

Probabilités discrètes – marches aléatoires

Valentin Féray, Université de Lorraine

18 novembre – 9 décembre 2021

1 Modèle et problématique

Soit $X_1, X_2, \dots$ 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 des. 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 :

- Estimation de divers événements : 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 $(S_k)_{k \leq n}$.

Applications/motivations :

- nous discuterons des conséquences sur les arbres aléatoires ;
- la limite des trajectoires des marches aléatoires est un processus Markovien continu, qui sera étudié dans la suite du cours (avec Denis Villemonais).

2 Préliminaires

2.1 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 défini) 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)

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

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 $]0; +\infty[$.

Théorème 2.1. *Sous les conditions ci-dessus, $\frac{S_n}{\sqrt{n}}$ converge en distribution vers une variable Gaussienne σZ centrée de variance σ^2 .*

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 aient 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 $\mathbb{R}$,

$$\varphi_n(t) = \left(1 - \frac{t^2\sigma^2}{2n}\mathbb{E}(X_1^2) + o(1/n)\right)^n \rightarrow \exp\left(-\frac{t^2\sigma^2}{2}\right) = \varphi_{\sigma Z}(t).$$

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

Exercice 2.1 (« loi des petits nombres »). Soit p_n une suite dans $]0; 1[$ vérifiant $np_n \rightarrow \lambda > 0$. 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 que $S_n = \sum_{i=1}^n X_{n,i}$ converge en distribution vers une loi de Poisson.

(On rappelle qu'une variable P_λ suit une loi de Poisson de paramètre λ si $\mathbb{P}[P_\lambda = k] = e^{-\lambda}\lambda^k/k!$; la fonction caractéristique de P_λ est alors donnée par $\mathbb{E}[\exp(itP_\lambda)] = \exp(\lambda(e^{it} - 1))$.)

3 Théorème local limite

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

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

Que nous dit le théorème centrale limite ? Par exemple pour $a = 0$, pour tout $\varepsilon > 0$

$$\mathbb{P}[S_n = 0] \leq \mathbb{P}\left[\frac{|S_n|}{\sqrt{n}} \leq \varepsilon\right]. \quad (1)$$

Le membre de droite converge vers

$$\mathbb{P}(\sigma Z \leq \varepsilon) = \frac{1}{\sqrt{2\pi}\sigma} \int_{-\varepsilon}^{\varepsilon} \exp(-x^2/(2\sigma^2)) dx$$

Comme (1) est vrai pour tout $\varepsilon > 0$, on peut faire tendre ε vers 0 et conclure que

$$\lim_{n \rightarrow \infty} \mathbb{P}[S_n = 0] = 0.$$

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 :

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

Cette formule a un équivalent continu :

Théorème 3.1 (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 φ_Y peut être raffiné ainsi.

Lemme 3.2. *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 surement ;*
- (ii) *$|\varphi_Y(2\pi/h)| = 1$.*

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

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

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 $b = \alpha h$. $\square$

Théorème 3.3 (Théorème local limite). . Soit $S_n = X_1 + \dots + X_n$ une marche aléatoire, avec X_1 à valeurs dans $\mathbb{Z}$ de fréquence 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

$$\sup_{a \in \mathbb{Z}} |\sqrt{n} \mathbb{P}[S_n = a] - n_\sigma(a/\sqrt{n})| \rightarrow 0, \quad (3)$$

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).

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 (2) et la deuxième du changement de variable $u = t/\sqrt{n}$ (et donc $du = dt/\sqrt{n}$). Par ailleurs, le théorème 3.1 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

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

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

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

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 2.1), 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 fréquence maximale de X_1 , le lemme 3.2 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

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

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 fréquence 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.

3.1 Exercices

Exercice 3.1. Calculer le nombre de suites $(x_1, \dots, x_n)$ dans $\{-1, 1\}^n$ telles que $\sum_{i=0}^n x_i = 0$. En déduire, pour la marche aléatoire **simple** sur $\mathbb{Z}$ (voir définition p. 1), 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 le théorème 3.3.

Exercice 3.2. Supposons que l'on a une suite S_n de variables aléatoires (pas nécessairement une marche aléatoire) vérifiant (3). Montrer que $\frac{S_n}{\sqrt{n}}$ converge en distribution vers σZ .

Exercice 3.3. 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

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

4 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,

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

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...

La limite du membre de droite peut être faite 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 ?).

4.1 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 4.1. *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 4.2), 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

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

on trouve

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

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 $\theta \mapsto \theta b - \Lambda(\theta)$ atteint son supremum pour $\theta \geq 0$. Cette fonction vaut 0 en 0, et pour $\theta < 0$, on a :

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

4. Voir exercice 4.2 pour comprendre la terminologie.

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 $\theta < 0$. 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. $\square$

4.2 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 4.2. *La fonction Λ est C^1 et convexe sur $\mathbb{R}$ et, pour tout θ dans $\mathbb{R}$,*

$$\Lambda'(\theta) = \frac{1}{M(\theta)} \mathbb{E}[X_1 e^{\theta X_1}]. \quad (4)$$

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 et 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, p.s.,

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

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

$$\mathbb{E}[X_1 e^{\theta X_1}] \geq K_1 \mathbb{E}[e^{\theta X_1} \mathbf{1}[X_1 \geq K_1]]. \quad (5)$$

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 5, 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 θ), ce qui conclut la démonstration. $\square$

Theorème 4.3. 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 limsup de $\frac{1}{n} \log (\mathbb{P}[S_n \geq bn])$ est au plus le membre de droite est une conséquence de la proposition 4.1. 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. Grâce à la formule (4), on peut écrire (cette formule sera utile ci-dessous) :

$$\frac{1}{M(\theta^*)} \mathbb{E}[X_1 e^{\theta^* X_1}] = b. \quad (6)$$

Notons P la loi de X_1 . On introduit une mesure Q par la formule $\frac{dQ}{dP}(x) = \frac{e^{\theta^* x}}{M(\theta^*)}$. On définit bien une loi de probas 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 (6),

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

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 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

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

ce qui finit la démonstration. $\square$

Note : la conclusion du théorème 4.3 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*.

4.3 Exercices

Exercice 4.1. 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 4.2. 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 : regarder le développement de $b\theta - \Lambda(b)$ autour de $b = 0$.)

Exercice 4.3. 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 .

5 Temps de premier passage et lemme cyclique

Dans cette section, on supposera que la variable X_1 prend uniquement des valeurs entières **supérieures** ou égales à -1 . On définit, pour $k > 0$,

$$T_{-k} = \inf\{j > 0, S_j = -k\},$$

le premier temps de passage à $-k$. La distribution de T_{-k} est a priori difficile à étudier car la définition fait intervenir tous les S_j . Heureusement, on dispose du résultat remarquable suivant.

Proposition 5.1 (formule de Kemperman). *Pour tous $k \leq n$, on a*

$$\mathbb{P}[T_{-k} = n] = \frac{k}{n} \mathbb{P}[S_n = -k].$$

Dans le membre de droite, la probabilité ne fait intervenir que S_n et est plus facile à analyser. En particulier, on peut utiliser le théorème local limite et on a

$$\mathbb{P}[T_{-k} = n] = \frac{k}{n^{3/2}} n_\sigma(k/\sqrt{n}) + o(k/n^{3/2}),$$

uniformément en k . Cette conséquence est un résultat asymptotique ; notons par opposition que la formule de Kemperman est une égalité exacte.

Démonstration. Soit $\mathbf{x} = (x_1, \dots, x_n)$ un vecteur de nombres entiers dans $\{-1, 0, 1, \dots\}$. Pour un tel vecteur $\mathbf{x}$, on note $S^{\mathbf{x}} = (S_1^{\mathbf{x}}, \dots, S_n^{\mathbf{x}})$ le vecteur des sommes partielles défini par $S_j^{\mathbf{x}} = x_1 + \dots + x_j$.

Pour $i \leq n$, on note $\mathbf{x}^{(i)}$ le décalé cyclique d'ordre i , c'est-à-dire $\mathbf{x}^{(i)} = (x_{i+1}, x_{i+2}, \dots, x_{i+n})$, où les indices sont interprétés modulo n . Le coeur de la démonstration est le résultat combinatoire suivant : *pour tout $\mathbf{x}$ tel que $\sum x_i = -k$, il y a exactement k valeurs de i pour lesquelles $S_j^{\mathbf{x}^{(i)}} > -k$ pour $j < n$ (i.e. les sommes partielles du décalé cyclique d'ordre i atteignent pour la première fois $-k$ au temps n).*

Avant de démontrer ce fait, expliquons pourquoi il implique la formule de Kemperman. Comme le vecteur $\mathbf{X} = (X_j)_{1 \leq j \leq n}$ est un vecteur de variables i.i.d., on a l'égalité en loi $\mathbf{X} \stackrel{\text{loi}}{=} \mathbf{X}^{(i)}$ pour tout i dans $\{0, \dots, n-1\}$. Par conséquent $\mathbf{X} \stackrel{\text{loi}}{=} \mathbf{X}^{(I)}$, où I est une variable aléatoire uniforme dans $\{0, \dots, n-1\}$. Or le résultat combinatoire implique que si $\mathbf{x}$ est un vecteur de somme $-k$, alors

$$\mathbb{P} \left[S_j^{\mathbf{X}^{(I)}} > -k \text{ pour } j < n \mid \mathbf{X} = \mathbf{x} \right] = \frac{k}{n}.$$

Ceci se réécrit

$$\mathbb{P} \left[\left(S_j^{\mathbf{X}^{(I)}} > -k \text{ pour } j < n \right) \wedge (\mathbf{X} = \mathbf{x}) \right] = \frac{k}{n} \mathbb{P}[(\mathbf{X} = \mathbf{x})].$$

En sommant sur tous les vecteurs $\mathbf{x}$ de somme $-k$, on obtient

$$\mathbb{P} \left[\left(S_j^{\mathbf{X}^{(I)}} > -k \text{ pour } j < n \right) \wedge (S_n^{\mathbf{X}} = -k) \right] = \frac{k}{n} \mathbb{P}[S_n^{\mathbf{X}} = -k].$$

On remplace dans le membre de gauche $S_n^{\mathbf{X}}$ par $S_n^{\mathbf{X}^{(I)}}$ (ces deux quantités sont clairement égales). L'aléatoire dans le membre de gauche ne dépend plus que de $\mathbf{X}^{(I)}$. Or on sait que $\mathbf{X} \stackrel{\text{loi}}{=} \mathbf{X}^{(I)}$, on peut donc remplacer $\mathbf{X}^{(I)}$ par $\mathbf{X}$. On obtient

$$\mathbb{P} \left[\left(S_j^{\mathbf{X}} > -k \text{ pour } j < n \right) \wedge (S_n^{\mathbf{X}} = -k) \right] = \frac{k}{n} \mathbb{P}[S_n^{\mathbf{X}} = -k].$$

L'évènement à gauche est exactement $T_{-k} = n$, la formule ci-dessus correspond donc la formule de Kemperman.

Il nous reste à démontrer le résultat combinatoire. Soit $(x_1, \dots, x_n)$ un vecteur de nombres entiers dans $\{-1, 0, 1, \dots\}$ de somme n . Pour $0 < h \leq k$, on note $T_{-h} = \min\{j : S_j = -h\}$. Comme S_j décroît au plus de 1 à chaque étape, on a $T_{-1} < T_{-2} < \dots < T_{-k}$. On vérifie facilement que $\mathbf{x}^{(i)}$ atteint $-k$ pour la première fois au temps n si et seulement si $i = T_{-h}$ pour un certain $h \leq k$, ce qui finit la démonstration. $\square$

5.1 Exercices

Exercice 5.1. Quelle est la probabilité que la marche aléatoire simple sur $\mathbb{Z}$ atteigne $-a$ pour la première fois au temps $2n + 1$?

Exercice 5.2. Montrer que les variables $T_{-i-1} - T_{-i}$ sont i.i.d. (avec la convention $T_0 = 0$).

Exercice 5.3. Soit $k \geq 0$. Montrer l'égalité suivante de distribution

$$\text{Loi}(T_{-1} \mid X_1 = k - 1) = \text{Loi}(T_{-k} + 1).$$

On pose $h(z) = \sum_n \mathbb{P}[T_{-1} = n] z^n$ et $g(z) = \sum_k \mathbb{P}[X_1 = k - 1] x^k$. Montrer (en utilisant l'exercice précédent) que $h(z) = z g(h(z))$.

6 Applications : arbre de Galton–Watson

Dans cette section, nous allons voir un modèle d'évolution de population, appelé processus de Galton–Watson. Ce processus est naturellement représenté par un arbre mathématique. Les marches aléatoires et les résultats ci-dessus donnent des informations utiles sur ces arbres aléatoires.

6.1 Le modèle

Considérons l'ensemble $\mathcal{U}_\infty = \bigsqcup_{h \geq 0} \mathbb{N}^h$ (par convention $0 \notin \mathbb{N}$). Si u est un élément de $\mathcal{U}_\infty$ et i un entier dans $\mathbb{N}$, on note ui l'élément de $\mathcal{U}_\infty$ obtenu en ajoutant i à la fin de u . On dira que u est le parent de ui . Cela munit $\mathcal{U}_\infty$ d'une structure d'« arbre généalogique » (où tous les nœuds ont un nombre infini dénombrable d'enfants).

Par définition, un arbre (plan enraciné) est un sous-ensemble T de $\mathcal{U}_\infty$ vérifiant les propriétés suivantes :

1. l'ensemble vide $\emptyset$ appartient à T (et est appelée racine de T) ;
2. si ui appartient à T , alors u appartient à T ainsi que tous les uj pour $j < i$.

Soit μ une mesure de probabilité sur $\{0, 1, 2, \dots\}$ et $(k_u)_{u \in \mathcal{U}_\infty}$ une collection de v.a. i.i.d. de loi μ indexée par $\mathcal{U}_\infty$. On considère alors l'arbre aléatoire $\mathbf{T}$ défini par

1. $\emptyset \in T$;
2. si u est dans T , alors, ui est dans T pour tout $i \leq k_u$.

Informellement, tous les sommets de l'arbre ont un nombre d'enfants k_u tiré au hasard selon la loi μ , de manière indépendante les uns des autres. On appelle $\mathbf{T}$ arbre de Galton–Watson de loi de reproduction μ . En définissant

$$Z_h = \#\{u \in \mathbf{T} : |u| = h\},$$

on obtient le « processus de Galton–Watson » de loi de reproduction μ .

Un résultat classique (obtenu par exemple avec la théorie des martingales) est le suivant (on exclut le cas trivial $\mu = \delta_1$) :

- Si $m := \sum_{k \geq 0} k \mu(k) \leq 1$, alors $\mathbf{T}$ est fini presque surement.
- Si $m := \sum_{k \geq 0} k \mu(k) > 1$, alors $\mathbf{T}$ est infini avec probabilité strictement positive.

6.2 Lien avec les marches aléatoires

Soit T un arbre. On ordonne les éléments de T en ordre lexicographique (parcours en profondeur de l'arbre). Considérons la suite $(k_u)_{u \in T}$, où k_u est le nombre d'enfants de u dans T . On peut montrer facilement⁵ qu'elle vérifie les propriétés suivantes :

1. Ses sommes partielles décalées sont positives, i.e., quelque soit v dans T ,

$$\sum_{u \in T; u < v} (k_u - 1) \geq 0.$$

2. La somme totale décalée est égale à -1 :

$$\sum_{u \in T} (k_u - 1) < 0.$$

Nous appellerons (*) la conjonction de ces deux propriétés.

Proposition 6.1. *L'application*

$$\begin{array}{ccc} \{\text{arbres de taille } n\} & \rightarrow & \{(k_i)_{i \leq n} \text{ vérifiant } (*)\} \\ T & \mapsto & (k_u)_{u \in T} \end{array}$$

est une bijection.

Démonstration. L'application est bien définie. Nous devons démontrer la surjectivité et l'injectivité. Pour cela, on part d'une suite $(k_i)_{i \leq n}$ vérifiant (*) et on montre qu'elle correspond à exactement un arbre T . Explications sur un exemple au tableau. $\square$

Corollaire 6.2. *Soit μ une loi de probabilité sur $\{0, 1, 2, \dots\}$, et $\mathbf{T}$ un arbre de Galton–Watson de loi de reproduction μ . Soit $X_i = K_i - 1$ une suite de v.a. i.i.d., telle que K_i a loi μ . Pour $j \geq 0$, on note $S_j = X_1 + \dots + X_j$. On a*

$$\mathbb{P}[|\mathbf{T}| = n] = \mathbb{P}[(\forall j < n, S_j \geq 0) \wedge (S_n = -1)].$$

Démonstration. Soit T_0 un arbre de taille n et $(k_i^0)_{i \leq n}$ la suite associée par la bijection de la proposition 6.1. On a

$$\begin{aligned} \mathbb{P}[\mathbf{T} = T_0] &= \mathbb{P}\left[\bigwedge_{u \in T_0} k_u = k_i^0\right] = \prod_{i \leq n} \mu(k_i^0) \\ &= \prod_{i \leq n} \mathbb{P}[K_i = k_i^0] = \mathbb{P}[(K_i)_{i \leq n} = (k_i^0)_{i \leq n}]. \end{aligned}$$

⁵. Pour le montrer, on se pose la question suivant : comment reconstruire l'arbre à partir de $(k_u)_{u \in T}$?

En effet, la première égalité provient de la définition d'un arbre de Galton–Watson, la seconde de l'indépendance et de la distribution des (k_u) , la troisième et la quatrième de la définition de la suite (K_i) . En sommant sur tous les arbres T_0 de taille n , ou de manière équivalente, sur les suites (k_i^0) vérifiant $(*)$, on a

$$\mathbb{P}[|\mathbf{T}| = n] = \mathbb{P}[(K_i)_{i \leq n} \text{ vérifie } (*)].$$

On finit la démonstration en observant que $(K_i)_{i \leq n}$ vérifie $(*)$ si S_j est positive jusqu'au temps $n - 1$ inclus, et vaut -1 au temps n . $\square$

Avec les résultats des sections précédentes, on déduit :

- Si $m := \sum_{k \geq 0} k \mu(k) \neq 1$, alors la marche aléatoire $S_j = X_1 + \dots + X_j$ n'est pas centrée. Si $m > 1$, on peut écrire

$$\mathbb{P}[|\mathbf{T}| = n] \leq \mathbb{P}(S_n = -1) \leq \mathbb{P}(S_n < 0),$$

qui décroît exponentiellement vite vers 0 par le principe de grande déviations. Quand $m < 1$, on écrit plutôt

$$\mathbb{P}[|\mathbf{T}| = n] \leq \mathbb{P}(S_n = -1) \leq \mathbb{P}(S_n \geq -\varepsilon n),$$

pour tout $\varepsilon > 0$ et on conclut également que cette quantité décroît exponentiellement vite vers 0 en choisissant $\varepsilon < 1 - m$.

- Si $m = 1$, alors la marche aléatoire $S_j = X_1 + \dots + X_j$ est centrée. On a

$$\mathbb{P}[|\mathbf{T}| = n] = \mathbb{P}[(\forall j < n, S_j \geq 0) \wedge (S_n = -1)] = \mathbb{P}[T_{-1} = n] = \frac{n^{-3/2}}{\sigma \sqrt{2\pi}} (1 + o(1)),$$

où $\sigma = \sum_{k \geq 0} (k - 1)^2 \mu(k)$, d'après les résultats de la section 5. En particulier, on note que

$$\mathbb{E}[|\mathbf{T}|] = \sum_n n \mathbb{P}[|\mathbf{T}| = n] = +\infty.$$

Conclusion : en combinant avec les résultats « rappelés » en fin de section 6.1, on distingue trois régimes :

regime sous-critique ($m < 1$) l'arbre est fini p.s., et la probabilité d'avoir taille n décroît exponentiellement vite vers 0.

regime critique ($m = 1$) l'arbre est fini p.s., mais la probabilité d'avoir taille n décroît seulement polynômialement vite vers 0 (en $n^{-3/2}$) pour être précis.

regime sur-critique ($m > 1$) l'arbre est infini avec probabilité strictement positive, et la probabilité d'avoir taille n décroît exponentiellement vite vers 0 (intuition : dans ce cas, si l'arbre survit durant quelques générations, alors on a une forte chance d'avoir un arbre infini, donc la probabilité d'avoir une taille fixée grande est faible).

6.3 Exercices

Exercice 6.1. Soit μ la loi géométrique de paramètre $1/2$ et $\mathbf{T}$ l'arbre de Galton–Watson associé. Soit T_0 un arbre à n sommets. Montrer que $\mathbb{P}[\mathbf{T} = T_0] = 2^{-2n+1}$. Calculer $\mathbb{P}[|\mathbf{T}| = n]$ et en déduire une formule pour le nombre d'arbres à n sommets.

Exercice 6.2. Soit μ la loi de distribution uniforme sur l'ensemble à deux éléments $\{0, 3\}$ et $\mathbf{T}$ l'arbre de Galton–Watson associé. Trouver

$$\lim_{n \rightarrow \infty} \frac{1}{n} \log (\mathbb{P}[|\mathbf{T}| = n]).$$

7 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 $[0, 1]$,

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

où $[nt]$ 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.

7.1 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-dimensionnelle de F . La proposition suivante montre que la loi de F est entièrement déterminée par ses lois fini-dimensionnelles.

Proposition 7.1. *Soient F et G deux processus aléatoires dans $\mathcal{C}([0, 1], \mathbb{R})$. Les assertions suivantes sont équivalentes :*

1. $F \stackrel{\text{loi}}{=} G$ en distribution.
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 $\pi_{t_1, \dots, t_d}$ 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 (stable par intersection finie), 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 distribution de F_n vers F (dans $\mathcal{C}([0, 1], \mathbb{R})$). Comme ci-dessus, il y a un sens facile.

Lemme 7.2. *Soient $(F_n)_{n \geq 1}$ et F des processus aléatoires dans $\mathcal{C}([0, 1], \mathbb{R})$. Supposons que F_n converge en distribution vers F . Alors, pour tout $d \geq 1$ et tous $t_1 < \dots < t_d$ dans $[0, 1]$, on a la convergence suivante en distribution 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 $\pi_{t_1, \dots, t_d}$ 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 on a le résultat suivant.

Proposition 7.3. *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 distribution.*
2. *Il existe t_0 tel que $F_n(t_0)$ soit stochastiquement borné, i.e., pour tout $\varepsilon > 0$, il existe $M > 0$ tel que $\mathbb{P}(|F_n(t_0)| \leq M) \geq 1 - \varepsilon$.*
3. *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*

$$\mathbb{E}(|F_n(s) - F_n(t)|^\alpha) \leq \beta |s - t|^{\gamma+1}. \quad (7)$$

Alors, F_n converge vers un processus aléatoire F en distribution.

De plus, si F_n est p.s. affine sur $[\frac{i-1}{n}; \frac{i}{n}]$ pour tout $i \leq n$ et si $\alpha \geq \gamma + 1$, il suffit de vérifier (7) pour s, t dans $\{0, \frac{1}{n}, \dots, \frac{n-1}{n}, 1\}$.

Démonstration. On admettra ce résultat. Il est basé sur plusieurs choses différentes : le théorème de Kolmogorov qui décrit les sous-ensembles compacts de mesures de probabilités, le théorème d'Arzelà-Ascoli qui décrit les sous-ensembles compacts de fonctions continues, la méthode des moments, ... $\square$

On note que l'hypothèse 1. de la proposition est la convergence des lois fini-dimensionnelles vers une limite *non spécifiée* – en particulier, on ne suppose pas que les limites soient elles-mêmes les lois fini-dimensionnelles d'un processus F donné (bien que ce soit une conséquence de la proposition).

Remark 7.4. Les conditions 1. et 2. de la proposition 7.3 sont aussi nécessaires pour que F_n aient une limite en distribution. La condition 3. ne l'est pas. On peut remplacer 3. par une condition plus technique (et plus dure à vérifier directement dans de nombreux cas) impliquant le module de continuité des F_n pour avoir une équivalence, voir par exemple [Bil99, Section 7].

7.2 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 7.5 (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 7.6) 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 7.3. 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

$$(W_n(t_1), W_n(t_2), \dots, W_n(t_d)). \quad (8)$$

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 (8) 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

$$(W_n(\frac{\lfloor nt_1 \rfloor}{n}), W_n(\frac{\lfloor nt_2 \rfloor}{n}) - W_n(\frac{\lfloor nt_1 \rfloor}{n}), \dots, W_n(\frac{\lfloor nt_d \rfloor}{n}) - W_n(\frac{\lfloor nt_{d-1} \rfloor}{n})). \quad (9)$$

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 (9) 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 .

La deuxième hypothèse de la Proposition 7.3 est triviale car $W_n(0) = 0$ p.s. Il reste à vérifier l'équation (7). 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 (7) avec $\alpha = 4$ et $\gamma = 1$ et la troisième hypothèse de la Proposition 7.3 est vérifiée.

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 7.6. *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 7.5 et le lemme 7.2, les lois fini-dimensionnelles de σB sont les limites de celles de W_n . Ces limites ont été déterminé (implicitement) dans la preuve du théorème 7.5, prouvant que B_t vérifient la propriété 3. ci-dessus. L'unicité découle de la proposition 7.1 $\square$

Remark 7.7. Construire directement un processus aléatoire B_t vérifiant les propriétés du corollaire 7.6 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 7.8. 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 (et le critère de convergence donné dans la Proposition 7.3 ne suffit pas).

7.3 Exercices

Exercice 7.1. 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 7.2. Soit S_n une marche aléatoire centrée et W_n défini comme ci-dessus. Soit t dans $]0; 1[$. Prouver que, conditionnellement à l'évènement $S_n = 0$, la quantité $W_n(t)$ converge vers une loi que l'on déterminera.

Conseil : utiliser le théorème local limite pour calculer

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

8 Sources

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

- le livre de Durrett « *Probability. Theory and examples* » [Dur19] pour le théorème local limite ;
- le livre de Dembo et Zeitouni « *Large deviations techniques and applications* » [DZ10] pour les grandes déviations ;
- le livre de Pitman « *Combinatorial stochastic processes* » [DZ10] pour le lemme cyclique ;

— et les notes de cours de Marckert « *Limits of random walks and random trees* » [Mar09] pour la partie arbres aléatoires et la convergence fonctionnelle.

Références

- [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, 2019.
- [DZ10] A. Dembo and O. Zeitouni. *Large deviations techniques and applications*, volume 38. Springer, 2010.
- [Mar09] J. F. Marckert. Limits of random walks and random trees, 2009. Notes de cours, disponible en ligne.