Courtesy of Naval Ordnance Test Station, Inyokern, California

By action of the Council of the Institute of Mathematical Statistics, the 1952 Volume of the Annals of Mathematical Statistics is dedicated to the memory of

ABRAHAM WALD

who contributed vitally to the advancement of mathematical statistics through his broad and fundamental research which will continue to influence the development of statistical theory and practice and who will long be remembered as an inspiring and esteemed teacher and colleague.

ABRAHAM WALD, 1902–1950

By J. Wolfowitz

Cornell University

In November, 1950, Abraham Wald left the United States, on leave from his post as Professor of Mathematical Statistics at Columbia University, for an invited tour of Indian universities and research centers. Accompanying him was Mrs. Wald. His book on statistical decision functions had recently been published, and he intended to teach the new theory to Indian statisticians. On December 13, 1950, an Air India plane, lost in a fog, crashed into a peak of the Nilgiris, killing all aboard, among them Professor and Mrs. Wald. Thus, cut off in the prime of activity, died this great statistician, whose work had changed the whole course and emphasis of modern statistics. The personal loss will be felt by his numerous friends, but all must mourn for the statistical discoveries yet unmade which were buried in the flaming wreckage on a mountain side in South India and which will slowly and painfully have to be made by others.

1. Abraham Wald was born in Cluj, Rumania, on October 31, 1902. His father was a small business man, but there was an intellectual atmosphere in the family. His grandfather was a famous rabbi, and the father had considerable intellectual interests. There were five other children in the family, and one brother, Martin, was considered as intellectually gifted as Abraham. Martin was an electrical engineer with many inventions to his credit, and Abraham rendered mathematical help to his brother in a few of the latter's researches. Wald's sisters, inventor brother, their spouses and children, his parents and other relatives, died in German crematoria and concentration camps. One brother only survived and is now in the United States.

Wald was not admitted to the local gymnasium because, as the son of an orthodox Jew, he would not attend school on Saturday, the Jewish Sabbath. He studied by himself and was admitted to the University of Cluj. After graduation from the local university he experienced considerable difficulty in entering the University of Vienna because of religious restrictions. He spent a year in the engineering school at Vienna, but finally was admitted to study mathematics at the University of Vienna.

At the University of Vienna Wald became acquainted with Menger and Hahn. Both soon recognized Wald's abilities and the former became his lifelong friend. Wald became a frequent contributor to, and assistant editor of, the regularly issued reports of the proceedings of Menger's colloquium, from which new results in mathematics issued steadily. Wald's initial work in mathematics was in geometry, his thesis dealing with a question of axiomatics.

Editorial Note: The first three papers published in this issue were presented at the Abraham Wald Memorial Session held September 7, 1951, at the Minneapolis meeting of the Institute.

In spite of his gifts, an academic post in Austria was practically barred to a man of Wald's background. Menger therefore advised Wald to work also in applied mathematics. He also put him in touch with a banker named Karl Schlesinger, who was interested in mathematical economics and to whom Wald gave lessons in mathematics. In his quest for a source of income and an opportunity to work in applied mathematics, Wald met Oskar Morgenstern, who was then director of the Institut für Konjunkturforschung. Morgenstern appreciated Wald's talents and increasingly employed Wald in his institute. Wald's book {24}[1] was one of the products of this employment. Like Menger, Morgenstern became a lifelong friend of Wald's. All three eventually emigrated to the United States. Morgenstern later became a coauthor with von Neumann, of their famous book [1][1] on the theory of games. Wald's most important work (whose basis was laid in {37} completely independently) was to have many points of contact with the theory of games, which in itself became one of Wald's later mathematical interests.

In addition to other papers in mathematical economics (e.g., {21}, {23}, {27}, {28}, {30}) which he wrote in Vienna, Wald also worked on the problem of consistency of the concept of a "Kollektiv" ({20}, {29}, {31}). As the problem was put to Wald and as he solved it, it was a difficult and noteworthy achievement. It was an important step in von Mises' axiomatization of probability and is often cited for this reason. However, the problem was difficult chiefly because of the way it was put. A simple consequence of the strong law of large numbers for identically distributed and independent variables already has as a consequence that almost every sequence of observations is a *Kollektiv*. Thus the modern measure-theoretic approach to the axiomatization of probability theory does away with the need for this pretty piece of work by Wald.

2. Wald came to the United States in the summer of 1938 as a fellow of the Cowles Commission for Research in Economics. In the fall of that year he was released by the Cowles Commission to accept a fellowship of the Carnegie Corporation which was obtained for him by Harold Hotelling, then at Columbia. Hotelling was already then one of the leading American teachers of the modern theory of statistics. His was one of the few voices in the wilderness proclaiming the importance of the new subject. Wald spent a very busy year learning modern statistics by reading and attending Hotelling's lectures. He also began that steady stream of papers in statistics which was never to cease until his death. It was in that academic year (1938–39) that what is probably his most important paper {37} was written, at a time when his knowledge of statistics was rather limited. Wald labored prodigiously, and most of his waking moments during this and the next several years were given to work.

When Wald took up residence at Columbia he knew little of what was then modern statistics or of the content of the courses he later taught. The great difficulty in learning statistics then was due to the obscure manner in which

[1] References in brackets (e.g. [5]) are listed at the end of this paper; references in braces (e.g. {26}) are listed in "The publications of Abraham Wald," pp. 29–33 of this issue.

much of the statistical literature was written. In spite of this, Wald mastered the subject so that the lectures which he gave at Columbia in 1939–40 were noted for their lucidity and mathematical rigor. Students not only flocked to them, but clamored for a record of them. Ralph J. Brookner, who later obtained his doctor's degree under Wald, took notes of the lectures. These were later reproduced for circulation among the students only. However, their fame had spread widely and requests for their purchase came from all over the country. Wald was always anxious to restrict their circulation. He did not want them considered as books (which they certainly were not) and reviewed as such. However, his good nature often prevented strict enforcement of the rule against circulating the notes outside of Columbia University. Thus many of the new generation of American statisticians learned the theory of the analysis of variance from the notes of his course.

These notes on the analysis of variance are typical of Wald's notes and many of his writings. They are rigorous, accurate, and clear, but some of the proofs are clumsy, and the organization of the notes could be improved. The original lectures were given under great time pressure, and there was no time to search for the most elegant proofs, or to plan the organization of the course long in advance. Wald seldom bothered to rework his writings for mathematical elegance or clarity—only new results interested him. Thus the original notes were allowed to stand without alteration, although he would depart from them in his class lectures.

As a lecturer, Wald was clear and lucid. His proofs were always carefully organized, his hypotheses carefully stated. Without omitting any essential detail he would wend his way logically and inexorably from hypothesis to conclusion. He seldom gave an intuitive justification of the theorems, probably because he himself needed it so little.

His lectures in 1939–40 marked the beginning of his teaching career at Columbia, which was terminated only by his death. Hotelling labored prodigiously to find him a permanent post at Columbia and the presence of Hotelling and Wald made Columbia a foremost center of mathematical statistics. Wald stayed at Columbia as a fellow of the Carnegie Corporation until 1941, when he was made assistant professor of economics. Outside recognition won him promotions to an associate professorship in 1943 and then to a professorship in 1944. When Hotelling left Columbia in 1946, a department of mathematical statistics was formed at Columbia with Wald as professor of mathematical statistics and its head. His fame was at its height and students came from all over the world to hear him.

3. Wald's greatest achievement was the theory of statistical decision functions, which includes almost all problems which are the raison d'être of statistics. Every science develops its own techniques, and the development of techniques often gives rise to difficult problems. However, we must never lose sight of the fact that, for example, the solution to a distribution problem, no matter how difficult, is only a device for enabling us to answer some statistical question, and is not of statistical importance per se. Wald brought to statistics a very high degree of mathematical ability and knowledge. Along with this, and in spite of his abstract and theoretical bent and predilections, he never, in any statistical investigation, lost sight of the fact that there was a question to be answered and a decision to be made. Practical people, in many cases dazzled by a mathematical approach they did not well understand, often did lose sight of the final goal, and submitted to having their problems forced into a framework into which they did not really fit. Wald never did this, and much of his statistical success was due to this fact. Wald not only posed his statistical problems clearly and precisely, but he posed them to fit the practical problem and to accord with the decisions the statistician was called on to make. This, in my opinion, was the key to his success—a high level of mathematical talent of the most abstract sort, and a true feeling for, and insight into, practical problems. The combination of the two in his person at such high levels was what gave him his outstanding character. Appropriately enough, his greatest achievement was the direct result of this combination.

What was probably Wald's most important paper ({37}) was written before he knew the details of modern statistical theory. Most of the important notions of his theory of decision functions are already present in this paper. It is true that observations are not taken seriatim but in one sample; this, however, is not too important. The notions of the decision space, of the weight and risk function, of a minimax solution, are all present. Wald proved that the risk function of a minimax solution is constant (under certain restrictions). He operated daringly with Bayes solutions for a priori distributions. At this time statisticians recoiled in horror from Bayes solutions, due to their earlier misuse, and under Fisher's great authority (rightly exercised, I think). Wald made use of Bayes solutions purely as a mathematical tool and without invoking any objectionable statistical connotations. At this time he already had the notions of an admissible test and of a least favorable a priori distribution. It is very saddening to leaf through the pages of this paper and to realize that its author is gone from us.

The paper went almost completely unnoticed. At this time there were few statisticians with the mathematical competence to read the paper. The use of Bayes solutions was a deterrent. Wald did not really emphasize that he was using Bayes solutions only as a tool. His ideas were startlingly new and far off the beaten track. Thus a reviewer of his paper (*Zentralblatt für Mathematik und ihre Grenzgebiete*, Vol. 55 (1941), p. 55) said: "... Es ist gerade der grosse Fortschritt gewesen, dass J. Neyman und E. S. Pearson im Gegensatz zu Th. Bayes ohne zusätzliche Annahmen ausgekommen sind. Die Einführung der beiden Hilfsfunktionen [i.e., the weight function and the a priori distribution function] wirkt demgegenüber wie ein Rückschritt. Die Sätze des Verfassers bleiben leere Theorie, die kurzen Beispiele sprechen keineswegs für die praktische Bedeutung seiner Ansätze."

Wald's personality was also a factor in the reception which his new ideas received. He made no effort to popularize his ideas or to make them accessible to a less mathematical public. I remember well conversations I had with him on the Columbia campus in the spring of 1939. I had recently begun the study of statistics, and tried to oppose his arguments with what I had recently learned. As he convinced me and I grew more and more enthusiastic over what I had just heard from him, Wald proposed several problems for us to work on together. These were all of the most abstract sort, and concerned weakening restrictions under which he proved the theorems of his paper. For example, he was irked by the restriction of compactness and wished to weaken it. Clearly such a program of research was not calculated to popularize his theory. What was required to do so was to apply the theory to outstanding problems. This incident illustrates Wald's personality—always ready to talk about mathematics, but uninterested in popularization and special applications. He was practical-minded in that he always kept statistical ends in sight when working in statistical theory. When the latter was finished to his satisfaction he was not interested in its special application to practical problems.

Wald did not resume serious work on decision functions until 1946. When he did begin work again on decision functions he was also spurred on by the connection between the newly announced results of [1] and his own theory, and by the general interest among economists and others aroused by the theory of games.

In the most general formulation of Wald's theory in {94} one deals with a sequence $X_1, X_2, \cdots$ of (not necessarily independent) chance variables, about whose joint distribution F the statistician in the beginning knows nothing except that F is a member of a given class Ω. There is given a space D of decisions d, one of which the statistician has to make. The loss due to decision d when F is actually the distribution is $W(F, d)$, where W is a given function (the "weight function"). (In {101} the loss is allowed to depend also upon the sequence of observations.) The total loss is the sum of the loss due to the decision made and the cost of the observations. A statistical decision function is a rule which at the ith stage ($i = 1, 2, \cdots$, the first stage is at the outset of the experiment before any observations have been taken) tells the statistician whether or not to take further observations (at the first stage, whether to take any observations), on which chance variables to take observations (if at all), and which decision to make (if no further observations are to be taken). At each stage the decision function is a function of the preceding observations, and is a probability distribution function over the various available possibilities. The actual decision is made by an independent chance mechanism governed by this distribution. The risk function (a function of F) is the expected value of the total loss. Naturally it depends upon the statistical decision function adopted; different decision functions in general have different risk functions.

Comparison of two decision functions is made on the basis of their risk functions. Let $r_1(F)$ and $r_2(F)$ be the risk functions of decision functions δ_1 and δ_2, respectively. If $r_1(F) \leq r_2(F)$ for all F in Ω, δ_1 is said to be at least as good as

δ_2. If the inequality sign holds for at least one F δ_1 is said to be uniformly better than δ_2. A class C of decision functions is said to be complete (essentially complete) if for any decision function δ not in C there exists a decision function δ' in C such that δ' is uniformly better than (at least as good as) δ.

If a probability measure ξ on the space Ω is given, then the decision function δ which minimizes

$$\int r_\delta(F)\, d\xi$$

is called a Bayes solution (for ξ). It used to be thought essential to assume some definite ξ, but the modern point of view rejects this notion. The importance of Bayes solutions for Wald lay in the fact that the totality of all Bayes solutions (or a class derived from this totality) constitutes, under certain circumstances, a complete or essentially complete class. A minimax decision function is one for which

$$\sup_F r_\delta(F)$$

is a minimum. A least favorable a priori distribution ξ_0 is one for which

$$\inf_\delta \int r_\delta(F)\, d\xi_0$$

is a maximum. If $r_\delta(F)$ be regarded as the pay-off function of a zero-sum two-person game between Nature and the statistician, the significance of a minimax decision function and of a least favorable a priori distribution, as minimax strategies in the game, becomes apparent.

The main results of {94} are existence theorems and complete class theorems. The chief tool for proving existence is Wald's Theorem 2.15, of which we shall speak below. Under his weaker restrictions Wald proves the existence of Bayes solutions and of a minimax solution. Under the stronger conditions he proves the existence of a least favorable a priori distribution. In both cases he gives complete and essentially complete classes in terms of Bayes solutions and their closures. It would be repetitious and occupy considerable space here to describe these results in any completeness. A relatively short and easy-to-read exposition is to be found in {102}.

The statistician who wants to apply the results of {94} to specific problems is likely to be disappointed. Except for special problems, the complete classes are difficult to characterize in a simple manner and have not yet been characterized. Satisfactory general methods are not yet known for obtaining minimax solutions. If one is not always going to use a minimax solution (to which serious objections have been raised) or a solution satisfying some given criterion, then the statistician should have the opportunity to choose from among "representative" decision functions on the basis of their risk functions. These are not available except for the simplest cases. It is clear that much remains to be done before the use of decision functions becomes common. The theory provides a rational basis for attacking almost any statistical problem, and, when some computational help is available and one makes some reasonable compromises in the interest of computational feasibility, one can obtain a practical answer to many problems which the classical theory is unable to answer or answers in an unsatisfactory manner. However, for this purpose a relatively simple exposition would suffice to instruct the reader in the rationale of such procedures and it is unnecessary for him to tackle the mathematical details of the theory. The principal value of Wald's book must therefore be for research workers, and the practicing statistician can probably content himself with a reading of the first, and perhaps parts of the last, chapters.

The book is in places not easy to read. Some of the longer arguments could, with some effort, be made more accessible. A number of very minor errors have crept in. As an example of the latter we may cite the following. A principal tool is Theorem 2.15. This theorem states that any sequence of probability measures on a compact metric space contains a subsequence which converges in the ordinary sense to a probability measure. Convergence in the ordinary sense means convergence for every open set whose boundary has probability zero according to the limit distribution. Theorem 2.15 was found by Wald in 1947 and in his book he states that it is related to a theorem of Krylov and Bogolyubov [2]. To describe the latter we first give the theorem of Helly-Bray. This theorem states that if the probability measures ξ_i approach the probability measure ξ_0 in the ordinary sense, and if $\phi(x)$ is any bounded continuous function, then

$$(*) \int \phi(x)\, d\xi_i(x) \to \int \phi(x)\, d\xi_0(x).$$

The theorem of Krylov and Bogolyubov states that, given a sequence of probability measures on a compact metric space, there exists a probability measure ξ_0 and a subsequence ξ_i such that (*) holds for any bounded continuous function of x. Actually it may be shown that Wald's Theorem 2.15 and the theorem of Krylov and Bogolyubov are identical.

It is possible to adopt a definition of convergence of decision functions to a limiting decision function ({94}, pp. 65–6) which will permit a single unified treatment of both absolutely continuous and discrete probability distributions ({94}, Ths. 3.1, 3.2). The elegance of the book would be enhanced by such a treatment.

The book marks the end of the first chapter of work on decision functions. It contains many of Wald's results to date, and in some respects the most general results. However, papers {87} and {90} contain results not in the book. The papers {91}, {97}, {99}, and {101} date after the book and contain entirely new results.

The results of {90} are basic for many purposes in decision theory and deserve some mention. They were obtained in January, 1948, in connection with work on the optimum character of the sequential probability ratio test {84}, but because of various delays were not published until much later (see also {85}). The paper {84} united two currents of Wald's thought and I shall describe it below when I speak of his work on sequential analysis. The result of {84} was obtained by studying Bayes solutions of sequential decision problems involving two decisions. After this paper was completed it was natural to attack the same problem for k decisions. It turns out, as so frequently happens in mathematics, that the proof of a fundamental convexity property, on which the whole proof of {84} rests, and which was very lengthy for $k = 2$ ({84}, Lemma 2), became extremely simple for general k ({90}, Theorem 3.9). From this one obtains partial characterizations of certain regions of decision in the space of a priori (and a posteriori) distributions (ibid., Th. 3.10 and Section 4). Theorem 3.7 gives a complete characterization of Bayes solutions. Theorem 3.8.1 gives a strong continuity result. Equation (3.10) describes a relation which governs the minimum Bayes risk. The contents of Chapter 4 of Wald's book {94} recapitulate part of this paper {90}. The methods of {84} and {90} are entirely in the spirit of Wald's work and their results should be regarded as achievements of his theory.

Wald in his book used decision functions which can be random at each stage, that is, when an observation is taken the decision made depends upon the outcome of a random experiment. Another method of randomization is to randomize once for all at the beginning of the sequence of observations and then to proceed in a nonrandom manner. In {99} the equivalence of the two methods under rather general conditions is shown.

In {97} it is proved, inter alia, that when the number of decisions and possible distribution functions is finite, randomization can be eliminated if the distributions are continuous. At least in this case the role of randomization is to break up "atoms" of probability. The proof rests on a general measure theoretic result proved in {96}, an extension of Lyapunov's theorem. In {101} one can see clearly the intuitive basis of Wald's theorem that the totality of Bayes solutions is complete. Characterizations of admissible solutions are given there. These papers are likely to interest chiefly the research worker.

This discussion would not be complete without a brief statement of Wald's attitude toward the minimax criterion. This attitude has been widely misunderstood. The question concerns a criterion for choosing a decision function from among those in the complete class. Wald often wondered how to give a criterion for choosing a member of the complete class in the absence of any information about which member of Ω is the true distribution. One possible criterion seemed to him to call for the choice of an admissible minimax decision function. This has the advantages of being a very conservative procedure, of being independent of any a priori distribution on Ω, and of having a constant risk function (under certain conditions). However, it would be wrong to assert that Wald strongly advocated the minimax criterion. Thus in his book {94} he states on page 27: "Nevertheless, since Nature's choice is unknown to the experimenter, it is perhaps not unreasonable for the experimenter to behave as if Nature wanted to maximize the risk." However, even this qualified endorsement is tempered by the next sentence: "But, even if one is not willing to take this attitude, the theory of games remains of fundamental importance. . . ." Wald was searching for other criteria, and his last joint work with this writer concerned this problem. He was dissatisfied with known results on the problem and had no great faith in the necessity for the minimax criterion.

The theory of statistical decision functions is the most brilliant part of Wald's work. It is a landmark in statistical theory. Ours is the era of decision functions and its end is nowhere in sight.

4. Wald's sequential probability ratio test was a great statistical achievement. M. Friedman and W. A. Wallis posed to Wald the problem of performing sequentially a test of a hypothesis. Wald's first step was to take the simplest test of all, testing the hypothesis that the frequency function of independent observations is $f_1(x)$, against the alternative that the frequency function is $f_2(x)$, with prescribed maximum probabilities of error. Wald's immediate achievements were twofold: (a) he conjectured that a test procedure based on *constant* limits for S_n, where $S_n = \sum_{i=1}^{n} [\log f_1(x_i) - \log f_2(x_i)]$, would minimize the expected number of observations under each of $f_1(x)$ and $f_2(x)$ (optimum property of the sequential probability ratio test); (b) he obtained simple and excellent approximations for these limits in terms of the prescribed probabilities of error. The first of these is, to my mind, a stroke of genius, and a rare and daring flight of intuition. It was proved only later ({84}) after a number of attempts, and by the methods of Wald's decision theory and the manipulation of Bayes solutions. This was the paper Wald himself liked best. The excellence and simplicity of Wald's approximations to the limits to be set on the cumulative sums are also typical of Wald. His intuition about approximations was uncanny and the rudest of methods in his hands struck gold. On many occasions in connection with other problems I would protest that the methods he proposed were too crude to yield good approximations, only to find that this was not so. In this instance of sequential analysis the availability of excellent approximations was extremely important for future progress. Also very important was the fact that Wald had properly chosen to begin with the simplest problem, because it is there that the results are most startling. From there he went on to problems of composite hypotheses, where the results were not so definitive.

Once fired with the idea, Wald labored incessantly at sequential analysis. For several months he did little else. Most of his own contributions to his book {76} (and these constitute most of the book itself) were discovered then. They were first published in the volume {55}, which was put in the "restricted" category, and made available only to authorized recipients. Wald chafed greatly under this restriction. The first paper he managed to publish was {62}. Here he develops certain basic theorems governing sums of random variables, the number of random variables being itself a chance variable n. Among these results is the distribution of n. The connection between Wald's sequential analysis and the random walk problem was being emphasized. This was followed by more results of this type in {70}. The goal always was to prove the optimum character of the sequential probability ratio test. In {70} Wald obtained a result which implied that if one neglects the excess of the cumulative sums over the prescribed limits then the sequential probability ratio test minimizes the expected number of observations under each distribution. While this is still far from the desired mathematical result it points up the practical value of the sequential probability ratio test. Even if no more than this could be said about the latter the test would still be valuable in practice for most cases. The results of {70} are now largely obsolete, partly through the result of {84} and partly through the work of others.

No proof of the optimum property of the sequential probability ratio test different from {84} has yet been given. In my opinion, an essentially different proof would be an interesting and worthwhile achievement. It is likely that important things might result from a comparison of two essentially different ideas of proof.

Claims have been made elsewhere, chiefly in England, to the invention of sequential analysis. We can clarify the matter simply as follows. The notion of taking observations sequentially was not Wald's. It is probably an old idea, although a number of people lay claim to it, no doubt having discovered it independently. The brilliant and difficult part is the invention of the sequential probability ratio test. This is solely the work of Wald, and no trace of this idea exists in the literature prior to Wald's work.

When permission was granted to Wald to publish on sequential analysis, he described his work in {68} and gave an elementary exposition in {69}. The contents of {68} do not differ much from the book {76}. The book is typical of Wald. Clear, lucid, most of the researches on the subject his own, it was put together hurriedly without too much thought of elegance or of reference to related fields.

Papers {72} and {73} continued his work on sequential analysis, and their titles tell the story of their contents. Papers {75} and {82} evince the interest in the theory of random walk to which he was drawn by his work on sequential analysis. They are based on a paper by Erdös and Kac [3]. Here too the titles tell the story. In {89}, written with his student Sobel, the problem considered is that of deciding sequentially in which of three intervals the mean of a normal distribution with known variance lies. The solution is a simple, practical, and approximate one, and consists of an adaptation of his original solution of the problem of two hypotheses.

In 1943, when he was wholly absorbed in sequential analysis and trying to prove the optimum property of the sequential probability ratio test, we often speculated on an intuitive explanation of this property, which we believed to be true. Wald would wonder whether one could similarly effect great economies in the process of estimation. He early decided that no saving could be effected in the case of estimating the mean (say θ) of a normal distribution with known variance. His criterion was to consider all procedures for which the infimum (with respect to θ) of the confidence coefficient was the same, and measure efficiency by the supremum (also with respect to θ) of the expected number of observations. Several memoranda written for the Statistical Research Group contained a proof of the fact that no saving could be effected, under restrictions which Wald found obnoxious. In 1946 he succeeded in removing these restrictions. The result was proved independently by his student C. M. Stein. Both proofs were complicated. A joint simplified proof was published in {77}. This was a very pretty and significant result. (Its proof has since been considerably simplified and the result itself extended in [4].)

5. Wald's remaining statistical work falls under many headings.

A. Theory of games. While this work is perhaps not strictly statistical it was motivated in part by his theory of decision functions. He became interested in the theory of games on reading [1] and recognized its connection with his own work. He was the first to prove (in {67}) that, if one player of a zero-sum two-person game possesses only a finite number of pure strategies, the game is determined. This result has long since been generalized by Wald himself in {78}, where he needed a similar theorem for his theory of decision functions; conditional compactness with respect to an "intrinsic" metric replaces finiteness. His most general result on the determinateness of a game is Theorem 2.23 of {94}; this result was obtained independently by Karlin [5].

B. Asymptotic results connected with the method of maximum likelihood and the likelihood ratio test. These include papers {43}, {46}, {48}, {57}, {81}, and {88}. The last of these contains the neatest and most expeditious of the rigorous proofs of the consistency of the maximum likelihood estimate yet available in the literature.

C. Work on nonparametric inference. Nonparametric inference is that branch of statistics which deals with the case where the unknown distributions cannot be specified by the values of a finite number of parameters. His first paper on the subject was {34}. Here a completely unknown (except that its continuity is assumed) distribution function is estimated by a "belt" which is a function of the observations. The estimation is made in the sense of Neyman [6]. The paper {34} was followed shortly by {40}. The title tells the story of the contents here; the test proposed was based on the number of runs. The notion of consistency of a test was introduced in a manner generalizing the notion of the consistency of a point estimate.

In {50} and {52} Wald took up the notion of tolerance intervals (due to Wilks [7]). A tolerance interval is an interval which is a function of the observations in one sample and which, with prescribed confidence coefficient, contains at least a specified fraction of another sample. By a very simple device using conditional distributions he extended Wilks' univariate solution to the multivariate case in {52}. Paper {50} is based on normal approximation theory.

In {59} and {64} tests for a number of nonparametric problems are based on permutations of the observations. Limiting distributions are obtained. Much further remains to be done. (See also [13].)

D. Miscellaneous statistical results.

1. {38} gives confidence limits for the intraclass correlation coefficient in what is now called Model II of the analysis of variance. The method is forth-

right and simple. This was one of Wald's earliest statistical papers, and the problem was put to him by Hotelling. The result is generalized in {45}.

2. {41} gives a consistent method of fitting a line when both variables are subject to error. (In ordinary least squares theory only one variable is subject to error.) For a recent result on this subject the reader is referred to [14] and [15].

3. {44}, written with his student Ralph Brookner, is one of his papers on multivariate analysis. It gives the distribution, under certain conditions, of the statistic derived by Wilks by use of the likelihood ratio to test the independence of groups of jointly normally distributed chance variables. The statistic is the quotient of products of determinants of sample correlation coefficients.

4. {49}, written with H. B. Mann, considers optimizing the procedure of testing goodness of fit by means of the χ^2 test. A metric is introduced into the totality of alternatives. The authors decide to maximize the smallest value of the power of the test for alternatives which are neither too far away from the null hypothesis (such alternatives are easy to detect) nor too near the null hypothesis (such alternatives it is almost hopeless to detect). They conclude that the class intervals should all have the same probability under the null hypothesis and when the sample size N is large should be of the order $N^{1/5}$ in number. (A precise result is described.) Although this is a problem of great importance little new has been added to their result.

5. {51} generalizes a result of P. L. Hsu [8] on the power function of the analysis of variance. This work has since been continued by others (Hunt and Stein, Lehmann, and the writer). The proof of {51} was shortened in [9].

6. {54} tries to give a measure for the efficiency of a design for testing a linear hypothesis. This is a pioneer effort. No further work on this important problem has to my knowledge been done.

7. {60} treats the problem of classifying an individual into two groups. Each group has the same p characteristics jointly normally distributed with a common covariance matrix, but with differing means. The statistic actually employed is a slight modification of the one given by the probability ratio test. The problem is one of finding a distribution, and is not solved completely. (Wald had to be urged to publish this paper.) The problem is one of great practical importance and work is being done on it at present (see [12]).

8. {65} was a venture of Wald's into statistical quality control. It was stimulated by a paper by Dodge [10]. It contains, inter alia, a sampling inspection scheme which guarantees a prescribed lower bound on the outgoing quality, and, in the case of statistical control, requires a minimum of inspection.

9. {83} contains some results on a problem first raised by Neyman and Scott [11], that of estimating a fixed number of (structural) parameters when each successive observation depends also upon a new (incidental) parameter. This is a problem of great importance which is still far from being solved.

The above list of papers is far from exhaustive. It illustrates the multitude of problems attacked by Wald and his amazing fertility. Many of the problems are of Wald's own formulation. His amazing gift for the "practical" theory, coupled with his mathematical powers, is to be seen everywhere in these papers.

6. Wald himself was a man rather completely immersed in his work. His other chief interest was his family. He had played the violin before coming to the United States, but under pressure of work had never resumed playing. Hiking was his chief diversion; he was an indefatigable walker and some of his joint papers were worked out on long hikes. He was helpful and kind to his students, and more than generous in sharing credit with others. To the end he was modest and unassuming, with an unusual aversion to all forms of controversy.

In 1941 he married Lucille Lang, who met her untimely end with him. Two children survive—Betty, born in 1943, and Robert, born in 1947. He was much absorbed in his family life and very devoted to the children.

It is difficult to see the limits of Wald's influence on our young science of mathematical statistics. How tragic it is that we have lost him in his prime, when there is so much yet to do, and when there are still so few who can follow in the trails he blazed.

REFERENCES

[1] J. von Neumann and O. Morgenstern, *Theory of Games and Economic Behavior*, Princeton University Press, 1944.

[2] N. Krylov and N. Bogolyubov, "La théorie générale de la mesure dans son application à l'étude des systèmes dynamiques de la mécanique non-linéaire," *Annals of Mathematics*, Vol. 38 (1937), pp. 65–113.

[3] P. Erdös and M. Kac, "On certain limit theorems of the theory of probability," *Bull. Am. Math. Soc.*, Vol. 52 (1946), pp. 292–302.

[4] J. Wolfowitz, "Minimax estimates of the mean of a normal distribution with known variance," *Annals of Math. Stat.*, Vol. 21 (1950), pp. 218–230.

[5] S. Karlin, "Operator treatment of minmax principle," *Contributions to the Theory of Games*, Princeton University Press, 1950.

[6] J. Neyman, "Outline of a theory of statistical estimation based on the classical theory of probability," *Philos. Trans. Roy. Soc. London, Ser. A*, Vol. 236 (1937), pp. 333–380.

[7] S. S. Wilks, "Determination of sample sizes for setting tolerance limits," *Annals of Math. Stat.*, Vol. 12 (1941), pp. 91–96.

[8] P. L. Hsu, "Analysis of variance from the power function standpoint," *Biometrika*, Vol. 32 (1941), pp. 62–69.

[9] J. Wolfowitz, "The power of the classical tests associated with the normal distribution," *Annals of Math. Stat.*, Vol. 20 (1949), pp. 540–551.

[10] H. F. Dodge, "A sampling inspection plan for continuous production," *Annals of Math. Stat.*, Vol. 14 (1943), pp. 264–279.

[11] J. Neyman and E. L. Scott, "Consistent estimates based on partially consistent observations," *Econometrica*, Vol. 16 (1948), pp. 1–32.

[12] H. L. Harter, "On the distribution of Wald's classification statistic," *Annals of Math. Stat.*, Vol. 22 (1951), pp. 58–67.

[13] G. E. Noether, "Asymptotic properties of the Wald-Wolfowitz test of randomness," *Annals of Math. Stat.*, Vol. 21 (1950), pp. 231–246

[14] J. Neyman and Elizabeth L. Scott, "On certain methods of estimating the linear structural relation," *Annals of Math. Stat.*, Vol. 22 (1951), pp. 352–361.

[15] Elizabeth L. Scott, "Note on consistent estimation of the linear structural relation between two variables," *Annals of Math. Stat.*, Vol. 21 (1950), pp. 284–288.