

Processus stochastiques discrets – Méthode de moments et marche aléatoire

Valentin Féray, Université de Lorraine

octobre 2022 – janvier 2023

Première partie Méthodes de moments

1 Introduction

Par définition, le moment d'ordre k d'une variable aléatoire X est $\mathbb{E}[X^k]$ (il peut être non défini, ou égal à $+\infty$). Considérons une suite de variables aléatoires X_n . Le but de cette partie du cours est de voir quelle information on peut obtenir sur le comportement asymptotique de X_n à partir de d'informations sur les moments. Ce type de techniques est très utile pour l'étude d'objets combinatoires aléatoires (mots, permutations; voir aussi le cours de Pascal Moyal au S10 sur les graphes aléatoires) et nous verrons de nombreuses applications dans ce cours.

2 Méthode du premier moment

Soit \mathbb{N} l'ensemble des entiers naturels.

Theorème 2.1. *Soit X une variable aléatoire à valeurs dans \mathbb{N} . Alors*

$$\mathbb{P}[X = 0] \geq 1 - \mathbb{E}[X].$$

En particulier, si X_n est une suite de v.a. à valeurs dans \mathbb{N} telle que $\mathbb{E}[X_n]$ tend vers 0, alors $\mathbb{P}[X_n = 0]$ tend vers 1.

Première démonstration.

$$\mathbb{E}(X) = \sum_{k \geq 0} k \mathbb{P}[X = k] \geq \sum_{k \geq 1} \mathbb{P}[X = k] = 1 - \mathbb{P}[X = 0]. \quad \square$$

Deuxième démonstration. On commence par rappeler l'inégalité de Markov : si $Y \geq 0$ p.s., on a, pour tout $a > 0$,

$$\mathbb{E}(Y) = \mathbb{E}(Y \mathbf{1}[Y \geq a]) + \mathbb{E}(Y \mathbf{1}[Y < a]) \geq a \mathbb{P}[Y \geq a],$$

ce qui se réécrit $\mathbb{P}[Y \geq a] \leq \frac{\mathbb{E}[Y]}{a}$.

On applique ce résultat à $Y = X$ et $a = 1$. On a

$$\mathbb{P}[X = 0] = 1 - \mathbb{P}[X \geq 1] \geq 1 - \mathbb{E}[X]. \quad \square$$

On utilise souvent ce théorème dans le contexte suivant : X_n compte le nombre d'occurrences d'une certaine structure dans un objet aléatoire, et si $\mathbb{E}[X_n]$ tend vers 0, on sait que, avec probabilité $1 - o(1)$, cette structure n'apparaît pas.

Exemple 2.2. Soit w_n un mot aléatoire uniforme dans $\{0, 1\}^n$, i.e. on tire n variables indépendantes de loi Bern(1/2). Fixons $\varepsilon > 0$, et posons X_n le nombre de blocs de 0 consécutifs de taille $\ell_n := \lfloor (1 + \varepsilon) \log_2(n) \rfloor$. Soit A_i l'évènement $w_i = \dots = w_{i+\ell_n-1} = 0$, i.e. " i est le début d'un bloc de ℓ_n zéros", alors

$$X_n = \sum_{i=1}^{n-\ell_n+1} \mathbf{1}[A_i].$$

On en déduit

$$\mathbb{E}[X_n] = \sum_{i=1}^{n-\ell_n+1} \mathbb{P}[A_i] = \sum_{i=1}^{n-\ell_n+1} 2^{-\ell_n} \leq n 2^{-\lfloor (1+\varepsilon) \log_2(n) \rfloor} \leq 2^{-\varepsilon \log_2(n)+1} = 2 n^{-\varepsilon}.$$

La borne supérieure tend vers 0, donc $\mathbb{E}[X_n]$ tend vers 0. On en déduit que, avec probabilité tendant vers 1, w_n ne contient pas de blocs de zéros de taille $\lfloor (1 + \varepsilon) \log_2(n) \rfloor$.

Nous allons voir maintenant un exemple d'application aux permutations. Pour cela, le lemme suivant sera utile. Notons S_n l'ensemble des permutations de n , i.e. l'ensemble des bijections de $\{1, \dots, n\}$ dans $\{1, \dots, n\}$.

Theorème 2.3. *Soit σ_n une permutation aléatoire uniforme de taille n . Soient $k \geq 1$ un entier et (i_1, \dots, i_k) et (j_1, \dots, j_k) deux listes d'entiers $\leq n$ sans répétitions (mais potentiellement avec une intersection non vide). Alors*

$$\mathbb{P}(\sigma(i_1) = j_1 \wedge \dots \wedge \sigma(i_k) = j_k) = \frac{1}{n(n-1)\dots(n-k+1)}. \quad (1)$$

Démonstration. Fixons une liste (i_1, \dots, i_k) sans répétition. Je prétends que le membre de gauche de (1) ne dépend pas de (j_1, \dots, j_k) . En effet, prenons deux suites (j_1, \dots, j_k) et (j'_1, \dots, j'_k) sans répétitions et choisissons une permutation τ telle que $\tau(j_1) = j'_1, \dots, \tau(j_k) = j'_k$ (note : τ n'est pas aléatoire). Les évènements

$$\sigma(i_1) = j_1 \wedge \dots \wedge \sigma(i_k) = j_k,$$

et

$$\tau \circ \sigma(i_1) = j'_1 \wedge \dots \wedge \tau \circ \sigma(i_k) = j'_k$$

sont alors identiques. Mais si σ est une permutation uniforme dans S_n , alors $\tau \circ \sigma$ est aussi uniforme. On a donc

$$\begin{aligned} \mathbb{P}(\sigma(i_1) = j_1 \wedge \cdots \wedge \sigma(i_k) = j_k) &= \mathbb{P}(\tau \circ \sigma(i_1) = j'_1 \wedge \cdots \wedge \tau \circ \sigma(i_k) = j'_k) \\ &= \mathbb{P}(\sigma(i_1) = j'_1 \wedge \cdots \wedge \sigma(i_k) = j'_k), \end{aligned}$$

ou la dernière égalité vient du fait que σ et $\tau \circ \sigma$ ont la même distribution (distribution uniforme sur S_n). Comme il y a $n(n-1)\dots(n-k+1)$ listes (j_1, \dots, j_k) sans répétition, on a

$$\mathbb{P}(\sigma(i_1) = j_1 \wedge \cdots \wedge \sigma(i_k) = j_k) = \frac{1}{n(n-1)\dots(n-k+1)},$$

comme annoncé. □

Exercice 2.1. On rappelle qu'un point fixe d'une permutation est un entier i tel que $\sigma(i) = i$. Montrer que, avec probabilité $1 - o(1)$, une permutation aléatoire σ_n ne contient pas deux points fixes consécutifs.

Voici un dernier exemple d'application de la méthode du premier moment. Une sous-suite croissante dans une permutation σ est une liste $(\sigma(i_1), \dots, \sigma(i_k))$ telle que

$$i_1 < \cdots < i_k \text{ et } \sigma(i_1) < \cdots < \sigma(i_k).$$

Par exemple, si $\sigma = 47351826$ (j'utilise ici la notation "en ligne" des permutations consistant à écrire $\sigma(1)\sigma(2)\dots$ sans ponctuation), alors nous avons plusieurs sous-suites croissantes de taille 3, parmi lesquelles 126, 358, ..., mais aucune de taille 4 ou plus.

Avant de donner un énoncé, rappelons/introduisons quelques notations :

- $\binom{n}{k} := \frac{n!}{k!(n-k)!}$ est le coefficient binomial "k parmi n" ;
- $\llbracket 1, n \rrbracket := \{1, \dots, n\}$ l'ensemble des entiers de 1 à n ;
- $\left(\llbracket 1, n \rrbracket\right)_k$ l'ensemble des sous-parties de $\llbracket 1, n \rrbracket$ à k éléments.

Proposition 2.4. *Soit σ_n une permutation aléatoire uniforme de taille n, et L_n la longueur de sa plus longue sous-suite croissante. Alors*

$$\lim_{n \rightarrow +\infty} \mathbb{P}[L_n \leq 3\sqrt{n}] = 0.$$

Démonstration. Soit X_n le nombre de sous-suites croissantes de taille $\lfloor 3\sqrt{n} \rfloor$: on a

$$X_n = \sum_{I \in \left(\llbracket 1, n \rrbracket\right)_{\lfloor 3\sqrt{n} \rfloor}} \mathbf{1}[\sigma/I \text{ croissante}],$$

ou la somme est prise sur les sous-ensembles de $\llbracket 1, n \rrbracket$ de taille $\lfloor 3\sqrt{n} \rfloor$. Fixons $I \in \left(\llbracket 1, n \rrbracket\right)_{\lfloor 3\sqrt{n} \rfloor}$. Alors,

$$\mathbb{P}[\sigma/I \text{ croissante}] = \frac{1}{n!} \#\{\sigma \in S_n : \sigma/I \text{ croissante}\}.$$

Le numérateur se calcule ainsi : on commence par choisir l'ensemble des valeurs de σ/I ($\binom{n}{\lfloor 3\sqrt{n} \rfloor}$ choix) ; comme σ/I est croissante, cela détermine entièrement σ/I , et on complète ensuite la permutation (il reste $n - \lfloor 3\sqrt{n} \rfloor$ éléments à envoyer sur $n - \lfloor 3\sqrt{n} \rfloor$ valeurs, soit $(n - \lfloor 3\sqrt{n} \rfloor)!$ choix). Cela donne

$$\mathbb{P}[\sigma/I \text{ croissante}] = \frac{\binom{n}{\lfloor 3\sqrt{n} \rfloor} (n - \lfloor 3\sqrt{n} \rfloor)!}{n!} = \frac{1}{(\lfloor 3\sqrt{n} \rfloor)!}$$

et finalement

$$\mathbb{E}(X_n) = \binom{n}{\lfloor 3\sqrt{n} \rfloor} \frac{1}{(\lfloor 3\sqrt{n} \rfloor)!} = \frac{n!}{(n - \lfloor 3\sqrt{n} \rfloor)! (\lfloor 3\sqrt{n} \rfloor)!^2}$$

En utilisant la formule de Stirling, on obtient

$$\mathbb{E}(X_n) \sim \frac{n^n e^{\lfloor 3\sqrt{n} \rfloor}}{(n - \lfloor 3\sqrt{n} \rfloor)^{n - \lfloor 3\sqrt{n} \rfloor} (9n)^{\lfloor 3\sqrt{n} \rfloor} (6\pi\sqrt{n})}$$

En remarquant que

$$(n - \lfloor 3\sqrt{n} \rfloor)^{n - \lfloor 3\sqrt{n} \rfloor} = n^{n - \lfloor 3\sqrt{n} \rfloor} (1 - 3n^{-1/2})^{n - \lfloor 3\sqrt{n} \rfloor} = n^{n - \lfloor 3\sqrt{n} \rfloor} e^{-\lfloor 3\sqrt{n} \rfloor} O(1),$$

cela implique

$$\mathbb{E}(X_n) \sim \frac{(e^2)^{\lfloor 3\sqrt{n} \rfloor}}{9^{\lfloor 3\sqrt{n} \rfloor} (6\pi\sqrt{n})} \rightarrow 0.$$

On conclut par le théorème 2.1 que $X_n = 0$ avec probabilité $1 - o(1)$. En d'autres termes, avec probabilité $1 - o(1)$ une permutation aléatoire uniforme ne contient pas de sous-suites croissantes de taille $\lfloor 3\sqrt{n} \rfloor$ \square

Vrai ou Faux

1. Si $X \geq 0$ p.s. et $\mathbb{E}[X] = 0$, alors $X = 0$ p.s.
2. Si $w = (w_1, \dots, w_n)$ est un mot aléatoire uniforme dans $\{1, \dots, k\}^n$, alors w_1 et w_2 sont indépendants ;
3. Si σ est une permutation aléatoire dans S_n , alors σ_1 et σ_2 sont indépendants.
4. Si $\mathbb{E}[X_n]$ tend vers $+\infty$, alors $\mathbb{P}[X_n = 0]$ tend vers 0.

Exercice 2.2. Soit X une v.a. à valeurs **réelles** et espérance finie. Montrer que, pour tout $\mu \geq \mathbb{E}[X]$, on a

$$\mathbb{P}[X \leq \mu] > 0.$$

En déduire, sans utiliser le théorème 2.1, que, si X a valeurs dans \mathbb{N} et $\mathbb{E}[X] < 1$, alors $\mathbb{P}[X = 0] > 0$.

Exercice 2.3. Soit v_1, \dots, v_n des vecteurs de norme 1 dans \mathbb{R}^d . Montrer qu'il existe $\varepsilon_1, \dots, \varepsilon_n$ dans $\{-1, 1\}$ telle que

$$\left\| \sum_{i=1}^n \varepsilon_i v_i \right\| \leq \sqrt{n}.$$

(Indice : calculer $\mathbb{E} \left[\left\| \sum_{i=1}^n X_i v_i \right\|^2 \right]$, en prenant les X_i i.i.d. avec $\mathbb{P}[X_1 = -1] = \mathbb{P}[X_1 = 1] = 1/2$, puis utiliser l'exercice 2.2.)

3 Méthode du second moment

La méthode du premier moment sert à prouver qu'une certaine variable X_n vaut 0 avec probabilité $1 - o(1)$. Pour prouver que ce n'est pas le cas, il n'est pas suffisant de regarder l'espérance de X_n , mais il faut regarder la variance. On utilise ensuite le résultat suivant, connu sous le nom d'inégalité de Bienaymé–Tchebychev.

Theorème 3.1. Soit X une variable aléatoire de variance σ^2 , supposée finie. Alors pour tout $\alpha > 0$, on a

$$\mathbb{P}(|X - \mathbb{E}[X]| \geq \alpha) \leq \frac{\sigma^2}{\alpha^2}.$$

Démonstration. On applique l'inégalité de Markov ($\mathbb{P}[Y \geq a] \leq \frac{\mathbb{E}[Y]}{a}$ pour $Y, a \geq 0$) à $Y = (X - \mathbb{E}[X])^2$ et à $a = \alpha^2$. \square

Ce type d'inégalité qui borne la probabilité que X soit loin de sa moyenne (parfois uniquement dans une direction) est appelée "inégalité de concentration". On en verra d'autres dans ce cours.

L'inégalité de Bienaymé–Tchebychev implique en particulier,

$$\mathbb{P}(X = 0) \leq \mathbb{P}(|X - \mathbb{E}[X]| \geq \mathbb{E}[X]) \leq \frac{\text{Var}(X)}{\mathbb{E}[X]^2}.$$

Prenons une suite de v.a. X_n avec un second moment fini.

— Si $\text{Var}(X_n) \leq C \mathbb{E}[X_n]^2$ pour un certain $C < 1$, alors

$$\limsup \mathbb{P}(X_n = 0) \leq C < 1,$$

i.e. $X_n \neq 0$ avec une probabilité asymptotiquement strictement positive.

— Si $\text{Var}(X_n) = o(\mathbb{E}[X_n]^2)$, alors $\lim \mathbb{P}(X_n = 0) = 0$, i.e. $X_n \neq 0$ avec probabilité $1 - o(1)$. Dans ce cas, on peut dire plus. Supposons que $\mathbb{E}[X_n] \neq 0$. Alors $\frac{X_n}{\mathbb{E}[X_n]} \rightarrow 1$ en probabilité. En effet, pour tout $\varepsilon > 0$, on a

$$\mathbb{P} \left[\left| \frac{X_n}{\mathbb{E}[X_n]} - 1 \right| \geq \varepsilon \right] = \mathbb{P} \left[|X_n - \mathbb{E}[X_n]| \geq \varepsilon \mathbb{E}[X_n] \right] \leq \frac{\text{Var}(X_n)}{\varepsilon^2 \mathbb{E}[X_n]^2},$$

et cette borne supérieure tend vers 0 par hypothèse.

Exemple 3.2. On considère comme ci-dessus, w_n un mot aléatoire uniforme dans $\{0, 1\}^n$. Fixons $\varepsilon > 0$, et posons Y_n le nombre de blocs de 0 consécutifs de taille $\ell_n := \lfloor (1 - \varepsilon) \log_2(n) \rfloor$. Soit B_i l'évènement " i est le début d'un bloc de ℓ_n zéros", alors

$$Y_n = \sum_{i=1}^{n-\ell_n+1} \mathbf{1}[B_i].$$

On calcule l'espérance comme dans l'exemple précédent :

$$\mathbb{E}[Y_n] = \sum_{i=1}^{n-\ell_n+1} \mathbb{P}[B_i] = \sum_{i=1}^{n-\ell_n+1} 2^{-\ell_n} = (n-\ell_n+1) 2^{-\lfloor (1-\varepsilon) \log_2(n) \rfloor} \sim 2^{\varepsilon \log_2(n)} = n^\varepsilon.$$

On voit que $\mathbb{E}[Y_n]$ tend vers $+\infty$, on ne peut rien conclure sur $\mathbb{P}[Y_n = 0]$. Regardons le second moment.

$$\text{Var}[Y_n] = \sum_{1 \leq i, j \leq n-\ell_n+1} \text{Cov}(B_i, B_j).$$

Si $|i - j| > \ell_n$, les évènements B_i et B_j concernent des lettres différentes du mot w_n et sont donc indépendants. Si $|i - j| \leq \ell_n$, on a

$$\text{Cov}(B_i, B_j) = \mathbb{E}(B_i B_j) - \mathbb{E}(B_i) \mathbb{E}(B_j) = 2^{-\ell_n - |i-j|} - 2^{-2\ell_n}.$$

En posant $r = i - j$, on obtient

$$\text{Var}[Y_n] \leq \sum_{i=1}^{n-\ell_n+1} \left(\sum_{r=-\ell_n}^{\ell_n} 2^{-\ell_n - |r|} - 2^{-2\ell_n} \right)$$

La somme entre parenthèse est

$$2^{-2\ell_n} \left(2^{\ell_n} - 1 + 2 \sum_{r=1}^{\ell_n} (2^{\ell_n - r} - 1) \right) = 2^{-2\ell_n} (2^{\ell_n} - 1 + 2 * (2^{\ell_n} - 1)) \leq 2^{-\ell_n + 2}.$$

Finalement on obtient

$$\text{Var}[Y_n] \leq n 2^{-\ell_n + 2} \leq 4n^\varepsilon.$$

Comme $\mathbb{E}[Y_n]^2 \sim n^{2\varepsilon}$, on a $\text{Var}[Y_n] = o(\mathbb{E}[Y_n])$ et $Y_n \neq 0$ avec probabilité $1 - o(1)$. En d'autres termes, avec probabilité $1 - o(1)$, w_n contient au moins une suite de 0 consécutifs de taille $(1 - \varepsilon) \log_2(n)$.

Exercice 3.1. Soit X_n le nombre de blocs 01 dans un mot aléatoire uniforme w_n dans $\{0, 1\}^n$. Montrer que, X_n/n tend vers en probabilité vers un nombre c , que l'on calculera.

Exercice 3.2. Soit X_n le nombre de points fixes d'une permutation aléatoire uniforme σ_n . Calculer $\mathbb{E}(X_n)$ et $\mathbb{E}(X_n^2)$.

Exercice 3.3. Soit X une variable aléatoire avec une variance finie, que l'on note σ^2 . Montrer que, pour $u > 0$,

$$\mathbb{P}[X \geq \mathbb{E}(X) + u] \leq \frac{\sigma^2}{u^2 + \sigma^2}.$$

Appliquer ce résultat pour borner $\mathbb{P}[X_n \geq 2]$, où X_n est le nombre de points fixes d'une permutation aléatoire.

(Indice pour la première partie : écrire, pour $t > -\mathbb{E}[X]$,

$$\mathbb{P}[X \geq \mathbb{E}[X] + u] \leq \mathbb{P}[(X + t)^2 \geq (\mathbb{E}[X] + u + t)^2],$$

utiliser l'inégalité de Markov, et trouver la valeur de t optimale.)

4 Moments et convergence en distribution

4.1 Convergence des moments et convergence en distribution

Soit (X_n) une suite de v.a. et une v.a. Z (candidat limite pour X_n), on veut comparer les deux propriétés suivantes

Convergence en distribution $X_n \xrightarrow{d} Z$.

Convergence des moments Pour tout r entier positif, $\mathbb{E}[X_n^r] \rightarrow \mathbb{E}[Z^r]$.

De la convergence en distribution vers convergence des moments.

La convergence en distribution **n'implique pas** la convergence des moments ; par exemple, si X_n est définie par

$$\mathbb{P}(X_n = n) = 1/n = 1 - \mathbb{P}(X_n = 0),$$

alors X_n converge en distribution vers 0, mais $\mathbb{E}(X_n) = 1$ pour tout $n \geq 1$.

Cependant, avec des hypothèses supplémentaires, on peut obtenir la convergence des moments.

Theorème 4.1. *Soit r un entier positif et s un réel avec $r < s$. On suppose que $X_n \xrightarrow{d} Z$ et que $\mathbb{E}[|X_n|^s]$ est une suite bornée. Alors $\mathbb{E}[X_n^r]$ tend vers $\mathbb{E}[Z^r]$.*

En particulier, si tous les moments de X_n sont bornés, alors la convergence en distribution implique la convergence des moments.

Démonstration. Rappelons que la convergence en distribution de X_n vers Z signifie que pour toute fonction f continue bornée, on a $\mathbb{E}[f(X_n)] \rightarrow \mathbb{E}[f(Z)]$. La difficulté est que la fonction $g : x \mapsto x^r$ n'est pas bornée sur \mathbb{R} . On va la remplacer par une approximation bornée. Pour $A > 0$ on pose

$$g_A(x) = \begin{cases} x^r & \text{si } |x| \leq A; \\ (-A)^r & \text{si } x < -A; \\ A^r & \text{si } x > A. \end{cases}$$

Pour x dans \mathbb{R} , on a

$$|x^r - g_A(x)| \leq |x|^r \mathbf{1}_{|x| > A} \leq \frac{|x|^s}{A^{s-r}}.$$

En prenant l'espérance, on obtient

$$|\mathbb{E}[X_n^r] - \mathbb{E}[g_A(X_n)]| \leq \mathbb{E}[|X_n^r - g_A(X_n)|] \leq \frac{\mathbb{E}[|X_n|^s]}{A^{s-r}} \leq \frac{M}{A^{s-r}},$$

où $M = \sup \mathbb{E}[|X_n|^s]$ est fini par hypothèse. Fixons $\varepsilon > 0$ et choisissons A_0 tel que $\frac{M}{A_0^{s-r}} \leq \varepsilon$. On a

$$|\mathbb{E}[X_n^r] - \mathbb{E}[g_{A_0}(X_n)]| \leq \varepsilon.$$

On cherche une approximation similaire pour le moment d'ordre r de la variable limite Z . Par le théorème de représentation de Skorokhod et le lemme de Fatou, on a

$$\mathbb{E}[|Z|^s] \leq \liminf_{n \rightarrow \infty} \mathbb{E}[|X_n|^s] \leq M.$$

À partir de là, le même argument que ci-dessus donne

$$|\mathbb{E}[Z^r] - \mathbb{E}(g_{A_0}(Z))| \leq \varepsilon.$$

En mettant ensemble les deux approximations obtenues, on a

$$|\mathbb{E}[X_n^r] - \mathbb{E}[Z^r]| \leq |\mathbb{E}(g_{A_0}(X_n)) - \mathbb{E}(g_{A_0}(Z))| + 2\varepsilon.$$

Comme $X_n \xrightarrow{d} Z$ et comme g_{A_0} est continue bornée, on a

$$\lim_{n \rightarrow \infty} \mathbb{E}(g_{A_0}(X_n)) - \mathbb{E}(g_{A_0}(Z)) = 0.$$

On en déduit que

$$\limsup_{n \rightarrow \infty} |\mathbb{E}[X_n^r] - \mathbb{E}[Z^r]| \leq 2\varepsilon.$$

Ceci est vrai pour tout $\varepsilon > 0$, donc la suite $\mathbb{E}[X_n^r]$ tend vers $\mathbb{E}[Z^r]$, comme annoncé. \square

Exercice 4.1. Donner une preuve directe du résultat suivant, plus faible que le Théorème 4.1. Soit X_n une suite de v.a. uniformément bornées, alors $\mathbb{E}[X_n^r]$ tend vers $\mathbb{E}[Z^r]$ (« uniformément borné » signifie : $\exists A : \forall n, |X_n| \leq A$ p.s.).

Exercice 4.2. Une suite de v.a. réelles X_n est dite uniformément intégrable (u.i.) si,

$$\lim_{A \rightarrow +\infty} \sup_n \mathbb{E}(|X_n| \mathbf{1}_{|X_n| \geq A}) = 0.$$

1. Montrer que X_n u.i. implique : $\forall \varepsilon > 0, \exists B > 0 : \mathbb{P}[|X_n| \leq B] \geq \varepsilon$ (on dit que X_n est tendue, voir définition ci-dessous).
2. Montrer que X_n u.i. implique $\sup_n \mathbb{E}(|X_n|) < +\infty$.

3. Montrer que $\sup_n \mathbb{E}(|X_n|^{1+\varepsilon}) < +\infty$ pour un certain $\varepsilon > 0$ implique que X_n est u.i.
4. Montrer que si X_n est u.i. et X_n converge vers Z en distribution, alors $\mathbb{E}(X_n) \rightarrow \mathbb{E}(Z) < +\infty$.
5. (*) On suppose $X_n \geq 0$ p.s., pour tout $n \geq 1$. Montrer que X_n converge vers Z en distribution et $\mathbb{E}(X_n) \rightarrow \mathbb{E}(Z) < +\infty$, alors X_n est u.i.

De la convergence des moments à la convergence en distribution

Dans l'autre sens, la convergence des moments implique-t-elle la convergence en distribution? Encore une fois, il faut une hypothèse supplémentaire.

Definition 4.2. Soit X une v.a. avec des moments finis. On dit que X est déterminé par ses moments si, pour toute v.a. Y ,

$$\left(\forall r \geq 1, \mathbb{E}[Y^r] = \mathbb{E}[X^r] \right) \Rightarrow (X \stackrel{d}{=} Y).$$

Theorème 4.3 (Méthode des moments). *Soit X_n et X des v.a. telles que, pour tout $r \geq 1$, on ait $\mathbb{E}(X_n^r) \rightarrow \mathbb{E}(X^r)$ (en particulier, on suppose tous les moments finis). Supposons de plus que X est déterminé par ses moments. Alors $X_n \xrightarrow{d} X$.*

La condition “ X déterminé par ses moments” est clairement nécessaire, donc ce résultat est optimal. Pour prouver le théorème nous aurons besoin de la notion de suite tendue de v.a.

Definition 4.4. Une suite (X_n) de v.a. à valeurs dans un espace E (supposé métrique, complet et séparable) est tendue si pour tout $\varepsilon > 0$, il existe une partie compacte $K \subset E$ tel que $\mathbb{P}(X_n \in K) \geq 1 - \varepsilon$ pour tout n .

Note : K dépend de ε mais pas de n . Si $E = \mathbb{R}$ (cas qui va nous intéresser dans ce chapitre), le compact K peut être choisi de la forme $[-A, A]$, et la condition est équivalente à

$$\forall \varepsilon > 0, \exists A > 0 \text{ t.q. } \mathbb{P}(|X_n| \leq A) \geq 1 - \varepsilon.$$

(Dans ce cas, on parle aussi de suites stochastiquement bornées.) La tension est une sorte de notion de compacité pour des suites de v.a. vis-à-vis de la convergence en distribution. En particulier, on a le résultat suivant (connu sous le nom de théorème de Prokhorov, admis dans ce cours).

Theorème 4.5. *Soit X_n une suite tendue de v.a. Alors il existe une sous-suite $X_{\varphi(n)}$ de X_n qui converge en distribution.*

Corollaire 4.6. *Soit X_n une suite tendue de v.a. et Z un candidat limite. On suppose que toute sous-suite $X_{\varphi(n)}$ de X_n qui converge en distribution converge vers Z . Alors X_n converge vers Z .*

Démonstration. On raisonne par l'absurde. Si X_n ne converge pas vers Z , alors il existe une fonction f continue bornée, $\varepsilon > 0$ et une sous suite $X_{\psi(n)}$ de X_n telles que $|\mathbb{E}(f(X_{\psi(n)}) - \mathbb{E}(f(Z)))| \geq \varepsilon$ pour n assez grand. Comme X_n est tendue, $X_{\psi(n)}$ tendue. Il existe donc une sous-suite $X_{\psi(\psi'(n))}$ de $X_{\psi(n)}$ qui converge en distribution. Par hypothèse, elle converge nécessairement vers Z . Mais c'est impossible car $|\mathbb{E}(f(X_{\psi(n)}) - \mathbb{E}(f(Z)))| \geq \varepsilon$, et donc en particulier $|\mathbb{E}(f(X_{\psi(\psi'(n))}) - \mathbb{E}(f(Z)))| \geq \varepsilon$ pour n assez grand. \square

Démonstration du théorème 4.3. On prouve d'abord que X_n est tendue. Par l'inégalité de Markov

$$\mathbb{P}[|X_n| \geq A] = \mathbb{P}[X_n^2 \geq A^2] \leq \frac{\mathbb{E}(X_n^2)}{A^2}.$$

Or $\mathbb{E}(X_n^2)$ converge (vers $\mathbb{E}(X)$, par hypothèse) et est donc bornée. Le fait que X_n est tendu en découle immédiatement.

Considérons maintenant une sous-suite $X_{\varphi(n)}$ de X_n qui converge en distribution vers Y . On veut prouver $Y \stackrel{d}{=} X$. Pour tout entier pair $s > 0$, la suite $\mathbb{E}[|X_{\varphi(n)}|^s] = \mathbb{E}[X_{\varphi(n)}^s]$ est borné (car elle converge vers $\mathbb{E}(X^s)$ par hypothèse). Donc, d'après le théorème 4.1, les moments de $X_{\varphi(n)}$ converge vers ceux de Y . Mais comme $X_{\varphi(n)}$ est une sous-suite de X_n , ils convergent aussi vers les moments de X . Par unicité des moments, on en déduit que Y et X ont les même moments, i.e. $\mathbb{E}[Y^r] = \mathbb{E}[X^r]$ pour tout $r \geq 1$. Comme X est déterminé par ses moments, $Y \stackrel{d}{=} X$. On a donc montré que toute sous-suite de X_n qui converge en distribution converge nécessairement vers X .

D'après le corollaire 4.6 (rappelons que X_n est tendue), on conclut que $X_n \xrightarrow{d} X$. \square

Vrai ou Faux

1. Soit X_n une suite de v.a. positives telle que X_n tend vers X en distribution. On suppose que $\mathbb{E}[X] = +\infty$. Alors $\mathbb{E}[X_n]$ tend vers $+\infty$.
2. Soit a_n une suite de nombres réels ayant une unique valeur d'adhérence ℓ . Alors a_n converge vers ℓ .
3. Si X_n converge en distribution vers Z , alors $\mathbb{E}(f(X_n))$ tend vers $\mathbb{E}(f(Z))$ pour toute fonction continue f .

Exercice 4.3. Montrer que, si X_n converge en distribution vers X , alors X_n est tendue.

Quand une mesure est-elle déterminée par ses moments ?

Il y a une condition suffisante facile à vérifier. Commençons par un lemme.

Lemme 4.7. *Soit X une v.a. aléatoire réelle avec des moments finis. Alors les conditions suivantes sont équivalentes :*

- (i) *il existe $C > 0$ tel que $|\mathbb{E}(X^r)| \leq C^r r!$ pour tout entier $r \geq 1$;*
- (ii) *il existe $C > 0$ tel que $\mathbb{E}(|X|^r) \leq C^r r!$ pour tout entier $r \geq 1$;*

(iii) il existe $u > 0$ tel que $\mathbb{E}(e^{u|X|}) < +\infty$;

(iv) il existe $\varepsilon > 0$ tel que $\mathbb{E}(e^{uX}) < +\infty$ pour tout $|u| < \varepsilon$;

Démonstration. (ii) implique (i) est trivial. Montrons (i) \Rightarrow (ii). Il suffit de considérer le cas où $r = 2k - 1$ est impair ($k \geq 2$). En utilisant l'inégalité $|x|^{2k-1} \leq 1 + x^{2k}$, valable pour tout x réel, on a

$$E(|X|^{2k-1}) \leq 1 + \mathbb{E}(X^{2k}) \leq 1 + C^{2k}(2k)! \leq (C')^{2k-1}(2k-1)!,$$

pour C' bien choisi.

Pour (ii) \Leftrightarrow (iii), on utilise le théorème de Fubini–Tonelli pour écrire

$$\mathbb{E}(e^{u|X|}) = \mathbb{E} \left(\sum_{k \geq 0} \frac{u^k |X|^k}{k!} \right) = \sum_{k \geq 0} \frac{u^k}{k!} \mathbb{E}(|X|^k).$$

L'implication (ii) \Rightarrow (iii) se prouve alors en choisissant $u < 1/C$ et en bornant $\frac{u^k}{k!} \mathbb{E}(|X|^k)$ par la série géométrique convergente $(uC)^k$. Dans le sens (iii) \Rightarrow (ii), on utilise le fait que le terme général d'une série convergente tend vers 0, et est donc borné. On a donc $\frac{u^k}{k!} \mathbb{E}(|X|^k) \leq A$ pour $u > 0$ et $A > 0$ bien choisis, ce qui implique (ii).

(iii) \Leftrightarrow (iv) est facile en écrivant $e^{uX} \leq e^{\varepsilon|X|}$ pour $|u| \leq \varepsilon$, et, dans l'autre sens, $e^{u|X|} \leq e^{uX} + e^{-uX}$. \square

Theorème 4.8. *Soit X une v.a. aléatoire réelle avec des moments finis. Si les conditions (i)-(iv) du lemme sont vérifiées, alors X est déterminé par ses moments.*

Première démonstration (sans analyse complexe). Soit Y une v.a. avec les mêmes moments que X , i.e. pour tout $r \geq 1$, on suppose que $\alpha_r := \mathbb{E}(X^r) = \mathbb{E}(Y^r) < \infty$. L'idée de la preuve consiste à exprimer la fonction caractéristique de X et Y en fonction de leurs moments, pour utiliser le fait que la fonction caractéristique détermine la loi. Informellement, on a

$$e^{itX} \sim \sum_{k=0}^{\infty} (itX)^k / k!$$

et donc, en prenant l'espérance :

$$\varphi_X(t) \sim \sum_{k=0}^{\infty} (it)^k \alpha_k / k!$$

Rigoureusement, on peut contrôler le terme d'erreur par une expansion de Taylor :

$$\left| e^{itX} - \sum_{k=0}^N (itX)^k / k! \right| \leq \frac{|tX|^{N+1}}{(N+1)!}$$

En prenant l'espérance

$$\left| \varphi_X(t) - \sum_{k=0}^N (it)^k \alpha_k / k! \right| \leq \frac{|t|^{N+1}}{(N+1)!} \mathbb{E}[|X|^{N+1}] \leq (C|t|)^{N+1},$$

où on a utilisé que X vérifie la condition (ii) ci-dessus. Pour $|t| < \rho := 1/C$, le membre de droite tend vers 0 et

$$\varphi_X(t) = \sum_{k=0}^{\infty} (it)^k \alpha_k / k!.$$

On peut faire le même raisonnement avec Y (comme Y a les mêmes moments que X , elle vérifie la condition (i) ci-dessus, et donc aussi la condition (ii), même si les moments de $|Y|$ ne sont a priori pas les mêmes que ceux de $|X|$). On obtient que $\varphi_X(t) = \varphi_Y(t)$ pour $|t| \leq \rho$. Ce n'est malheureusement pas suffisant pour conclure que $X \stackrel{d}{=} Y$, il faut étendre l'égalité à \mathbb{R} tout entier.

Écrivons maintenant, pour t_0 dans \mathbb{R} fixé,

$$\left| e^{it_0 X} \left(e^{itX} - \sum_{k=0}^N (itX)^k / k! \right) \right| \leq \frac{|tX|^{N+1}}{(N+1)!}$$

de sorte que, pour $|t| < \rho$

$$\varphi_X(t_0 + t) = \sum_{k=0}^{\infty} \mathbb{E}(e^{it_0 X} X^k) (it)^k / k!. \quad (2)$$

Notons $c_{t_0, k}(X)$ l'espérance dans le membre de droite, et adoptons une notation similaire avec X remplacé par Y . L'équation (3) est un développement de Taylor de φ_X autour de t_0 : les $c_{t_0, k}(X)$ (resp. $c_{t_0, k}(Y)$) sont donc les dérivées successives de φ_X (resp. φ_Y) en t_0 , divisées par i^k . Pour $|t_0| < \rho$, on $\varphi_X \equiv \varphi_Y$ localement autour de t_0 , et donc

$$c_{t_0, k}(X) = c_{t_0, k}(Y)$$

En revenant à (2) et à l'égalité analogue pour Y , on en déduit que pour $|t_0| < \rho$ et $|t| < \rho$,

$$\varphi_X(t_0 + t) = \varphi_Y(t_0 + t).$$

Autrement dit, φ_X et φ_Y coïncident sur $(-2\rho, 2\rho)$. En itérant le raisonnement, on montre que φ_X et φ_Y coïncident sur $(-3\rho, 3\rho)$, puis sur $(-4\rho, 4\rho)$, etc. Finalement, les fonctions caractéristiques φ_X et φ_Y coïncident sur toute la droite réelle, ce qui prouve $X \stackrel{d}{=} Y$. \square

Deuxième démonstration (avec analyse complexe). Comme X vérifie la condition (ii) ci-dessus, il existe C tel que $\mathbb{E}[|X|^k] < C^k k!$. Notons que pour t_0 dans \mathbb{R} et z dans \mathbb{C} , on a

$$e^{(z+it_0)X} = e^{it_0 X} e^{zX} = \sum_{k \geq 0} \frac{z^k (e^{it_0 X} X^k)}{k!}.$$

Lorsque $|z| < 1/C$ la série

$$\sum_{k \geq 0} \mathbb{E} \left[\left| \frac{z^k (e^{it_0 X} X^k)}{k!} \right| \right] \leq \sum_{k \geq 0} \frac{|z|^k \mathbb{E}[|X|^k]}{k!} \leq \sum_{k \geq 0} (C|z|)^k$$

est sommable et on peut écrire

$$\mathbb{E}[e^{(z+it_0)X}] = \sum_{k \geq 0} \frac{z^k \mathbb{E}[e^{it_0 X} X^k]}{k!}. \quad (3)$$

Cela montre que la fonction $\omega \mapsto \mathbb{E}[e^{\omega X}]$ est analytique sur le disque de centre it_0 et de rayon $1/C$. Ceci étant vrai pour tout t_0 réel, la fonction $\omega \mapsto \mathbb{E}[e^{\omega X}]$ est analytique sur la bande $\{|\Re \omega| < 1/C\}$.

Soit Y une v.a. avec les mêmes moments que X , i.e. pour tout $r \geq 1$, on suppose que $\alpha_r := \mathbb{E}(X^r) = \mathbb{E}(Y^r) < \infty$. Comme X vérifie la condition (i) ci-dessus, Y aussi. On en déduit que Y vérifie aussi la condition (ii) du lemme, i.e. il existe C' tel que $\mathbb{E}[|Y|^k] < (C')^k k!$. Le raisonnement ci-dessus implique que $\omega \mapsto \mathbb{E}[e^{\omega Y}]$ est analytique sur la bande $\{|\Re \omega| < 1/C'\}$.

De plus les développements en série de $\mathbb{E}[e^{\omega X}]$ et $\mathbb{E}[e^{\omega Y}]$ autour de 0 sont donnés par l'équation (3) pour $t_0 = 0$ (et son équivalent pour Y) :

$$\mathbb{E}[e^{zX}] = \sum_{k \geq 0} \frac{z^k \mathbb{E}[X^k]}{k!}, \quad \mathbb{E}[e^{zY}] = \sum_{k \geq 0} \frac{z^k \mathbb{E}[Y^k]}{k!}.$$

Les moments $\mathbb{E}[X^k]$ et $\mathbb{E}[Y^k]$ étant égaux, les fonctions $\mathbb{E}[e^{\omega X}]$ et $\mathbb{E}[e^{\omega Y}]$ coïncident sur un disque autour de 0. Comme ce sont des fonctions analytiques, elles coïncident sur l'intersection de leur domaine qui est connexe. En particulier elle coïncide sur $i\mathbb{R}$, i.e. $\mathbb{E}[e^{itX}] = \mathbb{E}[e^{itY}]$ pour tout t réel. Comme la fonction caractéristique d'une v.a. réelle détermine sa loi, on en déduit que X et Y ont la même loi, ce qui conclut la preuve du théorème. \square

Exercice 4.4. Montrer que les v.a. Gaussiennes, de Poisson, exponentielles, et géométriques vérifient le critère ci-dessus, quelque soient leur paramètres, ainsi que toutes les lois bornées.

Si Z est une Gaussienne centrée réduite, à quelle condition Z^k vérifie-t-elle le critère ci-dessus ?

Notons que la méthode des moments ne s'applique pas aux lois dont certains moments sont infinis, comme la loi de Cauchy $\mathbb{P}(dx) = \frac{1}{\pi} \frac{1}{x^2+1}$ (dans ce cas, tous les moments sont infinis). Il existe aussi des mesures dont tous les moments sont finis, et qui ne sont pas déterminés par les moments. L'exercice suivant montre que $X = e^Z$, où Z est une loi normale, en est un exemple.

Exercice 4.5. Soit

$$f(x) = \begin{cases} \frac{1}{\sqrt{2\pi}} \frac{1}{x} e^{-(\log(x))^2/2} & \text{if } x > 0, \\ 0 & \text{if } x \leq 0, \end{cases}$$

$$g(x) = f(x) (1 + \sin(2\pi \log(x)))$$

1. Montrer que si Z est une loi normale standard, alors f est la densité de $X = e^Z$.
2. Montrer que f et g sont des densités et que les variables aléatoires correspondantes X et Y ont tous leurs moments finis.
3. Montrer que pour tout entier $k \geq 1$, on a $\mathbb{E}[X^k] = \mathbb{E}[Y^k]$.
(Indice : écrire $\mathbb{E}[X^k] - \mathbb{E}[Y^k]$ comme une intégrale et poser $\log(x) = s+k$.)

Exercice 4.6. Soit X une v.a. bornée. On veut montrer directement que X est déterminé par ses moments. Soit Y une v.a. ayant les mêmes moments que X .

1. Montrer que $(\mathbb{E}[|Y|^k]^{1/k})_{k \geq 1}$ est une suite bornée. En déduire que Y est essentiellement bornée (i.e. il existe A tel que $|Y| \leq A$ p.s.).
2. En utilisant le théorème de Stone–Weierstrass, montrer que pour toute fonction f continue à support compact, $\mathbb{E}[f(X)] = \mathbb{E}[f(Y)]$.
3. Conclure en utilisant le théorème de représentation de Riesz (aussi appelé de Riesz–Markov).

4.2 Quelques applications

4.2.1 Théorème centrale limite

Comme première application de la méthode des moments, donnons une preuve du théorème centrale limite dans le cas de v.a. dont tous les moments sont finis.

On commence par calculer les moments d’une variance Gaussienne.

Théorème 4.9. *Soit Z une v.a. Gaussienne centrée réduite. Alors*

$$\mathbb{E}[Z^k] = \begin{cases} 1 \cdot 3 \cdots (k-1) & \text{si } k \text{ est pair;} \\ 0 & \text{sinon.} \end{cases}$$

Démonstration. On a

$$\mathbb{E}[Z^k] = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{+\infty} x^k e^{-x^2/2} dx.$$

Pour k impair, l’intégrale vaut 0 par symétrie. Pour k pair, on fait une intégration par partie : on pose $u = x^{k-1}$, ce qui implique $du = (k-1)x^{k-2}dx$, et $dv = xe^{-x^2/2}dx$, soit $v = -e^{-x^2/2}$. On trouve

$$\int_{-\infty}^{+\infty} x^k e^{-x^2/2} dx = (k-1) \int_{-\infty}^{+\infty} x^{k-2} e^{-x^2/2} dx,$$

(les termes aux bords sont nuls), ce qui implique $\mathbb{E}[Z^k] = (k-1)\mathbb{E}[Z^{k-2}]$. En utilisant l’égalité triviale $\mathbb{E}[Z^0] = 1$, une récurrence immédiate prouve que, pour k pair,

$$\mathbb{E}[Z^k] = (k-1)\mathbb{E}[Z^{k-2}] = (k-1)(k-3)\mathbb{E}[Z^{k-4}] = \cdots = (k-1)(k-3) \cdots 3 \cdot 1. \quad \square$$

Theorem 4.10. Soit X_1, X_2, \dots des v.a. i.i.d. telles que, pour tout $r \geq 1$, $\mathbb{E}[|X_1|^r] < +\infty$. Posons $\sigma^2 = E[X_1^2]$ et $S_n = \sum_{i=1}^n X_i$. Alors $\tilde{S}_n := \frac{S_n - \mathbb{E}[S_n]}{\sigma\sqrt{n}}$ converge en distribution vers une loi Gaussienne centrée réduite Z .

Comme vous l'avez probablement déjà vu dans un cours précédent, le théorème est valide sous l'hypothèse plus faible $\mathbb{E}[X_1^2] < +\infty$.

Démonstration. On écrit

$$\mathbb{E}[\tilde{S}_n^k] = \frac{1}{\sigma^k n^{k/2}} \sum_{i_1, \dots, i_k} \mathbb{E} \left(\prod_{j=1}^k (X_{i_j} - \mathbb{E}(X_{i_j})) \right). \quad (4)$$

À une suite $\mathbf{i} = (i_1, \dots, i_k)$, on associe de manière unique une partition $\pi_{\mathbf{i}}$ de l'ensemble $\llbracket 1, k \rrbracket$ par la condition $i_j = i_h$ si et seulement si j et h sont dans la même part de π . De plus, pour une liste \mathbf{i} et une partition π , on écrit $\mathbf{i} \in \mathcal{I}_\pi$ si $\pi_{\mathbf{i}} = \pi$. En notant m_1, \dots, m_ℓ les tailles des parts de π , le caractère i.i.d. des X_i implique que, si $\mathbf{i} \in \mathcal{I}_\pi$,

$$\mathbb{E} \left(\prod_{j=1}^k (X_{i_j} - \mathbb{E}(X_{i_j})) \right) = \prod_{t=1}^{\ell} \mathbb{E} (X_1 - \mathbb{E}(X_1))^{m_t}. \quad (5)$$

En particulier cela ne dépend que de π . On peut réécrire (4) sous la forme

$$\mathbb{E}[\tilde{S}_n^k] = \frac{1}{\sigma^k n^{k/2}} \sum_{\pi} \left[\#\{\mathbf{i} : \mathbf{i} \in \mathcal{I}_\pi\} \prod_{t=1}^{\ell} \mathbb{E} (X_1 - \mathbb{E}(X_1))^{m_t} \right]. \quad (6)$$

Le comportement de la quantité entre crochets dépend de la combinatoire de la partition π .

- Si π contient un singleton, i.e. $m_t = 1$ pour un certain t , alors

$$\mathbb{E} (X_1 - \mathbb{E}(X_1))^{m_t} = 0$$

et le terme correspondant à π dans (6) vaut 0.

- Le nombre de suites \mathbf{i} telle que $\mathbf{i} \in \mathcal{I}_\pi$ est $n(n-1)\dots(n-\ell+1) = O(n^\ell)$. De plus on peut borner le produit des espérances en utilisant la croissance des normes L^p :

$$\begin{aligned} \left| \prod_{t=1}^{\ell} \mathbb{E} [(X_1 - \mathbb{E}(X_1))^{m_t}] \right| &\leq \prod_{t=1}^{\ell} \|X_1 - \mathbb{E}(X_1)\|_{m_t}^{m_t} \\ &\leq \prod_{t=1}^{\ell} \|X_1 - \mathbb{E}(X_1)\|_k^{m_t} = \|X_1 - \mathbb{E}(X_1)\|_k^{\sum_{t=1}^{\ell} m_t} = \|X_1 - \mathbb{E}(X_1)\|_k^k. \end{aligned}$$

Cette borne supérieure est indépendante de π et de n , le terme associé à π dans (6) est donc $O(n^\ell)$. Après division par $n^{k/2}$, les contributions des termes avec $\ell < k/2$ sont donc négligeables.

Il suffit donc de considérer les partitions sans singletons avec au moins $k/2$ parts, c'est à dire les partitions de $\llbracket 1, k \rrbracket$ en paires. Si k est impair, il n'y en a pas et on a $\mathbb{E}[\tilde{S}_n^k] = o(1)$. Si k est pair, il y en a $(k-1)(k-3)\cdots 3\cdot 1$, et la contribution de chacune de ces partitions est

$$n(n-1)\cdots(n-\frac{k}{2}+1)\prod_{t=1}^{k/2}\mathbb{E}(X_1-\mathbb{E}(X_1))^2\sim n^{k/2}\sigma^k.$$

On obtient donc

$$\lim_{n\rightarrow+\infty}\mathbb{E}[\tilde{S}_n^k]=(k-1)(k-3)\cdots 3\cdot 1.$$

Dans les deux cas (k pair et k impair), on a prouvé

$$\lim_{n\rightarrow+\infty}\mathbb{E}[\tilde{S}_n^k]=\mathbb{E}[Z^k],$$

ce qui implique $\tilde{S}_n \rightarrow Z$, en distribution, par la méthode des moments. \square

Exercice 4.7. Adapter la preuve ci-dessus pour montrer le résultat suivant. Soit X_1, \dots, X_n des v.a. i.i.d. centrées telles que, pour tout $r \geq 1$, $\mathbb{E}[|X_1|^r] < +\infty$. Posons $\sigma^2 = E[X_1^2]$ et $T_n = \sum_{i=1}^{n-1} X_i X_{i+1}$. Alors $\tilde{T}_n := \frac{T_n}{\sigma^2 \sqrt{n}}$ converge en distribution vers une loi Gaussienne centrée réduite Z .

4.2.2 Exemple de convergence vers une loi de Poisson

Au lieu de montrer la convergence des moments, dans le cas d'une loi de Poisson, c'est souvent plus facile de montrer la convergence des moments factoriels, définis par $\mathbb{E}[X(X-1)\cdots(X-k+1)]$.

Exercice 4.8. Soit X_n et Y des v.a. réelles, dont tous les moments sont finis.

1. Montrer que les deux assertions suivantes sont équivalentes

- (i) Pour tout k , $\mathbb{E}[X_n^k] \rightarrow \mathbb{E}[Y^k]$.
- (ii) Pour tout k ,

$$\mathbb{E}[X_n(X_n-1)\cdots(X_n-k+1)] \rightarrow \mathbb{E}[Y(Y-1)\cdots(Y-k+1)].$$

2. On suppose que Y suit une loi de Poisson de paramètre λ . Calculer $\mathbb{E}[Y(Y-1)\cdots(Y-k+1)]$.
3. On suppose que $X_n = \sum_{i=1}^N \mathbf{1}[A_i]$ pour certains événements A_i . Montrer que

$$\mathbb{E}[X_n(X_n-1)\cdots(X_n-k+1)] = \sum_{\substack{i_1, \dots, i_k \leq N \\ \text{distincts}}} \mathbb{P}[A_{i_1} \wedge \cdots \wedge A_{i_k}].$$

Indice : comme X_n est à valeurs entières, on a, pour toute fonction f ,

$$\mathbb{E}[f(Y)] = \sum_{r \geq 0} f(r) \mathbb{P}[X_n = r].$$

Théorème 4.11. Soit X_n le nombre de points fixes d'une permutation aléatoire σ_n . Alors X_n converge vers une loi de Poisson de paramètre 1.

Démonstration. D'après l'exercice précédent (questions 1 et 2), il suffit de montrer que, si Y suit une loi de Poisson de paramètre 1, on a, pour tout $k \geq 1$,

$$\mathbb{E}[X_n(X_n - 1) \cdots (X_n - k + 1)] \rightarrow \mathbb{E}[Y(Y - 1) \cdots (Y - k + 1)] = 1.$$

Mais, en utilisant la question 3 de l'exercice précédent, on a

$$\mathbb{E}[X_n(X_n - 1) \cdots (X_n - k + 1)] = \sum_{\substack{i_1, \dots, i_k \leq N \\ \text{distincts}}} \mathbb{P}[A_{i_1} \wedge \cdots \wedge A_{i_k}], \quad (7)$$

où A_i est l'évènement « $\sigma_n(i) = i$ ». Or, quand i_1, \dots, i_k sont distincts,

$$\mathbb{P}[A_{i_1} \wedge \cdots \wedge A_{i_k}] = \mathbb{P}[\sigma(i_1) = i_1 \wedge \cdots \wedge \sigma(i_k) = i_k] = \frac{1}{n(n-1) \cdots (n-k+1)},$$

où la dernière égalité est une application du théorème 2.3. On en déduit que les $n(n-1) \cdots (n-k+1)$ termes de la somme de l'équation (7) sont tous égaux à $\frac{1}{n(n-1) \cdots (n-k+1)}$ et on obtient

$$\mathbb{E}[X_n(X_n - 1) \cdots (X_n - k + 1)] = n(n-1) \cdots (n-k+1) \frac{1}{n(n-1) \cdots (n-k+1)} = 1,$$

ce qui finit la démonstration. \square

Note : dans cet exemple, les X_n et la loi limite Y prennent leurs valeurs dans un ensemble discret \mathbb{Z} , la convergence en distribution est donc équivalente à : pour tout j dans \mathbb{Z} , la probabilité $\mathbb{P}[X_n = j]$ tend vers $\mathbb{P}[Y = j] = e^{-1}/j!$.

Exercice 4.9. Prouver par la méthode des moments que, si S_n suit une loi binomiale $\text{Bin}(n, \lambda/n)$, alors S_n converge en loi vers une loi de Poisson de paramètre λ .

Exercice 4.10 (Secret Santa). On souhaite organiser un Père Noël Secret entre $2n$ individus, en couple 2 par 2, de telle sorte que personne ne s'offre de cadeaux à lui même, ni à son conjoint. Mathématiquement on veut tirer une permutation aléatoire uniforme, conditionnée à l'évènement :

$$\begin{aligned} & \sigma_n(i) \neq i \text{ pour tout } i \leq 2n, \\ & \text{et } \sigma_n(2k-1) \neq 2k \text{ et } \sigma_n(2k) \neq 2k-1 \text{ pour tout } k \leq n. \end{aligned}$$

Une méthode naïve consiste à tirer σ_n uniformément au hasard, et à recommencer si elle ne satisfait pas la condition voulue (méthode de rejet). Combien de fois doit-on recommencer en moyenne, à la limite $n \rightarrow \infty$?

Exercice 4.11 (Moments d'une loi de Poisson). 1. Soit $k > 0$. Montrer qu'il existe des nombres $(b_{i,k})_{i \leq k}$ tels que, pour tout entier $n \geq 1$, on ait

$$n^k = \sum_{i=1}^k b_{i,k} n(n-1) \cdots (n-i+1),$$

et donner une interprétation combinatoire des $b_{i,k}$ en termes de partitions d'ensembles.

(Indice : on comptera de deux manières différentes les suites (i_1, \dots, i_k) d'éléments dans $\llbracket 1, n \rrbracket$).

2. En déduire une formule combinatoire pour $\mathbb{E}[X^k]$, ou X suit une loi de Poisson de paramètre λ .
3. Soit B_n le nombre de partitions de l'ensemble $\llbracket 1, n \rrbracket$ (appelé n -ième nombre de Bell). Montrer que

$$B_n = \frac{1}{e} \sum_{k \geq 1} \frac{k^n}{k!}.$$

(Ce résultat est connu sous le nom de formule de Dobinski.)

4.2.3 Nombre de diviseurs d'un entier aléatoire

Soit $n \geq 1$, prenons x_n uniformément au hasard dans $\llbracket 1, n \rrbracket$. On s'intéresse au nombre de diviseurs premiers de x_n , que l'on notera $\nu(x_n)$. On commence par un résultat technique.

Lemme 4.12. *La quantité $\tilde{H}_n := \sum_{\substack{p \leq n \\ p \text{ premier}}} \frac{1}{p}$ tend vers l'infini.*

Esquisse de preuve. Supposons que \tilde{H}_n soit borné. Alors la série

$$\sum_{p \text{ premier}} \log\left(1 - \frac{1}{p}\right)$$

est convergente, et donc le produit $\prod_p \frac{1}{1 - \frac{1}{p}}$ est convergent aussi. Mais, d'après la formule d'Euler pour la fonction zeta, $\prod_p \frac{1}{1 - \frac{1}{p}} = \sum_n \frac{1}{n}$, et donc $\sum_n \frac{1}{n}$ serait convergent : on aboutit à une contradiction. \square

Note : en fait, il est connu que $\tilde{H}_n = \log(\log(n)) + M + o(1)$, pour un certain M réel, appelé constante de Meissel-Mertens ($M \approx 0.2615$).

Theorème 4.13 (Erdős-Kac theorem).

$$\frac{\nu(x) - \tilde{H}_n}{\sqrt{\tilde{H}_n}} \rightarrow_d \mathcal{N}(0, 1).$$

Démonstration. Posons $Y_n = \nu(x_n)$, on a $Y_n = \sum_{p \leq n} \mathbf{1}[A_p]$, où A_p est l'événement « p divise x_n ». Ici, et dans ce qui suit, les sommes sur des indices p sont implicitement prises sur les nombres premiers p . Notons que

$$\mathbb{P}[A_p] = \frac{1}{n} \sum_{x=1}^n \mathbf{1}[p|x] = \frac{1}{n} \left\lfloor \frac{n}{p} \right\rfloor = \frac{1}{p} + O(n^{-1}),$$

avec un grand O uniforme en p (l'erreur est toujours inférieur à $1/n$). En sommant sur p , on obtient $\mathbb{E}[Y_n] = \tilde{H}_n + O(1)$.

Pour calculer les moments d'ordre supérieur, il faut tronquer la somme, sinon les termes d'erreurs vont être trop grands. Soit ε_n une suite qui tendant vers l'infini, moins vite que $\sqrt{\tilde{H}_n}$. On approche Y_n par $Z_n := \sum_{p \leq n^{1/\varepsilon_n}} \mathbf{1}[A_p]$. On a $\mathbb{E}[Z_n] = \tilde{H}_{n^{1/\varepsilon_n}} + O(1)$.

Comme il ne peut y avoir plus de ε_n nombres premiers $p \geq n^{1/\varepsilon_n}$ divisant x_n , on a $Z_n \leq Y_n \leq Z_n + \varepsilon_n$. Cela implique

$$\mathbb{E}[Z_n] - \mathbb{E}[Y_n] = \tilde{H}_n - \tilde{H}_{n^{1/\varepsilon_n}} + O(1) \leq \varepsilon_n.$$

Comme $\varepsilon_n = o(\tilde{H}_n)$, cela signifie que

$$\mathbb{E}[Z_n] = \tilde{H}_{n^{1/\varepsilon_n}} + O(1) = \tilde{H}_n + O(\varepsilon_n).$$

Par conséquent, montrer le théorème revient à montrer que

$$\tilde{Z}_n := \frac{Z_n - \mathbb{E}[Z_n]}{\sqrt{\tilde{H}_{n^{1/\varepsilon_n}}}}$$

converge en distribution vers une Gaussienne.

On écrit

$$Z_n - \mathbb{E}[Z_n] = \sum_{p \leq n^{1/\varepsilon_n}} (\mathbf{1}[A_p] - \frac{1}{p}).$$

Soit B_p des variables de Bernoulli indépendantes de paramètre $1/p$, on considère aussi leur version centrée $B'_p = B_p - \frac{1}{p}$. Pour $p_1, \dots, p_k \leq n$ donnés, on a

$$\begin{aligned} \mathbb{E}[\mathbf{1}[A_{p_1}] \cdots \mathbf{1}[A_{p_k}]] &= \frac{1}{n} \sum_{x=1}^n \mathbf{1}[(p_1|x) \wedge \cdots \wedge (p_k|x)] = \frac{1}{n} \lfloor \frac{n}{p_1 \cdots p_k} \rfloor \\ &= \frac{1}{p_1 \cdots p_k} + O(n^{-1}) = \mathbb{E}[B_{p_1} \cdots B_{p_k}] + O(n^{-1}). \end{aligned}$$

On en déduit une égalité similaire pour les versions centrées :

$$\mathbb{E}\left[\left(\mathbf{1}[A_{p_1}] - \frac{1}{p_1}\right) \cdots \left(\mathbf{1}[A_{p_k}] - \frac{1}{p_k}\right)\right] = \mathbb{E}\left[B'_{p_1} \cdots B'_{p_k}\right] + O(n^{-1}).$$

Cela donne, pour $k \geq 1$,

$$\begin{aligned} \mathbb{E}[(Z_n - \mathbb{E}[Z_n])^k] &= \sum_{p_1, \dots, p_k \leq n^{1/\varepsilon_n}} \mathbb{E}\left[\left(\mathbf{1}[A_{p_1}] - \frac{1}{p_1}\right) \cdots \left(\mathbf{1}[A_{p_k}] - \frac{1}{p_k}\right)\right] \\ &= \sum_{p_1, \dots, p_k \leq n^{1/\varepsilon_n}} \mathbb{E}\left[B'_{p_1} \cdots B'_{p_k}\right] + O(n^{-1} n^{k/\varepsilon_n}). \end{aligned}$$

Le terme d'erreur est négligeable (note : si on n'avait pas tronqué la somme, comme il y a environ $n/\log(n)$ nombres premiers $\leq n$, on aurait un terme d'erreur de l'ordre $O(n^{k-1}/\log(n)^k)$, bien supérieur au terme principal).

Pour la terme principal, comme dans la preuve du théorème 4.10, on coupe selon la partition π de $\llbracket 1, k \rrbracket$ décrivant l'égalité des p . Pour π donnée avec ℓ parts de tailles m_1, \dots, m_ℓ ,

$$\sum_{\substack{p_1, \dots, p_k \leq n^{1/\varepsilon_n} \\ (p_1, \dots, p_k) \in \mathcal{I}_\pi}} \mathbb{E} \left[B'_{p_1} \cdots B'_{p_k} \right] = \sum_{\substack{q_1, \dots, q_\ell \leq n^{1/\varepsilon_n} \\ \text{distinct}}} \mathbb{E} \left[(B'_{q_1})^{m_1} \right] \cdots \mathbb{E} \left[(B'_{q_\ell})^{m_\ell} \right].$$

Ici, q_1, \dots, q_ℓ sont des nombres premiers correspondant aux p_1, \dots, p_k de la somme précédente, après effacement des répétitions. Comme précédemment si un des m_i vaut 1, alors tous les termes de la somme valent 0. Pour $m \geq 2$, on peut estimer $\mathbb{E} \left[(B'_q)^m \right] = \frac{1}{q} \left(1 - \frac{1}{q}\right)^m - \left(1 - \frac{1}{q}\right) \frac{1}{q^m} \sim 1/q$ quand q tend vers l'infini, de telle sorte que

$$\sum_{q_1, \dots, q_\ell \leq n^{1/\varepsilon_n}} \mathbb{E} \left[(B'_{q_1})^{m_1} \right] \cdots \mathbb{E} \left[(B'_{q_\ell})^{m_\ell} \right] \sim \sum_{q_1, \dots, q_\ell \leq n^{1/\varepsilon_n}} \frac{1}{q_1 \cdots q_\ell} = \tilde{H}_{n^{1/\varepsilon_n}}^\ell.$$

De plus, on peut montrer que les suites q_1, \dots, q_ℓ avec répétition ont une contribution négligeable. On a donc

$$\mathbb{E} \left[(Z_n - \mathbb{E}[Z_n])^k \right] = \sum_{\pi} \tilde{H}_{n^{1/\varepsilon_n}}^{\#\pi} (1 + o(1)) + O(n^{-1} n^{k/\varepsilon_n}),$$

où la somme est prise sur les partitions π de $\llbracket 1, k \rrbracket$ et $\#\pi$ désigne le nombre de parts de π . Les termes contribuant le plus sont les partitions sans singletons avec le plus de parts possibles, i.e. les partitions en paires. En se rappelant qu'il y a $1 \cdot 3 \cdots (k-1)$ telles partitions pour k pair, et aucune pour k impair, on obtient

$$\mathbb{E} \left[(Z_n - \mathbb{E}[Z_n])^k \right] = \begin{cases} 1 \cdot 3 \cdots (k-1) \tilde{H}_{1/\varepsilon_n}^{k/2} & \text{si } k \text{ est pair;} \\ o(\tilde{H}_{1/\varepsilon_n}^{k/2}) & \text{si } k \text{ est impair.} \end{cases}$$

En divisant par $H_{1/\varepsilon_n}^{k/2}$, on voit que les moments de \tilde{Z}_n tendent vers ceux d'une Gaussienne centrée réduite, ce qui conclut la démonstration. \square

Deuxième partie

Marches aléatoires et fonction caractéristique

5 Modèle et problématique

Soient X_1, X_2, \dots des variables aléatoires i.i.d. de loi μ à valeurs dans \mathbb{Z} . On définit $S_n = X_1 + X_2 + \dots$. On parle de marche aléatoire sur \mathbb{Z} : pour passer de

S_n à S_{n+1} , on fait un « pas » de loi μ . Le cas où $\mathbb{P}[X_1 = -1] = \mathbb{P}[X_1 = 1] = 1/2$ est appelé marche aléatoire simple.

La loi des grands nombres et le théorème centrale limite nous donnent le comportement typique de S_n quand n tend vers l'infini. Il est parfois utile d'avoir des résultats complémentaires :

- Qualité de l'approximation Gaussienne (vitesse de convergence)
- Estimation de la probabilité de divers événements « rares » (i.e. dont la probabilité tend vers 0) : arriver en un point précis au temps n (variante : y arriver pour la première fois au temps n), arriver loin de la valeur typique, ... ;
- Étude de l'ensemble de la trajectoire $(S_k)_{k \leq n}$ (probabilité qu'elle soit toujours positive, limite en tant que fonction \rightarrow mouvement Brownien).

Principal outil ici : transformée de Fourier. Beaucoup de résultats peuvent se généraliser à des suites de v.a., qui ne sont pas nécessairement des sommes de variable i.i.d., mais dont on peut calculer/estimer la transformée de Fourier (en général c'est plus difficile que d'estimer les moments, mais les résultats obtenus sont plus précis).

Lien avec la recherche : les marches aléatoires classiques ne sont plus un objet de recherche, mais

- les résultats présentés ici sont parfois utiles, aussi bien dans des problèmes théoriques qu'appliqués ;
- prouver des analogues des résultats présentés ici (grande déviation, vitesse de convergence, convergence fonctionnelle, ...) sur des processus plus compliqués (avec mémoire, environnement aléatoire, ...) est une question classique en recherche de nos jours.

5.1 Préliminaires

Quelques notations : On travaillera implicitement sur un espace de probabilité $(\Omega, \mathcal{A}, \mathbb{P})$. Pour une variable aléatoire $X : \Omega \rightarrow \mathbb{R}$ ou \mathbb{C} , on notera $\mathbb{E}[X]$ son espérance (si elle est bien définie) et $\varphi_X(t) := \mathbb{E}[\exp(itX)]$ sa transformée de Fourier.

Dans tout le document Z est une variable Gaussienne centrée réduite (c'est-à-dire de moyenne 0 et variance 1). Pour $\sigma > 0$, le produit σZ est alors une variable Gaussienne centrée de variance σ^2 , a pour densité

$$n_\sigma(x) = \frac{1}{\sigma\sqrt{2\pi}} \exp(-x^2/(2\sigma^2))$$

et pour transformée de Fourier (ou fonction caractéristique)

$$\varphi_{\sigma Z}(t) := \mathbb{E}[\exp(it\sigma Z)] = \exp(-t^2\sigma^2/2).$$

Rappel : théorème central limite. Soit $m = \mathbb{E}(X_1)$, qu'on supposera fini. Quitte à remplacer X_i par $X_i - m$ on peut supposer $m = 0$; on parle de marche aléatoire centrée. On suppose aussi que $\sigma^2 := \mathbb{E}(X_1)^2$ appartient à $]0; +\infty[$.

Theorème 5.1. *Sous les conditions ci-dessus, $\frac{S_n}{\sqrt{n}}$ converge en distribution vers une variable Gaussienne σZ centrée de variance σ .*

Esquisse de preuve. La fonction caractéristique de $\frac{S_n}{\sqrt{n}}$ est

$$\varphi_n(t) := \mathbb{E}[\exp(itS_n/\sqrt{n})] = \prod_{i=1}^n \varphi_{X_i}(t/\sqrt{n}) = \varphi_{X_1}(t/\sqrt{n})^n,$$

où on a utilisé l'indépendance des X_i à la deuxième étape et le fait qu'ils ait tous la même distribution à la troisième. Mais on peut montrer que, sous les hypothèses ci-dessus quand u tend vers 0,

$$\varphi_{X_1}(u) = \mathbb{E}[\exp(iuX_1)] = 1 - \frac{u^2}{2}\sigma^2 + o(u^2).$$

Donc, pour t fixé dans \mathbb{R} ,

$$\varphi_n(t) = \left(1 - \frac{t^2\sigma^2}{2n} + o(1/n)\right)^n \rightarrow \exp\left(-\frac{t^2\sigma^2}{2}\right) = \varphi_{\sigma Z}(t).$$

Le théorème de continuité de Lévy¹ implique que $\frac{S_n}{\sqrt{n}}$ converge en distribution vers $\psi_\sigma(t)$. □

Exercice 5.1 (« loi des petits nombres »). Soit p_n une suite dans $]0; 1[$ vérifiant $n p_n \rightarrow \lambda > 0$. Pour chaque n , on considère une suite de variables aléatoires de Bernoulli $X_{n,1}, \dots, X_{n,n}$ indépendantes de paramètre p_n . Montrer, en utilisant le théorème de continuité de Lévy, que $S_n = \sum_{i=1}^n X_{n,i}$ converge en distribution vers une loi de Poisson.

(Note : c'est le même résultat que l'exercice 4.9, mais la méthode demandée est différente.)

6 Théorème local limite

Dans cette section, on regarde des probabilités ponctuelles de S_n , i.e. des évènements du type $\mathbb{P}[S_n = a]$, avec a dans \mathbb{Z} .

Que nous dit le théorème centrale limite ? Par exemple pour $a = 0$, le théorème Portmanteau implique

$$\mathbb{P}[S_n = 0] = \mathbb{P}\left[\frac{|S_n|}{\sqrt{n}} = 0\right] \longrightarrow \mathbb{P}[Z = 0] = 0. \quad (8)$$

1. Théorème de continuité de Lévy : X_n converge en distribution vers X si et seulement si $\varphi_{X_n}(t)$ converge vers $\varphi_X(t)$ pour tout réel t

En effet, l'ensemble $\{0\}$ est un Borélien de \mathbb{R} , dont la frontière, $\{0\}$, a mesure nulle pour la distribution de la variable limite Z .

Problème : trouver un équivalent asymptotique de $\mathbb{P}[S_n = 0]$.

Préparation : Si Y est une variable à valeurs dans \mathbb{Z} , on a

$$\varphi_Y(t) = \mathbb{E}[\exp(itY)] = \sum_{k \in \mathbb{Z}} \mathbb{P}[Y = k] \exp(itk).$$

On voit que φ_Y est 2π -périodique et le membre de droite est son développement en série de Fourier. La formule d'inversion de Fourier donne alors :

$$\mathbb{P}[Y = k] = \frac{1}{2\pi} \int_{-\pi}^{\pi} \varphi_Y(t) \exp(-itk) dt. \quad (9)$$

Cette formule a un équivalent continu :

Theorème 6.1 (Théorème d'inversion de Fourier pour des distributions régulières). *Soit Y une variable aléatoire à valeurs réelles tel que φ_Y est intégrable, alors Y a une densité continue et bornée f_Y donnée par*

$$f_Y(x) = \frac{1}{2\pi} \int_{-\infty}^{\infty} e^{-itx} \varphi_Y(t) dt.$$

De plus, le lien entre l'ensemble des valeurs de Y et périodicité de φ_Y peut être raffiné ainsi.

Lemme 6.2. *Soit Y une variable aléatoire réelle et h dans \mathbb{R} . Les deux assertions suivantes sont équivalentes :*

- (i) *Il existe b tel que $Y \in b + h\mathbb{Z}$ presque sûrement ;*
- (ii) *$|\varphi_Y(2\pi/h)| = 1$.*

Dans ce cas, h est appelée période de Y . On peut montrer qu'une variable aléatoire non dégénérée² a une unique période maximale. En particulier, si Y a période maximale 1, on a $|\varphi_Y(2\pi)| = 1$, mais $|\varphi_Y(t)| < 1$ pour t dans $]0; 2\pi[$.

Démonstration. Supposons (i). Alors $\exp(2i\pi Y/h) = \exp(2i\pi b/h)$ p.s., et donc $\varphi_Y(2\pi/h) = \exp(2i\pi b/h)$, ce qui prouve (ii).

Supposons (ii). On écrit $\varphi_Y(2\pi/h) = \exp(2i\pi\alpha)$ pour un certain α dans \mathbb{R} . On a

$$\mathbb{E}[\exp(2i\pi(Y/h - \alpha))] = 1.$$

Or la partie réelle de $\exp(2i\pi(Y/h - \alpha))$ est inférieure à 1 p.s. Ceci indique qu'elle est égale à 1 p.s., forçant $Y/h - \alpha \in \mathbb{Z}$. On a prouvé (i) avec $b = \alpha h$. \square

2. Une v.a. Y est dégénérée s'il existe b tel que $Y = b$ p.s.

Théorème 6.3 (Théorème local limite). . Soit $S_n = X_1 + \dots + X_n$ une marche aléatoire centrée, avec X_1 à valeurs dans \mathbb{Z} de période maximale $h = 1$. On suppose que $\mathbb{E}(X_1) = 0$ et que $\sigma^2 := \mathbb{E}(X_1)^2$ appartient à $]0; +\infty[$. Alors

$$\sup_{a \in \mathbb{Z}} |\sqrt{n} \mathbb{P}[S_n = a] - n_\sigma(a/\sqrt{n})| \rightarrow 0, \quad (10)$$

quand n tends vers l'infini.

Autrement dit,

$$\mathbb{P}[S_n = a] = \frac{n_\sigma(a/\sqrt{n})}{\sqrt{n}} + o(n^{-1/2}),$$

avec un terme d'erreur *uniforme pour tout a dans \mathbb{Z}* (on peut donc en particulier prendre a dépendant de n). En particulier pour a constant,

$$\mathbb{P}[S_n = a] = \frac{1}{\sigma\sqrt{2\pi n}} + o(n^{-1/2}). \quad (11)$$

Démonstration. Soit a dans \mathbb{Z} . On pose $x = a/\sqrt{n}$ et on écrit

$$\begin{aligned} \mathbb{P}[S_n = a] &= \frac{1}{2\pi} \int_{-\pi}^{\pi} \varphi_{S_n}(u) \exp(-iua) du \\ &= \frac{1}{2\pi\sqrt{n}} \int_{-\pi\sqrt{n}}^{\pi\sqrt{n}} \varphi_{S_n}(t/\sqrt{n}) \exp(-itx) dt, \end{aligned}$$

où la première égalité vient de (9) et la deuxième du changement de variable $u = t/\sqrt{n}$ (et donc $du = dt/\sqrt{n}$). Par ailleurs, le théorème 6.1 appliqué à σZ implique

$$n_\sigma(x) = \frac{1}{2\pi} \int_{-\infty}^{\infty} e^{-itx} \varphi_{\sigma Z}(t) dt.$$

Les deux équations ci-dessus impliquent :

$$\begin{aligned} |\sqrt{n} \mathbb{P}[S_n = a] - n_\sigma(a/\sqrt{n})| &\leq \frac{1}{2\pi} \int_{-\pi\sqrt{n}}^{\pi\sqrt{n}} |\varphi_{S_n}(t/\sqrt{n}) - \varphi_{\sigma Z}(t)| dt \\ &\quad + \frac{1}{2\pi} \int_{]-\infty; -\pi\sqrt{n}[\cup]\pi\sqrt{n}; +\infty[} |\varphi_{\sigma Z}(t)| dt \end{aligned}$$

On veut montrer que le membre de gauche tend vers 0 uniformément en a . Comme le membre de droite ne dépend pas de a , il suffit de montrer qu'il tend vers 0. La deuxième intégrale tend clairement vers 0 car $\varphi_{\sigma Z}$ est intégrable. Pour la première, on fixe $A, \delta > 0$ (que l'on spécifiera plus tard) et on coupe l'intégrale en trois morceaux $I_1(n)$ sur $[-A; A]$, $I_2(n)$ sur $[-\delta\sqrt{n}; -A] \cup [A; \delta\sqrt{n}]$ et $I_3(n)$ sur $[-\pi\sqrt{n}; -\delta\sqrt{n}] \cup [\delta\sqrt{n}; \pi\sqrt{n}]$.

Commençons par $I_2(n)$. On rappelle que

$$\varphi_{S_n}(u) = \varphi_{X_1}(u)^n, \text{ avec } \varphi_{X_1}(u) = 1 - \frac{u^2}{2}\sigma^2 + o(u^2).$$

En particulier il existe $\delta > 0$ tel que pour tout $|u| \leq \delta$, on a

$$|\varphi_{X_1}(u)| \leq 1 - \frac{u^2}{4}\sigma^2.$$

Ceci implique que pour $t \leq \delta\sqrt{n}$, on a

$$|\varphi_{S_n}(t/\sqrt{n})| \leq \left(1 - \frac{t^2\sigma^2}{4n}\right)^n \leq \exp\left(-\frac{t^2\sigma^2}{4n}\right)^n = \exp\left(-\frac{t^2\sigma^2}{4}\right),$$

où dans la deuxième inégalité, on a utilisé l'inégalité $1 - x \leq e^{-x}$ valable pour x in $[0; 1]$. On a donc

$$|I_2(n)| \leq 2 \int_A^{+\infty} \left(\exp\left(-\frac{t^2\sigma^2}{4}\right) + \exp\left(-\frac{t^2\sigma^2}{2}\right) \right) dt$$

Le membre de droite peut être rendu plus petit que n'importe quel seuil fixé $\varepsilon > 0$, en choisissant A suffisamment grand.

Regardons maintenant $I_1(n)$, avec A fixé par l'argument ci-dessus. Comme $\varphi_{S_n}(t/\sqrt{n})$ converge simplement vers $\varphi_{\sigma Z}(t)$ (voir la preuve du théorème 5.1), et que ces quantités sont bornées, le théorème de convergence dominée implique,

$$\lim_{n \rightarrow \infty} I_1(n) = \int_{-A}^A \left(\lim_{n \rightarrow \infty} |\varphi_{S_n}(t/\sqrt{n}) - \varphi_{\sigma Z}(t)| \right) dt = 0.$$

Pour finir, considérons $I_3(n)$ (avec δ fixé par l'argument ci-dessus). La fonction $u \mapsto \varphi_{X_1}(u)$ est continue sur $[\delta; \pi]$ et atteint donc son maximum (en module) sur cet intervalle. Comme 1 est la période maximale de X_1 , le lemme 6.2 implique que $|\varphi_{X_1}(u)| < 1$ sur cette intervalle. La même chose est vraie sur $[-\pi; -\delta]$. Il existe donc $\eta < 1$ telle que $|\varphi_{X_1}(u)| \leq \eta$ pour tout u dans $[-\pi; -\delta] \cup [\delta; \pi]$. On conclut

$$I_3(n) \leq 2 \int_{\delta\sqrt{n}}^{\pi\sqrt{n}} \left(\eta^n + \exp\left(-\frac{t^2\sigma^2}{2}\right) \right) dt \leq 2\pi\sqrt{n}\eta^n + 2 \int_{\delta\sqrt{n}}^{+\infty} \exp\left(-\frac{t^2\sigma^2}{2}\right) dt.$$

Le membre de droite tend clairement vers 0 quand n tends vers l'infini, ce qui conclut la preuve du théorème. \square

Une conséquence : la marche aléatoire S_n peut être vu comme une chaîne de Markov. Dans le cas non centré (i.e. $\mathbb{E}(X_1) \neq 0$), la loi des grands nombres montrent que, p.s., on a $S_n \neq 0$ pour n assez grand (avec un seuil dépendant de ω dans l'espace de probabilités), ce qui implique que 0 est un état transient. Pour une marche centrée, avec X_1 de période maximale 1, on peut appliquer le théorème limite local avec $a = 0$:

$$\mathbb{P}[S_n = 0] \sim \frac{n^{-1/2}}{\sigma\sqrt{2\pi}}.$$

This implies that $\sum \mathbb{P}[S_n = 0] = +\infty$, ce qui est prouvé que 0 est un état récurrent.

Remarque : le théorème local limite implique le théorème central limite. En effet, supposons que (10) soit valide. On a, pour des réels $b < c$,

$$\mathbb{P}\left[\frac{S_n}{\sqrt{n}} \in [b, c]\right] = \sum_{a=b\sqrt{n}}^{c\sqrt{n}} \mathbb{P}[S_n = a] = \sum_{a=b\sqrt{n}}^{c\sqrt{n}} \left(\frac{n_\sigma(a/\sqrt{n})}{\sqrt{n}} + o(n^{-1/2}) \right),$$

avec un terme d'erreur uniforme en a . Les termes d'erreurs somment à $o(1)$, et le terme principal est une somme de Riemann et converge vers l'intégrale $\int_b^c n_\sigma(x) dx$. Finalement, on a, que pour tous réels $b < c$,

$$\lim \mathbb{P}\left[\frac{S_n}{\sqrt{n}} \in [b, c]\right] \rightarrow \int_b^c n_\sigma(x) dx,$$

ce qui implique la convergence en distribution de $\frac{S_n}{\sqrt{n}}$ vers σZ , i.e. le théorème central limite.

6.1 Exercices

Exercice 6.1. Calculer le nombre de suites (x_1, \dots, x_n) dans $\{-1, 1\}^n$ telles que $\sum_{i=1}^n x_i = 0$. En déduire, pour la marche aléatoire **simple** sur \mathbb{Z} (voir définition p. 21), la probabilité $\mathbb{P}[S_n = 0]$. Utiliser la formule de Stirling pour trouver un équivalent asymptotique de $\mathbb{P}[S_n = 0]$. Comparer avec l'équation (11).

Exercice 6.2 (théorème local limite pour une variable continue). On considère ici une variable aléatoire X_1 à valeurs dans \mathbb{R} . On suppose $\mathbb{E}(X_1) = 0$ et que $\mathbb{E}(X_1^2) = \sigma^2$ est non nulle et finie. On suppose³ aussi que φ_X est intégrable, i.e. $\int_{-\infty}^{\infty} |\varphi_X(t)| dt < +\infty$. Soit x_n une suite telle que $\lim_{n \rightarrow \infty} x_n/\sqrt{n} = x$ et on prend $a < b$. En utilisant le théorème d'inversion de Fourier pour les distributions régulières, trouver un équivalent asymptotique pour

$$\mathbb{P}(S_n \in [x_n + a, x_n + b]).$$

L'exercice suivant concerne le théorème d'inversion de Fourier et sert de préparation pour la section suivante

Exercice 6.3. Soit $X = U_1 + U_2$ où U_1 et U_2 sont des variables uniformes dans $[-\frac{1}{2}, \frac{1}{2}]$. Calculer la densité de X et sa fonction caractéristique. En utilisant le théorème d'inversion de Fourier, construire une variable Y dont la fonction caractéristique est à support compact.

3. En fait on peut montrer le même résultat en supposant seulement $|\varphi_X(t)| < 1$ pour tout t dans \mathbb{R} , mais la preuve est plus difficile...

7 Vitesse de convergence

Reprenons le cadre du théorème central limite. Si les X_i sont des variables i.i.d. centrées de variance $\sigma^2 < +\infty$, alors $S_n/\sigma\sqrt{n}$ converge en loi vers une loi Gaussienne centrée Z . Concrètement pour tout x dans \mathbb{R} on a :

$$\lim_{n \rightarrow +\infty} \mathbb{P}\left[\frac{S_n}{\sigma\sqrt{n}} \leq x\right] = \mathbb{P}[Z \leq x].$$

La question est la vitesse de cette convergence. On va montrer ici que, en supposant que $\rho := \mathbb{E}[|X|^3] < +\infty$, le terme d'erreur est $O(n^{-1/2})$, uniformément en x . Plus précisément, on va montrer le théorème suivant, connu sous le nom d'inégalité de Berry–Esseen :

Théorème 7.1. *Il existe une constante C qui vérifie la propriété suivante. Pour toute suite X_1, X_2, \dots de v.a. réelles i.i.d. centrées de variance $\sigma^2 < +\infty$, et de troisième moment absolu $\rho := \mathbb{E}[|X_1|^3] < +\infty$, on a*

$$\sup_{x \in \mathbb{R}} \left| \mathbb{P}\left[\frac{S_n}{\sigma\sqrt{n}} \leq x\right] - \mathbb{P}[Z \leq x] \right| \leq \frac{C\rho}{\sigma^3\sqrt{n}}.$$

Il a été démontré (en 2010) que l'on peut choisir $C = 0,4785$. La preuve donnée ici établit le théorème avec une valeur de $C = 3,8$.

Encore une fois, l'outil principal est la fonction caractéristique. Pour faire le lien entre vitesse de convergence et fonction caractéristique, commençons par considérer le cas de fonctions caractéristiques intégrables.

Lemme 7.2. *Soit V et W des v.a. d'espérance finie dont les fonctions caractéristiques φ_V et φ_W sont dans L^1 . Alors, pour tout x dans \mathbb{R} , on a*

$$\mathbb{P}[V \leq x] - \mathbb{P}[W \leq x] = \frac{-1}{2\pi} \int_{\mathbb{R}} e^{-itx} \frac{\varphi_V(t) - \varphi_W(t)}{it} dt.$$

Démonstration. Par le théorème d'inversion de Fourier (Théorème 6.1), V a densité

$$f_V(u) = \frac{1}{2\pi} \int_{\mathbb{R}} e^{-itu} \varphi_V(t) dt,$$

et donc, pour tout $a < x$,

$$\mathbb{P}[a \leq V \leq x] = \frac{1}{2\pi} \int_a^x \int_{\mathbb{R}} e^{-itu} \varphi_V(t) dt du.$$

En utilisant une formule similaire pour W et en échangeant les intégrales par le théorème de Fubini (comme φ_V et φ_W sont dans L^1 , la quantité à intégrer est absolument intégrable sur $[a, x] \times \mathbb{R}$), on obtient

$$\mathbb{P}[a \leq V \leq x] - \mathbb{P}[a \leq W \leq x] = \int_{\mathbb{R}} (\varphi_V(t) - \varphi_W(t)) \left(\int_a^x e^{-itu} du \right) dt.$$

L'intégrale intérieure vaut $\frac{-1}{it}(e^{-itx} - e^{-ita})$ et donc

$$\begin{aligned} \mathbb{P}[a \leq V \leq x] - \mathbb{P}[a \leq W \leq x] \\ = \int_{\mathbb{R}} \frac{\varphi_V(t) - \varphi_W(t)}{it} e^{-itx} dt - \int_{\mathbb{R}} \frac{\varphi_V(t) - \varphi_W(t)}{it} e^{-ita} dt. \end{aligned} \quad (12)$$

Une précision s'impose. Comme les variables V et W ont une espérance finie, φ_V et φ_W sont dérivables en 0, et $\varphi'_V(0) = i\mathbb{E}[V]$, $\varphi'_W(0) = i\mathbb{E}[W]$. Comme $\varphi_V(0) = \varphi_W(0) = 1$, le quotient $\frac{\varphi_V(t) - \varphi_W(t)}{it}$ tend vers $\mathbb{E}[V] - \mathbb{E}[W]$. Par conséquent, la fonction $t \mapsto \frac{\varphi_V(t) - \varphi_W(t)}{it}$ peut être prolongée en une fonction continue sur \mathbb{R} (rappelons que φ_V et φ_W sont continues, comme toute fonction caractéristique). La fonction ainsi obtenue est aussi bornée et intégrable, car φ_V et φ_W le sont. Les intégrales du membre de droite de (12) ont donc bien un sens.

Par ailleurs, il est classique que pour toute fonction f continue intégrable,

$$\lim_{b \rightarrow +\infty} \int_{\mathbb{R}} f(t) e^{itb} dt = 0.$$

En faisant tendre a vers $-\infty$ dans l'équation (12), on obtient le lemme. \square

On ne peut pas appliquer ce lemme à $V = \frac{S_n}{\sigma\sqrt{n}}$ et $W = Z$ car l'hypothèse φ_V dans L^1 n'est pas vérifiée en générale (en particulier, si X_1 et donc S_n sont des variables discrètes, φ_V dans L^1 est impossible, à cause du théorème d'inversion de Fourier, Théorème 6.1). Le lemme suivant diminue la régularité demandée sur V .

Lemme 7.3. *Soit V et W des v.a. d'espérance finie. On suppose que W a une densité $g(x)$ par rapport à la mesure de Lebesgue sur \mathbb{R} , et que g est bornée par une constante λ . Alors, pour tout x dans \mathbb{R} et tout $L > 0$, on a*

$$\mathbb{P}[V \leq x] - \mathbb{P}[W \leq x] \leq \frac{1}{\pi} \int_{-L}^L \frac{|\varphi_V(t) - \varphi_W(t)|}{t} dt + \frac{24\lambda}{\pi L}.$$

Démonstration. L'idée est d'approximer V et W par des « convolutions » $V + Y_L$ et $W + Y_L$ avec Y_L bien choisi, indépendant de V et W . On rappelle que $\varphi_{V+Y_L} = \varphi_V \varphi_{Y_L}$ donc pour que φ_{V+Y_L} soit intégrable, il faut que φ_{Y_L} décroisse vite à l'infini. On va en fait choisir Y_L telle que φ_{Y_L} soit à support compact.

D'après l'exercice 6.3, il existe une v.a. Y de densité $\frac{2}{\pi} \frac{\sin(t/2)^2}{t^2}$ sur \mathbb{R} et de fonction caractéristique

$$\varphi_Y(s) = \begin{cases} 1 - |s| & \text{si } |s| \leq 1; \\ 0 & \text{si } |s| \geq 1. \end{cases}$$

On pose $Y_L = Y/L$ de sorte que

$$\varphi_{Y_L}(s) = \varphi_Y(s/L) = \begin{cases} 1 - |s/L| & \text{si } |s| \leq L; \\ 0 & \text{si } |s| \geq L. \end{cases}$$

Par ailleurs Y_L a densité $f_L(y) := \frac{2}{L\pi} \frac{\sin(yL/2)^2}{y^2}$. Notons que pour L grand, Y_L se concentre autour de 0. Plus précisément, on a : pour tout $\delta > 0$,

$$\mathbb{P}[|Y_L| \geq \delta] \leq \int_{\mathbb{R} \setminus [-\delta, \delta]} f_L(y) dy \leq \frac{2}{L\pi} \int_{\mathbb{R} \setminus [-\delta, \delta]} \frac{dy}{y^2} = \frac{4}{\delta L\pi}, \quad (13)$$

ce qui nous servira plus tard dans la démonstration.

Les variables $V + Y_L$ et $W + Y_L$ ont des fonction caractéristiques $\varphi_V \varphi_{Y_L}$ et $\varphi_W \varphi_{Y_L}$ continues à support compact (inclus dans $[-L, L]$), donc intégrable. En utilisant le lemme 7.2 et l'inégalité $|\varphi_{Y_L}(t)| \leq \mathbf{1}[-L, L]$, on obtient, que, pour tout u dans \mathbb{R}

$$\begin{aligned} \left| \mathbb{P}[V + Y_L \leq u] - \mathbb{P}[W + Y_L \leq u] \right| &\leq \left| \frac{-1}{2\pi} \int_{\mathbb{R}} e^{-itu} \varphi_{Y_L}(t) \frac{\varphi_V(t) - \varphi_W(t)}{it} dt \right| \\ &\leq \frac{1}{2\pi} \int_{-L}^L \frac{|\varphi_V(t) - \varphi_W(t)|}{|t|} dt. \end{aligned} \quad (14)$$

Par ailleurs, on peut écrire

$$\mathbb{P}[V + Y_L \leq u] = \mathbb{E}[\mathbb{P}[V \leq u - Y_L | Y_L]] = \mathbb{E}(F_V(u - Y_L)),$$

où $F_V(x) = \mathbb{P}[V \leq x]$ est la fonction de répartition de V . La même chose est vraie en remplaçant V par W , d'où en faisant la différence

$$\Delta_L(u) := \mathbb{P}[V + Y_L \leq u] - \mathbb{P}[W + Y_L \leq u] = \mathbb{E}[\Delta(u - Y_L)],$$

où $\Delta(x) = \mathbb{P}[V \leq x] - \mathbb{P}[W \leq x]$ est la quantité que l'on cherche à borner.

Notons que Δ est continue à droite, fait uniquement des sauts positifs et tend vers 0 en $\pm\infty$. Elle atteint donc son maximum en un certain x_0 et « atteint » son minimum en un certain x_1^- (limite à gauche en un certain x_1 , où Δ est potentiellement discontinue). Supposons sans perte de généralité que $\Delta(x_0) \geq -\Delta(x_1^-)$ (dans le cas contraire, on remplace V par $-V$ et W par $-W$). En utilisant que $x \mapsto \mathbb{P}[W \leq x]$ est λ -Lipschitzienne, on a, pour $h \geq 0$,

$$\begin{aligned} \Delta(x_0 + h) &= \mathbb{P}[V \leq x_0 + h] - \mathbb{P}[W \leq x_0 + h] \\ &\geq \mathbb{P}[V \leq x_0] - (\mathbb{P}[W \leq x_0] + \lambda h) = \Delta(x_0) - \lambda h. \end{aligned}$$

Ceci nous permet de borner inférieurement Δ_L en un point bien choisi. En notant $\delta = \Delta(x_0)/(2\lambda)$, on a

$$\Delta_L(x_0 + \delta) = \mathbb{E}[\Delta(u - Y_L)]. \quad (15)$$

Pour Y_L dans $[-\delta, \delta]$, on applique la borne ci-dessus et on a

$$\mathbb{E}[\Delta(u - Y_L) \mathbf{1}[|Y_L| \leq \delta]] \geq \mathbb{E}[(\Delta(x_0) - \lambda(\delta - Y_L)) \mathbf{1}[|Y_L| \leq \delta]].$$

Par symétrie $\mathbb{E}[Y_L \mathbf{1}[|Y_L| \leq \delta]] = 0$. En remarquant que $\Delta(x_0) - \lambda\delta = \Delta(x_0)/2$, on conclut que

$$\mathbb{E}[\Delta(u - Y_L) \mathbf{1}[|Y_L| \leq \delta]] \geq \frac{\Delta(x_0)}{2} \mathbb{P}[|Y_L| \leq \delta] \geq \frac{\Delta(x_0)}{2} \left(1 - \frac{4}{\delta L\pi}\right), \quad (16)$$

où on a utilisé (13) à la dernière inégalité. Pour Y_L hors de $[-\delta, \delta]$, on utilise $\Delta(y) \geq \Delta(x_1^-) \geq -\Delta(x_0)$ et on obtient

$$\begin{aligned} \mathbb{E}[\Delta(u - Y_L)\mathbf{1}[|Y_L| > \delta]] &\geq -\Delta(x_0)\mathbb{P}[|Y_L| > \delta] \\ &\geq -\Delta(x_0)\mathbb{P}[|Y_L| > \delta] \geq -\Delta(x_0)\frac{4}{\delta L\pi}, \end{aligned} \quad (17)$$

où on a utilisé (13) encore une fois. En combinant (15), (16) et (17), on a

$$\Delta_L(x_0 + \delta) \geq \frac{\Delta(x_0)}{2} - \frac{6\Delta(x_0)}{\delta L\pi} = \frac{\Delta(x_0)}{2} - \frac{12\lambda}{L\pi},$$

en rappelant qu'on avait choisi $\delta = \Delta(x_0)/(2\lambda)$. Ceci se réécrit

$$\Delta(x_0) \leq 2\Delta_L(x_0 + \delta) + \frac{24\lambda}{L\pi},$$

ce qui prouve le lemme en utilisant l'équation (14). \square

Démonstration du théorème 7.1. On peut supposer $\sigma = 1$ sans perte de généralité (quite à remplacer X_i par $X_i/\sqrt{\sigma}$). Comme Z a une densité bornée par $1/\sqrt{2\pi}$ sur \mathbb{R} , on peut appliquer le lemme 7.3 à $V := S_n/\sqrt{n}$ et $W := Z$. On obtient, pour tout x dans \mathbb{R} et tout $L > 0$,

$$|\mathbb{P}(S_n/\sqrt{n} \leq x) - \mathbb{P}(Z \leq x)| \leq \frac{1}{\pi} \int_{-L}^L \frac{|\varphi_{S_n/\sqrt{n}}(t) - \varphi_Z(t)|}{t} dt + \frac{12\sqrt{2}}{\pi^{3/2}L}. \quad (18)$$

On choisit $L = \sqrt{n}/\rho$, de façon à ce que le deuxième terme soit inférieur à $C_1\rho/\sqrt{n}$ pour $C_1 = \frac{12\sqrt{2}}{\pi^{3/2}}$. Par ailleurs, on a

$$\varphi_{S_n/\sqrt{n}}(t) = \varphi_{S_n}(t/\sqrt{n}) = \varphi_{X_1}(t/\sqrt{n})^n \text{ et } \varphi_Z(t) = \exp(-t^2/2n)^n. \quad (19)$$

Pour borner la différence de ces termes, on utilise l'inégalité

$$|a^n - b^n| \leq n|a - b| \max(|a|, |b|)^{n-1} \quad (20)$$

qui découle de l'identité $a^n - b^n = (a - b)(a^{n-1} + a^{n-2}b + \dots + b^{n-1})$. Par le théorème de Taylor, on a, p.s.,

$$|e^{itX/\sqrt{n}} - (1 - \frac{t^2 X^2}{2n})| \leq \frac{|t|^3 |X|^3}{6n^{3/2}},$$

d'où on déduit, en prenant l'espérance,

$$|\varphi_X(t/\sqrt{n}) - (1 - \frac{t^2}{2n})| \leq \frac{|t|^3 \rho}{6n^{3/2}}. \quad (21)$$

Pour $|t| \leq \sqrt{n}/\rho$, le terme principal $1 - \frac{t^2}{2n}$ est positif (rappelons que $\rho \geq \sigma^3 = 1$), et le terme d'erreur inférieur à $\frac{t^2}{6n}$. On a donc

$$|\varphi_X(t/\sqrt{n})| \leq 1 - \frac{t^2}{2n} + \frac{t^2}{6n} \leq 1 - \frac{t^2}{3n},$$

d'où on déduit pour $n \geq 4$,

$$\max(|\varphi_X(t/\sqrt{n})|, \exp(-t^2/2n))^{n-1} \leq \exp(-\frac{t^2(n-1)}{3n}) \leq \exp(-\frac{t^2}{4}). \quad (22)$$

Par ailleurs, en remarquant que $|e^{-x} - (1-x)| \leq x^2/2$ pour $x \geq 0$ (c'est à nouveau une application du théorème de Taylor), on a, pour $|t| \leq \sqrt{n}/\rho$

$$\begin{aligned} |\varphi_X(t/\sqrt{n}) - \exp(-t^2/2n)| &\leq |\varphi_X(t/\sqrt{n}) - (1 - \frac{t^2}{2n})| + |\exp(-t^2/2n) - (1 - \frac{t^2}{2n})| \\ &\leq \frac{|t|^3 \rho}{6n^{3/2}} + \frac{|t|^4}{8n^2} \leq \frac{|t|^3 \rho}{3n^{3/2}}, \end{aligned} \quad (23)$$

où on a utilisé (21) et $\rho \geq 1$. En utilisant les équations (19), (20), (22) et (23), on obtient, pour $t \leq \sqrt{n}/\rho$, pour $n \geq 4$,

$$|\varphi_{S_n/\sqrt{n}}(t) - \varphi_Z(t)| \leq \frac{|t|^3 \rho}{3n^{1/2}} \exp\left(-\frac{t^2}{4}\right).$$

En utilisant cette inégalité dans l'équation (18) (pour $L = \sqrt{n}/\rho$), on obtient

$$|\mathbb{P}(S_n/\sqrt{n} \leq x) - \mathbb{P}(Z \leq x)| \leq \frac{\rho}{3\pi n^{1/2}} \int_{-L}^L |t|^2 \exp(-\frac{t^2}{4}) dt + \frac{C_1 \rho}{n^{1/2}} \leq \frac{C_2 \rho}{\sqrt{n}},$$

avec

$$C_2 = \frac{1}{3\pi} \int_{-\infty}^{+\infty} |t|^2 \exp(-\frac{t^2}{4}) dt + \frac{12\sqrt{2}}{\pi^{3/2}} \approx 3,80. \quad \square$$

7.1 Exercices

Exercice 7.1. Soit $p = 1/2$ et S_n une variable aléatoire de loi Binomial(n, p). Soit $\alpha > 0$ et z_α l'unique réel tel que $\mathbb{P}(Z \leq z_\alpha) = 1 - \alpha$. Un résultat classique de statistique indique que

$$I_\alpha := \left] -\infty; \frac{S_n}{n} + \frac{z_\alpha}{\sqrt{n}} \right]$$

est un intervalle de confiance de niveau **asymptotique** $1 - \alpha$ pour le paramètre p , c'est-à-dire

$$\lim_{n \rightarrow +\infty} \mathbb{P}(p \in I_\alpha) = 1 - \alpha.$$

En utilisant l'inégalité de Berry–Esseen, donner un encadrement du niveau réel pour $n = 1000$. Même question avec $p = 0.1$.

(On pourra prendre $C = 1/2$ dans le théorème de Berry–Esseen et approximer le calcul de ρ si besoin.)

Exercice 7.2. Soit S_n une marche aléatoire simple sur \mathbb{Z} . En utilisant l'exercice 6.1, trouver un équivalent asymptotique pour

$$\mathbb{P}[S_n \leq 0] - \mathbb{P}[Z \leq 0].$$

En déduire une minoration pour la constante C du théorème 7.1.

Note : Esseen (1956) a montré que $C \geq 0.4$.

8 Grandes déviations

Comme ci-dessus, soit $S_n = X_1 + \dots + X_n$ une marche aléatoire. On suppose que $\sigma^2 := \mathbb{E}(X_1)^2$ appartient à $]0; +\infty[$. Le théorème central limite nous dit que S_n est typiquement $n\mathbb{E}[X_1] + O(\sqrt{n})$. Le but de cette section est d'estimer la probabilité que S_n soit très différent de sa valeur typique. Plus précisément, soit $b > \mathbb{E}[X_1]$, on cherche à estimer $\mathbb{P}[S_n \geq bn]$.

Encore une fois le théorème central limite implique que cette quantité tend vers 0. En effet, quelque soit $A > 0$, on a, pour n suffisamment grand,

$$\mathbb{P}[S_n \geq bn] \leq \mathbb{P}[S_n - \mathbb{E}[S_n] \geq A\sqrt{n}] \rightarrow \frac{1}{\sigma\sqrt{2\pi}} \int_A^{+\infty} \exp(-\sigma^2 t^2/2) dt.$$

La limite du membre de droite peut être rendue aussi petite que voulu en faisant grandir A , donc $\mathbb{P}[S_n \geq bn]$ tend vers 0.

La question est de savoir à quelle vitesse $\mathbb{P}[S_n \geq bn]$ tend vers 0 (polynomiale? exponentielle?).

8.1 Borne supérieure : inégalité de Chernoff

Considérons la *fonction génératrice des moments*⁴ de X_1 défini, pour θ dans \mathbb{R} , par

$$M(\theta) = \mathbb{E}[\exp(\theta X_1)].$$

L'espérance est bien définie car $\exp(\theta X_1)$ est une v.a. positive, mais elle peut être égale à $+\infty$. On notera aussi $\Lambda(\theta) = \log(M(\theta))$. On note que $M(0) = 1$, impliquant $\Lambda(0) = 0$.

Proposition 8.1. *Soit $S_n = X_1 + \dots + X_n$ une somme de v.a. i.i.d. Pour tout réel $b \geq \mathbb{E}[X_1]$, on a*

$$\mathbb{P}[S_n \geq bn] \leq \exp \left[-n \left(\sup_{\theta \in \mathbb{R}} (\theta b - \Lambda(\theta)) \right) \right]$$

La fonction $b \mapsto \sup_{\theta \in \mathbb{R}} (\theta b - \Lambda(\theta))$ est appelée transformée de Legendre-Fenchel de Λ et est notée Λ^* . Pour tout b , on a $\Lambda^*(b) \geq 0$ (obtenu en regardant $\theta = 0$). Si $\Lambda^*(b) > 0$ (ce qui est généralement le cas quand $b > \mathbb{E}[X_1]$, voir exercice 8.3), ce résultat montre que $\mathbb{P}[S_n \geq bn]$ tend exponentiellement vite vers 0, et donne une borne sur le taux de décroissance.

Démonstration. On rappelle l'inégalité de Markov. Si Y est une v.a. positive, alors, quelque soit A , on a $\mathbb{P}[Y \geq A] \leq \mathbb{E}[Y]/A$. Pour tout $\theta \geq 0$, on applique cette inégalité à $Y = \exp(\theta S_n)$ et $A = \exp(\theta bn)$. En remarquant que

$$\mathbb{E}[\exp(\theta S_n)] = \mathbb{E}[\exp(\theta X_1)]^n = M(\theta)^n,$$

4. Voir exercice 8.3 pour comprendre la terminologie.

on trouve

$$\begin{aligned}\mathbb{P}[S_n \geq bn] &= \mathbb{P}[\exp(\theta S_n) \geq \exp(\theta bn)] \leq \mathbb{E}[\exp(\theta S_n)] \exp(-\theta bn) \\ &\leq M(\theta)^n \exp(-\theta bn) \leq \exp[-n(\theta b - \Lambda(\theta))]\end{aligned}$$

Comme ceci est valable pour tout $\theta \geq 0$ et que le membre de gauche ne dépend pas de θ , on peut optimiser la borne supérieure sur θ et on a

$$\mathbb{P}[S_n \geq bn] \leq \exp \left[-n \left(\sup_{\theta \geq 0} (\theta b - \Lambda(\theta)) \right) \right].$$

Il reste à montrer que la fonction $F(\theta) := \theta b - \Lambda(\theta)$ atteint son supremum pour $\theta \geq 0$. Cette fonction vaut 0 en 0, et pour $\theta < 0$, on a :

$$\Lambda(\theta) = \log(\mathbb{E}[\exp(\theta X_1)]) \geq \mathbb{E}(\theta X_1) \geq \theta b,$$

où la première inégalité est une application de l'inégalité de Jensen (pour l'espérance d'une fonction convexe d'une variable) et la seconde vient des hypothèses $b \geq \mathbb{E}[X_1]$ et $\theta < 0$. Ceci implique $F(\theta) \leq 0 = F(0)$ pour $\theta \leq 0$. On a donc

$$\sup_{\theta \geq 0} (\theta b - \Lambda(\theta)) = \sup_{\theta \in \mathbb{R}} (\theta b - \Lambda(\theta)),$$

et la proposition est démontrée. □

Note : par symétrie, on a, pour $b \leq \mathbb{E}[X]$,

$$\mathbb{P}[S_n \leq bn] \leq \exp \left[-n \left(\sup_{\theta \in \mathbb{R}} (\theta b - \Lambda(\theta)) \right) \right]$$

(Dans ce cas, le sup est le même si on restreint au $\theta \leq 0$.)

8.2 Borne inférieure : changement de mesure

Nous allons voir dans cette section que la borne supérieure ci-dessus est quasi optimale. Pour limiter les difficultés techniques, on supposera que X_1 est borné p.s. et on notera K son supremum essentiel, défini par

$$K = \inf\{A : X_1 \leq A \text{ p.s.}\}.$$

Dans ce cadre, les fonctions M et Λ sont bien définies et finies pour tout θ dans \mathbb{R} . On commence par un préliminaire technique.

Lemme 8.2. *La fonction Λ est C^1 et convexe sur \mathbb{R} et, pour tout θ dans \mathbb{R} ,*

$$\Lambda'(\theta) = \frac{1}{M(\theta)} \mathbb{E}[X_1 e^{\theta X_1}]. \quad (24)$$

De plus, on a $\Lambda'(0) = \mathbb{E}[X_1]$ et $\lim_{\theta \rightarrow +\infty} \Lambda'(\theta) = K$.

Démonstration. Soit θ_1 et θ_2 dans \mathbb{R} et α dans $[0; 1]$. On a

$$\begin{aligned}\Lambda(\alpha\theta_1 + (1-\alpha)\theta_2) &= \log\left(\mathbb{E}\left[(e^{\theta_1 X_1})^\alpha (e^{\theta_2 X_1})^{1-\alpha}\right]\right) \\ &\leq \log\left(\mathbb{E}[e^{\theta_1 X_1}]^\alpha \mathbb{E}[e^{\theta_2 X_1}]^{1-\alpha}\right) = \alpha\Lambda(\theta_1) + (1-\alpha)\Lambda(\theta_2),\end{aligned}$$

où l'inégalité centrale est une application de l'inégalité de Hölder (avec $p = 1/\alpha$ et $q = 1/(1-\alpha)$). Ceci prouve la convexité de Λ .

Pour le caractère C^1 de Λ et la formule donnant sa dérivée, on rappelle que $M(\theta) = \mathbb{E}[e^{\theta X_1}]$. Pour θ dans n'importe quel intervalle borné, la quantité $e^{\theta X_1}$ et sa dérivée $X_1 e^{\theta X_1}$ sont continues et bornées. On peut donc appliquer un théorème de dérivation sous l'espérance (une espérance est un cas particulier d'intégrale), et on obtient que M est continue et dérivable sur \mathbb{R} et que

$$M'(\theta) = \mathbb{E}[X_1 e^{\theta X_1}].$$

La formule pour $\Lambda'(\theta)$ suit immédiatement. On observe en particulier que $\Lambda'(0) = \mathbb{E}(X_1)$.

Il reste à montrer que $\lim_{\theta \rightarrow +\infty} \Lambda'(\theta) = K$. Clairement on a

$$\mathbb{E}[X_1 e^{\theta X_1}] \leq K \mathbb{E}[e^{\theta X_1}],$$

ce qui implique $\Lambda'(\theta) \leq K$. Dans l'autre sens, choisissons K_1 et K_2 vérifiant $K_1 < K_2 < K$. On écrit

$$\mathbb{E}[X_1 e^{\theta X_1}] \geq K_1 \mathbb{E}[e^{\theta X_1} \mathbf{1}[X_1 \geq K_1]]. \quad (25)$$

On compare maintenant $\mathbb{E}[e^{\theta X_1} \mathbf{1}[X_1 \geq K_1]]$ à $\mathbb{E}[e^{\theta X_1}]$. La différence peut être bornée ainsi :

$$\left| \mathbb{E}[e^{\theta X_1}] - \mathbb{E}[e^{\theta X_1} \mathbf{1}[X_1 \geq K_1]] \right| = \left| \mathbb{E}[e^{\theta X_1} \mathbf{1}[X_1 < K_1]] \right| \leq e^{\theta K_1}.$$

D'un autre côté, par l'inégalité de Markov, on a :

$$\mathbb{E}[e^{\theta X_1}] \geq \mathbb{P}[X_1 \geq K_2] e^{\theta K_2}.$$

La probabilité ci-dessus est non nulle (par minimalité de K). Comme $K_1 < K_2$, on en déduit que

$$\lim_{\theta \rightarrow +\infty} \frac{\mathbb{E}[e^{\theta X_1} \mathbf{1}[X_1 < K_1]]}{\mathbb{E}[e^{\theta X_1}]} = 0,$$

et donc

$$\lim_{\theta \rightarrow +\infty} \frac{\mathbb{E}[e^{\theta X_1} \mathbf{1}[X_1 \geq K_1]]}{\mathbb{E}[e^{\theta X_1}]} = 1.$$

En reprenant la formule pour Λ' et l'équation 25, on trouve

$$\liminf_{\theta \rightarrow +\infty} \Lambda'(\theta) \geq K_1.$$

Comme ceci est vrai pour tout $K_1 < K$, cette limite inférieure vaut au moins K . On savait déjà que la limite supérieure est au plus K (car $\Lambda'(\theta) \leq K$ pour tout λ), ce qui conclut la démonstration. \square

Théorème 8.3. Soit $S_n = X_1 + \dots + X_n$ une somme de v.a. i.i.d bornées. Pour tout réel b dans l'intervalle $]\mathbb{E}(X_1); K[$ on a

$$\lim_{n \rightarrow +\infty} \frac{1}{n} \log (\mathbb{P}[S_n \geq bn]) = - \sup_{\theta \in \mathbb{R}} (\theta b - \Lambda(\theta)).$$

Démonstration. Le fait que la limite supérieure de $\frac{1}{n} \log (\mathbb{P}[S_n \geq bn])$ est au plus le membre de droite est une conséquence de la proposition 8.1. On va établir ici la borne inférieure.

Notons tout d'abord que Λ' est une fonction continue croissante valant $\mathbb{E}(X_1)$ en 0 et tendant vers K en l'infini. Comme on a pris b dans l'intervalle $]\mathbb{E}(X_1); K[$, il existe θ^* (non nécessairement unique, mais prenons un arbitrairement; θ^* dépend de b) tel que $\Lambda'(\theta^*) = b$. Clairement, θ^* annule la dérivée de la fonction concave $\theta \mapsto \theta b - \Lambda(\theta)$ et est donc la position d'un maximum de cette fonction. Montrer le théorème revient donc à montrer que

$$\liminf_{n \rightarrow +\infty} \frac{1}{n} \log (\mathbb{P}[S_n \geq bn]) \geq -(\theta^* b - \Lambda(\theta^*)).$$

Grâce à la formule (24), on peut écrire (cette formule sera utile ci-dessous) :

$$\frac{1}{M(\theta^*)} \mathbb{E}[X_1 e^{\theta^* X_1}] = b. \quad (26)$$

Notons P la loi de X_1 . On introduit une mesure Q sur \mathbb{R} par la formule $\frac{dQ}{dP}(x) = \frac{e^{\theta^* x}}{M(\theta^*)}$. La mesure Q est en fait une distribution de probabilités car

$$Q(\mathbb{R}) = \int_{\mathbb{R}} dQ(x) = \frac{1}{M(\theta^*)} \int_{\mathbb{R}} e^{\theta^* x} dP(x) = \frac{1}{M(\theta^*)} \mathbb{E}[e^{\theta^* X_1}] = 1.$$

Soit Y_1, \dots, Y_n des variables i.i.d. de loi Q . On a, en utilisant (26),

$$\mathbb{E}[Y_1] = \int_{\mathbb{R}} x dQ(x) = \frac{1}{M(\theta^*)} \int_{\mathbb{R}} x e^{\theta^* x} dP(x) = \frac{1}{M(\theta^*)} \mathbb{E}[X_1 e^{\theta^* X_1}] = b,$$

Par ailleurs, Y_1 est borné et non dégénéré (car X_1 a ces deux propriétés) et donc de variance finie non nulle σ_Y^2 . On considère la somme $T_n = Y_1 + \dots + Y_n$. Le théorème central limite nous dit que

$$\frac{T_n - \mathbb{E}(T_n)}{\sqrt{n}} \rightarrow \sigma_Y Z$$

en distribution. En particulier

$$\lim_{n \rightarrow \infty} \mathbb{P}(T_n - \mathbb{E}(T_n) \in [0; n^{3/4}]) = \mathbb{P}[Z \geq 0] = \frac{1}{2}.$$

Or $\mathbb{E}[T_n] = n\mathbb{E}[Y_1] = nb$. En notant $I_n = [bn; bn + n^{3/4}]$, on a donc montré que $\mathbb{P}(T_n \in I_n)$ tend vers 1/2 et est donc supérieur à 1/3 pour n suffisamment grand.

L'idée est maintenant de relier $\mathbb{P}(T_n \in I_n)$ et $\mathbb{P}(S_n \in I_n)$. On a

$$\begin{aligned} \mathbb{P}(T_n \in I_n) &= \int_{\mathbb{R}^n} \mathbf{1}(y_1 + \dots + y_n \in I_n) dQ(y_1) \dots dQ(y_n) \\ &= \frac{1}{M(\theta^*)^n} \int_{\mathbb{R}^n} \mathbf{1}(x_1 + \dots + x_n \in I_n) e^{\theta^*(x_1 + \dots + x_n)} dP(x_1) \dots dP(x_n). \end{aligned}$$

On note que la quantité à intégrer est plus petite que

$$\mathbf{1}[x_1 + \dots + x_n \in I_n] e^{\theta^*(bn + n^{3/4})}.$$

On a donc

$$\begin{aligned} \mathbb{P}(T_n \in I_n) &\leq \frac{e^{\theta^*(bn + n^{3/4})}}{M(\theta^*)^n} \int_{\mathbb{R}^n} \mathbf{1}(x_1 + \dots + x_n \in I_n) dP(x_1) \dots dP(x_n) \\ &= e^{n(\theta^*b - \Lambda(\theta^*)) + n^{3/4}\theta^*} \mathbb{P}(S_n \in I_n). \end{aligned}$$

En combinant ce résultat avec le fait que $\mathbb{P}[T_n \geq bn] \geq 1/3$ pour n assez grand, on obtient que, pour n assez grand,

$$\mathbb{P}(S_n \geq bn) \geq \mathbb{P}(S_n \in I_n) \geq \frac{1}{3} e^{-n(\theta^*b - \Lambda(\theta^*)) - n^{3/4}\theta^*}.$$

Ceci implique

$$\liminf_{n \rightarrow +\infty} \frac{1}{n} \log(\mathbb{P}[S_n \geq bn]) \geq -(\theta^*b - \Lambda(\theta^*)),$$

et finit la démonstration. \square

Note : la conclusion du théorème 8.3 est en fait vraie sans l'hypothèse que X_1 est bornée. On peut aussi considérer des événements du type « $S_n/n \in A$ » et obtenir la limite de $\frac{1}{n} \log(\mathbb{P}[S_n/n \in A])$ (ou seulement des bornes, selon la topologie de A). Cela s'appelle la *théorie des grandes déviations*.

8.3 Exercices

Exercice 8.1. Considérons la marche aléatoire simple $S_n = X_1 + \dots + X_n$. En utilisant la formule de Stirling, trouver un équivalent asymptotique de $\mathbb{P}[S_n = n/2]$ (on suppose que n est un multiple de 4). Comparer avec l'équivalent de $\log(\mathbb{P}[S_n \geq n/2])$ donné ci-dessus.

Exercice 8.2. Soit P une variable aléatoire suivant une loi de Poisson de paramètre 1. Trouver une borne supérieure pour $\mathbb{P}[P \geq A + 1]$, qui décroît exponentiellement vite en A .

Exercice 8.3. On utilise les notations de cette section. Supposons que $M(\theta) < +\infty$ pour tout θ dans un intervalle $[-\varepsilon; \varepsilon]$ ($\varepsilon > 0$). Montrer que, pour tout $p > 0$, on a $\mathbb{E}[|X_1|^p] < +\infty$ et

$$E[X_1^p] = M^{(p)}(0).$$

En déduire que si $b > \mathbb{E}[X_1]$, alors $\theta^* > 0$ et $\Lambda^*(b) > 0$.

(Indice : pour la première partie, on pourra développer en série $\mathbb{E}[e^{zX_1}]$ autour de $z = 0$; pour la seconde partie, regarder le développement de $b\theta - \Lambda(b)$ autour de $b = 0$.)

Exercice 8.4. Soit X_1, X_2, \dots des v.a. i.i.d. centrées bornées à valeurs dans \mathbb{Z} de période maximale 1. On utilise les notations du théorème 8.3 et de sa preuve, en particulier on a choisi b dans $] \mathbb{E}[X_1], K[$. On supposera dans la suite nb entier.

1. Comparer $\mathbb{P}[S_n = nb]$ et $\mathbb{P}[T_n = nb]$.
2. En déduire un équivalent asymptotique pour $\mathbb{P}[S_n = nb]$.

9 Convergence fonctionnelle

Comme précédemment, soit X_i une suite de v.a. i.i.d. et $S_n = X_1 + \dots + X_n$ pour $n \geq 0$. On veut voir la marche aléatoire $(S_n)_{n \geq 0}$ comme une fonction du « temps » n . Pour ça, on définit, pour t dans $[0, 1]$,

$$W_n(t) = \frac{1}{\sqrt{n}} (S_{[nt]} + \{nt\}(S_{[nt]+1} - S_{[nt]})),$$

ou $[nt]$ et $\{nt\}$ désignent les parties entière et fractionnaire de nt , respectivement. En d'autres termes, $W_n(t) = \frac{1}{\sqrt{n}} S_{nt}$ si nt est entier, et W_n est affine sur chaque segment $[i/n; (i+1)/n]$.

Le but de cette section est de trouver une limite (aléatoire) de W_n en tant que fonction continue, i.e. dans l'espace $\mathcal{C}([0, 1], \mathbb{R})$ des fonctions continues à valeurs dans \mathbb{R} munie de la norme infinie.

9.1 Processus aléatoires et lois fini-dimensionnelles

Soit F une fonction aléatoire dans $\mathcal{C}([0, 1], \mathbb{R})$ (on parlera aussi de processus aléatoire) et $t_1 < \dots < t_d$ des points dans $[0, 1]$. Le vecteur $(F(t_1), \dots, F(t_d))$ est alors un vecteur aléatoire dans \mathbb{R}^d . Les lois de tels vecteurs (quand d et $t_1 < \dots < t_d$ varient) sont appelées lois fini-dimensionnelles de F . La proposition suivante montre que la loi de F est entièrement déterminée par ses lois fini-dimensionnelles. La preuve est basée sur un résultat standard de théorie de la mesure, qui dit que deux mesures qui coïncident sur un π -système (ensemble d'ensembles mesurables stable par intersection finie) coïncide sur la tribu engendrée par ce π -système.

Proposition 9.1. *Soient F et G deux processus aléatoires dans $\mathcal{C}([0, 1], \mathbb{R})$. Les assertions suivantes sont équivalentes :*

1. $F \stackrel{\text{loi}}{=} G$.
2. Pour tout $d \geq 1$ et tous $t_1 < \dots < t_d$ dans $[0, 1]$, on a l'égalité en distribution

$$(F(t_1), \dots, F(t_d)) \stackrel{\text{loi}}{=} (G(t_1), \dots, G(t_d)).$$

Démonstration. 1. implique 2. est immédiat car si on fixe $d \geq 1$ et $t_1 < \dots < t_d$ dans $[0; 1]$, l'application π_{t_1, \dots, t_d} qui à F associe $(F(t_1), \dots, F(t_d))$ est continue et donc mesurable sur $\mathcal{C}([0, 1], \mathbb{R})$.

Montrons que 2. implique 1. Soient F et G deux processus aléatoires vérifiant 2. Les distributions P_F et P_G coïncident alors sur les ensembles du type $\pi_{t_1, \dots, t_d}^{-1}(B)$, où B est un Borélien de \mathbb{R} . Comme ceci forme un π -système, P_F et P_G coïncident sur la tribu \mathcal{A} engendrée par ces ensembles. Or pour f une fonction continue et $R > 0$,

$$\begin{aligned} \overline{B(f, R)} &= \{g : \|g - f\|_\infty \leq R\} = \bigcap_{q \in \mathbb{Q}} \{g : |g(q) - f(q)| < R\} \\ &= \bigcap_{q \in \mathbb{Q}} \pi_q^{-1}([f(q) - R; f(q) + R]) \\ &= \bigcap_{d \geq 1} \pi_{q_1, \dots, q_d}^{-1}([f(q_1) - R; f(q_1) + R] \times \dots \times [f(q_d) - R; f(q_d) + R]), \end{aligned}$$

où, dans la dernière expression, q_1, q_2, \dots est une énumération des rationnels. On en déduit que \mathcal{A} contient les boules fermées de l'espace métrique $\mathcal{C}([0, 1], \mathbb{R})$, et donc \mathcal{A} est la tribu des Boréliens sur cet ensemble. Comme P_F et P_G coïncident sur \mathcal{A} , on a $P_F = P_G$, comme annoncé. \square

Considérons maintenant des processus aléatoires $(F_n)_{n \geq 0}$ et F . Nous allons comparer la convergence des lois fini-dimensionnelles des F_n vers celles de F à la convergence en loi de F_n vers F (dans $\mathcal{C}([0, 1], \mathbb{R})$). Comme ci-dessus, il y a un sens facile.

Lemme 9.2. *Soient $(F_n)_{n \geq 1}$ et F des processus aléatoires dans $\mathcal{C}([0, 1], \mathbb{R})$. Supposons que F_n converge en loi vers F . Alors, pour tout $d \geq 1$ et tous temps $t_1 < \dots < t_d$ dans $[0; 1]$, on a la convergence suivante en loi dans \mathbb{R}^d :*

$$(F_n(t_1), \dots, F_n(t_d)) \xrightarrow{\text{loi}} (F(t_1), \dots, F(t_d)).$$

Démonstration. Immédiat, car, pour tout $d \geq 1$ et tous $t_1 < \dots < t_d$ dans $[0; 1]$, l'application π_{t_1, \dots, t_d} est continue. \square

La réciproque n'est pas vraie en général (pour des fonctions déterministes, la convergence point par point n'entraîne pas la convergence uniforme!), mais elle l'est si on suppose que la suite F_n est une suite de v.a. tendue dans $\mathcal{C}([0, 1], \mathbb{R})$. On rappelle (voir Section 4) que une suite (F_n) est dite tendue si, pour tout $\varepsilon > 0$, il existe un sous-ensemble compact K de $\mathcal{C}([0, 1], \mathbb{R})$ tel que

$$\text{pour tout } n, \mathbb{P}[F_n \in K] \geq 1 - \varepsilon.$$

Par ailleurs, le théorème de Prokhorov nous dit que si une suite F_n est tendue, alors elle admet une sous-suite $F_{\varphi(n)}$ qui converge en loi.

Proposition 9.3. *Soient $(F_n)_{n \geq 1}$ des processus aléatoires dans $\mathcal{C}([0, 1], \mathbb{R})$. Supposons :*

1. Pour tout $d \geq 1$ et tous $t_1 < \dots < t_d$ dans $[0; 1]$, le vecteur $(F_n(t_1), \dots, F_n(t_d))$ converge en loi.

2. F_n est tendue.

Alors il existe un processus aléatoire F tel que F_n converge en loi vers F .

Démonstration. On montre que la suite F_n n'a qu'une valeur d'adhérence, i.e. qu'il existe F tel que toute sous-suite $F_{\varphi(n)}$ qui converge doit converger vers F .

Comme F_n est tendue, elle admet une sous-suite $F_{\psi(n)}$ qui converge vers un certain F . Prenons maintenant une autre sous-suite $F_{\varphi(n)}$ qui converge vers un certain G . Soit $d \geq 1$ et $t_1 < \dots < t_d$ dans $[0; 1]$. D'après le lemme 9.2, on a les convergences suivantes (en loi dans \mathbb{R}^d)

$$\begin{aligned} (F_{\psi(n)}(t_1), \dots, F_{\psi(n)}(t_d)) &\rightarrow (F(t_1), \dots, F(t_d)) \\ (F_{\varphi(n)}(t_1), \dots, F_{\varphi(n)}(t_d)) &\rightarrow (G(t_1), \dots, G(t_d)) \end{aligned}$$

Mais $(F_{\psi(n)}(t_1), \dots, F_{\psi(n)}(t_d))$ et $(F_{\varphi(n)}(t_1), \dots, F_{\varphi(n)}(t_d))$ sont des sous-suites de la même suite $(F_n(t_1), \dots, F_n(t_d))$, que l'on a supposé convergente en loi (hypothèse 1). Elles ont donc la même limite en loi (celle de la suite complète), i.e.

$$(F(t_1), \dots, F(t_d)) \stackrel{\text{loi}}{=} (G(t_1), \dots, G(t_d)).$$

Comme ceci est vrai pour tout $d \geq 1$ et $t_1 < \dots < t_d$, la proposition 9.1 implique que F et G ont la même loi.

Cela montre que F_n a un unique point d'accumulation F . Comme F_n est tendue, on peut utiliser le corollaire 4.6 et conclure que F_n converge en loi vers F . \square

9.2 Compacité et tension dans $\mathcal{C}([0, 1], \mathbb{R})$

La difficulté pour appliquer la proposition 9.3 est que la définition de suite tendue de fonctions aléatoires n'est pas très explicite. Nous allons donner ici des conditions suffisantes impliquant la tension.

La première étape consiste à décrire les sous-ensembles compacts de $\mathcal{C}([0, 1], \mathbb{R})$. Pour cela rappelons que pour f dans $\mathcal{C}([0, 1], \mathbb{R})$ et ε , on définit

$$\omega(f; \delta) = \sup_{x, y: |x-y| \leq \delta} |f(x) - f(y)|.$$

La fonction $\delta \mapsto \omega(f, \delta)$ est appelée « module de continuité de f ». Comme f est uniformément continu (car continu sur un compact), $\omega(f, \delta)$ tend vers 0 quand δ tend vers 0. On admettra le résultat d'analyse fonctionnelle suivant.

Theorème 9.4 (Arzelà-Ascoli). *Soit \mathcal{F} un sous-ensemble de $\mathcal{C}([0, 1], \mathbb{R})$. Sa clôture $\overline{\mathcal{F}}$ est compacte si et seulement si*

- (i) $\{f(0) : f \in \mathcal{F}\}$ est borné ;
- (ii) $\sup_{f \in \mathcal{F}} \omega(f, \delta)$ tend vers 0 quand δ tend vers 0.

On en déduit aisément un premier critère de tension.

Proposition 9.5. *Soit $(F_n)_{n \geq 1}$ une suite de fonctions aléatoires dans $\mathcal{C}([0, 1], \mathbb{R})$. On suppose que F_n vérifie les propriétés suivantes :*

- (a) *La suite de v.a. réelles $F_n(0)$ est tendue ;*
- (b) *Pour tout $\varepsilon > 0$, la quantité $\sup_{n \geq 1} \mathbb{P}[\omega(F_n, \delta) \geq \varepsilon]$ tend vers 0 quand δ tend vers 0.*

Alors la suite $(F_n)_{n \geq 1}$ est tendue dans $\mathcal{C}([0, 1], \mathbb{R})$.

Démonstration. Soit $\varepsilon > 0$. Comme $F_n(0)$ est tendue, il existe $A > 0$ tel que, pour tout $n \geq 0$,

$$\mathbb{P}[|F_n(0)| \leq A] \geq 1 - \frac{\varepsilon}{2}.$$

D'autre part, d'après (b), pour tout $p \geq 1$, il existe $\delta_p > 0$ tel que, pour tout n ,

$$\mathbb{P}[\omega(F_n, \delta_p) \leq 1/p] \geq 1 - \frac{\varepsilon}{2^{p+1}}.$$

Regardons l'ensemble

$$K := \left\{ f \in \mathcal{C}([0, 1], \mathbb{R}) : |f(0)| \leq A \text{ et } \omega(f, \delta_p) \leq 1/p \text{ pour tout } p \geq 1. \right\}.$$

Pour tout n , on a $\mathbb{P}[F_n \in K] \geq 1 - \varepsilon$. Par ailleurs, d'après le théorème d'Arzelà-Ascoli, K est un sous-ensemble compact de $\mathcal{C}([0, 1], \mathbb{R})$. Comme on peut trouver un tel K pour tout $\varepsilon > 0$, la suite F_n est tendue. \square

Nous allons maintenant établir un critère de moments pour la tension.

Théorème 9.6. *Soit F_n une suite de fonctions aléatoires vérifiant les propriétés suivantes.*

- (i) *la suite $F_n(0)$ est tendue dans \mathbb{R} ;*
- (ii) *Il existe $\alpha > 1$, $\beta > 0$ et $\gamma > 0$ tel que, pour tout s, t dans $[0, 1]$ et pour tout $n \geq 1$, on ait*

$$\mathbb{E}(|F_n(s) - F_n(t)|^\alpha) \leq \beta |s - t|^{\gamma+1}. \quad (27)$$

Alors F_n est une suite tendue dans $\mathcal{C}([0, 1], \mathbb{R})$.

Pour démontrer ce théorème, on va utiliser la proposition 9.5. La première hypothèse dans ces deux énoncés étant identiques, il faut montrer que sous les hypothèses du théorème, l'hypothèse (b) de la proposition 9.5 est vérifiée.

Pour cela on va utiliser l'inégalité de Markov (comme dans la méthode des moments) et écrire :

$$\mathbb{P}[\omega(F_n; \delta) \geq \varepsilon] \leq \frac{\mathbb{E}[\omega(F_n; \delta)^\alpha]}{\varepsilon^\alpha}.$$

et on voudrait relier $\mathbb{E}[\omega(F_n; \delta)^\alpha]$ aux quantités $\mathbb{E}(|F_n(s) - F_n(t)|^\alpha)$. En général, contrôler l'espérance d'un supremum/maximum est difficile. Dans le cas d'un maximum sur un nombre fini de variables, une borne supérieure triviale, mais cependant souvent utile, est l'espérance de la somme.

On commence donc par le lemme suivant qui permet de ramener le supremum dans le calcul du module de continuité à un nombre fini de valeurs.

Lemme 9.7. Pour $r \geq 0$ et f une fonction continue, soit

$$U_p(f) := \sup_{0 \leq q < 2^p} \left| f\left(\frac{q}{2^p}\right) - f\left(\frac{q+1}{2^p}\right) \right|,$$

où le supremum est pris sur les entiers q entre 0 et 2^n . Alors

$$\omega(f; 2^{-r}) \leq 2 \sum_{p \geq r} U_p(f),$$

où le membre de droite peut être infini (auquel cas l'énoncé est bien sûr trivialement vrai).

Démonstration. Soit t dans $[0, 1]$. Pour $m \geq r$, on définit $t_m = \lfloor \frac{t2^m}{2^m} \rfloor$, i.e. t_m est la meilleure approximation de t par en-dessous par un nombre de la forme $q/2^m$. On a

$$|f(t_m) - f(t_n)| \leq \sum_{p=r}^{m-1} |f(t_{p+1}) - f(t_p)| \leq \sum_{p=r}^{m-1} U_{p+1}(f),$$

où la dernière inégalité est dû au fait que t_{p+1} et t_p sont soit égaux (dans le cas où $t - t_p < 2^{p+1}$), soit deux nombres consécutifs de la forme $q/2^{p+1}$ (si $t - t_p \geq 2^{p+1}$).

Quand m tend vers l'infini, t_m tend vers t . En utilisant la continuité de f , on a

$$|f(t) - f(t_n)| \leq \sum_{p=r}^{+\infty} U_{p+1}(f).$$

Prenons maintenant s et t dans $[0, 1]$ vérifiant $|s - t| \leq 2^{-n}$. Les approximations s_r et t_r de s et t sont soit égales, soit deux nombres consécutifs de la forme $q/2^r$, ce qui implique $|f(s_r) - f(t_r)| \leq U_r(f)$. On a donc

$$|f(t) - f(s)| \leq |f(t) - f(t_r)| + |f(s) - f(s_r)| + |f(s_r) - f(t_r)| \leq 2 \sum_{p \geq r} U_p(f).$$

Comme cette inégalité est vraie pour tout s et t dans $[0, 1]$ vérifiant $|s - t| \leq 2^{-r}$, le lemme est démontré. \square

Démonstration du théorème 9.6. Soit F_n une suite de fonctions aléatoires vérifiant les hypothèses du théorème. On veut montrer que l'item (b) de la proposition 9.5 est vérifiée.

On rappelle que $\|X\|_\alpha = \mathbb{E}[|X|^\alpha]^{1/\alpha}$ définit une norme sur les v.a. réelles d'un espace de proba. Pour δ fixé, $\omega(F_n, \delta)$ et $U_p(F_n)$ sont des v.a. réelles est le lemme 9.7 implique que, p.s., on a

$$\omega(F_n; 2^{-r}) \leq 2 \sum_{p \geq r} U_p(F_n).$$

En prenant la α -norme, et en appliquant l'inégalité triangulaire, on obtient

$$\|\omega(F_n; 2^{-r})\|_\alpha \leq 2 \sum_{p \geq r} \|U_p(F_n)\|_\alpha. \quad (28)$$

Or, p.s. on a

$$U_p(F_n)^\alpha \leq \sum_{0 \leq q < 2^p} \left| f\left(\frac{q}{2^p}\right) - f\left(\frac{q+1}{2^p}\right) \right|^\alpha,$$

ce qui implique :

$$\|U_p(F_n)\|_\alpha^\alpha = \mathbb{E}[U_p(F_n)^\alpha] \leq \sum_{0 \leq q < 2^p} \mathbb{E}\left[\left| f\left(\frac{q}{2^p}\right) - f\left(\frac{q+1}{2^p}\right) \right|^\alpha\right].$$

En utilisant l'hypothèse (30), on obtient :

$$\|U_p(F_n)\|_\alpha^\alpha \leq 2^p \beta 2^{-p(\gamma+1)} = \beta 2^{-p\gamma}. \quad (29)$$

En revenant à (28), cela donne

$$\|\omega(F_n; 2^{-r})\|_\alpha \leq 2\beta^{1/\alpha} \sum_{p \geq r} 2^{-p\gamma/\alpha} = \frac{2^{1-\frac{r\gamma}{\alpha}} \beta^{1/\alpha}}{1 - 2^{-\gamma/\alpha}}.$$

L'inégalité de Markov implique ensuite que, pour tout $\varepsilon > 0$

$$\mathbb{P}[\omega(F_n; 2^{-r}) \geq \varepsilon] \leq \frac{\mathbb{E}[\omega(F_n; 2^{-r})^\alpha]}{\varepsilon^\alpha} \leq \varepsilon^{-\alpha} \left(\frac{2^{1-\frac{r\gamma}{\alpha}} \beta^{1/\alpha}}{1 - 2^{-\gamma/\alpha}} \right)^\alpha.$$

Cette inégalité est vraie pour tout $n, r \geq 1$. Ce qui est important est que le membre de droite est indépendant de n , et tend vers 0 quand r tend vers l'infini (à ε fixé). On a donc

$$\sup_{n \geq 1} \mathbb{P}[\omega(F_n; 2^{-r}) \geq \varepsilon] \xrightarrow{r \rightarrow +\infty} 0,$$

ce qui implique l'hypothèse (b) de la proposition 9.5 (car $\omega(F_n; \delta)$ est croissant en δ). \square

Finalement, voici un dernier critère de tension, spécifique a des fonctions affines par morceaux.

Proposition 9.8. *Soit F_n une suite de fonctions aléatoires. On suppose que pour tout n , F_n est p.s. affine sur chacun des intervalles $[(i-1)/n; i/n]$ et que les hypothèses suivantes sont satisfaites :*

- (i) *la suite $F_n(0)$ est tendue dans \mathbb{R} ;*
- (ii) *Il existe $\alpha > 1$, $\beta > 0$ et $\gamma > 0$ tel que $\alpha \geq 1 + \gamma$, pour tout s, t dans $[0, 1]$ et pour tout $n \geq 1$ tels que sn et tn soient entiers, on ait*

$$\mathbb{E}(|F_n(s) - F_n(t)|^\alpha) \leq \beta |s - t|^{1+\gamma}. \quad (30)$$

Alors F_n est une suite tendue dans $\mathcal{C}([0, 1], \mathbb{R})$.

Démonstration. Il suffit de montrer que (30) est en fait valable pour tout s et t dans $[0, 1]$, quitte à modifier la constante β . Soit $s < t$ dans $[0, 1]$. On distingue deux cas :

- Si s et t sont dans un même intervalle $[\frac{i}{n}, \frac{i+1}{n}]$, alors comme F_n est affine sur cet intervalle, on a

$$|F_n(t) - F_n(s)| \leq n(t-s)|F_n(\frac{i+1}{n}) - F_n(\frac{i}{n})|.$$

Cela implique, en utilisant (30) aux points $\frac{i}{n}$ et $\frac{i+1}{n}$,

$$\mathbb{E}(|F_n(s) - F_n(t)|^\alpha) \leq n^\alpha(t-s)^\alpha \beta n^{-1-\gamma} = \beta(t-s)^{1+\gamma}(n(t-s))^{\alpha-1-\gamma}.$$

Comme $t-s \leq 1/n$, le dernier facteur est inférieur ou égal à 1, et (30) est vérifiée.

- Considérons maintenant le cas où $\lfloor ns \rfloor < \lfloor nt \rfloor$. En utilisant l'équivalence des normes $\|\cdot\|_1$ et $\|\cdot\|_\alpha$ sur \mathbb{R}^3 , on sait qu'il existe une constante C_α telle que, pour tout $a, b, c \geq 0$, on a

$$(a+b+c)^\alpha \leq C_\alpha(a^\alpha + b^\alpha + c^\alpha).$$

En particulier,

$$\begin{aligned} |F_n(s) - F_n(t)|^\alpha &\leq C_\alpha \left(|F_n(s) - F_n(\lceil ns \rceil)|^\alpha \right. \\ &\quad \left. + |F_n(\lceil ns \rceil) - F_n(\lfloor nt \rfloor)|^\alpha + |F_n(\lfloor nt \rfloor) - F_n(t)|^\alpha \right). \end{aligned}$$

On prend maintenant l'espérance, et on peut appliquer (30) aux trois termes du membre de droite : pour le terme central, c'est possible car on considère des valeurs de F_n en des points $i/n, j/n$ (avec i, j entiers), alors que pour les premier et dernier termes, c'est possible grâce au cas traité ci-dessus. On obtient

$$\mathbb{E}(|F_n(s) - F_n(t)|^\alpha) \leq C_\alpha \beta (|s - \lceil ns \rceil|^{1+\gamma} + |\lceil ns \rceil - \lfloor nt \rfloor|^{1+\gamma} + |\lfloor nt \rfloor - t|^{1+\gamma}).$$

En invoquant une nouvelle fois l'équivalence des normes sur \mathbb{R}^3 , on sait qu'il existe une constante $c_{1+\gamma}$ telle que, pour tout $a, b, c \geq 0$, on a

$$(a^{1+\gamma} + b^{1+\gamma} + c^{1+\gamma}) \leq c_{1+\gamma}(a+b+c)^{1+\gamma}.$$

Ceci permet de conclure

$$\mathbb{E}(|F_n(s) - F_n(t)|^\alpha) \leq C_\alpha \beta c_{1+\gamma} |s - t|^{1+\gamma}.$$

Cette dernière inégalité est bien de la forme (30) avec une constante β différente, ce qui conclut la démonstration. \square

9.3 Application à la marche aléatoire – théorème de Donsker

On rappelle qu'on a défini

$$W_n(t) = \frac{1}{\sqrt{n}} (S_{\lfloor nt \rfloor} + \{nt\}(S_{\lfloor nt \rfloor + 1} - S_{\lfloor nt \rfloor})),$$

ou $S_n = X_1 + \dots + X_n$ est une marche aléatoire centrée. On suppose $\mathbb{E}[X_1^4] < +\infty$. Cela implique en particulier que $\sigma^2 := \mathbb{E}[X_1^2] < +\infty$.

Théorème 9.9 (Theoreme de Donsker). *Avec les hypothèses ci-dessus, il existe une processus aléatoire $B = (B_t)_{t \in [0;1]}$ telle que W_n converge vers σB en distribution dans $\mathbb{C}([0;1], \mathbb{R})$.*

La notation B_t au lieu de $B(t)$ pour l'image de t par la fonction (aléatoire) B peut surprendre mais elle est standard dans ce contexte. On verra (voir Corollaire 9.10) que B ne dépend pas de la distribution du pas X_1 de la marche aléatoire considérée. B est appelé mouvement Brownien sur $[0;1]$.

Démonstration. Il suffit de montrer que W_n vérifie les hypothèses de la proposition 9.3, i.e. la convergence des lois finis dimensionnelles et la tension.

Commençons par la convergence des lois fini-dimensionnelles. Soient $d \geq 1$ et $t_1 < \dots < t_d$ dans $[0;1]$. On veut montrer la convergence de

$$(W_n(t_1), W_n(t_2), \dots, W_n(t_d)). \quad (31)$$

On remarque que, pour tout t , la différence $W_n(t) - W_n(\frac{\lfloor nt \rfloor}{n})$ tend vers 0 en distribution. On peut donc regarder à la place de (31) le vecteur

$$(W_n(\frac{\lfloor nt_1 \rfloor}{n}), W_n(\frac{\lfloor nt_2 \rfloor}{n}), \dots, W_n(\frac{\lfloor nt_d \rfloor}{n})).$$

En appliquant une simple transformation linéaire, on voit qu'il est équivalent de prouver la convergence du vecteur

$$(W_n(\frac{\lfloor nt_1 \rfloor}{n}), W_n(\frac{\lfloor nt_2 \rfloor}{n}) - W_n(\frac{\lfloor nt_1 \rfloor}{n}), \dots, W_n(\frac{\lfloor nt_d \rfloor}{n}) - W_n(\frac{\lfloor nt_{d-1} \rfloor}{n})). \quad (32)$$

Or, pour tout i (avec la convention $t_0 = 0$), on a

$$W_n(\frac{\lfloor nt_i \rfloor}{n}) - W_n(\frac{\lfloor nt_{i-1} \rfloor}{n}) = \frac{1}{\sqrt{n}} (S_{\lfloor nt_i \rfloor} - S_{\lfloor nt_{i-1} \rfloor}) = \frac{1}{\sqrt{n}} \left(\sum_{j=\lfloor nt_{i-1} \rfloor + 1}^{\lfloor nt_i \rfloor} X_j \right).$$

Pour des i différents, les ensembles de X_j intervenant dans la formule sont disjoints et les $(W_n(\frac{\lfloor nt_i \rfloor}{n}) - W_n(\frac{\lfloor nt_{i-1} \rfloor}{n}))$ sont donc indépendants. Prouver la convergence du vecteur (32) revient donc à prouver la convergence de chaque composante. Par ailleurs, par le théorème central limite (pour des tableaux triangulaires), la quantité

$$\frac{1}{\sqrt{t_i - t_{i-1}}} (W_n(\frac{\lfloor nt_i \rfloor}{n}) - W_n(\frac{\lfloor nt_{i-1} \rfloor}{n})) = \frac{1}{\sqrt{n(t_i - t_{i-1})}} \left(\sum_{j=\lfloor nt_{i-1} \rfloor + 1}^{\lfloor nt_i \rfloor} X_j \right)$$

tend vers σZ . Ceci conclut la preuve de la convergence des lois fini-dimensionnelles de W_n .

On veut maintenant prouver que W_n est une suite tendue de processus aléatoire. Comme, pour tout n , on a $W_n(0) = 0$ p.s., il suffit de vérifier l'équation (30). Pour $s = i/n < t = j/n$, on a

$$\begin{aligned} \mathbb{E}(|W_n(s) - W_n(t)|^4) &= \frac{1}{n^2} \mathbb{E} \left(\left| \sum_{k=i+1}^j X_k \right|^4 \right) \\ &= \frac{1}{n^2} \left((j-i) \mathbb{E}(X_1^4) + 6 \binom{j-i}{2} \mathbb{E}(X_1^2)^2 \right), \end{aligned}$$

ou la deuxième égalité s'obtient en développant $\left(\sum_{k=i+1}^j X_k \right)^4$ et en observant que seuls les termes de la forme X_j^4 ou $X_j^2 X_j^2$, ont une espérance non nulle. En substituant $i = sn$ et $j = tn$ et en utilisant $|t-s| \geq 1/n$ on obtient

$$\begin{aligned} \mathbb{E}(|W_n(s) - W_n(t)|^4) &\leq \frac{1}{n} (t-s) \mathbb{E}(X_1^4) + 6 \mathbb{E}(X_1^2)^2 (t-s)^2 \\ &\leq (\mathbb{E}(X_1^4) + 6 \mathbb{E}(X_1^2)^2) |t-s|^2. \end{aligned}$$

Cette inégalité est bien de la forme (30) avec $\alpha = 4$ et $\gamma = 1$ et le critère de tension de la Proposition 9.8 est vérifié.

La suite de processus aléatoires W_n tend donc en distribution vers un processus W . En posant $B = W/\sigma$, on obtient l'énoncé du théorème. \square

Corollaire 9.10. *Le processus aléatoire B_t est caractérisé (à égalité en distribution près) par les propriétés suivantes :*

- $t \mapsto B_t$ est continu p.s. ;
- $B_0 = 0$ p.s. ;
- pour tout $t_1 < t_2 < \dots < t_d$ dans $[0; 1]$, les quantités $B_{t_i} - B_{t_{i-1}}$ sont des Gaussiennes indépendantes de variances respectives $t_i - t_{i-1}$ (pour i dans $\{1, \dots, d\}$ avec la convention $t_0 = 0$).

Démonstration. Le fait que B_t vérifie 1. et 2. est clair par construction. D'après le théorème 9.9 et le lemme 9.2, les lois fini-dimensionnelles de σB sont les limites de celles de W_n . Ces limites ont été déterminées (implicitement) dans la preuve du théorème 9.9, prouvant que B_t vérifient la propriété 3. ci-dessus. L'unicité découle de la proposition 9.1 \square

Remark 9.11. Construire directement un processus aléatoire B_t vérifiant les propriétés du corollaire 9.10 n'est pas évident. De manière remarquable, on n'a pas besoin de le faire : la preuve du théorème de Donsker assure l'existence d'une tel processus.

Remark 9.12. L'hypothèse $\mathbb{E}[X_1^4] < +\infty$ est superflue. Le théorème de Donsker nécessite en fait seulement $\sigma^2 := \mathbb{E}[X_1^2] < +\infty$, mais la preuve est plus complexe (elle utilise le critère de tension via le module de continuité).

9.4 Exercices

Exercice 9.1. Soit F et G les fonctions aléatoires définies ainsi :

- Avec probabilité $1/2$, on a $F(x) \equiv 0$. Sinon $F(x) \equiv x - \frac{1}{2}$. (\equiv signifie « égal pour tout x dans $[0, 1]$ »)
 - Avec probabilité $1/2$, on a $G(x) \equiv \min(x - \frac{1}{2}, 0)$. Sinon $G(x) \equiv \max(x - \frac{1}{2}, 0)$.
1. Soit t dans $[0, 1]$. Montrer que $F(t) \stackrel{\text{loi}}{=} G(t)$.
 2. Trouver $s < t$ dans $[0, 1]$ tels que $(F(s), F(t)) \stackrel{\text{loi}}{\neq} (G(s), G(t))$.

Exercice 9.2. Soit $d > 0$ et $(B_i)_{i \leq 2d}$ des v.a. i.i.d. de distribution

$$\mathbb{P}[B_i = -1] = \mathbb{P}[B_i = 1] = \frac{1}{2}.$$

On définit des fonctions aléatoires H_d et K_d telles que

$$\begin{cases} H_d(\frac{i}{2d}) = K_d(\frac{i}{2d}) = B_i \text{ pour } 1 \leq i \leq 2d; \\ H_d(0) = -K_d(0) = \prod_{i=1}^{2d} B_i. \end{cases}$$

et H_d et K_d sont affines sur chacun des segments $[\frac{i-1}{2d}, \frac{i}{2d}]$ (pour $1 \leq i \leq 2d$).

1. Soit j dans $\{0, 1, \dots, 2d\}$. Montrer que

$$\left(H_d\left(\frac{i}{2d}\right) \right)_{\substack{0 \leq i \leq 2d \\ i \neq j}} \stackrel{\text{loi}}{=} \left(K_d\left(\frac{i}{2d}\right) \right)_{\substack{0 \leq i \leq 2d \\ i \neq j}}.$$

(Notez bien l'ensemble d'indices des vecteurs que l'on veut comparer ; on prend les entiers de 0 à $2d$, y compris 0, auxquels on retire j).

2. En déduire que H_d et K_d ont les mêmes lois fini-dimensionnelles de dimension d , i.e. pour tous $t_1 < \dots < t_d$ dans $[0, 1]$, on a

$$\left(H_d(t_1), \dots, H_d(t_d) \right) \stackrel{\text{loi}}{=} \left(K_d(t_1), \dots, K_d(t_d) \right).$$

3. Montrer que

$$\left(H_d(0), H_d\left(\frac{1}{2d}\right), \dots, H_d\left(\frac{2d-1}{2d}\right), H_d(1) \right) \stackrel{\text{loi}}{\neq} \left(K_d(0), K_d\left(\frac{1}{2d}\right), \dots, K_d\left(\frac{2d-1}{2d}\right), K_d(1) \right).$$

Exercice 9.3. 1. Soit f_n une suite de fonctions (déterministes) dans $\mathcal{C}([0, 1], \mathbb{R})$. Montrer que si $f_n(x)$ converge pour tout x vers $f(x)$ et si

$$\lim_{\delta \rightarrow 0} \sup_{n \geq 1} \omega(f_n; \delta) = 0, \tag{33}$$

alors f_n converge uniformément vers f .

2. Supposons qu'il existe ℓ tel que f_n soit ℓ -Lipshitzienne pour tout n . Montrer que (33) est satisfaite.
3. Soit F_n une suite de processus aléatoires définis sur un même espace de probabilités. On suppose qu'il existe une v.a. L telle que, pour tout n , p.s., F_n est L -Lipshitzienne. Montrer que, pour tout $\varepsilon > 0$, la quantité $\sup_{n \geq 1} \mathbb{P}[\omega(F_n, \delta) \geq \varepsilon]$ tend vers 0 quand δ tend vers 0.

Exercice 9.4. Soit S_n une marche aléatoire centrée avec $\sigma^2 := \mathbb{E}[X_1^2] < +\infty$. Pour tout $n \geq 1$, on pose

$$M_n = \max_{0 \leq k \leq n} S_k.$$

Montrer que M_n/\sqrt{n} converge en distribution vers une variable aléatoire réelle, que l'on décrira à l'aide du mouvement Brownien.

Exercice 9.5. Soit S_n une marche aléatoire centrée à valeurs dans \mathbb{Z} et telle que X_1 a une période maximale 1. On définit W_n comme ci-dessus. Soit t dans $]0, 1[$.

1. Soit a_n dans \mathbb{Z} avec $a_n \sim x\sqrt{n}$ pour x dans \mathbb{R} . Utiliser le théorème local limite pour trouver un équivalent asymptotique de

$$\mathbb{P}[\sqrt{n} W_n(\lfloor \frac{nt}{n} \rfloor) = a_n \mid S_n = 0].$$

2. Sans preuve formelle, pouvez-vous déterminer la loi limite de $W_n(t)$, conditionnellement à $S_n = 0$?

Exercice 9.6. Soit δ dans $]0, 1[$. On rappelle que fonction f de $[0, 1]$ dans \mathbb{R} est dite δ -Hölderienne s'il existe C telle que, pour tout x, y dans $[0, 1]$ telle que

$$|f(x) - f(y)| \leq C|x - y|^\delta.$$

Le but de l'exercice est de montrer que si F_n converge en loi vers F dans $\mathcal{C}([0, 1], \mathbb{R})$ et si F_n vérifie (30) (i.e. le critère de tension par les moments), alors F est p.s. δ -Hölderienne pour tout $\delta < \gamma/\alpha$ (avec une constante C intervenant dans la définition de δ -Hölderienne aléatoire).

1. En utilisant les notations du lemme 9.7, montrer que

$$\sum_{p \geq 1} 2^{p\delta\alpha} \mathbb{E}[U_p(F)^\alpha] < +\infty.$$

(Indice : utiliser l'équation (29).)

2. En déduire qu'il existe une variable aléatoire S telle que, p.s., pour tout $p \geq 1$,

$$U_p(F) \leq S 2^{-p\delta}$$

3. Prouver que F est p.s. δ -Hölderienne.
4. En déduire que le mouvement Brownien est p.s. δ -Hölderienne pour tout $\delta < 1/4$.
5. En cherchant une égalité du type (30) pour tout $\alpha = 2k$, montrer que le mouvement Brownien est en fait p.s. δ -Hölderienne pour tout $\delta < 1/2$.

10 Sources

Pour préparer ce cours, j'ai utilisé les documents suivants (où le lecteur pourra trouver des discussions plus détaillées et parfois des résultats plus généraux) :

- le livre de Billingsley [Bil95] (Chapitre 30) pour la méthode des moments ; voir aussi [JŁR00, Chapter 3 and 6] pour les méthodes des premier et deuxième moments, et pour les moments factoriels dans un contexte de graphes aléatoires.
- le livre de Durrett « *Probability. Theory and examples* » [Dur19, Chapter 3] pour le théorème local limite et l'inégalité de Berry–Esseen ;
- le livre de Dembo et Zeitouni « *Large deviations techniques and applications* » [DZ10, Section 2.2] pour les grandes déviations ;
- et les notes de cours de Marckert « *Limits of random walks and random trees* » [Mar09, Chapter 5] pour la convergence fonctionnelle (avec les livres de Billingsley [Bil99, Section 7] et Kallenberg [Kal02, Corollary 16.9] en références secondaires).

Références

- [Bil95] P. Billingsley. *Probability and measure*. Wiley, 1995. 3rd edition.
- [Bil99] P. Billingsley. *Convergence of probability measures*. Wiley, New York, 1999. 2nd edn.
- [Dur19] R. Durrett. *Probability. Theory and examples*, volume 49 of *Camb. Ser. Stat. Probab. Math.* Cambridge University Press, 5th edition, 2019.
- [DZ10] A. Dembo and O. Zeitouni. *Large deviations techniques and applications*, volume 38. Springer, 2010.
- [JŁR00] S. Janson, T. Łuczak, and A. Ruciński. *Random graphs*, volume 45 of *Wiley Series in Discrete Mathematics and Optimization*. Wiley-Interscience, 2000.
- [Kal02] O. Kallenberg. *Foundations of Modern Probability*. Probability and Its Applications. Springer, 2002. 2nd edn.
- [Mar09] J. F. Marckert. Limits of random walks and random trees, 2009. Notes de cours, disponible en ligne.