TD 10 – Méthode probabiliste (suite) et chaines de Markov (corrigé)

Exercice 1. *Lemme local de Lovasz*

Soit $k > 6$. On se donne une famille $(A_i)_{i\in I}$ de sous-ensembles d'un ensemble fini F telle que :

1. Pour tout $i \in I$, $|A_i| = k$,
2. Pour tout $x \in F$, $|\{i \in I \mid x \in A_i\}| \leq \frac{2^k}{8k}$.

En utilisant le lemme local de Lovász[1], montrer qu'il existe une partition $F = F_1 \cup F_2$ telle que

$$\forall i \in I, \qquad A_i \cap F_1 \neq \emptyset \qquad \text{et} \qquad A_i \cap F_2 \neq \emptyset.$$

Lemme Local de Lovász (rappel) : *Soient n, d des entiers, $0 \leq p \leq 1$ et $A_1, \ldots, A_n$ des événements tels que :*

1. *pour tout $1 \leq i \leq n$, on a $\mathbf{P}\{A_i\} \leq p$,*
2. *les événements $(A_i)_{1\leq i\leq n}$ admettent un graphe de dépendance de degré $\leq d$,*
3. *on a $4dp \leq 1$.*

Alors on a $\mathbf{P}\left\{\overline{A_1} \cap \ldots \cap \overline{A_n}\right\} > 0$.

☞ On partitionne F au hasard en décidant indépendamment et uniformément si chaque $x \in F$ est dans F_1 ou F_2.
Pour tout $i \in I$, soit E_i l'événement $\{A_i \cap F_1 = \emptyset\} \cup \{A_i \cap F_1 = \emptyset\}$.
Chaque élément $x \in A_i$ appartient à au maximum $\frac{2^k}{k} - 1$ autres A_j par hypothèse 2. Du plus A_i a au maximum k éléments par hypothèse 1, donc finalement E_i est indépendant de tous les autres E_j sauf au plus $\frac{2^k}{8}$ d'entre eux.
On peut alors appliquer le lemme local de Lovász avec :

- $\mathbf{P}\{E_i\} \leq 2 \times \frac{1}{2^k}$ donc $p := \frac{1}{2^{k-1}}$,
- $d := \frac{2^k}{8}$,
- on vérifie qu'on a bien $4dp = 1$.

Et alors $\mathbf{P}\left\{\bigcap_{i\in I} \overline{E_i}\right\} > 0$.

Remarque : le $k > 6$ vient de la condition 2 où on veut $\frac{2^k}{8k} > 1$ (sinon pour tout x, x n'appartient à aucun A_i, donc tous les A_i sont vides, contradiction avec la condition 1).

Exercice 2. *Partition de graphe*

Soit $G = (V, E)$ un graphe non dirigé avec n sommets et m arrêtes.
Montrer qu'il existe une partition de V en deux ensembles disjoints A et B telle que au moins la moitié des arrêtes de G relie un sommet de A et un sommet de B.

☞ On construit une partition $G = A \uplus B$ au hasard, en choisissant pour chaque sommet e s'il appartient à A ou B uniformément et indépendamment. Pour toute arrête $e_i \in E$ de G, on note X_i l'événement « e_i relie A et B ». Alors $\mathbf{P}\{X_i = 1\} = \frac{1}{2}$, donc $\mathbf{E}[X_i] = \frac{1}{2}$.
Notons $M = \sum_{i=1}^m X_i$ la v.a. qui compte le nombre d'arrêtes reliant A et B. Son espérance est

$$\mathbf{E}[M] = \sum_{i=1}^m \mathbf{E}[X_i] = \frac{m}{2}.$$

Mais on a toujours $\mathbf{P}\{M \geq \mathbf{E}[M]\} > 0$ (et $\mathbf{P}\{M \leq \mathbf{E}[M]\} > 0$), donc en particulier $\mathbf{P}\left\{M \geq \frac{m}{2}\right\} > 0$, d'où le résultat.

Exercice 3. *Meteo*

Dans un certain pays, il ne fait jamais beau deux jours de suite. Si un jour il fait beau, alors le lendemain il peut neiger ou pleuvoir avec autant de chances. Si un jour il pleut ou il neige, il y a une chance sur deux qu'il y ait changement de temps le lendemain, et s'il y a changement, il y a une chance sur deux que ce soit pour du beau temps.

1. Former, à partir de cela, une chaîne de Markov que l'on représentera à la fois sur sa forme graphique et sous la forme matrice de transition.

   ☞ voir `http://w3.mathinfolmd.univ-tlse2.fr/membres/chabriac/M1process/exopoly2c.pdf`
   L'ensemble des états est $E = \{B, P, N\}$ pour "Beau", "Pluie", Neige". Le temps pour un jour ne dépend que du temps du jour précédent, indépendamment de tout le reste. On a donc bien un chaîne de Markov, dont la matrice de transition est :

$$P = \begin{pmatrix} 0 & 1/2 & 1/2 \\ 1/4 & 1/2 & 1/4 \\ 1/4 & 1/4 & 1/2 \end{pmatrix}$$

1. László Lovász (né en 1948), mathématicien hongrois.

2. Si un jour il fait beau, quel est le temps le plus probable pour le surlendemain ?

☞ Pour le temps du surlendemain, il faut déterminer P^2. Ici, seulement la première ligne nous intéresse donc on s'économise les deux autres lignes :

$$P^2 = \begin{pmatrix} 1/4 & 3/8 & 3/8 \\ & & \end{pmatrix}$$

Conclusion : si un jour il fait beau, le temps le plus probable pour le surlendemain est la pluie et ou la neige, avec même probabilité.

3. Peut-on supposer maintenant que l'on a deux deux états, un pour "Beau temps" et un pour "Mauvais Temps" ? Déterminer la nouvelle matrice de transition.

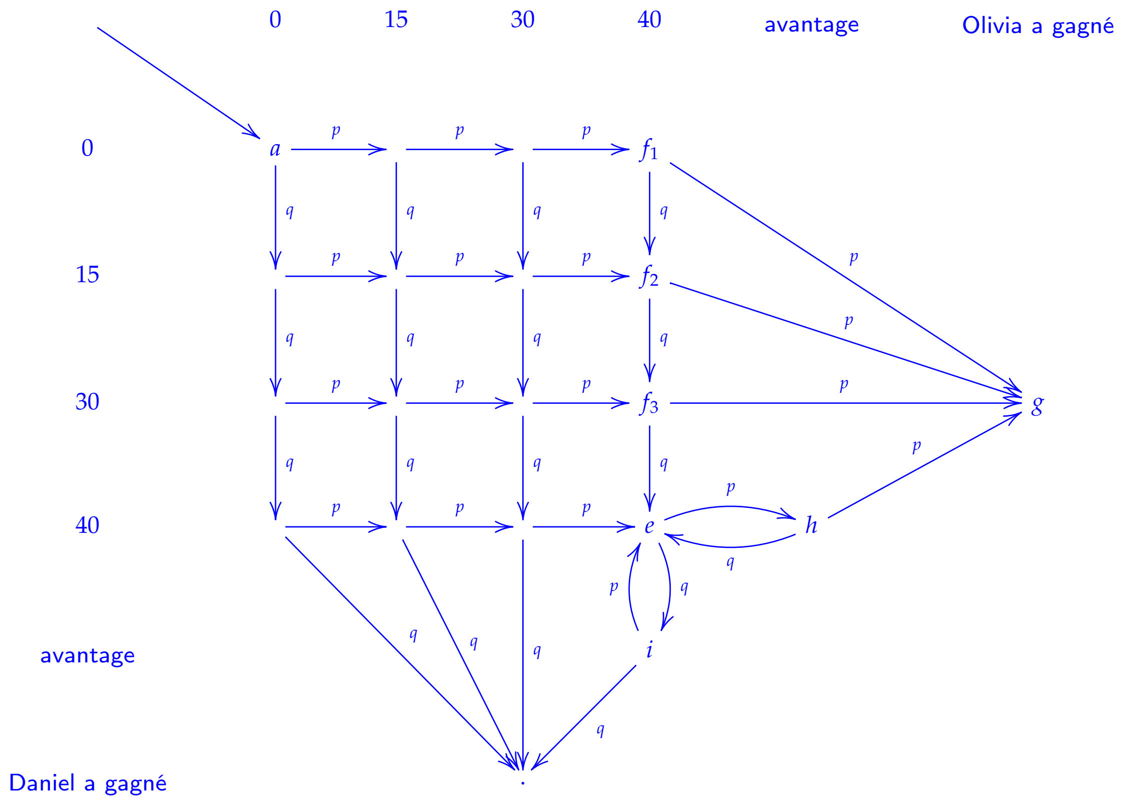
☞ Oui, on peut car les états pluie et neige se comportent de la meme facon. On a donc maintenant :

$$P' = \begin{pmatrix} 0 & 1 \\ 1/4 & 3/4 \end{pmatrix}$$

Exercice 4. *Tennis*

1. On considère un jeu classique du tennis (pas un jeu décisif) de Daniel contre Olivia. Olivia gagne chaque point avec probabilité p. Modéliser par une chaîne de Markov et donner la probabilité pour Olivia de gagner le jeu.

☞ On va noter $q = 1 - p$. Voici la chaîne de Markov de ce jeu de tennis :



La probabilité d'arriver en g est la somme sur tous les chemins menant à g de la probabilité de chaque chemin. On compte ces chemins selon le dernier état rencontré :

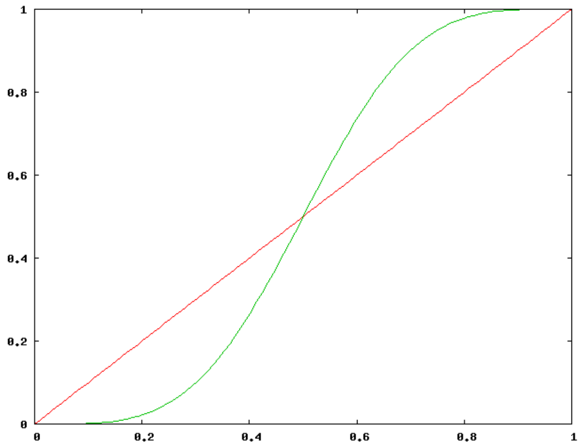
- s'il s'agit de f_1 la probabilité du chemin est p^4 ;
- pour f_2, il y a C_4^1 chemins possibles, chacun de probabilité p^4q ;
- pour f_3, il y a de même C_5^2 chemins, chacun de probabilité p^4q^2 ;
- s'il s'agit de h, on décompose le chemin en trois étapes :
  - la portion $a...e$, il y a C_6^3 chemins, chacun de probabilité p^3q^3,
  - un certain nombre de boucles ehe ou eie, la probabilité de faire n boucles est $2^n(pq)^n$,
  - la portion ehg, de probabilité p^2.

En sommant toutes les possibilités, la probabilité d'arriver en g est

$$p^4 + 4p^4q + 10p^4q^2 + 20p^3q^3 \times \underbrace{\sum_{i=0}^{\infty}(2pq)^n}_{\frac{1}{1-2pq}} \times p^2.$$

2. Lorsque $p \approx 0$, donner l'équivalent asymptotique de cette probabilité. Commenter.

☞ Si $p \approx 0$, alors $q \approx 1$ et c'est p^4 qui domine. L'équivalent est alors $p^4 + 4p^4 + 10p^4 = 15p^4 << p$. Les joueurs forts sont avantagés. Cela ce voit aussi sur un dessin. Voici en vert la probablité qu'Olivia gagne en fonction de p :



On a augmenté le contraste : le systéme de comptage des points avantage le plus fort joueur.

Exercice 5. *Marche aléatoire sur Z biaisée*

Soit $p \in]0,1[$, et considérons la chaîne de Markov d'espace d'états $\mathbb{Z}$ et de matrice de transition

$$P(i,j) = \begin{cases} p \text{ si } j = i+1 \\ 1-p \text{ si } j = i-1 \\ 0 \text{ sinon} \end{cases}$$

1. Cette chaîne est-elle irréductible ?

☞ Oui, le graphe associé est connexe.

2. Dans cette question on suppose $p \neq 1/2$, montrer que tous les états de cette chaîne sont transients. *Indication : on pourra utiliser le lemme de Borel-Cantelli. On rappelle également l'équivalent de Stirling :* $n! \sim \sqrt{2\pi n}(n/e)^n$.

☞ On rappelle le lemme de Borel-Cantelli : si A_n sont des événements tels que $\sum_n \mathbf{P}\{A_n\} < \infty$, alors la probabilité qu'une infinité de A_n se réalisent simultanément est nulle.

Pour montrer que les états sont transients, comme la chaîne est irréductible, il suffit de montrer que l'état 0 est transient. C'est à dire qu'il suffit de montrer que $\Pr(N_0 < \infty | X_0 = 0) = 1$, où $N_0 = \sum_n 1_{X_n=0}$. Notons A_n l'événement "$X_n = 0$". la variable N_0 est alors le nombre de A_n qui se réalisent simultanément. On va montrer avec Borel-Cantelli que ce nombre est fini presque sûrement.

Pour cela, calculons $\Pr(A_n | X_0 = 0) = \Pr(X_n = 0 | X_0 = 0)$. Si n est impair, alors X_n ne peut pas être égal à 0 (il faut faire un nombre pair de pas pour revenir au départ. Si $n = 2k$, alors $\Pr(X_n = 0 | X_0 = 0) = \binom{2k}{k}(p(1-p))^k$ (on obtient cette probabilité en sommant la probabilité $((p(1-p)^k)$ de tous les $\binom{2k}{k}$ chemins possibles).

L'équivalent de Stirling nous donne $\binom{2k}{k} = O(4^n)$. Comme $p \neq 1/2$, on a $p(1-p) < 1/4$. On conclut que la série des $\binom{2k}{k}(p(1-p))^k = O((4/(p(1-p)))^k)$ converge. Donc, par Borel-Cantelli, on a $\Pr(N_0 < \infty | X_0 = 0) = 1$, ce qui conclut la preuve.

Remarque : on peut alternativement prouver ce résultat avec la loi forte des grands nombres. On peut écrire $X_n = \sum_{i=1}^n Y_i$, avec $Y_i = 1$ avec probabilité p et $Y_i = -1$ avec probabilité $1-p$, et les Y_i indépendantes. La loi forte des grands nombre nous dit alors que presque sûrement, X_n/n converge vers $\mathbf{E}[Y_i] = 2p - 1$. Mais alors $X_n \sim (2p-1)n$ (car $2p - 1 \neq 0$ par hypothèses). Donc X_n tend vers $\pm\infty$, et on en déduit qu'il ne passe qu'un nombre fini de fois par 0.

Exercice 6. *Marche aléatoire sur Z non biaisée*

Soit $\{X_k\}$ des variables aléatoires discrétes indépendantes et identiquement distribuées. Chaque X_k prend la valeur 1 avec probabilité $1/2$ et -1 avec probabilité $1/2$. On définit alors une marche aléatoire dans R par $S_n = \sum_{k=1}^n X_k$. On s'intéresse à la probabilité d'un retour à l'origine (en un temps fini).

1. S'il y a eu un retour à l'origine au temps m, que peut-on dire de m ? Montrer qu'un retour à l'origine au temps $2n$ arrive avec une probabilité $u_{2n} = \binom{2n}{n}2^{-2n}$.

☞ Déjà, on ne peut revenir à l'origine que si le temps m est pair.

Ensuite, on revient à l'origine au temps $2n$ si on a fait n pas à gauche parmi les $2n$ pas. Cela donne bien $u_{2n} = \binom{2n}{n}2^{-2n}$.

2. On définit de même la probabilité f_{2n} qu'un premier retour à l'origine se fasse au temps $2n$. Montrer que pour $n > 0$ les probabilités $\{f_{2k}\}$ et $\{u_{2k}\}$ vérifient la relation $u_{2n} = f_0 u_{2n} + f_2 u_{2n-2} + \cdots + f_{2n} u_0$ (on pose $u_0 = 1$ et $f_0 = 0$).

☞ On a pour tout $n > 0$:

$$\begin{aligned} u_{2n} &= \mathbf{P}\{S_{2n} = 0\} = \mathbf{P}\left\{ \bigcup_{0<k\leq n} S_{2n} = 0 \cap S_{2k} = 0 \cap S_i \neq 0 \text{ pour } i < 2k \right\} \\ &= \mathbf{P}\{S_{2n} = 0 \cap S_i \neq 0 \text{ pour } i < 2n\} \\ &\quad + \sum_{k=1}^{n-1} S_{2n} = 0 \cap S_{2k} = 0 \cap S_i \neq 0 \text{ pour } i < 2k \qquad \text{(par indépendance)} \\ &= u_0 f_{2n} + \sum_{k=1}^{n-1} u_{2(n-k)} f_{2k} \qquad \text{(par indépendance + décalage)} \\ &= u_0 f_{2n} + u_2 f_{2n-2} + \ldots + u_{2n-2} f_2 + f_0 u_{2n}. \end{aligned}$$

3. On définit les fonctions génératrices :

$$U(x) = \sum_{m=0}^{\infty} u_{2m} x^m \text{ et } F(x) = \sum_{m=0}^{\infty} f_{2m} x^m$$

Déduire de la question précédente une relation simple entre $U(x)$ et $F(x)$.

☞ On a $U(x) = 1 + U(x) \cdot F(x)$.

Attention : la formule de la question précédente n'est pas vraie pour $n = 0$.

4. Montrer que $U(x) = \frac{1}{\sqrt{1-x}}$. En déduire que $F(x) = 1 - \sqrt{1-x}$.

Indication : on rappelle que $(1+x)^\alpha = 1 + \sum_{k=1}^{+\infty} \frac{\alpha(\alpha-1)\ldots(\alpha-k+1)}{k!} x^k$.

☞ On a :

$$\begin{aligned} \frac{1}{\sqrt{1-4x}} &= (1-4x)^{-1/2} = 1 + \sum_{k=0}^{+\infty} \frac{-\frac{1}{2}(-\frac{1}{2}-1)\ldots(-\frac{1}{2}-k+1)}{k!} (-4x)^k \\ &= 1 + \sum_{k=1}^{+\infty} \frac{1 \cdot 3 \ldots (2k-1)}{(-2)^k k!} (-2)^k 2^k x^k = 1 + \sum_{k=1}^{+\infty} \frac{(2k)!}{k!k!} x^k = \sum_{k=0}^{+\infty} \binom{2k}{k} x^k \end{aligned}$$

On en déduit par changement de variable que

$$U(x) = \sum_{k=0}^{+\infty} \binom{2k}{k} 2^{-2k} x^k = \sum_{k=0}^{+\infty} \binom{2k}{k} \left(\frac{x}{4}\right)^k = \frac{1}{\sqrt{1-x}}.$$

Ainsi :

$$F(x) = \frac{U(x) - 1}{U(x)} = 1 - \sqrt{1-x}.$$

5. Montrer que $f_{2m} = \frac{\binom{2m}{m}}{(2m-1)2^{2m}}$.

Indication : considérer F'.

☞ On a $F'(x) = \frac{1}{2\sqrt{1-x}} = \frac{U(x)}{2}$.

Ainsi, $m \cdot f_{2m} = \frac{u_{2(m-1)}}{2} = \binom{2m-2}{m-1} 2^{-2m+1} = \frac{m}{2(2m-1)} \frac{\binom{2m}{m}}{2^{2m-1}}$, d'où la formule annoncée.

6. Définissons w_n la probabilité qu'un retour à l'origine se fasse au plus tard au temps n. Notre but est de savoir si l'on va revenir en un temps fini, c'est-à-dire déterminer $w_* = \lim_{n\to\infty} w_n$. Montrer que $w_* = F(1)$. Conclure.

☞ On a $w_n = \sum_{k=1}^{\lfloor n/2 \rfloor} f_{2k} \to w_* = F(1) = 1$. On retourne presque sûrement à l'origine en un temps fini.