

Pour une juste interprétation des données économiques

par Jean-Marie Gogue

Résumé

L'interprétation des données numériques donne lieu à des erreurs systématiques dont le poids sur l'économie est considérable. L'auteur propose de tirer les conséquences de la théorie des probabilités et des méthodes statistiques, en restant à un niveau élémentaire, afin de réduire les risques d'erreur, tant dans la prévision que dans la décision.

Une première étape consiste à distinguer les deux types d'erreur qui sont indissolublement liés à la méthode expérimentale. Il faut comprendre que l'opération fondamentale de l'interprétation des chiffres est la comparaison. La source de toutes les erreurs réside dans cette opération. Certaines méthodes statistiques ont pour but de donner aux résultats des comparaisons le maximum de pertinence, donc de minimiser le risque d'erreur. Il s'agit là d'erreurs au sens strict. L'auteur aborde ensuite la question des pseudo-erreurs, c'est-à-dire des conclusions fondées, pour une raison ou une autre, sur des interprétations volontairement fausses.

Au même titre que l'informatique, les méthodes statistiques d'interprétation des données devraient quitter le domaine spécialisé dans lequel elles ont pris naissance pour s'étendre à un large public. Cette démocratisation mettrait sans doute un terme à l'usage abusif des statistiques en faveur de propositions partisans.

o o o

Dans un pays moderne, de nombreux aspects de la vie quotidienne sont déterminés par l'interprétation de quelques données numériques auxquelles on attribue une importance particulière en certaines circonstances. C'est le cas notamment des indicateurs économiques qui alimentent les débats politiques, des résultats financiers qui traduisent les performances d'une entreprise, ou encore des notes qui servent à évaluer les connaissances d'un élève sur son bulletin scolaire. Souvent sujets à discussion, ces chiffres servent, d'une part, à faire des prévisions que l'avenir dira justes ou fausses, d'autre part, à prendre des décisions qui aboutiront à des résultats plus ou moins satisfaisants. Par exemple si un élève de terminale obtient de bonnes notes tout au long de l'année, il est facile de prévoir qu'il aura son bac, sauf accident. De même, on peut prévoir qu'un commerçant fera de bonnes affaires pendant les mois qui viennent s'il a fait de bonnes affaires au cours des mois précédents. Quand les chiffres obtenus sont réguliers, les prévisions qui en découlent ont de fortes chances d'être vérifiées. Dans le cas contraire, elles sont aléatoires.

Nous sommes tentés de croire qu'il existe une seule manière d'interpréter correctement les données numériques : une manière rationnelle, confirmée par les experts, et qui rejoint le sens commun. Pourtant, les chiffres font souvent l'objet de longues discussions qui n'aboutissent pas toujours à des décisions pertinentes, notamment dans les entreprises ; j'en fus jadis témoin. C'est faute de savoir interpréter correctement les chiffres de leur tableau de bord que de grandes entreprises françaises, et non des moindres, sont incapables de fixer une ligne stratégique claire et cohérente. Alcatel, qui fut le leader français des télécommunications pendant toute la seconde moitié du vingtième siècle, en est un bon exemple. Au terme d'un parcours sinueux, jalonné de restructurations, cette société s'est alliée au leader américain Lucent en 2006 afin de tirer un trait sur le passé et de mettre son avenir en des mains plus sûres. On se souvient qu'Alcatel était le pôle « courant faible » de la Compagnie Générale d'Électricité dans les années 60, avant d'être nationalisée en 1981, d'absorber Thomson Téléphone en 1986, puis d'être privatisée en 1993. Les politiques ont donc joué un rôle non négligeable dans le management calamiteux de cette grande entreprise ; il n'en reste pas moins que tous les hommes qui se sont succédé à sa tête ont fait preuve d'une affligeante myopie. La plupart de leurs prévisions à long terme étaient fausses. On observe la même situation dans un grand nombre de petites entreprises auxquelles il arrive parfois de faire faillite et de mettre des gens au chômage. Il en résulte finalement, au niveau national, une perte économique considérable, difficile à évaluer. Certes, je reconnais qu'il est difficile de fournir la preuve qu'un échec est imputable à une erreur de décision ; les responsables des faillites d'entreprises, et plus généralement les responsables des échecs commerciaux, industriels ou économiques, invoquent toujours de bonnes raisons pour leur défense. Mais je montrerai ici que beaucoup de décisions inopportunes sont fondées sur une

mauvaise interprétation des chiffres disponibles, ce qui est une preuve indirecte du lien de causalité entre l'échec et le mode de décision. Il est donc d'une importance capitale que les nouvelles générations parviennent à un consensus sur des principes généraux d'interprétation des chiffres, afin de réduire les risques d'erreur de prévision et de décision dans la limite du possible. Nous en sommes encore loin.

Pour commencer, il faut distinguer deux types d'erreurs dans l'interprétation des chiffres. L'erreur du premier type consiste à considérer une donnée nouvelle comme une information assez importante pour justifier une action rapide, alors qu'elle est véritablement insignifiante : une non-information pourrait-on dire. L'erreur du second type, c'est l'inverse. Elle consiste à ne pas comprendre l'importance d'une donnée nouvelle, alors qu'elle annonce, en réalité, l'arrivée d'un problème ou d'une chance à saisir. Partant de cette définition, on comprend que les deux types d'erreurs sont inévitables, à des degrés divers, car comment évaluer l'importance d'une donnée nouvelle ?

L'erreur du premier type provient d'une sorte de nervosité permanente provoquée par un sentiment d'impatience et de malaise quand on reste dans l'inaction. À chaque instant, il doit se passer quelque chose. C'est de cette manière que l'on voit se comporter de nombreux journalistes, souvent par déformation professionnelle. S'ils ne montaient pas en épingle quelques non-événements piochés chaque jour parmi les faits divers, ils pourraient craindre de ne pas avoir assez de choses à dire. L'obstination avec laquelle certains éditorialistes commentent sur les chaînes de radio les moindres variations des cours de la bourse, heure par heure, frise même le ridicule. Dans l'industrie, le besoin instinctif d'agir en toute circonstance fait des ravages. Voici quelques années, le directeur d'un site de production qui comptait un millier de personnes dans l'est de la France assistait tous les matins à la réunion des contremaîtres. On y examinait avec un soin d'apothicaire tous les problèmes rencontrés la veille, afin de trouver des solutions. On y faisait un suivi méthodique des actions correctives. Malgré tous ces efforts, le niveau des problèmes ne baissait pas, car les causes présumées n'étaient pas les bonnes. Les problèmes étaient inhérents au système. Le plus triste dans cette histoire, c'est que les prétendues actions correctives se soldaient par une activité parasite qui consommait la totalité des bénéfices. Mais il peut arriver aussi qu'une erreur du premier type soit commise délibérément, par intérêt personnel. Par exemple, pour se débarrasser d'un collaborateur qui, pour une raison quelconque - il s'est peut-être montré trop arrogant - est tombé en disgrâce, un patron prendra comme prétexte une baisse de son résultat trimestriel, alors qu'il sait parfaitement que cette baisse n'est pas significative. « Je regrette, mais tu n'as pas tenu ton objectif. Va chercher du travail ailleurs. » Le même argument sera utilisé pour fermer une usine. On peut estimer que le quart des fermetures d'usines en France, au cours des dix dernières années, n'avait aucune justification économique.

L'erreur du second type provient au contraire d'une attitude qui semble pécher par excès de patience. Quand on ne veut pas reconnaître l'importance d'un chiffre, c'est souvent par crainte des conséquences de l'action qui devrait en résulter ; ou bien c'est pour retarder cette action. Par exemple, dans une société de service, un ordinateur donne des signes de faiblesse. Le directeur ne veut pas le remplacer immédiatement par un matériel plus moderne car c'est un investissement qui n'est pas prévu au budget, et le travail de bureau risquerait d'être perturbé pendant quelques jours, en pleine période des fêtes. Mauvais calcul : le matériel tombe en panne, plusieurs fichiers sont détériorés, et la société de service perd beaucoup plus de temps et d'argent que si le directeur avait tenu compte des chiffres annonciateurs de la panne. Le crash du Concorde à Gonesse en juillet 2000 est un exemple particulièrement dramatique d'erreur du second type. En analysant les causes de l'accident, on voit bien que la cause prédominante n'est pas l'objet métallique perdu sur la piste quelques minutes auparavant par un avion de la Continental Airlines, mais la vulnérabilité du Concorde roulant au décollage sur ce type d'objet. Plusieurs incidents survenus au cours des années précédentes avaient mis le fait en évidence, mais le constructeur n'en avait pas tiré les conséquences, peut-être pour des raisons politiques. C'est une situation qui se présente aujourd'hui trop souvent en économie, voire en écologie, à une échelle incomparablement plus grande. Elle a inspiré à un humoriste anonyme la courte fable de la grenouille dans l'eau chaude. « Mettez une grenouille dans une casserole d'eau froide et chauffez lentement. Au début, la grenouille est contente ; puis, bien au chaud, elle s'endort. À la fin, elle meurt ébouillantée. »

Dans les exemples que nous avons donnés, s'agit-il d'erreurs au sens strict, ou de pseudo-erreurs ? Lorsqu'une interprétation abusive des chiffres est instrumentalisée par les passions, les intérêts ou ce qu'on voudra, l'erreur est consciente ; son but est de tromper. C'est pour la distinguer d'une erreur involontaire, imputable à la méthode d'analyse, que nous la nommerons *pseudo-erreur*. Le directeur d'usine obsédé par les actions correctives commet une erreur au sens strict ; le patron qui licencie un collaborateur qu'il n'aime pas sous prétexte de mauvais résultats commet une pseudo-erreur. Le cas du Concorde est

ambigu. La méthode d'analyse n'était peut-être pas bonne, mais les ingénieurs avaient peut-être un secret désir de se tromper eux-mêmes. La plupart des exemples qui vont suivre comportent un mélange d'erreurs et de pseudo-erreurs.

Quand des chiffres font l'objet d'un débat, notamment dans la vie politique, comment ne pas tomber dans l'une ou l'autre de ces deux erreurs d'interprétation ? Les intervenants se rangent habituellement en deux camps, que nous pourrions nommer les « progressistes » et les « conservateurs ». Mais leur attitude n'est pas la même en toute circonstance. Tel progressiste sur des questions financières sera conservateur sur des questions de sécurité, ou inversement. En les écoutant, on comprend bien que l'attitude de chacun est dictée principalement par son intérêt personnel et celui de son clan. C'est pourquoi les débats politiques, quelle que soit la qualité des intervenants, échappent généralement à toute logique ; ils sont déterminés essentiellement par des rapports de forces, et l'intérêt commun y trouve rarement son compte. Les pseudo-erreurs dominent le débat, avec en plus la caution morale des statistiques. Dans une entreprise, dont le but est l'efficacité économique, les erreurs d'interprétation proviennent plutôt de la conjoncture et du caractère du dirigeant. Si l'argent manque, la tendance sera de reporter les investissements à plus tard : risque d'erreur du premier type. Si le dirigeant veut se montrer énergique, la tendance sera de modifier constamment la structure et de pousser le personnel à plus de mobilité : risque d'erreur du second type, d'autant plus important que, dans le monde du management, la mode actuelle est à la « réactivité ». Les effets de ces erreurs s'ajoutent pour produire un énorme gaspillage, à la fois financier et humain. Ce n'est pas en corrigeant les tendances des uns et des autres qu'on peut le combattre, car on voit bien que, dans une situation donnée, le nombre d'erreurs du second type augmentera si l'on fait en sorte de diminuer le nombre d'erreurs du premier type, et inversement. Il est impossible d'éliminer toutes les erreurs ; le seul moyen d'en réduire le taux est de fixer des règles de traitement des données, valables pour tous, qui conduisent à un équilibre optimum entre les deux types d'erreurs.

Dans l'interprétation des chiffres, l'opération fondamentale, beaucoup plus importante que les quatre opérations élémentaires, c'est la *comparaison*. « Les impôts ont encore augmenté ; mes notes sont inférieures à la moyenne ; etc. » On ne peut tirer aucune conclusion à partir d'un chiffre sans le comparer, consciemment ou pas, à d'autres chiffres ; c'est une évidence. Dans cet ordre d'idées, il faut distinguer deux sortes de comparaisons :

- comparaison dans le temps pour une même entité (une personne, un groupe, etc.)
- comparaison entre plusieurs entités pendant la même période (mois, année, etc.)

Dans le premier cas (comparaison dans le temps), le problème est d'apprécier la stabilité d'une performance ou d'une donnée et, le cas échéant, d'en connaître le sens de variation. Dans le second cas (comparaison dans l'espace), il s'agit d'établir un classement. En politique, ce sont des questions qui reviennent constamment : on nous parle de la diminution du nombre d'accidents de la route, des différences de taux d'endettement entre les États européens, etc.

Chacun comprend intuitivement que le résultat de cette opération est d'autant plus juste qu'elle met en jeu un plus grand nombre de données numériques. Par exemple, si nous apprenons que l'indice Dow Jones de l'acier est passé de 331 en janvier à 332 en février, rien ne nous permet de dire que la tendance est à la hausse. Nous ne pourrions en juger que si nous voyons un graphique représentant les variations de l'indice sur plusieurs mois. De même, si un joueur de golf marque 80 un jour et 72 le lendemain, il ne faut pas en déduire qu'il a fait des progrès. On ne le saura qu'après avoir observé tous ses scores sur une période suffisamment longue. De même, si nous regardons les notes obtenues par deux élèves d'une même classe à un contrôle de français, 12 pour l'un et 8 pour l'autre, il ne faut pas en déduire aussitôt que l'un est meilleur que l'autre, car les résultats du contrôle suivant peuvent être inversés. Ce n'est qu'à la fin de l'année scolaire qu'on peut porter un tel jugement.

Poussons plus loin le raisonnement. Quand un jugement est porté à la suite d'une comparaison entre deux ensembles de données numériques, c'est toujours avec un risque d'erreur, mais un risque d'autant plus faible que les données sont nombreuses ; l'erreur zéro n'existe pas. Le hasard intervient en partie dans toute interprétation. Les mathématiques permettent même de calculer le risque d'erreur en fonction du nombre de données ; les instituts de sondages connaissent parfaitement cela. Il faudrait donc édicter une première règle, assez simple pour être comprise par tout le monde, qui servirait à déterminer le nombre de données souhaitable pour une telle opération. Une deuxième règle servirait ensuite à donner une réponse positive ou négative à la question : « La différence entre ces deux ensembles de données numériques est-elle assez importante pour justifier une action ? »

Le problème n'est pas de savoir s'il y a une différence, car on ne trouve jamais deux ensembles identiques. Il y a toujours une différence ; le problème est d'en connaître l'importance. À cet égard, les mathématiques ont fait un bond considérable au siècle dernier avec l'introduction du concept de variance. Utilisée couramment dans les calculs statistiques au niveau du lycée, la variance est un paramètre de dispersion associé à la moyenne. Sa racine carrée est l'écart-type, représenté par la lettre grecque sigma. Supposons par exemple qu'un directeur commercial, dans une entreprise, cherche à savoir si le dernier chiffre mensuel des ventes est en hausse significative par rapport à la moyenne des 12 mois précédents (il s'agit cette fois de comparer un chiffre isolé à un ensemble de 12 données). Supposons que la hausse soit de 11 %. On pourrait fixer arbitrairement le seuil de signification à 10 %, et dans ce cas la réponse serait positive ; ou bien à 12 %, et la réponse serait négative. On comprend donc bien la faiblesse de la méthode qui consiste à prendre pour critère un pourcentage donné, comme le célèbre 3 % de Maastricht. C'est une méthode qu'il faut abandonner. L'écart-type permet au contraire de fixer un seuil parfaitement logique, adapté à la situation, pouvant varier au cours du temps. Dans l'exemple précédent, le directeur commercial l'obtiendra facilement en entrant les résultats des 12 mois précédents sur une calculatrice de poche et en tapant sur la touche de l'écart-type. Le seuil de signification sera placé à trois écarts-type au-dessus de la moyenne. On peut dire en quelque sorte qu'il n'est pas soumis à la volonté de quelqu'un, mais imposé par le système lui-même. Si cette méthode avait été universellement adoptée lorsque fut signé le protocole de Kyoto, elle aurait certainement permis d'obtenir un plus large consensus sur l'importance du dérèglement climatique. Il y a plus de dix ans, l'augmentation de l'effet de serre était un phénomène avéré, mais les experts n'étaient pas tous d'accord pour en attribuer la cause à l'émission de gaz carbonique et de méthane, ce que l'étude des variations, graphiques à l'appui, permet de démontrer formellement. L'administration américaine, hostile au protocole, n'a pas réussi à mettre la méthode en défaut.

Nous avons donné ici un bref aperçu d'une méthode susceptible de nombreuses applications en économie, dans les entreprises et dans la vie courante. Une autre méthode, dérivée de la même théorie statistique, permet de savoir si dans un ensemble formé d'éléments tels que des personnes, des villes, des régions, des écoles, des hôpitaux ou des entreprises, un ou plusieurs individus se distinguent de l'ensemble par des performances sortant de l'ordinaire. C'est l'analyse de variance. Elle consiste, une fois encore, à comparer chaque performance individuelle à la moyenne des autres performances en prenant l'écart-type comme critère. Elle s'oppose à une pratique courante qui consiste à classer systématiquement les individus pour exhorter les derniers à faire des efforts afin de regagner des places dans le classement. C'est une pratique irrationnelle, injuste, profondément nuisible, ayant pour seul effet d'entretenir le désordre et la confusion, car chacun sait que le hasard intervient toujours avec une intensité plus ou moins grande dans l'attribution des performances. Le ministère de l'Éducation Nationale, qui publie chaque année le classement des lycées par ordre de performances afin, soi-disant, d'améliorer le système, est l'un des champions français de cette pratique, importée de Grande-Bretagne. Au contraire, l'analyse de variance présente l'immense avantage de réduire l'effet du hasard en mettant tous les individus au même niveau, sauf quelques-uns dont les performances traduisent une différence significative, en plus ou en moins, par rapport à l'ensemble. Les efforts d'amélioration gagnent en efficacité lorsque toute l'attention est portée sur les causes possibles de ces quelques différences. Cette méthode mériterait donc d'être généralisée dans l'enseignement, la recherche et l'industrie.

L'abondance des informations à base de statistiques qui se trouvent dans la presse audiovisuelle et écrite pourrait faire croire que le public est entraîné à l'interprétation des données économiques, d'autant plus que l'Éducation Nationale a introduit des chapitres de probabilités et de statistiques dans les programmes de mathématiques à partir de la classe de seconde. Mais beaucoup croient encore que les méthodes statistiques sont réservées à quelques spécialistes, tandis que des experts au service des observatoires économiques et des instituts de sondages utilisent sans vergogne les statistiques comme un moyen d'impressionner le public. « Les prêtres ne sont pas ce qu'un vain peuple pense. Notre crédulité fait toute leur science », disait Voltaire. Au même titre que l'informatique, les méthodes statistiques peuvent être utiles à chacun de nous, en de nombreux domaines... à condition de savoir s'en servir.

Août 2007