BOOK REVIEWS

Mathematical Methods of Statistics. *Harald Cramér.* Uppsala, Sweden: Almqvist and Wiksell, 1945. pp. xvi, 575. (Princeton, N. J.: Princeton University Press, 1946. $6.00)

Reviewed by Will Feller

Cornell University

This book represents a contribution of a novel kind to the statistical literature and will render valuable services both as textbook and reference book. Of its three parts the first one (134 pages) is entitled *Mathematical Introduction* and develops the necessary formal mathematical tools. The second part (186 pages) is devoted to *Random Variables and Probability Distributions*, that is to say, to a chapter of the modern theory of probability. The third, and main, part of the book (some 233 pages) is entitled *Statistical Inference*. Ordinarily these three topics would require consultation of three or more books, and these would rarely be found on the same shelf. However, the masterly exposition succeeds in creating the impression of natural unity and harmony. The ideas are developed with elegance and apparent ease as if the line of presentation followed a well explored path. The uninitiated will not notice how unconventional the treatment is and how the very selection of topics depends on the author's scientific personality.

It is hardly necessary to point out that Cramér's book fills an urgent need. The emergence of statistical theory and methodology as an exact science, firmly grounded in mathematical probability, is only of recent date. Its rapid development went hand in hand with an extraordinary increase of the number and importance of its various applications. Under such circumstances there was naturally little time for an exposition of the theoretical foundations and ramifications. Modern statistical inference has its roots in the classical limit theorems of probability. Now classical probability used to consist of a bewildering collection of special and mutually uncorrelated problems; unified guiding principles and methods are a rather new development and have not yet found expression in the textbook literature. The original investigations are usually written in an exceedingly abstract language and the existing close ties to applications are not apparent. Consequently, there is no easy access either to probability or statistics and it is often difficult to establish whether, or to what extent, various assertions have actually been proved. The present book therefore closes a serious gap in the literature and will greatly facilitate both teaching and research.

Of the 12 chapters of the *Mathematical Introduction* 9 are devoted to the theory of measure and integration. The antiquated theory of the so-called Riemann integral (kept alive by elementary textbooks) considered only point functions $y = f(x)$, where the independent variable is a point. The temperature at a given point or the velocity at a given moment are typical examples. Many mathematical considerations simplify greatly if from the very beginning also set functions $y = F(A)$ are introduced, where the independent variable is a set. Typical examples are mass in mechanics, the amount of heat or of electricity, area or wealth of a geographic region, and the probability of events (i.e. sets in sample space). The Lebesgue-Stieltjes theory frees the concept of integral from artificial devices and reduces it to the natural notion of mean values with respect to set functions. In a simile, believed to be due to Lebesgue, the Riemann integral corresponds to the procedure of a grocer who computes the day's receipts by actually adding the several amounts in the order as they had come in. The Lebesgue procedure imitates the more intelligent grocer who orders his cash in piles of notes and coins of equal denomination and counts them. The analogy with the customary procedure of computing mathematical expectation is clear. The Lebesgue-Stieltjes integral is conceptually simpler than the Riemann integral and can be presented in as simple a way with rigor adequate for elementary textbooks. It has become an indispensable tool in probability, statistics, physics, and other applied fields. Since it has, unfortunately, not found its way into calculus textbooks, physicists are compelled to use the less flexible notion of the Dirac δ-function, and the formal mathematical apparatus in general becomes unnecessarily clumsy. It is a curious anomaly that so many calculus textbooks profess to be written with a view to applications and yet completely disregard the most obvious practical needs and that the teaching of practical mathematics should remain uninfluenced by the great developments of the last fifty years.

In such circumstances the chapters on integration will be particularly welcome to statisticians as probably the only place in the literature where they will find easy access to the theory. Of course, this exposition leads far beyond what the average statistician will require under ordinary circumstances and beyond the necessary prerequisites of the main body of the book. Of the 88 pages roughly half can be omitted at first reading in accordance with detailed instructions given in the Preface. The remaining half will form a valuable reference book for theorems and tools used occasionally in connection with more delicate parts of statistical theory. The mathematical introduction contains also a chapter on Fourier integrals (characteristic functions), one on matrices and quadratic forms, and finally miscellaneous complements such as orthogonal polynomials, Euler's summation formula, beta and gamma functions, etc.

The title to the second part, *Random Variables and Probability Distributions*, is the same as that of the author's well-known Cambridge Tract of 1937. Both start with a discussion of the foundations along axiomatic lines. The new treatment does not differ essentially from the old one, but some changes are introduced which are regrettable in the reviewer's opinion (in particular axiom 3). Otherwise there is practically no overlap between the two expositions. The 1937 booklet devoted much space to the asymptotic expansions connected with the central limit theorem which are due to the author himself. This topic is not touched upon in the present book. This is a judicious procedure since the 1937-booklet is generally accessible (although at present sold out). Instead we now find a detailed study of some univariate distributions such as χ^2, Student's t, Fisher's z, the Pearson system, etc., none of which were mentioned in the Cambridge tract. Similarly, there is now a section on correlation and regression, and the normal distributions in several variables. The theory of probability is developed only to the extent of the formal theory of distribution functions. This implies that even so important a notion as stochastic convergence is treated only summarily while the strong law of large numbers falls completely outside the framework of the book. This is regrettable inasmuch as the strong law is of greater importance than the classical weak law (whose fame rests essentially on a classical misunderstanding). It should be mentioned that this second part of the book contains some 39 well chosen illustrative exercises the solution of which is left to the reader.

In the main part of the book, entitled *Statistical Inference*, the outer form changes inasmuch as the text there is accompanied by numerous practical examples. However, the exposition remains mathematical in nature and the main emphasis rests on exact formulations; much attention is paid to the establishment of the precise conditions of validity of the individual theorems, their logical interrelations and their connections with general probability. The expert will find many minor and major improvements in formulations and proofs. They are too numerous to be listed here. Suffice it to point out, as a typical example, the theorem on pp. 426–27 concerning the limiting form of the χ^2 distribution with estimated parameters; this theorem appears to be more general than usually stated and also the proof seems to be novel. The topics treated in the statistical part of the book will be seen from the following list of titles to the chapters. 25. Preliminary Notions on Sampling. 26. Statistical Inference (general orientation). 27. Characteristics of Sampling Distributions (moments, semi-invariants, corrections for grouping, etc.). 28. Asymptotic Properties of Sampling Distributions (moments, extreme values, range, etc.). 29. Exact Sampling Distributions (degrees of freedom, Student, Fisher, correlation and regression coefficients, partial and multiple correlations, generalized variance, etc.). 30. Tests of Goodness of Fit and Allied Tests (treating mostly applications of χ^2). 31. Tests of Significance for Parameters. 32. Classification of Estimates (sufficient, efficient and asymptotically efficient estimates; minimum variance, etc.). 33. Methods of Estimation (method of moments, maximum likelihood, χ^2-minimum methods). 34. Confidence Regions. 35. General Theory of Testing Statistical Hypotheses. 36. Analysis of Variance. 37. Some Regression Problems. There follow tables of the normal distribution, the χ^2 and the t-distributions, and a long list of references.

If an expression of wishes for a second edition were permitted, most statisticians would probably give first choice to non-parametric and sequential tests. It is needless to point out that the latter became public only after completion of the Swedish edition of the present book

Even this short account will show the extremely wide range of topics and theories covered in the book, from abstract integration to randomized experiments. They are all presented with uniform lucidity. The exposition throughout is formal, and yet inspiring, rigorous and yet never pedantic. It will serve as an example worthy of imitation and is an achievement on which the author deserves our sincere congratulations.

The Advanced Theory of Statistics. Vols. I and II. *Maurice G. Kendall.* London: C. Griffin and Co., Ltd. Vol. I. Second ed. revised, 1945; pp. xii, 457, 50 shillings. Vol. II. 1946; pp. viii, 521; 42 shillings.

REVIEWED BY M. S. BARTLETT

Cambridge University and *The University of North Carolina*

With the recent appearance of the second volume, it is now possible to review as one work this comprehensive treatise. To quote the author's opening remarks to the Preface to Volume I: "The need for a thorough exposition of the theory of statistics has been repeatedly emphasized in recent years. The object of this book is to develop a systematic treatment of that theory as it exists at the present time." An outline of the contents, which in the two volumes make up just on a thousand pages, will indicate that this formidable task has been squarely faced by the author, who, when a tentative co-operative venture of writing such a treatise was upset by the outbreak of the war, continued alone with the project.

Volume I contains sixteen chapters. The first six introduce the concept of frequency distributions via observational data on groups and aggregates, and their mathematical representation (Ch. 1), measures of location and dispersion (Ch. 2) and moments and cumulants in general (Ch. 3), characteristic functions (Ch. 4), and ending with a description of the standard distribution functions, such as the binomial, Poisson, hypergeometric and normal distributions, and the Pearson and Gram-Charlier systems. The next section opens with probability (Ch. 7) and proceeds to sampling theory (Chs. 8–11), including a chapter (Ch. 10) on exact sampling distributions, many of the standard sampling distributions being used in this chapter to illustrate the mathematical methods available for obtaining sampling distributions. Chapter 11 deals with the general sampling theory of cumulants, including a useful reference list of formulae and a demonstration, due to the author, of the validity of Fisher's combinatorial rules for obtaining these formulas. The section concludes with a chapter on the Chi-square distribution and some of its applications. The last four chapters of Volume I deal with association and contingency, correlation, including partial and multiple correlation, and rank correlation; this last chapter being a comprehensive treatment including comparatively recent results of the author.

It will be convenient to list also the contents of Volume II before any critical comment on either volume. The first section of the second volume comprises four chapters on the theory of estimation, including a derivation of the properties of the maximum likelihood estimate (Ch. 17) and separate chapters on Fisher's theory of fiducial probability and Neyman's theory of confidence intervals. The