TD 10 – Marche aléatoire simple

Dans tout le TD, on notera $(S_k)_{0\leq k\leq \rho}$ une marche aléatoire simple symétrique de longueur ρ aussi grande que nécessaire.

1 – Petites questions

1. Fermat et Pascal jouent à croix ou pile. Étant bons amis, ils décident de jouer jusqu'à ce que chacun ait retrouvé sa fortune initiale. Combien de fois en moyenne vont-ils lancer la pièce ?
2. Fermat et Pascal décident à présent de passer une nuit blanche à jouer à croix ou pile sans s'arrêter. Il jouent finalement pendant 8 heures, en lançant une pièce par seconde. À la fin de la partie, Fermat perd son calme et s'écrie : “Mon cher Pascal, après seulement une demi-heure de jeu, tu es passé en tête et nous ne sommes plus jamais revenus à égalité depuis ! Tu dois avoir usé de quelque fourberie, pour qu'un tel biais en ta faveur apparaisse !” Pouvez-vous aider Pascal à convaincre son compère qu'il n'est pas un tricheur ?

Corrigé.

1. On s'intéresse donc à l'espérance de T le premier instant de retour en 0, qui est une variable aléatoire à valeur dans les entiers pairs strictement positifs.
   Première méthode. Pour $k \geq 1$, on a
   $$\mathbb{P}(T \geq 2k+2) = \mathbb{P}(T \geq 2k+1) = \mathbb{P}(S_1 \neq 0, \ldots, S_{2k} \neq 0) = \mathbb{P}(S_{2k} = 0) = \frac{1}{2^{2k}}\binom{2k}{k} \sim \frac{1}{\sqrt{\pi k}}$$
   par le lemme fondamental. Or on a
   $$\mathbb{E}[T] = \sum_{i\geq 1} \mathbb{P}(T \geq i) = \sum_{k \geq 1} 2\mathbb{P}(T \geq 2k),$$
   et la série de terme général $1/\sqrt{k}$ diverge, donc $\mathbb{E}[T] = \infty$.
   Deuxième méthode. D'après le cours, on connaît explicitement la loi du r^{e} retour en 0 et on a
   $$\mathbb{P}(T = 2k) = \frac{2^{-(2k-1)}}{2k-1}\binom{2k-1}{k} = \frac{2^{-(2k-1)}}{2k-1}\binom{2k-1}{k-1} = \frac{2 \cdot 2^{-2k}}{2k-1}\frac{k}{2k}\binom{2k}{k} \sim \frac{1}{2\sqrt{\pi}k^{3/2}},$$
   en prenant les équivalents quand $k \to \infty$. Ainsi $\mathbb{E}[T] = \infty$.
2. On approche la loi du dernier retour en 0 par la loi de l'arcsinus continue. Alors la probabilité que le dernier retour en 0 ait lieu avant un quart d'heure de jeu est environ
   $$\frac{2}{\pi}\arcsin\left(\sqrt{\frac{30}{60*8}}\right) \simeq 0,16086 \simeq 16\%.$$
   Il y avait 16% de chance que l'un des joueurs domine à partir de 30 minutes de jeu, et 8% de chance que ce soit Pascal : ce n'est pas si faible.

Pour toute question, n'hésitez pas à m'envoyer un mail à michel.pain@ens.fr, ou bien à venir me voir au bureau V2.

Si l'on ne veut pas approximer par la loi de l'arcsinus continue, on peut aussi calculer, avec $n = 8 * 3600$ et $k = 30 * 60$,

$$\mathbb{P}\left(\forall j > k, S_j > 0\right) = \frac{1}{2}\mathbb{P}\left(\forall j > k, S_j \neq 0\right) = \frac{1}{2}\sum_{i=0}^{k/2}\mathbb{P}\left(\max\{j \in [\![0,n]\!] : S_j = 0\} = 2i\right)$$
$$= \frac{1}{2}\sum_{i=0}^{k/2}\frac{1}{2^n}\binom{2i}{i}\binom{n-2i}{\frac{n}{2}-i} \simeq \frac{1}{2}\cdot 0,16093 \simeq 8\%,$$

donc l'approximation fonctionne assez bien ici.

2 – Théorème du scrutin

Exercice 1.

1. (*Théorème du scrutin*) Soit $n, a \in \mathbb{N}^*$. Montrer que
$$\mathbb{P}(S_1 > 0, S_2 > 0, \ldots, S_{n-1} > 0, S_n = a) = \frac{a}{n}\mathbb{P}(S_n = a).$$
2. Lors de l'élection du COF, la ProbabiListe gagne face à la DétermiListe avec 300 voix contre 200. On suppose que le dépouillement a été fait par une seule personne, qui note le score après chaque bulletin de vote ouvert. Avec quelle probabilité la ProbabiListe a-t-elle été en tête tout au long du dépouillement ?

Corrigé.

1. On a
$$\begin{aligned}&\#\{\text{chemins de } (0,0) \text{ à } (n,a) \text{ ne retouchant pas l'axe des abscisses après le point de départ}\}\\ &= \#\{\text{chemins de } (1,1) \text{ à } (n,a) \text{ ne touchant jamais l'axe des abscisses}\}\\ &= \#\{\text{chemins de } (1,1) \text{ à } (n,a)\} - \#\{\text{chemins de } (1,1) \text{ à } (n,a) \text{ touchant l'axe des abscisses}\}\\ &= \#\{\text{chemins de } (1,1) \text{ à } (n,a)\} - \#\{\text{chemins de } (1,-1) \text{ à } (n,a)\},\end{aligned}$$
par le principe de réflexion. En notant, $n = p+q$ et $a = p-q$ aec $p,q \in \mathbb{N}$ (si cela n'est pas possible alors on ne peut pas avoir $S_n = a$ donc le résultat est immédiat), on obtient
$$\mathbb{P}(S_1 > 0, \ldots, S_{n-1} > 0, S_n = a) = \frac{1}{2^n}\left(\binom{p+q-1}{p-1} - \binom{p+q-1}{p}\right) = \frac{1}{2^n}\left(\frac{p}{p+q}\binom{p+q}{p} - \frac{q}{p+q}\binom{p+q}{q}\right)$$
$$= \frac{1}{2^n}\frac{p-q}{p+q}\binom{p+q}{p} = \frac{a}{n}\mathbb{P}(S_n = a).$$
2. On pose $n = p+q$ et $a = p-q$. La question revient à déterminer la probabilité que $S_1 > 0, \ldots, S_{n-1} > 0$ sachant que $S_n = a$ qui est égale à
$$\frac{\mathbb{P}(S_1 > 0, \ldots, S_{n-1} > 0, S_n = a)}{\mathbb{P}(S_n = a)} = \frac{a}{n} = \frac{1}{5}.$$
C'est donc assez peu probable, l'issue du vote n'a donc pas de raison d'être claire dès le début du dépouillement.

3 – Changements de signe

Exercice 2. On dit qu'il y a un *changement de signe* à l'instant $k \geq 1$ si S_{k-1} et S_{k+1} sont de signes différents (alors $S_k = 0$ et k est pair). On note ξ_{2n+1} le nombre de changements de signe qui ont lieu avant l'instant $2n+1$. L'objectif de cet exercice est de déterminer la loi de ξ_{2n+1}.

1. De manière analogue, on note ξ'_{2n} le nombre de fois que le chemin croise le niveau -1 avant l'instant $2n$. Montrer que ξ_{2n+1} et ξ'_{2n} ont même loi.
2. Montrer que $\mathbb{P}(\xi_{2n+1}=0)=2\mathbb{P}(S_{2n+1}=1)$.
3. Montrer que, pour tout $r\in\mathbb{N}$, $\mathbb{P}(\xi_{2n+1}=r)=2\mathbb{P}(S_{2n+1}=2r+1)$.
4. En déduire que le nombre de changements de signe le plus probable est 0, quelle que soit la longueur de la marche aléatoire considérée.
5. Calculer $\mathbb{E}[\xi_{2n+1}]$ et en donner un équivalent quand $n\to\infty$.

Corrigé.

1. On remarque tout d'abord que par symétrie du problème on peut se contenter de compter les chemins passant par $(1,1)$: on a, pour tout $r\in\mathbb{N}$,

$$\#\{\xi_{2n+1}=r\}=\#\{\xi_{2n+1}=r,S_1=1\}+\#\{\xi_{2n+1}=r,S_1=-1\}=2\cdot\#\{\xi_{2n+1}=r,S_1=1\}.$$

Les chemins partant de $(1,1)$ et croisant r fois le niveau 0 avant l'instant $2n+1$ correspondent aux chemins partant de $(0,0)$ et croisant r fois le niveau -1 avant l'instant $2n$ (en ramenant le point de départ de $(1,1)$ à $(0,0)$), on a donc

$$\mathbb{P}(\xi_{2n+1}=r)=\frac{2\cdot\#\{\xi_{2n+1}=r,S_1=1\}}{2^{2n+1}}=\frac{\#\{\xi'_{2n}=r\}}{2^{2n}}\mathbb{P}(\xi'_{2n}=r),$$

et ainsi ξ_{2n+1} et ξ'_{2n} ont même loi.

2. *Méthode agréable.* On a

$$\#\{\xi_{2n+1}=0\}=\#\{S_1\geq 0,\dots,S_{2n+1}\geq 0\}+\#\{S_1\leq 0,\dots,S_{2n+1}\leq 0\}=2\#\{S_1\geq 0,\dots,S_{2n+1}\geq 0\},$$

par symétrie. Remarquons alors que

$$\begin{aligned}\#\{S_1>0,\dots,S_{2n+2}>0\}&=\#\{\text{chemins partant de }(1,1)\text{ restant supérieur ou égal à 1 sur }[\![1,2n+2]\!]\}\\&=\#\{S_1\geq 0,\dots,S_{2n+1}\geq 0\}.\end{aligned}$$

Ainsi, on obtient

$$\mathbb{P}(\xi_{2n+1}=0)=\frac{1}{2^{2n+1}}\cdot 2\cdot 2^{2n+2}\mathbb{P}(S_1>0,\dots,S_{2n+2}>0)=2\mathbb{P}(S_1\neq 0,\dots,S_{2n+2}\neq 0)=2\mathbb{P}(S_{2n+2}=0),$$

par le lemme fondamental. Or on a $\mathbb{P}(S_{2n+2}=0)=\frac{1}{2}(\mathbb{P}(S_{2n+1}=-1)+\mathbb{P}(S_{2n+1}=1))=\mathbb{P}(S_{2n+1}=1)$, donc on obtient le résultat souhaité.
Méthode bourrin. Par la question 1., il suffit de calculer $\mathbb{P}(\xi'_{2n}=0)$. Un chemin issu de $(0,0)$ ne croise pas le niveau -1 si et seulement si il ne touche jamais le niveau -2, donc

$$\begin{aligned}\mathbb{P}(\xi'_{2n}=0)=\mathbb{P}(S_1>-2,\dots,S_{2n}>-2)&=\sum_{k=0}^{n}\mathbb{P}(S_1>-2,\dots,S_{2n-1}>-2,S_{2n}=2k)\\&=\sum_{k=0}^{n}\frac{1}{2^n}\#\{\text{chemins de }(0,2)\text{ à }(2n,2k+2)\text{ ne touchant pas l'axe des abcisses}\}\\&=\frac{1}{2^n}\sum_{k=0}^{n}\#\{\text{chemins de }(0,2)\text{ à }(2n,2k+2)\}-\#\{\text{chemins de }(0,-2)\text{ à }(2n,2k+2)\},\end{aligned}$$

par le principe de réflexion. Ainsi, on obtient

$$\mathbb{P}(\xi'_{2n}=0)=\sum_{k=0}^{n}\mathbb{P}(S_{2n}=2k)-\mathbb{P}(S_{2n}=2k+4)=\mathbb{P}(S_{2n}=0)+\mathbb{P}(S_{2n}=2),$$

les autres termes se simplifiant entre eux. Enfin, on conclut en remarquant que $\mathbb{P}(S_{2n+1}=1)=\frac{1}{2}(\mathbb{P}(S_{2n}=0)+\mathbb{P}(S_{2n}=2))$.

3. On procède par récurrence, toujours en travaillant avec ξ'_{2n}. Soit $r \geq 1$, on suppose que, pour tout $n \in \mathbb{N}$,
$$\mathbb{P}(\xi'_{2n} = r-1) = 2\mathbb{P}(S_{2n+1} = 2r-1).$$
On note $T := \min\{k \in [\![1,2n]\!] : S_k = -2\}$, qui correspond à l'instant suivant la 1[re] traversée du niveau -1. Sur l'événement $\{\xi'_{2n} = r\}$, T est bien défini et on a
$$\begin{aligned}\mathbb{P}(\xi'_{2n} = r) &= \sum_{k=1}^{n} \mathbb{P}(T = 2k, \xi'_{2n} = r)\\ &= \sum_{k=1}^{n} \mathbb{P}(S_1 > -2, \dots, S_{2k-1} > -2, S_{2k} = -2, \xi'_{2n} = r)\\ &= \sum_{k=1}^{n} \frac{1}{2^{2n}} \#\{S_1 > -2, \dots, S_{2k-1} > -2, S_{2k} = -2\}\\ &\quad \cdot \#\{\text{chemins partant de } (2k,-2) \text{ croisant } r-1 \text{ fois le niveau } -1 \text{ avant l'instant } 2n\}.\end{aligned}$$
On a, par symétrie par rapport à -1 puis en déplaçant le point de départ,
$$\begin{aligned}&\#\{\text{chemins partant de } (2k,-2) \text{ croisant } r-1 \text{ fois le niveau } -1 \text{ avant l'instant } 2n\}\\ &= \#\{\text{chemins partant de } (2k,0) \text{ croisant } r-1 \text{ fois le niveau } -1 \text{ avant l'instant } 2n\}\\ &= \#\{\xi'_{2n-2k} = r-1\} = \#\{S_{2n-2k+1} = 2r-1\},\end{aligned}$$
en utilisant l'hypothèse de récurrence. On obtient ainsi
$$\begin{aligned}\mathbb{P}(\xi'_{2n} = r) &= \sum_{k=1}^{n} \frac{1}{2^{2n}} \#\{S_1 > -2, \dots, S_{2k-1} > -2, S_{2k} = -2\} \#\{S_{2n-2k+1} = 2r-1\}\\ &= \sum_{k=1}^{n} \frac{1}{2^{2n}} \#\{S_1 > -2, \dots, S_{2k-1} > -2, S_{2k} = -2\} \#\{\text{chemins de } (2k,-2) \text{ à } 2r-1\}\\ &= \sum_{k=1}^{n} \frac{1}{2^{2n}} \#\{T = 2k, S_{2n} = 2r-1\}\\ &= \frac{1}{2^{2n}} \#\{S_{2n} = 2r-1, \exists k \in [\![0, 2n+1]\!] : S_k = -2\},\\ &= \frac{1}{2^{2n}} \#\{\text{chemins de } (0,2) \text{ à } (2n, 2r-3) \text{ touchant l'axe des abcisses}\}\\ &= \frac{1}{2^{2n}} \#\{\text{chemins de } (0,-2) \text{ à } (2n, 2r-3)\} = \mathbb{P}(S_{2n+1} = 2r+1),\end{aligned}$$
en utilisant le principe de réflexion.

4. Par la question 3., on a que, pour tout $n \in \mathbb{N}$, $(\mathbb{P}(\xi_{2n+1} = r))_{r \in \mathbb{N}}$ est une suite décroissante.

5. On a, en utilisant quelques relations sur les coefficients binomiaux dont $q\binom{p}{q} = p\binom{p-1}{q-1}$,

$$\begin{aligned}
\mathbb{E}[\xi_{2n+1}] &= \sum_{r=0}^{n} 2\mathbb{P}(S_{2n+1} = 2r+1)r = \frac{1}{2^{2n+1}} \sum_{r=0}^{n} 2\binom{2n+1}{r+n+1} r \\
&= \frac{1}{2^{2n+1}} \sum_{r=0}^{n} 2\binom{2n+1}{r+n+1}(r+n+1) - \frac{1}{2^{2n+1}} \sum_{r=0}^{n} 2\binom{2n+1}{r+n+1}(n+1) \\
&= \frac{2n+1}{2^{2n+1}} \sum_{r=0}^{n} 2\binom{2n}{r+n} - \frac{n+1}{2^{2n+1}} \sum_{r=0}^{n} 2\binom{2n+1}{r+n+1} \\
&= \frac{2n+1}{2^{2n+1}} \sum_{r=0}^{n} \left(\binom{2n}{r+n} + \binom{2n}{n-r}\right) - \frac{n+1}{2^{2n+1}} \sum_{r=0}^{n} \left(\binom{2n+1}{r+n+1} + \binom{2n+1}{n-r}\right) \\
&= \frac{2n+1}{2^{2n+1}} \left(\binom{2n}{n} + \sum_{k=0}^{2n} \binom{2n}{k}\right) - \frac{n+1}{2^{2n+1}} \sum_{k=0}^{2n+1} \binom{2n+1}{k} \\
&= \frac{2n+1}{2^{2n+1}} \binom{2n}{n} + \frac{2n+1}{2} - (n+1) = \frac{2n+1}{2^{2n+1}} \binom{2n}{n} - \frac{1}{2}.
\end{aligned}$$

Pour obtenir l'équivalent, on utilise que $\binom{2n}{n} \sim 2^{2n}/\sqrt{\pi n}$, ce qui donne

$$\mathbb{E}[\xi_{2n+1}] \sim \sqrt{\frac{n}{\pi}}.$$

Ainsi, le nombre moyen de changements de signe ne coit qu'en $\sqrt{n}$ alors que l'on pourrait intuitivement croire qu'il croît linéairement en n.

4 – Dualité

Exercice 3. On considère une marche aléatoire simple de longueur $2n$ et on définit l'instant de la 1re visite au point terminal par $T := \min\{k \in [\![0, 2n]\!] : S_k = S_{2n}\}$. Déterminer la loi de T.

Corrigé. On remarque tout d'abord que T est pair. Ensuite, on note $(S_k^*)_{0 \le k \le 2n}$ la marche duale à $(S_k)_{0 \le k \le 2n}$ (on rappelle que $S_k^* = S_{2n} - S_{2n-k}$). Alors on a, pour tout $k \in [\![0, n]\!]$,

$$\begin{aligned}
\{T = 2k\} &= \{S_0 < S_{2n}, \dots, S_{2k-1} < S_{2n}, S_{2k} = S_{2n}, S_{2k+1} \le S_{2n}, \dots, S_{2n-1} \le S_{2n}\} \\
&= \{S_{2n}^* > 0, \dots, S_{2n-2k+1}^* > 0, S_{2n-2k}^* = 0, S_{2n-2k-1}^* \ge 0, \dots, S_1^* \ge 0\} \\
&= \{\text{la dernière visite en } 0 \text{ pour } (S_k^*)_{0 \le k \le 2n} \text{ a lieu à l'instant } 2k\}.
\end{aligned}$$

Or $(S_k^*)_{0 \le k \le 2n}$ a la loi d'une marche aléatoire simple, donc l'instant de dernière visite en 0 pour $(S_k^*)_{0 \le k \le 2n}$ suit une loi de l'arcsinus discrète. On en conclut que T suit une loi de l'arcsinus discrète de longueur $2n$.

5 – Compléments (hors TD)

Exercice 4. (*Retours en 0*) On note N le nombre de retours en 0 de la marche avant l'instant $2n$ inclus.

1. Soit $r \in \mathbb{N}$. Montrer que $\mathbb{P}(N = r) = \mathbb{P}(N \ge r, S_{2n} = 0)$.
2. En utilisant la loi du r^{e} retour en 0 et la question précédente, calculer la loi de N.

Corrigé.

1. On note T_r l'instant du r^{e} retour en 0, qui est bien défini dès que $N \geq r$. On a

$$\begin{aligned}
\mathbb{P}(N=r) &= \sum_{k=1}^{n} \mathbb{P}(T_r = 2k, N=r) \\
&= \mathbb{P}(T_r = 2n) + \sum_{k=1}^{n-1} \mathbb{P}(T_r = 2k, S_{2k+1} \neq 0, \ldots, S_{2n} \neq 0) \\
&= \mathbb{P}(T_r = 2n) + \sum_{k=1}^{n-1} \frac{1}{2^{2n}} \#\{T_r = 2k\} \#\{S_1 \neq 0, \ldots, S_{2n-2k} \neq 0\} \\
&= \mathbb{P}(T_r = 2n) + \sum_{k=1}^{n-1} \frac{1}{2^{2n}} \#\{T_r = 2k\} \#\{S_{2n-2k} = 0\},
\end{aligned}$$

par le lemme fondamental. On obtient alors

$$\begin{aligned}
\mathbb{P}(N=r) = \mathbb{P}(T_r = 2n) &+ \sum_{k=1}^{n-1} \mathbb{P}(T_r = 2k, S_{2n} = 0) \\
&= \mathbb{P}(N = r, S_{2n} = 0) + \mathbb{P}(N \geq r+1, S_{2n} = 0) = \mathbb{P}(N \geq r, S_{2n} = 0).
\end{aligned}$$

2. On remarque que $N \leq n$ et donc, en utilisant la question précédente, on a

$$\begin{aligned}
\mathbb{P}(N=r) = \sum_{i=r}^{n} \mathbb{P}(N = i, S_{2n} = 0) &= \sum_{i=r}^{n} \mathbb{P}(T_i = 2n) \\
&= \sum_{i=r}^{n} \frac{i}{2^{2n-i}(2n-i)} \binom{2n-i}{n}
\end{aligned}$$

d'après le cours. On obtient ensuite

$$\begin{aligned}
\mathbb{P}(N=r) &= \sum_{i=r}^{n} \frac{1}{2^{2n-i}} \left(1 - \frac{2(n-i)}{2n-i}\right) \binom{2n-i}{n} = \sum_{i=r}^{n} \frac{1}{2^{2n-i}} \binom{2n-i}{n} - \sum_{i=r}^{n} \frac{1}{2^{2n-i}} \frac{2(n-i)}{2n-i} \binom{2n-i}{n-i} \\
&= \sum_{i=r}^{n} \frac{1}{2^{2n-i}} \binom{2n-i}{n} - \sum_{i=r}^{n} \frac{1}{2^{2n-i-1}} \binom{2n-i-1}{n-i-1} = \frac{1}{2^{2n-r}} \binom{2n-r}{n}
\end{aligned}$$

car les deux sommes se télescopent.

Exercice 5. On note $X_1, \ldots, X_{2n} \in \{-1, 1\}$ les sauts de la marche aléatoire simple $(S_k)_{0 \leq k \leq 2n}$.

1. On définit $X'_1 := X_{2n}$, $X'_k := X_{k-1}$ pour $k \in [\![2, 2n]\!]$ et $S'_k := \sum_{i=1}^{k} X'_i$ pour $k \in [\![0, 2n]\!]$. Montrer que $(S'_k)_{0 \leq k \leq 2n}$ est une marche aléatoire simple, qui a même point terminal que $(S_k)_{0 \leq k \leq 2n}$.
2. On note T l'instant où la marche $(S_k)_{0 \leq k \leq 2n-1}$ atteint son maximum (si le maximum est atteint en plusieurs instants, T est choisi uniformément parmi ceux-là). Montrer que, pour $k \in [\![0, 2n-1]\!]$,

$$\mathbb{P}(T = k, S_{2n} = 0) = \frac{1}{2n} \mathbb{P}(S_{2n} = 0),$$

c'est-à-dire que la loi de T sachant que $S_{2n} = 0$ est la loi uniforme sur $[\![0, 2n-1]\!]$.

Corrigé.

1. On note φ la fonction qui à un chemin s associe le chemin s' obtenu en appliquant le cycle $(1, \ldots, 2n)$ aux incréments de s (c'est la fonction telle que $S' = \varphi(S)$ est une bijection car on a $\varphi^n = \text{id}$. Donc φ préserve la mesure uniforme sur l'ensemble des chemins de longueur $2n$ partant de 0. Or la loi de $S' = \varphi(S)$ est la mesure image par φ de la loi de S, c'est-à-dire la loi uniforme. Donc S' a même loi que S. Il est clair que $S_n = S'_n$.

2. On considère l'ensemble $\mathcal{C}_0$ des chemins de longueur $2n$ se terminant en 0. La restriction de φ à $\mathcal{C}_0$ défini une bijection de $\mathcal{C}_0 \to \mathcal{C}_0$.
Soit $s \in \mathcal{C}_0$. Remarquons que $\max_{0\le k\le 2n-1} s_k = \max_{0\le k\le 2n} s_k$ (car $s_n = 0 \le s_0$) et on le note dorénavant $\max s$. On note $s' := \varphi(s)$, $I := \{k \in [\![0, 2n-1]\!] : s_k = \max s\}$ et I' défini de la même manière pour s'. On note θ la permutation circulaire $(0, \dots, 2n-1)$ (c'est-à-dire $\theta(k) = k+1$ sauf $\theta(2n-1) = 0$). Montrons que $I' = \theta(I)$.
Cas 1 : $s_{2n-1} = 1$. Alors $s_{2n} - s_{2n-1} = -1$ et $\max s' = \max s - 1$. En outre, s' atteint son maximum en tout point de $I+1$ et éventuellement en 0 (si et seulement si $\max s = 1$). Ainsi, pour $k \in [\![0, 2n-2]\!]$, on a $k \in I$ si et seulement si $k+1 \in I$. Si $\max s \ge 2$, alors $0 \notin I'$ et $2n-1 \notin I$: ainsi $I' = \theta(I)$. Si $\max s = 1$, alors $0 \in I'$ et $2n-1 \in I$: ainsi $I' = \theta(I)$. Enfin $\max s = 0$ n'est pas possible car $s_{2n-1} = 1$.
Cas 2 : $s_{2n-1} = -1$. On procède de même, mais c'est encore plus simple car on ne peut pas avoir $2n-1 \in I$ ni $0 \in I'$ (car $\max s' = \max s + 1 \ge 1$).
Comme $I' = \theta(I)$ pour tout $s \in \mathcal{C}_0$, on a

$$\mathbb{P}(T = k, S_{2n} = 0) = \mathbb{P}(T' = \theta(k), S'_{2n} = 0) = \mathbb{P}(T = \theta(k), S_{2n} = 0)$$

car S et S' ont même loi. Ainsi, on obtient

$$\mathbb{P}(T = k, S_{2n} = 0) = \frac{1}{2n}\sum_{i=0}^{2n-1} \mathbb{P}\Big(T = \theta^i(k), S_{2n} = 0\Big) = \frac{1}{2n}\sum_{j=0}^{2n-1} \mathbb{P}(T = j, S_{2n} = 0) = \frac{1}{2n}\mathbb{P}(S_{2n} = 0),$$

en utilisant que $\{\theta^i(k) : i \in [\![0, 2n-1]\!]\} = [\![0, 2n-1]\!]$.

Exercice 6. (*Théorèmes limites*)

1. (*Théorème local limite*) Soit $\varepsilon > 0$. Montrer que, quand $n \to \infty$, on a, uniformément en $a \in \mathbb{Z}$,

$$\mathbb{P}(S_n = a) = \left(\frac{2\mathrm{e}^{-a^2/2n}}{\sqrt{2\pi n}} + O\Big(n^{\varepsilon - \frac{3}{2}}\Big)\right)1_{a \in n+2\mathbb{Z}}.$$

2. (*Théorème central limite*) Soit $x < y$. Montrer que

$$\mathbb{P}\left(\frac{S_n}{\sqrt{n}} \in [x, y]\right) \xrightarrow[n\to\infty]{} \int_x^y \frac{\mathrm{e}^{-z^2/2}}{\sqrt{2\pi}}\,\mathrm{d}z.$$

3. Soit $x \in \mathbb{R}$. En utilisant l'exercice 2, déterminer la limite, quand $n \to \infty$, de la probabilité qu'il y ait au moins $x\sqrt{n}$ changements de signes avant l'instant $2n+1$.

Corrigé.

1. Soit a et n de même parité, on suppose tout d'abord que $|a| \le n^\alpha$ avec $0 < \alpha < 1$ que l'on choisira plus tard, de sorte que $a/n \to 0$, $n - a \to \infty$ et $n + a \to \infty$. On rappelle la version suivante de la formule de Stirling :

$$n! = \sqrt{2\pi n}\left(\frac{n}{\mathrm{e}}\right)^n\left(1 + O\Big(\frac{1}{n}\Big)\right).$$

On obtient, uniformément en $a \in n + 2\mathbb{Z}$,

$$\begin{aligned}
\mathbb{P}(S_n = a) &= \frac{1}{2^n}\binom{n}{(n-a)/2} \\
&= \frac{1}{2^n}\frac{\sqrt{2\pi n}\left(\frac{n}{\mathrm{e}}\right)^n\left(1 + O\big(\frac{1}{n}\big)\right)}{\sqrt{\pi(n-a)}\left(\frac{n-a}{2\mathrm{e}}\right)^{(n-a)/2}\left(1 + O\big(\frac{1}{n-n^\alpha}\big)\right)\sqrt{\pi(n+a)}\left(\frac{n+a}{2\mathrm{e}}\right)^{(n+a)/2}\left(1 + O\big(\frac{1}{n-n^\alpha}\big)\right)} \\
&= \sqrt{\frac{2}{\pi}}\sqrt{\frac{n}{n^2 - a^2}}\left(1 - \frac{a}{n}\right)^{-(n-a)/2}\left(1 + \frac{a}{n}\right)^{-(n+a)/2}\left(1 + O\Big(\frac{1}{n}\Big)\right).
\end{aligned}$$

Or, on a, toujours uniformément en $a \in n + 2\mathbb{Z}$,

$$\begin{aligned}\left(1+\frac{a}{n}\right)^{-(n+a)/2} &= \exp\left(-\frac{n+a}{2}\ln\left(1+\frac{a}{n}\right)\right) = \exp\left(-\frac{n+a}{2}\left(\frac{a}{n}-\frac{a^2}{2n^2}+\frac{a^3}{3n^3}+O\left(\frac{a^4}{n^4}\right)\right)\right)\\ &= \exp\left(-\frac{1}{2}\left(a+\frac{a^2}{n}-\frac{a^2}{2n}-\frac{a^3}{2n}+\frac{a^3}{3n^2}+O\left(n^{4\alpha-3}\right)\right)\right)\\ &= \exp\left(-\frac{a}{2}-\frac{a^2}{4n}+\frac{a^3}{4n}-\frac{a^3}{6n^2}+O\left(n^{4\alpha-3}\right)\right),\end{aligned}$$

donc, les termes d'ordre impair en a se simplifiant, on obtient

$$\left(1-\frac{a}{n}\right)^{-(n-a)/2}\left(1+\frac{a}{n}\right)^{-(n+a)/2} = \exp\left(-\frac{a^2}{2n}+O\left(n^{4\alpha-3}\right)\right) = \mathrm{e}^{-a^2/2n}\left(1+O\left(n^{4\alpha-3}\right)\right).$$

D'autre part, on a

$$\sqrt{\frac{n}{n^2-a^2}} = \frac{1}{\sqrt{n}}\left(1-\left(\frac{a}{n}\right)^2\right)^{-1/2} = \frac{1}{\sqrt{n}}\left(1+O\left(\frac{n^{2\alpha}}{n^2}\right)\right)$$

et donc en regroupant tout on en arrive à

$$\mathbb{P}(S_n = a) = \sqrt{\frac{2}{\pi n}}\mathrm{e}^{-a^2/2n}\left(1+O\left(\frac{1}{n}\right)+O\left(n^{2\alpha-2}\right)+O\left(n^{4\alpha-3}\right)\right).$$

On choisit alors $\alpha = 1/2+\varepsilon/4$. On a ainsi, pour tout $|a| \leq n^\alpha$ de même parité que n,

$$\left|\mathbb{P}(S_n = a)-\sqrt{\frac{2}{\pi n}}\mathrm{e}^{-a^2/2n}\right| \leq \sqrt{\frac{2}{\pi n}}\mathrm{e}^{-a^2/2n}\frac{C}{n^{1-\varepsilon}} \leq \frac{C'}{n^{\frac{3}{2}-\varepsilon}},$$

avec C et C' des constantes. D'autre part, pour tout $a > n^\alpha$ de même parité que n (le cas $a < -n^\alpha$ en découle par symétrie), on a, en utilisant la décroissance de $a \mapsto \mathbb{P}(S_n = a)$ est décroissant sur les entiers positifs a de même parité que n,

$$\begin{aligned}\left|\mathbb{P}(S_n = a)-\sqrt{\frac{2}{\pi n}}\mathrm{e}^{-a^2/2n}\right| &\leq \mathbb{P}(S_n = a)+\sqrt{\frac{2}{\pi n}}\mathrm{e}^{-a^2/2n} \leq \mathbb{P}(S_n = \lfloor n^\alpha\rfloor)+\sqrt{\frac{2}{\pi n}}\mathrm{e}^{-n^{2\alpha}/2n}\\ &\leq \sqrt{\frac{2}{\pi n}}\mathrm{e}^{-(\lfloor n^\alpha\rfloor-1)^2/2n}+\frac{C'}{n^{\frac{3}{2}-\varepsilon}}+\sqrt{\frac{2}{\pi n}}\mathrm{e}^{-n^{\varepsilon/2}/2},\end{aligned}$$

en appliquant le résultat précédent à $a = \lfloor n^\alpha\rfloor$ ou $a = \lfloor n^\alpha\rfloor - 1$ selon la parité de n. Comme $\mathrm{e}^{-(\lfloor n^\alpha\rfloor-1)^2/2n}$ et $\mathrm{e}^{-n^{\varepsilon/2}/2}$ décroissent plus vite que tout polynôme en n, on en conclut que, pour tout $a > n^\alpha$ de même parité que n,

$$\left|\mathbb{P}(S_n = a)-\sqrt{\frac{2}{\pi n}}\mathrm{e}^{-a^2/2n}\right| \leq \frac{C''}{n^{\frac{3}{2}-\varepsilon}},$$

avec C'' une constante. Cela montre le théorème local limite.

2. On note I l'ensemble des entiers appartenant à $[x\sqrt{n}, y\sqrt{n}]$ et de même parité que n. On a

$$\mathbb{P}\left(\frac{S_n}{\sqrt{n}} \in [x,y]\right) = \sum_{a\in I}\mathbb{P}(S_n = a) = O\left(\frac{\#I}{n^{\frac{3}{2}-\varepsilon}}\right)+\sum_{a\in I}\frac{2\mathrm{e}^{-a^2/2n}}{\sqrt{2\pi n}},$$

en utilisant la question précédente avec uniformité en a. On choisit $\varepsilon < 1$ de sorte que $\#I/n^{\frac{3}{2}-\varepsilon} \to 0$ car $\#I \sim (y-x)\sqrt{n}/2$. Enfin, en notant $f : z \mapsto \mathrm{e}^{-z^2/2}/\sqrt{2\pi}$, on remarque que la dernière somme sur $a \in I$ est une somme de Riemann pour f sur $[x,y]$ avec pour pas de subdivision $2/\sqrt{n}$ (seuls les deux segments extrémaux sont de longueur différente mais leur contribution est négligeable) : ainsi on obtient

$$\sum_{a\in I} \frac{2\mathrm{e}^{-a^2/2n}}{\sqrt{2\pi n}} \xrightarrow[n\to\infty]{} \int_x^y f(z)\,\mathrm{d}z$$

Cela montre le théorème central limite

Remarque. Pour une marche aléatoire quelconque, avec des sauts de variance finie, le théorème central limite sera démontré dans le cours, d'une manière moins douloureuse qu'ici. Cependant, le théorème local limite dans ce cadre de généralité est lui plus difficile à démontrer et ne sera pas vu dans le cadre de ce cours (en outre, on ne peut pas avoir un terme correctif aussi précis).

3. Montrons tout d'abord que le théorème central limite peut aussi s'écrire

$$\mathbb{P}\left(\frac{S_n}{\sqrt{n}} \geq x\right) \xrightarrow[n\to\infty]{} \int_x^\infty \frac{\mathrm{e}^{-z^2/2}}{\sqrt{2\pi}}\,\mathrm{d}z.$$

Par symétrie, on a

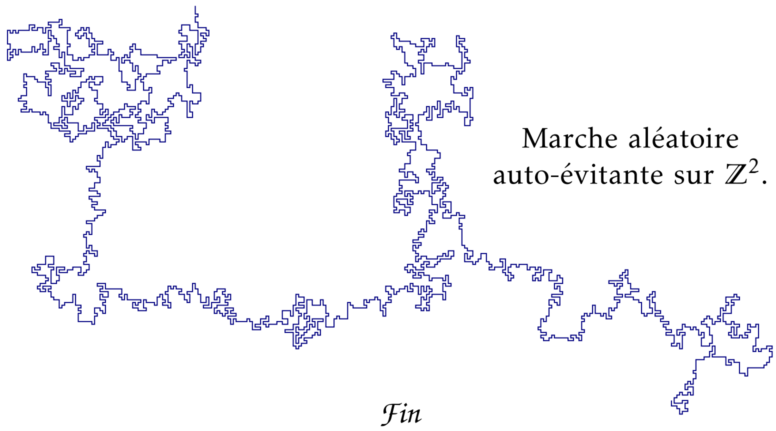
$$\begin{aligned}
\left|\mathbb{P}\left(\frac{S_n}{\sqrt{n}} \geq x\right) - \int_x^\infty \frac{\mathrm{e}^{-z^2/2}}{\sqrt{2\pi}}\,\mathrm{d}z\right| &= \frac{1}{2}\left|\mathbb{P}\left(\frac{|S_n|}{\sqrt{n}} \geq x\right) - \int_{]-\infty,x]\cup[x,\infty[} \frac{\mathrm{e}^{-z^2/2}}{\sqrt{2\pi}}\,\mathrm{d}z\right| \\
&= \frac{1}{2}\left|\left(1-\mathbb{P}\left(\frac{|S_n|}{\sqrt{n}} \geq x\right)\right) - \left(1-\int_{]-\infty,x]\cup[x,\infty[} \frac{\mathrm{e}^{-z^2/2}}{\sqrt{2\pi}}\,\mathrm{d}z\right)\right| \\
&= \frac{1}{2}\left|\mathbb{P}\left(\frac{S_n}{\sqrt{n}} \in]-x,x[\right) - \int_{-x}^x \frac{\mathrm{e}^{-z^2/2}}{\sqrt{2\pi}}\,\mathrm{d}z\right| \\
&\xrightarrow[n\to\infty]{} 0,
\end{aligned}$$

par la question précédente, car on a $\mathbb{P}(S_n = x\sqrt{n}) = \mathbb{P}(S_n = -x\sqrt{n}) \to 0$ (par le théorème local limite par exemple, ou par le théorème central limite appliqué à $x = y$).

On note ξ_{2n+1} le nombre de changements de signe qui ont lieu avant l'instant $2n+1$ et on rappelle que, pour tout $r \in \mathbb{N}$, $\mathbb{P}(\xi_{2n+1} = r) = 2\mathbb{P}(S_{2n+1} = 2r+1)$ (cf exercice 2). Ainsi, pour $x \in \mathbb{R}$, on a

$$\mathbb{P}\left(\xi_{2n+1} \geq x\sqrt{n}\right) = 2\mathbb{P}\left(S_{2n+1} \geq 2x\sqrt{n}+1\right) = 2\mathbb{P}\left(\frac{S_{2n+1}}{\sqrt{2n+1}} \geq \frac{2x\sqrt{n}+1}{\sqrt{2n+1}}\right) \xrightarrow[n\to\infty]{} 2\int_x^\infty \frac{\mathrm{e}^{-z^2/n}}{\sqrt{2\pi}}\,\mathrm{d}z,$$

en utilisant la nouvelle version du théorème central limite montrée juste avant.



Marche aléatoire auto-évitante sur $\mathbb{Z}^2$.

Fin