The Canadian Journal of Statistics
Vol. 18, No. 3, 1990, Pages 245–253
La Revue Canadienne de Statistique

Modifications of the Mack-Wolfe umbrella tests for a generalized Behrens-Fisher problem

Y.I. CHEN and D.A. WOLFE

Ohio State University

Key words and phrases: Distribution-free test; generalized Behrens-Fisher problem; k-sample location problem; Jonckheere-Terpstra test; Mack-Wolfe tests; ordered alternatives; umbrella alternatives.
AMS 1985 subject classifications: 62E20, 62E25, 62G05, 62G10, 62G30.

ABSTRACT

We are concerned with testing procedures for umbrella alternatives in the k-sample location problem without making the assumption that the underlying populations have the same shape. Modifications of the Mack-Wolfe tests are proposed for the cases when the peak of the umbrella is known or unknown. The proposed procedures are exactly distribution-free when the continuous populations have the same shape. The modified test for peak-known umbrella alternatives remains asymptotically distribution-free when the continuous populations are symmetric, but not necessarily with the same shape.

RÉSUMÉ

On s'intéresse à des procédures pour tester l'égalité de k paramètres de position versus des alternatives de type parapluie ($\theta_1 \leq \cdots \leq \theta_p \geq \ldots \geq \theta_k$, pour un p, avec au moins une inégalité stricte), cela sans supposer que les populations sous-jacentes ont la même forme. On propose des modifications aux tests de Mack-Wolfe dans le cas où le sommet du parapluie est connu et aussi dans le cas où il est inconnu. Lorsque les lois des populations sous-jacentes sont de type continu et ont la même forme, les procédures proposées sont indépendantes de celles-ci. Dans le cas d'alternatives avec sommet connu, le test modifié demeure asymptotiquement indépendant des populations sous- jacentes si celles-ci sont de type continu et symétriques, mais pas nécessairement de la même forme.

1. INTRODUCTION

Suppose that $X_{i1}, \ldots, X_{in_i}$, $i = 1, \ldots, k$, are k independent random samples from populations with continuous distribution functions $F_1(x), \ldots, F_k(x)$, respectively. For each $i = 1, \ldots, k$, let θ_i be the unique median of the ith population. In this article, we consider testing the null hypothesis $\mathcal{H}_0 : (\theta_1 = \cdots = \theta_k)$ against the umbrella alternatives $\mathcal{H}_A : (\theta_1 \leq \cdots \leq \theta_p \geq \ldots \geq \theta_k$, for some p, with at least one strict inequality) without assuming the same shapes for the k populations. Since we are concerned with testing for location parameters without making the assumption that the underlying populations have the same shape, this problem can be regarded as a generalization of the Behrens-Fisher problem.

Nonparametric tests for differences between two medians in the generalized Behrens-Fisher problem have been extensively studied; see Fligner and Policello (1981) for detailed references. For a k-sample setting with symmetric underlying populations having possibly different shapes, Rust and Fligner (1984) suggested an asymptotically distribution-free test for general location alternatives based on a modification of the Kruskal-Wallis statistic (Kruskal and Wallis 1952). They also noted that their modified Kruskal-Wallis test is distribution-free when the populations are identical.

The ordinary nonparametric tests for umbrella alternatives, such as the Mack-Wolfe tests (Mack and Wolfe 1981), require the assumption that the continuous populations have the same shape to ensure the distribution-free property. However, the levels of these tests will not necessarily be preserved when the populations have different shapes or scale parameters. In this paper, rank-based modifications of the Mack-Wolfe tests are proposed which are exactly distribution-free when the continuous populations have the same shape. In addition, the modified Mack-Wolfe test for a peak-known umbrella alternative is still asymptotically distribution-free when the continuous populations are assumed symmetric, even if they differ in shape.

In Section 2 we review the Mack-Wolfe tests for umbrella location alternatives with either known or unknown umbrella peak. In Section 3 we modify the Mack-Wolfe statistics to obtain tests in the generalized Behrens-Fisher problem for both the setting where the peak of the umbrella is known and that where it is unknown. In Section 4 we present and discuss the results of a substantial Monte Carlo level and power study.

2. MACK-WOLFE TESTS

For testing $\mathcal{H}_0$ against an arbitrary peak-known (p) umbrella alternative $\mathcal{H}_A$, Mack and Wolfe (1981) suggested rejecting $\mathcal{H}_0$ for large values of

$$A_p = \sum_{i=1}^{p-1} \sum_{j=i+1}^{p} U_{ij} + \sum_{i=p}^{k-1} \sum_{j=i+1}^{k} U_{ji}, \tag{2.1}$$

where U_{ij} is the usual Mann-Whitney statistic (Mann and Whitney 1947) corresponding to the number of observations in sample j that exceed observations in sample i. In particular, the test based on A_k is the Jonckheere-Terpstra test (Jonckheere 1954, Terpstra 1952) for ordered location altneratives. Moreover, suppose that $N \longrightarrow \infty$ in such a way that $n_i/N \longrightarrow \lambda_i$, with $0 < \lambda_i < 1$, $i = 1, \ldots, k$. Mack and Wolfe also noted that, under $\mathcal{H}_0$, the statistic

$$A_p^* = \frac{A_p - \mathcal{E}_0 A_p}{(\mathcal{V}ar_0\, A_p)^{\frac{1}{2}}} \tag{2.2}$$

has an asymptotic $(N \longrightarrow \infty)$ distribution that is standard normal, where

$$\mathcal{E}_0 A_p = \frac{1}{4}\left(N_1^2 + N_2^2 - \sum_{i=1}^{k} n_i^2 - n_p^2\right) \tag{2.3}$$

and

$$\mathcal{V}ar_0\, A_p = \frac{1}{72}\left(2(N_1^3 + N_2^3) + 3(N_1^2 + N_2^2) - \sum_{i=1}^{k} n_i^2(2n_i + 3) - n_p^2(2n_p + 3) \right.$$
$$\left. + 12 n_p N_1 N_2 - 12 n_p^2 N\right), \tag{2.4}$$

with $N_1 = \sum_{i=1}^{p} n_i$ and $N_2 = \sum_{i=p}^{k} n_i$, are the mean and variance, respectively, of A_p when the F_i's are identical.

For the more general unknown-peak alternatives and $t = 1, \ldots, k$, let

$$Z_t = \sum_{i \neq t} U_{it} \tag{2.5}$$

and set

$$Z_t^* = \frac{Z_t - \mathcal{E}_0 Z_t}{(\mathcal{V}ar_0\, Z_t)^{\frac{1}{2}}},$$

where

$$\mathcal{E}_0 Z_t = \frac{n_t(N - n_t)}{2} \tag{2.6}$$

and

$$\mathcal{V}ar_0\, Z_t = \frac{n_t(N - n_t)(N + 1)}{12}, \tag{2.7}$$

with $N = \sum_{t=1}^{k} n_t$, are the respective mean and variance of Z_t when $F_1(x) = \cdots = F_k(x)$. For the unknown-peak alternative, Mack and Wolfe then proposed to reject $\mathcal{H}_0$ for large values of

$$A_{\hat{p}}^* = \frac{A_{\hat{p}} - \mathcal{E}_0 A_{\hat{p}}}{(\mathcal{V}ar_0\, A_{\hat{p}})^{\frac{1}{2}}}, \tag{2.8}$$

where $\hat{p}$ is a sample estimate of the unknown peak p such that $Z_{\hat{p}}^* = \max\{Z_t^*, t = 1, \ldots, k\}$. It was noted, however, that there is a positive probability of observing several (say r) populations tied for the largest value of Z_t^*. In this case, the values of $A_{\hat{p}}^*$ is set equal to the average of those standardized p known statistics corresponding to peaks at each of the r samples tied for the maximum Z_t^*.

3. MODIFICATIONS OF MACK-WOLFE TESTS

When the underlying populations are symmetric, the problem considered in this paper is in fact that of testing the null hypothesis $\mathcal{H}_0^*$: ($\pi_{ij} = \frac{1}{2}$ for all pairs of i and j) against the class of alternatives $\mathcal{H}_A^*$: ($\pi_{ij} \geq \frac{1}{2}$, $1 \leq i < j \leq p$, and $\pi_{ij} \leq \frac{1}{2}$, $p \leq i < j \leq k$, for some p, with at least one strict inequality), where $\pi_{ij} = \mathrm{pr}(X_{j1} \geq X_{i1}) = \int F_i \, dF_j$, $i \neq j = 1, \ldots, k$. It is obvious that, under $\mathcal{H}_0^*$, the expected values in (2.3) and (2.6) remain the same. However, when the underlying populations have different shapes, the variances in (2.4) and (2.7) are changed even under $\mathcal{H}_0^*$. To modify the Mack-Wolfe statistics for testing umbrella location alternatives with fewer assumptions on the shapes of the populations, we therefore need to first find the respective variances of Z_t (2.5) and A_t (2.1), $t = 1, \ldots, k$, under a general setting. Let

$$\phi_{ijt} = \int F_i F_j \, dF_t - \left(\int F_i \, dF_t \right) \left(\int F_j \, dF_t \right), \qquad i, j, t = 1, \ldots, k.$$

From the results of Birnbaum and Klose (1957), we have, for $i \neq j = 1, \ldots, k$,

$$\mathcal{E}\, U_{ij} = n_i n_j \pi_{ij} \tag{3.1}$$

and

$$\mathcal{V}ar\;\, U_{ij} = n_i n_j \{(n_j - 1)\phi_{jji} + (n_i - 1)\phi_{iij} + \pi_{ij}\pi_{ji}\}. \tag{3.2}$$

After some algebraic manipulations, we also have the following result:

$$Cov(U_{ij}, U_{rs}) = \begin{cases} n_i n_j n_s \phi_{jsi} & \text{for} \quad i = r, \quad j \neq s, \\ n_i n_j n_r \phi_{irj} & \text{for} \quad i \neq r, \quad j = s, \\ -n_i n_j n_s \phi_{isj} & \text{for} \quad j = r, \quad i \neq s, \\ -n_i n_j n_r \phi_{jri} & \text{for} \quad j \neq r, \quad i = s, \\ 0 & \text{if } i, j, r, s \text{ are distinct.} \end{cases} \tag{3.3}$$

By using the results in (3.1), (3.2), and (3.3) we obtain, after some straightforward computations, that, for $t = 1, \ldots, k$,

$$\mathcal{V}ar\ Z_t = \sum_{i \neq t} n_i n_t \{(n_i - 1)\phi_{iit} + (n_t - 1)\phi_{tti} + \pi_{it}\pi_{ti}\} + 2 \sum\sum_{i<j;\, i\neq t;\, j\neq t} n_i n_j n_t \phi_{ijt} \tag{3.4}$$

and

$$\begin{aligned}\mathcal{V}ar\ A_t = & \sum_{i=1}^{t-1}\sum_{j=i+1}^{t} n_i n_j \{(n_i - 1)\phi_{iij} + (n_j - 1)\phi_{jji} + \pi_{ij}\pi_{ji}\} \\ & + \sum_{i=t}^{k-1}\sum_{j=i+1}^{k} n_i n_j \{(n_i - 1)\phi_{iij} + (n_j - 1)\phi_{jji} + \pi_{ij}\pi_{ji}\} \\ & + 2\Bigg(\sum_{i=1}^{t-2}\sum_{j=i+1}^{t-1}\sum_{s=j+1}^{t} n_i n_j n_s (\phi_{ijs} + \phi_{jsi} - \phi_{isj}) \\ & + \sum_{i=t}^{k-2}\sum_{j=i+1}^{k-1}\sum_{s=j+1}^{k} n_i n_j n_s (\phi_{ijs} + \phi_{jsi} - \phi_{isj}) \\ & + n_t \sum_{i=1}^{t-1}\sum_{j=t+1}^{k} n_i n_j \phi_{ijt}\Bigg).\end{aligned} \tag{3.5}$$

In what follows we find consistent estimators of $N^{-\frac{3}{2}}\ \mathcal{V}ar\ Z_t$ and $N^{-\frac{3}{2}}\ \mathcal{V}ar\ A_t$, $t = 1, \ldots, k$. Following the suggestion of Fligner and Policello (1981), we estimate the π_{ij}'s and ϕ_{ijt}'s by replacing the F_i's with their sample distribution function analogues F_{n_i}. For $i \neq j = 1, \ldots, k$, let

$$P_{ij}^{v} = n_i F_{n_i}(X_{jv}) = \sum_{u=1}^{n_i} \psi(X_{jv} - X_{iu}), \qquad v = 1, \ldots, n_j,$$

and

$$\bar{P}_{ij} = \sum_{v=1}^{n_j} P_{ij}^{v} / n_j,$$

where

$$\psi(a) = \begin{cases} 1 & \text{for} \quad a \geq 0, \\ 0 & \text{for} \quad a < 0. \end{cases}$$

Note that the statistic P_{ij}^{v} is actually the placement of X_{jv} with respect to the ith sample (Orban and Wolfe 1982). We then estimate π_{ij} and ϕ_{ijt} by, respectively,

$$\hat{\pi}_{ij} = \bar{P}_{ij} / n_i$$

and

$$\begin{aligned}\hat{\phi}_{ijt} &= \int F_{n_i} F_{n_j}\, dF_{n_t} - \left(\int F_{n_i}\, dF_{n_t}\right)\left(\int F_{n_j}\, dF_{n_t}\right) \\ &= \frac{1}{n_i n_j n_t} \sum_{v=1}^{n_t} (P_{it}^{v} - \bar{P}_{it})(P_{jt}^{v} - \bar{P}_{jt}).\end{aligned}$$

Now we estimate the exact variances, $\mathcal{V}ar\ Z_t$ and $\mathcal{V}ar\ A_t$, by replacing the involved π_{ij}'s and ϕ_{ijt}'s with the $\hat{\pi}_{ij}$'s and $\hat{\phi}_{ijt}$'s, respectively. However, in order to simplify the computation of the estimators, we set

$$w_{ijt} = \sum_{v=1}^{n_t} (P_{it}^v - \bar{P}_{it})(P_{jt}^v - \bar{P}_{jt}), \qquad i,j,t = 1,\ldots,k,$$

and replace the $(n_i - 1)$'s with n_i's. The estimators of $\mathcal{V}ar\ Z_t$ and $\mathcal{V}ar\ A_t$ are then given by

$$\widehat{\mathcal{V}ar}\ Z_t = \sum_{i \neq t} (w_{iit} + w_{tti} + \bar{P}_{it}\bar{P}_{ti}) + 2 \sum \sum_{i<j;\, i\neq t;\, j\neq t} w_{ijt} \tag{3.6}$$

and

$$\begin{aligned} \widehat{\mathcal{V}ar}\ A_t = {} & \sum_{i=1}^{t-1} \sum_{j=i+1}^{t} (w_{iij} + w_{jji} + \bar{P}_{ij}\bar{P}_{ji}) + \sum_{i=t}^{k-1} \sum_{j=i+1}^{k} (w_{iij} + w_{jji} + \bar{P}_{ij}\bar{P}_{ji}) \\ & + 2\Bigg(\sum_{i=1}^{t-2} \sum_{j=i+1}^{t-1} \sum_{s=j+1}^{t} (w_{ijs} + w_{jsi} - w_{isj}) \\ & + \sum_{i=t}^{k-2} \sum_{j=i+1}^{k-1} \sum_{s=j+1}^{k} (w_{ijs} + w_{jsi} - w_{isj}) + \sum_{i=1}^{t-1} \sum_{j=t+1}^{k} w_{ijt} \Bigg), \end{aligned} \tag{3.7}$$

respectively, $t = 1,\ldots,k$. Consequently, we propose rejecting $\mathcal{H}_0{}^*$ in favor of the peak-known (p) umbrella alternative $\mathcal{H}_A{}^*$ for large values of

$$\hat{A}_p^* = \frac{A_p - \mathcal{E}_0 A_p}{(\widehat{\mathcal{V}ar}\ A_p)^{\frac{1}{2}}}, \tag{3.8}$$

where A_p, $\mathcal{E}_0 A_p$, and $\widehat{\mathcal{V}ar}\ A_p$ are given in Equations (2.1), (2.3), and (3.7), respectively. For the more realistic practical setting where the peak is unknown, we first choose the group $\hat{p}$ such that $\hat{Z}_{\hat{p}}^* = \max\{\hat{Z}_t^*,\ t = 1,\ldots,k\}$, where $\hat{Z}_t^* = (Z_t - \mathcal{E}_0 Z_t)/(\widehat{\mathcal{V}ar}\ Z_t)^{\frac{1}{2}}$, $t = 1,\ldots,k$, with Z_t, $\mathcal{E}_0 Z_t$, and $\widehat{\mathcal{V}ar}\ Z_t$ given by (2.5), (2.6), and (3.6), respectively. The null hypothesis $\mathcal{H}_0{}^*$ is then rejected for large values of

$$\hat{A}_{\hat{p}}^* = \frac{A_{\hat{p}} - \mathcal{E}_0 A_{\hat{p}}}{(\widehat{\mathcal{V}ar}\ A_{\hat{p}})^{\frac{1}{2}}}. \tag{3.9}$$

For the situation where two ore more groups are tied for having the largest $\hat{Z}_t^*$ sample values, let χ be the set of the groups tied for the maximum $\hat{Z}_t^*$. We then take the value of $\hat{A}_{\hat{p}}$ as the average of the $\hat{A}_t^*$'s for those t in the set χ.

Suppose that $N \longrightarrow \infty$ in such a way that $n_i/N \longrightarrow \lambda_i$, with $0 < \lambda_i < 1$, $i = 1,\ldots,k$. From the results of Archambault, Mack, and Wolfe (1977), we observe that the random variable $(A_p - \mathcal{E}_0 A_p)/(\mathcal{V}ar\ A_p)^{\frac{1}{2}}$ has an asymptotic $(N \longrightarrow \infty)$ null $(\mathcal{H}_0{}^*)$ distribution that is standard normal, where A_p, $\mathcal{E}_0 A_p$, and $\mathcal{V}ar\ A_p$ are given in Equations (2.1), (2.3), and (3.5), respectively. Furthermore, applying the Glivenko-Cantelli theorem [see, for example, Theorem 2.1.4A of Serfling (1980)], it follows that F_{n_i} converges uniformly to F_i with probability one for $i = 1,\ldots,k$. Using this result we obtain that $(\mathcal{V}ar\ A_p)/(\widehat{\mathcal{V}ar}$

A_p) converges to one almost surely as $N \longrightarrow \infty$. This implies that the statistic $\hat{A}_p^*$ (3.8) has an asymptotic ($N \longrightarrow \infty$) null ($\mathcal{H}_0^*$) distribution that is standard normal. Therfore, we observe that the test based on $\hat{A}_p^*$ is asymptotically distribution-free under $\mathcal{H}_0^*$.

Note that since the $\widehat{Var}\ Z_t$'s and $\widehat{Var}\ A_t$'s involve ranks only, the tests based on $\hat{A}_p^*$ and $\hat{A}_{\hat{p}}^*$ are both exactly distribution-free when the populations are identical. In addition, for the case $k = 2$, the test based on either $\hat{A}_1^*$ or $\hat{A}_2^*$ is the same as the modified Mann-Whitney test proposed by Fligner and Policello (1981) for differences between two medians.

4. MONTE CARLO STUDY

4.1. Discussion of Study.

To compare tests based on the modified Mack-Wolfe statistics $\hat{A}_p^*$ (3.8) and $\hat{A}_{\hat{p}}^*$ (3.9) with those based on the original Mack-Wolfe statistics A_p^* and $A_{\hat{p}}^*$ given in (2.2) and (2.8), respectively, we conducted a Monte Carlo study. For these simulations, we selected three families of distributions: normal, contaminated normal, and Cauchy. Appropriate normal and Cauchy deviates were generated by the International Mathematical and Statistical Libraries (IMSL) routines rnnor and rnchy. The contaminated normal distribution utilized was a mixture of the standard normal distribution and a normal distribution with mean zero and standard deviation 5 in proportions 0.9 and 0.1, respectively.

To study the effect that heteroscedasticity has on the significance levels of the test procedures, we considered distributions with the same medians but different scale parameters, namely, $F_1(x), \ldots, F_k(x)$ with $F_i(x) = F(x/\sigma_i)$, $i = 1, \ldots, k$, and $F(0) = \frac{1}{2}$. Several choices of $\sigma_2/\sigma_1, \ldots, \sigma_k/\sigma_1$ in combination with the three distributions mentioned above were studied. Note that for the case of known umbrella peak (p) the level performance of the test based on A_p^* relative to that of the test based on $\hat{A}_p^*$ is similar for $p = 1, \ldots, k$. Therefore, we simply considered the case $p = k$ in this study. The estimated levels are presented in Tables 1 and 2.

The results of a second Monte Carlo study, designed to compare the powers of the modified tests with the original tests for a variety of umbrella location alternatives when the populations are otherwise the same, are presented in Tables 3 and 4. Specifically, we considered distribution functions $F_i(x) = F(x - \theta_i)$, $i = 1, \ldots, k$, for various choices of $\theta_2 - \theta_1, \ldots, \theta_k - \theta_1$ and F being normal, contaminated normal, or Cauchy.

Both the level and power studies were conducted for $k = 3$ and $k = 4$ populations with $n_1 = \cdots = n_k = 10$ observations per sample. For each setting we used 10,000 replications, and the estimated level or power was obtained by computing the frequency of the test statistic falling in the level-0.10 critical region. Since we took 0.10 as the nominal level of the tests, the standard deviation of the estimated levels in Tables 1 and 2 is $0.003 = \{(0.10)(0.90)/10{,}000\}^{\frac{1}{2}}$. We then indicate, by + (−) signs, whenever the estimated level is two or more standard deviations above (below) 0.10.

4.2. Discussion of Results.

It can be seen from Tables 1 and 2 that the tests based on the statistics A_k^* and $A_{\hat{p}}^*$ do not hold their levels when the populations have different scale parameters, while those based on the modifications $\hat{A}_k^*$ and $\hat{A}_{\hat{p}}^*$ hold their levels quite well across all situations. These findings also demonstrate the fact that the modified tests are exactly distribution-free when the distributions are identical. Following the results in Section 3, we have, for

TABLE 1: Estimated levels for nominal $\alpha = 0.10$ when $k = 3$ and $n_1 = n_2 = n_3 = 10$.[a]

Distribution	σ_2/σ_1	σ_3/σ_1	A_k^*	$\hat{A}_k^*$	$A_{\hat{p}}^*$	$\hat{A}_{\hat{p}}^*$
Normal	1	1	0.103	0.103	0.100	0.101
	1	2	0.127+	0.107+	0.119+	0.101
	2	1	0.058−	0.096	0.083−	0.096
	2	2	0.088−	0.103	0.094−	0.103
	1	3	0.135+	0.109+	0.136+	0.105
	3	1	0.036−	0.086−	0.082−	0.094−
	3	3	0.087−	0.103	0.098	0.104
Contaminated normal	1	1	0.104	0.103	0.091−	0.092−
	1	2	0.125+	0.108+	0.116+	0.101
	2	1	0.067−	0.099	0.077−	0.092−
	2	2	0.094−	0.103	0.092−	0.100
	1	3	0.127+	0.102	0.128+	0.101
	3	1	0.042−	0.083−	0.081−	0.093−
	3	3	0.094−	0.106+	0.097	0.102
Cauchy	1	1	0.098	0.100	0.098	0.099
	1	2	0.112+	0.102	0.110+	0.103
	2	1	0.073−	0.096	0.088−	0.100
	2	2	0.090−	0.100	0.095	0.099
	1	3	0.119+	0.103	0.118+	0.104
	3	1	0.062−	0.092−	0.087−	0.099
	3	3	0.088−	0.099	0.097	0.101

[a]+: At least two standard deviations above 0.10, computed as if $\alpha = 0.10$. −: At least two standard deviations below 0.10, computed as if $\alpha = 0.10$.

TABLE 2: Estimated levels for nominal $\alpha = 0.10$ when $k = 4$ and $n_1 = n_2 = n_3 = n_4 = 10$.[a]

Distribution	σ_2/σ_1	σ_3/σ_1	σ_4/σ_1	A_k^*	$\hat{A}_k^*$	$A_{\hat{p}}^*$	$\hat{A}_{\hat{p}}^*$
Normal	1	1	1	0.100	0.101	0.101	0.101
	1	1	2	0.125+	0.103	0.126+	0.101
	1	2	3	0.124+	0.104	0.127+	0.099
	2	1	3	0.115+	0.103	0.115+	0.097
	3	3	1	0.032−	0.091−	0.078−	0.093−
	1	3	5	0.135+	0.105	0.138+	0.102
	5	3	3	0.060−	0.096	0.082−	0.096
Contaminated normal	1	1	1	0.097	0.097	0.100	0.101
	1	1	2	0.111+	0.100	0.120+	0.100
	1	2	3	0.110+	0.099	0.121+	0.101
	2	1	3	0.105	0.100	0.114+	0.098
	3	3	1	0.055−	0.094−	0.074−	0.090−
	1	3	5	0.118+	0.100	0.133+	0.101
	5	3	3	0.072−	0.097	0.084−	0.098
Cauchy	1	1	1	0.099	0.100	0.103	0.101
	1	1	2	0.115+	0.101	0.117+	0.100
	1	2	3	0.113+	0.102	0.115+	0.100
	2	1	3	0.110+	0.102	0.111+	0.102
	3	3	1	0.055−	0.096	0.082−	0.098
	1	3	5	0.121+	0.104	0.121+	0.099
	5	3	3	0.073−	0.098	0.086−	0.097

[a]+: At least two standard deviations above 0.10, computed as if $\alpha = 0.10$. −: At least two standard deviations below 0.10, computed as if $\alpha = 0.10$.

TABLE 3: Estimated powers for nominal $\alpha = 0.10$ when $k = 3$ and $n_1 = n_2 = n_3 = 10$.

Distribution	θ_2-θ_1	θ_3-θ_1	A_p^*	$\hat{A}_p^*$	$A_{\hat{p}}^*$	$\hat{A}_{\hat{p}}^*$
Normal	0.0	1.0	0.799	0.806	0.601	0.620
	0.5	1.0	0.808	0.806	0.610	0.613
	1.0	2.0	0.999	0.999	0.988	0.988
	1.0	0.0	0.883	0.881	0.749	0.749
	1.0	0.5	0.709	0.711	0.574	0.585
	2.0	1.0	0.988	0.990	0.971	0.976
Contaminated normal	0.0	1.0	0.692	0.699	0.474	0.486
	0.5	1.0	0.714	0.705	0.491	0.490
	1.0	2.0	0.986	0.983	0.932	0.923
	1.0	0.0	0.781	0.774	0.616	0.613
	1.0	0.5	0.601	0.602	0.472	0.482
	2.0	1.0	0.946	0.944	0.881	0.882
Cauchy	0.0	1.0	0.445	0.443	0.248	0.250
	0.5	1.0	0.458	0.450	0.274	0.267
	1.0	2.0	0.804	0.786	0.612	0.585
	1.0	0.0	0.515	0.503	0.353	0.349
	1.0	0.5	0.392	0.388	0.285	0.286
	2.0	1.0	0.698	0.690	0.566	0.564

TABLE 4: Estimated powers for nominal $\alpha = 0.10$ when $k = 4$ and $n_1 = n_2 = n_3 = n_4 = 10$.

Distribution	θ_2-θ_1	θ_3-θ_1	θ_4-θ_1	A_p^*	$\hat{A}_p^*$	$A_{\hat{p}}^*$	$\hat{A}_{\hat{p}}^*$
Normal	0.0	0.0	1.0	0.754	0.772	0.503	0.518
	0.0	0.5	1.0	0.860	0.864	0.649	0.645
	0.5	1.0	1.0	0.861	0.862	0.698	0.690
	0.5	1.0	1.5	0.985	0.983	0.909	0.905
	0.0	1.0	0.0	0.870	0.872	0.693	0.700
	0.0	1.0	0.5	0.764	0.769	0.607	0.616
	0.5	1.0	0.5	0.772	0.767	0.613	0.601
	0.5	1.0	0.0	0.874	0.871	0.741	0.731
Contaminated normal	0.0	0.0	1.0	0.664	0.671	0.410	0.410
	0.0	0.5	1.0	0.774	0.770	0.530	0.521
	0.5	1.0	1.0	0.776	0.773	0.589	0.576
	0.5	1.0	1.5	0.941	0.935	0.809	0.791
	0.0	1.0	0.0	0.773	0.767	0.565	0.565
	0.0	1.0	0.5	0.658	0.659	0.504	0.508
	0.5	1.0	0.5	0.669	0.662	0.513	0.498
	0.5	1.0	0.0	0.782	0.772	0.618	0.605
Cauchy	0.0	0.0	1.0	0.415	0.420	0.205	0.203
	0.0	0.5	1.0	0.500	0.494	0.269	0.260
	0.5	1.0	1.0	0.503	0.499	0.326	0.317
	0.5	1.0	1.5	0.698	0.680	0.460	0.435
	0.0	1.0	0.0	0.508	0.499	0.309	0.307
	0.0	1.0	0.5	0.420	0.418	0.284	0.279
	0.5	1.0	0.5	0.435	0.424	0.298	0.282
	0.5	1.0	0.0	0.513	0.504	0.356	0.347

large N,

$$\mathrm{pr}(A_p^* \geq z_\alpha | \mathcal{H}_0^*) \simeq 1 - \Phi\left(z_\alpha \left(\frac{\mathcal{V}ar_0\, A_p}{\mathcal{V}ar\, A_p} \right)^{\frac{1}{2}} \right),$$

where $\Phi(z_\alpha) = 1 - \alpha$, with Φ being the standard normal distribution function, and A_p^*, $\mathcal{V}ar_0\, A_p$, and $\mathcal{V}ar\, A_p$ are given by (2.2), (2.4), and (3.5), respectively. This means that the asymptotic level of the test based on A_p^* depends on the value of $(\mathcal{V}ar_0\, A_p / \mathcal{V}ar\, A_p)$. This, in some sense, explains the evidence presented in Tables 1 and 2 that for some choices of $\sigma_2/\sigma_1, \ldots, \sigma_k/\sigma_1$ the level of the based on A_k^* is inflated, while for the other choices its level is deflated.

The power study presented in Tables 3 and 4 shows that, for small sample sizes, the estimated powers of the modified tests are sometimes slightly lower than those of the corresponding original tests. However, these small power differences do not seem too high a price to pay for holding the levels over the broader null hypothesis.

In conclusion, we recommend use of the modified Mack-Wolfe tests for two reasons. First, since the modified tests are strictly distribution-free when the populations are identical, the levels of these tests are exactly controlled for different distributional types, just as with the original procedures. Second, the levels of the modified tests are also maintained when the populations have different scales, and for small sample sizes there is no appreciable loss of power relative to the associated unmodified tests when the populations differ only in locations.

REFERENCES

Archambault, W.A.T., Jr., Mack, G.A., and Wolfe, D.A. (1977). K-sample rank tests using pair-specific scoring functions. *Canad. J. Statist.*, 5, 195–207.

Birnbaum, Z.W., and Klose, O.M. (1957). Bounds for the variance of the Mann-Whitney statistic. *Ann. Math. Statist.*, 28, 933–945.

Fligner, M.A., and Policello, G.E., II (1981). Robust rank procedures for the Behrens-Fisher problem. *J. Amer. Statist. Assoc.*, 76, 162–168.

Jonckheere, A.R. (1954). A distribution-free k-sample test against ordered alternatives. *Biometrika*, 41, 133–145.

Kruskal, W.H., and Wallis, W.A. (1952). Use of ranks in one-criterion variance analysis. *J. Amer. Statist. Assoc.*, 47, 583–621.

Mack, G.A., and Wolfe, D.A. (1981). K-sample rank tests for umbrella alternatives. *J. Amer. Statist. Assoc.*, 76, 175–181.

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

Orban, J., and Wolfe, D.A. (1982). A class of distribution-free two-sample tests based on placements. *J. Amer. Statist. Assoc.*, 77, 666–672.

Rust, S.W., and Fligner, M.A. (1984). A modification of the Kruskal-Wallis statistic for the generalized Behrens-Fisher problem. *Comm. Statist. Theory Methods*, A13(16), 2013–2027.

Serfling, R.J. (1980). *Approximation Theorems of Mathematical Statistics.* Wiley, New York.

Terpstra, T.J. (1952). The asymptotic normality and consistency of Kendall's test against trend, when ties are present in one ranking. *Indag. Math.*, 14, 327–333.

Received 1 February 1989
Revised 31 January 1990
Accepted 3 February 1990

Department of Statistics
Ohio State University
141 Cockins Hall
1958 Neil Avenue
Columbus, OH 43210-1247
U.S.A.