Biom. J. **33** (1991) 3, 281–290 Akademie Verlag

Notes on the Mack-Wolfe and Chen-Wolfe Tests for Umbrella Alternatives

Y. I. Chen

Graduate Institute of Statistics, National Central University, ROC

Summary

In this paper we find a class of umbrella alternatives for which the Mack-Wolfe (1981) peak known test is optimal in the sense of Pitman efficiency. The asymptotic null distribution of the Chen-Wolfe (1989) statistic for the peak-unknown umbrella alternatives problem is obtained. Some percentiles of the asymptotic distribution computed by simulation are presented.

Key words: Asymptotic null distribution; Pitman efficiency; Ordered alternatives; Umbrella alternatives.

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 $F_i(x)=F(x-\vartheta_i)$, $i=1, \ldots, k$. In this paper we consider the problem of testing the null hypothesis $H_0: [\vartheta_1=\ldots=\vartheta_k]$ against the class of alternatives $H_A: [\vartheta_1 \leqq \ldots \leqq \vartheta_p \geqq \ldots \geqq \vartheta_k$, for some p, with at least one strict inequality]. Since the location parameters, ϑ's, have an up-down ordering, H_A is referred to as an umbrella alternative (see, for example, Mack and Wolfe 1981). The point p which separates the location parameters into the two different ordering groups is called the peak or point of the umbrella.

In a drug study, for instance, the investigator is usually interested in testing the equality of the effects of increasing dosage levels. Suppose the investigator believes that if the treatment effects are not identical, then, in general, the higher the dose of the drug applied, the better will be the resulting treatment effect. However, he is also aware that the subject may actually succumb to toxic effects at high doses, thereby decreasing the treatment effects. In this case, the prior information about the umbrella pattern treatment effects is available and then such umbrella alternatives are appropriate.

A variety of nonparametric tests have been developed for umbrella alternatives in the k-sample setting. See Chen and Wolfe (1989) for more detailed references. When the peak of the umbrella is known, Mack and Wolfe (1981) generalized the Jonckheere (1954)–Terpstra (1952) test for ordered alternatives and provided a solution to this problem. HETTMANSPERGER and NORTON (1987) suggested a linear rank test which has maximum Pitman efficacy when the peak-known umbrella alternative specifies equal spacings. In this article, the MACK-WOLFE (1981) peak known test is proved to have Pitman efficiency 1 relative to a class of linear rank tests for umbrella alternatives. We, therefore, find a class of umbrella alternatives for which the MACK-WOLFE (1981) peak known test is optimal. This also provides insight into the comparison of the HETTMANSPERGER-NORTON (1987) test and the MACK-WOLFE (1981) test for umbrella alternatives with peak known.

CHEN and WOLFE (1989) proposed an extension of the MACK-WOLFE (1981) peak known test to the unkown peak setting. However, no properties of the test were given. In this paper, to use the CHEN-WOLFE (1989) statistic when sample sizes are large, the asymptotic null distribution of the statistic is derived and the corresponding percentiles computed by simulation are presented.

2. Description of Previous Tests

Let R_{ij} be the rank of X_{ij} among the $N=\sum_{i=1}^{k} n_i$ observations and let $\bar{R}_i=\sum_{j=1}^{n_i} R_{ij}/n_i$ be the average rank of the ith sample. Set $\lambda_{Ni}=n_i/N$, $i=1, \ldots, k$. For testing an arbitrary peak-known (p) umbrella alternative H_A, MACK and WOLFE (1981) proposed to reject 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 corresponding to the number of observations in sample j that exceed observations in sample i, while HETTMANSPERGER and NORTON (1987) suggested to reject H_0 for large values of the statistic

$$V_p=\sum_{i=1}^{p} \lambda_{Ni}(i-\bar{c}_w)\,\bar{R}_i+\sum_{i=p+1}^{k} \lambda_{Ni}\,(2p-i-\bar{c}_w)\,\bar{R}_i\,, \tag{2.2}$$

where

$$\bar{c}_w=\sum_{i=1}^{p} i\lambda_{Ni}+\sum_{i=p+1}^{k} (2p-i)\,\lambda_{Ni}\,.$$

If the peak of the umbrella is unknown, CHEN and WOLFE (1989) viewed the alternative H_A as a union of k individual umbrella alternatives with the peak at groups $1, \ldots, k$, respectively, and obtained a natural extension of the Mack-Wolfe test based on A_p (2.1) to the unknown peak setting. This natural extension corresponds to rejecting H_0 for large values of

$$A^*_{\max}=\max\,(A^*_1, \ldots, A^*_k) \tag{2.3}$$

where $A_t^* = [A_t - E_0(A_t)]/[\mathrm{Var}_0\,(A_t)]^{1/2}$ and

$$E_0(A_t) = \left[N_t^2 + (N - N_{t-1})^2 - \sum_{i=1}^{k} n_i^2 - n_t^2\right]/4$$

and

$$\begin{aligned}\mathrm{Var}_0\,(A_t) = \Bigg\{ & 2\,[N_t^3 + (N - N_{t-1})^3] + 3\,[N_t^2 + (N - N_{t-1})^2] \\ & - \sum_{i=1}^{k} n_i^2\,(2n_i + 3) - n_t^2\,(2n_t + 3) + 12 n_t N_t \\ & \times (N - N_{t-1}) - 12 n_t^2 N \Bigg\}/72\,,\end{aligned}$$

with $N_t = \sum_{i=1}^{t} n_i$, are the mean and variance, respectively, of A_t, $t = 1, \ldots, k$.

3. A Class of Umbrella Alternatives for which the Mack-Wolfe Peak Known Test is Optimal

For each $i = 1, \ldots, k$, let $X_{i1}, \ldots, X_{in_i}$ be a random sample from a population with an absolutely continuous distribution function $F_i(x)$ and associated density $f_i(x)$. Assume that $N \to \infty$ in such a way that $\lambda_{Ni} \to \lambda_i$, with $0 < \lambda_i < 1$, $i = 1, \ldots, k$. In this section, we consider testing the null hypothesis

$$H_0 : [f_1(x) = \ldots = f_k(x)]$$

against the Pitman sequence of alternatives

$$\begin{aligned}H_{AN} : [f_i(x) = f\,(x - \vartheta_i/\sqrt{N}), \quad \text{for} \quad & \vartheta_1 \leqq \ldots \leqq \vartheta_p \geqq \ldots \geqq \vartheta_k \\ & \text{with at least one strict inequality}]\,,\end{aligned}$$

where f is the density function with finite Fisher information; that is,

$$I(f) = \int_{-\infty}^{\infty} \left[\frac{f'(x)}{f(x)}\right]^2 f(x)\,dx < \infty\,.$$

Denote a sequence of score functions by $a_N(i)$, $i = 1, \ldots, N$, where $a_N\,(1 + [uN]) \to \varphi(u)$, $0 < u < 1$, as $N \to \infty$, with $0 < \int_0^1 (\varphi(u) - \bar{\varphi})^2\,du < \infty$, $\bar{\varphi} = \int_0^1 \varphi(u)\,du$ and $[v]$ being the greatest integer equal to or less than v. Let L be a class of linear rank statistics

$$T = N^{-1/2} \sum_{i=1}^{k} \delta_{Ni} \sum_{j=1}^{n_i} a_N(R_{ij})\,, \tag{3.1}$$

where $\delta_{N1} \leqq \ldots \leqq \delta_{Np} \geqq \ldots \geqq \delta_{Nk}$. Following the results in section 7.2.1 of HAJÊK and SIDAK (1967), an asymptotically efficient test of H_0 versus H_{AN} among tests of the form (3.1) is obtained by setting $\delta_{Ni}=a\vartheta_i+b$, $a>0$, and

$$\varphi(u)_f=\varphi(u)=\frac{d\log(f(x))}{dx}\bigg|_{x=F^{-1}(u)}.$$

Furthermore, if $\delta_{Ni}\to\delta_i$ as $N\to\infty$, the Pitman efficiency of a test of the form (3.1) relative to the asymptotically efficient test (AET) is given by

$$\mathrm{ARE}\,(T,\mathrm{AET})=\frac{\left[\sum_{i=1}^{k}(\delta_i-\bar\delta)\,(\vartheta_i-\bar\vartheta)\right]^2}{\sum_{i=1}^{k}(\delta_i-\bar\delta)^2\sum_{i=1}^{k}(\vartheta_i-\bar\vartheta)^2}\times\frac{\left[\int_0^1(\varphi(u)-\bar\varphi)\,(\varphi_f(u)-\bar\varphi_f)\,du\right]^2}{\int_0^1(\varphi(u)-\bar\varphi)^2\,du\int_0^1(\varphi_f(u)-\bar\varphi_f)^2\,du},$$

where $\bar\delta=\sum_{i=1}^{k}\lambda_i\delta_i$, $\bar\vartheta=\sum_{i=1}^{k}\lambda_i\vartheta_i$ and $\varphi_f=\int_0^1\bar\varphi_f(u)\,du$. It was noted in Fairley and Fligner (1987) that the optimal test among tests of the form (3.1) can be obtained by setting $\delta_{Ni}=a\vartheta_i+b$, with $a>0$, $i=1,\ldots,k$, even though it may not be the asymptotically efficient test.

We consider in this section the subclass of L corresponding to $a_N(i)=i/N$, $i=1,\ldots,N$. The associated test statistics are

$$T_p=N^{-1/2}\sum_{i=1}^{k}\lambda_{Ni}\delta_{Ni}\bar R_i\,, \tag{3.2}$$

where $\delta_{N1}\leqq\ldots\leqq\delta_{Np}\geqq\ldots\geqq\delta_{Nk}$. Noted that when $f(\cdot)$ is the density of a logistic distribution, the optimal test of the form (3.2) is also the asymptotically efficient test in the larger class L.

It can be seen that the Hettmansperger-Norton (1987) test for a peak known umbrella alternative based on V_p (2.2) is of the form (3.2) and is optimal for equally spaced alternatives since

$$\begin{aligned}N^{-1/2}V_p&=N^{-1/2}\left\{\sum_{i=1}^{p}\lambda_{Ni}\,(i-\bar c_w)\,\bar R_i+\sum_{i=p+1}^{k}\lambda_{Ni}\,(2p-i-\bar c_w)\,\bar R_i\right\}\\&=N^{-1/2}\sum_{i=1}^{k}\lambda_{Ni}\delta_{Ni}R_i\,,\end{aligned}$$

where $\delta_{Ni}=(i-\bar c_w)$, $i=1,\ldots,p$, and $\delta_{Ni}=(2p-i-\bar c_w)$, $i=p+1,\ldots,k$. The Mack-Wolfe (1981) test based on the statistic A_p (2.1) is not of the form (3.2). However, from the results of Koziol and Reid (1977), we see that the statistic $N^{-3/2}A_p$ has the same Pitman efficiency as does the statistic

$$T'_p=N^{-1/2}\left[\sum_{i=1}^{p-1}\sum_{j=i+1}^{p}\lambda_{Ni}\lambda_{Nj}\,(\bar R_j-\bar R_i)+\sum_{i=p}^{k-1}\sum_{j=i+1}^{k}\lambda_{Ni}\lambda_{Nj}\,(\bar R_i-\bar R_j)\right]$$

$$= N^{-1/2}\left\{\left[\sum_{i=1}^{p-1}\lambda_{Ni}\sum_{j=i+1}^{p}\lambda_{Nj}\bar{R}_j-\sum_{i=1}^{p-1}\lambda_{Ni}\bar{R}_i\sum_{j=i+1}^{p}\lambda_{Nj}\right]\right.$$
$$\left.+\left[\sum_{i=p}^{k-1}\lambda_{Ni}\bar{R}_i\sum_{j=i+1}^{k}\lambda_{Nj}-\sum_{i=p}^{k-1}\lambda_{Ni}\sum_{j=i+1}^{k}\lambda_{Nj}\bar{R}_j\right]\right\}$$
$$= N^{-1/2}\left\{\left[\sum_{j=2}^{p}\lambda_{Nj}\bar{R}_j\sum_{i=1}^{j-1}\lambda_{Ni}-\sum_{i=1}^{p-1}\lambda_{Ni}\bar{R}_i\sum_{j=i+1}^{p}\lambda_{Nj}\right]\right.$$
$$\left.+\left[\sum_{i=p}^{k-1}\lambda_{Ni}\bar{R}_i\sum_{j=i+1}^{k}\lambda_{Nj}-\sum_{j=p+1}^{k}\lambda_{Nj}\bar{R}_j\sum_{i=p}^{j-1}\lambda_{Ni}\right]\right\}$$
$$= N^{-1/2}\left[\lambda_{N1}(\lambda_{N1}+2\Lambda_{N0}-\Lambda_{Np})\bar{R}_1+\sum_{i=2}^{p-1}\lambda_{Ni}(\lambda_{Ni}+2\Lambda_{N(i-1)}-\Lambda_{Np})\bar{R}_i\right.$$
$$\left.+\lambda_{Np}(1-\lambda_{Np})\bar{R}_p+\sum_{j=p+1}^{k}\lambda_{Nj}(1-\lambda_{Nj}-2\Lambda_{N(j-1)}+\Lambda_{N(p-1)}\bar{R}_j\right],$$

where $\Lambda_{N0}=0$ and $\Lambda_{Ni}=\lambda_{N1}+\ldots+\lambda_{Ni}$, $i=1, \ldots, k$. Therefore, we have the following theorem and corollary.

Theorem 3.1: Suppose that $\lambda_{Ni}\to\lambda_i$ as $N\to\infty$, with $0<\lambda_i<1$, $i=1, \ldots, k$. Then, among tests of the form (3.2), the Mack-Wolfe procedure based on A_p (2.1) is optimal for testing the sequence of alternatives H_{AN} with $\vartheta_i=c\delta^*_{Ni}+d$, $i=1, \ldots, k$, where $c>0$ and

$$\delta^*_{Ni}=\begin{cases}\lambda_{Ni}+2\Lambda_{N(i-1)}-\Lambda_{Np}; & \text{if} \quad 1\leqq i\leqq p-1\\ 1-\lambda_{Ni}; & \text{if} \quad i=p\\ 1-\lambda_{Ni}-2\Lambda_{N(i-1)}+\Lambda_{N(p-1)}; & \text{if} \quad p+1\leqq i\leqq k\,.\end{cases}$$

Corollary 3.2: Suppose that $\lambda_{Ni}\to 1/k$ as $N\to\infty$, $i=1, \ldots, k$. Then, among tests of the form (3.2), the test based on A_p is optimal for the alternatives H_{AN} with $\vartheta_i=c\delta'_i+d$, $i=1, \ldots, k$, where $c>0$ and

$$\delta'_i=\begin{cases}2i-p-1; & \text{if} \quad 1\leqq i\leqq p-1\\ k-1; & \text{if} \quad i=p\\ k+p-2i; & \text{if} \quad p+1\leqq i\leqq k\end{cases}$$

Note that the results obtained in this section are in agreement with the findings of the Monte Carlo power comparison between A_p and V_p, as discussed in CHEN and WOLFE (1989).

4. Asymptotic Null Distribution of the Chen-Wolfe statistic

Note that the rank-based statistic $A^*_{\max}$(2.3) proposed by CHEN and WOLFE (1989) is distribution-free under H_0. Therefore, the null distribution of the statistic can be computed by evaluating the statistic for every possible arrangement of the ranks. Since each of these arrangements is equally likely under H_0, to compute the null distribution one only needs to count the number of arrangements which lead to each value of the statistic. However, the number of arrangements becomes prohibitively large very rapidly as each of the sample sizes $n_1, \dots, n_k$ gets large. To use the Chen-Wolfe test based on $A^*_{\max}$ (2.3) when the sample sizes are large, the asymptotic null distribution is needed.

Theorem 4.1: Suppose that $\lambda_{Ni} \to \lambda_i$ as $N \to \infty$, with $0 < \lambda_i < 1$, $i = 1, 2, \dots, k$. Then, under H_0, the limiting $(N \to \infty)$ distribution of the random vector

$$\underset{\sim}{A}^*_{k-1} = (A^*_2, \dots, A^*_k)$$

is a $(k-1)$-variate normal distribution with zero mean vector.

Proof: Following the results of ARCHAMBAULT, MACK and WOLFE (1977) and MACK (1977), we obtain that, under H_0, any linear combination of the components of the random vector $\underset{\sim}{A}^*_{k-1}$ has a limiting $(N \to \infty)$ univariate normal distribution. It implies that $\underset{\sim}{A}^*_{k-1}$ has a limiting $(k-1)$-variate normal distribution with zero mean vector.

Theorem 4.2: Suppose that $\lambda_{Ni} \to \lambda_i$ as $N \to \infty$, with $0 < \lambda_i < 1$, $i = 1, \dots, k$. Then, under H_0, the limits $(N \to \infty)$ of the elements of the covariance matrix, $\sum$, corresponding to the random vector $\underset{\sim}{A}^*_{k-1} = (A^*_2, \dots, A^*_k)$ are given by

$$\lim_{N \to \infty} \operatorname{Var}_0 (A^*_t) = 1, \quad t = 2, \dots, k,$$

and

$$\lim_{N \to \infty} \operatorname{Cov}_0 (A^*_s, A^*_t) = 3a_k(s, t)\, [b_k(s)\, b_k(t)]^{-1/2}, \quad 2 \leqq s \leqq t \leqq k\,,$$

where

$$a_k(s, t) = \sum_{i=2}^{s} \lambda_i \Lambda_{i-1} \Lambda_i + \sum_{j=t}^{k-1} \lambda_j\, (1 - \Lambda_j)\, (1 - \Lambda_{j-1}) - \sum_{i=s}^{t} \lambda_i \Lambda_{i-1} \Lambda_i + \Lambda_{s-1}\, (\Lambda_t - \Lambda_{s-1})$$

and

$$b_k(s) = \Lambda^3_s + (1 - \Lambda_{s-1})^3 - \sum_{i=1}^{s} \lambda^3_i - \sum_{j=s}^{k} \lambda^3_j + 6\lambda_s \Lambda_{s-1}\, (1 - \Lambda_s)\,,$$

with $\Lambda_0 = 0$ and $\Lambda_v = \sum_{u=1}^{v} \lambda_u$, $v = 1, \dots, k$.

This theorem is obtained by deriving the relevant covariances and then finding their limiting values. For a proof, see the Appendix.

Theorem 4.3: Suppose that $\lambda_{Ni} \to \lambda_i$ as $N \to \infty$, with $0 < \lambda_i < 1$, $i = 1, 2, \dots, k$. Then, under H_0, the statistic

$$A^*_{\max} = \max\, (A^*_1, A^*_2, \dots, A^*_k)$$

converges in distribution to $\max\,(-W_k, W_2, \dots, W_k)$ as $N \to \infty$, where the random vector $(W_2, \dots, W_k)$ has a $(k-1)$-variate normal distribution with zero mean vector and covariance matrix, $\sum$, given in theorem 4.2.

Proof: This theorem follows from the fact that $A_1^* = -A_k^*$ and an application of the Cramer-Wold theorem (see MACK (1977)).

Corollary 4.4: Suppose that $\lambda_{Ni} \to \lambda_i$ as $N \to \infty$, with $0 < \lambda_i < 1$, $i = 1, 2$. Then, under H_0, the statistic

$$A_{\max}^* = \max(A_1^*, A_2^*)$$

converges in distribution to $|W|$, where the random variable W has a standard normal distribution.

When the limiting sample size proportions are all equal, some critical values for the test based on $A_{\max}^*$ were estimated by simulation and are presented in Table 1. For each value of k, the number of populations, the International Mathematical and Statistical Libraries (IMSL) routine RNMVN was used to generate appropriate $(k-1)$-variate normal random vectors for which the statistic $A_{\max}^*$ was evaluated. Proceeding in this fashion, we obtained an empirical cumulative distribution of $A_{\max}^*$ based on a sample of size 10,000 from the corresponding true distribution. The estimated critical values for the $A_{\max}^*$ test then correspond to the appropriate percentiles of this empirical distribution. When $k=4$, for example, the estimated 95th percentile for the asymptotic null distribution of $A_{\max}^*$ is 2.20.

We will now illustrate how to conduct the proposed approximate test based on $A_{\max}^*$ by giving a numerical example. In in vitro mutagenicity assays experimental organisms may succumb to toxic effects at high doses of the test agent, thereby reducing the number of organisms at risk of mutation and causing a downturn in the dose-response curve (MARGOLIN, KAPLAN and ZEIGER (1981)). The observations (fictional data is used) given in the following are the numbers of visible revertant colonies observed on plates, containing Salmonella bacteria of strain TA 98 and exposed to various doses of Acid Red 114.

Revertant Colonies for Acid Red 114, TA98, Hamster Liver Activation

Dose (μg/ml)				
0	100	333	1000	3333
24	67	78	82	44
22	59	43	58	33
17	27	98	45	28
19	23	37	50	21
35	54	36	60	30

We begin by calculating the Mack-Wolfe statistics A_t's.

$$A_1 = U_{21} + U_{31} + U_{41} + U_{51} + U_{32} + U_{42} + U_{52} + U_{43} + U_{53} + U_{54}$$
$$= 3 + 0 + 0 + 6 + 9 + 9 + 17 + 9 + 22 + 25 = 100\ ,$$
$$A_2 = U_{12} + U_{32} + U_{42} + U_{52} + U_{43} + U_{53} + U_{54}$$
$$= 22 + 9 + 9 + 17 + 9 + 22 + 25 = 113\ ,$$

$$A_3 = U_{12} + U_{13} + U_{23} + U_{43} + U_{53} + U_{54} = 22 + 25 + 16 + 9 + 22 + 25 = 119,$$

$$A_4 = U_{12} + U_{13} + U_{14} + U_{23} + U_{24} + U_{34} + U_{54}$$
$$= 22 + 25 + 25 + 16 + 16 + 16 + 25 = 145$$

and

$$A_5 = U_{12} + U_{13} + U_{14} + U_{15} + U_{23} + U_{24} + U_{25} + U_{34} + U_{35} + U_{45}$$
$$= 22 + 25 + 25 + 19 + 16 + 16 + 8 + 16 + 3 + 0 = 150 .$$

To obtain the value of $A^*_{\max}$, we need to find the $E_0(A_t)$'s and $\mathrm{Var}_0\,(A_t)$'s.

$$E_0(A_1) = [(5)^2 + (25)^2 - 5(5)^2 - (5)^2]/4 = 500/4 = 125$$

and

$$\mathrm{Var}_0\,(A_1) = \{2\,[(5)^3 + (25)^3] + 3\,[(5)^2 + (25)^2] - 5(5)^2\,(13) - (5)^2\,(13)$$
$$+ 12(5)\,(5)\,(25) - 12(5)^2\,(25)\}/72 = 31500/72 = 437.5\,,$$

$$E_0(A_2) = [(10)^2 + (20)^2 - 5(5)^2 - (5)^2]/4 = 350/4 = 87.5$$

and

$$\mathrm{Var}_0\,(A_2) = \{2\,[(10)^3 + (20)^3] + 3\,[(10)^2 + (20)^2] - 5(5)^2\,(13) - (5)^2\,(13)$$
$$+ 12(5)\,(10)\,(20) - 12(5)^2\,(25)\}/72 = 22050/72 = 306.25\,,$$

$$E_0(A_3) = [(15)^2 + (15)^2 - 5(5)^2 - (5)^2]/4 = 300/4 = 75$$

and

$$\mathrm{Var}_0\,(A_3) = \{2\,[(15)^3 + (15)^3] + 3\,[(15)^2 + (15)^2] - 5(5)^2\,(13) - (5)^2\,(13)$$
$$+ 12(5)\,(15)\,(15) - 12(5)^2\,(25)\}/72 = 12150/72 = 168.75\,.$$

Since $n_1 = \ldots = n_5 = 5$, we also have

$$E_0(A_4) = E_0(A_2) = 87.5 \quad \text{and} \quad \mathrm{Var}_0\,(A_4) = \mathrm{Var}_0\,(A_2) = 306.25$$

and

$$E_0(A_5) = E_0(A_1) = 125 \quad \text{and} \quad \mathrm{Var}_0\,(A_5) = \mathrm{Var}_0\,(A_1) = 437.5\,.$$

It implies that

$$A_1^* = [A_1 - E_0(A_1)]/[\mathrm{Var}_0\,(A_1)]^{1/2} = (100 - 125)/(437.5)^{1/2} = -1.20\,,$$
$$A_2^* = [A_2 - E_0(A_2)]/[\mathrm{Var}_0\,(A_2)]^{1/2} = (113 - 87.5)/(306.25)^{1/2} = 1.46\,,$$
$$A_3^* = [A_3 - E_0(A_3)]/[\mathrm{Var}_0\,(A_3)]^{1/2} = (119 - 75)/(168.75)^{1/2} = 3.39\,,$$
$$A_4^* = [A_4 - E_0(A_4)]/[\mathrm{Var}_0\,(A_4)]^{1/2} = (145 - 87.5)/(306.25)^{1/2} = 3.29$$

and

$$A_5^* = [A_5 - E_0(A_5)]/[\mathrm{Var}_0\,(A_5)]^{1/2} = (150 - 125)/(437.5)^{1/2} = 1.20\,.$$

Therefore, $A^*_{\max} = \max\,(A_1^*, \ldots, A_5^*) = \max\,(-1.20,\ 1.46,\ 3.39,\ 3.29,\ 1.20) = 3.39$.

With $k=5$ and equal sample size proportions, we know, from Table 1, that the estimated 99th percentile for the asymptotic null distribution of $A^*_{\max}$ is 2.89. Thus there is a significant evidence for an umbrella patterned dose-response curve.

Table 1
Estimated percentiles for the asymptotic null distribution $A^*_{\max}$ when $\lambda_1 = \dots = \lambda_k = 1/k$

Level	Number of populations (k)											
α	2	3	4	5	6	7	8	9	10	15	20	30
.10	1.65	1.82	1.91	1.97	2.02	2.06	2.09	2.13	2.13	2.20	2.20	2.21
.05	1.96	2.13	2.20	2.27	2.30	2.34	2.36	2.42	2.41	2.50	2.51	2.53
.01	2.58	2.70	2.80	2.89	2.94	2.93	2.94	2.97	3.00	3.07	3.08	3.13

Acknowledgement

This paper is part of the author's Ph. D. thesis written at the Ohio State University in USA under the supervision of Dr. DOUGLAS A. WOLFE. I would like to thank Dr. WOLFE for his helpful suggestions and Dr. MICHAEL A. FLIGNER for drawing my attention to the problem in section 3.

Appendix

Proof of Theorem 4.2: Let

$$Q_i = \sum_{u=1}^{i-1} U_{ui} \quad \text{and} \quad Q'_j = \sum_{v=j+1}^{k} U_{vj}$$

for $i=2, \dots, k$ and $j=1, \dots, k-1$. Then

$$A_t = \sum_{i=2}^{t} Q_i + \sum_{j=t}^{k-1} Q'_j, \quad t=2, \dots, k-1, \quad \text{and} \quad A_k = \sum_{i=2}^{k} Q_i .$$

Since the Q_i is the two-sampleMann-Whitney statistic computed between the ith sample and the preceding $i-1$ samples combined, we obtain

$$\mathrm{Var}_0\,(Q_i) = n_i N_{i-1}\,(N_i+1)/12 ,$$

where $N_i = \sum_{u=1}^{i} n_u$, $i=2, \dots, k$. Similarly, we have

$$\mathrm{Var}_0\,(Q'_j) = n_j\,(N-N_j)\,(N-N_{j-1}+1)/12 , \quad j=1, \dots, k-1 .$$

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

$$\mathrm{Cov}_0\,(Q_i, Q'_j) = \begin{cases} -n_i n_j\,(N+1)/12 & \text{for} \quad i>j \\ n_i N_{i-1}\,(N-N_i)/12 & \text{for} \quad i=j \\ 0 & \text{for} \quad i<j . \end{cases}$$

Furthermore, we see, directly from TERPSTRA (1952), that, if H_0 is true, the random variables $Q_2, \dots, Q_k$ are independent and the random rariables $Q'_1, \dots, Q'_{k-1}$ are also independent. Therefore, we obtain, after some straightforward computations, the following result:

$$\mathrm{Cov}_0(A_s, A_t) = \frac{1}{12}\Bigg[\sum_{i=2}^{s} n_i N_{i-1}\,(N_i+1) + \sum_{j=t}^{k-1} n_j\,(N-N_j)\,(N-N_{j-1}+1) - \sum_{i=s}^{t} n_i N_{i-1}\,(N_i+1) + N_{s-1}\,(N_t-N_{s-1})\,(N+1)\Bigg], \quad 2 \leq s < t \leq k .$$

Since

$$\mathrm{Cov}_0\,(A_s^*, A_t^*) = [\mathrm{Var}_0\,(A_s)\;\mathrm{Var}_0\,(A_t)]^{-1/2}\;\mathrm{Cov}_0\,(A_s, A_t)$$
$$= [N^{-3}\,\mathrm{Var}_0\,(A_s)\;N^{-3}\,\mathrm{Var}_0\,(A_t)]^{-1/2}\;(N^{-3}\,\mathrm{Cov}_0\,(A_s, A_t))\,,$$

the limits are easily obtained by substituting $\lambda_i N$ for n_i, $i=1, \ldots, k$, and replacing $N^{-3}\,\mathrm{Var}_0\,(A_s)$, $N^{-3}\,\mathrm{Var}_0\,(A_t)$ and $N^{-3}\,\mathrm{Cov}_0\,(A_s, A_t)$ by their limiting values.

References

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

Chen, Y. I. and Wolfe, D. A., 1989: A study of distribution free tests for umbrella alternatives. Biom. J. **32**, 47–57.

Fairley, D. and Fligner, M., 1987: Linear rank statistics for the ordered alternatives problem. Comm. Statist.-Theor. Meth. A**16**, 1–16.

Hajek, J. and Sidak, Z., 1967: *Theory of rank tests.* New York: Academic Press.

Hettmansperger, T. P. and Norton, R. M., 1987: Tests for patterned alternatives in *k*-sample problems. J. Amer. Statist. Assoc. **82**, 292–299.

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

Koziol, J. A. and Reid, N. 1977: On the asymptotic equivalence of two ranking methods for *k*-sample linear rank statistics. Ann. Statist. **5**, 1099–1106.

Mack, G. A., 1977: A *k*-sample Wilcoxon rank test for the umbrella alternatives. II. Point of the umbrella unknown. Doctoral Dissertation, The Ohio State University.

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

Margolin, B. H., Kaplan, N. and Zeiger, E., 1981: Statistical analysis of the Ames Salmonella/microsome test. Proceedings of the National Academy of Science, USA, **78**, 3779–3783.

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

Received Febr. 1990
Revised April 1990

Yuh-Ing Chen
Graduate Institute of Statistics
National Central University
Chung-Li, Taiwan 32054, R.O.C.