Revue d'histoire des mathématiques,
7 (2001), p. 7–65.

MAURICE FRÉCHET STATISTICIEN, ENQUÊTEUR ET AGITATEUR PUBLIC

Michel ARMATTE (*)

RÉSUMÉ. — Le mathématicien Maurice Fréchet (1878–1973) est surtout connu pour sa contribution à l'analyse et à la topologie, notamment ses travaux sur les espaces métriques et les espaces abstraits. Son rôle dans la renaissance d'une école française de calcul des probabilités à la suite d'Émile Borel et aux côtés de Paul Lévy a été largement étudié. Ici c'est la facette Fréchet statisticien qui est explorée, en prenant pour principal fil conducteur une campagne menée entre 1934 et 1936 à l'Institut International de Statistique (IIS) contre les usages abusifs du coefficient de corrélation. Cette campagne prend la forme inhabituelle d'une enquête d'opinion auprès de nombreux collègues du monde entier, d'une série d'articles, de rapports de commission et de motions incendiaires dans les organes de l'IIS. Elle met en lumière les difficultés d'un fondement mathématique solide d'une des notions les plus élémentaires de la statistique mathématique, à un moment clé de son déploiement comme discipline scientifique. De plus la corrélation est une notion largement mobilisée comme instrument de preuve dans plusieurs domaines des sciences d'observation. Maurice Fréchet ne dédaignera pas d'apporter, après la guerre, sa pierre à cet édifice de refondation mathématique de la corrélation, à partir de ses propres travaux sur la notion de distance. Et, parce qu'il a une conception très respectueuse des mathématiques appliquées, il usera encore de la même technique d'enquête par correspondance et de la même pugnacité pour interroger la cohérence et la pertinence d'autres méthodes statistiques, comme l'estimation des paramètres d'une distribution théorique et l'application des mathématiques aux questions économiques et sociales.

ABSTRACT. — MAURICE FRÉCHET STATISTICIAN, INVESTIGATOR AND PUBLIC AGITATOR. — The mathematician Maurice Fréchet (1878–1973) is well known for his contributions to analysis and topology, and particularly his works on

---

(*) Texte reçu le 11 juillet 2001, révisé le 28 octobre 2001.

M. ARMATTE, UFR d'économie appliquée, Université Paris-Dauphine, 75775 Paris CEDEX 16 et Centre Alexandre Koyré, Pavillon Chevreul, 57 rue Cuvier, 75005 Paris.
Tél. : 01 44 05 40 16, fax : 01 44 05 47 33.
Courrier électronique : michel.armatte@dauphine.fr.

Mots clés : Fréchet, statistique, probabilité, corrélation, distance, histoire, enquête, correspondance.

Classification AMS : 62-03, 01A60, 01A90.

metric and abstract spaces. His role, following Émile Borel and together with Paul Lévy, in the renaissance of the French school of probability theory has been amply studied. In this paper, Fréchet the statistician will be explored, in particular the campaign sustained from 1934 to 1936 at the International Institute of Statistics (IIS) against improper uses of the correlation coefficient. This campaign took the unusual form of a survey sent to colleagues all over the world as well as a series of papers, committee reports, and censure motions within the IIS. It sheds light on the difficulties of giving a mathematically sound foundation to one of the most elementary notions of mathematical statistics, at the key moment when it was developing into an autonomous scientific discipline. Moreover, correlation is a notion largely used as an instrument of proof in several domains of observational sciences. Maurice Fréchet did not disdain contributing his own stone to the building of these mathematical foundations via his work on the notion of distance. And, because he respected applied mathematics, he used the same surveying technique – and with the same pugnacity – to examine the consistency and relevance of other statistical methods, like the estimation of the parameters of a theoretical distribution and the application of mathematics to economic and social questions.

INTRODUCTION : LA VÉRITÉ MATHÉMATIQUE SERAIT-ELLE UNE QUESTION D'OPINION ?

On dit souvent[1] que la science mathématique n'est plus, depuis le milieu du XIX[e] siècle, définie par ses objets — la quantité et l'espace — mais par sa méthode, déductive, organisée autour de la notion de preuve et de démonstration. Pour Charles S. Peirce, «*mathematics is the science of making necessary conclusions*». Et si la notion de preuve est elle-même ambiguë, sauf à la définir mathématiquement, comme l'ont tenté Russell, Peano ou Tarski, on ne s'attend guère à trouver dans la littérature mathématique autre chose que des énoncés établis par des preuves logiques et des démonstrations formelles, dans le style argumentatif couramment admis dans la communauté des mathématiciens.

Le dossier que nous présentons ici tranche curieusement avec cette vision standard de la discipline mathématique et de ses modes d'expression considérés comme légitimes. Certes, il s'agit d'une branche un peu particulière de la mathématique qu'on appelle la statistique mathématique, toute jeune au moment des faits qui seront évoqués, puisqu'elle ne prend ce nom qu'au début du XX[e] siècle, même si les premiers principes en ont déjà été établis fin XVIII[e]–début XIX[e] par des mathématiciens de renom

[1] Par exemple [Davis et Hersh 1981, p. 6].

comme Laplace et Gauss. Le cadre de la controverse que nous allons étudier n'est pas celui de la communauté des mathématiciens au sens le plus général, mais celui d'une respectable société savante internationale créée en 1885, l'Institut International de Statistique (IIS), qui tient son congrès tous les deux ans et qui édite un *Bulletin* (*BIIS*) ainsi qu'une *Revue* (*RIIS*). L'auteur des textes que l'on va lire n'est pas un quelconque statisticien amateur, mais un professeur titulaire des chaires de mathématiques générales, puis de calcul différentiel et intégral à l'université de Paris, auteur de recherches fondamentales sur les espaces abstraits, une des bases de la topologie.

Or cet auteur suscite en 1934 une commission de l'IIS (dont il sera le rapporteur) sur l'étude des propriétés d'une formule mathématique vieille de 40 ans, lance une vaste enquête par correspondance sur cette question auprès d'une vingtaine de statisticiens de tous pays, « particulièrement compétents en statistique mathématique », publie dans la *Revue* et le *Bulletin* de l'IIS plusieurs articles au vitriol, qui dénoncent les usages abusifs de cette formule, et suggère de mettre fin à ces mésusages redondants : « il y a des morts qu'il faut tuer plusieurs fois » nous dit-on [Fréchet 1935a, p. 4]. Pour finir, ce même auteur propose aux participants à la session le vote d'une motion de défiance, qui nie que quoi que ce soit ait été « démontré » quant à la capacité de ladite formule. Le débat ainsi lancé et relayé par les structures de l'Institut durera plus de deux ans. De plus, loin d'être un fait isolé, cette même technique d'enquête et d'agitation est réitérée par notre prestigieux mathématicien sur trois ou quatre questions centrales de la statistique mathématique dans les quinze ans qui suivent.

Alors quoi ? Une dénonciation, une commission, une enquête, une motion ... est-ce bien là le langage habituel de la science, et plus encore, celui des mathématiques ? Ne s'agit-il pas plutôt des ingrédients trop connus du débat idéologique, celui-là même que la communauté mathématicienne exècre, et qui ne saurait donc être érigé en méthode légitime ? La vérité mathématique serait-elle une question d'*opinion* ? Pourrait-elle ainsi résulter d'un sondage et d'un vote en assemblée, sans se déconsidérer, sans risquer de détruire la rationalité propre de sa démarche, que celle-ci soit tendue vers la pure vérité, ou qu'elle s'applique à des questions socio-économiques et serve ouvertement à éclairer la décision publique ?

Qu'est-ce qui fait courir ainsi un mathématicien chevronné ? Qu'a-t-il à faire d'une opinion publique dans une question scientifique ?

Nous pouvons faire l'hypothèse que ce comportement résulte d'une complexion caractérielle propre à ce mathématicien et qui le porterait à la critique et au doute systématique. Nous essaierons donc de comprendre quel type d'homme et quel type de savant était Maurice Fréchet, puisque c'est de lui qu'il s'agit. Il est possible aussi que ces faits relèvent d'une caractéristique spécifique du champ de la statistique mathématique, agité à cette période par une série de controverses fondamentales, dont celle-ci sur les usages du coefficient de corrélation. Et nous apprendrons beaucoup sur l'histoire de cette branche des mathématiques en dressant l'état des positions antagoniques sur le contenu même des différentes controverses. Il se peut également que ces enquêtes d'opinion nous informent davantage sur l'organisation sociale de la controverse, sur l'importance des écoles nationales, ou des différents paradigmes de la discipline, et sur le rôle des institutions internationales comme l'IIS. Il est enfin à peu près certain que la forme même de règlement de la controverse, et en particulier le recours à l'enquête, nous apprenne quelque chose sur le statut particulier de la connaissance en mathématique appliquée.

Il nous a semblé que les matériaux de ces enquêtes, parce qu'ils combinaient les approches biographiques, cognitives, sociales et épistémologiques, pouvaient être d'une grande richesse historiographique pour dresser un tableau de la statistique mathématique entre 1930 et 1950, c'est-à-dire au moment crucial de l'élaboration de ses fondements théoriques[2] et de

[2] La statistique comme discipline a longtemps évolué entre plusieurs statuts. Elle fut d'abord au XVII^e siècle une pratique administrative éclairée, adossée à une réflexion politique sur ce qui fait la force des États et les ressources du territoire, puis elle s'est un moment identifiée à une science des populations et de leur renouvellement, avant de se consacrer presque exclusivement aux activités économiques et sociales, en liaison avec la seconde industrialisation et la construction des identités nationales. Au début du XX^e siècle, son statut est devenu celui d'une discipline majeure de la logique inductive, permettant d'inférer d'observations et de mesures en grand nombre des régularités, voire des lois, dans n'importe quel domaine – social ou naturel – des sciences. Comme science de l'État, elle s'est institutionnalisée dans des bureaux de statistique qui ont mis en place une méthodologie de l'investigation, de l'enregistrement et de la publication, produisant ainsi au début du XIX^e siècle une avalanche de chiffres. Comme prototype de bon nombre de sciences sociales, elle s'est articulée aux questions du paupérisme, de la dénatalité, du travail ou de la crise économique, en fonction des agendas gouvernementaux. Comme méthode de l'inférence inductive, elle s'est appuyée sur les développements du calcul des probabilités dans les trois moments forts que celui-

sa constitution comme discipline académique[3].

FRÉCHET STATISTICIEN : ÉLÉMENTS BIOGRAPHIQUES

De la biographie de Maurice Fréchet, nous ne donnerons qu'un rapide survol, juste ce qui est nécessaire pour suivre l'intrigue, et avec un biais systématique qui consiste à privilégier le statisticien, parmi tous les rôles qu'a joués notre savant. C'est en effet dans cette discipline que va se nouer la problématique dont nous voulons rendre compte. De plus, c'est une facette du personnage et de sa carrière que les historiens des mathématiques n'ont guère étudiée[4].

(René) Maurice Fréchet est né le 2 septembre 1878 à Maligny (Yonne), le quatrième des six enfants d'une famille d'origine ardéchoise montée à Paris. Son père était directeur de l'école protestante de l'Oratoire et sa mère tenait une pension de famille, mais on la dit aussi couturière. Maurice a fait ses études au Lycée Buffon où il fut l'élève de Jacques Hadamard qui restera son père spirituel, puis à Saint-Louis. Comme ses

ci a connus à la fin du XVII[e] siècle (Pascal, Bernoulli), au début du XIX[e] siècle (Laplace, Gauss) et au début du XX[e] siècle (Pearson, Von Mises, Kolmogorov, Borel, Lévy et Fréchet). On considère en général que la statistique mathématique moderne est issue des travaux de Ronald Fisher sur la théorie de l'estimation et des tests statistiques entre 1920 et 1935, complétés par ceux de Neyman sur les intervalles de confiance et les sondages.

[3] Des chaires de statistique mathématique apparaissent à la fin des années 1920 aux États-Unis (rôle essentiel de Harold Hotelling), en Angleterre (première chaire de statistique mathématique à Cambridge en 1932) et en France (fondation de l'Institut de Statistique de l'Université de Paris, l'ISUP, en 1922). À la suite de l'*American Statistical Association* (ASA) fondée en 1839, les sociétés savantes et les revues consacrées à la statistique mathématique – comme l'*Econometric Society* (ES, 1930) et l'*Institute of Mathematical Statistics* (IMS, 1935) – se multiplient dans les années trente. Stephen Stigler [1999, p. 157] défend l'idée que la statistique mathématique commence en 1933, date du départ à la retraite de Karl Pearson, mais aussi de l'abandon par l'ASA du soutien aux *Annals of Mathematical Statistics* (fondées par Harry C. Carver), qui deviendront la publication officielle de l'IMS après 1935.

[4] Ni les notices de Angus E. Taylor du *Dictionary of Scientific Biography* [DSB] ou de [Charle et Telkès 1989], ni les études plus complètes de Taylor [1982–1987] ou de Luis C. Arboleda [1980] ne consacrent le moindre développement à l'œuvre statistique de Fréchet. Les très courtes notices de Daniel Dugué [1974] dans le *Journal de la Société Statistique de Paris* (la SSP, dont Fréchet devient membre en 1936 et qu'il préside en 1948) et dans la revue de l'IIS (que Fréchet avait intégré en 1931) n'y font que des références anecdotiques. La thèse de Sébastien Hertz [1997] sur Gumbel et celle de Bernard Locker, qui est en cours sous la direction de Bernard Bru, corrigent en partie ce manque.

trois frères, il choisit la carrière d'enseignant, mais il passe par la grande porte de l'École normale supérieure (1900–1903). Agrégé de mathématique en 1903, docteur-ès-sciences en 1906, il épouse en 1908 une béarnaise qui lui donnera quatre enfants. L'aînée, Hélène, épousera Edgar Lederer, professeur de biochimie à Paris, puis à Orsay.

La carrière de Maurice Fréchet commence en province dans les classes préparatoires de Besançon et de Nantes, puis dans les facultés des sciences de Rennes et de Poitiers. C'est là qu'il obtient sa première chaire de mécanique rationnelle en 1910. Mobilisé en août 1914, il passera une bonne partie de la guerre dans l'armée britannique, entre autres comme lieutenant interprète. À l'armistice, il est mis à la disposition du service d'Alsace-Lorraine à l'Université de Strasbourg qui ouvre ses portes (françaises) en 1919. Il y occupera la chaire d'analyse supérieure de 1919 à 1928. Son discours inaugural du 17 novembre 1919 a été publié dans la *Revue du mois* [Fréchet 1920]. Fréchet poursuit les recherches, engagées par sa thèse, sur ce qui deviendra le calcul fonctionnel, dont il est considéré comme un des principaux fondateurs avec Henri Lebesgue et David Hilbert. Il développe sous le nom d'espaces abstraits une généralisation de la topologie des ensembles de points [Fréchet 1921], inventant au passage la notion d'espace métrique. L'école russe de topologie des espaces abstraits (Aleksandrov et Urysohn) se réclame essentiellement de Fréchet et Hausdorff.

Cette première partie de sa vie scientifique, fortement associée à sa contribution à l'analyse et à la topologie, est bien connue depuis la thèse de Luis Carlos Arboleda [1980] et les publications de Angus E. Taylor [1982–1987].

Après 1924, s'ouvre une autre période. Fréchet, qui avait suivi le cours de statistique de la faculté de droit vers 1910, est chargé d'un enseignement de statistique et d'assurances à l'Institut commercial d'enseignement supérieur de Strasbourg (1924–1929). C'est dans ce cadre qu'il commence à lire et à écrire au sujet de la statistique, croisant le chemin de Maurice Halbwachs, professeur à la faculté des lettres et chargé d'un cours de statistique au même institut. Tous deux signent en 1924 un petit ouvrage sur « le calcul des probabilités à la portée de tous » [Fréchet et Halbwachs 1924], qui recevra le prix Montyon de statistique en 1925. Comme le fait remarquer Olivier Martin [2000], cette collaboration intéressante d'un mathématicien et d'un sociologue durkheimien fut sans lendemain. Et Fréchet, qui s'attribue le beau rôle dans la préface[5] et dans une lettre à Lazarsfeld, ne suivra pas les travaux d'Halbwachs, ne citant même plus la thèse complémentaire de celui-ci [Halbwachs 1912] quand il entreprendra sa « réhabilitation de la notion statistique d'homme moyen » [Fréchet 1950a]. Cet ouvrage a néanmoins marqué son temps comme une des premières tentatives d'introduction du calcul des probabilités auprès de populations intéressées par les phénomènes sociaux — la préface en appelle « aux médecins, démographes, économistes, actuaires et agents d'assurance [...] qui veulent posséder la théorie de leur art » — et « dont les connaissances mathématiques ne dépassent point le niveau de l'enseignement élémentaire » [Fréchet et Halbwachs 1924, p. vi]. Il devait également servir, selon la même préface, à la préparation des examens de l'Institut des actuaires, de la Statistique générale de la France et du Contrôle des assurances.

En 1928, Maurice Fréchet, appelé par Émile Borel pour enseigner le calcul des probabilités, revient à Paris, où il est successivement professeur sans chaire (de 1928 à 1933), puis professeur de mathématiques générales (de 1933 à 1949) et de calcul différentiel et intégral (de 1935 à 1941) à la faculté des sciences. À la suite de Borel, il occupe la chaire de « calcul des probabilités et physique mathématique » de 1940 à 1949, Georges Darmois lui succédant. Il sera parallèlement directeur d'études à la section de mathématiques de l'École pratique des hautes études, ainsi que professeur d'analyse et de mécanique à l'École normale de Saint-Cloud. Dans cette période parisienne, Fréchet s'intéresse surtout aux probabilités et se trouve très vite placé au centre de l'école française. Il publie deux ouvrages de recherches théoriques modernes sur le calcul des probabilités [Fréchet 1936b, 1938a] dans le *Traité de calcul des probabilités* de Borel et dirige la majorité des thèses dans ce domaine. Il prononce une conférence générale aux Congrès internationaux de mathématiques qui se sont tenus à Bologne (en 1928) et à Oslo (en 1936). Il préside et introduit le colloque

[5] Selon Fréchet, Halbwachs n'aurait été qu'un assistant à ses cours, auquel il aurait demandé de rédiger ses notes. La préface met en avant une collaboration entre le « mathématicien de profession » et le « sociologue et statisticien », qui « tire son origine d'une série de leçons sur les probabilités faites à l'université de Strasbourg en 1921 par l'un des auteurs, rédigées, complétées, remaniées et mises au point en collaboration avec l'autre » [Fréchet et Halbwachs 1924, p. vii].

international de Genève (1937) sur la théorie des probabilités [Fréchet 1938]. Nous renvoyons à la thèse de Bernard Locker quant au travail de fondement du calcul des probabilités, en particulier dans leur liaison avec les travaux de Paul Lévy sur les sommes de variables aléatoires. Une application importante de ces recherches concernera la théorie des erreurs, un domaine dans lequel ses publications sont nombreuses [Fréchet 1922, 1924c, 1925a, 1925b, 1928a, 1940b, 1943c, 1947e, 1947f, 1948c, 1949e, 1950a]. Il faut retenir en particulier son combat contre les dogmes de la moyenne et de la loi de Laplace-Gauss, ainsi que sa tentative de réhabilitation de la médiane et de la première loi de Laplace. Sa participation à la statistique des extrêmes [Fréchet 1927b] a donné lieu à de vifs échanges de correspondance avec Paul Lévy et d'autres plus sereins avec son protégé Emil Julius Gumbel, qui se trouvent reproduits et analysés dans la thèse de Sébastien Hertz [1997][6]. On notera en particulier la polémique lyonnaise qui oppose Fréchet et Gumbel à Galbrun. Celle-ci caractérise assez bien l'opposition de certains milieux (ici les actuaires) au travail mathématique sur le calcul des probabilités.

La polémique qui va nous intéresser porte sur la signification et les usages du coefficient de corrélation. Elle se déroule dans plusieurs cénacles — principalement l'Institut international de statistique et la Société statistique de Paris — et révèle l'intérêt de Fréchet pour la statistique mathématique. Elle va lui conférer une certaine notoriété qui trouvera son point d'orgue dans sa présidence de la SSP en 1948. Depuis, cet intérêt ne s'est pas démenti, comme le montre sa collaboration régulière aux sessions de l'IIS, qui commence dès 1925 (Rome) et se poursuit aux sessions de Londres (1934), Mexico (1935), Athènes (1936) et Rio (1955).

Mais auparavant, la seconde guerre mondiale, qui donne l'occasion à son maître Émile Borel d'exercer d'autres fonctions, libère une chaire à la faculté des sciences de Paris, que Fréchet occupera[7]. De plus, en 1942, il prend la direction du laboratoire de calcul et de statistique, que Borel et Valiron avaient fondé deux ans plus tôt à l'Institut Henri Poincaré,

---

[6] Voir en particulier les sections 3.3 et 3.5 ainsi que les correspondances Fréchet-Gumbel en annexe, p. 7–60.

[7] « Certaines circonstances nées de la nouvelle guerre me contraignent à consacrer maintenant la plus large part de mon temps à la Statistique Mathématique » écrit-il en 1940 dans le JSSP [Fréchet 1940b, p. 68].

puis cherche à le fusionner avec le laboratoire de calcul expérimental et graphique de Joseph Pérès, ainsi que celui de calcul mécanique de Louis Couffignal[8]. Le centre de gravité de ses activités se déplace alors vers le calcul numérique, qui se trouve placé au carrefour d'intérêts nouveaux pour la statistique, pour les technologies des machines à calculer et pour quelques champs d'application privilégiés comme l'économie. L'essentiel de son activité reste, pour sa fin de carrière, consacré au calcul des probabilités[9]. Correspondant de l'Académie des sciences depuis 1952, il est élu le 14 mai 1956, à 77 ans et après quatre essais infructueux, à la section de géométrie de l'Académie des sciences, au fauteuil libéré par le décès d'Émile Borel. Maurice Fréchet est décédé à Paris le 4 juin 1973. Il avait 95 ans.

UNE BRÈVE HISTOIRE DU COEFFICIENT DE CORRÉLATION

Le coefficient de corrélation, dit de Bravais-Pearson, mais appelé de Bravais-Galton par Fréchet, est une invention de l'école biométrique anglaise, [Stigler 1986] et [Armatte 1995, 2001], qu'il faudrait plutôt attribuer à Galton et Pearson, si l'on donne au mot invention tout son sens moderne et constructiviste d'innovation socio-technique. Les mathématiciens et astronomes du début du XIX[e] siècle, en tout premier lieu Laplace dans son mémoire de 1810 et dans le paragraphe 21 de la *Théorie analytique* [Laplace 1820][10], mais aussi Gauss [1826], avaient déjà mis en évidence un terme qui apparaît dans la forme quadratique exprimant la distribution de Laplace-Gauss à deux dimensions. Mais aucun d'eux n'avait identifié le coefficient de ce terme à une quelconque

---

8 Archives de l'Académie des sciences, carton F4.2.

9 Il entre à l'Institut des actuaires français en 1945, devient président de la Société de biométrie et *fellow* en 1951 de la Société internationale d'économétrie. Il préside le colloque international de Lyon (28 juin-3 juillet 1949) du Centre National de la Recherche Scientifique (CNRS) sur le calcul des probabilités et il est le président d'honneur du Colloque international des probabilités organisé par le Congrès international de mathématiques d'Amsterdam en 1954.

10 Laplace [1820] est amené à écrire la densité de probabilité du couple de variables u et u' sous la forme

$$P(u,u') = \frac{kI}{4k''a^2\pi\sqrt{E}}\,\mathrm{e}^{k(Fu^2+2Guu'+Hu'^2)/4k''a^2E}\,\mathrm{d}u\,\mathrm{d}u'.$$

mesure de corrélation, ni distingué ce paramètre d'une loi de probabilité de son estimation empirique à partir des données observées et supposées produites par ce modèle. Une telle distinction ne sera pas explicite avant les travaux de Ronald Fisher vers 1920.

Auguste Bravais [1846, p. 263], étudiant l'erreur sur un point du plan dont les coordonnées x, y sont fonctions linéaires de n grandeurs indépendantes entachées d'erreurs gaussiennes, écrit : « La coexistence des mêmes variables $m, n, p, \ldots$ dans les équations simultanées en x et y amène une corrélation telle que les modules hx, hy cessent de représenter la possibilité des valeurs simultanées de (x, y) sous le vrai point de vue de la question ». Il reconnaît que la surface de probabilité du couple (x, y) a pour courbes de niveau des ellipses d'équation $ax^2 + 2cxy + by^2 = \textit{constante}$, mais le terme c n'est pas identifié comme mesure de cette « corrélation » bien que le terme soit prononcé. Jules Bienaymé à Paris, William Donkin à Oxford, Friedrich-Robert Helmert en Allemagne et P. Pizzetti à Gênes ont largement anticipé les travaux de Karl Pearson sur la loi d'un vecteur aléatoire gaussien. Les artilleurs, à la suite des travaux de Poisson publiés dans le *Mémorial d'artillerie* en 1830 et 1837, ont fait un large usage de cette loi de distribution dans le plan pour rendre compte des phénomènes de balistique extérieure, comme par exemple à l'Ecole de Metz et chez Isidore Didion pour les impacts d'obus[11].

Alors, comment expliquer que Karl Pearson [1920] revendique pour lui et Galton l'invention de la corrélation dans les années 1880 et 1890 ? C'est précisément, parce que les variables corrélées x et y ne sont pas, ni chez Gauss ni chez Bravais, des quantités directement et organiquement observées et que tout au contraire celles-ci sont chez eux indépendantes. Ce qui renvoie chez Pearson à une sorte de coupure sémantique entre l'approche classique de la théorie des erreurs et celle de la biométrie qu'il inaugure. S'y ajoute une coupure épistémologique, due au rôle de la contingence qu'il a lui même placée, dans la *Grammaire de la science*, au centre d'une philosophie des sciences radicalement phénoménaliste, caractérisée par une notion de loi qui n'exprime plus chez lui une idée de nécessité, mais celle d'un simple « résumé sténographique » de notre perception du phénomène[12].

[11] Voir à ce sujet les travaux de Bernard Bru [1995] et Pierre Crépel [1994].

[12] « Pas à pas les hommes de science ont reconnu que les phénomènes n'ont pas

Dans ses études sur l'hérédité, et en liaison étroite avec son programme eugéniste, Francis Galton, cousin de Darwin, avait associé la mesure de corrélation entre deux variables statistiques à l'homoscédasticité[13] des sous-populations de parents de même taille, puis à la pente de la droite des médianes conditionnelles des tailles des enfants (ce qui formalise la « *regression towards mediocrity* » observée dans le phénomène héréditaire) et enfin à l'équation des ellipses d'équiprobabilité repérées sur le tableau croisé des tailles enfants-parents [Galton 1889]. Karl Pearson, son successeur au London College, donne à cette mesure de corrélation un statut mathématique et philosophique un peu différent et beaucoup plus précis : il établit en 1896 une formule du coefficient de corrélation empirique r dite du « moment produit » directement calculable sur la base des n observations de x et y (voir notre encadré), dont il montre qu'elle est l'estimation « naturelle », par la méthode des moments, du coefficient de corrélation théorique ρ qui apparaît dans la densité d'une loi normale à deux dimensions. Ce coefficient r devient alors le meilleur résumé possible de la contingence, cette notion intermédiaire entre indépendance et liaison fonctionnelle que Pearson considère comme première et plus générale que la notion de loi. Ce coefficient résume « une routine de perception » et peut devenir ainsi le moyen rhétorique idéal de la preuve, dans une approche de la science qui rejette toute raison *a priori*, comme la physique de Mach dont Pearson s'inspire, ou comme la théorie eugénique des populations qu'il cherche à développer.

pour origine un mécanisme, mais l'interprétation mécanique constitue seulement la sténographie conceptuelle à l'aide de laquelle ils peuvent décrire brièvement et résumer les phénomènes » [Pearson 1912, p. xiii].

[13] Cette propriété signifie que les sous-populations ont la même variance.

Propriétés syntaxiques et sémantiques attribuées à la mesure de corrélation

Chez Galton (1889). — C'est en étudiant la façon dont la taille (y) des enfants varie en fonction de la taille moyenne (x) de leurs parents sur les pois de senteur (1877) puis sur les humains (1885), que Galton découvre ce qu'il appelle d'abord réversion puis régression : il constate que la médiane des tailles y des enfants de parents de même taille x est une fonction linéaire de cette taille x.

Le coefficient r apparaît comme la pente de cette droite si on centre et réduit les données. Elle est inférieure à l'unité, ce qui lui fait dire que l'hérédité n'est pas totale : les plus grands font des enfants grands mais pas aussi grands qu'eux. Il y a « régression vers la médiocrité ».

Il découvre la décomposition de la variabilité (erreur probable) de y en deux parties : variabilité des médianes de groupe « expliquée » par x et variabilité intra-groupe. r^2 apparaît alors comme le rapport de la variabilité « expliquée » à la variabilité totale de y.

Son collègue mathématicien lui montre que les courbes qu'il découvre sur son tableau en joignant les cases de même fréquence ne sont rien d'autres que les ellipses d'égale densité d'une loi de Laplace-Gauss à deux dimensions.

En 1888, il découvre que le même r peut servir de mesure symétrique du degré de corrélation de deux caractères x et y mesurés sur n individus.

Chez Pearson (1896). — La notion de contingence est première par rapport à celle de nécessité. La loi normale multidimensionnelle qui exprime cette distribution s'écrit

$$z = \frac{n}{2\pi\sigma_x\sigma_y\sqrt{1-\rho^2}}\,\mathrm{e}^{-\frac{1}{2}\left\{\frac{x^2}{\sigma_x^2(1-\rho^2)} - \frac{2\rho xy}{\sigma_x\sigma_y(1-\rho^2)} + \frac{y^2}{\sigma_y^2(1-\rho^2)}\right\}}.$$

Le terme qui apparaît comme paramètre de la forme quadratique est le coefficient de corrélation théorique entre les deux variables.

Le terme ρ apparaît comme paramètre de la forme quadratique qui y figure.

La linéarité de la régression s'en déduit par l'espérance conditionnelle de y (en données centrées) :

$$E(y/x) = \rho\frac{\sigma_y}{\sigma_x}x$$

En cherchant une estimation de ρ qui maximise la vraisemblance de l'échantillon observé, Pearson obtient la formule de r (dite de Bravais-Pearson) qui ne dépend que de la moyenne et de l'écart-type des observations

$$r = \frac{\operatorname{cov}(x,y)}{\sigma_x\sigma_y} = \frac{\sum(x_i-\bar{x})(y_i-\bar{y})}{\sqrt{\sum(x_i-\bar{x})^2}\sqrt{\sum(y_i-\bar{y})^2}}.$$

Il établit que cette statistique a un écart-type

$$\sigma_r = \frac{1 - r^2}{\sqrt{n}}\cdot$$

Chez Yule (1897, 1911). — On ne fait plus l'hypothèse d'une distribution normale des observations.

Si la régression vue comme lieu des moyennes de y pour une valeur x_i donnée est linéaire, son équation $\bar{y}_i = ax_i + b$ coïncide avec celle de la droite des moindres carrés de Legendre et Gauss. Sinon, celle-ci en donne le meilleur ajustement.

r^2 mesure la part de variance totale de y « expliquée » par la relation précédente :

$$V(\bar{y}_i) = a^2 V(x_i) = r^2 V(y).$$

r est donné par la formule de Pearson et s'interprète comme la pente de la droite de régression en données centrées et réduites :

$$\frac{y_i - \bar{y}}{\sigma_y} = r\,\frac{x_i - \bar{x}}{\sigma_x}\cdot$$

Des formules similaires (en échangeant x et y) fournissent les équations de la droite de régression de x en y.

Chez Fisher (1915). — Un échantillon x est un point ou vecteur de l'espace $\mathbb{R}^n$; sa moyenne $\bar{x}$ est la projection de ce vecteur sur la droite des constantes, et son écart-type σ_x est sa longueur euclidienne. Si y est un autre échantillon-vecteur de $\mathbb{R}^n$ le cosinus de l'angle $(x - \bar{x}), (y - \bar{y})$ n'est autre que le coefficient de corrélation r.

La distribution exacte de la statistique d'échantillonnage r n'est pas normale. Il faut lui substituer une transformation dite de Fisher-Snedecor dont la loi a été tabulée.

Ces deux considérations permettent de tester une hypothèse de nullité de la valeur théorique de ρ, c'est-à-dire de tester la significativité d'une valeur empirique r.

Nous avons étudié ailleurs et en détail [Armatte 1995, 2001] dans quelles conditions la notion de corrélation et sa mesure avaient été adoptées et transformées par les économistes au début du XX$^{\text{e}}$ siècle. Nous y avons beaucoup insisté sur le fait que ce transfert à d'autres champs et d'autres problématiques avait nécessité une succession de redéfinitions des propriétés mathématiques du coefficient r, ou encore une sélection parmi ces propriétés de celles qui doivent être mises en avant. Ce travail sur les propriétés syntaxiques du coefficient s'accompagne à chaque fois d'une forte inflexion de ses propriétés sémantiques vers des significations pertinentes pour le domaine et l'épistémologie nouvelle considérée, ainsi que d'une série de recommandations sur les usages à privilégier. Le transfert du coefficient de corrélation de la biométrie à l'économie est un bon exemple de ces transformations.

George Udny Yule est, avec R.H. Hooker et Arthur L. Bowley, le principal introducteur de la méthodologie statistique de la régression et de la corrélation en économie, mais il choisit une approche de la corrélation qui est aux antipodes de celle de Pearson. À la contingence Yule [1897, 1909] substitue une notion de causalité statistique qui renoue avec la philosophie et le vocabulaire des astronomes : 1°) il n'y a plus alors de différence entre la théorie de la corrélation de Galton et Pearson d'une part et la théorie des erreurs plus ancienne de Gauss et Laplace de l'autre. La droite de régression se confond avec la droite des moindres carrés ; 2°) « il est important d'obtenir la formule de corrélation et ses propriétés sans avoir recours à la distribution de fréquence » [Yule 1909, p. 268], autrement dit sans l'hypothèse de normalité qui fonde la théorie de la corrélation chez Pearson.

La nature même des données économiques et du programme de recherche, qui tourne principalement autour de la question des cycles économiques, impose également une réinterprétation complète de ce que signifie la corrélation de deux séries de nombres quand celles-ci prennent une dimension chronologique. Comme on le voit clairement dans la controverse qui s'empare des études sur la covariation des taux de mariage et des prix du blé, controverse qui se généralise ensuite aux différents types de baromètres économiques ([Armatte 1992], [Morgan 1995]), la mesure de corrélation devient à proprement parler insensée, ininterprétable, si l'on s'en tient aux séries brutes. Il faut, comme le propose Hooker en 1901, décomposer chaque série en une tendance et des cycles par des filtres appropriés (moyennes mobiles, différences, analyse harmonique) avant que de pouvoir calculer des corrélations qui soient interprétables, puis il faut rechercher le décalage (*lag*) temporel entre les séries (par exemple de pluviométrie et de rendement agricole) qui maximise la mesure de leur corrélation. Mais faut-il encore que ces techniques de filtrage n'introduisent pas des liaisons artefactuelles ! Or c'est précisément ce que George U. Yule et Eugen Slutzky vont montrer dans les années vingt, ruinant l'espoir que ces techniques de filtrage permettent enfin d'interpréter correctement la corrélation entre deux séries chronologiques. Celle-ci ne serait-elle pour finir qu'un pur artefact, conséquence du type de leur autocorrélation propre ? C'est ce que suggère Yule en 1926. Encore faudrait-il aussi que l'optimisation du *lag* par la maximisation du coefficient de corrélation soit licite, or c'est justement ce que va contester Fréchet, bien que cette méthode soit très répandue, y compris en France, et serve de fondement aux propriétés prévisionnelles des baromètres économiques les plus célèbres. Celui par exemple, qui a été élaboré à Harvard par Warren Persons et qui sert de modèle aux différents baromètres européens, repose sur une avance du cycle du marché boursier par rapport aux cycles des marchés des biens industriels et de la monnaie. Mais c'est très exactement la permanence de ce mécanisme qui est questionnée après 1925 (les courbes divergent) et qui est totalement remise en cause par la crise de 1929.

PREMIÈRE ACCROCHE AVEC LA CORRÉLATION

C'est dans cette situation de crise de la corrélation que s'inscrit l'intervention du mathématicien Fréchet. Il est temps, maintenant, de la décrire plus finement. Fréchet entre dans le débat en rédigeant le compte rendu [Fréchet 1927a] d'un ouvrage de statistique mathématique de Henry Lewis Rietz publié la même année par la *Mathematical Association of America*[14] et qui connaîtra cinq éditions. Ce professeur[15] de l'Université d'Iowa avait, selon D.B. Owen [1974], une connaissance importante des auteurs européens (58 références contre 10 pour les Américains dans ce petit ouvrage) et affirmait sa volonté de développer la statistique mathématique sur les bases rigoureuses du calcul des probabilités et, plus précisément encore, selon le principe attribué à Borel et exigeant que toute table de données statistiques soit modélisée par un schéma d'urne produisant quasiment la même série. Cette conception était loin

[14] « Elle se distingue de l'*American Mathematical Society* par un caractère pédagogique plus marqué et par une collaboration plus importante du personnel de l'enseignement secondaire » dit Fréchet [1927a].

[15] H.L. Rietz sera le premier président de l'*Institute of Mathematical Statistics* fondé en septembre 1935. Walter Shewhart, inventeur avec Egon Pearson des méthodes statistiques du contrôle de qualité, en fut le vice-président et Harold Hotelling un membre influent du comité de direction.

d'être partagée par les statisticiens économistes américains ou européens, à commencer par les producteurs de baromètres comme Warren Persons à Harvard ou Lucien March[16] à Paris, qui souhaitaient construire une statistique mathématique non probabiliste. Comme le montrent bien les articles sur la statistique de l'*Encyclopédie des sciences mathématiques* dans ses deux versions allemande et française [Molk 1908], le début du XX^{e} siècle est marqué par une renaissance du probabilisme qui prend ses distances avec cette tendance, mais aussi avec le modèle bernoullien de l'urne constante, les dogmes de la loi normale et de l'homme moyen, pour explorer, avec Wilhelm Lexis, Emile Borel et Ludwig von Bortkiewicz, des modèles de la variabilité et de l'hétérogénéité. Rietz, comme Fréchet, s'engagent dans cette direction. Le compte-rendu que le second consacre à l'ouvrage du premier est donc très favorable.

Fréchet [1927a] souligne « que cet ouvrage se distingue des ouvrages de statistique mathématique pullulant dans les pays de langue anglaise, par le souci d'expliquer et de justifier les règles que ces ouvrages se contentent généralement de poser [...] montrant que ces règles empiriques n'ont rien d'absolu (et) que l'appareil mathématique les accompagnant ne doit pas faire illusion sur la portée des définitions conventionnelles ».

Et il poursuit : « Je pense qu'on ne saurait trop s'élever contre l'usage qui est fait du coefficient de corrélation comme mesurant l'étroitesse de la relation. On éviterait probablement mieux de fausses interprétations en l'appelant *coefficient de linéarité* puisqu'il atteint sa valeur maximum $+1$, non quand x et y sont simplement fonction l'un de l'autre, mais seulement quand x et y sont liés par une relation linéaire. »

Georges Darmois, qui vient de publier un important traité de statistique mathématique [Darmois 1928], s'est associé dans un article de *Metron* à cette critique et à cette tentative de rebaptiser le coefficient de corrélation : il n'est pas absurde, dit-il, d'utiliser le coefficient de Pearson pour résumer une relation plus ou moins linéaire entre deux séries chronologiques, mais « ce qui serait tout à fait absurde, c'est de penser qu'une telle relation linéaire, assez grossièrement vérifiée, valable pour un petit intervalle de temps soit une base suffisante pour l'induction d'une loi reliant objectivement les deux grandeurs [...] C'est pourquoi j'estime, étant donné la

[16] Directeur de la Statistique Générale de la France (SGF) et fondateur de l'ISUP.

force de suggestion du seul mot corrélation, qu'il vaudrait mieux renoncer dans ces questions à l'emploi du qualificatif coefficient de corrélation. En l'appelant indice de linéarité, ou quelque chose d'analogue, nous le ramenons à un rôle plus modeste et formel » [Darmois 1929, p. 226]. Darmois, bien au courant de la littérature anglaise qu'il cite précisément, s'appuie en particulier sur les travaux de Yule, Slutzky et Anderson cités plus haut, pour émettre de telles réserves.

Mais la fin des années vingt et le début des années trente restent marqués par de nombreuses études, en particulier dans le domaine économique, où le coefficient de corrélation est utilisé de façon intensive pour faire la preuve d'une relation. On ne sait pas exactement ce que Fréchet a eu entre les mains à cette période, mais toute la littérature sur les baromètres économiques est marquée par cette approche[17]. L'article qu'il publie dans la *RIIS* « sur le coefficient *dit* de corrélation » est motivé, dit-il, par « la lecture de mémoires où, malgré les conseils de prudence qui leur ont été donnés de divers côtés, de nombreux statisticiens continuent à employer sans aucune précaution le coefficient dit de corrélation pour mesurer l'étroitesse de la dépendance fonctionnelle entre deux variables statistiques » [Fréchet 1934a, p. 16]. L'usage dénoncé est principalement celui qui consiste à induire d'une faible valeur du coefficient de corrélation r l'absence de relation fonctionnelle ou encore l'indépendance des variables en question. Fréchet malheureusement se trompe lorsqu'il fournit un contre-exemple[18] d'une relation fonctionnelle (non linéaire) pour laquelle $r = 0$, mais il est facile d'en inventer d'autres. S'il considère comme préférable d'utiliser le rapport de corrélation de Pearson pour mesurer l'intensité d'une relation non linéaire, il rappelle que son calcul nécessite l'observation de plusieurs valeurs de x pour une même valeur de y, ce qui est rarement le cas dans les usages statistiques.

Enfin, il s'appuie sur une décomposition du coefficient de corrélation qu'il a publiée dans les *Comptes rendus … de l'Académie des sciences*

[17] On pense en particulier aux travaux de Warren Persons sur le baromètre de Harvard, mais aussi aux premiers travaux de Tinbergen, ou encore, dans un environnement plus proche de Fréchet, aux recherches faites à l'ISUP depuis sa création par ses professeurs, [March 1905, 1910, 1928], [Bunle 1911], ou les premiers thésards [Machali 1931].

[18] Le contre-exemple proposé par Fréchet est $y = 0, 1, 2, 3, 4, 5, 6, 7, 8$; $x = 0, 1, 2, 3, 8, 7, 6, 5, 4$. Il ne conduit pas à $r = 0$. Il faudrait remplacer par exemple la série x par $x = 0, 1, 2, 3, 4, 3, 2, 1, 0$ pour obtenir $r = 0$.

[Fréchet 1933a] pour montrer que l'on confond trop souvent derrière une mesure de corrélation les deux propriétés de linéarité et d'intensité de la dépendance : le coefficient de corrélation s'exprime en effet comme le produit du coefficient de corrélation ρ des couples $(\bar{y}_i, x_i)$ exprimant la linéarité de la régression et du coefficient η de Pearson mesurant la concentration autour de la ligne de régression (ou plus exactement le rapport de la variance expliquée par les moyennes liées à la variance totale de y), soit

$$r = \rho\eta \quad \text{avec} \quad \eta = \sqrt{\frac{\sum n_i(\bar{y}_i - \bar{y})^2}{\sum n_j(y_j - \bar{y})^2}} = \sqrt{\frac{\operatorname{var}\bar{y}_i}{\operatorname{var}\bar{y}_i + \operatorname{var}\lambda_{ij}}}$$

où les λ_{ij} sont les fluctuations de y autour de $\bar{y}_i$ telles que $y_j = \bar{y}_i + \lambda_{ij}$ et

$$\rho = \frac{\sum n_i(\bar{x}_i - \bar{x})(\bar{y}_i - \bar{y})}{\sqrt{\sum n_i(x_i - \bar{x})^2}\sqrt{\sum n_i(y_i - \bar{y})^2}}\cdot$$

Or, dans ce produit, figure le facteur ρ, qui est étranger à la mesure de la dépendance entre x et y, et qui faussera une induction directe basée sur r. On sait que, dans le cas d'une distribution normale, la régression est linéaire et $|r| = \eta$, donc la confusion disparaît, ce qui fait dire à certains statisticiens plus exigeants que la mesure de la dépendance par r n'est valide que sous cette hypothèse de normalité. Et pour finir, même la substitution de η à r pour mesurer la dépendance entre x et y peut se révéler tout à fait illégitime, si le nombre d'observations de y pour une même valeur de x est trop faible pour donner lieu à une dispersion autour de la moyenne liée (voire égale à 1 comme très souvent dans des données chronologiques). En tous cas, conclut-il, « un grand nombre des applications courantes du coefficient de corrélation [qui ne tiennent pas compte de ce qui précède] sont très hasardées » [Fréchet 1934, p. 21] et en particulier celles qui déterminent le retard d'une série sur une autre qui lui serait liée par une maximisation du coefficient de corrélation pour différents retards. Or c'est justement la méthode standard des baromètres, utilisée entre autres par Hooker, Persons, Bunle et Machali.

MAURICE FRÉCHET ENGAGE UNE CAMPAGNE[19] CONTRE LE COEFFICIENT DE CORRÉLATION

Si Fréchet en était resté à cette publication, nous aurions admis qu'il était dans son rôle de mathématicien rappelant à l'ordre les usagers de la statistique. Mais il va s'engager dans une vaste campagne qu'il s'agit maintenant de relater. Membre de l'IIS depuis 1931, Fréchet va mobiliser cette structure à l'occasion de sa 22e session qui se tient à Londres en avril 1934.

L'Institut international de statistique a été créé en 1885 lors du cinquantième anniversaire de la *Royal Statistical Society* pour prendre la suite du Congrès international fondé par Quetelet en 1851 et dont la dernière réunion à Saint-Pétersbourg en 1872 avait montré les limites, en particulier l'hostilité des gouvernements à assumer des décisions prises en son sein. Ramené à la seule fonction académique de confrontation scientifique internationale des méthodes et des résultats de la statistique, l'IIS n'a plus, en principe, de lien politique direct avec les gouvernements et peut tout au plus émettre des vœux sur la politique des Etats. Les sessions ont lieu tous les deux ans[20]. Les statuts, refondus à Londres en 1934, définissent l'IIS comme « une association internationale autonome qui a pour but de favoriser le progrès de la science administrative et scientifique ». Le financement de l'institut provient des cotisations des membres titulaires (limités en 1934 à 225 et 25 par nation), du produit des abonnements, des dons et des fondations. Un office permanent, créé en 1913, gère la documentation et les publications dans les inter-sessions. Chaque session se clôt par une assemblée générale qui procède à l'élection du bureau (président, vice-présidents, secrétaire général et trésorier). Le mémoire de Zahn [1935] permet de se faire une idée de l'organisation du travail de préparation des sessions de l'IIS, jusqu'aux années 1930. Lors de ces sessions, diverses sections fonctionnent en parallèle. En général au nombre de

[19] Selon ses propres termes dans le supplément à la notice sur ses travaux [Fréchet 1951, p. 9].

[20] Londres 1885, Rome 1887, Paris 1889, Vienne 1891, Chicago 1893, Saint-Pétersbourg 1897, Christiania 1899, Budapest 1901, Berlin 1903, Londres 1905, Copenhague 1907, Paris 1909, La Haye 1911, Vienne 1913. Elles ont repris après la guerre : Bruxelles 1923, Rome 1925, Le Caire 1927, Varsovie 1929, Tokyo 1930, Madrid 1931, Mexico 1933, Londres 1934.

trois, elles sont consacrées aux statistiques démographiques, économiques et sociales (ou encore « administratives et judiciaires »), mais on leur a adjoint une section de statistique mathématique en 1935. Dans chaque section, des comités s'emparent des questions mises à l'ordre du jour par le bureau, tiennent lors des sessions des réunions séparées qui se bornent à un examen préliminaire de chaque question, préparent une discussion en séance plénière et mettent en place des commissions chargées entre les sessions d'approfondir l'étude ou l'enquête sur la question soulevée, d'appliquer les décisions et motions de l'assemblée générale, ainsi que de préparer un rapport pour la prochaine session. Les séances plénières combinent les communications originales et les discussions des rapports des commissions.

En 1934, Maurice Fréchet obtient du bureau la mise à l'ordre du jour du débat sur le coefficient de corrélation et la constitution d'un comité qui préparera la séance, puis se transformera en une commission. Celle-ci est composée des statisticiens Ronald Fisher (Londres), Corrado Gini (Rome), Warren Persons (Harvard), Gaetano Pietra (Rome), Frank Zizek (Vienne) et Maurice Fréchet, rapporteur. Plus tard s'y adjoindra le mathématicien Edwin B. Wilson de Harvard, à la demande du bureau de l'IIS, mais il se déclarera « personnellement très peu intéressé par les indices de corrélation »[21]. Maurice Fréchet, au nom de cette commission, prend l'initiative de mobiliser son réseau de correspondants et met en place une enquête auprès d'une vingtaine de statisticiens « particulièrement compétents en statistique mathématique »[22], pour recueillir leur opinion sur les usages licites du coefficient de corrélation. La liste des correspondants ainsi transformés en « faiseurs d'opinion scientifique » est assez impressionnante à la fois par ce qu'elle témoigne des relations épistolaires de Fréchet et par sa représentativité internationale[23]. Tous, bien sûr avec

[21] Lettre du 21/02/1936 à Fréchet, Archive ISUP, carton Fréchet.

[22] « ceux avec qui nous avions été en relations personnelles ou épistolaires les plus fréquentes ou les plus récentes » précise-t-il [Fréchet 1935a, p. 4].

[23] Arthur Bowley (London School of Economics), Edwin B. Wilson (Harvard), Georges Darmois (ISUP Paris), Edward Huntington (Harvard), Henry L. Rietz (Iowa), Emil Julius Gumbel (Institut Supérieur de Finances et Assurances, Lyon), Harold Hotelling (Columbia), R. von Mises (Vienne et Istanbul), René Risser (Centre National des Arts et Métiers, Paris), Charles Jordan (Budapest), Steffensen (Oslo), Paul Lévy

des nuances qui sont intéressantes pour l'historien des statistiques[24], abondent dans le sens de Fréchet en ce qui concerne la signification d'une faible corrélation. Les différentes réponses à cette enquête sont publiées avec le rapport de la commission dans le second tome des comptes rendus de la session de Londres [Fréchet 1935c, BIIS, 28-2][25] et dans la *Revue de l'IIS* [Fréchet 1935a] précédés d'une communication personnelle de Fréchet, « Sur l'usage du soi-disant coefficient de corrélation », qui enfonce le clou :

« Le nombre des publications s'élevant — avec preuves à l'appui — contre l'emploi sans précaution du coefficient de corrélation est assez grand pour qu'on puisse légitimement en conclure qu'un tel abus doit être en voie de disparition. Il n'en est malheureusement rien. Un nombre considérable de travaux continuent à paraître où l'on utilise sans aucune réserve le coefficient de corrélation

[...] Il n'est plus possible de laisser passer de pareilles erreurs, sans mettre en danger l'estime en laquelle doivent être tenues les conclusions des statisticiens. Puisque des voix isolées, quoiqu'assez nombreuses, n'ont pas réussi à se faire entendre de la masse des statisticiens trop confiants, nous proposons de faire entrer en jeu la grande autorité de l'Institut International de Statistique. Nous proposerons à sa prochaine session le vote d'une motion qu'on trouvera plus loin et dont les termes seront jugés sans doute par beaucoup comme trop faibles. Si nous nous sommes contentés de lui donner la forme d'un avertissement, c'est afin de tenir compte de l'objection bien légitime que la vérité d'une assertion scientifique ne peut être décidée par un vote » [Fréchet 1935a, p. 4].

(Polytechnique Paris), Corrado Gini (Rome), A.G. Guldberg (Oslo), Ragnar Frisch (Oslo).

[24] L'école anglo-saxonne par exemple (Fisher, Hotelling) met en évidence les études qui ont été faites en son sein de la distribution de r sous l'hypothèse d'indépendance et les tests d'hypothèse que cela permet. Et il semble clair que ni Risser, ni Darmois, ni Fréchet ni les Italiens n'ont à cette date une idée claire de la théorie des tests d'hypothèses, peu diffusée avant la traduction française en 1947 de *Statistical Methods for Research Workers* [Fisher 1925].

[25] Les travaux des sessions sont publiés en général en deux livraisons du bulletin, la première consacrée aux comptes rendus et publiée avec quelques années de retard (1938 pour la session de Londres 1934) et la seconde consacrée aux rapports et communications publiée plus rapidement (1935 pour la session de Londres), d'où la nécessité de préciser les tomaisons et les années.

C'est bien une motion que Fréchet fait voter à la session de Londres. Le décalage entre la session et la publication des actes lui permet de la reproduire[26] dans les publications du BIIS et de la RIIS de 1935, avec le rapport de la commission et les amendements qui résultent de la discussion. Fréchet a beau écrire dans son rapport de la séance du 17 avril 1934 que « les termes de la motion ne tendent nullement à faire décider par un vote une vérité scientifique »[27], il n'empêche que l'objectif de la motion est bien d'infléchir ce qui doit être considéré comme vrai : « on ne saurait considérer cette assertion comme une vérité mathématiquement démontrée » dit le texte de la motion à propos de l'équivalence entre (quasi) nullité de r et inexistence d'une relation. Les collègues de Fréchet ne s'y trompent pas et certains comme le Polonais M.L. Hersch refusent pour cette raison de procéder au vote[28].

Si la motion est votée à la majorité, on peut néanmoins repérer certaines réticences des statisticiens à suivre Fréchet dans sa condamnation du coefficient de corrélation. Plusieurs lui font remarquer qu'une forte valeur de

---

[26] Le texte complet de la motion proposé (avant amendements) est le suivant :

L'Institut International de Statistique, dans sa session d'avril 1934,

1) Considérant que de nombreux statisticiens emploient le coefficient de corrélation, dit de Bravais-Galton, en croyant – ou comme s'ils croyaient – mathématiquement démontré que ce coefficient est, en tout état de cause, une bonne mesure de la rigueur de la dépendance fonctionnelle entre deux variables statistiques ;

attire l'attention sur les résultats d'une enquête menée auprès de quelques uns des statisticiens les plus connus par leur compétence mathématique ;

et, observant que parmi les réponses à cette enquête, il n'en est aucune soutenant qu'en l'absence de tout autre renseignement, il suffit de savoir que le coefficient de deux variables statistiques X et Y est voisin de zéro pour conclure à l'inexistence d'une relation fonctionnelle entre X et Y, qu'au contraire la plupart s'élèvent contre cette assertion ;

constate qu'on ne saurait considérer cette assertion comme une vérité mathématiquement démontrée, ou, à tout le moins, reconnue universellement comme telle ;

2) en ce qui concerne le cas plus délicat où l'usage du coefficient de corrélation serait limité en ce qui concerne le degré de dépendance, aux observations où la courbe des moyennes est presque linéaire ou au cas plus restreint encore où l'on suppose que les variables X et Y vérifient presque exactement la loi normale,

constatant que certains auteurs élèvent des objections, même dans ce cas particulier, contre l'assertion ci-dessus mentionnée ;

décide que le mandat de la commission sera prorogé pour étudier ces cas particuliers.

[27] *BIIS* 28-1, [Fréchet 1938c, p. 83].

[28] Voir sa réponse à l'enquête communiquée après la session de Londres et reproduite dans le *BIIS* 28-2 [Fréchet 1935c, p. 48–51].

la corrélation est toujours significative et que la linéarité peut être obtenue par un changement de variable, de sorte que, « s'il y a des morts qu'il faut tuer plusieurs fois », c'est « qu'ils doivent être singulièrement viables » et qu'il convient peut-être « de ne pas tenter de les tuer » [intervention de M.L. Hersch, dans Fréchet 1935c, p. 51].

La motion recueille un consensus maximum sur la proposition de proroger le mandat de la commission pour étudier les diverses interprétations que l'on peut attribuer au coefficient de corrélation. Fréchet reprend à son compte l'idée émise par Neyman d'une sorte de bêtisier des usages de la corrélation. Mais comme le montre l'échange de correspondance[29] avec les membres de la commission entre les sessions de Londres (avril 1934) et d'Athènes (septembre 1936), il ne réussit pas à faire condamner publiquement les pratiques des statisticiens en matière de corrélation :

« Sur l'invitation du Bureau de l'Institut, je viens vous prier de donner votre avis sur une question où M. Gini et moi-même ne sommes pas d'accord. [...] Vous vous rappelez peut-être qu'à Londres, j'ai fait à l'Assemblée Générale, la proposition suivante : confier à une commission le soin de préparer un ouvrage rassemblant quelques unes des erreurs les plus typiques commises dans l'usage des indices de corrélation et indiquant comment les éviter. [...] M. Gini pense que l'Assemblée n'a pas ratifié ma proposition. Je pense que l'Assemblée n'ayant pas voté contre, elle l'a ainsi adoptée. »[30]

En fait, Fisher, Persons, Gini et Pietra s'opposent à un tel recensement des « horreurs statistiques » au prétexte qu'il n'a point été décidé en assemblée et que « cela éveillerait des susceptibilités », mais surtout parce qu'une telle entreprise est jugée contre-productive dans les sciences et qu'il vaut mieux « que les auteurs se consacrent à la rédaction d'ouvrages ayant un caractère uniquement scientifique ou un caractère de divulgation » qu'à la dénonciation de mésusages[31]. Fréchet envoie une dernière circulaire en forme de « vote par correspondance » sur l'opportunité d'une telle publication et de l'adjonction de statisticiens non membres de l'IIS (comme Neyman) à la commission, et il demande à ses collègues « de lui envoyer

[29] Voir le carton « corrélation » des archives Fréchet de l'ISUP.

[30] Lettre du 2 novembre 1934 à « Messieurs les membres de la Commission d'étude de l'usage du coefficient de corrélation », Archives ISUP, carton Fréchet.

[31] Lettre de Pietra à Fréchet du 6 avril 1935.

toute suggestion susceptible de l'aider à dresser un questionnaire destiné à orienter les travaux de la Commission »[32]. Au vu des réponses de ses collègues, il doit donc renoncer à la publication de son recueil d'erreurs[33] dans le cadre de l'IIS[34] et se contenter d'annoncer la publication d'un quatrième article, « à titre individuel », dans la *Revue de l'IIS* [Fréchet 1936a].

Dans une première partie de cet article[35], Fréchet revient sur son idée que r peut résoudre la question de déterminer la forme de la relation (question d'ajustement linéaire), mais ne permet pas de repérer une dépendance. En particulier si l'on n'a qu'une valeur de y pour chaque valeur de x (critique de l'exemple de Hersch), la relation fonctionnelle est un artefact. Quant au rapport η de Pearson, il est trop sensible au regroupement des valeurs de x nécessaire au calcul des moyennes liées de y, du moins quand la régression n'est pas linéaire ou quand le nombre de valeurs observées dans une classe de x est trop faible et peut par exemple être nul pour $y = \pm\sqrt{1-x^2}$. Fréchet insiste beaucoup sur la fréquence des mésusages du coefficient de corrélation, qui se trouve confirmée par un article de Furfey et Daly [1935] : ceux-ci ont recensé les emplois de la corrélation dans cinq périodiques de leur discipline et ils ont constaté que sur 268 articles étudiés, 84 ont fait mention ou usage de r (calculé 1354 fois !), mais sept fois seulement une information sur le caractère normal ou linéaire de la relation a été donnée. Dans une seconde partie du même article, ainsi que dans celui paru dans *Barometro economico* [Fréchet 1935b], Fréchet donne une méthode générale de formation de nouveaux indices de dépendance sur laquelle nous reviendrons parce qu'elle a été reprise et développée dans son article d'*Econometrica* en 1947.

Dans la même lettre circulaire de février 1936, Fréchet avait réitéré sa demande d'une note critique sur l'usage du coefficient r pour « repérer la rigueur d'une dépendance fonctionnelle ». Les réponses qu'il reçoit de Fisher (19/02/36) et de Persons (26/02/36) sont parfaitement en

[32] Lettres de Fréchet des 26 et 27 août 1935, ainsi que les réponses de Gini, Pietra, Zizek, Persons , Archives ISUP, carton Fréchet.

[33] Lettre de Fréchet du 8 février 1936, à en-tête de l'IIS, Archives ISUP, carton Fréchet.

[34] Il annonce l'intention de le faire en dehors de l'IIS, mais cette promesse ne semble pas avoir été tenue.

[35] Notons le subtil décalage des titres depuis l'article publié deux ans plus tôt dans la même revue : « sur le coefficient dit de corrélation » est devenu « sur le coefficient de linéarité dit de corrélation ».

opposition : le premier lie totalement la question de la signification de r avec sa significativité, au sens de la théorie des tests d'hypothèse qu'il a lui-même développée[36]. Fréchet a d'ailleurs besoin d'un échange de correspondance spécifique avec Fisher et de l'aide de Georges Darmois pour comprendre ce qu'est cette théorie des tests et pour repérer le rôle qu'y joue l'hypothèse de normalité, ce qui nous offre un témoignage intéressant sur la lenteur de la diffusion en France de la théorie des tests. À l'opposé, Warren Persons affirme que « la corrélation n'est qu'un simple degré de similarité et que l'erreur probable n'a que très peu, voire pas du tout, de signification » (lettre du 26/02/36, Archives ISUP, carton Fréchet). Fréchet fait alors l'hypothèse[37] que cette opposition tient aux domaines des applications faites par l'un aux données agronomiques, par l'autre aux données économiques, qui sont respectivement bien et mal modélisées par une loi normale. Mais cette interprétation est contestée par Persons : sa critique essentielle — une mesure de corrélation ne prouve rien, hors de la distribution et de son graphique — n'est pas, dit-il, restreinte aux données économiques. Zizek ne se prononce pas sur les propriétés du coefficient r, mais sur les limites d'une interprétation hors d'un schéma causal. Gaetano Pietra enfin s'appuie sur ses publications et sur celles de Corrado Gini [1914–1917] pour rappeler la distinction importante faite par l'école italienne entre les deux concepts de *connexion* et de *concordance* que l'on confond trop souvent dans une mesure de corrélation, le η de Pearson et le r de Bravais appartenant respectivement à ces deux classes de mesure [Pietra, 1925]. Cette position est détaillée dans les deux articles de Gini [1936] et Pietra [1936] inclus dans les troisième et quatrième livraisons du *Bulletin de l'IIS*.

Gini, avec le ton quelque peu méprisant qu'il a souvent affiché, est d'accord avec Fréchet pour préférer η à r lorsqu'il s'agit de mesurer

[36] « *... since in the business of discovery, we are only concerned with significance and not necessarily with estimation, or even with knowing what we want to estimate, the correlation coefficient may be legitimately used in this way, even when its numerical value is without scientific meaning beyond that given it by its mathematical definition as a calculable function of observable quantities* » (Lettre du 19/02/1936 à Fréchet, Archives ISUP, carton Fréchet). « *It is in fact only the test of significance in its entirety which is of value to the experimenter, and not the particular statistic in which he happens to find it convenient to carry out the test* » (Lettre du 5/03/1936 à Fréchet, Archives ISUP, carton Fréchet).

[37] Lettre du 7 avril 1936, Archives ISUP, carton Fréchet.

l'étroitesse d'une dépendance statistique, mais ce n'est pas une raison pour abandonner r. Il faut selon lui multiplier les indices — « coefficient de corrélation, indice d'homophilie, indice d'association, coefficient de colligation, coefficient de contingence, indice d'attraction, indice de ressemblance » — si on veut repérer « quelque autre caractéristique des relations statistiques qui semble échapper à M. Fréchet » [Gini 1936, p. 357]. Cette multiplicité des indices est d'ailleurs une tradition de l'école statistique italienne de Messedaglia à Benini, Pietra et lui-même. Gini dégage d'un indice de corrélation plusieurs fonctions autres que la mesure de dépendance (ou connexion), en particulier la caractérisation d'un degré de concordance entre deux séries, quelle que soit la forme de la relation statistique. C'est déjà ce qu'avait observé Lucien March [1905] en dérivant le coefficient de corrélation de Pearson d'une élaboration du coefficient de Fechner[38]. Un indice de connexion, ε, est présenté comme une simple moyenne arithmétique des écarts absolus $|\bar{y}_i - \bar{y}|$ rapportés à leur valeur maximale, tandis que le η de Pearson est une moyenne quadratique de ces mêmes écarts et r une moyenne arithmétique des écarts algébriques pondérés par $|x_i - \bar{x}|$. Pour Gini la différence importante entre r et η (ou ε) est que seul le premier tient compte du signe des écarts, propriété indépendante de la forme linéaire ou non de la relation. Gini conclut donc : « Les multiples applications que l'on a faites et que l'on fait tous les jours du coefficient de corrélation, souvent sans se préoccuper de la forme de la relation statistique, ne sont pas dénuées de sens. Elles ont au contraire un sens bien précis que nous avons mis en évidence » [Gini 1936, p. 365]. Pietra résume de la même façon sa pensée dans la livraison suivante de la *RIIS* : « Le coefficient de corrélation, interprété correctement comme mesure de la concordance, peut être encore employé utilement et sur une grande échelle, indépendamment de la circonstance que la relation soit linéaire ou non » [Pietra 1936, p. 500]. La controverse est donc portée à son comble par les deux auteurs italiens.

La session d'Athènes de l'IIS comporte pour la première fois une quatrième section de statistique mathématique. Après la naissance difficile

[38] Le coefficient de Fechner est fondé sur le nombre de concordances (c) et de discordances (d) des signes des séries centrées (ou encore de leur différence première) : $i = (c-d)/(c+d)$. En pondérant ces concordances et discordances par les variations v et v' des séries et en normalisant l'expression pour la rendre invariante à une homothétie près, March obtient son coefficient de covariation similaire à celui de Karl Pearson.

de l'*Institute of Mathematical Statistics* et l'abandon, par le *Journal of the Royal Statistical Society*, de la publication d'un résumé des travaux mathématiques (devenus trop nombreux), on pourrait y voir un troisième signe en faveur de la thèse de Stephen Stigler[39] d'une irruption de la statistique mathématique au milieu des années trente. Cette session héberge tout naturellement la suite des travaux de la commission animée par Fréchet, dont l'objectif est d'étudier l'usage du coefficient de corrélation. Le 29 septembre 1936, celui-ci y présente un rapport qui semble être un compromis négocié avec l'école italienne[40].

« L'Assemblée Générale de l'Institut International de Statistique, prenant acte du rapport de la Commission d'étude de l'usage du coefficient de corrélation,

constate que le coefficient de corrélation peut être employé indépendamment de la forme de la relation statistique — linéaire ou non — pour repérer le degré de concordance, c'est à dire lorsque des relations statistiques on considère non seulement l'intensité mais aussi la direction de la dépendance, et observant que de nombreux exemples lui ont été soumis, où d'une valeur absolue plus grande du coefficient r de Bravais-Galton on a inféré une dépendance plus grande même dans les cas où les courbes de régression ne sont pas linéaires

attire l'attention des statisticiens sur les objections faites à une telle pratique de différents côtés, et en particulier par plusieurs membres de la dite commission. »

Seuls les deux derniers alinéas étaient proposés par Fréchet. Le second est bien directement inspiré des thèses italiennes. Le rapport plus complet qui est présenté dans la troisième livraison du *BIIS* [Fréchet 1937a] fait la part belle à la théorie de la connexion de Gini et Pietra. Cachant mal l'échec de Fréchet à susciter la rédaction d'un ouvrage consacré aux horreurs de la corrélation, le rapport annonce son intention de reprendre ce projet à titre privé. Nous n'en trouverons pas trace. Fréchet reconnaît enfin avoir tenu compte d'objections légitimes en n'exprimant

---

[39] «*I propose to advance and defend the claim that mathematical statistics began in 1933*» [Stigler 1999, p. 157].

[40] Si l'on sait lire entre les lignes du compte rendu de séance (*BIIS* 29-1, p. 130–131) et du rapport présenté en assemblée générale (*BIIS* 29-1, p. 193–194). Selon le diagnostic d'un membre français – M. Dieulefait – on y voit mélangés les deux points de vues de Gini et de Galton.

aucun jugement d'ordre scientifique dans cette motion. Il n'est plus question de mélanger expressément vérité scientifique et opinion. On peut donc considérer que la controverse se referme — provisoirement — sur un demi-échec de Fréchet : s'il a réussi à poser le problème de la signification précise qu'il faut attacher à un coefficient de corrélation et à dénoncer les usages abusifs de ce coefficient, il n'a pas réussi à les faire condamner officiellement par l'institution représentative de la communauté statisticienne, ni même à dégager en son sein un consensus sur les interprétations licites à donner à ce coefficient. L'IIS peut se transformer en une tribune efficace, mais n'est pas prêt à jouer le rôle d'une instance de normalisation, de légitimation ou de condamnation.

LES RECHERCHES DE FRÉCHET SUR DE NOUVEAUX INDICES DE CORRÉLATION

Fréchet ne s'est pas contenté cependant de dénoncer les usages abusifs du coefficient de corrélation. Dans la seconde partie de son article de la *RIIS* [Fréchet 1936a], il a déjà posé les jalons d'une « méthode de formation de nouveaux indices de dépendance », dont l'exposé le plus complet est publié en 1947 dans *Econometrica* [Fréchet 1947a]. L'idée, déjà clairement formulée en 1936, consiste à rattacher la construction d'un indice de corrélation au problème plus général de la mesure d'une distance dans un espace de fonctions. C'est une idée dont Fréchet reconnaît qu'elle lui vient de sa fréquentation de la théorie des ensembles[41]. Si en effet on définit l'indépendance de Y et de X par le fait que « l'ensemble E_X des valeurs que peut prendre Y (ou encore par extension la loi de probabilité de Y) pour X donné est indépendant de X » alors :

« Pour mesurer ou au moins repérer le degré de dépendance de Y et de X il faudra repérer le degré de différence de ces deux lois. Ces deux lois peuvent être représentées par deux fonctions et les théories mathématiques modernes peuvent encore ici nous être utiles en nous suggérant de repérer ce degré de différence par ce que l'Analyse générale a conduit à introduire sous le nom de distance de ces deux fonctions. Ce sera un certain

[41] « C'est seulement l'esprit nouveau résultant de la connaissance de cette théorie qui par une sorte d'endosmose, s'étend à ceux qui ne la connaissent pas mais sont en contact avec ceux qui la connaissent – qui conduit naturellement à cette définition [de l'indépendance] » [Fréchet 1936a, p. 372].

nombre d_x dont nous indiquerons plusieurs expressions possibles. Ceci fait, on pourra, provisoirement, estimer le degré de dépendance de Y et de X par une valeur moyenne ou plus généralement par une valeur typique w des valeurs prises par d_x quand x varie [...] Mais on aura intérêt à substituer à w un indice I variant de 0 à 1 » [Fréchet 1936a, p. 372–373].

Fréchet impose donc que I soit égal à 0 si et seulement si les deux lois sont indépendantes et $I = 1$ si et seulement si Y est une fonction uniforme de X. Dans ce cas, w atteint un maximum W. Dans ce même article Fréchet propose trois solutions pratiques (d, w, W) au problème tel qu'il vient d'être posé en toute généralité.

La première solution consiste à prendre pour d_x l'écart absolu entre moyennes liées et moyennes de y, pour w leur moyenne quadratique, ce qui conduit à choisir pour I le coefficient η de Pearson.

La seconde solution consiste à prendre pour d_x une distance qui combine écart des moyennes et écart des variances. Elle conduit à un indice original publié par Fréchet dans la revue italienne *Barometro economico* [Fréchet 1935b] :

$$I_1 = \sqrt{\frac{M[(\mu^2 - \mu_x^2)^2] + [M(\mu^2 - \mu_x^2)]^2}{M[(\mu^2 - \mu_x^2)^2] + \mu^4}}$$

dans lequel M représente l'opérateur moyenne et μ^2 une variance.

La troisième solution prend en compte toutes les valeurs des deux distributions à comparer, pas seulement leurs moyennes ou leurs variances, et s'appuie sur l'indice de disssemblance J_x de Gini, dont le maximum est la différence moyenne D_x[42]. Le rapport de leurs moyennes w et W pour différentes valeurs de x produit ce que Gini appelle l'indice de connexion. Fréchet lui reconnaît donc les propriétés d'une distance, d'une plus grande richesse d'information que celle de Pearson. Fondée sur des écarts absolus, cette distance repose sur une norme plus robuste que la norme euclidienne. Cette remarque inaugure un second combat très important de Fréchet pour la médiane[43] et la norme L^1, qui se traduira par plusieurs études sur

[42] C'est-à-dire la moyenne arithmétique des valeurs absolues des mn différences de tous les couples constitués d'un élément de X (parmi m) et d'un élément de Y (parmi n).

[43] On rappelle que la quantité $\frac{1}{n}\sum(x_i - a)^2$ est minimale pour $a = \bar{x}$ (moyenne) et représente la variance (carré de la norme euclidienne L^2) tandis que $\frac{1}{n}\sum|x_i - a|$ est minimale pour $a = m$ (médiane) et représente la norme L^1 .

la médiane et une révision de la notion d'homme moyen chère à Quetelet sous la forme moins paradoxale d'un homme médian [Fréchet 1950a][44].

Dans le journal *Econometrica*, Fréchet [1947a] ajoute un quatrième exemple qui est l'indice de Charles Jordan [1938], fondé sur la remarque que le carré moyen de contingence de Karl Pearson, calculé sur une table de contingence[45], atteint son maximum ($m-1$ ou $n-1$) quand Y est une fonction univalente de X. Le dénominateur du carré moyen de contingence de Karl Pearson est modifié en conséquence :

$$J^2 = \frac{\chi^2}{N(m-1)} \quad \text{ou} \quad J^2 = \frac{\chi^2}{N(n-1)}$$

selon que le nombre n de valeurs distinctes de X est supérieur ou inférieur au nombre m de valeurs distinctes de Y.

Ce n'est pas la moindre des surprises que de découvrir ainsi un lien direct entre les questions très empiriques de la statistique de la corrélation qui mobilisent Fréchet au milieu des années trente et les préoccupations très théoriques de ses travaux sur les espaces abstraits. L'absence de relation que nous supposions *a priori* entre ses travaux mathématiques et son intérêt pour la statistique — qui aurait été en quelque sorte sa danseuse — est ici prise en défaut. Un rapport s'est construit entre les deux activités. Fréchet s'en explique d'ailleurs très directement dans un article du *JSSP* intitulé « les espaces abstraits et leur utilité en statistique théorique et même en statistique appliquée » [Fréchet 1947h]. Le recours à la notion de distance permet non seulement de construire une infinité d'indices de corrélation en l'appliquant à des séries de nombres, mais cette notion peut aussi s'appliquer à des fonctions aléatoires, à des surfaces aléatoires et à des ensembles aléatoires. Il convient alors d'étendre les notions de lois, de convergence stochastique[46] et de moments (valeurs

[44] Voir aussi la discussion qui fait suite à l'article intitulé « Sur une limitation très générale de la dispersion de la médiane » [Fréchet 1940b] : « La médiane sera toujours préférable à la moyenne quand il s'agira d'obtenir un spécimen typique d'une qualité non mesurable mais seulement repérable... (comme une couleur). »

[45] $\chi^2 = \frac{1}{N}\frac{\Sigma(Nn_{ij}-n_in_j)}{n_in_j}$ est la célèbre distance du Chi-2 dans laquelle $n_{ij}n_in_j$ et N sont respectivement l'effectif observé pour $X = x_i$ et $Y = y_j$, les effectifs marginaux correspondants et l'effectif total.

[46] C'est dans ce cadre que Fréchet mobilise la notion de convergence presque certaine développée par Borel et qu'il considère comme « la découverte la plus importante du calcul des probabilités des cinquante dernières années » [Fréchet 1947h, p. 414].

typiques et dispersions) à ces espaces distanciés d'objets mathématiques de plus en plus généraux. Le cas particulier des espaces vectoriels de Wiener et Banach dans lesquels les combinaisons linéaires d'objets ont un sens permet une généralisation des intégrales définissant les moments. Fréchet renvoie ici à ses travaux théoriques comme par exemple ses « Recherches théoriques modernes sur le calcul des probabilités » [Fréchet 1936b] ou l'article qu'il a donné aux *Annales de l'Institut Henri Poincaré* [Fréchet 1948a].

Au bout du détour par les espaces abstraits, Fréchet entrevoit déjà explicitement de nombreuses applications nouvelles de la statistique à des objets techniques comme la forme d'un fil noué jeté sur une table, mais aussi à des objets sociaux comme le contour des villes, la forme des crânes, etc. Pour certains de ces objets, Fréchet entreprendra d'ailleurs des expérimentations et des calculs très laborieux qui visent à comparer les différentes mesures de valeurs typiques ou de corrélation que donnent les applications de différentes formules de distances.

Il faut rappeler ici que Maurice Fréchet s'est donné les moyens d'une expertise et d'une ingénierie des calculs numériques parfois très lourds que nécessitent ces déterminations, sous la forme du laboratoire de calcul de l'Institut Henri Poincaré[47]. Cet institut avait déjà hébergé entre les deux guerres un laboratoire de calcul établi par Émile Borel. L'histoire parallèle des laboratoires de calcul et des laboratoires d'économétrie s'accélère dans les années quarante. Dès l'été 1939, le directeur du Centre National de la Recherche Scientifique Appliquée (CNRSA)[48] avait chargé Louis Couffignal et François Divisia de construire avec René Roy, Eugène Morice et Férignac un centre de recherche en économétrie. Ce projet n'aboutira pas, mais un comité d'économétrie animé par Divisia est institué en mai 1940. Celui-ci s'appuie sur différents instituts (ISUP, ISRES[49], SGF, Ministère des Finances) et sur les laboratoires de Darmois et Couffignal à l'IHP [Bungener et Joël 1989]. Émile Borel, avec le double concours du CNRS et du Ministère de l'éducation, a lui-même fondé en 1940 son propre laboratoire de calcul à l'IHP. C'est à Pâques 1942 que Fréchet succède à

[47] Le carton F4.2 des archives Fréchet de l'Académie des sciences contient de nombreux documents relatifs à ce laboratoire entre 1940 et 1945.

[48] Le CNRSA a précédé la création du CNRS en octobre 1939 [Bungener et Joel 1989].

[49] Institut scientifique de recherche économique et sociale, de Charles Rist.

Valiron à la tête de ce laboratoire, qui prend le nom de Laboratoire de calcul et de statistique (LCS). En mars 1944, Fréchet élabore le projet d'une action concertée entre le LCS et trois autres laboratoires de calcul : le Centre d'Études et de Mathématiques en vue de leurs Applications (CEMA) dirigé par Joliot et Louis de Broglie ; le Laboratoire de calcul expérimental et graphique de Joseph Pérès et L. Malavard ; le Laboratoire de calcul mécanique de Louis Couffignal[50]. Ce projet, complété par une demande d'achat d'un bâtiment commun, aboutira deux années plus tard à la création en 1946 de l'Institut Blaise Pascal. À la même date la section « théories physiques, probabilités et applications » du CNRS crée une commission d'économétrie[51] à laquelle Fréchet participe et qui mettra en place un centre d'économétrie dirigé par Darmois et chargé de superviser les séminaires de Roy, Allais et Eyraud.

Pour l'heure, ces laboratoires de calcul mobilisent dans les années quarante des techniques bien différentes : tandis que Pérès et Malavard travaillent sur des systèmes analogiques électriques ou hydrauliques, Couffignal teste des machines logiques et Fréchet utilise les machines commerciales usuelles du calcul mécanique. Sa direction du LCS le met en contact avec des spécialistes de ce calcul. En témoignent des correspondances avec W.J. Eckert et Leavens de l'*Econometric Society* ou avec Couffignal, qui le mettent au courant des premiers balbutiements de l'informatique, comme par exemple la machine de 51 pieds construite par IBM à Harvard. En attendant, il faut faire des calculs astronomiques avec des machines électro-mécaniques de type Brunsviga, Monroe ou Astra, pour lesquelles le budget temps d'un calcul de médiane d'une loi binomiale pour quelques valeurs des paramètres n et p dépasse les 70 heures[52].

Les archives révèlent également la multiplicité des contacts avec les porteurs d'une demande de calcul issue de disciplines très diverses. Fréchet est sollicité par la Caisse d'épargne pour le calcul de « permanences

[50] Lettre de Fréchet au directeur du CNRS, datée de mars 1944 [Académie des Sciences, F4.2].

[51] Elle est constituée de MM. Allais, Bunle, Darmois, Divisia, Dubourdieu, Dumontier, Fréchet, Hubert, Lutfalla et Roy.

[52] Commande de calcul du 22 septembre 1944, Archives de l'Académie des Sciences, carton F4.2. Voir aussi la lettre de Gumbel à Fréchet du 22 août 1945 au sujet de la machine automatique *Friden* [Hertz 1997, Annexe 1, p. 30].

statistiques », par Couffignal pour des calculs d'intégrales, par les démographes pour des ajustements de tables de mortalité par la formule de Makeham, par Stoetzel tracassé par les fondements théoriques des travaux de Gallup sur les sondages, par le Musée de l'homme pour des études de crânes, etc.

Fréchet, dans l'article de synthèse publié dans *Econometrica* [Fréchet 1947a], articule parfaitement la théorie des espaces abstraits distanciés, les réponses que peut fournir cette approche pour la construction d'anciens ou de nouveaux indices de corrélation ayant les bonnes propriétés limites (en 0 et en 1) et le calcul numérique de ces indices sur des données empiriques. Ces trois approches, et les trois compétences dont elles témoignent, se révèlent parfaitement complémentaires, en particulier dans les deux études qui illustrent cet article. La première s'empare de la fameuse question des co-variations de séries économiques et du décalage temporel optimal entre les séries qui avait servi de prétexte à la campagne de Fréchet à l'IIS. Ici Fréchet simule deux séries $X(t)$ et $Y(t) = f[X(t-\alpha)] + \xi(t)$, dont on sait d'avance pour quel retard α elles présentent une corrélation très étroite (dont l'intensité ne dépend que de l'aléa $\xi(t)$ introduit) et il cherche à retrouver ce retard en leur appliquant ensuite les coefficients de corrélation r de Bravais-Pearson, g de Gini, d de Fréchet et j de Jordan. Curieusement ses résultats n'invalident aucun des coefficients, pas même celui de Bravais-Pearson qu'il avait tant critiqué, puisque les quatre indices donnent un minimum pour le même retard α. Cependant les calculs de sensibilité montrent bien que le coefficient de corrélation linéaire est le moins efficace dans la discrimination des retards.

Le second exemple est assez surprenant puisqu'il part de données de contour d'une section crânienne observée pour une centaine d'individus des Canaries. Ces contours sont donnés par une fonction empirique définie en coordonnées polaires que l'on ajuste par un développement en séries de Fourier. On calcule les premiers coefficients a_k, b_k de ce développement et l'on s'intéresse à la corrélation[53] de certains couples a_k, b_j de coefficients pour la centaine de contours crâniens. Fréchet, au bout de calculs assez fastidieux faits au Laboratoire de calcul et de statistique[54], arrive aux

[53] Cet exemple offre un raccourci saisissant de l'histoire de la corrélation qui avait commencé par un emprunt de l'école biométrique anglaise à la biologie et qui s'incarne ici dans une application à la craniométrie.

[54] On peut même trouver dans le carton F4.3 de l'Académie des sciences une

mêmes résultats quant à la sensibilité comparée des quatre indices quand on passe d'un couple à un autre.

Revenons maintenant à la question que nous avions posée au début : qu'est-ce qui fait courir ainsi le mathématicien Fréchet dans son combat contre les mésusages du coefficient de corrélation et qu'est-ce qui justifie une telle campagne à l'IIS à base d'enquêtes, de rapports et de motions ? Une première réponse évidente résulte des travaux que nous venons d'analyser. Fréchet, en bon mathématicien, armé de la notion de distance dans un espace abstrait, est persuadé que les outils de la statistique comme le coefficient de corrélation, mais aussi la moyenne et l'écart-type, sont utilisés et interprétés bien au delà de ce qu'autoriserait un respect de leur domaine de validité. Cette critique, écrit Fréchet [1940b, p. 68], est « un cri d'alarme mais en même temps un cri d'espoir », car il est vraisemblable qu'en adaptant au mieux les outils statistiques aux circonstances, on arrive à la fois à une meilleure validation des outils existants dans certaines conditions et à un plus grand éclectisme dans les autres cas, voire à une « diminution de l'ampleur des calculs ». C'est donc bien le souci d'un fondement mathématique solide des pratiques du statisticien qui anime Maurice Fréchet dans ses campagnes.

Cependant ce souci fondamentaliste ne serait pas suffisant en soi pour justifier l'ampleur et la ferveur de ces interventions. Pourquoi en effet se soucier des pratiques plus ou moins approximatives des statisticiens et pourquoi ne pas faire comme la majorité des mathématiciens : les ignorer purement et simplement et cultiver le champ proprement labouré de la mathématique pure et des espaces abstraits, sans se soucier davantage de tous ceux qui pataugent dans le marécage des questions pratiques, techniques ou politiques. Pour en arriver à un tel niveau d'engagement dans le débat statistique, il fallait que Fréchet ne prenne pas la posture classique du mathématicien dégagé des contingences matérielles de la mathématique appliquée. Il fallait qu'il eût au contraire une philosophie particulière du rapport — épistémique et social — entre mathématiques pures et mathématiques appliquées qui lui fasse voir des enjeux plus intéressants lorsqu'on les articule que lorsqu'on les oppose.

---

comparaison des durées de calcul de chaque indice (de 30 à 100 minutes).

LA PHILOSOPHIE DES SCIENCES MATHÉMATIQUES DE FRÉCHET

Déjà dans sa préface du *Calcul des probabilités à la portée de tous*, Fréchet se présentait comme « un mathématicien de profession, [qui juge] que le savant le plus préoccupé de recherches spéculatives ne doit pas se désintéresser de la pratique, et qu'il est d'ailleurs utile au progrès de la science d'en diffuser (nous ne disons point : d'en vulgariser) les résultats » [Fréchet et Halbwachs 1924, p. vii]. Fréchet ne se départira jamais de cette position, sans doute héritée du culte de son maître Émile Borel pour « la valeur pratique et philosophique des probabilités », déjà présent dans *Le hasard* (1914) et donnant son titre au dernier des 16 fascicules du grand *Traité de calcul des probabilités et de ses applications* publié par Borel entre 1925 et 1939. Une lecture même rapide des deux ouvrages consacrés par Fréchet [1965, 1967] à Borel suffit à nous convaincre que le premier partageait avec le second une même conviction : si l'on pouvait développer une théorie purement mathématique des probabilités en chassant par la porte toute idée d'application, alors c'était précisément dans la phase de ses applications aux phénomènes réels que les difficultés surgissaient et réintroduisaient le réel par la fenêtre du laboratoire.

Cette position est précisée par Fréchet lui-même en de nombreuses occasions — discours, conférences, introductions — et dans le corps même de plusieurs textes mathématiques. L'ouvrage *Les mathématiques et le concret*, qu'il publie à la fin de sa carrière [Fréchet 1955b], reprend plusieurs de ces textes philosophiques et s'organise presque entièrement autour de ce thème. Le premier texte « sur les mathématiques en général », issu d'une conférence à Berne [Fréchet 1925e], s'orne même d'un sous-titre provocateur — « sur une désaxiomatisation de la science » — qui en dit long de sa méfiance vis-à-vis d'une approche purement axiomatique. Non pas qu'il la rejette, puisqu'il concède même s'en être servi avec persévérance dans ses travaux, mais il trouve indispensable de la compléter « par un travail inverse de désaxiomatisation [...] pour les sciences ayant déjà atteint un haut degré d'abstraction » [Fréchet 1955b, p. 3]. Comme il le précise dans le texte suivant du même recueil sur « les origines des notions mathématiques », le rapport entre concret et abstrait est bien plus subtil que celui d'une application au réel de théories établies de façon déductive. Parce que les hypothèses et axiomes choisis comme prémisses ne sont pas créés *ex-nihilo* :

« La théorie déductive n'est pas une création spontanée, elle doit sa naissance à l'œuvre collective préliminaire, dispersée, trop souvent oubliée et pourtant capitale des savants qui, de la complexité des choses se sont efforcés peu à peu de dégager les idées simples constituant la base des théories. M. Destouches en donnant un nom à cette œuvre préliminaire, celui de synthèse inductive, en a fait ressortir le caractère particulier. Ses observations, qui s'adressent à la théorie physique sont en bien des points applicables à la théorie mathématique » [Fréchet 1955b, p. 27].

Même une science purement déductive comme les mathématiques a pour Fréchet des fondements expérimentaux « et nous ne sommes pas loin de conclure que la logique elle-même est un produit de notre expérience » [Fréchet 1955b, p. 22]. Ce qui ne l'empêche nullement de réclamer « pour les mathématiciens, le droit de créer dans l'abstraction des notions nouvelles, qui ne soient pas nécessairement formées à l'image de faits rencontrés dans le monde sensible » [Fréchet 1955b, p. 35] et de reconnaître un rôle essentiel à la méthode du « prolongement » pour « porter une notion du dehors au dedans des Mathématiques » [Fréchet 1955b, p. 37]. Il suffit simplement de le savoir et de distinguer, comme le faisait Jacques Rueff, un des pères de l'économie mathématique, les sciences euclidiennes, qui ont un objet empirique correspondant à leurs énoncés, des sciences non euclidiennes, qui n'en ont pas (ou pas encore). Évidemment cette position expérimentaliste de Fréchet n'est pas du goût de tous les mathématiciens et philosophes. Elle heurte un platonicien comme Enriques qui lui répond que « la véritable signification des Mathématiques, c'est d'étudier des êtres qui n'appartiennent pas au monde extérieur, des êtres intelligibles au sens de Platon » [cité dans Fréchet 1955b, p. 47]. Mais la position de Fréchet n'est jamais dogmatique. Il dit seulement que la méthode hypothético-déductive n'est qu'une phase de la démarche mathématique :

« En résumé nous croyons à la nécessité de la décomposition de chaque branche des mathématiques (comme des sciences physiques) en quatre parties : synthèse inductive, dégagement à partir de celle-ci d'un ensemble d'axiomes portant sur des termes primitifs, théorie déductive basée sur ces axiomes et ces termes, vérification des conséquences de cette théorie quand on substitue aux notions abstraites qui y figurent, les notions concrètes qu'elles ont pour but de représenter schématiquement » [Fréchet 1955b, p. 22].

Une des conséquences immédiates de cette vision des mathématiques est le rôle que Fréchet accorde à l'enseignement des mathématiques et, dans l'enseignement des mathématiques, aux approches expérimentales :

« Il est absolument indispensable que le professeur établisse un lien entre la définition expérimentale et la définition abstraite [...] J'ai déjà eu d'ailleurs l'occasion de critiquer la méthode dogmatique dans ma leçon inaugurale de l'Université de Strasbourg » [Fréchet 1955b, p.7].

Le projet d'introduction du calcul des probabilités dans l'enseignement moyen, défendu par Fréchet à la commission Langevin de 1947 [Fréchet 1947k, p. 286], prévoit de faire précéder cet enseignement par celui de la statistique mathématique « pour faire mieux apprécier par les élèves l'étendue et l'importance des applications du calcul des probabilités », décision qui obtient l'assentiment de Allais, Amy, Chapelon, Dumas, Fortet, Hénon, Morice, mais pas de Baticle et Borel qui souhaitent commencer par les jeux !

Retracer la position expérimentaliste de Fréchet sur les mathématiques n'est pas anecdotique pour notre sujet. Cela permet de comprendre un peu mieux comment un mathématicien rompu aux théories des espaces abstraits, base importante de la topologie, peut aussi s'intéresser à la statistique. L'attention de Fréchet porte non seulement sur les applications éventuelles des notions de mesure et de distance élaborées dans les espaces de fonctions, mais aussi sur les problèmes que se posent les statisticiens et qui constituent ce qu'il appelle la phase de *synthèse inductive*. Très clairement, c'est parce que les statisticiens dégagent de leur problématique de corrélation une idée de ressemblance-dissemblance-liaison, qui peut s'exprimer comme une distance, que le mathématicien peut retravailler en toute généralité cette idée de distance et trouver des réponses nouvelles, originales et bien adaptées à la nature des situations analysées.

Pourtant cette réponse philosophique ne suffit pas totalement à éclairer l'engagement de Fréchet dans une campagne contre le « soi-disant coefficient de corrélation » [Fréchet 1935a, titre]. Certes, ce positionnement lui interdit le repli sur la mathématique pure et explique son intérêt pour les questions expérimentales posées par les statisticiens. Mais pourquoi considère-t-il que l'enjeu de cette question est considérable au point de porter le débat sur la place publique, dans l'arène de la communauté statisticienne ? Fréchet y voit un enjeu de société qui dépasse le seul débat philosophique. Et de cela il s'explique effectivement tout au long de la controverse.

Déjà dans son intervention à la session de l'IIS à Londres, il avait reconnu que sa motion n'avait pas pour seul objectif le rétablissement de la vérité scientifique. Le second objectif de cette motion, comme de l'enquête, était d'agir sur les pratiques mêmes de la communauté scientifique, car « la plupart des statisticiens n'ont ni le goût ni le loisir de discuter le bien fondé mathématique des notions qu'ils utilisent [et] ils s'en remettent tout naturellement à leurs collègues mathématiciens, de même que les mathématiciens s'en remettent aux médecins pour les traitements de leurs maladies » [Fréchet 1935a, p. 4]. Derrière cette pratique de la motion pointe donc une sorte d'analyse sociale de l'activité statistique, qui distingue entre praticiens et théoriciens, et qui interprète les mésusages du coefficient de corrélation comme le signe d'une coupure sociale assez dangereuse entre les deux communautés des statisticiens et des mathématiciens. Ce à quoi L. Hersch avait répondu que « l'Institut aura fait assez d'honneur aux statisticiens qui ignorent le sens véritable des procédés essentiels qu'ils emploient eux-mêmes ou qu'ils enseignent aux autres » [Fréchet 1935c, p. 48–49].

Dans le texte qu'il publie « à titre individuel » dans la *RIIS*, Maurice Fréchet revient sur l'opportunité de sa campagne et de la motion qu'il a fait voter. Il s'en prend vivement à ceux qui, comme Hersch, minimisent le rôle des erreurs faites par les statisticiens et qui soutiennent soit que « la connaissance de la théorie de la corrélation est assez répandue pour qu'il n'y ait pas à revenir sur des vérités bien connues », soit que « la méthode de la corrélation est appliquée de façon incorrecte uniquement par des auteurs qui ne méritent pas qu'on tienne compte de leurs écrits » [Fréchet 1936a, p. 369]. Ces deux propositions contradictoires, toutes deux fondées sur un certain mépris des statisticiens, ne sont pas du goût de Fréchet qui développe une approche autrement plus intéressante du rapport entre les deux communautés :

« Il ne faut pourtant pas oublier que les statisticiens sont d'origines très diverses, qu'un nombre important d'entre eux sont des hommes très distingués qui : médecins, économistes, psychologues, etc., se sont fait un renom dans leurs spécialités respectives sans avoir eu nécessairement une forte éducation mathématique. Ils ont fait cependant l'effort de se remettre à l'école pour devenir capables d'utiliser les instruments de la statistique mathématique. Ils ont fait confiance aux mathématiciens pour la démonstration des règles à employer, de même que les mathématiciens font confiance aux médecins quant aux raisons qui dictent leurs ordonnances. Les écrits de ces statisticiens un peu improvisés peuvent être très importants en ce qui concerne le traitement statistique de certaines questions de médecine, d'économie politique, de psychologie, etc., parfois beaucoup plus important que ceux des statisticiens de profession traitant les mêmes sujets. Il est donc nécessaire, non pas de les critiquer, encore moins de les ignorer, mais au contraire, sans détruire leur confiance légitime dans les méthode statistiques, de les avertir des réserves qu'elle comporte. Il faut détruire l'impression que laissent certains ouvrages que la Statistique mathématique est d'un usage automatique, qu'elle est aussi commode à employer qu'un barème, une table de logarithmes, ou une machine à calculer » [Fréchet 1936a, p. 369–370].

Cette analyse du rapport social entre deux communautés de spécialistes est tout à fait moderne. Elle préfigure une vision sociale de la science, aujourd'hui plus banale, dans laquelle la compréhension des acteurs, leur positionnement, leurs objectifs, leurs interactions, sont aussi importantes que la compréhension des transformations cognitives des objets qu'ils manipulent. Les producteurs de formalismes et les utilisateurs-adaptateurs de ces mêmes formalismes, placés dans des contextes disciplinaires très variés mais suffisamment particuliers pour justifier l'existence de spécialistes (comme en économétrie, en psychométrie...), n'ont ni les mêmes bagages, ni les mêmes objectifs. La qualité des énoncés scientifiques produits, aussi bien que le progrès qui peut en résulter pour la société dans son ensemble, sont terriblement dépendants du type de contrôle qu'ils exercent les uns sur les autres, les statisticiens en fournissant des problèmes aux mathématiciens, les mathématiciens en jouant sérieusement leur rôle d'experts auprès des statisticiens.

La leçon n'est pas isolée. Dans le rapport sur l'enquête relative à l'estimation statistique (présenté *infra*), on trouve une mise en garde équivalente et la précision suivante :

« En particulier, les biologistes, les psychologues, les ingénieurs, etc., qui usent des méthodes d'estimation statistique des paramètres ne sont pas toujours très bien informés par les manuels de statistique mathématique de la signification exacte des tests de signification, des intervalles de confiance, et surtout des coefficients de confiance des probabilités fiduciaires. En outre, ils ne savent pas qu'il y a entre les mathématiciens de graves divergences sur ces sujets, qu'il n'y a même pas deux écoles tranchées, mais une variété infinie d'opinions » [Fréchet 1947c, p. 364].

Cette analyse des rapports entre des groupes sociaux différents qui constituent la communauté statisticienne se double chez Fréchet d'une vision historique de la mathématique appliquée très percutante, qui renforce l'enjeu d'une vigilance vis-à-vis des usages d'outils mathématiques par les statisticiens. En effet nous avons vu que de nombreux signes objectifs[55] militent en faveur d'une hypothèse de rupture dans la méthodologie statistique et d'irruption de la statistique mathématique dans le début des années 1930. Maurice Fréchet en est un acteur conscient. Prenant la présidence de la Société statistique de Paris en 1948, il remarque [Fréchet 1948c] qu'il est après Borel (1922) et Darmois (1938) le troisième mathématicien à présider cette société plus traditionnellement considérée comme le fief des économistes libéraux. Il prononce un discours résolument moderniste et met l'accent sur les progrès énormes de la méthode statistique dans la dernière décennie, qu'il situe dans une perspective historique de long terme : la statistique, d'abord associée à — voire confondue avec — la démographie et l'économie politique, prenant son autonomie comme méthode à la fin du XIX$^{\text{e}}$ siècle, puis se tournant vers les mathématiques dès les années 1920. Ce mouvement historique irrésistible vers la formalisation s'est accompagné d'une demande de plus en plus forte d'instruments de la preuve faite aux statisticiens, dans le cadre de l'explosion des sciences sociales et des sciences expérimentales. Ce second mouvement l'emportant sur le premier, les statisticiens ont « répondu de leur mieux » à cette demande, mais leurs méthodes ont manqué le plus souvent « de rigueur dans la démonstration, de précision dans l'énoncé des résultats et surtout des conventions simplement plausibles sur lesquelles elles étaient fondées » [Fréchet 1948c, p. 88]. Cela justifie alors aux yeux de Fréchet un certain interventionnisme des mathématiciens. D'autant plus qu'il repère des mouvements assez divergents des écoles nationales : « L'envolée de

[55] Comme la part des articles méthodologiques dans les revues statistiques, la classification des traités de la discipline, l'existence d'une nouvelle section « mathématique » à l'IIS, la rupture entre l'ASA et l'IMS, etc.

l'École anglaise » s'est faite parfois dans une tradition de pragmatisme qui a oblitéré tout souci de rigueur, mais heureusement l'École italienne a montré l'exemple « en ne suivant qu'avec prudence et des réserves parfois excessives » cette envolée [Fréchet 1948c, p. 88]. Il place aussi tous ses espoirs dans la nouvelle École américaine (sans doute pense-t-il à Rietz) et dans la tâche entreprise en France par Borel, Darmois, ... et lui même, dans une tradition rationaliste-patriote qu'il n'hésite pas à reprendre : « il est rationnellement de la vocation de la France d'apporter ordre et lumière là où régnaient confusion et obscurité[56] » [Fréchet 1948c, p. 89].

L'enjeu de la bataille sur la corrélation apparaît maintenant dans toute son épaisseur. C'est à la fois un enjeu conceptuel (mesurer quoi exactement ?), un enjeu de la pratique statistique (quel indice choisir dans la multiplicité des solutions techniques ?), un enjeu pour la science mathématique (à quelle mesure, dans quel espace, muni de quelle distance faut-il rattacher la mesure empirique choisie ?), un enjeu d'organisation de la communauté scientifique (il y va du contrôle mutuel des statisticiens et des mathématiciens), un enjeu historique de l'institution d'une discipline nouvelle (la statistique mathématique écartelée entre hyper-formalisme et empirisme) et un enjeu de société (à qui appartient l'expertise et comment la contrôler à la fois scientifiquement et démocratiquement ?).

Voilà qui justifie grandement de considérer que la question de la bonne méthode en statistique est suffisamment importante pour ne l'abandonner ni aux spécialistes, ni aux seuls usagers. Voilà qui justifie un débat d'opinion, une enquête, un rapport et une motion. La question, scientifique, est en effet aussi politique, relevant de la politique de la science comme de celle de la décision.

Cette méthode des enquêtes et du débat d'opinion n'est pas attachée à la seule controverse de la corrélation : c'est encore elle, soutenue par la même philosophie, que Fréchet a mobilisée dans le cas de l'estimation et dans celui de l'application des mathématiques aux sciences sociales. Et bien que nous ne puissions dans cet article les aborder avec le même détail

[56] On est à une époque où le patriotisme a des raisons de s'exprimer avec emphase : Fréchet rappelle dans son discours que l'ancien président de la SSP devait s'associer sans broncher à la devise « Travail, Famille, Patrie » et que lui même fut censuré par les autorités allemandes pour un article dans lequel il avait osé utiliser des statistiques sur les rivières de Tchécoslovaquie et les îles de Yougoslavie, « des pays que les armées allemandes avaient fait disparaître » [Fréchet 1848c, p. 92].

que la querelle sur la corrélation, les éléments qui suivent renforceront la thèse d'une unité stratégique des positions scientifiques et politiques de Maurice Fréchet.

L'ENQUÊTE DE FRÉCHET SUR L'ESTIMATION

Nous avons déjà fait allusion aux raisons qui ont poussé Fréchet à faire une enquête sur les méthodes d'estimation en citant longuement son introduction au rapport présenté à la réunion commune de la section de méthodologie statistique de l'IIS et de l'*Econometric Society*, qui eut lieu à Washington en septembre 1947. Ici aussi l'enquête a été décidée par le bureau de l'IIS, Fréchet en est l'organisateur et le rapporteur, mais les conditions ont cependant bien changé.

À la sortie de la guerre, après quasiment dix années d'interruption, la session de Washington marque un tournant important. Comme le détaille le rapport du bureau à la session suivante (Berne 1949), l'IIS a perdu un quart de ses membres et l'usage d'une bonne partie de ses subsides, bloqués par la mort de son trésorier depuis 1941. Il a subi de plein fouet la concurrence des nouvelles agences statistiques liées à la Société des Nations, puis à l'Organisation des Nations Unies[57], et il a perdu l'exclusivité en matière de conventions et d'accords intergouvernementaux dans le domaine de la statistique. De nouvelles élections, un nouveau budget et de nouveaux statuts seront mis en place au cours de cette session de Washington. Par exemple le rôle de l'affiliation directe à l'IIS de sociétés constituées change la donne. L'IIS s'est ouvert aux autres disciplines en admettant dans la *Revue* des articles d'auteurs non membres de l'Institut. Le bureau a dissout 19 des 28 commissions d'avant-guerre. Pour la première fois, l'Institut admet un membre de sexe féminin. Il n'y a pas que l'organisation qui a changé : la science statistique a fait un bond gigantesque, en particulier dans les domaines militaires, industriels, biométriques, économiques et dans la mesure de l'opinion publique. Il convient donc de minimiser le rôle de coordination des activités statistiques des différentes nations et de renforcer la dimension

[57] Il est vrai que le Conseil économique et social de l'ONU a reconnu à l'IIS un statut consultatif de catégorie B qui lui permet encore de faire valoir ses positions dans les organismes internationaux.

« académie internationale » d'échange scientifique, puis aussi de mettre sur pied un enseignement large de cette nouvelle discipline.

C'est dans ce cadre rénové que Fréchet lance ses nouvelles enquêtes. À vrai dire, celle qui porte sur l'estimation a commencé bien avant par une abondante correspondance internationale sur la théorie des tests et la méthode fiduciaire de Ronald Fisher. On en retrouve la trace sous forme d'une trentaine de lettres datant du début de la guerre (août 1939–avril 1940) et figurant dans différentes archives Fréchet[58]. Comme il l'écrit dans une lettre circulaire[59], le prétexte de cette enquête par correspondance est le cours de statistique mathématique dont il est chargé en raison de la guerre et sans doute du départ de Darmois à Londres. Il tombe alors sur plusieurs écrits de Ronald Fisher, dont la 11e édition de *Statistical Methods for Research Workers*, qui lui inspirent les plus grands doutes, doutes qui se voient renforcés par une critique de Gini publiée dans la *Rivista di politica economica* en 1939. Ces doutes proviennent principalement de la confusion qu'il repère entre constante et variable d'échantillonnage dans la théorie fiduciaire que Fisher tente d'imposer à partir de 1935, afin de relayer l'arbitraire de ce qu'il appelle la méthode (bayesienne) de la probabilité inverse[60]. Cette théorie ne survivra du reste pas au paradoxe de Behrens et à la théorie concurrente des intervalles de confiance et des tests d'hypothèse de Jerzy Neyman. Les lettres de Fisher (de plus en plus agacé), et les publications qu'il y joint, ne convainquent pas Fréchet (il le dit dans une lettre du 12 avril à Fisher), ni d'ailleurs nombre de ses correspondants, de la validité de l'argument de la « fiducial probability ». Cette correspondance de Fréchet avec Fisher renseigne aussi sur un contexte dans lequel les statisticiens américains sont de plus en plus engagés dans les travaux de guerre.

N'entrons pas davantage dans le fond de la querelle, car ce qui importe pour notre argument est uniquement de voir Fréchet, en pleine guerre,

58 Ces lettres proviennent des trois fonds d'archives de l'Académie des sciences, de l'ISUP et de la chaire de calcul des probabilités de la faculté des sciences. Je remercie Bernard Bru de m'en avoir communiqué une copie il y a une dizaine d'années pour un exposé sur l'histoire de la notion d'estimation fiduciaire chez Fisher.

59 Lettre circulaire de janvier 1940, carton F4.4, envoyée à Bernstein, Kolmogorov, Slutzky, Khintchine, Wilson, Doob, Dodd, Fry, Lévy, Darmois, Borel, Dieulefait, Cantelli, de Finetti, Cramer, Anderson, Castelnuovo, Steffenson.

60 Voir notre article [Armatte, 1988] sur la théorie de l'estimation chez Fisher.

poursuivre de ses questions, parfois naïves, parfois insolentes, celui que l'on peut encore considérer comme le maître de la statistique mathématique et prendre le reste de la communauté à témoin de son manque de rigueur dans le maniement du calcul probabiliste, allant même jusqu'à réclamer la liste des axiomes qui définissent la probabilité chez Fisher !

L'enquête publiée en 1947 par l'IIS, porte de fait sur un sujet très voisin de l'estimation : « les raisonnements présentés sont souvent obscurcis par l'intervention des calculs mathématiques complexes nécessités par les exemples traités, calculs qui pourtant ne touchent pas au fond des questions en litige, lesquelles ne relèvent que des principes du calcul des probabilités et du sens commun ». C'est pourquoi Fréchet propose un problème relativement simple : « Un événement E a été observé r fois dans n épreuves. Sachant (ou admettant) que E y avait une probabilité constante (mais inconnue) p, que peut-on dire sur la valeur inconnue de p, connaissant seulement le nombre r de répétitions de E (ou sa fréquence $f = r/n$) ? Pour simplifier la réponse et la comparaison des réponses, on ne demande d'examiner que le cas où $n = 1$ et $r = 1$ » [Fréchet 1947c, p. 363].

C'est vraiment un piège dangereux tendu aux statisticiens que de leur demander une estimation d'une proportion sur la base d'une observation unique ! Mais il faut reconnaître que la science statistique est mal partie si les réponses divergent pour ce cas d'école (sans intérêt pratique) : « Si on ne peut résoudre ce problème, on ne peut en résoudre aucun » répond Ville [Fréchet 1947c, p. 365]. D'autres pourraient arguer que la science des grands nombres n'a que faire d'une telle question. Jacques Bernoulli aurait, en son temps, répondu $r/n = 1$, tandis que Laplace (1774) et Condorcet[61] auraient eu recours à la « règle de succession » déduite du calcul de Bayes : $\frac{r+1}{n+2} = \frac{2}{3}$. À moins qu'ils aient tous refusé de répondre pour $n = 1$.

Fréchet obtiendra 16 réponses[62] dont aucune convergence ne se dégagera.

[61] Dans les manuscrits édités par B. Bru et P. Crépel, *Condorcet. Arithmétique politique*, Paris : INED, 1994, ou dans l'*Essai* de 1785 (Problème I). En fait, Condorcet lui-même propose aussi d'autres solutions en fonction d'hypothèses spécifiques sur la stabilité du phénomène [Crépel 1988].

[62] Ces réponses sont celles de R. von Mises, R. Fortet, B. de Finetti, H. Eyraud, M. Dumas, M.S. Bartlett, J. Neyman, G. Darmois, L.H.C. Tippett, M. Kendall, R.C. Geary, L. Féraud, O. Anderson, J.B.D. Derksen, C. Jordan, H. Cramer.

Von Mises répond « rien », Neyman « toute valeur entre 0 et 1 », pour Darmois « E est possible », Eyraud donne $\frac{2}{3}$, Dumas 1, etc. Fréchet reporte de manière détaillée les réponses, les embarras et les contradictions, puis fait le lien avec son enquête de 1940 sur la probabilité inverse et la probabilité fiducaire. La preuve est faite de la divergence des positions et de ce que « la théorie de l'estimation est essentiellement conditionnée par l'opinion que l'on a sur la nature de la probabilité » [René Fortet cité par Fréchet 1947c, p. 392].

« Non seulement les réponses ne sont pas unanimes, mais elles ne se sont même pas réparties en faveur de deux jugements distincts entre lesquels il aurait suffi d'écarter le faux. Ce n'est même pas une gamme ascendante, c'est un réseau où d'abord la pensée du lecteur hésite, embarrassée [...] Ce n'est là qu'un début dans la discussion des méthodes d'estimation, méthodes que des nécessités pressantes ont suscitées mais dont la partie mathématique, malgré sa solidité propre, ne doit pas faire illusion sur la fréquente fragilité de leurs bases », dit Fréchet dans son discours d'entrée à la présidence de la SSP [Fréchet 1948c, p. 89]. Le but de l'enquête, qui était d'attirer l'attention sur le manque de consensus dans la communauté des mathématiciens s'occupant de statistique, est atteint. L'impression de confusion doit alors laisser place à un véritable travail d'éclaircissement conceptuel. Contrairement à ce qui s'est passé pour la corrélation, Fréchet s'est contenté ici de révéler un état de fait sans contribuer lui-même à la mise au point d'une nouvelle méthode d'estimation. Peut-être parce que l'estimation statistique ne peut pas se ramener à la topologie, mais constitue plutôt un problème de logique inférentielle (Fisher) ou de décision (Neyman), comme les travaux des années 1950 le montreront.

L'ENQUÊTE DE FRÉCHET SUR MATHÉMATIQUE ET SCIENCE SOCIALE

À la réunion de Washington, toujours à la requête du bureau de l'IIS, Fréchet a présenté les résultats d'une seconde enquête (la quatrième en fait) sur un autre sujet brûlant de l'immédiat après-guerre : « Dégager les possibilités et les limites de l'application des sciences mathématiques et en particulier du calcul des probabilités, à l'étude de phénomènes économiques et sociaux » [Fréchet 1946b, p. 16]. C'est le titre exact — qui ressemble plus à un sujet de baccalauréat qu'à une enquête scientifique —

de la question que Fréchet a posée à un aréopage de savants. En fait ce sont deux enquêtes qu'il avait prévues.

« C'est au secrétaire général de l'Institut International de Statistique, M. Methorst qu'est due l'idée heureuse d'organiser une enquête sur le sujet ci-dessus. Sujet très controversé où s'affrontent les points de vue les plus divers, mais où ceux qui prennent part à la discussion ne se comprennent pas toujours très bien. Il m'a paru qu'on pourrait essayer de sortir de cette situation en organisant successivement deux enquêtes. Celle-ci s'adresserait de préférence aux économistes ayant reçu une formation mathématique avancée[63], l'autre aux économistes classiques[64] » [Fréchet 1946b, p. 16].

Nous n'avons trouvé aucune trace de cette seconde enquête et ne savons pas si elle a été réalisée. Quant à la première, on trouve la communication de Fréchet dans les *Proceedings* de Washington, le rapport dans le *BIIS* [Fréchet 1947b] et le dossier complet de l'enquête dans la *RIIS* [Fréchet 1946b] avec la publication intégrale des réponses. L'intérêt de ces matériaux est évident : les réponses, que nous n'analyserons pas ici en détail[65], offrent un remarquable panorama des conceptions de ceux que Fréchet appelle des « économistes ayant reçu une formation mathématique avancée » en matière de mathématiques appliquées à l'économie, au moment même où la poignée de chercheurs de la Cowles Commission — dont Jacob Marschak est ici le représentant — est en train d'imposer au monde entier sa conception de l'économétrie, et plus précisément de la modélisation structurelle stochastique[66]. La diversité

[63] Fréchet a obtenu les réponses de Harold Hotelling (University of North Caroline), Oscar Anderson (Kiel), Irving Fisher (New Haven, Connecticut), J.B.D. Derksen (La Haye et ONU), Henri Eyraud (ISFA Lyon), Bruno de Finetti (Trieste), Jacques Rueff (Paris), Jan Tinbergen (La Haye), L.V. Furlan (Bâle), Luigi Amoroso (Rome), V. Rouquet la Garrigue (Bordeaux), Jacob Marschak (Chicago), René Roy (Paris), Maurice Allais (Paris), Georges Darmois (Paris), Lucien Féraud (Genève), K.-G. Hagstroem (Stockholm), ce qui représente un joli panel des économistes mathématiciens et statisticiens, auxquels il faut ajouter les citations de Bernard Chait et de Georges et Édouard Guillaume.

[64] « Nous proposons en outre que la seconde enquête soit organisée maintenant en dehors de l'Institut International de Statistique par une association internationale suffisamment représentative des seuls économistes et sociologues classiques » dit Fréchet [1946b, p. 16].

[65] Voir [Armatte 1995, chap. 13].

[66] Cette conception repose sur l'idée qu'aucune mise en relation d'une théorie

des réponses obtenues cache mal le désarroi des économistes avertis, dans cette époque de transition, où ni leur méthodologie ni leur philosophie de la modélisation ne sont unifiées.

Ici encore, la forme de l'enquête et le point de vue de l'enquêteur sont plus intéressants que celui de l'enquêté. On peut penser que Maurice Fréchet n'a eu que très peu de contacts avec les économistes à la faculté des sciences ou à l'IHP[67], avant qu'il ne soit associé en 1946 aux tentatives du CNRS de fonder un centre d'économétrie. Mais, ses multiples voyages et son abondante correspondance l'ont bien souvent mis en rapport avec des économistes statisticiens. Enfin il était impossible que l'onde de choc de la révolution probabiliste en économie ne l'atteigne pas. D'ailleurs ses propres travaux sur la répartition des revenus [Fréchet 1917, 1925c, 1928d, 1939, 1945b] ont été une occasion de mêler calcul des probabilités et question économique. Cette onde de choc, Fréchet l'a bien perçue et accompagnée. Il a compris que le texte de Haavelmo [1944] intitulé *The Probability Approach in Econometrics* traçait la voie d'un changement radical en économie, bien que son adhésion à la Société internationale d'économétrie comme *fellow* ne date que de 1951. Sa fréquentation de la théorie des jeux au travers des travaux de Borel est ancienne et constitue certainement son point principal d'entrée sur le terrain des économistes. La conférence de Bruxelles [Fréchet 1952] sur *le calcul des probabilités dans les sciences sociales* ne fait que confirmer cette impression de contacts multiples mais superficiels avec les questions fondamentales de la méthodologie économétrique, que Fréchet a très peu pratiquée à l'exception de la statistique des revenus, pour laquelle sa contribution est importante[68].

L'enquête scientifique n'est plus, dans ce quatrième cas, un instrument d'intervention dans le champ scientifique de la statistique. Il n'y a d'ailleurs aucune « motion » présentée par Fréchet à l'IIS sur ce sujet de la

économique et de données statistiques n'est possible sans la médiation d'un modèle stochastique, constitué d'équations qui traduisent des relations structurelles de l'économie et de termes aléatoires qui représentent les erreurs de mesure et de spécification. Mise sous cette forme, la théorie devient un jeu d'hypothèses probabilistes testables. Le manifeste de cette école de la Cowles Commission est [Haavelmo 1944].

[67] Citons tout de même un débat avec B. Nogaro en 1939, dont on trouve trace dans le dossier F4.4 des Archives de l'Académie des sciences.

[68] Elle est analysée dans [Armatte 1995, chap. 10].

mathématique économique. L'enquête scientifique est devenue un simple instrument de connaissance des représentations et des pratiques collectives sur un sujet dans une population spécifique. Il n'est plus question de rétablir une quelconque vérité, mais d'établir une photographie d'une sorte d'opinion publique. D'ailleurs n'est-ce pas l'une des plus importantes manifestations de la statistique sociale que cette mesure des opinions publiques qui a jailli du croisement des travaux des probabilistes et des initiatives de Gallup et Stoetzel juste avant guerre ?

CONCLUSION

Les enquêtes de Maurice Fréchet offrent un support historiographique privilégié à l'historien des mathématiques appliquées. Par leur forme bâtarde, entre enquête statistique, correspondance scientifique, débat public et action politique concertée, elles offrent de multiples points de vue sur les controverses qui marquent la naissance des concepts et des outils de la statistique mathématique. Elles montrent comment celle-ci naît par une lente et douloureuse extraction d'une gangue nourricière complexe faite de méthodologies simples mais hétérogènes, mal assurées dans leurs fondements et que la mathématique permet quelquefois — mais pas toujours — d'unifier. Elles suggèrent une certaine complexité des rapports formels et des rapports sociaux entre mathématique pure et mathématiques appliquées, qui d'ailleurs disqualifie le terme d'application, trop univoque.

Elles montrent qu'entre vérité mathématique démontrée et opinion collective, il n'y a pas toujours l'opposition que l'on peut croire, mais parfois l'épaisseur d'un papier lu à un colloque et de quelques éléments de rhétorique. Elles révèlent que, derrière l'image d'un mathématicien au sens le plus fort de celui qui a démontré des théorèmes ou construit une théorie, il peut y avoir un homme soucieux d'histoire et de philosophie des sciences, qui s'intéresse de près à cette science imparfaite de la preuve quantitative qu'est la statistique.

Enfin, ces enquêtes nous rappellent encore aujourd'hui les enjeux d'un véritable débat public sur les usages des outils et modèles mathématiques : il y va tout simplement du contrôle social, collectif et démocratique, d'une expertise qui à tous les moments et dans tous les champs de la pratique sociale, détermine notre rapport au monde, présent et à venir. L'expertise statistique est en effet régulièrement convoquée pour construire un consensus social sur les risques de toutes sortes, qui accompagnent presque toutes les entreprises humaines, aujourd'hui plus que jamais au centre de nos préoccupations. Nous pouvons remercier Maurice Fréchet d'avoir mis en lumière la nécessité pour les mathématiciens et les experts statisticiens de dominer à la fois les propriétés syntaxiques et sémantiques de leurs outils, d'avoir rétabli un lien fort entre ces propriétés et les usages qu'elles autorisent, et d'avoir donné l'exemple d'une vigilance sans relâche vis-à-vis des utilisations sociales de l'outil mathématique :

« Il importe que les membres de la première catégorie [les mathématiciens] ne donnent pas en toute occasion à ceux de la seconde [les experts] l'impression de sécurité à laquelle conduit souvent l'emploi des mathématiques, sécurité qui existe, en effet, mais seulement du passage des hypothèses aux conclusions, mais qui cesse aux deux extrémités. Or très souvent, cette réserve n'est pas mise suffisamment en lumière » [Fréchet 1947c, p. 364].

Remerciements

L'auteur remercie vivement les deux rapporteurs qui l'ont empêché d'écrire ou de reproduire quelques erreurs présentes dans la première version.

BIBLIOGRAPHIE

Certainement non exhaustive, cette bibliographie de l'œuvre statistique de Fréchet a été faite à partir de la notice [Fréchet 1933a] et de son supplément [Fréchet 1951a], qui fournissent les références Nxxx. Celles-ci ont été complétées par celles du volume bibliographique publié par *Econometrica* en 1957. Nous avons par ailleurs dépouillé les principaux journaux statistiques européens de la période 1925–1955 : le *Journal de la Société Statistique de Paris*, le *Bulletin de l'IIS*, la *Revue de l'IIS*, *Econometrica* et *Metron*. Fréchet lui-même chiffre à 32 le nombre de ses publications en statistique et à 71 celui de ses publications en calcul des probabilités entre 1903 et 1951, sur un total de 294. Le fonds Fréchet de l'Académie des sciences (une trentaine de caisses) a fait l'objet d'un classement en 56 dossiers par L.C. Arboleda à la fin des années soixante-dix. Nous avons consulté les cartons de la série F4.x contenant des documents et manuscrits de probabilités, statistique et mathématique appliquée. Les archives de l'ISUP comprennent un certain nombre d'articles tirés à parts de Fréchet et surtout un carton (qui n'était pas identifié) comprenant la correspondance relative à l'affaire du coefficient de corrélation.

1. Œuvre statistique de Maurice Fréchet

FRÉCHET (Maurice) & HALBWACHS (Maurice)

[1924] (N76) *Le Calcul des probabilités à la portée de tous*, Paris : Dunod, 1924, 297 p.

FRÉCHET (Maurice) & ROULLET (H.)

[1928] (N119) *Nomographie. Pratique et construction des abaques*, Paris : A. Colin, 1928, 208 p. ; 2e éd. 1933 ; 3e éd. 1946 ; 4e éd. 1952.

FRÉCHET (Maurice) & ROMANN (R.)

[1930] *Représentation des lois empiriques par des formules approchées à l'usage des chimistes, des physiciens, des ingénieurs, et des statisticiens*, Paris : Eyrolles, 1930, 302 p.

FRÉCHET (Maurice) & FAN (K.)

[1946] *Introduction à la topologie combinatoire*, Paris : Vuibert, 1946, 88 p.

FRÉCHET (Maurice)

[1905] (N7BIS) *Leçons sur les fonctions de variables réelles et les développements en séries de polynômes, professées à l'École normale supérieure par É. Borel*, Paris : Gauthier-Villars, 1905.

[1906] (N22) Sur quelques points de calcul fonctionnel, thèse de doctorat publ., dans *Rendiconti del Circolo matematico di Palermo*, 22 (1906), p. 1–74.

[1917] Une nouvelle représentation analytique de la répartition des revenus, *Bulletin de l'Institut International de Statistique*, 22 (1917), p. 547 et sq.

[1920] (N66) Les mathématiques à l'Université de Strasbourg (Leçon d'ouverture du cours d'Analyse supérieure), *Revue du mois*, 1920, p. 337–362 ; extrait repris dans [Fréchet 1955b, p. 368–388] sous le titre : *Biographie du mathématicien alsacien Arbogast.*

[1921a] (N62) Remarques sur les probabilités continues, *Bulletin des sciences mathématiques*, 45 (1921), p. 87–88.

[1921b] (N63) Sur divers modes de convergence d'une suite de fonctions d'une variable, *Bulletin of the Calcutta Mathematical Society*, 2 (1921), p. 187–206.

[1921c] (N71) Sur les ensembles abstraits, *Annales scientifiques de l'École normale supérieure*, 38 (1921), Paris.

[1922] (N95a) *Sur la théorie des erreurs d'observation*, conférence non publiée de Berne.

[1923a] (N73) Une expression élémentaire approchée de la loi des grands nombres, *Revue générale des sciences*, 34 (1923), p. 211–212.

[1923b] (N77) Sur la distance de deux ensembles, *Comptes rendus des séances hebdomadaires de l'Académie des sciences*, 176 (1923), p. 1123–1124.

[1924a] (N73BIS) L'organisation scientifique (I. Les congrès nationaux), *Revue générale des sciences*, 35 (1924), p. 107–108.

[1924b] (N93) Sur l'ajustement des tables de mortalité par la méthode Tchebichef, *Compte rendu du Congrès des sociétés savantes*, Dijon 1924, p. 41–53.

[1924c] (N95) Sur la loi des erreurs d'observation, *Bulletin de la Société mathématique de Moscou*, 33 (1924), p. 5–8.

[1924d] (N123) Sur une formule générale pour le calcul des primes pures d'assurance sur la vie, *Compte rendu du Congrès international de mathématiques de Toronto*, 2 (1924), p. 857–865.

[1925a] (N96) Sur la loi des erreurs d'observation, *C. R. Acad. sci. Paris*, 181 (1925), p. 204–206.

[1925b] (N107) Sur la loi des erreurs d'observation, *Bull. Soc. math. Moscou*, 32 (1925), p. 704–710.

[1925c] (N110) Une nouvelle représentation analytique de la répartition des revenus, *BIIS*, 22-3 (1925), p. 547.

[1925d] (N111) Sur une formule générale pour le calcul des primes pures, *Compte rendu du Congrès des sociétés savantes*, Paris, 1925, p. 78–81.

[1925e] *Sur une désaxiomatisation de la science*, conférence faite à Berne et reproduite dans [Fréchet 1955b, p. 1–10].

[1926] *Sur l'ajustement des tables de mortalité à l'aide de la méthode de Tchebichef*, en collaboration avec Perrenoud et Mahrer, Paris : Dulac frères, 1926.

[1927a] Rietz (H.-L.), Mathematical Statistics (1927), Compte rendu 22, *Bulletin des sciences*, 1927.

[1927b] (N115) Sur la loi de probabilité de l'écart maximum, *Annales de la Société polonaise de mathématiques*, 5 (1927), p. 93–116.

[1928a] (N114) Sur l'hypothèse de l'additivité des erreurs partielles, *Bulletin des sciences mathématiques*, 2e série, 52 (1928), p. 203–216.

[1928b] (N124) *Les espaces abstraits et leur théorie considérée comme introduction à l'analyse générale*, xi-296 p., Paris : Gauthier-Villars, 1928 ; 2e tirage 1943.

[1928c] (N125) Sur l'existence d'un indice de désirabilité des biens indirects, *C. R. Acad. sci. Paris*, 187 (1928), p. 589–591.

[1928d] (N125a) *Comparaison des divers indices proposés pour mesurer l'inégalité de la répartition des revenus*, conférence non publiée tenue à Bruxelles.

[1930a] Le calcul des probabilités, conférence à l'École supérieure des Postes et Télégraphes, *Congrès de l'Association française pour l'avancement des sciences*, repris dans [Fréchet 1955b, p. 127–152].

[1930b] Nouvelles expressions de la distance de deux variables aléatoires et de la distance de deux fonctions mesurables, *Ann. Soc. pol. math.*, (1930), p. 45–49.

[1933a] *Notice sur les travaux scientifiques de M. Fréchet*, Paris : Hermann, 1933, 104 p.

[1933b] Sur le coefficient dit de corrélation, *C. R. Acad. sci. Paris*, 197 (1933), p. 268–269.

[1934] Sur le coefficient dit de corrélation et sur la corrélation en général, *RIIS*, 1933-4 (janvier 1934), p. 16–23.

[1934b] L'arithmétique de l'infini, dans *Exposés d'analyse générale*, Paris : Hermann, 1934, 38 p.

[1935a] Sur l'usage du soi-disant coefficient de corrélation, *RIIS*, 1934-4 (janvier 1935), p. 3–26.

[1935b] Les indices statistiques de dépendance fonctionnelle, *Barometro economico*, VII-4 (1935), p. 231–232.

[1935c] Sur l'usage du coefficient de corrélation, *Comptes rendus de la 22e session de l'IIS à Londres*, BIIS, 28-2 (1935), p. 25–52.

[1936a] Sur le coefficient de linéarité dit de corrélation, *RIIS*, 1935-4 (janvier 1936), p. 365–379.

[1936b] (N272) Recherches théoriques modernes sur le calcul des probabilités, 1er livre : Généralités sur les probabilités. Éléments aléatoires avec notes de P. Lévy, dans Borel (Émile), éd., *Traité de calcul des probabilités et de ses applications*, t. I, fasc. III, livre I, 1e éd., xvi-308 p., Paris : Gauthier-Villars, 1936 ; 2e éd. augmentée, *ibid.*, xvi-355 p., 1948.

[1937a] Rapport de la commission d'étude de l'usage du coefficient de corrélation, *Comptes rendus de la 23e session de l'IIS*, Athènes, BIIS, 29-3 (1937), p. 318–322.
[1937b] Mélanges mathématiques, *Comptes rendus du Congrès international des mathématiciens*, Oslo : A.W. Broggers Boktrykkeri, 1937, p. 269–283.
[1938a] (N284) Recherches théoriques modernes sur le calcul des probabilités, 2e livre : La méthode des fonctions arbitraires. Les évènements en chaîne dans le cas d'un nombre fini de cas possibles, dans Borel (Émile), éd., *Traité de calcul des probabilités et de ses applications*, t. I, fasc. III, livre II, x-315p, 1938 ; 2e éd. augmentée d'un supplément et d'une note de M. Paul Lévy, 374 p., 1952.
[1938b] Exposé et discussion de quelques recherches récentes sur les fondements du calcul des probabilités, dans *Actes du Colloque de Genève sur la théorie des probabilités*, 10–15 oct. 1937, présidé par M. Fréchet, Paris : Hermann, 1938.
[1938c] Compte rendu de la séance du 17 avril 1934. L'usage du coefficient de corrélation, *BIIS*, 28-1 (1938), p. 82–84.
[1939] Sur les formules de répartition des revenus, *RIIS*, 1939-1, p. 32–38.
[1940a] Comparaison des différentes mesures de dispersion, *RIIS*, 1940-1.
[1940b] Sur une limitation très générale de la dispersion de la médiane, *JSSP*, 81 (1940), p. 67–79.
[1940c] *Les probabilités associées à un système d'événements compatibles et dépendants*, 1er fasc. : *Événements en nombre fini*, Paris : Hermann, 1940, viii-80 p.
[1941a] *L'analyse générale et la question des fondements : rapports aux entretiens de Zurich sur les fondements et la méthode des sciences mathématiques de décembre 1938*, Zurich : S.A. Leeman frères, 1941, repris sous le titre *Les origines des notions mathématiques* dans [Fréchet 1955b, p. 11–51].
[1941b] Sur la loi de répartition de certaines grandeurs géographiques, *JSSP*, 1941, p. 114–122.
[1942] *Leçons de statistique mathématique : 1er cahier : Introduction. Exposé préliminaire du calcul des probabilités*, Cours de la Sorbonne, Centre de documentation universitaire, 1942, 119 p.
[1943a] *Aide-mémoire de statistique mathématique*, Paris : Eyrolles, 1943.
[1943b] *Les probabilités associées à un système d'événements compatibles et dépendants*, 2e partie : *Cas particuliers. Applications*, Paris : Hermann, 1943.
[1943c] Sur l'expression simple et approchée de la loi de probabilité des erreurs d'observation, *JSSP*, 1943, p. 52–70.
[1943d] *Les probabilités associées à un système d'événements compatibles et dépendants*, 2e fasc. : *Cas particuliers et applications*, Paris : Hermann, 1943, 131 p.
[1945a] *Cours de mathématiques générales*, 2e éd. revue et augmentée, t. I : *Calcul différentiel*, Paris : CDU, 1945.
[1945b] (N239) Nouveaux essais d'explication de la répartition des revenus, *RIIS*, 1945, p. 294–297.
[1946a] (N248) Fondements des méthodes statistiques d'estimation, *Portugaliae Mathematica*, 5 (1946), p. 137–141.
[1946b] Dégager les possibilités et les limites de l'application des sciences mathématiques à l'étude des phénomènes économiques et sociaux, Rapport pour la session de Washington, *RIIS*, 1946, p. 16–51.

[1946d] Les définitions courantes de la probabilité, *Revue philosophique*, 1946, p. 129–169 ; repris dans [Fréchet 1955b, p. 157–204].

[1946e] Nouvelles définitions de la valeur moyenne et des valeurs équiprobables d'un nombre aléatoire, *Annales de l'université de Lyon*, 3e sér., 1946, section A, p. 5–26.

[1946f] Une expression approchée des lois de mortalité valable pour la vie entière, *JSSP*, 1946, p. 242.

[1946g] (N262) A general method of constructing correlation indices, *Proceedings of the Matematical and Physical Society of Egypt*, 3 (1946), p. 13–20 ; *Additional notes*, 1948, p. 73–74.

[1947a] Anciens et nouveaux indices de corrélation. Leur application au calcul des retards économiques, *Econometrica*, 15 (janvier 1947), p. 1–30 ; *Errata*, octobre 1947, p. 374–375.

[1947b] (N243) Rapport sur une enquête concernant les possibilités et limites de l'emploi des mathématiques dans les sciences sociales, *BIIS*, 14e année (1947), p. 1–15.

[1947c] (N243Bis) Rapport sur une enquête internationale relative à l'estimation statistique des paramètres, *Proceedings of the IIS Conference*, Washington, Calcutta : EKA Press, 6–18 sept. 1947, vol. III-A (1947), p. 363–420.

[1947d] (N244) Sur les expressions analytiques de la mortalité valables pour la vie entière, *JSSP*, 1947, p. 261–284.

[1947e] (N250) Definition of the probable deviation, *Annals of Mathematical Statistics*, 18 (1947), p. 288–290.

[1947f] (N250Bis) The general relation between the mean and the mode for a discontinuous variable, *Ann. Math. Stat.*, 18 (1947), p. 290–293.

[1947g] (N251) Le coefficient de connexion statistique de Gini-Salvemini, *Mathematica*, 23 (1947–1948), p. 46–51.

[1947h] (N252) Les espaces abstraits et leur utilité en statistique théorique et même en statistique appliquée, *JSSP*, 1947, p. 410–421.

[1947i] (N266) Contributions françaises récentes au calcul des probabilités et à la statistique mathématique, *Intermédiaire des recherches mathématiques* et *Compte rendu du Congrès de l'Association française pour l'avancement des sciences pour 1945*, Paris, 1947.

[1947j] Buffon, philosophe des mathématiques, *Bulletin de l'Institut d'Egypte*, 28 (1945–1946), p. 185–202.

[1947k] Projet d'introduction du calcul des probabilités dans l'enseignement moyen, suivi d'une discussion par dix spécialistes, *JSSP*, 1947, p. 286–297 ; repris dans [Fréchet 1955b, p. 291–300].

[1948a] (N253) Les éléments aléatoires de nature quelconque dans un espace distancié, *Annales de l'IHP*, 14 (1948), p. 215–310.

[1948b] (N254) Sur une nouvelle définition des positions typiques d'un élément aléatoire abstrait, *C. R. Acad. sci. Paris*, 226 (1948), p. 1419–1420.

[1948c] (N255) Allocution présidentielle d'entrée en fonction, *JSSP*, 89e année (1948), Procès verbal de séance du 17 mars 1948, p. 85–92.

[1948c] (N258) Les valeurs typiques d'ordre nul ou infini d'un nombre aléatoire, *RIIS*, 16 (1948), p. 1–22.

[1948d] (N256) Sur l'estimation statistique, *Ann. Soc. polonaise math.*, 21(1948), p. 207–213.

[1948e] (N257) Positions typiques d'un élément aléatoire de nature quelconque, *Ann. sci. École normale sup.*, 61(1948), p. 211–237.

[1948f] (N260) Les mathématiques en économie politique, *Revue de sciences économiques*, 23 (1948), p. 147–160.

[1948g] (N261) On two new chapters on the calculus of probability, *Mathematical Magazine*, 22 (1948–1949), p. 1–12.

[1948h] (N263) Mathematics and the social sciences, *Math. Mag.*, 21 (1948), p. 199–212.

[1948i] *Cours de mathématiques générales*, 2e éd. revue et augmentée, t. II : *Géométrie*.

[1948j] La possibilité et les limitations de l'application des mathématiques aux sciences sociales (résumé), *Econometrica*, 16 (1948), p. 51–53.

[1948k] L'estimation statistique des paramètres (résumé), *Econometrica*, 16 (1948), p. 60–62.

[1949a] (N264) Statistical self renewing aggregates, Simaika (J.B.), éd., *A Course of Lectures*, Fouad I Univiversity Press, 1949, 125 p.

[1949b] (N265) Les mathématiques dans les sciences humaines, *Conférences du Palais de la découverte*, 27 p. ; repris dans [Fréchet 1955b, p. 89–117].

[1949c] (N267) Allocution du président sortant, *JSSP*, 90 (1949), procès-verbal de séance du 16 mars 1949, p. 161–167.

[1949d] La statistique, ses buts, ses applications, son enseignement, conférence d'ouverture d'un cycle de cours sur la statistique et ses applications, organisé par la Faculté des sciences de Madrid, *Archives de l'Académie des sciences Paris*, carton F4.4 ; repris dans [Fréchet 1955b, p. 273–290].

[1949e] Les valeurs typiques d'ordre nul ou infini et leur généralisation, dans *Le calcul des probabilités et ses applications*, Colloque de Lyon, Paris, Éditions du CNRS, p. 47–51.

[1950a] (N269) Réhabilitation de la notion statistique de l'homme moyen, *Conférence du Palais de la découverte*, 15 octobre 1949, 24 p. ; réédité dans [Fréchet 1955, p. 317–341].

[1950b] (N270-271-286-287) *Leçons de statistique mathématique* : 2e cahier : recueil d'exercices de calcul des probabilités et de statistique mathématique, 1950, 52 p. ; 3e cahier : solutions des problèmes du second cahier ; 4e cahier : les ensembles statistiques renouvelés (cas d'une durée maximum d'usure), 1950, 168 p. ; 5e cahier : les ensembles statistiques renouvelés (cas général) ; 6e cahier : en marge des traités (sur les valeurs typiques, la dispersion, les indices de corrélation et l'estimation statistique), 240 p. ; 7e cahier : en marge des traités (Sur les lois de probabilités et sur les distributions statistiques), 238 p. Polycopiés du Centre de documentation universitaire, Paris : Constans.

[1950c] (N275) La moyenne réduite converge 'légalement' mais non 'en probabilité', *Ann. Univ. Lyon*, 3e sér., 1950, section A, p. 33–36.

[1950d] (N277) Sur diverses définitions de la moyenne d'un élément aléatoire de nature quelconque, *Proceedings of the Second Symposium on Mathematical Statistics and Probabilities*, Berkeley : University of California Press, 1950.

[1950e] (N278) Sur un essai illégitime de sauver le coefficient classique dit de corrélation, *RIIS*, 18 (1950).

[1951a] Supplément pour la période 1934–1951 à la notice pour les travaux scientifiques de M. Maurice Fréchet, Paris : Hermann, 13 p.

[1951b] (N274) Généralisation de la loi de probabilité de Laplace, *Annales de l'IHP*, 12 (1951), p. 1–29.

[1951c] (N276) Une propriété générale des valeurs typiques d'un nombre aléatoire, *Publication de l'ISUP*, 1(1951), p. 1–48.

[1951d] (N279) Sur une application de la statistique mathématique à la biologie, *Biometrika*, 7 (1951), p. 180–184.

[1951e] (N280) Rapport général sur les travaux de la section de calcul des probabilités, *Actes du congrès international de philosophie des sciences (Paris 1949)*, Paris : Hermann, 1951 ; repris dans [Fréchet 1955b, p. 205–241], avec note de De Finetti.

[1951f] (N282) Elementos aleatorios de naturaleza cualquiera, *Trabajos de estatistica*, 2 (1951), Madrid, p. 157–187.

[1951g] (N289) Sur les fonctions du bien-être, *Revue d'économie politique*, 16 (1951), p. 702–707 ; repris dans [Fréchet 1955b, p. 118–126].

[1951h] (N290) Sur les tableaux de corrélation dont les marges sont données, *Ann. Univ. Lyon*, 3e sér., 14 (1951), section A, p. 53–77.

[1952] (N285) Le calcul des probabilités dans les sciences sociales, conférence à Bruxelles en 1951, publiée dans *Théorie des probabilités ; Exposé sur ses fondements et ses applications*, Paris : Gauthier-Villars, 1952, p. 149–168 ; repris dans [Fréchet 1955b, p. 250–272].

[1953a] *Cours de mathématiques générales*, 2e éd. revue et augmentée, t. III : Intégrales simples, multiples et curvilignes, Paris 1953.

[1953b] (N294) Émile Borel, Initiator of the theory of psychological games and its applications ; commentary of the Borel Notes, *Econometrica*, 21 (1953), p. 95–96 et 118–124.

[1953c] *Pages choisies d'analyse générale*, Paris : Gauthier-Villars, 1953, 213 p.

[1954–1955] La statistique et le calcul des probabilités dans l'enseignement technique, *Technique, art, science*, 8 (1954), cahier 7, p. 35–37 ; 10 (1955), cahier 2, p. 23–31 et cahier 3, p. 19–27.

[1954a] Buffon comme mathématicien, statisticien, et philosophe des mathématiques, dans *Corpus général des philosophes français*, Paris : PUF, 1954 ; repris dans [Fréchet 1955b, p. 344–367].

[1954b] Calcul des probabilités et statistique, dans *Clarté ; Encyclopédie du temps présent*, vol. 16, *Les lois de la pensée*, 1954, 17 p.

[1954c] Interdépendance du centre et du rayon empiriques de variation de n observations indépendantes, *Studies in Mathematics and Mechanics*, 1954, p. 285–294.

[1954d] Sur les statistiques municipales françaises, *JSSP*, 95 (1954), p. 73–75.

[1954e] (N309) Un problème psychologique sur les probabilités irrationnelles, *Journal de psychologie*, 1954, p. 431–438.

[1955a] La statistique dans les sciences sociales, dans [Fréchet 1955b, p. 301–316].

[1955b] *Les mathématiques et le concret*, Paris : PUF, 1955, 438 p.

[1955c] Remarques sur l'article de M. Finetti, *Journal de psychologie*, 1955, p. 260–261.

[1955d] (N310) Sur l'importance en économétrie de la distinction entre les probabilités rationnelles et irrationnelles, *Econometrica*, 23 (1955), p. 303–306.

[1956-1957] Les inégalités de Minkowski dégénérées et leurs applications au calcul des probabilités, *Annales de l'IHP*, 15 (1956–1957).

[1956a] Émile Borel (1871–1956), *Revue philosophique*, 1956, p. 158–160.

[1956b] Sur les tableaux de corrélation dont les marges sont données, *C. R. Acad. Sci. Paris*, 242 (1956), p. 2426–2428.

[1957a] Les vrais et faux dangers de la méthode mathématique en économie politique et les insuffisances de la méthode non mathématique, *Économie appliquée*, 1957.

[1957b] Sur la distance de deux lois de probabilité, *C. R. Acad. Sci. Paris*, 1244 (1957), p. 689–692.

[1957c] Les tableaux de corrélation dont les marges et les bornes sont données, *Ann. Univ. Lyon*, 3e sér., 20 (1957), section A, p. 13–31.

[1959a] *Cours de mathématiques générales*, 2e éd. revue et augmentée, t. IV : Équations différentielles, équations linéaires aux dérivées partielles. Énoncés d'examen, Paris, 1959.

[1959b] Le rôle d'Émile Borel dans la théorie des jeux, *Revue d'économie politique*, 1959, p. 139–167.

[1965] *La vie et l'œuvre d'Émile Borel*, Genève : Kundig, 1965, 97 p.

[1967] *Émile Borel philosophe*, Paris : Gauthier-Villars, 1967, 407 p.

2. Autres références

Arboleda (Luis Carlos)

[1977] *Sur la vie et l'œuvre de M. Fréchet*, Mémoire de DEA, EHESS, Paris, 1977.

[1979] Les débuts de l'école topologique soviétique : Note sur les lettres de Paul S. Alexandroff et Paul S. Urysohn à Maurice Fréchet, *Archive for History of Exact Sciences*, 20 (1979), p. 73–89.

[1980] *Contribution à l'étude des premières recherches topologiques, d'après la correspondance et les publications de Maurice Fréchet (1904–1928)*, Thèse EHESS, Paris, 1980.

Armatte (Michel)

[1988] La construction des notions d'estimation et de vraisemblance chez Ronald A. Fisher, *Journal de la Société statistique de Paris*, 129 (1988), no 1–2, dans Mairesse (D.J.), dir., *Estimation et sondages, cinq contributions à l'histoire de la statistique*, Paris : Economica, 1988, p. 69–96.

[1992] Conjonctions, conjoncture et conjecture. Les baromètres économiques, *Histoire et mesure*, VII, 1–2 (1992), p. 99–149.

[1995] *Histoire du modèle linéaire. Formes et usages en statistique et en économétrie jusqu'en 1945*, Thèse soutenue à l'EHESS, sous la dir. de J. Mairesse, Paris.

[2001a] Developments in statistical reasoning and their links with mathematics, dans Bottazzini (Umberto) et Dahan (Amy), éd., *Changing Images in Mathematics, From the French Revolution to the new Millenium*, Actes du colloque international d'histoire des mathématiques, Luminy, sept. 1997, London : Harwood Academic Publisher, p. 137–162.

[2001b] Le statut changeant de la corrélation en économétrie (1914-1944), *Revue économique*, 52-3 (2001), p. 617–631.

Bravais (Auguste)

[1846] Analyse mathématique sur les probabilités des erreurs de situation d'un point, présenté à l'Académie en 1838, rapporté par Poisson et Savary, *C. R. Acad. Sci. Paris*, 7 (1846), p. 77–78, publié dans *Mémoires présentés par divers savants*, 9.

Bru (Bernard)

[1995] Problème de l'efficacité du tir à l'École de Metz, *Cahiers du CAMS*, Série 'Histoire du calcul des probabilités', no 22, Paris, 1995.

Bungener (Martine) & Joël (Marie-Ève)

[1989] L'essor de l'économétrie au CNRS, *Cahiers pour l'histoire du CNRS*, 1989-4, p. 45–78.

Bunle (Henri)

[1911] Relations entre les variations des indices économiques et le mouvement des mariages, *JSSP*, 1911, p. 80–91.

Charle (Christophe) & Telkes (Eva)

[1989] Fréchet (René, Maurice), dans *Les professeurs de la faculté des sciences*, Paris 1989, p. 334–340.

Crépel (Pierre)

[1988] Condorcet et l'estimation statistique, dans Mairesse (D.J.), dir., *Estimation et sondages, cinq contributions à l'histoire de la statistique*, Paris : Economica, 1988, p. 47-68.

[1994] Le calcul des probabilités : de l'arithmétique sociale à l'art militaire, dans Belhoste (Bruno), Dahan (Amy) et Picon (Antoine), éd., *La formation polytechnicienne*, Paris : Dunod, 1994, p. 197–216.

Darmois (Georges)

[1928] *Statistique mathématique*, Paris : Librairie Douin, 1928.

[1929] Analyse et comparaison de séries statistiques qui se développent dans le temps, *Metron*, VIII, 1–2, mai 1929, p. 211–250.

Davis (Philip J.) & Hersh (Reuben)

[1981] *The Mathematical Experience*, Boston : Birkhäuser, 1981.

Doss (Shafik)

[1948] Sur la moyenne d'un élément aléatoire dans un espace distancié, *Bulletin sci. math.*, 1948.

Dugué (Daniel)

[1974] Maurice Fréchet (1878–1973), Ancien président de la société de statistique de Paris, *Journal de la Société de statistique de Paris*, 1974, p. 2.

[1974] Maurice Fréchet (1878–1973), *International Statistical Review*, 42 (1974), p. 113–114.

Feron (R.)

[1947] Mérites comparés des différents indices de corrélation, *JSSP*, 1947, p. 328 *et sq.*

Fisher (Ronald)

[1925] *Statistical Methods for Research Workers*, 14e éd., Edinburgh : Oliver and Boyd, 1925 ; trad. fr. par J. Bertrand de la 10e éd., Paris, 1947.

Furfey (P.H.) & Daly (J.F.)

[1935] Product-moment correlation as a research technique in education, *Journal of educational psychology*, 1935, p. 206–211.

Galton (Francis)

[1889] *Natural Inheritance*, London : Macmillan, 1889.

Gauss (Charles)

[1926] Theoria Combinationis..., dans *Méthode des moindres carrés. Mémoire sur la combinaison des observations*, trad. J. Bertrand, Paris : Mallet-Bachelier, 1955 ; trad. anglaise *Theory of the Combination of Observations Least Subject to Errors*, trad. by G.W. Stewart, Philadelphia : SIAM, 1995.

Gini (Corado)

[1914–1915] Di una misura della dissomiglianza tra due gruppi di quantita e delle suoi applicazioni allo studio delle relazioni statistiche, *Actes de l'Institut royal vénitien des sciences, lettres et arts*, t. 74, Venise, p. 185–213.

[1914–1915] Indici di omofila e di rassomiglianza e loro relazioni col coefficiente di correlazione e con gli indici di attrazione, *Actes de l'Institut royal vénitien des sciences, lettres et arts*, t. 74, Venise.

[1914–1915] Nuovi contributi alla teoria delle relazioni statistiche, *Actes de l'Institut royal vénitien des sciences, lettres et arts*, t. 75, Venise.

[1915–1916] Sul criterio di concordanza tra due caratteri, *Actes de l'Institut royal vénitien des sciences, lettres et arts*, t. 75, Venise.

[1915–1916] Indici di concordanza, *Actes de l'Institut royal vénitien des sciences, lettres et arts*, t. 76, Venise.

[1926] Contributions of Italy to modern statistical methods, *JRSS*, vol. 89-4, July 1926.

[1936] Sur le coefficient de corrélation, *RIIS*, 1936-3, p. 355–366.

Gispert (Hélène)

[1980] La correspondance de Fréchet (1907–1926) et son apport à la théorie de la dimension (avec trois lettres de Brouwer à Baire), *Cahiers du séminaire d'histoire des mathématiques*, 1 (1980), p. 69–120.

Haavelmo (Trygve)

[1944] The Probability Approach in Econometrics, *Supplément à Econometrica*, 12 (1944), 115 p.

Halbwachs (Maurice)

[1913] *La théorie de l'homme moyen. Essai sur Quételet et la statistique morale*, Paris : Alcan, 1913.

Hertz (Sébastien)

[1997] *Emil Julius Gumbel (1891–1966) et la statistique des extrêmes*, Thèse de doctorat, dir. P. Crépel, Université de Lyon I, 1997.

Jordan (Charles)

[1938] Critique de la corrélation du point de vue des probabilités, *Actes du colloque de calcul des probabilités*, Paris : Hermann, 1938, p. 15–33.

Laplace (Pierre-Simon)

[Œuvres] *Œuvres complètes de Laplace*, 14 vol., Paris : Gauthier-Villars, 1878–1912.

[1820] Théorie analytique des probabilités, 3e éd. ; *Œuvres* VII + suppl. dans *Œuvres* VIII, p. 98–120.

Machali (M.)

[1931] *Prévision du rendement du blé d'après les éléments météorologiques*, Thèse ISUP, Paris : Librairie Marcel Rivière, 1931, 92 p.

Martin (Olivier)

[1999] Raison statistique et raison sociologique chez Maurice Halbwachs, *Revue d'histoire des sciences humaines*, 1 (1999), p. 69–101.

Morgan (Mary)

[1990] *The History of Econometric Ideas*, Cambridge : Cambridge University Press, 1990.

[1995] Searching for Casual Relations in Economic Statistics : Reflections from History, *LSE*, Discussion Paper.

Owen (D.B.)

[1974] *On the History of Statistics and Probability*, Symposium Dallas 1974, New York : M. Dekker Ed., 1974.

Pearson (Karl)

[1912] *La grammaire de la science*, 1re éd. française 1892, trad. par L. March de la 3e éd. (1911), Paris : Alcan, 1912.

[1920] Notes on the history of correlation, *Biometrika*, 13 (1920), 25–45 ; repris dans *Studies in the History of Statistics and Probability*, éd. Kendall & Pearson, 1970, p. 185–206.

PIETRA (Gaetano)

[1925] The theory of statistical relations, with special reference to cyclical series, *Metron*, 4 (1925).

[1935] Note di statistica methodologica : A proposito della misura della variabilità e della concentrazione dei caratteri, *Nuovi problemi di politica, storia ed economia*, Anno VI, fasc. 1–6, Ferrara, 1935.

[1936] De la concordance et du coefficient de corrélation, *RIIS*, 1936-4, p. 500–514.

RIETZ (Henry-Lewis), ed.

[1924] *Handbook of Mathematical Statistics*, Cambridge : Riverside Press, Houghton Mifflin Cy, 1924.

RIETZ (Henry-Lewis)

[1927] *Mathematical Statistics*, Chicago : Open Court Publishing Cy, 1927, 81 p. ; 5e éd. 1947.

RISSER (René) & TRAYNARD (C.-E.)

[1933] Les principes de la statistique mathématique, dans Borel (Émile), éd., *Traité du calcul des probabilités et de ses applications*, t. 1 ; fasc. IV, Paris : Gauthier-Villars, 1933.

RISSER (René)

[1941] Note bibliographique sur Fréchet : Méthodes des fonctions arbitraires. Théorie des éléments en chaîne, *JSSP*, 1941, p. 151–160.

STIGLER (Stephen M.)

[1986] *The History of Statistics. The Measurement of Uncertainty before 1900*, Cambridge : Harvard University Press, 1986.

[1999] *Statistics on the Table. The History of Statistical Concepts and Methods*, Cambridge : Harvard University Press, 1999.

TAYLOR (Angus E.)

[DSB] Fréchet, René Maurice, *Dictionary of Scientific Biography*, vol. 17, Supplement II, 1970–1990, p. 309–311.

[1982] A study of Maurice Fréchet, I : His early work on point set theory and the theory of functionals, *Archive for History of Exact Sciences*, 27 (1982), p. 233–295.

[1985] A study of Maurice Fréchet, II : Mainly about his work on general topology, *Archive for History of Exact Sciences*, 34 (1985), p. 279–380.

[1987] A study of Maurice Fréchet, III : Fréchet as analyst, *Archive for History of Exact Sciences*, 37 (1987), p. 25–76.

YULE (G. Udny)

[1897] On the theory of correlation, *Journal of the Royal Statistical Society*, 60 (1897), p. 812–854.

[1909] Les applications de la méthode de corrélation aux statistiques sociales et économiques, *BIIS*, VIII-1 (juillet 1909), p. 265–277.

ZAHN (F.)

[1935] Au sujet de la méthode de travail de notre Institut, *RIIS*, 3 (1935), p. 2–8.