**Mathématiques et sciences humaines**
Mathematics and social sciences

**166 | 2004**
**Varia**

# Les débuts de la statistique mathématique en Espagne (1914-1936)

*The rise of mathematical statistics in Spain (1914-1936)*

**Jose-Maria Arribas**

**Édition électronique**
URL : http://msh.revues.org/2895
DOI : 10.4000/msh.2895
ISSN : 1950-6821

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

**Édition imprimée**
Date de publication : 1 mars 2004
ISSN : 0987-6936

**Référence électronique**
Jose-Maria Arribas, « Les débuts de la statistique mathématique en Espagne (1914-1936) », *Mathématiques et sciences humaines* [En ligne], 166 | Été 2004, mis en ligne le 05 mars 2006, consulté le 13 octobre 2016. URL : http://msh.revues.org/2895 ; DOI : 10.4000/msh.2895

Ce document est un fac-similé de l'édition imprimée.

© École des hautes études en sciences sociales

*Math. & Sci. hum. / Mathematics and Social Sciences* (42e année, n° 166, *2004*(2), p. 25-46)

# LES DEBUTS DE LA STATISTIQUE MATHEMATIQUE EN ESPAGNE (1914-1936)[1]

José M. ARRIBAS[2]

RESUME – *Il y a aujourd'hui un consensus pour reconnaître que les années vingt ont été décisives dans l'histoire de la statistique mathématique. Quelques auteurs on parlé d'une nouvelle ère dans le développement de la statistique théorique. L'article essaie d'expliquer en quoi consiste cette nouvelle ère qui touche tous les pays européens, et de quelle statistique il s'agit. Ces mêmes années vingt sont, à tous points de vue, une période de grande vitalité pour l'Espagne. Comme dans les autres pays européens, la statistique mathématique s'y développe dans deux directions : la physique mathématique et l'économétrie. On y retrouve également tous les problèmes liés à l'institutionnalisation de la discipline ou au changement de paradigme associé aux noms de F. Y. Edgeworth et de A. Bowley. L'hypothèse de départ est en effet que la nouvelle statistique a pour noyau la théorie des échantillons. Bien que son application rencontre encore des problèmes techniques et méthodologiques importants, les conditions sociales et politiques de cette période permettent déjà d'envisager la naissance d'une nouvelle science appelée statistique mathématique.*

MOTS-CLES – Statistique mathématique, Histoire de la statistique, Méthodologie, Sociologie de la science

SUMMARY – The rise of mathematical statistics in Spain (1914-1936)

*There is a consensus amongst the historians of statistics about the importance of the twenties as starting "a new era" in the theoretical statistics. The article tries to give account of this period and to show the characteristics of this new form of statistics. For Spain, this period was very exciting in may respects. As in the rest of the countries, Mathematical Statistics is developed here in two directions: mathematical physics and econometrics. We can find all the problems related to the institutionalisation of the discipline and the change of paradigm associated to the British mathematicians F.Y. Edgeworth and A. Bowley. The hypothesis is that the new statistics has its central point in the theory of sampling. In spite of the technical and methodological problems that its application still entails, the social and political conditions allow for a new science called Mathematical Statistics to arise.*

KEYWORDS – Mathematical Statistics, History of Statistics, Methodology, Sociology of Science

## 1. INTRODUCTION

Il y a aujourd'hui un consensus pour reconnaître que les années vingt ont été décisives dans l'histoire de la statistique mathématique. Ainsi Donald Mackenzie conclut son étude devenue classique "*Statistics in Britain*" en affirmant que «□usqu'au milieu des années vingt, il y avait des signes clairs du commencement d'une nouvelle ère dans le

[1] Article reçu le 20.06.2003, révisé le 17.09.2003; accepté le 28 janvier 2004

[2] Departamento de Sociologia I, UNED, Madrid, jarribas@poli.uned.es

développement de la statistique théorique en Grand Bretagne[3]. En quoi consiste cette nouvelle ère qui touche tous les pays européens[4], et de quelle statistique s'agit-il ?

Ces mêmes années vingt sont, à tous points de vue, une période de grande vitalité pour l'Espagne. Comme dans les autres pays européens, la statistique mathématique s'y développe dans deux directions : la physique mathématique d'un côté, l'économétrie de l'autre. On y retrouve également tous les problèmes liés à l'institutionnalisation de la discipline ou au changement de paradigme associé aux noms de F. Y. Edgeworth et de A. Bowley. Notre hypothèse est en effet que la nouvelle statistique a pour noyau la théorie des échantillons. Bien que son application rencontre encore des problèmes techniques et méthodologiques importants, les conditions sociales et politiques de cette période permettent déjà d'envisager la naissance d'une nouvelle science appelée statistique mathématique.

Ce travail comprendra trois parties. Dans la première nous montrons l'influence de la statistique anglaise et surtout de la voie représentée par F. I. Edgeworth et A.Bowley, ainsi que les problèmes liés à l'application des nouvelles techniques. C'est au sein d'une autre nouvelle discipline, la sociologie, que se produisent des critiques importantes à ce propos. Dans la seconde partie, nous présenterons les données historiques et les acteurs espagnols de la statistique mathématique, et dans un troisième temps, nous verrons quelques problèmes méthodologiques liés à l'enseignement de la statistique et à la confusion entre deux paradigmes : le modèle linéaire, déjà très bien développé et la théorie des échantillons qui arrive.

## 2. QU'EST-CE QUE LA STATISTIQUE MATHEMATIQUE ?

Au milieu des années vingt[5], la statistique joue un rôle important dans l'agriculture, la production industrielle, l'industrie militaire ou la création de services publics. Quelques instruments statistiques sont déjà arrivés à un certain point de sophistication mathématique. C'est à cette époque qu'apparaissent aussi les premiers manuels de statistique mathématique, par exemple, Schultz publie en 1927 *Mathematical Economics and the quantitative methods,* G. Darmois publie en 1928 *Statistique Mathématique*, les traductions et rééditions des manuels de A. Bowley (1901) et U.Yule (1911) se suivent, le premier manuel espagnol apparaît en 1924[6], et la statistique commence à entrer dans les programmes d'enseignement des facultés de sciences. A. Bowley occupe en 1919 la première chaire de statistique à l'université de Londres, et en 1928 on inaugure l'Institut Henri Poincaré à l'intérieur duquel s'installe l'ISUP, l'Institut de Statistique de l'université de Paris, créé en 1922[7].

---

[3] D.A. Mackenzie, 1981, p. 213.

[4] Un autre pays important du point de vue du développement statistique est la Russie, spécialement en la personne de A.Tchouprov (1874-1926) et A. G. Kovalevski (1892-1933). Ce dernier publie *Fondements de la théorie de la méthode du sondage,* Saratov (Russie*)*, 1924, (cf. [BLUM., MESPOULET, 2003]).

[5] Les années vingt sont une époque de vitalité intellectuelle si extraordinaire que certains en arrivent à la considérer comme la plus importante depuis l'époque de la pensée grecque. Les interprétations élaborées par l'institut de physique de Copenhague entre 1925 et 1927 autour de Bohr sur l'espace, la causalité et le temps, en sont l'illustration, ainsi que les articles que l'on peut trouver dans la *Revue Philosophique* sur la notion de cause (Meyerson), sur la physique quantique (Goblot) ou sur la statistique (Halbwachs). *Revue Philosophique de la France et de l'Etranger*, XCVI, Paris, Félix Alcan, juillet à décembre 1923.

[6] A. De Miguel, *Metodología Estadística. Fundamentos de Estadística matemática,* Madrid, 1924.

[7] Voir le n° 67, mars 2002 de *Gérer & Comprendre*, Annales des Mines, particulièrement l'article de F.

Bien que la statistique mathématique apparaisse dans des manuels d'enseignement des années vingt, il est évident qu'elle est en gestation dans les décennies antérieures avec ce qui touche à la prévision des cycles économiques et l'évolution des prix [Armatte, 1992]. Les méthodes de la corrélation développées par Bravais (1846), Galton (1888), Edgeworth, Pearson et Yule, étaient déjà généralisées avant l'institutionnalisation de l'enseignement de la statistique mathématique. Nous savons par Michel Armatte qu'à partir de 1880 on commence à définir une statistique dont les principes et les techniques sont différents de la statistique et de la théorie des moyennes caractéristiques du XIX[e] siècle, et dont l'objet était de présenter différentes mesures sur un mode comparatif[8]. Edgeworth, publie en 1881 *Mathematical Psychics,* où il tente d'appliquer les mathématiques aux sciences sociales, notamment à la psychologie, livre qui aura beaucoup d'influence sur les économistes marginalistes. En 1906, les mesures avec des moyennes, tout comme les idées de médiane, quartiles, mode, moyenne arithmétique, nombres indices, écart type, etc. faisaient déjà partie du savoir des statisticiens avancés, et ils étaient utilisés couramment dans les statistiques officielles en Grande-Bretagne et aux États-Unis.

Pour mieux comprendre cette irruption de la statistique mathématique, nous allons nous servir du rapport de 1906, adressé par A. L. Bowley à la section d'économie et de statistique de l'Association Britannique pour l'Avancement de la Science, et publié cette même année dans le bulletin de la *Royal Statistical Society*[9].

Il semble qu'au début du siècle, l'enseignement de la statistique en Grande-Bretagne était presque inexistant. Bowley nous dit que les mathématiques occupent une petite place dans les programmes d'économie des universités britanniques[10], et qu'il n'est pas évident que les applications statistiques et la théorie de la probabilité soient inclues dans d'autres programmes d'étude. Il n'était pas non plus facile de trouver en 1906 des mathématiciens capables d'appliquer leurs connaissances aux affaires publiques, et c'est pour cela que Bowley propose de développer un vaste programme qui lie mathématiques et sciences sociales, un programme similaire à celui qui sera aussi mis en place par les fondations nords-américaines.

Comment envisager alors ce programme scientifique ? Une première approche est d'en délimiter le champ de travail, et Bowley trace bien une limite qui sépare les pratiques statistiques anciennes des pratiques et des théories qui caractérisent la nouvelle science statistique. Quelle est la caractéristique fondamentale d'une science ? C'est *la méthode*[11], bien sûr : *une méthode de mesure qui permettra de différencier le vrai du faux, dans tout raisonnement basé sur la construction de tableaux statistiques.* À ce propos, il va reprendre toute la tradition britannique qui depuis le XVII[e] siècle

---

Pavé et les entretiens de G.Th. Guilbaud et de B. Bru.

[8] M. Armatte a remarqué que « *dans les années 1880 se redessine une statistique inférentielle dont les principes et les techniques sont tout différents de la théorie des moyennes, et qui rompt avec le programme déterministe de la première moitié du siècle aussi bien qu'avec son programme probabiliste* », M. Armatte, 1992, p. 119.

[9] A. L. Bowley, 1906, p. 540-558.

[10] Depuis 1856, la section F de l'Association reçoit « le nom curieux » de Science Économique et Statistique. Cette association entre statistique et économie se produit au moment où Quetelet développe à Bruxelles un vaste programme de construction de la statistique comme science sociale positive.

[11] La méthode devient déjà dans le XVII[e] siècle, le noyau dur de la nouvelle science expérimentale. F. Bacon, Descartes, Hobbes, Hooke et tant d'autres avaient une confiance absolue en la capacité de comprendre la nature à condition que l'esprit soit discipliné par la « bonne méthode » (cf. [SHAPIN, 1998]).

présente la méthode expérimentale comme source principale de légitimation de la science.

Les techniques de recueil d'information restent ainsi : *art statistique*, et la nouvelle statistique mathématique devient : *science statistique*[12]. La statistique mathématique ressemble aux sciences naturelles dans le sens où ses développements théoriques ont aussi des applications pratiques, comme peuvent le montrer les études sur le nivellement des impôts ou l'estimation du coût de la vie des travailleurs. A. Bowley présente tout un programme pour faire de la statistique une discipline qui inclut l'idée de *précision*, et pour cela il emprunte à l'astronomie le concept d'*erreur probable*. Il faut dire, en tout cas, que la position de A. Bowley quant aux possibilités de la statistique mathématique est bien modérée :

> *Il faut reconnaître que beaucoup de statistiques sont nécessairement approximatives. En statistique, l'exactitude et la précision consistent à estimer les bornes de l'erreur probables et possibles, et la fausse apparence de la prétendue exactitude mathématique doit être abandonnée*[13].

La mesure statistique est cependant approximative et provisoire, et de même que les autres sciences comme la physique ont recours aux expérimentations pour améliorer leur consistance, la consistance des estimations statistiques peut s'améliorer par des *échantillons* soigneusement sélectionnés. Apparemment, Bowley situe la méthode des échantillons au cœur de la nouvelle science statistique ; il était très conscient des possibilités de cette méthode et de sa capacité à se transformer en puissante arme de recherche scientifique.

Il y a d'autres éléments importants qui vont contribuer au développement de la nouvelle science statistique : l'intérêt croissant des classes populaires pour la statistique, surtout les syndicats et les organisations ouvrières, ainsi que celui de la presse pour la publication des données statistiques. La presse commence à jouer à cette époque un rôle décisif dans l'élaboration du nouveau discours statistique, bien que Bowley condamne *la confiance aveugle et souvent équivoque que l'on concède aux déclarations statistiques*.

Bien que la théorie se soit développée rapidement, et que quelques méthodes aient déjà été utilisées par les biologistes et les botanistes, il y avait peu d'applications relatives aux problèmes pratiques jusqu'au moment où on a commencé à considérer sérieusement le rapport entre les fréquences de la déviation et la loi de l'erreur. Bowley ajoute que la méthode des échantillons disposait de matériaux suffisants depuis les travaux réalisés par Edgeworth en 1885, mais que ceux-ci ont été totalement ignorés.

Le contrôle de la mesure suppose l'estimation la plus probable, prenons l'exemple, d'une moyenne de salaires. Dans ce cas, il propose des énoncés du type 24s.± 6d., adoptant l'écart type comme mesure de sécurité (*dans une courbe normale de*

---

[12] Le terme «□scientifique□», nous dit Shapin, ne sera inventé qu'à partir du XIX^e^ siècle et ne passera dans la langue commune qu'au début du XX^e^ siècle. Au XVII^e^ siècle, le mot «□science□» (scientia) signifiait «□connaissance□» ou «□ntelligence□» et désignait n'importe quel ensemble de connaissances établies.

[13] "*It must be recognised that most statistics are necessarily approximatives; and just as in other scientific measurements the quantity is given as correct to so many significant figures, so in statistics the possible and probable limits of error should be estimated, and the false show of so-called mathematical accuracy given up*", A. Bowley, *op. cit.*, p.543.

*fréquences, deux tiers de l'aire sont dans l'écart type*) Il ajoute: quand cela est applicable, la mesure est très précise.

Il propose deux voies pour appliquer la nouvelle méthode: celle de Pearson, c'est-à-dire, une formule empirique et l'ajustement des observations en déterminant une courbe appropriée de fréquences pour assigner la probabilité des observations ou la voie d'Edgeworth qui consiste à accepter la loi généralisée des grands nombres et à déterminer *a priori* les phénomènes où l'on peut utiliser cette loi. Un exemple avec des données de l'*Investor's Record* et de *l'Almanach Nautique* lui sert à montrer que tous les éléments de l'ensemble doivent avoir les mêmes possibilités de sélection, et que la précision ne dépendra pas de la taille du groupe à échantillonner mais seulement de *sa nature* et de la taille des échantillons. La précision peut cependant être aussi grande que l'on veut, et l'erreur probable et possible, aussi petite que l'on veut, à condition toutefois d'augmenter seulement la taille de l'échantillon.

Il conclut en assurant que, tant que la théorie n'est pas totalement terminée, l'usage de la méthode devra rester pour un temps entre les mains de spécialistes, et qu'il s'agit d'une formule d'application de la théorie de la probabilité parmi d'autres. Il faut cependant remarquer qu'à partir du colloque de l'Institut International de Statistique organisé à Rome en 1925, elle devient la méthode presque universelle dans la nouvelle statistique mathématique.

## STATISTIQUE MATHEMATIQUE ET SOCIOLOGIE

La statistique, trouve finalement sa méthode[14] dans l'application de la théorie des erreurs aux problèmes sociaux, tels le chômage ou les conditions de vie de la classe ouvrière. La Grande-Bretagne devient aussi le foyer privilégié de cette nouvelle science statistique, peut-être par la combinaison d'une riche tradition philosophique inductive qui donne la primauté aux méthodes empiriques et de quelques conditions sociales et politiques. La Grande-Bretagne, par exemple, est le premier pays à créer les indemnités de chômage (1911), ce qui produit des conditions suffisantes à l'application de la méthode des échantillons. Certainement, ce n'était pas la même chose de découvrir les lois des séries des prix que de connaître le nombre de chômeurs à un moment et à un endroit donnés. En 1923, lorsque le Ministère du Travail anglais avait besoin d'une description détaillée des presque deux millions deux cent cinquante mille chômeurs enregistrés, John Hilton, professeur à l'université de Cambridge et alors directeur des statistiques du Ministère du Travail britannique, s'est rendu compte de l'utilité de la méthode représentative et de l'urgence d'avoir des entretiens directement avec les chômeurs. Les enquêtes sur les conditions de vie de la classe ouvrière rendent alors possible la construction de la nouvelle méthode statistique.

Les protagonistes de ce processus, par contre, ne seront pas des sociologues, mais des économistes, ou plus exactement, des mathématiciens devenus économistes. Pour comprendre ce processus à l'intérieur du champ de la sociologie, il faut se tourner vers les États-Unis et y découvrir deux personnages clés, F. Ogburn et Samuel A. Stouffer. Le premier, auteur de nombreuses publications sur l'opinion publique, ouvre le débat sur

[14] À partir du XVII$^{e}$ siècle, la méthode devient le mythe de la science ; capable de transformer tout ce qu'elle touche en science, elle est probablement à l'origine de plusieurs œuvres emblématiques comme celle de Durkheim *Les règles de la méthode sociologique*, parue en 1885. Vid. note 21. La statistique, considérée durant la seconde moitié du XIX$^{e}$ siècle comme science sociale, devient une science indépendante au début du XX$^{e}$ siècle sous l'influence des statisticiens britanniques.

la méthode statistique en 1929, dans une séance de *l'American Sociological Society*, dans laquelle il lance un appel à l'abandon des procédures «non scientifiques», pour une pratique statistique quantitative et objective en matière de prédiction et d'ingénierie sociales. Le second, docteur en 1930 à l'Université de Chicago, est formé sous l'influence de Thurstone et Ogburn, mais fait aussi un stage post-doctoral à l'université de Londres avec K. Pearson et R. Fisher. Durant la seconde guerre mondiale, il dirige le *Research Branch Information and Education Division* au Département de la Guerre, où il est responsable d'une des enquêtes les plus impressionnantes qu'on ait jamais faite. À cette recherche a participé une autre étoile montante de la sociologie statistique, le jeune astrophysicien viennois, devenu sociologue : Paul Lazarsfeld.

En Espagne, les enquêtes sur les conditions de vie de la classe ouvrière sont effectuées à l'Institut de Réformes Sociales entre 1903 et 1924 par des sociologues ayant une formation juridique, mais qui exécutent des travaux importants dans la lignée des monographies de Leplay, et établissent des indices du coût de la vie des ouvriers. Parmi les travaux des sociologues, on retiendra simplement l'enquête démographique faite par Severino Aznar à l'aide de la méthode représentative, avec les données du recensement de 1920[15].

Plus riche, sans doute, est la situation en France, avec la figure de M. Halbwachs, récemment sortie de l'oubli par la *Revue d'Histoire des Sciences Humaines*[16]. Néanmoins, l'idée s'est répandue qu'Halbwachs avait une attitude ambivalente vis-à-vis de la statistique ; ainsi, O. Martin, dans son article *Raison Statistique et Raison Sociologique chez Maurice Halbwachs*[17] écrit que bien qu'il soit le sociologue français le plus intéressé par les mathématiques, et celui qui avait la meilleure formation statistique, il donne aux mathématiques une valeur relative. Selon O. Martin, pour Halbwachs, les mathématiques ne servent «qu'à établir des faits que le sociologue doit ensuite interpréter et expliquer, et qui doivent être replacés dans leur contexte social précis»[18]. Nous pensons pour notre part, qu'Halbwachs critique seulement les excès et les abus qui commencent à se succéder dans le domaine des sciences sociales à partir de l'introduction des techniques mathématiques. Son intérêt pour la statistique mathématique est déjà suffisamment justifié par les études qu'il a faites sur les conditions de vie de la classe ouvrière et par son approche des problèmes de calcul des probabilités[19].

---

[15] C'est une analyse démographique comparative à Madrid. À partir des registres de la *Diputación Provincial* ont été sélectionnées 7 775 fiches de familles de classe moyenne (artisans, petits industriels, commerçants et agriculteurs) ; 18 670 fiches de familles de professions libérales (employés publics, privés, écrivains, journalistes, acteurs, professeurs, avocats) ; 524 fiches de familles de milieux aisés (ceux qui ont un salaire entre 4 000 et 5 000 pts, qui payent des impôts entre 7 501 et 10 000 pts., ou loyer de 3 000 à 8 000 pts. ; 387 fiches de classes nobles (familles qui apparaissent dans le guide de la *Sociedad de Madrid* et de la *Grandeza de España*, dont le chef a un titre de noblesse ou appartient à un ordre militaire). Il s'agit d'une étude de style darwiniste sur la reproduction des classes sociales, moyennes et aisées ; en fait, l'échantillon est utilisé pour calculer le pourcentage différentiel des naissances, mortalités, et de la reproductibilité des classes sociales (cf. S. Aznar, « La familia vista por un demógrafo », *Estudios Demográficos* nº V, Instituto Balmes de Sociología, CSIC, Madrid, 1962, p. 111-112.

[16] Revue d'Histoire des Sciences Humaines, nº 1, 1999, *Maurice Halbwachs et les sciences humaines de son temps,* Septentrion, Presses Universitaires. Voir aussi, le nº 6, 2002, *Mathématiques et sciences sociales au cours du XX^e siècle*.

[17] Revue d'Histoire des Sciences Humaines, 1999, p. 69-103.

[18] *Ibidem,* p.69.

[19] Nous ne devons pas oublier qu'en 1912 il présente une thèse complémentaire à l'Université de Paris sur Quételet et la statistique morale où, en plus de critiquer la théorie de l'homme moyen, il fait une

En 1935, il participe au colloque sur la statistique organisé à Paris par le Centre International de Synthèse[20], afin de présenter les problèmes consécutifs à l'application de la statistique aux faits sociaux. Il commence un exposé intitulé *La Statistique en Sociologie* en disant que la statistique a été découverte dans le domaine des sciences sociales et a été définie comme la science des moyennes. À ce propos, il fait la différence entre les moyens utilisés par les sciences de la nature et ceux des sciences sociales où la statistique s'applique. Il utilise la définition de Durkheim du fait social et donne l'exemple de la durée de vie moyenne qui lui permet de faire la première affirmation importante□ *tout dénombrement n'est pas une statistique* ; il est nécessaire que le groupe présente une certaine consistance. La moyenne des tailles, des prix recueillis au hasard, des salaires, des loyers, ne sont pas vraiment des statistiques ; en cela il suit les idées de Simiand.

Il s'interroge sur les différences entre les observations faites dans le domaine de la biologie, où les êtres vivants sont des entités réelles, et les observations faites dans le domaine des faits sociaux. Les faits d'observation d'êtres vivants sont constitués par des éléments presque identiques, à la différence des phénomènes sociaux : *Au fond, il n'y a d'ensembles réels que les groupes sociaux, précisément, parce qu'ils sont constitués par des éléments différents. Tous les autres ensembles sont des collections*[21].

Ici le problème des statistiques devient aussi complexe□ la population d'un département n'est pas un groupe social ; une période quinquennale n'est pas une période socialement définie. D'après les propositions de Simiand d'étudier le phénomène de façon continue, il se demande si les groupes d'âge, tels que les statisticiens les utilisent, sont une réalité sociale. Nous pouvons penser d'abord que l'origine de la méfiance de Halbwachs vis-à-vis de la statistique est l'utilisation qui en est faite par certains mathématiciens. Il parle par exemple, de l'excessive *rigueur arithmétique* qui commence à se répandre dans le domaine des sciences sociales, et qu'il considère comme *un peu artificielle et arbitraire*.

Halbwachs considère qu'il faut rendre compte de la diversité des classes sociales, des professions, des différences entre milieu rural et urbain, et se demande si la pyramide des âges est la même en milieu urbain et rural, ou pour les riches et les pauvres. Par exemple, la proportion d'adultes aux États-Unis est la même qu'en France ; ce n'est pas que la natalité y soit faible comme en France, mais beaucoup d'immigrés débarquent dans ce pays à l'âge adulte.

Après toutes ces considérations, voyons en quoi consistent les réserves de Halbwachs concernant l'utilisation que font les mathématiciens de la statistique. Nous

---

présentation du calcul des probabilités avec des références à Quételet (*Instructions populaires sur le calcul des probabilités*), à Borel (*Eléments de la théorie des probabilités*), à Poincaré (*Science et Méthode*), et à Bertrand (*Calcul des Probabilités*).

[20] *La Statistique. Ses applications. Les problèmes qu'elles soulèvent*, Septième Semaine Internationale de Synthèse, 1935, publié en 1944 sous l'occupation allemande, Presses Universitaires de France.

[21] *Tendances, croyances et pensées collectives sont représentées inégalement, diversement, par chaque individu : chacun n'en pressent qu'une partie, ou qu'un aspect. C'est pourquoi il n'y a qu'un moyen d'atteindre l'état collectif : c'est de rassembler toutes ses parties, de les dénombrer intégralement, de façon à ne négliger aucune d'entre elles, et de recomposer l'ensemble. Au fond, il n'y a d'ensembles réels que les groupes sociaux, précisément parce qu'ils sont constitués par des éléments différents. Tous les autres ensembles sont des collections. Les espèces vivantes sont des collections d'organismes, les organes et tissus sont des collections de cellules, et l'organisme lui-même n'est qu'un individu. Les groupes sociaux qui sont plus encore, et autre chose*□ *op. cit.*, p. 116.

avons déjà vu la définition de la statistique comme science des moyennes et des courbes ; mais il s'agit d'un point de départ pour une définition satisfaisante. Á la différence des sciences physiques où le phénomène ne se produit pas si on supprime quelques aspects fondamentaux durant l'expérience dans la statistique, *les chiffres se laissent toujours combiner avec des chiffres*[22]. Ce qui pose une fois de plus la question du rapport entre les modèles mathématiques et la réalité… *de même que l'homo oeconomicus, un tel homo demographicus est une abstraction trop soigneusement détachée de la réalité pour nous apprendre quoi que ce soit sur le réel*[23].

Concernant le respect de l'utilisation des courbes et graphiques, la position de Halbwachs est aussi claire : *Quant aux courbes, elles doivent subir tous les replis du phénomène, le représenter dans toutes ses phases, mais aussi l'embrasser dans toute son étendue, et en toutes ses parties*[24]. Mais surtout, ce qui nous intéresse est la critique qu'il fait de l'utilisation de la courbe la plus fameuse, parce qu'il cite R. Gibrat, ancien élevé de l'École Polytechnique et économiste réputé à l'époque, qui dit à propos des inégalités : *Notre loi est essentiellement statistique. Elle ramène simplement les courbes des répartitions économiques à une courbe célèbre, la courbe en cloche, dite aussi de Gauss ou des erreurs*[25].

Halbwachs se demande si c'est *bien vraiment l'idéal de la recherche statistique que de ramener les faits économiques et sociaux, leurs mouvements et leurs variations, à telle ou telle courbe avec laquelle les mathématiciens sont familiers ?*[26] Il anticipe cette pratique qui a fini par se généraliser et, à mon avis, c'est ce qui a empêché une interprétation correcte de la méthode des échantillons pendant plusieurs années et qui a contribué à la confusion des méthodes à l'intérieur des sciences sociales. De surcroît, ces mauvaises pratiques ont contribué aux disputes académiques entre spécialistes en sciences sociales ayant une formation mathématique, et ceux qui n'en ont pas :

> *Certes, l'ingéniosité des algébristes et des géomètres est grande. Ils disposent des nombreux procédés d'ajustement, ils savent introduire dans leurs formules divers paramètres, si bien qu'en faisant les conventions nécessaires il est toujours possible de ramener les courbes observées à des courbes théoriques de forme connue*[27].

Il critique aussi U. Yule, qui présente les courbes de population de pays aussi différents que l'Angleterre, la France et les États-Unis, et les transforme, par des ajustements mathématiques, en une seule courbe théorique avec une seule formule. Halbwachs est mordant à ce sujet : *des courbes de ce type ne nous apprennent rien sur le mécanisme interne*[28]. Les rapports entre différentes séries nous proposent une image

---

[22] *Op.cit.,* p. 124.
[23] *Op.cit.*, p. 123.
[24] *… c'est ainsi qu'on représentera les mouvements des salaires par plusieurs courbes, continues autant que possible, juxtaposées, aussi nombreuses qu'il y a de données correspondant à des groupes différents, agriculture, industrie, et diverses espèces et formes d'industrie, grandes villes, villes moyennes, petites villes. Cela exige un effort d'attention multiple, à la fois abstraite et concrète. Mais cette méthode empirique est la seule qui permette de rester en contact aussi étroit que possible avec la réalité. Op.cit*, p. 124.
[25] *Ibidem.*
[26] *Ibidem.*
[27] *Op cit.*, p. 125.
[28] *Ibidem.*

approximative, mais c'est *avec l'étude de ce jeu de rapports, que commence, et commence seulement, la recherche positive*[29].

Il poursuit sa critique avec les *Cours de statistique* de A. Aftalion avec lequel il commence une polémique à propos des altérations de prix produits depuis la fin du XIX[e] siècle. A. Aftalion soutient que deux siècles et demi sont insuffisants pour établir une régularité. F. Simiand, au contraire, parce qu'il regarde le mécanisme complexe, a soutenu qu'une ou deux répétitions suffisent pour établir la réalité de ces grands cycles. De cette façon, deux expériences peuvent suffire quand elles montrent toute une complexité de facteurs mesurés avec précision⊡

> *C'est, à vrai dire, d'une toute autre façon qu'il faut concevoir l'utilisation des méthodes statistiques en sociologie quantitative. Elles ne nous apportent pas des théories, mais des instruments d'observation et de comparaison, à la fois précis et objectifs, et c'est dans cette direction qu'il faudra de plus en plus les développer*[30].

À ce propos, il recommande d'analyser *sans entrer dans les détails techniques* les coefficients de dépendance et de corrélation.

## 3. LA STATISTIQUE MATHEMATIQUE EN ESPAGNE :1914-1936

Le 14 avril 1931 est proclamée la seconde république espagnole et, cinq mois plus tard, le 15 septembre 1931, l'Institut International de la Statistique (I.I.S.) célèbre sa vingtième session dans un Madrid déjà républicain. La réunion était prévue depuis 1923 et, l'on peut supposer qu'au moins depuis cette date, la statistique espagnole avait déjà une certaine présence internationale. Du point de vue des apports théoriques, la réunion n'était pas importante, elle n'aura pas la portée de celles tenues à Paris en 1909 ou à Rome en 1925, mais elle nous permet de présenter quelques figures principales de la statistique mathématique espagnole.

Des personnalités de la statistique internationale y assistent, un détail qui ne passe pas inaperçu auprès des autorités et qui sera mis à profit pour la présentation internationale du nouveau régime républicain. Les sessions furent inaugurées par le président de la République D. Niceto Alcalà Zamora. Y participèrent entre autres : A. Delatour, Président de l'I.I.S., M. Huber, directeur de la statistique générale de France, les Français L. March, F. Simiand et M. Girard, professeurs de l'École des Sciences Politiques de Paris ; C. Gini, professeur à l'université de Rome, M. A. Julin, chef de la délégation statistique belge, J. Hilton, professeur de l'université de Cambridge et directeur des statistiques du Ministère du Travail britannique, A. L. Bowley, professeur de statistique à l'université de Londres, et A Wilcox, professeur d'économie et de statistique à l'Université de Cornell.

Le groupe des Espagnols était formé de H. de Castro, Directeur de l'institut géographique statistique et président du comité organisateur, de J. Gichot, chef du service de statistique du Ministère du Commerce, de O. F. Baños, vice-président du colloque, professeur de géométrie à l'Université de Saint-Jacques-de-Compostelle et sous-directeur du nouveau service d'études de la Banque d'Espagne, de A. de Miguel,

---

[29] *Ibidem.*
[30] *Op. cit.*, p. 127.

mathématicien et directeur des services statistiques de la direction générale de la dette publique, ainsi que de J. A. Vandellos, directeur de l'Institut d'Études de Barcelone.

La plupart d'entre eux – de jeunes professionnels que nous pouvons situer dans la génération dite de 1927, même si certains plus âgés sont de celle de 1914 – sont déjà liés aux activités de la *Junta de Ampliación de Estudios*[31] et de la *Residencia de Estudiantes*, dans un milieu intellectuel qui rend possible l'apparition de personnalités comme celle du philosophe Ortega y Gasset, du physicien Esteban Terradas (génération de 1914), du mathématicien Rey Pastor (génération de 1914), des jeunes Fernandez Baños ou Antonio de Miguel, et de célébrités aussi importantes que le poète Garcia Lorca, le peintre Dali, le cinéaste Luis Buñuel, le biologiste Severo Ochoa, etc., et tant d'autres, représentatifs de cette brillante génération de 1927[32]. Bien que très tôt perturbée par la guerre civile, la Seconde République représente alors l'apogée d'un processus, de modernisation de la société et des institutions espagnoles.

Deux générations emblématiques y participent : celle de 1898 – l'année de la perte de Cuba et des Philippines dans une lutte inégale contre les États Unis – début d'un vaste mouvement intellectuel de régénération du pays qui compte dans ses rangs des poètes brillants comme Antonio Machado et le philosophe Miguel de Unamuno ; celle de 1927, l'autre génération qui jouera un rôle dans le changement général du pays, et à mon avis, la mieux préparée pour affronter les changements de type technique et professionnel dont le pays avait besoin.

Les années vingt représenteront, alors, une conjoncture très favorable aux changements, grâce au statut de neutralité dont avait profité le pays durant la guerre européenne. La production industrielle et économique augmente alors de manière considérable, la population s'accroît de deux millions entre 1920 et 1930[33], et 42 % de la population vit déjà dans des noyaux urbains de plus de 10 000 habitants ; Madrid approche alors le million d'habitants et Barcelone le dépasse. Le début des réformes sociales en matière de logement, d'assurances, de conditions de travail, de régulations de conflits, etc. permet d'envisager le futur avec un certain optimisme.

C'est dans ce contexte que se produit tout un mouvement de profonde rénovation générale de la mathématique[34] impulsée par le mathématicien Rey Pastor, où apparaît la

---

[31] *La Junta de Ampliación de Estudios* et la *Residencia de Estudiantes* sont des institutions créées pour développer la recherche scientifique, et constituent l'antécédent du CSIC (on pourrait dire avec beaucoup de prudence, l'équivalent actuel du CNRS français).

[32] La fin de la première guerre mondiale signifia un changement d'époque, et les années vingt représentent la naissance d'un modèle de société qui va caractériser tout le XX$^{e}$ siècle. Du point de vue de l'économie, cela impliquait de substituer, aux marchés autorégulateurs, des politiques économiques dans lesquelles l'État commençait à intervenir sur différents facteurs (la terre, les marchés agricoles, le travail, les réglementations du travail, et le capital) participant activement à la création de monopoles de services comme l'électricité, les communications, le pétrole, etc., tout cela, coïncide avec l'apparition de la puissance économique nord-américaine. C'est aussi le début de la consommation de masse et de ce que l'on nommera plus tard la société de consommation. Pour l'Espagne, cf. J. M. Arribas, «Antecedentes de la Sociedad de Consumo en España : de la Dictadura de Primo de Rivera a la II República», *Política y Sociedad* 16, 1994, p. 149-168.

[33] La baisse de la mortalité et de la natalité caractérisent les nouvelles tendances démographiques. Le pays commence à envisager des transformations sociales que certains historiens comme Slomo Ben Ami ont considérées comme équivalentes à celles des années soixante.

[34] Sur la situation des mathématiques en Espagne cf. M. Hormigon, *Las matemáticas en España en el primer tercio del siglo XX,* en Sánchez Ron JM, *Ciencia y Sociedad en España*, Ediciones El arquero,

statistique mathématique espagnole. Des associations et des revues mathématiques voient le jour, le laboratoire de mathématique de la *Junta de Ampliación de Estudios* est créé, et c'est le début de vastes contacts internationaux[35].

La Société Mathématique Espagnole avait été créée en 1911 dans cet esprit de renouvellement. Dans le premier numéro de sa revue□ la *Revista de la Sociedad Matemática Española* sont publiés déjà des articles d'Esteban Terradas, la première figure importante de la statistique mathématique espagnole. Cette revue, bien qu'ayant 423 abonnés et une économie saine, disparaît six ans plus tard avec le premier voyage de Rey Pastor en Argentine[36]. Après son retour en 1918, il reprend le projet, et l'année suivante lance un nouveau projet éditorial : la *Revista de Matematica Hispano Americana* qui, malgré de notables collaborations argentines, sera fondamentalement espagnole, et où la jeune génération commence à publier. Associée à la signature de certains mathématiciens déjà vétérans comme Terradas ou Alvarez Ude, apparaissent dans les pages de la revue de plus jeunes auteurs comme Fernandez Baños, Santalo, Orts Aracil ou des étrangers comme Hadamar, Hilbert, Klein, Levi-Civita, etc. La *Revista* obtient une reconnaissance internationale et, avec l'arrivée de la seconde République, la création d'une seconde revue devient possible : *Matemática Elemental*, destinée aux cercles d'étudiants en mathématiques d'Argentine et d'Espagne. Le laboratoire des mathématiques de la *Junta de Ampliación de Estudios* sera donc l'endroit privilégié d'où vont sortir la plupart des mathématiciens spécialistes en statistique mathématique.

Le premier mathématicien s'occupant de statistique dans cette période est sans doute Esteban Terradas. Avec une formation de physicien et d'ingénieur, il est le principal initiateur de la physique nucléaire en Espagne (physique de l'atome, disait-on à l'époque). Terradas est co-directeur avec Rey Pastor du laboratoire de mathématique, et possède un curriculum extraordinaire. Professeur de l'École Supérieure Aéronautique, directeur général de la Compagnie Téléphonique, directeur des travaux souterrains du chemin de fer de Barcelone, professeur à la Faculté des sciences, membre de l'Assemblée Nationale... voilà quelques-unes des activités développées par cette personnalité surprenante. Sa capacité est aussi saluée par des historiens de la Mathématique comme Mariano Hormigon□*première figure importante de la statistique mathématique espagnole, un des scientifiques espagnols les plus importants de la première moitié du siècle,...*[37].

---

CSIC, 1998.

[35] À propos du programme de rénovation des mathématiques mis en place par Rey Pastor, Sixto Rios, dans l'hommage qu'il lui rend, remarque que *vingt-cinq ans de travail exemplaire suffirent pour que l'anathème qui paraissait exister sur la capacité de* «□homo hispanicus□ *à faire des mathématiques soit aboli*. On peut affirmer objectivement que dans les années trente en Espagne, il existe déjà une culture mathématique contemporaine avec des apports originaux au niveau européen. C'est alors que Rey Pastor et ses disciples directs et indirects publient des travaux importants dans les principales revues internationales : *Comptes Rendus de l'Académie des Sciences* de Paris, *Acta Matematica, Ergebnisse eines Mathematischen kolloquiums* de Vienne, *Abhandlungen* de Hambourg, *Rendiconti* de Palerme, *Mathemastische Zeitschrift,* etc. Nombre de théorèmes et de théories issus de ces travaux passent dans les livres des universités européennes, comme dans les travaux de Doetch, S. Mandelbrojt, A. Denjoy, W. Hurewizc, Wilder, W. Blaschke, Menger, etc. R. Pastor Selecta, Fundacion Banco Exterior, 1988.

[36] Le jeune mathématicien, élève d'Eduardo Torroja et de Zoel Garcia de Galdeano, commence à fustiger à partir de 1915 (à l'âge de 27 ans) les consciences de ses collègues mathématiciens en proposant un vaste programme de rénovation des mathématiques qui, à partir de 1918 commencera à donner d'importants résultats, cf. M. Hormigon, *op. cit.*

[37] M. Hormigon, *op. cit.*, p. 274

Il naît à Barcelone en 1883, cinq ans avant Rey Pastor et fait ses études secondaires près de Berlin, ce qui lui permet de dominer assez tôt la langue allemande. Il revient à Barcelone et obtient la licence d'ingénieur industriel et celle de sciences physico-mathématiques. En 1904, il obtient à Madrid deux doctorats en sciences exactes et en sciences physiques, le premier sur les *mouvements des fils selon les courbes* et le second sur *l'absorption de la lumière par les corps cristallins*. Cette même année, il devient professeur auxiliaire à l'université Centrale de Madrid, et l'année suivante obtient une chaire de Mécanique rationnelle à Saragosse. En 1907, il obtient la chaire d'optique et d'acoustique de l'Université de Barcelone où il travaillera pendant vingt ans. Il partage son activité académique et scientifique avec la direction et la gestion de projets d'ingénierie. Il crée également un séminaire physico-mathématique et un Institut d'Electricité et de Mécanique à l'Université de Barcelone, édite des monographies scientifiques, publie dans l'Encyclopédie *Espasa* des articles scientifiques (par exemple la rubrique «□robabilité□) qui sont de vrais traités.

À partir de 1926, il commence à donner des cours dans les écoles d'ingénieur de Madrid, et en 1927, en Argentine, en Uruguay, au Chili et au Pérou. Cette même année, il est nommé membre de l'Assemblée Nationale par le gouvernement de Primo de Rivera, ce qui l'oblige à passer plus de temps à Madrid. Dans la capitale, il donne des cours et des conférences à l'Académie des sciences, à la Société mathématique espagnole ; il participe aux activités de la *Junta de Ampliación de Estudios* et de la Résidence d'étudiants où il invite des professeurs étrangers comme Levi-Civita, Weyl et Einstein. Membre des associations scientifiques d'Europe et d'Amérique, il suscite des commentaires élogieux de la part d'Einstein, et H. Weyl lui dédie son œuvre *Mathematische Analyse des Raumproblems.*

Quant à ses orientations idéologiques, il se consacre à l'action sociale et à l'instruction des jeunes ouvrières dans des institutions comme le patronage de la *Sagrada Familia* de Barcelone et le patronage de *San José* de Madrid, ce qui indique un engagement en direction du catholicisme social.

Concernant sa contribution à l'enseignement de la statistique mathématique, il est le premier à donner des cours à l'université centrale et a suscité la création de la première chaire de Statistique mathématique de la faculté des Sciences, qu'obtient Fernandez Baños en 1934. En 1931-32, il donne un cours à l'université centrale dans lequel il utilise les bibliographies de G. Darmois, ainsi que les textes de Von Mises. Cette même année, il est élu membre de l'Union Mathématique Internationale, et en 1933, donne des cours à la faculté de Droit de Madrid sur *la théorie des échantillons*.

## OLEGARIO FERNANDEZ BAÑOS ET LA NAISSANCE DE L'ÉCONOMÉTRIE.

L'économie mathématique est l'autre domaine important où va se développer la statistique mathématique, et surtout, les études sur les fluctuations du change de la peseta jointes au rapport de la Commission de l'étalon d'or. Olegario Fernandez Baños est ici la figure principale, non la seule, mais la plus emblématique.

Fils de modestes paysans de la Rioja, région du nord de l'Espagne, il naît en 1886, et étudie au séminaire catholique de Logroño de 1902 à 1906 comme il était d'usage chez la plupart des agriculteurs qui n'avaient pas les moyens de payer les études de leurs enfants. Il poursuit ses études de théologie, mais n'ayant aucune vocation religieuse, il abandonne les études ecclésiastiques. En 1908, il s'inscrit à l'Institut général et technique de Logroño pour préparer le baccalauréat, et en 1909 il réussit les examens

pour travailler dans le corps des télégraphistes. Sa nouvelle situation professionnelle lui permet de venir à Madrid, s'installant définitivement dans la capitale une fois terminés les cours de formation de télégraphiste. Il passe le baccalauréat et commence une carrière de sciences exactes à l'Université Centrale en 1910.

À Madrid, il fait la connaissance de Rey Pastor, disciple des mathématiciens Zoel García de Galdeano, Ventura Reyes Prosper et Eduardo Torroja. Avec Rey Pastor, il partage ses repas et des discussions mathématiques dans une petite pension ; de ce fait, celui-ci devient son protecteur et son principal mentor. Après avoir passé en tant qu'étudiant deux ans à Barcelone, où il fait la connaissance de Torroja et où il travaille aussi comme télégraphiste, il revient à Madrid en 1914 comme assistant de classes pratiques d'analyse mathématique.

En 1915, il commence à enseigner l'algèbre et les mathématiques supérieures à l'École industrielle de Valladolid et, cette même année, il obtient le diplôme de docteur en sciences exactes avec une thèse dans la ligne de la géométrie fondamentale de Hilbert intitulée *Construction des espaces complexes contenus dans* $E_n$ *et leurs représentations réelles*.

Deux ans plus tard, il part à Zurich avec une bourse de *la Junta de Ampliación de Estudios*. En 1917, il travaille au Séminaire mathématique de l'Institut Polytechnique de Zurich et s'inscrit à un cours de géométrie projective, à un deuxième de théorie des assurances et un troisième d'équations différentielles. Il assiste à plusieurs colloques sur des sujets d'actualité de l'époque comme la physique mathématique et la théorie de la relativité. Cela le conduit à rencontrer les plus grands de la mathématique de cette époque comme H. Weyl, l'assistant de F. Klein et O. Hilbert, qui lui recommandent d'étudier la physique mathématique (mécanique et électromécanique). Finalement, il convainc Rey Pastor de la nécessité d'aller à Bologne pour étudier la géométrie non-euclidienne avec le professeur F. Enriquez.

À son retour, en janvier 1918, il continue son enseignement à Valladolid et commence à préparer le concours de la chaire de géométrie analytique de l'Université de Saint-Jacques-de-Compostelle, qu'il obtiendra en 1921. Sa résidence en Galice ne l'empêche pas de participer aux activités du Laboratoire Mathématique de la *Junta de Ampliación de Estudios* où il avait publié sa thèse.

En 1920, il séjourne à nouveau en Italie, à Bologne,où il assiste aux réunions scientifiques avec des mathématiciens, des juristes et des sociologues, des historiens des mathématiques, etc. Commençant à s'intéresser à la politique et à la sociologie, il assiste aux congrès du Parti populaire et du Parti socialiste.

En 1923, il obtient une bourse pour étudier l'économie mathématique et l'économie financière à Paris et à Bordeaux et, en 1923-1924, fait un autre voyage en Italie pour étudier la statistique mathématique et le calcul des probabilités. Son mémoire final, qui justifie son séjour à Rome, a pour titre *Notes pour une première étude de la théorie du risque*. Ce passage des mathématiques à l'économie étonne aujourd'hui les économistes qui travaillent sur l'histoire de la pensée économique. Salvador Almenar[38] soutient que c'est l'installation définitive de Rey Pastor en Argentine et les disputes avec les mathématiciens plus conservateurs de l'université qui aurait décidé Fernández Baños à s'orienter vers le champ de l'économie, mais ce n'est pas évident. Il est plus

[38] S. Almenar, 2002, p. 592.

satisfaisant de penser que Fernández Baños n'a jamais abandonné les mathématiques et est devenu statisticien à une époque où la statistique mathématique émerge avec deux champs d'application fondamentaux : la physique mathématique et l'économétrie. Esteban Terradas, vice-directeur du Laboratoire de mathématique, jouera un rôle fondamental dans le domaine de la physique, autant que Fernandez Baños dans le domaine de l'économie. Pour comprendre le glissement vers le social de quelques personnalités de cette époque, il suffira de considérer les convulsions historiques du moment : la première guerre mondiale, la naissance de l'État soviétique, la création des partis communistes, l'apparition du fascisme et des fronts populaires, etc. Un exemple de l'intérêt pour l'économie est la création de centres de recherche économique quantitative, surtout l'Institut de recherches économiques de Barcelone et le Service d'études de la Banque d'Espagne créés en 1930.

Les premières productions économiques de Fernandez Baños datent de 1924-25[39] : « Notions fondamentales d'économétrie mathématique et quelques-unes de ses applications » et « Note sur la décomposition des courbes représentatives des phénomènes économiques ». La première est une synthèse de l'économie marginaliste néoparetienne, et la seconde une application de l'analyse harmonique basée sur les séries de Fourier[40]. Dans un article de 1926, il aborde la distribution des revenus, un thème lié à l'intérêt des états européens pour l'organisation de la fiscalité ; d'où l'importance que l'on donne aux débats théoriques sur la distribution des revenus (V. Pareto). Dans un autre article sur le « problème fiscal », Fernández Baños défend la progressivité de l'impôt sur les revenus, et se sert de la loi de redistribution de Pareto, arrivant à la conclusion que les politiques de distribution sans croissance économique peuvent aggraver la pauvreté et l'inégalité.

Entre octobre et décembre 1927, il fait un autre séjour en Suisse et en Italie. Il commence à publier des travaux sur la peseta, les prix et la parité économique qui lui serviront à obtenir en 1930 le poste de sous-directeur de la Banque d'Espagne. Cette même année, l'administration de l'État va lui offrir d'autres postes de responsabilité importants : il est nommé chef du service des indices économiques du Ministère de l'Économie, et l'année suivante, il est chargé au Ministère du Travail et de la Prévision de former le personnel du corps national de la statistique pour l'élaboration des tables de mortalité.

À partir de 1927, il se rapproche de « façon accélérée et surprenante » de la littérature anglo-saxonne (Henri L. Moore, Irving Fisher, Henry Schultz, Griffith C. Evans, Roos)[41]. Cette année-là est publié « Recientes progresos de la ciencia económica » (Récents progrès de la science économique), inspiré, d'après Almenar, du travail d'Henry Schultz « Mathematical economics and the quantitative methods » et du texte de Charles F. Roos « A dynamical theory of economics ». L'économie mathématique dynamique et la statistique constituent à partir de là les piliers de l'économie, ainsi que d'autres axes d'intérêt scientifique comme la cinématique économique (théorie monétaire de Fisher), et l'analyse des oscillations économiques (baromètres de Harvard, nombres index...).

---

[39] En 1924, Antonio de Miguel publie une *Introducción a la metodología estadística* et Jose Antonio Artigas introduit l'enseignement de la statistique mathématique à l'École des Ingénieurs Industriels.
[40] S. Almenar, *op cit.*, p. 594.
[41] *Op. cit.*, p. 603.

Les études sur les fluctuations du change de la peseta annexées au rapport de la Commission de l'étalon or constituent la naissance de l'économétrie en Espagne, et les années 1930-1931 marquent une vraie rupture. En effet, si les textes antérieurs de Fernandez Baños sont des études pratiques qui tentent d'expliquer l'évolution du type de change de la peseta en utilisant des instruments statistiques de traitement et de décomposition des séries chronologiques, à partir de cette date, il s'agit de réflexions et de considérations méthodologiques sur la corrélation, d'études de la distribution des erreurs d'estimation dans le cas des ajustements de séries temporelles et de la notion de causalité en essayant d'utiliser des variables économiques[42].

En 1932, il participe au Colloque de l'*Econometric Society* à Paris avec le rapport *Contribution aux Index Numbers*. Il n'est pas prouvé qu'il était membre fondateur de la société, comme il le prétendait, mais en tout cas, il apparaît dans les documents de 1934 avec Antonio de Miguel et Jose Antonio Vandellos.

## 4. L'ENSEIGNEMENT ET QUELQUES PROBLEMES METHODOLOGIQUES

En Espagne, l'enseignement du calcul des probabilités avait déjà commencé à l'intérieur des académies militaires (D. Diego Ollero) et des écoles d'ingénieurs (José Antonio de Artigas), mais c'est à la Faculté des sciences de Madrid que va commencer l'enseignement de la statistique mathématique. En 1931, Esteban Terradas commence à donner des cours et, en 1934, Olegario Fernandez Baños obtient par concours la chaire de statistique mathématique de la Faculté des sciences.

Dans son mémoire écrit pour passer le concours, il utilise deux lettres de grands statisticiens comme Corrado Gini et Ronald Fisher où nous pouvons voir les deux conceptions de la discipline : la considération de la statistique comme science, la position anglo-saxonne, et la considération de la statistique comme ensemble d'outils mathématiques pour toutes les sciences, c'est plus la position française et italienne.

Pour Gini, ce domaine est surtout constitué d'un ensemble d'outils techniques, en tout cas, il utilise le terme comme acquis : à *mon avis, il n'existe pas de discipline qui puisse s'appeler statistique mathématique. Il existe seulement une statistique méthodologique qui nécessite le recours aux mathématiques pour traiter à fond beaucoup de ces problèmes*[43]. Il fait la différence entre les étudiants des facultés des sciences – par exemple mathématique, physique et chimie – avec lesquels on peut utiliser beaucoup de mathématiques, mais il ajoute qu'il serait *très dangereux de traiter seulement les problèmes mathématiques que présentent les statistiques en les séparant des problèmes logiques et des notions techniques et pratiques indispensables à la statistique*. Pour les étudiants en sciences humaines et en sciences sociales, il recommande un cours d'introduction mathématique à la statistique.

Ronald Fisher, dans la lignée de Bowley, souligne la méthode inductive de la statistique, et sa vocation à être science exacte⊡ *La statistique mathématique a des points communs avec d'autres études mathématiques qui tendent à développer une méthode de raisonnement exacte, concise et générale*[44]. La méthode déductive propre

[42] S. Almenar, *op. cit.*, p. 608.
[43] O. Fernandez Baños, *Programa, concepto, método y fuentes de Estadística Matemática*. Talleres Gráficos Marsiega, Madrid, 1941, p. 21.
[44] *Op. cit.,* p. 22.

aux mathématiques peut être utilisée *lorsque notre pensée va de la multitude à l'individu*, mais toujours de *façon subalterne*.

Il propose de familiariser les étudiants avec cette distinction fondamentale à travers les tests de signification comme le *t* de Student, le khi 2 et le *z*. C'est, en général, un programme que Fernandez Baños va suivre dans ses cours, et que nous pouvons voir dans son *Tratado de Estadistica,* publié après la guerre. Fisher commence par la description des phénomènes collectifs comme principal objet de la statistique, pour passer immédiatement au calcul des probabilités. Il n'y a encore aucune partie qui explique la théorie, ou la technique des échantillons comme dans le manuel de 1944, mais il y a des titres comme *statistiques adéquates pour estimer les paramètres collectifs*, *consistance des estimations statistiques,* et *précision, exactitude et quantité d'information appliquées à l'estimation statistique*.

En général, Fisher, comme Gini, considère qu'il n'est pas nécessaire dans les programmes de physique et chimie de consacrer beaucoup de temps à la statistique, à l'exception de la mécanique statistique. Il termine par une recommandation surprenante :

> *Pour l'étudiant de mathématiques, il est particulièrement important que le champ de ses expériences comprenne un matériel biologique et sociologique, sachant que le raisonnement inductif est celui qui s'emploie le plus dans les sciences biologiques et sociales.*[45]

La question de savoir si la statistique est une science ou simplement une technique auxiliaire de toutes les sciences est un débat sur le «□ong terme□. Durant la seconde moitié du XIX[e] siècle, la statistique était synonyme de science sociale, mais quand elle devient discipline mathématique, indispensable pour la biologie et la physique, elle commence à revendiquer le statut de science indépendante. La polémique va continuer jusqu'à nos jours, bien que ce soit aujourd'hui un débat stérile. La conception anglo-saxonne est plus proche de considérer la statistique comme science – on peut le voir dans la lettre de Fisher, mais c'est aussi la conception de Bowley – au contraire la conception française[46] et celle de Corrado Gini penchent vers une considération de la statistique mathématique comme simple outil des autres sciences. La position de Fernandez Baños est plus proche de la conception de Fisher, à qui il dédiera son *Tratado de Estadistica.*

En Espagne, le premier manuel de statistique mathématique est écrit en 1924 par Antonio de Miguel. Licencié en sciences mathématiques de l'université de Madrid, et statisticien de l'État, il va jouer avec Fernandez Baños un rôle fondamental dans l'émergence de l'économétrie espagnole. Velarde Fuertes[47], professeur d'économie, lui adresse un grand éloge malgré les commentaires que Flores de Lemus avait faits à propos de son manuel. En plus de cela et de ses travaux sur le revenu de l'Espagne, il va s'occuper de faire le calcul de la conversion de la peseta républicaine en peseta du régime de Franco.

---

[45] *Ibidem.*

[46] Henri Berr, dans les Actes du 7[e] colloque international de synthèse publié en 1944 sous l'Occupation, commence par ces mots : *La statistique n'est pas une science : c'est une méthode, une méthode applicable aux objets les plus divers.*

[47] AHEPE, (2002) *Historia de la Probabilidad y de la Estadística*, Ias Jornadas de Historia de la Estadística organizadas por la Asociación de Historia de la Estadística y de la Probabilidad de España, Madrid, p. 300.

Avant le manuel d'Antonio de Miguel, il existait des traités de statistique administrative et des traités de calcul de probabilités. Merino Melchor, directeur de l'Observatoire de l'Académie de Madrid, et secrétaire de l'Académie des Sciences, avait fait en 1868 son discours de réception à l'Académie des Sciences *Sur l'application du calcul des probabilités aux événements humains*. Parmi les manuels les plus répandus, on trouvait le *Tratado de cálculo des probabilidades* d'Ollero, publié en 1879, et celui de Gabriel Galan Ruiz (1869-1938), *Cálculo de Probabilidades*, publié en 1923 mais qui avait obtenu le prix de l'Académie des Sciences dès 1909.

Ces manuels utilisaient des auteurs classiques, surtout de langue française et allemande. Par contre, le manuel d'Antonio de Miguel a toutes les caractéristiques des manuels de statistique mathématique de style anglo-saxon. Il utilise surtout les références de Bowley et Yule ; c'est probablement pour cela que Flores de Lemus, l'un des économistes les plus prestigieux de l'Espagne, ne l'a pas beaucoup apprécié. En introduction du livre, Flores utilise le terme «□brita□, terme péjoratif en espagnol (littéralement «□euvrette□, «□petit opuscule□), pour se référer au manuel – ce qui n'est pas très juste, parce qu'à mon avis c'est un bon manuel, et il ne faut pas oublier que c'est le premier traité espagnol de statistique mathématique. Toutefois, Flores fait des remarques pertinentes, par exemple□*la connexion entre les méthodes statistiques et les principes fondamentaux de la théorie de la probabilité n'est pas assez mise en lumière*[48]. Il ne faut pas oublier que la reconnaissance officielle de la méthode des échantillons n'aura lieu qu'en 1925 lors du colloque de l'I.I.S. Peut-être la controverse entre les économistes n'ayant pas de formation mathématique (Flores de Lemus était juriste d'origine), et les mathématiciens qui commencent à faire de l'économie, peut éclairer cette étrange présentation d'un bon manuel.

Flores conclut : *nous sommes très loin d'avoir une systématisation satisfaisante de cette magnifique création de l'esprit humain* qu'est la statistique. Il cite le manuel de Yule, mais critique le fait de rejeter à la fin les formules sur l'erreur. Il fait aussi la critique du travail de Davenport : *Statistical Methods* (1904) qu'il accuse d'être le premier responsable de l'utilisation de *l'artifice* qui consiste à *introduire des formules pour chaque série de problèmes en indiquant les conditions d'application,* formules avec lesquelles les manuels de statistiques deviennent une collection d'instruments pour des problèmes spécifiques de nature mystérieuse. Et même s'il fait des éloges de Pearson, Shuster et Fourier, la conclusion est claire :

> *la possibilité de montrer au public non-mathématique les conditions d'application des méthodes, et les instructions de son usage, sont très éloignées.*[49]

## LA CONFUSION DE PARADIGMES

En 1933, Fernandez Baños publie dans la revue *Economía Española*, un article qui illustre la portée des critiques faites par M. Halbwachs et Flores de Lemus. L'article s'intitulait : *Aplicación del análisis estadístico a un problema económico*[50] (Application de l'analyse statistique à un problème économique), et c'est dans celui-ci que Fernandez Baños tâchait de montrer la portée de la méthode statistique. Le motif apparent du travail était l'analyse du changement et de la parité économique de la peseta avec le

[48] A. De Miguel, 1924, p. V.
[49] *Ibidem*.
[50] O. Fernandez Baños, 1933.

franc français, mais, comme l'annonçait le titre, il s'agissait de montrer le fonctionnement des nouvelles techniques de la statistique mathématique[51].

Trois tableaux constituent la base empirique : un premier donnant les indices des prix de gros à Barcelone (qu'il prend comme représentatifs de ceux de l'Espagne), un deuxième, les indices des prix de gros en France, et un troisième, la valeur du franc en pesetas. Dans chaque tableau apparaissent les indices relatifs aux douze mois (il manque les dix derniers mois de 1933) de l'année, depuis 1913 jusqu'en 1933 et, à partir des ces trois tableaux, il en élabore un quatrième avec les «□différences entre le change et la parité exprimés en %□, un calcul qui, même si ce n'est pas explicite dans l'article, est obtenu en divisant l'indice espagnol par le français (les deux se référant à la même base), puis en multipliant le quotient par une constante égale au change moyen des deux monnaies en 1913 (l'année de référence).

Une fois présentée la base des données, Fernandez Baños calcule les statistiques fondamentales : moyenne, médiane, écart type, pour faire une première observation : il ne s'agit pas d'une distribution au hasard puisque la moyenne et la médiane ne coïncident pas. Les valeurs théoriques, comme les paramètres $K_1$, $K_2$ de Pearson qui indiquent le degré de déviation de la courbe empirique, montrent que la distribution empirique est proche de la courbe normale. De là découle toute une série d'observations qui aujourd'hui nous semblent un peu surprenantes, mais qui à l'époque constituent *l'habitus* de la statistique mathématique. Certainement, la période est cruciale□en effet, nous sommes à un moment de transition où l'on passe de la statistique mathématique basée sur les modèles linéaires et la corrélation à un autre modèle où le noyau de la discipline glisse vers la théorie des échantillons et vers la théorie de la décision.

Voyons comment se produit cette transition dans le discours de notre protagoniste. Déjà la première opération nous renvoie à la critique que Halbwachs faisait à Gibrat[52] : *seulement en faisant abstraction de cinq observations sporadiques, dont nous parlerons, et qui sont les cinq plus grands écarts positifs* (…) *il apparaît clairement que la courbe normale de Gauss est tout à fait indiquée*. Mais la courbe normale ne s'utilise pas comme le fait Bowley dans le texte de 1906 pour vérifier si le choix des échantillons donne des éléments fiables par rapport à l'ensemble dont ils ont été tirés. On est encore dans la recherche de lois qui rendent possible la prévision des tendances futures. Fernandez Baños l'annonce dans la note 3, et le souligne dans ses conclusions : *ce que l'on cherche est une loi empirique*.

Le procédé se trouve être très baroque : les paramètres de la courbe théorique servent à conclure *qu'il ne s'agit pas d'une loi pure de probabilité*, après quoi l'on cherche les causes d'un tel comportement dans des évènements de nature politique et sociale qui peuvent parfaitement s'entrevoir dans les représentations graphiques antérieures à l'utilisation de la courbe de Gauss. Les graphiques des pages 6 et 7 de l'article reflètent le changement de paradigme qui est en train de se produire : le passage du modèle linéaire à la théorie des échantillons. Le graphique de la page 6 est construit

[51] En 1936, Bowley publie un article : «□The application of sampling to economic and sociological problems□, *Journal of the American Statistical Associatio*n, September 1936, vol. 31, nº 195, où il raconte des expériences sur l'application de la méthode des échantillons. Voir *Empiria*, nº 5, 2002.

[52] *Notre loi est essentiellement statistique. Elle ramène simplement les courbes des répartitions économiques à une courbe célèbre, la courbe en cloche, dit aussi de Gauss, ou d'erreurs,* cf. note 24.

avec la logique (pourrait-on dire *le style de pensée* ?[53]) du modèle linéaire, alors que la page 7, annonce la théorie des échantillons, même si elle s'applique encore à un ensemble d'éléments dont nous connaissons déjà toutes les données.

Ainsi deux logiques se mêlent, donnant lieu à un discours relativement confus :

> *La raison fondamentale pour rejeter une loi de probabilité pure, – c'est-à-dire, une loi qui corresponde à une urne de composition telle qu'en faisant une extraction, il y ait une certaine probabilité d'obtenir un écart donné – c'est que les écarts que nous donne l'expérience viennent parfois en véritables rafales, c'est-à-dire, en blocs ou en grappes*(...). *Nous sommes devant un phénomène «*□'héritage□ *proprement dit, avec toutes les difficultés que cela implique.*

L'*héritage* ici n'est autre que les altérations de prix dues à des phénomènes comme la guerre, la stabilisation de la monnaie réalisée par un pays, ou un autre type de facteurs de nature sociopolitique.

Même Fernandez Baños a des doutes sur la signification réelle de tels ajustements :

> *Que signifient les deux constantes des termes de correction considérées? Il est évident qu'en augmentant suffisamment le nombre de tels paramètres (de la série de Charlier) on peut réussir n'importe quelle approximation désirée. Abstraction faite de l'erreur qui leur est inhérente, n'est-il pas évident que, tant que nous ignorons la signification des constantes, nous n'aurons rien obtenu de concret ou d'efficace, hormis un formalisme sans contenu ?*

Cela nous ramène de nouveau aux problèmes de connexion entre modèles mathématiques et réalité déjà posés par Halbwachs. Nous sommes cependant à une époque d'euphorie méthodologique qui fixe les propriétés de la méthode statistique au centre de laquelle se trouve la courbe de Gauss :

> *Tout cela indique, à notre avis, que nous avons trouvé une première approximation très acceptable quant au phénomène que nous étudions, une approximation qui trouve son expression dans une loi de fréquence normale.*

La nécessité de vérifier que ce sont bien les phénomènes sociopolitiques qui produisent les plus grandes altérations dans la courbe des prix pousse Fernandez Baños vers la corrélation, mais son objet principal est encore de découvrir des lois qui permettent des prévisions futures :

> *Cependant, il ne faut pas se faire d'illusion sur les valeurs données par le □2 parce que nous ne sommes pas face à un problème avec des probabilités a priori, ni devant un collectif où nous disposons de nombreuses expériences faites de manière à ce que nous ayons une probabilité empirique. L'expérience connue la plus importante est, peut être, celle que nous pressentons ici, et sa portée est réellement très grande, à notre avis, mais pas assez pour être sûr que nous soyons en possession d'une loi empirique.*

Pour l'explication des lois, on a de nouveau recours à la corrélation multiple : la façon de progresser dans l'explication de cette classe de phénomènes consiste à

[53] Ian Hacking propose parler de style de raisonnement scientifique. Cours donné au Collège de France, Paris, 2003.

introduire de nouvelles variables, et de trouver le pourcentage de l'explication qu'elles nous fournissent sur le phénomène. Ce sont des pratiques qui ont cours encore aujourd'hui, même si leur efficacité s'est montrée relativement douteuse dans certaines occasions.

Une autre pratique qui nous indique une compréhension encore insuffisante de la méthode des échantillons est le calcul du coefficient de corrélation linéaire sur la base de seulement vingt données annuelles. Si on utilise une parabole de troisième degré comme courbe théorique, le coefficient R2 s'élève de 0,95 à 0,96 (*sic*), même si pour une explication totale du phénomène, il nous manque l'introduction des nouvelles variables.

La conclusion finale ne peut être plus prétentieuse et abstraite :

> *Que pour l'Espagne, nous avons obtenu une magnifique explication des variations du change par le biais des variations de prix (et réciproquement), à condition de bien employer la technique statistique : la valeur de la corrélation par rapport à la courbe de fréquences trouvée, autorise à parler d'uniformité, ce qui constitue la première approximation d'une loi [trouvée] en Espagne.*

Halbwachs n'avait pas tort en dénonçant *l'excès de la rigueur*, un peu *artificielle et arbitraire* qui commençait à se répandre dans le milieu des sciences sociales ainsi que le goût des abstractions qui, souvent, reflètent très peu la réalité.

## 5. CONCLUSIONS

La statistique mathématique devient science statistique durant les années vingt. C'est au cours de cette décennie qu'apparaissent les manuels de statistique mathématique de style anglo-saxon. La statistique mathématique comme science, bien que marquée par les travaux de Pearson qui donne sa place aux généralisations du modèle linéaire, est liée à la théorie des échantillons et aux travaux d'Edgeworth et d'Arthur Bowley. Durant les années 1920 se produit un changement de paradigme : la statistique mathématique, qui avait commencé avec la recherche de lois empiriques et surtout avec l'analyse des problèmes de conjoncture économique, centre peu à peu son attention sur des phénomènes comme le chômage et l'opinion publique, ce qui rend possible l'émergence de la théorie des échantillons.

La statistique mathématique a été développée par des mathématiciens intéressés à l'économie et aux questions sociales. Pour comprendre son application à la sociologie, il faut considérer ce qui se passe alors aux États-Unis où des personnalités comme F. Ogburn et Samuel A. Stouffer vont développer les études d'opinion publique. En Grande-Bretagne, les études sur les conditions de vie de la classe ouvrière, sur le chômage, permettent à Arthur Bowley et à John Hilton de développer la méthode des échantillons.

En Espagne, la statistique mathématique se développe au milieu d'un vaste renouvellement des mathématiques mené par Rey Pastor. Il y a deux branches : la physique mathématique, développée par Esteban Terradas, et l'économétrie, dont la figure principale est Fernandez Baños. Le premier manuel espagnol de statistique mathématique paraît en 1924. Il est influencé par les manuels de A. Bowley et U. Yule et suppose une rupture avec les traités antérieurs consacrés au calcul des probabilités.

L'enseignement de la statistique mathématique commence à l'intérieur des écoles d'ingénieurs et à la Faculté des sciences de Madrid, où la première chaire de statistique mathématique est créée en 1934. L'application des méthodes mathématiques soulève quelques problèmes méthodologiques, qui sont mis en évidence par Bowley, Simiand et Halbwachs et, à sa façon, par les difficultés dans lesquelles se trouve Fernandez Baños. La généralisation du modèle linéaire et la recherche de lois empiriques ont retardé la reconnaissance de la méthode des échantillons ; le processus de construction de la statistique mathématique comme discipline scientifique avec ses rivalités professionnelles, a contribué à la généralisation de pratiques confuses et à une certaine difficulté de l'enseignement de la discipline.

## BIBLIOGRAPHIE


AHEPE, *Historia de la Probabilidad y de la Estadística*, Ias Jornadas de Historia de la Estadística organizadas por la Asociación de Historia de la Estadística y de la Probabilidad de España, Madrid, 2002.

ALMENAR S., «Olegario Fernandez Baños : de la geometría a la econometría», E. Fuentes Quintana (ed.), *Economía y economistas españoles,* vol. 6, *La modernización de los estudios de economía,* Madrid, Fundación de las Cajas de Ahorros Confederadas, Círculo de Lectores, 2002.

ARMATTE M., «Conjonctions, conjoncture et conjecture. Les barometres économiques (1885-1930)», *Histoire et Mesure*, VII, 1-2, 1992.

ARMATTE M., «El coeficiente de correlación y los economistas (1910-1940)», J.M. Arribas, M. Barbut (eds.), *Estadística y Sociedad,* UNED, 2002.

ARRIBAS J.M., «Antecedentes de la Sociedad de Consumo en España : de la Dictadura de Primo de Rivera a la II República», *Política y Sociedad*, 16, 1994.

ARRIBAS J.M., BARBUT M., *Estadística y Sociedad*, UNED, 2002.

BLUM A., MESPOULET M., *L'anarchie bureaucratique. Statistique et pouvoir sous Staline*, Paris, La Découverte, 2003.

BOWLEY A.L., «Presidential Adress to the Economic Section of the British Association for the Advanced Sciences», *Journal of the Royal Statistical Society*, 1906.

BOWLEY A.L., «The application of sampling to economic and sociological problems», *Journal of the American Statistical Association*, vol. 31, September 1936, nº 195. [Traduction en espagnol : E*mpiria*, nº 5, 2002].

ISI, *El congreso de Madrid en 1931*, INE, Madrid, 1983.

FERNANDEZ BAÑOS O., «Aplicación del análisis estadístico a un problema económico», *Economía Española*, Oct.-Nov. 1933, Dic., nº 10,11,12.

FERNANDEZ BAÑOS O., S*obre la correlación, medida de enlace directo o indirecto en los fenómenos economicos,* Comunicacion al IV Congreso de la Econometric Society, Stresa, Septiembre 1934, Banco de España.

FERNANDEZ BAÑOS O., *Programa, concepto, método y fuentes de Estadística Matemática*, Madrid, Talleres Gráficos Marsiega, 1941.

HALBWACHS M., *La Statistique. Ses applications. Les problèmes qu'elles soulèvent*, 1944, septième Semaine Internationale de Synthèse, 1935.

HORMIGON M., «Las matemáticas en España en el primer tercio del siglo XX», J.M. Sánchez Ron (ed.), *Ciencia y Sociedad en España*, Ediciones El arquero, CSIC, 1998.

MACKENZIE D.A., *Statistics in Britain 1865-1930*, Edinburgh University Press, 1981.

MARTINEZ LOPEZ V., *Olegario Fernandez Baños*, Graficas Ochoa S.A., 1995.

MIGUEL A. (de), «Metodología Estadística», *Fundamentos de Estadística matemática*, Madrid, 1924.

Revue d'Histoire des Sciences Humaines, *Maurice Halbwachs et les sciences humaines de son temps*, Septentrion, Presses Universitaires, nº 1, 1999.

REY PASTOR SELECTA, Fundación Banco Exterior, 1988.

SHAPIN S., *La Révolution scientifique*, Flammarion 1998.