By action of the Council of the Institute of Mathematical Statistics the 1979 *volume of the Annals of Statistics is dedicated to the memory of*

HENRY SCHEFFÉ

The Annals of Statistics
1979, Vol. 7, No. 6, 1149–1161

HENRY SCHEFFÉ 1907–1977

BY C. DANIEL AND E. L. LEHMANN
Rhinebeck, New York, and University of California, Berkeley

Henry Scheffé was born on April 11, 1907, in New York City. His father and mother were German, originally from Alsace. He went to elementary school in New York and graduated from high school in Islip, Long Island in 1924. Following graduation, he entered the Cooper Union Free Night School for the Advancement of Science and Art to study electrical engineering, and in 1925 became a student at the Polytechnic Institute of Brooklyn. There his grades were all A's except for a D in mechanical engineering which on reexamination was changed to an E! During this time he also worked as a technical assistant at the Bell Telephone Laboratories and took a training course there.

In 1928 he was admitted to the University of Wisconsin with advanced standing to study mathematics. The record shows but one course related to statistics. It was "Theories of probabilities (sic) and methods of least squares" given by Warren Weaver. Scheffé was an intercollegiate wrestler at Wisconsin. (Contrary to widespread rumor, he acquired his broken nose at the age of three, not in wrestling.) He received his B.A. with high honors in 1931 and remained at Wisconsin for graduate study, obtaining his Ph.D. in 1935. His thesis, written under R. E. Langer, on asymptotic solutions of certain differential equations, was published in 1936 [S1][1].

During the years 1935-1938, Scheffé taught mathematics at the University of Wisconsin, at Oregon State University, and at Reed College. His first statistical paper [S3] was written at Oregon State and appeared in the American Mathematical Monthly in 1942.

At some time in 1940-1941, he decided that the field of mathematical statistics promised more interesting opportunities for research than analysis. (Harry Goheen remembers that when Scheffé found that parts of his thesis subject had already been worked on by Gauss he felt that he should move to less well-travelled paths.) This resulted in his going to Princeton in 1941 to teach (as Instructor from 1941-1943 and as Lecturer during 1943-1944) and to do research with the statistics group in the Mathematics Department. In the company of Ted Anderson, George Brown, Bill Cochran, Will Dixon, Alex Mood, Fred Mosteller, John Tukey, Sam Wilks and Charlie Winsor it is easy to imagine his growing attachment to statistics. From 1943 to 1946, he worked as consultant and Senior Mathematics Officer at the Office of Scientific Research and Development, under a contract of this office with

Received October 1978.

[1]The numbers [S1], . . . refer to the list of Scheffé's publications at the end of the article.

Princeton University. In this capacity, he wrote a number of reports, which we have not seen but which describe work carried out under the general heading "Effects of impact and explosion."

In 1944–45 Scheffé taught statistics at Syracuse University and from 1946-1948 served as Associate Professor of Engineering at UCLA. Actually, the first year he was on leave as a Guggenheim Fellow, part of it in Berkeley where he began his joint work with E. Lehmann. In 1948 Scheffé moved once more, this time to Columbia University, where he remained as Associate Professor of Mathematical Statistics until 1953, serving as executive officer of the department for 1951-1953.

Scheffé joined the faulty of the University of California at Berkeley as Professor of Statistics and Assistant Director of the Statistical Laboratory in 1953 and remained there until his retirement in 1974. In 1954 he was President of the Institute of Mathematical Statistics of which he had been elected Fellow ten years earlier, and from 1954 to 1956 he was a Vice President of the American Statistical Association. The year 1962-1963 was spent at the University of London on a Fulbright Research Award. From 1965 to 1968 he was Chairman of the Berkeley department. It was a period of great unrest at the University, which put a heavy strain on him. But in spite of violently conflicting attitudes by different groups of faculty and students, he managed to hold the department together and keep the atmosphere within the department pleasant. His fairmindedness was greatly valued by all members of the department.

Scheffé enjoyed teaching and would recall the better students in his courses for many years. He preferred not only to correct his own examinations, even when the classes were rather large, but also to conduct the laboratory sections of his courses himself without the aid of teaching assistants. He was equally conscientious and sympathetic as a referee for several journals. An associate editor of one of these (Technometrics) tells us that of all the refereeing he has seen, Henry's was the most patient and helpful.

Along with his teaching and research, Scheffé managed a schedule of daily bicycling and swimming and he engaged in frequent summer snorkeling and backpacking. He was a dedicated tourist, especially to Mexico and France, who often returned with small works of art chosen with a sure and strongly individual taste. He was sensitive to the beauties of nature and had a particular enthusiam for desert country. He loved music and as an adult learned to play the recorder and treasured the opportunity this brought of playing chamber music with friends. A few months before his death, he had just finished reading all of Trollope's novels.

In 1974 Scheffé retired from Berkeley and accepted a three-year appoinment as Professor of Mathematics at the University of Indiana in Bloomington. In June of 1977 he returned to Berkeley, which he considered home and where he was planning to prepare a new edition of his book. This was not to be—he died on July 5, 1977 from injuries sustained in a bicycle accident earlier that day.

He is survived by his wife, Miriam, by his daughter Molly, now a mathematician working near Boston, and by his son Michael, a commericial artist and designer in Los Angeles. Friends and colleagues all over the world join them in mourning his loss.

Work in Statistics. Scheffé's first major statistical investigation was concerned with two-sided tests and confidence intervals for the ratio of two normal variances. However, in studying this problem he found that determination of the optimum procedures he was seeking required first an extension of the Neyman-Pearson theory of optimum tests. In [9], these authors had developed a theory of uniformly most powerful (UMP) and (in the presence of nuisance parameters) UMP similar tests. This theory was successful in dealing with an important class of one-sided problems but typically was not applicable to two-sided alternatives. To overcome this difficulty, Neyman and Pearson in [10] introduced the notion of unbiased tests, but only for one-parameter situations, that is, when there are no nuisance parameters. The main result of Scheffé's paper [S4] fills in this gap by providing conditions for UMP unbiased tests in the presence of nuisance parameters.

These results were applied to the problem of obtaining optimum tests for the hypothesis of equality of two normal variances σ^2 and τ^2 and corresponding confidence intervals for their ratio $\theta = \tau^2/\sigma^2$ in the succeeding statistical publication [S6]. In the first part of this paper it is taken for granted that the hypothesis will be rejected when the appropiate F-statistic T is either too small or too large, so that the associated confidence intervals have the form $T/B \leqslant \theta \leqslant T/A$, and a number of choices for A and B are considered: (i) minimizing $E[\log(T/A) - \log(T/B)]$ or equivalently B/A to obtain the "logarithmically shortest" confidence intervals; (ii) the reciprocal limits, for which $A = 1/B$; (iii) the limits corresponding to the likeihood ratio test; (iv) those determined by the equal-tails test. In particular, there is a numerical comparison of (i) with the more convenient limits (iv).

The second part of the paper uses the results of [S4] to show that the test (i) is in fact UMP among all similar tests that are unbiased. This optimum test is now part of the classical literature. It is a sign of the times that the paper does not mention the disastrous sensitivity of the test to the assumption of normality. Many years later, Scheffé devoted considerable attention to the problem of robustness in his book [S29].

Having "disposed" of the problem of the equality of variances in the presence of unknown means, it was natural for Scheffé to turn next to the dual problem of testing the equality of two normal means when both variances are unknown, the so-called Behrens-Fisher problem. Unfortunately, the conditions under which he had derived his optimum tests did not apply in the new situation. Even more seriously, there was now no obvious statistic corresponding to the F-statistic T in the earlier case, which would provide a class of "natural" similar tests of the hypothesis. In [S7] Scheffé set himself the task of providing such a class. His starting point was a test mentioned by Neyman and attributed by him to Bartlett, for the case of equal sample sizes. If the two samples are then denoted by $X_1, \cdots, X_n$ and $Y_1, \cdots, Y_n$ and their expectations by ξ and η, the differences $Z_i = Y_i - X_i$ constitute a sample of n from a normal distribution with mean $\zeta = \eta - \xi$ and variance $\sigma^2 + \tau^2$. Bartlett's suggestion was to use Student's one-sample t-statistic calculated form the Z's to test the hypothesis $\zeta = 0$, i.e., $\eta = \xi$. When the sample sizes are unequal, $m < n$ say, a statistic with a t-distribution can of course be obtained by discarding n-m of the Y's. Since this is clearly wasteful, Scheffé considered a more general class of statistics which under the hypothesis $\eta = \xi$ have a t-distribution and determined the member of this class which leads to the confidence intervals for $\eta = \xi$ with smallest expected length. He compared his solution with the optimum confidence intervals when the value of $\theta = \tau^2/\sigma^2$ is known, and showed that there is very little loss of efficiency.

Scheffé's solution of the Behrens-Fisher problem received some criticism because of its lack of symmetry, more precisely, because it is not invariant under permutations of the X's among themselves and the Y's among themselves. In the case $m = n$ for example, the denominator of the t-statistic is proportional to $\Sigma(Y_i - X_i)^2$, which is clearly changed by such permutations. Scheffé reacted to this criticism in a note [S12] in the succeeding volume of the Annals by showing that there exists no symmetric t-statistic with the properties he had postulated.

This note is however not his last word on the subject. In 1970, more than twenty-five years later, he returned once more to the Behrens-Fisher problem in [S35] to join his earlier critics and disavow his t-solution, which he discussed in a section entitled "An impractical solution." Referring to [S7] and [S12], he wrote: "These articles were written before I had much consulting experience, and since then I have never recommended the solution in practice. The reason is that the estimate s_d [the estimate of the standard deviation in the denominator of the t-statistic] requires putting in random order the elements of the larger sample, and the value of s_d and hence the length of the interval depends very much on the result of this randomization of the data. The effect of this in practice would be deplorable." He then proceeded to discuss some alternative solutions of the problem, in particular the Welch-Aspin test and Welch's t-test with estimated degrees of freedom, neither of which has the similarity property, which earlier had seemed so essential. The question of the existence of a similar test of reasonable structure, which was initiated by Scheffé in [S7] and [S12], was answered negatively in the 1960's by Linnik and his students. (An account of this work is given by Linnik: "Lecons sur les problèmes de statistique analytique." Gauthier-Villars, Paris, 1967.)

The problem of similar two-sample tests is also the subject of the paper following [S7], this time in a nonparametric context. If $X_1, \cdots, X_m$ and $Y_1, \cdots, Y_n$ are samples from F and G, Scheffé sought in [S8] to determine the totality of level α tests of the hypothesis $F = G$ when this common distribution is completely unknown, except possibly restricted to be continuous or to have a density. Expressed in modern terminology, his necessary and sufficient condition is essentially that the conditional rejection probability given the $m + n$ order statistics $z_{(1)} < \cdots < z_{(m+n)}$ of the combined sample is α a.e.

[S8] was Scheffé's first contribution to nonparametric statistical inference which, particularly as the result of the recent work of Wald and Wolfowitz, was beginning to become a "hot" topic. Scheffé himself contributed significantly to the heating up process by following [S8] by the first comprehensive review of the new field [S9].

The longest section of his review paper is again devoted to the problem of similar tests. It is pointed out how the structure of these tests, which had been uncovered in [S8], extends to other nonparametric problems and encompasses the two important special classes of nonparametric tests to be found in the literature; randomization tests and rank tests. The paper then goes on to describe the tests in these classes which had been proposed for problem of randomness, the two-sample problem, the hypothesis of independence, and the analysis of variance.

The second part of the paper deals with estimation. There is a brief section on point estimation, which formulates the problem as that of estimating a functional defined over the nonparametric class of distributions in question, and points out how the concepts of unbiasedness and consistency continue to apply in this new setting. The survey then turns to confidence sets and discusses confidence intervals for a population median or the difference of two medians, and confidence bands for an unknown distribution function. This second part of the paper concludes with an account of the work of Wald and Wilks on nonparametric tolerance limits.

The third and last part of the paper is entitled, "Toward a general theory." It discusses the consistency criterion and the modified likelihood principle of Wolfowitz and contains a short description of Wald's decision theoretic approach to statistics.

In his introduction to this survey, Scheffé wrote: "Only a very small fraction of the extensive literature of mathematical statistics is devoted to the nonparametric case, and most of this is of the last decade. We may expect this branch to be rapidly explored however: The prospects of a theory freed from specific assumptions about the form of the population distribution should excite both the theoretician and the practitioner since such a theory might combine elegance of structure with wide applicability. "

The paper made an important contribution to the burgeoning new field by providing it with a solid foundation for the spectacular development which Scheffé had so accurately predicted, but it leaves the present-day reader with the question of why the author himself did not participate in this development. He saw the attraction of the new methods, and by detailing what had been accomplished made clear some of the outstanding gaps. Why did he not try to fill some of them? A partial answer may be found in his view of the principal task that lay ahead, which is outlined in the concluding sentences of his paper: " . . . we have in the parametric case a large body of constructive theory which may be applied in particular examples to yield the optimum tests or estimates; thus we have the Fisher theory of maximum likelihood statistics for point estimation, and the constructive theorems of the Neyman-Pearson theory for the existence of critical regions of types A, A_1, B, B_1 and the related types of "best" confidence intervals. The contrasting lack of any constructive general methods challenges us in the nonparametric theory". This program probably was not feasible then and in fact has not yet been carried out. Scheffé himself later came to a quite different view of the needs of statistics. By that time he no longer thought so highly of nonparametric methods and hence did not return to that subject.

During the next few years Scheffé wrote only a few shorter notes, some of which may have grown out of his war work. Two of the these were his first directed to applied statisticians. The first of these [S14] shows some of the relations between control charts, the analysis of variance, and chi-square tests. The latter [S17] combines an imaginative understanding of the problems of quality control engineers with his typical thoroughness in expounding the operating characteristics of the usual mean and range charts.

In spite of these more applied concerns, his heart at that time was still in optimum theory and the problem of similar tests. In these interests he was joined in 1946 by one of us (EL); the collaboration led to a preliminary note [S16] and then to the papers [S18] and [S24]. This work is concerned with two statistical problems, determining the totality of similar tests for a given family of distributions and finding unbiased estimators with uniformly minimum variance. A test represented by a critical function ϕ is similar at level α with respect to a family $\mathcal{P}$ of distribution P if

$$EP\phi(X) = \alpha \quad \text{for all } P \in \mathcal{P},$$

while δ is an unbiased estimator of a real-valued functional g defined over $\mathcal{P}$ if

$$EP\delta(X) = g(P) \quad \text{for all } P \in \mathcal{P}.$$

The tool which provides an easy key to the two problems in a large class of cases is that of a complete family of distributions. If T is a sufficient statistic for $\mathcal{P}$, then T or more accurately the induced family $\mathcal{P}^T$ of distributions of T, is said to be complete if

$$EPf(T) = 0 \quad \text{for all } P \in \mathcal{P} \quad \text{implies } f(t) = 0 \quad (\text{a.e. } \mathcal{P}^T).$$

In the presence of such a complete sufficient statistic one easily obtains the results:

(i) A necessary and sufficient condition for ϕ to be similar at level α is that

$$E[\phi(X)|t] = \alpha \quad (\text{a.e. } \mathcal{P}^T).$$

and

(ii) Any functional g which has an unbiased estimator, has an unbiased estimator with uniformly minimum variance, namely the unique unbiased estimator depending only on T.

These results, together with various extensions and ramifications, provide simple and unified proofs of most of the special results that had been obtained earlier by various authors on these two problems. It must however be emphasized that the basic ideas underlying (i) and (ii) were not really new. In particular, the method (i) of obtaining all similar tests had been used by Hsu in [5] while the method underlying (ii) had been employed by Rao in [12], [13] and [14]. The main contribution appears to be the isolation of the concept of *completeness*, which plays an important role in mathematical statistics even beyond its impact through the results (i) and (ii). Another concept which was introduced and investigated in the course of this work is that of *minimal* sufficient statistic. This concept was introduced independently in the year following the publication of [S18] by Dynkin [3]. The two concepts are related by the fact that a sufficient statistic cannot be complete unless it is minimal.

With this series of papers, Scheffé's concern with similar tests and optimality theory, which had started with his first major statistical paper in 1942, essentially came to an end. However, before abandoning the subject, he wrote one final paper related to the problem of optimum tests, which grew out of the thesis of one of his Ph.D. students, Stanley Isaacson. Scheffé's interest had been primarily in the problem indicated in the title of [S4], testing hypotheses with one constraint, that is, which specify the value of a single real-valued parameter. Isaacson [6], at Scheffé's suggestion, had studied hypotheses specifying the values of two or more parameters. For this purpose, he required an extension of the fundamental lemma of Neyman and Pearson, and in [6] established sufficient conditions for a solution to his optimization problem. In [S20], Chernoff and Scheffé generalized Isaacson's result, and proved that their condition is not only sufficient but also necessary.

The work described so far, which constitutes the first phase of Scheffé's research, was dominated by issues belonging to the domain of mathematical statistics. The second phase begins with [S21]. It is devoted primarily to problems relating to the analysis of variance and is concerned with issues belonging to the domain of statistical methodology. The extent of his change of attitude is indicated by the fact that in his comprehensive and rigorous book on *The Analysis of Variance*, the optimality properties are compressed into five pages. The shift in point of view is seen perhaps most clearly by an entry in the subject index. Under optimality properties we find the sub-entry: "Of lesser importance than robustness". The page to which this entry refers contains the following footnote: "When I became aware that the nominal probability of type 1 error for the standard test of the equality of variances of two populations is invalidated by non-normality to the same order of magnitude as found in Table 10.2.1, I found little consolation in the optimum properties someone once established for that test (Scheffé, 1942)."

This development was not due to any relaxation in the intensity, depth or rigor of his work. It reflected, rather , the growth of a well trained but extremely modest mathematician into a more self-confident—though still modest—scientist who wanted to be useful to a wider circle of research workers.

A short expository account for chemical engineers on the background of statistical methods [S19] was his first effort in the new direction. As a consultant for

Consumer's Union he became involved in the problem of "organoleptic" preference tests. The resulting paper was on an analysis of variance for paired comparisons [S21] which extends the general methods of the analysis of variance to situations in which only subjective degrees of preference between pairs of samples are recorded. This becomes an analysis of *differences*, which is not transformable into the more familiar partitioning of direct measurements. The designs and methods of analysis produced have been widely used in many areas ever since.

The next paper, on judging all contrasts in the analysis of variance [S22], was motivated in large part by the problems of his colleagues in applied statistics. There are of course many settings in which experimenters—and hence statisticians—are willing to commit themselves, before the results are in, to a specified set of comparisons among the results. But there are at least as many other settings in which the experimenter rightly refuses this restriction. He must see all that the data tell him, and this may well require him to make comparisons that were not foreseen.

The problem of multiple comparisons has been a source of concern to statisticians and to others for a long time. It was J. W. Tukey who first devised a method of simultaneously estimating by confidence intervals *all* contrasts among a set of means like those usually produced in a planned experiment, all being independent with the same variance. The method he chose—called the T-Method by Scheffé—gives priority to the $K(K-1)/2$ pairwise comparisons among K means, using the critical point of the Studentized range as a multiplier of the standard error of each mean, to guarantee with predetermined confidence coefficient that all intervals cover their corresponding parametric values. These intervals are the shortest possible (in expectation) for the differences between pairs of population means. But the T-method pays a price for this in requiring larger intervals than need be for more complex groupings, for example, for the comparison of the average of two means with the average of a different set of three.

Scheffé set himself the problem of making allowances, which would control the probability of no errors of coverage at preassigned confidence level, for the experimenter who reserves the right to make *any* comparison among the means, even those suggested by long mulling of data. He saw and proved that the confidence ellipsoid in $(K-1)$-dimensional space is equivalent to the intersection of the sets between all pairs of parallel tangent hyperplanes that support the ellipsoid. This proved that the confidence ellipsoid provides a confidence set (at the same level) for all contrasts among the means (from one line of an analysis of variance table). The size of the ellipsoid depends only on the value of the F-statistic at the chosen level of confidence, and on K. Scheffé showed that the (infinite) set of confidence interval statements given by the S-method for a given set of K means is equivalent to the usual F-test in the sense that at least one confidence interval on a contrast among the means fails to cover zero if and only if the F-test is significant at the α-level. (He later discovered a forerunner of this result in a paper by Working and Hotelling 1929.)

The allowance for any observed contrast $\hat{\psi} = \Sigma_{i=1}^{K} c_i X_i$ ($\Sigma c_i = 0$) is the interval from $\hat{\psi} - A$ to $\hat{\psi} + A$ where A is a constant times the estimated standard error of $\hat{\psi}$. The constant is calculated from the F-distribution with $K - 1$ and ν (the degrees of freedom for error) degrees of freedom. For the case $K = 6$, $\nu = 18$ and $\alpha = .05$, for example, the constant is 3.72. This (typical) value is disturbingly large to statisticians accustomed to using a tabular t-value of, say, 2.101 for such cases. The paper (in Tables 5 and 6) gives some numerical consequences of the latter practice. Scheffé often said that he consoled himself with the thought that users who did not mind using the F-test for overall tests of null hypotheses about a set of K means, should not mind the values produced by the S-method since they were equivalent to the F-test. He also suggested some relaxation of the size of the critical region by setting $\alpha = .1$, which of course results in some shortening of the expected values of the intervals. In the example just given, the effect is to reduce the constant from 3.72 to 3.32.

One of us (CD) takes pride in the fact that his copy of this paper is inscribed "To the guy who hounded me into this." This required listening to long disquistions throughout more than a year, on how messy the distribution problem was, climaxed one evening at 11:30 with "Guess what; it's the F-distribution."

Extensions, limitations, alternatives continue to be matters of active research and controversy. R. G. Miller's book [7] and recent paper [8] give an excellent overview and bibliography up to 1977. Judging from the frequency of its citation, the S-method is one of the most widely used devices in interpreting F-ratios found in the analysis of variance and in regression.

An objection to the S-method, or at least to one interpretation of it, was raised by Olshen [11]. Suppose that the S-intervals are reported only when a preliminary F-test rejects the overall null hypothesis H. Olshen showed that the conditional coverage probability (given that H has been rejected) which seems then to be the probability of interest, is at least in some situations always smaller than the stated confidence probability. The issue is further discussed in a note by Scheffé [S37] together with a comment by Olshen and a rejoinder by Scheffé.

In this note Scheffé distinguished between two situations. In one only a test of H is to be carried out, with no further statements in case of rejection. (This is basically the case in which the parameters being tested are nuisance parameters, which one hopes to discard). In the other, in which the parameters being tested are of primary interest, rejection of H is to be followed by some confidence statements for significant contrasts. Scheffé pointed out that in the latter type of situation, the S-procedure (interpreted as consisting of an initial statement that H is accepted or rejected, supplemented in the latter case by more detailed confidence statements) has an overall probability of a correct statement equal to that claimed by his method.

Finally, in his rejoinder following Olshen's comment, Scheffé noted that the difficulty is shared by many standard statistical procedures, since in fact studies are usually published only when they report statistically significant results. The debate thus joins an important on-going discussion of differences between statistical theory and practice.

The paper [S22] on the S-method was followed in 1954 by a study of partially hierarchal (crossed and nested) models in the analysis of variance [S23]. Its purpose was to describe as compactly as possible the effects of several process variables and several sources of random variation on the corrosion resistance of sheet steel used in making tin cans. A companion paper by the experimenter and the applied statisticians Vaurio and Daniel [16] gave the details of the system and the experiment. In a later letter to the same journal (see [S23]) Scheffé withdrew the model he had used in favor of that of Wilk and Kempthorne [17].

His work on mixed and alternative models for analysis of variance resulted in two papers [S25, S26], both published in 1956. These reflected his efforts to develop "Model II" which he considered less well understood than "Model I", the so-called fixed effects model. They constituted part of his program for the book *The Analysis of Variance* that was taking shape during this period.

The paper on fitting straight lines when one variable is controlled [S27] contains first a detailed explication of Berkson's celebrated 1950 paper [1] and then a generalization to sets of straight lines with possibly different parameters. The paper solved an industrial consulting problem which he believed to be of frequent occurrence.

The two papers on experiments with mixtures [S28, S32] were responses to real problems in chemical and petrochemical industry. The former appears to be the fundamental work in the field. It is extended, discussed and applied in a large number of later papers by other authors. Scheffé devised symmetrical designs[2] for q mixture components, x_i, each at $(m + 1)$ levels, called $\{q, m\}$ lattices, and gave symmetrical polynomial fitting equations of degrees 1, 2, and 3. The linear equation (which contains no constant term) is:

$$\eta = \Sigma_1^q \beta_i x_i, \qquad i = 1, \ldots, q.$$

The quadratic equation (which contains no constant or squared terms) is:

$$\eta = \Sigma \beta_i x_i + \Sigma \beta_{ij} x_i x_j, \qquad \text{with } 1 \leqslant i < j \leqslant q.$$

A canonical cubic equation is also given.

For each fitting equation, simple expressions for the least-squares estimates of the coefficients are derived in terms of the observed responses, as well as formulas for their variances. (The coefficient $-4/3$ in his equation (4.8) should be -12; the -3 should be $+3$.) Extensions to equations of the fourth order, to correspond to gasoline blending problems, were made by Gorman and Hinman [4]. They also included a useful explanation of the equations of lower order.

Scheffé's second paper in mixture designs [S32] was also written in response to applied problems. It is more difficult to read, even though as G. A. Barnard

[2]Optimality properties of these designs were established by Kiefer in *Ann. Math. Statist.* **32** (1961) 298-325.

complained in the discussion following its presentation, the paper does not conform to Royal Statistical Society usage, since "it leaves too few loose ends for discussants to jump on." The title does not indicate that the paper handles experiments with some mixture-variables (constrained by $\Sigma x_i = 1$) and some "process variables (unconstrained)." Perhaps it still awaits a more elementary exposition to complete the bridge to physical application.

The Analysis of Variance [S29] is Scheffé's magnum opus. It has established itself as the basic work on the theory and application of that form of analysis. The review in the Journal of the Royal Statistical Society [2] by D. R. Cox starts: "The first adjective that comes to mind to describe this book is *professional.*" The review ends:

> Altogether this is a most important book, deserving to be widely read. Statisticians working in fields where the analysis of variance is used extensively are likely to find the book extremely valuable in consolidating their knowledge of the theoretical side of the subject.

L. J. Savage, writing in Mathematical Reviews [15], concluded his review:

> Whatever may be said in criticism is overshadowed by the merits of this book which is unique in its field and will be indispensible to all who seriously do, study, or teach statistics in connection with experimentation.

The long life of the first edition has moreover justified the judgments of the reviewers. This must be attributed to its combination of thoroughness and generality (e.g., in Chapters 1 and 2) with its intuitive and practical insights (especially in Chapter 10).

The note on separation of variables [S34] extends a result on transformability to additivity from the two-variable case given in his book (page 95 ff.) to any number of independent (continuous) variables. It includes a number of examples, some of them with surprising outcomes.

In his last major paper "A statistical theory of calibration" Scheffé attacked a problem of great practical importance. The result was a thirty-four page treatment of characteristic thoroughness but with no numerical example. (He was working at the end on an expository paper that was to include several sets of real calibration.) One of the inhibiting factors to its wide use must be the eleven-page set of tables (of over 13,000 entries) giving values needed for actual use in calibration. Only two values from the tables are needed for a single calibration problem. The tables are extraordinarily condensed in that all entries are one- or two-digit numbers, less than 18 in the first and less than 7 in the second. He gives a procedure that for given α and δ, and for every possible sequence of later observations (after the calibration curve is made), guarantees that the probability is $\geqslant (1 - \delta)$ that the proportion of true statements about coverage that are made is in the long run $\geqslant (1 - \alpha)$.

Scheffé made path-breaking contributions both to statistical theory and to methodology. Again and again he showed us that it is possible to solve problems posed by statistical practice on the firm basis of carefully stated assumptions. He thus served as a model for younger statisticians, and his influence will long be felt in the statistical world.

We gratefully acknowledge the assistance of the following persons who provided us with information, and with comments and suggestions about earlier drafts of this obituary: D. Brunk, C. Eisenhart, K. Green, W. Kruskal, D. Lowry, F. Mosteller, I. Olkin, M. Scheffé, S. Stigler, J. Tukey, G. Watson.

The publications of Henry Scheffé

[1] Asymptotic solutions of certain linear differential equations in which the coefficient of the parameter may have a zero. *Trans. Amer. Math. Soc.* **40** 127–154 (1936).

[2] At what level of rigor should advanced calculus for undergraduates be taught. *Amer. Math. Monthly* **47** 635–640 (1940).

[3] An inverse problem in correlation theory. *Amer. Math. Monthly* **49** 99–104 (1942).

[4] On the theory of testing composite hypotheses with one constraint. *Ann. Math. Statist.* **13** 280–293 (1942).

[5] Linear differential equations with two-term recurrence formulas. *J. Math. Physics* **21** 240–249 (1942).

[6] On the ratio of the variances of two normal populations. *Ann. Math. Statist.* **13** 371–388 (1942).

[7] On solutions of the Behrens-Fisher problem, based on the *t*-distribution. *Ann. Math. Statist.* **41** 35–44 (1943).

[8] On a measure problem arising in the theory of non-parametric tests. *Ann. Math. Statist.* **14** 227–233 (1943).

[9] Statistical inference in the non-parametric case. *Ann. Math. Statist.* **14** 305–332 (1943).

[10] A formula for sample sizes for population tolerance limits. (with J. W. Tukey) *Ann. Math. Statist.* **15** 217 (1944).

[11] Note on the use of the tables of percentage points of the incomplete beta function to calculate small sample confidence intervals for a binomial *p*. *Biometrika* **33** 181 (1944).

[12] A note on the Behrens-Fisher problem. *Ann. Math. Statist.* **15** 430–431 (1944).

[13] Non-parametric estimation. I. Validation of order statistics. (with J. W. Tukey) *Ann. Math. Statist.* **16** 187–192 (1945).

[14] The relation of control charts to analysis of variance and chi-square tests. *J. Amer. Statist. Assoc.* **42** 425–431 (1947).

[15] A useful convergence theorem for probability distributions. *Ann. Math. Statist.* **18** 434–438 (1947).

[16] On the problem of similar regions. (with E. L. Lehmann) *Proc. Natl. Acad. Sci. USA* **33** 382–386 (1947).

[17] (Book review). Operating characteristics of average and range charts. *Indust. Quality Control* **5** 13–18 (May 1949).

[18] Completeness, similar regions, and unbiased estimation. Part I. (with E. L. Lehmann) *Sankhyā* **10** 305–340 (1950). *Corrigendum, Sankhyā* **17** 250 (1950).

[19] Theoretical backgrounds of the statistical methods. *Indust. Engin. Chem.* **43** 1292–1294 (1951).

[20] A generalization of the Neyman-Pearson fundamental lemma. (with H. Chernoff) *Ann. Math. Statist.* **23** 213–225 (1952).

[21] An analysis of variance for paired comparisons. *J. Amer. Statist. Assoc.* **47** 381–400 (1952).

[22] A method for judging all contrasts in the analysis of variance. *Biometrika* **40** 87–104 (1953).

[23] Statistical methods for evaluation of several sets of constant and several sources of variability. *Chem. Engin. Progress* **50** 200–205 (1954). Correction, March 1955, same journal.

[24] Completeness, similar regions, and unbiased estimation. Part II. (with E. L. Lehmann) *Sankhyā* **15** 219–236 (1955).

[25] A mixed model for the analysis of variance. *Ann. Math. Statist.* **27** 23–36 (1956).

[26] Alternative models for the analysis of variance. *Ann. Math. Statist.* **27** 251–271 (1956).

[27] Fitting straight lines when one variable is controlled. *J. Amer. Statist. Assoc.* **53** 106–117 (1958).

[28] Experiments with mixtures. *J. Roy Statist. Soc. B* **20** 344–360 (1958).

[29] *The Analysis of Variance*. Wiley, New York, 1959.

[30] Simultaneous interval estimates of linear functions of parameters. *Bull. Int. Statist. Inst.* **38** 245–253 (1961).

[31] Reply to Mr. Quenouille's comments about my paper on mixtures. *J. Roy Statist. Soc. B* **23** 171–172 (1961).

[32] The simplex-centroid design for experiments with mixtures. (including discussion) *J. Roy. Statist. Soc. B* **25** 235–263 (1963).

[33] Multiple testing versus multiple estimation. Improper confidence sets. Estimation of directions and ratios. *Ann. Math. Statist.* **41** 1–29 (1970).

[34] A note on separation of variables. *Techometrics* **12** 388–393 (1970).

[35] Practical solutions of the Behrens-Fisher problem. *J. Amer. Statist. Assoc.* **65** 1501–1508 (1970).

[36] Statistical theory of calibration. *Ann. Math. Statist.* **1** 1–37 (1973).

[37] A note on a reformulation of the *S*-method of multiple comparison. *J. Amer. Statist. Assoc.* **72** 143–146 (1977).

REFERENCES

[1] BERKSON, J. (1950). Are there two regressions? *J. Amer. Statist. Assoc.* **45** 164–180.

[2] COX, D. R. (1960). Scheffé: The analysis of variance. *J. Roy. Statist. Soc. A* **123** 482.

[3] DYNKIN, E. B. (1951). Necessary and sufficient statistics for a family of probability distributions. *Uspekhi Mat. Nauk* (*N. S.*) **6** (41), 68–90. (Also Selected Transl. Math. Statist. Prob. 17–40).

[4] GORMAN, J. W. and HINMAN, J. E. (1962). Simplex lattice designs for multicomponent systems. *Techometrics* **4** 462–487.

[5] HSU, P. L. (1941). Analysis of variance from the power function standpoint. *Biometrika* **32** 62–69.

[6] ISAACSON, S. L. (1951). On the theory of unbiased tests of simple statistical hypotheses specifying the values of two or more parameters. *Ann. Math. Statist.* **22** 217–234.

[7] MILLER, R. G., JR. (1966). *Simultaneous Statistical Inference*. McGraw Hill, New York.

[8] MILLER, R. G., JR. (1977). Developments in multiple comparisons. *J. Amer. Statist. Assoc.* **72** 779–788.

[9] NEYMAN, J. and PEARSON, E. S. (1933). On the problem of the most efficient tests of statistical hypotheses. *Phil. Trans. Roy. Soc. Ser. A* **231** 289–337.

[10] NEYMAN, J. and PEARSON, E. S. (1936, 1938). Contributions to the theory of testing statistical hypotheses. *Statist. Res. Memoirs* **1** 1–37 and **2** 25–57.

[11] OLSHEN, R. A. (1973). The conditional level of the *F*-test. *J. Amer. Statist. Assoc.* **68** 692–698.

[12] RAO, C. R. (1945). Information and accuracy attainable in the estimation of statistical parameters. *Bull. Calcutta Math. Soc.* **37** 81–91.

[13] RAO, C. R. (1947). Minimum variance and the estimation of several parameters. *Proc. Camb. Phil. Soc.* **43** 280–283.

[14] RAO, C. R. (1949). Sufficient statistics and minimum variance estimates. *Proc. Camb. Phil. Soc.* **45** 213–218.

[15] SAVAGE, L. J. (1961). Scheffé: The analysis of variance. *Math. Rev.* **22** 661–7152.

[16] VAURIO, V. W. and DANIEL, C. (1954). Evaluation of several sets of constants and several sources of variability. *Chem. Engin. Progress* **50** 81–86.

[17] WILK, M. B. and KEMPTHORNE, O. (1955). Fixed, mixed, and random models. *J. Amer. Statist. Assoc.* **50** 1144–1167.

[18] WORKING, H. and HOTELLING, H. (1929). Application of the theory of error to the interpretation of trends. *J. Amer. Statist. Assoc.* (March Suppl.) 73–85.

BOX 313, R.D. 2 RHINEBECK
NEW YORK 12572

DEPARTMENT OF STATISTICS
UNIVERSITY OF CALIFORNIA
BERKELEY CA. 94720