

Marion Selz et Florence Maillochon, *Le raisonnement statistique en sociologie*, Paris, Presses Universitaires de France, série Licence (Socio), 2009, 313 pages.

Marion Selz et Florence Maillochon, *Le raisonnement statistique en sociologie*, Paris, Presses Universitaires de France, série Licence (Socio), 2009, 313 pages.

Marc Barbut

**Édition électronique**

URL : <http://journals.openedition.org/msh/11217>

DOI : 10.4000/msh.11217

ISSN : 1950-6821

Éditeur

Centre d'analyse et de mathématique sociales de l'EHESS

Édition imprimée

Date de publication : 31 décembre 2009

Pagination : p. 69-71

ISSN : 0987-6936

Référence électronique

Marc Barbut, « Marion Selz et Florence Maillochon, *Le raisonnement statistique en sociologie*, Paris, Presses Universitaires de France, série Licence (Socio), 2009, 313 pages. », *Mathématiques et sciences humaines* [En ligne], 188 | Hiver 2009, mis en ligne le 15 mars 2010, consulté le 25 septembre 2020.

URL : <http://journals.openedition.org/msh/11217> ; DOI : <https://doi.org/10.4000/msh.11217>

ANALYSE BIBLIOGRAPHIQUE

Marion Selz et Florence Maillochon, *Le raisonnement statistique en sociologie*, Paris, Presses Universitaires de France, série Licence (Socio), 2009, 313 pages.

Lorsque j'étais étudiant sur les bancs de l'Institut Henri Poincaré, ce haut lieu des mathématiques, Georges Darmois commençait son cours de statistique mathématique en nous tenant à peu près ce langage : « Désormais, oubliez que vous êtes des apprentis mathématiciens. Soyez des physiciens ».

Il voulait dire par là, et il y revenait tout au long de l'année, que la statistique est avant tout une science d'observation et une *pratique* faute de laquelle sa théorie (qui, elle, est mathématique) serait vide de sens.

Cette leçon, je ne l'ai jamais oubliée, et me suis toujours efforcé de m'y conformer. C'est dire à quel point j'abonde dans le parti pris des auteurs du *raisonnement statistique en sociologie* de présenter celui-ci à partir de la démarche scientifique du sociologue.

Par contre, là où je ne les suis plus du tout, c'est dans leur autre parti pris : celui de délibérément ignorer la formulation mathématique des techniques statistiques qui mettent en œuvre le raisonnement.

Probablement ce parti pris n'est-il que le reflet de la réalité *actuelle* des rapports entre sociologie et statistique – encore qu'il y ait quelques notables exceptions dont témoignent notamment certains des articles publiés dans notre revue. Mais cette réalité est affligeante. En quelques décennies, la situation a bien changé. Comment en est-on arrivé là ? J'y reviendrai.

L'ouvrage est divisé en quatre parties. Les parties II et III, respectivement intitulées « Quantifier pour objectiver » et « Collecter des données » concernent les méthodes de l'observation et de l'enquête sociologiques. Ce qui y est exposé emporte l'adhésion et, pour l'essentiel, relève d'ailleurs du simple « Bon sens » dont les auteurs ont bien raison de se prévaloir. Étant incompetent en sociologie, je ne me risque pas à commenter plus avant ces deux parties du livre.

La quatrième partie est intitulée « Le raisonnement statistique ». Nous arrivons donc là à ce qui est le cœur de l'ouvrage. En quoi un « Raisonnement statistique » consiste-t-il ? Les auteurs le soulignent bien : en inférences. Seulement, voilà, ces inférences, au moyen d'observations partielles, quant à une population tout entière, elles sont *probabilistes*, toujours probabilistes, et seulement probabilistes. C'est d'ailleurs bien ce qui ressort de la lecture de M. Selz et F. Maillochon. Deux mots reviennent constamment dans leur texte : probabilité, aléatoire.

Mais ils ne sont jamais définis. Et pour cause. Car probabilité (d'un événement) ou variable aléatoire ne sont définies que par des règles de calcul (les règles du calcul des probabilités) dont justement on ne veut dire mot dans cet ouvrage. Moyennant quoi, tout ce qui est expliqué quant aux techniques et aux méthodes de la statistique, ne peut que rester incompris du lecteur non averti.

À titre d'exemple : comment comprendre qu'un échantillon de taille donnée, 1.000, par exemple, (p. 255) donne autant d'information sur une population de 10^{27} individus que sur une population de 100.000, si l'on ne connaît pas – ne serait-ce que de façon rudimentaire – le *calcul* de l'intervalle de confiance et de sa probabilité ?

Autre exemple, à propos du khi-deux (p. 275 et sq.). Qui peut comprendre que ce même mot recouvre trois notions bien distinctes, et à soigneusement distinguer, à savoir un écart, une distribution de probabilité de cet écart, et un test statistique résultant des deux premières, sans passer par leur représentation mathématique

Mais non. Pour les auteurs, c'est inutile, puisque les logiciels de statistique vous fournissent le résultat numérique des calculs. Ce n'est pas comme «auparavant» (note p. 276) «maintenant» le recours à la boîte noire nous évite la fatigue de raisonner et d'essayer de comprendre. C'est vraiment magique

Car nos auteurs, faute de mathématiques, renvoient aux logiciels à propos de chacune des techniques évoquées dans leur ouvrage. Ce qui est exactement le contraire de la démarche dont elles ont fort pertinemment accompli une première étape sa formulation discursive. La deuxième, sa formulation mathématique, est indispensable à mon sens.

Quelques remarques, qui ne sont pas que de détail sur cette dernière partie du livre.

- La taille quelque peu mythique de 30 pour un échantillon (p. 250) n'est nullement suffisante pour que la «loi des grands nombres» s'applique. Ce à quoi elle suffit, c'est à approximer convenablement une loi binomiale (voire hypergéométrique) par la loi de Laplace-Gauss.
- La note de la p. 255 censée définir la variance est à supprimer celle-ci est correctement définie par la note de la p. 269.
- À propos des «tableaux de contingence» (p. 272), pourquoi ne pas dire tout simplement qu'ici «contingence» est un anglicisme, et qu'il faut comprendre «tableaux de dépendance»
- En français, «odds-ratio» (p. 241) se dit «bi-rapport». Mais il est vrai, les logiciels salvateurs ne connaissent que l'anglais.

Comment en est-on arrivé là À cette question que je posais plus haut, la principale réponse est dans la première partie du livre, «La place et le rôle des statistiques en sociologie», bien meilleure, à mon sens, et nettement plus convaincante que la dernière. Que ce soit à propos de «l'enseignement des statistiques» (et non de la «statistique», la nuance est importante) ou des «attentes, illusions, ...» suscités, dans l'histoire de la sociologie, par la statistique et ses outils.

M. Selz et F. Maillochon sont persuadées (cf. p. 35) que l'étudiant en sociologie et le sociologue sont irrémédiablement rebelles aux mathématiques, d'où leur parti pris de les passer par profits et pertes.

Or, il n'en a pas toujours été ainsi. Au début des «années 1920», à Strasbourg, un mathématicien et un sociologue, les deux grands Maurice (Fréchet et Halbwachs) enseignèrent de concert la statistique. Il en résulta un ouvrage «à quatre mains» vaillamment intitulé *Le calcul des probabilités à la portée de tous* (Paris, Dunod, 1924). À cette époque, le dit calcul était donc loin de faire peur au sociologue.

Bien plus tard, en 1966, lors de la création du D.U.E.L. (et non du D.E.U.G. qui n'existera qu'en 1974, cf. p. 37) de sociologie, le parti pris voulu par les sociologues, Jean Stoetzel en tête, a été d'y inclure un enseignement obligatoire et exigeant de mathématiques. Les programmes, élaborés en collaboration, consistaient en des rudiments d'algèbre linéaire d'une part, de calcul des probabilités d'autre part les deux piliers sur lesquels repose la statistique, dont l'enseignement proprement dit ne devait

intervenir qu'à partir de la troisième année (Licence et Maîtrise). Et ça *a marché*, nous sommes encore quelques survivants à pouvoir en témoigner. Oserai-je rappeler que des étudiants de la «Corpo» (affiliée à l'U.N.E.F.) de sociologie de la Sorbonne rédigeaient en 1966-1968 mon cours hebdomadaire, me le soumettaient pour corrections (il y en avait fort peu) et le diffusaient ensuite à leurs condisciples.

Pourquoi cela ne *marche-t-il* plus ? Les auteurs nous le disent très explicitement (p. 45 et 46 notamment) : «Plusieurs logiciels ont été mis sur le marché au cours de ces trente dernières années...» «L'existence de ces logiciels a également fait évoluer l'enseignement de la statistique, pour le meilleur mais aussi, parfois pour le pire», etc...

Au «Parfois» près, auquel je substituerai volontiers «Presque toujours», ce diagnostic est le bon.

Un regret pour terminer cette (trop) longue note de lecture. C'est que n'aient été mentionnés ni dans le corps du texte, ni dans sa bibliographie, deux textes admirables concernant le propos du livre (le raisonnement statistique en sociologie).

Il s'agit d'une part du chapitre VI, tome 1 du *Traité de sociologie* de Georges Gurcitch (Paris, Presses Universitaires de France, 1959), chapitre intitulée «Les problèmes de la statistique» par son auteur Georges Th. Guilbaud. Ce chapitre a été réédité dans notre revue en 1996 (*Mathématiques et Sciences humaines* 135, p. 33-50) et est donc accessible et téléchargeable sur notre site.

Il s'agit d'autre part de l'article publié par Emile Borel dans *L'année psychologique*, t. 4, 1908, p. 125-151) «Le calcul des probabilités et la méthode des majorités». Ce texte est actuellement disponible sur le site (www.jehps.net) du *Journal Electronique d'Histoire de la Probabilité et de la Statistique*, n° 5-2, décembre 2009.

En résumé, *Le raisonnement statistique en sociologie* est un livre à lire, il y a beaucoup de bon à y prendre, malgré quelques scories dont j'ai relevé plus haut certaines.

Mais cette lecture doit conduire à suivre le conseil donné par les auteurs dans leur note 1, p. 250-251, ou celle de la p. 253 : n'en restez pas là et approfondissez vos connaissances en étudiant des cours de statistique «classique».

Comme l'aurait dit G. Darmon, le statisticien travaille avec sa tête et avec ses mains : le recours aux logiciels fait qu'il n'y a plus ni tête, ni mains. Il n'y a qu'un seul moyen de s'affranchir de cette servitude. C'est de suivre la voie qu'on préconisait au Siècle des Lumières : *Studia la mathematica*.

Marc Barbut