ON THE THEORY OF RANK-ORDER TESTS AND ESTIMATES FOR LOCATION IN THE MULTIVARIATE ONE SAMPLE PROBLEM

by

Pranab K. Sen and Madan L. Puri

University of North Carolina

Institute of Statistics Mimeo Series No. 488

August 1966

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

DEPARTMENT OF BIOSTATISTICS
UNIVERSITY OF NORTH CAROLINA
Chapel Hill, N. C.
ON THE THEORY OF RANK-ORDER TESTS AND ESTIMATES FOR LOCATION
IN THE MULTIVARIATE ONE SAMPLE PROBLEM*

by
Pranab K. Sen¹ and Madan L. Puri

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

0. Summary. In the multivariate one-sample location problem, the theory of permutation distribution under sign-invariant transformations is extended to a class of rank-order statistics, and is utilized in the formulation of a genuinely distribution free class of rank-order tests for location. Asymptotic properties of these permutation rank order tests are studied and certain stochastic equivalence relations with a similar class of multivariate extensions of one-sample Chernoff-Savage-Hajék type tests are derived. The power properties of these tests are studied. Finally, Hodges-Lehmann (1963) technique of estimating shift parameters through rank-order tests is extended to the multivariate case which includes among other results, the results obtained by Bickel (1964 and 1965).


¹On leave of absence from Calcutta University.
1. Introduction. Let $\mathbf{x}_\alpha = (x_\alpha^{(1)}, \ldots, x_\alpha^{(p)})$ be a $p$-variate random variable having a continuous $p$-variate cdf (cumulative distribution function) $F(\mathbf{x}, \mathbf{\theta})$ where $\mathbf{x} = (x^{(1)}, \ldots, x^{(p)})$, $\mathbf{\theta} = (\theta_1, \ldots, \theta_p)$, and where $p$ is a positive integer which in the sequel will be assumed to be greater than unity.

Let then $x_1, \ldots, x_N$ be $N$ i.i.d.r.v.'s (independent and identically distributed random variables) distributed according to the common cdf $F(\mathbf{x}, \mathbf{\theta})$. Let $\Omega$ be a set of all continuous $p$-variate cdf's, and it is assumed that $F \in \Omega$.

Let now $\omega$ denote the set of all continuous $p$-variate cdf's which are symmetric about some known origin which we may without any loss of generality take to be $\mathbf{x} = \mathbf{0}$, the symmetry being defined in the following manner.

Let $\mathbf{\mathcal{L}}$ be any set of points $P = P(x_1, \ldots, x_p)$ in the $p$-dimensional Euclidean space $\mathbb{E}_p$. Any point $P$ belonging to $\mathbf{\mathcal{L}}$ is denoted by $P \in \mathbf{\mathcal{L}}$, and the image of the point $P$ is denoted by $P^* = P(x_1^*, \ldots, x_p^*)$ where $x_i^* = -x_i$, $i = 1, \ldots, p$.

The image set $\mathbf{\mathcal{L}}^*$ of $\mathbf{\mathcal{L}}$ is denoted by

\[(1.1) \quad \mathbf{\mathcal{L}}^* = \{x : P(x^*) \in \mathbf{\mathcal{L}}\}.\]

Then, if for any $\mathbf{\mathcal{L}} \subset \mathbb{E}_p$

\[(1.2) \quad P[\tilde{x} \in \mathbf{\mathcal{L}} | F] = P[\tilde{x} \in \mathbf{\mathcal{L}}^* | F] \]

we say that $F$ is symmetric about $\mathbf{0}$.

If, in addition, $F$ is absolutely continuous having a density function $f$, the symmetry of $F$ may also be defined
by the invariance of the density function under simultaneous changes of signs of all the coordinate variates. For the convenience of terminology, (1.2) may also be termed as the sign-invariance. Excepting with the later section dealing with the asymptotic normality and other properties of a proposed class of estimates and tests, we will drop the assumption of absolute continuity of $F$.

Thus, for the time being let us frame the null hypothesis $H_0$ of sign-invariance as

$$H_0 : F \in \omega \subset \Omega.$$  

The two particular classes of alternative in which we may usually be interested are

$$H_1 : F(\xi) \text{ is symmetric about some } \xi \neq \xi_0.$$  

This will be the multivariate extension of the well known one sample location problem.

$$H_2 : F(\xi) \text{ is asymmetric about } \xi = \xi_0 \text{ though it may have the location vector } \xi_0.$$  

This is the multivariate extension of the univariate symmetry problem.

In the univariate case, there is a third problem whose multivariate extension would be the estimation of location vector $\xi_0$, assuming $F(\xi)$ to be symmetric about $\xi = \xi_0$.

Tests of the type mentioned above in the univariate case are due to Wilcoxon (1949), Govindarajulu (1960) and the
classical sign test. Estimates in the univariate case are due to Hodges and Lehmann (1963) and Sen (1963). However, in the multivariate case there are only a few suitable tests and estimates.

In the bivariate case, Hodges (1955) has considered an association invariant sign test on which some further work has been done by Klotz (1964), and Joffe and Klotz (1962). The test is strictly distribution-free but neither the null nor the non-null distribution (for any suitable sequence of alternative hypotheses) of the test-statistic appears to be simple, and so there is little scope for studying the Pitman-efficiency of such a test. The only study of the efficiency of this test is made by Klotz (1962) and that is the Bahadur Efficiency. Further, being a sign test it is anticipated that as in the univariate case, for nearly normal cdf's, the efficiency of this test may be appreciably low. Some sign-tests are due to Blumen (1958) and Bennet (1962). Bluman's test is also strictly distribution-free but it is subject to the same drawback as Hodges' test. Bennet's test is distribution-free only asymptotically. Some further asymptotically distribution-free tests are due to Bickel. His tests are based on generalizations of the univariate sign-statistic and Wilcoxon's signed rank statistic and are distribution free only asymptotically. However, his study of the asymptotic power properties of these two tests provides some useful information about the performance characteristics of them. Very recently, Chatterjee (1966) has considered a
very interesting permutationally distribution-free bivariate sign test which asymptotically agrees with Bickel's and Bennet's sign-tests. In this investigation this permutation argument will be extended not only to the p-variate case for any \( p \geq 2 \) but also to a much wider class of rank-order tests which includes Bickel's tests as special cases, and which are strictly distribution free under permutation arguments. This extension also covers the normal scores and the other type of multivariate score tests on which practically there is no work in the literature. In the problem of estimation too, Bickel (1964) has extended Hodges and Lehmann (1963) type of estimates to the multivariate case using only the median and rank sum procedures. In this investigation a much wider class of rank order statistics will be used, and this will be a further generalization of Bickel's work.

2. The Basic Permutational Argument.

Let us denote the sample point by

\[
Z_N = (x_1, \ldots, x_N)
\]

and the sample space by \( \mathcal{Z} \). Then under the null hypothesis to be tested the joint distribution of \( Z_N \) remains invariant under the following finite group \( G_N \) of transformations \( g_n \) given by

\[
g_nZ_N = ((-1)^{j_1}x_1, \ldots, (-1)^{j_N}x_N), \quad j_1 = 0,1; \quad i = 1, \ldots, N,
\]
where \((-1)^{X_{\alpha}} = (-X_{\alpha}^{(1)}, \ldots, -X_{\alpha}^{(p)})\). Thus there are \(2^N\) such possible \(\{z_N\}\) under which the distribution of \(Z_N\) remains invariant. Hence, given \(Z_N\), we can generate a set \(S_N\) of \(2^N\) points \(Z_N\) in \(Z_N\), such that all these \(2^N\) points will have the same probability. Hence, conditioned on the generated set \(S_N\), there are \(2^N = M\) points and the permutational probability measure associated with each point is the same, viz., \(2^{-N}\), when \(H_0\) given by (1.3) is true.

Let us now denote by \(\phi(Z_N)\) a test function which with each \(Z_N\) associates a probability of rejecting \(H_0\). Now if we use the known permutational probability measure of \(Z_N\) on \(S_N\), then it readily follows from the well known result of Lehmann and Stein (1949) that a test function \(\phi(Z_N)\) based on this permutational probability measure will possess the structure \(S(\epsilon)\) (cf. Lehmann and Stein (1949)), and hence will be a strictly distribution-free test. This characterizes the existence of strictly distribution-free test of size \(\epsilon\), \(0 < \epsilon < 1\), for the hypothesis \(H_0\) given by (1.3).

Now in actual practice \(\phi(Z_N)\) has to be constructed with special attention to the class of alternatives in mind, and in most of the multivariate problems, \(\phi(Z_N)\) depends on \(Z_N\) through a single-valued test statistic \(T_N = T(Z_N)\) in order to have a comprehensive test. Due to these reasons we shall consider the following procedure.

Let us rank the \(N\) values of \(|X_{\alpha}^{(1)}|, \alpha = 1, \ldots, N\), in order of magnitude, and let \(R_{\alpha}^{(1)}\) denote the rank of \(|X_{\alpha}^{(1)}|\) in this set.
Then \((R_1^{(1)}, \ldots, R_N^{(1)})\) is a permutation of the numbers \(1, 2, \ldots, N\). This procedure is adopted separately for each \(i = 1, \ldots, p\).

Thus, conditioned on a given \(\mathcal{Z}_N\), an observation \(\mathcal{X}_\alpha\) having \(p\) values is converted into

\[
R_\alpha = (R^{(1)}_\alpha, \ldots, R^{(p)}_\alpha), \quad \alpha = 1, \ldots, N.
\]  

(2.3)

Let now \(\{E^{(1)}_{N,\alpha}; \alpha = 1, \ldots, N\}\) be a sequence of \(N\) real numbers; the sequence may of course vary over \(i = 1, \ldots, p\). The value of \(E^{(1)}_{N,\alpha}\) corresponding to the value \(\alpha = R^{(1)}_\alpha\) is denoted by \(E^{(1)}_{N,R^{(1)}_\alpha}; \alpha = 1, \ldots, N; i = 1, \ldots, p\).

Thus, if \(\alpha\{E^{(1)}_{N,\alpha}; \alpha = 1, \ldots, N\}; i = 1, \ldots, p, \) are selected corresponding to \(R_\alpha\) in (2.3) we have a \(p\)-tuple of real values

\[
E^{(1)}_{N,\alpha} = (E^{(1)}_{N,R^{(1)}_\alpha}, \ldots, E^{(p)}_{N,R^{(p)}_\alpha}); \quad \alpha = 1, \ldots, N,
\]  

(2.4)

and, we denote by \(\mathcal{E}_N\) the following \(p\) by \(N\) matrix

\[
\mathcal{E}_N = \begin{pmatrix}
E^{(1)}_{N,1} \\
\vdots \\
E^{(p)}_{N,N}
\end{pmatrix}
= \begin{pmatrix}
E^{(1)}_{N,R^{(1)}_1} & \ldots & E^{(1)}_{N,R^{(1)}_N} \\
\vdots & \ddots & \vdots \\
E^{(p)}_{N,R^{(p)}_1} & \ldots & E^{(p)}_{N,R^{(p)}_N}
\end{pmatrix}
\]  

(2.5)

\(E_N\) is really a stochastic matrix, its \(i\)-th row being a random permutation of \((E^{(1)}_{N,1}, \ldots, E^{(1)}_{N,N})\) for \(i = 1, \ldots, p\).

Thus there are \((N!)^p\) possible values of \(\mathcal{E}_N\). Two realizations \(\mathcal{E}_N\) and \(\mathcal{E}_N'\) are said to be equivalent when it is possible to arrive
at $E_N$ from $E_N^1$ by only a finite number of inversions of columns of $E_N^1$. Thus, we may always consider the first row of $E_N$ in natural order. There will be thus $(N!)^{p-1}$ non-equivalent realizations of $E_N$. We denote this set of $(N!)^{p-1}$ realizations by $E_N^C$. $E_N$ will be termed a **collection rank order matrix** and $E_N^C$ will be termed as the **space of collection matrices**.

Let us now consider a sequence of $N$ diagonal matrices $[C_\alpha, \alpha=1,\ldots,N]$ each of order $p$ by $p$, where

\begin{equation}
C_\alpha = \text{diag} \left( C_\alpha^{(1)}, \ldots, C_\alpha^{(p)} \right)
\end{equation}

(2.6)

\begin{equation}
C_\alpha^{(1)} = (-1)^{\alpha_1}; \quad \alpha_1 = 0,1; \quad \alpha = 1,\ldots,N \quad ;
\end{equation}

(2.7)

\begin{equation}
\alpha_1 = 0 \text{ if } X_\alpha^{(1)} > 0, \quad \text{and} \quad \alpha_1 = 1 \text{ if } X_\alpha^{(1)} < 0 .
\end{equation}

(2.8)

Thus, each $C_\alpha$ is a stochastic diagonal matrix, which can have $2^p$ realizations. Let us now define a $p$ by $pN$ matrix

\begin{equation}
C_{(N)} = (C_1, \ldots, C_N)
\end{equation}

(2.9)

which we term the **collection sign matrix**. This is again a stochastic matrix which can have $2^{pN}$ realizations. The set of $2^{pN}$ possible realizations of $C_{(N)}$ is denoted by $E_N^C$, and is termed the **space of sign-matrices**, while $C_{(N)}$ will be termed as the **collection sign-matrix**. This is again a stochastic matrix which can have $2^{pN}$ possible realizations. The set of $2^{pN}$ possible realizations of $C_{(N)}$ is denoted by $E_{(N)}$ and will be called the **space of sign-matrices**, while $C_{(N)}$ will be termed the **collection sign-matrix**.
Let us now consider the pair of stochastic variables $(E_n, \xi(N))$ which lie in the product space $(\xi_N, \Gamma_N)$. The probability distribution of $(E_n, \xi(N))$ on this product space (defined on an additive class of (product) $\sigma$-fields of subsets of $(\xi_N, \Gamma_N)$) depends on the unknown $F(\xi)$, even when $F=0$, that is, even when $H_0$ is true.

It follows from (2.2) that if we consider the finite group $G_N$ of transformations $g_n$ (containing $2^N$ elements), then the probability distribution of $\xi(N)$ or $\Gamma_N$ remains invariant under the transformations in $G_N$, that is, if

\begin{equation}
E_{N,N}(N) = \left((-1)^{j_1}C_1, \ldots, (-1)^{j_N}C_N\right), \quad \xi_N \in G_N,
\end{equation}

where $(-1)^{j_\alpha}C_\alpha = \text{diag}((-1)^{j_\alpha}C_\alpha(1), \ldots, (-1)^{j_\alpha}C_\alpha(p))$, $\alpha=1, \ldots, N$
(with $j_\alpha = 0, 1$ for $\alpha = 1, \ldots, N$), then the probability distribution of $\xi(N)$ remains invariant under any such transformations $g_N$. Again by definition of $R^{(i)}_{\alpha}$, $i=1, \ldots, p$, $\alpha = 1, \ldots, N$, and of $E_N$ (in (2.3), (2.4) and (2.5)) we see that $E_N$ remains invariant under any transformation $g_N \in G_N$ or the original vector $Z_N \in Z_N$. Hence if we consider the stochastic matrix

\begin{equation}
E_{N}^{*} = \left(E_{N1}(1), E_{N1}(\xi(1)), \ldots, E_{NN}(1), E_{NN}(\xi(N))\right)
\end{equation}

where $E_N^{(i)} = (-1)^{j_1}; j_i = 0, 1; i=1, \ldots, N$; then it readily follows that under $H_0$ in (1.4), the joint distribution of $E_N^{*}$ conditioned on the given $Z_N$ (which implies that given $(E_N, \xi(N))$ remains invariant under $g_N \in G_N$. Thus, if we consider the set of $2^N$ values
(2.12) \[ \{E_{N}^{*} : e_{N} \in G_{N}\} \]

we get a set of \(2^{N}\) permutationally equivalent (probabilistic) points, and the permutational probability measure on this set is denoted by \(\rho_{N}\), which is a conditional measure given \(Z_{N}\), that is given \((E_{N}, \sim_{N})\).

We shall now develop a strictly distribution free test procedure under this permutational probability measure.

Let us define

(2.13) \[ T_{N}^{(j)} = \sum_{\alpha=1}^{N} E_{NR_{\alpha}}^{(j)} C_{\alpha}^{(j)} , \quad j = 1, \ldots, p ; \]

(2.14) \[ T_{N} = (T_{N}^{(1)}, \ldots, T_{N}^{(p)}) . \]

Thus, \(T_{N}^{(j)}\) is the difference of the sum of the rank functions \([E_{N\alpha}^{(j)}]\) for which \(X_{\alpha}^{(j)} > 0\) and the sum over the remaining set. It follows easily that

(2.15) \[ E(T_{N}^{*} | \rho_{N}) = \varnothing = (0, \ldots, 0) \]

for any given \((E_{N}, \sim_{N})\). Let us now define for each combination \(j \neq k (= 1, \ldots, p)\) four subsets \(S_{jk}^{(1)}, \ldots, S_{jk}^{(4)}\), where

\[
S_{jk}^{(1)} = \{E_{N\alpha}^{(j)} : C_{\alpha}^{(j)} < 0, C_{\alpha}^{(k)} < 0\}
\]

(2.16) \[
S_{jk}^{(2)} = \{E_{N\alpha}^{(j)} : C_{\alpha}^{(j)} < 0, C_{\alpha}^{(k)} > 0\}
\]

\[
S_{jk}^{(3)} = \{E_{N\alpha}^{(j)} : C_{\alpha}^{(j)} > 0, C_{\alpha}^{(k)} < 0\}
\]

\[
S_{jk}^{(4)} = \{E_{N\alpha}^{(j)} : C_{\alpha}^{(j)} > 0, C_{\alpha}^{(k)} > 0\} .
\]
Of course, if \( j = k \), then \( s_{jk}^{(2)} \) and \( s_{jk}^{(3)} \) will be the null sets. Since under the transformation \( g_N \) on \( \mathcal{C}_\alpha \), both \( c_{\alpha}^{(j)} \) and \( c_{\alpha}^{(k)} \) change signs simultaneously, the product \( E(j) / E(k) \) will have a positive sign for \( S_{jk}^{(1)} \) and \( S_{jk}^{(4)} \) and a negative sign for \( S_{jk}^{(2)} \) and \( S_{jk}^{(3)} \), and the same will remain invariant under \( g_N \in \mathcal{G}_N \). Hence it can easily be shown that if we define

\[
(2.17) \quad v_{N, jk} = \frac{1}{N} \left( \sum_{S_{jk}^{(1)}} E(j) / E(k) + \sum_{S_{jk}^{(4)}} E(j) / E(k) - \sum_{S_{jk}^{(2)}} E(j) / E(k) - \sum_{S_{jk}^{(3)}} E(j) / E(k) \right), \quad j \neq k = 1, \ldots, p;
\]

\[
(2.18) \quad v_{N, jj} = \frac{1}{N} \sum_{\alpha=1}^N \left[ E_{N\alpha} \right]^2, \quad j = 1, \ldots, p;
\]

then

\[
(2.19) \quad \text{cov} \left\{ (T_N^{(j)}, T_N^{(k)}) | \mathcal{P}_N \right\} = N \cdot v_{N, jk}, \quad j, k = 1, \ldots, p.
\]

Let us now denote by

\[
(2.20) \quad \mathcal{V}_N = (v_{N, jk})_{j,k=1,\ldots,p}
\]

and assume (for the time being) that \( \mathcal{V}_N \) is positive definite. (\( \mathcal{V}_N \) being a covariance matrix will be positive semidefinite at least. If \( \mathcal{V}_N \) is singular, then we can work with the highest order non-singular minor of \( \mathcal{V}_N \) and work with the
corresponding variates). Later we will show that under certain conditions $Y_N$ will be positive definite. Thus, we consider the following positive semi-definite quadratic form

\begin{equation}
S_N = \frac{1}{N} \{ T_N V_N^{-1} T_N' \}
\end{equation}

where $V_N^{-1}$ is the inverse matrix of $V_N$.

If the null hypothesis is true, the center of gravity of $T_N$ will be zero, and conditionally on given $Z_N$, that is, on given $(E_N, S_N)$, there will be $2^N$ permutationally equally likely (not necessarily all distinct) values of $S_N$. On the other hand, if $H_0$ is not true, it can be shown, as in the univariate case that $\frac{1}{N} T_N^{(j)}$ converges in probability to some non-zero quantity for at least one $j = 1, \ldots, p$ and hence $S_N/N$ will converge to a non-null quantity. Thus, $S_N$ will be stochastically large. This suggests the use of the right hand side tail of the permutation distribution of $S_N$ as the suitable critical region. Thus the critical function

\begin{equation}
\phi(z_N) = \begin{cases} 
1 & \text{if } S_N > S_N, \varepsilon(z_N) \\
a_N, \varepsilon & \text{if } S_N = S_N, \varepsilon(z_N) \\
0 & \text{if } S_N < S_N, \varepsilon(z_N) 
\end{cases}
\end{equation}

where $S_N, \varepsilon(z_N)$ and $a_N, \varepsilon$ are chosen such that

\begin{equation}
E[\phi(z_N) | \mathcal{P}_N] = \varepsilon, \quad 0 < \varepsilon < 1,
\end{equation}

provides a strictly size $\varepsilon$ similar region.

For small samples, $S_N, \varepsilon(z_N)$ and $a_N, \varepsilon$ are to be determined from the actual permutational cdf of $S_N$, while for large
samples, we have proved in the next section that

\[(2.23) \quad S_{N,\epsilon}(z_N) \text{ and } a_{N,\epsilon} \text{ converge in probability to } \chi^2_{p,\epsilon} \]

and zero respectively as \( N \to \infty \), where \( \chi^2_{p,\epsilon} \) is the \( 100(1-\epsilon)\% \) point of the \( \chi^2 \) distribution with \( p \) degrees of freedom.

3. Properties of \( V_N \) and Asymptotic Permutation Distribution of \( S_N \).

Let us denote the marginal cdf of \( x^{(j)}_\alpha \) and of \((x^{(j)}_\alpha, x^{(k)}_\alpha)\)
by \( F_j(x;\theta_j) \) and \( F_{j,k}(x,y;\theta_j,\theta_k) \) respectively, and let

\[(3.1) \quad H_j(x;\theta_j) = F_j(x;\theta_j) - F_j(-x;\theta_j), \quad x \geq 0 \]

\[(3.2) \quad H_{j,k}(x,y;\theta_j,\theta_k) = F_j(x,y;\theta_j,\theta_k) - F_{j,k}(x,-y;\theta_j,\theta_k) - F_{j,k}(-x,y;\theta_j,\theta_k) + F_{j,k}(-x,-y;\theta_j,\theta_k), \quad 0 \leq x,y \leq \infty \]

Set

\[(3.3) \quad F_{N,j}(x) = (\text{number of } x^{(j)}_\alpha \leq x)/N \]

\[(3.4) \quad F_{N,j,k}(x,y) = (\text{number of } (x^{(j)}_\alpha, x^{(k)}_\alpha) \leq (x,y))/N \]

\[(3.5) \quad H_{N,j}(x) = F_{N,j}(x) - F_{N,j}(-x) \]

\[(3.6) \quad H_{N,j,k}(x,y) = F_{N,j,k}(x,y) - F_{N,j,k}(x,-y) - F_{N,j,k}(-x,y) + F_{N,j,k}(-x,-y) . \]
Finally, as in Chernoff and Savage, we write

$$E_{N\alpha}^{(j)} = J_N, j(\alpha/(N+1)) = J_N, j\left(\frac{N}{N+1} H_N, j(x)\right),$$

$$\alpha = 1, \ldots, N, \quad j = 1, \ldots, p,$$

where $J_N, j$ though defined only at $1/(N+1), \ldots, N/(N+1)$ may have its domain of definition extended to $(0,1)$ by letting it have constant value over $(\alpha/(N+1), (\alpha+1)/(N+1))$. Furthermore, we make the following assumptions:

**Assumption 1.** $\lim_{N \to \infty} J_N, j(u) = J_j(u)$ exists for $0 < u < 1$ and is not a constant.

**Assumption 2.** $\int_{-\infty}^{\infty} \left[ J_N, j\left(\frac{N}{N+1} H_N, j\right) - J_j\left(\frac{N}{N+1} H_j\right) \right] dF_N, j(x) = o_p \left( N^{-1/2} \right).$

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

$$|J_j^1(u)| = \left| \frac{d(1)J_j(u)}{du(1)} \right| \leq K[u(1-u)]^{5-1-\frac{1}{2}}$$

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

**Assumption 4.**

$$\int_{0}^{\infty} \int_{0}^{\infty} \left[ J_N, j\left(\frac{N}{N+1} H_N, j\right) J_N, k\left(\frac{N}{N+1} H_N, k\right) - J_j\left(\frac{N}{N+1} H_j\right) J_k\left(\frac{N}{N+1} H_k\right) \right] dH_N, j, k(x,y) = o_p(1).$$

Let us now define

$$\nu_{jk} = \int_{00}^{\infty} J_j[H_j(x; \theta_j)] J_k[H_k(y; \theta_k)] dH_j, k(x,y; \theta_j, \theta_k),$$

$$j, k = 1, \ldots, p.$$
We may remark that the assumptions 1 to 3 are needed to prove the joint asymptotic normality of the permutation distribution of $\tau_N$ given by (2.14). Assumption 4 is required to establish the stochastic convergence of the permutation covariance matrix $V_N$ to $\gamma$ which we shall require in the sequel.

**Theorem 3.1.** Under the assumptions 1 to 4, $V_N \to \gamma$ in probability as $N \to \infty$, where $\gamma = (\gamma_{jk})$ is given by

$$
\gamma_{jk} = \int_0^\infty \int_0^\infty \int J_j[H_j(x;\theta_j)] J_k[H_k(y;\theta_k)] dH_j,k(x,y;\theta_j,\theta_k).
$$

**Proof:** From (2.17), using (3.6) and (3.7) and the assumption 2, we find that

$$
V_{N,j,k} = \int_0^\infty \int_0^\infty \int J_j[N+1 \ H_j(x)] J_k[N+1 \ H_k(y)] dH_{N,j,k}(x,y)
$$

$$
= \int_0^\infty \int_0^\infty \int J_j[N+1 \ H_j(x)] J_k[N+1 \ H_k(y)] dH_{N,j,k}(x,y) + O_p(1).
$$

Now, writing

$$
J_j[N+1 \ H_j(x)] = J_j(H_j) + (H_j - H_j)J_j'(H_j) - \frac{H_j}{N+1} J_j'(H_j)
$$

$$
+ \left[J_j[N+1 \ H_j(x)] - J_j(H_j) - (N+1) H_j - J_j'(H_j)\right]
$$
\[ J_k \left[ \frac{N}{N+1} H_N, k(y) \right] = J_k(H_k) + (H_N, k - H_k) J_k'(H_k) \]

\[ - \frac{H_N, k}{N+1} J_k'(H_k) + \{ J_k(\frac{N}{N+1} H_N, k), k(H_k) \} \]

\[ - \left( \frac{N}{N+1} H_N, k - H_k \right) J_k'(H_k) \]

\[ dH_{N,j,k}(x,y) = d(H_N, k(x,y)) + H_{N,j,k}(x,y) - H_{j,k}(x,y) \]

and proceeding exactly as in [14] and [26], we find, omitting the details of computations that,

\[ \nu_{N,j,k} = \int J_j[H_j(x; \Theta_j)] J_k[H_k(y; \Theta_k)] dH_{j,k}(x,y; \Theta_j, \Theta_k) + o_p(1) \]

**Corollary 3.1:** If

1. \[ F_j(x; \Theta_j) = F_j(x - \frac{\Theta_j}{\sqrt{N}}), \quad F_{j,k}(x,y; \Theta_j, \Theta_k) = F_{j,k}(x - \frac{\Theta_j}{\sqrt{N}}, y - \frac{\Theta_k}{\sqrt{N}}) \]

where \( F_j \) and \( F_{j,k} \) are symmetric about zero.

2. The assumptions of Theorem 3.1 are satisfied, then \( \nu_{N} \rightarrow \nu^* \) in probability as \( N \rightarrow \infty \), where \( \nu^* = (\nu^*_{j,k}) \)

is given by

\[ (3.16) \quad \nu^*_{j,k} = \int \int J_j[2F_j(x)-1] J_k[2F_k(y)-1] \ dF_{j,k}(x,y) \]

**Theorem 3.2:** Under the assumptions 1 to 4, \( N^{-1/2} T_N \) has asymptotically a \( p \)-variate normal distribution with mean vector zero and covariance matrix \( \Gamma = (\gamma_{j,k}) \) where \( (\gamma_{j,k}) \)
is given by (3.10).
Proof: To prove this theorem, it suffices to show that for any arbitrary constant vector \( \tilde{\alpha} = (\tilde{\alpha}_1, \ldots, \tilde{\alpha}_p) \), the distribution of \( N^{-1/2} \tilde{\alpha} T_N' \) is asymptotically normal under the permutational probability measure \( P_N \). From (2.13) we notice that

\[
(3.17) \quad \tilde{\alpha} \cdot T_N' = \sum_{\alpha=1}^{N} \sum_{j=1}^{P} \tilde{\alpha}_j E_{NR_{\alpha}}^{(j)} C_{\alpha}^{(j)} = \sum_{\alpha=1}^{N} a_{\alpha\alpha}(E_N, C(N), \tilde{\alpha})
\]

where by definition \( a_{\alpha\alpha}(E_N, C(N), \tilde{\alpha}) \) depends upon \( E_N, C_N \) and \( \tilde{\alpha} \).

Let us now consider the permutation distribution generated by \( 2^N \) transformations \( g_N \) in \( G_N \) defined in (2.2). It then follows from earlier discussion that \( E_N \) and \( \tilde{\alpha} \) remain invariant under \( G_N \) or \( Z_N \), while \( C(N) \) may have \( 2^N \) permutationally equally probable values:

\[
(3.18) \quad C_{\alpha} = \text{diag} \left( C_{\alpha}^{(1)}, \ldots, C_{\alpha}^{(p)} \right)
\]

which can only assume two values

\[
(3.19) \quad \text{diag} \left( C_{\alpha}^{(1)}(-1)^{i_{\alpha}}, \ldots, C_{\alpha}^{(p)}(-1)^{i_{\alpha}} \right), \quad i_{\alpha} = 0, 1,
\]

with equal probability \( \frac{1}{2} \) for \( \alpha = 1, \ldots, N \). Thus, conditionally on the given \( \widetilde{Z}_N \), that is, on the given \( (E_N, C_N) \),

\[ 5 \cdot E_{NR_{\alpha}}^{(j)}, \quad j = 1, \ldots, p; \quad \alpha = 1, \ldots, N \]

remain fixed, while \( \widetilde{C}(\alpha) \) \( C_{\alpha}^{(i_{\alpha})} \) can only change sign simultaneously (with in each \( \alpha \)) for \( \alpha = 1, \ldots, N \). Thus, for the \( 2^N \) \( g_N \in G_N \), we have the corresponding \( 2^N \) values of \( \tilde{\alpha} \cdot T_N' \) which may be written as
\begin{equation}
\sum_{\alpha = 1}^{N} \left| a_{N,\alpha}(E_{N}^{\infty}, C(N), \xi) \right| d_{N,\alpha}
\end{equation}

where \(\{d_{N,\alpha}\}\) are mutually independent (under the permutational probability measure) and \(d_{N,\alpha} = +1\) or \(-1\) with equal probability \(1/2\); and where conditionally on the given \(Z_{N}\), that is, on the given \((E_{N}^{\infty}, C(N))\) and \(\xi\), \(\left| a_{N,\alpha}(E_{N}^{\infty}, C(N), \xi) \right|, \alpha = 1, \ldots, N\)
are all fixed. It is also easily seen that the permutational average of (3.20) is zero, and the variance \(V_{N} \xi \xi'\).

So if we define \(W_{N,\alpha} = \left| a_{N,\alpha}(E_{N}^{\infty}, C(N), \xi) \right| d_{N,\alpha}\), then under \(\mathcal{O}_{N}\), \(\{W_{N,\alpha}\}\) are independent random variables with means zero and \(W_{N,\alpha}\) can assume only two values \(\pm |W_{N,\alpha}|\) each with probability \(1/2\). We will now apply the central limit theorem to the sequence \(\{W_{N,\alpha} : \alpha = 1, \ldots, N\}\) under the Liapounoff's condition, viz.,

\begin{equation}
\lim_{N \to \infty} \frac{\sum_{\alpha = 1}^{N} \mathbb{E}[|W_{N,\alpha}|^{2+\delta}] \mathcal{O}_{N}}{\sum_{\alpha = 1}^{N} \mathbb{E}[|W_{N,\alpha}|^2] \mathcal{O}_{N}}^{1+\delta/2 = 0}
\end{equation}

which reduces to showing that for some \(r > 2\),

\begin{equation}
\lim_{N \to \infty} N^{-(\frac{r}{2})/2} \left( \frac{1}{N} \sum_{\alpha = 1}^{N} \left| a_{N,\alpha}(E_{N}^{\infty}, C(N), \xi) \right|^r \right) = 0
\end{equation}

\begin{equation}
\left( \frac{1}{N} \sum_{\alpha = 1}^{N} a_{N,\alpha}^2(E_{N}^{\infty}, C(N), \xi) \right)^{r/2}
\end{equation}

The denominator of the second factor of (3.22) is \(\xi V_{N} \xi'\) which by Theorem 3.1 converges to some positive quantity for any non-null \(\xi\) as \(\xi\) is assumed to be positive definite. So we require only
to show that for some $r > 2$, the numerator of the second factor of $(3.22)$ is bounded in probability. Since $a_{N, \alpha}(E_{N, \alpha}, \xi)$ is a linear function of $E_{N, \alpha}$ and $\xi$ (with the sign of the coefficients being a minus or plus depending on $\xi(N)$), on applying the well known inequality that

$$|\sum_{i=1}^{p} a_i|^r \leq p^{r-1} \sum_{i=1}^{p} |a_i|^r$$

(Cf. Loève [23], pp. 155), we get

$$(3.23) \quad \frac{1}{N} \sum_{\alpha=1}^{N} |a_{N, \alpha}(E_{N, \alpha}, \xi)|^r \leq p^{r-1} \sum_{i=1}^{p} |\xi_i|^r \cdot \left( \frac{1}{N} \sum_{\alpha=1}^{N} |E_{N, \alpha}|^r \right).$$

Since, by using Chernoff-Savage conditions, we have for $2 < r < 2+\delta$, $\delta > 0$,

$$\frac{1}{N} \sum_{\alpha=1}^{N} |E_{N, \alpha}|^r < \infty , \quad \text{for } i = 1, \ldots, p,$$

we see that $(3.22)$ converges to zero in probability as $N \to \infty$. The proof follows.

**Theorem 3.3.** Under the assumptions 1 to 4, the limiting permutational distribution of $S_N = \frac{1}{N} T_N V_N^{-1} T_N'$ under the probability measure $\mathcal{P}_N$, is central chi square with $p$ degrees of freedom.

The proof of this theorem is an immediate consequence of Theorem 3.2 and is therefore omitted.

Let us now define

$$(3.22') \quad \xi_i = \int_{-\infty}^{+\infty} J_i[F_1(x) - F_1(-x)] dF_{[1]}(x); \quad i = 1, \ldots, p,$$
and $\xi = (\xi_1, \ldots, \xi_p)$.

**Theorem 3.4.** The test based on $S_N$ is consistent against the set of alternatives that $\xi \neq 0$. In particular, if $F_1(x)$ is symmetric about some $\delta_1$, $\delta = (\delta_1, \ldots, \delta_p)$, then $\xi \neq 0$ for any $\delta \neq 0$, and hence the test is consistent against any shift in location of a symmetric distribution.

**Proof:** It is easy to show that $\frac{1}{N} T^{(1)}_N \rightarrow \xi_1$ in probability as $N \to \infty$ for all $i = 1, \ldots, p$. Thus, $\frac{1}{N} S_N \rightarrow \xi \sim \frac{1}{N} \xi'$ in probability as $N \to \infty$, where $\chi$ is positive definite by assumption. Consequently $S_N$ will be stochastically indefinitely large as $N \to \infty$. Again, $S_N, \varepsilon \sim \chi^2_{p, \alpha}$ as $N \to \infty$. Hence

$$P(S_N \geq S_N, \varepsilon | \xi) \geq P(\frac{1}{N} S_N \geq \frac{1}{N} S_N, \varepsilon | \xi)$$

and since the right hand inequality tends to one as $N \to \infty$, the test is consistent against $\xi \neq 0$. Furthermore, if $F_1(x)$ is symmetric about some $\delta_1 \neq 0$, then it is easy to show that $\xi_1 \neq 0$. Hence the theorem.

4. **Asymptotic Normality of $T_N$ for Arbitrary $F$.**

In this section we shall prove that under a certain set of conditions the random vector $T_N = (T^{(1)}_N, \ldots, T^{(p)}_N)$ has a joint normal distribution in the limit. For the sake of convenience we shall consider the statistic

$$T_N^* = (T_N, 1, \ldots, T_N, p)$$

where
\[ T_{N,j} = \frac{1}{N} \sum_{i=1}^{N} E_{N,1}^{(j)} Z_{N,1}^{(j)} \]

where, \( Z_{N,1}^{(j)} = 1 \) if \( X_{1}^{(j)} > 0 \), and \( Z_{N,1}^{(j)} = 0 \) otherwise.

\( E_{N,i}^{(j)}; i = 1, \ldots, N; j = 1, \ldots, p \), are the constants satisfying the assumptions (1) to (3) of the previous section. The reader may notice that the vector \( \frac{T_{N}^{*}}{E_{N}} \) is related to the vector \( T_{N} \) by the relation \( \frac{T_{N}}{E_{N}} = 2N \frac{T_{N}^{*}}{E_{N}} - N \frac{E_{N}}{E_{N}} \) where

\[ E_{N} = \frac{1}{N} \left( \sum_{i=1}^{N} E_{N,1}^{(1)}, \ldots, \sum_{i=1}^{N} E_{N,1}^{(p)} \right) \]

and so it suffices to consider the equivalent statistic \( \frac{T_{N}}{E_{N}} \).

The main theorem of this section is the following:

**Theorem 4.1.** Under the assumptions (1) to (3) of Section 3, the random vector \( N^{1/2} \left( \frac{T_{N}^{*}}{E_{N}} - \xi_{N}(\theta) \right) \) has asymptotically a p-variate normal distribution with mean vector zero, and covariance matrix \( \sigma_{N,jk} \), where

\[ \mu_{N} = (\mu_{N1}, \ldots, \mu_{Np}); \mu_{N,j} = \int_{0}^{\infty} J_{j}(H_{j}(x; \theta_{j})) dP_{j}(x; \theta_{j}) \]

\[ j = 1, \ldots, p, \]

and where \( (\sigma_{N,jk}) \) is given by (4.14) and (4.15) respectively.

**Proof:** Writing \( J_{N,j}^{(N+1)} H_{N,j} \) as

\[ (J_{N,j}^{(N+1)} H_{N,j}) - J_{j}^{(N+1)} H_{N,j} + J_{j}^{(N+1)} H_{N,j}, \]

and

\[ J_{j}^{(N+1)} H_{N,j} \]

as

\[ J_{j}(H_{j}) + (H_{N,j} - H_{j})J_{j}(H_{j}) - \frac{N+1}{N+1} J_{j}(H_{j}) + [J_{j}^{(N+1)} H_{N,j} - J_{j}^{(N+1)} H_{j}] - \]

\[ \frac{N}{N+1} (H_{N,j} - H_{j}) J_{j}^{(N+1)} H_{j} \]
and,

$$dF_{N,j} \text{ as } d(F_{N,j} - F_j + F_j),$$

and, simplifying, we can express $T_{N,j}$ as

$$(4.4) \quad T_{N,j} = \mu_{N,j} + B_{1N,j} + B_{2N,j} + \sum_{i=1}^{N} C_{1N,j}$$

where

$$(4.5) \quad \mu_{N,j} = \int J_j(H_j(x, \theta_j)) dF_j(x, \theta_j)$$

$$(4.6) \quad B_{1N,j} = \int J_j(H_j(x, \theta_j)) d(F_{N,j}(x) - F_j(x, \theta_j))$$

$$(4.7) \quad B_{2N,j} = \int (H_n,j - H_j) J_j(H_j) dF_j(x)$$

$$(4.8) \quad C_{1N,j} = -\frac{1}{N+1} \int H_n,j J_j'(H) dF_{N,j}(x)$$

$$(4.9) \quad C_{2N,j} = \int (H_n,j - H_j) J_j'(H_j) d(F_{N,j}(x) - F_j(x))$$

$$(4.10) \quad C_{3N,j} = \int [J_j(H_n,j) - J_j(H_j) - (\frac{N}{N+1} H_n,j - H_j) J_j'(H_j)] dF_{N,j}$$

$$(4.11) \quad C_{4N,j} = \int [J_n,j (\frac{N}{N+1} H_n,j - J(\frac{N}{N+1} H_n,j)) - J_n,j] dF_{N,j}(x).$$

Proceeding as in Govindrajulu et al. (1965) it can be shown that the $C$-terms are all $o_p(1/N^{1/2})$. Hence the difference $\sqrt{N}(T_{N,j} - \mu_{N,j}) - \sqrt{N}(B_{1N,j} + B_{2N,j})$ tends to zero in probability. Thus to prove this theorem, it suffices to show that for any real $\delta_i$, $i = 1, \ldots, p$, not all zero, $N^{1/2} \sum_{j=1}^{p} \delta_j (B_{1N,j} + B_{2N,j})$ has normal distribution in the limit.

Now proceeding as in [4], we can express $N^{1/2} \sum_{j=1}^{p} \delta_j (B_{1N,j} + B_{2N,j})$
as a sum of independent and identically distributed random variables having finite first two moments. The proof follows.

To compute the variance-covariance matrix of $B_{1N,j} + B_{2N,j}$, we note from (4.9) and (4.10) that

$$
(4.12) \quad B_{1N,j} + B_{2N,j} = \int (F_{N,j}(x) - F_j(x; \theta_j)) J_j'(H_j(x; \theta_j)) \ dF_j(-x; \theta_j)
$$

$$
- \int (F_{N,j}(-x) - F_j(-x; \theta_j)) J_j'(H_j(x; \theta_j)) \ dF_j(x; \theta_j)
$$

This has mean zero, and variance

$$
= E \left\{ \int (F_{N,j}(x) - F_j(x; \theta_j)) J_j'(H_j(x; \theta_j)) \ dF_j(-x; \theta_j)
$$

$$
- \int (F_{N,j}(-x) - F_j(-x; \theta_j)) J_j'(H_j(x; \theta_j)) \ dF_j(x; \theta_j) \right\}^2
$$

$$
= \frac{2}{N} \int_0^\infty \int_0^\infty F_j(x; \theta_j)[1 - F_j(y; \theta_j)] J_j'(H_j(x; \theta_j)) J_j'(H_j(y; \theta_j)) dF_j(-x; \theta_j)
$$

$$
+ \frac{2}{N} \int_0^\infty \int_0^\infty F_j(-y; \theta_j)[1 - F_j(-x; \theta_j)] J_j'(H_j(x; \theta_j)) J_j'(H_j(y; \theta_j)) dF_j(x; \theta_j) dF_j(y; \theta_j)
$$

$$
- \frac{2}{N} \int_0^\infty \int_0^\infty F_j(-y; \theta_j)[1 - F_j(x; \theta_j)] J_j'(H_j(x; \theta_j)) J_j'(H_j(y; \theta_j)) dF_j(-x; \theta_j) dF_j(y; \theta_j)
$$

(4.13)

[Note that the application of Fubini's theorem permits the interchange of integral and expectation.]

Hence

$$
(4.14) \quad \sigma_{N,j}^2 = 2 \int_0^\infty \int_0^\infty F_j(x; \theta_j)[1 - F_j(y; \theta_j)] J_j'(H_j(x; \theta_j)) J_j'(H_j(y; \theta_j)) dF_j(-x; \theta_j) dF_j(-y; \theta_j)
$$

$$
+ 2 \int_0^\infty \int_0^\infty F_j(-y; \theta_j)[1 - F_j(-x; \theta_j)] J_j'(H_j(x; \theta_j)) J_j'(H_j(y; \theta_j)) dF_j(x; \theta_j) dF_j(y; \theta_j)
$$

$$
- 2 \int_0^\infty \int_0^\infty F_j(-y; \theta_j)[1 - F_j(x; \theta_j)] J_j'(H_j(x; \theta_j)) J_j'(H_j(y; \theta_j)) dF_j(-x; \theta_j) dF_j(y; \theta_j)
$$
Similarly,

\[ (4.15) \sigma_{N,jk} = \int_{x=0}^{\infty} \int_{y=0}^{\infty} [F_{j,k}(x,y;\theta_j,\theta_k) - F_j(x;\theta_j)F_k(y;\theta_k)] J'_k[H_j(x;\theta_j)] \]

\[ J'_k[H_k(y;\theta_k)] \, dF_j(-x;\theta_j) \, dF_k(-y;\theta_k) \]

\[ - \int_{x=0}^{\infty} \int_{y=0}^{\infty} [F_{j,k}(x,-y;\theta_j,\theta_k) - F_j(x;\theta_j)F_k(-y;\theta_k)] J'_j[H_j(x;\theta_j)] \]

\[ J'_k[H_k(y;\theta_k)] \, dF_j(x;\theta_j) \, dF_k(y;\theta_k) \]

\[ + \int_{x=0}^{\infty} \int_{y=0}^{\infty} [F_{j,k}(-x,-y;\theta_j,\theta_k) - F_j(-x;\theta_j)F_k(-y;\theta_k)] J'_j[H_j(x;\theta_j)] \]

\[ J'_k[H_k(y;\theta_k)] \, dF_j(x;\theta_j) \, dF_k(y;\theta_k) . \]

5. **Asymptotic Distribution Under Translation Alternatives.**

From this section onward, we shall concern ourselves with a sequence of admissible alternative hypothesis \( H_N^P \), which specifies that for each \( j,k = 1,\ldots,p \),

\[ (5.1) \quad F_j(x;\theta_j) = F_j(x - \frac{\theta_j}{\sqrt{N}}); \quad F_{j,k}(x,y;\theta_j,\theta_k) = F_{j,k}(x - \frac{\theta_j}{\sqrt{N}}, y - \frac{\theta_k}{\sqrt{N}}) \]

where \( F_j(x) \) and \( F_{j,k}(x,y) \) are symmetric about zero, and \( \theta = (\theta_1,\ldots,\theta_p) \) is unknown.

We shall also assume that the constant \( E_{n,1}^{(j)}, i=1,\ldots,N; j=1,\ldots,p \) is the expected value of the \( i \)-th order statistic
of a sample of size $N$ from a distribution function $\psi_j(x)$ given by

$$
(5.2) \quad \psi_j(x) = \psi_j^*(x) - \psi_j^*(-x) ; \quad x \geq 0 ,
$$

where $\psi_j^*(x)$ is a distribution function either symmetric about zero or uniform over $(0,1)$. It may be noted that the above definition of $E_{N,j}^{(j)}$ implies that the function

$$
(5.3) \quad J_j(x) = \psi_j^{-1}(x) = \psi_j^*\left(\frac{1+x}{2}\right).
$$

We first prove the following lemma.

Lemma 5.1. If $J_j = \psi_j^{-1}$, then

$$
(5.1) \quad 4 \int_0^\infty \int_0^\infty F_j(x)[1-F_j(y)]J_j^'(2F_j(x)-1)J_j^'(2F_j(y)-1) \, dF_j(x) \, dF_j(y) \\
+ 2 \int_0^\infty \int_0^\infty (1-F_j(x))(1-F_j(y))J_j^'(2F_j(x)-1)J_j^'(2F_j(y)-1) \, dF_j(x) \, dF_j(y) \\
= \frac{1}{2} \int_0^\infty H_j(x)[1-H_j(y)]J_j^'(2F_j(x)-1)J_j^'(2F_j(y)-1) \, dH_j(x) \, dH_j(y) \\
+ \frac{1}{4} \left( \int_0^1 J_j(x) \, dx \right)^2 \\
= \frac{1}{4} \int_0^\infty J_j^2(x) \, dx.
$$

$$
(5.2) \quad \int_0^\infty \int_{-\infty}^\infty [F_{j,k}(x,y)-F_j(x)F_k(y)]J_j^'(2F_j(x)-1)J_k^'(2F_k(y)-1) \, dF_j(-x) \, dF_k(-y) \\
- \int_0^\infty \int_{-\infty}^\infty [F_{j,k}(x,-y)-F_j(x)F_k(-y)]J_j^'(2F_j(x)-1)J_k^'(2F_k(y)-1) \, dF_j(-x) \, dF_k(y) \\
- \int_0^\infty \int_{-\infty}^\infty [F_{j,k}(x,0)-F_j(x)F_k(0)]J_j^'(2F_j(x)-1)J_k^'(2F_k(0)-1) \, dF_j(-x) \, dF_k(0) \\
- \int_0^\infty \int_{-\infty}^\infty [F_{j,k}(0,y)-F_j(0)F_k(y)]J_j^'(2F_j(0)-1)J_k^'(2F_k(y)-1) \, dF_j(0) \, dF_k(y).
$$
\[
- \int_{x=0}^{\infty} \int_{y=0}^{\infty} \left[ F_{j,k}(-x,y) - F_j(-x)F_k(y) \right] J'_j(2F_j(x)-1) J'_k(2F_k(y)-1) \, dF_j(x) \, dF_k(-y) \\
+ \int_{x=0}^{\infty} \int_{y=0}^{\infty} \left[ F_{j,k}(-x,-y) - F_j(-x)F_k(-y) \right] J'_j(2F_j(x)-1) J'_k(2F_k(y)-1) \, dF_j(x) \, dF_k(y) \\
= \frac{1}{4} \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} J_j(2F_j(x)-1) K_k(2F_k(y)-1) \, dF_j, k(x,y), \quad j \neq k, \quad J_j = \psi_j^{-1}.
\]

**Proof:** To establish the first equality of (5.1), we note that \( H_j(x) = 2F_j(x)-1 \), and so the right hand side of the first equality can be written as

\[
\frac{1}{4} \iint_{0<x<y<\infty} (2F_j(x)-1)(1-F_j(y))J'_j(2F_j(x)-1)J'_j(2F_j(y)-1) \, dF_j(x) \, dF_j(y) \\
+ \frac{1}{4} \left( \int_{0}^{1} J_j(x) \, dx \right)^2,
\]

which equals

\[
\frac{1}{4} \iint_{0<x<y<\infty} F_j(x)[1-F_j(y)]J'_j(2F_j(x)-1)J'_j(2F_j(y)-1) \, dF_j(x) \, dF_j(y) \\
- \frac{1}{4} \iint_{0<x<y<\infty} (1-F_j(x))(1-F_j(y))J'_j(2F_j(x)-1)J'_j(2F_j(y)-1) \, dF_j(x)dF_j(y) \\
+ \frac{1}{4} \left( \int_{0}^{1} J_j(x) \, dx \right)^2
\]

that is,
\[ 4 \int \int F_j(x)[1-F_j(y)]J'_j(2F_j(x)-1)J'_j(2F_j(y)-1) \, dF_j(x) \, dF_j(y) \]
\[ + 2 \int \int (1-F_j(x))(1-F_j(y))J'_j(2F_j(x)-1)J'_j(2F_j(y)-1) \, dF_j(x) \, dF_j(y) \]
\[ - 4 \int \int (1-F_j(x))(1-F_j(y))J'_j(2F_j(x)-1)J'_j(2F_j(y)-1) \, dF_j(x) \, dF_j(y) \]
\[ + \frac{1}{4} \left( \int J_j(x) \, dx \right)^2. \]

The last two integrals cancel out, and this proves the first equality. To prove the second equality, note that we can rewrite the second equation as

\[ \frac{1}{2} \int \int x(1-y)J'_j(x)J'_j(y) \, dx \, dy + \frac{1}{4} \left( \int J_j(x) \, dx \right)^2, \]

and the result follows as in Chernoff-Savage (1958).

To prove (5.2), we note that when \( \psi \) is uniform over \((0,1)\), the left hand side reduces to

\[ \int \int F_{j,k}(x,y) \, dF_j(-x) \, dF_k(-y) - \frac{9}{64} \]
\[ - \int \int F_{j,k}(x,y) \, dF_j(-x) \, dF_k(y) - \frac{3}{64} \]
\[ - \int \int F_{j,k}(-x,y) \, dF_j(x) \, dF_k(y) - \frac{3}{64} \]
\[ + \int \int F_{j,k}(-x,-y) \, dF_j(x) \, dF_k(y) - \frac{1}{64}. \]

Noting that \( F_j(x) = 1 - F_j(-x) \), making simple transformations, and integrating by parts, the above expressions equal
\[
\int_{-\infty}^{+\infty} \int_{-\infty}^{+\infty} F_j(x) F_k(y) \, dF_{j,k}(x,y) - \frac{1}{4}
\]

The proof of the case when \( J_j(x) = \psi_j^{-1}(x) \) where \( \psi_j(x) \) is given by (5.2) follows by using the relations

\[
J'_j(2F_j(x)-1) = \frac{1}{2} J'_j(F_j(x)) = J'_j(1-2F_j(x)),
\]

\( j = 1, \ldots, p, \quad J_j = \psi_j^{-1}, \quad J_j^* = \psi_j^*{-1}, \)

and proceeding as above.

**Theorem 5.1.** If, for every \( j = 1, \ldots, p, \)

(1) the conditions of Theorem 4.1 are satisfied

(2) \( F_j(x;\theta_j) = F_j(x-\frac{\xi_j}{\sqrt{N}}); \quad F_j(x,y;\theta_j,\theta_k) = F_j(x-\frac{\xi_j}{\sqrt{N}},y-\frac{\xi_k}{\sqrt{N}}) \)

then \( N^{1/2}(\tau^*_N-k_N) \) has asymptotically a \( p \)-variate normal distribution with mean vector zero, and covariance matrix

\[
\sum^*_N = (\tau_{jk}) \text{ where}
\]

\[
\tau_{jk} = \begin{cases} 
\frac{1}{4} \int_{-\infty}^{+\infty} J_j^2(x) \, dx, & j = k; \quad J_j = \psi_j^{-1}; \\
\frac{1}{4} \int_{-\infty}^{+\infty} \int_{-\infty}^{+\infty} J_j(2F_j(x)-1)J_k(2F_k(y)-1) \, dF_{j,k}(x,y), & j \neq k, \quad J_j = \psi_j^{-1}.
\end{cases}
\]

(5.3)

The proof of this theorem is a direct consequence of Theorem 4.1 and Lemma 5.1 and is therefore omitted.

It may be noted that the limiting distribution of \( N^{1/2}(\tau^*_N-k_N) \) is nonsingular if and only if the functions \( J_j \) and \( F_j \) are such that the moment matrix of \( \{J_j(2F_j(x)-1), j=1, \ldots, p\} \) is non-singular.
The following theorem gives the conditions under which the limiting distribution of $N^{1/2}(T_{N}^{*}-\mu_{N})$ is singular.

Theorem 5.2. Let $X_{j}$, $j = 1, \ldots, N$, be random vectors satisfying the assumptions of Theorem 5.1. Then the asymptotic distribution of $N^{1/2}(T_{N}^{*}-\mu_{N})$ is $k$-variate normal, $k < p$ if and only if a.s. F.

(5.4) \[ J_{j}(2F_{j}(x)-1) = \sum_{k \neq j} \alpha_{k}J_{k}(2F_{k}(y)-1) + \text{constant}, \]

[$\alpha_{k}$ are some constants].

The proof of this theorem is analogous to that of Theorem 3.3 of Bickel (1964) and is therefore omitted.


We now assume that the covariance matrix $\sum^{*} = (\tau_{jk})$ defined by (5.3) is non-singular. Then for testing the hypothesis $H_{0}$ given by (1.3), we propose to consider the test statistics $S_{N}^{*}$ defined as

(6.1) \[ S_{N}^{*} = N(T_{N}^{*}-\mu_{N}) \sum^{*-1} (T_{N}^{*}-\mu_{N})(0)), \]

where $T_{N}^{*}$ is defined by (4.1) and (4.2); $\sum^{*-1}$ is a consistent estimator of $\sum^{*-1}$, and

(6.2) \[ \mu_{N}(0) = (\mu_{N,1}(0), \ldots, \mu_{N,p}(0)) \]

where

(6.3) \[ \mu_{N,j}(0) = \int_{x=0}^{\infty} J_{j}(2F_{j}(x)-1) \, dF_{j}(x), \quad J_{j} = \Psi_{j}^{-1}. \]
Lemma 6.1. If

(i) the conditions of Theorem 5.1 are satisfied,

(ii) the assumptions of Lemma 7.2 of Puri (1964) hold for each function $F_j$ and $J_j; j = 1, \ldots, p$

then, $N^{1/2}(\sum_t^* \mu_N(0))$ has asymptotically a $p$-variate normal distribution with mean vector,

\[(6.4)\]
\[a = (c_1 \theta_1, \ldots, c_p \theta_p)\]

where

\[(6.5)\]
\[c_j = -2 \int_0^\infty J_j'(2F_j(x)-1)f_j(x) \, dF_j(x)\]

and covariance matrix $\sum_t^* = (\tau_{jk})$ given by (5.3). [Here $f_j(x)$ is the density of $F_j(x)$.] The proof of the lemma follows by noting (that by assumption (ii))

\[(6.6)\]
\[\lim_{N \to \infty} N^{1/2}(\mu_N,J_j(0) - \mu_N,J_j(0)) = c_j \theta_j\]

and then applying Slutski's theorem ([12], p. 254).

A consequence of this lemma is the following:

Theorem 6.1. If the assumptions of Lemma 6.1 are satisfied, then for $N \to \infty$, the limiting distribution of

\[
N(\sum_t^* \mu_N(0))' = \sum_t^* \mu_N(0)'
\]

is non-central chi square with $p$ degrees of freedom and noncentrality parameter

\[(6.7)\]
\[b(S_N^*) = a \sum_t^* a' = \theta \sum_t^* \theta'\]
where \( a \) is given by \((6.4)\) and \((6.5)\), and \( \sum_e^{**} = (\tau_{jk}) \) is given by

\[
(6.8) \quad \tau_{jk}^* = \tau_{jk}/c_{jk}
\]

Now since \( \sum_e^{*-1} \) is a consistent estimator of \( \sum_e^{*-1} \), we have the following

**Theorem 6.2.** If the assumptions of Lemma 6.1 are satisfied, then for \( N \to \infty \), the limiting distribution of the statistic \( S_N^* \) is non-central chi square with \( p \) degrees of freedom, and non-centrality parameter \( \delta(S_N^*) \) given by \((6.7)\).

From Theorem 6.2, it is clear that the choice of \( \sum_e^{*-1} \) is of no importance in the limit. Any consistent estimator of \( \sum_e^{*-1} \) will preserve the asymptotic distribution of the statistic \( S_N^* \). One such consistent estimator is provided by the permutational covariance elements in Section 3.

We may also use a theorem by Bhuchongkul(1964) to propose a similar class of estimates. However here conditions are relatively more restrictive (since they relate to asymptotic normality) than the ones in Theorem 3.1, and for our purpose, we need not bother about the conditions as we simply require here the asymptotic convergence (not normality) of the estimates.

**Theorem 6.3.** The permutation test based on \( S_N \) given by \((2.21)\) and the asymptotically non-parametric test based on \( S_N^* \) given
by (6.1) are asymptotically power equivalent for the sequence of translation type of alternatives defined by (5.1).

**Proof:** From Theorem 5.1, it follows that for any such translation type of alternatives, the covariance matrix of the proposed vector of rank order statistics converges asymptotically to $\sum^*$, of which $\hat{\sum}^*$ is a consistent estimator. Again by Corollary 3.1, it readily follows that under any such sequence of translation type of alternatives, the permutation covariance matrix in Theorem 3.2 also converges to $\tilde{\sum}^*$. Thus by looking at (2.21) and Theorem 6.2, we observe that under (5.1), $S_N \overset{P}{\sim} S^*_N$ where $\overset{P}{\sim}$ means stochastically equivalent. The proof follows.

By virtue of the stochastic equivalence of the tests $S_N$ and $S^*_N$, we shall only consider the asymptotic properties of the unconditional test based on $S_N^*$ in the next section.

**Special Cases:**

A. Let $J_j(u) = u$, then the $S^*_N$ test reduces to the rank-sum $S^*_N(R)$ test (which may be regarded as a p-variate version of the univariate one sample test) considered in detail by Bickel (1965). In this case the non-centrality parameter (6.7) reduces to

$$5(R) = \varphi \gamma^{-1} $$

where $\gamma = (\gamma_{jk})$ is given by
\[ \gamma_{jk} = \begin{cases} 
1/12 ( \int f_j^2(x) dx )^2, & j = k \\
\int_{-\infty}^{+\infty} \int_{-\infty}^{+\infty} F_j(x) F_k(y) \, dF_{j,k}(x,y) - 1/4 \\
\left[ \int_{-\infty}^{+\infty} f_j^2(x) dx \right] \left[ \int_{-\infty}^{+\infty} f_k^2(y) dy \right] & j \neq k 
\end{cases} \]

B. Let \( J_j \) be the inverse of a chi-distribution with one degree of freedom. Then the \( S^*_N \) test reduces to the normal scores \( S^*_N (\Phi) \) test which may be regarded as a \( p \)-variate version of the univariate one-sample normal scores test.

In this case the non-centrality parameter (6.7) reduces to

\[ \delta(\Phi) = \Theta^{-1} \Theta' \]  

where \( \lambda = (\lambda_{jk}) \) is given by

\[ \lambda_{jk} = \begin{cases} 
\frac{1}{\left( \int_{-\infty}^{+\infty} \frac{f_j^2(x)}{\Phi^{-1}(F_j(x))} dx \right)^2} & j = k \\
\int_{-\infty}^{+\infty} \int_{-\infty}^{+\infty} \frac{1}{\Phi^{-1}(F_j(x))} \frac{1}{\Phi^{-1}(F_k(y))} \, dF_{j,k}(x,y) \\
\int_{-\infty}^{+\infty} f_j^2(x) dx \int_{-\infty}^{+\infty} f_k^2(y) dy & j \neq k 
\end{cases} \]
7. Asymptotic Relative Efficiency.

The asymptotic efficiency of one test relative to another in the present context can easily be obtained by a method given explicitly by Hanan (1956). It may be stated roughly as follows. If the two test statistics have, under the alternative hypothesis, non-central chi square distribution with the same number of degrees of freedom, the asymptotic relative efficiency of one test with respect to the other test is equal to the ratio of their non-centrality parameters after the alternatives have been set equal. For details, the reader is also referred to Bickel (1965), Puri (1964) and Sen (1965) among others. Our main interest will be to study the relative performance of (i) the $S^*_N(\Phi)$ test, (ii) the $S^*_N(R)$ test and (iii) the Hotelling's $T^2_N$ test, which too has asymptotically a non-central chi square distribution with $p$ degrees of freedom, and non-centrality parameter

\begin{equation}
5(T^2_N) = \varrho \prod \varrho' \\prod^{-1}
\end{equation}

where $\prod = (\sigma_{jk})$ is the covariance matrix of the underlying distribution function $F$.

Thus, applying the definition of Hanan, and denoting $e_{T,T^*}$ as the asymptotic efficiency of a test $T$ relative to $T^*$, we have (cf. (6.7))

\begin{equation}
e_{S^*_N,T^2_N} = \left( \varrho \sum_{\sim}^{**} \varrho' \right) / \left( \varrho \prod \varrho' \right)
\end{equation}

where $\sum_{\sim}^{**} = (\tau^*_{jk})$ is given by (6.8) and $\prod = (\sigma_{jk})$ is the covariance matrix of $F$. 
Special Cases:

(a) **Normal Scores and Rank Sum Tests.**

From (7.2), we find that the efficiencies of the normal scores \( S_N^*(\mathbf{Y}) \) test and the rank sum \( S_N^*(R) \) test relative to \( T_N^2 \)-test are

\[
e_{S_N^*(\mathbf{Y}), T_N^2} = \left( \mathbf{\Theta} \left( \lambda^{-1} \mathbf{\Theta}' \right) \right) / \left( \mathbf{\Theta} \mathbf{T}^{-1} \mathbf{\Theta}' \right)
\]

where \( \lambda = (\lambda_{jk}) \) is given by (6.12) and \( \mathbf{T} = (\sigma_{jk}) \).

\[
e_{S_N^*(R), T_N^2} = \left( \mathbf{\Theta} \left( \gamma^{-1} \mathbf{\Theta}' \right) \right) / \left( \mathbf{\Theta} \mathbf{T}^{-1} \mathbf{\Theta}' \right)
\]

where \( \gamma = (\gamma_{jk}) \) is given by (6.11).

We may remark that the expression (7.4) is the same as the one obtained by Bickel (1965) for the p-variate one-sample problem, and by Chatterjee and Sen (1964) for the bivariate two-sample problem. For the study of the various aspects of the efficiency (7.4) in special situations the reader is referred to the interesting papers of Bickel (1965) and Chatterjee and Sen (1964).

(b) **Totally Symmetric Case.** A bivariate random vector \((X, Y)\) is said to be totally symmetric if \((X, Y), (X, -Y), (-X, Y)\) and \((-X, -Y)\) have the same distribution function. It can be shown following the lines of the argument of Bickel (1965), that a sufficient condition for the asymptotic independence of the components of \( T_N^* \) is the total symmetry of \((X_{1}^{(j)}, X_{1}^{(k)})\) for every pair \((j, k)\).
Thus in the event of totally symmetric case $\lambda, \Pi, \gamma$ are all diagonal matrices, and hence we have

\begin{align*}
(7.5) & \quad e_{S^*_{(\Phi)}, T_N^2} = \sum_{j=1}^{p} \frac{\theta_j^2 (\int_{-\infty}^{+\infty} \frac{f_j^2(x) \, dx}{\phi[\Phi^{-1}(f_j(x))]} )^2}{\sum_{j=1}^{p} \theta_j^2 / \sigma_j^2} \\
(7.6) & \quad e_{S^*_{(R)}, T_N^2} = 12 \sum_{j=1}^{p} \theta_j^2 (\int_{-\infty}^{+\infty} f_j^2(x) \, dx)^2 / \sum_{j=1}^{p} \theta_j^2 / \sigma_j^2 \\
(7.7) & \quad e_{S^*_{(\Omega)}, S^*_{(R)}} = \sum_{j=1}^{p} \theta_j^2 (\int_{-\infty}^{+\infty} \frac{f_j^2(x) \, dx}{\phi[\Phi^{-1}(f_j(x))]} )^2 / 12 \sum_{j=1}^{p} \theta_j^2 (\int_{-\infty}^{+\infty} f_j^2(x) \, dx)^2}
\end{align*}

Applying a theorem of Courant (cf. footnote) Bickel (1965) proved that

\begin{equation}
(7.8) \quad \inf_{\Theta} \inf_{S^*_{(R)}, T_N^2} e_{S^*_{(R)}, T_N^2} \leq 0.86
\end{equation}

where $\mathcal{F}$ is the family of all totally symmetric $p$-variate distributions whose marginal densities exist.

While the authors were working on the study of bounds of the efficiencies (7.5) and (7.7), the results of Bhattacharyya (1966) who was considering a two-sample

\begin{footnote}
Theorem (Courant) The maximal and minimal values of $xAx' / xBx'$, where $A$ and $B$ are non-negative definite, and $B$ is non-singular, are given by the maximal and minimal eigenvalues of $AB^{-1}$.
\end{footnote}
multivariate problem were made public. It turns out that the efficiency expressions (7.5) and (7.7) are the same as in the case of the corresponding tests for the multivariate two sample problem. Thus for the ease of reference we give below some of the results for the proofs of which the reader is referred to [2].

\[
\inf_{F \in \mathcal{F}} \inf_{\Theta} e_{S_{N}^{*}(\bar{X}), T_{N}^{2}} = 1
\]

(7.9)

\[
\inf_{F \in \mathcal{F}} \inf_{\Theta} e_{S_{N}^{*}(\bar{X}), S_{N}^{*}(R)} = \pi/6.
\]

(7.10)

In the passing we may remark that when the components of F are totally symmetric as well as identically distributed, then the expressions (7.5), (7.6) and (7.7) become independent of \( \Theta \)'s, and the results are the same as in the case of the corresponding univariate one-sample problems.

(c) Normal Case.

Let us now assume that the underlying distribution function F is a non-singular p-variate normal with mean vector zero and covariance matrix \( \Sigma = (\sigma_{jk}) \). Then it can easily be checked that

\[
e_{S_{N}^{*}(\bar{X}), T_{N}^{2}} = 1.
\]

(7.11)

This means that in the case of normal distributions, the property of the univariate normal scores test relative to the student's t test is preserved in the multivariate case.
This is interesting in the sense that the same is not the case with the multivariate rank sum test as Bickel (1965) has shown that

\[(7.12) \quad \inf_{\Phi} \inf_{\Theta} e_{S_N(R), T_N} = 0 \quad \text{for } p \geq 3\]

and

\[(7.13) \quad \frac{3}{\pi} \leq e_{S_N(R), T_N} \leq 0.965 \quad \text{for } p = 2.\]

From (7.11) and (7.12), it follows that

\[(7.14) \quad \sup_{\Phi} \sup_{\Theta} e_{S_N(\Phi), S_N(R)} = \infty \quad \text{for } p \geq 3\]

and Bhattacharyya (1966) has proved that

\[(7.15) \quad \inf_{\Phi} \inf_{\Theta} e_{S_N(\Phi), S_N(R)} \geq 1 \quad \text{for } p \geq 2.\]

[\(\Phi\) is the family of all nonsingular \(p\)-variate normal distributions.]

The above results indicate that in case the underlying distribution is normal, the normal scores test is always preferable to the rank-sum test.

8. Estimation Problem.

One Sample Case:

As before, let \(X = (X_1, \ldots, X_p)\), \(\alpha = 1, \ldots, N\) be a sample from a \(p\)-variate cdf \(F(x_1 - \Theta_1, \ldots, x_p - \Theta_p)\) where \(\Theta = (\Theta_1, \ldots, \Theta_p)\) is unknown, \(F\) is symmetric about zero, and \(F\) is continuous. Our aim to find suitable estimates for the parameter \(\Theta\). Suppose that the Hodges-Lehmann statistic \(h\) \([(5.1)\) of [18]] is calculated for every univariate
sample $X_1^{(j)}, \ldots, X_N^{(j)}$; $j = 1, \ldots, p$. We shall write
$h_j = h_j(x_1^{(j)}, \ldots, x_N^{(j)})$ for the value obtained from sample
$X_1^{(j)}, \ldots, X_N^{(j)}$ on the $j$-th variate. Thus we have

$$(8.1) \quad h_j = \frac{1}{N} \sum_{i=1}^N E_{N,i}^{(j)} Z_{N,i}^{(j)}$$

where $E_{N,i}^{(j)}$ is the expected value of the $i$-th order statistic
of a sample of size $N$ from the distribution function given by

$$\psi_j^*(x) = \psi_j(x) - \psi_j(-x) \quad \text{for } x \geq 0.$$ 

Let

$$(8.3) \quad \Theta_j^* = \sup \{ \Theta_j : h_j(x_1^{(j)} - \Theta_j, \ldots, x_N^{(j)} - \Theta_j) > \mu \}$$
and

$$\Theta_j^{**} = \inf \{ \Theta_j : h_j(x_1^{(j)} - \Theta_j, \ldots, x_N^{(j)} - \Theta_j) < \mu \}$$

where $\mu$ is a point of symmetry of the distribution of
$h_j$ when $\Theta_j = 0$. It was shown in [18] that the estimate
$(\Theta_j^* + \Theta_j^{**})/2 = \hat{\Theta}_{1,N,j}$ of $\Theta_j$ has more robust efficiency than
the classical estimate $\bar{X}_j = \sum_{\alpha=1}^N x_{\alpha,j}/N$.

In this section we shall consider the properties
of the estimate

$$(8.4) \quad \hat{\Theta}_N = (\hat{\Theta}_{1,N}, \ldots, \hat{\Theta}_{N,p})$$

of $\Theta$, and compare it with the normal theory estimate
$
\bar{X} = (\bar{X}_1, \ldots, \bar{X}_p)$ where $\bar{X}_j = \sum_{\alpha=1}^N x_{\alpha,j}/N$, and the estimates
considered by Bickel [964].
8(a). Regularity Properties.

The following theorems are immediate consequences of the theorems proved by Hodges and Lehmann (1963) for the case \( p = 1 \), and are therefore stated without proofs.

**Theorem 8.1.** The distribution of the estimate \( \hat{\Theta}_N = (\hat{\Theta}_{N,1}, \ldots, \hat{\Theta}_{N,p}) \) is (absolutely) continuous if \( F \) is (absolutely) continuous.

**Theorem 8.2.** \( \hat{\Theta}_N(x_1+a, \ldots, x_N+a) = \hat{\Theta}_N(x_1, \ldots, x_N) + a \)

where \( a = (a_1, \ldots, a_p) \) is any \( 1 \times p \) vector of constants.

**Theorem 8.3.** The distribution of \( \hat{\Theta}_N \) is symmetric about \( \Theta \), if

(i) \( F \) is symmetric about zero,

(ii) \( h_j(x_1^{(j)}, \ldots, x_N^{(j)}) + h_j(-x_1^{(j)}, \ldots, -x_N^{(j)}) = 2\mu \).

**Theorem 8.4.** For every vector \( a = (a_1, \ldots, a_p) \)

\[
P[h_j(x_1^{(j)} - a_j, \ldots, x_N^{(j)} - a_j) < \mu, j = 1, \ldots, p]
\]

\[\leq \ P[\hat{\Theta}_N < a] \leq P[h_j(x_1^{(j)} - a_j, \ldots, x_N^{(j)} - a_j) \leq \mu], \ j = 1, \ldots, p.\]

8(b). Asymptotic Normality.

Consider a sequence of sample sizes \( N = 1, 2, \ldots, \)

and let \( \Theta_N \) be a sequence of values of \( \Theta \). We shall indicate the dependence of \( h_j \) and \( \mu \) on \( N \) by writing \( h_{jN} \) and \( \mu_N \).

**Theorem 8.5.** Let \( a = (a_1, \ldots, a_p) \) be a one by \( p \) vector of constants. Let \( c_1, \ldots, c_N, \ldots \) be constants, and let

\( \Theta_N = -a/c_N = -(a_1/c_N, \ldots, a_p/c_N) \).

Let \( \zeta \) be a continuous \( p \)-variate distribution function
whose marginals have means zero and variances one, and suppose

\[ \lim_{N \to \infty} P_N \{ C_N (h_j - \mu_N) \leq u_j ; \ j = 1, \ldots, p \} = \mathcal{G} \left( \frac{u_1 + a_1 B_1}{A_1}, \ldots, \frac{u_p + a_p B_p}{A_p} \right) \]

where \( P_N \) indicates that the probability is computed for the parameter values \( \theta_N \) and where \( h_j \) stands for \( h_j (x_1^{(j)}, \ldots, x_N^{(j)}) \). Then for any fixed \( \Theta \)

\[ \lim_{N \to \infty} P_\Theta \{ C_N (\hat{\Theta}_N - \Theta) \leq a \} = \mathcal{G} \left( \frac{a_1 B_1}{A_1}, \ldots, \frac{a_p B_p}{A_p} \right) \]

**Proof:** By Theorems (8.2) and (8.4)

\[
\lim_{N \to \infty} P_\Theta \{ C_N (\hat{\Theta}_N - \Theta) \leq a \}
= \lim_{N \to \infty} P_\Theta \{ C_N \hat{\Theta}_N \leq a \}
= \lim_{N \to \infty} P_\Theta \{ h_j (x_1^{(j)} - \frac{a_1}{C_N}, \ldots, x_N^{(j)} - \frac{a_p}{C_N}) \leq \mu_N \}
= \lim_{N \to \infty} P_\Theta \{ h_j (x_1^{(j)}, \ldots, x_N^{(j)}) - \mu_N \leq 0 \}
= \mathcal{G} \left( \frac{a_1 B_1}{A_1}, \ldots, \frac{a_p B_p}{A_p} \right).
\]

Now combining Theorems (8.5) and (4.1) we have

**Theorem 8.6.** Under the assumptions of Theorem 5.1, \( N^{1/2} \hat{\Theta}_N - \Theta \) has asymptotically a \( p \)-variate normal distribution with mean vector zero and covariance matrix

\[ \Sigma^{**} = (\tau_{jk}^*) \] given by (6.8).

**Remark 1.** Let \( J_j = \psi_j^{-1} \) where \( \psi_j \) is \( R(0,1) \), then the estimate \( \hat{\Theta}_N \) will be denoted by \( \hat{\Theta}_N(R) \) and in this case the
covariance matrix $\sum^{**} = (\tau_{jk}^*)$ reduces to $\chi = (\gamma_{jk})$
defined by (6.11). [See also Bickel (1964).]

Remark 2. Let $J_j = \psi_j^{-1}$ where $\psi_j$ is given by (5.2);
then the estimate $\hat{\Theta}_N$ will be denoted by $\hat{\Theta}_N,\overline{\theta}$ and in
this case, the covariance matrix $\sum^{**} = (\tau_{jk}^*)$ reduces
to $\lambda = (\lambda_{jk})$ defined by (6.12).

8(c). Asymptotic Efficiency.

To obtain an idea of the relative performance of one estimator with respect to another, we employ the notion
of "the generalized variances" a concept introduced by
Wilks (see Wilks (1962), p. 547). The "generalized variance"
of a $p$-variate random vector $(X_1, ..., X_p)$ with non-singular
covariance matrix $\sum = (\rho_{jk} \sigma_j \sigma_k)$ is defined to be
$\text{var } X = \sigma_1^2 ... \sigma_p^2 \text{ det } \|\rho_{jk}\|$\nwhere "det" denotes the determinant.

Suppose that the two asymptotically unbiased estimates
$T$ and $T'$ of $\Theta$ with asymptotically non-singular matrices
$\sum^T = (\rho_{jk}^T \sigma_j T \sigma_k)$ and $\sum^{T'} = (\rho_{jk}^{T'} T' T' \sigma_j T' \sigma_k)$
require $N$ and $N'$ observations to achieve equal asymptotic generalized
variances. Then the asymptotic relative efficiency of $T'$
with respect to $T$ is defined as

$$
e_{T',T} = \lim_{N \to \infty} \frac{N'}{N} = \left[ \left( \frac{\sigma_1^T \cdots \sigma_p^T}{\sigma_1 T' \cdots \sigma_p T'} \right)^2 \text{det} \|\rho_{jk}^T\| \right]^D$$

Now from (6.8), the generalized variance of $\hat{\Theta}_N$ is
\[
\text{var} \left( \hat{\theta}_N \right) = \prod_{j=1}^{p} \frac{1}{4} \int_{0}^{1} J_j^2(x) \, dx \det \| \rho_{jk}^{***} \| \cdot 4 \left( \int_{x=0}^{\infty} J_j'(2F_j(x) - 1) f_j(x) \, dx \right)^2
\]

where
\[
\rho_{jk}^{***} = \begin{cases} 
\int_{-\infty}^{+\infty} \int_{-\infty}^{+\infty} J_j(2F_j(x) - 1) J_k(2F_k(y) - 1) \, dF_{j,k}(x,y) & j \neq k \\
1 & j = k,
\end{cases}
\]

and, the generalized variance of the mean estimator \( \bar{X}_N = (\bar{X}_{N,1}, \ldots, \bar{X}_{N,p}) \) is

\[
\text{var} \left( \bar{X}_N \right) = \prod_{j=1}^{p} \sigma_j^2 \cdot \det \| \rho_{jk} \|
\]

where \((\sigma_{jk}, \sigma_{jk}, \sigma_{jk})\) is the covariance matrix of \(X_j\).

Hence, we have

\[
\mathbb{E} \left( \hat{\theta}_N \mid \bar{X}_N \right) = \prod_{j=1}^{p} \frac{1}{16 \sigma_j^2 \left( \int_{x=0}^{\infty} J_j'(2F_j(x) - 1) f_j^2(x) \, dx \right)^2} \left[ \begin{array}{c} \frac{1}{4} \int_{0}^{1} J_j^2(x) \, dx \\ \end{array} \right]^{p} \begin{bmatrix} \det \| \rho_{jk}^{***} \| \\ \det \| \rho_{jk} \| \end{bmatrix}.
\]

Our main interest is the study of the relative performance of the estimators \(\hat{\theta}_N, \bar{X}, \hat{\theta}_{N,R}\) and \(\bar{X}_N\). [For the efficiency comparisons of \(\hat{\theta}_{N,R}, \bar{X}_N\) and the estimates based on the medians of the observations, the reader is referred to Bickel (1964).]

From (8.11) it is easy to deduce
\[
\begin{align*}
(8.12) & \quad e_{\hat{\theta}_{N,R}, \bar{X}_N} = \left[ \prod_{j=1}^{p} \frac{1}{\delta^2_j \left( \int_{-\infty}^{\infty} r_j^2(x) \, dx \right)^2} \right]^{p} \left[ \frac{\det \rho_{jk}^*}{\det \rho_{jk}^{**}} \right]^{1/p} \\
& \quad \text{where} \\
& \quad \rho_{jk}^* = \begin{cases} \\
12 \left[ \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} F_j(x) F_k(y) \, dF_{j,k}(x,y) - \frac{1}{4} \right] & \text{if } j \neq k \\
1 & \text{if } j = k.
\end{cases} \\
(8.13) & \quad e_{\hat{\theta}_{N,\bar{X}}, \bar{X}_N} = \left[ \prod_{j=1}^{p} \sigma_j^2 \left( \int_{-\infty}^{\infty} \frac{r_j^2(x) \, dx}{\phi(\Phi^{-1}(F_j(x)))} \right)^2 \right]^{p-1} \left[ \frac{\det \rho_{jk}^*}{\det \rho_{jk}^{**}} \right]^{1/p} \\
& \quad \text{where} \\
& \quad \rho_{jk}^{**} = \begin{cases} \\
\int_{-\infty}^{\infty} \int_{-\infty}^{\infty} \Phi^{-1}[F_j(x)]^{-1} \Phi^{-1}[F_k(y)] \, dF_{j,k}(x,y) & \text{if } j \neq k \\
1 & \text{if } j = k.
\end{cases}
\end{align*}
\]

[When \( p = 1 \), the efficiencies (8.12), (8.14) and (8.16) reduce to the familiar expressions in the one-sample univariate case.] We shall study the above efficiencies in different situations.

**Case 1. Total Symmetry.** Let us assume that the estimators \( \hat{\theta}_{N,\bar{X}}, \hat{\theta}_{N,R} \) and \( \bar{X}_N \) have pairwise independent coordinates. In this situation the covariance matrices reduce to diagonal matrices, and we have
\[
\begin{align*}
(8.17) \quad e_{\hat{\Theta}_{N,R};\bar{X}_N} & = 12 \left[ \prod_{j=1}^{p} \sigma_j^2 \left( \int_{-\infty}^{+\infty} f_j^2(x) \, dx \right)^2 \right]^{1/p} \\
(8.18) \quad e_{\hat{\Theta}_{N,\Phi};\bar{X}_N} & = \left[ \prod_{j=1}^{p} \sigma_j^2 \left( \int_{-\infty}^{+\infty} f_j^2(x) \frac{1}{\phi[\Phi^{-1}(F_j(x))]^2} \right) \right]^{1/p} \\
(8.19) \quad e_{\hat{\Theta}_{N,\Phi};\hat{\Theta}_{N,R}} & = \frac{1}{12} \left[ \prod_{j=1}^{p} \left( \int_{-\infty}^{+\infty} f_j^2(x) \frac{1}{\phi[\Phi^{-1}(F_j(x))]^2} \right) \right]^{1/p}
\end{align*}
\]

Hence from the results of Hodges-Lehmann (1961), Chernoff-Savage (1958) and Mikulski (1963), we have

\[
(8.20) \quad \inf_{F \in \mathcal{F}} e_{\hat{\Theta}_{N,R};\bar{X}_N} = 0.864
\]

\[
(8.21) \quad \inf_{F \in \mathcal{F}} e_{\hat{\Theta}_{N,\Phi};\bar{X}_N} = 1
\]

\[
(8.22) \quad \inf_{F \in \mathcal{F}} e_{\hat{\Theta}_{N,\Phi};\hat{\Theta}_{N,R}} = \frac{\pi}{6}
\]

where \( \mathcal{F} \) is the set of all totally symmetric absolutely continuous \( p \)-variate distributions.

Now suppose that the components of \( F \) are totally symmetric as well as identically distributed, then the efficiencies are independent of \( p \), and hence the same as in the case of one-sample univariate situations.
Case 2. Equally Correlated Distributions.

Let us now assume that the distribution of \((X^{(j)}_\alpha, X^{(k)}_\alpha)\) is independent of \(j\) and \(k\) when \(\alpha = 0\). Then, it turns out that

$$
\begin{align*}
\hat{\theta}_{N,R}^2 &= 12\sigma_1^2 \left( \int_{-\infty}^{+\infty} f_1^2(x) dx \right)^2 \left[ \frac{1+(p-1)\rho_{12}}{1+(p-1)\rho_{12}} \right]^{1/p} \left[ \frac{1-\rho_{12}^*}{1-\rho_{12}} \right]^{1-1/p} \\
\hat{\theta}_{N,F;X_N}^2 &= \sigma_1^2 \left( \int_{-\infty}^{+\infty} \frac{f_1^2(x) dx}{\phi[\Phi^{-1} F_1(x)]} \right)^2 \left[ \frac{1+(p-1)\rho_{12}^*}{1+(p-1)\rho_{12}^*} \right]^{1/p} \left[ \frac{1-\rho_{12}^*}{1-\rho_{12}^*} \right]^{1-1/p} \\
\hat{\theta}_{N,F;\hat{\theta}_{N,R}}^2 &= \left( \int_{-\infty}^{+\infty} \frac{f_1^2(x) dx}{\phi[\Phi^{-1} F_1(x)]} \right)^2 \left[ \frac{1+(p-1)\rho_{12}^*}{1+(p-1)\rho_{12}^*} \right]^{1/p} \left[ \frac{1-\rho_{12}^*}{1-\rho_{12}^*} \right]^{1-1/p}
\end{align*}
$$

Case 3. Normal Case.

Let the underlying distribution function \(F\) be a nonsingular \(p\)-variate normal distribution function with mean vector zero, and covariance matrix \(-\gamma = (\rho_{jk}\sigma_j\sigma_k)\) then, from (8.12), (8.14) and (8.16) we obtain

$$
\hat{\theta}_{N,R}^2 = \frac{3}{\pi} \left[ \frac{\det ||\rho^*_jk||}{\det ||\rho_{jk}||} \right]^p
$$

where

$$
\rho^*_{jk} = \int_{-\infty}^{+\infty} \int_{-\infty}^{+\infty} \Phi_j \left[ \frac{x}{\sigma_j} \right] \Phi_k \left[ \frac{y}{\sigma_k} \right] d\Phi_j, k(x,y)
$$

\[ (8.28) \quad \hat{\Theta}_{N,\Phi; \hat{X}_N} = 1 \]

\[ (8.29) \quad \hat{\Theta}_{N,\Phi; \hat{\Theta}_{N,R}} = \frac{\pi}{3} \left[ \frac{\text{det}\|\rho_{jk}^*\|}{\text{det}\|\rho_{jk}\|} \right]^{1/p}, \]

where \( \rho_{jk}^* \) is defined by (8.27).

Now Bickel (1964) has proved that for \( p > 2 \),

\[ (8.30) \quad \inf_{F \in \Phi} \hat{e}_{N,R; \hat{X}_N} = 0 \]

where \( \Phi \) is the set of all nonsingular \( p \)-variate normal distributions and, since

\[ (8.31) \quad \inf_{F \in \Phi} \hat{e}_{N,\Phi; \hat{X}_N} = 1 \quad \text{for all} \quad p \]

it follows that

\[ (8.32) \quad \sup_{F \in \Phi} \hat{e}_{N,\Phi; \hat{\Theta}_{N,R}} = \infty. \]

Thus, we find that when the underlying distribution function is normal, the multivariate normal scores estimator \( \hat{\Theta}_{N,\Phi} \) (which is as good as the means estimator) can be infinitely better than the multivariate Rank-sum estimator \( \hat{\Theta}_{N,R} \) for \( p > 2 \). This leads to the question of examining the relative performance of these two estimators viz. \( \hat{\Theta}_{N,\Phi} \) and \( \hat{\Theta}_{N,R} \) for the case when the underlying distribution function \( F \) is bivariate normal \( N(\underline{0}, \sigma_1^2, \sigma_2^2, \rho) \). The efficiency behavior of \( \hat{\Theta}_{N,\Phi} \) and \( \hat{\Theta}_{N,R} \) is given by the following
Theorem 8.7. The efficiency of $\hat{\vartheta}_{N,R}$ with respect to $\hat{\vartheta}_{N,\overline{X}}$ is independent of $\sigma_1$ and $\sigma_2$ and is given by

$$e_{\hat{\vartheta}_{N,R}; \hat{\vartheta}_{N,\overline{X}}} = \frac{3}{\pi} \left[ \frac{1 - \rho^2}{1 - g(1 - \frac{2}{\pi} \cos^{-1}\frac{\rho}{2})^2} \right]^{1/2}$$

$$= \frac{3}{2} \left[ - \frac{(1 + 2 \cos \frac{2}{3} (\pi + u))}{u(\pi - u)} \right]^{1/2}$$

where $u$ is determined by $\rho = 2 \cos (u \pi)/3$. The function $e_{\hat{\vartheta}_{N,R}; \hat{\vartheta}_{N,\overline{X}}}$ is monotone increasing for $0 \leq \rho < 1$ and symmetric about $\rho = 0$ and hence unimodal. Finally,

$$\lim_{|\rho| \to 1} e_{\hat{\vartheta}_{N,R}; \hat{\vartheta}_{N,\overline{X}}} = \left( \frac{3}{\pi \sin \frac{\pi}{3}} \right)^{1/2} = 0.91.$$
References


Asymptotic Normality of a Certain Non-parametric Test

Princeton University Press.

Asymptotic Efficiency for One-Sample Non-parametric
University of Minnesota, Dept. of Statistics.

Generalization of Chernoff-Savage Theorems on Asymptotic
Normality of Non-Parametric Test Statistics.

227-233.

Math. Statist. 26 523-527.

Comparison of Normal Scores and Wilcoxon Tests.

Estimates of Location Based on Rank Tests.

Bahadur Efficiency of the Hodges Bivariate Sign Test.


