INSTITUT DE FRANCE
Académie des sciences

# HASARD? VOUS AVEZ DIT HASARD?

Pour maîtriser l'inconnu en « mesurant » le hasard, les scientifiques ont élaboré des objets mathématiques, les matrices aléatoires. **Alice Guionnet** en expose la genèse et montre comment cette découverte majeure irrigue d'autres sciences et de nombreux domaines de la société.

**MATHÉMATIQUES**

Dans le cadre de notre partenariat avec l'Académie des sciences, des académiciens nouvellement élus présentent un éclairage sur l'actualité de la recherche scientifique dans leur discipline à travers leur expérience personnelle.

« Comment oser parler des lois du hasard ? Le hasard n'est-il pas l'antithèse de toute loi ? (…) La probabilité est opposée à la certitude ; c'est donc ce qu'on ignore et par conséquent semble-t-il ce qu'on ne saurait calculer. » C'est ainsi que le mathématicien Poincaré commence son célèbre texte sur le hasard. Pourtant, le calcul des probabilités a envahi de nombreuses sciences et la société, notamment par le biais des statistiques. Et moi-même et mes collègues passons nos journées à effectuer ces calculs !

Comment ça marche ? Tout commence au XVI[e] siècle dans les travaux de l'Italien Cardano. Cardano est scientifique et médecin. On lui doit l'invention du cardan, pièce mécanique toujours utilisée dans les automobiles. Mais il est aussi un grand mathématicien. Il est le premier à avoir l'idée que le hasard peut être mesuré. L'histoire dit que Cardano arrondit ses fins de mois grâce aux jeux de hasard et, dans cet esprit, il essaye de comprendre quelle est sa chance d'effectuer un double six. Même si le mouvement d'un dé est régi par des lois physiques connues, un lancer dépend de trop de paramètres (position initiale, direction du lancer, etc.) pour qu'on puisse prédire le résultat exactement. Mais la chance – ou probabilité – d'obtenir une face comme le six peut être calculée. Cardano remarque que toutes les faces doivent avoir la même probabilité d'apparaître, donc une chance sur six. Quand il lance deux dés, la face sur laquelle tombe un dé ne doit pas influencer la face sur laquelle va tomber l'autre dé : ces événements sont indépendants ! Il doit donc y avoir une probabilité de 1/36 d'effectuer un double six.

### LA LOI DES GRANDS NOMBRES

Ces constatations peuvent aujourd'hui sembler bien évidentes, mais elles sont à la base du calcul des probabilités. Elles permettent de mesurer ce qui est inconnu, incalculable, en le décrivant par une probabilité. Et cela permet de prédire, avec une certaine probabilité, ce qui peut arriver. À partir de ce moment, la théorie des probabilités va rapidement se développer. Une des premières questions est de pouvoir vérifier que nous ne nous sommes pas trompés en donnant une probabilité de 1/6 à une face d'apparaître dans le lancer de dés.

Regardons un cas plus simple : une pièce. A priori les deux faces ont la même chance d'apparaître : 1/2. Mais comment le vérifier ? Le bon sens nous dit que si on jette 1 000 fois notre pièce, le pile devrait apparaître environ 500 fois. Mais à partir de quelle erreur suspecterons-nous que la pièce est truquée et notre modèle mauvais ?

La réponse va venir de Bernoulli et de de Moivre au XVIII[e] siècle. Ils vont montrer qu'il y a plus de 65 % de chances que le côté pile apparaisse entre 460 et 540 fois, mais une probabilité inférieure à 2 % qu'il apparaisse moins de 400 fois. Le dernier cas est très


FRÉDÉRIC BELLAY

**PROFIL**

Alice Guionnet, mathématicienne, membre de l'Académie des sciences, est directrice de recherches au CNRS dans l'unité des mathématiques pures et appliquées de l'ENS Lyon. Connue pour ses travaux sur les probabilités et les grandes matrices aléatoires, elle est notamment titulaire de la médaille d'argent du CNRS et du prix Loève.

improbable et, si nous l'observons, nous nous demanderons si notre modèle est mauvais. Peut-être que notre pièce est biaisée et a plutôt une probabilité de 0,4 de tomber sur pile ? Nous pourrons valider cette nouvelle hypothèse en effectuant de nouveaux lancers. Le théorème de Bernoulli s'appelle la loi des grands nombres : il montre que si on lance une pièce un grand nombre de fois, alors le côté pile doit apparaître environ la moitié des fois. De Moivre quantifie l'erreur qu'on peut commettre par rapport à cette moyenne : elle est de l'ordre de la racine du nombre de lancers (ici environ 33) et est décrite par la courbe en cloche. C'est cette dernière qui nous dit qu'on a une probabilité de moins de 2 % de voir apparaître seulement 400 piles.

### 1,05 GARÇON POUR 1 FILLE

Laplace montre bientôt que ces théorèmes ne se limitent pas aux pièces de monnaie mais s'appliquent dans de nombreuses autres circonstances, dès lors qu'elles font intervenir des événements indépendants. Cela lui permet de les utiliser pour répondre à des questions de la vie civile comme la démographie. Par exemple, il étudie le nombre de garçons par rapport au nombre de filles dans différentes régions de France. Dans la plupart, la moyenne est similaire : environ 1,05 garçon pour 1 fille. L'erreur par rapport à cette moyenne est de l'ordre de la racine du nombre de personnes recensées, la courbe en cloche réapparaît ! Mais que s'est-il passé pour que certaines régions aient un ratio différent ? Il nous faut en tout cas identifier ce phénomène qui n'est, lui, plus lié au hasard ! Ce jeu-là est-il truqué ?

La courbe en cloche est aussi utilisée pour effectuer les sondages, par exemple pour déterminer la taille de l'échantillon sondé pour une erreur donnée. Comme on l'imagine, le hasard devient un fourre-tout où on peut mettre de nombreuses choses qu'on ignore exactement. Il a donc envahi de nombreuses sciences et les objets mathématiques probabilistes se sont multipliés. Par exemple, en physique, la mécanique statistique cherche à expliquer le comportement de la matière en modélisant les interactions des petites particules qui la composent. À cette échelle, de nombreux phénomènes nous sont inconnus ou difficilement calculables : nous les modélisons donc par des particules aléatoires. Comme il y a beaucoup de particules par gramme de matière, on s'attend, comme dans la loi des grands nombres, à voir émerger un comportement moyen qui n'a plus rien d'aléatoire. Nous pouvons alors vérifier que nos modèles sont bons si ce comportement moyen est conforme à nos observations. On peut ainsi comprendre le phénomène d'aimantation du fer en le décrivant par les petits moments magnétiques qui le composent. Cette démarche s'étend au vivant : on modélise le cerveau par des réseaux de neurones. Ces réseaux sont aujourd'hui au cœur de l'intelligence artificielle.

### DÉTERMINER DES CORRÉLATIONS

L'objet probabiliste qui m'intéresse est apparu au début du XX$^{e}$ siècle dans les travaux du statisticien Wishart. Celui-ci considère alors un grand tableau de données et se demande si elles sont corrélées. Par exemple, imaginons que notre tableau représente les performances de nombreux athlètes dans différents sports. Comment savoir si ces résultats sont corrélés et, par exemple, que si on est bon au foot on devrait l'être aussi en natation ? Si les performances de tous les athlètes dans ces deux sports sont proportion- »»

**« Le bon sens nous dit que si on jette 1 000 fois une pièce, le pile devrait apparaître environ 500 fois. Mais à partir de quelle erreur suspecterons-nous que la pièce est truquée ? »**

))) nelles, on conclura assez vite que ces résultats sont reliés… Mais comment trouver des critères plus fins, et comment savoir si ce qu'on observe est lié au hasard, qui pourrait avoir rendu ces données peu représentatives, ou à une raison biologique ?

### INFORMATIQUE, PHYSIQUE ATOMIQUE…

C'est ainsi que Wishart se mit à utiliser de grands tableaux dont les coefficients étaient choisis au hasard : c'est ce qu'on appelle une grande matrice aléatoire. Comme dans le jeu du pile ou face, l'idée est de voir si les données étudiées ont les mêmes propriétés que les matrices aléatoires, et sinon de chercher les hypothèses supplémentaires qui ont pu engendrer nos observations. Pour cela, les mathématiciens étudient toutes sortes de matrices aléatoires afin d'en comprendre les propriétés typiques.

Aujourd'hui encore, les matrices aléatoires sont utilisées pour analyser des données. Avec l'avènement du big data, ces tableaux deviennent de plus en plus grands, et comme les mathématiciens n'ont peur de rien, ils supposent que leur taille tend vers l'infini !

Mais les matrices aléatoires ne sont pas restées cantonnées aux tableaux de données. Elles se sont d'abord répandues en physique atomique pour décrire la physique des noyaux lourds. Même si celle-ci est donnée par la mécanique quantique, elle donne lieu à des équations si compliquées qu'on ne sait pas les résoudre. Dans les années 1950, le physicien Wigner proposa un modèle plus simple à analyser, basé sur des matrices aléatoires. Et il trouva que les niveaux d'énergies ainsi prédits étaient proches de ceux observés ! Quand cela n'est pas le cas, Dyson propose de rechercher les propriétés physiques qui n'ont pas été prises en compte dans cette modélisation pour arriver à un modèle plus proche de la réalité.

Les matrices aléatoires sont apparues dans de nombreuses autres sciences, comme dans l'informatique avec la compression de données, ou avec les problèmes de données sous-déterminées comme Netflix qui recommande des films à ses clients sur la base d'un très faible nombre de « like ». Moi-même, je suis venue aux matrices aléatoires par le biais des réseaux de neurones. Bref, comme l'a titré récemment le magazine « Pour la science », les matrices aléatoires sont partout !

### DE NOUVEAUX HORIZONS

Mais elles sont tout d'abord un magnifique objet qui a ouvert un nouveau champ des mathématiques. Elles interviennent non seulement en probabilités et en analyse, mais aussi dans d'autres domaines des mathématiques comme la théorie des nombres ou les algèbres d'opérateurs. En tant que mathématiciens, nous cherchons à comprendre les propriétés fondamentales de ces objets : quelles sont leurs propriétés typiques ? Quelle est la probabilité qu'ils aient un comportement différent ?

Quand je dis que je suis mathématicienne, on me demande souvent s'il y a encore des choses à trouver en mathématique. En vérité, il y a de plus en plus de choses à chercher. D'une part, le développement des sciences pose tous les jours de nouvelles questions mathématiques. D'autre part, l'avancée de notre savoir nous ouvre de nouveaux horizons. Les matrices aléatoires en sont un parfait exemple : impensable sans le concept même de probabilité apparu finalement bien tardivement, leur étude se développe aujourd'hui rapidement, à la fois grâce à leurs applications, mais aussi par la richesse de la théorie qu'elles engendrent. »

SHUTTERSTOCK / JNT VISUAL

**« On me demande souvent s'il y a encore des choses à trouver en mathématique. En vérité, il y a de plus en plus de choses à chercher. Les matrices aléatoires, que l'on retrouve partout, en sont un parfait exemple. »** ALICE GUIONNET

## EN SAVOIR PLUS

**Le site de l'Académie des sciences : www.academie-sciences.fr**

« Pour la science », n° 487, mai 2018 : **« L'énigme des matrices aléatoires. Pourquoi ces objets mathématiques sont-ils partout ? »** (en vente sur www.pourlascience.fr).

**« Vieillissement : un mouvement perpétuel ? »** conférence d'Alice Guionnet devant les lycéens à l'Académie des sciences, à l'occasion de la Semaine des mathématiques, le 13 mars 2018 : www.academie-sciences.fr/fr/Seances-publiques/vieillissement-un-mouvement-perpetuel.html

**« La plus grande valeur propre de matrices de covariance empirique »,** de Sandrine Péché, 2006, en ligne sur : images.math.cnrs.fr/La-plus-grande-valeur-propre-de.html (site du CNRS dédié à la vulgarisation des mathématiques).

**« An Introduction to Random Matrices »,** de Greg W. Anderson, Alice Guionnet et Ofer Zeitouni, Cambridge University Press, 2010. En ligne : cims.nyu.edu/~zeitouni/cupbook.pdf