ENS de Lyon
Master 1 Mathématiques 2016-2017

Graphes aléatoires
E. Baur & M. Maazoun

Corrigé TD 6 : Vacances

1 Exercice 1

Soit Z le nombre de composantes unicycliques.

$$\begin{aligned}\mathbb{E}[Z] &= \sum_{H \subset K_n \text{ unicyclique}} \mathbb{P}(H \text{ composante}) \\ &= \sum_{k=3}^{n} \#\{H \subset [n] \text{ unicyclique}, |H| = k\} p^k (1-p)^{\binom{k}{2} - k + (n-k)k}.\end{aligned}$$

On majore $\#\{H \subset [n] \text{ unicyclique}, |H| = k\}$ par $\binom{n}{k} k^{k-2} \binom{k}{2}$. On minore $\binom{k}{2} - k + (n-k)k$ par $nk/4$. On obtient

$$\begin{aligned}\mathbb{E}[Z] &\leq \sum_{k=3}^{n} \binom{n}{k} k^{k-2} \binom{k}{2} p^k (1-p)^{nk/4} \\ &\leq \sum_{k=3}^{n} \frac{n^k e^k}{k^k} k^{k-2} k^2 p^k \exp(\log(1-p) nk/4) \\ &\leq \sum_{k=3}^{n} (npe)^k \exp(-pnk/4) = \sum_{k=3}^{n} (npe \exp(-np/4))^k \leq \frac{(npe \exp(-np/4))^3}{1 - npe \exp(-np/4)}.\end{aligned}$$

Dans la deuxième inégalité on a utilisé $\binom{n}{k} \leq (\frac{ne}{k})^k$, dans la troisième on a utilisé $\log(1+x) \leq x$. Maintenant, vu que $np \to \infty$, $npe \exp(-np/4) = o(1)$. Donc $\mathbb{E}[Z] = \frac{o(1)}{1-o(1)} = o(1)$.

2 Exercice 2

$F_r = \Delta(\Delta - 1) \cdots (\Delta - r)$ est le nombre de suites de r triangles distincts (mais pas disjoints !),

$$\mathbb{E}[F_r] = \sum_{\substack{T_1, \ldots, T_r, \\ T_1 \neq \cdots \neq T_r}} \mathbb{P}(T_1, T_2, \ldots T_r \in G_{n,p}).$$

On veut montrer, comme souvent dans une méthode du second moment, que le terme qui fait intervenir des arêtes disjointes (le terme en p^{3r}) domine les autres. Soit N_j le nombre de suites de r triangles distincts dans $[n]$ avec $3r-j$ arêtes disjointes en total. On obtient

$$\mathbb{E}[F_r] = \sum_{j=0}^{3r} N_j p^{3r-j}.$$

On va montrer les estimées suivantes :

$$N_0 \sim \frac{n^{3r}}{6^r}, \tag{1}$$

$$\forall j \geq 1, \quad N_j = O(n^{3r-j-1}). \tag{2}$$

Ces estimées impliquent que $\mathbb{E}[F_r] \to \left(\frac{c^3}{6}\right)^r$, ce qui prouve alors que Δ converge en loi vers une variable de loi $\mathrm{Po}(c^3/6)$.

Reste à prouver les estimées (1) et (2). N_0 est majoré par le nombre de suites quelconques de r triangles dans $[n]$, qui est $\binom{n}{3}^r \leq \frac{n^{3r}}{6^r}$. Par ailleurs N_0 est minoré par le nombre de suites de r triangles qui occupent des sommets distincts, qui est $\frac{n(n-1)\cdots(n-3r+1)}{6} \sim \frac{n^{3r}}{6}$. Cet encadrement prouve (1).

Pour N_j quand $j \geq 1$, considérons une suite de r triangles distincts où j arêtes sont superposés. L'union de tous ces triangles forme un sous-graphe H à $3r-j$ arêtes, cherchons à majorer son nombre de sommets. Chaque composante connexe $\mathcal{C}$ de H contient au moins un cycle (parce qu'on considère une union de triangle). On a donc $|V(\mathcal{C})| \leq |E(\mathcal{C})|$. Par ailleurs, il existe forcément dans notre union deux triangles distincts non disjoints, qui donnent lieu à une composante à au moins deux cycles, pour laquelle $|V(\mathcal{C})| \leq |E(\mathcal{C})| - 1$. En sommant sur toutes les composantes, on obtient

$$|V(H)| \leq |E(H)| - 1 = 3r - j - 1.$$

On déduit que N_r est majoré par le nombre de choix de $3r-j-1$ sommets dans $[n]$ puis de r triangles parmi ces sommets :

$$N_r \leq \binom{n}{3r-j-1}\binom{3r-j-1}{3}^r = O(n^{3r-j-1}).$$

Ceci prouve l'estimée (2).

3 Exercice 3

Soit $p = c/n$, $c > 1$. Soit $L_1 \geq L_2 \geq \cdots \geq L_n$ l'énumération des tailles des composantes connexes de $G_{n,p}$ (on pose $L_i = 0$ pour $j < i \leq n$ si $G_{n,p}$ a exactement j composantes connexes). Soient a et b deux sommets choisis uniformément et independemment dans $G_{n,p}$. Nous écrivons $a \leftrightarrow b$ si il existe un chemin entre a et b (de manière équivalente si a et b sont dans la même composante connexe).

1. Soient $C_1, \ldots, C_n$ les composantes connexes triées par taille décroissante.

$$\mathbb{P}\left(a \leftrightarrow b \mid L_1, L_2, \ldots\right) = \sum_{i=1}^{n} \mathbb{P}\left(a, b \in C_i \mid L_1, L_2, \ldots\right).$$

   Pour calculer cette dernière probabilité, prenons d'abord un conditionnement plus fort : on conditionne par tout $G_{n,p}$.

$$\begin{aligned}\mathbb{P}\left(a, b \in C_i \mid G_{n,p}\right) &= \sum_{j=1}^{n}\sum_{k=1}^{n} \mathbb{E}[\mathbb{1}_{a=j}\,\mathbb{1}_{b=k}\,\mathbb{1}_{j,k\in C_i}\,|G_{n,p}] \\ &= \sum_{j=1}^{n}\sum_{k=1}^{n} \mathbb{E}[\mathbb{1}_{a=j}\,\mathbb{1}_{b=k}]\,\mathbb{1}_{j,k\in C_i} \\ &= \sum_{j=1}^{n}\sum_{k=1}^{n} \frac{1}{n^2}\,\mathbb{1}_{j,k\in C_i} = \frac{L_i^2}{n^2}.\end{aligned}$$

   On reconditionne par $L_1, \ldots, L_n$ pour obtenir l'égalité demandée.
2. La convergence en probabilité de variables bornées implique la convergence de tous les moments. Donc comme $L_1/n \to (1 - x/c)^2$, le premier terme converge vers $(1-x/c)^2$. On majore le deuxième terme par $\mathbb{P}(L_2 > K \log n)$, qui tend vers zéro par le théorème 7. Le troisième terme est borné lui par $\sum_{i=2}^{n} \frac{\log(n)^2}{n} = O(\log(n)^2/n) = o(1)$. On obtient bien la convergence demandée.

4 Exercice 4

Soit Z le nombre de chemins de taille $k = \lfloor \beta \log(n) \rfloor$. On a

$$\mathbb{E}[Z] = \frac{1}{2}n(n-1)\cdots(n-k)p^k \sim \frac{1}{2}n^{k+1}\left(\frac{c}{n}\right)^k = \frac{1}{2}nc^k.$$

On déduit que pour $\beta > 1/\log(1/c)$, $\mathbb{E}[Z] \to 0$, alors que pour $\beta < 1/\log(1/c)$, $\mathbb{E}[Z] \to \infty$. Calculons maintenant le second moment. On commence à avoir l'habitude :

$$\mathbb{E}[Z^2] = \sum_{j=0}^{k} N_j p^{2k-j},$$

où N_j est le nombre de paires de chemins de longueur k dans K_n qui ont j arêtes en commun. N_0 est majoré par $n^{2k+2}/4$, et minoré par $n(n-1)\cdots(n-2k-1)/4 \sim n^{2k+2}/4$. On en déduit un équivalent pour le premier terme $N_0 p^{2k} \sim \frac{1}{4}n^2c^2k \sim \mathbb{E}[Z]^2$. Donc

$$\begin{aligned}\mathbb{E}[Z^2] - \mathbb{E}[Z]^2(1+o(1)) = \sum_{j=1}^{k} N_j p^{2k-j} &\leq \text{cte} \sum_{j=1}^{k} n^{2k+1-j}k^4\left(\frac{c}{n}\right)^{2k-j} \\ &\leq \text{cte} \times k \times nk^4c^k = O(k^5nc^k) = o((nc^k)^2),\end{aligned}$$

ce qui permet d'appliquer la méthode du second moment. On a utilisé l'inégalité donnée en indication : pour $j \geq 1$, $N_j \leq \text{cte} \times n^{2k+1-j}k^4$.

Prouvons maintenant cette inégalité. Considérons deux chemins, γ_1 et γ_2, de longueur k, partageant j arêtes en commun. Fixons un sens de parcours pour chacun (cela multiplie par 4 le nombre de chemins). L'union des deux chemins forme un graphe connexe H a $2k - j$ arêtes. Le nombre de sommets de H est $2k + 1 - j - l$, où l est le nombre de cycles distincts (en fait disjoints vu la structure) de H. Écrivons $x_1, \ldots, x_{k+1}$ les sommets de γ_1 dans leur ordre de parcours, et $y_1, \ldots, y_{k-j-l}$ les sommets de $\gamma_2 \setminus \gamma_1$, dans leur ordre de parcours.

On constate que le nombre de composantes connexes de $\gamma_2 \cap \gamma_1$ vaut $l + 1$. On voit ces composantes connexes $\sigma_1, \ldots, \sigma_{l+1}$ comme des sous-chemins ordonnés (éventuellement de longueur 0) de γ_1, d'après leur ordre de parcours par γ_2. Le chemin γ_2 alterne entre les sommets $y_1, \ldots, y_{k-j-l}$ et les sous-chemins $\sigma_1, \ldots, \sigma_{l+1}$. On déduit de tout ça que

$$N_j \leq 4 \sum_{l=0}^{2k-j} n^{2k+1-j-l}(k + 1 - j - l)^{l+1}((k + 1)^2)^{l+1}.$$

Le premier facteur compte les choix pour l'ensemble de sommets $x_0, \ldots, y_{k-j-l}$, le second compte les choix des interstices entre les y_i où caser les chemins $\sigma_1, \ldots, \sigma_{l+1}$, et le dernier compte les choix des sous-chemins $\sigma_1, \ldots, \sigma_{l+1}$ (un sous-chemin est caractérisé par son début et sa fin).

Pour k assez grand tel que $k^3 \leq n$, on obtient $N_j \leq \text{cte} \times n^{2k+1-j}k^4$.