BOOK REVIEWS

Mathematical methods of statistics. By H. Cramér. (Princeton Mathematical Series, no. 9.) Uppsala, Almqvist and Wiksells, 1945; Princeton University Press, 1946. 16+575 pp. $6.00.

This book has long been needed, and its effect on the future development of mathematical statistics in both teaching and research will be sharp and lasting. The achievement of the author is to present the first treatise on statistics in which the mathematical developments are carried through according to standards of rigor comparable to those now customary in pure mathematics and excelling those in many fields of applied mathematics. He has accomplished this by integrating modern statistical concepts developed by the English and American schools of statisticians, with the mathematical methods of the modern work in probability of the French and Russians. While uniting these two streams he has added important contributions of his own.

The book is divided into three parts. Part I, called *Mathematical introduction*, consists mainly of an exposition of measure and integration, but contains also needed parts of matrix theory, theory of Fourier integrals, and some other topics. Part II, called *Random variables and probability distributions*, develops general concepts associated with random variables and applies them to some specific distributions such as the binomial, normal, chi-square, and so on. Part III, called *Statistical inference*, deals also with asymptotic and exact sampling distributions. Actually only a little over a fourth of the book deals specifically with statistical testing and estimation, and here the treatment is not exhaustive but limited to fundamentals. Much recent work of a fundamental character in statistical inference is omitted because the author's contact with foreign publications was curtailed by the war; for example, there is no reference to any of Wald's work. The main value of the book will be to supply mathematical background for those working in advanced mathematical statistics.

In his preface the author states that he has tried to make his book accessible to the reader with "a good working knowledge of the elements of the differential and integral calculus, algebra, and analytic geometry." However, a student not previously familiar with ϵ-δ analysis would find the book too difficult. On the other hand, the level of difficulty is much below that of the author's Cambridge Tract. The reviewer would guess that American students beginning graduate mathematics should be able to handle the book. At its chosen level of difficulty it is a masterwork of exposition. The student is led with skill and charm from the familiar to the new: For example, beginning on the real line, Lebesgue measure is introduced as a natural generalization of the notion of length of an interval. Using upper and lower Darboux sums, the Lebesgue integral of a bounded function on a set of finite measure is revealed as a generalization of the Riemann integral, the allowable subdivisions of the set in the Darboux sums being now less restricted. With a fixed function for the integrand and a variable set, the Lebesgue integral furnishes a concrete example of a completely additive set function, and this leads to the consideration of more general measures called "P-measures." Next a simple modification of the Lebesgue integral, in which Lebesgue measure is replaced by P-measure, yields the Lebesgue-Stieltjes integral. With this one-dimensional development behind him the student is then prepared for a k-dimensional development using fewer steps, namely: Lebesgue measure, P-measure, Lebesgue-Stieltjes integrals.

Among subjects usually developed in a mathematically questionable way but here presented with rigor we may mention (1) the asymptotic normality of large classes of statistics, (2) the asymptotic distribution of Karl Pearson's chi-square statistic for goodness of fit in cases where parameters of the theoretical distribution are estimated from the sample, (3) the optimum properties of R. A. Fisher's maximum likelihood estimates. All of these developments are enriched by original contributions of the author; especially exciting is the new small sample theory of efficiency of estimates created by him in relation to (3).

One hesitates to make any adverse criticisms of this book for fear of giving a wrong impression of the balance of good and bad: the weight of the mathematically praiseworthy is overwhelming, while the few features which are possibly objectionable are mainly so for nonmathematical and subjective reasons based on one's statistical taste and philosophy. Of the half dozen such features which troubled the reviewer two may be worth mentioning. A correct exposition is given of the meaning of Neyman's confidence intervals, in which the interval is a random interval such that the probability that it cover the unknown parameter value θ is a fixed preassigned number α, independent of the true value of θ. But after this the author gives a new twist to the method. A fixed interval (a, b) is preassigned. A "confidence coefficient" α for the statement $a \leqq \theta \leqq b$ is then calculated from the sample, so that α is a random variable. While there seem to be no definitely wrong statements made, nevertheless the result may be confusion of the reader as to the meaning of confidence intervals.

The simple frequency interpretation is lost: Thus with Neyman's confidence intervals if the statistical assumptions were satisfied in the applications, the user would know that of all statements he makes with confidence coefficients $\alpha \geqq \alpha_0$, the fraction of these statements which are correct will be not less than α_0 "in the long run," regardless of what the true values of the parameters happened to be in the situations encountered. The user of the Cramér "confidence intervals" has no such assurance: It becomes obvious that in a sequence of confidence statements *all* may be wrong, by thinking of a sequence of experiments in which θ is fixed and outside the interval (a, b).

Another criticism, which may be of a controversial nature, is that the book shows insufficient concern with the question whether data generated in a given way may safely be assumed to be results of trials on the same random variable, or in Shewhart's language, whether the underlying chance cause system is constant. This manifests itself in several ways: (1) Among the illustrations of "random experiments" are some, such as commodity prices, or quality characteristics of factory output, whose "randomness" is highly questionable. (2) In the first description given of statistical tests, the order of the observations is disregarded, that is, it is assumed that the critical region of the test is always a part of the sample space completely symmetrical in all coordinates. (3) No tests of randomness are mentioned.

Clearly an oversight is the statement that in using the analysis of variance to test the hypothesis of equality of a set of unknown means the appropriate critical region of Fisher's statistic z is $|z| > z_p$, and that the z-tables are calculated for such a two-tailed test.

The number of misprints is so amazingly low as to deserve comment.

Here are some suggestions to the author for additions to a future edition of this fine work, or to teachers for supplementing the present edition: (1) It would be useful in view of the recent development of sequential analysis to have some material on measure and integration in the space of infinite sequences $(x_1, x_2, \cdots)$. (2) It would be enlightening to relate the estimation and testing problems arising in the analysis of variance, the analysis of covariance, and linear regression, as examples of a single theory for "linear hypotheses." (3) Even though the treatment of statistical inference is not intended to be extensive, a brief mention at the end of the book of Wald's formulation of the general problem of statistical inference and his concept of statistical decision function might be justified. (4) The value of the book to most users would be increased by augmenting the 39 problems.

HENRY SCHEFFÉ