JOHN E. WALSH, *Handbook of Nonparametric Statistics*. D. Van Nostrand Company, Inc., Princeton, New Jersey, 1962. $15.00, 549 pp.

Review by LINCOLN E. MOSES

Stanford University

The author's goal is "to present statements of the more important nonparametric procedures in a concise but understandable form." By a nonparametric procedure he means one having "properties which are satisfied to a reasonable approximation when some assumptions that are at least of a moderately general nature hold." This statement of objective leaves the intended scope of the book a vague matter; partly this is unavoidable.

The material included in this volume comprises tests of randomness, Tchebycheff-type inequalities, point and interval estimation of parameters, tolerance regions, tests and estimates relating to the cdf (and to the pdf) and a last chapter contains "Sequential, Decision, and Categorical Data Results for Distributions." Generally, the author has striven to include coverage of all the relevant literature through and including 1957. The coverage of subject matter is very wide; the literature cited is comprehensive. (I remarked only one reference which seemed to me an important omission.)

The preface states that a second volume[1] will cover material (excluded from this one) that is "concerned with the two-sample problem, the several-sample problem, analysis of variance, regression and discrimination, multivariate analysis, matching and comparison problems, and tests of symmetry and extreme observations." Actually, there are a few two-sample tests included in the book, as well as some tests of symmetry.

The author has intended to write a handbook for applications, a purpose emphasized in the preface as follows: "All results given are designed to use data from actual statistical experiments." In collecting a very large body of literature between two covers, the author has done somewhat more than merely assemble, edit, and summarize; he has offered some opinions on conditions in which particular methods are applicable, (e.g., necessary sample sizes), difficulty of application, sensitivity to assumptions, etc.

It appears that doubtful cases of practicality have been resolved by deciding to include rather than exclude. For example, bounds on probabilities in terms of moment matrices smack less of data analysis than of theorem-proving; they are none the less welcome inclusions in the text. Similarly, Stein put forward his contribution to the *Third Berkeley Symposium* as an indication of how the nonparametric hypothesis testing problem may become (asymptotically) no more difficult than the same problem formulated in parametric terms. Stein somewhat shrugs off the practical utility of his construction; Walsh includes it as a tool for use. There are other such cases. Again, the over-inclusiveness is not harmful.

The book begins with four introductory chapters, totalling 53 pages, followed by seven more chapters, each containing a collection of related statistical procedures; these chapters occupy another 426 pages. There follow tables of the

[1] The author now projects two further volumes to cover these topics and some additional ones.

Normal, t, χ^2, and F distributions, and finally an ample bibliography, giving 602 books and papers, each listing the pages on which it is cited in the *Handbook*—a valuable feature. There is a long and satisfactory subject-index. The table of contents is sixteen pages long, facilitating reference; regrettably, its format frightens the eye.

The main body of the book consists of the seven topical chapters. These all comprise two phases; first, introductory and explanatory material, at the beginning of each chapter and at the beginning of each subsection of a chapter; second, rigidly standardized, highly terse "presentations" of methods. Both phases are frustrating to read and to use—so acutely so as to gravely limit the book's value. The introductory and explanatory material is frequently diffuse, inexact, repetitious, and vague. The "presentations", on the other hand, are for brevity written with abbreviated words, incomplete sentences, and concise (often intricate) notation. It is ironical that of the 479 pages of text fewer than 179 are devoted to these over-terse "presentations," which are the core of the book, and that more than 300 pages is spent in generalized verbalization about the subject matter.

Again and again a reader wishes that he didn't have to decode the "presentations," that some words of motivation accompanied the test, or even only the notation for it. Again and again he wishes that definitions were crisp enough to grab onto. These remarks require illustration. Below is given a photographic reproduction of a passage from page 184:

Approximations to Binomial Distribuiion

Description 1 Approximations to binomial prob. distribution; also exact expression in terms of F-dist. **2** $P[nM(i) \leqslant k]$, $0 \leqslant k \leqslant n$ (k integer), where $nM(i)$ = observed no. of successes **Assumptions 1** Random sample **3** $0 < P[X(i)=1] < 1$ **Results 1** $p = P[X(i)=1]$, $q = 1-p$ **2** *Approximations. For* $p \leqslant .5$ *and* $np \leqslant .8$, $P[nM(i) < k] \doteq \frac{1}{2}p(k-np)e^{-np}(np)^k/k! + e^{-np}\sum_{u=0}^{k}(np)^u/u!$ *For* $p \leqslant .5$ *and* $np \geqslant .8$, $P[nM(i) \leqslant k] \doteq \Phi(k_1/3\sqrt{k_2})$, where $k_1 = [(n-k)p/(k+1)q]^{1/3}[9-1/(n-k)] + 1/(k+1) - 9$ and $k_2 = [(n-k)p/(k+1)q]^{2/3}/(n-k) + 1/(k+1)$. *For* $p \geqslant .5$ *and* $nq \leqslant .8$, $P[nM(i) \leqslant k] \doteq 1 - \frac{1}{2}(n-k-1-nq)e^{-nq}(nq)^{n-k-1}/(n-k-1)! - e^{-nq}\sum_{u=0}^{n-k-1}(nq)^u/u!$. *For* $p \geqslant .5$ *and* $nq \geqslant .8$, $P[nM(i) \leqslant k] \doteq 1 - \Phi(k'_1/3\sqrt{k'_2})$, where $k'_1 = [(k+1)q/(n-k)p]^{1/3}[9-1/(k+1)] + 1/(n-k) - 9$ and $k'_2 = [(k+1)q/(n-k)p]^{2/3}/(k+1) + 1/(n-k)$. *Exact expression*: $P[nM(i) \leqslant k] = P[F > (n-k)p/(k+1)(1-p)]$, where F has F-dist. with $2(k+1)$ and $2(n-k)$ degrees of freedom **3** Applied by use of tables of Poisson, normal, and F-dists. **4** For $npq \geqslant 9$, $nM(i)$ approx. normally distributed with mean $np + (p - \frac{1}{2})$ and variance npq (continuity correction of $\frac{1}{2}$); also for $npq \geqslant 15$, $2\sin^{-1}\sqrt{M(i)}$ approx. normally distributed with mean $2\sin^{-1}\sqrt{p}$ and variance $1/n$. For np sufficiently small (say, at most $\frac{3}{2}$), or p sufficiently small (say, at most .03), dist. of $nM(i)$ approx. Poisson with mean np. Exact expression based on F-dist. of somewhat approx. nature since determination of α for given $F_\alpha(k_1, k_2)$ usually based on interpolation **Characteristics 1** Low **2** Approximations have excellent accuracy **3** Used symmetrically **4** Not pertinent **5** Not pertinent **6** None **Bibliography** Ref. 7, 87, 115, 168, 191, 224 (pages 673–75, 698), 234, 368, 421.

This passage, which deals with material already familiar to any likely reader of the book, is of less notational complexity than the majority of such presentations. It fairly exemplifies the use of the word "decode."[2]

On really complicated procedures sheer impossibility faces some readers, in certain cases, perhaps every reader. It is surely true that "presentations" which exceed two pages in length, and which rest (as many do) primarily on a single paper, are useless if one is near a library which has, or can get, the original publication.

In expository writing one desires to avoid unnecessary formality, but it can be overdone. If "soft" definitions result, too high a price is paid. On page 30 consistent estimates are introduced (defined?) in these words:

> As the number of observations increases, any reasonable type of point estimate should tend to take on values which approach the value of the probability constant estimated. This limiting requirement for a reasonable type of estimate is referred to as estimate consistency. An estimate is consistent if it converges in a probability sense to the value of the constant estimated as the number of observations increases indefinitely. That is, the limiting probability that a value exceeds any fixed deviation from the value of the constant approaches zero even though the fixed deviation may be as small as desired.

It seems like caviling, perhaps, to object that estimate consistency is a *property* of an estimator (or sequence of estimators) rather than a *requirement*. But such approximate use of language can give real trouble and does when we are told [Page 426] that "a loss function depends on the hypothesis selected and the true hypothesis." This is a misleading way to say that these are the arguments of the loss function—i.e., that its values depend upon the hypothesis selected and the true hypothesis. Again, we are told on the same page that "The risk is determined for each decision of the class considered on the basis of the loss function, the *a priori* distribution for the eligible decisions [!?], and the sample values." It is simply not true that the *a priori* probabilities are defined over the action space, and only harm can come of saying so, even though frequently the action space and the set of states of nature do have a perfect correspondence.

Efficiency, a topic appearing throughout the book, is introduced (defined?) in the following words:

> The efficiency of a procedure furnishes a measure of how much of the total available information obtainable on the basis of the observations is yielded by this procedure. Suppose that a specified class of procedures is considered for investigating a stated population property and, for the situations of interest, one of these procedures always furnishes at least as much "information" about this property as any of the others. This "best" procedure evidently can be considered to have a 100 per cent efficiency. For the case

[2] The same passage happens to exhibit some expository sloppiness which amounts to outright error. First, the continuity correction is not correctly given, as is obvious from the fact that even if $p = \frac{1}{2}$, correction needs to be made. Second, and more important, the Poisson approximation is said to be applicable if $np \leqq \frac{3}{2}$, or if $p \leqq .03$; actually, the first condition will not suffice to give even passable approximations (e.g., take $n = 2$, $p = \frac{3}{4}$), nor is either condition necessary for obtaining excellent approximation (e.g., choose $p = .04$, $n = 1{,}000$ and find an excellent approximation).

> of a random sample, the efficiencies of the other procedures of the class can be evaluated by increasing their sample sizes until the information furnished is equivalent to that given by the best procedure. Here, fractional sample sizes are allowed, whose information content is determined by interpolation from that for integral sample sizes. Then the efficiency of a procedure is taken to be the actual number of sample values divided by the number of sample values for which the given procedure is information equivalent to the best procedure. Thus an efficiency of 100γ per cent implies that the given procedure based on $1/\gamma$ as many sample values has an information content equivalent to that of the best test based on the actual sample size.

This seems to say that efficiency is a "good" property related to information, and is somehow quantifiable. It is hard to see what further content these 200 words carry.

In Chapter 5 an idea of I. R. Savage to the effect that certain symmetric uses of data lead to results stochastically independent of the results of certain tests of randomness applied to the same data is alluded to seven times and not once defined precisely. This leads to error. On page 60 the penultimate of these seven references to Savage's idea leads into the remarkable statement, "that is, if the two-sample test supports the hypothesis of the same population, the two samples can be combined into a single sample and any procedure using the values of this combined sample symmetrically is independent of the two-sample test." The statement as it stands is false. (For example, let the test depend on the difference between the two sample means and let the "procedure" be the computation of the variance of the combined sample.) Now what true thing did the author mean to say? He appears to have been misled by his own vague treatment of Savage's idea.

Nine procedures were checked in detail, back to the original references. Two of these had misprints which might be confusing. On page 69 $\sigma_{22} = 1.126n - 1.630$ should be replaced by $\sigma_{22} = .1126n - .1630$, and $\sigma_x^2 = (n + 1)/2$ should be replaced by $\sigma_x^2 = (n + 1)/12$. Again, on page 74 the reader is instructed to reject a null hypothesis if $D > D_\alpha = E_D - \frac{1}{2} + K_\alpha \sigma_D$. This should be changed to reject if $D < D_\alpha = E_D - K_\alpha \sigma_D + \frac{1}{2}$. Although these nine procedures were not chosen at random, but rather in a vague, subjective way, this frequency of error (together with the errors in the cited passage of page 184) give grounds for some alarm about the detailed reliability of the text.

The book is a monument, in a sense. It shows enormous perseverance and desire to serve the profession. It has faults of design and execution which, regrettably, go far indeed toward totally vitiating it. The principal fault of design is in attempting to take a thing so conceptually rich as a statistical inference procedure and to encapsulate it in a terse set format, with a minimum of motivation, and with no worked examples. The principal fault of execution is the failure to bring sufficiently precise and clear writing to an overwhelmingly difficult task. The book will be a useful reference to a person about to study a new topic area who wishes a preview of what papers exist in that area, if it is one covered in the *Handbook*. This may be the book's greatest contribution to its owner.