

LA STATISTIQUE ENTRE LE LANGAGE DE LA SCIENCE ET CELUI DE L'ACTION

OU COMMENT DISCUTER L'INDISCUTABLE ?

ALAIN DESROSIÈRES

Membre de l'Institut National de la Statistique et des Études Économiques, Alain DESROSIÈRES est chercheur au Centre de Recherche en Économie et Statistique (CREST-INSEE) et participe aux activités du Groupe de Sociologie Politique et Morale (École des Hautes Études en Sciences Sociales et CNRS). Il a publié notamment Les catégories socioprofessionnelles, en collaboration avec Laurent THÉVENOT (Paris : La Découverte : 1992) et La politique des grands nombres. Histoire de la raison statistique (Paris : La Découverte, 1993).

LE STATISTICIEN ET SES LUNETTES

Ma réflexion sur la raison statistique est redevable, à l'origine, de l'enseignement de Pierre Bourdieu. Celui-ci s'employait à sensibiliser les apprentis statisticiens aux "lunettes de l'observation", autrement dit aux conditions de production de la connaissance. Les techniques et les outils mobilisés pour identifier, décrire et expliquer sont historiquement et socialement situés. Lorsque l'on regarde le monde, il arrive un moment où il est utile d'enlever les lunettes et d'en examiner les foyers, pour rompre avec une

épistémologie réaliste, toujours vivace, qui postule naïvement que le statisticien compte le mieux qu'il peut.

L'interrogation nécessaire sur l'outil trouvait sa justification immédiate dans l'intérêt porté aux classes sociales comme catégorie explicative. Dès lors que l'on rompt avec l'épistémologie naïve, l'étude des classes sociales en termes statistiques passe inévitablement par une mise en question du dispositif censé les saisir, autrement dit le code des C.S.P. (catégories socioprofessionnelles).

Cette posture réflexive sur les statistiques a donné une impulsion forte et originale à l'époque. Mais *a posteriori* j'en perçois toutes les limites. En effet, au regard du propos sur les conditions de production de la connaissance, l'attitude de P. BOURDIEU par rapport à la science pose problème. Celui-ci se réclamait de la Science avec un grand "S", affirmant bien haut la scientificité de la sociologie. Cependant, sa théorie du champ scientifique conduit à s'interroger sur le rapport entre *la* Science et les sciences. Elle présente l'inconvénient de ne pas vraiment prendre au sérieux le contenu des activités scientifiques. Que ce soit en matière de science, de religion, de droit ou d'art, domaines dans lesquels des travaux éclairants ont été réalisés, le dénominateur commun de ces champs serait de mettre en opposition les pratiques des intervenants

et les représentations qu'ils s'en font. Sans minimiser l'apport décisif du courant sociologique, force est de considérer qu'une telle démarche ne favorise pas une mise en perspective du développement de l'activité scientifique.

Apparu plus récemment, le courant de la sociologie des sciences et des techniques a joué un rôle non moins important.

DE LA SOCIOLOGIE DES SCIENCES À LA MONTÉE EN GÉNÉRALITÉ

La sociologie des sciences existait déjà de longue date mais était enfermée dans une sorte de partage des tâches consistant dans une séparation de l'interne et de l'externe. Certains travaux, de type *externaliste*, étudiaient les conditions sociales de possibilité d'une science à l'occidentale. Leurs interrogations portaient sur l'environnement et les facteurs susceptibles d'expliquer le déploiement différencié de l'activité scientifique. Ainsi, pouvait-on se demander pourquoi la science s'est développée d'une certaine façon en Europe occidentale et non pas en Chine ou dans le Monde arabe, qui, à d'autres époques, faisaient montre d'une avance en ce domaine. Les travaux *internalistes*, quant à eux, portaient sur la connaissance elle-même ses instruments, ses théorèmes et leur démonstration. Tandis que l'approche *externaliste* était surtout le fait d'historiens et de sociologues, la démarche *internaliste* relevait des spécialistes des disciplines scientifiques concernées : les mathématiciens traitaient de l'histoire interne des mathématiques et les physiciens de celle de la physique.

Ce partage a été remis en cause, durant les années 1970 par des chercheurs anglo-saxons, qui prônaient la combinaison des deux approches, en soulignant la nécessité d'observer simultanément la façon dont les scientifiques bricolent du social, constituent des réseaux d'alliance pour renforcer leurs positions, et bricolent leurs objets, autrement dit leurs théorèmes et leurs instruments¹. Cette abolition de la distinction entre objets sociaux et objets techniques a débouché sur un programme de sociologie de la science, dit *programme fort*. Elle a donné lieu à débat parmi les scientifiques dans la mesure où elle mettait entre parenthèses la question de la vérité. La sociologie de la science, en focalisant son approche sur les

mécanismes sociaux qui érigent des résultats en faits avérés, ne niait-elle pas la possibilité de la vérité et, partant, de la scientificité ?

Commentant le "renversement" de la sociologie de la science, Alain Caillé note : "Plutôt que de se demander ce que la société explique de la science, on se demandera ce que l'existence de la science nous apprend sur la société au sein de laquelle elle est possible"². Pour ma part, je me situe plutôt du côté du deuxième terme de l'alternative. Ce qui m'intéresse n'est pas seulement la sociologie de la statistique, mais ce que cette sociologie nous apprend de la société. Le problème de la vérité est un problème d'une autre nature, qui n'est pas dépourvu d'intérêt mais qui n'est pas l'objet de mon propos.

Du *programme fort* de la sociologie de la science, j'ai surtout retenu ce qui a été importé et développé en France par Michel CALLON³ et Bruno LATOUR⁴. Ces deux enseignants-chercheurs de l'École des Mines de Paris ont notamment écrit sur ce qu'ils appellent "la sociologie des laboratoires", une activité qui consiste à observer la science "en train de se faire" au travers de ses opérations pratiques sur les lieux même de la production de connaissances. Ils conduisent des travaux de type microsociologique portant sur les interactions fines qui sous-tendent les processus de recherche.

Une autre méthode de travail fort intéressante, mise en avant par ce courant et particulièrement par M.CALLON et B.LATOUR, se rapporte à l'idée de controverse : la *controverse* est un moment privilégié qui donne à voir comment les scientifiques travaillent. Avec la *controverse*, le sociologue des sciences est à même d'observer des individus qui argumentent et recourent à des ressources de type varié, scientifiques, sociales et techniques.

Personnellement, j'ai été séduit par cette manière de procéder. Contrairement à la démarche initiée par P.BOURDIEU, elle a le mérite de prendre au sérieux les contenus, qu'elle considère au même titre que les caractéristiques sociales. Elle opère des reconstitutions extrêmement vivantes, fouillées, souvent drôles, d'affaires, de réseaux et de controverses.

Je dois ajouter que malgré tout ce que je dois à ce courant méthodologique, j'en fais aujourd'hui la critique. Je lui reprocherais ce que j'appelle son manque d'intérêt pour les ressources de "montée en

généralité". On est parti, il y a trente ou quarante ans, d'une sociologie privilégiant la généralité pour passer d'une extrême à l'autre avec une sociologie de type ethno-méthodologique faisant table rase de toute généralité pour se cantonner dans l'observation d'acteurs qui bricolent des objets. Le *programme fort* laisse dans l'ombre le fait que ces acteurs puissent être reliés à des ressources permettant de monter en généralités. Or l'idée de construction de généralités, d'objectivation me paraît particulièrement importante. Plutôt que de considérer, à la manière des objectivistes, que l'objectivité ou la généralité préexistent en tant que telles, il importe de s'arrêter sur les façons de faire de la généralité, de construire des objectivités. Chaque facteur se bricole, mais on ne part jamais de zéro, on recourt à des ressources de montée en généralités. La statistique figure parmi ces ressources. Du moins, tel est le programme dans lequel j'essaie de me situer et qui fait apparaître la mise en place d'un appareil statistique comme un objet particulièrement intéressant.

Sans doute, la sociologie de la science appliquée aux sciences dures est-elle sans cesse confrontée à la question de savoir, par exemple, si pour expliquer la structure de l'atome il importe de connaître l'organisation d'un laboratoire ou, *a fortiori*, les origines d'Einstein. Il est loisible de soutenir que la structure de l'électron est totalement indépendante des considérations sociales. En revanche, si l'on observe la société à travers ses instruments statistiques, on aborde un domaine qui illustre pleinement la pertinence du programme postulant la non séparation des outils cognitifs et politiques. La statistique, les probabilités en tant que projet mathématique recouvrent une activité scientifique où il est strictement impossible d'autonomiser les formalismes et les réseaux sociaux en présence. On pourrait dire que la probabilité statistique est le domaine de rêve pour une sociologie des sciences et des techniques. D'autant plus qu'elle s'avère typiquement une façon de monter en généralité. Introduire de la statistique, c'est introduire un élément de généralité.

STATISTIQUE ET MISE EN FORME DU SOCIAL

Lorsqu'on parle de statistique, on utilise un même

mot pour désigner des postures, sinon des objets, différents. La statistique des instituts de statistique consiste dans la collecte de données chiffrées pour décrire des caractéristiques sociales et économiques. Elle connote des opérations et des instruments tels que, par exemple, les recensements ou les indices des prix. Sa visée est descriptive alors que celle de la statistique enseignée est formaliste. Dans les départements universitaires de statistiques on ne parle pas, à vrai dire, de données mais de méthodes mathématiques, de statistique inférentielle, de tests, de méthodes d'estimation. Cette ambivalence de la statistique, qui renvoie dans un cas à la description de la société et dans l'autre au formalisme, procède de deux histoires relativement autonomes et différentes : d'une part, l'histoire des institutions, c'est-à-dire la mise en place de structures administratives, mobilisant des catégories du droit, des normes et des standards au travers des recensements, enquêtes, fichiers, *etc.*, et d'autre part, une histoire, de type plus internaliste, des méthodes mathématiques. Qu'est-ce qui fait tenir ensemble ces deux histoires ?

Il y a là une question philosophique tout à fait fondamentale, celle du réalisme. On est en présence d'une façon particulière de mettre en scène des objets. De la même façon que Molière écrit une pièce qui donne à voir Sganarelle - Sganarelle existe, Molière le met en scène - la statistique met en scène des objets, à savoir l'indice des prix, le taux de chômage, une mesure institutionnelle, une moyenne, une corrélation. Ces objets sont fabriqués mais, une fois fabriqués, ils vivent leur propre vie. Si on les soumet à la méthode d'enquête chère à M.CALLON et B.LATOUR, on constate une structure de *controverse* qui réapparaît régulièrement et qui a trait à leur réalité. Ainsi, les objets *classe ouvrière*, *chômeurs* ou *agriculteurs* procèdent de catégorisations : on fabrique des catégories à partir d'une définition des ouvriers, des chômeurs ou des agriculteurs et en vue d'évaluer leur nombre. Les catégorisations donneront lieu à dénonciation par référence aux critères retenus qui pourront être jugés insuffisants ou trop généraux. Tiennent-elles compte de toute la diversité de la condition ouvrière, des formes multiples du chômage ou de la différenciation régionale de l'agriculture ? La catégorie ou classe d'équivalence et sa dénonciation relèvent d'une structure de *controverse* qu'on ne cesse de rencontrer et qui affecte également les outils mathématiques. La moyenne, méthode de calcul la

plus élémentaire, a fait l'objet pendant tout le XIX^e siècle de débats d'une extrême importance. Calculer une moyenne c'est supposer qu'il existe un objet commun qui justifie l'opération. Souvent on assistera à une espèce d'opposition stéréotypée entre une position *réaliste* qui consiste à parler du taux de chômage ou de l'indice des prix comme si ces objets existaient par eux-mêmes et une position anti-réaliste, qualifiable selon les cas de *nominaliste* ou de *relativiste*, qui consiste à déconstruire. Cette opposition se traduit par un phénomène plus troublant dans la mesure où un certain nombre de personnes pourront, à un moment donné, tenir un discours réaliste et, en d'autres circonstances, adopter une attitude rigoureusement anti-réaliste. La mesure de la population active nous en offre un exemple avec l'ouvrage publié par deux statisticiens français sous le titre *Deux siècles de population active*⁵. Les auteurs ont reconstitué une série de la population active sur deux siècles par des méthodes très indirectes. La structure de la population active du XIX^e siècle n'a rien à voir avec ce qu'elle est aujourd'hui. La question posée en l'occurrence n'est pas de savoir si cette reconstitution est juste ou fautive. Elle relève davantage d'une sociologie de la rhétorique de l'usage de la statistique et pourrait se formuler de la manière suivante : dans quelles circonstances utilise-t-on tel ou tel type d'arguments ?

Il est frappant de constater que le discours dont la statistique fait l'objet n'a pratiquement plus bougé depuis trente ans, depuis la diffusion des techniques quantitatives dans les sciences sociales, entre les années 30 et 70. Les débats se sont figés dans la mesure où l'on a persisté à considérer l'objet statistique uniquement comme un moyen au lieu de le réintégrer dans une sociologie des modalités de construction de la généralité. Les statistiques se sont diffusées dans les domaines de l'administration et de la gestion en tant qu'outil de rationalité parmi d'autres, la scientificité devenant ainsi partie prenante du langage de l'action. Elles ont par la suite trouvé également une instrumentation dans les sciences sociales, notamment en économie, en sociologie et en histoire. Sans doute y aurait-il place pour une étude sur la réception différentielle entre ces trois disciplines. Mais force est de constater qu'après avoir été à la mode, le quantitatif est désormais critiqué de façon stéréotypée. On pourrait même parler à cet égard d'un refus, encore que celui-ci n'emprunte pas les mêmes formes en France et dans les pays anglo-saxons et varie suivant

les disciplines. Il reste que le sujet s'est étiolé et que la statistique n'est pas considérée comme un sujet de réflexion noble. Or, à mes yeux, la raison statistique, par le fait même qu'elle relève simultanément du langage de l'action et du projet scientifique, constitue en elle-même un objet d'investigation particulièrement éclairant sur la mise en forme du social. Envisagée sous cet angle, elle implique l'adoption d'une perspective historique.

DIMENSION SOCIALE ET DIMENSION COGNITIVE DE LA STATISTIQUE

Le rapport entre l'histoire et la statistique va bien au-delà de l'histoire sérielle, qui s'attache à rechercher et ordonner des statistiques relevant de domaines préalablement délimités. Il se noue de manière plus ambitieuse dans une confrontation des problématiques respectives de l'histoire des sciences et de la statistique. Cette confrontation a été amorcée, il y a une vingtaine d'années à l'occasion d'une rencontre entre historiens et statisticiens. Elle s'est développée en France avec la conduite de travaux sur l'histoire des administrations, des enquêtes, des recensements et des nomenclatures. Avec celles-ci, on retrouve le thème des lunettes du statisticien, cher à P.BOURDIEU. Dans ce cadre, pour ma part, j'ai travaillé sur l'histoire des CSP françaises. Un autre courant issu de l'histoire des sciences proprement dites s'est cristallisé durant les années 80 dans les pays anglo-saxons sur l'histoire des probabilités, autrement dit sur l'histoire des statistiques en tant qu'instrument mathématique. Pendant une année, une dizaine de chercheurs - philosophes, historiens, économistes, physiciens, psychologues, sociologues - se sont constitués en collectif de recherche sur ce thème. Il en est résulté une série de livres qui ont fait avancer le sujet de façon tout à fait impressionnante⁶.

L'ouvrage que j'ai consacré à *La politique des grands nombres. Histoire de la raison statistique* se situe au carrefour de ces différentes approches. Il procède d'un projet tendant à les faire tenir ensemble. Il m'importait d'adopter une démarche générale qui, tout en intégrant mes recherches antérieures sur les classifications⁷, fasse place aux apports de chercheurs venus d'autres horizons. A titre d'exemple, je citerai le nom du philosophe I.HACKING, auteur, entre autres

ouvrages, d'un livre dont le titre, *The Taming of chance* (Apprivoiser le hasard), illustre bien le propos sur le rôle joué par les probabilités⁸. Avec celles-ci, on se trouve à la charnière entre une démarche cognitive et une démarche sociale. Le développement des travaux sur la moyenne au XIX^e siècle, avec notamment l'astronome belge Adolphe QUETELET⁹, rend compte de ce trait. La question de la moyenne, de "l'homme moyen", était axée sur l'idée de régularité des phénomènes macro-sociaux, par opposition à la volatilité et à l'imprévisibilité des comportements individuels. La régularité des moyennes mettait à nu la contradiction entre une sorte de déterminisme macro-social et une supposée, ou revendiquée, liberté individuelle. De la sorte, le débat sur la moyenne combinait un aspect technique et une dimension philosophique relative au libre-arbitre.

Au regard de ces deux volets - la classification et l'outil mathématique - la controverse réalisme/anti-réalisme pose une question tout à la fois cognitive et pratique qui porte sur l'existence de l'objet ainsi que sur la qualité et les objectifs de ceux qui s'opposent.

On pourrait aller plus loin dans l'analyse et considérer le formalisme mathématique ou statistique d'un triple point de vue : syntaxique (quelles sont les règles de fonctionnement ?), sémantique (ce qu'on fait dire à la moyenne ; comment on l'interprète ?), pragmatique (le rapport à l'action ; ce qu'on compte en faire). Si l'on envisage le formalisme statistique sur une longue période, la relative stabilité des syntaxes contraste avec les mutations des sémantiques. Ainsi, la statistique mathématique anglaise, avec Francis GALTON¹⁰ et Karl PEARSON¹¹ au XIX^e siècle, dans un contexte fortement inspiré par Darwin, a poursuivi un projet inspiré par des préoccupations philosophiques et politiques sur l'hérédité, les aptitudes humaines et l'eugénisme. A cette sémantique, a correspondu une syntaxe des plus productives avec pour outillage : la corrélation, la régression, le test du *chi-deux* et l'analyse multivariée. Tous ces instruments figurent aujourd'hui en bonne place dans les manuels de statistiques alors que l'épistémologie et la ligne politique qui ont présidé à leur élaboration sont rejetées et presque oubliées. Les outils se sont autonomisés, la syntaxe survivant à une sémantique et une pragmatique initiales. De la même manière, une thèse récente sur l'histoire du modèle linéaire montre

que la sémantique autour de cette syntaxe a changé du tout au tout¹². C'est là un programme de recherche qui, en construisant un espace de variation de l'objet à partir de sa syntaxe, de sa sémantique et de sa pragmatique peut aller beaucoup plus loin que les approches de P.BOURDIEU ou de B.LATOUR.

Reste à savoir si l'objet objectivé tient toujours. Un axe de recherche fructueux a été développé en ce domaine par Laurent THÉVENOT avec *l'investissement de formes*¹³. Il s'agit d'appliquer à la fabrication des outils statistiques une structure de raisonnement utilisée en économie à propos de l'investissement. La fabrication d'une machine destinée à produire des objets standards suppose un investissement coûteux dans la perspective d'un gain ultérieur. De même, avec l'élaboration d'un outil statistique, on investit dans des formes techniques et sociales qui permettent d'assembler des choses distinctes. Cet investissement a un coût dans la mesure où il conduit à sacrifier quelque chose (la multiplicité des cas singuliers) mais il génère un profit ultérieur avec la stabilisation de formes standards, mémorisables, transmissibles et réutilisables.

STATISTIQUE, MATHÉMATIQUES ET SCIENCES SOCIALES

On le voit, les statistiques ne sauraient être considérées comme un simple prolongement des mathématiques. Sur ce point, l'histoire des relations entre les statistiques et l'économie est éclairante.

A la fin du XIX^e siècle, les économistes se répartissaient entre deux pôles. Les uns, dans le sillage de l'école historique allemande, se plaçaient dans une perspective descriptive, s'attachant à la singularité d'une société ou d'une économie et utilisant la statistique comme un instrument, parmi d'autres, aux fins de mesure et de quantification. Ce courant, illustré à l'origine par Ernest ENGEL, admettait que la statistique puisse faire apparaître des régularités empiriques. En revanche, bien que par ironie de l'histoire le nom de Engel soit associé à une "loi" sur la relation entre revenus et dépenses alimentaires, il récusait l'idée que la statistique prétende établir des "lois" analogues à celles de la physique. A l'opposé, d'autres économistes se réclamaient d'une théorie économique, littéraire ou mathématique de type

hypothético-déductif, réservant peu de place aux statistiques. Il a fallu attendre les années 30 pour assister au mariage de la statistique et de la théorie économique, avec aux Etats-Unis le programme de la *Cowles Commission* et la création de la Société d'économétrie. L'appareil statistique a alors acquis une légitimité par son alliance avec l'économie théorique et l'économie appliquée. Cette évolution a été rythmée par des controverses entre les institutionnalistes, héritiers des historicistes, et les économètres. Parmi ces derniers, T. KOOPMANS a conduit, en 1947, une critique des plus radicales à l'encontre des statisticiens économistes descriptifs, accusés de ne référer leurs investigations statistiques à aucune théorie et, partant, de ne pouvoir généraliser leurs observations en vue de la prévision. Son article "Measurement without theory"¹⁴ témoignerait, si nécessaire, de ce que les enjeux de la statistique ne sont pas uniquement d'ordre mathématique mais qu'ils recouvrent la conception que l'on se fait des sciences sociales.

Dans un article souvent cité, Carlo GINZBURG, historien de la *microstoria*, a mis en cause le *paradigme galiléen* pour lui opposer le *paradigme indiciaire*.¹⁵ Il entendait signifier que les sciences sociales en général et l'histoire en particulier se sont représenté leur horizon de connaissance en fonction de lignes de force inspirées de l'astronomie galiléenne. A l'encontre de ce point de vue macro-social et formaliste, qui est celui de l'objectivation, il préconisait une approche interprétative fondée sur les *traces* et *indices*. En écho de cette critique du paradigme galiléen s'est diffusée une vulgate suivant laquelle la statistique et les mathématiques constitueraient, pour ainsi dire, les termes d'une même équation. Or, le classement et la quantification ne relèvent pas nécessairement ni exclusivement du *paradigme galiléen*. A plus d'un titre, ils participent des *traces*.

ALAIN DESROSIÈRES

NOTES

- 1 BLOOR D.- *Socio-logie de la logique ou les limites de l'épistémologie*.- Paris : Éditions Pandore, 1982.
- 2 CAILLÉ A.- *La démission des clercs. La politique des sciences sociales et l'oubli du politique*. - Paris : La Découverte, 1993
- 3 CALLON M., éd.- *La science et ses réseaux. Genèse et circulation des faits scientifiques*.- Paris : La Découverte, 1989.
- 4 LATOUR B.- *La science en action*. Paris : La Découverte, 1989.
- 5 MARCHAND O. et THÉLOT C.- *Deux siècles de marché du travail en France*.- INSEE, 1991.
- 6 A titre d'exemples : GIGERENZER G. et alii. - *The Empire of Chance. How probability changed science and everyday life*.- Cambridge : Cambridge-University Press, 1989.
- 7 DESROSIÈRES A. et THÉVENOT L.- *Les catégories socioprofessionnelles*.- Paris : La Découverte, 1992.
- 8 HACKING I.- *The Taming of chance*.- Cambridge : Cambridge University Press, 1990.
- 9 QUETELET A.- *Sur l'homme et le développement de ses facultés ou essai de physique sociale*.- Paris : Bachelier, 1835.
- 10 GALTON F.- *Hereditary genius. An inquiry into its laws and consequences*.- London : Macmillan, 1869.
- 11 PEARSON K.- *La grammaire de la science*.- Paris, Alcan, 1912 (trad. L. March).
- 12 ARMATTE M.- *Histoire du modèle linéaire. Formes et usages en statistique et économétrie jusqu'en 1945*.- Thèse de doctorat, EHESS, Paris.
- 13 THÉVENOT L.- "Les investissements de forme" in *Les Conventions économiques, Cahiers du Centre d'Étude de l'emploi*, 29, 1986, pp. 21-71.
- 14 KOOPMANS T.- "Measurement without theory", *Review of Economic Statistics*, 29, 1947, pp. 161-172.
- 15 GINZBURG C.- "Traces : racines d'un paradigme indiciaire" in *Mythes, emblèmes, traces, morphologie et histoire*.- Paris : Flammarion, 1989.