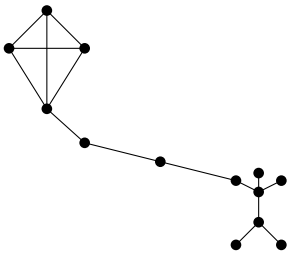
Introduction aux graphes aléatoires

(et à la méthode probabiliste)

Seconde édition
corrigée et complétée

Roger Mansuy



Fichier mis à disposition par l'auteur sur
`https://rogermansuy.fr/`

Préface

Quels que soient les progrès des connaissances humaines, il y aura toujours place pour l'ignorance et par suite pour le hasard et la probabilité.
Émile Borel, Le hasard

Les graphes aléatoires sont l'objet d'un passionnant domaine mathématique à l'intersection de l'étude des graphes et des probabilités (discrètes). Chacun des sujets est abordé en premier cycle et l'on a même croisé des exercices sur ce sujet dans les oraux de concours (comme on peut le constater en lisant [KM18] ou [BMV17] ou dans le sujet écrit de l'épreuve 2 du concours Mines-Ponts 2024). Or, les graphes, comme beaucoup de mathématiques discrètes, sont étudiés dans les cours d'informatique et les probabilités sont trop souvent associées à des questions de modélisation : les étudiants ont par conséquent rarement l'occasion de percevoir tout le charme de cette théorie à l'interface.

Ce petit volume vise à corriger ce manque en fournissant une introduction accessible à n'importe quel étudiant en mathématiques (fin de première année, que cela soit à l'université ou en classes préparatoires) curieux de découvrir de beaux objets et les résultats intéressants et parfois contre-intuitifs que l'on peut facilement obtenir en les manipulant. Fidèle à cet objectif, j'ai conclu chaque chapitre par une liste raisonnable d'exercices corrigés et commentés.

– Plan de l'ouvrage –

Le texte se découpe en une dizaine de chapitres rassemblés en quatre parties.

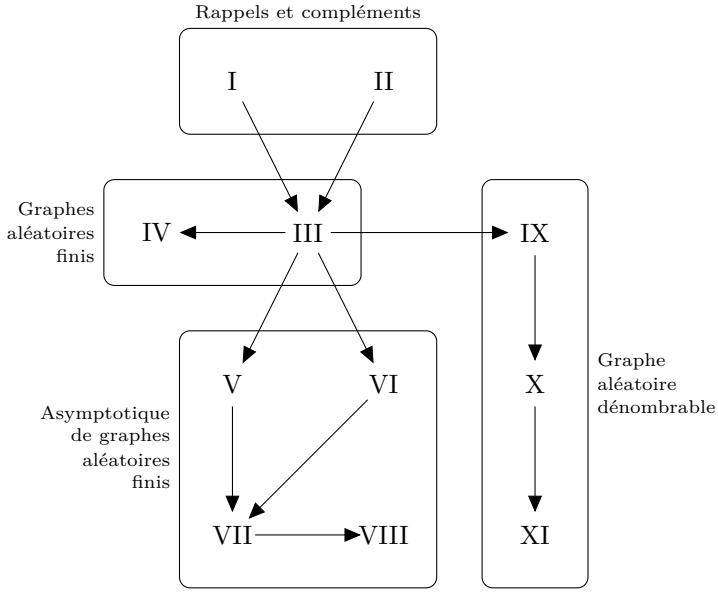
La première partie précise quelques points des deux théories sous-jacentes : les graphes et les probabilités. Il ne s'agit pas ici d'un cours complet (pour lequel il existe de très nombreux traités), mais de compléments permettant de lire sans encombre le reste de l'ouvrage. On insiste par exemple sur l'importance des variables aléatoires indicatrices, si utiles pour manipuler les variables de comptage.

La deuxième partie présente les graphes aléatoires finis. Si l'on se concentre d'abord sur le modèle binomial (quelquefois appelé de manière ambiguë graphe d'Erdős-Rényi), on développe tout de même le modèle uniforme et les applications obtenues avec la méthode probabiliste (c'est-à-dire des résultats déterministes sur les graphes obtenus en passant par les probabilités et des graphes aléatoires).

La troisième partie répond à une question simple : comment évolue un graphe aléatoire lorsque sa taille tend vers $+\infty$? Plus précisément, on met en évidence des comportements distincts (selon la valeur de paramètres) pour la limite de probabilités calculées sur le graphe aléatoire et l'on étudie un intrigant effet de « seuil » ou de « transition de phase ». Cette partie exploite judicieusement des calculs de moments (espérance et variance notamment) de variables aléatoires et les inégalités probabilistes classiques. Dans cette seconde édition, un chapitre sur le cas « limite » a été ajouté suite aux remarques des lecteurs et lectrices.

Pour la dernière partie, on abandonne les graphes aléatoires finis et leurs limites pour considérer directement un graphe aléatoire infini dénombrable. On vérifie alors que ce graphe jouit de nombreuses propriétés d'universalité, d'homogénéité et de robustesse. La propriété d'extension garantit notamment son unicité (à isomorphisme près) et permet de le réinterpréter à partir de constructions déterministes variées, notamment arithmétiques.

Voici le « graphe » des dépendances entre chapitres et parties :



On remarque que cette structure oriente naturellement vers les chapitres IV, VII/VIII et XI, points d'orgue de ce petit ouvrage : les résultats qu'ils contiennent apporteront sans aucun doute de grandes satisfactions mathématiques aux lecteurs.

En revanche, il n'apparaît pas clairement dans ce schéma que certaines idées ou méthodes (comme, par exemple, l'utilisation fréquente des variables aléatoires indicatrices) traversent l'ensemble des chapitres, c'est pourtant un élément important pour toute personne qui travaille avec cet outil.

La rédaction de cet ouvrage repose sur le principe de chapitres courts, rythmés par des sections d'exercices qui complètent logiquement le contenu du cours. Il est donc conseillé de chercher activement ces exercices avant de lire leurs corrections détaillées [1].

1. Au moment des relectures, voici mes exercices favoris : I-8.10, IV-5.4, IV-5.7, V-4.3, X-6.4 et XI-4.3.

– Remerciements –

Comme pour mes autres publications, j'ai eu la chance de recevoir de nombreux commentaires de lecteurs et de lectrices au sujet de la première édition de ce livre [2]. Qu'ils soient remerciés de leurs efforts. Ces retours signalant tantôt des imprécisions, tantôt des changements possibles ont, je l'espère, permis d'améliorer la qualité du texte pour cette seconde édition.

J'ai corrigé les coquilles, ajouté un court chapitre sur l'approximation poissonnienne (le chapitre VIII) pour répondre à certaines de leurs interrogations et complété le texte initial avec de nouveaux exercices et des illustrations supplémentaires.

Il va de soi que j'assume l'entière responsabilité des éventuelles erreurs qui resteraient ou qui auraient été introduites avec les ajouts. N'hésitez pas à me faire part de vos commentaires et questions sur cet ouvrage à l'adresse suivante :

`roger.mansuy@ac-paris.fr`

Bonne lecture !

Bagneux, août 2026

2. La première édition, désormais épuisée, était publiée par Calvage-et-Mounet en 2020.

Table des matières

Rappels et compléments

0. Quelques notations

I. Compléments de théorie des graphes

1. Définitions 9
2. Graphes courants 11
3. Voisins, connexité 13
4. Degré d'un sommet 14
5. Graphes induits, sous-graphes 15
6. Arbres 16
7. Autres invariants 17
8. Exercices 19

II. Compléments de probabilités

1. Variables aléatoires indicatrices 31
2. Variables aléatoires de comptage 32
3. Espérance 35
4. Variance 36
5. Lemmes de Borel et Cantelli 38
6. Exercices 40

Graphes aléatoires finis

III. Modèle binomial $\mathscr{G}_{n,p}$
1. Définition 51
2. Indépendance des arêtes 54
3. Dénombrements en moyenne 57
4. Modèle uniforme 61
5. Exercices 63

IV. Méthode probabiliste
1. Propriété de k-extension 73
2. Nombres de Ramsey 77
3. Nombre de chemins de longueur 3 79
4. Nombre de croisements 81
5. Exercices 84

Asymptotique de graphes aléatoires finis

V. Méthode du premier moment
1. Inégalité de Markov 101
2. Asymptotiquement sans sommets isolés 103
3. Asymptotiquement sans triangle 104
4. Exercices 105

VI. Méthode du deuxième moment
1. Inégalité de Bienaymé et Tchebychev 119
2. Asymptotiquement avec sommets isolés 121
3. Asymptotiquement avec triangle 122
4. Exercices 124

VII. Fonctions de seuil
1. Propriétés croissantes 131
2. Fonction de seuil 133
3. Théorème d'existence 136
4. Exercices 137
5. Récapitulatif 143

VIII. Probabilité au seuil
1. Moments factoriels 147
2. Cas des variables de comptage 151
3. Théorème de convergence 153
4. Nombres de triangles au seuil 156
5. Exercices 158

Graphe aléatoire dénombrable

IX. Graphe de Rado
1. Définition 163
2. Diamètre du graphe de Rado 164
3. Sous-graphes finis du graphe de Rado 166
4. Exercices 168

X. Propriété d'extension
1. Définition et premières propriétés 171
2. Construction par récurrence 173
3. Construction « binaire » 176
4. Construction par résidus quadratiques 177
5. Construction probabiliste 179
6. Exercices 179

XI. « Le » graphe aléatoire dénombrable
1. Théorème d'unicité 185
2. Le graphe aléatoire 187
3. Propriété d'universalité 188
4. Exercices 189

Bibliographie **197**
Index **199**

Rappels et compléments

Chapitre 0

Quelques notations

Notations ensemblistes

0.1. Notation. Les notations $\mathbb{N}$ et $\mathbb{N}^*$ désignent respectivement l'ensemble des entiers naturels et l'ensemble des entiers naturels non nuls.

0.2. Notation. Pour tous entiers $m \leqslant n$, $[\![m, n]\!]$ désigne l'ensemble $\{m, m+1, m+2, \ldots, n\}$ des entiers entre m et n au sens large. Pour tout entier naturel n non nul, $[\![n]\!]$ désigne l'ensemble $\{1, 2, \ldots, n\}$ des entiers entre 1 et n, soit encore $[\![n]\!] = [\![1, n]\!]$.

0.3. Notation. Pour tout ensemble E, $\mathscr{P}(E)$ désigne l'ensemble des parties de E et, pour tout entier naturel k, $\mathscr{P}_k(E)$ désigne l'ensemble des parties de E de cardinal k.

Dénombrement

0.4. Notation. Si E est un ensemble de cardinal n, alors le cardinal de l'ensemble des permutations de E (ou de manière équivalente des n-uplets d'éléments distincts de E) est égal à la factorielle de n, notée

$$n! = n(n-1)\cdots 1.$$

0.5. Notation. Si E est un ensemble de cardinal n, alors le cardinal de l'ensemble des k-uplets d'éléments distincts de E est égal à la k-ième factorielle descendante de n (aussi appelée symbole de Pochhammer), notée

$$(n)_k = \tfrac{n!}{(n-k)!} = n(n-1)\cdots(n-(k-1)).$$

On remarque que $(n)_n = n!$.

0.6. Notation. Si E est un ensemble de cardinal n, alors le cardinal de $\mathscr{P}_k(E)$, l'ensemble des parties de E de cardinal k, est le coefficient binomial « k parmi n », ce que l'on note :

$$|\mathscr{P}_k(E)| = \tbinom{n}{k} = \tfrac{n!}{k!(n-k)!} = \tfrac{(n)_k}{k!}.$$

On rappelle que

$$|\mathscr{P}(E)| = 2^n.$$

Relations de comparaison de Landau

0.7. Notation. Une suite réelle $(u_n)_n$ est négligeable devant une suite $(v_n)_n$ si

$$\forall \varepsilon > 0,\ \exists N \in \mathbb{N},\ \forall n \geqslant N, \qquad |u_n| \leqslant \varepsilon |v_n|.$$

On note alors $u_n = o(v_n)$.

0.8. Remarque. Cette définition se simplifie dans le cas d'une suite $(v_n)_n$ qui ne s'annule pas à partir d'un certain rang : une suite $(u_n)_n$ est négligeable devant $(v_n)_n$ si, et seulement si, la suite quotient $(\frac{u_n}{v_n})_n$ tend vers 0.

0.9. Notation. Une suite $(u_n)_n$ est dominée par une suite $(v_n)_n$ si

$$\exists C > 0,\ \exists N \in \mathbb{N},\ \forall n \geqslant N, \qquad |u_n| \leqslant C |v_n|.$$

On note alors $u_n = O(v_n)$ (ou plus rarement $u_n = \Omega(v_n)$).

0.10. Remarque. Dans le cas d'une suite $(v_n)_n$ qui ne s'annule pas à partir d'un certain rang, cette définition équivaut au caractère borné de la suite quotient $(\frac{u_n}{v_n})_n$.

0.11. Notation. Des suites réelles $(u_n)_n$ et $(v_n)_n$ sont du même ordre de grandeur si $u_n = O(v_n)$ et $v_n = O(u_n)$.

On note alors $u_n = \Theta(v_n)$ ou, de manière équivalente, $v_n = \Theta(u_n)$ puisque la relation est symétrique.

0.12. Notation. Les suites réelles $(u_n)_n$ et $(v_n)_n$ sont équivalentes si $u_n - v_n = o(u_n)$ (ou $u_n - v_n = o(v_n)$ de manière équivalente). On note alors $u_n \sim v_n$.

0.13. Remarque. Deux suites équivalentes sont du même ordre de grandeur. La réciproque est fausse.

On utilisera fréquemment les deux équivalents classiques suivants.

0.14. Proposition. *Pour tout $k \in \mathbb{N}$, $(n)_k \sim n^k$.*

0.15. Proposition (formule de Stirling).

$$n! \sim \sqrt{2\pi n}\left(\tfrac{n}{e}\right)^n.$$

0.16. Remarque. Lorsque l'on n'a pas besoin de toute la précision de cet équivalent, on peut exploiter l'encadrement suivant plus élémentaire : pour tout entier non nul n,

$$\left(\tfrac{n}{e}\right)^n \leqslant n! \leqslant en\left(\tfrac{n}{e}\right)^n.$$

Chapitre I

Compléments de théorie des graphes

Ce chapitre explicite quelques éléments de théorie des graphes nécessaires pour comprendre l'ensemble de ce livre. On précise les principales définitions, les conventions et notations sans pour autant entrer dans des détails superflus pour la suite. On renvoie vers la bibliographie (et son commentaire) pour des références proposant un panorama plus complet.

1. Définitions

La notion de graphe utilisée dans cet ouvrage est la plus élémentaire possible (non orienté, simple, sans boucle, sans multi-arête).

1.1. Définition. Un graphe est un couple (S, A) avec S un ensemble et A une partie de $\mathscr{P}_2(S)$. Les éléments de S sont appelés sommets du graphe et ceux de A arêtes.

Dans la suite de ce document, l'ensemble de sommets S est un ensemble fini ou dénombrable (à partir de la quatrième partie).

1.2. Exemple. Le graphe divisoriel d'ordre 7 est le graphe $(\llbracket 7 \rrbracket, A)$ avec pour A l'ensemble suivant

$$\big\{\{1,2\},\{1,3\},\{1,4\},\{1,5\},\{1,6\},\{1,7\},\{2,4\},\{2,6\},\{3,6\}\big\}.$$

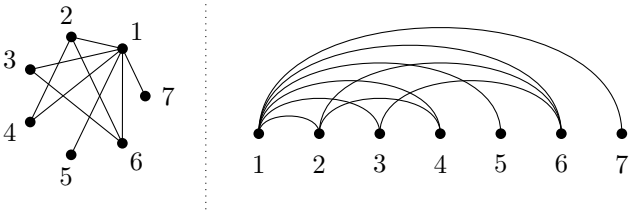
On peut décrire ce graphe en indiquant qu'une arête joint deux sommets si, et seulement si, l'un des deux divise l'autre.

1.3. Remarque. Pour simplifier la manipulation et la compréhension, on représente parfois un graphe par un dessin :

- ▷ les sommets sont des points distincts du plan ;
- ▷ les arêtes sont des courbes continues joignant les sommets.

Il n'y a évidemment pas unicité de la représentation d'un graphe et l'on se méfiera d'attribuer des propriétés à un graphe d'après l'un de ses dessins.

Voici par exemple deux dessins du graphe divisoriel d'ordre 7 de l'exemple précédent :



1.4. Définition. La réunion des graphes (S, A) et (S', A') est le graphe $(S \cup S', A \cup A')$.

1.5. Définition. Deux graphes $G = (S, A)$ et $G' = (S', A')$ sont isomorphes s'il existe une bijection $f : S \to S'$ telle que, pour tous sommets distincts x, $y \in S$, $\{x, y\} \in A$ si, et seulement si, $\{f(x), f(y)\} \in A'$. L'application f est alors appelée isomorphisme. En d'autres termes, un isomorphisme entre graphes est une bijection qui préserve les arêtes.

Un automorphisme d'un graphe $G = (S, A)$ est un isomorphisme de $G = (S, A)$ dans lui-même.

1.6. Remarque. Deux graphes sont isomorphes s'ils sont identiques aux noms des sommets près, c'est-à-dire s'ils partagent les mêmes dessins non étiquetés.

2. Graphes courants

2.1. Définition. Soit $n \geqslant 3$ un entier naturel. Un graphe est un chemin de longueur n s'il est isomorphe au graphe $P_n = \big(\llbracket n \rrbracket, A\big)$ où

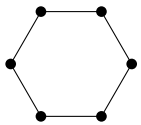
$$A = \big\{\{k, k+1\},\ k \in \llbracket n-1 \rrbracket\big\}.$$

Voici un dessin d'un chemin de longueur 6 :

2.2. Définition. Soit n un entier naturel non nul. Un graphe est un cycle de longueur n s'il est isomorphe au graphe $C_n = \big(\llbracket n \rrbracket, A\big)$ où

$$A = \big\{\{k, k+1\},\ k \in \llbracket n-1 \rrbracket\big\} \bigcup \big\{\{n, 1\}\big\}.$$

Voici un dessin d'un cycle de longueur 6 :



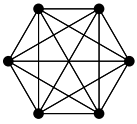
Un cycle de longueur 3 est par la suite (et pour des raisons évidentes) appelé triangle.

2.3. Définition. Soit n un entier naturel non nul. Un graphe est complet d'indice n s'il est isomorphe au graphe

$$K_n = \big(\llbracket n \rrbracket, \mathscr{P}_2(\llbracket n \rrbracket)\big).$$

Il s'agit donc d'un graphe à n sommets qui comporte toutes les arêtes possibles entre les sommets.

Voici un dessin d'un graphe complet d'indice 6 :

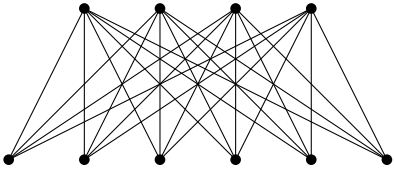


2.4. Définition. Soit n, m deux entiers naturels non nuls. Un graphe est biparti complet d'indices n et m s'il est isomorphe au graphe $K_{n,m}$ d'ensemble de sommets $\big([\![n]\!] \times \{0\}\big) \cup \big([\![m]\!] \times \{1\}\big)$ et d'ensemble d'arêtes

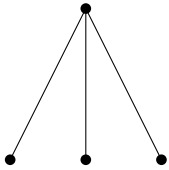
$$\big\{\{(k,0),(l,1)\},\ k \in [\![n]\!],\ l \in [\![m]\!]\big\}.$$

En d'autres termes, un tel graphe est composé de $n+m$ sommets et chacun des n premiers est relié à tous les m derniers (et il n'y a aucune autre arête).

Voici un dessin d'un graphe biparti complet d'indices 4 et 6 :



Le graphe $K_{1,3}$ est par la suite appelé griffe. En voici un dessin :



2.5. Remarque. Plus généralement, un graphe est biparti s'il existe une partie A de l'ensemble des sommets telle que chaque arête joigne un sommet appartenant à A et un sommet hors de A. Par exemple, un graphe-cycle de longueur paire est biparti (en prenant pour A un sommet sur deux lors du parcours du cycle).

3. Voisins, connexité

3.1. Définition. Soit $G = (S, A)$ un graphe. Deux sommets x et $y \in S$ sont voisins si $\{x, y\} \in A$, c'est-à-dire s'ils sont les extrémités d'une même arête.

Un chemin d'extrémités x et y (ou joignant x à y) est une suite de sommets $s_0, \ldots, s_N \in S$ (où N est un entier non nul) telle que
- ▷ $s_0 = x$, $s_N = y$,
- ▷ pour tout $k \in [\![N]\!]$, $\{s_{k-1}, s_k\} \in A$.

La longueur du chemin est l'entier N.

Un cycle dans un graphe est un chemin joignant un sommet à lui-même.

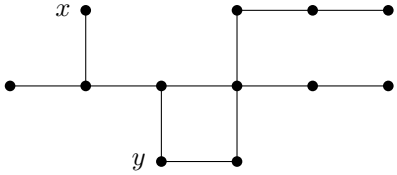
3.2. Définition. Soit $G = (S, A)$ un graphe. La composante connexe (ou classe de connexité) d'un sommet x est l'ensemble des sommets y tels qu'il existe un chemin d'extrémités x et y.

Le graphe G est connexe s'il n'existe qu'une seule composante connexe, c'est-à-dire que, pour toute paire de sommets distincts, il existe un chemin joignant ces deux sommets.

3.3. Définition. Soit $G = (S, A)$ un graphe connexe. La distance entre deux sommets x et $y \in S$ est la longueur minimale d'un chemin joignant x à y.

Le diamètre de G est la plus grande distance entre deux de ses sommets.

3.4. Exemple. Dans le graphe dessiné ci-dessous, les sommets x et y sont à distance 3 mais il existe aussi un chemin de longueur 5 les joignant. Le diamètre du graphe est 6 mais il existe un chemin de longueur 9 ne passant pas deux fois par la même arête.



Pour le graphe divisoriel d'ordre 7 déjà évoqué à l'exemple I-1.2, le diamètre est 2.

4. Degré d'un sommet

4.1. Définition. Soit $G = (S, A)$ un graphe. Le degré d'un sommet $x \in S$ est le nombre d'arêtes $a \in A$ dont x est extrémité, c'est-à-dire telles que $x \in a$. On le note $\deg(x)$. Autrement dit, le degré d'un sommet est le nombre de ses sommets voisins.

Un sommet isolé de G est un sommet $x \in S$ vérifiant $\deg(x) = 0$.

4.2. Exemple. Dans le graphe biparti complet $K_{n,m}$ pour des entiers naturels n, m non nuls et distincts, on compte n sommets de degré m et m sommets de degré n.

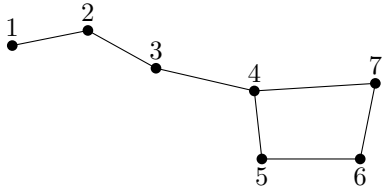
4.3. Proposition. *Un isomorphisme entre graphes envoie un sommet de degré d sur un sommet de degré d.*

Démonstration. Soit f un isomorphisme entre les graphes (S, A) et (S', A'). Fixons $x \in S$. Alors, par définition d'isomorphisme,

- ▷ pour tout $y \in S$ voisin de x, $f(y)$ est voisin de $f(x)$;
- ▷ pour tout $y' \in S'$ voisin de $f(x)$, $f^{-1}(y')$ est voisin de x.

Ainsi, x et $f(x)$ ont le même nombre de voisins, c'est-à-dire même degré. ♢

4.4. Exemple. Le graphe en « grande ourse » représenté par le dessin suivant admet deux automorphismes.



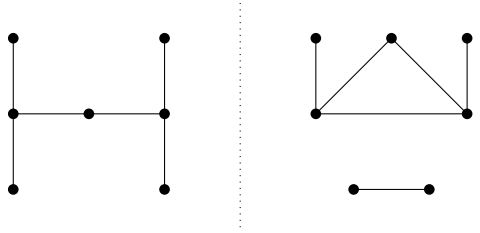
En effet, si σ est un automorphisme de ce graphe, alors $\sigma(4) = 4$ (car 4 est le seul sommet de degré 3) et $\sigma(1) = 1$ (car 1 est le seul

sommet de degré 1). L'unique voisin de 1 est envoyé sur l'unique voisin de $\sigma(1) = 1$ donc $\sigma(2) = 2$. Ensuite, les voisins de 2 sont envoyés sur les voisins de $\sigma(2) = 2$, à savoir 1 et 3 : comme σ est bijective, $\sigma(3) = 3$.

De même, les voisins de 4 sont envoyés sur les voisins de $\sigma(4) = 4$, à savoir 3, 5 et 7 : on obtient alors deux possibilités $\sigma(5) = 7$ et $\sigma(7) = 5$ ou l'inverse $\sigma(5) = 5$ et $\sigma(7) = 7$: dans les deux cas, on obtient ensuite $\sigma(6) = 6$.

En conclusion, les seuls automorphismes de ce graphe sont l'identité et la transposition (5 7).

4.5. Remarque. La suite des degrés ne suffit pas à caractériser un graphe à isomorphisme près. Par exemple, les deux graphes suivants à sept sommets ont chacun un sommet de degré 2, deux sommets de degré 3 et quatre sommets de degré 1, mais ne sont pas isomorphes : le premier est connexe, pas le second ; or, un isomorphisme préserve les chemins donc la connexité.



5. Graphes induits, sous-graphes

5.1. Définition. Soit $G = (S, A)$ un graphe et S' une partie de S. Le graphe induit par S' sur G est le graphe, noté $G[S']$, défini par

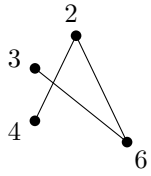
$$\big(S', A \cap \mathscr{P}_2(S')\big).$$

En d'autres termes, on obtient le graphe induit par S' sur G en ne conservant de G que les sommets dans S' et toutes les arêtes dont les deux extrémités appartiennent à S'.

5.2. Exemple. Le graphe induit par $\{2,3,4,6\}$ sur le graphe divisoriel d'ordre 7 comporte donc trois arêtes :

$$\big\{\{2,4\},\{2,6\},\{3,6\}\big\}.$$

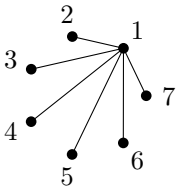
On remarque qu'il s'agit d'un graphe chemin de longueur 3 (en passant successivement du sommet 3 au sommet 6 puis au sommet 2 et enfin au sommet 4). Voici son dessin :



5.3. Définition. Soit $G = (S, A)$ un graphe. Un sous-graphe de G est un graphe $G' = (S', A')$ avec $S' \subset S$ et $A' \subset A \cap \mathscr{P}_2(S')$.

En d'autres termes, on obtient un sous-graphe de G en ne conservant de G que les sommets dans une partie S' et certaines des arêtes dont les deux extrémités sont dans S'.

5.4. Remarque. Un graphe induit est un sous-graphe ; la réciproque est fausse. Voici par exemple un sous-graphe du graphe divisoriel d'ordre 7 qui n'est pas un graphe induit (par exemple, on a conservé les sommets 2 et 4 mais pas l'arête les reliant) :



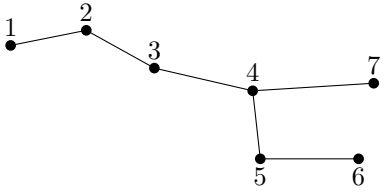
6. Arbres

6.1. Définition. Un arbre est un graphe à la fois connexe et sans cycle (c'est-à-dire tel qu'aucun sous-graphe n'est un cycle).

6.2. Exemple. Soit m un entier naturel non nul. Le graphe biparti complet $K_{1,m}$ est un arbre.

6.3. Définition. Un arbre couvrant d'un graphe $G = (S, A)$ est un sous-graphe de G qui est un arbre et dont l'ensemble des sommets est S.

6.4. Exemple. Le graphe en « grande ourse » déjà considéré admet quatre arbres couvrants. En voici un (chacun se déduit du graphe initial en enlevant l'une des arêtes du cycle de sommets 4, 5, 6 et 7).



6.5. Proposition. *Tout graphe connexe fini admet un arbre couvrant.*

Démonstration. Si le graphe est un arbre, il est son propre arbre couvrant. Sinon, le graphe admet un cycle et en supprimant une arête de ce cycle, on obtient un graphe connexe avec les mêmes sommets et une arête de moins. En recommençant, on obtient un graphe sans cycle après un nombre fini d'étapes et connexe (car la propriété est préservée à chaque étape) qui admet le même ensemble de sommets que le graphe de départ : c'est donc un arbre couvrant du graphe initial. ♢

7. Autres invariants

On a vu que la liste des degrés des sommets ou le diamètre sont des quantités invariantes quand on passe d'un graphe à un graphe isomorphe. D'autres invariants sont parfois utilisés. En voici quelques uns qui apparaîtront dans les exercices de ce livre.

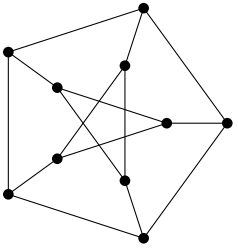
7.1. Définition. Soit $G = (S, A)$ un graphe.

▷ Un ensemble indépendant (ou stable) est une partie I de S telle que $\mathscr{P}_2(I) \cap A = \varnothing$. En d'autres termes, un ensemble de sommets est indépendant s'il n'y a aucune arête joignant deux de ces sommets. Le plus grand cardinal d'un ensemble indépendant de G est notée $\alpha(G)$.

▷ Un ensemble dominant est une partie D de S telle que, pour tout $s \in S$, il existe $x \in D$ tel que $\{x, s\} \in A$. En d'autres termes, un ensemble de sommets est dominant si tout sommet appartient à cet ensemble ou est voisin d'un sommet de cet ensemble. Le plus petit cardinal d'un ensemble dominant est noté $\gamma(G)$.

▷ La maille de G est quand elle existe la plus petite longueur d'un cycle de G de longueur au moins 3. On la note $g(G)$.

7.2. Exemple. Le dessin suivant représente le graphe de Petersen [1] à 10 sommets et 15 arêtes :



Pour ce graphe, on vérifie que les invariants ci-dessus prennent les valeurs $\alpha = 4$, $\gamma = 3$ et $g = 5$.

7.3. Remarque. De manière générale, on peut vérifier que les ensembles indépendants maximaux pour l'inclusion sont les ensembles dominants minimaux pour l'inclusion.

1. Ce graphe cubique (car tous les sommets sont de degré 3) sert très souvent de contre-exemple pour les conjectures naïves sur les graphes. Il porte le nom du danois Julius Petersen mais avait déjà été étudié avant ce mathématicien.

7.4. Définition. Soit $G = (S, A)$ un graphe.

▷ Soit c un entier non nul. Un coloriage de G avec c couleurs est une application $\varphi : S \to [\![c]\!]$ telle que, pour toute arête $\{x, y\} \in A$, $\varphi(x) \neq \varphi(y)$.

▷ Le nombre chromatique de G est le plus petit entier c tel qu'il existe un coloriage de G avec c couleurs. On le note $\chi(G)$.

7.5. Exemple. ▷ Le nombre chromatique du graphe complet à n sommets est n.

▷ Le nombre chromatique d'un cycle est 2 si son nombre de sommets est pair, 3 sinon.

▷ Le nombre chromatique du graphe de Petersen ci-dessus est 3.

7.6. Proposition. *Soit $G = (S, A)$ un graphe. Alors,*

$$\chi(G)\alpha(G) \geqslant |S|.$$

Démonstration. Considérons un coloriage φ de G avec $\chi(G)$ couleurs et remarquons que, pour tout $k \in [\![\chi(G)]\!]$, $\varphi^{-1}(\{k\})$ est un ensemble indépendant donc de cardinal inférieur à $\alpha(G)$. En partitionnant S selon les images par φ, on obtient

$$|S| = \sum_{k=1}^{\chi(G)} |\varphi^{-1}(\{k\})| \leqslant \sum_{k=1}^{\chi(G)} \alpha(G) = \chi(G)\alpha(G).$$

♢

8. Exercices

8.1. Exercice. *Soit $n \geqslant m$ des entiers naturels non nuls et G un graphe biparti complet d'indices m et n. Déterminer les valeurs de a et b pour lesquelles il existe un graphe induit de G biparti complet d'indices a et b.*

Solution. Soit A et B des parties de cardinal à m et n éléments de l'ensemble des sommets de G correspondant à la propriété de graphe biparti, U et V les parties d'un graphe induit H de G biparti complet

d'indices a et b. D'après la propriété de graphe biparti, on a $A \subset U$ et $B \subset V$ ou l'inverse ; par conséquent, $\max(a,b) \leqslant n$ et $\min(a,b) \leqslant m$.

La réciproque est immédiate : si $a \geqslant b$, il suffit de prendre a sommets parmi les m et b sommets parmi les n pour obtenir un graphe induit de G isomorphe à $K_{a,b}$.

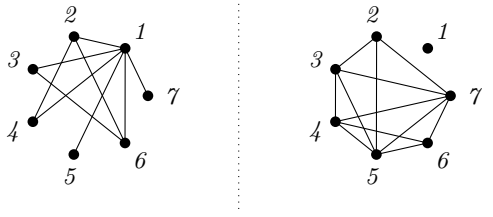
8.2. Exercice. *Soit $n \geqslant m$ des entiers naturels non nuls et G un graphe biparti complet d'indices m et n. Déterminer les valeurs de a pour lesquelles il existe un sous-graphe de G qui est un cycle de longueur a.*

Solution. Soit A et B des parties de cardinal à m et n éléments de l'ensemble des sommets de G correspondant à la propriété de graphe biparti. Un cycle de G comporte alternativement un élément de A et un élément de B. Ainsi, G admet un cycle induit de longueur a si, et seulement si, a est pair et si l'on peut trouver $\frac{a}{2}$ éléments de A et $\frac{a}{2}$ éléments de B.

En conclusion, il existe un cycle de longueur a dans G pour tout a pair inférieur ou égal à $2m$.

8.3. Exercice. *Le complémentaire d'un graphe $G = (S, A)$ est le graphe de même ensemble de sommets $\overline{G} = \big(S, \mathscr{P}_2(S) \backslash A\big)$. En d'autres termes, deux sommets sont voisins dans $\overline{G}$ si, et seulement si, il ne sont pas voisins dans G.*

Par exemple, voici le graphe divisoriel d'ordre 7 et son complémentaire



Montrer que tout automorphisme d'un graphe G est aussi un automorphisme de son complémentaire $\overline{G}$.

Solution. Soit f un automorphisme du graphe $G = (S, A)$.
Remarquons tout d'abord que, par définition, f réalise une bijection

de S dans S et, pour tous $x, y \in S$, $\{x, y\} \in A$ si, et seulement si, $\{f(x), f(y)\} \in A$. Ce deuxième point se récrit : pour tous $x, y \in S$, $\{x, y\} \notin A$ si, et seulement si, $\{f(x), f(y)\} \notin A$: en conclusion, f est un automorphisme de $\overline{G}$.

8.4. Exercice. *Soit n, m des entiers naturels non nuls distincts. Déterminer les automorphismes du graphe chemin P_n, du graphe cycle C_n, du graphe complet K_n et du graphe complet biparti $K_{n,m}$.*

Solution. ▷ Soit σ un automorphisme du graphe chemin P_n. Comme 1 et n sont les seuls sommets de degré 1, $\sigma(1) \in \{1, n\}$. Une fois cette valeur fixée, $\sigma(2)$ est l'unique voisin de $\sigma(1)$ puis $\sigma(3)$ le voisin de $\sigma(2)$ non encore atteint par σ et ainsi de suite. Il y a donc au plus deux automorphismes de P_n :

— l'identité,

— l'application $k \mapsto n + 1 - k$ (qui correspond à « retourner » le chemin).

On vérifie immédiatement que ces deux applications sont bien des automorphismes de P_n.

▷ Soit σ un automorphisme du graphe cycle C_n. Notons $k = \sigma(1)$. Alors, $\sigma(2)$ est un voisin de k, donc appartient à $\{k - 1, k + 1\}$ (avec des calculs modulo n). Une fois $\sigma(2)$ fixé, la valeur de $\sigma(3)$ est le voisin de $\sigma(2)$ différent de k. Par suite, σ est entièrement déterminée par le choix de $\sigma(1)$ et $\sigma(2)$: on vérifie que cette construction donne bien $2n$ automorphismes de C_n.

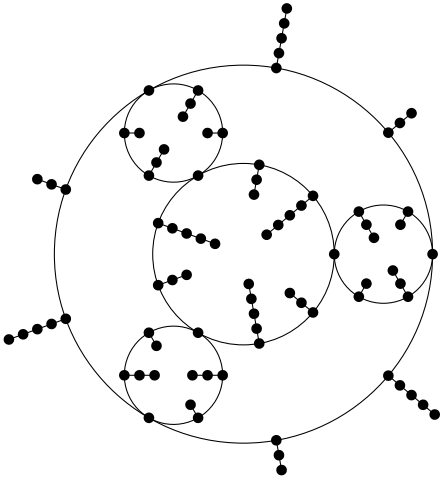
▷ On vérifie immédiatement que chaque permutation de $[\![n]\!]$ définit un automorphisme du graphe complet K_n : il y en a donc $n!$.

▷ Soit σ un automorphisme du graphe biparti complet $K_{n,m}$. Notons A l'ensemble des n premiers sommets et B l'ensemble des m suivants.

Pour des raisons de degré avec la condition $n \neq m$, $\sigma(A) = A$ et $\sigma(B) = B$. Réciproquement, on vérifie que toute permutation obtenue à partir d'une permutation de A et d'une permutation de B est un automorphisme de $K_{n,m}$. Il y a donc $n! \times m!$ automorphismes.

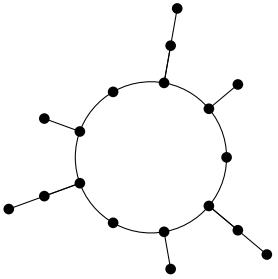
On vérifie que dans les différents cas traités dans cet exercice, les groupes des automorphismes sont respectivement isomorphes au groupe cyclique à 2 éléments, au groupe diédral à $2n$ éléments, au groupe symétrique S_n et au groupe produit $S_n \times S_m$.
Un célèbre théorème dû au mathématicien germano-chilien Roberto

Frucht indique que tout groupe fini est le groupe des automorphismes d'un graphe. Toutefois les preuves de ce résultat fournissent des exemples de taille peu raisonnable : voici par exemple le graphe associé au groupe symétrique S_3 par la preuve la plus élémentaire du théorème de Frucht.



L'exercice suivant donne un exemple simple issu de la preuve de Frucht.

8.5. Exercice. *Déterminer le groupe des automorphismes d'un graphe G correspondant au dessin suivant :*



Solution. Le graphe G n'admet que trois sommets x_1, x_2, x_3 de degré 2 dont les deux voisins sont de degré 3. Comme un automorphisme préserve le degré, tout automorphisme de G induit donc une permutation de l'ensemble $\{x_1, x_2, x_3\}$. Une fois les images de ces trois sommets

fixées, tous les autres sommets ont leur image déterminée uniquement : en effet, parmi leurs deux voisins, un seul a un voisin de degré 1. le voisin de x_1 ayant aussi un voisin de degré 1 est envoyé sur l'unique voisin de l'image de x_1 ayant aussi un voisin de degré 1. En poursuivant ainsi, on obtient que tout automorphisme de G se déduit de manière unique d'une permutation de l'ensemble $\{x_1, x_2, x_3\}$ (sur le dessin ci-dessus, il s'agit simplement de l'identité et des rotations d'angle $\frac{2\pi}{3}$ et $\frac{4\pi}{3}$).
En conclusion, le groupe d'automorphismes de G est cyclique d'ordre 3.

8.6. Exercice. *Soit G un graphe d'ensemble de sommets $[\![n]\!]$ et d'ensemble d'arêtes A. Considérons, pour toute permutation $\sigma \in S_n$, le graphe image $\sigma(G)$ de même ensemble de sommets et dont l'ensemble des arêtes est*

$$\big\{\{\sigma(i), \sigma(j)\},\ \{i,j\} \in A\big\}.$$

1. *Montrer que, pour σ, $\rho \in S_n$, $\sigma(G) = \rho(G)$ si, et seulement si, la permutation $\rho^{-1} \circ \sigma$ est un automorphisme de G.*
2. *En déduire que le nombre de sous-graphes de K_n isomorphes à G est $n!$ divisé par le nombre d'automorphismes de G.*

Solution.

1. Soit σ, $\rho \in S_n$. Remarquons que les graphes $\sigma(G)$ et $\rho(G)$ coïncident si, et seulement si, ils ont le même ensemble d'arêtes. Si $\sigma(G) = \rho(G)$, alors pour toute arête $\{i,j\} \in A$, il existe une arête $\{i',j'\} \in A$ telle que $\{\sigma(i), \sigma(j)\} = \{\rho(i'), \rho(j')\}$, et donc

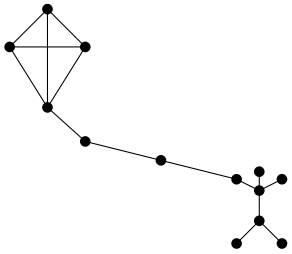
   $$\{\rho^{-1}(\sigma(i)), \rho^{-1}(\sigma(j))\} = \{i',j'\} \in A.$$

   Ainsi, $\rho^{-1} \circ \sigma$ est un automorphisme de G.

   La réciproque est immédiate.
2. Considérons l'application $\sigma \mapsto \sigma(G)$ de S_n dans l'ensemble des sous-graphes de K_n isomorphes à G (ce qui assure la surjectivité). D'après la question précédente, chaque élément atteint admet autant d'antécédents qu'il y a d'automorphismes de G, c'est-à-dire $|\mathrm{Aut}(G)|$. Ainsi, le cardinal de l'ensemble d'arrivée est

   $$\tfrac{n!}{|\mathrm{Aut}(G)|}.$$

Le graphe suivant à 13 sommets admet 24 automorphismes[2] *donc est présent* $\frac{13!}{24} = 259459200$ *fois dans le graphe complet* K_{13}.



8.7. Exercice. *Soit* $G = (S, A)$ *un graphe et* k *un entier naturel. Montrer que*

$$\sum_{\{x,y\}\in A} \big(\deg(x)^k + \deg(y)^k\big) = \sum_{s\in S} \deg(s)^{k+1}.$$

Solution. Comme $\{x, y\} = \{y, x\}$, on compte deux fois chaque arête en sommant selon les sommets x et y.

Ainsi,

$$\begin{aligned}
&\sum_{\{x,y\}\in A} \big(\deg(x)^k + \deg(y)^k\big) \\
&\qquad = \tfrac{1}{2} \sum_{x\in S}\sum_{y\in S} \big(\deg(x)^k + \deg(y)^k\big)\mathbb{1}_{\{x,y\}\in A} \\
&\qquad = \tfrac{1}{2} \sum_{x\in S} \sum_{\substack{y\in S\\ \{x,y\}\in A}} \deg(x)^k + \tfrac{1}{2} \sum_{y\in S} \sum_{\substack{x\in S\\ \{x,y\}\in A}} \deg(y)^k \\
&\qquad = 2\tfrac{1}{2} \sum_{x\in S} \deg(x)\deg(x)^k \\
&\qquad = \sum_{s\in S} \deg(s)^{k+1}.
\end{aligned}$$

8.8. Exercice. *Soit* k *un entier supérieur à* 2 *et* n *un entier supérieur à* $3k - 1$. *Considérons* G *un graphe obtenu comme réunion de* k *sous-graphes distincts isomorphes à* K_3 *(c'est-à-dire des triangles) tels qu'au*

2. Six choix pour la permutation des sommets haut du « cerf-volant », deux pour la permutation « bras libre-tête » et deux pour les « jambes », soit 6×2×2 = 24.

moins deux d'entre eux partagent au moins un sommet commun. Montrer que le nombre d'arêtes de G est strictement supérieur au nombre de sommets de G.

Solution. Procédons par récurrence sur k.

▷ Pour $k = 2$, les deux triangles composant G ont soit un, soit deux sommets communs. Dans le premier cas, il y a six arêtes pour cinq sommets ; dans le second, cinq arêtes pour quatre sommets.

▷ Soit $k \geqslant 2$ tel que le résultat soit obtenu pour toute réunion de k triangles tels qu'au moins deux d'entre eux partagent au moins un sommet commun. Considérons G la réunion de $k+1$ triangles tels qu'au moins deux d'entre eux partagent au moins un sommet commun ; notons F la réunion de k de ces triangles tels qu'au moins deux d'entre eux partagent au moins un sommet commun. D'après l'hypothèse de récurrence, le nombre a d'arêtes de F et le nombre s de sommets vérifient $a > s$.

Discutons selon le nombre p de sommets communs entre F et le dernier triangle :

- ▷ si $p = 0$, alors la réunion admet $s + 3$ sommets et $a + 3$ arêtes (aucune des arêtes du triangle ne peut appartenir à F, car elles n'ont pas d'extrémités dans F) : $a + 3 > s + 3$,
- ▷ si $p = 1$, alors la réunion admet $s + 2$ sommets et $a + 3$ arêtes (aucune des arêtes du triangle ne peut appartenir à F, car elles n'ont au plus qu'une extrémité dans F) : $a + 3 > s + 2$,
- ▷ si $p = 2$, alors la réunion admet $s + 1$ sommets et au moins $a + 2$ arêtes (on ajoute, au moins, les arêtes dont une extrémité est le sommet qui n'appartient pas à F) : $a + 2 > s + 1$,
- ▷ si $p = 3$, alors la réunion admet s sommets et au moins a arêtes : $a > s$.

Dans tous les cas, le nombre d'arêtes de G est strictement plus grand que le nombre de ses sommets.

8.9. Exercice. *Soit $G = (S, A)$ un graphe fini à n sommets et m arêtes.*

1. *Montrer que, si G est connexe, alors $m \geqslant n - 1$.*
2. *Montrer que, si G est sans cycle, alors G admet un sommet de degré au plus 1, puis $m \leqslant n - 1$.*
3. *Montrer que G est un arbre si, et seulement si, il est connexe et $m = n - 1$.*

Solution.

1. Procédons par récurrence sur le nombre d'arêtes m.

   ▷ L'inégalité est vérifiée si $m = 0$, car le graphe ne peut-être connexe que si $n \leqslant 1$.

   ▷ Soit m un entier naturel non nul. Supposons le résultat établi pour tous les graphes connexes à au plus m arêtes et considérons $G = (S, A)$ un graphe connexe à $m + 1$ arêtes. Soit enfin a une arête et $G' = (S, A \setminus \{a\})$ le graphe déduit de G en enlevant cette arête.
   - ▷ Si G' est connexe, alors $m + 1 = |A| \geqslant |A \setminus \{a\}| \geqslant n - 1$, d'après l'hypothèse de récurrence pour G'.
   - ▷ Sinon, G' admet deux composantes connexes (S_1, A_1) et (S_2, A_2) (où S_1 et S_2 forment une partition de S). Alors, l'hypothèse de récurrence appliquée à chaque composante donne

   $$\begin{aligned} m + 1 = |A| &= |A_1| + |A_2| + 1 \\ &\geqslant |S_1| - 1 + |S_2| - 1 + 1 \\ &= |S| - 1. \end{aligned}$$

2. ▷ Pour le premier point, raisonnons par l'absurde et supposons que tout sommet est de degré au moins 2. Construisons alors une suite infinie de sommets distincts par récurrence.

   ◦ Fixons un sommet $x_0 \in S$ de degré au moins 2.

   ◦ Soit $n \in \mathbb{N}$ tel qu'il existe une suite $x_0, \ldots, x_n$ de sommets distincts vérifiant

   $$\forall k \in [\![n]\!], \qquad \{x_{k-1}, x_k\} \in A.$$

   Or, le sommet x_n ne peut être de degré 1 par hypothèse, donc il existe $x_{n+1} \in S$ tel que $\{x_n, x_{n+1}\} \in A$ et $x_{n+1} \neq x_{n-1}$. Mais, comme G est sans cycle, le sommet x_{n+1} est aussi distinct des sommets déjà considérés : $x_0, \ldots, x_{n-2}$.

On dispose ainsi d'une suite infinie de sommets distincts ce qui contredit l'hypothèse selon laquelle le graphe G (donc son ensemble de sommets) est fini. En conclusion, il existe au moins un sommet de degré au plus 1.

▷ Montrons le second point par récurrence sur le nombre de sommets n.

∘ L'inégalité est immédiate pour $n = 1$ puisqu'il faut au moins deux sommets pour avoir une arête.

∘ Soit n un entier naturel non nul. Supposons le résultat établi pour tous les graphes sans cycle à au plus n sommets et considérons $G = (S, A)$ un graphe sans cycle à $n + 1$ sommets. D'après le point précédent, il existe un sommet x de degré au plus 1. Le graphe induit $G' = G[S \setminus \{x\}]$ est sans cycle, donc son nombre d'arêtes est inférieur ou égal à $|S \setminus \{x\}| - 1 = n - 1$, par hypothèse de récurrence. Comme le sommet x est de degré au plus 1, il existe au plus une arête de G non présente dans G' : ainsi,

$$\begin{aligned} m &\leqslant (n-1) + 1 \\ &= (n+1) - 1. \end{aligned}$$

3. ($\Rightarrow$) Le sens direct est une conséquence immédiate des deux premières questions.

   ($\Leftarrow$) Supposons G connexe et $m = n - 1$. Raisonnons par l'absurde : si G admet un cycle, alors le graphe déduit de G en enlevant une arête de ce cycle demeure connexe mais comporte n sommets et $m - 1$ arêtes.

   D'après la première question, on a $m-1 \geqslant n-1$, ce qui contredit l'égalité supposée.

   Ainsi, G est connexe sans cycle, donc un arbre.

On pourrait montrer de même que G est un arbre si, et seulement si, il est sans cycle et admet exactement $n - 1$ arêtes. Autrement dit, un arbre est un graphe connexe de taille minimale ou encore un graphe sans cycle de taille maximale (dans les deux cas, le nombre de sommets dépasse d'une unité le nombre d'arêtes).

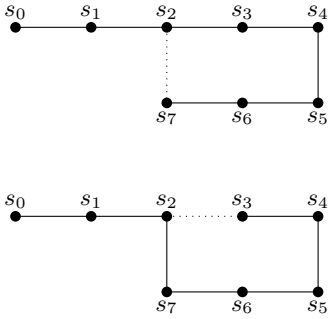
8.10. Exercice (Lemme de Posá, [Pós76]).

Soit $G = (S, A)$ un graphe. Étant donné un chemin $(s_0, s_1 \ldots, s_N)$ de sommets distincts de G et une arête $\{s_k, s_N\}$ partant de la seconde extrémité du chemin, on définit le chemin déduit par la « transformation autorisée » associée à l'arête $\{s_k, s_N\}$ comme le chemin de G suivant :

$$(s_0, \ldots, s_k, s_N, s_{N-1}, \ldots, s_{k+1}).$$

Intuitivement, ce nouveau chemin suit le chemin initial de s_0 à s_k, puis file directement à l'extrémité s_N et enfin parcourt la fin du chemin initial « à rebours ».

Voici une illustration du chemin $(s_0, s_1, s_2, s_3, s_4, s_4, s_6, s_7)$ et du chemin qui en est déduit par la transformation autorisée correpondant à l'arête $\{s_2, s_7\}$.



Fixons un sommet s_0 et un chemin $\mathscr{C} = (s_0, s_1, \ldots, s_N)$ (à sommets deux à deux distincts) de G de longueur maximale. Considérons les ensembles suivants

- ▷ *l'ensemble E des secondes extrémités (celle distincte de s_0) des chemins déduits de $\mathscr{C}$ par transformations autorisées successives ;*
- ▷ *l'ensemble V des voisins des sommets de E dans le chemin initial $\mathscr{C}$ (si un sommet s_k appartient à E, les sommets s_{k-1} et s_{k+1} appartiennent à V) ;*
- ▷ *l'ensemble R des sommets restants de $\mathscr{C}$, c'est-à-dire qui ne sont ni dans E ni dans V.*

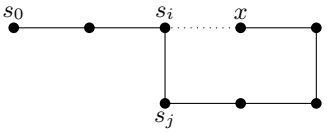
1. *Justifier qu'un sommet de E ne peut être relié dans G à un sommet hors de $\mathscr{C}$.*
2. *Montrer qu'il n'existe pas d'arête de G reliant un sommet de R à un sommet de E.*

3. *En déduire que l'ensemble des sommets de E et de leurs voisins dans G est de cardinal inférieur ou égal à $3|E| - 1$.*

Solution.

1. Soit $s \in E$. Par définition, il existe un chemin $\mathscr{C}'$ (obtenu par transformations autorisées successives depuis $\mathscr{C}$) issu de s_0, de même longueur que $\mathscr{C}$ et de seconde extrémité s. Si s est relié dans G à un sommet hors de $\mathscr{C}$, donc hors de $\mathscr{C}'$ (car les deux chemins ont même ensemble de sommets), il serait possible de prolonger le chemin $\mathscr{C}'$, ce qui contredit la maximalité du chemin initial $\mathscr{C}$.
2. Supposons qu'il existe des sommets $s_i \in R$ et $s_j \in E$ reliés par une arête de G. Comme $s_j \in E$, il existe un chemin
$$\mathscr{C}' = (s_0, \dots, s_i, x, \dots, s_j)$$
obtenu par transformations autorisées successives depuis $\mathscr{C}$.

   ▷ Si x est un sommet voisin de s_i dans le chemin initial $\mathscr{C}$ (c'est-à-dire s_{i-1} ou s_{i+1}), alors en effectuant la transformation autorisée associée à l'arête $\{s_i, s_j\}$ dans $\mathscr{C}'$, on obtient que x est une extrémité donc dans E ; par suite, s_i appartient à V, ce qui contredit l'hypothèse de s_i dans R.



   ▷ Si x n'est pas un sommet voisin de s_i dans le chemin initial $\mathscr{C}$, cela signifie que l'une des arêtes (au moins) de $\mathscr{C}$ issue de s_i a été effacée lors des transformations autorisées permettant de passer de $\mathscr{C}$ à $\mathscr{C}'$. Or, lorsqu'une arête est effacée, l'une des extrémités de cette arête devient la nouvelle extrémité du chemin : ainsi, soit s_i, soit l'un de ses voisins appartient à E, ce qui contredit à nouveau l'hypothèse de s_i dans R.
3. Remarquons que les voisins des sommets de E appartiennent à $\mathscr{C}$, d'après la première question et donc à V d'après la deuxième question. On y trouve par conséquent le sommet s_{N-1} (le seul voisin de s_N dans $\mathscr{C}$) et des sommets parmi $\{s_{k-1}, s_{k+1}\}$ pour

tout sommet $s_k \in E$ différent de s_N. Ainsi,

$$\begin{aligned} |E \cup V| &\leqslant |E| + 1 + \sum_{\substack{s \in E \\ s \neq s_N}} 2 \\ &= |E| + 1 + 2(|E| - 1) \\ &= 3|E| - 1. \end{aligned}$$

Cette question indique que la partie E associée au chemin $\mathscr{C}$ admet un « petit taux d'extension » (au sens où elle est reliée à « peu » d'autres sommets).

Chapitre II

Compléments de probabilités

Ce chapitre n'a pas vocation a remplacer un cours de probabilités : on suppose connues les notions d'espace probabilisé, d'événements, de mesure de probabilité (avec notamment les propriétés de σ-additivité ou de continuité monotone), de variables aléatoires, d'indépendance). L'objectif est triple :

- préciser les règles de manipulation des variables aléatoires obtenues par dénombrement d'événements aléatoires (et donc l'utilisation de variables indicatrices),
- rappeler les propriétés de l'espérance,
- démontrer les lemmes de Borel et Cantelli.

Les inégalités de concentration (de Markov ou de Bienaymé et Tchebytchev) joueront un rôle crucial dans ce livre mais ne seront énoncées que dans les chapitres V et VI.

1. Variables aléatoires indicatrices

1.1. Définition. L'indicatrice d'un événement A est la variable

aléatoire $\mathbb{1}_A$ à valeurs dans $\{0,1\}$ définie par

$$\mathbb{1}_A : \omega \mapsto \begin{cases} 1 & \text{si } \omega \in A, \\ 0 & \text{sinon.} \end{cases}$$

L'indicatrice d'un événement A suit la loi de Bernoulli de paramètre $\mathbf{P}(A)$.

On obtient immédiatement quelques résultats opératoires.

1.2. Proposition. *Soit A et B des événements. Alors,*

$$\mathbb{1}_A \mathbb{1}_B = \mathbb{1}_{A \cap B}, \qquad \mathbb{1}_{\overline{A}} = 1 - \mathbb{1}_A.$$

D'après les définitions d'indépendance pour les événements et pour les variables aléatoires, on déduit le résultat suivant.

1.3. Proposition. *Deux événements A et B sont indépendants si, et seulement si, les variables indicatrices $\mathbb{1}_A$ et $\mathbb{1}_B$ sont indépendantes.*

1.4. Remarque. En pratique, les variables indicatrices s'avèrent fort utiles pour rédiger une disjonction de cas : pour étudier une variable aléatoire X en discutant selon qu'un événement A est ou non réalisé, il suffit d'écrire la décomposition $X = X\mathbb{1}_A + X\mathbb{1}_{\overline{A}}$ puis d'étudier chacun des termes séparément.

2. Variables aléatoires de comptage

2.1. Définition. Soit $A_1, \ldots, A_n$ des événements. La variable de comptage de ces événements est la variable aléatoire N égale au nombre d'événements réalisés ; elle se décompose avec les variables indicatrices de la façon suivante

$$N = \sum_{k=1}^{n} \mathbb{1}_{A_k}.$$

2.2. Remarque. Cela est simplement la récriture du cardinal d'une partie F d'un ensemble fini E comme somme d'indicatrices (déterministes cette fois-ci) :

$$|F| = \sum_{x \in E} \mathbb{1}_{x \in F}.$$

Rappelons le résultat suivant typique du schéma de Bernoulli [1].

2.3. Proposition. *Soit $A_1, \ldots, A_n$ des événements mutuellement indépendants de même probabilité p. Alors, la variable de comptage de ces événements suit la loi binomiale de paramètres n et p.*

2.4. Exemple. Lançons n dés ordinaires discernables numérotés de 1 à n ; notons $X_1, \ldots, X_n$ les variables aléatoires correspondant à leurs résultats, puis N le nombre de paires de dés dont la somme est paire ; par exemple, avec quatre dés A, B, C et D qui donnent respectivement $1, 3, 4, 3$, il y a trois telles paires : $\{A, B\}$, $\{A, D\}$ et $\{B, D\}$. Alors, on obtient la décomposition suivante de N :

$$N = \sum_{i<j} \mathbb{1}_{X_i \equiv X_j[2]}.$$

Soit $i < j$; la probabilité de l'événement $X_i \equiv X_j[2]$ s'obtient par la formule des probabilités totales (c'est-à-dire en discutant selon les valeurs de X_j par exemple) :

$$\begin{aligned}
\mathbf{P}(X_i \equiv X_j[2]) &= \sum_{x=1}^{6} \mathbf{P}(X_i \equiv X_j[2] \,\big|\, X_j = x)\mathbf{P}(X_j = x) \\
&= \sum_{x=1}^{6} \mathbf{P}(X_i \equiv x[2])\mathbf{P}(X_j = x) \\
&= \sum_{x=1}^{6} \tfrac{1}{2}\tfrac{1}{6} = \tfrac{1}{2}.
\end{aligned}$$

1. Le « schéma de Bernoulli » est un modèle probabiliste où l'on considère une suite d'expériences aléatoires mutuellement indépendantes de même paramètre de réussite ; le résultat suivant correspond au nombre de réussites en n expériences.

On peut ensuite vérifier que les événements $X_i \equiv X_j[2]$ pour $i < j$ sont indépendants deux à deux : c'est évident si les ensembles d'indices des deux événements considérés sont disjoints.

Par ailleurs, pour tous i, j, k distincts,
$\mathbf{P}(X_i \equiv X_j[2], X_k \equiv X_j[2])$

$$= \sum_{x=1}^{6} \mathbf{P}(X_i \equiv X_j[2], X_k \equiv X_j[2] | X_j = x)\mathbf{P}(X_j = x)$$
$$= \sum_{x=1}^{6} \mathbf{P}(X_i \equiv x[2], X_k \equiv x[2])\mathbf{P}(X_j = x)$$
$$= \sum_{x=1}^{6} \mathbf{P}(X_i \equiv x[2])\mathbf{P}(X_k \equiv x[2])\mathbf{P}(X_j = x)$$
$$= \sum_{x=1}^{6} \tfrac{1}{2}\tfrac{1}{2}\tfrac{1}{6} = \tfrac{1}{4} = \mathbf{P}(X_i \equiv X_j[2])\mathbf{P}(X_k \equiv X_j[2]).$$

Enfin, constatons que ces événements ne sont pas mutuellement indépendants : pour tous i, j, k distincts,
$\mathbf{P}(X_i \equiv X_j[2], X_k \equiv X_j[2], X_i \equiv X_k[2])$

$$= \mathbf{P}(X_i \equiv X_j[2], X_k \equiv X_j[2]) = \tfrac{1}{4}$$
$$\neq \tfrac{1}{8} = \mathbf{P}(X_i \equiv X_j[2])\mathbf{P}(X_k \equiv X_j[2])\mathbf{P}(X_i \equiv X_k[2]).$$

2.5. Proposition. *Soit $A_1, \ldots, A_n$ des événements et N la variable de comptage de ces événements. Alors,*

$$N(N-1) = \sum_{k=1}^{n} \sum_{\substack{l=1 \\ l \neq k}}^{n} \mathbb{1}_{A_k} \mathbb{1}_{A_l}.$$

Démonstration. Si E est un ensemble de cardinal a, alors le nombre de couples d'éléments distincts de E est $a(a-1)$. Ainsi, la quantité $N(N-1)$ est le nombre de couples d'événements réalisés et d'indices distincts parmi $A_1, \ldots, A_n$. Alors,

$$N(N-1) = \sum_{\substack{(k,l) \\ k \neq l}} \mathbb{1}_{A_k \cap A_l} = \sum_{k=1}^{n} \sum_{\substack{l=1 \\ l \neq k}}^{n} \mathbb{1}_{A_k} \mathbb{1}_{A_l}.$$

♢

2.6. Remarque. En conservant les notations de cette proposition, ce résultat s'étend naturellement à toutes les factorielles descendantes de N.
Pour tout entier naturel non nul q, la variable $(N)_q$ est la somme sur tous les q-uplets $(B_1, \dots, B_q)$ d'événements deux à deux distincts parmi $A_1, \dots, A_n$ des indicatrices des intersections des composantes du q-uplet : $\mathbb{1}_{B_1 \cap \dots \cap B_q}$.

3. Espérance

3.1. Définition. L'espérance d'une variable aléatoire X discrète réelle d'ensemble de valeurs $\{x_i,\ i \in I\}$ est, quand elle existe, la somme

$$\sum_{i \in I} x_i \mathbf{P}(X = x_i).$$

On la note $\mathbf{E}(X)$.

3.2. Remarque. Pour une variable aléatoire discrète positive, la somme définissant l'espérance est soit un réel positif, soit diverge vers $+\infty$: cela permet de définir l'espérance comme un élément de $\mathbb{R}_+ \cup \{+\infty\}$ et d'étendre les propriétés suivantes avec les manipulations habituelles de $+\infty$.

3.3. Remarque. Soit X une variable aléatoire discrète réelle admettant une espérance finie. Alors, les probabilités $\mathbf{P}\big(X \leqslant \mathbf{E}(X)\big)$ et $\mathbf{P}\big(X \geqslant \mathbf{E}(X)\big)$ sont strictement positives.

Cette remarque est tout à fait intuitive : l'espérance se comprend comme une moyenne des valeurs prises par X (avec les probabilités comme poids), et cette variable prend donc des valeurs en-dessous et au-dessus (au sens large) de l'espérance.

3.4. Exemple. D'après le calcul pour la loi de Bernoulli, l'espérance de la variable indicatrice d'un événement A est le réel $\mathbf{P}(A)$.

L'espérance est linéaire ; précisons le cas particulier pour des variables positives en effectuant le calcul est dans $\mathbb{R}_+ \cup \{+\infty\}$ (d'où l'absence d'hypothèse de finitude des espérances).

3.5. Proposition. *Soit $X_1, \dots, X_n$ des variables aléatoires discrètes positives. Alors,*

$$\mathbf{E}\Big(\sum_{k=1}^{n} X_k\Big) = \sum_{k=1}^{n} \mathbf{E}(X_k).$$

3.6. Remarque. Ce résultat est particulièrement utile pour les sommes d'indicatrices. En effet, si une variable est de la forme

$$N = \sum_{k=1}^{n} \mathbb{1}_{A_k},$$

alors son espérance est, par linéarité,

$$\mathbf{E}(N) = \sum_{k=1}^{n} \mathbf{E}(\mathbb{1}_{A_k}) = \sum_{k=1}^{n} \mathbf{P}(A_k).$$

On utilisera également la propriété de croissance de l'espérance.

3.7. Proposition. *Soit X, Y des variables aléatoires discrètes admettant une espérance et telles que $\mathbf{P}(X \leqslant Y) = 1$. Alors, $\mathbf{E}(X) \leqslant \mathbf{E}(Y)$.*

L'indépendance des variables permet également de calculer l'espérance d'un produit.

3.8. Proposition. *Soit $X_1, \dots, X_n$ des variables aléatoires discrètes mutuellement indépendantes admettant une espérance. Alors, la variable $\prod_{k=1}^{n} X_k$ admet une espérance et*

$$\mathbf{E}\Big(\prod_{k=1}^{n} X_k\Big) = \prod_{k=1}^{n} \mathbf{E}(X_k).$$

4. Variance

4.1. Définition. ▷ La variance d'une variable aléatoire X discrète réelle est l'espérance de la variable positive $\big(X - \mathbf{E}(X)\big)^2$. On la note $\mathbf{V}(X)$.

▷ La covariance de deux variables aléatoires X et Y discrètes réelles est, quand elle existe, l'espérance de la variable

$$\big(X - \mathbf{E}(X)\big)\big(Y - \mathbf{E}(Y)\big).$$

On la note $\mathbf{Cov}(X, Y)$.

4.2. Remarque. La covariance de deux variables indépendantes X et Y est nulle ; en effet, comme les variables $X - \mathbf{E}(X)$ et $Y - \mathbf{E}(Y)$ sont indépendantes puis centrées,

$$\begin{aligned}\mathbf{Cov}(X, Y) &= \mathbf{E}\big(\big(X - \mathbf{E}(X)\big)\big(Y - \mathbf{E}(Y)\big)\big)\\ &= \mathbf{E}\big(X - \mathbf{E}(X)\big)\mathbf{E}\big(Y - \mathbf{E}(Y)\big) = 0.\end{aligned}$$

La variance ne jouit pas de la propriété de linéarité, mais on peut toutefois calculer la variance d'une somme par l'identité remarquable suivante.

4.3. Proposition. *Soit $X_1, \dots, X_n$ des variables aléatoires discrètes admettant une variance. Alors,*

$$\begin{aligned}\mathbf{V}\Big(\sum_{k=1}^{n} X_k\Big) &= \sum_{k=1}^{n}\sum_{l=1}^{n} \mathbf{Cov}(X_k, X_l)\\ &= \sum_{k=1}^{n} \mathbf{V}(X_k) + \sum_{k=1}^{n}\sum_{\substack{l=1\\ l\neq k}}^{n} \mathbf{Cov}(X_k, X_l).\end{aligned}$$

Démonstration. Notons, pour tout k, $\tilde{X}_k = X_k - \mathbf{E}(X_k)$. Alors,

$$\begin{aligned}\mathbf{V}\Big(\sum_{k=1}^{n} X_k\Big) &= \mathbf{E}\Big(\Big(\sum_{k=1}^{n} X_k - \sum_{k=1}^{n} \mathbf{E}(X_k)\Big)^2\Big) = \mathbf{E}\Big(\Big(\sum_{k=1}^{n} \tilde{X}_k\Big)^2\Big)\\ &= \mathbf{E}\Big(\sum_{k=1}^{n}\sum_{l=1}^{n} \tilde{X}_k\tilde{X}_l\Big) = \sum_{k=1}^{n}\sum_{l=1}^{n} \mathbf{E}(\tilde{X}_k\tilde{X}_l)\\ &= \sum_{k=1}^{n} \mathbf{E}(\tilde{X}_k^2) + \sum_{k=1}^{n}\sum_{\substack{l=1\\ l\neq k}}^{n} \mathbf{E}(\tilde{X}_k\tilde{X}_l).\end{aligned}$$

En se ramenant aux variables X_k, on constate que les deux dernières égalités correspondent exactement à la proposition. ♢

Cette formule se simplifie dans le cas de variables deux à deux indépendantes (l'indépendance mutuelle n'est pas requise, car il suffit à chaque fois d'annuler les covariances de deux variables).

4.4. Corollaire. *Soit $X_1, \dots, X_n$ des variables aléatoires discrètes deux à deux indépendantes admettant une variance. Alors,*

$$\mathbf{V}\Big(\sum_{k=1}^{n} X_k\Big) = \sum_{k=1}^{n} \mathbf{V}(X_k).$$

5. Lemmes de Borel et Cantelli

Dans cette section, nous abordons un outil puissant, qui sera utile pour certaines études sur les graphes aléatoires dénombrables à partir du chapitre IX.

5.1. Remarque. Soit $(A_n)_{n\geqslant 1}$ une suite d'événements et N la variable de comptage associée à valeurs dans $\mathbb{N} \cup \{+\infty\}$. D'après l'écriture suivante de N avec les indicatrices

$$N = \sum_n \mathbb{1}_{A_n},$$

la linéarité de l'espérance et le passage à la limite, on obtient

$$\mathbf{E}(N) = \sum_n \mathbf{P}(A_n),$$

avec la convention $\mathbf{E}(N) = +\infty$ si la série du membre de droite diverge. Il semble alors légitime de s'interroger sur l'éventuelle réalisation de l'événement « $N = +\infty$ » lorsque la série de terme général $\mathbf{P}(A_n)$ diverge. C'est l'objet du second lemme de Borel et Cantelli. Pour en obtenir un énoncé cohérent, introduisons la notion de limite supérieure d'une suite d'événements.

5.2. Définition. Soit $(A_n)_n$ une suite d'événements. La limite supérieure de cette suite est l'événement

$$\limsup A_n = \bigcap_p \Big(\bigcup_{n\geqslant p} A_n \Big).$$

Attention, contrairement aux apparences, $\limsup A_n$ ne dépend pas de l'entier n.

5.3. Remarque. D'après la définition, l'événement $\limsup A_n$ est réalisé lorsque, pour tout $p \in \mathbb{N}$, il existe $n \geqslant p$ tel que l'événement A_n est réalisé. Ainsi, l'événement $\limsup A_n$ est réalisé lorsqu'il existe une infinité d'indices k tels que l'événement A_k est réalisé.

Autrement dit, la limite supérieure d'une suite est réalisée lorsque les événements de la suite sont réalisés infiniment souvent.

5.4. Proposition (Premier lemme de Borel et Cantelli). *Soit $(A_n)_n$ une suite d'événements telle que la série de terme général $\mathbf{P}(A_n)$ converge. Alors, $\mathbf{P}\big(\limsup A_n\big) = 0$.*

Démonstration. Remarquons que, pour tout $p \in \mathbb{N}$,

$$\limsup A_n \subset \bigcup_{k \geqslant p} A_k.$$

Par conséquent,

$$\mathbf{P}(\limsup A_n) \leqslant \mathbf{P}\Big(\bigcup_{k \geqslant p} A_k\Big) \leqslant \sum_{k \geqslant p} \mathbf{P}(A_k).$$

Comme la série de terme général $\mathbf{P}(A_k)$ converge, le majorant tend vers 0 lorsque $p \to +\infty$. Par comparaison, $\mathbf{P}(\limsup A_n) = 0$. $\diamondsuit$

5.5. Exemple. Soit $(X_n)_n$ une suite de variables mutuellement indépendantes suivant la loi de Bernoulli de paramètre $\frac{1}{2}$. Notons, pour tout n, A_n l'événement

$$X_n = X_{n+1} = \cdots = X_{2n-1}.$$

Alors, pour tout n, $\mathbf{P}(A_n) = 2^{1-n}$ (en discutant selon la valeur commune des n variables intervenant dans A_n). Ainsi, la série de terme général $\mathbf{P}(A_n)$ converge : d'après le premier lemme de Borel et Cantelli, il n'existe presque sûrement qu'un nombre fini d'événements A_n réalisés.

Le résultat suivant est complémentaire du précédent : on remarquera toutefois l'hypothèse additionnelle d'indépendance.

5.6. Proposition (Second lemme de Borel et Cantelli). *Soit une suite d'événements mutuellement indépendants $(A_n)_n$ telle la série de terme général $\mathbf{P}(A_n)$ diverge. Alors, $\mathbf{P}(\limsup A_n) = 1$.*

Démonstration. Considérons le complémentaire de $\limsup A_n$, c'est-à-dire

$$\overline{\limsup A_n} = \bigcup_p \Big(\bigcap_{n \geqslant p} \overline{A_n} \Big).$$

Remarquons que, pour tous $p \leqslant q$,

$$\begin{aligned} \mathbf{P}\big(\bigcap_{n=p}^{q} \overline{A_n} \big) &= \prod_{n=p}^{q} \mathbf{P}\big(\overline{A_n}\big) \\ &= \prod_{n=p}^{q} \big(1 - \mathbf{P}(A_n)\big) \\ &\leqslant \prod_{n=p}^{q} \exp\big(- \mathbf{P}(A_n)\big) \qquad \text{car } e^{-x} \geqslant 1 - x \\ &= \exp\Big(- \sum_{n=p}^{q} \mathbf{P}(A_n)\Big). \end{aligned}$$

Ainsi, en faisant tendre $q \to +\infty$ avec la divergence de la série (vers $+\infty$) et la continuité décroissante de $\mathbf{P}$, on obtient

$$\mathbf{P}\Big(\bigcap_{n \geqslant p} \overline{A_n} \Big) = 0.$$

Ainsi, $\overline{\limsup A_n}$ est une réunion dénombrable d'événements de probabilité nulle, donc est de probabilité nulle. En repassant au complémentaire, on en déduit le résultat annoncé. ♢

6. Exercices

6.1. Exercice. *Soit X une variable aléatoire discrète et a un réel. Montrer que si la variable X admet une espérance inférieure ou égale à a, alors $\mathbf{P}(X \leqslant a) > 0$.*

Solution. Raisonnons par contraposition en supposant que la probabilité $\mathbf{P}(X \leqslant a)$ est nulle ; notons $(x_n)_n$ une énumération des valeurs prises par X.

Alors, pour tout n tel que $x_n \leqslant a$, l'événement $X = x_n$ est inclus dans l'événement $X \leqslant a$ (qui est par hypothèse de probabilité nulle), donc est de probabilité nulle. Par conséquent, la somme définissant l'espérance s'écrit uniquement avec des valeurs x_n strictement supérieures à a :

$$\begin{aligned}\mathbf{E}(X) &= \sum_{\substack{n\\ x_n > a}} x_n \mathbf{P}(X = x_n)\\ &> \sum_{\substack{n\\ x_n > a}} a\mathbf{P}(X = x_n)\\ &= a\mathbf{P}(X > a)\\ &= a\big(1 - \mathbf{P}(X \leqslant a)\big) = a.\end{aligned}$$

Ainsi, $\mathbf{E}(X) > a$, ce qui achève le raisonnement.

6.2. Exercice. *Soit X une variable aléatoire discrète et m, n deux entiers tels que $m \leqslant n$.*

1. *Montrer que $X^m \leqslant 1 + X^n$.*
2. *En déduire que si X^n admet une espérance (finie), alors X^m admet une espérance (finie).*

Solution.

1. Utilisons la décomposition de 1 comme somme de variables indicatrices $1 = \mathbb{1}_{X\leqslant 1} + \mathbb{1}_{X>1}$, puis majorons X^m (respectivement par 1 et X^n selon les cas) :

$$\begin{aligned}X^m &= X^m \mathbb{1}_{X\leqslant 1} + X^m \mathbb{1}_{X>1}\\ &\leqslant \mathbb{1}_{X\leqslant 1} + X^n \mathbb{1}_{X>1}.\end{aligned}$$

   Il ne reste plus qu'à majorer les indicatrices par 1 pour obtenir le résultat.
2. Par hypothèse la variable $1 + X^n$ admet une espérance. On conclut par domination que la variable X^m admet elle aussi une espérance.

Ce résultat, très classique, sert ici à illustrer l'utilisation des variables indicatrices pour mettre en place une discussion. On retiendra notamment que si une variable admet un moment d'ordre 2, alors elle admet une espérance (ce qui était implicite dans la définition de la variance).

6.3. Exercice. *Soit X_1, X_2 deux variables aléatoires à valeurs dans $\mathbb{N}$. La distance en variations totales entre les lois de X_1 et X_2 est la quantité*

$$d_{VT}(X_1, X_2) = \sup\{|\mathbf{P}(X_1 \in A) - \mathbf{P}(X_2 \in A)|,\ A \subset \mathbb{N}\}.$$

1. *Justifier que, pour tout $A \subset \mathbb{N}$,*

$$\mathbf{P}(X_1 \in A) - \mathbf{P}(X_2 \in A) = \mathbf{E}\big(\mathbb{1}_{X_1 \neq X_2}(\mathbb{1}_{X_1 \in A} - \mathbb{1}_{X_2 \in A})\big).$$

2. *En déduire que $d_{VT}(X_1, X_2) \leqslant 2\mathbf{P}(X_1 \neq X_2)$.*

Solution.

1. Soit $A \subset \mathbb{N}$. Introduisons les variables indicatrices des événements $X_1 \in A$ et $X_2 \in A$

$$\begin{aligned}\mathbf{P}(X_1 \in A) - \mathbf{P}(X_2 \in A) &= \mathbf{E}(\mathbb{1}_{X_1 \in A}\big) - \mathbf{E}\big(\mathbb{1}_{X_2 \in A}) \\ &= \mathbf{E}(\mathbb{1}_{X_1 \in A} - \mathbb{1}_{X_2 \in A}) \\ &= \mathbf{E}\big(\mathbb{1}_{X_1 \neq X_2}(\mathbb{1}_{X_1 \in A} - \mathbb{1}_{X_2 \in A})\big) \\ &\quad + \mathbf{E}\big(\mathbb{1}_{X_1 = X_2}(\mathbb{1}_{X_1 \in A} - \mathbb{1}_{X_2 \in A})\big) \\ &= \mathbf{E}\big(\mathbb{1}_{X_1 \neq X_2}(\mathbb{1}_{X_1 \in A} - \mathbb{1}_{X_2 \in A})\big)\end{aligned}$$

car la variable $\mathbb{1}_{X_1 = X_2}(\mathbb{1}_{X_1 \in A} - \mathbb{1}_{X_2 \in A})$ est nulle.

2. Partons du résultat de la question précédente. Pour toute partie $A \subset \mathbb{N}$, on obtient (en utilisant que la valeur absolue de l'espérance est inférieure à l'espérance de la valeur absolue) :

$$\begin{aligned}&|\mathbf{P}(X_1 \in A) - \mathbf{P}(X_2 \in A)| \\ &\qquad \leqslant \mathbf{E}\big(|\mathbb{1}_{X_1 \neq X_2}(\mathbb{1}_{X_1 \in A} - \mathbb{1}_{X_2 \in A})|\big) \\ &\qquad \leqslant \mathbf{E}\big(\mathbb{1}_{X_1 \neq X_2}|\mathbb{1}_{X_1 \in A} - \mathbb{1}_{X_2 \in A}|\big) \\ &\qquad \leqslant 2\mathbf{E}(\mathbb{1}_{X_1 \neq X_2}) = 2\mathbf{P}(X_1 \neq X_2).\end{aligned}$$

En passant à la borne supérieure sur A, on obtient la majoration attendue.

6.4. Exercice. *Soit $(A_i)_{i \in [\![n]\!]}$ une suite d'événements croissante pour l'inclusion telle que*

$$\forall i \in [\![n]\!], \qquad \mathbf{P}(A_i) = \tfrac{i}{n}.$$

Déterminer la loi de la variable de comptage N de ces événements.

Solution. Définissons les événements $B_1 = A_1$ et, pour les indices $i \geqslant 2$, $B_i = A_i \setminus A_{i-1}$. Par construction, pour tout indice i, A_i est la réunion disjointe des B_k pour $k \in [\![i]\!]$, donc

$$\mathbb{1}_{A_i} = \sum_{k=1}^{i} \mathbb{1}_{B_k}.$$

Par suite,

$$\begin{aligned} N &= \sum_{i=1}^{n} \mathbb{1}_{A_i} \\ &= \sum_{i=1}^{n} \sum_{k=1}^{i} \mathbb{1}_{B_k} \\ &= \sum_{k=1}^{n} \sum_{i=k}^{n} \mathbb{1}_{B_k} \\ &= \sum_{k=1}^{n} (n+1-k) \mathbb{1}_{B_k}. \end{aligned}$$

Remarquons maintenant que $\mathbf{P}(B_i) = \mathbf{P}(A_i) - \mathbf{P}(A_{i-1}) = \frac{1}{n}$ pour tout indice i, et que les événements $(B_k)_k$ sont deux à deux disjoints (donc que, pour chaque réalisation, au plus une des indicatrices vaut 1), de réunion A_n de probabilité 1. Ainsi, pour tout $j \in [\![n]\!]$, l'événement $N = j$ correspond à l'événement B_{n+1-j}, qui est de probabilité $\frac{1}{n}$.

En conclusion, la variable N suit la loi uniforme sur $[\![n]\!]$.

6.5. Exercice. *Soit X une variable aléatoire à valeurs dans $\mathbb{N}$ et c une constante strictement positive.*

1. *Justifier les inégalités*

$$c \sum_{n=0}^{+\infty} n \mathbb{1}_{X \in [cn, c(n+1)[} \leqslant X \leqslant c \sum_{n=0}^{+\infty} (n+1) \mathbb{1}_{X \in [cn, c(n+1)[}.$$

2. *Montrer que si l'espérance de X est finie, alors la série de terme général*

$$n\mathbf{P}\big(X \in [cn, c(n+1)[\big)$$

converge. En déduire que la série de terme général $\mathbf{P}(X \geqslant cn)$ converge.

3. *Supposons que la série de terme général $\mathbf{P}(X \geqslant cn)$ converge. Montrer que la variable X admet une espérance finie.*

Solution.

1. Comme les ensembles $\big([cn, c(n+1)[\big)_n$ forment une partition de $\mathbb{R}_+$ et que la variable X est à valeurs positives, on obtient immédiatement la décomposition de la constante 1 comme somme des indicatrices $\mathbb{1}_{X\in[cn,c(n+1)[}$:

$$\begin{aligned} X &= X\sum_{n=0}^{+\infty} \mathbb{1}_{X\in[cn,c(n+1)[} \\ &= \sum_{n=0}^{+\infty} X\mathbb{1}_{X\in[cn,c(n+1)[}. \end{aligned}$$

Pour conclure, il suffit d'exploiter l'encadrement qui définit l'indicatrice :

$$cn\mathbb{1}_{X\in[cn,c(n+1)[} \leqslant X\mathbb{1}_{X\in[cn,c(n+1)[} < c(n+1)\mathbb{1}_{X\in[cn,c(n+1)[}.$$

2. ▷ D'après la question précédente, pour tout N,

$$c\sum_{n=0}^{N} n\mathbb{1}_{X\in[cn,c(n+1)[} \leqslant X$$

soit, en passant à l'espérance qui est linéaire :

$$c\sum_{n=0}^{N} n\mathbf{P}\big(X\in[cn,c(n+1)[\big) \leqslant \mathbf{E}(X).$$

Ainsi, la série de terme général $n\mathbf{P}\big(X\in[cn,c(n+1)[\big)$ converge et sa somme est inférieure ou égale à $\frac{1}{c}\mathbf{E}(X)$.

▷ Notons, pour tout n, A_n l'événement « $X \geqslant cn$ ». Par construction, pour tout n,

$$\mathbf{P}\big(X\in[cn,c(n+1)[\big) = \mathbf{P}(A_n) - \mathbf{P}(A_{n+1}),$$

donc

$$\begin{aligned} \sum_{n=0}^{N} n\mathbf{P}\big(X\in[cn,c(n+1)[\big) &= \sum_{n=0}^{N} \big(n\mathbf{P}(A_n) - n\mathbf{P}(A_{n+1})\big) \\ &= \sum_{n=0}^{N} n\mathbf{P}(A_n) - \sum_{n=0}^{N} n\mathbf{P}(A_{n+1}) \\ &= \sum_{n=1}^{N} \mathbf{P}(A_n) - N\mathbf{P}(A_{N+1}). \end{aligned}$$

Remarquons enfin que $0 \leqslant N\mathbf{P}(A_{N+1}) \leqslant \mathbf{E}(X\mathbb{1}_{X\geqslant c(N+1)})$ et que ce dernier majorant tend vers 0 comme reste de la série convergente définissant l'espérance de X.

Par suite, la série de terme général $\mathbf{P}(A_n)$ converge.

3. En reprenant le calcul de la question précédente, on obtient pour tout N,

$$\sum_{n=0}^{N}(n+1)\mathbf{P}\big(X \in [cn, c(n+1)[\big) = \sum_{n=1}^{N}\mathbf{P}(A_n) - (N+1)\mathbf{P}(A_{N+1}) \leqslant \sum_{n=1}^{N}\mathbf{P}(A_n) \leqslant \sum_{n=1}^{+\infty}\mathbf{P}(A_n).$$

Ainsi, la série de terme général $(n+1)\mathbf{P}\big(X \in [cn, c(n+1)[\big)$ converge et, par linéarité de l'espérance,

$$\mathbf{E}(X) \leqslant \sum_{n=0}^{+\infty}(n+1)\mathbf{P}\big(X \in [cn, c(n+1)[\big).$$

On a ainsi établi l'équivalence entre l'existence de l'espérance de X et la convergence de la série de terme général $\mathbf{P}(X \geqslant cn)$. Dans le cas de convergence pour $c = 1$, les inégalités établies permettent également d'obtenir l'égalité

$$\mathbf{E}(X) = \sum_{n=1}^{+\infty}\mathbf{P}(X \geqslant n).$$

6.6. Exercice. *Soit $A_1, \ldots, A_n$ des événements et N la variable de comptage de ces événements. Exprimer $\mathbf{E}(N)$ et $\mathbf{E}(N(N-1))$ à l'aide des probabilités des différents événements.*

Solution. Rappelons que $N = \sum_{k=1}^{n}\mathbb{1}_{A_k}$ et donc, par linéarité de l'espérance,

$$\mathbf{E}(N) = \sum_{k=1}^{n}\mathbf{P}(A_k).$$

De même, on peut écrire

$$N(N-1) = \sum_{k=1}^{n} \sum_{\substack{l=1 \\ l \neq k}}^{n} \mathbb{1}_{A_k \cap A_l},$$

et donc, toujours par linéarité de l'espérance,

$$\mathbf{E}(N(N-1)) = \sum_{k=1}^{n} \sum_{\substack{l=1 \\ l \neq k}}^{n} \mathbf{P}(A_k \cap A_l).$$

On déduit de ce calcul la variance de N en utilisant la formule de König et Huygens :

$$\mathbf{V}(N) = \mathbf{E}\big(N(N-1)\big) + \mathbf{E}(N) - \mathbf{E}(N)^2.$$

On aurait pu directement calculer $\mathbf{E}(N^2)$ avec l'écriture de N comme somme d'indicatrices.
Le résultat de cet exercice est généralisé à la proposition VIII-2.1 avec le calcul des moments factoriels pour les variables aléatoires de comptage.

6.7. Exercice. *Soit σ une variable aléatoire de loi uniforme sur le groupe symétrique S_n et N la variable aléatoire égale au nombre de points fixes de σ. Calculer $\mathbf{E}(N)$ et $\mathbf{E}(N^2)$.*

Solution. ▷ Pour calculer l'espérance de N, commençons par écrire N comme somme d'indicatrices :

$$N = \sum_{i=1}^{n} \mathbf{1}_{\sigma(i)=i}.$$

Par linéarité de l'espérance,

$$\mathbf{E}(N) = \sum_{i=1}^{n} \mathbf{P}(\sigma(i) = i).$$

Or, pour tout $i \in [\![n]\!]$, il y a $(n-1)!$ permutations admettant i comme point fixe, donc $\mathbf{P}(\sigma(i) = i) = \frac{(n-1)!}{n!} = \frac{1}{n}$.
En reportant,

$$\mathbf{E}(N) = \sum_{i=1}^{n} \tfrac{1}{n} = 1.$$

▷ Utilisons à nouveau la linéarité de l'espérance pour calculer l'espérance de N^2,

$$\begin{aligned}\mathbf{E}(N^2) &= \mathbf{E}\Big(\sum_{i=1}^{n}\sum_{j=1}^{n}\mathbf{1}_{\sigma(i)=i}\mathbf{1}_{\sigma(j)=j}\Big)\\ &= \sum_{i=1}^{n}\mathbf{P}(\sigma(i)=i) + \sum_{i=1}^{n}\sum_{j=1,\ j\neq i}^{n}\mathbf{P}(\sigma(i)=i,\sigma(j)=j)\\ &= n\tfrac{(n-1)!}{n!} + n(n-1)\tfrac{(n-2)!}{n!} = 2.\end{aligned}$$

6.8. Exercice. *Soit σ une variable aléatoire de loi uniforme sur le groupe symétrique S_n. Considérons l'ensemble aléatoire des records de σ :*

$$R = \big\{k,\ \forall j \in [\![k]\!],\ \sigma(j) \leqslant \sigma(k)\big\}.$$

1. *Calculer, pour tout $k \in [\![n]\!]$, la probabilité $\mathbf{P}(k \in R)$.*
2. *En déduire $\mathbf{E}(|R|)$ le cardinal moyen de R.*

Solution. Remarquons tout d'abord que $k \in R$ si, et seulement si, pour tout $j < k$, $\sigma(j) < \sigma(k)$, car σ est bijective.

1. Fixons $k \in [\![n]\!]$. Dénombrons les permutations f pour lesquelles $f(j) < f(k)$ pour tout $j \in [\![k]\!]$. Pour cela, discutons selon la valeur l de $f(k)$. Une fois cette valeur fixée, f est entièrement déterminée par
   - ▷ la partie $A = f([\![k-1]\!])$ dans $\mathscr{P}_{k-1}([\![l-1]\!])$,
   - ▷ la bijection induite entre $[\![k-1]\!]$ et A,
   - ▷ la bijection induite entre $[\![n]\!] \setminus [\![k]\!]$ et $[\![n]\!] \setminus (A \cup \{l\})$.

   Ainsi, comme σ suit la loi uniforme sur S_n, on obtient

   $$\mathbf{P}(k \in R) = \tfrac{1}{n!}\sum_{l=1}^{n}\tbinom{l-1}{k-1}(k-1)!(n-k)!$$

   En utilisant alors la formule suivante sur les coefficients binomiaux (que l'on peut établir par simple récurrence sur n) :

   $$\forall a, n \in \mathbb{N},\quad \sum_{b=a}^{n}\tbinom{b}{a} = \tbinom{n+1}{a+1},$$

   on conclut que

   $$\mathbf{P}(k \in R) = \tfrac{(k-1)!(n-k)!}{n!}\tbinom{n}{k} = \tfrac{1}{k}.$$

2. Remarquons que le cardinal $|R|$ est exactement la somme des indicatrices suivantes

$$|R| = \sum_{k=1}^{n} \mathbb{1}_{k \in R}.$$

Par linéarité de l'espérance, on obtient

$$\mathbf{E}(|X|) = \sum_{k=1}^{n} \mathbf{P}(k \in R) = \sum_{k=1}^{n} \tfrac{1}{k}.$$

L'exercice suivant (ou plutôt la résolution proposée) illustre un principe de couplage que l'on retrouve dans l'étude asymptotique des graphes aléatoires.

6.9. Exercice (Exemple de couplage). *Soit n un entier naturel non nul et $p \in]0,1[$, X et Y des variables suivant les lois binomiales de paramètres n et p, et $n+1$ et p respectivement.*

Montrer que, pour tout entier $k \in \mathbb{N}$, $\mathbf{P}(X \geqslant k) \leqslant \mathbf{P}(Y \geqslant k)$.

Solution. L'idée de la solution repose sur un « couplage » : on va construire X' et Y' des variables aléatoires telles que

- ▷ X' suit la même loi que X,
- ▷ Y' suit la même loi que Y,
- ▷ $Y' \geqslant X'$.

Pour cela, considérons $n+1$ variables mutuellement indépendantes $I_1, \ldots, I_{n+1}$ suivant la loi de Bernoulli de paramètre p ; notons X' la somme des n premières et Y' la somme de toutes.

Alors, les variables X' et Y' suivent les lois binomiales de paramètres n et p, et $n+1$ et p respectivement et l'on a $Y' = X' + I_{n+1} \geqslant X'$.

Pour conclure, il suffit alors de remarquer que, pour tout $k \in \mathbb{N}$, l'événement $X' \geqslant k$ est inclus dans l'événement $Y' \geqslant k$, puis d'en déduire

$$\mathbf{P}(X \geqslant k) = \mathbf{P}(X' \geqslant k) \leqslant \mathbf{P}(Y' \geqslant k) = \mathbf{P}(Y \geqslant k).$$

On veillera à ne pas confondre l'événement « $Y' \geqslant X'$ », toujours réalisé, et l'événement « $Y \geqslant X$ », dont on ne connaît pas la probabilité puisque l'on ne connaît pas la loi jointe de (X, Y). On ne sait d'ailleurs même pas si les variables X et Y sont définies sur le même espace probabilisé (ce qui serait le minimum à exiger pour pouvoir considérer l'événement « $Y \geqslant X$ »).

Graphes aléatoires finis

Chapitre III

Modèle binomial $\mathscr{G}_{n,p}$

L'objectif de ce chapitre est d'introduire des « graphes aléatoires » ou plus précisément des lois pour les variables aléatoires à valeurs sur un ensemble de graphes finis. Le premier modèle, dit binomial, est peut-être le plus simple ; il a été introduit en 1959 par Edgar Gilbert [Gil59] et indépendamment par Paul Erdős et Alfréd Rényi [ER59]. Il est souvent baptisé graphe d'Erdős-Rényi mais cette appellation est ambiguë puisque dans leur article de 1959, ces auteurs développent également le modèle uniforme.

Dans cette première approche, nous insisterons sur l'indépendance des arêtes et les dénombrements en moyenne qui seront utiles pour le chapitre suivant et les chapitres de la troisième partie.

1. Définition

1.1. Définition. Soit n un entier naturel non nul et $p \in]0,1[$. Une variable aléatoire X à valeurs graphes suit le modèle binomial de paramètres n et p si la probabilité que X soit égale à un graphe (S,A) dont l'ensemble des sommets est $S = [\![n]\!]$ est

$$\mathbf{P}\big(X = (S,A)\big) = p^{|A|}(1-p)^{\binom{n}{2}-|A|}.$$

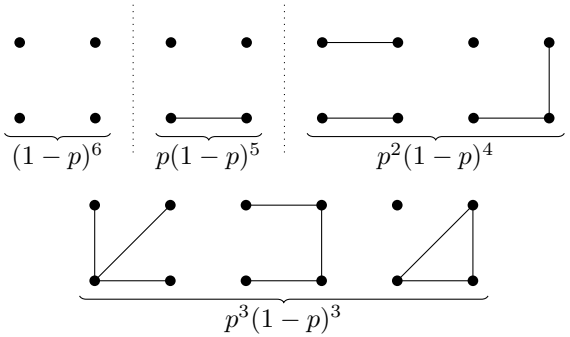
On note alors $X \hookrightarrow \mathscr{G}_{n,p}$ et on dit que X est un graphe aléatoire.

1.2. Remarque. Cela définit bien une loi de probabilité puisque l'on dénombre $\binom{\binom{n}{2}}{k}$ graphes d'ensemble de sommets $S = [\![n]\!]$ et d'ensemble d'arêtes A de cardinal k et que, en regroupant les graphes selon leur nombre d'arêtes et en utilisant la formule du binôme, on vérifie que

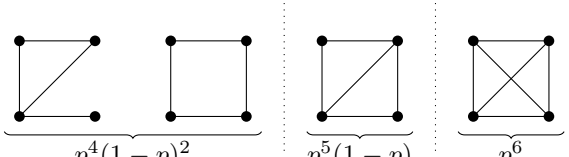
$$\sum_A \mathbf{P}\big(X = (S,A)\big) = \sum_{k=0}^{\binom{n}{2}} \binom{\binom{n}{2}}{k} p^k (1-p)^{\binom{n}{2}-k}$$
$$= (p+1-p)^{\binom{n}{2}} = 1.$$

1.3. Remarque. Le cas particulier $p = \frac{1}{2}$ a été étudié par Paul Erdős dès 1947 [Erd47]. Pour cette valeur spécifique, les différentes probabilités ci-dessus valent toutes $2^{-\binom{n}{2}}$: ainsi, la variable X suit la loi uniforme sur l'ensemble (fini) des graphes d'ensemble de sommets $S = [\![n]\!]$.

1.4. Exemple. Il existe exactement $2^6 = 64$ graphes dont l'ensemble des sommets est $[\![4]\!]$; mais, à isomorphisme près, il n'y a que onze cas à considérer [1] pour définir la loi $\mathscr{G}_{4,p}$. Le schéma ci-dessous indique les probabilités associées à chaque graphe selon sa classe d'isomorphisme :



1. Deux graphes isomorphes ont le même nombre d'arêtes, donc la même probabilité d'apparition ; la réciproque est fausse.



Pour simplifier les écritures, on adoptera quelquefois la notation suivante.

1.5. Notation. Pour un graphe aléatoire X d'ensemble de sommets $[\![n]\!]$, on note A_X son ensemble d'arêtes qui est donc une variable aléatoire à valeurs dans $\mathscr{P}_2([\![n]\!])$.

1.6. Exemple. Soit n un entier naturel non nul, $p \in]0,1[$ puis un graphe aléatoire X suivant le modèle binomial de paramètres n et p. Alors, le graphe complémentaire $([\![n]\!], \mathscr{P}_2([\![n]\!]) \setminus A_X)$ suit le modèle binomial de paramètres n et $1-p$.

1.7. Proposition. *Soit n un entier naturel non nul, $p \in]0,1[$ puis un graphe aléatoire X suivant le modèle binomial de paramètres n et p.*

Alors, pour tout $m \in [\![n]\!]$, le graphe induit $X[[\![m]\!]]$ par X sur l'ensemble de sommets $[\![m]\!]$ suit le modèle binomial de paramètres m et p.

Démonstration. Considérons un graphe $G = ([\![m]\!], A)$ à m sommets et calculons $\mathbf{P}(X[[\![m]\!]] = G)$.

Pour cela, pour $k \geqslant |A|$, dénombrons les graphes $G' = ([\![n]\!], A')$ à n sommets et $|A'| = k$ arêtes dont l'induit $G'[[\![m]\!]]$ est G. Cela correspond aux manières d'ajouter $k - |A|$ arêtes qui ne relient pas deux éléments de $[\![m]\!]$: il y en a donc

$$\binom{\binom{n}{2} - \binom{m}{2}}{k - |A|}.$$

Ainsi, la probabilité $\mathbf{P}(X[[\![m]\!]] = G)$ est égale à la somme des probabilités $\mathbf{P}(X = G')$ pour ces graphes G'.

Par conséquent,

$$\begin{aligned}
&\sum_{k\geqslant|A|}\binom{\binom{n}{2}-\binom{m}{2}}{k-|A|}p^k(1-p)^{\binom{n}{2}-k}\\
&\qquad=p^{|A|}(1-p)^{\binom{m}{2}-|A|}\sum_{l\geqslant 0}\binom{\binom{n}{2}-\binom{m}{2}}{l}p^l(1-p)^{\binom{n}{2}-\binom{m}{2}-l}\\
&\qquad=p^{|A|}(1-p)^{\binom{m}{2}-|A|}.
\end{aligned}$$

♢

2. Indépendance des arêtes

2.1. Proposition. *Soit $X=(\llbracket n\rrbracket, A_X)$ un graphe aléatoire suivant le modèle binomial de paramètres n et p. Pour tout $a\in\mathscr{P}_2(\llbracket n\rrbracket)$, l'événement $a\in A_X$ (l'arête a existe dans le graphe aléatoire X) est réalisé avec probabilité p.*

Démonstration. Dénombrons les graphes $G=(\llbracket n\rrbracket, A')$ tels que l'arête a appartienne à A' et $|A'|=k$: il y en a autant qu'il y a d'ensembles à $k-1$ éléments parmi $\mathscr{P}_2(\llbracket n\rrbracket)\setminus\{a\}$, c'est-à-dire

$$\binom{\binom{n}{2}-1}{k-1}.$$

La probabilité recherchée est, encore avec la formule du binôme,

$$\begin{aligned}
\mathbf{P}(a\in A_X)&=\sum_{k\geqslant 1}\binom{\binom{n}{2}-1}{k-1}p^k(1-p)^{\binom{n}{2}-k}\\
&=p\sum_{l\geqslant 0}\binom{\binom{n}{2}-1}{l}p^l(1-p)^{\binom{n}{2}-1-l}=p.
\end{aligned}$$

♢

2.2. Proposition. *Soit $X=(\llbracket n\rrbracket, A_X)$ un graphe aléatoire suivant le modèle binomial de paramètres n et p. Alors, les différents événements $a\in A_X$ pour $a\in\mathscr{P}_2(\llbracket n\rrbracket)$ sont mutuellement indépendants.*

Autrement dit, dans le modèle binomial, chacune des $\binom{n}{2}$ arêtes existe indépendamment des autres avec la même probabilité p.

Démonstration. Considérons des arêtes $a_1, \dots, a_j \in \mathscr{P}_2(\llbracket n \rrbracket)$ deux à deux distinctes et dénombrons les graphes $G = (\llbracket n \rrbracket, A')$ tels que $|A'| = k$ et $\{a_1, \dots, a_j\} \subset A'$. Comme précédemment, il y en a autant que de parties à $k - j$ éléments parmi

$$\mathscr{P}_2(\llbracket n \rrbracket) \setminus \{a_1, \dots, a_j\},$$

c'est-à-dire

$$\binom{\binom{n}{2} - j}{k - j}.$$

La probabilité recherchée est alors (encore avec la formule du binôme)

$$\begin{aligned}\mathbf{P}(a_1 \in A_X \cap \cdots \cap a_j \in A_X) &= p^j \\ &= \mathbf{P}(a_1 \in A_X) \cdots \mathbf{P}(a_j \in A_X).\end{aligned}$$

Ainsi, les différents événements $a \in A_X$ sont mutuellement indépendants. ♢

Cela permet d'envisager une construction équivalente du graphe binomial arête par arête (et la démonstration est un simple calcul).

2.3. Proposition. *Soit n un entier naturel non nul, $p \in]0, 1[$ puis $(I_a)_{a \in \mathscr{P}_2(\llbracket n \rrbracket)}$ une famille de variables aléatoires mutuellement indépendantes de loi de Bernoulli de paramètre p. Définissons l'ensemble aléatoire d'arêtes suivant :*

$$A = \{a \in \mathscr{P}_2(\llbracket n \rrbracket),\ I_a = 1\}.$$

Alors, la variable aléatoire (à valeurs graphes) $X = (\llbracket n \rrbracket, A)$ suit le modèle binomial de paramètres n et p.

Illustrons cette nouvelle description du modèle binomial avec le résultat important suivant concernant la réunion de deux graphes aléatoires indépendants.

2.4. Proposition (Réunion). *Soit n un entier naturel non nul, p, $p' \in]0,1[$, X et X' des variables aléatoires indépendantes suivant le modèle binomial de paramètres n et p et n et p' respectivement.*

Alors, le graphe obtenu en effectuant la réunion des graphes X et X', c'est-à-dire la réunion des ensembles d'arêtes correspondants puisque les deux graphes ont le même ensemble de sommets $[\![n]\!]$, suit le modèle binomial de paramètres n et

$$1-(1-p)(1-p') = p+p'-pp'.$$

Démonstration. Notons A_X et $A_{X'}$ les ensembles d'arêtes des graphes X et X'.

▷ Soit $a \in \mathscr{P}_2([\![n]\!])$. En utilisant l'indépendance de X et X', donc des événements $a \notin A_X$ et $a \notin A_{X'}$, on obtient

$$\begin{aligned}\mathbf{P}(a \in A_X \cup A_{X'}) &= 1-\mathbf{P}(a \notin A_X, a \notin A_{X'})\\ &= 1-\mathbf{P}(a \notin A_X)\mathbf{P}(a \notin A_{X'})\\ &= 1-(1-p)(1-p')\\ &= p+p'-pp'.\end{aligned}$$

▷ Montrons que les événements $a \in A_X \cup A_{X'}$ (ou de manière équivalente leurs complémentaires $a \in \overline{A_X} \cap \overline{A_{X'}}$) pour $a \in \mathscr{P}_2([\![n]\!])$ sont mutuellement indépendants.

Soit $a_1, \dots, a_k \in \mathscr{P}_2([\![n]\!])$ des arêtes deux à deux distinctes. Alors,

$$\begin{aligned}\mathbf{P}\Big(\bigcap_{i=1}^{k}(a_i \in \overline{A_X} \cap \overline{A_{X'}})\Big) &= \mathbf{P}\Big(\bigcap_{i=1}^{k}(a_i \in \overline{A_X})\bigcap(a_i \in \overline{A_{X'}})\Big)\\ &= \mathbf{P}\Big(\bigcap_{i=1}^{k}(a_i \in \overline{A_X})\bigcap\bigcap_{i=1}^{k}(a_i \in \overline{A_{X'}})\Big)\\ &= \mathbf{P}\Big(\bigcap_{i=1}^{k}(a_i \in \overline{A_X})\Big)\mathbf{P}\Big(\bigcap_{i=1}^{k}(a_i \in \overline{A_{X'}})\Big),\end{aligned}$$

en utilisant l'indépendance des deux graphes X et X'.

Utilisons ensuite l'indépendance des arêtes dans chaque graphe :

$$\begin{aligned}\mathbf{P}\Big(\bigcap_{i=1}^{k}(a_i \in \overline{A_X} \cap \overline{A_{X'}})\Big) &= \Big(\prod_{i=1}^{k}\mathbf{P}(a_i \in \overline{A_X})\Big)\Big(\prod_{i=1}^{k}\mathbf{P}(a_i \in \overline{A_{X'}})\Big)\\ &= \prod_{i=1}^{k}\mathbf{P}(a_i \in \overline{A_X})\mathbf{P}(a_i \in \overline{A_{X'}})\\ &= \prod_{i=1}^{k}\mathbf{P}(a_i \in \overline{A_X} \cap \overline{A_{X'}})\end{aligned}$$

En conclusion, les arêtes de la réunion de X et de X' sont mutuellement indépendantes et apparaissent avec probabilité $p+p'-pp'$, d'où la propriété annoncée. ♢

2.5. Remarque (Couplage). Ce résultat est important puisqu'il permet le « couplage » de graphes aléatoires suivant le modèle binomial avec le même premier paramètre mais avec des probabilités différentes. Pour être plus précis, si $p \leqslant \tilde{p}$, on peut construire deux graphes aléatoires X et $\tilde{X}$ suivant le modèle binomial de paramètres n et p et n et $\tilde{p}$ respectivement tels que l'ensemble d'arêtes de X soit inclus dans l'ensemble d'arêtes de $\tilde{X}$.

En effet, il suffit de construire $\tilde{X}$ en prenant la réunion de X avec un graphe aléatoire suivant le modèle binomial de paramètres n et $p' = \frac{\tilde{p}-p}{1-p}$, indépendant de X : d'après la proposition précédente, le graphe obtenu suit le modèle binomial de paramètres n et

$$p + p' - pp' = p + \tfrac{\tilde{p}-p}{1-p}(1-p) = \tilde{p}.$$

Cela sera particulièrement intéressant lorsqu'il s'agira d'étudier, au chapitre VII, la monotonie de certaines propriétés lorsque le paramètre p varie.

3. Dénombrements en moyenne

Dans cette section, on cherche à regarder quelques caractéristiques en moyenne du graphe aléatoire suivant le modèle binomial.

3.1. Proposition. *Soit n un entier naturel non nul et $p \in]0,1[$. Le nombre moyen d'arêtes d'un graphe aléatoire X suivant le modèle binomial de paramètres n et p est $\binom{n}{2}p$.*

Démonstration. La variable aléatoire N égale au nombre d'arêtes de X est la somme des indicatrices des événements $a \in A_X$ pour toute arête $a \in \mathscr{P}_2([\![n]\!])$. Par linéarité de l'espérance,

$$\begin{aligned}\mathbf{E}(N) &= \sum_{a\in\mathscr{P}_2([\![n]\!])} \mathbf{E}(\mathbb{1}_{a\in A_X}) \\ &= \sum_{a\in\mathscr{P}_2([\![n]\!])} \mathbf{P}(a \in A_X) = \tbinom{n}{2}p.\end{aligned}$$

♢

3.2. Proposition. *Soit n un entier naturel non nul et $p \in]0,1[$. Le nombre moyen de sommets isolés (c'est-à-dire de degré 0) d'un graphe aléatoire X suivant le modèle binomial de paramètres n et p est $n(1-p)^{n-1}$.*

Démonstration. La variable aléatoire N égale au nombre de sommets isolés de X est la somme sur le sommet i des variables aléatoires

$$\prod_{j\neq i} \mathbb{1}_{\{i,j\}\notin A_X},$$

qui indiquent qu'il n'y a pas d'arête issue de i.

Par linéarité de l'espérance puis par indépendance des variables indicatrices des arêtes,

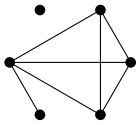
$$\begin{aligned}\mathbf{E}(N) &= \sum_{i=1}^{n} \mathbf{E}\Big(\prod_{j\neq i} \mathbb{1}_{\{i,j\}\notin A_X}\Big) \\ &= \sum_{i=1}^{n} (1-p)^{n-1} = n(1-p)^{n-1}.\end{aligned}$$

♢

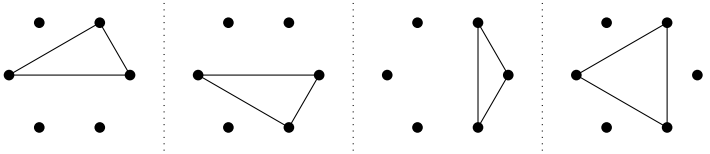
Dans la proposition suivante, on dénombre les « triangles » d'un graphe aléatoire, c'est-à-dire les ensembles de trois sommets reliés

deux à deux par des arêtes (ou pour le dire plus rigoureusement, le nombre de graphes induits isomorphes à K_3).

Par exemple, le graphe suivant comporte quatre triangles



dont voici les tracés « isolés »



3.3. Proposition. *Soit n un entier naturel non nul et $p \in]0,1[$. Le nombre moyen de triangles d'un graphe aléatoire X suivant le modèle binomial de paramètres n et p est $\binom{n}{3}p^3$.*

Démonstration. La variable aléatoire N égale au nombre de triangles dans X est la somme sur les $\{i,j,k\} \in \mathscr{P}_3(\llbracket n \rrbracket)$ des variables aléatoires indicatrices

$$\mathbb{1}_{\{i,j\}\in A_X}\mathbb{1}_{\{j,k\}\in A_X}\mathbb{1}_{\{i,k\}\in A_X}.$$

Par linéarité de l'espérance,

$$\begin{aligned}\mathbf{E}(N) &= \sum_{\{i,j,k\}\in\mathscr{P}_3(\llbracket n\rrbracket)} \mathbf{E}(\mathbb{1}_{\{i,j\}\in A_X}\mathbb{1}_{\{j,k\}\in A_X}\mathbb{1}_{\{i,k\}\in A_X}) \\ &= \sum_{\{i,j,k\}\in\mathscr{P}_3(\llbracket n\rrbracket)} p^3 = \binom{n}{3}p^3.\end{aligned}$$

♢

3.4. Remarque. On peut immédiatement généraliser le résultat à de plus grands graphes induits : le nombre moyen d'occurrences d'un graphe complet à k sommets (donc à $\binom{k}{2}$ arêtes) dans un graphe aléatoire X suivant le modèle binomial de paramètres n et p est $\binom{n}{k}p^{\binom{k}{2}}$.

3.5. Proposition. *Soit n, k deux entiers naturels non nuls et $p \in]0,1[$. Le nombre moyen d'occurrences du graphe cyclique C_k dans un graphe aléatoire X suivant le modèle binomial de paramètres n et p est*

$$\frac{n(n-1)\cdots(n-k+1)}{2k} p^k.$$

Attention, plusieurs k-uplets de sommets de $[\![n]\!]$ correspondent à un même graphe cyclique dans K_n. En effet, en agissant par permutation circulaire du k-uplet ou par retournement de l'ordre, on conserve la structure d'un cycle (non orienté) ; il y a donc $2k$ k-uplets pour un même cycle, d'où le dénominateur dans cette formule.

Démonstration. Soit N la variable aléatoire égale au nombre d'occurrences de C_k dans X. Alors, $2N$ est la somme sur les k-cycles $\sigma = (x_1\ x_2\ \cdots\ x_k)$ de $[\![n]\!]$ des variables aléatoires indicatrices

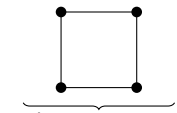
$$\prod_{j=1}^{k} \mathbb{1}_{\{x_j,\sigma(x_j)\}\in A_X}.$$

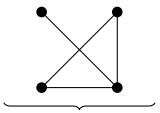
Par indépendance des arêtes, chacune de ces variables est de probabilité p^k. Ensuite, par linéarité de l'espérance,

$$\mathbf{E}(2N) = \tbinom{n}{k}(k-1)!p^k = \tfrac{n(n-1)\cdots(n-k+1)}{k} p^k,$$

d'où l'on déduit la formule attendue pour $\mathbf{E}(N)$. ♢

3.6. Remarque. Le nombre moyen d'occurrences d'un graphe donné dans un graphe aléatoire ne dépend pas uniquement du nombre de sommets et d'arêtes mais du nombre de façons distinctes de réaliser ce graphe dans le graphe complet K_n. Par exemple, dans K_4, les graphes suivants n'ont pas la même fréquence d'apparition


trois occurrences


douze occurrences

Par conséquent, les variables aléatoires de comptage de leurs occurrences n'auront pas la même espérance (même si individuellement les occurrences ont la même probabilité d'apparition puisqu'elles comportent le même nombre d'arêtes).

De manière plus savante, le nombre moyen d'occurrences d'un graphe donné G dans un graphe aléatoire suivant le modèle binomial dépend du nombre d'automorphismes du graphe G (voir l'exercice I-8.6 pour le calcul exact).

4. Modèle uniforme

Ce modèle très proche du précédent en constitue une alternative.

4.1. Définition. Soit n un entier naturel non nul et M un entier naturel inférieur à $\binom{n}{2}$. Une variable aléatoire X à valeurs graphes suit le modèle uniforme de paramètres n et M si, pour tout graphe (S, A) d'ensemble de sommets $S = [\![n]\!]$ et tel que $|A| = M$,

$$\mathbf{P}\big(X = (S, A)\big) = \frac{1}{\binom{\binom{n}{2}}{M}}.$$

4.2. Remarque. Montrons que, dans ce modèle, les variables indicatrices des arêtes ne sont plus indépendantes.

Considérons $X = ([\![n]\!], A_X)$ un graphe aléatoire suivant le modèle uniforme de paramètres n et M, a_1, a_2 des éléments distincts de $\mathscr{P}_2([\![n]\!])$.

▷ Comme il y a $\binom{\binom{n}{2}-1}{M-1}$ parties de $\mathscr{P}_2([\![n]\!])$ de cardinal M contenant a_1,

$$\mathbf{P}(a_1 \in A_X) = \binom{\binom{n}{2}-1}{M-1} \Big/ \binom{\binom{n}{2}}{M} = \frac{M}{\binom{n}{2}}.$$

On obtient de même $\mathbf{P}(a_2 \in A_X)$.

▷ Comme il y a $\binom{\binom{n}{2}-2}{M-2}$ parties de $\mathscr{P}_2([\![n]\!])$ de cardinal M contenant a_1 et a_2,

$$\mathbf{P}(a_1 \in A_X, a_2 \in A_X) = \frac{\binom{\binom{n}{2}-2}{M-2}}{\binom{\binom{n}{2}}{M}} = \frac{M(M-1)}{\binom{n}{2}\left(\binom{n}{2}-1\right)}.$$

Ainsi, $\mathbf{P}(a_1 \in A_X, a_2 \in A_X) \neq \mathbf{P}(a_1 \in A_X)\mathbf{P}(a_2 \in A_X)$.

Cette propriété de dépendance constitue un véritable frein à l'utilisation de ce modèle pour des calculs (par opposition au modèle binomial).

4.3. Proposition. *Soit n un entier naturel non nul, $p \in]0,1[$ et X un graphe aléatoire suivant le modèle binomial de paramètres n et p.*

Alors, pour tout graphe (S,A) d'ensemble de sommets $S = [\![n]\!]$ et tel que $|A| = M$,

$$\mathbf{P}\big(X = (S,A) \big| |A_X| = |A|\big) = \frac{1}{\binom{\binom{n}{2}}{|A|}}.$$

En d'autres termes, conditionnellement à $|A_X| = M$, la variable X suit le modèle uniforme de paramètres n et M.

Démonstration. D'après la définition de probabilité conditionnelle, il suffit de calculer les quotients des probabilités

$$\begin{aligned}\mathbf{P}\big(X = (S,A), |A_X| = |A|\big) &= \mathbf{P}\big(X = (S,A)\big) \\ &= p^{|A|}(1-p)^{\binom{n}{2}-|A|}\end{aligned}$$

et

$$\begin{aligned}\mathbf{P}\big(|A_X| = |A|\big) &= \sum_{\substack{B \in \mathscr{P}(S) \\ |B|=|A|}} \mathbf{P}\big(X = (S,B)\big) \\ &= \binom{\binom{n}{2}}{|A|} p^{|A|}(1-p)^{\binom{n}{2}-|A|}.\end{aligned}$$

Le résultat en découle immédiatement. ♢

5. Exercices

5.1. Exercice. *Soit n un entier naturel non nul, X un graphe aléatoire suivant le modèle binomial de paramètres n et $\frac{1}{2}$, S_1 et S_2 deux parties de $[\![n]\!]$ non vides et de même cardinal. Considérons l'événement E_1 (respectivement E_2) défini par « le graphe induit $X[S_1]$ (respectivement $X[S_2]$) ne comporte aucune arête. »*

Montrer que E_1 et E_2 sont indépendants si, et seulement si,

$$|S_1 \cap S_2| \leqslant 1.$$

Solution. Notons k le cardinal commun de S_1 et S_2 et l le cardinal de l'intersection $S_1 \cap S_2$.
Notons d'une part qu'il y a $\binom{k}{2}$ arêtes possibles à sommets dans S_1 (respectivement S_2) donc autant d'arêtes qui doivent être absentes dans le graphe induit $X[S_1]$ (respectivement $X[S_2]$) pour la réalisation de l'événement E_1 (respectivement E_2), d'où

$$\mathbf{P}(E_1) = \mathbf{P}(E_2) = 2^{-\binom{k}{2}}.$$

D'autre part, Le nombre d'arêtes à sommets dans S_1 ou à sommets dans S_2 est (car il y a l sommets dans $S_1 \cap S_2$ donc $\binom{l}{2}$ arêtes comptées deux fois)

$$\binom{k}{2} + \binom{k}{2} - \binom{l}{2} = 2\binom{k}{2} - \binom{l}{2}.$$

Par conséquent,

$$\mathbf{P}(E_1 \cap E_2) = 2^{-2\binom{k}{2}+\binom{l}{2}}.$$

L'égalité $\mathbf{P}(E_1 \cap E_2) = \mathbf{P}(E_1)\mathbf{P}(E_2)$ qui définit l'indépendance des événements E_1 et E_2 équivaut donc à $\binom{l}{2} = 0$, soit $l(l-1) = 0$, ou encore $l \in \{0, 1\}$.

Il semble relativement intuitif que les comportements de deux graphes induits soient indépendants lorsqu'il n'y a aucune arête commune à ces graphes.

Un rapide calcul montre que le résultat demeure vrai pour un paramètre $p \neq \frac{1}{2}$ ou des parties S_1 et S_2 de cardinaux différents.

5.2. Exercice. *Soit n un entier naturel non nul, $p \in]0, 1[$, et X un graphe aléatoire suivant le modèle binomial de paramètres n et p.*

Montrer que le nombre N d'arêtes de X suit une loi binomiale dont on précisera les paramètres.

Solution. Donnons deux solutions.

▷ Remarquons que, d'après la construction alternative du modèle binomial par le choix d'arêtes mutuellement indépendantes (c'est-à-dire d'après la proposition III-2.3), le nombre N d'arêtes de X s'obtient comme somme de $\binom{n}{2}$ variables de Bernoulli de paramètre p mutuellement indépendantes donc suit la loi binomiale de paramètres $\binom{n}{2}$ et p.

▷ Si l'on souhaite revenir à la définition du modèle binomial, on décompose, pour tout $k \leqslant \binom{n}{2}$, l'événement $N = k$ selon les différents graphes d'ensembles de sommets $[\![n]\!]$ et admettant exactement k arêtes. Ainsi,

$$\begin{aligned}\mathbf{P}(N=k) &= \sum_{A\in\mathscr{P}_k([\![n]\!])} \mathbf{P}\big(X=([\![n]\!],A)\big)\\ &= \sum_{A\in\mathscr{P}_k([\![n]\!])} p^{|A|}(1-p)^{\binom{n}{2}-|A|}\\ &= \tbinom{n}{k}p^k(1-p)^{\binom{n}{2}-k}.\end{aligned}$$

On retrouve bien que la variable N suit la loi binomiale de paramètres $\binom{n}{2}$ et p.

5.3. Exercice. *Soit n un entier naturel non nul et M un entier naturel inférieur à $\binom{n}{2}$ puis X un graphe aléatoire qui suit le modèle uniforme de paramètres n et M.*

Déterminer la loi du degré d'un sommet de X.

Solution. Soit $s \in [\![n]\!]$ un sommet de X. Pour tout entier k entre 0 et $n-1$, le degré de s est égal à k si, et seulement si, il existe dans le graphe X

- ▷ k arêtes tracées parmi les $n-1$ issues du sommet s,
- ▷ $M-k$ arêtes tracées parmi les $\binom{n}{2}-(n-1)$ qui n'admettent pas le sommet s comme extrémité.

Ainsi, il existe

$$\binom{n-1}{k}\binom{\binom{n}{2}-(n-1)}{M-k}$$

graphes à M arêtes tels que $\deg(s)=k$.

Comme X suit le modèle uniforme, on obtient la probabilité cherchée

$$\mathbf{P}(\deg(s)=k)=\frac{\binom{n-1}{k}\binom{\binom{n}{2}-(n-1)}{M-k}}{\binom{\binom{n}{2}}{M}}.$$

On reconnaît ici que la variable $\deg(s)$ suit une loi hypergéométrique (alors que dans le modèle binomial, elle suit la loi binomiale de paramètres $n-1$ et p).

On peut remarquer que les variables $\deg(s)$ *et* $\deg(s')$ *pour deux sommets distincts* s *et* s' *ne sont indépendantes ni dans le modèle binomial, ni dans le modèle uniforme puisque, dans les deux cas,*

$$\mathbf{P}\big(\deg(s)=0,\deg(s')=n-1\big)=0,$$

(il n'est pas possible qu'un sommet soit relié à tous les autres et, en même temps, qu'un sommet soit isolé), mais $\mathbf{P}\big(\deg(s)=0\big)\neq 0$ *et* $\mathbf{P}\big(\deg(s')=n-1\big)\neq 0$.

5.4. Exercice. *Soit* $p\in\,]0,1[$. *Notons, pour tout* $n\geqslant 1$, u_n *la probabilité qu'un graphe aléatoire suivant le modèle binomial de paramètres* n *et* p *soit connexe.*

Montrer que, pour tout $n\geqslant 1$,

$$1-u_n=\sum_{k=1}^{n-1}\tbinom{n-1}{k-1}u_k(1-p)^{k(n-k)}.$$

Solution. Soit n un entier naturel non nul, et X un graphe aléatoire suivant le modèle binomial de paramètres n et p et Y la variable aléatoire à valeurs graphes égale à la composante connexe de 1 dans X.

Partitionnons l'événement complémentaire à celui considéré, à savoir « X n'est pas connexe », comme la réunion disjointe des événements « $Y=A\cup\{1\}$ » avec $A\in\mathscr{P}([\![2,n]\!])$ telle que $A\neq[\![2,n]\!]$. Alors,

$$1-u_n=\sum_{k=1}^{n-1}\sum_{A\in\mathscr{P}_{k-1}([\![2,n]\!])}\mathbf{P}(Y=A).$$

Pour une partie $A\in\mathscr{P}_{k-1}([\![2,n]\!])$ fixée, l'événement « $Y=A$ » est l'intersection des événements indépendants suivants :

- ▷ « le graphe $X[A \cup \{1\}]$ est connexe » ; comme $|A \cup \{1\}| = k$, la variable $X[A \cup \{1\}]$ suit le modèle binomial de paramètres k et p (quitte à renuméroter les sommets) ; la probabilité de cet événement est donc u_k,
- ▷ « dans X, il n'existe pas d'arête entre les k sommets de la partie $A \cup \{1\}$ et les $n-k$ sommets de son complémentaire », événement de probabilité $p^{k(n-k)}$.

Combinant tout cela, on obtient

$$\begin{aligned} 1 - u_n &= \sum_{k=1}^{n-1} \sum_{A \in \mathscr{P}_{k-1}([\![2,n]\!])} u_k (1-p)^{k(n-k)} \\ &= \sum_{k=1}^{n-1} \tbinom{n-1}{k-1} u_k (1-p)^{k(n-k)}. \end{aligned}$$

L'exercice suivant prolonge la démarche de la proposition III-2.4 en remplaçant la réunion de deux graphes aléatoires suivant le modèle binomial par leur intersection : les différentes étapes de calcul sont proches. En revanche, le résultat obtenu est de moindre portée puisqu'il ne s'interprète pas directement comme un « couplage ».

5.5. Exercice. *Soit n un entier naturel non nul, p, $p' \in]0,1[$, X et X' des variables aléatoires indépendantes suivant le modèle binomial de paramètres n et p et n et p' respectivement.*

Déterminer la loi de l'intersection de X et X', c'est-à-dire du graphe d'ensemble de sommets $S = [\![n]\!]$ et d'ensemble d'arêtes l'intersection des ensembles d'arêtes de X et X'.

Solution. ▷ Soit $a \in \mathscr{P}_2([\![n]\!])$. Notons A_X et $A_{X'}$ les ensembles d'arêtes des graphes X et X'. Alors, en utilisant l'indépendance des variables X et X' donc celle des deux événements $a \notin A_X$ et $a \notin A_{X'}$, on obtient la probabilité que l'arête a appartienne au graphe intersection :

$$\begin{aligned} \mathbf{P}(a \in A_X \cap A_{X'}) &= \mathbf{P}(a \in A_X, a \in A_{X'}) \\ &= \mathbf{P}(a \in A_X)\mathbf{P}(a \in A_{X'}) = pp'. \end{aligned}$$

▷ Montrons que les événements $a \in A_X \cap A_{X'}$ pour $a \in \mathscr{P}_2([\![n]\!])$ sont mutuellement indépendants. Soit $a_1, \ldots, a_k \in \mathscr{P}_2([\![n]\!])$ des arêtes deux à deux distinctes.

Alors, en utilisant tour à tour l'indépendance de X et X', puis celle des arêtes au sein de chacun de ses graphes, on obtient :

$$\begin{aligned}
\mathbf{P}\Big(\bigcap_{i=1}^{k}(a_i \in A_X \cap A_{X'})\Big) &= \mathbf{P}\Big(\bigcap_{i=1}^{k}(a_i \in A_X) \cap (a_i \in A_{X'})\Big) \\
&= \mathbf{P}\Big(\bigcap_{i=1}^{k}(a_i \in A_X)\bigcap\bigcap_{i=1}^{k}(a_i \in A_{X'})\Big) \\
&= \mathbf{P}\Big(\bigcap_{i=1}^{k}(a_i \in A_X)\Big)\mathbf{P}\Big(\bigcap_{i=1}^{k}(a_i \in A_{X'})\Big) \\
&= \Big(\prod_{i=1}^{k}\mathbf{P}(a_i \in A_X)\Big)\Big(\prod_{i=1}^{k}\mathbf{P}(a_i \in A_{X'})\Big) \\
&= \prod_{i=1}^{k}\mathbf{P}(a_i \in A_X)\mathbf{P}(a_i \in A_{X'}) \\
&= \prod_{i=1}^{k}\mathbf{P}(a_i \in A_X \cap A_{X'}).
\end{aligned}$$

En conclusion, les arêtes de l'intersection de X et de X' sont mutuellement indépendantes et apparaissent avec probabilité pp' : ce graphe suit le modèle binomial de paramètres n et pp'.

On aurait également pu passer au complémentaire et utiliser la proposition III-2.4 en se rappelant que le complémentaire d'un graphe suivant le modèle binomial de paramètres n et p suit le modèle binomial de paramètres n et $1-p$.

5.6. Exercice. *Soit n un entier naturel non nul, p, $p' \in]0,1[$, X et X' des variables aléatoires indépendantes suivant le modèle binomial de paramètres n et p et n et p' respectivement.*

Calculer la probabilité de l'événement « X est un sous-graphe de X' ».

Solution. Utilisons la formule des probabilités totales pour discuter selon la valeur prise par l'ensemble A_X des arêtes de X. Plus précisément, la discussion porte tour à tour sur le nombre k d'arêtes puis sur la partie A des k arêtes effectivement réalisées.

$$
\begin{aligned}
&\mathbf{P}(A_X \subset A_{X'}) \\
&\quad= \sum_{k=0}^{\binom{n}{2}} \sum_{A \in \mathscr{P}_k([\![n]\!])} \mathbf{P}(A_X \subset A_{X'} \mid A_X = A)\mathbf{P}(A_X = A) \\
&\quad= \sum_{k=0}^{\binom{n}{2}} \sum_{A \in \mathscr{P}_k([\![n]\!])} \mathbf{P}(A \subset A_{X'})\mathbf{P}(A_X = A) \\
&\quad= \sum_{k=0}^{\binom{n}{2}} \binom{\binom{n}{2}}{k} p'^k p^k (1-p)^{\binom{n}{2}-k} \\
&\quad= (1-p+pp')^{\binom{n}{2}}.
\end{aligned}
$$

On peut vérifier la cohérence du résultat en testant les cas limites : pour la valeur extrême $p' = 1$, X' est le graphe complet et l'on retrouve que la probabilité recherchée est 1 ; pour l'autre limite $p' = 0$, X' est le graphe sans arêtes, donc la probabilité est dans ce cas celle que X soit sans arêtes, à savoir

$$(1-p)^{\binom{n}{2}}.$$

5.7. Exercice. *Soit n un entier naturel non nul, $p \in]0,1[$ puis X un graphe aléatoire qui suit le modèle binomial de paramètres n et p.*

1. *Déterminer le nombre moyen de sommets qui ne sont reliés qu'à un seul autre sommet.*
2. *Soit $k \in [\![n-1]\!]$. Déterminer le nombre moyen de sommets de degré k.*

Solution. Notons A l'ensemble d'arêtes de X.

1. ▷ Commençons par calculer la probabilité qu'un sommet i fixé soit de degré 1. Cet événement se récrit

$$\bigcup_{j \neq i} \{\{i,j\} \in A\} \bigcap \bigcap_{k \neq i,j} \{\{i,k\} \notin A\}$$

donc sa probabilité est, avec l'indépendance entre arêtes,

$$\mathbf{P}\big(\deg(i) = 1\big) = \sum_{j \neq i} p(1-p)^{n-2} = (n-1)p(1-p)^{n-2}.$$

En poursuivant ce calcul, on retrouve que la variable $\deg(i)$ *suit la loi binomiale de paramètres* $n-1$ *et* p.

▷ Soit N_1 la variable aléatoire égale au nombre de sommets qui ne sont reliés qu'à un seul autre sommet. Écrivons N_1 comme somme de variables indicatrices

$$N_1 = \sum_{i=1}^{n} \mathbb{1}_{\deg(i)=1},$$

puis utilisons la linéarité de l'espérance

$$\mathbf{E}(N_1) = \sum_{i=1}^{n} \mathbf{P}\big(\deg(i)=1\big) = n(n-1)p(1-p)^{n-2}.$$

2. On procède de même en remarquant tout d'abord que, pour un sommet i fixé,

$$\mathbf{P}\big(\deg(i)=k\big) = \tbinom{n-1}{k} p^k (1-p)^{n-1-k}.$$

Par linéarité, on obtient que l'espérance de la variable aléatoire N_k égale au nombre de sommets de degré k vaut

$$\mathbf{E}(N_k) = n\tbinom{n-1}{k} p^k (1-p)^{n-1-k}.$$

On peut vérifier rapidement que

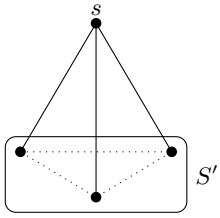
$$\sum_{k=0}^{n-1} \mathbf{E}(N_k) = n \sum_{k=0}^{n-1} \tbinom{n-1}{k} p^k (1-p)^{n-1-k} = n,$$

ce qui est cohérent avec l'identité élémentaire $\sum_{k=0}^{n-1} N_k = n$.

5.8. Exercice. *Soit* n *un entier naturel non nul,* $p \in]0,1[$ *puis* X *un graphe aléatoire qui suit le modèle binomial de paramètres* n *et* p. *Déterminer le nombre moyen de graphes induits par* X *isomorphes à la « griffe »* $K_{1,3}$.

Solution. Notons N la variable aléatoire égale au nombre de griffes de X.

Remarquons qu'une griffe est décrite par un couple (s, S'), où s est un sommet et S' un ensemble de trois sommets distincts de s tel que toutes les arêtes entre s et S' existent mais qu'aucune arête entre sommets de S' ne soit tracée.



Ainsi,

$$N = \sum_{s=1}^{n} \sum_{S' \in \mathscr{P}_3(\llbracket n \rrbracket \setminus \{s\})} \mathbb{1}_{(s,S') \text{ griffe de } X}.$$

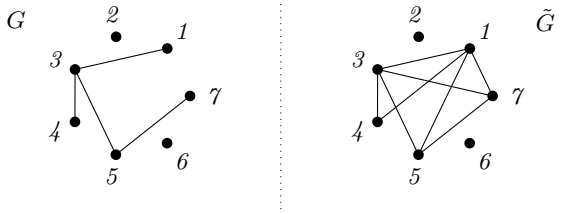
Par linéarité de l'espérance,

$$\begin{aligned}\mathbf{E}(N) &= \sum_{s=1}^{n} \sum_{S' \in \mathscr{P}_3(\llbracket n \rrbracket \setminus \{s\})} \mathbf{P}((s,S') \text{ griffe de } X) \\ &= \sum_{s=1}^{n} \sum_{S' \in \mathscr{P}_3(\llbracket n \rrbracket \setminus \{s\})} p^3(1-p)^3 \\ &= n\tbinom{n-1}{3} p^3(1-p)^3.\end{aligned}$$

5.9. Exercice (Graphe de Barak-Erdős [BE84]). *Soit n un entier naturel non nul. Pour tout graphe G d'ensemble de sommets $\llbracket n \rrbracket$ et d'ensemble d'arêtes A, définissons le graphe $\tilde{G}$ clôture transitive d'ensemble de sommets $\llbracket n \rrbracket$ dans lequel une arête relie deux sommets i et j tels que $j > i$ dans $\tilde{G}$ si, et seulement si, il existe une suite finie croissante*

$$i_0 = i < i_1 < \cdots < i_N = j$$

telle que, pour tout $k \in \llbracket N \rrbracket$, $\{i_{k-1}, i_k\} \in A$.
Autrement dit, il existe une arête entre i et j dans $\tilde{G}$ s'il existe dans G un chemin joignant i à j et rencontrant les sommets dans l'ordre croissant.



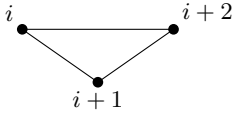
Dans l'exemple dessiné ci-avant, on trouve par exemple une arête entre 1 *et* 5 *dans* $\tilde{G}$*, car il y a des arêtes entre* 1 *et* 3 *et entre* 3 *et* 5 *dans* G*. En revanche, il n'y a pas d'arête entre* 4 *et* 5 *dans* $\tilde{G}$*, même s'il y a des arêtes entre* 4 *et* 3 *et entre* 3 *et* 5 *dans* G *car ce chemin ne respecte pas la condition de « croissance » des sommets rencontrés.*

1. *Déterminer le graphe clôture transitive du graphe divisoriel* D_n *défini par* $(\llbracket n \rrbracket, A)$ *où*
$$A = \big\{\{k, l\},\ k \text{ divise } l\big\}.$$
2. *Déterminer le graphe clôture transitive du graphe chemin* P_n *défini par* $(\llbracket n \rrbracket, A)$ *où*
$$A = \big\{\{k, k+1\},\ k \in \llbracket n-1 \rrbracket\big\}.$$
3. *Fixons* $p \in]0, 1[$ *puis considérons* X *un graphe aléatoire qui suit le modèle binomial de paramètres* n *et* p*.*

   Soit $i \in \llbracket n-2 \rrbracket$*. Déterminer la probabilité que l'arête* $\{i, i+2\}$ *apparaisse dans la clôture transitive* $\tilde{X}$*.*

Solution.

1. Soit $i < j$. Alors, l'arête $\{i, j\}$ figure dans $\tilde{D}_n$ si, et seulement si, il existe $i_0 = i < i_1 < \cdots < i_N = j$ tels que i_{k-1} divise i_k pour tout $k \in \llbracket N \rrbracket$: cela entraîne en particulier que i divise j, donc que $\{i, j\} \in A$. Par conséquent, $\tilde{D}_n = D_n$.
2. Soit $i < j$. Alors, l'arête $\{i, j\}$ figure dans $\tilde{P}_n$ si, et seulement si, il existe $i_0 = i < i_1 < \cdots < i_N = j$ tels que $i_k = i_{k-1} + 1$ pour tout $k \in \llbracket N \rrbracket$: cela entraîne que toute arête $\{i, j\}$ existe dans $\tilde{P}_n$, donc que le graphe obtenu est complet.
3. Notons, pour tout $i < j$, $E_{i,j}$ l'événement « l'arête $\{i, j\}$ existe dans X ». L'événement E considéré dans cette question est réalisé lorsque l'une des deux conditions suivantes l'est :
   - ▷ il y a une arête joignant i et $i+2$ dans X ;
   - ▷ il y a un chemin « croissant » de longueur au moins 2 joignant i à $i+2$ dans X (et la condition de croissance impose que ce chemin passe successivement par les sommets i, $i+1$ et $i+2$).



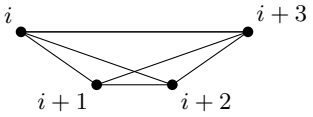
Il s'écrit donc $E_{i,i+2} \cup (E_{i,i+1} \cap E_{i+1,i+2})$. Alors, en utilisant la formule pour la réunion d'événements puis l'indépendance des différents événements $(E_{i,j})_{i<j}$, on obtient

$$\begin{aligned}\mathbf{P}(E) &= \mathbf{P}(E_{i,i+2}) + \mathbf{P}(E_{i,i+1} \cap E_{i+1,i+2}) \\ &\qquad - \mathbf{P}(E_{i,i+2} \cap E_{i,i+1} \cap E_{i+1,i+2}) \\ &= p + p^2 - p^3.\end{aligned}$$

On peut poursuivre le calcul dans $\tilde{X}$: par exemple, avec la formule du crible, toujours en distinguant les différents chemins possibles, on trouve que, pour tout $i \in [\![n-3]\!]$, l'événement « l'arête $\{i, i+3\}$ apparaît dans la clôture transitive $\tilde{X}$ » se récrit comme réunion des événements

$$E_{i,i+3},\ E_{i,i+1} \cap E_{i+1,i+3},\ E_{i,i+2} \cap E_{i+2,i+3},$$
$$E_{i,i+1} \cap E_{i+1,i+2} \cap E_{i+2,i+3},$$

en discutant les différents chemins « croissants » possibles joignant i et $i+3$.



Sa probabilité est donc, après application de la formule du crible,

$$p + 2p^2 - p^3 - 4p^4 + 4p^5 - p^6.$$

Chapitre IV

Méthode probabiliste

L'objectif de ce chapitre est d'illustrer une technique, appelée « méthode probabiliste », qui exploite des arguments de probabilités afin d'établir un résultat où les probabilités n'apparaissent pas ; ici, on utilise un graphe aléatoire (et plus précisément des calculs d'espérances de variables associées à ce graphe) afin d'obtenir des propriétés déterministes de graphes ou d'assurer l'existence de graphes vérifiant certaines propriétés.

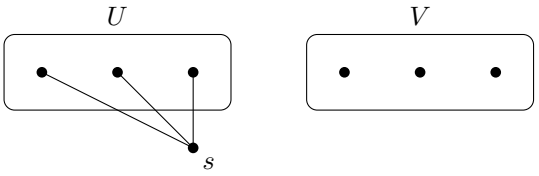
Les exemples des sections suivantes sont totalement indépendants.

1. Propriété de k-extension

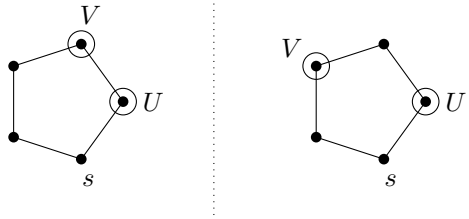
1.1. Définition. Soit k un entier naturel non nul.
Un graphe (S, A) vérifie la propriété de k-extension si pour toutes parties U et V de S, disjointes et de (même) cardinal k, il existe un sommet $s \in S \setminus (U \cup V)$ tel que

- ▷ pour tout $x \in U$, $\{x, s\} \in A$, autrement dit s est voisin de tous les sommets de U,
- ▷ pour tout $x \in V$, $\{x, s\} \notin A$, autrement dit s n'est voisin d'aucun sommet de V.

Voici une illustration pour $k = 3$:

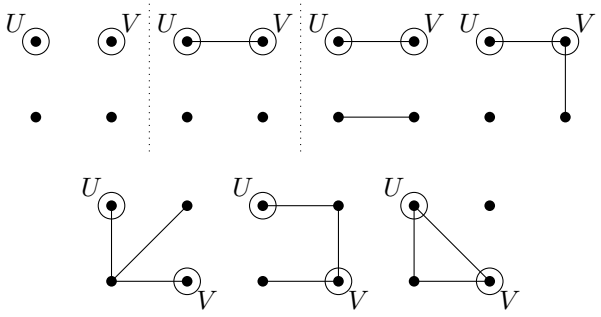


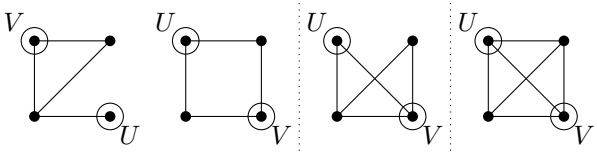
1.2. Exemple. Le graphe cycle à cinq sommets vérifie la propriété de 1-extension ; en effet, quels que soient les sommets choisis pour U et V, il existe un voisin de U qui n'est pas voisin de V :



1.3. Remarque. Il n'existe pas de graphe à quatre sommets satisfaisant la propriété de 1-extension (ce qui indique que C_5 est le plus petit graphe vérifiant cette propriété).

Il suffit de trouver pour chaque classe d'isomorphisme de graphes à quatre sommets des singletons U et V ne satisfaisant pas la propriété ; sur le schéma suivant des différentes classes, on a indiqué des singletons U et V problématiques :





L'objectif est désormais de montrer le résultat suivant, mais sans fournir explicitement un graphe correspondant à la propriété de k-extension.

1.4. Proposition. *Soit k un entier naturel non nul. Alors, il existe un graphe fini vérifiant la propriété de k-extension.*

Schématiquement, la démonstration s'appuie sur la remarque informelle suivante : si une variable aléatoire associée au modèle binomial $\mathscr{G}_{n,p}$ mesure un paramètre et admet une petite espérance, cela signifie qu'il y a des réalisations du graphe aléatoire pour lesquelles la valeur de cette variable est petite, donc qu'il existe des graphes correspondant à une petite valeur du paramètre.

Plus précisément, voici un lemme élémentaire qui traduit cette remarque.

1.5. Lemme. *Soit N une variable aléatoire à valeurs dans $\mathbb{N}$ telle que $\mathbf{E}(N) < 1$. Alors, $\mathbf{P}(N = 0) > 0$.*

Démonstration. Le résultat découle directement de l'inégalité suivante

$$\begin{aligned}\mathbf{E}(N) &= \sum_{k\geqslant 1} k\mathbf{P}(N = k)\\ &\geqslant \sum_{k\geqslant 1} \mathbf{P}(N = k) = 1 - \mathbf{P}(N = 0).\end{aligned}$$

♢

Revenons à la preuve de la proposition IV-1.4.

Démonstration. Soit n un entier naturel non nul et X un graphe aléatoire suivant le modèle binomial de paramètres n et $\frac{1}{2}$.

Notons N la variable aléatoire de comptage des couples (U, V) de parties de $[\![n]\!]$, disjointes et de cardinal k telles qu'il n'existe aucun sommet $s \notin U \cup V$ relié dans X à tous les sommets de U et à aucun de V, c'est-à-dire des « contre-exemples » à la propriété de k-extension. L'événement $N = 0$ correspond donc à « X vérifie la propriété de k-extension ».

Pour tout couple (U, V) de parties disjointes et de cardinal k, un des $n - 2k$ sommets hors de $U \cup V$ est relié dans X à tous les sommets de U et aucun de V avec probabilité 2^{-2k} (car cela impose des conditions, absence ou présence, sur $2k$ arêtes). La probabilité qu'il n'existe pas de tel sommet, c'est-à-dire que (U, V) soit un contre-exemple, est donc, en utilisant l'indépendance des arêtes joignant aux $n - 2k$ sommets hors de $U \cup V$,

$$\left(1 - 2^{-2k}\right)^{n-2k}.$$

Par linéarité de l'espérance, avec l'écriture suivante comme somme de variables indicatrices,

$$N = \sum_{U \in \mathscr{P}_k([\![n]\!])} \sum_{V \in \mathscr{P}_k([\![n]\!] \setminus U)} \mathbb{1}_{(U,V) \text{ contre-exemple}},$$

on obtient

$$\begin{aligned} \mathbf{E}(N) &= \tbinom{n}{k}\tbinom{n-k}{k}\left(1 - 2^{-2k}\right)^{n-2k} \\ &\leqslant n^k (n-k)^k \left(1 - 2^{-2k}\right)^{n-2k} \\ &\leqslant n^{2k} \left(1 - 2^{-2k}\right)^{n-2k}. \end{aligned}$$

Lorsque $n \to +\infty$, ce majorant tend vers 0 donc, pour n assez grand, $\mathbf{E}(N) < 1$. Avec le lemme précédent, on en déduit que (toujours pour n assez grand) $\mathbf{P}(N = 0) > 0$, et donc qu'il existe au moins un graphe à n sommets satisfaisant la propriété de k-extension. ♢

1.6. Remarque. Comme toutes les démonstrations de ce chapitre, celle-ci n'est pas constructive.

Au mieux, en étudiant précisément l'inégalité

$$n^{2k}\left(1-2^{-2k}\right)^{n-2k}<1,$$

on peut obtenir une valeur de n (pas forcément la plus petite) pour laquelle il existe un graphe à n sommets vérifiant la propriété de k-extension. Il reste toutefois beaucoup de travail pour exhiber parmi tous les graphes à n sommets un graphe qui convient.

2. Nombres de Ramsey

2.1. Proposition. *Soit k un entier supérieur ou égal à 3 et n un entier tels que*

$$\binom{n}{k}<2^{\binom{k}{2}-1}.$$

Alors, il existe un graphe $G=(S,A)$ à n sommets tel que, dans tout ensemble de k sommets de S, on peut en trouver à la fois deux reliés par une arête et deux non reliés par une arête.

En d'autres termes, ce graphe n'admet pas de graphe induit à k sommets isomorphe au graphe complet ou au graphe vide.

Démonstration. Soit X un graphe aléatoire qui suit le modèle binomial de paramètres n et $\frac{1}{2}$.
Considérons N la variable aléatoire indiquant le nombre d'occurrences de K_k dans X (c'est-à-dire le nombre de sous-graphes de X isomorphes au graphe complet K_k), et N' la variable aléatoire indiquant le nombre d'occurrences de K_k dans le graphe complémentaire de X qui suit également le modèle binomial de paramètres n et $\frac{1}{2}$. D'après les calculs en moyenne du chapitre précédent, les espérances de N et N' coïncident et valent

$$\mathbf{E}(N)=\mathbf{E}(N')=\binom{n}{k}2^{-\binom{k}{2}}.$$

En particulier, avec l'inégalité en hypothèse, $\mathbf{E}(N+N')<1$ puis avec le lemme IV-1.5, $\mathbf{P}(N+N'=0)>0$: ainsi, il existe au moins une réalisation du graphe X pour laquelle $N+N'=0$, c'est-à-dire pour laquelle $N=N'=0$. En conclusion, il existe bien un graphe satisfaisant la condition de l'énoncé. ♢

2.2. Remarque. Un célèbre résultat de la théorie de Ramsey indique que, pour tout k, il existe un entier $R(k,k)$ tel que tout graphe avec au moins $R(k,k)$ sommets contienne soit k sommets tous reliés entre eux deux à deux, soit k sommets entre lesquels il n'existe aucune arête. Toutefois la démonstration de ce résultat n'est pas constructive et l'on ne connaît que quelques valeurs : $R(3,3) = 6$, $R(4,4) = 18$, les suivantes demeurent inconnues. Par ailleurs, il est peu envisageable de calculer une valeur comme $R(6,6)$ en un temps raisonnable, même en utilisant d'immenses ressources informatiques, car il y a bien trop de graphes à tester. Faute de pouvoir les calculer, on cherche donc à encadrer précisément ces quantités.

Le résultat précédent fournit donc une minoration de la forme $R(k,k) > A$ en indiquant un graphe à A sommets ne vérifiant pas la propriété recherchée. Les meilleurs encadrements actuels pour les valeurs suivantes sont $R(5,5) \in [\![43,46]\!]$, $R(6,6) \in [\![102,160]\!]$...

Plus précisément, avec un peu de calculs et de ténacité, on obtient le résultat suivant.

2.3. Corollaire. *Pour tout entier $k \geqslant 3$,*

$$R(k,k) > \tfrac{1}{\sqrt{2}e} 2^{-1/k} k 2^{k/2}.$$

Démonstration. Avec la proposition précédente, il suffit d'établir que l'entier $n = \frac{1}{\sqrt{2}e} 2^{-1/k} k 2^{k/2}$ vérifie l'inégalité en hypothèse. Or,

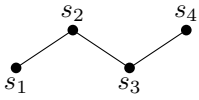
$$\begin{aligned}\tbinom{n}{k} 2^{-\binom{k}{2}+1} &< \tfrac{n^k}{k!} 2^{-\binom{k}{2}+1} \\ &< e^{-k} 2^{-k/2-1} k^k 2^{k^2/2} \tfrac{e^k}{k^k} 2^{-\binom{k}{2}+1} \\ &< 1,\end{aligned}$$

en utilisant la minoration $k! > \left(\frac{k}{e}\right)^k$. ♢

2.4. Remarque. On peut préférer la minoration plus maniable : pour tout entier $k \geqslant 3$, $R(k,k) > k 2^{k/2-2}$.

3. Nombre de chemins de longueur 3

3.1. Définition. Soit $G = (S, A)$ un graphe. Un chemin géométrique de longueur 3 est un ensemble de quatre sommets distincts s_1, s_2, s_3 et $s_4 \in S$ tels que $\{s_1, s_2\}$, $\{s_2, s_3\}$ et $\{s_3, s_4\} \in A$.



Dans cette notion de chemin « géométrique », on choisit d'exclure les cycles et de ne pas considérer l'orientation du chemin.

Commençons par démontrer un résultat (tout à fait déterministe) sur les graphes concernant les chemins géométriques de longueur 3.

3.2. Proposition. *Soit $n \geqslant 3$. Dans un graphe à n sommets et n arêtes qui n'est pas réunion de graphes C_3 disjoints, il existe (au moins) un chemin géométrique de longueur* 3.

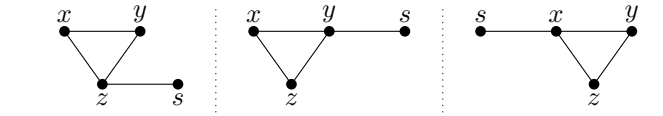
Démonstration. Soit G un graphe avec n sommets et n arêtes.

Quitte à enlever les composantes connexes de G isomorphes à C_3 (ce qui ne change pas l'égalité entre le nombre d'arêtes et le nombre de sommets), on peut supposer que G ne comporte pas de composantes connexes isomorphes à C_3.

Ensuite, G admet (au moins) une composante connexe avec un nombre d'arêtes supérieur ou égal à son nombre de sommets : on se limite désormais à celle-ci.

Considérons un arbre couvrant de cette composante connexe. Le nombre d'arêtes de cette composante connexe est supérieur ou égal à son nombre de sommets, donc on dispose d'une arête $\{x, y\}$ de la composante qui n'appartienne pas à l'arbre couvrant. Dans l'arbre couvrant, il existe un unique chemin joignant x à y (car un arbre est connexe ce qui entraîne l'existence, et sans cycle ce qui entraîne l'unicité) et il est de longueur au moins 2 (car l'arête $\{x, y\}$ n'appartient pas à l'arbre couvrant.

- ▷ Si ce chemin $x, x_1, x_2, \ldots, x_N = y$ est de longueur au moins trois : on obtient un chemin géométrique de longueur 3 avec les sommets x, y, x_1, x_2 et les arêtes $\{y, x\}$, $\{x, x_1\}$, $\{x_1, x_2\}$.
- ▷ Sinon, il est de longueur 2 et il existe un sommet z dans l'arbre tel que $\{x, y\}$, $\{z, y\}$ et $\{x, y\}$ sont des arêtes de la composante connexe : comme celle-ci n'est pas isomorphe à C_3, un autre sommet s est relié à l'un des trois sommets déjà considérés et l'on obtient facilement un chemin géométrique de longueur 3 en discutant sur ce point de rattachement comme indiqué sur le schéma suivant :



$\{s, z\}, \{z, x\}, \{x, y\}$ $\{s, y\}, \{y, x\}, \{x, z\}$ $\{s, x\}, \{x, y\}, \{y, z\}$

♢

On remarque que la construction précédente donne un chemin géométrique de longueur 3 pour chaque arête ajoutée à l'arbre couvrant choisi (en fait la forêt si l'on prend un arbre couvrant pour chaque composante connexe). En considérant les différentes arêtes hors de l'arbre couvrant, on obtient le résultat suivant plus général.

3.3. Corollaire. *Soit $m \geqslant n \geqslant 3$. Dans un graphe à n sommets et m arêtes, il existe au moins $m - n$ chemins géométriques de longueur* 3.

Passons maintenant à un résultat obtenu par la méthode probabiliste.

3.4. Proposition. *Soit $n \geqslant 3$ et m tels que $2m > 3n$. Dans un graphe à n sommets et m arêtes, il existe au moins $\frac{4m^3}{27n^2}$ chemins géométriques de longueur* 3.

Démonstration. Soit G un tel graphe ; notons $S = \{s_1, \dots, s_n\}$ son ensemble de sommets et c le nombre de chemins géométriques de longueur 3 de G.

Soit $p \in]0,1[$ (la valeur sera fixée plus loin), $X_1, \dots, X_n$ des variables aléatoires mutuellement indépendantes suivant la loi de Bernoulli de paramètre p. Considérons l'ensemble (aléatoire) de sommets

$$S' = \{s_i,\ X_i = 1\},$$

puis le graphe aléatoire induit $G' = G[S']$.

Notons N, M et C les variables aléatoires indiquant le nombre de sommets, d'arêtes et de chemins géométriques de longueur 3 de G'. D'après le corollaire précédent, $C \geqslant M - N$ puis, par croissance de l'espérance, $\mathbf{E}(C) \geqslant \mathbf{E}(M) - \mathbf{E}(N)$.

Par construction, $\mathbf{E}(N) = np$ et $\mathbf{E}(M) = mp^2$ (car il faut conserver les deux sommets aux extrémités d'une arête pour conserver l'arête). Par ailleurs, on obtient C en comptant les chemins géométriques de longueur 3 conservés, d'où $\mathbf{E}(C) = cp^4$ (il faut conserver les quatres sommets d'un chemin géométrique de longueur 3 pour le conserver).

Combinant tous ces résultats, on obtient l'inégalité

$$cp^4 \geqslant mp^2 - np.$$

En étudiant la fonction $x \mapsto mx^2 - nx^3$ sur $]1, +\infty[$ (la variable x correspond à $\frac{1}{p}$ dans l'inégalité), on trouve que cette fonction est maximale pour $x = \frac{2m}{3n}$ (qui est bien strictement plus grand que 1 d'après l'hypothèse).

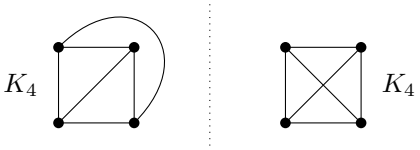
Ainsi, en fixant $p = \frac{3n}{2m}$, on obtient la minoration attendue pour le nombre de chemins $c \geqslant \frac{4m^3}{27n^2}$. ♢

4. Nombre de croisements

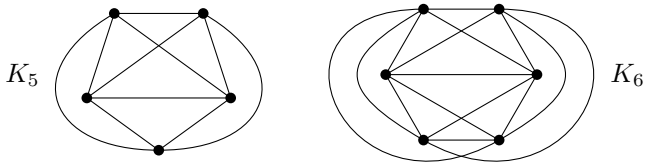
4.1. Définition. Le nombre de croisements d'un graphe G est le nombre minimal d'intersections de paires d'arêtes dans un dessin de G.

Un graphe G est planaire si son nombre de croisements est nul.

4.2. Exemple. Pour le graphe K_4 complet à quatre sommets, le nombre de croisements est 0 (donc K_4 est planaire), comme on peut le voir avec le dessin gauche (alors que ce n'est pas apparent sur celui, plus classique, de droite).



Le nombre de croisements de K_5 est 1, de K_6 est 3 et voici des dessins de ces graphes avec des nombres de croisements minimaux :



La formule suivante

$$\tfrac{1}{4}\left\lfloor \tfrac{n}{2} \right\rfloor \left\lfloor \tfrac{n-1}{2} \right\rfloor \left\lfloor \tfrac{n-2}{2} \right\rfloor \left\lfloor \tfrac{n-3}{2} \right\rfloor$$

pour le nombre de croisements de K_n (vérifiée pour de petites valeurs de n et approchée asymptotiquement) constitue la conjecture de Hill [1].

On établit classiquement le résultat suivant par récurrence sur le nombre d'arêtes (en s'appuyant sur la formule d'Euler qui indique que pour un graphe sans croisement, $s - a + f = 2$ où s, a et f désignent respectivement le nombre de sommets, d'arêtes et de faces du graphe).

1. Anthony Hill (1930-2020) est un artiste anglais du mouvement constructiviste, qui avoue une grande fascination pour les mathématiques. Il a été élu membre de la *London Mathematical Society* pour sa contribution à la théorie des graphes.

4.3. Proposition. *Soit $m \geqslant n \geqslant 3$. Pour un graphe à n sommets et m arêtes, le nombre de croisements est supérieur ou égal à*

$$m - 3n + 6.$$

Passons maintenant à un résultat obtenu à partir de cette proposition par la méthode probabiliste.

4.4. Proposition. *Soit $n \geqslant 3$ et m tels que $2m > 9n$. Dans un graphe à n sommets et m arêtes, le nombre de croisements est supérieur ou égal à $\frac{4m^3}{243n^2}$.*

Démonstration. Soit G un tel graphe ; notons $S = \{s_1, \ldots, s_n\}$ son ensemble de sommets et c le nombre de croisements de G.

Soit $p \in]0,1[$ (la valeur sera fixée plus loin), $X_1, \ldots, X_n$ des variables aléatoires mutuellement indépendantes suivant la loi de Bernoulli de paramètre p. Considérons l'ensemble (aléatoire) de sommets

$$S' = \{s_i,\ X_i = 1\},$$

puis le graphe aléatoire induit $G' = G[S']$.

Notons N, M et C les variables aléatoires indiquant le nombre de sommets, d'arêtes et de croisements de G'. D'après la proposition précédente, $C \geqslant M - 3N + 6$ puis, par croissance de l'espérance,

$$\mathbf{E}(C) \geqslant \mathbf{E}(M) - 3\mathbf{E}(N) + 6 \geqslant \mathbf{E}(M) - 3\mathbf{E}(N).$$

Par construction, $\mathbf{E}(N) = np$ et $\mathbf{E}(M) = mp^2$. Par ailleurs, en partant d'un dessin de G avec le nombre minimal de croisements, on obtient un majorant de C en comptant les croisements conservés sur ce dessin, d'où $\mathbf{E}(C) \leqslant cp^4$ (il faut conserver les quatres sommets aux extrémités des deux arêtes qui se croisent pour conserver un croisement).

Combinant tous ces résultats, on obtient l'inégalité

$$cp^4 \geqslant mp^2 - 3np.$$

En étudiant la fonction $x \mapsto mx^2 - 3nx^3$ sur $]1, +\infty[$ (la variable x correspond encore à $\frac{1}{p}$ dans l'inégalité), on trouve que cette fonction est maximale pour $x = \frac{2m}{9n}$.

Fixons alors le paramètre p à la valeur $\frac{9n}{2m}$ (qui est bien strictement plus petite que 1 d'après l'hypothèse) : en reportant on obtient la minoration $c \geqslant \frac{4m^3}{243n^2}$. ♢

4.5. Remarque. Ce résultat est quelquefois présenté avec l'hypothèse plus faible $m \geqslant 4n$. Par la même méthode, on obtient toujours l'inégalité $cp^4 \geqslant mp^2 - 3np$ et, en prenant cette fois-ci $p = \frac{4n}{m}$ (qui est bien strictement plus petit que 1, d'après la nouvelle hypothèse), on obtient la minoration plus faible $c \geqslant \frac{m^3}{64n^2}$.

5. Exercices

5.1. Exercice. *Soit x_1, …, $x_n \in \mathbb{R}$ non tous nuls tels que $\sum_{k=1}^n x_n = 0$. Montrer qu'il existe $\sigma \in S_n$ tel que $\sum_{k=1}^n x_{\sigma(k)} x_{\sigma(k+1)} < 0$ (avec la convention $\sigma(n+1) = \sigma(1)$).*

Solution. Soit Σ une variable aléatoire de loi uniforme sur S_n. Considérons la variable aléatoire réelle $S = \sum_{k=1}^n x_{\Sigma(k)} x_{\Sigma(k+1)}$. Par linéarité de l'espérance,

$$\begin{aligned}
\mathbf{E}(S) &= \sum_{k=1}^{n} \mathbf{E}(x_{\Sigma(k)} x_{\Sigma(k+1)}) \\
&= n\mathbf{E}(x_{\Sigma(1)} x_{\Sigma(2)}) \\
&= n \tfrac{1}{n(n-1)} \sum_{i=1}^{n} \sum_{j \neq i} x_i x_j \\
&= \tfrac{1}{n-1} \Big(\Big(\sum_{i=1}^{n} x_i \Big)^2 - \sum_{i=1}^{n} x_i^2 \Big) \\
&= -\tfrac{1}{n-1} \sum_{i=1}^{n} x_i^2 \\
&< 0.
\end{aligned}$$

Ainsi, il existe une réalisation $S(\omega) < 0$; en notant $\sigma = \Sigma(\omega)$, on obtient le résultat attendu.

5.2. Exercice. *Soit E un espace préhilbertien, $v_1, \dots, v_n$ des vecteurs unitaires de E et $X_1, \dots, X_n$ des variables aléatoires mutuellement indépendantes, de même loi donnée par*

$$\mathbf{P}(X_1 = 1) = \mathbf{P}(X_1 = -1) = \tfrac{1}{2}.$$

Considérons la variable aléatoire $U = X_1 v_1 + \cdots + X_n v_n$.

1. *Calculer l'espérance de la variable aléatoire réelle $\|U\|^2$. En déduire qu'il existe une famille $(\varepsilon_1, \dots, \varepsilon_n) \in \{-1, 1\}^n$ telle que*

$$\|\varepsilon_1 v_1 + \cdots + \varepsilon_n v_n\| \leqslant \sqrt{n}.$$

2. *Montrer que la famille $(v_1, \dots, v_n)$ est orthogonale si, et seulement si, pour toute famille $(\varepsilon_1, \dots, \varepsilon_n) \in \{-1, 1\}^n$,*

$$\|\varepsilon_1 v_1 + \cdots + \varepsilon_n v_n\| = \sqrt{n}.$$

3. *Montrer que si la famille $(v_1, \dots, v_n)$ n'est pas orthogonale, il existe une famille $(\varepsilon_1, \dots, \varepsilon_n) \in \{-1, 1\}^n$ telles que*

$$\|\varepsilon_1 v_1 + \cdots + \varepsilon_n v_n\| > \sqrt{n}.$$

Solution.

1. ▷ Commençons par récrire la variable $\|U\|^2$ en exploitant la bilinéarité du produit scalaire :

$$\begin{aligned}\|U\|^2 &= \Big\langle \sum_{i=1}^n X_i v_i \Big| \sum_{j=1}^n X_j v_j \Big\rangle \\ &= \sum_{i=1}^n \sum_{j=1}^n \langle v_i | v_j \rangle X_i X_j.\end{aligned}$$

Appliquons ensuite la linéarité de l'espérance

$$\mathbf{E}\big(\|U\|^2\big) = \sum_{i=1}^n \sum_{j=1}^n \langle v_i | v_j \rangle \mathbf{E}(X_i X_j).$$

Notons enfin que, pour des entiers distincts i et j, on obtient

$$\mathbf{E}(X_i X_j) = \mathbf{E}(X_i)\mathbf{E}(X_j) = 0$$

(car les variables sont indépendantes et centrées) et que, pour tout entier i, $\mathbf{E}(X_i^2) = \mathbf{E}(1) = 1$. En conclusion,

$$\mathbf{E}\big(\|U\|^2\big) = \sum_{i=1}^n \|v_i\|^2 = n.$$

▷ La variable $\|U\|^2$ est d'espérance n, donc il existe une réalisation donnant une valeur inférieure ou égale à n, c'est-à-dire une réalisation $\varepsilon_1, \dots, \varepsilon_n$ des variables $X_1, \dots, X_n$ telle que

$$\|\varepsilon_1 v_1 + \cdots + \varepsilon_n v_n\|^2 \leqslant n.$$

Il suffit de passer à la racine carrée pour conclure.

2. ($\Rightarrow$) Il s'agit d'une simple application du théorème de Pythagore à la famille orthogonale $(\varepsilon_1 v_1, \dots, \varepsilon_n v_n)$:

$$\begin{aligned}\|\varepsilon_1 v_1 + \cdots + \varepsilon_n v_n\|^2 &= \|\varepsilon_1 v_1\|^2 + \cdots + \|\varepsilon_n v_n\|^2 \\ &= \|v_1\|^2 + \cdots + \|v_n\|^2 = n,\end{aligned}$$

car les vecteurs $v_1, \dots, v_n$ sont unitaires.

($\Leftarrow$) Soit $i \neq j$ et $v = \sum_{k \notin \{i,j\}} v_k$. Alors, d'après une identité remarquable,

$$\begin{aligned}4\langle v_i | v_j + v\rangle &= \|v_i + v_j + v\|^2 - \|v_i - v_j - v\|^2 \\ &= n - n \\ &= 0,\end{aligned}$$

$$\begin{aligned}4\langle v_i | v_j - v\rangle &= \|v_i + v_j - v\|^2 - \|v_i - v_j + v\|^2 \\ &= n - n \\ &= 0.\end{aligned}$$

D'où, en additionnant ces deux résultats, $8\langle v_i | v_j\rangle = 0$.

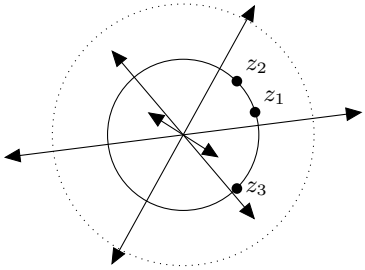
Par conséquent, la famille $(v_1, \dots, v_n)$ est orthogonale.

3. D'après la question précédente, la variable $\|U\|^2$ n'est pas la constante égale à $n = \mathbf{E}\big(\|U\|^2\big)$. Ainsi, il existe au moins une réalisation strictement supérieure à $n = \mathbf{E}\big(\|U\|^2\big)$ (et une réalisation strictement inférieure).

Dans le cas particulier des nombres complexes, la dernière question se récrit : pour $n \geqslant 3$ nombres complexes $z_1, \dots, z_n$ de module 1, il existe des signes $\varepsilon_1, \dots, \varepsilon_n$ tels que

$$|\varepsilon_1 z_1 + \cdots + \varepsilon_n z_n| > \sqrt{n}.$$

On pourra tester ce résultat en cherchant des signes correspondants à cette inégalité pour trois complexes de module 1.



5.3. Exercice. *Soit $n \leqslant m$ des entiers naturels non nuls, $k \in [\![n]\!]$.*

1. *Montrer qu'il existe un coloriage des arêtes du graphe complet K_n avec deux couleurs tel que l'on dénombre au plus*

$$\binom{n}{k} 2^{1-\binom{n}{2}}$$

graphes induits isomorphes à K_k dont toutes les arêtes sont de la même couleur.

2. *Montrer qu'il existe un coloriage avec deux couleurs des arêtes du graphe biparti complet $K_{n,m}$ tel que l'on dénombre au plus*

$$\binom{n}{k}\binom{m}{k} 2^{1-k^2}$$

graphes induits isomorphes à $K_{k,k}$ dont toutes les arêtes sont de la même couleur.

Solution. La première question n'est qu'une récriture de la deuxième section de ce chapitre.

1. Colorions chaque arête de K_n avec probabilité $\frac{1}{2}$ en bleu ou en rouge (indépendamment les unes des autres). Notons N (respectivement N') la variable aléatoire égale au nombre de graphes induits isomorphes à K_k dont toutes les arêtes sont coloriées en bleu (respectivement en rouge). Alors, comme dans la première section,

$$\mathbf{E}(N) = \mathbf{E}(N') = \binom{n}{k} 2^{-\binom{n}{2}}.$$

Ainsi, $\mathbf{E}(N + N') = \binom{n}{k} 2^{1-\binom{n}{2}}$. Il existe une réalisation de ce coloriage aléatoire pour laquelle la variable aléatoire $N+N'$, qui dénombre le nombre total de graphes induits isomorphes à K_k monochromes, est inférieure ou égale à $\mathbf{E}(N + N')$. Donc, il existe (au moins) un coloriage vérifiant la propriété recherchée.

2. Colorions chaque arête de $K_{n,m}$ avec probabilité $\frac{1}{2}$ en bleu ou en rouge (indépendamment les unes des autres). Notons N (respectivement N') la variable aléatoire égale au nombre de graphes induits isomorphes à $K_{k,k}$ dont toutes les arêtes sont coloriées en bleu (respectivement en rouge). Sachant que $K_{n,m}$ comporte $\binom{n}{k}\binom{m}{k}$ graphes induits isomorphes à $K_{k,k}$ (il suffit de choisir les k sommets dans chaque partie) et que chacun de ces graphes induits est entièrement colorié en bleu (ou entièrement en rouge) avec probabilité 2^{-k^2} (puisqu'il y a k^2 arêtes à colorier indépendamment), on obtient

$$\mathbf{E}(N) = \mathbf{E}(N') = \binom{n}{k}\binom{m}{k}2^{-k^2}.$$

Ainsi, $\mathbf{E}(N+N') = \binom{n}{k}\binom{m}{k}2^{1-k^2}$. Par conséquent, il existe une réalisation de ce coloriage aléatoire pour laquelle la variable aléatoire $N+N'$ est inférieure ou égale à $\mathbf{E}(N+N')$. Donc, il existe (au moins) un coloriage vérifiant la propriété recherchée.

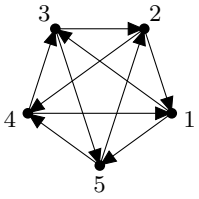
5.4. Exercice (Théorème de Szele [Sze43]). *Soit n un entier naturel non nul.*

Une orientation du graphe complet K_n est un ensemble

$$A = \{(x,y) \in [\![n]\!]^2,\ x \neq y\}$$

tel que, pour tous $x \neq y$, A contient (x,y) ou (y,x) mais pas les deux[2]*. On représente la partie A sur le dessin de K_n par des flèches. Voici, par exemple, un dessin de K_5 avec l'orientation*

$$A = \{(1,3),(1,5),(2,1),(2,4),(3,2),(3,5),(4,1),(4,3),(5,2),(5,4)\}.$$



2. Un graphe complet muni d'une orientation est souvent appelé « tournoi ».

Un chemin hamiltonien relatif à A est un chemin $(x_1, x_2, \dots, x_n)$ passant par tous les sommets de $[\![n]\!]$ et tel que, pour tout $k < n$, l'arête orienté (x_k, x_{k+1}) appartienne à A.

Dans l'exemple ci-dessus d'orientation de K_5, le chemin $(1, 3, 5, 2, 4)$ est hamiltonien, mais $(1, 3, 5, 4, 2)$ ne l'est pas (car $(4, 2) \notin A$).

Montrer qu'il existe une orientation de K_n qui admet un nombre de chemins hamiltoniens supérieur ou égal à $n!2^{1-n}$.

Solution. Considérons une orientation aléatoire A du graphe K_n où chaque arête prend l'une des deux orientations avec probabilité $\frac{1}{2}$ indépendamment des autres.

Un chemin de longueur n passant par tous les sommets respectent les orientations de A avec probabilité 2^{1-n} puisqu'il comporte $n-1$ arêtes.

La variable aléatoire comptant les chemins hamiltoniens relatif à A est la somme des indicatrices sur tous les $n!$ chemins injectifs de longueur n de l'événement « ce chemin respecte A ». La linéarité de l'espérance donne alors

$$\mathbf{E}(N) = n!2^{1-n}.$$

Par conséquent, il existe une réalisation de l'orientation aléatoire A pour laquelle la variable aléatoire N est supérieure ou égale à $n!2^{-1-n}$, ce qui était attendu.

5.5. Exercice. *Soit $G = (S, A)$ un graphe à m arêtes. Considérons l'ensemble aléatoire S' obtenu en conservant chaque sommet de G avec probabilité $\frac{1}{2}$ (indépendamment des autres) et l'ensemble aléatoire A' des arêtes de G joignant un sommet de S' à un sommet de $S \setminus S'$.*

1. *Montrer que $\mathbf{E}(|A'|) = \frac{1}{2}m$.*
2. *En déduire qu'il existe une partition $S = S_1 \cup S_2$ tel qu'il y ait au moins $\frac{m}{2}$ arêtes joignant un sommet de S_1 à un sommet de S_2.*

Solution.

1. Considérons une arête $\{x, y\} \in A$. Alors,

$$\begin{aligned}\mathbf{P}\big(\{x, y\} \in A'\big) &= \mathbf{P}(x \in S', y \notin S') + \mathbf{P}(x \notin S', y \in S') \\ &= \tfrac{1}{2}\tfrac{1}{2} + \tfrac{1}{2}\tfrac{1}{2} = \tfrac{1}{2}.\end{aligned}$$

   Ensuite, en décomposant $|A'|$ comme somme des variables indi-

catrices $\mathbb{1}_{\{x,y\}\in A'}$ puis en utilisant la linéarité de l'espérance,

$$\mathbf{E}(|A'|) = \sum_{\{x,y\}\in A} \mathbf{P}\big(\{x,y\} \in A'\big) = \tfrac{m}{2}.$$

2. Considérons une réalisation de l'expérience aléatoire précédente telle que $|A'| \geqslant \frac{m}{2}$; notons S_1 la réalisation correspondante de S' et S_2 son complémentaire dans S. Par construction, les arêtes joignant les sommets de S_1 aux sommets de S_2 sont exactement celles de A', d'où le résultat.

En d'autres termes, on vient d'établir qu'il existe un sous-graphe de G qui est biparti et qui conserve au moins la moitié des arêtes. On précise ce résultat dans l'exercice suivant.

5.6. Exercice. *Soit $G = (S, A)$ un graphe tel que $|S| = 2n$ et $|A| = m$. Montrer qu'il existe une partie $S' \subset S$ de cardinal n telle que le nombre d'arêtes joignant un sommet de S' à un sommet de $S\backslash S'$ soit strictement plus grand que $\frac{m}{2}$.*

Solution. Soit T une variable aléatoire de loi uniforme sur $\mathscr{P}_n(S)$ et N la variable aléatoire de comptage des arêtes joignant T à $S \setminus T$. Alors,

$$N = \sum_{a\in A} \mathbb{1}_{|a\,\cap\, T|=1}.$$

Ainsi, par linéarité de l'espérance,

$$\begin{aligned}\mathbf{E}(N) &= \sum_{a\in A} \mathbf{P}(|a \cap T| = 1)\\ &= \sum_{a\in A} \tfrac{2\binom{2n-2}{n-1}}{\binom{2n}{n}}.\end{aligned}$$

En effet, une partie de cardinal n ne rencontrant l'arête a qu'en un point est uniquement déterminée par

- ▷ l'extrémité de a qu'elle contient, soit deux choix,
- ▷ les $n-1$ autres éléments qui n'appartiennent pas à l'arête, soit $\binom{2n-2}{n-1}$ choix.

Ainsi,

$$\mathbf{E}(N) = \sum_{a\in A} \tfrac{n}{2n-1} = m\tfrac{n}{2n-1} > \tfrac{m}{2}.$$

En conclusion, il existe une réalisation de la variable T telle que $N > \frac{m}{2}$, et donc un graphe vérifiant la conclusion de l'énoncé.

L'inégalité n'est par exemple pas optimale pour des graphes comme le graphe cycle C_{2n} ou le graphe complet K_{2n}.
En revanche, elle l'est pour le graphe à six sommets composé de deux cycles d'indice 3 disjoints : trois sommets qui n'appartiennent pas à la même composante connexe (entourés sur le dessin suivant) sont reliés par quatre arêtes aux trois autres sommets

5.7. Exercice. *Soit n, d des entiers naturels non nuls et un graphe $G = (S, A)$ à n sommets tous de degré au moins d.*

Montrer qu'il existe un ensemble dominant de cardinal inférieur ou égal à

$$n\frac{\ln(d+1)+1}{d+1}.$$

On pourra remarquer que $\frac{\ln(d+1)}{d+1} \in]0,1[$, et que cette quantité peut donc être choisie comme probabilité d'apparition pour les sommets d'un graphe aléatoire.

Solution. Considérons le graphe aléatoire déduit de G en conservant chaque sommet (indépendamment des autres) avec probabilité

$$p = \frac{\ln(d+1)}{d+1}.$$

Notons S' l'ensemble aléatoire des sommets conservés et S'' l'ensemble des sommets à la fois non conservés et non voisins d'au moins un sommet de S'. La variable aléatoire N' égale au cardinal de S' est d'espérance np. Considérons la variable aléatoire N'' égale au cardinal de S'' et écrivons

$$N'' = \sum_{x \in S} \mathbb{1}_{x \in S''}.$$

Or, pour tout $x \in S$, $\mathbf{P}(x \in S'')$ est la probabilité que ni x, ni l'un de ses au moins d voisins ne soient conservés. Ainsi,

$$\mathbf{P}(x \in S'') = (1-p)(1-p)^{\deg x} \leqslant (1-p)^{d+1}.$$

Alors, par linéarité de l'espérance,

$$\mathbf{E}(N'') = \sum_{x \in S} \mathbf{P}(x \in S'') \leqslant n(1-p)^{d+1}.$$

Utilisons ensuite la majoration élémentaire

$$(1-p)^{d+1} \leqslant \exp(-p(d+1)) = \tfrac{1}{d+1}$$

pour obtenir une majoration de $\mathbf{E}(N'')$, d'où l'on déduit :

$$\begin{aligned}\mathbf{E}(N'+N'') &\leqslant np + n(1-p)^{d+1} \\ &\leqslant n\tfrac{\ln(d+1)}{d+1} + n\tfrac{1}{d+1}.\end{aligned}$$

Il existe donc une réalisation (c'est-à-dire un choix de sommets) pour laquelle $N'+N'' \leqslant n\frac{\ln(d+1)+1}{d+1}$.

Remarquons que l'ensemble $S' \cup S''$ est un ensemble dominant du graphe (on a ajouté à S' tous les sommets qui n'étaient pas voisins de S') et, pour cette réalisation, son cardinal vérifie la majoration souhaitée.

5.8. Exercice. *Soit n, k des entiers naturels non nuls.*

Posons $m = kn$ et considérons un graphe G admettant n sommets et m arêtes.

1. *Montrer qu'il existe un graphe induit sur G tel que la différence entre le nombre de sommets et son nombre d'arêtes est supérieure ou égale à $\frac{n}{4k}$.*
2. *Montrer qu'il existe un ensemble indépendant de cardinal $\frac{n}{4k}$.*

Solution.

1. Soit $p \in]0,1[$ (la valeur sera fixée plus loin). Considérons le graphe aléatoire déduit de G en conservant chaque sommet avec probabilité p (indépendamment des autres). Notons N et M les variables aléatoires égales au nombre de sommets et d'arêtes de ce nouveau graphe. Alors, $\mathbf{E}(N) = np$ et $\mathbf{E}(M) = mp^2$ donc, par linéarité, $\mathbf{E}(N-M) = np - mp^2$.

   Choisissons alors $p = \frac{n}{2m}$ de sorte à obtenir

   $$\mathbf{E}(N-M) = \tfrac{n^2}{4m} = \tfrac{n}{4k}.$$

   Il existe donc une réalisation du graphe aléatoire (induit de G) telle que la différence entre le nombre de ses sommets et le nombre de ses arêtes est supérieure ou égale à $\mathbf{E}(N-M)$ (car l'espérance est une « moyenne » des valeurs prises).

2. Notons n' et m' le nombre de sommets et le nombre d'arêtes du graphe obtenu à la question précédente. Alors,

$$n' \geqslant m' + \tfrac{n}{4k}.$$

En enlevant une extrémité de chacune des m' arêtes, on obtient un ensemble de $n' - m'$ (supérieur à $\frac{n}{4k}$) sommets entre lesquels il n'y a aucune arête.

5.9. Exercice. *Soit k un entier supérieur ou égal à 3 et $\varepsilon \in]0, \frac{1}{k}[$. Notons, pour tout entier n non nul, $p_n = n^{\varepsilon - 1}$ et considérons le graphe aléatoire de paramètres n et p_n.*

1. *Soit N_n le nombre de cycles de longueur inférieure ou égale à k dans le graphe aléatoire de paramètres n et p_n. Montrer que*

$$\mathbf{P}(N_n \geqslant \tfrac{n}{2}) \leqslant (k-2)n^{\varepsilon k - 1}.$$

2. *Soit α_n le cardinal maximal d'un ensemble indépendant du graphe aléatoire de paramètres n et p_n. Montrer que*

$$\mathbf{P}(\alpha_n \geqslant \tfrac{n}{2k}) \leqslant \tbinom{n}{\lceil n/(2k) \rceil} (1 - n^{\varepsilon - 1})^{n(n-2k)/(8k^2)}.$$

3. *En déduire qu'il existe un graphe de maille et de nombre chromatique strictement supérieurs à k.*
On pourra considérer n de sorte à avoir $\mathbf{P}(N \geqslant \frac{n}{2}) < \frac{1}{2}$ et $\mathbf{P}(\alpha \geqslant \frac{n}{2k}) < \frac{1}{2}$, puis utiliser le résultat I-7.6.

Solution.

1. Pour tout l entre 3 et k, il existe au plus $\frac{1}{2}\binom{n}{l}(l-1)!$ cycles de longueur l (on choisit les l sommets composant le support du cycle puis l'ordre de parcours de ces sommets et faire attention à l'orientation du parcours) dans le graphe aléatoire de paramètres n et p_n ; chacun d'entre eux comporte l arêtes, et apparaît donc avec la probabilité p_n^l.

Calculons l'espérance de la variable N_n :

$$\begin{aligned}\mathbf{E}(N_n) &= \sum_{l=3}^{k} \tfrac{1}{2}\tbinom{n}{l}(l-1)!p_n^l \\ &\leqslant \sum_{l=3}^{k} \tfrac{n(n-1)\dots(n-l+1)}{2l} p_n^l \\ &\leqslant \sum_{l=3}^{k} (np_n)^l \\ &\leqslant (k-2)n^{k\varepsilon}.\end{aligned}$$

Avec l'inégalité de Markov pour la variable positive N_n,

$$\mathbf{P}(N_n \geqslant \tfrac{n}{2}) \leqslant \tfrac{2}{n}\mathbf{E}(N_n) \leqslant (k-2)n^{k\varepsilon-1}.$$

2. ▷ Soit $m \in [\![n]\!]$. L'événement $\alpha_n \geqslant m$ est inclus dans la réunion

$$\bigcup_{E\in\mathscr{P}_m([\![n]\!])} (E \text{ est un ensemble indépendant}).$$

Par conséquent,

$$\begin{aligned}\mathbf{P}(\alpha_n \geqslant m) &\leqslant \sum_{E\in\mathscr{P}_m([\![n]\!])} \mathbf{P}(E \text{ est un ensemble indépendant}) \\ &\leqslant \sum_{E\in\mathscr{P}_m([\![n]\!])} (1-p_n)^{\binom{m}{2}} \\ &\leqslant \tbinom{n}{m}(1-p_n)^{\binom{m}{2}}.\end{aligned}$$

▷ Appliquons le point précédent à l'entier $m = \lceil n/2k \rceil$ avec l'inégalité

$$(1-p_n)^{m(m-1)/2} \leqslant (1-p_n)^{n(n-2k)/(8k^2)}.$$

On conclut en remplaçant p_n par son expression.

3. D'après les deux questions précédentes, le choix de ε et le lemme d'encadrement, quand $n \to +\infty$,

$$\mathbf{P}(N_n \geqslant \tfrac{n}{2}) \to 0, \qquad \mathbf{P}(\alpha_n \geqslant \tfrac{n}{2k}) \to 0.$$

Choisissons n de sorte à ce que

$$\mathbf{P}(N_n \geqslant \tfrac{n}{2}) < \tfrac{1}{2}, \qquad \mathbf{P}(\alpha_n \geqslant \tfrac{n}{2k}) < \tfrac{1}{2},$$

soit, en passant au complémentaire,

$$\mathbf{P}(N_n < \tfrac{n}{2}) > \tfrac{1}{2}, \qquad \mathbf{P}(\alpha_n < \tfrac{n}{2k}) > \tfrac{1}{2}.$$

Par conséquent,

$$\begin{aligned}\mathbf{P}&((N_n < \tfrac{n}{2}) \cap (\alpha_n < \tfrac{n}{2k}))\\ &= \mathbf{P}(N_n < \tfrac{n}{2}) + \mathbf{P}(\alpha_n < \tfrac{n}{2k}) - \mathbf{P}((N_n < \tfrac{n}{2}) \cup (\alpha_n < \tfrac{n}{2k}))\\ &> \tfrac{1}{2} + \tfrac{1}{2} - 1 = 0.\end{aligned}$$

Ainsi, on dispose d'un graphe G d'ensemble de sommets $[\![n]\!]$ qui admet strictement moins de $\frac{n}{2}$ cycles de longueur au plus k et dont le plus grand ensemble indépendant est de cardinal strictement inférieur à $\frac{n}{2}$.

Considérons un graphe H déduit de G en enlevant un sommet dans chacun de ses cycles (et les arêtes qui en sont issues). Par conséquent, le graphe H compte au moins $\frac{n}{2}$ sommets.

Par ailleurs, H ne contient pas de cycles de longueur au plus k donc sa maille est strictement supérieure à k.

Enfin, avec le résultat I-7.6, on obtient que le nombre chromatique de H vérifie

$$\chi(H) \geqslant \tfrac{|H|}{\alpha(H)} \geqslant \tfrac{n/2}{\alpha(G)} > k.$$

5.10. Exercice. *Soit $G = ([\![n]\!], A)$ un graphe sans cycle. Considérons une énumération aléatoire $(s_k)_{k\in[\![n]\!]}$ de l'ensemble des sommets $S = [\![n]\!]$ et définissons l'ensemble aléatoire suivant*

$$S' = \big\{s_k,\ \forall j < k,\ \{s_j, s_k\} \notin A\big\}.$$

En d'autres termes, S' est l'ensemble des sommets qui sont placés avant tous leurs voisins dans G dans l'énumération aléatoire.

1. *Montrer que, pour tout $s \in S$, $\mathbf{P}(s \in S') = \frac{1}{\deg s+1}$.*
2. *En déduire l'espérance du nombre $|S'|$ de sommets dans S'.*
3. *Justifier que le graphe induit $G[S']$ est sans arêtes.*
4. *Montrer qu'il existe un graphe induit de G sans arêtes dont le nombre de sommets est supérieur ou égal à*

$$\sum_{s\in S} \tfrac{1}{\deg s+1}.$$

5. *Montrer qu'il existe un graphe induit de G complet dont le nombre de sommets est supérieur ou égal à*

$$\sum_{s\in S} \tfrac{1}{n-\deg s}.$$

Solution.

1. Fixons un sommet $s \in S$, et dénombrons les énumérations des sommets telles que $s \in S'$. Pour obtenir une énumération telle que $s \in S'$, il faut
   - ▷ choisir une partie $A \subset [\![n]\!]$ de cardinal $\deg s + 1$ qui correspond aux positions de s et de ses $\deg s$ voisins,
   - ▷ attribuer aux $\deg s$ voisins de s les positions correspondant aux éléments de A sauf le plus petit qui est dévolu à s,
   - ▷ attribuer aux sommets pas encore considérés (ni s, ni ses voisins) les positions du complémentaire de A.

   Ainsi, il existe

$$\tbinom{n}{\deg s+1}(\deg s)!(n-1-\deg s)! = \tfrac{n!}{\deg s+1}$$

   énumérations telles que $s \in S'$. On en déduit la probabilité recherchée en divisant par le cardinal de l'ensemble des énumérations (avec l'hypothèse implicite que le choix « aléatoire » de l'énumération s'avère uniforme).
2. Décomposons $|S'|$ comme somme de variables indicatrices

$$|S'| = \sum_{s\in S'} \mathbb{1}_{s\in S'},$$

   puis utilisons la linéarité de l'espérance

$$\begin{aligned}\mathbf{E}(|S'|) &= \sum_{s\in S'} \mathbf{P}(s \in S') \\ &= \sum_{s\in S'} \tfrac{1}{\deg s+1}.\end{aligned}$$

3. Supposons par l'absurde qu'il existe deux sommets voisins appartenant à S' : le second dans l'ordre de l'énumération aléatoire admet donc un sommet voisin placé avant lui dans cette énumération, ce qui contredit la définition de S'.

   Ainsi, le graphe aléatoire induit $G[S']$ est sans arêtes.

4. D'après le calcul de l'espérance de la question précédente, il existe une réalisation de l'énumération aléatoire telle que
$$|S'| \geqslant \sum_{s\in S} \tfrac{1}{\deg s+1}.$$
Pour ce choix de l'énumération, le graphe induit $G[S']$ est sans arêtes et admet bien le nombre annoncé de sommets.
5. Il suffit d'appliquer le résultat de la question précédente au graphe G' complémentaire de G en remarquant que
   - ▷ le degré d'un sommet s dans G' est $n-1-\deg_G s$, donc $\deg_G s + 1 = n - \deg_{G'} s$,
   - ▷ un graphe induit $G[S']$ est complet si, et seulement si, le graphe induit $G'[S']$ est sans arêtes.

En exploitant la convexité de la fonction $x \mapsto \frac{1}{n-x}$, on obtient alors le célèbre théorème suivant (que l'on pourra rapprocher de l'exercice précédent).

5.11. Proposition (Turán). *Un graphe à n sommets n'admettant pas de graphe induit isomorphe à K_{r+1} admet au plus $\frac{(r-1)n^2}{2r}$ arêtes.*

Asymptotique de graphes aléatoires finis

Chapitre V

Méthode du premier moment

L'objectif des deux chapitres suivants est d'étudier les limites lorsque $n \to +\infty$ de probabilités associées à un graphe aléatoire suivant le modèle binomial de paramètres n et p_n (avec une hypothèse asymptotique sur p_n).

On distingue alors deux méthodes :

- ▷ celle reposant sur le calcul d'une espérance et l'inégalité de Markov, que l'on trouve dans ce chapitre ;
- ▷ celle reposant sur le calcul d'une variance et l'inégalité de Bienaymé et Tchebychev, qui se trouve dans le chapitre suivant.

Les deux chapitres sont complémentaires et, pour tout dire, miroirs l'un de l'autre.

1. Inégalité de Markov

Commençons par un résultat très classique, sous sa forme la plus simple, et qui concerne les probabilités associées à des variables aléatoires positives.

1.1. Proposition. *Soit X une variable aléatoire à valeurs positives. Alors, pour tout $a > 0$,*

$$\mathbf{P}(X \geqslant a) \leqslant \tfrac{\mathbf{E}(X)}{a}.$$

Démonstration. Écrivons la décomposition de X selon que la valeur prise est en-deçà ou au-delà de a :

$$X = X\mathbb{1}_{X \geqslant a} + X\mathbb{1}_{X<a}.$$

Comme X est à valeurs positives, le second terme est positif, et donc

$$X \geqslant X\mathbb{1}_{X \geqslant a} \geqslant a\mathbb{1}_{X \geqslant a}.$$

Par croissance puis linéarité de l'espérance,

$$\mathbf{E}(X) \geqslant \mathbf{E}(a\mathbb{1}_{X \geqslant a}) = a\mathbf{E}(\mathbb{1}_{X \geqslant a}) = a\mathbf{P}(X \geqslant a).$$

♢

En particulier, avec $a = 1$, on obtient le corollaire suivant.

1.2. Corollaire. *Soit X une variable aléatoire à valeurs positives. Alors, $\mathbf{P}(X \geqslant 1) \leqslant \mathbf{E}(X)$.*

Cela est particulièrement intéressant pour les variables à valeurs dans $\mathbb{N}$. En effet, en utilisant que $\mathbf{P}(X = 0) = 1 - \mathbf{P}(X \geqslant 1)$, on obtient la minoration suivante.

1.3. Corollaire. *Soit X une variable aléatoire à valeurs dans $\mathbb{N}$. Alors, $\mathbf{P}(X = 0) \geqslant 1 - \mathbf{E}(X)$.*

1.4. Remarque. Les exemples qui suivent dans ce chapitre partagent tous la même structure.

— On considère une suite $(N_n)_n$ de variables aléatoires à valeurs dans $\mathbb{N}$.
— On montre que $\mathbf{E}(N_n) \to 0$ lorsque $n \to +\infty$.
— On applique le corollaire V-1.3 pour en déduire, par encadrement, que $\mathbf{P}(N_n = 0) \to 1$ lorsque $n \to +\infty$.

Dans ce chapitre et les suivants, on adopte la notation simplificatrice suivante.

1.5. Notation. Dans la suite, $X_{n,p}$ désigne un graphe aléatoire (c'est-à-dire une variable aléatoire à valeurs graphes) suivant le modèle binomial de paramètres $n \in \mathbb{N}^*$ et $p \in]0,1[$.

2. Asymptotiquement sans sommets isolés

2.1. Proposition. *Soit $(p_n)_n$ une suite à valeurs dans $]0,1[$ satisfaisant les conditions suivantes*

- $p_n \to 0$,
- $\frac{np_n}{\ln n} \to +\infty$, *c'est-à-dire* $\frac{\ln n}{n} = o(p_n)$.

Notons N_n la variable de comptage des sommets isolés (c'est-à-dire de degré 0) du graphe aléatoire X_{n,p_n}. Alors, $\mathbf{P}(N_n = 0) \to 1$.

Démonstration. ▷ Commençons par remarquer qu'un sommet fixé est isolé dans X_{n,p_n} avec probabilité $(1-p_n)^{n-1}$ puisque, pour qu'il soit isolé, il faut (et il suffit) qu'aucune des $n-1$ arêtes mutuellement indépendantes qui en sont issues ne soit présente dans X_{n,p_n}.

▷ La variable N_n est la somme des variables indicatrices $\mathbb{1}_{i \text{ est isolé}}$ donc son espérance est, par linéarité,

$$\mathbf{E}(N_n) = \sum_{i=1}^{n} \mathbf{P}(i \text{ est isolé}) = n(1-p_n)^{n-1}.$$

▷ Étudions le comportement asymptotique de $\mathbf{E}(N_n)$ avec les hypothèses sur p_n.

Ainsi, on obtient l'estimation asymptotique de $\ln\big(\mathbf{E}(N_n)\big)$:

$$\begin{aligned}\ln(\mathbf{E}(N_n)) &= \ln n + (n-1)\ln(1-p_n)\\ &= \ln n + (n-1)\big(-p_n + o(p_n)\big)\\ &= -np_n + o(np_n)\end{aligned}$$

car $\ln n$ et p_n sont des termes négligeables devant np_n.

Par conséquent, $\ln(\mathbf{E}(N_n)) \to -\infty$, c'est-à-dire $\mathbf{E}(N_n) \to 0$.

▷ Appliquons le corollaire V-1.3 de l'inégalité de Markov à la variable N_n à valeurs dans $\mathbb{N}$:

$$1 \geqslant \mathbf{P}(N_n = 0) = 1 - \mathbf{E}(N_n).$$

Par encadrement, $\mathbf{P}(N_n = 0) \to 1$. ♢

2.2. Remarque. La démonstration et le résultat peuvent être étendus au cas où les hypothèses sur $(p_n)_n$ sont remplacées par l'équivalence $p_n \sim c\frac{\ln n}{n}$ pour une constante $c > 1$; dans ce cas $\ln(\mathbf{E}(N_n)) \sim (1-c)\ln n$ et donc $\mathbf{E}(N_n) \to 0$.

3. Asymptotiquement sans triangle

3.1. Proposition. *Soit $(p_n)_n$ une suite à valeurs dans $]0,1[$ telle que $np_n \to 0$. Notons N_n la variable de comptage des triangles (c'est-à-dire de graphes induits isomorphes à K_3) du graphe aléatoire X_{n,p_n}. Alors, $\mathbf{P}(N_n = 0) \to 1$.*

Démonstration. ▷ Commençons par écrire que la variable de comptage N_n est la somme des indicatrices de présence des différents triangles envisageables. Ensuite, remarquons qu'un triangle donné est présent dans le graphe X_{n,p_n} avec probabilité p_n^3 (il faut la présence de trois arêtes mutuellement indépendantes).

Par linéarité de l'espérance, $\mathbf{E}(N_n) = \binom{n}{3}p_n^3$, d'où l'on déduit l'équivalent suivant $\mathbf{E}(N_n) \sim \frac{1}{6}(np_n)^3$. Par suite, $\mathbf{E}(N_n) \to 0$ avec l'hypothèse asymptotique sur $(p_n)_n$.

▷ Appliquons le corollaire V-1.3 de l'inégalité de Markov à la variable N_n à valeurs dans $\mathbb{N}$,

$$1 \geqslant \mathbf{P}(N_n = 0) = 1 - \mathbf{E}(N_n).$$

Par encadrement, $\mathbf{P}(N_n = 0) \to 1$. ♢

Un calcul analogue permet de montrer le résultat suivant en partant de la probabilité qu'un ensemble de k sommets induisent un graphe isomorphe au graphe complet K_k dans un graphe aléatoire.

3.2. Proposition. *Fixons k un entier supérieur ou égal à 3. Soit $(p_n)_n$ une suite à valeurs dans $]0,1[$ telle que*

$$p_n = o\big(n^{-2/(k-1)}\big).$$

Notons N_n la variable de comptage des graphes induits isomorphes à K_k du graphe aléatoire X_{n,p_n}. Alors, $\mathbf{P}(N_n = 0) \to 1$.

Démonstration. ▷ La probabilité que k sommets fixés induisent dans X_{n,p_n} un graphe isomorphe à K_k est $p_n^{k(k-1)/2}$ puisqu'il y a exactement $\binom{k}{2} = \frac{k(k-1)}{2}$ arêtes entre ces sommets.
Par linéarité de l'espérance, $\mathbf{E}(N_n) = \binom{n}{k} p_n^{k(k-1)/2}$. Par conséquent, on en déduit l'équivalent

$$\begin{aligned}\mathbf{E}(N_n) &\sim \tfrac{1}{k!} n^k p_n^{k(k-1)/2} \\ &\sim \tfrac{1}{k!} \big(n^{(k-1)/2} p_n\big)^{k(k-1)/2}.\end{aligned}$$

Ainsi, d'après l'hypothèse asymptotique sur $(p_n)_n$, $\mathbf{E}(N_n) \to 0$ et on conclut une fois encore avec le corollaire V-1.3 de l'inégalité de Markov. ◇

On retrouve bien le résultat précédent pour le cas des triangles en prenant $k = 3$.

3.3. Remarque. On peut facilement étendre ce résultat pour étudier la présence d'un graphe H à s sommets et a arêtes. Soit $(p_n)_n$ une suite à valeurs dans $]0,1[$ telle que $n^{s/a} p_n \to 0$. Alors, le nombre N_n de graphes induits isomorphes à H du graphe aléatoire X_{n,p_n} vérifie bien $\mathbf{P}(N_n = 0) \to 1$.

4. Exercices

4.1. Exercice. *Soit $p \in]0,1[$ et k un entier naturel fixé. On rappelle que l'on a établi à l'exercice III-5.7 page 68 que, pour tout entier naturel non nul n, le nombre moyen de sommets de degré j dans $X_{n,p}$ est*

$$n\tbinom{n-1}{j} p^j (1-p)^{n-1-j}.$$

Montrer que la probabilité que le degré minimal d'un sommet de $X_{n,p}$ soit strictement supérieur à k tend vers 1 lorsque n tend vers $+\infty$.

Solution. Notons, pour tout n, N_n la variable de comptage des sommets de degré au plus k dans le graphe aléatoire $X_{n,p}$. D'après le résultat rappelé dans l'énoncé, le nombre moyen de sommets de degré au plus k dans $X_{n,p}$ est

$$\begin{aligned}\mathbf{E}(N_n) &= \sum_{j=0}^{k} n\tbinom{n-1}{j} p^j (1-p)^{n-1-j} \\ &\leqslant \sum_{j=0}^{k} n^{j+1}(1-p)^n.\end{aligned}$$

Chaque terme de cette somme est de limite nulle lorsque n tend vers $+\infty$. Par conséquent, $\mathbf{E}(N_n) \to 0$.

Or, d'après le corollaire V-1.3 de l'inégalité de Markov,

$$1 \geqslant \mathbf{P}(N_n = 0) \geqslant 1 - \mathbf{E}(N_n),$$

donc $\mathbf{P}(N_n = 0) \to 1$. Comme l'événement $N_n = 0$ correspond à l'événement « $X_{n,p}$ est de degré minimal strictement supérieur à k » (car il n'y a aucun sommet de degré inférieur ou égal à k), on obtient le résultat annoncé.

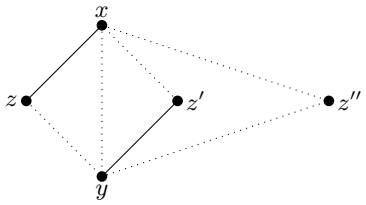
4.2. Exercice. *Soit $p \in]0,1[$. Notons N_n la variable de comptage des paires de sommets à distance strictement supérieure à 2 dans le graphe aléatoire $X_{n,p}$.*

1. *Justifier que deux sommets fixés sont à distance strictement supérieure à 2 dans $X_{n,p}$ avec probabilité*
$$(1-p)(1-p^2)^{n-2}.$$
2. *En déduire que $\mathbf{P}(N_n = 0) \to 1$ puis que la probabilité que $X_{n,p}$ soit de diamètre au plus 2 tend vers 1.*

Solution.

1. Fixons x, $y \in [\![n]\!]$ deux sommets distincts.

   L'événement « x et y sont à distance strictement supérieure à 2 » se décompose comme intersection des événements suivants :
   - ▷ les sommets x et y ne sont pas voisins, c'est-à-dire l'arête $\{x,y\}$ n'existe pas dans $X_{n,p}$,
   - ▷ pour tout autre sommet z, l'une (au moins) des deux arêtes $\{x,z\}$ ou $\{y,z\}$ n'existe pas dans $X_{n,p}$.



En notant $X_{n,p} = (S, A)$, la probabilité recherchée est donc celle de l'événement

$$\big(\{x,y\} \notin A\big) \bigcap \bigcap_{z \in S \setminus \{x,y\}} \big(\{x,z\} \notin A \cup \{y,z\} \notin A\big).$$

Par indépendance des arêtes, la probabilité de cet événement est

$$\big(1 - \mathbf{P}(\{x,y\} \in A)\big) \prod_{z \in S \setminus \{x,y\}} \big(1 - \mathbf{P}(\{x,z\} \in A, \{y,z\} \in A)\big)$$

c'est-à-dire $(1-p)(1-p^2)^{n-2}$.

2. ▷ En utilisant tour à tour l'écriture de N_n comme somme d'indicatrices sur toutes les paires de sommets, la linéarité de l'espérance et le résultat de la première question, on obtient

$$\mathbf{E}(N_n) = \tbinom{n}{2}(1-p)(1-p^2)^{n-2}.$$

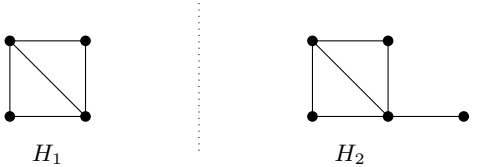
Par conséquent, $\mathbf{E}(N_n) \to 0$ lorsque $n \to +\infty$. Or, d'après le corollaire V-1.3 de l'inégalité de Markov,

$$1 \geqslant \mathbf{P}(N_n = 0) \geqslant 1 - \mathbf{E}(N_n),$$

donc $\mathbf{P}(N_n = 0) \to 1$.

▷ Pour conclure, remarquons que l'événement « $N_n = 0$ » correspond à « la distance maximale entre deux sommets de $X_{n,p}$ est inférieure ou égale à 2 » soit encore « le diamètre de $X_{n,p}$ est inférieur ou égal à 2 ».

4.3. Exercice (Ceci n'est pas un paradoxe). *Considérons deux graphes H_1 et H_2 respectivement isomorphes aux graphes représentés ci-dessous*



H_1 H_2

Fixons un réel $\alpha \in \left]\frac{4}{5}, \frac{5}{6}\right[$ *et définissons pour tout entier naturel* n *non nul,* $p_n = n^{-\alpha}$.

1. *Montrer que la probabilité que le graphe* X_{n,p_n} *admette un sous-graphe isomorphe à* H_1 *tend vers* 0 *lorsque* $n \to +\infty$.
2. *Soit* $N_n^{(2)}$ *la variable de comptage des sous-graphes de* X_{n,p_n} *isomorphes à* H_2. *Montrer que* $\mathbf{E}(N_n^{(2)}) \to +\infty$.
3. *Expliquer l'apparent paradoxe entre ces deux résultats.*

Solution.

1. Notons $N_n^{(1)}$ la variable de comptage des sous-graphes de X_{n,p_n} isomorphes à H_1. Écrivons $N_n^{(1)}$ comme la somme des indicatrices des différentes apparitions possibles de H_1. Chaque occurrence apparaît avec probabilité p_n^5 (car il y a cinq arêtes).

   Dénombrons ces occurrences : il convient de choisir les quatre sommets puis de considérer le nombre c_1 de façons d'obtenir un sous-graphe isomorphe à H_1 à partir de ces quatre sommets (qui a été calculé à partir de l'ordre du groupe d'automorphismes à l'exercice I-8.6). Ainsi,

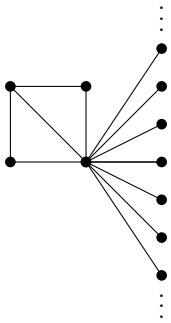
   $$\begin{aligned}\mathbf{E}\big(N_n^{(1)}\big) &= c_1 \tbinom{n}{4} p_n^5 \\ &\sim \tfrac{c_1}{4!} n^{4-5\alpha},\end{aligned}$$

   donc $\mathbf{E}\big(N_n^{(1)}\big) \to 0$ d'après la condition $\alpha > \frac{4}{5}$. D'après le corollaire V-1.3 de l'inégalité de Markov, $\mathbf{P}\big(N_n^{(1)} = 0\big) \to 1$.
2. En procédant comme à la question précédente et en notant c_2 le nombre de façons d'obtenir un sous-graphe isomorphe à H_2 à partir de 5 sommets, on obtient

   $$\begin{aligned}\mathbf{E}\big(N_n^{(2)}\big) &= c_2 \tbinom{n}{5} p_n^6 \\ &\sim \tfrac{c_2}{5!} n^{5-6\alpha}.\end{aligned}$$

   On observe ainsi que $\mathbf{E}\big(N_n^{(2)}\big) \to +\infty$, car $\alpha < \frac{5}{6}$.

3. Comme le graphe H_1 se retrouve dans le graphe H_2, toute apparition d'un graphe H_2 entraîne l'apparition d'un graphe H_1 ; il peut donc sembler paradoxal que la probabilité d'apparition du graphe H_1 tende vers 0 alors que le nombre moyen d'apparitions de H_2 tend vers $+\infty$.



Pour lever cette incompréhension, il faut réaliser que chaque occurrence de H_1 peut entraîner jusqu'à $2(n-4)$ occurrences de H_2 (deux choix pour le point de raccord et $n-4$ pour l'autre extrémité de la nouvelle arête) : la probabilité de voir apparaître une occurrence de H_1 diminue mais le nombre d'occurrences de H_2 qui prolonge chaque occurrence diverge...
L'intuition est alors trompée par le calcul de la limite du produit d'un terme de limite nulle avec un terme de limite $+\infty$.

En anticipant les résultats du chapitre VII, on peut préciser cette explication : des fonctions de seuil pour l'apparition des graphes H_1 et H_2 dans le graphe aléatoire sont respectivement $p^{4/5}$ et $p^{5/6}$. On s'est donc placé ici strictement entre ces deux seuils ce qui permet d'obtenir les résultats annoncés.

4.4. Exercice. *Soit $(p_n)_n$ une suite à valeurs dans $]0,1[$ telle que $n^2 p_n \to 0$. Notons N_n la variable de comptage des arêtes isolées du graphe aléatoire X_{n,p_n}, c'est-à-dire des arêtes tracées qui ne partagent pas d'extrémité avec une autre arête tracée dans le graphe X_{n,p_n}.*

Montrer que $\mathbf{P}(N_n = 0) \to 1$.

Solution. Commençons par calculer la probabilité qu'une arête donnée soit isolée : pour cela, il faut à la fois qu'elle existe et que les $n-2$

autres arêtes issues de chacune de ses extrémités n'existent pas ; cette probabilité est donc

$$p_n(1-p_n)^{2(n-2)}.$$

Écrivons ensuite N_n comme la somme des indicatrices que chacune des arêtes soit isolée puis, en utilisant la linéarité de l'espérance,

$$\mathbf{E}(N_n) = \tbinom{n}{2} p_n(1-p_n)^{2(n-2)}.$$

Comme $\binom{n}{2} \sim \frac{n^2}{2}$, $\mathbf{E}(N_n) = O(n^2 p_n)$, et donc $\mathbf{E}(N_n) \to 0$.

D'après le corollaire V-1.3 de l'inégalité de Markov, $\mathbf{P}(N_n = 0) \to 1$. *On peut remarquer avec ce calcul que l'hypothèse $p_n = o(\frac{1}{n^2})$ est loin d'être optimale.*

4.5. Exercice. *Soit $(p_n)_n$ une suite à valeurs dans $]0,1[$ telle que $np_n \to 0$. Montrer que la probabilité que le graphe X_{n,p_n} contienne un cycle tend vers* 0.

Solution. Pour tout l entre 3 et n, il existe au plus $\frac{1}{2}\binom{n}{l}(l-1)!$ cycles de longueur l (il faut choisir les l éléments composant le support du cycle puis une permutation circulaire de ce support et faire attention à l'ordre de parcours) dans X_{n,p_n} ; chacun d'entre eux comporte l arêtes, et apparaît donc avec la probabilité p_n^l dans X_{n,p_n}.

Par conséquent, la variable N_n dénombrant les cycles dans X_{n,p_n}, c'est-à-dire la variable de comptage obtenue comme somme des indicatrices de présence de chacun de ces cycles, admet comme espérance

$$\mathbf{E}(N_n) = \sum_{l=3}^{n} \tfrac{1}{2}\tbinom{n}{l}(l-1)!p_n^l.$$

Majorons cette espérance :

$$\begin{aligned}\mathbf{E}(N_n) &= \sum_{l=3}^{n} \tfrac{n(n-1)\dots(n-l+1)}{2l} p_n^l \\ &\leqslant \sum_{l=3}^{n} (np_n)^l \leqslant \tfrac{(np_n)^3}{1-np_n}.\end{aligned}$$

D'après l'hypothèse asymptotique sur $(p_n)_n$, $\mathbf{E}(N_n) \to 0$.

On conclut encore avec le corollaire V-1.3 de l'inégalité de Markov que $\mathbf{P}(N_n = 0) \to 1$, ce qui traduit que la probabilité que le graphe X_{n,p_n}

ne contienne aucun cycle est de limite 1, ou par passage au complémentaire, que la probabilité le graphe X_{n,p_n} contienne un cycle tend vers 0.

L'exercice suivant est le pendant de la proposition V-3.2 pour pour les graphes aléatoires suivant le modèle uniforme.

4.6. Exercice. *Fixons k un entier supérieur ou égal à 3.*
Soit $(p_n)_n$ une suite à valeurs dans $]0,1[$ telle que

$$p_n = o\big(n^{-2/(k-1)}\big).$$

Notons N_n la variable de comptage des graphes induits isomorphes à K_k d'un graphe aléatoire X qui suit le modèle uniforme de paramètres n et

$$M = \big\lfloor p_n \tbinom{n}{2} \big\rfloor.$$

Remarquons que le choix du paramètre M correspond en moyenne (approximativement à cause de la partie entière) au nombre d'arêtes d'un graphe binomial de paramètres n et p_n.

1. *Montrer que $\mathbf{E}(N_n) \leqslant \binom{n}{k} p_n^{k(k-1)/2}$.*
2. *En déduire que $\mathbf{P}(N_n = 0) \to 1$.*

Solution.

1. Écrivons N_n comme somme des indicatrices correspondant à chaque partie de k sommets. Par ailleurs, pour une partie A de k sommets donnée, il existe

$$\binom{\binom{n}{2} - \binom{k}{2}}{M - \binom{k}{2}}$$

graphes à M arêtes contenant toutes les arêtes entre sommets de A (on a déjà $\binom{k}{2}$ arêtes et il faut choisir les $M - \binom{k}{2}$ autres).

Ainsi, par définition du modèle uniforme, la probabilité que le graphe induit sur X par la partie A soit complet est

$$\frac{\binom{\binom{n}{2} - \binom{k}{2}}{M - \binom{k}{2}}}{\binom{\binom{n}{2}}{M}} = \prod_{l=0}^{\binom{k}{2}-1} \frac{M-l}{\binom{n}{2}-l}.$$

Or, par définition de M, chaque fraction du produit est inférieure à p_n : la probabilité que le graphe induit sur X par la

partie A soit complet est donc majorée par $p_n^{k(k-1)/2}$. En sommant sur les $\binom{k}{2}$ parties de cardinal k (ce qui revient à exploiter la linéarité de l'espérance et l'écriture de N_n comme somme d'indicatrices), on obtient la majoration souhaitée pour $\mathbf{E}(N_n)$.

2. Le majorant dans l'inégalité précédente est l'espérance du nombre de graphes induits isomorphes à K_k du graphe aléatoire qui suit le modèle binomial de paramètres n et p. On a déjà établi que cette espérance tend vers 0. Par encadrement, $\mathbf{E}(N_n) \to 0$.

   D'après le corollaire V-1.3 de l'inégalité de Markov, on obtient comme attendu $\mathbf{P}(N_n = 0) \to 1$.

L'exercice suivant montre l'analogue de la proposition V-2.1 pour les graphes aléatoires suivant le modèle uniforme.

4.7. Exercice. *Soit $(m_n)_n$ une suite d'entiers naturels non nuls telle que*

- ▷ *pour tout n, $m_n \leqslant \binom{n-1}{2}$, ce qui assure notamment $m_n = O(n^2)$,*
- ▷ *$m_n = \frac{1}{2} n \ln n + n u_n$ avec $u_n \to +\infty$.*

Notons N_n la variable de comptage des sommets isolés dans un graphe aléatoire qui suit le modèle uniforme de paramètres n et m_n.

Montrer que $\mathbf{P}(N_n = 0) \to 1$ (en d'autres termes, que ce graphe est asymptotiquement sans sommets isolés).

Solution. ▷ Calculons tout d'abord la probabilité qu'un sommet fixé s soit isolé dans un graphe suivant le modèle uniforme de paramètres n et m_n : cet événement correspond au choix des m_n arêtes parmi les arêtes n'admettant pas s comme extrémité ; il y a

$$\binom{\binom{n-1}{2}}{m_n}$$

ensembles de m_n arêtes n'admettant pas s comme extrémité (c'est-à-dire les arêtes joignant les $n-1$ autres sommets) et, d'après la définition du modèle uniforme, chacun d'entre eux est l'ensemble des arêtes avec la probabilité

$$\frac{1}{\binom{\binom{n}{2}}{m_n}}.$$

Ainsi, la probabilité que s soit isolé est

$$\frac{\binom{\binom{n-1}{2}}{m_n}}{\binom{\binom{n}{2}}{m_n}} = \prod_{k=0}^{m_n-1} \frac{\binom{n-1}{2}-k}{\binom{n}{2}-k}.$$

En décomposant les coefficients binomiaux en factorielles et en mettant en facteur les termes $\frac{(n-1)(n-2)}{2}$ au numérateur et $\frac{n(n-1)}{2}$ au dénominateur, cette probabilité se récrit

$$\left(\tfrac{n-2}{n}\right)^{m_n} \prod_{k=0}^{m_n-1} \frac{1-\frac{2k}{(n-1)(n-2)}}{1-\frac{2k}{n(n-1)}},$$

donc est majorée par

$$\left(\tfrac{n-2}{n}\right)^{m_n}.$$

▷ La variable N_n est la somme sur les sommets s des indicatrices de l'événement « s est isolé » donc par linéarité de l'espérance, $\mathbf{E}(N_n)$ est égale à n fois la probabilité calculée ci-dessus et donc

$$\mathbf{E}(N_n) \leqslant n\left(\tfrac{n-2}{n}\right)^{m_n}.$$

En particulier, le logarithme du membre de droite est, avec la définition de m_n en fonction de u_n,

$$\begin{aligned} m_n \ln\left(1-\tfrac{2}{n}\right) + \ln n &= -\tfrac{2m_n}{n} + \ln n + O\left(\tfrac{m_n}{n^2}\right) \\ &= -2u_n + O(1). \end{aligned}$$

Ainsi, par hypothèse sur u_n, ce logarithme tend vers $-\infty$ et $\mathbf{E}(N_n)$ vers 0.

▷ On conclut à nouveau d'après le corollaire V-1.3 de l'inégalité de Markov que $\mathbf{P}(N_n = 0) \to 1$.

La résolution de l'exercice suivant s'appuie sur l'exercice I-8.10 ; il en reprend les notations et le résultat à partir de la troisième question.

4.8. Exercice (Théorème de Posá [Pós76].). *Soit $(p_n)_n$ la suite définie, pour tout $n \geqslant 1$, par $p_n = c\frac{\ln n}{n}$ pour une constante $c > 24$. Considérons le graphe aléatoire X_{n,p_n}.*

1. *Notons N_n la variable de comptage des couples (U, V) de parties disjointes de l'ensemble des sommets du graphe X_{n,p_n} vérifiant les conditions suivantes :*

- ▷ *le cardinal k de U se situe entre 1 et $\frac{n}{4}$,*
- ▷ *le cardinal de V est exactement $n-3k+1$,*
- ▷ *aucune arête de X_{n,p_n} ne relie un sommet de U et un sommet de V.*

Montrer que $\mathbf{P}(N_n \geqslant 1) \leqslant \frac{1}{n^2}$.

Soit B l'événement « il existe une partie U de l'ensemble des sommets de X_{n,p_n} de cardinal k entre 1 et $\frac{n}{4}$ telle que l'ensemble des sommets appartenant à U ou voisins d'un sommet de U est de cardinal inférieur ou égal à $3k-1$ ».

2. *Montrer que $\mathbf{P}(B) \leqslant \frac{1}{n^2}$.*

   Pour tout sommet s, considérons $G(s)$ le graphe induit par X_{n,p_n} en supprimant le sommet s et définissons B_s l'événement « il existe une partie U de l'ensemble des sommets du graphe $G(s)$ de cardinal k entre 1 et $\frac{n-1}{4}$ telle que l'ensemble des sommets de U ou voisins d'un sommet de U est de cardinal cardinal inférieur ou égal à $3k-1$ ».

3. *Soit s un sommet et $\mathscr{C}$ un chemin de longueur maximale dans le graphe $G(s)$.*

   Montrer que, sachant $\overline{B_s}$, la probabilité que s ne soit pas relié à une extrémité d'un chemin déduit de $\mathscr{C}$ par transformations autorisées est majorée par

$$(1-p_n)^{(n-1)/4}.$$

4. *En déduire que la probabilité que qu'un sommet s n'appartienne pas à tous les plus longs chemins de X_{n,p_n} est majorée par*

$$(1-p_n)^{(n-1)/4} + \mathbf{P}(B_s),$$

   puis qu'elle est négligeable devant $\frac{1}{n}$.

5. *Montrer que la probabilité que tous les plus longs chemins de X_{n,p_n} passent par tous les sommets tend vers 1.*

Solution.

1. ▷ Remarquons tout d'abord qu'un couple (U,V) de parties disjointes de l'ensemble des sommets du graphe aléatoire X_{n,p_n} vérifie la dernière condition avec probabilité $(1-p_n)^{|U||V|}$ puisqu'il y a $|U||V|$ arêtes entre ces parties.

▷ Écrivons ensuite N_n comme somme d'indicatrices sur tous les couples (U,V), calculons son espérance par linéarité et regroupons les différents termes selon le cardinal k de la partie U ; ainsi,

$$\begin{aligned}\mathbf{E}(N_n) &= \sum_{k=1}^{\lfloor n/4\rfloor} \tbinom{n}{k}\tbinom{n-k}{n-3k+1}(1-p_n)^{k(n-3k+1)}\\ &= \sum_{k=1}^{\lfloor n/4\rfloor} \tbinom{n}{k}\tbinom{n-k}{2k+1}(1-p_n)^{k(n-3k+1)}.\end{aligned}$$

Majorons le terme général de cette somme :

$$\begin{aligned}&\tbinom{n}{k}\tbinom{n-k}{2k+1}(1-p_n)^{k(n-3k+1)}\\ &\qquad\leqslant n^k(n-k)^{2k+1}(1-p_n)^{k(n-3k+1)}\\ &\qquad\leqslant n^{3k}(1-p_n)^{kn/4}\\ &\qquad\leqslant n^{3k}\exp\big(-k\tfrac{n}{4}p_n\big)\\ &\qquad\leqslant n^{k(3-c/4)}\leqslant n^{3-c/4}.\end{aligned}$$

Ainsi,

$$\begin{aligned}n^2\mathbf{E}(N_n) &\leqslant n^2\sum_{k=1}^{\lfloor n/4\rfloor} n^{3-c/4}\\ &\leqslant n^{6-c/4}.\end{aligned}$$

Par conséquent, comme $c>24$ (et donc $6-\frac{c}{4}<0$),

$$\mathbf{E}(N_n)\leqslant n^{-2}.$$

D'après l'inégalité de Markov V-1.2, $\mathbf{P}(N_n\geqslant 1)\leqslant n^{-2}$.

2. Soit U une partie de cardinal k entre 1 et $\frac{n}{4}$ vérifiant la condition de l'événement B et k son cardinal : il y a alors $n-3k+1$ sommets (au moins) qui n'appartiennent ni à U, ni à l'ensemble des voisins de U. Considérons une partie V formée de $n-3k+1$ de ces sommets. Le couple (U,V) vérifie alors les hypothèses de la question précédente.

   Ainsi, l'événement B est inclus dans l'événement « $N_n\geqslant 1$ », ce qui entraîne la majoration souhaitée avec la première question.

   Intuitivement, on indique que toute partie de petit cardinal admet (asymptotiquement) un grand « taux d'extension » (elle se connecte à beaucoup d'autres sommets).

3. Notons E l'ensemble des extrémités des chemins déduits de $\mathscr{C}$ par transformations autorisées.

   D'après l'exercice I-8.10, l'ensemble obtenu par réunion de E et de ses voisins dans $G(s)$ est de cardinal inférieur ou égal à $3|E|-1$. Comme l'on suppose l'événement $\overline{B_s}$ réalisé, cela entraîne que E est de cardinal supérieur à $\frac{n-1}{4}$.

   Par conséquent, la probabilité que s ne soit pas relié à E est majorée par la probabilité d'absence de $\frac{n-1}{4}$ arêtes, soit

   $$(1-p_n)^{(n-1)/4}.$$

4. ▷ Remarquons qu'un plus long chemin du graphe X_{n,p_n} qui ne passe pas par s est, de fait, un plus long chemin de $G(s)$; par ailleurs, un tel chemin ne peut être modifié (par transformations autorisées) en un chemin prolongeable par s (sinon, on contredit la propriété de plus long chemin de X_{n,p_n}).

   Notons C_s l'événement « s n'appartient pas à tous les plus longs chemins de X_{n,p_n} » et utilisons la formule des probabilités totales :

   $$\begin{aligned}\mathbf{P}(C_s) &= \mathbf{P}\big(C_s|\overline{B_s}\big)\mathbf{P}\big(\overline{B_s}\big) + \mathbf{P}(C_s|B_s)\mathbf{P}(B_s)\\ &\leqslant \mathbf{P}\big(C_s|\overline{B_s}\big) + \mathbf{P}(B_s)\\ &\leqslant (1-p_n)^{(n-1)/4} + \mathbf{P}(B_s),\end{aligned}$$

   en utilisant la remarque et la question précédente.

   ▷ Utilisons les hypothèses sur p_n et la question 2, pour contrôler ce majorant par un terme immédiatement négligeable devant $\frac{1}{n}$:

   $$\begin{aligned}\mathbf{P}(C_s) &\leqslant \exp\big(-\tfrac{n-1}{4}c\tfrac{\ln n}{n}\big) + \tfrac{1}{(n-1)^2}\\ &= n^{-c(n-1)/4n} + (n-1)^{-2}.\end{aligned}$$

5. Soit N'_n la variable de comptage des sommets n'appartenant pas à tous les plus longs chemins de X_{n,p_n}.

   Par linéarité de l'espérance, on obtient

   $$\mathbf{E}(N'_n) = \sum_{s=1}^{n} \mathbf{P}(C_s),$$

   ce qui, avec le contrôle de $\mathbf{P}(C_s)$ établi ci-dessus, entraîne

   $$\mathbf{E}(N'_n) = o(1).$$

Ainsi, avec l'inégalité de Markov V-1.2,

$$\mathbf{P}(N'_n \geqslant 1) = o(1);$$

en d'autres termes, la probabilité que tous les plus longs chemins dans X_{n,p_n} passent par tous les sommets tend vers 1.

On a notamment obtenu le corollaire suivant : pour ce choix de p_n, la probabilité d'existence d'un chemin hamiltonien dans X_{n,p_n} tend vers 1.

Chapitre VI

Méthode du deuxième moment

Dans ce chapitre, on reprend les exemples du chapitre précédent en remplaçant l'inégalité de Markov par l'inégalité de Bienaymé et Tchebychev, donc en majorant la quantité $\mathbf{P}(X > 0)$ au lieu de la minorer. Cela donne alors des résultats de « présence » là où l'on obtenait jusqu'ici des résultats d'« absence ».

1. Inégalité de Bienaymé et Tchebychev

Commençons par rappeler une inégalité classique (qui sert notamment à établir la loi faible des grands nombres pour des variables admettant un moment d'ordre 2).

1.1. Proposition. *Soit X une variable aléatoire réelle admettant un moment d'ordre 2. Alors, pour tout $a > 0$,*

$$\mathbf{P}\big(|X - \mathbf{E}(X)| \geqslant a\big) \leqslant \tfrac{\mathbf{V}(X)}{a^2}.$$

Démonstration. Récrivons l'événement considéré avec la croissance de la fonction $x \mapsto x^2$ sur $\mathbb{R}_+$ puis appliquons l'inégalité de Mar-

kov à la variable positive $(X - \mathbf{E}(X))^2$:

$$\begin{aligned}\mathbf{P}\big(|X - \mathbf{E}(X)| \geqslant a\big) &= \mathbf{P}\big((X - \mathbf{E}(X))^2 \geqslant a^2\big) \\ &\leqslant \tfrac{\mathbf{E}((X-\mathbf{E}(X))^2)}{a^2} = \tfrac{\mathbf{V}(X)}{a^2}.\end{aligned}$$

♢

1.2. Corollaire. *Soit X une variable aléatoire non nulle à valeurs dans $\mathbb{N}$ admettant un moment d'ordre 2. Alors,*

$$\mathbf{P}(X = 0) \leqslant \tfrac{\mathbf{V}(X)}{\mathbf{E}(X)^2}.$$

Démonstration. Remarquons dans un premier temps que l'événement $X = 0$ est inclus dans l'événement $|X - \mathbf{E}(X)| \geqslant \mathbf{E}(X)$. Ensuite, par croissance de $\mathbf{P}$ puis par l'inégalité de Bienaymé et Tchebychev,

$$\mathbf{P}(X = 0) \leqslant \mathbf{P}\big(|X - \mathbf{E}(X)| \geqslant \mathbf{E}(X)\big) \leqslant \tfrac{\mathbf{V}(X)}{\mathbf{E}(X)^2}.$$

♢

1.3. Remarque. Comme dans le chapitre précédent, on développe à travers différents exemples une unique méthode dont voici le schéma général

- On considère une suite $(N_n)_n$ de variables aléatoires à valeurs dans $\mathbb{N}$.
- On montre que $\frac{\mathbf{V}(N_n)}{\mathbf{E}(N_n)^2} \to 0$ lorsque $n \to +\infty$.
- On applique le corollaire VI-1.2 pour en déduire, par encadrement, que $\mathbf{P}(N_n = 0) \to 0$ lorsque $n \to +\infty$.

Le point le plus délicat dans le cheminement précédent est celui qui consiste à estimer $\mathbf{V}(N_n)$ asymptotiquement pour la deuxième étape ; ce n'est plus, comme dans le chapitre précédent, une simple utilisation de la linéarité de l'espérance : il faut désormais exploiter les éventuelles hypothèses d'indépendance.

2. Asymptotiquement avec sommets isolés

Complétons le résultat asymptotique sur les sommets isolés énoncé à la proposition V-2.1 et obtenu par la méthode du premier moment.

2.1. Proposition. *Soit $(p_n)_n$ une suite à valeurs dans $]0,1[$ telle que $\frac{np_n}{\ln n} \to 0$, c'est-à-dire telle que $p_n = o(\frac{\ln n}{n})$.*

Notons N_n la variable de comptage des sommets isolés, c'est-à-dire de degré 0, du graphe aléatoire X_{n,p_n}. Alors, $\mathbf{P}(N_n = 0) \to 0$.

Démonstration. Rappelons que $\mathbf{E}(N_n) = n(1-p_n)^{n-1}$ puis calculons le moment d'ordre 2 de N_n. Pour cela, utilisons l'écriture de $N_n(N_n-1)$ comme somme d'indicatrices puis la linéarité de l'espérance :

$$\begin{aligned}\mathbf{E}\big(N_n(N_n-1)\big) &= \mathbf{E}\Big(\sum_{k=1}^{n}\sum_{\substack{l=1\\l\neq k}}^{n} \mathbb{1}_{k \text{ est isolé}}\mathbb{1}_{l \text{ est isolé}}\Big)\\ &= \sum_{k=1}^{n}\sum_{\substack{l=1\\l\neq k}}^{n} \mathbf{P}(k,l \text{ sont isolés})\\ &= n(n-1)(1-p_n)^{2n-3},\end{aligned}$$

car il a exactement $2n-3$ arêtes dont au moins l'une des extrémités est k ou l.

Par conséquent,

$$\begin{aligned}\mathbf{E}\big(N_n(N_n-1)\big) &= \big(n(1-p_n)^{n-1}\big)^2 \tfrac{n-1}{n(1-p_n)}\\ &= \mathbf{E}(N_n)^2 \tfrac{n-1}{n(1-p_n)}.\end{aligned}$$

Comme $p_n \to 0$, $\mathbf{E}\big(N_n(N_n-1)\big) \sim \mathbf{E}(N_n)^2$.

Remarquons alors que

$$\begin{aligned}\ln\big(\mathbf{E}(N_n)\big) &= \ln n + (n-1)\ln(1-p_n)\\ &= \ln n + (n-1)\big(-p_n + o(p_n)\big)\\ &= \ln n - np_n + o(np_n)\\ &= \ln n\Big(1 - \tfrac{np_n}{\ln n} + o\big(\tfrac{np_n}{\ln n}\big)\Big).\end{aligned}$$

Ainsi, $\ln\big(\mathbf{E}(N_n)\big) \sim \ln n$, donc $\mathbf{E}(N_n) \to +\infty$. Fort de ce résultat et de la formule de König et Huygens, on obtient

$$\begin{aligned}\tfrac{\mathbf{V}(N_n)}{\mathbf{E}(N_n)^2} &= \tfrac{\mathbf{E}\left(N_n(N_n-1)\right)+\mathbf{E}(N_n)-\mathbf{E}(N_n)^2}{\mathbf{E}(N_n)^2}\\ &= \tfrac{\mathbf{E}\left(N_n(N_n-1)\right)}{\mathbf{E}(N_n)^2} + \tfrac{\mathbf{E}(N_n)}{\mathbf{E}(N_n)^2} - 1\\ &\to 1 + 0 - 1 = 0.\end{aligned}$$

On conclut alors par encadrement avec le corollaire VI-1.2 de l'inégalité de Bienaymé et Tchebychev pour la variable N_n à valeurs dans $\mathbb{N}$ que $\mathbf{P}(N_n = 0) \to 0$. ♢

2.2. Remarque. À nouveau, le résultat peut être étendu au cas où l'hypothèse de négligeabilité sur $(p_n)_n$ est remplacée par la condition $p_n \sim c\frac{\ln n}{n}$ pour une constante $c < 1$. L'effet « seuil » de la valeur $\frac{\ln n}{n}$ sera expliqué au chapitre suivant.

3. Asymptotiquement avec triangle

Complétons le résultat asymptotique sur les triangles obtenu au chapitre précédent avec la méthode du premier moment (proposition V-3.1).

3.1. Proposition. *Soit $(p_n)_n$ une suite à valeurs dans $]0,1[$ telle que*

- ▷ $p_n \to 0$,
- ▷ $np_n \to +\infty$.

Notons N_n la variable de comptage des triangles (c'est-à-dire de graphes induits isomorphes à K_3) du graphe aléatoire X_{n,p_n}.
Alors, $\mathbf{P}(N_n = 0) \to 0$.

Démonstration. ▷ Rappelons que l'on a établi au chapitre précédent que $\mathbf{E}(N_n) = \binom{n}{3} p_n^3$. Estimons la variance de N_n. Pour cela, notons, pour tout $A \in \mathscr{P}_3(\llbracket n \rrbracket)$, « A triangle » l'événement « le graphe induit par A dans le graphe X_{n,p_n} est isomorphe à K_3 ». Alors, en utilisant l'écriture de N_n comme somme d'indicatrices, on obtient :

$$\begin{aligned}\mathbf{V}(N_n) &= \mathbf{V}\Big(\sum_{A \in \mathscr{P}_3(\llbracket n \rrbracket)} \mathbb{1}_{A \text{ triangle}} \Big) \\ &= \sum_{A \in \mathscr{P}_3(\llbracket n \rrbracket)} \sum_{B \in \mathscr{P}_3(\llbracket n \rrbracket)} \mathbf{Cov}(\mathbb{1}_{A \text{ triangle}}, \mathbb{1}_{B \text{ triangle}}).\end{aligned}$$

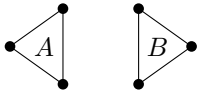
Remarquons que, pour tous $A, B \in \mathscr{P}_3(\llbracket n \rrbracket)$, la covariance des variables $\mathbb{1}_{A \text{ triangle}}$ et $\mathbb{1}_{B \text{ triangle}}$ est égale à

$$\mathbf{E}(\mathbb{1}_{A \text{ triangle}} \, \mathbb{1}_{B \text{ triangle}}) - \mathbf{E}(\mathbb{1}_{A \text{ triangle}}) \mathbf{E}(\mathbb{1}_{B \text{ triangle}})$$

soit

$$\mathbf{P}(A, B \text{ triangles}) - \mathbf{P}(A \text{ triangle}) \mathbf{P}(B \text{ triangle}).$$

▷ Si les triangles définis par A et B n'admettent aucune arête en commun, les variables $\mathbb{1}_{A \text{ triangle}}$ et $\mathbb{1}_{B \text{ triangle}}$ sont indépendantes, donc leur covariance est nulle.



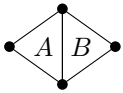
▷ Si les triangles définis par A et B admettent exactement une arête en commun, alors, par indépendance des arêtes,

$$\begin{aligned}\mathbf{P}(A, B \text{ triangles}) &= p_n^5, \\ \mathbf{P}(A \text{ triangle}) = \mathbf{P}(B \text{ triangle}) &= p_n^3,\end{aligned}$$

d'où

$$\mathbf{Cov}(\mathbb{1}_{A \text{ triangle}}, \mathbb{1}_{B \text{ triangle}}) = p_n^5 - p_n^6.$$

Il y a $\binom{n}{4} \times \binom{4}{2} \times 2 = 12\binom{n}{4}$ tels couples (A, B) (choix des quatre sommets, de l'arête commune puis du troisième sommet de A)



▷ Si les triangles définis par A et B admettent au moins deux arêtes en commun, alors $A = B$ et

$$\mathbf{P}(A, B \text{ triangles}) = \mathbf{P}(A \text{ triangle}) = \mathbf{P}(B \text{ triangle}) = p_n^3,$$

d'où

$$\mathbf{Cov}(\mathbb{1}_{A \text{ triangle}}, \mathbb{1}_{B \text{ triangle}}) = p_n^3 - p_n^6.$$

Il y a $\binom{n}{3}$ tels couples (A, B) (car il y a $\binom{n}{3}$ choix d'un ensemble de trois sommets).

En additionnant ces différentes contributions, on obtient

$$\mathbf{V}(N_n) = 0 + 12\tbinom{n}{4}(p_n^5 - p_n^6) + \tbinom{n}{3}p_n^3(1 - p_n^3)$$

et donc la majoration

$$\mathbf{V}(N_n) \leqslant 12\tbinom{n}{4}p_n^5 + \tbinom{n}{3}p_n^3.$$

▷ Fort des calculs précédents, on obtient

$$\begin{aligned}\frac{\mathbf{V}(N_n)}{\mathbf{E}(N_n)^2} &\leqslant \frac{12\binom{n}{4}p_n^5 + \binom{n}{3}p_n^3}{\binom{n}{3}^2 p_n^6} \\ &\leqslant \frac{18(n-3)}{n(n-1)(n-2)p_n} + \frac{6}{n(n-1)(n-2)p_n^3}.\end{aligned}$$

Remarquons alors que le premier terme est négligeable par rapport à $\frac{1}{np_n}$ et que le second est équivalent à $\frac{6}{(np_n)^3}$, donc les deux tendent vers 0 d'après l'hypothèse $np_n \to +\infty$. En conclusion,

$$\frac{\mathbf{V}(N_n)}{\mathbf{E}(N_n)^2} \to 0.$$

▷ On conclut à nouveau avec le corollaire VI-1.2 de l'inégalité de Bienaymé et Tchebychev que $\mathbf{P}(N_n = 0) \to 0$. ♢

4. Exercices

Ce premier exercice montre un raffinement de l'inégalité de Bienaymé et Tchebychev.

4.1. Exercice (Inégalité de Cantelli). *Soit X une variable aléatoire réelle admettant un moment d'ordre 2 et $a > 0$.*

1. *Montrer que*
$$\mathbf{P}\big(X \geqslant \mathbf{E}(X) + a\big) \leqslant \tfrac{\mathbf{V}(X)}{a^2}.$$

2. *En majorant la variable $a + \mathbf{E}(X) - X$ par sa partie positive, montrer l'inégalité de Cantelli suivante :*
$$\mathbf{P}\big(X \geqslant \mathbf{E}(X) + a\big) \leqslant \tfrac{\mathbf{V}(X)}{a^2 + \mathbf{V}(X)}.$$

Solution.

1. L'événement « $X - \mathbf{E}(X) \geqslant a$ » est inclus dans l'événement « $|X - \mathbf{E}(X)| \geqslant a$ », d'où une première inégalité. Majorons ensuite la probabilité de ce nouvel événement avec l'inégalité de Bienaymé et Tchebychev :
$$\begin{aligned}\mathbf{P}\big(X - \mathbf{E}(X) \geqslant a\big) &\leqslant \mathbf{P}\big(|X - \mathbf{E}(X)| \geqslant a\big)\\ &\leqslant \tfrac{\mathbf{V}(X)}{a^2}.\end{aligned}$$

2. Remarquons tout d'abord que
$$(a + \mathbf{E}(X) - X) \leqslant (a + \mathbf{E}(X) - X)\mathbb{1}_{a+\mathbf{E}(X)-X>0},$$
donc, par croissance de l'espérance,
$$\begin{aligned}a &= \mathbf{E}\big(a + \mathbf{E}(X) - X\big)\\ &\leqslant \mathbf{E}\big((a + \mathbf{E}(X) - X)\mathbb{1}_{a+\mathbf{E}(X)-X>0}\big).\end{aligned}$$
Appliquons ensuite l'inégalité de Cauchy et Schwarz :
$$a^2 \leqslant \mathbf{E}\big((a + \mathbf{E}(X) - X)^2\big)\mathbf{P}\big(a + \mathbf{E}(X) - X > 0\big).$$
Notons enfin que
$$\begin{aligned}\mathbf{E}\big((a + \mathbf{E}(X) - X)^2\big) &= a^2 + 2a\mathbf{E}(\mathbf{E}(X) - X) + \mathbf{V}(X)\\ &= a^2 + \mathbf{V}(X).\end{aligned}$$
On en déduit en combinant ces résultats l'inégalité suivante, dont découle immédiatement la formule de l'énoncé :
$$\tfrac{a^2}{a^2+\mathbf{V}(X)} \leqslant 1 - \mathbf{P}\big(X \geqslant \mathbf{E}(X) + a\big).$$

4.2. Exercice. *Pour cet exercice, on admet la formule de Cayley : pour tout entier $n \geqslant 2$, le graphe complet K_n admet exactement n^{n-2} arbres couvrants.*

Soit $(p_n)_n$ une suite à valeurs dans $]0,1[$. Notons N_n la variable de comptage des arbres couvrants du graphe aléatoire X_{n,p_n}.

1. *Montrer que si $np_n \to 0$, alors $\mathbf{P}(N_n = 0) \to 1$.*

   Supposons dorénavant que $np_n \to \ell > 1$.
2. *Montrer que désormais $\mathbf{E}(N_n) \to +\infty$.*
3. *Montrer que la probabilité que le graphe X_{n,p_n} soit connexe tend vers 0.*
4. *En déduire que $\mathbf{E}(N_n^2)$ n'est pas équivalent à $\mathbf{E}(N_n)^2$.*

Solution.

1. Appliquons la méthode du premier moment.

   Chacun des n^{n-2} arbres couvrants de K_n admet $n-1$ arêtes (car c'est un arbre à n sommets ; voir l'exercice I-8.9), donc apparaît dans X_{n,p_n} avec probabilité p_n^{n-1}. Par linéarité de l'espérance,

   $$\mathbf{E}(N_n) = n^{n-2} p_n^{n-1} = p_n (np_n)^{n-2}.$$

   Ainsi, $\mathbf{E}(N_n) \to 0$. Par le corollaire V-1.3 de l'inégalité de Markov, $\mathbf{P}(N_n = 0) \to 1$.
2. En reprenant le calcul de la première question avec la nouvelle hypothèse asymptotique sur $(p_n)_n$, on obtient successivement que $\ln\big(\mathbf{E}(N_n)\big) \to +\infty$, puis $\mathbf{E}(N_n) \to +\infty$.
3. Par hypothèse, $p_n = o\big(\frac{\ln n}{n}\big)$, donc la probabilité qu'il existe au moins un sommet isolé tend vers 1 (d'après la proposition VI-2.1). Par passage au complémentaire, la probabilité qu'il n'existe pas de sommets isolés tend vers 0 et comme cet événement contient l'événement « X_{n,p_n} soit connexe », on en déduit que la probabilité que le graphe X_{n,p_n} soit connexe tend vers 0.
4. En récrivant que l'événement « X_{n,p_n} soit connexe » comme « $N_n = 0$ » (car un graphe est connexe si, et seulement si, il admet un arbre couvrant), on obtient d'après la question précédente que $\mathbf{P}(N_n = 0) \to 1$.

Supposons par l'absurde que $\mathbf{E}(N_n^2) \sim \mathbf{E}(N_n)^2$; d'après la formule de König et Huygens pour la variance,

$$\frac{\mathbf{V}(N_n)}{\mathbf{E}(N_n)^2} = \frac{\mathbf{E}(N_n^2)}{\mathbf{E}(N_n)^2} - 1 \to 0.$$

D'après le corollaire de l'inégalité de Bienaymé et Tchebychev, on obtient $\mathbf{P}(N_n = 0) \to 0$, ce qui contredit le point précédent.

4.3. Exercice (Présence d'un graphe équilibré). *Soit H un graphe à s sommets et a arêtes « équilibré », c'est-à-dire tel que tout sous-graphe H' de H admettant s' sommets et a' arêtes vérifie l'inégalité suivante $\frac{a}{s} \geqslant \frac{a'}{s'}$ (ce qui signifie que H est plus dense en nombre d'arêtes que n'importe lequel de ses sous-graphes).*

Considérons une suite $(p_n)_n$ à valeurs dans $]0,1[$ et notons N_n la variable de comptage des sous-graphes de X_{n,p_n} isomorphes à H.

1. *Justifier que si $n^{s/a}p_n \to 0$, alors $\mathbf{P}(N_n = 0) \to 1$.*
2. *Montrer que la variance de N_n s'écrit sous la forme finie d'une somme indexée par les sous-graphes de H.*

Supposons désormais que $n^{s/a}p_n \to +\infty$.

3. *En déduire que $\mathbf{P}(N_n = 0) \to 0$.*

Solution. Notons $H_1, \dots, H_r$ les occurrences de sous-graphes de K_n isomorphes à H. Le nombre r est déterminé par le choix des s sommets puis par le nombre c de façons d'obtenir un sous-graphe isomorphe à H à partir de ces s sommets (quantité qui ne dépend pas de n et qui a été calculée à l'exercice I-8.6). Ainsi, $r = c\binom{n}{s}$.

Notons, pour tout $i \in [\![r]\!]$, E_i l'événement « le sous-graphe H_i est réalisé dans X_{n,p_n} ».

1. Écrivons N_n comme la somme des indicatrices des différentes apparitions possibles de H :

$$N_n = \sum_{i=1}^{r} \mathbb{1}_{E_i}.$$

Donc, par linéarité de l'espérance,

$$\mathbf{E}(N_n) = \sum_{i=1}^{r} \mathbf{P}(E_i).$$

Or, chaque occurrence apparaît avec probabilité p_n^a (car elle compte a arêtes). Ainsi, $\mathbf{E}(N_n) = c\binom{n}{s}p_n^a \sim \frac{c}{s!}(n^{s/a}p_n)^a$, et

donc $\mathbf{E}(N_n) \to 0$ d'après la condition asymptotique sur p_n. D'après le corollaire V-1.3 de l'inégalité de Markov,

$$\mathbf{P}(N_n = 0) \to 1.$$

2. Pour obtenir la variance de N_n, repartons de l'écriture comme somme d'indicatrices :

$$\mathbf{V}(N_n) = \sum_{i=1}^{r} \sum_{j=1}^{r} \mathbf{Cov}(\mathbb{1}_{E_i}, \mathbb{1}_{E_j}).$$

Calculons la covariance de deux indicatrices des événements E_i et E_j. Plus précisément, en notant $a_{i,j}$ le nombre d'arêtes de l'intersection de H_i et H_j, on obtient

$$\begin{aligned}\mathbf{Cov}(\mathbb{1}_{E_i}, \mathbb{1}_{E_j}) &= \mathbf{E}(\mathbb{1}_{E_i}\mathbb{1}_{E_j}) - \mathbf{E}(\mathbb{1}_{E_i})\mathbf{E}(\mathbb{1}_{E_j}) \\ &= p_n^{2a-a_{i,j}} - p_n^{2a}.\end{aligned}$$

Ainsi,

$$\mathbf{V}(N_n) = \sum_{i=1}^{r} \sum_{j=1}^{r} (p_n^{2a-a_{i,j}} - p_n^{2a}).$$

Rassemblons dans cette double somme tous les couples (H_i, H_j) qui donnent la même intersection. Soit H' un sous-graphe de H. Pour dénombrer les couples (H_i, H_j) dont l'intersection est isomorphe à H', il faut choisir les $2s - s_{H'}$ sommets puis considérer le nombre $c_{H'}$ de façons d'obtenir une configurations correcte à partir de ces sommets. Par conséquent,

$$\mathbf{V}(N_n) = \sum_{H' \subset H} c_{H'} \tbinom{n}{2s-s_{H'}} (p_n^{2a-a_{H'}} - p_n^{2a}).$$

3. ▷ Montrons dans un premier temps que $\frac{\mathbf{V}(N_n)}{\mathbf{E}(N_n)^2} \to 0$.

Pour cela, remarquons que, pour tout H' sous-graphe de H,

$$\begin{aligned}\frac{\binom{n}{2s-s_{H'}}(p_n^{2a-a_{H'}} - p_n^{2a})}{\mathbf{E}(N_n)^2} &= O\Big(\frac{s!^2 n^{2s-s_{H'}} p_n^{2a-a_{H'}}}{c^2(2s-s_{H'})! n^{2s} p_n^{2a}}\Big) \\ &= O\Big(\frac{s!^2}{c^2(2s-s_{H'})!} \frac{1}{n^{s_{H'}} p_n^{a_{H'}}}\Big).\end{aligned}$$

Or, la quantité $n^{s_{H'}} p_n^{a_{H'}} = (n^{s_{H'}/a_{H'}} p_n)^{a_{H'}}$ tend vers $+\infty$, d'après l'hypothèse d'équilibrage de H ($s_{H'}/a_{H'} \geqslant s/a$) et l'hypothèse asymptotique sur $(p_n)_n$.

En conclusion, la fraction considérée tend vers 0 et, par somme finie, $\frac{\mathbf{V}(N_n)}{\mathbf{E}(N_n)^2} \to 0$.

▷ Avec le corollaire VI-1.2 de l'inégalité de Bienaymé et Tchebychev, on obtient que $\mathbf{P}(N = 0) \to 0$.

Le résultat s'étend à un graphe H non équilibré en considérant son sous-graphe H' le plus dense.

Chapitre VII

Fonctions de seuil

Lors des deux chapitres précédents, les exemples ont montré une « transition de phase » pour certains motifs : de part et d'autre d'un certain ordre de grandeur pour p_n, la limite de leurs probabilités d'apparition est soit 0, soit 1. Cela s'avère être un résultat général pour toutes les propriétés monotones, et l'objectif de ce chapitre est de l'établir dans un large cadre.

On rappelle la notation suivante.

0.1. Notation. Soit n un entier naturel non nul, $p \in]0,1[$. Dorénavant, $X_{n,p}$ désigne une variable aléatoire suivant le modèle binomial de paramètre n et p.

1. Propriétés croissantes

1.1. Définition. Soit n un entier naturel non nul. Une propriété croissante sur les graphes d'ensemble de sommets $[\![n]\!]$ est une partie $\mathscr{A}_n$ de l'ensemble des graphes d'ensemble de sommets $[\![n]\!]$ qui vérifie

- ▷ si G appartient à $\mathscr{A}_n$, alors tout graphe isomorphe à G d'ensemble de sommets $[\![n]\!]$ appartient à $\mathscr{A}_n$,
- ▷ si G appartient à $\mathscr{A}_n$, alors tout graphe déduit de G par ajout d'arêtes appartient à $\mathscr{A}_n$.

1.2. Exemple. Il y a deux propriétés croissantes dites triviales sur l'ensemble de sommets $[\![n]\!]$: celle ne contenant aucun graphe et celle les contenant tous. D'ordinaire, on les exclut lorsque l'on énonce des propriétés sur les propriétés croissantes.

1.3. Exemple. ▷ L'ensemble $\mathscr{A}_n$ des graphes d'ensemble de sommets $[\![n]\!]$ admettant un graphe induit isomorphe à K_3 (c'est-à-dire les graphes avec triangle) définit une propriété croissante.

▷ En revanche, son complémentaire n'en est pas une (on pourrait dire qu'il s'agit d'une propriété décroissante) : en ajoutant des arêtes à un graphe qui n'admet pas de graphe induit en triangle, on peut faire apparaître un triangle.

1.4. Proposition. *Soit $\mathscr{A}_n$ une propriété croissante sur les graphes d'ensemble de sommets $[\![n]\!]$.*

Alors, la fonction $p \mapsto \mathbf{P}\big(X_{n,p} \in \mathscr{A}_n\big)$ est continue et croissante sur $]0,1[$.

Démonstration. ▷ En énumérant les éléments de $\mathscr{A}_n$, on obtient que, pour tout p,

$$\mathbf{P}\big(X_{n,p} \in \mathscr{A}_n\big) = \sum_{G \in \mathscr{A}_n} \mathbf{P}\big(X_{n,p} = G\big).$$

Or, chaque terme $\mathbf{P}(X_{n,p} = G)$ est polynomial en p ; la fonction $p \mapsto \mathbf{P}\big(X_{n,p} \in \mathscr{A}_n\big)$ est donc également polynomiale et, en particulier, continue.

▷ Pour la croissance, utilisons un argument de couplage tel qu'il est explicité à la proposition III-2.4. Soit $p \leqslant \tilde{p}$; on peut construire deux graphes aléatoires $X \hookrightarrow \mathscr{G}_{n,p}$ et $\tilde{X} \hookrightarrow \mathscr{G}_{n,\tilde{p}}$ tels que l'ensemble d'arêtes de X soit inclus dans l'ensemble d'arêtes de $\tilde{X}$ et donc tels que l'événement $X \in \mathscr{A}_n$ soit inclus dans l'événement $\tilde{X} \in \mathscr{A}_n$ (par croissance de $\mathscr{A}_n$ puisque $\tilde{X}$ est obtenu en ajoutant des arêtes à X). Par conséquent,

$$\begin{aligned}\mathbf{P}\big(X_{n,p} \in \mathscr{A}_n\big) &= \mathbf{P}\big(X \in \mathscr{A}_n\big)\\ &\leqslant \mathbf{P}\big(\tilde{X} \in \mathscr{A}_n\big) = \mathbf{P}\big(X_{n,\tilde{p}} \in \mathscr{A}_n\big).\end{aligned}$$

Ainsi, la fonction $p \mapsto \mathbf{P}\big(X_{n,p} \in \mathscr{A}_n\big)$ est croissante. ♢

1.5. Remarque. On peut compléter la proposition précédente en remarquant que les limites en 0 et en 1 sont respectivement 0 (sauf pour le cas où $\mathscr{A}_n$ contient le graphe sans arêtes, donc contient tous les graphes) et 1 (sauf pour le cas où $\mathscr{A}_n$ est vide).

1.6. Définition. Une suite de propriétés croissantes est une suite $(\mathscr{A}_n)_n$ telle que, pour tout entier naturel n non nul, $\mathscr{A}_n$ est une propriété croissante sur les graphes d'ensemble de sommets $[\![n]\!]$.

2. Fonction de seuil

2.1. Définition. Soit $(\mathscr{A}_n)_n$ une suite de propriétés croissantes. Une fonction de seuil est une suite $(s_n)_n$ à valeurs dans $[0,1]$ telle que

- ▷ pour toute suite $(p_n)_n$ à valeurs dans $[0,1]$ négligeable devant $(s_n)_n$,
$$\mathbf{P}\big(X_{n,p_n} \in \mathscr{A}_n\big) \to 0,$$
- ▷ pour toute suite $(p_n)_n$ à valeurs dans $[0,1]$ telle que $(s_n)_n$ est négligeable devant $(p_n)_n$,
$$\mathbf{P}\big(X_{n,p_n} \in \mathscr{A}_n\big) \to 1.$$

2.2. Remarque. Il n'y a pas unicité de la fonction de seuil : si $(s_n)_n$ est une fonction de seuil pour une suite de propriétés croissantes $(\mathscr{A}_n)_n$, alors toute suite $(s'_n)_n$ telle qu'il existe des constantes A, B strictement positives vérifiant pour tout entier n assez grand,
$$As_n \leqslant s'_n \leqslant Bs_n,$$
c'est-à-dire $s'_n = \Theta(s_n)$, est aussi une fonction de seuil pour la suite $(\mathscr{A}_n)_n$.

Il suffit pour cela de remarquer qu'une suite est négligeable devant (respectivement domine) $(s_n)_n$ si, et seulement si, elle est négligeable devant (respectivement domine) $(s'_n)_n$.

2.3. Remarque. On définit quelquefois une notion plus fine de seuil étroit : une suite $(s_n)_n$ est un seuil étroit pour la suite $(\mathscr{A}_n)_n$ de propriétés croissantes si, pour tout $\varepsilon > 0$,

- ▷ pour toute suite $(p_n)_n$ telle que $p_n \leqslant (1-\varepsilon)s_n$ à partir d'un certain rang,
$$\mathbf{P}\big(X_{n,p_n} \in \mathscr{A}_n\big) \to 0,$$
- ▷ pour toute suite $(p_n)_n$ telle que $p_n \geqslant (1+\varepsilon)s_n$ à partir d'un certain rang,
$$\mathbf{P}\big(X_{n,p_n} \in \mathscr{A}_n\big) \to 1.$$

Fidèle à l'objectif introductif de ce texte, on ne cherchera pas à établir si les seuils étudiés par la suite sont ou non étroits.

2.4. Exemple. Notons, pour tout entier n non nul, $\mathscr{A}_n$ la propriété qu'un graphe d'ensemble de sommets $[\![n]\!]$ admette au moins une arête. En passant par l'événement complémentaire, on obtient immédiatement, pour tout $p \in]0,1[$,

$$\mathbf{P}(X_{n,p} \in \mathscr{A}_n) = 1 - (1-p)^{\binom{n}{2}}.$$

En utilisant que $\binom{n}{2} \sim \frac{n^2}{2}$ et donc que, pour une suite $(p_n)_n$ de limite nulle,

$$\ln\big((1-p_n)^{\binom{n}{2}}\big) = \tbinom{n}{2}\ln(1-p_n) \sim -\tfrac{n^2 p_n}{2},$$

on déduit que

- ▷ pour toute suite $(p_n)_n$ telle que $n^2 p_n \to 0$,
$$\mathbf{P}(X_{n,p_n} \in \mathscr{A}_n) \to 0,$$
- ▷ pour toute suite $(p_n)_n$ telle que $n^2 p_n \to +\infty$,
$$\mathbf{P}(X_{n,p_n} \in \mathscr{A}_n) \to 1.$$

Ainsi, une fonction de seuil est obtenue avec les valeurs $\frac{1}{n^2}$.

2.5. Exemple. Notons, pour tout entier n non nul, $\mathscr{A}_n$ la propriété qu'un graphe d'ensemble de sommets $[\![n]\!]$ admette un graphe induit triangle. Alors, d'après les chapitres précédents, une fonction de seuil est obtenue avec les valeurs $\frac{1}{n}$.

Plus généralement, une fonction de seuil pour qu'un graphe d'ensemble de sommets $[\![n]\!]$ admette un graphe induit isomorphe à K_k est obtenue avec les valeurs $n^{-2/(k-1)}$.

Énonçons quelques résultats sur les probabilités associées aux propriétés croissantes, qui seront utilisés pour la preuve du théorème d'existence à la section suivante.

2.6. Lemme. *Soit n un entier naturel non nul, $\mathscr{A}_n$ une propriété croissante, $p \in]0,1[$ et k un entier naturel tel que $kp \in]0,1[$. Alors,*

$$\mathbf{P}(X_{n,kp} \in \mathscr{A}_n) \geqslant 1 - \big(1 - \mathbf{P}(X_{n,p} \in \mathscr{A}_n)\big)^k.$$

Démonstration. Soit $X^{(1)}, \dots, X^{(k)}$ des variables aléatoires mutuellement indépendantes suivant le modèle binomial de paramètres n et p.

Alors, d'après la Proposition III-2.4, la variable aléatoire définie comme la réunion de ces graphes suit le modèle binomial de paramètres n et $1-(1-p)^k$. Remarquons la majoration suivante [1] :

$$1 - (1-p)^k \leqslant 1 - (1 - kp) = kp.$$

Ainsi, d'après la propriété de croissance de la propriété $\mathscr{A}_n$,

$$\begin{aligned}
\mathbf{P}(X_{n,kp} \in \mathscr{A}_n) &\geqslant \mathbf{P}\Big(\bigcup_{i=1}^{k} X^{(i)} \in \mathscr{A}_n\Big) \\
&\geqslant \mathbf{P}(\exists i \in [\![k]\!],\ X^{(i)} \in \mathscr{A}_n) \\
&= 1 - \mathbf{P}(\forall i \in [\![k]\!],\ X^{(i)} \notin \mathscr{A}_n) \\
&= 1 - \mathbf{P}\big(X^{(1)} \notin \mathscr{A}_n\big)^k \\
&= 1 - \big(1 - \mathbf{P}(X_{n,p} \in \mathscr{A}_n)\big)^k.
\end{aligned}$$

♢

1. Il s'agit simplement de l'inégalité de Bernoulli : pour tout entier naturel n non nul et tout réel $x \geqslant -1$, $(1+x)^n \geqslant 1 + nx$.

2.7. Corollaire. *Soit n un entier naturel non nul, $\mathscr{A}_n$ une propriété croissante, $p \in]0,1[$ et k un entier naturel non nul. Alors,*

$$\mathbf{P}(X_{n,p/k} \in \mathscr{A}_n) \leqslant 1 - \big(1 - \mathbf{P}(X_{n,p} \in \mathscr{A}_n)\big)^{1/k}.$$

Démonstration. D'après le lemme précédent en partant de la probabilité $\frac{p}{k}$,

$$\mathbf{P}(X_{n,p} \in \mathscr{A}_n) \geqslant 1 - \big(1 - \mathbf{P}(X_{n,p/k} \in \mathscr{A}_n)\big)^{k}.$$

Il suffit alors d'exprimer le terme de droite à partir du terme de gauche pour obtenir

$$\big(1 - \mathbf{P}(X_{n,p/k} \in \mathscr{A}_n)\big)^{k} \geqslant 1 - \mathbf{P}(X_{n,p} \in \mathscr{A}_n),$$

d'où découle l'inégalité souhaitée. ♢

3. Théorème d'existence

3.1. Théorème (Bollobás-Thomason [BT87]). *Toute suite $(\mathscr{A}_n)_n$ de propriétés croissantes non triviales admet une fonction de seuil.*

Démonstration. Soit n un entier non nul. La fonction

$$p \mapsto \mathbf{P}(X_{n,p} \in \mathscr{A}_n) = \sum_{([\![n]\!],A)\in\mathscr{A}_n} p^{|A|}(1-p)^{\binom{n}{2}-|A|}$$

est

- continue (car polynomiale), croissante de $[0,1]$ dans $[0,1]$ (proposition VII-1.4),
- vaut 0 en 0 (car $\mathscr{A}_n$ étant non triviale, le graphe sans arête n'appartient pas à $\mathscr{A}_n$),
- vaut 1 en 1.

Ainsi, d'après le théorème des valeurs intermédiaires, il existe s_n un élément de $]0,1[$ tel que

$$\mathbf{P}(X_{n,s_n} \in \mathscr{A}_n) = \tfrac{1}{2}.$$

Montrons que la suite $(s_n)_n$ fournit une fonction seuil.

▷ Soit une suite $(p_n)_n$ à valeurs dans $[0,1]$ telle que $s_n = o(p_n)$: il existe une suite $(k_n)_n$ d'entiers divergeant vers $+\infty$ telle que, pour tout n, $p_n \geqslant k_n s_n$. Alors, pour tout n, la croissance de $\mathscr{A}_n$, le lemme VII-2.6 et la définition de s_n donnent successivement

$$\begin{aligned}\mathbf{P}(X_{n,p_n} \in \mathscr{A}_n) &\geqslant \mathbf{P}(X_{n,k_n s_n} \in \mathscr{A}_n)\\ &\geqslant 1 - \big(1 - \mathbf{P}(X_{n,s_n} \in \mathscr{A}_n)\big)^{k_n}\\ &= 1 - 2^{-k_n}.\end{aligned}$$

Ce minorant tend vers 1 donc, par encadrement (une probabilité est toujours inférieure à 1), $\mathbf{P}(X_{n,p_n} \in \mathscr{A}_n) \to 1$.

▷ Soit une suite $(p_n)_n$ à valeurs dans $[0,1]$ telle que $p_n = o(s_n)$: il existe une suite $(k_n)_n$ d'entiers divergeant vers $+\infty$ telle que, pour tout n, $s_n \geqslant k_n p_n$. On procède comme dans le premier point mais en remplaçant cette fois-ci le lemme par son corollaire.

Pour tout n, on obtient

$$\begin{aligned}\mathbf{P}(X_{n,p_n} \in \mathscr{A}_n) &\leqslant \mathbf{P}(X_{n,s_n/k_n} \in \mathscr{A}_n)\\ &\leqslant 1 - \big(1 - \mathbf{P}(X_{n,s_n} \in \mathscr{A}_n)\big)^{1/k_n}\\ &= 1 - 2^{-1/k_n}.\end{aligned}$$

Ce majorant tend vers 0 donc, par encadrement (une probabilité est toujours positive), $\mathbf{P}(X_{n,p_n} \in \mathscr{A}_n) \to 0$. ♢

3.2. Remarque. Le choix de $(s_n)_n$ dans la démonstration précédente semble accorder un rôle particulier à la probabilité $\frac{1}{2}$: il n'en est rien ! On aurait pu choisir n'importe quel $q \in]0,1[$ puis définir, pour tout n, $s_n \in]0,1[$ tel que $\mathbf{P}(X_{n,s_n} \in \mathscr{A}_n) = q$. La démonstration aurait ensuite été identique.

4. Exercices

4.1. Exercice. *Parmi les ensembles suivants, lesquels sont des propriétés croissantes :*

1. *les graphes d'ensemble $[\![n]\!]$ de sommets connexes,*

2. *les graphes d'ensemble $[\![n]\!]$ de sommets connexes de diamètre au plus d,*
3. *les graphes d'ensemble $[\![n]\!]$ de sommets connexes de diamètre au moins $d > 1$.*

Solution.

1. Les graphes connexes d'ensemble de sommets $[\![n]\!]$ forment bien une propriété croissante : en effet, tout graphe isomorphe à un graphe connexe est connexe, et ajouter des arêtes à un graphe connexe ne lui enlève pas sa connexité (le chemin qui joignait deux sommets dans le graphe initiale existe encore dans le graphe augmenté).
2. Cet ensemble de graphes est une propriété croissante : elle est clairement invariante par isomorphisme. Par ailleurs, rajouter des arêtes ne peut que faire diminuer la distance entre deux sommets donnés (en créant de nouveaux chemins les joignant) donc faire baisser le diamètre.

   Ainsi, tout graphe déduit d'un graphe de diamètre au plus d par l'ajout d'arêtes est de diamètre au plus d.
3. Cet ensemble de graphes n'est pas une propriété croissante : considérons un graphe de diamètre d et ajoutons toutes les arêtes manquantes ; on obtient ainsi un graphe complet de diamètre $1 < d$ (puisque chaque sommet est à distance 1 de tous les autres), donc qui n'appartient plus à l'ensemble considéré.

L'exercice suivant remplace la proposition VII-1.4 pour le modèle uniforme et permet ensuite de développer une théorie de fonctions de seuil dans le cadre de ce modèle.

4.2. Exercice. *Soit $\mathscr{A}_n$ une propriété croissante sur les graphes d'ensemble de sommets $[\![n]\!]$. Notons, $N = \binom{n}{2}$ et, pour tout entier $m \leqslant N$, $X_{n,m}$ une variable aléatoire suivant le modèle uniforme de paramètres n et m.*

Montrer que la suite finie $\big(\mathbf{P}(X_{n,m} \in \mathscr{A}_n)\big)_m$ est croissante.

Solution. Soit $m \leqslant m'$ deux entiers inférieurs ou égaux à N. L'idée est alors de « coupler » les variables $X_{n,m}$ et $X_{n,m'}$. Pour cela, partons de $X_{n,m}$ et considérons un graphe aléatoire G d'ensemble de sommets $[\![n]\!]$ et admettant $m' - m$ arêtes uniformément parmi les arêtes qui n'appartiennent pas à $X_{n,m}$.

▷ Montrons que le graphe X' obtenu comme réunion de $X_{n,m}$ et G, ou plutôt de leurs ensembles d'arêtes notés A_m et A_G respectivement, suit le modèle binomial de paramètres n et m'.

En effet, pour toute partie B de m' arêtes, la probabilité que le graphe X' admette B comme ensemble d'arêtes est

$$\begin{aligned}
&\mathbf{P}(A_m \cup A_G = B) \\
&\qquad = \sum_{C \in \mathscr{P}_m(B)} \mathbf{P}(A_G = B \setminus C | A_m = C)\mathbf{P}(A_m = C) \\
&\qquad = \sum_{C \in \mathscr{P}_m(B)} \tfrac{1}{\binom{N-m}{m'-m}} \tfrac{1}{\binom{N}{m}} \\
&\qquad = \tbinom{m'}{m} \tfrac{1}{\binom{N-m}{m'-m}\binom{N}{m}} = \tfrac{1}{\binom{N}{m'}}.
\end{aligned}$$

▷ Ensuite, exploitons la croissance de la propriété $\mathscr{A}_n$ pour obtenir l'inégalité attendue :

$$\begin{aligned}
\mathbf{P}\big(X_{n,m} \in \mathscr{A}_n\big) &\leqslant \mathbf{P}\big(X' \in \mathscr{A}_n\big) \\
&= \mathbf{P}\big(X_{n,m'} \in \mathscr{A}_n\big).
\end{aligned}$$

4.3. Exercice (Seuil pour la non planarité). *Un graphe G est planaire s'il existe un dessin de G dans lequel les arêtes ne s'intersectent qu'en leurs extrémités communes, c'est-à-dire si le nombre de croisements (tel que défini en IV-4.1) est nul.*

1. *Justifier que l'ensemble $\mathscr{A}_n$ des graphes d'ensemble de sommets $[\![n]\!]$ non planaires définit une propriété croissante.*

   Soit $(p_n)_n$ une suite à valeurs dans $]0,1[$.

2. *Supposons dans cette question $p_n = o(\frac{1}{n})$. Montrer que*

$$\mathbf{P}\big(X_{n,p_n} \in \mathscr{A}_n\big) \to 0.$$

3. *Supposons dans cette question $\frac{1}{n} = o(p_n)$. Montrer que*

$$\mathbf{P}(\chi(X_{n,p_n}) \geqslant 7) \to 1,$$

   puis

$$\mathbf{P}\big(X_{n,p_n} \in \mathscr{A}_n\big) \to 1.$$

   On pourra s'inspirer de l'exercice IV-5.9 et utiliser sans dé-

monstration[2] *qu'un graphe planaire a un nombre chromatique inférieur ou égal à* 6.

4. *En déduire une fonction seuil pour la propriété de non planarité.*

Solution.

1. Soit G un graphe d'ensemble de sommets $[\![n]\!]$ non planaire et G' le graphe déduit de G en ajoutant une arête a. Si G' est planaire, on dispose d'un dessin de G' sans croisement ; alors en enlevant l'arête a, on obtient un dessin de G sans croisement, ce qui contredit l'hypothèse. Ainsi, G' appartient bien à $\mathscr{A}_n$. En conclusion, $\mathscr{A}_n$ est une propriété croissante.
2. Remarquons qu'un graphe sans cycle est planaire. Par conséquent, l'événement $X_{n,p_n} \in \mathscr{A}_n$ est inclus dans la réunion des événements "le graphe X_{n,p_n} admet un cycle de longueur l" lorsque l parcourt les entiers entre 3 et n. Ainsi,

$$\begin{aligned}
\mathbf{P}\big(X_{n,p_n} \in \mathscr{A}_n\big) &\leqslant \sum_{l=3}^{n} \mathbf{P}\big(X_{n,p_n} \text{ admet un } l\text{-cycle}\big) \\
&\leqslant \sum_{l=3}^{n} \tbinom{n}{l} \tfrac{1}{2}(l-1)! p_n^l \\
&\leqslant \sum_{l=3}^{n} \tfrac{n(n-1)\cdots(n-l+1)}{2l} p_n^l \\
&\leqslant \sum_{l=3}^{n} (np_n)^l \\
&\leqslant \sum_{l=3}^{+\infty} (np_n)^l \\
&\leqslant \tfrac{(np_n)^3}{1-np_n}.
\end{aligned}$$

   Comme $np_n \to 0$, $\mathbf{P}\big(X_{n,p_n} \in \mathscr{A}_n\big) \to 0$ avec le lemme d'encadrement.
3. D'après le résultat admis, $(\chi(X_{n,p_n}) \geqslant 7) \subset \big(X_{n,p_n} \in \mathscr{A}_n\big)$, donc
$$\mathbf{P}\big(X_{n,p_n} \in \mathscr{A}_n\big) \geqslant \mathbf{P}(\chi(X_{n,p_n}) \geqslant 7).$$
   D'après la proposition I-7.6, $\chi(X_{n,p_n})\alpha(X_{n,p_n}) \geqslant n$, d'où l'in-

2. Ce résultat qui se démontre simplement par récurrence n'est pas optimal.

clusion $(n/\alpha(X_{n,p_n}) \geqslant 7) \subset (\chi(X_{n,p_n}) \geqslant 7)$, puis

$$\mathbf{P}\big(X_{n,p_n} \in \mathscr{A}_n\big) \geqslant \mathbf{P}(\alpha(X_{n,p_n}) \leqslant n/7).$$

Or,

$$\begin{aligned}\mathbf{P}(\alpha(X_{n,p_n}) > n/7) &= \mathbf{P}(\alpha(X_{n,p_n}) > \lfloor n/7 \rfloor) \\ &\leqslant \tbinom{n}{\lfloor n/7 \rfloor}(1-p_n)^{\binom{\lfloor n/7 \rfloor}{2}} \\ &\leqslant 2^n \exp\big(-p_n\tbinom{\lfloor n/7 \rfloor}{2}\big) \\ &\leqslant \exp\big(n\ln 2 - p_n\tbinom{\lfloor n/7 \rfloor}{2}\big).\end{aligned}$$

Comme l'argument de l'exponentielle est de limite $-\infty$, ce majorant tend vers 0. Par encadrement, $\mathbf{P}(\alpha(X_{n,p_n}) > n/7) \to 0$, donc $\mathbf{P}(\alpha(X_{n,p_n}) \leqslant n/7) \to 1$. En conclusion,

$$\mathbf{P}\big(X_{n,p_n} \in \mathscr{A}_n\big) \to 1.$$

4. D'après les deux questions précédentes, $\frac{1}{n}$ définit une fonction de seuil pour la suite de propriétés croissantes $(\mathscr{A}_n)_n$ de non planarité.

4.4. Exercice (Seuil pour la connexité).

1. *Justifier que, pour une suite $(p_n)_n$ à valeurs dans $]0,1[$ telle que $p_n = o\big(\frac{\ln n}{n}\big)$, la probabilité que le graphe aléatoire X_{n,p_n} soit connexe tend vers 0.*
2. *Montrer que, pour tout $p \in]0,1[$,*

$$\mathbf{P}\big(X_{n,p} \text{ non connexe}\big) \leqslant \sum_{k=1}^{\lfloor n/2 \rfloor} \tbinom{n}{k}(1-p)^{k(n-k)}.$$

3. *Justifier que, pour tout entier k entre 1 et $\frac{n}{2}$,*

$$k(n-k) \geqslant \tfrac{1}{2}\big(k(n-2)+n\big).$$

4. *En déduire que, pour $p_n = c\frac{\ln n}{n}$ avec $c > 2$, la probabilité que le graphe aléatoire X_{n,p_n} soit connexe tend vers 1.*
5. *Notons, pour tout n, $\mathscr{A}_n$ l'ensemble des graphes connexes d'ensemble de sommets $[\![n]\!]$.*

   Déterminer une fonction de seuil pour la suite de propriétés croissantes $(\mathscr{A}_n)_n$.

Solution.

1. Nous avons déjà établi lors de l'étude des sommets isolés (et cela est détaillé à la question 3 de l'exercice VI-4.2) que si une suite $(p_n)_n$ à valeurs dans $]0,1[$ vérifie $p_n = o\big(\frac{\ln n}{n}\big)$, alors la probabilité que le graphe X_{n,p_n} soit connexe tend vers 0.
2. Remarquons que le graphe $X_{n,p}$ n'est pas connexe si, et seulement si, il admet plusieurs composantes connexes, donc s'il existe une composante connexe de cardinal au plus $\frac{n}{2}$. Par conséquent, le graphe $X_{n,p}$ n'est pas connexe si, et seulement si, il existe un entier $k \leqslant \frac{n}{2}$ et une partie A de cardinal k qui n'est pas reliée à son complémentaire.

   D'après cette caractérisation, on obtient la majoration :
   $\mathbf{P}\big(X_{n,p_n}$ non connexe$\big)$

$$\begin{aligned}
&\leqslant \sum_{k=1}^{\lfloor n/2 \rfloor} \sum_{A \in \mathscr{P}_k([\![n]\!])} \mathbf{P}\big(A \text{ non reliée à } \overline{A}\big) \\
&= \sum_{k=1}^{\lfloor n/2 \rfloor} \sum_{A \in \mathscr{P}_k([\![n]\!])} (1-p)^{k(n-k)} \\
&= \sum_{k=1}^{\lfloor n/2 \rfloor} \tbinom{n}{k} (1-p)^{k(n-k)},
\end{aligned}$$

   où l'on a utilisé qu'il existait $k(n-k)$ arêtes entre une partie de cardinal k et son complémentaire de cardinal $n-k$.
3. Remarquons que, pour les entiers considérés, $(k-1)(\frac{n}{2}-k) \geqslant 0$ et que cette inégalité se récrit $k(n-k) \geqslant \frac{1}{2}(k(n-2)+n)$.
4. En reportant dans la formule de la question 2, on obtient que
   $\mathbf{P}\big(X_{n,p_n}$ non connexe$\big)$

$$\begin{aligned}
&\leqslant (1-p_n)^{n/2} \sum_{k=1}^{\lfloor n/2 \rfloor} \tbinom{n}{k} (1-p_n)^{k(n-2)/2} \\
&\leqslant (1-p_n)^{n/2} \big(1 + (1-p_n)^{(n-2)/2}\big)^n.
\end{aligned}$$

   Étudions désormais la limite du majorant en séparant les deux facteurs du produit.

▷ Tout d'abord, le premier facteur est

$$(1-p_n)^{n/2} = \exp\left(\tfrac{n}{2}\ln(1-p_n)\right)$$

et l'argument de l'exponentielle est équivalent à $-\frac{np_n}{2}$, donc à $-\frac{c}{2}\ln n$: ce premier facteur est de limite nulle.

▷ Le logarithme du second facteur est

$$n\ln\left(1+(1-p_n)^{(n-2)/2}\right) \sim n(1-p_n)^{(n-2)/2}.$$

Enfin, le logarithme du membre de droite de cet équivalent est

$$\begin{aligned}\ln n + \tfrac{n-2}{2}\ln(1-p_n) &= \ln n + \tfrac{n-2}{2}\left(-c\tfrac{\ln n}{n} + o\left(\tfrac{\ln n}{n}\right)\right)\\ &= \left(1-\tfrac{c}{2}\right)\ln n + o(\ln n).\end{aligned}$$

Comme $c > 2$, la limite de ce terme est $-\infty$ et par suite

$$n(1-p_n)^{(n-2)/2} \to 0$$

puis

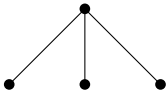
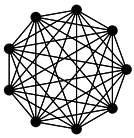
$$\left(1+(1-p_n)^{(n-2)/2}\right)^n \to 1.$$

▷ Combinant ces deux résultats, on en déduit par encadrement que $\mathbf{P}\big(X_{n,p_n}$ non connexe$\big) \to 0$, c'est-à-dire la probabilité que le graphe aléatoire X_{n,p_n} soit connexe tend vers 1.

5. D'après la question précédente et avec la propriété de croissance, on obtient que, pour toute suite $(p_n)_n$ à valeurs dans $]0,1[$ et telle que $\frac{np_n}{\ln n} \to +\infty$, la probabilité que le graphe aléatoire X_{n,p_n} soit connexe tend vers 1. En combinant ce résultat avec la première question, on en déduit que $\frac{\ln n}{n}$ définit une fonction de seuil pour la suite de propriétés croissantes $(\mathscr{A}_n)_n$.

5. Récapitulatif

À travers le texte et les exercices précédents, on a exhibé un certain nombre de fonctions de seuil. Par exemple, un arbre à k sommets (et donc $k-1$ arêtes d'après l'exercice I-8.9) est un graphe équilibré : on obtient une fonction de seuil avec l'exercice VI-4.3. Dressons un petit tableau récapitulatif en « triant » les fonctions de seuil selon la relation de négligeabilité.

La suite …	est un seuil pour …
$\frac{1}{n^2}$	l'apparition d'une arête
$\frac{1}{n^{3/2}}$	l'apparition de deux arêtes adjacentes (c'est-à-dire d'un chemin injectif de longueur 2)
$\frac{1}{n^{4/3}}$	l'apparition d'une griffe $K_{1,3}$ 
$\frac{1}{n^{k/(k-1)}}$	l'apparition d'une arbre donné à $k \geqslant 4$ sommets
$\frac{1}{n}$	l'apparition d'un triangle K_3
	l'apparition d'un cycle (indépendamment de la longueur)
	la disparition de la planarité
$\frac{1}{n^{2/(k-1)}}$	l'apparition d'un K_k 

La suite …	est un seuil pour …
$\frac{\ln n}{n}$	la disparition des sommets isolés
	l'apparition de la connexité

5.1. Remarque. On constate en particulier qu'il existe asymptotiquement de grands arbres avant qu'il n'existe le moindre cycle. Notons également qu'un seuil d'apparition de la connexité est également un seuil de disparition des points isolés.

Chapitre VIII

Probabilité au seuil

Pour comprendre plus finement ce qui se passe pour une propriété croissante « au seuil », il faut dire quelques mots de l'approximation poisonnienne, ou de manière équivalente, d'un phénomène de convergence vers la loi de Poisson.

La principale méthode consiste à calculer les moments factoriels d'une variable aléatoire de comptage, puis de voir comment ils peuvent servir pour établir la convergence d'une suite de variables aléatoires vers la loi de Poisson.

Dans ce court chapitre, on précise quelques propriétés élémentaires puis on illustre comment le calcul des moments factoriels est pertinent pour des propriétés croissantes considérées « au seuil ».

1. Moments factoriels

1.1. Définition. Soit X une variable aléatoire à valeurs dans $\mathbb{N}$. Pour tout $k \in \mathbb{N}^*$, le k-ième moment factoriel de X est, quand elle existe, l'espérance

$$\mathbf{E}((X)_k) = \mathbf{E}(X(X-1)\cdots(X-(k-1))).$$

Par convention, le 0-ième moment factoriel est $\mathbf{E}((X)_0) = 1$.

1.2. Remarque. Pour une variable aléatoire X à valeurs dans $\mathbb{N}$, on peut montrer que l'existence du k-ième moment factoriel équivaut à la dérivabilité k fois en 1 de sa fonction génératrice G_X ; dans ce cas, $\mathbf{E}((X)_k) = G_X^{(k)}(1)$.

Le calcul fondamental pour ce chapitre est celui des moments factoriels de la loi de Poisson.

1.3. Proposition. *Soit $\lambda \in \mathbb{R}_+^*$ et X une variable aléatoire suivant la loi de Poisson de paramètre λ. Alors, X admet des moments factoriels de tout ordre et, pour tout $k \in \mathbb{N}$,*

$$\mathbf{E}((X)_k) = \lambda^k.$$

Démonstration. Soit $k \in \mathbb{N}$. La série de terme général (pour $j \geqslant k$)

$$j(j-1)\cdots(j-k+1)\mathbf{P}(X=j) = \lambda^k \tfrac{\lambda^{j-k}}{(j-k)!} e^{-\lambda}$$

est convergente donc X admet un moment factoriel d'ordre k.

D'après la formule de transfert pour la variable X et la fonction $x \mapsto x(x-1)\cdots(x-k+1)$,

$$\begin{aligned}\mathbf{E}((X)_k) &= \sum_{j=0}^{+\infty} j(j-1)\cdots(j-k+1)\mathbf{P}(X=j)\\ &= \lambda^k \sum_{j=k}^{+\infty} \tfrac{\lambda^{j-k}}{(j-k)!} e^{-\lambda}\\ &= \lambda^k.\end{aligned}$$

♢

Avec les deux propositions suivantes, on explicite les liens entre les moments factoriels d'une variable et la loi de celle-ci.

1.4. Proposition. *Soit X une variable aléatoire à valeurs dans $\mathbb{N}$ admettant des moments de tout ordre. Alors, pour tout $k \in \mathbb{N}$,*

$$\mathbf{E}((X)_k) = \sum_{j=0}^{+\infty} \tfrac{(k+j)!}{j!} \mathbf{P}(X = k+j).$$

Démonstration. Appliquons la formule de transfert pour la variable X et la fonction positive $x \mapsto x(x-1)\cdots(x-k+1)$:

$$\begin{aligned}\mathbf{E}((X)_k) &= \sum_{l=0}^{+\infty} l(l-1)\cdots(l-k+1)\mathbf{P}(X=l)\\ &= \sum_{l=k}^{+\infty} \tfrac{l!}{(l-k)!}\mathbf{P}(X=l)\\ &= \sum_{j=0}^{+\infty} \tfrac{(k+j)!}{j!}\mathbf{P}(X=k+j).\end{aligned}$$

♢

1.5. Proposition. *Soit X une variable aléatoire bornée et à valeurs dans $\mathbb{N}$. Alors, pour tout $k \in \mathbb{N}$,*

$$\mathbf{P}(X=k) = \sum_{l=0}^{+\infty} \tfrac{(-1)^l}{k!l!}\mathbf{E}((X)_{k+l}).$$

Démonstration. Fixons $k \in \mathbb{N}$. Étudions le membre de droite en utilisant l'expression des moments factoriels de la proposition précédente.
La famille $\left(\frac{(-1)^l(k+l+j)!}{k!l!j!}\mathbf{P}(X=k+l+j)\right)_{l,j}$ est finie (car X est bornée) donc sommable. Appliquons ensuite le théorème de sommation par paquets pour les familles sommables en posant $s = k+l+j$:

$$\begin{aligned}\sum_{l=0}^{+\infty} \tfrac{(-1)^l}{k!l!}\mathbf{E}((X)_{k+l}) &= \sum_{l=0}^{+\infty} \tfrac{(-1)^l}{k!l!}\sum_{j=0}^{+\infty} \tfrac{(k+l+j)!}{j!}\mathbf{P}(X=k+l+j)\\ &= \sum_{s=k}^{+\infty} \tbinom{s}{k}\mathbf{P}(X=s)\sum_{l=0}^{s-k}\tbinom{s-k}{l}(-1)^l\\ &= \sum_{s=k}^{+\infty} \tbinom{s}{k}\mathbf{P}(X=s)\delta s-k,0\\ &= \mathbf{P}(X=k).\end{aligned}$$

♢

1.6. Remarque. On vérifie immédiatement que ces relations sont aussi valables pour les variables de Poisson : si X une variable aléatoire suivant la loi de Poisson de paramètre λ et $k \in \mathbb{N}$,

$$\begin{aligned}\mathbf{P}(X=k) &= \tfrac{\lambda^k}{k!}e^{-\lambda} \\ &= \tfrac{\lambda^k}{k!}\sum_{l=0}^{+\infty}\tfrac{(-1)^l\lambda^l}{l!} \\ &= \sum_{l=0}^{+\infty}\tfrac{(-1)^l}{k!l!}\mathbf{E}((X)_{k+l}).\end{aligned}$$

1.7. Proposition. *Soit X une variable aléatoire bornée et à valeurs dans $\mathbb{N}$. Alors, pour tout $p \in \mathbb{N}$,*

$$\Big|\mathbf{P}(X=0) - \sum_{l=0}^{p}\tfrac{(-1)^l}{l!}\mathbf{E}((X)_l)\Big| \leqslant \tfrac{1}{(p+1)!}\mathbf{E}((X)_{p+1}).$$

Démonstration. ▷ Commençons par un calcul sur les variables aléatoires : pour tout $p \in \mathbb{N}$,

$$\begin{aligned}\sum_{l=0}^{p}\tfrac{(-1)^l}{l!}(X)_l - \mathbb{1}_{X=0} &= \sum_{l=1}^{p}\tfrac{(-1)^l}{l!}(X)_l + \mathbb{1}_{X\geqslant 1} \\ &= \mathbb{1}_{X\geqslant 1}\Big(\sum_{l=1}^{p}\tfrac{(-1)^l}{l!}(X)_l + 1\Big) \\ &= \mathbb{1}_{X\geqslant 1}\sum_{l=0}^{p}\tfrac{(-1)^l}{l!}(X)_l \\ &= \mathbb{1}_{X\geqslant 1}\sum_{l=0}^{p}(-1)^l\tbinom{X}{l} \\ &= \mathbb{1}_{X\geqslant 1}(-1)^p\tbinom{X-1}{p},\end{aligned}$$

en utilisant le résultat déterministe (facilement obtenu par récurrence) que pour tous n et p, $\sum_{l=0}^{p}(-1)^l\binom{n}{l} = (-1)^p\binom{n-1}{p}$.

Passons désormais à l'espérance : pour tout $p \in \mathbb{N}$, la quantité $\sum_{l=0}^{p}\frac{(-1)^l}{l!}\mathbf{E}((X)_l) - \mathbf{P}(X=0)$ est du signe de $(-1)^p$.

▷ Les suites extraites d'indices pairs et d'indices impairs de la suite $\left(\sum_{l=0}^{p} \frac{(-1)^l}{l!}\mathbf{E}((X)_l)\right)_p$ sont adjacentes et leur limite commune est $\mathbf{P}(X=0)$: le résultat d'encadrement de la proposition en découle immédiatement. ♢

2. Cas des variables de comptage

Avant de passer aux propriétés croissantes sur les graphes, il faut également comprendre les moments factoriels des variables de comptage.

2.1. Proposition. *Soit $(A_j)_{1\leqslant j\leqslant n}$ des événements et N la variable de comptage associée (c'est-à-dire $N=\sum_{j=1}^{n}\mathbb{1}_{A_j}$). Alors, pour tout $k\in\mathbb{N}^*$, la variable $(N)_k$ est la variable de comptage des intersections de k événements de la famille, à savoir*

$$(N)_k = \sum_{j_1,\dots,j_k\in[\![1,n]\!] \text{ distincts}} \mathbb{1}_{\bigcap_{l=1}^{k} A_{j_l}}.$$

En particulier, X admet des moments factoriels de tout ordre et, pour tout $k\in\mathbb{N}^*$,

$$\mathbf{E}((N)_k) = \sum_{j_1,\dots,j_k\in[\![1,n]\!] \text{ distincts}} \mathbf{P}\Big(\bigcap_{l=1}^{k} A_{j_l}\Big).$$

Démonstration. ▷ Soit $\omega\in\Omega$, $J=\{j\in[\![1,n]\!],\ \omega\in A_j\}$ et m le cardinal de J. Fixons un entier $k\in\mathbb{N}^*$.
∘ D'une part, $N(\omega)=m$ et donc $(N)_k(\omega)=(m)_k$.
∘ D'autre part, la variable X de comptage des intersections de k événements distincts parmi $(A_j)_{1\leqslant j\leqslant n}$ prend en ω la valeur du nombre de k-uplets d'indices distincts dans $[\![1,n]\!]$ correspondants à des événements réalisés par ω, c'est-à-dire le nombre de k-uplets d'indices distincts dans l'ensemble J qui est de cardinal m, soit

$$X(\omega) = m(m-1)\cdots(m-(k-1)) = (m)_k.$$

Par conséquent, $(N)_k = X$.

▷ La variable N est finie donc admet des moments (factoriels) de tout ordre. On obtient l'expression désirée des moments factoriels par linéarité de l'espérance avec l'expression ci-dessus. ♢

Retrouvons désormais la formule du crible à partir des résultats sur les moments factoriels de la première section.

2.2. Proposition. *Soit $(A_j)_{1\leqslant j\leqslant n}$ des événements. Alors,*

$$\mathbf{P}\Big(\bigcup_{j=1}^{n} A_j\Big) = \sum_{k=1}^{n}(-1)^{k-1} \sum_{j_1<\ldots<j_k} \mathbf{P}\Big(\bigcap_{l=1}^{k} A_{j_l}\Big).$$

Démonstration. Appliquons la proposition VIII-1.5 avec N la variable de comptage des événements $(A_j)_{1\leqslant j\leqslant n}$ et $k=0$:

$$\begin{aligned}\mathbf{P}(N\neq 0) &= 1-\mathbf{P}(N=0)\\ &= \sum_{l=1}^{+\infty} \tfrac{(-1)^{l-1}}{l!}\mathbf{E}((N)_l).\end{aligned}$$

Remarquons que l'événement $N\neq 0$ est $\bigcup_{j=1}^{n} A_j$. Par ailleurs, avec l'expression des moments factoriels obtenue ci-dessus, pour tout k,

$$\begin{aligned}\mathbf{E}((N)_k) &= \sum_{j_1,\ldots,j_k\in[\![1,n]\!] \text{ distincts}} \mathbf{P}\Big(\bigcap\nolimits_{l=1}^{k} A_{j_l}\Big)\\ &= k!\sum_{j_1<\ldots<j_k} \mathbf{P}\Big(\bigcap_{l=1}^{k} A_{j_l}\Big).\end{aligned}$$

La formule du crible en découle immédiatement en remplaçant les différents termes dans l'expression ci-dessus de $\mathbf{P}(N\neq 0)$. ♢

En introduisant une variable de comptage de la même manière, l'encadrement de la proposition VIII-1.7 (en fait celui qui apparaît dans sa preuve) fournit immédiatement les inégalités de Bonferroni.

2.3. Proposition. *Soit $(A_j)_{1\leqslant j\leqslant n}$ des événements. Alors, pour tout $p \in \mathbb{N}^*$,*

$$\mathbf{P}\Big(\bigcup_{j=1}^{n} A_j\Big) \geqslant \sum_{k=0}^{2p}(-1)^{k+1} \sum_{j_1<\ldots<j_k} \mathbf{P}\Big(\bigcap_{l=1}^{k} A_{j_l}\Big),$$

et

$$\mathbf{P}\Big(\bigcup_{j=1}^{n} A_j\Big) \leqslant \sum_{k=0}^{2p-1}(-1)^{k+1} \sum_{j_1<\ldots<j_k} \mathbf{P}\Big(\bigcap_{l=1}^{k} A_{j_l}\Big).$$

3. Théorème de convergence

Montrons comment l'étude des moments factoriels fournit un moyen d'obtenir la convergence vers une loi de Poisson.

3.1. Théorème. *Soit $(N_n)_n$ une suite de variables aléatoires de comptage telle qu'il existe $\lambda \in \mathbb{R}_+^*$ vérifiant, pour tout $k \in \mathbb{N}$,*

$$\mathbf{E}((N_n)_k) \to \lambda^k.$$

Alors,

$$\mathbf{P}(N_n = 0) \to e^{-\lambda}.$$

3.2. Remarque. Avec ce théorème, on énonce un résultat minimal qui correspond à l'objectif de ce chapitre. En fait, on aurait pu directement conclure avec les mêmes hypothèses que, pour tout $k \in \mathbb{N}$, $\mathbf{P}(N_n = k) \to \frac{\lambda^k}{k!}e^{-\lambda}$. Cet énoncé plus général indique alors que la convergence des moments factoriels de la suite de variables $(N_n)_n$ vers les moments factoriels d'une loi de Poisson entraîne la convergence en loi de $(N_n)_n$ vers cette même loi de Poisson.

Ce résultat provient essentiellement du fait que la loi de Poisson est caractérisée par ses moments (donc par ses moments factoriels).

Démonstration. Fixons $\varepsilon \in \mathbb{R}_+^*$. Alors, pour tout $n \in \mathbb{N}$ et tout $p \in \mathbb{N}^*$, en séparant les sommes au rang p,

$$
\begin{aligned}
&|\mathbf{P}(N_n = 0) - e^{-\lambda}| \\
&\qquad = \Big|\mathbf{P}(N_n = 0) - \sum_{l=0}^{+\infty} \tfrac{(-1)^l \lambda^l}{l!}\Big| \\
&\qquad \leqslant \sum_{l=p}^{+\infty} \tfrac{\lambda^l}{l!} + \sum_{l=0}^{p-1} \tfrac{1}{l!}\big|\mathbf{E}((N_n)_l) - \lambda^l\big| + \\
&\qquad\qquad + \Big|\mathbf{P}(N_n = 0) - \sum_{l=0}^{p-1} \tfrac{(-1)^l}{l!}\mathbf{E}((N_n)_l)\Big| \\
&\qquad \leqslant \sum_{l=p}^{+\infty} \tfrac{\lambda^l}{l!} + \sum_{l=0}^{p-1} \tfrac{1}{l!}\big|\mathbf{E}((N_n)_l) - \lambda^l\big| + \tfrac{1}{p!}\mathbf{E}((N_n)_p),
\end{aligned}
$$

en utilisant la majoration du troisième terme fournie par la proposition VIII-1.7.

Considérons p tel que $\sum_{l=p}^{+\infty} \frac{\lambda^l}{l!} \leqslant \varepsilon$ et $\frac{2\lambda^p}{p!} \leqslant \varepsilon$.

Comme $\mathbf{E}((N_n)_p) \to \lambda^p$, il existe un rang n_1 tel que, pour tout $n \geqslant n_1$, $\mathbf{E}((N_n)_p) \leqslant 2\lambda^p$.

Comme pour tout $l \in [\![0, p-1]\!]$, $\mathbf{E}((N_n)_l) - \lambda^l \to 0$, il existe un rang n_2 tel que, pour tout $n \geqslant n_2$,

$$\sum_{l=0}^{p-1} \tfrac{1}{l!}\big|\mathbf{E}((N_n)_l) - \lambda^l\big| \leqslant \varepsilon.$$

Combinant ces différentes majorations, on obtient que, pour tout $n \geqslant \max(n_1, n_2)$,

$$|\mathbf{P}(N_n = 0) - e^{-\lambda}| \leqslant 3\varepsilon.$$

♢

3.3. Exemple. Soit $\lambda \in \mathbb{R}_+^*$. Pour tout $n \geqslant \lambda$, considérons n événements indépendants de même probabilité $\frac{\lambda}{n}$ et N_n la variable aléatoire de comptage associée à ces événements.
Alors, pour tout $k \in \mathbb{N}^*$ et tout $n \geqslant \lambda$, la proposition VIII-2.1 donne

$$\begin{aligned}\mathbf{E}((N_n)_k) &= \sum_{j_1,\dots,j_k\in[\![1,n]\!] \text{ distincts}} \left(\tfrac{\lambda}{n}\right)^k \\ &= \left(\tfrac{\lambda}{n}\right)^k (n)_k.\end{aligned}$$

Ainsi, pour tout $k \in \mathbb{N}$, $\mathbf{E}((N_n)_k) \to \lambda^k$.

D'après le théorème VIII-3.1, on obtient que $\mathbf{P}(N_n = 0) \to e^{-\lambda}$.

Ce résultat correspond à la convergence en loi (bien connue) de variables binomiales de paramètres n et $\frac{\lambda}{n}$ vers la loi de Poisson de paramètre λ.

L'exemple précédent permet de traiter le cas de variables de comptage associées à des événements à la fois rares et indépendants. Le suivant correspond à un cas où il n'y a plus tout à fait indépendance.

3.4. Exemple. Notons, pour tout $n \in \mathbb{N}^*$, Σ_n une permutation aléatoire de loi uniforme sur S_n et N_n la variable aléatoire de comptage des points fixes de Σ_n. Alors, pour tout $k \in [\![1, n]\!]$, la proposition VIII-2.1 donne

$$\begin{aligned}\mathbf{E}((N_n)_k) &= \sum_{j_1,\dots,j_k\in[\![1,n]\!] \text{ distincts}} \mathbf{P}\Big(\bigcap\nolimits_{l=1}^{k} (\Sigma_n(j_l) = j_l)\Big) \\ &= \sum_{j_1,\dots,j_k\in[\![1,n]\!] \text{ distincts}} \tfrac{(n-k)!}{n!} \\ &= 1.\end{aligned}$$

et, pour tout $k > n$, $\mathbf{E}((N_n)_k) = 0$ car N_n est à valeurs dans l'ensemble $[\![0, n]\!]$.

Par conséquent, pour tout $k \in \mathbb{N}$, $\mathbf{E}((N_n)_k) \to 1$ (en fait la suite est stationnaire à 1). D'après le théorème VIII-3.1 avec $\lambda = 1$, on obtient que

$$\mathbf{P}(N_n = 0) \to e^{-1}.$$

Pour des études de graphes aléatoires au seuil, on va chercher à appliquer le résultat d'approximation poissonnienne à des fonctions

de comptage pour des événements qui ne sont pas mutuellement indépendants mais tels que un nombre conséquent d'entre eux sont mutuellement indépendants, une sorte de presque indépendance en quelque sorte.

4. Nombres de triangles au seuil

Soit $c \in \mathbb{R}_+^*$. Notons, pour tout entier $n > c$, $p_n = \frac{c}{n}$ et N_n la variable aléatoire de comptage des triangles (c'est-à-dire de graphes induits isomorphes à K_3) du graphe aléatoire X_{n,p_n}.

4.1. Proposition. *Pour tout $k \in \mathbb{N}^*$, le k-ième moment factoriel de N_n vérifie*

$$\mathbf{E}((N_n)_k) = \tbinom{n}{3}\tbinom{n-3}{3}\cdots\tbinom{n-3(k-1)}{3} p_n^{3k} + o(1).$$

Démonstration. Fixons $k \in \mathbb{N}^*$. Calculons l'espérance $\mathbf{E}((N_n)_k)$ avec la proposition VIII-2.1 : il s'agit de la somme sur tous les k-uplets de triangles distincts $(T_1, \ldots, T_k)$ de la probabilité d'apparition des triangles T_1, ..., T_k dans le graphe aléatoire X_{n,p_n}. Décomposons cette somme en deux :

- d'une part, la somme Σ_1 sur les k-uplets de triangles distincts et n'ayant aucun sommet en commun,
- d'autre part, la somme Σ_2 sur tous les autres k-uplets de triangles distincts.

On obtient immédiatement en dénombrant les k-uplets de triangles à sommets distincts que

$$\Sigma_1 = \tbinom{n}{3}\tbinom{n-3}{3}\cdots\tbinom{n-3(k-1)}{3} p_n^{3k}.$$

Partageons la somme Σ_2 selon le graphe G' correspondant à la réunion des triangles du k-uplet (donc un graphe admettant au plus $3k-1$ sommets d'après la définition de Σ_2). Pour chaque tel graphe G', notons

- s' le nombre de sommets de G' et a' son nombre d'arêtes,
- n' le nombre de k-uplets de triangles dont la réunion est G'.

Avec toutes ces notations,

$$\begin{aligned}\Sigma_2 &= \sum_{G'} n' \mathbf{P}\big(G' \text{ apparaît dans } X_{n,p_n}\big)\\ &= \sum_{G'} n' p_n^{a'}.\end{aligned}$$

Remarquons ensuite que $n' = O(1)$ (car le nombre n' dépend de G', donc de k mais pas de la variable n pour laquelle on considère le comportement asymptotique) puis que s' se situe entre 4 et $3k-1$.

Par conséquent,

$$\Sigma_2 = O\Big(\sum_{s'=4}^{3k-1} \tbinom{n}{s'} p_n^{a'}\Big),$$

Par construction de G' (voir par exemple l'exercice I-8.8), le nombre d'arêtes a' est strictement plus grand que le nombre de sommets s', donc $p_n^{a'} \leqslant p_n^{s'+1}$. Ainsi,

$$\Sigma_2 = O\Big(\sum_{s'=4}^{3k-1} n^{s'} p_n^{s'+1}\Big)$$

avec l'estimation $\binom{n}{s'} = O(n^{s'})$. Par conséquent, on obtient

$$\Sigma_2 = O\Big(p_n \sum_{s'=4}^{3k-1} (np_n)^{s'}\Big).$$

Comme $np_n \to c$,

$$\sum_{s'=4}^{3k-1} (np_n)^{s'} \to (3k-4)c^{s'}.$$

Par suite, $\Sigma_2 = O(p_n)$ et donc $\Sigma_2 = o(1)$: le résultat sur $\mathbf{E}((N_n)_k)$ suit immédiatement. ♢

4.2. Corollaire. *Pour tout $k \in \mathbb{N}^*$, $\mathbf{E}((N_n)_k) \to \big(\frac{c^3}{6}\big)^k$.*

Démonstration. Remarquons que, pour tout entier $l \in [\![0, n]\!]$ fixé,

$$\binom{n-l}{3} = \frac{(n-l)_3}{3!} \sim \frac{n^3}{6}.$$

Par conséquent, avec la proposition précédente,

$$\mathbf{E}((N_n)_k) \sim \left(\frac{(np_n)^3}{6}\right)^k.$$

Par hypothèse sur p_n, cette quantité tend vers $\left(\frac{c^3}{6}\right)^k$. ♢

On peut désormais conclure avec le théorème VIII-3.1 sur la probabilité d'avoir un graphe sans triangle avec une probabilité au seuil.

4.3. Proposition. $\mathbf{P}(N_n = 0) \to \frac{c^3}{6}$.

5. Exercices

L'objectif de ce premier exercice est de revisiter la proposition VIII-1.5 en montrant une identité entre variables aléatoires.

5.1. Exercice. *Soit X une variable aléatoire à valeurs dans $\mathbb{N}$.*

1. *Montrer que*
$$\mathbb{1}_{X=0} = \sum_{k=0}^{X} \frac{(-1)^k}{k!}(X)_k.$$
2. *Trouver une formule analogue pour la variable $\mathbb{1}_{X=n}$ avec n un entier fixé.*

Solution.

1. Remarquons que $\mathbb{1}_{X=0} = (1-1)^X$ donc avec la formule du binôme,
$$\begin{aligned}\mathbb{1}_{X=0} &= \sum_{k=0}^{X} \binom{X}{k}(-1)^k \\ &= \sum_{k=0}^{X} \frac{(-1)^k}{k!}(X)_k.\end{aligned}$$

2. Soit $n \in \mathbb{N}$. Reprenons la stratégie de la première réponse en partant de l'identité $\mathbb{1}_{X=n} = \binom{X}{n}(1-1)^{X-n}$. La formule du binôme donne alors

$$\begin{aligned}\mathbb{1}_{X=n} &= \tbinom{X}{n} \sum_{k=0}^{+\infty} \tbinom{X-n}{k}(-1)^k \\ &= \sum_{k=0}^{+\infty} (-1)^k \tbinom{X}{n+k}\tbinom{n+k}{n} \\ &= \tfrac{1}{n!} \sum_{k=0}^{+\infty} \tfrac{(-1)^k}{k!}(X)_{n+k}.\end{aligned}$$

5.2. Exercice. *Soit $(p_n)_n$ une suite à valeurs dans $[0,1]$. Définissons, pour tout $n \in \mathbb{N}^*$, la variable N_n de comptage du nombre de sommets isolés dans le graphe aléatoire X_{n,p_n}.*

1. *Montrer que, pour tout $k \in \mathbb{N}^*$,*

$$\mathbf{E}((N_n)_k) = (n)_k (1-p_n)^{k(n-1)-k(k-1)/2}.$$

2. *Soit $c \in \mathbb{R}_+^*$ tel que $p_n = \frac{1}{n}(\ln n + c + o(1))$. Déterminer $\lim \mathbf{P}(N_n = 0)$.*

Solution.

1. Fixons $k \in \mathbb{N}^*$. Calculons l'espérance $\mathbf{E}((N_n)_k)$ avec la proposition VIII-2.1 :

$$\begin{aligned}\mathbf{E}((N_n)_k) &= \sum_{j_1,\dots,j_k \in [\![1,n]\!] \text{ distincts}} \mathbf{P}(j_1,\dots,j_k \text{ isolés dans } X_{n,p_n}) \\ &= \sum_{j_1,\dots,j_k \in [\![1,n]\!] \text{ distincts}} (1-p_n)^{k(n-1)-k(k-1)/2},\end{aligned}$$

   car, pour k sommets fixés, il y a $k(n-1)$ arêtes issues de ces sommets dont $\binom{k}{2}$ qui joignent deux de ces sommets.

   En dénombrant les k-uplets de sommets, on obtient la formule attendue.

2. ▷ Remarquons tout d'abord qu'avec le développement asymptotique de p_n de l'énoncé, on obtient

$$\begin{aligned} n(1-p_n)^n &= \exp(\ln n + n\ln(1-p_n)) \\ &= \exp(\ln n - np_n + O(np_n^2)) \\ &= \exp(\ln n - \ln n - c + o(1)) \\ &\to e^{-c}. \end{aligned}$$

▷ En reportant ce résultat dans le calcul de la première question,

$$\begin{aligned} \mathbf{E}((N_n)_k) &\sim n^k(1-p_n)^{kn}(1-p_n)^{-k-k(k-1)/2} \\ &\sim (n(1-p_n)^n)^k \\ &\to e^{-kc}. \end{aligned}$$

D'après le théorème VIII-3.1, $\mathbf{P}(N_n = 0) \to \exp\big(-\exp(-c)\big)$.

Paul Erdős et Alfréd Rényi montrent dans [ER59] le résultat plus fort suivant.

5.3. Proposition. *Soit $c \in \mathbb{R}_+^*$ et $(p_n)_n$ une suite à valeurs dans $[0,1]$ telle que $p_n = \frac{1}{n}(\ln n + c + o(1))$. Alors, la probabilité que le graphe aléatoire X_{n,p_n} soit connexe converge vers $\exp\big(-\exp(-c)\big)$.*

Il est possible au prix de quelques efforts de le déduire essentiellement du résultat de cet exercice sur les sommets isolés.

Graphe aléatoire dénombrable

Chapitre IX

Graphe de Rado

Jusqu'à présent, nous considérions des graphes aléatoires finis ou leurs limites quand le nombre de sommets augmentait. Dans ce chapitre, nous adoptons une approche « complémentaire » en considérant directement un graphe aléatoire dénombrable. Cette première description de ce graphe est un tantinet « naïve » et les résultats obtenus seront revisités par la suite.

La construction proposée est une nouvelle fois due à Paul Erdős et Alfréd Rényi [ER63]. Toutefois, comme cela apparaîtra dans les chapitres suivants, d'autres mathématiciens ont étudié ce graphe indépendamment : notamment, Wilhelm Ackermann [Ack37] et Richard Rado [Rad64]. On a choisi d'utiliser l'appellation « graphe de Rado » afin d'éviter toute ambiguïté.

1. Définition

1.1. Définition. Soit $p \in]0,1[$, puis $(I_a)_{a\in\mathscr{P}_2(\mathbb{N}^*)}$ une famille de variables aléatoires mutuellement indépendantes suivant toutes la loi de Bernoulli de paramètre p. Définissons l'ensemble aléatoire d'arêtes suivant :

$$A = \{a \in \mathscr{P}_2(\mathbb{N}^*),\ I_a = 1\}.$$

Le graphe de Rado de paramètre p est le graphe aléatoire $(\mathbb{N}^*, A)$.

Un premier lien avec les chapitres précédents apparaît naturellement avec la proposition suivante.

1.2. Proposition. *Soit $p \in]0,1[$, G le graphe aléatoire de Rado de paramètre p. Alors, pour tout entier naturel n non nul, le graphe induit $G[[\![n]\!]]$ suit le modèle binomial de paramètres n et p.*

Démonstration. Il suffit de remarquer que l'ensemble des arêtes du graphe induit est défini à partir de la famille $(I_a)_{a\in\mathscr{P}_2([\![n]\!])}$ de variables aléatoires mutuellement indépendantes de loi de Bernoulli de paramètre p. On conclut avec la proposition III-2.3. ♢

1.3. Remarque. L'ensemble $[\![n]\!]$ ne joue pas un rôle particulier dans ce résultat : pour toute partie S de $\mathbb{N}^*$ de cardinal n, le graphe induit $G[S]$ est isomorphe à un graphe aléatoire suivant le modèle binomial de paramètres n et p (il suffit de renommer les sommets).

Dans les sections suivantes, on développe quelques propriétés élémentaires (presque sûres) du graphe de Rado, dont on obtiendra une meilleure compréhension à la suite du théorème XI-1.1.

2. Diamètre du graphe de Rado

2.1. Proposition. *Soit $p \in]0,1[$ et G le graphe aléatoire de Rado de paramètre p. Alors, deux sommets fixés de G sont presque sûrement à distance au plus 2.*

Démonstration. Soit $i \neq j$ deux entiers naturels non nuls (ou de manière équivalente deux sommets de G). Écrivons l'événement « i et j sont à distance strictement plus grande que 2 » sous la forme suivante

$$(\{i,j\} \notin A) \bigcap \bigcap_{k\notin\{i,j\}} \big((\{i,k\} \notin A) \cup (\{k,j\} \notin A)\big).$$

Remarquons que, d'après la définition du graphe de Rado,

$$\mathbf{P}(\{i,j\} \notin A) = 1 - \mathbf{P}(\{i,j\} \in A) = 1 - p.$$

De même, pour tout sommet k différent de i et j,

$$\begin{aligned}\mathbf{P}\big((\{i,k\} \notin A) \cup (\{k,j\} \notin A)\big) \\ &= 1 - \mathbf{P}\big(\{i,k\} \in A, \{k,j\} \in A\big) \\ &= 1 - \mathbf{P}\big(\{i,k\} \in A\big)\mathbf{P}\big(\{k,j\} \in A\big) \\ &= 1 - p^2.\end{aligned}$$

Ainsi, par indépendance des arêtes, la probabilité que i et j soient à distance strictement plus grande que 2 dans le graphe induit de G par l'ensemble de sommets $[\![n]\!]$ (avec $n \geqslant \max(i,j)$) tend vers

$$(1-p)\big(1-p^2\big)^{n-2}.$$

Enfin, par utilisation de la continuité décroissante de $\mathbf{P}$, la probabilité que i et j soient à distance strictement plus grande que 2 est

$$\lim (1-p)\big(1-p^2\big)^{n-2} = 0.$$

En conclusion, la distance de i à j est inférieure ou égale à 2 avec probabilité 1. ♢

En particulier, pour chaque paire de sommets, il existe avec probabilité 1 un chemin (de longueur au plus 2) les joignant. On en déduit le corollaire suivant.

2.2. Corollaire. *Soit $p \in]0,1[$ et G le graphe aléatoire de Rado de paramètre p. Alors, G est presque sûrement connexe.*

Démonstration. L'événement « G est connexe » contient l'intersection pour $i < j$ des événements « il existe un chemin joignant i à j dans G ». Comme tous ces événements sont de probabilité 1 et que l'intersection est dénombrable, l'intersection est de probabilité 1 (passer au complémentaire s'il faut redémontrer ce résultat : une réunion dénombrable d'événements négligeables est négligeable), donc G est presque sûrement connexe. ♢

2.3. Proposition. *Soit $p \in]0,1[$ et G le graphe aléatoire de Rado de paramètre p. Alors, le diamètre de G est presque sûrement égal à 2.*

Démonstration. D'après la première proposition, le diamètre de G est presque sûrement 1 ou 2. Si G était de diamètre 1, cela signifierait que deux sommets quelconques sont à distance 1, autrement dit que G est presque sûrement complet.

Comme précédemment, on calcule la probabilité que le graphe induit de G par l'ensemble de sommets $[\![n]\!]$ soit complet :

$$\prod_{\substack{i,j\leqslant n\\ i\neq j}} \mathbf{P}(\{i,j\}\in A) = p^{\binom{n}{2}}.$$

En passant à la limite décroissante en n, on obtient que la probabilité que G soit complet est nulle. ♢

3. Sous-graphes finis du graphe de Rado

Commençons avec un cas élémentaire.

3.1. Proposition. *Soit $p\in\,]0,1[$ et G le graphe aléatoire de Rado de paramètre p. Alors, G admet presque sûrement un graphe induit isomorphe à K_3 (c'est-à-dire un triangle).*

Démonstration. Passons au complémentaire et calculons la probabilité de l'événement « G n'admet pas de graphe induit isomorphe à K_3 ». Pour cela, remarquons que cet événement est inclus dans la réunion suivante, où l'on se limite aux triangles composés avec des sommets consécutifs (donc de la forme $3k+1$, $3k+2$ et $3k+3$) :

$$E = \bigcup_{k\in\mathbb{N}^*} (G[\{3k+1, 3k+2, 3k+3\}] \text{ est isomorphe à } K_3).$$

En passant au complémentaire puis en utilisant l'indépendance des arêtes et la continuité décroissante, on obtient que la probabilité de ce dernier événement est

$$\mathbf{P}(E) = \lim(1-p^3)^n = 0.$$

Ainsi, la probabilité que G n'admette pas de graphe induit isomorphe à K_3 est inférieure à $\mathbf{P}(E)=0$, donc nulle. ♢

En soignant un peu l'argument précédent, on peut montrer le corollaire suivant.

3.2. Corollaire. *Soit $p \in]0,1[$, G le graphe aléatoire de Rado de paramètre p. Alors, G admet presque sûrement une infinité de graphes induits isomorphes à K_3.*

Démonstration. Notons, pour tout $k \in \mathbb{N}$, A_k l'événement « le graphe induit

$$G[\{3k+1, 3k+2, 3k+3\}]$$

est isomorphe à K_3 », et remarquons que

- ▷ les événements $(A_k)_k$ sont mutuellement indépendants,
- ▷ pour tout k, $\mathbf{P}(A_k) = p^3$, donc la série de terme général $\mathbf{P}(A_k)$ est (grossièrement) divergente.

Ainsi, d'après le second lemme de Borel et Cantelli (proposition II-5.6), il existe presque sûrement une infinité d'événements parmi les A_k qui sont réalisés et donc G admet presque sûrement une infinité de graphes induits isomorphes à K_3. ♢

Revenons au cas général.

3.3. Proposition. *Soit $p \in]0,1[$, G le graphe aléatoire de Rado de paramètre p et H un graphe fini. Alors, G admet presque sûrement un graphe induit isomorphe à H.*

Démonstration. Notons n le nombre de sommets de H et m son nombre d'arêtes. La probabilité de l'événement A_k : « le graphe

$$G[\{kn+1, kn+2, \ldots, (k+1)n\}]$$

est isomorphe à H » est de la forme

$$cp^m(1-p)^{\binom{n}{2}-m},$$

avec c une constante qui ne dépend que de H (en fait de son groupe d'automorphismes d'après l'exercice I-8.6).

Alors,

- ▷ les événements $(A_k)_k$ sont mutuellement indépendants (car les graphes induits considérés ont des ensembles de sommets, et donc d'arêtes, deux à deux disjoints),
- ▷ la série de terme général $\mathbf{P}(A_k)$ diverge (car $\mathbf{P}(A_k)$ est une constante non nulle par rapport à k).

D'après le second lemme de Borel et Cantelli (proposition II-5.6), il existe presque sûrement une infinité d'indices k tels que l'événement A_k soit réalisé : donc le graphe G admet presque sûrement une infinité de graphes induits isomorphes à H. ♢

On peut bien entendu obtenir le même corollaire que précédemment en remplaçant K_3 par H.

4. Exercices

4.1. Exercice. *Soit $p \in]0,1[$ et G le graphe aléatoire de Rado de paramètre p. Montrer que, pour tout entier N, le graphe G' induit par G avec l'ensemble de sommets $\{k \in \mathbb{N},\ k \geqslant N\}$ est isomorphe à G.*

Solution. L'application $k \mapsto k + 1 - N$ réalise une bijection de $\{k \in \mathbb{N},\ k \geqslant N\}$ dans $\mathbb{N}^*$. Ainsi, le graphe G' est isomorphe à un graphe d'ensemble de sommets $\mathbb{N}^*$ pour lequel chaque arête apparaît indépendamment des autres avec la même probabilité p : il est donc isomorphe au graphe de Rado G.

4.2. Exercice. *Soit $p \in]0,1[$ et G le graphe aléatoire de Rado de paramètre p. Montrer que chaque sommet est presque sûrement de degré infini.*

Solution. Soit i un entier naturel non nul fixé.

Notons, pour tout $n \neq i$, A_n l'événement « le sommet i est voisin du sommet n ». Alors,

- ▷ les événements $(A_n)_{n \neq i}$ sont mutuellement indépendants,
- ▷ pour tout $n \neq i$, $\mathbf{P}(A_n) = p$: la série de terme général $\mathbf{P}(A_n)$ est donc divergente.

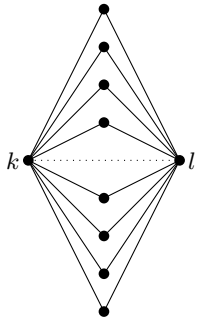
Ainsi, d'après le second lemme de Borel et Cantelli (proposition II-5.6), il existe presque sûrement une infinité d'événements A_n réalisés, c'est-à-dire le sommet i est presque sûrement de degré infini.

4.3. Exercice. *Soit $p \in]0,1[$ et G le graphe aléatoire de Rado de paramètre p.*

1. *Soit $i \neq j$ deux entiers naturels non nuls et G' le graphe déduit de G en enlevant l'arête $\{i,j\}$ de G. Montrer que le graphe G' est presque sûrement connexe.*
2. *Soit A une partie finie de l'ensemble des paires de $\mathbb{N}^*$ et G' le graphe déduit de G en enlevant les arêtes de A. Montrer que le graphe G' est presque sûrement connexe.*

Solution.

1. ▷ Soit $k \neq l$ deux sommets. D'après la section 2 de ce chapitre, il existe presque sûrement une infinité de chemins de longueur 2 reliant k à l dans G (conséquence du second lemme de Borel et Cantelli). Au plus un de ces chemins exploite l'arête $\{i,j\}$; par conséquent, il existe presque sûrement un chemin de longueur 2 reliant k à l dans G'.



   ▷ L'événement « G' est connexe » est la réunion (dénombrable) sur tous les paires $\{k,l\}$ de sommets des événements presque sûrs « les sommets k et l sont reliés dans G' », donc est presque sûr.

   En termes plus savants, le graphe aléatoire de Rado de paramètre p n'admet presque sûrement pas d'isthme.
2. On reprend le raisonnement ci-dessus : comme il y a presque sûrement une infinité de chemins de longueur 2 joignant dans G

deux sommets donnés, il en existe presque sûrement un (en fait une infinité) qui joint ces sommets dans G'. Les autres étapes de la correction sont identiques.

Ce résultat sera une conséquence immédiate et élémentaire du dernier chapitre.

Chapitre X

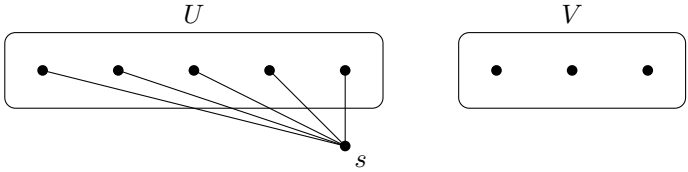
Propriété d'extension

Dans ce chapitre, on définit la propriété d'extension pour un graphe dénombrable. Ensuite, on passe en revue différentes constructions (exploitant des arguments mathématiques de natures différentes) de graphes vérifiant cette propriété.

1. Définition et premières propriétés

1.1. Définition. Un graphe d'ensemble de sommets S dénombrable et d'ensemble d'arêtes A vérifie la propriété d'extension si, pour tout couple (U, V) de parties de S finies et disjointes, il existe un sommet $s \in S \setminus (U \cup V)$ tel que

- ▷ pour tout $u \in U$, $\{s, u\} \in A$: autrement dit, s est voisin de tous les sommets appartenant à U,
- ▷ pour tout $v \in V$, $\{s, v\} \notin A$: autrement dit, s n'est voisin d'aucun sommet appartenant à V.



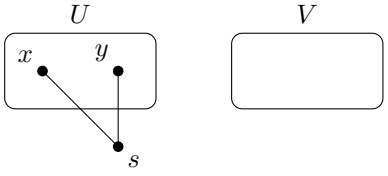
1.2. Remarque. Il est immédiat qu'aucun graphe fini ne peut vérifier cette propriété (prendre U l'ensemble de tous les sommets et V l'ensemble vide). En revanche, rien n'impose l'hypothèse de dénombrabilité qui figure ici dans la définition, et qui est dorénavant une hypothèse implicite.

Admettons un instant qu'un tel graphe existe et montrons une propriété élémentaire sur son diamètre.

1.3. Proposition. *Soit G un graphe vérifiant la propriété d'extension. Alors, la distance entre deux sommets de G est au plus* 2.

En d'autres termes, le diamètre de G est inférieur ou égal à 2.

Démonstration. Soit x et y deux sommets distincts et non voisins (sinon leur distance est inférieure ou égale à 1). Appliquons la propriété d'extension aux parties $U = \{x, y\}$ et $V = \varnothing$: il existe un sommet s qui est voisin de tous les sommets de U, donc voisin commun à x et y.



Par conséquent, la distance entre x et y est inférieure ou égale à 2 (par le chemin x, s, y). ♢

En particulier, le résultat suivant est établi.

1.4. Corollaire. *Un graphe vérifiant la propriété d'extension est connexe.*

1.5. Proposition. *Dans un graphe G vérifiant la propriété d'extension, il n'existe pas de sommet voisin de tous les autres.*

En d'autres termes, le diamètre de G est supérieur ou égal à 2.

Démonstration. Il suffit de remarquer que si un tel sommet x existait, cela contredirait la propriété d'extension avec $U = \varnothing$ et $V = \{x\}$. ♢

En combinant les deux propositions, on obtient le corollaire suivant.

1.6. Corollaire. *Un graphe vérifiant la propriété d'extension est de diamètre* 2*.*

Les sections suivantes donnent des constructions de graphes satisfaisant la propriété d'extension.

2. Construction par récurrence

Construisons par récurrence une suite de graphes $(G_k = (S_k, A_k))_k$ croissante pour l'inclusion dont la réunion vérifie la propriété d'extension. Par croissance, on indique que, pour tout entier naturel k, $S_k \subset S_{k+1}$ et $A_k \subset A_{k+1}$. En fait, on aura même la propriété suivante plus forte

$$A_k = A_{k+1} \cap \mathscr{P}_2(S_k).$$

Ainsi, on n'ajoute pas d'arête entre sommets de S_k à une étape postérieure à k.

▷ À la première étape, considérons le graphe $G_0 = (\{1\}, \varnothing)$.

▷ Soit $k \geqslant 0$. Supposons le graphe $G_k = (S_k, A_k)$ construit ; pour chaque partie $U \subset S_k$, ajoutons un nouveau sommet exactement relié à tous les sommets de U et à aucun autre de S_k puis étiquetons-le avec le plus petit entier non encore utilisé pour étiqueter un sommet. Notons G_{k+1} le graphe obtenu après tous ces ajouts.

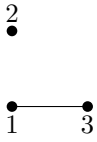
2.1. Remarque. Notons, pour tout $k \in \mathbb{N}$, u_k le nombre de sommets du graphe G_k. Alors, $u_0 = 1$, et, pour tout entier k,

$$u_{k+1} = u_k + 2^{u_k}.$$

Cette suite diverge très rapidement vers $+\infty$.

Le dessin du graphe G_1 est immédiat

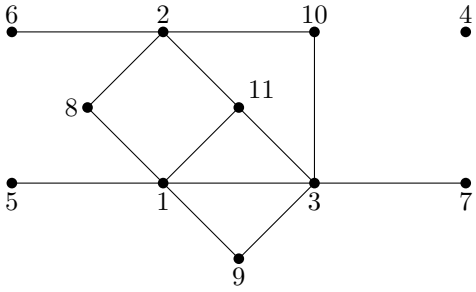
Pour le graphe G_2, on ajoute à la fois un sommet 2 uniquement relié à $U = \varnothing$, et un sommet 3 uniquement relié à $U = \{1\}$; on obtient



À l'étape suivante, on obtient G_3 en ajoutant successivement

- ▷ un sommet 4 uniquement relié à $U = \varnothing$ (donc isolé),
- ▷ des sommets 5, 6, 7 uniquement reliés respectivement aux singletons $U = \{1\}$, $U = \{2\}$, et $U = \{3\}$,
- ▷ des sommets 8, 9, 10 uniquement reliés respectivement aux paires $U = \{1, 2\}$, $U = \{1, 3\}$, et $U = \{2, 3\}$,
- ▷ un sommet 11 uniquement relié à $U = \{1, 2, 3\}$.

On obtient :



Il ne semble pas envisageable de poursuivre le dessin à l'étape suivante puisque le graphe G_4 compte déjà 2059 sommets : il faut rajouter $2^{11} = 2048$ sommets associés au 2^{11} parties de $[\![11]\!]$.

2.2. Remarque. Par construction, les suites des ensembles de sommets $(S_k)_k$ et des ensembles d'arêtes $(A_k)_k$ sont croissantes pour l'inclusion.

On peut donc considérer les « limites », c'est-à-dire les réunions, de ces suites croissantes

$$S = \bigcup_{k\in\mathbb{N}} S_k, \qquad A = \bigcup_{k\in\mathbb{N}} A_k.$$

Par ailleurs, toutes les arêtes de A entre deux sommets de l'ensemble des sommets S_k appartiennent à l'ensemble des arêtes A_k : on ne rajoute jamais d'arêtes entre deux sommets préalablement construits.

2.3. Proposition. *Le graphe $G = (S, A)$ obtenu comme réunion des graphes G_k pour $k \in \mathbb{N}$ vérifie la propriété d'extension.*

Démonstration. Considérons deux parties U et V finies et disjointes de $S = \mathbb{N}^*$. Soit k un entier tel que $U \cup V \subset S_k$, c'est-à-dire que k est l'indice d'une étape où tous les sommets de U ou V ont déjà été construits.

Par construction, le sommet associé à l'étape $k+1$ à la partie $U \subset S_k$ est relié dans G_{k+1} (donc dans G) à tous les sommets de U mais à aucun de V. ♢

En regardant un peu plus soigneusement la construction, on obtient le résultat plus précis suivant.

2.4. Corollaire. *Soit $G = (S, A)$ le graphe obtenu comme réunion des graphes G_k pour $k \in \mathbb{N}$ et (U, V) un couple de parties finies et disjointes de $S = \mathbb{N}^*$. Alors, il existe un entier $s > \max(U \cup V)$ tel que l'ensemble des voisins plus petits que s est exactement U, c'est-à-dire*

$$\{x < s,\ \{x, s\} \in A\} = U.$$

Démonstration. Soit k le plus petit entier tel que $U \cup V \subset S_k$. Par construction de G_{k+1}, il existe un sommet s qui est relié à U et à aucun autre sommet de S_k ni aucun sommet ajouté à cette étape $k+1$: par conséquent, les voisins de s avec une étiquette inférieure à s sont exactement les éléments de U. ♢

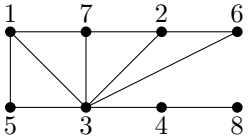
3. Construction « binaire »

Proposons une construction alternative sans récurrence.

3.1. Proposition. *Soit G le graphe d'ensemble de sommets $\mathbb{N}^*$ et tel qu'il existe une arête entre les sommets x et y avec $x < y$ si, et seulement si, le x-ième (en partant de la droite) chiffre de l'écriture binaire de y est un* 1*. Alors, G vérifie la propriété d'extension.*

3.2. Exemple. Il existe une arête entre 5 et 26, car l'entier 26 s'écrit en binaire $1\,1010_2$ mais pas entre 5 et 37, car 37 s'écrit en binaire $1\,0\,0101_2$ (donc admet un 0 en 5-ième position).

Voici le graphe induit par G sur l'ensemble $[\![8]\!]$



Démonstration. Considérons deux parties U et V finies et disjointes de $\mathbb{N}^*$. Soit M un entier strictement plus grand que tous les éléments de $U \cup V$ et considérons l'entier s dont l'écriture binaire comporte le chiffre 1 aux positions appartenant à l'ensemble $U \cup \{M\}$ (ce qui est possible puisque cette partie est finie) et des 0 ailleurs.

▷ Comme, pour tout $u \in U$, $u < s$ et s comporte le chiffre 1 en u-ième position de son écriture binaire, il existe bien une arête joignant u à s dans G.

▷ Comme, pour tout $v \in V$, $v < s$. En effet, à cause du chiffre 1 en M-ième position dans l'écriture binaire de s, on obtient

$$s \geqslant 2^{M-1} \geqslant M > v.$$

Par ailleurs, s comporte le chiffre 0 en v-ième position de son écriture binaire (car $U \cup \{M\}$ et V sont disjoints), donc il n'existe pas d'arête $\{v, s\}$ dans G. ♢

3.3. Remarque. Une erreur dans la preuve précédente serait de considérer pour un couple (U, V) de parties finies et disjointes de $\mathbb{N}^*$ l'entier s dont l'écriture binaire comporte le chiffre 1 exactement aux positions dans l'ensemble U : en effet, il pourrait alors y avoir un sommet $v \in V$ plus grand que s et tel que v admet le chiffre 1 en s-ième position de son écriture binaire (par exemple, avec les parties $U = \{1\}$, $V = \{4\}$ et le sommet $s = 2$).

4. Construction par résidus quadratiques

Proposons une troisième construction utilisant de l'arithmétique un tantinet plus élaborée dont le théorème de la progression arithmétique de Dirichlet dont voici un énoncé.

4.1. Théorème. *Pour tout entier naturel n non nul et tout entier naturel r premier avec n, il existe une infinité de nombres premiers congrus à r modulo n.*

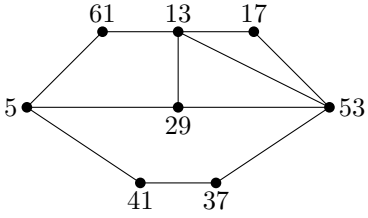
Dans la construction suivante l'ensemble des sommets est l'ensemble des nombres premiers congrus à 1 modulo 4 qui est infini dénombrable d'après le théorème X-4.1 ; quitte à énumérer cet ensemble, on peut se ramener comme précédemment à $\mathbb{N}^*$ pour ensemble de sommets : on ne détaillera pas cette étape élémentaire dans cette section.

4.2. Proposition. *Soit G le graphe d'ensemble de sommets les nombres premiers congrus à 1 modulo 4 et tel qu'il existe une arête entre les sommets x et y si, et seulement si, x est un carré modulo y. Alors, G vérifie la propriété d'extension.*

4.3. Remarque. D'après la loi de réciprocité quadratique, pour des nombres premiers x et y congrus à 1 modulo 4, x est un carré modulo y si, et seulement si, y est un carré modulo x. Cela permet de bien justifier le « si, et seulement si » qui apparaît dans la définition de G.

4.4. Exemple. Par exemple, une arête relie 29 à 13, car 29 est un carré modulo 13 ($29 \equiv 4^2[13]$ ou, comme indiqué dans la remarque précédente, $13 \equiv 10^2[29]$) mais il n'y a pas d'arête entre 13 et 5, car $13 \equiv 3[5]$ et 3 n'est pas un carré modulo 5.

Dessinons le graphe induit par G sur l'ensemble des sommets associés aux huit plus petits premiers congrus à 1 modulo 4, c'est-à-dire sur $\{5, 13, 17, 29, 37, 41, 53, 61\}$.



Revenons à la preuve de la proposition.

Démonstration. Considérons deux parties U et V finies et disjointes de l'ensemble des nombres premiers congrus à 1 modulo 4.

Pour tout $v \in V$, considérons un entier n_v qui n'est pas un carré modulo v. Le système diophantien suivant

$$\begin{cases} x \equiv 1 \quad [4] & \\ x \equiv 1 \quad [u] & \text{pour tout } u \in U, \\ x \equiv n_v \quad [v] & \text{pour tout } v \in V, \end{cases}$$

admet, d'après le théorème chinois, comme solution tous les éléments d'une certaine classe modulo

$$4 \prod_{p \in U \cup V} p.$$

Par ailleurs, cette classe contient (au moins) un nombre premier s d'après le théorème de Dirichlet (théorème X-4.1). Par construction, s est bien un sommet du graphe (car premier et congru à 1 modulo 4), relié à chaque élément de U (car égal à 1^2 modulo chaque élément de U donc bien un résidu quadratique), mais à aucun de V (d'après le choix des n_v pour $v \in V$).

♢

5. Construction probabiliste

5.1. Théorème. *Le graphe de Rado de paramètre $p \in]0,1[$ vérifie presque sûrement la propriété d'extension.*

Démonstration. Soit (U,V) un couple de parties finies et disjointes de l'ensemble des sommets. Notons $(x_k)_k$ une énumération des sommets en dehors de $U \cup V$.

Un sommet $x \notin U \cup V$ vérifie la propriété d'extension pour le couple (U,V) avec probabilité $p^{|U|}(1-p)^{|V|}$.

Par indépendance, la probabilité qu'aucun des sommets x_k pour $k \in [\![n]\!]$ ne vérifie la propriété d'extension pour (U,V) est par conséquent

$$\left(1-p^{|U|}(1-p)^{|V|}\right)^n.$$

Par continuité décroissante de $\mathbf{P}$, la probabilité qu'aucun sommet de la suite $(x_k)_k$ ne vérifie la propriété d'extension pour (U,V) est

$$\lim \left(1-p^{|U|}(1-p)^{|V|}\right)^n,$$

et comme $1-p^{|U|}(1-p)^{|V|} \in]0,1[$, cette limite est nulle. Ainsi, presque sûrement, il existe un sommet vérifiant la propriété d'extension pour le couple (U,V).

En effectuant la réunion sur l'ensemble dénombrable des couples (U,V) de parties finies et disjointes des événements presque sûrs « il existe un sommet $s \notin U \cup V$ vérifiant la propriété d'extension pour (U,V) », on obtient que l'événement « le graphe de Rado vérifie presque sûrement la propriété d'extension ». ♢

6. Exercices

6.1. Exercice. *Le graphe G d'ensemble de sommets $\mathbb{N}^*$ et d'ensemble d'arêtes $A = \{\{i,j\},\ i \wedge j \neq 1\}$ vérifie-t-il la propriété d'extension ?*

Solution. Rappelons que, pour tout nombre premier p et tous entiers naturels n, k non nuls, $n \wedge p = 1$ si, et seulement si $n \wedge p^k = 1$.

On obtient donc un contre-exemple à la propriété d'extension pour G en considérant les parties finies et disjointes $U = \{2\}$ et $V = \{4\}$ (et,

plus généralement, si U et V contiennent tous les deux une puissance d'exposant non nul d'un même nombre premier).

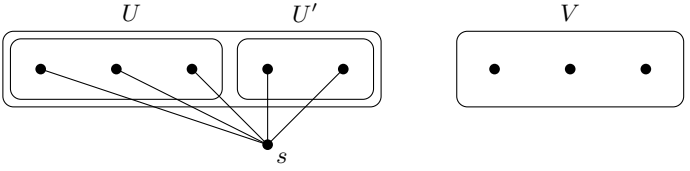
6.2. Exercice. *Soit G un graphe vérifiant la propriété d'extension.*

1. *Soit (U, V) un couple de parties finies et disjointes de l'ensemble des sommets. Montrer qu'il existe une infinité de sommets s reliés à tous les sommets de U et à aucun de V.*
2. *Montrer qu'un sous-graphe déduit de G en supprimant un nombre fini d'arêtes vérifie la propriété d'extension.*
3. *Montrer qu'un graphe induit par G en supprimant un nombre fini de sommets (et les arêtes qui en sont issues) vérifie la propriété d'extension.*

Solution.

1. Raisonnons par l'absurde. Supposons qu'il n'y ait qu'un nombre fini de tels sommets (on sait qu'il en existe au moins un d'après la propriété d'extension de G) et notons U' leur ensemble.

   Appliquons la propriété d'extension de G au couple de parties finies $(U \cup U', V)$ (qui sont bien disjointes) : il existe un sommet $s \notin U \cup U' \cup V$ relié à tous les sommets de $U \cup U'$ et à aucun sommet de V : en particulier, s est relié à tous les sommets de U et à aucun sommet de V, donc appartient à U' : contradiction.



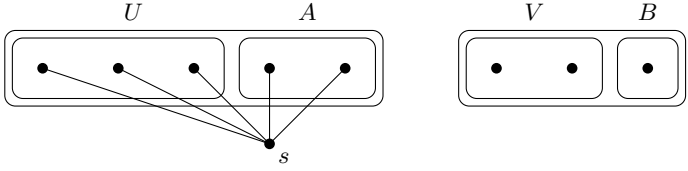
   On peut montrer, en poussant le raisonnement un peu plus loin, que G contient une infinité de fois chaque graphe fini.
2. Soit G' un tel sous-graphe et (U, V) un couple de parties finies et disjointes de l'ensemble des sommets. D'après la première question, il existe une infinité de sommets s qui sont reliés (dans G) à tous les sommets de U et à aucun de V. Or, les arêtes supprimées sont en nombre fini, et l'un de ses sommets demeure donc relié dans G' à tous les sommets de U. Ainsi, G' vérifie la propriété d'extension.

3. Soit G' un tel graphe induit et E l'ensemble fini des sommets supprimés. Soit (U, V) un couple de parties finies et disjointes de l'ensemble des sommets de G'. D'après la première question, il existe une infinité de sommets s qui sont reliés (dans G) à tous les sommets de U et à aucun de V. Or, E est une partie finie donc, parmi tous ses sommets s, il en existe au moins un qui n'appartient pas à E donc qui reste sommet de G'. Ainsi, le graphe G' vérifie la propriété d'extension.

Les deux dernières questions montrent la robustesse des graphes vérifiant la propriété d'extension : toute modification finie préserve cette propriété.

6.3. Exercice. *Soit G un graphe vérifiant la propriété d'extension, (A, B) un couple de parties finies et disjointes de l'ensemble des sommets et S l'ensemble des sommets (hors de $A \cup B$) reliés à tous les sommets de A et à aucun de B. Montrer que le graphe induit $G[S]$ vérifie la propriété d'extension.*

Solution. Soit (U, V) un couple de parties finies et disjointes de l'ensemble des sommets de $G[S]$. Appliquons la propriété d'extension du graphe G au couple $(U \cup A, V \cup B)$: il existe un sommet s de G qui est relié à tous les sommets de $U \cup A$ et à aucun de $V \cup B$.



Comme recherché, s est donc un sommet de $G[S]$ relié à tous les sommets de A et à aucun de B.

Ce dernier exercice montre une autre construction probabiliste, beaucoup plus récente, d'un graphe vérifiant la propriété d'extension.

6.4. Exercice (Modèle de Biró et Darji [BD18]). *Soit $(d_n)_{n\in\mathbb{N}^*}$ une suite d'entiers telle que, pour tout $n \in \mathbb{N}^*$, $0 \leqslant d_n \leqslant n$.*

Construisons un graphe aléatoire G d'ensemble de sommets $\mathbb{N}^*$ en choisissant, pour tout $n \in \mathbb{N}^*$, les voisins du sommet n indexés par un entier inférieur comme une partie de cardinal d_n de $[\![0, n-1]\!]$ choisie uniformément.

1. *Montrer que, si $d_n = 1$ pour tout n, le graphe obtenu G est presque sûrement un arbre, c'est-à-dire connexe et sans cycle.*

   La question suivante, un petit peu technique, est utile pour la suite ; elle est indépendante de toute considération sur les graphes ou les probabilités.

2. *Soit $(u_n)_n$ une suite à valeurs dans $]0,1[$. Justifier que la série de terme général u_n converge si, et seulement si, il existe $\ell > 0$ tel que*
$$\prod_{k=1}^{n}(1-u_k) \to \ell.$$
   Fixons U et V deux parties disjointes de $\mathbb{N}$ de cardinal k et N un entier strictement plus grand que $\max(U \cup V)$.

3. *Déterminer la probabilité qu'un sommet $n \geqslant N$ soit voisin dans G de tous les sommets appartenant à U mais d'aucun sommet appartenant à V.*

4. *Montrer que si la série de terme général*
$$\frac{(d_n)_k (n-d_n)_k}{(n)_{2k}}$$
   diverge, alors il existe presque sûrement une infinité de sommets voisins dans G de tous les sommets appartenant à U et d'aucun appartenant à V.

   Indication : on pourra utiliser le second lemme de Borel et Cantelli (proposition II-5.6).

5. *Supposons qu'il existe $\alpha \in]0,1[$ tel que $d_n = \lfloor \alpha n \rfloor$ pour tout entier n non nul. Montrer que le graphe G satisfait la propriété d'extension.*

Solution.

1. • Montrons, par récurrence sur $n \in \mathbb{N}$, que le graphe G_n induit de G avec l'ensemble des sommets inférieurs ou égaux à n est un arbre.

   ▷ Le résultat est immédiat au rang $n = 0$ puisque G_0 est un graphe avec un unique sommet et donc aucune arête.

   ▷ Soit $n \in \mathbb{N}$ tel que G_n est un arbre. Le sommet $n+1$ est alors voisin d'un unique sommet entre 0 et n (car $d_{n+1} = 1$) choisi uniformément. Ainsi, ce nouveau sommet ne peut appartenir à

un cycle et est relié à tous les sommets déjà présents par son sommet de rattachement. Ainsi, G_{n+1} est un arbre.

• Déduisons désormais que G est bien un arbre en montrant successivement la connexité et le caractère sans cycle.

▷ Deux sommets i et j sont reliés par un chemin dans $G_{\max(i,j)}$, donc dans G : ainsi, G est connexe ;

▷ Si G admet un cycle, alors le graphe intermédiaire G_n où n est le plus grand sommet du cycle admet également un cycle, ce qui contredit le caractère acyclique de G_n déjà établi.

De manière plus générale, si $d_n \in \{0,1\}$ pour tout n, on obtient une « forêt », c'est-à-dire une réunion d'arbres disjoints, et le nombre de composantes connexes est le nombre de termes nuls dans la suite $(d_n)_n$.

2. Quitte à passer le produit au logarithme, il s'agit d'établir que les séries de terme général positif u_n et $-\ln(1-u_n)$ sont de même nature.

   Or, dès que l'une de ces séries converge, le terme général tend vers 0 donc $u_n \to 0$: cette remarque entraîne immédiatement que $-\ln(1-u_n) \sim u_n$ puis, par comparaison de séries à termes positifs, la convergence de l'autre série.

3. Il suffit de dénombrer les parties $A \subset [\![0, n-1]\!]$ à d_n éléments qui contiennent U et ne rencontrent pas V, soit encore dénombrer les parties $A \setminus U \subset [\![0, n-1]\!] \setminus (U \cup V)$ à $d_n - k$ éléments de qui est de cardinal $n-2k$ (car les parties U et V sont à la fois disjointes, de cardinal k et incluses dans cet ensemble par l'hypothèse $n \geqslant N > \max(U \cup V)$) : il y en a donc $\binom{n-2k}{d_n-k}$.

   La probabilité recherchée est donc

$$\frac{\binom{n-2k}{d_n-k}}{\binom{n}{d_n}} = \frac{(n-2k)!\,d_n!\,(n-d_n)!}{(d_n-k)!\,(n-d_n-k)!\,n!}$$
$$= \frac{(d_n)_k\,(n-d_n)_k}{(n)_{2k}}.$$

4. Notons, pour tout $n \geqslant N$, A_n l'événement étudié à la question précédente. Alors, d'une part, la suite d'événements $(A_n)_{n\geqslant N}$ est composée d'événements mutuellement indépendants ; d'autre part, la série de terme général $\mathbf{P}(A_n)$ diverge. D'après le second lemme de Borel et Cantelli, il existe presque sûrement une infinité d'entiers $n \geqslant N$ tels que l'événement A_n est réalisé. Ainsi,

il existe une infinité de sommets voisins dans G de tous les sommets appartenant à U mais d'aucun sommet appartenant à V.

On peut poursuivre la question avec le premier lemme de Borel et Cantelli (proposition II-5.4) : si la série converge, il existe au plus un nombre fini de sommets voisins dans G de tous les sommets appartenant à U et d'aucun appartenant à V.

5. • Montrons que la condition de la question précédente est satisfaite par ce choix de d_n. Un rapide calcul donne

$$d_n \sim \alpha n, \qquad (n - d_n) \sim (1 - \alpha)n,$$

donc la limite suivante lorsque $n \to +\infty$:

$$\tfrac{(d_n)_k (n-d_n)_k}{(n)_{2k}} \to \big(\alpha(1-\alpha)\big)^k.$$

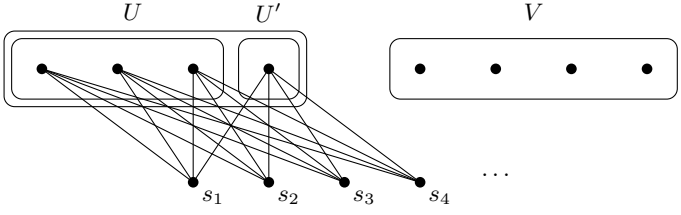
Ainsi, le terme général $\frac{(d_n)_k (n-d_n)_k}{(n)_{2k}}$ converge vers une limite non nulle, et la série associée diverge grossièrement.

Par conséquent, pour toutes parties disjointes U et V de cardinal k, il existe une infinité de sommets voisins de tous les sommets appartenant à U mais d'aucun sommet appartenant à V.

• Soit U et V deux parties finies et disjointes de $\mathbb{N}$.

▷ Si $|U| = |V|$, le résultat précédent permet de conclure avec la question 4.

▷ Sinon, supposons que $|U| < |V|$ (le cas $|V| < |U|$ se traite de manière analogue) et considérons une partie U' disjointe de $U \cup V$ telle que $|U \cup U'| = |V|$; on applique alors la propriété précédente aux parties $U \cup U'$ et V : il existe une infinité de sommets voisins de tous les sommets appartenant à $U \cup U'$ (donc des sommets appartenant à U) mais d'aucun sommet appartenant à V.



Chapitre XI

« Le » graphe aléatoire dénombrable

Dans ce chapitre, on montre un résultat plutôt contre-intuitif : il n'y a, à isomorphisme près, qu'un seul graphe dénombrable vérifiant la propriété d'extension : ainsi, toutes les constructions du chapitre précédent donnent le même graphe à isomorphisme près.

1. Théorème d'unicité

1.1. Théorème. *Deux graphes (dénombrables) vérifiant la propriété d'extension sont isomorphes.*

On se limite dans cette démonstration (sans perte de généralité) à des graphes d'ensembles de sommets $\mathbb{N}^*$.

Démonstration. Soit G et G' deux graphes d'ensemble de sommets $\mathbb{N}^*$ vérifiant la propriété d'extension.

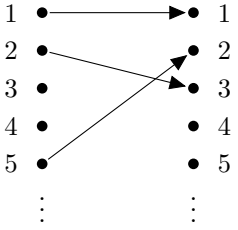
Construisons par récurrence une application $f : \mathbb{N}^* \to \mathbb{N}^*$ qui réalise un isomorphisme entre G et G'.

• Posons $f(1) = 1$ (et donc $f^{-1}(1) = 1$).

Avant de passer à la phase d'hérédité, détaillons ce qui se passe à la deuxième étape.

— *On cherche l'image de* 2 *par* f *: si les sommets* 1 *et* 2 *sont voisins dans* G*, on choisit pour* $f(2)$ *un voisin de* $1 = f(1)$ *dans* G' *; sinon, on choisit pour* $f(2)$ *un sommet non voisin de* $1 = f(1)$ *dans* G'*.*

— *On cherche l'image de* 2 *par* f^{-1}*, c'est-à-dire l'antécédent de* 2 *par* f *: si les sommets* 1 *et* 2 *sont voisins dans* G'*, on choisit pour* $f^{-1}(2)$ *un voisin de* $1 = f^{-1}(1)$ *dans* G *; sinon, on choisit pour* $f^{-1}(2)$ *un sommet non voisin de* $1 = f^{-1}(1)$ *dans* G*.*



• Soit $k \in \mathbb{N}^*$. Supposons que les valeurs $f(i)$ et $f^{-1}(i)$ sont définies pour tout $i \in [\![k]\!]$ de sorte que f réalise un isomorphisme entre les sous-graphes induits $G[S_k]$ et $G'[f(S_k)]$, où S_k désigne l'ensemble des sommets dont on a déjà fixé l'image par f, à savoir

$$S_k = [\![k]\!] \cup \big\{f^{-1}(i),\ i \in [\![k]\!]\big\}.$$

On cherche à définir $f(k+1)$ et $f^{-1}(k+1)$.

▷ Si $k+1 \notin S_k$, c'est-à-dire que l'on a pas encore défini $f(k+1)$, considérons U l'ensemble des sommets de S_k auxquels $k+1$ est connecté dans G et V le complémentaire de U dans S_k, puis $U' = f(U)$ et $V' = f(V)$; les parties U' et V' sont à la fois finies et disjointes.

D'après la propriété d'extension de G', il existe donc un sommet $s \notin U' \cup V' = f(S_k)$ relié à tous les sommets U' et à aucun de V'. Posons $f(k+1) = s$.

▷ Si $k+1 \notin f(S_k \cup \{k+1\})$, c'est-à-dire que l'on a pas encore défini $f^{-1}(k+1)$, considérons U' l'ensemble des sommets de

$f(S_k \cup \{k+1\})$ auxquels $k+1$ est connecté dans G' et V' le complémentaire de U' dans $f(S_k \cup \{k+1\})$, puis $U = f^{-1}(U')$ et $V = f^{-1}(V')$; les parties U et V sont à la fois finies et disjointes.
D'après la propriété d'extension de G, il existe donc un sommet $s \notin U \cup V = S_k$ relié à tous les sommets U et à aucun de V. Posons $f^{-1}(k+1) = s$.

Par construction, l'isomorphisme réalisé par f entre les graphes induits $G[S_k]$ et $G'[f(S_k)]$ est prolongé en un isomorphisme entre les graphes induits $G[S_{k+1}]$ et $G'[f(S_{k+1})]$.

L'application f ainsi définie réalise bien une bijection entre les sommets de G et ceux de G' et forme, par construction, un isomorphisme entre ces deux graphes. ♢

1.2. Remarque. Cette démonstration s'appuie sur une méthode dite d'allers-retours successifs : à chaque étape, on définit l'image du plus petit sommet encore disponible dans G et l'antécédent du plus petit sommet disponible dans G' : ce choix assure notamment la surjectivité de l'application finalement construite.

1.3. Remarque. Le résultat d'unicité indique que les graphes déterministes construits au chapitre précédent pour disposer de la propriété d'extension se comportent globalement comme un graphe aléatoire suivant le modèle de Rado, alors que rien dans leurs constructions ne laisse apparaître d'aléa.

2. Le graphe aléatoire

2.1. Proposition. *Soit p, $p' \in]0,1[$, X, X' des variables aléatoires suivant le modèle de Rado de paramètres p et p' respectivement. Alors, X et X' sont presque sûrement isomorphes.*

Démonstration. Soit G un graphe déterminisme vérifiant la propriété d'extension.

D'après la proposition précédente et le théorème d'unicité pour les graphes vérifiant la propriété d'extension, X est presque sûrement isomorphe à G.

De même, X' est presque sûrement isomorphe à G donc, par composition, X et X' sont presque sûrement isomorphes. ♢

2.2. Remarque. Cette dernière proposition indique qu'il existe essentiellement un seul graphe aléatoire et que le choix du paramètre $p \in]0,1[$ n'importe pas.

Des résultats du chapitre précédent (dans le cours ou les exercices), on déduit les propriétés suivantes du graphe aléatoire.

2.3. Corollaire. *Soit X une variable aléatoire suivant le modèle de Rado de paramètre $\frac{1}{2}$. Alors,*

- ▷ *X est presque sûrement connexe,*
- ▷ *le diamètre de X est presque sûrement 2,*
- ▷ *toute modification finie (c'est-à-dire n'affectant qu'un nombre fini de sommets et d'arêtes) de X est presque sûrement isomorphe à X,*
- ▷ *parmi les graphes obtenus par partition (finie) de X, il en existe au moins un presque sûrement isomorphe à X.*

3. Propriété d'universalité

3.1. Définition. Un graphe $G = (S, A)$ est universel si, pour tout graphe H à ensemble de sommets fini ou dénombrable, il existe une partie $S' \subset S$ tel que le graphe induit $G[S']$ est isomorphe au graphe H.

3.2. Proposition. *Un graphe vérifiant la propriété d'extension est universel.*

La preuve de cette proposition repose sur la construction récursive de l'isomorphisme. On se limite au cas où le graphe recherché est d'ensemble de sommets dénombrable (le cas fini s'obtient en arrêtant la récurrence après un nombre fini d'étapes).

Démonstration. Soit G un graphe d'ensemble de sommets $\mathbb{N}^*$ vérifiant la propriété d'extension et H un graphe quelconque d'ensemble de sommets $\mathbb{N}^*$. Construisons par récurrence une application $f : \mathbb{N}^* \to \mathbb{N}^*$ qui réalise un isomorphisme de H dans $G[f(\mathbb{N}^*)]$.

▷ Posons $f(1) = 1$.

▷ Soit $k \in \mathbb{N}^*$. Supposons f définie sur $[\![k]\!]$ réalisant un isomorphisme entre les graphes induits $H[[\![k]\!]]$ et $G[f([\![k]\!])]$.

Notons U l'ensemble des sommets de $[\![k]\!]$ reliés à $k+1$ dans H et V son complémentaire dans $[\![k]\!]$. Notons $U' = f(U)$ et $V' = f(V)$. Les parties U' et V' étant finies et disjointes, il existe d'après la propriété d'extension un sommet s de G relié à tous les sommets de U' et à aucun de V'. Posons $f(k+1) = s$.

Par construction, f réalise un morphisme injectif de H dans G, donc un isomorphisme de H dans $G[f(\mathbb{N}^*)]$. ♢

En appliquant ce résultat au graphe aléatoire, on obtient le corollaire suivant.

3.3. Corollaire. *Soit X une variable aléatoire suivant le modèle de Rado de paramètre $\frac{1}{2}$ et H un graphe fini ou dénombrable. Alors, presque sûrement, il existe une partie $S \subset \mathbb{N}^*$ telle que le graphe induit $X[S]$ est isomorphe à H.*

4. Exercices

4.1. Exercice. *Montrer qu'il existe un graphe universel qui ne vérifie pas la propriété d'extension.*

Solution. Considérons un graphe G d'ensemble de sommets $\mathbb{N}^*$ et d'ensemble d'arêtes A qui vérifie la propriété d'extension, et considérons le graphe dénombrable G' obtenu en rajoutant un sommet 0 relié à tous les autres, c'est-à-dire

$$G' = \big(\mathbb{N}, A \cup \big\{\{0, n\},\ n \in \mathbb{N}^*\big\}\big).$$

▷ Pour tout graphe H fini ou dénombrable, il existe $S \subset \mathbb{N}^*$ tel que le graphe induit $G[S]$ soit isomorphe à H (conséquence de la propriété

d'extension de G) ; or, $G[S] = G'[S]$ (car $S \subset \mathbb{N}^*$). En conclusion, G' est universel.

▷ En revanche, le sommet 0 est relié à tous les autres sommets, ce qui contredit la propriété d'extension (par exemple avec les deux parties finies et disjointes $U = \{1\}$ et $V = \{0\}$).

4.2. Exercice (Graphe de Henson [Hen71]). *Considérons G un graphe (déterministe) d'ensemble de sommets $\mathbb{N}^*$ vérifiant la propriété d'extension.*

Notons T l'ensemble des entiers $k \geqslant 3$ tels qu'il existe $i < j < k$ vérifiant que le graphe induit $G[\{i, j, k\}]$ est isomorphe au graphe complet K_3.

Considérons enfin le graphe induit $G_3 = G[\mathbb{N}^* \setminus T]$, appelé graphe de Henson d'indice 3. En d'autres termes, on a enlevé le sommet de plus grande étiquette au sein de chaque triangle présent dans G.

1. *Montrer que G_3 est sans triangle, c'est-à-dire qu'il n'y a pas de graphe induit de G_3 isomorphe à K_3.*
2. *Soit H un graphe fini ou dénombrable sans triangle. Montrer qu'il existe un graphe induit de G_3 isomorphe à H.*

Solution.

1. Supposons qu'il existe trois sommets $i < j < k$ de G_3 tels que le graphe induit sur ces sommets soit un triangle. Alors, $G[\{i, j, k\}]$ est aussi un triangle, et donc $k \in T$, ce qui entre en contradiction avec le fait que k soit un sommet de G_3.
2. Soit H un graphe sans triangle d'ensemble de sommets $\mathbb{N}^*$ (on procéderait de même pour un graphe fini). Construisons par récurrence une application $f : \mathbb{N}^* \to \mathbb{N}^*$ qui réalise un isomorphisme de H dans $G_3[f(\mathbb{N}^*)]$.

   ▷ Posons $f(1) = \min \mathbb{N}^* \setminus T$.

   ▷ Soit $k \in \mathbb{N}^*$. Supposons f définie sur $[\![k]\!]$ réalisant un isomorphisme entre les graphes induits $H[[\![k]\!]]$ et $G_3[f([\![k]\!])]$.

   Notons U l'ensemble des sommets de $[\![k]\!]$ reliés à $k+1$ dans H et V son complémentaire dans $[\![k]\!]$.

   Définissons $U' = f(U)$ et $V' = f(V)$. Ces parties U' et V' étant finies et disjointes, il existe d'après la propriété d'extension (telle qu'elle a été obtenue avec la première construction) un sommet s

supérieur à $\max(U' \cup V')$ tel que

$$\{x < s,\ \{x, s\} \in A\} = U'.$$

Si $s \in T$, il existe $f(u_1)$ et $f(u_2) \in U'$ tel que les trois sommets $f(u_1)$, $f(u_2)$ et s forment un triangle de G. Par construction de f (hypothèse de récurrence), cela indique que $\{u_1, u_2\}$ est une arête de H ; or, u_1 et u_2 sont tous deux voisins de $k+1$ et il y a un triangle dans H : contradiction !

Ainsi, $s \notin T$, et l'on peut poser $f(k+1) = s$.

Par construction, f réalise un morphisme injectif de H dans G, donc un isomorphisme de H dans $G[f(\mathbb{N}^*)]$.

Ce résultat se généralise de manière analogue pour trouver un graphe G_p « universel » pour les graphes sans graphe induit isomorphe à K_p pour un entier $p \geqslant 3$ fixé.

4.3. Exercice. *Soit G un graphe d'ensemble de sommet $\mathbb{N}^*$ vérifiant la propriété d'extension, et A, B une partition de $\mathbb{N}^*$.*

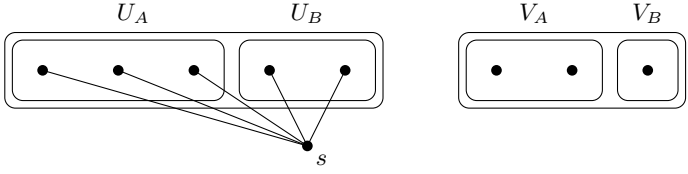
Montrer qu'au moins l'un des deux graphes induits $G[A]$ et $G[B]$ est isomorphe à G.

Solution. • Montrons que l'un des deux graphes induits vérifie la propriété d'extension.

Pour cela raisonnons par l'absurde en supposant qu'il existe

- ▷ un couple (U_A, V_A) de parties finies et disjointes de A tel qu'aucun sommet de A n'est relié à tous les sommets de U_A et à aucun de V_A,
- ▷ un couple (U_B, V_B) de parties finies et disjointes de B tel qu'aucun sommet de B n'est relié à tous les sommets de U_B et à aucun de V_B.

Appliquons alors la propriété d'extension de G au couple de parties finies et disjointes $(U_A \cup U_B, V_A \cup V_B)$: il existe un sommet s qui est relié à tous les sommets de $U_A \cup U_B$ et à aucun de $V_A \cup V_B$.



En discutant selon que $s \in A$ ou $s \in B$ (car A et B forment une partition de l'ensemble des sommets), on obtient une contradiction avec le choix des différentes parties finies.

• Le graphe induit vérifiant la propriété d'extension est isomorphe à G, d'après le théorème d'unicité.

Cette propriété de partition n'est pas caractéristique des graphes vérifiant la propriété d'extension puisque les graphes « triviaux »

$$\big(\mathbb{N}^*, \mathscr{P}_2(\mathbb{N}^*)\big) \text{ et } \big(\mathbb{N}^*, \varnothing\big)$$

la vérifient également.

4.4. Exercice. *Soit G un graphe vérifiant la propriété d'extension, S et S' deux parties finies de l'ensemble des sommets et $f : S \to S'$ réalisant un isomorphisme entre les graphes induits $G[S]$ et $G[S']$.*

Montrer que la bijection f peut être prolongée en un automorphisme de G.

Solution. Étendons f par allers-retours successifs : on ne rédige ici que l'étape d'hérédité de la récurrence.

- ▷ Soit x le plus petit sommet de G qui n'appartient pas à S. Notons U l'ensemble des sommets de S voisins de x et V son complémentaire dans S puis $U' = f(U)$ et $V' = f(V)$. D'après la propriété d'extension de G avec les parties finies et disjointes U' et V', il existe un sommet $x' \notin U' \cup V'$ voisin de tous les éléments de U' et d'aucun de V'. Posons $f(x) = x'$.
- ▷ Soit y' le plus petit sommet de G qui n'appartient pas à $S' \cup \{x'\}$. Notons U' l'ensemble des sommets de $S' \cup \{x'\}$ voisins de y' et V' son complémentaire dans $S' \cup \{x'\}$.
  Posons $U = f^{-1}(U')$ et $V = f^{-1}(V')$. D'après la propriété d'extension de G avec les parties finies et disjointes U et V, il existe un sommet $y \notin U \cup V$ voisin de tous les éléments de U et d'aucun de V. Posons $f(y) = y'$.

Par construction, f réalise désormais un isomorphisme entre les graphes induits $G[S \cup \{x, y\}]$ et $G[S' \cup \{x', y'\}]$. En poursuivant la construction par récurrence, on obtient un automorphisme de G (le caractère surjectif provient une fois encore du choix des plus petits éléments disponibles dans l'étape de prolongement de f par allers-retours).

Ce résultat indique essentiellement que le graphe vérifiant la propriété d'extension est « homogène », c'est-à-dire hautement symétrique.

Pour aller plus loin

Je ne suis pas collectionneur mais il y a dans ma bibliothèque quelques livres fort rares et que bien peu de personnes possèdent. Je me flatte en outre de les avoir écrits.
Éric Chevillard, L'autofictif

Le texte dont vous venez d'achever la lecture avait pour unique ambition de vous faire découvrir un sujet mathématique passionnant ; comme toute tentative d'introduction, il est loin d'être exhaustif et, désormais, vous avez sûrement envie d'en savoir davantage. Aussi ai-je rassemblé quelques références plus poussées afin de vous permettre de poursuivre l'aventure. Il s'agit souvent d'ouvrages avancés (en anglais) abordant de nombreux sujets avec des listes d'exercices (cette fois-ci non corrigés) et s'adressant à un public plus expérimenté. Je les ai choisis selon mes goûts personnels parmi une littérature plutôt fournie puis je les ai triés selon les parties de cet ouvrage.

Rappels et compléments

▷ Il existe de très nombreux livres de théorie des graphes, certains très mathématiques, d'autres se focalisant sur des sujets plus algorithmiques. Pour poursuivre dans la veine proposée ici, on pourra

consulter les classiques [BM08] (les graphes aléatoires sont évoqués dans le chapitre 13) ou [Die05] (l'avant dernier chapitre de ce livre a pour sujet les graphes aléatoires). Ces (lourds) volumes contiennent toutes les bases nécessaires sur ce sujet et peuvent être lus linéairement (si l'on dispose de suffisamment de temps et de courage) mais aussi être consultés plus ponctuellement pour répondre à un besoin spécifique.

▷ Pour les probabilités, il y a également beaucoup de références disponibles selon les goûts de chaque lecteur. Je confesse une affection particulière pour les deux ouvrages élémentaires de Sheldon Ross [Ros12] et [Ros00]. Pour poursuivre avec des questions plus élaborées (théorèmes limites ou processus), on pourra consulter le joli texte de Richard Durrett [Dur10].

Graphes aléatoires finis

Pour la méthode probabiliste, la situation est plus simple : il existe une somme incontournable, un ouvrage à lire sur ce sujet, [AS16]. Les auteurs y passent en revue de nombreuses questions de combinatoire grâce aux probabilités. Le spectre de l'ouvrage est large et inégalé.

Asymptotique de graphes aléatoires finis

L'étude asymptotique est souvent le commencement des ouvrages sur les graphes aléatoires : il ne s'agit que d'un ou deux chapitres (sur 26) dans [FK15] et d'un chapitre dans [Bol01]. En revanche, on trouve dans ces références bien d'autres propriétés des graphes aléatoires, des modèles étendus, des graphes aléatoires obtenus par d'autres processus de construction, des applications à la modélisation...

L'ouvrage [Gri10] mérite d'être cité même si l'auteur adopte un point de vue différent : faire des probabilités (par exemple des

marches aléatoires) sur un graphe déterministe ; à l'exception, notable, du chapitre 11, il est disjoint de notre volume.

Plus récent, l'excellent livre [Cur25] issu des cours avancés de Nicolas Curien s'impose comme une référence incontournable. La modeste introduction que vous avez entre les mains est complètement reprise dans les premières pages de la partie II de cette somme.

Graphe aléatoire dénombrable

Le graphe aléatoire dénombrable n'est pas à proprement parler le sujet d'une monographie. Toutefois, les excellents textes de Peter Cameron [Cam97] et [Cam01] offrent une synthèse très complète, qui dépasse, de loin, les aspects traités dans la partie correspondante de notre volume. On y trouvera notamment les liens avec la logique et la théorie des ensembles.

Bibliographie

[Ack37] W. Ackermann. Die Widerspruchsfreiheit der allgemeinen Mengenlehre. *Mathematische Annalen*, 114(1) :305–315, 1937.

[AS16] N. Alon and J. H. Spencer. *The Probabilistic Method*. Wiley, 4th edition, 2016.

[BD18] C. Biró and U. Darji. Generating Infinite Random Graphs. *Proceedings of the Edinburgh Mathematical Society*, 61(3) :847–868, 2018.

[BE76] B. Bollobás and P. Erdős. Cliques in Random Graphs. *Mathematical Proceedings of the Cambridge Philosophical Society*, 80(3) :419–427, 1976.

[BE84] A. B. Barak and P. Erdős. On the Maximal Number of Strongly Independent Vertices in a Random Acyclic Directed Graph. *SIAM J. Algebraic Discrete Methods*, 5 :508–514, 1984.

[BM08] J. A. Bondy and U. S. R. Murty. *Graph Theory*. Springer, 2008.

[BMV17] I. Belghiti, R. Mansuy, and J-J. Vie. *Les clés pour l'info*. Calvage et Mounet, 2017.

[Bol01] B. Bollobás. *Random Graphs*. Cambridge University Press, 2001.

[BT87] B. Bollobás and A. G. Thomason. Threshold Functions. *Combinatorica*, 7(1) :35–38, 1987.

[Cam97] P. J. Cameron. The Random Graph. In *The Mathematics of Paul Erdős II*, pages 333–351. Springer, 1997.

[Cam01] P. J. Cameron. The Random Graph Revisited. In *European Congress of Mathematics*, pages 267–274. Birkhäuser Basel, 2001.

[Cur25] N. Curien. *A random walk among random graphs*. Société Mathématique de France, 2025.

[Die05] R. Diestel. *Graph Theory*. Springer, 2005.

[Dur10] R. Durrett. *Probability : Theory and Examples*. Cambridge University Press, 2010.

[ER59] P. Erdős and A. Rényi. On Random Graphs I. *Publicationes Mathematicae Debrecen*, 6 :290–297, 1959.

[ER60] P. Erdős and A. Rényi. On the Evolution of Random Graphs. In *Publication of the Mathematical Institute of the Hungarian Academy of Sciences*, pages 17–61, 1960.

[ER63] P. Erdős and A. Rényi. Asymmetric Graphs. *Acta Mathematica Academiae Scientiarum Hungarica*, 14(3) :295–315, 1963.

[Erd47] P. Erdős. Some Remarks on the Theory of Graphs. *Bulletin of the American Mathematical Society*, 53(4) :292–294, 1947.

[FK15] A. Frieze and M. Karoński. *Introduction to Random Graphs*. Cambridge University Press, 2015.

[Gil59] E. N. Gilbert. Random Graphs. *Ann. Math. Statist.*, 30(4) :1141–1144, 1959.

[Gri10] G. Grimmett. *Probability on Graphs : Random Processes on Graphs and Lattices*. Cambridge University Press, 2010.

[Hen71] C. W. Henson. A Family of Countable Homogeneous Graphs. *Pacific J. Math.*, 38(1) :69–83, 1971.

[KM18] I. Kortchemski and R. Mansuy. *Probabilités*. Vuibert, 2018.

[Pós76] L. Pósa. Hamiltonian Circuits in Random Graphs. *Discrete Math.*, 14(4) :359–364, 1976.

[Rad64] R. Rado. Universal Graphs and Universal Functions. *Acta Arithmetica*, 9(4) :331–340, 1964.

[Ros00] S. Ross. *Introduction to Probability Models*. Academic Press, 2000.

[Ros12] S. Ross. *A First Course in Probability*. Pearson, 2012.

[Sze43] T. Szele. Kombinatorikai vizsgalatok az iranyitott teljes graffal kapcsolatban. *Mat. Fiz. Lapok*, 50 :223–256, 1943.

Index

arbre, 16
    couvrant, 17

cerf-volant, 24
chemin hamiltonien, 89, 117
couplage, 48, 56, 138

ensemble
    dominant, 91
    indépendant, 92
    dominant, 18
    indépendant, 18
espérance, 35

fonction de seuil, 133, 143, 156
formule
    de Cayley, 126
    d'Euler, 82
    de Stirling, 7
    du crible, 152

graphe
    biparti complet, 12
    chemin, 11
    complet, 12
    connexe, 13
    cycle, 11
    de Barak-Erdős, 71
    de Henson, 190
    de Rado, 163
    diamètre d'un, 13
    divisoriel, 10
    griffe, 12
    induit, 15
    sous-, 16

inégalité
    de Markov, 102
    de Bienaymé et Tchebychev, 119
    de Bonferroni, 153
    de Cantelli, 125
isomorphisme
    de graphe, 10

lemme
    de Borel et Cantelli, 39, 40
    de Posá, 28

maille
    d'un graphe, 18
modèle
    binomial, 51
    uniforme, 61
moment factoriel
    d'une variable de comptage, 151
    d'une variable à valeurs entières, 147

nombre
    chromatique, 19
    d'automorphismes, 23
    de croisements, 81
    de Ramsey, 78

propriété
    d'extension, 73, 171
    d'homogénéité, 192
    d'universalité, 188
    de partition, 191
propriété croissante, 131

sommet
    degré d'un, 14
    isolé, 14, 58, 103, 121

théorème
    de la double exponentielle d'Erdős–Rényi, 160
    de Posá, 114
    de Szele, 88
    de Turán, 97

variable
    de comptage, 32, 151
    indicatrice, 32
variance, 36