \varphi_k > x - A + \epsilon).$$

Ici, on a utilisé le fait que les φ_k sont i.i.d et vérifient $P(\varphi_1 > x) = P(S_{\tau_1} > x | \tau_1 < \infty)$ et les vecteurs aléatoires

$$(\tau_1, S_{\tau_1}), (\tau_2 - \tau_1, S_{\tau_2 - \tau_1}), ..., (\tau_k - \tau_{k-1}, S_{\tau_k} - S_{\tau_{k-1}})$$

sont iid sous la condition τ_k est fini.

On utilise finalement les asymptotiques démontrés précédemment, pour obtenir

$$\limsup_{x\to+\infty} \frac{P(M > x)}{\overline{F}_I(x)} \leq \frac{1}{(a-\epsilon)(1-\gamma)^2}$$

On envoie A en $+\infty$, ce qui fait tendre $\gamma = P(\tau_1 < \infty)$ vers 0 (d'après la loi des grands nombres comme expliqué un peu plus haut) ; on envoie ensuite ϵ vers 0 pour obtenir :

$$\limsup_{x\to+\infty} \frac{P(M > x)}{\overline{F}_I(x)} \leq \frac{1}{a}.$$

□

Chapitre 3

Aire sous la première excursion positive

Sommaire

3.1 Introduction

Dans cette partie, ajoutée au contenu du stage en cours de route, nous allons étudier l'aire sous la première excursion positive d'une marche aléatoire dont les incréments sont à queues lourdes et avec une déviation négative. La preuve d'un des gros résultats de cette partie a grandement posé problème, et n'est donnée ici que partiellement et à titre indicatif, bien malheureusement.

Plus précisement, on note comme précédemment $(X_k)_{k\geq 1}$ les incréments indépendants et identiquement distribués à queue lourde, $\overline{F}$ leur fonction de survie, ainsi que leur espérance commune $\mathbb{E}(X_1) = -a$ où $a > 0$. On note $(S_n)_{n\geq 0}$ notre marche aléatoire, et $\tau = \min\{n \geq 1, S_n \leq 0\}$. On sait déjà que τ est presque sûrement fini, grâce à la loi des grands nombres.

On note A_τ l'aire sous la première excursion, c'est à dire $A_\tau = \displaystyle\sum_{k=1}^{\tau} S_k$. On sait que cette aire est presque sûrement fini car τ l'est.

Nous allons ici étudier $P(A_\tau > x)$ quand x est grand.

Il a déjà été montré dans [3] que si $\overline{F}(x) = x^{-\alpha}L(x)$ avec $\alpha > 1$, L une fonction à variation lente, on a :

$$P(A_\tau > x) \underset{x\to+\infty}{\sim} \mathbb{E}(\tau)\overline{F}(\sqrt{2ax})$$

Ce résultat s'explique assez bien. Heuristiquement, nos incréments étant à queue lourde, on s'attend à ce qu'une grande aire soit le fait d'un grand saut (c'est le principe du grand saut). Pour que l'aire soit le plus grosse possible, on s'attend à ce que ce saut soit au début de la marche (la redescente est ainsi plus longue). Mettons que ce saut prenne la valeur y. D'après la loi des grands nombres, la redescente se fera le long de la droite $y - an$. Ainsi, notre excursion devrait durer environ $\frac{y}{a}$. Ainsi, on s'attend à ce que l'aire soit d'ordre $\frac{y^2}{2a}$ (c'est l'aire d'un triangle rectangle). Ainsi, si notre aire vaut $x = \frac{y^2}{2a}$, on obtient $y = \sqrt{2ax}$.

En conclusion, si on note M_τ le maximum atteint par la marche aléatoire avant le premier passage sous l'axe des abscisses, on a $P(A_\tau > x) \underset{x \to +\infty}{\sim} P(M_\tau > \sqrt{2ax})$

Une preuve rigoureuse de ce résultat est proposée dans [3]. Nous allons proposer une preuve alternative.

3.2 Quelques définitions

Définition 12 (Fonction à variation lente). *Une fonction L est dite à variation lente si pour tout $a > 0$,*

$$\lim_{x \to +\infty} \frac{L(ax)}{L(x)} = 1$$

Définition 13 (Fonction à variation régulière). *Une fonction F est dite à variation régulière si*

$$F(x) = x^{-\alpha} L(x),$$

où L est une fonction à variation lente.

3.3 Ce que nous allons montrer

Théorème 5. *Si $\overline{F}(x) = P(X_1 > x) = x^{-\alpha} L(x)$, où L est à variations lente, avec $\alpha > 1$, et $X_1 \in L^1$, alors, uniformément en $y \in [\epsilon\sqrt{x}; \sqrt{2ax}]$,*

$$P(A_\tau > x, M_\tau > y) \underset{x \to +\infty}{\sim} \mathbb{E}(\tau)\overline{F}(\sqrt{2ax}).$$

Théorème 6. *Avec les mêmes hypothèses sur F,*

$$P(A_\tau > x) \underset{x \to +\infty}{\sim} \mathbb{E}(\tau)\overline{F}(\sqrt{2ax}).$$

3.4 Des premiers résultats

Propriété 22. *Si $\overline{F}$ est à variation régulière, alors F est sous-exponentielle.*

Démonstration. On écrit $\overline{F}(x) = x^{-\alpha}l(x)$ où l est à variations lente. Soit X_1 et X_2 des lois indépendantes et identiquement distribuées de fonction de survie $\overline{F}$. Alors :

$$P(X_1+X_2 > x) = P(X_1 < \frac{x}{2}; X_1+X_2 > x)+P(X_2 < \frac{x}{2}; X_1+X_2 > x)+P(X_1 > \frac{x}{2}; X_2 > \frac{x}{2}).$$

On en déduit :

$$\overline{F}^{*2}(x) = 2\int_0^{\frac{x}{2}} \overline{F}(x-y)dF(y) + \overline{F}^{*2}(\frac{x}{2}).$$

On divise le tout pour x assez grand (afin que le dénominateur ne soit pas nul), et on obtient :

$$\frac{\overline{F}^{*2}(x)}{\overline{F}(x)} = 2\int_0^{\frac{x}{2}} \frac{\overline{F}(x-y)}{\overline{F}(x)}dF(y) + \frac{\overline{F}^{*2}(\frac{x}{2})}{\overline{F}(x)}.$$

Le dernier terme tend vers zéro car $\overline{F}$ est à variations régulières donc

$$\overline{F}(\frac{x}{2})\frac{\overline{F}(\frac{x}{2})}{\overline{F}(x)} = \overline{F}(\frac{x}{2})(\frac{x}{2x})^{-\alpha}\frac{l(\frac{x}{2})}{l(x)} \underset{x\to+\infty}{\sim} (\frac{1}{2})^{\alpha}\overline{F}(\frac{x}{2})$$

qui tend bien vers 0 quand x tend vers $+\infty$.

Quant à l'autre terme, l'intégrande vérifie $\frac{\overline{F(x-y)}}{\overline{F}(x)} \leq \frac{\overline{F}(\frac{x}{2})}{\overline{F}(x)}$ pour $0 \leq y \leq \frac{x}{2}$, qui est intégrable. Donc le théorème de convergence dominée s'applique et l'intégrale tend vers 1. On conclut donc bien que notre loi est sous-exponentielle.

□

Propriété 23. *On suppose que* $\overline{F}(x) = x^{-\alpha}L(x)$ *où* $\alpha > 1$,
alors, pour tout $k \in \mathbb{N}$, *on a :*

1. $P(M_k > y) \underset{y\to+\infty}{\sim} P(S_k > y) \underset{y\to+\infty}{\sim} k\overline{F}(y)$,
2. $P(\max_{n\leq\tau\wedge k} S_n > y) \underset{y\to+\infty}{\sim} \mathbb{E}(\tau\wedge k)\overline{F}(y)$,
3. $P(M_\tau > y) \underset{y\to+\infty}{\sim} \mathbb{E}(\tau)\overline{F}(y)$,
4. $P(\tau > n) \underset{y\to+\infty}{\sim} \mathbb{E}(\tau)\overline{F}(an)$.

Démonstration. La preuve de ce théorème utilise un résultat difficile qui est le résultat de principal de [4], qui nous ramène à montrer que $\int_0^{+\infty} \overline{F}(y)dy < \infty$ et que :

$$\int_0^x \overline{F}(y)\overline{F}(x-y)dy \underset{x\to+\infty}{\sim} 2\overline{F}(x)\int_0^{+\infty} \overline{F}(y)dy$$

On dit alors que $\overline{F} \in \mathcal{S}^*$.

Nous nous contenterons donc de montrer que $\overline{F}$ appartient à $\mathcal{S}^*$; le lecteur curieux pourra aller lire [4] pour en savoir plus.

Dans un premier temps, on remarque que $\overline{F}(x) \underset{x\to+\infty}{\sim} \overline{F}(x+y)$ pour tout $y > 0$ car $\overline{F}$ est à variation régulière donc, en remarquant que :

$$\int_0^x \frac{\overline{F}(y)\overline{F}(x-y)}{\overline{F}(x)}dy = 2\int_0^{\frac{x}{2}} \frac{\overline{F}(y)\overline{F}(x-y)}{\overline{F}(x)}dy,$$

où l'on a effectué un changement de variable. On peut alors appliquer le théorème de convergence dominée, et on obtient alors le résultat.

□

On définit maintenant $\sigma_y = \inf\{n < \tau, S_n > y\}$.

Alors, pour $k \geq 1$, on a :

$$P(\sigma_y = k|M_\tau > y) = \frac{P(\sigma_y = k \cap M_\tau > y)}{P(M_\tau > y)} = \frac{P(\sigma_y = k)}{P(M_\tau > y)},$$

d'où

$$P(\sigma_y = k|M_\tau > y) = \frac{P(\max_{n\leq\tau\wedge k} S_n > y) - P(\max_{n\leq\tau\wedge(k-1)} S_n > y)}{P(M_\tau > y)}.$$

D'après les points 2 et 3 de la 23, on a alors :

Propriété 24. $\lim\limits_{y\to+\infty} P(\sigma_y = k|M_\tau > y) = \dfrac{P(\tau > k-1)}{\mathbb{E}(\tau)}$

Démonstration. Avec les points (2) et (3) de 23, et les calculs précédents, on obtient :

$$\lim_{y\to+\infty} P(\sigma_y = k|M_\tau > y) = \frac{\mathbb{E}(\tau\wedge k) - \mathbb{E}(\tau\wedge(k-1))}{\mathbb{E}(\tau)} = \frac{P(\tau > k-1)}{\mathbb{E}(\tau)},$$

où on a obtenu la dernière égalité avec un téléscopage en écrivant la définition de l'espérance.

□

On notera dorénavant $q_k = \dfrac{P(\tau > k-1)}{\mathbb{E}(\tau)}$. On remarque en particulier que $\displaystyle\sum_{k=1}^{+\infty} q_k = 1$ (c'est la formule de l'espérance $E(X) = \displaystyle\sum_{n=1}^{+\infty} P(X \geq n)$).

3.5 Preuve du théorème 6

La preuve du théorème 6 à partir du théorème 5 est très simple, en particulier, elle est plus simple que la preuve originale de [3]. La voici :

Démonstration. On écrit d'abord :

$$[A_\tau > x] = [A_\tau > x, M_\tau > y] \cup [A_\tau > x, M_\tau \leq y].$$

Occupons nous du seconde terme : On remarque que :

$$[A_\tau > x, M_\tau \leq y] \subset [\tau < \frac{x}{y}].$$

Une démonstration rapide consiste en un joli dessin.
Alors : $P(A_\tau > x, M_\tau > y) \leq P(A_\tau > x) \leq P(A_\tau > x, M_\tau > y) + P(\tau > \frac{x}{y})$.
Or, d'après la 12 ; qui s'applique car F est sous-exponentielle d'après la 22 on montre que :

$$\frac{P(\tau > t)}{\overline{F}(at)} \underset{t\to+\infty}{\sim} E(\tau).$$

On choisit alors $y = \epsilon\sqrt{x}$, et on obtient, grâce au fait que $\overline{F}$ est à variations régulières ; que :

$$P(\tau > \frac{x}{y}) = P(\tau > \frac{\sqrt{x}}{\epsilon}) \underset{x\to+\infty}{\sim} \epsilon^\alpha \mathbb{E}(\tau)\overline{F}(\sqrt{x}).$$

Or, d'après 5,

$$P(A_\tau > x, M_\tau > \epsilon x) \underset{x\to+\infty}{\sim} \mathbb{E}(\tau)\overline{F}(\sqrt{2ax}).$$

On en déduit, en repartant de :

$$P(A_\tau > x, M_\tau > y) \leq P(A_\tau > x) \leq P(A_\tau > x, M_\tau > y) + P(\tau > \frac{x}{y}),$$

et en utilisant nos deux équivalents, que :

$$\mathbb{E}(\tau)\overline{F}(\sqrt{2ax})(1 + o(1)) \leq P(A_\tau > x) \leq \mathbb{E}(\tau)\overline{F}(\sqrt{2ax})(1 + \frac{\epsilon^\alpha}{(2\alpha)^{\frac{\alpha}{2}}}) + o(1)).$$

On conclut alors en faisant tendre ϵ vers 0. $\square$

3.6 Démonstration partielle

Voici le début de la démonstration du 5, qui est elle-même très succinte. Le reste de la preuve pose des problèmes essentiellement techniques et calculatoires... Soit $N \geq 1$. Alors :

$$P(A_\tau > x, M_\tau > y) = \sum_{k=1}^{N} P(A_\tau > x, \sigma_y = k, M_\tau > y) + P(A_\tau > x, \sigma_y > N, M_\tau > y),$$

où l'on a utilisé la loi des probabilités totales.

Occupons nous du dernier terme.

$P(A_\tau > x, \sigma_y > N, M_\tau > y) \leq P(\sigma_y > N, M_\tau > y) = P(M_\tau > y)P(\sigma_y > N | M_\tau > y)$.

Or, $\lim\limits_{y \to +\infty} P(\sigma_y > N | M_\tau > y) = \sum\limits_{k=N+1}^{+\infty} q_k$, d'après la propriété 13. Ainsi, grâce à la troisième proposition de la propriété 12, on a :

$P(M_\tau > y)P(\sigma_y > N | M_\tau > y) \underset{y \to +\infty}{\sim} \mathbb{E}(\tau)\overline{F}(y) \sum\limits_{k=N+1}^{+\infty} q_k = \mathbb{E}(\tau)\overline{F}(y) \frac{1}{\mathbb{E}(\tau)} \sum\limits_{k=N+1}^{+\infty} P(\tau > k-1) = \overline{F}(y) \sum\limits_{k=N+1}^{+\infty} P(\tau > k-1)$.

Ainsi, on obtient donc :$P(A_\tau > x, \sigma_y > N, M_\tau > y) \leq \epsilon_N \overline{F}(y)$ avec $\lim\limits_{N \to +\infty} \epsilon_N = 0$.

Maintenant, on regarde les termes dans la somme.

Soit k un entier naturel fixé. On regarde donc $P(A_\tau > x, \sigma_y = k, M_\tau > y)$. Par une inclusion d'événements, on a déjà :

$$P(A_\tau > x, \sigma_y = k, M_\tau > y) = P(A_\tau > x, \sigma_y = k).$$

On remarque d'abord que $S_j < y$ pour $j < k$. On en déduit que

$$P(A_\tau > x, \sigma_y = k) \leq P\Big(\sum_{j=k}^{\tau-1} S_j > x - (k-1)y, \sigma_y = k\Big)$$

où l'on a grossièrement majoré les $k-1$ premiers termes par y dans la somme.

De la même manière, en les minorant par zéro :

$$P(A_\tau > x, \sigma_y = k) \leq P\Big(\sum_{j=k}^{\tau-1} S_j > x, \sigma_y = k\Big).$$

Maintenant, soit $z > 0$. Par propriété de Markov :

$$P\Big(\sum_{j=k}^{\tau-1} S_j > z, \sigma_y = k\Big) = \int_y^{+\infty} P(S_k \in dv, \sigma_y = k) P(A_\tau > z | S_0 \in dv)$$

En fait, comme on s'intéresse seulement aux termes d'indices supérieurs à k, on s'y ramène en "réindéxant" par propriété de Markov forte k à zéro.

On pose

$$B_v = \left\{ v - \delta v \leq S_n + na \leq v + \delta v, \quad \forall n \leq \frac{v + \delta v}{a} \right\}.$$

Alors la loi des grands nombres montre que $P(B_v | S_1 = v)$ tend vers 1 quand $v \to +\infty$.

On en déduit que

$$P\Big(\sum_{j=k}^{\tau-1} S_j > z, \sigma_y = k\Big) = \int_y^{+\infty} P(S_k \in dv, \sigma_y = k) P(A_\tau > z \cap B_v | S_0 \in dv) + o(P(\sigma_y = k)).$$

Si B_v est réalisé pour v assez grand, alors :

$$\frac{(v-\delta v)^2}{2a} \leq A_\tau \leq \frac{(v+\delta v)^2}{2a}.$$

Un dessin permet de s'en convaincre.

On peut donc conclure que :

$$P(A_\tau > z \cap B_v | S_0 = v) = P(B_v) \quad \text{si} \quad v - \delta v \geq \sqrt{2az}$$

et

$$P(A_\tau > z \cap B_v | S_0 = v) = 0 \quad \text{si} \quad v + \delta v < \sqrt{2az}$$

Donc, pour v assez grand, avec ce qui précède :

$$P\Big(\sum_{j=k}^{\tau-1} S_j > z, \sigma_y = k\Big) \leq \int_{\sqrt{2az}-\delta\sqrt{2az}}^{+\infty} P(S_k \in dv, \sigma_y = k) + o(P(\sigma_y = k)).$$

D'où :

$$P\Big(\sum_{j=k}^{\tau-1} S_j > z, \sigma_y = k\Big) \leq P(S_{\sigma_y} > \sqrt{2az}(1-\delta), \sigma_y = k) + o(P_{\sigma_y = k})$$

De la même manière, on montre que :

$$P\Big(\sum_{j=k}^{\tau-1} S_j > z, \sigma_y = k\Big) \leq \int_{\sqrt{2az}+\delta\sqrt{2az}}^{+\infty} P(S_k \in dv, \sigma_y = k) + o(P(\sigma_y = k)).$$

D'où :

$$P\Big(\sum_{j=k}^{\tau-1} S_j > z, \sigma_y = k\Big) \leq P(S_{\sigma_y} > \sqrt{2az}(1+\delta), \sigma_y = k) + o(P(\sigma_y = k)).$$

où la seconde inégalité est la réécriture de la première.

La suite de la démonstration n'a pas été bien étudiée car présente des difficultés techniques et n'est pas très claire (cette partie a été étudiée en autonomie pendant les vacances d'été, et n'était pas prévue à l'origine dans le stage). Néanmoins, elle est trouvable dans [2].

Bibliographie

[1] *An Introduction to Heavy-tailed and Subexponential Distributions*, Sergey FOSS, Dmitry KORSHUNOV et Stan ZACHARY, 2009

[2] *Tail Asymptotics for the area under the excursion of a random walk with heavy-tailed increments*, Denisov, Perfilev, Wachtel.

[3] *On the integral of the workload process of the single server queue*, Borovkov, A.A., Boxmla, O.J et Palmowski, J. Appl. Probab., 2003.

[4] *The probability of exceeding a high boundary on a random time interval for a heavy tailed random walk*, J. Appl. Probab, Foss, Palmowski, Zachary, 1936-1957, 2005.