ON A MODIFICATION OF CERTAIN RANK TESTS

By Ryoji Tamura

Shimane University

1. Introduction. This note is concerned with a modification of certain well-known rank tests based on the theory of Chernoff-Savage [2]. Let $X_1, X_2, \cdots, X_m$ and $Y_1, Y_2, \cdots, Y_n$ be two independent samples from populations with continuous distribution functions F and G. Our interest is to test the hypothesis $H_0:F \equiv G$. For the problem for location, we assume $G(x) = F(x + \theta)$ and want to test $H_0:\theta = 0$ against the alternative $H_1:\theta > 0$. On the other hand, the problem for scale is to test $H_0:\theta = 1$ against $H_2:\theta > 1$ under the restrictions $G(x) = F(x/\theta)$ and $F(0) = \frac{1}{2}$. There are well-known test statistics, Wilcoxon's U [4] for the location problem and Mood's M [3] and Ansari-Bradley's S [1] for the scale problem. We shall modify these statistics to raise the asymptotic efficiency of the corresponding tests.

2. Test statistics. We adopt the following statistics for $k > 0$,

$$mU_k = \sum_{i=1}^{N} (i/N)^k Z_i, \qquad N = m + n, \tag{1}$$

$$mS_k = \sum_{i=1}^{p} (i/N)^k Z_i + \sum_{i=p+1}^{N} \{(N + 1 - i)/N\}^k Z_i, \quad p = [\tfrac{1}{2}(N + 1)], \tag{2}$$

$$mM_k = \sum_{i=1}^{N} |i/N - (N + 1)/2N|^k Z_i, \tag{3}$$

where Z_i is 1(0) if the ith smallest among the combined sample is $X(Y)$. We may directly apply the theorem of Chernoff-Savage for U_k and S_k and show their asymptotic normality with the mean $\mu_k(\theta)$ and the variance $\sigma_k^2(\theta)$,

$$\begin{aligned} \mu_k(\theta) &= \int_{-\infty}^{\infty} H(x)^k \, dF(x), && \text{for } U_k \\ &= \int_{-\infty}^{0} H(x)^k \, dF(x) + \int_0^{\infty} \{1 - H(x)\}^k \, dF(x), && \text{for } S_k \end{aligned} \tag{4}$$

and under the hypothesis H_0

$$\begin{aligned} N\sigma_k^2 &= [(1 - \lambda_N)/\lambda_N]k^2/(2k + 1)(k + 1)^2, && \text{for } U_k, \\ &= [(1 - \lambda_N)/\lambda_N]k^2/4^k(2k + 1)(k + 1)^2, && \text{for } S_k, \end{aligned} \tag{5}$$

where $H(x)$ is the combined population distribution function and $\lambda_N = m/N$. Though the assumption (4) of their theorem (see [2]) is not satisfied for M_k, we may similarly prove its asymptotic normality from the fact that the following

Received October 25, 1962; revised March 19, 1963.

Institute of Mathematical Statistics is collaborating with JSTOR to digitize, preserve, and extend access to *The Annals of Mathematical Statistics.* www.jstor.org

relations hold

$$J_N(\tfrac{1}{2}) = o(N^{\frac{1}{2}}),$$
$$|J^{(i)}(H)| \leqq K[H(\tfrac{1}{2} - H)]^{-i-\frac{1}{2}+\delta}, \qquad \text{for } 0 < H < \tfrac{1}{2},$$
$$\leqq K[(H - \tfrac{1}{2})(1 - H)]^{-i-\frac{1}{2}+\delta}, \quad \text{for } \tfrac{1}{2} < H < 1,$$

for $i = 0, 1, 2$ and some $\delta > 0$, where $J_N(H) = |H - (N + 1)/2N|^k$ and $J(H) = \lim_{N\to\infty} J_N(H)$. Then we may get the following expression for M_k

$$\mu_k(\theta) = \int_{-\infty}^{0} (\tfrac{1}{2} - H)^k \, dF + \int_{0}^{\infty} (H - \tfrac{1}{2})^k \, dF \tag{6}$$

and under H_0

$$N\sigma_k^2 = [(1 - \lambda_N)/\lambda_N]k^2/(2k + 1)(k + 1)^2 4^k. \tag{7}$$

3. Asymptotic efficacy. Let the alternative be $H_1{:}\theta_N = \gamma/N^{\frac{1}{2}}$ for the location problem and $H_2{:}\theta_N = 1 + \gamma/N^{\frac{1}{2}}$ for the scale problem, $\gamma > 0$. Then it is easy to check the regularity conditions for Pitman efficiency for $k > 0$ from the relations (4), (5), (6) and (7). We shall denote the efficacy of test by R^2 that is defined as $[(\mu_k'(\theta))^2/\sigma_k^2 \mid H_0]$ and $\lim_{N\to\infty} R^2/N = E$. We may compare the relative asymptotic efficiency of tests by knowing the values of E. Thus we get

$$\begin{aligned}
E &= \lambda(1 - \lambda)(2k + 1)(k + 1)^2 \left[\int_{-\infty}^{\infty} fF^{k-1} \, dF\right]^2, && \text{for } U_k, \\
&= \lambda(1 - \lambda)4^k(2k + 1)(k + 1)^2 \left[-\int_{-\infty}^{0} xfF^{k-1} \, dF \right. \\
&\qquad \left. + \int_{0}^{\infty} xf(1 - F)^{k-1} \, dF\right]^2, && \text{for } S_k, \\
&= \lambda(1 - \lambda)4^k(2k + 1)(k + 1)^2 \left[\int_{-\infty}^{0} xf(\tfrac{1}{2} - F)^{k-1} \, dF \right. \\
&\qquad \left. - \int_{0}^{\infty} xf(F - \tfrac{1}{2})^{k-1} \, dF\right]^2, && \text{for } M_k,
\end{aligned} \tag{8}$$

where $\lambda = \lim_{N\to\infty}\lambda_N$.

(a) U_k *test.* We may easily get the following formulas from (8) when the underlying distribution is uniform and exponential,

$$E(u) = \lambda(1 - \lambda)(2k + 1)(k + 1)^2/k^2, \tag{9}$$

$$E(e) = \lambda(1 - \lambda)(2k + 1)/k^2, \tag{10}$$

where $*$ in $E(*)$ denotes the initial letter of the underlying distribution. Thus the smaller $k > 0$ we take, the larger value of E may be obtained. We have also Table I in the normal case after some numerical integration. $E^* = E/\lambda(1 - \lambda)$ and the value when $k = 0$ expresses $\lim_{k\to 0} E$, etc.

(b) S_k *test.* In this case, it is easily shown for the double exponential and uni-

TABLE I

k	0	$\frac{1}{4}$	$\frac{1}{2}$	$\frac{3}{4}$	1	2	3	4
$E^*(n)$	0.81	0.91	0.95	0.96	0.95	0.90	0.82	0.76

TABLE II

k	0	$\frac{1}{4}$	$\frac{1}{2}$	$\frac{3}{4}$	1	2
$E^*(n)$	1.93	1.76	1.55	1.37	1.22	0.81

TABLE III

k	0	1	2	3	4	5	6
$E^*(e)$	0.10	0.75	0.87	0.91	0.926	0.925	0.92
$E^*(n)$	0.56	1.22	1.52	1.77	1.84	1.85	1.83

form distributions that

$$E(u) = \lambda(1 - \lambda)(2k + 1)/k^2 \tag{11}$$

$$E(e) = \lambda(1 - \lambda)(2k + 1)/(k + 1)^2. \tag{12}$$

In the normal case, we get Table II.

(c) M_k *test.* We may get the expression in the uniform case

$$E(u) = \lambda(1 - \lambda)(2k + 1) \tag{13}$$

and also Table III for the double exponential and normal distributions.

In conclusion, it is said that we may constitute the tests with larger efficacy than the existing tests by choosing some suitable value of k.

4. Acknowledgment. I should like to express my gratitude to Dr. M. Okamoto and Mr. N. Sugiura, Osaka University, and the referee for much help and criticism.

REFERENCES

[1] Ansari, A. R. and Bradley, R. A. (1960). Rank-sum tests for dispersions. *Ann. Math. Statist.* **31** 1174–1189.

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

[3] Mood, A. M. (1954). On the asymptotic efficiency of certain nonparametric two sample tests. *Ann. Math. Statist.* **25** 514–522.

[4] Wilcoxon, F. (1945). Individual comparison by ranking methods. *Biometrics Bull.* **1** 80–83.