ASYMPTOTIC EFFICIENCY OF CERTAIN LOCALLY MOST POWERFUL RANK TESTS

By Jack Capon

Federal Scientific Corporation

1. Introduction and Summary. We are given independent random samples $X_1, \cdots, X_m$ and $Y_1, \cdots, Y_n$ from populations with unknown cumulative distribution functions (cdf's) F_X and F_Y, respectively. It is desired to test

$$H_o: F_X = F_Y$$

against

$$H_1: F_X = G_\theta, \qquad F_Y = G_\phi, \qquad \theta, \phi \,\varepsilon\, R,$$

where G_θ is a specified family of cdf's (one for each θ), R is an interval on the real line, θ and ϕ are specified and very close to some specified value ϕ_o, and $\theta \neq \phi$.

A theorem of Hoeffding is used to show that the locally most powerful rank test (L.M.P.R.T.) of H_o against H_1 is based on a linear rank statistic

$$T_N = (m)^{-1} \sum_{i=1}^{N} a_{Ni} Z_{Ni},$$

where $Z_{Ni} = 1$ when the ith smallest of $N = m + n$ observations is an X, and $Z_{Ni} = 0$, otherwise, and the a_{Ni} are given numbers. In a recent paper, Chernoff and Savage established the asymptotic normality of the test statistic T_N, subject to some weak restrictions.

The concept of asymptotic relative efficiency (A.R.E.) was introduced by Pitman to compare sequences of tests. It was pointed out by Chernoff and Savage that the asymptotic efficiency of a sequence of tests can be established by means of a likelihood ratio test. Using this method, in conjunction with the theorem of Chernoff and Savage on asymptotic normality, it is shown that the L.M.P.R.T. of H_o against H_1 is asymptotically efficient. Several applications to Cauchy, exponential, and normal populations are given.

2. The Locally Most Powerful Rank Test. In our ensuing discussion we shall need the following regularity conditions:

(i) $G_\theta(x)$ has a density function $g_\theta(x)$, which, along with $\partial g_\theta(x)/\partial\theta$, is continuous with respect to θ for $\phi_o - a \leqq \theta \leqq \phi_o + a$, $a > 0$, for almost all x; there exist functions $M_0(x)$ and $M_1(x)$, integrable over $(-\infty, \infty)$, such that

$$g_\theta(x) \leqq M_0(x), \qquad |\partial g_\theta(x)/\partial\theta| \leqq M_1(x), \qquad \phi_o - a \leqq \theta \leqq \phi_o + a,$$

(ii) $g_\theta(x) > 0$ if and only if $g_\phi(x) > 0$,

(iii) $$|J^{(i)}(H)| = |d^iJ/\,dH^i| \leqq K(H(1-H))^{-i-\frac{1}{2}+\delta},$$

Received July 20, 1959; revised September 1, 1960.

for $i = 0, 1, 2$, and for some $\delta > 0$, where K is a constant, and where

$$J(G_{\phi_o}(x)) = \frac{\partial}{\partial \theta} \ln g_\theta(x)|_{\theta=\phi_o},$$

(iv) $$0 < \lim_{N\to\infty} m/n = r < \infty.$$

Condition (iii) has been termed a smoothness condition by Chernoff and Savage [1], and is essential in applying their theorem on the asymptotic normality of linear rank statistics.

If $\beta_N(\theta, \phi)$ denotes the power of a size-α rank test of H_o against H_1, then we define the rank test with power $\beta_N^*(\theta, \phi)$ as the L.M.P.R.T. for $\theta > \phi$, if

$$\beta_N^*(\theta, \phi) \geqq \beta_N(\theta, \phi)$$

uniformly in N, and for all θ and ϕ in some sufficiently small neighborhood of ϕ_o; i.e., $\theta, \phi \epsilon [\phi_o - \epsilon_N, \phi_o + \epsilon_N]$, $\epsilon_N > 0$. The L.M.P.R.T. for $\theta < \phi$ is defined in a similar manner. We note that the neighborhood of ϕ_o is allowed to vary with N.

If we arrange $X_1, \cdots, X_m, Y_1, \cdots, Y_n$ in order of increasing magnitude, and replace the ith smallest of this combined sample by a one if it is an X and by a zero otherwise, then we obtain a sequence of zeros and ones, $Z_{N1}, \cdots, Z_{NN}$. Such a sequence is termed an ordering. It was shown by Hoeffding [2] that, if condition (ii) is satisfied, the probability, under H_1, of obtaining an ordering $Z_{N1} = z_{N1}, \cdots, Z_{NN} = z_{NN}$, is

$$\begin{aligned} P_{\theta\phi}(Z_{N1} = z_{N1}, \cdots, Z_{NN} = z_{NN}) &= \binom{N}{n}^{-1} E_{\phi\phi}\left\{\prod_{i=1}^{N}\left[\frac{g_\theta(Z_i)}{g_\phi(Z_i)}\right]^{z_{Ni}}\right\} \\ &= \binom{N}{n}^{-1} E_{\theta\theta}\left\{\prod_{i=1}^{N}\left[\frac{g_\phi(Z_i)}{g_\theta(Z_i)}\right]^{y_{Ni}}\right\}, \end{aligned} \tag{1}$$

where Z_i is the ith smallest of the N observations, $y_{Ni} = 1 - z_{Ni}$, and $E_{\phi\phi}$ indicates that the expectation is taken under the assumption that $F_X = F_Y = G_\phi$. The probability of such an ordering, under H_o, is

$$P_{\phi_o\phi_o} = \binom{N}{n}^{-1}$$

We can expand $P_{\theta\phi}$ as

$$\begin{aligned} P_{\theta\phi} = P_{\phi_o\phi_o} &+ (\theta - \phi_o)\left.\frac{\partial P_{\theta\phi}}{\partial\theta}\right|_{\theta=\phi=\phi_o} + (\phi - \phi_o)\left.\frac{\partial P_{\theta\phi}}{\partial\phi}\right|_{\theta=\phi=\phi_o} \\ &+ o(|\theta - \phi_o| + |\phi - \phi_o|), \end{aligned} \tag{2}$$

provided that $P_{\theta\phi}$ has continuous partial derivatives in a neighborhood of $\theta = \phi = \phi_o$. It can be shown that the latter follows from condition (i).

As a consequence of condition (i) and a well known theorem ([9], p. 67), we may interchange differentiation and expectation to obtain from Eq. (1)

$$\left.\frac{\partial P_{\theta\phi}}{\partial\theta}\right|_{\theta=\phi=\phi_o} = \binom{N}{n}^{-1} E_{\phi_o\phi_o}\left\{\sum_{i=1}^{N}\frac{(\partial/\partial\theta)g_\theta(Z_i)\,|_{\theta=\phi_0}}{g_{\phi_o}(Z_i)} z_{Ni}\right\}$$

$$(3) \qquad = \binom{N}{n}^{-1}\sum_{i=1}^{N} E_{\phi_o\phi_o}\left(\frac{\partial}{\partial\theta}\ln g_\theta(Z_i)\Big|_{\theta=\phi_o}\right) z_{Ni}$$

$$= \binom{N}{n}^{-1}\sum_{i=1}^{N} a_{Ni}z_{Ni},$$

and similarly

$$(4) \qquad \left.\frac{\partial P_{\theta\phi}}{\partial\phi}\right|_{\theta=\phi=\phi_o} = \binom{N}{n}^{-1}\sum_{i=1}^{N} a_{Ni}y_{Ni},$$

where

$$(5) \qquad a_{Ni} = E_{\phi_o\phi_o}\left(\frac{\partial}{\partial\theta}\ln g_\theta(Z_i)\Big|_{\theta=\phi_0}\right).$$

Substituting Eqs. (3) and (4) in (2), we obtain

$$P_{\theta\phi} = \binom{N}{n}^{-1}$$

$$\cdot\left(1 + (\theta-\phi)\sum_{i=1}^{N} a_{Ni}z_{Ni} + (\phi-\phi_o)\sum_{i=1}^{N} a_{Ni} + o(|\theta-\phi_o| + |\phi-\phi_o|)\right).$$

We observe that $\sum_{i=1}^{N} a_{Ni}$ depends on N, but does not depend on the ordering $z_{N1}, \cdots, z_{NN}$, so that it may be considered a constant as far as any particular hypothesis testing problem is concerned.

Using the Neyman-Pearson fundamental lemma ([15], p. 65), we have that the most powerful rank test rejects H_o when

$$1 + (\theta-\phi)\sum_{i=1}^{N} a_{Ni}z_{Ni} + (\phi-\phi_o)\sum_{i=1}^{N} a_{Ni} + o(|\theta-\phi_o| + |\phi-\phi_o|) > c.$$

If $\theta > \phi$, the test is to reject H_o when

$$\sum_{i=1}^{N} a_{Ni}z_{Ni} + o(1) > c,$$

and, if $\theta < \phi$, the test is to reject H_o when

$$\sum_{i=1}^{N} a_{Ni}z_{Ni} + o(1) < c,$$

where c is a constant chosen to give the test size α, and is not necessarily the same from one line to the next. Since there are a finite number of orderings that can be obtained, namely $\binom{N}{n}$, we have that the L.M.P.R.T. rejects H_o, if $\theta > \phi$, when

$$(6) \qquad T_N = \frac{1}{m}\sum_{i=1}^{N} a_{Ni}z_{Ni} > c,$$

and rejects H_o, if $\theta < \phi$, when

$$T_N < c. \tag{7}$$

The statistic T_N defined in Eq. (6) is known as a linear rank statistic. The constant a_{Ni} is the expected value of a certain function, $\partial \ln g_\theta(x)/\partial\theta|_{\theta=\phi_o}$, of the ith smallest observation of a sample of size N from the cdf G_{ϕ_o}. The purpose of condition (iii) is to insure that this expected value exists for all i.

We define the function $h_{\theta\phi}(u)$ as $G_\theta(x) = h_{\theta\phi}(G_\phi(x))$. This function can always be obtained, since we can always think of G_θ and G_ϕ as related by $h_{\theta\phi}$, no matter how complicated $h_{\theta\phi}$ may be. Since both $g_\theta(x)$ and $g_\phi(x)$ exist for all x, we may write

$$g_\theta(x)/g_\phi(x) = h'_{\theta\phi}(G_\phi(x)), \tag{8}$$

where

$$h'_{\theta\phi}(u) = \partial h_{\theta\phi}(u)/\partial u.$$

We have from Eq. (8) that

$$\frac{\partial}{\partial\theta} \ln g_\theta(x)\,|_{\theta=\phi_o} = \frac{\partial}{\partial\theta} h'_{\theta\phi}(G_\phi(x))\,|_{\theta=\phi=\phi_o},$$

and hence

$$\begin{aligned} a_{Ni} &= E_{\phi_o\phi_o}\left\{\frac{\partial}{\partial\theta} h'_{\theta\phi}(G_\phi(Z_i))\,|_{\theta=\phi=\phi_o}\right\} \\ &= E\left\{\frac{\partial}{\partial\theta} h'_{\theta\phi}(U_i)\,|_{\theta=\phi=\phi_o}\right\}, \end{aligned} \tag{9}$$

where U_i is the ith smallest observation of a sample of size N from the uniform distribution on (0, 1). Thus, a_{Ni} is the expected value of a certain function, $\partial h'_{\theta\phi}(u)/\partial\theta|_{\theta=\phi=\phi_o}$, of the ith smallest observation of a sample of size N from the uniform distribution. This expected value exists, since we have assumed condition (iii) to be satisfied.

It is observed that Eqs. (5) and (9) are identical. The appropriate one to use depends on the ease of application. If $h_{\theta\phi}(u)$ is a complicated function, Eq. (5) is used, and, if $h_{\theta\phi}(u)$ is a simple function, Eq. (9) is used. In the latter case we say that we are dealing with a functional alternative.

Similar results have been obtained by Lehmann [3] for various functional alternatives, by Hoeffding [2] and Terry [4] when $G_\theta(x)$ is a normal cdf, and by Savage [5] when $G_\theta(x)$ is an exponential cdf. A generalized approach, somewhat different than ours, but which leads to essentially the same results presented above has been given by Pyke [6].

We obtain from Theorem 1 of Chernoff and Savage [1], and a simple extension of their Theorem 2, that the linear rank statistic T_N has asymptotically a normal distribution, if the smoothness condition (iii) is satisfied; i.e.,

$$\lim_{N\to\infty} \text{Prob}\left\{\frac{T_N - E_{\theta\phi}(T_N)}{\sigma_{\theta\phi}(T_N)} \leqq t\right\} = \int_{-\infty}^{t} (2\pi)^{-\frac{1}{2}} \exp(-x^2/2)\,dx,$$

where

$$E_{\theta\phi}(T_N) = \int_{-\infty}^{\infty} J(H_{\theta\phi}(x))\, dG_\theta(x), \tag{10}$$

$$\begin{aligned} N\sigma^2_{\theta\phi}(T_N) = \frac{2n}{N}\Bigg\{ \iint_{-\infty<x<y<\infty} & G_\phi(x)(1 - G_\phi(y)) \\ & \cdot J'(H_{\theta\phi}(x))J'(H_{\theta\phi}(y))\, dG_\theta(x)\, dG_\theta(y) \\ + \frac{n}{m}\iint_{-\infty<x<y<\infty} & G_\theta(x)(1 - G_\theta(y))J'(H_{\theta\phi}(x))J'(H_{\theta\phi}(y))\, dG_\phi(x)\, dG_\phi(y)\Bigg\}, \end{aligned} \tag{11}$$

and where

$$\begin{aligned} H_{\theta\phi}(x) &= (n/N)G_\phi(x) + (m/N)G_\theta(x), \\ J(G_{\phi_o}(x)) &= \partial \ln g_\theta(x)/\partial\theta|_{\theta=\phi_o}, \\ J'(u) &= dJ(u)/\,du, \end{aligned}$$

providing $\sigma_{\theta\phi}(T_N) \neq 0$.

The variance of the limiting distribution of T_N under H_o is obtained from Eq. (11) by letting $\theta = \phi = \phi_o$; thus, if we denote this variance by $\sigma^2_{\phi_o}(T_N)$, we have

$$\frac{mN}{n}\sigma^2_{\phi_o}(T_N) = 2\iint_{0<x<y<1} x(1 - y)J'(x)J'(y)\, dx\, dy \tag{12}$$

$$= \int_o^1 J^2(x)\, dx - \left(\int_o^1 J(x)\, dx\right)^2. \tag{13}$$

As was pointed out by Chernoff and Savage [1], Eq. (13) can be obtained from Eq. (12) by interpreting the double integral in Eq. (12) as

$$\iiiint_{0<u<x<y<v<1} J'(x)J'(y)\, du\, dx\, dy\, dv,$$

and integrating with respect to y first and x second.

If we now use the definition for $J(G_{\phi_o}(x))$, we obtain from Eq. (13) that

$$\begin{aligned} (mN/n)\sigma^2_{\phi_o}(T_N) = E_{\phi_o}&\left(\left(\frac{\partial}{\partial\theta}\ln g_\theta(X)\right)^2\Bigg|_{\theta=\phi_o}\right) \\ & - \left(E_{\phi_o}\left(\frac{\partial}{\partial\theta}\ln g_\theta(X)\right)\Bigg|_{\theta=\phi_o}\right)^2, \end{aligned} \tag{14}$$

where E_{ϕ_o} indicates that the expectation is taken under the assumption that the cdf of X is G_{ϕ_o}.

As a consequence of condition (i) we may interchange differentiation and integration ([9], p. 67), to obtain

$$\begin{aligned} E_{\phi_o}\left(\left(\frac{\partial}{\partial\theta}\ln g_\theta(X)\,|_{\theta=\phi_o}\right)\right) &= \int_{-\infty}^{\infty}\frac{\partial}{\partial\theta} g_\theta(x)\,|_{\theta=\phi_o}\, dx \\ &= \frac{\partial}{\partial\theta}\int_{-\infty}^{\infty} g_\theta(x)\, dx\,|_{\theta=\phi_o} = \frac{\partial(1)}{\partial\theta}\,|_{\theta=\phi_o} = 0. \end{aligned} \tag{15}$$

Hence Eq. (14) becomes

$$(mN/n)\sigma^2_{\phi_o}(T_N) = \inf G_{\phi_o}, \tag{16}$$

where

$$\inf G_{\phi_o} = E_{\phi_0}\left(\left(\frac{\partial}{\partial\theta}\ln g_\theta(X)\right)^2\bigg|_{\theta=\phi_o}\right).$$

The quantity $\inf G_{\phi_o}$ is known as the information of the cdf G_θ evaluated at $\theta = \phi_o$, and will be used repeatedly in our discussions.

In a subsequent analysis we shall need to establish that $\sigma^2_{\theta\phi}(T_N)$ is continuous at the point (ϕ_o, ϕ_o), uniformly in N. This is true if the two integrals in Eq. (11) are continuous at (ϕ_o, ϕ_o). That the first integral in (11) is continuous at (ϕ_o, ϕ_o) can be seen as follows. If we let

$$u = G_\theta(x), \qquad v = G_\theta(y),$$

$$G^*_{\theta\phi}(u) = G_\phi(G_\theta^{-1}(u)), \qquad H^*_{\theta\phi}(u) = (n/N)u + (m/N)G^*_{\theta\phi}(u),$$

the integral can be written as

$$\int\int_{0<u<v<1} R(\theta, \phi, u, v)\, du\, dv,$$

where $R(\theta, \phi, u, v) = G^*_{\theta\phi}(u)(1 - G^*_{\theta\phi}(v))J'(H^*_{\theta\phi}(u))J'(H^*_{\theta\phi}(v))$. It follows from conditions (i) and (iii) that $R(\theta, \phi, u, v)$ is continuous with respect to (θ, ϕ) at (ϕ_o, ϕ_o) for almost all u, v. It follows from a well known theorem ([9], p. 67), that it is sufficient to show that $|R(\theta, \phi, u, v)|$ is bounded by a function integrable over $0 < u < v < 1$, which is independent of θ and ϕ. Since

$$G^*_{\theta\phi} \leqq (N/m)H^*_{\theta\phi} \quad \text{and} \quad 1 - G^*_{\theta\phi} \leqq (N/m)(1 - H^*_{\theta\phi}),$$

we obtain, in conjunction with condition (iii), that

$$|R(\theta, \phi, u, v)| \leqq K^2(N/m)^2H^*_{\theta\phi}(u)^{-\frac{1}{2}+\delta}(1 - H^*_{\theta\phi}(v))^{-\frac{1}{2}+\delta}(1 - H^*_{\theta\phi}(u))^{-\frac{3}{2}+\delta}H^*_{\theta\phi}(v)^{-\frac{3}{2}+\delta}.$$

With no loss of generality we may assume that $\delta < \frac{1}{2}$; since $H^*_{\theta\phi}(u) \geqq (n/N)u$, and $1 - H^*_{\theta\phi}(u) \geqq (n/N)(1 - u)$, we have

$$|R(\theta, \phi, u, v)| \leqq K^2(N/m)^2\,(N/n)^{4-4\delta}u^{-\frac{1}{2}+\delta}(1 - v)^{-\frac{1}{2}+\delta}(1 - u)^{-\frac{3}{2}+\delta}v^{-\frac{3}{2}+\delta}. \tag{17}$$

The bound in (17) is independent of θ and ϕ and is easily seen to be integrable over $0 < u < v < 1$. The proof for the second integral in (11) is analogous.

3. Asymptotic Relative Efficiency of Test Procedures. We are now in a position to use the Pitman [7], [8], criterion for finding efficiencies of test procedures based on sequences of statistics $\{W_N\}$. We let $\Delta = \theta - \phi$, and we assume that the following conditions are true in some neighborhood of $\Delta = 0$, $\theta = \phi = \phi_o$:

(a) $\mathfrak{L}((W_N - E_{\theta\phi}(W_N))/\sigma_{\theta\phi}(W_N)) \Rightarrow N(0, 1)$,

(b) for the sequence of alternatives $\{\Delta_N\}$, where $\Delta_N = \theta_N - \phi_N = kN^{-\frac{1}{2}}$, k is a non-zero constant, and $(\phi_N - \phi_o)/(\theta_N - \phi_o) = -m/n$,

$$\lim_{N\to\infty} \frac{\sigma_{\theta_N\phi_N}(W_N)}{\sigma_{\phi_o}(W_N)} = 1,$$

and

$$E_W = \lim_{N\to\infty} \left\{\frac{E_{\theta_N\phi_N}(W_N) - E_{\phi_o\phi_o}(W_N)}{\Delta_N(mn/N)^{\frac{1}{2}}\sigma_{\phi_o}(W_N)}\right\}^2 = \lim_{N\to\infty} \left\{\frac{(\partial/\partial\Delta)E_{\theta\phi}(W_N)}{(mn/N)^{\frac{1}{2}}\sigma_{\phi_o}(W_N)}\right\}^2_{\substack{\Delta=o\\ \theta=\phi=\phi_o}}$$

exists, and is independent of k.

The quantity E_W has been termed the efficacy of the test procedure based on the sequence of statistics $\{W_N\}$. When we compare two sequences of tests, say $\{W_N\}$ and $\{W_N^*\}$, for the same pair of near alternatives given in (b), we find that the two tests will have the same power only when the corresponding sample sizes, N and N^* satisfy the relationship

$$\lim_{N\to\infty} \frac{N^*}{N} = \frac{E_W}{E_{W^*}} = E_{W,W^*}, \tag{18}$$

if $E_{W^*} \neq 0$, and $\lim_{N\to\infty} m/n = \lim_{N^*\to\infty} m^*/n^* = r$. E_{W,W^*} is called the A.R.E. of the $\{W_N\}$-test with respect to the $\{W_N^*\}$-test.

Chernoff and Savage have pointed out (see footnote on p. 983 of [1]) that no invariant test of $\Delta = \Delta_N$ vs. $\Delta = 0$ can have greater efficacy than the likelihood ratio test for testing $\Delta = \Delta_N$ (when the densities of X and Y are g_{θ_N} and g_{ϕ_N}, respectively) against $\Delta = 0$ (when the densities of X and Y are both equal to g_{ϕ_o}). The test is to reject H_o when

$$\prod_{i=1}^{m} g_{\theta_N}(X_i)/g_{\phi_o}(X_i) \prod_{i=1}^{n} g_{\phi_N}(Y_i)/g_{\phi_o}(Y_i) > c$$

$$\sum_{i=1}^{m} \ln \frac{g_{\theta_N}(X_i)}{g_{\phi_o}(X_i)} + \sum_{i=1}^{n} \ln \frac{g_{\phi_N}(Y_i)}{g_{\phi_o}(Y_i)} > c \tag{19}$$

$$L_N = \frac{1}{m}\sum_{i=1}^{m} \frac{\partial}{\partial\theta} \ln g_\theta(X_i)\,|_{\theta=\phi_o} - \frac{1}{n}\sum_{i=1}^{n}\left(\frac{\partial}{\partial\phi} \ln g_\phi(Y_i)\,|_{\phi=\phi_o} + o_N(1)\right) > c,$$

where $\lim_{N\to\infty} o_N(1) = 0$. We see that, except for the $o_N(1)$ term, L_N is equal to the difference of two sums of independent and identically distributed random variables. If $\partial \ln g_\theta(X)/\partial\theta|_{\theta=\phi_o}$ and $\partial \ln g_\phi(Y)/\partial\phi|_{\phi=\phi_o}$ have finite variances in some neighborhood of $\Delta = 0$, $\theta = \phi = \phi_o$, then L_N is asymptotically normal and condition (a) is satisfied. We have

$$\begin{aligned} E_{\theta\phi}(L_N) &= E_\theta\left(\frac{\partial}{\partial\theta} \ln g_\theta(X)\,|_{\theta=\phi_o}\right) - E_\phi\left(\frac{\partial}{\partial\phi} \ln g_\phi(Y)\,|_{\phi=\phi_o}\right) + o_N(1) \\ &= (\theta - \phi_o)E_{\phi_o}\left(\frac{\partial}{\partial\theta} \ln g_\theta(X)\,|_{\theta=\phi_o}\right)^2 \\ &\qquad - (\phi - \phi_o)E_{\phi_o}\left(\frac{\partial}{\partial\phi} \ln g_\phi(Y)\,|_{\phi=\phi_o}\right)^2 + o_N(1), \end{aligned}$$

$$\frac{\partial}{\partial\Delta} E_{\theta\phi}(L_N)\,|_{\substack{\Delta=0\\ \theta=\phi=\phi_o}} = \inf G_{\phi_o} + o_N(1),$$

and

$$\sigma^2_{\phi_o}(L_N) = (m^{-1} + n^{-1})(\inf G_{\phi_o} + o_N(1)).$$

Hence

$$E_L = \inf G_{\phi_o} . \tag{20}$$

We shall say that a test is asymptotically efficient, for the sequence of alternatives $\{\Delta_N\}$, if its efficacy achieves the upper bound in Eq. (20), namely $\inf G_{\phi_o}$. We shall now show that the L.M.P.R.T. is asymptotically efficient.

4. Asymptotic Efficiency of the Locally Most Powerful Rank Test. We have already seen that if the smoothness condition (iii) is satisfied, then T_N is asymptotically normal. This implies that condition (a) is satisfied. Since we have shown that $\sigma^2_{\theta\phi}(T_N)$ is continuous at the point (ϕ_o, ϕ_o), uniformly in N, we obtain

$$\lim_{N\to\infty} \frac{\sigma_{\theta_N\phi_N}(T_N)}{\sigma_{\phi_o}(T_N)} = 1,$$

which shows that the first part of condition (b) is satisfied. We shall show that the rest of condition (b) is fulfilled by calculating the efficacy E_T, and showing that it exists.

We may use the mean value theorem to write

$$H_{\theta\phi}(x) = G_\theta(x) + \frac{n}{N}(G_\phi(x) - G_\theta(x)) = G_\theta(x) - \frac{n\Delta}{N}\frac{\partial}{\partial u} G_u(x)\,|_{u=\hat{\phi}}, \tag{21}$$

where $\hat{\phi}$ is between θ and ϕ. If we use Eq. (21) in (10) we obtain

$$E_{\theta\phi}(T_N) = \int_{-\infty}^{\infty} J\left(G_\theta(x) - \frac{n\Delta}{N}\frac{\partial}{\partial u} G_u(x)\,|_{u=\hat{\phi}}\right) dG_\theta(x). \tag{22}$$

Using the mean value theorem still one more time we can rewrite Eq. (22) as

$$\begin{aligned} E_{\theta\phi}(T_N) &= \int_{-\infty}^{\infty} J(G_\theta(x))\, dG_\theta(x) \\ &\quad - \frac{n\Delta}{N}\int_{-\infty}^{\infty}\frac{\partial}{\partial u} G_u(x)\,|_{u=\hat{\phi}}\, J'(G_u(x))\,|_{G_u=\hat{G}_\theta}\, dG_\theta(x), \end{aligned} \tag{23}$$

where $\hat{G}_\theta$ is between G_θ and $G_\theta - (n/N)\Delta\partial G_u/\partial u|_{u=\hat{\phi}}$.

The first integral in Eq. (23) can be evaluated as

$$\begin{aligned} \int_{-\infty}^{\infty} J(G_\theta(x))\, dG_\theta(x) &= E_{\phi_o}\left(\frac{\partial}{\partial\theta}\ln g_\theta(X)\,|_{\theta=\phi_o}\right) = \int_{-\infty}^{\infty}\frac{\partial}{\partial\theta} g_\theta(x)\,|_{\theta=\phi_o}\, dx \\ &= \frac{\partial}{\partial\theta}\int_{-\infty}^{\infty} g_\theta(x)\, dx\,|_{\theta=\phi_o} = 0. \end{aligned} \tag{24}$$

Hence we obtain from Eq. (23) that

$$
\begin{aligned}
\frac{\partial}{\partial\Delta} E_{\theta\phi}(T_N) \Big|_{\substack{\Delta=0 \\ \theta=\phi=\phi_o}} &= -\left(\frac{n}{N}\right) \int_{-\infty}^{\infty} \left(\frac{\partial}{\partial\theta} G_\theta(x)\right) J'(G_\theta(x))\, dG_\theta(x) \big|_{\theta=\phi_o} \\
&= -\left(\frac{n}{N}\right) \int_{-\infty}^{\infty} \left(\frac{\partial}{\partial\theta} J(G_\theta(x))\right) dG_\theta(x) \big|_{\theta=\phi_o} \\
&= -\left(\frac{n}{N}\right) E_{\phi_o} \left(\frac{\partial^2}{\partial\theta^2} \ln g_\theta(X) \big|_{\theta=\phi_o}\right) \qquad (25) \\
&= \left(\frac{n}{N}\right) E_{\phi_o} \left(\frac{\partial}{\partial\theta} \ln g_\theta(X) \big|_{\theta=\phi_o}\right)^2 \\
&= \left(\frac{n}{N}\right) \inf G_{\phi_o}.
\end{aligned}
$$

Therefore, we have from Eqs. (16) and (25) that the efficacy of the L.M.P.R.T. is

$$E_T = \inf G_{\phi_o}. \qquad (26)$$

Thus, the L.M.P.R.T. of H_o against H_1 is asymptotically efficient. Chernoff and Savage [1] have obtained the same result for the particular case when $G_\theta(x) = F(x - \theta)$; i.e., translation alternatives.

Our results, of course, hold true for the case when there is a simple functional relationship between G_θ and G_ϕ; i.e., in the case of functional alternatives the L.M.P.R.T. is asymptotically efficient. In this case the efficacy may also be expressed as

$$E_T = E\left(\frac{\partial}{\partial\theta} h'_{\theta\phi}(U) \big|_{\theta=\phi=\phi_o}\right)^2 = \int_0^1 \left(\frac{\partial}{\partial\theta} h'_{\theta\phi}(u) \big|_{\theta=\phi=\phi_o}\right)^2 du. \qquad (27)$$

5. Applications. We now give applications of our results to some specific cases. In each example a straightforward calculation shows that the regularity conditions (i)–(iii) are satisfied. It is noted that in each case the form of the test does not depend on ϕ_o.

(A) *Exponential Case* (*Scalar*). We let

$$
\begin{aligned}
G_\theta(x) &= 1 - \exp(-\theta x), && x \geqq 0 \\
&= 0, && x < 0, \quad \theta, \phi, \phi_o \in (0, \infty),
\end{aligned}
$$

so that

$$
\begin{aligned}
g_\theta(x) &= \theta \exp(-\theta x), && x \geqq 0 \\
&= 0, && x < 0, \\
\frac{\partial}{\partial\theta} \ln g_\theta(x) \big|_{\theta=\phi_o} &= \phi_o^{-1} - x, && x \geqq 0 \\
&= 0, && x < 0, \\
J(v) &= \phi_o^{-1}(1 + \ln(1 - v)), && 0 \leqq v \leqq 1.
\end{aligned}
$$

Hence, the L.M.P.R.T. for this problem is based on a linear rank statistic S_N, with a_{Ni} given by $a_{Ni} = E(\eta_i)$, where η_i is the ith smallest observation of a sample of size N from the exponential distribution $G_1(x)$. It can be shown [10] that in this case a_{Ni} is given by

$$a_{Ni} = \sum_{j=N-i+1}^{N} j^{-1}.$$

Therefore, we have that

$$S_N = \frac{1}{m} \sum_{i=1}^{N} \left(\sum_{j=N-i+1}^{N} j^{-1} \right) Z_{Ni} .$$

If $\theta > \phi$, the test is to reject H_o when $S_N < c$, and if $\theta < \phi$ the test is to reject H_o when $S_N > c$.

The statistic S_N is asymptotically normal, and the $\{S_N\}$-test is asymptotically efficient. The S_N-test was proposed originally by Savage [5], who showed that it was the L.M.P.R.T. of H_o against H_1. The asymptotic normality of S_N has been shown by Chernoff and Savage [1].

(B) *Normal Case* (*Translation*). Set

$$g_\theta(x) = (2\pi\sigma^2)^{-\frac{1}{2}} \exp\left(-\frac{1}{2}\left(\frac{x-\theta}{\sigma}\right)^2 \right), \qquad \theta, \phi, \phi_o \,\varepsilon\, (-\infty, \infty).$$

We have

$$\frac{\partial}{\partial\theta} \ln g_\theta(x) \,|_{\theta=\phi_o} = \frac{x - \phi_o}{\sigma},$$

$$J(v) = \Phi^{-1}(v), \qquad 0 \leqq v \leqq 1,$$

where

$$\Phi(x) = (2\pi)^{-\frac{1}{2}} \int_{-\infty}^{x} \exp\left(-\tfrac{1}{2}y^2\right) dy,$$

and $\Phi^{-1}(v)$ represents the inverse function to $\Phi(x)$; i.e., $v = \Phi(x)$. Thus, the L.M.P.R.T. is based on the linear rank statistic c_N defined as

$$c_N = \frac{1}{m} \sum_{i=1}^{N} E(\xi_i) Z_{Ni},$$

where ξ_i is the ith smallest observation of a sample of size N from the $N(0, 1)$ distribution. If $\theta > \phi$, the test is to reject H_o when $c_N > c$, and if $\theta < \phi$, the test is to reject H_o when $c_N < c$. This test is also known as the c_1-test.

The statistic c_N is asymptotically normal, and the $\{c_N\}$-test is asymptotically efficient. The c_1-test was proposed originally by Fisher and Yates [11], and shown to be locally most powerful by Hoeffding [2] and Terry [4]. The asymptotic normality of c_N, and the asymptotic efficiency of the $\{c_N\}$-test have been established by Hoeffding (pp. 289–292 of [12]) and by Chernoff and Savage [1]. In addition, it is shown in [1] that the A.R.E. of the c_1-test with respect to the t-test, for non-normal translation alternatives, is strictly greater than one.

(C) *Normal Case* (*Scalar*). Let

$$g_\theta(x) = (2\pi\theta)^{-\frac{1}{2}} \exp[-\tfrac{1}{2}(x-u)^2/\theta], \qquad \theta, \phi, \phi_o \, \varepsilon \, (0, \infty),$$

so that

$$\frac{\partial}{\partial\theta} \ln g_\theta(x)_{\theta=\phi_o} = -\frac{1}{2}\phi_o^{-1} + \frac{1}{2}\left(\frac{x-u}{\phi_o}\right)^2,$$

$$J(v) = \tfrac{1}{2}\phi_o^{-1}((\Phi^{-1}(v))^2 - 1), \qquad 0 \leqq v \leqq 1.$$

Hence the L.M.P.R.T. is based on the linear rank statistic F_N defined as

$$F_N = \frac{1}{m}\sum_{i=1}^{N} E(\xi_i^2) Z_{Ni},$$

where ξ_i is the ith smallest observation of a sample of size N from the $N(0, 1)$ distribution.

The statistic F_N is asymptotically normal, and the $\{F_N\}$-test is asymptotically efficient. If $\theta > \phi$, the test is to reject H_o when $F_N > c$, and if $\theta < \phi$, the test is to reject H_o when $F_N < c$.

(D) *Cauchy Case* (*Translation*). Let

$$g_\theta(x) = [\pi(1 + (x-\theta)^2]^{-1}, \qquad \theta, \phi, \phi_o \, \varepsilon \, (-\infty, \infty),$$

so that

$$\frac{\partial}{\partial\theta} \ln g_\theta(x) \,|_{\theta=\phi_o} = \frac{2(x-\phi_o)}{1 + (x-\phi_o)^2},$$

$$J(v) = \frac{2\tan\pi(v - \frac{1}{2})}{1 + \tan^2\pi(v - \frac{1}{2})}, \qquad 0 \leqq v \leqq 1.$$

Thus the L.M.P.R.T. is based on the linear rank statistic Q_N defined as

$$Q_N = \frac{1}{m}\sum_{i=1}^{N} E\left(\frac{\mu_i}{1+\mu_i^2}\right) Z_{Ni},$$

where μ_i is the ith smallest of N observations from the Cauchy distribution $G_0(x)$.

If $\theta > \phi$, the test is to reject H_o when $Q_N > c$, and if $\theta < \phi$, the test is to reject H_o when $Q_N < c$. The statistic Q_N is asymptotically normal, and the $\{Q_N\}$-test is asymptotically efficient.

(E) *Cauchy Case* (*Scalar*). Let

$$g_\theta(x) = \theta/[\pi(1 + \theta^2 x^2)], \qquad \theta, \phi, \phi_o \, \varepsilon \, (0, \infty),$$

so that

$$\frac{\partial}{\partial\theta} \ln g_\theta(x) \,|_{\theta=\phi_o} = \phi_o^{-1} - \frac{2\phi_o x^2}{1 + \phi_o^2 x^2},$$

$$J(v) = \phi_o^{-1}\left[1 - \frac{2\tan^2\pi(v - \frac{1}{2})}{1 + \tan^2\pi(v - \frac{1}{2})}\right], \qquad 0 \leqq v \leqq 1.$$

Therefore the L.M.P.R.T. is based on the linear rank statistic R_N, defined as

$$R_N = \frac{1}{m} \sum_{i=1}^{N} E\left(\frac{\mu_i^2}{1 + \mu_i^2}\right) Z_{Ni},$$

where μ_i is the ith smallest of N observations from the Cauchy distribution $G_1(x)$.

If $\theta > \phi$, the test is to reject H_o when $R_N < c$, and if $\theta < \phi$, the test is to reject H_o when $R_N > c$. The statistic R_N is asymptotically normal, and the $\{R_N\}$-test is asymptotically efficient.

(F) *The Mann-Whitney-Wilcoxon Test.* In this case we have a functional alternative, where

$$h_{\theta\phi}(u) = (1 - \theta + \phi))u + (\theta - \phi)u^2,$$

$$\theta, \phi, \phi_o \, \varepsilon \, (-\infty, \infty), \quad 0 < \theta - \phi < 1,$$

so that

$$J(u) = \frac{\partial}{\partial\theta} h'_{\theta\phi}(u) \,|_{\theta=\phi=\phi_o} = 2u - 1, \qquad 0 \leqq u \leqq 1.$$

The L.M.P.R.T. is based on the statistic V_N defined as

$$V_N = \frac{1}{m} \sum_{i=1}^{N} E(U_i) \, Z_{Ni},$$

where U_i is the ith smallest of N observations from the uniform distribution on $(0, 1)$. It can be shown [10] that

$$E(U_i) = i/(N + 1),$$

and hence V_N can be written as

$$V_N = \frac{1}{m(N + 1)} \sum_{i=1}^{N} i \, z_{Ni},$$

Since $\theta > \phi$, the test is to reject H_o when $V_N > c$.

This result was obtained originally by Lehmann [3], who also pointed out that the V_N-test is equivalent to the Mann-Whitney-Wilcoxon test [13], [14].

The statistic V_N is asymptotically normal, and the $\{V_N\}$-test is asymptotically efficient.

Acknowledgment. The author would like to thank Professor Wassily Hoeffding and the referee for many helpful comments and criticisms concerning this work. In particular, the author is indebted to Professor Hoeffding for certain clarifying remarks about the regularity conditions, and for the proof of the fact that $\sigma_{\theta\phi}^2(T_N)$ is continuous at (ϕ_o, ϕ_o). This work is based on results that were obtained in a dissertation entitled "Nonparametric methods for the detection of signals in noise," which was submitted in partial fulfillment of the requirements for the Ph.D. in electrical engineering at Columbia University, in 1959. This dissertation was supported by the U. S. Air Force under Contract No. AF

19(604)-4140, and monitored by the office of Scientific Research, Air Research and Development Command.

REFERENCES

[1] H. Chernoff and I. R. Savage, "Asymptotic normality and efficiency of certain nonparametric test statistics," *Ann. Math. Stat.*, Vol. 29 (1958), pp. 972–994.

[2] W. Hoeffding, " 'Optimum' nonparametric tests," *Proceedings of the Second Berkeley Symposium on Probability and Statistics*, University of California Press, Berkeley, 1951, pp. 83–92.

[3] E. L. Lehmann, "The power of rank tests," *Ann. Math. Stat.*, Vol. 24 (1953), pp. 23–43.

[4] M. E. Terry, "Some rank-order tests which are most powerful against specific parametric alternatives," *Ann. Math. Stat.*, Vol. 23 (1952), pp. 346–366.

[5] I. R. Savage, "Contributions to the theory of rank-order statistics: two-sample case," *Ann. Math. Stat.*, Vol. 27 (1956), pp. 590–616.

[6] R. Pyke, *Lecture Notes on Nonparametric Statistical Inference*, Columbia University, Fall 1958.

[7] E. J. G. Pitman, *Lecture notes on nonparametric statistical inference*, Columbia University, Spring 1948.

[8] G. E. Noether, "On a theorem of Pitman," *Ann. Math. Stat.*, Vol. 26 (1955), pp. 64–68.

[9] H. Cramér, *Mathematical Methods of Statistics*, Princeton University Press, Princeton, N. J., 1951.

[10] E. J. Gumbel, *Statistics of Extremes*, Columbia University Press, New York, 1958.

[11] R. A. Fisher and F. Yates, *Statistical Tables for Biological, Agricultural and Medical Research*, 3rd Ed., Hafner Pub. Co., New York, 1959.

[12] D. A. S. Fraser, *Nonparametric Methods in Statistics*, John Wiley and Sons, New York, 1957.

[13] H. B. Mann and D. R. Whitney, "On a test of whether one of two random variables is stochastically larger than the other," *Ann. Math. Stat.*, Vol. 18 (1947), pp. 50–60.

[14] F. Wilcoxon, "Individual comparisons by ranking methods," *Biometric Bull.*, Vol. 1 (1945), pp. 80–83.

[15] E. L. Lehmann, *Testing Statistical Hypotheses*, John Wiley and Sons, New York, 1959.