Composition de Mathématiques

Le 3 février 2016 – De 13 heures à 17 heures

Si, au cours de l'épreuve, un candidat repère ce qui lui semble être une erreur d'énoncé, il le signale sur sa copie et poursuit sa composition en expliquant les raisons des initiatives qu'il est amené à prendre.

Les calculatrices et téléphones portables sont interdits.

❖ I – Problème ❖

1. Déterminer une condition nécessaire et suffisante pour que la matrice

$$A = \begin{pmatrix} 0 & 1 \\ y-4 & 2x \end{pmatrix}$$

soit diagonalisable dans $\mathfrak{M}_2(\mathbb{R})$.

2. On pose $E_1 = \{u \in \mathbb{R}_+ \ : \ u^2 \notin \mathbb{N}\}$ et $E_2 = \mathbb{R}_+ \setminus E_1$. Démontrer que l'ensemble E_2 est dénombrable.

3. Soit $f : \mathbb{R}_+ \to \mathbb{R}$, l'application définie par

$$\forall\, u \in E_1, \quad f(u) = 0 \quad \text{et} \quad \forall\, u \in E_2, \quad f(u) = \frac{\lambda}{2^{u^2}}.$$

Déterminer le réel λ pour lequel il existe une variable aléatoire discrète X définie sur un espace probabilisé $(\Omega, \mathcal{A}, \mathbf{P})$ telle que

$$\forall\, u \in \mathbb{R}_+, \quad \mathbf{P}(X = u) = f(u).$$

4. Déterminer la loi et l'espérance de X^2.

5. Déterminer la fonction génératrice de X^2 et retrouver ainsi l'espérance de X^2.

6. Soit Y, une variable aléatoire définie sur Ω, indépendante de X, telle que

$$\forall\, u \in \mathbb{R}_+, \quad \mathbf{P}(Y = u) = \left| \begin{array}{ll} 0 & \text{si } u \notin \mathbb{N}, \\ \dfrac{1}{2^{u+1}} & \text{sinon.} \end{array} \right.$$

Déterminer la fonction génératrice de la variable aléatoire

$$Z = X^2 + Y.$$

En déduire la loi de Z.

7. Calculer la probabilité pour que la matrice aléatoire

$$A = \begin{pmatrix} 0 & 1 \\ Y-4 & 2X \end{pmatrix}$$

soit diagonalisable.

❖ II – Problème ❖

Partie A. Préliminaires

1. Pour tout entier $n \in \mathbb{N}$, on pose

$$u_n = \binom{2n}{n}.$$

1. a. Vérifier que $u_n \leqslant 4^n$ pour tout $n \in \mathbb{N}$.

1. b. Déduire de la formule de Stirling un réel $C > 0$ tel que

$$u_n \sim C\frac{4^n}{\sqrt{n}}$$

lorsque n tend vers $+\infty$.

2. On *admet* que

$$\forall\, 0 \leqslant x < 1/4, \quad \frac{1}{\sqrt{1-4x}} = \sum_{n=0}^{+\infty} u_n x^n.$$

On pose en outre

$$\forall\, n \in \mathbb{N}, \quad v_n = \sum_{k=0}^{n} u_k u_{n-k}.$$

Calculer la somme

$$\sum_{n=0}^{+\infty} v_n x^n$$

pour $0 \leqslant x < 1/4$. En déduire la valeur de v_n en fonction de n.

Partie B. Une marche aléatoire

On considère une suite $(X_n)_{n\geqslant 1}$ de variables aléatoires définies sur $(\Omega, \mathcal{A}, \mathbf{P})$, indépendantes et de même loi, avec

$$\mathbf{P}(X_i = 1) = \mathbf{P}(X_i = -1) = 1/2.$$

On pose alors $S_0 = 0$ et

$$\forall\, n \in \mathbb{N}^*, \quad S_n = \sum_{i=1}^{n} X_i.$$

La suite $(S_n)_{n\in\mathbb{N}}$ est une **marche aléatoire** : elle modélise la succession des positions occupées par un mobile qui se déplace au hasard sur $\mathbb{Z}$.

3. Calculer l'espérance et la variance de S_1.

4. Caractériser la loi de S_2 et calculer son espérance.

5. Démontrer que

$$\forall\, n \geqslant 1, \quad \mathbf{P}(S_{2n} = 0) = \frac{u_n}{4^n}.$$

☞ *On pourra relier la loi de S_{2n} à une loi binomiale bien choisie.*

6. Pour tout entier $i \geqslant 1$, on définit la variable aléatoire

$$Z_i = \mathbb{1}_{[S_{2i}=0]}$$

et, pour tout entier $n \geqslant 1$, on définit

$$Y_n = \sum_{i=1}^{n} Z_i.$$

6. a. Que représente la variable Y_n pour la marche aléatoire $(S_n)_{n\in\mathbb{N}}$?
6. b. Démontrer que

$$\mathbf{E}(Y_n) \sim \sum_{k=1}^{n} \frac{1}{\sqrt{k\pi}}$$

lorsque n tend vers $+\infty$. En comparant la somme à une intégrale, donner un équivalent simple de $\mathbf{E}(Y_n)$ lorsque n tend vers $+\infty$.

Partie C. Une autre marche aléatoire

On fixe un entier naturel $N \geqslant 1$. On considère une variable aléatoire A définie sur $(\Omega, \mathcal{A}, \mathbf{P})$ qui suit la loi uniforme sur l'ensemble E_N des parties de $E = \{1, \dots, 2N\}$ dont le cardinal est égal à N :

– Pour tout $\omega \in \Omega$, il existe des entiers

$$1 \leqslant i_1 < i_2 < \cdots < i_N \leqslant 2N$$

tels que $A(\omega) = \{i_1, i_2, \dots, i_N\} \in E_N$.

– Pour toute partie $\{i_1, i_2, \dots, i_N\} \in E_N$,

$$\mathbf{P}\big(A = \{i_1, \dots, i_N\}\big) = \frac{1}{\#(E_N)}.$$

7. Quel est le cardinal de E_N ?

Pour tout entier $1 \leqslant i \leqslant 2N$ et tout $\omega \in \Omega$, on pose alors

$$X'_i(\omega) = \left| \begin{array}{ll} 1 & \text{si } i \in A(\omega), \\ -1 & \text{si } i \notin A(\omega). \end{array} \right.$$

On pose ensuite $S'_0 = 0$ et

$$\forall\, 1 \leqslant n \leqslant 2N, \quad S'_n = \sum_{k=1}^{n} X'_k.$$

Ainsi, si $N = 3$ et $A(\omega) = \{1, 2, 5\}$, alors

$$X'_1(\omega) = X'_2(\omega) = 1,\ X'_3(\omega) = X'_4(\omega) = -1,$$
$$X'_5(\omega) = 1, X'_6(\omega) = -1$$

et par conséquent

$$S'_1(\omega) = 1,\ S'_2(\omega) = 2,\ S'_3(\omega) = 1,\ S'_4(\omega) = 0,$$
$$S'_5(\omega) = 1,\ S'_6(\omega) = 0.$$

8. Démontrer que les fonctions $X'_i : \Omega \to \mathbb{Z}$ sont des variables aléatoires sur $(\Omega, \mathcal{A})$.
9. Démontrer que les fonctions $S'_n : \Omega \to \mathbb{Z}$ sont des variables aléatoires sur $(\Omega, \mathcal{A})$.
10. Démontrer que les variables X'_i suivent toutes la loi uniforme sur $\{-1; 1\}$:

$$\forall\, 1 \leqslant i \leqslant 2N, \quad \mathbf{P}(X'_i = 1) = \mathbf{P}(X'_i = -1) = 1/2.$$

11. Que dire de la variable aléatoire S'_{2N} ? Quelle est sa variance ?
12. Les variables aléatoires X'_i sont-elles indépendantes ?
13. On pose alors

$$Y'_N = \sum_{i=1}^{N} \mathbb{1}_{[S'_{2i}=0]}.$$

13. a. Démontrer que la fonction $Y'_N : \Omega \to \mathbb{N}$ est une variable aléatoire sur $(\Omega, \mathcal{A})$.
13. b. Démontrer que

$$\mathbf{E}(Y'_N) = \frac{v_N}{u_N} - 1$$

et en déduire un équivalent de $\mathbf{E}(Y'_N)$ lorsque N tend vers l'infini.

Partie D. Modélisation

14. Proposer un modèle d'urne pour chacune des deux marches aléatoires : on considérera une urne contenant des boules de deux couleurs différentes (par exemple, céladon et rose thé) et on associera les variables aléatoires X_i et X'_i à des protocoles de tirage précis. On interprétera les variables aléatoires Y_N et Y'_N en fonction de ces modèles.

❖ III – Problème ❖

Contexte

Un horticulteur doit greffer un nombre $R \geqslant 1$ de rosiers.

Une semaine après chaque greffe, on peut savoir si la greffe a pris. Si la greffe n'a pas pris, on recommence l'opération jusqu'à ce qu'elle prenne effectivement.

Notations

Pour tout entier $1 \leqslant k \leqslant R$, on note X_k, le nombre de tentatives nécessaires pour que la greffe du rosier k prenne. On note X, le nombre de tentatives pour que toutes les greffes prennent.

On note S, le nombre de semaines nécessaires pour que toutes les greffes prennent.

On pose $Y_0 = 0$ et, pour tout $n \in \mathbb{N}$, on note Y_n, le nombre de rosiers dont la greffe a pris au cours des n premières semaines et on pose $Z_n = Y_{n+1} - Y_n$.

Rappel théorique

Pour toute loi de probabilité $(p_n)_{n\in\mathbb{N}}$ sur $\mathbb{N}$ et pour tout ensemble dénombrable I, il existe un espace probabilisé $(\Omega, \mathcal{A}, \mathbf{P})$ et une famille $(U_i)_{i\in I}$ de variables aléatoires définies sur $(\Omega, \mathcal{A})$, qui sont indépendantes et telles que

$$\forall\, i \in I,\ \forall\, n \in \mathbb{N}, \quad \mathbf{P}(X_i = n) = p_n.$$

Travail à effectuer

Proposer un modèle probabiliste simple (c'est-à-dire une famille de variables aléatoires $(U_i)_{i\in I}$ dont la loi sera décrite avec précision) pour lequel X, S, les X_k, les Y_n et les Z_n sont des variables aléatoires.

Pour chacune de ces variables aléatoires, on énoncera le théorème qui permet de déduire sa loi du modèle ou on indiquera précisément (mais sans entrer dans le détail des calculs) une méthode permettant de calculer sa loi.

Solution I ❀ Une matrice aléatoire

1. Le polynôme caractéristique de A est égal à

$$X^2 - (\operatorname{tr} A)X + (\det A) = X^2 - 2xX + (4 - y).$$

Son discriminant réduit est égal à $x^2 + y - 4$.

❧ Si le discriminant est strictement négatif, la matrice A n'a pas de valeur propre réelle, donc n'est pas diagonalisable dans $\mathfrak{M}_2(\mathbb{R})$.

❧ Si le discriminant est nul, alors la matrice A admet une valeur propre unique et comme ce n'est pas une homothétie, alors elle n'est pas diagonalisable dans $\mathfrak{M}_2(\mathbb{R})$.

❧ Si le discriminant est strictement positif, alors la matrice A admet deux valeurs propres réelles distinctes et comme elle appartient à $\mathfrak{M}_2(\mathbb{R})$, elle est donc diagonalisable.

Par conséquent, la matrice A est diagonalisable si, et seulement si,

$$x^2 + y > 4.$$

2. L'ensemble E_2 est par définition l'image de $\mathbb{N}$, partie dénombrable de $\mathbb{R}_+$, par l'application injective $\left[x \mapsto \sqrt{x}\right]$, donc E_2 est bien une partie dénombrable de $\mathbb{R}_+$.

$$E_2 = \left\{\sqrt{n},\ n \in \mathbb{N}\right\}$$

3. On sait qu'il existe un espace probabilisé $(\Omega, \mathcal{A}, \mathbf{P})$ et une variable aléatoire discrète X définie sur $(\Omega, \mathcal{A}, \mathbf{P})$ telle que

$$\forall\, u \in \mathbb{R}_+, \quad \mathbf{P}(X = u) = f(u)$$

si, et seulement si, la famille $\big(f(u)\big)_{u \in \mathbb{R}_+}$ est une famille sommable de réels positifs dont la somme est égale à 1.

D'après **2.**, il s'agit donc de trouver les valeurs de $\lambda \in \mathbb{R}$ pour lesquelles la série $\sum \lambda/2^n$ est une série de terme général positif dont la somme est égale à 1.

Comme $0 < 1/2 < 1$, on reconnaît une série géométrique convergente dont le terme général est positif si, et seulement si, $\lambda \in \mathbb{R}_+$ et dont la somme est égale à

$$\sum_{n=0}^{+\infty} \frac{\lambda}{2^n} = \frac{\lambda}{1 - 1/2} = 2\lambda.$$

Donc le seul réel λ convenable est égal à $1/2$.

4. Comme l'ensemble des valeurs prises par X est E_2, alors l'ensemble des valeurs prises par X^2 est égal à $\mathbb{N}$ et comme $E_2 \subset \mathbb{R}_+$,

$$\forall\, n \in \mathbb{N}, \quad [X^2 = n] = [X = \sqrt{n}]$$

donc

$$\forall\, n \in \mathbb{N}, \quad \mathbf{P}(X^2 = n) = \Big(\frac{1}{2}\Big)^{n+1}.$$

(C'est presque une loi géométrique !)

❧ La variable X^2 est d'espérance finie si, et seulement si, la série $\sum 2^{-(n+1)}n$ est convergente. On sait que le rayon de convergence de la série entière $\sum x^n$ est égal à 1 et comme il est strictement positif, on peut dériver terme à terme sur l'intervalle ouvert de convergence, donc

$$\forall\, x \in \left]-1, 1\right[, \quad \sum_{n=1}^{+\infty} n x^{n-1} = \frac{d}{dx}\Big(\frac{1}{1-x}\Big) = \frac{1}{(1-x)^2}.$$

En particulier, pour $x = 1/2$, la série $\sum 2^{-(n+1)}n$ est convergente et

$$\sum_{n=0}^{+\infty} n(1/2)^{n+1} = (1/2)^2 \sum_{n=1}^{+\infty} n(1/2)^{n-1} = \frac{1/4}{1/4} = 1$$

(le terme en $n = 0$ est nul), donc $\mathbf{E}(X^2) = 1$.

5. Comme X^2 est une variable aléatoire à valeurs dans $\mathbb{N}$, sa fonction génératrice F est par définition la somme de la série entière $\sum \mathbf{P}(X^2 = n)x^n$, série entière dont le rayon de convergence est au moins égal à 1. Donc

$$\begin{aligned}\forall\, x \in [-1, 1], \quad F(x) &= \sum_{n=0}^{+\infty} \frac{1}{2^{n+1}} x^n = \frac{1}{2} \sum_{n=0}^{+\infty} (x/2)^n \\ &= \frac{1}{2} \cdot \frac{1}{1 - x/2} = \frac{1}{2 - x}.\end{aligned}$$

On remarque en passant que le rayon de convergence de la série génératrice de X^2 est égal à 2, ce qui prouve que la fonction génératrice est de classe $\mathscr{C}^\infty$ sur l'intervalle ouvert $\left]-2, 2\right[$ et, en particulier, dérivable en $x = 1$. Cela prouve à nouveau que X^2 est une variable aléatoire d'espérance finie et que

$$\mathbf{E}(X^2) = F'(1) = \frac{1}{(2-1)^2} = 1.$$

6. En tant que somme de deux variables aléatoires à valeurs dans $\mathbb{N}$, la fonction Z est encore une variable aléatoire sur $(\Omega, \mathcal{A}, \mathbf{P})$ à valeurs dans $\mathbb{N}$, donc sa fonction génératrice G est bien définie sur $[-1, 1]$ au moins.

La variable aléatoire Y suit la même loi que X^2, donc X^2 et Y ont même fonction génératrice. Comme X^2 et Y sont indépendantes, alors les variables aléatoires t^{X^2} et t^Y sont indépendantes et comme ces variables sont d'espérance finie pour $|t| < 2$, alors $t^Z = (t^{X^2})(t^Y)$ est aussi une variable aléatoire d'espérance finie pour $t \in \left]-2, 2\right[$ et

$$\begin{aligned}\forall\, t \in \left]-2, 2\right[, \quad G(t) &= \mathbf{E}(t^Z) = \mathbf{E}(t^{X^2})\,\mathbf{E}(t^Y) = [F(t)]^2 \\ &= \frac{1}{(2-t)^2} = F'(t).\end{aligned}$$

Comme le rayon de convergence de la série génératrice de X^2 est strictement positif, on peut dériver terme à terme. Sachant que

$$\forall\, t \in \left]-2, 2\right[, \quad F(t) = \sum_{n=0}^{+\infty} \frac{t^n}{2^{n+1}}$$

on en déduit que

$$\forall\, t \in \left]-2, 2\right[, \quad G(t) = \sum_{n=1}^{+\infty} \frac{n t^{n-1}}{2^{n+1}} = \sum_{n=0}^{+\infty} \frac{n+1}{2^{n+2}} \cdot t^n$$

et par unicité du développement en série entière (le rayon de convergence étant strictement positif), on en déduit que

$$\forall\, n \in \mathbb{N}, \quad \mathbf{P}(Z = n) = \frac{n+1}{2^{n+2}}.$$

7. D'après **1.**, la matrice A est diagonalisable si, et seulement si, la variable aléatoire $Z = X^2 + Y$ est strictement supérieure à 4. On cherche donc

$$\mathbf{P}(Z > 4) = 1 - \mathbf{P}(Z \leqslant 4) = 1 - \sum_{k=0}^{4} \mathbf{P}(Z = k).$$

D'après la question précédente,

$$\mathbf{P}(Z=0) = \frac{1}{4} = \frac{16}{64} \qquad \mathbf{P}(Z=1) = \frac{2}{8} = \frac{16}{64}$$
$$\mathbf{P}(Z=2) = \frac{3}{16} = \frac{12}{64} \qquad \mathbf{P}(Z=3) = \frac{4}{32} = \frac{8}{64}$$
$$\mathbf{P}(Z=4) = \frac{5}{64}$$

donc la probabilité pour que la matrice A soit inversible est égale à

$$1 - \frac{16+16+12+8+5}{64} = \frac{7}{64}.$$

Solution II ❀ Marches aléatoires

Partie A. Préliminaires

1. a. D'après la formule du binôme,

$$4^n = (1+1)^{2n} = \sum_{k=0}^{2n} \binom{2n}{k} 1^k 1^{2n-k} \geqslant \binom{2n}{n}$$

puisque tous les termes sont positifs.

1. b. La formule de Stirling nous dit que

$$(u_n)! \sim \sqrt{2\pi u_n} e^{-u_n} u_n^{u_n}$$

pour toute suite d'entiers $(u_n)_{n\in\mathbb{N}}$ qui tend vers $+\infty$. On en déduit que

$$\binom{2n}{n} = \frac{(2n)!}{n!n!} \sim \frac{4^n}{\sqrt{\pi n}}.$$

2. La propriété admise par l'énoncé prouve que le rayon de convergence de la série entière $\sum u_n x^n$ est supérieur à $1/4$. (Cela pourrait également être déduit de l'équivalent établi en **1.b.**)

En particulier, pour tout $0 \leqslant x < 1/4$, la série $\sum u_n x^n$ est *absolument* convergente. La série $\sum v_n x^n$, produit de Cauchy de $\sum u_n x^n$ par elle-même, est donc elle aussi absolument convergente et de plus

$$\sum_{n=0}^{+\infty} v_n x^n = \left(\sum_{n=0}^{+\infty} u_n x^n\right)^2 = \frac{1}{1-4x}.$$

☙ On en déduit (série *géométrique*) que

$$\forall\, 0 \leqslant x < 1/4, \quad \sum_{n=0}^{+\infty} v_n x^n = \frac{1}{1-4x} = \sum_{n=0}^{+\infty} 4^n x^n$$

et donc que

$$\forall\, n \in \mathbb{N}, \quad v_n = 4^n$$

(unicité du développement en série entière).

Partie B. Une marche aléatoire

3. La variable aléatoire $S_1 = X_1$ est bornée, donc elle admet des moments de tout ordre : en particulier, elle est d'espérance finie et admet une variance.

Par définition de l'espérance,

$$\mathbf{E}(S_1) = (-1)\,\mathbf{P}(S_1 = -1) + (+1)\,\mathbf{P}(S = 1) = 0.$$

La variable S_1 est donc *centrée*.

De plus, la variable S_1^2 est constante, égale à 1, donc $\mathbf{E}(S_1^2) = 1$. D'après la formule de Koenig-Huyghens,

$$\mathbf{V}(S_1) = \mathbf{E}(S_1^2) - [\mathbf{E}(S_1)]^2 = 1.$$

La variable S_1 est donc *réduite*.

4. On connaît les lois de X_1 et X_2 et on sait que ces variables aléatoires sont indépendantes. On connaît donc la loi du couple (X_1, X_2) et on peut en déduire la loi de la somme $X_1 + X_2$.

x_1	x_2	$\mathbf{P}[(X_1,X_2) = (x_1,x_2)]$	S_2
-1	-1	$1/4$	-2
-1	1	$1/4$	0
1	-1	$1/4$	0
1	1	$1/4$	2

Par conséquent, la loi de S_2 est portée par $\{-2; 0; 2\}$ et

$$\mathbf{P}(S_2 = -2) = \mathbf{P}(S_2 = 2) = \frac{1}{4}, \ \mathbf{P}(S_2 = 0) = \frac{1}{2}.$$

5. On déduit de la loi de X_i que la variable aléatoire

$$B_i = \frac{X_i + 1}{2}$$

suit la loi de Bernoulli de paramètre $1/2$. Comme les variables X_1, …, X_{2n} sont indépendantes par hypothèse, alors les variables $B_1 = f(X_1)$, …, $B_{2n} = f(X_{2n})$ sont indépendantes elles aussi et, en tant que somme de $2n$ variables de Bernoulli indépendantes, la variable

$$\Sigma_{2n} = \frac{S_{2n} + 2n}{2} = \sum_{i=1}^{2n} B_i$$

suit la loi binomiale $\mathcal{B}(2n, 1/2)$:

$$\forall\, 0 \leqslant k \leqslant 2n, \quad \mathbf{P}(\Sigma_{2n} = k) = \binom{2n}{k} (1/2)^k (1/2)^{2n-k} = \frac{1}{4^n}\binom{2n}{k}.$$

Comme $[S_{2n} = 0] = [\Sigma_n = n]$, on en déduit que

$$\mathbf{P}(S_{2n} = 0) = \frac{1}{4^n} u_n. \tag{1}$$

6. a. La marche aléatoire part de 0 puisque $S_0 = 0$. La variable Z_i prend la valeur 1 lorsque $S_{2i} = 0$, c'est-à-dire lorsque le mobile passe à nouveau par 0. La variable Y_n est donc le nombre de passages en 0 entre les instants 1 et n, c'est-à-dire le *nombre de retours à l'origine* qui ont lieu avant l'instant n.

6. b. La variable Y_n est une somme de n variables aléatoires de Bernoulli. Elle est donc presque sûrement bornée et donc d'espérance finie.

Par linéarité de l'espérance,

$$\mathbf{E}(Y_n) = \sum_{i=1}^{n} \mathbf{E}(Z_i) = \sum_{k=1}^{n} \mathbf{P}(S_{2k} = 0).$$

D'après (1),

$$\mathbf{E}(Y_n) = \sum_{k=1}^{n} \frac{u_k}{4^k}$$

et d'après **1.b.**,

$$\frac{u_k}{4^k} \sim \frac{1}{\sqrt{\pi k}}.$$

La série de terme général *positif* $1/\sqrt{\pi k}$ est *divergente* (Riemann), donc les sommes partielles $\mathbf{E}(Y_n)$ tendent vers $+\infty$ et plus précisément

$$\mathbf{E}(Y_n) \sim \sum_{k=1}^{n} \frac{1}{\sqrt{\pi k}}$$

lorsque n tend vers $+\infty$ (théorème de sommation des équivalents).

❧ La fonction

$$f = \left[t \mapsto \frac{1}{\sqrt{\pi t}}\right]$$

est continue et strictement décroissante sur $]0, +\infty[$. On en déduit *(figure correctement légendée obligatoire sur la copie)* que

$$\int_1^n f(t)\,dt + f(n) \leqslant \sum_{k=1}^{n} f(k) \leqslant f(1) + \int_1^n f(t)\,dt$$

c'est-à-dire

$$2\frac{\sqrt{n}-1}{\sqrt{\pi}} + \frac{1}{\sqrt{\pi n}} \leqslant \sum_{k=1}^{n} f(k) \leqslant \frac{1}{\sqrt{\pi}} + 2\frac{\sqrt{n}-1}{\sqrt{\pi}}$$

pour tout $n \in \mathbb{N}$ et donc que

$$\mathbf{E}(Y_n) \sim 2\sqrt{\frac{n}{\pi}}$$

lorsque n tend vers $+\infty$.

Partie C. Une autre marche aléatoire

7. Il y a $\binom{2N}{N}$ parties de E constituées de N éléments (par définition du coefficient binomial).

8. Comme $A : \Omega \to E_N$ est une variable aléatoire discrète sur l'espace probabilisable $(\Omega, \mathcal{A})$, alors

$$\forall\, B \subset E_N, \quad [A \in B] \in \mathcal{A}. \tag{2}$$

Pour $1 \leqslant i \leqslant 2N$, notons E_N^i, l'ensemble des parties de E constituées de N éléments et qui contiennent i :

$$B \in E_N^i \iff \begin{cases} B \in E_N \\ i \in B \end{cases}.$$

Par définition,

$$[X'_i = 1] = [A \in E_N^i] \quad \text{et} \quad [X'_i = -1] = [A \in E_N^i]^c \tag{3}$$

donc $[X'_i = 1] \in \mathcal{A}$ et $[X'_i = -1] \in \mathcal{A}$ par (2).

D'autre part, $[X'_i = k] = \varnothing \in \mathcal{A}$ pour tout $k \in \mathbb{Z}$ distinct de ± 1.

Finalement,

$$\forall\, k \in \mathbb{Z}, \quad [X'_i = k] \in \mathcal{A},$$

donc $X'_i : \Omega \to \mathbb{Z}$ est bien un variable aléatoire discrète sur $(\Omega, \mathcal{A})$.

9. En tant que somme d'un nombre fini de variables aléatoires discrètes sur $(\Omega, \mathcal{A})$, S'_n est une variable aléatoire discrète sur $(\Omega, \mathcal{A})$.

10. On a déjà justifié que la loi de X_i était portée par l'ensemble $\{\pm 1\}$ et que, par conséquent,

$$\mathbf{P}(X'_i = -1) = 1 - \mathbf{P}(X'_i = 1).$$

❧ Comme la loi de A est la loi uniforme sur E_N, on déduit de (3) que

$$\mathbf{P}(X'_i = 1) = \frac{\#(E_N^i)}{\#(E_N)}.$$

L'application $[B \mapsto B \cup \{i\}]$ réalise une bijection de l'ensemble des parties de cardinal $(N-1)$ prises dans l'ensemble

$$\{1, \dots, i-1, i+1, \dots, 2N\}$$

sur l'ensemble E_N^i, donc

$$\#(E_N^i) = \binom{2N-1}{N-1}$$

et par conséquent,

$$\mathbf{P}(X'_i = 1) = \frac{(2N-1)!/[(N-1)!N!]}{(2N)!/[N!N!]} = \frac{N}{2N} = \frac{1}{2},$$

ce qui prouve que X'_i suit la loi uniforme sur $\{\pm 1\}$.

11. Par hypothèse, $\#[A(\omega)] = N$.

Il y a donc N indices $1 \leqslant i \leqslant 2N$ qui appartiennent à $A(\omega)$ et N autres indices $1 \leqslant i \leqslant 2N$ qui n'appartiennent pas à $A(\omega)$.

Il y a donc N indices $1 \leqslant i \leqslant 2N$ tels que $X'_i(\omega) = 1$ et N indices $1 \leqslant i \leqslant 2N$ tels que $X'_i(\omega) = -1$. Par conséquent,

$$\forall\, \omega \in \Omega, \quad S'_{2N}(\omega) = N - N = 0.$$

La variable aléatoire S'_{2N} est donc identiquement nulle (loi de Dirac) et, comme pour toute variable aléatoire presque sûrement constante, sa variance est nulle.

12. Les variables aléatoires X'_i sont toutes de même loi et leur variance commune est égale à 1 (calculée au **3.**). Si ces variables aléatoires étaient indépendantes, on aurait alors

$$\mathbf{V}(S'_{2N}) = \mathbf{V}\left(\sum_{i=1}^{2N} X'_i\right) = \sum_{i=1}^{2N} \mathbf{V}(X'_i) = 2N.$$

On a démontré à la question précédente que $\mathbf{V}(S'_{2N}) = 0$, ce qui prouve que les variables X'_1, ..., X'_{2N} ne sont pas indépendantes.

13. a. Comme S'_{2i} est une variable aléatoire sur $(\Omega, \mathcal{A})$, alors $[S'_{2i} = 0]$ est un événement : $[S'_{2i} = 0] \in \mathcal{A}$. Par conséquent, son indicatrice $\mathbb{1}_{[S'_{2i}=0]}$ est une variable aléatoire de Bernoulli sur $(\Omega, \mathcal{A})$ et Y'_N, en tant que somme d'un nombre fini de variables aléatoires, est encore une variable aléatoire sur $(\Omega, \mathcal{A})$.

13. b. Par linéarité de l'espérance,

$$\mathbf{E}(Y'_N) = \sum_{i=1}^{N} \mathbf{E}(\mathbb{1}_{[S'_{2i}=0]}) = \sum_{i=1}^{N} \mathbf{P}(S'_{2i} = 0).$$

La variable S'_{2i} s'annule lorsque i variables parmi $X_1, \dots, X_{2i}$ prennent la valeur 1 (et les autres prennent la valeur -1). Autrement dit,

$$[S'_{2i} = 0] = \bigsqcup_{1 \leqslant k_1 < k_2 < \dots < k_i \leqslant 2i} \big[A \cap \{1, \dots, 2i\} = \{k_1, \dots, k_i\}\big].$$

Comme la loi de A est uniforme, il reste à compter le nombre de cas favorables. La condition

$$A(\omega) \cap \{1, \dots, 2i\} = \{k_1, \dots, k_i\}$$

signifie que les entiers $k \leqslant 2i$ qui appartiennent à $A(\omega)$ sont connus : il reste donc à choisir les $(N-i)$ entiers $2i < k \leqslant 2N$ qui appartiennent à $A(\omega)$. Par définition du coefficient binomial, il y a $\binom{2N-2i}{N-i}$ possibilités et par σ-additivité de $\mathbf{P}$,

$$\mathbf{P}(S'_{2i} = 0) = \sum_{1 \leqslant k_1 < k_2 < \dots < k_i \leqslant 2i} \frac{\binom{2N-2i}{N-i}}{u_N}.$$

Il y a (toujours par définition du coefficient binomial) $\binom{2i}{i}$ termes dans cette somme, donc

$$\mathbf{P}(S'_{2i} = 0) = \frac{\binom{2i}{i}\binom{2N-2i}{N-i}}{u_N}$$

et par définition de v_N,

$$\mathbf{E}(Y'_N) = \frac{v_N}{u_N} - 1.$$

(Le terme en $i = 0$ manque à l'appel !)

❧ On déduit alors de **1.b.** et **2.** que

$$\mathbf{E}(Y'_N) \sim \sqrt{\pi N}$$

lorsque N tend vers $+\infty$.

Partie D. Modélisation

14. On considère une urne contenant N boules céladon et N boules rose thé.

Dans le premier cas, on tire les boules *avec* remise et compte $+1$ pour chaque boule rose et -1 pour chaque boule verte. Il est raisonnable dans ces conditions de supposer que chaque couleur est tirée avec une probabilité $1/2$ et que les tirages sont indépendants.

Dans le second cas, on tire N boules *sans* remise parmi les 2N boules présentes dans l'urne. Il est raisonnable de supposer que chaque ensemble de N tirages a la même probabilité d'être réalisé (loi uniforme sur l'ensemble E_N des tirages possibles). Les variables X'_i qui indiquent la couleur chaque boule tirée ne sont plus indépendantes **[12.]** car la composition de l'urne varie avec chaque boule tirée. Ce qui est plus surprenant est que chaque couleur a la même probabilité d'apparaître à chaque tirage, quelle que soit l'évolution de la composition de l'urne **[10.]**.

Les variables Y_N et Y'_N comptent le nombre de fois où, au cours de la succession de N tirages, on a retiré autant de boules roses que de boules vertes.

Solution III ❁ Modèle de greffe de rosiers

L'ensemble $I = \{1, \dots, R\} \times \mathbb{N}^*$ est dénombrable (produit cartésien d'un ensemble fini par un ensemble dénombrable). D'après le théorème de Kolmogorov, il existe un espace probabilisé $(\Omega, \mathcal{A}, \mathbf{P})$ et une famille $(U_i)_{i \in I}$ de variables aléatoires définies sur $(\Omega, \mathcal{A})$, indépendantes et qui suivent toutes la loi de Bernoulli de paramètre $0 < p < 1$:

$$\forall\, 1 \leqslant k \leqslant R,\ \forall\, n \geqslant 1, \qquad \begin{cases} \mathbf{P}(U_{k,n} = 1) = p \\ \mathbf{P}(U_{k,n} = 0) = q \end{cases}$$

avec $q = 1 - p$ comme de coutume.

L'événement $[U_{k,n} = 1]$ signifie que la n-ième tentative de greffe du k-ième rosier a pris.

❧ On pose, *par convention,*

$$[X_k = 0] = \bigcap_{n \in \mathbb{N}^*} [U_{k,n} = 0] \in \mathcal{A}$$

pour envisager le cas d'une greffe qui ne prendrait jamais. On peut vérifier que cet événement est négligeable (par continuité décroissante).

On envisage maintenant le cas d'une greffe qui prend du premier coup :

$$[X_k = 1] = [U_{k,1} = 1] \in \mathcal{A}$$

et plus généralement d'une greffe qui finit par prendre : pour tout $n \geqslant 2$,

$$[X_k = n] = [U_{k,1} = \dots = U_{k,n-1} = 0, U_{k,n} = 1] \in \mathcal{A}.$$

On en déduit que les X_k sont bien des variables aléatoires sur $(\Omega, \mathcal{A})$ qui suivent toutes la loi géométrique de paramètre p et, par coalition, ces variables sont indépendantes.

❧ Par définition,

$$X = X_1 + \dots + X_R.$$

Cette fonction X est une variable aléatoire discrète en tant que somme d'un nombre fini de variables aléatoires discrètes. Comme ces variables sont indépendantes et de même loi, on peut calculer facilement la fonction génératrice de X :

$$\forall\, t \in [0,1], \quad \mathbf{E}(t^X) = \big[\mathbf{E}(e^{tX_1})\big]^R$$

et en déduire la loi de X.

❧ Dire que toutes les greffes ont prises en moins de n semaines signifie que le nombre de greffes nécessaires pour chaque rosier est inférieur à n, soit

$$\forall\, n \in \mathbb{N}^*, \quad [S \leqslant n] = [X_1 \leqslant n, \dots, X_R \leqslant n] \in \mathcal{A}.$$

On en déduit que

$$[S = 1] = [X_1 = 1, \dots, X_R = 1] \in \mathcal{A}$$

et que, pour tout $n \geqslant 2$,

$$[S = n] = [S \leqslant n] \cap [S \leqslant n-1]^c \in \mathcal{A}.$$

Pour que la fonction $S : \Omega \to \mathbb{N}$ soit bien définie, il reste à envisager le cas où aucune greffe ne prend jamais, soit

$$[S = 0] = [X_1 = \cdots = X_R = 0] \in \mathcal{A},$$

événement négligeable (puisque chaque X_k est presque sûrement supérieure à 1).

Cela prouve que la fonction

$$S = \max\{X_1, \ldots, X_R\}$$

est bien une variable aléatoire sur $(\Omega, \mathcal{A})$ et qu'on peut calculer sa loi (puisqu'on sait que les X_k sont indépendantes et que leurs lois sont connues).

❧ Comme les X_k sont des variables aléatoires indépendantes et de même loi, les événements

$$\big([1 \leqslant X_k \leqslant n]\big)_{1 \leqslant k \leqslant R}$$

sont indépendants et de même probabilité, de même que les événements

$$\big([X_k = n]\big)_{1 \leqslant k \leqslant R}.$$

On en déduit que les fonctions définies par

$$Y_n = \sum_{k=1}^{R} \mathbb{1}_{[1 \leqslant X_k \leqslant n]} \quad \text{et par} \quad Z_n = \sum_{k=1}^{R} \mathbb{1}_{[X_k = n]}$$

sont des variables aléatoires qui suivent des lois binomiales (en tant que somme de variables aléatoires de Bernoulli indépendantes et de même paramètre).

En particulier, puisque

$$\forall\, 1 \leqslant k \leqslant R,\ \forall\, n \geqslant 1, \quad \mathbf{P}(1 \leqslant X_k \leqslant n) = 1 - q^n$$

la variable Y_n suit la loi binomiale $\mathscr{B}(R, 1-q^n)$ et on en déduit que

$$\mathbf{P}(Y_n = k) = \binom{R}{k}(1-q^n)^k q^{n(R-k)},$$

ce qui montre que

$$\lim_{n \to +\infty} \mathbf{P}(Y_n = k) = \left|\begin{array}{ll} 0 & \text{si } 0 \leqslant k < R, \\ 1 & \text{si } k = R. \end{array}\right.$$