Journal of the KSQC
Vol.21, No.2, Dec 1993

1993년 12월 韓國品質管理學會誌 제21권 제2호

Modifications of the Weighted Mack - Wolfe Umbrella Tests for a Generalized Behrens - Fisher Problem

Hee Moon Park*

ABSTRACT

Modifications of the weighted Mack-Wolfe tests are proposed for both cases when the peak of umbrella is known and unknown. The modified weighted Mack-Wolfe tests are exactly distribution-free when the continuous populations have the same shape. For the case of peak-known umbrella alternatives, the modified weighted Mack-Wolfe tests remain asymptotically distribution-free when the continuous populations are symmetric, but not necessarily with the same shape. For the case of peak-unknown umbrella alternatives, the maximum of standardized modified weighted Mack-Wolfe tests with peak-known umbrella alternatives was used. The simulation results show that the modified weighted Mack-Wolfe tests are more recommended than the modified Mack-Wolfe tests for various patterns.

1. Introduction

Suppose that $X_{i1}, \cdots, X_{in_i}$, $i=1, \cdots, k$, are k-independent random samples from

* Department of Computer Science and Statistics, Seoul National University, Seoul, Korea

populations with continuous distribution functions $F_i(x)$, $i=1, \cdots, k$. Let θ_i be the unique median of the i-th population. In this paper, we are interested in testing the null hypothesis.

$$H_0 : \theta_1 = \cdots = \theta_k$$

against the umbrella alternatives

$$H_u : \theta_1 \leq \cdots \leq \theta_l \geq \cdots \geq \theta_k,$$

for some l, with at least one strict inequality. Suppose that the shapes of the underlying k populations are not all the same. Since we are interested in testing the location parameters without assumption that the underlying populations have the same shape, this problem can be regarded as a generalization of the Behrens-Fisher problem.

Various nonparametric tests for a generalized Behrens-Fisher problem in testing the differences of location parameters have been extensively studied. Fligner and Policello(1981) proposed a robust rank procedures for the difference of two medians using the modification of Mann-Whitney-Wilcoxon test. For the several sample location problems with continuous symmetric populations having different shapes, Rust and Fligner(1984) proposed an asymptotically distribution-free tests using the modification of well-known Kruskal-Wallis test for the general alternatives. They also stated that the modified Kruskal-Wallis test is distribution-free when the underlying populations have the same shape.

For the umbrella alternatives, the usual nonparametric tests such as the Mack-Wolfe tests(Mack and Wolfe, 1981) and the weighted Mack-Wolfe tests(Park, 1993) require the assumption that the continuous populations have the same shape to ensure the distribution-free property. However, the significance levels of these tests will not necessarily be preserved when the populations have different shapes or scale parameters. Chen and Wolfe(1990b) proposed the modifications of Mack-Wolfe tests which are exactly distribution-free when the continuous populations have the same shape. Also, the modified Mack-Wolfe tests for a peak-known umbrella alternatives is still asymptotic distribution-free.

In this paper, the rank-based modifications of weighted Mack-Wolfe tests are proposed. The modified weighted Mack-Wolfe tests are satisfied with the properties of the modified Mack-Wolfe tests. In Section 2, the weighted Mack-Wolfe tests(Park, 1993) for umbrella alternatives are reviewed when the peak of umbrella is known or unknown. In Section 3, the rank-based modifications of weighted Mack-Wolfe tests in a generalized Behrens-Fisher problem are proposed for both cases when the peak is known and unknown. In Section 4, a small-sample Monte-Carlo simulation results for the comparison between the modification of Mack-Wolfe and those of weighted Mack-Wolfe tests are presented and discussed.

2. Weighted Mack-Wolfe Tests

When the peak l is known, for testing the null hypothesis H_0 against the umbrella alternatives H_u, Park(1993) considered some weighted Mack-Wolfe test statistic which is expected to have good performance on power properties. The proposed test statistic is defined by

$$A_{lw}=\sum_{i=1}^{l-1}\sum_{j=i+1}^{l} w_{ij}U_{ij}+\sum_{i=l}^{k-1}\sum_{j=i+1}^{k} w_{ji}U_{ji}, \tag{2. 1}$$

where w_{ij} are weights and U_{ij} are the usual Mann-Whitney counts between the i-th and j-the samples. If $w_{ij}=1$ for $1\le i<j\le l$ and $w_{ji}=1$ for $l\le i<j\le k$, then the weighted Mack-Wolfe test A_{lw} is equal to the Mack-Wolfe(1981) test defined by

$$A_l=\sum_{i=1}^{l-1}\sum_{j=i+1}^{l} U_{ij}+\sum_{i=l}^{k-1}\sum_{j=i+1}^{k} U_{ji}\ . \tag{2. 2}$$

A_l used the combination of an ordinary Jonckheere-Terpstra(Jonckheere(1954), Terpstra(1952)) and a reversed Jonckheere-Terpstra statistic according to the direction of the peak l. Note that A_{lw} includes the Jonckheere-Terpstra statistic A_k for testing the ordered alternatives. Suppose that $N\to\infty$ in such way that $n_i/N\to\lambda_i\in(0,\ 1)$, $i=1,\ \cdots, k$. Then under H_0,

$$A_{lw}^{*}=\frac{A_{lw}-E_0(A_{lw})}{\{Var_0(A_{lw})\}^{\frac{1}{2}}} \tag{2. 3}$$

has a limiting standard normal distribution, where

$$E_0(A_{lw})=\frac{1}{2}\{\sum_{i=1}^{l-1}\sum_{j=i+1}^{l} w_{ij}n_in_j+\sum_{i=l}^{k-1}\sum_{j=i+1}^{k} w_{ji}n_in_j\} \tag{2. 4}$$

and

$$Var_0(A_{lw}) = \frac{1}{12}[\sum_{i=2}^{l}\{(\sum_{u=1}^{i-1} w_{ui}n_u)^2 + \sum_{u=1}^{i-1} w_{ui}^2 n_u(n_i+1)\}n_i$$
$$+\sum_{i=l}^{k-1}\{(\sum_{u=i+1}^{k} w_{ui}n_u)^2 + \sum_{u=i+1}^{k} w_{ui}^2 n_u(n_u+1)\}n_i$$
$$+2\sum\sum_{2\le s<t\le l}\{\sum_{u=1}^{s-1} w_{us}(w_{ut}-w_{st})n_u\}n_s n_t$$
$$+2\sum\sum_{l\le s<t\le k-1}\{\sum_{u=t+1}^{k} w_{ut}(w_{us}-w_{ts})n_u\}n_s n_u$$
$$+2\sum_{s=1}^{l-1}\sum_{t=l+1}^{k} w_{sl}w_{tl}n_s n_t n_l] \tag{2.5}$$

are the mean and variance, respectively, of A_{lw} when the F_i's are identical.

In order to improve the power property of A_{lw}, Park(1993) used constant weights $W_{ij}=l-1$ for $1\le i<j\le l$ and $w_{ji}=k-l$ for $1\le i<j\le k$, or weights $w_{ij}=|j-i|$ for all $i\ne j$ depending on the differences of sample numbers. Thus the test statistics proposed by Park(1993) are the weighted Mack-Wolfe statistics defined by the following two types.

$$A_{l1}=(l-1)\sum_{i=1}^{l-1}\sum_{j=i+1}^{l} U_{ij}+(k-l)\sum_{i=l}^{k-1}\sum_{j=i+1}^{k} U_{ji} \tag{2.6}$$

and

$$A_{l2}=\sum_{i=1}^{l-1}\sum_{j=i+1}^{l}(j-i)U_{ij}+\sum_{i=l}^{k-1}\sum_{j=i+1}^{k}(j-i)U_{ji}\,. \tag{2.7}$$

Then under H_0, the means and variances of A_{l1}, A_{l2}, when the F_i's are identical, are in the following way.

$$E_0(A_{l1})=\frac{1}{4}\{(l-1)(N_1^2-\sum_{i=1}^{l}n_i^2)+(k-l)(N_2^2-\sum_{i=l}^{k}n_i^2)\}, \tag{2.8}$$

$$E_0(A_{l2})=\frac{1}{2}\{\sum_{i=1}^{l-1}\sum_{j=i+1}^{l}(j-i)n_i n_j+\sum_{i=l}^{k-1}\sum_{j=i+1}^{k}(j-i)n_i n_j\}, \tag{2.9}$$

$$Var_0(A_{l1})=\frac{1}{72}[(l-1)^2\{2N_1^3+3N_1^2-\sum_{i=1}^{l}n_i^2(2n_i+3)\}$$
$$+(k-l)^2\{2N_2^3+3N_2^2-\sum_{i=l}^{k}n_i^2(2n_i+3)\}$$
$$+12(l-1)(k-l)n_l(N_1N_2-n_lN)] \tag{2.10}$$

and

$$Var_0(A_{l2})=\frac{1}{12}[\sum_{i=2}^{l}\{(\sum_{u=1}^{i-1}(i-u)n_u)^2+\sum_{u=1}^{i-1}(i-u)^2n_u(n_i+1)\}n_i$$

$$+\sum_{i=l}^{k-1}\{(\sum_{u=i+1}^{k}(u-i)n_u)^2+\sum_{u=i+1}^{k}(u-i)^2 n_u(n_i+1)\}n_i$$
$$+2\sum\sum_{2\leq s<t\leq l}\sum_{u=1}^{s-1}(s-u)^2 n_u n_s n_t$$
$$+2\sum\sum_{l\leq s<t\leq k-1}\sum_{u=t+1}^{k}(s-u)^2 n_u n_s n_t$$
$$+2\sum_{s=1}^{l-1}\sum_{t=l+1}^{k}(l-s)(t-l)n_s n_t n_l]. \tag{2.11}$$

Now to compare the asymptotic relative efficiency(ARE) between these tests, the notion of efficacy in Noether(1955) is used. Consider the translation alternatives defined by

$$H_{uN}:\ \theta_j=\begin{cases} j\theta & j=1,\cdots,l \\ (2l-j)\theta/\sqrt{N}, & j=l+1,\cdots,k \text{ with } \theta>0 \end{cases}$$

and efficacy of T_N defined by

$$eff(T_N)=\lim_{N\to\infty}\frac{\left\{\frac{d}{d\theta}E(T_N:\theta)|_{\theta=0}\right\}^2}{N\sigma_0^2(T_N)}.$$

According to Noether's(1955) theorem, the ARE of S_n relative to T_n, denoted by $ARE(S,T)$, can be evaluated by $ARE(S,T)=eff(S)/eff(T)$. After some algebraic calculations, we obtain the following efficacies, for $n_i=n\to\infty$, $i=1,\cdots,k$,

$$k\cdot eff(A_l)=\frac{I_2\cdot\{l^3-l+(k-l+1)^3-(k-l+1)\}^2}{l^3-l+(k-l+1)^3-(k-l+1)+6(l-1)(k-l)}, \tag{2.12}$$

$$k\cdot eff(A_{l1})=\frac{I_2\cdot[(l-1)(l^3-l)+(k-l)\{(k-l+1)^3-(k-l+1)\}]^2}{(l-1)^2(l^3-l)+(k-l)^2\{(k-l+1)^3-(k-l+1)\}+6(l-1)^2(k-l)^2}, \tag{2.13}$$

$$k\cdot eff(A_{l2})=\frac{I_2\cdot\{l^4-l^2+(k-l+1)^4-(k-l+1)^2\}^2}{l^3(l^2-l)+(k-l+1)^3\{(k-l+1)^2-(k-l+1)\}+6(l-1)l(k-l)(k-l+1)} \tag{2.14}$$

where $I_2=[\int f^2(x)\,dx]^2$. The $ARE(A_{lw},A_l)$'s under H_{uN} are tabulated in Table 1. Table 1 shows that the weighted Mack-Wolfe test statistics have greater(or equal) efficiency than the Mack-Wolfe test statistic for equally spaced umbrella patterns H_{uN}.

When the peak l is unknown, The alternative H_u can be viewed as a union of k- individual umbrella alternatives with the peak at group $1,\cdots,k$, respectively. Set $H_u=\cup_{t=1}^{k}H_{ut}$, where $H_{ut}:\ \theta_1\leq\cdots\leq\theta_t\geq\cdots\geq\theta_k$, with at least one strict inequality. Chen and

Table I : $ARE(\cdot, A_l)$ (equally spaced patterns)

	l	$k=3$	$k=4$	$k=5$	$k=6$	$k=7$	$k=8$
A_{l1}	1	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000
	2	1.0000	1.0800	1.1119	1.1037	1.0876	1.0729
	3			1.0000	1.0377	1.0760	1.0887
	4					1.0000	1.0191
A_{l2}	1	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000
	2	1.0000	1.0553	1.0857	1.0841	1.0739	1.0629
	3			1.0000	1.0295	1.0624	1.0758
	4					1.0000	1.0162

Wolfe(1990a) showed that the test $A^*_{\max}(=\max_{1\le t\le k} A^*_t)$ is more powerful than the unknown peak estimated test $A^*_{\hat{l}}$, where A^*_t is the standardized statistic of A_t with peak t-known and $\hat{l}$ is the sample estimate of unknown peak l from the observations. For more detailed properties of A^*_t, see Mack and Wolfe(1981). Park(1993) suggested to use $\max_{1\le t\le k} A^*_{tw}$. If $w_{ij}=1$ for $1\le i<j\le l$ and $w_{ji}=1$ for $l\le i<j\le k$, then the Park test $\max_{1\le t\le k} A^*_{tw}$ is equal to the Chen-Wolfe(1990a) test $A^*_{\max}$ which is a natural extension of the Mack-Wolfe test A_l for the unknown peak setting. For the case of unknown peak, the statistics proposed by Park(1993) are

$$A^*_{\max i} = \max_{i\le t\le k} A^*_{ti} \quad i=1,\ 2. \tag{2.15}$$

Under H_0, when the F_i's are identical, $A^*_{\max i}$, $i=1, 2$, have asymptotically the same distribution as that of the maximum of components of multivariate normal distribution.

3. Modifications of Weighted Mack-Wolfe Tests

When the underlying populations are symmetric, the problem considered in this paper is equivalent to testing the null hypothesis

$$H_0^* : \pi_{ij}=\frac{1}{2} \quad \text{for all pairs of } i \text{ and } j \tag{3.1}$$

against the umbrella alterntives

$$H_u^* : \pi_{ij} \geq \frac{1}{2},\ 1 \leq i < j \leq l \quad \text{and} \quad \pi_{ji} \geq \frac{1}{2},\ l \leq i < j \leq k, \tag{3.2}$$

for some l, with at least one strict inequality, where

$$\pi_{ij} = pr(X_{i1} < X_{j1}) = \int F_i dF_j,\ i \neq j = 1, \cdots, k$$

Under H_0^*, the expected values in (2.4), (2.8) and (2.9) remains the same. But, when the underlying populations have different shapes or scale parameters, the variance in (2.5), (2.10) and (2.11) are changed even under H_0^*.

Our objective is to develop a modification of the weighted Mack-Wolfe statistics A_{lw} that is distribution-free for testing the umbrella alternatives when the populations have the same shapes and asymptotically distribution free when the populations have different shapes and scale parameters. First we need to find the variance of A_{lw} when the populations have different shapes or scale parameters. Let

$$\pi_{ijt} = \int F_i F_j dF_t - (\int F_i dF_t)(\int F_j dF_t), \quad i, j, t = 1, \cdots, k$$

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

$$E(U_{ij}) = n_i n_j \pi_{ij}, \tag{3.3}$$

$$Var(U_{ij}) = n_i n_j \{(n_j - 1)\pi_{jji} + (n_i - 1)\pi_{iij} + \pi_{ij}\pi_{ji}\} \tag{3.4}$$

and

$$Cov(U_{ij}, U_{rs}) = \begin{cases} n_i n_j n_s \pi_{jsi} & \text{for } i = r,\ j \neq s, \\ n_i n_j n_r \pi_{irj} & \text{for } i \neq r,\ j = s, \\ -n_i n_j n_s \pi_{isj} & \text{for } j = r,\ i \neq s, \\ -n_i n_j n_r \pi_{jri} & \text{for } j \neq r,\ i = s, \\ 0 & \text{if } i, j, r, s \text{ are distinct.} \end{cases} \tag{3.5}$$

By using the results in (3.4) and (3.5) we obtain, after some straightforward computations, that, for $\ell = 1, \cdots, k$,

$$E(A_{lw}) = \sum_{i=1}^{l-1} \sum_{j=i+1}^{l} w_{ij} n_i n_j \pi_{ij} + \sum_{i=l}^{k-1} \sum_{j=i+1}^{k} w_{ji} n_i n_j \pi_{ji} \tag{3.6}$$

and

$$
\begin{aligned}
Var(A_{lw}) = & \sum_{i=1}^{l-1} \sum_{j=i+1}^{l} w_{ij}^2 n_i n_j \{(n_i-1)\pi_{iij} + (n_j-1)\pi_{jji} + \pi_{ij}\pi_{ji}\} \\
& + \sum_{i=l}^{k-1} \sum_{j=i+1}^{k} w_{ji}^2 n_i n_j \{(n_i-1)\pi_{iij} + (n_j-1)\pi_{jji} + \pi_{ij}\pi_{ji}\} \\
& + 2\sum_{i=1}^{l-2} \sum_{j=i+1}^{l-1} \sum_{s=j+1}^{l} n_i n_j n_s (w_{is}w_{js}\pi_{ijs} + w_{ij}w_{is}\pi_{jsi} - w_{ij}w_{js}\pi_{isj}) \\
& + 2\sum_{i=l}^{k-2} \sum_{j=i+1}^{k-1} \sum_{s=j+1}^{k} n_i n_j n_s (w_{si}w_{sj}\pi_{ijs} + w_{ji}w_{si}\pi_{jsi} - w_{ji}w_{sj}\pi_{isj}) \\
& + 2n_l \sum_{i=1}^{l-1} \sum_{j=l+1}^{k} n_i n_j w_{il} w_{jl} \pi_{ijl}. \qquad (3.7)
\end{aligned}
$$

Second, to find the consistent estimator of $N^{-\frac{3}{2}} Var(A_{lw})$, we use the suggestion of Fligner and Policello(1987). Fligner and Policello estimated the π_{ij}'s and π_{ijt}'s by replacing the populations F_i's with their sample distribution functions F_{n_i}. From Chen and Wolfe(1990b), for $i \neq j = 1, \cdots, k$, let

$$P_{ij}^{v} = n_i F_{n_i}(X_{jv}) = \sum_{u=1}^{n_i} \phi(X_{jv} - X_{iu}), \quad v = 1, \cdots, n_j$$

and

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

where

$$\phi(x) = \begin{cases} 1 & \text{for } x \geq 0 \\ 0 & \text{for } x < 0 \end{cases}$$

P_{ij}^{v} is the number of observation in the i-th sample that less than or equal to X_{jv} and $n_j \bar{P}_{ij}$ is the Mann-Whitney count between the i-th and j-th sample. We refer to P_{ij}^{v} as the placement of X_{jv} with respect to the i-th sample (see Orban and Wolfe(1982)). Thus we estimate π_{ij} and π_{ijt} by, respectively,

$$\hat{\pi}_{ij} = \bar{P}_{ij} / n_i = \sum_{u=1}^{n_i} \sum_{v=1}^{n_j} \phi(X_{jv} - X_{iu}) / n_i n_j$$

and

$$\hat{\pi}_{ijt} = \int F_{n_i} F_{n_j} dF_{n_t} - \left(\int F_{n_i} dF_{n_t}\right)\left(\int F_{n_j} dF_{n_t}\right)$$

$$= \frac{1}{n_i n_j n_t} \#(X_{tw} - X_{iu} \geq 0 \text{ and } X_{tw} - X_{jv} \geq 0) - \hat{\pi}_{it}\hat{\pi}_{jt}$$

$$= \frac{1}{n_i n_j n_t} \sum_{v=1}^{n_t} (P_{it}^v - \bar{P}_{it})(P_{jt}^v - \bar{P}_{jt}),$$

where $\#(a)$ is the number satisfying the condition a.

Third, estimate the exact variance $Var(A_{lw})$, by replacing the π_{ij}'s and π_{ijt}'s with the $\hat{\pi}_{ij}$'s and $\hat{\pi}_{ijt}$'s, respectively. Let

$$\zeta_{ijt} = \sum_{v=1}^{n_t} (P_{it}^v - \bar{P}_{it})(P_{jt}^v - \bar{P}_{it}) = n_i n_j n_t \hat{\pi}_{ijt}, \quad i, j, t = 1, \cdots, k$$

and replace the $(n_i - 1)$'s with n_i's. Thus we obtain the estimator of $Var(A_{lw})$ given by

$$\begin{aligned} \widehat{Var}(A_{lw}) = & \sum_{i=1}^{l-1} \sum_{j=i+1}^{l} w_{ij}^2 (\zeta_{iij} + \zeta_{jji} + \bar{P}_{ij}\bar{P}_{ji}) + \sum_{i=l}^{k-1} \sum_{j=i+1}^{k} w_{ji}^2 (\zeta_{iij} + \zeta_{jji} + \bar{P}_{ji}\bar{P}_{ij}) \\ & + 2\sum_{i=1}^{l-2} \sum_{j=i+1}^{l-1} \sum_{s=j+1}^{l} (w_{is}w_{js}\zeta_{ijs} + w_{ij}w_{is}\zeta_{jsi} - w_{ij}w_{js}\zeta_{isj}) \\ & + 2\sum_{i=l}^{k-2} \sum_{j=i+1}^{k-1} \sum_{s=j+1}^{k} (w_{si}w_{sj}\zeta_{ijs} + w_{ji}w_{si}\zeta_{jsi} - w_{ji}w_{sj}\zeta_{isj}) \\ & + 2\sum_{i=1}^{l-1} \sum_{j=l+1}^{k} w_{il}w_{jl}\zeta_{ijl}. \end{aligned} \tag{3.8}$$

At last, The estimators of $Var(A_{l1})$ and $Var(A_{l2})$ are also obtained by replacing the weights w_{ij}'s with the suitable forms.

Hence when the peak l of umbrella is known, the modifications of weighted Mack-Wolfe test statistics A_{lw} for testing the umbrella alternatives H_u^*, lead to rejecting H_0^* for large values of

$$\hat{A}_{lw}^* = \frac{A_{lw} - E_0(A_{lw})}{\{\widehat{Var}(A_{lw})\}^{\frac{1}{2}}}, \tag{3.9}$$

where A_{lw}, $E_0(A_{lw})$ and $\widehat{Var}(A_{lw})$ are given in (2.1), (2.4) and (3.8), respectively.

When the peak l of umbrella is unknown, The alternative H_u^* can be viewed as a union of k-individual umbrella alternatives with the peak at group $1, \cdots, k$, respectively. Set $H_u^* = \cup_{t=1}^{k} H_{ut}^*$, where $H_{ut}^* : \pi_{ij} \geq \frac{1}{2}$, $1 \leq i \leq j \leq t$ and $\pi_{ji} \geq \frac{1}{2}$, $t \leq i < j \leq k$, for some t, with at least one strict inequality. The natural extension of the weighted Mack-Wolfe tests similar to Chen-Wolfe(1990a), which is defined by

$$\hat{A}_{\max w}^* (= \max_{1 \leq t \leq k} \hat{A}_{tw}^*) \tag{3.10}$$

is used for testing the umbrella alternatives H_u^*.

Suppose that $N\to\infty$ in such a way that $n_i/N\to\lambda_i\in(0,\ 1)$, $i=1,\ \cdots,\ k$. We know that for $t=1,\ \cdots,\ k$, $(A_{tw}-E_0(A_{tw}))/\{Var(A_{tw})\}^{\frac{1}{2}}$ has an asymptotic ($N\to\infty$) null (H_0^*) distribution that is standard normal from Hettmansperger(1984). Also from Glivenko-Cantelli theorem (see Serfling(1980)), we know that $Var(A_{tw})/\widehat{Var}(A_{tw})$ converges to one almost surely as $N\to\infty$. This implies that the modified weighted Mack-Wolfe test statistics $\hat{A}_{lw}^*$(3.9) with peak l known, have a limiting standard normal distribution under H_0^*. Moreover, when the peak l is unknown, the modified weighted Mack-Wolfe test statistics $\hat{A}_{\max w}^*$(3.10) have asymptotically the same distribution as that of the maximum of components of multivariate normal distribution. Therefore, we observe that the modified Mack-Wolfe test statistics based on $\hat{A}_{tw}$ are asymptotically distribution-free under H_0^*. Since $\widehat{Var}(A_{tw})$ depends only on the ranks (or placements), the tests based on $\hat{A}_{tw}$ are exactly distribution-free when the populations have the same shapes (identical) and asymptotically distribution-free when the populations have different shapes or scale parameters.

4. Monte-Carlo Study

For both cases when the peak l is known and unknown, a Monte-Carlo study is performed to compare the modified weighted Mack-Wolfe tests ($\hat{A}_{li}^*$, $\hat{A}_{\max i}^*$, $i=1, 2$), the weighted Mack-Wolfe tests (A_{li}^*, $A_{\max i}^*$, $i=1, 2$), the Mack-Wolfe tests (A_l^*, $A_{\max}^*$) and the modified (natural extended) Mack-Wolfe tests ($\hat{A}_l^*$, $\hat{A}_{\max}^*$), where subindex $_{\max}$ denotes the test for the unknown peak umbrella alternative. For these simulations, only the case of equal sample sizes are considered. The selected three populations are normal, contaminated normal and Cauchy. The contaminated normal distribution is a mixture of the standard normal distribution and a normal distribution with mean zero and standard deviation 5 in proportion 0.9 and 0.1, respectively. These distributions can be described as the medium tailed and heavy tailed distributions. The normal and Cauchy random variates are generated from the IMSL subroutine RNNOR and RNCHY, respectively. The contaminated normal random variates are generated by using RNNOR and RNUN.

To study the effects that heteroscedasticity have on the significance levels, all the distributions mentioned above are centered to be symmetric about zero but have different scale parameters, namely, $F_i(x)=F(x/\sigma_i)$, $i=1,\ \cdots,\ k$, and $F(0)=\frac{1}{2}$. Several selections of $\sigma_2/\sigma_1,\ \cdots,\ \sigma_k/\sigma_1$ were studied with the three distributions mentioned above. Since the performance of tests with peak group $l=1,\ \cdots,\ k$ are similar for the levels, the tests corresponding to the peak k were used. The critical values by the normal approximtion are used when the peak l is known and the simulated critical values by some sort of interpolations are used when the peak l is unknown. The estmated levels for the significance level $\alpha=0.10$ are tabulated in Tables 2 and 3. To compare the powers for the various umbrella alternatives, the underlying populations having the same shapes were only considered. Several choices of $\theta_2-\theta_1, \cdots, \theta_k-\theta_1$ were studied with the three distributions mentioned above. The estimated powers are tabulated in Table 4 through Table 7.

Both the level and power studies were performed for $k=3$ and $k=4$ populations with equal observations $n_1=\cdots=n_k=10$. For each setting, 1,000 replications were used. The values in Tables were obtaind by counting the number which falls in the level-0.10 critical region. The sign + (—) in Tables 2 and 3 indicates that the estimated level is two or more standard deviations above (below) 0.10. The results from Tables 2 and 3 show that the tests based on the Mack-Wolfe (weighted) statistics A_l^* and $A_{\max}^*$ (A_{lw}^* and $A_{\max w}^*$) do not hold their nominal significance levels when the populations have different scale parameters, while the tests based on the modified (weighted) Mack-Wolfe statistics $\hat{A}_l^*$ and $\hat{A}_{\max}^*$($\hat{A}_{lw}^*$ and $\hat{A}_{\max w}^*$) hold their levels quite well across all the situations. These facts illustrate that the tests based on the modified statistics are exactly distribution-free when the populations have the same shapes. Also for large N, since the asymptotic level of the test based on A_l^* depends on the value of $Var_0(A_{..})/$

Table 2 : Estimated levels for nominal $\alpha=0.10$, when $k=3$ and $n_1=n_2=n_3=10$(l : known)

Distri.	σ_2/σ_1	σ_3/σ_1	A_k^*	A_{k1}^*	A_{k2}^*	$\hat{A}_k^*$	$\hat{A}_{k1}^*$	$\hat{A}_{k2}^*$
Normal	1	1	.110	.110	.110	.110	.110	.108
	1	2	.126+	.126+	.113	.100	.110	.103
	2	1	.054—	.054—	.062—	.082	.082	.094
	2	2	.092	.092	.094	.108	.108	.114
	1	3	.138+	.138+	.124+	.107	.107	.106
	3	1	.038—	.038—	.048—	.078—	.078—	.088
	3	3	.098	.098	.110	.116	.116	.125+
Contam Normal	1	1	.110	.110	.109	.103	.103	.110
	1	2	.131+	.131+	.125+	.115	.115	.119
	2	1	.062—	.062—	.077—	.098	.098	.109
	2	2	.089	.089	.094	.104	.104	.104
	1	3	.143+	.143+	.135+	.118	.118	.115
	3	1	.040—	.040—	.056—	.093	.093	.098
	3	3	.081—	.081—	.088	.091	.091	.096
Cauchy	1	1	.128+	.128+	.122+	.126+	.126+	.129+
	1	2	.110	.110	.104	.100	.100	.102
	2	1	.076—	.076—	.082	.100	.100	.106
	2	2	.082	.082	.086	.091	.091	.094
	1	3	.101	.101	.102	.093	.093	.093
	3	1	.065—	.065—	.080—	.101	.101	.105
	3	3	.092	.092	.098	.113	.113	.111

Table 3 : Estimated levels for nominal $\alpha=0.10$,
when $k=4$ and $n_1=n_2=n_3=n_4=10$ (l : unknown)

Distri.	σ_2/σ_1	σ_3/σ_1	σ_4/σ_1	A_m^*	A_{m1}^*	A_{m2}^*	$\hat{A}_m^*$	$\hat{A}_{m1}^*$	$\hat{A}_{m2}^*$
Normal	1	1	1	.087	.094	.094	.089	.094	.099
	1	1	2	.107	.110	.096	.089	.094	.088
	1	2	3	.137+	.139+	.132+	.106	.111	.116
	2	1	3	.110	.116	.113	.093	.105	.107
	3	3	1	.079−	.063−	.088	.111	.107	.120+
	1	3	5	.141+	.143+	.130+	.105	.104	.099
	5	3	3	.085	.070−	.089	.099	.101	.106
Contam Normal	1	1	1	.108	.111	.114	.118	.122+	.121+
	1	1	2	.117	.118	.108	.103	.106	.101
	1	2	3	.125+	.135+	.119+	.102	.105	.106
	2	1	3	.106	.106	.098	.089	.102	.100
	3	3	1	.075−	.062−	.086	.101	.097	.110
	1	3	5	.123+	.130+	.111	.097	.105	.098
	5	3	3	.090	.078−	.092	.104	.108	.105
Cauchy	1	1	1	.093	.089	.092	.102	.102	.103
	1	1	2	.102	.105	.095	.084	.088	.085
	1	2	3	.106	.105	.104	.102	.109	.109
	2	1	3	.112	.115	.113	.108	.118	.111
	3	3	1	.076−	.064−	.081−	.098	.092	.103
	1	3	5	.112	.119+	.111	.103	.101	.108
	5	3	3	.075−	.075−	.086	.099	.101	.106

Table 4 : Estimated powers for nominal $\alpha=0.10$,
when $k=3$ and $n_1=n_2=n_3=10$ (l : known)

Distri.	$\theta_2-\theta_1$	$\theta_3-\theta_1$	A_l^*	A_{l1}^*	A_{l2}^*	$\hat{A}_l^*$	$\hat{A}_{l1}^*$	$\hat{A}_{l2}^*$
Normal	0.0	1.0	.788	.788	.790	.794	.794	.811
	0.5	1.0	.819	.819	.824	.819	.819	.826
	1.0	2.0	.998	.998	.997	.998	.998	.998
	1.0	0.0	.877	.877	.877	.892	.892	.892
	1.0	0.5	.672	.672	.672	.703	.703	.703
	2.0	1.0	.984	.984	.984	.991	.991	.991
Contam Normal	0.0	1.0	.747	.747	.754	.750	.750	.758
	0.5	1.0	.776	.776	.780	.775	.775	.775
	1.0	2.0	.997	.997	.998	.998	.998	.997
	1.0	0.0	.816	.816	.816	.837	.837	.837
	1.0	0.5	.667	.667	.667	.695	.695	.695
	2.0	1.0	.977	.977	.977	.985	.985	.985
Cauchy	0.0	1.0	.432	.432	.426	.424	.424	.424
	0.5	1.0	.447	.447	.441	.441	.441	.449
	1.0	2.0	.811	.811	.801	.785	.785	.783
	1.0	0.0	.504	.504	.504	.533	.533	.533
	1.0	0.5	.372	.372	.372	.398	.398	.398
	2.0	1.0	.676	.676	.676	.699	.699	.699

Table 5 : Estimated powers for nominal $\alpha=0.10$,
when $k=4$ and $n_1=n_2=n_3=n_4=10$ (l : known)

Distri.	$\theta_2-\theta_1$	$\theta_3-\theta_1$	$\theta_4-\theta_1$	A^*_l	A^*_{l1}	A^*_{l2}	$\hat{A}^*_l$	$\hat{A}^*_{l1}$	$\hat{A}^*_{l2}$
Normal	0.0	0.0	1.0	.741	.741	.758	.758	.758	.780
	0.0	0.5	1.0	.821	.821	.834	.834	.834	.839
	0.5	1.0	1.0	.855	.855	.867	.861	.861	.865
	0.5	1.0	1.5	.982	.982	.984	.983	.983	.984
	0.0	1.0	0.0	.859	.838	.850	.870	.848	.857
	0.0	1.0	0.5	.761	.781	.781	.780	.790	.791
	0.5	1.0	0.5	.755	.781	.774	.765	.788	.782
	0.5	1.0	0.0	.869	.841	.855	.869	.844	.856
Contam Normal	0.0	0.0	1.0	.703	.703	.717	.726	.726	.740
	0.0	0.5	1.0	.810	.810	.819	.819	.819	.825
	0.5	1.0	1.0	.827	.827	.834	.833	.833	.834
	0.5	1.0	1.5	.974	.974	.976	.973	.973	.979
	0.0	1.0	0.0	.817	.800	.815	.832	.810	.827
	0.0	1.0	0.5	.716	.741	.742	.727	.764	.759
	0.5	1.0	0.5	.722	.747	.743	.731	.759	.752
	0.5	1.0	0.0	.825	.787	.806	.833	.797	.820
Cauchy	0.0	0.0	1.0	.417	.417	.432	.425	.425	.439
	0.0	0.5	1.0	.497	.497	.505	.500	.500	.495
	0.5	1.0	1.0	.456	.456	.461	.455	.455	.464
	0.5	1.0	1.5	.699	.699	.690	.688	.688	.686
	0.0	1.0	0.0	.505	.488	.502	.510	.493	.499
	0.0	1.0	0.5	.442	.459	.462	.457	.469	.469
	0.5	1.0	0.5	.403	.411	.408	.413	.424	.425
	0.5	1.0	0.0	.537	.516	.522	.539	.503	.521

Table 6 : Estimated powers for nominal $\alpha=0.10$,
when $k=3$ and $n_1=n_2=n_3=10$ (l : unknown)

Distri.	$\theta_2-\theta_1$	$\theta_3-\theta_1$	A^*_m	A^*_{m1}	A^*_{m2}	$\hat{A}^*_m$	$\hat{A}^*_{m1}$	$\hat{A}^*_{m2}$
Normal	0.0	1.0	.593	.606	.606	.613	.615	.619
	0.5	1.0	.618	.635	.627	.612	.612	.620
	1.0	2.0	.992	.993	.992	.988	.988	.989
	1.0	0.0	.712	.714	.715	.728	.728	.726
	1.0	0.5	.588	.595	.591	.620	.623	.619
	2.0	1.0	.979	.979	.980	.984	.984	.983
Contam Normal	0.0	1.0	.561	.582	.565	.556	.558	.556
	0.5	1.0	.582	.598	.589	.578	.579	.586
	1.0	2.0	.978	.981	.980	.977	.977	.977
	1.0	0.0	.676	.678	.676	.699	.699	.700
	1.0	0.5	.572	.582	.582	.605	.605	.603
	2.0	1.0	.951	.953	.956	.963	.963	.964
Cauchy	0.0	1.0	.229	.244	.239	.227	.227	.233
	0.5	1.0	.308	.328	.310	.289	.290	.290
	1.0	2.0	.659	.671	.647	.596	.599	.590
	1.0	0.0	.331	.338	.334	.343	.343	.350
	1.0	0.5	.274	.280	.282	.288	.290	.291
	2.0	1.0	.604	.614	.625	.622	.623	.626

Table 7 : Estimated powers for nominal $\alpha=0.10$,
when $k=4$ and $n_1=n_2=n_3=n_4=10$ (l : known)

Distri.	$\theta_2-\theta_1$	$\theta_3-\theta_1$	$\theta_4-\theta_1$	A^*_{m1}	A^*_{m2}	$\hat{A}^*_m$	$\hat{A}^*_{m1}$	$\hat{A}^*_{m2}$	
Normal	0.0	0.0	1.0	.476	.485	.489	.507	.505	.524
	0.0	0.5	1.0	.663	.674	.678	.672	.675	.683
	0.5	1.0	1.0	.749	.760	.753	.735	.742	.749
	0.5	1.0	1.5	.925	.931	.925	.924	.926	.932
	0.0	1.0	0.0	.703	.653	.687	.718	.677	.695
	0.0	1.0	0.5	.645	.642	.651	.664	.655	.664
	0.5	1.0	0.5	.627	.632	.635	.629	.629	.637
	0.5	1.0	0.0	.728	.700	.718	.725	.711	.725
Contam Normal	0.0	0.0	1.0	.441	.452	.449	.470	.470	.496
	0.0	0.5	1.0	.603	.609	.609	.610	.615	.625
	0.5	1.0	1.0	.701	.716	.710	.699	.699	.710
	0.5	1.0	1.5	.893	.896	.897	.882	.884	.896
	0.0	1.0	0.0	.645	.618	.641	.662	.637	.653
	0.0	1.0	0.5	.582	.582	.598	.605	.606	.615
	0.5	1.0	0.5	.576	.576	.580	.571	.576	.574
	0.5	1.0	0.0	.689	.668	.677	.682	.659	.669
Cauchy	0.0	0.0	1.0	.193	.200	.196	.197	.197	.206
	0.0	0.5	1.0	.279	.287	.281	.269	.270	.279
	0.5	1.0	1.0	.342	.347	.334	.322	.329	.331
	0.5	1.0	1.5	.450	.465	.447	.423	.438	.440
	0.0	1.0	0.0	.304	.282	.294	.312	.291	.309
	0.0	1.0	0.5	.278	.276	.286	.300	.290	.301
	0.5	1.0	0.5	.308	.302	.302	.297	.296	.306
	0.5	1.0	0.0	.337	.326	.340	.346	.335	.350

$Var(A_{..})$, it can be inflated (or deflated) according to the choices of $\sigma_2/\sigma_1, \cdots, \sigma_k/\sigma_1$ (Chen and Wolfe, 1990b).

The powers tabulated in Table 4 through Table 7 show that, for small-sample sizes, the differences between the estimated powers of the tests based on the modified statistics and those of the tests based on the original statistics are small. However, these small differences do not seem too high a price to pay for holding the levels over the various null hypothesis. The results also show that the tests based on (modified) weighted Mack-Wolfe statistics are more powerful than the tests based on (modified) Mack-Wolfe statistics for the cases excluding the location umbrella alternatives ($\theta_1, \theta_2, \theta_3, \theta_4$) = (0.0, 0.0, 1.0, 0.0) and (0.0, 0.5, 1.0, 0.0).

In conclusion, we recommend the tests based on the modified weighted Mack-Wolfe statistics for the following three reasons. First, they are exactly distribution-free when the populations are identical and the levels are controlled when the populations have different scale parameters. Second, for a small-sample sizes, the differences of the powers of the tests based on those two statistics are small-sample sizes, the differences of the powers of the tests based on those two statistics are small when the location parameters of the underlying populations are only different. Third, the modified weighted Mack-Wolfe tests are more powerful than the modified Mack-Wolfe tests for the various umbrella alternatives.

REFERENCES

1. Birnbaum, Z. W. and Klose, O. M. (1957). Bounds for the variance of the Mann-Whitney statistic. *The Annals of Mathematical Statistics*, 28, 933-945.
2. Chen, Y. I. and Wolfe, D. A. (1990a). A study of distribution-free tests for umbrella alternatives, *Biometrical Journal*, 32(1), 47-57.
3. Chen, Y. I. and Wolfe, D. A. (1990b). Modification of the Mack-Wolfe umberlla tests for a generalized Behrens-Fisher problem, *The Canadian Journal of Statistics*, 18, 245-253.
4. Fligner, M. A. and Policello, G. E. (1981). Robust rank procedures for the Behrens-Fisher problem, *Jounal of the American Statistical Association*, 76, 162-168.
5. Hettmansperger, T. P. (1984). *Statistical Inference Based on Ranks*, John Wiley & Sons, New York.
6. Jonckheere, A. R. (1954). A distribution-free k-sample test against ordered alternatives, *Biometrika*, 41, 133-145.
7. Mack, G. A. and Wolfe, D. A. (1981). K-Sample rank tests for umbrella alternatives, *Journal of the American Statistical Association*, 76, 175-181.
8. Noether, G. E. (1955). On a Theorem of Pitmann, *The Annals of Mathematical Statistics*, 26, 64-68.
9. Orban, J. and Wolfe, D. A. (1982). A class of distribution-free two-sample tests based on placements, *Journal of the American Statistical Association*, 77, 666-672.
10. Park, H. M. (1993). A study on distribution-free tests for k-sample location umbrella alternatives, *Ph. D. Dissertation, Department of Computer Science and Statistics, Seoul National University, Seoul, Korea.*
11. Ruster, S. W. and Fligner, M. A. (1984). A modification of the Kruskal-Wallis statistic for the generalized Behrens-Fisher problem, *Communications in Statistics-Theory and Methodology*, A13(16), 2013-2027.
12. Serfling, R. J. (1980). *Approximation Theorems of Mathematical Statistics*, John Wiley & Sons, New York.
13. Terpstra, T. J. (1952). The asymptotic normality and consistency of Ken-dall's test against trend, When ties are present, *Indagationes Mathematicae*, 14, 327-333.