*Nonparametric Tests for the Multisample Multivariate Location Problem**

SHOUTIR KISHORE CHATTERJEE AND

PRANAB KUMAR SEN[1], *University of Calcutta*

SUMMARY

In this paper, the univariate nonparametric analysis of variance tests by Kruskal and Wallis and by Brown and Mood have been extended to the general case of $p(>1)$ variates and $c(>2)$ samples. This has been accomplished through a conditional approach which makes the tests distribution-free under the null hypothesis. Various properties of the proposed tests have been studied and their asymptotic powers compared.

1. INTRODUCTION

During the past three decades the pace of development of nonparametric inference procedures has been tremendous. But, from the point of view of applications, this development has been confined mostly to univariate problems in the case of single

* Part of the work was completed with the partial support of the National Science Foundation Grant, GP-2593, at the Statistical Laboratory, University of California, Berkeley, while the second author was there.

[1] Now at The University of North Carolina at Chapel Hill, U.S.A.

as well as several samples and bivariate problems in the case of single sample. Multivariate problems have received comparatively much less attention; so that, at present there are few nonparametric contenders to standard methods of parametric multivariate analysis based on the assumption of multinormal parent distributions. Among the few multivariate nonparametric tests available, mention may be made of the bivariate sign-tests by Hodges (1955), Blumen (1958) and Bennett (1962, 1964); for a comparative study of these tests, the reader is referred to Chatterjee (1966). These tests actually relate to the single sample case. In the several sample case, a permutation test based on Hotelling's T^2-statistic was suggested by Wald and Wolfowitz (1944); but the test suffers from such shortcomings as are common to all tests based on permutations of values. In course of a series of lectures at the Calcutta University in 1962, S. N. Roy referred to a two sample bivariate location test developed by Roy, Bhapkar and Sathe; but the test is based on a step-down procedure in which the roles of the two variates do not appear to be symmetric.

In an earlier paper [7], the present authors considered two nonparametric tests for the bivariate two sample location problem, these being the generalizations of two well-known univariate two sample location tests, namely, Wilcoxon-Mann-Whitney rank-sum test and Mood's median test. Later on, the same principle has been used to formulate certain two sample nonparametric tests for testing the identity of association of two bivariate distributions. The object of the present investigation is to generalize the results of [7] to the general case of $p(\geqslant 1)$ variates and $c(\geqslant 2)$ samples, and develop nonparametric procedures for testing the identity of locations of several multivariate distributions. From one point of view, these results generalize the tests by Kruskal and Wallis (1952) and Brown and Mood (1950) to the multivariate case, and from another, they are the nonparametric analogues of the parametric tests by Roy (1942) and others [*see* Anderson (1958, Chapter 8)] for testing the equality of mean vectors of several multinormal distributions.

2. THE PROBLEM

Let $\boldsymbol{X}_\alpha^{(k)} = (X_{1\alpha}^{(k)}, \ldots, X_{p\alpha}^{(k)})$, $\alpha = 1, \ldots, n_k$ be n_k independent and identically distributed (vector valued) random variables (i.i.d.r.v.) having $p(\geqslant 1)$ variate continuous cumulative distribution functions (cdf) $F_k(\boldsymbol{x})$, $k = 1, \ldots, c$, and let the c samples be distributed independently. Let Ω be the set of all c-tuples of p variate continuous cdf's, and it is asumed that $\boldsymbol{F} = (F_1, \ldots, F_c)$ belongs to Ω. Later, some mild restrictions will have to be imposed on Ω, and these will be stated as occasions arise. On the basis of these c independent samples we desire to test the null hypothesis

$$H_0 : F_1(\boldsymbol{x}) = \ldots = F_c(\boldsymbol{x}), \text{ a. e.} \tag{2.1}$$

In testing this null hypothesis, we are particularly interested in detecting those alternatives which imply any difference in locations among the c distributions; the phrase *difference in locations* will be interpreted differently for the two tests to be considered here. For the rank-sum tests this will mean that for at least one of the p variates, the c sample observations are not all stochastically equal; the stochastic equality of two random variables X and Y being defined by the equality

$$P\{X > Y\} + \tfrac{1}{2}P\{X = Y\} = \tfrac{1}{2}. \tag{2.2}$$

Precisely, this means that for some $i(= 1, \ldots, p)$ there is at least one pair of (k, q) such that

$$P\{X_{i\alpha}^{(k)} > X_{i\beta}^{(q)}\} + \tfrac{1}{2}P\{X_{i\alpha}^{(k)} = X_{i\beta}^{(q)}\} \neq \tfrac{1}{2}, \ (k \neq q = 1, \ldots, c) \tag{2.3}$$

For the median test, let us denote by $\mu_i^{(k)}$ the median (assumed to be uniquely defined) of the random variable $X_{i\alpha}^{(k)}$, for $i = 1, \ldots, p$; $k = 1, \ldots, c$. Then the difference in location means that the following $pc(c-1)$ differences

$$\mu_i^{(k)} - \mu_i^{(q)}, \quad i = 1, \ldots, p; \quad k \neq q = 1, \ldots, c \tag{2.4}$$

are not all zero. There will be a special class of alternatives that will be of interest for both the rank-sum and median tests. This will be the class of *translation type* alternatives, and may be sketched as follows:

$$(2.5) \quad F_k(x) = F(x+\boldsymbol{\theta}^{(k)}),\ \boldsymbol{\theta}^{(k)} = (\theta_1^{(k)}, \ldots, \theta_p^{(k)}),\ k = 1, \ldots, c,$$

where $\theta_i^{(k)}$, $i = 1, \ldots, p$; $k = 1, \ldots, c$ are all real constants. Thus the c cdf's may be regarded to differ only by shifts in the locations. The proposed tests are valid for the types of alternatives considered above.

3. PRELIMINARY NOTIONS

Let us rank the N observations $X_{i\alpha}^{(k)}$, $\alpha = 1, \ldots, n_k$, $k = 1, \ldots, c$ on the ith variate in an increasing order of magnitude and let the rank of $X_{i\alpha}^{(k)}$ so obtained be denoted by $I_{i\alpha}^{(k)}$. Since the cdf's are all continuous, we can take the absence of ties for granted, in probability. The observation vector $\boldsymbol{X}_\alpha^{(k)}$ thus gives rise to the rank-vector

$$\boldsymbol{I}_\alpha^{(k)} = (I_{1\alpha}^{(k)}, \ldots, I_{p\alpha}^{(k)}),\ \alpha = 1, \ldots, n_k, k = 1, \ldots, c.$$

The N rank vectors corresponding to the N observations can be represented by the rank matrix (of order $p \times N$)

$$(3.1) \quad \boldsymbol{I}_N = \begin{pmatrix} I_{11}^{(1)} & \ldots & I_{1n_1}^{(1)} & \ldots & I_{11}^{(c)} & \ldots & I_{1n_c}^{(c)} \\ \cdot & \ldots & \cdot & \ldots & \cdot & \ldots & \cdot \\ I_{p1}^{(1)} & \ldots & I_{pn_1}^{(1)} & \ldots & I_{p1}^{(c)} & \ldots & I_{pn_c}^{(c)} \end{pmatrix}$$

Each row of this random matrix is a permutation of the numbers $1, \ldots, N$. Therefore the matrix $\boldsymbol{I}_N$ can have $(N!)^p$ possible realizations. We shall say that two rank matrices of the form (3.1) represent the same collection of rank vectors if one can be obtained from the other by a rearrangement of the columns. Specifically, the collection corresponding to (3.1) can be represented by permuting the columns so that the first row becomes $(1, \ldots, N)$. Let the ith row of the matrix so obtained be $(\lambda_{i1}, \ldots, \lambda_{iN})$, $i = 2, \ldots, p$. If we write

$$(3.2) \quad \lambda_{1\alpha} = \alpha, \quad \alpha = 1, \ldots, N,$$

the collection of rank vectors corresponding to (3.1) can be represented by the *collection matrix*

$$\mathbf{\Lambda}_N = \begin{pmatrix} \lambda_{11} & \cdots & \lambda_{1N} \\ \cdots & \cdots & \cdots \\ \lambda_{p1} & \cdots & \lambda_{pN} \end{pmatrix} \tag{3.3}$$

The matrix $\mathbf{\Lambda}_N$ being based on the random matrix $\boldsymbol{I}_N$ is itself random. As each row of $\mathbf{\Lambda}_N$, other than the first, can realize all the permutations of the numbers $1, \ldots, N$, $\mathbf{\Lambda}_N$ can have $(N!)^{p-1}$ possible realizations. We shall denote this set of possible realizations of $\mathbf{\Lambda}_N$ by $\mathscr{L}_N$. Typically, we shall write $\boldsymbol{L}_N$ for a realization of $\mathscr{L}_N$ where

$$\boldsymbol{L}_N = \begin{pmatrix} l_{11} & \cdots & l_{1N} \\ \cdots & \cdots & \cdots \\ l_{p1} & \cdots & l_{pN} \end{pmatrix} \tag{3.4}$$

$$l_{1\alpha} = \alpha, \qquad \alpha = 1, \ldots, N. \tag{3.5}$$

The probability distribution of $\mathbf{\Lambda}_N$ over $\mathscr{L}_N$ would, of course, depend on the distributions $F_1, \ldots, F_c$. However, given a particular realization $\boldsymbol{L}_N$ of $\mathbf{\Lambda}_N$, the conditional distribution of $\boldsymbol{I}_N$ over the $N!$ permutations of the columns of $\boldsymbol{L}_N$ would be uniform under H_0, whatever the common parent distribution may be. In this paper, we shall propose different functions of the elements of $\boldsymbol{I}_N$ as statistics for testing H_0. *From the observation made above, it follows that the conditional distribution of any such statistic given $\mathbf{\Lambda}_N = \boldsymbol{L}_N$ would be distribution free when H_0 is true.*

4. MULTIVARIATE MULTISAMPLE RANK-SUM TESTS

On the basis of the rank matrix $\boldsymbol{I}_N$ given by (3.1), let us find the mean ranks

$$\bar{I}^{(k)} = (1/n_k) \sum_{\alpha=1}^{n_k} I_{i\alpha}, \; i = 1, \ldots, p; \; k = 1, \ldots, c, \tag{4.1.1}$$

for the p variates in each of the c samples. We shall formulate a test for H_0 on the basis of these mean ranks. For this we shall derive first the expressions for the first and second order moments of the conditional joint distribution (under H_0) of the pc random variables defined in (4.1.1), and given $\mathbf{\Lambda}_N = \mathbf{L}_N$, where $\mathbf{\Lambda}_N$ is given by (3.2) and (3.3), and $\mathbf{L}_N \in \mathscr{L}_N$ is a particular realization of $\mathbf{\Lambda}_N$ given by (3.4) and (3.5).

Let the random variables $Z_\alpha^{(k)}$, $\alpha = 1, \ldots, N$; $k = 1, \ldots, c$, be defined as

$$(4.1.2) \qquad Z_\alpha^{(k)} = \begin{cases} 1, \text{ if } \alpha = I_{11}^{(k)}, \ldots, I_{1n_k}^{(k)} \\ 0, \text{ otherwise.} \end{cases}$$

We can then write

$$(4.1.3) \qquad \bar{I}_i^{(k)} = (1/n_k) \sum_{\alpha=1}^{n_k} Z_\alpha^{(k)} l_{i\alpha} \text{ for } i = 1, \ldots, p;\ k = 1, \ldots, c.$$

When H_0 is true

$$\begin{pmatrix} Z_1^{(1)} & \ldots & Z_N^{(1)} \\ . & \ldots & . \\ Z_1^{(c)} & \ldots & E_N^{(c)} \end{pmatrix}$$

would be a matrix obtained by permuting randomly the columns of

$$\begin{pmatrix} \overbrace{1 \ldots 1}^{n_1} & \overbrace{0 \ldots 0}^{n_2} & \ldots & \overbrace{0 \ldots 0}^{n_c} \\ 0 \ldots 0 & 1 \ldots 1 & \ldots & 0 \ldots 0 \\ . \ldots . & . \ldots . & \ldots & . \ldots . \\ 0 \ldots 0 & 0 \ldots 0 & \ldots & 1 \ldots 1 \end{pmatrix}$$

and this will be true independently of $\mathbf{\Lambda}_N$. Hence we obtain

$$(4.1.4) \qquad E(Z_\alpha^{(k)} \mid H_0) = n_k/N,$$

$$(4.1.5) \qquad \operatorname{Cov}(Z_\alpha^{(k)}, Z_\beta^{(q)} \mid H_0) = [n_k(\delta_{\alpha\beta}N-1)/N(N-1)][\delta_{kq}-n_q/N];$$

for α, $\beta = 1, \ldots, N$; k, $q = 1, \ldots, c$, where $\delta'_{\alpha\beta}$s and δ'_{kq}s are Kronecker deltas. Hence it is easy to show that

$$(4.1.6) \qquad E(\bar{I}_i^{(k)} \mid \mathbf{L}_N, H_0) = (N+1)/2, i = 1, \ldots, p, k = 1, \ldots, c;$$

$$(4.1.7) \qquad \operatorname{Cov}(\bar{I}_i^{(k)}, \bar{I}_j^{(q)} \mid \mathbf{L}_N, H_0) = \frac{(N+1)(\delta_{kq}N-n_k)}{12n_k} r_{ij}(\mathbf{L}_N),$$

for $i, j = 1, \ldots, p$; $k, q = 1, \ldots, c$, and where

$$(4.1.8) \quad r_{ij}(\boldsymbol{L}_N) = 12 \sum_{\alpha=1}^{N} \left(\ell_{i\alpha} - \frac{N+1}{2}\right) \left(\ell_{j\alpha} - \frac{N+1}{2}\right) \Big/ N(N^2-1),$$

for $i, j = 1, \ldots, p$, (it may be noted that $r_{ii}(\boldsymbol{L}_N) = 1$, for all $i = 1, \ldots, p$).

Now under the null hypothesis, the apportionment of the numbers $1, \ldots, N$ to the c sets $(I_{i1}^{(k)}, \ldots, I_{in_k}^{(k)})$, $k = 1, \ldots, c$ is likely to be equitable for each $i\,(= 1, \ldots, p)$, and hence we would expect that for each $i\,(= 1, \ldots, p)$, the mean ranks $\bar{I}_i^{(k)}$, $k = 1, \ldots, c$ would be close to each other and as a result to $(N+1)/2$. Since, only $p(c-1)$ of these mean ranks are linearly independent, it seems that we may base our test for H_0 on the set of $p(c-1)$ contrasts

$$(4.1.9) \qquad \left(\bar{I}_i^{(k)} - \frac{N+1}{2}\right), \; i = 1, \ldots, p; \; k = 1, \ldots, c-1.$$

Again, for practical convenience, it would be necessary to formulate the test on the basis of a single function of the elements in (4.1.9), and for this we are to choose a function that would reflect the numerical largeness of any of these contrasts. A positive definite quadratic form in these contrasts seems to be the most natural answer. Now, if we write

$$(4.1.10) \qquad \boldsymbol{R}(\boldsymbol{L}_N) = (r_{ij}(\boldsymbol{L}_N))\; i, j = 1, \ldots, p,$$

by (4.1.6), (4.1.7) and (4.1.8), the conditional dispersion matrix (under H_0) of the random variables in (4.1.9) (taken in that order), is readily deduced to be

$$(4.1.11) \qquad \frac{N+1}{12} \left(\frac{N}{n_k}\,\delta_{kq} - 1\right)_{k,\, q=1, \ldots, c-1} \otimes \boldsymbol{R}(\boldsymbol{L}_N),$$

where $\otimes$ stands for the Kronecker product [*see* Anderson (1958, p. 347)]. Also, it may be easily verified that

$$(4.1.12) \qquad \left(\frac{N}{n_k}\delta_{kq} - 1\right)^{-1}_{k,\, q=1, \ldots, c-1} = \left(\frac{n_k}{N}\,\delta_{kq} + \frac{n_k n_q}{n_c N}\right)_{k, q=1, \ldots, c-1}$$

If now $\boldsymbol{R}(\boldsymbol{L}_N)$ is assumed to be positive definite and its inverse matrix is denoted by

$$(4.1.13) \qquad \boldsymbol{R}^{-1}(\boldsymbol{L}_N) = (r^{ij}(\boldsymbol{L}_N)),$$

then the inverse of the dispersion matrix in (4.1.11) would be given by

$$(4.1.14) \qquad \frac{12}{N+1}\,(\delta_{kq}n_k/N + n_k n_q/n_c N)_{k,q=1,\ldots,c-1} \otimes \boldsymbol{R}^{-1}(\boldsymbol{L}_N).$$

Now we are in a position to formulate the test statistic. Whenever, the collection matrix $\boldsymbol{\Lambda}_N$ of the pooled sample is such that $\boldsymbol{R}(\boldsymbol{\Lambda}_N)$ is positive definite, we take as our test statistic

$$(4.1.15) \qquad W_N = \frac{12}{N+1} \sum_{k=1}^{c} \frac{n_k}{N} \sum_{i=1}^{p} \sum_{j=1}^{p} r^{ij}(\boldsymbol{\Lambda}_N) \left[\bar{T}_i^{(k)} - \frac{N+1}{2}\right]\left[\bar{T}_j^{(k)} - \frac{N+1}{2}\right],$$

which is a symmetric expression in $\{\bar{T}_i^{(k)}, i = 1, , p; k = 1, \ldots, c\}$. When $\boldsymbol{R}(\boldsymbol{\Lambda}_N)$ is not positive definite and is of rank $p' < p$, we may choose a subset of p' variables for which the rank correlation matrix would be positive definite, and write down for W_N an expression similar to (4.1.15) but involving only the p' variables chosen. However, as we shall see in Section 4.2, under some mild restriction on Ω, $\boldsymbol{R}(\boldsymbol{\Lambda}_N)$ will be positive definite, in probability.

From the remark made at the end of section 3, it follows that the conditional distribution of W_N given $\boldsymbol{\Lambda}_N = \boldsymbol{L}_N$ would be the same under H_0, whatever the cdf's $F_1 = \ldots = F_c = F$ may be. From this conditionally known distribution of W_N it is possible to construct the test function $\phi(W_N)$:

$$(4.1.16) \qquad \phi(W_N) = \begin{cases} 1, & \text{if } W_N > W_{N,\varepsilon}(\boldsymbol{\Lambda}_N), \\ A_{N,\varepsilon}(\boldsymbol{\Lambda}_N), & \text{if } W_N = W_{N,\varepsilon}(\boldsymbol{\Lambda}_N), \\ 0, & \text{otherwise,} \end{cases}$$

where the constants $W_{N,\varepsilon}(\mathbf{\Lambda}_N)$ and $A_{N,\varepsilon}(\mathbf{\Lambda}_N)$ (which may depend on $\mathbf{\Lambda}_N$,) are so chosen that

$$E\{\phi(W_N)\,|\,\boldsymbol{L}_N,\ H_0\} = \epsilon \tag{4.1.17}$$

The last equation implies

$$E\{\phi(W_N)\,|\,H_0\} = \epsilon. \tag{4.1.18}$$

Thus we can always construct a size ϵ test for H_0. In practice, the use of this exact test is forbidden (unless n and p are very small) because of the prohibitive amount of labor involved in the numerical evaluation of the permutation distribution of W_N. Therefore, we have to consider approximations to the exact permutation distribution of W_N that will be satisfactory at least for large samples. We discuss this in the following sub-section.

4.2. Large sample permutation test

Here we shall assume that N is adequately large, and for each N, there is a set $n_1, \ldots, n_c$ such that

$$\sum_{k=1}^{c} n_k = N,\ \lim_{N=\infty} n_k/N = \nu_k, \quad k = 1, \ldots, c; \tag{4.2.1}$$

where $\nu_1, \ldots, \nu_c$ are c fixed numbers lying in the open interval (0,1) and adding to unity.

Theorem 4.2.1. *If $\{\boldsymbol{L}_N : \boldsymbol{L}_N \epsilon\ \mathscr{L}_N\}$ is a sequence such that the corresponding sequence of matrices $\{\boldsymbol{R}(\boldsymbol{L}_N)\}$, defined by* (4.1.8) *and* (4.1.10), *has a positive definite limit, then the sequence of conditional distributions of W_N (defined by* (4.1.15), *) given $\mathbf{\Lambda}_N = \boldsymbol{L}_N$, converges to a chi-square distribution with $p(c-1)$ degrees of freedom.*

Proof. By virtue of (4.1.6), (4.1.7) and (4.1.15), it seems sufficient to show that the set of $p(c-1)$ contrasts in (4.1.9) has jointly (under the conditional probability measure induced by the N equally likely permutations of the columns of $\boldsymbol{L}_N$) asymptotically a multinormal distribution. For this it suffices to show that any nontrivial linear compound of the form

$$N^{\frac{1}{2}} \sum_{k=1}^{c} \sum_{i=1}^{p} t_i^{(k)}\, \bar{T}_i^{(k)} \quad \left(\sum_{k=1}^{c} t_i^{(k)}\, n_k = 0,\ i = 1, \ldots, p\right) \tag{4.2.2}$$

is asymptotically normal with zero mean and a nondegenerate variance. Using (4.1.2), (4.1.3) and some simple algebraic manipulations, we can rewrite (4.2.2) in the form

$$T_N = N^{\frac{1}{2}} \sum_{\alpha=1}^{N} \left(\sum_{i=1}^{p} t_{i\alpha}\, l_{i\beta_\alpha} \right), \tag{4.2.3}$$

where we define $l_{i\alpha}$ as in (3.4) and (3.5), and write

$$I_{1\alpha} = \beta_\alpha, \quad \alpha = 1, \ldots, N; \tag{4.2.4}$$

and where n_k of the $t_{i\alpha}$ have the common value $(t_i^{(k)}/n_k)$, for $k = 1, \ldots, c$, $i = 1, \ldots, p$. Thus writing

$$\sum_{i=1}^{p} t_{i\alpha}\, l_{i\alpha'} = b_{\alpha\alpha'}^{(N)}, \ \alpha, \alpha' = 1, \ldots, N \tag{4.2.5}$$

we have

$$N^{-\frac{1}{2}}\, T_N = \sum_{\alpha=1}^{N} b_{\alpha\beta_\alpha}^{(N)}$$

By (4.2.4), $(\beta_1, \ldots, \beta_N)$ represent a random permutation of the numbers $(1, \ldots, N)$. Therefore, we can use a combinatorial central limit theorem by Hoeffding (1951), and require only to show that on subtituting

$$d_{\alpha\alpha'}^{(N)} = \sum_{i=1}^{p} \left(t_{i\alpha} - \frac{1}{N} \sum_{\alpha=1}^{N} t_{i\alpha} \right) \left(l_{i\alpha'} - \frac{N+1}{2} \right), \tag{4.2.6}$$

the following condition is satisfied

$$\lim_{N=\infty} \frac{\max\limits_{1 \leq \alpha, \alpha' \leq N} \{d_{\alpha\alpha'}^{(N)}\}^2}{\frac{1}{N} \sum\limits_{\alpha=1}^{N} \sum\limits_{\alpha'=1}^{N} \{d_{\alpha\alpha'}^{(N)}\}^2} = 0, \tag{4.2.7}$$

Now, by (4.1.8) and (4.2.6), we have

$$\frac{1}{N} \sum_{\alpha, \alpha'=1}^{N} \{d_{\alpha\alpha'}^{(N)}\}^2 \tag{4.2.8}$$

$$= \frac{N^2-1}{12} \sum_{i,j=1}^{p} r_{ij}(L_N) \sum_{\alpha=1}^{N} \left(t_{i\alpha} - \frac{1}{N} \sum_{\alpha=1}^{N} t_{i\alpha} \right) \left(t_{j\alpha'} - \frac{1}{N} \sum_{\alpha=1}^{N} t_{j\alpha} \right).$$

Hence, by (4.2.1) we get after some simple adjustments that

$$\text{(4.2.9)} \qquad \lim_{N=\infty} \sum_{\alpha, \alpha'=1}^{N} \{d_{\alpha\alpha'}^{(N)}\}^2/N^2$$

$$= \frac{1}{12} \sum_{k=1}^{c} \nu_k \sum_{i,j=1}^{p} \left(\nu_k t_i^{(k)} - \sum_{q=1}^{c} t_i^{(q)}\right)\left(\nu_k t_j^{(k)} - \sum_{q=1}^{c} t_j^{(q)}\right) r_{ij}(\boldsymbol{L}_N).$$

As the matrix $\boldsymbol{R}(\boldsymbol{L}_N)$ is positive definite (by assumption, for N adequately large), the right hand side of (4.2.9) will be positive unless

$$\nu_k t_i^{(k)} - \sum_{q=1}^{c} t_i^{(q)} = 0 \text{ for all } i = 1, \ldots, p,\ k = 1, \ldots, c,$$

which holds only in the trivial case $t_i^{(k)} = 0$ for $i = 1, \ldots, p$, $k = 1, \ldots, c$. Thus, as $N \to \infty$

$$\text{(4.2.10)} \qquad \frac{1}{N} \sum_{\alpha, \alpha'=1}^{N} \{d_{\alpha\alpha'}^{(N)}\}^2 = O(N).$$

Again, from (4.2.6), as $\boldsymbol{L}_N \epsilon \mathscr{L}_N$,

$$|d_{\alpha\alpha'}^{(N)}| \leqslant \frac{N-1}{2} \sum_{i=1}^{b} \left| t_{i\alpha} - \frac{1}{N} \sum_{\alpha=1}^{N} t_{i\alpha} \right| \leqslant \frac{N-1}{2} \sum_{i=1}^{p} \left\{ \max_{1 \leqslant k \leqslant c} \left| \frac{t_i^{(k)}}{n_k} - \frac{1}{N} \sum_{k=1}^{c} t_i^{(k)} \right| \right\} \quad \text{for } \alpha, \alpha' = 1, \ldots, N,$$

and hence,

$$\text{(4.2.11)} \qquad \max_{1 \leqslant \alpha, \alpha' \leqslant N} |d_{\alpha\alpha'}^{(N)}| \leqslant \frac{N-1}{2N} \sum_{i=1}^{P} \left\{ \max_{1 \leqslant k \leqslant c} \left| \frac{N}{n_k} t_i^{(k)} - \sum_{k=1}^{c} t_i^{(k)} \right| \right\} = O(1)$$

as by (4.2.1), the right hand side of (4.2.11) converges to a finite quantity as $N \to \infty$. (4.2.10) and (4.2.11) together imply that the sets of numbers (4.2.5) satisfy the condition (4.2.10). The rest of the proof of the theorem follows by routine methods and hence is omitted.

We need to prove a lemma before the main theorem of this section is proved. We shall write

$$\bar{F}(x_1, \ldots, x_p) = \sum_{k=1}^{c} \nu_k F_k(x_1, \ldots, x_p) \tag{4.2.12}$$

and denote by $(\eta_1, \ldots, \eta_p)$ a set of random variables following the distribution law (4.2.12) $\bar{\psi}_{ij}$ will stand for the grade correlation coefficient between η_i and η_j, and

$$\bar{\Psi} = (\bar{\psi}_{ij})_{i, j=1, \cdots, p}. \tag{4.2.13}$$

will stand for the grade correlation matrix of $(\eta_1, \ldots, \eta_p)$.

Here as well as later, we shall use the following notations for the univariate marginal cdf's associated with the distributions F_k, $k = 1, \ldots, c$ and $\bar{F}$. Let $F_{k[i]}(x_i)$ and $\bar{F}_{[i]}(x_i)$ stand for the marginal cdf of the ith variate $X_i^{(k)}$ and η_i respectively, for $i = 1, \ldots, p$ and $k = 1, \ldots, c$. In terms of these notations $\bar{\psi}_{ij}$ is explicitly given by

$$\psi_{ij} = 3 \int_{E^p} \{2\bar{F}_{[i]}(x_i) - 1\}\{2\bar{F}_{[i]}(x_j) - 1\} d\bar{F}(x_1, \ldots, x_p), \tag{4.2.14}$$

$$i, j = 1, \ldots, p,$$

where E^p stands for the p-dimensional real (Euclidean) space.

Lemma 4.2.2. *As $N \to \infty$, $\boldsymbol{R}(\boldsymbol{\Lambda}_N)$ converges in probability to $\bar{\boldsymbol{\Psi}}$.*

The proof follows as a simple extension of a similar lemma (in the particular case of $p = c = 2$) by the authors ([7], pp. 29–31) and hence is not reproduced here.

For the further development of the theory, we shall have to assume that Ω satisfies a mild restriction. This we state below :

Condition A. *Ω is such that for all points $(F_1, \ldots, F_c) \in \Omega$ and for all ν_k, $k = 1, \ldots, c$, $0 < \nu_k < 1$, $\sum_{k=1}^{c} \nu_k = 1$, the grade correlation matrix $\bar{\Psi}$ defined by* (4.2.13), (4.2.14) *is positive definite.*

Throughout the rest of Section 4.2, condition A will be assumed. This assumption, however, is not very restrictive, because $\bar{\Psi}$ is positive definite, unless one or more of the random variables $\{2\bar{F}_{[i]}(\eta_i)-1\}$, $i = 1, \ldots, p$ can be expressed linearly in terms of the others with probability one. This, by our assumption of continuity of $F_1, \ldots, F_c$ for all points in Ω implies that one or more of the random variables $\eta_1, \ldots, \eta_p$ can almost surely be expressed as functions of the remaining variables, the function being monotonic in each of the arguments involved. So, if this form of singularity is excluded, condition A will always be satisfied. We now prove the main theorem of this subsection.

Theorem 4.2.3. *Under condition A on Ω, the conditional null distribution of W_N given Λ_N converges, in probability, to chi-square distribution with $p(c-1)$ degrees of freedom (d.f.).*

The proof of this theorem follows readily from theorem 4.2.1. and lemma 4.2.2. For the intended brevity of the paper, the details are omitted.

Corollary. For any $\epsilon : 0<\epsilon< 1$, let $W_{N,\epsilon}(\boldsymbol{L}_N)$ and $A_{N,\epsilon}(\boldsymbol{L}_N)$ be defined as in (4.1.16) and (4.1.17) and let $\chi^2_p(c-1), \epsilon$ be the upper $100\epsilon\%$ point of a χ^2 distribution with $p(c-1)$d.f. Then as $N \to \infty$,

$$W_{N,\epsilon}(\boldsymbol{\Lambda}_N) \xrightarrow{P} \chi^2_{p(c-1),\epsilon} \text{ and } A_{N,\epsilon}(\boldsymbol{\Lambda}_N) \xrightarrow{P} 0.$$

The proof is omitted (*see* [13, p. 171]).

Remark. In the theorem and corollary above, Λ_N corresponds to $\boldsymbol{F} = (F_1, \ldots, F_c)$, where $\boldsymbol{F}\epsilon\Omega$, but H_0 may or may not be true. Further, it is not essential for the above theorem and corollary that $\boldsymbol{\Lambda}_N$ should be the collection matrix corresponding to samples taken from a fixed point $\boldsymbol{F}\epsilon\Omega$. If $\boldsymbol{\Lambda}_N$ is any sequence of random matrices, $\boldsymbol{\Lambda}_N$ assuming values of $\mathscr{L}_N$, such that $\boldsymbol{R}(\boldsymbol{\Lambda}_N)$ is positive definite, in probability, then the theorem and corollary would hold.

Again, by an adaptation of the method of proof of corollary 3 to theorem 3.2.2. of [7], and using (4.2.12), we obtain the following theorem on the unconditional distribution of W_N (under H_0).

Theorem 4.2.4. *Under H_0, the statistic W_N, defined by* (4.1.15), *has asymptotically a chi-square distribution with* $p(c-1)$ *d.f.*

Theorem 4.2.4. enables us to suggest an asymptotically distribution free test based on the same statistic W_N, defined by (4.1.15). This test consists in rejecting the null hypothesis when W_N exceeds the value $\chi^2_{p(c-1),\varepsilon}$. This test will be termed hereafter as the *asymptotically distribution-free* rank-sum test.

4.3. Consistency of the tests

As in the preceding section, we define by $F_{[i]}(x_i)$ the marginal cdf of the ith variate η_i, $i = 1, \ldots, p$, where $(\eta_1, \ldots, \eta_p)$ has the cdf $\overline{F}$, defined by (4.2.12). Also let

$$d_i^{(k)} = \int_{E^p} \{2\overline{F}_{[i]}(x_i) - 1\} dF_k(x_1, \ldots, x_p), \qquad i = 1, \ldots, p; \quad k = 1, \ldots, c. \tag{4.3.1}$$

Theorem 4.3.1. *The exact test* (4.1.16) *and the asymptotically distribution-free test for* H_0, *are both consistent against the set of alternatives that* $d_i^{(k)}(k = 1, \ldots, c;\ i = 1, \ldots, p)$ *are not all zero. Thus, for translation type of alternatives in* (2.5), *these tests are consistent against the set of alternatives that* $\boldsymbol{\theta}^{(1)}, \ldots, \boldsymbol{\theta}^{(c)}$ *are not all null vectors.*

The proof of this theorem is straight-forward and is omitted. By virtue of the consistency, for any fixed alternative (deviating from the null hypothesis (2.1),) the power of the tests will be asymptotically equal to unity. In the next subsection we shall consider the usual Pitman's type of shift alternatives and study the power proporties of the tests.

4.4. Asymptotic power of the tests

Let us consider the sequence of alternatives $\{H_N\}$, where H_N is represented by $(F_{1N}, \ldots, F_{cN})$:

$$F_{kN}(\boldsymbol{x}) = F(\boldsymbol{x} + N^{-1/2}\,\boldsymbol{\theta}^{(k)}),\ k = 1, \ldots, c; \tag{4.4.1}$$

where $\boldsymbol{\theta}^{(k)} = (\theta_1^{(k)}, \ldots, \theta_p^{(k)})$, $k = 1, \ldots, c$, are c vectors not all of which are identical.

Just as in Sections 4.2 and 4.3 we define the marginal cdf of the ith variate associated with $F(x_1, \ldots, x_p)$ by $F_{[i]}(x_i)$, for $i = 1, \ldots, p$, and write

$$(4.4.2) \qquad \psi_{ij} = 3 \int_{E^p} \{2F_{[i]}(x_i)-1\}\{2F_{[j]}(x_j)-1\}dF(x_1, \ldots, x_p), \quad j = 1, \ldots, p;$$

$$\Psi = (\psi_{ij}), \Psi^{-1} = (\psi^{ij}) = (\psi_{ij})^{-1}.$$

The grade correlation matrix Ψ and its reciprocal Ψ^{-1} will be positive definite by virtue of condition A, stated in subsection 4.2. Further, to simplify the expression for the asymptotic power of the tests, we shall assume that the marginal cdf $F_{[i]}(x)$ is absolutely continuous (for all $i = 1, \ldots, p$) and write $f_{[i]}(x)$ as the corresponding density function (assumed to be continuous), for $i = 1, \ldots, p$. Further, the integrals

$$(4.4.3) \qquad h_i = \int_{-\infty}^{\infty} \{f_{[i]}(x)\}^2 \, dx, \quad i = 1, \ldots, p;$$

will be assumed to exist and the limiting relations

$$(4.4.4) \qquad \lim_{N=\infty} \int_{-\infty}^{\infty} f_{[i]}(x+0(1/\sqrt{N}))dF_{[i]}(x) = h_i, i = 1, \ldots, p$$

will be assumed to hold. Finally, (4.2.1) will be implicit through out this subsection. For each N, $(X^{(k)}_{1(N)}, \ldots, X^{(k)}_{p(N)})$ will denote a set of random variables following the distribution F_{kN}, defined by (4.4.1), there being n_k observations from this distribution, for $k = 1, \ldots, c$. For notational simplicity, we shall keep the subscript N understood, and denote the N observations by

$$(4.4.5) \qquad (X^{(k)}_{1\alpha}, \ldots, X^{(k)}_{p\alpha}), \quad \alpha = 1, \ldots, n_k; \quad k = 1, \ldots, c.$$

Also, we define the mean ranks $\bar{T}^{(k)}_i$, $i = 1, \ldots, p$, $k = 1, \ldots, c$ as in (4.1.1) and the statistic W_N as in (4.1.15).

Theorem 4.4.1. *Under* (4.4.1) *through* (4.4.5), W_N *has asymptotically a noncentral chi-square distribution with* $(p(c-1)$ *d.f. and the noncentrality parameter*

$$(4.4.6) \qquad \Delta_R = 12 \sum_{k=1}^{c} \nu_k \sum_{i=1}^{p} \sum_{j=1}^{p} \psi^{ij}(\theta^{(k)}_i-\bar{\theta}_i)(\theta^{(k)}_j-\bar{\theta}_j)h_i h_j,$$

where $\bar{\theta}_i = \sum_{k=1}^{c} \nu_k \theta^{(k)}_i$, *for* $i = 1, \ldots, p$.

Proof. We shall prove first that $\sqrt{N}\left(\bar{I}_i^{(k)} - \frac{N+1}{2}\right)$, $i=1,\dots,p$; $k = 1, \dots, c$, have asymptotically a multinormal distribution. For this, let us define the usual sign-function $s(x)$ to be $+1\,(-1)$ according as $x > (<) 0$, and let it be 0, otherwise. Also write

$$U_{iN}^{(k,q)} = \frac{1}{n_k n_q} \sum_{\alpha=1}^{n_k} \sum_{\beta=1}^{n_q} s(X_{i\alpha}^{(k)} - X_{i\beta}^{(q)}) \tag{4.4.7}$$

for $k \neq q = 1, \dots, c$; $i = 1, \dots, p$. Then we have

$$\sqrt{N}\left(\bar{I}_i^{(k)} - \frac{N+1}{2}\right) = \frac{1}{2} \sum_{\substack{q=1\\ \neq k}}^{c} \frac{n_q}{N} \sqrt{N}\, U_{iN}^{(k,q)}. \tag{4.4.8}$$

Now, by a well-known technique used in studying the asymptotic distributions of U-statistics (*see*, Hoeffding (1948), and Andrews (1954),) and following some simple but essentially lengthy steps, it can be shown that $\{\sqrt{N} U_{iN}^{(k,q)}, k \neq q = 1, \dots, c;\ i = 1, \dots, p,\}$ has asymptotically a multinormal distribution which is essentially singular and is of rank $p(c-1)$ when Ψ is positive definite. Hence, using (4.4.8) and some routine calculations, it can be shown that under (4.4.1) through (4.4.5) the vector

$$N^{1/2}\left(\bar{I}_1^{(1)} - \frac{N+1}{2}, \dots, \bar{I}_p^{(1)} - \frac{N+1}{2}, \dots, \bar{I}_1^{(c-1)} - \frac{N+1}{2}, \dots, \bar{I}_p^{(c-1)} - \frac{N+1}{2}\right)$$

will have asymptotically a multinormal distribution with mean vector

$$(m_1^{(1)}, \dots, m_p^{(1)}, \dots, m_1^{(c-1)}, \dots, m_p^{(c-1)}),$$

and dispersion matrix $\frac{1}{12}\left(\frac{1}{\nu_k}\delta_{kq} - 1\right)_{k,q=1,\dots,c-1} \otimes \Psi$, where

$$m_i^{(k)} = -h_i(\theta_i^{(k)} - \bar{\theta}_i) \text{ for } i = 1, \dots, p;\ k = 1, \dots, c. \tag{4.4.9}$$

Hence, by simple reasonings, we may conclude that

$$\frac{12}{N} \sum_{k=1}^{c-1} \sum_{q=1}^{c-1} \sum_{i=1}^{p} \sum_{j=1}^{p} \left(\nu_k \delta_{kq} + \frac{\nu_k \nu_q}{\nu_c}\right) \psi^{ij} \left(\bar{I}_i^{(k)} - \frac{N+1}{2}\right)\left(\bar{I}_i^{(q)} - \frac{N+1}{2}\right) \tag{4.4.10}$$

will be asymptotically distributed as a non-central χ^2 with d.f. $p(c-1)$ and non-centrality parameter

$$\Delta_R = 12 \sum_{k=1}^{c-1} \sum_{q=1}^{c-1} \sum_{i=1}^{p} \sum_{j=1}^{p} \left(\nu_k \delta_{kq} + \frac{\nu_k \nu_q}{\nu_c}\right) \psi^{ij} m_i^{(k)} m_j^{(q)}$$

$$(4.4.11) \quad = 12 \sum_{i=1}^{p} \sum_{j=1}^{p} \psi^{ij} \left\{ \sum_{k=1}^{c-1} \nu_k m_i^{(k)} m_j^{(k)} + \frac{1}{\nu_c} \sum_{k=1}^{c-1} \sum_{q=1}^{c-1} \nu_k \nu_q m_i^{(k)} m_j^{(q)} \right\}$$

$$= 12 \sum_{k=1}^{c} \nu_k \sum_{i=1}^{p} \sum_{j=1}^{p} \psi^{ij} (\theta_i^{(k)} - \bar{\theta}_i)(\theta_j^{(k)} - \bar{\theta}_j) h_i h_j,$$

Again, it follows from lemma 4.2.2, (4.2.14), (4.4.1) and (4.4.2) that $\boldsymbol{R}(\boldsymbol{\Lambda}_N) \xrightarrow{P} \boldsymbol{\Psi}$. Hence the statistic W_N, given by (4.1.15) is seen to be asymptotically equivalent, in probability, to the statistic (4.4.10), and the theorem follows.

With the help of theorem 4.4.1, corollary to theorem 4.2.3 and a result due to Hoeffding (1952, p. 171), we readily arrive at the following.

Theorem 4.4.2. *Both the permutation test and the asymptotically distribution-free test (based on the statistic W_N, defined by* (4.1.15),) *are asymptotically power-equivalent and have asymptotically (under the sequence of alternatives in* (4.4.1),) *a noncentral chi-square distribution with $p(c-1)$ d.f. and the noncentrality parameter Δ_R, defined by* (4.4.6).

5. MULTIVARIATE MULTISAMPLE MEDIAN TESTS

We shall now generalize the well-known univariate median test by Brown and Mood (cf. [16]) to the $p(\geqslant 1)$ variate case, following the same conditional approach as in section 3. We adopt the same notations as in the previous sections. Among the N values of $\{I_{i\alpha}^{(k)},\ \alpha = 1, \ldots, n_k;\ k = 1, \ldots, c\}$ there are exactly $a = [N/2]$ ($[s]$ being the largest integer contained in s) values which do not exceed a, for $i = 1, \ldots, p$. Let us then define a system of 2^p-mutually exclusive and exhaustive cells $\left\{J_{r_1 \cdots r_p} : r_i = 0, 1, i = 1, \ldots, p\right\}$ by the convention that if for any cell $J_{r_1 \cdots r_p} r_i = 1$ (or 0) then for any observation belonging to it $I_i^{(k)} >$ (or $\leqslant$) a, for $i = 1, \ldots, p$. Let among the N random rank p-tuplets $C_{r_1 \cdots r_p}$ observations belong to the cell $J_{r_1 \cdots r_p}$, for all $\left(r_{1, \ldots, r_p}\right)$. Then $\{C_{r_1 \cdots r_p}\}$ is a set of random variables, depending on $\mathbf{\Lambda}_N$, and we may note that

$$\sum_{(S_{j\alpha})} C_{r_1 \cdots r_p} = a \text{ (if } \alpha = 0) \text{ or } N - a \text{ (if } \alpha = 1), \tag{5.1.1}$$

where the summation $S_{j\alpha}$ extends over all $r_1, \ldots, r_{j-1}, r_{j+1}, \ldots, r_p$ (for a given $r_j = \alpha$), $j = i, \ldots, p$. Thus we have a 2^p-contingency table for the pooled sample of size N, and we term this as the *basic table*. Now, corresponding to the kth sample of size n_k, we have, by reference to the same system of cells $\left\{J_{r_1 \cdots r_p}\right\}$, another 2^p-table, whose entries are denoted by

$$n_{k(r_1 \cdots r_p)}, \; r_i = 0, 1, \; i = 1, \ldots, p; \tag{5.1.2}$$

this table will be termed as *table* k, for $k = 1, \ldots, c$. Obviously

$$\sum_{k=1}^{c} n_{k(r_1 \cdots r_p)} = C_{r_1 \cdots r_p}, \quad \text{for all } (r_1, \ldots, r_p). \tag{5.1.3}$$

Now under the null hypothesis that the c cdf's $F_1, \ldots, F_c$ are all identical, these c (2^p)-contingency tables should be statistically homogeneous and consistent with the basic table. By virtue of (5.1.3), it is thus sufficient to test for the agreement of the c tables (i.e., table k, $k = 1, \ldots, c$). Since the test to be proposed is based on the system of cells demarcated by the pooled sample medians, by analogy with the univariate case (cf. [16]), it is termed the multivariate multisample median test.

Now, under the null hypothesis $H_0 : F_1 = \ldots F_c = F$ (say), and given $\mathbf{\Lambda}_N = \mathbf{L}_N$, all possible partitioning of the N rank p-tuplets into the c subsets of sizes $n_1, \ldots, n_c$, respectively, are equally likely, each having the conditional probability $(\Pi_1^c n_k!/N!)$ which we conventionally put as $\binom{N}{[n_k]}^{-1}$. Let now for the given $\mathbf{\Lambda}_N = \mathbf{L}_N$, the realized values of $\{C_{r_1 \cdots r_p}\}$ be $\{c_{r_1 \cdots r_p}\}$. Then by simple logic we derive the conditional probability function of $\left\{n_{k(r_1\cdots r_p)};\ r=0,1;\ i=1,\ldots,p,\ \ k=1,\ldots,c\right\}$ conditioned on $\Lambda_N = \boldsymbol{L}_N$ as given by

$$(5.1.4)\qquad \binom{N}{[n_k]}^{-1}\left\{\prod_S \binom{c_{r_1\cdots r_p}}{[n_{k(r_1\cdots r_p)}]}\right\}$$

Thus for any $(N, \boldsymbol{L}_N)$, the expression (5.1.4) is a completely known function. Here also, we shall use a quadraric form in the variables $\left\{n_{k(r_1\cdots r_p)},\ r_i=0,1,\ i=1,\ldots,p;\ k=1,\ldots,c\right\}$ to formulate our test statistic, and for this we consider first the first and second order moments of these random variables. It follows directly from (5.1.4) that

$$(5.1.5)\qquad E\left\{n_{k(r_1\cdots r_p)}\,\middle|\,\boldsymbol{L}_N, H_0\right\} = \frac{n_k}{N}\,c_{r_1\cdots r_p},\ \text{and}$$

$$(5.1.6)\qquad \operatorname{Cov}\left\{n_{k(r_1\cdots r_p)}, n_{q(r'_1,\ldots,r'_p)}\,\middle|\,\boldsymbol{L}_N, H_0\right\}$$

$$= n_k(\delta_{kq}\,N-n_q)c_{r_1\cdots r_p}\left(\delta_{rr'}N-c_{r'_1\cdots r'_p}\right)\Big/N^2(N-1),$$

where $\delta_{kq}=1$ if $k=q$, $\delta_{kq}=0$ if $k\neq q$, and $\delta_{rr'}=1$ if $(r_1\ldots r_p) \equiv (r'_1\ldots r'_p)$ and $\delta_{rr'}$ is zero otherwise.

5.2. The classification of median tests

The 2^pc random variables $\left\{n_{k(r_1\cdots r_p)},\ r_i=0,1,\ i=1,\ldots,p,\ k=1,\ldots,c\right\}$ are subject to 2^p+c-1 constraints, namely (5.1.3) and

$$(5.2.1)\qquad \sum_S n_{k(r_1\cdots r_p)} = n_k,\ \text{for}\ k=1,\ldots,c;$$

(where the summation S extend over all $(r_1,\ldots,r_p)$,) and where (5.1.3) and (5.2.1) satisfy in turn

$$(5.2.2)\qquad \sum_{k=1}^{c} n_k = N = \sum_S C_{r_1\cdots r_p}.$$

Thus, the effective number of degrees of freedom of this set of random variables is $(2^p-1)(c-1)$. Of course, study of all these d.f. may not be necessary for detecting differences in locations only. For this purpose, we define

(5.2.3) $$n^*_{k(i)} = \sum_{S_{i0}} n_{k(r_1 \ldots r_p)},\ i = 1, \ldots, p,\ k = 1, \ldots, c;$$

where the summation S_{i0} extends over all $(r_1, \ldots, r_p)$ for which $r_{i0} = 0$, i.e., $n^*_{k(i)}$ is the number of observations in the kth sample whose ith variate values do not exceed that of the ath order statistic of pooled sample ith variate values, for $i = 1, \ldots, p$, $k = 1, \ldots, c$. Using (5.1.4) it can be readily shown that

(5.2.4) $$E\{n^*_{k(i)} \mid \boldsymbol{L}_N, H_0\} = an_k/N,\ i = 1, \ldots, p,\ k = 1, \ldots, c.$$

Thus, generalizing the univariate median test [cf. Andrews (1954)], and keeping in mind shifts in locations only, we may base our test on the $p(c-1)$ d.f. carried by the discrepancies $\{N^{-\frac{1}{2}}(n^*_{k(i)} - an_k/N),\ i = 1, \ldots, p,\ k = 1, \ldots, c\}$. Such a median test would be termed a *Type A median test.* It can be shown that the remaining $(2^p - p - 1)(c-1)$ d.f. are sensitive not only to shifts in locations but also to any heterogeneity of the association patterns of the different cdf's. Tests based on these $(2^p - p - 1)(c-1)$ d.f. may be used to test for the identity of association patterns (assuming the identity of locations) (cf. Chatterjee and Sen (1965), for $p = c = 2$). Such a test would be termed the *Type B median test.* Finally, the test based on all the $(2^p - 1)(c-1)$ d.f. is consistent not only against shifts in the median vectors but also against any difference of the association patterns; such a test will be termed the *Type C median test.* Since we are mainly interested in the location problem, we shall consider only type A median test and append very briefly the case of type C median test.

To formulate the test statistic, we write

(5.2.5) $$C^{\alpha,\beta}_{(i,j)} = \sum_{S_{\alpha,\beta,i,j}} C_{r_1 \ldots r_p}$$

where the summation $S_{\alpha,\beta,i,j}$ extends over all possible $r_1, \ldots, r_p$ for which $r_i = \alpha$, $r_j = \beta$; $\alpha, \beta = 0, 1$; and let

(5.2.6) $$d_{ij} = N^{-1} \sum_{\alpha=0}^{1} \sum_{\beta=0}^{1} (-1)^{\alpha+\beta} C^{\alpha,\beta}_{(i,j)},\ \text{for } i, j = 1, \ldots, p;$$

(5.2.7) $$\boldsymbol{D} = (d_{ij}).$$

If all $\left\{C_{r_1\cdots r_p}\right\}$ are positive, it can be shown that $\boldsymbol{D}$ is a positive definite matrix. The positive-definiteness of $\boldsymbol{D}$ may also follow when some of $\left\{C_{r_1\cdots r_p}\right\}$ are equal to zero. However, it will be seen later on that under some mild restrictions on Ω, $\boldsymbol{D}$ will be positive definite, in probability. If $\boldsymbol{D}$ is not positive definite, as in the rank-sum test, we may work with the highest order positive definite principle minor of it. Thus, on denoting by $\boldsymbol{D}^{-1}=(d^{ij})$ the reciprocal matrix of $\boldsymbol{D}$, we may formulate (following some simple but somewhat lengthy deductions involving the use of (5.1.5), (5.1.6) and (5.2.3), the type A median test-satistic as

$$(5.2.8)\quad M_N = 4\sum_{k=1}^{c}\sum_{i=1}^{p}\sum_{j=1}^{p} d^{ij}\{n^*_{k(i)}-an_k/N\}\{n^*_{k(j)}-an_k/N\}/n_k.$$

For small values of $\left(N,\left\{C_{r_1\cdots r_p}\right\}\right)$, the matrix $\boldsymbol{D}$ can be readily evaluated and the exact null distribution of M_N (given $\boldsymbol{\Lambda}_N=\boldsymbol{L}_N$) may be derived with the aid of (5.1.4). The test function will be essentially similar to (4.1.16), (4.1.17) and (4.1.18). The large sample approach is considered below.

5.3. Large simple distribution of M_N under H_0

In this section we make N indefinitely large, subject to (4.2.1). We define $\overline{F}$ as in (4.2.12), and let $\boldsymbol{\eta}=(\eta_1,\ldots,\eta_p)$ follow the cdf $\overline{F}$. Let the medians of $\overline{\eta}_1,\ldots,\overline{\eta}_p$ be denoted by $\bar{\mu}_1,\ldots,\bar{\mu}_p$ and are assumed to be uniquely defined. Also we determine 2^p-cells $\left\{J^0_{r_1\cdots r_p}\right\}$ by the rule that for any observation belonging to the cell $J^0_{r_1\cdots r_p}$, r_i is equal to zero (or one) according as $\eta_i \leqslant$ (or $>$) $\bar{\mu}_i$, for $i=1,\ldots,p$. Let then

$$(5.3.1)\qquad P\left\{\boldsymbol{\eta}\in J^0_{r_1\cdots r_p}\right\}=\overline{P}^0_{r_1\cdots r_p},\ \text{ for all } (r_1,\ldots,r_p).$$

In the rest of Section 5, we assume the following condition :

Condition B. *Whatever* $\nu_1, \ldots, \nu_c$ $(0 < \nu_k < 1,\ \sum_1^c \nu_k = 1)$, *for all* $F \epsilon \Omega$,

$$\operatorname*{Inf.}_{(r_1 \cdots r_p) \epsilon S} \bar{P}^0_{r_1 \cdots r_p} > 0.$$

If $X_k = (X_1^{(k)}, \ldots, X_p^{(k)})$ follows the distribution F_k, and

(5.3.2) $$P\left\{X_k \epsilon\, J^0_{r_1 \cdots r_p}\right\} = P^{(k)}_{r_1 \cdots r_p},$$

then $\bar{P}^0_{r_1 \cdots r_p} = \sum_{k=1}^{c} \nu_k \bar{P}^{(k)}_{r_1 \cdots r_p}$, and hence, the condition B is equivalent to: whatever $\nu_1, \ldots, \nu_c$

(5.3.3) $$\operatorname*{Inf}_{(r_1 \cdots r_p) \epsilon S} \operatorname*{Inf}_{k} \left\{P^{(k)}_{r_1 \cdots r_p}\right\} > 0.$$

If as in Section 2, we write $\mu_1^{(k)}, \ldots, \mu_p^{(k)})$ for the median point of the distributiion F_k, then (5.3.4) will be satisfied, if and only if, none of the 2^p cells formed by taking any arbitrary point within the simplex spanned over the c points $(\mu_1^{(k)}, \ldots, \mu_p^{(k)})$, $k = 1, \ldots, c$ and drawing lines parallel to the axis through that point have zero probability content with respect to all the distributions $F_1, \ldots, F_c$. Then we have the following.

Theorem 5.3.1. *Under* H_0 *in* (2.1) *and subject to the condition* (4.2.1) *and condition* B *in* (5.3.1), *the statistic* M_N, *defined by* (5.2.8), *has asymptotically, in probability, a chi-square distribution with* $p(c-1)$ *d.f.*

Proof. Avoiding the details of proof, we say that under the condition B, as N is increased (subject to (4.2.1), all the $\left\{\left(C_{r_1 \cdots r_p}\right)/N\right\}$ can be made strictly positive, in probability. Consequently, we can apply Stirling's approximations to all the factorials in (5.1.4) if N is taken adequately large. This will lead to the asymptotic multinormality of the set of random variables $\left\{n_k^{-\frac{1}{2}}\left(n_{k(r_1 \cdots r_p)} - (n_k/N)\, c_{r_1 \cdots r_p}\right)\right.$, for all $(r_1, \ldots, r_p)$ and $k = 1, \ldots, c\}$, (under the conditional probability measure, given $\Lambda_N = L_N$.) Since $\{n^*_{k(i)}\}$ are all linear functions of these 2^p-c random variables, by means of linear transformations of variables it can be shown that $[n_k^{-\frac{1}{2}}(n^*_{k(i)}-an_k/N),\ i=1,\ldots,p,\ k=1,\ \ldots,\ c]$ has (jointly) asymptotically a multinormal distribution with a null mean vector and dispersion matrix $\boldsymbol{D}$, defined by (5.2.6) and (5.2.7). The rest of the proof is simple and is omitted. Hence the theorem.

Corollary 5.3.1. It follows by the same technique that on defining the type C median test statistic as

$$\text{(5.3.4)} \qquad M^*_N = \sum_{k=1}^{c} (N/n_k) \sum_{S} \left\{ n_{k(r_1\cdots r_p)} - (n_k/N)C_{r_1\cdots r_p} \right\}^2 \Big/ C_{r_1\cdots r_p},$$

under H_0 in (2.1) and conditioned on $\boldsymbol{\Lambda}_N = \boldsymbol{L}_N$, the statistic M^*_N has asymptotically a chi-square distribution with $(2^p-1)(c-1)$ d.f.

By virtue of theorem 5.3.1. the large sample type A median test may be formulated as follows. Let $\chi^2_{t,\varepsilon}$ be the upper $100\epsilon\%$ point of a chi-square distribution with t d.f. Then, if

$$\text{(5.3.5)} \qquad M_N \begin{cases} \geqslant \chi^2_{(p-1)(c-1),\,\varepsilon}, & \text{reject } H_0 \text{ in (2.1)} \\ < \chi^2_{(p-1)(c-1),\,\varepsilon}, & \text{accept } H_0. \end{cases}$$

(For the type C median test, replace M_N by M^*_N and the d.f. $(p-1)(c-1)$ by $(2^p-1)(c-1)$, respectively).

5.4. Consistency of the tests

We define the population medians as in section 2 (cf. (2.3),) and consider the following theorem whose proof is omitted.

Theorem 5.4.1. *The type A median test is consistent against the set of alternatives that the pc(c—1) differences in* (2.3) *are not all zero. The type C median test is consistent against the set of alternatives*

$$\text{(5.4.1)} \qquad \sup_{(k,\,q)} \left[\sup_{(r_1\cdots r_p)} \left| P^{(k)}_{r_1\cdots r_p} - P^{(q)}_{r_1\cdots r_p} \right| \right] > 0.$$

and as such, it will be consistent not only to shifts in the median vectors but also to the difference of the association patterns.

By virtue of theorem 5.4.1 for any fixed alternative different from H_0 in (2.1) the power of the tests will be asymptotically equal to unity. Thus, for the study of the asymptotic power properties of the tests we shall again consider the sequence of alternatives in (4.4.1), and this study is made in the following subsection.

5.5. Asymptotic power function of the median tests

As in section 4.4, we shall conceive of a sequence of values of N, and for each N, c subsequences of random variables. The condition (4.2.1) will also be implicit in this section. $F(\boldsymbol{x})$ will be assumed to be absolutely continuous and the corresponding (continuous) density function will be denoted by $f(\boldsymbol{x})$. Further, without any loss of generality, we assume that the median vector of $F(\boldsymbol{x})$ is a null vector. As in the bivariate case (cf. [7]), we require the following conditions :

$$(5.5.1) \qquad \int_{-\infty}^{\infty} \cdots \int_{-\infty}^{\infty} f\left(x_1, \ldots, x_{p-1}, \frac{1}{N}\right) dx_1, \ldots, dx_{p-1}$$

$$= \int_{-\infty}^{\infty} \cdots \int_{-\infty}^{\infty} f(x_1, \ldots, x_{p-1}, 0) dx_1 \ldots dx_{p-1} + 0(1/N),$$

and similarly for each of the other $p-1$ coordinates in $f(x_1, \ldots, x_p)$. Let then

$$(5.5.2) \qquad \begin{aligned} \alpha_1 &= \int_{-\infty}^{\infty} \cdots \int_{-\infty}^{\infty} f(0, x_2, \ldots, x_p) dx_2 \ldots dx_p \\ &\cdots \qquad \cdots \qquad \cdots \\ &\cdots \qquad \cdots \qquad \cdots \\ \alpha_p &= \int_{-\infty}^{\infty} \cdots \int_{-\infty}^{\infty} f(x_1, \ldots, x_{p-1}, 0) dx_1 \ldots dx_p; \end{aligned}$$

We define $P^0_{r_1 \cdots r_p}$ as in (5.3.1) with the only change that η has the cdf $F(\boldsymbol{x})$ instead of $\bar{F}(\boldsymbol{x})$. Also we rewrite these as

$$P^0_{0,\cdots 0} = P^0_1; \quad P_{0 \cdots 0,1} = P^0_2, P^0_{0,\cdots 1,0} = P^0_3,$$

$$P^0_{0,\cdots 1,1} = P^0_4, \ldots, P^0_{11\ldots 1} = P^0_{2^p}.$$

Let us then define a set of p vectors $\gamma_1, \ldots, \gamma_p$ as

(5.5.3) $\gamma_i = \left(e_{i1}\sqrt{P_1^0}, e_{i2}\sqrt{P_2^0}, \ldots e_{i2}\sqrt{P_{2^p}^0}\right); j = 1, \ldots, p_j$

$$\text{where} \quad e_{ik} = \begin{cases} 1 \text{ if for } P_k^0, r_i = 0 \\ = -1 \text{ if for } P_k^0, r_i = 1 \end{cases} \text{ for } \begin{array}{l} i = 1, \ldots, p \\ k = 1, \ldots, 2^p. \end{array}$$

Then obviously $\gamma_1 \ldots \gamma_p$ are linearly independent of each other and are all of unit length. Also let

$$\varphi_{ij} = \gamma_i \cdot \gamma_j, \text{ for } i, j = 1, \ldots, p; \text{ and} \tag{5.5.4}$$
$$\Phi = (\varphi_{ij}), \Phi^{-1} = (\varphi_{ij})^{-1} = (\varphi^{ij}).$$

Then it is easily shown that Φ or Φ^{-1} are positive definite. Then we have the following.

Theorem 5.5.1. *Under the sequence of alternative in* (4.4.1), *the statistic M_N has asymptotically a non-central χ^2 distribution with* $p.(c-1)$ *d.f. and with the non-centrality parameter*

$$\Delta_M = 4 \sum_{i=1}^{p} \sum_{j=1}^{p} \phi^{ij}\left(\sum_{k=1}^{c} \nu_k \,\delta_{ik}\, \delta_{jk}\right) \alpha_i\, \alpha_j.$$

Proof. Let $\tilde{X}_1 \ldots \tilde{X}_p$ be the pooled sample medians (a-th largest value; $a = [N/2]$) of the p variates based on the samples (4.4.5) taken from the distributions F_{kN}, $k = 1, \ldots, c$. We define the cells $\left\{J_{r_1 \cdots r_p}, r_i = 0, 1; i = 1, \ldots, p\right\}$ as in the beginning of section 5, i.e., by means of the sample median vector $(\tilde{X}_1, \ldots, \tilde{X}_p)$. Also, let $\mathbf{X}_N^{(k)}$ denote a random vector following the cdf F_{kN}, for $k = 0, \ldots, c$, where $F_{0N} = F$. Then we write $p_{r_1 \cdots r_p}^{(k)}$ as the probability content of the cell $J_{r_1 \cdots r_p}$ with respect to the cdf F_{kN}, for $k = 0, \ldots, c$, and the successive derivatives of these probabilities (with respect to $\tilde{X}_i$, $i = 1, \ldots, p$) are denoted by $p_{r_1 \cdots r_p}^{(k); i}$ $p_{r_1 \cdots r_p}^{(k); i, j}$, $\ldots$, $p_{r_1 \cdots r_p}^{(k); 1, \cdots, p}$ $(i, j, \ldots = 1, \ldots, p)$, for $k = 0, \ldots, c$; $(r_1 \ldots r_p) \epsilon S$. Then there are 2^p values of $\left\{p_{r_1 \cdots r_p}^{(k)}\right\}$ for each $k = 0, \ldots, c$, $p2^p$ values of $\left\{p_{r_1 \cdots r_p}^{(k); i}\right\}$ for each $k = 0, \ldots, c$, and so on. We now adopt Mood's (1941) technique of finding the joint distribution of the sample medians in a sample from a multivariate population with an extension to the multisample case, and get that the joint probability function of $\{n_{k(r_1\cdots r_p)}$, $k=1, \ldots, c, (r_1 \ldots r_p)\epsilon S\}$ and (the density function of) $\tilde{X}_1, \ldots, \tilde{X}_p$ is the sum of the following terms :

(i) If the pooled sample median vector is determined by p different observations of the pooled sample, then the contribution will be

$$(5.5.5) \quad \left\{\prod_{k=1}^{c}\left[\left(\begin{matrix} n_k \\ \left[n_{k(r_1\cdots r_p)}\right]\end{matrix}\right)\prod_{S}\left(p^{(k)}_{r_1\cdots r_p}\right)^{n_{k(r_1\cdots r_p)}}\right]\right\}.$$

$$\left\{\sum_{S_p}\sum_{S^*}\left[\prod_{j=1}^{p}\left\{n_{k_j(r_1^*\cdots r_p^*)}-\sum_{l=1}^{j-1}\delta_{k_jk_l}\right\}p^{(k_j);j}_{r_1^*\cdots r_p^*}\Big/p^{(k_j)}_{r_1^*\cdots r_p^*}\right]\right\}\prod_{j=1}^{p}d\tilde{X}_j,$$

where $\{\delta_{k_jk_l}\}$ are Kronecker deltas with $\delta_{k_jk_0}=0$; the product S extends over all the 2^p terms $(r_1 \ldots r_p)\epsilon S$, the sum S_p over the c^p terms : $k_j = 1, \ldots, c$; $j = 1, \ldots, p$; and the sum S^* over all the possible terms like the one within the third bracket succeeding S^*. In the particular case of $c = p = 2$, reference may be made to [7] for some simplification of this procedure.

(ii) If the pooled sample median vector is determined by less than p observations, say $p-h$ observations ($h \geqslant 1$), then proceeding precisely on the same line as in Mood (1941), it can be shown that the corresponding probabilility terms would amount to a term of the order N^{-h}, as compared to (5.5.5). Thus, for $h \geqslant 1$, the contributions of the probability terms may be neglected. Further, it follows from the results on the univariate median test (cf. Mood (1954)] that $N^{\frac{1}{2}}\tilde{X}_j$ is bounded in probability, for $j = 1, \ldots, p$. Hence, it can be shown by simple but somewhat lengthy algebraic manipulations that under the sequence of alternatives in (4.4.1)

$$\text{(i)} \quad n_{k(r_1\cdots r_p)}\Big/\left[n_k\, p^{(0)}_{r_1\cdots r_p}\right] = 1+O_p(N^{-\frac{1}{2}}),$$

(ii) $\left| p^{(k)}_{r_1\cdots r_p} - p^{(0)}_{r_1\cdots r_p} \right| = O_p(N^{-\frac{1}{2}})$,

(iii) $\left| p^{(k);\,i}_{r_1\cdots r_p} - p^{(0);\,i}_{r_1\cdots r_p} \right| = O_p(N^{-\frac{1}{2}})$, for $i = 1, \ldots, p$;

(iv) $\left| p^{(0)}_{r_1\cdots r_p} - p^{(0)}_{1-r_1\cdots 1-r_p} \right| = O_p(N^{-\frac{1}{2}})$,

for all $(r_1 \ldots r_p)\epsilon S$, $k = 1, \ldots, c$. Hence, it can be shown by some simple but somewhat lengthy algebraic manipulations that

(5.5.6) $$\sum_{S_i} p^{(0);\,i}_{r_1\cdots r_p} = [\alpha_i + O_p(N^{-\frac{1}{2}})], \quad \text{for } i = 1, \ldots, p;$$

where the summation S_i extends over all possible 2^{p-1} values of $(r_1 \ldots r_p)$; $r_j = 0, 1, j = 1, \ldots, p$ $(\neq i)$, for $i = 1, \ldots, p$, and where α_j's are defined by (5.5.2). Hence, from (5.5.5), (5.5.6) and some simplifications, the joint probability function of the set of random variables $\left\{ n_{k(r_1\cdots r_p)}\,;\, k = 1, \ldots, c, (r_1 \ldots r_p)\epsilon S \right\}$ and $\tilde{X}_1, \ldots, \tilde{X}_p$ asymptotically reduces to

(5.5.7) $$N^p \alpha_1 \ldots \alpha_p \left[\prod_{k=1}^{c} \left\{ \left(\begin{matrix} n_k \\ [n_{k(r_1\cdots r_p)}] \end{matrix} \right) \prod_S \left(p^{k}_{r_1\cdots r_p} \right)^{n_{1(r_1\cdots r_p)}} \right\} \right] d\tilde{X}_1 \ldots d\tilde{X}_p + O_p(N^{-1/2}).$$

Again, by using (multivariate) Taylor's expansion it can be shown that

(i) $$N^{\frac{1}{2}} \left| \left(p^{(k)}_{r_1\cdots r_p} - P^{(k)}_{r_1\cdots r_p} \right) - \left(p^{(0)}_{r_1\cdots r_p} - P^{(0)}_{r_1\cdots r_p} \right) \right| \xrightarrow{P} 0, k = 1, \ldots, c;$$

(ii) $\left| P^{(0)}_{r_1\cdots r_p} - P^{(0)}_{r_1\cdots r_p} \right| = O_p(N^{-\frac{1}{2}})$,

(iii) $p^{(0)}_{r_1\cdots r_p} = \sum_{k=1}^{c} \nu_k \left[p^{(k)}_{r_1\cdots r_p} \right] + O_p(N^{-\frac{1}{2}})$,

for all $(r_1 \ldots r_p)\epsilon S$.

Hence, (5.5.7) can be shown to be reducible asymptotically to

$$N^p\alpha_1 \dots \alpha_p \left| |A|/(2\pi)^{2p-1} \right|^{c/2}$$

(5.5.8)
$$\exp\left\{ -\frac{1}{2} \sum_{k=1}^{c} \sum_{S} \frac{\left[n_{k(r_1\cdots r_p)} - n_k p^{(k)}_{r_1\cdots r_p} \right]^2}{n_k P^{(k)}_{r_1\cdots r_p}} \right\}.$$

$$\prod_{i=1}^{p} d\tilde{X}_i \prod_{k=1}^{c} \prod_{l=1}^{2^p-1} dZ_{k(l)},$$

where A has the principal minors $A_{ii} = 1/P_i^{(0)} + 1/P^{(0)}_{2^p}$, $i = 1, \dots, 2^p-1$, and the off-diagonal minors $A_{ij} = 1/P^{(0)}_{2^p}$, $i \neq j = 1, \dots, 2^p-1$, $Z_{k(j)} = n_k^{-\frac{1}{2}} \; (n_{k(j)} - n_k \, p_j^{(k)})$, $j = 1, \dots, 2^p-1$, $k = 1, \dots, c$; the number i or j ($= 1, \dots, 2^p$), being attached to $(r_1 \dots r_p)$ in the same convention as in just before (5.5.3). We now write

(5.5.9)
$$d^k_{r_1\cdots r_p} = n_k^{\frac{1}{2}} \left(p^{(k)}_{r_1\cdots r_p} - p^{(0)}_{r_1\cdots r_p} \right), \; k = 1, \dots, c, \;\; (r_1 \dots r_p) \epsilon S,$$

(5.5.10)
$$\xi^k_{(l)} = \left\{ n_{k(l)} - \frac{n_k}{N} \; C_{(l)} \right\} / \sqrt{n_k}$$

for $l = 1, \dots, 2^p-1$; $k = 1, \dots, c$; and

$$W_{(l)} = \{C_{(l)} - N^{(0)}_{p(l)}\}/\sqrt{N} \text{ for } l = 1, \dots, 2^p-1.$$

Then, we have

(5.5.11)
$$Z_{k(l)} = \xi^k_{(l)} - d^k_{(l)} + \sqrt{n_k/N} \; W_{(l)}$$
$$\text{for } l = 1, \dots, 2^p-1; \;\; k = 1, \dots, c;$$

(5.5.12)
$$\sum_{k=1}^{c} \sum_{l=1}^{2^p} Z^2_{k(l)} = \sum_{k=1}^{c} \sum_{l=1}^{2^p} [\xi^k_{(l)} - d^k_{(l)}]^2 / P^0_{(l)}$$
$$+ \sum_{l=1}^{2^p} W^2_{(l)}/P^0_{(l)} + o_p(1).$$

By noting that $d^k_{r_1\cdots r_p} = n_k^{\frac{1}{2}} \left(p^{(k)}_{r_1\cdots r_p} - p^{(0)}_{r_1\cdots r_p} \right) + o_p(1)$, for all $k = 1, \dots, c$ and all $(r_1 \dots r_p) \epsilon S$, we get from (5.5.8) through (5.5.12) and on integrating over the range of the variables $\tilde{X}_1, \ldots, \tilde{X}_p$, that the joint distribution of $\xi_{(l)}^{(k)}$, $l = 1, \ldots, 2^p-1$, $k = 1, \ldots, c$ reduces to

$$\left\{ \left| |A|/(2\pi)^{2^{p-1}} \right|^{(c-1)/2} \right. \tag{5.5.13}$$

$$\left. \exp\left\{ -\frac{1}{2} \sum_{k=1}^{c} \sum_{S} \frac{\left[\xi^{k}_{r_1\cdots r_p} - d^{k}_{r_1\cdots r_p}\right]^2}{P^0_{r_1\cdots r_p}} \right\} \prod_{k=1}^{c=1} \prod_{i=1}^{2^p-1} d\xi^{k}_{(i)} \right\}.$$

Hence it follows by some routine algebra that

$$T_N = \sum_{k=1}^{c} \sum_{i=1}^{p} \sum_{j=1}^{p} \phi^{ij} 4\{n^*_{k(i)} - an_k/N\}\{n^*_{k(j)} - an_k/N\}/n_k \tag{5.5.14}$$

has asymptotically a noncentral chi-square distribution with $p(c-1)$ d.f. and the noncentrality parameter

$$\Delta_M = 4 \sum_{k=1}^{c} \nu_k \sum_{i=1}^{p} \sum_{j=1}^{p} \varphi^{ij} \alpha_i \alpha_j (\theta_i^{(k)} - \bar{\theta}_i)(\theta_j^{(k)} - \bar{\theta}_j), \tag{5.5.15}$$

where α_j's are defined by (5.5.2), $\boldsymbol{\theta}^{(k)}$, $k = i, \ldots, c$ by (4.4.1), and $\boldsymbol{\theta}$ by (4.4.6). Further, it is easy to see that under (4.4.1).

$$\boldsymbol{D} \xrightarrow{P} \boldsymbol{\Phi}, \text{ i.e., } \boldsymbol{D}^{-1} \xrightarrow{P} \boldsymbol{\Phi}^{-1} \tag{5.5.16}$$

(as $\boldsymbol{\Phi}$ is assumed to be nonsingular), and hence, from (5.2.8), (5.5.14) and (5.5.16) it follows that under (4.4.1),

$$M_N \overset{P}{\sim} T_N. \tag{5.5.17}$$

The rest of the proof follows by some standard procedure and is therefore omitted.

6. ASYMPTOTIC POWER-EFFICIENCY OF THE TWO TESTS

Since, under $\{H_N\}$ in (4.4.1), both the test statistics have noncentral chi-square distributions, differing only in the noncentrality parameters, a comparison of the two noncentrality parameters will reveal their asymptotic relative efficiency (A.R.E.). However, such an A.R.E. depends on the vectors $\boldsymbol{\theta}^{(k)}$, $k = 1, \ldots, c$ as well as on the two matrices $\boldsymbol{\Psi}$ and $\boldsymbol{\Phi}$ entering into the expressions (4.4.6) and (5.5.15). Moreover, interpreted as the *Pitman-efficiency*, it also depends on the level of significance ϵ. Thus, unlike the univariate case, no single measure of efficiency (independent of $\theta^{(k)}$, $k = 1, \ldots, c$ and ϵ) usually exists (unless the p variates in $F(\boldsymbol{x})$ are independent or totally symmetric), and we may have to be satisfied with the assessment of various bounds for the A.R.E. (for any specified $F(\boldsymbol{x})$, which are independent of $\boldsymbol{\theta}^{(k)}$, $k = 1, \ldots, c$. In the particular case of $F(\boldsymbol{x})$ being a p variate normal distribution with a null mean vector, unit variances and a correlation matrix $\boldsymbol{\rho} = (\rho_{ij})$, it follows from well-known results that

$$(6.1) \qquad \psi_{ij} = (6/\pi)\sin^{-1}({}_{ij}/2),\ \phi_{ij} = (2/\pi)\sin^{-1}(\rho_{ij}), \quad \text{for } i, j = 1, \ldots, p;$$

and hence, the A.R.E. becomes a function of $\boldsymbol{\theta}^{(k)}$, $k = 1, \ldots, c$ and $\boldsymbol{\rho}$. In the particular case of $p = 2$, it has been shown by the present authors [7] that the A.R.E. of the rank-sum test with respect to the median test is uniformly greater than unity (i.e., $\geqslant 1$, for all $\boldsymbol{\theta}^{(k)}$, $k = 1, 2$ and ϵ). In the multivariate case of $p \geqslant 3$, really the bounds for this A..R.E. depend on the characteristic roots of $\boldsymbol{\Psi}\boldsymbol{\Phi}^{-1}$. Some result of this type has been considered by Bickel (1964) in the multivariate one sample problem, and it appears that the same bounds are also applicable in our case. We therefore omit these results. Finally, comparison of the rank-sum test or the median test with the parametrically optimum test (viz., the likelihood ratio test) for multinormal parent distribution, requires the comparison of the noncentrality parameters in (4.4.6) or (5.5.15) with that of

$$(6.2) \qquad \Delta_H = \sum_{k=1}^{c} \nu_k \sum_{i=1}^{p} \sum_{j=1}^{p} \rho^{ij}(\theta_i^{(k)} - \bar{\theta}_i)(\theta_j^{(k)} - \bar{\theta}_j),\ (\rho^{ij}) = \boldsymbol{\rho}^{-1}.$$

For the case of $p = 2$, the present authors [7] have shown that the A.R.E. of either of the two proposed tests with respect to the Hotelling's T^2 test (equivalent to the likelihood ratio test) is uniformly less than one, (for bivariate normal distributions). Bounds similar to the one given in Bickel's one sample paper, can again be considered for the multisample case when $p > 2$.

7. CONCLUDING REMARKS

In this paper, we have considered the p variate c sample case, for $p,\ c \geqslant 2$. For the case of p variates and 2 samples, the expressions for the two statistics in (4.1.15) and (5.2.8) can be simplified as

$$(7.1)\qquad W_N = \frac{12}{(N+1)n_1n_2}\sum_{i=1}^{p}\sum_{j=1}^{p}\psi^{ij}(\mathbf{\Lambda}_N)\left\{R_i-\frac{n_1(N+1)}{2}\right\}\left\{R_j-\frac{n_1(N+1)}{2}\right\}$$

$$M_N = \frac{4N}{n_1n_2}\sum_{i=1}^{p}\sum_{j=1}^{p} d^{ij}\{n^*_{1(i)}-\tfrac{1}{2}n_1\}\{n^*_{1(j)}-\tfrac{1}{2}n_1\},$$

where R_i is the sum of ranks (with respect to the ith variate) of the first sample observations, for $i = 1, \ldots, p$. Further, if $p = 2$, these expressions can be more simplified as in [7].

The two nonparametric tests developed in this paper have an important role in multivariate analysis. Not only the assumption of multinormality of the parent distribution is waved here, but also the scope of the inference procedures is increased to data where the observations may be available on an ordinal scale, (thus creating much difficulties to the applicability of usual methods); in the field of Psychometry there are numerous instances of this type of data. The authors believe that the proposed tests are also more robust than the usual parametric tests, though the detailed study of this aspect is appreciably computer-dependent and is left to such interested readers.

References

[1] Anderson, T. W. (1958). *An Introduction to Multivariate Analysis,* John Wiley, New York.

[2] Andrews, F. C. (1954). "Asymptotic Behaviour of Some Rank Tests for Analysis of Variance," *Ann. Math. Statist.*, **25**, 724–736.

[3] Bennett, B. M. (1962). "On Multivariate Sign Tests," *J. Roy. Statist. Soc., Ser. B.*, **24**, 159–161.

[4] Bennett, B. M. (1964). "A Bivariate Signed-Rank Test," *J. Roy. Statist. Soc. Ser. B.*, **26**, 457–461.

[5] Bickel, P. J. (1964). "On Some Alternative Estimates for Shift in the p-Variate One Sample Problem," *Ann. Math. Statist.*, **35**, 1079–1090.

[6] Chatterjee, S. K. (1966). "A Bivariate Sign Test for Location," *Ann. Math. Statist.*, **37**, 1771–1782.

[7] Chatterjee, S. K. and Sen, P. K. (1964). "Nonparametric Tests for the Bivariate Two Sample Location Problem,' *Calcutta Statist. Assoc. Bull.*, **13**, 18–58.

[8] Chatterjee, S. K. and Sen, P. K. (1965). "Some Nonparametric Tests for the Two Sample Bivariate Association Problem," *Calcutta Statist. Assoc. Bull.*, **14**, 14–35.

[9] Hodges, J. L. (1955). "A Bivariate Sign Test," *Ann. Math. Statist.*, **26**, 102–106.

[10] Hoeffding, W. (1948). "A Class of Statistics with Asymptotically Normal Distribution," *Ann. Math. Statist.*, **19**, 293–325.

[11] Hoeffding, W. (1951). "A Combinatorial Central Limit Theorem," *Ann. Math. Statist.*, **22**, 558–566.

[12] Hoeffding, W. (1952). "The Large Sample Power of Tests Based on Permutations of Observations," *Ann. Math. Statist.*, **23**, 169–192.

[13] Kruskal, W. H. and Wallis, W. A. (1952). "Use of Ranks in One Criterion Analysis of Variance," *J. Amer. Statist. Assoc.*, **47**, 583–621.

[14] Mood, A. M. (1941). "On the Joint Distribution of the Medians in Samples From a Multivariate Population," *Ann. Math. Statist.*, **12**, 268–278.

[15] Mood, A. M. (1950). *Introduction to the Theory of Statistics*, McGraw-Hill, New York.

[16] Mood, A. M. (1954). "On the Asymptotic Efficiency of Certain Nonparametric Two Sample Tests," *Ann. Math. Statist.*, **25**, 514–522.

[17] Roy, S. N. (1942). "Analysis of Variance for Multivariate Normal Populations; The Sampling Distribution of the Requisite p-Statistic on the Null and Non-Null Hypotheses," *Sankhyā*, **6**, 35–50.

[18] Wald, A. and Wolfowitz, J. (1944). "Statistical Tests Based on Permutations of Observations," *Ann. Math. Statist.*, **15**, 358–372.

(*Received Jan. 1, 1966.*)