

Entre réalisme métrologique et conventions d'équivalence : les ambiguïtés de la sociologie quantitative*

(Article initialement publié dans *Genèses*, 43, juin 2001, pp. 112-127, reproduit avec l'aimable autorisation du comité de rédaction de la revue *Genèses* (<http://www.iresco.fr/revues/geneses/>))

Alain Desrosières

INSEE

79 rue Daguerre, 75014 Paris

Alain.Desrosieres@insee.fr

La “ méthodologie statistique ” utilisée par les sciences sociales quantitatives est en général présentée en deux parties bien distinctes, rarement reliées entre elles : d'une part, la construction des données, puis, d'autre part, leur traitement et leur interprétation. Souvent, entre ces deux mondes, se dresse la “ banque de données ”, qui fonctionne comme un sas de passage de l'un à l'autre. Or le monde de la “ construction ” est lui-même tendu entre deux façons de rendre compte de ses pratiques : la *mesure*, issue du langage des sciences de la nature, le *codage conventionnel*, inspiré, selon les cas, du droit, des sciences politiques, ou des sciences cognitives. Cette ambiguïté, caractéristique des sciences sociales quantitatives, peut être analysée à travers une mise en perspective historique, qui permet aussi de relier les questions soulevées par la construction des données à celles qui résultent de la diversité des modes d'analyse : l'exemple des différences entre “ régression logistique ” et “ analyse des données à la française ” sera notamment présenté ici.

L'expression même de “ méthodologie statistique ” implique une division du travail entre, d'une part, des “ experts ” de l'outil statistique en tant que tel et, d'autre part, des “ usagers ” de celui-ci, économistes, sociologues, historiens ou psychologues. Dans cette perspective, les progrès et les innovations de la statistique auraient un impact sur chacune de ces sciences, perçues comme des domaines d'application de formalisations élaborées en dehors d'elles. Il est vrai que l'autonomisation croissante de la recherche et de l'enseignement des statistiques mathématiques tend à conforter cette impression. Pourtant l'histoire et la sociologie de la statistique offrent de nombreux exemples d'effets *de sens inverse*, où des questions posées par des sociologues, des économistes ou des spécialistes d'autres disciplines ont profondément transformé la méthodologie statistique. Un exemple important en est fourni, à la fin du XIX^e siècle, par les travaux de Francis Galton et Karl Pearson sur l'hérédité humaine et l'eugénisme, qui ont conduit à formaliser la régression linéaire, la corrélation et le test du chi-deuxⁱ. À leur suite, les recherches agronomiques de Ronald Fisher conduisent à la statistique inférentielle moderne et à l'analyse de la variance.

Ce sont bien sûr les développements de l'économétrie qui, depuis les années 1930 et 1940, offrent les exemples les plus connus des interactions réciproques entre une discipline scientifique et la théorie statistique. C'est même à l'occasion des controverses de cette période cruciale de la quantification des sciences sociales, qu'ont été distinguées trois façons différentes d'envisager les relations entre “ théorie ” et “ données observées ”ⁱⁱ. Pour la première, une théorie, de type hypothético-déductif (Léon Walras) peut, au mieux, être illustrée par des données, mais elle est logiquement supérieure à celles-ci. Pour la seconde en