The Annals of Mathematical Statistics
1971, Vol. 42, No. 3, 1138–1141

BOOK REVIEWS

(*Correspondence concerning reviews should be addressed to the Book Review Editor, Professor James F. Hannan, Department of Statistics, Michigan State University, East Lansing, Michigan* 48823.)

DAVID, H. A. *Order Statistics*. Wiley, New York, 1970. xi + 272 pp.

Review by I. RICHARD SAVAGE
Center for Advanced Study in the Behavioral Sciences

A lively presentation is given of properties of order statistics, such as exact sampling theory, methods for computing moments, and bounding of distributions. Estimates, tests, and methods for handling outliers based on order statistics are also presented in detail.

About 1500 references are made to the bibliography of over 700 items and half the items appeared after 1962, the publication date for Sarhan and Greenberg *Contributions to Order Statistics*, Wiley, New York. These references are used in a variety of ways. David has prepared a very large collection of problems which usually refer the reader to the original paper. He also presents many concise summaries of very specialized subjects. In particular, an appendix summarizes the sources of tables to facilitate work with order statistics. The text is tightly written but most conclusions and mathematical techniques are supported by citations of the research literature.

David wishes his book to have the "features of a textbook and of a guide through the research literature." The mathematics is at the level of advanced calculus and some work in mathematical statistics. At this level a great wealth of techniques are utilized. Working through text and problems with frequent trips to the library would give the student courage to solve hard problems on the job or in future research. Even if a course in order statistics was not required, this book would be useful in a problem course.

I am sure that users of order statistics will find this a very useful guide. Of course, much research is being done. Perhaps future guides will include new areas, such as, Bayesian statistics, decision theory and large sample theory (which David did not consider in detail).

David deserves thanks from the profession, but he does not need it since clearly the book was a labor of love.

KENDALL, M. G. and STUART, A. *The Advanced Theory of Statistics*, **3**, 2nd ed. Hafner Publishing Company, New York, 1968. x + 557.

Review by G. S. WATSON
Princeton University

The first edition of the concluding volume (III) of Kendall and Stuart's epic work was not reviewed in the journal. The appearance of the second edition provides an opportunity to do so and to make a few comments on their achievement, now that they have said their "final words with a considerable sense of relief." Those ghoulish readers who expect a repeat of Kiefer's penetrating review of Volume II will be disappointed—I have neither the skill nor the stomach. I find it very useful to possess and use all three volumes, though Volume I is the most dog-eared. If read with care, they are great books.

The topics covered in Volume III are Analysis of Variance (chapters 35–37), Design of Experiments (chapter 38), Sample Survey Theory (chapters 39, 40), Multivariate Analysis (chapters 41–44) and Time Series Analysis (chapters 45–50). These, unlike the contents of Volumes I and II, are each covered separately by many books at many levels and usually at much greater length.

The treatment here is mostly a rather condensed survey with some most enlightening asides. The latter would be more helpful to a student than the theory, generally speaking.

Except for Sample Survey Theory, the topics are closely linked mathematically but this opportunity is not exploited very much for a simple reason—and it is my main criticism of the book. Vector spaces are never used although they are now in everyone's undergraduate curriculum. Matrix algebra is well used in Analysis of Variance and Design of Experiments, though not elsewhere. However, the endless calculations obscure the simplicity of the ideas. To see a Helmert transformation written out in full when any old orthogonal matrix with the right first row will do . . .! Curiously, multivariate distribution theory suffers least. For ANOVA it is disastrous. There all models are reduced to $\mathbf{y} = \mathbf{X}\boldsymbol{\beta}+\mathbf{u}$, X full rank, by linear restrictions. The Yates "single degree of freedom" approach is not exploited. The method of getting SS for testing H by differencing the residual SS, with or without H, comes later than it should. Chapter 35 covers Model I and chapter 36 Model II and randomization models. Chapter 37 gives a good discussion of the role of assumptions in ANOVA, particularly of transformations. Chapter 38 begins with some unconvincing remarks on the mystery of randomization. I wonder whether this *is* "the most important and (the most) influential of his (Fisher's) many achievements in statistics." I used to think so but it has been a mystery for too long. The bulk of the chapter is concerned with blocking. It is a pity that more space is not devoted to designing experiments to measure the response function well, with a quote from Box of the *many* criteria to keep in mind.

Chapters 39 and 40 are very compact with a minimum of motivation and a maximum of content, though I do not know the field. The derivations are often very neat and non-obvious. Efforts to draw Sample Survey Theory into the main stream of statistics were only beginning when this was written. If our Theory of Inference is no good for a finite population, there does not seem to be much future for it. I also do not agree with K. and S. that Survey Theory has nothing to do with the theory of design of experiments.

The Multivariate Analysis section begins with distribution theory and goes then to testing where the first numerical examples of this volume appear. K. and S. are realistic about the great difficulties of applying the theory. Some of their examples are old friends (Barnard's skulls, Fisher's irises and Rao's neurotics). Craddock's temperature series is a newcomer but it is not clear how canonical analysis helped to identify the changing temperature pattern. This is basically a Time Series problem. Nevertheless, chapters 43 and 44 on Canonical Analysis, Discrimination and Classification contain much interesting discussion. While Discrimination with qualitative data is mentioned, nothing much is presented. The Air Force work of Hodges and Fix is not featured, perhaps because it did not appear in a regular journal. The device illustrated by Fisher in his "Statistical Methods" of running a regression on arbitrary scores for each category and choosing the scores to maximize the multiple correlation certainly deserves a place. This writer once used it successfully to relate dietary factors and a subjective judgement on the state of the tongue where it led the clinician to change his evaluation. This is more of a classification problem. Classification, Cluster Analysis, Pattern Recognition have been successively very O.K. words which might embrace all our problems. As K. and S. say, "The subject is far from exhausted."

Chapters 45 to 50 are concerned with Time Series. Processes with "multidimensional time," and point processes are not considered but multiple series are briefly mentioned. Chapters 45 and 46 provide a survey of old methods and ideas on randomness, trend and seasonality which would be hard to find elsewhere. In chapter 48, the sampling theory of serial correlations is dealt with in great detail. The remaining chapters give a very readable introduction to the statistical theory of stationary processes. There are very few accounts of spectrum analysis and fitting rational processes which may be read with profit by someone with only a solid undergraduate knowledge of statistics. Usually the treatment is more mathematical or more loaded with practical jargon and both are invariably longer than this one. As is usual with these volumes, the commonsensical treatment allows one to read those chapters 47, 49, 50 without reading much of the other ninety-four percent of the treatise.

An opportunity for giving some general results on the distribution of quadratic forms and their ratios is lost in chapter 48 and that episode in our subject seems even more isolated in consequence. It seems a pity that trends seasonality and time series regression are not all viewed in the light of spectrum theory. The latter is mentioned in sections 50.39, 5.40 with several *faux pas* and without any insights. Forecasting is only briefly treated although it is a subject that the authors could have done very well and that badly needs a sensible discussion. But by then the end was in sight and the endurance of K. and S., though colossal, is bounded. We have every reason to be very grateful for their labors and to wish them a well earned rest!

Viewing the whole ambitious project in retrospect, one can see that it *does* review the whole subject in the spirit and with the methods of about 1955. Of course many of the details and references go up to about 1967, provided they fit into the earlier schema. Given that the subject has its ephemeral fads, this is not such a bad thing. Naturally no encyclopaedic work can be "up to date." Of course the fact that methods and ideas have endured for twenty years is no guarantee of immortality in our subject—although as a student in 1950 it seemed quite likely!

Most textbooks nowadays have a very careful development but they are inclined to stick too narrowly to their title so that the fertilizing cross-connections between different topics are not mentioned. Unless one studies them from page one onwards, they are inclined to be very hard to dip into. Moreover there are very few happy turns of phrase to indicate that they are written by fallible human beings about a subject with shaky foundations. On these three counts, K. and S. is virtually unique. Although very different, Feller's Volume 1 comes to mind in this connection.

I can only see a 1970 encyclopaedia arising as a joint effort of many authors with someone like Keifer as Editor. Whether the contributors could be persuaded to write for the common man and not their peers would be his main task. But doctrinal difficulties would surely defeat the project now. I will resist drawing out the humourous aspects one can see in such a project.

In conclusion while K. and S. is not the ultimate authority on each topic, it is an invaluable mine of formulae, facts and commentary and will not have a competitor for a long time to come.

CHIANG, CHIN LONG. *Introduction to Stochastic Processes in Biostatistics*. Wiley, New York, 1968. xvi+313 pp. $13.96.

Review by PETER BROCKWELL
Michigan State University

This book is divided into two parts. Part 1 contains a review of the basic probability theory needed later in the book, a chapter on probability generating functions, and six chapters devoted to solutions of the Kolmogorov differential equations for a variety of continuous-time Markov chains. Two of these chapters are devoted to illness-death processes and a third allows for the additional complications of immigration and emigration. Part 2 deals with the construction of life-tables from empirical data with particular emphasis on the estimation of survival probabilities and life expectancies. The author's life-table techniques are then developed and applied to the analysis of mortality data classified by cause of death. The aim here is to estimate such quantities as the increases in survival probabilities and life-expectancies which would be achieved by the elimination of a particular disease. The last chapter shows how the methods can also be applied to the statistical analysis of medical follow-up studies.

The book will be primarily of interest to demographers, biostatisticians and those who are interested in the applications rather than the general theory of stochastic processes. In fact the treatment is almost exclusively confined to continuous time Markov chains and the emphasis is on finding explicit solutions for the Kolmogorov differential equations. The solution $P(t) = e^{Vt}$ for the transition matrix of a homogeneous finite state-space Markov chain with infinitesimal generator V is given in Chapter 6. The evaluation of e^{Vt} is discussed only when V has real distinct eigenvalues ρ_i, $i = 1, \ldots, n$, in which case the author expresses the transition probabilities $P_{ij}(t)$ in terms of cofactors of the matrices $(\rho_i I - V)$, $i = 1, \ldots, n$, where I is the $n \times n$ identity matrix. The assumption of real distinct eigenvalues is very strong however and a more general discussion of e^{Vt} is needed. The fact that irreducibility of the chain implies ergodicity (i.e. $\lim_{t\to\infty} P_{ij}(t)$