

Sondages, statistiques : l'information manipulée

A partir de deux termes cousins pour exprimer des méthodes de recherche et de compilation d'informations sont appliqués des traitements qui vont permettre d'exprimer une opinion pour le sondage, un enseignement pour la statistique.

Exploités à trop forte doses les sondages et les statistiques peuvent avoir des effets pervers de manipulation : « *Les médias ne disent pas aux gens ce qu'ils doivent penser, mais à quoi ils doivent penser* ».

Les définitions :

Il n'y a pas de définition officielle du sondage.

- **Sondage** : réalisé dans les règles de l'art le sondage est un outil statistique permettant d'évaluer en proportions les caractéristiques ou les préférences (et donc les opinions et les pratiques) d'une population donnée à partir de l'étude d'un échantillon de population.

Le terme « sondage » inventé par Jean Stoetze désigne deux termes anglo-saxons à la fois : les enquêtes de presse qui imitent le référendum (les polls) et les enquêtes de type universitaires sur les « attitudes » (les surveys)

« Recettes » de fabrication d'un sondage : Les échantillons sont constitués selon la méthode des quotas. Elle consiste à respecter, au sein de la fraction de la population interrogée, les proportions de la population que l'on souhaite étudier sur un certain nombre de critères particulièrement caractéristiques : le sexe, l'âge, la catégorie socioprofessionnelle et le lieu de résidence, auxquels certains instituts ajoutent le niveau de diplôme.

Les enquêtes d'intentions de vote sont des sondages visant à quantifier non pas des opinions mais une intention d'agir, un projet non définitif jusqu'au jour de l'élection: la certitude d'aller voter (se déplacer au bureau de vote) et le choix d'un candidat (le bulletin déposé dans l'urne).

Redressement : Les méthodes de redressement sont perçues, par les professionnels du secteur, comme des [http://junon.univ-cezanne.fr/u3iredic/?p=8813 - _ftn1« recettes de fabrication »](http://junon.univ-cezanne.fr/u3iredic/?p=8813_-_ftn1«_recettes_de_fabrication_»), qui doivent être tenues confidentielles. Elles sont couvertes par le secret industriel.

- **Statistique** : Étude méthodique des faits économiques et sociaux par des classements, des inventaires chiffrés, des recensements, etc. *La statistique apprendra (...) que la population s'est accrue, que le prix des denrées a baissé ou augmenté.* (Cournot, *Fond. conaiss.*, 1851, p. 497).

Si elle est descriptive : la statistique est un ensemble des méthodes utilisables pour mettre en valeur les caractéristiques extérieures d'une série de chiffres. (Combe 1971)

C'est plus aujourd'hui l'ensemble des méthodes qui ont pour objet la collecte, le traitement et l'interprétation de ces données.

La statistique est la science dont l'objet est de recueillir, de traiter et d'analyser des données issues de l'observation de phénomènes aléatoires, c'est-à-dire dans lesquels le hasard intervient. L'analyse des données est utilisée pour d'écrire les phénomènes étudiés, faire des prévisions et prendre des décisions à leur sujet. En cela, la statistique est un outil essentiel pour la compréhension et la gestion des phénomènes complexes.

Origines.

Pour les sondages : La quantification des faits sociaux a de tout temps intéressé les pouvoirs centraux. Au Moyen-Âge, des recensements de soldats et d'habitants étaient effectués afin de lever des armées et de collecter les impôts. L'administration de Louis XV réalise en 1745 ce qui est considéré comme la première forme des grandes enquêtes d'opinion. Dans ses directives aux grands intendants de province, le contrôleur général explique qu'il s'agit d'établir un état de la richesse et de la pauvreté des peuples des provinces, de dénombrer les garçons sujets à la milice et capables de porter les armes, de dresser l'inventaire des ressources de chaque ville, et surtout, ce qui est nouveau, de « Semer les bruits dans les villes franches de votre département d'une augmentation d'un tiers sur le droit des entrées, et de la levée d'une future milice de deux hommes dans chaque paroisse. Vous recueillerez avec soin ce qu'en diront les habitants et vous en ferez mention dans l'état que le roi vous demande. » Si la propagation de rumeurs n'est pas un instrument nouveau, la consignation systématique des effets de leur diffusion par un réseau d'observateurs disséminés dans tout le pays est en revanche inédite.

L'arrivée des sondages en France doit beaucoup à la fascination et à l'importation des méthodes américaines, avant et surtout après la deuxième guerre mondiale. La situation sociale et politique de la cinquième république, qui rend possible le « face-à-face » d'un président avec le peuple, y sont également pour beaucoup.

Pour les statistiques : la création du terme « statistique » est attribué à un professeur de Göttingen, G. Achenwall, qui aurait créé en 1746 le mot *Statistik*. En fait, l'activité correspondante de recueil de

données remonte à la lointaine antiquité. On cite, d'une part, l'empereur chinois Yao, organisant le recensement des productions agricoles en 2238 avant J.-C. et l'institution du [cadastre](#) et du cens (redevance) chez les Égyptiens, en 1700 avant J.-C. L'importance sociale de la statistique était reconnue, puisqu'il advint que le pharaon Amasis édicta la peine de mort contre ceux qui refusaient de déclarer leurs nom, profession et moyens de subsistance.

Plus loin encore, à l'origine du stockage de l'information, on a trouvé depuis une centaine d'années un certain nombre d'objets gravés, principalement sur des os ou des bois de rènes au Paléolithique supérieur (environ -35 000 ans en Europe et -60 000 ans en Afrique).

L'établissement du cadastre (cadastre de Rome 578 AC) est dans un premier temps une des pierres sur laquelle a été construite la statistique et ses outils. Elle est passée d'un instrument essentiellement destiné à dénombrer les richesses, les sujets et à prélever des impôts au statut d'instrument prévisionnel, on peut dire qu'elle a outrepassé les statistiques pour atteindre à « la statistique ». On peut penser sur la base de certaines pratiques que ces outils, qui appartiennent assurément à la statistique descriptive, ont une certaine fonction prévisionnelle et annoncent, de ce fait, une information comportant une conjecture sur le futur une espèce de « statistique active ».

Techniques

Pour les sondages :

1. **L'échantillonnage** : La **méthode aléatoire** consiste à effectuer un tirage au hasard de l'échantillon parmi la population, afin de n'introduire aucun biais (sous- ou surreprésentation de catégories) par rapport à la population réelle (à partir d'un fichier complet de population type INSEE) ; La **méthode des quotas** est la plus employée par les instituts de sondage. Il s'agit de reconstituer une population en miniature, c'est-à-dire de construire un échantillon dans lequel les individus sont répartis selon les mêmes proportions que dans le groupe à étudier.
2. **Les incertitudes liées à l'élaboration du questionnaire** : Une des principales incertitudes quant à la validité scientifique des sondages provient de la façon dont est élaboré le questionnaire. L'intention de celui qui pose la question (ou la fait poser par un institut de sondage) influe sur la méthode utilisée pour interroger des gens ainsi que sur la signification des résultats.
3. **Les méthodes d'enquête** : Les enquêteurs se heurtent à de nombreuses difficultés qui peuvent induire des biais dans la composition de l'échantillon ainsi que dans les réponses. Selon l'heure à laquelle s'effectue l'enquête, on a beaucoup plus de chance de trouver telle ou telle catégorie socio-professionnelle ensuite la façon dont les enquêtés perçoivent les questions qu'on leur pose ainsi que la perception de l'enquêteur lui-même peuvent influencer les réponses. L'enquêté n'est pas une machine à répondre mais une personne prise dans un contexte social, réagissant non seulement à des questions mais à un enquêteur et à l'image qu'il en a, et capable de stratégies.
4. **Le traitement des données** : La méthode du redressement recouvre deux pratiques de traitement des données, selon que l'on souhaite corriger un biais structurel de l'échantillon ou un biais observé a posteriori dans les réponses (typique des sondages politiques). Le traitement des données c'est la porte ouverte à tous les excès.
Si « un bon sondage est un sondage bien corrigé », dans quelle mesure cela permet-il « non de figer des choix mais au contraire de suivre les évolutions » de l'opinion ?

Pour les statistiques :

Les techniques de collecte des statistiques et de leur traitement ont considérablement évolué avec le temps, la circulation rapide de l'information et les outils mathématiques et informatiques.

L'informatique et la statistique sont deux éléments du traitement de l'information : l'informatique acquiert et traite l'information tandis que la statistique l'analyse. Les deux disciplines sont donc étroitement liées. En particulier, l'augmentation considérable de la puissance des ordinateurs et la facilité de transmission des données par internet ont rendu possible l'analyse de très grandes masses de données, ce qui nécessite l'utilisation de méthodes de plus en plus sophistiquées, connues sous le nom de data mining ou fouille de données. Enfin, l'informatique décisionnelle ou business intelligence regroupe les outils d'aide à la décision devenus essentiels dans la gestion des entreprises. Ces outils nécessitent un recours important aux méthodes statistiques.

Manipulation de l'information

Pour les sondages : Un rapide coup d'œil sur la fiche de l'enquête IFOP Ifop-JDD (8 janvier 2010) concernant les Français et les retraites aurait permis pourtant de relever la non représentativité de l'échantillon et d'économiser quelques verdicts infondés : deux questions ont été soumises à une partie de l'échantillon initial, à savoir 522 personnes seulement. La ventilation par âge des réponses à ces deux mêmes questions recèle par ailleurs une erreur grossière (56% des répondants ayant plus de 65 ans pensent

qu'ils seront contraints de partir à la retraite entre...60 et 64 ans, et 18% des mêmes déclarent être prêts à travailler jusqu'à...64 ans pour toucher une bonne retraite). Cela contribue forcément à accroître les résultats favorables à un allongement de l'âge de la retraite.

1. L'exemple de la profusion des sondages politiques en période préélectorale (jusqu'à la saturation) paraît contre productif de la fonction du sondage. Et finalement semble montrer que les intentions de vote ne sont pas si fiables. De plus en comparant les résultats des sondages publiés avant une élection aux résultats de la dite élection on fausse la véritable évolution du vote des électeurs.
2. Février 2014 « Un sondage BVA pour i-Télé et Le Parisien dit que les "français" sont prêts pour une diminution des prestations et allocation dues au chômeurs, et évidemment leur radiation en cas de refus de 2 offres d'emplois ? » Mais ces questions sont posées à des personnes qui ne sont pas au chômage ... !?
3. « Touche pas à mon église » *C'est ce que répondent (très) massivement les Français à Dalil Boubakeur. Selon notre sondage exclusif Ifop-Valeurs actuelles, pas moins de 67% d'entre eux près de sept sur dix ! - répondent, en effet, non à sa proposition de "recourir aux églises vides ou abandonnées pour accueillir les fidèles musulmans" "* (Valeurs Actuelles, 8 juillet 2015). Quelle était la question ? Celle-ci était présentée par une explication : *Pour répondre à la pénurie de lieux de culte musulmans en France, Dalil Boubakeur, le président du Conseil français du culte musulman, a proposé de recourir aux églises vides ou abandonnées pour accueillir les fidèles musulmans".* La question proprement dite suivait : *"Vous personnellement, seriez-vous favorable ou opposé à ce que des églises vides ou abandonnées soient transformées en mosquées ?"*. En somme, la réponse était dans la question. Le sondeur a pris bien soin d'annoncer que c'était une proposition musulmane, d'ailleurs contestée, de manière à obtenir une large majorité contre cette proposition.
4. Ou encore : "Etes vous favorables au démantèlement des camps illégaux de Roms ?" Réponse : oui puisqu'ils sont illégaux.

Pour les statistiques :

Quatre grands types de manipulation sont usités dans la production de statistiques publiques :

1. La première méthode consiste à «décrire un phénomène multidimensionnel à l'aide d'un indicateur unique». On la retrouve pour la mesure du pouvoir d'achat, qui est souvent basée uniquement sur l'indice des prix alors qu'il faut également prendre en compte le revenu.
2. L'exploitation des moyennes, qui se prêtent particulièrement aux raisonnements fallacieux dès lors qu'elles résument «une variable dont les valeurs sont très dispersées.»
3. Autre moyen de jouer avec les statistiques est «le changement de catégorisation des faits sociaux et de leur mode d'enregistrement.» Pour faire baisser le nombre de chômeurs comptabilisés, un moyen simple et efficace consiste à ne plus considérer comme chômeurs tout un ensemble de personnes qui l'étaient auparavant.
4. L'entretien d'une confusion entre la mesure de l'efficacité d'un service et la mesure de l'évolution d'un phénomène social dont il s'occupe est une autre façon de fausser les résultats d'une politique.

Enfin les modes de représentation des résultats tant d'un sondage que de séries statistiques auront de fortes influences sur la compréhension des phénomènes communiqués.

Quelques exemples

Pour les sondages :

- Est-ce à l'Elysée de financer des enquêtes sur la grossesse de Rachida Dati ou la rumeur de mariage entre Nicolas Sarkozy et Carla Bruni ? De commander des études sur les têtes de liste en Ile-de-France aux régionales de 2010, sur les candidats probables du PS en 2012 ? Pendant cinq ans, les moindres faits et gestes de Ségolène Royal et de Dominique Strauss-Kahn ont été observés à la loupe, décortiqués. François Hollande, pour sa part, a très peu intéressé l'Elysée. *"L'UMP aurait dû prendre à sa charge les sondages relatifs au PS, au FN ou aux écologistes, argumente M^e Karsenti. Le contribuable n'a pas à assumer [...] des sondages [...] partisan[s]."*
- Les méthodes utilisées par l'IFOP ont été dénoncées dans le scandale du faux sondage du JDD ([lien](#)), un sondage bidonné commandé par Publicis (la société de Maurice Lévy, 3^e patron le mieux payé de France), publié par le JDD (journal d'Arnaud Lagardère, proche de Nicolas Sarkozy) et réalisé par l'IFOP (appartenant à Laurence Parizot, présidente du MEDEF).
- Intentions de vote pour les présidentielles de 2017 [sondage IFOP](#). Certains prétendent que valoriser Le FN pour créer la peur serait un artifice pour décourager les électeurs à voter FN (hypothèse dangereuse).

Pour les statistiques :

- Paradoxe : Il y a plus d'arbres qu'observé précédemment mais quand même une forte diminution est constatée (50%) !!???.
Jusqu'à présent, les données satellitaires permettaient d'estimer à 400,25 milliards le nombre d'arbres sur Terre. C'est très en dessous des nouvelles évaluations qui tablent sur plus de 3 040 milliards d'arbres ! Soit environ 422 arbres par personne (7,2 milliards d'habitants en 2014). Ce nouveau recensement a été mené par une équipe dirigée par Thomas Crowther à l'université de Yale (États Unis). Malheureusement, ce nombre est en forte diminution à cause de la déforestation qui se poursuit. Les auteurs de l'étude rappellent qu'environ 15,3 milliards d'arbres sont coupés chaque année, particulièrement dans les zones tropicales, leur berceau. Ainsi, depuis les débuts de l'agriculture, il y a 12 000 ans environ, le nombre d'arbres sur Terre a chuté de 46 % ! Et avec eux, toute une biodiversité, des espèces animales et végétales...
"Nous avons presque réduit par deux le nombre d'arbres sur la planète, ce qui a affecté le climat et la santé. Cette étude souligne à quel point d'autres efforts sont nécessaires si nous voulons restaurer des forêts saines à travers le monde", explique Thomas Crowther.
- Publié dimanche 4 septembre 2015 : « Selon les projections de la Banque Mondiale, quelque 702 millions de personnes, soit 9,6% de la population mondiale, vivent sous le seuil de pauvreté, que l'institution a d'ailleurs relevé de 1,25 à 1,90 dollar par jour pour tenir notamment compte de l'inflation. En 2012, date des données disponibles les plus récentes, les plus défavorisés de la planète étaient 902 millions, soit près de 13% de la population mondiale, une proportion qui atteignait encore 29% en 1999.
- Dans le calcul de l'indice des prix à la consommation en France l'alimentaire et les boissons non alcoolisées ne rentrent que pour 15% dans la composition du panier, l'éducation pour 0% et la santé pour 10% ... Une analyse fine du contenu du panier pourrait expliquer pourquoi le pouvoir d'achat de la majorité des français baisse alors que l'inflation est presque nulle.

Sources : Loïc Blondiaux, histoire des sondages (<http://www.qualiquanti.com/pdfs/histoiresondages.pdf>) ; notre-planete.info, <http://www.notre-planete.info/actualites/4332-nombre-arbres-monde> ; Grenoble INP, [Ensimag](#) ; INSEE, indice des prix à la consommation ([technique](#)) ; [Indice des prix, Pourquoi ? Comment ?](#) ; [Panier représentatif](#) INSEE ; La Société Française de Statistique (SFdS) met en ligne des textes pédagogiques sur les sondages qui sont faciles à lire et édifiants http://www.sfds.asso.fr/228-Sondages_electoraux