Chapitre 9

La méthode probabiliste

La méthode probabiliste est une technique efficace pour prouver l'existence d'objets combinatoires ayant des propriétés spécifiques. Elle est basée sur la théorie de la probabilité mais, surprenamment, elle peut être utilisée pour prouver des théorèmes n'ayant rien à voir avec les probabilités. Dans ce chapitre, nous donnons les outils et techniques les plus usuels de cette méthode.

9.1 Petit rappel de probabilités

Un *espace de probabilités (fini)* est un couple $(\Omega, \mathbf{Pr})$ où Ω est un ensemble fini appelé *ensemble échantillon* ou *ensemble des observables*, et $\mathbf{Pr}$ est une fonction, appelée *fonction de probabilité*, de Ω dans $[0, 1]$ telle que $\sum_{\omega\in\Omega} \mathbf{Pr}(w) = 1$. Nous considérerons souvent une *distribution uniforme* pour laquelle $\mathbf{Pr}(\omega) = \frac{1}{|\Omega|}$ pour tout $\omega \in \Omega$.

On peut voir l'ensemble des graphes étiquetés à n sommets $\mathcal{G}_n$ comme l'espace échantillon d'un espace de probabilité $(\mathcal{G}_n, \mathbf{Pr})$. Le résultat de la sélection d'un élément G de cet ensemble échantillon suivant la fonction de probabilité $\mathbf{Pr}$ est appelé *graphe aléatoire*.

L'exemple le plus simple d'un tel espace de probabilité est celui pour lequel tous les graphes $G \in \mathcal{G}_n$ ont la même probabilité d'être choisis. Comme $|\mathcal{G}_n| = 2^{\binom{n}{2}}$, dans ce cas, la fonction de probabilité est $\mathbf{Pr}(G) = 2^{-\binom{n}{2}}$ pour tout $G \in \mathcal{G}_n$. Une manière naturelle de voir cet espace de probabilité est de considérer les arêtes de K_n une par une et de choisir chacune d'entre elles pour $E(G)$ avec probabilité $\frac{1}{2}$, ces choix étant faits indépendamment les uns des autres. Le résultat d'une telle procédure est un graphe couvrant de K_n avec tous les $G \in \mathcal{G}_n$ équiprobables. Un espace de probabilité sur $\mathcal{G}_n$ plus élaboré peut être obtenu en fixant un réel p entre 0 et 1 et en choisissant chaque arête avec probabilité p, ces choix étant faits indépendamment les uns des autres. Ainsi $1 - p$ est la probabilité qu'une arête particulière ne soit pas choisie. La fonction de probabilité est alors $\mathbf{Pr}(G) = p^m(1 - p)^{\binom{n}{2}-m}$ pour tout $G \in \mathcal{G}_n$, avec $m = |E(G)|$. Cet espace est noté $\mathcal{G}_{n,p}$. Par exemple, $\mathcal{G}_{3,p}$ a pour espace échantillon des $2^{\binom{3}{2}} = 8$ sous-graphes couvrants de K_3 représentés Figure 9.1 avec la fonction de probabilité indiquée.

Notons que plus p est petit, plus la probabilité est grande d'avoir un graphe avec peu d'arêtes.

A chaque propriété de graphe, comme la connexité par exemple, correspond l'ensemble des graphes de $\mathcal{G}_n$ vérifiant cette propriété. La probabilité qu'un graphe aléatoire satisfasse cette propriété particulière est alors la somme des probabilités des graphes de ce sous-ensemble de $\mathcal{G}_n$. Par exemple, la probabilité qu'un graphe aléatoire de $\mathcal{G}_{3,p}$ soit connexe vaut $3p^2(1 - p) + p^3 = p^2(3 - 2p)$, la probabilité qu'il soit biparti est $(1 - p)^3 + 3p(1 - p)^2 + 3(1 - p)p^2 = (1 - p)(1 + p + p^2)$, et la probabilité qu'il soit à la fois connexe et biparti est $3p^2(1 - p)$.

Ceci est capturé par la notion d'événement. Dans un espace de probabilité $(\Omega, \mathbf{Pr})$, un *événement* est un sous-ensemble A de Ω. La *probabilité* d'un événement A est définie par

$$\mathbf{Pr}(A) = \sum_{\omega\in A} \mathbf{Pr}(\omega).$$

Par définition, nous avons

$$\mathbf{Pr}(A \cup B) = \mathbf{Pr}(A) + \mathbf{Pr}(B) - \mathbf{Pr}(A \cap B) \tag{9.1}$$

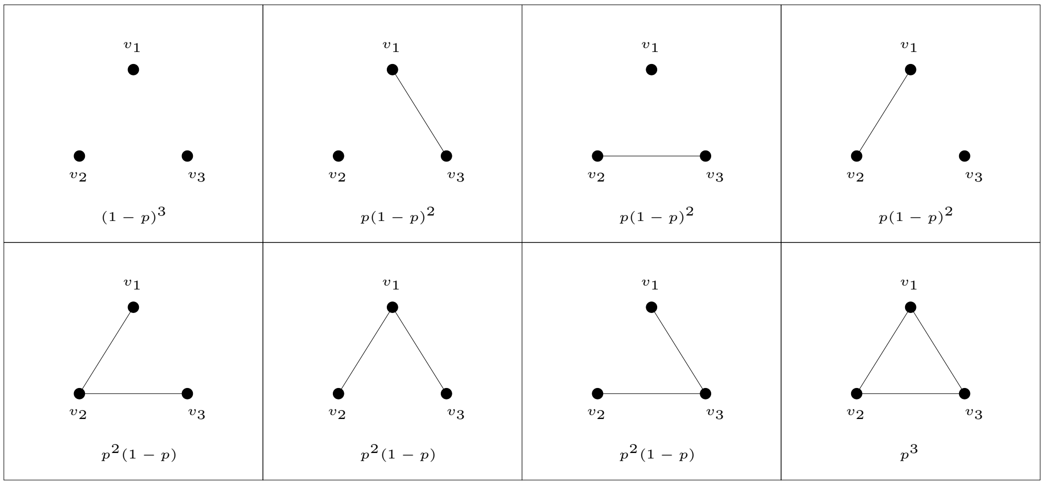

FIG. 9.1 – L'espace de probabilité $\mathcal{G}_{3,p}$

En particulier, $\mathbf{Pr}(A \cup B) \leq \mathbf{Pr}(A) + \mathbf{Pr}(B)$ et plus généralement,

$$\mathbf{Pr}(A_1 \cup \cdots \cup A_n) \leq \mathbf{Pr}(A_1) + \ldots \mathbf{Pr}(A_n).$$

Ces dernière inéquation est appelée la **Sous-Additivité des Probabilités**.

9.1.1 Probabilité conditionnelle et indépendance

Pour deux événements A et B, la *probabilité (conditionnelle) de A sachant B*, notée $\mathbf{Pr}(A|B)$ est, de manière informelle, la probabilité que A se produise si on sait que B se produit. Formellement, $\mathbf{Pr}(A|B) = \frac{\mathbf{Pr}(A \cap B)}{\mathbf{Pr}(B)}$ (si $\mathbf{Pr}(B) = 0$ alors $\mathbf{Pr}(A|B) = \mathbf{Pr}(A)$).

Intuitivement, A est indépendant de B si savoir si B se produit ou pas ne change pas pas la probabilité que A se produise. Formellement, A est *indépendant* de B si $\mathbf{Pr}(A|B) = \mathbf{Pr}(A)$, c'est à dire si $\mathbf{Pr}(A \cap B) = \mathbf{Pr}(A)\,\mathbf{Pr}(B)$. Remarquons que cela implique que $\mathbf{Pr}(B|A) = \mathbf{Pr}(B)$, i.e. B est indépendant de A. Ainsi, nous pouvons parler d'*événements indépendants*. Deux événements non-indépendants sont dits *dépendants*.

Par exemple, si A est l'événement 'G est connexe' et B l'événement 'G est biparti' dans l'espace de probabilité $\mathcal{G}_{3,p}$, alors (sauf si $p = 0$ ou $p = 1$)

$$\mathbf{Pr}(A)\,\mathbf{Pr}(B) = p^2(3 - 2p)(1 - p)(1 + p + p^2) \neq 3p^2(1 - p) = \mathbf{Pr}(A \cap B).$$

Donc ces deux événements sont dépendants. En d'autres termes, la connaissance que G est connexe a une influence sur la probabilité d'être biparti et vice-versa.

Soit A_i, $i \in I$ un ensemble (fini) d'événements. Ils sont *deux à deux indépendants* si pour tout $i \neq j$, A_i et A_j sont indépendants. Ils sont *mutuellement indépendants* si pour tout $i \in I$ et $J \subset I \setminus \{i\}$ alors

$$\mathbf{Pr}\left(A_i | \bigcap_{j \in J} A_j\right) = \mathbf{Pr}(A_i).$$

Il est facile de voir que si les A_i, $i \in I$ sont mutuellement indépendants alors pour tout sous-ensemble S de I,

$$\mathbf{Pr}\left(\bigcap_{i \in S} A_i\right) = \prod_{i \in S} \mathbf{Pr}(A_i).$$

Il est important de noter qu'un ensemble d'événements peuvent ne pas être mutuellement indépendants tout en étant deux à deux indépendants. Voir l'Exercice 9.3.

Un événement A est *mutuellement indépendant* d'un ensemble d'événements $\{B_j \mid j \in J\}$ si, pour tout sous-ensemble S de J, $\mathbf{Pr}\left(A | \bigcap_{j\in S} B_j\right) = \mathbf{Pr}(A)$. Il est facile de cette condition est équivalente à la condition plus forte suivante : pour tout $J_1 J_2 \subset J$ tels que $J_1 \cap J_2 = \emptyset$ alors $\mathbf{Pr}\left(A | \bigcap_{j\in J_1} B_j \cap \bigcap_{j\in J_2} \overline{B_j}\right) = \mathbf{Pr}(A)$.

9.1.2 Variables aléatoires et espérance

Une grande partie de la théorie des graphes concerne l'étude de paramètres basiques tels que la connectivité, la taille maximale d'une clique oun d'un stable, le nombre chromatique. Comme nous l'avons vu, les valeurs de ces paramètres donnent beaucoup de renseignements sur un graphe et ses propriétés. Dans le contexte des graphes aléatoires, de telles fonctions sont appelées des variables aléatoires, car elles dépendent du graphe qui est sélectionné. Plus généralement, une *variable aléatoire* sur un espace de probabilité $(\Omega, \mathbf{Pr})$ est une fonction de l'espace échantillon à valeur dans $\mathbb{R}$.

L'*espérance* d'une variable aléatoire X est :

$$\mathbf{E}(X) = \sum_{\omega\in\Omega} \mathbf{Pr}(\omega) X(\omega).$$

Intuitivement, $\mathbf{E}(X)$ est la valeur attendue si l'on effectue un grand nombre de tirages aléatoires et que l'on prend la moyenne sur les valeurs obtenues pour X.

Par exemple, si X est une variable aléatoire qui vaut 1 avec probabilité p et 0 avec probabilité $1-p$, alors l'espérance de X est $p \times 1 + (1-p) \times 0 = p$.

Linéarité de l'Espérance *Si* $X = \sum_{i=1}^n X_i$ *alors* $\mathbf{E}(X) = \sum_{i=1}^n \mathbf{E}(X_i)$.

Démonstration.

$$\begin{aligned}
\mathbf{E}(X) = \mathbf{E}\left(\sum_{i=1}^n X_i\right) &= \sum_{\omega\in\Omega} \mathbf{Pr}(\omega) \sum_{i=1}^n X_i(\omega) \\
&= \sum_{i=1}^n \left(\sum_{\omega} \mathbf{Pr}(\omega) \times X_i(\omega)\right) \\
&= \sum_{i=1}^n \mathbf{E}(X_i)
\end{aligned}$$

□

La Linéarité de l'Espérance est un outil très utile pour les calculs d'espérances. Par exemple, puisque dans un graphe de $\mathcal{G}_{n,p}$, il y a $\binom{n}{2}$ arêtes, chacune étant présente avec probabilité p alors l'espérance du nombre d'arêtes est $p\binom{n}{2}$.

Ceci est un cas particulier d'un paradigme général qui apparaît fréquemment. De nombreuses variables sont des sommes de variables 0-1 et donc leur espérance peut-être calculée comme la somme des espérances de ces variables 0-1. La variable aléatoire qui est la somme de n variables 0-1 valant 1 avec probabilité p et 0 avec probabilité p est notée $BIN(n,p)$. La linéarité de l'espérance nous donne : $\mathbf{E}(BIN(n,p)) = np$.

9.2 Méthode de base

La principe général de la méthode probabiliste est le suivant. On désire montrer l'existence d'un objet combinatoire satisfaisant une propriété donnée. Pour cela, on considère un objet aléatoire dans un espace de probabilité bien choisi et on calcule la probabilité que cet objet ait la propriété en question. Si on prouve que cette probabilité est non-nulle, alors on en déduit qu'un tel objet existe : en effet, si aucun objet ne satisfaisait la propriété alors la probabilité vaudrait zéro.

9.2.1 2-coloration d'hypergraphes

Un *hypergraphe* est un graphe généralisé dans lequel les arêtes peuvent contenir plus de deux sommets. Ainsi une *hyperarête* peut être n'importe quel ensemble d'au moins 2 sommets. Un hypergraphe est k-*uniforme* si chaque arête est incidente à k sommets. Ainsi un hypergraphe 2-uniforme est simplement un graphe.

Une 2-*coloration propre* d'un hypergraphe est une application de l'ensemble des sommets dans un ensemble de deux couleurs telle qu'aucune arête ne soit *monochromatique*, i.e. dont tous les sommets reçoivent la même couleur. Un hypergraphe est 2-*colorable* s'il admet une 2-coloration propre.

Théorème 9.1. *Si H est un hypergraphe k-uniforme avec moins de 2^{k-1} hyperarêtes alors H est 2-colorable.*

Démonstration. Colorons les sommets aléatoirement de manière indépendante avec la couleur rouge avec probabilité 1/2 et la couleur bleue avec probabilité 1/2. Autrement dit, considérons une 2-coloration aléatoire uniforme. Pour toute hyperarête e, soit A_e l'événement que e est monochromatique. Alors $\mathbf{Pr}(A_e) = 2 \times 2^{-k} = 2^{1-k}$ car la probabilité d'être monochromatique bleue (ou rouge) est 2^{-k}. Par la Sous-Additivité des Probabilités, la probabilité qu'au moins un de ces événements se produise est $\sum_{e \in E(H)} \mathbf{Pr}(A_e) \leq |E(H)| \times 2^{1-k} < 1$. Donc la probabilité qu'aucune arête ne soit monochromatique est non-nulle. □

9.2.2 Nombres de Ramsey

Le Théorème de Ramsey affirme que tout graphe de taille suffisamment grande contient un stable ou une clique de taille donnée. Ainsi quels que soient deux entiers k et l, le plus petit entier n tel que tout graphe à n sommets contient soit une clique de taille k soit un stable de taille l, existe donc : c'est le *nombre de Ramsey* $R(k,l)$. Cependant la valeur de $R(k,l)$ est uniquement connu dans un petit nombre de cas. Nous donnons maintenant une borne inférieure de $R(k,k)$.

Théorème 9.2. *Pour tout $k \geq 3$, $R(k,k) > 2^{k/2-1}$.*

Démonstration. Soit G un graphe aléatoire de $\mathcal{G}_{n,\frac{1}{2}}$. Soit S ensemble S de k sommets. On note A_S l'événement que S soit un stable ou une clique. La probabilité que S soit un stable est $2^{-\binom{k}{2}}$, de même que celle d'être une clique. Donc $\mathbf{Pr}(A_S) = 2 \times 2^{-\binom{k}{2}}$. Comme il y a $\binom{n}{k}$ ensembles de k sommets, la Sous-Additivité des Probabilités nous donne

$$\mathbf{Pr}(G \text{ contient un stable ou une clique de taille } k) = \mathbf{Pr}\left(\bigcup A_S\right) \leq 2\binom{n}{k}2^{-\binom{k}{2}}.$$

Il nous faut montrer que cette expression est strictement inférieure à 1. Comme $\binom{n}{k} \leq n^k$, il suffit de montrer que $2n^k < 2^{k(k-1)/2}$. Ceci est certainement vrai si $n \leq 2^{k/2-1}$. Ainsi il existe des graphes d'ordre $\lfloor 2^{k/2-1} \rfloor$ qui ne contiennent ni clique de taille k, ni stable de taille k. □

9.3 Méthode du Premier Moment

La Méthode du Premier Moment est l'outil le plus fondamental de la méthode probabiliste. Elle repose sur les deux énoncés simples et surprenamment puissants :

Principe du Premier Moment *Si $\mathbf{E}(X) \leq t$ alors $\mathbf{Pr}(X \leq t) > 0$.*

Démonstration. De manière intuitive, l'espérance de X peut être vue comme la moyenne de X sur tous les résultats possibles. Si chaque résultat est plus grand que t alors la moyenne est nécessairement plus grande que t.

Plus formellement, comme l'espace est fini, X ne peut prendre qu'un ensemble fini I de valeurs. Ainsi, $\mathbf{E}(X) = \sum_{i \in I} i \times \mathbf{Pr}(X = i)$, si $\mathbf{Pr}(X \leq t) = 0$ nous avons $\mathbf{E}(X) = \sum_{i > t} i \times \mathbf{Pr}(X = i) > t \times \sum_i > t\,\mathbf{Pr}(X = i) = t$.

□

Inégalité de Markov Pour tout variable aléatoire positive X,

$$\mathbf{Pr}(X \geq t) \leq \frac{\mathbf{E}(X)}{t}.$$

Démonstration. Puisque $\mathbf{E}(X) = \sum_i i \times \mathbf{Pr}(X = i)$ et comme X est toujours positive, $\mathbf{E}(X) \geq \sum_{i \geq t} i \times \mathbf{Pr}(X = i) \geq t \times \mathbf{Pr}(X \geq t)$. □

Pour appliquer la Méthode du Premier Moment, il faut faire un choix judicieux de la variable aléatoire X et faire le calcul (souvent aisé) de son espérance. Souvent X est positive et entière et on montre que $\mathbf{E}(X)$ est inférieure à 1, prouvant ainsi que $\mathbf{Pr}(X = 0)$ est strictement positive. Typiquement, on utilise l'Inégalité de Markov lorsque X est positive et entière et $\mathbf{E}(X)$ est inférieure à 1. On a alors $\mathbf{Pr}(X > 0) = \mathbf{Pr}(X \geq 1) \leq \mathbf{E}(X)$.

Puisque $\mathbf{E}(X) = \sum_i i \times \mathbf{Pr}(X = i)$, on pourrait penser à première vue que pour calculer $\mathbf{E}(X)$, il faille calculer $\mathbf{Pr}(X = i)$ pour tous les i, ce qui est au moins aussi difficile que de calculer directement $\mathbf{Pr}(X \leq t)$. Cependant la Linéarité de l'Espérance permet souvent de calculer $E(X)$ sans calculer tous les $\mathbf{Pr}(X = i)$.

9.3.1 2-coloration d'hypergraphes

Le Théorème 9.1 peut se montrer en utilisant le Principe du Premier Moment au lieu de la Sous-Additivité des Probabilités.

Considérons une 2-coloration aléatoire uniforme. Pour toute hyperarête e, soit X_e la variable aléatoire valant 1 si e est monochromatique et 0 sinon. Alors $X = \sum_{e \in E(H)} X_e$ est le nombre d'hyperarêtes monochromatiques de H. Comme chaque hyperarête est monochromatique avec probabilité 2^{1-k}, pour tout $e \in E(H)$, $\mathbf{E}(X_e) = 2^{1-k}$. Donc par la Linéarité de l'Espérance, $\mathbf{E}(X) = \sum_{e \in E(H)} \mathbf{E}(X_e) = |E(H)| \times 2^{1-k} < 1$. Ainsi, par le Principe du Premier Moment, la probabilité que $X = 0$, c'est-à-dire qu'il n'y ait pas d'hyperarête monochromatique, est non-nulle.

En fait, la Sous-Addivité des Probabilités est un cas particulier du Principe du Premier Moment. (Voir Exercice 9.4). Ce dernier permet cependant parfois d'obtenir des résultats plus généraux. Par exemple, le Théorème 9.1 peut être généralisé de la manière suivante :

Théorème 9.3. *Un hypergraphe k-uniforme H avec m hyperarêtes peut être 2-coloré de manière à ce qu'au plus $\frac{m}{2^{r-1}}$ hyperarêtes soient monochromatiques.*

Démonstration. Si X est la variable aléatoire du nombre d'hyperarêtes monochromatiques de H, le calcul précédent nous donne $\mathbf{E}(X) = m \times 2^{1-k}$. Ainsi par le Principe du Premier Moment, $\mathbf{Pr}(X \leq m \times 2^{1-k}) > 0$. □

En fait, on peut légérement améliorer ce résultat. En effet, il existe des 2-colorations qui ont strictement plus de $\mathbf{E}(X) = m \times 2^{1-k}$ hyperarêtes monochromatiques. Par exemple, les deux 2-colorations pour lesquelles tous les sommets sont d'une même couleur ont toutes les hyperarêtes monochromatiques. Ainsi il faut des 2-colorations pour "compenser" celles-ci.

Théorème 9.4. *Un hypergraphe k-uniforme H avec m hyperarêtes peut être 2-coloré de manière à ce que strictement moins de $\frac{m}{2^{r-1}}$ hyperarêtes soient monochromatiques.*

Démonstration. $\mathbf{E}(X) = \sum_{i < \mathbf{E}(X)} i \times \mathbf{Pr}(X = i) + \sum_{i = \mathbf{E}(x)} i \times \mathbf{Pr}(X = i) + \sum_{i > \mathbf{E}(X)} i \times \mathbf{Pr}(X = i)$. Ainsi si $\mathbf{Pr}(X > E(X)) > 0$ alors $\sum_{i > \mathbf{E}(X)} i \times \mathbf{Pr}(X = i) > 0$ et donc $\sum_{i < \mathbf{E}(X)} i \times \mathbf{Pr}(X = i) > 0$ d'où $\mathbf{Pr}(X < E(X)) > 0$. □

9.3.2 Graphes de grande maille et grand nombre chromatique

Le Théorème 6.4 nous dit qu'il existe des graphes de maille 4 et de nombre chromatique aussi grand que l'on veut. Ce théorème peut être généralisé pour des graphes de maille arbitrairement grande.

Théorème 9.5 (Erdös, 1959). *Pour toute paire d'entiers naturels g et k, il existe un graphe G telle que $maille(G) > g$ et $\chi(G) > k$.*

Démonstration. D'après la Proposition 6.5, $\chi(G) \geq \frac{|V(G)|}{\alpha(G)}$. Ainsi pour montrer que $\chi(G) > k$, il suffit de montrer que $\alpha(G) < \frac{|V(G)|}{k}$.

Soit $\theta < \frac{1}{g}$. Soit n suffisamment grand et G un graphe aléatoire de $\mathcal{G}_{n,p}$ avec $p = n^{\theta-1}$. Soit X le nombre de cycles de G de longueur au plus g. Il y a $\frac{n!}{2i(n-i)!}$ cycles potentiels de longueur i car il y a $\frac{n!}{(n-i)!}$ suites de i sommets distincts et un cycle de taille i correspond à $2i$ de ces suites. De plus, chaque cycle existe avec probabilité p^i. D'après la Linéarité de l'Espérance

$$\mathbf{E}(X) = \sum_{i=3}^{g} \frac{n!}{2i(n-i)!} p^i \leq \sum_{i=3}^{g} \frac{n^{\theta i}}{2i} = o(n),$$

puisque $\theta g < 1$. D'après l'Inégalité de Markov,

$$\mathbf{Pr}\left(X \geq \frac{n}{2}\right) \leq \frac{2\,\mathbf{E}(X)}{n} = o(1).$$

Posons $x = \left\lceil \frac{3}{p} \ln(n) \right\rceil$. Nous avons

$$\mathbf{Pr}(\alpha(G) \geq x) \leq \binom{n}{x} (1-p)^{\binom{x}{2}} < \left(ne^{-p(x-1)/2}\right)^x = o(1).$$

Soit n suffisamment grand pour que les deux événements aient probabilité inférieure à $\frac{1}{2}$. Alors il existe un graphe G avec moins de $n/2$ cycles de longueur au plus g et avec $\alpha(G) < x < 3n^{1-\theta}\ln(n)$. Enlevons de G un sommet par cycle de taille au plus g. Nous obtenons un graphe G' sans cycle de longueur au plus g et qui a au moins $n/2$ sommets. Ainsi G' a maille supérieure à g et $\alpha(G') \leq \alpha(G)$. Ainsi

$$\chi(G') \geq \frac{|V(G')|}{\alpha(G')} \geq \frac{n/2}{3n^{1-\theta}\ln(n)} = \frac{n^\theta}{6\ln(n)}.$$

En prenant n suffisamment grand, ceci est plus grand que k.

□

9.4 Lemme Local

Dans les preuves typiques de résultats combinatoires, on doit généralement prouver que la probabilité d'un certain événement est non-nulle. Dans la plupart des cas, nous avons un ensemble fini de "mauvais événements" $A_i, i \in I$ et nous voulons montrer qu'avec une probabilité non-nulle aucun de ces événements ne se produit, c'est-à-dire $\mathbf{Pr}\left(\bigcap_{i\in I} \overline{A_i}\right) > 0$. Une des manières de faire est la Sous-Additivité des Probabilités qui implique le résultat si $\sum_{i\in I} \mathbf{Pr}(\overline{A_i}) < 1$. Dans le cas où $\sum_{i\in I} \mathbf{Pr}(\overline{A_i}) \geq 1$, on peut également conclure si les événements sont mutuellement indépendants (et de probabilité < 1). En effet, dans ce cas

$$\mathbf{Pr}\left(\bigcap_{i\in I} \overline{A_i}\right) = \prod_{i\in I} \mathbf{Pr}(\overline{A_i}) = \prod_{i\in I} (1 - \mathbf{Pr}(A_i)) > 0.$$

Cependant, en général, les événements considérés ne sont pas mutuellement indépendants. Mais tout n'est pas perdu. Le Lemme Local, établi par Erdös et Lovasz en 1975, montre que la probabilité $\mathbf{Pr}\left(\bigcap_{i\in I} \overline{A_i}\right) > 0$ est toujours strictement positive si les événements A_i sont de probabilité faible et, dans une certaine mesure, suffisamment indépendants les uns des autres.

Il existe différentes versions du Lemme Local. Nous ne présentons ici qu'une des plus simples, dans laquelle les probabilités des A_i ont une borne supérieure commune.

Théorème 9.6 (Lemme Local Symétrique). *Soient $A_i, i \in I$ des événements d'un espace de probabilité $(\Omega, \mathbf{Pr})$ tels que*

- *il existe p tel que pour tout $i \in I$, $\mathbf{Pr}(A_i) \leq p$,*
- *pour tout $i \in I$, l'événement A_i est mutuellement indépendant de tous les événements sauf au plus d.*

Si $4pd \leq 1$ alors $\mathbf{Pr}\left(\bigcap_{i\in I} \overline{A_i}\right) > 0$.

9.4.1 Coloration d'hypergraphes

Théorème 9.7. *Si chaque hyperarête d'un hypergraphe k-uniforme intersecte au plus 2^{k-3} autres hyperarêtes, alors l'hypergraphe est 2-colorable.*

Démonstration. Considérons une coloration aléatoire de l'hypergraphe pour laquelle chaque sommet est coloré bleu ou rouge de manière indépendante avec probabilité $\frac{1}{2}$. Pour toute hyperarête e de l'hypergraphe, soit A_e l'événement que e est monochromatique. Alors $\mathbf{Pr}(A_e) = 2^{1-k} = p$.
Assertion : *Chaque A_e est mutuellement indépendant de tous les événements A_f tels que $e \cap f = \emptyset$, donc de tous les événements sauf au plus $d = 2^{k-3}$.*

Cette assertion entraine directement la preuve du théorème par le Lemme Local car $4pd < 1$.

Reste donc l'assertion à prouver. Intuitivement cela semble clair mais prenons soin de la prouver car l'intuition débouche parfois sur de mauvaises surprises.

Supposons que les sommets sont ordonnés $v_1, v_2, \dots, v_n$ et que $e = \{v_1, \dots, v_t\}$. Considérons des hyperarêtes $f_1, \dots, f_r, g_1, \dots, g_s$ qui n'intersectent pas e. Soit Γ l'ensemble des 2-colorations pour lesquelles l'événement $B = A_{f_1} \cap \dots \cap A_{f_r} \cap \overline{A_{g_1}} \cap \dots \cap \overline{A_{g_s}}$ se produit. Pour toute 2-coloration c de $G - V(e)$, soit T_c l'ensemble des 2^t 2-colorations différentes de G qui étendent c. Il est facile de vérifier que pour toute c, Γ contient soit toutes les colorations T_c soit aucune. En d'autres termes, il existe l et des 2-colorations $c_1, \dots, c_l$ tels que Γ est l'union disjointe $T_{c_1} \cup \dots T_{c_l}$. Ainsi $\mathbf{Pr}(B) = \frac{2^t l}{2^n}$.

Maintenant dans chaque T_c, il y a exactement deux 2-colorations dans laquelle e est monochromatique, et donc $\mathbf{Pr}(A_e \cap B) = \frac{2l}{2^n}$. Ainsi $\mathbf{Pr}(A_e | B) = (\frac{2l}{2^n}) / \mathbf{Pr}(B) = 2^{-(t-1)} = \mathbf{Pr}(A_e)$ comme annoncé. □

L'assertion de la preuve précédente est un cas particulier d'un principe plus général qui permet de montrer l'indépendance mutuelle. Dans la très grande majorité des cas, l'indépendance mutuelle dérive de ce principe.
Principe de l'Indépendance Mutuelle *Supposons que $\mathcal{X} = X_1, \dots, X_m$ soit une suite de tirages aléatoires indépendants. Supposons de plus que $A_1, \dots, A_N$ est un ensemble d'événements, tels que chaque A_i est déterminé par $F_i \subset \mathcal{X}$. Si $F_i \cap \left(\bigcup_{j \in J} F_j\right) = \emptyset$ alors A_i est mutuellement indépendant de $\{A_j \ / \ j \in J\}$.*

La preuve de ce lemme est similaire à celle de l'assertion précédente et est laissée en Exercice 9.14.

9.4.2 Coloration par listes

Théorème 9.8. *Soit L une l-assignation de listes d'un graphe G. Si pour tout sommet v chaque couleur $i \in L(v)$ apparait dans au plus $\frac{l}{8}$ listes des voisins de v, alors G est L-colorable.*

Démonstration. Considérons une coloration aléatoire dans laquelle chaque sommet indépendamment se voit attribué une couleur de sa liste uniformément (avec probabilité $1/l$). Pour chaque arête $e = xy$ et chaque couleur $i \in L(x) \cap L(y)$, soit $A_{e,i}$ l'événement que x et y sont tous deux colorés i. Notons $\mathcal{E}$ l'ensemble de tels événements. Nous allons utiliser le Lemma Local afin de montrer qu'avec probabilité non-nulle aucun des événements de $\mathcal{E}$ ne se produit et donc la coloration est propre.

Considérons tout d'abord la probabilité d'$A_{e,i}$, clairement $\mathbf{Pr}(A_{e,i}) = \frac{1}{l^2} = p$. Comsidérons maintenant la dépendance des événements. Si $e = xy$, alors $A_{e,i}$ dépend uniquement des couleurs attribuées à x et y. Ainsi posant $\mathcal{E}_x = \{A_{f,j} \mid x \text{ est une extrémité de } f, j \in L(x)\}$ et $\mathcal{E}_y = \{A_{f,j} \mid y \text{ est une extrémité de } f, j \in L(y)\}$, par le Principe d'Indépendance Mutuelle, $A_{e,i}$ est indépendant de $\mathcal{E} \setminus (\mathcal{E}_x \cup \mathcal{E}_y)$. Maintenant comme $L(x)$ a l éléments et pour tout $i \in L(x)$, x a au plus $\frac{l}{8}$ voisins ayant la couleur i, nous avons $|\mathcal{E}_x| \leq \frac{l^2}{8}$. De même, $|\mathcal{E}_y| \leq \frac{l^2}{8}$. Ainsi $A_{e,i}$ est indépendant de tous les événements sauf au plus $d = \frac{l^2}{4}$. Comme $pd \leq \frac{1}{4}$, le Lemme Local donne le résultat. □

Comme c'est souvent le cas avec le Lemme Local, une fois que les mauvais événements sont choisis, la preuve est directe. Cependant, le choix des mauvais événements appropriés peut parfois s'avérer astucieux. Par exemple, dans l'Exercice 9.13, nous donnons deux définitions naturelles d'événements mauvais pour la preuve du Théorème 9.8 qui ne permettent pas d'obtenir une preuve.

9.5 Exercices

Exercice 9.1. Calculer l'espérance du nombre de sommets isolés (incidents à aucune arête) dans un graphe aléatoire de $\mathcal{G}_{n,p}$.

Exercice 9.2. Soit $G \in \mathcal{G}_{n,\frac{1}{2}}$. Pour tout sous-ensemble S de $V(G)$, soit A_S l'événement que S est un stable de G. Montrer que si S et T sont deux ensembles distincts de k sommets alors $A(S)$ et $A(T)$ sont indépendants si $|S \cap T| \leq 1$, et dépendants sinon.

Exercice 9.3. Une *k-coloration aléatoire* d'un graphe $G = (V, E)$ est un élément de l'espace de probabilité $(\Omega, \mathbf{Pr})$ où Ω est l'ensemble des k-colorations de G $(V_1, V_2, \ldots, V_k)$, toutes ces colorations étant équiprobables (donc se produisant avec probabilité k^{-n}). Pour une arête e de G, soit A_e l'événement que les deux extrémités de e reçoivent la même couleur. Montrer que :

a) quelles que soient deux arêtes e et f de G, les événements A_e et A_f sont indépendants.

b) si e, f et g sont trois arêtes d'un triangle de G, les événements A_e, A_f et A_g sont dépendants.

Exercice 9.4. Montrer la Sous-Additivité des Probabilités à partir du Principe du Premier Moment.

Exercice 9.5. Un *tournoi* T est un digraphe tel que pour toute paire de sommets $x \neq y$ exactement un des arcs xy et yx est dans $E(T)$. En d'autres termes, un tournoi est l'orientation d'un graphe complet. Le nom de tournoi est naturel, puisque l'on peut voir $V(T)$ comme un ensemble de joueurs qui s'affrontent tous deux à deux dans un match unique, et l'arc xy veut dire que x bat y. Un tournoi a la *propriété P_k* si pour tout ensemble de k sommets, il existe un sommet qui les dominent tous.

a) Montrer que si $\binom{n}{k}(1 - 2^{-k})^{n-k} < 1$ alors il existe un tournoi à n sommets ayant la propriété P_k.

b) En déduire qu'il existe un tournoi à $\lceil 4k^2 2^k \rceil$ sommets ayant la propriété P_k.

Exercice 9.6.

a) Montrer que l'espérance du nombre de chemins hamiltoniens d'un tournoi aléatoire est $n! \cdot 2^{-(n-1)}$.

b) Déduire que, pour tout $n \geq 1$, il y a un tournoi à n sommets qui a au moins $n! \cdot 2^{-(n-1)}$ chemins hamiltoniens.

Exercice 9.7. Une *variable booléenne* est une variable qui prend comme valeur soit $Vrai$ soit $Faux$. Pour toute variable booléenne x, il y a deux *litéraux* correspondant x et $\overline{x}$, où $\overline{x}$ veut dire "NON x" et a la valeur opposée à celle de x. Une formule booléenne en *Forme Normale Conjonctive* (FNC) consiste en une suite de *clauses* reliées par "$\wedge$" (ET), dans laquelle chaque clause est un ensemble de litéraux reliés par "$\vee$" (OR). La formule est *satisfiable* s'il existe une affectation de valeurs à ses variables telle que la formule entière vale $Vrai$, i.e. une affectation telle que chaque clause contienne au moins un litéral avec la valeur $Vrai$. Pour tout entier naturel k, une instance de k-SAT est une formule FNC dans laquelle chaque clause contient exactement k litéraux.

a) Montrer que toute instance de k-SAT avec strictement moins de 2^k clauses est satisfiable.

b) Construire une instance de k-SAT avec exactement 2^k clauses qui n'est pas satisfiable.

Exercice 9.8. Soit G un graphe avec n sommets et m arêtes. Son degré moyen est donc $Ad = \frac{2m}{d}$. Soit XS un sous-ensemble de sommets de G aléatoire où chaque sommet est dans S indépendamment avec probabilité p.

1) Montrer que l'espérance de la différence entre le nombre de sommets et le nombre d'arêtes de de $G[S]$ est $A(p) = np(1 - \frac{1}{2}dp)$.

2) Montrer que G a un stable de taille $A(p)$.

3) En déduire que $\alpha(G) \geq \frac{n}{2d}$.

Exercice 9.9. Montrer que pour n suffisamment grand, il existe un graphe à n sommets tel que $\chi(G) \geq \frac{n}{2}$ et $\omega(G) \leq n^{3/4}$. (Indication : Que peut-on dire du nombre chromatique du complémentaire d'un graphe sans triangle ?)

Exercice 9.10. Montrer par le Principe du Premier Moment que tout graphe ayant un couplage de taille m a un sous-graphe biparti avec au moins $\frac{1}{2}(|E(G)| + m)$ arêtes. (Indication : comment peut-on choisir une partition aléatoire telle que les arêtes du couplage aient leurs extrémités dans des parties opposées ?)

Exercice 9.11. Une coloration des arêtes d'un graphe G est dite *k-bornée* si aucune couleur n'est utilisée plus de k fois. Montrer que si $m \geq kn^2 2^n$ toute coloration k-bornée de l'hypercube Q_m contient une copie de Q_n dans laquelle toutes les arêtes sont de couleurs différentes.

Exercice 9.12. Montrer qu'il existe une fonction $g(d)$ tendant vers l'infini avec d telle que si G est un graphe biparti de degré minimum au moins d alors sa choisisabilité est au moins $g(d)$.

Exercice 9.13. Montrer ce qui ne marche pas si on essaie de prouver le Theorem 9.8 en colorant chaque sommet de manière indépendante avec une couleur aléatoire de sa liste de manière uniforme et en appliquant le Lemme Local avec les mauvais événements suivant :

1) Pour chaque sommet v, A_v est l'événement que v reçoive la même couleur qu'un de ses voisins.
2) Pour chaque arête e, A_e est l'événement que les deux extrémités de e reçoivent la même couleur.

Exercice 9.14. Montrer le Principe d'Indépendance Mutuelle.

Exercice 9.15. Soit $\mathcal{F}$ une instance de k-SAT telle que chaque variable est dans au plus $2^{k-2}/k$ clauses Montrer que $\mathcal{F}$ est satisfiable.

Exercice 9.16. Montrer que tout digraphe k-régulier D a une collection de $\left\lfloor \frac{k}{3\ln(k)} \right\rfloor$ cycles disjoints. On pourra montrer que pour une partition aléatoire en $\left\lfloor \frac{k}{3\ln(k)} \right\rfloor$ parties, avec probabilité non-nulle chaque sommet à un voisin externe dans sa partie.

Exercice 9.17. Soit D un digraphe de degré sortant minimum δ^+ et degré entrant maximum Δ^-. montrer que pour tout entier $k \leq \frac{\delta^+}{\ln(4\delta^+(\Delta^-+1))}$ D contient un cycle de longueur divisible par k.

Exercice 9.18. Soit $G = (V, E)$ un graphe et $V_1, \ldots V_k$ une partition de G. Montrer que si pour tout $1 \leq i \leq k$, $|V_i| \geq 8\Delta(G)$, alors il existe un stable qui contient exactement un sommet de chacun des V_i.