2 | Marches aléatoires

Les marches aléatoires sont largement utilisées pour modéliser des phénomènes réels. Par exemple, la fortune d'un joueur qui joue une suite de parties (par exemple à gain ou perte fixe) contre un adversaire est une marche aléatoire, on trouve aussi les application de marche aléatoire sur un graphe (Recherche sur Google). Les variations temporelles du cours d'une action sur les marchés financiers paraissent aussi bien modélisées par des marches aléatoires ; depuis les années 1970, les modèles probabilistes se sont généralisés en finance. Au début du siècle, le célèbre statisticien Pearson utilisait des marches aléatoires pour décrire des migrations de moustiques. En remontant encore dans l'histoire jusqu'au siècle, on voit le botaniste Brown décrire le mouvement erratique et aléatoire d'une toute petite particule, en l'occurrence un grain de pollen, posée dans un liquide. Einstein en construit un modèle mathématique qui explique ce mouvement par l'agitation thermique. Ce modèle n'est pas à proprement parler une marche aléatoire.

2.1 Marche aléatoire simple

2.1.1 Sur la droite

Définition 2.1.1. *La loi de Rademacher est une loi de probabilité discrète ayant une probabilité* $\frac{1}{2}$ *d'obtenir 1 et* $\frac{1}{2}$ *d'obtenir* -1.

Définition 2.1.2. *La marche aléatoire simple sur la droite* $\mathbb{Z}$ *est la suite* $(X_n)_{n\geqslant 0}$ *de variables aléatoires définie par la relation récursive (ou autorégressive)*

$$X_{n+1} = X_n + \varepsilon_{n+1} = X_0 + \varepsilon_1 + \cdots + \varepsilon_{n+1}$$

pour tout $n \geqslant 0$, *où* $(\varepsilon_n)_{n\geqslant 1}$ *est une suite de v.a.r. i.i.d. de loi de Rademacher, indépendante de* X_0. *On pose*

$$p = \mathbb{P}(\varepsilon_n = 1) = 1 - \mathbb{P}(\varepsilon_n = -1) \in]0, 1[.$$

La suite $(X_n)_{n\geqslant 0}$ est une chaîne de Markov d'espace d'états $\mathbb{Z}$ et de noyau de transition donné pour tous $x, y \in \mathbb{Z}$ par

$$\mathbf{P}(x, y) = p\mathbf{1}_{\{y=x+1\}} + (1-p)\mathbf{1}_{\{y=x-1\}}$$

et la matrice $\mathbf{P}$ est tridiagonale. Une variable aléatoire B suit la loi de Bernoulli $(1-p)\delta_0+p\delta_1$ si et seulement si la variable aléatoire $2B-1$ suit la loi de Rademacher $(1-p)\delta_{-1}+p\delta_1$. Ainsi, pour tout $n \geqslant 0$,

$$\frac{X_n - X_0 + n}{2} \sim \mathrm{Bin}(n,p) \qquad \text{à vérifier numériquement.}$$

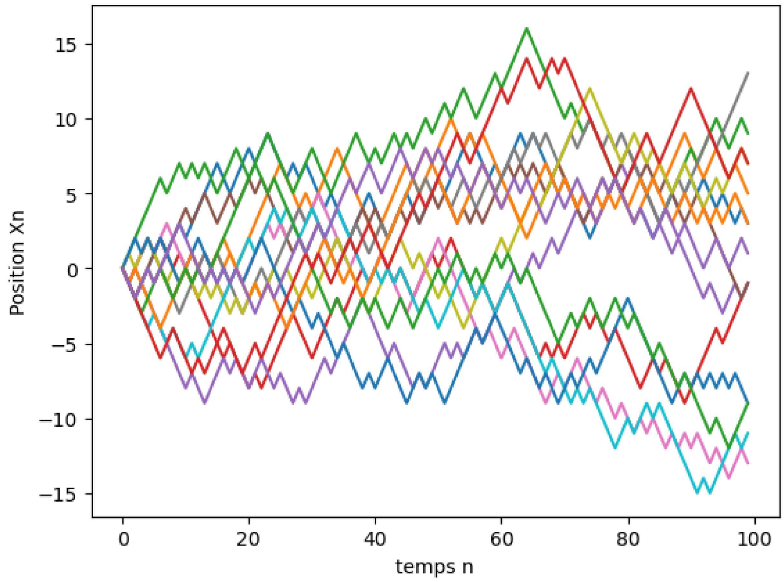


FIGURE 2.1 – Trajectoires issues de 0 de la marche aléatoire simple sur $\mathbb{Z}$.

théorème 2.1.3. *La marche aléatoire simple sur $\mathbb{Z}$ est récurrente nulle si $p = 1/2$ (marche aléatoire simple symétrique) et transitoire si $p \neq 1/2$ (marche aléatoire simple asymétrique).*

Démonstration. Rappelons qu'un état x est récurent lorsque la marche issue de x y revient p.s. ou de manière équivalente lorsque cette chaîne visite une infinité de fois x. Un critère utile est le suivant : l'état x est récurrent si et seulement si $\sum_n \mathbf{P}^n(x,x) = \infty$. La chaîne est de période 2 et donc $\mathbf{P}^{2n+1}(x,x) = 0$ pour tout $n \geqslant 0$. Comme la chaîne est irréductible, tous les états ont même nature, et on peut donc se ramener à l'état 0 . La formule binomiale donne

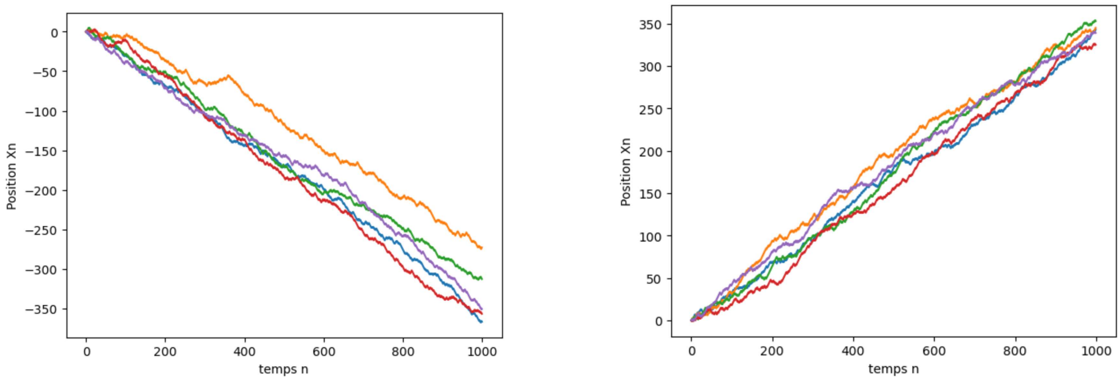
$$\mathbf{P}^{2n}(0,0) = \mathbb{P}\left(X_{2n} = 0 \mid X_0 = 0\right) = \binom{2n}{n} p^n(1-p)^n.$$

La formule de Stirling $n! \sim \sqrt{2\pi n}(n/e)^n$ donne, en notant $\rho = 4p(1-p)$,

$$\mathbf{P}^{2n}(0,0) \sim \frac{\rho^n}{\sqrt{\pi n}}$$

À présent, si $p = 1/2$ alors $\rho = 1$ et la chaîne est récurrente tandis que si $p \neq 1/2$ alors $\rho < 1$ et la chaîne est transitoire. Lorsque $p = 1/2$ la mesure de comptage est symétrique (car la matrice $\mathbf{P}$ est symétrique) donc invariante, et comme il ne s'agit pas d'une mesure finie, la chaîne est récurrente nulle. $\square$

On peut alternativement traiter le cas $p \neq 1/2$ avec un peu plus d'intuition probabiliste. En effet, la loi forte des grands nombres affirme qu'on a presque sûrement $X_n = X_0 + n\left(2p - 1 + o_{n\to\infty}(1)\right)$, donc $X_n \to +\infty$ p.s. si $p > 1/2$ et $X_n \to -\infty$ p.s. si $p < 1/2$, ce qui interdit toute récurrence.



(a) $p = \frac{1}{3}$ (b) $p = \frac{2}{3}$

FIGURE 2.2 – Le cas d'une marche aléatoire transitoire

2.1.2 En deux dimensions

Le cas d'une marche aléatoire en 2 dimension, il y a 4 possibilités pour chaque $X_i \in \{(1,0), (0,1), (-1,0), (0,-1)\}$ chacune avec une portabilité $\frac{1}{4}$, (Voir 2.3).

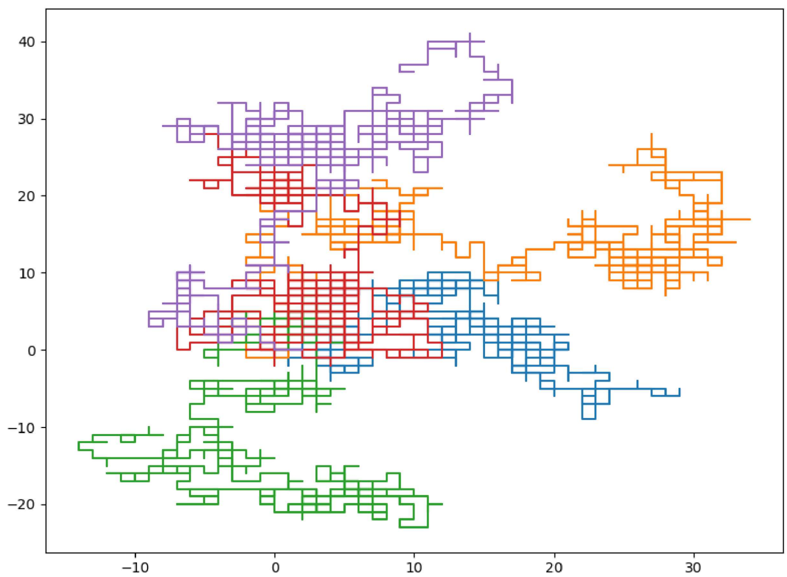


FIGURE 2.3 – 5 Marches aléatoires sur $\mathbb{Z}^2$

2.1.3 En trois dimensions

Le cas d'une marche aléatoire en 2 dimension, il y a 6 possibilités pour chaque $X_i \in \{(0,0,1), (0,0,-1), (0,1,0), (0,-1,0), (1,0,0), (-1,0,0)\}$ chacune avec une portabilité $\frac{1}{6}$ (Voir 2.4).

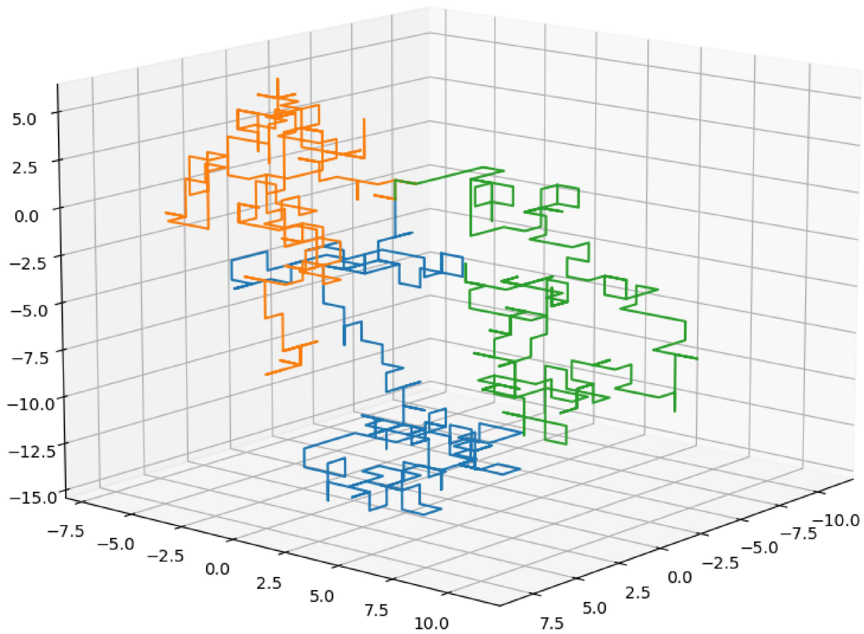


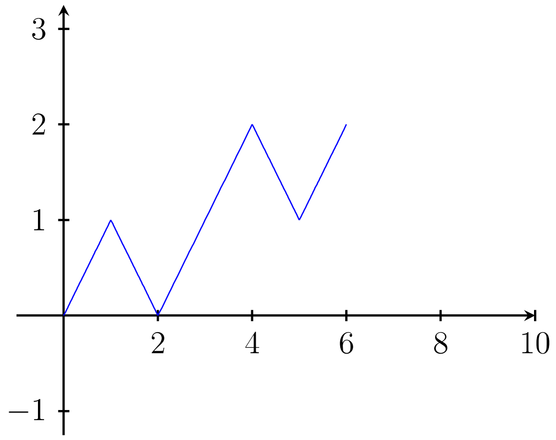
FIGURE 2.4 – 3 Marches aléatoires sur $\mathbb{Z}^3$

théorème 2.1.4. *La marche aléatoire simple symétrique sur $\mathbb{Z}^d$ est transitoire si $d \geq 3$.*

Démonstration. Voir [1] □

2.2 Problème du scrutin

Le problème du scrutin traite le cas lors d'une élection avec deux candidats V et P et n votants, dans laquelle V obtient p votes et P obtient $q \leqslant p$ votes, la probabilité que V soit devant P tout le long du dépouillement des n bulletins de vote est $(p-q)/(p+q)$. On appelle V le vainqueur et P le vaincu. On associe au déroulement du dépouillement un chemin de la façon suivante : on suppose que la position initiale est le point $(0,0)$. Dès qu'on ouvre un bulletin du vainqueur, on avance d'un pas à droite, et on monte d'un pas. Et lorsqu'on ouvre un bulletin du vaincu, on avance à droite, et on descend d'un pas. Par exemple si le début du dépouillement donne V-P-V-V-P-V, alors le chemin est :



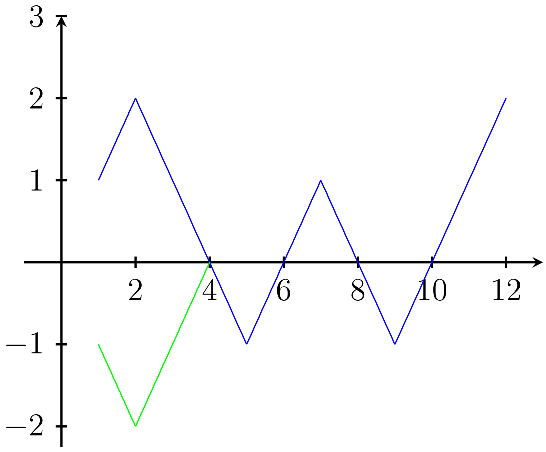
Si le point de coordonnées (x, y) est sur le chemin, x représente le nombre de bulletins dépouillés, et y la différence entre le nombre de voix obtenus par le vainqueur et le vaincu à cet instant du dépouillement. Le premier point du chemin est $(0, 0)$, le dernier le point (n, s) avec $n = p + q$, $s = p - q$. Notre problème est de dénombrer le nombre de chemins allant de (0,0) à (n, s), et se situant, à part le point (0,0), strictement au-dessus de l'axe horizontal. Le premier segment d'un tel chemin amène forcément $(0, 0)$ en $(1, 1)$, et il suffit de compter le nombre de chemins qui vont de $(1, 1)$ à (n, s) sans toucher ou traverser l'axe horizontal. On réalise ce calcul en 3 étapes :

— Etape 1 : on compte le nombre total de chemins qui vont de $(1, 1)$ à $(p + q, p - q)$. Il s'agit, pour $(p + q - 1)$ voix dépouillées, de choisir le moment où les $(p - 1)$ voix sont données au vainqueur. Il faut donc choisir une partie à $p - 1$ éléments parmi $(p + q - 1)$: il y a $\begin{pmatrix} p+q-1 \\ p-1 \end{pmatrix}$ tels chemins.

— Etape 2 : on compte le nombre de chemins de $(1, 1)$ à $(p+q, p-q)$ qui touchent ou traversent l'axe horizontal. On utilise le principe de réflexion suivant :

Lemme 2.2.1. *Le nombre de chemins qui vont de $(1, 1)$ à $(p + q, p - q)$ qui touchent ou traversent l'axe horizontal vaut le nombre total de chemins allant de $(1, -1)$ à $(p + q, p - q)$.*

Démonstration. On suit une démarche géométrique. Soit C un chemin qui va de $(1, 1)$ à $(p + q, p - q)$ en touchant l'axe. On note T le premier point de contact avec l'axe horizontal. On note C_1 la portion de chemin qui va de $(1, 1)$ à T, et C_2 la portion de chemin qui va de T à $(p+q, p-q)$. On construit C' un chemin de la façon suivante : C' est la réunion de C'_1 et C'_2, où C'_1 est le symétrique de C_1 par rapport à l'axe horizontal :



Alors, C' est un chemin qui va de $(-1,1)$ à $(p+q,p-q)$, et la correspondance C donne C' est une bijection.

Pour compter le nombre de chemins qui vont de $(1,1)$ à $(p+q,p-q)$ qui touche l'axe horizontal, il suffit de compter le nombre de chemins qui vont de $(1,-1)$ à $(p+q,p-q)$. On calcule ce nombre comme à l'étape 1 , et on trouve $\begin{pmatrix} p+q-1 \\ p \end{pmatrix}$ tels chemins. □

— Etape 3 : Conclusion. Le nombre de chemin recherché est égal au nombre total de chemins moins le nombre de chemins qui traversent l'axe, soit :

$$\begin{pmatrix} p+q-1 \\ p-1 \end{pmatrix} - \begin{pmatrix} p+q-1 \\ p \end{pmatrix} = \frac{p-q}{p+q} \begin{pmatrix} p+q \\ p \end{pmatrix}.$$

Maintenant, il y a $\begin{pmatrix} p+q \\ p \end{pmatrix}$ dépouillements possibles (on compte les moments où les p voix sont données au vainqueur parmi $p+q$). La probabilité pour que le vainqueur soit toujours en tête est donc :

$$\frac{p-q}{p+q}$$

théorème 2.2.2. (*Scrutin*) *Si* $n = p+q$ *et* $k = p-q$ *avec* $0 \leqslant q \leqslant p$ *alors*

$$\mathbb{P}(S_1 > 0, \ldots, S_n > 0 \mid S_0 = 0, S_n = k) = \frac{k}{n} = \frac{p-q}{p+q}.$$

Démonstration. à refaire. □