La statistique : la maîtrise des chiffres... ou leur manipulation ?¹

Il vaut mieux débattre avec ses manques et ses interrogations qu'avec ses certitudes

I – Des idées justes à partir de chiffres qui ne le sont pas... toujours

ou comment découvrir « l'ordre qui se cache dans les chiffres » (Süssmilch)

l'histoire

- 1- l'information de l'Etat
- 2- la découverte de la permanence statistique, les lois, la méthode
- 3- le sondage, enfin !
- 4- l'irruption de l'informatique
- 5- la statistique pour tous avec internet

l'usage de la statistique

- 1 – Connaître, se connaître
- 2 – Prévoir, pour décider
- 3 - Evaluer

II – Gouverner, c'est prévoir : la statistique publique

- 1 – l'information de base (recensements, enquêtes emploi...)
- 2 – des principes et des institutions – un code européen
- 3 – les difficultés de la comparaison internationale - Eurostat
- 4 – des bases de données en ligne

III – Pour voir mieux, voir peu : les sondages

- 1 – des débuts difficiles –
- 2 - l'échantillon : tirage aléatoire ou par quotas
- 3 – la collecte (face-à-face, téléphone, internet...)
- 4 – critique des sondages

IV – L'art de mentir avec précision : la statistique, indispensable, mais tellement décriée

- 1 – des chiffres toujours inadaptés aux attentes...
- 2 – des nomenclatures impossibles
- 3 – des domaines inaccessibles ou presque (du PIB au BIB)
- 4 – des usages abusifs ?
- 5 – des querelles d'experts en tous genres

¹ Document rédigé à la suite d'une conférence donnée le 20 octobre 2011 à l'ESAD (Ecole Supérieure des Arts Décoratifs de Strasbourg)

I - les étapes de la construction de la statistique

On peut résumer l'histoire de la statistique en quatre étapes.

1 – Les balbutiements

Connaître les ressources des pays, souvent des empires. On renvoie aux premiers recensements effectués il y a plusieurs millénaires (à Sumer, en Chine...) Il s'agit d'estimer les ressources (objectifs fiscaux et militaires). Rome réalise un recensement tous les cinq ans (cf. celui de la Judée). Signalons la grande opération de recensement anglais, le Domesday Book (1086). Quelques comptages dans les villes au Moyen-âge.

2- La réflexion

Le calcul des probabilités au XVIIème (Fermat, Pascal) ouvre la porte à une réflexion et nourrit l'étude des lois qui régissent les chiffres. Le terme de statistique apparaît en Allemagne et en France pour désigner les rapports rédigés à l'intention des gouvernements de l'époque. En Angleterre, Graunt et Petty estiment la population d'une ville, d'un pays à l'aide de coefficients multiplicateurs : début de la *statistique inférentielle*. Au XVIIIème, on parle d'arithmétique politique (Baron de Bielefeld, Laplace, Condorcet...). Le recensement de la population est inscrit dans la constitution américaine (1787). En France, la série des recensements commence en 1801.

La statistique est en effervescence au XIXème : exploitation systématique des documents administratifs, réflexions sur la *permanence statistique*, élaboration de lois (Gauss et sa loi normale). Mais les statisticiens étant des gens sérieux, ils veillaient à assurer l'exhaustivité de leurs observations (il fallait n'oublier personne !), d'où une certaine résistance à l'idée de travailler sur des échantillons.

3 – Les sondages, enfin !

Des journaux américains réalisent des sondages préélectorales dès 1824 (des votes de paille), le plus souvent avec succès, mais de façon empirique. Quelques enquêtes publiques par sondage sont menées aux USA et surtout en Russie. Il faut attendre le Congrès international de la statistique (Berne 1895) pour que l'idée de juger du tout d'après l'examen d'une partie bénéficie d'un début de reconnaissance (Kiaer). Les querelles entre partisans et adversaires des sondages se terminent au début des années 1930. Mais la véritable reconnaissance des sondages date de l'élection américaine (1936) : selon la revue *Literary Digest* qui interroge des millions de personnes (ses lecteurs), Roosevelt est donné battu. Mais il est donné gagnant selon Gallup et quelques autres instituts de sondage qui, de leur côté, interrogent un échantillon représentatif de quelques milliers d'électeurs seulement. En France, les sondages se développent sous l'impulsion de Jean Stoetzel qui fonde l'IFOP peu avant la guerre. Mais c'est avec la Vème république (qui instaure le régime présidentiel) et la publication d'un fameux sondage dans *France-Soir* qui annonce, la veille des élections, la défaite de De Gaulle que les sondages devaient prendre une place croissante dans la vie politique à travers leur publication dans les médias.

4 – L'irruption des nouvelles technologies

Les grandes enquêtes, les documents administratifs sont très lourds à dépouiller, à codifier puis à traiter. Les machines traditionnelles (machines à cartes perforées) sont progressivement remplacées à partir des années 1960 par des ordinateurs qui vont révolutionner la pratique statistique. Plus tard, avec la généralisation de l'internet, les directives européennes imposent aux instituts nationaux de statistique de

mettre en ligne l'essentiel de leur patrimoine national. Ainsi les citoyens n'ont plus à payer pour obtenir l'information qu'ils ont contribué à produire par leurs impôts. Seules les exploitations spécifiques font l'objet d'une facturation. Par ailleurs, les capacités de stockage permettent maintenant à tout un chacun de traiter des fichiers de données aussi volumineux que ceux d'un recensement. D'où la nécessité de développer des savoir-faire.

II- Gouverner, c'est prévoir : la statistique publique

- « *Non seulement les chiffres nous gouvernent, mais encore ils montrent comment le monde est gouverné* » Goethe
- *La statistique est un puissant réducteur d'incertitude* (Yves Franchet, ancien directeur général d'Eurostat)
- « *La statistique économique et sociale est un nouveau facteur de production, à l'instar du facteur travail et du facteur capital* » Jacques Delors

Disposer de données chiffrées, adaptées, fiables et reconnues par tous permet, du moins en principe, au débat démocratique de s'engager sur des bases saines. Il importe donc de créer et d'organiser un système élaboré de recueil et de diffusion de statistiques. De surcroît les données doivent être comparables au plan international. La construction de l'Union européenne passe aussi par la mise en place de deux institutions aux rôles et aux fonctions tout à fait différentes exercent pour le compte de la Commission une activité statistique, à savoir Eurostat et Eurobaromètre.

EUROSTAT – Cette institution située à Luxembourg est d'abord chargée de fournir à la Commission les statistiques nécessaires à la mise en œuvre des politiques dont l'Union a la compétence (agriculture, industrie, circulation des personnes...). Elle se doit également d'assurer des fonctions d'animation et de coordination, mais aussi de formation les cadres des instituts nationaux de statistique (INS). Elle publie un grand nombre de documents et gère la diffusion d'un vaste ensemble de bases de données.

L'application sans doute trop laxiste des critères de Maastricht a affaibli la confiance dans les chiffres. Le pouvoir de contrôle d'Eurostat sur les instituts nationaux, d'abord faible, a dû être grandement renforcé afin d'améliorer la crédibilité des statistiques utilisées pour la prise de décision. Eurostat a imposé un code européen des bonnes pratiques de la statistique (15 articles, le premier exigeant l'indépendance des INS par rapport aux gouvernements). Eurostat ne réalise pas d'enquête statistique mais les délègue aux INS.

Les instituts nationaux de statistique (INS) - Depuis plus de deux siècles, chaque pays dispose de ses propres institutions chargées de fournir les données dont il a besoin (en France, l'INSEE). L'organisation de ces systèmes statistiques est un héritage de l'histoire (système centralisé ou non), mais aussi de la culture du pays et des préoccupations qui émanent de la société. Celles-ci sont examinées par un comité consultatif, en France le Conseil National de la statistique publique, un petit « parlement » où siègent des représentants des différentes composantes de la société et l'Insee. Chaque année le CNIS élabore un programme d'enquêtes d'intérêt public devant être exécutées par un organisme public (INSEE ou un service public d'un ministère). Pour autant qu'ils se conforment aux directives européennes et qu'ils respectent le code des bonnes pratiques, les INS ont, selon le principe de subsidiarité, la liberté d'effectuer les travaux et études qu'ils souhaitent.

l'Eurobaromètre

Il s'agit d'un service de la Commission qui est chargé de faire exécuter (par des instituts privés) une série d'enquêtes par sondage auprès des ménages et des entreprises sur des thèmes ayant trait de près ou de loin à la construction européenne. L'Eurobaromètre en diffuse largement les résultats. Il n'a pas d'équivalent en France où les sondages d'opinion, y compris ceux qui sont réalisés pour le compte du gouvernement, sont généralement effectués par des entreprises privées.

Avec des séries internationales remontant dans certains cas à l'année 1973, et malgré les critiques, parfois justifiées, qui peuvent lui être faites, l'Eurobaromètre reste une base de données de références unique.

III– Pour voir mieux, voir peu : les sondages

- *« Mais comment n'aurais-je pas appris que ce qui est salubre pour le pays ne va pas sans blâme dans la nation ni sans pertes dans l'élection ? » De Gaulle, phrase conclusive de ses Mémoires d'espoir.*
- *Je suis leur chef, donc je les suis (X)*
- *Les sondages n'annoncent pas la révolution (Pierre Viansson-Ponté)*
- *Le sondage est à l'homme politique ce que le miroir est aux alouettes*

Effectuer un sondage, c'est procéder à une mesure sur un échantillon et à l'inférer à l'ensemble de la population. En ce sens le terme sondage s'oppose au terme recensement (observation exhaustive).

On distingue des acceptions assez différentes du mot sondage qui méritent d'être soulignées : le sondage comme méthode scientifique et le sondage comme outil d'analyse de l'opinion.

A - le sondage comme méthode scientifique

Il résulte du principe de non exhaustivité un certain nombre de conséquences, avantages et inconvénients par rapport à un recensement que l'on peut résumer ainsi :

- rapidité, coût plus faible de l'enquête,
- investigation plus poussée
- une plus grande qualité de la collecte.

- incertitude dans les résultats dus à la méthode, mais cette incertitude est maîtrisée, puisque l'on sait calculer un intervalle de confiance (fourchette)
- impossibilité de réaliser des croisements multiples de variables
- les catégories peu représentées sont ignorées

La réalisation d'un sondage implique que soit défini un plan de sondage (soit une méthode de tirage et un modèle de redressement). L'échantillon peut être tiré de façon raisonnée (méthode des quotas) ou de façon aléatoire (nombreuses façons de procéder, pouvant être combinées entre elles). La collecte des données se fait de façon directe (interview avec enquêteur en face-à-face, si possible à domicile, sur support papier ou, de plus en plus, sur tablettes informatiques) ou de façon indirecte (téléphone, internet). Le sondage trouve de très nombreuses applications, notamment dans le domaine scientifique (épidémiologie...), dans l'industrie, etc. A l'Insee on utilisait hier la méthode des sondages pour traiter les documents administratifs (avant le transfert automatique de ces documents à l'INSEE) ou pour exploiter et

contrôler les recensements. Aujourd'hui on l'utilise pour tirer l'échantillon du « Nouveau recensement ». En l'occurrence, le terme *recensement* représente, dans une certaine mesure, un abus de langage.

b – les sondages d'opinion comme outil d'analyse de l'opinion

Il s'agit d'une application spécifique de la méthode qui conduit souvent à s'éloigner des critères scientifiques qui fondent la méthode, ce qui n'exclut nullement une volonté de rigueur et de probité. Le sondage d'opinion est relativement peu appliqué à l'INSEE (enquêtes de conjoncture et certaines enquêtes auprès des ménages).

Les sondages d'opinion sont réalisés surtout par des instituts privés. Les critiques à l'égard de ces sondages sont nombreuses, mais ils sont d'un apport considérable à la connaissance de la société et au fonctionnement de la démocratie.

IV - La statistique, indispensable, mais tellement décriée

- *Les bonnes causes ne font pas toujours de bonnes statistiques*
- *Il se sert de la statistique comme un ivrogne se sert d'un lampadaire, pour s'y appuyer et non pour s'éclairer (Andrew Young)² –*

1 – La statistique, comme l'humour, fait défaut quand on en a le plus besoin

Les chiffres sont inadaptés aux attentes et sont fréquemment dépassés avant d'être publiés. Les débats convoquent des chiffres qui, souvent, ne servent pas vraiment l'argumentation, etc... Souvent ils ne convainquent pas car ils ne correspondent pas au ressenti et les producteurs de chiffres ne sont pas toujours sensibles à cette critique (l'indice des prix, les indicateurs de chômage, la mesure de l'immigration...).

Les meilleures institutions ne sont pas à l'abri d'erreurs ou de faux-pas. Il est alors tentant d'essayer de les masquer ou de les sous-estimer (les défauts de collecte au recensement de 1999 par exemple). Par ailleurs, l'inévitable fréquentation du pouvoir et les ambitions personnelles des responsables ne facilitent pas toujours la diffusion d'une statistique irréprochable. On ne peut que souhaiter une plus grande transparence, en conformité avec les prescriptions du code des bonnes pratiques européen.

2 – des nomenclatures impossibles

Classer est indispensable pour faire de bonnes statistiques, mais mettre les personnes dans des cases est forcément réducteur : ex comment classer les ouvriers-paysans qui gèrent une exploitation agricole ? dans quelle catégorie économique mettre les intérimaires ? Que faire de la population, toujours plus nombreuse, ayant plusieurs résidences, etc. Qu'est-ce qu'un cadre supérieur, que représentent au juste les classes moyennes ?

Certes, on fait des choix, on adopte des conventions que ceux qui en aval font les analyses ne connaissent pas, oublient ou négligent. Or on sait par expérience que, dans le débat médiatique, on a besoin d'idées

² *Maintenir l'indépendance de la statistique publique est une condition nécessaire – elle n'est évidemment pas suffisante - si l'on veut donner aux chiffres une crédibilité suffisante. On pourra se référer à l'article suivant (le point de vue de Stéphane Jugnot, statisticien économique, publié sur le blog de Médiapart)*
<http://blogs.mediapart.fr/edition/en-defense-de-la-statistique-publique/article/181010/lindependance-de-la-statistique-pu>

simples – donc de statistiques-massues - et malheureusement, on se garde d'apporter les nuances qui affaibliraient le propos³.

3 – des domaines inaccessibles, ou presque, à la statistique

On crée des agrégats qui sont des indicateurs remarquables en tant que constructions de la pensée (la statistique quantifie en quelque sorte des conventions qui se proposent de construire la réalité). Ils peuvent être internationalement reconnus puisque utilisés dans tous les pays, comme le PIB (Produit intérieur brut). Et pourtant on ne peut que constater leur inadéquation à rendre compte de la réalité. Les indicateurs de substitution ne manquent pas (Indice de développement humain par exemple) ; on cherche ainsi aujourd'hui à remplacer le PIB dont on connaît les insuffisances (en matière de prise en compte de l'environnement par exemple) par un autre qu'on appellerait le Bonheur national brut (BNB), mais cet indicateur, qui introduit la notion de bien-être, a du mal à s'imposer (rapport Stiglitz).

Et comment mesurer la nouvelle économie, quel sens a encore la mesure du temps de travail quand il s'agit de mesurer l'emploi intellectuel, des emplois où la création, l'innovation jouent un rôle majeur?

4 – des usages inappropriés des statistiques

Créer un système statistique dans un but d'évaluation (des politiques publiques, de l'efficacité des salariés...) conduit assez naturellement à des effets pervers. Ainsi la statistique renseigne souvent davantage sur l'activité statistique des acteurs (les chiffres de la délinquance sont souvent contestés pour cette raison au moins) plus qu'elle n'exprime la réalité qu'elle est censée mesurer. Au bilan, la statistique est dévaluée elle s'auto-affaiblit.

5 – Des querelles en tous genres

La science n'est pas à l'abri des effets de l'idéologie ou d'egos surdimensionnés, comme en témoignent trop souvent de lamentables débats entre experts... Trop souvent on s'engage dans des querelles fondées sur des chiffres qui ne correspondent à rien. Dans d'autres cas, les chiffres sont incontestables mais la manière de les interpréter conduit à des interprétations multiples. Exemple : le débat sur la baisse de la fécondité dans les années 1970. La réalité est complexe et ne saurait se réduire à quelques indicateurs, aussi séduisants fussent-ils.

Bibliographie

Ouvrages généraux sur la statistique

INSEE - Pour une histoire de la statistique – Economica 1987

M.L. LEVY – Comprendre les statistiques – Seuil - 1979

A. DESROSIERES - La politique des grands nombres - La Découverte 1993

M.VOLLE - Le métier de statisticien - Economica 1984

J. KLATZMANN - Attention Statistiques - Cahiers libres, la Découverte 1995

E.DIDIER – En quoi consiste l'Amérique ? Les statistiques, le New Deal et la démocratie – La Découverte

³ *Les journalistes ont certes des circonstances atténuantes : s'ils devaient à chaque fois accompagner leur diffusion de très longues mises en garde méthodologiques, les chiffres ne seraient plus audibles...*

Sur la statistique européenne

INSEE - Courrier des statistiques - dossier sur la statistique européenne - n° 121-122

Les catégories sociales

A. DESROSIERES, L. THEVENOT - Les catégories socioprofessionnelles - La Découverte 1988
M.OBERTI, A KIEFFER, E PRETECEILLE - Harmonisation des statistiques socioprofessionnelles en Europe - Sociétés contemporaines, n°45

Sur les sondages

L. BLONDIAUX - La fabrique de l'opinion - Science politique - Seuil 1998
H. MEYNAUD, D.DUCLOS - les sondages d'opinion - la Découverte 1985
A. GARRIGOU - L'ivresse des sondages - La Découverte - novembre 2006

Sofres - L'état de l'opinion, annuel depuis 1984, Seuil

ESOMAR (Association européenne pour les études d'opinion et de marketing) : code des bonnes pratiques
sur l'opinion européenne

B.CAUTRES et D.REYNIE - L'opinion européenne - Presses de Sciences Po - 2000 ...

D.REYNIE et alii - L'Opinion européenne - Odile Jacob - 2003 ...

Y.DELOYE et alii - Dictionnaire des élections européennes - Economica -2005

O.GALLAND et B.ROULET - Les jeunes européens et leurs valeurs - La Découverte 2005