JOURNAL OF MULTIVARIATE ANALYSIS 2, 219–229 (1972)

Nonparametric Tests for Ordered Alternatives in the Bivariate Case*

RICHARD A. JOHNSON

University of Wisconsin, Madison, Wisconsin

AND

KISHAN G. MEHROTRA

Syracuse University, Syracuse, New York

Communicated by P. K. Sen

A class of bivariate rank tests are developed for the two-sample problem of testing equality of distributions against certain one-sided alternatives. These tests are the nonparametric analogs of a normal theory test proposed by Schaafsma and Smid. These tests are shown to be unbiased and the asymptotic distributions are obtained under the null distribution and local alternatives. Some asymptotic efficiency comparisons are also made.

INTRODUCTION

Let $\mathbf{Z} = (X, Y)$ has a continuous *df* $F(x, y)$ and $\mathbf{Z}' = (X', Y')$ has a continuous *df* $G(x, y)$. Suppose m independent copies $\mathbf{Z}_1, ..., \mathbf{Z}_m$ of $\mathbf{Z}$ and n independent copies $\mathbf{Z}_1', ..., \mathbf{Z}_n'$ of $\mathbf{Z}'$ are given. In Section 2 of the present paper, nonparametric tests are developed for testing the null hypothesis $F = G$ against a class of one-sided alternatives defined later in this paper. These tests are nonparametric analogs of the normal theory test of Schaafsma and Smid [8]. The special case of rank scores relates to their test in the same manner that the rank test of Chatterjee and Sen [3] corresponds to T^2 with multisided alternatives. Consequently, the distribution theory in [5], which generalizes [3], plays an important role.

Received June 11, 1971; revised November 29, 1971.

AMS 1970 subject classification: Primary 62H15; Secondary 62G10.

Key words and phrases: Multivariate; nonparametric tests; ordered alternatives.

* This research was supported by the Air Force Office of Scientific Research under Grant AFOSR-69-1789.

From a Pitman relative efficiency viewpoint the rank score test is compared with the layer rank test proposed by Bhattacharyya and Johnson [1] and the parametric test obtained by Schaafsma and Smid [8] for bivariate normal populations.

1. An Invariance Argument

Suppose $\mathbf{Z}$ and $\mathbf{Z}'$ are as defined earlier. Set

$$\bar{F}(x, y) = 1 - F(x, \infty) - F(\infty, y) + F(x, y)$$

and a similar expression for $\bar{G}(x, y)$. We are interested in the problem of testing $H_0 : F = G$ against the class of ordered alternatives $H_1 : \{(F, G) : F \neq G, F(x, y) \geqslant G(x, y), \bar{F}(x, y) \leqslant \bar{G}(x, y)$ for all $(x, y)\}$, which was considered by Bhattacharyya and Johnson [1]. It is also shown in their paper that the class $\{(F, G): G(x, y) = F(x - \theta_1, y - \theta_2), \boldsymbol{\theta} \geqslant \mathbf{0}, \boldsymbol{\theta} \neq \mathbf{0}\}$ is a subset of the above class H_1.

Suppose $\mathscr{G}$ is a group of transformations g such that $g(\mathbf{Z}) = \{g_1(x), g_2(y)\}$, where g_1 and g_2 are continuous and strictly increasing. Suppose $\mathscr{H}$ is the group consisting of the identity transformation and h, where $h(\mathbf{Z}) = h(x, y) = (y, x)$. Let $\mathscr{G}^*$ be the smallest group containing elements generated by two subgroups $\mathscr{G}$ and $\mathscr{H}$.

Lemma 1.[1] *The problem of testing the hypothesis $F = G$ against $(F, G) \in H_1$ on the basis of random samples $\mathbf{Z}_1, \ldots, \mathbf{Z}_m$ from F and $\mathbf{Z}_{m+1}, \ldots, \mathbf{Z}_N$, $N = m + n$, from G remains invariant with respect to the group $\mathscr{G}^*$, and the maximal invariant is given by the set C^* containing vectors of the coordinatewise ranks.*

Proof. Lemma 2.1 of Bhattacharyya and Johnson [1] shows that the problem remains invariant. To obtain the maximal invariant, we denote the observations by $\mathbb{Z} = (\mathbf{Z}_1, \ldots, \mathbf{Z}_N)$ and set $\mathbf{R}_1 = (R_{11}, \ldots, R_{1N})$, where R_{1j} is the rank of X_j among $X_1, \ldots, X_N$ and, similarly, $\mathbf{R}_2 = (R_{21}, \ldots, R_{2N})$, where R_{2j} is the rank of Y_j among $Y_1, \ldots, Y_N$. First consider $\mathscr{G}$. Clearly, $\binom{\mathbf{R}_1}{\mathbf{R}_2}$ is invariant under $\mathscr{G}$ which consist of coordinatewise monotone transformations. Further, the method of Example 3 in Chapter 6 of Lehmann [4] applied to each coordinate separately, shows that $\binom{\mathbf{R}_1}{\mathbf{R}_2}$ is a maximal invariant. Identifying $\mathscr{G}$ with D in Theorem 2, Chapter 6 of Lehmann [4], the proof of the lemma will be complete if we are able to show (1) that the rank matrix of $h(\mathbf{z}_1), \ldots, h(\mathbf{z}_N)$ is the same as that of $h(\mathbf{z}_1'), \ldots, h(\mathbf{z}_N')$ whenever the rank matrix of $\mathbf{z}_1, \ldots, \mathbf{z}_N$ is the same as that of $\mathbf{z}_1', \ldots, \mathbf{z}_N'$; and (2) that the subgroup $\mathscr{H}$, acting on the rank matrix, gives the set of row vectors as a maximal invariant.

[1] The group $\mathscr{G}^*$ was suggested by Dr. James Bondar.

Under the group of transformations $\mathscr{H}$ the rows of the matrix $\mathbf{R}$ are permuted and the collection of row vectors of $\mathbf{R}$, namely $C^* = \{(R_{11},..., R_{1N}); (R_{21},..., R_{2N})\}$, is clearly invariant. That this is maximal invariant can be seen as follows. Suppose two such collections C_1^* and C_2^* are obtained from matrices $\mathbf{R}^1$ and $\mathbf{R}^2$, respectively, and suppose C_1^* and C_2^* are equal, i.e., $C_1^* \equiv \{\mathbf{R}_1^1, \mathbf{R}_2^1\} = \{\mathbf{R}_1^2, \mathbf{R}_2^2\} \equiv C_2^*$ This implies that either

(i) $\quad \mathbf{R}_1^1 = \mathbf{R}_1^2$ and $\mathbf{R}_2^1 = \mathbf{R}_2^2$

or

(ii) $\quad \mathbf{R}_1^1 = \mathbf{R}_2^2$ and $\mathbf{R}_2^1 = \mathbf{R}_1^2$

However, in the case of (i) $\mathbf{R}^1 = \mathbf{R}^2$ and in the case (ii) $\mathbf{R}^1 = h(\mathbf{R}^2)$ Thus, by definition, C^* is maximal invariant.

Hence, under $\mathscr{G}^*$, using Theorem 2, Chapter 6, of Lehmann [4], the maximal invariant is given by C^*.

Remark. The above lemma also applies to the case of alternatives $F \neq G$. Thus in the two-sample bivariate case, the rank test proposed by Chatterjee and Sen [3] as well as rank tests based on other scores are a function of the maximal invariant under $\mathscr{G}^*$.

2. A Permutation Test

Assume that the combined sample is ranked coordinatewise, and the rank matrix is denoted by $\mathbf{R}_N$ and $\mathbf{R}_N^*$ denotes the matrix obtained by permuting the columns of $\mathbf{R}_N$ so that the first row is $(1, 2,..., n)$.

Further, following [5, p. 184], we set $\mathbf{E}_N$ equal to the matrix of scores

$$\mathbf{E}_N = \left[\begin{array}{ccc|ccc} E_{N,R_{11}} & \cdots & E_{N,R_{1m}} & E_{N,R_{1m+1}} & \cdots & E_{N,R_{1N}} \\ E_{N,R_{21}} & \cdots & E_{N,R_{2m}} & E_{N,R_{1m+1}} & \cdots & E_{N,R_{2N}} \end{array}\right], \tag{2.1}$$

where the scores $E_{N,\alpha} = J_N(\alpha/(N+1))$, $1 \leqslant \alpha \leqslant N$, with J_N satisfying the Chernoff–Savage conditions (see [5, p. 95]) and $E_{N\alpha_1} \leqslant E_{N\alpha_2}$ for $\alpha_1 < \alpha_2$.

We define a class of test statistics of the form

$$T_{J_N} = T_N^1 + T_N^2, \tag{2.2}$$

where

$$T_N^k = \frac{1}{m} \sum_{j=1}^{N} E_{N,R_{kj}}\,; \qquad k = 1, 2.$$

This definition is motivated from asymptotic considerations since T_{J_N} is the rank version of the normal theory test of Schaafsma and Smid [8] which is most stringent among tests which are somewhere most powerful.

For the null hypothesis, the conditional moments are given in [5, p. 185] as

$$\begin{aligned} E_{H_0}(T_N{}^k \mid P_N) &= \frac{1}{N} \sum_{\alpha=1}^{N} E_{N,\alpha} = \bar{E}_N, \qquad k = 1, 2, \\ \operatorname{cov}_{H_0}(T_N{}^k, T_N^{k'} \mid P_N) &= (N\delta_{kk'} - m)\, v_{ij}(\mathbf{R}_N{}^*)/m(N-1), \qquad k, k' = 1, 2, \end{aligned} \tag{2.3}$$

where $\delta_{kk'}$ is the Kronecker delta,

$$v_{ij}(\mathbf{R}_N{}^*) = \frac{1}{N} \sum_{\alpha=1}^{N} E_{N,R_{i\alpha}} E_{N,R_{j\alpha}} - \bar{E}_N{}^2. \tag{2.4}$$

In the special case where $E_{N\alpha} = \alpha/(N+1)$, the covariance term involves Spearman's rank correlation. Another case of interest is the normal scores test.

The exact conditional distribution of T_{J_N}, given $\mathbf{R}_N{}^*$, can be obtained by permuting the columns of $\mathbf{R}_N{}^*$ and calculating the values of the test statistics. The permutation test can be characterized by the critical function ϕ

$$\phi(R_N) = 1, \qquad \gamma_N \text{ or } 0 \text{ as } T_{J_N} <, = \text{ or } > T_{J_N}(R_N{}^*), \tag{2.5}$$

where γ_N is selected to make the conditional level α. Thus the unconditional level of significance is also α.

3. Unbiasedness

In this section, we prove that a permutation test defined by (2.5), is unbiased for a certain subclass of the alternatives K which we denote by K^*. The class K^* contains pairs $(F, G) \in K$ such that if F is the *df* of (X, Y) then G is the *df* of $[h_1(X, Y), h_2(X, Y)]$, where h_1 and h_2 are continuous maps $R^2 \to R$ satisfying

$$h_1(x, y) \geqslant x, \qquad h_2(x, y) \geqslant y, \tag{3.1}$$

and

$$(x_1, y_1) \geqslant (x_2, y_2) \Rightarrow h_i(x_1, y_1) \geqslant h_i(x_2, y_2) \qquad i = 1, 2. \tag{3.2}$$

A strict inequality for the pairs (x, y) implies the same for $h_i(x, y)$'s. The subclass K^* has been studied in detail by Bhattacharyya and Johnson [1].

THEOREM 3.1. *The permutation test given by* (2.5) *is unbiased for all alternatives* $(F, G) \in K^*$.

Proof. Let $\mathscr{C}$ be the class of maps of R^{2N} into itself such that, for a given set of observations $\mathbf{Z}_1, \mathbf{Z}_2, ..., \mathbf{Z}_N$, (all of which have $df F$), $C(\mathbf{Z}_1, ..., \mathbf{Z}_N) = (\mathbf{Z}_1, ..., \mathbf{Z}_m; \mathbf{Z}'_{m+1}, ..., \mathbf{Z}_N')$, where $\mathbf{Z}_i' = [h_1(\mathbf{Z}_i), h_2(\mathbf{Z}_i)]$ $i = m + 1, ..., N$ and h_1, h_2 satisfy (3.1) and (3.2). If $\mathbf{R}$ and $\mathbf{R}'$ denote the rank matrices of $\mathbb{Z} = (\mathbf{Z}_1, ..., \mathbf{Z}_N)$ and $\mathbb{Z}' = C(\mathbf{Z}_1, ..., \mathbf{Z}_N)$, respectively, the relations (3.1) and (3.2) imply that

$$\begin{aligned} R'_{ij} &\leqslant R_{ij} \qquad j = 1, 2, ..., m, \qquad i = 1, 2, \\ &\geqslant R_{ij} \qquad j = m + 1 \cdots N, \qquad i = 1, 2. \end{aligned} \tag{3.3}$$

From a result of Lehmann [4] (cf. Problem 30, p. 256), it is then sufficient to prove that $\phi(\mathbb{Z}) \leqslant \phi(\mathbb{Z}')$ for all $\mathbb{Z} \in R^{2N}$ and $C \in \mathscr{C}$.

In the permutation test defined in (2.5), the $N!$ possible values are obtained by permuting the columns of the matrix $\mathbf{R}$. Of these values, only $\binom{N}{m}$ are possibly distinct and each distinct value occurs with frequency $N!/\binom{N}{m}$. Let us denote this collection of possibly distinct values of the statistic by $S(\mathbb{Z})$ and $S(\mathbb{Z}')$, respectively, for $\mathbb{Z}$ and $\mathbb{Z}'$. Now the matrices $\mathbf{R}$ and $\mathbf{R}'$ associated with $\mathbb{Z}$ and $\mathbb{Z}'$ are

$$\mathbf{R} = \left[\begin{array}{ccc|c} R_{11} & R_{12} \cdots R_{1m} & & R_{1m+1} \cdots R_{1N} \\ R_{21} & R_{22} \cdots R_{2m} & & R_{2m+1} \cdots R_{2N} \end{array}\right] = [\mathbf{R}_{\mathrm{I}} \mid \mathbf{R}_{\mathrm{II}}]$$

and

$$\mathbf{R}' = \left[\begin{array}{ccc|c} R'_{11} & R'_{12} \cdots R'_{2m} & & R'_{2m+1} \cdots R'_{1N} \\ R'_{21} & R'_{22} \cdots R'_{2m} & & R'_{2m+1} \cdots R'_{2N} \end{array}\right] = [\mathbf{R}_{\mathrm{I}}' \mid \mathbf{R}'_{\mathrm{II}}].$$

In order to obtain a member of $S(\mathbb{Z})$, other than the statistic obtained for the particular sample, k columns of $\mathbf{R}_{\mathrm{I}}$ are exchanged with an equal number of columns of $\mathbf{R}_{\mathrm{II}}$, $0 < k \leqslant \min(m, n)$. Denote the statistic obtained after this operation by $T_{J_N,k}$. If exactly the same operation is performed on columns of $\mathbf{R}_1'$ and $\mathbf{R}_2'$, and if the statistic obtained is denoted by $T'_{J_N,k}$, then from (3.3) and the fact that the J_N are nondecreasing we obtain that

$$T'_{J_N,k} \geqslant T_{J_N,k}. \tag{3.4}$$

The above is true for all values of k and an arbitrary choice of the k columns. It also follows from (3.3) again that when there is no exchange of columns the statistic obtained from $\mathbf{R}$ is larger than the statistic obtained from $\mathbf{R}'$. This implies that the relative position of statistic obtained from $\mathbf{R}'$ within $s(\mathbb{Z}')$ can only move to the left compared to the position of statistic obtained from $\mathbf{R}$ within $S(\mathbb{Z})$ which in turn implies that increases when going the value of ϕ from $\mathbb{Z}$ to $\mathbb{Z}'$. This completes the proof.

4. Asymptotic Normality and Efficiency

Application of a permutation test is difficult in the case of large sample sizes. An approximation is desirable for such cases. To study the permutation test defined in Section 2, one can follow the same steps as given in Puri and Sen [5]. In Theorem 4.1 below a result is given which is used in the calculation of asymptotic Pitman efficiency.

We consider translation alternatives of the form

$$G(x, y) = f(x - \delta_1 N^{-1/2}, y - \delta_2 N^{-1/2}) \qquad \delta_i \geqslant 0, \quad i = 1, 2, \tag{4.1}$$

where $F(x, y)$ is an absolutely continuous *cdf* with marginal distributions F_1 and F_2. Similar results hold for changes in scale.

THEOREM 4.1. *Let* $(m/N) \to \lambda$, $0 < \lambda < 1$ *and let* $\mathbf{v}(\mathbf{R}_N{}^*) = (v_{ij}(\mathbf{R}_N{}^*))$ *be defined by* (2.4) *Then*

(i) $\mathbf{v}(\mathbf{R}_N{}^*) \xrightarrow{P} \mathbf{v}(F)$,

(ii) $\mathscr{L}[N^{1/2}(T_N{}^1, T_N{}^2)] \Rightarrow \Phi[(\delta_1\eta_1, \delta_2\eta_2), \Sigma]$, (4.2)

where, under the alternatives defined by (4.1), $\Sigma = (\sigma_{ij})$ *with* $\sigma_{11}^2 = \sigma_{22}^2 = (1 - \lambda)\lambda^{-1}v_{11}(F)$, $\sigma_{12} = (1 - \lambda)\lambda^{-1}v_{12}(F)$. *Here* $v_{11}(F) = \int_0^1 J^2(u)\,du - \mu^2$, $\mu = \int_0^1 J(u)\,du$ *and* $v_{12}(F) = \iint J(F_1)\,J(F_2)\,dF - \mu^2$. *Also*

$$\eta_i = (1 - \lambda) \int_{-\infty}^{\infty} J'(F_i)\,dF_i, \qquad i = 1, 2.$$

Thus

(iii) $\mathscr{L}[N^{1/2}T_{J_N}] \Rightarrow \Phi\{\delta_1\eta_1 + \delta_2\eta_2, 2\lambda^{-1}(1 - \lambda)[v_{11}(F) + v_{12}(F)]\}$. (4.3)

Proof. The proof of (i) and (ii) is contained in Puri and Sen [5], mainly Theorem 5.6.1 and pp. 204–205, and (iii) follows from the continuity of $T_{J_N} = T_N{}^1 + T_N{}^2$.

Remark. From (4.3) we observe that the null asymptotic distribution of T_{J_N} is not distribution-free, because it contains $v_{12}(F)$. However, from (i) of the above theorem we observe that $v_{12}(\mathbf{R}_N{}^*)$ is a consistent estimate of $v_{12}(F)$. Hence an asymptotically distribution-free statistic can be defined using this estimate to standardize T_{J_N}.

Employing Theorem 4.1, we find that the efficacy of T_{J_N} is

$$\frac{\lambda(1 - \lambda)[\delta_1 B_1(F_1) + \delta_2 B_2(F_2)]^2}{2v_{11}(F)[1 + \rho_J]}, \tag{4.4}$$

where $\rho_J = v_{12}(F)/v_{11}(F)$ and $B_i(F_i) = \int_{-\infty}^{\infty} J'(F_i)\, dF_i$, $i = 1, 2$ when the same score function is used for both variates. From (4.4) we are able to obtain the ARE of two tests of the form (25) based on different score functions. This will, in general, depend on δ_1 and δ_2.

In particular, if we denote the test based on ranks by U_N and let ρ_F denote the grade correlation, the ARE of T_{N_J} with respect to U_N is

$$\frac{(1+\rho_F)}{12v_{11}(1+\rho_J)} \left\{ \frac{\delta_1 B(F_1) + \delta_2 B(F_2)}{\delta_1 \int f_1^2 + \delta_2 \int f_2^2} \right\}^2 = e(F, \delta_1, \delta_2), \tag{4.5}$$

which depends on δ_1 and δ_2. However, writing

$$e_i = B^2(F_i) \Big/ \left\{ 12v_{11} \left(\int f_i^2 \right)^2 \right\}, \qquad i = 1, 2,$$

we have

$$e(F, \delta_1, \delta_2) = \left(\frac{1+\rho_F}{1+\rho_J} \right) \left\{ \frac{\delta_1(\int f_1^2)\sqrt{e_1} + \delta_2(\int f_2^2)\sqrt{e_2}}{\delta_1(\int f_1^2) + \delta_2(\int f_2^2)} \right\}^2,$$

$$\left(\frac{1+\rho_F}{1+\rho_J} \right) \min(e_1, e_2) \leqslant e(F, \delta_1, \delta_2) \leqslant \left(\frac{1+\rho_F}{1+\rho_J} \right) \max(e_1, e_2). \tag{4.6}$$

One could use (4.6) to investigate bounds on efficiency using the known univariate results regarding the e_i.

In the rest of this section, we concentrate on the test U_N.

a. *Efficiency of U_N Versus the Layer Rank Statistic*

Bhattacharyya and Johnson [1] obtained the efficacy of the layer rank statistic S_N which is given by expression (4.7).

$$\frac{\lambda(1-\lambda)}{\sigma_F^2} \left[\delta_1 \iint \frac{\partial F(x, y)}{\partial x}\, dF(x, y) + \delta_2 \iint \frac{\partial F(x, y)}{\partial y}\, dF(x, y) \right]^2, \tag{4.7}$$

where

$$\sigma_F^2 = \iint F^2\, dF - \left(\iint F\, dF \right)^2.$$

Therefore, the efficiency of the layer rank test with respect to U_N is given by (4.8).

$$\mathscr{E}_{S_N:U_N} = \frac{1+\rho_F}{6\sigma_F^2} \left[\frac{\delta_1 \iint [\partial F(x, y)/\partial x]\, dF(x, y) + \delta_2 \iint [\partial F(x, y)/\partial y]\, dF(x, y)}{\delta_1 \int f_1^2(x)\, dx + \delta_2 \int f_2^2(x)\, dx} \right]^2. \tag{4.8}$$

LEMMA 4.1. *If $F(x, y)$ has continuous density $f(x, y)$ then*

$$\iint \frac{\partial F(x, y)}{\partial x} dF(x, y) = 1/2 \int f_1^2(x)\, dx. \tag{4.9}$$

Proof. By the definition of conditional distribution, for every x,

$$F(x, y) = \int_{-\infty}^{x} F(y \mid u) f_1(u)\, du,$$

where $F(y \mid u)$ denotes the conditional distribution of Y given $x = u$. Thus we obtain

$$\partial F(x, y)/\partial x = F(y \mid x) f_1(x) \quad \text{a.e.}$$

Therefore

$$\begin{aligned}\iint \frac{\partial F(s, y)}{\partial x} dF(x, y) &= \iint F(y \mid x) f_1(x) f(x, y)\, dx\, dy \\ &= \int \left[\int F(y \mid x) f(y \mid x)\, dy\right] f_1^2(x)\, dx \\ &= 1/2 \int f_1^2(x)\, dx,\end{aligned}$$

because $\int F(y \mid x) f(y \mid x)\, dy = 1/2$ a.e. being the expected value of a continuous distribution function.

Using (4.9), (4.8) reduces to

$$\mathscr{E}_{S_N : U_N} = (1 + \rho_F)/24\sigma_F^2 \tag{4.10}$$

which shows that the above efficiency is independent of δ_i $i = 1, 2$. In the case X and Y are independent, this efficiency is equal to 6/7. Thus the performance of U_N is slightly better than the layer rank statistics in the asymptotic sense.

In general, unlike the univariate case, the exact evaluation of σ_F^2 is difficult in the bivariate situation. In the following we will obtain an upper bound on σ_F^2. Suppose that the density $f(x, y)$ exists. Further if $F(x, y)$ is a *df*, the function $H(x, y) = F^n(x, y)$ is also a distribution function. Thus for an integer n,

$$1 = \iint dH(x, y) = \iint dF^n(x, y).$$

This gives us

$$1 = \iint n(n-1) F^{n-2} \left[\frac{\partial}{\partial x} F(x, y)\right]\left[\frac{\partial}{\partial y} F(x, y)\right] dx\, dy$$

$$+ \iint nF^{n-1}(x, y) f(x, y)\, dx\, dy$$

$$= n(n-1) \iint F^{n-2} \left(\frac{\partial}{\partial x} F(x, y)\right)\left(\frac{\partial}{\partial y} F(x, y)\right) dx\, dy$$

$$+ nE_F(F^{n-1}(x, y)).$$

Consequently

$$E_F(F^{n-1}) = \frac{1}{n} - (n-1) \iint F^{n-2} \left(\frac{\partial}{\partial x} F(x, y)\right)\left(\frac{\partial}{\partial y} F(x, y)\right) dx\, dy. \quad (4.11)$$

In particular

$$E_F(F) = 1/2 - \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} \left(\frac{\partial}{\partial x} F(x, y)\right)\left(\frac{\partial}{\partial y} F(x, y)\right) dx\, dy$$

$$= 1/2 - \alpha, \quad (4.12)$$

where

$$\alpha = \iint \left[\frac{\partial}{\partial x} F(x, y) \cdot \frac{\partial}{\partial y} F(x, y)\right] dx\, dy$$

and

$$E(F^2) = 1/3 - 2 \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} F(x, y) \left[\frac{\partial}{\partial x} F(x, y)\right]\left[\frac{\partial}{\partial y} F(x, y)\right] dx\, dy. \quad (4.13)$$

But $0 \leqslant F(x, y) \leqslant F_1(x)$ implies $0 < E_F(F(x, y)) \leqslant E_F(F_1(x)) = E_{F_1}(F_1(x)) = 1/2$. As a consequence of this result $0 \leqslant \alpha < 1/2$. Further the equality $F(x, y) = F_1(x)$ holds for all y only in the degenerate case when Y is a monotone increasing function of X; and then $\alpha = 0$. In this particular case, from (4.13), $E(F^2) = 1/3$ and therefore $\mathrm{var}(F) = 1/12$. The asymptotic relative efficiency $\mathscr{E}_{S_N:U_N}$ is simply 1. In the other extreme case when Y is a monotone decreasing function of X, $\rho_F = -1$ and $\sigma_F{}^2 = 0$ so that (4.10) does not apply directly.

A general upper bound on $\sigma_F{}^2$ is obtained by making use of the fact that

$$E[F^2(x, y)] \leqslant E[F_1(x) \cdot F_2(y)]$$

so that $E(F^2) \leqslant (\rho_F/12) + (1/4)$. Therefore $\sigma_F{}^2 \leqslant (\rho_F/12) + \alpha(1-\alpha) \leqslant (\rho_F/12) + (1/4)$. Thus

$$\mathscr{E}_{S_N:U_N} = \frac{1 + \rho_F}{24\sigma_F{}^2} \geqslant \frac{1 + \rho_F}{6 + 2\rho_F}. \quad (4.14)$$

Substituting $\rho_F = -1$ in the above expression we observe that $\mathscr{E}_{S_N:U_N} \geqslant 0$ and is really not indeterminant.

b. *Efficiency of U_N Versus a Parametric Test*

It would be difficult to compare U_N with the LR statistic, from an asymptotic relative efficiency viewpoint, because the LR statistic has a nonnormal asymptotic distribution, see Mehrotra [7].

The somewhere most powerful, most stringent test of Schaafsma and Smid [8], for the two-sample case, would be based on $V_N = (\bar{X}_2 - \bar{X}_1 + \bar{Y}_2 - \bar{Y}_1)$. We compare U_N with this statistic. Note that U_N is the rank version of the statistic V_N.

Let the covariance matrix corresponding to the *df* $F(x, y)$ exist and be denoted by

$$\Lambda = \begin{pmatrix} \Lambda_{11} & \Lambda_{12} \\ \Lambda_{12} & \Lambda_{22} \end{pmatrix}, \qquad \Lambda > 0.$$

By the multivariate central limit theorem, the asymptotic distribution of V_N, under the sequence of alternatives (4.1), is given by

$$\mathscr{L}(\sqrt{N}\, V_N) \to \Phi \left[(\delta_1 + \delta_2), \frac{1}{\lambda(1-\lambda)} (\Lambda_{11} + \Lambda_{22} + 2\Lambda_{12}) \right].$$

As a consequence, the efficiency of V_N with respect to U_N has the expression

$$\mathscr{E}_{V_N:U_N} = \frac{2(1 + \rho_F)(\delta_1 + \delta_2)^2}{12\{\delta_1 \int f_1^2(x)\, dx + \delta_2 \int f_2^2(x)\, dx\}^2 (\Lambda_{11} + \Lambda_{21} + 2\Lambda_{12})}. \tag{4.15}$$

In case $F(x, y)$ is symmetric in (x, y), the above efficiency simplifies to

$$(1 + \rho_F) \Big/ \left[12 \left\{ \int f_1^2(x)\, dx \right\}^2 (\Lambda_{11} + \Lambda_{12}) \right] \tag{4.16}$$

which is independent of (δ_1, δ_2). Finally, for the bivariate normal whose covariance matrix is $\begin{pmatrix} 1 & \tau \\ \tau & 1 \end{pmatrix}$, (4.16) becomes

$$\{1 + (6/\pi) \sin^{-1}(\tau/2)\} / \{(3/\pi)(1 + \tau)\}.$$

The expressions (4.15) and (4.16) were obtained independently by Fryer [6].

When $\tau = 0$ and $\tau = 1$, this reduces to $\pi/3$, whereas for $\tau = -1$, after using the L'Hopital rule, this gives $2/\sqrt{3}$. From Bickel [2] we observe that the efficiency of the Hotelling T^2 compared to the multisided test of Chatterjee and Sen [3] is also $\pi/3$, when the underlying distribution is an independent bivariate normal.

It is clear that one expects the correspondence between the efficiency results for one-sided and two-sided tests to be preserved. The normal score test would have higher efficiency than the normal theory test for nonnormal distributions.

Acknowledgment

The authors would like to thank the referee for helpful suggestions leading to an improvement of both the results and the presentation.

References

[1] Bhattacharyya, G. K. and Johnson, R. A. (1970). A layer rank test for ordered bivariate alternatives. *Ann. Math. Statist.* **41** 1296–1309.

[2] Bickel, P. J. (1965). On some asymptotically nonparametric competitors of Hotelling's T^2. *Ann. Math. Statist.* **36** 160–173.

[3] Chatterjee, S. K. and Sen, P. K. (1964). Non-parametric tests for the bivariate two-sample location problem. *Calcutta Statist. Assoc. Bull.* **13** 18–57.

[4] Lehmann, E. L. (1959). *Testing Statistical Hypotheses.* John Wiley and Sons, New York.

[5] Puri, M. L. and Sen, P. K. (1971). *Nonparametric Methods in Multivariate Analysis.* John Wiley and Sons, New York.

[6] Fryer, J. G. (1970). On the nonparametric tests of David and Fix for the bivariate two-sample location problem. *J. Amer. Statist. Assoc.* **65** 1297–1307.

[7] Mehrotra, K. G. (1970). A non-parametric approach to the problem of ordered alternatives with emphasis on locally most powerful rank tests for the two-sample problem with censoring. Doctoral Dissertation, University of Wisconsin, Madison, Wisconsin.

[8] Schaafsma, W. and Smid, L. J. (1966). Most stringent somewhere most powerful tests against alternatives restricted by a number of linear inequalities. *Ann. Math. Statist.* **37** 1161–1172.