

Les dérapages incontrôlés des maths

LE MONDE SCIENCE ET TECHNO | 14.09.2015 à 16h18 • Mis à jour le 14.09.2015 à 19h10 | Par David Larousserie

(journaliste/david-larousserie) et Nathaniel Herzberg (journaliste/nathaniel-herzberg/)

Stéphan Cléménçon se souvient parfaitement de son arrivée à Télécom Paris Tech, en 2007. A sa grande surprise, ce spécialiste en statistiques découvrait un département en souffrance. « *Il fallait remplir les classes avec des cours de maths financières. Un comble pour une école d'ingénieurs en télécom! Alors j'ai bataillé un peu pour convaincre que les statistiques pouvaient servir à bien des choses, surtout avec le développement des gigantesques bases de données des géants d'Internet. Aujourd'hui, nous refusons du monde!* »



La Bourse de Wall Street (New York) enregistre, le 26 août, son plus fort gain depuis quatre ans. LUCAS JACKSON / REUTERS

Une preuve de plus que les maths sont partout. Mais pas toujours pour le meilleur! Une sonde martienne se disloque à cause d'une erreur d'unité métrique de la NASA en 1999. Chômage et dettes des Etats explosent depuis 2008 après l'explosion de bulles financières entretenues par des équations douteuses. L'outil de prédiction de la grippe de Google, incarnation de la toute-puissance des algorithmes, surestime de près du double l'épidémie de fin 2012 aux Etats-Unis. En janvier 2013, le Fonds monétaire international concède que son modèle a sous-estimé de moitié l'effet des réductions budgétaires sur la croissance de pays en crise. Ce mois-ci, une équipe annonce n'avoir pu reproduire que 39 résultats de psychologie sur 100 étudiés: les autres n'étaient pas statistiquement significatifs. Et que dire de la dizaine d'erreurs judiciaires imputées à des calculs probabilistes incorrects recensés par Leila Schneps et Coralie Colmez dans leur livre *Les Maths au tribunal* (Seuil, 288 p., 20 euros) ? Simples étourderies, mauvaise utilisation de formules, ou problèmes plus sérieux concernant les mathématiciens eux-mêmes ?

Un peu de tout cela, en vérité. La responsabilité des mathématiciens semble engagée. Le temps est révolu où un Godfrey Hardy, grand mathématicien anglais du début du XX^e siècle, pouvait se vanter que ses théories ne serviraient à rien. Quelques années plus tard, ses idées sur les nombres nourrissaient les méthodes de chiffrement des militaires... « *Puisque nous pensons faire des choses intéressantes, il est normal que cela finisse par intéresser des gens* », constate Jean-Pierre Kahane, mathématicien de l'Académie des sciences et membre du Comité consultatif national d'éthique. « *C'est une illusion de penser que nos travaux ne serviront pas. Le savoir est là, il peut échapper. Mais il ne faut pas perdre de vue le champ d'applications et les hypothèses de ce savoir* », estime Jean-Pierre Bourguignon, mathématicien, actuel président du Conseil européen de la recherche.

Car c'est bien plus souvent du mésusage des mathématiques qu'il s'agit. « *Avec les sondages, on continue à commettre des erreurs mathématiques de base*, explique Jean Chiche, statisticien et chercheur au Centre de recherches politiques de Sciences Po (Cevipof). *Par exemple, considérer comme indépendantes des variables qui ne le sont pas: niveau d'études, profession, lieu d'habitation... On écrit "toutes choses égales par ailleurs", mais les choses ne sont bien souvent pas égales par ailleurs.* »

Devant l'autorité naturelle liée au chiffre et à sa prétendue objectivité et neutralité, il est néanmoins tentant de s'emparer de quelques formules magiques. Nous avons choisi trois domaines – connu pour l'un, largement ignorés pour les deux autres –, afin d'en mesurer les conséquences dramatiques.

Prévoir les risques financiers

« *Des professeurs de maths enseignent à leurs étudiants comment faire des coups boursiers. Ce qu'ils font relève, sans qu'ils le sachent, du crime contre l'humanité.* » Lorsque l'ancien premier ministre Michel Rocard lance cette accusation dans *Le Monde*, le 3 novembre 2008, beaucoup de mathématiciens sont choqués. Les vendeurs des produits financiers ayant conduit à la crise ne sont-ils pas plus responsables que les équations complexes utilisées pour en estimer le risque? Pourtant, la crise financière a réveillé les consciences et remis sur le tapis les questions de la responsabilité des scientifiques dans l'usage de leurs découvertes.

Ouvrons la boîte noire. A l'origine de la crise, il y a des produits financiers complexes comme les CDO (*collateralized debt obligations*), portefeuilles d'obligations d'entreprises et de prêts bancaires. Pour donner un prix à ce mélange hétérogène, les experts ont ressorti un vieux théorème disant que peu ou prou, on pouvait mêler simplement les différentes valeurs de risques de chacun des titres, faisant fi de certaines corrélations dévastatrices, en vertu desquelles le défaut sur un actif pouvait entraîner la chute de l'autre. Le tout adoubi par les agences de notation.

La suite est connue: peu chers, abondants et bien notés, les CDO ont eu du succès. Mais l'explosion de défaut de remboursement de crédits immobiliers les a fait soudain chuter; les prix calculés par les formules hasardeuses n'étant pas corrects, tout le monde a voulu s'en débarrasser, accentuant la chute.

Autre erreur commise: une mauvaise perception de l'aléa. Les banques doivent évaluer leur pire perte en cas d'accident « rare » (une chance sur 100 ou sur 1 000 par exemple). Et constituer des réserves, qui immobilisent des fonds. Des modèles ont été proposés, tenant compte des fluctuations des titres, mais escamotant parfois l'amplitude de ces variations. Comme si le constructeur d'une digue évaluait la fréquence d'une crue sans s'intéresser à la hauteur des vagues. Evidemment, cela a marché tant que les variations étaient peu importantes. Mais celles de 2007-2008, exceptionnelles, ont balayé toutes ces estimations.

« *Il fallait des réponses rapides et efficaces, on s'est donc borné à utiliser des modèles un peu frustes. Les responsables ont fait comme si les maths étaient solides alors qu'elles reposaient sur des hypothèses non vérifiées* », note Stéphan Cléménçon, directeur de la chaire « Apprentissage statistique » à Télécom ParisTech. Charles-Albert Lehalle, chercheur à Capital Fund Management, ancien responsable de la recherche dans la banque d'investissement du Crédit agricole, précise: « *Les bases mathématiques de ces modèles de risque étaient bonnes, c'est leur application dans des contextes incorrects qui a posé problème.* »

Libérer les détenus

Prévenir la récidive. Tous les pays occidentaux ont fait de cet objectif une priorité. Aux Etats-Unis, c'est même une obsession. Au point de mettre en doute la capacité de jugement des juges. Ainsi est né, dans les années 1960, le souci d'objectiver les risques de chaque condamné. Désormais, la plupart des Etats américains disposent de systèmes dits actuariels qui dressent un portrait du criminel. La nature des faits et leur nombre y côtoient l'âge, le sexe, les revenus, la situation de famille, le passé criminel des parents, le statut dans l'emploi... De vingt à quarante données sont ainsi enregistrées. L'ordinateur applique alors un modèle statistique dit de « *régression logistique* ». Et « score » le risque de récidive... « *Les juges s'en inspirent pour fixer une peine ou pour accorder une libération conditionnelle*, explique Sonja Starr, professeure de droit à l'université du Michigan, dans la revue *Law and Economics Working Papers*. *Mais cette décision est prise en fonction de l'histoire du criminel, pas de son histoire criminelle.* »

« *Ça donne un sentiment d'objectivité, mais en réalité, cela pose des problèmes innombrables, notamment éthiques* », insiste Virginie Gautron, maître de conférences à l'université de Nantes.

Ainsi, un condamné qui habite dans un quartier à fort niveau de criminalité est considéré comme plus susceptible de récidiver. Il en va de même si ses parents ont été condamnés. « *Pire, se désole la chercheuse: comme on sait que les délinquants sexuels ont souvent été d'abord abusés, avoir été victime devient un handicap dans le profil. C'est la double peine.* »

Les problèmes sont également d'ordre pratique, ajoute Seena Fazel, professeur de psychiatrie légale à l'université d'Oxford. Le psychiatre a passé en revue les nombreux algorithmes utilisés. « *Leurs résultats sont assez bons pour prédire la faible récidive. Mais très mauvais avec les prétendus individus dangereux: pas loin de 50 % d'erreurs.* »

Faut-il alors les écarter? « *Je l'ai cru mais, j'ai changé d'avis. Parce qu'aujourd'hui, nous n'avons plus les moyens budgétaires de bien évaluer tous les délinquants par entretien. Il faut toutefois réserver ce profilage aux cas où une erreur n'aura pas de conséquence grave pour celui qui la subit. Donc l'interdire pour l'attribution d'une peine ou d'une remise en liberté, mais l'autoriser pour déterminer l'ampleur d'un suivi psychiatrique, d'un pointage au commissariat ou d'une cure de désintoxication.* »

C'est la position qu'a retenue le syndicat des personnels de l'administration pénitentiaire, le Snepap-FSU. La France teste dans six départements, et avec la plus grande discrétion, des logiciels de « scoring » élaborés au Canada. Pour le syndicat, l'outil doit rester « *une béquille* » à l'usage des éducateurs, donc interdite aux juges. Sera-t-il entendu?

Lire aussi : « Des calculs de probabilité ont produit des erreurs judiciaires » (sciences

[/article/2015/09/14/des-calculs-de-probabilite-ont-produit-des-erreurs-judiciaires_4756815_1650684.html](http://article/2015/09/14/des-calculs-de-probabilite-ont-produit-des-erreurs-judiciaires_4756815_1650684.html))

Evaluer les profs

La place des mathématiques agite depuis longtemps le monde de l'école. Les littéraires dénoncent leur rôle dans la sélection des meilleurs élèves. Les physiciens, chimistes et biologistes leur reprochent de transformer leur matière en simple objet de problèmes à résoudre.

Aux Etats-Unis, un tout autre aspect de la science des nombres a créé le tumulte: l'évaluation des enseignants. Les parents réclamaient de « bons profs »? Les mathématiciens ont proposé leurs modèles dits « à valeur ajoutée ». Le principe en est assez simple, explique la mathématicienne Cathy O'Neil: « *Un élève est soumis chaque année à des tests normalisés. On analyse ses résultats précédents, le niveau de sa classe, sa situation familiale et sociale, etc. Et un logiciel très perfectionné, appuyé sur la méthode dite de régression quantile, lui prédit un score. Puis compare avec son résultat au test. Il ne reste plus qu'à rapprocher tous les élèves d'un même prof pour voir l'influence d'un enseignant.* »

Washington, New York, Los Angeles, Chicago... Une à une, les grandes villes américaines ont adopté le modèle dans les années 2000, pour attribuer des primes ou titulariser les meilleurs enseignants et sanctionner les moins bons. « *Sauf que ces tests sont biaisés et même inopérants* », poursuit Cathy O'Neil, qui, sur son blog, Mathbabe, chasse les abus de sa science préférée.

Son confrère John Ewing, président de l'association Math for America, a, de son côté, dressé une liste de dysfonctionnements. Il a notamment constaté qu'il n'y avait aucune corrélation entre les résultats d'un professeur avec une classe et avec une autre; pas davantage dans ses performances d'une année sur l'autre. Or, conséquence dommageable, devant l'importance de cette évaluation, bien des enseignants renoncent aux apprentissages non notés, ceux qui nourrissent la curiosité et le travail en groupe.

Conscients que « *les résultats aux tests ne traduisent pas l'accomplissement d'un élève* », pour reprendre la formule de John Ewing, les enseignants de Chicago ont fait grève en 2012 pour dénoncer le poids des modèles mathématiques. Sans grand résultat. Dans les pays anglo-saxons, l'évaluation par valeur ajoutée gagne du terrain. En France, il n'en est, pour l'heure, pas question.

Quels remèdes ?

Promis, juré, les nouvelles règles de la finance vont corriger les errements passés. Il est vrai que les garanties demandées aux banques ont été relevées et que des audits plus sérieux des modèles de risque sont diligentés. « *Mais la formation continue pêche toujours. Il serait bon aussi de favoriser la mixité entre les jeunes et les plus expérimentés. Trop souvent, les jeunes devenaient opérationnels sans vraiment connaître les produits sur lesquels ils travaillaient* », suggère le chercheur Charles-

Albert Lehalle.

« A la suite de la multiplication des problèmes liés aux statistiques dans des articles scientifiques, les revues ont monté leur niveau d'exigence pour la publication : augmentation de la taille des échantillons, seuil de significativité plus élevé, dépôt des données brutes et des logiciels utilisés pour répliquer l'expérience... », note aussi Jean-François Royer, de la Société française de statistique.

Leila Schneps voudrait aussi relever le niveau de l'expertise en justice et dans la police scientifique. « Il faudrait systématiquement et dès l'instruction faire appel à un statisticien. Et établir une liste de tests statistiques et de logiciels acceptables dans les enquêtes et les procès », affirme-t-elle. L'institut Newton, de Cambridge, consacra aux mathématiques en sciences criminelles un cycle d'un semestre, en 2016. « Avec une centaine de mathématiciens du monde entier », souligne Leila Schneps. Mais pour l'heure, aucun Français.

D'autres, enfin, font remarquer que les mathématiques peuvent être la solution aux problèmes qu'elles ont pu engendrer. En effet, analyser les pannes ou les crises demande aussi des outils complexes...

C'est loin d'être terminé. Car une vague nouvelle se lève, gourmande en maths : le big data. Ce terme générique désigne le traitement et l'analyse d'une grande masse d'informations (traces numériques laissées sur Internet, ensembles de gènes et de protéines, capteurs divers dans l'industrie ou la santé...). Or il n'est plus possible de travailler sur ces nouveaux objets avec de « vieux » outils. Les étudiants l'ont compris, qui plébiscitent les formations de « scientifiques des données » et leur large panel d'applications pour la sécurité (prévision de la délinquance, lutte antiterroriste), la santé et le bien-être (médicaments personnalisés, recherche de nouvelles molécules...), ou l'économie en général (assurances, commerce...). De plus en plus, des automates prendront des décisions ou anticiperont des phénomènes, sans qu'on sache vraiment ce que contiennent ces boîtes noires.

« Lors des dernières élections américaines, le marketing politique a été très loin, relève Cathy O'Neil. En disposant de l'adresse IP [numéro d'identification attribué à chaque appareil connecté à un réseau informatique] de chaque individu cible, vous pouvez adapter votre site à son profil... C'est l'asymétrie complète : le candidat sait tout de vous, vous ne savez rien de lui. Mathématiquement, c'est solide, mais pour la démocratie, c'est un danger. »

Les dérapages arriveront forcément. Bien entendu, les mathématiciens n'en seront pas les seuls responsables. Mais il leur sera impossible de faire comme si ces outils leur avaient juste échappé.

Lire aussi : Maths : quelques pièges à éviter ([/sciences/article/2015/09/14/maths-quelques-pieges-a-eviter_4756829_1650684.html](http://sciences/article/2015/09/14/maths-quelques-pieges-a-eviter_4756829_1650684.html))