Examen final

26 mai 2025

Exercice 1 Méthode probabiliste

Soit $G = (V, E)$ un graphe (non orienté, sans boucle ni arête multiple). A l'aide de la méthode probabiliste, montrer qu'il existe une partition $V = V_1 \cup V_2$ telle que le nombre d'arêtes reliant un sommet de V_1 à un sommet de V_2 est $\geqslant |E|/2$.

Exercice 2 Chiffre des unités de la somme de $n \gg 1$ dés

Soit (X_n) une suite de variables aléatoires i.i.d. de loi uniforme sur $\{1, \ldots, 6\}$. On note Y_n la valeur de $X_1 + \cdots + X_n$ modulo 10. A l'aide d'un théorème sur les chaînes de MARKOV, déterminer pour tout $k \in \{0, \ldots, 9\}$ la valeur de

$$\lim_{n\to\infty} \mathbf{P}(Y_n = k).$$

Exercice 3 Valeurs extrêmes pour le temps de retour

Dans cet exercice les graphes sont supposés connexes, non orientés, sans boucle ni arête multiple. On fixe un entier $n \geqslant 1$.

Étant donné un graphe $G = (V, E)$ à n sommets et $x \in V$, on s'intéresse à la quantité $\mathbf{E}_x[T_x]$, le temps moyen de retour en x, pour la marche aléatoire sur le graphe G.

1. Quel choix de G et de x minimise la quantité $\mathbf{E}_x[T_x]$?
2. Quel choix de G et de x maximise la quantité $\mathbf{E}_x[T_x]$?

Exercice 4 Lecture

Vous décidez de lire les livres de la bibliothèque de l'ENS selon l'algorithme suivant : après avoir lu un livre, ce qui vous prend une semaine, vous lisez le suivant dans l'ordre alphabétique (sans revenir en arrière si jamais des livres ont été ajoutés entre temps). La bibliothèque achète un nombre fixe $k \in \mathbf{N}^*$ de livres par semaine. Le but de cet exercice est de calculer la probabilité de lire le dernier livre dans l'ordre alphabétique.

On modélise le problème en identifiant l'ensemble des livres à $[0, 1]$ et en supposant que les achats sont indépendants et de loi uniforme. Soient $(Z_n^j)_{n\geqslant 1, 1\leqslant j\leqslant k}$ une famille de variables aléatoires i.i.d. de loi uniforme dans $[0, 1]$. L'ensemble $L_n = \{Z_m^j : m \leqslant n, 1 \leqslant j \leqslant k\}$ décrit la bibliothèque à la $n$ème semaine. On définit par récurrence une suite (X_n) par $X_0 = 0$ et

$$X_{n+1} = \begin{cases} \text{le minimum de l'ensemble }]X_n, 1] \cap L_{n+1} & \text{si ce dernier est non vide} \\ 1 & \text{sinon} \end{cases}$$

La v.a. $X_n \in [0, 1]$ correspond au livre lu à la $n$ème semaine. On veut calculer $\mathbf{P}(\exists n : X_n = 1)$, la probabilité d'avoir lu le dernier livre.

1. Montrer que si $Z_1, \ldots, Z_N$ sont des variables aléatoires i.i.d. de loi uniforme dans $[0, 1]$ alors $\mathbf{E}[\min(Z_1, \ldots, Z_N)] = \frac{1}{N+1}$. On pourra utiliser le fait que si X est une variable aléatoire positive, alors $\mathbf{E}[X] = \int_0^\infty \mathbf{P}(X > t)\,\mathrm{d}t$.
2. On pose $Y_n = 1 - X_n$. Montrer que pour tout $n \geqslant 1$, $\mathbf{E}[Y_n] \leqslant \left(1 - \frac{1}{kn+1}\right) \mathbf{E}[Y_{n-1}]$.
3. En déduire que $\lim_{n\to\infty} \mathbf{E}[Y_n] = 0$. On admet que, à l'aide du théorème de convergence dominée, cela implique que (Y_n) tend presque sûrement vers 0.
4. Soit $n \in \mathbf{N}$. On considère l'événement $A_n = \{Y_n < 1/k\}$. Quel théorème implique que, conditionnellement à A_n, on a presque sûrement $Y_m = 0$ pour m assez grand ?
5. Conclure.

Exercice 5 Algorithme probabiliste pour 3-SAT

0. Rappeler le principe de l'algorithme probabiliste pour 2-SAT vu en cours (on demande uniquement les grandes lignes du raisonnement, pas de refaire tous les calculs en détail).

On va adapter cet algorithme au problème 3-SAT. On considère une formule 3-SAT en N variables booléennes $x_1, \ldots, x_N$, c'est-à-dire une conjonction de disjonctions (appelées clauses) dans lesquelles 3 variables exactement (ou leur négation) apparaissent. On suppose qu'une clause ne contient jamais deux fois la même variable, ni une variable et sa négation. Un exemple de telle formule est

$$(x_1 \vee \overline{x_2} \vee \overline{x_6}) \wedge (x_3 \vee \overline{x_1} \vee \overline{x_6}) \wedge (\overline{x_4} \vee x_2 \vee \overline{x_6}) \wedge (x_5 \vee \overline{x_2} \vee x_1)$$

On considère l'algorithme probabiliste suivant, où $m \in \mathbf{N}$ est un paramètre, afin de décider si une formule 3-SAT est satisfiable.

```
RANDOM 3-SAT
  1 Initialiser avec une affectation des variables choisie uniformément au hasard.
  2 Répéter m fois, en s'arrêtant si toutes les clauses sont satisfaites
    (a) Choisir arbitrairement une clause non satisfaite.
    (b) Choisir uniformément au hasard une des 3 variables de cette clause et la remplacer par
        sa négation.
  3 Si une affectation valide a été trouvée, la retourner.
    Sinon, répondre que la formule n'est pas satisfiable.
```

1. Supposons qu'il existe une affectation S des variables satisfaisant la formule. On note $X_n \in \{0, \ldots, N\}$ le nombre de variables ayant la même valeur que dans S au $n$ème tour de boucle de l'algorithme. Étant donnés $i \in \{0, \ldots, N-1\}$ et $j \in \{0, \ldots, N\}$, que pouvez-vous dire de

$$\mathbf{P}(X_{n+1} = j | X_n = i) \ ?$$

2. On considère la chaîne de Markov $(Y_n)_{n \geqslant 0}$ d'espace d'états $\{0, \ldots, N\}$ donnée par la matrice de transition

$$Q(0,1) = Q(N,N) = 1$$
$$Q(i, i-1) = 2/3,\ Q(i, i+1) = 1/3 \qquad \text{si } 0 < i < N,$$

les autres coefficients étant nuls.

   (a) Quels sont les états récurrents ou transitoires ?
   (b) On note $T_N = \inf\{n \geqslant 0 : Y_n = N\}$ le temps d'atteinte de l'état N. Pour $0 \leqslant j \leqslant N$, on note $h_j = \mathbf{E}_j[T_N]$ (où $\mathbf{E}_j$ dénote l'espérance pour la chaîne issue de j). Donner une relation de récurrence pour h_j.
   (c) En déduire que $h_j = 2^{N+2} - 2^{j+2} - 3(N-j)$.

3. Comment faut-il choisir le paramètre m (en fonction de N) de l'algorithme `RANDOM 3-SAT` pour que la probabilité d'erreur soit inférieure à 2^{-100} ?
4. Pour $0 \leqslant j \leqslant N$, on pose $q(j) = \mathbf{P}_j(T_N \leqslant 3N)$. Montrer que

$$q(N-j) \geqslant \binom{3j}{j} \left(\frac{2}{3}\right)^j \left(\frac{1}{3}\right)^{2j} \geqslant \frac{1}{(3j+1)2^j}$$

5. Montrer que si $m = 3N$, la probabilité que l'algorithme `RANDOM 3-SAT` (appelé sur une formule satisfiable) trouve une affectation valide des variables est $\geqslant \frac{1}{3N+1}\left(\frac{3}{4}\right)^N$.
6. En déduire un algorithme probabiliste pour 3-SAT de meilleure complexité que celle obtenue à la question 3.
7. Comment cela ce généralise-t-il au problème k-SAT pour $k > 3$?