Graphes aléatoires, sporz

Exercice 1 Graphes aléatoires

Soit $G = (V, E)$ un graphe non orienté. On appelle ***triangle de G*** un sous-ensemble de $\{x, y, z\}$ de trois sommets distincts de G tel que $\{x, y\} \in E$, $\{x, z\} \in E$ et $\{y, z\} \in E$. Pour rappel, on appelle ***distance*** de deux sommets du graphe le nombre d'arêtes du plus court chemin entre ces deux sommets. Le ***diamètre*** du graphe est le maximum des distances entre deux sommets du graphe.

Le modèle de graphe aléatoire d'Erdős-Rényi $G(n, p)$ est le graphe $G(V, E)$ tel que $|V| = n$ et $\forall x, y \in V$, $\mathbb{P}(\{x, y\} \in E) = p$. Souvent p dépend de n, mais dans ce TD il sera constant.

▷ 1. Soit x et y deux sommets distincts. Calculer $\mathbb{P}(\exists z, \{z, x\} \in E \wedge \{z, y\} \in E)$.

▷ 2. Calculer la probabilité de l'évènement « $\{x, y\}$ est l'un des côtés d'un triangle ».

▷ 3. Soit E_x l'évènement $\exists y \neq x \quad \forall z \quad \neg(\{z, x\} \in E \wedge \{z, y\} \in E)$.
Calculer $\lim_{n\to+\infty} \mathbb{P}(E_x)$.

▷ 4. Calculer $\lim_{n\to+\infty} \mathbb{P}(G(n, p)$ contient au moins un triangle$)$.

▷ 5. Quel est le nombre de triangles que l'on peut espérer dans un tel graphe ?

▷ 6. Calculer $\lim_{n\to+\infty} \mathbb{P}(\text{diamètre}(G(n, p)) \leqslant 2)$.

▷ 7. Calculer $\lim_{n\to+\infty} \mathbb{P}(\text{diamètre}(G(n, p)) = 2)$.

Solution

▷ 1. Soit $A_{x,y}$ l'évènement « $\exists z,\ \{z, x\} \in E \ \wedge\ \{z, y\} \in E$ » où x et y sont distincts. L'évènement $\neg A_{x,y}$ est donc « $\forall z,\ \{z, x\} \notin E \vee \{z, y\} \notin E$ ». Soit z distinct de x et y.

$$\begin{aligned}\mathbb{P}(\{z, x\} \notin E \vee \{z, y\} \notin E) &= 1 - \mathbb{P}(\{z, x\} \in E \wedge \{z, y\} \in E)\\ &= 1 - p^2\end{aligned}$$

Or $\neg A_{x,y}$ est l'intersection pour z différent de x et y des évènements $\{z, x\} \notin E \vee \{z, y\} \notin E$, qui sont tous indépendants (ils ne font pas intervenir d'arête commune). D'où $\mathbb{P}(\neg A_{x,y}) = (1 - p^2)^{n-2}$ et donc $\mathbb{P}(A_{x,y}) = 1 - (1 - p^2)^{n-2}$.

▷ 2. Soit $B_{x,y}$ l'évènement « $\{x, y\}$ est l'un des côtés d'un triangle ». $B_{x,y} = A_{x,y} \wedge \{x, y\} \in E$. Ces évènements sont indépendants donc $\mathbb{P}(B_{x,y}) = p\left(1 - \left(1 - p^2\right)^{n-2}\right)$.

▷ 3. On remarque que $E_x = (\exists y \neq x,\ \neg A_{x,y}) = \bigcup_{x\neq y} \neg A_{x,y}$.

$$\begin{aligned}\mathbb{P}(E_x) &= \mathbb{P}\left(\bigcup_{y\neq x} \neg A_{x,y}\right)\\ &\leqslant \sum_{y\neq x} \mathbb{P}(\neg A_{x,y})\\ &\leqslant (n-1)(1-p^2)^{(n-2)}\end{aligned}$$

Dès que $p \neq 0$, $\lim_{n\to\infty}(n-1)(1-p^2)^{(n-2)} = 0$ donc $\lim_{n\to\infty} \mathbb{P}(E_x) = 0$.

▷ 4. Soit C l'événement « $G(n, p)$ contient au moins un triangle ». Nous allons montrez que $\exists x, \neg E_x \Rightarrow C$. Supposons donc qu'il existe x tel que $\neg E_x$; i.e. tel que $\forall y \neq x, \exists z, \{z, x\} \in E \wedge \{z, y\} \in E$.
Soit $y \neq x$, il existe z tel que $\{z, x\} \in E$. En particulier $z \neq x$ donc on peut appliquer la propriété avec z : il existe $z' \neq x$ tel que $\{z', x\} \in E \wedge \{z', z\} \in E$. Et donc $\{x, z, z'\}$ est un triangle.
On en déduit que $\mathbb{P}(C) \geqslant \mathbb{P}\left(\bigcup_{x\in V} \neg E_x\right) \geqslant \max_{x\in V} \mathbb{P}(\neg E_x)$. Comme $\forall x,\ \lim_{n\to\infty} \mathbb{P}(\neg E_x) = 1$, $\lim_{n\to\infty} \mathbb{P}(C) = 1$.

Autre méthode Numérotons les sommets de V : $V = \{x_1, \ldots, x_n\}$. Soit C_i l'événement « $\{x_{3i}, x_{3i+1}, x_{3i+2}\}$ est un triangle ». $\bigcup_{3i+2\leqslant n} C_i \subset C$ donc $\neg C \subset \bigcap_{3i+2\leqslant n} \neg C_i$. Or les C_i sont indépendants donc

$$\mathbb{P}(\neg C) \leqslant \mathbb{P}\left(\bigcap_{3i+2\leqslant n} \neg C_i\right) = \prod_{3i+2\leqslant n} \mathbb{P}(\neg C_i) = \left(1 - p^3\right)^{\lfloor n/3\rfloor} \xrightarrow[n\to\infty]{} 0$$

Donc $\mathbb{P}(C) = 1$.

▷ 5. Soit $\{x, y, z\}$ trois sommets distincts, $\mathbb{P}(\{x, y, z\}$ est un triangle$) = p^3$. Il y a $\binom{n}{3}$ triangles possibles et donc par linéarité de l'espérance, l'espérance du nombre de triangle est $\binom{n}{3}p^3$.

▷ 6. Soit D l'évènement « diamètre$(G(n, p)) \leqslant 2$ ». $\neg D = \exists x,\ \exists y \neq x,\ \{x, y\} \notin E \wedge \neg A_{x,y}$. On en déduit que

$$\begin{aligned}\mathbb{P}(\neg D) &= \mathbb{P}\left(\bigcup_{x\neq y} \{x, y\} \notin E \wedge \neg A_{x,y}\right)\\ &\leqslant \sum_{x\neq y} \mathbb{P}(\{x, y\} \notin E \wedge \neg A_{x,y})\\ &\leqslant \binom{n}{2}(1-p)(1-p^2)^{n-2}\end{aligned}$$

Or $\lim_{n\to\infty} \binom{n}{2}(1-p)(1-p^2)^{n-2} = 0$, donc $\lim_{n\to\infty} \mathbb{P}(\neg D) = 0$ et $\lim_{n\to\infty} \mathbb{P}(D) = 1$.

▷ 7. Soit E_i l'évènement « diamètre$(G(n, p)) = i$ ». $E_2 = D \setminus E_1$ et $E_1 \subset D$ d'où $\mathbb{P}(E_2) = \mathbb{P}(D) - \mathbb{P}(E_1) = \mathbb{P}(D) - p^{\frac{n(n-1)}{2}}$. Finalement, $\lim_{n\to\infty} \mathbb{P}(E_2) = 1$.

Exercice 2 (***) Sporz

La difficulté de cet exercice réside dans la modélisation.

Les joueurs de sporz(.fr) on parfois à révéler une information au groupe tout en restant anonyme. Ils ont mis au point le protocole suivant. Un joueur qui souhaite révéler une information fait un tirage au sort (par exemple en regardant discrètement le nombre de secondes de sa montre) ;

- avec probabilité p, il publie l'information à voix haute[1] ;
- sinon, il choisit au hasard un joueur, lui transmet discrètement l'information, et lui demande de faire pareil (regarder sa montre, etc.).

* ▷ 1. Quelle est la probabilité a_k de l'événement $A_k :=$ « Le k^{e} joueur recevant discrètement l'information est l'auteur. » ?

** ▷ 2. Vous êtes l'un des n joueurs, et vous entendez quelqu'un publier à voix haute une information, quelle est la probabilité qu'il en soit l'auteur ?

*** ▷ 3. Quelqu'un vient de vous transmettre discrètement une information, quelle est la probabilité qu'il en soit l'auteur ?

Solution

▷ 1. $a_1 = 0$ car le premier joueur la reçoit de l'auteur. Pour que la $(k+1)^{\text{e}}$ personne soit l'auteur, il faut et suffit que :

- la k^{e} ne soit pas l'auteur, ce qui arrive avec probabilité $1 - a_k$;
- et elle choisisse la $(k+1)^{\text{e}}$, ce qui arrive avec probabilité $\frac{1}{n-1}$ (elle ne se choisit pas).

Donc $a_{k+1} = \frac{1-a_k}{n-1}$, ce qui se résoud en, $a_k = \frac{1}{n} - \frac{1}{n}\left(\frac{-1}{n-1}\right)^{k-1}$. En particulier, $a_k \to \frac{1}{n}$: on oublie rapidement qui est à l'origine de la chaîne.

▷ 2. Notons K le nombre de personnes par qui l'information s'est transmise avant d'être publiée (auteur exclu). En d'autres termes, K est la longueur de la chaîne. L'évènement qui nous intéresse est « $K = 0$ ou $K \geqslant 1$ et A_K ». Il est de probabilité $p + \sum_{k=1}^{\infty}(1-p)^k p\, a_k = p + \frac{p}{n}\left(\sum_{k=1}^{\infty}(1-p)^k\right) + p\frac{n-1}{n}\left(\sum_{k=1}^{\infty}(-\frac{1-p}{n-1})^k\right) = p + \frac{1-p}{n} - \frac{n-1}{n}\frac{p(1-p)}{n-p}$.
(Ceci vaut environ $p + \frac{1}{n}$ pour p petit et n grand, et une bonne approximation dès $n = 4$ est $p + \frac{1}{n} - \frac{2p}{n}$.)

▷ 3. Soient

- B_k l'événement « $K \geqslant k$ », i.e. « L'information a été transmise (et non publiée) au moins k fois. »,
- B'_k l'événement « B_k et sur les k premières transmissions vous n'avez jamais reçu l'information. ».

$\mathbb{P}(B_k) = (1-p)^k$ et $\mathbb{P}(B'_k) = \mathbb{P}(B_k)\left(\frac{n-2}{n-1}\right)^k$.
On convient que le « 0^{e} » joueur à recevoir l'information est l'auteur. L'évènement qui nous intéresse est « $\bigcup_{k\geqslant 0}[B'_k$ et A_k et le k^{e} joueur à la recevoir choisit de vous la transmettre], sachant B_1. ».

Soit $c_k := \mathbb{P}(A_k \mid B'_k)$. $c_1 = 0$ car le premier joueur reçoit l'information de l'auteur, donc il n'est pas l'auteur. De même qu'à la question 1, l'évènement « A_{k+1} et B'_{k+1} » est « $\overline{A_k}$ et B'_k et le k^{e} joueur choisit de transmettre à l'auteur. ».
Donc $\mathbb{P}\left(A_{k+1} \text{ et } B'_{k+1}\right) = \mathbb{P}\left(\overline{A_k} \text{ et } B'_k\right)\frac{1-p}{n-1}$.
Or $\mathbb{P}\left(\overline{A_k} \text{ et } B'_k\right) = \mathbb{P}\left(\overline{A_k} \,\middle|\, B'_k\right)\mathbb{P}(B'_k) = (1 - \mathbb{P}(A_k \mid B'_k))\mathbb{P}(B'_k) = (1-c_k)\mathbb{P}(B'_k)$.
Donc $c_{k+1} = \frac{\mathbb{P}\left(A_{k+1} \text{ et } B'_{k+1}\right)}{\mathbb{P}\left(B'_{k+1}\right)} = \frac{1-p}{n-1}(1-c_k)\frac{\mathbb{P}\left(B'_k\right)}{\mathbb{P}\left(B'_{k+1}\right)} = \frac{1-c_k}{n-2}$. Ceci se résoud de même en $c_k = \frac{1}{n-1} - \frac{1}{n-1}\left(\frac{-1}{n-2}\right)^{k-1}$.

Remarquons que « B'_k et A_k » est équivalent à « B'_k et (A_k sachant B'_k) ». On peut maintenant calculer la probabilité de l'évènement qui nous intéresse. C'est

$$\left(\frac{1-p}{n-1} + \sum_{k\geqslant 1}\left((1-p)\frac{n-2}{n-1}\right)^k c_k \frac{1-p}{n-1}\right)/(1-p) = \frac{pn - 3p + 2}{(np+1-2p)(n-p)}$$

Ceci vaut environ $\frac{np+2}{n(np+1)}$ pour p petit, ce qui est une approximation à 0.03 près pour $n \geqslant 5$ et $p \leqslant 0.5$.

[1] Le jeu a une phase où les joueurs circulent dans la pièce et se parlent librement.