Statistica Sinica 8(1998), 785-800

AFFINE INVARIANT MULTIVARIATE RANK TESTS FOR SEVERAL SAMPLES

T. P. Hettmansperger, J. Möttönen and Hannu Oja

Pennsylvania State University, Tampere University of Technology and University of Jyväskylä

Abstract: Affine invariant analogues of the two-sample Mann-Whitney-Wilcoxon rank sum test and the c-sample Kruskal-Wallis test for the multivariate location model are introduced. The definition of a multivariate (centered) rank function in the development is based on the Oja criterion function. This work extends bivariate rank methods discussed by Brown and Hettmansperger (1987a,b) and multivariate sign methods by Hettmansperger and Oja (1994). The asymptotic distribution theory is developed to consider the Pitman asymptotic efficiencies and the theory is illustrated by an example.

Key words and phrases: Kruskal-Wallis test, multivariate rank test, Oja median, permutation test, Wilcoxon test.

1. Introduction

Let $\boldsymbol{x}_1, \ldots, \boldsymbol{x}_m$ and $\boldsymbol{x}_{m+1}, \ldots, \boldsymbol{x}_{m+n}$, $N = m+n$ be two independent samples from k-variate distributions with cumulative distribution functions $F(\boldsymbol{x}-\boldsymbol{\mu})$ and $F(\boldsymbol{x}-\boldsymbol{\mu}-\Delta)$, respectively. We assume that $F(\boldsymbol{x})$ is absolutely continuous with probability density function $f(\boldsymbol{x})$ and that the centre (the multivariate Oja median, for example) of F is $\mathbf{0}$. In this paper we develop a multivariate affine invariant two-sample rank test for testing $H_0: \ \Delta = \mathbf{0}$, a multivariate analogue of the Mann-Whitney-Wilcoxon rank sum test. The work extends the affine invariant bivariate rank tests proposed by Brown and Hettmansperger (1987a,b) and is related to the affine invariant multivariate sign tests by Hettmansperger, Nyblom and Oja (1994) and Hettmansperger and Oja (1994). The corresponding estimates are also discussed. Further, c-sample extensions are provided.

Underlying the development of sign and rank methods is the L_1 criterion. Note first that the c-sample problem is a special case of the general k-variate linear model case where $\boldsymbol{X}$ is an $N \times k$ response matrix with rows $\boldsymbol{x}_i^T$, $\boldsymbol{Z}$ is the $N \times p$ design matrix (p regressors) and $\boldsymbol{\beta}$ the $p \times k$ matrix of regression coefficients. The rows of the residual matrix $\boldsymbol{R} = \boldsymbol{X} - \boldsymbol{Z}\boldsymbol{\beta}$ are denoted by $\boldsymbol{r}_i^T$, i.e., $\boldsymbol{r}_i$ is the residual vector for the ith observation. For estimating the parameter matrix $\boldsymbol{\beta}$ and for constructing corresponding tests, Brown and Hettmansperger

(1987a) described three possible extensions of the L_1 criterion functions to the multivariate setting (D_1 for generalizing sign methods and D_2 for generalizing rank methods):

(1) The objective functions ("Manhattan distance")

$$D_1(\boldsymbol{\beta}) = \Sigma(|r_{i1}|+\cdots+|r_{ik}|) \quad \text{and} \quad D_2(\boldsymbol{\beta}) = \Sigma\Sigma(|r_{i1}-r_{j1}|+\cdots+|r_{ik}-r_{jk}|)$$

(2) the objective functions ("Euclidean distance")

$$D_1(\boldsymbol{\beta}) = \Sigma(r_{i1}^2+\cdots+r_{ik}^2)^{1/2} \text{ and } D_2(\boldsymbol{\beta}) = \Sigma\Sigma((r_{i1}-r_{j1})^2+\cdots+(r_{ik}-r_{jk})^2)^{1/2}$$

(3) the objective functions

$$D_1(\boldsymbol{\beta})=\Sigma_{i_1<\cdots<i_k}V(\mathbf{0},\boldsymbol{r}_{i_1},\ldots,\boldsymbol{r}_{i_k}) \text{ and } D_2(\boldsymbol{\beta})=\Sigma_{i_1<\cdots<i_{k+1}}V(\boldsymbol{r}_{i_1},\ldots,\boldsymbol{r}_{i_{k+1}}),$$

where V is the volume of the k-variate simplex with $k+1$ vertices given as arguments. Note that all these three pairs of objective functions reduce to the same L_1 criterion functions in one dimension. D_1 yields a family of sign tests and median-type estimates and D_2 a family of rank tests and Hodges-Lehmann type estimates with one- and two- sample sign, signed-rank and rank tests as special cases. In all three cases the objective functions can be rewritten as

$$D_1(\boldsymbol{\beta}) = \sum \boldsymbol{S}^T(\boldsymbol{r}_i)\boldsymbol{r}_i \text{ and } D_2(\boldsymbol{\beta}) = \sum \boldsymbol{R}_N^T(\boldsymbol{r}_i)\boldsymbol{r}_i,$$

where $\boldsymbol{S}(\boldsymbol{r})$ and $\boldsymbol{R}_N(\boldsymbol{r})$ are vector valued sign and centered rank functions, respectively.

In the first criterion function case (1), the sign and rank vectors obtained are just the vectors of componentwise signs and centered ranks. The score type tests are then combinations of the univariate componentwise tests and the estimates are the vectors of the componentwise univariate estimates. In the two-sample location case, for example, the obtained estimates of the shift parameter Δ (corresponding to D_1 and D_2) are the difference of marginal sample medians and the vector of marginal two-sample Hodges-Lehmann shift estimates. The methods are scale equivariant/invariant but unfortunately not rotation equivariant/invariant. Chakraborty and Chaudhuri (1996, 1998) utilized a transformation and retransformation approach to construct affine equivariant versions of the estimates. (For similar techniques to find invariant tests, see Dietz (1982) and Chaudhuri and Sengupta (1993)). The marginal efficiencies of the estimates agree with univariate efficiencies: In the multivariate normal case, for example, the efficiencies are .637 for the sign methods and .955 for the rank methods. The Pitman efficiencies of the tests and 'global efficiencies' of the estimate vectors (measured by the Wilks' generalized variance) may become really poor in the case of highly correlated components. (See Chakraborty and Chaudhuri (1998) for a discussion about the connection between affine equivariance and asymptotic efficiency. See also Bickel (1964, 1965) for efficiency properties and Puri and Sen (1971) for a detailed description of these methods.)

In the second case (2), the sign vector of $\boldsymbol{r}_i$ is the unit vector in the direction of $\boldsymbol{r}_i$ and, also, analogously with the univariate case, the centered rank for $\boldsymbol{r}_i$ is the sum of signs of $\boldsymbol{r}_i - \boldsymbol{r}_j,\ j = 1, \ldots, N$. In the one-sample location case the estimate is the well known spatial median and in the two-sample case the median-type estimate of the shift parameter Δ is the difference of the sample spatial medians. Brown (1983), Chaudhuri (1992) and Möttönen and Oja (1995), for example, discussed one-sample and multisample spatial tests and estimates. (For the general multivariate linear model case, see Rao (1988)). These methods are rotation equivariant/invariant but not scale equivariant/invariant. (For affine equivariant/invariant versions, see Rao (1988) and Chakraborty, Chaudhuri and Oja (1997)). Brown (1983) considered the efficiency of the spatial median (and the spatial sign test) and showed that in the multivariate spherical normal case the efficiency increases with the dimension, being, for example, .785 in 2 dimensions, .849 in 3 dimensions, .920 in 6 dimensions and going to 1 as the dimension increases. Chaudhuri (1992) gave the formulae for calculating the efficiency of the spatial HL-estimate in the multivariate spherical normal case, and using his formula we get .967 in 2 dimensions, .973 in 3 dimensions and .984 in 6 dimensions. Note that all these efficiencies dominate the corresponding univariate efficiency. (For the t-distribution case, see Möttönen, Oja and Tienari (1997)). Rescaling one of the components, i.e. moving from a spherical case to an elliptic case, may however highly reduce the efficiency (Brown (1983), Chakraborty, Chaudhuri and Oja (1997)). This is due to the lack of scale invariance property.

In this paper the third case (3) is considered. In the one-sample case, the median type estimate is called the Oja median (Oja (1983)). Brown and Hettmansperger (1987b, 1989) introduced the corresponding bivariate sign and rank tests and Hodges-Lehmann estimate. Unlike the methods above, these estimates/tests are automatically affine equivariant/invariant. Sign tests are as efficient as the spatial sign tests in spherical cases but strictly better in other elliptic cases (Oja and Niinimaa (1985), Niinimaa and Oja (1995)). (For a detailed description of the sign tests and their efficiency properties, see Hettmansperger, Nyblom, and Oja (1994) and Hettmansperger and Oja (1994).)

The tests listed above are not strictly, but only conditionally and asymptotically distribution-free. Randles (1989) introduced an affine invariant sign test, based on so-called interdirections, which has a distribution-free property over a broad class of distributions with elliptical directions. Later (conditionally and asymptotically distribution-free) extensions to the two-sample case (Randles (1992)) as well as corresponding one-sample (Peters and Randles (1990), Jan and Randles (1994)) and two-sample rank tests (Randles and Peters (1990)) were developed. Liu and Singh (1993) introduced a strictly distribution-free affine invariant multivariate two-sample rank test ranking the depths (Liu (1990)) of the original observations in an additional 'reference' sample.

Our plan is as follows. In Section 2 we introduce the multivariate centered rank function based on the Oja (1983) criterion function (the case (3) above) and discuss its properties. The multivariate analogue of the two-sample Mann-Whitney-Wilcoxon statistic with corresponding shift estimate is studied in Section 3. Asymptotic distribution theory is developed in Section 4 and multisample extensions are discussed in Section 5. We conclude with an example and some final remarks in Section 6.

2. Multivariate Centered Rank Function

We start with the definition of the multivariate Oja median (1983). Denote the volume of the k-variate simplex with $k+1$ vertices $\boldsymbol{x}_1, \ldots, \boldsymbol{x}_k, \boldsymbol{x}$ by

$$V(\boldsymbol{x}_1, \ldots, \boldsymbol{x}_k, \boldsymbol{x}) = \frac{1}{k!}\text{abs}\Big\{\det\Big(\begin{matrix} 1 & \cdots & 1 & 1 \\ \boldsymbol{x}_1 & \cdots & \boldsymbol{x}_k & \boldsymbol{x} \end{matrix}\Big)\Big\}.$$

Let $\boldsymbol{x}_1, \ldots, \boldsymbol{x}_N$ be a random sample from a k-variate distribution. The **multivariate Oja sample median** (Oja (1983)) $\hat{\boldsymbol{\mu}}$ is then the choice of $\boldsymbol{\mu}$ to minimize the objective function

$$D_N(\boldsymbol{\mu}) = \binom{N}{k}^{-1} \sum_{i_1 < \cdots < i_k} V(\boldsymbol{x}_{i_1}, \ldots, \boldsymbol{x}_{i_k}, \boldsymbol{\mu}).$$

To simplify the notation, let $P = \{p = (i_1, \ldots, i_k) \ : \ 1 \leq i_1 \ < \cdots < i_k \leq N\}$ be the set of $N_P = N!/[k!(N-k)!]$ different k-tuples of the index set $\{1, \ldots, N\}$. Index $p \in P$ then refers to a k-subset of the original observations. Note that p also refers to the **hyperplane** going through the k observations listed in p.

The volume of the simplex formed by $\boldsymbol{x}$ and k observations listed in p is then

$$V_p(\boldsymbol{x}) = \frac{1}{k!}\text{abs}\Big\{\det\Big(\begin{matrix} 1 & 1 & \cdots & 1 & 1 \\ \boldsymbol{x}_{i_1} & \boldsymbol{x}_{i_2} & \cdots & \boldsymbol{x}_{i_k} & \boldsymbol{x} \end{matrix}\Big)\Big\} = \frac{1}{k!}|d_{0p} + \boldsymbol{x}^T \boldsymbol{d}_p|,$$

where d_{0p} and d_{jp}, $j = 1, \ldots, k$, are the cofactors according to the last column of the matrix above. The values d_{0p} and $\boldsymbol{d}_p = (d_{1p}, \ldots, d_{kp})^T$ characterize the hyperplane p as follows. The vector $\boldsymbol{d}_p$ is normal to hyperplane p and we use the direction of $\boldsymbol{d}_p$ to define the positive or upper side of p. $V_p = [(k-1)!]^{-1}||\boldsymbol{d}_p||$ is the volume of the $(k-1)$-dimensional subsimplex determined by k observations listed in p. The indicator $S_p(\boldsymbol{x}) = \mathrm{sgn}(d_{0p} + \boldsymbol{x}^T \boldsymbol{d}_p)$ tells whether $\boldsymbol{x}$ is above or below the hyperplane p and $|d_{0p} + \boldsymbol{x}^T \boldsymbol{d}_p|/||\boldsymbol{d}_p||$ is the distance of point $\boldsymbol{x}$ from the plane p. The vector $S_p(\boldsymbol{x})\boldsymbol{d}_p$, normal to p and pointing towards $\boldsymbol{x}$ (if originating from plane p), is called the **sign of $\boldsymbol{x}$ with respect to $\boldsymbol{p}$**. Note that $S_p(\boldsymbol{0}) = \mathrm{sgn}(d_{0p})$ tells in which direction the origin is. (See also Hettmansperger, Möttönen, and Oja (1997, 1998).) Using this new notation, the objective function for the multivariate median is $D_N(\boldsymbol{\mu}) = N_P^{-1} \sum_p V_p(\boldsymbol{\mu})$.

Definition 2.1. The gradient of the objective function $k!\ D_N(\boldsymbol{\mu})$ with respect to $\boldsymbol{\mu}$ at $\boldsymbol{x}$,

$$\boldsymbol{R}_N(\boldsymbol{x}) = N_P^{-1} \sum_{p \in P} \boldsymbol{Q}_p(\boldsymbol{x}) = N_P^{-1} \sum_{p \in P} S_p(\boldsymbol{x}) \boldsymbol{d}_p,$$

is called the (empirical) vector valued **multivariate centered rank** of $\boldsymbol{x}$ with respect to the sample $\boldsymbol{x}_1, \ldots, \boldsymbol{x}_N$. If $\boldsymbol{x}_1, \ldots, \boldsymbol{x}_N$ is a random sample from a k-variate distribution with cdf F then the corresponding **theoretical centered rank** of $\boldsymbol{x}$ is the expected value $\boldsymbol{R}(\boldsymbol{x}) = E_F(\boldsymbol{R}_N(\boldsymbol{x})) = E_F(\boldsymbol{Q}_p(\boldsymbol{x}))$.

Analogously to the univariate case, $N_P \boldsymbol{R}_N(\boldsymbol{x})$ is nothing but the sum of signs of $\boldsymbol{x}$ w.r.t. observation hyperplanes p. In the univariate case, $R_N(x) = 2(F_N(x) - 1/2)$ where F_N is the sample cdf. Note that $\boldsymbol{R}_N(\boldsymbol{x})$ is piecewise constant with jumps $2N_P^{-1}\boldsymbol{d}_p$ when crossing the hyperplane p. The centered rank function $\boldsymbol{R}_N(\boldsymbol{x})$ is affine equivariant in the sense that if $\boldsymbol{R}_N^*(\boldsymbol{x})$ is the centered rank function for transformed observations $\boldsymbol{C}\boldsymbol{x}_1 + \boldsymbol{d}, \ldots, \boldsymbol{C}\boldsymbol{x}_N + \boldsymbol{d}$ then $\boldsymbol{R}_N^*(\boldsymbol{C}\boldsymbol{x} + \boldsymbol{d}) = \boldsymbol{C}^* \boldsymbol{R}_N(\boldsymbol{x})$ where $\boldsymbol{C}^* = \mathrm{abs}(\det(\boldsymbol{C}))(\boldsymbol{C}^{-1})^T$. If $\boldsymbol{C}$ is orthogonal then $\boldsymbol{C}^* = \boldsymbol{C}$. Consequently, the squared version of the rank test statistic introduced in Section 3 is affine invariant. Now we are ready to give the main results of this section. The proofs are postponed to the appendix.

Theorem 2.1. *The sum of the centered ranks of the observations is the zero vector, i.e.* $\sum_{i=1}^N \boldsymbol{R}_N(\boldsymbol{x}_i) = \boldsymbol{0}$.

The following result extends Theorem 1 in Brown and Hettmansperger (1987a). (Recall the discussion in the introduction.)

Theorem 2.2. $\sum_{i=1}^N \boldsymbol{R}_N^T(\boldsymbol{x}_i)\boldsymbol{x}_i = k \sum_{i_1 < \cdots < i_{k+1}} \{k!\ V(\boldsymbol{x}_{i_1}, \ldots, \boldsymbol{x}_{i_{k+1}})\}$.

3. Multivariate Two-Sample Rank Tests

Consider the multivariate two-sample location case, i.e., assume that $\boldsymbol{x}_1, \ldots, \boldsymbol{x}_m$ and $\boldsymbol{x}_{m+1}, \ldots, \boldsymbol{x}_N$, $N = m + n$, are two independent random samples from k-variate distributions with cumulative distribution functions $F(\boldsymbol{x} - \boldsymbol{\mu})$ and $F(\boldsymbol{x} - \boldsymbol{\mu} - \Delta)$ respectively. We wish to test the null hypothesis of no difference $H_0 : \Delta = \boldsymbol{0}$.

Consider first the test statistics of the form $\sum_{i=m+1}^{N}\boldsymbol{R}_N(\boldsymbol{x}_i)$ and $\sum_{i=m+1}^{N}\boldsymbol{R}_m(\boldsymbol{x}_i)$ where $\boldsymbol{R}_N$ is constructed using the combined sample and $\boldsymbol{R}_m$ using the first sample only. The second statistic is a placement type statistic (Orban and Wolfe (1982)) where the observations of the second sample are 'placed' or ranked among the observations in the first sample. A symmetrized version of the placement statistic is

$$\frac{m}{N}\sum_{j=m+1}^{N}\boldsymbol{R}_m(\boldsymbol{x}_j)-\frac{n}{N}\sum_{i=1}^{m}\boldsymbol{R}_n(\boldsymbol{x}_i),$$

where $\boldsymbol{R}_m$ and $\boldsymbol{R}_n$ are constructed using the first and second sample, respectively. In the univariate case all three statistics determine the same test but in the multivariate case the tests differ. The first version seems convenient when constructing tests based on the permutation argument and the second and third versions are useful when introducing the corresponding estimate, the two-sample Hodges-Lehmann shift estimate. Another version of the first statistic $\sum_{i=m+1}^{N}\boldsymbol{R}_N(\boldsymbol{x}_i)$, yielding the same conditional test but being more convenient for determining the permutational distribution, is given by the following

Definition 3.1. The two-sample rank test statistic for testing the null hypothesis $H_0:\ \Delta=\boldsymbol{0}$ is $\boldsymbol{T}_N=\sum_{i=1}^{N}a_i\boldsymbol{R}_N(\boldsymbol{x}_i)$ where

$$a_i=\begin{cases}-\lambda, i=1,\ldots,m\\ 1-\lambda, i=m+1,\ldots,N\end{cases}$$

with $\lambda=n/N$.

Consider now the conditional distribution of $\boldsymbol{T}_N$, conditioned on the observation set $\{\boldsymbol{x}_1,\ldots,\boldsymbol{x}_N\}$. If the null hypothesis is true then all the N observations come from the same distribution, and assigning ranks $\boldsymbol{R}_N(\boldsymbol{x}_i)$ to 'treatments' a_i is random. There are $\binom{N}{n}$ different permutations of $(a_1,\ldots,a_N)$, i.e. different random shufflings of the m $-\lambda$'s and n $(1-\lambda)$'s to ranks $\boldsymbol{R}_N(\boldsymbol{x}_1),\ldots,\boldsymbol{R}_N(\boldsymbol{x}_N)$. Hence, under the null hypothesis, these permutations are equiprobable and we get $E(a_i)=0$, $E(a_i^2)=\lambda(1-\lambda)$ and $E(a_ia_j)=-(N-1)^{-1}\lambda(1-\lambda)$ and consequently, conditionally, $E_0(\boldsymbol{T}_N)=\boldsymbol{0}$ and $\mathrm{Cov}_0(\boldsymbol{T}_N)=N\lambda(1-\lambda)\boldsymbol{B}_N$ where

$$B_N=\frac{1}{N-1}\sum\boldsymbol{R}_N(\boldsymbol{x}_i)\boldsymbol{R}_N^T(\boldsymbol{x}_i).$$

The approximate null distribution of the affine invariant multivariate two-sample rank test statistic $N^{-1/2}\boldsymbol{T}_N$ is k-variate normal with zero mean vector and covariance matrix $\lambda(1-\lambda)\boldsymbol{B}$ where $\boldsymbol{B}$ is the probability limit of $\boldsymbol{B}_N$ (under H_0) and $\lambda=\lim(n/N)$. The limiting distribution of $(N\lambda(1-\lambda))^{-1}\boldsymbol{T}_N^T\boldsymbol{B}_N^{-1}\boldsymbol{T}_N$ is then χ_k^2. (See Puri and Sen (1971).)

A natural two-sample shift estimate is obtained using the symmetrized placement test statistic. The estimate is then an Oja median computed on the linked pairwise differences of the second sample and first sample observations and is the extension of the univariate Hodges-Lehmann shift estimate. The bivariate extension was earlier given by Brown and Hettmansperger (1987a).

Definition 3.2. In the multivariate two-sample case, the Hodges-Lehmann shift estimate $\hat{\Delta}_N$ is the choice of Δ to minimize

$$\left[n\binom{m}{k}\right]^{-1} \sum_{j=m+1}^{N} \sum_{p \in P} V_p(\boldsymbol{x}_j - \Delta) + \left[m\binom{n}{k}\right]^{-1} \sum_{i=1}^{m} \sum_{p' \in P'} V_{p'}(\boldsymbol{x}_i + \Delta),$$

where p (p') goes over all different first (second) sample hyperplanes listed in P (P'). The estimating equations are

$$(1-\lambda) \sum_{j=m+1}^{N} \boldsymbol{R}_m(\boldsymbol{x}_j - \hat{\Delta}_N) - \lambda \sum_{i=1}^{m} \boldsymbol{R}_n(\boldsymbol{x}_i + \hat{\Delta}_N) = \boldsymbol{0}.$$

4. Limiting Distributions and Efficiency

In this section the asymptotic distribution of our test statistic $\boldsymbol{T}_N$ is found under the null hypothesis as well as under a sequence of contiguous alternatives. In this presentation the design variables from Definition 3.1 for sample size N, say $a_{N1}, \ldots, a_{NN}$, are fixed and $\lambda_N = n_N/N \to \lambda$ as $N \to \infty$.

We start the discussion by first considering the two-sample score test statistics of the general form

$$\boldsymbol{U}_N = \sum_{i=1}^{N} a_i \boldsymbol{r}_i = \sum_{i=1}^{N} a_i \boldsymbol{R}(\boldsymbol{x}_i)$$

for a fixed vector $(k \times 1)$ valued function $\boldsymbol{R}(\boldsymbol{x})$. (Later in our application $\boldsymbol{R}(\boldsymbol{x})$ is the **theoretical rank function**.) $\boldsymbol{R}$ is centered so that the expected value of $\boldsymbol{U}_N$ under the null hypothesis is the zero vector, i.e., $E_0(\boldsymbol{r}) = E_0(\boldsymbol{R}(\boldsymbol{x})) = \boldsymbol{0}$. It is well known that the asymptotically best choice for the score function $\boldsymbol{R}(\boldsymbol{x})$ is the optimal score $\boldsymbol{L}(\boldsymbol{x})$, i.e., the gradient vector of the logarithm of $f(\boldsymbol{x} - \boldsymbol{\mu})$ w.r.t. $\boldsymbol{\mu}$ at the origin. Let

$$\boldsymbol{V}_N = \sum_{i=1}^{N} a_i \boldsymbol{l}_i = \sum_{i=1}^{N} a_i \boldsymbol{L}(\boldsymbol{x}_i)$$

be this optimal test statistic. Note that $\boldsymbol{R}(\boldsymbol{x}) = \boldsymbol{x}$ gives a test which is asymptotically equivalent with the two-sample Hotelling's T^2 test and optimal under multinormality.

Consider now the limiting distribution of U_N under the following contiguous alternative sequences $\{H_N\}$: The first sample and the second sample come from distributions with k-variate density functions $f(\boldsymbol{x})$ and $f(\boldsymbol{x} - N^{-1/2}\boldsymbol{\delta})$, respectively. We assume that under the null hypothesis

$$\sum_{i=m+1}^{m+n} \{\ln f(\boldsymbol{x}_i - N^{-1/2}\boldsymbol{\delta}) - \ln f(\boldsymbol{x}_i)\} = N^{-1/2}\Big(\sum_{i=m+1}^{m+n} \boldsymbol{L}(\boldsymbol{x}_i)\Big)^T \boldsymbol{\delta} - \frac{\lambda}{2}\boldsymbol{\delta}^T \boldsymbol{I}_0 \delta + \boldsymbol{o}_p(1),$$

where $\boldsymbol{I}_0 = E_0(\boldsymbol{l}\boldsymbol{l}^T)$ is the expected Fisher information matrix for a single observation at $\boldsymbol{\mu} = \mathbf{0}$.

LeCam's Third Lemma (Hajek and Sidak (1967)) then gives

Theorem 4.1. *Under the alternative sequences* $\{H_N\}$,

$$N^{-1/2}\binom{\boldsymbol{U}_n}{\boldsymbol{V}_n} \to_D N_{2k}\Big(\lambda(1-\lambda)\binom{\boldsymbol{A}\boldsymbol{\delta}}{\boldsymbol{I}_0\boldsymbol{\delta}}, \qquad \lambda(1-\lambda)\begin{pmatrix} \boldsymbol{B} & \boldsymbol{A} \\ \boldsymbol{A}^T & \boldsymbol{I}_0 \end{pmatrix}\Big),$$

where $\boldsymbol{B} = E_0(\boldsymbol{r}\boldsymbol{r}^T)$, $\boldsymbol{A} = E_0(\boldsymbol{r}\boldsymbol{l}^T)$ and $\boldsymbol{I}_0 = E_0(\boldsymbol{l}\boldsymbol{l}^T)$.

Unfortunately, the affine invariant two-sample rank test statistic is not of the above form, since $\boldsymbol{T}_N = \sum a_i \boldsymbol{R}_N(\boldsymbol{x}_i)$, with empirical rank function $\boldsymbol{R}_N$. As in Brown, Hettmansperger, Nyblom and Oja (1992), Hettmansperger, Nyblom and Oja (1994) and Möttönen, Oja, and Tienari (1997) we first show that $\boldsymbol{R}_N$ is a uniformly weakly convergent estimate of the corresponding theoretical signed-rank function $\boldsymbol{R}(\boldsymbol{x})$. Test statistics $N^{-1/2}\boldsymbol{T}_N$ and $N^{-1/2}\boldsymbol{U}_N = N^{-1/2}\sum \boldsymbol{R}(\boldsymbol{x}_i)$ are then shown to be asymptotically equivalent with the same asymptotic properties and one can just apply the above formulae for $\boldsymbol{A}$ and $\boldsymbol{B}$ utilizing the "limit score function" $\boldsymbol{R}$. First we give the following two results.

Theorem 4.2. *Under the sequence of contiguous alternatives* $\{H_N\}$, $\sup_{\mathrm{X}} |\boldsymbol{R}_N(\boldsymbol{x}) - \boldsymbol{R}(\boldsymbol{x})| \to_P 0$.

Theorem 4.3. *Under the sequence of contiguous alternatives* $\{H_N\}$, $N^{-1/2}\boldsymbol{T}_N - N^{-1/2}\boldsymbol{U}_N = \boldsymbol{o}_P(1)$.

The main result concerning the limiting distribution of the affine invariant two-sample rank test statistic then follows from Theorems 4.1 and 4.3.

Theorem 4.4. *Under the sequence of contiguous alternatives* $\{H_N\}$, $N^{-1/2}\boldsymbol{T}_N$ *is asymptotically k-variate normal with mean vector* $\lambda(1-\lambda)\boldsymbol{A}\boldsymbol{\delta}$ *and covariance matrix* $\lambda(1-\lambda)\boldsymbol{B}$ *given in Theorem* 4.1.

Moreover, the limiting distribution of the test statistic $[N\lambda(1-\lambda)]^{-1}\boldsymbol{T}_N^T \boldsymbol{B}_N^{-1}\boldsymbol{T}_N$ under the sequence of contiguous hypotheses H_N is noncentral chi-square with k degrees of freedom and noncentrality parameter $\lambda(1-\lambda)\boldsymbol{\delta}^T\boldsymbol{A}^T\boldsymbol{B}^{-1}\boldsymbol{A}\boldsymbol{\delta}$.

Note also that $[\lambda(1-\lambda)]^{-1}\boldsymbol{A}^{-1}\boldsymbol{B}(\boldsymbol{A}^T)^{-1}$ is the asymptotic covariance matrix of the Hodges-Lehmann shift estimate $\hat{\Delta}$ given in Definition 3.2:

Theorem 4.5. *Assume that Δ_0 is the true shift parameter vector. Under mild regularity conditions, the limiting distribution of $N^{1/2}(\hat{\Delta}_N-\Delta_0)$ is k-variate normal with mean vector zero and covariance matrix $(\lambda(1-\lambda))^{-1}\boldsymbol{A}^{-1}\boldsymbol{B}(\boldsymbol{A}^T)^{-1}$.*

For the one-sample HL-estimate, see Hettmansperger, Möttönen, and Oja (1997). The efficiency factors for Hotelling's test, for the multivariate affine invariant sign test, for the multivariate spatial sign test and for the multivariate spatial rank test depend similarly on the inverses of the corresponding one-sample location estimates (mean vector, Oja median, affine invariant multivariate HL-estimate, spatial median and spatial HL-estimate).

Theorem 4.6. *The Pitman asymptotic relative efficiency of the multivariate affine invariant two-sample rank test with respect to Hotelling's two sample T^2 test is*

$$\frac{\boldsymbol{\delta}^T\boldsymbol{A}^T\boldsymbol{B}^{-1}\boldsymbol{A}\boldsymbol{\delta}}{\boldsymbol{\delta}^T\Sigma^{-1}\boldsymbol{\delta}},$$

$\boldsymbol{A}$ and $\boldsymbol{B}$ given in Theorem 4.1.

See Hettmansperger, Möttönen, and Oja (1997) and Möttönen, Hettmansperger, Oja and Tienari (1998) for thorough discussions about the efficiency of affine invariant rank methods as compared to Hotelling's tests and spatial sign and rank tests. In the k-variate normal case the efficiencies (e_k) with respect to the Hotelling's T^2 test are $e_1 = 0.955$, $e_2 = 0.937$, $e_3 = 0.934$, $e_6 = 0.947$ and $e_{10} = 0.961$, for example.

5. Multivariate c-Sample Rank Test

Let $\boldsymbol{x}_1,\ldots,\boldsymbol{x}_{n_1}$, $\boldsymbol{x}_{n_1+1},\ldots,\boldsymbol{x}_{n_1+n_2},\ldots$ and $\boldsymbol{x}_{n_1+\cdots n_{c-1}+1},\ldots,\boldsymbol{x}_{n_1+\cdots+n_c}$ be c independent samples from distributions $F(\boldsymbol{x}-\boldsymbol{\mu})$, $F(\boldsymbol{x}-\boldsymbol{\mu}-\Delta_2),\ldots,F(\boldsymbol{x}-\boldsymbol{\mu}-\Delta_c)$, correspondingly. Write $N=n_1+\cdots+n_c$ and $\lambda_j=n_j/N$ and $N_j=n_1+\cdots+n_j$, $j=1,\ldots,c$. We wish to test the null hypothesis that the distributions are the same, i.e., $H_0:\ \Delta_2=\cdots=\Delta_c=\mathbf{0}$.

Let $\boldsymbol{T}_{Nj}$ be the two-sample rank test statistic for testing whether the jth sample differs from the others, i.e., $\boldsymbol{T}_{Nj}=\sum_{i=1}^N a_{ji}\boldsymbol{R}_N(\boldsymbol{x}_i)$ where

$$a_{ji}=\begin{cases} 1-\lambda_j, & \text{if } i\in\{N_{j-1}+1, N_{j-1}+2,\ldots,N_j\}\\ -\lambda_j, & \text{otherwise}\end{cases}$$

with $\lambda_j=n_j/N$. Write also $H_j=[N\lambda_j(1-\lambda_j)]^{-1}\boldsymbol{T}_{Nj}^T\boldsymbol{B}_N^{-1}\boldsymbol{T}_{Nj}$. Under the null hypothesis, the limiting distribution of H_j is chisquare with k degrees of freedom. Finally, the combined test statistic (the Lawley-Hotelling trace statistic for testing $H_0 : \Delta_2 = \cdots = \Delta_c = \mathbf{0}$) $H = \sum_{j=1}^{c}(1-\lambda_j)H_j$ can be approximated by a χ^2-distribution with $k(c-1)$ degrees of freedom. In the univariate case, the test becomes the well known Kruskal-Wallis test for the one-way lay-out. The test is similar in structure to the combination of componentwise rank tests which is not affine invariant, however. (See Puri and Sen (1971).) Multivariate multisample spatial rank tests can be constructed similarly using two-sample spatial rank tests (Möttönen and Oja (1995)). For affine invariant multivariate multisample sign test statistics with a similar structure, see Hettmansperger and Oja (1994).

6. An Example

We include here a brief example illustrating the computation entailed by the use of the two-sample rank test and the estimation of shift. The data, carapace measurements for painted turtles, consists of two samples of size 10 taken from Table 6.5 of Johnson and Wichern (1988) (see Table 1 below).

Table 1. Carapace measurement data: Two independent samples

First sample			Second sample		
98	81	38	93	74	37
109	88	44	94	78	35
123	92	50	96	80	35
133	99	51	101	84	39
133	102	51	107	82	38
133	102	51	114	86	40
153	107	56	120	89	40
155	115	63	127	96	45
159	118	63	128	95	45
162	124	61	135	106	47

We treat observations as samples from trivariate distributions that differ at most in a shift. We first consider a test for H_0: $\Delta = \mathbf{0}$ where Δ is the trivariate shift vector separating the two distributions. To carry out the test, we compute $\boldsymbol{T}_N$ from Definition 3.1 along with $\boldsymbol{B}_N$. Then the test statistic is $[N\lambda(1-\lambda)]^{-1}\boldsymbol{T}_N^T\boldsymbol{B}_N^{-1}\boldsymbol{T}_N$. Then

$$\boldsymbol{T}_N = (84.4,\ 25.0,\ -340.7)^T,$$
$$\boldsymbol{B}_N = \begin{pmatrix} 1086.9 & -1277.6 & -654.0 \\ -1277.6 & 2610.5 & -844.3 \\ -654.0 & -844.3 & 2982.9 \end{pmatrix}$$

and finally $[N\lambda(1-\lambda)]^{-1}\boldsymbol{T}_N^T\boldsymbol{B}_N^{-1}\boldsymbol{T}_N = 13.91$. We see immediately that when compared to the percentiles of a chi-square distribution with 3 degrees of freedom that the asymptotic P-value is .003 and we easily reject the null hypothesis H_0: $\Delta=\mathbf{0}$. Note that it is possible to approximate the permutation P-value also. In this case, the two samples of size 10 are combined, shuffled, and the test statistic is recomputed. Based on 50,000 shuffles the permutation P-value is estimated to be .0005. Assuming multinormality, the popular Hotelling's T^2 gives the P-value .0003. The permutation test based on T^2 (using 50,000 shuffles) yields an estimated P-value .0002.

Since we reject $H_0 : \Delta = \mathbf{0}$ we next wish to estimate Δ. We apply Definition 3.2 and find the Hodges-Lehmann estimate to be $\hat{\Delta} = (-21.8, -14.1, -11.7)^T$. This can be compared to the vector of differences of component means $\bar{\mathrm{x}}_1 - \bar{\mathrm{x}}_2 = (-24.3, -15.8, -12.7)^T$. SAS/IML macros and S-PLUS functions to compute centered rank vectors, covariance matrix, test statistic, asymptotic P-value, and a simulated (or exact) permutation P-value are available on request from the authors.

Acknowledgements

The authors are thankful to an anonymous referee and an associate editor for careful reading and helful comments.

Appendix A: Proofs of the Theorems

Proof of Theorem 2.1. (a) First note that

$$\sum_i \boldsymbol{R}_N(\boldsymbol{x}_i) = N_P^{-1} \sum_{i_1<\cdots<i_{k+1}} \Big\{ \sum_{j=1}^{k+1} \mathrm{sgn}(d_{0p_j} + \boldsymbol{x}_{i_j}^T \boldsymbol{d}_{p_j})\boldsymbol{d}_{p_j} \Big\},$$

where $p_j = (i_1, \ldots, i_{j-1}, i_{j+1}, \ldots, i_{k+1})$. But the inner sum $\sum_{j=1}^{k+1} \mathrm{sgn}(d_{0p_j} + \boldsymbol{x}_{i_j}^T \boldsymbol{d}_{p_j})\boldsymbol{d}_{p_j}$ is just $k+1$ times the sum of centered ranks constructed for the subsample of $k+1$ observations $\boldsymbol{x}_{i_1}, \ldots, \boldsymbol{x}_{i_{k+1}}$. It is therefore enough to show that the proposition is true for $N = k+1$.

(b) Consider the sample of $N = k+1$ observations $\boldsymbol{x}_1, \ldots, \boldsymbol{x}_{k+1}$. The observations then determine a simplex. Write $p_i = (1, \ldots, i-1, i+1, \ldots, k+1)$, $i = 1, \ldots, k+1$. Now

$$\boldsymbol{R}_N(\boldsymbol{x}) = \frac{1}{k+1} \sum_{i=1}^{k+1} \mathrm{sgn}(d_{0p_i} + \boldsymbol{x}^T \boldsymbol{d}_{p_i})\boldsymbol{d}_{p_i}.$$

Let $\boldsymbol{x}_0$ be any fixed interior point of the simplex which means that $\boldsymbol{R}_N(\boldsymbol{x}_0) = \mathbf{0}$. But then also

$$\sum_{i=1}^{N} \boldsymbol{R}_N(\boldsymbol{x}_i) = \frac{1}{k+1} \sum_{i=1}^{k+1} \mathrm{sgn}(d_{0p_i} + \boldsymbol{x}_i^T \boldsymbol{d}_{p_i})\boldsymbol{d}_{p_i}$$

$$= \frac{1}{k+1} \sum_{i=1}^{k+1} \mathrm{sgn}(d_{0p_i} + \boldsymbol{x}_0^T \boldsymbol{d}_{p_i}) \boldsymbol{d}_{p_i} = \boldsymbol{0}$$

since $\boldsymbol{x}_0$ and $\boldsymbol{x}_i$ are always on the same side of p_i.

Proof of Theorem 2.2. As in the proof of Theorem 2.1, it is enough to consider the case $N = k+1$ only. Let $\boldsymbol{x}_1, \ldots, \boldsymbol{x}_{k+1}$ again be a sample of size $N = k+1$ and $p_i = (1, \ldots, i-1, i+1, \ldots, k+1)$, $i = 1, \ldots, k+1$. Clearly

$$\frac{1}{k!} \sum_{i=1}^{k+1} \mathrm{sgn}(d_{0p_i} + \boldsymbol{x}_i^T \boldsymbol{d}_{p_i})[d_{0p_i} + \boldsymbol{x}_i^T \boldsymbol{d}_{p_i}]$$
$$= \frac{1}{k!} \sum_{i=1}^{k+1} |d_{0p_i} + \boldsymbol{x}_i^T \boldsymbol{d}_{p_i}| = (k+1) V(\boldsymbol{x}_1, \ldots, \boldsymbol{x}_{k+1})$$

and

$$\begin{aligned}\frac{1}{k!} \sum_{i=1}^{k+1} \mathrm{sgn}(d_{0p_i} + \boldsymbol{x}_i^T \boldsymbol{d}_{p_i}) d_{0p_i} &= \frac{1}{k!} \mathrm{sgn}(d_{0p_1} + \boldsymbol{x}_1^T \boldsymbol{d}_{p_1}) \sum_{i=1}^{k+1} (-1)^{k+1} d_{0p_i} \\ &= \frac{1}{k!} \mathrm{sgn}\left(\det\begin{pmatrix} 1 & \cdots & 1 \\ \boldsymbol{x}_1 & \cdots & \boldsymbol{x}_{k+1} \end{pmatrix}\right) \cdot \det\begin{pmatrix} 1 & \cdots & 1 \\ \boldsymbol{x}_1 & \cdots & \boldsymbol{x}_{k+1} \end{pmatrix} \\ &= V(\boldsymbol{x}_1, \ldots, \boldsymbol{x}_{k+1})\end{aligned}$$

$(\mathrm{sgn}(d_{0p_{i+1}} + \boldsymbol{x}_{i+1}^T \boldsymbol{d}_{p_{i+1}}) = -\mathrm{sgn}(d_{0p_i} + \boldsymbol{x}_i^T \boldsymbol{d}_{p_i}),\ i = 1, \ldots, k)$, which gives the desired result

$$\frac{1}{k!} \sum_{i=1}^{k+1} \boldsymbol{R}_N^T(\boldsymbol{x}_i) \boldsymbol{x}_i = \frac{1}{k!} \sum_{i=1}^{k+1} \mathrm{sgn}(d_{0p_i} + \boldsymbol{x}_i^T \boldsymbol{d}_{p_i}) \boldsymbol{x}_i^T \boldsymbol{d}_{p_i} = k V(\boldsymbol{x}_1, \ldots, \boldsymbol{x}_{k+1}).$$

Proof of Theorem 4.2. Consider the extension of $\boldsymbol{R}_N(\boldsymbol{x})$ with the new domain of definition R^{k+1} defined by

$$\boldsymbol{R}_N^*(\boldsymbol{z}) = \frac{1}{N_P} \sum_{p \in P} \boldsymbol{S}_p^*(\boldsymbol{z}) \boldsymbol{d}_p, \quad \boldsymbol{z} \in R^{k+1},$$

where

$$\boldsymbol{S}_p^*(\boldsymbol{z}) = \mathrm{sgn}\{z_1 d_{0p} + z_2 d_{1p} + \cdots + z_{k+1} d_{kp}\}.$$

For each $\boldsymbol{z}$, $\boldsymbol{R}_N^*(\boldsymbol{z})$ is a U-statistic with expected value $\boldsymbol{R}^*(\boldsymbol{z}) = E(\boldsymbol{R}_N^*(\boldsymbol{z})) = E(\boldsymbol{S}_p^*(\boldsymbol{z}) \boldsymbol{d}_p)$. The functions $\boldsymbol{R}_N^*(\boldsymbol{z})$ and $\boldsymbol{R}^*(\boldsymbol{z})$ depend on $\boldsymbol{z}$ only through its direction $|\boldsymbol{z}|^{-1}\boldsymbol{z}$ and

$$R_N(\boldsymbol{x}) = \boldsymbol{R}_N^*\left(\begin{pmatrix} 1 \\ \boldsymbol{x} \end{pmatrix}\right) \text{ and } R(\boldsymbol{x}) = \boldsymbol{R}^*\left(\begin{pmatrix} 1 \\ \boldsymbol{x} \end{pmatrix}\right).$$

It is then obvious that

$$\sup_{\boldsymbol{x}\in R^k} |\boldsymbol{R}_N(\boldsymbol{x}) - \boldsymbol{R}(\boldsymbol{x})| \leq \sup_{\boldsymbol{z}\in R^{k+1}} |\boldsymbol{R}^*_N(\boldsymbol{z}) - \boldsymbol{R}^*(\boldsymbol{z})| = \sup_{|\boldsymbol{z}|\leq 1} |\boldsymbol{R}^*_N(\boldsymbol{z}) - \boldsymbol{R}^*(\boldsymbol{z})|.$$

Now use the construction described in Hettmansperger, Nyblom, and Oja (1994), Proof of Proposition 2: The symmetric kernel for $\boldsymbol{R}^*_N(\boldsymbol{z})$ is the sum of 2^{k+1} asymmetric but monotone (in z_i's) kernels $\boldsymbol{h}_b(\boldsymbol{x}_{i_1}, \ldots, \boldsymbol{x}_{i_k}; \boldsymbol{z}) = \boldsymbol{S}^*_p(\boldsymbol{z})(\pi^k_{i=0} I_{\{b_i d_{ip}>0\}})\boldsymbol{d}_p$ where $\boldsymbol{b} = (b_0, \ldots, b_k)$ goes through all possible 2^{k+1} $(\pm 1, \ldots, \pm 1)$.

Therefore $\boldsymbol{R}^*_N(\boldsymbol{z})$ can be represented as the sum of 2^{k+1} U-statistics with asymmetric kernels each being monotone in z_i's and each converging (under quite general assumptions) pointwise in probability to a continuous limit functions. Consequently, $\boldsymbol{R}^*_N(\boldsymbol{z})$ as a sum of uniformly convergent functions on the compact set $\{\boldsymbol{z} \in R^{k+1} : |\boldsymbol{z}| \leq 1\}$ converges uniformly in probability to the limit function $\boldsymbol{R}^*(\boldsymbol{z})$.

Proof of Theorem 4.3. It is enough to show that the proposition is true for the jth cordinate under the null hypothesis. First note that under the null hypothesis,

$$\text{Var}\,(R_{Nj}(\boldsymbol{x})) \leq \frac{k}{N} E(||\boldsymbol{d}_p||^2) \to 0, \text{ uniformly in } \boldsymbol{x}$$

and also

$$\text{Cov}(R_{Nj}(\boldsymbol{x}), R_{Nj}(\boldsymbol{y})) \leq \frac{k}{N} E(||\boldsymbol{d}_p||^2) \to 0, \text{ uniformly in } \boldsymbol{x} \text{ and } \boldsymbol{y}.$$

But this means that (under H_0) constants $v_{Nj} = \text{Var}\,(R_{Nj}(\boldsymbol{x}_i)) = E(\text{Var}\,(R_{Nj}(\boldsymbol{x}_i))|\boldsymbol{x}_i)$ and $c_{Nj} = \text{Cov}(R_{Nj}(\boldsymbol{x}_i), R_{Nj}(\boldsymbol{x}_k)) = E(\text{Cov}(R_{Nj}(\boldsymbol{x}_i), R_{Nj}(\boldsymbol{x}_k))|\boldsymbol{x}_i, \boldsymbol{x}_k)$ converge to zero as $N \to \infty$. Finally note that

$$\text{Var}\,\Big(\frac{1}{\sqrt{N}} \sum_{i=1}^N a_i \{R_{Nj}(\boldsymbol{x}_i) - R_j(\boldsymbol{x}_i)\}\Big) = \frac{v_{Nj}}{N} \sum_{i=1}^N a_i^2 + \frac{c_{Nj}}{N} \sum_{i\neq k} a_i a_k$$

which goes to zero since

$$\frac{1}{N} \sum a_i \to \lambda(1-\lambda) \text{ and} \frac{1}{N} \sum_{i\neq k} a_i a_k \to -\lambda(1-\lambda)$$

as $m, n \to \infty$ and $n/N \to \lambda$.

Proof of Theorem 4.5. First write

$$D_N(\Delta) = N^{-1}\Big[(1-\lambda)\binom{m}{k}^{-1} \sum_{j=m+1}^N \sum_{p\in P} V_p(\boldsymbol{x}_j - \Delta) + \lambda \binom{n}{k}^{-1} \sum_{i=1}^m \sum_{p'\in P'} V_p(\boldsymbol{x}_i + \Delta)\Big].$$

One can assume that the right shift parameter value is $\mathbf{0}$. Then under general assumptions, with $c = \frac{1}{2}$,

$$\begin{aligned} V_N(\Delta) &= N[D_N(N^{-1/2}\Delta) - D_N(\mathbf{0})] \\ &= -N^{-1/2}\Big[(1-\lambda)\sum_{j=m+1}^{N} \boldsymbol{R}_m(\boldsymbol{x}_j - cN^{-1/2}\Delta) \\ &\qquad -\lambda\sum_{i=1}^{m} \boldsymbol{R}_n(\boldsymbol{x}_i + cN^{-1/2}\Delta)\Big]^T \Delta + \boldsymbol{o}_p(1). \end{aligned}$$

As in Theorem 4.4 one can show that the limiting distribution of $V_N(\Delta)$ is the distribution of $V(\Delta) = -(\boldsymbol{y} - \frac{1}{2}\lambda(1-\lambda)\boldsymbol{A}\Delta)^T\Delta$ where $\boldsymbol{y}$ is $N_k(\mathbf{0}, \lambda(1-\lambda)\boldsymbol{B})$-distributed. As $V_N(\Delta)$ and $V(\Delta)$ are convex processes, $V_N(\cdot) \to_d V(\cdot)$ (Theorem 10.8 in Rockafellar (1970)). Lemma 2.2 in Davis, Knight, and Liu (1992) then gives the result. See also the proof of Theorem 4.4 in Hettmansperger, Möttönen and Oja (1998).

References

Bickel, P. J. (1964). On some alternative estimates for shift in the p-variate one sample problem. *Ann. Math. Statist.* **35**, 1079-1090.

Bickel, P. J. (1965). On some asymptotically non-parametric competitors of Hotelling's T^2. *Ann. Math. Statist.* **36**, 160-173.

Brown, B. M. (1983). Statistical uses of the spatial median. *J. Roy. Statist. Soc. Ser. B* **45**, 25-30.

Brown, B. M. and Hettmansperger, T. P. (1987a). Affine invariant rank methods in the bivariate location model. *J. Roy. Statist. Soc. Ser. B* **49**, 301-310.

Brown, B. M. and Hettmansperger, T. P. (1987b). Invariant tests in bivariate models and the L_1 criterion. In *Statistical Data Analysis Based on the L_1-norm and Related Methods* (Edited by Y. Dodge), 333-344. North Holland, Amsterdam, New York.

Brown, B. M. and Hettmansperger, T. P. (1989). An affine invariant bivariate version of the sign test. *J. Roy. Statist. Soc. Ser. B* **51**, 117-125.

Brown, B. M., Hettmansperger, T. P., Nyblom, J. and Oja, H. (1992). On certain bivariate sign tests and medians. *J. Amer. Statist. Assoc.* **87**, 127-135.

Chakraborty, B. and Chaudhuri, P. (1996). On a tranformation and retransformation technique for constructing affine equivariant multivariate median. *Proc. Amer. Math. Soc.* **124**, 2539-2547.

Chakraborty, B. and Chaudhuri, P. (1998). On an adaptive transformation-retransformation estimate of multivariate location. *J. Roy. Statist. Soc. Ser. B* **60**, 145-157.

Chakraborty, B., Chaudhuri, P. and Oja, H. (1998). Operating transformation retransformation on spatial median and angle test. *Statist. Sinica* **8**, 767-784.

Chaudhuri, P. (1992). Multivariate location estimation using extension of R-estimates through U-statistics type approach. *Ann. Statist.* **20**, 897-916.

Chaudhuri, P. and Sengupta, D. (1993). Sign tests in multidimension: Inference based on the geometry of the data cloud. *J. Amer. Statist. Assoc.* **88**, 1363-1370.

Davis, R. A., Knight, K. and Liu, J. (1992). M-estimation for autoregressions with infinite variance. *Stochastic Process. Appl.* **40**, 145-180.

Dietz, E. J. (1982). Bivariate nonparametric tests for the one-sample location problem. *J. Amer. Statist. Assoc.* **77**, 163-169.

Hajek, J. and Sidak, Z. (1967). *Theory of Rank Tests.* Academic Press, New York.

Hettmansperger, T. P., Möttönen, J. and Oja, H. (1997). Affine invariant multivariate one sample signed-rank tests. *J. Amer. Statist. Assoc.* **92**, 1591-1600.

Hettmansperger, T. P., Möttönen, J. and Oja, H. (1998). The geometry of the affine invariant multivariate sign and rank methods. Conditionally accepted in *J. Nonparamet. Statist.*

Hettmansperger, T. P. Nyblom, J. and Oja, H. (1994). Affine invariant multivariate one-sample sign tests. *J. Roy. Statist. Soc. Ser. B* **56**, 221-234.

Hettmansperger, T. P. and Oja, H. (1994). Affine invariant multivariate multisample sign tests. *J. Roy. Statist. Soc. Ser. B* **56**, 235-249.

Jan, S.-L. and Randles, R. H. (1994). A multivariate signed sum test for the one-sample location problem. *J. Nonparamet. Statist.* **4**, 49-63.

Johnson, R. A. and Wichern, D. W. (1988). *Applied Multivariate Statistical Analysis* (2nd edition). Prentice-Hall, Englewood Cliffs, N.J.

Liu, R. Y. (1990). On a notion of data depth based on random simplices. *Ann. Statist.* **18**, 405-414.

Liu, R. Y. and Singh, K. (1993). A quality index based on data depth and multivariate rank tests. *J. Amer. Statist. Assoc.* **88**, 252-260.

Möttönen, J., Hettmansperger, T. P., Oja, H. and Tienari, J. (1998). On the efficiency of affine invariant multivariate rank tests. To appear in *J. Multivariate Anal.*

Möttönen, J. and Oja, H. (1995). Multivariate spatial sign and rank tests. *J. Nonparamet. Statist.* **5**, 201-213.

Möttönen, J., Oja, H. and Tienari, J. (1997). On the efficiency of multivariate spatial sign and rank methods. *Ann. Statist.* **25**, 542-552.

Niinimaa, A. and Oja, H. (1995). On the influence functions of certain bivariate medians. *J. Roy. Statist. Soc. Ser. B* **57**, 565-574.

Oja, H. (1983). Descriptive statistics for multivariate distributions. *Statist. Probab. Lett.* **1**, 327-332.

Oja, H. and Niinimaa, A. (1985). Asymptotic properties of the generalized median in the case of multivariate normality. *J. Roy. Statist. Soc. Ser. B* **47**, 372-377.

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.

Peters, D. and Randles, R. H. (1990). A multivariate signed-rank test for the one-sample location problem. *J. Amer. Statist. Asoc.* **85**, 552-557

Puri, M. L. and Sen, P. K. (1971). *Nonparametric Methods in Multivariate Analysis.* Wiley, New York.

Randles, R. H. (1989). A distribution-free multivariate sign test based on interdirections. *J. Amer. Statist. Assoc.* **84**, 1045-1050.

Randles, R. H. (1992). A two sample extension of the multivariate interdirection sign test. In *L_1-Statistical Analysis and Related Methods* (Edited by Y. Dodge), 295-302. North-Holland, Amsterdam.

Randles, R. H. and Peters, D. (1990). Multivariate rank tests for the two-sample location problem. *Comm. Statist. Theory Methods* **19**, 4225-4238.

Rao, C. R. (1988). Methodology based on the L_1-norm in statistical inference. *Sankhyā Ser. A* **50**, 289-313.

Rockafellar, R. T. (1970). *Convex Analysis.* Princeton University Press, Princeton.

Department of Statistics, The Pennsylvania State University, University Park, PA 16802-1913, U.S.A.
E-mail: tph@stat.psu.edu

Signal Processing Laboratory, Tampere University of Technology, FIN-33101 Tampere, Finland.
E-mail: mottonen@cs.tut.fi

Department of Statistics, University of Jyväskylä, FIN-40351 Jyväskylä, Finland.
E-mail: ojahannu@jyu.fi

(Received November 1996; accepted June 1997)