

be called excessive. It is fervently to be hoped that not only libraries, but many young mathematicians will be able to acquire them and profit from them. Eisenstein tells us that his love for mathematics came from studying first Euler and Lagrange, then Gauss; studying the great work of the past is still the best education.

ANDRÉ WEIL

BULLETIN OF THE
AMERICAN MATHEMATICAL SOCIETY
Volume 82, Number 5, September 1976

Sequential statistical procedures, by Z. Govindarajulu, Academic Press, New York, 1975, xvi + 565 pp., \$39.50.

Although the idea of sequential statistical procedures did not originate with Abraham Wald, it was he who pushed the subject in a few years to great heights. Some of his work in the later years was done in collaboration with J. Wolfowitz. By the time of Wald's premature death in 1950 sequential analysis had been established as an important new and exciting branch of mathematical statistics. It gave rise to numerous new problems, both in probability and in statistics. No wonder then that many researchers have taken up where Wald left off. But most of their results are scattered throughout the literature, and very few books have been written that attempt to put all or some of this together.

Clearly then, there is a need for a comprehensive book on sequential analysis. Govindarajulu's book is such an attempt to fill that void. According to his own words, in the Preface to his book, he has been mostly interested in gathering in one place what has been done to date in the field of sequential estimation. The last (fourth) and longest chapter is devoted to that subject. But sequential testing of hypotheses has also been treated extensively. Chapter 2 is on the sequential probability ratio test (SPRT) for simple hypotheses or for a one-parameter family of distributions. Chapter 3 deals with composite hypotheses and some multiple decision problems. Certain other topics have purposely been omitted. But what has been included constitutes a very large proportion of what has been done in sequential analysis from its beginning to the present. Also, the book has a long list of references, and each reference is followed by the numbers of the pages in the book where the reference has been made—a useful feature. Another useful feature is the large number of problems sprinkled throughout the text. The author disclaims completeness, but there can be no denying that the book is reasonably exhaustive. As a result, I think the book will be mostly useful as a reference work: one can now easily find out what has been done in a particular area, and by whom. However, in spite of the comprehensive treatment of testing and estimation, a few, in my eyes, serious omissions have been committed. I shall return to this point later in the review.

Will Govindarajulu's book serve another purpose besides reference? In the Preface the author states that he also has tried to serve the needs of students, and recommends his book as a text in a course in sequential analysis. Here I sharply disagree. While the book may serve the instructor, and the problems will be useful for the student, I am of the opinion that the book is totally

unsuitable as a text to learn from. If this judgment sounds unduly harsh, consider what a good textbook ideally should be like. Since it is intended to teach something to the novice, it should select the topics that seem to the author most important and give a lucid and self-contained exposition of those. That implies, of course, that the author has digested the material thoroughly before he reproduces it in his words. In contrast, Govindarajulu is exhaustive rather than selective; in fact he has collected just about everything he could lay his hands on, with no indication whatsoever of relative importance. Furthermore, the exposition leaves everything to be desired. The main trouble I see here is that (with a few exceptions) Govindarajulu has copied from the original sources with very little change and without any attempt at digesting them. I think it is as important for a writer as it is for a teacher in the classroom to have mastered the material fully, extracted the main ideas, and then to reproduce those ideas in his or her own words. That way teaching can be a creative and enjoyable experience.

This brings me to what I regard as the most disturbing aspect of Govindarajulu's book. If one compares the text in the book with any of the articles on which the text is based, then very soon it becomes clear that Govindarajulu has copied the original source (or part of it) essentially verbatim, with only minor changes (if any), e.g. in the notation. Here is an example. In the paper by Anderson and Friedman (1960), p. 60, one reads: "The result may be applied directly to problems in which it is desired solely to specify the risk of acceptance or rejection at a single point. For example, it might be desired solely to specify the risk of accepting a product with a stated percentage of defectives –the consumer's risk –and not to specify the producer's risk. For such a problem, a curtailed single-sampling plan C would be indicated." Compare this with Govindarajulu's rendition on p. 6: "Theorem 1.2.1 may be applied directly to problems in which it is desired solely to specify the risk of acceptance or rejection at a single point. For example, it might be desired solely to specify the risk of accepting a product with a stated percentage of defectives –the consumer's risk –and not to specify the producer's risk. For such a problem, a curtailed single-sampling plan C would be appropriate."

I wish I could say that this is an isolated instance. But unfortunately, on the contrary, it is the same situation in every case that I have checked (and I checked quite a few). Never before have I read a book in which such a shockingly large fraction was copied from other sources. I do not want to give the impression that I am accusing Govindarajulu of plagiarism. That is certainly not the case, for Govindarajulu always identifies his sources very carefully. But copying it is, nevertheless.

In taking material from another book obviously greater care had to be exercised (I presume in order to avoid violation of copyrights). Compare B. K. Ghosh (1970), p. 255: "So far we have discussed the theory of tests which discriminate between *two* hypotheses. There are, of course, practical situations where the experimenter may be interested in more than two distinct hypotheses, and indeed he may not be willing to reformulate his courses of action into the simplified version of (2.26)." This is adapted by Govindarajulu on p. 163 to read: "In the preceding sections we have discussed a theory of sequential tests that is appropriate for distinguishing between two simple (or composite) hypotheses. However, there are many practical situations wherein a choice

among three or more courses of action is required and the theory described earlier is not appropriate.” Govindarajulu then goes on to describe a different example from Ghosh’s. Here there is only a similarity rather than an identity between the two passages. But a little farther the two texts are uncomfortably close. Ghosh: “Consider a given family $[F_n(E_n; \theta), \theta \in \Theta]$ of models, and $k > 2$ hypotheses $H_j: \theta \in \omega_j$, for $j = 0, 1, \dots, k - 1$, where $\{\omega_j\}$ are disjoint nonempty subsets of Θ . A statistical test S_k of (5.76) is defined by a set of rules, based on $\{E_n\}$ ($n = 1, 2, \dots$), to accept one and only one of the k hypotheses.” Compare this with Govindarajulu: “Consider a given family, $[F_n(\mathbf{X}; \theta), \theta \in \Theta]$ of distributions and $k > 2$ hypotheses: $H_j: \theta \in \omega_j, j = 0, 1, \dots, k - 1$, where $\{\omega_j\}$ are disjoint nonempty subsets of Θ . A statistical test δ_k of (3.8.1) is a set of rules based on $\{\mathbf{X}_n\}$ ($n = 1, 2, \dots$) to accept one and only one of the k hypotheses.”

Comparing Govindarajulu’s and B. K. Ghosh’s books I could easily find the correspondence. For instance, pp. 121–122 in Govindarajulu correspond to pp. 239–240 in Ghosh, 123–126 to 247–251, etc. Altogether about 36 pp. of Ghosh’s book found their way into Govindarajulu’s. If I were the author of a book I doubt whether I would appreciate someone else using my labor to write his book.

I am sure that Govindarajulu’s formula for book writing reduces the writing time considerably, since it removes the worry about how to write the material down once it has been collected. But the effect of this kind of writing on the exposition cannot help but be disastrous. Besides the lack of digestion of the material (about which I complained earlier) it is obviously not even always understood or checked for accuracy. For instance, there is an obvious error on p. 247 of Blackwell and Girshick (1954): “Consider now the conditional expected value of $\xi_{j+1} \dots$ ”. Here ξ_{j+1} should have been $g(\xi_{j+1})$. This error is copied faithfully by Govindarajulu on p. 344 (he has g_{j+1} instead of $R(g_{j+1})$). This is such an obvious error that I am convinced anyone who would have read the sentence with understanding would have caught it. On p. 72 there is a big confusion between n and N , copied from the original. Sometimes symbols that occur in the original source are inadvertently left unchanged. For instance, in §4.3 n was never changed to N which makes Conditions I and II rather puzzling.

As a result of certain omissions or change of order of presentation from the original, some statements do not make sense. Thus, Theorem 3.13.2 makes no sense because of the unexplained symbols m_i and t_i . They are belatedly introduced in the proof. In the original paper by Darling and Robbins (1967b) the proof comes before the statement of the theorem, so these things do make sense! Remark 2.4.1.2 is very puzzling until one reads the original paper by Chow, Robbins, and Teicher (1965) in which the term “genuine stopping variable” is explained. On p. 82 out of nowhere come the symbols u and v . In the original paper the notational change from B, A to u, v had been announced. On p. 344–345 the symbols E_0, E_j , and quantities $R_j(a; x)$ and $R(g, \mathcal{E}_n)$ are undefined. They had been introduced by Blackwell and Girschick in an earlier section that was omitted by Govindarajulu.

There are several passages in the book that lead me to believe that Govindarajulu does not always fully understand what he is writing. On p. 18,

second paragraph, Govindarajulu asks the question whether given $\theta_0 < \theta_1$ there is a SPRT that coincides with plan C . But that question had already been answered in the negative in the preceding paragraph. What Theorem 2.1.1 proves is something very different: given $\theta_0 < \theta_1$ there is no SPRT with the same α and β as C , unless C happens to be a SPRT. The interest in this theorem lies in its connection with the optimum property of the SPRT, but that subject has not been discussed yet in §2.1 (thus, Theorem 2.1.1 comes at the wrong time).

Theorem 4.9.1 is stated for arbitrary estimator T . This is entirely incorrect. The theorem as stated in Blackwell and Girshick is for T that incorporates the best terminal action (assuming the “best” exists; actually, the theorem can and is stated in Blackwell and Girshick without that assumption). The discussion preceding the statement of the theorem gives as an example the estimation of the mean of a normal distribution if the prior is uniform over the real line! Of course, there is no such prior. For a normal prior with mean 0 the claim would be correct. On p. 79 it is stated that part of the proof of Theorem 2.9.1 consists in showing that given a SPRT, and given prior π and cost per observation c , there exist losses w_0 and w_1 such that the SPRT is Bayes. Govindarajulu then refers for details of the proof of Lehmann (1959, pp. 104–110) or B. K. Ghosh (1970, pp. 93–98). But it has escaped him that in those treatments (relying on a lemma by LeCam) it is not c that is fixed. Rather, one sets $w_1 = w$, $w_0 = 1 - w$, and it is the constants w and c whose existence has to be shown. The approach with fixed c and w_1, w_2 varying independently was used in a paper by Burkholder and myself (1963). That paper contains no less than two proofs, but neither is mentioned by Govindarajulu.

The conclusion in the statement of Theorem 2.2.1 is incomplete and must be extremely puzzling to the novice: “Then, Wald’s SPRT terminates finitely with probability one provided $P(Z = 0) < 1$.” What is P ? It should have been clearly stated that P is the true distribution of Z , and need not be given by either f_0 or f_1 . The second and third sentences in the proof do not clear this point up (instead, they cause more confusion). The second sentence is irrelevant: Z could be 0 a.e. (P) without being 0 everywhere. Thus, f_0 and f_1 could easily be essentially different while at the same time $P(Z = 0) = 1$. The failure to recognize for which distribution P an assertion is made about N also persists in Chapter 3. On p. 137 Govindarajulu remarks: “For Example 3.6.7 B. K. Ghosh (1970) points out that the z_i are iid and hence all the properties of an SPRT are valid.” This is true if the X_i are normal, but not in general. Thus, Govindarajulu misses the point that the distribution P of the X_i is completely arbitrary and need not be normal, even though the model giving rise to the SPRT is normal.

The treatment in §2.4 of the equations $ES_N = ENEZ$ and $ES_N^2 = ENEZ^2$ displays an inexcusable ignorance of the history of this subject. These equations are due to Wald and were proved by him in the more restricted setting of the SPRT. They are often referred to as Wald’s first and second equations. It is true that others after Wald made important generalizations, such as N. L. Johnson (1959) for the first moment of S_N , and Chow, Robbins, and Teicher (1965) for higher moments. But the equations originated with Wald, and Wald is not even mentioned in §2.4! Incidentally, Wald’s first equation is suddenly called “Blackwell’s theorem” on p. 56 (copied from M. N. Ghosh’s (1960) paper).

There are some places where Govindarajulu could easily have improved upon the original source. Thus, the condition in Lemma 2.6.1 that Z be absolutely continuous is absolutely unnecessary. Of course, the proof then has to be given without differentiation. In Theorem 2.6.2 the unpleasant restriction (2.6.9) is also unnecessary. Any random walk with negative drift is absorbed by a left boundary at a time that has finite expectation.

Govindarajulu claims in the Preface that the book is made self-contained by providing most of the proofs. That is far from true. Many proofs are not included, or only partly. The proof of Theorem 2.5.3 is an example of the latter. Also in the proof of Theorem 4.9.1 and its preliminaries repeated reference had to be made to Blackwell and Girshick for important parts in the argument that Govindarajulu does not wish to present. There are many more instances like these. I do not see much sense in presenting part of a proof. If one has to go to the source anyway for the missing arguments, one may as well learn it all (and better) from that source.

Granted that Govindarajulu has striven for reasonable completeness, I miss certain very well-known classical results. Perhaps the most glaring omission is any reference to the papers by Wald: *Foundations of a general theory of sequential decision functions*, *Econometrica* **15** (1947), 279–313, and Wald and Wolfowitz: *Bayes solutions of sequential decision problems*, *Ann. Math. Statist.* **21** (1950), 82–99. In the first of these papers Wald sketched the nature of Bayes procedures in general, and in the second paper Wald and Wolfowitz specialized this to iid observations and constant cost per observation. Although their approach (which, by-the-way, for the case of testing simple hypotheses is concisely reproduced in Lehmann (1959), pp. 104–106) runs into measure-theoretic difficulties which are avoided by doing it the Arrow-Blackwell-Girshick (1949) way, it has considerable intuitive appeal, and one should know it.

As a whole the subject of sequential Bayes procedures has received short shrift from Govindarajulu. §3.12 treats some special cases; the scope of §4.9 is very limited. Yet, the nature of Bayes procedures in general is basic to various problems in sequential analysis, both in testing and in estimation. For instance, it is vital to most proofs of the optimum property of SPRT's. It is also a necessary preliminary to Theorem 4.9.1 (which is then an easy corollary). Parts of the derivation of the nature of Bayes procedures, taken from Blackwell and Girshick, are now hidden in rather incoherent form in §4.9. Govindarajulu makes it appear as if Theorem 4.9.1 is the basic result, whereas the basic result is really the stuff hidden in the preliminaries to the proof. A rigorous proof of the nature of Bayes procedures is a rather tricky business. It requires the union of Arrow, Blackwell, and Girshick (1949), and Blackwell and Girshick (1954) (each separately does not quite suffice), where the basic method is truncation and backward induction. (By the way, it is strange that in Appendix 8 on backward induction Govindarajulu never even mentions Arrow, Blackwell, and Girshick.) To the best of my knowledge this has not appeared in print in a unified way and it seems to me that Govindarajulu missed an opportunity here. (Alternative ways, employing optimal stopping methods, can be found in Chow, Robbins, and Siegmund (1971).)

I also missed reference to Blyth's (1951) results on minimax and admissible

sequential estimation of the mean of a normal distribution under rather general loss and cost function. Yet, it is a classical paper whose method for proving admissibility is often quoted. His work on the mean of a rectangular population is misstated on p. 342.

If I have given the impression that Govindarajulu's book only contains accounts of other people's work, with nothing added of his own, then that is not quite fair. A small amount of his own research is incorporated. The problems he supplies have been mentioned earlier. Furthermore, Govindarajulu did catch a few mistakes in the sources from which he borrowed. For instance, he points out a rather bad error in a footnote on p. 72. On p. 214 he points out an error in B. K. Ghosh's book which invalidates Ghosh's argument. Unfortunately, when trying to correct the computation and the argument he also commits an error (in the differentiation of the function h) and so his expression for $h'(\alpha)$ is incorrect. (It is claimed that $h'' < 0$ in the interval $(0, \frac{1}{2})$, but in fact $h''(\alpha) > 0$ for α close to $\frac{1}{2}$. I have no doubt that $h > 0$ in $(0, \frac{1}{2})$, but I have not seen a proof yet.)

Theorems 2.4.2, 2.4.3, and 2.4.4 seem to be new (except, of course, part (i) of Theorem 2.4.2). Unfortunately, the assumptions are not completely stated. Worse is that no proofs are supplied. There is a lot of manipulation, which is valid provided the interchange of summation and expectation can be justified, but Govindarajulu never provides this justification. Dominated convergence does not seem to work. I am much obliged to Professor Tze Leung Lai for showing me a martingale proof of Theorem 2.4.2(ii). So at least that result seems to be true, even though not proved in the book. Another question is what the formulas in Theorems 2.4.2–2.4.4 are good for. In a remark on p. 36 Govindarajulu says that in the case of Wald's SPRT Theorem 2.4.2 leads to approximate expressions for EN and $\text{Var } N$. But all formulas (also the ones in Theorems 2.4.3 and 2.4.4) contain only the first moment of N , so it is hard to see what $\text{Var } N$ would follow from. Govindarajulu continues on p. 36 with some more puzzling remarks about the conditional expectation of N given that the hypothesis is accepted (or rejected). How (2.4.15)–(2.4.17) are going to be used for that purpose is a mystery to me.

In conclusion, I would like to transmit to Professor Govindarajulu my sincere regrets that this review turned out to be so negative. But in my mind a prerequisite for reaching an audience, be it by spoken or by written word, is a deep concern with the manner in which the thoughts are going to be conveyed. In my opinion the book fails to display that kind of concern.

ROBERT A. WIJSMAN

BULLETIN OF THE
AMERICAN MATHEMATICAL SOCIETY
Volume 82, Number 5, September 1976

Sobolev spaces, by Robert Adams, Academic Press, New York, 1975, xviii + 268 pp., \$24.50.

This monograph is devoted to the study of real valued functions u defined on an open set Ω in Euclidean n -space R^n having the property that u and all its distribution derivatives up to (and including) order m are functions that are p th power summable. Here $1 \leq p < \infty$ and m is a positive integer. The set of all such functions u is denoted by $W^{m,p}(\Omega)$ and when endowed with an appropriate norm, for example,