Examen partiel
2 avril 2025

On pourra utiliser les variantes suivantes des inégalités de Chernoff.

(I) Si X est la somme de n variables aléatoires i.i.d. de loi de Bernoulli de paramètre $1/2$, alors pour tout $a > 0$

$$\mathbf{P}(X \geqslant \frac{n}{2} + a) \leqslant \exp(-2a^2/n)$$
$$\mathbf{P}(X \leqslant \frac{n}{2} - a) \leqslant \exp(-2a^2/n)$$

(II) Si X est une somme de variables aléatoires indépendantes ayant chacune une loi de Bernoulli, alors en posant $\mu = \mathbf{E}[X]$, pour tout $0 < \varepsilon < 1$,

$$\mathbf{P}(|X - \mu| \geqslant \varepsilon\mu) \leqslant 2\exp(-\mu\varepsilon^2/3)$$

Exercice 1 Moments et déviations

1. Soit X une variable aléatoire d'espérance nulle. On suppose que $0 < \mathbf{E}[X^k] < \infty$ pour un entier pair k. Montrer que pour tout réel $t > 0$,

$$\mathbf{P}\left(|X| > t\right) \leqslant \frac{\mathbf{E}[X^k]}{t^k} .$$

On a $\mathbf{P}\left(|X| > t\right) = \mathbf{P}(X^k > t^k) \leqslant \frac{\mathbf{E}[X^k]}{t^k}$ par l'inégalité de Markov.

2. Soit (Y_i) une suite de variables aléatoires i.i.d. vérifiant $\mathbf{E}[Y_i] = 0$, $\mathbf{E}[Y_i^2] = 1$ et $\mathbf{E}[Y_i^4] \leqslant 3$. On pose $S_n = Y_1 + \cdots + Y_n$. Montrer que pour tout entier $n \geqslant 1$ et tout réel $s > 0$,

$$\mathbf{P}\left(|S_n| \geqslant s\sqrt{n}\right) \leqslant \frac{3}{s^4}$$

On a $\mathbf{E}[S_n^4] = \sum_{i,j,k,l} \mathbf{E}[Y_iY_jY_kY_l]$. Par indépendance, puisque $\mathbf{E}[Y_i] = 0$, on a $\mathbf{E}[Y_iY_jY_kY_l] = 0$ dès lors qu'un entier apparaît une seule fois parmi $\{i,j,k,l\}$. Il s'ensuit que

$$\mathbf{E}[S_n^4] = \sum_i \mathbf{E}[Y_i^4] + 3\sum_{i<j} \mathbf{E}[Y_i^2Y_j^2] \leqslant 3n + 3n(n-1) = 3n^2$$

et donc d'après la question 1

$$\mathbf{P}(|S_n| \geqslant s\sqrt{n}) \leqslant \frac{\mathbf{E}[X^4]}{(s\sqrt{n})^4} \leqslant \frac{3n^2}{s^4n^2} = \frac{3}{s^4}.$$

Exercice 2 Urnes et boules

Considérons une urne contenant initialement 2 boules : une rouge et une verte. On répète la procédure suivante jusqu'à ce que l'urne contienne n boules : à chaque étape, on tire une boule depuis l'urne, uniformément au hasard, et on la remet dedans ainsi qu'une nouvelle boule de la même couleur. Montrer que le nombre final de boules rouges suit la loi uniforme sur $\{1, \ldots, n-1\}$.

Soit X_k la v.a. définie comme le nombre de boules rouges après l'ajout de k boules. On a donc $X_0 = 1$ presque sûrement ; on va montrer par récurrence sur k que X_k suit la loi uniforme sur $\{1, \ldots, k+1\}$. Pour $1 \leqslant j \leqslant k+2$, l'événement $\{X_{k+1} = j\}$ est la réunion des événements $\{X_{k+1} = j \text{ et } X_k = j\}$ et $\{X_{k+1} = j \text{ et } X_k = j-1\}$, donc

$$\begin{aligned}\mathbf{P}(X_{k+1} = j) &= \mathbf{P}(X_{k+1} = j | X_k = j)\mathbf{P}(X_k = j) + \mathbf{P}(X_{k+1} = j | X_k = j-1)\mathbf{P}(X_k = j-1) \\ &= \frac{k+2-j}{k+2}\frac{1}{k+1} + \frac{j-1}{k+2}\frac{1}{k+1} \\ &= \frac{1}{k+2}\end{aligned}$$

d'où le résultat.

Exercice 3 Coupe d'un graphe aléatoire en deux parties égales

Soit n un entier pair et soit $G = (V, E)$ un graphe (non orienté, sans boucle ni arête multiple) aléatoire de loi $\mathsf{G}(n, 1/2)$, c'est-à-dire que les événements $(A_{ij})_{1 \leqslant i < j \leqslant n}$ définis par

$$A_{ij} = \text{« l'arête } \{i, j\} \text{ est dans } E \text{ »}$$

sont indépendants et de probabilité $1/2$.

1. Soit $W \subset V$ un sous-ensemble de cardinal $n/2$. On note X_W le nombre d'arêtes reliant un élément de W à un élément de $V \setminus W$. Déterminer la loi et l'espérance de la variable aléatoire X_W. On a $X_W = \sum_{x \in W, y \in V \setminus W} \mathbf{1}_{xy}$ donc X_W suit la loi binomiale $\mathsf{B}(n^2/4, 1/2)$. En particulier, on a $\mathbf{E}[X_W] = n^2/8$.
2. Montrer que pour tout $\delta > 0$,

$$\lim_{n \to \infty} \mathbf{P}\left(X_W \in \left[(1-\delta)\frac{n^2}{8}, (1+\delta)\frac{n^2}{8}\right] \text{ pour tout } W \subset V \text{ de cardinal } n/2\right) = 1.$$

Par l'inégalité de Chernoff I, pour tout W de cardinal $n/2$,

$$\begin{aligned}\mathbf{P}(|X_W - n^2/8| > \delta n^2/8) &= \mathbf{P}(X_W > (1+\delta)n^2/8) + \mathbf{P}(X_W < (1-\delta)n^2/8) \\ &\leqslant 2\exp(-2(\delta n^2/8)^2/(n^2/4))) \\ &= 2\exp\left(-\frac{\delta^2 n^2}{8}\right)\end{aligned}$$

Puisque $\binom{n}{n/2} \leqslant 2^n$, par la borne de l'union,

$$\mathbf{P}(\exists W \,:\, |X_W - n^2/8| > \delta n^2/8) \leqslant 2\exp(-\delta^2 n^2/8)\binom{n}{n/2} \leqslant \exp(-\delta^2 n^2/8) \cdot 2^n$$

qui tend vers 0 quand n tend vers $+\infty$.

Exercice 4 Version récursive de l'algorithme de Karger

1. Soit T un arbre binaire complet de hauteur h (c'est-à-dire ayant 2^h feuilles, voir Figure 1). On colorie chacun des sommets indépendemment en rouge ou en bleu avec probabilité $1/2$. On note $P(h)$ la probabilité qu'il existe un chemin de la racine à une feuille dont toutes les sommets soient rouges (en particulier, on a $P(0) = \frac{1}{2}$).
   (a) Donner une relation de récurrence reliant $P(h+1)$ à $P(h)$. Soit T un arbre binaire complet de hauteur $h+1$ enraciné en $\emptyset$, et soient T_0 et T_1 les sous-arbres binaires complets de hauteur h enracinés en les enfants (notés 0 et 1) de la racine. Considérons les événements $A =$«il existe un chemin rouge dans T», $B =$«le sommet $\emptyset$ est rouge» et $C_i =$«il existe un chemin rouge dans T_i» pour $i = 1$ ou 2. Alors $A = B \cap (C_0 \cup C_1)$. Par le lemme des coalitions, les événements B, C_0 et C_1 sont indépendants et donc

   $$\mathbf{P}(A) = \mathbf{P}(B)\mathbf{P}(C_0 \cup C_1) = \mathbf{P}(B)(1 - \mathbf{P}(\overline{C_0})\mathbf{P}(\overline{C_1})) = \mathbf{P}(B)(\mathbf{P}(C_0) + \mathbf{P}(C_1) - \mathbf{P}(C_0)\mathbf{P}(C_1))$$

   et donc $P(h+1) = P(h) - \frac{1}{2}P(h)^2$
   (b) En déduire que $P(h) \geqslant \frac{1}{h+2}$ pour tout $h \geqslant 0$. La fonction $x \mapsto x - \frac{1}{2}x^2$ est croissante sur $[0, 1]$, donc par récurrence : si $P(h) \geqslant \frac{1}{h+2}$ alors

   $$P(h+1) \geqslant \frac{1}{h+2} - \frac{1}{2(h+2)^2} = \frac{2h+3}{2(h+2)^2} \geqslant \frac{1}{h+3},$$

   la dernière inégalité étant justifiée car $\frac{2h+3}{2(h+2)^2} - \frac{1}{h+3} = \frac{h+1}{2(h+2)^2(h+3)} \geqslant 0$.

Dans cet exercice, on considère des graphes non orientés, sans boucle, ayant éventuellement des arêtes multiples. Une *coupe* d'un graphe $G = (V, E)$ est un sous-ensemble $C \subset E$ tel que le graphe $(V, E \setminus C)$ n'est pas connexe. On note $\mathrm{mincut}(G)$ le cardinal minimal d'une coupe de G. On décrit une variante récursive de l'algorithme de Karger pour calculer $\mathrm{mincut}(G)$.

Étant donnée une arête $e = \{x, y\} \in E$, la contraction de G selon e, notée G/e, est le graphe obtenu en identifiant les sommets x et y (pour obtenir un nouveau sommet noté xy), en remplaçant les arêtes $\{x, z\}$ ou $\{y, z\}$ par $\{xy, z\}$ et en effaçant les boucles éventuellement créées.

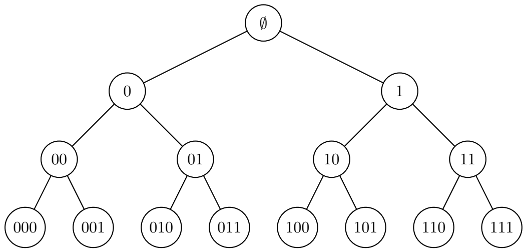

FIGURE 1 – Un arbre binaire de hauteur 3

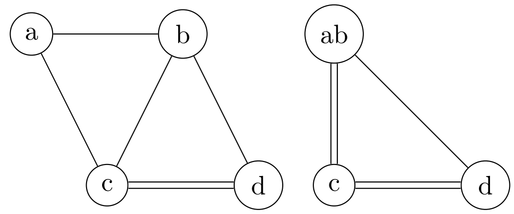

FIGURE 2 – Un graphe G (à gauche) et sa contraction G/e pour $e = \{a, b\}$ (à droite)

2. On pose $n = |V|$ et on fixe une coupe C_0 telle que $|C_0| = \mathrm{mincut}(G)$. On considère une suite de graphes définie par $G_1 = G$ et (pour $1 \leqslant i \leqslant n-2$) $G_{i+1} = G_i/e_i$ où l'arête e_i est choisie uniformément parmi les arêtes de G_i, tous les choix étant indépendants. Soit A_i l'événement «l'arête e_i choisie à la $i$ème étape est dans C_0» et $B_i = \overline{A_i}$.
   (a) Montrer que $\mathbf{P}(A_1) \leqslant \frac{2}{n}$, Soit $k = \mathrm{mincut}(G)$. Alors $\mathbf{P}(A_1) = \frac{k}{|E|}$. Comme tout sommet a degré $\geqslant k$, on a $2|E| \geqslant kn$ et donc $\mathbf{P}(A_1) \leqslant \frac{2}{n}$.
   (b) Montrer que pour tout $1 \leqslant i \leqslant n-2$

$$\mathbf{P}(B_1 \cap B_2 \cap \cdots \cap B_i) \geqslant \frac{(n-i)(n-i-1)}{n(n-1)},$$

   Si l'événement B_1 est vérifié, alors le graphe G_2 est un graphe de coupe minimale k, et une coupe de taille k est donnée par les arêtes provenant de C_0. On a donc, comme à la question (a), $\mathbf{P}(A_2|B_1) \leqslant \frac{2}{n-1}$. On calcule donc

$$\begin{aligned}\mathbf{P}(B_1 \cap B_2 \cap \cdots \cap B_i) &= \mathbf{P}(B_1)\mathbf{P}(B_2|B_1)\mathbf{P}(B_3|B_1 \cap B_2) \cdots \mathbf{P}(B_i|B_1 \cap \cdots \cap B_{i-1}) \\ &\geqslant \frac{n-2}{n} \cdot \frac{n-3}{n-1} \cdot \frac{n-4}{n-2} \cdots \frac{n-i-1}{n-i+1} \\ &= \frac{(n-i)(n-i-1)}{n(n-1)}\end{aligned}$$

   (c) En déduire que si $i \leqslant n - \frac{n}{\sqrt{2}}$, alors $\mathbf{P}(\mathrm{mincut}(G_{i+1}) = \mathrm{mincut}(G)) \geqslant \frac{1}{2}$. Si $B_1 \cap \cdots \cap B_i$ est réalisé, alors $\mathrm{mincut}(G_{i+1}) = \mathrm{mincut}(G)$. On conclut en remarquant que si $i \leqslant n - \frac{n}{\sqrt{2}}$ alors

$$\frac{(n-i)(n-i-1)}{n(n-1)} \geqslant \frac{\frac{n}{\sqrt{2}}\left(\frac{n}{\sqrt{2}}-1\right)}{n(n-1)} \geqslant \frac{1}{2}$$

   Oups, la dernière inégalité va dans le mauvais sens, il y avait une petite erreur dans l'énoncé. Il fallait montrer l'inégalité sous l'hypothèse que $i \leqslant n - \frac{n}{\sqrt{2}} - 1$ et non $i \leqslant n - \frac{n}{\sqrt{2}}$

Considérons l'algorithme suivant

Algorithm 1 KargerStein(G) où G est un graphe à n sommets

Si $n < 6$, calculer $\mathrm{mincut}(G)$ par recherche exhaustive.
Si $n \geqslant 6$,
  pour i de 1 à 2,
    Contracter aléatoirement des arêtes jusqu'à obtenir un graphe G^i ayant $\leqslant \frac{n}{\sqrt{2}} + 1$ sommets
    puis calculer récursivement $X_i = \mathrm{KargerStein}(G^i)$.
  Renvoyer $\min(X_1, X_2)$.

3. Montrer que cet algorithme calcule la valeur correcte de $\mathrm{mincut}(G)$ avec probabilité $\Omega(\frac{1}{\log n})$. Soit $h = O(\log n)$ la profondeur de la récursion. On considère un arbre binaire complet T de hauteur h dont les sommets sont les appels à l'algorithme. Comme dans la question 2, on fixe une coupe minimale C_0 et on colorie un sommet v de T en rouge quand les arêtes contractées par l'appel

v ne sont pas dans V, et en bleu sinon. Par la question 2, chaque sommet est colorié en rouge avec probabilité $\geqslant 1/2$, indépendamment. S'il existe un chemin rouge de la racine à une feuille de T, alors l'algorithme est correct. Par la question 1, la probabilité que cela se produise est $\geqslant P(h) = \frac{1}{O(\log n)}$

4. Donner un algorithme probabiliste calculant $\mathrm{mincut}(G)$ avec probabilité d'erreur $\leqslant 2^{-100}$. Quelle est sa complexité ? On admettra que l'opération de contraction d'une arête aléatoire dans un graphe à n sommets peut s'effectuer en temps $O(n)$. L'algorithme de la question précédente est correct avec probabilité $\geqslant 1 - \frac{1}{c \log n}$. Répétons N fois cet algorithme en renvoyant le minimum de toutes les valeurs trouvées. La probabilité d'erreur est alors $\leqslant \left(1 - \frac{1}{c \log n}\right)^N \leqslant \exp(-\frac{N}{c \log n})$ qui est 2^{-100} pour $N = 100c \log_2 n$.

   Soit $C(n)$ la complexité de l'algorithme de la question 3. On a

$$C(n) = 2C\left(\left\lfloor \frac{n}{\sqrt{2}} + 1 \right\rfloor\right) + O(n^2)$$

   donc par le master theorem, $C(n) = O(n^2 \log n)$. L'algorithme de la question 4 a donc une complexité $O(n^2 \log^2 n)$.