The Annals of Statistics
1984, Vol. 12, No. 2, 406–415

THE RESEARCH OF JACK KIEFER OUTSIDE THE AREA OF EXPERIMENTAL DESIGN

BY LAWRENCE D. BROWN[1]

Cornell University

This article is a survey of Jack Kiefer's published works in areas other than design of experiments. The approximately 50 articles can be divided into categories by subject matter. These categories will be discussed individually in the following review. At times Jack worked in several areas simultaneously and ideas from one area sometimes influenced work in another. Some of these influences are noted in the following survey, as the categories are discussed in roughly the order Jack began work in them.

Collaborators. The majority of Jack's earlier papers were written jointly with his former thesis advisor, Jack Wolfowitz. Six of these were a three-way collaboration including Aryeh Dvoretzky. Jack was brimful of curiosity and good ideas and a desire to assist and share. In all, this led to joint work with 21 different collaborators throughout his career. The extent of this collaboration can be noted from the bibliography. No attempt will be made here to unravel the who-did-what-and-why of these collaborations.

Foundations and decision theory. Paper [7] distills the first half of Jack's 200 page Ph.D. thesis into a 6 page article. The main results here establish the existence of minimax and admissible procedures under regularity conditions weaker than those assumed in the fundamental book of Wald (1950). Jack's methods and results have since become antiquated by those of LeCam and others. See especially LeCam (1955).

What still remains very much of note is that this work signals a commitment to the philosophy of Wald's decision theoretic approach. This philosophy transcends the important concepts of loss, risk, admissibility, minimaxity, etc. It holds that one must weigh carefully—in a precise mathematical formulation—the frequential consequences of any statistical course of action.

A quotation from Dvoretzky, Kiefer, and Wolfowitz [2], describes this philosphy in relation to the loss function though it applies equally to the formulation of a probabilistic model, specification of the action space, and so forth:

> "It may be objected that our method requires one to specify the [loss function] and that this function may be unknown or difficult to give. We wish to emphasize that the need for a [loss] function, W, is inevitable in the sense that any method which does not explicitly use a function W simply uses one implicitly. Thus one who selects a method of solving the $\cdots$ problem which ostensibly has the advantage of not requiring the specification of W is simply relinquishing control of W and may be implicitly using a W of which he would disapprove $\cdots$. It is difficult to see what advantage can accrue $\cdots$ from deliberately burying [one's] intellectual head in the sand."

Received December 1983.
[1] Preparation of this manuscript supported in part by N.S.F. Grant MCS 8200315.

Institute of Mathematical Statistics is collaborating with JSTOR to digitize, preserve, and extend access to *The Annals of Statistics.* www.jstor.org

The minimax principle is one feature of Wald's theory. Kiefer, as well as Wolfowitz, accepted it only as, "a *possible* principle ··· [which] ··· might be the course of a very conservative statistician." (Italic, mine. Quotation from Wolfowitz (1951) page 461.) Indeed, Jack's first published article, [1], consists of three relatively elementary examples of problems where one would clearly prefer not to use the minimax procedure.

As I will note, the minimax principle reappears many times in various contexts in Jack's work, sometimes in rather surprising ways.

The minimax principle is, of course, a central element of [19]. Theorems establishing the existence (under general regularity conditions) of minimax admissible procedures are called Hunt-Stein theorems. This paper unifies and extends the contemporaneous Hunt-Stein theorems. It remains a classic in spite of the fact that its methods—though not its results—can be considered inferior to those developed later by Huber. (Huber's work on the Hunt-Stein theorem remains largely unpublished. A brief sketch of the method in relation to testing problems appears in [45]. Another such sketch and some important related results appear in Bondar and Milnes (1981).)

The fact that this paper extends the Hunt-Stein theorem to sequential problems should be noted for its own merit, and also as a measure of Jack's wide interests and his determination to extend theoretical results to their statistically natural boundaries.

Jack's insistence on establishing clearly the frequential consequences of any statistical action persisted. In the series of papers [67], [69], [70] (with C. Brownie), and [71] he applied this insistence to the frustrating dilemma of conditional confidence. Jack's results leave an optimistic feeling that a frequential discussion of conditional confidence is possible, useful, and even perhaps natural. But, as expressed in Brown (1978), I think that a finished satisfactory frequential theory will have to diverge at some point from that proposed in these papers.

This immersion in Wald's theory had earlier led almost naturally to a climax in Jack's career. Given an insistence on establishing a priori the frequential consequences of statistical procedures, it was natural to look at problems of experimental design, as indeed Wald (1943) himself had done.

The first paper to result was the remarkable paper, [22], which presents the possibility of using randomized designs and discusses optimality properties in this context. After this began the climactic series of papers—several written jointly with Jack Wolfowitz. Kiefer and Wolfowitz skillfully brought together and refined various existing optimality criteria for judging experimental designs. In the course of this they discovered a formal parallel with the classical minimax theorems. The development of this parallel and the years of interesting consequences belong to the essay to follow by Henry Wynn about Jack's work on experimental design. Here, I want only to emphasize how the initial achievements in design related naturally to Jack's general statistical outlook and to his formal decision theoretic background.

Before turning to the next major area, I will mention Jack's commentary papers, [48] and [73], for these express a side of Jack's statistical personality very familiar to those who knew him, but not so evident in his published papers.

He was extremely concerned with the question of how well procedures justified by elegant theories actually work in real life. For example, how well do asymptotically optimal procedures perform for experiments involving realistic sample sizes?—or, what happens in design and analysis when fitting a linear or quadratic regression to a response curve which is not linear or quadratic?

Sequential analysis. Wald's decision theory was developed in tandem with his theory of sequential statistical procedures. Analogously, the second half of Jack's doctoral thesis was an attempt to construct optimal sequential procedures for problems whose parameter spaces have three or more points. It turns out that there is no precise characterization possible here like that of the familiar Wald-Wolfowitz theorem for the sequential probability ratio test (SPRT) for testing a simple null hypothesis versus simple alternative. Thus, this second part of the thesis was never directly adapted for publication. But, what Jack discovered there appears in a very significant way in Kiefer and Weiss [18], which concerns properties of the generalized SPRT, and also in Kiefer and Sacks [37], which is discussed in more detail below.

Aside from work in inventory theory which will be described later, Jack's first published paper in sequential analysis involved a problem of a very different sort from those in his thesis. Kiefer and Wolfowitz [4] proposed what has come to be known as the "Kiefer-Wolfowitz procedure". This is a consistent procedure for sequentially locating the maximum of an unknown unimodal regression function. The stimulus for this work was of course the earlier paper of Robbins and Munro (1951) describing a consistent sequential procedure for locating the zero of a regression function.

The problem for the Kiefer-Wolfowitz procedure is statistical—the values of the regression function are observed with statistical error. What happens if the regression function can be observed exactly? Jack's involvement with this question goes back to his master's thesis, written at M.I.T., but the final results did not appear until his paper [8], with generalizations in [20]. There is still a problem to be solved—namely, where to place the observations in order to have the most precise final statement concerning the location of the maximum. This is not a classical type of statistical problem. Nevertheless, Jack was able to give it a minimax formulation (again the minimax principle!) and to use some clever analysis (as usual) to produce a precise description, involving the Fibonacci numbers, of ε-minimax rules and of asymptotically optimum rules. Even now there is no comparably complete solution to the statistical problem attacked by Kiefer and Wolfowitz—indeed, it is not even known whether in their general formulation there can exist a procedure which achieves the optimal stochastic rate of approach to the unknown maximum. (If not, maybe it is necessary to restrict the formulation somewhat?)

The pair of papers, [9] and [10], by Dvoretzky, Kiefer, and Wolfowitz, considers statistical problems concerning continuous time stochastic processes—for example, a Wiener process with unknown drift or a Poisson process with unknown intensity. The first paper contains a proof of the optimum property of the SPRT

for such problems. The second paper treats questions of estimation. Primarily, it characterizes those situations where a fixed sample size procedure is optimal among all sequential procedures.

A feature of these papers which is of note—apart from the important basic results themselves—is the understanding that discrete time processes can be profitably approximated by continuous time ones. Such approximations are now an essential part of modern methodology for attacking sequential problems. Dvoretzky, Kiefer and Wolfowitz wrote, "There are many cases $\cdots$ in which an exact determination of the optimal procedure is possible in the continuous case [but not in the discrete case]. Thus even when treating the discrete case the continuous case $\cdots$ may be used to derive approximations." ([9, page 255]). My impression is that this understanding was an important motivating force and building block for a second climax in Jack's work, which I'll describe later.

An earlier place where this realization proved useful in Kiefer's work was in the monograph [102] on ranking and selection by Bechhofer, Kiefer, and Sobel. This manuscript was more than 10 years in preparation. It was one of the few projects on which Jack felt himself to be distinctly the junior author, but he worked on it very industriously, especially to create and perfect sequential arguments involving approximation by continuous time stochastic processes.

This brings us to the monumental paper of Kiefer and Sacks, [37]. This paper was not the first to consider asymptotic sequential properties—there were important precursors by Chernoff (1959) and by Schwarz (1962)—but it was the first comprehensive treatment. Within the models considered, it clearly settled the issue of the asymptotic relationship between expected sample size and probability of error for asymptotically optimal procedures. In fact it constructed such procedures—that is, procedures which are asymptotically Bayes (as the cost per observation approaches zero) simultaneously for all priors whose support is the full parameter space.

It is characteristic of Jack's work that this paper began as an attempt to settle an even broader question, the construction of asymptotically optimal procedures in problems involving a sequential choice of design. After they began work on this broader question, Kiefer and Sacks realized that the existing formulation and results for the basic (nondesign) problem were inadequate for their purposes. After solving this basic problem they then turned to the original design question and in the second half of [37] they solved this problem as well.

Here is one main result from the first half of this paper. Consider the problem of sequentially testing a composite null hypothesis, Ω_0, versus a composite alternative, Ω_1, with the cost of each observation being c. The potential observations are independently and identically distributed (i.i.d.) with density f_ω, some $\omega \in \Omega_0 \cup \Omega_1$. Introduce the Kullback-Liebler distance

$$\lambda_i(\omega) = \inf\{E_\omega(\log(f_\omega(X)/f_{\omega_j}(X))): \omega \in \Omega_i,\ \omega_j \in \Omega_j,\ j \neq i\}, \quad i = 0, 1. \tag{1}$$

Assume

$$\lambda_i = \inf\{\lambda_i(\omega): \omega \in \Omega_i\} > 0, \quad i = 0, 1. \tag{2}$$

Then, under suitable mild regularity conditions there is a procedure satisfying

$$\sup\{P_\omega(\text{terminal error}): \omega \in \Omega_0 \cup \Omega_1\} = o(c \mid \log c \mid) \tag{3}$$

and

$$E_\omega(\text{terminal sample size}) \leq (1 + o(1)) \mid \log c \mid / \lambda_i(\omega) \quad \text{uniformly for } \omega \in \Omega_i,\ i = 1, 2. \tag{4}$$

This procedure is a modification of the weight function procedure introduced in Wald (1947). If G is any prior distribution supported on all of $\Omega_0 \cup \Omega_1$ then the Bayes procedure for G achieves exactly the bounds on the right of (3) and (4). It follows that the modified weight function procedure is asymptotically Bayes.

The power and elegance of such a result often leads one to overlook its deficiencies. The primary deficiency is that the given asymptotically optimal procedure does not seem to be satisfactorily near optimality for practically realistic values of c. Kiefer and Sacks noted this deficiency. Others have since attempted to provide a theoretical basis for defining an improved asymptotically optimal procedure. Lorden (1972), (1976) and Zerdy (1980), both former doctoral students of Kiefer, have investigated related questions and have made some progress on this issue. Another possible deficiency lies in the assumption (2) which specifies that the null and alternative hypothesis be effectively separate. Analogous results without this assumption have not been proved, but some progress may be seen in Bickel and Yahav (1972) and works cited there, and from a somewhat different perspective in Brown, Cohen, and Samuel-Cahn (1983).

Writing style. In terms of general style, [37] (considered above) is an excellent example of Jack's craftsmanship—so this seems a good place to pause and consider some general characteristics of Jack's writing.

Jack was always bursting with useful ideas. In writing a paper he tried to share them all and to pass along a wealth of information on the relationship of results and methods to other important related works, on possible generalizations, and on alternate formulations and methods of proof.

This is immediately obvious to anyone who glances at this work, or at almost any other of Jack's major works. Some other sides of Jack's mathematical personality are not so immediately apparent. Jack organized long and complicated mathematical arguments in a startlingly logical and concise way. He had a special skill for pulling apart complex proofs and putting them together in a step by step fashion, and for condensing routine arguments to their bare outlines while, on the other side, displaying and explaining fully all technical or conceptual innovations. All this requires careful reading to see, mainly because difficult mathematics, even when presented brilliantly, is still difficult and demands careful study.

Inventory theory and queueing theory. Many types of problems are essentially sequential in character. Jack's background in both Wald's decision

theory and sequential analysis prepared him well to look at some of these problems, and he did so with important consequences jointly with Dvoretzky and Wolfowitz. The following description of the work in inventory theory is borrowed from the introduction in [103].

Dvoretzky, Kiefer, and Wolfowitz [2], on the inventory problem, suggest a development which applies to the broader class of settings now referred to as discounted dynamic programming, in which an optimal policy is sought for adjusting a chance process so that the sum of discounted rewards over many time periods is maximized. A particularly simple form of policy suggested in earlier work of Arrow, Harris, and Marschak (1951) is shown in Dvoretzky, Kiefer, and Wolfowitz [11] to be optimum under certain conditions. The case in which the chance law that governs the process is unknown and thus, in effect, has to be estimated as time periods pass, is treated in Dvoretzky, Kiefer, and Wolfowitz [3].

Kiefer and Wolfowitz [12] contains the first general systematic treatment of the multiserver queue. In this paper, "general results are obtained on the convergence in probability of waiting times and other quantities of interest in queueing systems." The methods yield results on random walks, such as characterization of random walks S_n based on i.i.d. summands for which $E(\max S_n)^k < \infty$ ([14]).

Nonparametric tests and related distribution theory. A colleague once told me that, of all his accomplishments, Jack was most proud of those related to the Kolmogorov-Smirnov problem. Whether or not Jack ever expressed such an opinion, he had every right to be proud of his work in this area.

This work began with the joint paper of Kac, Kiefer, and Wolfowitz [13]. The paper first develops the distribution theory for tests of normality based on the Kolmogorov-Smirnov statistic and the von Mises statistic. Then the minimax principle again appears. It is convincingly demonstrated—via an asymptotic minimax formulation—that these tests are vastly preferable to the standard χ^2-test. Roughly, if the χ^2 test of size $\alpha < \frac{1}{2}$ requires N observations (N large) to obtain a certain power β against all alternatives at a distance (depending on α, β, N) from the null hypothesis, then the proposed Kolmogorov-Smirnov type test requires only $O(N^{4/5})$ observations.

The minimax principle is also a cornerstone of Dvoretzky, Kiefer, and Wolfowitz [15] and Kiefer and Wolfowitz [25]. These papers establish the asymptotic minimax property of the sample cumulative distribution function (CDF) as an estimate of the true CDF, in terms of the Kolmogorov-Smirnov distance and a variety of other measures of loss. Paper [15] deals with univariate problems; paper [25] treats the multivariate case and, in passing, develops a streamlined heuristic argument and proof which is useful also in the univariate case. In a sense hinted in these papers as well as in Kiefer and Wolfowitz [16], and made much more transparent by contemporary research, the sample CDF plays the role for this nonparametric problem that the maximum likelihood estimator plays for nice parametric problems.

This asymptotic minimaxity result for estimation is of a vastly superior nature compared to the crude order of magnitude argument of [13] for the testing

problem. At the present time there is still no completely satisfactory asymptotic minimaxity result for the testing problem.

Kiefer and Wolfowitz returned to this asymptotic question in [65] and [68] where they give analogous results for the situation when the cumulative distribution function involved is known to be concave or to be convex, or to satisfy certain other similar conditions.

In the univariate case the Kolmogorov-Smirnov statistic is distribution free (so long as the underlying distribution is continuous). In the multivariate case it is not. Kiefer and Wolfowitz were not the first to note this fact, but they were the first to do something concrete about it. In [21] they established a uniform bound on the tail of the Kolmogorov-Smirnov statistic.

Here is their crude bound: The (one-sided) statistic is $T_n = \sup\{\sqrt{n}\ (F_n(x) - F(x)): x \in R^k\}$, where F denotes the underlying CDF on R^k, and F_n denotes the sample CDF. Then, for some $\alpha = \alpha(k) > 0$ and $c = c(k) < \infty$,

$$P(T_n > r) \leq ce^{-\alpha r^2} \quad \forall n, r, F. \tag{5}$$

This bound is certainly very crude but it did, at least, suffice to prove the existence of the limiting distribution of T_n as a Gaussian process. Although Kiefer and Wolfowitz established the existence of this limiting distribution, no one as yet has found any explicit form for it. Indeed, there is only one nontrivial case where reasonably accurate bounds are known. This is when $F = U_2$, the uniform distribution on the unit square. Here the limiting distribution is that of a pinned Brownian sheet, and fairly close lower and upper bounds on the limiting distribution appear in Goodman (1976) and Cabaña and Wschebor (1982), respectively. In a classic paper, [32], Kiefer greatly improved this bound. He proved that for all $\varepsilon > 0$ there is a $c = c(k, \varepsilon)$ such that

$$P(T_n > r) \leq ce^{-2(1-\varepsilon)r^2} \quad \forall n, r, F. \tag{6}$$

This bound, (6) has since been widely generalized to other Gaussian processes (see, for example Marcus and Shepp (1972) or Révész (1976a, b)), but not until now basically improved, except in the case $F = U_2$ mentioned above. This bound is of interest on its own merits and also because, as Jack noted, it enables one to establish a law of the iterated logarithm for the multivariate Kolmogorov-Smirnov statistic.

As a footnote I'm pleased to mention, because Jack would have been pleased to hear, that Robert Adler and I were recently able to capitalize on a remark in [21] to refine (6) to the bound: There is a $c = c(k)$ such that

$$P(T_n > r) \leq cr^{2(k-1)}e^{-2r^2} \quad \forall n, r, F. \tag{7}$$

In a series of articles, more probabilistic than statistical in nature, Jack then pursued various distributional questions involving the one-dimensional sample CDF or, in [24] and [30] (with Blum and Rosenblatt), the distance between several independent one-dimensional CDF's.

Jack was clearly fascinated as well by the search for the best methodology. Paper [24] is apparently the first nonparametric test for equality of several

sample CDF's. (The test uses an appropriate generalization of the Kolmogorov-Smirnov statistic.) In this paper the methodology is to reduce the limiting question to one about several independent Brownian bridges. Paper [30] works out the limiting distribution of a statistic for testing independence, and this time the methodology is a characteristic function argument. In [27], Jack published a largely expository paper describing how similar results could be derived via an argument based on the differential generator of the process.

Next comes an important paper, [47], on laws of the iterated logarithm for distributions involving the sample quantiles. Here is one basic result from this paper. Let $Y_{p,n}$ denote the sample pth quantile based on a sample of size n from a uniform distribution on (0, 1). Let $R_n(p) = Y_{p,n} - p + (F_n(p) - p)$, where F_n denotes the sample CDF. Let $\sigma_p^2 = p(1-p)$. Then

$$\text{(8)} \qquad \limsup_{n\to\infty}[R_n(p)/(32/27)^{1/4}\sigma_p^{1/2}n^{-3/4}(\log\log n)^{3/4}] = 1, \quad \text{w.p.1.}$$

Related questions were profitably pursued also in [51], [52], and [54].

In the course of studying sample quantiles Jack first encountered still another methodology for investigating the sample CDF—the Skorohod-Strassen embedding, or strong invariance principle. Paper [50] contains important results about this embedding.

Jack was methodologically prepared for a second major climax in his career. In addition, his statistical background provided the problem: the ordinary Skorohod embeddings supply a satisfactory approximation to the sample CDF for a single large n, but they do not provide the right joint distribution for several different large n at once. Jack's familiarity with the asymptotics of sequential analysis emphasized the importance of this joint distribution.

Here is the background for Jack's formulation. Let $X_1, \cdots$, be i.i.d. real random variables with continuous CDF, F. Take F to be uniform on (0, 1) without loss of generality. Let $T_n(x) = n^{1/2}(F_n(x) - x)$. Let $\{B(x)\}$ denote the Brownian bridge on [0, 1] and $\{W(x_1, x_2)\}$ the Brownian sheet on $[0, \infty) \times [0, \infty)$. Note that $B(x) = n^{-1/2}[W(x, n) - xW(1, n)]$ for each $n \in (0, \infty)$. Brillinger (1969) had earlier proved that, for each fixed n,

$$\text{(9)} \qquad |B(x) - T_n(x)| = O_p(n^{-1/4}(\log n)^{1/2}(\log\log n)^{1/4}), \quad \forall x \in [0, 1].$$

Jack now introduced in [56] the Gaussian process

$$\text{(10)} \qquad K(x, t) = W(x, t) - xW(1, t)$$

on $[0, 1] \times [0, \infty)$. This process has since been called, "the Kiefer process." (Note that $B(x) = n^{-1/2}K(x, n)$.) The process, K, can be understood as a Brownian sheet conditioned to be zero on $\{1\} \times [0, \infty)$.) He then used all of his accumulated methodological knowledge and ingenuity. In a paper exploding with significant detail he established the useful uniform bound

$$\text{(11)} \qquad |n^{-1/2}K(x, n) - T_n(x)| = O_p(n^{-1/6}(\log n)^{2/3}), \quad \forall\, x, n \in [0, 1) \times [0, \infty).$$

Jack conjectured the validity of a much better bound. A few years later Kómlos,

Major, Tusnády (1975) verified that

$$(12) \quad |n^{-1/2}K(x, n) - T_n(x)| = O_p(n^{-1/2}(\log n)^2), \quad \forall\, x, n \in [0, 1) \times [0, \infty).$$

(This result incidentally also improves on (9).) An analogous result for multivariate random variables (whose study was initiated in [21], as has been described) appears in Révész (1976a, b).

Multivariate analysis. Jack's activity in this area covers only a brief fragment of his career, although his work here has roots into, and connections with, other of his interests. These connections are especially strong in the important paper by Giri and Kiefer [38] in which the local optimality criteria discussed are intimately related to those Jack had exploited in design papers such as [23] and [31]. The connections are perhaps weakest in the startling result in Kiefer and Schwartz [42], which establishes the Bayes character (for rather unusual appearing priors) of various standard multivariate tests.

Giri, Kiefer and Stein [36] and Giri and Kiefer [38] present very complex proofs of minimaxity over invariant shells of Hotelling's T^2 and of R^2 only in the very simplest nontrivial case. The question for other sample sizes and dimensions resisted Jack's attempts. A partial solution appears in Šalaevskiĭ (1968, 1969). It is necessary to approach these multivariate problems on an individual basis because they necessarily fall outside the scope of a general Hunt-Stein theorem. For example, as Jack pointed out, it is still an open question whether the best invariant (or, "equivariant", depending on terminology) estimator of the population CDF is a minimax estimator, even though it was proved already in Dvoretzky, Kiefer, and Wolfowitz [15] to be asymptotically minimax.

Additional comments. I have tried to describe how the minimax principle gently guided much of Jack's work. It could be said that Jack's professional life was also characterized by a form of this principle—he accomplished a maximum amount in a minimum of time.

His published works are a remarkable accomplishment. But it is not primarily through these that I remember him, and I know that my feelings in this are shared by many others.

Jack was above all a loyal friend, a delightful comrade and colleague, a statistician of unerringly high standards and aspirations for himself and the profession, and most important, a deeply and sincerely warm and nice person.

REFERENCES

ARROW, K., HARRIS, T. and MARSCHAK, J. (1951). Optimal inventory policy. *Econometrica* **19** 250–272.

BICKEL, P. and YAHAV, Y. (1972). On the Wiener process approximation to Bayesian sequential testing problems. *Proc. Sixth Berkeley Symp. Math. Statist. Probab.* **1** 57–84.

BONDAR, J. V. and MILNES, P. (1981). Amenability: A survey for statistical applications of Hunt-Stein and related conditions on groups. *Z. Wahrsch. verw. Gebiete.* **57** 103–128.

BRILLINGER, D. (1969). An asymptotic representation of the sample d.f. *Bull. Amer. Math. Soc.* **75** 545–547.

BROWN, L. D. (1978). A contribution to Kiefer's theory of conditional confidence procedures. *Ann. Statist.* **6** 59–71.

BROWN, L. D., COHEN, A. and SAMUEL-CAHN, E. (1983). A sharp necessary condition for admissibility of sequential tests. *Ann. Statist.* **11** 640–653.

CABAÑA, E. M. and WSCHEBOR, M. (1982). The two parameter Brownian bridge: Kolmogorov inequalities and upper and lower bounds for the distribution of the maximum. *Ann. Probab.* **10** 289–302.

CHERNOFF, H. (1959). Sequential design of experiments. *Ann. Math. Statist.* **30** 755–770.

GOODMAN, V. (1976). Distribution estimates for functionals of the two parameter Wiener process. *Ann. Probab.* **4** 977–982.

KÓMLOS, J., MAJOR, P. and TUSNÁDY, G. (1975). An approximation of partial sums of independent R.V.'s and the sample DF, I. *Z. Wahrsch. verw. Gebiete* **32** 111–131.

LECAM, L. (1955). An extension of Wald's theory of statistical decision functions. *Ann. Math. Statist.* **26** 69–81.

LORDEN, G. (1976). 2-SPRT's and the modified Kiefer-Weiss problem of minimizing an expected sample size. *Ann. Statist.* **4** 281–292.

LORDEN, G. (1972). Likelihood ratio tests for sequential k-decision problems. *Ann. Math. Statist.* **43** 1412–1427.

MARCUS, M. B. and SHEPP, L. A. (1972). Sample behavior of Gaussian processes. *Proc. Sixth Berkeley Symp. Math. Statist. Probab. II*, 423–439.

RÉVÉSZ, P. (1976a). On strong approximations of the multidimensional empirical process. *Ann. Probab.* **4** 729–743.

RÉVÉSZ, P. (1976b). Three theorems of multivariate empirical processes. *Lecture Notes in Math.* **566** *Empirical Distributions and Processes* 106–126. Selected papers from a meeting at Oberwolfach, March 28–April 13, 1976. (Ed. by A. Dold and B. Eckmann). Springer, New York.

ROBBINS, H. and MUNRO, S. (1951). A stochastic approximation method. *Ann. Math. Statist.* **22** 400–407.

ŠALAEVSKIĬ, O. V. (1968). The minimax character of Hotelling's T^2-test. *Dokl. Akad. Nauk SSSR* **180** 1048–1050.

ŠALAEVSKIĬ, O. V. (1969). Minimax character of Hotellings T^2-test I. *Zap. Naučn. Sem. Leningrad Otdel. Mat. Inst. Steklov* (*LOMI*) **13** 138–182.

SCHWARZ, G. (1962). Asymptotic shapes of Bayes sequential testing regions. *Ann. Math. Statist.* **33** 224–236.

WALD, A. (1943). On the efficient design of statistical investigations. *Ann. Math. Statist.* **14** 134–140.

WALD, A. (1947). *Sequential Analysis.* Wiley, New York.

WALD, A. (1950). *Statistical Decision Functions.* Wiley, New York.

WOLFOWITZ, J. (1951). On ε-complete classes of decision functions. *Ann. Math. Statist.* **22** 461–465.

ZERDY, G. (1980). Risk of asymptotically optimum sequential tests. *Ann. Statist.* **8** 1110–1122.

DEPARTMENT OF MATHEMATICS
CORNELL UNIVERSITY
ITHACA, NEW YORK 14853