

## DÉSORDRE MATHÉMATIQUE CONTRE DÉRAISON HUMAINE

Le déluge de sondages annonçant les élections nous rappelle que la France est le plus gros consommateur mondial d'enquêtes d'opinion, pour plusieurs milliards d'euros annuels. Si les français, en majorité, jugent les sondages indispensables, ils pensent aussi qu'ils sont commentés à tort et à travers, et influencent le choix des électeurs — tout en estimant, à une majorité encore plus forte, qu'eux-mêmes échappent à cette influence !

Derrière les sondages, il y a d'abord une base scientifique aussi simple que puissante : la prédictibilité du hasard. Comme l'a montré Pierre-Simon de Laplace dès 1800, les fluctuations de la moyenne d'un grand nombre d'observations statistiques indépendantes sont gouvernées par une loi universelle, la fameuse courbe des erreurs de Gauss, ou courbe en cloche. Adolphe Quételet et Francis Galton découvriront avec émerveillement que cette loi mathématique s'applique aussi aux sciences humaines; Boltzmann et Maxwell l'introduiront dans la physique des gaz. Cette "loi suprême du désordre et de la déraison" selon l'expression de Galton, permet de calculer aussi bien les fluctuations de la taille des individus ou des vitesses des molécules de l'air, que celles des sondages.

Pourtant les sondeurs doivent encore affronter trois obstacles considérables. Le premier est la taille de l'échantillon : l'erreur obéit à un schéma universel, mais elle décroît lentement, comme l'inverse d'une racine carrée ! Pour 1000 sondés, ne pas espérer une précision meilleure que 3%; et si l'on veut passer à 1%, c'est plutôt de l'ordre de 10 000 réponses qu'il faut obtenir, et pour cela interroger peut-être 100 000 personnes ! Alors contentons-nous de cette marge de l'ordre de 3%, tout en nous souvenant qu'une tendance à la hausse de 1% aura plus de chances d'être due au hasard qu'à une quelconque réalité.

Le deuxième obstacle est beaucoup plus traître : comment garantir que les sondés sont bien choisis au hasard ? Qu'on les contacte par téléphone, Internet ou toute autre méthode, le mode d'interrogation influera sur l'échantillon; il se peut même que ceux qui acceptent de répondre aient tendance à voter davantage pour tel candidat — comme lors du fameux sondage américain géant de 1936, qui malgré plusieurs millions de participants s'avéra incapable de prédire l'élection de Franklin Roosevelt. La méthode des quotas contourne ce problème en cherchant à construire un modèle réduit de la population au regard de certains critères; mais le choix des critères est sujet à caution, et l'analyse mathématique de la méthode est délicate et incertaine.

Pourtant le pire est encore à venir : les fausses réponses des sondés ! Erreurs ou mensonges, conscients ou inconscients, par exemple pour un vote perçu comme naïf ou extrémiste. Ces facteurs peuvent ruiner un sondage, comme lors des élections présidentielles en France en 2002, ou en Côte d'Ivoire en 2010. Comment extrapoler, deviner les intentions qui se cachent derrière les déclarations, en tenant compte du profil du vote ou de l'actualité ? Que conserver des statistiques antérieures, quelles questions poser ? Autant d'interrogations qui mettent à contribution les sciences "dures" comme les sciences "humaines", ici inextricablement mêlées.

Cédric Villani, Professeur de l'Université de Lyon, Directeur de l'Institut Henri Poincaré (CNRS/UPMC)  
— Carte blanche du supplément *Sciences & Technologie* du Monde, 13 avril 2012