BOOK REVIEWS

Correspondence concerning reviews should be addressed to the Book Review Editor, Professor William Kruskal, Department of Statistics, University of Chicago, Chicago 37, Illinois.

MAURICE G. KENDALL AND ALAN STUART, *The Advanced Theory of Statistics, Volume 2*, "*Inference and Relationship*." Hafner Publishing Company, New York, 1961. $21.00, 132 shillings, x + 676 pp.

Review by J. KIEFER

Cornell University

The preface to this second volume of the three-volume work (hereafter referred to as K-S) states that the volume bears little resemblance to the original Volume 2 of Kendall (1946) (hereafter referred to as K), and that it was "planned and written practically *ab initio*, owing to the rapid development of the subject over the past fifteen years." Even superficial page counting shows that the three new volumes will contain much more material than did the old two, and the list of references indicates the extent to which the authors have updated the work.

A chief asset of the work is this large content, which will probably not be surpassed by any other reference work in statistics for many years. Also on the positive side is the successful presentation of the material of some of the more classical and unmathematical chapters. My main criticism of the book is that it has a very high density of errors in statements and proofs. Another negative aspect is the exclusion, in a work of this encyclopaedic nature, of much of the content and almost all of the spirit of modern mathematical statistics (as typified, for example, by the papers in these *Annals* which have elicited the most interest in the last two decades). I shall expand on these assessments in the next few paragraphs, and shall then list more detailed comments. My attempt will be to criticize this book in terms of what its aims appear to be, although this attempt may not be successful, since the aims are not stated.

One notices almost immediately that the book is written on an unfortunate mixture of mathematical levels, especially in view of the promise of its title. Unquestionably, one could produce a valuable reference work which lists statistical topics, models, procedures, etc., without giving proofs. One could also write a book which proves precisely stated results, but which keeps the level down by doing this only under restricted but stated conditions. But it seems uneven to make mention of sets of measure zero on one page and then, repeatedly, to give proofs which are only stated to hold "under regularity conditions" which are not specified. (A footnote on page 8 is evidently meant to justify this loose approach, but it is unsatisfactory both because the necessary regularity conditions vary, and also because they are sometimes assumed to hold where they do not.) Reading is also made difficult by the authors' deriving incorrect equalities or giving unsatisfactory proofs, and only *afterward* stating that the equality is only approximate or the proof only heuristic! The large number of errors, some quite fundamental, makes matters worse. Thus, I believe that it would be difficult for a student to use this book as a text in a first graduate course if he really wanted to understand the proofs and many of the deeper ideas. Referring to Hoeffding's (1962) comprehensive review of the recent book of Wilks (1962), I would say that, despite Hoeffding's well-founded criticisms, that book has a uniformity of level and a mathematical preciseness which the present volume does not touch, and Wilks' book therefore seems to me more appropriate as a text or reference for a first course for students with advanced calculus. The comparison with Lehmann (1959), on its own special topic (which occupies at least four chapters and a hundred pages of K-S, not counting preliminaries and applications in other sections), is therefore even more severe on K-S. As a reference work, K-S is much broader than Wilks, and in some areas it contains modern material not mentioned in the latter. I do not think, however, that K-S has succeeded in capturing the spirit of major parts of recent statistical theory, so that the present owner of K who buys K-S will gain more in the updating of the more classical statistical developments than in the addition of important modern developments. This is not a book on decision theory, and it is not and should not necessarily be written in that framework; but lack of even mention of risk functions, admissibility, complete classes, multiple decision procedures, etc., is a noticeable omission in a book of this encyclopaedic nature. What is more, as should become apparent in this review, the book could gain enormously in terms of presentation from the clarity and preciseness associated with modern statistics, without ever adopting its general framework and abstract results or mentioning its applications.

Some of the chapters on the more classical and less mathematical topics, such as the first two of the four chapters on statistical relations (26 and 27) and the chapter on categorized data (33), contain comprehensive and well-presented treatments. A few chapters (especially 33) are admirable in their inclusion of many practical examples. Some of the chapters on topics which have undergone considerable advancement in recent years, such as those on tests of fit (30) and robustness and distribution-free procedures (31), are notably up-to-date; but the deficient mathematical developments are all the more noticeable in such chapters.

Thus, the book's main usefulness will be as a reference work (which, although incomplete, is by far the most complete of its kind) and as a source-book of problems for a student who is able to rectify the errors.

It would require an undue amount of space to list all the book's errors. As a selection, I shall therefore go into more (although not complete) detail in the first chapters (17 and 18) and two others (29 and 30) than in other chapters.

The first three chapters (17, 18, 19) cover point estimation in about one hundred pages. It is disappointing that a work aiming for such completeness should omit mention here of Bayes procedures, completeness, invariance, and the minimax principle, all of which are treated in more elementary books. (The first three of these are given inadequate coverage in other contexts in later sections.) One notes at once an unfortunate choice of notation which makes reading more difficult than necessary throughout the book: the writing of E or P instead of something which evidences the underlying probability law, such as E_θ or P_θ. The terminology "consistent estimator" instead of "consistent family (or sequence) of estimators" seems unfortunate. (The terminology for tests on page 240 is similar.) On page 5 it is incorrectly stated that "a consistent estimator with finite mean value must tend to be unbiased in large samples." This is repeated on pages 42, 55, and on page 19 in a further discussion which is made precise only in the discussion of super-efficiency on page 44. (The corresponding discussion of consistency and unbiasedness in testing on page 445 is also incorrect.) The limiting variance and variance of the limiting d.f. are not distinguished. Although there is no discussion of the *local* character of best unbiased estimators, the discussion of (globally) best unbiased estimators is reasonable up to page 16, at which point a function of an unbiased estimator of a parameter is discussed as though it is an unbiased estimator of the function of the parameter. In Example 17.14 the "estimator" obtained depends on the unknown parameter. Sufficiency is unfortunately introduced (pages 22 and 27) in terms of estimators. The entire discussion of sufficiency, "single sufficient statistics", etc., would benefit greatly from a precise definition of "statistic", which from its usage seems usually (e.g., on p. 25) but not always (e.g., p. 29) to mean "continuously differentiable real function." The discussion of Equation (17.73) et seq. treats the likelihood function for a rectangular family as though the range $0 < x < \theta$ in which the functional form θ^{-n} holds is not part of the functional definition. Regarding what they consider ambiguities in choosing a sufficient statistic, the authors state that "we simply choose a function t which is a consistent estimator, and usually also an unbiased estimator, of θ." Exercise 17.13 is incorrect as stated; for example, if $X_1, X_2, \cdots X_n$ are independent with common Lebesgue density $c(\theta, \phi) \exp \{-x\theta - x^2\phi - x^3\theta\phi\}$ for $x > 0$ where $\theta, \phi > 0$, then there is a "single sufficient statistic" for θ if ϕ is known and for ϕ if θ is known, but no sufficient pair when both are unknown.

The appraisal of ML on pages 38 and 61–62 is good as far as it goes. The proof of consistency of ML which is presented as "a simplified form of Wald's proof," misses the point of Wald's compactness argument by incorrectly concluding that $\hat{\theta}_n \rightarrow \theta_0$ (the true value) in probability from the fact that $P_{\theta_0}\{\hat{\theta}_n \rightarrow \theta^*\} = 0$ for *each* $\theta^* \neq \theta_0$. (The "proof" is made all the more difficult by its incorrect conclusion that $n^{-1} \log [L(x \mid \theta_0)/L(x \mid \theta^*)] > 0$ "for large n with probability unity" from the corresponding conclusion about expectations, and by the writing of the event of convergence as "$\hat{\theta} = \theta_0$".) The discussion of Hurzubazar's work does not consider the possibility of inconsistent solutions of the likelihood equation. Example 18.4 reflects lack of identifiability rather than lack of uniqueness of ML estimation itself. On page 45 differentiation under an integral sign is incorrectly justified by the fact that the derivative of the integrand fails to exist at only one point. The "efficiency" of ML estimation for several parameters is presented in terms of generalized variance; more satisfactory is the fact that the limiting covariance matrix of any sufficiently regular sequence of estimators, minus that of the ML sequence, is nonnegative definite; this implies, for example, the (one-dimensional) efficiency of ML estimation of any regular real function of the parameters. (On page 81, LS estimators receive a similar treatment.) Exercise 18.33 is incorrect: if the two variances are $(\log n)^{-1}$ and $(\log n - \log 2)^{-1}$, they are of the same order $(\log n)^{-1}$, but the relative efficiency, as defined on the top of page 20, is $\frac{1}{2}$. The ML discussion admirably includes references to such recent work as LeCam's and Bahadur's but regretably omits the v. Mises-Kolmogorov-LeCam-Wolfowitz asymptotic Bayes result.

A more geometric approach to LS estimation would have enhanced some of the discussion; identifiability is discussed on page 87, but the linear subspace of identifiable linear parametric functions is unfortunately not described in settings where some parameters are unidentifiable. On page 91 it is incorrectly stated that "the addition to $\hat{\theta}$ of an arbitrary function of the observations, which tends to zero in probability, will make no difference to its asymptotic properties." On the positive side, such topics as BAN estimators are discussed.

Chapters 20 and 21 discuss interval estimation. The presentation, like that of K, gives one chapter to confidence intervals and one to fiducial intervals (plus Bayesian intervals), but includes somewhat more criticism than did K. The treatment would be enhanced if confidence *sets* were considered from the outset. On page 115 shortness (referring to length) is discussed as though it is a global rather than local property. Neyman shortness is also considered, but the relation between the two concepts is not. Asymptotic, but no small sample, optimality results are given (in which the super-efficiency mentioned two sections earlier is forgotten). In the discussion of randomization, the use of the actual sample values to aid in randomization based on the sufficient statistic, which is of practical importance, is omitted. Tolerance intervals are discussed briefly.

The chapter on fiducial intervals contains several additions from that in K, and is quite lucid in some of its criticism. The problem of two normal means is given extensive treatment, including work of Scheffé, Wald, and Welch. The Bayesian approach is based on infinite measures (following Jeffreys to some extent) without justification; moreover, when the normal mean and variance are both unknown, $d\mu\, d\sigma/\sigma$ rather than $d\mu\, d\sigma/\sigma^2$ is used. Invariance of either the Jeffreys or non-Bayesian variety is not mentioned. On page 153 there is a confusing passage regarding "imaginary" confidence intervals and the "need for sufficiency," and a discussion of Bayesian inference which is too brief and which unfortunately does not discuss the question of where the prior distribution comes from. An example on that page is misleading; after computing the a posteriori law of a normal mean μ which is assumed to be uniformly distributed on $(0, 1)$, the authors criticize the confidence interval approach by ignoring the ways in which it can properly make use of such information limiting the domain of

μ; this last appears to be at least partly due to an overemphasis on "exactness" (corresponding to similarity for tests).

Chapters 22 through 25 are concerned with hypothesis testing. The definitions of "parameters" and "simple" on page 162 are rather loose. In the discussion of simple hypotheses on page 166, Equation (22.5) supposes $L(x \mid H_0) > 0$ whenever $L(x \mid H_1) > 0$. The conclusion (22.7) that the ratio can equal a constant only on a set of measure zero is false. The result "proved" on page 173 regarding nonexistence of two-sided UMP tests for regular distributions with range independent of parameter value is false, as the density $[1 + \theta^6 x^{2+\text{sgn}\,\theta}]/2$, $-1 < x < 1$, $-1 < \theta < 1$, shows for $\theta_0 = 0$; an error in the proof is that θ^*, θ^{**} depend on x, θ_1, so that $\theta_1 - \theta_0$ need not change sign as claimed. The reference on page 174 to the rectangular distribution is irrelevant in view of the nondifferentiability in that case. Example 22.9, which is used in Example 22.10 and Exercises 22.10 and 22.11, is incorrect; the critical region should also include $\{\min_i x_i \leqq \theta_0\}$, and thus does not depend only on $\bar{x}$. Thus, Example 22.10 is not an example of what it claims to be (and, if the critical region *were* $\{\bar{x} \leqq c_\alpha\}$ as incorrectly claimed, Condition (2) of Section 22.21 *would* be satisfied, which it is correctly stated not to be). The criticism of conventional significance levels is good. The reason given for sufficiency in Excerise 22.9 is a wrong one.

In Chapter 23, page 187, it is incorrect to conclude that there is a critical *region*, rather than a randomized test, depending only on a sufficient statistic. (Several parts of the book would receive easier treatment in terms of randomized test functions.) Feller's example on page 188 is not exactly relevant to the discussion there, since the number of parameters increases with n. Example 23.2 is misleading, since the *minimal* sufficient statistic is complete. The statement on page 196 about H_1 being reduced to a simple alternative is false. The proof on page 208 of a generalized NP Lemma (e.g., Lemma 8.2 of Lehmann and Scheffé's work on completeness) is somewhat garbled. In the geometric interpretation on page 216, the condition is sufficient but not necessary as stated; a function can maximize without its derivative maximizing. In the discussion of ancillary statistics, the phrase "$T_r \mid T_s$ is sufficient for θ_k" sets the tone for an unprecise presentation which later mentions "the two statistics $(T_r \mid T_s)$ and T_s." The function h does not necessarily disappear as stated when $r + s = n$. In Exercise 23.1, the value $\alpha = \frac{1}{2}$ should be excluded. Exercise 23.4 incorrectly states that $\text{cov}\,(Z, Y) = 0$ implies $E(Y \mid Z) = 0$ (in a context $Y = \partial \log L(x \mid \theta)/\partial\theta$ where a counterexample is provided by $Z = X$ normal with mean 0, variance θ^2) and incorrectly ascribes the statement to Neyman; fortunately, Exercises 23.6–23.9 do not depend on this statement as suggested. Read $(x_1, \cdots, x_n)$ for θ at the end of Exercise 23.27.

Chapter 24 deals with LR tests and the general linear hypothesis. The sloppy "proof" of the asymptotic LR statistic distribution in Section 24.7 could well have been omitted. The "asymptotic sufficiency" proof at the beginning of that section is also inadequate. The study of the nonregular case is an interesting nonstandard inclusion, but there are errors in it; for example, in Section 24.15, the LR statistic is 0 with probability approaching 1 as $n_1 \to \infty$ when $\theta^* > \theta_0$, and does not have the stated chi-square distribution; the statement about bias is also wrong, as Example 22.6 shows. On page 243 there is another confusing development concerning conditional statistics; for example, a conditional statistic $t \mid L$ is integrated with respect to L to obtain t. The LR approach receives fair criticism. A total of about one page is spent in the book's main effort on invariance (there being slight reference later in conjunction with rank-order tests). The Hunt-Stein theorem is never mentioned, nor is any of Stein's recent work.

It is a pity that part of the general discussion of Chapter 25, "The Comparson of Tests," did not begin the treatment of hypothesis testing, to present earlier developments in a better-directed framework. Most of the development here is, however, asymptotic. On page 265 it is concluded from the fact that often power $\to 1$ at a fixed alternative as $n \to \infty$, that one "must" consider a sequence of alternative approaching the null hypothesis θ_0; this ignores another well known approach, which is based on log $(1 -$ Power). Mention is made on pages 267 and 270 of the fact that "no case where $m > 2$ seems to be known," where $(\theta - \theta_0)^m$ is the first nonconstant term of the power function; reparametrization in standard examples yields such an m, and families of competing tests with any two values of m can be obtained easily. The discussion of asymptotic power of tests at the point of greatest difference seems quite useful.

Chapters 26–29 cover statistical relationships, and the coverage is more extensive than that of earlier chapters, relative to the existing literature. Chapter 26 deals with linear regression and correlation. The subject receives a good classical treatment, as do partial and multiple correlation in the next chapter. (The notation convention introduced on page 323 regarding secondary subscripts is somewhat confusing.) The development is continued in Chapter 28 ("The General Theory of Regression"). The demonstration of Section 28.5 is incorrect; the authors speak of "completeness" of a single density g and use it in the manner of Chapter 23; what is needed is completeness of the family $\{e^{it_1x}g(x)\}$, and the factor e^{it_1x} should be deleted from (28.29) (which, as it stands, could never follow from (28.28)). In Section 28.15, the question, "how should the elements of $\mathbf{X}$ be chosen so that the estimators $\hat{\beta}_i$ are uncorrelated?" and the subsequent discussion could be misleading, since such a choice is not always possible. The precise quantitative basis for deciding which degree polynomials "evidently . . . are good fits" on page 361 is unfortunately not given. The development on page 368 of confidence regions for a regression line is unconvincing in its statement that the functional I cannot be effectively minimized without further restriction on the form of g; there is no mention of why or whether the form chosen loses nothing. In Exercise 28.18 the matrix should be replaced by its inverse. It is disappointing to see exercises such as 28.22 in which the testing theory developed earlier is completely forgotten, and tests which could have been derived from that theory are instead devised entirely on intuitive grounds.

Chapter 29, "Functional and Structural Relationship," is the least successful of these four chapters. It begins with a statement of the need for a "clear terminology and notation"; the reader can decide for himself whether Section 29.2 achieves this. (See also footnote, page 383.) Since (Section 29.3) errors are assumed uncorrelated and not independent, the conclusion of Section 29.9 regarding joint normality is incorrect. In the second paragraph on page 380 are some more misleading comments on identifiability; only one of the six given parameters is identifiable, but many combinations are. The idea of treating the four cases of the structural relationship under normality with various combinations of parameters known is a good one, but the treatment of the "overidentified" case as "a somewhat embarrassing position" which is resolved by changing the model, rather than merely by using ML or some other method of estimation on the original model (with the given underlying restrictions on the parameter space), is very bad. The pages which follow again indicate how much more satisfactory it would have been to discuss identifiability per se at the outset of the chapter, rather than to drag it in where it arises because the ML method breaks down. (The alternatives presented in Section 29.27 reflect the same spirit.) The properties incorrectly described for various procedures later in the chapter would be more clearly understood if the lack of identifiability for normally distributed incidental parameter (Reiersøl) were kept in mind. It seems unwise to say (page 388) that an estimator has "performed rather well," on the evidence of one experiment based on 9 observations. "Imaginary confidence regions" arise again on page 391. The properties stated for Geary's method are not very appealing; in fact, except in trivial parametric examples, no satisfactory estimation procedure is described for the linear structural relation in the whole chapter; Wolfowitz's minimum distance method and the methods of Neyman and Scott and Stein are not described. Since normality of errors has been assumed in the development leading up to page 399, one may question the logarithmic transformation discussed there (as do the authors, but on page 413). The discussion of Wald's procedure on page 400 is inadequate; the use of the instrumental variables is not the same as in (29.94), since they are no longer independent, and the procedure (since Neyman's critique) is well-known not to have the indicated validity; the treatment through page 405 suffers accordingly. Section 29.42 deserves similar comment; the subject is much more delicate than indicated in this chapter, and the assumption that "the values x are so far spread out compared with error variances that the series of observed ξ's is in the same order as the series of unobserved x's" is not a very meaningful one. Criticism similar to the above applies to the material on controlled variables, curvilinear regression, etc.

Chapters 30–32, and to some extent 33, are concerned with nonparametric problems. The first of these chapters, entitled "Tests of Fit," is introduced by a commentary on the point of formulating the problem in a certain manner, which omits mention of the most important aspect, the infinite dimensional character of the alternatives. In all of the asymptotic developments here and elsewhere, $O(n^{-\frac{1}{2}})$ is used without explanation to mean also $O_p(n^{-\frac{1}{2}})$. The choice of the usual chi-square one-tail critical region receives an interesting discussion on page

422, but the asymptotic optimum power properties are not mentioned as reasons. The developments of this chapter are otherwise admirably up to date in the inclusion of work of Chernoff and Lehmann, Watson, etc., although the execution is sloppy. The discussion of Section 30.22 (see also 30.44) is vague about where the power should be maximized and the sense in which it is maximized by Mann and Wald; the derivation on page 438 of the results of these authors is given in terms of a different criterion than theirs (which can be, but is not, shown to be equivalent); in this derivation (30.66) is false for large k, and when $k \to \infty$ (as it does) the asymptotic normality results cannot be introduced to obtain (30.70) as stated. (Some such oversights, which have been investigated by T. Taylor in a Cornell master's essay, appear in the original, but they are not justified by the irrelevant reason on page 440 that "an upper limit to k is provided by the fact that the multinormal approximation to the multinomial distribution cannot be expected to be satisfactory if the np_{oi} are very small.") As an example of the style and attitude toward proofs which is sometimes infuriating, we cite the fact that, in the *following* numbered section, it is stated that Mann and Wald use "a much more sophisticated and rigorous argument . . . our own heuristic argument . . ." (the last referring also to a generalization to composite hypotheses, which is not as "clear" as it is stated to be). Pity again the poor student who spends hours trying to understand the "proof" which precedes, and *then* reads this! The avoidance of the noncentral chi-square distribution in the asymptotic discussion of Section 30.24 is mysterious. The argument on the top of page 436 is not quite as obvious as stated; for example, (30.59) requires that $(E_1 - E_0)(x^2 - x)I_R > 0$ where I_R is the indicator of the critical region. The top of page 440 makes a conclusion, "clearly," on the basis of a single sampling experiment. The discussion of the "more conservative" equal probability case on page 440 seems incorrect, since the expected minimum frequency is greater there than for unequal probabilities. Section 30.34 and Exercise 30.7 assume H_0. Section 30.36 and Exercise 30.10 are incorrect unless t_1 and t_2 are assumed appropriately invariant. The comparison of the p_k^2 tests and chi-square test on page 448 is not very meaningful: the alternatives differ. On page 451, the normalization $n\omega^2$ is cited as different from the $n^{\frac{1}{2}}$ of the CLT; of course, they are really the same normalization. Feller's derivation of the Kolmogorov-Smirnov law is given essentially in toto except for justifying taking the limit of the generating function; an equation on page 456 reads $p_n(0) \to (2\pi n)^{-\frac{1}{2}}$(as $n \to \infty$). In Exercise 30.9 the χ^2-distribution is only asymptotic.

Chapter 31, on robust and distribution-free procedures, is one of the relatively better chapters in converage of what it sets out to cover, although point estimation (Winsorisation, etc.) receives little attention in this or the next chapter. The relationship of various problems given in Section 31.14 is somewhat misleading. The inclusion of much material on rank order tests, including, for example, the Chernoff-Savage results on the Fisher-Yates statistic, is to be commended. Section 31.69 seems slightly overenthusiastic about the Wilcoxon and Fisher-Yates two-sample tests relative to all others; the statement that "although little is known of the power of these tests" (such as the Kolmogorov-

Smirnov and other two-sample tests), "it is clear that they are less efficient" certainly requires at least a limitation of alternatives from the general (31.111) that had been discussed. The reduction from (3.1.160) to (3.1.161) enlarges H_0, since symmetry about 0 of the d.f. of $x_1 - x_2$ does not imply that $F(x_1, x_2) = F(x_2, x_1)$. On page 509, it is incorrectly stated that the method of breaking ties affects limiting distributions in the *normal* case (where ties have probability zero); it is not clear what the authors meant to refer to. Chapter 32, which covers uses of order statistics, is brief but fairly adequate, except again on point estimation. (The parametric use of linear combinations of order statistics was treated briefly in Chapter 19.)

Chapter 33 covers categorized data comprehensively. This chapter, in its many numerical examples, seemed different in spirit from, and more successful than, most other chapters. A more critical development is needed in some places. For example, Yule's approach is developed without comment by using a questionable "invariance principle" on page 546. Three possible degrees of "fixing" in 2×2 tables are explained well; the principle of conditioning could have received longer comment. On page 556, the sense in which Lancaster's pooling prescription is "best" is unclear. The canonical analysis of page 569 receives a good start, but (Sections 33.46 and 33.50) sample and population properties are not sufficiently distinguished. The analysis of Section 33.57 does not mention the extent to which the Poisson dispersion test is really a test of smallness of variance/mean. Formula (33.127) is complete nonsense.

The final Chapter 34 covers sequential methods. In describing excess over the boundaries, we read (page 597), "there is no exact probability of reaching the boundary at M—and, in fact, this point is inaccessible." The consideration of the OC and ASN function for other states than the two simple hypotheses which have been under consideration, should be motivated. The explanation of why optimum properties of SPRT's "will not, then, come as a surprise" does not make it clear why nonconstant bounds should not yield such properties. The "proof" of Wald's equation, ascribed to Johnson (who later published a correction), is a sloppy version of Wolfowitz's proof; it incorrectly does not use the finiteness of En, but justifies the interchanges E and $\sum_1^n$ (with n a random variable) as "being legitimate since $E|z_i|$ exists and is finite." Similar remarks apply to Section 34.31. On page 611, the authors do not specify the sense in which SPRT's for composite hypotheses "can be used in the ordinary way" when weighting functions are used and the tests therefore are not based simply on the product of independent factors. This is one of the least precisely written chapters, and page 613 exemplifies this: the passage to the limit in using nonintegrable weight functions is never justified; a statement about the distribution of s^2 and $(\bar{x} - \mu_0)/s$ should instead refer to the joint distribution over all stages; the reason given for monotonicity oversimplifies the situation greatly; "optimality" is used in a sense which has no visible relationship to the ASN; "excess" is neglected everywhere (as it was announced earlier that it would be); the restriction here and on page 612 to tests with given bounds A and B and the weight function-LR structure is never given meaning. The statement on page 616 that "it cannot be said that, in general, sequential procedures are very satisfactory for estimation" is never explained. In Exercise 34.4, replace g by $-g$ in the last two sentences. Exercises 34.5, 34.6, and 34.7 require unstated conditions. The differentiation in Exercise 34.9 is never justified. The results of Exercises 34.11, 34.12, and 34.13 only hold asymptotically. In Exercise 34.17, an upper bound which depends on the unknown parameter is given on the choice of sample size.

Finally, on the last page of text, one-half page is spent on "decision functions." I feel obliged to repeat that, while I do not feel that the ideal three-volume work on the advanced theory of statistics is a three-volume work on decision theory, I question the relegation of this topic to one-half page of a 676 page volume which is part of a three-volume work of such supposedly broad coverage. Why include such detail on some theoretical aspects of hypothesis testing and measures of association (for example), and never mention admissibility, multiple decision procedures, etc.? The general results of abstract decision theory are probably not so important for inclusion in a work of this sort. But, even if the general theory and interesting recent applications are omitted from the book, the sampling of criticisms included in this review certainly shows that K-S would benefit greatly from absorption of some of the spirit of modern statistics, in the careful way in which the latter looks at problems and compares procedures. Clarity and precise statements in the motivation and definitions, and in the development of solutions, does not entail mathematical abstraction or the rigid use of oversimplified decision-theoretic models. Nor does the complex nature of many practical statistics problems mean that one should not try to be as clear and accurate as possible in formulating them; if anything, the importance of such care is greatest in such settings which are farthest from the simplest decision-theoretic models.

In summary, this useful reference would be much improved if it contained fewer invalid proofs and misleading motivations, but greater precision and clarity.

REFERENCES

[1] Hoeffding, W. (1962). Book Reviews. *Ann. Math. Statist.* **33** 1467.
[2] Kendall, M. G. (1946). *Advanced Theory of Statistics*, **2**. Griffin, London.
[3] Lehmann, E. L. (1959). *Testing Statistical Hypotheses*. Wiley, New York.
[4] Wilks, S. S. (1962). *Mathematical Statistics*. Wiley, New York.