Deux applications du principe de réflexion

Mon développement

Soit $(S_n)_{n\in\mathbb{N}}$ une marche aléatoire symétrique sur $\mathbb{Z}$. Pour tout $n \in \mathbb{N}$, on note

$$T_n = \max\{k \in [\![0, n]\!] \mid S_k = 0\}$$

le dernier temps de passage en 0 avant l'instant n.
Dans la suite, on utilisera la correspondance usuelle entre les marches aléatoires et les chemins.
Pour tous $n \in \mathbb{N}^*$, $a, b \in \mathbb{Z}$, on note $N_n(a, b)$, resp. $N_n^0(a, b)$, $N_n^{\neq 0}(0, b)$, le nombre de chemins allant de $(0, a)$ à (n, b), resp. allant de $(0, a)$ à (n, b) en passant par 0 sans considérer les extrémités, allant de $(0, 0)$ à (n, b) sans passer par 0 sauf éventuellement aux extrémités.

Proposition (Principe de réflexion). $\forall n \in \mathbb{N}^*$, $\forall a, b \in \mathbb{N}$, on a : $N_n^0(a, b) = N_n(-a, b)$.

Soient $n \in \mathbb{N}^*$, $a, b \in \mathbb{N}$.
Soit C un chemin allant de $(0, a)$ à (n, b) passant par 0 sans considérer les extrémités. L'instant $t = \min\{k \in [\![1, n-1]\!] \mid S_k = 0\}$ du premier retour en 0 est bien défini. On note C_1, resp. C_2, le sous-chemin de C allant de $(0, a)$ à $(t, 0)$, resp. de $(t, 0)$ à (n, b), on note $\widetilde{C_1}$ le chemin symétrique de C_1, et on note $\varphi(C)$ la concaténation des chemins $\widetilde{C_1}$ et C_2.
On définit ainsi une application $\varphi : C \mapsto \varphi(C)$ de l'ensemble des chemins de $(0, a)$ à (n, b) passant par 0 sans considérer les extrémités sur l'ensemble des chemins allant de $(0, -a)$ à (n, b).
φ est clairement injective et surjective, en réalisant le même type de construction. Donc φ est bijective, et donc $N_n^0(a, b) = N_n(-a, b)$. $\square$

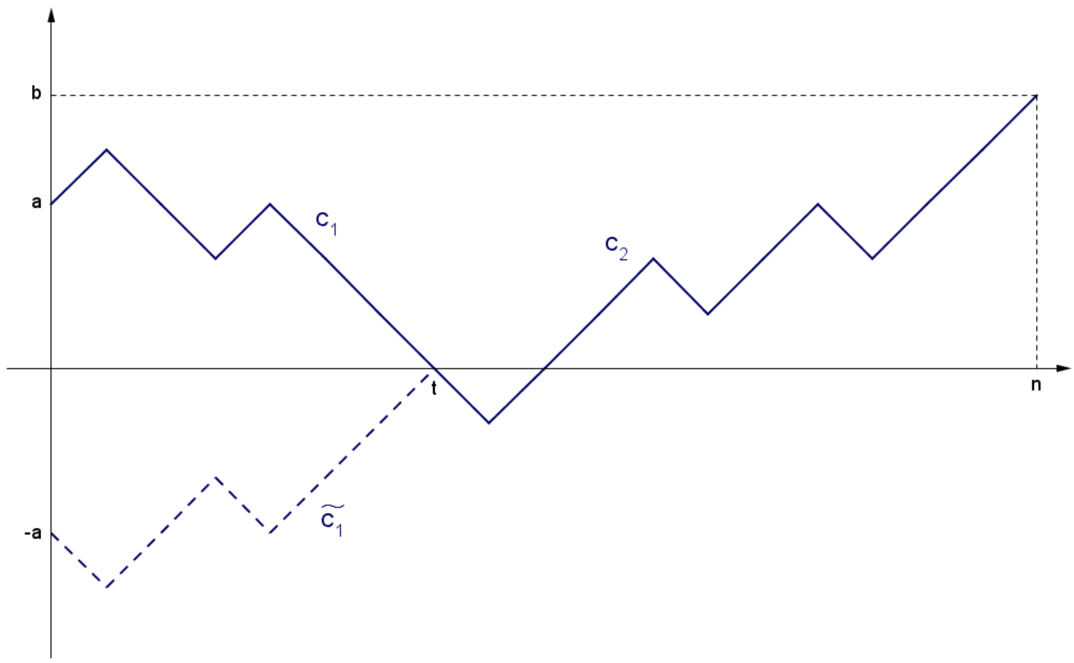

FIGURE 1 – Illustration du principe de réflexion

Rappel. Soient $a_1, a_2 \in \mathbb{N}$, $b_1, b_2 \in \mathbb{Z}$. Le nombre de chemins allant de (a_1, b_1) à (a_2, b_2) est égal à

$$\binom{a_2 - a_1}{\frac{1}{2}(a_2 - a_1 + b_2 - b_1)} .$$

Soient $a_1, a_2 \in \mathbb{N}$, $b_1, b_2 \in \mathbb{Z}$.
Soit C un chemin allant de (a_1, b_1) à (a_2, b_2). On note m et d les nombres de montées et de descentes dans C. On a donc $\begin{cases} a_2 - a_1 = m + d \\ b_2 - b_1 = m - d \end{cases}$, donc $m = \frac{1}{2}(a_2 - a_1 + b_2 - b_1)$. Ainsi, pour tout chemin de (a_1, b_1) à (a_2, b_2), $\frac{1}{2}(a_2 - a_1 + b_2 - b_1)$ montées sont effectuées. Choisir un chemin de (a_1, b_1) à (a_2, b_2) revenant à choisir les instants des montées, on a $\binom{a_2 - a_1}{\frac{1}{2}(a_2 - a_1 + b_2 - b_1)}$ possibilités. $\square$

Proposition (Problème du scrutin). $\forall n \in \mathbb{N}^*$, $\forall b \in \mathbb{N}^*$, on a :

$$N_n^{\neq 0}(0, b) = \frac{b}{n} \binom{n}{\frac{n+b}{2}} .$$

Soient $n \in \mathbb{N}^*$, $b \in \mathbb{N}^*$.
Comme $b > 0$, un chemin issu de 0 ne touchant pas l'axe est un chemin dont le premier pas est $+1$. On a donc : $N_n^{\neq 0}(0, b) = N_{n-1}^{\neq 0}(1, b) = N_{n-1}(1, b) - N_{n-1}^{0}(1, b) = N_{n-1}(1, b) - N_{n-1}(-1, b)$ d'après le principe de réflexion. D'après le rappel, on a :

$$\begin{aligned}
N_n^{\neq 0}(0, b) &= \binom{n-1}{\frac{1}{2}(n+b) - 1} - \binom{n-1}{\frac{1}{2}(n+b)} \\
&= \frac{(n-1)!}{(\frac{1}{2}(n+b) - 1)!(\frac{1}{2}(n-b))!} - \frac{(n-1)!}{(\frac{1}{2}(n+b))!(\frac{1}{2}(n-b) - 1)!} \\
&= \frac{(n-1)!}{(\frac{1}{2}(n+b))!(\frac{1}{2}(n-b))!} \left(\frac{n+b}{2} - \frac{n-b}{2} \right) \\
&= \frac{1}{n} \binom{n}{\frac{1}{2}(n+b)} b
\end{aligned}$$

ce qui est la formule annoncée. $\square$

Application numérique : au second tour de l'élection présidentielle 2012, François Hollande a reçu 18 000 668 voix, alors que Nicolas Sarkozy en a reçu 16 860 685. La probabilité que François Hollande ait été en tête tout au long du scrutin est donc égale à $\frac{18000668}{34861353} \simeq 0,516$.

Lemme. Soit $n \in \mathbb{N}^*$. On a : $\mathbb{P}(S_1 \neq 0, \dots, S_{2n} \neq 0) = \mathbb{P}(S_{2n} = 0)$.

Soit $n \in \mathbb{N}^*$. On a :

$$\begin{aligned}
\mathbb{P}(S_1 \neq 0, \ldots, S_{2n} \neq 0) &= 2\,\mathbb{P}(S_1 > 0, \ldots, S_{2n} > 0) \quad \text{(par symétrie)} \\
&= 2\sum_{r=1}^{+\infty} \mathbb{P}(S_1 = 1, S_2 > 0, \ldots, S_{2n} = 2r) \quad \text{(évènements disjoints)} \\
&= 2\sum_{r=1}^{+\infty} \frac{1}{2}\frac{1}{2^{2n-1}} N_{2n-1}^{\neq 0}(1, 2r) \\
&= 2\sum_{r=1}^{+\infty} \frac{1}{2^{2n}}(N_{2n-1}(1,2r) - N_{2n-1}(-1,2r)) \quad \text{(comme précédemment)} \\
&= 2\sum_{r=1}^{+\infty} \frac{1}{2^{2n}}(N_{2n-1}(0,2r-1) - N_{2n-1}(0,2r+1)) \\
&= \frac{2}{2^{2n}} N_{2n-1}(0,1) = \frac{1}{2^{2n}} N_{2n}(0,0) = \mathbb{P}(S_{2n} = 0)
\end{aligned}$$

en utilisant des arguments de comptage de chemins. □

Proposition (Loi de l'arcsinus). Soient $0 < \alpha < \beta < 1$. On a :

$$\lim_{n \to +\infty} \mathbb{P}(2n\alpha \leq T_{2n} \leq 2n\beta) = \frac{2}{\pi}(\mathrm{Arcsin}\sqrt{\beta} - \mathrm{Arcsin}\sqrt{\alpha}) \ .$$

Soient $n \in \mathbb{N}^*$, $k \in [\![0, n]\!]$. On a :

$$\begin{aligned}
\mathbb{P}(T_{2n} = 2k) &= \mathbb{P}(S_{2k} = 0, S_{2k+1} \neq 0, \ldots, S_{2n} \neq 0) \\
&= \mathbb{P}(S_{2k+1} \neq 0, \ldots, S_{2n} \neq 0 |\ S_{2k} = 0)\,\mathbb{P}(S_{2k} = 0) \\
&= \mathbb{P}(S_1 \neq 0, \ldots, S_{2n-2k} \neq 0)\,\mathbb{P}(S_{2k} = 0) \\
&= \mathbb{P}(S_{2n-2k} = 0)\,\mathbb{P}(S_{2k} = 0)
\end{aligned}$$

en utilisant respectivement un argument de comptage et le lemme pour les deux dernières égalités. On a donc :

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

Par ailleurs, d'après la formule de Stirling, quand $p \to +\infty$, $\binom{2p}{p} \sim \frac{2^{2p}}{\sqrt{\pi p}}$.
Soit $\varepsilon \in]0, 1[$. Il existe donc $N \in \mathbb{N}$ tel que $\forall p \geq N$:

$$\frac{2^{2p}}{\sqrt{\pi p}}(1 - \varepsilon) \leq \binom{2p}{p} \leq \frac{2^{2p}}{\sqrt{\pi p}}(1 + \varepsilon) \ . \tag{1}$$

Soit $N' \in \mathbb{N}$ tel que $\forall n \geq N'$, $n\alpha \geq N$ et $n(1 - \beta) \geq N$ (un tel N' existe bien). Pour tout entier $k \in [n\alpha, n\beta]$, k et $n - k$ sont supérieurs à N. En multipliant les deux inégalités (1) pour k et $n - k$, on obtient :

$$\frac{2^{2n}(1-\varepsilon)^2}{\pi\sqrt{k(n-k)}} \leq \binom{2n-2k}{n-k}\binom{2k}{k} \leq \frac{2^{2n}(1+\varepsilon)^2}{\pi\sqrt{k(n-k)}}$$

donc, en divisant par 2^{2n} et en sommant :

$$\sum_{k:\ n\alpha \leq k \leq n\beta} \frac{(1-\varepsilon)^2}{\pi\sqrt{k(n-k)}} \leq \sum_{k:\ n\alpha \leq k \leq n\beta} \mathbb{P}(T_{2n} = 2k) \leq \sum_{k:\ n\alpha \leq k \leq n\beta} \frac{(1+\varepsilon)^2}{\pi\sqrt{k(n-k)}} \ . \tag{2}$$

Or la fonction $x \mapsto \frac{1}{\pi\sqrt{x(1-x)}}$ est continue sur $[\alpha, \beta]$ donc, d'après le théorème des sommes de Riemann,

$$\lim_{n\to+\infty} \sum_{k:\ n\alpha\le k\le n\beta} \frac{1}{\pi\sqrt{k(n-k)}} = \int_\alpha^\beta \frac{dx}{\pi\sqrt{x(1-x)}} \ .$$

De plus, en effectuant le changement de variables "$u = \sqrt{x}$", on obtient que

$$\int_\alpha^\beta \frac{dx}{\pi\sqrt{x(1-x)}} = \int_{\sqrt{\alpha}}^{\sqrt{\beta}} \frac{2u}{\pi u\sqrt{1-u^2}}\, du = \left[\frac{2}{\pi}\mathrm{Arcsin}(u)\right]_{\sqrt{\alpha}}^{\sqrt{\beta}} \ .$$

D'après la relation (2), on a d'une part

$$\limsup \mathbb{P}(2n\alpha \le T_{2n} \le 2n\beta) \le \frac{2}{\pi}(1+\varepsilon)^2(\mathrm{Arcsin}\sqrt{\beta} - \mathrm{Arcsin}\sqrt{\alpha})$$

et d'autre part

$$\frac{2}{\pi}(1-\varepsilon)^2(\mathrm{Arcsin}\sqrt{\beta} - \mathrm{Arcsin}\sqrt{\alpha}) \le \liminf \mathbb{P}(2n\alpha \le T_{2n} \le 2n\beta) \ .$$

Ceci étant vrai pour tout $\varepsilon \in]0,1[$, on obtient la limite annoncée en faisant tendre $\varepsilon \to 0$. $\square$

Références

J'ai utilisé [CDGCM11, pp. 18, 22], [Les01, p. 51].

Leçons correspondantes

J'utilise ce développement pour les leçons 145, 249.

Remarques

- Presque sûrement, T_n est défini à partir d'un certain rang. Pour n'avoir aucune ambigüité, on peut choisir de noter $T_n = 0$ pour les instants n auxquels la marche n'est pas encore passée par 0.
- On autorise les arguments quelconques pour les coefficients binomiaux, avec la convention $\binom{b}{a} = 0$ si on n'a pas $b \in \mathbb{N}$ et $a \in [\![0, b]\!]$.
- Tout le développement peut se faire à l'aide de [Les01] uniquement mais les arguments ne sont pas toujours les mêmes. Par exemple, dans [Les01], $\mathbb{P}(T_{2n} = 2k)$ est calculée par récurrence.
- À la place de certains arguments de comptage, il est possible d'invoquer la propriété de Markov faible.
- Comme l'ensemble des trois propriétés est assez long, dans la leçon de combinatoire, il faut plus détailler le comptage de chemins et la première application, et dans la leçon de probabilités, on peut choisir de détailler plus la seconde application par exemple.

Questions possibles

1. Justifier que quand $p \to +\infty$,
$$\binom{2p}{p} \sim \frac{2^{2p}}{\sqrt{\pi p}} \ .$$

2. Énoncer le théorème des sommes de Riemann.
3. Calculer $\mathbb{P}(S_1 > 0, \dots, S_{2n} > 0)$.
4. Soit pour tout $n \in \mathbb{N}^*$, $U_n = \mathrm{Card}\{k \in [\![1, n]\!] \mid S_k = 0\}$ le nombre de retours en zéro avant l'instant n. Calculer la loi de U_{2n} pour tout $n \in \mathbb{N}^*$.

Références

[CDGCM11] Marie Cottrell, Christian Duhamel, Valentine Genon-Catalot et Thierry Meyre : *Exercices de probabilités.* Cassini, 2011.

[Les01] Emmanuel Lesigne : *Pile ou face.* Ellipses, 2001.