ENS de Lyon — Graphes aléatoires
Master 1 Mathématiques 2016-2017 — E. Baur & M. Maazoun

Corrigé TD 7 : Degrés v2

Exercice 1. *Degrés dans la composante géante*

On se place dans $G = G_{n,p}$ avec $p = c/n$ et $c > 1$. Pour $d \geq 1$ on veut calculer l'asymptotique du nombre Z_d de sommets de degré d dans la composante géante (que l'on note C_1). On pose $f(n) \to \infty$ arbitrairement lentement.

1. On considère le sommet 1, et on veut calculer la probabilité qu'il soit à la fois de degré d et hors de la composante géante. On dénote G' le graphe induit par G sur $[\![2, n]\!]$, et C'_1 sa composante géante. On dénote $\mathcal{V}_G(x)$ le voisinage d'un sommet x dans le graphe G, et $\mathcal{C}_G(x)$ sa composante.

   (a) Pour $x_1, \ldots, x_d$ des sommets disjoints de G', encadrer

   $$\mathbb{P}(\mathcal{V}_G(1) = \{x_1, \ldots, x_d\}, |\mathcal{C}_G(1)| \leq f(n)\log(n)).$$

   On considère l'inclusion d'évènements suivante :

   $$\begin{aligned}&\left\{\mathcal{V}_G(1) = \{x_1, \ldots, x_d\},\, \forall_{i=1}^d, |\mathcal{C}_{G'}(x_i)| \leq \tfrac{f(n)}{d}\log(n)\right\}\\ &\qquad\subset \{\mathcal{V}_G(1) = \{x_1, \ldots, x_d\}, |\mathcal{C}_G(1)| \leq f(n)\log(n))\}\\ &\qquad\qquad\subset \left\{\mathcal{V}_G(1) = \{x_1, \ldots, x_d\},\, \forall_{i=1}^d, |\mathcal{C}_{G'}(x_i)| \leq f(n)\log(n).\right\}\end{aligned}$$

   On obtient par indépendance

   $$\begin{aligned}&p^d(1-p)^{n-1-d}\,\mathbb{P}\left(\forall_{i=1}^d, |\mathcal{C}_{G'}(x_i)| \leq \tfrac{f(n)}{d}\log(n)\right)\\ &\qquad\leq \mathbb{P}\left(\mathcal{V}_G(1) = \{x_1, \ldots, x_d\}, |\mathcal{C}_G(1)| \leq f(n)\log(n)\right)\\ &\qquad\qquad\leq p^d(1-p)^{n-1-d}\,\mathbb{P}\left(\forall_{i=1}^d, |\mathcal{C}_{G'}(x_i)| \leq f(n)\log(n)\right).\end{aligned}$$

   (b) On estime les deux bornes grâce au Théorème 7 du cours.

   $$\mathbb{P}\left(\forall_{i=1}^d, |\mathcal{C}_{G'}(x_i)| \leq \tfrac{f(n)}{d}\log(n)\right) \geq \mathbb{P}\left(\forall_{i=1}^d, x_i \notin C'_1\right) - o(1)$$

   et

   $$\begin{aligned}\mathbb{P}\left(\forall_{i=1}^d, |\mathcal{C}_{G'}(x_i)| \leq f(n)\log(n)\right) &\leq \mathbb{P}\left(\forall_{i=1}^d, x_i \notin C'_1\right) + \mathbb{P}\left(|C'_1| \leq f(n)\log(n)\right)\\ &= \mathbb{P}\left(\forall_{i=1}^d, x_i \notin C'_1\right) + o(1).\end{aligned}$$

   Par ailleurs, les quantités considérées ne dépendent pas du choix de $x_1, \ldots, x_d$, donc quand on somme, les $o(1)$ sont uniformes, ce qui permet d'écrire

$$\begin{aligned}
&\mathbb{P}(\deg(1)=d, 1\notin C_1)\\
&=\mathbb{P}(\deg(1)=d, 1\notin C_1, |\mathcal{C}_G(1)|<f(n)\log(n))+o(1)\\
&=o(1)+\sum_{x_1,\dots,x_d}\mathbb{P}\left(\mathcal{V}_G(1)=\{x_1,\dots,x_d\}, |\mathcal{C}_G(1)|\le f(n)\log(n)\right)\\
&=o(1)+o(1)+e^{-c}\frac{c^d}{d!}\times\frac{1}{n^d}\sum_{2\le x_1,\dots,x_d\le n \text{ distincts}}\mathbb{P}(\forall_{i=1}^d, x_i\notin C_1').
\end{aligned}$$

(c) On a

$$\frac{1}{n^d}\sum_{2\le x_1,\dots,x_d\le n \text{ distincts}}\mathbb{P}(\forall_{i=1}^d, x_i\notin C_1')=\mathbb{E}\left[\left(\frac{|G'\setminus C_1'|}{n}\right)^d\right].$$

Par l'indication et le Théorème 7 du cours, on a bien la convergence vers $(x/c)^d$, où $x(c)$ désigne la solution dans $(0,1)$ de $xe^{-x}=ce^{-c}$.

2. On a alors

$$\begin{aligned}
\mathbb{E}[Z_d]&=n\left(\mathbb{P}(\deg(1)=d)-\mathbb{P}(\deg(1)=d, 1\notin C_1)\right)\\
&=n\left(e^{-c}\frac{c^d}{d!}+o(1)-\left(\frac{x}{c}\right)^d e^{-c}\frac{c^d}{d!}+o(1)\right)\sim ne^{-c}\frac{c^d-x^d}{d!}.
\end{aligned}$$

Exercice 2. *Méthode des moments factoriels pour des variables entières*

1. On considère

$$\left(\sum_{r=k}^{k+m}(-1)^{r-k}\frac{(x)_r}{(r-k)!k!}\right)-\mathbb{1}_{\{x=k\}}.$$

Si $x<k$, cette quantité vaut 0. Si $x=k$, on obtient $(-1)^0\frac{(k)_k}{k!}-1=0$. Si $x>k$, on calcule

$$\begin{aligned}
\sum_{r=k}^{k+m}(-1)^{r-k}\frac{(x)_r}{(r-k)!k!}&=\sum_{r=k}^{k+m}(-1)^{r-k}\binom{x}{r}\binom{r}{k}=\sum_{r=k}^{k+m}(-1)^{r-k}\binom{x}{k}\binom{x-k}{r-k}\\
&=\binom{x}{k}\sum_{l=0}^{m}(-1)^l\binom{x-k}{l}
\end{aligned}$$

qui est du signe de $(-1)^m$ d'après l'indication, et nul quand $m>x-k$. Donc la quantité considérée est bien (faiblement) du signe de $(-1)^m$.

On déduit que pour X une variable aléatoire entière positive,

$$S_m^{(k)}(X)=\sum_{r=k}^{k+m}(-1)^{r-k}\frac{\mathbb{E}[(X)_r]}{(r-k)!k!} \tag{1}$$

alterne autour de $\mathbb{P}(X=k)$.

2. On suppose maintenant que pour tout $r \geq 1$, $\mathbb{E}[(X_n)_r] \to \mathbb{E}[(X)_r]$ et que pour tout $k \geq 0$, $\mathbb{E}[(X)_r]/(r-k)!$ tend vers 0 quand $r \to \infty$. On fixe $k \geq 0$ et on veut montrer que $\mathbb{P}(X_n = k) \to \mathbb{P}(X = k)$. Soit $\epsilon > 0$.
   On commence par fixer m pair tel que $\frac{\mathbb{E}[(X)_{k+m}]}{(k+m-k)!k!} < \epsilon/2$. Alors
$$0 \leq S_m^{(k)}(X) - \mathbb{P}(X = k) \leq S_m^{(k)}(X) - S_{m-1}^{(k)}(X) < \epsilon/2.$$
$$0 \leq \mathbb{P}(X = k) - S_{m-1}^{(k)}(X) \leq S_m^{(k)}(X) - S_{m-1}^{(k)}(X) < \epsilon/2.$$
   Mais alors, il existe n_0 tel que pour $n \geq n_0$,
$$|S_m^{(k)}(X_n) - S_m^{(k)}(X)| \leq \epsilon/2 \quad ; \quad |S_{m-1}^{(k)}(X_n) - S_{m-1}^{(k)}(X)| \leq \epsilon/2.$$
   D'où $|\,\mathbb{P}(X = k) - S_m^{(k)}(X_n)| < \epsilon$, et $|\,\mathbb{P}(X = k) - S_{m-1}^{(k)}(X_n)| < \epsilon$ pour $n \geq n_0$. Alors en considérant à nouveau la série alternée (1), on obtient
$$\mathbb{P}(X = k) - \epsilon \leq S_{m-1}^{(k)}(X_n) \leq \mathbb{P}(X_n = k) \leq S_m^{(k)}(X_n) \leq \mathbb{P}(X = k) + \epsilon,$$
   d'où le résultat.
3. Il suffit pour déduire le Théorème 5 de vérifier que le k-ième moment factoriel de $\mathrm{Po}(\lambda)$ est λ^k, et de vérifier que pour tout k, $\lambda^r = o((r-k)!)$. C'est immédiat.

Exercice 3.

On a
$$\mathbb{P}(\sum_{d=0}^{\infty} |\mu_n(d) - e^{-c}\frac{c^d}{d!}| \geq \epsilon)$$
$$\leq \mathbb{P}(\sum_{d=g(n)}^{\infty} |\mu_n(d) - e^{-c}\frac{c^d}{d!}| \geq \epsilon/2) + \sum_{d=0}^{g(n)-1} \mathbb{P}(|\mu_n(d) - \mathbb{E}[X_d]/n| \geq \epsilon/(4g(n)))$$
$$+ \sum_{d=0}^{g(n)-1} \mathbb{1}(|\,\mathbb{E}[X_d]/n - e^{-c}\frac{c^d}{d!}| \geq \epsilon/(4g(n))).$$

Pour borner le premier terme on utilise le fait qu'avec grande probabilité il n'y a aucun terme de degré $\geq g$ quand $g(n) \gg \log(n)$. Donc
$$\mathbb{P}(\sum_{d=g(n)}^{\infty} |\mu_n(d) - e^{-c}\frac{c^d}{d!}| \geq \epsilon/2) = o(1) + \mathbb{P}(\sum_{d=g(n)}^{\infty} e^{-c}\frac{c^d}{d!} \geq \epsilon/2) = o(1)$$
vu que $g(n) \to \infty$.

Par la propriété du cours, on montre que le deuxième terme est majoré par
$$\sum_{d=0}^{g(n)-1} \mathbb{P}(|X_d - e^{-c}\frac{c^d}{d!}| \geq \epsilon n/(4g(n))) \leq \sum_{d=0}^{g(n)-1} \frac{C\epsilon g(n)^2}{\epsilon^2 n} = O(g(n)^3/n) = o(1)$$

dès que $g(n) \ll n^{1/3}$.

Pour le dernier terme, considérons

$$\begin{aligned}
\mathbb{E}[X_d]/n - e^{-c}\frac{c^d}{d!} &= \binom{n}{d}(c/n)^d(1-c/n)^{n-1-d} - e^{-c}\frac{c^d}{d!} \\
&= e^{-c}\frac{c^d}{d!}\left(\frac{(n)_d}{n^d}\exp(c-(n-1-d)\log(1-c/n)) - 1\right) = e^{-c}\frac{c^d}{d!}O(d/n)
\end{aligned}$$

pour un certain $O(\cdot)$ uniforme en (d,n), obtenu par des développements limités. Alors si par exemple $g(n) = n^{1/4}$, le troisième terme est nul à partir d'un certain n.