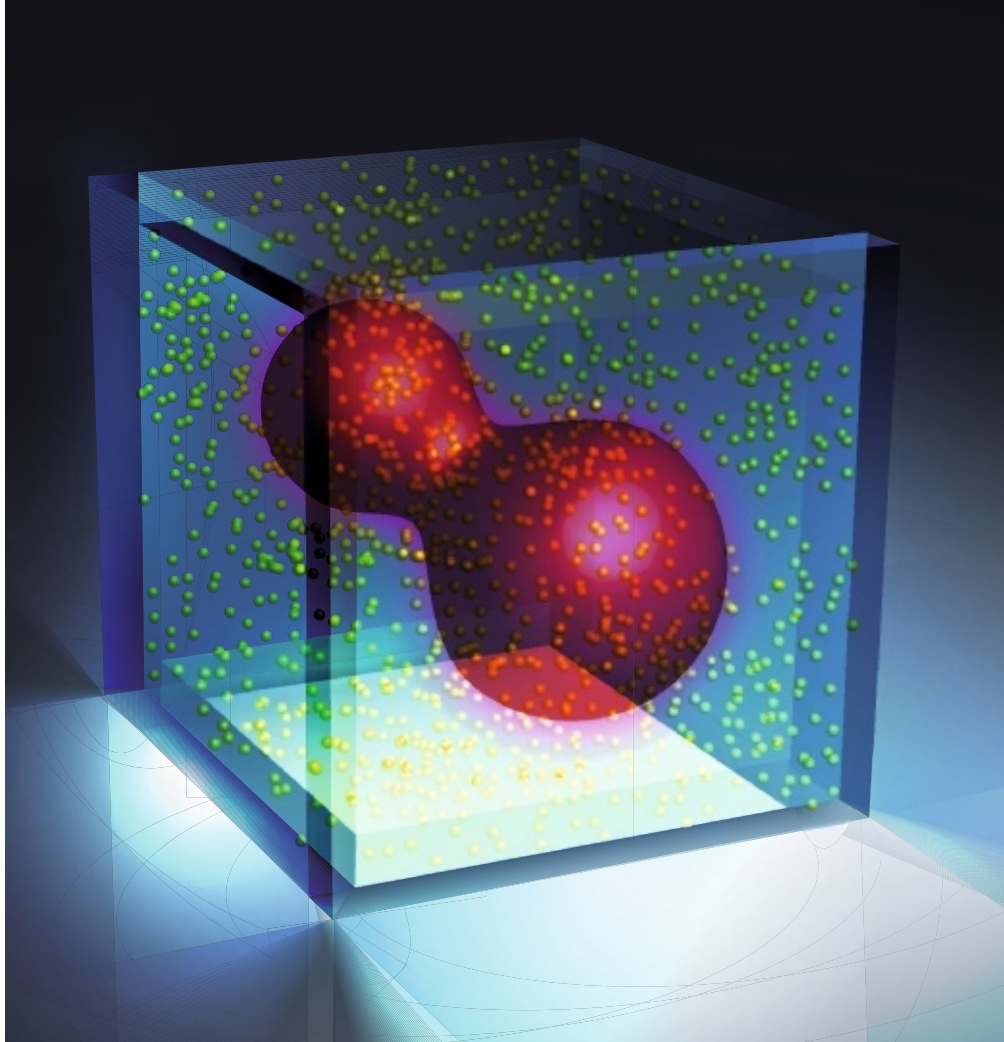


Les méthodes de Monte Carlo font intervenir des nombres pseudo-aléatoires pour effectuer de lourds calculs numériques. Mais des nombres dits quasi-aléatoires, où le hasard a moins de place, se révèlent parfois plus efficaces.



Excursions quasi-

Au début des années 1990, Spassimir Paskov, étudiant en thèse à l'Université Columbia, commençait à analyser un instrument financier exotique nommé CMO (*Collateralized Mortgage Obligation*, obligation garantie par une hypothèque) introduit par la banque d'investissements *Goldman Sachs*. Il cherchait à estimer la valeur actuelle d'un CMO, d'après les possibles encaissements à venir correspondant à des milliers d'emprunts-logement sur 30 ans. Il ne s'agissait pas simplement d'appliquer la formule usuelle pour les intérêts composés. Beaucoup d'emprunts-logement sont liquidés en avance quand le logement est vendu ou refinancé; certains emprunts ne sont pas honorés; les taux d'intérêt montent et baissent. Ainsi, la valeur actuelle d'un CMO sur 30 ans dépend de 360 (30×12) flux de trésorerie mensuels incertains et interdépendants. Il

s'agissait donc plutôt d'évaluer une intégrale dans un espace à 360 dimensions.

Il n'y avait aucun espoir de faire une évaluation exacte. S. Paskov et son directeur de thèse, Joseph Traub, ont décidé d'essayer une technique d'approximation quelque peu obscure, la méthode de «quasi-Monte Carlo». Un calcul de Monte Carlo ordinaire prend des échantillons aléatoires dans l'ensemble de toutes les solutions possibles. La variante «quasi» effectue un échantillonnage d'un type différent, non aléatoire, mais pas tout à fait régulier non plus. S. Paskov et J. Traub ont trouvé que certains de leurs programmes de quasi-Monte Carlo tournent mieux et plus vite que la technique classique. Leur découverte permettrait à un banquier ou un investisseur d'estimer la valeur d'un CMO en quelques minutes de calcul seulement, au lieu de plusieurs heures.

Cela ferait une belle histoire si l'on pouvait alors raconter que la période «d'exubérance irrationnelle» qui a suivi sur les marchés financiers (la frénésie des transactions de produits dérivés complexes, et le triste cortège de crise, effondrement, récession, chômage) avait pour origine une innovation mathématique dans l'évaluation des intégrales de dimension élevée. Mais ce n'est pas le cas; il y avait d'autres causes à cette folie.

Cependant, les travaux de S. Paskov et J. Traub ont bien eu un effet: ils ont suscité un extraordinaire regain d'intérêt pour la méthode de quasi-Monte Carlo. Des résultats théoriques antérieurs avaient suggéré que les modèles de quasi-Monte Carlo commenceraient à s'essouffler quand le nombre de dimensions dépasserait une dizaine ou une vingtaine, en tout cas longtemps avant d'atteindre 360. Aussi le succès de

l'expérience a-t-il été une surprise, que les mathématiciens ont essayé d'expliquer tant bien que mal. Une question clef est de savoir si la même approche fonctionnera pour d'autres problèmes.

Tout cela met en lumière le rôle curieusement ambivalent du hasard en informatique. Les algorithmes sont par définition strictement déterministes, et pourtant beaucoup d'entre eux exploitent une petite dose de hasard : la possibilité, de temps à autre, de faire un choix en tirant à pile ou face.

Pseudo-aléatoire ou quasi-aléatoire ?

En pratique, cependant, les nombres aléatoires fournis aux programmes informatiques ne sont jamais vraiment aléatoires. Ils sont pseudo-aléatoires – ce sont des faux astucieux, qui semblent aléatoires et qui passent les tests statistiques, mais qui sont issus d'une source déterministe. Chose curieuse, les faux nombres aléatoires semblent fonctionner très bien, en tout cas pour la plupart des tâches.

Les nombres quasi-aléatoires vont un pas plus loin. Ils ne font même pas l'effort

de la feuille et celle du panneau. Si l'on convient que l'aire du panneau est égale à 1, l'aire estimée de la feuille sera donc égale à n/N .

Lancer des fléchettes au hasard est non seulement difficile, mais aussi dangereux. Mieux vaut remplacer ces projectiles par des points et laisser l'ordinateur se charger de les parsemer au hasard. La mesure est alors facile, et l'expérience fonctionne remarquablement bien.

J'ai cueilli une feuille d'érable, je l'ai photographiée et j'ai placé son image verte sur un champ carré de 1024×1024 pixels blancs. J'ai alors écrit un programme qui jette 1024 points aléatoires (en fait des points pseudo-aléatoires) sur l'image numérisée. La première fois que j'ai fait tourner le programme, 429 des 1024 points sont tombés sur la feuille, d'où une superficie estimée de 0,4189. La superficie réelle, définie par comptage des pixels verts, était de 0,4185. L'approximation des points aléatoires était meilleure que ce à quoi je m'attendais (un coup de chance, mais rien d'extraordinaire cependant). Quand j'ai répété l'expérience un millier de fois, l'estimation moyenne de la superficie de la feuille était

-aléatoires

de se faire passer pour aléatoires. Et pourtant, ils semblent eux aussi très efficaces dans de nombreux cas où l'on fait habituellement appel au hasard. Ils pourraient même faire mieux que leurs cousins pseudo-aléatoires dans certaines circonstances.

Un petit problème amusant aide à préciser les différences entre pseudo-aléatoire et quasi-aléatoire. Supposons que l'on veuille estimer la superficie d'un objet de forme compliquée, une feuille d'érable par exemple (voir les figures 1 et 2). Une astuce bien connue permet de résoudre ce problème avec l'aide d'un peu de hasard. Placez la feuille sur un panneau de superficie connue, et lancez des fléchettes dessus au hasard, sans viser. Si un total de N fléchettes touche le panneau, et que n d'entre elles atterrissent sur la feuille, alors le rapport n/N est approximativement égal au rapport entre la super-

ficie de la feuille et celle du panneau. Si l'on convient que l'aire du panneau est égale à 1, l'aire estimée de la feuille sera donc égale à n/N .

L'idée de base des calculs de Monte Carlo est de reformuler un problème mathématique (tel que calculer la surface d'une feuille) comme un jeu de hasard, où les gains attendus d'un joueur sont les réponses au problème. Dans les jeux simples, on peut calculer la probabilité exacte de chaque résultat, et les gains attendus peuvent donc également être déterminés exactement. Là où de tels calculs sont infaisables, l'alternative consiste à se lancer et jouer au jeu, et voir ce qui ressort. C'est la stratégie d'une simulation de Monte Carlo. L'ordinateur joue au jeu de nombreuses fois, et prend comme estimation de la valeur attendue réelle le résultat moyen.

Pour le problème de la feuille sur le carré, la valeur attendue est le rapport de la superficie de la feuille à celle du carré. En choisissant N points aléatoires et en

L'ESSENTIEL

✓ De nombreux calculs numériques sont réalisés à l'aide d'un échantillonnage aléatoire de points : c'est le principe des méthodes dites de Monte Carlo.

✓ Dans certaines circonstances qui restent à préciser, un échantillonnage plus régulier, dit quasi-aléatoire, conduit à un calcul plus efficace.

✓ Les méthodes de calcul correspondantes, dites méthodes de quasi-Monte Carlo, connaissent aujourd'hui un regain d'intérêt.

© Shutterstock/Kundin : Pour la Science

Brian Hayes

1. POUR CALCULER LE VOLUME
d'une forme irrégulière (*en rouge*), le principe de la méthode de Monte Carlo est de prendre un échantillon aléatoire de points uniformément distribués dans un domaine plus grand et de volume connu (*ici un cube*), puis de mesurer la proportion des points se trouvant à l'intérieur du volume recherché. Cette proportion est une approximation du volume relatif, d'autant plus précise que le nombre de points est grand.