JOURNAL DE LA SOCIÉTÉ STATISTIQUE DE PARIS

Vie de la société. Soirée de la statistique du 6 décembre. Séance solennelle du 30 janvier 1996 : prix du statisticien de langue française

Journal de la société statistique de Paris, tome 137, n° 3 (1996), p. 3-49

<http://www.numdam.org/item?id=JSFS_1996__137_3_3_0>

I

VIE DE LA SOCIÉTÉ

SOIRÉE DE LA STATISTIQUE

6 décembre 1995

La Société a pensé qu'il serait intéressant d'organiser une réunion un peu longue en fin d'après-midi consacrée à l'enseignement de la statistique.

Malheureusement cette réunion a eu lieu au moment des grèves de décembre dernier et faute de moyens de transport, la participation a été très faible ! D'où l'avantage du **Journal** *qui permet de prendre connaissance des exposés :*
celui, introductif de Félix ROSENFELD
et celui, technique, de Marie BERRONDO-AGRELL.

Le troisième exposé, celui du Professeur TENENHAUS de HEC-ISA ne figure pas, car il ne se prête pas à publication. En effet, il s'est agi d'une "étude de cas" sur ordinateur, démonstration passionnante de la manière dont les logiciels actuels permettent d'enseigner en "temps réel" les méthodes d'analyse d'enquêtes.

Journal de la Société de Statistique de Paris, tome 137, n° 3, 1996.

RAPPEL HISTORIQUE DE L'ENSEIGNEMENT DE LA STATISTIQUE

Félix ROSENFELD
Président d'honneur de la
Société de Statistique de France

I. Introduction

La statistique étudie, dénombre, mesure les caractéristiques de populations de toutes sortes – vivantes, matérielles (ensembles d'objets), immatérielles (suites d'événements, de mesures) –, les compare, recherche les relations éventuelles entre elles, examine leur évolution.

C'est dire que la statistique est plongée dans les faits, dans le monde, qu'elle doit en suivre l'évolution et adapter constamment ses méthodes à cette évolution. Enseigner la statistique, c'est enseigner une science en perpétuelle évolution, continuellement modifiée par la recherche, continuellement sollicitée par de nouvelles applications.

En d'autres termes, l'enseignement de la statistique ne peut être dissocié de la recherche qui la concerne, ni de ses applications. C'est ce qui apparaîtra clairement dans ce rappel historique.

Nous ne ferons pas remonter ce rappel au Déluge, bien que Noé se soit signalé par la réalisation d'un excellent échantillon représentatif de tout ce qui vivait à l'époque sur cette Terre, nous ne remonterons pas non plus aux grands prédécesseurs déjà lointains, Pascal, Descartes, Bernoulli, Gauss ou Quételet, notre propos sera suffisamment nourri en le limitant à ce qui s'est passé au cours de notre siècle.

Période 1900-1945. Les fondements

C'est incontestablement en **Angleterre** que du début du siècle jusqu'à la Seconde Guerre mondiale la recherche et l'enseignement de la statistique ont été le plus avancés. Les universités de Cambridge, d'Oxford et de Londres se sont partagé les professeurs et les enseignements les plus marquants.

Au Laboratoire de Biométrie de l'University College de Londres, *Karl Pearson*, prenant la suite de *Francis Galton*, développe, dès 1896-1900, la théorie de l'hérédité et celle de la corrélation ; il formule ensuite l'utilisation des tests de signification et invente le coefficient Chi2. Il publie dans deux gros volumes de *Biometrika* ses *Tables for Statisticians and Biometricians*, utilisées par toutes les générations successives de statisticiens dans le monde entier.

G. Udny Yule, élève de K. Pearson, a également été professeur à l'University College de Londres. Son ouvrage, *An Introduction to the Theory of Statistics*, publié dès 1911, a servi de base à tous les étudiants en statistique jusqu'à la veille de la dernière guerre mondiale. En 1937, l'ouvrage était réédité avec le concours de *M.G. Kendall*, lequel publia plus tard, en 1943, un cours magistral sous le titre de *The Advanced Theory of Statistics* ; ce cours fait encore autorité.

Tout statisticien connaît les *Tables de Student*. Celui-ci s'appelait en réalité *William S. Gosset*. Industriel, il travaillait dans les Brasseries Guinness. Très porté vers la statistique, il passa un an auprès de Karl Pearson et publia en 1908, dans *Biometrica*, des articles sur les erreurs d'estimation de la moyenne et du coefficient de corrélation. Ses tables sur la signification des erreurs d'observation ont été publiées dans la revue *METRON*, en 1925. Il a été un précurseur dans l'application de la statistique à l'industrie.

Ronald A. Fisher a été une autre figure dominante de la statistique dans la première moitié du siècle ; sa recherche et son enseignement ont fait la notoriété de la station expérimentale agricole de *Rothamsted*. Son ouvrage, *The Design of Experiments*, dont la première édition date de 1925, a eu un impact considérable sur le recherche et l'enseignement de la statistique pendant plusieurs décennies, dans le monde entier. On lui doit les méthodes de base concernant les plans d'expériences, les carrés latins, mais aussi les premières approches de l'analyse de la variance. *Frank Yates*, son disciple, a pris sa suite à Rothamsted et ne s'est éteint que récemment.

En 1938, Fisher avait été invité à donner des conférences à Paris ; M. Darmois, dont j'étais assistant, me chargea de lui montrer quelques sites intéressants. Nous devions déjeuner à Saint-Cloud. Avant de traverser le pont, il me fit remarquer que nous étions sur la bonne voie, puisqu'un poteau portait l'indication "St-Cloud $0,^{k}2$". Il lisait en effet St-Cloud OK, et même St-Cloud OK au carré.

Egon S. Pearson, fils de Karl Pearson, est surtout connu comme un des promoteurs de l'application de la statistique au contrôle des fabrications, en raison de son livre *The Application of Statistical Methods to Industrial Standardization and Quality Control* (1935). En réalité, il a joué également un rôle important dans les recherches sur les tests de vérification des hypothèses, en rapport avec J. Neyman, et dans l'utilisation des nombres aléatoires, ainsi que dans le domaine nouveau de la robustesse des estimations.

Bien d'autres statisticiens britanniques ont marqué cette première moitié de siècle : *A.L. Bowley, M.S. Bartlett, F.Y. Edgeworth, H. Jeffreys, W.F. Sheppard, L.H.C. Tippett*, parmi les plus connus.

N'oublions évidemment pas aussi que *J.M. Keynes*, qui s'est rendu célèbre parmi les économistes par sa *Théorie générale sur l'emploi, l'intérêt et la monnaie*, publiée en 1936, s'est montré également un excellent probabiliste, en produisant en 1921 un *Traité sur les probabilités*. En tant qu'économiste, son équation générale stipulant que :

$$\text{production} + \text{importations} = \text{consommation} + \text{investissement} + \text{exportations}$$

est un élément fondamental de la comptabilité nationale, telle qu'elle a été développée plus tard.

Saisissons l'occasion pour signaler que les années trente ont connu d'autres apports importants à la méthodologie de la comptabilité nationale, avec notamment le tableau des relations interindustrielles élaboré aux Etats-Unis par *W. Leontief* (tableaux input-output) et, en Angleterre, les travaux du professeur australien *Colin Clark* (National Income and Outlay, 1937).

Revenons aux **Etats-Unis** pour constater que peu de statisticiens éminents se sont signalés dans la première partie du siècle, alors même que l'enseignement de la discipline était dispensé dans de nombreuses universités. Il est vrai que dans ce pays la statistique était plutôt orientée vers l'économie, les affaires et l'industrie. *Irving Fisher* avait publié en 1922 un ouvrage sur les nombres-indices, *The Making of Index Numbers*, qui a longtemps fait référence ; de même, *M. Ezekiel* publiait en 1930 un ouvrage qui devint très classique : *Correlation Analysis.* A l'Université d'Ames, dans l'Iowa, *G.W. Snedecor* publiait, dans les années trente, des travaux sur l'analyse de la variance, ainsi que sur les applications de la statistique à l'expérimentation agricole. Une importante innovation a été apportée par *W.A. Shewhart* dans le domaine du contrôle statistique dans la qualité des fabrications, en publiant en 1931 le premier ouvrage en la matière, *Economic Control of Quality of Manufactured Products.* Ce n'est que plus tard, avec l'effort de guerre et l'impulsion d'après-guerre, que la statistique américaine s'est propulsée dans les premiers rangs mondiaux.

Sans atteindre l'importance de l'Ecole anglaise, l'**Italie** a été au cours de ce premier demi-siècle relativement riche en statisticiens de renom ; ils ont été surtout orientés vers la démographie et le calcul des probabilités. De nombreuses universités ont bénéficié de leur enseignement ; celle de Rome a été associée à l'Institut Central de Statistique, et l'est toujours. Si *Corrado Gini* en est la figure la plus marquante, le lot n'en comprend pas moins de nombreux noms prestigieux tels que : *M. Boldrini, F.P. Cantelli, G. Castelnuovo, Bruno de Finetti, Livio Livi, G. Mortara, A. Niceforo, Vito Volterra.*

Dans les pays de l'**Est** et du **Nord de l'Europe**, l'enseignement et la recherche en matière de calcul des probabilités et de statistique, en différentes universités, a été important. Il suffit pour s'en rendre compte de consulter la bibliographie de l'ouvrage de M.G. Kendall. On y trouve notamment les références importantes concernant *Harald Cramér, Herman Wold* pour la Suède, *Ragnar Frisch, T. Haavelmo* pour la Norvège, *Gerhardt Tintner, Abraham Wald* pour l'Autriche, *Jerzy Neyman* en Pologne, le Hollandais *Jan Tinbergen*, les Allemands *E.J. Gumbel, R. von Mises*, de nombreux Russes : *A. Khintcine, A. Kolmogoroff, A. Liapounoff, A.A. Markow, A.A. Tchouprow.* Beaucoup de ces noms nous sont familiers.

Avant le déclenchement de la guerre, plusieurs de ces professeurs émigrèrent vers l'Ouest, fuyant les régimes totalitaires. *Jerzy Neyman* en est sans doute le plus connu. Dès 1934, il rejoignit Egon Pearson à l'University College de Londres et apporta tout son talent à la théorie de la vérification des hypothèses. Il quittera Londres en 1938 pour l'Université de Berkeley, en Californie, qui devint avec lui l'un des principaux foyers mondiaux de la recherche statistique.

E.J. Gumbel, spécialiste le plus connu de l'étude des valeurs extrêmes des distributions statistiques, a aussi rejoint l'Angleterre, après avoir enseigné avant les hostilités à l'Université de Lyon et donné quelques conférences à la Sorbonne.

Signalons, pour lui rendre hommage, le cas d'un jeune et brillant probabiliste juif allemand, *Wolfgang Doeblin*, réfugié à Paris en 1936 et poursuivant études et recherches à l'Institut Henri Poincaré, qui trouva la mort dans les combats de 1940. La bibliographie de Kendall cite quatre de ses articles, datés de 1936 à 1939, sur les probabilités en chaîne, sur les puissances des lois de probabilité et sur l'addition d'un grand nombre de variables aléatoires.

En **France**, ce sont les grands mathématicienss qui ont contribué aux progrès et à l'enseignement de la statistique théorique. *Emile Borel* avait dès 1895 publié une *Introduction à la théorie des nombres*, et en 1914, un ouvrage, *Le Hasard*. Il a été de 1924 à 1938 le chef de file d'une cohorte de professeurs et de spécialistes qui ont contribué à la réalisation d'un grand *Traité du calcul des probabilités et de ses applications*, complété par une *Collection de monographies des probabilités*. Le tout comprenait plus de vingt-cinq volumes. Y ont notamment contribué : *Paul Lévy (Théorie de l'addition des variables aléatoires), Maurice Fréchet (Recherches théoriques sur les probabilités et les variables aléatoires), Francis Perrin (Mécanique statistique quantique), Robert Deltheil (Probabilités géométriques), René Risser (Applications de la statistique à la démographie et à la biologie).*

L'enseignement de la statistique à proprement parler est demeuré néanmoins assez limité. Sur le plan universitaire, une seule chaire existait en France, qui préparait le certificat de *Calcul des probabilités et Statistique mathématique*, dont le titulaire *Emile Borel* était assisté de *Georges Darmois* pour la partie Statistique mathématique. Le Professeur Darmois a été le maître incontesté et très estimé de la statistique mathématique en France, depuis la publication de son cours *Statistique mathématique*, en 1928, jusqu'à sa mort, en 1960. Ses caractéristiques étaient la clarté et la précision ; il utilisait une notation vectorielle dans l'hyperespace à n dimensions, dont sans doute s'est inspiré plus tard Benzecri dans la création de l'Analyse des données.

Sous l'impulsion conjointe des universitaires, des dirigeants de la Statistique Générale de la France (prédécesseur de l'actuel INSEE), dont la plupart étaient des polytechniciens, et des compagnies d'assurances, l'Institut de Statistique de l'Université de Paris (ISUP) fut créé à la Sorbonne en 1922. Dans les années trente, les principaux cours portaient sur les matières suivantes :

- Méthode statistique (appelée aujourd'hui Statistique descriptive), par *L. Dugé de Bernonville* et *Michel Huber* ;
- Calcul des probabilités et statistique mathématique, par *Georges Darmois* ;
- Démographie, par *Henri Bunle* ;
- Economie, par *René Roy* ;
- Mathématiques financières, par *Alfred Barriol* ;
- Théorie des assurances, par M. *François* ;

– Hygiène, médecine et assistance sociale, par le Dr *Ichok*.

Daniel Dugué, après un séjour à Rothamsted, chez R.A. Fisher, était chargé des travaux pratiques. Il devait devenir directeur de l'ISUP beaucoup plus tard.

Il existait aussi au Conservatoire national des Arts et Métiers, le CNAM, un cours d'Economie industrielle et statistique, dispensé magistralement en deux ans par *François Divisia*. Divisia s'intéressait beaucoup aux nombres-indices, ainsi qu'aux applications de la statistique à l'économie, et tenait avec ses étudiants des séances de travail sur ces questions.

Alfred Sauvy animait à l'Ecole polytechnique un groupe de recherche sur les fluctuations économiques, appelé *X-Crise*, fréquenté par de nombreux chercheurs en la matière.

Un tout premier institut privé de sondages d'opinions, l'IFOP (Institut Français de sondages d'Opinions), a été créé à Paris en 1938 par notre regretté collègue *Jean Stoetzel*, à l'image de l'Institut Gallup des Etats-Unis.

Signalons enfin, pour terminer cette première partie de notre exposé, qu'Emile Borel, directeur de l'Institut Henri Poincaré, et Georges Darmois, qui y était aussi installé, créèrent fin 1937 au sein de cet institut un *Laboratoire de Statistique* destiné à favoriser les recherches et les applications. Parmi les travaux qui y étaient conduits, mentionnons des tests de signification concernant des médicaments et des traitements médicaux, pour le compte d'instituts voisins, notamment l'Institut Curie, ainsi que les premières applications en France des méthodes de contrôle statistique des fabrications.

Période 1946-1970. Le tournant d'après-guerre

Les principaux thèmes développés dans la statistique et son enseignement dès la fin de la Seconde Guerre mondiale ont été *les sondages*, avec leurs applications aux recensements, aux enquêtes, à la recherche agricole, au contrôle des fabrications, *la comptabilité nationale*, les applications à *l'économétrie*, et *la recherche opérationnelle*, qu'il vaut mieux appeler *théorie de la décision*.

Pour ce qui concerne **les sondages**, une lignée de spécialistes s'est épanouie aux **Etats-Unis**, avec notamment *W.E. Deming, M.H. Hansen, W.N. Hurwitz, R.J. Jessen, B.J. Tepping*, auxquels il faut ajouter le Canadien *Nathan Keyfitz*, et même *Tore Dalenius* venu de Suède. Leurs recherches et leur enseignement ne se sont pas limités aux milieux universitaires, mais ont pris place aussi dans les services administratifs spécialisés comme le Bureau of the Census, le Department of Agriculture (USDA), le National Bureau of Economic Research (NBER).

La recherche et l'enseignement de la statistique au niveau universitaire se sont largement étendus et intensifiés, notamment à Berkeley, Columbia (New York), Princeton, Stanford. A Berkeley, la présence de *J. Neyman* a été un grand atout. Il a été rejoint après les hostilités par *Michel Loève*, venu d'Egypte à l'Institut Henri Poincaré, qui avait eu la malchance d'être envoyé par les nazis en camp de concentration. Deux autres statisticiens-mathématiciens français ayant aussi rallié plus tard les Etats-Unis et qui méritent d'être mentionnés, sont nos collègues *Benoît Mandelbrot* et *Lucien Le Cam.*

L'Ecole statistique de l'**Inde**, déjà bien connue avant guerre avec *P.C. Mahalanobis, V.G. Panse, P.V. Sukhatme*, tous formés en Angleterre, s'est affirmée et largement développée, surtout dans la théorie des sondages et les applications à l'agriculture.

L'**Italie** a continué sur sa lancée; de nouveaux noms se sont distingués tels *G. d'all'Aglio, B. Barberi, F. Brambilla, V. Castellano, G. Pompilj, T. Salvemini.*

Aussitôt après la fin de la Seconde Guerre mondiale, l'**Espagne**, qui sortait elle-même d'une longue et sanglante guerre civile, a connu l'éclosion d'une pléiade de statisticiens de qualité : *Angel Anos, Francisco Azorin, Enrique Cansado, Jose Ros Jimeno, Sixto Rios, J. Torrens-Ibern.* Ils enseignaient aux universités de Madrid et de Barcelone, et avaient l'appui de l'Institut National de Statistique ainsi que celui du Conseil Supérieur de la Recherche Scientifique. Pour les applications à l'industrie et au contrôle des fabrications, l'appui est venu de *Jose Antonio de Artigas*, de l'Escuela Central de Ingenieros Industrial de Madrid, qui de 1937 à 1939 avait fait des recherches à l'Institut Henri Poincaré auprès de Georges Darmois.

En **France**, après la Libération, l'enseignement du calcul des probabilités et de la statistique s'est étendu progressivement à plusieurs universités : Bordeaux, Lyon, Nancy, Rennes, Strasbourg, Toulouse, que ce soit dans les facultés des sciences, les facultés de médecine et de pharmacie ou les facultés de droit, devenues facultés de droit et de sciences économiques. Il en a été de même pour les grandes écoles d'ingénieurs, de commerce et d'affaires.

Un pas important a été franchi par la création de l'*Ecole d'Application de l'INSEE*, lequel, en 1946, sous l'impulsion de *F.L. Closon*, avait pris la succession de la Statistique Générale de la France, devenue pendant la guerre Service National de Statistique. Cette école, créée en 1942 mais démarrant véritablement en 1946 sous la direction d'*Eugène Morice*, était destinée, en collaboration avec l'ISUP, à former le personnel technique de l'INSEE. Elle a pris par la suite le nom d'*Ecole Nationale de Statistique et d'Administration Economique (ENSAE)*, et son enseignement s'est ouvert à une gamme plus large d'étudiants. Dans l'immédiat après-guerre, les cours ont été assurés par *Georges Darmois* pour la statistique mathématique et pour plusieurs spécialités par de jeunes normaliens ou polytechniciens, ayant pour la plupart transité par l'IHP ou par l'ISUP, tels *Fernand Chartier* pour les méthodes d'estimation, *Pierre Thionet* pour les sondages, ou *Jean Mothes* pour les applications industrielles. Tous les cours ont été rédigés et publiés.

En annexe de l'ISUP, G. Darmois a créé en 1952 un *Centre d'Enseignement et de Recherche de Statistique Appliquée*, le *CERESTA*, pour former aux méthodes du contrôle statistique les ingénieurs des entreprises industrielles. D'abord animé par *Jean Mothes*, ce centre a été dirigé dès 1955 par *André Vessereau*, auquel a succédé en 1990 *J.-P. Therme*. L'activité de formation du centre vient d'être reprise par un autre organisme, le *CISIA*.

Aussitôt après sa création, le CERESTA a entrepris la publication de la *Revue de Statistique Appliquée, R.S.A.*, animée d'abord par E. Morice et G. Morlat. Par la suite, A. Vessereau et J.-P. Therme en ont assuré la publication, avant qu'elle ne soit transmise à la Société de Statistique de France. Celle-ci en a confié l'édition à l'A.S.U. ; *Pierre Cazes* en est le rédacteur en chef actuel.

Toujours en liaison avec l'ISUP, un organisme spécialisé dans l'enseignement de la statistique dans le domaine médical, le CESAM, a été créé et animé par *Daniel Schwartz*, véritable pionnier en la matière depuis les années cinquante. Il faut remarquer néanmoins que la statistique médicale et l'épidémiologie avaient déjà fait l'objet de travaux en France dans la seconde moitié du XIX[e] siècle par *Villermé, Broca* et *Gavarre.*

Pour revenir aux années qui ont suivi la Seconde Guerre mondiale, deux noms doivent être cités : *Bradford Holl* et *Richard Dole*, qui ont inauguré l'évaluation scientifique des traitements médicaux, concernant notamment le cancer, et l'épidémiologie.

L'après-guerre a vu se créer de nombreux **Centres internationaux de formation statistique**, sous l'égide des Nations Unies et d'institutions internationales spécialisées. Le premier a sans doute été le *Centre Européen d'Application statistique agricole et démographique* qui s'est tenu à Paris de septembre à décembre 1949, organisé par la FAO, les Nations Unies, l'UNESCO et le gouvernement français. La FAO a réalisé en 1953 le premier centre africain de statistique agricole, à Ibadan en Nigéria, auquel ont participé pour l'enseignement en français Jacques Desabie, Gérard Théodore, Jules Leveugle, Jacques Royer et moi-même.

A l'initiative de l'Office de Statistique des Communautés Européennes, a été créé à Paris, avec l'assistance pédagogique de l'INSEE, le *Centre Européen de Formation des Statisticiens économistes des Pays en Développement (CESD).* Ce Centre, initialement dirigé par *Guy Le Hégarat*, a formé 360 statisticiens de niveau supérieur de ces pays avant d'être transféré à Abidjan en 1993. Pour la formation des adjoints aux travaux statistiques, le transfert s'est effectué dès 1978 sur Abidjan, Kigali et Yaoundé.

Bien d'autres centres de formation ou de perfectionnement ont suivi à travers le monde, et ce mouvement se poursuit, surtout pour contribuer à former les statisticiens des pays en développement, ou en phase de transition. Signalons aussi que l'*Institut International de Statistique* joue maintenant un rôle important dans ce domaine, en particulier avec sa section spécialisée, l'*Association Internationale pour l'Enseignement de la Statistique (IASE).*

Il faut maintenant dire quelques mots de la recherche opérationnelle, de l'informatique, de l'économétrie et de la comptabilité nationale, qui ont connu de très importants développements dans l'immédiat après-guerre.

La recherche opérationnelle trouve son origine dans les calculs effectués pendant la guerre, en Grande-Bretagne et aux Etats-Unis, en vue d'obtenir le maximum de réussite dans des opérations de caractère ou d'intérêt militaire, avec la mise en œuvre du minimum de moyens, ou en essayant de courir le minimum de risque ou de subir le moins possible de pertes. *"Operations Research"*, c'était la recherche en vue de mener au mieux les opérations. Combien fallait-il, par exemple, d'avions pour être certain d'atteindre une cible, ou pour avoir une certaine probabilité de l'atteindre, avec un minimum de pertes en hommes et en avions ? Par quelle route, ou quelles routes, acheminer depuis les Etats-Unis les Liberty Ships pour avoir telle probabilité de faire parvenir hommes et matériels en Angleterre, en dépit des sous-marins allemands qui les guettaient dans l'Atlantique ? C'est ainsi qu'ont été développés les méthodes mini-max, ou maxi-mini, les systèmes de programmation mathématique, la théorie des graphes.

La paix venue, tout cela trouvait naturellement de larges applications dans la recherche scientifique, dans l'industrie, les transports, les affaires. En France, c'est *Jacques Lesourne* qui a créé, en 1958, la première et plus importante société dans ce domaine, la *SEMA, Société d'Economie et de Mathématiques Appliquées.* Dans ses équipes figuraient de nombreux statisticiens et même des scientifiques : *Robert Fortet, Bernard Roy.* Une filiale était créée par *Jacques Antoine*, chargée des enquêtes par sondages, la *SOFRÈS, Société Française d'Enquêtes par Sondages*, qui est encore aujourd'hui la plus importante société française de sondages.

Robert Lattès créait une autre filiale de la SEMA, la *Société d'Informatique Appliquée*, qui par la suite a été absorbée par la SEMA, dont elle est devenue la principale activité. Transformée en *SEMA-Group*, l'entreprise est devenue franco-britannique, et ses actions sont cotées à la fois sur la Bourse de Londres et sur celle de Paris.

La SEMA a largement contribué à l'enseignement de matières liées à la statistique, au moyen de nombreux séminaires et sessions de formation.

Pour le statisticien, **l'informatique** a été la grande aventure du siècle. *I.B.M.* s'appelait à l'origine la *Société Hollerith*, du nom de l'ingénieur norvégien qui inventa les machines à cartes perforées. Jusqu'à la veille de la dernière guerre, elles servaient essentiellement à dépouiller les recensements, à faire d'autres travaux statistiques lourds (tels les relevés des statistiques douanières, ou ceux du trafic des chemins de fer) et à tenir la comptabilité de très grandes entreprises. La filiale française de Hollerith s'appelait la *Société Electro-Comptable de France.* Son siège se trouvait place Vendôme, à Paris ; il y est toujours.

La chaîne des machines à statistique, ou comptables, comprenait les perforatrices, les trieuses et les tabulatrices-imprimantes. Travail fastidieux, lent et bruyant, nécessitant en fin de course une mise en page et un passage à une véritable imprimerie.

Les laboratoires de recherche statistique étaient équipés de machines à calculer tournant à la main, et quelques machines électriques commençaient à venir, mais elles étaient bruyantes et se mettaient souvent en panne.

Pendant la guerre, les Américains ont construit les premiers véritables ordinateurs, par exemple l'UNIVAC, fonctionnant toujours avec des cartes perforées, mais ayant une grande capacité et rapidité de travail. L'utilisation de la numération binaire a tout changé et a permis de trier et de mémoriser les données au moyen de circuits oscillants. Tout a été facilité après la guerre avec l'emploi des semi-conducteurs, ou transistors, à la place des diodes et triodes de ces circuits, d'où le développement rapide des ordinateurs de grande puissance.

La SEMA avait pu obtenir dès 1964 l'ordinateur le plus important d'Europe, un CDC 3600 (Control Data), puis un CDC 7200. Mais il fonctionnait toujours au départ avec des cartes perforées. En fin de course seulement, les données pouvaient être stockées sur des rubans magnétiques. Pour faire fonctionner la machine, il fallait tout programmer, utilisant pour cela les langages Cobol ou Fortran. Les engins occupaient un hall de plus de quarante mètres de long, entièrement climatisé. C'était déjà un progrès considérable, mais cela a un goût de préhistoire, comparé aux moyens dont on dispose aujourd'hui.

L'ordinateur a largement facilité l'application des méthodes de la recherche opérationnelle. A celle-ci il faut d'ailleurs rattacher la gestion des stocks, qui a connu d'énormes progrès, l'ordonnancement des fabrications complexes, lesquelles font appel à la théorie des graphes, l'évaluation des projets d'investissement, la gestion des portefeuilles, etc.

Disons maintenant un mot de **la comptabilité nationale**. A part ceux de Leontief, les travaux effectués dans les années trente par Colin Clark en Angleterre, par Kuznets aux Etats-Unis, par d'autres au Canada ou en Italie, visaient surtout à déterminer l'agrégat principal, à savoir le *revenu national.* Pendant les années de guerre et surtout après la guerre, le besoin s'est fait sentir de décomposer ce revenu en ses différents constituants, et surtout de connaître la part due aux différents intervenants dans l'activité nationale : entreprises, ménages, administrations, secteur financier.

Sous l'impulsion de l'économiste britannique *J.R.N. Stone*, l'OCDE publiait en 1950 un système simplifié de comptes nationaux, devenu en 1952 un système normalisé. De leur côté, les Nations Unies présentaient en 1953 leur *Système de Comptabilité Nationale.* Cela restait insatisfaisant quant à l'articulation des comptes des différents groupes d'acteurs.

Les chercheurs français se sont aussi attaqués à ce problème : *L.-A. Vincent* et *J. Dumontier* à l'INSEE, *Claude Gruson* au Ministère des Finances, *François Perroux* et *Jean Marczewski* à l'ISEA (Institut de Sciences Economiques Appliquées). Il en est résulté un système de comptes nationaux relativement complet qui a grandement contribué à établir le système des Nations Unies tel qu'il existe actuellement. Le système de comptes nationaux donne maintenant aux services centraux de statistique un cadre général dans lequel s'insère l'ensemble des statistiques d'ordre économique.

Ce système comptable national a permis, notamment à l'INSEE, d'élaborer des *modèles économétriques* représentant l'ensemble de l'économie. A partir de tels modèles, il est possible de faire des simulations, d'essayer de tirer des prévisions moyennant un ensemble d'hypothèses et de contribuer ainsi à l'orientation rationnelle de certaines décisions de la part des autorités responsables.

Disons qu'au préalable les études et l'enseignement en matière d'**économétrie** s'étaient largement développées, en particulier avec les apports de *Maurice Allais*, d'*Edmond Malinvaud* et *Charles Prou.* Allais enseignait l'économie à l'Ecole des Mines depuis 1944, et, en 1954, il prenait la succession de

René Roy à l'ISUP, pour y enseigner l'économie théorique. Malinvaud professait l'économétrie à l'ENSAE et assura la direction de cette école avant de prendre la direction générale de l'INSEE ; sous son impulsion, l'enseignement de l'économie a pris une part beaucoup importante à l'ENSAE. Quant à Prou, il a dirigé le CEPE, Centre d'Etudes des Programmes Economiques.

Au Conservatoire national des Arts et Métiers, *Jean Fourastié* a pris la succession de François Divisia dans la chaire d'Economie industrielle et statistique ; il a également enseigné à l'Institut d'Etudes politiques et à l'Ecole pratique des Hautes Etudes.

Après 1970. L'épanouissement

Les développements dans la recherche, les méthodes, les moyens, les applications et l'enseignement de la statistique ont été spectaculaires après les années soixante. Nous n'en évoquerons que quelques aspects.

L'analyse des données, lancée par *Jean-Paul Benzecri* dès 1964 par son cours à la Faculté des Sciences de Rennes, et par la publication de son ouvrage en 1973, a ouvert une nouvelle voie à l'analyse statistique : analyse factorielle, analyse multidimensionnelle, en composantes principales, recherche des correspondances, etc. Tous les cours de statistique dans les facultés et les grandes écoles placent maintenant ces méthodes dans leurs programmes.

Les applications de ces méthodes ont été mises à la portée de tous les chercheurs par l'irruption du **micro-ordinateur** dès 1980, grâce au développement extraordinaire de sa vitesse de calcul et de ses capacités de mémoire. Bien entendu, les *logiciels*, de plus en plus performants, ont contribué, et contribuent toujours davantage, à la fête. L'enseignement bénéficie tout naturellement des nouveaux et puissants moyens informatiques.

Ces moyens permettent aussi d'intensifier la recherche dans de nombreux autres domaines; citons à titre d'exemple celui de **l'analyse des séries chronologiques**, pour laquelle une nouvelle méthodologie a du être créée.

L'enseignement, maintenant doté de matériel informatique, a continué à s'étendre, dans les universités comme dans les grandes écoles, ainsi que dans les instituts universitaires de technologie (IUT), et dans de nouveaux instituts dits "professionalisants". Les sociétés de consultants et les associations scientifiques ou professionnelles multiplient les cours et les séminaires. Et il en est de même sur le plan international où, après la formation dans les pays en développement, l'effort se porte maintenant en direction des pays "en transition" vers l'économie de marché.

En matière de **recherche**, il se produit aujourd'hui un véritable bouillonnement, que ce soit dans des domaines anciens, tels la *statistique bayesienne*, ou les *mouvements browniens (théorie du chaos)*, ou dans des domaines entièrement nouveaux, comme l'*apprentissage*, ou la *reconnaissance des formes*. Il faut maintenant que les choses puissent se décanter et laisser plus tard, à d'autres, le soin de reconnaître les principaux acteurs et les résultats déterminants.

Bibliographie

A.S.U. (1991) *L'enseignement de la Statistique en France*, Paris, A.S.U.

CORSINI C. (1989) *Per una Storia della Statistica Italiana*, Pise, Pacini Editore.

DROESBEKE J.-J. & TASSI PH. (1989) *Histoire de la statistique*, Paris, P.U.F.

KENDALL M.G. (1948) *The Advanced Theory of Statistics*, vol. II, Londres, Ch. Griffin and Co Ltd.

UNESCO (1957) *Les sciences sociales dans l'enseignement supérieur. La Statistique*, Paris, UNESCO.

NOUVELLE DÉMARCHE DANS L'ENSEIGNEMENT DU CALCUL DES PROBABILITÉS

Marie BERRONDO-AGRELL
Maître de Conférences à l'IUT d'Evry-Brétigny
et problémiste à *Valeurs Actuelles*

RÉSUMÉ

Le calcul des probabilités s'est développé historiquement à partir de l'expérience. Ce n'est qu'au cours du XX^e siècle qu'il perd "ce relent d'astronomie ou d'alchimie" dont parle KEYNES, au profit d'une lourdeur axiomatique paralysante [8]. Nous proposons ici de démocratiser son enseignement sans revenir pour autant à une approche purement expérimentale. Nous utiliserons pour cela les résultats récents développés par Anthony EDWARS [5], permettant d'établir une analyse graphique systématique du calcul des probabilités. C'est la "Syntaxe des diagrammes de Venn généralisés" [1]. Nous mettrons ici en évidence l'équivalence des deux méthodes. Puis, à l'aide de quelques exemples, nous montrerons la simplification didactique qui en résulte, ce que nous analyserons dans la conclusion.

SUMMARY

Probability calculus has developped (in history) thanks to experiment. It is only in this century that Probability Theory exchanges "this track of Astronomy and Alchimy", described by KEYNES, to a paralysing heavy axiomatic [8]. We present here a way to democratize its teaching, without coming back to the experimental old way. We use the recent works of Anthony EDWARDS [5], allowing to establish a "Syntax of Generalized Venn Diagramms" [1]. We show here the equivallence of the two methods. Then we use a few examples to show the easiness of solving. Probability problems in that way. This will be analized in the conclusion.

I. INTRODUCTION ET APERÇU HISTORIQUE

Les jeux de hasard existent depuis des temps immémoriaux. C'est en les analysant que PASCAL et FERMAT ont, dans leur correspondance, au XVII^{e} siècle, établi la base de la "Géométrie du hasard". Les développements ultérieurs de cette science ont pris le nom de "Calcul des probabilités"[4].

En 1897, le célèbre probabiliste Emile BOREL imaginait la très abstraite théorie des ensembles mesurables, qui fut exploitée ensuite dans l'intégrale de LEBESGUE (1901). Cela ne changea en rien sa propre présentation du calcul des probabilités. Tout au contraire, il ne rêvait que de démocratiser son enseignement, par une théorie qui se prête à une pédagogie facile. Mais en 1933, à Moscou, KOLMOGOROFF reprit la théorie des ensembles mesurables pour fonder de façon parfaitement axiomatique le calcul des probabilités dans son célèbre traité "*Grundbegriffe de Warscheinlichkeitsrechnung*".[8]

Les probabilités perdaient certes ainsi ce "relent d'astronomie ou d'alchimie" dont parlait KEYNES en 1921, mais elles perdaient aussi la possibilité d'être comprises par l'immense majorité de la population. Cela n'empêcha pas l'exportation progressive de ces idées vers les pays de l'Ouest, et vers la France en particulier où elles furent progressivement unanimement appliquées.

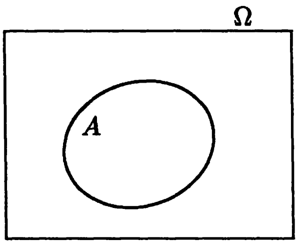
Il est donc aujourd'hui **théoriquement impossible à la population française, sauf à une faible minorité de mathématiciens, de COMPRENDRE le CALCUL DES PROBABILITÉS.**

Certes, beaucoup l'abordent néanmoins, en apprenant des séries de théorèmes par cœur, et les utilisent ensuite intuitivement, souvent en dépit du bon sens. Les professeurs ayant tout de même l'obligation de se trouver devant un pourcentage décent de copies correctes, ont pris l'habitude d'axer la difficulté sur l'algèbre combinatoire, ou l'analyse, au sens mathématique du terme, qui ne sont en réalité que des disciplines-outils pour le calcul des probabilités.

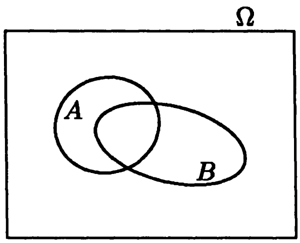
Il est grand temps que cette situation évolue, selon le propre désir d'Emile BOREL. Laissons KOLMOGOROFF aux étudiants spécialisés en mathématiques, et développons une théorie compréhensible par la majorité. Les Français méritent l'accès officiel à la sagesse probabiliste et par là même à une véritable assimilation de la statistique. Mais ne repartons pas pour autant vers un enseignement exclusivement basé sur l'expérience comme autrefois. Il faut reprendre la théorie des ensembles et l'utiliser dans ce qu'elle a de simplificateur et non pas d'inhibant. La probabilité se définit en effet comme une mesure sur l'ensemble des possibilités, tout comme l'aire sur un ensemble de surfaces. L'analogie entre surfaces et possibilités, entre aire et probabilité, permettra donc à l'individu ordinaire d'entrer sans le savoir dans le monde des tribus [7], puis d'appliquer automatiquement et à bon escient, les théoréorèmes fondamentaux du calcul des probabilités, qu'il s'agisse de l'axiome des probabilités totales ou des probabilités composées, des théorèmes de Bayes, de Poincaré, ... Il abordera ensuite en douceur les variables aléatoires sans se sentir paralysé par leur seule définition. Il s'aidera pour cela d'images de compréhension, analysées par la théorie des graphes [9].

II. IMAGES DE COMPRÉHENSION ET ANALYSE GRAPHIQUE

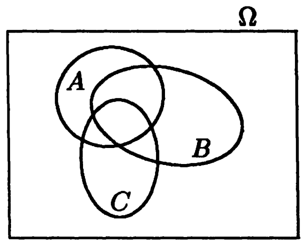
Pour présenter nos images de compréhension, il nous faut repartir aux découvertes de John VENN, alors recteur au Caïus Collège de Cambridge, en 1881 [10].
Ce sont les fameux diagrammes de VENN que voici :



Un contour fermé pouvant représenter un événement A.
$2^1 = 2$ surfaces distinctes représentant 2 possibilités.



Deux contours fermés représentant deux événements A et B.
$2^2 = 4$ surfaces distinctes représentant 4 possibilités.



Trois contours fermés représentant trois événements A, B et C.
$2^3 = 8$ surfaces distinctes représentant 8 possibilités.

Remarque : Nous parlons ici de surface distincte ou connexe.
La définition de la connexité se retrouve en topologie ou en théorie des graphes. On peut vulgariser cette notion en disant qu'une surface connexe est la surface découpée en remplaçant les coups de crayon par des coups de ciseaux.

On observe ici un blocage au niveau 3. Ce blocage fut rapidement repoussé au niveau 4 par Lewis Carroll (l'auteur d'*Alice au pays des Merveilles*) et le physicien KARNAUGH de la manière suivante :

Diagramme de Caroll

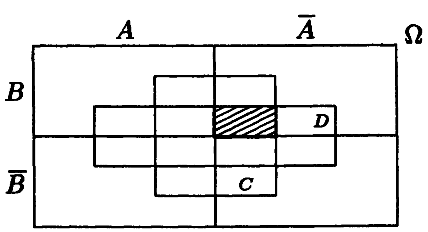
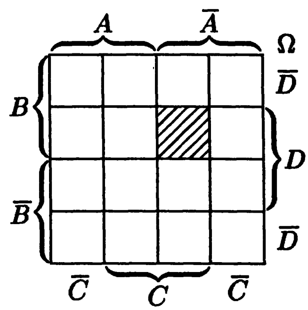


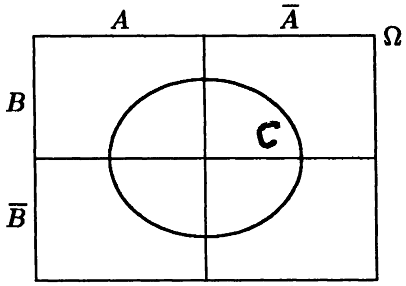
Diagramme de Karnaugh



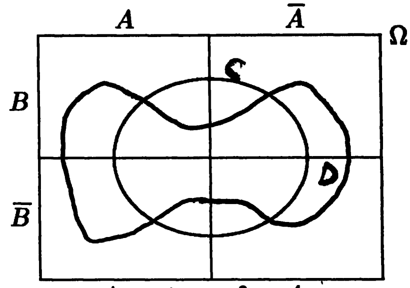
Dans un cas comme dans l'autre, nous observons 4 contours fermés partiellement tangents au rectangle de référence **R** représentant 4 événements, et $2^4 = 16$ surfaces distinctes représentant 16 possibilités.

Nous avons ici hachuré, à titre d'exemple, la partie connexe représentant la réalisation de B, C, D et du contraire de A.

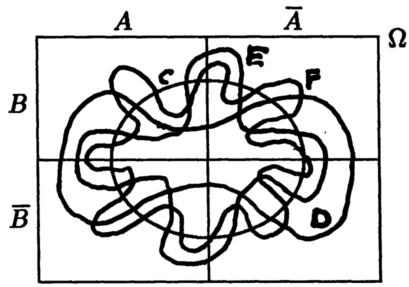
Ce blocage au niveau 4 interdisait toute récurrence basée sur de tels diagrammes. Mais cent ans plus tard, Anthony EDWARDS, professeur d'architecture au Caïus College de Cambridge, fut chargé de créer un vitrail pour commémorer l'œuvre de John VENN. C'est alors que fut découvert comment créer un système de n contours fermes dans un rectangle, pour former 2^n surfaces distinctes ou connexes, et ceci quel que soit n : /[5].



3 contours fermés,
$2^3 = 8$ surfaces connexes.



4 contours fermés,
$2^4 = 16$ surfaces connexes.

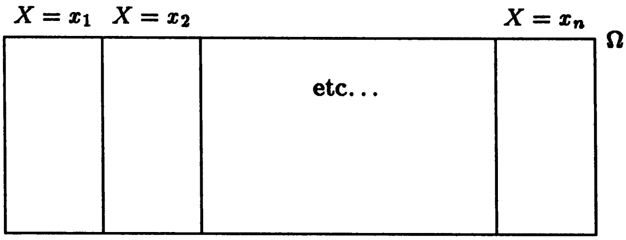


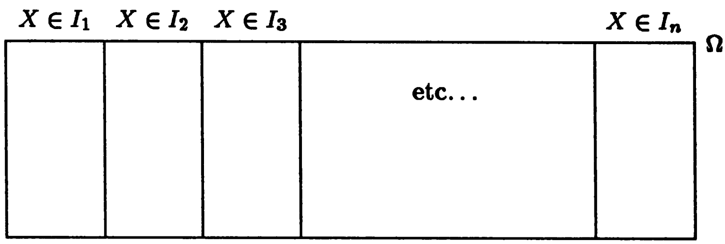
6 contours fermés,
$2^6 = 64$ surfaces connexes.

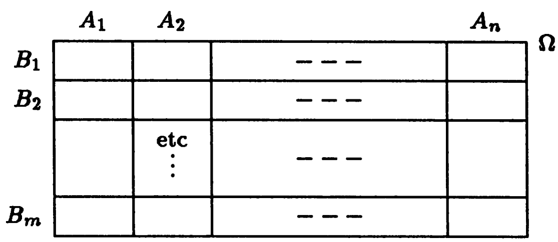
Comme on peut, plus ou moins clairement, le voir sur les figures ci-dessus, il s'agit d'un système de roues dentées de plus en plus fines, qui fut présenté en 1988 dans la revue *The scientific American* [5]. L'ordinateur peut en dessiner une infinité.

Dès lors, de tels diagrammes peuvent servir de base à des récurrences, et illustrer n événements quelconques à la fois, toutes les intersections possibles s'y trouvant représentées.

Nous avons ensuite joint une généralisation de ces diagrammes pour l'étude des variables aléatoires simples et doubles pouvant servir également à la présentation de variables statistiques ainsi qu'à des variables qualitatives diverses. Ce sont les diagrammes structurés dont voici quelques exemples.

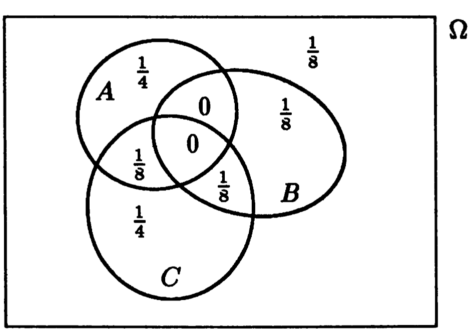






Les diagrammes précédents (Venn, Carroll, Karnaugh ou Edwards) seront alors qualifiés de "directs", par opposition aux "structurés"[1] (en lignes, colonnes, ou lignes et colonnes).

Il nous restait à pouvoir indiquer les probabilités correspondant à chaque possibilité, tantôt numériquement, tantôt à l'aide d'inconnues algébriques.
Ce sont les diagrammes "complets" ou "algébriquement complets" dont voici quelques exemples :



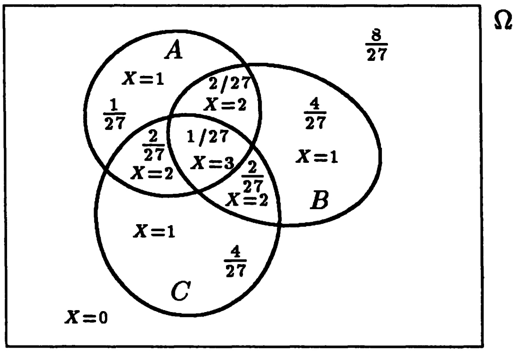
	A_1	A_2	A_3	A_4
B_1	x	0	$\frac{1}{4}$	x
B_2	0	$2x$	$\frac{1}{8}$	$\frac{1}{2}$

Ω

1. ou partitifs

La somme des valeurs indiquées sur les différentes parties connexes fera 1 s'il s'agit de probabilités. Mais le même type de diagrammes peut être utilisé pour indiquer des aires, des effectifs, des pourcentages ou des proportions sur des ensembles quelconques. C'est même de l'analogie avec les aires que viendra la simplification didactique.

Remarque : On peut aussi sur un diagramme introduire une (ou plusieurs) variable aléatoire en indiquant sa (ou ses) valeur dans chaque partie connexe. On parlera alors de "diagramme surchargé". En voici deux exemples :



avec X = nombre d'événements réalisés parmi A, B et C.

	$X=1$	$X=2$	$X=3$
$Y=1$	$Z=1$ 1/12	$Z=2$ 1/12	$Z=3$ 1/12
$Y=2$	$Z=2$ 1/12	$Z=2$ 1/12	$Z=3$ 1/12
$Y=3$	$Z=3$ 1/12	$Z=3$ 1/12	$Z=3$ 1/12
$Y=4$	$Z=4$ 1/12	$Z=4$ 1/12	$Z=4$ 1/12

Ω

avec $Z = \mathrm{Sup}(X, Y)$.

Diagrammes correspondants

pour X :

$X=0$	$X=1$	$X=2$	$X=3$
$\frac{8}{27}$	$\frac{12}{27}$	$\frac{6}{27}$	$\frac{1}{27}$

Ω

pour Z :

$Z=1$	$Z=2$	$Z=3$	$Z=4$
$\frac{1}{12}$	$\frac{3}{12}$	$\frac{5}{12}$	$\frac{3}{12}$

Ω

Nous avons ainsi des diagrammes directs ou structurés, complets ou surchargés. L'art de les utiliser pour résoudre un problème de probabilités forme la "syntaxe des diagrammes de Venn généralisés" (1 et 2).

III. EQUIVALENCE ENTRE LES 2 APPROCHES

Voici donc un tableau-résumé des principales notions de base du calcul des probabilités, avec leur équivalent-graphique.

Notion	Définition théorique	Equivalent graphique (dessinons, selon le contexte, un diagramme direct ou structuré)
Eventualité s_i	Une description possible de l'avenir dans un contexte donné.	Une surface connexe (morceau de papier distinct si on remplace les coups de crayon par des coups de ciseau).
Univers Ω	Un ensemble d'éventualités Ω tel que une et une seule d'entre elles va se réaliser.	L'ensemble des surfaces connexes formant le rectangle de base du diagramme.
Evénement	Un ensemble comportant 0, une ou plusieurs éventualités.	Un ensemble de 0, une ou plusieurs surfaces connexes qu'on pourra hachurer d'une façon particulière.
Tribu	Une partie $\mathcal{A}$ de $\mathcal{P}(\Omega)$ telle que : *i*) $\Omega \in \mathcal{A}$; *ii*) $(\forall A \in \mathcal{A})(\overline{A} \in \mathcal{A})$; *iii*) $(\forall i / A_i \in \mathcal{A})(\bigcup_{i=1,n} A_i \in \mathcal{A})$.	Ensemble de toutes les parties hachurables (événements). (Les axiomes écrits à gauche sont vérifiés.)
Probabilité	Une application P de $\mathcal{A}$ dans $[0,1]$ telle que : *i*) $P(\Omega) = 1$; *ii*) $(A \cap B = \varnothing)$ implique $(P(A \cup B) = P(A) + P(B) - P(A \cap B))$.	Ecrivons sur chaque surface connexe s_i une valeur positive ou nulle p_i dont la somme fait 1. La probabilité P d'un événement A quelconque est la somme des valeurs p_i écrites sous les hachures correspondant à cet événement. (Les axiomes écrits à gauche sont vérifiés.)
Espace probabilisable	Le triplet $(\Omega, \mathcal{A}, P)$: Univers, Tribu, Probabilité.	Un diagramme complet.
Variable aléatoire X	Application X de $\mathcal{A}$ dans $\mathbb{R}$ telle que : *i*) $(\Omega, \mathcal{A}, P)$ espace probabilisable ; *ii*) $(\forall I)(X^{-1}(I) \in \mathcal{A})$.	Valeur numérique X indiquée sur chacune des surfaces connexes d'un diagramme complet (avec ou sans précision), c'est-à-dire un diagramme complet-surchargé.

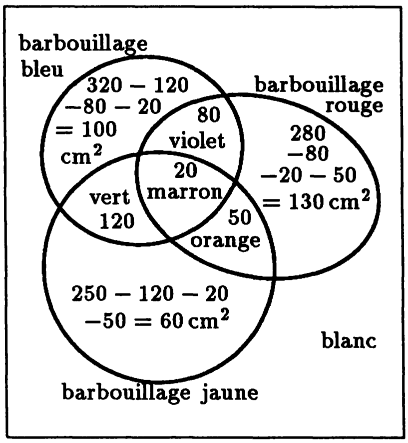
IV. DIDACTIQUE APPLIQUÉE

Lola et la peinture à l'eau.

Lola adore sa boîte de peinture à l'eau. Sur de belles feuilles blanches de format 21×27, elle barbouille $320\,\text{cm}^2$ de bleu, puis $280\,\text{cm}^2$ de rouge et $250\,\text{cm}^2$ de jaune. Après ces trois barbouillages, Lola observe qu'elle a ainsi obtenu $120\,\text{cm}^2$ de vert, $80\,\text{cm}^2$ de violet, $50\,\text{cm}^2$ d'orange et $20\,\text{cm}^2$ de marron (surperposition des trois barbouillages).

Une mouche se pose au hasard sur la feuille. Quelle est la probabilité pour que ce soit sur la partie restée blanche après la peinture à l'eau de Lola ?

▶ Diagramme direct, de degré 3, forme historique de Venn, complété par l'aire :



Ω

← Feuille de papier 21×27

Aire totale de la feuille :
$21 \times 27 = 567\,\text{cm}^2$.

Aire peinte :
$100 + 80 + 20 + 120 + 130 + 60 + 50 = 560\,\text{cm}^2$.

Il ne reste que $7\,\text{cm}^2$ de blanc.

Probabilité pour que la mouche se pose sur la partie restée blanche :

$$\frac{7}{21 \times 27} = \frac{1}{81}.$$

Safarinocepopotam.

Il y a ici 2 fois plus d'hippopotames que de rhinocéros. Un hippopotame sur 2 est gros (poids supérieur à 600 kilos), les autres sont aussi souvent petits (poids inférieur à 400 kilos) que moyens. Parmi les rhinocéros, on trouve 5 petits pour 1 gros et 2 fois moins de gros que de moyens. Voici au loin un animal gris, de poids moyen. Je tire. Je le tue. Quelle est la probabilité d'avoir ainsi supprimé un rhinocéros ?

▶ Diagramme doublement structuré, complété algébriquement selon la mesure "probabilité". Nous supposons donc choisir un animal au hasard.

	Il est petit	Il est moyen	Il est gros	Ω
C'est un rhinocéros	$5y = 5\,\frac{x}{8}$	$2y = \frac{x}{4}$	$y = \frac{x}{8}$	$8y = x$
C'est un hippopotame	$\frac{x}{2}$	$\frac{x}{2}$	x	$2y$
		$3\,\frac{x}{4}$		

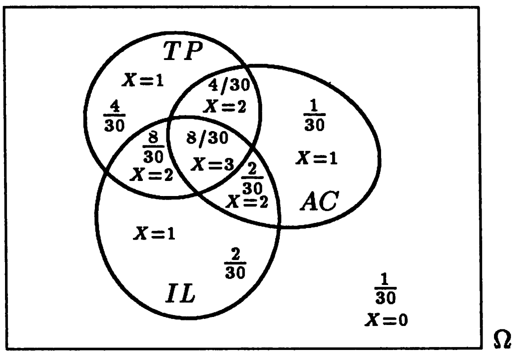
Probabilité pour qu'un animal de poids moyen soit un rhinocéros : $\dfrac{x/4}{3x/4} = \dfrac{1}{3}$.

Anne-Cécile ou le percepteur.

Je rentre de voyage. Je sais qu'il y a 4 chances sur 5 pour que le tiers provisionnel soit arrivé pendant ce temps, 2 chances sur 3 pour que les impôts locaux m'aient été adressés, 1 chance sur 2 pour que ma bien-aimée, Anne-Cécile, m'ait écrit, et que personne d'autre ne peut m'avoir envoyé de lettre. J'ouvre ma boîte aux lettres : 2 lettres m'attendent.

Quelle est alors la triste probabilité pour que ce soit Anne-Cécile et non l'un de mes percepteurs qui m'ait oublié ?

▶ Diagramme direct, degré 3, forme historique de Venn, complet et surchargé selon la variable aléatoire X = nombre de lettres reçues.



TP = "Je reçois le tiers provisionnel",
IL = "Je reçois les impôts locaux",
AC = "Anne-Cécile m'a écrit".

$$P(X=2) = \frac{14}{30},$$

$$P(\overline{AC}/X=2) = \frac{8/30}{14/30} = 0,57.$$

Au centre commercial.

Ce matin, j'ai rendez-vous chez le coiffeur à 9 h 30. Une fois sur deux, il me garde une heure ; sinon, c'est aussi souvent trois quarts d'heure qu'une heure un quart. Vingt-cinq minutes après, je serai au supermarché. Il me faut tantôt une demi-heure, tantôt trois-quarts d'heure, tantôt une heure pour faire mes courses, ces trois durées étant aussi fréquentes les unes que les autres.

Quand j'aurai fini, j'irai retrouver mon amie Patricia à la cafétéria à côté. Nous avons rendez-vous à midi moins le quart. Elle arrive toujours à l'heure précise. Laquelle de nous deux a-t-elle la probabilité la plus forte d'attendre l'autre ?

▶ Diagramme doublement structuré, complété par la probabilité, et surchargé par la variable aléatoire $X = 9\,\text{h}\,30 + Y + Z + 25\,\text{mm}$.

On a $P(X > 11\,\text{h}\,45) = \frac{4}{12}$.

En conclusion, elle a 4 chances sur 12 (soit 1 chance sur 3) de devoir m'attendre, tandis que j'ai 2 chances sur 3 de devoir l'attendre.

Temps chez le coiffeur en mm

Temps au supermarché en mm	$Y = 45$	$Y = 60$	$Y = 75$	Ω
$Z = 30$	$X = 11\,\text{h}\,10$ 1/12	$X = 11\,\text{h}\,25$ 2/12	$X = 11\,\text{h}\,40$ 1/12	1/3
$Z = 45$	$X = 11\,\text{h}\,25$ 1/12	$X = 11\,\text{h}\,40$ 2/12	$X = 11\,\text{h}\,55$ 1/12	1/3
$Z = 60$	$X = 11\,\text{h}\,40$ 1/12	$X = 11\,\text{h}\,55$ 2/12	$X = 12\,\text{h}\,10$ 1/12	1/3
	1/4	1/2	1/4	

V. Conclusion

Nous venons de voir de façon très concrète comment l'utilisation des diagramme de VENN généralisés nous permettait de donner des solutions exactes sans avoir consciemment appliqué les principaux axiomes des probabilités, le théorème de POINCARÉ, celui de BAYES, la définition d'une variable aléatoire et bien d'autres merveilles. Seule la définition d'une probabilité conditionnelle $P(A/B)$ est à joindre à l'analogie avec les aires. Les efforts de mémoire se trouvent ainsi considérablement diminués, d'autant plus que les données du problème sont en permanence exposées par cette "mise à plat". Les efforts d'imagination, nécessairement associés à la recherche du meilleur raisonnement, se trouvent également presque supprimés.

Notons qu'un programme informatique a été récemment mis au point par un groupe d'élèves de l'Ecole Centrale de Paris (laboratoire du Professeur DEJAX) pour résoudre des exercices de probabilité, et même de logique, en imageant sur écran, selon cette technique, les événements correspondants.

Cette pédagogie se trouve respectivement appliquée et développée dans mes récents ouvrages (1994) : [6] et [2] en collaboration avec Jacqueline FOURASTIÉ (CNAM). Nous y insistons sur la seule difficulté restante : le choix d'un diagramme adapté à chaque problème, selon le type d'information. Il s'agit autrement dit, du passage du langage courant au langage technique et du retour du langage technique au langage courant. Et nous sommes heureuses d'appliquer ainsi la véritable pensée de Borel [3] :

"Le calcul des probabilités est une des branches les plus attrayantes et les moins ardues de la mathématique".

Bibliographie

[1] AGRELL Per et BERRONDO-AGRELL Marie (1992) "Vers une syntaxe des diagrammes de Venn, lutte contre un mythe", *Journal de Statistiques de Paris,* N°1/2.

[2] BERRONDO-AGRELL Marie et FOURASTIÉ Jacqueline (Sept. 1994) *Pour comprendre les probabilités*, Hachette, Collection "Les Fondamentaux".

[3] BOREL Emile et DELTHEIL Robert (Mai 1923) *Probabilités - Erreurs,* Collection Armand Colin.

[4] DROESBEKE J.-J. et TASSI Philippe (Août 1990) *Histoire de la statistique*, Collection Que sais-je / P.U.F.

[5] EDWARDS Anthony (1989) "Venn Diagrammes for many Sets", *New Scientist,* Jan. 9th 1989.

[6] EURÊKA (Juin 1994) *Faîtes vos jeux*, Dunod.

[7] FRUGIER Gérard (Février 1994) *Formules Ordinaires des Probabilités et de Statistiques*, Ellipse.

[8] KOLMOGOROV (1993) *Gründbegriffe der Walscheinlichkeitsrechnung.*,

[9] LEGUEN Monique (Décembre 1995) "Statistique, Imagerie et Sciences Cognitives", *Bulletin de Sociologie méthodologique.*

[10] VENN John (1981) *Symbolic Logic*, Mac Millan - London.

Séance solennelle du 30 janvier 1996

PRIX DU STATISTICIEN de langue française

Allocution du président sortant Jean-Louis BODIN

Messieurs les Présidents,

Mes chers Confrères,

Je suis heureux de vous accueillir pour cette séance solennelle de notre Société. Permettez-moi tout d'abord, puisque nous sommes encore pour vingt-quatre heures dans une période où, en France, il est encore temps de souhaiter à tous ceux qui vous sont chers une excellente année, de vous présenter, à toutes et à tous, mes voeux de bonheur et de prospérité pour 1995, dans vos vies professionnelles et personnelles.

Mon premier, et des plus agréables, devoir ce soir est de remettre à nos deux lauréats du Prix du Statisticien d'expression française au titre de 1995 la médaille que leur a décernée en décembre 1994 le Jury formé, comme vous le savez, des anciens présidents de notre Société. Cette médaille aurait dû leur être remise l'année dernière, conformément au millésime qu'elle porte, mais des circonstances imprévues nous en ont empêché. Mais c'est avec plaisir que je le fais ce soir.

Comme vous le savez, notre Prix est décerné chaque année, sur proposition des anciens présidents de la Société, et par roulement, à un statisticien français confirmé, puis à un statisticien étranger d'expression française, puis enfin à un jeune statisticien. Pour 1995, il s'agissait de distinguer un jeune statisticien ; le Jury formé de ces anciens présidents n'a pas pu trancher entre deux brillants candidats, et a donc décidé de partager le Prix et de leur remettre à chacun la médaille qui rappelle cette distinction. C'est ainsi que j'ai l'honneur ce soir de remettre cette médaille – par ordre alphabétique ! – à Daniel PIERRE-LOTI-VIAUD et à Christian ROBERT. Permettez-moi de vous les présenter, avant de leur demander de nous exposer très rapidement – nous ne disposons, hélas ! que de peu de temps ce soir – leurs travaux scientifiques.

Daniel PIERRE-LOTI-VIAUD a trente-huit ans. Agrégé de Mathématiques à vingt-trois ans, il a soutenu avec succès deux thèses, l'une de troisième cycle et l'autre d'Etat, sur des sujets liés à ce qui constitue l'un des fils conducteurs de ses travaux, l'étude des processus. Il enseigne aujourd'hui à l'ISUP où il est responsable de la filière Actuariat-Assurances, et il est également responsable des stages du DEA de Statistiques à l'Université de Paris-VI. Ses travaux et publications portent pour l'essentiel sur les processus de Galton-Watson dépendant de l'effectif de la population, et sur les techniques d'approximations fortes de la fonction de répartition empirique et l'étude des processus gaussiens associés.

Christian ROBERT a trente-quatre ans. Titulaire à vingt-quatre ans à la fois d'un DEA de Mathématiques pures décerné par l'Université de Paris-VI, et du diplôme de statisticien-économiste de l'ENSAE, il est actuellement détaché au laboratoire de statistique du CREST, c'est-à-dire du Centre de Recherche en Economie et en Statistique du groupe des écoles nationales d'Economie et de Statistiques de l'INSEE. Malgré son âge, Christian Robert a à son actif un total impressionnant de trois ouvrages et de trente-huit publications ; il est «référé» pour quatorze revues comptant parmi les plus prestigieuses dans notre profession, telles que : *Annals of Mathematical Statistics, Econometrica, Journal of the American Statistical Association* ou *Journal of the Royal Statistical Society.* Il a été professeur invité dans deux des plus prestigieuses universités américaines, Cornell, puis Purdue.

Avant de passer la parole à nos deux lauréats, je voudrais vous révéler le nom du lauréat 1996 sur lequel le Jury des anciens présidents vient de délibérer. Il s'agissait cette année de récompenser un statisticien français confirmé. Le choix du Jury s'est porté sur Pierre DUCIMETIÉRE. Pierre DUCIMETIÉRE est un épidémiologiste, ancien élève de l'Ecole polytechnique (promotion 1962), directeur de recherche à l'INSERM dont il a été pendant quatre années membre du Conseil scientifique et vice-président du Haut Comité de la Santé publique (dont le président est le ministre de la Santé). Il faut noter que c'est la première fois que cette fonction est confiée à un non-médecin. Ses travaux ont fait l'objet de plus de cent publications et il est rédacteur en chef de la *Revue d'Epidémiologie et de Santé publique.* Au moment où nous cherchons à unifier toutes les composantes de la Statistique française ou francophone autour de la SSF, ce qui pourrait se concrétiser d'ailleurs prochainement par le rapprochement de l'Association des Epidémiologistes de langue française et de la SSF, il a paru intéressant à notre Jury d'attribuer notre Prix à un représentant de cette spécialité.

PRINCIPAUX TRAVAUX DE RECHERCHE EN PROBABILITÉ ET STATISTIQUE

Daniel PIERRE LOTI VIAUD

Professeur à l'Université Paris-VI

ISUP-LSTA

Nous présentons succinctement les différents thèmes de recherche sur lesquels nous avons travaillé. Nous nous sommes efforcés de respecter l'ordre chronologique dans lequel nous les avons abordés.

1. Processus de Bienaymé-Galton-Watson

L'essentiel de nos travaux est une contribution à l'étude des processus de BIENAYMÉ-GALTON-WATSON, ou processus de ramification, et de leurs généralisations. Ces processus sont des modèles stochastiques de l'évolution dans le temps et l'espace de populations éventuellement soumises à diverses influences. L'un des plus simples de ces modèles peut se décrire de la manière suivante. Ici on suppose que les individus de la population ont une durée de vie égale à 1, ainsi le temps n désigne la $n^{ième}$ génération de la population et on appelle Z_n l'effectif de la $n^{ième}$ génération. L'effectif de la $(n+1)^{ième}$ génération est alors obtenu en comptant tous les enfants des individus de la $n^{ième}$ génération :

$$Z_{n+1} = \sum_{i \in \{1,\dots,Z_n\}} \xi_{n,i} \,. \tag{1}$$

Dans cette équation $\xi_{n,i}$ est le nombre de descendants de l'individu i de la $n^{ième}$ génération. Lorsque les $\xi_{n,i}$ sont des variables aléatoires indépendantes et équidistribuées (v.a. i.i.d.), ces processus ont été étudiés depuis le siècle dernier. Ils sont maintenant bien connus depuis la fin des années 1960. La généralisation de l'étude de ces processus lorsqu'on fait l'hypothèse que la loi du nombre d'enfants d'un individu dépend de l'effectif de la population au moment où cet individu se reproduit est au centre de nos travaux sur les processus de BIENAYMÉ-GALTON-WATSON. Dans ce cas, conditionnellement à Z_n, les variables aléatoires $\xi_{n,i}$, $i = 1, 2, \dots$, sont des réalisations indépendantes et équidistribuées d'une loi qui dépend de Z_n. On notera $m(k) = \mathrm{E}(\xi_{n,i}/Z_n = k)$, pour $k = 0, 1, \dots$, le nombre moyen d'enfants d'un individu lorsque l'effectif de la population est égal à k. On notera m le nombre moyen d'enfants lorsqu'il ne dépend pas de k.

1-a. Paramètre de Malthus

Notons $\alpha = \log m$. Rappelons que, dans le cas classique où les variables aléatoires $\xi_{n,i}$, $i = 1, 2, \dots$, $n = 0, 1, \dots$, sont indépendantes et équidistribuées, et lorsque m est strictement plus grand que 1 , alors on observe le comportement suivant pour le processus $Z_0, Z_1, \dots$:

$$\mathcal{P}\Big(\{\exists N \text{ avec } Z_N = Z_{N+1} = \cdots = 0\} \bigcup \{\lim_{n\to\infty} \frac{Z_n}{e^{n\alpha}} = W > 0\}\Big) = 1 ,$$

c'est-à-dire, ou bien il y a extinction de la population, ou bien l'effectif de la population croît avec une vitesse exponentielle. α est le paramètre de MALTHUS, il donne le taux de la croissance exponentielle de l'effectif de la population.

Lorsque la loi du nombre d'enfants d'un individu dépend de l'effectif de la population, et si on suppose en outre que $\lim_{n\to\infty} m(k) = m_\infty$, où m_∞ est strictement plus grand que 1 , ce résultat a été en partie généralisé par KUSTER (1985) et KLEBANER (1985). Il restait cependant à démontrer que la variable aléatoire W est strictement positive sur l'ensemble des trajectoires où $Z_0, Z_1, \dots$ ne s'annule pas. En nous appuyant sur un résultat de KERSTING (1992), c'est ce que nous avons obtenu dans un travail paru au *Journal of Applied Probability* en 1994. Nous y avons également étudié la vitesse de convergence du processus normalisé en démontrant par des techniques de martingales un théorème limite central qui étend les résultats connus dans le cas classique où les $\xi_{n,i}$ sont des v.a. i.i.d..

1-b. Cadre multitype : convergence du vecteur des proportions

Lorsqu'il y a différents types d'individus dans la population le processus de BIENAYMÉ-GALTON-WATSON multitype classique est bâti sur l'hypothèse que la loi du vecteur $\xi_{n,i}$, donnant les nombres d'enfants des différents types d'un individu, ne dépend que du type de cet individu. Dans ce cas, on sait depuis le résultat de KESTEN et STIGUM (1966) que le vecteur V_n des proportions des différents types d'individus dans la $n^{ième}$ génération de la population converge presque sûrement sur l'ensemble où l'effectif de la population croît vers l'infini. Nous avons étudié la généralisation de ce résultat lorsque l'on suppose que la loi du vecteur $\xi_{n,i}$ donnant les nombres d'enfants d'un individu de la $n^{ième}$ génération dépend du type de cet individu mais aussi de V_n . D'une part, pour un modèle à deux types, en temps continu (i.e. les individus ont une durée de vie aléatoire, elle est indépendante d'un individu à l'autre et tirée suivant une loi exponentielle dans le cas que nous avons considéré), dans un article publié aux *Annales de l'Institut Henri Poincaré* en 1985. D'autre part, pour le modèle en temps discret présenté ici, dans un travail récent publié au *Journal of Applied Probability* en 1995. Dans les deux cas, nous utilisons et développons des techniques de la théorie des algorithmes stochastiques et des systèmes dynamiques pertubés. Le second travail permet de compléter les résultats de

KLEBANER (1989), et conduit, au moins dans certains cas, à une démonstration simplifiée des trés jolis mais plutôt compliqués résultats de KESTEN (1972). De plus, il est énoncé dans le cadre général des systèmes dynamiques perturbés et, ainsi, il s'applique dans d'autres domaines, notamment les extensions stochastiques de l'algorithme EM étudiées par CELEUX et DIEBOLT (1992) et BISCARAT (1993). Nous l'avons appliqué pour montrer la convergence de la proportion d'individus infectés dans une population se renouvelant comme un processus de BIENAYMÉ-GALTON-WATSON et soumise à l'influence d'une épidémie récurrente.

1-c. Grandes déviations

Pour simplifier nous considérons ici des processus de BIENAYMÉ-GALTON-WATSON à un seul type d'individus. L'idée suivante est à la base d'une autre partie de notre travail. Notons $\theta_n = Z_n/Z_0$ pour $n = 0, 1, \dots,$ et $T(\theta) = \theta m(\theta)$ pour $\theta \geq 0$. L'équation (1) est alors facilement transformée en :

$$(2) \qquad \theta_{n+1} = T(\theta_n) + \varepsilon_{n+1} \quad \text{où} \quad \varepsilon_{n+1} = \theta_n \frac{1}{Z_n} \sum_{i \in \{1, \dots Z_n\}} (\xi_{n,i} - m(\theta_n))$$

(ici on a implicitement fait l'hypothèse que la loi du nombre d'enfants d'un individu dépend de θ_n en lieu et place de Z_n). Sur l'ensemble où $Z_0, Z_1, \dots$ converge vers l'infini, par la loi des grands nombres, ε_{n+1} agit comme une petite perturbation qui s'ajoute au système dynamique déterministe $x_{n+1} = T(x_n)$. En faisant tendre Z_0 vers l'infini, on obtient ainsi une famille de systèmes dynamiques perturbés dont l'étude à été introduite par une école russe à la fin des années soixante (LABKOVSKII [1972]). Dans un travail effectué en collaboration avec Frédéric PORTAL (CNRS-Université Paris XI) et publié au *Bulletin des Sciences Mathématiques* en 1993, nous avons proposé une nouvelle méthode permettant de montrer un principe de grandes déviations pour des systèmes dynamiques perturbés, en temps discret, plus généraux que ceux donnés par (2). De manière très schématique nous pouvons résumer un tel principe de grandes déviations sous la forme suivante, où N appartenant à $\{1, 2, \dots\}$ est fixé :

$$\mathcal{P}\Big((\theta_0, \dots, \theta_N) = (x_0, \dots, x_N)\Big) \approx \exp(-Z_0\, I(x_0, \dots, x_N)),$$

quand Z_0 est au voisinage de l'infini. Ici $I(.)$ désigne la fonctionnelle des grandes déviations associée à notre processus, elle se calcule en fonction des transformées de CRAMER des lois des $\xi_{n,i}$. D'autres méthodes avaient précédemment été proposées par MORROW et SAWYER (1989) et KIFER (1990). Partant des travaux en temps continu de FREIDLIN et WENTZELL (1984), notre méthode nous permet d'étendre dans le cadre multitype les travaux de LABKOVSKII (1972), ceci sous un minimum d'hypothèses. Dans un travail paru dans *Mathematical Biosciences* en 1993 nous avons appliqué cette méthode à de nouveaux modèles plus réalistes d'épidémies de type SIR (susceptible $\rightarrow$ infected $\rightarrow$ removed) avec renouvellement des individus susceptibles dans la population. Dans ce dernier cas, nous obtenons, dans l'asymptotique Z_0 converge vers l'infini, des estimations du temps d'extinction d'une épidémie récurrente. Pour une autre approche de ce problème voir KRYSCIO et LEFÈVRE (1989).

Antérieurement à ce travail nous avions démontré un principe de grandes déviations d'une forme inhabituelle pour une autre transformation du processus $Z_0, Z_1, \dots$. Indiquons que cette transformation du processus $Z_0, Z_1, \dots$ est celle utilisée par BIGGINS (1979), RUGET (1981) et ROUAULT (1985) pour étudier une généralisation spatiale du processus de BIENAYMÉ-GALTON-WATSON. Fixons $T > 0$. Notons $X_t = \log(Z_{[t \log Z_0]}) / \log Z_0$ pour $0 < t < T$, où $[.]$ désigne la partie entière d'un nombre réel. Nous considérons toujours l'asymptotique Z_0 converge vers l'infini. Dans ce cas l'équation (1) prend la forme :

$$\frac{1}{1/\log Z_0} \left(X_{t+(1/\log Z_0)} - X_t \right) = \log m(X_t) + \varepsilon_t , \tag{3}$$

où

$$\varepsilon_t \approx \frac{1}{m(X_t)\exp(X_t \log Z_0)} \sum_{i \in \{1,\dots,[\exp(X_t \log Z_0)]\}} (\xi_{[t \log Z_0],i} - m(X_t)) ,$$

quand Z_0 est au voisinage de l'infini. Remarquons que, dans ce cas, le nombre de termes figurant dans la somme définissant le terme ε_t est de l'ordre de $Z_0^{X_t}$ avec X_t qui converge (dans (2) ε_{n+1} contient un nombre de termes de l'ordre de $Z_0 \theta_n$ avec θ_n qui converge). Ainsi, les propriétés asymptotiques des familles de processus associées à (2) et à (3) sont complétement différentes. Dans notre thèse d'état nous avons démontré une loi des grands nombres, un théorème limite central et un principe de grandes déviations pour la famille de processus associée à (3). Le principe de grandes déviations a fait l'objet d'une publication aux *Annales de l'Institut Henri Poincaré* en 1991. De manière trés schématique ce principe de grandes déviations se résume sous la forme suivante :

$$\mathcal{P}(X_t = x_t\, , \, 0 \leq t \leq T) \approx \exp\left(- Z_0^{I(x_t, 0 \leq t \leq T)} \right) ,$$

quand Z_0 est au voisinage de l'infini. Ici encore la fonctionnelle des grandes déviations I se calcule en fonction des transformées de CRAMER des lois des $\xi_{n,i}$.

1-d. Le cas spatial

Supposons ici que les individus de la population vivent sur $(-\infty, \infty)$. Les $\xi_{n,i}$ sont alors des processus ponctuels qui fixent le nombre et le lieu de naissance des enfants d'un individu. Pour chaque n, Z_n est un processus ponctuel qui décrit la place des individus de la $n^{ième}$ génération. Des résultats de type loi des grands nombres ont été montrés par BIGGINS (1979) et ROUAULT (1985) pour l'extension spatiale de la famille de processus associée à (3). BIGGINS (1979) considère le cas où les $\xi_{n,i}$ sont des v.a. i.i.d. et ROUAULT étend en partie les résultats de BIGGINS lorsque la loi des $\xi_{n,i}$ dépend de l'instant et du lieu où l'individu se reproduit. Dans notre thèse d'état nous considérons le cas où la loi des $\xi_{n,i}$ dépend de l'effectif de la population à l'endroit où l'individu se reproduit. En particulier, lorsque l'espace se réduit à $\mathbb{Z}$, nous avons étudié la convergence du système déterministe associé à la famille de processus (obtenu en prenant $\varepsilon = 0$ dans l'analogue de (3)). Ce système déterministe se comporte comme un schéma numérique approximant les solutions d'une équation d'HAMILTON-JACOBI. Dans un travail publié aux *Comptes rendus de l'Académie des Sciences* en 1988 nous avons démontré la convergence de ce schéma numérique, une vitesse de convergence pour ce schéma numérique a également été donnée.

2. Processus empiriques

Etant donnée une suite $U_1, \ldots, U_n$ de v.a. i.i.d. de loi uniforme sur $(0,1)$, notons :

$$F_n(s) = \frac{1}{n} \sum_{i \in \{1,\ldots,n\}} \mathbb{1}_{\{U_i \le s\}} \text{ et } \alpha_n(s) = \sqrt{n}(F_n(s) - s) \text{ pour } 0 \le s \le 1,$$

la fonction de répartition empirique et le processus empirique uniforme. Notons aussi :

$$G_n(s) = \inf\{t \ge 0\,;\, F_n(t) \ge s\} \text{ et } \beta_n(s) = \sqrt{n}(G_n(s) - s) \text{ pour } 0 \le s \le 1,$$

la fonction des quantiles empirique et le processus des quantiles empirique. BAHADUR (1966) et KIEFER (1967, 1970) ont établi, parmi d'autres résultats, que, pour s fixé, presque sûrement :

$$\limsup_{n\to\infty} n^{1/4}(2\log\log n)^{-3/4}|\alpha_n(s) + \beta_n(s)| = 2^{1/2}3^{-3/4}(s(1-s))^{1/4}.$$

Récemment DEHEUVELS, DEHEUVELS et MASON (1990-2) ont donné une nouvelle démonstration de ce résultat et ont développé des résultats analogues pour d'autres processus de type BAHADUR-KIEFER. Dans un travail paru au *Journal of Multivariate Analysis* en 1993 nous avons étendu ce résultat

au cas de l'estimateur de KAPLAN-MEIER de la fonction de répartition empirique que l'on utilise lorsque les observations $U_1, \ldots, U_n$ sont censurées par des réalisations indépendantes et équidistribuées d'une autre loi qui sont indépendantes des observations. Ce résultat a été obtenu en construisant une approximation forte pour notre processus de type BAHADUR-KIEFER par un processus brownien itéré (du type $t \rightarrow X(Y(t))$ où X et Y sont deux mouvement browniens standard indépendants) et en utilisant une loi fonctionnelle du logarithme itéré pour ce processus brownien itéré.
Dans la continuité de ce travail, un article écrit en collaboration avec ZHAN SHI et YUE HU (L.S.T.A. et U.R.A. 1321, Université Paris 6) et paru au *Journal of Theoretical Probability* en 1995, présente différentes lois fonctionnelles du logarithme itéré pour le processus brownien itéré.

3. Mathématiques actuarielles

Pour des raisons professionnelles nous nous sommes récemment intéressés aux mathématiques actuarielles. Plus particulièrement aux systèmes de *bonus-malus* basé sur l'hypothèse que la loi des fréquences de sinistres est un mélange de lois de POISSON, ou de lois binomiales, ou de lois binomiales négatives. Dans un travail trés récent, et soumis pour publication, nous étudions un modèle statistique paramétré contenant les trois types précédents de mélanges de lois. Nous construisons aussi un test, basé sur la méthode du bootstrap, de l'hypothèse la loi des fréquences de sinistres est un mélange de lois de Poisson.

Principales références de l'auteur

1985 – "Processus de branchements dépendant de la densité, markovien en temps continu". *Annales de l'Institut Henri Poincaré,* vol. 21, 3, 289-303.

1988 – "Dynamique de population et équation d'Hamilton-Jacobi". *Comptes rendus de l'Académie des Sciences,* vol. 307, 1, 807-812.

1990 – *Processus de Galton-Watson dépendant de l'effectif de la population.* Thèse d'état, Université Paris XI.

1991 – "Grandes déviations pour une famille de processus de Galton-Watson dépendant de l'effectif de la population". *Annales de l'Institut Henri Poincaré,* vol. 27, 2, 141-179.

1993 – "A Pointwise almost Sure Bahadur-Kiefer-type Representation for the Product Limit Estimator". *Journal of Multivariate Analysis,* vol. 45, 1-8.

1993 – "Large Deviations for Random Perturbations of Discrete Time Dynamical Systems", en collaboration avec F. Portal. *Bulletin des Sciences Mathématiques,* vol. 117, 333-355.

1993 – "Large Deviations for Discrete-Time Epidemic Models". *Mathematical Biosciences,* vol. 117, 197-210.

1994 – "A Strong Law and a Central Limit Theorem for Controlled Galton-Watson Processes". *Journal of Applied Probability,* vol. 31, 22-37.

1995 – "Laws of the Iterated Logarithm for Iterated Wiener Processes", en collaboration avec Y. HU et Z. SHI. *Journal of Theoretical Probability,* vol 8, 303-319.

1995 – "Random Perturbations of Recursive Sequences with an Application to an Epidemic Model". *Journal of Applied Probability,* vol. 32, 559-578.

1996 – *Mélanges de lois sur* $\mathbb{N}$ *et système de bonus-malus.* Soumis pour publication.

PRIX DU STATISTICIEN DE LANGUE FRANÇAISE

FAUT-IL ACCEPTER OU REJETER LES P-VALUES ?

Christian P. ROBERT[1]
INSEE-CREST-Laboratoire de Statistique - Malakoff

1. Introduction

Au contraire de la théorie de l'estimation ponctuelle, la théorie des tests et la théorie des régions de confiance qui en découle ont connu une formalisation classique relativement tardive, sous la direction de NEYMAN, PEARSON et WALD. A présent, cette formalisation, exposée et approfondie dans LEHMANN [1986], est devenue l'approche "classique" des tests d'hypothèses, en particulier pour les praticiens.

De nombreuses critiques, théoriques (bayésiennes, conditionnelles) et pratiques, demeurent néanmoins sans réponse et le fossé avec la théorie bayésienne est beaucoup plus marquée que pour l'estimation ponctuelle –où les estimateurs de BAYES *moins favorables* engendrent les estimateurs *minimax*. D'autre part, l'usage des *p-values* (ou *niveaux de signification*) s'est progressivement répandu parmi les statisticiens appliqués sans qu'une étude théorique ne vienne véritablement valider l'utilisation de ces indicateurs. En effet, bien qu'ils soient généralement déterminées à partir de tests optimaux, on ne disposait pas jusqu'à présent de critère d'évaluation de ces instruments.

Il est à noter que la théorie des tests engendre une méthode de construction de *régions de confiance* (voir LEHMANN, 1986). Ces régions de confiance ne conservent pas l'optimalité des tests dont elles sont issues puisque HWANG et CASELLA [1982] ont exhibé, dans le cas normal, un *effet Stein* pour la région de confiance usuelle, i.e. une domination en probabilité par la sphère de même rayon recentrée en un estimateur de JAMES-STEIN. (Pour une revue des développements en ce domaine, voir ROBERT et SALEH [1989] et ROBERT [1992]). Outre son intérêt intrinsèque –proposer une meilleure région de confiance pour la moyenne d'une loi normale–, ce résultat appelait une évaluation indépendante des estimateurs par régions de confiance. (CASELLA, HWANG et ROBERT [1993] reprennent et généralisent les résultats dans ce domaine, en proposant des fonctions de coût pour les régions de confiance.)

1. Cet article est partiellement issu de l'article de HWANG *et al.* (1992), rédigé avec J. HWANG, G. CASELLA, M. WELLS et R. FARRELL lors d'un séjour à l'Université Cornell (E.U).

De même, il semble donc nécessaire d'évaluer les p-values *per se* en fonction d'un nouveau critére, favorisant l'émergence d'estimateurs –tests– qui répondent aux critiques passées de NEYMAN-PEARSON tout en englobant l'approche bayésienne. Le coût proposé dans cet article est une solution possible, bien que d'autres alternatives restent à examiner.

En Section 2, nous rappelons les diverses approches de la théorie des tests, en soulignant leur antagonismes. La Section 3 présente le critère alternatif qui conduit à des estimateurs plus adaptatifs, tout en englobant critiques et résultats précédents, et la Section 4 contient quelques remarques suppléementaires.

2. Approches fréquentiste et bayésienne des tests

2.1 La théorie de Neyman-Pearson

2.1.1 Définitions

Etant donnée une observation x d'une distribution $f(x|\theta)$, $\theta \in \Omega$, on considère deux sous-ensembles, Θ_0 et Θ_1, de Ω et l'on cherche à tester la validité de l'hypothèse (*nulle*), $H_0 : \theta \in \Theta_0$, contre l'hypothèse (*alternative*), $H_1 : \theta \in \Theta_1$.

Exemple 1. Soit $x \sim \mathcal{N}(\theta, 1)$, $\Theta_0 =]-\infty, 0]$ et $\Theta_1 =]0, +\infty[$. On teste alors l'hypothèse nulle $H_0 : \theta \leq 0$ contre $H_1 : \theta > 0$. $\triangle$

Exemple 2. Soit $x \sim \mathcal{B}(n, \theta)$, $\Theta_0 = \{1/2\}$ et $\Theta_1 = \Theta_0^c$. On teste alors l'hypothèse ponctuelle $H_0 : \theta = 1/2$ contre $H_1 : \theta \neq 1/2$. $\triangle$

Choisissant un *seuil* préalable, α, qui permet de borner supérieurement l'erreur sous H_0, la méthode de Neyman-Pearson cherche à déterminer un *test*, φ, i.e. un estimateur à valeurs dans $\{0, 1\}$, qui minimise le risque d'erreur sous H_1, les valeurs "0" et "1" étant respectivement associées aux réponses "oui" et "non" à la question "H_0 est-elle vraie ?". Sous la contrainte

$$\sup_{\Theta_0} E_\theta[\varphi(x)] \leq \alpha , \tag{2.1}$$

on cherche donc à maximiser $E_\theta[\varphi(x)]$ (ou *puissance*) pour $\theta \in \Theta_1$. Les tests remplissant cette condition sont dits *uniformément plus puissants* (ou UMP). (Voir LEHMANN (1986, Chap.3) pour des développements.)

Dans le cas de deux hypothèses *simples*, i.e. $\Theta_0 = \{\theta_0\}$ et $\Theta_1 = \{\theta_1\}$, le *lemme de Neyman-Pearson* donne la forme des tests UMP,

$$\varphi(x) = \begin{cases} 1 & \text{si} \quad f(x|\theta_1) > kf(x|\theta_0), \\ 0 & \text{si} \quad f(x|\theta_1) < kf(x|\theta_0), \end{cases} \tag{2.2}$$

où k est fixé en fonction de α. Dans les cas plus généraux où les hypothèses ne sont pas simples, il arrive que des tests UMP n'existent pas. Il faut alors se restreindre à la classe des tests *uniformément plus puissants sans biais* (UMPU), qui satisfont

$$\sup_{\theta_0} E_\theta[\varphi(x)] \leq \inf_{\theta_1} E_\theta[\varphi(x)] . \tag{2.3}$$

Exemple 1. (*suite*) Un test UMP au niveau $\alpha = 0.05$ est donné par

$$\varphi(x) = \begin{cases} 1 & \text{si } x > 1.65 , \\ 0 & \text{sinon.} \end{cases}$$

Par contre, si l'on teste $H_0 : \theta \neq 0$, un test UMP n'existe pas et un test UMPU à $\alpha = 5\%$ est

$$\varphi(x) = \begin{cases} 1 & \text{si } |x| > 1.96, \\ 0 & \text{sinon} . \end{cases}$$

$\triangle$

2.1.2 Critiques

Les critiques à l'égard de la formalisation ci-dessus sont nombreuses et concernent surtout l'aspect rigide des tests UMP (ou UMPU).

Une critique usuelle des tests UMP est liée au choix de la constante α. Le seuil doit-il être choisi parmi quelques valeurs "étalons" ou doit-il y avoir équilibrage entre seuil et puissance, comme le conseille LEHMANNN (1986) ? La détermination du seuil étant en général complexe, il est fréquent, en pratique, de recourir au "sacro-saint" 5%, malgré les dangers potentiels de cette approche, illustrés par les exemples suivants :

Exemple 3. (*Berger et Sellke, 1987*) Un astronome teste des hypothèses normales ponctuelles et les rejette si l'écart est supérieur à 2. Puis il apprend fortuitement que la borne à 5% est en fait 1.96 et décide de réexaminer les résultats passés dont il a pu vérifier la véracité. En supposant que la proportion d'hypothèses nulles vraies soit 50%, il découvrira que, parmi les écarts compris entre 1.96 et 2, au moins 22% correspondaient à des hypothèses vraies...

$\triangle$

Exemple 4. (*Berger, 1985*) Du temps de KEPLER, la précision des instruments et le nombre d'observations sur la trajectoire des planétes lui auraient permis de conclure au seuil $\alpha = 5\%$ que la trajectoire était effectivement une ellipse. Le même test aujourd'hui rejetterait cette hypothèse.

$\triangle$

En second lieu, le fait que φ soit à valeurs dans $\{0, 1\}$ empêche une réponse adaptative et graduelle. Par exemple, dans le cas normal, pour un test à 5%, la réponse sera la même si $|x|$ vaut 1.97 ou 10000, bien que l'une des deux observations supporte beaucoup plus fortement l'hypothèse alternative. L'approche de NEYMAN-PEARSON ne permet donc pas une réactualisation du seuil de confiance en fonction de l'observation. Cette critique est à relier à l'*approche conditionnelle* de la théorie des tests, appuyée par KIEFER (1977) et ROBINSON (1979), qui suggère de conditionner par rapport à des statistiques ancillaires pour éviter des comportements paradoxaux comme dans l'exemple suivant.

Exemple 5. (*Berger et Wolpert, 1988*) Soient x_1 et x_2 deux observations i.i.d. de la distribution

$$P(X = \theta - 1) = P(X = \theta + 1) = 1/2 .$$

Si

$$T(x_1, x_2) = \begin{cases} \dfrac{x_1 + x_2}{2} & \text{si } x_1 \neq x_2 , \\ x_1 - 1 & \text{sinon} , \end{cases}$$

un test à 25% de $H_0 : \theta = 0$ contre $H_1 : \theta \neq 0$ est

$$\varphi(x_1, x_2) = \begin{cases} 1 & \text{si } T(x_1, x_2) \neq 0, \\ 0 & \text{sinon} . \end{cases}$$

Bien entendu, ce test est en fait au niveau 0% pour $|x_1 - x_2| = 2$ et au niveau $\alpha = 50\%$ sinon. On voit alors l'intérêt de conditionner par rapport à la v.a. ancillaire $|x_1 - x_2|$. $\triangle$

Malheureusement, le choix de la "bonne" statistique ancillaire dans des cas plus concrets est généralement moins aisé.

Laissant de côté les cas où les tests UMPU n'existent pas et les situations où l'on doit recourir à la *randomisation* pour atteindre le niveau désiré (comme dans l'exemple 2) –ce qui contredit à la fois le bon sens et le *principe de vraisemblance* (BERGER ET WOLPERT, 1988)—, une critique plus fondamentale est que la théorie de NEYMAN-PEARSON ne répond pas à la question posée. JEFFREYS (1961) résume ce défaut par :

> *A hypothesis that may or may not be true is rejected because a greater departure from the trial value is impossible, i.e. it has not predicted something that has not happened...*

En d'autres termes, le conservatisme trop poussé des tests UMP est en partie dû au fait qu'ils ne prennent en compte que l'évènement $\{T(X) \geq T(x)\}$, plutôt que $\{T(X) = T(x)\}$. Il y a donc perte d'information au détriment de l'hypothèse nulle, ce qui explique les valeurs plus optimistes des réponses bayésiennes, comme le montre l'exemple suivant (voir aussi Section 2.2).

Exemple 6. (*Paradoxe de Lindley-Jeffreys*) – Soit $\bar{x}_n \sim \mathcal{N}(\theta, 1/n)$ et $\theta \sim \mathcal{N}(0, 1)$, situation correspondant à des observations répétées de la même distribution. Le test de $H_0 : \theta = 0$ contre $H_1 : \theta \neq 0$ dépend uniquement de la valeur de $z_n = \bar{x}_n/\sqrt{n}$. Pour $z_n = 1.97$, le test classique donne une réponse négative à 5% pour tout n, tandis que la réponse bayésienne, i.e. le *Bayes factor*, tend vers 1 lorsque n tend vers $+\infty$.

$\triangle$

2.1.3 P-values

Les *p-values* (ou niveaux de signification) constituent une réponse partielle aux critiques précédentes tout en restant dans un cadre fréquentiste. On les définit généralement comme le plus petit niveau α pour lequel H_0 serait rejetée en utilisant un test UMP ou UMPU. Pour l'exemple 1, si l'on observe $x = 2.1$, la p-value correspondante sera donc $p(x) = 0.04$. Lorsqu'il n'existe pas de test UMP ni UMPU, on recours généralement à un test du *rapport de vraisemblance.*

Exemple 2. (*suite*) Dans ce cas, on prend

$$p(x) = P_{1/2}[f(X|1/2) \leq f(x|1/2)]\,.$$

$\triangle$

Il est certain que ces statistiques présentent une approche plus graduelle que les tests UMP, en proposant un degré d'évidence en faveur de H_0. Elles évitent également le problème du choix du seuil de confiance mais évaluent toujours le "mauvais" objet, i.e. la probabilité de dépasser la valeur observée – la critique du paradoxe de LINDLEY-JEFFREYS s'applique tout aussi bien dans ce contexte–. De plus, elles constituent une "adhocquerie" car, construites de manière empirique à partir des tests UMP, elles ne sont validées par aucun critère d'évaluation. Enfin, difficiles à interpréter (ou à "calibrer"), les p-values sont aussi potentiellement dangereuses, car elles donnent l'illusion de pouvoir parler de la probabilité $P(H_0|x)$ sans avoir à passer par une modélisation bayésienne.

2.2 Solutions bayésiennes

Etant donnée la dichotomie imposée par la division hypothèse nulle –hypothèse alternative, on peut représenter la modélisation bayésienne sous la forme suivante : π_0 (resp. π_1) est la probabilité *a priori* que $\theta \in \Theta_0$ (resp. $\theta \in \Theta_1$) et g_0 (resp. g_1) est la densité de probabilité de θ sur Θ_0 (resp. θ_1). Deux mesures bayésiennes sont usuellement associées à cette distribution *a priori* :

(i) le *Bayes factor*,

$$B = \frac{P(\theta \in \Theta_0|x)}{P(\theta \in \Theta_1|x)} \Big/ \frac{x_0}{\pi_1}\,, \tag{2.4}$$

(ii) la probabilité *a posteriori*, $P = P(\theta \in \Theta_0|x)$.

La seconde solution est sans doute plus naturelle pour un bayésien mais implique une modélisation supplémentaire si $\Theta_1 \neq \Theta_0^c$. La statistique (2.4) est souvent défendue comme plus "objective" car elle élimine l'effet des poids π_0 et π_1 –il reste cependant à choisir g_0 et g_1. Elle est donc l'évidence apportée par les observations "seules", car le rapport produit une réévaluation de la balance *a priori* entre H_0 et H_1. L'éloignement de (2.4) de 1 peut ainsi être interprété comme "facteur de vraisemblance" pour l'acceptation ou le rejet de H_0.

Une critique habituellement avancée par les fréquentistes est que la détermination des densités g_0 et g_1 est difficile, voire impossible. Afin de robustifier l'approche bayésienne, on peut donc proposer des bornes inférieures sur P et B, $\underline{P}$ et $\underline{B}$, en travaillant sur des classes de densités pouvant être acceptées comme raisonnables par des non-bayésiens. Ainsi, BERGER et SELLKE (1987) et BERGER et DELAMPADY (1988) ont étudié plusieurs classes de lois pour des hypothèses nulles ponctuelles, $H_0 : \Theta_0 = \{\theta_0\}$.

Exemple 1. (*suite*) Si l'on prend g_1 dans la famille

$$G_1 = \{\text{densités unimodales symétriques autour de} \, \theta_0\},$$

les bornes sur B et P sont données dans le tableau suivant :

dim	1	3	5
$\underline{P}$	0.109	0.083	0.076
	0.392	0.350	0.339
$\underline{B}$	0.123	0.090	0.082
	0.644	0.540	0.531

pour des p-values de 0.01 et 0.1 (BERGER et DELAMPADY, 1988). La différence entre réponses bayésiennes et p-values est donc nette. On peut même élargir la classe à

$$G_2 = \{\text{toutes les distributions}\},$$

et cependant obtenir la table suivante (en dimension 1) (BERGER et SELLKE, 1987)

p − value	0.10	0.05	0.01	0.001
$\underline{P}$	0.205	0.128	0.035	0.004
$\underline{B}$	0.236	0.146	0.036	0.004

Δ

L'opposition entre réponses bayésiennes et p-values ne se limite pas au cas normal mais est aussi présente pour la plupart des distributions, en particulier les distributions discrètes.

Exemple 2. (*suite*) La notion de distribution symétrique pour tester H_0 : $p = p_0$ se traduit en l'ensemble de densités *a priori*

$$G_3 = \left\{ \text{distributions symétriques en } \frac{p - p_0}{\sqrt{p(1-p)}} \right\}$$

et donne le tableau suivant (BERGER ET DELAMPADY, 1988)

p − value	0.0093	0.0507	0.1011
$\underline{P}$	0.0794	0.2210	0.2969

Δ

Ainsi, dans le cas des tests ponctuels et, plus généralement ***bilatères***, les p-values n'appartiennent pas au champ de variation des réponses bayésiennes, donc sont trop conservatrices, en partie pour les raisons évoquées en 2.1. CASELLA ET BERGER (1987) viennent pondérer cette conclusion car ils obtiennent pour des hypothèses unilatères, des résultats diamétralement opposés, i.e. que les p-values peuvent être atteintes par des réponses bayésiennes et que, pour certaines distributions, notamment celles possédant la propriété du ***rapport de vraisemblance monotone*** (voir Lehmann, 1986), elles forment en fait la borne inférieure du domaine de variation de ces réponses.

Exemple 7. Soit x suivant une distribution de CAUCHY avec paramètre de position θ et soit à tester $H_0 : \theta \leq 0$ contre $H_1 : \theta > 0$. Lorsque g, densité –inconditionnelle– de θ, appartient à la famille

$$G_4 = \{\text{distributions symétriques autour de } 0\},$$

les valeurs respectives des p-values et des bornes inférieures sont

p − value	.437	.102	.063	.013	.004
$\underline{P}$	.429	.077	.044	.007	.002

Δ

Ces valeurs soulignent l'aspect conservatif des p-values, puisqu'elles correspondent généralement à la borne inférieure des probabilités ***a posteriori*** et que l'identité avec une réponse bayésienne et "ponctuelle" (i.e. la p-value ne correspond pas la même probabilité pour toutes les observations). Néanmoins, ils légitimisent partiellement l'usage des p-values dans le cas unilatère. La dichotomie observée ci-dessus entre les deux types d'hypothèses sera conservée et expliquée dans la section suivante.

3. Un critère d'évaluation des tests

On peut considérer un test, à valeurs dans $[0,1]$, comme un estimateur de l'indicateur $I_{\Theta_0}(\theta)$, sous l'hypothèse alternative naturelle $H_1 : \theta \notin \Theta_0$. La restriction à $\{0,1\}$ adoptée par NEYMAN et PEARSON est néfaste car elle introduit une "non-convexité" inutile dans le problème. Au contraire, considérer des estimateurs à valeurs dans $[0,1]$ autorise une modulation de la réponse, suivant le poids de l'observation. En suivant une approche Théorie de la Décision, on peut alors évaluer les estimateurs au moyen d'un coût, $L(I_{\Theta_0}(\theta) - \varphi(x))$ et utiliser les notions usuelles de minimaxité et d'admissibilité. Deux coûts seulement sont à envisager :

$$L_1(t) = |t|, \qquad L_2(t) = t^2 \tag{3.1}$$

En effet, le premier choix correspond globalement à l'approche de NEYMAN-PEARSON tandis que le second permet d'évaluer les p-values. De plus, ces deux coûts sont les plus consistants avec une approche bayésienne.

Remarque 1. Les estimateurs considérés dans cette partie seront donc les complémentaires des tests introduits auparavant, puisque, si φ est un test de NEYMAN-PEARSON, $1 - \varphi$ sera un estimateur de $I_{\Theta_0}(\theta)$.

Remarque 2. Ces deux coûts sont symétriques en H_0 et H_1, au sens où un estimateur φ pour H_0 donne un estimateur $1 - \varphi$ pour H_1. La théorie de NEYMAN-PEARSON n'admet pas cette symétrie en général.

Remarque 3. Une approche alternative serait de prendre en compte la distance entre θ et la frontière de Θ_0, afin de moduler le coût en fonction de cette distance. Le problème fondamental lié à cette approche concerne l'utilisation des résultats et implique une reformulation globale des tests d'hypothèses et leurs motivations (voir ROBERT ET CASELLA, 1994).

Les estimateurs de BAYES associés aux deux coûts (3.1) correspondent aux mesures données en 2.2. Sous le coût L_1, l'estimateur de BAYES est donné par

$$\varphi_1^\pi = \begin{cases} 0 & \text{si } \pi_0 \int_{\Theta_0} f(x|\theta)g_0(\theta)d(\theta) < \pi_1 \int_{\Theta_1} f(x|\theta)g_1(\theta), \\ 1 & \text{sinon .} \end{cases} \tag{3.2}$$

Ce résultat illustre la rigidité du coût L_1, puisque même les estimateurs de BAYES ne donnent pas une solution adaptative et ne peuvent prendre que les valeurs 0 et 1. On ne retrouve pas exactement le Bayes factor, car le test de BAYES fait aussi intervenir les probabilités π_0 et π_1 dans la décision.

Les estimateurs de BAYES associés au coût L_2 sont en fait les probabilités *a posteriori*,

$$\begin{aligned} \varphi_2^\pi(x) &= P^\pi(\theta \in \Theta_0|x) \\ &= \frac{\pi_0 \displaystyle\int_{\Theta_0} f(x|\theta)g_0(\theta)d(\theta)}{\pi_0 \displaystyle\int_{\Theta_0} f(x|\theta)g_0(\theta)d(\theta) + \pi_1 \displaystyle\int_{\Theta_1} f(x|\theta)g_1(\theta)d(\theta)}, \end{aligned} \tag{3.3}$$

et ce coût donne donc une réponse bayésienne adaptative, comme il était souhaité. Nous allons voir, de plus, qu'il permet d'évaluer les p-values.

La notion de minimaxité ne s'avère pas être d'une grande utilité dans ce contexte car, pour le coût L_2 –et tout coût strictement convexe–, le test "neutre", $\varphi = 1/2$, est l'*unique estimateur minimax*, donc est admissible. Pour le coût L_1, ce test est toujours minimax mais uniformément dominé par

la p-value, elle-même dominée par le test UMP de seuil $\alpha = 50\%$, admissible d'après le lemme de NEYMAN-PEARSON.
L'examen des estimateurs admissibles pour le coût L_2 nous ramène à la section précédente, car elle renforce les conclusions de BERGER ET SELLKE (1987) et CASELLA ET BERGER (1987), opposant les cas unilatère et bilatère. Dans le cas unilatère, pour le test de $H_0 : \theta \leq \theta_0$ contre $H_1 : \theta > \theta_0$, on a pu exhiber l'admissibilité de la p-value dans les cas suivants –par la méthode de BLYTH (voir BERGER, 1985), à partir des distributions conjuguées–

(a) si $x \sim \mathcal{N}(\theta, 1)$, $p(x) = 1 - \Phi(x - \theta_0)$;

(b) si $x \sim \mathcal{B}(n, \theta)$, $p(x) = P_{\theta_0}(x \geq \theta_0)$;

(c) si $x \sim \mathcal{P}(\theta)$, $p(x) = P_{\theta_0}(x \geq \theta_0)$.

Cette propriété d'admissibilité des p-values est certainement généralisable à des nombreuses familles de distributions. Ainsi le coût L_2 justifie et légitime l'usage des p-values pour les tests unilatères[2]. Au contraire, comme l'impliquaient les résultats de BERGER et SELLKE (1987), les p-values ne sont pas performantes pour les tests bilatères. En effet, on montre, pour les tests de $H_0 : \theta \in [\theta_0, \theta_1]$, qu'un test φ est admissible *si et seulement si il est de Bayes généralisé*, i.e.

$$\varphi_2^\pi(x) = \frac{\displaystyle\int_{\theta_0}^{\theta_1} f(x|\theta) g_0(\theta) d(\theta)}{\displaystyle\int_{\theta_0}^{\theta_1} f(x|\theta) g_0(\theta) d(\theta) + \int_{[\theta_0,\theta_1]^c} f(x|\theta) g_1(\theta) d(\theta)},$$

où g_1 est une mesure sur $[\theta_0, \theta_1]^c$. Ce théorème de classe complète est comparable aux résultats classiques de l'estimation ponctuelle qui montrent que tout estimateur admissible est estimateur de BAYES ou limite de BAYES (voir BERGER [1985] et BROWN [1986]). Caractérisant les estimateurs admissibles, il permet aussi d'éliminer les p-values, qui ne peuvent être égales à un estimateur de BAYES généralisé.

Il faut noter que, dans le cadre de l'estimation ponctuelle, les estimateurs inadmissibles sont parfois conservés lorsqu'ils sont presque admissibles et simples à calculer. C'est par exemple le cas de l'estimateur de JAMES-STEIN tronqué. Une question demeurant en suspend pour ce modèle est donc la "valeur" des p-values, i.e. l'amplitude de l'amélioration apportée par un meilleur estimateur. Les écarts constatés par BERGER et SELLKE (1987) et BERGER et DELAMPADY (1988) tendent à réfuter la "presqu'admissibilité" des p-values mais l'exemple suivant vient tempérer cette impression.

Exemple 1 (*suite*) Dans le cadre normal bilatère, la p-value ne peut être dominée par un estimateur de BAYES (au sens strict).

2. Un problème demeurant en suspend est l'utilité ultime des tests unilatères. En pratique, tout test d'hypothèse peut-il et doit-il se ramener à un test bilatère ?

Cependant, considérant à nouveau le cadre de l'estimation ponctuelle, ce résultat n'implique pas nécessairement des propriétés de quasi-optimalité pour l'estimateur en question. Ainsi STRAWDERMAN (1973) a montré que l'estimateur du maximum de vraisemblance de la moyenne d'une loi normale ne pouvait être dominé par un estimateur de BAYES propre si la dimension était inférieure à 5 (voir BERGER, 1985). Il est néanmoins possible de l'améliorer de manière appréciable à partir de la dimension 3. Les conséquences de l'exemple ci-dessus sont donc encore mal définies et seule l'obtention d'un estimateur dominant la p-value permetterait de régler ce problème.

4. Conclusions

Cette étude a mis en avant les critiques de la théorie de NEYMAN-PEARSON et a proposé le coût L_2 comme outil d'évaluation de leur substitut empirique, les p-values. Etant strictement convexe, ce coût conduit à des estimateurs plus progressifs que les tests $0-1$ de NEYMAN-PEARSON, donc à des évaluateurs plus pertinents de la vraisemblance d'une hypothèse. De plus, ce coût conserve la cohérence avec l'approche bayésienne, puisqu'il redonne les probabilités *a posteriori* comme estimateurs de BAYES. Cette propriété le distingue d'autres alternatives, comme celles intégrant la distance à la frontière entre hypothèse nulle et hypothèse alternative, qui produisent des estimateurs difficilement interprétables.

L'approche L_2 de la théorie des tests conduit également à une séparation entre tests et régions de confiance. Ce divorce est, de notre point de vue, salutaire car la méthode de NEYMAN-PEARSON d'inversion des tests UMP et UMPU est trompeuse, comme l'ont montré HWANG et CASELLA (1982). L'évaluation des régions de confiance, de même que celle des p-values, doit donc être indépendante, une possibilité étant l'approche théorie de la décision développée dans CASELLA *et al.* (1992). Une alternative, plus dans l'esprit du présent article, est celle de l'*estimation du coût*, i.e. pour une région donnée C_x, l'estimation de $I_{C_x}(\theta)$, suivant le coût quadratique

$$(I_{C_x}(\theta) - \gamma(x))^2 .$$

Cette approche a été suivie dans LU et BERGER (1989) pour les régions recentrées en les estimateurs de JAMES-STEIN tronqués (voir aussi CASELLA et ROBERT, 1993). Suivant cette approche, les régions de confiance permettant de développer les estimateurs γ les plus précis et les plus adaptatifs seraient les plus intéressantes.

RÉFÉRENCES

BERGER, J.O (1985) *Statistical Decision Theory and Bayesian Analysis*, Springer, N.Y.

BERGER, J.O and DELAMPADY, M. (1988) "Testing Precise Hypotheses", *Stat. Science,* **2**, 317-52.

BERGER, J.O and SELLKE, T. (1987) "Testing a Point Null Hypothesis : the Irreconcibility of p-values and Evidence", *J.A.S.A.,* **82**, 112-22.

BERGER, J.O and WOLPERT, R. (1988) *The Likelihood Principle,* (2d edition) IMS Monograph Series.

BROWN, L. (1986) *Fundamentals of Statistical Exponential Families*, IMS Monograph Series.

CASELLA, G. and BERGER, R. (1987) "Reconciling Evidence in the One-sided Testing Problem", *J.A.S.A.,* **82**, 106-11.

CASELLA, G., HWANG, J.T. and ROBERT, C.P. (1993) "Loss Functions for Set Estimation", *in Stat. Dec. Theory and Related Topics*, 237-257.

CASELLA, G. and ROBERT, C.P. (1993) "Domination of the Constant Confidence Statement for the Usual Normal Confidence Set", *in Stat. Dec. Theory and Related Topics*, 351-368.

HWANG, J.T. and CASELLA, G. (1982) "Minimax Confidence Sets for the Mean of Multivariate Normal Distribution", *Ann. Stat.* **10**, 868-81.

HWANG, J.T., CASELLA, G., ROBERT, C.P., WELLS, M. and FARELL, R. (1992) "Estimation of Accuracy in Testing.", *Ann. Stat.* **20**, 490-509.

JEFFREYS, H. (1961) *Theory of Probability*, (3d edition) Oxford University Press, London.

KIEFER, J. (1977) "Conditional Confidence Statements and Confidence Estimators", *J.A.S.A.,* **72**, 789-808.

LEHMANN, E. (1986) *Testing Statistical Hypotheses*, J. Wiley, N.Y.

LU, K. and BERGER, J.O. (1989) "Estimated Confidence Procedures for Multidimensional Normal Means", *J.S.P.I.,* **23**, 1-20.

ROBINSON, G. (1979) "Conditional Properties of Statistical Procedures", *Ann. Stat.* **7**, 742-55.

ROBERT, C.P. (1992) *L'Analyse Statistique Bayésienne*, Economica, Paris.

ROBERT C.P. and CASELLA, G. (1994) "Distance Penalized Losses for Testing and Confidence Set Evaluation", *Test* **3**, 167-1982.

ROBERT, C.P. and SALEH, A. (1989) "Recentered Confidence Sets : a Review", Tech. Report, L.S.T.A., Université Paris VI.