ASYMPTOTICALLY MINIMAX DISTRIBUTION-FREE PROCEDURES[1]

By Kjell Doksum

University of California, Berkeley

1. Introduction and summary. The power and efficiency of robust procedures have been considered for parametric alternatives by Tukey (1946), Hoeffding (1951), Lehmann (1953), (1959), Chernoff and Savage (1958), Capon (1961), and others. For instance, it has been shown that the two-sample normal scores statistic is optimal for normal translation alternatives in that for these alternatives, it is locally most powerful [12], [15], [4] in the class of all rank statistics and it is asymptotically efficient [6], [4] in the class of all statistics. However, for other types of translation alternatives, the normal scores statistic does not have these optimal properties.

In this paper, optimality properties for non-parametric classes of alternatives are treated. In particular, statistics that maximize the minimum power asymptotically over classes of non-parametric alternatives will be considered. If one takes (1-power) as the risk function, these statistics are asymptotically minimax. It turns out that in this sense the Wilcoxon statistic is asymptotically minimax over a class of one-sided alternatives (F, G) with $\rho(F, G) \geqq \Delta$, where ρ is the Kolmogorov distance, while the normal scores statistic is asymptotically minimax for a non-parametric class of translation type alternatives.

For the problem of estimating differences in location, the normal scores estimate of Hodges and Lehmann (1963) is shown to be minimax when the risk function is asymptotic variance. Finally, the results are used to obtain asymptotic efficiencies that are defined for non-parametric classes of alternatives.

2. Minimax results in the two-sample case. Let $X_1, \cdots, X_m$ and $Y_1, \cdots, Y_n$ be two independent random samples with distributions F and G respectively. Consider the problem of testing $H_0: F = G$ against one sided alternatives $H_1: G < F$. If $\mathfrak{J}$ is a class of level α tests for this problem, and Ω is a class of alternatives (F, G) with $G < F$, then $\varphi \varepsilon \mathfrak{J}$ is said to be *minimax* over Ω and $\mathfrak{J}$ iff it maximizes the minimum power, i.e. iff

$$\inf_{(F,G)\varepsilon\Omega} \beta_\varphi(F, G) = \sup_{\psi\varepsilon\mathfrak{J}} [\inf_{(F,G)\varepsilon\Omega} \beta_\psi(F, G)]. \tag{2.1}$$

It is clear thar Ω can not be taken to be all (F, G) with $G < F$ since for this class, the infima in (2.1) would both be α and all tests in $\mathfrak{J}$ would be minimax. Thus the alternatives in Ω must be "separated." Birnbaum (1953), Chapman (1958), Bell, Moser and Thompson (1966), and others have considered alternatives separated by the Kolmogorov distance, i.e. alternatives (F, G) with $G < F$ and $\sup_x [F(x) - G(x)] \geqq \Delta$. Here, $\Omega(\Delta)$ will denote the class of (F, G) with $G < F$, $\sup_x [F(x) - G(x)] \geqq \Delta$ and F continuous.

Received 15 June 1965; revised 20 January 1966.

[1] This research was partially supported by the General Electric Corporation.

A class of statistics considered extensively in the literature (e.g. [6], [15], [4]) is the class of rank statistics that can be put in the form

$$T_{J_N} = m^{-1} \sum J_N[r_i/(N + 1)], \tag{2.2}$$

where r_i is the rank of x_i in the combined sample of x's and y's, $N = m + n$, and J_N is a function on the unit interval. Lehmann (1959), Capon (1961) and others have shown that these are the only statistics that are invariant for H_0 against H_1 and are locally most powerful for a specific parametric alternative. For instance, for normal translation alternatives, the normal scores statistic

$$V = m^{-1} \sum E[Z(r_i) | \Phi]$$

is locally most powerful and is known [6] to be of the form (2.2) ($E[Z(i) | \Phi]$ denotes the expected value of the ith order statistic in a random sample of size N from a standard normal distribution). Similarly, for logistic translation alternatives, the Wilcoxon statistic $W = m^{-1} \sum r_i/(N + 1)$ is locally most powerful and is of the form (2.2) with $J_N(u) = u$.

The class of tests for which asymptotic minimax results are first obtained is the class $\mathfrak{J}$ of level α tests φ_{J_N} with critical region of the form $[T_{J_N} < c]$ with J_N converging to some continuously differentiable function J on $(0, 1)$, and with J and J_N satisfying the conditions of Theorem 4.1 or Theorem 4.2 of Govindarajulu, LeCam and Raghavachari (1965). The conditions of these authors are generalizations of those of Theorem 1 of Chernoff and Savage (1958). The above conditions are used because they yield asymptotic normality and they imply that

$$J \text{ has a continuous derivative } J' \text{ on } (0, 1), \text{ and } \int_0^1 J^2(u)\, du < \infty. \tag{2.3}$$

It is assumed throughout that

$$0 < \lim (m/N) = \lambda < 1. \tag{2.4}$$

THEOREM 2.1. (i) *The Wilcoxon test* φ_0 *is asymptotically minimax over* $\Omega(\Delta)$ *and* $\mathfrak{J}$ *in the sense that*

$$\lim_{N\to\infty} [\inf_{(F,G)\varepsilon\Omega(\Delta_N)} \beta_{\varphi_0}(F, G)] = \sup_{\varphi\varepsilon\mathfrak{J}} \{\lim\sup_{N\to\infty} [\inf_{(F,G)\varepsilon\Omega(\Delta_N)} \beta_{\varphi}(F, G)]\} \tag{2.5}$$

for each sequence $\{\Delta_N\}$ *satisfying*

(I) $\alpha < \lim_{N\to\infty} [\inf_{(F,G)\varepsilon\Omega(\Delta_N)} \beta_{\varphi_0}(F, G)] = \beta < 1$.

(ii) (I) *holds* iff $\Delta_N N^{\frac{1}{2}} \to c$ *for some* $c > 0$.

(iii) φ_0 *is asymptotically uniquely minimax in the sense that if* φ_{J_N} *satisfies* (i), *then* $J_N(u) \to bu + d$ *on* $(0, 1)$ *for some* b *and* d.

PROOF. The result follows at once from the following two lemmas. Let $\beta_{\varphi_0}(\Delta_N)$ denote $\inf \{\beta_{\varphi_0}(F, G) : (F, G) \varepsilon \Omega(\Delta_N)\}$ and let Φ denote the standard normal distribution, then

LEMMA 2.1. *$\beta_{\varphi_0}(\Delta_N) \to \Phi(k_\alpha + [3\lambda(1-\lambda)]^{\frac{1}{2}}c^2)$ with $\Phi(k_\alpha) = \alpha$ iff $\Delta_N N^{\frac{1}{2}} \to c$, $c > 0$.*

PROOF. Using the arguments of Birnbaum (1953), Chapman (1958), and Bell, Moser and Thompson (1966), one finds distributions $(U, G_{\Delta,a})$ that are least favorable in the sense that

$$\beta_{\varphi_0}(\Delta) = \inf_{0 \leqq a \leqq 1-\Delta} \beta_{\varphi_0}(U, G_{\Delta,a}), \tag{2.6}$$

where U is the standard uniform distribution, and $G_{\Delta,a}$ is the Birnbaum alternative defined by

$$\begin{aligned} G_{\Delta,a}(x) &= 0 && \text{if} \quad x \leqq 0 \\ &= x && \text{if} \quad 0 \leqq x \leqq a \quad \text{or} \quad a + \Delta \leqq x \leqq 1 \\ &= a && \text{if} \quad a \leqq x < a + \Delta \\ &= 1 && \text{if} \quad x \geqq 1. \end{aligned}$$

Note that $G_{\Delta,a}$ is not continuous. Since most theorems on asymptotic normality require continuity, the following remark by Thompson (1964) on continuous least favorable distributions will be used:

$$\beta_{\varphi_0}(U, G_{\Delta,a}) = \beta_{\varphi_0}(F_{\Delta,a}, G^*_{\Delta,a}), \tag{2.7}$$

where $F_{\Delta,a}$ and $G^*_{\Delta,a}$ are the Thompson alternatives defined by

$$\begin{aligned} F_{\Delta,a}(x) &= 0 && \text{if} \quad x \leqq 0 \\ &= x && \text{if} \quad 0 \leqq x \leqq a + \Delta \\ &= a + \Delta && \text{if} \quad a + \Delta \leqq x \leqq a + 2\Delta \\ &= x - \Delta && \text{if} \quad a + 2\Delta \leqq x \leqq 1 + \Delta \\ &= 1 && \text{if} \quad x \geqq 1 + \Delta \end{aligned}$$

and

$$\begin{aligned} G^*_{\Delta,a}(x) &= 0 && \text{if} \quad x \leqq 0 \\ &= x && \text{if} \quad 0 \leqq x \leqq a \\ &= a && \text{if} \quad a \leqq x \leqq a + \Delta \\ &= x - \Delta && \text{if} \quad a + \Delta \leqq x \leqq 1 + \Delta \\ &= 1 && \text{if} \quad x \geqq 1 + \Delta. \end{aligned}$$

Equation (2.7) follows since r_i equals the rank of $F_{\Delta,a}(x_i)$ among

$$F_{\Delta,a}(x_1), \cdots, F_{\Delta,a}(x_m); \qquad F_{\Delta,a}(y_1), \cdots, F_{\Delta,a}(y_n)$$

and the distributions of $F_{\Delta,a}(X_1)$ and $F_{\Delta,a}(Y_1)$ are U and $G_{\Delta,a}$ respectively.

Let $\beta_{\varphi_0}(\Delta, a)$ denote $\beta_{\varphi_0}(F_{\Delta,a}, G^*_{\Delta,a})$. Corollary 1 of Chernoff and Savage and a few computations yield

(2.8) $\sup_{0 \leqq a \leqq 1-\Delta_N} |\beta_{\varphi_0}(\Delta_N, a) - \Phi(k_\alpha + [3\lambda(1-\lambda)]^{\frac{1}{2}}\Delta_N{}^2 N^{\frac{1}{2}})| \to 0,$

provided that $\Delta_N \to 0$ as $N \to \infty$. Hence,

(2.9) $|\inf_{0 \leqq a \leqq 1-\Delta_N} \beta_{\varphi_0}(\Delta, a) - \Phi(k_\alpha + [3\lambda(1-\lambda)]^{\frac{1}{2}}\Delta_N{}^2 N^{\frac{1}{2}})| \to 0,$

if $\Delta_N \to 0$ as $N \to \infty$. Using (2.6) and (2.7), Lemma 2.1 follows.

LEMMA 2.2. *Let $\varphi_{J_N} \varepsilon \mathfrak{J}$, if Δ_N is such that $\Delta_N N^{\frac{1}{4}} \to c$, and if $J_N \to J$ with $J(u)$ not of the form $bu + d$ (i.e. φ_{J_N} is not asymptotically equivalent to φ_0), then*

(2.10) $\limsup_{N\to\infty} [\inf_{(F,G)\varepsilon\Omega(\Delta_N)} \beta_{\varphi_{J_N}}(F, G)] < \Phi(k_\alpha + [3\lambda(1-\lambda)]^{\frac{1}{2}}c^2).$

PROOF. If $J(u)$ is not of the form $bu + d$, it must give less weight to some parts of (0, 1) than to others. The proof consists of choosing $(F_N, G_N) \varepsilon \Omega(\Delta_N)$ such that the most "mass" falls in the parts of (0, 1) to which $J(u)$ gives the least weight. The details are as follows:

Since $(F_{\Delta,a}, G^*_{\Delta,a}) \varepsilon \Omega(\Delta)$,

(2.11) $\inf_{(F,G)\varepsilon\Omega(\Delta)} \beta_{\varphi_{J_N}}(F, G) \leqq \beta_{\varphi_{J_N}}(F_{\Delta,a}, G^*_{\Delta,a}).$

Let $\beta_{J_N}(\Delta, a)$ denote the right hand side of (2.11), then one can use the results of Chernoff and Savage (1958) to compute

(2.12) $\beta_{J_N}(\Delta_N, a) \to \Phi[k_\alpha + [\lambda(1-\lambda)]^{\frac{1}{2}} J'(a)c^2/2\sigma_J],$

where $\sigma_J{}^2 = \int_0^1 J^2(x)\,dx - (\int_0^1 J(x)\,dx)^2$. If $J'(a) \leqq 0$ for some a in (0, 1), then (2.12) implies that $\lim_{N\to\infty} \beta_{J_N}(\Delta_N, a) \leqq \alpha$ and the result follows from (2.11). The only remaining case is when $J'(a) > 0$ for all a in (0, 1). Then J is the inverse of some distribution function H (say) with variance $\sigma_J{}^2$ and a density h. One can now write (2.12) in the form

(2.13) $\beta_{J_N}(\Delta_N, a) \to \Phi[k_\alpha + [\lambda(1-\lambda)]^{\frac{1}{2}}c^2/2\sigma_J h[J(a)]].$

Among all densities h with variance $\sigma_J{}^2$ the uniform density has the smallest supremum $\sup_x h(x)$, namely $1/(2\cdot 3^{\frac{1}{2}}\sigma_J)$. Since J is not of the form $bu + d$ (and thus $h(x)$ cannot be uniform), one concludes that there exists an $a_0 \varepsilon (0, 1)$ such that $h[J(a_0)] > 1/(2\cdot 3^{\frac{1}{2}}\sigma_J)$. Lemma 2.2 now follows from this remark together with (2.11) and (2.13).

From (2.12), it is seen that the condition (2.3) is not necessary for Theorem 2.1. Let $\mathfrak{J}_1$ denote the class of tests obtained from $\mathfrak{J}$ by replacing (2.3) with the condition: $\int_0^1 J^2 < \infty$ and there exists a_0 in (0, 1) such that J has a continuous derivative $J'(a_0)$ at a_0 satisfying $J'(a_0) \leqq 2\cdot 3^{\frac{1}{2}}\sigma_J$; then the above arguments show that $\mathfrak{J} \subset \mathfrak{J}_1$ and

COROLLARY 2.1. *The results of Theorem 2.1 hold if $\mathfrak{J}$ is replaced by $\mathfrak{J}_1$.*

Next, asymptotic minimax results will be considered for a different class of nonparametric alternatives. It is the class $\Gamma(\Delta, \gamma)$ of translation type alternatives (F, G) with F continuous, $G(x) \leqq F(x - \Delta)$, and such that F has a density f and variance $\sigma_F^2 \leqq \gamma$. The assumption $\sigma_F^2 \leqq \gamma$ is a convenient way to keep infima of power functions away from α.

Let φ_1 denote the level α normal scôres test whose rejection region is of the form $[V < c]$ and let $\mathfrak{I}^*$ be the class of all level α tests. Using the results of Chernoff and Savage (1958), one now easily proves

THEOREM 2.2. (i) *The normal scores test* φ_1 *is minimax asymptotically over* $\Gamma(\Delta, \gamma)$ *and* $\mathfrak{I}^*$ *in the sense that*

$$\text{(2.14)} \quad \inf [\liminf_{N\to\infty} \beta_{\varphi_1}(F, G_N)] = \sup_{\varphi\varepsilon\mathfrak{I}^*} \{\inf [\limsup_{N\to\infty} \beta_\varphi(F, G_N)]\}$$

for infima over all sequences of alternatives $\{(F, G_N)\}$ *with* $(F, G_N)\ \varepsilon\ \Gamma(\Delta_N, \gamma)$ *and for each* $\{\Delta_N\}$ *satisfying*

(II) $\alpha < \inf [\liminf_{N\to\infty} \beta_{\varphi_1}(F, G_N)] = \beta < 1$ *and* $\lim \Delta_N N^{\frac{1}{2}}$ *exists.*

(ii) (II) *holds* iff $\Delta_N N^{\frac{1}{2}} \to c$ *for some* $c > 0$.

(Note that Theorem 2.1 gives minimax properties in terms of *limits of minimum powers*, while Theorem 2.2 gives minimax properties in terms of *minimums of limits of powers*. Of these two quantities, the first seems to be the most natural, but it is also the hardest to deal with as far as translation type alternatives are concerned.)

PROOF. Since J_N is non-decreasing when $\varphi_{J_N} = \varphi_1$, then one can use the arguments of Chapman (1958), Lehmann (1959), p. 187, and Bell, Moser and Thompson (1966) about monotone tests to conclude that for φ_1, the infimum in (2.14) may be computed over sequences $\{(F, F_N^*)\}$ with $F_N^*(x) = F(x - \Delta_N)$ and $\sigma_F^2 = \gamma = 1$. (As far as minimax results are concerned, no generality is lost by assuming $\gamma = 1$.) For such $\{(F, F_N^*)\}$, Chernoff and Savage (1958) have essentially shown that

$$\text{(2.15)} \qquad \inf_F [\liminf_{N\to\infty} \beta_{\varphi_1}(F, F_N^*)] = \Phi[k_\alpha + [\lambda(1 - \lambda)]^{\frac{1}{2}} c]$$

and that the infimum is attained iff $F = \Phi$. Since $(\Phi, \Phi_N^*)\ \varepsilon\ \Gamma(\Delta_N, 1)$, then

$$\text{(2.16)} \qquad \inf [\limsup_{N\to\infty} \beta_\varphi(F, G_N)] \leqq \limsup_{N\to\infty} \beta_\varphi(\Phi, \Phi_N^*)$$

for any test φ. However, it is well known (e.g. Chenoff and Savage (1958)) that φ_1 is asymptotically efficient for the normal translation alternatives (Φ, Φ_N^*); i.e. it has greater asymptotic power than any other test. Hence, if φ is any test, then

$$\text{(2.17)} \quad \limsup_{N\to\infty} \beta_\varphi(\Phi, \Phi_N^*) \leqq \lim_{N\to\infty} \beta_{\varphi_1}(\Phi, \Phi_N^*) = \Phi[k_\alpha + [\lambda(1 - \lambda)]^{\frac{1}{2}} c].$$

Using (2.15) and (2.16), Theorem 2.2 follows.

Asymptotic uniqueness of the minimax solution in Theorem 2.2 can readily be proved under additional assumptions. Let $\Gamma_1(\Delta, \gamma)$ be the class of alternatives $(F, F^*)\ \varepsilon\ \Gamma(\Delta, \gamma)$ with $F^*(x) = F(x - \Delta)$ and such that

$$h(x) = (\partial/\partial\Delta) \log f(x - \Delta)|_{\Delta=0}$$

exists and satisfies $\int h^2(x) f(x)\, dx < \infty$. Moreover, let $\mathfrak{I}_1^*$ denote the class of tests φ_{T_N} that have rejection regions of the form $[T_N > c]$ or $[T_N < c]$ with T_N

satisfying the conditions of van Eeden (1963) when $F = \Phi$, where T_N is a statistic such that $V - bT_N$ is distribution-free with respect to $\Gamma_1(\Delta, \gamma)$ for all b, i.e. $P[(V - bT_N) \leqq t \mid (F, F)]$ is independent of F for all F such that

$$(F, F^*) \;\varepsilon\; \Gamma_1(\Delta, \gamma).$$

Note that $\mathfrak{I}_1^*$ contains the class of all rank tests whose rejection region is of the form $[T_N > c]$ or $[T_N < c]$ with T_N satisfying the conditions of van Eeden (1963) when $F = \Phi$.

THEOREM 2.3. (i) *The normal scores test* φ_1 *is asymptotically minimax over* $\Gamma_1(\Delta, \gamma)$ *and* $\mathfrak{I}_1^*$ *in the sense of Theorem* 2.2.

(ii) *It is asymptotically uniquely minimax for* $\Gamma_1(\Delta, \gamma)$ *and* $\mathfrak{I}_1^*$ *in that if* $\varphi_{T_N} \;\varepsilon\; \mathfrak{I}_1^*$ *is any other asymptotically minimax tests, then there exist constants* b_N *and* d_N *such that* $N^{\frac{1}{2}}[V - (b_N T_N + d_N)] \to 0$ *in probability under* (F, F_N^*) *for all* $(F, F_N^*) \;\varepsilon\; \Gamma_1(c/N^{\frac{1}{2}}, \gamma)$, $c > 0$.

PROOF. Since $(\Phi, \Phi^*) \;\varepsilon\; \Gamma_1(\Delta, \gamma)$, then (i) follows from the proof of Theorem 2.2. For (ii), note that if φ_{T_N} is asymptotically minimax, then T_N must have asymptotic efficiency one with respect to V for normal translation alternatives. Using the results of van Eeden (1963), this implies that

$$E(N[V - (b_N T_N + d_N)]^2 \mid \Phi, \Phi) \to 0,$$

where b_N and d_N are regression coefficients. However, since $V - (b_N T_N + d_N)$ is distribution-free, then this implies that $E(N[V - (b_N T_N + d_N)]^2 \mid F, F) \to 0$ for all F such that $(F, F^*) \;\varepsilon\; \Gamma_1(\Delta, \gamma)$. Using the theory of contiguity developed by LeCam (1960), Hájek (1962), and Matthes and Truax (1965), the result follows.

Next, the arguments proving Theorem 2.2 will be applied to find a minimax solution for the problem of estimating differences in location. It is now assumed that $G(t) = F(t - \Delta)$, where Δ is to be estimated. A class of robust estimates has been proposed for this problem by Hodges and Lehmann (1963). Using their notation, let $h_{J_N}(x, y) = m^{-1} \sum J_N(r_i/N)$ with $J_N \to J$. It is assumed that when $\Delta = 0$, the distribution of $h_{J_N}(X, Y)$ is symmetric about a fixed point μ. Let $\Delta^*(J_N) = \sup \{\Delta : h_{J_N}(x, y - \Delta) > \mu\}$ and $\Delta^{**}(J_N) = \inf \{\Delta : h_{J_N}(x, y - \Delta) < \mu\}$, then the Hodges-Lehmann estimate is $\hat{\Delta}(J_N) = [\Delta^*(J_N) + \Delta^{**}(J_N)]/2$. Hodges and Lehmann (1963), Theorem 5, have shown that $\hat{\Delta}(J_N)$ is asymptotically normal with asymptotic variance,

$$V_F[\hat{\Delta}(J_N)] = \sigma_J^2/\lambda(1 - \lambda)[\int \{dJ[F(x)]/dx\} \, dF(x)]^2, \tag{2.18}$$

where $\sigma_J^2 = \int_0^1 J^2(t) \, dt - (\int_0^1 J(t) \, dt)^2$.

Let $\tilde{\Delta}$ be the normal scores estimate obtained from $\hat{\Delta}(J_N)$ by letting h_{J_N} be the normal scores statistic, and let $\mathfrak{F}$ be the class of continuous distributions F that have densities f and variances $\sigma_F^2 \leqq \gamma$ and that satisfy the condition (b) of Lemma 3 of Hodges and Lehmann (1961). Finally, let $\mathcal{S}$ be the class of translation invariant estimates $\hat{\Delta}$ such that $N^{\frac{1}{2}}(\hat{\Delta} - \Delta)$ has mean zero and an asymptotic normal distribution with asymptotic variance $V_\Phi(\hat{\Delta})$ for normal alternatives. The arguments proving Theorem 2.2 yield

THEOREM 2.4. *The normal scores estimate $\tilde{\Delta}$ is minimax over $\mathfrak{F}$ and $\mathcal{S}$ in the sense that it minimizes the maximum asymptotic variance*, i.e.

$$\sup_{F\varepsilon\mathfrak{F}} V_F(\tilde{\Delta}) = \inf_{\hat{\Delta}\varepsilon\mathcal{S}} [\sup_{F\varepsilon\mathfrak{F}} V_F(\hat{\Delta})]. \tag{2.19}$$

Note that $V_F(\hat{\Delta})$ is not necessarily available for each $F \varepsilon \mathfrak{F}$. However, (2.19) should be interpreted as meaning that for each $\hat{\Delta} \varepsilon \mathcal{S}$, there exists an $F_0 \varepsilon \mathfrak{F}$ such that $V_{F0}(\hat{\Delta}) \geqq \sup_{F\varepsilon\mathfrak{F}} V_F(\tilde{\Delta})$. The arguments proving uniqueness in Theorem 2.3 can not be used here since $\tilde{\Delta}$ is not distribution-free.

3. Non-parametric asymptotic efficiencies. An immediate extension of the notion of Pitman efficiency to classes of distribution functions is obtained by considering the limit of the ratio M_1/M_2 of the smallest number of observations required by two tests to have power greater than or equal to β *for all* alternatives in a class of distributions.

More precisely, the asymptotic efficiencies are defined as follows: $\Omega(\Delta)$ will be considered first. Since the Wilcoxon test is asymptotically minimax for this class of alternatives and the class of tests considered, the asymptotic efficiency of these tests will be defined with respect to the Wilcoxon test. Fix $0 < \alpha < \beta < 1$ and let φ_N denote the Wilcoxon test. From Section 3, it is known that there exists a sequence $\{\Delta_N\}$ such that $\inf \{\beta_{\varphi_N}(F, G):(F, G) \varepsilon \Omega(\Delta_N)\}$ tends to β as $N \to \infty$. For some other test ψ_N, suppose there exists a sequence $\{k_N\}$ such that $\inf \{\beta_{\psi_{k_N}}(F, G); (F, G) \varepsilon \Omega(\Delta_N)\}$ tends to β as $N \to \infty$; then define the *asymptotic efficiency of ψ over* $\Omega(\Delta)$ by

$$e(\psi) = \lim_{N\to\infty} (N/k_N) \tag{3.1}$$

whenever this limit exists and is well defined. Moreover, define $e(\psi) \leqq e$ to mean that there exists a sequence $\{\beta_N(\Delta)\}$ such that $\inf \{\beta_{\psi_N}(F, G):(F, G) \varepsilon \Omega(\Delta)\} \leqq \beta_N(\Delta)$ for all $\Delta \varepsilon$ [0, 1] and such that $\lim (N/k_N') = e$ for each sequence $\{k_N'\}$ such that $\lim \beta_{k_N'}(\Delta) = \beta$. In other words, $e(\psi) \leqq e$ means that ψ is *dominated* by a sequence with efficiency e.

It now follows from (2.12) and the arguments proving Lemma 2 of Hodges and Lehmann (1961), that

THEOREM 3.1. *If $\psi_{J_N} \varepsilon \mathfrak{J}$, then its asymptotic efficiency satisfies*

$$e(\psi_J) \leqq \inf \{|J'(u)/2\cdot 3^{\frac{1}{2}}\sigma_J|^2 : u \varepsilon (0, 1)\} \leqq 1. \tag{3.2}$$

Moreover, $e(\psi_J) = 1$ iff $J(u)$ *is of the form* $bu + d$, i.e., iff *the test ψ_{J_N} is asymptotically equivalent to the Wilcoxon test.*

COROLLARY 3.1. *If $\psi_{J_N} \varepsilon \mathfrak{J}$ and J is not strictly increasing, then* $e(\psi_J) = 0$.

PROOF. If J is strictly decreasing, then it follows from the results of Lehmann (1959), p. 187, that $\beta_{\psi_{J_N}}(F, G) \leqq \alpha$ for all $(F, G) \varepsilon \Omega(\Delta)$ and the result follows. If $J'(u_0) = 0$ for some $u_0 \varepsilon (0, 1)$, the result follows from Theorem 3.1.

Example 3.1. For the normal scores test, $J = \Phi^{-1}$ and $\inf \{J'(x); x \varepsilon (0, 1)\} = (2\Pi)^{\frac{1}{2}}$, thus the asymptotic efficiency e' of the normal scores test satisfies $e' \leqq \Pi/6 \doteq .52$. Hence the normal scores test is at most about half as efficient over $\Omega(\Delta)$ as the asymptotically minimax Wilcoxon test. The author conjectures that $e' = \Pi/6$.

Next consider non-parametric efficiencies for the translation type alternatives $\Gamma(\Delta, \gamma)$ for which the normal scores test is minimax. Fix $0 < \alpha < \beta < 1$ and let φ_N^* denote the normal scores test. It is known from Section 2 that there exists a sequence Δ_N^* such that $\inf [\liminf_{N\to\infty} \beta_{\varphi_N}^*(F, G_N)] = \beta$, where the infimum is over sequences $\{(F, G_N)\}$ with $(F, G_N) \varepsilon \Gamma(\Delta_N^*, \gamma)$. For this same infimum and some other test ψ_N, suppose there exists a sequence $\{k_N^*\}$ such that

$$\inf [\liminf_{N\to\infty} \beta_{\psi_{k_N^*}}(F, G_N)] = \beta,$$

then define the *asymptotic efficiency of* ψ *over* $\Gamma(\Delta, \gamma)$ to equal

$$e^*(\psi) = \lim_{N\to\infty} (N/k_N^*) \tag{3.3}$$

whenever this limit exists and is well defined.

A result similar to Theorem 3.1 can now be deduced from the arguments of Section 2 in that statistics in $\mathfrak{I}^*$ can be shown to have efficiencies bounded by unity. In particular, one has

Example 3.2. The Wilcoxon test φ_N has the asymptotic efficiency $e'' = 108/125 \doteq .86$ over $\Gamma(\Delta, \gamma)$. This follows from Theorem 1 of Hodges and Lehmann (1956) and Theorem 3 of Chernoff and Savage (1958) in the same way that Theorems 2.2 and 2.3 follow from Theorem 3 of Chernoff and Savage.

For the problem of estimation treated in Section 2, a non-parametric efficiency can similarly be defined over the class $\mathfrak{F}$. Let $\bar{V}(\hat{\Delta}) = \sup \{V_F(\hat{\Delta}): F \varepsilon \mathfrak{F}\}$ denote the maximum asymptotic variance over $\mathfrak{F}$ of the estimate $\hat{\Delta} \varepsilon \mathcal{S}$. Define the asymptotic efficiency of $\hat{\Delta}$ over $\mathfrak{F}$ to equal

$$\tilde{e}(\hat{\Delta}) = \bar{V}(\tilde{\Delta})/\bar{V}(\hat{\Delta}). \tag{3.4}$$

It is clear from [11] that if $\hat{\Delta}$ is the estimate corresponding to the rank test ψ, then $\tilde{e}(\hat{\Delta}) = e^*(\psi)$. Thus if $\hat{\Delta}_0$ is the estimate obtained from the Wilcoxon statistic, then $\tilde{e}(\hat{\Delta}_0) = 108/125 \doteq .86$.

4. Extensions to the one-sample location problem. In this section, the results of the previous two sections will be extended to the one-sample problem. The author hopes to give extensions to other problems in a later paper. The one-sample problem is the same as that of Section 2 with one exception: F is known, so that no X-sample is needed. The class of tests for this problem which corresponds to the class (2.2) considered for the two-sample problem is the class $\mathfrak{I}'$ of tests with critical regions of the form $[\sum J[F(Y_i)] > c]$ with J satisfying (2.3) (see [1]). From the central limit theorem and the arguments of Section 2, it follows that the "uniform scores" test with critical region $[\sum F(Y_i) > c]$ is asymptotically minimax over $\Omega(\Delta)$ and $\mathfrak{I}'$ in the sense of Theorem 2.1. Moreover, since J does not depend on n, it is uniquely asymptotically minimax in that any other test that is asymptotically minimax coincides with it.

Similarly, the "normal scores" test with critical region $[\sum \Phi^{-1}[F(Y_i)] > c]$ is minimax asymptotically in the sense of Theorem 2.2.

Suppose that $G(x) = F(x - \theta)$ where θ is to be estimated. Let $T_J(Y) = \sum J[F(Y_i)]$ have a symmetric distribution about a point μ when $\theta = 0$, and let J be strictly increasing. Following Hodges and Lehmann (1963), one can now define an estimate $\hat{\theta}$ of θ to be the unique number satisfying $T_J(y - \hat{\theta}) = \mu$, where $y = (y_1, \cdots, y_n)$ and $y - \hat{\theta} = (y_1 - \hat{\theta}, \cdots, y_1 - \hat{\theta})$. It now follows from the arguments of Hodges and Lehmann (1963) and Section 2 that the normal scores estimate based on the normal scores statistic $\sum \Phi^{-1}[F(Y_i)]$ is asymptotically minimax (in the sense of Theorem 2.4) over $\mathfrak{F}$ (of Section 2).

Finally, the efficiency results of Section 3 carry over to this one-sample problem with the uniform scores statistic $\sum F(Y_i)$ taking the place of the Wilcoxon statistic and the one-sample normal scores statistic taking the place of the two-sample normal scores statistic.

5. Acknowledgment. I wish to express my gratitude to Professor Erich Lehmann for his guidance and advice throughout the preparation of this paper.

REFERENCES

[1] Bell, C. B. and Doksum, K. D. (1966). "Optimal" one-sample distribution-free statistics and their two-sample extensions. *Ann. Math. Statist.* **37** 120–132.

[2] Bell, C. B., Moser, J. M. and Thompson, R. (1966). Goodness criteria for two-sample distribution-free tests. *Ann. Math. Statist.* **37** 133–142.

[3] Birnbaum, Z. W. (1953). On the power of a one-sided test of fit for continuous distribution functions. *Ann. Math. Statist.* **24** 284–289.

[4] Capon, J. (1961). Asymptotic efficiency of certain locally most powerful rank tests. *Ann. Math. Statist.* **29** 972–994.

[5] Chapman, D. G. (1958). A comparative study of several one-sided goodness-of-fit tests. *Ann. Math. Statist.* **29** 655–674.

[6] Chernoff, H. and Savage, I. R. (1958). Asymptotic normality and efficiency of certain nonparametric tests. *Ann. Math. Statist.* **29** 972–994.

[7] Govindarajulu, Z., LeCam, L. and Raghavachari, M. (1965). Generalizations of Chernoff-Savage theorems on asymptotic normality of nonparametric test statistics. *Proc. Fifth Berkeley Symp. Math. Statist. and Prob.* To appear.

[8] Hájek, J. (1962). Asymptotically most powerful rank-order tests. *Ann. Math. Statist.* **32** 1124–1147.

[9] Hodges, J. L. Jr. and Lehmann, E. L. (1956). The efficiency of some nonparametric competitors of the t-test. *Ann. Math. Statist.* **27** 324–335.

[10] Hodges, J. L. Jr. and Lehmann, E. L. (1961). Comparison of the normal scores and Wilcoxon tests. *Proc. Fourth Berkeley Symp. Math. Statist. and Prob.* **1** 307–317. Univ. of California Press.

[11] Hodges, J. L. Jr. and Lehmann, E. L. (1963). Estimates of location based on rank tests. *Ann. Math. Statist.* **34** 598–611.

[12] Hoeffding, W. (1951). "Optimum" nonparametric tests. *Proc. Second Berkeley Symp. Math. Statist. and Prob.* 83–92. Univ. of California Press.

[13] LeCam, L. (1960). Locally asymptotically normal families of distributions. *University of California Publications in Statistics* **3** 37–98.

[14] LEHMANN, E. L. (1953). The power of rank tests. *Ann. Math. Statist.* **24** 23–43.

[15] LEHMANN, E. L. (1959). *Testing Statistical Hypotheses.* Wiley, New York.

[16] MATTHES, T. K. and TRUAX, D. R. (1965). Optimal invariant rank tests for the k-sample problem. *Ann. Math. Statist.* **36** 1207–1222.

[17] PURI, M. L. (1964). Asymptotic efficiency of a class of c-sample tests. *Ann. Math. Statist.* **35** 102–121.

[18] SIEGEL, S. and TUKEY, J. W. (1960). A nonparametric sum of ranks procedure for relative spread in unpaired samples. *J. Amer. Statist. Assoc.* **55** 429–444.

[19] THOMPSON, RORY (1964). Simple alternatives with max and min power. San Diego State College. Unpublished.

[20] TUKEY, J. W. (1946). Sampling from contaminated distributions. *Ann. Math. Statist.* **17** 501.

[21] VAN EEDEN, C. (1963). The relation between Pitman's asymptotic relative efficiency of two tests and the correlation coefficient between their test statistics. *Ann. Math. Statist.* **34** 1442–1451.