TD 4 : Galton-Watson v2

Dans le premier exercice nous prenons un autre point de vue sur les arbres de Galton-Watson, en les découvrant sommet par sommet selon le parcours en largeur, plutôt que brutalement, génération par génération.

Exercice 1. *Arbres codés par des marches entières*

Soit $G = (V, E)$ un graphe, on suppose que V est totalement ordonné, et on prend $x \in V$. On définit par récurrence l'algorithme de parcours en largeur du graphe. On pose $A^0 = (x)$, $B_0 = x$ puis pour $k \geq 0$, si $A^k \neq ()$,

- On pose $x_k = A_1^k$.
- On pose $\xi_k = |\mathcal{V}_V(x_k) \setminus B_k|$ le nombre de voisins de x_k non encore explorés.
- On pose $c_1(x_k) < \ldots < c_{\xi_k}(x_k)$ le réordonnement croissant de ces voisins $\mathcal{V}_V(x_k) \setminus B_k$.
- $A^{k+1} = (A_2^k, \ldots, A_{|A^k|}^k, c_1(x_k), \ldots, c_{\xi_k}(x_k))$.
- $B_{k+1} = B_k \cup \mathcal{V}_V(A_1^k)$.

On arrête l'algorithme quand $k = N = \inf\{k \geq 0, A^k = ()\}$. On peut montrer (et on admettra) les assertions suivantes :

- L'ensemble des sommets explorés $B_N = \{x_k, 0 \leq k \leq N-1\}$ est $\mathcal{C}(x)$, la composante connexe qui contient x. Ceci implique $|\mathcal{C}(x)| = N$.
- L'ensemble des arêtes explorées $\{(x_k, c_i(x_k)), 0 \leq k \leq N-1, 1 \leq i \leq \xi_k\}$ forme un arbre couvrant de $\mathcal{C}(x)$, dit arbre de parcours en largeur.
- En particulier, si G est un arbre, alors toutes les arêtes sont explorées. La suite $\xi_0, \ldots, \xi_{N-1}$ caractérise alors l'arbre généalogique G à isomorphisme près.

1. Soit $S_0 = 1$, $S_{k+1} = S_k + \xi_k - 1$ pour $0 \leq k \leq N-1$. Justifier par récurrence que $S_k = |A^k|$ pour $0 \leq k \leq N$. En déduire que $N = \inf\{k \geq 0, S_k = 0\}$.
2. On pose $\tau_0 = 0$, $\tau_1 = 1$,$Z_0 = 1$, et par récurrence pour $i \geq 0$, $Z_{i+1} = \sum_{k=\tau_i}^{\tau_{i+1}-1} \xi_k$, et $\tau_{i+2} = \tau_{i+1} + Z_{i+1}$. Montrer que $Z_i = |\{y \in G, d(y, x) = i\}|$.
3. Etant donné une loi μ sur $\mathbb{N}$, on admet pouvoir générer un arbre éventuellement infini $G = T$, appelé arbre de Galton-Watson de loi μ, tel que $(\xi_k)_k$ est une suite i.i.d. de loi μ.
   (a) Montrer qu'alors $(S_k)_{0 \leq i \leq N}$ est une marche aléatoire de loi $\mu(\cdot + 1)$ arrêtée à son premier temps d'atteinte de 0.
   (b) Montrer également que $(Z_i)_{i \geq 0}$ est un processus de branchement de loi μ, c'est-à-dire
   $$\mathbb{P}(Z_{i+1} = a | Z_1, \ldots, Z_i) = \mu^{\star Z_i}(a).$$
   On pourra montrer d'abord pour le conditionnement à $(\tau_1, \ldots \tau_i)$.
   (c) Montrer l'égalité d'évènements
   $$\{|T| < \infty\} = \{N < \infty\} = \{\exists i, Z_i = 0\}.$$

4. Ces deux derniers points permettent d'utiliser la théorie des processus de branchements que vous connaissez déjà. Se rappeler que si $\rho_i = \mathbb{P}(Z_i = 0)$, alors $\rho_0 = \varphi(0)$, et $\rho_{i+1} = \varphi(\rho_i)$, où $\varphi(s) = \mathbb{E}[s^{\xi_1}]$. Déduire que $\rho = \mathbb{P}(T < \infty)$ est le plus petit point fixe de φ dans $[0,1]$.

Exercice 2. *Cas sous-critique mieux fait que dans le TD 2*

On s'intéresse au parcours en largeur de $G_{n,p}$ pour $p = \frac{c}{n}$, $c < 1$.

1. Que vaut $|B_k|$ en fonction des $(\xi_i)_i$? En déduire que conditionnellement à $\xi_0, \dots, \xi_{k-1}$,

$$\xi_k \sim \mathrm{Bi}(n - 1 - \sum_{i=0}^{k-1} \xi_i, p).$$

2. En déduire que l'on peut construire un arbre de Galton-Watson T' de loi $\mathrm{Bi}(n,p)$ tel que $N' = |T'| \geq |\mathcal{C}(x)|$.
3. Soit $s > 1$. Soit $(\xi'_k)_k$ une suite i.i.d. de loi $\mathrm{Bi}(n,p)$. Calculer $\varphi(s) = \mathbb{E}[s^{\xi'_1}]$, et montrer que $\varphi(s)/s \leq e^{c(s-1)-\log(s)}$. Optimiser en $s > 1$ pour avoir une borne aussi petite que possible.
4. Dès maintenant on suppose que la marche S' associée à T' continue après le temps N', ce qui simplifie les calculs. Noter que ça ne change pas la loi de N'. Montrer que $D_k = s^{S'_k+k}\varphi(s)^{-k}$ est une martingale positive partant de 1.
5. Appliquer l'arrêt optionnel en $N' \wedge k$ pour montrer que $\mathbb{E}[(\varphi(s)/s)^{-k\wedge N'}] \leq 1$, et en déduire par convergence monotone que $\mathbb{E}[e^{\alpha N'}] \leq 1$, où $\alpha = c - 1 - \log c$.
6. Conclure qu'avec grande proba il n'y a aucune composante de taille $\geq \frac{(1+\epsilon)}{\alpha} \log n$.

Cet exercice montre une sorte d'analogue dans Galton-Watson de la dualité entre $c > 1$ hors de la composante géante et $c < 1$ dans Erdös-Rényi.

Exercice 3. *Dualité dans Galton-Watson*

On considère un arbre de Galton-Watson de loi μ. On utilise la notation du premier exercice, et on se place dans le cas où $0 < \rho < 1$ (cas sur-critique).

1. Soit $\mathbb{P}_x$ la loi sous laquelle $S_0 = x$. Montrer en utilisant la propriété de Markov forte, que $\mathbb{P}_x(N < \infty) = \rho^x$.
2. Soit $\overline{\mathbb{P}}_x$ la loi conditionnelle à l'extinction :

$$\overline{\mathbb{P}}_x(A) = \frac{1}{\rho^x} \mathbb{P}_x(A \cap (N < \infty)).$$

Calculer $\overline{\mathbb{P}}_{x_0}(S_{i+1} - S_i = k | S_0 = x_0, \dots, S_i = x_i)$ en utilisant la propriété de Markov simple, pour montrer que sous $\overline{\mathbb{P}}_x$, la suite $(S_i)_{i\geq 0}$ est toujours une marche aléatoire arrêtée en 0.

3. En déduire la loi $\overline{\mu}$ des ξ_i sous $\overline{\mathbb{P}}_x$. Que vaut la fonction génératrice $\overline{\varphi}(s) = \sum_{i=0}^{\infty} \overline{\mu}(i) s^i$? A quel régime appartient $\overline{\mu}$?