
SOPHIE ROSSI

Trajectoires Browniennes

Mémoire d'initiation à la recherche

Sous la direction de **Julien Randon-Furling**



Cycle Pluridisciplinaire d'Etudes Supérieures
Paris Sciences et Lettres
Troisième année (L3)
Janvier-Juin 2019

Table des matières

Introduction	2
Motivation	2
1 Construction du mouvement Brownien et premières définitions	3
2 Temps de premier passage et barrières horizontales	4
2.1 Densité du temps de premier passage à travers une barrière horizontale $z = a$	4
2.2 Densité du temps de premier passage d'un mouvement confiné entre deux barrières horizontales	6
2.3 Simulation du mouvement brownien	8
3 Temps de premier passage et cône délimité par deux barrières affines symétriques	9
4 Marche Aléatoire Symétrique Simple (MASS) et conditionnement	11
Théorème du scrutin	13
Retour sur les trajectoires browniennes et la ségrégation urbaine . . .	14
Conclusion	16
Remerciements	16
Bibliographie	16
Annexe	17

Introduction

Au cours de mon stage, j'ai étudié le comportement du mouvement Brownien et ses trajectoires, et plus particulièrement lorsqu'on lui impose des contraintes. Ces contraintes prennent la forme de barrières qui vont séparer l'espace en question, et donc confiner le mouvement brownien. Dans ce cadre, le temps de première sortie de l'espace délimité par les contraintes, est une grandeur pertinente et complexe à déterminer explicitement.

Je vais tout d'abord expliquer brièvement comment le mouvement brownien est construit, pour ensuite étudier le temps de premier passage à travers des barrières horizontales et obliques. Enfin, je reviendrai sur une marche aléatoire discrète et sur le théorème du scrutin, ainsi que sur le lien que l'on peut établir avec les trajectoires browniennes.

Motivation

Ce stage s'inscrit dans la continuité de mon stage de L2, au cours duquel j'avais étudié comment modéliser les phénomènes de ségrégation urbaine en évitant d'établir des seuils arbitraires. L'idée était de comparer les données des villes réelles à un modèle nul. Les villes sont représentées par des grilles, où chaque case est une unité spatiale, et le modèle nul probabiliste constitue, le plus souvent, une ville uniforme.

Plus particulièrement, l'article *From urban segregation to multifocal patterns detection*, construit des trajectoires en agrégeant successivement toutes les unités spatiales de la ville. Ainsi, afin de construire un modèle probabiliste de ville "uniforme" (dont les groupes de population sont bien mélangés), il est pertinent de se concentrer sur des marches aléatoires symétriques simples ainsi que sur le mouvement brownien.

1 Construction du mouvement Brownien et premières définitions

Dans cette section, nous allons sommairement revenir sur la construction du mouvement brownien, de sorte que son existence et sa définition seront bien légitimes dans les sections suivantes.

On se place sur un espace de probabilité (Ω, F, P) .

Définition 1.1. Soit G une mesure gaussienne sur \mathbb{R}_+ (absolument continue par rapport à la mesure de Lebesgue). Le pré-mouvement brownien est le processus $(B_t)_{t \in \mathbb{R}_+} : B_t = G(1_{[0,t]})$

Définition 1.2. Si $X(t)_{t \in T}$ est un processus aléatoire à valeurs dans un espace E , les trajectoires de X sont les applications $t \in T \mapsto X_t(\omega) \in E$ obtenues en fixant $\omega \in \Omega$.

Si (E, d) est un espace métrique, on dit que les trajectoires sont höldériennes d'exposant α si il existe $C \in \mathbb{R}$ tel que, $\forall t, s, d(X_s(\omega), X_t(\omega)) \leq C|s - t|^\alpha$

Remarque 1.3. Si une application entre des espaces métriques vérifie la condition de Hölder (ci-dessus) alors elle est uniformément continue.

Le passage du pré-mouvement brownien au mouvement brownien exige la propriété additionnelle de continuité des trajectoires, qui est assurée par le lemme classique de Kolmogorov que l'on admet. On a le résultat suivant :

Théorème 1.4. Soit $B = (B_t)_{t \geq 0}$ un pré-mouvement brownien. Le processus B a une modification dont les trajectoires sont continues, et même localement höldériennes d'exposant $\frac{1}{2} - \delta$, pour tout $\delta \in]0, \frac{1}{2}[$

Définition 1.5. Un processus $(B(t))_{t \geq 0}$ est un mouvement brownien si :

1. $(B(t))_{t \geq 0}$ est un pré-mouvement brownien
2. Les trajectoires de B , ie les applications $t \mapsto B_t(\omega)$ pour $\omega \in \Omega$, sont toutes continues.

Remarque 1.6. L'existence du mouvement brownien découle du théorème précédent.

Définition 1.7. Le propagateur standard du Brownien $p(x, y, t)$ correspond à la probabilité de passer de x à y en un temps t avec le mouvement brownien. Il est donné par :

$$p(x, y, t) = \frac{1}{\sqrt{2\pi t}} \exp\left(-\frac{(x - y)^2}{2t}\right)$$

2 Temps de premier passage et barrières horizontales

2.1 Densité du temps de premier passage à travers une barrière horizontale $z = a$

Soit $a \in \mathbb{R}$.

On considère un mouvement brownien qui part de $x > 0$ et qui arrive en $y > 0$. Quitte à faire une translation de $-a$, on peut se ramener à chercher le temps de premier passage en 0, ce qui revient à mettre une barrière horizontale $z(t) = 0$.

L'idée est de prendre toutes les trajectoires possibles (ce qui correspond à une densité de probabilité donnée par le propagateur), et on enlève la probabilité des trajectoires qui partent de $-x$. Ainsi, on enlève la probabilité de passer en 0.

En effet, comme $p(x, y, t)$ donne la probabilité de tous les chemins allant de x à y , $p(-x, y, t)$ donne la probabilité de tous les chemins allant de $-x$ à y , et donc la probabilité des chemins qui passent par 0.

On pose alors :

$$q(x, y, t) = p(x, y, t) - p(-x, y, t)$$

Soit $t > 0$ fixé.

Pour t fixé, en intégrant $q(x, y, t)$ par rapport à y on obtient la probabilité que le mouvement brownien ne soit pas passé par 0 jusqu'à t . On définit :

$$I := \int_0^\infty q(x, y, t) dy$$

Alors :

$$\begin{aligned} I &= \int_0^\infty \frac{1}{\sqrt{2\pi t}} \left(\exp\left(-\frac{(x-y)^2}{2t}\right) - \exp\left(-\frac{(x+y)^2}{2t}\right) \right) dy \\ &= \int_0^\infty \frac{1}{\sqrt{2\pi t}} \exp\left(-\frac{(x-y)^2}{2t}\right) dy - \int_0^\infty \frac{1}{\sqrt{2\pi t}} \exp\left(-\frac{(x+y)^2}{2t}\right) dy \\ &= \frac{1}{\sqrt{2\pi t}} \left(\int_{-\frac{x}{2t}}^\infty \sqrt{2t} \exp(-z^2) dz - \int_{\frac{x}{2t}}^\infty \sqrt{2t} \exp(-z^2) dz \right) \\ &= \frac{1}{\sqrt{\pi}} \int_{-\frac{x}{2t}}^{\frac{x}{2t}} \exp(-z^2) dz \end{aligned}$$

Par parité de $z \mapsto \exp(-z^2)$, on obtient finalement $I = \operatorname{erf}\left(\frac{x}{\sqrt{2t}}\right)$, où erf est la fonction erreur définie par : $\forall x \in \mathbb{R}, \operatorname{erf}(x) := \frac{2}{\sqrt{\pi}} \int_0^x \exp(-z^2) dz$.

Simulation d'un mouvement brownien avec une barrière horizontale $y=0.2$

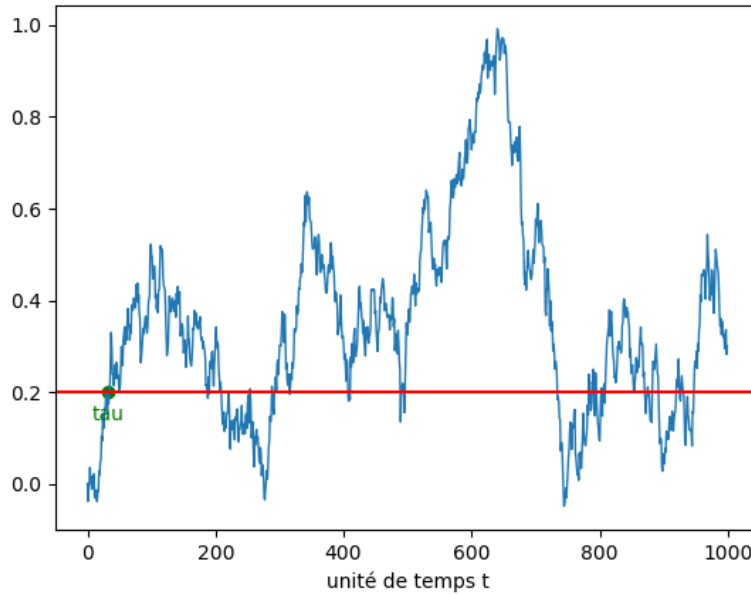


FIGURE 1 – Simulation d'un mouvement brownien avec une barrière horizontale et temps de premier passage τ

Revenons à la définition de I : I donne la probabilité que le mouvement brownien ne soit pas passé par 0 jusqu'à t . En notant f_{τ_0} la densité du temps de premier passage en 0 et F_{τ_0} sa fonction de répartition, on a :

$$\begin{aligned} I &= \mathbb{P}(\forall u \in [0, t], x(u) > 0) \\ &= 1 - F_{\tau_0}(t) \\ &= \int_t^{\infty} f_{\tau_0}(u) du \end{aligned}$$

Ainsi, pour trouver f_{τ_0} , il suffit de dériver I par rapport à t . Finalement, on trouve :

$$f_{\tau_0}(t) = \frac{|x|}{\sqrt{2\pi t^3}} \exp\left(-\frac{x^2}{2t}\right)$$

2.2 Densité du temps de premier passage d'un mouvement confiné entre deux barrières horizontales

Dans ce paragraphe, nous étudions une autre approche possible pour le traitement du mouvement brownien. Il s'agit d'une approche plus physicienne ; c'est pourquoi on introduit un formalisme utilisé en mécanique quantique.

Définition 2.1. L'équation de diffusion du brownien est : $\frac{\partial p(x,t)}{\partial t} = D \frac{\partial^2 p(x,t)}{\partial x^2}$, où D est le coefficient de diffusion ($D = \frac{1}{2}$ pour le brownien standard).

Remarque 2.2. Ici on note $p(x,t) = p(L,x,t)$ le propagateur du mouvement brownien.

On étudie un mouvement brownien confiné entre deux barrières horizontales $y = 0$ et $y = 2L$ (où $L \in \mathbb{R}_+^*$).

Ceci revient à ajouter un potentiel $V(x) = \begin{cases} \infty & \text{si } x \notin]0, 2L[\\ 0 & \text{si } x \in]0, 2L[\end{cases}$

D'après la formule de Feynman-Kac, qui fait le lien entre le formalisme de la mécanique quantique (bra-ket) et les processus stochastiques, le propagateur peut s'écrire :

$$p(x,t) = \langle x_0 | \exp(-\hat{H}T) | x(T) \rangle$$

où \hat{H} est l'hamiltonien associé à notre problème :

$$\hat{H} = -\frac{1}{2} \frac{\partial^2}{\partial x^2} + V(x)$$

Remarque 2.3. $x_0 = x(t=0)$ correspond à la position initiale de la particule

Remarque 2.4. Pour une particule libre (dont le mouvement n'est pas contraint par des barrières), $V(x) = 0 \forall x \in \mathbb{R}$.

La forme du propagateur dans le cas de 2 barrières horizontales est connue :

$$p(x,t) = \frac{1}{L} \sum_{n=1}^{\infty} \sin \frac{n\pi x_0}{2L} \sin \frac{n\pi x}{2L} \exp -\frac{n^2 \pi^2 t}{8L^2}$$

On cherche le temps τ_0 de première sortie du mouvement brownien à travers l'une des deux barrières : $\tau_0 = \inf\{s > 0, X(s) \notin]0, 2L[\}$. On trouve donc aisément la loi de τ_0 , la position initiale x_0 étant connue :

$$\mathbb{P}(\tau_0 \geq t) = \int_0^{2L} p(x,t) dx$$

Dans notre cas, on obtient :

$$\begin{aligned}
\mathbb{P}(\tau_0 \geq t) &= \frac{1}{L} \int_0^{2L} \sum_{n=1}^{\infty} \sin \frac{n\pi x_0}{2L} \sin \frac{n\pi x}{2L} \exp - \frac{n^2 \pi^2 t}{8L^2} dx \\
&= \frac{1}{L} \sum_{n=1}^{\infty} \sin \frac{n\pi x_0}{2L} \exp - \frac{n^2 \pi^2 t}{8L^2} \underbrace{\int_0^{2L} \sin \frac{n\pi x}{2L} dx}_{= \frac{2L}{n\pi} (1 - \cos n\pi) = \frac{4L}{n\pi} \text{ si } n \text{ impair}} \\
&= \frac{4}{\pi} \sum_{k \geq 0} \frac{1}{2k+1} \sin \frac{(2k+1)\pi x_0}{2L} \exp - \frac{(2k+1)^2 \pi^2 t}{8L^2}
\end{aligned}$$

En dérivant par rapport à t , on obtient donc la densité du temps de premier passage à travers l'une des deux barrières :

$$f_{\tau_0}(t) = \frac{\pi}{2L^2} \sum_{k \geq 0} (1 - 2k) \sin \frac{(2k+1)\pi x_0}{2L} \exp - \frac{(2k+1)^2 \pi^2 t}{8L^2}$$

On rappelle que si Z est une variable aléatoire à densité f_Z telle que $Z \geq 0$ p.s., alors $\mathbb{E}[Z] = \int_0^{\infty} \mathbb{P}(Z \geq u) du$. En effet, par Fubini-Tonelli :

$$\begin{aligned}
\int_0^{\infty} \mathbb{P}(Z \geq u) du &= \int_0^{\infty} \int_0^s f_Z(s) ds du \\
&= \int_0^{\infty} \left(\int_0^s f_Z(s) du \right) ds \\
&= \int_0^{\infty} s f_Z(s) ds \\
&= \mathbb{E}[Z]
\end{aligned}$$

Ici, τ_0 est bien une variable aléatoire positive à densité. On peut alors calculer explicitement le temps moyen de sortie :

$$\begin{aligned}
\mathbb{E}[\tau_0] &= \int_0^{\infty} \mathbb{P}(\tau_0 \geq t) dt \\
&= \frac{4}{\pi} \sum_{k \geq 0} \frac{1}{2k+1} \sin \frac{(2k+1)\pi x_0}{2L} \underbrace{\int_0^{\infty} \exp - \frac{(2k+1)^2 \pi^2 t}{8L^2} dt}_{= \frac{8L^2}{(2k+1)^2 \pi^2}} \\
&= 4 \frac{(2L)^2}{\pi^3} \sum_{k \geq 0} \frac{1}{(2k+1)^3} \sin \frac{(2k+1)\pi x_0}{2L}
\end{aligned}$$

On obtient donc une série de Fourier.

En particulier, pour $x_0 = L$, on a :

$$\sin \frac{(2k+1)\pi x_0}{2L} = (-1)^k$$

Simulation d'un mouvement brownien avec deux barrières horizontales

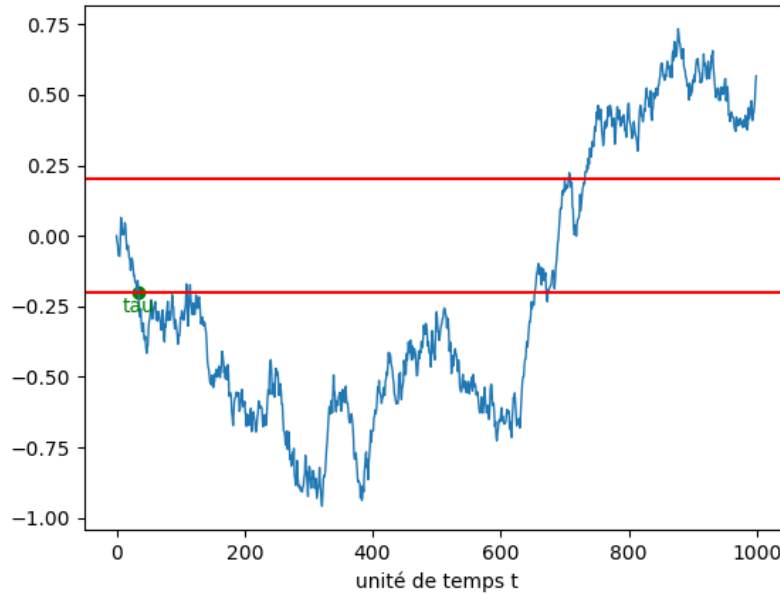


FIGURE 2 – Simulation d'un mouvement brownien avec deux barrières horizontales et temps de première sortie τ

Finalement, pour $x_0 = L$:

$$\mathbb{E}[\tau_0] = 4 \frac{(2L)^2}{\pi^3} \underbrace{\sum_{k \geq 0} \frac{(-1)^k}{(2k+1)^3}}_{\frac{\pi^3}{32}}$$

$$\boxed{\mathbb{E}[\tau_0] = \frac{L^2}{2} = \frac{L^2}{4D}}$$

2.3 Simulation du mouvement brownien

Les graphiques des trajectoires browniennes dans les différentes situations étudiées sont obtenus à partir de la simulation du mouvement brownien décrite dans le TP *Simulation de processus gaussiens* de A. Grafort et C. Malinaud. Il ont été codés en Python, à partir de la méthode des incréments.

L'idée derrière la simulation des trajectoires browniennes est de générer des nombres aléatoires à l'aide du module random de Python : on construit le mouvement brownien à partir de la loi gaussienne centrée réduite. Le premier code

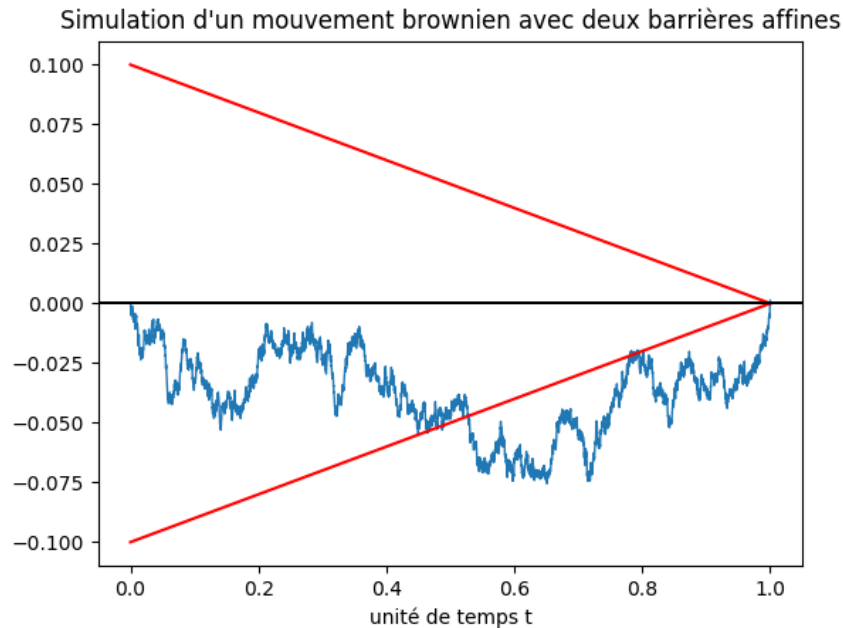


FIGURE 3 – Simulation d'un pont brownien partant et arrivant en 0 avec deux barrières affines

joint en annexe construit une trajectoire brownienne partant de 0.

Ensuite, afin de trouver le temps de premier passage à travers une (ou plusieurs) barrière(s), on étudie simplement le signe de la différence entre le mouvement brownien B et la barrière $y = a$. Le point où advient le premier changement de signe correspond au temps de premier passage.

Le deuxième code en annexe trouve et trace le temps de premier passage à travers la barrière $y = 0.2$ d'un mouvement brownien partant de 0.

Remarque 2.5. Dans le cas de deux barrières on cherche τ_0 pour chacune des barrières, et ensuite on prend le minimum.

3 Temps de premier passage et cône délimité par deux barrières affines symétriques

On considère ici un pont brownien qui part et revient en 0. On renormalise le domaine d'étude de sorte à se placer sur l'intervalle $[0, 1]$ et on prend le temps en sens inverse, c'est-à-dire que l'on considère $1 - t$ au lieu de t . Par propriété du mouvement brownien, le processus étudié reste un mouvement brownien.

On veut confiner le mouvement brownien dans le cône délimité par les droites affines $v(t) = \pm\alpha(1-t)$ où α est le coefficient directeur, que l'on sera amené à faire varier par la suite.

Le propagateur du brownien devient donc : $p(0, \alpha(1-t), 1-t) = \frac{1}{\sqrt{2\pi(1-t)}} \exp(-\alpha^2 \frac{1-t}{2})$

Dans leur article *On hitting times of affine boundaries by reflecting Brownian motion and Bessel processes*, M. Yor et P. Salminen ont calculé la densité du temps de premier passage d'un mouvement brownien entre deux barrières affines :

$$f_{\tau_0}(u) = \frac{1}{\sqrt{2\pi u^3}} \sum_{k=-\infty}^{\infty} (-1)^k \alpha(1-u-2k) \exp(-\alpha^2 k^2) \exp \frac{-\alpha^2(1+u-2k)^2}{2u}$$

Dans notre cas, on force le mouvement brownien à partir de 0 et à arriver en 0, on étudie donc la densité finale f (définie à une constante d'intégration près) :

$$\boxed{f(u)\alpha f_{\tau_0}(u)p(0, \alpha(1-u), 1-u)}$$

L'expression de f étant complexe et difficilement intégrable de manière explicite, on s'intéresse au comportement asymptotique de f lorsque $t \rightarrow 1$. En effet, le temps de dernier passage correspond, symétriquement, au temps de premier passage lorsque le temps est inversé.

On pose donc $\varepsilon = 1-u$ et on étudie f lorsque $\varepsilon \rightarrow 0$.

Ainsi, avec les équivalents usuels en 0, on a :

$$\begin{aligned} f(\varepsilon) &= \frac{1}{\sqrt{2\pi(1-\varepsilon)^3}} \sum_{k=-\infty}^{\infty} (-1)^k \alpha(\varepsilon-2k) \exp(-\alpha^2 k^2) \exp \frac{-\alpha^2(2-\varepsilon-2k)^2}{2(1-\varepsilon)} \frac{1}{\sqrt{2\pi(\varepsilon)}} \exp(-\alpha^2 \frac{\varepsilon}{2}) \\ &\underset{0}{=} \frac{1}{2\pi} \frac{1}{\sqrt{\varepsilon}} \underbrace{\left(1 + \frac{3}{2}\varepsilon + o(\varepsilon)\right) \left(1 - \frac{\alpha^2}{2}\varepsilon + o(\varepsilon)\right)}_{\xrightarrow{\varepsilon \rightarrow 0} 1} \underbrace{\sum_{k=-\infty}^{\infty} (-1)^{k+1} 2k\alpha \exp(-\alpha^2(3k^2 - 2 + k))}_{O(1)} \end{aligned}$$

En effet, d'après le critère spécial des séries alternées (la suite $(U_n)_n$ définie par $U_n = 2n\alpha \exp(-\alpha^2(3n^2 - 2 + n))$ est décroissante et tend vers 0 en ∞), le terme de droite est une série convergente. Elle est donc en $O(1)$ et bornée en valeur absolue par une constante $K > 0$.

Finalement, pour $\varepsilon \rightarrow 0$, on a $|f(\varepsilon)| \leq \frac{\tilde{K}}{\sqrt{\varepsilon}} \rightarrow \infty$. Ce résultat n'est pas satisfaisant, et le terme qui pose problème est celui en $\frac{1}{\sqrt{\varepsilon}}$ qui vient du propagateur (on a $p(x, 0) = \delta(0)$). Or, on peut remarquer que lorsque u est proche de

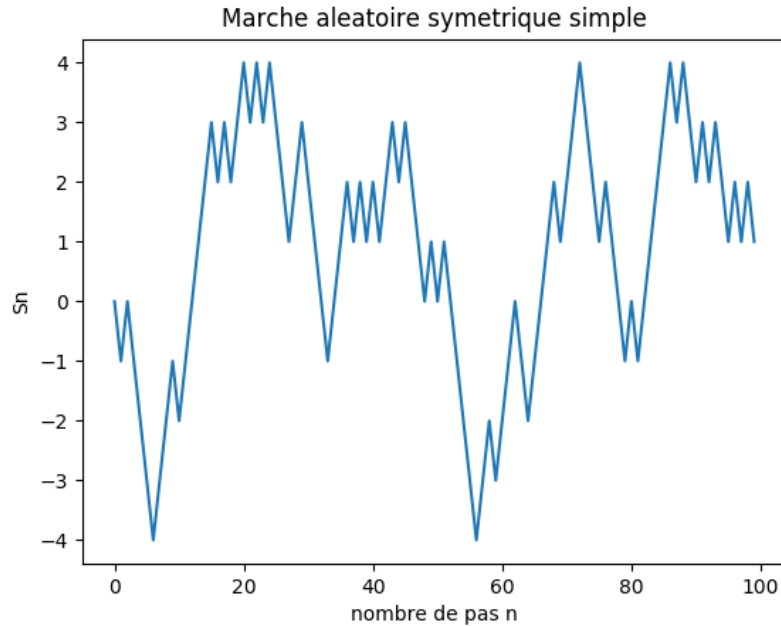


FIGURE 4 – Simulation d’une marche aléatoire symétrique simple

1, la contrainte indiquée par le propagateur est forcément vérifiée (on ne peut pas trop s’éloigner de 0). Ainsi, on peut considérer uniquement la limite pour $\varepsilon \rightarrow 0$ de $f_{\tau_0}(\varepsilon)$, qui est bien finie, d’après ce qui précède.

Ainsi, la décomposition de chemin $f(u)\alpha f_{\tau_0}(u)p(0, \alpha(1-u), 1-u)$ que l’on a choisie ne permet pas ici d’obtenir un résultat. Il faudra procéder autrement.

4 Marche Aléatoire Symétrique Simple (MASS) et conditionnement

Soit $(S_n)_n$ une MASS, ie $\forall n \geq 1, S_n = \sum_{k=0}^n X_k$ et $S_0 = 0$, où les $(X_k)_k$ sont indépendantes et identiquement distribuées telles que $\forall k, \mathbb{P}(X_k = 1) = p$ et $\mathbb{P}(X_k = -1) = 1 - p$. Donc, pour tout n, S_n représente la position après n pas.

On considère N pas et on suppose que S_N est connu, de la forme : $S_N = (2\rho - 1)N$ où $\rho \in [\frac{1}{2}, 1]$.

Soit M une variable aléatoire discrète uniforme sur $\{1, \dots, N\}$, ie $\mathbb{P}(M = k) = \frac{1}{N} \forall k \in [1, N]$.

On cherche la probabilité que $S_M > 0$ sachant $S_N = (2\rho - 1)N$.

Dans les calculs ci-dessous, on a introduit la notation $\mathbb{E}_\rho[\cdot] = \mathbb{E}[\cdot | S_N = (2\rho - 1)N]$

$$\begin{aligned}
\mathbb{P}(S_M > 0 | S_N = (2\rho - 1)N) &= \sum_{k=1}^{\infty} \mathbb{P}(S_M = k | S_N = (2\rho - 1)N) \\
&= \sum_{k=1}^{\infty} \mathbb{E}[\mathbb{1}_{\{S_M=k\}} | S_N = (2\rho - 1)N] \\
&= \sum_{k=1}^{\infty} \mathbb{E}_\rho[\mathbb{1}_{\{S_M=k\}}] \\
&= \sum_{k=1}^{\infty} \mathbb{E}_\rho[\mathbb{E}_\rho[\mathbb{1}_{\{S_M=k\}} | M]] \tag{1}
\end{aligned}$$

Ainsi, $\mathbb{E}_\rho[\mathbb{1}_{\{S_M=k\}} | M = m] = \phi(m)$, et donc, par définition de l'espérance conditionnelle, $\mathbb{E}_\rho[\mathbb{1}_{\{S_M=k\}} | M] = \phi(M)$.

On cherche donc $\phi(m)$:

$$\begin{aligned}
\mathbb{E}_\rho[\mathbb{1}_{\{S_M=k\}} | M = m] &= \mathbb{P}(S_m = k | M = m, S_N = (2\rho - 1)N) \\
&= \frac{\mathbb{P}(S_m = k, M = m, S_N = (2\rho - 1)N)}{\mathbb{P}(S_N = (2\rho - 1)N, M = m)} \\
&= \mathbb{P}(S_m = k | S_N = (2\rho - 1)N) \\
&=: \phi(m) \tag{2}
\end{aligned}$$

Ainsi (1) devient :

$$\begin{aligned}
(1) &= \sum_{k=1}^{\infty} \mathbb{E}_\rho[\phi(M)] \\
&= \sum_{k=1}^{\infty} \sum_{m=1}^N \phi(m) \underbrace{\mathbb{P}(M = m | S_N = (2\rho - 1)N)}_{= \frac{1}{N} \text{ par indépendance}}
\end{aligned}$$

D'autre part, pour l'équation (2) on a :

$$\begin{aligned}
\phi(m) &= \frac{\mathbb{P}(S_m = k, S_N = (2\rho - 1)N)}{\mathbb{P}(S_N = (2\rho - 1)N)} \\
&= \frac{\mathbb{P}(S_m = k) \mathbb{P}(S_N = (2\rho - 1)N - k)}{\mathbb{P}(S_N = (2\rho - 1)N)}
\end{aligned}$$

En effet, on peut voir S_m et S_N comme deux processus indépendants (propriété de Markov) et donc $\mathbb{P}(S_m = k, S_N = (2\rho - 1)N)$ peut être vue comme la probabilité de 2 processus indépendants :

- un premier processus qui part de 0 et arrive en k après m pas
- un deuxième qui part de 0 et qui arrive en $(2\rho - 1)N - k$ en $N - m$ pas (ce qui est équivalent à un processus qui partirait de k et qui arriverait en S_N en $N - m$ pas).

Il reste à calculer la loi de S_n :

$$\begin{aligned}\mathbb{P}(S_n = k) &= \mathbb{P}(qX_1 = 1, mX_1 = -1) \\ &= \binom{n}{q} p^q (1-p)^m\end{aligned}$$

$$\text{où : } \begin{cases} -1 * m + 1 * q = k \\ m + q = n \end{cases} \text{ ie } \begin{cases} q = \frac{k+n}{2} \\ m = \frac{n-k}{2} \end{cases}$$

Finalement :

$$\mathbb{P}(S_n = k) = \binom{n}{\frac{k+n}{2}} p^{\frac{k+n}{2}} (1-p)^{\frac{n-k}{2}}$$

En réinjectant dans (2), on trouve :

$$\phi(m) = \left(\frac{1}{2}\right)^m \binom{m}{\frac{k+m}{2}} \binom{N}{\frac{2\rho N - k}{2}} \frac{1}{\binom{N}{\rho N}}$$

Ainsi :

$$\begin{aligned}\mathbb{P}(S_M > 0 | S_N = (2\rho - 1)N) &= \sum_{k=1}^{\infty} \mathbb{E}_{\rho}[\phi(M)] \\ &= \sum_{k=1}^{\infty} \sum_{m=1}^N \phi(m) \mathbb{P}(M = m | S_N = (2\rho - 1)N)\end{aligned}$$

$$\mathbb{P}(S_M > 0 | S_N = (2\rho - 1)N) = \sum_{k=1}^{\infty} \sum_{m=1}^N \frac{1}{N} \left(\frac{1}{2}\right)^m \binom{m}{\frac{k+m}{2}} \binom{N}{\frac{2\rho N - k}{2}} \frac{1}{\binom{N}{\rho N}}$$

Théorème du scrutin

On admet le théorème suivant :

Théorème 4.1 (Théorème du scrutin). *Soient A et B les deux seuls candidats à une élection, qui récoltent respectivement a et b votes lors du scrutin. On note n le nombre total de votes ($a + b = n$).*

On suppose que A gagne ($a > b$).

La probabilité que, au cours du dépouillement des votes, A soit toujours en tête est : $\frac{a-b}{a+b}$.

Si on note α et β la proportion des votes de A et B, cette probabilité devient $\alpha - \beta$ ie $2\alpha - 1$.

On peut alors appliquer ce résultat à une MASS, où le candidat A correspond à un pas +1 et le candidat B à un pas -1.

On note $\rho > \frac{1}{2}$ la proportion de pas +1.

D'après le théorème du scrutin, en reprenant les notations précédentes :

$$\mathbb{P}(\forall 0 < k \leq N, S_k > 0 | S_N = (2\rho - 1)N) = 2\rho - 1$$

Dans ce cas, on observe tous les points, allant de S_0 à S_N , et on veut savoir quelle est la probabilité qu'ils soient tous positifs, sachant que le point final S_N est positif.

On peut alors se poser la question suivante : si on prend un seul point d'observation M au hasard (uniformément distribué), quelle est la probabilité qu'il soit positif, sachant que le point final $S_N = (2\rho - 1)N$ est positif ?

Cette probabilité est exactement celle qu'on a calculée précédemment. On a donc démontré une généralisation du théorème du scrutin.

Théorème 4.2 (Théorème du scrutin pour des trajectoires observées partiellement). *Pour un point d'observation M uniformément réparti entre 1 et N , la probabilité que $S_M > 0$ sachant que $S_N = (2\rho - 1)N > 0$ est :*

$$\mathbb{P}(S_M > 0 | S_N = (2\rho - 1)N) = \sum_{k=1}^{\infty} \sum_{m=1}^N \frac{1}{N} \left(\frac{1}{2}\right)^m \binom{m}{\frac{k+m}{2}} \binom{N}{\frac{2\rho N - k}{2}} \frac{1}{\binom{N}{\rho N}}$$

Retour sur les trajectoire browniennes et la ségrégation urbaine

Afin de mesurer la ségrégation urbaine, une approche possible consiste à construire des trajectoires en partant d'une unité spatiale choisie au hasard et en agrégeant successivement toutes les unités spatiales de la ville.

Par exemple, si on mesure la proportion de HLM, on peut modéliser ce problème par une MASS, où on fait un pas -1 si l'unité spatiale contient un HLM, et +1 sinon. On connaît à l'avance la proportion de HLM dans la ville étudiée. Ce que l'on cherche à savoir c'est comment la catégorie étudiée (ici les HLM) est répartie dans la ville, afin de déterminer si la ville est ségrégée.

Durant mon stage je n'ai pas travaillé avec des données réelles, mais j'ai étudié et construit des "modèles nuls". Un modèle nul est un "modèle de référence" pour une ville "bien mélangée", dans laquelle la répartition des différentes catégories (sociales, de logements...) est réalisée à l'aide de distributions aléatoires uniformes.

L'idée simple est de suivre la valeur d'une variable (par exemple le taux de hlm, ou bien la fraction de gens appartenant à tel ou tel groupe social) depuis sa valeur locale en un point de la ville jusqu'à sa valeur sur la ville entière, en agrégeant un à un les points de la ville autour du point de départ choisi. En

faisant cela pour chaque point de la ville, on obtient un faisceau de trajectoires qui convergent toutes vers la valeur sur la ville entière (de la variable choisie).

Si la ville est "bien mélangée" du point de vue de cette variable, alors une trajectoire choisie au hasard est modélisable par un processus stochastique dérivé d'une marche aléatoire (ce processus est une marche en score/pourcentage). Et les "instants" de convergence d'une trajectoire vers la moyenne de la ville correspondent à des temps de premier passage pour une marche aléatoire.

Et donc, dans ce cas, pour une trajectoire au hasard, si la ville est bien mélangée, on devrait retrouver la probabilité donnée par le théorème du scrutin. Sinon, la ville est ségréguée. De plus, comme en pratique il est rare d'avoir des données sur tous les points de la ville, étudier les trajectoires à partir d'un seul point d'observation est très pratique. Cela revient donc à utiliser la généralisation du théorème du scrutin.

Finalement, comme on a souvent un grand nombre de données, il est pertinent d'étudier le comportement de $\mathbb{P}(S_M > 0 | S_N = (2\rho - 1)N)$ lorsque $N \rightarrow \infty$. Ceci revient donc à faire une approximation brownienne de la trajectoire, et ensuite à étudier et calculer des temps de premier passage.

Conclusion

Les résultats exposés dans ce rapport, et plus particulièrement les formules explicites des densités de temps de premier passage, sont très généraux. Ils peuvent être utilisés dans de multiples applications, allant de la physique, dans l'étude du mouvement de particules, à la modélisation mathématique des villes et mesure de la ségrégation.

La dernière section revient sur un processus discret : l'étape successive sera de transformer la MASS en un mouvement brownien afin d'exploiter le résultat trouvé dans le cadre de trajectoires browniennes continues.

Remerciements

Je tiens à remercier M. Randon-Furling qui m'a accompagnée et aiguillée au cours de mes stages de L2 et L3, me faisant ainsi explorer et découvrir les processus stochastiques. Je remercie également le département SAMM de l'Université Paris 1 Panthéon-Sorbonne ainsi que tous les professeurs du CPES, qui nous offrent la possibilité de réaliser des stages de recherche.

Bibliographie

- *TP1 MMFI Simulation de processus gaussiens*, A. GRAMFORT, C. MAILINAUD, 2005
- *Mouvement brownien, martingales et calcul stochastique*, J-F. Le Gall, 2013, Springer.
- *Processus stochastiques et mouvement brownien*, P. Lévy, 1965, Editions Jacques Gabay
- *Brownian Motion, Cambridge Series in Statistical and Probabilistic Mathematics*, P. Mörters, Y. Peres, 2010, Cambridge University Press.
- Olteanu, M., Randon-Furling, J., Clark, W. A. (2019). Segregation through the multiscalar lens. *Proceedings of the National Academy of Sciences*, 116(25), 12250-12254.
- Randon-Furling, J., Olteanu, M., Lucquiaud, A. (2018). From urban segregation to spatial structure detection. *Environment and Planning B : Urban Analytics and City Science*, 2399808318797129.
- Julien Randon-Furling, *Statistiques d'extrêmes du mouvement brownien et applications*, Université Paris Sud - Paris XI, 2009.
- *On hitting times of affine boundaries by reflecting brownian motion and besse processes*, P. Salminen, M. Yor, 2011.

Annexe : Codes Python simulant le mouvement brownien

Le code ci-dessous construit une trajectoire brownienne partant de 0.

```
import math
import numpy as np
import matplotlib.pyplot as plt
from random import *

B=[]
B.append(0)
N=5000 #nombre de points pour chaque trajectoire
T=1 #une seule trajectoire
for n in range(N-1):
    b=B[n]+math.sqrt(T)/math.sqrt(N)*np.random.normal(0,1)
    B.append(b)
lx=[i for i in range(N)]

plt.plot(lx,B,linewidth=1)
plt.title("Simulation d'un mouvement brownien partant de 0")
plt.xlabel("unite de temps t")
plt.show()
```

Le code ci-dessous trouve et trace le temps de premier passage à travers la barrière $y = 0.2$ d'un mouvement brownien partant de 0 :

```
diff=[]
for i in range(N-1):
    d=B[i]-0.2 #difference entre les 2 courbes
    diff.append(d)
```

```
Ltau=[]
for i in range(N-2):
    if np.sign(diff[i])!=np.sign(diff[i+1]):
        Ltau.append(i)
tau=Ltau[0]
print("Le temps de premier passage est ",tau)
```

```
#graphique
plt.plot(lx,B,linewidth=1)
plt.axhline(y=0.2,color='r') #droite y=0.2
plt.title("Simulation d'un mouvement brownien avec une barriere horizontale y=0.2")
plt.xlabel("unite de temps t")
plt.scatter(tau,0.2,color='g') #trace tau
```

```
plt.text(tau,0.15, 'tau', horizontalalignment = 'center', verticalalignment  
'center', color='g')  
plt.show()
```