POFM: Combinatoire Probabiliste

Julien Portier

Février 2021

Contact : julien.portier@polytechnique.edu

Rappel : Soit X une variable aléatoire. On définit son espérance $\mathbb{E}[X]$ par :

$$\mathbb{E}[X] = \sum_x x\mathbb{P}(X = x)$$

Théorème : Soit $(X_i)_{1\leq i\leq n}$ une suite de variables aléatoires, pas forcément indépendantes. Alors :

$$\mathbb{E}\left[\sum_{i=1}^{n} X_i\right] = \sum_{i=1}^{n} \mathbb{E}[X_i]$$

Théorème : Soit X une variable aléatoire. Alors $\mathbb{P}(X \geq \mathbb{E}[X]) > 0$ et $\mathbb{P}(X \leq \mathbb{E}[X]) > 0$

Exemple 1 : Borne inférieure des nombres de Ramsey diagonaux

Rappel : On définit le nombre de Ramsey $R(k,l)$ comme étant le plus petit entier naturel n tel que pour tout $N \geq n$, tout coloriage des arêtes de K_N en rouge ou bleu contient un sous-graphe K_k dont toutes les arêtes sont rouges, ou un sous-graphe K_l dont toutes les arêtes sont bleues.

Nous allons établir une borne inférieure des nombres de Ramsey diagonaux $R(k,k)$. Soit N un entier naturel, on considère un coloriage aléatoire des arêtes de K_N : chaque arête est coloriée, indépendamment des autres arêtes, bleue avec une probabilité $\frac{1}{2}$, et rouge avec une probabilité $\frac{1}{2}$. Un k-clique est dit "mauvais" si toutes ses arêtes sont bleues, ou si toutes ses arêtes sont rouges. On numérote les k-cliques arbitrairement de 1 à $\binom{N}{k}$.
Soit X le nombre de mauvais k-cliques, et soit X_i pour $1 \leq i \leq \binom{N}{r}$ la variable aléatoire qui vaut 1 si le k-clique numéro i est mauvais, et 0 sinon. On a alors $X = \sum_i X_i$, d'où :

$$E[X] = E[\sum_i X_i] = \sum_i E[X_i] = \binom{N}{k} E[X_1] = \binom{N}{k} \frac{2}{2^{\binom{k}{2}}}$$

Par linéarité de l'espérance, puis parce que les X_i ont la même loi, et enfin par indépendance de la coloration des arêtes. Mais :

$$\begin{aligned}\binom{N}{k}\frac{2}{2^{\binom{k}{2}}} &\leq \frac{N^k}{k!}\frac{2}{2^{\binom{k}{2}}}\\ &\leq \frac{N^k}{k^k}\frac{k^k}{k!}\frac{2}{2^{\binom{k}{2}}}\end{aligned}$$

Comme $e^k = \sum_{i=0}^{\infty} \frac{k^i}{i!}$, on a $\frac{k^k}{k!} \leq e^k$, ce qui nous amène à $\mathbb{E}[X] \leq 2\left(\frac{Ne}{k2^{\frac{k-1}{2}}}\right)^k$

Ainsi, tant que $N < (\sqrt{2})^k \frac{\sqrt{2}k}{e2^{\frac{1}{k}}}$, on a $\mathbb{E}[X] < 1$, et donc $\mathbb{P}(X = 0) = \mathbb{P}(X \leq \mathbb{E}[X]) > 0$

Ce qui signifie qu'il existe une configuration sans mauvais k-cliques, et que donc $R(k,k) \geq (\sqrt{2})^k \frac{\sqrt{2}k}{e2^{\frac{1}{k}}}$

Exercice 1 : Soit n réels $x_1, x_2, \ldots, x_n$ non tous nuls, dont la somme est 0. Montrer qu'on peut les réarranger $a_1, a_2, \ldots, a_n$ tel que :

$$a_1a_2 + a_2a_3 + \cdots + a_{n-1}a_n < 0$$

Solution : Soit une permutation aléatoire uniforme $a_1, a_2, \ldots, a_n$. On a alors pour tout $1 \leq i \leq n-1$:

$$2\binom{n}{2}\mathbb{E}[a_ia_{i+1}] = \sum_{i\neq j} x_ix_j = \sum_{i,j} x_ix_j - \sum_i x_i^2 = \left(\sum_i x_i\right)^2 - \sum_i x_i^2 < 0$$

Ainsi, $\mathbb{E}[\sum_{i=1}^{n-1} a_ia_{i+1}] < 0$ donc il existe une permutation telle que cette somme est strictement négative.

Exercice 2 : Soit une compétition telle que chacun des 6 exercices a été résolu par au moins 120 des 200 participants. Montrer qu'on peut trouver 2 participants tel que chaque exercice a été résolu par au moins un des deux.

Solution : Soit X une paire aléatoire uniforme de participants, et soit S le nombre d'exercices qu'au moins un des deux a résolu. Soit, pour $1 \leq i \leq 6$, S_i la variable aléatoire valant 1 si au moins un des membres de X a résolu l'exercice i, et 0 sinon. On a alors :

$$\mathbb{E}[S] = \sum_{i=1}^{6} \mathbb{E}[S_i] \geq 6\left(1 - \frac{\binom{80}{2}}{\binom{200}{2}}\right) = 6\left(1 - \frac{80 \times 79}{200 \times 199}\right) \geq 6\left(1 - \frac{80^2}{200^2}\right) = 6\left(1 - \frac{4}{25}\right) = \frac{126}{25} > 5$$

Il existe donc une paire telle que $S > 5$, c'est-à-dire, $S = 6$, ce qui résout l'exercice.

Exercice 3 : (Test Sélection Iran 2008) On considère 799 équipes de football dans un tournoi tel qu'il y a un match entre chaque paire d'équipes, qui ne peut pas se finir en égalité. Montrer qu'il existe deux ensembles A et B chacun de cardinal 7 tels que chaque équipe de A a battu chaque équipe de B.

Solution : On considère le graphe dont les sommets sont les 799 équipes, et on trace une arête de u à v si et seulement si l'équipe u a battu l'équipe v. Pour tout ensemble A, on considère l'ensemble X_A des équipes qui ont été battues par toutes les équipes composant A. L'exercice est alors équivalent à trouver A de cardinal 7 tel que $|X_A| \geq 7$. Soit A la variable aléatoire associée à un sous-ensemble aléatoire uniforme de cardinal 7. Soit X la variable aléatoire associée à l'ensemble des équipes battues par toutes les équipes de A. On a :

$$\begin{aligned} E[|X|] &= E\left[\sum_{i=1}^{799} 1_{i\in X}\right] \\ &= \sum_{i=1}^{799} E\left[1_{i\in X}\right] \\ &= \sum_{i=1}^{799} \mathbb{P}(i \in X) \\ &= \sum_{i=1}^{799} \frac{\binom{{d_i}^-}{7}}{\binom{799}{7}} \end{aligned}$$

Où ${d_i}^-$ est le nombre de sommets u tel qu'il y a une arête orientée de u à i. Or, d'après l'inégalité de Jensen,

$$\sum_{i=1}^{799} \frac{\binom{{d_v}^-}{7}}{\binom{799}{7}} \geq 799 \frac{\binom{399}{7}}{\binom{799}{7}} = \frac{\prod_{i=393}^{399} i}{\prod_{j=793}^{798} j} > 6$$

Ainsi, $\mathbb{E}[X] > 6$ ce qui conclut la preuve.

Des raisonnements un peu plus complexes sont possibles, comme nous allons le voir avec ce théorème :

Exemple 2 : (Théorème de Szemeredi-Trotter) Soit m droites et n points. Le nombre de couples (d_i, p_j) tel que le point p_j est situé sur la droite d_i est appelé I le nombre d'incidences. On a alors pour une constante $C > 0$:

$$I \leq C \max\{n, m, (nm)^{\frac{2}{3}}\}$$

Nous allons tout d'abord prouver ce lemme :

Lemme des intersections : Soit un graphe G avec n sommets et m arêtes. Soit t le nombre minimum d'intersections d'arêtes dans toutes les représentations possibles de G dans le plan. Si $m \geq 6n$, alors $t \geq \dfrac{m^3}{72n^2}$.

Preuve du lemme : Tout d'abord, remarquons qu'on peut enlever t arêtes à G pour obtenir un graphe planaire. Or, pour tout graphe planaire, on a la formule d'Euler : $n' - m' + f' \geq 2$. Or, $3f' \leq 2m'$, donc en substituant cette inégalité dans la première inégalité, il vient : $m' \leq 3n' - 6$. Comme nous l'avons dit plus haut, on peut enlever t arêtes à G pour obtenir un graphe planaire, ainsi $m - t \leq 3n - 6$, donc $t \geq m - 3n + 6$ pour tout graphe G.

On considère maintenant le graphe G' obtenu comme le sous-graphe de G tel que chaque sommet appartient à G' avec probabilité p (d'une valeur que l'on choisira plus tard), de façon indépendante des autres sommets. On a alors $\mathbb{E}[n'] = pn$, $\mathbb{E}[m'] = p^2m$ et $\mathbb{E}[t'] = p^4t$, ce qui donne, en injectant dans l'inégalité plus haut : $p^4t \geq p^2m - 3pn + 6$

Pour $p = \frac{6n}{m}$, il vient :

$$t \geq \frac{m^3}{72n^2}$$

Revenons au théorème de Szemeredi-Trotter, et considérons le graphe K tel que les sommets sont les n points, et l'on trace une arête entre 2 sommets si et seulement si il existe une droite sur laquelle ces 2 sommets sont consécutifs sur cette droite. Le nombre d'arêtes est alors $I - m$. Tant que $I - m \geq 6n$, on a :

$$t \geq \frac{(I-m)^3}{72n^2}$$

Sauf que $t \leq \binom{m}{2} < \dfrac{m^2}{2}$, d'où :

$$I \leq m + 36^{\frac{1}{3}}(nm)^{\frac{2}{3}}$$

Avec rajoutant le cas $I - m \leq 6n$, on a bien, pour $C = 7$:

$$I \leq C \max\{n, m, (nm)^{\frac{2}{3}}\}$$

Exercice 4 : (Théorème de Sperner) Soit X un sous-ensemble de $\mathbb{P}([n])$ tel que pour tous $A, B \in X$ avec $A \neq B$, A n'est pas inclus dans B. Montrer que $|X| \leq \binom{n}{\frac{n}{2}}$ avec égalité si et seulement n est pair et $X = [n]^{\frac{n}{2}}$ ou $n = 2m+1$ est impair et $X = [n]^m$ ou $X = [n]^{m+1}$ ($[n]^r$ étant l'ensemble des sous-ensembles de $[n]$ de cardinal r)

Solution : Soit C la chaîne $C = \{\{\}, \{x_1\}, \{x_1, x_2\}, \ldots, \{x_1, \ldots, x_n\}\}$ pour une permutation aléatoire uniforme $x_1, \ldots, x_n$ de $[n]$. Soit Z la variable aléatoire associée au nombre d'éléments de X appartenant à C. Par définition de X, $|Z| \leq 1$, donc $\mathbb{E}[Z] \leq 1$. D'autre part :

$$\begin{aligned} E[Z] &= \sum_{A \in X} E[1_{A \in C}] \\ &= \sum_{A \in X} P(A \in C) \end{aligned}$$

Mais $P(A \in C) = \dfrac{1}{\binom{n}{|A|}}$ car $A \in C$ correspond au cas où la permutation qui construit C, admet comme premiers éléments les $|A|$ éléments de A. Cela donne, pour X_k le sous-ensemble de X constitué des éléments de cardinal k :

$$1 \geq E[Z] = \sum_{k=0}^{n} \frac{|X_k|}{\binom{n}{k}} \geq \frac{|X|}{\binom{n}{\frac{n}{2}}}$$

Ce qui conclut pour l'inégalité. Considérons maintenant le cas d'égalité : si n est pair, on a évidemment $X = [n]^{\frac{n}{2}}$. Si $n = 2m+1$ est impair, on a $X \subset [n]^m \cup [n]^{m+1}$, donc $X = X_m \cup X_{m+1}$. Quitte à remplacer chaque élément de X par son complémentaire, on peut supposer que $|X_m| \geq |X_{m+1}|$. Soit X'_m le sous-ensemble de $[n]^{m+1}$ constitué des éléments qui contiennent au moins un élément de X_m. Par définition de X, on a $X'_m \cap X_{m+1} = \emptyset$, donc $|X'_m| \leq \binom{n}{m+1} - |X_{m+1}| = |X_m|$. Considérons le graph bipartite entre X_m et X'_m tel qu'il y a une arête entre x et y si et seulement si x est inclu dans y, ou l'inverse. Comptons de deux manières différentes le nombre d'arêtes de ce graphe : d'une part, il y a $(m+1)|X_m|$ arêtes puisque chaque élément de X_m admet exactement $m+1$ éléments de $[n]^{m+1}$ qui le contiennent. D'autre part, chaque élément de X'_m contient au plus $m+1$ éléments de X_m, donc il y a au plus $(m+1)|X'_m|$ arêtes. Ainsi, $|X'_m| \geq |X_m|$ ce qui donne au total $|X'_m| = |X_m|$. De plus, comme on a le cas d'égalité dans le dernier raisonnement, cela signifie que si l'on part de n'importe quel $A \in X_m$, qu'on ajoute un élément à A, puis qu'on en retire un, alors on retombe dans X_m. Cela signifie que $X_m = [n]^m$ ce qui conclut l'exercice.

Exercice 5 : Soit G un graphe à n sommets et de degré moyen d. Montrer que l'on peut trouver un ensemble S de sommets indépendant de cardinal au moins $\frac{n}{d+1}$ (un ensemble de sommets est dit indépendant si chaque paire de sommets n'admet pas d'arêtes)

Solution : On a envie d'appliquer un algorithme glouton ici : on considère une permutation des sommets, et pour chaque sommet, on l'ajoute dans S si et seulement si aucun de ses voisins n'y est. Le problème de cet algorithme est qu'on peut se retrouver avec un mauvais choix de sommet dès le début qui va nous handicaper pour la suite... Mais on peut résoudre ce problème en considérant une permutation aléatoire uniforme sur laquelle on va effectuer notre algorithme glouton ! Ecrivons cela rigoureusement :
On considère une permutation aléatoire uniforme des sommets $x_1, x_2, \ldots, x_n$, et on considère l'algorithme glouton suivant : on parcourt les sommets $x_1, x_2, \ldots, x_n$ dans cet ordre, et au sommet x_i, si S ne contient aucun des voisins de x_i, on ajoute x_i à S, sinon on ne fait rien, puis on recommence avec le sommet x_{i+1}. On a alors :

$$\mathbb{E}[|S|] = \mathbb{E}[\sum_{i=1}^{n} 1_{i \in S}] = \sum_{i=1}^{n} \mathbb{P}(i \in S)$$

Or, la probabilité que le sommet i soit dans S est au moins $\frac{1}{d_i+1}$, puisque i est ajouté à S s'il est situé avant tous ses voisins dans la permutation aléatoire uniforme. Ainsi :

$$\mathbb{E}[|S|] \geq \sum_{i=1}^{n} \frac{1}{d_i+1} \geq \frac{n}{d+1}$$

D'après l'inégalité de Jensen, ce qui conclut.

Example 3 : (IMO 2014/6) On dit qu'un ensemble de droites est en position générale si deux droites de cet ensemble ne sont jamais parallèles et trois droites de cet ensemble ne s'intersectent jamais en un point. Un ensemble de droites en position générale partagent le plan en régions, celles qui ont une aire finie sont appelées les "régions finies". Montrer qu'il existe $c > 0$ tel que pour tout n suffisamment large, pour tout ensemble de n droites en position générale, on peut colorier au moins $c\sqrt{n}$ droites en bleu tel qu'aucune des régions finies n'a son périmètre totalement colorié en bleu.

Solution : Comme les droites sont en position générale, il y a exactement $\binom{n}{2}$ intersections. Or chaque intersection ne peut être le sommet que de deux triangles au maximum (faire un dessin).

Ainsi, il y a au plus $\frac{2}{3}\binom{n}{2} < \frac{1}{3}n^2$ régions finies qui sont des triangles. On colorie alors chaque droite bleue avec probabilité p, de façon indépendante des autres droites. Soit d le nombre de droites coloriées en bleu, t le nombre de triangles coloriés entièrement en bleu, et p le nombre de polygones qui ne sont pas des triangles et entièrement coloriés en bleu. On a $\mathbb{E}[d] = pn$, $\mathbb{E}[t] < \frac{p^3}{3}n^2$ et $\mathbb{E}[p] < \frac{p^4}{2}n^2$ (car il y a au plus $\frac{n^2}{2}$ régions finies).
Pour chaque région finie qui est entièrement coloriée en bleu, on décolore une droite qui forme le périmètre, ainsi une obtient au moins r droites coloriées en bleu, avec $r \geq d - t - p$, donc :

$\mathbb{E}[r] \geq pn - \frac{p^3}{3}n^2 - \frac{p^4}{2}n^2 = pn\left(1 - \dfrac{p^2 n}{3} - \dfrac{p^3 n}{2}\right)$

Pour $p = \frac{1}{\sqrt{n}}$, on obtient $\mathbb{E}[r] \geq \dfrac{2}{3}\sqrt{n} - \dfrac{1}{2}$ ce qui prouve l'énoncé pour tous $c < \frac{2}{3}$

Théorème : Soit $A_1, \dots, A_n$ des événements.

$$\mathbb{P}(A_1 \cup \dots \cup A_n) \leq \sum_{i=1}^{n} \mathbb{P}(A_i)$$

Preuve : $n = 1$ est évident, pour $n = 2$ on a $\mathbb{P}(A \cup B) = \mathbb{P}(A) + \mathbb{P}(B) - \mathbb{P}(A \cap B) \leq \mathbb{P}(A) + \mathbb{P}(B)$, et le cas général se traite facilement par récurrence.

Théorème : Soit $A_1, \dots, A_n$ des événements. Si $\sum_{i=1}^n \mathbb{P}(\bar{A}_i) < 1$, alors $\mathbb{P}(A_1 \cap \dots \cap A_n) > 0$

Preuve : Il s'agit seulement de remarquer que $\mathbb{P}(A_1 \cap \dots \cap A_n) = 1 - \mathbb{P}(\bar{A}_1 \cup \dots \cup \bar{A}_n) \geq 1 - \sum_{i=1}^n \mathbb{P}(\bar{A}_i)$

Théorème (Inégalité de Markov) Soit X une variable aléatoire positive, alors $\mathbb{P}(X \geq k) \leq \dfrac{\mathbb{E}[X]}{k}$

Preuve : $\mathbb{E}[X] \geq \mathbb{E}[X 1_{X \geq k}] \geq k\mathbb{E}[1_{X \geq k}] = k\mathbb{P}(X \geq k)$

Exercice 6 : Soit n un graphe tel que chaque sommet a une degré d'au moins $d = \sqrt{n}$. Montrer que pour tout n suffisamment grand, il existe un ensemble A d'au plus $2\sqrt{n}\log(n)$ sommets tel que tout sommet n'appartenant pas à A est voisin d'au moins un élément de A.

Solution : Pour chaque sommet v, on tire une variable aléatoire indépendante des autres sommets telle que v appartient à A avec probabilité p.
Soit A_i l'événement $\{$Le sommet i a un de ses voisins dans A$\} \cup \{i \in A\}$ et soit B l'événement $\{|A| \leq 2\sqrt{n}\log(n)\}$. Il nous suffit de montrer que $\mathbb{P}(\bar{B}) + \sum_{i=1}^n \mathbb{P}(\bar{A}_i) < 1$, et le théorème ci-dessus conclura sur l'existence d'un ensemble A de sommets ayant les propriétés demandées. D'après l'inégalité de Markov :

$\mathbb{P}(\bar{B}) = \mathbb{P}(|A| > 2\sqrt{n}\log(n)) \leq \dfrac{\mathbb{E}[|A|]}{2\sqrt{n}\log(n)} = \dfrac{pn}{2\sqrt{n}\log(n)} = \dfrac{p\sqrt{n}}{2\log(n)}$

$\mathbb{P}(\bar{A}_i) \leq (1-p)^{\sqrt{n}} \leq e^{-p\sqrt{n}}$ puisque $1 + x \leq e^x$

Ainsi, on a $\mathbb{P}(\bar{B}) + \sum_{i=1}^n \mathbb{P}(\bar{A}_i) \leq \dfrac{p\sqrt{n}}{2\log(n)} + ne^{-p\sqrt{n}} = \dfrac{p\sqrt{n}}{2\log(n)} + e^{\log(n) - p\sqrt{n}}$

Pour $p = \frac{3}{2}\frac{\log(n)}{\sqrt{n}}$ (qui est bien ≤ 1 pour n suffisamment grand), on a :

$\dfrac{p\sqrt{n}}{2\log(n)} + e^{\log(n) - p\sqrt{n}} = \frac{3}{4} + e^{-\frac{1}{2}log(n)} = \frac{3}{4} + \frac{1}{\sqrt{n}} < 1$ pour n suffisamment grand, ce qui conclut cet exercice.

Lemme Local de Lovasz : Soit $A_1, \dots, A_n$ tels que la probabilité de chacun de ces événements est au plus p, et chaque événement est indépendant de tous les autres, sauf au plus d événements. Si $epd \leq 1$, alors $\mathbb{P}(\bar{A}_1 \cap \dots \cap \bar{A}_n) > 0$

Exercice 7 : Soit $11n$ points sur un cercle que l'on colorie avec n couleurs de telle sorte que chaque couleur soit utilisée pour exactement 11 points. Montrer que l'on peut choisir un point de chaque couleur sans avoir deux points adjacents.

Solution : On numérote les points $x_1, \dots, x_{11n}$. Pour chaque couleur, on choisit uniformément au hasard un des 11 points de cette couleur, et on définit pour tous $1 \leq i \leq 11n$, $A_i = \{x_i \text{ et } x_{i+1} \text{ ont été choisis}\}$ ($x_{11n+1} = x_1$). Chaque A_i est réalisé avec probabilité au plus $p = \frac{1}{11^2}$ et est indépendant de tous les autres A_j sauf au plus $d = 42$: au plus 20 autres points ont la même couleur que x_i ou x_{i+1}, donc 40 événements au plus comportent un point de la même couleur que x_i ou x_{i+1}, donc ne sont pas forcément indépendants de A_i, plus A_{i-1} et A_{i+1}, ce qui donne bien 42.
Or, $epd \leq \frac{2.8 \times 42}{121} = \frac{117.6}{121} \leq 1$, et le lemme local de Lovasz nous permet de conclure.

Exercice 8 (Théorème d'Erdos-Ko-Rado) Soit $r \leq 2n$, et soit $A \subset [n]^r$ tel que toute paire d'éléments de A a une intersection non vide. Montrer que $|A| \leq \binom{n-1}{r-1}$

Solution : On considère σ une permutation aléatoire uniforme de $[n]$ et on écrit sur un cercle les éléments de $[n]$ dans l'ordre de σ. On dit qu'un ensemble C est contenu dans la permutation σ si l'on peut lire tous les éléments de C à la suite sur le cercle de la permutation σ. Ainsi, il est direct que pour tout élément C de A, $\mathbb{P}(C \text{ contenu dans } \sigma) = \frac{n}{\binom{n}{r}}$, d'où pour X la variable aléatoire associée au nombre d'éléments de A contenu dans σ : $\mathbb{E}[X] = \frac{n|A|}{\binom{n}{r}}$
Bornons maintenant $\mathbb{E}[X]$ pour borner $|A|$: Soit C un élément de A contenu dans σ, disons $x_1, \dots, x_r$ alors on sait que chaque autre élément C' de A contenu dans σ doit intersecter C, et donc contenir un des x_i pour $1 \leq i \leq r$. Pour chaque x_i avec $1 \leq i \leq r$, on lui associe $\rightarrow$ si $\{x_i, x_{i+1}, \dots, x_{i+r-1}\}$ appartient à A, et $\leftarrow$ si $\{x_{i-r}, x_{i-r+1}, \dots, x_{i-1}\}$ appartient à A. Notons qu'on ne peut pas avoir à la fois $\rightarrow$ et $\leftarrow$ à un point x_i fixé, sinon cela contredirait la propriété que deux éléments de A ont toujours une intersection non vide. De plus, il est facile de voir que le nombre de flèches est le nombre d'éléments de A contenu dans σ. Il y a au plus r flèches pour chaque configuration, donc $\mathbb{E}[X] \leq r$, ce qui donne en reprenant l'égalité au-dessus : $|A| \leq \frac{n\binom{n}{r}}{r} = \binom{n-1}{r-1}$

Remarque : On peut également montrer que le cas d'égalité correspond au cas où tous les éléments de A s'intersectent en un point x.

Exercice 9 : Soit G un graphe tel que chaque sommet a un degré compris entre 50 et 100. Montrer que l'on peut colorier les sommets de ce graphe avec au plus 114 couleurs tel que chaque sommet a au moins 20 voisins coloriés de couleurs distinctes.

Solution : On va colorier les sommets du graphe aléatoirement de façon indépendantes les uns des autres avec une des C couleurs.
L'événement $A_i = \{$ Les voisins du sommet i sont coloriés avec au plus 19 couleurs $\}$. Si les sommets i et j ne sont ni voisins ni ne partagent un sommet voisin commun, alors A_i et A_j sont indépendants. Ainsi, comme i a au plus 100 voisins, A_i est indépendant de tous les A_j sauf au plus $100 + 99 \times 100 = 100^2 = d$ événements. Maintenant la probabilité de A_i est au plus $\binom{C}{19}(\frac{19}{C})^{50} < (\frac{eC}{19})^{19}(\frac{19}{C})^{50} = e^{19}(\frac{19}{C})^{31} = p$

Le lemme local de Lovasz conclut alors, à condition que $epd \leq 1$, c'est-à-dire $e^{20}(\frac{19}{C})^{31}100^2 \leq 1$ qui est équivalent à $C \geq 19e^{\frac{20}{31}}100^{2/31}$. Or $19e^{\frac{20}{31}}100^{2/31} \leq 19e10^{1/4} \leq 19 \times 3 \times 2 = 114$, donc $C = 114$ convient.

Remarque : Attention à ne pas faire n'importe quoi : il est très facile de commettre des confusions avec des raisonnements probabilistes, notamment sur l'indépendance des événements. Par exemple, on peut tenter d'introduire des méthodes probabilistes en théorie des nombres, en considérant, pour N une variable aléatoire, l'événement $\{p$ divise $N\}$ pour p premier. Malheureusement $\{p_1$ divise $N\}$ et $\{p_2$ divise $N\}$ ne sont en général pas indépendants : par exemple pour $p_1 \geq \sqrt{N}$ et $p_2 \geq \sqrt{N}$, si p_1 divise N, alors il n'y a aucune chance que p_2 divise N... Donc il faut à chaque fois se demander si ce que l'on fait est bien logique et rigoureux !

Références :

1. Expected Uses of Probability, Evan Chen, 2014

2. Probabilistic Methods in Combinatorics, Po-Shen Loh, 2009