



UNIVERSITE KASDI MERBAH  
OUARGLA

Faculté des mathématiques et sciences de la  
matière



DEPARTEMENT DE MATHÉMATIQUES

MASTER

Spécialité : Mathématiques

Option : Probabilité et Statistique

Par : REMMA KHAOULA

Thème

Quelques propriétés des graphes aléatoires

Soutenu publiquement le : 30/05/2017

Devant le jury composé de :

BAHEDDI Aissa	Professeur. UKMOuargla	Président
MEDDI Fatima	M. C. B. UKMOuargla	Examineur
BOUANANE Khadra	M. C. B. UKMOuargla	Rapporteur

---

## *Remerciements*

---

Mes remerciements vont premièrement à **Allah** le tout puissant pour la volonté, la santé et la patience qu'il m'a donnée pour achever ce travail.

Je tiens exprimer ma profonde gratitude, mes sincères et chaleureux remerciements mon encadreur Mme **Bouanane Khadra**, pour la confiance qu'elle m'a accordée en acceptant d'encadrer ce travail et pour ses multiples conseils.

Je remercie également tous mes enseignants du département de mathématiques à l'université de Ouargla.

Je remercie également tous les membres de jury pour le temps consacré pour examiner et évaluer ce travail.

Enfin, je me permets également de remercier mes parents, et toute ma famille qui a supporté toutes les difficultés pour me soutenir tout au long de mes études.

---

## Table des matières

---

<b>Introduction générale</b>	<b>5</b>
<b>1 Théorie des Graphes</b>	<b>7</b>
1.1 Notion de graphe . . . . .	7
1.2 Quelques types de graphes . . . . .	11
1.3 Graphe complémentaire, biparti complémentaire . . . . .	15
1.4 Couplage, stable et clique . . . . .	15
1.5 Nombre chromatique . . . . .	15
<b>2 Méthode Probabiliste</b>	<b>17</b>
2.1 Préliminaires mathématiques . . . . .	17
2.2 Méthode probabiliste . . . . .	25
<b>3 Graphes Aléatoires</b>	<b>32</b>
3.1 Introduction . . . . .	32
3.2 Modèle d'Erdős-Rényi . . . . .	32
3.3 Connexité . . . . .	36
3.4 Evolution d'un graphe aléatoire . . . . .	42
3.5 Cycle hamiltonien . . . . .	45
3.6 Stabilité d'un graphe aléatoire . . . . .	46
3.7 Nombre chromatique d'un graphe aléatoire . . . . .	48
<b>Conclusion générale</b>	<b>53</b>
<b>Bibliographie</b>	<b>54</b>

---

## Table des figures

---

1.1	Un graphe non orienté. . . . .	7
1.2	Matrice d'adjacence du graphe de la Figure 1.1. . . . .	8
1.3	Matrice d'incidence. . . . .	8
1.4	Graphe avec $\Delta(G) = 5$ et $\delta(G) = 3$ . . . . .	9
1.5	Graphe $G$ , Graphe partiel et Sous-graphe de $G$ . . . . .	11
1.6	Graphe complet $k_6$ . . . . .	12
1.7	Graphe biparti complet $K_{2,3}$ . . . . .	12
1.8	Graphe planaire topologique(une carte). . . . .	13
1.9	$W = \{v_3\}$ est un séparateur minimal dans $G$ . . . . .	13
1.10	Arbre. . . . .	14
1.11	Fôret. . . . .	14
1.12	Graphe $G$ . . . . .	14
1.13	Un arbre couvrant. . . . .	14
1.14	Graphe admettant une 4-coloration avec $\chi(G) = 4$ . . . . .	15
3.1	Exemple d'un graphe aléatoire de type Erdős-Rényi. . . . .	33
3.2	Étapes I et II. Les graphes $G(1000, 0.0005)$ et $G(1000, 0.00095)$ sont représentés. [36] . . . . .	43
3.3	Étapes III et IV. Les graphes $G(1000, 0.001)$ et $G(1000, 0.0015)$ sont représentés. La composante géante est née.[36] . . . . .	44
3.4	Étapes V et VI. Les graphes $G(1000, 0.004)$ et $G(1000, 0.007)$ sont représentés. Le graphe final est connecté.[36] . . . . .	44

---

## Introduction

---

La théorie des graphes est un très vaste domaine, en évolution constante tant du point de vue de la recherche fondamentale que celui des applications. Son développement est d'ailleurs assez similaire à celui de la théorie des probabilités dont beaucoup de résultats sont dûs à l'effort de compréhension des jeux de hasard.

L'introduction de probabilités dans les graphes n'a vu le jour qu'à la fin des années 1950 avec les travaux des mathématiciens hongrois Paul Erdős et Alfréd Rényi, qui étudient en 1959 un modèle qui porte leur nom depuis (Erdős and Rényi, 1959).

Un graphe aléatoire est un ensemble de points, ou de sommets, avec des lignes ou des arêtes, en reliant des paires au hasard. L'étude des graphes aléatoires a une longue histoire. La théorie des graphes aléatoire est devenue l'un des piliers des mathématiques discrètes modernes et a produit un nombre prodigieux de résultats, dont beaucoup sont très ingénieux, en décrivant les propriétés statistiques des graphes, tels que la distribution de la taille des composants, l'existence et la taille d'un composant géant.

Les graphes aléatoires ont été introduits pour prouver certains résultats combinatoires sur les graphes, tel qu'une borne inférieure sur les nombres de Ramsey par exemple. Les graphes aléatoires peuvent être également utilisés pour évaluer la complexité en moyenne d'algorithmes ou encore pour modéliser de vrais réseaux de la vie de tous les jours (réseaux sociaux, internet, etc).

L'étude des graphes aléatoires a suscité de nombreuses recherches et une très dense littérature. On peut citer à titre non exhaustif [11, 7, 9, 17, 19, 27]. Notre objectif est donc d'introduire les graphes aléatoires d'Erdős et Rényi, et de présenter quelques unes des diverses propriétés qui existent dans la littérature, en mettant en avant l'utilisation des outils probabilistes comme techniques de preuve de ces propriétés.

Ce travail se divise en trois parties capitales.

Nous présentons dans le premier chapitre les définitions et notations des concepts de la

théorie des graphes qui nous seront utiles pour la bonne compréhension de ce travail.

Dans le deuxième chapitre, De plus des définitions des notions élémentaires de la théorie des probabilités, nous introduisons la méthode probabiliste en illustrant son utilisation comme technique de preuve de plusieurs résultats combinatoires.

Le troisième chapitre, est consacré à l'étude de graphes aléatoires d'Erdős et Rényi. Nous présentons certains résultats théoriques qui existent dans la littérature, en mettant en avant les techniques de preuve utilisées pour montrer ces résultats.

# Chapitre 1

---

## Théorie des Graphes

---

Dans ce chapitre, nous utilisons les notations et définitions données dans [12, 31] afin d'introduire quelques concepts de la théorie des graphes qui sont nécessaires à la bonne compréhension de ce manuscrit.

### 1.1 Notion de graphe

Un *graphe non orienté*  $G$  est défini par deux ensembles disjoints notés  $V(G)$  et  $E(G)$  ( $V$  et  $E$  s'il n'y a aucune confusion) où  $E$  est constitué de paires (non ordonnées) de  $V$ .

Les éléments de  $V$  et  $E$  sont appelés *sommets* et *arêtes* de  $G$ . On appelle alors  $|V| = n$  et  $|E| = m$ , *l'ordre* et la *taille* du graphe  $G$  respectivement.

On notera par  $uv$  au lieu de  $\{u, v\}$  une arête de  $E$ . Dans ce cas, les sommets  $u$  et  $v$  constituent les *extrémités* de l'arête  $uv$ .

Une *boucle* est une arête de type  $uu$ . On appelle *arêtes multiple* des arêtes qui relient les mêmes sommets  $u$  et  $v$ . Un graphe  $G$  est dit *simple* s'il ne contient ni boucles ni arêtes multiples.

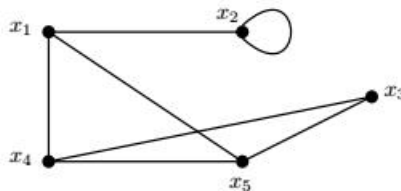


FIGURE 1.1 – Un graphe non orienté.

### 1.1.1 Adjacence et incidence

Deux sommets  $u$  et  $v$  sont *adjacents* s'il existe une arête qui les relie. De même, deux arêtes sont *adjacentes* si elles ont exactement une extrémité en commun. On dira qu'un sommet est *incident* à une arête ou une arête est *incidente* à un sommet, s'il constitue une de ses extrémités.

**Matrice d'adjacence.** On peut représenter un graphe par une *matrice d'adjacence*, notée  $A$ , qui est une matrice de dimension  $(n \times n)$ , où chaque élément  $A_{ij}$  désigne le nombre d'arêtes reliant les sommets  $i$  et  $j$ .

$$\sum_i A_{ij} = \sum_i A_{ji} = d(j)$$

$$M = \begin{pmatrix} 0 & 1 & 0 & 1 & 1 \\ 1 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 1 \\ 1 & 0 & 1 & 0 & 1 \\ 1 & 0 & 1 & 1 & 0 \end{pmatrix}$$

FIGURE 1.2 – Matrice d'adjacence du graphe de la Figure 1.1.

**Matrice d'incidence.** À tout graphe non orienté  $G = (V, E)$ , on peut associer une matrice  $M$  de  $n$  lignes et  $m$  colonnes. Où  $n$  est le nombre de sommets dans  $G$  et  $m$  est le nombre d'arêtes dans  $G$ .

$M_{ij}$  représente le nombre de fois où le sommet  $i$  est incident à l'arête  $j$ . Les éléments de  $M$  sont dans  $(0, 1, 2)$ .

Si deux colonnes  $j_1$  et  $j_2$  sont identiques alors les arêtes  $j_1$  et  $j_2$  sont parallèles. Si un élément  $M_{ij} = 2$  alors l'arête  $j$  est une boucle.

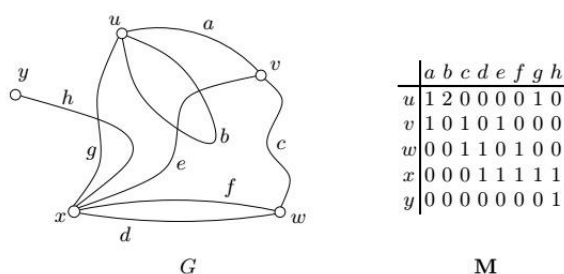


FIGURE 1.3 – Matrice d'incidence.



### 1.1.2 Degré d'un sommet

Le *degré d'un sommet*  $v$  dans un graphe  $G$ , noté  $d_G(v)$ , est le nombre d'arêtes de  $G$  incidentes avec  $v$ , chaque boucle comptant comme deux arêtes. En particulier, si  $G$  est un graphe simple,  $d_G(v)$  est le nombre de voisins de  $v$  dans  $G$ . Un sommet de degré zéro est appelé *sommet isolé*. On désigne par  $\delta(G)$  et  $\Delta(G)$ , les *degrés minimum* et *maximum* des sommets de  $G$ , et par  $d(G)$  leur degré moyen :

$$\delta(G) = \min \{d(v), v \in V\}$$

$$\Delta(G) = \max \{d(v), v \in V\}$$

$$d(G) = \frac{1}{n} \sum_{v \in V} d(v)$$

#### Exemple 1 .

Le graphe de la Figure 1.4 ci-dessous, est de degré maximum  $\Delta(G) = 5$ , de degré minimum  $\delta(G) = 3$ , son degré moyen est égal à  $10/3$ .

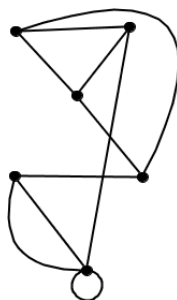


FIGURE 1.4 – Graphe avec  $\Delta(G) = 5$  et  $\delta(G) = 3$  .

Le théorème suivant établit une relation fondamentale entre les degrés des sommets d'un graphe et son nombre d'arêtes.

#### Théorème 1.1.1 .

Pour tout graphe  $G$ ,

$$\sum_{v \in V} d(v) = 2m$$

*Preuve.*

Considérons la matrice d'incidence  $M$  de  $G$ . La somme des entrées dans la ligne correspondant au sommet  $v$  est exactement  $d(v)$ . Par conséquent  $\sum_{v \in V} d(v)$  est la somme de toutes les entrées de  $M$ . Mais cette somme vaut également  $2m$ , car la somme des entrées de chacune des  $m$  colonnes

de  $M$  vaut 2, une arête ayant deux extrémités. □

### 1.1.3 Voisins et voisins communs

On appelle *ensemble de voisins* du sommet  $u$ , noté  $N_G(u)$  ( $N(u)$  s'il n'y a pas confusion), l'ensemble de tous les sommets différents de  $u$  qui lui sont adjacents.

Si  $G$  est simple alors  $|N(u)| = d_G(u) \leq n - 1$ .

On appelle *ensemble de voisins communs* du sous-ensemble de sommets  $S = \{u_1, u_2, \dots, u_k\}$ , noté  $N_G(S)$  ou  $N(S)$  (s'il n'y a pas confusion), l'ensemble de sommets qui sont voisins communs à tous les sommets de l'ensemble  $S$ .

Autrement dit,  $N(S) = N(u_1) \cap N(u_2) \cap \dots \cap N(u_k)$ .

### 1.1.4 Chaînes et cycles

Une *chaîne* entre deux sommets  $u$  et  $v$  de longueur  $p$  est une séquence de  $p$  sommets  $u_0, u_1, \dots, u_p$ , tels que deux sommets consécutifs sont adjacents, avec  $u = u_0$  et  $v = u_p$ .

Une chaîne est dite *élémentaire*, si elle ne passe pas deux fois par un même sommet, elle est dite simple si elle n'utilise pas deux fois une même arête.

Une chaîne dont les extrémités sont confondues est appelée *cycle*.

Dans un graphe  $G$ , on appelle circonférence de  $G$ , notée  $\ell(G)$ , la longueur d'un plus petit cycle de  $G$ .

### 1.1.5 Connexité

Un graphe  $G = (V, E)$  est dit *connexe* si chaque paire de ses sommets est reliée par une chaîne.

Pour un graphe quelconque, il est facile de remarquer que la relation binaire  $R$  définie sur  $V \times V$  par :  $uRv \iff u = v$  ou il existe une chaîne entre  $u$  et  $v$ , est une relation d'équivalence. Les sous-graphes de  $G$  induits par les classes d'équivalence de cette relation sont appelés composantes connexes de  $G$ . Autrement dit, un graphe est connexe s'il possède une seule classe d'équivalence.

### 1.1.6 Connectivité

Un graphe  $G$  est dit, *k-connexe* (resp. *k-arête connexe*) si la suppression de moins de  $k$  sommets (resp. arêtes) ne déconnecte pas  $G$ .

On appelle alors *connectivité* (resp. *arête-connectivité*) d'un graphe connexe  $G$ , notée  $\kappa(G)$  (resp.  $\lambda(G)$ ), le plus petit nombre de sommets (resp. d'arêtes) dont la suppression déconnecte le graphe  $G$ .

Rappelons l'inégalité de Whitney [38] reliant la connectivité, l'arête-connectivité et le degré minimum d'un graphe  $G$  donné.

$$\kappa(G) \leq \lambda(G) \leq \delta(G).$$

Un *ensemble d'articulation minimum* est alors, le plus petit ensemble de sommets qu'on doit supprimer pour déconnecter le graphe  $G$ .

### 1.1.7 Sous-graphe

Soient  $G = (V, E)$  et  $H = (W, F)$  deux graphes. On dira que  $H$  est un *sous-graphe* de  $G$ , si  $W \subset V$  et  $F \subset E$ .

Si  $H$  est un sous-graphe de  $G$  tel que  $F$  est constitué de toutes les arêtes de  $G$  reliant les sommets de  $W$ , alors  $H$  est appelé *sous-graphe induit* par l'ensemble de sommets  $W$ .

On dira que  $H$  est un *graphe partiel* de  $G$  si  $W = V$ .

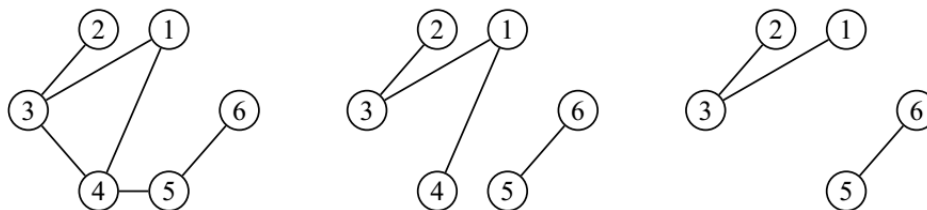


FIGURE 1.5 – Graphe  $G$ , Graphe partiel et Sous-graphe de  $G$ .

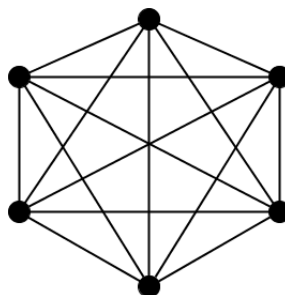
## 1.2 Quelques types de graphes

### 1.2.1 Graphe régulier

Un graphe est *régulier*, si tous ses sommets ont le même degré. C'est-à-dire  $\delta(G) = \Delta(G) = k$ , on dit alors que le graphe est régulier de degré  $k$  ou  $k$ -régulier.

### 1.2.2 Graphe complet

Un graphe *complet* d'ordre  $n$ , noté  $K_n$ , est un graphe simple tel que chaque paire de sommets sont adjacents. Autrement dit, on a  $d_G(u) = n - 1$  pour tout sommet  $u \in V$ .

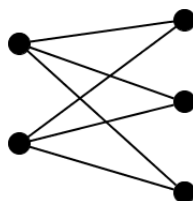
FIGURE 1.6 – Graphe complet  $k_6$ .

### 1.2.3 Graphe biparti

Un graphe  $G = (V, E)$  est dit *biparti*, si  $V$  admet une partition en deux sous-ensembles  $V_1$  et  $V_2$ , tels que chaque arête de  $E$  a une extrémité dans  $V_1$  et une extrémité dans  $V_2$ .

Un graphe biparti est caractérisé par l'inexistence de cycles de longueur impaire.

Un graphe biparti complet noté  $K_{p,q}$ , est un graphe biparti où chaque sommet de  $V_1$  est adjacent à chaque sommet de  $V_2$ .

FIGURE 1.7 – Graphe biparti complet  $K_{2,3}$ .

#### Théorème 1.2.1 .

*Un graphe est biparti si et seulement s'il ne contient aucun cycle de longueur impaire.*

### 1.2.4 Graphe planaire

On dit qu'un graphe est *planaire* si on peut le dessiner dans le plan de sorte que ses arêtes ne se croisent pas. Rappelons que les arêtes ne sont pas forcément rectilignes. Une carte, ou graphe planaire topologique, est une représentation particulière d'un multigraphe planaire fini. On dit qu'une carte est connexe si son graphe l'est. Une carte divise le plan en plusieurs régions. Le degré d'une région  $r$ , noté  $d(r)$ , est la longueur de la chaîne fermée minimum passant par tous les sommets qui délimitent cette région.

Par exemple, la carte ci-dessous, avec sept sommets et neuf arêtes, divise le plan en quatre régions (A,B,C,D). Trois régions sont limitées alors que la quatrième (D), extérieure au diagramme, ne l'est pas.

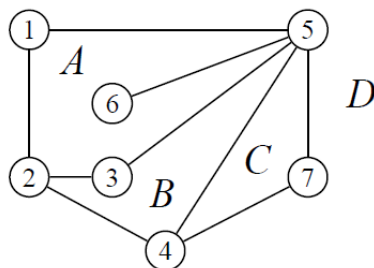


FIGURE 1.8 – Graphe planaire topologique (une carte).

### 1.2.5 Graphes hamiltoniens

On appelle *cycle hamiltonien* d'un graphe  $G$  un cycle passant une et une seule fois par chacun des sommets de  $G$ . Un graphe est dit *hamiltonien* s'il possède un cycle hamiltonien.

On appelle *chaîne hamiltonienne* d'un graphe  $G$  une chaîne passant une et une seule fois par chacun des sommets de  $G$ . Un graphe ne possédant que des chaînes hamiltoniennes est *semi-hamiltonien*.

### 1.2.6 Graphes triangulés

Un graphe est *triangulé* si tous ses cycles de plus de 3 sommets contiennent au moins une corde (arête reliant deux sommets non adjacents d'un cycle).

Un *séparateur* est un sous-ensemble  $W$  de sommets dans un graphe connexe  $G = (V, E)$  tel que le graphe  $G[V - W]$  est non connexe. Dans le graphe de la Figure 1.9,  $W = \{v_1, v_4\}$  est un séparateur,  $W = \{v_3\}$  est un séparateur minimal.

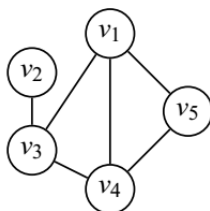


FIGURE 1.9 –  $W = \{v_3\}$  est un séparateur minimal dans  $G$ .

### 1.2.7 Arbres

On appelle *arbre* tout graphe connexe sans cycle. Un graphe sans cycle mais non connexe est appelé une *forêt*.

Une feuille ou sommet pendant est un sommet de degré 1.

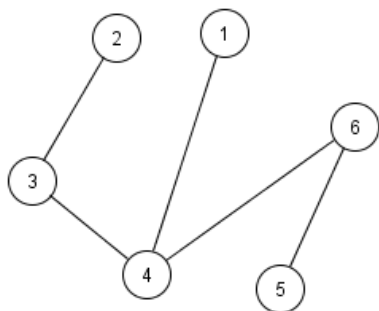


FIGURE 1.10 – Arbre.

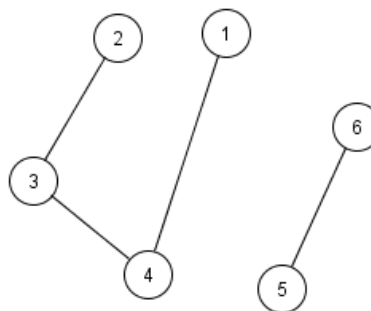


FIGURE 1.11 – Forêt.

#### Théorème 1.2.2 .

Les assertions suivantes sont équivalentes pour tout graphe  $G$  à  $n$  sommets.

1.  $G$  est un arbre
2.  $G$  est sans cycle et connexe.
3.  $G$  est sans cycle et comporte  $n - 1$  arêtes.
4.  $G$  est connexe et comporte  $n - 1$  arêtes.
5. chaque paire  $u, v$  de sommets distincts est reliée par une seule chaîne simple (et le graphe est sans boucle).

**Arbres couvrants.** Un *arbre couvrant* d'un graphe  $G$  est un graphe partiel qui est aussi un arbre.

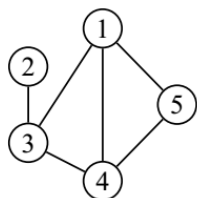


FIGURE 1.12 – Graphe  $G$ .

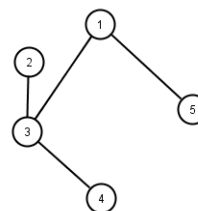


FIGURE 1.13 – Un arbre couvrant.

### 1.3 Graphe complémentaire, biparti complémentaire

Soit  $G = (V, E)$  un graphe simple. Le *graphe complémentaire* de  $G$ , est le graphe simple  $\overline{G} = (V, \overline{E})$ , avec  $\overline{E} = \{e \mid e \notin E\}$ .

Soit  $G = (V_1 \cup V_2, E)$  un graphe biparti. Le graphe *biparti complémentaire* de  $G$ , noté  $\overline{G}_B$ , est un graphe biparti tel que pour tout couple de sommets  $(v_1, v_2) \in V_1 \times V_2$ ,  $v_1$  et  $v_2$  sont non adjacents dans  $G$  si et seulement s'ils sont adjacents dans  $\overline{G}_B$ .

### 1.4 Couplage, stable et clique

Dans un graphe  $G = (V, E)$ , un *couplage* noté  $M$ , est un sous ensemble d'arêtes qui sont deux à deux non adjacentes. Chaque sommet  $u$  incident à une arête de  $M$  est dit saturé par  $M$ . Un couplage qui sature tous les sommets de  $G$  est appelé *couplage parfait*.

Un *stable* (resp. clique) de  $G$  est un sous ensemble de sommets qui sont deux à deux non adjacents (resp. adjacents).

La *stabilité* de  $G$ , notée  $\alpha(G)$ , est alors le cardinal d'un stable maximum de  $G$ .

### 1.5 Nombre chromatique

On appelle *k-coloration* d'un graphe  $G = (V, E)$ , une application  $\xi : V \rightarrow \{1, 2, \dots, k\}$ , telle que pour toute paire de sommets adjacents  $u$  et  $v$ , on a  $\xi(u) \neq \xi(v)$ .

On dit que  $G$  est *k-colorable* s'il admet une *k-coloration*.

Le *nombre chromatique* du graphe  $G$ , noté  $\chi(G)$ , est alors le plus petit entier  $k$  tel que  $G$  est *k-colorable*.

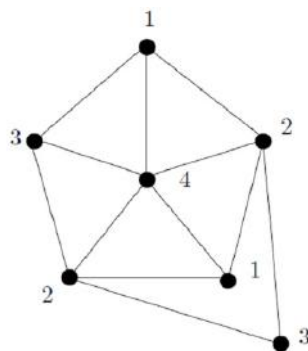


FIGURE 1.14 – Graphe admettant une 4-coloration avec  $\chi(G) = 4$ .

On peut également voir le nombre chromatique comme le plus petit nombre de stables

nécessaires pour couvrir (ou partitionner) les sommets de  $G$ . On a donc

$$\chi(G) = \frac{|V|}{\alpha(G)}.$$



## Chapitre 2

---

### Méthode Probabiliste

---

La méthode probabiliste, popularisée par Paul Erdős, désigne l'utilisation de techniques probabilistes pour traiter des problèmes "déterministes" en mathématiques discrètes. Souvent, il s'agit, quand on veut montrer l'existence d'un objet satisfaisant certaines propriétés, de choisir cet objet de manière aléatoire et de montrer qu'il vérifie les propriétés voulues avec probabilité strictement positive.

Dans ce chapitre, on introduira des outils probabilistes et leurs propriétés puis les applications de la méthode probabiliste. Nous nous contenterons ici d'en donner quelques illustrations.

### 2.1 Préliminaires mathématiques

#### 2.1.1 Espace probabilisé

**Définition 2.1.1 (Tribu).** Soit  $\Omega$  un ensemble quelconque, Une Tribu (notée  $\mathcal{F}$ ) ou  $\sigma$ -algèbre est un sous-ensemble de  $\mathbb{P}(\Omega)$  vérifiant les propriétés suivantes :

1.  $\emptyset \in \mathcal{F}$
2.  $\forall A \in \mathcal{F}, \bar{A} \in \mathcal{F}$
3. Pour toute suite  $(A_n)$  d'éléments de  $\mathcal{F}$ ,  $\bigcup_{n=0}^{+\infty} A_n \in \mathcal{F}$ . (stabilité par union dénombrable).

**Définition 2.1.2 (Probabilité).** Soient  $\Omega$  un ensemble fini non vide,  $\mathcal{F}$  une famille d'événements observables sur  $\Omega$ . On appelle probabilité sur  $(\Omega, \mathcal{F})$  toute fonction  $\mathbb{P} : \mathcal{F} \rightarrow [0, 1]$  vérifiant :

1.  $\mathbb{P}(\Omega) = 1$ .
2. Pour toute suite  $(A_i)_{i \in \mathbb{N}}$  d'événements de  $\mathcal{F}$  deux à deux incompatibles (ou disjoints) :

$$\mathbb{P}\left(\bigcup_{j \in \mathbb{N}} A_j\right) = \sum_{j=1}^{+\infty} \mathbb{P}(A_j) \text{ (Axiome de } \sigma\text{-additivité).}$$

Le triplet  $(\Omega, \mathcal{F}, \mathbb{P})$  s'appelle **espace probabilisé**.

**Proposition 2.1.1.** 1.  $\mathbb{P}(\emptyset) = 0$ .

2. Pour tout  $A \in \mathcal{F}$ ,  $\mathbb{P}(A^c) = 1 - \mathbb{P}(A)$ .

3. (Additivité) Pour tout  $A, B \in \mathcal{F}$  tels que  $A \cap B = \emptyset$ ,

$$\mathbb{P}(A \cup B) = \mathbb{P}(A) + \mathbb{P}(B).$$

4. Pour tout  $A \subseteq B \in \mathcal{F}$ ,

$$\mathbb{P}(B) = \mathbb{P}(A) + \mathbb{P}(B \setminus A) \geq \mathbb{P}(A).$$

5. Pour tout  $A, B \in \mathcal{F}$ ,

$$\mathbb{P}(A \cup B) = \mathbb{P}(A) + \mathbb{P}(B) - \mathbb{P}(A \cap B).$$

6. Plus généralement,  $\forall A_1, A_2, \dots, A_n \in \mathcal{F}$ ,

$$\begin{aligned} \mathbb{P}\left(\bigcup_{i=1}^n A_i\right) &= \sum_{i=1}^n \mathbb{P}(A_i) - \sum_{1 \leq j < i \leq n} \mathbb{P}(A_i \cap A_j) \\ &\quad + \sum_{1 \leq j < i < k \leq n} \mathbb{P}(A_i \cap A_j \cap A_k) - \dots + (-1)^{n+1} \mathbb{P}(A_1 \cap A_2 \cap \dots \cap A_n). \end{aligned}$$

7. (Sous- $\sigma$ -additivité) Pour toute collection  $\forall A_1, A_2, \dots, A_n \in \mathcal{F}$ ,

$$\mathbb{P}\left(\bigcup_{i=1}^n A_i\right) \leq \sum_{i=1}^{+\infty} \mathbb{P}(A_i).$$

### 2.1.2 Equiprobabilité

**Définition 2.1.3.** Dans le cas d'un espace probabilisable fini, on dit qu'il y a *équiprobabilité* lorsque les probabilités de tous les événements élémentaires sont égales. On parle aussi de probabilité uniforme.

#### **Théorème 2.1.1 .**

*S'il y a équiprobabilité, pour tout événement  $A$ , on a :*

$$\mathbb{P}(A) = \frac{\text{card}(A)}{\text{card}(\Omega)}$$

On parle aussi du quotient :  $\frac{\text{nombre de cas favorables}}{\text{nombre de cas possibles}}$ .

**Définition 2.1.4 (Probabilités Conditionnelles).** Soit  $(\Omega, \mathcal{F}, \mathbb{P})$  un espace probabilisé,  $B$  un événement de probabilité non nulle. On appelle *probabilité conditionnée* par  $B$ , l'application notée  $\mathbb{P}_B$  de  $(\Omega, \mathcal{F})$  dans  $R_+$  définie par :

$$\mathbb{P}(A/B) = \frac{\mathbb{P}(A \cap B)}{\mathbb{P}(B)}.$$

### 2.1.3 Événements indépendants

Soit  $(\Omega, \mathcal{F}, \mathbb{P})$  un espace probabilisé. Nous dirons que deux événements  $A$  et  $B$  sont indépendants si, et seulement si, on peut écrire :

$$\mathbb{P}(A \cap B) = \mathbb{P}(A)\mathbb{P}(B).$$

Plus généralement, une famille d'événements  $(A_i)_{i \in I}$  est indépendante sous  $\mathbb{P}$  si :

$$\mathbb{P}\left(\bigcap_{i \in J} A_i\right) = \prod_{i \in J} \mathbb{P}(A_i),$$

pour tous les sous-ensembles finis  $J$  de  $I$ .

#### 2.1.3.1 Tribus indépendantes

On dit que les tribus  $\tau_1, \dots, \tau_n$  sont indépendantes si :

$$\forall (A_1, \dots, A_n) \in \tau_1 \times \dots \times \tau_n, \mathbb{P}\left(A_1 \cap \dots \cap A_n\right) = \prod_{i=1}^n \mathbb{P}(A_i).$$

Une suite de tribus est dite indépendante si toute sous-famille finie l'est.

### 2.1.4 Variable aléatoire

**Définition 2.1.5.** Une variable aléatoire est une application  $X$  de  $\Omega$  dans  $\mathbb{R}^+$ . On notera  $\mathbb{P}(X = a)$  pour  $\mathbb{P}\{\omega \in \Omega \mid X(\omega) = a\}$ , et de même pour  $X \leq a$ .

### 2.1.4.1 Loi d'une variable aléatoire

Soit  $X$  une variable aléatoire sur un espace  $\Omega$ .

**Définition 2.1.6.** La loi de  $X$  est définie par :

$$p(x) = \mathbb{P}[X = x] = \sum_{X(\omega)=x} \mathbb{P}[\omega], \quad \text{pour } x \in \mathbb{R}.$$

On dit aussi que  $p$  est la loi image de  $\mathbb{P}$  par  $X$ .

### 2.1.4.2 Variables aléatoires indépendantes

**Définition 2.1.7.**  $X$  et  $Y$  sont des variables indépendantes si :

$$\forall x, y \in \Omega, \quad \mathbb{P}[X = x \text{ et } Y = y] = \mathbb{P}[X = x] \mathbb{P}[Y = y].$$

### 2.1.5 Espérance

**Définition 2.1.8.** L'espérance d'une variable aléatoire  $X$  est notée  $\mathbb{E}[X]$  sur un espace  $\Omega$  et à valeurs réelles est sa valeur moyenne :

$$\mathbb{E}[X] = \sum_{\omega \in \Omega} \mathbb{P}[\omega] X(\omega).$$

Si  $X$  est une variable discrète à valeurs dans l'ensemble infini  $D = \{x_i : i \geq 1\}$ , lorsque la somme est bien définie, son espérance est :

$$\mathbb{E}[X] = \sum_{i=1}^{\infty} x_i \mathbb{P}(X = x_i).$$

Lorsqu'une variable  $X$  vérifie  $\mathbb{E}[X] = 0$ , on dit que la variable est **centrée**.

#### Propriétés

1. Si  $X \geq 0$ , alors  $\mathbb{E}[X] \geq 0$ .
2. Si  $X \leq Y$ , alors  $\mathbb{E}[X] \leq \mathbb{E}[Y]$ .

#### 2.1.5.1 Linéarité de l'espérance

Soient  $X_1, \dots, X_n$  des variables aléatoires,  $X = c_1 X_1 + \dots + c_n X_n$ . La linéarité des espérance indique que :

$$\mathbb{E}[X] = c_1 \mathbb{E}[X_1] + \dots + c_n \mathbb{E}[X_n].$$

Le pouvoir de ce principe vient du fait qu'il n'y a aucune restriction sur la dépendance ou l'indépendance du  $X_i$ . Dans de nombreux cas,  $\mathbb{E}[X]$  peut facilement être calculé par une décomposition judicieuse en variables aléatoires simples (souvent indicateurs)  $X_i$ .

Soit  $\sigma$  une permutation aléatoire sur  $\{1, \dots, n\}$ , uniformément choisie. Soit  $X(\sigma)$  le nombre de points fixes de  $\sigma$ . Pour trouver  $\mathbb{E}[X]$  nous décomposons  $X = X_i + \dots + X_n$  où  $X_i$  est la variable aléatoire indicatrice de l'événement  $\sigma(i) = i$ , alors :

$$\mathbb{E}[X_i] = \mathbb{P}[\sigma(i) = i] = \frac{1}{n},$$

pour que

$$\mathbb{E}[X] = \frac{1}{n} + \dots + \frac{1}{n} = 1.$$

### 2.1.6 Variance

**Définition 2.1.9.** La variance et l'écart-type d'une variable aléatoire sont définis respectivement par :

$$\text{Var}(X) = \mathbb{E}[(X - \mathbb{E}(X))^2] = \mathbb{E}[X^2] - \mathbb{E}[X]^2.$$

**Définition 2.1.10.** La covariance de deux variables aléatoires est :

$$\text{cov}[X, Y] = \mathbb{E}[(X - \mathbb{E}[X])(Y - \mathbb{E}[Y])] = \mathbb{E}[XY] - \mathbb{E}[X]\mathbb{E}[Y].$$

**Lemme 2.1.1 .**

la variance du somme des variables aléatoires est :

$$\text{var}\left[\sum_{i=1}^n X_i\right] = \sum_{i=1}^n \text{var}[X_i] + \sum_{i \neq j} \text{cov}[X_i, X_j].$$

*Preuve.*

$$\begin{aligned} \text{var}\left[\sum_{i=1}^n X_i\right] &= \mathbb{E}\left[\sum_{i=1}^n X_i \sum_{j=1}^n X_j\right] - \mathbb{E}\left[\sum_{i=1}^n X_i\right]\mathbb{E}\left[\sum_{j=1}^n X_j\right] \\ &= \sum_{i=1}^n \mathbb{E}[X_i^2] + \sum_{i \neq j} \mathbb{E}[X_i X_j] - \sum_{i=1}^n (\mathbb{E}[X_i])^2 - \sum_{i \neq j} \mathbb{E}[X_i]\mathbb{E}[X_j] \\ &= \sum_{i=1}^n \text{var}[X_i] + \sum_{i \neq j} \text{cov}[X_i, X_j]. \end{aligned}$$

□

### 2.1.7 Moments d'ordre k

Si  $X$  est une variable aléatoire, on appelle moment d'ordre  $k$ , s'il existe, le nombre  $\mathbb{E}(X^k)$ . L'espérance d'une variable aléatoire est son moment d'ordre 1. Si  $X$  est une variable aléatoire discrète, son moment d'ordre  $k$  se calcule par la formule :

$$m_k = \sum_{i=1}^q x_i^k P(X = x_i^k).$$

Le **moment centré** d'ordre  $k$  :

$$\mu_k = \mathbb{E} \left[ (X - \mathbb{E}(X))^k \right].$$

La variance d'une variable aléatoire est donc son moment centré d'ordre 2.

#### **Théorème 2.1.2 (Inégalité de Markov).**

Soit  $X$  une variable aléatoire positive et  $t$  un réel strictement positif. Alors :

$$\mathbb{P}(X \geq a) \leq \frac{\mathbb{E}(X)}{a}.$$

*Preuve.*

$$\begin{aligned} \mathbb{E}(X) &= \sum \{X(\omega)\mathbb{P}(\omega) : \omega \in \Omega\} \geq \sum \{X(\omega)\mathbb{P}(\omega) : \omega \in \Omega, X(\omega) \geq a\} \\ &\geq \sum \{a\mathbb{P}(\omega) : \omega \in \Omega, X(\omega) \geq a\} = a \sum \{\mathbb{P}(\omega) : \omega \in \Omega, X(\omega) \geq a\} = a\mathbb{P}(X \geq a). \end{aligned}$$

En divisant les deux membres par  $a$ , nous obtenons l'inégalité souhaitée. □

#### **Théorème 2.1.3 (Inégalité de Chebyshev).**

Soit  $X$  une variable aléatoire à valeurs réelles,  $a > 0$ , on a :

$$\mathbb{P} [|X - \mathbb{E}[X]| \geq a] \leq \frac{Var[X]}{a^2}.$$

*Preuve.*

Posons  $Y = X - m$ . on a  $\mathbb{E}[Y] = 0$  et  $Var[X] = Var[Y]$ . En posant  $A = \{|Y| \geq a\}$ , on a :

$$Var[Y] = \sum_{\omega \in A} |Y(\omega)|^2 \times \mathbb{P}[\omega] \geq \sum_{\omega \in A} a^2 \mathbb{P}[\omega] = a^2 \mathbb{P}[A].$$

□

### 2.1.8 Quelques lois de probabilité discrètes

- **Loi de Bernoulli**

La loi d'une variable aléatoire  $X : \Omega \rightarrow \{0, 1\}$ , avec  $P(X = 1) = p$ ,  $P(X = 0) = 1 - p$ ,  $p \in [0, 1]$ , est appelée loi de Bernoulli de paramètre  $p$ . On écrit  $X \sim \text{bernoulli}(p)$ . On parle souvent d'épreuve de Bernoulli, et les événements  $\{X = 1\}$  et  $\{X = 0\}$  sont respectivement appelés **succès** et **échec**.

**Moments :**

$$\mathbb{E}(X) = p$$

$$\text{Var}(X) = p(1 - p).$$

- **Loi binomiale**

Répétons  $n$  fois de manière indépendante une épreuve de Bernoulli de paramètre  $p$ , et notons  $X$  la variable aléatoire représentant le nombre de succès obtenus à l'issue des  $n$  épreuves. La loi de  $X$  est appelée loi binomiale de paramètres  $n$  et  $p$ ;  $X \sim \text{binom}(n; p)$ . Puisqu'il y a  $\binom{n}{k}$  façons d'obtenir  $k$  succès sur  $n$  épreuves, on voit que la fonction de masse associée à cette loi est donnée par

$$\mathbb{P}(X = k) = \binom{n}{k} p^k (1 - p)^{n-k}, \quad k \in \{0, \dots, n\}.$$

**Moments :**

$$\mathbb{E}(X) = np$$

$$\text{Var}(X) = np(1 - p).$$

- **Loi uniforme**

La loi uniforme sur  $[1, n]$  est la loi de probabilité d'une variable aléatoire  $X$  prenant chaque valeur de l'ensemble  $\{1, 2, \dots, n\}$  avec la même probabilité :

$$p(X = k) = \frac{1}{n} \text{ pour tout entier } k \text{ compris entre } 1 \text{ et } n.$$

Plus généralement, soit  $\Omega$  un ensemble fini de cardinal  $n$ . La loi de probabilité équidistribuée ou uniforme sur  $\Omega$  est la loi définie sur  $\Omega$  par la probabilité :

$$p(\omega) = 1/n \text{ pour tout élément } \omega \text{ de } \Omega.$$

**Moments**

$$\mathbb{E}(X) = \frac{1}{n} \sum_{k=1}^n k = \frac{n+1}{2}.$$

$$\text{Var}(X) = \frac{1}{n} \sum_{k=1}^n k^2 - \frac{n+1}{2} = \frac{n^2-1}{12}.$$

- **Loi géométrique**

La loi géométrique est la loi de la variable  $Y$  " loi du nombre d'essais nécessaires " pour qu'un événement de probabilité  $p$  apparaisse pour la première fois, les hypothèses étant les mêmes que pour la loi binomiale, en particulier, la probabilité  $p$  est constante au cours des essais :

$$\mathbb{P}(Y = k) = p(1-p)^{k-1} \quad k \in \mathbb{N}^*.$$

**Moments**

$$\mathbb{E}(Y) = \frac{1}{p}$$

$$\text{Var}(Y) = \frac{1-p}{p^2}.$$

- **Loi de Pascal**

La loi de Pascal est la loi de la variable  $Z$  " loi du nombre d'essais nécessaires " pour obtenir exactement  $k$  fois un événement de probabilité  $p$ , les hypothèses étant les mêmes que pour la loi binomiale (la probabilité  $p$  est constante au cours des essais).

$$p(Z = z) = p \binom{k-1}{z-1} p^{z-1} (1-p)^{k-z}.$$

**Moments**

$$\mathbb{E}(Z) = \frac{k}{p}$$

$$\text{Var}(Z) = \frac{k(1-p)}{p^2}.$$

- **Loi de Poisson**

La loi de Poisson de paramètre  $\lambda$  est la loi d'une variable aléatoire discrète réelle  $X$ , prenant toutes les valeurs entières non négatives, avec les probabilités :

$$p(X = k) = \frac{e^{-\lambda} \lambda^k}{k!} \quad k \in [0, +\infty[.$$

**Moments**

$$\mathbb{E}(X) = \text{Var}(X) = \lambda.$$



## 2.2 Méthode probabiliste

Nous voudrions prouver l'existence d'un objet combinatoire avec des propriétés spécifiques. Malheureusement, la construction explicite d'un tel bon objet ne semble pas possible, et peut-être nous n'avons même pas besoin d'un exemple spécifique, Nous voulons juste prouver que quelque chose de bon existe. On peut alors considérer un objet aléatoire à partir d'un espace de probabilité approprié et calculer la probabilité qu'il satisfasse à nos conditions. Si nous prouvons que cette probabilité est strictement positive, nous concluons qu'il doit exister un "bon" objet, si tous les objets étaient mauvais, la probabilité serait nulle. Cette technique a été utilisée pour montrer bon nombre de propriétés combinatoires dans plusieurs domaines. Nous donnons dans ce qui suit, quelques unes de ces applications.

### 2.2.1 Nombre de Ramsey

**Définition 2.2.1.** Le nombre de Ramsey  $R(k, l)$  est le plus petit entier  $n$  tel que pour toute coloration des arêtes de  $K_n$  en deux couleurs, disons bleu et rouge, il y ait soit un  $K_k$  rouge (c'est-à-dire un sous-graphe complet à  $k$  sommets dont toutes les arêtes sont coloriées en rouge), soit un  $K_l$  bleu.

Il peut également être défini comme le plus petit entier  $n$ , pour lequel un graphe  $G$  à  $n$  sommets, contienne une clique de cardinalité  $k$  et un stable de cardinalité  $l$

Ramsey a montré en 1930 que ces nombres sont tous finis. En 1947, Erdős utilise les probabilités pour obtenir des bornes inférieures pour les nombres de Ramsey diagonaux  $R(k, k)$ . C'est le premier problème qu'il traite avec la méthode probabiliste.

**Théorème 2.2.1** [4].

(Erdős, 1947) Si les entiers positifs  $n$  et  $k$  satisfont à l'inégalité  $\binom{n}{k} 2^{1-\binom{k}{2}} < 1$ , alors  $R(k, k) \geq n$ .

*Preuve.*

Tout d'abord, nous remarquons que prouver ce théorème revient à montrer l'existence d'une coloration des arêtes de  $K_n$  sans  $K_k$  rouge ni  $K_k$  bleu.

Supposons que chaque arête est coloriée soit en bleu, soit en rouge avec une probabilité  $\frac{1}{2}$  et ceci indépendamment des autres arêtes. Formellement, nous venons de créer un espace de probabilités  $(\Omega, \mathcal{F}, \mathbb{P})$  tel que les éléments de  $\Omega$  sont les  $2^{\binom{n}{2}}$  colorations possibles des arêtes de  $K_n$  en deux couleurs. Ces éléments sont tous équiprobables.

Pour tout ensemble  $S$  de  $k$  sommets, soit  $A_S$  l'évènement que le sous-graphe de  $k_n$  induit par  $S$  est unicolore. Clairement pour chaque  $S$ ,

$$\mathbb{P}(A_S) = 2\left(\frac{1}{2}\right)^{\binom{k}{2}} = 2^{1-\binom{k}{2}}.$$

Soit  $B$  l'événement qu'il y ait au moins un  $K_k$  unicolore dans une coloration de  $K_n$ . Alors :

$$\mathbb{P}(B) = P\left(\bigcup_{\substack{S \subseteq V \\ |S|=k}} A_S\right) \leq \sum_{\substack{S \subseteq V \\ |S|=k}} P(A_S) = \binom{n}{k} 2^{1-\binom{k}{2}},$$

et le dernier terme est strictement inférieur à 1 par hypothèse. Par conséquent,

$$\mathbb{P}(B^c) = 1 - \mathbb{P}(B) > 0.$$

Cela assure l'existence d'un point de notre espace de probabilité pour lequel l'événement  $B^c$  se réalise. Autrement dit, il existe une coloration de  $K_n$  pour laquelle il n'y a aucun  $k_k$  unicolore. Ceci termine la preuve.  $\square$

**Corollaire 2.2.1** [26].

$R(k, k) > 2^{k/2}$  pour tout  $k \geq 3$ .

*Preuve.*

Le résultat découle directement du Théorème 2.2.1 en choisissant  $n = \lceil 2^{k/2} \rceil$ . En effet, dans ce cas, on obtient :

$$\binom{n}{k} 2^{1-\binom{k}{2}} < \frac{n^k}{k!} 2^{1-\frac{k(k-1)}{2}} = \frac{2^{1+k/2}}{k!} \frac{n^k}{2^{k^2/2}} < 1.$$

$\square$

**Théorème 2.2.2 .**

Quels que soient deux entiers  $k \geq 2$  et  $l \geq 2$ ,

$$r(k, l) \leq r(k, l-1) + r(k-1, l)$$

De plus, si  $r(k, l-1)$  et  $r(k-1, l)$  sont tous les deux pairs, l'inégalité précédent est stricte.

*Preuve.*

Soit  $G$  un graphe à  $r(k, l-1) + r(k-1, l)$  sommets, et soit  $v \in V$ . Nous distinguons deux cas :

1. le sommet  $v$  est non-adjacent à un ensemble  $S$  d'au moins  $r(k, l-1)$  sommets.
2. le sommet  $v$  est adjacent à un ensemble  $T$  d'au moins  $r(k-1, l)$  sommets.

Notons que l'un des deux cas doit se produire, car le nombre de sommets auxquels  $v$  n'est pas adjacent plus le nombre de sommets auxquels  $v$  est adjacent est égal à  $r(k, l-1) + r(k-1, \ell) - 1$ .

Dans le cas 1,  $G[S]$  contient soit une clique de  $k$  sommets, soit un stable de  $l-1$  sommets, et par conséquent  $G[S \cup \{v\}]$  contient soit une clique de  $k$  sommets, soit un stable de  $l$  sommets.

De manière similaire, dans le cas 2,  $G[T \cup \{v\}]$  contient soit une clique de  $k$  sommets, soit un stable de  $l$  sommets. Comme un des deux cas se produit, il s'ensuit que  $G$  contient soit une clique de  $k$  sommets, soit un stable de  $l$  sommets. Ceci achève la preuve.  $\square$

On va maintenant montrer que le résultat du Théorème 2.2.1 peut être amélioré.

**Théorème 2.2.3** [26].

$R(k, k) > n - \binom{n}{k} 2^{1-\binom{k}{2}}$  quels que soient les entiers  $n$  et  $k$ .

*Preuve.*

nous considérons une coloration aléatoire des arêtes de  $K_n$  obtenue en coloriant indépendamment chaque arête soit en bleu, soit en rouge. Chaque couleur a une probabilité  $\frac{1}{2}$  d'être choisie. Pour tout ensemble  $S$  de  $k$  sommets, soit  $X_S$  la variable aléatoire indicatrice définie par :

$$X_S = \begin{cases} 1 & \text{si le sous-graphe de } K_n \text{ induit par } S \text{ est unicolore} \\ 0 & \text{sinon} \end{cases}$$

Soit  $X$  la variable aléatoire qui compte le nombre de  $K_k$  unicolores dans une coloration de  $K_n$ .

$$X = \sum_{\substack{S \subseteq V \\ |S|=k}} X_S.$$

En calculant l'espérance de  $X$ , nous obtenons

$$\begin{aligned} \mathbb{E}[X] &= \sum_{\substack{S \subseteq V \\ |S|=k}} \mathbb{E}[X_S] \\ &= \sum_{\substack{S \subseteq V \\ |S|=k}} \mathbb{P}(\text{le sous-graphe de } K_n \text{ induit par } S \text{ est unicolore}) \\ &= \mu, \end{aligned}$$

où  $\mu = \binom{n}{k} 2^{1-\binom{k}{2}}$ . Ceci certifie l'existence d'une coloration des arêtes de  $K_n$  pour laquelle  $X \leq \mu$ . Fixons une telle coloration. Ensuite, supprimons de  $K_n$  un sommet dans chaque  $K_k$

unicolore. Au plus  $\mu$  sommets sont supprimés. Après leur suppression, il nous reste un graphe complet d'au moins  $n - \mu$  sommets. De plus aucun  $K_k$  unicolore n'apparaît dans sa coloration. Nous pouvons alors en conclure que  $R(k, k) > n - \mu$ .  $\square$

### 2.2.2 Hamiltonicité

**Définition 2.2.2.** Un *tournoi*, noté  $T$  est défini comme une orientation d'un graphe complet. Un *chemin hamiltonien* est alors une orientation d'une chaîne hamiltonienne où tous les arcs ont le même sens.

Alors que l'existence d'un chemin hamiltonien dans un tournoi a été prouvée par Rédei en 1934, Le théorème suivant utilise la propriété de la linéarité de l'espérance pour estimer le nombre de chemins Hamiltoniens dans un tournoi :

**Théorème 2.2.4** [29].

Il y a un tournoi  $T$  avec  $n$  joueurs et au moins  $\frac{n!}{2^{(n-1)}}$  chemins Hamiltoniens.

*Preuve.*

Dans le jeu aléatoire, soit  $X$  le nombre de chemins hamiltoniens. Pour chaque permutation  $\sigma$ , soit  $X_\sigma$  la variable aléatoire indicative pour  $\sigma$  donnant un chemin hamiltonien ; C'est-à-dire satisfaisant  $(\sigma(i), < \sigma(i + 1)) \in T$  pour  $1 \leq i < n$ . Alors  $X = \sum X_\sigma$  et

$$\mathbb{E}[X] = \sum_{\sigma} \mathbb{E}[X_\sigma] = \frac{n!}{2^{(n-1)}}.$$

Ainsi, certains tournois ont au moins  $\mathbb{E}[X]$  chemins Hamiltoniens.  $\square$

### 2.2.3 Graphes de séparation

Un autre exemple illustrant l'usage de la méthode probabiliste en théorie des graphes :

**Théorème 2.2.5** [4].

Soit  $G(V, E)$  un graphe avec des  $n$  sommets et des  $e$  arêtes. Alors  $G$  contient un sous-graphe biparti avec au moins  $e/2$  arêtes.

*Preuve.*

Soit  $T \subseteq V$  un sous-ensemble aléatoire donné par  $\mathbb{P}[x \in T] = 1/2$ . Ces choix sont mutuellement indépendants. Ensemble  $B = V - T$ . Appelez un arête  $\{x, y\}$  traversant si exactement un de  $x, y$

est dans  $T$ . Soit  $X$  le nombre d'arêtes croisés,

$$X = \sum_{\{x,y\} \in E} X_{xy},$$

où  $X_{xy}$  est la variable aléatoire de l'indicateur pour  $\{x, y\}$  étant le croisement. alors

$$\mathbb{E}[X_{xy}] = 1/2,$$

Comme deux jetons de monnaie juste ont probabilité  $1/2$  d'être différent. alors

$$\mathbb{E}[X] = \sum_{\{x,y\} \in E} \mathbb{E}[X_{xy}] = \frac{e}{2}.$$

Ainsi  $X \geq e/2$  pour un certain choix de  $T$  et l'ensemble de ces arêtes croisés forment un graphe biparti.  $\square$

**Théorème 2.2.6** [4].

Si  $G = (V, E)$  a  $2n$  sommets et  $e$  arêtes alors il contient un sous-graphe biparti avec au moins  $\frac{en}{2n-1}$  arêtes. Si  $G$  a  $2n + 1$  sommets et  $e$  arêtes alors il contient un sous-graphe biparti avec au moins  $\frac{e(n+1)}{2n-1}$  arêtes.

*Preuve.*

Lorsque  $G$  a  $2n$  sommets, on choisit uniformément  $T$  parmi tous les sous-ensembles de  $n$  éléments de  $V$ . Toute arête  $\{x, y\}$  a maintenant une probabilité  $n/(2n - 1)$  de croisement. Lorsque  $G$  a  $2n + 1$ , les sommets choisissent  $T$  uniformément parmi tous les sous-ensembles de  $n$ -éléments de  $V$  et la preuve est similaire.  $\square$

### 2.2.4 Théorème d'Erdős-Ko-Rado

Une famille d'ensembles  $\mathcal{F}$  est dite intersectante si  $A \cap B \neq \emptyset$ .

**Théorème 2.2.7** [37].

Si  $|X| = n, n \geq 2k$  et  $\mathcal{F}$  une famille intersectante de sous-ensembles de  $k$  éléments de  $X$ , alors

$$|\mathcal{F}| \leq \binom{n-1}{k-1}.$$

**Lemme 2.2.1** .

Pour  $0 \leq s \leq n - 1$ , on pose  $A_s = \{s, s + 1, \dots, s + k - 1\}$ , l'addition étant modulo  $n$ . Alors  $\mathcal{F}$  ne peut contenir plus de  $k$  ensembles  $A_s$ .

*Preuve.*

Supposons que  $A_l \in \mathcal{F}$ . A part  $A_l$  lui-même, exactement  $2k-2$  des ensembles  $A_s$  intersectent  $A_l$ . Ceux-ci peuvent être répartis en  $k-1$  paires d'ensembles disjoints. Puisque  $\mathcal{F}$  ne peut contenir qu'au plus un membre de chacune de ces paires, le lemme est démontré.  $\square$

*preuve du Théorème 2.2.7.* On tire au hasard une permutation  $\sigma$  de  $\{0, \dots, n-1\}$  et un élément  $i \in \{0, \dots, n-1\}$ , tous deux de façon uniforme, et indépendamment l'un de l'autre. Soit  $A = \{\sigma(i), \sigma(i+1), \dots, \sigma(i+k-1)\}$  (la somme étant toujours prise modulo  $n$ ). Il est clair (et facilement démontré) que la loi de  $A$  est uniforme sur l'ensemble des sous-ensembles à  $k$  éléments de  $\{0, \dots, n-1\}$ . En particulier,  $\mathbb{P}(A \in \mathcal{F}) = |\mathcal{F}| / \binom{n}{k} = nk$ . D'un autre côté,

$$\mathbb{P}(A \in \mathcal{F}) = \sum_{\sigma} \mathbb{P}(A \in \mathcal{F} \mid \sigma) \mathbb{P}(\sigma).$$

Conditionnellement à  $\sigma$ ,  $A$  suit la loi uniforme sur les  $n$  sous-ensembles de  $k$  éléments consécutifs de l'ensemble ordonné  $\{\sigma(0), \sigma(1), \dots, \sigma(n-1)\}$  (consécutifs au sens du Lemme). Par conséquent, le lemme (appliqué à l'ensemble  $\{\sigma(0), \sigma(1), \dots, \sigma(n-1)\}$ ) implique que  $\mathbb{P}(A \in \mathcal{F} \mid \sigma) \leq k/n$ , pour chaque permutation  $\sigma$ , et donc  $\mathbb{P}(A \in \mathcal{F}) \leq k/n$ . On doit donc avoir

$$\frac{|\mathcal{F}|}{\binom{n}{k}} \leq \frac{k}{n}.$$

$\square$

### 2.2.5 Estimation du coefficient binomial moyen

Parmi les coefficients binomiaux  $\binom{2m}{k}$ ,  $k = 0, 1, \dots, 2m$ ,  $\binom{2m}{m}$  est le plus grand et il apparaît souvent dans diverses formules (par exemple, dans les numéros catalans, qui contiennent des arbres binaires et bien d'autres choses). La méthode du deuxième moment fournit un moyen simple de délimiter  $\binom{2m}{m}$  par le bas. Il existe plusieurs autres approches, dont certaines donnent des estimations beaucoup plus précises, mais l'astuce simple avec l'inégalité de Chebyshev donne l'ordre de grandeur correct.

**Proposition 2.2.1** [29].

Pour tout  $m \geq 1$ , nous avons  $\binom{2m}{m} \geq 2^{2m} / (4\sqrt{m} + 2)$ .

*Preuve.*

Considérer la variable aléatoire  $X = X_1 + X_2 + \dots + X_{2m}$ , où  $X_i$  sont indépendants et Chacun d'entre eux atteint les valeurs 0 et 1 avec la probabilité 1/2. nous avons  $\mathbb{E}[X] = m$  et  $\text{Var}[X] =$

$\frac{m}{2}$ . L'inégalité de Chebyshev à  $a = \sqrt{m}$  donne :

$$\mathbb{P}[|X - m| < \sqrt{m}] \geq \frac{1}{2}.$$

La probabilité que  $x$  atteigne une valeur spécifique  $m + k$ , où  $|k| < \sqrt{m}$ , est  $\binom{2m}{m+k} 2^{-2m} \leq \binom{2m}{m} 2^{-2m}$  (Parce que  $\binom{2m}{m}$  est le coefficient binomial le plus grand). Donc nous avons

$$\frac{1}{2} \leq \sum_{|k| < \sqrt{m}} \mathbb{P}[X = m + k] \leq (2\sqrt{m} + 1) \binom{2m}{m} 2^{-2m}.$$

□

## Chapitre 3

---

### Graphes Aléatoires

---

#### 3.1 Introduction

La théorie des graphes aléatoires a été fondée par Erdős et Rényi en 1959. Ce modèle consistait à choisir équitablement un graphe parmi les  $2^{\binom{n}{2}}$  graphes que l'on peut former à partir de  $n$  sommets. En d'autres termes, son modèle se basait sur l'espace de probabilité  $(\Omega, \mathcal{F}, \mathbb{P})$ , où l'ensemble  $\Omega$  contient tous les graphes possédant  $n$  sommets, la tribu  $\mathcal{F}$  contient tous les sous-ensembles de  $\Omega$ , et la probabilité  $P$  est définie de manière à ce que pour tout  $\omega \in \Omega$ , on ait

$$\mathbb{P}(\omega) = 2^{-\binom{n}{2}}. \quad (1)$$

Dans le passé, deux modèles ont été effectivement associés au nom d'Erdős-Rényi, le modèle uniforme [17] avec un nombre fixe  $M \in \mathbb{N}^*$  d'arêtes, noté  $G(n, M)$ , et le modèle binomial [24], noté  $G(n, p)$  où  $p \in [0, 1]$  est la probabilité de connexion entre deux sommets quelconques, et est fixe. Aujourd'hui, le nom Erdős-Rényi est plutôt donné au modèle binomial.

#### 3.2 Modèle d'Erdős-Rényi

Le modèle d'Erdős-Rényi est parfois appelé le graphe aléatoire au sens où, dans ce modèle, la topologie du réseau est tirée uniformément au hasard. Historiquement deux modèles sont en fait associés au nom d'Erdős-Rényi, le modèle uniforme à nombre d'arêtes  $M$  fixé, noté  $G(n, M)$ , et le modèle binomial, où la probabilité  $p$  de connexion entre deux sommets quelconques est fixée, noté  $G(n, p)$ .

L'avantage du modèle de d'Erdős-Rényi est qu'il est relativement simple à comprendre et à étudier. Par exemple on peut facilement s'intéresser à la loi de probabilité des degrés des sommets.



Ces modèles ne sont pas formellement équivalents, mais leurs caractéristiques sont très similaires les unes aux autres dans un cadre asymptotique. Ils ont été étudiés sous toutes leurs formes, le second plus intensément. [7, 15, 3].

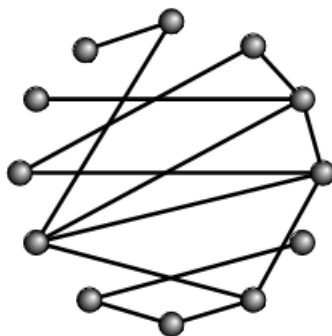


FIGURE 3.1 – Exemple d'un graphe aléatoire de type Erdős-Rényi.

### 3.2.1 Graphe aléatoire uniforme

Soit  $M$  un nombre entier compris entre 0 et  $\binom{n}{2}$ . Le graphe aléatoire uniforme, noté  $G(n, M)$ , peut alors être défini en prenant pour  $\Omega$  l'ensemble des graphes ayant  $n$  sommets et  $M$  arêtes, et en posant, pour tout  $G \in \Omega$ ,

$$\mathbb{P}(G) = \binom{\binom{n}{2}}{M}^{-1}.$$

Cette probabilité est donc identique pour tous les graphes appartenant à  $\Omega$ . On peut d'ailleurs étendre ce modèle à toute famille  $F$  de graphes ayant une même propriété. Par exemple on peut considérer l'ensemble de tous les graphes  $r$ -réguliers (i.e. les graphes dont chaque sommet est de degré  $r$ ), ou encore l'ensemble de tous les arbres à  $n$  sommets.

### 3.2.2 Graphe aléatoire binomial

On se donne un nombre réel  $p$  compris entre 0 et 1. Le *graphe aléatoire binomial*, noté  $G(n, p)$ , peut être défini en prenant pour  $\Omega$  l'ensemble de tous les graphes possédant  $n$  sommets, et en posant, pour tout  $G \in \Omega$ ,

$$\mathbb{P}(G) = p^{e_G} (1 - p)^{\binom{n}{2} - e_G}. \quad (2)$$

Il peut donc être vu comme le résultat de  $\binom{n}{2}$  expériences indépendantes, une pour chaque paire de sommets distincts, ayant chacune une probabilité de succès égale à  $p$ . On peut remarquer

qu'en prenant  $p = 1/2$ , on retrouve l'équation (1).

De manière plus générale, on peut prendre pour  $\Omega$  l'ensemble des graphes partiels d'un graphe  $F$  donné, et on aurait alors  $P(G) = p^{e_F} (1-p)^{e_F - e_G}$ . En prenant  $F = K_n$ , on retrouve l'équation (2). On peut aussi considérer le graphe aléatoire biparti  $G(m, n, p)$  en posant  $F = K_{m,n}$ .

L'avantage de ce modèle est que la présence ou non d'une arête est un événement indépendant de celle des autres arêtes. Mais l'inconvénient est que le nombre d'arêtes n'est pas fixé.

Il varie aléatoirement selon une loi binomiale dont l'espérance vaut  $\binom{n}{2}p$ .

Nous appelons un tel graphe aléatoire, un graphe aléatoire binomial et le désignons par  $G_{n,p} = ([n], E_{n,p})$ .

Puisque la distribution binomiale peut être approchée par la distribution de Poisson dans le cas  $np \rightarrow \lambda$ , on pose  $p = \frac{\lambda}{n}$  et on donne le résultat final que la distribution en degrés du graphe aléatoire d'Erdős-Rényi est Poisson avec le paramètre  $\lambda$  :

$$\mathbb{P}(X = k) = \frac{\lambda^k}{k!} e^{-\lambda}.$$

### 3.2.3 Lien entre les deux modèles

Le Lien entre ces deux modèles est notamment établi dans [36]. Ces modèles sont liés dans le cadre asymptotique particulier où le nombre d'arêtes  $M$  pour l'un, la probabilité de connexion  $p$  pour l'autre, dépendent de la taille  $n$  du graphe. En effet le nombre d'arêtes de  $G(n, p)$  suit une loi binomial qui par l'inégalité de Hoeffding se concentre très vite autour de son espérance  $Np$ . Ainsi, intuitivement si on pose  $p_n = M/N$ ,  $G(n, p_n)$  se comporte de la même façon que  $G(n, M)$  quand  $n$  tend vers l'infini. Janson et al. [36] éclairent ceci : " la plupart du temps, une propriété de graphe est asymptotiquement vérifiée dans  $G(n, p_n)$  si et seulement si elle l'est dans  $G(n, M)$ ".

**Lemme 3.2.1** [23].

*Un graphe aléatoire  $G(n, p)$ , avec  $m$  arêtes, est également susceptible d'être l'un quelconque des  $\binom{n}{M}$  graphes avec  $m$  arêtes.*

*Preuve.*

Soit  $G_0$  un graphe marqué de  $m$  arêtes. On a :

$$\{G(n, p) = G_0\} \subseteq \{|E_{n,p}| = m\},$$

nous avons

$$\begin{aligned}
 \mathbb{P}(G(n, p) = G_0 \mid |E_{n,p}| = m) &= \frac{\mathbb{P}(G(n, p) = G_0, |E_{n,p}| = m)}{\mathbb{P}(|E_{n,p}| = m)} \\
 &= \frac{\mathbb{P}(G_{n,p} = G_0)}{\mathbb{P}(|E_{n,p}| = m)} \\
 &= \frac{p^m (1-p)^{\binom{n}{2}-m}}{\binom{\binom{n}{2}}{M} p^m (1-p)^{\binom{n}{2}-m}} \\
 &= \left( \frac{\binom{n}{2}}{M} \right)^{-1}.
 \end{aligned}$$

□

### 3.2.4 Exemple d'application du modèle d'Erdős-Rényi

Un exemple d'application du modèle d'Erdős-Rényi en épidémiologie a été présenté dans [15]. Pour illustrer comment modéliser un problème avec ce modèle et quelles hypothèses pratiques de modélisation sont en jeu, les auteurs ont introduit le modèle Reed-Frost et montré l'équivalence avec Erdős-Rényi.

Le modèle Reed-Frost appartient à la classe des modèles épidémiques appelés SIR (pour Susceptible, Infecté, Récupéré). Dans ces modèles, chaque individu de la population se trouve dans l'un des états S, I ou R, selon qu'il n'a jamais été infecté par la maladie et susceptible d'être, ou est actuellement infecté, ou récupéré, c'est-à-dire immunisé. Le modèle Reed-Frost est une version ce genre de modèles. Un seul individu est infecté à l'origine de temps. La dynamique est alors suit :

1. La population est fermée : la maladie ne peut être transmise que dans les  $n$  individus.
2. La maladie peut être transmise par un seul type de contact et la probabilité  $p$  avoir ce contact dans une unité de temps est le même entre deux individus et à tout moment unité.
3. Les contacts entre individus sont indépendants.
4. Une personne infectée ne l'est que pour une unité de temps et est alors définitivement immunisée.

Barbour et Mollison [6] ont d'abord introduit la construction mathématique du Reed-Frost modèle du modèle Erdős-Rényi.  $X$  est supposé être un graphe construit sous le modèle  $G(n, p)$ .

Soit  $i \in [n]$  l'individu infecté à l'origine.

On note  $\delta_X$  la distance dans le graphe  $X$ . L'état  $e_t(j)$  de l'individu  $j \in [n]$  au temps  $t \in \mathbb{N}$  est défini par :

$$e_t(j) = \begin{cases} S & \text{si } \delta_X(i, j) > t \\ I & \text{si } \delta_X(i, j) = t \\ R & \text{si } \delta_X(i, j) < t \end{cases}$$

Ainsi, les contacts ayant donné lieu à une infection (si aucun des deux sommets n'était pas déjà retiré) sont matérialisés par les arêtes du graphe. L'indépendance des contacts correspond à l'indépendance des variables de Bernoulli dans le modèle  $G(n, p)$  et l'égalité de la probabilité d'avoir un contact infectieux entre tout les individus, à l'égalité du paramètre  $p$  de ces variables.

L'épidémie se propage du sommet  $i$  vers ses voisins directs, puis itérativement vers la couche de voisins directement à chaque nouvelle unité de temps. La composante connexe de  $i$  dans  $X$  représente alors l'ensemble des sommets qui sont retirés à la fin, c'est-à-dire ceux qui ont été un jour infectés. L'étude des composantes connexes du modèle  $G(n, p)$  et de leur taille peut donc être interprétée en termes épidémiques et répond à la question de la proportion de la population touchée par la maladie.

### 3.2.5 Distribution des degrés

On remarque immédiatement que si  $d(v)$  désigne le degré d'un sommet fixe dans  $G(n, p)$ , alors  $d(v)$  est une variable aléatoire binomialement distribuée, avec les paramètres  $n - 1$  et  $p$ , c'est-à-dire pour  $d = 0, 1, 2, \dots, n - 1$ ,

$$\mathbb{P}(d(v) = d) = \binom{n-1}{d} p^d (1-p)^{n-1-d}.$$

Tandis que dans  $G(n, M)$ , la distribution de  $d(v)$  est hypergéométrique, c'est-à-dire,

$$\mathbb{P}(d(v) = d) = \frac{\binom{n-1}{d} \binom{M-d}{n-1-d}}{\binom{M}{n-1}}.$$

## 3.3 Connexité

### 3.3.1 Disparition des sommets isolés

Soit  $X$  le nombre de sommets isolés dans  $G(n, p)$ .

$$\begin{aligned} X &= \sum_{i=1}^n \mathbb{I}_{i \text{ est isolé}} \\ \mathbb{E}(X) &= n(1-p)^{n-1} \end{aligned}$$

Lorsque  $p = o(1)$ , ce qu'on supposera dans la suite,

$$\mathbb{E}(X) = e^{\log n - np - np^2/2 + o(np^2)}.$$

Si  $np - \log n$  tend vers  $+\infty$ ,  $G(n, p)$  ne contient pas de sommet isolé, a.p.s. Si  $np - \log n$  tend vers  $-\infty$ ,  $\mathbb{E}(X)$  tend vers  $+\infty$ . Examinons le second moment :

$$\begin{aligned} \mathbb{E}(X^2) &= \mathbb{E}(X) + \sum_{i \neq j} \mathbb{P}(i, j \text{ sont isolés}), \\ &= \mathbb{E}(X) + (n)_2 (1-p)^{2n-3}, \\ &= \mathbb{E}(X) + \mathbb{E}(X)^2 (1 + o(1)). \end{aligned}$$

Donc si  $E(X)$  tend vers l'infini,  $\mathbb{E}(X^2) \sim \mathbb{E}(X)^2$  et on peut appliquer la méthode du second moment. On a montré que :

**Proposition 3.3.1** [35].

*Si  $np - \log n$  tend vers  $-\infty$ ,  $G(n, p)$  contient au moins un sommet isolé, a.p.s. Si  $np - \log n$  tend vers  $+\infty$ ,  $G(n, p)$  ne contient pas de sommet isolé, a.p.s.*

On peut imaginer que dans l'intervalle de seuil,  $X$  a un comportement poissonien, vu la faible corrélation entre les indicatrices la composant. Calculons le  $k$ -ème moment factoriel de  $X$  pour  $k$  fixé :

$$\begin{aligned} \mathbb{E}[(X)_k] &= \sum_{\substack{(i_1, \dots, i_k) \\ 2 \text{ à } 2 \text{ distincts}}} \mathbb{P}((i_1, \dots, i_k) \text{ sont isolés}) \\ &= (n)_k (1-p)^{\binom{k}{2} + k(n-k)}, \\ &\sim \mathbb{E}[X]^k. \end{aligned}$$

Donc si  $E(X)$  converge vers une constante  $\lambda$ ,  $X$  converge en loi vers la loi de Poisson de paramètre  $\lambda$ . Or, si  $np - \log n$  converge vers  $c \in \mathbb{R}$ ,  $\mathbb{E}(X)$  converge vers  $e^{-c}$ . On en déduit le résultat suivant.

**Théorème 3.3.1** [7].

*Soit  $X$  le nombre de sommets isolés dans  $G(n, p)$ . Si  $p = \frac{\log n}{n} + \frac{c}{n} + o\left(\frac{1}{n}\right)$  avec  $c \in \mathbb{R}$ ,  $X$  converge vers une loi de Poisson de paramètre  $e^{-c}$ . En particulier, la probabilité que  $G(n, p)$  possède au moins un sommet isolé converge vers  $e^{-c}$ .*

### 3.3.2 Composante géante dans le modèle $G(n, p)$

L'une des propriétés cruciales pour comprendre la structure d'un graphe d'Erdős-Rényi est la présence d'une composante dite "géante", contenant une fraction constante des sommets du graphe (cela n'a de sens qu'asymptotiquement). On connaît beaucoup de choses sur cette composante géante, et notamment le seuil critique associé à sa présence dans le graphe [4].

#### Lemme 3.3.1 .

Posons  $p(n) = \frac{c}{n}$ , où  $c > 0$ .

— Pour  $c < 1$ , avec probabilité tendant vers 1, la plus grande composante connexe de  $G(n, p(n))$  est de taille  $O(\log n)$  quand  $n \rightarrow \infty$ .

— Pour  $c > 1$ , il existe avec probabilité tendant vers 1 une unique composante dite géante, de taille  $Kn(1+o(1))$ , où  $K$  est une constante dépendant de  $c$ . Les autres composantes connexes sont de taille  $O(\log n)$ .  $K$  est en outre l'unique solution non nulle de l'équation  $x = 1 - e^{-cx}$ .

Ainsi, il y a une transition dans le comportement du graphe. Si  $p$  est assez faible, toutes les composantes sont très petites. Mais lorsqu'on dépasse le seuil  $p = \frac{1}{n}$ , il existe une unique composante géante, toutes les autres étant de nouveau très petites. Puis, à partir de  $p = \frac{\log n}{n}$ , le graphe est même connexe (avec probabilité tendant vers 1).

La plupart des propriétés des graphes du modèle d'Erdős-Rényi possèdent une transition de phase par rapport au paramètre  $p$ , que ce soit le fait d'avoir une composante géante ou d'être connexe, d'avoir un certain diamètre, de contenir au moins un motif donné, etc.

On aborde ici la question de la composante géante. On se place dans une asymptotique poissonnienne, c'est-à-dire que l'on considère la suite de modèles  $G(n, p)$ , où le degré moyen  $\lambda = np_n$  est un paramètre fixe. Le nom de cette asymptotique vient en particulier du fait que les degrés convergent alors en loi vers une loi de poisson de paramètre  $\lambda$ . on note  $\mathcal{C}_{(1)}, \mathcal{C}_{(2)} \dots$  Les composantes connexes rangées dans l'ordre décroissant de taille. Le principal résultat est la transition de phase que subit la taille de la plus grande composante  $\mathcal{C}_{(1)}$  par rapport au paramètre  $\lambda$  [15].

#### Théorème 3.3.2 [15].

Dans le régime sous-critique ( $\lambda < 1$ ). Il existe une constante  $A_1 \geq 0$  dépendant de  $\lambda$  telle que :

$$\lim_{n \rightarrow \infty} P_{G(n, p_n)} (|\mathcal{C}_{(1)}| \leq A_1 \log(n)) = 1.$$

Dans le régime sur-critique ( $\lambda > 1$ ). Soit  $p_{GW}(\lambda)$  l'unique solution de  $x = e^{-\lambda(1-x)}$  dans

$]0, 1[$ . Il existe une constante  $A_2 > 0$  dépendant de  $\lambda$  telle que pour tout  $\varepsilon > 0$  :

$$\lim_{n \rightarrow \infty} P_{G(n, p_n)} \left( \left| \frac{|\mathcal{C}_{(1)}|}{n} - (1 - p_{GW}(\lambda)) \right| \leq \varepsilon, |\mathcal{C}_{(2)}| \leq A_2 \log(n) \right) = 1.$$

Dans le régime critique ( $\lambda = 1$ ). Il existe une constante  $A_3 \geq 0$  telle que pour tout  $b > 0$  :

$$P_{G(n, p_n)} \left( |\mathcal{C}_{(1)}| \geq bn^{2/3} \right) \leq \frac{A_3}{b^2}.$$

Dans le régime sous-critique, les composantes sont asymptotiquement toutes de taille au plus logarithmique en  $n$ . Dans un régime sur-critique, elles sont aussi asymptotiquement toutes de taille au plus logarithmique, sauf la plus grande, qui englobe une proportion non négligeable de sommets : il y a donc une composante géante asymptotiquement presque sûrement. Cette proportion est  $1 - p_{GW}(\lambda)$ , où  $p_{GW}(\lambda)$  est en fait la probabilité d'extinction d'un processus de Galton-Watson dont la loi de descendance est la loi de poisson de paramètre  $\lambda$ . On pouvait noter que ce processus apparaissait déjà dans la formulation du modèle de Reed-Frost, qui construit en fait un arbre de descendance dans le graphe. Le régime critique est plus compliqué que les deux autres ; Nous ne donnons que cette borne supérieure simple qui permet d'avoir une idée de l'ordre de grandeur  $n^{2/3}$  de la plus grande composante [7, 27].

### 3.3.3 Composantes connexes

Les composantes connexes de  $G(n, p)$  ne peuvent être que de deux types : des sommets isolés et une composante "géante", de taille strictement supérieure à  $n/2$  (il ne peut contenir plus d'une composante connexe de taille strictement supérieure à  $n/2$ ). On en déduira donc que a.p.s.  $G(n, p)$  est connexe si et seulement si il ne contient pas de sommet isolé, ce qui donne, avec le théorème

**Théorème 3.3.3** [7].

Si  $p = \frac{\log n}{n} + \frac{c}{n} + o\left(\frac{1}{n}\right)$  avec  $c \in \mathbb{R}$ , la probabilité que  $G(n, p)$  soit connexe converge vers  $e^{-c}$ .

**Lemme 3.3.2** .

Si  $np \geq \alpha \log n$  avec  $\alpha > 2/3$ , alors a.p.s.  $G(n, p)$  ne possède pas de composante connexe d'ordre appartenant à  $[2, n/2]$ .

*Preuve.*

Supposons que  $p = o(1)$  et  $np \geq \alpha \log n$  avec  $\alpha > 1/2$ . Examinons séparément le cas des

composantes connexes d'ordre 2 (les arêtes isolées). La probabilité que  $G(n, p)$  contienne une composante connexe d'ordre 2 est inférieure à :

$$\binom{n}{2} p (1-p)^{n-2} = \Theta(pn^2 e^{-2pn}) = O\left(n \log n e^{-2\alpha \log n}\right) = o(1).$$

Toute composante connexe d'ordre  $r$  contient un arbre couvrant dont tous les sommets sont isolés des sommets en dehors de la composante connexe. Le nombre d'arbres étiquetés sur  $r$  sommets est égal à  $r^{r-2}$ , d'après la formule de Cayley. Donc, la probabilité que  $G(n, p)$  contienne une composante connexe d'ordre appartenant à  $[3, n/2]$  est inférieure à :

$$\sum_{r=3}^{[n/2]} \binom{n}{r} r^{r-2} p^{r-1} (1-p)^{r(n-r)}.$$

Pour borner cette quantité, on peut utiliser la formule de Stirling. Il existe une constante universelle  $C$  telle que :

$$\binom{n}{r} \leq C \sqrt{\frac{n}{r(n-r)}} \left(\frac{n}{r}\right)^r \left(\frac{n}{n-r}\right)^{n-r}.$$

Lorsque  $r \leq n/2$ , on obtient :

$$\binom{n}{r} r^{r-2} \leq \frac{C'}{r^{5/2}} n^r e^r.$$

D'autre part, lorsque  $r \leq n/2$ ,

$$(1-p)^{r(n-r)} \leq e^{-npr/2}.$$

Ainsi,

$$\binom{n}{r} r^{r-2} p^{r-1} (1-p)^{r(n-r)} \lesssim \frac{1}{p} \left[enpe^{-np/2}\right]^r.$$

Comme  $np$  tend vers l'infini,  $enpe^{-np/2}$  tend vers 0 et donc, pour  $n$  assez grand, la probabilité que  $G(n, p)$  contienne une composante connexe d'ordre appartenant à  $[3, n/2]$  est dominée par :

$$\sum_{r=3}^{[n/2]} \frac{1}{p} \left[enpe^{-np/2}\right]^r \lesssim \frac{1}{p} \left[enpe^{-np/2}\right]^3 = ne^{-3np/2} (np)^2.$$

Pour  $np \geq \alpha \log n$  avec  $\alpha > 2/3$  ceci tend vers 0. □



### 3.3.4 Connexité et sommets isolés dans le modèle $G(n, p)$

Un des premiers résultats portant sur la connexité dans le modèle  $G(n, p)$  date de l'article de Gilbert [24], où il y donne une formule exacte de la probabilité qu'un graphe aléatoire soit connexe sous le modèle  $G(n, p)$ , notée  $\pi_{n,p}$ . Il utilise une méthode courante en combinatoire, basée sur les séries génératrices ; il écrit d'abord  $\pi_{n,p}$  en fonction de  $C_{n,l}$ , défini comme le nombre de graphes connexes à  $n$  sommets et  $l$  arêtes. On l'a vu plus tôt, si  $N = \binom{n}{2}$ , tout graphe à  $n$  noeuds et  $l$  arêtes a pour probabilité d'apparition  $p^l q^{N-l}$  dans ce modèle, où  $q = 1 - p$ . D'où :

$$\pi_{n,p} = \sum_{l=n-1}^N C_{n,l} p^l q^{N-l}.$$

On note qu'un graphe à  $n$  sommets connexe a au moins  $n - 1$  arêtes (s'il en a exactement  $n - 1$ , c'est même un arbre), et donc  $C_{n,l} = 0$  pour tout  $l < n - 1$ . Gilbert décrit une série génératrice pour la suite double  $(C_{n,l})_{n \in \mathbb{N}^*, l \in \mathbb{N}}$  et la convertit en série génératrice pour  $\pi_{n,p}$ . Par des développements en série formelle et en notant  $\mathcal{R}_n$  l'ensemble des partitions <sup>(1)</sup> de l'entier  $n$ . Gilbert obtient :

$$\pi_{n,p} = n! \sum_{(r_1, \dots, r_n) \in \mathcal{R}_n} \frac{(-1)^s (s-1)! q^{(n^2 - 1^2 r_1 - \dots - n^2 r_n)/2}}{r_1! \dots r_n! (1!)^{r_1} \dots (n!)^{r_n}},$$

où dans la somme,  $s = r_1 + \dots + r_n$ .  $\pi_{n,p}$  s'écrit donc comme une somme dont le nombre de termes est le cardinal de  $\mathcal{P}_n$ .

Combien d'arêtes avons-nous besoin pour faire un graphe aléatoire connexe ? Le seuil de connectivité en  $G(n, m)$  a été donné par Erdős et Rényi en 1960 dans un document séminal sur les graphes aléatoires. La preuve consiste à estimer le nombre prévu de petites composantes.

**Théorème 3.3.4** [19].

Si  $m = \frac{1}{2}n(\log n + c_n)$ , et  $c_n \rightarrow c$  quand  $n \rightarrow \infty$ , alors

$$\mathbb{P}(G(n, m) \text{ est connecté}) \rightarrow e^{-e^{-c}} \text{ quand } n \rightarrow \infty.$$

Il s'ensuit que si  $c_n \rightarrow -\infty$  alors la probabilité  $\rightarrow 0$ , et si  $c_n \rightarrow \infty$  alors la probabilité  $\rightarrow 1$ .

---

(1). Une partition de l'entier  $n$  est un  $n$ -uplet d'entiers naturels  $(r_1, \dots, r_n)$  tel que  $r_1 + 2r_2 + \dots + nr_n = n$

### 3.4 Evolution d'un graphe aléatoire

Ici nous résumons les étapes distinctes de l'évolution du graphe aléatoire Erdős-Rényi.

**Étape I :**  $p = o(1/n)$

Le graphe aléatoire  $G(n, p)$  est l'union disjointe des arbres. En fait, comme on vous demande de prouver dans l'un des problèmes de l'examen, il n'y a pas d'arbres d'ordre  $k$  si  $pn^{k/(k-1)} = o(1)$ . En outre, pour  $p = cn^{k/(k-1)}$  et  $c > 0$ , la distribution de probabilité du nombre d'arbres d'ordre  $k$  tend vers la distribution de Poisson avec le paramètre  $\lambda = c^{k-1}k^{k-2}/k!$ . Si  $1/(pn^{k/(k-1)}) = o(1)$  et  $pkn - \log n - (k-1)\log \log n \rightarrow \infty$ , alors il y a des arbres de n'importe quel ordre a.a.s. Si  $1/(pn^{k/(k-1)}) = o(1)$  et  $pkn - \log n - (k-1)\log \log n \sim x$  alors les arbres d'ordre  $k$  répartis asymptotiquement par Poisson Loi avec le paramètre  $\lambda = e^{-x}/(kk!)$ .

**Étape II :**  $p \sim c/n$  pour  $0 < c < 1$

Des cycles de n'importe quelle taille apparaissent. Tous les composants connexes de  $G(n, p)$  sont soit des arbres, soit des composants monocycles (arbres avec une arête supplémentaire). Presque tous les sommets des composantes qui sont des arbres ( $n - o(n)$ ). La plus grande composante connexe est un arbre et a environ  $\alpha^{-1}(\log n - 2.5 \log \log n)$  vertices, où  $\alpha^{-1} = c - 1 - \log c$ . La moyenne du nombre de composants connexes est  $n - p\binom{n}{2} + \mathcal{O}(1)$ , c'est-à-dire, l'addition d'un nouvel arête diminue le nombre de composantes connectées d'une unité. La distribution du nombre de cycles sur  $k$  sommets est approximativement une distribution de Poisson avec  $\lambda = c^k/(2k)$ .

**Étape III :**  $p \sim 1/n + \mu n$ , le double saut

Apparition du composant géant. Lorsque  $p < 1/n$ , alors la taille de la plus grande composante est  $\mathcal{O}(\log n)$  et la plupart des sommets appartiennent aux composantes de la taille  $\mathcal{O}(1)$ , alors que pour  $p > 1/n$  la taille de la plus grande composante est  $\mathcal{O}(n)$ , les autres composantes sont toutes petites, la plus grande étant de l'ordre de  $\mathcal{O}(\log n)$ . Tous les composants autres que le géant sont des arbres ou unicycliques, bien que le composant géant ait une structure complexe (il ya des cycles de n'importe quelle période). La question naturelle est de savoir comment le plus grand composant se développe si rapidement. Erdős et Rényi montrent qu'il se produit en fait en deux étapes, d'où le terme de double saut. Si  $\mu < 0$ , la plus grande composante a la taille  $(\mu - \log(1 + \mu))^{-1} \log n + \mathcal{O}(\log \log n)$ . Si  $\mu = 0$  alors la plus grande composante a la taille d'ordre  $n^{2/3}$ , et pour  $\mu > 0$  la composante géante a la taille  $\alpha n$  pour une certaine constante  $\alpha$ .

**Étape IV :**  $p \sim c/n$  où  $c > 1$

Sauf pour un composant géant tous les composants sont petits, et la plupart d'entre eux sont des arbres. L'évolution du graphe aléatoire ici peut être décrite comme fusionnant les plus petits composants avec le géant, l'un après l'autre. Plus la composante est petite, plus la chance de survie est grande. Le temps de survie d'un arbre d'ordre  $k$  est approximativement exponentiellement distribué avec la valeur moyenne  $n/(2k)$ .

**Étape V :**  $p = c \log n/n$  avec  $c \geq 1$

Le graphe aléatoire devient connecté. Pour  $c = 1$ , il n'y a que la composante géante et les sommets isolés.

**Étape VI :**  $p = \omega(n) \log n/n$  où  $\omega(n) \rightarrow \infty$  quand  $n \rightarrow \infty$ . Dans cette plage, le graphe aléatoire n'est pas seulement connecté, mais aussi les degrés de tous les sommets sont asymptotiquement égaux.

Voici l'illustration numérique de l'évolution du graphe aléatoire. Pour le présenter, j'ai fixé  $n = 1000$  le nombre de sommets et généré  $\binom{n}{2}$  variables aléatoires, uniformément réparties en  $[0, 1]$ . Ainsi, chaque arête obtient son seul nombre  $p_j \in [0, 1]$ ,  $j = 1, \dots, \binom{n}{2}$ . Pour tout  $p \in (0, 1)$  fixe, je trace seulement les arêtes pour lesquelles  $p_j \leq p$ . Par conséquent, de cette manière, je peux observer comment l'évolution du graphe aléatoire se produit pour  $G(n, p)$ .

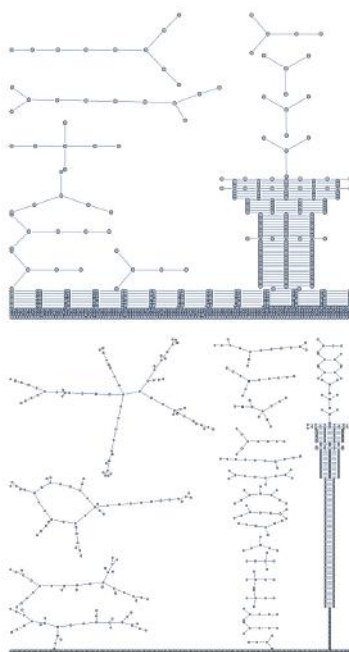
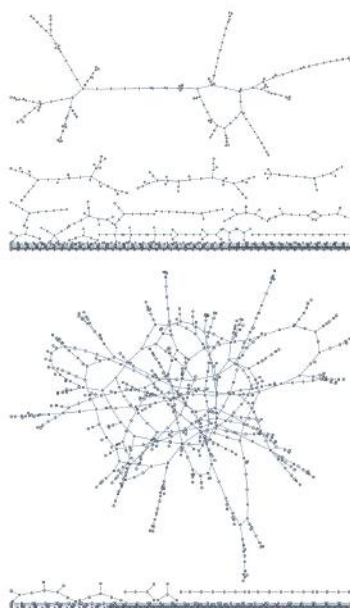
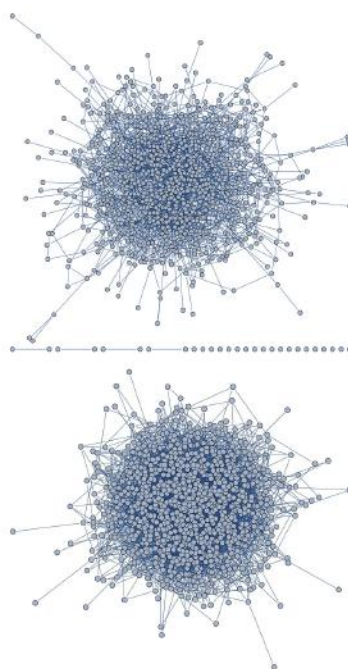


FIGURE 3.2 – Étapes I et II. Les graphes  $G(1000, 0.0005)$  et  $G(1000, 0.00095)$  sont représentés. [36]



**FIGURE 3.3** – Étapes III et IV. Les graphes  $G(1000, 0.001)$  et  $G(1000, 0.0015)$  sont représentés. La composante géante est née.[36]



**FIGURE 3.4** – Étapes V et VI. Les graphes  $G(1000, 0.004)$  et  $G(1000, 0.007)$  sont représentés. Le graphe final est connecté.[36]

### 3.5 Cycle hamiltonien

La détermination du seuil pour l'existence d'un cycle hamiltonien était un problème ouvert majeur depuis de nombreuses années. L'idée-clé s'est révélée être la rotation-extension. Supposons que nous avons un chaîne de longueur  $k$ , avec des sommets  $v_0, v_1, \dots, v_k$ . Si  $v_0$  ou  $v_k$  a un voisin qui n'est pas dans le chaîne, alors nous pouvons étendre le chaîne en ajoutant simplement un tel voisin. À défaut, supposons que  $v_k$  ait un voisin  $v_i$ , où  $0 \leq i \leq k - 2$ . Si  $i = 0$  et qu'il existe un arête  $v_i w$  joignant le cycle  $v_0, v_1, \dots, v_k, v_0$  au reste du graphe, alors la chaîne  $w, v_i, v_{i+1}, \dots, v_k, v_0, \dots, v_{k-1}$  est de longueur  $k + 1$ . Si  $i \neq 0$ , alors en effectuant une rotation, la chaîne  $v_0, \dots, v_i, v_k, v_{k-1}, \dots, v_{i+1}$  est construite; Cette chaîne a toujours la longueur  $k$ , mais a un point final différent  $v_{i+1}$  à partir duquel nous pouvons à nouveau chercher une extension. En utilisant des rotations et des extensions, Komlos et Szemerédi [11] ont finalement résolu le problème du seuil en 1983.

Le graphe aléatoire  $G(n, m)$  à également des sommets  $v_1, \dots, v_2$ , chaque graphe de ce type avec exactement  $m$  arête est également probable. Son comportement très similaire à celui de  $G(n, p)$  avec  $p = m/\binom{n}{2}$ .

**Théorème 3.5.1** [11].

Si  $m = \frac{1}{2}n(\log n + \log \log n + c_n)$ , et  $c_n \rightarrow c$  quand  $n \rightarrow \infty$ , alors

$$\mathbb{P}(G(n, m) \text{ est Hamiltonien}) \rightarrow e^{-e^{-c}} \text{ quand } n \rightarrow \infty.$$

Il existe en effet un algorithme de complexité polynomiale, basé sur l'idée de la rotation-extension, qui trouve un cycle hamiltonien avec la probabilité indiquée dans le Théorème 3.5.1 [11]. Il a été aussi montré que pour tout  $r \geq 3$  fixe, un graphe  $r$ -régulier aléatoire sur  $n$  sommets est a.s. Hamiltonien [33, 34].

**Théorème 3.5.2** [23].

Soit  $m = \lfloor cn \log n \rfloor$ . Alors, un graphe aléatoire  $G_{n,m} \in G(n, m)$  presque sûrement contient un cycle Hamiltonien.

En d'autres termes, pour un certain  $c$  constant, presque chaque graphe avec  $n$  sommets et au moins  $cn \log n$  arêtes est hamiltonien. En effet,  $c$  est choisi suffisamment grand pour garantir que si  $G$  est un graphe avec  $n$  sommets et les arêtes sont construites indépendamment avec probabilité  $\frac{c \log n}{n}$ , alors  $\mathbb{P}[G \text{ a un cycle Hamiltonien}] \rightarrow 1$  quand  $n \rightarrow \infty$ .

Pour la détermination de la fonction seuil pour l'existence d'un cycle hamiltonien, notons

tout d'abord que si  $G$  est hamiltonien, alors  $\delta(G) \geq 2$ , donc :

$$\mathbb{P}[G_m \text{ Hamiltonien}] \leq \mathbb{P}[\delta(G_m) \geq 2].$$

On peut montrer que  $\mathbb{P}[\delta(G_m) \geq 2] \rightarrow 1$  si  $\omega(n) = 2m/n - \log n - \log \log n \rightarrow \infty$ . Par conséquent, si a.e.  $G_m$  est hamiltonien, on doit avoir  $\omega(n) \rightarrow \infty$ . Il s'avère que cette condition est également suffisante, et a été d'abord prouvée par Komlós et Szemerédi en utilisant des rotations et des extensions de cycles [28].

Pour le graphe aléatoire du modèle binomial  $G(n, p)$ , On énoncer le résultat suivant :

**Théorème 3.5.3** [8, 1].

Soit  $p = \frac{\log n + \log \log n + c_n}{n}$ . Alors

$$\lim_{n \rightarrow \infty} \mathbb{P}(G(n, p) \text{ un cycle Hamiltonien}) = \begin{cases} 0 & \text{if } c_n \rightarrow -\infty \\ e^{-e^{-c}} & \text{if } c_n \rightarrow c \\ 1 & \text{if } c_n \rightarrow \infty \end{cases}$$

De plus,

$$\lim_{n \rightarrow \infty} \mathbb{P}(G(n, p) \text{ un cycle Hamiltonien}) = \lim_{n \rightarrow \infty} \mathbb{P}(\delta(G(n, p)) \geq 2).$$

Le théorème suivant donne le hamiltonien de a.e.  $G_p$  et a.e.  $G_m$ .

**Théorème 3.5.4** .

Soit  $\omega(n) \rightarrow \infty$ ,  $p = (1/n)(\log n + \log \log n + \omega(n))$ , et

$m(n) = \lfloor (n/2)(\log n + \log \log n + \omega(n)) \rfloor$ . Alors a.e.  $G_p$  et a.e.  $G_m$  sont hamiltoniens.

## 3.6 Stabilité d'un graphe aléatoire

**Théorème 3.6.1** [21].

Un graphe aléatoire de  $G(n, p)$  a presque sûrement une stabilité d'au plus  $\lceil 2p^{-1} \log n \rceil$ .

Afin de prouver ce théorème, nous avons besoin du corollaire suivant :

**Corollaire 3.6.1** .

Soit  $X_n$  une variable aléatoire à valeurs réelles positives dans un espace de probabilité  $(\Omega_n, \mathbb{P}_n)$ ,  $n \geq 1$ . Si  $\mathbb{E}(X_n) \rightarrow 0$  quand  $n \rightarrow \infty$ , alors  $\mathbb{P}(X_n = 0) \rightarrow 1$  quand  $n \rightarrow \infty$ .

*preuve du Théorème 3.6.1.* Soit  $G \in G(n, p)$  et soit  $S$  un ensemble de  $k + 1$  sommets de  $G$ , avec  $k \in \mathbb{N}$ . La probabilité que  $S$  soit un stable de  $G$  est  $(1 - p)^{\binom{k+1}{2}}$ , car c'est la probabilité qu'aucune des  $\binom{k+1}{2}$  paires de sommets de  $S$  ne soit une arête du graphe aléatoire  $G$ .

Soit  $A_S$  l'évènement que  $S$  soit un stable de  $G$ , et soit  $X_S$  la variable aléatoire caractéristique de cet évènement. Nous avons :

$$\mathbb{E}(X_S) = \mathbb{P}(X_S = 1) = \mathbb{P}(A_S) = (1 - p)^{\binom{k+1}{2}}. \quad (3)$$

Soit  $X$  le nombre de stables de cardinal  $k + 1$  dans  $G$ . Alors

$$X = \sum \{\mathbb{E}(X_S) : S \subseteq V, |S| = k + 1\},$$

et donc, par linéarité de l'espérance et l'équation 3,

$$\mathbb{E}(X) = \sum \{\mathbb{E}(X_S) : S \subseteq V, |S| = k + 1\} = \binom{n}{k+1} (1 - p)^{\binom{k+1}{2}}.$$

Nous majorons le membre droit à l'aide de deux inégalités élémentaires :

$$\binom{n}{k+1} \leq \frac{n^{k+1}}{(k+1)!} \text{ et } 1 - p \leq e^{-p}.$$

Cela donne la borne supérieure suivante pour  $\mathbb{E}(X)$ ,

$$\mathbb{E}(X) \leq \frac{n^{k+1} e^{-p \binom{k+1}{2}}}{(k+1)!} = \frac{(ne^{-pk/2})^{k+1}}{(k+1)!}. \quad (4)$$

Supposons maintenant que  $k = c$ . Alors  $k \geq 2p^{-1} \log n$ , donc  $ne^{-pk/2} \leq 1$ . Comme  $k$  croît au moins aussi vite que le logarithme de  $n$ , l'inégalité 4 implique que  $\mathbb{E}(X) \rightarrow 0$  quand  $n \rightarrow \infty$ . Comme  $X$  est à valeurs entières positives, nous déduisons du Corollaire 3.6.1 que  $\mathbb{P}(X = 0) \rightarrow 1$  quand  $n \rightarrow \infty$ . Par conséquent, un graphe aléatoire de  $G(n, p)$  est presque sûrement de stabilité au plus  $\lceil 2p^{-1} \log n \rceil$ .  $\square$

Dans le cas où  $p = \frac{1}{2}$ , une borne sur  $\alpha(G)$ , légèrement meilleure que celle fournie par le Théorème 3.6.1 a été obtenue par ...comme l'indique le résultat suivant :

**Théorème 3.6.2 [10].**

Si  $G \in G(n, 1/2)$ , alors presque sûrement  $\alpha(G) \leq \lceil 2 \log_2 n \rceil$ .

*Preuve.*

Soit  $n$  un entier strictement positif. Pour  $0 \leq k \leq n$ , on pose  $f(k) := \binom{n}{k} 2^{-\binom{k}{2}}$ . Soit  $k^* = k^*(n)$  la plus petite valeur de  $k$  pour laquelle  $f(k)$  est inférieur à un. On montre que :

$$\begin{aligned} k^* &\leq \lceil 2 \log_2 n \rceil \leq k^* + \log_2 k^* - 1, \\ f(k^* + 1) &\rightarrow 0 \text{ quand } n \rightarrow \infty. \end{aligned}$$

Le résultat en découle. □

## 3.7 Nombre chromatique d'un graphe aléatoire

### 3.7.1 Ordre de grandeur du nombre chromatique

Le théorème suivant est dû à Bollobàs :

**Théorème 3.7.1** [9].

Si  $G \in G(n, 1/2)$ , alors presque sûrement  $\chi(G) \sim \frac{n}{2 \log_2 n}$ .

*Preuve.*

Nous venons de voir que  $\chi(G) \sim \frac{n}{2 \log_2 n} (1 + o(1))$ . Il nous suffit donc de montrer l'inégalité inverse.

Posons  $s = \lfloor n / \log_2 n \rfloor$ . Pour tout ensemble  $S$  de  $s$  sommets le sous-graphe induit  $G[S]$  a la distribution de  $G(s, 1/2)$ . Soit  $k' = k'(s) = k(s)^* - 4$  avec  $k^* = k^*(n)$  la plus petite valeur de  $k$  pour laquelle  $f(k)$  est inférieur à un. Nous avons

$$k' \sim 2 \log_2 s \sim \log_2 n.$$

on peut alors montrer que

$$\mathbb{P}(\alpha(G[S]) < k') < e^{-s^{2+o(1)}}.$$

Il y a  $\binom{n}{s} < 2^n = 2^{s^{1+o(1)}}$  tels ensembles  $S$ . D'où

$$\mathbb{P}(\alpha(G[S]) < k' \text{ pour un ensemble } S \text{ de cardinal } s) < 2^{s^{1+o(1)}} e^{-s^{2+o(1)}} = o(1).$$

Ainsi, presque sûrement, tous les ensembles de  $s$  sommets de  $G$  contiennent un stable de cardinal  $k'$ .

Maintenant supposons qu'un graphe  $G$  ait cette propriété. On peut enlever des stables de cardinal  $k$  et leurs donner des couleurs différentes jusqu'à ce qu'il reste moins de  $s$  sommets. On peut alors donner une couleur différente à chacun des sommets restants. Cette procédure donne



$$\begin{aligned}
\chi(G) &\leq \frac{n-s}{k'} + s \leq \frac{n}{k'} + s \\
&\leq \frac{n}{2 \log_2 n} (1 + o(1)) + o\left(\frac{n}{\log_2 n}\right) \\
&\leq \frac{n}{2 \log_2 n} (1 + o(1)).
\end{aligned}$$

□

### 3.7.2 Concentration du nombre chromatique

Le nombre chromatique d'un graphe aléatoire, comme beaucoup d'autres paramètres de graphes, est lui aussi concentré autour de son espérance. D'ailleurs, on peut montrer cette concentration sans même connaître la valeur de l'espérance.

**Théorème 3.7.2 (Inégalité d'Azuma) [5].**

Soit  $X_1, \dots, X_p$  une martingale telle que  $X_1 = 0$  et  $|X_{i+1} - X_i| \leq 1$  pour tout  $1 \leq i < p$ . Pour tout réel  $t$  strictement positif,

$$\mathbb{P}(X_p > t) < e^{-\frac{t^2}{2p-2}}.$$

Un paramètre de graphe  $f$  est lipschitzien (sur les sommets) si quels que soient deux graphes  $G$  et  $H$  qui diffèrent d'un seul sommet alors  $|f(H) - f(G)| \leq 1$ . De nombreux paramètres de graphes sont lipschitziens. C'est en particulier le cas de la stabilité  $\alpha(G)$  et du nombre chromatique  $\chi(G)$ .

En effet, supposons que  $H = G - v$  pour un certain sommet  $v$ . Toute coloration propre de  $G$  induit une coloration propre de  $H$ . À l'inverse, toute  $k$ -coloration propre de  $H$  peut être étendue en une  $(k+1)$ -coloration propre de  $G$  en donnant à  $v$  une nouvelle couleur. Donc  $\chi(H) \leq \chi(G) \leq \chi(H) + 1$ .

On peut montrer que si  $f$  est un paramètre de graphe lipschitzien sur les sommets, alors la martingale de révélation des sommets correspondante  $X_1, \dots, X_n$  satisfait  $|X_{i+1} - X_i| \leq 1$  pour tout  $1 \leq i < n$ . Intuitivement, cela veut dire que si la connaissance d'un sommet ne change la valeur de  $f$  que de 1 au plus, alors révéler un sommet ne peut pas changer l'espérance de  $f$  de plus que 1. Considérons alors la martingale  $X'_1, X'_2, \dots, X'_n$ , définie par  $X'_i = X_i - E(f(G))$ . On a  $X'_0 = 0$  et  $|X'_{i+1} - X'_i| \leq 1$  pour tout  $1 \leq i < n$ . On peut donc appliquer l'Inégalité d'Azuma sur cette martingale et obtenir le théorème suivant :

**Théorème 3.7.3** [9].

Soit  $G \in G(n, p)$ . Si  $f$  satisfait la condition de Lipschitz (sur les sommets), alors

$$\mathbb{P}(|f(G) - \mathbb{E}f(G)| > t) < 2e^{-\frac{t^2}{2n-2}}.$$

**Remarque 1 .**

Le Théorème 3.7.3 possède un analogue arête. La condition de Lipschitz porte alors sur la différence d'une arête et la martingale de révélation des sommets est remplacée par la martingale de révélation d'arêtes.

### 3.7.3 Existence d'un graphe ayant un nombre chromatique donné

En utilisant la méthode probabiliste, Erdős avait montré le théorème suivant qui affirme l'existence d'un graphe  $G$  avec des conditions sur les valeurs de la circonférence  $\ell(G)$  et du nombre chromatique  $\chi(G)$ .

**Théorème 3.7.4** [18].

(Erdős, 1959). Pour tous  $k, \ell$ , il existe un graphe  $G$  tel que  $\ell(G) > \ell$  et  $\chi(G) > k$ .

*Preuve.*

Soit  $G \sim G(n, p)$ . On va ajuster la valeur de  $p$  en fonction de  $n$  (tout en prenant  $n$  grand) de manière que :

- On puisse trouver, avec une probabilité supérieure à  $1/2$ , un graphe tel que  $\ell(G) > \ell$ .
- On puisse trouver, avec une probabilité supérieure à  $1/2$ , un graphe tel que  $\chi(G) > k$ .

Il n'est pas très difficile de s'arranger pour qu'il y ait beaucoup de graphes qui possèdent peu de circuits de longueur inférieure ou égale à  $k$ . Une fois un tel graphe trouvé, on manipulera ce graphe de telle manière à ce qu'il n'ait plus du tout de cycles de longueur inférieure ou égale à  $k$ .

Soit en effet  $X$  le nombre de cycles de longueur au plus  $l$ . Le nombre de (classes d'équivalence de) cycles de longueur  $i$  est  $\binom{n}{i} 2i$ . On en déduit que

$$\mathbb{E}X = \sum_{i=3}^l \binom{n}{i} 2ip^i \leq \sum_{i=3}^l \frac{1}{2i \cdot i!} (np)^i \leq k (np)^l.$$

Ainsi

$$\mathbb{P}(X \geq n/2) \leq \frac{2\mathbb{E}X}{n} \leq 2K \frac{(np)^l}{n}.$$

Il faudra que cette probabilité tende vers zéro, mais on peut voir tout de suite qu'il est nécessaire que  $p_n$  tende vers 0.

Pour minorer le nombre chromatique, on va majorer  $\alpha(G)$ . Si on note  $Y$  le nombre de stables de cardinal  $x$ , on a

$$\mathbb{P}(\alpha(G) \geq x) = \mathbb{P}(Y \geq 1) \leq \mathbb{E}Y.$$

Or

$$\mathbb{E}Y = \binom{n}{x} (1-p)^{\binom{x}{2}} \leq n^x (1-p)^{\binom{x}{2}},$$

de telle sorte que

$$\ln \mathbb{E}Y \leq x \ln n + \frac{x(x-1)}{2} \ln(1-p) = x \left( \ln n + \frac{x-1}{2} \ln(1-p) \right)$$

Si  $x_n$  tend vers  $+\infty$  et  $p_n$  vers 0, on a

$$\frac{x_n - 1}{2} \ln(1 - p_n) \sim \frac{x_n}{2p_n}$$

De la sorte, prendre une suite d'entiers telle que  $\frac{x_n}{2p_n} \sim K \ln n$ , avec  $K > 1$  est un choix assez naturel pour avoir  $\lim_{n \rightarrow +\infty} P((G(n, p_n))x_n) = 0$ . On pourra par exemple prendre pour  $x_n$  la partie entière de  $\frac{3 \ln n}{p_n}$ . (Bien sûr, il faudra vérifier que  $x_n$  tend vers  $+\infty$ ).

Prenons  $n$  suffisamment grand pour que  $\mathbb{P}((G(n, p_n))x_n) < 1/2$  et que  $\mathbb{P}(X \geq n/2) < 1/2$ . Il existe un graphe à  $n$  sommets qui possède au plus  $n/2$  cycles de longueurs inférieure ou égale à  $l$  et tel que  $\alpha(G) < x_n$ .

Construction à partir de  $G$  un nouveau graphe  $G'$  en retirant un sommet dans chaque cycle de longueur inférieure ou égale à  $l$ . On a maintenant  $\ell(G') \geq l$  et  $|G'| \geq |G| - n/2 = n/2$ . Comme on construit  $G'$  en enlevant des points à  $G$ , tout stable de  $G'$  est un stable de  $G$  donc  $\alpha(G') \leq \alpha(G) < x_n$ . On en déduit

$$\chi(G') \geq \frac{|G'|}{\alpha(G')} \geq \frac{n/2}{x_n}.$$

Si  $\frac{n/2}{x_n} > k$ , on a gagné : on veut donc que  $\frac{n}{x_n} \sim \frac{3np_n}{\ln n}$  tende vers  $+\infty$ .

Il faut maintenant choisir  $p_n$  : on veut  $\frac{np_n}{\ln n} \rightarrow +\infty$ ,  $\frac{\ln n}{p_n} \rightarrow +\infty$ ,  $p_n \rightarrow 0$  et  $\frac{(np_n)^l}{n} \rightarrow 0$ . On cherche naturellement  $p_n$  sous la forme  $p_n = n^{-\beta}$  et on constate facilement que tout  $\beta > \frac{l}{l-1}$  convient.  $\square$

---

## Conclusion

---

L'étude des graphes aléatoires est un domaine de recherche actif qui combine la théorie des graphes et la théorie des probabilités, qui a été introduit pour prouver certains résultats sur les graphes.

Dans ce travail, on a présenté le modèle de graphes aléatoires d'Erdős-Rényi, qui se basait sur l'espace de probabilité  $(\Omega, \mathcal{F}, \mathbb{P})$ .

Après avoir présenté quelques concepts de base de la théorie des graphes, dans le premier chapitre, nous avons introduit dans le second, la méthode probabiliste qui consiste on ensemble d'outils de probabilité élémentaire, utilisés pour prouver certaines propriétés combinatoires.

Dans le dernier chapitre, nous nous sommes consacrés à l'étude de deux modèles des graphes aléatoires à savoir, le modèle binomial et le modèle uniforme. Quelques propriétés des graphes aléatoires ont alors été présentées en mettant en évidence les techniques de preuve de la méthode probabiliste.

---

## Bibliographie

---

- [1] M. Ajtai, J. Komlós, and E. Szeherédi. First occurrence of hamilton cycles in random graphs. *North-Holland Mathematics Studies*, 115 :173–178, 1985.
- [2] D. Aldous. Brownian excursions, critical random graphs and the multiplicative coalescent. *The Annals of Probability*, pages 812–854, 1997.
- [3] D. Aldous and R. Durrett. Random graph dynamics, 2010.
- [4] N. Alon and J. Spencer. *The probabilistic Method*.
- [5] K. Azuma. Weighted sums of certain dependent random variables. (02) :357–367.
- [6] A. Barbour and D. Mollison. Epidemics and random graphs. In *Stochastic processes in epidemic theory*, pages 86–89. Springer, 1990.
- [7] B. Bollobás. Random graphs. *Cambridge University Press, Cambridge*.
- [8] B. Bollobás. The evolution of random graphs. *Transactions of the American Mathematical Society*, 286(1) :257–274, 1984.
- [9] B. Bollobás. The chromatic number of random graphs. *Combinatorica*, 8(1) :49–55, 1988.
- [10] B. Bollobás and P. Erdős. Cliques in random graphs. In *Mathematical Proceedings of the Cambridge Philosophical Society*, volume 80, pages 419–427. Cambridge Univ Press, 1976.
- [11] B. Bollobás, T. I. Fenner, and A. M. Frieze. An algorithm for finding hamilton paths and cycles in random graphs. *Combinatorica*, 7(4) :327–341, 1987.
- [12] J. Bondy and U. Murty. *Graph Theory*.
- [13] A. Channarond. Random graph models : an overview of modeling approaches. *Journal de la Société Française de Statistique*, 156(3) :56–94, 2015.
- [14] L. Doin-IME-USP and R. Bissacot-IME-USP. Legitimate coloring of finite projective planes via entropy compression.

- 
- [15] M. Draief and L. Massoulié. Epidemics and rumours in complex networks, volume 369 of london mathematical society lecture notes. *Cambridge University Press, Cambridge*, 1(52) :103, 2010.
- [16] R. Durrett. *Random graph dynamics*, volume 200. Cambridge university press Cambridge, 2007.
- [17] P. Erdős and A. Rényi. On random graphs, i. *Publicationes Mathematicae (Debrecen)*, 6 :290–297, 1959.
- [18] P. Erdős and A. Rényi. On random graphs, i. *Publicationes Mathematicae (Debrecen)*, 6 :290–297, 1959.
- [19] P. Erdős and A. Rényi. On the evolution of random graphs. *Pul. Math. Inst. Hungar. Acad. Sci.*, (5) :17–61, 1960.
- [20] P. Erdős and A. Rényi. On the evolution of random graphs. *Bull. Inst. Internat. Statist*, 38(4) :343–347, 1961.
- [21] P. Erdős and A. Rényi. On the strength of connectedness of a random graph. *Acta Mathematica Hungarica*, 12(1-2) :261–267, 1961.
- [22] P. Erdős and A. Rényi. On the evolution of random graphs. *Selected Papers of Alfréd Rényi*, 2 :482–525, 1976.
- [23] A. Frieze and M. Karoński. *Introduction to random graphs*. Cambridge University Press, 2015.
- [24] E. N. Gilbert. Random graphs. *The Annals of Mathematical Statistics*, 30(4) :1141–1144, 1959.
- [25] M. Gould. Ramsey theory. *University of Oxford*, pages 1–15, 2010.
- [26] M. Heyvaert and F. T. Bruss. La méthode probabiliste. *Gazette des Mathématiciens*, (124) :15–29.
- [27] S. Janson, D. E. Knuth, T. Łuczak, and B. Pittel. The birth of the giant component. *Random Structures & Algorithms*, 4(3) :233–358, 1993.
- [28] J. Komlós and E. Szemerédi. Limit distribution for the existence of hamiltonian cycles in a random graph. *Discrete Mathematics*, 43(1) :55–63, 1983.
- [29] J. Matoušek and J. Vondrák. The probabilistic method. *Lecture Notes, Department of Applied Mathematics, Charles University, Prague*, 2001.
- [30] C. McDiarmid. Clutter percolation and random graphs. *Combinatorial Optimization II*, pages 17–25, 1980.

- 
- [31] D. Müller. *Introduction à la théorie des graphes*. Commission romande de mathématique, 2012.
- [32] L. Pósa. Hamiltonian circuits in random graphs. *Discrete Mathematics*, 14(4) :359–364, 1976.
- [33] R. W. Robinson and N. C. Wormald. Almost all cubic graphs are hamiltonian. *Random Structures and Algorithms*, 3(2) :117–126, 1992.
- [34] R. W. Robinson and N. C. Wormald. Almost all regular graphs are hamiltonian. *Random Structures & Algorithms*, 5(2) :363–374, 1994.
- [35] R. Rossignol. Graphes aléatoires. 2012.
- [36] T. S. Janson and V. Kolchin. Random graphs, 2000.
- [37] Y. Velenik. Probabilités et statistique. *Université de Geneve*, 2011.
- [38] H. Whitney. Congruent graphs and the connectivity of graphs. *American Journal of Mathematics*, 1932.



## Résumé

Un graphe aléatoire est défini sur  $n$  sommets et  $m$  arêtes, où toute arête existe avec une probabilité  $p$ . Les graphes aléatoires ont été introduits par Erdős et Rényi en 1959 et sont largement étudiés depuis, puisqu'ils permettent de modéliser des réseaux concrets tels que les réseaux de communication et les réseaux sociaux où les sommets sont aléatoirement reliés.

Plusieurs modèles de graphes aléatoires ont été définis dont les plus connus sont le modèle binomial et le modèle uniforme. La méthode, dite probabiliste, est alors utilisée pour étudier leurs propriétés.

L'objectif de ce travail est de présenter les différentes techniques probabilistes utilisées pour montrer quelques propriétés des graphes aléatoires.

**Mots clés** : sommets, arêtes, graphe aléatoire, méthode probabiliste.

## Abstract

A random graph is defined on  $n$  vertices and  $m$  edges, where any edge exists with a probability  $p$ . The random graphs were introduced by Erdős and Rényi in 1959 and have been widely studied, since they allow to model concrete networks such as communication networks and social networks where the nodes are randomly linked.

Several models of random graphs have been defined, the best known of which are the binomial model and the uniform model. The probabilistic method is then used to study their properties.

The objective of this work is to present the different probabilistic techniques used to show certain properties of random graphs.

**Key words** : vertices, edges, random graph, probabilistic method.

## المخلص

يعرف الرسم البياني العشوائي بـ  $n$  عقدة و  $m$  ضلع، حيث كل الأضلاع موجودة باحتمال  $p$ . أنشأت الرسومات العشوائية من قبل اردوس و ريني في 1959، وامتدت دراستها على نطاق واسع، لأنها تسمح بنمذجة الشبكات المتماسكة كشبكات التواصل والشبكات الاجتماعية، حيث تكون الأضلاع موصولة عشوائياً. وقد تم تحديد عدة نماذج من الرسومات البيانية العشوائية، منها نموذج برنولي ونموذج بينوميال وهما الأكثر شهرة. وتستعمل الطريقة الاحتمالية لدراسة خصائص الرسومات العشوائية.

الهدف من هذا العمل هو تقديم مختلف الاساليب الاحتمالية المستعملة لبرهنة بعض خصائص الرسومات البيانية العشوائية.

**الكلمات المفتاحية** : العقد، الأضلاع، الرسومات العشوائية، الطريقة الاحتمالية.