THE NONEXISTENCE OF CERTAIN STATISTICAL PROCEDURES IN NONPARAMETRIC PROBLEMS[1]

BY R. R. BAHADUR AND LEONARD J. SAVAGE

The University of Chicago

1. Introduction. It seems plausible that if the population distribution of a real random variable is entirely unknown, then a sample from the population can yield little or no information about the tails of the distribution, even if the sample is obtained according to a sequential procedure. This paper gives evidence supporting and clarifying this proposition.

The paper treats in some detail problems of inference concerning the population mean μ. It is shown that there is neither an effective test of the hypothesis that $\mu = 0$, nor an effective confidence interval for μ, nor an effective point estimate of μ. These conclusions concerning μ flow from the fact that μ is sensitive to the tails of the population distribution; parallel conclusions hold for other sensitive parameters, and they can be established by the same methods as are here used for μ.

It is also shown that there exists no confidence band for the population distribution function such that the upper and lower limits of the band are themselves distribution functions; that is, no confidence band fits very well.

2. Theorems. Let $\mathfrak{F}$ be a given set of distribution functions F, G, $\ldots$ of a real variable. Some of the theorems to be proved would be of interest even if $\mathfrak{F}$ were required to be the class of all distributions or perhaps all distributions F with finite mean μ_F. But it is helpful to recognize that the proofs require only that $\mathfrak{F}$ have a certain richness. Specifically, Theorem 1 and Corollaries 1 through 4 depend on the following three hypotheses:

(i) For every $F \varepsilon \mathfrak{F}$, $\mu_F = \int_{-\infty}^{\infty} z \, dF$ exists and is finite.
(ii) For every real m, there is an $F \varepsilon \mathfrak{F}$ with $\mu_F = m$.
(iii) $\mathfrak{F}$ is convex; that is, if $F \varepsilon \mathfrak{F}$, $G \varepsilon \mathfrak{F}$, π is a positive fraction, and $H = \pi F + (1 - \pi)G$, then $H \varepsilon \mathfrak{F}$.

Theorem 2 depends on hypotheses (iii) and the following:

(iv) $\mathfrak{F}$ is closed under translation; that is, if $F \varepsilon \mathfrak{F}$, and $G(z) = F(z - h)$ for all z and some h, then $G \varepsilon \mathfrak{F}$.
(v) $\mathfrak{F}$ is nonvacuous.

Some obvious examples of sets satisfying all four conditions are the sets of all distribution functions F such that μ_F is finite; the points of increase of F are a bounded set, or are a finite set; F is absolutely continuous and dF/dz vanishes

Received October 11, 1955.

[1] Research sponsored in part by the Office of Naval Research.

outside of a bounded interval; as z approaches ∞, $1 - F(z) + F(-z) = 0(z^{-r})$ for an $r > 1$.

Since the theorems to be proved are theorems of nonexistence, it is appropriate that they be stated and proved for mixed (i.e., randomized) procedures—sampling, estimating, testing. They are, of course, true a fortiori for the smaller class of pure procedures. The technique of working with mixed procedures is presented in detail in certain publications, for example [1] and [2]. We feel free, therefore, to handle mixed procedures rather informally, to save space and tedium.

Let $X_1, X_2, \cdots$ denote an infinite sequence of independent random variables, each distributed according to F; that is, $\Pr(X_i \leqq z) = F(z)$. Suppose that a (randomized, sequential) sampling procedure is given, that is, a set of rules for observing $X_1, X_2, \cdots$ one by one up to a certain stage N such that at each stage the decision whether to continue depends (randomly) on the observed values in hand at that stage. The given procedure, which will remain fixed throughout the discussion, is naturally assumed to be closed, that is,

$$P_F(N < \infty) = 1 \tag{1}$$

for each $F \varepsilon \mathfrak{F}$. Except for this condition, the sampling procedure is arbitrary.

Denote the total outcome of the sampling procedure, regarded as a random variable, by V, that is, $V = (X_1, X_2, \cdots, X_N)$. As already exemplified in (1), for any event A defined on the sample space of V, $P_F(A)$ will denote the probability of A when F obtains, that is to say, when each X_i is distributed according to F. If φ is a real valued function of V, $E_F[\varphi]$ will denote the expected value of φ (if it exists) when F obtains.

For any real number m, let $\mathfrak{F}_m$ denote the set of all $F \varepsilon \mathfrak{F}$ with $\mu_F = m$.

THEOREM 1. *For each bounded real valued function φ on the sample space of V, $\inf_{F \varepsilon \mathfrak{F}_m} E_F[\varphi]$ and $\sup_{F \varepsilon \mathfrak{F}_m} E_F[\varphi]$ are independent of m.*

The proofs of this theorem and of Theorem 2 below are postponed to the next section. Theorem 1 states, in effect, that even if μ_F is known to equal one of two given values m_1 and m_2, the sample V cannot provide effective discrimination between the two hypothetical values. The following Corollaries 1 through 4 exploit the close relations between discrimination, testing, and estimation to make explicit some consequences of Theorem 1 in problems of inference concerning μ_F. As was mentioned in the introduction, analogues of Theorem 1 (and therewith of Corollaries 1 through 4) are valid for parameters other than the mean, and these analogues can be proved by the same method as is used in the next section to prove Theorem 1.

Let H be the hypothesis that $\mu_F = 0$ (i.e., $F \varepsilon \mathfrak{F}_0$). For any test t, let $\beta_F(t)$ denote the probability of rejecting H in using t when F obtains, in short, the power function of t. Call t a somewhere unbiased level-α test if $\beta_F \leqq \alpha$ for $F \varepsilon \mathfrak{F}_0$ and, for some m different from zero, $\beta_F \geqq \alpha$ for $F \varepsilon \mathfrak{F}_m$. Call t a similar level-α test if $\beta_F = \alpha$ for each $F \varepsilon \mathfrak{F}_0$.

Taking $\varphi(V)$ to be the probability prescribed by t of rejecting H on observation of V yields this corollary.

COROLLARY 1. *If t is a somewhere unbiased level-α test of H, or a similar level-α test of H, then $\beta_F(t) = \alpha$ for all $F \varepsilon \mathfrak{F}$.*

Corollary 1 asserts the failure, in certain senses, of all tests of the value of μ, assuming that μ exists. It would be interesting to know whether, in comparable nonparametric situations, tests of the *existence* of μ are equally unsuccessful. To be precise, suppose for example that $\mathfrak{F}$ is the set of all distribution functions. Let $\mathfrak{F}^*$ be the subset of $\mathfrak{F}$ on which μ_F exists finitely, and let H^* denote the hypothesis that F is in $\mathfrak{F}^*$. Then, does Corollary 1 hold with H replaced by H^* and "somewhere unbiased" replaced by "unbiased"?

Next, let I be a confidence set for μ_F, that is, I is a (randomized) function of V, that has Borel subsets of the real line for its values. For any real m, let $C[m]$ denote the event that I covers m.

COROLLARY 2. *If $P_F(C[\mu_F]) \geqq 1 - \alpha$ for all $F \varepsilon \mathfrak{F}$, then $P_F(C[m]) \geqq 1 - \alpha$ for all m and all $F \varepsilon \mathfrak{F}$.*

PROOF. For each m, let $p_m(V)$ be the conditional probability of $C[m]$ given V, $0 \leqq p \leqq 1$. Consider a fixed m. By hypothesis, $E_F[p_m] \geqq 1 - \alpha$ for $F \varepsilon \mathfrak{F}_m$. Hence, $P_F(C[m]) = E_F[p_m] \geqq 1 - \alpha$ for all $F \varepsilon \mathfrak{F}$, by Theorem 1. Since m is arbitrary, the corollary is proved

COROLLARY 3. *Suppose that there exists at least one $F \varepsilon \mathfrak{F}$ such that $P_F(I$ is a set bounded from below$) = 1$. Then $\inf_{F \varepsilon \mathfrak{F}} \{P_F(C[\mu_F])\} = 0$.*

PROOF. For each $n = 1, 2, \cdots$, let B_n denote the event that I is contained in the interval $[-n, \infty)$, and let $\bar{B}_n$ denote the complement of B_n. For each n, let $q_n(V)$ denote the probability of B_n given V; $0 \leqq q_n \leqq q_{n+1} \leqq 1$.

Now let F be a distribution in $\mathfrak{F}$ such that I is bounded from below with probability 1 when F obtains. By Lebesgue's theorem for monotone sequences, $E_F[\lim_n q_n] = \lim_n E_F[q_n] = \lim_n P_F(B_n) = P_F(I$ is bounded from below$) = 1$. Consequently, $\lim_n q_n(V) = 1$ except on a set of points V of P_F-measure zero. Since, for any $m < -n$, $p_m(V) = \Pr(m \varepsilon I \mid V) \leqq \Pr(\bar{B}_n \mid V) = 1 - q_n(V)$, it follows that, except on a P_F-null set,

$$\lim_{m \to -\infty} p_m(V) = 0. \tag{2}$$

Since $P_F(C[m]) = E_F[p_m]$ for all m, it follows from (2), by Lebesgue's theorem for boundedly convergent sequences, that

$$\lim_{m \to -\infty} P_F(C[m]) = 0. \tag{3}$$

Now, Corollary 2 states in effect that $\inf_{G \varepsilon \mathfrak{F}} \{P_G(C[\mu_G])\} = \inf_{G \varepsilon \mathfrak{F}} \inf_m \{P_G(C[m])\}$. It follows from (3) that the common value of these infima is zero. This completes the proof.

Of course, "set bounded from above," and, a fortiori, "bounded set" can be substituted for "set bounded from below" in the statement of Corollary 3. But the following example shows that it would not be enough to say "set bounded from above or from below." For all V, let $I = (-\infty, 0]$ with probability $\frac{1}{2}$ and $I = (0, \infty)$ with probability $\frac{1}{2}$; then $P_F(C[m]) = \frac{1}{2}$ for all m and all F.

Next, consider the problem of constructing a suitable point estimator for μ_F. Let M be an estimator, that is, a real valued (randomized) function of V. Suppose that when F obtains, the expected loss in using M is $E_F[L(M - \mu_F)] = r_F(M)$, where $L(m)$ is bounded from below and $\lim_{m\to\infty} L(m) = \infty$ or $\lim_{m\to-\infty} L(m) = \infty$ (e.g., $L(m) = |m|$, $L(m) = m^2$, $L(m) = (2 + \sin m)e^m$).

Let $\rho(F)$ be a real valued functional on $\mathfrak{F}$. Say that ρ is uncontrollable (from above) if there exists no real valued (randomized) function of V, say S, such that $\inf_{F\varepsilon\mathfrak{F}} \{P_F(\rho(F) < S)\} > 0$.

The following corollary shows that there is no estimator M for which the expected loss $r_F(M)$ is bounded in F, nor even one for which the sample gives any clue as to the possible expected loss.

Corollary 4. *For any estimator M, $r_F(M)$ is uncontrollable.*

Proof. There is no loss in generality in assuming that $\lim_{m\to\infty} L(m) = \infty$. Replacing $L(m)$ by $L(m) - \inf_a L(a)$, there is also no loss in assuming L non-negative, with $\inf_m L(m) = 0$. Consider a fixed estimator M. Write $L_F = L(M - \mu_F)$. Since $L_F \geqq 0$, it is easily seen (a la Tchebycheff) by considering the cases $r_F = 0$, $0 < r_F < \infty$, and $r_F = \infty$ separately that $P_F(L_F \leqq \alpha r_F) \geqq 1 - (1/\alpha)$ for all $\alpha > 0$ and all F.

Suppose, contrary to Corollary 4, that there exists a random variable S with distribution determined by V, and a positive constant β, such that $P_F(r_F < S) \geqq \beta$ for all $F \varepsilon \mathfrak{F}$. There is no loss of generality in assuming that S is always positive. Choose and fix an $\alpha > 0$ such that $\beta - (1/\alpha) > 0$. Let $Y = \sup\{m: L(m) \leqq \alpha S\}$ and define I to be the random interval $[M - Y, \infty)$. Then $P_F(I$ is bounded from below$) = 1$ for each F. Also, for each $F \varepsilon \mathfrak{F}$,

$$\begin{aligned} P_F(C[\mu_F]) &= P_F(M - \mu_F \leqq Y) \\ &\geqq P_F(L_F \leqq \alpha S) \\ &\geqq P_F(L_F \leqq \alpha S, r_F < S) \\ &\geqq P_F(L_F \leqq \alpha r_F, r_F < S) \\ &\geqq P_F(L_F \leqq \alpha r_F) + P_F(r_F < S) - 1 \\ &\geqq 1-(1/\alpha)+\beta-1 \\ &> 0. \end{aligned}$$

This contradiction to Corollary 3 establishes Corollary 4.

The preceding proof consists in showing that if M is an estimator such that $r_F(M)$ is controllable, then μ_F is controllable, contrary to Corollary 3. This argument can also be used to show the uncontrollability of certain parameters. Simple examples of such parameters are the variance of F, the difference between the mean and median values of F, and the supremum of the points of increase of F. Note that while the unboundedness of these parameters is evident when assumptions such as (iii) and (iv) hold, verification that they are uncontrollable is less trivial even in the case when V consists of a single observation.

Finally, let $A(z)$ be a (randomized) function of V taking values in the set of all distribution functions of z. Let $C^*[F]$ denote the event that $A(z) \geqq F(z)$ for all z.

THEOREM 2. $\inf_{F \in \mathfrak{F}} \{P_F(C^*[F])\} = 0$.

Application of Theorem 2 to $-X_i$ yields with little effort a similar theorem, dual to Theorem 2, concerning the probability that $A(z) \leqq F(z)$ for all z. Obviously these two theorems together imply a two-sided version of Theorem 2.

3. Proofs of the theorems. The proofs of the theorems depend on the fact that a given distribution function F can be so modified that, while the probability distribution of the X_i's (and therewith of V) is perturbed only slightly, parameters such as the mean suffer arbitrary displacements. This modification is described in the following paragraphs, before undertaking the proofs of Theorems 1 and 2.

Let Φ denote the class of all functions φ of V with $0 \leqq \varphi \leqq 1$, and (for any two distribution functions F and G) define the familiar absolute-variational distance between F and G by

$$\delta(F, G) = \sup_{\varphi \in \Phi} | E_F[\varphi] - E_G[\varphi] | . \tag{4}$$

Given F, let H be an arbitrary distribution function and π an arbitrary constant, $0 < \pi < 1$, and define the distribution function G thus:

$$G(z) = \pi F(z) + (1 - \pi)H(z). \tag{5}$$

The following lemma shows that if the given sampling procedure is closed for F in the sense of (1), and if π is sufficiently close to 1, then, *no matter what H may be*, the probability distributions of V under F and G are not very distant from one another. It may clarify the meaning and proof of the lemma to remark that it is for this application of the lemma, not for the lemma itself, that the sampling procedure must be closed for F.

LEMMA. *$\delta(F, G) \leqq 1 - \pi^k P_F(N \leqq k)$ for each positive integer k.*[2]

PROOF. Choose and fix a positive integer k. Let $R^{(k)}$ denote the space of all points $z^{(k)} = (z_1, z_2, \cdots, z_k)$ with $-\infty < z_i < \infty$ for $i = 1, 2, \cdots, k$. For any univariate distribution function $F(z)$, write $F^{(k)}(z^{(k)}) = \prod_{i=1}^k F(z_i)$.

It will be shown first that if F and G are related according to (5), then, for any nonnegative function f on $R^{(k)}$,

$$\int_{R^{(k)}} f \, dG^{(k)} \geqq \pi^k \int_{R^{(k)}} f \, dF^{(k)}. \tag{6}$$

To verify this inequality, let $Y_1, Y_2, \cdots, Y_k, Z_1, Z_2, \cdots, Z_k$, and $U_1, U_2, \cdots, U_k$ be independent random variables such that each Y_i is distributed according to F, each Z_i according to H, and $P(U_i = 1) = 1 - P(U_i = 0) = \pi$

[2] These inequalities are a considerable improvement of the corresponding ones in an earlier version of the lemma, and the proof is somewhat similar. The authors are indebted to Professor W. Hoeffding for these improvements.

for each U_i. Write $W_i = U_iY_i + (1 - U_i)Z_i$ for $i = 1, 2, \cdots, k$. Then W_1, $W_2, \cdots, W_k$ are independent random variables, each distributed according to G as defined by (5). Let B denote the event that $U_i = 1$ for all $i = 1, 2, \cdots, k$ and let $\bar{B}$ denote the complement of B. Now it is straightforward to show (6); thus,

$$\begin{aligned}\int_{R^{(k)}} f\,dG^{(k)} &= E[f(W_1, \cdots, W_k)]\\ &= P(B)E[f(W_1, \cdots, W_k) \mid B] + P(\bar{B})E[f(W_1, \cdots, W_k) \mid \bar{B}]\\ &\geqq P(B)E[f(W_1, \cdots, W_k) \mid B]\\ &= \pi^k E[f(Y_1, \cdots, Y_k) \mid B]\\ &= \pi^k E[f(Y_1, \cdots, Y_k)]\\ &= \pi^k \int_{R^{(k)}} f\,dF^{(k)}.\end{aligned} \tag{7}$$

Consider the space of all sequences $X^{(\infty)} = (X_1, X_2, \cdots \text{ad inf})$. Since V, the observed sample, is by definition a (randomized) function of $X^{(\infty)}$, it makes sense to speak of the conditional distribution of V given $X^{(k)} = (X_1, \cdots, X_k)$. It follows from a well known property of conditional expectation that, for any function $h(V)$ and any F,

$$E_F[h] = \int_{R^{(k)}} E_F[h \mid X^{(k)}]\,dF^{(k)}, \tag{8}$$

provided that $E_F[h]$ exists.

Next, let φ be a function of V such that $0 \leqq \varphi \leqq 1$. Define $\psi(V) = 1$ if $N \leqq k$ and $\psi(V) = 0$ if $N > k$. It is easy to see that there exists a function f on $R^{(k)}$ such that $0 \leqq f \leqq 1$, and

$$E_F[\varphi\cdot\psi \mid X^{(k)}] = f(X^{(k)}), \tag{9}$$

for *all* F. The function f depends, of course, on the given φ and the given sampling procedure.

Suppose, now, that F and G are two distribution functions related according to (5). Then,

$$\begin{aligned}E_G[\varphi] &\geqq E_G[\varphi\cdot\psi]\\ &= \int_{R^{(k)}} E_G[\varphi\cdot\psi \mid X^{(k)}]\,dG^{(k)} && \text{by (8)}\\ &= \int_{R^{(k)}} f\,dG^{(k)} && \text{by (9)}\\ &\geqq \pi^k \int_{R^{(k)}} f\,dF^{(k)} && \text{by (6)}\end{aligned}$$

$$= \pi^k \int_{R(k)} E_F[\varphi \cdot \psi \mid X^{(k)}]\, dF^{(k)} \qquad \text{by (9)}$$

$$= \pi^k E_F[\varphi \cdot \psi] \qquad \text{by (8)}$$

$$= E_F[\varphi] - E_F[\varphi(1 - \pi^k \psi)]$$

$$\geqq E_F[\varphi] - E_F[1 - \pi^k \psi]$$

$$= E_F[\varphi] - 1 + \pi^k P_F(N \leqq k).$$

Thus,

$$E_F[\varphi] - E_G[\varphi] \leqq 1 - \pi^k P_F(N \leqq k) \tag{10}$$

for all φ in Φ. Since $\varphi \varepsilon \Phi$ implies $1 - \varphi \varepsilon \Phi$, it follows from (10) that

$$-E_F[\varphi] + E_G[\varphi] \leqq 1 - \pi^k P_F(N \leqq k) \tag{11}$$

for all φ in Φ. In view of (10), (11), and the definition (4) of δ, $\delta(F, G) \leqq 1 - \pi^k P_F(N \leqq k)$ for all k, as was to be proved.

Proof of Theorem 1. Let m and m' be real numbers, and let $\epsilon > 0$ be given. Consider a fixed F in $\mathfrak{F}_m$. Choose and fix a positive integer k such that

$$P_F(N > k) < \epsilon. \tag{12}$$

The existence of such a k is assured by (1). Now choose and fix a π such that $0 < \pi < 1$ and

$$(1 - \pi^k) < \epsilon. \tag{13}$$

Let H be a distribution function in $\mathfrak{F}$ such that $\pi m + (1 - \pi)\mu_H = m'$ (see assumption (ii)), and let G be defined by (5). Then, by assumption (iii), G is in $\mathfrak{F}$, and since $\mu_G = \pi\mu_F + (1 - \pi)\mu_H = m'$, G is in $\mathfrak{F}_{m'}$. Since $1 - \pi^k P_F(N \leqq k) \leqq (1 - \pi^k) + P_F(N > k)$, it follows from (12), (13), and the Lemma that $\delta(F, G) < 2\epsilon$.

Since ϵ and F are arbitrary, $\inf_{G \epsilon \mathfrak{F}_{m'}} \{\delta(F, G)\} = 0$ for each $F \epsilon \mathfrak{F}_m$. In other words, $\mathfrak{F}_{m'}$ is everywhere dense in $\mathfrak{F}_m$, under the metric δ. Since m and m' are arbitrary, it follows (see assumption (i)) that, for each m, $\mathfrak{F}_m$ is everywhere dense in $\mathfrak{F}$. This conclusion, together with the observation that $E_F[\varphi]$ is continuous in F for any bounded φ, yields Theorem 1.

Proof of Theorem 2. Before the proof proper we present a line of argument that may be of some interest in suggesting a heuristic connection between this theorem and Theorem 1, though this line of argument makes assumptions that are actually gratuitous. It assumes in fact that (i) obtains and that, for some F, the mean of A almost always exists and is finite.

Suppose, then, that the random distribution function A is such that, for some $F \varepsilon \mathfrak{F}$;

$$-\infty < \int_{-\infty}^{\infty} z\, dA < \infty \tag{14}$$

except for a P_F-null event. Define $I = [\int_{-\infty}^{\infty} z\, dA, \infty)$ whenever (14) is satisfied, and $I = (-\infty, \infty)$ otherwise. Now, the event $A(z) \geqq F(z)$ for all z (that is, $C^*[F]$) implies the event $\infty > \mu_F \geqq \int_{-\infty}^{\infty} z\, dA$ (that is, $C[\mu_F]$) provided only that μ_F exists and is finite. Hence $P_F(C^*[F]) \leqq P_F(C[\mu_F])$ for each $F \varepsilon \mathfrak{F}$. The desired conclusion now follows from Corollary 3.

Now, dropping the assumptions (i) and (14), turn to the proof proper. Choose and fix an ϵ, $0 < \epsilon < 1$, and an $F \varepsilon \mathfrak{F}$ such that $F(0) > 0$. The existence of such an F is assured by assumptions (iv) and (v). For each z, let $J(z) = \inf \{u: P_F(A(z) \leqq u) \geqq 1 - \epsilon\}$. It is not difficult to see that J is a nondecreasing function of z, with $\lim_{z \to -\infty} J(z) = 0$, $\lim_{z \to \infty} J(z) = 1$, and that J is also continuous from the right, so that it is actually a distribution function. Also,

$$P_F\{A(z) > J(z)\} \leqq \epsilon \tag{15}$$

for each z.

Now choose k such that (12) holds, choose π such that (13) holds, and choose λ such that $J(\lambda) < (1 - \pi)F(0)$. Let G be defined by (5), with $H(z) \equiv F(z - \lambda)$. Then G is in $\mathfrak{F}$, by assumptions (iii) and (iv), and

$$J(\lambda) < G(\lambda), \tag{16}$$

by the choice of λ and the definition of G. Hence

$$\begin{aligned} P_G(C^*[G]) &= P_G(A(z) \geqq G(z) \text{ for all } z) \\ &\leqq P_G(A(\lambda) \geqq G(\lambda)) \\ &\leqq P_G(A(\lambda) > J(\lambda)) && \text{by (16)} \\ &\leqq P_F(A(\lambda) > J(\lambda)) + P_F(N > k) + (1 - \pi^k) \\ & && \text{by the lemma} \\ &\leqq 3\epsilon && \text{by (12), (13), (15).} \end{aligned} \tag{17}$$

Since ϵ is an arbitrary positive fraction, the theorem is proved.

The proof of Theorem 2 does not use quite the full force of (iii) and (iv). It is enough that for some $F \varepsilon \mathfrak{F}$ and two sequences of numbers $\alpha_i (0 \leqq \alpha_i < 1)$ and β_j, such that $\alpha_i \to 1$ and $\beta_j \to \infty$, the distributions G_{ij} such that

$$G_{ij}(z) = \alpha_i F(z) + (1 - \alpha_i)F(z + \beta_j) \tag{18}$$

are in $\mathfrak{F}$. For the dual of Theorem 2, it is required instead that $\beta_j \to -\infty$.

REFERENCES

[1] A. DVORETZKY, A. WALD, AND J. WOLFOWITZ, "Elimination of randomization in certain statistical decision procedures and zero sum two-person games," *Ann. Math. Stat.*, Vol. 22 (1951), pp. 1–21.

[2] ABRAHAM WALD, *Statistical Decision Functions*, New York, John Wiley and Sons, 1950.