SMALL SAMPLE POWER AND EFFICIENCY FOR THE ONE SAMPLE WILCOXON AND NORMAL SCORES TESTS[1]

By Jerome Klotz[2]

University of California, Berkeley

0. Summary. Small sample power and efficiency are computed for the one sample Wilcoxon and normal scores tests for normal shift alternatives. A recursive scheme is given which reduces the problem of power computation permitting investigations up to sample size $N = 10$. Local efficiencies for the two nonparametric tests are computed for small samples using the values of the normal scores statistic. In addition, efficiencies for large shifts are obtained by comparing the exponential rate of convergence to zero of the type two error.

1. Introduction and notation. Let $X_1, \cdots, X_N$ denote a sample with cumulative distribution function F which has median μ. The one sample Wilcoxon test [7], and the normal scores test (Fraser 1957) for the hypothesis that F is symmetric about $\mu = 0$ against shift alternatives are based upon the respective statistics

$$W_+ = \sum_{k=1}^{N} kZ_{Nk} \quad \text{and} \quad S = \sum_{k=1}^{N} E_{Nk}Z_{Nk} \tag{1.1}$$

where $Z_{Ni} = 1(0)$ if the ith smallest observation in magnitude is non-negative (negative). E_{Ni} are numbers equal to the expected value of the ith smallest order statistic from a sample of N absolute normal (chi-one degree of freedom) variables. Large values of the statistics are significant for one sided shift alternatives ($\mu > 0$). If we call the vector $Z_N = (Z_{N1}, \cdots, Z_{NN})$ the ordering, power for the tests can be computed by summing $P[Z_N = z_N]$ over those ordering values z_N which lie in each test's rejection region.

2. Power calculations. In order to compute power for normal shift alternatives it is sufficient to compute expressions of the form

$$P[Z_N = z_N] = N! \int_{0<t_1<\cdots<t_N<\infty} \cdots \int \prod_{i=1}^{N} \varphi(t_i - s_i\,\mu)\, dt_i \tag{2.1}$$

with $s_i = 2z_{Ni} - 1$ the sign of the variable which is ith smallest in magnitude, and φ the standard normal density. Fortunately, the problem of integration over N dimensions can be reduced to the problem of evaluating N successive one dimensional indefinite integrals by adapting a two sample scheme of Hodges to the one sample problem (see for example [4], p. 502). If we denote $A_{z_N}(u) =$

Received August 17, 1962.

[1] This paper was prepared with the partial support of the Office of Ordnance Research, U.S.A. Grant DA-ARO (D)-31-124-G-83.

[2] Now at Harvard University.

$P[Z_N = z_N$ and $|X_i| \leqq u$ for all $i]$ then the desired probabilities are $A_{z_N}(\infty)$. For normal alternatives,

$$(2.2)\qquad A_{(1)}(u) = \Phi(u - \mu) - \Phi(-\mu), \qquad A_{(0)}(u) = \Phi(u + \mu) - \Phi(\mu)$$

where Φ is the standard normal cumulative. If we denote $(z_N, 1) = (z_{N1}, \cdots, z_{NN}, 1)$ and define $(z_N, 0)$ similarly then probabilities for z_{N+r} can be obtained inductively using

$$(2.3a)\qquad A_{(z_N,1)}(x) = (N + 1) \int_0^x A_{z_N}(u)\varphi(u - \mu)\, du$$

$$(2.3b)\qquad A_{(z_N,0)}(x) = (N + 1) \int_0^x A_{z_N}(u)\varphi(u + \mu)\, du.$$

The expression (2.3a) is obtained by conditioning with respect to the largest X_i in magnitude and noting that it is positive for $(z_N, 1)$ while it is negative for the ordering value $(z_N, 0)$ in (2.3b). When suitably modified, the formulas are applicable to densities other than the normal. The scheme was programmed for a digital computer (I.B.M. 704) with the probabilities for all orderings z_N, with $1 \leqq N \leqq 10$, computed at alternatives $\mu = 0(.25)1.5(.5)3.0$.

TABLE 1

Selected values of power and efficiency of the one sample Wilcoxon for normal shift alternatives

	N	α slope	$\mu = .25$	.50	.75	1.00	1.25	1.50	2.0	2.5	3	∞	d
Power	5	.06250 .23586	.1450	.2780	.4496	.6286	.7809	.8877	.9808	.9981	.9999	1.	2.
e_{wt}		.988	.986	.984	.982	.981	.980	.979	.979				.881
	6	.04688 .20689	.1248	.2630	.4503	.6464	.8071	.9114	.9887	.9992	1.0000	1.	2.
		.986	.983	.980	.977	.975	.973	.972	.970				.784
	7	.05469 .26104	.1552	.3322	.5557	.7603	.8974	.9655	.9981	1.0000		1.	3.
		.986	.983	.981	.978	.976	.975	.974	.974				.797
	8	.05469 .28277	.1672	.3689	.6139	.8180	.9356	.9831	.9995	1.0000		1.	3.42857
		.982	.980	.977	.975	.973	.972	.971					.728
	8	.07422 .36340	.2118	.4364	.6827	.8654	.9581	.9905	.9998	1.0000		1.	3.5
		.977	.975	.972	.970	.969	.969	.968					.714
	9	.02734 .16704	.1035	.2713	.5146	.7499	.9038	.9728	.9992	1.0000		1.	3.42857
		.985	.982	.980	.977	.975	.974	.973					.757
	9	.03711 .21803	.1323	.3256	.5812	.8038	.9329	.9835	.9996	1.0000		1.	3.5
		.983	.981	.978	.976	.975	.973	.972					.727
	9	.04883 .27494	.1637	.3791	.6398	.8455	.9525	.9897	.9998	1.0000		1.	3.55556
		.978	.976	.973	.971	.970	.969	.969					.697
	10	.00098 .00779	.00592	.0250	.0765	.1777	.3274	.5009	.7944	.9396	.9866	1.	1.
		.938	.929	.920	.908	.896	.882	.869	.844	.816	.794		.651
	10	.00977 .07048	.0480	.1576	.3617	.6124	.8195	.9367	.9967	1.0000		1.	3.
		.983	.980	.976	.973	.970	.967	.964	.959				.744
	10	.02441 .16119	.1021	.2822	.5437	.7847	.9276	.98294	.9997	1.0000		1.	3.55556
		.984	.981	.979	.976	.974	.972	.971					.706
	10	.05273 .31066	.1844	.4274	.7013	.8914	.9734	.9957	1.0000			1.	5.
		.970	.968	.967	.966	.965	.965	.964					.756
	10	.09668 .50504	.2862	.5669	.8153	.9476	.9904	.9989	1.0000			1.	5.
		.961	.960	.959	.957	.956	.955						.685

3. Power and efficiency of the Wilcoxon test. Given an alternative μ, the sample size N, and a non-randomized significance level $\alpha = k/2^N$ for the Wilcoxon test, power is obtained by adding $P[Z_N = z_N]$ over the k ordering values z_N which give significant values for W_+. All power values to accuracy believed to be 4 decimal places were computed for $\alpha \leqq .10$, $N = 1(1)10$, and $\mu = 0(.25)1.5(.5)3.0$. Selected values are given in Table 1.

Efficiency of the Wilcoxon relative to the one sample t-test was computed using the randomized definition of Hodges and Lehmann (1956). For fixed N, α, μ of the Wilcoxon test, power of the t-test was computed at the same values of α, μ for two sample sizes N', N'' $(= N' + 1)$ needed to bracket the Wilcoxon power. Interpolating linearly with the power gives $N^* = \lambda N' + (1 - \lambda)N''$ $(0 \leqq \lambda \leqq 1)$ and efficiency $e_{W,t}(N, \alpha, \mu) = N^*/N$. A computer program was used to obtain critical values and power values of the t at the non-standard levels $\alpha = k/2^N$. The efficiency values given in Table 1 are extremely high in the region of interest. For small α, the efficiency appears roughly to decrease with α which is consistent with the fact that the sign test (which has generally lower efficiency for normality) and the Wilcoxon test coincide for the smallest level $\alpha = 1/2^N$. The efficiency appears also to decrease with μ which is consistent with the results of Section 6.

4. The normal scores test. The number of different values of the normal scores statistic appears in general to be 2^N corresponding to the 2^N values of the ordering vector Z_N. The values are symmetrically placed about the expected value $N/(2\pi)^{\frac{1}{2}}$ on the closed interval $[0, N(2/\pi)^{\frac{1}{2}}]$. The symmetric pairs are obtained by interchanging 0's and 1's in each z_N. Power for the normal scores test was obtained in the same manner as for the Wilcoxon. It is necessary to compute the weights E_{Ni} to determine which z_N belong to the rejection region. A program was written to evaluate the weights E_{Ni} and to sort the ordering values z_N according to the values of the statistic S. The values E_{Ni} to 5 decimals are given

TABLE 2

Expected values of absolute normal order statistics

E_{Ni}	$i = 1$	2	3	4	5	6	7	8	9	10
$N = 1$	.79788									
2	.46739	1.12838								
3	.33490	.73236	1.32639							
4	.26208	.55336	.91136	1.46473						
5	.21569	.44764	.71195	1.04430	1.56983					
6	.18344	.37695	.58903	.83487	1.14902	1.65400				
7	.15967	.32605	.50420	.70212	.93444	1.23485	1.72385			
8	.14141	.28752	.44163	.60849	.79575	1.01765	1.30726	1.78337		
9	.12693	.25728	.39335	.53820	.69636	.87526	1.08884	1.36966	1.83508	
10	.11515	.23289	.35485	.48317	.62075	.77198	.94412	1.15086	1.42436	1.88071

in Table 2 for $N = 1(1)10$. A small table of percentage points of the null distribution along with the exact probabilities is given in Table 3.

The normal scores test has power properties identical with those of the Wilcoxon test for N and α sufficiently small. The rejection region for the normal scores test coincides with that of the Wilcoxon for $\alpha \leqq .10$ and $N \leqq 5$. For $N = 6, 7$, $\alpha \leqq .10$ it coincides at the natural α levels of the Wilcoxon (the normal scores having more α values). Finally, the normal scores test is equivalent with the Wilcoxon at natural levels of the Wilcoxon for $N = 8$ and $\alpha \leqq .05469$,

TABLE 3

Upper percentage points of the normal scores null distribution

s	nominal levels						
$P[S \geqq s]$	.001	.005	.010	.025	.050	.075	.100
$N = 4$					3.19153		2.92945
					.06250		.12500
5				3.98941	3.77372		3.54177
				.03125	.06250		.09375
6			4.78731	4.60387	4.41036	4.19828	4.01484
			.01563	.03125	.04688	.07813	.09375
7		5.58518	5.42551	5.25913	4.92131	4.65074	4.49107
		.00781	.01563	.02344	.04688	.07813	.10156
8		6.38308	6.09556	5.80004	5.44592	5.19155	4.97884
		.00391	.01172	.02344	.05078	.07422	.10156
9	7.18096	6.92368	6.78761	6.35767	5.96432	5.70704	5.51022
	.00195	.00586	.00977	.02539	.05078	.07422	.09961
10	7.97884	7.62399	7.35809	6.88809	6.50397	6.20723	5.98996
	.00098	.00488	.00977	.02539	.04980	.07520	.09961

TABLE 4

Selected values of power and efficiency of the one sample normal scores test for normal shift alternatives

	N	α	slope	$\mu =$ .25	.50	.75	1.00	1.25	1.50	2.0	2.5	3	∞	d
Power	8	.07422	.36430	.2122	.4375	.6839	.8662	.9583	.9905	.9998	1.0000		1.	3.42857
$e_{s,t}$		.981		.978	.975	.973	.971	.970	.968					.680
	9	.03711	.21815	.1323	.3256	.5811	.8033	.9323	.9831	.9996	1.0000		1.	3.42857
		.984		.980	.978	.975	.974	.971	.969					.719
	9	.04883	.27639	.1641	.3804	.6416	.8470	.9532	.9899	.9998	1.0000		1.	3.5
		.986		.980	.978	.976	.973	.971	.971					.691
	10	.02441	.16135	.1022	.2826	.5443	.7851	.9277	.98287	.9997	1.0000		1.	3.5
		.985		.983	.980	.977	.975	.972	.971					.701
	10	.05273	.31300	.1857	.4310	.7055	.8940	.9743	.9959	1.0000			1.	3.6
		.982		.979	.977	.975	.973	.971	.969					.623
	10	.09668	.50962	.2884	.5716	.8199	.9499	.9910	.9990	1.0000			1.	4.55556
		.977		.975	.972	.970	.968	.966						.641

See Table 1 for $N \leqq 7$ and $\alpha \leqq .10$, $N = 8$ and $\alpha \leqq .05469$, $N = 9$ and $\alpha \leqq .02734$, $N = 10$ and $\alpha \leqq .01855$.

$N = 9$ and $\alpha \leqq .02734$, and $N = 10$ and $\alpha \leqq .01855$. For larger α values the normal scores test is generally more powerful for smaller shift, with the Wilcoxon improving and sometimes becoming more powerful for larger shift in the region of interest. Selected power and efficiency values are given in Table 4.

5. Local efficiency. At the hypothesis $\mu = 0$ the power of the parametric and non-parametric tests is the same being equal to the level α. Hence interpolation with respect to the slopes of the power curves at $\mu = 0$ for the two one sided tests can be used to define local efficiency for small samples. As pointed out by Professor Lehmann,

$$(5.1) \qquad \frac{d}{d\mu} P[Z_N = z_N] \bigg|_{\mu=0} = \sum_{i=1}^{N} s_i \, E_{Ni}/2^N = [2S - n(2/\pi)^{\frac{1}{2}}]/2^N$$

where s_i is given in (2.1) and E_{Ni} and S are given in (1.1). Thus the derivative of the power for either non-parametric test at $\mu = 0$ can be obtained from

$$(5.2) \qquad \sum_{z_N \varepsilon R} [2S(z_N) - n(2/\pi)^{\frac{1}{2}}]/2^N$$

where R is the rejection region of values of Z_N. Thus summing the values of the normal scores statistic over the rejection regions for W_+ and for S gives the respective slopes of the two tests. The slope for the t-test using N observations is given by

$$(5.3) \qquad \left(\frac{N}{2\pi}\right)^{\frac{1}{2}} \left(1 + \frac{t_\alpha^2}{N-1}\right)^{-(N-1)/2}$$

where t_α is the upper α percentage point for the t statistic. Randomization gives efficiencies for $\mu = 0$ as given in Tables 1 and 4. The limit of the local efficiency ($N \to \infty$) gives the Pitman values ($3/\pi$ for W_+ and 1 for S). The local Wilcoxon efficiency appears, roughly, to decrease towards its limiting Pitman value.

It is interesting to note that similar methods were used and similar results obtained by H. R. van der Vaart (1950) for the two sample Wilcoxon test. In addition, the two sample results of Witting might also be noted for comparison.

6. Far distant efficiency. To obtain the other end of the efficiency curve ($\mu = \infty$) for the non-parametric tests relative to the t-test, a comparison of the rates of convergence to zero of the type two error (β) can be used. Following the methods of Hodges and Lehmann [3] in the two sample problem we note for normal shift alternatives that the type two error of the t-test converges to zero at the rate $\exp(-d_t\mu^2/2)$ for large μ neglecting terms of smaller order (see for example Nicholson, 1954). The constant d_t is given by

$$(6.1) \qquad d_t = N/[1 + t_\alpha^2/(N-1)].$$

The following theorem which is a modification to the one sample case of a theorem by Hodges and Lehmann [3], permits comparison of exponential rates of convergence for the non-parametric tests.

THEOREM. *The probability of the ordering vector z_N is given, up to terms of smaller order, by* $\exp[-d(z_N)\mu^2/2]$ *for large μ where X_i are independent normal with mean μ and variance* 1. *The constant d, which depends upon z_N, is determined in the following way:*

(i) *from the vector z_N, construct a walk in the x, y plane by taking a step in the y direction for each* 1 *and a step in the x direction for each* 0*-proceeding in sequence (see Figure* 1);

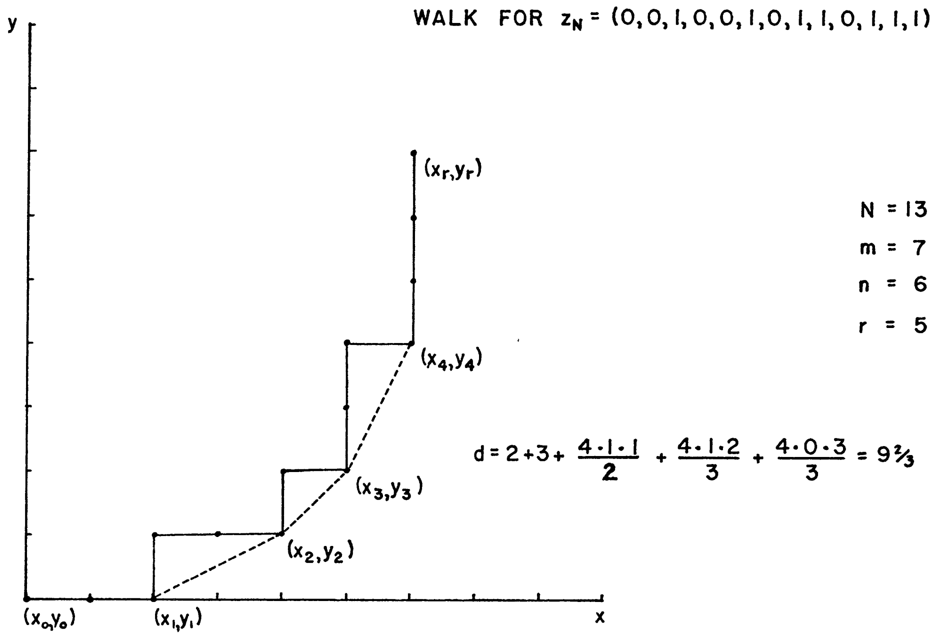

FIG. 1.

(ii) *determine $r + 1$ corner points (x_i, y_i) $i = 0, 1, 2, \cdots, r$ as the smallest set of points which determines the lower convex hull of the graph.* $(x_0, y_0) = (0, 0)$ *and* $(x_r, y_r) = (n, m)$ *where m is the number of* 1's *in z_N*;

(iii) *define*

$$d_i = \Delta x_i + \Delta y_i \text{ if } \Delta y_i < \Delta x_i,$$
$$= 4(\Delta x_i)(\Delta y_i)/(\Delta x_i + \Delta y_i) \text{ if } \Delta y_i \geqq \Delta x_i,$$

with $\Delta x_i = x_i - x_{i-1}$, $\Delta y_i = y_i - y_{i-1}$ *for* $i = 1, 2, \cdots, r$;

(iv) $d(z_N) = \sum_{i=1}^{r} d_i$.

PROOF. The proof uses transformations similar to those given in [3]. The transformations are of two types according as $\Delta y_i/\Delta x_i < 1$ or $\geqq 1$. Considering the expression (2.1) we transform the t_i variables to a collection of v and w variables by breaking the walk into two regions. The first $N_a = x_a + y_a$ variables determine the first region where $0 \leqq \Delta y_i/\Delta x_i < 1$ for $i = 1, 2, \cdots, a$ (where a is the largest such value). The transformation for this collection is

$$\begin{aligned} t_1 &= v_0 \\ t_2 &= t_1 + w_{0,1}/\mu \\ &\cdots \\ t_{N_a} &= t_{N_a-1} + w_{0,N_a-1}/\mu \quad \text{(in Figure 1, } N_a = 5,\ a = 2). \end{aligned} \tag{6.2}$$

If there are no corner points for which $\Delta y_i/\Delta x_i < 1$ we omit this transformation ($N_a = 0$). We next determine the remaining $N - N_a$ variables for the region $1 \leqq \Delta y_i/\Delta x_i \leqq \infty$ $i = a + 1, \cdots, r$ by the transformation

$$\begin{aligned} t_{N_a+1} &= v_1 + \left(\frac{\Delta y_{a+1} - \Delta x_{a+1}}{\Delta N_{a+1}}\right) \\ t_{N_a+2} &= t_{N_a+1} + w_{1,1}/\mu \\ &\cdots \\ t_{N_{a+1}} &= t_{N_{a+1}-1} + w_{1,\Delta N_{a+1}-1}/\mu \\ t_{N_{a+1}+1} &= v_2 + \left(\frac{\Delta y_{a+2} - \Delta x_{a+2}}{\Delta N_{a+2}}\right)\mu \\ t_{N_{a+2}} &= t_{N_{a+2}-1} + w_{2,\Delta N_{a+2}-1}/\mu \\ &\cdots \\ t_N &= t_{N-1} + w_{r-a,\Delta N_r-1}/\mu \end{aligned} \tag{6.3}$$

where $\Delta N_i = \Delta x_i + \Delta y_i$ and $N_r = N$. If there are no corner points satisfying the conditions on $\Delta y_i/\Delta x_i$ omit this transformation and take $a = r$.

The transformation which defines the $N = N_a + \sum_{k=1}^{r-a} \Delta N_{a+k}$ variables v_0, w_{0i} $i = 1, 2, \cdots N_a - 1$; v_k, w_{ki}, $i = 1, 2, \cdots \Delta N_{a+k} - 1$, $k = 1, 2, \cdots r - a$, is linear with Jacobian $|\partial(v, w)/\partial t| = \mu^{N-r+a-1}$. Without the constant factor $-\frac{1}{2}$ the exponent of the integrand of (2.1) equals $\sum_{i=1}^{N} (t_i - s_i\mu)^2$ which can be expressed in terms of the (v, w) variables using (6.2) and (6.3). After some algebra the expression is:

$$\begin{aligned} &N_a \mu^2 + v_0 \mu c + N_a v_0^2 + \sum_{j=1}^{N_a-1} w_{0j} c_{0j} + O(1/\mu) \\ &\quad + \sum_{k=1}^{r-a} \left[\left(\frac{4\Delta x_{a+k} \Delta y_{a+k}}{\Delta N_{a+k}}\right)\mu^2 + \Delta N_{a+k} v_k^2 + \sum_{j=1}^{\Delta N_{a+k}-1} w_{kj} c_{kj} + O(1/\mu)\right] \end{aligned} \tag{6.4}$$

where the constants are given by the expressions

$$c = 2(\Delta x_a - \Delta y_a), \qquad c_{0j} = -2\sum_{k=j+1}^{N_a} s_k$$

$$c_{kj} = 2\sum_{i=j+1}^{\Delta N_{a+k}} \left(\frac{\Delta y_{a+k} - \Delta x_{a+k}}{\Delta N_{a+k}} - s_{N_{a+k-1}+i}\right) \quad \text{for} \quad k = 1, 2, \cdots, r - a$$

and are positive because of the conditions on the two regions of the walk. $O(1/\mu)$

here denotes the terms in v, w which are multiplied by $1/\mu$ or $1/\mu^2$. Noting that we can express d as the sum of the constants multiplying μ^2 in (6.4) we write (2.1) in the form $\exp(-d\mu^2/2)g(\mu)$ where

$$(6.5) \qquad g(\mu) = N!\mu^{N-r+a-1} \int_{(v,w)\varepsilon R} \cdots \int e^{-\frac{1}{2}Q(v,w,\mu)} \prod_i dv_i \prod_{ij} dw_{ij},$$

$Q(v, w, \mu)$ is the expression given in 6.4 minus terms in μ^2, and R is the region of integration for the (v, w) variables determined from the region $0 < t_1 < \cdots < t_N < \infty$. Dropping the factor $\exp[-(c\mu v_0 + O(1/\mu))/2]$ and integrating over the larger region $R_1 = \{0 < v_0 < \cdots < v_{r-a}, w_{ij} > 0\}$ we get an upper bound (since terms in $1/\mu^2$ are positive) of the form $K_1\mu^{N-r+a-1}$ where K_1 is a positive constant independent of μ. Next if we replace R in (6.5) by $R_0 = R \cap \{0 < v_0 < 1\}$, and replace the factor $\exp(-cv_0\mu/2)$ by $\exp(-c\mu/2)$, and integrate, we get a lower bound for $g(\mu)$ of the form $K_0(\mu)\mu^{N-r+a-1}\exp(-c\mu)$ where $K_0(\mu) \to K_0$, $0 < K_0 < \infty$ as $\mu \to \infty$ using the dominated convergence theorem. Thus it follows that $\lim_{\mu\to\infty} (-2 \ln P[Z_N = z_N]/\mu^2) = d$ which completes the proof.

Applying the theorem, it follows for large μ, that the type two error for each non-parametric test is dominated by the probability of that z_N in the acceptance region with the smallest d value (probabilities for other d values being of smaller order). Randomization between sample sizes for the t-test (linear interpolation in d_t to match the minimum d of W_+ or of S) can then be used to define far distant efficiency ($e_{W,t}(N, \alpha, \infty)$ and $e_{S,t}(N, \alpha, \infty)$). Selected values are given in Tables 1 and 4 under $\mu = \infty$. The minimum d value in the Wilcoxon acceptance region is generally larger than that of the normal scores test for given (N, α) with a resulting greater far distant efficiency. This general statement appears to be true even for smaller μ in Tables 1 and 4 for a few isolated α values. For example for $N = 10$, $\alpha = 25/1024$ the power of the normal scores test falls below that of the Wilcoxon at $\mu = 1.5$. This may be explained by the fact that the normal scores test has one different ordering in its rejection region ($z_N = (1, 0, 0, 0, 1, 1, 1, 1, 1, 1)$ with $d(z_N) = 4$ for S and $z_N = (1, 1, 1, 1, 1, 1, 1, 0, 1, 1)$ with $d(z_N) = \frac{7}{2}$ for W_+). This gives a minimum d in the acceptance region of $\frac{7}{2}$ for S and $\frac{32}{9}$ (the next smallest d value after $\frac{7}{2}$) for the Wilcoxon.

7. Conclusion. Because of the extremely high efficiency of the non-parametric tests relative to the t in the region of interest, it is the author's opinion that the non-parametric tests would be preferred to the t in almost all practical situations. The exactness of the null distribution, good power for a wide class of shift alternatives, and the negligible loss in efficiency on the home ground of the t-test support this conclusion.

The normal scores test although most powerful locally and usually more powerful in the region covered becomes less powerful for large shift when compared to the Wilcoxon. In any case for the sample sizes covered the difference in power is somewhat academic. The greater number of significance levels must be balanced against the more difficult null distribution when comparing the normal scores test with the Wilcoxon.

8. Acknowledgments. I wish to thank Professor Hodges and Professor Lehmann for extremely helpful discussions and, in particular, for pointing out the local and far distant efficiency methods. In addition, I wish to thank the University of California Computer Center for providing machine time, and Mr. Raghavachari of the statistical laboratory for help in checking some of the numerical values.

REFERENCES

[1] Fraser, D. A. S. (1957). Most powerful rank type tests. *Ann. Math. Statist.* **28** 1040–1043.

[2] 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.

[3] Hodges, J. L., Jr. and Lehmann, E. L. (1962). Probabilities of rankings for two widely separated normal distributions. *Studies in Mathematical Analysis and Related Topics; Essays in Honor of George Pólya* 146–151. Stanford Univ. Press.

[4] Klotz, J. H. (1962). Nonparametric tests for scale. *Ann. Math. Statist.* **33** 498–512.

[5] Nicholson, W. L. (1954). A computing formula for the power in the analysis of variance test. *Ann. Math. Statist.* **25** 607–610.

[6] van der Vaart, H. R. (1950). Some remarks on the power function of Wilcoxon's test for the problem of two samples. *Nederl. Akad. Wetensch. Proc. Ser. A* **53** 494–520.

[7] Wilcoxon, R. (1945). Individual comparisons by ranking methods. *Biometrics Bull.* **1** 80–83.

[8] Witting, H. (1960). A generalized Pitman efficiency for nonparametric tests. *Ann. Math. Statist.* **31** 405–414.