

résultat quasi miraculeux confirme que C/n est théoriquement la plus équitable attribution de poids aux entiers. La densité $D_e(A) = \sum \{1/k^e, k \text{ dans } A\} / \sum \{1/k^e, k \text{ entier}\}$ est toujours définie si $e > 1$, car maintenant le dénominateur est fini (car la série de terme $1/k^e$ est de somme finie quand $e > 1$).

Comme $1/k^e$ est proche de $1/k$ si e est proche de 1, pour contourner la divergence de la série $\sum 1/k$, on définit la densité $D_z(A)$ de l'ensemble A d'entiers comme la limite, si elle existe, de $D_e(A)$ quand e tend vers 1. On la nomme parfois « densité zêta » (à cause du dénominateur qui est la fonction zêta de Riemann), d'où le « z ». Les ensembles d'entiers A dont la densité D_z existe sont les mêmes que ceux pour lesquels la densité logarithmique (voir l'encadré 3) existe, et alors les deux densités coïncident. Ce résultat remarquable est publié dans la thèse de Persi Diaconis (voir la bibliographie).

Les deux façons envisageables d'étendre la notion de densité naturelle en se fondant sur l'idée d'attribuer (autant que possible) une probabilité $1/n$ à l'entier n fonctionnent et conduisent de manière inattendue à la même notion. Le sens de ces théorèmes mathématiques est que, moyennant un petit passage par des limites pour contourner la divergence de $\sum 1/n$, il est possible de formuler une définition de « choix aléatoire équitable » entre entiers. Ainsi, la loi de Zipf est dans une position privilégiée par rapport à toutes les autres lois envisageables sur les entiers.

Benford retrouvé

Un dernier théorème s'ajoute aux arguments pratiques et mathématiques. La loi de Benford apparaît quand on examine le premier chiffre d'une série de données : la probabilité pour qu'un nombre entier pris dans un ensemble assez grand (par exemple dans le tableau des longueurs de tous les fleuves mesurées en kilomètres) commence par le chiffre i ($i = 1, 2, \dots, 9$) est $\log_{10} \{1 + 1/i\}$ et il y a donc plus de 30 % de chances (car $\log_{10} \{2\} = 0,30103$) que le premier chiffre d'un tel nombre soit 1 (voir l'article de janvier 2007 dans cette rubrique).

Une question s'est posée depuis longtemps : la liste des entiers vérifie-t-

elle la loi de Benford ? Autrement dit, le quotient (Nombre d'entiers $< n$ commençant par le chiffre i)/ n a-t-il pour limite $\log_{10} \{1 + 1/i\}$? Il se trouve que non : ce quotient n'a pas de limite quand n tend vers l'infini, mais oscille sans cesse, car les nombres commençant par i n'ont pas de densité naturelle. Maintenant que nous savons qu'on peut étendre la densité naturelle, reformulons la question : l'ensemble des entiers commençant par i ($i = 1, 2, \dots, 9$) a-t-il une densité logarithmique (ou, ce qui revient au même, une densité D_z) ? Comme par miracle, la réponse est oui, et la densité trouvée est $\log_{10} \{1 + 1/i\}$. Les entiers, traités équitablement, suivent la loi de Benford et il n'est donc pas étonnant que cette loi se rencontre si fréquemment.

À défaut de mesure de probabilité uniforme sur les entiers, la loi de Zipf joue le rôle d'une mesure naturelle. On le voit en observant la prévalence de cette loi pour des données concrètes (telles les fréquences d'utilisation des mots dans un texte). On le voit quand on s'aperçoit que la loi de Zipf est liée à la complexité de Kolmogorov des entiers. On le voit aussi quand on recherche les densités qui, pour les ensembles d'entiers, jouent le rôle de probabilités uniformes.

La preuve mathématique qu'avec ces densités en C/n , les entiers eux-mêmes vérifient la loi de Benford (et c'est vrai aussi des nombres premiers), renforce la conviction que ces densités tirées de la loi de Zipf sont les bonnes façons de choisir au hasard des nombres entiers. Le fait que la loi de Benford redonne en un certain sens la loi de Zipf est un argument supplémentaire, car il existe des explications directes de la loi de Benford, telle celle proposée par Nicolas Gauvrit et l'auteur de cette rubrique, et qu'elles sont donc aussi des explications de la loi de Zipf.

Le monde mathématique est déconcertant : l'infini dénombrable, le plus simple de tous, semble interdire qu'on en pioche les éléments au hasard équitablement, alors que le continu de l'intervalle $[0, 1]$, plus gros et plus compliqué que l'infini dénombrable, l'autorise. Heureusement, la loi de Zipf, à sa façon, joue ce rôle de probabilité uniforme sur les entiers. ■

L'AUTEUR



J.-P. DELAHAYE est professeur à l'Université de Lille et chercheur au Laboratoire d'informatique fondamentale de Lille (LIFL).

BIBLIOGRAPHIE

N. Gauvrit, J.-P. Delahaye et H. Zenil, *Sloane's Gap : Do mathematical and social factors explain the distribution of numbers in the OEIS ?*, *Journal of Humanistic Mathematics*, à paraître, 2012.

N. Gauvrit et J.-P. Delahaye, *Pourquoi la loi de Benford n'est pas mystérieuse*, *Mathématiques et Sciences Humaines*, vol. 182-2, pp. 7-15, 2008 (<http://msh.revues.org/10363>).

M. Newman, *Power laws, Pareto distributions and Zipf's law*, *Contemporary Physics*, vol. 46, pp. 323-351, 2005.

G. Tenenbaum, *Introduction to Analytic and Probabilistic Number Theory*, Cambridge University Press, 1995.

P. Diaconis, *Weak and Strong Averages in Probability and Theory of Numbers*, thèse de l'Université Harvard, 1974, http://www-stat.stanford.edu/ffcgates/PERSI/Persi_PhDdiss.pdf [c'est le document à lire sur les densités étendant la densité naturelle].

G. Zipf, *Human Behaviour and the Principle of Least Effort*, Addison-Wesley Publishing, 1949.

SCIENCE & FICTION

La Force des Jedis

Dans l'univers de *La guerre des étoiles*, elle serait conférée par des organismes microscopiques, les midichloriens. Qui sont ces « supports » de la Force ?

Jean-Sébastien STEYER et Roland LEHOUCQ

« **Q**ue la Force soit avec toi ! »
La Force est cette source de pouvoir des chevaliers Jedis dans l'univers de *Star Wars*. On apprend dans

le premier épisode, *La menace fantôme* (1999), que la Force a un support matériel, d'étranges organismes microscopiques nommés midichloriens. Une discussion entre le jeune Anakin Skywalker et son mentor Qui-Gon Jinn, nous en précise la nature :

« Les midichloriens sont des formes de vie microscopiques qui résident dans toutes les cellules vivantes et communiquent avec la Force. [Ils vivent] dans tes cellules. Nous vivons en symbiose avec les midichloriens. »

En introduisant les midichloriens dans l'histoire, le réalisateur George Lucas tente alors, au grand dam de ses fans, de donner une explication rationnelle à la Force. Approche scientifique ? Loin s'en faut si l'on regarde de plus près qui sont ces midichloriens venus de l'espace. Éliminons d'abord une éventuelle nature virale, les virus étant considérés comme des parasites intracellulaires, alors que les midichloriens sont des symbiotes, d'après la définition. Remarquons que le mot « midichlorien », forgé par G. Lucas, semble être un assemblage des mots « mitochondrie » et « chloroplaste », deux organites présents dans les cellules terriennes. Les mitochondries, localisées dans le cytoplasme, fournissent les molécules d'ATP (adénosine triphosphate), qui permet aux cellules de stocker de l'énergie. La mitochondrie est l'usine énergétique des cellules à noyau, dites eucaryotes, aussi bien animales que végétales.



1. ANAKIN SKYWALKER, qui deviendra Dark Vador dans les épisodes ultérieurs de *La guerre des étoiles*, a une affinité marquée avec la Force. Cette dernière lui est conférée par les midichloriens, des micro-organismes présents notamment dans son sang.

Quant au chloroplaste, c'est un organite spécifique des cellules végétales qui permet la synthèse de la chlorophylle dans les algues et les plantes. Les mitochondries et les chloroplastes ont leur propre ADN, et seraient issus de bactéries (cyanobactéries pour les chloroplastes et protéobactéries pour les mitochondries) qui se seraient « alliées », il y a 1,5 à 2 milliards d'années, avec des cellules pour finir par y vivre en symbiose. Cette hypothèse de l'endosymbiose a été proposée en 1967. Ainsi, les mitochondries et les chloroplastes sont synonymes d'énergie vitale tant pour les animaux que pour les végétaux. L'assemblage des deux confère apparemment des pouvoirs hors du commun aux porteurs de midichloriens.

Mais, même s'ils semblent présenter des propriétés symbiotiques identiques à celles des mitochondries ou des chloroplastes, les midichloriens sont, toujours selon la description de Maître Qui-Gon Jinn, des organismes et non des organites. Ils peuvent même être considérés comme des organismes pathogènes ou mutagènes à part entière. Dès lors, du point de vue microbiologique, les Jedis seraient des mutants contaminés, et une analyse de sang permet d'identifier le stade de contamination, voire le degré de contagiosité ! D'ailleurs, pour tester l'affinité du jeune Anakin avec la Force, Qui-Gon Jinn lui fait une prise de sang visant à déterminer la quantité de midichloriens présents dans son organisme. Le fait que la présence de midichloriens soit quantifiable semble leur donner une base scientifique.

Résultat : 20 000. En quelle unité ? On l'ignore... Est-ce le corps du jeune homme qui en contient cette quantité ? Est-ce seulement