DM 16
À rendre sur cahier-de-prépa pour, au plus tard, le 30 avril

Formules.

1. le minimum : problème 1, questions 1-5 ; problème 2, questions 1-4. Temps conseillé : 2h30
2. premier prolongement : un peu plus de probas élémentaires avec la fin du problème 1
3. deuxième prolongement (+ intéressant) : la fin du problème 2
4. troisième prolongement : l'exercice de la toute fin

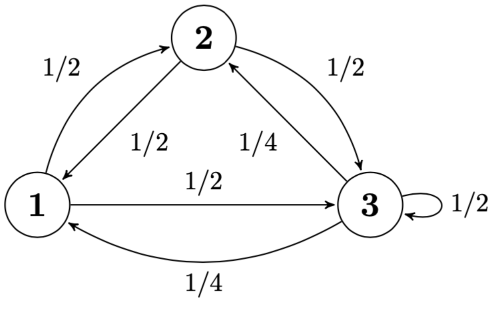
Problème 1. Marches sur des graphes

On appellera « graphe » tout dessin du type de l'un des deux exemples G_1 (ci-dessous) et G_2 (en page 2). Les sommets du graphe sont les cercles numérotés (de 1 à 3 dans le premier exemple, de 1 à 5 dans le second). Les flèches du graphe sont les flèches reliant deux sommets. On remarquera les points suivants :

- entre deux sommets distincts i et j, on peut avoir une flèche de i vers j et une de j vers i
- certains couples de sommets ne sont pas reliés par une flèche (par exemple 1 et 4 dans le graphe G_2)
- une flèche peut relier un sommet à lui-même (c'est le cas du sommet 3 de G_1) ;
- pour tout couple (i, j) de sommets, la flèche allant de i à j est étiquetée par un réel $s_{i,j} \in [0, 1]$, représentant une probabilité de saut (par exemple $s_{3,1} = \dfrac{1}{4}$ et $s_{3,3} = \dfrac{1}{2}$ dans le graphe G_1)
- pour tout sommet i, la somme des probabilités étiquetant les flèches partant de i vaut 1.

Une particule est placée à l'instant $n = 0$ sur le sommet i d'un graphe G. Elle saute aléatoirement à l'instant $n = 1$ sur un autre sommet de G en suivant une des flèches partant de i, la probabilité qu'elle suive la flèche de i vers j étant égale à $s_{i,j}$. On poursuit ainsi le processus, la particule sautant à chaque instant suivant, $n = 2, 3, 4\dots$ du sommet du graphe où elle se trouve vers un nouveau sommet (éventuellement le même) en suivant aléatoirement l'une des flèches selon les probabilités indiquées.

On suppose que les sommets du graphe G sont numérotés de 1 à m. Le processus décrit ci-dessus définit une suite $\mathcal{A} = (X_n)_{n\in\mathbb{N}}$ de variables aléatoires à valeurs dans $\{1, \dots, m\}$ telle que pour tout n on a $X_n = k$ si la particule se trouve sur le sommet k du graphe G après le $n^{\text{ème}}$ saut. $\mathcal{A}$ est appelée marche aléatoire sur le graphe G.
On étudie dans cette question quelques propriétés de la marche aléatoire sur le graphe G_1 ci-dessous.



Graphe G_1

1. On considère la marche aléatoire sur G_1 partant de $X_0 = 1$. Pour $n \in \mathbb{N}$, on note Y_n le vecteur colonne $Y_n = \begin{pmatrix} P(X_n = 1) \\ P(X_n = 2) \\ P(X_n = 3) \end{pmatrix}$. En appliquant la formule des probabilités totales au système complet d'événements associé à la variable aléatoire X_n, établir une relation matricielle entre Y_{n+1} et Y_n de la forme $Y_{n+1} = AY_n$ où A est une matrice de $\mathcal{M}_3(\mathbb{R})$

2. En déduire que, pour tout n entier naturel, $Y_n = A^n Y_0$.

3. Donner, pour $n \geqslant 1$, $P(X_n = 3)$. Justifier votre réponse.

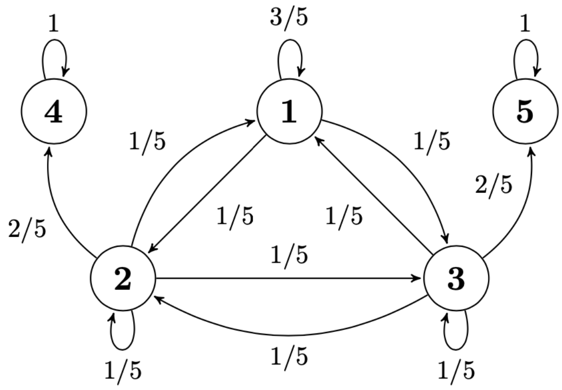
4. (a) Calculer A^2, puis $A^2 \times (2A - I)$ (où I est la matrice identité d'ordre 3). En déduire la relation $A^3 = \dfrac{1}{2}A^2 + \dfrac{1}{2}A$.

(b) En déduire l'existence de deux suites $(u_n)_{n\in\mathbb{N}^*}$ et $(v_n)_{n\in\mathbb{N}^*}$ telles que, pour tout entier naturel n non nul : $A^n = u_n A^2 + v_n A$.

(c) Calculer, pour tout entier naturel n non nul, u_n et v_n.

5. Déterminer, pour tout entier naturel n non nul la loi de X_n.

On considère maintenant une marche aléatoire $\mathcal{A}_2 = (X_n)_{n\in\mathbb{N}}$ sur le graphe G_2 ci-dessous.



Graphe G_2

On constate que lorsque la particule atteint le sommet 4 ou le sommet 5 , elle y reste ensuite indéfiniment avec une probabilité 1 : ces deux sommets sont dits absorbants. On dit que la marche aléatoire est absorbée en 4 ou en 5 lorsqu'elle atteint le sommet correspondant. On s'intéresse ici à la probabilité pour $\mathcal{A}_2$ d'être absorbée en 4 ou en 5 en fonction de son point de départ. Pour tout couple d'entiers (i, j) avec $1 \leqslant i \leqslant 5$ et $4 \leqslant j \leqslant 5$, on note $a_{i,j}$ la probabilité que $\mathcal{A}_2$ soit absorbée en j sachant que $X_0 = i$.

6. Donner les valeurs des $a_{i,j}$ dans le cas particulier $(i, j) \in \{4, 5\}^2$.

7. En distinguant les cas selon le résultat du prochain saut de la particule, montrer que $(x, y, z) = (a_{1,4}, a_{2,4}, a_{3,4})$ est un triplet solution du système

$$\begin{cases} \dfrac{3}{5}x + \dfrac{1}{5}y + \dfrac{1}{5}z = x \\[2ex] \dfrac{1}{5}x + \dfrac{1}{5}y + \dfrac{1}{5}z + \dfrac{2}{5} = y \\[2ex] \dfrac{1}{5}x + \dfrac{1}{5}y + \dfrac{1}{5}z = z \end{cases}$$

Résoudre le système.

8. Donner, par un argument sur la géométrie du graphe ne nécessitant aucun calcul supplémentaire, les valeurs de $(a_{1,5}, a_{2,5}, a_{3,5})$

9. Montrer que la probabilité que $\mathcal{A}_2$ soit absorbée (en 4 ou en 5 indifféremment) est égale à 1, quel que soit son point de départ.

10. On suppose que la loi de X_0 est uniforme sur $\{1, 2, 3\}$, et on constate que $\mathcal{A}_2$ est absorbée en 4 . Quelle est la probabilité qu'elle soit partie du sommet 3?.

Problème 2. Inégalités de concentration

Dans ce problème, on considère $(X_n)_{n\geqslant 1}$ une suite de variables aléatoires mutuellement indépendantes et de même loi que X où X est une variable aléatoire à valeurs réelles, d'espérance μ, de variance $\sigma^2 > 0$. De plus, pour tout entier $n > 0$, et pour tout réel x, on suppose

$$\mathbb{P}\left(\sum_{i=1}^{n}(X_i - \mu) \leqslant x\right) = \mathbb{P}\left(\sum_{i=1}^{n}(X_i - \mu) \geqslant -x\right).$$

A. Questions préliminaires

1. Donner l'espérance et la variance de la variable aléatoire $\dfrac{1}{n}\displaystyle\sum_{i=1}^{n} X_i$. Puis démontrer que pour tout $\varepsilon > 0$,

$$\lim_{n\to+\infty} \mathbb{P}\left(\left|\frac{1}{n}\sum_{i=1}^{n}(X_i - \mu)\right| \geqslant \varepsilon\right) = 0.$$

2. Dans cette question uniquement, on suppose que $\mathbb{P}(X = 1) = \mathbb{P}(X = -1) = \dfrac{1}{2}$. Soit $n \in \mathbb{N}$. Montrer que la variable $S_n = \displaystyle\sum_{i=1}^{n} \frac{X_i + 1}{2}$ suit une loi binomiale de paramètres n et $\dfrac{1}{2}$. En déduire que

$$\lim_{n\to+\infty} \frac{1}{n} \ln\left[\mathbb{P}\left(\left|\frac{1}{n}\sum_{i=1}^{n} X_i\right| \geqslant 1\right)\right] = -\ln(2)$$

B. Une loi des grands nombres logarithmique

On revient maintenant au cas général. On considère deux entiers strictement positifs n et q et deux réels strictement positifs ε_1 et ε_2.

3. Justifier que les deux variables aléatoires $\displaystyle\sum_{i=1}^{n} X_i$ et $\displaystyle\sum_{i=n+1}^{n+q} X_i$ sont indépendantes, puis établir que

$$\mathbb{P}\left(\sum_{i=1}^{n+q} X_i \geqslant \varepsilon_1 + \varepsilon_2\right) \geqslant \mathbb{P}\left(\sum_{i=1}^{n} X_i \geqslant \varepsilon_1\right) \mathbb{P}\left(\sum_{i=n+1}^{n+q} X_i \geqslant \varepsilon_2\right)$$

4. Soit $\varepsilon > 0$. On pose pour tout $n \geqslant 1$,

$$u_n = \mathbb{P}\left(\frac{1}{n}\sum_{i=1}^{n}(X_i - \mu) \geqslant \varepsilon\right).$$

Prouver que pour tout $n \geqslant 1$ et tout $q \geqslant 1$, $u_{n+q} \geqslant u_n u_q$.

Dans toute la suite de cette partie, on suppose qu'il existe $\varepsilon > 0$ tel que $\mathbb{P}(X - \mu \geqslant \varepsilon) > 0$.

5. Montrer que, pour tout $n \geqslant 1$, $u_n > 0$. On pose alors $\alpha_n = -\ln(u_n)$ pour tout $n \geqslant 1$. Montrer que $(\alpha_n)_{n\in\mathbb{N}}$ est sous-additive, c'est-à-dire que pour tout $n \geqslant 1$ et tout $q \geqslant 1$,

$$0 \leqslant \alpha_{n+q} \leqslant \alpha_n + \alpha_q.$$

6. Soit un entier $q \geqslant 1$, on pose $\beta_q = \sup\{\alpha_r, 1 \leqslant r < q\}$. Montrer que, pour tout entier $k \geqslant q$, on a

$$\frac{\alpha_k}{k} \leqslant \frac{\alpha_q}{q} + \frac{\beta_q}{k}.$$

Indication : on pourra utiliser la division euclidienne de k par q.

7. En déduire que la suite $\left(\frac{\alpha_n}{n}\right)_{n\geqslant 1}$ converge vers $\inf\left\{\frac{\alpha_k}{k}, k \geqslant 1\right\}$.
On reviendra à la définition epsilonesque de la borne inférieure.

8. Déduire des questions précédentes que la suite de terme général

$$-\frac{1}{n}\ln\left[\mathbb{P}\left(\left|\frac{1}{n}\sum_{i=1}^{n}(X_i - \mu)\right| \geqslant \varepsilon\right)\right]$$

admet une limite lorsque n tend vers $+\infty$. Puis comparez ce résultat avec celui des questions préliminaires.

C. Cas borné

Pour finir, on suppose que X est à valeurs dans un intervalle $[a, b]$. On note toujours μ son espérance.

9. Soit Y une variable aléatoire centrée, à valeurs dans $[-1, 1]$. Démontrer que pour tout $\gamma \in [-1, 1]$ et tout $t \in \mathbb{R}$,

$$e^{t\gamma} \leqslant \frac{1}{2}(1-\gamma)e^{-t} + \frac{1}{2}(1+\gamma)e^{t},$$

puis en déduire que $\mathbb{E}(e^{tY}) \leqslant \mathrm{ch}(t)$. On admet que, par une étude de fonctions, $\mathrm{ch}(t) \leqslant e^{\frac{t^2}{2}}$. On a donc $\mathbb{E}(e^{tY}) \leqslant e^{\frac{t^2}{2}}$.

10. En considérant $Y = \dfrac{X - \mu}{b - a}$, vérifier que Y est centrée et à valeurs dans $[-1, 1]$ et en déduire que pour tout t réel,

$$\mathbb{E}(e^{t(X-\mu)}) \leqslant e^{\frac{t^2(b-a)^2}{2}}.$$

11. Soit $\varepsilon > 0$. En utilisant le résultat de la question précédente, appliquée à la variable aléatoire $\displaystyle\sum_{i=1}^{n}(X_i - \mu)$, et en choisissant un t convenable, montrer que

$$\mathbb{P}\left(\frac{1}{n}\sum_{k=1}^{n}(X_k - \mu) \geqslant \varepsilon\right) \leqslant \exp\left(-\frac{n\varepsilon^2}{2(b-a)^2}\right),$$

puis majorer $\mathbb{P}\left(\left|\frac{1}{n}\sum_{k=1}^{n}(X_k - \mu)\right| \geqslant \varepsilon\right)$. Comparer avec le résultat de la question 8.

Exercice 1. *Court, mais festif*. Soit $(G, *)$ un groupe abélien fini, X une variable aléatoire uniforme sur G, Y une variable aléatoire quelconque sur G, indépendante de X. Démontrer que $X * Y$ est uniforme sur G.