

L'apparition des sondages aléatoires et des statisticiens agricoles aux États-Unis, 1920-1940.

Emmanuel Didier, CESDIP-CNRS. Journées d'Histoire de la Statistique, INSEE, 15 et 16 février 2006.

Il y a quelque chose d'ironique à demander, dans le cadre de ces journées d'histoire de la statistique, si l'usage des probabilités par certains statisticiens serait une « affaire de profession ». En effet, cette question, de sens commun, laisserait présupposer que ce serait l'état de la profession de statisticien au moment où nous l'observons qui, pour une raison ou une autre, aurait conditionné le recours à cette branche particulière des statistiques que sont les probabilités. La structure de la profession serait la condition des outils statistiques utilisés. Or l'ironie réside en ce que les organisateurs de ces journées, et je pense à Alain Desrosières et à Laurent Thévenot en particulier, ont par le passé bravé le sens commun et montré l'importance du rapport inverse entre statistique et profession : ils ont insisté sur le fait que la statistique était une *condition* de l'existence même d'une profession, et non son effet. Ils ont montré que, pour qu'on ait des catégories sociales bien établies comme celles des « cadres », des « ouvriers », ou des « infirmières », il leur fallait d'abord s'établir fermement dans les nomenclatures socioprofessionnelles, dans la taxinomie, outil statistique s'il en est, qui apparaît du même coup comme un des déterminants essentiels de la forme que prend un groupe professionnel. Autrement dit, la représentation statistique d'un groupe est un des principaux vecteurs permettant sa représentation sociale globale, et non l'inverse. Ce que nos prédécesseurs nous ont enseigné, c'est que ce sont les statistiques qui conditionnent les professions, et non les professions qui conditionnent les statistiques : le contraire de la question que ces auteurs posent aujourd'hui.

Mais, mais, mais voilà ... nous voici presque 20 ans après leur fait d'arme (la première version de leur livre, *Les catégories socioprofessionnelles* date de 1988), l'histoire des statistiques a elle-même une histoire, des styles successifs (ceci aussi, Alain Desrosières l'a montré), et le sens commun finit toujours par reprendre le dessus, de sorte qu'on ne se demande plus comment les statistiques participent à la construction des professions mais au contraire, comme il y a quarante ans, quel rôle joue la profession dans le recours à tel ou tel outil statistique. Mais alors, cette question d'aujourd'hui, « Les statistiques avec ou sans probabilités, une affaire de profession ? » ne mène-t-elle à rien d'autre qu'à un repentir tardif, à la dénégation de tous les acquis des travaux antérieurs ? Non. Trois fois non, et voici pourquoi.

Nous allons travailler sur un Bureau du ministère de l'Agriculture américain que nous appellerons pour simplifier – parce que son nom a en fait beaucoup changé - « la Division » (il s'appelle aujourd'hui le NASS : National Agricultural Statistical Service) qui depuis sa création en 1860 était chargée d'estimer, chaque année, la production totale d'un grand nombre de produits agricoles : céréales, bétail, etc. Ces gens qui pratiquaient la statistique au début de la période que nous étudions, c'est-à-dire jusque pendant les années 1920, peuvent être doublement caractérisés.

D'une part, initialement, la Division construisait ses statistiques selon des méthodes parfaitement étrangères aux probabilités.

Elle s'appuyait sur une longue liste de Correspondants qui étaient répartis dans tout le pays de façon à ce qu'ensemble ils lui donnent de l'information sur l'exhaustivité du territoire agricole national. Chaque Correspondant était chargé de renseigner la Division sur la production dans son petit bout de territoire à lui, et la division sommait cette multitude d'informations pour estimer les productions totales. Il n'était pas alors question d'aléatoire ni de probabilité, pour au moins quatre raisons : d'abord l'information fournie par les enquêtés était *exhaustive*, sa fiabilité ne pouvait donc être évaluée à l'aune d'une éventuelle erreur d'échantillonnage, deuxièmement les correspondants étaient *volontaires*, au sens où ils se déclaraient volontaires pour répondre, et se voyaient alors inscrits sur de longues listes de Correspondants, et non pas du tout tirés au sort, troisièmement, ils étaient *fidèles*, c'est-à-dire qu'ils n'étaient pas censés répondre à une enquête ponctuelle mais systématiquement à toutes les enquêtes de la Division, ce qui encore une fois aurait compliqué le tirage au sort et enfin, ils étaient eux-mêmes à la fois les enquêteurs et les enquêtés de la Division : ils devaient se renseigner par eux-mêmes pour estimer les quantités pertinentes dans tout leur voisinage. Par conséquent, point d'aléa ni de probabilités.

D'autre part, pendant cette période, nul n'aurait songé à parler d'une « profession » de statisticiens agricoles. Ainsi, premier exemple, les vrais noms de « La Division » ne contenaient-ils même pas le mot de « statistique » en tant que tel : elle s'appelait par exemple la « Division of Crop and Livestock Estimates » ou ensuite le « Bureau of Agricultural Economics », mais point de « Statistics » nulle part : les praticiens n'avaient donc pas de domaine de compétence reconnu. Second exemple, les individus que nous rencontrons n'ont pas été formés aux statistiques mais à l'administration, et pour cause puisque dans les universités il n'y avait que très exceptionnellement de service statistique séparé. Autrement dit, la profession que nous étudions, au sens classique de domaine réservé et de formation unifiée, *n'existait pas* au début de la période sous examen.

Or, si au début de notre période il n'y avait ni probabilité, ni profession de statisticiens, à la fin de la période qui nous intéresse, en 1945, les statisticiens agricoles avaient mis au point un instrument encore en usage aujourd'hui et qui a eu d'importantes répercussions, y compris en France puisqu'il a été réutilisé par l'INSEE ; cet instrument c'est « l'Echantillon maître », une méthode de sondage qui repose sur l'échantillonnage aléatoire, en particulier sur les techniques de sondage aréolaire inventées à cette occasion, et donc sur des outils probabilistes.

De même, au sortir de la Seconde Guerre mondiale, les statisticiens étaient organisés professionnellement (l'American Statistical Association existait dès le XIX^{ème} siècle, mais elle n'était devenue incontournable qu'après la seconde guerre mondiale), ils partageaient des formations (l'exemple du « Stat Lab » de l'Iowa State University avait été mille fois imité) et ils avaient établis un certain nombre de tours de mains propres à leur profession, bref la profession – toujours dans ce sens classique de formation identifiée et de corpus de connaissance partagé – existait.

Par conséquent, notre argument consistera alors à montrer comment la profession de statisticien et les probabilités se sont « co-construites » comme aime à le dire Alain Desrosières ou, pour le dire autrement, ont exprimé leurs caractères mutuels. Nous voudrions montrer comment, la profession de statisticien agricole permet de comprendre comment les acteurs ont eu recours aux probabilités et comment ils les ont adaptées à leurs besoins, mais nous verrons aussi comment, dans le même temps, la profession était elle-même durcie, autonomisée par le recours aux probabilités. Nous verrons comment profession et probabilités ont interagi et se sont donc mutuellement façonnées.

Ainsi, nous allons explorer ce que signifie « l'apparition » ou « la naissance » dans le cas et d'une méthode statistique (nous parlons ici des sondages aréolaires, qui sont probabilistes), et d'une profession. On verra que nous avons recours à la notion d'« expression » : les statisticiens ont peu à peu *exprimé* une nouvelle méthode probabiliste et inversement, les probabilités ont *exprimé* certaines caractéristiques du métier de statisticien. Il ne sera alors pas possible de dire qui, de l'organisation sociale (la profession) ou de la technique statistique (les probabilités) est la cause de l'autre, mais au contraire il le sera d'observer comment l'une a produit les caractéristiques de l'autre, et vice versa. Ainsi, la question qui nous est aujourd'hui posée n'est pas un renoncement aux acquis, mais au contraire leur prolongement, un prolongement qui au passage nous force à nous questionner sur un objet central de la méthode historique : l'apparition.