



LOGIQUE & CALCUL

Les entiers ne naissent pas égaux

Il est impossible de définir une loi de probabilité uniforme sur l'ensemble des nombres entiers. Ce fait est étroitement lié à la loi de Zipf, une loi statistique dont les manifestations sont innombrables.

Jean-Paul DELAHAYE

Dans un dé à six faces, le « 1 », comme chacune des autres faces, a la probabilité $1/6$ de « sortir » et en lançant le dé on a une probabilité de $3/6 = 1/2$ d'obtenir un nombre pair. Sur un ensemble ayant n éléments, une probabilité est équilibrée (on dit aussi « uniforme ») quand chaque élément, comme une face d'un dé, a la même probabilité. La probabilité d'un sous-ensemble de p éléments est alors p/n , le quotient du nombre de « cas favorables » sur le nombre total de cas. Cette probabilité p/n est l'outil de base pour comprendre la roulette des casinos et les tirages du loto. La question est : que se passe-t-il quand le nombre total d'éléments est infini ?

Pour l'ensemble infini des nombres réels compris entre 0 et 1, il est facile de définir une probabilité uniforme. On obtient la mesure de Lebesgue découverte par Henri Lebesgue (1875-1941) au début du XX^e siècle. Cette mesure sur $[0, 1]$ est fondée sur la longueur des intervalles $[a, b]$ de $[0, 1]$ auxquels on attribue la probabilité $b-a$. Elle donne un sens à l'opération « tirer un nombre réel équitabement entre 0 et 1 ». Elle indique par exemple que la probabilité pour qu'un tel nombre se trouve à moins d'un centième de

$1/\pi$ ou de $1/e$ est exactement $1/25$, car la zone favorable est composée des deux morceaux disjoints $[1/\pi - 1/100, 1/\pi + 1/100]$ et $[1/e - 1/100, 1/e + 1/100]$ dont la longueur totale est $4/100 = 1/25$ (voir le schéma en bas de la page).

La mesure de Lebesgue montre aussi que les nombres rationnels sont négligeables : la probabilité de tirer au hasard un nombre rationnel est nulle et la probabilité de tirer un nombre irrationnel est égale à un. En effet, on sait enfermer les rationnels, qui sont numérotés par des entiers (ils sont dénombrables), dans une réunion d'intervalles de longueurs $c/2^n$, de longueur cumulée c que l'on peut rendre aussi petite que possible.

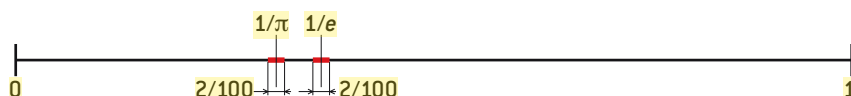
Étrangement, bien que l'infini dénombrable (celui des nombres entiers $1, 2, \dots, n, \dots$) précède et soit plus simple que l'infini des réels de l'intervalle $[0, 1]$ (dénommé « continu »), il est impossible d'y définir une probabilité uniforme qui jouerait, pour les entiers, le rôle que la mesure de Lebesgue joue pour $[0, 1]$. C'est très ennuyeux, car cela interdit de donner un sens mathématique au choix aléatoire équitabement d'un entier. « Tirer un nombre réel au hasard entre 0 et 1 sans en favoriser aucun » est mathématiquement possible et la théorie moderne des probabilités déduit de la solution de Lebesgue des résultats concernant

la physique, l'économie et toutes sortes de problèmes concrets. « Tirer un entier au hasard, sans en favoriser aucun » n'a, en revanche, mathématiquement pas de sens pour quiconque veut rester dans le cadre de la théorie des probabilités fixée en 1933 par Andreï Kolmogorov. Nous allons voir cependant qu'il existe de subtiles et merveilleuses solutions de remplacement, nommées densités, mises au point pour les besoins de l'arithmétique et qui sont importantes pour comprendre les distributions statistiques concrètes.

Probabilités non uniformes

Pour déterminer la probabilité d'un ensemble de nombres entiers positifs en oubliant temporairement l'exigence d'uniformité, nous prenons une série $\{a_n\}$ de nombres positifs ou nuls dont la somme est égale à 1, par exemple $a_n = 1/2^n$ qui vérifie bien $1/2 + 1/4 + \dots + 1/2^n + \dots = 1$. Une fois $\{a_n\}$ choisi, la probabilité d'un ensemble d'entiers A est la somme des termes a_n dont l'indice n est un élément de A .

Avec le choix $a_n = 1/2^n$, nous trouvons que l'ensemble I des nombres impairs $\{1, 3, \dots, 2n+1, \dots\}$ a pour probabilité $1/2 + 1/8 + 1/32 + \dots + 1/2^{2n+1} + \dots = 2/3$ et que l'ensemble P des nombres pairs $\{2, 4, \dots, 2n, \dots\}$ a pour probabilité $1/3$. Ce n'est guère satisfaisant : cela ne semble





1. Probabilités des entiers dans l'encyclopédie de Neil Sloane

Il est impossible d'associer une probabilité uniforme aux entiers (probabilité qui donnerait à chaque entier la même probabilité). En effet, la probabilité d'un sous-ensemble des entiers est égale à la somme des probabilités de chaque élément : si la probabilité d'un entier est non nulle, la probabilité de l'ensemble des entiers est infinie, ce qui ne convient pas pour une probabilité ; si elle est nulle, alors la probabilité de l'ensemble des entiers est nulle aussi, ce qui ne convient pas mieux. Pour traiter des probabilités d'ensembles d'entiers, il faut attribuer des probabilités différentes aux entiers, de la meilleure façon possible. L'étude des entiers présents dans l'encyclopédie de Sloane est un guide pour trouver cette meilleure attribution.

L'encyclopédie des suites numériques de Neil Sloane contient un

grand nombre de fois l'entier 1, un peu moins souvent l'entier 2, etc. La courbe, notons-la $Sloane(n)$, donnant le nombre d'occurrences de n en fonction de n ressemble à un nuage de points globalement décroissants. Classés par nombre d'occurrences décroissantes, on a une loi de type $C/n^{1,3}$, où C est une constante. Le nuage de Sloane est lié à la complexité de Kolmogorov et cela justifie que l'on retrouve une loi de Zipf. La courbe $Sloane(n)$ est une version approchée de la courbe $C/2^{K(n)}$, où $K(n)$ désigne la complexité de Kolmogorov de l'entier n , qui vaut typiquement $\log_2(n)$: quand n est « complexe », $K(n)$ est grand (n est impossible à définir simplement), $1/2^{K(n)}$ est petit et donc le point correspondant de $Sloane(n)$ est bas. C'est pourquoi aussi $1024 = 2^{10}$ est situé au-dessus des autres points de même ordre de grandeur : le fait

d'être une puissance de 2 a pour conséquence que sa complexité de Kolmogorov est faible et donc que son nombre d'occurrences dans la base (proportionnel à $1/2^{K(n)}$) est grand.

Cette représentation concrète de la complexité de Kolmogorov par le nuage de Sloane est légèrement biaisée par un effet culturel et social qu'on voit clairement sous la forme d'une zone claire, le fossé de Sloane, séparant le nuage en une partie inférieure et une partie supérieure.

Ce fossé étudié récemment (voir la bibliographie) est dû à ce que la communauté mathématique qui contribue à l'encyclopédie s'intéresse aux nombres entiers ayant une faible complexité de Kolmogorov (ce sont les nombres simples à définir et ayant plusieurs définitions), mais qu'elle s'y intéresse

d'une manière qui augmente la fréquence des nombres simples.

À cause des phénomènes culturels d'entraînements et de mode, les mathématiciens concentrent leur attention sur les nombres les plus simples, ce qui pousse les points correspondants de la courbe vers le haut, créant une zone évidée dans le graphe : le fossé.

