STATISTICAL INFERENCE IN INCOMPLETE BLOCKS DESIGNS

MADAN L. PURI† AND HAROLD D. SHANE

SUMMARY

In an earlier paper (Shane and Puri, 1969), the authors developed a class of asymptotically nonparametric tests for a bivariate paired comparison model. This paper unifies and complements the results of the previous paper by deriving a class of genuinely distribution free tests for the same problem but under the more general framework of $p(\geqq 2)$-variate situations. This is done by exploiting the theory of permutation distribution under sign invariant transformations to a class of rank order statistics. Asymptotic ,properties of these permutation rank order tests are studied and certain stochastic equivalence relationship with a similar class of multisample extensions of the p-variate one sample rank order tests proposed by Sen and Puri (1967) are derived. The asymptotic power properties of these tests are also studied.

1 INTRODUCTION

Although a great deal of work has appeared in recent years in the area of univariate paired comparisons, very little attention has been paid to its development in the multivariate situations, that is, those situations, in which there is interest in several characteristics or attributes. Recently, Sen and David (1968) and Bradley and Davidson (1969) have considered the situations where the responses to the paired comparisons on each of the characteristics are obtained, and these responses are qualitative in nature. Both these tests are generalizations of the univariate sign test, and are asymptotically equivalent. Their ARE (Asymptotic Relative Efficiency) can be as low as zero relative to the normal theory test when the parent population is normal. Shane and Puri (1969) considered a class of asymptotic tests based on the ranks of the observations for the bivariate situations, and established the asymptotic superiority of some of the members of the proposed class over the Sen and David procedure for a wide class of alternatives. The present work extends the previous paper in two directions. (i) It provides the exact tests for the problem considered and this fills the gap left in the previous paper, and (ii), it

† Work supported under National Science Foundation Grant No. GP-12462.

extends the results of the previous paper to the p-variate situations. The theory is thus unified and an attempt is made to give it a complete form. We now formulate the problem.

Let us consider t treatments in an experiment which yields paired observations, namely, $(\mathbf{X}_{il}, \mathbf{Y}_{jl})$, $l = 1, \ldots, N_{il}$, obtained by N_{ij} independent paired comparisons for each pair (i,j) of treatments, $1 \leqslant i < j \leqslant t$. We assume that N_{ij} difference scores $\mathbf{Z}_{ij,l} = \mathbf{Y}_{jl} - \mathbf{X}_{il} = (X^{(1)}_{ij,l}, \ldots, X^{(p)}_{ij,l})$, $l = 1, \ldots, N_{ij}$, have a common absolutely continuous cdf (cumulative distribution function) $\Pi_{ij}(\mathbf{z}), \mathbf{z} = (x_1, \ldots, x_p)'$. This is the situation, for example, if in the analysis of incomplete blocks experiments with each block of size 2, one makes the assumption of additivity in the usual (multivariate) analysis of variance model. We are interested in testing the hypothesis

$$H: \Pi_{ij}(\mathbf{z}) = \Pi(\mathbf{z}) \quad \text{for all} \quad 1 \leqslant i < j \leqslant t, \tag{1.1}$$

where $\Pi(\mathbf{z})$ is diagonally symmetric about $\mathbf{z} = \mathbf{0}$ (that is, its density $\pi(\mathbf{z})$ is invariant under simultaneous changes of signs of all the coordinate variates), against the alternative that not all Π_{ij} are identical.

2 PERMUTATIONALLY DISTRIBUTION-FREE RANK TESTS

Let $c = \binom{t}{2}$ denote the number of all possible pairs, and label the pair (i,j) by

$$\alpha = (i-1)t + j - \binom{i+1}{2} \quad (1 \leqslant i < j \leqslant t).$$

Then $\mathbf{Z}_{\alpha 1}, \ldots, \mathbf{Z}_{\alpha N_\alpha}$ are independent observations corresponding to the αth pair, and they are distributed according to the absolutely continuous cdf

$$\Pi_\alpha(z) = \Pi_\alpha(x_1, \ldots, x_p)' \quad (\alpha = 1, \ldots, c).$$

Let

$$N = \sum_{\alpha=1}^{c} N_\alpha = \frac{1}{2} \sum_{i=1}^{t} \sum_{\substack{j=1 \\ i \neq j}}^{t} N_{ij}.$$

Under the null hypothesis H of no treatment difference,

$$\text{(i)} \ \mathbf{Z}_{\alpha r}, r = 1, \ldots, N_\alpha, \alpha = 1, \ldots, c$$

are N independent and identically distributed random variables each having the cdf $\Pi(\mathbf{z})$, and (ii) $\mathbf{Z}_{\alpha r}$ and $(-1)\mathbf{Z}_{\alpha r}$ have the same cdf $\Pi(\mathbf{z})$ for all $r = 1, \ldots, N_\alpha$, $\alpha = 1, \ldots, c$. Let us denote the sample point by

$$\mathbf{Y}_N = (\mathbf{Z}_{11}, \ldots, \mathbf{Z}_{1N_1}, \ldots, \mathbf{Z}_{cN_c}, \quad \mathbf{Z}_{\alpha r} = (X^{(1)}_{\alpha r}, \ldots, X^{(p)}_{\alpha r})', \tag{2.1}$$

and the sample space by y_N. Let $G^{(1)}_N$ be the group of transformations

$g_N^{(1)}$ given by

$$g_N^{(1)} Y_N = ((-1)^{j_1} Z_{11}, \ldots, (-1)^{j_{N_1}} Z_{1N_1}, \ldots, (-1)^{j_{N_c}} Z_{cN_c})$$
$$(j_i = 0, 1;\ i = 1, \ldots, N). \quad (2.2)$$

Where

$$(-1)\,\mathbf{Z}'_{\alpha r} = (-X_{\alpha r}^{(1)}, \ldots, -X_{\alpha r}^{(p)}) \quad (r = 1, \ldots, N_\alpha, \alpha = 1, \ldots, c).$$

Let $G_N^{(2)}$ be another group of transformations $g_N^{(2)}$ given by

$$g_N^{(2)} \mathbf{Y}_N = \text{some permutation of the columns of } \mathbf{Y}_N. \quad (2.3)$$

Finally, let
$$G_N = G_N^{(1)} G_N^{(2)}$$
be the direct product of the groups. Thus G_N is a finite group of $N!2^N$ transformations which map the sample space onto itself. Under H the joint distribution of $\mathbf{Y}_N$ remains invariant under G_N. Thus for every point $\mathbf{Y}_N$ of $\mathbf{y}_N$, there exists a set $S(\mathbf{Y}_N)$ of $N!2^N$ points which are obtained by operating G_N on $\mathbf{Y}_N$. Hence, under H, given $\mathbf{Y}_N$, all the $N!2^N$ sample points generated by G_N are equiprobable, each having the conditional probability $(N!2^N)^{-1}$. Let us denote this conditional probability measure defined over $\{g_N \mathbf{Y}_N, g_N \in G_N\}$ by $\mathscr{P}_N$. Then, if we consider any test function $\phi(\mathbf{Y}_N)$ depending on the completely specified permutational probability measure P_N, it follows that such a test will be strictly distribution free. Now in actual practice $\phi(\mathbf{Y}_N)$ has to be constructed with special attention to the class of alternatives in mind; and in most of the problems, $\phi(\mathbf{Y}_N)$ depends on $\mathbf{Y}_N$ through a single-valued statistic $S_N = S(\mathbf{Y}_N)$ formulated in a suitable manner. With this end in view, we consider a class of rank order tests for H in (1.1).

Let us rank the N elements in each row of $\mathbf{Y}_N$ in the increasing order of their absolute values, and denote the rank of $|X_{\alpha r}^{(i)}|$ by $R_{\alpha r}^{(i)}$. We obtain a $p \times N$ rank matrix

$$\mathbf{R}_N = \begin{pmatrix} R_{11}^{(1)} & \ldots & R_{1N_1}^{(1)} & \ldots & R_{cN_c}^{(1)} \\ \cdots & \cdots & \cdots & \cdots & \cdots \\ R_{11}^{(i)} & \ldots & R_{1N_1}^{(i)} & \ldots & R_{cN_c}^{(i)} \\ \cdots & \cdots & \cdots & \cdots & \cdots \\ R_{11}^{(p)} & \ldots & R_{1N_1}^{(p)} & \ldots & R_{cN_c}^{(p)} \end{pmatrix}, \quad (2.4)$$

where by virtue of the assumed continuity of Π_{ij}, the possibility of ties is neglected in probability. Let now

$$\{E_{N,\alpha}^{(i)}, \alpha = 1, \ldots, N, i = 1, \ldots, p\}$$

be a sequence of real numbers, and denote the value of $E_{N,\alpha}^{(i)}$ corresponding to $\alpha = R_{\alpha r}^{(i)}$ by $E_{N,R_{\alpha r}^{(i)}}^{(i)}$. We get a $p \times N$ matrix of general scores $\mathbf{E}_N$

corresponding to $\mathbf{R}_N$:

$$\mathbf{E}_N = \begin{pmatrix} E^{(1)}_{N,R^{(1)}_{11}} & \cdots & E^{(1)}_{N,R^{(1)}_{1N_1}} & \cdots & E^{(1)}_{N,R^{(1)}_{cN_c}} \\ \cdots & \cdots & \cdots & \cdots & \cdots \\ E^{(i)}_{N,R^{(i)}_{11}} & \cdots & E^{(i)}_{N,R^{(i)}_{1N_1}} & \cdots & E^{(i)}_{N,R^{(i)}_{cN_c}} \\ \cdots & \cdots & \cdots & \cdots & \cdots \\ E^{(p)}_{N,R^{(p)}_{11}} & \cdots & E^{(p)}_{N,R^{(p)}_{1N_1}} & \cdots & E^{(p)}_{N,R^{(p)}_{cN_c}} \end{pmatrix}. \tag{2.5}$$

Later on, in § 3, we shall impose certain restrictions on $\mathbf{E}_N$. Consider now the univariate rank order statistics co-ordinatewise:

$$T_{Nk(\alpha)} = N_\alpha^{-1} \sum_{r=1}^{N_\alpha} E^{(k)}_{N,R^{(k)}_{\alpha r}} C^{(k)}_{\alpha r} \quad (k = 1, \ldots, p, \alpha = 1, \ldots, c), \tag{2.6}$$

where

$$C^{(k)}_{\alpha r} = +1 \quad \text{if} \quad X^{(k)}_{\alpha r} > 0 \quad \text{and} \quad C^{(k)}_{\alpha r} = -1 \quad \text{if} \quad X^{(k)}_{\alpha r} < 0. \tag{2.7}$$

To consider the test statistic we first find the conditional mean and dispersion matrix of $T_{Nk(\alpha)}$, $k = 1, \ldots, p$, $\alpha = 1, \ldots, c$ under $\mathscr{P}_N$. Omitting the routine computations, they are

$$E[T_{Nk(\alpha)} | \mathscr{P}_N] = 0 \quad (k = 1, \ldots, p, \alpha = 1, \ldots, c). \tag{2.8}$$

$$\begin{aligned} \operatorname{cov}[T_{Nk(\alpha)}, T_{Nl(\beta)} | \mathscr{P}_N] &= N_\alpha^{-1} N_\beta^{-1} \sum_{r=1}^{N_\alpha} \sum_{s=1}^{N_\beta} E[E^{(k)}_{N,R^{(k)}_{\alpha r}} E^{(l)}_{N,R^{(l)}_{\beta s}} C^{(k)}_{\alpha r} C^{(l)}_{\beta s} | \mathscr{P}_N] \\ &= N_\alpha^{-1} a^*_{N,kl\delta_{\alpha\beta}} \end{aligned} \tag{2.9}$$

where

$$a^*_{N,kl} = N^{-1} \sum_{\alpha=1}^{c} \sum_{r=1}^{N_\alpha} E^{(k)}_{N,R^{(k)}_{\alpha r}} E^{(l)}_{N,R^{(l)}_{\alpha l}} C^{(k)}_{\alpha r} C^{(l)}_{\alpha r}. \tag{2.10}$$

The last equality of (2.9) follows from the fact that

$$E[E^{(k)}_{N,R^{(k)}_{\alpha r}} E^{(l)}_{N,R^{(l)}_{\beta s}} C^{(k)}_{\alpha r} C^{(l)}_{\beta s} | \mathscr{P}_N] = \delta_{\alpha\beta} \delta_{rs} a^*_{N,kl}, \tag{2.11}$$

where $\delta_{\alpha\beta}$ and δ_{rs} are the usual Kronecker deltas.

Now let us revert back to the original notation and write the pair (i,j) for α. Denote

$$U^{(k)}_{N,i} = \sum_{\substack{j=1 \\ j \neq i}}^{p} N^{\frac{1}{2}}_{ij} T_{Nk(ij)} \quad (i \neq j = 1, \ldots, p, \; k = 1, \ldots, p). \tag{2.12}$$

Then

$$E[U^{(k)}_{N,i} | \mathscr{P}_N] = 0, \tag{2.13}$$

$$\operatorname{cov}[(U^{(k)}_{N,i}, U^{(l)}_{N,i'})] = (t\delta_{ii'} - 1) a^*_{N,kl} \quad (i, i' = 1, \ldots, p; \; k, l = 1, \ldots, p). \tag{2.14}$$

Denote

$$\mathbf{A}^*_N = ((a^*_{N,kl})) \tag{2.15}$$

and assume that $\mathbf{A}^*_N$ is positive definite. (If $\mathbf{A}^*_N$ is singular, then we can work with the highest order non-singular minor of $\mathbf{A}^*_N$ and work with the corresponding variates). Then we consider the following statistic for testing H:

$$\mathscr{L}_N^* = \frac{1}{t} \sum_{i=1}^{p} \mathbf{U}_{N,i} \mathbf{A}_N^{*-1} \mathbf{U}'_{N,i}, \tag{2.16}$$

where

$$\mathbf{U}_{N,i} = (U_{N,i}^{(1)}, \ldots, U_{N,i}^{(p)}) \quad \text{and} \quad \mathbf{A}_N^{*-1} \text{ is the inverse of } \mathbf{A}_N^*. \tag{2.17}$$

From the remarks made earlier, it follows that given $\mathbf{Y}_N$, the permutation distribution of $\mathscr{L}_N^*$ would be strictly distribution free under H, and hence an exact size α test can be constructed using this permutation distribution of $\mathscr{L}_N^*$. However, to apply the test in practice, we would require to study all the $N!\,2^N$ possible permuted values of $\mathscr{L}_N^*$ for any given $\mathbf{Y}_N$. Naturally, the labor involved in this procedure 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 $\mathscr{L}_N^*$ by some simple law and to reduce the computational labor by that. This study will be taken up in the next section. In passing we may remark that if $t = 2$, then the statistic $\mathscr{L}_N^*$ is equivalent to a class of multivariate one-sample statistics considered in Sen and Puri (1967) which includes as special cases the multivariate one-sample Wilcoxon and the normal scores statistics among others. Thus the present procedures may be regarded as the multi-sample analogues of the multivariate one sample tests developed by Sen and Puri (1967). [For the corresponding univariate theory, the reader is referred to Puri and Sen (1969).]

3 ASYMPTOTIC DISTRIBUTION THEORY OF PERMUTATION RANK ORDER STATISTICS

We first introduce a few notations.

Let $F_k^{(\alpha)}(x)$, $H_k^{(\alpha)}(x)$ and $F_{kl}^{(\alpha)}(x, y)$ denote the marginal cdf's of $X_{\alpha r}^{(k)}$, $|X_{\alpha r}^{(k)}|$ and $(X_{\alpha r}^{(k)}, X_{\alpha r}^{(l)})$ respectively, $k, l = 1, \ldots, p$; $\alpha = 1, \ldots, c$.

Let $F_{k,N}^{(\alpha)}(x)$ and $H_{k,N}^{(\alpha)}(x)$ be the sample cdf's of $X_{\alpha r}^{(k)}$ and

$$|X_{\alpha r}^{(k)}| \quad (r = 1, \ldots, N_\alpha)$$

respectively for $\alpha = 1, \ldots, c$ and $k = 1, \ldots, p$.

Denote $\rho_N^{(\alpha)} = N_\alpha / N$ and let $\rho_N^{(\alpha)} \to \rho_0^{(\alpha)}$ as $N \to \infty$, and assume that

$$0 < \rho_0^{(\alpha)} < 1 \quad \text{for} \quad \alpha = 1, \ldots, c.$$

Let
$$F_{k,N}(x) = \sum_{\alpha=1}^{c} \rho_N^{(\alpha)} F_{k,N}^{(\alpha)}(x), \quad \Pi_k(x) = \sum_{\alpha=1}^{c} \rho_0^{(\alpha)} F_k^{(\alpha)}(x),$$

$$H_{k,N}(x) = \sum_{\alpha=1}^{c} \rho_N^{(\alpha)} H_{k,N}^{(\alpha)}(x), \quad H_k(x) = \sum_{\alpha=1}^{c} \rho_0^{(\alpha)} H_k^{(\alpha)}(x).$$

Thus $F_{k,N}(x)$ and $H_{k,N}(x)$ are the combined sample cdf's of $X_{\alpha r}^{(k)}$ and $|X_{\alpha r}^{(k)}|, r = 1, \ldots, N_\alpha$, $\alpha = 1, \ldots, c$ respectively whose population cdf's are $\Pi_k(x)$ and $H_k(x)$ respectively for each $k = 1, \ldots, p$.

Let $\Pi_{kl,N}^{(\alpha)}(x, y)$ be the sample cdf of $(X_{\alpha r}^{(k)}, X_{\alpha r}^{(l)}), r = 1, \ldots, N_\alpha$ and denote

$$\Pi_{kl,N}(x, y) = \sum_{\alpha=1}^{c} \rho_N^{(\alpha)} \Pi_{kl,N}^{(\alpha)}(x, y), \quad F_{kl}(x, y) = \sum_{\alpha=1}^{c} \rho_N^{(\alpha)} F_{kl}^{(\alpha)}(x, y).$$

Thus $\Pi_{kl,N}(x, y)$ is the combined sample cdf of $(X_{\alpha r}^{(k)}, X_{\alpha r}^{(l)})$ for $r = 1, \ldots, N_\alpha$; $\alpha = 1, \ldots, c$ whose population cdf is $F_{kl}(x, y)$. Finally denote

$$H_{kl}(x, y) = F_{kl}(x, y) + F_{kl}(-x, y) + F_{kl}(x, -y) + F_{kl}(-x, -y)$$

and

$$H_{kl,N}(x, y) = \Pi_{kl,N}(x, y) + \Pi_{kl,N}(-x, y) + \Pi_{kl,N}(x, -y) + \Pi_{kl,N}(-x, -y).$$

Note that under H, $\Pi_\alpha(\mathbf{Z}) = \Pi(\mathbf{Z})$ and the marginal cdf of $X_{\alpha r}^{(k)}$ will be denoted by $F_k(x)$, and that of $(X_{\alpha r}^{(k)}, X_{\alpha r}^{(l)})$ by $F_{kl}^0(x, y)$. $\Pi_\alpha(\mathbf{z})$ may depend upon N, although this is not stated explicitly for notational convenience.

Using the above notations, we can express $T_{Nk(\alpha)}$ defined by (2.6) equivalently as

$$T_{Nk(\alpha)} = \int_0^\infty J_{Nk}\left[\frac{N}{N+1} H_{k,N}(x)\right] d[F_{k,N}^{(\alpha)}(x) + F_{k,N}^{(\alpha)}(-x)]$$

$$(k = 1, \ldots, p;\ \alpha = 1, \ldots, c), \quad (3.1)$$

where $$J_{Nk}\left[\frac{r}{N+1}\right] = E_{Nk,r} \quad (r = 1, \ldots, N;\ k = 1, \ldots, p).$$

Although the functions

$$J_{Nj}\left[\frac{r}{N+1}\right] \text{ are defined only at } \frac{1}{N+1}, \ldots, \frac{N}{N+1},$$

we may extend their domain of definition to (0, 1) by letting them have constant values over

$$\left(\frac{r}{N+1}, \frac{r+1}{N+1}\right) \quad (r = 0, 1, \ldots, N).$$

Furthermore, we make the following assumptions:

Assumption 1. $$J_{Nk}(0) = 0 \quad (k = 1, \ldots, p).$$

Assumption 2. $\lim_{N\to\infty} J_{Nk}(u) = J_k(u)$ exists for $0 < u < 1$ and is not constant for all $k = 1, \ldots, p$.

Assumption 3. $J_k(u)$ is absolutely continuous, and

$$|J_k^{(i)}(u)| = |dJ_k(u)/du^i| \leq K[u(1-u)]^{\delta-i-\frac{1}{2}} \quad (i = 0, 1, 2),$$

for some K and some $\delta > 0$.

Assumption 4. $$\int_0^\infty \left\{J_{Nk}\left[\frac{1}{N+1} H_{k,N}(x)\right] - J_k\left[\frac{N}{N+1} H_{k,N}(x)\right]\right\} + d[F_{k,N}^{(\alpha)}(x) + F_{k,N}^{(\alpha)}(-x)] = o_p(N^{-\frac{1}{2}}) \quad (k = 1, \ldots, p).$$

Assumption 5. $$\int_{x=0}^\infty \int_{y=0}^\infty \left\{J_{Nk}\left[\frac{N}{N+1} H_{k,N}(x)\right] J_{Nl}\left[\frac{N}{N+1} H_{l,N}(y)\right] - J_k\left[\frac{N}{N+1} H_{k,N}(x)\right] J_l\left[\frac{N}{N+1} H_{l,N}(y)\right]\right\} dH_{kl,N}(x,y) = o_p(1).$$

Let us now define

$$a_{kl}^* = \begin{cases} \int_{x=0}^\infty \int_{y=0}^\infty J_k[H_k(x)]\, J_l[H_l(y)]\, dH_{kl}(x,y) \quad (k \neq l = 1, \ldots, p), \\ \int_{x=0}^\infty J_k^2[H_k(x)]\, dH_k(x) \quad (k = l = 1, \ldots, p). \end{cases} \tag{3.2}$$

and denote $$\mathbf{A}^* = ((a_{kl}^*)) \quad (k, l = 1, \ldots, l). \tag{3.3}$$

Assumption 6. $\mathbf{A}^*$ is positive definite.

Remark. Since $H_k(x), H_l(y)$ and $H_{kl}(x,y)$ may depend on N, the matrix $\mathbf{A}^*$ may depend on N. However we shall suppress this fact whenever there is no confusion. Assumptions 2 to 4 are needed to prove the asymptotic normality of the permutation distribution of $N^{\frac{1}{2}}[T_{Nk(\alpha)}, k = 1, \ldots, p, \alpha = 1, \ldots, c]$. Assumption 5 is required only for the asymptotic convergence of the permutation covariance matrix $\mathbf{A}_N^*$ defined in (2.15). Assumption 1 is unessential (*see* Puri and Sen (1969*a*)), but is introduced here to simplify certain computations.

Theorem 3.1. *Under assumptions* 1 *to* 4, $|\mathbf{A}_N^* - \mathbf{A}^*|$ *converges to* $\mathbf{0}$ *in probability as* $N \to \infty$, *where* $\mathbf{0}$ *is a null matrix of order* $p \times p$ *and* $\mathbf{A}_N^*$ *and* $\mathbf{A}^*$ *are defined by* (2.15) *and,* (3.3) *respectively.*

Proof. By virtue of assumptions 4 and 5, it follows that

$$\begin{aligned} a_{N,kl}^* &= \int_{x=0}^\infty \int_{y=0}^\infty J_{Nk}\left[\frac{N}{N+1} H_{k,N}(x)\right] J_{Nl}\left[\frac{N}{N+1} H_{l,N}(y)\right] dH_{kl,N}(x,y) \\ &= \int_{x=0}^\infty \int_{y=0}^\infty J_k\left[\frac{N}{N+1} H_{k,N}(x)\right] J_l\left[\frac{N}{N+1} H_{l,N}(y)\right] dH_{kl,N}(x,y) + o_p(1). \end{aligned}$$

Now proceeding precisely as in Puri and Sen (1966), it follows after omitting the details of computations, that

$$a^*_{N,kl} = \int_{x=0}^{\infty}\int_{y=0}^{\infty} J_k[H_k(x)]\, J_l[H_l(y)]\, dH_{kl}(x,y) + o_p(1).$$

The proof follows.

Corollary 3.1. *If* (i) $\Pi_\alpha(\mathbf{z}) = \Pi(\mathbf{z} + \boldsymbol{\mu}_\alpha N^{-\frac{1}{2}})$ *where* $\mathbf{z} = (x_1, \ldots, x_p)'$,

$$\boldsymbol{\mu}_\alpha = (\mu_{1\alpha}, \ldots, \mu_{p\alpha})',$$

and $\Pi(\mathbf{z})$ *is a fixed absolutely continuous* cdf *diagonally symmetric about* $\mathbf{0}$, *and* (ii) *the conditions of Theorem* 3.1 *are satisfied, then* $\mathbf{A}^*_N \to \mathbf{A}$ *in probability, as* $N \to \infty$, *where* $\mathbf{A} = ((a_{kl}))$ *is given by*

$$a_{kl} = \begin{cases} \displaystyle\int_{x=0}^{\infty}\int_{y=0}^{\infty} J_k[2F_k(x)-1]\, J_l[2F_{[l]}(y)-1]\, dH^{(0)}_{kl}(x,y) \\ \qquad\qquad\qquad\qquad\qquad \textit{if} \quad (k \neq l = 1, \ldots, p), \\ \displaystyle\int_{x=0}^{1} J_k^2(x)\, dx \quad \textit{if} \quad (k = l = 1, \ldots, p), \end{cases} \tag{3.4}$$

where $H^{(0)}_{kl}(x,y)$ *is the value of* $H_{kl}(x,y)$ *when* $\boldsymbol{\mu}_\alpha = \mathbf{0}$.

The proof of this corollary is an immediate consequence of the above theorem, and is therefore omitted.

Theorem 3.2. *Under the assumptions* 1 *to* 4, *the random vectors*

$$[N_\alpha^{\frac{1}{2}}\mathbf{T}_{N(\alpha)}, \alpha = 1, \ldots, c] \quad \textit{where} \quad \mathbf{T}_{N(\alpha)} = (T_{N1(\alpha)}, \ldots, T_{Np(\alpha)}),$$

are under $\mathscr{P}_N$, *asymptotically independent and identically distributed* p*-variate normal random variables each having mean* $\mathbf{0}$, *and covariance matrix* $\mathbf{A}_N$.

Proof. From (2.6) and assumption 4,

$$T_{Nk(\alpha)} = T^*_{Nk(\alpha)} + o_p(N^{-\frac{1}{2}}) \quad (k = 1, \ldots, p, \alpha = 1, \ldots, c), \tag{3.5}$$

where

$$T^*_{Nk(\alpha)} = N_\alpha^{-1} \sum_{r=1}^{N_\alpha} C^{(k)}_{\alpha r} J_k[R^{(k)}_{\alpha r}/(N+1)] \quad (k = 1, \ldots, p, \alpha = 1, \ldots, c). \tag{3.6}$$

Define

$$S_{Nk(\alpha)} = N_\alpha^{-1} \sum_{r=1}^{N_\alpha} C^{(k)}_{\alpha r} J_k[U^{(k)}_{\alpha r}], \tag{3.7}$$

where, for each $k = 1, \ldots, p$, $U^{(k)}_{11}, \ldots, U^{(k)}_{1N_\alpha}, \ldots, U^{(k)}_{cN_c}$ are independent and identically distributed random variables having uniform distribution over $[0, 1]$. Denote

$$\mathbf{U}_{\alpha r} = (U^{(1)}_{\alpha r}, \ldots, U^{(p)}_{\alpha r})'. \tag{3.8}$$

Then $\mathbf{U}_{\alpha r}, r = 1, \ldots, N_\alpha, \alpha = 1, \ldots, c$ are independent and identically distributed random vectors distributed according to some p-variate cdf, say $G(\mathbf{x}), \mathbf{x} \in R^p$. (Note that if $\mathbf{Z}_{\alpha r} = (Z_{\alpha r}^{(1)}, \ldots, Z_{\alpha r}^{(p)})'$ has the cumulative distribution function $H(\mathbf{x}), \mathbf{x} \in R^p$, and if we let $\mathbf{Y} = (Y_1, \ldots, Y_p)'$, where

$$Y_k = F_k(Z_{\alpha r}^{(k)}), \tag{3.9}$$

F_k being the marginal cumulative distribution function of $Z_{\alpha r}^{(k)}$, then (i) the marginal cdf of each Y_k is uniform over $[0, 1]$, and (ii) the cdf $G(\mathbf{x})$ of $\mathbf{U}_{\alpha r}$ is the same as that of the $\mathbf{Y} = (Y_1, \ldots, Y_p)$.)

Using Theorem 4.3.1 of Ghosh (1969), it follows that

$$N_\alpha^{\frac{1}{2}} |T_{Nk(\alpha)}^* - S_{Nk(\alpha)}| = o_p(1) \quad (k = 1, \ldots, p, \alpha = 1, \ldots, c). \tag{3.10}$$

Thus from (3.5) and (3.10), the limit distribution of $N_\alpha^{\frac{1}{2}} \mathbf{T}_{N(\alpha)}, \alpha = 1, \ldots, c$, is the same as that of $N_\alpha^{\frac{1}{2}} \mathbf{S}_{N(\alpha)}$, where $\mathbf{S}_{N(\alpha)} = (S_{N1(\alpha)}, \ldots, S_{Np(\alpha)})'$, if the latter exists. Thus to prove the asymptotic normality of $N_\alpha^{\frac{1}{2}} \mathbf{T}_{N(\alpha)}$, $\alpha = 1, \ldots, c$, it suffices to show that, for any real

$$\delta_{k\alpha}, k = 1, \ldots, p, \alpha = 1, \ldots, c,$$

not all zero, the linear combination

$$W_N = \sum_{\alpha=1}^{c} N_\alpha^{\frac{1}{2}} \sum_{k=1}^{p} \delta_{k\alpha} S_{Nk(\alpha)} \tag{3.11}$$

is asymptotically normal.

Using (3.7), we can rewrite W_N as

$$W_N = \sum_{\alpha=1}^{c} N_\alpha^{\frac{1}{2}} \sum_{k=1}^{p} \delta_{k\alpha} N_\alpha^{-1} \sum_{r=1}^{N_\alpha} C_{\alpha r}^{(k)} J_k [U_{\alpha r}^{(k)}], \tag{3.12}$$

$$= \sum_{\alpha=1}^{c} \left[N_\alpha^{-\frac{1}{2}} \sum_{r=1}^{N_\alpha} V_{\alpha r} \right], \tag{3.13}$$

where

$$V_{\alpha r} = \sum_{k=1}^{p} \delta_{k\alpha} C_{\alpha r}^{(k)} J_k [U_{\alpha r}^{(k)}]. \tag{3.14}$$

(3.13) represents c-summations which involved independent random variables $V_{\alpha r}$ each of which (by repeated use of C_r-inequality) can be shown to have finite absolute moment of order $2+\delta'$, $0 < \delta' \leqslant 1$. Hence by the central limit theorem (Liapounoff), each sum properly normalised has normal distribution in the limit, with the result that the sum of c-summations will have normal distribution in the limit. This proves the joint asymptotic normality of $\mathbf{T}_{N(\alpha)}, \alpha = 1, \ldots, c$. Now using (2.9), (2.10), (3.5), (3.7) and (3.10), it can easily be verified that the covariances of the limit distribution is zero for $\alpha \neq \beta$, $k, l = 1, \ldots, p$. This coupled with the joint asymptotic normality of the variables $\mathbf{T}_{N(\alpha)}$, $\alpha = 1, \ldots, c$ establishes their asymptotic independence. Furthermore, since $\mathbf{T}_{N(\alpha)}$, $\alpha = 1, \ldots, c$ have identical dispersion matrices, the proof follows.

Theorem 3.3. *If the assumptions* 1 *to* 4 *and* 6 *hold, then under* H_0 *in* (1.1) *the permutation distribution of* $\mathscr{L}_N^*$ *defined by* (2.6) *is asymptotically* (*as* $N \to \infty$) *the chi-square with* $p(t-1)$ *degrees of freedom.*

The proof of this theorem follows as a consequence of Theorems 3.1 and 3.2, and is therefore omitted.

By virtue of Theorem 3.3, the permutation test procedure based on $\mathscr{L}_N^*$ simplifies in large samples to the following rule.

$$\begin{aligned} \mathscr{L}_N^* &\geqslant \chi^2_{\alpha, p(r-1)} \quad \text{reject } H, \\ &< \chi^2_{\alpha, p(r-1)} \quad \text{accept } H, \end{aligned}$$

where $\chi^2_{\alpha, p(t-1)}$ is $100(1-\alpha)\%$ point of the chi-square distribution with $p(t-1)$ degrees of freedom.

In order to study the asymptotic power properties of the test considered above, we require to study the unconditional asymptotic distribution of $\mathscr{L}_N^*$ under appropriate classes of alternatives. This, in turn, requires the study of joint asymptotic distribution of the rank order statistics defined in (2.6). For the case of $p = 2$, the same has been studied in Shane and Puri (1969). Since the methods for the case of any finite p are analogous for those for the case of $p = 2$, the details will be skipped as far as possible. However the efficiency results will be discussed in detail.

4 THE PROPOSED CLASS OF ASYMPTOTICALLY DISTRIBUTION FREE TESTS

To derive a class of asymptotically distribution free tests, we first state the following theorem.

Theorem 4.1. *Under the assumptions* 1 *to* 4 *of* §3, *the random variables* $\surd(N_\alpha)\,(T_{N1(\alpha)} - \mu_{N1(\alpha)}, \ldots, T_{Np(\alpha)} - \mu_{Np(\alpha)})$ $\alpha = 1, \ldots, c$ *have asymptotically a multivariate normal distribution with zero means and covariance matrix*

$$\operatorname{Cov}[\surd(N_\alpha)\,(T_{Nk(\alpha)} - \mu_{Nk(\alpha)}), \quad \surd(N_\beta)\,(T_{Nl(\beta)} - \mu_{Nl(\beta)})] = \sigma_{N, kl, \alpha\beta},$$

where $((\sigma_{N, kl, \alpha\beta}))$, $k, l = 1, \ldots, p$; $\alpha, \beta = 1, \ldots, c$ *are given by* (6.2), (6.3), (6.4) *and* (6.5) *respectively, and*

$$\mu_{Nk(\alpha)} = \int_{x=0}^{\infty} J[H_k(x)]\, d[F_k^{(\alpha)}(x) + F_k^{(\alpha)}(-x)]. \tag{4.1}$$

The proof of this theorem is long and involved, and is therefore not given here. For the case of $p = 2$, it is briefly sketched in (Shane and Puri, 1969, Theorem 3.1). For the case of any finite p, the details are given in Shane (1968). In fact it is established in Shane (1968), that subject to certain weak assumptions, the asymptotic normality holds uniformly with respect to $\Pi_\alpha(\mathbf{z})$, and $\rho_N^{(\alpha)}, \alpha = 1, \ldots, c$.

We now consider a sequence of admissible alternative hypotheses H_N, which specify that for each $\alpha = 1, \ldots, c$,

$$\Pi_\alpha(\mathbf{z}) = \Pi\left(\mathbf{z} + \frac{\mu_\alpha}{\sqrt{N}}\right),$$

where $\mathbf{z} = (x_1, \ldots, x_p)'$, $\mu_\alpha = (\mu_{1\alpha}, \ldots, \mu_{p\alpha})'$ and $\Pi(\mathbf{z})$ is a fixed continuous p-variate cdf diagonally symmetric about $\mathbf{0}$. We shall also assume that the constant $E_{Ni,r}$, $i = 1, \ldots, p$; $r = 1, \ldots, N$ is the expected value of the rth order statistics of a sample of size N from a distribution

$$\Psi_i(x) = \begin{cases} \Psi_i^*(x) - \Psi_i^*(-x) & \text{if} \quad x \geqslant 0, \\ 0 & \text{otherwise,} \end{cases} \tag{4.2}$$

where $\Psi_i^*(x)$ is a cdf symmetric about zero. If we denote by $J_i = \Psi_i^{-1}$ and $J_i^* = \Psi_i^{*-1}$ then the above definition of $E_{Ni,r}$ implies that

$$J_i(u) = J_i^*\left(\frac{u+1}{2}\right) \quad (0 < u < 1), \quad J_i^*(u) = J_i(2u-1) \quad (\tfrac{1}{2} < u < 1) \tag{4.3}$$

and $$J_i^*(u) = -J_i^*(1-u).$$

Remark. Two cases of special interest which we shall study in greater detail later are when (i) $\Psi_i^*(x)$ is the rectangular distribution over $(-1, 1)$ and (ii) $\Psi_i^*(x)$ is the standard normal distribution function. For the case (i) the corresponding statistic will be termed the rank sum statistic, and for the case (ii), the corresponding statistic will be termed the absolute normal scores statistic.

The following theorem, the proof of which follows from Theorem 4.1, plays a central role in deriving a class of asymptotically distribution free tests as well as in studying the efficiency properties.

Theorem 4.2. *If* (i) $\rho_N^{(\alpha)} \to \rho_\alpha$ *as* $N \to \infty$ *and* $0 < \rho_\alpha < 1$, $\alpha = 1, \ldots, c$.

(ii) *The conditions of Theorem* 4.1 *are satisfied.*

(iii) *For each fixed* N, *the hypothesis* H_N *is true. Then the random variables* $[\sqrt{(N_\alpha)}\,(T_{Nj(\alpha)} - u_{Nj(\alpha)}), j = 1, \ldots, p; \alpha = 1, \ldots, c]$ *have a limiting multivariate normal distribution as* $N \to \infty$ *with means zero and covariance matrix* $\boldsymbol{\tau} = (\tau_{ij,\alpha\beta})$, $i, j = 1, \ldots, p$; $\alpha, \beta = 1, \ldots, c$, *where*

$$\tau_{jj,\alpha\beta} = \delta_{\alpha\beta} A_j^2 \quad (j = 1, \ldots, p; \alpha, \beta = 1, \ldots, c), \tag{4.4}$$

$$\tau_{ij,\alpha\beta} = \tau_{ji\,\alpha\beta} = \delta_{\alpha\beta}\mathscr{S}_{ij} \quad (i \neq j = 1, \ldots, p; \alpha, \beta, = 1, \ldots, c), \tag{4.5}$$

$$A_i^2 = \int_0^1 J_i^2(u)\, du = \int_0^1 [J_i^*(u)]^2\, du \quad (i = 1, \ldots, p), \tag{4.6}$$

$$\mathscr{S}_{ij} = \int_{x=-\infty}^{+\infty} \int_{x=-\infty}^{+\infty} J_i^*(F_i(x))\, J_j^*(F_j(y))\, dH_{ij}^0(x, y), \tag{4.7}$$

$\delta_{\alpha\beta}$ is the Kronecker Delta and $H^0_{ij}(x, y)$ is the value of $H_{ij}(x, y)$ under H in (1.1).

Corollary 4.2.1. *Suppose that the hypothesis H_0 is true. Then under the conditions of Theorem* 4.1, *the random variables*

$$[\sqrt{(N_\alpha)}\, T_{Nj(\alpha)}, j = 1, \dots, p;\ \alpha = 1, \dots, c]$$

have a limiting multivariate normal distribution with means zero and covariance matrix $\boldsymbol{\tau} = (\tau_{ij,\alpha\beta}), i, j = 1, \dots, p; \alpha, \beta = 1, \dots, c$ given by (4.4) *and* (4.5).

Reverting back to our original notation, we have:

Corollary 4.2.2. *Under the assumptions of Theorem* 4.1, *the random variables $[\sqrt{(N_{ij})}\,(T_{Nk(i,j)} - \mu_{Nk(i,j)}), k = 1, \dots, p;\ 1 \leqslant i < j \leqslant t]$, where*

$$\mu_{Nk(i,j)} = \int_{x=0}^{\infty} J_k(H_k(x))\, d[F_k^{(ij)}(x) + F_k^{(ij)}(-x)] \qquad (k = 1, \dots, p;\ 1 \leqslant i < j \leqslant t) \quad (4.8)$$

have, in the limit as $N \to \infty$, multivariate normal distribution with zero mean and covariance matrix $\tau = (\tau_{kl(ij)(i'j')}), k, l = 1, \dots, p;\ 1 \leqslant (i, i') < (j, j') \leqslant t$ defined by

$$\tau_{kl(ij)(i'j')} = \delta_{(ij)(i'j')} A_k^2 \quad (k = 1, 2), \qquad (4.9)$$

$$\tau_{kl(ij)(i'j')} = \tau_{lk(ij)(i'j')} = \delta_{(ij)(i'j')} \mathscr{S}_{kl} \quad (k \neq l) \qquad (4.10)$$

where A_k^2 and $\mathscr{S}_{kl}$ are given by (4.11) *and* (4.12) *respectively and*

$$\delta_{(ij)(i'j')} = \begin{cases} +1 & \textit{if} \quad i = i', \quad j = j', \\ -1 & \textit{if} \quad i = j', \quad j = i', \\ 0 & \textit{otherwise.} \end{cases}$$

Now let

$$U_{Ni}^{(k)} = \sum_{\substack{j=1 \\ j \neq i}}^{t} \sqrt{(N_{ij})}\, T_{Nk(ij)}, \mu_{Nk(i\cdot)} = \sum_{\substack{j=1 \\ j \neq i}}^{t} \sqrt{(N_{ij})}\, \mu_{Nk(ij)} \qquad (k = 1, \dots, p;\ i = 1, \dots, t) \quad (4.11)$$

$$\eta_{k(ij)} = \lim_{N \to \infty} \sqrt{(N_{ij})}\, \mu_{Nk(ij)}, \quad \eta_{k(i\cdot)} = \sum_{\substack{j=1 \\ j \neq i}}^{t} \eta_{k(ij)} \qquad (4.12)$$

and assume that $\eta_{k(ij)}$ exists and is finite for $k = 1, \dots, p$ and $1 \leqslant i < j \leqslant t$. Let us introduce the following partitioned vectors

$$\mathbf{U}_N = (\mathbf{U}_N^{(1)}, \mathbf{U}_N^{(2)}, \dots, \mathbf{U}_N^{(p)}), \qquad (4.13)$$

where $$\mathbf{U}_N^{(k)} = (U_{N1}^{(k)}, \ldots, U_{Nt}^{(k)}) \quad (k = 1, \ldots, p); \tag{4.14}$$

$$\boldsymbol{\mu}_{N(\cdot)} = (\boldsymbol{\mu}_{N1(\cdot)}, \ldots, \boldsymbol{\mu}_{Np(\cdot)}) \tag{4.15}$$

$$\boldsymbol{\mu}_{Nk(\cdot)} = (\mu_{Nk(1\cdot)}, \ldots, \mu_{Nk(t\cdot)}) \quad (k = 1, \ldots, p); \tag{4.16}$$

$$\boldsymbol{\eta}_{(\cdot)} = (\boldsymbol{\eta}_{1(\cdot)}, \ldots, \boldsymbol{\eta}_{p(\cdot)}), \tag{4.17}$$

where $$\boldsymbol{\eta}_{k(\cdot)} = (\eta_{k(1\cdot)}, \ldots, \eta_{k(t\cdot)}) \quad (k = 1, \ldots, p). \tag{4.18}$$

Then we have the following theorem.

Theorem 4.3. *Under the assumptions of Theorem* 4.1, *the random vector* $\mathbf{U}_N = (\mathbf{U}_N^{(1)}, \ldots, \mathbf{U}_N^{(p)})$ *has asymptotically a multivariate normal distribution with mean vector* $\boldsymbol{\eta}_{(\cdot)} = (\boldsymbol{\eta}_{1(\cdot)}, \ldots, \boldsymbol{\eta}_{p(\cdot)})$ *and covariance matrix*

$$\mathbf{M} = (\mathbf{M}_{kl}) = \lim_{N\to\infty} E[(\mathbf{U}_N - \boldsymbol{\mu}_{N(\cdot)})' (\mathbf{U}_N - \boldsymbol{\mu}_{N(\cdot)})],$$

where $$\mathbf{M}_{kl} \quad (k, l = 1, \ldots, p);$$

is a $(t \times t)$ *matrix with entries*

$$\mathbf{M}_{kl.ij} = (t\delta_{ij} - 1)\, a_{kl} \quad (k, l = 1, \ldots, p;\ i, j = 1, \ldots, t), \tag{4.19}$$

where $$a_{kl} = \begin{cases} A_k^2 & \text{if} \quad k = l \\ \mathscr{S}_{kl} & \text{if} \quad k \neq l \end{cases}$$

and the rank of $\mathbf{M}$ *is* $p(t-1)$ *if and only if the matrix*

$$\mathbf{A} = (a_{kl}). \tag{4.20}$$

is nonsingular.

Proof. The asymptotic normality follows directly from Theorem 4.1 and Corollary 4.2.2. To compute the covariance matrix, we note that

$$\mathbf{M}_{kl} = \lim_{N\to\infty} E[\mathbf{U}_N^{(k)} - \mu_{Nk(\cdot)})' (\mathbf{U}_N^{(l)} - \mu_{Nl(\cdot)})]$$

so that
$$\begin{aligned} \mathbf{M}_{kl,ij} &= \lim_{N\to\infty} E[(U_{Ni}^{(k)} - \mu_{Nk(i\cdot)})(U_{Nj}^{(l)} - \mu_{Nl(j\cdot)})] \\ &= \lim_{N\to\infty} \operatorname{Cov}[U_{Ni}^{(k)}, U_{Nj}^{(l)}] \\ &= \lim_{N\to\infty} \operatorname{Cov}\left[\sum_{\substack{r=1 \\ r\neq i}}^{t} \sqrt{(N_{ir})}\, T_{Nk(ir)}, \sum_{\substack{s=1 \\ s\neq j}}^{t} \sqrt{(N_{js})}\, T_{Nl(js)}\right] \\ &= \sum_{\substack{r=1 \\ r\neq i}}^{t} \sum_{\substack{s=1 \\ s\neq j}}^{t} \lim_{N\to\infty} \sigma_{N,kl(ir)(js)} = \sum_{\substack{r=1 \\ r\neq i}}^{t} \sum_{\substack{s=1 \\ s\neq j}}^{t} \tau_{kl(ir)(js)} \\ &= (\tau\delta_{ij} - 1)\, a_{kl} \end{aligned}$$

by comparison with (4.9) and (4.10). Let

$$\mathbf{C} = (C_{ij}) = (t\delta_{ij} - 1) \quad (i,j = 1, \ldots, t) \tag{4.21}$$

Then

$$\mathbf{M} = \begin{pmatrix} a_{11}\mathbf{c} & a_{12}\mathbf{c} & \ldots & a_{1p}\mathbf{c} \\ \vdots & & & \\ a_{p1}\mathbf{c} & a_{p2}\mathbf{c} & \ldots & a_{pp}\mathbf{c} \end{pmatrix}. \tag{4.22}$$

Let $\|\mathbf{M}\|$ = Determinant $\mathbf{M}$. In $\|\mathbf{M}\|$ subtracting row 1 from each of the rows numbered $2, \ldots, t$, row $(t+1)$ from each of the rows $(t+2), \ldots, 2t$, and so on leaves $\|\mathbf{M}\|$ unchanged. Then add columns $2, \ldots, t$ to column 1; $(t+2), \ldots, 2t$ to $(t+1)$ etc. and the resulting determinant has columns $1, t+1, 2t+1, \ldots, (p-1)t+1$ all zero. Hence the rank of $\mathbf{M}$ is at most $p(t-1)$.

Now striking the 1st, $(t+1)$st, $\ldots, [(p-1)t+1]$st rows and columns leaves the minor,

$$\|\mathbf{M}^*\| = \begin{vmatrix} ta_{11}\mathbf{I} & \ldots & ta_{1p}\mathbf{I} \\ \vdots & & \\ ta_{p1}\mathbf{I} & \ldots & ta_{pp}\mathbf{I} \end{vmatrix},$$

where $\mathbf{I}$ is a $(t-1)\times(t-1)$ identity matrix. The fact that $\|\mathbf{M}^*\| = 0$ if and only if $\|\mathbf{A}\| = 0$ follows from

Lemma 4.1. *Let* $\mathbf{A}$ *be any symmetric matrix and let*

$$\mathbf{M}^* = \begin{pmatrix} a_{11}\mathbf{I} & \ldots & a_{1p}\mathbf{I} \\ \vdots & & \\ a_{p1}\mathbf{I} & \ldots & a_{pp}\mathbf{I} \end{pmatrix},$$

where I is a $(k\times k)$ identity matrix. Then $\|\mathbf{M}^*\| = \|\mathbf{A}\|^k$.

The proof being trivial, is omitted.

Remark: We have proved that the rank of $\mathbf{M}$ is $p(t-1)$ if and only if the matrix $\mathbf{A}$, which is the dispersion matrix of $\{J_1^*(F_1(X^{(1)})), \ldots, J_p^*(F_p(X^{(p)}))\}$ is non-singular. In what follows, we make the assumption that the distribution function $\Pi(\mathbf{z})$ and the score function $J_k, k = 1, \ldots, p$ are such that the moment matrix A is non-singular. The moment matrix will be singular if and only if there exists one or more relationships of the form

$$\sum_{j=1}^{p} a_j J_j^*(F_j(X^{(j)})) = \text{b. a.s. } \Pi.$$

Let us define

$$\mathscr{L}_N^{**} = \frac{1}{t}\sum_{i=1}^{t} (U_{Ni}^{(1)}, \ldots, U_{Ni}^{(p)})\,\mathbf{A}^{-1}(U_{Ni}^{(1)}, \ldots, U_{Ni}^{(p)})'. \tag{4.23}$$

We now have,

Theorem 4.4. *Under the assumptions of Theorem* 4.1

$$\mathscr{L}_N^{**} = \frac{1}{t}\sum_{i=1}^{t}(U_{Ni}^{(1)}, \ldots, U_{Ni}^{(p)})\,\mathbf{A}^{-1}(U_{Ni}^{(1)}, \ldots, U_{Ni}^{(p)})'$$

has asymptotically a non-central χ^2 *distribution with* $p(t-1)$ *degrees of freedom and non-centrality parameter*

$$\Delta_{\mathscr{L}} = \frac{1}{t}\sum_{i=1}^{t}(\eta_{1(i\cdot)}, \eta_{2(i\cdot)}, \ldots, \eta_{p(i\cdot)})\,\mathbf{A}^{-1}(\eta_{1(i\cdot)}, \ldots, \eta_{p(i\cdot)})'. \qquad (4.24)$$

The proof of this theorem follows as an application of Sverdrup's (1952) Theorem and the well known property (cf. Rao (1965), p. 443 (viii)) of the multivariate normal distribution. [For the case of $p = 2$, cf. Shane and Puri (1969).]

Corollary 4.3.1. *Suppose that the hypothesis* H_0 *is true. Then under the assumption of Theorem* 4.2, $\mathscr{L}_N^{**}$ *has the limiting central chi-square distribution with* $p(t-1)$ *degrees of freedom.*

Now let $\hat{\mathbf{A}}$ be a consistent estimator of $\mathbf{A}$, and denote

$$\mathscr{L}_N = \frac{1}{t}\sum_{i=1}^{t}(U_{Ni}^{(1)}, \ldots, U_{Ni}^{(p)})\,\hat{\mathbf{A}}^{-1}(U_{Ni}^{(1)}, \ldots, U_{Ni}^{(p)}). \qquad (4.25)$$

Then, it follows that $\mathscr{L}_N - \mathscr{L}_N^{**}$ converges to zero in probability as $N \to \infty$. Hence $\mathscr{L}_N$, too, has the limiting central chi-square distribution $p(t-1)$ degrees of freedom and so the critical function

$$\Phi(\mathscr{L}_N) = \begin{cases} 1 & \text{if} \quad \mathscr{L}_N \geqslant \chi^2_{p(t-1),\alpha}, \\ 0 & \text{if} \quad \mathscr{L}_N < \chi^2_{p(t-1),\alpha}, \end{cases}$$

where $\chi^2_{r,\alpha}$ is the $100(1-\alpha)\%$ point of the chi-squared distribution with r degrees of freedom, provides an asymptotically level α test of H.

From Theorem 4.3, it is clear that any consistent estimator of $\mathbf{A}^{-1}$ will preserve the asymptotic distribution of the test statistic. However to establish the existence of the test statistic $\mathscr{L}_N$, at least one consistent estimation of $\mathbf{A}^{-1}$ has to be proposed. From (2.10) and Theorem 3.1, it follows that the permutational covariance matrix $\mathbf{A}_N^{-1*}$ defined in (2.15) can be taken as a consistent estimator of $\mathbf{A}^{-1}$.

Theorem 4.5. *The permutation test based on* $\mathscr{L}_N$ *given by* (2.16) *and the asymptotically nonparametric test based on* $\mathscr{L}_N$ *given by* (4.25) *are asymptotically power equivalent for sequence of alternatives* H_N *defined above* (4.2).

The proof follows from Corollary (3.1), Theorem 4.2 and 4.3. Thus by virtue of the stochastic equivalence of the tests $\mathscr{L}_N^*$ and $\mathscr{L}_N$, we shall consider only the asymptotic properties of the unconditional test based on $\mathscr{L}_N$.

In most cases, the quantities $\eta_{j,\alpha} = \lim_{N\to\infty} N_\alpha^{\frac{1}{2}}\mu_{Nj(\alpha)}$ take on simple forms through the help of the following lemma similar to Lemma 7.2 of Puri (1964).

Lemma 4.2. *If* (i) $F_j, j = 1, \ldots, p$ *is continuous* cdf *differentiable in each of the open intervals* $[(0, a_1^{(j)}), \ldots, (a_s^{(j)}, \infty)]$ *and the derivative of* F_j *is bounded in each of the intervals and* $F_j(x)+F_j(-x) = 1$. (ii) *The function*

$$\frac{d}{dx} J_j\left[F_j(x)-F_j(-x)\right] \quad (j = 1, \ldots, p)$$

is bounded as $x \to +\infty$, $x \to 0+$. (iii) $J_j = \Psi_j^{-1}$ *and* Ψ_j^* *defined by* (4.2) *is symmetric and unimodal with density* Ψ_j^*; $j = 1, \ldots, p$. *Then*

$$\left.\begin{aligned} \eta_{j(\alpha)} &= \lim_{N\to\infty} \surd(N_\alpha) \int_{x=0}^{\infty} J_j\left[\sum_{\beta=1}^{c} \rho_N^{(\beta)}\left\{F_j\left(x+\frac{\mu_{j\beta}}{\surd N}\right) - F_j\left(-x+\frac{\mu_{j\beta}}{\surd N}\right)\right\}\right], \\ & d\left[F_j\left(x+\frac{\mu_{j\alpha}}{\surd N}\right)+F_j\left(-x+\frac{\mu_{j\alpha}}{\surd N}\right)\right] \\ &= -2\surd(\rho_\alpha)\mu_{j\alpha}\int_{x=0}^{\infty}\frac{d}{dx}J_j\left[2F_j(x)-1\right]dF_j(x). \end{aligned}\right\} \quad (4.26)$$

In case the conditions of Lemma 4.2 are satisfied, then

$$\Delta_{\mathscr{L}} = \frac{1}{t}\sum_{i=1}^{t} \mathcal{O}_i \mathbf{B}^{*-1}\mathcal{O}_i', \quad (4.27)$$

where
$$\mathcal{O}_i = (\mathcal{O}_{1i}, \ldots, \mathcal{O}_{pi}); \quad \mathcal{O}_{ki} = \sum_{\substack{j=1 \\ j\neq i}}^{t} \rho_{ij}^{\frac{1}{2}}\mu_{kij} \quad (k = 1, \ldots, p), \quad (4.28)$$

$$B^* = a_{kl/a_k^* a_l^*} \quad (k, l = 1, \ldots, p), \quad (4.29)$$

$$a_k^* = -2\int_{x=0}^{\infty}\frac{d}{dx}J_k[2F_k(x)-1]\,dF_k(x) \quad (k = 1, \ldots, p) \quad (4.30)$$

and a_{kl} is defined by (4.20).

Introducing the notations

$$\mathscr{L}_N = \mathscr{L}_N(R) \text{ when } J_k^*(x) = 2x-1 \ (k = 1, \ldots, p) \text{ (rank sum test)} \quad (4.31)$$

and

$$\mathscr{L}_N = \mathscr{L}_N(\Phi) \text{ when } J_k^*(x) = \Phi^{-1}(x) \ (k = 1, \ldots, p) \text{ (absolute normal scores test).} \quad (4.32)$$

We obtain from (4.27),

$$\Delta_{\mathscr{L}(R)} = \frac{1}{t}\sum_{i=1}^{t} \mathcal{O}_i \mathbf{\Lambda}_R^{-1} \mathcal{O}_i', \quad \mathbf{\Lambda}_R = ((\lambda_{kl}^{(R)})), \tag{4.33}$$

where
$$\lambda_{kl}^{(R)} = \begin{cases} \dfrac{1}{12}\left[\displaystyle\int_{-\infty}^{+\infty} f_k^2(x)\,dx\right]^{-2} & \text{if} \quad k = l, \\ \dfrac{\displaystyle\int_{-\infty}^{+\infty}\int_{-\infty}^{+\infty} F_k(x)\,F_l(y)\,dF_{kl}(x,y) - \frac{1}{4}}{\displaystyle\int_{-\infty}^{+\infty} f_k^2(x)\,dx \int_{-\infty}^{+\infty} f_l^2(x)\,dx} & \text{if} \quad k \neq l, \end{cases} \tag{4.34}$$

$$\Delta_{\mathscr{L}(\Phi)} = \frac{1}{t}\sum_{i=1}^{t} \mathcal{O}_i \mathbf{\Lambda}_\Phi^{-1} \mathcal{O}_i', \quad \mathbf{\Lambda}_\Phi = ((\lambda_{kl}^{(\Phi)})), \tag{4.35}$$

where

$$\lambda_{kl}^{(\Phi)} = \begin{cases} \left[\displaystyle\int_{-\infty}^{+\infty} \frac{f_k^2(x)\,dx}{\phi[\Phi^{-1}(F_k(x))]}\right]^{-2} & \text{if} \quad k = l, \\ \dfrac{\displaystyle\int_{-\infty}^{+\infty}\int_{-\infty}^{+\infty} \Phi^{-1}(F_k(x))\,\Phi^{-1}(F_l(y))\,dF_{kl}(x,y)}{\displaystyle\int_{-\infty}^{+\infty} \frac{f_k^2(x)\,dx}{\phi[\Phi^{-1}(F_k(x))]}\int_{-\infty}^{+\infty} \frac{f_l^2(x)\,dx}{\phi([\Phi^{-1}(F_l(x))]}} & \text{if} \quad k \neq l. \end{cases} \tag{4.36}$$

Sen and David (1968) have also considered a permutationally distribution free test statistic D_N which is a p-variate extension of the univariate sign test. They have shown that under the sequence of alternatives H_N, D_N has the limiting noncentral chi-square distribution with $p(t-1)$ degrees of freedom and non-centrality parameter Δ_D which in the notation of the present paper is

$$\Delta_D = \frac{1}{t}\sum_{i=1}^{t} \mathcal{O}_i \mathscr{D}^{-1} \mathcal{O}_i'^{-1}, \tag{4.37}$$

where $\mathscr{D} = (d_{kl})$ is given by

$$d_{kl} = \begin{cases} \dfrac{\Pi_{kl}(0,0) - \frac{1}{4}}{f_k(0) f_l(0)} & \text{if} \quad k \neq l, \\ \dfrac{1}{4 f_k(0) f_l(0)} & \text{if} \quad k = l. \end{cases} \tag{4.38}$$

In addition, we shall consider the likelihood ratio $\mathscr{F}$-test obtained by assuming that $\mathbf{Z}_{ijl}$ is distributed according to a normal distribution with mean $\boldsymbol{\mu}_i - \boldsymbol{\mu}_j$ and covariance matrix $\mathbf{\Sigma}$. This statistic can be shown to be asymptotically equivalent to

$$\mathscr{F} = \frac{1}{t}\sum_{i=1}^{t} (\hat{\mu}_i^{(1)} \ldots \hat{\mu}_i^{(p)})\,\hat{\Sigma}^{-1}(\hat{\mu}_i^{(1)} \ldots \hat{\mu}_i^{(p)})', \tag{4.39}$$

where
$$\hat{\mu}_i^{(k)} = \sum_{\substack{j=1 \\ j \neq i}}^{t} \frac{1}{N_{ij}} \sum_{l=1}^{N_{ij}} X_{ijl}^{(k)} \tag{4.40}$$

and $\hat{\mathbf{\Sigma}}$ is the parametric estimator of the covariance matrix

$$\mathbf{\Sigma} = (\sigma_{kl}) \quad (k, l = 1, \ldots, p), \tag{4.41}$$

where
$$\sigma_{kl} = \mathrm{cov}\,(X^{(k)}, X^{(l)}). \tag{4.42}$$

Under H_N, $\mathscr{F}$ has asymptotically a non-central chi-square distribution with $p(t-1)$ degrees of freedom and non-centrality parameter

$$\Delta_{\mathscr{F}} = \frac{1}{t} \sum_{i=1}^{t} \mathcal{O}_i \mathbf{\Sigma}^{-1} \mathcal{O}_1'. \tag{4.43}$$

5 ASYMPTOTIC RELATIVE EFFICIENCY

It is well known (Puri, 1964) that in the situations we are considering the asymptotic efficiency of one statistic relative to another is equal to the ratio of their non-centrality parameters. Thus, denoting e_{T_1,T_2} as the asymptotic efficiency of T_1 relative to T_2, we have

$$\left.\begin{aligned} e_{\mathscr{L}_N(\Phi),\mathscr{F}} &= \Delta_{\mathscr{L}(\Phi)}/\Delta_{\mathscr{F}} : e_{\mathscr{L}_N(R),\mathscr{F}} = \Delta_{\mathscr{L}(R)}/\Delta_{\mathscr{F}}, \\ e_{\mathscr{L}_N(\Phi),D_N} &= \Delta_{\mathscr{L}(\Phi)}/\Delta_D, \\ \text{and} \qquad \Delta_{\mathscr{L}_N(\Phi),\mathscr{L}_N(R)} &= \Delta_{\mathscr{L}(\Phi)}/\Delta_{\mathscr{L}(R)}. \end{aligned}\right\} \tag{5.1}$$

The above expressions do not lend themselves to easy analysis. They depend upon the underlying distribution $\Pi(\mathbf{z})$ and also on the vectors $\mathcal{O}_i$ and t, the number of treatments. Useful information may be obtained, however, for certain special cases.

Case 1. *Multivariate normal case.* Let us assume that the underlying distribution is a non-singular p-variate normal distribution with mean vector zero and covariance matrix $\mathbf{\Sigma} = ((\sigma_{kl})), k, l = 1, \ldots, p$. Then

$$\lambda_{kl}^{(\Phi)} = \sigma_{kl}, \quad \lambda_{kl}^{(R)} = 2\sigma_k \sigma_l \sin^{-1}\frac{\rho_{kl}}{2} \tag{5.2}$$

where
$$\rho_{kl} = \begin{cases} 1 & \text{if} \quad k = l, \\ \sigma_{kl}\sigma_k^{-1}\sigma_l^{-1} & \text{if} \quad k \neq l. \end{cases} \tag{5.3}$$

In such a case
$$e_{\mathscr{L}_N(\Phi),\mathscr{F}} = 1. \tag{5.4}$$

This means that for underlying non-singular normal distribution the property of the univariate normal scores test relative to the Student's t-test (or the Analysis of Variance $\mathscr{F}$-test) is preserved in the multivariate case.

Next from 6.2 we note that for any i, the expression

$$\frac{\mathcal{O}_i \Lambda_R^{-1} \mathcal{O}_i'}{\mathcal{O}_i \Sigma^{-1} \mathcal{O}_i'} \tag{5.5}$$

is the asymptotic relative efficiency of one sample p-variate rank sum test relative to Hotelling's T^2-test [cf. Sen and Puri (1969)]. Denoting by ϕ^* the class of all non-singular p-variate normal distributions we obtain (Sen and David, 1968; Shane, 1968; Shane and Puri, 1969)

$$\left.\begin{aligned} \inf_{\Pi\in\phi^*} \inf_{\mathcal{O}} e_{\mathscr{L}_N(R), \mathscr{F}} &= 0 \quad \text{for} \quad p \geqslant 3, \\ 3/\pi \leqslant e_{\mathscr{L}_N(R), \mathscr{F}} &\leqslant 0{\cdot}965 \quad \text{for} \quad p = 2, \end{aligned}\right\} \tag{5.6}$$

and from (5.4) and (5.6) to follow that

$$\sup_{\Pi\in\phi^*} \sup_{\mathcal{O}} e_{\mathscr{L}_N(\Phi), \mathscr{L}_N(R)} = \infty \quad \text{for} \quad p \geqslant 3.$$

Proceeding as in Bhattacharyya (1967),† we obtain

$$\inf_{\Pi\in\phi^*} \inf_{\mathcal{O}} e_{\mathscr{L}_N(\Phi), \mathscr{L}_N(R)} \geqslant 1 \quad \text{for} \quad p \geqslant 2.$$

The above results indicate that when the underlying distribution is normal, the $\mathscr{L}_N(\Phi)$ test is always preferably to the $\mathscr{L}_N(R)$ test. Similarly

$$\sup_{\Pi\in\phi^*} \sup_{\mathcal{O}} e_{\mathscr{L}_N(\Phi), D_N} = \infty \quad \text{for} \quad p \geqslant 3.$$

Case 2. *Independent co-ordinates.* Let $\Pi(\mathbf{z})$ have independent co-ordinates, then

$$\sigma_{kl} = \Lambda_{kl}^{(\Phi)} = \Lambda_{kl}^{(R)} = 0 \quad (k \neq l).$$

In such a case,

$$e_{\mathscr{L}_N(\Phi), \mathscr{F}} = \frac{\sum_{i=1}^{t} \sum_{k=1}^{p} \mathcal{O}_{k,i}^2 \Lambda_{kk}^{(\Phi)}}{\sum_{i=1}^{t} \sum_{k=1}^{p} \mathcal{O}_{k,i}^2 \sigma_k^{-2}}. \tag{5.7}$$

To obtain bounds for (5.7) we consider

$$e_A = \frac{\sum_{k=1}^{p} \mathcal{O}_{k,i}^2 \Lambda_{k,k}^{(\Phi)}}{\sum_{k=1}^{p} \mathcal{O}_{k,i}^2 \sigma_k^{-2}}. \tag{5.8}$$

Applying Courant's Theorem‡ we see that

$$\inf_{\mathcal{O}} e_A = \min\,[\sigma_k^2 \Lambda_{k,k}^{(\Phi)}, k = 1, \ldots, p]. \tag{5.9}$$

† *Ann. Math. Statist.* (1967), **38**, 1753–8.

‡ The maximal and minimal values of $\mathbf{x'Ax/x'Bx'}$ where A and B are non-negative definite and B is nonsingular are given by the maximal and minimal eigenvalues of AB^{-1}.

However, $\sigma_k^2 \Lambda_{kk}^{(\Phi)}$ is simply the asymptotic relative efficiency of the univariate one sample normal scores test relative to Student's t-test for the distribution F_k. As such it is known to be greater than or equal to 1 and is 1 only for F_k normal. Hence

$$\inf_{\Theta} e_A = 1, \tag{5.10}$$

$$\inf_{\Pi \in \mathscr{F}_0} \inf_{\Theta} e_{\mathscr{L}_N(\Phi), \mathscr{F}} = 1, \tag{5.11}$$

where $\mathscr{F}_0$ is the class of all absolutely continuous cdf's having independent co-ordinates.

In the same way, using the results of Chatterjee and Sen (1964), and Sen and Puri (1967), it can be shown that

$$\inf_{\Pi \in \mathscr{F}_0} \inf_{\Theta} e_{\mathscr{L}_N(\Phi)\ \mathscr{L}_N(R)} = \Pi/6, \tag{5.12}$$

$$\inf_{\Pi \in \mathscr{F}_0} \inf_{\Theta} e_{\mathscr{L}_N(R), \mathscr{F}} = 0{\cdot}864, \tag{5.13}$$

$$\inf_{\Pi \in \mathscr{F}_0^*} \inf_{\Theta} e_{D_N, \mathscr{F}} = 0{\cdot}33, \tag{5.14}$$

where $\mathscr{F}_0^*$ is the class of all absolutely continuous diagonally symmetric unimodal distributions with independent co-ordinates.

Case 3. *Identical equicorrelated marginals.* Let us assume

$$F_{ij}(x,y) = F(x,y) \quad \text{for} \quad 1 \leqslant i \neq j \leqslant p.$$

Proceeding as earlier, we obtain the

$$\sup_{\Theta} e_{\mathscr{L}_N(\Phi), \mathscr{F}} = \{\sigma_1^2 a_\Phi^2\} \max \left\{ \frac{1-\rho}{1-\rho_\sigma}, \frac{1+(p-1)\rho}{1+(p-1)\rho_\Phi} \right\}, \tag{5.15}$$

$$\inf_{\Theta} e_{\mathscr{L}_N(\Phi), \mathscr{F}} = \{\sigma_1^2 a_\Phi^2\} \min \left\{ \frac{1-\rho}{1-\rho_\Phi}, \frac{1+(p-1)\rho}{1+(p-1)\rho_\Phi} \right\}, \tag{5.16}$$

and similarly, one can evaluate $\inf_{\Theta} e_{\mathscr{L}_{N(R)}, \mathscr{F}}, \sup_{\Theta} e_{\mathscr{L}_{N(R)}, \mathscr{F}}$ etc. The bounds for these eigenvalues are not in general known. However, for specific distributions, their values may be computed.

6 APPENDIX

The expressions for $((\sigma_{N,kl,\alpha\beta}))$, $k, l = 1, \ldots, p; \alpha, \beta = 1, \ldots, c$.

In this section we give the expressions for $\sigma_{N,k\,l,\alpha\beta}$. For the details of computations, *see* Shane (1968). Denote

$$\left.\begin{aligned}
A_k^{(\alpha)}(x,y) &= F_k^{(\alpha)}(x)\,[1-F_k^{(\alpha)}(y)], \quad B_{kl}(x,y) = dJ_k[H_k(x)]\,dJ_l[H_l(y)],\\
C_{kl}^{(\alpha)}(x,y) &= d[F_k^{(\alpha)}(x)+F_k^{(\alpha)}(-x)]\,[F_l^{(\alpha)}(y)+F_l^{(\alpha)}(-y)],\\
D_{kl}^{(\alpha)}(x,y) &= dH_k(x)\,d[F_l^{(\alpha)}(y)+F_l^{(\alpha)}(-y)], \quad M_k^{(\alpha)}(x) = F_k^{(\alpha)}(x)+F_k^{(\alpha)}(-x),\\
U_{kl}(x,y) &= J_k'[H_k(x)]\,J_l'[H_l(y)],\\
E_{kl}^{(\alpha)}(x,y) &= F_{kl}^{(\alpha)}(x,y)+F_{kl}^{(\alpha)}(x,-y)-F_{kl}^{(\alpha)}(-x,y)-F_{kl}^{(\alpha)}(-x,-y),\\
G_{kl}^{(\alpha)}(x,y) &= F_{kl}^{(\alpha)}(x,y)-F_{kl}^{(\alpha)}(x,-y)-F_{kl}^{(\alpha)}(-x,y)+F_{kl}^{(\alpha)}(-x,-y),\\
L_{kl}^{(\alpha)}(x,y) &= F_{kl}^{(\alpha)}(x,y)+F_{kl}^{(\alpha)}(-x,y)-F_{kl}^{(\alpha)}(x,-y)-F_{kl}^{(\alpha)}(-x,-y),\\
N_{kl}^{(\alpha)}(x,y) &= F_{kl}^{(\alpha)}(x,y)+F_{kl}^{(\alpha)}(x,-y)-F_{kl}^{(\alpha)}(-x,y)-F_{kl}^{(\alpha)}(-x,-y),\\
P_{kl}^{(\alpha)}(x,y) &= F_{kl}^{(\alpha)}(x,y)-F_{kl}^{(\alpha)}(x,-y)-F_{kl}^{(\alpha)}(-x,y)+F_{kl}^{(\alpha)}(-x,-y).
\end{aligned}\right\} \tag{6.1}$$

Then

$$\begin{aligned}
\sigma_{N,kk,\alpha\alpha} = 2\bigg[&\iint_{0<x<y<\infty}[A_k^{(\alpha)}(x,y)+A_k^{(\alpha)}(-y,-x)]\,B_{kk}(x,y)\\
&+\int_{x=0}^{\infty}\int_{y=0}^{\infty}A_k^{(\alpha)}(-y,x)\,B_{kk}(x,y)\bigg]+2\rho_N^{(\alpha)}\sum_{r=1}^{c}\rho_N^{(r)}\\
&\times\bigg[\iint_{0<x<y<\infty}\{A_k^{(r)}(x,y)+A_k^{(r)}(-y,-x)\}\,U_{kk}(x,y)\,C_{kk}^{(\alpha)}(x,y)\\
&-\int_{x=0}^{\infty}\int_{y=0}^{\infty}A_k^{(r)}(-y,x)\,U_{kk}(x,y)\,C_{kk}^{(\alpha)}(x,y)\bigg]\\
&-2\rho_N^{(\alpha)}\bigg[\iint_{0<x<y<\infty}\{A_k^{(\alpha)}(x,y)+A_k^{(\alpha)}(-x,y)-A_k^{(\alpha)}(-y,x)\\
&-A_k^{(\alpha)}(-y,-x)\}\,U_{kk}(x,y)\,D_{kk}^{(\alpha)}(x,y)\\
&+\iint_{0<x<y<\infty}\{A_k^{(\alpha)}(y,x)+A_k^{(\alpha)}(-x,y)-A_k^{(\alpha)}(-y,x)\\
&-A_k^{(\alpha)}(-x,-y)\}\,U_{kk}(x,y)\,D_{kk}^{(\alpha)}(x,y)\bigg].
\end{aligned} \tag{6.2}$$

$$\begin{aligned}
\sigma_{N,kl,\alpha\alpha} = &\int_{x=0}^{\infty}\int_{y=0}^{\infty}\{E_{kl}^{(\alpha)}(x,y)-G_{kl}^{(\alpha)}(x,y)\}\,B_{kl}(x,y)\\
&-\rho_N^{(\alpha)}\int_{x=0}^{\infty}\int_{y=0}^{\infty}[\{L_{kl}^{(\alpha)}(x,y)-M_k^{(\alpha)}(x)\,H_k^{(\alpha)}(y)\}\,B_{kl}(x,y)\\
&-\{N_{kl}^{(\alpha)}(x,y)-M_l^{(\alpha)}(y)\,H_k^{(\alpha)}(x)\}\,U_{kl}(x,y)\,D_{kl}^{(\alpha)}(x,y)]-\rho_N^{(\alpha)}\sum_{r=1}^{c}\rho_N^{(r)}\\
&+\int_{x=0}^{\infty}\int_{y=0}^{\infty}\{P_{kl}^{(r)}(x,y)-H_k^{(r)}(x)\,H_l^{(r)}(y)\}\,U_{kl}(x,y)\,C_{kl}^{(\alpha)}(x,y).
\end{aligned} \tag{6.3}$$

$$\sigma_{N,kl,\alpha\beta} = -\sqrt{(\rho_N^{(\alpha)}\rho_N^{(\beta)})}\Bigg[\int_{x=0}^{\infty}\int_{y=0}^{\infty}\{N_{kl}^{(\alpha)}(x,y) - M_k^{(\alpha)}(x)\,H_l^{(\alpha)}(y)\}\,U_{kl}(x,y)$$

$$\times D_{lk}^{(\beta)}(x,y) + \{N_{kl}^{(\beta)}(x,y) - M_l^{(\beta)}(y)\,H_k^{(\beta)}(x)\}\,U_{kl}(x,y)\,D_{kl}^{(\alpha)}(x,y)$$

$$+\sum_{r=1}^{c}\rho_N^{(r)}\int_{x=0}^{\infty}\int_{y=0}^{\infty}\{P_{kl}^{(r)}(x,y) - H_{kl}^{(r)}(x)\,H_l^{(r)}(y)\}$$

$$\times U_{kl}(x,y)\,dM_k^{(\alpha)}(x)\,dM_l^{(\beta)}(y)\Bigg]. \quad (6.4)$$

$$\sigma_{N,kk,\alpha\beta} = -\sqrt{(\rho_N^{(\alpha)}\rho_N^{(\beta)})}\iint_{0<x<y<\infty}\{A_k^{(\alpha)}(x,y) - A_k^{(\alpha)}(-y,-x)$$

$$+A_k^{(\alpha)}(-y,x) - A_k^{(\alpha)}(-x,y)\}\,U_{kk}(x,y)\,D_{kk}^{(\beta)}(x,y)$$

$$-\iint_{0<y<x<\infty}\{A_k^{(\alpha)}(y,x) - A_k^{(\alpha)}(-x,-y) - A_k^{(\alpha)}(-y,x)$$

$$+A_k^{(\alpha)}(-x,y)\}\,U_{kk}(x,y)\,D_{kk}^{(\beta)}(x,y)$$

$$-\iint_{0<x<y<\infty}\{A_k^{(\beta)}(x,y) - A_k^{(\beta)}(-y,-x) - A_k^{(\beta)}(-x,y)$$

$$+A_k^{(\beta)}(-y,x)\}\,U_{kk}(x,y)\,D_{kk}^{(\alpha)}(x,y)$$

$$-\iint_{0<y<x<\infty}\{A_k^{(\beta)}(y,x) - A_k^{(\beta)}(-x,-y) - A_k^{(\beta)}(-x,y)$$

$$+A_k^{(\beta)}(-y,x)\}\,U_{kk}(x,y)\,D_{kk}^{(\alpha)}(x,y) - \sum_{r=1}^{c}\rho_N^{(r)}$$

$$\times\Bigg[\iint_{0<x<y<\infty}H_k^{(r)}(x)\,[1-H_k^{(r)}(y)]\,U_{kk}(x,y)\,dM_k^{(\alpha)}(x)\,dM_k^{(\beta)}(y)$$

$$+\iint_{0<y<x<\infty}H_k^{(r)}(y)\,[1-H_k^{(r)}(x)]\,U_{kk}(x,y)\,dM_k^{(\alpha)}(x)\,dM_k^{(\beta)}(y)\Bigg]. \quad (6.5)$$

ACKNOWLEDGEMENT

The authors would like to express their sincere thanks to Professor P. K. Sen for the careful examination of the paper, and for supplying a new proof of Theorem 3.2 which simplified considerably the author's original proof. They are also grateful to Professor H. T. David for pointing out an error in an original version of the paper.

REFERENCES

Anderson, T. W. (1958). *An Introduction to Multivariate Statistical Analysis*. New York: John Wiley and Sons.

Chatterjee, S. K. and Sen, P. K. (1964). Nonparametric tests for the bivariate two sample location problem. *Calcutta Statist. Ass. Bull.* **13**, 18–58.

David, H. A. (1963). *The Method of Paired Comparisons.* Charles Griffin and Co., Ltd.

Davidson, R. R. and Bradley, R. A. (1970). Multivariate paired comparisons: some large-sample results on estimation and tests of equality of preference. *Nonparametric. Techniques in Statistical Inference.* Cambridge University Press, pp. 111–25.

Durbin, J. (1951). Incomplete blocks in ranking experiments. *Brit. J. Psych.* **4**, 85–90.

Essen, C. G. (1945). Fourier analysis of distribution functions. *Acta. Math.* **77**, 1–125.

Ghosh, M. (1969). Asymptotically optimal nonparametric tests for miscellaneous problems of linear regression. Ph.D. Thesis, University of North Carolina.

Hájek, Jaroslav (1961). Some extensions of the Wald–Wolfowitz–Noether Theorem. *Ann. Math. Statist.* **32**, 506–25.

Noether, G. E. (1960). Remarks about a paired comparison model. *Psychometrika* **25**, 357–67.

Puri, M. L. (1964). Asymptotic efficiency of a class of *c*-sample tests. *Ann. Math. Statist.* **35**, 102–21.

Puri, M. L. and Sen, P. K. (1966). On a class of multivariate-multisample rank order tests. *Sankhyā* **28**, 353–76.

Puri, M. L. and Sen, P. K. (1960). On the asymptotic normality of one sample rank order test statistics. *Teoria veroyatnostey iee Primenenyia* **14**, 167–72.

Puri, M. L. and Sen, P. K. (1967). On some optimum nonparametric procedures in two-way layouts. *J. Am. statist. Ass.* **63**, 1214–29.

Puri, M. L. and Sen, P. K. (1969). The asymptotic theory of rank order tests for experiments involving paired comparisons. *Ann. Inst. Statist. Math.* **21**, 163–73.

Rao, C. R. (1965). *Linear Statistical Inference and its Applications.* New York: John Wiley and Sons.

Scheffe, H. (1952). An analysis of variance for paired comparisons. *J. Am. Statist. Ass.* **47**, 381–400.

Sen, P. K. and David, H. A. (1968). Paired comparisons for paired characteristics. *Ann. Math. Statist.* **39**, 200–8.

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**, 1216–28.

Shane, H. D. (1968). Multivariate multisample analogues of some one sample tests. Ph.D. Thesis, New York University.

Shane, H. D. and Puri, M. L. (1969). Rank Order Tests for Multivariate Paired Comparisons. *Ann. Math. Statist.* **40**, 2101–17.

Sverdrup, E. (1952). The limit distribution of a continuous function of several random variables. *Skand. Aktuariedskrift* **35**, 1–10.

Wald, A. and Wolfowitz, J. (1944). Statistical tests based on permutations of the observations. *Ann. Math. Statist.* **15**, 358–72.

DISCUSSION ON PURI AND SHANE'S PAPER

H. T. DAVID

This paper, as much as any of the Symposium, shows the importance of nonparametric method for non-subjective statistics. Nonparametric methods require a minimum of probabilistic modelling—only the discrete uniform distribution. In this they seem to meet 'data analyzers' halfway and others, outside the profession, wary of detailed probabilistic assumptions. Yet, as is so well demonstrated in this paper, nonparametric methods manage to compete successfully in the world of parametric hypotheses; there is in addition the attractive deductive link of permutation tests to experimental randomization, pointed out by Fisher and since stressed by Kempthorne. In this connection note that Puri and Shane's group G_N arises from the following experimental randomization $\mathscr{R}$: Randomly assign the N treatment pairs to the N blocks (randomization $\mathscr{R}_1$), then, independently, the individual treatments of a particular pair to the two plots within a block (randomization $\mathscr{R}_2$). Let H_N indicate permutation of paired vectors. Then the implied joint null permutation distribution of the $2N$ vector observations (joint distribution of $2pN$ scalars in all) is uniform over the relevant orbit of $H_N G_N^{(2)} \equiv G_N^*$. The induced joint null permutation distribution of the N vector differences $\mathbf{Z}\alpha, l$ (joint distributions of pN scalars in all) is then uniform over the orbit of G_N containing the observed Z. Thus $\mathscr{R}$ leads naturally to an intra-block permutation analysis, say based on the permutation distribution of (4.39) under G_N. This suggests several comments. To begin with, suppose that the individual vector observations $\mathbf{X}\alpha, l$ and $\mathbf{Y}\alpha, l$ are indeed available, in addition to the differences $\mathbf{Z}\alpha, l = \mathbf{Y}\alpha, l - \mathbf{X}\alpha, l$. Let now T_i be the average of *all* vector observations from all blocks featuring treatment i. Then G_N^* leads equally to an inter-block permutation analysis, say based (at least when the N_{ij} all are equal) on the permutation distribution under G_N^* (equivalently, $G_N^{(2)}$) of a suitable one-way MANOVA statistic involving the t vectors T_i. It is not clear how best to combine the intra- and inter-block analyses; however, replacing an intra-block statistic such as (4.39) by its conditional expectation given $g_N^{(1)}$ at least insures the conditional independence of the two analyses. A further point is that, for blocks of size $k > 2$, G_N^*, and in particular H_N, remains of course appropriate under $\mathscr{R}$. The group applicable to paired treatment differences within blocks is then pinpointed as the cyclic group of order $\binom{k}{2}$ multiplied by a modification of $G_N^{(1)}$ whose elements reverse the order of the $\binom{k}{2}$ difference vectors, in addition to the change of sign. A final point is that some experimenters, faced with the prospect of constant N_{ij} (say $N_{ij} = m > 1$) might not opt for $\mathscr{R}$ at all, but rather for separate randomization in each of m replicates of c blocks. I imagine that this or other alternatives to $\mathscr{R}$ (and hence to G_N^*) would be gauged by parametric models with random rep effects; these in addition to the random block effects that would be introduced in any power computations for the inter-block analysis.