A CLASS OF RANK-ORDER TESTS FOR A GENERAL LINEAR HYPOTHESIS

by

Madan Lal Puri and Pranab Kumar Sen

Courant Institute of Mathematical Sciences,
New York University and University of North Carolina

Institute of Statistics Mimeo Series No. 551

September 1967

Research sponsored by the Office of Naval Research, Contract Nonr-285(38), and by the National Institute of Health, Public Health Service, Grant GM-12868-03.

DEPARTMENT OF BIOSTATISTICS

UNIVERSITY OF NORTH CAROLINA

Chapel Hill, N. C.

A CLASS OF RANK-ORDER TESTS FOR A GENERAL LINEAR HYPOTHESIS*

By Madan Lal Puri and Pranab Kumar Sen

Courant Institute of Mathematical Sciences, New York University and University of North Carolina at Chapel Hill

Summary. For a general multivariate linear hypothesis testing problem a class of permutationally distribution-free rank-order tests is proposed and studied. Along with a multivariate cum multistatistics generalization of Hájek's (1966) results on the asymptotic normality of rank-order statistics (for regression alternatives), the asymptotic distribution theory of the proposed tests and their asymptotic power properties are studied. Asymptotic optimality of the proposed tests for specific alternatives is established and a characterization of the multivariate multisample location problem [cf. Puri and Sen (1966)] in terms of the proposed linear hypothesis is also considered.

* Research sponsored by the Office of Naval Research, Contract Nonr-285(38), and by the National Institute of Health, Public Health Service, Grant GM-12868-03. Reproduction in whole or in part is permitted for any purpose of the United States Government.

1. Introduction. Let us consider a sequence of stochastic matrices $\underset{\sim}{E}_\nu = (\underset{\sim}{X}_{\nu 1}, \ldots, \underset{\sim}{X}_{\nu N_\nu})$, $1 \leq \nu < \infty$, where $\underset{\sim}{X}_{\nu i} = (X_{\nu i}^{(1)}, \ldots, X_{\nu i}^{(p)})'$, $i = 1, \ldots, N_\nu$ are independent stochastic vectors having continuous cumulative distribution functions (cdf) $F_{\nu i}(\underset{\sim}{x})$, $\underset{\sim}{x} \in R^p$, $i = 1, \ldots, N_\nu$, respectively. It is assumed that

$$(1.1) \quad F_{\nu i}(\underset{\sim}{x}) = F(\underset{\sim}{x} - \underset{\sim}{\alpha} - \underset{\sim}{\beta}\underset{\sim}{c}_{\nu i}) \ , \quad 1 \leq i \leq N \ , \quad 1 \leq \nu < \infty \ ,$$

where $\underset{\sim}{\alpha}' = (\alpha_1, \ldots, \alpha_p)$ and $\underset{\sim}{\beta} = ((\beta_{jk}))_{j=1,\ldots,p;\ k=1,\ldots,q}$ are unknown parameters and $\underset{\sim}{c}_{\nu i} = (c_{\nu i}^{(1)}, \ldots, c_{\nu i}^{(q)})'$, $i = 1, \ldots, N_\nu$ are known non-stochastic vectors. Having observed $\underset{\sim}{E}_\nu$ and assuming some conditions on $\underset{\sim}{c}_{\nu i}$'s (to be stated in Section 2), we want to test the null hypothesis

$$(1.2) \qquad H_o: \underset{\sim}{\beta} = \underset{\sim}{0}^{p \times q} \ ,$$

against the set of alternatives that $\underset{\sim}{\beta}$ is non-null.

A variety of tests for H_o in (1.2), based on the assumed normality of $F(\underset{\sim}{x})$, is available in the literature [cf. Anderson (1958, Chapter 8) and Rao (1965, Chapter 8)]; the likelihood ratio test is one of the mostly adopted ones. In this paper, the assumption of multinormality of $F(\underset{\sim}{x})$ is relaxed and a class of Chernoff-Savage-Hájek type of rank-order tests is proposed and studied. These tests are valid for all continuous

$F(\underset{\sim}{x})$. In Section 2, along with the preliminary notions, these rank-order statistics are defined. Section 3 is concerned with certain permutational invariance structure of the joint distribution of $\underset{\sim}{E}_\nu$ (when H_0 in (1.2) holds), and this is then utilized for the construction of permutationally distribution-free tests for H_0 in (1.2). In Section 4, by a generalization of Hájek's (1966) results, the asymptotic joint normality of the proposed rank-order statistics (for nearby alternatives) is established. Section 5 deals with the asymptotic power properties of the proposed tests for nearby alternatives, and Section 6 is concerned with the asymptotic optimality of the proposed tests, granted certain conditions on $F(\underset{\sim}{x})$. In the last Section, relationship of the multivariate multisample location problem with the linear hypothesis in (1.1) and (1.2) is studied.

2. *Preliminary notions and fundamental assumptions*. Let $R_{\nu i}^{(j)}$ be the rank of $X_{\nu i}^{(j)}$ among the N_ν observations $X_{\nu 1}^{(j)}, \ldots, X_{\nu N_\nu}^{(j)}$, for $i = 1, \ldots, N_\nu$; by virtue of the assumed continuity of $F(\underset{\sim}{x})$ ties may be ignored, in probability, for all $j = 1, \ldots, p$. Let $\phi_{\nu,j}(t)$ be a function defined on (0,1) and taking constant values over intervals $[i/N_\nu, (i+1)/N_\nu)$, $0 \leq i \leq N_\nu$, i.e., there is some function $\phi_j(t)$ defined for $0 < t < 1$, for which

$$(2.1) \qquad \phi_{\nu,j}(t) = \phi_j\left(\frac{i}{N_\nu+1}\right) \text{ for } \frac{i-1}{N_\nu} < t \leq \frac{i}{N_\nu}, \; i = 1, \ldots, N_\nu; \; j = 1, \ldots, p .$$

Define

$$\mathcal{Q}_{\nu}^{(j)}(\alpha) = E\{\phi_j(U_{\nu(\alpha)})\} , \quad \alpha = 1,\dots, N_\nu , \quad j = 1,\dots, p , \tag{2.2}$$

where $U_{\nu(1)} \leq \cdots \leq U_{\nu(N_\nu)}$ denote an ordered sample of N_ν observations from the rectangular distribution over (0,1). (We may also define $\mathcal{Q}_{\nu}^{(j)}(\alpha) = \phi_j(\alpha/N_\nu+1)$ for $\alpha = 1,\dots, N_\nu$ and $j = 1,\dots, p$. However, as it is known that the two definitions lead to asymptotically equivalent results and the theory follows on almost parallel lines, for simplicity of notations and for intended brevity of discussion, we will consider only the scores defined by (2.2).) Concerning $\phi_{\nu,j}$'s, we assume that

(2.3)(I) $\lim_{\nu\to\infty} \phi_{\nu,j}(t) = \phi_j(t)$ exists for all $0 < t < 1$ and constant, for $j = 1,\dots, p$, (where of course $N_\nu \longrightarrow \infty$ as $\nu \longrightarrow \infty$), and

(2.4)(II) $\lim_{\nu\to\infty} \int_0^1 [\phi_{\nu,j}(t) - \phi_j(t)]^2 dt = 0$ for all $j = 1,\dots, p$.

Consider now the random variables

$$S_{\nu,jk} = \sum_{i=1}^{N_\nu} c_{\nu i}^{(k)} \mathcal{Q}_{\nu}^{(j)}(R_{\nu i}^{(j)}) \quad \text{for } j = 1,\dots, p; \quad k = 1,\dots, q . \tag{2.5}$$

Our proposed test for H_o in (1.2) is based on the stochastic matrix

$$(2.6) \qquad \underset{\sim}{S}_\nu = ((S_{\nu,jk})) .$$

Concerning $\underset{\sim}{c}_{\nu i}$'s, the following assumptions are made:

$$(2.7)(III) \qquad \sum_{i=1}^{N_\nu} c_{\nu i}^{(k)} = 0 \quad \text{for } k = 1,\ldots, q \quad \text{and} \quad 1 \leq \nu < \infty ,$$

(this can always be done by suitably adjusting $\underset{\sim}{\alpha}$ in (1.1)),

$$(2.8)(IV) \quad \lim_{\nu \to \infty} \{ \max_{1 \leq i \leq N_\nu} [c_{\nu i}^{(k)}]^2 \Big/ \sum_{i=1}^{N_\nu} [c_{\nu i}^{(k)}]^2 \} = 0 \quad \text{for all } k = 1,\ldots, q;$$

$$(2.9)(V) \qquad \operatorname{Sup}_\nu \{\sum_{i=1}^{N_\nu} [c_{\nu i}^{(k)}]^2\} < \infty \quad \text{for all } k = 1,\ldots, q .$$

Let then $\underset{\sim}{C}_\nu = ((C_{\nu,kk'}))_{k,k'=1,\ldots,q}$ where

$$(2.10)(VI) \qquad \text{Rank of } \underset{\sim}{C}_\nu = q \geq 1 .$$

(In fact, by reparameterization in (1.1), $\underset{\sim}{C}_\nu$ can always be made of full rank. Hence, (2.11) is no restriction to our model.) Let us also define

$$(2.12) \quad \mathcal{M}_{\nu,jj'} = \frac{1}{N_\nu - 1} \{\sum_{i=1}^{N_\nu} a_\nu^{(j)}(R_{\nu i}^{(j)}) a_\nu^{(j')}(R_{\nu i}^{(j')})$$

$$- \frac{1}{N_\nu} (\sum_{i=1}^{N_\nu} a_\nu^{(j)}(i))(\sum_{i=1}^{N_\nu} a_\nu^{(j')}(i))\} ,$$

$$(2.13) \qquad \underset{\sim}{M}_\nu = ((\mathcal{M}_{\nu,jj'}))_{j,j'=1,\ldots,p} .$$

(2.14)(VII) $$\text{Rank of } \underset{\sim}{M}_\nu = p \geq 1 \; .$$

(Later on, we shall consider a theorem which establishes that under certain conditions on $F(\underset{\sim}{x})$, (2.14) holds.) Finally, let

(2.15) $$\underset{\sim}{D}_\nu = \underset{\sim}{M}_\nu \otimes \underset{\sim}{C}_\nu = ((d_{\nu,jk;j'k'})) \; ,$$

(where $\otimes$ refers to the Kronecker-product of two matrices). By (2.9) and (2.14), we obtain that

(2.16) $$\text{Rank of } \underset{\sim}{D}_\nu = pq \; .$$

We denote by $\underset{\sim}{D}_\nu^{-1} = ((d_{\nu,jk;j'k'}))^{-1} = ((d_\nu^{jk,j'k'}))$ the reciprocal matrix of $\underset{\sim}{D}_\nu$. We may define formally our proposed test statistic as

(2.17) $$\mathcal{L}_\nu = \sum_{j=1}^{p} \sum_{j'=1}^{p} \sum_{k=1}^{q} \sum_{k'=1}^{q} d_\nu^{jk,j'k'} S_{\nu,jk} S_{\nu,j'k'} \; ,$$

and its rationality will be made clear in the next section. Before that, we would like to introduce the following notations. Let $H(\underset{\sim}{x})$ be a p-variate cdf whose univariate marginals are denoted by $H_{[j]}(x)$ $(j = 1, \ldots, p)$ and the bivariate (joint) marginals by $H_{[j,j']}(x,y)$ for $j \neq j' = 1, \ldots, p$. Further, let

(2.18) $$\bar{\phi}_j = \int_0^1 \phi_j(t)dt \; , \qquad j = 1, \ldots, p \; ,$$

(2.19) $$\eta_{jj} = \int_0^1 \phi_j^2(t)dt - \bar{\phi}_j^2 \; , \quad j = 1, \ldots, p \; ;$$

(2.20) $\eta_{jj'}(H)$

$$= \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} \phi_j(H_{[j]}(x))\phi_{j'}(H_{[j']}(y))dH_{[jj']}(x,y) - \bar{\phi}_j\bar{\phi}_{j'} ,$$

$j \neq j' = 1,\ldots, p$;

(2.21) $$\underset{\sim}{\Lambda}(H) = ((\eta_{jj'}(H)))_{j,j'=1,\ldots,p} .$$

Finally, let

(2.22) $$\underset{\sim}{\Gamma}_\nu(H) = \underset{\sim}{\Lambda}(H) \otimes \underset{\sim}{C}_\nu \quad \text{and} \quad \underset{\sim}{\Gamma}(H) = \lim_{\nu\to\infty} \underset{\sim}{\Gamma}_\nu(H) .$$

It follows from (2.9), (2.11) and (2.11) that under the assumption (2.23) (VIII) $\underset{\sim}{\Lambda}(H)$ is of rank p and is finite, $\underset{\sim}{\Gamma}_\nu(H)$ will also be of rank pq, and $\underset{\sim}{\Gamma}(H)$ is finite, provided $\lim_{\nu\to\infty} \underset{\sim}{C}_\nu$ exists. However, $\text{Sup}_\nu \|\underset{\sim}{\Gamma}_\nu(H)\|$ will be finite, by (2.9).

3. Permutation distribution of S_ν and the rationality of $\mathcal{L}_\nu$.

Under H_o in (1.2), $\underset{\sim}{E}_\nu$ is composed of N_ν independent and identically distributed random vectors. Hence, the joint distribution of $\underset{\sim}{E}_\nu$ remains invariant under the $N_\nu!$ permutation of the N_ν vectors $\underset{\sim}{X}_{\nu 1},\ldots, \underset{\sim}{X}_{\nu N_\nu}$ among themselves when (1.2) holds. We consider now the basic rank-permutation principle, which is essentially due to Chatterjee and Sen (1966) and discussed also by Puri and Sen (1966). We define a $p \times N_\nu$ rank matrix $\underset{\sim}{R}_\nu$

$$(3.1) \qquad \underset{\sim}{R}_\nu = \begin{pmatrix} R_{\nu 1}^{(1)} & \cdots & R_{\nu N_\nu}^{(1)} \\ \vdots & & \vdots \\ R_{\nu 1}^{(p)} & \cdots & R_{\nu N_\nu}^{(p)} \end{pmatrix} = (\underset{\sim}{R}_{\nu 1}, \ldots, \underset{\sim}{R}_{\nu N_\nu}) ,$$

where $\underset{\sim}{R}_{\nu i} = (R_{\nu i}^{(1)}, \ldots, R_{\nu i}^{(p)})'$ for $i = 1, \ldots, N_\nu$. Each row of $\underset{\sim}{R}_\nu$ is a permutation of the numbers $1, \ldots, N_\nu$, there being thus in all $(N_\nu!)^p$ possible realizations of $\underset{\sim}{R}_\nu$. Let us rearrange the columns of $\underset{\sim}{R}_\nu$ in such a way that the first row has elements $1, \ldots, N_\nu$ in the natural order, and denote the corresponding matrix by $\underset{\sim}{R}_\nu^*$. $\underset{\sim}{R}_\nu$ is said to be permutationally equivalent to a matrix $\underset{\sim}{R}_\nu^{(1)}$, if it is possible to obtain $\underset{\sim}{R}_\nu^{(1)}$ from $\underset{\sim}{R}_\nu$ only by permutations of the columns of the latter. Thus, corresponding to each $\underset{\sim}{R}_\nu^*$, there will be a set $\Sigma(\underset{\sim}{R}_\nu^*)$ of $N_\nu!$ possible realizations of $\underset{\sim}{R}_\nu$, such that any member of this set will be permutationally equivalent to $\underset{\sim}{R}_\nu^*$. Now, the probability distribution of $\underset{\sim}{R}_\nu$ over the $(N_\nu!)^p$ possible realizations will depend on $F(\underset{\sim}{x})$, even when H_0 in (1.2) holds (unless $F(\underset{\sim}{x})$ has coordinate-wise independent marginals). Thus, rank-statistics, like $\underset{\sim}{S}_\nu$ in (2.6), are, in general, not distribution-free under H_0 in (1.2). However, given a particular set $\Sigma(\underset{\sim}{R}_\nu^*)$ (of $N_\nu!$ realizations), the conditional distribution of $\underset{\sim}{R}_\nu$ over the $N_\nu!$ over the $N_\nu!$ permutations of the columns of $\underset{\sim}{R}_\nu^*$ would be uniform under H_0 in (1.2), i.e.,

$$(3.2) \qquad P\{\underset{\sim}{R}_\nu = \underset{\sim}{r}_\nu \mid \Sigma(\underset{\sim}{R}_\nu^*), H_0\} = \frac{1}{N_\nu!} \quad \text{for all } \underset{\sim}{r}_\nu \in \Sigma(\underset{\sim}{R}_\nu^*) ,$$

(2.20) $\eta_{jj'}(H)$

$$= \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} \phi_j(H_{[j]}(x))\phi_{j'}(H_{[j']}(y))dH_{[jj']}(x,y) - \bar{\phi}_j\bar{\phi}_{j'} ,$$

$$j \neq j' = 1,\ldots, p ;$$

(2.21) $$\underset{\sim}{A}(H) = ((\eta_{jj'}(H)))_{j,j'=1,\ldots,p} .$$

Finally, let

(2.22) $$\underset{\sim}{\Gamma}_\nu(H) = \underset{\sim}{A}(H) \otimes \underset{\sim}{C}_\nu \quad \text{and} \quad \underset{\sim}{\Gamma}(H) = \lim_{\nu\to\infty} \underset{\sim}{\Gamma}_\nu(H) .$$

It follows from (2.9), (2.11) and (2.11) that under the assumption (2.23) (VIII) $\underset{\sim}{A}(H)$ is of rank p and is finite, $\underset{\sim}{\Gamma}_\nu(H)$ will also be of rank pq, and $\underset{\sim}{\Gamma}(H)$ is finite, provided $\lim_{\nu\to\infty} \underset{\sim}{C}_\nu$ exists. However, $\text{Sup}_\nu \|\underset{\sim}{\Gamma}_\nu(H)\|$ will be finite, by (2.9).

3. Permutation distribution of S_ν and the rationality of $\mathcal{L}_\nu$.

Under H_o in (1.2), $\underset{\sim}{E}_\nu$ is composed of N_ν independent and identically distributed random vectors. Hence, the joint distribution of $\underset{\sim}{E}_\nu$ remains invariant under the N_ν! permutation of the N_ν vectors $\underset{\sim}{X}_{\nu 1},\ldots, \underset{\sim}{X}_{\nu N_\nu}$ among themselves when (1.2) holds. We consider now the basic rank-permutation principle, which is essentially due to Chatterjee and Sen (1966) and discussed also by Puri and Sen (1966). We define a $p \times N_\nu$ rank matrix $\underset{\sim}{R}_\nu$

$$(3.1)\qquad \underset{\sim}{R}_\nu = \begin{pmatrix} R_{\nu 1}^{(1)} & \cdots & R_{\nu N_\nu}^{(1)} \\ \vdots & & \vdots \\ R_{\nu 1}^{(p)} & \cdots & R_{\nu N_\nu}^{(p)} \end{pmatrix} = (\underset{\sim}{R}_{\nu 1}, \ldots, \underset{\sim}{R}_{\nu N_\nu}) ,$$

where $\underset{\sim}{R}_{\nu i} = (R_{\nu i}^{(1)}, \ldots, R_{\nu i}^{(p)})'$ for $i = 1, \ldots, N_\nu$. Each row of $\underset{\sim}{R}_\nu$ is a permutation of the numbers $1, \ldots, N_\nu$, there being thus in all $(N_\nu!)^p$ possible realizations of $\underset{\sim}{R}_\nu$. Let us rearrange the columns of $\underset{\sim}{R}_\nu$ in such a way that the first row has elements $1, \ldots, N_\nu$ in the natural order, and denote the corresponding matrix by $\underset{\sim}{R}_\nu^*$. $\underset{\sim}{R}_\nu$ is said to be permutationally equivalent to a matrix $\underset{\sim}{R}_\nu^{(1)}$, if it is possible to obtain $\underset{\sim}{R}_\nu^{(1)}$ from $\underset{\sim}{R}_\nu$ only by permutations of the columns of the latter. Thus, corresponding to each $\underset{\sim}{R}_\nu^*$, there will be a set $\Sigma(\underset{\sim}{R}_\nu^*)$ of $N_\nu!$ possible realizations of $\underset{\sim}{R}_\nu$, such that any member of this set will be permutationally equivalent to $\underset{\sim}{R}_\nu^*$. Now, the probability distribution of $\underset{\sim}{R}_\nu$ over the $(N_\nu!)^p$ possible realizations will depend on $F(\underset{\sim}{x})$, even when H_o in (1.2) holds (unless $F(\underset{\sim}{x})$ has coordinate-wise independent marginals). Thus, rank-statistics, like $\underset{\sim}{S}_\nu$ in (2.6), are, in general, not distribution-free under H_o in (1.2). However, given a particular set $\Sigma(\underset{\sim}{R}_\nu^*)$ (of $N_\nu!$ realizations), the conditional distribution of $\underset{\sim}{R}_\nu$ over the $N_\nu!$ over the $N_\nu!$ permutations of the columns of $\underset{\sim}{R}_\nu^*$ would be uniform under H_o in (1.2), i.e.,

$$(3.2)\qquad P\{\underset{\sim}{R}_\nu = \underset{\sim}{r}_\nu \mid \Sigma(\underset{\sim}{R}_\nu^*), H_o\} = \frac{1}{N_\nu!} \quad \text{for all} \quad \underset{\sim}{r}_\nu \in \Sigma(\underset{\sim}{R}_\nu^*) ,$$

whatever be $F(\underset{\sim}{x})$. Let us denote by $\mathcal{P}_\nu$ the permutational (conditional) probability measure generated by the conditional law in (3.2). Then, by routine computations (along the lines of Puri and Sen (1966)), we obtain that

$$(3.3) \qquad E\{\underset{\sim}{S}_\nu | \mathcal{P}_\nu\} = \underset{\sim}{0}^{p\times q} .$$

$$(3.4) \qquad E\{S_{\nu,jk}\, S_{\nu,j'k'} | \mathcal{P}_\nu\} = d_{\nu,jk;j'k'} \quad \text{for } j,j' = 1,\ldots, p, \; k,k' = 1,\ldots, q,$$

where $d_{\nu,jk;j'k'}$'s are defined by (2.15).

Since $\underset{\sim}{S}_\nu$ is a stochastic matrix, for actual test construction it is more convenient to use a real valued function of $\underset{\sim}{S}_\nu$ as a test-statistics. By an adaption of the same arguments as in Chatterjee and Sen (1966) and Puri and Sen (1966), we may work with a positive-semidefinite quadratic form in the pq elements of $\underset{\sim}{S}_\nu$, where the discriminant of the quadratic form is the inverse of the matrix $\underset{\sim}{D}_\nu$, which has the elements $d_{\nu,jk;j'k'}$ in (3.2). This leads to the test statistic $\mathcal{L}_\nu$, defined by (2.17). $\mathcal{L}_\nu$ will be stochastically large if $\underset{\sim}{S}_\nu$ is stochastically different from $\underset{\sim}{0}$. For small values of ν (i.e., N_ν), the conditional distribution of $\mathcal{L}_\nu$, given $\Sigma(R_\nu^*)$, can be obtained with the aid of (3.2), and a conditionally distribution-free test for H_0 in (1.2) based on $\mathcal{L}_\nu$ can be constructed. This, however, requires the evaluation of the $N_\nu!$ realizations of $\underset{\sim}{S}_\nu$ (under $\mathcal{P}_\nu$), while $\underset{\sim}{D}_\nu$ is $\mathcal{P}_\nu$-invariant. The task becomes prohibitively laborious and for large values of ν, we are forced to consider the following large sample approach in which we approximate the true permutation distribution of $\mathcal{L}_\nu$ (under $\mathcal{P}_\nu$) by that of a chi-square distribution with pq degrees of freedom (d.f.). As a basis of this study, we consider the following.

THEOREM 3.1. *Let* $H_\nu(\underset{\sim}{x}) = (1/N_\nu) \sum_{i=1}^{N_\nu} F_{\nu i}(\underset{\sim}{x})$, *and let* $\underset{\sim}{\Lambda}_\nu = \Lambda(H)|_{H \equiv H_\nu}$ *be defined as in* (2.21). *Then, under the assumptions* (I) *to* (V) *of Section* 2, $[\underset{\sim}{M}_\nu - \underset{\sim}{\Lambda}_\nu] \xrightarrow{p} \underset{\sim}{0}^{p\times p}$ *as* $\nu \longrightarrow \infty$. *Thus, under the assumptions* (VI) *and* (VII), $\underset{\sim}{M}_\nu$ *is positive definite, in probability, and* $\underset{\sim}{D}_\nu - \underset{\sim}{\Gamma}_\nu(H_\nu) \xrightarrow{p} 0^{pq\times pq}$ *as* $\nu \longrightarrow \infty$.

Proof: The proof that $\underset{\sim}{M}_\nu - \underset{\sim}{\Lambda}_\nu \xrightarrow{p} \underset{\sim}{0}^{p\times p}$ as $\nu \longrightarrow \infty$ follows more or less on the same line as in the proof of Theorem 4.2 of Puri and Sen (1966), and hence, for intended brevity, the details are omitted. The second part of the theorem is a consequence of the first part and the results in (2.9) through (2.23). Q.E.D.

THEOREM 3.2. *Under* $\mathcal{P}_\nu$, *the joint distribution of the elements of* $\underset{\sim}{S}_\nu$ *converges, in probability, to a multivariate normal distribution with null means and covariances given by* (3.4).

Proof: The computation of the means and covariances of the elements of $\underset{\sim}{S}_\nu$ follows directly from (3.3) and (3.4). Since $\underset{\sim}{D}_\nu$ is positive definite, in probability (by the preceding theorem), the asymptotic multinormality of the permutation distribution of $\underset{\sim}{S}_\nu$ follows readily by an appeal to the multivariate extension of the well-known Wald-Wolfowity-Noether-Hoeffding-Motoo-Hájek permutational central limit theorems [cf. Hájek (1961, p. 522)]. Hence, the theorem.

By virtue of the preceding theorem, we arrive at the following theorem through a few simple steps.

THEOREM 3.3. *Under the conditions* (I) *to* (VIII) *of Section* 2, $\mathcal{L}_\nu$, *defined by* (2.17), *has a permutation distribution (under* $\mathcal{P}_\nu$) *asymptotically converging, in probability, to a chi-square distribution with* pq *degrees of freedom*.

Hence we have the following large sample test procedure:

$$\text{(3.5)} \qquad \text{if } \mathcal{L}_\nu \begin{cases} \geq \chi^2_{pq,\alpha}\ , & \text{reject } H_o \text{ in (1.2);} \\ < \chi^2_{pq,\alpha}\ , & \text{accept } H_o, \end{cases}$$

where $P\{\chi^2_t \geq \chi^2_{t,\alpha}\} = \alpha (0 < \alpha < 1)$, the level of significance.

4. *Asymptotic multinormality of* $\underset{\sim}{S}_\nu$. We introduce the following notations. Let

$$\text{(4.1)} \qquad H_\nu(\underset{\sim}{x}) = (1/N_\nu) \sum_{i=1}^{N_\nu} F_{\nu i}(\underset{\sim}{x})\ ,$$

and let $H_{\nu[j]}(x)$ be the marginal of the jth variate of $H_\nu(\underset{\sim}{x})$, for j = 1,..., p. Also, let $F_{\nu i[j]}(x)$ and $F_{\nu i[j,j']}(x,y)$ be the univariate (jth) and bivariate ((j,j')th) marginals corresponding to $F_{\nu i}(\underset{\sim}{x})$, for $j \neq j' = 1,\ldots, p$. We define

$$(4.2) \quad A_{jj(i;rs)} = \iint_{-\infty<x<y<\infty} F_{\nu i[j]}(x)[1-F_{\nu i[j]}(y) \cdot \phi_j'(H_{\nu[j]}(x))\phi_j'(H_{\nu[j]}(y))dF_{\nu r[j]}(x)dF_{\nu s[j]}(y) + \iint_{-\infty<x<y<\infty} F_{\nu i[j]}(x)[1-F_{\nu i[j]}(y)] \cdot \phi_j'(H_{\nu[j]}(x))\phi_j'(H_{\nu[j]}(y))dF_{\nu s[j]}(x)dF_{\nu r[j]}(y) ,$$

for i, r, s = 1,..., N_ν and j = 1,..., p;

$$(4.3) \quad A_{jj'(i;rs)} = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} [F_{\nu i[j,j']}(x,y)-F_{\nu i[j]}(x)F_{\nu i[j']}(y)] \cdot \phi_j'(H_{\nu[j]}(x))\phi_{j'}'(H_{\nu[j']}(y))dF_{\nu r[j]}dF_{\nu s[j']}(y),$$

for $j \neq j' = 1,\ldots, p$ and i, r, s = 1,..., N_ν; the subscript ν in (4.2) and (4.3) is understood. Let then

$$(4.4) \quad \sigma_{\nu,jk;j'k'} = (1/N_\nu^2) \sum_{i=1}^{N_\nu}\sum_{r=1}^{N_\nu}\sum_{s=1}^{N_\nu} (c_{\nu r}^{(k)}-c_{\nu i}^{(k)})(c_{\nu s}^{(k')}-c_{\nu i}^{(k')}) \cdot A_{jj'(i;rs)} ,$$

for j, j' = 1,..., p and k, k' = 1,..., q; the corresponding pq×pq matrix is denoted by $\underset{\sim}{\Sigma}_\nu$. Throughout this section it will be assumed that (i) $F(\underset{\sim}{x})$ in (1.1) is absolutely continuous having a continuous density function $f(\underset{\sim}{x})$, $\underset{\sim}{x} \in R^p$, and (ii)

(4.5) $\underset{\sim}{\Sigma}_\nu$ <u>is positive definite and finite</u>.

Before presenting the main theorem of this section, we consider the following lemma.

LEMMA 4.1. <u>Under</u> (1.1), (2.8) <u>and</u> (2.9), $\underset{\sim}{\Sigma}_\nu - \underset{\sim}{\Lambda}(F) \otimes \underset{\sim}{C}_\nu \longrightarrow \underset{\sim}{0}^{pq\times pq}$ <u>as</u> $\nu \longrightarrow \infty$. <u>Hence under</u> (2.11) <u>and</u> (2.23), $\underset{\sim}{\Sigma}_\nu$ <u>is positive definite in the limit as</u> $\nu \longrightarrow \infty$.

<u>Proof</u>: By virtue of (2.8) and (2.9), $\max_{1\le i\le N_\nu} \max_{1\le k\le q} |c_{\nu i}^{(k)}| \longrightarrow 0$ as $\nu \longrightarrow \infty$, and hence

$$(4.6) \qquad \lim_{\nu\to\infty} \{ \max_{1\le i\le N_\nu} \sup_{\underset{\sim}{x}\in R^p} |F_{\nu i}(\underset{\sim}{x}) - F(\underset{\sim}{x})|\} = 0 .$$

From (4.2), (4.3), (4.6) and some routine computations, it follows that

$$(4.7) \qquad \lim_{\nu\to\infty} \{ \max_{1\le i,r,s\le N_\nu} \max_{1\le j,j'\le p} [A_{jj'}(i;rs) - \eta_{jj'}(F)]\} = 0 ,$$

where $\eta_{jj'}(F)$ is defined by (2.20). From (4.4), (4.7), (2.8) and (2.9), we obtain after some simplifications that

$$(4.8) \qquad \sigma_{\nu,jk;j'k'} - C_{\nu,kk'}\eta_{jj'}(F) \longrightarrow 0 \quad \text{as } \nu \longrightarrow \infty ,$$

for all $j,j' = 1,\ldots, p$; $k,k' = 1,\ldots, q$. This completes the first part of the proof. The second part follows from

(2.11), (2.23) (with $H \equiv F$), (4.8) and some well-known results on limits of sequences. Q.E.D.

Let us now introduce the following notations. Let

(4.9) $$B_{j(i',i)}(X_{\nu i}^{(j)}) = \int_{-\infty}^{\infty} [u(x-X_{\nu i}^{(j)})-F_{\nu i[j]}(x)]\phi_j'(H_{\nu[j]}(x))dF_{\nu i'[j]}(x) ,$$

(where $u(x)$ is equal to 1 or 0 according as x is positive or not), for $j = 1,\ldots, p$, $i,i' = 1,\ldots, N_\nu$. Also let

(4.10) $$Z_{\nu i(jk)} = (1/N_\nu) \sum_{r=1}^{N_\nu} (c_{\nu r}^{(k)}-c_{\nu i}^{(k)})B_{j(r,i)}(X_{\nu i}^{(j)}) ,$$

for $j = 1,\ldots, p$; $k = 1,\ldots, q$ and $i = 1,\ldots, N_\nu$. Finally let

(4.11) $$Z_{\nu,jk} = \sum_{i=1}^{N_\nu} Z_{\nu i(jk)} \quad \text{for } j = 1,\ldots, p; \; k = 1,\ldots, q.$$

Straight-forward but somewhat lengthy computations yield that

(4.12) $$\operatorname{Cov}\{Z_{\nu,jk}, Z_{\nu,j'k'}\} = \sigma_{\nu,jk;j'k'} ,$$

$$\text{for all } j,j' = 1,\ldots,p; \quad k,k' = 1,\ldots,q,$$

where $\sigma_{\nu,jk,j'k'}$'s are defined by (4.4).

In this context, we bring now the findings of Hájek (1966), who considered the special case of $p = q = 1$ and discussed the asymptotic normality of the corresponding rank-order statistic (which is just any one of the $S_{\nu,jk}$'s). Thus referring to his elegant paper and thereby avoiding the details, we obtain that

$$\lim_{\nu\to\infty} E\{[S_{\nu,jk}-E(S_{\nu,jk})] - Z_{\nu,jk}\}^2/Var(Z_{\nu,jk}) = 0 , \tag{4.13}$$

for all $j = 1,\ldots, p$; $k = 1,\ldots, q$. We shall utilize (4.13) to establish a comparatively stronger result. With this end in view, we first consider the following lemma.

LEMMA 4.2. *Let* $\underset{\sim}{a}_\nu = (a_{\nu 1},\ldots, a_{\nu t})$ *and* $\underset{\sim}{b}_\nu = (b_{\nu 1},\ldots, b_{\nu t})$ $t \geq 1$, *be two sequences of stochastic vectors, such that* $V(a_{\nu j} - b_{\nu j})/V(a_{\nu j}) \longrightarrow 0$ *as* $\nu \longrightarrow \infty$ *for all* $j = 1,\ldots, t$. *Then*

$$[Cov(a_{\nu j},a_{\nu \ell}) - cov(b_{\nu j},b_{\nu \ell})]/\{V(a_{\nu j})V(a_{\nu \ell})\}^{1/2} \longrightarrow 0 \quad \text{as } \nu \longrightarrow \infty$$

for all $j, \ell = 1,\ldots, t$.

The proof follows by expressing $cov(a_{\nu j},a_{\nu \ell})$ as $cov(b_{\nu j},b_{\nu \ell}) + cov(b_{\nu j},a_{\nu \ell}-b_{\nu \ell}) + cov(a_{\nu j}-b_{\nu j},b_{\nu \ell})$ $+ cov(a_{\nu j}-b_{\nu j},a_{\nu \ell}-b_{\nu \ell})$, applying Cauchy-Schwarz inequality to the second, third and fourth terms on the right hand side and making use of the conditions stated in the lemma.

By virtue of (4.12), (4.13) and lemma 4.2, it follows that

$$\lim_{\nu\to\infty} [cov\{S_{\nu,jk},S_{\nu,j'k'}\} - \sigma_{\nu,jk,j'k'}]/\{\sigma_{\nu,jk;jk} \cdot \sigma_{\nu,j'k';j'k'}\}^{1/2} = 0 , \tag{4.14}$$

for all $j, j' = 1, \ldots, p$; $k, k' = 1, \ldots, q$.

An immediate consequence of Lemma 4.2 is the following.

LEMMA 4.3. If $\underset{\sim}{a}_\nu$ has asymptotically a multinormal distribution with mean vector $\underset{\sim}{\alpha}_\nu$ and dispersion matrix $\underset{\sim}{A}_\nu$ and if for each $j(= 1, \ldots, t)$ $V(a_{\nu j} - b_{\nu j})/V(a_{\nu j}) \longrightarrow 0$ as $\nu \longrightarrow \infty$, then b_ν has also asymptotically the same multinormal law.

Thus it follows from (4.13) and Lemmas 4.2 and 4.3 that for proving the asymptotic normality of $\underset{\sim}{S}_\nu$, it is sufficient to show that $\{Z_{\nu, jk}, j = 1, \ldots, p; k = 1, \ldots, q\}$ has asymptotically a multinormal distribution. This will be accomplished by showing that any arbitrary linear combination of $Z_{\nu, jk}$'s has asymptotically a normal distribution. With this end in view, we define

$$(4.15) \qquad Z_\nu = \sum_{j=1}^{p} \sum_{k=1}^{q} \ell_{jk} Z_{\nu, jk} ,$$

where ℓ_{jk}'s are real and finite and not all of them are zero. From (4.10), (4.11) and (4.15), we obtain that

$$(4.16) \qquad Z_\nu = \sum_{i=1}^{N_\nu} g_i(\underset{\sim}{X}_{\nu i}) ,$$

where

$$(4.17) \quad g_i(\underset{\sim}{X}_{\nu i}) = \frac{1}{N_\nu} \sum_{r=1}^{N_\nu} \sum_{j=1}^{p} \sum_{k=1}^{q} \ell_{jk} (c_{\nu i}^{(k)} - c_{\nu i}^{(k)}) B_{j(r,i)}(X_{\nu i}^{(j)}) ,$$

for $i = 1, \ldots, N_\nu$. By definition, $g_i(\underset{\sim}{X}_{\nu i})$ are independent random variables. Further

$$(4.18)\quad \sum_{i=1}^{N_\nu} \mathrm{Var}\{g_i(\underset{\sim}{X}_{\nu i})\} = \sum_{j=1}^{p} \sum_{k=1}^{q} \sum_{j'=1}^{p} \sum_{k'=1}^{q} \ell_{jk}\ell_{j'k'}\sigma_{\nu,jk;j'k'}$$

is finite and positive by (4.5). Proceeding then precisely on the same line as in Hájek (1966, Section 5 specifically and using a linear combination such as (4.15) with his Y_ν's), it can be shown that under (4.5), the random variables $\{g_i(\underset{\sim}{X}_{\nu i})\}$ satisfy the Lindeberg condition for the applicability of the central limit theorem. Hence, we arrive at the following.

THEOREM 4.1. *Under the assumptions (I) - (V) of Section 2 and (4.5), the elements of $\underset{\sim}{S}_\nu - E\underset{\sim}{S}_\nu$ has asymptotically a multivariate normal distribution with zero means and dispersion matrix $\underset{\sim}{\Sigma}_\nu$.*

Now, under the model (1.1) and granted the assumptions made in Section 2, it follows from routine computations that

$$(4.19)\quad E\{S_{\nu,jk}\} = \left(\sum_{\ell=1}^{q} (\beta_{j\ell}C_{\nu,k\ell}\right)\left(\int_{-\infty}^{\infty} \frac{d}{dx}\, \phi_j(F_{[j]}(x))dF_{[j]}(x)\right),$$

for all $j = 1,\ldots, p$; $k = 1,\ldots, q$; and by Lemma 4.1, the covariance matrix $\underset{\sim}{\Sigma}_\nu$ is asymptotically equivalent to $\underset{\sim}{\Lambda}(F)\otimes \underset{\sim}{C}_\nu$. In the next section, we shall make use of this result for studying the asymptotic properties of the test based on $\mathcal{L}_\nu$ in (2.17).

5. Asymptotic properties of the test. We denote by

$$(5.1)\quad \nu_{jj'}(F) = \eta_{jj'}(F)/\left(\int_{-\infty}^{\infty} \frac{d}{dx}\,\phi_j(F_{[j]}(x))\,dF_{[j]}(x)\right) \cdot \left(\int_{-\infty}^{\infty} \frac{d}{dx}\,\phi_{j'}(F_{[j']}(x))\,dF_{[j']}(x)\right),$$

for all $j, j' = 1, \ldots, p$, where $\eta_{jj'}$'s are defined by (2.19), (2.20) and (2.21). Then, using Theorem 4.1, Lemma 4.1, (4.19) and some well-known result on the limiting distributions of quadratic forms in asymptotically normally distributed random variables, it follows that

$$(5.2)\qquad \mathcal{L}_\nu^* = \sum_{j=1}^{p}\sum_{j'=1}^{p}\sum_{k=1}^{q}\sum_{k'=1}^{q} S_{\nu,jk}S_{\nu,j'k'}\ \eta^{jj'}(F)\cdot C_\nu^{kk'}$$

has asymptotically (under (1.1) and the assumptions of Section 2) a non-central chi-square distribution with pq degrees of freedom and non-centrality parameter

$$(5.3)\qquad \Delta_{\mathcal{L}} = \sum_{j=1}^{p}\sum_{j'=1}^{p}\sum_{k=1}^{q}\sum_{k'=1}^{q} \nu^{jj'}(F)\beta_{jk}\beta_{j'k'}C_{\nu,kk'},$$

where

$$(5.4)\qquad \underset{\sim}{\nu}^{-1}(F) = ((\nu(F)))^{-1} = ((\nu^{jj'}(F))).$$

Further, using Theorems 3.1 and 4.1, it is straight-forward to show that under the assumptions made in Section 2,

$$(5.5)\qquad \mathcal{L}_\nu - \mathcal{L}_\nu^* \xrightarrow{p} 0 \quad \text{as } \nu \longrightarrow \infty .$$

Thus, we arrive at the following theorem.

THEOREM 5.1. Under (1.1) and the assumptions made in Section 2, the statistic $\mathcal{L}_\nu$ has asymptotically a non-central chi-square distribution with pq degrees of freedom and non-centrality parameter $\Delta_{\mathcal{L}}$, defined by (5.3).

Theorem 5.1 may be used to study the asymptotic power properties of the test in (3.5) and to compare its efficiency relative to standard parametric tests. Since in general linear hypothesis testing problems, the different test-criteria do not have uniformly good or bad performances, [cf. Anderson (1958, Chapter 8)], it may be quite involved to give a neat picture of the relative-efficiency of the proposed tests with respect to all the parametric ones. Among the notable parametric tests, Roy's largest characteristic root-criterion is admissible and powerful against certain alternatives, while the likelihood ratio test is so for certain other alternatives. However, borrowing the ideas of Wald (1943), it can be shown that in the sense of Wald, the likelihood ratio test has asymptotically the best average power (on suitable ellipsoids in the parameters β_{jk}'s). In this sense, the likelihood ratio test may be regarded as optimum. As such, we will consider the asymptotic relative efficiency of our proposed test with respect to the likelihood ratio test.

Following the notations of Anderson (1958, Chapter 8), we define the log-likelihood criterion (adjusted by a suitable multiplying factor) by T_ν (say,) and conclude from his Theorem 8.6.2 that T_ν has asymptotically (under H_0 in (1.2)) a chi-square distribution with pq degrees of freedom. Hence, the same test procedure as in (3.5) applies to the likelihood ratio test (with $\mathcal{L}_\nu$ replaced by T_ν). We denote the covariance matrix (assumed to be finite) of $F(\underset{\sim}{x})$ by $\underset{\sim}{\zeta} = ((\zeta_{jk}(F)))_{j,k} = 1,\ldots,p$ and assume it to be positive definite. The corresponding reciprocal matrix is denoted by $\underset{\sim}{\zeta}^{-1} = ((\zeta^{jk}(F)))$. Since the likelihood ratio-test criterion is exclusively based on linear estimators of $\underset{\sim}{\beta}$, generalizing the method of approach of Eicker (1963) and following essentially some routine steps it can be shown that under the assumptions made in Section 2, T_ν has asymptotically a non-central chi-square distribution with pq degrees of freedom and noncentrality parameter

$$(5.6) \qquad \Delta_T = \sum_{j=1}^{p} \sum_{j'=1}^{p} \sum_{k=1}^{q} \sum_{k'=1}^{q} \zeta^{jj'}(F)\beta_{jk}\beta_{j'k'}C_{\nu,kk'} .$$

It is seen that in multivariate situations, the asymptotic relative efficiency (A.R.E.) of the test based on $\mathcal{L}_\nu$ with respect to the one based on T_ν, as measured by $\Delta_{\mathcal{L}}/\Delta_T$, depends not only on $((\nu_{jj'}(F)))$ and $((\zeta_{jj'}(F)))$ but also on $\underset{\sim}{\beta}$ and $\underset{\sim}{C}_\nu$. The following points are worth noting in this context.

(i) If $p = 1$, no matter whatever be $\underset{\sim}{\beta}$ and $\underset{\sim}{C}_\nu$

$$\Delta_{\mathcal{L}}/\Delta_T = \zeta_{11}(F)/\nu_{11}(F) , \tag{5.7}$$

which happens to coincide with the usual efficiency expressions for the two sample location problem. A few particular cases may be worth mentioning here. First, if $\phi(u) = u$ (i.e., the Wilcoxon scores case), (5.7) reduces to $12\zeta_{11}(F)(\int_{-\infty}^{\infty} f^2(x)dx)^2$, which has known values (as well as lower bounds) for various $F(x)$. Secondly, if $\phi(x)$ is the inverse of the standard normal cumulative distribution function (i.e., the normal scores case), (5.7) is always greater than or equal to unity, the equality sign holds only when F is normal. This clearly indicates the asymptotic efficiency of the proposed tests for $p = 1$.

(5.8) (ii) If $F(\underset{\sim}{x}) \equiv \prod_{j=1}^{p} F_{[j]}(x_j), \quad \underset{\sim}{x} \in R^p$,

it readily follows that both $\underset{\sim}{\zeta}(F)$ and $\underset{\sim}{\nu}(F)$ are diagonal matrices and as such

$$\frac{\Delta_{\mathcal{L}}}{\Delta_T} = \frac{\sum_{j=1}^{p} \frac{1}{\nu_{jj}(F)} \sum_{k=1}^{q} \sum_{k'=1}^{q} \beta_{jk}\beta_{jk'}C_{\nu,kk'}}{\sum_{j=1}^{p} \frac{1}{\zeta_{jj}(F)} \sum_{k=1}^{q} \sum_{k'=1}^{q} \beta_{jk}\beta_{jk'}C_{\nu,kk'}} . \tag{5.9}$$

Hence, if we use the normal scores for $\phi_{\nu j}$'s, again (5.9) becomes greater than or equal to unity, where the equality sign holds only when F(x) is multinormal.

(iii) If we use the normal scores for $\phi_{\nu j}$'s, and if the parent cdf $F(\underset{\sim}{x})$ is a multinormal cdf, it readily follows from (2.19), (2.20) and (5.1) that $\underset{\sim}{\nu}(F) = \underset{\sim}{\zeta}(F)$, and hence, from (5.3), (5.6) that $\Delta_{\mathcal{L}}/\Delta_{T} = 1$. Or in otherwords, for parent normal distributions, the use of normal scores leads to asymptotically most efficient tests.

(iv) In general, for arbitrary $F(\underset{\sim}{x})$, $\Delta_{\mathcal{L}}/\Delta_{T}$ is bounded below and above by the minimum and maximum characteristic roots of $\underset{\sim}{\zeta}(F)\underset{\sim}{\nu}^{-1}(F)$, (the proof follows by a straight-forward application of a theorem by Courant (cf. [11]) on the bounds of the ratio of two quadratic forms). Now, the bounds of $\underset{\sim}{\zeta}(F)\underset{\sim}{\nu}^{-1}(F)$ have been studied in detail by the authors (Sen and Puri (1967)) in connection with the multivariate one sample location problem. As such, to avoid repetition, the details are omitted.

6. Asymptotic optimality of $\mathcal{L}_\nu$ for certain $F(\underset{\sim}{x})$. We shall study the asymptotic optimality of $\mathcal{L}_\nu$ for a class of parent distributions. We assume that $F(\underset{\sim}{x})$ has the absolutely continuous density function $f(\underset{\sim}{x})$, and define,

$$(6.1)\quad f(x_i|x_1,\ldots, x_{i-1}, x_{i+1},\ldots, x_p) = f(x_1,\ldots, x_p)/\int_{-\infty}^{\infty} f(x_1,\ldots, x_p)dx_i ,$$

$$(6.2) \qquad f_{[i]}(x) = \frac{d}{dx} F_{[i]}(x) , \quad \text{for } i = 1,\ldots, p .$$

Let then

$$(6.3) \qquad f_i'(x_i|x_1,\ldots, x_p) = \frac{d}{dx_i} f(x_i|x_1,\ldots, x_p) ,$$

$$(6.4) \qquad f_{i[i]}'(x) = \frac{d}{dx} f_{[i]}(x) , \quad \text{for } i = 1,\ldots, p .$$

We define the statistics

$$(6.5) \qquad U_{\nu,jk} = \sum_{i=1}^{N_\nu} c_{\nu i}^{(k)} [f_j'(X_{\nu i}^{(j)}|\underset{\sim}{X}_{\nu i}^{(1)},\ldots, \underset{\sim}{X}_{\nu i}^{(p)})/f(\underset{\sim}{X}_{\nu i})] ,$$

for $j = 1,\ldots, p$ and $k = 1,\ldots, q$. We also define

$$(6.6) \quad \xi_{jj'} = E\{f_j'(X_{\nu i}^{(j)}|\underset{\sim}{X}_{\nu 1},\ldots, \underset{\sim}{X}_{\nu p})f_{j'}'(X_{\nu i}^{(j')}|\underset{\sim}{X}_{\nu 1},\ldots, \underset{\sim}{X}_{\nu p})/f^2(\underset{\sim}{X}_{\nu i})|H_o\}$$

for all $j,j' = 1,\ldots, p$, and let $\underset{\sim}{T} = ((\tau_{jk,j'k'}))$ be the $pq \times pq$ matrix whose elements are $c_{\nu,kk'} \times \xi_{jj'}$, where $\underset{\sim}{C}_\nu$ is defined by (2.10). Finally let $\underset{\sim}{T}_\nu^{-1} = ((\tau_\nu^{jk,j'k'}))$ be the reciprocal matrix of $\underset{\sim}{T}_\nu$, and

$$(6.7) \qquad U_\nu^* = \sum_{j=1}^{p} \sum_{j\leq 1}^{p} \sum_{k=1}^{q} \sum_{k'=1}^{q} \tau_\nu^{jk,j'k'} U_{\nu,jk} U_{\nu,j'k'} .$$

Then by an adaptation of the line of approach of Wald (1943), we can show that (i) under H_o in (1.2), U_ν^* has asymptotically a chi-square distribution with pq degrees of freedom (provided $((\xi_{jj'}))$ is finite and positive definite) and (ii) the test based on U_ν^* has asymptotically the best average power (on the family of ellipsoids

$$(6.8) \quad \sum_{j=1}^{p} \sum_{j'=1}^{p} \sum_{k=1}^{q} \sum_{k'=1}^{q} \xi^{jj'} c_{\nu,kk'} \beta_{jk} \beta_{jk'} = a^2 > 0) .$$

We shall show that under certain conditions on $F(\underset{\sim}{x})$, the test based on $\mathcal{L}_\nu$ has also asymptotically the best average power on the same family of ellipsoids, provided ϕ_j's are chosen suitably. Our treatment deals with a set of sufficient conditions for this asymptotic optimality and the authors are not aware of necessary conditions for the same.

Regarding $F(\underset{\sim}{x})$ we assume that

$$(6.9) \quad \frac{f_j'(x_j | x_1, \ldots, x_p)}{f(x_1, \ldots, x_p)} = \sum_{j'=1}^{p} h_{jj'} f'_{j'[j']}(x_j)/f_{[j']}(x_j) ,$$

where $h_{jj'}$'s are real constants, not all zero, for $j = 1, \ldots, p$. (6.9) holds for (i) the multivariate normal distribution, (ii) any coordinate-wise independent distribution, and it may also hold for many other distributions.

Let us now define

$$(6.10) \quad V_{\nu,jk} = \sum_{i=1}^{N_\nu} c_{\nu i}^{(k)} f'_{j[j]}(X_{\nu i}^{(j)})/f_{[j]}(X_{\nu i}^{(j)}) ,$$

for $j = 1, \ldots, p$, $k = 1, \ldots, q$. If (6.9) holds, it follows that $U_{\nu,jk}$'s are linear functions of $V_{\nu,jk}$'s, and consequently, the quadratic form based on $V_{\nu,jk}$'s (analogous to U_ν^*), will also have the same properties as that of U_ν^*.

We now define

$$(6.11) \quad \phi_j(u) = f'_{j[j]}(F_{[j]}^{-1}(u))/f_{[j]}(F_{[j]}^{-1}(u)) , \quad 0 < u < 1, \quad j = 1, \ldots, p ,$$

and define $S_{\nu,jk}$'s as in (2.5). Following then Hájek's (1962) elegant approach it is seen that $S_{\nu,jk} - V_{\nu,jk}$ converges in quadratic mean to zero for all $j = 1,\ldots, p$, $k = 1,\ldots, q$. Consequently, on using Lemma 4.2 and some routine computations, we obtain that for ϕ_j's given by (6.11), $\mathcal{L}_\nu$ in (2.17) is stochastically equivalent to the quadratic form in $V_{\nu,jk}$'s, under (1.1) and the assumptions of Section 2. Since, under (6.9), this quadratic form in $V_{\nu,jk}$'s is also equal to U_ν^* in (6.7), it follows that under (1.1), the assumptions made in Section 2 and (6.9), (6.11),

$$\mathcal{L}_\nu \overset{p}{\sim} U_\nu^* . \tag{6.12}$$

Consequently, we arrive at the following.

THEOREM 6.1. *Under (i) the model (1.1), (ii) the assumptions made in Section 2, and (iii) (6.9) and (6.10) $\mathcal{L}_\nu$ has asymptotically the best average power on the family of ellipsoids in β_{jk}'s, defined by (6.8).*

In particular, if $F(\underset{\sim}{x})$ is normal, the condition (6.9) is satisfied and (6.11) lead to the coordinate-wise normal scores. Hence, $\mathcal{L}_\nu$ based on normal scores statistics leads to asymptotically best test for the family of ellipsoids in (6.8). A special case of this theorem (that is, for $p = 1$) is dealt in an interesting power of Matthes and Truax (1965).

7. *A characterization of the multivariate multisample location problem.* Let $X_{k1},\ldots, X_{kn_k}$ be n_k independent and identically distributed p-variate random variables having a continuous p-variate cumulative distribution function $F_k(x)$ for $k=1,\ldots,c(\geq 2)$. In the multivariate multisample location problem [cf. Puri and Sen (1966) and the other references cited therein], we may let

$$(7.1) \qquad F_k(x) = F_1(x - \theta_k),\ K = 2,\ldots,\ c,\ \theta_1 = 0\ .$$

We define $N = n_1 + \ldots + n_c$, and consider a sequence of N p-vectors $Z_1,\ldots, Z_N$, of which the first n_1 observations are from the first sample, the next n_2 from the second sample and so on. Thus, (7.1) will correspond to (1.1), if we let $q = c - 1$, $\beta_{jk} = \theta_{j,k+1}$ for $k = 1,\ldots, c-1$, and where $c_{\nu i}$'s are null vectors for an $1 \leq i \leq n_1$, $c_{\nu i} = (0, 1, 0,\ldots, 0)'$ for $n_1 + 1 \leq i \leq n_1 + n_2$, and so on. Thus, the results derived in this paper, also generalizes the results of Puri and Sen (1966). The statistic $\mathcal{L}_\nu$ in (2.17) can be shown to be identical with the corresponding $\mathcal{L}_\nu$ in Puri and Sen (1966), when (7.1) holds. Consequently, the efficiency considerations made in Sections 5 and 6 also applies to the multisample multivariate location problem.

REFERENCES

[1] Anderson, T. W. An introduction to multivariate statistical analysis. New York: John Wiley, 1958.

[2] Chatterjee, S. K., and Sen, P. K. (1966). Nonparametric tests for the multivariate multisample location problem S. N. Roy memorial volume. (Edited by R. C. Bose et al). University of North Carolina (in press).

[3] Chernoff, H., and Savage, I. R. (1958). Asymptotic normality and efficiency of certain nonparametric test statistics. Ann. Math. Statist. 29, 972 - 994.

[4] Eicker, F. (1963). Asymptotic normality and consistency of the least squares estimators for families of linear regressions. Ann. Math. Statist. 34, 447 - 456.

[5] Hájek, J. (1961). Some extensions of the Wald-Wolfowitz-Noether theorem. Ann. Math. Statist. 32, 506 - 523.

[6] Hájek, J. (1962). Asymptotically most powerful rank order tests. Ann. Math. Statist. 33, 1124 - 1147.

[7] Hájek, J. (1966) Asymptotic normality of simple linear rank statistics under alternatives, I. (to be published).

[8] Matthes, T. K. and Truax, D. R. (1965). Optimal Invariant Rank Tests for the k-Sample Problem. Ann. Math. Statist. 36, 1207 - 1222.

[9] Puri, M. L., and Sen, P. K. (1966). On a class of multivariate multisample rank-order tests. Sankhya, Ser A. 28, 353 - 376.

[10] Rao, C. R., _Linear statistical inference and its applications_. New York: John Wiley, 1965.

[11] Sen, P. K., and Puri, M. L. (1967). On the theory of rank order tests for location in the multivariate one sample problem. _Ann. Math. Statist_. _38_, 1216-1228.

[12] Wald, A. (1943). Tests of statistical hypotheses concerning several parameters when the number of observations is large. _Trans Amer. Math. Soc_. _45_, 426 - 482.