

Marches aléatoires

Ces travaux dirigés suivent Grimmett G., Stirzaker D., Probability and Random Processes,, section 3.9 pages 71 et suivantes. Nous suivrons la présentation élémentaire (sans le formalisme martingales et chaînes de Markov). On montre tous les résultats par des raisonnements élémentaires, le principe de réflexion et un usage habile de la fonction génératrice. Cette présentation élémentaire est à connaître. L'option probabilité verra des résultats complémentaires, et retrouvera les mêmes, grâce aux martingales et aux chaînes de Markov. Le nombre des questions qu'on peut se poser sur le jeu de pile ou face est immense. En fait, la liste des questions et réponses présentée dépend aussi de l'accessibilité technique du résultat. Un ouvrage entier classique, en deux tomes, le Feller, est entièrement consacré au jeu de pile ou face. Voici une liste succincte des questions que l'on se pose avec une indication des réponses démontrées dans ce chapitre.

Soit X_n une suite de variables de Bernoulli i.i.d. centrées de paramètre p , c'est-à-dire $\mathbb{P}(X_n = 1) = p$ et $\mathbb{P}(X_n = -1) = 1 - p = q$. On pose $S_n = S_0 + \sum_{i=1}^n X_i$. On peut interpréter S_n comme la fortune d'un joueur de roulette ou de pile ou face. La probabilité de gagner pour le joueur est p et sa mise est supposée unitaire. X_n désigne le n -ième tirage. Par exemple si le joueur joue à rouge ou noir, sa probabilité est $\frac{18}{37}$, car il y a 36 chiffres colorés dont 18 noirs et 18 rouges, plus le zéro. La banque ramasse tout quand le zéro sort. (Lire *Le Joueur* de Dostoïevsky pour tous les détails, probablement la plus belle analyse littéraire et psychologique du jeu de roulette mais avec une suite d'observations concrètes). La suite S_n s'appelle une *marche aléatoire simple*. Elle correspond aussi au mouvement aléatoire d'une particule sur \mathbb{N} , progressant de $+1$ avec probabilité p et reculant de -1 avec probabilité q . La marche est appelée symétrique si $p = q = \frac{1}{2}$.

– S_n est-elle une chaîne de Markov (le futur ne dépend que du présent ?). Oui, calculs élémentaires de probabilité conditionnelle.

– Ruine du joueur dans le cas symétrique : le joueur commence avec une fortune de k , la banque a une fortune de N : le jeu s'arrête-t-il en temps fini ? (oui, technique de temps d'arrêt élémentaire).

Espérance du temps de jeu : $k(N - k)$.

Probabilité de ruine du joueur : $p_k = 1 - \frac{k}{N}$.

Ces résultats peuvent/doivent faire l'objet d'une discussion (la meilleure illustration est le film de Luigi Comencini, *L'argent de la vieille*).

– Le retour à l'origine partant de l'origine en dimension 1. Probabilité égale à $1 - |p - q|$.
Technique : usage intensif des fonctions génératrices de distributions discrètes. Egalement par la même technique : calcul de la probabilité de retour en exactement k pas et calcul des probabilités d'atteindre 1, 2, r partant de 0.

Recurrence en dimension inférieure ou égal à 3 : Fourier et intégrabilité de $1/\|x\|^2$ en 0 en dimension d , combinatoire et divergence de la série $\sum 1/k^\alpha$.

toujours recurrent nulle, le temps se lit sur la queue de distribution

- Probabilité d'aller de $a > 0$ en $b > 0$ sans passer par 0 (c'est-à-dire sans se ruiner). Ce calcul utilise l'astuce remarquable du **principe de réflexion**.

Une suite de résultats sur la probabilité des longueurs d'excursions sans ruine permet d'établir le **théorème de ballotage ou de scrutin**, qui donne la probabilité (étonnamment élevée) que lors d'un dépouillement de votes le résultat du dépouillement partiel contredise le résultat final.

La **loi de l'arcsinus** évalue la probabilité que dans un jeu un des joueurs soit gagnant ($S_n > 0$) sur une très longue période. Elle est étonnamment longue : si on note T_n le dernier passage par zéro avant l'instant n ,

$$\mathbb{P}(T_n \leq xn) \xrightarrow{n \rightarrow \infty} \frac{2}{\pi} \text{Arcsin}(\sqrt{x}).$$

L'interprétation de ces résultats permet de comprendre mieux le phénomène du jeu de hasard, l'existence des casinos, les suspenses paradoxaux provoqués par le jeu.

Lors d'une leçon sur pile ou face, il est essentiel d'interpréter chaque résultat donné.

1 Propriétés élémentaires

Exercice 1.1. 1) Représentation graphique standard : tirer à pile ou face 10 fois. Noter les résultats. Représenter les résultats avec en abscisse n et en ordonnée la valeur de S_n , avec $S_0 = 0$. Joindre les points par des segments de droite.

2) Homogénéité spatiale. Montrer que

$$\mathbb{P}(S_n = j | S_0 = a) = \mathbb{P}(S_n = j + b | S_0 = a + b). \quad (1.1)$$

Indication : Montrer que les deux cotés sont égaux à $\mathbb{P}(\sum_1^n X_i = j - a)$.

3) Homogénéité temporelle Montrer que

$$\mathbb{P}(S_n = j | S_0 = a) = \mathbb{P}(S_{m+n} = j | S_m = a). \quad (1.2)$$

Indication : Le membre de gauche est $\mathbb{P}(\sum_1^n X_i = j - a)$ et le membre de droite est $\mathbb{P}(\sum_{m+1}^{m+n} X_i = j - a)$.

4) Les marches aléatoires simples ont la propriété de Markov

$$\mathbb{P}(S_{m+n} = j | S_0, S_1, \dots, S_m) = \mathbb{P}(S_{m+n} = j | S_m), \quad n \geq 0. \quad (1.3)$$

D'abord, comprenons bien ce que veut dire cette relation, à savoir que pour toutes les valeurs possibles de a_0, a_1, \dots, a_m , on a

$$\mathbb{P}(S_{m+n} = j | S_0 = a_0, S_1 = a_1, \dots, S_m = a_m) = \mathbb{P}(S_{m+n} = j | S_m = a_m). \quad (1.4)$$

La "démonstration" de [Grimett Strizacker] est la suivante "If one knows the value of S_m , then the distribution of S_{m+n} depends only on the jumps X_{m+1}, \dots, X_{m+n} and cannot depend on further information concerning the values of S_0, S_1, \dots, S_{m-1} . Vous paraît-elle suffisante?"

5) Montrer (1.3) en menant les calculs grâce à (1.4) et en utilisant la définition de la probabilité conditionnelle.

Indication :

$$\begin{aligned} \mathbb{P}(S_{m+n} = j | (S_0 \dots S_m) = (a_0 \dots a_m)) &= \mathbb{P}\left(\sum_{m+1}^n X_i = j - a_m | (S_0 \dots S_m) = (a_0 \dots a_m)\right) = \\ &= \frac{\mathbb{P}(\sum_{m+1}^n X_i = j - a_m) \mathbb{P}((S_0 \dots S_m) = (a_0 \dots a_m))}{\mathbb{P}((S_0 \dots S_m) = (a_0 \dots a_m))} = \mathbb{P}\left(\sum_{m+1}^n X_i = j - a_m\right) = \\ &= \mathbb{P}(S_{m+n} = j | S_m = a_m). \end{aligned}$$

2 La ruine du joueur

Exercice 2.1. Pile ou face-La ruine du joueur, Grimmett Stirzacker, 1.7 p. 17 et 3.9 p. 74. Un joueur a une fortune de k et il désire atteindre en jouant à pile ou face une fortune de N , $0 < k < N$. Il parie toujours sur pile, qui a probabilité p . La question est de savoir avec quelle probabilité il se ruinera avant d'atteindre son but. On désigne par A_k l'événement "ruine du joueur, partant d'une fortune $S_0 = k$ " et par B l'événement que le résultat du premier tirage est pile. On munit l'espace Ω des suites infinies de tirages de la probabilité \mathbb{P} . On note $\omega = (\omega_1, \omega_2, \dots, \omega_n, \dots)$ un élément de Ω . On note également $\omega = (\omega_1, \omega^1)$ avec $\omega^1 = (\omega_2, \dots, \omega_n, \dots)$.

0) Montrer que le joueur se ruine ou atteint son but p.s.

1) Donner une définition formelle de l'événement "ruine du joueur". On utilisera le temps d'arrêt

$$t_k(\omega) = \inf\{n, k + \omega_1 + \dots + \omega_n = 0 \text{ ou } N\},$$

qui est bien défini grâce à la question précédente. Montrer que t_k est mesurable. Montrer que pour $1 < k < N - 1$,

$$\mathbb{P}(A_k) = p\mathbb{P}(A_{k+1}) + q\mathbb{P}(A_{k-1}).$$

On commencera par interpréter cette formule. Vérifier que cette formule est encore valable pour $k = 1$ ou $N - 1$ en remarquant que $\mathbb{P}(A_0) = 1$ et $\mathbb{P}(A_N) = 0$

Indication : montrer que pour $\omega \in \{\omega_1 = 1\}$, on a $t_k(\omega) = t_{k+1}(\omega^1) + 1$. Attention : ni Grimmett-Stirzacker ni Feller ne jugent bon de donner cette preuve. Le formalisme de Grimmett-Stirzacker est discutable : ils notent $\mathbb{P}_k(A)$ ce que nous notons $\mathbb{P}(A_k)$, l'événement A étant alors un événement absolu appelé "ruine du joueur". L'ennui, c'est que cet événement dépend du point de départ k et donc, A a de toutes manières une signification différente quand k change. Il vaut mieux le noter A_k et garder le même espace de probabilité Ω pour toutes les marches aléatoires. Pour le reste, nous suivons Grimmett Stirzacker [?].

2) On pose $p_k = \mathbb{P}(A_k)$. Si $p = \frac{1}{2}$, déduire que $p_k = \frac{1}{2}(p_{k+1} + p_{k-1})$, $0 < k < N$.

3) En utilisant les valeurs de p_0 et p_N , montrer que

$$\mathbb{P}(A_k) = 1 - \frac{k}{N}.$$

4) Interpréter cette relation. On pourra s'inspirer comme suggéré par Grimmett-Stirzaker de "Millionaires should always gamble, poor men never" (J.M. Keynes). Et voir *L'argent de la vieille* de Luigi Comencini.

5) Plus généralement, on a l'équation $p_k = p \cdot p_{k+1} + q p_{k-1}$, $1 \leq k \leq N-1$, avec conditions au bord $p_0 = 1$, $p_N = 0$. Montrer que

$$p_k = \frac{\left(\frac{q}{p}\right)^k - \left(\frac{q}{p}\right)^N}{1 - \left(\frac{q}{p}\right)^N}.$$

6) En déduire q_k , la probabilité que le joueur fasse fortune.

7) Vérifier que $p_k + q_k = 1$. Interpréter.

8) Montrer que le temps d'arrêt est presque sûrement fini.

9) On pose $D_k = \mathbb{E}(t_k)$. Grimmett-Stirzaker n'écrit pas cette définition mais décrit D_k comme "the mean number of steps before the particle hits one of the absorbing barriers, starting from k ". Justifier la relation proposée par ces auteurs

$$D_k = p(1 + D_{k+1}) + q(1 + D_{k-1}), \quad 1 \leq k \leq N-1, \quad D_0 = D_N = 0.$$

Pour cela, on appliquera au temps d'arrêt les formules de l'espérance conditionnelle pour des variables aléatoires discrètes et des événements B que l'on rappelle :

$$\mathbb{E}(Y) = \sum_x \mathbb{E}(Y|X = x) \mathbb{P}(X = x), \quad (2.1)$$

$$\mathbb{E}(Y|B) = \sum_l l \mathbb{P}(Y = l|B). \quad (2.2)$$

10) En déduire que

$$D_k = \frac{1}{q-p} \left[k - N \left(\frac{1 - \left(\frac{q}{p}\right)^k}{1 - \left(\frac{q}{p}\right)^N} \right) \right] \quad \text{si } p \neq \frac{1}{2},$$

$$D_k = k(N-k) \quad \text{si } p = \frac{1}{2}.$$

11) Interpréter ces résultats.

Solutions :

0) Quel que soit l'état initial $i \in \{0, \dots, N\}$, la probabilité d'être ruiné ou d'atteindre son but au bout d'une séquence de N tirages consécutifs est plus grande qu'une constante positive $a > 0$. Par indépendance de ces séquences successives, le nombre de séquence pour être ruiné ou atteindre son but est dominé par une loi géométrique de paramètre a , qui est finie p.s.

1) On a

$$A_k = \{t_k(\omega) = \inf\{n, k + \omega_1 + \dots + \omega_n = 0 \text{ ou } N\} < \infty \text{ et } k + \omega_1 + \dots + \omega_{t_k(\omega)} = 0\}.$$

On écrit $A_k = (A_k \cap \{\omega_1 = 1\}) \cup (A_k \cap \{\omega_1 = -1\})$. Mais pour $\omega \in \{\omega_1 = 1\}$, on a

$$t_k(\omega) = \inf\{n, k + 1 + \omega_2 + \dots + \omega_n = 0 \text{ ou } N\} < \infty = t_{k+1}(\omega^1) + 1.$$

On en déduit que

$$A_k \cap \{\omega_1 = 1\} = \{t_{k+1}(\omega^1) < \infty \text{ et } k + 1 + \omega_2 + \dots + \omega_{t_{k+1}(\omega^1)} = 0\} \cap \{\omega_1 = 1\}.$$

Comme les X_i sont i.i.d., on obtient

$$\mathbb{P}(A_k \cap \{\omega_1 = 1\}) = p\mathbb{P}(A_{k+1}).$$

On conclut aisément.

2), 3) 4) Immédiat, simple calcul. Le résultat nous dit que plus on est pauvre, moins on a de chances de devenir riche. L'interprétation est toutefois facilitée si on considère le jeu de manière symétrique : le premier joueur possède k et le second possède $N - k$. On peut alors reformuler la règle du jeu ainsi : le jeu s'arrête quand l'un des deux joueurs est ruiné. La probabilité de ruine du premier joueur est $\frac{N-k}{N}$ et celle du second $\frac{k}{N}$. L'une tend vers 1 et l'autre vers 0 quand $\frac{N}{k}$ tend vers l'infini. Les riches gagnent donc presque toujours. Mais, quand ils perdent, il perdent gros, comme on peut le vérifier en remarquant que l'espérance des gains des deux joueurs au moment de l'arrêt du jeu sont les mêmes et égales à zéro. En effet, l'espérance du gain du premier joueur est $-k\mathbb{P}(A_k) + (N - k)(1 - \mathbb{P}(A_k)) = 0$ et on vérifie immédiatement que celle du second est aussi zéro. Ce résultat est connu sous le nom de théorème d'arrêt de Doob (Williams page 100), mais pour appliquer directement ce théorème il faudrait d'abord démontrer que le temps d'arrêt $t_k(\omega)$ est presque sûrement fini. Une autre manière de calculer $\mathbb{P}(A_k)$ est donc, pour ceux qui connaissent les martingales, de :

- montrer que le temps d'arrêt du jeu est presque sûrement fini.
- utiliser le théorème d'arrêt qui dit que sous cette condition, $\mathbb{E}S_{t_k} = \mathbb{E}S_0 = k$, ce qui veut exactement dire que l'espérance de gain est nulle.

Il est quand même préférable de savoir que l'on peut tout faire avec des moyens élémentaires !

5) Calculs classiques avec la formule de récurrence.

6) La valeur de q_k s'interprète comme une probabilité de ruine en échangeant les rôles de p et q et de k et $N - k$. On obtient donc $q_k = \frac{\left(\frac{p}{q}\right)^{N-k} - \left(\frac{p}{q}\right)^N}{1 - \left(\frac{p}{q}\right)^N}$.

7) 8) La relation est vite vérifiée. On en déduit que presque sûrement, le joueur se ruine ou fait fortune en temps fini, ce qui veut implique que t_k est presque sûrement fini.

9) On a en utilisant (2.2),

$$\begin{aligned} \mathbb{E}(t_k | X_1 = -1) &= \sum_l l \frac{\mathbb{P}(t_k = l; X_1 = -1)}{\mathbb{P}(X_1 = -1)} = \sum_l l \frac{\mathbb{P}(\inf\{n, k - 1 + X_2 + \dots + X_n = 0 \text{ ou } N\} = l)q}{q} \\ &= \sum_l l \mathbb{P}(\inf\{n, k - 1 + X_1 + \dots + X_{n-1} = 0 \text{ ou } N\} = l) \\ &= \sum_l l \mathbb{P}(\inf\{n - 1, k - 1 + X_1 + \dots + X_{n-1} = 0 \text{ ou } N\} = l - 1) \\ &= 1 + \sum_l (l - 1) \mathbb{P}(t_{k-1} = l - 1) = 1 + \mathbb{E}t_{k-1} = 1 + D_{k-1}. \end{aligned}$$

On a utilisé le fait que $X_2 + \dots + X_n$ et $X_1 + \dots + X_{n-1}$ ont la même loi, l'indépendance de X_1 et des X_i pour $i \geq 2$ et le fait que t_k est presque sûrement fini. On a de même

$$\mathbb{E}(t_k | X_i = +1) = 1 + D_{k+1}.$$

En utilisant la formule (2.1), appliquée à $Y = t_k$ et $X = X_1$, on obtient comme demandé

$$D_k = p(1 + D_{k+1}) + q(1 + D_{k-1}).$$

10)-11) Simple calcul. L'interprétation n'est pas très excitante si p est très différent de q , car alors l'un des joueurs a beaucoup plus de chances que l'autre. On va donc interpréter le cas $p = 1/2$ et remarquer qu'aux jeux de casino, p est assez proche de $1/2$ et la première formule donnée pour D_k peut alors s'interpréter dans la même ligne que la seconde. Le point important qui a été démontré est que

le temps de jeu $D_k = k(N - k)$ est proportionnel au produit des fortunes des deux joueurs.

Ceci explique pourquoi les casinos existent. On doit d'une part avoir $N - k$ très grand pour que le casino ne perde jamais. Cela garantit de plus un temps moyen de jeu élevé pour tous les joueurs et donc permet de les "accrocher". La fortune du casino étant grande et fixée, on remarque que la même formule donne un temps de jeu à peu près proportionnel à la fortune du joueur. Pour s'amuser longtemps dans un casino, il vaut mieux faire de petites mises, et, à fortune égale, le temps de jeu est grosso modo double si les mises sont deux fois plus petites. On voit surgir d'intéressantes questions de stratégie pour le casino : il faut imposer des mises pas trop petites, car sinon le client risque de ressortir par simple ennui, avant d'avoir tout perdu. Inversement, une mise minimale trop grosse va chasser trop vite les petits clients qui auront manqué les plaisirs des "hauts et bas" qui font le charme du jeu. Il y a donc un problème d'optimisation de la mise minimale imposée en fonction de la fortune moyenne des clients et de l'estimation du temps de jeu optimal (une fin de soirée). L'idéal est que les clients aient le temps de s'amuser avant de sortir les poches vides.

3 Temps de retour des marches aléatoires

Exercice 3.1. Le retour à l'origine

On considère une marche aléatoire sur \mathbb{Z} ,

$$S_n = X_1 + \dots + X_n, \quad S_0 = 0$$

où les X_i sont des Bernoulli indépendantes prenant la valeur 1 avec probabilité p et la valeur -1 avec la probabilité $1 - p = q$.

On pose $p_0(n) = \mathbb{P}(S_n = 0)$ et pour $n \geq 1$, $f_0(n) = \mathbb{P}(S_1 \neq 0, \dots, S_{n-1} \neq 0, S_n = 0)$ la probabilité que le premier retour en 0 se fasse après n pas. On note les fonctions génératrices associées

$$P_0(s) = \sum_0^{\infty} p_0(n)s^n, \quad F_0(s) = \sum_{n=1}^{\infty} f_0(n)s^n.$$

On va montrer le théorème suivant

Théorème 3.1. *On a*

(a) $P_0(s) = 1 + P_0(s)F_0(s),$

(b) $P_0(s) = (1 - 4pqs^2)^{-\frac{1}{2}},$

(c) $F_0(s) = 1 - (1 - 4pqs^2)^{\frac{1}{2}}.$

1) Soit $A = \{S_n = 0\}$ et B_k l'évènement que le premier retour à l'origine se fait à k pas. Remarquer que

$$\mathbb{P}(A) = \sum_{k=1}^n \mathbb{P}(A|B_k)\mathbb{P}(B_k).$$

2) Montrer que $\mathbb{P}(A|B_k) = p_0(n - k).$

3) Dédire que

$$p_0(n) = \sum_{k=1}^n p_0(n - k)f_0(k) \tag{3.1}$$

si $n \geq 1$ et obtenir (a).

4) Montrer par un simple argument combinatoire que $p_0(n) = C_n^{\frac{n}{2}}(pq)^{\frac{n}{2}}$ si n est pair, $= 0$ sinon et déduire (b) et (c).

5) Dédire des résultats précédents que :

- la probabilité de retour à l'origine est $1 - |p - q|$. (Commencer par définir cette probabilité)
- cette probabilité n'est égale à 1 que si $p = \frac{1}{2}$ et alors le temps moyen de retour est infini. (Commencer par définir le temps moyen de retour).

Solutions :

1) *C'est la règle des probabilités disjointes car les B_k sont disjointes.*

2) *On a*

$$\begin{aligned} \mathbb{P}(A \cap B_k) &= \mathbb{P}(S_n = 0, S_k = 0, S_{k-1} \neq 0, \dots, S_1 \neq 0) \\ &= \mathbb{P}(X_{k+1} + \dots + X_n = 0, S_k = 0, S_{k-1} \neq 0, \dots, S_1 \neq 0) \\ &= \mathbb{P}(X_{k+1} + \dots + X_n = 0)\mathbb{P}(B_k) = \mathbb{P}(S_{n-k} = 0)\mathbb{P}(B_k) \end{aligned}$$

car les X_i sont i.i.d. et les lois de $X_1 + \dots + X_{n-k}$ et de $X_{k+1} + \dots + X_n$ sont donc égales. (On peut aussi utiliser l'homogénéité temporelle)

3) *La première relation découle de 1) et 2). La seconde découle du fait que le calcul de la fonction génératrice transforme une convolution de suites en produit. Plus précisément, en utilisant la relation du 3) :*

$$P_0(s)F_0(s) = \sum_1^\infty s^n \left[\sum_{k=1}^n f_0(k)p_0(n - k) \right] = \sum_1^\infty s^n p_0(n) = P_0(s) - 1.$$

4) *$S_n = 0$ si et seulement si la particule fait un nombre égal à $\frac{n}{2}$ de pas vers la droite et un nombre égal vers la gauche. Ceci peut se faire de $C_n^{\frac{n}{2}}$ manières et chacune de ces marches a probabilité $(pq)^{\frac{n}{2}}$.*

Si n est impair il est impossible que $S_n = 0$ et donc $p_0(n) = 0$. Pour (b) : à vos développements en série ! La relation (c) découle de (a) et (b).

5) La probabilité de retour à l'origine est $\sum_{n=1}^{\infty} f_0(n) = F_0(1) = 1 - |p - q|$. Le retour à l'origine n'est certain que si $F_0(1) = 1$ et donc $p = \frac{1}{2}$. Le temps moyen de retour se définit comme l'espérance du temps de premier retour, soit $\mathbb{E}T_0 = \sum_{n=1}^{\infty} n f_0(n) = F_0'(1) = +\infty$. (La seconde égalité vient du théorème de convergence monotone).

Exercice 3.2. Le temps de première visite en 1

Soit T_1 le premier temps (éventuellement infini) pour lequel la marche atteint 1. Soit F_1 la fonction génératrice de $1_{\{T_1 < +\infty\}} T_1$ et pour tout entier n positif ou nul $f_1(n) = \mathbb{P}(T_1 = n)$.

1a) Montrer que $f_1(1) = p$ et que pour tout entier $n \geq 2$, on a la relation de récurrence $f_1(n) = q \sum_{\nu=2}^{n-1} f_1(\nu - 1) f_1(n - \nu)$

1b) En déduire que l'on a $F_1(s) = qsF_1(s)^2 + ps$, puis que F_1 satisfait $F_1(s) = \frac{1 - \sqrt{1 - 4pqs^2}}{2qs}$

2) En déduire la distribution de T_1 :

$$\mathbb{P}(T_1 = 2k) = 0, \quad \mathbb{P}(T_1 = 2k - 1) = \frac{(-1)^{k-1}}{2q} C_{1/2}^k (4pq)^k.$$

Solutions :

1a) Le résultat pour $n = 1$ étant trivial, soit donc $n \geq 2$. On a :

$$\begin{aligned} \mathbb{P}(T_1 = n) &= \mathbb{P}(S_n = 1, S_{n-1} \leq 0, \dots, S_1 \leq 0) \\ &= \sum_{\nu=2}^{n-1} \mathbb{P}(X_1 = -1, \inf\{j \geq 2, -1 + X_2 + \dots + X_j = 0\} = \nu, \inf\{k \geq \nu + 1, X_{\nu+1} + \dots + X_k = 1\} = n) \\ &= q \sum_{\nu=2}^{n-1} \mathbb{P}(\inf\{j \geq 2; -1 + X_2 + \dots + X_j = 0\} = \nu) \mathbb{P}(\inf\{k \geq \nu + 1, X_{\nu+1} + \dots + X_k = 1\} = n) \\ &= q \sum_{\nu=2}^{n-1} \mathbb{P}(\inf\{j \geq 1; X_1 + \dots + X_j = 1\} = \nu - 1) \mathbb{P}(\inf\{k \geq 1, X_1 + \dots + X_k = 1\} = n - \nu) \\ &= q \sum_{\nu=2}^{n-1} f_1(\nu - 1) f_1(n - \nu). \end{aligned}$$

La première égalité résulte du fait que l'on est en dimension 1 et que la marche est plus proche voisin. Puisque $n \geq 2$, c'est que la marche est partie à gauche (sinon $T_1 = 1$) et puisque la marche est plus proche voisin, la marche ne peut redevenir strictement positive sans passer par 1.

La deuxième égalité repose sur l'observation suivante : puisque $n \geq 2$, la marche part à gauche, reste strictement négative jusqu'au temps $\nu - 1$, redevient nulle à l'instant ν (pour la première fois depuis l'instant 0) et au temps n devient strictement positive égale à 1 (et pour la première fois !).

La troisième et quatrième égalité sont de simples conséquences de l'indépendance et du caractère équidistribué des X_k .

1b) On a :

$$\begin{aligned} \sum_{n=1}^{\infty} f_1(n)s^n &= qs \sum_{n=2}^{\infty} \left(\sum_{j=1}^{n-1} f_1(j)f_1(n-1-j) \right) s^{n-1} + ps \\ &= qs \sum_{n=1}^{\infty} \left(\sum_{j=0}^n f_1(j)f_1(n-j) \right) s^n + ps \end{aligned}$$

soit

$$F_1(s) = qsF_1(s)^2 + ps$$

On obtient donc une équation du second degré pour $F_1(s)$ et on a :

$$F_1(s) = \frac{1 \pm \sqrt{1 - 4pqs^2}}{2qs}$$

et puisque $F_1(s)$ tend vers 0 quand $s \rightarrow 0$, on a donc :

$$F_1(s) = \frac{1 - \sqrt{1 - 4pqs^2}}{2qs}.$$

2) En utilisant le développement en série entière de $\sqrt{1-x}$, on obtient le résultat annoncé dans la question 2. Remarquer qu'il était évident que $\mathbb{P}(T_1 = 2k) = 0$ (on doit faire un nombre impair de pas pour de 0 aller en 1).

Exercice 3.3. Le temps de première visite en r

Soit T_r le temps de première visite au point r . On pose $f_r(n) = \mathbb{P}(S_1 \neq r, \dots, S_{n-1} \neq r, S_n = r)$, c'est-à-dire la probabilité que la première visite se produise au n -ième pas. On note sa fonction génératrice $F_r(s) = \sum_{n=1}^{\infty} f_r(n)s^n$.

1) Montrer que $f_r(n) = \sum_{k=1}^{n-1} f_{r-1}(n-k)f_1(k)$ si $r > 1$.

2) En déduire que pour $r \geq 1$,

$$F_r(s) = [F_1(s)]^r.$$

3) Montrer que la probabilité que la marche aléatoire entre dans $\mathbb{N}^* = \{1, 2, \dots, n, \dots\}$ est

$$F_1(1) = \frac{1 - |p - q|}{2q} = \min\left(1, \frac{p}{q}\right).$$

(Cette probabilité est la même que celle de passer par 1, qui est le vestibule de \mathbb{N}^* .)

4) Démontrer, en considérant les probabilités conditionnelles par rapport aux valeurs possibles de X_1 , que $f_0(n) = qf_1(n-1) + pf_{-1}(n-1)$ et que donc $F_0(s) = qsF_1(s) + psF_{-1}(s)$.

5) Montrer que $F_{-1}(s) = \frac{1 - (1 - 4pqs^2)^{\frac{1}{2}}}{2ps}$. Pour cela, on appliquera une *méthode de réflexion* : Soit ω une marche quelconque arrivant en -1 et ω^* la marche symétrique par rapport à 0, qui arrive donc en 1. Soit $\mathbb{P}_{p,q}(\pi)$ la probabilité de π . Alors $\mathbb{P}_{p,q}(\pi) = \mathbb{P}_{q,p}(\pi^*)$.

6) En déduire à nouveau $F_0(s)$: par le principe de réflexion on peut donc déduire directement F_0 de F_1 .

Solutions : 1) On reprendra point par point l'argument qui donnait la relation (3.1).

2) On multiplie par s^n la relation du 1) et on somme sur n pour obtenir $F_r(s) = F_{r-1}(s)F_1(s) = F_1(s)^r$.

4) L'argument suggéré conduit à échanger les rôles de p et q dans la formule donnant F_1 pour obtenir F_{-1} .

5) Simple calcul.

Exercice 3.4. Marche aléatoire simple symétrique dans \mathbb{Z}^d . On utilise la notion de chaîne de Markov (et plus spécifiquement de récurrence transience).

On considère la marche aléatoire issue de 0

$$S_n = X_1 + \dots + X_n$$

où les X_i sont des v.a.i.i.d. de loi commune μ . Le support de μ est inclus dans \mathbb{Z}^d et en notant $e_i = (0, \dots, 0, 1, 0, \dots)$ les vecteurs de la base canonique de \mathbb{R}^d :

$$\mu(e_i) = \mu(-e_i) = \frac{1}{2d}.$$

C'est à dire que le marche aléatoire passe (sans mémoire) d'un point de \mathbb{Z}^d à un point voisin choisi au hasard de manière équiprobable.

On cherche à savoir si la marche va presque sûrement toujours revenir à son point de départ (récurrence) ou si elle peut s'échapper à l'infini.

1. Discuter intuitivement de l'effet de la dimension sur cette propriété de récurrence.
2. Justifier que $(S_n : n \in \mathbb{N})$ est une chaîne de Markov et donner sa matrice de transition.

On appelle ici L^1 l'espace des fonctions sommables $f : \mathbb{Z}^d \rightarrow \mathbb{R}$, i.e. $\sum_{x \in \mathbb{Z}^d} |f(x)| < \infty$. Pour ces fonctions, on définit la transformée de Fourier

$$\hat{f}(\theta) = \sum_{x \in \mathbb{Z}^d} f(x) e^{i\langle \theta, x \rangle} \quad (\theta \in \mathbb{R}^d)$$

et le produit de deux fonctions par

$$f \star g(x) = \sum_{y \in \mathbb{Z}^d} f(y) g(x - y).$$

3. Montrer que si $f, g \in L^1$ alors $\widehat{f \star g} = \hat{f} \hat{g}$.
4. Montrer que si $f \in L^1$ alors \hat{f} est une fonction bornée et on a la formule d'inversion

$$f(x) = \frac{1}{(2\pi)^d} \int_{]-\pi, \pi]^d} \hat{f}(\theta) e^{-i\langle \theta, x \rangle} d\theta \quad (x \in \mathbb{Z}^d).$$

5. Montrer que la transformée de Fourier $\hat{\mu}$ de $x \rightarrow \mu(x)$ vaut

$$\frac{1}{d} \sum_{k=1}^d \cos(\theta_k).$$

Montrer ensuite que $\hat{\mu}(\theta) < 1$ pour $\theta \in]-\pi, \pi]^d \setminus \{0\}$ et que

$$1 - \hat{\mu}(\theta) \sim \frac{|\theta|^2}{2d} \quad \text{quand } \theta \rightarrow 0.$$

6. On définit pour $|\lambda| \leq 1$:

$$u_\lambda(x) = \sum_{n=0}^{\infty} \lambda^n \mathbb{P}(S_n = x).$$

(a) Montrer que $u_\lambda(x)$ croît vers $u(x)$ quand λ croît vers 1.

(b) Montrer que si $|\lambda| < 1$, alors $u_\lambda \in L^1$ et

$$\hat{u}_\lambda(\theta) = \frac{1}{1 - \lambda \hat{\mu}(\theta)}.$$

(c) En déduire dans quel cas $u_1(0)$ est finie.

7. Conclure sur la récurrence de la chaîne.

1. *Quand la dimension augmente, le nombre de chemins augmente et les chances de revenir au point de départ s'amenuisent. Une trajectoire peut avoir son projeté qui retourne au point initial sans elle même le faire. La propriété de récurrence de la chaîne peut donc être perdue en augmentant la dimension.*

2. $p(x, y) = \mu(y - x)$. Autrement dit :

$$p(x, y) = \begin{cases} \frac{1}{2d} & \text{si } |x - y| = 1 \\ 0 & \text{si } |x - y| \neq 1. \end{cases}$$

3. On vérifie qu'on peut appliquer le théorème de Fubini (somme finie en valeur absolue) :

$$\begin{aligned} \widehat{f \star g}(\theta) &= \sum_{x \in \mathbb{Z}^d} \left[\sum_{y \in \mathbb{Z}^d} f(y)g(x - y) \right] e^{i\langle \theta, x \rangle} = \sum_{y \in \mathbb{Z}^d} f(y) \sum_{x \in \mathbb{Z}^d} g(x - y) e^{i\langle \theta, x \rangle} \\ &= \sum_{y \in \mathbb{Z}^d} f(y) e^{i\langle \theta, y \rangle} \sum_{x \in \mathbb{Z}^d} g(x - y) e^{i\langle \theta, x - y \rangle} = \hat{f}(\theta) \hat{g}(\theta). \end{aligned}$$

4. On a $|\hat{f}(\theta)| \leq \sum_{x \in \mathbb{Z}^d} |f(x)| < \infty$ donc \hat{f} est bornée. Calculons

$$\int_{]-\pi, \pi]^d} \hat{f}(\theta) e^{-i\langle \theta, x \rangle} d\theta = \int_{]-\pi, \pi]^d} \sum_{y \in \mathbb{Z}^d} f(y) e^{i\langle \theta, y \rangle} e^{-i\langle \theta, x \rangle} d\theta = \sum_{y \in \mathbb{Z}^d} f(y) \int_{]-\pi, \pi]^d} e^{i\langle \theta, y - x \rangle} d\theta.$$

On conclut en observant que

$$\int_{]-\pi, \pi]^d} e^{i\langle \theta, y - x \rangle} d\theta = (2\pi)^d \delta_{xy}.$$

5. La transformée de Fourier s'écrit

$$\hat{\mu}(\theta) = \sum_{x \in \mathbb{Z}^d} \mu(x) e^{i(\theta, x)} = \frac{1}{2^d} \sum_{k=1}^d (e^{i\theta_k} + e^{-i\theta_k}) = \frac{1}{d} \sum_{k=1}^d \cos(\theta_k).$$

Si $\theta \in (-\pi, \pi]^d \setminus \{0\}$ alors $\cos(\theta_k) < 1$ et donc $\hat{\mu}(\theta) < 1$. Pour $\theta \rightarrow 0$, on fait un développement limité :

$$1 - \hat{\mu}(\theta) = \frac{1}{d} \sum_{k=1}^d [1 - \cos(\theta_k)] \sim \frac{1}{2d} \sum_{k=1}^d \theta_k^2 = \frac{|\theta|^2}{2d}.$$

6. (a) C'est le théorème de convergence monotone pour les séries de fonctions (ou pour la mesure de comptage).
 (b) On remarque que $\mathbb{P}(S_n = x) = \mathbb{P}(X_1 + \dots + X_n = x) = \mu^{*n}(x)$ (récurrence immédiate une fois que l'on a $\mathbb{P}(X + Y = x) = \lambda \star \mu(x)$ pour deux variables indépendantes X et Y de lois λ et μ .) Ensuite, en utilisant 3)

$$\hat{u}_\lambda(\theta) = \sum_{n=0}^{\infty} \lambda^n \widehat{\mathbb{P}(S_n = \cdot)}(\theta) = \sum_{n=0}^{\infty} \lambda^n (\hat{\mu}(\theta))^n,$$

et l'on reconnaît une série géométrique.

(c) Par la question 4),

$$u_\lambda(0) = \frac{1}{(2\pi)^d} \int_{]-\pi, \pi]^d} \frac{1}{1 - \lambda \hat{\mu}(\theta)} d\theta.$$

et par convergence monotone on obtient

$$u_1(0) = \frac{1}{(2\pi)^d} \int_{]-\pi, \pi]^d} \frac{1}{1 - \hat{\mu}(\theta)} d\theta.$$

La fonction intégrée est bornée sauf au voisinage de zéro où elle se comporte comme $1/|\theta|^2$. Elle est donc intégrable en zéro ssi $d \geq 3$ (pour $d = 1$, on intègre $1/x^2$ en 0 ; pour $d = 2$, un changement de coordonnées polaires conduit à intégrer $1/r$ en 0 ; pour $d = 3$ un changement de coordonnées sphériques conduit à intégrer 1 en 0.)

7. On observe que $u_1(0) = \sum_{n \in \mathbb{N}} \mathbb{P}(S_n = 0) = \mathbb{E}(\#\{n \in \mathbb{N} : S_n = 0\})$. Pour un chaîne irréductible, le fait que cette espérance soit infinie caractérise le fait que la chaîne est récurrente. En effet si 0 est récurrent, le nombre de retour en 0 est p.s. infini et donc l'espérance est infinie. Si 0 est transitoire, alors le nombre de retours suit une loi géométrique de paramètre < 1 (propriété de Markov) et donc son espérance est finie. Donc la chaîne est récurrente ssi $d \leq 2$.

4 Le principe de réflexion et ses conséquences

Grimmett G., Stirzaker D., Probability and Random Processes, 3.10, pages 75 à 82.

Dans la suite, on pose $S_0 = a$, $S_n = a + \sum_{i=1}^n X_i$ où X_i sont des variables de Bernoulli i.i.d. de paramètre p : $\mathbb{P}(X_1 = 1) = p$, $\mathbb{P}(X_1 = -1) = 1 - p = q$. On représente une marche aléatoire par la suite des couples temps-position, $\{(n, S_n), n \geq 0\}$. On note $N_n(a, b)$ le nombre de chemins distincts allant de $(0, a)$ à (n, b) et $N_n^0(a, b)$ le nombre de chemins allant de $(0, a)$ à (n, b) et passant au moins une fois par 0 : de tels chemins contiennent $(k, 0)$ pour un k tel que $0 \leq k \leq n$.

Exercice 4.1. 1) Démontrer que

$$\mathbb{P}(S_n = b) = \begin{cases} C_n^{\frac{1}{2}(n+b-a)} p^{\frac{1}{2}(n+b-a)} q^{\frac{1}{2}(n-b+a)} & \text{si } \frac{1}{2}(n+b-a) \in \{0, 1, \dots, n\} \\ 0 & \text{sinon} \end{cases}$$

2) *Le principe de réflexion.* On suppose que $S_0 = a$ et $S_n = b$. Montrer que si $a, b > 0$, alors $N_n^0(a, b) = N_n(-a, b)$. On appliquera le principe de réflexion, qui consiste à associer à tout chemin de $(0, -a)$ à (n, b) passant pour la première fois en 0 à l'instant k le chemin de $(0, a)$ à (n, b) obtenu en changeant en leur opposé tous les points avant k . (Faire un dessin où l'axe des x représente le temps et l'axe des y la position S_n . La réflexion se fait par rapport à l'axe Ox .)

3) Montrer que pour tous $n \in \mathbb{N}$, $a < b$, tels que $0 \leq \frac{1}{2}(n+b-a) \leq n$, on a $N_n(a, b) = C_n^{\frac{1}{2}(n+b-a)}$, et $N_n(a, b) = 0$ si cette condition sur n, a, b n'est pas satisfaite.

4) Montrer le *théorème de scrutin* : Si $b > 0$, alors le nombre de chemins de $(0, 0)$ à (n, b) qui ne repassent pas par 0 est égal à $\frac{b}{n} N_n(0, b)$.

5) On prend $a = 0$. Montrer que

$$\mathbb{P}(S_1 S_2 \dots S_n \neq 0, S_n = b) = \frac{|b|}{n} \mathbb{P}(S_n = b)$$

6) En déduire que

$$\mathbb{P}(S_1 S_2 \dots S_n \neq 0) = \frac{1}{n} \mathbb{E}|S_n|.$$

7) Interpréter ces résultats.

Solutions : 1) Si $S_n = b$, et s'il y a eu r pas vers la droite et l pas vers la gauche, on a $a+r-l = b$ et $r+l = n$. On trouve $r = \frac{1}{2}(n+b-a)$ et $l = \frac{1}{2}(n-b+a)$. Le nombre de chemins distincts est le nombre de manières de placer les r pas vers la droite parmi les n pas, soit $C_n^r = C_n^{\frac{1}{2}(n+b-a)}$. Tous les chemins ont la même probabilité $p^r q^l$, d'où le résultat.

2) La réflexion définit une bijection entre les chemins de a à b passant par 0 et les chemins de $-a$ à b . Leurs nombres sont donc égaux.

3) Ce raisonnement a déjà été fait au 1).

4) Comme tous les chemins ne repassant pas par 0 vont d'abord en $(1, 1)$, leur nombre est $N_{n-1}(1, b) - N_{n-1}^0(1, b) = N_{n-1}(1, b) - N_{n-1}(-1, b)$, par le 2). Un calcul facile utilisant 3) donne le résultat.

5) Prenons d'abord $b > 0$. Le nombre de chemins dans l'événement considéré est par le théorème de scrutin $\frac{b}{n}N_n(0, b)$. et chaque chemin a $\frac{1}{2}(n+b)$ pas vers la droite et $\frac{1}{2}(n-b)$ pas vers la gauche. Donc

$$\mathbb{P}(S_1 S_2 \dots S_n \neq 0, S_n = b) = \frac{b}{n} N_n(0, b) p^{\frac{1}{2}(n+b)} q^{\frac{1}{2}(n-b)} = \frac{b}{n} \mathbb{P}(S_n = b).$$

Un calcul du même type traite le cas $b < 0$.

6) Il suffit de sommer la relation du 5) sur toutes les valeurs de b .

7) Les formules du 5) et du 1) sont intéressantes pour le problème de la fortune du joueur : elle permettent de calculer la probabilité de fortune ($S_n = b$) sous la condition que le joueur ne se ruine pas ($S_0 \dots S_n \neq 0$).

Exercice 4.2. Interprétation du théorème de scrutin Ce théorème ("Ballot theorem") est parfois traduit improprement par "théorème de ballotage". En fait, il s'agit plutôt du théorème de scrutin et il concerne le dépouillement des votes. La question est de savoir : supposons que le candidat A a plus de votes que le candidat B : quelle est la probabilité que cette supériorité soit vérifiée tout au long du dépouillement ?

1) Supposons que le candidat A ait α votes et que le candidat B ait β votes, avec $\alpha > \beta$. Montrer en utilisant le théorème de scrutin que la probabilité que le candidat A ait tout au long du dépouillement un score supérieur à celui de B est égale à $\frac{\alpha-\beta}{\alpha+\beta}$.

2) Interpréter.

Solutions :

1) Le dépouillement se fait bulletin par bulletin. On pose pour $i = 1, \dots, \alpha + \beta$ $X_i = 1$ si le bulletin est pour A et -1 s'il est pour B . On suppose les X_i i.i.d. On a $S_0 = 0$ et S_n représente l'avantage de A sur B . Les chemins tels que S_k soit toujours différent de 0 pour $k \geq 1$ sont en nombre $N_n^0(0, b) = \frac{b}{n} N_n(0, b)$, et comme tous les chemins arrivant en $\alpha + \beta$ sont équiprobables, la probabilité recherchée est la proportion de chemins ne repassant pas par zéro, soit $\frac{b}{n} = \frac{\alpha-\beta}{\alpha+\beta}$.

2) Remarquons d'abord que le résultat ne dépend pas de la connaissance de p , la probabilité de voter A . Dans les démocraties, α est en général proche de β (1% de différence) et un ordre de grandeur de la probabilité précédente est donc $p = \frac{1}{100}$. Il est donc très probable que le dépouillement soit incertain. Toutefois, on peut se demander si cette incertitude ne se produit pas, en fait, au début du dépouillement, et s'il n'y a pas stabilisation du résultat bien avant la fin du dépouillement. Peut-on utiliser directement la loi des grands nombres, le théorème central limite ou des inégalités de grande déviation pour démontrer une stabilisation des proportions ? La réponse est "non", car nous avons une information a priori supplémentaire, à savoir $S_{\alpha+\beta} = \alpha - \beta$ et par contre nous ignorons la probabilité p de vote pour A .

Exercice 4.3. La méthode d'inversion du temps

On va montrer :

Théorème 4.1. La probabilité $f_b(n)$ qu'une marche aléatoire S_n partant de 0 arrive en $b > 0$ pour la première fois au temps n vérifie

$$f_b(n) = \frac{|b|}{n} \mathbb{P}(S_n = b), \quad n \geq 1.$$

Cette conclusion ressemble beaucoup à celle du théorème de scrutin et s'en déduit grâce à la méthode d'inversion du temps. On associe à la marche aléatoire

$$\{0, S_1, S_2, \dots, S_n\} = \{0, X_1, X_1 + X_2, \dots, \sum_1^n X_i\}$$

la marche aléatoire à rebours

$$\{0, T_1, T_2, \dots, T_n\} = \{0, X_n, X_n + X_{n-1}, \dots, \sum_1^n X_i\}.$$

- 1) Montrer que les deux marches ont la même loi (quelle que soit la valeur de p).
- 2) Montrer que la marche directe ne repasse pas par 0 et vérifie $S_n = b$ si et seulement si la marche inverse vérifie $T_n = b$ et que la première visite de T_n en b se produit au temps n (faire un dessin). En déduire le théorème.

Solutions :

Le fait que les marches aient même loi est immédiat, puisque les X_i sont i.i.d. On $S_n = b$ et $S_1 S_2 \dots S_n \neq 0$ si et seulement si $T_n = b$ et $T_n - T_{n-i} = X_1 + \dots + X_i > 0$ pour tout $i \geq 1$, ce qui veut bien dire que la première visite de T_n en b se fait au temps n .

5 Lois de l'arc sinus pour les marches aléatoires

Grimmett Stirzacker, 3.10, pages 80-82.

Pour comprendre vraiment ce qui est démontré dans cette section, il faut absolument que vous fassiez quelques simulations longues de marches aléatoires et que vous visualisiez la fonction $n \rightarrow S_n$. On observe notamment que la marche aléatoire a tendance à s'éloigner de zéro pour de longues périodes, et qu'il est en particulier assez probable qu'une marche partant de zéro ne revienne en zéro qu'après un temps assez long. Ceci ajoute une explication de plus au mystère de l'existence des casinos, qui peut se formuler ainsi : comment peut-on aller dans un endroit où l'on est statistiquement sûrs de perdre son argent ? La réponse est là encore dans la durée du plaisir obtenu à peu de frais : une mise minimale donne une chance importante de jouer longtemps.

Exercice 5.1. Soit une marche aléatoire $S_n = X_1 + \dots + X_n$ où les X_i sont i.i.d. et Bernouilli. On pose, pour indiquer le point de départ, $S_0 = 0$. On suppose que $\mathbb{P}(X_i = 1) = \mathbb{P}(X_i = -1) = \frac{1}{2}$. On va montrer le théorème suivant.

Théorème 5.1. *La probabilité que la dernière visite à 0 avant (et y compris) le temps $2n$ se produise au temps $2k$ est $\mathbb{P}(S_{2k} = 0)\mathbb{P}(S_{2n-2k} = 0)$.*

Tout d'abord, interprétons cette relation.

1) On appelle $\alpha_{2n}(2k)$ la probabilité précédente. On a donc $\alpha_{2n-2k} = u_{2k}u_{2n-2k}$, en posant $u_{2k} = \mathbb{P}(S_{2k} = 0)$. On rappelle que $\mathbb{P}(S_{2k} = 0) = \frac{C_{2k}^k}{2^{2k}}$. Montrer en utilisant la formule de Stirling $n! \approx n^n e^{-n} \sqrt{2\pi n}$ que

$$\alpha_{2n}(2k) \approx \frac{1}{\pi \sqrt{k(n-k)}} = \frac{1}{\pi \sqrt{k(n-k)}} \left(1 + \varepsilon\left(\frac{1}{k}, \frac{1}{n-k}\right)\right),$$

cette équivalence étant valide quand $n \rightarrow \infty$, $k \rightarrow \infty$, et $n - k \rightarrow \infty$ simultanément.

2) On note T_{2n} le temps de la dernière visite en 0 avant le temps $2n$. Montrer que

$$\mathbb{P}(\varepsilon n \leq T_{2n} \leq 2xn) \approx \sum_{\varepsilon n \leq k \leq xn} \frac{1}{\pi \sqrt{k(n-k)}} \approx \int_{\varepsilon n}^{xn} \frac{1}{\pi \sqrt{u(n-u)}} du = \frac{2}{\pi} \arcsin \sqrt{x} - \frac{2}{\pi} \arcsin \sqrt{\varepsilon}.$$

En déduire que

$$\liminf_{n \rightarrow +\infty} \mathbb{P}(T_{2n} \leq 2xn) \geq \frac{2}{\pi} \arcsin \sqrt{x}.$$

3) En déduire qu'il y a, quand n devient grand, plus d'une chance sur deux que la marche aléatoire ne repasse pas par 0 entre n et $2n$ et plus d'une chance sur 5 qu'elle ne repasse pas par 0 entre $\frac{n}{5}$ et $2n$.

4) Montrer la loi de l'arcsinus :

$$\lim_n \mathbb{P}(T_{2n} \leq 2xn) = \frac{2}{\pi} \arcsin \sqrt{x}.$$

Solutions :

1), 2) et 3) viennent tous seuls. Montrons juste le 4). On a

$$\lim_n \mathbb{P}(2\varepsilon n \leq T_{2n} \leq 2xn) = \frac{2}{\pi} (\arcsin \sqrt{x} - \arcsin \sqrt{\varepsilon}) \rightarrow 1,$$

quand $x \rightarrow 1$ et $\varepsilon \rightarrow 0$. Donc, pour tout ε' il existe η tel que pour $x \geq 1 - \eta$ et $\varepsilon \leq \eta$ on ait

$$\lim_n (\mathbb{P}(T_{2n} \leq 2\varepsilon n) + \mathbb{P}(T_{2n} \geq 2xn)) \leq \varepsilon' \text{ et donc}$$

$$\limsup_n \mathbb{P}(T_{2n} \leq 2\varepsilon n) \leq \varepsilon'.$$

On en déduit que

$$\begin{aligned} & \limsup_n |\mathbb{P}(T_{2n} \leq 2xn) - \frac{2}{\pi} \arcsin \sqrt{x}| \leq \\ & \limsup_n |\mathbb{P}(2\varepsilon n \leq T_{2n} \leq 2xn) - \frac{2}{\pi} \arcsin \sqrt{x} + \frac{2}{\pi} \arcsin \sqrt{\varepsilon} - \frac{2}{\pi} \arcsin \sqrt{\varepsilon} + \mathbb{P}(T_{2n} \leq 2\varepsilon n)| \\ & \leq \limsup_n |\mathbb{P}(2\varepsilon n \leq T_{2n} \leq 2xn) - \frac{2}{\pi} \arcsin \sqrt{x} + \frac{2}{\pi} \arcsin \sqrt{\varepsilon}| + \frac{2}{\pi} \arcsin \sqrt{\varepsilon} + \varepsilon' \\ & = \frac{2}{\pi} \arcsin \sqrt{\varepsilon} + \varepsilon'. \end{aligned}$$

Cette dernière quantité étant arbitrairement petite, on obtient le résultat.

Exercice 5.2. Preuve de la loi en arc sinus pour les retours à l'origine.

1) Montrer que $\alpha_{2n}(2k) = \mathbb{P}(S_{2k} = 0) \mathbb{P}(S_1 S_2 \dots S_{2n-2k} \neq 0)$. On pose dans la suite $m = 2n - 2k$.

2) On rappelle que

$$\mathbb{P}(S_1 S_2 \dots S_n \neq 0, S_n = b) = \frac{|b|}{n} \mathbb{P}(S_n = b)$$

et que $\mathbb{P}(S_{2m} = 2j) = C_{2m}^{m+j} \left(\frac{1}{2}\right)^{2m}$. En déduire que

$$\mathbb{P}(S_1 S_2 \dots S_{2m} \neq 0) = \mathbb{P}(S_{2m} = 0).$$

Indication : vérifier et utiliser $\frac{2k}{2m} C_{2m}^{m+k} = C_{2m-1}^{m+k-1} - C_{2m-1}^{m+k}$.

Solutions :

- 1) $\alpha_{2n}(2k) = \mathbb{P}(S_{2k} = 0) \mathbb{P}(S_{2k+1} S_{2k+2} \dots S_{2n} \neq 0 | S_{2k} = 0) = \mathbb{P}(S_{2k} = 0) \mathbb{P}(S_1 S_2 \dots S_{2n-2k} \neq 0)$.
- 2) Par la loi des probabilités disjointes,

$$\begin{aligned} \mathbb{P}(S_1 S_2 \dots S_{2m} \neq 0) &= 2 \sum_{k=1}^m \frac{2k}{2m} \mathbb{P}(S_{2m} = 2k) \\ &= 2 \sum_{k=1}^m \frac{2k}{2m} C_{2m}^{m+k} \left(\frac{1}{2}\right)^{2m} = 2 \left(\frac{1}{2}\right)^{2m} \sum_{k=1}^m \left[C_{2m-1}^{m+k-1} - C_{2m-1}^{m+k} \right] \\ &= 2 \left(\frac{1}{2}\right)^{2m} C_{2m-1}^m = C_{2m}^m \left(\frac{1}{2}\right)^{2m} = \mathbb{P}(S_{2m} = 0). \end{aligned}$$

Exercice 5.3. Loi de l'arc sinus pour les périodes de gain On dit que la marche aléatoire passe l'intervalle $(k, k+1)$ à droite de l'origine si $S_k > 0$ ou $S_{k+1} > 0$, ou les deux. On va montrer le théorème suivant :

Théorème 5.2. *La probabilité qu'une marche aléatoire symétrique partant de 0 passe exactement $2k$ intervalles de temps à droite de l'origine jusqu'au temps $2n$ est égale à $\mathbb{P}(S_{2k} = 0) \mathbb{P}(S_{2n-2k} = 0)$.*

- 1) Interpréter le résultat dans les mêmes lignes que la loi arc sinus pour les retours à l'origine.
- 2) Lire la solution page 81-82 du Grimmett. Attention, cette solution donne les idées mais a quelques petites erreurs à corriger.

Par les mêmes équivalents que pour la loi de l'arc sinus pour les retours à l'origine, on obtient que la probabilité que la marche passe $2xn$ unités de temps ou moins à droite de l'origine est de l'ordre de $\frac{2}{\pi} \arcsin \sqrt{x}$. Si on prend par exemple $x = \frac{1}{5}$, cette probabilité est de l'ordre de $\frac{1}{5}$. Du point de vue du jeu de pile ou face, il y a donc une chance sur cinq pour que l'un des joueurs soit gagnant pendant plus des quatre cinquièmes du temps.