# recent developments in statistics

j.r. barra, f. brodeau, g. romier
and b. van cutsem
editors

north-holland

# *RECENT DEVELOPMENTS IN STATISTICS*

Proceedings of the European Meeting of Statisticians
Grenoble, 6-11 September, 1976

Edited by

J. R. BARRA
F. BRODEAU
G ROMIER
B. VAN CUTSEM
*University of Grenoble*

1977

NORTH-HOLLAND PUBLISHING COMPANY
AMSTERDAM • NEW YORK • OXFORD

North-Holland ISBN: 0 7204 0751 6

Published by:
NORTH-HOLLAND PUBLISHING COMPANY
AMSTERDAM • OXFORD • NEW YORK

Distributors for the U.S.A. and Canada:
Elsevier/North-Holland, Inc.
52 Vanderbilt Avenue
New York, N.Y. 10017

PRINTED IN THE NETHERLANDS

# FOREWORD

In 1976, the European Meeting of Statisticians was held at the University of Grenoble (France) from September 6-10, and was sponsored by the Institute of Mathematical Statistics and the Bernoulli Society for Mathematical Statistics and Probability. This conference brought together more than 600 statisticians, from many countries, who came here to listen to about 150 talks.

The subjects dealt with were those of classical statistics : theory of estimation and testing hypotheses, non-parametric problems, statistical inference in the classical case and for stochastic processes, experimental designs, multivariate analysis, stochastic approximation, asymptotic expansions, time series, linear models, etc. Moreover, numerous applications of statistics were presented concerning the medical, social and teaching sciences, demography, genetics, physics, etc. Finally, you will notice the great number of papers devoted to data analysis and to the relations between computer science and statistics. Nearly all the invited talks, as well as the most representative ones among the contributed papers, have been combined in this volume and arranged around those principal themes, hoping that this would facilitate the use of this book.

On the occasion of publishing these Proceedings, we would like to thank particularly the Organizing Committee of the European Meetings of Statisticians who agreed to the 1976 Meeting being held at our University, as well as the French authorities who contributed their financial support. Besides, we have the pleasure of saluting the birth of the Bernoulli Society, as this Meeting was the first to be held under its sponsorship. We also feel greatly obliged to Professor H. CRAMER who honoured our Meeting with his presence, as well as to the professors GANI, DURBIN, VAN ZWET and BARNDORFF-NIELSEN for their contributions to the organization of the Meeting and more specifically to the scientific program. Last but not least our thanks are due to the North-Holland Publishing Company, to whom we owe the present publication of the Proceedings.

The Editors - February 1977

CONTENTS

Recent Developments in Statistics
J.R.Barra et al., editors


# L'HERITAGE DE BERNOULLI

Grenoble 6 Septembre 1976

H. CRAMER, Suède

Devant cette assemblée, qui s'est réunie sous le patronage de la Bernoulli Society et de l'Institute of Mathematical Statistics, il m'a semblé naturel de dire quelques mots sur L'HERITAGE DE BERNOULLI.

Il va sans dire que c'est de Jacques Bernoulli que je vais parler, et de ce qu'il nous a laissé en notre qualité de statisticiens et probabilistes. Je vais donc passer sous silence les contributions de quelques autres membres de cette famille célèbre, ainsi que les travaux de Jacques Bernoulli lui-même appartenant à d'autres domaines scientifiques.

Quand Jacques Bernoulli mourut en 1705, à l'âge de seulement 51 ans, son grand traité de la théorie des probabilités, auquel il avait travaillé depuis une vingtaine d'années, restait toujours inachevé. Son manuscrit, écrit en latin, fut publié huit ans après sa mort, tel qu'il l'avait laissé, sous le titre ARS CONJECTANDI. Autant que je sache, il n'existe qu'une seule traduction complète de cet ouvrage en une langue moderne, à savoir une édition allemande de l'année 1899, qui d'ailleurs semble être assez difficilement accessible.

L'Ars conjectandi consiste en quatre parties, dont les trois premières n'ont guère un grand intérêt pour nous, même si elles contiennent une étude générale de l'analyse combinatoire, et son application à un grand nombre de questions concernant les jeux de hasard. C'est la quatrième partie, intitulée "Applications de la théorie précédente aux questions sociales, morales et économiques", qui contient les idées fondamentalement nouvelles de l'auteur. Malheureusement c'est aussi cette partie qu'il a due laisser inachevée. Telle que nous l'avons, elle contient des principes pour ces applications, mais pas les applications elle-mêmes.

Je veux essayer de rappeler très brièvement quelques parties essentielles de son raisonnement, en me servant autant que possible de ses propres mots, exprimés dans une terminologie un peu modernisée.

En discutant le caractère de nécessité des événements futurs, il prend une attitude parfaitement déterministe. Le temps qu'il fera demain est parfaitement déterminé par l'état présent de l'atmosphère, de même que les éclipses qui auront lieu l'année prochaine sont déterminées par les positions actuelles des corps célestes. Le résultat d'un coup de dé est déterminé par l'état de mouvement du dé et par sa distance de la table au moment où il sort de la main.

Et pourtant, Bernoulli fait observer, nous tenons à regarder le temps du demain et le résultat d'un coup de dé comme des événements fortuits, dépendant du hasard, tandis que les éclipses sont des événements certains, qu'on peut prédire par des calculs. C'est dans notre ignorance qu'il trouve la raison de cette contradiction apparente : l'état actuel de l'atmosphère, qui en fait est exactement déterminé, ne nous est pas connu avec une précision suffisante pour nous permettre de prédire le temps du demain avec certitude. Et la même situation se présente pour le dé. - Ce n'est pas sans intérêt de trouver de telles pensées si clairement exprimées déjà à cette époque.

Pour un événement de ce caractère fortuit, Bernoulli admet qu'il existe une probabilité, qu'il envisage comme un degré de certitude, ayant une valeur entre zéro et un.

Un événement qui a une probabilité très voisine de l'unité doit être regardé comme moralement certain, tandis qu'une probabilité très petite correspond à un événement moralement impossible. Un détail amusant dans ce contexte est que Bernoulli pense que les gouvernements devraient fixer une limite supérieure officielle pour les probabilités des événements qu'il serait légitime de regarder comme pratiquement impossibles. - Peut-être ça serait-il quelque chose pour nos politiciens contemporains ?

L'art d'estimer les probabilités des événements est ce qu'il appelle Ars conjectandi sive stochastice. Je crois que c'est ici la première fois qu'on rencontre dans cette matière le mot stochastique, si bien connu aujourd'hui.

Il suppose que, pour un événement fortuit, il existe toujours un ensemble de cas possibles, dont les uns sont favorables, les autres défavorables pour que l'événement ait lieu. Si l'on connaît le nombre et les poids relatifs de tous ces cas, on peut trouver la probabilité de l'événement. Et il rappelle que c'est en fait cette méthode qu'on applique dans le domaine des simples jeux de hasard.

Mais, dit Bernoulli, comment serait-il possible de connaître toutes les maladies qui pourraient causer la mort d'un individu donné, ou toutes les variations dans l'atmosphère qui vont déterminer l'évolution du temps ?
Puisque ces événements - et bien d'autres analogues - dépendent de causes infiniment nombreuses et infiniment compliquées, il lui semble impossible de déterminer leurs probabilités par une énumération de tous les cas possibles.

Mais c'est sur ce point qu'une idée nouvelle et féconde fait son entrée dans son raisonnement. Quand une probabilité inconnue ne peut pas être déterminée a priori, par une énumération des cas possibles, il dit qu'on peut souvent la trouver au moins approximativement a posteriori, c'est-à-dire en faisant usage des observations d'un grand nombre d'événements analogues. Si, par exemple, on a observé que, parmi 400 hommes du même âge et constitution que Titius, 300 sont décédés avant dix ans, on peut conclure que le nombre des cas où Titius va mourir avant dix ans est à peu près trois fois plus grand que le nombre des cas où il survivra ; sa probabilité de mourir est donc approximativement 3/4. De même si l'on a observé le temps pendant un grand nombre de jours, on peut former des conclusions p. ex. sur la probabilité d'avoir une pluie demain, et si l'on a noté les résultats d'une longue série de parties du jeu de paume entre deux joueurs Pierre et Paul, on pourra estimer la probabilité que Pierre va gagner la partie suivante.

Il s'agit donc ici d'une méthode générale d'employer des observations statistiques pour déterminer les probabilités des événements. Bernoulli fait observer que la méthode n'est pas nouvelle - en fait, il dit, c'est ainsi qu'on agit dans la vie ordinaire. Sans doute cela est vrai, mais c'est à lui qu'on doit l'introduction systématique des observations statistiques comme un instrument scientifique, intimement lié à la théorie des probabilités.

Cependant il dit que, pour donner une confirmation scientifique de cette méthode déjà appliquée par tant de gens pratiques, il a trouvé nécessaire de se poser d'abord la question fondamentale suivante : Est-il possible, en faisant un assez grand nombre d'observations, d'être moralement certain d'obtenir une approximation aussi bonne que l'on veut de la probabilité cherchée ?

C'est la réponse affirmative à cette question qui constitue le célèbre théorème qu'il annonce comme le résultat principal de ses études.

"Voilà donc", il dit, "ce problème que je me suis déterminé de publier ici, après l'avoir retenu pendant vingt années. Sa nouveauté, sa grande utilité et sa difficulté sont destinées à donner une importance plus élevée à toute cette doctrine."

Et puis il procède à énoncer et à démontrer ce théorème classique qui aujourd'hui fait partie de tous nos traités de la théorie des probabilités.

Il faut admettre que Bernoulli n'exagère point en parlant de la nouveauté de son théorème. En effet, c'est ici la première fois qu'on envisage une distribution de probabilité comme un objet mathématique dépendant d'un paramètre, et qu'on étudie son comportement lorsque ce paramètre tend vers l'infini.

Dans cet ouvrage posthume de Jacques Bernoulli, on trouve donc à la fois une première indication d'un programme pour la statistique mathématique, et la première application de l'analyse infinitésimale à l'étude d'une distribution de probabilité. - Ce n'est pas là un pauvre héritage !

Eh bien : qu'est-ce qu'on a fait - qu'est-ce que nous avons fait - pour développer cet héritage ? Jetons un coup d'oeil rapide, et nécessairement très incomplet, sur ce qui s'est passé en notre domaine après la publication de l'Ars conjectandi.

Pendant tout ce temps de plus de deux cent soixante ans, il y a eu une série continue de travaux qui ont pris leur point de départ directement du théorème original de Bernoulli, tout en le précisant et le généralisant dans une multitude de directions différentes. Ces travaux nous ont donné ce qu'on appelle les théorèmes-limites de la théorie des probabilités.

La première contribution importante à cette série est dûe à De Moivre. Bernoulli avait démontré son théorème par des moyens tout à fait élémentaires, en étudiant la croissance et la décroissance des termes du développement binomial. De Moivre, ayant à sa disposition la formule asymptotique de Stirling, pouvait montrer que la probabilité considérée par Bernoulli est, à la limite, donnée par la distribution normale qui, plus tard, devrait être connue sous les noms de Gauss et de Laplace.

On sait bien que le théorème original de Bernoulli, ainsi que sa forme précisée dûe à De Moivre, peut s'exprimer comme une propriété d'une somme de variables aléatoires indépendantes d'une structure particulièrement simple.
Laplace, dans son grand ouvrage "Théorie analytique des probabilités", a étudié une somme analogue, formée de variables aléatoires beaucoup plus générales, mais toujours mutuellement indépendantes. Pour une telle somme dûment normée, il a énoncé le théorème qu'elle a la même distribution limite exponentielle qui avait été rencontrée par De Moivre dans le simple cas de Bernoulli.

C'est ce théorème pour qui le nom "théorème-limite central de la théorie des probabilités" a été proposé par Georges Polya. C'est un théorème d'une importance capitale, qui contient les résultats antérieurs de Bernoulli et De Moivre comme des cas très particuliers. Mais malheureusement la démonstration donnée par Laplace à son grand théorème n'était point satisfaisante.
Bien que plusieurs des mathématiciens du XIXème siècle se soient efforcés de combler la lacune, c'était seulement Tchebychev qui savait indiquer un chemin qui devait conduire au but. Mais une démonstration vraiment complète et rigoureuse ne fut donnée que dans les premières années de notre siècle, par Liapounov.

La série de ces théorèmes-limites dont le premier était le théorème de Bernoulli s'est prolongée jusqu'à nos jours. Les conditions données par Liapounov pour la validité du théorème central furent généralisées par Lindeberg, et enfin Feller a donné des conditions à la fois nécessaires et suffisantes. On a aussi étudié le passage à la limite de la distribution considérée par ces auteurs plus en détail, en obtenant des développements asymptotiques, dont la distribution normale forme le premier terme, et en considérant ce qu'on appelle les "grandes déviations" qui échappent au théorème de Laplace. Lévy et Khintchine ont apporté des généralisations importantes, en étudiant des cas où il y a une distribution limité appartenant à une classe plus générale que la loi normale, à savoir la classe des distributions infiniment divisibles. D'intéressantes recherches, qui ne sont pas encore terminées, ont été consacrées à l'étude des propriétés limites de sommes de variables aléatoires interdépendantes. Et finalement, toutes ces recherches ont été généralisées aux variables aléatoires multi-dimensionnelles.

Pour tous ces travaux cherchant à généraliser et à approfondir le théorème original de Bernoulli, il était nécessaire d'avoirs recours à des méthodes analytiques nouvelles et puissantes. Lagrange, Laplace et Cauchy avaient employé une première forme de l'outil analytique connu aujourd'hui sous le nom de "fonctions caractéristiques". C'était aussi de cette méthode que s'était servi Liapounov pour sa démonstration du théorème limite central. Mais une théorie détaillée des fonctions caractéristiques ne fut donnée que plus tard, dans les travaux de Paul Lévy, qui a mis en évidence leur haute valeur. Son théorème de continuité pour les fonctions caractéristiques a eu une importance fondamentale pour les recherches sur les théorèmes-limites de la théorie des probabilités. - Une autre méthode importante pour l'étude de ces problèmes est la méthode des moments, employée par exemple par Tchebychev. D'ailleurs il y a des relations intéressantes entre ces deux méthodes.

Les vingt années de paix entre les deux guerres mondiales ont été pour nos domaines une période d'un développement énorme. D'une part, les statisticiens anglais et américains - Fisher, Neyman, Pearson et bien d'autres - ont crée pendant cette époque une nouvelle science, la statistique mathématique, qui s'occupe des relations entre les observations statistiques et les modèles probabilistiques dont on veut se servir pour l'analyse des observations, et pour des prédictions. Ils ont introduit la théorie de l'estimation statistique, la théorie des tests pour des hypothèses statistiques, et des méthodes nouvelles d'échantillonnage.

L'objet principal de cette science est de trouver des méthodes pour déduire des conclusions valides en partant de données statistiques. C'est là précisément le problème que s'était posé Bernoulli dans le cas le plus élémentaire. Nous savons tous combien le développement vigoureux de cette science a continuellement accéléré jusqu'à l'heure actuelle, et qu'elle a trouvé un nombre toujours croissant de domaines d'application.

Il y a ici, comme on sait, deux points de vue opposés. On peut, comme Bernoulli, considérer une probabilité - ou quelque autre paramètre associé à une distribution - comme une constante fixée, quoique inconnue, et se proposer de trouver quelque information sur sa valeur au moyen des observations statistiques. Or, on peut aussi la regarder comme une variable aléatoire, assujettie à une distribution <u>a priori</u> et appliquer un raisonnement du type de Bayes. C'est pourtant toujours la même théorie des probabilités qui fournit l'instrument mathématique dont on se sert. Et sans doute il y a, parmi le nombre immense des champs d'application, pour chacune des deux manières de voir, des cas où elle est convenable.

D'autre part, pendant ces mêmes vingt années entre les guerres, les mathématiciens ont de leur côté développé la théorie des probabilités comme une branche des mathématiques pures. J'ai déjà mentionné quelques-uns des travaux se rattachant aux théorèmes-limites pour les sommes de variables aléatoires. Toute cette ligne de recherche peut être considérée comme un développement naturel du théorème original de Bernoulli. Mais les mathématiciens ont introduit aussi des idées d'un ordre différent, indiquant une ère nouvelle où les relations avec l'héritage de Bernoulli, bien que sans doute encore présentes, n'auront plus le même caractère d'évidence. La théorie des probabilités classique, telle qu'on la connaissait il y a un demi-siècle, a été complètement transformée par l'introduction de ces nouvelles idées.

D'abord ce n'était qu'un développement tout naturel, quand certaines applications statistiques ont conduit à l'étude de suites de variables aléatoires plus générales que de simples sommes, par exemple des suites obtenues en observant les valeurs d'une variable aléatoire à des points de temps équidistants. Encore un pas en avant s'imposait quand, dans les applications, on rencontrait des fonctions du temps telles que leur valeur à tout instant semblait dépendre plus ou moins du hasard - des fonctions aléatoires, comme on dit maintenant. Un exemple célèbre est le mouvement Brownien, étudié en 1906 par Einstein. Se basant sur des arguments physiques, il avait obtenu certains résultats, mais c'était évident que les méthodes de la théorie classique des probabilités ne suffisaient pas pour un traitement rigoureux d'un tel problème.

C'est Norbert Wiener qui, dans une note publiée en 1923, a d'abord attaqué cette difficulté avec succès. Il y a introduit une distribution de probabilité dans un espace fonctionnel, ce qui était alors une innovation remarquable. De cette manière il a pu donner un modèle mathématique du mouvement Brownien où l'on se sert de fonctions aléatoires du temps qui sont continues mais n'ont pas une dérivée. C'était la première fois qu'on avait donné une théorie rigoureuse d'un processus stochastique à temps continu.

C'était encore seulement un cas très particulier de cette idée nouvelle, mais quelques années plus tard Kolmogorov a publié ses grands travaux, où toutes ces questions sont traitées d'une manière générale et profonde.
En montrant l'identité formelle de la probabilité avec la notion purement mathématique de mesure, il a donné un fondement nouveau et un nouvel élan aux recherches dans ce domaine. En particulier, la notion d'une distribution de probabilité se présente chez lui sous une forme infiniment plus générale qu'auparavant, comme une distribution dans un espace tout à fait arbitraire. La distribution dans le "différentiel space" considérée par Wiener apparaît comme un cas de ce qu'on appelle maintenant un processus de Markov, c'est-à-dire un processus tel que la loi de probabilité de son développement futur est bien déterminée par sa valeur au moment actuel.
L'étude de cette classe de processus a été commencée par Kolmogorov en 1931. Il a démontré que leurs distributions satisfont à des équations fonctionnelles qui, sous certaines conditions de continuité, se réduisent à des équations différentielles. On sait bien qu'il y a maintenant une littérature très étendue sur les processus de Markov.

Une autre classe importante de processus stochastiques, les processus stationnaires, a été introduite en 1934 par Khintchine. Pour ces processus on suppose que leurs lois de probabilité satisfont à une condition d'invariance dans le temps. L'entière préhistoire est alors relevante pour la prédiction de leur développement futur. Leur théorie comprend une analyse spectrale qui a conduit à des résultats d'une haute importance pour un grand nombre d'applications.

La théorie des processus stochastiques, dont je n'ai pu rappeller ici que ces deux exemples, constitue aujourd'hui une des parties les plus importantes et les plus caractéristiques de la théorie des probabilités. Elle emploie les méthodes les plus avancées de la mathématique moderne, mais une partie de ses racines se trouvent dans les théorèmes-limites classiques donnés par

Bernoulli et par ses successeurs.

Nous avons donc devant nos yeux ces deux grandes lignes de recherche scientifique : la ligne statistique et la ligne probabilistique. Un parcours du programme de ce congrès met en évidence combien le rapprochement des points de vue et des méthodes de travail de ces deux lignes a été utile pour le progrès de la science. Leurs différentes manières d'attaquer les problèmes se sont mutuellement fertilisées. La pénétration des méthodes statistiques au moyen de l'analyse mathématique a permis la construction d'une théorie statistique logiquement rigoureuse, au même temps que les applications statistiques ont posé de nouveaux problèmes aux mathématiciens, et ont suggéré des méthodes nouvelles pour leur résolution.

L'histoire de chacune de ces deux lignes peut être suivie à travers des siècles passés jusqu'à ces premières indications dans l'Ars conjectandi que j'ai rappelées en commençant. C'est en développant ces indications, et en y ajoutant continuellement des idées nouvelles, qu'on a pu édifier nos branches de la sciences contemporaine, basées sur l'héritage de Bernoulli.

## PART ONE: Invited Papers

Recent Developments in Statistics
J.R.Barra et al., editors

Invited Paper

# ASYMPTOTIC EXPANSIONS FOR BAYES PROCEDURES

Helmut Strasser [1)]

## 1. Introduction

### a) The general case

Let $(\Omega,\mathcal{A})$ be a measurable space and $P_\vartheta|\mathcal{A}, \vartheta\in\Theta\subseteq\mathbb{R}$, a family of probability measures. Assume that the family $\{P_\vartheta\}_{\vartheta\in\Theta}$ is dominated by a $\sigma$-finite measure $\mu|\mathcal{A}$ and let $h_\vartheta := \frac{dP_\vartheta}{d\mu}, \vartheta\in\Theta$. $\mathcal{B}_\Theta$ denotes the Borel-$\sigma$-algebra of $\Theta$.

Many statistical procedures are based on the likelihood functions

$$\vartheta \mapsto \prod_{i=1}^{n} h_\vartheta(\omega_i), \vartheta\in\Theta, \underline{\omega}=(\omega_1,\ldots,\omega_n)\in\Omega^n,$$

(e.g. maximum likelihood estimates). Bayes procedures are based on averages of likelihood functions, i.e. on set functions of the form

$$B \mapsto \int_B \prod_{i=1}^{n} h_\vartheta(\omega_i)p(\vartheta)d\vartheta, \ B\in\mathcal{B}_\Theta, \underline{\omega}\in\Omega^n,$$

where $p:\Theta\to\mathbb{R}^+$ is a weight function. If such a set function is a finite measure then the normalized set function is called "posterior" distribution with respect to the (not necessarily finite) "prior" measure defined by p.

The classical theory of asymptotic optimality ("first order efficiency") shows that among optimum procedures there are both Bayes procedures and other ones which seemingly are of non-Bayesian type,

---

1) H.Strasser, Mathematisches Institut, Justus-Liebig-Universität, Arndtstraße 2, D-6300 Gießen, Federal Republic of Germany.