Grandes déviations pour des sommes de variables aléatoires indépendantes

Shengjun Zhang, Yulai Huang

4 juin 2023

Résumé

Dans ce mémoire, on présente un phénomène important dans la théorie de probabilité : les grandes déviations. Tout d'abord, on présente le principle de grandes déviations, et sous certains conditions on a le théorème de Cramér, qui estime la probabilité qu'une somme de variables aléatoires i.i.d. soit anormalement grande. Cependant, lorsque la condition de Cramér n'est pas vérifiée (la fonction génératrice des moments est finie dans un voisinage de l'origine), on peut observer des phénomènes différents. Par exemple, on a étudié les grandes déviations sous la condition de sous-exponentialité et on a observé le phénomène de grand saut, c'est-à-dire qu'il existe une variable aléatoire qui prend une grande valeur, tandis que les autres variables aléatoires reste petites. Ensuite, on étudie aussi comment les variables aléatoires sous-exponentielles sont distribuées lorsque la grande déviation du somme est observée. Et puis, sous une autre condition, on a le principe du moins de grands sauts, on voit que c'est un certain nombre de variables aléatoires qui prennent des grandes valeurs. À la fin de ce mémoire, en adaptant la démonstration de l'article de référence [KM22], on a obtenu quelques résultats similaires un peu plus compliqués.

1 Théorème de Cramér

Tout d'abord, on donne une motivation pour les grandes déviations. La référence principale de cette section est le livre [Dem09]. De plus, on omet la preuve des résultats de cette section, vous trouverez les preuves dans le livre [Dem09].

Soit $X_1, X_2, \cdots X_n$ une suite de variables aléatoires gaussiennes i.i.d d'espérance 0 et de variance 1, on note $S_n = \frac{1}{n}\sum_{i=1}^{n} X_i$, alors S_n est une variable aléatoire gaussienne d'espérance 0 et de variance $\frac{1}{n}$, on a donc que pour tout $\delta > 0$, $\mathbb{P}(|S_n| \geq \delta) \underset{n\to\infty}{\to} 0$. De plus, pour tout intervalle A, on a que $\mathbb{P}(\sqrt{n}S_n \in A) \underset{n\to\infty}{\to} \frac{1}{\sqrt{2\pi}}\int_A e^{-\frac{x^2}{2}}dx$.

Notons maintenant que :

$$\mathbb{P}(|S_n| \geq \delta) = 1 - \frac{1}{\sqrt{2\pi}} \int_{-\delta\sqrt{n}}^{\delta\sqrt{n}} e^{-\frac{x^2}{2}} dx$$

et donc $\frac{1}{n} \log \mathbb{P}(|S_n| \geq \delta) \underset{n\to\infty}{\to} -\frac{\delta^2}{2}$.

Alors, on peut dire que la valeur de S_n est environ d'ordre $\frac{1}{\sqrt{n}}$, mais avec une petite probabilité, $|S_n|$ peut être large.

Une question naturelle est l'existence de la limite de $\frac{1}{n} \log \mathbb{P}(|S_n| \geq \delta)$, c'est ce qu'on va discuter maintenant.

1.1 Le principe de grandes déviations

Le principe de grandes déviations caratérise le comportement de la limite d'une famille de mesures de probabilité $\{\mu_\epsilon\}$ sur $(X, \mathcal{B})$ lorsque $\epsilon \to 0$, en termes de la fonction de taux. Dans cette section, X est un espace topologique, avec $\mathcal{B}_X$ le tribu borélien complété.

Définition 1.1. Une fonction de taux est une application inférieurement semicontinue $I : X \to [0, \infty]$. Une bonne fonction de taux est une fonction de taux pour laquelle tous les ensembles de niveau $\Psi_I(\alpha) \stackrel{\text{def}}{=} \{x | I(x) \leq \alpha\}$ sont des compacts de X. Le domaine effectif de I, noté D_I, est l'ensemble de points de X de taux fini, c'est-à-dire, $D_I = \{x | I(x) < \infty\}$.

Définition 1.2. Soit μ_ϵ une famille de mesures de probabilité sur $(X, \mathcal{B})$, on dit que μ_ϵ satisfie le principe de grandes déviations avec une fonction de taux I, si pour tout $\Gamma \in \mathcal{B}$, on a :

$$-\inf_{x\in\Gamma^\circ} I(x) \leq \liminf_{\epsilon\to 0} \epsilon \log \mu_\epsilon(\Gamma) \leq \limsup_{\epsilon\to 0} \epsilon \log \mu_\epsilon(\Gamma) \leq -\inf_{x\in\overline{\Gamma}} I(x).$$

Dans la suite, on dit souvent que μ_ϵ satisfie le PGD pour remplacer la phrase : $\{\mu_\epsilon\}$ satisfie le principe de grandes déviations, avec fonction de taux I.

Dans le cas où $\mathcal{B}_X \subset \mathcal{B}$, le PGD est équivalent à dire que :

1. Pour tout fermée $F \subset X$, $\limsup_{\epsilon\to 0} \epsilon \log \mu_\epsilon(F) \leq -\inf_{x\in F} I(x)$;
2. Pour tout ouvert $G \subset X$, $\liminf_{\epsilon\to 0} \epsilon \log \mu_\epsilon(G) \geq -\inf_{x\in G} I(x)$.

1.2 Théorème de Cramér

Après avoir défini le principe de grandes déviations, on peut énoncer le théorème de Cramér. Tout d'abord, on va donner quelques définitions. Dans la suite, on va supposer que $X = \mathbb{R}^d$.

Soit $X_1, X_2, \cdots, X_n, \cdots$ une suite de variables aléatoires i.i.d à valeurs dans $\mathbb{R}^d$, où la loi de X_1 est $\mu \in M_1(\mathbb{R}^d)$. On note $S_n = \frac{1}{n}\sum_{j=1}^n X_j$, et on note μ_n pour la loi de S_n.

On définit $\Lambda(\lambda) \stackrel{\text{def}}{=} \log M(\lambda) \stackrel{\text{def}}{=} \log \mathbb{E}[e^{\langle \lambda, X_1 \rangle}]$, avec $\langle \lambda, x \rangle \stackrel{\text{def}}{=} \sum_{j=1}^{d} \lambda^j x^j$. De plus, dans $\mathbb{R}^d$, on utilise le norme $|x| = \sqrt{\langle x, x \rangle}$.

Remarquons que $\Lambda(0) = 0$, et pour tout λ, $\Lambda(\lambda) > 0$, tandis qu'on peut avoir $\Lambda(\lambda) = \infty$ pour un certain λ.

Lorsque $\bar{x} \stackrel{\text{def}}{=} \mathbb{E}[X_1]$ existe et est fini, et $\mathbb{E}[|X_1 - \bar{x}|^2] < \infty$, on a que $S_n \underset{n\to\infty}{\to} \bar{x}$ en probabilité, étant donné que

$$\mathbb{E}[|X_1 - \bar{x}|^2] = \frac{1}{n} \sum_{j=1}^{n} \mathbb{E}[|X_j - \bar{x}|^2] = \frac{1}{n^2} \mathbb{E}[|X_1 - \bar{x}|^2] \underset{n\to\infty}{\to} 0.$$

En conséquence, pour tout fermé F tel que $\bar{x} \notin F$, on a que $\mu_n(F) \underset{n\to\infty}{\to} 0$. Le théorème de Cramér caractérise le taux de cette convergence par la fonction de taux suivante :

Définition 1.3. La tranformation de Fenchel-Legendre de $\Lambda(\lambda)$ est

$$\Lambda^*(x) \stackrel{\text{def}}{=} \sup_{\lambda \in \mathbb{R}^d} \{\langle \lambda, x \rangle - \Lambda(\lambda)\}.$$

1.2.1 Le théorème de Cramér pour la dimension 1

Soient $D_\Lambda \stackrel{\text{def}}{=} \{\lambda | \Lambda(\lambda) < \infty\}$ et $D_{\Lambda^*} \stackrel{\text{def}}{=} \{\lambda | \Lambda^*(\lambda) < \infty\}$. Le théorème de Cramér pour $\mathbb{R}$ est vrai même si $\bar{x}$ n'existe pas.

D'abord, on donne quelques propriétés pour la démonstration de ce théorème :

Lemme 1.4. *Pour Λ et Λ^*, on a les propriétés suivantes :*

1. *Λ est une fonction convexe et Λ^* est une fonction de taux convexe ;*
2. *Si $D_\Lambda = \{0\}$, alors Λ^* est zéro. Si $\Lambda(\lambda) < \infty$ pour un certain $\lambda > 0$, alors $\bar{x} < \infty$, de plus, pour $x \geq \bar{x}$, on a que la fonction $\Lambda^*(\lambda) = \sup_{\lambda \geq 0}[\lambda x - \Lambda(\lambda)]$ est non-decroissante pour $x > \bar{x}$. Si $\Lambda(\lambda) < \infty$ pour un certain $\lambda < 0$, alors $\bar{x} > -\infty$, de plus, pour $x \leq \bar{x}$, on a que la fonction $\Lambda^*(\lambda) = \sup_{\lambda \leq 0}[\lambda x - \Lambda(\lambda)]$ est non-croissante pour $x < \bar{x}$. Lorsque $\bar{x}$ est fini, alors $\Lambda^*(\bar{x}) = 0$ et $\inf_{x \in \mathbb{R}} \Lambda^*(\bar{x}) = 0$;*
3. *Λ est différentiable sur D_Λ° avec $\Lambda'(\eta) = \frac{1}{M(\eta)} \mathbb{E}[X_1 e^{\eta X_1}]$. De plus, $\Lambda'(\eta) = y \Rightarrow \Lambda^*(y) = \eta y - \Lambda(\eta)$.*

Avec ces propriétés, on peut démontrer le théorème :

Théorème 1.5. *Soit $X_1, X_2, \cdots, X_n, \cdots$ une suite de variables aléaoires i.i.d à valeurs dans $\mathbb{R}$, notons que $S_n = \frac{1}{n} \sum_{j=1}^{n} X_j$, et que μ_n pour la loi de S_n. Alors, la suite des mesures $\{\mu_n\}$ satisfie le PGD avec la fonction de taux convexe Λ^*, c'est-à-dire :*

1. *Pour tout fermée $F \subset \mathbb{R}$, on a que*
$$\limsup_{n\to\infty} \frac{1}{n} \log \mu_n(F) \leq - \inf_{x\in F} \Gamma^*(F).$$
2. *Pour tout ouvert $G \subset \mathbb{R}$, on a que*
$$\liminf_{n\to\infty} \frac{1}{n} \log \mu_n(G) \geq - \inf_{x\in G} \Gamma^*(G).$$

1.2.2 Le théorème de Cramér pour la dimension d

Dans le cas générale, on suppose que $D_\Lambda = \mathbb{R}^d$, i.e., $\Lambda(\lambda) < \infty$ pour tout λ. En particulier, $\mathbb{E}[|X_1|^2] < \infty$ et $S_n \underset{n\to\infty}{\to} \bar{x}$ en probabilité.

De même, on énonce quelques propriétés de Λ et Λ^* :

Lemme 1.6. *Pour Λ et Λ^*, on a les propriétés suivantes :*

1. *Λ est convexe et différentiale partout, Λ^* est une bonne fonction de taux ;*
2. *$y = \nabla\Lambda(\eta) \Rightarrow \Lambda^*(\eta) = \langle \eta, y\rangle - \Lambda(y)$.*

Avec ces propriétés, on peut démontrer le théorème :

Théorème 1.7. *Soit $X_1, X_2, \cdots, X_n, \cdots$ une suite de variables aléaoires i.i.d à valeurs dans $\mathbb{R}^d$, notons que $S_n = \frac{1}{n}\sum_{j=1}^n X_j$, et que μ_n pour la loi de S_n. Supposons de plus que $D_\Lambda = \mathbb{R}^d$. Alors, la suite des mesures $\{\mu_n\}$ satisfie le PGD sur $\mathbb{R}^d$ avec la fonction de taux convexe Λ^*,*

1.2.3 Le théorème de Gärtner-Ellis

Dans le théorème de Cramér, on suppose que les variables aléatoires sont i.i.d., en fait, on peut faire quelques extensions pour qu'on ait un résultat pour les variables aléatoirs non-i.i.d, c'est ce qu'on va présenter maintenant.

Considérons une suite de variables aléatoires $Z_n \in \mathbb{R}^d$, où la loi de Z_n est μ_n, notons que $\Lambda_n(\lambda) \stackrel{\text{def}}{=} \log \mathbb{E}[e^{\langle \lambda, Z_n\rangle}]$.

Dans toute cette section, on fait les hypothèses suivantes :

1. Pour tout $\lambda \in \mathbb{R}^d$, la limit $\lim_{n\to\infty} \frac{1}{n}\Lambda_n(n\lambda)$ existe dans $[-\infty, +\infty]$, et on note cette limite $\Lambda(\lambda)$;
2. Notons que $D_\Lambda \stackrel{\text{def}}{=} \{\lambda \in \mathbb{R}^d | \Lambda(\lambda) < \infty\}$, alors $0 \in D_\Lambda^\circ$.

Soit Λ^* la transformation de Fenchel-Legendre de Λ, avec $D_{\Lambda^*} \stackrel{\text{def}}{=} \{x \in \mathbb{R}^d | \Lambda^* < \infty\}$. Pour la suite, le but est d'énoncer les conditions sous laquelles μ_n satisfie le PGD avec la fonction de taux Λ^*.

Définition 1.8. On dit que $y \in \mathbb{R}^d$ est un point exposé de Λ^*, s'il existe un $\lambda \in \mathbb{R}^d$ tel que pour tout $x \neq y$, on a que
$$\langle \lambda, y\rangle - \Lambda^*(y) > \langle \lambda, x\rangle - \Lambda^*(x).$$
ici, λ s'appelle un hyperplan exposant.

Définition 1.9. On dit qu'une fonction convexe $\Lambda : \mathbb{R}^d \to]-\infty, \infty]$ est essentiellement lisse, si :

1. D_Λ° est non vide ;
2. Λ est différentiable sur D_Λ° ;
3. Pour toute suite $\{\lambda_n\}$ dans D_Λ° convergeant vers un point du bord de D_Λ°, on a que $\lim_{n\to\infty} |\nabla\Lambda(\lambda_n)| = \infty$.

Le résultat principal est le théorème suivant :

Théorème 1.10. *Pour les notations ci-dessus, on a que :*

1. *Pour tout fermée $F \subset \mathbb{R}^d$, on a que*
$$\limsup_{n\to\infty} \frac{1}{n} \log \mu_n(F) \leq -\inf_{x\in F} \Gamma^*(F).$$
2. *Pour tout ouvert $G \subset \mathbb{R}^d$, on a que*
$$\liminf_{n\to\infty} \frac{1}{n} \log \mu_n(G) \geq -\inf_{x\in G\cap\mathcal{F}} \Gamma^*(G).$$
où $\mathcal{F}$ est l'ensemble des points exposés de Λ^* dont l'hyperplan exposant appartient à D_Λ°.
3. *Si Λ est essentiellement lisse et semicontinue inférieurement, alors μ_n satisfie le PGD avec la fonction de taux Λ^*.*

2 La grande déviation sous la subexponentialité

On peut étudier le comportement de grande déviation même lorsque la condition de Cramér n'est pas valide. Si l'on restreint sur une classe spéciale de distributions (i.e. distribution à localement sous-exponentielle), on peut observer le phénomène de grand saut. C'est à dire que la probabilité que la somme soit dans une intervalle est approximativement égale à la probabilité que l'un d'entre eux soit dans cette intervalle. La référence principale de cette section est l'article [DDS08].

Soit $X_1, X_2, \cdots, X_n, \cdots$ une suite des variables aléaoires i.i.d. à valeur dans $\mathbb{R}$. On note $S_n = X_1 + \cdots + X_n$ et $\Delta =]0, s]$ où $0 < s \leq \infty$. Notons que la notation utilisée ici diffère de celle de la section précédente. Notons F la fonction de répartition de X_1, et l'on utilise aussi la notation $F(x + \Delta) = \mathbb{P}(x < X_1 \leq x + s)$.

Définition 2.1. F est appelée à **localement queue longue** si
$$F(x + y + \Delta) \overset{x\to\infty}{\sim} F(x + \Delta), \ \forall y \in \mathbb{R}.$$
De plus, F est appelée à **localement subexponentielle** si F est à queue longue et on a
$$F^{*2}(x + \Delta) \overset{x\to\infty}{\sim} 2F(x + \Delta).$$

Intuitivement, F est à localement sous-exponentielle si la probabilité que $S_2 = X_1 + X_2$ soit dans la intervalle $x + \Delta$ est approximativement égale à la probabilité que au moins un X_k soit dans la intervalle $x + \Delta$.

Pour décrire le phénomène de grand saut, on doit d'abord introduire quatre suites liées à la distribution à localement sous-exponentielle.

Définition 2.2. On dit qu'une suite de nombres réels positifs $\{b_n\}$ est à **échelle naturelle** si $\{\frac{S_n}{b_n}\}$ est tendue.

On a le lemme suivant :

Lemme 2.3. *Soit $\{b_n\}$ une suite à échelle naturelle. Alors il existe $C > 0$ tel que pour $n \geq 1, c \geq 1, x \geq 0$ on a*

$$\mathbb{P}(S_n > x, X_1 \leq cb_n, \cdots, X_n \leq cb_n) \leq Ce^{-\frac{x}{cb_n}}$$
$$\mathbb{P}(|S_n| > x, |X_1| \leq cb_n, \cdots, |X_n| \leq cb_n) \leq Ce^{-\frac{x}{cb_n}}.$$

Définition 2.4. Soit $\{b_n\}$ une suite convergeant vers l'infini, on dit que $\{I_n\}$ est une suite de $\{b_n\}$**-insensibilité**, si $I_n \gg b_n$ et

$$\sup_{x \geq I_n} \sup_{0 \leq t \leq b_n} \left| \frac{F(x - t + \Delta)}{F(x + \Delta)} - 1 \right| \to 0.$$

Le lemme suivant donne une motivation pour la définition précédente.

Lemme 2.5. *Soit $\{b_n\}$ une suite convergeant vers l'infini. Alors F est à localement queue longue si et seulement s'il existe une suite de $\{b_n\}$-insensibilité pour F.*

La démonstration de la suffisance dans ce lemme est évidente. La démonstration de la nécessité utilise les propriétés des fonctions à variation lente, pour obtenir une certaine forme de l'uniformité.

Définition 2.6. Soit $\{b_n\}$ une suite convergeant vers l'infini, on dit que $\{h_n\}$ est une suite de $\{b_n\}$**-troncature** pour F si

$$\lim_{K \to \infty} \limsup_{n \to \infty} \sup_{x \geq h_n} \frac{n\mathbb{P}(S_2 \in x + \Delta, X_1, X_2 \in] - \infty, -Kb_n[\cup]h_n, \infty[)}{F(x + \Delta)} = 0.$$

On a aussi une description de localement sous-exponentielle en termes de la définition précédente.

Lemme 2.7. *Soit $\{b_n\}$ une suite à échelle naturelle. Alors F est à localement sous-exponentielle si et seulement si F est à localement queue longue et il existe une suite de $\{b_n\}$-troncature pour F.*

Définition 2.8. Étant donné $\{h_n\}$, on dit que $\{J_n\}$ est une suite de $\{h_n\}$**-petit-pas** pour F si

$$\lim_{n \to \infty} \sup_{x \geq J_n} \sup_{z \geq x} \frac{\mathbb{P}(S_n \in z + \Delta, X_1, \cdots, X_n \leq h_n)}{nF(x + \Delta)} = 0.$$

Étant donné toutes les suites définies précédemment, on a le phénomène de grand saut, dont vous trouverez la preuve dans l'article [DDS08].

Théorème 2.9. *Soient $\{b_n\}$ une suite à échelle naturelle, $\{I_n\}$ une suite de $\{b_n\}$-insensibilité, $\{h_n\}$ une suite de $\{b_n\}$-troncature et $\{J_n\}$ une suite de $\{h_n\}$-petit-pas. Si $h_n = O(b_n)$ et $h_n \leq J_n$, on a*

$$\lim_{n\to\infty} \sup_{x \geq I_n + J_n} \left| \frac{\mathbb{P}(S_n \in x + \Delta)}{nF(x+\Delta)} - 1 \right| = 0.$$

3 Distribution conditionée par une grande déviation

Dans cette section, on étudie comment les variables aléatoires sous-exponentielles sont distribuées lorsque la grande déviation du somme est observée. La référence principale de cette section est l'article [AL11].

Supposons que X_n, S_n et Δ soient définies comme la section précédente. On considère la loi de $(X_1, \cdots, X_n)$ conditionée à l'événement de grande déviation de leur somme. Plus précisément, on considère la mesure sur $\mathbb{R}^n$

$$\mu_{n,x}(A) := \mathbb{P}\left((X_1, \cdots, X_n) \in A | S_n \in x + \Delta\right).$$

On note la mesure induite de X_1 par μ. Supposons que μ soit à localement queue longue, et que pour tout $n \geq 1$ on ait

$$\lim_{x\to\infty} \frac{\mathbb{P}(S_n \in x + \Delta)}{n\mu(x+\Delta)} = 1.$$

On peut aussi dit que la distribution μ ici est à localement sous-exponentielle. Notons que la définition de subexponentialité ici est plus précise par rapport à la section précédente. Dans le cas où $\Delta =]0, \infty[$, les deux définitions sont équivalentes [Nor98].

On définit T: $\bigcup \mathbb{R}^n \to \bigcup \mathbb{R}^n$, où $T(X_1, \cdots, X_n)$ échange le premier élément maximum dans ce vecteur et X_n. Intuitivement, lorsque le phénomène de grand saut se produit, les autres variables en dehors de l'élément maximum sont indépendantes de la loi μ. Pour préciser cette intuition, on a les deux théorèmes suivants :

Théorème 3.1. *Il existe une suite $\{q_n\}$ telle que*

$$\lim_{n\to\infty} \sup_{x \geq q_n} \sup_{A \in \mathcal{B}(\mathbb{R}^{n-1})} \left| \mu_{n,x} T^{-1}(A \times \mathbb{R}) - \mu^{n-1}(A) \right| = 0.$$

Théorème 3.2. *On suppose que $\Delta =]0, \infty[$ ici. Alors il existe une suite $\{q_n\}$ telle que*

$$\lim_{n\to\infty} \sup_{x \geq q_n} \|\mu_{n,x} \circ T^{-1} - \mu^{n-1} \times \nu_x\|_{TV} = 0$$

où $\nu_x(\cdot) = \mathbb{P}(X_1 \in \cdot | X_1 > x)$ et $\|\cdot\|_{TV}$ signifie la variation d'une mesure signée.

Avant de nous plonger dans les preuves, on fait une observation. Si μ est une mesure à localement queue longue, $x \mapsto \mu(\log x + \Delta)$ est une fonction à variation lente. Donc les limites

$$\lim_{x\to\infty} \frac{\mu(x+y+\Delta)}{\mu(x+\Delta)} = 1$$

sont uniformes lorsque y est dans un ensemble compact. En conséquence, pour toute suite $\{b_n\}$ il existe une suite $\{l_n\}$ telle que

$$D_n \overset{\text{def}}{=} \sup_{x\geq l_n} \sup_{|y|\leq b_n} \left| \frac{\mu(x+y+\Delta)}{\mu(x+\Delta)} - 1 \right| \to 0,\ n\to\infty.$$

Preuve de théorème 3.1. Soit $X = (X_1, \cdots, X_n)$. Notons $M_X = \max_{1\leq k\leq n} X_k$. On note le premier indice j tel que $X_j = M_X$ par m_X. De plus, on note $\sigma^j X = (X_1, \cdots, X_{j-1}, X_n, X_{j+1} \cdots, X_{n-1}, X_j)$ et $X^j = (X_1, \cdots, X_j)$.

On introduit une suite $\{b_n\}$ à déterminer. Supposons que $\{l_n\}$ et D_n aient été choisis comme ci-dessus. Pour tout ensemble borélien A, on a

$$\begin{aligned}
\mathbb{P}(TX \in A\times\mathbb{R}, S_n \in x+\Delta) &= \sum_{k=1}^n \mathbb{P}(TX\in A\times\mathbb{R}, S_n\in x+\Delta, m_X = k)\\
&\geq \sum_{k=1}^n \mathbb{P}(\sigma^k X \in A\times\mathbb{R}, S_n\in x+\Delta, m_{\sigma^k X} = n)\\
&= n\mathbb{P}(X\in A\times\mathbb{R}, S_n\in x+\Delta, m_X = n)\\
&\geq n\mathbb{P}(X\in A\times\mathbb{R}, S_n \in x+\Delta,\\
&\qquad\qquad |S_{n-1}|\leq b_n, M_{X^{n-1}} \leq x - b_n)\\
&= n\int_{X^{n-1}\in A\bigcap G} \mu(x - S_{n-1} + \Delta) X_*^{n-1}(d\mathbb{P})
\end{aligned}$$

où $G = \{(X_1, \cdots, X_{n-1}) \in \mathbb{R}^{n-1} : |\sum_{k=1}^{n-1} X_k| \leq b_n, M_{(X_1,\cdots,X_{n-1})} \leq x - b_n\}$. Comme μ est à localement queue longue, on a

$$\mu(x - S_{n-1} + \Delta) \geq \mu(x+\Delta)(1 - D_n)$$

pour tout $x \geq l_n$. Donc on a

$$\mu_{n,x} T^{-1}(A\times\mathbb{R}) \geq (1 - D_n)\frac{n\mu(x+\Delta)}{\mathbb{P}(S_n\in x+\Delta)} \mathbb{P}(X^{n-1}\in A\cap G).$$

Alors

$$\begin{aligned}
\mu_{n,x}T^{-1}(A\times\mathbb{R}) - \mu^{n-1}(A) &\geq \left((1-D_n)\frac{n\mu(x+\Delta)}{\mathbb{P}(S_n\in x+\Delta)} - 1\right)\mathbb{P}(X^{n-1}\in A\cap G)\\
&\quad - \mathbb{P}(X^{n-1}\in A - G)\\
&\geq -\left|\frac{n\mu(x+\Delta)}{\mathbb{P}(S_n\in x+\Delta)} - 1\right| - \left|D_n\frac{n\mu(x+\Delta)}{\mathbb{P}(S_n\in x+\Delta)}\right|\\
&\quad - \mathbb{P}(X^{n-1}\notin G).
\end{aligned}$$

On peut remplacer A par A^c et obtenir l'estimation de l'autre côté. Pour conclure, on a

$$\begin{aligned}|\mu_{n,x}T^{-1}(A \times \mathbb{R}) - \mu^{n-1}(A)| \leq &\left|\frac{n\mu(x+\Delta)}{\mathbb{P}(S_n \in x+\Delta)} - 1\right| + \left|D_n \frac{n\mu(x+\Delta)}{\mathbb{P}(S_n \in x+\Delta)}\right| \\ &+ \mathbb{P}(X^{n-1} \notin G)\end{aligned}$$

pour tout A et tout $x \geq l_n$. Comme μ est à localement subexponentielle, on peut choisir $\{r_n\}$ telle que

$$\lim_{n\to\infty} \sup_{x \geq r_n} \left|\frac{n\mu(x+\Delta)}{\mathbb{P}(S_n \in x+\Delta)} - 1\right| = 0.$$

On définit $q_n = \max\{r_n, l_n, 2b_n\}$. Pour tout $x \geq q_n$, on a

$$\begin{aligned}\mathbb{P}(X^{n-1} \notin G) &\leq \mathbb{P}(S_{n-1} > b_n) + \mathbb{P}(M_{X^{n-1}} > x - b_n) \\ &\leq \mathbb{P}(S_{n-1} > b_n) + (n-1)\mathbb{P}(X_1 > b_n).\end{aligned}$$

On peut donc choisir $\{b_n\}$ a priori telle que les quantités ci-dessus tendent vers 0. Alors la suite $\{q_n\}$ satisfie la condition. □

Preuve de théorème 3.2. Par la décomposition de Jordan des mesures signées, on a

$$\|\mu_{n,x} \circ T^{-1} - \mu^{n-1} \times \nu_x\|_{TV} = 2 \sup_U |\mu_{n,x} \circ T^{-1}(U) - \mu^{n-1} \times \nu_x(U)|$$

où U est l'ensemble borélien de $\mathbb{R}^n$. Par des calculs similaires à la preuve précédente, on a

$$\begin{aligned}\mu_{n,x} \circ T^{-1}(U) &= \mathbb{P}(TX \in U | S_n > x) \\ &\geq \frac{n}{\mathbb{P}(S_n > x)} \mathbb{P}(X \in U, S_n > x, m_X = n) \\ &\geq \frac{n}{\mathbb{P}(S_n > x)} \mathbb{P}(X \in U, X_n > x + b_n, M_{X^{n-1}} < x + b_n, |S_{n-1}| < b_n) \\ &= \frac{n\mu(x+\Delta)}{\mathbb{P}(S_n > x)} \mu^{n-1} \times \nu_x(U \cap G \times I).\end{aligned}$$

où $G = \{(X_1, \cdots, X_{n-1}) \in \mathbb{R}^{n-1} : |\sum_{k=1}^{n-1} X_k| \leq b_n, M_{(X_1,\cdots,X_{n-1})} \leq x + b_n\}$ et $I = (x + b_n, \infty)$. Notons que l'égalité finale est vérifiée pour les rectangles, et que l'on peut l'utiliser pour approcher le cas général (car toutes les mesures de probabilité boréliennes sur l'espaces polonais sont les mesures de Radon). On a donc

$$|\mu_{n,x} \circ T^{-1}(U) - \mu^{n-1} \times \nu_x(U)| \leq \left|\frac{n\mu(x+\Delta)}{\mathbb{P}(S_n > x)} - 1\right| + \mu^{n-1} \times \nu_x((G \times I)^c).$$

Comme μ est à localement subexponentielle, on peut choisir $\{r_n\}$ telle que

$$\lim_{n\to\infty} \sup_{x\geq r_n} \left| \frac{n\mu(x+\Delta)}{\mathbb{P}(S_n > x)} - 1 \right| = 0.$$

En outre, on a

$$\mu^{n-1} \times \nu_x((G \times I)^c) \leq \mu^{n-1}(G^c) + \nu_{n,x}(I^c).$$

Pour le premier terme, on a

$$\mu^{n-1}(G^c) \leq \mathbb{P}(|S_{n-1}| > b_n) + n\mathbb{P}(X_1 > x + b_n).$$

Choisissons $\{b_n\}$ telle que $\lim_{n\to\infty} \sup_{x\geq 0} \mu^{n-1}(G^c) = 0$. Pour le deuxième terme, on a

$$\nu_{n,x}(I^c) = 1 - \frac{\mu(x + b_n + \Delta)}{\mu(x+\Delta)}.$$

Comme μ est à queue longue, il existe une suite $\{l_n\}$ telle que

$$\lim_{n\to\infty} \sup_{x\geq l_n} \nu_{n,x}(I^c) = 0.$$

Alors $q_n = \max\{r_n, l_n\}$ satisfie la condition. □

4 Le principe du moins de grands sauts

Dans cette section, nous prouvons un principe de grandes déviations pour la somme de n variables aléatoires indépendantes à queue lourds, qui sont soumises à un bord de coupure mobile à l'emplacement n. Lorsque la somme soit grande à l'échelle n, nous montrerons qu'un nombre fini de variables aléatoires prend des valeurs proches du bord de coupure, tandis que les variables aléatoires restants obéssent encore à la loi des grands nombres.

De plus, ce qu'on fait est qu'on a mondifié le démontration de l'article [KM22] pour obtenir un autre résultat similaire, on va le discuter à la fin de cette section.

Maintenant, on considère un tableau de variables aléatoires, dont les variables aléatoires de la n-ième ligne $(W_k^{(n)}, 1 \leq k \leq n)$ sont i.i.d. et sont distribuées identiquement à une variable aléatoire $W^{(n)}$, pour laquelle on a les hypothèses suivantes :

On a que $0 \leq W^{(n)} \leq n$ et qu'il existe $\alpha > 1$ tel que

1. $(W^{(n)})_n$ est L^p-borné pour tout $1 \leq p < \alpha$,
2. Il existe $\mu \geq 0$ tel que $\mu_n := \mathbb{E}[W^{(n)}] \to \mu$,
3. Il existe $\eta_n \downarrow 0$ et une fonction $h :]0,1[\to [0,\infty[$, qui est continue et intégrable sur des intervalles bornés à partir de zero, tels que pour

tout $\epsilon > 0$, lorsque $\epsilon \leq a_n < b_n$ avec $b_n \leq 1 - \eta_n$ ou $b_n = 1$, et avec $\eta_n \leq b_n - a_n \to 0$, on a

$$\mathbb{P}(a_n n \leq W^{(n)} < b_n) = (1 + o(1))n^{-\alpha} \int_{a_n}^{b_n} h(x)dx,$$

où $o(1)$ est une fonction tendant vers 0, telle que pour tout $\epsilon > 0$ fixé et $\Delta_n \downarrow 0$, la convergence est uniforme pour tous les choix possibles de a_n, b_n avec $b_n - a_n \leq \Delta_n$.

Maintenant, on peut énoncer le théorème principale dans cette section :

Théorème 4.1. *Soient $W_1^{(n)}, \cdots, W_n^{(n)}$ des variables aléatoires indépendantes satisfaisant les hypothèses ci-dessus, on note $S_n = W_1^{(n)} + \cdots W_n^{(n)}$ la somme de ces variables aléatoires. Soit $\rho > 0$ un réel qui n'est pas d'entier et on note $k \in \mathbb{N}$ l'entier tel que $k - 1 < \rho < k$.*

Alors, il existe $\delta_n \downarrow 0$ tels que pour toute suite $\rho_1(n) \uparrow \rho$ et $\rho_2(n) \downarrow \rho$ avec $\rho_2(n) - \rho_1(n) \gg \delta_n$, on a que

$$\mathbb{P}(n(\rho_1(n)+\mu) \leq S_n \leq n(\rho_2(n)+\mu)) = (1+o(1))\binom{n}{k}(\rho_2(n)-\rho_1(n))n^{-\alpha k}K_\rho,$$

où $K_\rho \stackrel{def}{=} h(\rho)$ si $0 < \rho < 1$, et pour $\rho > 1$,

$$K_\rho \stackrel{def}{=} \int_{]0,1[^{k-1}} h(x_1) \cdots h(x_{k-1})h(\rho - (x_1 + \cdots x_{k-1}))dx_1 \cdots dx_k.$$

Dans la suite, on fait une esquisse de la démonstration de ce théorème avec les deux lemmes suivants, on va montrer le deuxième lemme et omet la démonstration du première lemme.

Lemme 4.2. *Il existe $\zeta_n \downarrow 0$ tel que $S_n \stackrel{def}{=} W_1^{(n)} + \cdots W_n^{(n)}$ satisfie :*

$$\mathbb{P}(|S_n - n\mu| > \zeta_n n) \leq \zeta_n.$$

On peut montrer qu'il faut qu'on ait $\zeta_n > n^{-1}$ et $\zeta_n > n^{1-\alpha}$ pour n grand.

Lemme 4.3. *Soit k un entier fixé tel que $k - 1 < \rho < k$ et on note $T_k = T_k(n) \stackrel{def}{=} W_1^{(n)} + \cdots + W_k^{(n)}$. Alors, pour toute suite $\sigma_1(n), \sigma_2(n) \to \rho$ avec $\sigma_2(n) - \sigma_1(n) \gg \eta_n$, on a que*

$$\mathbb{P}(n\sigma_1(n) \leq T_k \leq n\sigma_2(n)) = (1 + o(1))(\sigma_2(n) - \sigma_1(n))n^{-\alpha k}K_\rho.$$

Démonstration. Fixons $\epsilon > 0$. Soit $\Delta_n \downarrow 0$ une suite satisfaisant $\eta_n \ll \Delta_n \ll \sigma_2(n) - \sigma_1(n)$ et $m = \frac{1-\epsilon}{\Delta_n} \in \mathbb{N}$. Fixons une partition $P_n = \{t_0, \cdots, t_m\}$ de

$[\epsilon, 1]$ satisfaisant $|t_i - t_{i+1}| = \Delta_n$ pour tout i. Alors, on a que

$$
\begin{aligned}
&\mathbb{P}(n\sigma_1(n) \le T_k \le n\sigma_2(n)) \\
&\quad \ge \sum_{(l_1,\cdots,l_k)\in[1,m]^k\cap\mathbb{N}^k} \prod_{j=1}^{k} \mathbb{P}(nt_{l_j-1} \le W_j^{(n)} < nt_{l_j}) \mathbb{1}_{\{\sigma_1(n)\le\sum_{j=1}^k t_{l_j-1}, \sum_{j=1}^k t_{l_j} \le \sigma_2(n)\}} \\
&\quad = (1+o(1))n^{-\alpha k} \sum_{\substack{(l_1,\cdots,l_k)\in[1,m]^k\cap\mathbb{N}^k \\ \sum_{j=1}^k t_{l_j-1}\ge\sigma_1(n), \sum_{j=1}^k t_{l_j}\le\sigma_2(n)}} \int_{t_{l_1-1}}^{t_{l_1}} dx_1 \cdots \int_{t_{l_k-1}}^{t_{l_k}} dx_k h(x_1)\cdots h(x_k) \\
&\quad \ge (1+o(1))n^{-\alpha k} \int_{]\epsilon,1[^k} h(x_1)\cdots h(x_k) \mathbb{1}_{\substack{\{x_1+\cdots x_k \le \sigma_2(n)-k\Delta_n \\ x_1+\cdots x_k \ge \sigma_1(n)+k\Delta_n\}}} d(x_1,\cdots,x_k).
\end{aligned}
$$

Comme h est continue sur $]0,1[$ et que $\sigma_2(n)-k\Delta_n$ et $\sigma_1(n)+k\Delta_n$ convergent tous vers $\rho > k-1$ lorsque n tend vers l'infini, on a que

$$
\begin{aligned}
&\int_{]\epsilon,1[^k} h(x_1)\cdots h(x_k) \mathbb{1}_{\substack{\{x_1+\cdots x_k \le \sigma_2(n)-k\Delta_n \\ x_1+\cdots x_k \ge \sigma_1(n)+k\Delta_n\}}} d(x_1,\cdots,x_k) \\
&= \int_{\epsilon}^{1} dx_1 \cdots \int_{\epsilon}^{1} dx_{k-1} h(x_1)\cdots h(x_k) \int_{\sigma_1(n)+k\Delta_n-(x_1+\cdots+x_{k-1})}^{\sigma_2(n)-k\Delta_n-(x_1+\cdots+x_k)} h(x_k)dx_k \\
&= (1+o(1))(\sigma_2(n)-\sigma_1(n)-2k\Delta_n) \\
&\int_{]\epsilon,1[^{k-1}} h(x_1)\cdots h(x_{k-1})h(\rho-(x_1+\cdots+x_{k-1}))d(x_1,\cdots,dx_{k-1}).
\end{aligned}
$$

Comme h est positive, par le théorème de convergence monotone, on sait que cette intégration converge vers K_ρ quand ϵ tend vers 0. De plus, par le choix de Δ_n, on a que $\sigma_2(n)-\sigma_1(n)-2k\Delta_n = (1+o(1))(\sigma_2(n)-\sigma_1(n))$.

Ensuite, soit $0<\epsilon<\rho-(k-1)$ et notons que pour n grand, $n\sigma_1(n) \le T_k \le n\sigma_2(n)$ implique que $W_i^{(n)} \ge \epsilon$ pour tout i, alors, on a que

$$
\begin{aligned}
&\mathbb{P}(n\sigma_1(n) \le T_k \le n\sigma_2(n)) \\
&\quad = \mathbb{P}(n\sigma_1(n) \le T_k \le n\sigma_2(n), \forall 1\le i\le k, W_i^{(n)} \ge \epsilon n) \\
&\quad \le \sum_{(l_1,\cdots,l_k)\in[1,m]^k\cap\mathbb{N}^k} \prod_{j=1}^{k} \mathbb{P}(nt_{l_j-1} \le W_j^{(n)} < nt_{l_j}) \mathbb{1}_{\{\sigma_2(n)\ge\sum_{j=1}^k t_{l_j-1}, \sum_{j=1}^k t_{l_j} \ge \sigma_1(n)\}} \\
&\quad = (1+o(1))n^{-\alpha k} \sum_{\substack{(l_1,\cdots,l_k)\in[1,m]^k\cap\mathbb{N}^k \\ \sum_{j=1}^k t_{l_j-1}\ge\sigma_2(n), \sum_{j=1}^k t_{l_j}\le\sigma_1(n)}} \int_{t_{l_1-1}}^{t_{l_1}} dx_1 \cdots \int_{t_{l_k-1}}^{t_{l_k}} dx_k h(x_1)\cdots h(x_k) \\
&\quad \le (1+o(1))n^{-\alpha k} \int_{]\epsilon,1[^k} h(x_1)\cdots h(x_k) \mathbb{1}_{\substack{\{x_1+\cdots x_k \ge \sigma_1(n)-k\Delta_n \\ x_1+\cdots x_k \le \sigma_2(n)+k\Delta_n\}}} d(x_1,\cdots,x_k).
\end{aligned}
$$

De plus, comme on a que

$$\int_{]\epsilon,1[^k} h(x_1)\cdots h(x_k)\mathbb{1}_{\substack{\{x_1+\cdots x_k\geq\sigma_1(n)-k\Delta_n\\ x_1+\cdots x_k\leq\sigma_2(n)+k\Delta_n\}}} d(x_1,\cdots,x_k)$$
$$\leq \int_{]0,1[^k} h(x_1)\cdots h(x_k)\mathbb{1}_{\substack{\{x_1+\cdots x_k\geq\sigma_1(n)-k\Delta_n\\ x_1+\cdots x_k\leq\sigma_2(n)+k\Delta_n\}}} d(x_1,\cdots,x_k),$$

on peut obtenir le resultat de manière similaire.

En combinant les deux parties ci-dessus, on obtient l'équation du lemme.

□

Avec ces deux lemmes, on peut démontrer notre théorème principal. L'idée de la démonstration est de découper l'événement en quelques événements et de les estimer séparément.

Preuve du théorème 4.1. Avec les notations des lemmes et des hypothèses, on prend une suite $\delta_n \downarrow 0$, tel que $\delta_n \gg \zeta_n \vee \eta_n$. De plus, on peut supposer que $\delta_n > n^{1-\alpha} \vee n^{-1}$ pour tous les $n \geq 0$. Soient $\rho_1(n) \uparrow \rho$ et $\rho_2(n) \downarrow \rho$ deux suites satisfaisant $\rho_2(n) - \rho_1(n) \gg \delta_n$. Fixons $0 < \epsilon < \frac{\rho-(k-1)}{k+\frac{2}{\alpha-1}}$ et on définit l'intervalle

$$I_n = I_n(\rho_1(n), \rho_2(n)) \stackrel{\text{def}}{=} [n(\rho_1(n)+\mu), n(\rho_2(n)+\mu)].$$

De plus, pour $1 \leq m \leq n$, on note $T_m = T_m(n) := W_1^{(n)} + \cdots + W_m^{(n)}$. Soient $(\rho_1^{\subset}(n))_n$ et $(\rho_2^{\subset}(n))_n$ deux suites convergeant vers ρ satisfaisant $\rho_1(n)+\delta_n < \rho_1^{\subset}(n) < \rho_1(n)+2\delta_n$ et $\rho_2(n)-2\delta_n < \rho_2^{\subset}(n) < \rho_2(n)-\delta_n$. Notons que $\rho_2^{\subset}(n) - \rho_1^{\subset}(n) = (1+o(1))(\rho_2(n)-\rho_1(n))$.

Maintenant on décompose $\{S_n \in I_n\}$ en des événements disjoints et étudier leurs probabilités séparément. On définit les événements suivants :

$$\begin{aligned}
\mathcal{E}_1 :& S_n \in I_n, \text{il existe exactement } k \text{ indices } i_1, \cdots i_k, \text{tels que}\\
& W_{i_1}^{(n)} > \epsilon n, \cdots, W_{i_k}^{(n)} > \epsilon n, \text{et } W_{i_1}^{(n)} + \cdots W_{i_k}^{(n)} \in [n\rho_1^{\subset}(n), n\rho_2^{\subset}(n)].\\
\mathcal{E}_2 :& S_n \in I_n, \text{il existe exactement } k \text{ indices } i_1, \cdots i_k, \text{tels que}\\
& W_{i_1}^{(n)} > \epsilon n, \cdots, W_{i_k}^{(n)} > \epsilon n, \text{et } W_{i_1}^{(n)} + \cdots W_{i_k}^{(n)} \notin [n\rho_1^{\subset}(n), n\rho_2^{\subset}(n)].\\
\mathcal{E}_3 :& S_n \in I_n, W_i^{(n)} > \epsilon n \text{ pour au moins } k+1 \text{ indices } i.\\
\mathcal{E}_4^j :& S_n \in I_n, \text{il existe exactement } j \text{ indices } i_1, \cdots i_j, \text{tels que}\\
& W_{i_1}^{(n)} > \epsilon n, \cdots, W_{i_k}^{(n)} > \epsilon n.
\end{aligned}$$

Avec ces événements, on a que

$$\{S_n \in I_n\} = \mathcal{E}_1 \cup \mathcal{E}_2 \cup \mathcal{E}_3 \cup \bigcup_{j=0}^{k-1} \mathcal{E}_4^j.$$

La démonstration consiste en montrer que $\mathcal{E}_1$ est l'événement dominant et que les autres événements ont des probabilités plus petites. De plus, on définit l'événements suivants :

$$\begin{aligned}
\mathcal{E}_{1i_1\cdots i_k} :& S_n \in I_n, W_{i_1}^{(n)} + \cdots W_{i_k}^{(n)} \in [n\rho_1^{\subsetneq}(n), n\rho_2^{\subsetneq}(n)] \text{ et} \\
& W_{i_1}^{(n)} > \epsilon n, \cdots, W_{i_k}^{(n)} > \epsilon n, W_{i_k}^{(n)} \leq \epsilon n \ \forall j \notin \{i_1, \cdots, i_k\}, \\
\mathcal{E}^*_{2i_1\cdots i_k} :& S_n \in I_n, W_{i_1}^{(n)} + \cdots W_{i_k}^{(n)} \notin [n\rho_1^{\subsetneq}(n), n\rho_2^{\subsetneq}(n)] \text{ et} \\
& W_{i_1}^{(n)} > \epsilon n, \cdots, W_{i_k}^{(n)} > \epsilon n, \\
\mathcal{D}_{i_1\cdots i_k i_{k+1}} :& S_n \in I_n, W_{i_1}^{(n)} > \epsilon n, \cdots, W_{i_{k+1}}^{(n)} > \epsilon n.
\end{aligned}$$

On a donc les relations suivantes :

$$\begin{aligned}
\mathbb{P}(\mathcal{E}_1) &= \tbinom{n}{k}\mathbb{P}(\mathcal{E}_{112\cdots k}), \\
\mathbb{P}(\mathcal{E}_2) &\leq \tbinom{n}{k}\mathbb{P}(\mathcal{E}^*_{212\cdots k}), \\
\mathbb{P}(\mathcal{E}_3) &\leq \tbinom{n}{k+1}\mathbb{P}(\mathcal{D}_{12\cdots kk+1}).
\end{aligned}$$

De plus, on peut montrer les équations suivantes dont on omet la démontration.

$$\begin{aligned}
&\mathbb{P}(\mathcal{E}_{112\cdots k}) = (1+o(1))\mathbb{P}(n\rho_1^{\subsetneq}(n) \leq T_k \leq n\rho_2^{\subsetneq}(n)). \\
&\mathbb{P}(\mathcal{E}^*_{212\cdots k}) = o((\rho_2(n) - \rho_1(n))n^{-\alpha k}). \\
&\mathbb{P}(\mathcal{D}_{12\cdots kk+1}) = o((\rho_2(n) - \rho_1(n))n^{-\alpha k - 1}. \\
&\mathbb{P}(\mathcal{E}_4^j) = o((\rho_2(n) - \rho_1(n))n^{-\alpha k}\tbinom{n}{k})).
\end{aligned}$$

Avec le deuxième lemme, on obtient :

$$\mathbb{P}(\mathcal{E}_1) = (1+o(1))\tbinom{n}{k}(\rho_2^{\subsetneq}(n) - \rho_1^{\subsetneq}(n))n^{-\alpha k}K_\rho.$$

Comme on a que $\rho_2^{\subsetneq}(n) - \rho_1^{\subsetneq}(n) = (1+o(1))(\rho_2(n) - \rho_1(n))$, on en déduit :

$$\mathbb{P}(S_n \in I_n) = (1+o(1))\mathbb{P}(\mathcal{E}_1) = (1+o(1))\tbinom{n}{k}(\rho_2(n) - \rho_1(n))n^{-\alpha k}K_\rho.$$

C'est ce qu'on veut démontrer. □

En modifiant un peu la démonstration de ce théorème, on a aboutit à démontrer un résultat similaire, après avoir lu l'article [KM22] donnant le théorème ci-dessus.

Théorème 4.4. *Soient $W_1^{(n)}, \cdots, W_n^{(n)}$ des variables aléatoires indépendantes satisfaisant les mêmes hypothèses, on note $S_n = W_1^{(n)} + \cdots W_n^{(n)}$ la somme de ces variables aléatoires. Soit $\rho > 0$ un réel qui n'est pas d'entier et on note $k \in \mathbb{N}$ l'entier tel que $k - 1 < \rho < k$.*

Alors, on a que

$$\mathbb{P}(S_n \geq n(\rho + \mu)) = (1+o(1))\tbinom{n}{k}n^{-\alpha k}C_\rho,$$

où $C_\rho \overset{def}{=} \int_\rho^1 h(x)dx$ *si* $0 < \rho < 1$, *et pour* $\rho > 1$,

$$C_\rho \overset{def}{=} \int_{]0,1[^{k-1}} h(x_1)\cdots h(x_{k-1})(\int_\rho^k h(x_k-(x_1+\cdots x_{k-1}))dx_k)dx_1\cdots dx_{k-1}.$$

Pour les deux lemmes pour la démontration, il faut seulement modifier le deuxième lemme :

Lemme 4.5. *Soit* k *un entier fixé tel que* $k-1 < \rho < k$ *et on note* $T_k = T_k(n) \overset{def}{=} W_1^{(n)} + \cdots + W_k^{(n)}$. *Alors, pour toute suite* $\sigma_1(n) \to \rho$, *on a que*

$$\mathbb{P}(T_k \geq n\sigma_1(n)) = (1+o(1))n^{-\alpha k}C_\rho.$$

Comme les hypothèses disent que $W^{(n)} \leq n$, on a remarqué que :

$$\mathbb{P}(T_k \geq n\sigma_1(n)) = \mathbb{P}(n\sigma_1(n) \leq T_k \leq kn).$$

Donc, en la modifiant un petit peu, on voit que la démonstration de cet article marche très bien pour notre résultat. Par exemple, dans la démonstration du deuxième lemme, on peut utiliser k pour remplacer $\sigma_2(n)$ et obtenir notre résultat.

Avec une méthode similaire, on a réussit à démontrer une proposition dont la démonstration n'est pas donnée par l'article :

Proposition 4.6. *Soient* $W_1^{(n)}, \cdots, W_n^{(n)}$ *des variables aléatoires indépendantes satisfaisant les hypothèses ci-dessus, on note* $S_n = W_1^{(n)} + \cdots W_n^{(n)}$ *la somme de ces variables aléatoires. Soit* $\rho > 0$ *un réel qui n'est pas d'entier et on note* $k \in \mathbb{N}$ *l'entier tel que* $k-1 < \rho < k$.

Alors, il existe $\delta_n \downarrow 0$ *tels que pour toute suite* $\rho_1(n) \uparrow \rho$ *et* $\rho_2(n) \downarrow \rho$ *avec* $\rho_2(n) - \rho_1(n) \gg \delta_n$, *on a que pour les* $a_1, \cdots, a_k \in]0,1]$,

$$\mathbb{P}(\textit{Pour les } k \textit{ plus grands } W_i^{(n)} : W_{i_1}^{(n)} \leq a_1 n, \cdots, W_{i_k}^{(n)} \leq a_k n | S_n \in I_n)$$
$$\to K_\rho^{-1} C_\rho(a_1, \cdots, a_k).$$

lorsque n *tend vers l'infini, où* $I_n = [n(\rho_1(n)+\mu), n(\rho_2(n)+\mu)]$ *et que* $C_\rho(a_1, \cdots, a_k)$ *est définit par :*

$$\int_0^{a_1} \cdots \int_0^{a_{n-1}} h(x_1)\cdots h(x_{k-1})h(\rho-(x_1+\cdots a_{k-1}))\mathbb{1}_{\{x_1+\cdots x_{k-1} \geq \rho - a_k\}}dx_1\cdots dx_{k-1}.$$

Autrement dit, on a une convergence en loi pour les k plus grands variables aléatoires parmi les $W_i^{(n)}$ et on a donné la loi de limite.

D'une façon similaire, il suffit de montrer le lemme suivant :

Lemme 4.7. *Soit* k *un entier fixé tel que* $k-1 < \rho < k$ *et on note* $T_k = T_k(n) := W_1^{(n)} + \cdots + W_k^{(n)}$. *Alors, pour toute suite* $\sigma_1(n), \sigma_2(n) \to \rho$ *avec* $\sigma_2(n) - \sigma_1(n) \gg \eta_n$, *on a que*

$$\mathbb{P}(n\sigma_1(n) \leq T_k \leq n\sigma_2(n), W_i^{(n)} \leq a_i n \ \forall 1 \leq i \leq k)$$
$$= (1+o(1))(\sigma_2(n) - \sigma_1(n))n^{-\alpha k}C_\rho(a_1, \cdots, a_k).$$

Démonstration. Dans la preuve, on suppose que $\forall i,\ a_i < 1$.

Fixons $\epsilon > 0$. Soit $\Delta_n \downarrow 0$ une suite satisfaisant $\eta_n \ll \Delta_n \ll \sigma_2(n)-\sigma_1(n)$ et $m = \frac{1-\epsilon}{\Delta_n} \in \mathbb{N}$. Fixons une partition $P_n = \{t_0, \cdots, t_m\}$ de $[\epsilon, 1]$ satisfaisant $|t_i - t_{i+1}| = \Delta_n$ pour tout i. Alors, on a que

$$
\begin{aligned}
&\mathbb{P}(n\sigma_1(n) \leq T_k \leq n\sigma_2(n), W_i^{(n)} \leq a_i n \ \forall 1 \leq i \leq k) \\
&\quad \geq \sum_{(l_1,\cdots,l_k)\in[1,m]^k\cap\mathbb{N}^k} \prod_{j=1}^{k} \mathbb{P}(nt_{l_j-1} \leq W_j^{(n)} < nt_{l_j}) \mathbb{1}_{\substack{\{\forall 1\leq j\leq k, t_{l_j}\leq a_j \\ \sigma_1(n)\leq \sum_{j=1}^k t_{l_j-1}, \sum_{j=1}^k t_{l_j}\leq \sigma_2(n)\}}} \\
&\quad = (1+o(1))n^{-\alpha k} \sum_{\substack{(l_1,\cdots,l_k)\in[1,m]^k\cap\mathbb{N}^k, \forall j \ \ t_{l_j}\leq a_j \\ \sum_{j=1}^k t_{l_j-1}\geq\sigma_1(n), \sum_{j=1}^k t_{l_j}\leq\sigma_2(n)}} \int_{t_{l_1-1}}^{t_{l_1}} dx_1 \cdots \int_{t_{l_k-1}}^{t_{l_k}} dx_k h(x_1)\cdots h(x_k) \\
&\quad \geq (1+o(1))n^{-\alpha k} \\
&\quad \int_{\epsilon}^{a_1-\Delta_n} dx_1 \cdots \int_{\epsilon}^{a_{k-1}-\Delta_n} dx_{k-1} \int_{\epsilon}^{1} dx_k h(x_1)\cdots h(x_k) \mathbb{1}_{\substack{\{x_k\leq a_k-\Delta_n \\ x_1+\cdots x_k\leq\sigma_2(n)-k\Delta_n \\ x_1+\cdots x_k\geq\sigma_1(n)+k\Delta_n\}}}.
\end{aligned}
$$

Comme h est continue sur $]0,1[$ et que $\sigma_2(n)-k\Delta_n$ et $\sigma_1(n)+k\Delta_n$ convergent tous vers $\rho > k-1$ lorsque n tend vers l'infini, on a que

$$
\begin{aligned}
&\int_{\epsilon}^{a_1-\Delta_n} dx_1 \cdots \int_{\epsilon}^{a_{k-1}-\Delta_n} dx_{k-1} \int_{\epsilon}^{1} dx_k h(x_1)\cdots h(x_k) \mathbb{1}_{\substack{\{x_k\leq a_k-\Delta_n \\ x_1+\cdots x_k\leq\sigma_2(n)-k\Delta_n \\ x_1+\cdots x_k\geq\sigma_1(n)+k\Delta_n\}}} \\
&\geq \int_{\epsilon}^{a_1-\Delta_n} dx_1 \cdots \int_{\epsilon}^{1} dx_k h(x_1)\cdots h(x_k) \mathbb{1}_{\substack{\{\sigma_2(n)-k\Delta_n-(x_1+\cdots+x_{k-1})\leq a_k-\Delta_n \\ x_1+\cdots x_k\leq\sigma_2(n)-k\Delta_n \\ x_1+\cdots x_k\geq\sigma_1(n)+k\Delta_n\}}} \\
&= \int_{\epsilon}^{a_1-\Delta_n} h(x_1)dx_1 \cdots \int_{\epsilon}^{a_{k-1}-\Delta_n} h(x_{k-1})dx_{k-1} \\
&\int_{\sigma_1(n)+k\Delta_n-(x_1+\cdots+x_k)}^{\sigma_2(n)-k\Delta_n-(x_1+\cdots+x_k)} h(x_k)\mathbb{1}_{\{x_1+\cdots x_{k-1}\geq\sigma_2(n)-a_k-(k-1)\Delta_n\}} dx_k \\
&= (1+o(1))(\sigma_2(n)-\sigma_1(n)-2k\Delta_n) \\
&\int_{\epsilon}^{a_1-\Delta_n} h(x_1)dx_1 \cdots \int_{\epsilon}^{a_{k-1}-\Delta_n} h(x_{k-1})dx_{k-1} h(\rho-(x_1+\cdots+x_{k-1}))\mathbb{1}_{\substack{\{x_1+\cdots x_{k-1}+a_k\geq \\ \sigma_2(n)-(k-1)\Delta_n\}}}.
\end{aligned}
$$

Cette intégration converge vers $C_\rho(a_1,\cdots,a_k)$ quand ϵ tend vers 0. De plus, par le choix de Δ_n, on a que $\sigma_2(n)-\sigma_1(n)-2k\Delta_n = (1+o(1))(\sigma_2(n)-\sigma_1(n))$.

D'une manière pareille, on peut obtenir une borne supérieure. $\square$

Après avoir obtenu la convergence en loi, on a essayé de montrer une convergence en variation totale, mais on n'a pas réussi à le démontrer.

On prend $A \in \mathcal{B}([0,1]^n), B \in \mathcal{B}([0,1])$ et pour $1 \leq i \leq k, A_i \in \mathcal{B}([0,1])$, on définit les mesures suivantes.

$$\nu_n(A) = \frac{\binom{n}{k}\mathbb{P}(S_n \in I_n, (W_1, \cdots, W_n) \in nA, W_1, \cdots, W_k \text{ sont les k plus grands})}{\mathbb{P}(S_n \in I_n)};$$

$$\mu_n(B) = \mathbb{P}(W^{(n)} \in nB);$$

$$\nu(A_1 \times \cdots \times A_k) =$$

$$\int_{A_1} \cdots \int_{A_{k-1}} h(x_1) \cdots h(x_{k-1}) h(\rho - (x_1 + \cdots a_{k-1})) \mathbb{1}_{\{\rho - (x_1 + \cdots x_{k-1}) \in A_k\}} dx_1 \cdots dx_{k-1}.$$

On veut montrer que

$$\lim_{n \to \infty} \|\nu_n - \nu \times \mu_n^{n-k}\|_{TV} = 0.$$

Mais en adaptant la démonstration de l'article [KM22], on peut seulement montrer une convergence uniforme pour des ensembles de la forme $]a_1, b_1[\times \cdots \times]a_k, b_k[\times B$, avec $]a_i, b_i[\subset [0,1]$ et $B \in \mathcal{B}([0,1]^{n-k})$.

Références

[AL11] Inés Armendáriz et Michail Loulakis. “Conditional distribution of heavy tailed random variables on large deviations of their sum”. In : *Stochastic processes and their applications* 121.5 (2011), p. 1138-1147.

[DDS08] Denis Denisov, Antonius Bernardus Dieker et Vsevolod Shneer. “Large deviations for random walks under subexponentiality : the big-jump domain”. In : (2008).

[Dem09] Amir Dembo. *Large deviations techniques and applications.* Springer, 2009.

[KM22] Céline Kerriou et Peter Mörters. “The fewest-big-jumps principle and an application to random graphs”. In : *arXiv preprint arXiv :2206.14627* (2022).

[Nor98] Ragnar Norberg. “P. Embrechts, C. Klüppelberg, T. Mikosch (1997) : Modelling Extremal Events for Insurance and Finance, Springer-Verlag. 645 pp (1.04 kg). ISSN 0172-4568, ISBN 3-540-60931-8.” In : *ASTIN Bulletin : The Journal of the IAA* 28.2 (1998), p. 285-286. doi : 10.2143/AST.28.2.519071.