TD 05 – Inégalité de Chernoff (2)

Exercice 1. *Interrupteurs*

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 $\mathbf{E}\left[X^2\right] \leq 3$, alors $\mathbf{P}\{X \geq 1/4\} \geq \gamma$. »

   Indication : définir la variable aléatoire $Y = \mathbf{1}_{X \geq 1/4}$ et se ramener à l'inégalité de Cauchy-Schwarz. $\mathbb{E}(XY) \leq \sqrt{\mathbb{E}(X^2)\mathbb{E}(Y^2)}$
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 $\mathbf{E}\left[Y^2\right]$ et $\mathbf{E}\left[Y^4\right]$ et en déduire que :
   $$\mathbf{E}\left[|X_1 + \cdots + X_n|\right] \geq \frac{\gamma}{2}\sqrt{n}.$$

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

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

qui est le nombre d'ampoules allumées moins le nombre d'ampoules éteintes.
Deux joueuses jouent au jeu suivant :

1. la joueuse 1 choisit la position des interrupteurs (a_{ij}),
2. puis la joueuse 2 choisit la position des interrupteurs (b_i) et (c_j).

La joueuse 1 veut minimiser $\mathbf{F}(a, b, c)$ et la joueuse 2 veut le maximiser. On considère donc :

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

3. Montrer que $\mathbf{V}(n) = \mathcal{O}(n^{3/2})$ en considérant le cas où la joueuse 1 joue au hasard.

4. La joueuse 2 applique la stratégie suivante : elle 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 *(indication : utiliser la question 2)* et en déduire que $\mathbf{V}(n) = \Omega(n^{3/2})$.

Exercice 2. *Estimer l'intersection avec un rectangle*

Soit $P \subset \mathbb{Z}^2$ un ensemble de n coordoonées. On veut répondre (rapidement) à des questions du type :

$\mathcal{Q}_r$: « Quel est la proportion de points de P qui se situent dans le rectangle $r = [a_1, b_1] \times [a_2, b_2]$? »

Pour n'importe quel rectangle r, on note $\mathbf{r}[P] = \frac{|P \cap r|}{n}$ la proportion recherchée.
Pour estimer $\mathbf{r}[P]$ de façon efficace, on considère $S \subseteq P$ un sous-ensemble de taille m de P (choisi aléatoirement) et on renvoie $\mathbf{r}[S] = \frac{|S \cap r|}{m}$.
On dit que S est une *ε-approximation* si pour tout r, on a $|\mathbf{r}[P] - \mathbf{r}[S]| \leq \varepsilon$.

1. Avec quelle taille m obtient-on une ε-approximation avec une probabilité $1 - \delta$?

   Indication : on peut considérer un ensemble S dont l'espérance de la taille est *m, plutôt que de taille exactement m.*

Exercice 3. *Graphe aléatoire bipartite*

Soit $0 < p < 1$ et $n \in \mathbb{N}^*$. On définit un graphe aléatoire non orienté $H_{2n,p}$ de la manière suivante : on se donne une famille $\{X_{i,j} \mid 1 \leq i \leq n,\ n+1 \leq j \leq 2n\}$ 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

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

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 $c > 0$.

i. Montrer que si $c > 1$, alors :

$$\lim_{n\to\infty} \mathbf{P}\left\{H_{2n,p} \text{ a un sommet isolé}\right\} = 0.$$

ii. Montrer que si $c < 1$, alors :

$$\lim_{n\to\infty} \mathbf{P}\left\{H_{2n,p} \text{ a un sommet isolé}\right\} = 1.$$

4. Dans cette question, on pose $p = 1/2$. Montrer qu'il existe une constante $C > 0$ telle que

$$\lim_{n\to\infty} \mathbf{P}\left\{\text{tous les sommets de } H_{2n,p} \text{ ont un degré inférieur à } \frac{n}{2} + C\sqrt{n\log n}\right\} =$$

Exercice 4. *Arrondi*

Soit U un ensemble à n éléments. On appelle recouvrement de U un ensemble $\mathcal{S} = (S_1, \cdots, S_m)$ de parties de U qui vérifie $\bigcup S_i = U$. Étant donné $\mathcal{S}$ un recouvrement de U, on note $\mathrm{OPT}(\mathcal{S})$ le cardinal minimal d'un sous-ensemble de $\mathcal{S}$ qui est encore un recouvrement de U.

1. Expliquer rapidement pourquoi $\mathrm{OPT}(\mathcal{S})$ est la solution du problème d'optimisation suivant :

$$\text{Minimiser } \sum_{i=1}^{m} x_i \text{ sous les contraintes } x_i \in \{0,1\} \text{ et } \sum_{i=1}^{m} x_i \mathbf{1}_{S_i} \geq \mathbf{1}_U \tag{1}$$

On considère le problème suivant qui est une relaxation de (1) :

$$\text{Minimiser } \sum_{i=1}^{m} z_i \text{ sous les contraintes } z_i \in [0,1] \text{ et } \sum_{i=1}^{m} z_i \mathbf{1}_{S_i} \geq \mathbf{1}_U \tag{2}$$

Alors que le problème (1) est NP-difficile, le problème (2) peut être résolu en temps polynomial par les méthodes de programmation linéaire.

2. Soit k un entier, et $(z_i)_{1\leq i\leq m}$ qui minimisent (2). Soient $(X_{i,j})_{1\leq i\leq m, 1\leq j\leq k}$ des variables aléatoires indépendantes vérifiant $\mathbf{P}\left\{X_{i,j} = 1\right\} = z_i$, $\mathbf{P}\left\{X_{i,j} = 0\right\} = 1 - z_i$. On définit un sous-ensemble $\mathcal{T} \subset \mathcal{S}$ par la condition :

$$S_i \in \mathcal{T} \iff \exists j \in \{1, \ldots, k\} \,:\, X_{i,j} = 1\,.$$

Montrer que :

$$\mathbf{E}\left[\#\mathcal{T}\right] \leq k\,\mathrm{OPT}(\mathcal{S})\,.$$

3. Déterminer une valeur de $c > 0$ telle que, si on pose $k = \lfloor c \log n \rfloor$, on ait :

$$\mathbf{P}\left\{\mathcal{T} \text{ est un recouvrement de } U\right\} \geq 1 - \frac{1}{n}\,.$$