ALTERNATIVE EFFICIENCIES FOR SIGNED RANK TESTS[1]

By Jerome Klotz[2]

Harvard University

0. Summary. Asymptotic efficiency curves for the one sample Wilcoxon and normal scores tests are obtained by comparing the exponential rate of convergence to zero of the type I error (α) while keeping the type II error (β) fixed ($0 < \beta < 1$). A wider than usual view of test performance consistent with small sample results is obtained. The Pitman efficiency value is derived as a special case when the alternative approaches the null hypothesis. Comparisons of the signed rank tests relative to $\bar{x}$ or t for normal location alternatives yield small efficiency values for extreme alternatives. The relative performance of the Wilcoxon is seen to be slightly better than the normal scores for normal alternatives with larger location parameter values despite the local (Pitman) optimality of the normal scores. Similar results hold for two other non-normal alternatives considered.

1. Notation and tail probabilities. Let $X_1, X_2, \cdots, X_n$ have continuous cdf F. For testing the hypothesis of symmetry about zero for F we consider the Wilcoxon [11] and normal scores [4] signed rank tests. The Wilcoxon statistic is equivalent to

$$(1.1)\quad W = \sum_{i=1}^{n} iU_i = 2\,(\text{number of positive } (X_i + X_j);\ 1 \leqq i \leqq j \leqq n) - n(n+1)/2$$

where U_i is the sign of the ith smallest observation when ordered by magnitude. Similarly, the normal scores statistic is equivalent to

$$(1.2)\quad \sum_{i=1}^{n} E_{ni}U_i$$

where the constants E_{ni} are expected values of the ith smallest order statistics from a sample of n absolute normal (chi — one degree of freedom) variables.

Under the null hypothesis of symmetry, $P[U_i = \pm 1] = \frac{1}{2}$ and the U_i are independent. Since (1.1) and (1.2) are sums of non-identical but independent random variables under H, we can obtain probabilities in the extreme tail of the null distribution by using Theorem 1 of Feller [3]. Specifically, we use the following particularization:

Theorem. *Let the numbers E_{ni} be expected values of the* ith *smallest order*

Received 29 March 1965.

[1] This research was started at San Diego State College under NSF grant G 12973 and continued with the support of the Office of Naval Research under contract Nonr 1866(37). Computations were performed at the Harvard Computing Center under NSF grant GP-683.

[2] Now at University of Wisconsin, Madison, Wisconsin.

statistics from a sample with cdf G *on* $(0, \infty)$ *with* $\int_0^\infty x^3\, dG(x) < \infty$. *If* U_i *are independent,* $P[U_i = \pm 1] = \frac{1}{2}$ *for* $i = 1, 2, \cdots, n$ *and* $S_n = \sum E_{ni} U_i$ *then for* $\rho_n \to \rho$ *where* $0 < \rho < \int_0^\infty x\, dG(x) < \infty$

$$\lim_n (-1/n) \ln P[S_n > \rho_n n] = e_s(\rho). \tag{1.3}$$

The constant $e_s(\rho) > 0$ *is evaluated from*

$$e_s(\rho) = h_\rho - \int_0^\infty \ln (\cosh (xh))\, dG(x) \tag{1.4}$$

and $h > 0$ *satisfies*

$$\int_0^\infty x \tanh (xh)\, dG(x) = \rho. \tag{1.5}$$

PROOF. Following the proof of Theorem 1 in Feller we particularize by considering sums of random variables taking only two values. Define, for $h > 0$ to be determined, independent random variables $U_i(h) = \pm 1$ with probability $\exp(\pm h E_{ni})/2 \cosh (h E_{ni})$ and let $S_n(h) = \sum E_{ni} U_i(h)$. Denote $\eta = ES_n(h)$, $\sigma^2 = \text{Var}(S_n(h))$. Then, we express the probability for S_n in terms of $S_n(h)$, as is done by Feller [3], p. 366,

$$P[S_n > \eta] = \int_\eta^\infty dF_{S_n}(t) = \prod_{i=1}^n \cosh (h E_{ni}) \int_\eta^\infty e^{-ht}\, dF_{S_n(h)}(t). \tag{1.6}$$

(1.6) can be shown to hold using induction on n. Making the substitution $Z = (S_n(h) - \eta)/\sigma$ for a normal approximation, (1.6) becomes

$$\prod_{i=1}^n \cosh (h E_{ni}) \int_0^\infty e^{-h(\eta + z\sigma)}\, dF_Z(z) \tag{1.6.1}$$

with

$$F_Z(z) = \Phi(z) + R_n(z). \tag{1.7}$$

The error term satisfies $|R_n(z)| \leqq c \sum E_{ni}^3/\sigma^3$ by the Berry-Essen theorem (see for example [5], p. 201). Applying a theorem of Hoeffding [8], we have

$$n^{-1} \sum_{i=1}^n E_{ni}^3 \to \int_0^\infty x^3\, dG(x) \quad (< \infty \text{ by assumption}), \tag{1.8}$$

$$n^{-1}\sigma^2 = n^{-1} \sum_i E_{ni}^2 \operatorname{sech}^2 (h E_{ni}) \to \int_0^\infty x^2 \operatorname{sech}^2 (hx)\, dG(x) < \infty.$$

Thus $|R_n(z)| = O(n^{-\frac{1}{2}})$. Substituting (1.7) in (1.6) with

$$\left|\int_0^\infty e^{-h\sigma z}\, dR_n(z)\right| = \left|-R_n(0) + h\sigma \int_0^\infty R_n(z) e^{-h\sigma z}\, dz\right| = O(n^{-\frac{1}{2}}),$$

integrating, and adjusting h so that $\rho_n = \eta/n$, we obtain

$$(-1/n) \ln P[S_n > \rho_n n] = (-1/n) \sum_i \ln(\cosh (h E_{ni})) + h\rho_n (-1/n) \ln [\exp ((h\sigma)^2/2)(1 - \Phi(h\sigma)) + O(n^{-\frac{1}{2}})]. \tag{1.9}$$

Using the Mills ratio approximation $(1 - \Phi(h\sigma))/\varphi(h\sigma) = (1 - \nu/(h\sigma)^2)/(h\sigma)$ where $0 < \nu < 1$ and $h\sigma = O(n^{\frac{1}{2}})$ by (1.5) and (1.8) we see that the last term on the r.h.s. of (1.9) goes to zero. Applying once more the theorem of Hoeffding to (1.9) we conclude (1.3) and (1.4) with (1.5) the limiting form of the equation $\rho_n = \eta/n$.

2. Exponent calculations. Denote the distributions specified under the null hypothesis by F_0 and under the alternative by F_μ. For a fixed alternative μ, we consider the problem of adjusting the critical values of the test statistics so as to obtain asymptotic type II error β where $0 < \beta < 1$. With the restriction $0 < \beta < 1$, we can use the normal approximation to obtain critical values as established by Govindarajulu [6]. Thus

$$\beta_n = P_\mu[(S_n - E_\mu S_n)/(\sigma_\mu(S_n)) < z_n] \to \beta \quad \text{when} \quad z_n \to z = \Phi^{-1}(\beta), \quad n \to \infty$$

and critical values are asymptotically given by

$$E_\mu S_n + z\sigma_\mu(S_n). \tag{2.1}$$

For the statistic $W/(n + 1)$, (2.1) becomes

$$[n(n - 1)/(n + 1)]P_\mu[X_1 + X_2 > 0] + [2n/(n + 1)]P_\mu[X_1 > 0] - \tfrac{1}{2}n + zO(n^{\frac{1}{2}}) = n\rho_n \sim n\rho$$

where

$$\rho = P_\mu[X_1 + X_2 > 0] - \tfrac{1}{2}. \tag{2.2}$$

Similarly for the normal scores ($n.s.$) statistic (1.2) the work of Govindarajulu ([6], p. 27) gives the corresponding

$$\begin{aligned}\rho &= \lim_n n^{-1}(E_\mu S_n + z\sigma_\mu(S_n)) \\ &= 2 \int_0^\infty \Phi^{-1}\{\tfrac{1}{2}[1 + F_\mu(x) - F_\mu(-x)]\} \, dF_\mu(x) - (2/\pi)^{\frac{1}{2}}.\end{aligned} \tag{2.3}$$

Applying the theorem to the Wilcoxon statistic with $S_n = W/(n + 1)$ we see that the conditions apply with $G(x) = x$, $(0 < x < 1)$. The type I error goes to zero at an exponential rate if $\rho > 0$ with

$$\begin{aligned}&\lim_n ((-1/n) \ln \alpha_n) \\ &\quad = \lim_n (-1/n) \ln P[\textstyle\sum U_i(i/(n + 1)) > n\rho_n] = e_W(\rho)\end{aligned} \tag{2.4}$$

where ρ and $e_W(\rho)$ are given by (2.2) and (1.4). Simplifying (1.5) for this G, we determine h by solving

$$\tfrac{1}{2} - (\pi^2/24h^2) + n^{-1} \ln (1 + e^{-2h}) + (1/2h^2) \textstyle\sum_{k=1}^{\infty} (-1)^{k+1}(e^{-2hk}/k^2) = \rho. \tag{2.5}$$

Similarly, the theorem applies directly to the normal scores statistic (1.2) with $G(x) = 2\Phi(x) - 1$ although less simplification occurs.

3. Efficiency values. For two sequences of tests we define a limiting efficiency at a fixed alternative by setting both asymptotic type II errors equal β, $0 < \beta < 1$ and adjusting the limiting ratio of sample sizes so that the type I errors will go to zero at the same rate. From (2.4) we have $-\ln \alpha_i = n_i e_i(\rho)(1 + o(1))$ for each test $(i = 1, 2)$. Equating $\ln \alpha_1 = -n_1 e_1(1 + o(1)) = -n_2 e_2(1 + o(1)) = \ln \alpha_2$ we obtain $\lim n_2/n_1 = e_1/e_2$ for the relative efficiency ($e_{1,2}$) of test 1 to test 2.

For normal alternatives, Bahadur [1] gives corresponding type I exponential rates of convergence to zero for the $\bar{X}$, sign, and t tests. If $F_\mu(x) = \Phi(x - \mu)$ then

$$e_{\bar{X}} = \mu^2/2 \tag{3.1}$$

$$e_S = p \ln 2p + q \ln 2q \quad \text{where} \quad p = 1 - q = \Phi(\mu) \tag{3.2}$$

$$e_t = (\ln (1 + \mu^2))/2. \tag{3.3}$$

Formulas (3.1, 3.2) can be derived using theorems for sums of independent identically distributed random variables as in [2] while (3.3) could be obtained with the aid of the tail formula of Pinkham and Wilk [10]. For purposes of comparison, Table I gives exponents for the Wilcoxon (W), normal scores ($n.s.$), sign (S), $\bar{X}$, and t at normal alternatives with variance 1 and means $\mu = 0(.125)3.000$. Under the assumption of normality, Expression (2.2) simplifies to give $\frac{1}{2} - \rho = 1 - \Phi(\sqrt{2}\mu)$ for the Wilcoxon. However, no simplification was found for the normal scores and (2.3), (1.5), and (1.4) were solved numerically.

TABLE I

Exponents for normal alternatives

μ	e_W	$e_{n.s.}$	e_S	$e_{\bar{X}}$	e_t
0	0	0	0	0	0
.125	.007416	.007752	.004956	.007813	.007752
.250	.02914	.03031	.01961	.03125	.03031
.375	.06368	.06577	.04336	.07031	.06579
.500	.1087	.1114	.07522	.1250	.1116
.625	.1615	.1642	.1139	.1953	.1649
.750	.2189	.2211	.1580	.2813	.2231
.875	.2779	.2793	.2058	.3828	.2843
1.000	.3360	.3365	.2557	.5000	.3466
1.125	.3910	.3908	.3062	.6328	.4089
1.250	.4416	.4409	.3558	.7813	.4705
1.375	.4867	.4859	.4034	.9453	.5307
1.500	.5262	.5254	.4478	1.1250	.5893
1.625	.5600	.5593	.4885	1.3203	.6461
1.750	.5884	.5880	.5250	1.5313	.7009
1.875	.6120	.6117	.5570	1.7578	.7538
2.000	.6311	.6309	.5846	2.0000	.8047
2.125	.6464	.6463	.6079	2.2578	.8538
2.250	.6585	.6483	.6272	2.5313	.9011
2.375	.6678	.6677	.6429	2.8203	.9466
2.500	.6749	.6747	.6554	3.1250	.9905
2.625	.6802	.6800	.6652	3.4453	1.0328
2.750	.6841	.6839	.6728	3.7813	1.0737
2.875	.6869	.6867	.6786	4.1328	1.1132
3.000	.6889	.6887	.6829	4.5000	1.1513
∞	ln 2	.6931	.6931	∞	∞

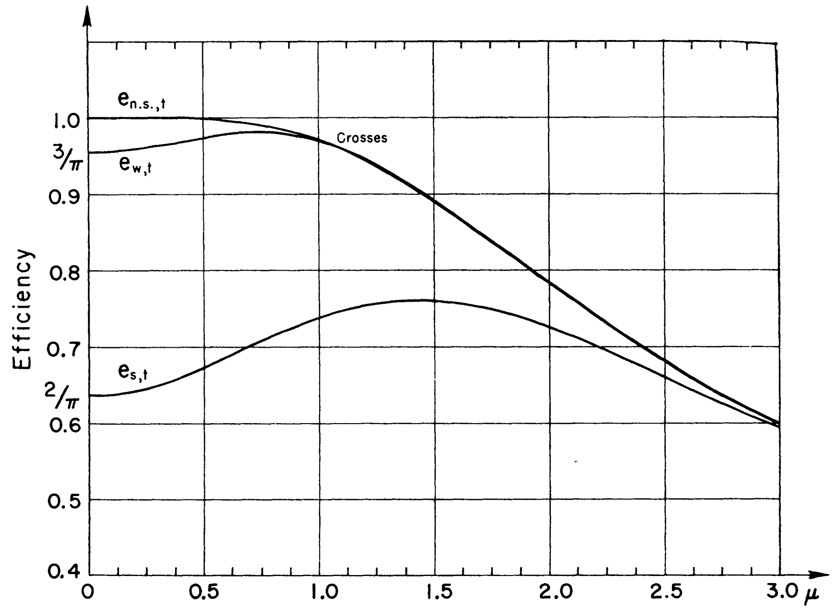

FIG. 1. Efficiencies for normal alternatives

Taking ratios of corresponding exponents gives efficiencies. For example, the efficiency of the Wilcoxon relative to the t for normal alternatives is

$$(3.4) \qquad e_{W,t} = [h\rho - \textstyle\int_0^1 \ln(\cosh xh)\,dx]/\frac{1}{2}(\ln(1+\mu^2))$$
$$= [2h\rho - \ln(\cosh(h))]/\tfrac{1}{2}(\ln(1+\mu^2))$$

where h, ρ satisfy (2.5) and (2.2). The limit of the above expression as μ goes to zero is the Pitman value $3/\pi \doteq .955$. For sufficiently large μ values, the rejection region consists of only the outcome with all signs positive which has null probability 2^{-n}. Thus, as can also be shown analytically, the numerator of (3.4) approaches ln 2 as μ becomes large and the ratio goes to zero. This points out the importance of the magnitudes of the observations for normal alternatives with large means. Figure 1 gives the normal efficiency curves relative to the t test for the Wilcoxon, normal scores, and sign tests. Although preferable locally, it is interesting that the normal scores efficiency falls very slightly below that of the Wilcoxon between $\mu = 1.000$ and $\mu = 1.125$. This behavior is consistent with small sample results obtained in [9] (see especially p. 631).

For non-normal comparisons the logistic and double exponential distributions with densities

$$(3.5) \qquad f(x) = e^{-x}/(1+e^{-x})^2 \qquad \text{(logistic)}$$
$$f(x) = \tfrac{1}{2}e^{-|x|} \qquad \text{(double exponential)}$$

TABLE II

Non-normal exponents

μ	Logistic alternatives			Double exponential alternatives		
	e_W	$e_{n.s.}$	e_S	e_W	$e_{n.s.}$	e_S
0	0	0	0	0	0	0
.125	.002598	.002481	.001949	.005772	.004916	.006919
.250	.01031	.009851	.007752	.02222	.01908	.02467
.375	.02290	.02190	.01727	.04742	.04113	.04972
.500	.03999	.03827	.03030	.07910	.06939	.07954
.625	.06109	.05853	.04654	.1151	.1022	.1123
.750	.08561	.08213	.06566	.1536	.1380	.1465
.875	.1129	.1085	.08726	.1931	.1754	.1813
1.000	.1424	.1370	.1109	.2324	.2133	.2158
1.125	.1733	.1670	.1363	.2707	.2508	.2496
1.250	.2050	.1979	.1628	.3074	.2872	.2823
1.375	.2371	.2293	.1902	.3420	.3222	.3136
1.500	.2689	.2606	.2181	.3744	.3551	.3434
1.625	.3001	.2914	.2461	.4044	.3862	.3715
1.750	.3304	.3215	.2738	.4321	.4149	.3979
1.875	.3593	.3504	.3012	.4575	.4417	.4226
2.000	.3869	.3781	.3278	.4806	.4660	.4456
2.125	.4129	.4043	.3536	.5017	.4886	.4670
2.250	.4373	.4290	.3784	.5208	.5088	.4868
2.375	.4600	.4522	.4021	.5382	.5277	.5050
2.500	.4810	.4737	.4246	.5539	.5443	.5219
2.625	.5005	.4938	.4459	.5681	.5598	.5374
2.750	.5184	.5123	.4659	.5809	.5733	.5516
2.875	.5349	.5293	.4847	.5924	.5861	.5647
3.000	.5500	.5450	.5023	.6028	.5970	.5766
∞	ln 2	.6931	.6931	.6931	.6931	.6931

were chosen for preliminary consideration because of convenience and optimal Pitman efficiency for the Wilcoxon and sign tests respectively. Using the location parameter family of alternatives $F(x - \mu)$ for $\mu = 0(.125)3$, we similarly compute Table II. Ratios of appropriate columns give relative efficiencies as in Figure 2. As before, limits of the ratios as μ goes to zero give Pitman values. On an overall basis the Wilcoxon does comparatively well for these two distributions with a similar reversal for the Wilcoxon efficiency relative to the sign test under double exponential alternatives.

4. Related problems. It would be interesting to have this type of efficiency comparison for other tests. In particular a similar comparison for the many two-sample non-parametric tests would be valuable. However, the two-sample problem appears more difficult because of dependence under the null hypothesis. Other limiting efficiencies obtained with different limits for α and β ($\beta \to 0$) such as studied by Chernoff [2], and Hodges and Lehmann [7] would be useful

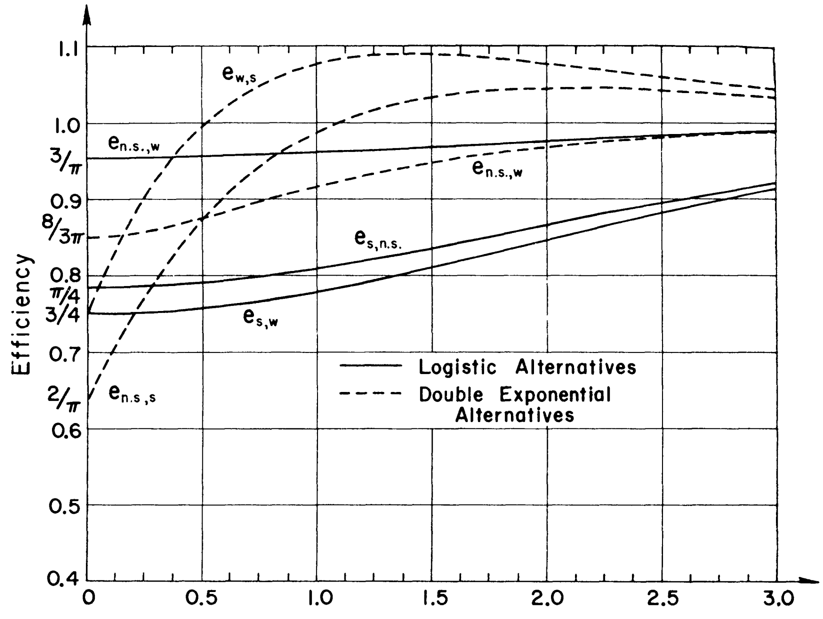

FIG. 2. Relative efficiencies for logistic and double exponential alternatives

for comparison. Their results for the sign test would seem to indicate a less rapid deterioration in normal efficiency values with large μ when using their limiting efficiency definitions.

5. Acknowledgments. I wish to thank Professors H. Chernoff and M. Katz for pointing out reference [3]. My thanks go also to R. Montgomery for programming assistance, members of the Mathematics Department at San Diego State College for stimulating discussions, and the referee for helpful corrections.

REFERENCES

[1] BAHADUR, R. R. (1960). Simultaneous comparison of the optimum and sign tests of a normal mean. *Contributions to Probability and Statistics; Essays in Honor of Harold Hotelling* 79–89. Stanford Univ. Press.

[2] CHERNOFF, H. (1952). A measure of asymptotic efficiency for tests of a hypothesis based on the sum of observations. *Ann. Math. Statist.* **23** 493–507.

[3] FELLER, W. (1943). Generalization of a probability limit theorem of Cramér. *Trans. Amer. Math. Soc.* **54** 361–372.

[4] FRASER, D. A. S. (1957). Most powerful rank type tests. *Ann. Math. Statist.* **27** 324–335.

[5] GNEDENKO, B. V. and KOLMOGOROV, A. N. (1954). *Limit Theorems for Sums of Independent Random Variables.* Addison-Wesley, New York.

[6] GOVINDARAJULU, Z. (1960). Central limit theorems and asymptotic efficiency for one sample nonparametric procedures. *Univ. of Minnesota Technical Report No. 11.*

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

[8] HOEFFDING, W. (1953). On the distribution of the expected values of the order statistics. *Ann. Math. Statist.* **24** 93–100.
[9] KLOTZ, J. H. (1963). Small sample power and efficiency for the one-sample Wilcoxon and normal scores tests. *Ann. Math. Statist.* **34** 624–632.
[10] PINKHAM, R. S. and WILK, M. B. (1963). Tail areas of the t-distribution from a Mills'-ratio-like expansion. *Ann. Math. Statist.* **34** 335–337.
[11] WILCOXON, F. (1945). Individual comparisons by ranking methods. *Biometrics Bull.* **1** 80–83.