DOMAINE : Combinatoire
NIVEAU : Avancé
CONTENU : Cours et exercices

AUTEUR : Thomas BUDZINSKI
STAGE : Montpellier 2014

Méthode probabiliste en combinatoire

1 Résumé du cours

Définition 1. Un *espace de probabilité* fini est un ensemble Ω muni d'une application $\mathbb{P}$ de l'ensemble des parties de Ω dans $\mathbb{R}^+$ vérifiant :
- Si $A \cap B = \varnothing$, alors $\mathbb{P}(A \cup B) = \mathbb{P}(A) + \mathbb{P}(B)$.
- $\mathbb{P}(\Omega) = 1$.

Les parties de Ω sont alors appelées *événements*.

Proposition 2.
- $\mathbb{P}(\varnothing) = 0$.
- $\mathbb{P}(A^c) = 1 - \mathbb{P}(A)$.
- Si $A \subset B$, alors $\mathbb{P}(A) \leqslant \mathbb{P}(B)$.
- $\mathbb{P}(A_1 \cup \cdots \cup A_k) \leqslant \mathbb{P}(A_1) + ... + \mathbb{P}(A_k)$.

Définition 3.
- Deux événements A et B sont dits *indépendants* si $\mathbb{P}(A \cap B) = \mathbb{P}(A)\mathbb{P}(B)$.
- k événements $A_1, ..., A_k$ sont dits *indépendants* si pour tout $I \subset [\![1, k]\!]$, on a :

$$\mathbb{P}\Big(\bigcap_{i \in I} A_i\Big) = \prod_{i \in I} \mathbb{P}(A_i)$$

Définition 4. Une *variable aléatoire* est une application X de Ω dans $\mathbb{R}^+$. On notera $\mathbb{P}(X = a)$ pour $\mathbb{P}\big(\{\omega \in \Omega | X(\omega) = a\}\big)$, et de même pour $X \leqslant a$, $X \in E$ etc...

Définition 5. Soit X une variable aléatoire. L'*espérance* de X est le nombre :

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

Remarque 6. Si $\mathbb{E}[X] \geqslant a$, alors il existe ω tel que $X(\omega) \geqslant a$.

Proposition 7. Soient X et Y deux variables aléatoires. Alors $\mathbb{E}[X+Y] = \mathbb{E}[X] + \mathbb{E}[Y]$ et, si $a \in \mathbb{R}$, $\mathbb{E}[aX] = a\mathbb{E}[X]$.

Proposition 8. Soit E un ensemble fini et A un sous-ensemble aléatoire de E. Alors :

$$\mathbb{E}[|A|] = \sum_{x \in E} \mathbb{P}(x \in A)$$

Théorème 9. (Inégalité de Markov) Soient X une variable aléatoire positive et $a > 0$. Alors :

$$\mathbb{P}(X \geqslant a) \leqslant \frac{\mathbb{E}[X]}{a}$$

2 Exercices

Exercice 1 Montrer qu'il est possible de colorier les entiers de 1 à 2014 en quatre couleurs de manière à n'avoir aucune progression arithmétique monochrome de longueur 11.

<u>*Solution de l'exercice 1*</u> On colorie aléatoirement chaque entier de 1 à 2014 : la proba qu'une progression arithmétique fixée de longueur 11 soit monochrome est $\frac{1}{4^{10}} < \frac{1}{1000000}$, et il y a au plus 203 choix pour la raison r (r est compris entre 1 et 201, car le dernier terme est au moins $1 + 10r$ donc $1 + 10r \leqslant 2014$) et 2014 choix pour le premier terme, donc la proba qu'il existe une suite monochrome est inférieure à $\frac{203 \times 2014}{1000000} < \frac{300 \times 3000}{1000000} < 1$, donc il existe un coloriage sans progression monochrome.

Exercice 2 Soient n et k dans $\mathbb{N}^*$. Au cours d'un tournoi entre n joueurs, chaque paire s'affronte exactement une fois. On dit que le tournoi est k-*indécis* si pour tout ensemble A de k joueurs, il existe un joueur qui a battu tous ceux de A.
Montrer que pour tout k, il existe un tournoi k-indécis à plus de k joueurs.

<u>*Solution de l'exercice 2*</u> On considère un tournoi aléatoire : pour chaque match, le résultat est tiré à pile ou face ! Pour un ensemble de k joueurs fixés, qu'un adversaire fixé les ait tous battus est $\frac{1}{2^k}$. La proba qu'aucun des $n-k$ joueurs restants ne les ait tous battus est donc $\left(1 - \frac{1}{2^k}\right)^{n-k}$. La proba qu'il existe k joueurs que personne n'a battu se majore donc par $\binom{n}{k}\left(1 - \frac{1}{2^k}\right)^{n-k}$.

Or, quand $n \longrightarrow +\infty$, cette quantité tend vers 0 car le dénominateur croît exponentiellement et le numérateur polynomialement. Si n est choisi assez grand, la proba que le tournoi ne soit pas k-indécis est donc strictement plus petite que 1, donc en particulier il existe un tournoi k-indécis.

Exercice 3 (Sperner) Soit $\mathcal{A}$ un ensemble de sous-ensembles de $[\![1, 2n]\!]$: on suppose que pour tous B et C distincts dans $\mathcal{A}$, B n'est pas inclus dans C. (On dit alors que $\mathcal{A}$ est une antichaîne.)

– Montrer l'inégalité suivante :

$$\sum_{B \in \mathcal{A}} \frac{1}{\binom{n}{|B|}} \leqslant 1$$

– En déduire $|\mathcal{A}| \leqslant \binom{n}{\lfloor \frac{n}{2} \rfloor}$.

<u>*Solution de l'exercice 3*</u>

– On prend une chaîne aléatoire uniforme (A_i) : on ajoute les éléments un par un tels qu'à chaque étape, tous les éléments restants ont la même probabilité d'être ajoutés, et on note A_i l'ensemble des éléments ajoutés aux i premières étapes. La proba que la chaîne contienne un B est égale à la probabilité que B soit le $|B|$-ième élément de la chaîne. Or, il y a $\binom{n}{|B|}$ sous-ensembles de taille $|B|$ et chacun a la même proba d'être le $|B|$-ième élément de la chaîne, donc $\mathbb{P}(B \in \mathcal{A}) = \frac{1}{\binom{n}{|B|}}$. On conclut car la chaîne contient au plus un élément de l'antichaîne.

– Chaque dénominateur se majore par $\binom{n}{\lfloor \frac{n}{2} \rfloor}$.

Exercice 4 Soient $p, q \geqslant 0$ avec $p + q = 1$ et $m, n \in \mathbb{N}^*$. Montrer l'inégalité :

$$(1 - p^m)^n + (1 - q^n)^m \geqslant 1$$

<u>*Solution de l'exercice 4*</u> On colorie chaque case d'un rectangle m sur n en blanc ou noir de manière indépendante aevc probas p et $1 - p$. $1 - p^m$ est la proba qu'une ligne fixée soit entièrement blanche, donc le premier terme est la proba qu'il n'y ait une pas de ligne blanche. De même, le deuxième est la proba qu'il n'y ait pas de colonne noire. Or, au moins un de ces deux évènements se produit car si il y a une ligne blanche et une colonne noire, alors elles s'intersectent et on obtient une contradiction.

Exercice 5 Soit G un graphe et, pour tout x sommet de G, $d(x)$ son degré, i.e le nombre de voisins de ce sommet dans le graphe. On dit qu'un ensemble A de sommets de G est *indépendant* si il ne contient pas deux sommets voisins. Montrer qu'il existe un ensemble indépendant de taille supérieure ou égale à :

$$\sum_x \frac{1}{1 + d(x)}$$

Solution de l'exercice 5 On colle sur chaque sommet une étiquette dans $[\![1,n]\!]$ de manière à ce que chaque entier de 1 à n apparaisse une unique fois, et ce de manière uniforme parmi les $n!$ manières possibles. On prend pour A l'ensemble des points dont l'étiquette est plus grande que celles de tous les voisins. Cet ensemble est forcément indépendant car parmi deux voisins, l'un a forcément une étiquette plus petite que l'autre et donc n'est pas dans A.

De plus, soit x un sommet du graphe. On note $V(x)$ l'ensemble formé par x et ses voisins : $\mathbb{P}(x \in A)$ est la probabilité que, dans $V(x)$ voisins, x soit le sommet avec la plus grande étiquette. Or, chacun des $d(x)+1$ sommets de $V(x)$ a la même probabilité d'être celui avec la plus grande étiquette donc $\mathbb{P}(x \in A) = \frac{1}{1+d(x)}$, d'où :

$$\mathbb{E}[|A|] = \sum_x \frac{1}{1+d(x)}$$

En particulier, il existe au moins une configuration où $|A|$ est plus grande que cette grandeur.

Exercice 6 (USAMO 2012) Soient $x_1, x_2, ..., x_n$ tels que $\sum_i x_i = 0$ et $\sum_i x_i^2 = 1$. Pour tout $A \subset [\![1,n]\!]$ on pose $S_A = \sum_{i \in A} x_i$. Soit aussi $\lambda > 0$.

Montrer que le nombre de A tels que $S_A \geqslant \lambda$ est inférieur ou égal à $\frac{2^{n-3}}{\lambda^2}$.

Solution de l'exercice 6 Tout d'abord, si A' est le complémentaire de A, comme $\sum_i x_i = 0$, la somme des éléments de A' est l'opposé de celle de A, donc il y a autant de sous-ensembles avec $S_A \geqslant \lambda$ que de sous-ensembles avec $S_A \leqslant -\lambda$, donc il suffit de montrer que le nombre de A tels que $S_A^2 \geqslant \lambda^2$ est inférieur ou égal à $\frac{2^{n-2}}{\lambda^2}$.

On choisit donc un sous-ensemble aléatoire uniforme A de $[\![1,n]\!]$, ce qui revient à choisir indépendamment avec proba $\frac{1}{2}$ pour chaque i si i est dans A. L'énoncé revient à montrer que $\mathbb{P}(S_A^2 \geqslant \lambda^2) \leqslant \frac{1}{4\lambda^2}$. D'après l'inégalité de Markov, cette proba se majore par $\frac{\mathbb{E}[S_A^2]}{\lambda^2}$, donc il suffit de montrer $\mathbb{E}[S_A^2] = \frac{1}{4}$.

On fait donc le calcul :

$$\begin{aligned}
\mathbb{E}[S_A^2] &= \mathbb{E}\Big[\Big(\sum_i x_i \mathbb{1}_{i\in A}\Big)^2\Big] \\
&= \mathbb{E}\Big[\sum_i x_i^2 \mathbb{1}_{i\in A} + \sum_{i\neq j} x_i x_j \mathbb{1}_{i\in A}\mathbb{1}_{j\in A}\Big] \\
&= \sum_i x_i^2 \mathbb{E}[\mathbb{1}_{i\in A}] + \sum_{i\neq j} x_i x_j \mathbb{E}[\mathbb{1}_{i\in A}\mathbb{1}_{j\in A}] \\
&= \sum_i x_i^2 \mathbb{P}(i\in A) + \sum_{i\neq j} x_i x_j \mathbb{P}(i \in A \text{ et } j \in A) \\
&= \frac{1}{2}\sum_i x_i^2 + \frac{1}{4}\sum_{i\neq j} x_i x_j \\
&= \frac{1}{4}
\end{aligned}$$

car $\sum_i x_i^2 = 1$ et $\sum_{i\neq j} x_i x_j = \big(\sum_i x_i\big)^2 - \sum_i x_i^2 = -1$, d'où le résultat.

Exercice 7 (IMO 1989) On dit qu'une permutation σ de $[\![1, 2n]\!]$ est *gentille* si il existe i tel que $|\sigma(i) - \sigma(i+1)| = n$. Elle est *méchante* sinon.

Montrer qu'il y a plus de permutations gentilles que de permutations méchantes.

Solution de l'exercice 7 Soit A l'ensemble des i vérifiant la condition demandée : Pour tout $i \leqslant 2n-1$, $\mathbb{P}(i \in A) = \frac{1}{2n-1}$ car pour chaque a il existe une unique valeur b différente de a telle que $|a-b| = n$. On a donc $\mathbb{E}[|A|] = 1$. On voudrait maintenant montrer que $|A|$ ne s'éloigne "pas trop" de sa valeur moyenne. Pour cela, on calcule sa variance :

$\mathbb{P}(i, j \in A)$ vaut $\frac{1}{2n-1}$ si $i = j$, 0 si $|i-j| = 1$ et $\frac{2n-1}{2n-3}$ sinon. Par un calcul similaire à l'exercice précédent, on obtient $\mathbb{E}[|A|^2] = 1 + \frac{2n-3}{2n-4} < 2$, soit $Var(|A|) < 1$, ce qui permet de conclure en bidouillant un peu...