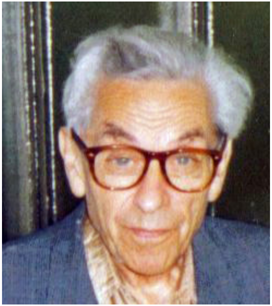
Grands graphes aléatoires

Raphaël Rossignol (Univ. Grenoble Alpes)

Journée régionale de l’APMEP, 9 mars 2022

Pál Erdős, Alfréd Rényi et Edgar Gilbert



Pál Erdős
(1913-1996)

Alfréd Rényi
(1921-1970)

Edgar Gilbert
(1923-2013)

Graphes : définition

Définition

Un graphe (non orienté, sans boucle) est un couple d'ensembles (V, E)*, où* E *est un ensemble de paires d'éléments de* V

Exemple

- $V = \{1, 2, 3, 4\}$
- $E = \{\{1, 2\}, \{2, 3\}, \{3, 1\}, \{1, 4\}\}$

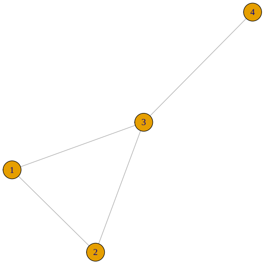
Graphes : définition

Définition

Un graphe (non orienté, sans boucle) est un couple d'ensembles (V, E)*, où* E *est un ensemble de paires d'éléments de* V

Exemple

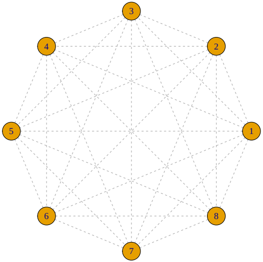
- $V = \{1, 2, 3, 4\}$
- $E = \{\{1, 2\}, \{2, 3\}, \{3, 1\}, \{1, 4\}\}$



Graphes d'Erdős-Rényi

Erdős & Rényi (1959) : $\mathcal{G}(N, M)$

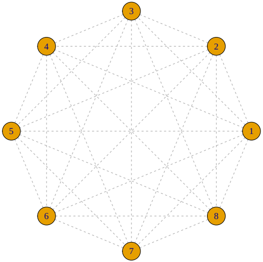
- Sommets : $V = \{1, \ldots, N\}$
- Arêtes possibles : $\mathcal{E}_N = \{$paires d'éléments de V$\}$
  $|\mathcal{E}_N| = \binom{N}{2} = \frac{N(N-1)}{2})$
- Arêtes de $\mathcal{G}(N, M)$: M arêtes choisies au hasard dans $\mathcal{E}_N$

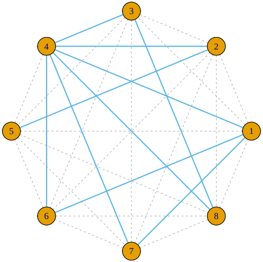


Graphes d'Erdős-Rényi

Erdős & Rényi (1959) : $\mathcal{G}(N, M)$

- Sommets : $V = \{1, \dots, N\}$
- Arêtes possibles : $\mathcal{E}_N = \{$paires d'éléments de V$\}$
  $|\mathcal{E}_N| = \binom{N}{2} = \frac{N(N-1)}{2})$
- Arêtes de $\mathcal{G}(N, M)$: M arêtes choisies au hasard dans $\mathcal{E}_N$

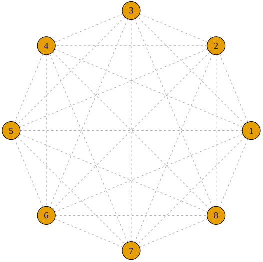




Graphes de Gilbert

Gilbert (1959) : $\mathcal{G}(n, p)$

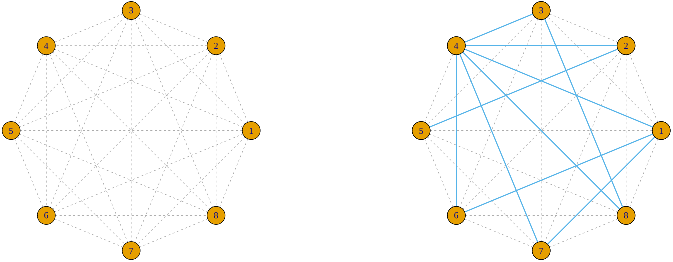
- $n \in \mathbb{N}^*$, $p \in [0, 1]$
- Sommets : $1, \ldots, n$
- Arêtes : $\binom{n}{2}$ arêtes possibles, indépendantes, présentes avec proba p



Graphes de Gilbert

Gilbert (1959) : $\mathcal{G}(n, p)$

- $n \in \mathbb{N}^*$, $p \in [0, 1]$
- Sommets : $1, \ldots, n$
- Arêtes : $\binom{n}{2}$ arêtes possibles, indépendantes, présentes avec proba p



Indépendance de v.a discrètes

On dit que des v.a $X_1, \dots, X_r$ sont indépendantes ssi :

$$\forall x_1, \dots, x_r, \quad \mathbb{P}(X_1 = x_1, \dots, X_r = x_r) = \mathbb{P}(X_1 = x_1) \dots \mathbb{P}(X_r = x_r)$$

Si $G = (\{1, \dots, n\}, E)$, $\mathbb{P}(\mathcal{G}(n,p) = G) = p^{|E|}(1-p)^{\binom{n}{2}-|E|}$

Proposition

Si $X_1, \dots, X_r$ sont indépendantes, $\mathbb{E}[X_1 \dots X_r] = \mathbb{E}[X_1] \dots \mathbb{E}[X_r]$

Indépendance de v.a discrètes

On dit que des v.a $X_1, \ldots, X_r$ sont indépendantes ssi :

$$\forall x_1, \ldots, x_r, \quad \mathbb{P}(X_1 = x_1, \ldots, X_r = x_r) = \mathbb{P}(X_1 = x_1) \ldots \mathbb{P}(X_r = x_r)$$

Si $G = (\{1, \ldots, n\}, E)$, $\mathbb{P}(\mathcal{G}(n,p) = G) = p^{|E|}(1-p)^{\binom{n}{2} - |E|}$

Proposition

Si $X_1, \ldots, X_r$ sont indépendantes, $\mathbb{E}[X_1 \ldots X_r] = \mathbb{E}[X_1] \ldots \mathbb{E}[X_r]$

Indépendance de v.a discrètes

On dit que des v.a $X_1, \ldots, X_r$ sont indépendantes ssi :

$$\forall x_1, \ldots, x_r, \quad \mathbb{P}(X_1 = x_1, \ldots, X_r = x_r) = \mathbb{P}(X_1 = x_1) \ldots \mathbb{P}(X_r = x_r)$$

Si $G = (\{1, \ldots, n\}, E)$, $\mathbb{P}(\mathcal{G}(n,p) = G) = p^{|E|}(1-p)^{\binom{n}{2} - |E|}$

Proposition

Si $X_1, \ldots, X_r$ sont indépendantes, $\mathbb{E}[X_1 \ldots X_r] = \mathbb{E}[X_1] \ldots \mathbb{E}[X_r]$

Nombre d'arêtes dans $\mathcal{G}(n, p)$ et n "grand"

Nombre d'arêtes :

$$\mathcal{N}_n = \sum_{\{i,j\} \in \mathcal{E}_n} \mathbb{1}_{i \sim j}$$

$\mathcal{N}_n$ est de loi $\mathcal{B}(\binom{n}{2}, p)$

$$\mathbb{P}(\mathcal{N}_n = k) = \binom{\frac{n(n-1)}{2}}{k} p^k (1-p)^{\frac{n(n-1)}{2} - k}$$

$$\mathbb{P}(\mathcal{N}_n = 0) = (1-p)^{\frac{n(n-1)}{2}} = e^{-p\frac{n^2}{2} + o(n^2 p)}$$

Si $p \ll \frac{1}{n^2}$,

$$\mathbb{P}(\mathcal{N}_n = 0) \xrightarrow[n \to +\infty]{} 1$$

$\mathcal{G}(n, p)$ n'a aucune arête, a.p.s

$\rightsquigarrow$ Dans la suite, $\frac{1}{n^2} \ll p$

Nombre d'arêtes dans $\mathcal{G}(n,p)$ et n “grand”

Nombre d'arêtes :

$$\mathcal{N}_n = \sum_{\{i,j\}\in\mathcal{E}_n} \mathbb{1}_{i\sim j}$$

$\mathcal{N}_n$ est de loi $\mathcal{B}(\binom{n}{2}, p)$

$$\mathbb{P}(\mathcal{N}_n = k) = \binom{\frac{n(n-1)}{2}}{k} p^k (1-p)^{\frac{n(n-1)}{2}-k}$$

$$\mathbb{P}(\mathcal{N}_n = 0) = (1-p)^{\frac{n(n-1)}{2}} = e^{-p\frac{n^2}{2}+o(n^2p)}$$

Si $p \ll \frac{1}{n^2}$,

$$\mathbb{P}(\mathcal{N}_n = 0) \xrightarrow[n\to+\infty]{} 1$$

$\mathcal{G}(n,p)$ n'a aucune arête, a.p.s

$\rightsquigarrow$ Dans la suite, $\frac{1}{n^2} \ll p$

Nombre d'arêtes dans $\mathcal{G}(n,p)$ et n “grand”

Nombre d'arêtes :

$$\mathcal{N}_n = \sum_{\{i,j\}\in\mathcal{E}_n} \mathbb{1}_{i\sim j}$$

$\mathcal{N}_n$ est de loi $\mathcal{B}(\binom{n}{2}, p)$

$$\mathbb{P}(\mathcal{N}_n = k) = \binom{\frac{n(n-1)}{2}}{k} p^k (1-p)^{\frac{n(n-1)}{2}-k}$$

$$\mathbb{P}(\mathcal{N}_n = 0) = (1-p)^{\frac{n(n-1)}{2}} = e^{-p\frac{n^2}{2}+o(n^2p)}$$

Si $p \ll \frac{1}{n^2}$,

$$\mathbb{P}(\mathcal{N}_n = 0) \xrightarrow[n\to+\infty]{} 1$$

$\mathcal{G}(n,p)$ n'a aucune arête, a.p.s

$\rightsquigarrow$ Dans la suite, $\frac{1}{n^2} \ll p$

Exemples de questions

Pour n grand,

- Pour un graphe H fixé, à partir de quelle valeur de p $\mathcal{G}(n,p)$ contient-il une copie de H ?
- Evolution des tailles des composantes connexes de $\mathcal{G}(n,p)$?

Exemples de questions

Pour n grand,

- Pour un graphe H fixé, à partir de quelle valeur de p $\mathcal{G}(n,p)$ contient-il une copie de H ?
- Evolution des tailles des composantes connexes de $\mathcal{G}(n,p)$?

Plan

Plan

Markov et Tchebychev

Univers : $\Omega = \{\text{Graphes de sommets } 1, \dots, n\}$

Espérance de X :

$$\begin{aligned}\mathbb{E}[X] &:= \sum_{G \in \Omega} X(G)\mathbb{P}(\mathcal{G}(n,p) = G) = \sum_{x \in X(\Omega)} x\mathbb{P}(\{G \in \Omega \,:\, X(G) = x\}) \\ &= \sum_{x \in X(\Omega)} x\mathbb{P}(X = x)\end{aligned}$$

Linéarité de l'espérance : si X et Y v.a, λ réel

$$\mathbb{E}[\lambda X] = \lambda \mathbb{E}[X] \quad \text{et} \quad \mathbb{E}[X + Y] = \mathbb{E}[X] + \mathbb{E}[Y]$$

Markov et Tchebychev

Univers : $\Omega = \{\text{Graphes de sommets } 1, \dots, n\}$

Espérance de X :

$$\begin{aligned} \mathbb{E}[X] &:= \sum_{G \in \Omega} X(G)\mathbb{P}(\mathcal{G}(n,p) = G) = \sum_{x \in X(\Omega)} x\mathbb{P}(\{G \in \Omega : X(G) = x\}) \\ &= \sum_{x \in X(\Omega)} x\mathbb{P}(X = x) \end{aligned}$$

Linéarité de l'espérance : si X et Y v.a, λ réel

$$\mathbb{E}[\lambda X] = \lambda \mathbb{E}[X] \quad \text{et} \quad \mathbb{E}[X + Y] = \mathbb{E}[X] + \mathbb{E}[Y]$$

Markov et Tchebychev

Si X est à valeurs positives et $a > 0$

$$\mathbb{P}(X \geq a) = \sum_{x \geq a} \mathbb{P}(X = x) \leq \sum_{x} \frac{x}{a} \mathbb{P}(X = x) = \frac{\mathbb{E}[X]}{a}$$

Inégalité de Markov : $\mathbb{P}(X \geq a) \leq \frac{\mathbb{E}[X]}{a}$

Si $a > 0$,

$$\mathbb{P}(|X - \mathbb{E}[X]| > a) = \mathbb{P}((X - \mathbb{E}[X])^2 > a^2) \leq \frac{\mathbb{E}[(X - \mathbb{E}[X])^2]}{a^2} = \frac{\mathrm{Var}(X)}{a^2}$$

Inégalité de Tchebychev : $\mathbb{P}(|X - \mathbb{E}[X]| \geq a) \leq \frac{\mathrm{Var}(X)}{a^2}$

Markov et Tchebychev

Si X est à valeurs positives et $a > 0$

$$\mathbb{P}(X \geq a) = \sum_{x \geq a} \mathbb{P}(X = x) \leq \sum_{x} \frac{x}{a} \mathbb{P}(X = x) = \frac{\mathbb{E}[X]}{a}$$

Inégalité de Markov : $\mathbb{P}(X \geq a) \leq \dfrac{\mathbb{E}[X]}{a}$

Si $a > 0$,

$$\mathbb{P}(|X - \mathbb{E}[X]| > a) = \mathbb{P}((X - \mathbb{E}[X])^2 > a^2) \leq \frac{\mathbb{E}[(X - \mathbb{E}[X])^2]}{a^2} = \frac{\mathrm{Var}(X)}{a^2}$$

Inégalité de Tchebychev : $\mathbb{P}(|X - \mathbb{E}[X]| \geq a) \leq \dfrac{\mathrm{Var}(X)}{a^2}$

Markov et Tchebychev

Si X est à valeurs positives et $a > 0$

$$\mathbb{P}(X \geq a) = \sum_{x \geq a} \mathbb{P}(X = x) \leq \sum_{x} \frac{x}{a} \mathbb{P}(X = x) = \frac{\mathbb{E}[X]}{a}$$

Inégalité de Markov : $\mathbb{P}(X \geq a) \leq \dfrac{\mathbb{E}[X]}{a}$

Si $a > 0$,

$$\mathbb{P}(|X - \mathbb{E}[X]| > a) = \mathbb{P}((X - \mathbb{E}[X])^2 > a^2) \leq \frac{\mathbb{E}[(X - \mathbb{E}[X])^2]}{a^2} = \frac{\mathrm{Var}(X)}{a^2}$$

Inégalité de Tchebychev : $\mathbb{P}(|X - \mathbb{E}[X]| \geq a) \leq \dfrac{\mathrm{Var}(X)}{a^2}$

Markov et Tchebychev

Si X est à valeurs positives et $a > 0$

$$\mathbb{P}(X \geq a) = \sum_{x \geq a} \mathbb{P}(X = x) \leq \sum_{x} \frac{x}{a} \mathbb{P}(X = x) = \frac{\mathbb{E}[X]}{a}$$

Inégalité de Markov : $\mathbb{P}(X \geq a) \leq \dfrac{\mathbb{E}[X]}{a}$

Si $a > 0$,

$$\mathbb{P}(|X - \mathbb{E}[X]| > a) = \mathbb{P}((X - \mathbb{E}[X])^2 > a^2) \leq \frac{\mathbb{E}[(X - \mathbb{E}[X])^2]}{a^2} = \frac{\mathrm{Var}(X)}{a^2}$$

Inégalité de Tchebychev : $\mathbb{P}(|X - \mathbb{E}[X]| \geq a) \leq \dfrac{\mathrm{Var}(X)}{a^2}$

Concentration du nombre d'arêtes

$$\mathcal{N}_n \sim \mathcal{B}\left(\binom{n}{2}, p\right)$$

- $\mathbb{E}[\mathcal{N}_n] = \binom{n}{2}p \gg 1$
- $\mathrm{Var}(\mathcal{N}_n) = \binom{n}{2}p(1-p) \leq \mathbb{E}[\mathcal{N}_n]$

Donc, pour tout $\varepsilon > 0$

$$\mathbb{P}\left(\left|\mathcal{N}_n - \binom{n}{2}p\right| > \varepsilon\binom{n}{2}p\right) \leq \frac{1}{\varepsilon^2\binom{n}{2}p} \ll 1$$

Concentration du nombre d'arêtes

$$\mathcal{N}_n \sim \mathcal{B}\left(\binom{n}{2}, p\right)$$

- $\mathbb{E}[\mathcal{N}_n] = \binom{n}{2} p \gg 1$
- $\mathrm{Var}(\mathcal{N}_n) = \binom{n}{2} p(1-p) \leq \mathbb{E}[\mathcal{N}_n]$

Donc, pour tout $\varepsilon > 0$

$$\mathbb{P}\left(\left|\mathcal{N}_n - \binom{n}{2} p\right| > \varepsilon \binom{n}{2} p\right) \leq \frac{1}{\varepsilon^2 \binom{n}{2} p} \ll 1$$

Evolution

Fonction seuil $s(n)$ pour "contenir une copie de H"

$$\text{Si } p \ll s,\ \mathbb{P}(\mathcal{G}(n,p) \text{ contient une copie de } H) \xrightarrow[n\to+\infty]{} 0$$

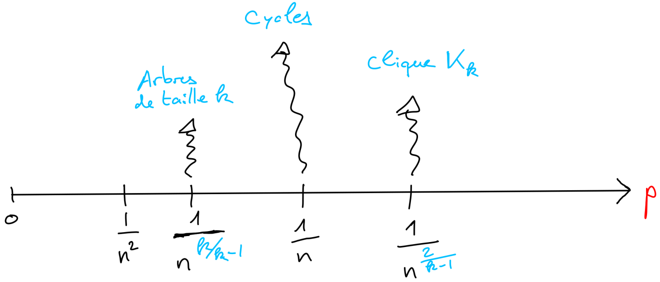
$$\text{Si } p \gg s,\ \mathbb{P}(\mathcal{G}(n,p) \text{ contient une copie de } H) \xrightarrow[n\to+\infty]{} 1$$

Evolution

Fonction seuil $s(n)$ pour "contenir une copie de H"

$$\text{Si } p \ll s,\ \mathbb{P}(\mathcal{G}(n,p) \text{ contient une copie de } H) \xrightarrow[n\to+\infty]{} 0$$

$$\text{Si } p \gg s,\ \mathbb{P}(\mathcal{G}(n,p) \text{ contient une copie de } H) \xrightarrow[n\to+\infty]{} 1$$



Triangles

Théorème

Soit X_n le nombre de triangles contenus dans $\mathcal{G}(n,p)$.

- *Si $p \ll \frac{1}{n}$,*

$$\mathcal{G}(n,p) \text{ ne contient pas de triangle a.p.s}$$

- *Si $p \gg \frac{1}{n}$,*

$$\mathcal{G}(n,p) \text{ contient au moins un triangle a.p.s}$$

$\frac{1}{n}$ est la fonction seuil de la propriété "contenir un triangle"

Triangles : démo pour $p \ll \frac{1}{n}$

$$X_n := \sum_{\{i,j,k\} \text{ distincts}} \mathbb{1}_{i\sim j}\mathbb{1}_{j\sim k}\mathbb{1}_{i\sim k}$$

Premier moment (Markov) :

$$\begin{aligned} \mathbb{P}(X_n \geq 1) &\leq \mathbb{E}[X_n] \\ &= \binom{n}{3} p^3 \\ &\xrightarrow[n\to+\infty]{} 0 \quad \text{si } p \ll \frac{1}{n} \end{aligned}$$

Triangles : démo pour $p \ll \frac{1}{n}$

$$X_n := \sum_{\{i,j,k\} \text{ distincts}} \mathbb{1}_{i\sim j}\mathbb{1}_{j\sim k}\mathbb{1}_{i\sim k}$$

Premier moment (Markov) :

$$\begin{aligned} \mathbb{P}(X_n \geq 1) &\leq \mathbb{E}[X_n] \\ &= \binom{n}{3} p^3 \\ &\xrightarrow[n\to+\infty]{} 0 \quad \text{si } p \ll \frac{1}{n} \end{aligned}$$

Triangles : démo pour $p \gg \frac{1}{n}$

Second moment (Tchebychev) :

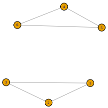
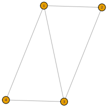
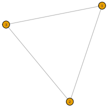
$$\begin{aligned}\mathbb{E}[(X_n)^2] &= \sum_{T_1, T_2} \mathbb{P}(T_1 \subset \mathcal{G}(n,p) \text{ et } T_2 \subset \mathcal{G}(n,p)) \\ &= A + B + C\end{aligned}$$

$A = \binom{n}{3}\binom{n-3}{3} p^6 \leq \mathbb{E}[X_n]^2$ donc $\mathrm{Var}(X_n) \leq B + C$

Triangles : démo pour $p \gg \frac{1}{n}$

Second moment (Tchebychev) :

$$\begin{aligned}\mathbb{E}[(X_n)^2] &= \sum_{T_1, T_2} \mathbb{P}(T_1 \subset \mathcal{G}(n,p) \text{ et } T_2 \subset \mathcal{G}(n,p)) \\ &= A + B + C\end{aligned}$$

$A : \lvert T_1 \cap T_2 \rvert = 0$	$B : \lvert E(T_1) \cap E(T_2) \rvert = 1$	$C : T_1 = T_2$
		

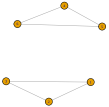
$A = \binom{n}{3}\binom{n-3}{3} p^6 \leq \mathbb{E}[X_n]^2$ donc $\mathrm{Var}(X_n) \leq B + C$

Triangles : démo pour $p \gg \frac{1}{n}$

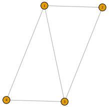
Second moment (Tchebychev) :

$$\begin{aligned} \mathbb{E}[(X_n)^2] &= \sum_{T_1,T_2} \mathbb{P}(T_1 \subset \mathcal{G}(n,p) \text{ et } T_2 \subset \mathcal{G}(n,p)) \\ &= A + B + C \end{aligned}$$

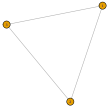
$A : |T_1 \cap T_2| = 0$



$B : |E(T_1) \cap E(T_2)| = 1$



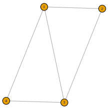
$C : T_1 = T_2$

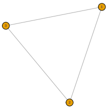


$A = \binom{n}{3}\binom{n-3}{3} p^6 \leq \mathbb{E}[X_n]^2$ donc $\mathrm{Var}(X_n) \leq B + C$

Triangles : démo pour $p \gg \frac{1}{n}$

$B : |E(T_1) \cap E(T_2)| = 1$ | $C : T_1 = T_2$





- $B \lesssim n^4 p^5$
- $C \lesssim n^3 p^3$

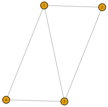
Si $p \gg \frac{1}{n}$,

$$B + C \ll \mathbb{E}[X_n]^2$$

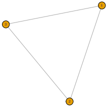
$$\mathbb{P}(X_n = 0) \leq \mathbb{P}(|X_n - \mathbb{E}[X_n]| \geq \mathbb{E}[X_n]) \leq \frac{\mathrm{Var}(X_n)}{\mathbb{E}[X_n]^2} \xrightarrow[n \to +\infty]{} 0$$

Triangles : démo pour $p \gg \frac{1}{n}$

$B : |E(T_1) \cap E(T_2)| = 1$



$C : T_1 = T_2$



- $B \lesssim n^4 p^5$
- $C \lesssim n^3 p^3$

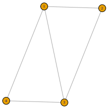
Si $p \gg \frac{1}{n}$,

$$B + C \ll \mathbb{E}[X_n]^2$$

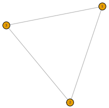
$$\mathbb{P}(X_n = 0) \leq \mathbb{P}(|X_n - \mathbb{E}[X_n]| \geq \mathbb{E}[X_n]) \leq \frac{\mathrm{Var}(X_n)}{\mathbb{E}[X_n]^2} \xrightarrow[n \to +\infty]{} 0$$

Triangles : démo pour $p \gg \frac{1}{n}$

$B : |E(T_1) \cap E(T_2)| = 1$



$C : T_1 = T_2$



- $B \lesssim n^4 p^5$
- $C \lesssim n^3 p^3$

Si $p \gg \frac{1}{n}$,

$$B + C \ll \mathbb{E}[X_n]^2$$

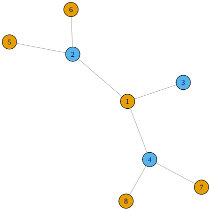
$$\mathbb{P}(X_n = 0) \leq \mathbb{P}(|X_n - \mathbb{E}[X_n]| \geq \mathbb{E}[X_n]) \leq \frac{\mathrm{Var}(X_n)}{\mathbb{E}[X_n]^2} \xrightarrow[n \to +\infty]{} 0$$

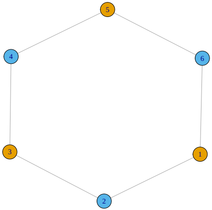
Plan

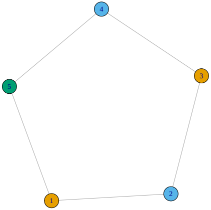
Coloriage d’un graphe G

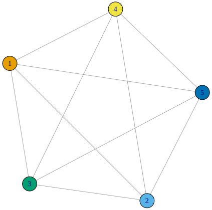
Colorier un graphe G : attribuer une couleur à chaque sommet de G de sorte que deux sommets adjacents n’aient pas la même couleur

Nombre chromatique $\chi(G)$: plus petit entier k tel qu’il existe un coloriage de G à k couleurs


6
5
2
3
1
4
7
8


5
4
6
3
1
2


4
5
3
1
2


4
1
5
3
2

Coloriage d'un graphe G

Application : Planifier n réunions
Chacune réquérant la présence d'un ensemble de participants déterminé
Chaque réunion dure une demi-journée.

$\leadsto$ Combien de demi-journées seront nécessaires ?

- Sommets : réunions
- Arête entre deux réunions ssi elles ont un participant commun
- Couleurs : demi-journées

En général, problème difficile (3-COL NP-complet) !

Coloriage d’un graphe G

Application :

Planifier n réunions
Chacune réquérant la présence d’un ensemble de participants déterminé
Chaque réunion dure une demi-journée.

$\rightsquigarrow$ Combien de demi-journées seront nécessaires ?

- Sommets : réunions
- Arête entre deux réunions ssi elles ont un participant commun
- Couleurs : demi-journées

En général, problème difficile (3-COL NP-complet) !

Coloriage d'un graphe G

Application :

Planifier n réunions
Chacune réquérant la présence d'un ensemble de participants déterminé
Chaque réunion dure une demi-journée.

$\rightsquigarrow$ Combien de demi-journées seront nécessaires ?

- Sommets : réunions
- Arête entre deux réunions ssi elles ont un participant commun
- Couleurs : demi-journées

En général, problème difficile (3-COL NP-complet) !

Des graphes à grand nombre chromatique sans triangles

Mycielski (1955) “Sur le coloriage des graphes”

SUR LE COLORIAGE DES GRAPHS

PAR

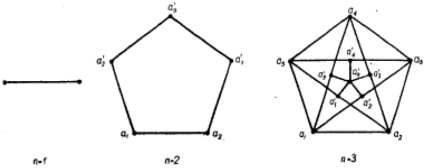
J. MYCIELSKI (WROCŁAW)

Un *graph fini* est un ensemble fini (situé dans l'espace euclidien à trois dimensions) de points dont certains sont joints par des arcs simples qui n'ont pas de points communs (sauf — peut-être — de points finaux)[1]. Un graph composé de trois points joints deux-à-deux s'appelle un *triangle*.

Colorier un graph au moyen de n couleurs — veut dire — peindre chacun de ses points par une de ces couleurs, les points joints par un arc simple étant peints de couleurs différentes.

Il est évident qu'un graph qui contient m points joints deux-à-deux ne peut pas être colorié par moins que m couleurs. Je me propose de démontrer dans cette note qu'un théorème inverse serait faux. En effet je vais prouver le théorème suivant:

THÉORÈME. *Pour chaque nombre naturel n il existe un graph fini ne contenant aucun triangle, qui ne peut pas être colorié au moyen de n couleurs.*



P 130. Existe-t-il un graph composé d'un nombre de points plus petit que $3\cdot 2^{n-1}-1$, qui ne contient aucun triangle et qui ne peut pas être colorié par n couleurs?

Appelons *circuit* une suite $a_1, a_2, \ldots, a_k$ de points d'un graph lorsque les points a_1 et a_2, a_2 et $a_3, \ldots, a_k$ et a_1 sont joints.

Le problème suivant est une généralisation du problème résolu dans ce travail:

P 131. Existe-t-il pour chaque couple de nombres naturels n et $m \geqslant 3$ un graph fini qui ne peut être colorié par n couleurs et qui ne contient, pour $k=3,4,\ldots,m$, aucun circuit à k points.

Des graphes à grand nombre chromatique et « localement arborescents »

Maille d'un graphe G :

$$\gamma(G) := \text{ longueur minimale d'un cycle}$$

Théorème (Erdős 1960)

Pour tous k et l entiers, il existe un graphe G tel que

$$\gamma(G) > l \quad \text{et} \quad \chi(G) > k$$

Des graphes à grand nombre chromatique et « localement arborescents »

Maille d'un graphe G :

$$\gamma(G) := \text{ longueur minimale d'un cycle}$$

Théorème (Erdős 1960)

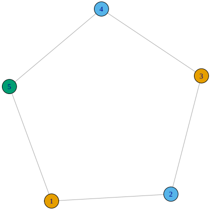
Pour tous k et l entiers, il existe un graphe G tel que

$$\gamma(G) > l \quad \text{et} \quad \chi(G) > k$$

Démonstration

Nombre d'indépendance

$\alpha(G) :=$ Taille du plus grand ensemble de sommets sans aucune arête



$$n \leq \chi(G)\alpha(G)$$

Démonstration

On pose $p = n^{\frac{1-\varepsilon}{l}-1}$ avec $\varepsilon \in]0,1[$

Premier moment de $X_l :=$ nombre de cycles de longueur $\leq l$ de $\mathcal{G}(n,p)$

$$\mathbb{E}[X_l] \leq \sum_{i=3}^{l} n^i p^i \leq l(np)^l = ln^{1-\varepsilon} \ll n$$

Donc (Markov)

$$\mathbb{P}(X_l \geq n/2) \leq \frac{2\mathbb{E}[X_l]}{n} \ll 1$$

Démonstration

On pose $p = n^{\frac{1-\varepsilon}{l}-1}$ avec $\varepsilon \in]0,1[$

Premier moment de $X_l :=$ nombre de cycles de longueur $\leq l$ de $\mathcal{G}(n,p)$

$$\mathbb{E}[X_l] \leq \sum_{i=3}^{l} n^i p^i \leq l(np)^l = ln^{1-\varepsilon} \ll n$$

Donc (Markov)

$$\mathbb{P}(X_l \geq n/2) \leq \frac{2\mathbb{E}[X_l]}{n} \ll 1$$

Démonstration

Contrôle de $\alpha(\mathcal{G}(n,p))$:

$$\begin{aligned}
\mathbb{P}(\alpha(\mathcal{G}(n,p)) \geq x) &\leq \sum_{A:|A|=x} \mathbb{P}(\forall i \neq j \in A,\ i \not\sim j) \\
&\leq \binom{n}{x}(1-p)^{\binom{x}{2}} \\
&\leq n^x e^{-px\frac{x-1}{2}} \\
&= e^{-x\left(\frac{p(x-1)}{2} - \log n\right)} \\
&\ll 1
\end{aligned}$$

si $x \geq \frac{3}{p}\log n$

Démonstration

Contrôle de $\alpha(\mathcal{G}(n,p))$:

$$\begin{aligned}
\mathbb{P}(\alpha(\mathcal{G}(n,p)) \geq x) &\leq \sum_{A:|A|=x} \mathbb{P}(\forall i \neq j \in A,\ i \not\sim j) \\
&\leq \binom{n}{x}(1-p)^{\binom{x}{2}} \\
&\leq n^x e^{-px\frac{x-1}{2}} \\
&= e^{-x\left(\frac{p(x-1)}{2} - \log n\right)} \\
&\ll 1
\end{aligned}$$

si $x \geq \frac{3}{p} \log n$

Démonstration

Contrôle de $\alpha(\mathcal{G}(n,p))$:

$$\begin{aligned} \mathbb{P}(\alpha(\mathcal{G}(n,p)) \geq x) &\leq \sum_{A:|A|=x} \mathbb{P}(\forall i \neq j \in A,\ i \not\sim j) \\ &\leq \binom{n}{x}(1-p)^{\binom{x}{2}} \\ &\leq n^x e^{-px\frac{x-1}{2}} \\ &= e^{-x\left(\frac{p(x-1)}{2} - \log n\right)} \\ &\ll 1 \end{aligned}$$

si $x \geq \frac{3}{p} \log n$

Démonstration

Pour n assez grand, il existe G tel que $X_l \leq n/2$ et $\alpha(G) \leq \frac{3}{p} \log n$.

En supprimant $n/2$ sommets de G on obtient G' avec :

- au moins $n/2$ sommets
- $\gamma(G') > l$
- $\chi(G') \geq \frac{n/2}{\alpha(G')} \geq \frac{n}{2\alpha(G)} \geq \frac{np}{6 \log n} = \frac{n^{\frac{1-\varepsilon}{l}}}{6 \log n} \gg 1$

$$\# \text{ sommets de } G' \leq \chi(G')^{\frac{l}{1-\varepsilon}}$$

Démonstration

Pour n assez grand, il existe G tel que $X_l \leq n/2$ et $\alpha(G) \leq \frac{3}{p} \log n$.

En supprimant $n/2$ sommets de G on obtient G' avec :

- au moins $n/2$ sommets
- $\gamma(G') > l$
- $\chi(G') \geq \frac{n/2}{\alpha(G')} \geq \frac{n}{2\alpha(G)} \geq \frac{np}{6 \log n} = \frac{n^{\frac{1-\varepsilon}{l}}}{6 \log n} \gg 1$

$$\# \text{ sommets de } G' \leq \chi(G')^{\frac{l}{1-\varepsilon}}$$

Démonstration

Pour n assez grand, il existe G tel que $X_l \leq n/2$ et $\alpha(G) \leq \frac{3}{p} \log n$.

En supprimant $n/2$ sommets de G on obtient G' avec :

- au moins $n/2$ sommets
- $\gamma(G') > l$
- $\chi(G') \geq \frac{n/2}{\alpha(G')} \geq \frac{n}{2\alpha(G)} \geq \frac{np}{6 \log n} = \frac{n^{\frac{1-\varepsilon}{l}}}{6 \log n} \gg 1$

$$\# \text{ sommets de } G' \leq \chi(G')^{\frac{l}{1-\varepsilon}}$$

Démonstration

Pour n assez grand, il existe G tel que $X_l \leq n/2$ et $\alpha(G) \leq \frac{3}{p} \log n$.

En supprimant $n/2$ sommets de G on obtient G' avec :

- au moins $n/2$ sommets
- $\gamma(G') > l$
- $\chi(G') \geq \frac{n/2}{\alpha(G')} \geq \frac{n}{2\alpha(G)} \geq \frac{np}{6 \log n} = \frac{n^{\frac{1-\varepsilon}{l}}}{6 \log n} \gg 1$

$$\# \text{ sommets de } G' \leq \chi(G')^{\frac{l}{1-\varepsilon}}$$

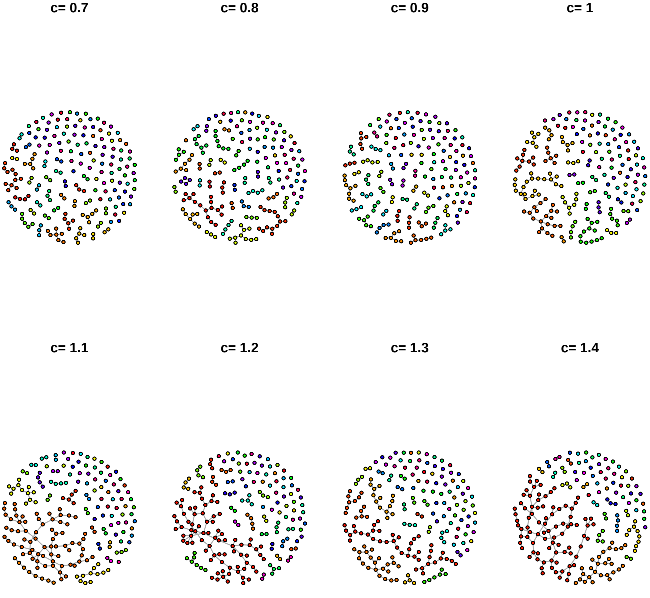
Plan

Evolution des petits sous-graphes

Nombre chromatique et “méthode probabiliste”

Composante géante

Applications ?


c= 0.7
c= 0.8
c= 0.9
c= 1
c= 1.1
c= 1.2
c= 1.3
c= 1.4

Naissance de la composante géante

L_i : taille de la i-ème plus grande composante connexe

Théorème

Soit $c > 0$ *et* $p = \frac{c}{n}$

- *Si* $c < 1$,

$$L_1 \simeq \log n \quad a.p.s$$

- *Si* $c = 1$ *et* $k \in \mathbb{N}$,

$$L_k \simeq n^{2/3} \quad a.p.s$$

- *Si* $c > 1$,

$$L_1 \simeq n \quad et \quad L_2 \simeq \log n \quad a.p.s$$

Taille de la composante géante en fonction de c

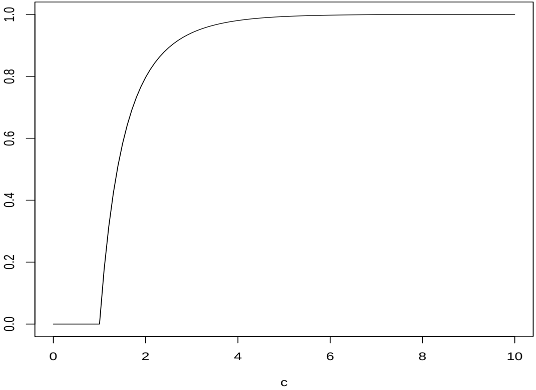
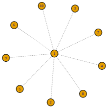


Figure – ($p = \frac{c}{n}$) – Limite de $\frac{L_1}{n}$ en fonction de c

Première heuristique

Soit D_n degré du sommet 1



$$D_n \sim \mathcal{B}(n-1, p)$$

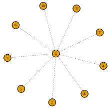
$$\mathbb{E}[D_{n-1}] = (n-1)p \sim c$$

$$\mathbb{E}[\#\text{sommets à distance 2}] \sim c^2$$

$$\mathbb{E}[\#\text{sommets à distance k2}] \sim c^k$$

Première heuristique

Soit D_n degré du sommet 1



$$D_n \sim \mathcal{B}(n-1, p)$$

$$\mathbb{E}[D_{n-1}] = (n-1)p \sim c$$

$$\mathbb{E}[\#\text{sommets à distance 2}] \sim c^2$$

$$\mathbb{E}[\#\text{sommets à distance k2}] \sim c^k$$

Loi des évènements rares

Proposition

Soit $c > 0$. Si $p = \frac{c}{n}$ et $k \in \mathbb{N}$

$$\mathbb{P}(D_n = k) \xrightarrow[n \to +\infty]{} e^{-c} \frac{c^k}{k!}$$

$$D_n \xrightarrow[n \to +\infty]{\mathcal{L}} \mathcal{P}(c)$$

Loi des évènements rares : démo

$$
\begin{aligned}
\mathbb{P}(D_n = k) &= \binom{n-1}{k} p^k (1-p)^{n-k} \\
&= \left(\frac{c}{n}\right)^k \frac{n(n-1)\dots(n-k+1)}{k!} \left(1 - \frac{c}{n}\right)^{n-k} \\
&\sim \frac{c^k}{k!} \left(1 - \frac{c}{n}\right)^{n-k} \\
&\sim \frac{c^k}{k!} \left(1 - \frac{c}{n}\right)^{n} \\
&\xrightarrow[n \to +\infty]{} e^{-c} \frac{c^k}{k!}
\end{aligned}
$$

Loi des évènements rares : démo

$$\begin{aligned}
\mathbb{P}(D_n = k) &= \binom{n-1}{k} p^k (1-p)^{n-k} \\
&= \left(\frac{c}{n}\right)^k \frac{n(n-1)\dots(n-k+1)}{k!} \left(1 - \frac{c}{n}\right)^{n-k} \\
&\sim \frac{c^k}{k!} \left(1 - \frac{c}{n}\right)^{n-k} \\
&\sim \frac{c^k}{k!} \left(1 - \frac{c}{n}\right)^{n} \\
&\xrightarrow[n \to +\infty]{} e^{-c} \frac{c^k}{k!}
\end{aligned}$$

Loi des évènements rares : démo

$$\begin{aligned}
\mathbb{P}(D_n = k) &= \binom{n-1}{k} p^k (1-p)^{n-k} \\
&= \left(\frac{c}{n}\right)^k \frac{n(n-1)\dots(n-k+1)}{k!} \left(1 - \frac{c}{n}\right)^{n-k} \\
&\sim \frac{c^k}{k!} \left(1 - \frac{c}{n}\right)^{n-k} \\
&\sim \frac{c^k}{k!} \left(1 - \frac{c}{n}\right)^{n} \\
&\xrightarrow[n\to+\infty]{} e^{-c} \frac{c^k}{k!}
\end{aligned}$$

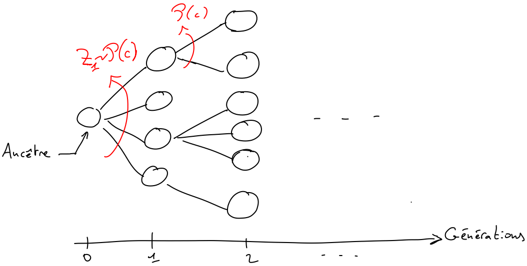
Loi des évènements rares : démo

$$
\begin{aligned}
\mathbb{P}(D_n = k) &= \binom{n-1}{k} p^k (1-p)^{n-k} \\
&= \left(\frac{c}{n}\right)^k \frac{n(n-1)\dots(n-k+1)}{k!} \left(1-\frac{c}{n}\right)^{n-k} \\
&\sim \frac{c^k}{k!}\left(1-\frac{c}{n}\right)^{n-k} \\
&\sim \frac{c^k}{k!}\left(1-\frac{c}{n}\right)^{n} \\
&\xrightarrow[n\to+\infty]{} e^{-c}\frac{c^k}{k!}
\end{aligned}
$$

Deuxième heuristique : Arbres de Galton-Watson

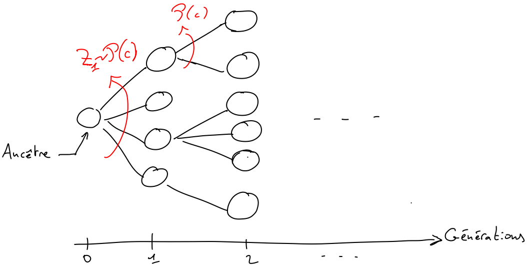
Z_n :=taille de la n-ème génération

$\mathbb{P}[Z_{n+1} = 0] = g(\mathbb{P}[Z_n = 0])$
où $g(z) = \mathbb{E}[z^{Z_1}] = e^{c(z-1)}$



Deuxième heuristique : Arbres de Galton-Watson

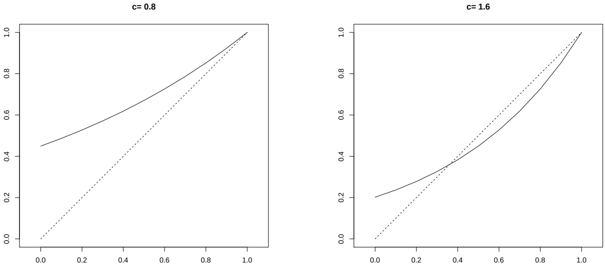
Z_n :=taille de la n-ème génération



$$\mathbb{P}[Z_{n+1} = 0] = g(\mathbb{P}[Z_n = 0])$$
où $g(z) = \mathbb{E}[z^{Z_1}] = e^{c(z-1)}$

Deuxième heuristique : Arbres de Galton-Watson

Donc $\mathbb{P}[\mathcal{T} \text{ est fini }] = \lim_{n\to+\infty} \mathbb{P}[Z_n = 0] = \lim_{n\to+\infty} g^{(n)}(0)$

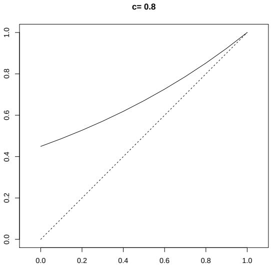


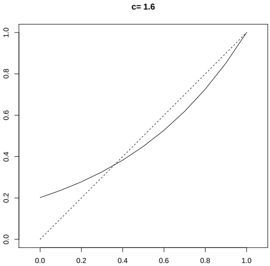
Théorème

- *Si $c \leq 1$ $\mathcal{T}$ est fini p.s*
- *Si $c > 1$, $\mathbb{P}(\mathcal{T}$ est fini $)$ plus petit point fixe de g (strictement inférieur à 1).*

Deuxième heuristique : Arbres de Galton-Watson

Donc $\mathbb{P}[\mathcal{T}$ est fini $] = \lim_{n\to+\infty} \mathbb{P}[Z_n = 0] = \lim_{n\to+\infty} g^{(n)}(0)$



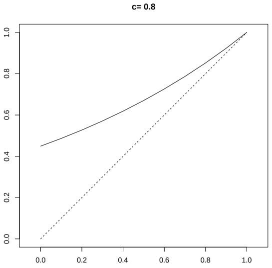


Théorème

- *Si $c \leq 1$ $\mathcal{T}$ est fini p.s*
- *Si $c > 1$, $\mathbb{P}(\mathcal{T}$ est fini $)$ plus petit point fixe de g (strictement inférieur à 1).*

Deuxième heuristique : Arbres de Galton-Watson

Donc $\mathbb{P}[\mathcal{T} \text{ est fini }] = \lim_{n\to+\infty} \mathbb{P}[Z_n = 0] = \lim_{n\to+\infty} g^{(n)}(0)$



Théorème

- *Si $c \leq 1$ $\mathcal{T}$ est fini p.s*
- *Si $c > 1$, $\mathbb{P}(\mathcal{T}$ est fini $)$ plus petit point fixe de g (strictement inférieur à 1).*

Retour à $\mathcal{G}(n, p)$: cas critique, aspect métrique

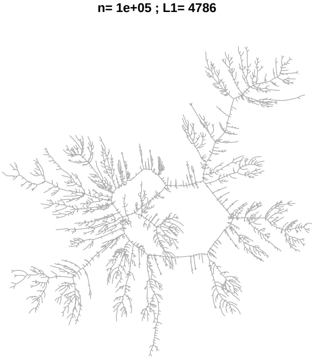
Théorème

On divise les distances par $n^{1/3}$, et on prend $p = \frac{1}{n}$. Les plus grandes composantes de $\mathcal{G}(n, p)$ convergent vers des graphes (aléatoires) "fractals" (et de dimension de Hausdorff $= 2$)

Retour à $\mathcal{G}(n, p)$: cas critique, aspect métrique

Théorème

On divise les distances par $n^{1/3}$, et on prend $p = \frac{1}{n}$. Les plus grandes composantes de $\mathcal{G}(n, p)$ convergent vers des graphes (aléatoires) "fractals" (et de dimension de Hausdorff $= 2$)



Plan

Stochastic Block Model : modèle de communautés

Paramètres des tailles des communautés : $q_1 + \ldots + q_r = 1$

- Étape 1 : Z_i, $i = 1 \ldots, n$ i.i.d, $\mathbb{P}(Z_i = j) = q_j$.
- Étape 2 :

$$\mathbb{P}[i \sim j | Z] = \kappa(Z_i, Z_j)$$

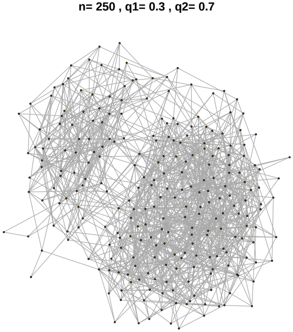
Stochastic Block Model : modèle de communautés

Paramètres des tailles des communautés : $q_1 + \ldots + q_r = 1$

- Étape 1 : Z_i, $i = 1 \ldots, n$ i.i.d, $\mathbb{P}(Z_i = j) = q_j$.
- Étape 2 :

$$\mathbb{P}[i \sim j | Z] = \kappa(Z_i, Z_j)$$

$$\kappa = \begin{pmatrix} .08 & .005 \\ .005 & .05 \end{pmatrix}$$



Objectifs statistiques :

- Tester l'existence de communautés distinctes (Hypothèse nulle : Erdős-Rényi)
- Trouver le nombre de communautés (r inconnu)
- Séparer la population en groupes homogènes
- Estimer les paramètres q et κ

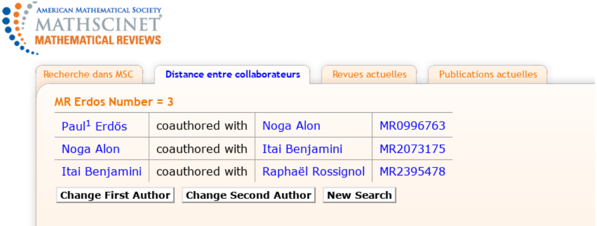
Généralisation : Latent space model

Objectifs statistiques :

- Tester l'existence de communautés distinctes (Hypothèse nulle : Erdős-Rényi)
- Trouver le nombre de communautés (r inconnu)
- Séparer la population en groupes homogènes
- Estimer les paramètres q et κ

Généralisation : Latent space model

Merci !


American Mathematical Society
MathSciNet
Mathematical Reviews
Recherche dans MSC
Distance entre collaborateurs
Revues actuelles
Publications actuelles
MR Erdos Number = 3
Paul[1] Erdős coauthored with Noga Alon MR0996763
Noga Alon coauthored with Itai Benjamini MR2073175
Itai Benjamini coauthored with Raphaël Rossignol MR2395478
Change First Author
Change Second Author
New Search