ON THE LARGE SAMPLE PROPERTIES OF A GENERALIZED WILCOXON-MANN-WHITNEY STATISTIC[1]

By A. P. Basu[2]

University of Minnesota and University of Wisconsin

0. Summary. Let there be two independent samples of sizes m and n respectively from two populations with continuous cdf's $F(x)$ and $G(y)$. To test the equality of the two populations Sobel [19], has proposed the statistic $V_r^{m,n}$ (to be defined later) based on the first r ordered observations only. In this paper the large sample properties of $V_r^{m,n}$ have been studied. The test is compared with other "r out of N" tests by computing the appropriate asymptotic relative efficiencies. The test statistic is found to be quite satisfactory in all the cases considered and is particularly suitable for location alternatives.

1. Introduction. Let $X_1, X_2, \cdots, X_m$ and $Y_1, Y_2, \cdots, Y_n$ be two independent samples of sizes m and n from two populations with continuous cumulative distribution functions (cdf's) $F(x)$ and $G(y)$, where F and G belong to the same family $\mathfrak{F}$ of distribution functions indexed by a parameter θ. We wish to test the hypothesis

$$H_0 : F = G \tag{1.1}$$

against the alternative that they are different.

Let all the $m + n = N$ observations be ordered in a sequence and suppose we want to base a decision on (at most) the first r of the combined set of N observations, i.e. we have a right-censored sample of size at most r.

Such a censored sample occurs naturally in many physical situations, as for example, in problems of life testing where we are interested in comparing the mean life of two physical systems, or in clinical trials or bio-assay problems where we want to compare the efficacy of two drugs but we can not afford to wait indefinitely to get information on all the sampling units put on test. For facility of discussion, we shall use the terminology of life testing. Any test based on the first r ordered observations (out of a combined sample of size N) will be termed an "r out of N" test.

For the above problem Sobel [19], has proposed a statistic $V_r^{m,n}$ which we now introduce. Let m_i and n_i be the number of x and y failures, respectively, among the first i ordered observations of the combined sample, so that

$$m_i + n_i = i, \qquad i = 1, 2, \cdots, r. \tag{1.2}$$

Received 27 June 1966; revised 8 December 1966.

[1] This paper is a part of the author's doctoral thesis submitted at the University of of Minnesota (1966) and has been supported by the National Science Foundation under Grant No. GP-3813.

[2] Now at the University of Wisconsin.

These observations (x's and y's) are the failure times in a life testing experiment. The proposed statistic is given by

$$\mathrm{V}_r^{m,n} = V_r^{(N)} = \textstyle\sum_{i=1}^{r} (nm_i - mn_i). \tag{1.3}$$

In [20] this statistic $V_r^{(N)}$ is shown to be related to the well-known Wilcoxon-Mann-Whitney statistic [13], [22] and the small sample properties of $V_r^{(N)}$ and its exact and asymptotic distribution under the null hypothesis are also discussed.

In view of the usefulness of the above statistic it seems desirable to explore the properties of this statistic further; the object of this paper is to establish some large-sample properties of the statistic $V_r^{(N)}$. In Section 3 we prove the asymptotic normality in the null and non-null case of a statistic (defined in Section 2) which is equivalent to $\mathrm{V}_r^{(N)}$. Consistency of the test statistic is established in Section 4. In Section 5 general expressions for the efficacy of the test are given and in Section 6 we derive the asymptotic relative efficiency (ARE) of the above test with respect to the likelihood ratio test for testing the scale parameter in the case of the exponential distribution. The performance of the test has been compared with other asymptotically most powerful rank tests for censored sampling in Section 7. In Section 8 a modified Sobel statistic [21] has also been studied.

Some studies in this direction have recently been made by Halperin [12] and Gehan [9]. However, they consider censoring schemes in which the experiment is terminated after a given period so that r, the number of uncensored observations, becomes a random variable.

2. Relation of $V_r^{(N)}$ to other statistics. To facilitate discussion we shall first define a new sequence $\{z_i\}$ ($i = 1, 2, \cdots, N$) derived from the combined ordered sample, always counting ordered observations from the left, as follows:

$$\begin{aligned} z_i &= 1 \quad \text{if the } i\text{th ordered observation is an } X \\ &= 0 \quad \text{otherwise.} \end{aligned} \tag{2.1}$$

Also let

$$\begin{aligned} \delta_x^*(i) &= i \quad \text{if the } (r+1-i)\text{th ordered observation is an } X \\ &= 0 \quad \text{otherwise,} \\ \delta_y^*(i) &= i \quad \text{if the } (r+1-i)\text{th ordered observation is a } Y \\ &= 0 \quad \text{otherwise } (i = 1, 2, \cdots, r). \end{aligned} \tag{2.2}$$

Clearly

$$\delta_x^*(r+1-i) = (r+1-i)z_i\,; \quad \delta_y^*(r+1-i) = (r+1-i)(1-z_i) \tag{2.3}$$

and

$$\delta_x^*(i) + \delta_y^*(i) = i \qquad (i = 1, 2, \cdots, r). \tag{2.4}$$

We now prove

LEMMA 2.1. *For any* $r \geqq 1$,

$$(2.5)\quad \text{(a)} \qquad \sum_{i=1}^{r} m_i = \sum_{i=1}^{r} \delta_x^*(i),$$
$$\text{(b)} \qquad \sum_{i=1}^{r} n_i = \sum_{i=1}^{r} \delta_y^*(i).$$

PROOF. Note that if the first observation is an x, it will contribute unity to each of m_i's ($i = 1, 2, \cdots, r$), whereas it contributes r to $\delta_x^*(r)$ only. In general. if the $(r + 1 - i)$th ordered observation is an x it contributes 1 to each of the last i terms $m_{r-i+1}, m_{r-i+2}, \cdots, m_r$ on the left side of (2.5a), whereas it contributes i to the right side of (2.5a). On the other hand, if the $(r + 1 - i)$th observation is a y it contributes nothing to either side of (2.5a), ($i = 1, 2, \cdots, r$). This proves (2.5a) and equation (2.5b) is proved similarly.

Let us define the statistic $T_r^{(N)}$ by

$$(2.6)\qquad T_r^{(N)} = \sum_{i=1}^{r} ((i - r - 1)/N)z_i + m(r + 1)^2/2N^2 = \sum_{i=1}^{N} e_i z_i,$$

where

$$(2.7)\qquad e_i = (i - r - 1)/N + (r + 1)^2/2N^2 \quad \text{if } 1 \leqq i \leqq r$$
$$= (r + 1)^2/2N^2 \quad \text{if } r < i \leqq N.$$

The equivalence of $V_r^{(N)}$ and $T_r^{(N)}$ is shown in the following:

THEOREM 2.1. *For testing* $H_0 : F = G$ *against one-sided (or two-sided) alternatives* $H_1 : F \neq G$, *the statistics* $V_r^{(N)}$ *and* $T_r^{(N)}$ *are equivalent.*

PROOF. Using (1.3), (2.4) and (2.5)

$$V_r^{(N)} = n \sum_{i=1}^{r} m_i - m \sum_{i=1}^{r} n_i$$
$$= n \sum_{i=1}^{r} \delta_x^*(i) - m \sum_{i=1}^{r} \delta_y^*(i) \qquad \text{(by Lemma 2.1)}$$
$$(2.8)\qquad = N \sum_{i=1}^{r} \delta_x^*(i) - mr(r + 1)/2$$
$$= N \sum_{i=1}^{r} (r + 1 - i)z_i - mr(r + 1)/2$$
$$= m(r + 1)/2 - N^2 T_r^{(N)}.$$

The relationship of $T_r^{(N)}$ (and hence that of $V_r^{(N)}$ with the Wilcoxon statistic W and the Mann-Whitney statistic U becomes clear by putting $r = N$ in (2.6), (that is, when the complete combined sample is available).

$$NT_N^{(N)} = \sum_{i=1}^{N} iz_i - m(N + 1) + m(N + 1)^2/2N$$
$$(2.9)\qquad = W - m(N^2 - 1)/2N$$
$$= U + m(m + 1)/2 - m(N^2 - 1)/2N,$$

where $W = \sum_{i=1}^{N} iz_i$ is the Wilcoxon statistic [22] and U is the Mann-Whitney statistic [13] defined for a sequence of m x's and n y's as the number of y's preceding each x_i, summed from $i = 1$ to m.

3. Asymptotic normality of $T_r^{(N)}$. The asymptotic normality of $T_r^{(N)}$ can be derived from a theorem (and its corollaries) of Hájek [11]. For details see Basu [1]. Referring to Hájek's notations, except for the fact that we interchange the use of the symbols N and ν, we show how his theorem applies in our case. In the present case $(X_{N1}, X_{N2}, \cdots, X_{N\nu_N})$ denotes the sequence consisting of $m_N(=m)$ X's and $n_N(=n)$ Y's and $\nu_N = N$. (We have attached the subscript N to indicate the dependence of m, n and r on N.) Let

$$\Delta_N = \nu_N^{-\frac{1}{2}} \tag{3.1}$$

that is, $C_{Ni} = (1 - z_i)/\nu_N$, $(i = 1, 2, \cdots, \nu_N)$ and assume that

$$\lim_{N\to\infty} m_N/\nu_N = \lim_{N\to\infty} m/N = \lambda. \tag{3.2}$$

From (3.1) and (3.2) it is easily verified that assumptions (3.2) and (3.3) of Hájek are satisfied. We now define $\varphi_N(u)$ explicitly by asserting that it is a step function which is constant on the intervals $(i-1)/\nu_N < u \leqq i/\nu_N$ and defined at the intermediate points $i/(\nu_N + 1)$ for $i = 1, 2, \cdots, \nu_N$ by

$$\begin{aligned} \varphi_N(i/(\nu_N+1)) &= (i - r_N - 1)/\nu_N + (r_N+1)^2/2\nu_N^2 \quad &\text{for} \quad i \leqq r_N \\ &= (r_N+1)^2/2\nu_N^2 \quad &\text{for} \quad i > r_N. \end{aligned} \tag{3.3}$$

From (3.1) and (3.3) we see that the rank order statistic S_N is related to $T_r^{(N)}$ by the following equation:

$$\begin{aligned} S_N &= \textstyle\sum_{i=1}^{N} (C_{Ni} - \bar{C}_N)\varphi_N(R_{Ni}/(\nu_N+1)), \\ &= m_N(r_N+1)/2\nu_N^{\frac{5}{2}} - \nu_N^{-\frac{1}{2}} T_r^{(N)}. \end{aligned} \tag{3.4}$$

By Corollary 2.2 of Hájek [11] the statistic S_N, and therefore $T_r^{(N)}$ is asymptotically normally distributed. For we can take the function $\varphi(u)$ to be

$$\begin{aligned} \varphi(u) &= u - p + p^2/2, \qquad &0 < u \leqq p, \\ &= p^2/2, &u > p, \end{aligned} \tag{3.5}$$

where

$$p = \lim_{N\to\infty} r_N/\nu_N = \lim_{N\to\infty} r/N$$

and we can find a distribution function $H(x)$ with density function given by

$$\begin{aligned} h(x) &= 2C^2 e^{-C(x+K)}/[1 + e^{-C(x+K)}]^2, \qquad &-\infty \leqq x \leqq X_0, \\ &= (p^2/2)e^{-xp^2/2}, &x > X_0; \end{aligned} \tag{3.6}$$

so that

$$\varphi(u) = -[h'(H^{-1}(u))/h(H^{-1}(u))]; \qquad 0 < u < 1. \tag{3.7}$$

Here $X_0 = H^{-1}(p)$ is the pth fractile of H.

It should be noted that the asymptotic ($m, N \to \infty$ with $m/N \to \lambda$) normality of $T_r^{(N)}$ also follows from the Chernoff-Savage theorem [2] using the relaxed sufficiency conditions given in [10]. Here, however, we assume that there exists a $\lambda_0 \leqq \frac{1}{2}$ such that $0 < \lambda_0 \leqq \lambda \leqq 1 - \lambda_0 < 1$ and $\lambda_N = m/N$. Thus

$$(3.8) \qquad \lim_{N\to\infty} P\{(T_r^{(N)}/m - \mu_N)/\sigma_N < t\} = (2\pi)^{-\frac{1}{2}} \int_{-\infty}^{t} e^{-u^2/2}\, du$$

where

$$(3.9) \qquad \mu_N = \int_{-\infty}^{\infty} \varphi[H(x)]\, \mathrm{dF(x)} = \int_{-\infty}^{\infty} \varphi[\lambda F(x) + (1-\lambda)G(x)]\, dF(x)$$

and assuming that, $G(x) = F(x - \theta_N)$ and $\theta_N \to 0$ as $N \to \infty$,

$$(3.10) \qquad \lim_{N\to\infty} N\lambda_N \sigma_N^2/(1 - \lambda_N) = \tfrac{1}{12}p^3(4 - 3p).$$

4. Consistency of the $T_r^{(N)}$ test. Consistency of the $T_r^{(N)}$ test of $H_0 : F(x) = G(x)$ against one-sided alternatives $H_1 : G(x) = F(x - \theta)$ for $\theta > 0$, say is shown by using the Chernoff-Savage theorem [2]. When the null hypothesis is not true, i.e., $\theta > 0$ it follows easily from (3.5), since $\lambda < 1$, that

$$(4.1) \qquad \varphi(F(x)) - \varphi(H(x)) = (1 - \lambda)(F - G) \neq 0,$$

for all x in the interval $-\infty < x \leqq \min\,[H^{-1}(p), F^{-1}(p)]$. It is also seen that

$$(4.2) \qquad \int_{-\infty}^{\infty} [\varphi(F(x)) - \varphi(H(x))] dF(x) \neq 0$$

by writing the integral in (4.2) as the sum of three separate integrals formed by the intervals with endpoints $\pm\infty$, $F^{-1}(p)$ and $H^{-1}(p)$; in fact if $\theta > 0$ the result is negative and if $\theta < 0$ the result is positive. Let $\sigma_N^2(\theta)$ denote the variance of $T_r^{(N)}$ under the alternative hypothesis $G(x) = \mathrm{F}(x - \theta)$. As $N \to \infty$ and $\theta_N \to \theta_0 = 0$ it follows from (3.10) that $\sigma_N^2(\theta_0) \to 0$ and $\sigma_N^2(\theta_N) \to 0$. Using Chebyshev's inequality we find that the $T_r^{(N)}$ test, and consequently the equivalent $V_r^{(N)}$ test, are consistent against one-sided alternatives of the form $G(x) = F(x - \theta)$.

The consistency of the $T_r^{(N)}$ test for two-sided alternatives can be proved in a similar manner.

5. Efficacy of the $T_r^{(N)}$ test. For comparing the large sample power of two sequences of tests, the concept of asymptotic relative efficiency (ARE) was developed by Pitman [15]. An exposition of his work with extensions is given by Noether [14], (see also Fraser [7]).

Considering that under the alternative hypothesis F and G differ only by a shift of location (or scale) parameter, i.e., assuming that $G(x) = F(x - \theta)$ for $\theta \neq 0$ (or $G(x) = F(\theta x)$ for $\theta \neq 1$), Pitman shows that the ARE $\mathrm{e}(T^*, T)$ of the test T^* with respect to the test T can be evaluated both for the one-sided and two-sided case by the formula

$$(5.1) \qquad e(T^*, T) = \lim_{N\to\infty} e(T_N^*)/e(T_N);$$

where $E(T_N)$ denotes the expectation of T_N under H_1 and $\sigma_0^2(T_N)$ denotes the variance of T_N under the null hypothesis and the efficacy $e(T_N)$ of a test T_N based on N observations is defined by

$$(5.2) \qquad e(T_N) = [dE(T_N)/d\theta\, |_{\theta=\theta_0}]^2/\sigma_0^2(T_N).$$

Thus any two tests can be compared if their efficacies (or limiting efficacies) are known. In this section we shall calculate the efficacy of the $T_r^{(N)}$ test.

For the statistic $T_r^{(N)}$ we have $ET_N = m\mu_N$ where μ_N is given by (3.9); hence for the location parameter problem

$$\lim_{N\to\infty} d\mu_N/d\theta = (1 - \lambda) \int_{-\infty}^{\infty} \varphi'[H(x)]f(x - \theta)\, dF(x) \tag{5.3}$$

where $\varphi[\mathrm{H}(x)]$ is given by (3.5). Using (3.10) and (5.3), the efficacy of $T_r^{(N)}$ is given by

$$e(T_r^{(N)}) = [12N\lambda(1 - \lambda)/p^3(4 - 3p)]\{\int_{-\infty}^{F^{-1}(p)} f^2(x)\, dx\}^2. \tag{5.4}$$

For $p = 1$, the above reduces to

$$12N\lambda(1 - \lambda)[\int_{-\infty}^{\infty} f^2(x)\, dx]^2, \tag{5.5}$$

which is the known value of the efficacy of the Wilcoxon statistic.

Similarly, for scalar alternatives if we assume $G(x) = F(\theta x)$, we have

$$[d\mu_N/d\theta|_{\theta=1}]^2 = (1 - \lambda)^2\{\int_{-\infty}^{F^{-1}(p)} xf^2(x)\, dx\}^2, \tag{5.6}$$

so that the efficacy of the $T_r^{(N)}$ statistic is given by

$$[12N\lambda(1 - \lambda)/p^3(4 - 3p)]\{\int_{-\infty}^{F^{-1}(p)} xf^2(x)\, dx\}^2. \tag{5.7}$$

6. ARE of $T_r^{(N)}$ with respect to the F-test for the scale parameter in the case of the exponential distribution. The $T_r^{(N)}$ test can also be used for testing the $H_0 : F = G$ against scalar alternatives $H_0 : G(x) = F(\theta x)$ where $\theta > 0$. This is possible for positive random variables since the scale parameter becomes the location parameter under logarithmic transformation and the rank tests remain invariant under any strictly increasing transformation of the original variables.

In this section we shall compute the ARE of the $T_r^{(N)}$ test with respect to the F-test (which is equivalent to the likelihood ratio test) when the underlying parent populations are exponential. The exponential distribution is the most widely used model in problems of life-testing. (See Epstein and Sobel [4], [5]). Considering only positive random variables we let

$$\mathrm{F}(x) = 1 - e^{-x} \quad \text{for} \quad x > 0 \tag{6.1}$$

and

$$\mathrm{G}(y) = 1 - e^{-\theta y} \quad \text{for} \quad y > 0 \quad \text{where} \quad \theta > 0. \tag{6.2}$$

The likelihood-ratio test of $\mathrm{H}_0 : \theta = \theta_0 = 1$ against alternative $\theta < 1$ is based on the first r ordered observations from a combined sample of $m + n$ observations, (m are x's and n are y's), where $m_r \leqq m$ observations are from (6.1) and $n_r \leqq n$ observations are from (6.2), so that $m_r + n_r = r$. This test reduces to the F test conditioned on m_r and n_r [6] for which the test statistic given by

$$R = m_r\{\textstyle\sum_{i=1}^{n_r} y_i + (n - n_r)y_{n_r}\}/n_r\{\sum_{i=1}^{m_r} x_i + (m - m_r)x_{m_r}\} \tag{6.3}$$

follows the F-distribution with $(2n_r, 2m_r)$ degrees of freedom under H_0 and that $R\theta$ has the same distribution under the alternative hypothesis $G(x) = F(\theta x)$.

However, for deriving the expression for the efficacy of R we need the unconditional asymptotic distribution of R. To this end, note that under H_0 (or under alternative hypothesis $\theta = \theta_N$ where $\theta_N \to \theta_0 = 1$) as $N \to \infty$ the ratio m_r/N has the binomial distribution with expectation rm/N^2 which tends to λp and has variance rmn/N^4 which tends to zero. Also $\sum_{i=1}^{n_r} y_i + (n - n_r)y_{n_r}$ and $\sum_{i=1}^{m_r} x_i + (m - m_r)x_{m_r}$ are independent and are sums of m_r and n_r independent variables when m_r and n_r are given. Using the results of Rényi [17] one concludes that $N^{\frac{1}{2}}(R - 1/\theta)$ tends in law to a normal variable with mean zero and variance $1/\theta^2 p\lambda(1 - \lambda)$ unconditionally. Thus the efficacy of R is given by

$$e(R) = N\lambda(1 - \lambda)p. \tag{6.4}$$

In the exponential case we also have

$$F^{-1}(p) = -\log(1 - p). \tag{6.5}$$

Hence using (5.7) the efficacy of the $T_r^{(N)}$ statistic is given by

$$e(T_r^{(N)}) = [3N\lambda(1 - \lambda)/4p^3(4 - 3p)]\{2(1 - p)^2 \log(1 - p) + 2p - p^2\}^2. \tag{6.6}$$

From (5.1), (6.4) and (6.6) the ARE of the $T_r^{(N)}$ test with respect to the R-test is given by

$$e(T_r^{(N)}, R) = [3/4p^4(4 - 3p)]\{2(1 - p)^2 \log(1 - p) + 2p - p^2\}^2. \tag{6.7}$$

In the special case $p = 1$, we obtain from (6.7)

$$e(T_r^{(N)}, R) = .75. \tag{6.8}$$

7. Comparison with other ampr tests from censored data. In a recent paper Gastwirth [8] has considered several rank tests based on well-known statistics for the case of censored data and he has derived the weight functions $J(u)$ for which these tests are asymptotically most powerful rank tests (amprt). His weight functions for the modified Wilcoxon test is

$$J(u) = u - \tfrac{1}{2}, \quad 0 \leqq u \leqq p, \qquad = p/2, \quad p < u \leqq 1, \tag{7.1}$$

where p has the same meaning as before; we denote the corresponding statistic for this test by $G_r^{(N)}$. From the discussion of the asymptotic normality of $T_r^{(N)}$ and from Theorem 1.1 of Hájek [11] it is clear that each of the statistics based on weight functions $\varphi(u)$ in (3.5) and $J(u)$ in (7.1) is an amprt with respect to a certain family of distributions. In each case the family has the form of a logistic distribution to the left of the censored percentile and has the exponential form to the right of this point. However, the two families differ in functional form, the first one being given by (3.6) and the second is given by letting $X_0 = H^{-1}(p)$,

$$h(x) = \tfrac{1}{2}e^{-\frac{1}{2}(x+k)}/(1 + e^{-\frac{1}{2}(x+k)})^2, \quad -\infty < x \leqq X_0, \qquad = \tfrac{1}{2}pe^{-px/2}, \quad x > X_0, \tag{7.2}$$

where k is a function of p.

It is of interest to find the ARE of one test when the other is the amprt. We can take this in either direction since the result is symmetric. From Section 6 of Hájek [11] it is known that the ARE of statistic T_1 with respect to T_2 is given by

$$e(T_1, T_2) = \rho^2(T_1, T_2) \tag{7.3}$$

where T_2 corresponds to the ampr test for the underlying distribution T_1 corresponds to any other rank test and $\rho(T_1, T_2)$ is the correlation coefficient between T_1 and T_2 and in our case is given by

$$\begin{aligned} e(T_r^{(N)}, G_r^{(N)}) &= \rho^2(T_r^{(N)}, G_r^{(N)} \mid p) \\ &= [\textstyle\int_0^1 \varphi(u)J(u)\,du]^2 / \int_0^1 \varphi^2(u)\,du \int_0^1 J^2(u)\,du \\ &= (3-2p)^2/(4-3p)(3-3p+p^2). \end{aligned} \tag{7.4}$$

It can be easily seen that $\rho^2(T_r^{(N)}, G_r^{(N)} \mid p)$ is an increasing function of p so that $\rho^2(T_r^{(N)}, G_r^{(N)} \mid p) \geqq \rho^2(T_r^{(N)}, G_r^{(N)} \mid 0) = 0.75$. The ARE $e(T_r^{(N)}, G_r^{(N)})$ has been computed for several values of p in Table I. In particular for $p = \frac{1}{2}$, $\rho^2(T^{(N)}, G^{(N)} \mid p) = .91$. We see from Table I that for all p the performance of $G_r^{(N)}$ and $T_r^{(N)}$ are roughly comparable.

A second statistic worth comparing with the $T_r^{(N)}$ statistic is a form of the statistic $S_N^{(N)}$ proposed by Savage [18] which is the amprt in the exponential as well as in the Weibull case. For a fair comparison we shall consider the modified test $S_r^{(N)}$ based on censored data as given by Gastwirth [8] with weight function

$$\begin{aligned} k(u) &= -\ln(1-u) - 1, && 0 \leqq u \leqq p, \\ &= -\ln(1-p), && p < u \leqq 1. \end{aligned} \tag{7.5}$$

The ARE of $T_r^{(N)}$ with respect to $S_r^{(N)}$ when the underlying population is exponential from zero to the point of censoring and again exponential to the right of the point of censoring is given by

$$\begin{aligned} e(T_r^{(N)}, S_r^{(N)}) &= \rho^2(T_r^{(N)}, S_r^{(N)}) \\ &= \{\textstyle\int_0^1 \varphi(u)k(u)\,du\}^2 / \int_0^1 \varphi^2(u)\,du \int_0^1 k^2(u)\,du. \end{aligned} \tag{7.6}$$

TABLE I

ARE *of different statistics with respect to Sobel statistic* $T_r^{(N)}$ *in the case of the exponential distribution for different values of* p

p	$e(T_r^{(N)}, R) = e(T_r^{(N)}, S_r^{(N)})$	$e(T_r^{(N)}, G^{(N)})$
.1	.7550	.78
.2	.7627	.81
.4	.7753	.88
.5	.7812	.91
.6	.7864	.94
.75	.7911	.9796
.8	.7908	.9879
.9	.7835	.9979
1.0	.7500	1.0000

Now, using (3.10) and (7.5) $\int_0^1 \varphi^2(u)\,du = \frac{1}{12}p^3(4-3p)$, $\int_0^1 k^2(u)\,du = p$ and

$$\int_0^1 \varphi(u)k(u)\,du = \tfrac{1}{4}[2(1-p)^2 \ln(1-p) + 2p - p^2].$$

Hence

$$e(T_r^{(N)}, S_r^{(N)}) = 3[2(1-p)^2 \ln(1-p) + 2p - p^2]^2/4p^4(4-3p)$$

which is exactly the same expression (6.7) we get for the ARE of $T_r^{(N)}$ with respect to the likelihood ratio test. If $p \to 1$ the above ARE $\to .75$ implying a correlation coefficient of $(.75)^{\frac{1}{2}} = .8660$ which agrees with Savage's result [18].

In Table I we have computed the ARE's of $T_r^{(N)}$ statistic with respect to the $G_r^{(N)}$ and $S_r^{(N)}$ or R statistics for different values of p when the latter statistics are optimal.

Rao, Savage and Sobel [16] have proposed another statistic $R_r^{(N)}$ for the case of censored data which is locally most powerful for the Lehmann family of alternatives (i.e., alternatives of the form $1 - G = (1-F)^\theta$) and is the same as $T_N^{(N)}$ when the complete sample is available. We shall not compare $T_r^{(N)}$ with $R_r^{(N)}$ since asymptotic properties of $R_r^{(N)}$ are not known.

8. A modified Sobel statistic. Recently Sobel [21] has proposed a new statistic V_r (which is a modification of the original $V_r^{(N)}$ statistic) because of some practical considerations. In our notation, the modified statistic V_r may be defined as

$$V_r = V_r^{(N)} + \tfrac{1}{2}(N - r - 1)(nm_r - mn_r). \tag{8.1}$$

The object of this section is to point out that the above statistic is asymptotically equivalent to the Gastwirth modification $G_r^{(N)}$ of the Wilcoxon statistic. This follows from the following:

THEOREM 8.1. *V_r is asymptotically equivalent to the statistic $G_r^{(N)}$.*

PROOF. Define the statistic G_r by,

$$G_r = \sum_{i=1}^N l_i z_i$$

where

$$\begin{aligned} l_i &= (2i - N - 1)/2N, & 1 \leqq i \leqq r, \\ &= r/2N, & r+1 \leqq i \leqq N. \end{aligned} \tag{8.2}$$

Following the line of Theorem 2.1, we then have

$$\begin{aligned} -G_r &= \textstyle\sum_{i=1}^r ((N + r + 1 - 2i)/2N)z_i - mr/2N \\ &= V_r/N^2. \end{aligned} \tag{8.3}$$

From the definition of G_r it is quite clear that the corresponding weight function $\varphi(u)$ will be same as $J(u)$ defined in (7.1).

From Section 7 and Theorem 8.1 it is clear that V_r possesses all the asymptotic properties which $V_r^{(N)}$ enjoys. That is, the two perform quite comparably in large samples. However, as Sobel has recently pointed out, from the exact sampling point of view V_r is preferable to $V_r^{(N)}$.

9. Conclusion. From the above discussions and from Table I we see that for testing for location alternatives the $V_r^{(N)}$ (as well as the V_r) test is quite satisfactory for most of the cases encountered. Even for testing for scalar alternatives (especially for distributions useful in life testing) the performance of the $V_r^{(N)}$ test is reasonably satisfactory; however Savage's statistic appears to be the most suitable one for the case of scalar alternatives. Properties of the Savage statistic are being studied further and the results will be communicated later. Several k-sample extensions of the $V_r^{(N)}$ statistic have been considered and will be reported later.

10. Acknowledgment. The author wishes to record his sincere appreciation to Professor Milton Sobel of The University of Minnesota under whose guidance this research was carried out. Thanks are also due to the referee and Professor Herman Chernoff of Stanford University for their helpful comments.

REFERENCES

[1] Basu, A. P. (1966). On some two-sample and K-sample rank tests with applications to life testing. Technical Report No. 77 Univ. of Minnesota.
[2] Chernoff, H. and Savage, I. R. (1958). Asymptotic normality and efficiency of certain non-parametric test statistics. *Ann. Math. Statist.* **29** 972–994.
[3] Cramér, H. (1946). *Mathematical Methods of Statistics.* Princeton Univ. Press.
[4] Epstein, B. and Sobel, M. (1953). Life testing. *J. Amer. Statist. Assoc.* **48** 486–502.
[5] Epstein, B. and Sobel, M. (1954). Some theorems relevant to life testing from an exponential distribution. *Ann. Math Statist.* **25** 373–381.
[6] Epstein, B. and Tsao, Chia Kuei (1953). Some tests based on ordered observations from two exponential populations. *Ann. Math. Statist.* **24** 458–466.
[7] Fraser, D. A. S. (1959). *Nonparametric Methods in Statistics.* Wiley, New York.
[8] Gastwirth, Joseph L. (1965). Asymptotically most powerful rank tests for the two sample problem with censored data. *Ann. Math. Statist.* **36** 1243–1247.
[9] Gehan, Edmund A. (1965). A generalized Wilcoxon test for comparing arbitrarily singly-censored samples. *Biometrika* **52** 203–223.
[10] Govindarajulu, Zakkula, Le Cam, L. and Raghavachari, M. (1966). Generalizations of Chernoff-Savage theorems on asymptotic normality of nonparametric statistics (preliminary report). *Ann. Math. Statist.* **37** 315.
[11] Hájek, Jaroslav (1962). Asymptotically most powerful rank-order tests. *Ann. Math. Statist.* **33** 1124–1147.
[12] Halperin, M. (1960). Extension of the Wilcoxon-Mann-Whitney test to samples censored at the same fixed point. *J Amer. Statist. Assoc.* **55** 125–138.
[13] Mann, H. B. and Whitney, D. R. (1947). On a test of whether one of the two random variables is stochastically greater than the other. *Ann. Math. Statist.* **18** 50–60.
[14] Noether, Gottfried E. (1955). On a theorem of Pitman. *Ann. Math. Statist.* **26** 64–68.
[15] Pitman, E. J. G. (1949). *Lecture notes on nonparametric statistical inference.* Columbia University.
[16] Rao, U. V. R., Savage, I. R. and Sobel M. (1960). Contributions to the theory of rank order statistics; the two-sample censored case. *Ann. Math. Statist.* **31** 415–426.
[17] Rényi, A. (1960). On the central limit theorem for the sum of a random number of independent random variables. *Acta Math. Acad. Sci. Hung.* **11** 97–102.

[18] SAVAGE, I. RICHARD (1956). Contributions to the theory of rank order statistics—the two-sample case. *Ann. Math. Statist.* **27** 590–615.

[19] SOBEL, MILTON (1957). On a generalized Wilcoxon statistic for life testing. Proc. Working conference on the theory of reliability (April 17–19, 1957), New York University and the Radio Corporation of America, pp. 8–13.

[20] SOBEL, M. (1965). On a generalization of Wilcoxon's rank sum test for censored data. Technical Report No. 69, University of Minnesota.

[21] Sobel, M. (1966). On a generalization of Wilcoxon's rank sum test for censored data. Technical Report No. 69 (Revised). University of Minnesota.

[22] Wilcoxon, F. (1945). Individual comparisons by rank methods. *Biometrics* **1** 80–83.