

TD 07 – Graphes aléatoires

Exercice 1.*Grphe 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 $\{X_{i,j} : 1 \leq i \leq n, n+1 \leq j \leq 2n\}$ de v.a. i.i.d. de loi de Bernoulli de paramètre p . On pose alors $H_{2n,p} = (V, E)$, avec $V = \{1, \dots, 2n\}$ et

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

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 $c > 0$.
 1. Montrer que si $c > 1$, alors

$$\lim_{n \rightarrow \infty} \mathbf{P}(H_{2n,p} \text{ a un sommet isolé}) = 0.$$

2. Montrer que si $c < 1$, alors

$$\lim_{n \rightarrow \infty} \mathbf{P}(H_{2n,p} \text{ a un sommet isolé}) = 1.$$

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

$$\lim_{n \rightarrow \infty} \mathbf{P} \left(\text{tous les sommets de } H_{2n,p} \text{ ont un degré inférieur à } \frac{n}{2} + C\sqrt{n \log n} \right) = 1.$$

Exercice 2.*K4*

Soit G un graphe aléatoire de loi $G_{n,p}$. L'objectif de cet exercice est de montrer qu'il y a un seuil $p_0 := n^{-2/3}$ tel que pour $p = o(p_0)$, le graphe G n'a pas de clique de taille 4 avec bonne probabilité, et que pour $p = \omega(p_0)$, le graphe G a au moins une clique de taille 4 avec bonne probabilité.

Rappels / définitions :

- Un graphe aléatoire G suit la loi $G_{n,p}$ s'il a n sommets et que chaque arête est présente dans G avec probabilité p ;
- une clique de taille 4 est un ensemble de 4 sommets tous reliés deux à deux par des arêtes ;
- $p = o(p_0)$ signifie $\frac{p}{p_0} \rightarrow 0$ quand $n \rightarrow +\infty$;
- $p = \omega(p_0)$ signifie $\frac{p_0}{p} \rightarrow 0$ quand $n \rightarrow +\infty$.

1. Pour p quelconque, calculer $\mathbf{E}[X]$, où X est le nombre de cliques du graphe G .
 2. Soit $p = o(p_0)$, montrer que $\Pr(X \neq 0) \rightarrow 0$ quand n tend vers l'infini.
- On suppose maintenant $p = \omega(p_0)$, et on veut montrer que $\mathbf{P}\{X = 0\} \rightarrow 0$ quand n tend vers l'infini.

3. Montrer que $\mathbf{P}\{X = 0\} \leq \frac{\mathbf{Var}[X]}{\mathbf{E}[X]^2}$. Il suffira donc de montrer que $\frac{\mathbf{Var}[X]}{\mathbf{E}[X]^2} \rightarrow 0$.

4. Soit X_i des variables aléatoires à valeur dans $0, 1$ (et non indépendantes). Montrer que

$$\mathbf{Var} \left[\sum_i X_i \right] \leq \mathbf{E} \left[\sum_i X_i \right] + \sum_{i \neq j} \mathbf{E} [(X_i - \mathbf{E}[X_i])(X_j - \mathbf{E}[X_j])].$$

5. En déduire que $\mathbf{Var}[X] = o(\mathbf{E}[X]^2)$ et conclure.

Exercice 3.

Clique Number

Le clique number $\omega(G)$ d'un graphe G est l'entier k maximal tel que G contient une clique de taille k (c'est-à-dire un ensemble de k sommets tous reliés deux à deux par des arêtes). Soit G un graphe aléatoire de loi $G_{n,1/2}$ (i.e. G a n sommets et chaque arête de G est présente avec probabilité $1/2$). L'objectif de cet exercice est de montrer que pour tout $\varepsilon > 0$,

$$\lim_{n \rightarrow +\infty} \mathbf{P} \{ (2 - \varepsilon) \log n \leq \omega(G) \leq (2 + \varepsilon) \log n \} = 1.$$

Asymptotiquement, $\omega(G)$ est donc de l'ordre de $2 \log n$.

Remarque : La méthode utilisée pour montrer ce résultat ressemble beaucoup à celle utilisée dans l'exercice K4.

1. Pour un entier k quelconque, on définit X_k la variable aléatoire comptant le nombre de cliques de taille k dans G . Calculer $\mathbf{E}[X_k]$
2. Soit $\varepsilon > 0$. Montrer que $\lim_{n \rightarrow +\infty} \mathbf{P} \{ \omega(G) \geq (2 + \varepsilon) \log n \} = 0$ (pour simplifier, on pourra supposer que $k = (2 + \varepsilon) \log n$ est un entier).
3. On considère maintenant $k = (2 - \varepsilon) \log n$ (encore une fois, on le suppose entier). Montrer que $\lim_{n \rightarrow +\infty} \mathbf{P} \{ \omega(G) \leq (2 - \varepsilon) \log n \} = 0$. *Indice : on pourra utiliser les questions 3 et 4 de l'exercice K4.*

Exercice 4.

Approximation de Poisson

On se place dans le modèle *Balls and Bins* où l'on jette m balles au hasard dans n paniers. Le problème est que les v.a. X_i représentant le nombre de balles dans le i -ème panier ne sont pas indépendantes (intuitivement, car $X_1 + \dots + X_n = m$). On voudrait approximer le modèle *Balls and Bins* par le modèle *Approximation de Poisson*, dans lequel Y_1, \dots, Y_n sont des variables aléatoires indépendantes qui suivent chacune une loi de Poisson de moyenne $\mu = m/n$ (la variable Y_i est donc pensée pour être la version simplifiée de X_i).

1. Montrer que $Y = \sum_{i=1}^n Y_i$ suit une loi de Poisson dont on précisera le paramètre.
2. Montrer que la distribution de (Y_1, \dots, Y_n) conditionnée au fait que $Y = m$ est la même que la distribution de (X_1, \dots, X_n) .

Note : on peut en fait obtenir un résultat légèrement plus général. Si (X_1, \dots, X_n) représente la charge de n paniers après avoir lancé au hasard k balles, et que les Y_i sont n v.a. indépendantes suivant chacune une loi de Poisson de paramètre m/n , alors la distribution de (Y_1, \dots, Y_n) conditionnée au fait que $Y = k$ est la même que la distribution de (X_1, \dots, X_n) , indépendamment de la valeur de m .

3. Soit f une fonction sur n variables, à valeurs réelles positives ou nulles. Prouver que

$$\mathbf{E}[f(X_1, \dots, X_n)] \leq e\sqrt{m} \mathbf{E}[f(Y_1, \dots, Y_n)].$$

On pourra prouver comme étape intermédiaire que $m! < e\sqrt{m} \left(\frac{m}{e}\right)^m$.

4. En déduire le corollaire suivant : soit \mathcal{E} un événement qui dépend de la charge des paniers. Supposons que \mathcal{E} arrive avec probabilité p dans l'*Approximation de Poisson*, c'est-à-dire si la charge des paniers est (Y_1, \dots, Y_n) . Alors \mathcal{E} arrive avec probabilité au plus $pe\sqrt{m}$ dans le modèle *Balls and Bins*, c'est-à-dire si la charge des paniers est (X_1, \dots, X_n) .
5. *Application :* On jette n balles dans n paniers selon le modèle *Balls and Bins*. Montrer qu'avec probabilité au moins $1 - 1/n$ (pour n assez grand), la charge maximale est $\geq \ln n / \ln \ln n$.