ENS LYON | Informatique
L3 Probabilités

Examen partiel du 21 mars 2018
Durée : 2 heures

Exercice 1 Question de cours

Démontrer la version suivante de l'inégalité de Chernoff : si $(X_1, \ldots, X_n)$ sont des v.a. i.i.d. de loi de Bernoulli de paramètre $1/2$, alors pour tout $a \geqslant 0$,

$$\mathbf{P}\left(X_1 + \cdots + X_n \geqslant \frac{n}{2} + a\right) \leqslant \exp(-2a^2/n).$$

On pourra utiliser sans justification l'inégalité $\cosh(t) \leqslant \exp(t^2/2)$ pour $t \in \mathbf{R}$.

Exercice 2 Graphe aléatoire bipartite

Soit $0 < p < 1$ et $n \in \mathbf{N}^*$. On définit un graphe aléatoire non orienté $H_{2n,p}$ de la manière suivante. On se donne une famille $\{X_{i,j} \,:\, 1 \leqslant i \leqslant n,\ n+1 \leqslant j \leqslant 2n\}$ de v.a. i.i.d. de loi de Bernoulli de paramètre p. On pose alors $H_{2n,p} = (V, E)$, avec $V = \{1, \ldots, 2n\}$ et

$$E = \{(i,j) \,:\, X_{i,j} = 1\} \subset \{1, \ldots, n\} \times \{n+1, \ldots, 2n\}.$$

1. Quelle est la loi du nombre d'arêtes de $H_{2n,p}$?
2. Quelle est l'espérance du nombre de sommets isolés de $H_{2n,p}$?
3. Dans cette question on pose $p = c\log(n)/n$ pour un nombre réel $c > 0$.
   (a) Montrer que si $c > 1$, alors
   $$\lim_{n\to\infty} \mathbf{P}(H_{2n,p} \text{ a un sommet isolé}) = 0.$$
   (b) Montrer que si $c < 1$, alors
   $$\lim_{n\to\infty} \mathbf{P}(H_{2n,p} \text{ a un sommet isolé}) = 1.$$
4. Dans cette question on pose $p = 1/2$. Montrer qu'il existe une constante $C > 0$ telle que
   $$\lim_{n\to\infty} \mathbf{P}\left(\text{tous les sommets de } H_{2n,p} \text{ ont un degré inférieur à } \frac{n}{2} + C\sqrt{n\log n}\right) = 1.$$

Exercice 3 Interrupteurs

Partie I :

1. Montrer qu'il existe une constante $\gamma > 0$ rendant l'énoncé suivant vrai : si une v.a. positive X vérifie $\mathbf{E}[X] = 1$ et $\mathbf{E}[X^2] \leqslant 3$, alors $\mathbf{P}(X \geqslant 1/4) \geqslant \gamma$.
2. Soient $(X_1, \ldots, X_n)$ des v.a. i.i.d. vérifiant $\mathbf{P}(X_i = 1) = \mathbf{P}(X_i = -1) = \frac{1}{2}$. On pose $Y = \frac{1}{\sqrt{n}}(X_1 + \cdots + X_n)$. Calculer $\mathbf{E}[Y^2]$ et $\mathbf{E}[Y^4]$ et en déduire que
   $$\mathbf{E}[|X_1 + \cdots + X_n|] \geqslant \frac{\gamma}{2}\sqrt{n}.$$

Partie II :

On considère une grille $n \times n$ d'ampoules ainsi que 3 séries d'interrupteurs : des interrupteurs $a = (a_{ij})_{1 \leqslant i,j \leqslant n}$ associés à chaque ampoule, des interrupteurs $b = (b_i)_{1 \leqslant i \leqslant n}$ associés à chaque ligne et des interrupteurs $c = (c_j)_{1 \leqslant j \leqslant n}$ associés à chaque colonne. Chaque interrupteur prend la valeur -1 ou 1. L'ampoule en position (i, j) est allumée si et seulement si $a_{ij} b_i c_j = 1$. On considère la quantité

$$F(a, b, c) = \sum_{i,j=1}^{n} a_{ij} b_i c_j$$

qui est le nombre d'ampoules allumées moins le nombre d'ampoules éteintes. Enfin, deux joueurs jouent au jeu suivant : le joueur 1 choisit la position des interrupteurs (a_{ij}), puis le joueur 2 choisit la position des interrupteurs (b_i) et (c_j). Le joueur 1 veut minimiser $F(a, b, c)$ et joueur 2 veut le maximiser. On considère donc

$$V(n) = \min_{a \in \{-1,1\}^{n \times n}} \max_{b,c \in \{-1,1\}^n} F(a, b, c).$$

1. Montrer que $V(n) = O(n^{3/2})$ en considérant le cas où le joueur 1 joue au hasard.
2. Le joueur 2 applique la stratégie suivante : il choisit b au hasard, puis ensuite choisit c de façon à allumer le maximum de lampes. Estimer le nombre moyen de lampes allumées par cette stratégie à l'aide de la question I.2 et en déduire que $V(n) = \Omega(n^{3/2})$.

Exercice 4 Algorithme probabiliste pour trouver la médiane

On étudie un algorithme probabiliste [1] pour déterminer la médiane d'un ensemble $E = \{x_1, \ldots, x_n\}$ de n nombres réels en temps $O(n)$. On rappelle que m est une médiane de E si au moins $\lceil n/2 \rceil$ des élements de E sont inférieurs ou égaux à m, et au moins $\lceil n/2 \rceil$ des élements de E sont supérieurs ou égaux à m. Pour simplifier on suppose n impair (ce qui fait que la médiane est unique) et on suppose aussi que les éléments de E sont tous distincts.

Voici comment fonctionne l'algorithme

(a) Soit $(Y_i)_{1 \leqslant i \leqslant n}$ une suite de v.a. i.i.d. de loi de Bernoulli de paramètre $n^{-1/4}$. On considère le sous-ensemble aléatoire de E défini par $F = \{x_i : Y_i = 1\}$. Si $\operatorname{card} F \leqslant \frac{2}{3} n^{3/4}$ ou $\operatorname{card} F \geqslant 2n^{3/4}$ on répond «`ERREUR 1`».

(b) On trie F et on appelle d le $\lfloor \frac{1}{2} n^{3/4} - \sqrt{n} \rfloor$ème plus petit élément de F, et u le $\lfloor \frac{1}{2} n^{3/4} - \sqrt{n} \rfloor$ème plus grand élément de F.

(c) On détermine le rang de d et de u dans E (l'élément minimal a rang 1, l'élément maximal a rang n), que l'on note respectivement r_d et r_u. Si $r_d > n/2$ ou $r_u < n/2$ on répond «`ERREUR 2`».

(d) On note $G = \{x_i \in E : d < x_i < u\}$. Si $\operatorname{card} G \geqslant 4n^{3/4}$ on répond «`ERREUR 3`».

(e) On trie G et on renvoie le $(\lceil n/2 \rceil - r_d)$ème élement de G.

1. Justifier pourquoi l'algorithme retourne la médiane en temps $O(n)$ lorsqu'il ne répond pas de message d'erreur.
2. Montrer que pour $i \in \{1, 2, 3\}$, on a

$$\lim_{n \to \infty} \mathbf{P}\,(\text{l'algorithme retourne «}\texttt{ERREUR } i\text{»}) = 0.$$

   Pour simplifier l'analyse et éviter d'écrire des symobles $\lfloor \cdot \rfloor$ ou $\lceil \cdot \rceil$, on pourra supposer implicitement que des nombres tels que $\sqrt{n}$, $\frac{1}{2} n^{3/4}$, ... sont des entiers

1. Remarque : il existe un algorithme déterministe de même performance