1 Lemmes de Borel-Cantelli et applications à l'étude des nombres premiers

Premier lemme

Lemme 1. *Soit* (A_n) *une suite d'évènements tels que* $\sum_{n\in\mathbb{N}} \Pr(A_n) < \infty$, *alors* $\mathbb{P}(A) = 0$ *en notant* $A = \limsup_n A_n$.

Proof. A est aussi égal à $\bigcap_{n\in\mathbb{N}} \bigcup_{k\geq n} A_k$. On a alors $\mathbb{P}(A) = \lim_{n\to\infty} \mathbb{P}\left(\bigcup_{k\geq n} A_k\right)$. Or, $\mathbb{P}\left(\bigcup_{k\geq n} A_k\right) \leq \sum_{k\geq n} \mathbb{P}(A_k)$. En passant à la limite, on obtient $\mathbb{P}(A) = 0$. $\square$

Une application du premier lemme est :

Proposition 1. *Soit* X_n *une suite de variables aléatoires et* X *une variable aléatoire discrète. On pose* $A_n(\epsilon) = \{|X_n - X| > \epsilon\}$. *Si pour tout* $\epsilon > 0$, $\sum_{n\in\mathbb{N}} \Pr(A_n(\epsilon)) < \infty$, *alors* X_n *converge presque sûrement vers* X.

Proof. Pour tout ϵ, on note $A(\epsilon) = \limsup_n A_n(\epsilon)$. D'après les hypothèses et le premier lemme de Borel-Cantelli, on a $\mathbb{P}(A(\epsilon)) = 0$ pour tout ϵ. On a alors

$$\mathbb{P}\left(\bigcup_{n\in\mathbb{N}} A\left(2^{-n}\right)\right) \leq \sum_{n\in\mathbb{N}} \mathbb{P}(A\left(2^{-n}\right)) = 0. \tag{1}$$

Par complémentaire, on a $\mathbb{P}(\bigcap_{n\in\mathbb{N}} \overline{A\left(2^{-n}\right)}) = 1$. Soit

$$\mathbb{P}(\bigcap_{n\in\mathbb{N}} \bigcup_{j\in\mathbb{N}} \bigcap_{k\geq j} \{|X_j - X| \leq 2^{-k}\}) = 1. \tag{2}$$

Ce qui permet de conclure que X_n converge presque sûrement vers X.

$\square$

Deuxième lemme

Lemme 2. *Soit* (A_n) *une suite d'évènements indépendants. Si* $\sum_{n\in\mathbb{N}} \mathbb{P}(A_n) = \infty$, *alors* $\mathbb{P}(A) = 1$ *en notant* $A = \limsup_n A_n$.

Proof. Pour commencer, on va considérer le complémentaire de A, que l'on va noter B, donc

$$B = \bigcup_{n\in\mathbb{N}} \bigcap_{k\geq n} \overline{A_k}. \tag{3}$$

Remarquons que $\left(\bigcap_{k\geq n} \overline{A_k}\right)$ est une suite croissante d'évènements. On a alors

$$\mathbb{P}\left(\bigcup_{n\in\mathbb{N}} \bigcap_{k\geq n} \overline{A_k}\right) = \lim_{n\to\infty} \mathbb{P}\left(\bigcap_{k\geq n} \overline{A_k}\right). \tag{4}$$

Or, $\mathbb{P}\left(\bigcap_{k\geq n}\overline{A_k}\right) = \prod_{k\geq n}\mathbb{P}(\overline{A_k}) = \prod_{k\geq n}(1-\mathbb{P}(A_k))$. On peut alors appliquer l'inégalité de convexité suivante : $1-x \leq e^{-x}$.

$$\prod_{k\geq n}\mathbb{P}(\overline{A_k}) \leq \prod_{k\geq n} e^{-\mathbb{P}(A_k)} = e^{-\sum_{k\geq n}\mathbb{P}(A_k)}. \tag{5}$$

On peut aussi remarquer que $\lim_n -\sum_{k\geq n}\mathbb{P}(A_k) = -\infty$ puisqu'il s'agit du reste d'une suite divergente. Ainsi,

$$\lim_n \prod_{k\geq n}\mathbb{P}(\overline{A_k}) = 0. \tag{6}$$

Et donc, $\mathbb{P}(B) = 0$, ce qui permet de conclure que $\mathbb{P}(A) = 1$. $\square$

On peut appliquer ce lemme pour démontrer la fameuse expérience de pensée des singes de Shakespeare.

Proposition 2. *Un singe tape au hasard sur un clavier (de 26 lettres) pour un temps infini. Alors on trouvera, presque sûrement, Hamlet dans le texte tapé.*

Proof. On suppose que la longueur de Hamlet est l. On note A_n l'évènement "Hamlet" est tapé entre les n et $n+l-1$ lettres. Les A_n sont indépendants et $\mathbb{P}(A_n) = 26^{-l}$. Et donc, $\sum_{n\in\mathbb{N}}\mathbb{P}(A_n) = \infty$. En utilisant le deuxième lemme, on a que $\mathbb{P}(A) = 1$ où $A = \limsup_n A_n$. C'est-à-dire que

$$\mathbb{P}\left(\bigcap_n \bigcup_{k\geq n} A_k\right) = 1. \tag{7}$$

Ce qui signifie que Hamlet sera presque sûrement tapé. $\square$

Proposition 3. *Il n'existe pas de probabilité $\mathbb{P}$ sur $(\mathbb{N}^*, \mathcal{P}(\mathbb{N}^*))$ telle que $\mathbb{P}(n\mathbb{N}^*) = \frac{1}{n}$.*

Proof. Notons $A_n = n\mathbb{N}^*$ Supposons qu'une telle probabilité existe. Remarquons que si p et q deux premiers distincts, alors $A_p \cap A_q = A_{pq}$. En effet, on a généralement que $A_p \cap A_q \subset A_{pq}$ mais puisque p et q sont premiers, on a que $A_{pq} \subset A_p \cap A_q$ (plus généralement, cela est vrai pour deux nombres premiers entre eux). Par définition,

$$\mathbb{P}(A_p \cap A_q) = \mathbb{P}(A_{pq}) = \frac{1}{pq} = \mathbb{P}(A_p)\mathbb{P}(A_q). \tag{8}$$

On remarquera que pour un nombre arbitraire de premiers distincts, les A_p seront indépendants. Notons $p_1, ..., p_n, ...$ les nombres premiers rangés dans l'ordre croissant.

On sait que $\sum_{k=1}^{\infty}\frac{1}{p_k} = \infty$ donc d'après le deuxième lemme, on a que

$$\mathbb{P}\left(\limsup_n A_{p_n}\right) = 1. \tag{9}$$

Remarquons qu'il n'existe qu'un seul entier multiples d'une infinité de nombre premiers, il s'agit de 0 (ici exclu). De plus, $\limsup_n A_{p_n} \subset \{$entier multiple d'une infinité de premiers $= \emptyset\}$. On obtient une contradiction

$$\mathbb{P}\left(\limsup_n A_{p_n}\right) = \mathbb{P}(\emptyset) = 1. \tag{10}$$

Il n'existe donc pas de probabilité $\mathbb{P}$ sur $(\mathbb{N}^*, \mathcal{P}(\mathbb{N}^*))$ telle que $\mathbb{P}(n\mathbb{N}^*) = \frac{1}{n}$. □

On rappelle une preuve de la divergence de la série des inverses de nombres premiers, autant le faire avec des probabilités.

Proposition 4. *La série* $\sum_{n\in\mathbb{N}} \frac{1}{p_n}$ *diverge.*

Proof. On définit la probabilité suivante pour $s > 1$:

$$\mathbb{P}(\{n\}) = \frac{1}{\zeta(s)}\frac{1}{n^s} \tag{11}$$

où ζ est la fonction zêta de Riemann. On peut alors montrer que $\mathbb{P}$ est bien une probabilité. On reprend la notation A_p pour $p\mathbb{N}^*$. Remarquons déjà que pour un ensemble B, on a

$$\mathbb{P}(B) = \sum_{n\in B} \mathbb{P}(\{n\}) = \frac{1}{\zeta(s)} \sum_{n\in B} \frac{1}{n^s}. \tag{12}$$

En particulier, pour A_p avec p premier, on a

$$\mathbb{P}(A_p) = \frac{1}{\zeta(s)} \sum_{n\in p\mathbb{N}^*} \frac{1}{n^s} = \frac{1}{\zeta(s)} \sum_{n\in\mathbb{N}^*} \frac{1}{(pn)^s} = \frac{1}{p^s}. \tag{13}$$

On considère $q_1, ..., q_k$ un ensemble de nombres de premiers distincts

$$\mathbb{P}\left(\bigcap_{i=1}^k A_{q_i}\right) = \sum_{n\in\bigcap_{i=1}^k A_{q_i}} \mathbb{P}(\{n\}) \tag{14}$$

$$= \frac{1}{\zeta(s)} \sum_{n\in\mathbb{N}^*} \frac{1}{(q_1 \dots q_k n)^s} \tag{15}$$

$$= \prod_{i=1}^k \frac{1}{q_i^s} \qquad = \prod \mathbb{P}(A_{q_i}). \tag{16}$$

On en déduit que les (A_{p_i}) sont indépendants. Notons B l'ensemble des entiers qui ne sont multiples d'aucun nombre premier.

$$\mathbb{P}(B) = \mathbb{P}(\{1\}) + \mathbb{P}(B \setminus \{1\}) = \frac{1}{\zeta(s)} \tag{17}$$

car $\mathbb{P}(B \setminus \{1\}) = \mathbb{P}(\emptyset)$. Mais par définition :

$$\mathbb{P}(B) = \mathbb{P}\left(\bigcap_{k\geq 1} \overline{A_{p_k}}\right) \tag{18}$$

d'où

$$\mathbb{P}(B) = \prod_{k\geq 1} 1 - \frac{1}{p_k^s} = \frac{1}{\zeta(s)}. \tag{19}$$

Appliquons un logarithme :

$$-\log(\zeta(s)) = \sum_{k\geq 1} \log\left(1 - \frac{1}{p_k^s}\right). \tag{20}$$

Or, $\log(1-x) \leq -x$ pour $x \in [0,1]$ donc

$$-\log(\zeta(s)) \geq -\sum_{k\geq 1} \frac{1}{p_k^s}. \tag{21}$$

On en déduit

$$\log(\zeta(s)) \leq \sum_{k\geq 1} \frac{1}{p_k^s} \leq \sum_{k\geq 1} \frac{1}{p_k}. \tag{22}$$

Or, $\zeta(s)$ tend vers l'infini quand s tend vers 1, donc la série $\sum_{k\geq 1} \frac{1}{p_k}$ diverge. $\square$