ON A CLASS OF MULTIVARIATE MULTISAMPLE RANK-ORDER TESTS

by

PRANAB KUMAR SEN
University of North Carolina
Chapel Hill, North Carolina

and

MADAN LAL PURI
Courant Institute of Mathematical Sciences,
New York University

Institute of Statistical Mimeo Series No. 474

May 1966

---

*Research sponsored by Sloan Foundation Grant for Statistics, U. S. Navy Contract Nonr-285(38), and by Research Grant, GM-12868, from the National Institute of Health, Public Health Service.

DEPARTMENT OF BIOSTATISTICS
UNIVERSITY OF NORTH CAROLINA
CHAPEL HILL, NORTH CAROLINA

ON A CLASS OF MULTIVARIATE MULTISAMPLE RANK-ORDER TESTS*

Madan Lal Puri* and Pranab Kumar Sen[1]

Courant Institute of Mathematical Sciences, New York University

and University of North Carolina, Chapel Hill

SUMMARY. A class of rank-order tests for the multivariate several sample location and scale problems is proposed and studied here. The principle of rank-permutation tests by Chatterjee and Sen ([6],[8]) is utilized to make these tests strictly distribution-free, and the well-known Chernoff-Savage Theorem (cf. [9], [11], [16]) is generalized here to the multivariate several sample problems. This takes care of the asymptotic power-properties of the tests.

1. INTRODUCTION

Very recently, some attention has been paid to the development of non-parametric tests in the multivariate several sample problems. The earliest test on this line is the permutation test based on Hotelling's T^2 statistic, proposed and studied by Wald and Wolfowitz [22]. However, this is a value permutation test and is subject to the usual limitations of this type of tests. Chatterjee and Sen ([6], [8]) have extended the idea of permutation tests to rank-permutation tests in the multivariate case and have considered a class of genuinely distribution-free tests for location in the bivariate two sample as well as p-variate c-sample case

*Research sponsored by Sloan Foundation Grant for Statistics, U. S. Navy Contract Nonr-285(38), and by Research Grant, GM-12868, from the National Institute of Health, Public Health Service.

[1]On leave of absence from Calcutta University.

$(p,c \geq 2)$. Anderson [3] has recently considered the application of the principle of statistically equivalent blocks (usually used in setting of tolerance limits) in multivariate non-parametric analysis. His procedure contains a little bit of arbitrariness of the choice of "cutting functions" and moreover, it is really difficult to study the properties of these tests when the null hypothesis is not true, even in the asymptotic case. On the other hand, Bhapkar [4] has considered a class of location tests for the multivariate several sample problem, which is distribution-free only asymptotically. Sen [21] has established the asymptotic power equivalence of one of the tests by Bhapkar [4] and a somewhat similar permutation test by Chatterjee and Sen [8]. A more general class of non-parametric tests for the various types of problems that may usually arise in the multivariate several sample case, has been considered by Sen ([19],[20],[21]). These tests are all permutation tests based on appropriate generalized U-statistics.

However, the above developments may not be regarded as fully adequate. First, mostly the above literature deals with the location problems, while the other problems seem to be somewhat neglected. Second, as in the univariate case, the rank-order tests are supposed to play a very vital role in the above development, and in the multivariate case, practically no work has been on progress on this line. The object of the present investigation is to consider a class of rank-order tests for location and scale in the multivariate multisample case. In this context, the well-known theorem by Chernoff and Savage [9] (also see [11], [16]) is extended here to the p-variate c-sample case, for $p,c \geq 2$. In view of the fact that for multivariate distributions, the distributions of rank-order statistics mostly fail to be strictly non-parametric even under the null hypothesis of the identity of the different distributions, the principle of rank-permutations by Chatterjee and Sen ([6],[8]) has been adopted here to achieve the desired distribution-freeness of the proposed class of tests. The asymptotic properties of these permutation rank-order tests (PROT) are studied here with the aid of the Wald-Wolfowitz-Noether-Hoeffding-Hájek theorems (cf. [12]), and certain stochastic equivalence relations of PROT with the multivariate extensions of Chernoff-Savage type of rank-order tests (CSROT) have also been established.

2. PRELIMINARY NOTIONS

Let $\underset{\sim}{X}_{\alpha}^{(k)} = (X_{1\alpha}^{(k)}, \ldots, X_{p\alpha}^{(k)})$, $\alpha = 1,\ldots,n_k$ be n_k independent vector valued random variables drawn from a p variate universe with a continuous cdf (cumulative distribution function) $F_k(\underset{\sim}{x})$, for $k = 1,\ldots,c$; all these c (≥ 2) samples being independent of each other. It is desired to test the null hypothesis

$$H_0: F_1 \equiv \ldots \equiv F_c \equiv F \tag{2.1}$$

where $F \in \mathcal{F}$, $\mathcal{F}$ being the class of all continuous p-variate cdf's. We are interested in translation and/or scale type of alternatives, which may be posed as follows. Let

$$F_k(\underset{\sim}{x}) = F(\underset{\sim}{x} + \underset{\sim}{\delta}_k) \tag{2.2}$$

where $\underset{\sim}{\delta}_1,\ldots,\underset{\sim}{\delta}_c$ real p-vectors. Then under the translation type of alternatives, we are interested in testing

$$H_0\colon \underset{\sim}{\delta}_1 = \cdots = \underset{\sim}{\delta}_c = \underset{\sim}{0} \tag{2.3}$$

against the set of alternatives that these c p-vectors are not all identical. Again, we let

$$F_k(\underset{\sim}{x}) = F(\underset{\sim}{x}_k^*), \quad \underset{\sim}{x}_k^* = (x_{k1}^*,\ldots,x_{kp}^*)\ , \ k = 1,\ldots,c; \tag{2.4}$$

where $x_{ki}^* = (x_i-\mu_i)/\sigma_i^{(k)}$, $k = 1,\ldots,c$; $i = 1,\ldots,p$; $\underset{\sim}{\mu} = (\mu_1,\ldots,\mu_p)$ is a real p-vector and $\underset{\sim}{\sigma}_k = (\sigma_1^{(k)},\ldots,\sigma_p^{(k)})$; $k = 1,\ldots,c$ are c non-negative p-vectors. Then under scale alternatives, we are interested in testing

$$H_0\colon \underset{\sim}{\sigma}_1 = \cdots = \underset{\sim}{\sigma}_c = \underset{\sim}{I} \tag{2.5}$$

(where $\underset{\sim}{I}$ is the p-vector with unit elements) against the set of alternatives that $\underset{\sim}{\sigma}_1,\ldots,\underset{\sim}{\sigma}_c$ are not all equal.

Let us now combine the c samples into a pooled sample of size

$$N = n_1 + \cdots + n_c \tag{2.6}$$

and denote these N p-variate observations by

$$\underset{\sim}{Z}_\alpha, \quad \alpha = 1,\ldots,N; \qquad \underset{\sim}{Z}_\alpha = (Z_{1\alpha},\ldots,Z_{p\alpha})\ . \tag{2.7}$$

We then arrange the N values $Z_{i\alpha}$, $\alpha = 1,\ldots,N$ in order of magnitude, and then denote by $R_{i\alpha}^{(k)}$ the rank of $X_{i\alpha}^{(k)}$ in this set, for $\alpha = 1,\ldots,n_k$; $k = 1,\ldots,c$; $i = 1,\ldots,p$. By virtue of the assumed continuity of $F_1,\ldots,F_c$ the possibility of ties can be ignored, in probability. Let now

$$\underset{\sim}{R}_{\alpha}^{(k)} = (R_{1\alpha}^{(k)},\ldots,R_{p\alpha}^{(k)})', \quad \alpha = 1,\ldots,n_k;\ k = 1,\ldots,c\ , \tag{2.8}$$

be the N random rank p-tuplets, and we define the collection (-rank) <u>Matrix</u> by

$$\underset{\sim}{R}_N = (\underset{\sim}{R}_1^{(1)},\ldots,\underset{\sim}{R}_{n_c}^{(c)}) \tag{2.9}$$

so that $\underset{\sim}{R}_N$ is a $p \times N$ random matrix:

$$\underset{\sim}{R}_N^{p\times N} = \begin{pmatrix} R_{11}^{(1)} & \cdots & R_{1n_1}^{(1)} & \cdots & R_{1n_c}^{(c)} \\ \vdots & \ddots & \vdots & \ddots & \vdots \\ R_{p1}^{(1)} & \cdots & R_{pn_1}^{(1)} & \cdots & R_{pn_c}^{(c)} \end{pmatrix} \tag{2.10}$$

where each row of this stochastic matrix is a random permutation of the numbers $1,\ldots,N$. Let then $Z_{iN,\alpha}^{(k)} = 1$, if the α^{th} smallest observation of the i^{th} variate values of the combined sample is from the k-th sample, and otherwise let $Z_{iN,\alpha}^{(k)} = 0$ for $\alpha = 1,\ldots,N$; $k = 1,\ldots,c$; $i = 1,\ldots,p$. Let $\{E_{N\alpha}^{(i)},\ \alpha = 1,\ldots,N;\ i = 1,\ldots,p\}$ be given numbers satisfying certain restrictions to be stated below (cf. Section 4). We then consider the pc random variables defined as

$$T_{Ni}^{(k)} = \frac{1}{n_k}\sum_{\alpha=1} E_{N\alpha}^{(i)} Z_{iN,\alpha}^{(k)}\ ; \quad k = 1,\ldots,c;\ i = 1,\ldots,p\ . \tag{2.11}$$

Let us finally define

$$\overline{E}_N^{(i)} = \frac{1}{N}\sum_{\alpha=1}^{N} E_{N\alpha}^{(i)} \; ; \quad i = 1,\ldots,p. \tag{2.12}$$

Then our proposed test is based on a statistic $\mathcal{L}_N$ which is a positive semi-definite quadratic form in the pc stochastic variables $T_{Ni}^{(k)} - \overline{E}_N^{(i)}$, $i = 1,\ldots,p$; $k = 1,\ldots,c$.
To formulate the test criterion in an appropriate way and the rationality of the test procedure, we consider the basic rank permutation principle, which is essentially discussed in detail in ([6],[8]), and hence is just sketched in brief.

Each row of the rank collection matrix $\underset{\sim}{R}_N$ defined in (2.10) is a permutation of the numbers $1,\ldots,N$, and thus $\underset{\sim}{R}_N$ is a stochastic matrix which can have $(N!)^p$ possible realizations. Two collection matrices, say R_N and R_N^*, are said to be permutationally equivalent when it is possible to arrive at one from the other by only a finite number of inversions of its columns. Based on this convention, the collection matrix $\underset{\sim}{R}_N$ defined in (2.10) will be permutationally equivalent to another matrix $\underset{\sim}{R}_N^*$ which has the same column vectors as in $\underset{\sim}{R}_N$ but these are so arranged that the first row of $\underset{\sim}{R}_N^*$ is comprised of the numbers $1,\ldots,N$ in the natural order, i.e.

$$\underset{\sim}{R}_N^* = \begin{pmatrix} 1 & 2 & \cdots & N \\ R_{12}^* & R_{22}^* & \cdots & R_{N2}^* \\ \vdots & \vdots & \vdots & \vdots \\ R_{1p}^* & R_{2p}^* & \cdots & R_{Np}^* \end{pmatrix}.$$

Thus, we have a set of N! collection matrices which are all permutationally equivalent to $\underset{\sim}{R}_N^*$, and let this set be denoted by $S(\underset{\sim}{R}_N^*)$.

Now $\underset{\sim}{R}_N^*$ can have $(N!)^{p-1}$ realizations and the set of all these realizations is denoted by $\underset{\sim}{\mathcal{R}}_N^*$. The probability distribution of $\underset{\sim}{R}_N^*$ over $\underset{\sim}{\mathcal{R}}_N^*$ will evidently depend on the parent distributions, even under H_0 defined by (2.1). However, given a particular realization $\underset{\sim}{R}_N^*$, the conditional distribution of $\underset{\sim}{R}_N$ over the N! permutations of the columns of $\underset{\sim}{R}_N^*$ would be uniform under H_0, i.e. if $\underset{\sim}{r}_N$ is any member of $S(\underset{\sim}{R}_N^*)$, then under H_0,

$$P\left\{\underset{\sim}{R}_N = \underset{\sim}{r}_N \mid S(\underset{\sim}{R}_N^*)\right\} = 1/N! \tag{2.13}$$

whatever the parent cdf F may be. Thus if we consider any test depending explicitly on the collection matrix $\underset{\sim}{R}_N$, and formulate a test function $\phi(\underset{\sim}{R}_N)$ depending on the completely specified permutational probability law (2.13), it follows that such a test will be a similar test for the hypothesis H_0. In the next two sections, we shall consider in detail such permutationally distribution-free rank order tests.

3. PERMUTATION RANK-ORDER TESTS (PROT)

In this section, we shall consider a genuinely distribution-free test. Let us denote by $\mathcal{P}_N$ the permutation probability measure generated by the N! possible permutations of the columns of $\underset{\sim}{R}_N^*$. Then, following a few simple steps, we get

$$E(T_{Ni}^{(k)} - \overline{E}_N^{(i)} \mid \mathcal{P}_N) = 0 \tag{3.1}$$

$$\operatorname{cov}\,(T_{Ni}^{(k)}, T_{Nj}^{(q)} \mid \mathcal{P}_N) = (\delta_{kq}N - n_k)v_{ij}(\underset{\sim}{R}_N^*)\big/ n_k(N-1) \tag{3.2}$$

for $i,j = 1,\ldots,p$; $k,q = 1,\ldots,c$. δ_{kq} is the Kronecker delta, and

$$v_{ij}(\underset{\sim}{R}_N^*) = \frac{1}{N}\sum_{k=1}^{c}\sum_{\alpha=1}^{n_k} E_{N\alpha,i}^{(k)}\, E_{N\alpha,j}^{(k)} - \bar{E}_N^{(i)}\bar{E}_N^{(j)}\,; \quad i,j = 1,\ldots,p \tag{3.3}$$

$E_{N\alpha,i}^{(k)}$ being the value of $E_{Ns}^{(i)}$ associated with the rank $s = R_{i\alpha}^{(k)}$. Following then the same arguments as in ([6],[8]), it seems reasonable to base a permutation test on the set of $p(c-1)$ contrasts

$$T_{Ni}^{(k)} - \bar{E}_N^{(i)}\,; \quad i = 1,\ldots,p;\ k = 1,\ldots,c-1. \tag{3.4}$$

If we denote by $\underset{\sim}{V}(\underset{\sim}{R}_N^*)$ the $p \times p$ covariance matrix with elements $v_{ij}(\underset{\sim}{R}_N^*)$ defined in (3.3), then it can be easily seen that the permutational covariance matrix of the $p(c-1)$ random variables in (3.4) (taken in that order) is

$$\frac{1}{(N-1)}\;\left(\left(\frac{N}{n_k}\,\delta_{kq} - 1\right)\right)_{k,q=1,\ldots,c-1} \otimes \underset{\sim}{V}(\underset{\sim}{R}_N^*), \tag{3.5}$$

where $\otimes$ stands for the Kronecker product of two matrices (cf. [2, p. 347]). Thus, if we assume $\underset{\sim}{V}(R_N^*)$ to be positive definite (for the given $\underset{\sim}{R}_N^*$), and denote its reciprocal by $\underset{\sim}{V}^{-1}(\underset{\sim}{R}_N^*) = ((v^{ij}(\underset{\sim}{R}_N^*)))_{i,j=1,\ldots,p}$ then the reciprocal of the $p(c-1) \times p(c-1)$ matrix in (3.5) comes out as

$$(N-1)\;\left(\left(\frac{n_k}{N}\,\delta_{kq} + \frac{n_k n_q}{n_c}\right)\right)_{k,q=1,\ldots,c-1} \otimes \underset{\sim}{V}^{-1}(\underset{\sim}{R}_N^*)\;. \tag{3.6}$$

Then, using (3.6), we get after some essentially simple steps that $\mathcal{L}_N$ can be taken as

$$\mathcal{L}_N = \sum_{k=1}^{c} n_k \left[(\underset{\sim}{T}_N^{(k)} - \overline{E}_N)' \underset{\sim}{V}^{-1}(\underset{\sim}{R}_N^*)(\underset{\sim}{T}_N^{(k)} - \underset{\sim}{\overline{E}}_N) \right] \tag{3.7}$$

which is a weighted sum of c quadratic forms in $(\underset{\sim}{T}_N^{(k)} - \underset{\sim}{\overline{E}}_N)$, $k = 1,\ldots,c$ with the same discriminant, where

$$\underset{\sim}{T}_N^{(k)} = (T_{N,1}^{(k)},\ldots,T_{N,p}^{(k)}) \ , \quad k = 1,\ldots,c \tag{3.8}$$

and

$$\underset{\sim}{\overline{E}}_N = (\overline{E}_N^{(1)},\ldots,\overline{E}_N^{(p)}) \ . \tag{3.9}$$

From the remark made at the end of Section 2, it follows that given R_N, the permutation distribution of $\mathcal{L}_N$ would be strictly distribution-free under H_0 defined by (2.1), and hence an exact size ε $(0 < \varepsilon < 1)$ test can be constructed using this permutation distribution of $\mathcal{L}_N$. (For details, see [6],[8] which deal specifically with the rank-sum and median tests.) In the above context, we have assumed that $\underset{\sim}{V}(\underset{\sim}{R}_N^*)$ is positive definite. As we shall see afterwards that under very mild conditions on $F \in \mathcal{F}$, this statement holds with a very high probability. However, if $\underset{\sim}{V}(\underset{\sim}{R}_N^*)$ happens to be singular, we can work with the highest order non-singular minor of $\underset{\sim}{V}(\underset{\sim}{R}_N^*)$, and proceed similarly with the corresponding subject of variates.

To apply the above test in practice, one would require to study all the N! possible permuted values of $\mathcal{L}_N$ (under $\mathcal{P}_N$), there being at most $N!/\prod_{k=1}^{c} n_k!$ distinct values of $\mathcal{L}_N$ for any given $\underset{\sim}{R}_N^*$. Naturally, the labor involved in this scheme increases prohibitively with the increase in the sample sizes. So in large samples, we are faced with the problem of approximating the true permutation distribution of $\mathcal{L}_N$ by some simple law and to reduce the computational labor by that. This study would be taken up in the next section.

4. ASYMPTOTIC PERMUTATION DISTRIBUTION OF PROT-STATISTIC.

We assume that for all N, the inequalities

$$0 < \lambda_0 \leq \lambda_N^{(k)} = n_k/N \leq 1 - \lambda_0 < 1 \tag{4.1}$$

for $k = 1,\ldots,c$, hold for some fixed $\lambda_0 \leq 1/c$.

Also we require to impose some regularity conditions on $E_{N\alpha}^{(i)}$ defined in (2.11). Extending the idea of Chernoff-Savage [9] to the multivariate case, we write

$$E_{N\alpha}^{(i)} = J_{N(i)}(\alpha/(N+1)); \ \alpha = 1,\ldots,N \ . \tag{4.2}$$

While $J_{N(i)}$ need be defined only at $1/(N+1),\ldots,N/(N+1)$, we shall find it convenient to extend its domain of definition to (0,1) by letting $J_{N(i)}$ to be constant on $(\alpha/(N+1), (\alpha+1)/N+1))$, $\alpha = 0,1,\ldots,N$.

Let us now define

$$F_{N[i]}^{(k)}(x) = \frac{1}{n_k} \text{ (number of } X_{i\alpha}^{(k)} \leq x); \ i=1,\ldots,p; k=1,\ldots,c. \tag{4.3}$$

$$H_{N[i]}(x) = \sum_{k=1}^{c} \lambda_N^{(k)} F_{N[i]}^{(k)}(x), \ i = 1,\ldots,p. \tag{4.4}$$

Also let

$$F_{N[i,j]}^{(k)}(x,y) = \frac{1}{n_k} \text{ (number of } (X_{i\alpha}^{(k)}, X_{j\alpha}^{(k)}) \leq (x,y)); \quad k=1,\ldots,c. \tag{4.5}$$

$$H_{N[i,j]}(x,y) = \sum_{k=1}^{c} \lambda_N^{(k)} F_{N[i,j]}^{(k)}(x,y) \ ; \quad i \neq j = 1,\ldots,p. \tag{4.6}$$

Finally, we define the marginal cdf of $X_{i\alpha}^{(k)}$ and of $(X_{i\alpha}^{(k)}, X_{j\alpha}^{(k)})$ by $F_{[i]}^{(k)}(x)$ and $F_{[i,j]}^{(k)}(x,y)$ respectively, and define

$$H_{[i]}(x) = \sum_{k=1}^{c} \lambda_N^{(k)} F_{[i]}^{(k)}(x) \; ; \quad i = 1,\ldots,p \tag{4.7}$$

$$H_{[i,j]}(x,y) = \sum_{k=1}^{c} \lambda_N^{(k)} F_{[i,j]}^{(k)}(x,y); \; i \neq j = 1,\ldots,p \tag{4.8}$$

$$F_N^{(k)}(\underset{\sim}{x}) = \frac{1}{n_k} (\text{number of } \underset{\sim}{X}_{\alpha}^{(k)} \leq \underset{\sim}{x}) \; , \; k = 1,\ldots,c \tag{4.9}$$

$$H_N(\underset{\sim}{x}) = \sum_{k=1}^{c} \lambda_N^{(k)} F_N^{(k)}(\underset{\sim}{x}) \tag{4.10}$$

$$H(\underset{\sim}{x}) = \sum_{k=1}^{c} \lambda_N^{(k)} F_k(\underset{\sim}{x}) \; . \tag{4.11}$$

Then for the study of the asymptotic permutation distribution of $\mathcal{L}_N$, we shall make the following assumptions:

(1) $J_{(i)}(H) = \lim_{N \to \infty} J_{N(i)}(H)$ exists for $0 < H < 1$ and is not a constant.

(2) $\int_{-\infty}^{\infty} [J_{N(i)}(\frac{N}{N+1} H_{N[i]}) - J_{(i)}(\frac{N}{N+1} H_{N(i)})] \, dF_{N[i]}^{(k)}(x) = o_p(N^{-1/2})$,
for all $k = 1,\ldots,c$ and $i = 1,\ldots,p$.

(3) $J_{(i)}(u)$ is absolutely continuous in u: $0 < u < 1$, and
$$\left|\frac{\partial^r}{\partial u^r} J_{(i)}(u)\right| = |J_{(i)}^{(r)}(u)| \leq K[u(1-u)]^{-r-1/2+\delta},$$
for $r = 0,1$, and some $\delta > 0$, where K is a constant, independent of $i = 1,\ldots,p$.

(4) $$\int_{-\infty}^{\infty}\int_{-\infty}^{\infty} [J_{N(i)}(\tfrac{N}{N+1} H_{N[i]}) J_{N(j)}(\tfrac{N}{N+1} H_{N[j]}) - J_{(i)}(\tfrac{N}{N+1} H_{N[i]}) \, J_{(j)}(\tfrac{N}{N+1} H_{N[j]})] \, dF_{N[i,j]}^{(k)}(x,y) = o_p(1),$$
for all $i \neq j = 1,\ldots,p$ and $k = 1,\ldots,c$.

It may be noted that condition (4) will be satisfied if for all $i = 1,\ldots,p$,

$$\int_{-\infty}^{\infty} [J_{N(i)}(\frac{N}{N+1} H_{N[i]}) - J_{(i)}(\frac{N}{N+1} H_{N[i]})]^2 \, dF_{N[i]}^{(k)}(x) = o_p(1) \quad (4.12)$$

for all $k = 1,\ldots,c$. In practice, it is easier to verify (4.12) than condition (4).

So long we have assumed that $F_k \in \mathcal{F}$, $k = 1,\ldots,c$, where $\mathcal{F}$ is the class of all continuous p-variate cdf's. We will now impose some restriction on $\mathcal{F}$. Let us define

$$\mu_i = \int_0^1 J_i(u)\, du \;; \quad i = 1,\ldots,p \quad (4.13)$$

$$\nu_{ij}(F) = \int_0^1 J_{(i)}^2(u)\, du - \mu_i^2 \,, \quad i=j=1,\ldots,p$$

$$= \int_{-\infty}^{+\infty}\int_{-\infty}^{+\infty} J_{(i)}(F_{[i]}(x))J_{(j)}(F_{[j]}(y))\, dF_{[i,j]}(x,y) - \mu_i\mu_j \quad (4.14)$$

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

and denote

$$\underset{\sim}{\nu}(F) = ((\nu_{ij}(F)))_{i,j=1,\ldots,p}. \quad (4.15)$$

Let $\mathcal{F}_0$ denote the class of all continuous p-variate cdf's for which $\underset{\sim}{\nu}(F)$ is positive definite and assume that

$$F_k(\underset{\sim}{x}) \in \mathcal{F}_0 \quad \text{for all } k = 1,\ldots,c. \quad (4.16)$$

It may be observed that (4.16) implies that

$$H(\underset{\sim}{x}) \in \mathcal{F}_0 \quad (4.17)$$

that is, $\underset{\sim}{\nu}(H)$ is positive definite uniformly in $\lambda_N^{(1)},\ldots,\lambda_N^{(c)}$ satisfying (4.1). Further, it may be noted that conditions (1), (2) and (3) are sufficient for the asymptotic multinormality of the joint permutation distribution of the p(c-1) random variables in (3.4). Condition (4.16) or (4.17) is required to ensure the non-singularity of the above asymptotic distribution, as the same is required, in turn, for the asymptotic distribution of $\mathcal{L}_N$ in (3.7). Condition (4) is required to establish the convergence (in probability) of the (random) permutation covariance matrix $\underset{\sim}{V}(\underset{\sim}{R}_N^*)$ to $\underset{\sim}{\nu}(H)$, which will be required in the sequel.

Before we proceed to consider the main theorem of this section, we present the following two theorems, as they will be required subsequently.

THEOREM 4.1. <u>If conditions</u> (1), (2) <u>and</u> (3) <u>hold and if</u> $\underset{\sim}{V}(\underset{\sim}{R}_N^*)$ <u>is asymptotically non-singular, then the joint permutation distribution of the</u> p(c-1) <u>random variables</u> $\{T_{Ni}^{(k)}-\bar{E}_N^{(i)};\ i = 1,\ldots,p;\ k = 1,\ldots,c-1\}$ <u>defined in</u> (2.11) <u>and</u> (2.12) <u>is asymptotically a</u> p(c-1)-<u>variate normal distribution</u>.

<u>PROOF</u>. It is sufficient to show that under the stated regularity conditions, any arbitrary linear function of $\underset{\sim}{T}_N^{(1)},\ldots,\underset{\sim}{T}_N^{(c-1)}$ (defined in (3.8)), when standardized, has a permutation distribution which is asymptotically a p variate normal one. If now $\beta_1,\ldots,\beta_{c-1},\beta_c$ (=0) be any c-1 arbitrary constants, then

$$\sum_{k=1}^{c} \beta_k \underset{\sim}{T}_N^{(k)} = \underset{\sim}{U}_N = (U_{N1},\ldots,U_{Np})', \tag{4.18}$$

where

$$U_{Ni} = \sum_{\alpha=1}^{N} c_{N\alpha}^{(i)} E_{N\alpha}^{(i)};\quad i = 1,\ldots,p; \tag{4.19}$$

$$c_{N\alpha}^{(i)} = \sum_{k=1}^{c} \beta_k Z_{iN,\alpha}^{(k)}\ ,\ \alpha = 1,\ldots,N;\quad i = 1,\ldots,p. \tag{4.20}$$

It is then easily shown that for any given $\beta_1,\ldots,\beta_c$; $\underset{\sim}{c}_N^{(i)} = (c_{N1}^{(i)},\ldots,c_{NN}^{(i)})$ satisfies the Wald-Wolfowitz condition ([10],p.236) of the Wald-Wolfowitz-Noether-Hoeffding-Hajek theorem ([12],p.522) for each $i = 1,\ldots,p$. Also, using the conditions (1), (2) and (3), it is readily seen that $\underset{\sim}{E}_N^{(i)} = (E_{N1}^{(i)},\ldots,E_{NN}^{(i)})$ satisfies the Noether-Hoeffding condition ([10],p.236) of the same theorem for each $i = 1,\ldots,p$. Hence, on applying the above mentioned theorem (as extended to the vector-case by Hajek ([12],p. 522) under the assumed non-singularity condition on $\underset{\sim}{V}(\underset{\sim}{R}_N^*))$, the desired result follows readily.

Hence, the theorem.

THEOREM 4.2. <u>If the conditions</u> (1), (2), (3) <u>and</u> (4) <u>hold, then for all</u> $F_k \in \mathcal{F}$, $k = 1,\ldots,c$, <u>and uniformly in</u> $\lambda_N^{(1)},\ldots,\lambda_N^{(c)}$ <u>(satisfying</u> (4.1))

$$\underset{\sim}{V}(\underset{\sim}{R}_N^*) \xrightarrow{P} \underset{\sim}{v}(H) \ .$$

(Note: We shall throughout use the notation $\xrightarrow{P}$ to mean convergence in probability as $N \to \infty$.)

<u>PROOF</u>. By virtue of conditions (2) and (3), and the well-known results in the univariate case (cf. [9]), it follows that

$$\bar{E}_N^{(i)} \xrightarrow{P} \int_0^1 J_{(i)}(u)\, du = \mu_i \quad \text{for all } i = 1,\ldots,c. \tag{4.21}$$

Hence, using the condition (4), it follows from (3.3) and (4.21) that

$$V_{ij}(\underset{\sim}{R}_N^*) + \mu_i\mu_j = \iint_{R^2} J_{(i)}(\tfrac{N}{N+1} H_{N[i]}) \; J_{(j)}(\tfrac{N}{N+1} H_{N[j]}) dH_{N[i,j]}(x,y) + o_p(1) \tag{4.22}$$

where R^2 denote the two dimensional real space. So our problem reduces to that of showing that the integral on the right hand side of (4.22) converges in probability to

$$A = \iint_{R^2} J_{(i)} \; (H_{[i]}) \; J_{(j)}(H_{[j]}) \; dH_{[i,j]}(x,y) \; . \tag{4.23}$$

Let us now define

$$\begin{aligned} I_{N(i)} &= \left\{x\colon \; N^{-1/3} \le H_{[i]}(x) \le 1 - N^{-1/3}\right\}; \; i = 1,\ldots,p \\ I_{N(i,j)} &= I_{N(i)} \cap I_{N(j)} \; , \qquad\qquad j = 1,\ldots,p. \end{aligned} \tag{4.24}$$

Then, using Cauchy-Schwarz inequality, (4.24) and the well-known result (cf. [9]) that for each $i = 1,\ldots,p$,

$$\sup_x N^{1/2}|\tfrac{N}{N+1} H_{N[i]}(x) - H_{[i]}(x)| \text{ is bounded in probability,} \tag{4.25}$$

we can write the right hand side of (4.22) as

$$A_N + B_{1N} + B_{2N} + o_p(1) \tag{4.26}$$

where

$$A_N = \iint_{I_{N(i,j)}} J_{(i)}(H_{[i]})J_{(j)}(H_{[j]}) \; dH_{N[i,j]}(x,y) \tag{4.27}$$

$$B_{1N} = \iint_{I_{N(i,j)}} J_{(i)}(\tfrac{N}{N+1} H_{N[i]})[J_{(j)}(\tfrac{N}{N+1} H_{N[j]}) - J_{(j)}(H_{[j]})] \; dH_{N[i,j]}(x,y) \tag{4.28}$$

$$B_{2N} = \iint_{I_{N(i,j)}} J_{(j)}(H_{[j]})[J_{(i)}(\tfrac{N}{N+1} H_{N[i]}) - J_{(i)}(H_{[i]})] \; dH_{N[i,j]}(x,y). \tag{4.29}$$

Now, using (4.25) and the condition (3), it is easily seen that for all points in $I_{N(i,j)}$

$$
\begin{aligned}
&|J_{(i)}(\tfrac{N}{N+1} H_{N[i]}) - J_{(i)}(H_{[i]})| \\
&= \left\{ N^{1/2} |\tfrac{N}{N+1} H_{N[i]} - H_{[i]}| \right\} \left\{ N^{-1/2} |J'_{(i)}(\phi \tfrac{N}{N+1} H_{N[i]} + \right. \\
&\quad \left. + (1-\phi) H_{[i]})| \right\}, \quad 0 < \phi < 1 \qquad (4.30) \\
&= o_p(1)
\end{aligned}
$$

uniformly in $x \in I_{N(i,j)}$. Also, noting that

$$\int_{I_{N(i)}} |J_{(i)}(\tfrac{N}{N+1} H_{N[i]})| \, dH_{N[i]}(x) < \infty \qquad (4.31)$$

for all $i = 1, \ldots, p$ (by condition (3)), we obtain from (4.28), (4.30) and (4.31) that

$$|B_{1N}| = o_p(1) \ . \qquad (4.32)$$

Similarly, it is easily seen that under the stated regularity conditions

$$\int_{I_{N(i)}} |J_{(i)}(H_{[i]}(x))| \, dH_{N[i]}(x) \qquad (4.33)$$

is bounded in probability for all $i = 1, \ldots, p$. Hence, from (4.29), (4.30) and (4.33) it follows that

$$|B_{2N}| = o_p(1) \ . \qquad (4.34)$$

Finally, A_N in (4.27) can be written as

$$
\begin{aligned}
&\sum_{k=1}^{c} \lambda_N^{(k)} \iint_{I_{N(i,j)}} J_{(i)}(H_{[i]}(x)) \, J_{(j)}(H_{[j]}(y)) \, dF_{N[i,j]}^{(k)}(x,y) \\
&= \sum_{k=1}^{c} \lambda_N^{(k)} \left\{ \frac{1}{n_k} \sum_{\alpha=1}^{n_k} g_N(X_{i\alpha}^{(k)}, X_{j\alpha}^{(k)}) \right\} \qquad (4.35)
\end{aligned}
$$

where

$$g_N(x,y) = J_{(i)}(H_{[i]}(x))J_{(j)}(H_{[j]}(y)) \quad \text{if } (x,y) \in I_{N(i,j)}$$
$$= 0 \qquad\qquad \text{otherwise} \tag{4.36}$$

and, where $\left\{(X_{i\alpha}^{(k)}, X_{j\alpha}^{(k)});\ \alpha = 1,\ldots,n_k\right\}$ are independent and identically distributed random variables (i.i.d.i.v.) having the bivariate cdf $F_{[i,j]}^{(k)}(x,y)$. Now by virtue of condition (3), it is easily seen (by using the elementary inequalities of ([9],p. 986)) that

$$\sup_N E\left\{|g_N(X_{i\alpha}, Y_{j\alpha})|^{1+\delta}\ |F_{[i,j]}^{(k)}\right\} < \infty \quad \text{for } k=1,\ldots,c, \tag{4.37}$$

where δ is a positive quantity defined in condition (3). Also, $\left\{g_N(X_{i\alpha}, X_{j\alpha}),\ \alpha = 1,\ldots,n_k\right\}$ are n_k independent and identically distributed random variables, and hence it can easily be shown by using (4.37) that

$$\frac{1}{n_k}\sum_{\alpha=1}^{n_k} g_N(X_{i\alpha}^{(k)},\ X_{j\alpha}^{(k)}) \xrightarrow{P} \iint_{R^2} J_{(i)}(H_{[i]}(x))J_{(j)}(H_{[j]}(y))\ dF_{[i,j]}^{(k)}(x,y) \tag{4.38}$$

for $k = 1,\ldots,c$.

Now from (4.8), (4.35) and (4.38) it follows that

$$A_N \xrightarrow{P} A \tag{4.39}$$

where A is given by (4.23).

From (4.22), (4.26), (4.32), (4.34) and (4.39), we arrive at the result

$$v_{ij}(\underset{\sim}{R}_N^*) \longrightarrow \nu_{ij}(H)\ , \qquad i \neq j = 1,\ldots,p.$$

Similarly, it can be shown precisely on the same lines as in the univariate case that $v_{ii}(\underset{\sim}{R}_N^*)$ is independent of $\underset{\sim}{R}_N^*$ and

$$v_{ii}(\underset{\sim}{R}_N^*) \xrightarrow{P} v_{ii}(H) \ , \quad i = 1,\dots,p.$$

Hence the theorem.

THEOREM 4.3. _If conditions_ (1), (2), (3) _and_ (4) _hold and if_ $F_k \in \mathcal{F}_0$ _for all_ $k = 1,\dots,c$, _then the permutation distribution of the statistic_ $\mathcal{L}_N$ _(defined by_ (3.7)) _asymptotically reduces to a chi-square distribution with_ $p(c-1)$ _degrees of freedom_.

PROOF. The proof follows more or less trivially from the preceding two theorems, and hence is omitted.

By virtue of Theorem 4.3, the permutation test procedure based on $\mathcal{L}_N$ simplifies in large samples to the following rule.

$$\text{If } \mathcal{L}_N \begin{cases} \geq \chi^2_{\varepsilon,p(c-1)} \ , & \text{reject } H_0, \\ < \chi^2_{\varepsilon,p(c-1)} \ , & \text{accept } H_0 \ , \end{cases} \tag{4.40}$$

where $\chi^2_{\varepsilon,r}$ is the $100(1-\varepsilon)\%$ point of a chi-square distribution with r d.f. (degrees of freedom).

In order to study the power properties of the test considered above, we require to study the unconditional distribution of $\mathcal{L}_N$, under appropriate classes of alternative hypotheses. This, in turn, requires the study of the joint distribution of the rank order statistics defined by (2.11) and the same will be considered in the next section. The problem of the choice of suitable $\{\underset{\sim}{E}_N^{(i)},\ i = 1,\dots,p\}$, will be considered in a later section.

5. ASYMPTOTIC MULTINORMALITY OF $T_{Ni}^{(k)}$; $k = 1,\ldots,c$, $i = 1,\ldots,p$, FOR ARBITRARY $F_1,\ldots,F_c$.

It follows from (2.11), (4.2) and (4.4) that we can write $T_{Ni}^{(k)}$ equivalently as

$$T_{Ni}^{(k)} = \int_{-\infty}^{+\infty} J_{N(i)}\left(\frac{N}{N+1} H_{N[i]}(x)\right) dF_{N[i]}^{(k)}(x) \quad . \tag{5.1}$$

THEOREM 5.1. *If* $F_k \in \mathcal{F}$ *for all* $k = 1,\ldots,c$, *and if* $J_{N(i)}$ *satisfies the conditions* (1), (2) *and* (3) (cf. *Section* 4) *for all* $i = 1,\ldots,p$, *then the* pc-*vector with stochastic elements* $[N^{1/2}(T_{Ni}^{(k)} - \mu_{Ni}^{(k)}),\ i = 1,\ldots,p;\ k = 1,\ldots,c]$ *has a limiting normal distribution with zero mean vector and covariance matrix* $\underset{\sim}{\Sigma} = ((\sigma_{N[i,j]}^{(k,q)}))$, *where*

$$\mu_{Ni}^{(k)} = \int_{-\infty}^{+\infty} J_{(i)}(H_{[i]}(x))\, dF_{[i]}^{(k)}(x) \ ; \ i = 1,\ldots,p;\ k=1,\ldots,c. \tag{5.2}$$

$$\begin{aligned}\sigma_{N[i,j]}^{(k,q)} = {} & 2 \sum_{\substack{r=1\\ r\neq k}}^{c} \lambda_N^{(r)} \iint_{-\infty<x<y<\infty} A_i^{(r)}(x,y)\, dF_{[i]}^{(r)}(x)\, dF_{[i]}^{(r)}(y) \\ & + \frac{2}{\lambda_N^{(k)}} \sum_{\substack{r=1\\ r\neq k}}^{c} [\lambda_N^{(r)}]^2 \iint_{-\infty<x<y<\infty} A_i^{(k)}(x,y)\, dF_i^{[r]}(x)\, dF_{[i]}^{(r)}(y) \\ & + \frac{1}{\lambda_N^{(k)}} \sum_{\substack{r=1\\ r\neq s, r\neq k}}^{c} \sum_{\substack{s=1\\ s\neq k}}^{c} \lambda_N^{(r)} \lambda_N^{(s)} \Bigg[\iint_{-\infty<x<y<\infty} A_i^{(k)}(x,y)\, dF_{[i]}^{(r)}(x)\, dF_{[i]}^{(s)}(y) \\ & \qquad + \iint_{-\infty<y<x<\infty} A_i^{(k)}(y,x)\, dF_{[i]}^{(r)}(x)\, dF_{[i]}^{(s)}(y) \Bigg] \end{aligned} \tag{5.3}$$

(if $k = q$, $i = j$);

$$= \sum_{r=1}^{c} \sum_{s=1}^{c} \frac{\lambda_N^{(r)} \lambda_N^{(s)}}{\lambda_N^{(k)}} \int_{-\infty}^{+\infty} \int_{-\infty}^{+\infty} B_{i,j}^{(k)}(x,y)\ dF_{[i]}^{(r)}(x)\ dF_{[j]}^{(s)}(y) \tag{5.4}$$

$$- \sum_{r=1}^{c} \lambda_N^{(r)} \int_{-\infty}^{+\infty} \int_{-\infty}^{+\infty} B_{i,j}^{(k)}(x,y)\ dF_{[i]}^{(r)}(x)\ dF_{[j]}^{(k)}(y)$$

$$- \sum_{r=1}^{c} \lambda_N^{(r)} \int_{-\infty}^{+\infty} \int_{-\infty}^{+\infty} B_{i,j}^{(k)}(x,y)\ dF_{[i]}^{(k)}(x)\ dF_{[j]}^{(r)}(y)$$

$$+ \sum_{r=1}^{c} \lambda_N^{(r)} \int_{-\infty}^{+\infty} \int_{-\infty}^{+\infty} B_{i,j}^{(r)}(x,y)\ dF_{[i]}^{(k)}(x)\ dF_{[j]}^{(k)}(y)$$

(if $k = q$, $i \neq j$);

$$= - \sum_{r=1}^{c} \lambda_N^{(r)} \Bigg[\iint\limits_{-\infty<x<y<\infty} A_i^{(k)}(x,y)\ dF_{[i]}^{(r)}(x)\ dF_{[i]}^{(q)}(y) \tag{5.5}$$

$$+ \iint\limits_{-\infty<y<x<\infty} A_i^{(k)}(y;x)\ dF_{[i]}^{(r)}(x)\ dF_{[i]}^{(q)}(y) \Bigg]$$

$$- \sum_{r=1}^{c} \lambda_N^{(r)} \Bigg[\iint\limits_{-\infty<x<y<\infty} A_i^{(q)}(x,y)\ dF_{[i]}^{(r)}(x)\ dF_{[i]}^{(k)}(y)$$

$$+ \iint\limits_{-\infty<y<x<\infty} A_i^{(q)}(y,x)\ dF_{[i]}^{(r)}(x)\ dF_{[i]}^{(k)}(y) \Bigg]$$

$$+ \sum_{r=1}^{c} \lambda_N^{(r)} \Bigg[\iint\limits_{-\infty<x<y<\infty} A_i^{(r)}(x,y)\ dF_{[i]}^{(k)}(x)\ dF_{[i]}^{(q)}(y)$$

$$+ \iint\limits_{-\infty<y<x<\infty} A_i^{(r)}(y,x)\ dF_{[i]}^{(k)}(x)\ dF_{[i]}^{(q)}(y) \Bigg]$$

(if $k \neq q$, $i = j$);

$$= -\sum_{r=1}^{c} \lambda_N^{(r)} \int_{-\infty}^{+\infty}\int_{-\infty}^{+\infty} B_{i,j}^{(k)}(x,y)\, dF_{[i]}^{(r)}(x)\, dF_{[j]}^{(q)}(y) \tag{5.6}$$

$$- \sum_{r=1}^{c} \lambda_N^{(r)} \int_{-\infty}^{+\infty}\int_{-\infty}^{+\infty} B_{i,j}^{(q)}(x,y)\, dF_{[i]}^{(k)}(x)\, dF_{[j]}^{(r)}(y)$$

$$+ \sum_{r=1}^{c} \lambda_N^{(r)} \int_{-\infty}^{+\infty}\int_{-\infty}^{+\infty} B_{i,j}^{(r)}(x,y)\, dF_{[i]}^{(k)}(x)\, dF_{[j]}^{(q)}(y)$$

(if $k \neq q$, $i \neq j$);

where

$$A_i^{(\alpha)}(t,u) = F_{[i]}^{(\alpha)}(t)\,[1- F_{[i]}^{(\alpha)}(t)]\, J'_{(i)}[H_{[i]}(t)]\, J'_{(i)}[H_{[i]}(u)] \tag{5.7}$$

$$B_{i,j}^{(\alpha)}(t,u) = [F_{[i,j]}^{(\alpha)}(t,u) - F_{[i]}^{(\alpha)}(t)F_{[j]}^{(\alpha)}(u)]\, J'_{(i)}[H_{[i]}(t)]J'_j[H_{[j]}(u)] \tag{5.8}$$

<u>Proof</u>: As in the univariate case (cf. [11, Theorem 5.1]) we rewrite $T_{Ni}^{(k)}$ as

$$T_{Ni}^{(k)} = \mu_{Ni}^{(k)} + B_{1N(i)}^{(k)} + B_{2N(i)}^{(k)} + \sum_{r=1}^{4} C_{rN(i)}^{(k)} \tag{5.9}$$

where $\mu_{Ni}^{(k)}$ is given by (5.2),

$$B_{1N(i)}^{(k)} = \int J_{(i)}(H_{[i]}(x))\, d[F_{N[i]}^{(k)}(x) - F_{[i]}^{(k)}(x)] \tag{5.10}$$

$$B_{2N(i)}^{(k)} = \int [H_{N[i]}(x) - H_{[i]}(x)]\, J'_{(i)}[H_{(i)}(x)]\, dF_{[i]}^{(k)}(x) \tag{5.11}$$

and the $C_{rN(i)}^{(k)}$, $r = 1,2,3,4$ are all $o_p(N^{-1/2})$; for $i = 1,\dots,p$ and $k = 1,\dots,c$. The difference $N^{1/2}(T_{Ni}^{(k)} - \mu_{Ni}^{(k)}) - N^{1/2}(B_{1N(i)}^{(k)} + B_{2N(i)}^{(k)})$ tends to zero in probability for all $i = 1,\ldots,p$ and $k = 1,\ldots,c$ and so the vectors

$$[N^{1/2}(T_{Ni}^{(k)} - \mu_{Ni}^{(k)}),\ i = 1,\ldots,p;\ k = 1,\ldots,c]$$

and

$$[N^{1/2}(B_{1N(i)}^{(k)} + B_{2N(i)}^{(k)}),\ i = 1,\ldots,p;\ k = 1,\ldots,c]$$

have the same limiting distribution if they have one at all. Thus, to prove this theorem, it suffices to show that for any real non-null $\delta = (\delta_1,\ldots,\delta_c)$, the random vector

$$N^{1/2}\sum_{k=1}^{c}\delta_k[\underset{\sim}{B}_{1N}^{(k)} + \underset{\sim}{B}_{2N}^{(k)}] \tag{5.12}$$

(where $\underset{\sim}{B}_{rN}^{(k)} = (B_{rN(1)}^{(k)},\ldots,B_{rN(p)}^{(k)})'$, $r = 1,2$), has asymptotically a p-variate normal distribution.

Now. proceeding as in the univariate c-sample case (cf. [16] and [[11], Section 5]), we can rewrite $B_{1N(i)}^{(k)} + B_{2N(i)}^{(k)}$ as

$$\sum_{\substack{q=1\\ \neq k}}^{c}\lambda_N^{(q)}\left\{\frac{1}{N_q}\sum_{\alpha=1}^{n_q}B_{kq(1)}^{(i)}(X_{i\alpha}^{(q)}) - \frac{1}{n_k}\sum_{\alpha=1}^{n_k}B_{kq(2)}^{(i)}(X_{i\alpha}^{(k)})\right\} \tag{5.13}$$

where

$$B_{kq(1)}^{(i)}(X_{i\alpha}^{(q)}) = \int_{-\infty}^{\infty}[F_{1[i]}^{(q)} - F_{(i)}^{(q)}]J'_{(i)}(H_{[i]})\,dF_{[i]}^{(k)}(x) \tag{5.14}$$

$$B_{kq(2)}^{(i)}(X_{i\alpha}^{(k)}) = \int_{-\infty}^{\infty}[F_{1[i]}^{(k)} - F_{[i]}^{(k)}]J'_{(i)}(H_{[i]})\,dF_{[i]}^{(q)}(x) \tag{5.15}$$

$F_{1[i]}^{(k)}$ being the empirical cdf of $X_{i\alpha}^{(k)}$, whose true cdf is $F_{[i]}^{(k)}(x)$, for $q,k = 1,\ldots,c$; $i = 1,\ldots,p$. If we now define

$$B_{0k}^{(i)}(X_{i\alpha}^{(k)}) = \sum_{\substack{q=1\\ \neq k}}^{c} \delta_q \, B_{qk(1)}^{(k)}(X_{i\alpha}^{(k)}) \tag{5.15}$$

$$B_{k0}^{(i)}(X_{i\alpha}^{(k)}) = \sum_{\substack{q=1\\ \neq k}}^{c} \lambda_N^{(q)} \, B_{kq(2)}^{(i)}(X_{i\alpha}^{(k)}) \tag{5.16}$$

$$B_{k}^{(i)}(X_{i\alpha}^{(k)}) = B_{0k}^{(i)}(X_{i\alpha}^{(k)}) - \frac{\delta_k}{\lambda_N^{(k)}} B_{k0}^{(i)}(X_{i\alpha}^{(k)}) \tag{5.17}$$

and

$$B_N^{(i)} = \frac{1}{N} \sum_{k=1}^{c} \sum_{\alpha=1}^{n_k} B_k^{(i)}(X_{i\alpha}^{(k)}) \;, \quad i = 1,\ldots,p, \tag{5.18}$$

$$\underset{\sim}{B}_N = (B_N^{(1)},\ldots,B_N^{(p)})' \tag{5.19}$$

then (5.12) can be written as $N^{1/2}\,\underset{\sim}{B}_N$, which is a stochastic p-vector with elements $B_N^{(i)}$ $(i = 1,\ldots,p)$ which involve c independent sums of independent and identically distributed random variables $B_k^{(i)}(X_{i\alpha}^{(k)})$, $\alpha = 1,\ldots,n_k$; $k = 1,\ldots,c$. Further, as in the univariate case (cf. [11, Section 5]) it can be easily shown that

$$E\left\{ |B_k^{(i)}(X_{i\alpha}^{(k)})|^{2+\delta} \right\} < \infty \tag{5.20}$$

for all $i = 1,\ldots,p$; $k = 1,\ldots,c$, where δ is defined under Condition 3 in Section 4.

Hence, the asymptotic multinormality of (5.12) readily follows from (5.18), (5.19) and (5.20) through an application of the (vector case) Central Limit theorem.

Further, it can be shown by somewhat lengthy algebraic manipulations that

$$\operatorname{cov}\left\{ B^{(i)}_{kq(1)}(X^{(q)}_{i\alpha}),\ B^{(j)}_{k'q'(1)}(X^{(q')}_{j\alpha}) \right\} \tag{5.21}$$

$$= 0 \qquad \text{if } q \neq q' ,$$

$$= \int_{-\infty}^{+\infty}\int_{-\infty}^{+\infty} B^{(q)}_{i,j}(x,y)\ dF^{(k)}_{[i]}(x)\ dF^{(k')}_{[j]}(y) \quad \text{if } q = q';\ i \neq j;$$

$$= \iint\limits_{-\infty<x<y<\infty} A^{(q)}_{i}(x,y)\ dF^{(k)}_{[i]}(x)\ dF^{(k')}_{[i]}(y)$$

$$+ \iint\limits_{-\infty<y<x<\infty} A^{(q)}_{i}(y,x)\ dF^{(k)}_{[i]}(x)\ dF^{(k')}_{[i]}(y) \quad \text{if } q = q',\ i = j,$$

where $A^{(q)}_{i}(x,y)(A^{(q)}_{i}(y,x))$ and $B^{(q)}_{i,j}(x,y)$ are given by (5.7) and (5.8) respectively.

$$\operatorname{cov}\left\{ B^{(i)}_{kq(2)}(X^{(k)}_{i\alpha}),\ B^{(j)}_{k'q'(2)}(X^{(k')}_{j\alpha}) \right\} \tag{5.22}$$

$$= 0 \qquad \text{if } k \neq k'$$

$$= \int_{-\infty}^{+\infty}\int_{-\infty}^{+\infty} B^{(k)}_{i,j}(x,y)\ dF^{(q)}_{[i]}(x)\ dF^{(q')}_{[j]}(y) \quad \text{if } k = k';\ i \neq j;$$

$$= \iint\limits_{-\infty<x<y<\infty} A^{(k)}_{i}(x,y)\ dF^{(q)}_{[i]}(x)\ dF^{(q')}_{[i]}(y)$$

$$+ \iint\limits_{-\infty<y<x<\infty} A^{(k)}_{i}(y,x)\ dF^{(q)}_{[i]}(x)\ dF^{(q')}_{[i]}(y) \quad \text{if } k = k';\ i = j.$$

Finally

$$\operatorname{cov}\left\{ B_{kq(1)}^{(i)}(X_{i\alpha}^{(k)}),\ B_{k'q'(2)}^{(j)}(X_{j\alpha}^{(k')}) \right\} \tag{5.23}$$

$$= 0, \qquad \text{if } q \neq k';$$

$$= \int_{-\infty}^{+\infty}\int_{-\infty}^{+\infty} B_{i,j}^{(q)}(x,y)\ dF_{[i]}^{(k)}(x)\ dF_{[j]}^{(q')}(y) \qquad \text{if } q = k';\ i \neq j;$$

$$= \iint_{-\infty<x<y<\infty} A_i^{(q)}(x,y)\ dF_{[i]}^{(k)}(x)\ dF_{[i]}^{(q')}(y)$$

$$+ \iint_{-\infty<y<x<\infty} A_i^{(q)}(y,x)\ dF_{[i]}^{(k)}(x)\ dF_{[i]}^{(q')}(y) \quad \text{if } q = k';\ i = j.$$

Using (5.13), (5.14), (5.15), (5.21), (5.22) and (5.23) and following somewhat lengthy algebraic computations we obtain the covariance terms given by (5.3), (5.4), (5.5) and (5.6). The theorem follows.

It may be noted that the asymptotic normal distribution derived in Theorem 5.1 is singular and the rank of this distribution can at most be equal to p(c-1). If the null hypothesis (2.1) is true, then it is readily seen that

$$\sigma_{N[i,j]}^{(k,q)} = (\delta_{kq}/\lambda_N^{(k)} - 1)\nu_{ij}(F), \tag{5.24}$$

where δ_{kq} is the kronecker delta and

$$\nu_{ij}(F) = \begin{cases} \int_0^1 J_{(i)}^2(x)\ dx - [\int_0^1 J_{(i)}(x)\ dx]^2 & \text{if } i = j \\ \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} J_{(i)}(F_{[i]}(x))J_{(j)}(F_{[j]}(y))\ dF_{[i,j]}(x,y) & \\ \quad - [\int_0^1 J_{(i)}(x)\ dx]\ [\int_0^1 J_{ij}(y)\ dy] & \text{if } i \neq j \end{cases} \tag{5.25}$$

$F_{[i]}$ and $F_{[i,j]}$ being the marginal cdf's X_i and (X_i, X_j) respectively when $X = (X_1, \ldots, X_p)$ has the cdf $F(\underset{\sim}{x})$.

Thus, if we assume that $F \in \mathcal{F}_0$, where $\mathcal{F}_0$ is the class of all continuous cdf's for which (4.15) is positive definite, then it readily follows that the asymptotic normal distribution in Theorem 5.1 is of the rank p(c-1), and by simple argument, we may conclude that

$$\mathcal{L}_N^* = \sum_{k=1}^{c} n_k[(\underset{\sim}{T}_N^{(k)} - \underset{\sim}{\bar{E}}_N)' \; \underset{\sim}{\nu}^{-1}(F)(\underset{\sim}{T}_N^{(k)} - \underset{\sim}{\bar{E}}_N)] \tag{5.26}$$

where $\underset{\sim}{\nu}^{-1}(F) = ((\nu_{ij}(F))^{-1} = ((\nu^{ij}))$ has under the conditions of Theorem 5.1, asymptotically a chi-square distribution with p(c-1) d.f. when H_0 defined by (2.1) is true.

If now $\hat{\underset{\sim}{\nu}} = ((\hat{\nu}_{ij}))$ is any consistent estimator of $((\nu_{ij}))$ and if $\hat{\underset{\sim}{\nu}}^{-1} = ((\hat{\nu}^{ij}))$, then it follows from (5.26) and some simple arguments that

$$\hat{\mathcal{L}}_N = \sum_{i=1}^{c} n_k[(\underset{\sim}{T}_N^{(k)} - \underset{\sim}{\bar{E}}_N)' \; \hat{\underset{\sim}{\nu}}^{-1}(\underset{\sim}{T}_N^{(k)} - \underset{\sim}{\bar{E}}_N)] \tag{5.27}$$

has also under the conditions of Theorem 5.1, asymptotically a chi-square distribution with p(c-1) d.f. when H_0 defined by (2.1) is true. Hence an asymptotic test may be based on the statistic $\hat{\mathcal{L}}_N$, and such a test will be termed as a CSROT (see the concluding part of Section 1).

6. THE LIMITING DISTRIBUTIONS OF PROT AND CSROT FOR SEQUENCES OF TRANSLATION AND SCALE TYPE OF ALTERNATIVES.

In this section, we shall concern ourselves with a sequence of admissible alternative hypotheses $\{H_N\}$ which specifies that for each $k = 1,\ldots,c$,

$$F_k(\underset{\sim}{x}) = F(x_{1N}^{(k)},\ldots,x_{pN}^{(k)}) \quad \text{with } F \in \mathcal{F}_0 \tag{6.1}$$

where

$$x_{iN}^{(k)} = (x_i + \theta_i^{(k)}N^{-1/2})(1 + \delta_i^{(k)}N^{-1/2}) \;,\; i = 1,\ldots,p. \tag{6.2}$$

We shall use the notation

$$\underset{\sim}{\theta}^{(k)} = (\theta_1^{(k)},\ldots,\theta_p^{(k)}) \quad \text{and} \quad \underset{\sim}{\delta}^{(k)} = (\delta_1^{(k)},\ldots,\delta_p^{(k)}) \tag{6.3}$$

and assume that at least one of the following two equalities

$$\underset{\sim}{\theta}^{(1)} = \cdots = \underset{\sim}{\theta}^{(c)}, \qquad \underset{\sim}{\delta}^{(1)} = \cdots = \underset{\sim}{\delta}^{(c)}$$

is not true. It may be noted that if $\underset{\sim}{\delta}^{(1)} = \cdots = \underset{\sim}{\delta}^{(c)}$, we have the usual location problem, and if $\underset{\sim}{\theta}^{(1)} = \cdots = \underset{\sim}{\theta}^{(c)}$, we have the scale problem; while if both the equalities hold, then $F_1 \equiv \cdots \equiv F_c$. Furthermore we shall denote as before the marginal cdf's of the i-th variate, and that of the i-th and j-th variate, corresponding to the distribution function $F(x_1,\ldots,x_p)$ by $F_{[i]}$ and $F_{[i,j]}$ respectively.

Next, let us denote

$$B_{(i)}(F) = \int_{-\infty}^{+\infty} \frac{d}{dx} J_{(i)}(F_{[i]}(x))\, dF_{[i]}(x) \;,\; i = 1,\ldots,p \tag{6.4}$$

$$C_{(i)}(F) = \int_{-\infty}^{+\infty} x \frac{d}{dx} J_{(i)}(F_{[i]}(x))\, dF_{[i]}(x) \ , \quad i = 1,\ldots,p \qquad (6.5)$$

$$\eta_i^{(k)} = B_{(i)}(F)\, \theta_i^{(k)} + c_i(F)\delta_i^{(k)} \ , \qquad k = 1,\ldots,c \qquad (6.6)$$

$$\underset{\sim}{\eta}^{(k)} = (\eta_1^{(k)},\ldots,\eta_p^{(k)}) \ , \qquad k = 1,\ldots,c \qquad (6.7)$$

Then, we have

THEOREM 6.1. _If_

(i) _for all_ $k = 1,\ldots,c$; $\lim_{N\to\infty} \lambda_N^{(k)} = \lambda^{(k)}$ _exists and is positive fraction less than unity,_

(ii) _conditions of Theorem_ 5.1 _are satisfied,_

then under the hypothesis $\{H_N\}$ _defined by_ (6.1) _and_ (6.2), _the random vector_ $[N^{1/2}(T_{Ni}^{(k)} - \mu_{Ni}^{(k)}),\ i = 1,\ldots,p;\ k = 1,\ldots,c]$ _has a limiting normal distribution with mean vector zero and covariance matrix_ $\underset{\sim}{\Sigma}^* = (((\sigma_{[i,j]}^{(k,q)}))$ _where_

$$\sigma_{[i,j]}^{(k,q)} = \left(\frac{\delta_{kq}}{\lambda^{(k)}} - 1\right)\nu_{ij}(F) \ , \quad i,j = 1,\ldots,p;\ k,q = 1,\ldots,c \qquad (6.8)$$

where δ_{kq} _is the usual kronecker delta and_ $\nu_{ij}(F)$ _is given by_ (5.25).

The proof of this theorem is an immediate consequence of Theorem 5.1 and some routine algebra, and is therefore omitted.

<u>THEOREM</u> 6.2. <u>If</u> $F \in \mathcal{F}_0$ <u>and if in addition to the conditions of Theorem</u> 6.1, <u>the conditions of Lemma</u> 7.2 <u>of Puri</u> [16] <u>are satisfied by</u> $F_{[i]}(x)$ <u>for all</u> $i = 1,\ldots,p$, <u>then under the sequence of alternatives</u> H_N <u>in</u> (6.1) <u>and</u> (6.2), $\mathcal{L}_N$ (<u>defined in</u> (3.7)) <u>has asymptotically a non-central chi-square distribution with</u> $p(c-1)$ <u>degrees of freedom and the non-centrality parameter</u>

$$\Delta_{\mathcal{L}} = \sum_{k=1}^{c} \lambda^{(k)}[\underset{\sim}{\eta}^{(k)})(\underset{\sim}{\nu}^{-1}(F))(\underset{\sim}{\eta}^{(k)})'] .$$

<u>PROOF</u>. It follows from Theorem 6.1 that under $\{H_N\}$ the rank of the asymptotic normal distribution in Theorem 5.1 is $p(c-1)$. Further, it follows from Theorem 4.2, (4.11), (4.14) and Theorem 6.1 that under the stated conditions

$$\underset{\sim}{\nu}(H) \longrightarrow \underset{\sim}{\nu}(F) \quad \text{as } N \to \infty , \tag{6.9}$$

$$\underset{\sim}{\nu}(\underset{\sim}{R}_N^*) \xrightarrow{P} \underset{\sim}{\nu}(H) \sim \underset{\sim}{\nu}(F) . \tag{6.10}$$

where $\sim$ indicates that the difference of the two sides tends to zero in probability. Finally under $\{H_N\}$ it is easily seen (by using (4.2) and (5.2)) that

$$[\mu_{Ni}^{(k)} - \overline{E}_N^{(i)}] \to \eta_i^{(k)} \quad \text{for all } i = 1,\ldots,p;\ k = 1,\ldots,c. \tag{6.11}$$

The rest of the proof of the theorem follows from Theorem 6.1 and by an application of the well-known results on the limit distributions of quadratic forms of asymptotic normal distributions, and hence is omitted.

Now since $\hat{\underset{\sim}{A}}^{-1}$ is a consistent estimator of $\underset{\sim}{A}^{-1}$, we have the following

THEOREM 6.3. <u>If the assumptions of Theorem</u> 6.2 <u>are satisfied, then for</u> $N \to \infty$, <u>the limiting distribution of the statistic</u> $\hat{\mathcal{L}}_N$ <u>defined by</u> (5.27) <u>is non-central chi-square with</u> $p(c-1)$ <u>d.f., and noncentrality parameter</u> $\Delta_{\mathcal{L}}$ <u>defined in Theorem</u> 6.2.

It now follows from Theorems 4.3, 6.1, 6.2 and 6.3 that under the sequence $\{H_N\}$ of alternative hypotheses

$$\mathcal{L}_N \overset{P}{\sim} \hat{\mathcal{L}}_N \tag{6.12}$$

and hence, using a well-known result of Hoeffding ([[14], p.172]), it follows that for testing H_0 against the set of alternatives $\{H_N\}$ the PROT and CSROT are asymptotically power-equivalent.

7. CHOICE OF $\{J_{(i)},\ i = 1,\dots,p\}$ AND RELATED ASYMPTOTIC POWER EFFICIENCY OF THE RANK-ORDER TESTS.

In this section, we shall consider specifically the problems of location and scale considered in Section 2, and for each problem, we shall consider some specific tests and study the resulting asymptotic efficiency.

(a) <u>Location problem</u>.

We shall consider here the following tests.

(i) <u>Median test</u>. This is characterized by

$$J_{N(i)}(\alpha/(N+1)) = \begin{cases} 1, & \text{if } \alpha \leq [(N+1)/2], \\ 0, & \text{otherwise}, \end{cases} \tag{7.1}$$

where $[s]$ denotes the largest integer $\leq s$. This implies that

$$J_{(i)}(u) = \begin{cases} 1 & \text{if } u \leq 1/2 \\ 0 & \text{if } u > 1/2 \end{cases}$$

The function $J_{(i)}(u)$ has a single point of discontinuity at $u = 1/2$, and if the cdf $F_{[i]}(x)$ has a uniquely defined median, the Lebesgue measure of this point will be equal to zero. If we work with (7.1) for all $i = 1,\ldots,p$, the test will be termed the <u>multivariate multisample median test</u>.

(ii) <u>Rank-sum test</u>. Let

$$J_{N(i)}(\alpha/(N+1)) = \alpha/(N+1);\ \alpha = 1,\ldots,N \tag{7.2}$$

which implies that $J_{(i)}(u) = u$: $0 < u < 1$. If this weight-function is used for all $i = 1,\ldots,p$ the corresponding test will be termed the <u>multivariate rank-sum test</u>.

Both the median and rank-sum tests have been studied in detail by Chatterjee and Sen ([6],[8]), and hence, here these will not be considered again.

(iii) <u>Ψ-Score test</u>. Let $\Psi(x)$ be any specified continuous cdf, and we define

$$J_{N(i)}(\alpha/(N+1)) = \Psi^{-1}(\alpha/(N+1));\ \alpha=1,\ldots,N \tag{7.3}$$

This implies that $J_{(i)}(u) = \Psi^{-1}(u)$: $0 < u < 1$. If we use (7.3) for all $i = 1,\ldots,p$, the corresponding test will be termed the <u>Ψ-score test</u>. In particular, if Ψ is a standardized normal cdf, the test is termed the <u>normal scores test</u>. It may be noted that often we define the Ψ-scores in a slightly different but asymptotically equivalent way. Let $a_{N,\alpha}$ be the expected value of the α-th order statistic in a sample of size N drawn from a universe having the cdf $\Psi(x)$. Then we take

$$J_{N(i)}(\alpha/(N+1)) = a_{N,\alpha} \; ; \quad \alpha = 1,\ldots,N. \tag{7.4}$$

For a wide class of cdf's [cf. Hoeffding [15]), the use of (7.3) and (7.4) leads to asymptotically equivalent results. The conditions (1), (2) and (3) of Section 4 are known to be satisfied by (7.3) and (7.4) for a wide class of $\Psi(x)$ (cf. [9], [11], [16]), and so we require only to show that condition (4) or (4.12) is also satisfied. To do this we rewrite (4.12) as

$$\frac{1}{n_k} \sum_{j=1}^{n_k} [a_{N,\alpha_j} - \xi_{N,\alpha_j}]^2 , \tag{7.5}$$

where $1 \leq \alpha_1 < \ldots < \alpha_{n_k} \leq N$ are any n_k distinct integers, and $\Psi(\xi_{N,\alpha}) = \alpha/(N+1)$, for $\alpha = 1,\ldots,N$. Now (7.5) cannot be larger than

$$(N/n_k)\left\{ (1/N) \sum_{\alpha=1}^{N} [a_{N,\alpha} - \xi_{N,\alpha}]^2 \right\}. \tag{7.6}$$

We shall then prove the following.

LEMMA 7.1. <u>Let</u> $\Psi(x)$ <u>be any continuous cdf, satisfying</u>

(i) $\int x \, d\Psi(x) = 0$, $\int x^2 \, d\Psi(x) < \infty$,

(ii) $\Psi(x)$ <u>is symmetric about</u> $x = 0$; $\Psi(x)$ <u>is convex for all</u> $x \leq 0$ <u>and concave for all</u> $x > 0$, (<u>or vice versa</u>). <u>Then</u>

$$\lim_{N \to \infty} (1/N) \sum_{i=1}^{N} [a_{N,i} - \xi_{N,i}]^2 = 0 . \tag{7.7}$$

PROOF. The left hand side of (7.7) can be written as

$$\frac{1}{N}\sum_{i=1}^{N}\alpha_{N,i}^{2}+\frac{1}{N}\sum_{i=1}^{N}\xi_{N,i}^{2}-\frac{2}{N}\sum_{i=1}^{N}\alpha_{N,i}\,\xi_{N,i}\,. \tag{7.8}$$

It has been shown by Ali and Chan [1] that if $\Psi(x)$ is concave (convex) for $x \geq 0$, then for $i \geq [(N+1)/2]$, $\alpha_{N,i} \geq (\leq)\ \xi_{N,i}$ and opposite inequalities hold for $i < [(N+1)/2]$. Thus, if $\Psi(x)$ is convex for $x \leq 0$ and concave for $x > 0$, we readily get from the above result that

$$\sum_{i=1}^{N}\xi_{N,i}^{2} \leq \sum_{i=1}^{N}\xi_{N,i}\,\alpha_{N,i} \leq \sum_{i=1}^{N}\alpha_{N,i}^{2} \tag{7.9}$$

and the opposite inequalities hold if $\Psi(x)$ is concave for $x \leq 0$ and convex for $x \geq 0$. Thus, from (7.7), we get that

$$\frac{1}{N}\sum_{i=1}^{N}(\alpha_{N,i}-\xi_{N,i})^{2} \leq \left|\frac{1}{N}\sum_{i=1}^{N}\alpha_{N,i}^{2}-\frac{1}{N}\sum_{i=1}^{N}\xi_{N,i}^{2}\right|. \tag{7.10}$$

Again it follows from Hoeffding's results [15] that

$$\frac{1}{N}\sum_{i=1}^{N}\alpha_{N,i}^{2} \longrightarrow \int x^{2}\,d\Psi(x) < \infty\ , \tag{7.11}$$

and by the fundamental theorem of integral calculus

$$\frac{1}{N}\sum_{i=1}^{N}\xi_{N,i}^{2} \longrightarrow \int x^{2}\,d\Psi(x) < \infty . \tag{7.12}$$

From (7.10), (7.11) and (7.12), we readily conclude that (7.7) converges to zero as $N \to \infty$.

Hence the lemma.

In particular, if $\Psi(x)$ is a standardized normal cdf, it is convex for $x \leq 0$ and concave for $x > 0$, and hence the result of Lemma 7.1 holds. The same lemma holds for logistic, double exponential, rectangular and many other cdf's.

Finally, we require to show that $\underset{\sim}{\nu}(F)$ in (4.15) is positive definite under certain conditions. $\underset{\sim}{\nu}(F)$ will cease to be positive definite when there exists one or more linear relations among the variables $J_{(i)}(F_{[i]}(x))$, $i = 1,\ldots,p$. This in turn requires that the scatter of the points of the distribution $F(x_1,\ldots,x_p)$ is confined to a (p-1) or lower dimensional) hyper-curve whose equation is

$$\sum_{i=1}^{p} c_i \, J_{(i)}(F_{[i]}(X_i)) = \text{constant} \tag{7.13}$$

(or more than one such equations). Thus, if the cdf $F(x_1,\ldots,x_p)$ is non-singular in the sense that there is no (p-1) or lower dimensional subspace of the p-dimensional Euclidean space E_p whose probability mass is exactly equal to unity, then (7.13) cannot hold, in probability, and hence (4.15) would be positive definite. In actual practice this is practically no restriction on F.

In the parametric case, under the assumption of non-singular multinormal distribution of each of $F_1,\ldots,F_c$, the likelihood ratio test for location is given in Anderson ([2], Sec. 8.8) and Rao ([18], Sec. 8d.4), and it is easily seen that under $\{H_N\}$ (when $\underset{\sim}{\delta}^{(k)} = 0$ for all $k = 1,\ldots,c$), this statistic has asymptotically a non-central χ^2 distribution with p(c-1) d.f. and the non-centrality parameter

$$\Delta_U = \sum_{k=1}^{c} \lambda^{(k)} \left[(\underset{\sim}{\theta}^{(k)}) \underset{\sim}{\Sigma}^{-1} (\underset{\sim}{\theta}^{(k)})' \right] \tag{7.14}$$

where $\underset{\sim}{\Sigma} = ((\sigma_{ij}))$ is the common covariance matrix of all the cdf's.

Thus the relative asymptotic power-efficiency of $\mathcal{L}_N$-test with respect to the likelihood ratio U-test is given by

$$e_{\{\mathcal{L}_N\};\{U\}} = \Delta \;/\Delta_U \; . \tag{7.15}$$

It follows from (7.3), (7.4), (4.15), Theorem 6.2 and (7.14) that if F is a p-variate non-singular normal cdf, then

$$\Delta_{\mathcal{L}} = \Delta_U$$

for all $\underset{\sim}{\theta}^{(k)}$, $k = 1,\ldots,c$. So asymptotically, for normal cdf's, the normal scores test is as efficient as the U-test. In the univariate case, Hodges and Lehmann [13] have shown that asymptotically normal scores test is always at least as good as the student's t-test for all cdf $F \in \mathcal{F}$. The generalization of such a statement to the multivariate case seems to quite difficult to prove and dubious too. It turns out that for the multivariate case, (7.15) depends not only on $B_i(F)$, $i = 1,\ldots,p$ (defined in (6.4)), but also on the matrix $\underset{\sim}{\nu}(F)$ in (4.15) as well as on $\underset{\sim}{\theta}^{(k)}$, $k = 1,\ldots,c$. Using the results of Hodges and Lehmann [13], it is easily seen that for the normal scores test

$$B_i(F)/\sigma_{ii} \geq 1 \quad \text{for all } i = 1,\ldots,p, \text{ and all } F \in \mathcal{F}_0. \tag{7.16}$$

So, if the p-variates in F are all uncorrelated, it is easily seen that (7.15) ≥ 1 for all $F \in \mathcal{F}_0$, independently of $\underset{\sim}{\theta}^{(k)}$, $k = 1,\ldots,c$. On the other hand, if these p variates are correlated, we may proceed precisely on the same line as in Bickel ([5, (4.3)]), and express (7.15) as the product of two factors, namely the univariate and the multivariate factors. The univariate factor is independent of $\bar{\underset{\sim}{\theta}}^{(k)}$, $k = 1,\ldots,c$, and is easily shown to be at least as large as 1, while the multivariate factor depends implicitly on the characteristic roots of $\underset{\sim}{\nu}^{-1}\underset{\sim}{\Sigma}$ and explicitly on $\underset{\sim}{\theta}^{(1)},\ldots,\underset{\sim}{\theta}^{(c)}$. In the bivariate two sample case, Chatterjee and Sen [6] have obtained some asymptotic power-inequalities of the median test and the rank-sum test, and Bickel [5] has obtained some bounds for the maximum and minimum power of the same tests in the single sample case. Essentially, the same technique may be applied to study the bounds for (7.15) in the particular case of $p = c = 2$. However, in the general case of $p,c \geq 2$, it seems considerably difficult to establish the bounds for (7.15) or to evaluate (7.15) exactly. In fact, for some non-normal multivariate cdf's (7.15) can be shown to be greater than one, but we believe that (7.15) cannot be greater than or equal to 1 for all $F \in \mathcal{F}_0$. Some counter-examples may be provided for that, especially in the limiting non-degenerate case, where one has essentially to follow Bickel's [5] technique with multisample extensions. This is however not considered here.

(b). <u>Scale problem</u>. Here, we shall consider the following rank-order tests.

(i) Let $J_{N(i)}(\alpha/(N+1)) = \left|\frac{\alpha}{N+1} - \frac{1}{2}\right|$ for $\alpha = 1,\ldots,N$; $i = 1,\ldots,p$ (7.17)

so that

$$J_{(i)}(u) = \left|u - \tfrac{1}{2}\right| \text{ for } 0 < u < 1 \text{ and } i = 1,\ldots,p.$$

(ii) Let $J_{N(i)}(\alpha/(N+1)) = [\frac{\alpha}{N+1} - \frac{1}{2}]^2$ for $\alpha = 1,\ldots,N; i=1,\ldots,p$ (7.18)

so that

$$J_{(i)} = (u - \tfrac{1}{2})^2 \quad \text{for } 0 < u < 1 \text{ and } i = 1,\ldots,p.$$

(iii) Let $\mathcal{a}_{N,\alpha}$ be defined as in Section 7 (a) and let

$$J_{N(i)}(\alpha/(N+1)) = \mathcal{a}^2_{N,\alpha}, \quad \alpha = 1,\ldots,N; i=1,\ldots,p \qquad (7.19)$$

so that

$$J_{(i)}(u) = [\Psi^{-1}(u)]^2, \qquad 0 < u < 1; i=1,\ldots,p. \quad (7.20)$$

In the univariate case, the rank-order test corresponding to (7.17) has been considered by Ansari and Bradley, and Siegel and Tukey (cf. [17]); the rank-order test corresponding to (7.18) by Mood and the one corresponding to (7.19) by Klotz, among others (cf. [17]). Here also, in the multivariate case, the last test (for the case when $\Psi(x)$ is the standard normal cdf) will be termed the multivariate normal score test for scale.

Proceeding then precisely on the same line as in the location case, it is easily seen that under essentially similar conditions, a strictly distribution-free permutation test and a large sample non-parametric test can be constructed for each of the three types of weight functions defined in (7.17), (7.18) and (7.19).

In the parametric case, under the assumption of non-singular normal distribution of each of $F_1,\ldots,F_c$, some optimum likelihood ratio test is available for testing the identity of dispersion matrices of $F_1,\ldots,F_c$. If, however, we are interested only in testing the identity of the variances for each of the p variates, we can have a similar test whose test-statistic (say V) has asymptotically under H_N in (7.1) and (7.2) (where $\underset{\sim}{\theta}^{(k)} = \underset{\sim}{0}$ for all $k = 1,\ldots,c$), a non-central χ^2 distribution with $p(c-1)$ d.f. and the non-centrality parameter

$$\Delta_V = \sum_{k=1}^{c} \lambda^{(k)}[(\underset{\sim}{\delta}^{(k)})\underset{\sim}{\Gamma}^{-1}(\underset{\sim}{\delta}^{(k)})'] \tag{7.21}$$

where $\underset{\sim}{\Gamma} = ((\gamma_{ij}))_{i,j=1,\ldots,p}$ with

$$\begin{aligned} \gamma_{ii} &= 2\sigma_{ii}^2, & i &= 1,\ldots,p \\ \gamma_{ij} &= 2(\sigma_{ij})^2, & i &\neq j = 1,\ldots,p, \end{aligned} \tag{7.22}$$

$\underset{\sim}{\Sigma} = ((\sigma_{ij}))$ being the covariance matrix of F.

Hence the asymptotic efficiency of the $\mathcal{L}_N$ test with respect to the V-test is given by

$$e_{\{\mathcal{L}_N\},\{V_N\}} = \Delta \;/\; \Delta_V . \tag{7.23}$$

In particular, when the underlying distribution function F is a non-singular p-variate normal, the multivariate normal scores test for scale is asymptotically as efficient as the V-test. The efficiency of the first two scale tests can be expressed as the product of two factors, the univariate factors are known (for the normal cdf, they are equal to 0.61 and 0.76 respectively), while the multivariate factors will depend on $\underset{\sim}{\Sigma}$ as well as on $\underset{\sim}{\delta}^{(k)}$, for $k = 1,\ldots,c$. As in the location test, nothing can be said about the lower bounds for the asymptotic efficiency of the normal scores test or the other two tests for the non-normal cdf's.

So far, we have considered rank-order tests for which $J_{N(i)}$ does not depend on i. Evidently, it is always possible to use a rank order test with different $J_{N(i)}$'s for different $i = 1,\ldots,p$. In the event when there are sufficient reasons to believe that for different i's the $F_{[i]}$'s are appreciably different from each other, some specific $J_{N(i)}$ will be suitable, such a test may be recommended.

ACKNOWLEDGMENT

Some of the ideas of this paper evolved in the course of collaboration of one of the authors with Dr. S. K. Chatterjee of Calcutta University, while they were studying the rank-permutation tests considered in ([6],[7],[8]). The authors would like to express their sincere thanks to Dr. J. Sethuraman for pointing out some errors in the original draft and making useful suggestions.

REFERENCES

[1] ALI, M., and CHAN, L. (1965). Some Bounds for Expected Values of Order Statistics. Ann. Math. Stat. 36, 1055-1057.

[2] ANDERSON, T. W. (1958). An Introduction to Multivariate Statistical Analysis. John Wiley. New York.

[3] ANDERSON, T. W. (1965). Some multivariate non-parametric procedures based on statistically equivalent blocks. Proc. International Symp. Multivariate Analysis. Dayton, Ohio. (in press).

[4] BHAPKAR, V. P. (1965). Some non-parametric test for the multivariate several sample location problem. Proc. International Symp. Multivariate Analysis. Dayton, Ohio. (in press).

[5] BICKEL, P. J. (1965). Some asymptotically non-parametric completitors of Hotelling's T^2. Ann. Math. Stat. 36, 160-173.

[6] CHATTERJEE, S. K. and SEN, P. K. (1964). Non-parametric tests for the bivariate two sample location problem. Calcutta Stat. Assoc. Bull. 13, 18-58.

[7] CHATTERJEE S. K. and SEN, P. K. (1965). Some non-parametric tests for the two sample bivariate association problem. Calcutta Stat. Assoc. Bull. 14, 14-35.

[8] CHATTERJEE, S. K. and SEN, P. K. (1965). Non-parametric tests for the multivariate multisample location problem. Ann. Math. Stat. (submitted).

[9] CHERNOFF, H. and SAVAGE, I. R. (1958). Asymptotic normality of certain nonparametric test statistics. Ann. Math. Stat. 29, 972-994.

[10] FRASER, D. A. S. (1957). Non-parametric Methods in Statistics. John Wiley and Sons, Inc., New York.

[11] GOVINDARAJULU, Z., LeCAM, L., and RAGHAVACHARI, M. (1965). Generalization of Chernoff-Savage theorems on asymptotic normality of non-parametric test statistics. Proc. Fifth Berkeley Symp. Math. Stat. Prob. (in press).

[12] HAJEK, J. (1961). Some extensions of the Wald-Wolfowitz-Noether theorem. Ann. Math. Stat. 32, 506-523.

[13] HODGES, J. L., Jr. and LEHMANN, E. L. (1960). Comparison of the normal score and Wilcoxon tests. Proc. Fourth Berkeley Symp. Math. Stat. Prob. 1, 307-317.

[14] HOEFFDING, W. (1952). The large sample power of tests based on permutations of observations. Ann. Math. Stat. 23, 169-192.

[15] HOEFFDING, W. (1953). On the Distribution of the Expected Values of the Order Statistics. Ann. Math. Stat. 24, 93-100.

[16] PURI, M. L. (1964). Asymptotic efficiency of a class of c-sample tests. Ann. Math. Stat. 35, 102-121.

[17] PURI, M. L. (1965). On some tests of homogeneity of variances. Ann. Inst. Stat. Math., 17, 323-330.

[18] RAO, C. R. (1965). Linear Statistical Inference and its Applications. John Wiley and Sons, Inc, New York.

[19] SEN, P. K. (1965). On some permutation tests based on U-statistics. Calcutta Stat. Assoc. Bull. 14, 106-126.

[20] SEN, P. K. (1965). On a class of two sample bivariate non-parametric tests. Proc. Fifth Berkeley Symp. Math. Stat. Prob. (in press).

[21] SEN, P. K. (1965). On a class of multisample multivariate non-parametric tests. Ann. Inst. Stat. Math. (submitted).

[22] WALD, A. and WOLFOWITZ, J. (1944). Statistical tests based on permutation of observations. Ann. Math. Stat. 15, 358-372.