JOURNAL OF MULTIVARIATE ANALYSIS 12, 219–229 (1982)

Multivariate Linear Rank Statistics for Profile Analysis*

VERNON M. CHINCHILLI AND PRANAB KUMAR SEN

Medical College of Virginia and University of North-Carolina

Communicated by P. R. Krishnaiah

For some general multivariate linear models, linear rank statistics are used in conjunction with Roy's Union–Intersection Principle to develop some tests for inference on the parameter (vector) when they are subject to certain linear constraints. More powerful tests are designed by incorporating the a priori information on these constraints. Profile analysis is an important application of this type of hypothesis testing problem; it consists of a set of hypothesis testing problem for the p responses q-sample model, where it is a priori assumed that the response–sample interactions are null.

1. INTRODUCTION

In multivariate nonparametric hypothesis testing theory, the role of Roy's [14] Union–Intersection (UI-) principle has not yet been examined fully. Chatterjee and De [5] considered some UI-rank tests for a bivariate, two-sample location problem with an orthant restriction. The current authors [7, 8] generalized this UI technique to develop asymptotically distribution-free (ADF) tests for a broad class of restricted alternative problems in multivariate analysis, with special emphasis on the orthant restriction problem in the multivariate case. The object of the present investigation is to examine another application of the theory developed in [7], namely, the *linear equality restriction* and *profile analysis*. Some preliminary results on multivariate linear rank statistics are introduced in Section 2 and these are then incorporated in Section 3 in the formulation of a class of UI-rank tests for the linear equality restriction problem. As a special case, the profile analysis problem is treated in Section 4. ADF tests for the profile analysis

Received March 12, 1980; revised September 17, 1981.

AMS subject classifications: 62G10, 62H15, 90C30.

Key words and phrases: General multivariate linear model, linear euality restriction, linear rank statistics, profile analysis, Union–Intersection principle.

* Work partially supported by the National Heart, Lung and Blood Institute, Contract NIH-NHLBI-71-2243-L from the National Institutes of Health.

0047-259X/82/020219-11$02.00/0

problem, based on U-statistics, were proposed by Bhapkar and Patterson [4]. Their statistics were computationally and analytically complex. The current approach not only provides comparatively simpler solutions but also more efficient ones (in the majority of the cases).

2. A Class of Multivariate Linear Rank Statistics

We consider the following multivariate (general) linear model:

$$\mathbf{X}_i = \boldsymbol{\beta}_0^{(N)} + \boldsymbol{\beta}(\mathbf{c}_{Ni} - \bar{\mathbf{c}}_N) + \boldsymbol{\varepsilon}_i, \qquad i = 1,..., N, \tag{2.1}$$

where $\boldsymbol{\beta} = (\boldsymbol{\beta}_1,..., \boldsymbol{\beta}_q)$ is a $p \times q$ matrix and $\boldsymbol{\beta}_0^{(N)}$ is a p-vector of unknown parameters, the regressors $\mathbf{c}_{Ni} = (c_{Ni1},..., c_{Niq})'$ are specified q-vectors, $\bar{\mathbf{c}}_N = N^{-1} \sum_{i=1}^N \mathbf{c}_i$, and the $\boldsymbol{\varepsilon}_i$ are independent and identically distributed random vectors (i.i.d.r.v.) with a continuous (unknown) p-variate distribution function (d.f.) F. As we shall see in Section 3, our main interest lies in testing hypotheses about $\boldsymbol{\beta}$, where there may be some a priori linear restriction on $\boldsymbol{\beta}$. Our proposed tests are based on some linear rank statistics which we present below. We form the adjusted constants

$$\mathbf{d}_{Ni} = N^{-1/2}(\mathbf{c}_{Ni} - \bar{\mathbf{c}}_N) = (d_{Ni1},..., d_{Niq})', \qquad i = 1,..., N; \tag{2.2}$$

$$\mathbf{D}_N = \sum_{i=1}^N \mathbf{d}_{Ni}\mathbf{d}'_{Ni} = ((d_{N(kk')}))_{k,k'=1,...,q}. \tag{2.3}$$

Also, we let $\mathbf{X}_i = (X_{i1},..., X_{ip})'$, $i = 1,..., N$, and let R_{ij} be the rank of X_{ij} among the set $(X_{1j},..., X_{Nj})$, for $i = 1,..., N$ and $j = 1,..., p$, so that we have a set of rankings for each of the p coordinates. Then for each j $(1 \leqslant j \leqslant p)$, we consider a set of scores $a_{Nj}(i)$, $1 \leqslant i \leqslant N$, defined in the following manner:

$$a_{Nj}(i) = \phi_j(i/(N+1)) \quad \text{or} \quad E\phi_j(U_{Ni}) \quad \text{or} \quad N\int_{(i-1)/N}^{i/N} \phi_j(t)\,dt, \tag{2.4}$$

where $U_{N1} < \cdots < U_{NN}$ are the ordered random variables of a sample of size N from the uniform (0, 1) d.f. and the ϕ_j are suitable *score functions*. Then in a manner similar to Puri and Sen [12], we construct the matrix of linear rank statistics

$$\mathbf{T}_N = (\mathbf{T}_{N1},..., \mathbf{T}_{Nq}) = ((T_{Njk}))_{j=1,...,p;k=1,...,q}, \tag{2.5}$$

where

$$T_{Njk} = \sum_{i=1}^N d_{Nik} a_{Nj}(R_{ij}) \qquad \text{for} \quad j = 1,..., p; \quad k = 1,..., q. \tag{2.6}$$

As additional notations, we let $F_{[j]}$ and $F_{[j,j']}$ be the univariate and bivariate marginals of F, respectively, for $j\ (\neq j') = 1,..., p$; similar notations hold for the probability density function f. Finally, we let $\dot{\mathbf{f}}(\mathbf{x}) = (\partial/\partial\mathbf{x}) f(\mathbf{x})$ denote the p-vector of partial derivatives of f, so that the Fisher information matrix, of order $p \times p$, for the density f is

$$\mathbf{I}(f) = \int [\{\dot{\mathbf{f}}(\mathbf{x})\}\{\dot{\mathbf{f}}(\mathbf{x})\}' \{f(\mathbf{x})\}^{-1} d\mathbf{x}. \tag{2.7}$$

The assumptions listed below are needed for the asymptotic distribution theory of $\mathbf{T}_N$ and for the development of a UI statistic.

Assumption I. The constants in (2.2) and (2.3) satisfy the following:

(i) $$\lim_{N\to\infty} \max_{1\leqslant k\leqslant q} \{|d_{Nik}|/(d_{N(kk)})^{1/2}\} = 0; \tag{2.8}$$

(ii) $$\lim_{N\to\infty} \mathbf{D}_N = \mathbf{D}, \text{ which exists and is of full rank } q. \tag{2.9}$$

Assumption II. (i) The distribution function F is absolutely continuous with an absolutely continuous density function f;

(ii) $\dot{\mathbf{f}}(\mathbf{x})$ exists and is continuous almost everywhere (a.e.);

(iii) the largest characteristic root of $\mathbf{I}(f)$ is finite.

Assumption III. The score functions ϕ_j, $1 \leqslant j \leqslant p$, in (2.4) are nondecreasing, square integrable, and absolutely continuous inside (0, 1).

Assumption IV. For each j, $1 \leqslant j \leqslant p$, at least one of the following two conditions is true:

(i) $\lim_{x\to\pm\infty} \phi_j\{F_{[j]}(x)\} f_{[j]}(x) = 0$;

(ii) $-\log\{f_{[j]}(x)\}$ is nondecreasing.

Whenever it is convenient, we shall roll out $\mathbf{T}_N$ and $\boldsymbol{\beta}$ into pq-vectors. Next, we define the $p \times p$ matrix $\boldsymbol{\nu} = ((\nu_{jj'}))$ by letting

$$\nu_{jj'} = \iint \phi_j\{F_{[j]}(x)\} \phi_{j'}\{F_{[j']}(y)\} dF_{[j,j']}(x, y) - \bar{\phi}_j\bar{\phi}_{j'} \tag{2.10}$$

for $j, j' = 1,..., p$, where

$$\bar{\phi}_j = \int_0^1 \phi_j(u)\, du \qquad \text{for} \quad j = 1,..., p. \tag{2.11}$$

Also, we define the $p \times p$ stochastic matrix $\mathbf{V}_N = ((v_{Njj'}))$, where

$$v_{Njj'} = (N-1)^{-1} \sum_{i=1}^{N} \{a_{Nj}(R_{ij}) - \bar{a}_{Nj}\}\{a_{Nj'}(R_{ij'}) - \bar{a}_{Nj'}\} \tag{2.12}$$

for $j, j' = 1,\ldots, p$, and

$$\bar{a}_{Nj} = N^{-1} \sum_{i=1}^{N} a_{Nj}(i) \qquad \text{for} \quad j = 1,\ldots, p. \tag{2.13}$$

For proofs of any of the following results in this section, we may refer to [7]. We shall only outline what is essential for the subsequent sections. Now under $H_0: \boldsymbol{\beta} = \mathbf{0}$,

$$\mathbf{V}_N \xrightarrow{p} \mathbf{v} \qquad \text{and} \qquad \mathbf{T}_N \xrightarrow{\mathscr{D}} \mathscr{N}_{pq}(\mathbf{0}, \mathbf{D} \otimes \mathbf{v}), \tag{2.14}$$

where $\otimes$ represents the Kronecker product. We reformulate the model in (2.1) as

$$\mathbf{X}_i = \boldsymbol{\beta}_0^{(N)} + \mathbf{d}_{Ni}\boldsymbol{\lambda} + \boldsymbol{\varepsilon}_i, \qquad i = 1,\ldots, N, \tag{2.15}$$

where the $\boldsymbol{\varepsilon}_i$ and $\boldsymbol{\beta}_0^{(N)}$ are defined as before and $\boldsymbol{\lambda} = N^{1/2}\boldsymbol{\beta}$. Then, we frame H_0: $\boldsymbol{\lambda} = \mathbf{0}$ against an alternative $\{K_N$: (2.15) holds for some fixed $\boldsymbol{\lambda} \neq 0\}$. The sequence of alternatives $\{K_N\}$ is contiguous to H_0, and under $\{K_N\}$,

$$\mathbf{V}_N \xrightarrow{p} \mathbf{v} \qquad \text{and} \qquad \mathbf{T}_N \xrightarrow{\mathscr{D}} \mathscr{N}_{pq}((\mathbf{D} \otimes \boldsymbol{\Gamma})\,\boldsymbol{\lambda}, \mathbf{D} \otimes \mathbf{v}), \tag{2.16}$$

where $\boldsymbol{\Gamma}$ is a diagonal matrix (of order $p \times p$) with elements

$$\gamma_{jj} = \int_0^1 \phi_j(u)\, \psi_j(u)\, du, \qquad j = 1,\ldots, p, \tag{2.17}$$

and where for each j $(= 1,\ldots, p)$ and $u \in (0, 1)$,

$$\psi_j(u) = -f'_{[j]}\{F_{[j]}^{-1}(u)\}/f_{[j]}\{F_{[j]}^{-1}(u)\}. \tag{2.18}$$

We also need to estimate $\boldsymbol{\Gamma}$, and for this purpose we borrow some results from Jurečková [9, 10]. We let $\mathbf{B} = ((b_{jk}))$ be a $p \times q$ matrix of real elements and let $\mathbf{X}_i(\mathbf{B}) = \mathbf{X}_i - \mathbf{B}(\mathbf{c}_{Ni} - \bar{\mathbf{c}}_N)$, $1 \leqslant i \leqslant N$; corresponding to these new variables, we define $R_{ij}(\mathbf{B})$ as the rank of $X_{ij}(\mathbf{B})$ among the set $(X_{1j}(\mathbf{B}),\ldots, X_{Nj}(\mathbf{B}))$, for $1 \leqslant i \leqslant N$ and $1 \leqslant j \leqslant p$. Then in (2.5) and (2.6) we replace the R_{ij} with the $R_{ij}(\mathbf{B})$ and denote the resultant statistics by $T_{Njk}(\mathbf{B})$, for $1 \leqslant j \leqslant p$ and $1 \leqslant k \leqslant q$, and $\mathbf{T}_N(\mathbf{B}) = ((T_{Njk}(\mathbf{B}))$. Finally, we let $\mathbf{E}_{jk}$ be the $p \times q$ matrix having 1 in the cell (j, k) and 0 elsewhere, for $1 \leqslant j \leqslant p$ and $1 \leqslant k \leqslant q$. The estimator of $\boldsymbol{\Gamma}$ we propose is $\hat{\boldsymbol{\Gamma}}_N = \mathrm{Diag}(\hat{\gamma}_{Njj}, 1 \leqslant j \leqslant p)$, where

$$\hat{\gamma}_{Njj} = q^{-1} \sum_{k=1}^{q} \{T_{Njk}(\mathbf{0}) - T_{Njk}(\mathbf{E}_{jk})\}. \tag{2.19}$$

If we let

$$\mathbf{U}_N = (\mathbf{D}_N^{-1} \otimes \hat{\mathbf{\Gamma}}_N^{-1})\,\mathbf{T}_N, \tag{2.20}$$

then under $H_0: \boldsymbol{\lambda} = \mathbf{0}$,

$$\mathbf{U}_N \xrightarrow{\mathscr{D}} \mathscr{N}_{pq}(\mathbf{0}, \mathbf{D}^{-1} \otimes (\mathbf{\Gamma}^{-1}\mathbf{v}\mathbf{\Gamma}^{-1})), \tag{2.21}$$

and under $\{K_N\}$,

$$\mathbf{U}_N \xrightarrow{\mathscr{D}} \mathscr{N}_{pq}(\boldsymbol{\lambda}, \mathbf{D}^{-1} \otimes (\mathbf{\Gamma}^{-1}\mathbf{v}\mathbf{\Gamma}^{-1})). \tag{2.22}$$

3. The Linear Equality Constraint Problem

We intend to develop a functional statistic form $\mathbf{U}_N$ in (2.20) to test for

$$H_0: \boldsymbol{\beta} = \mathbf{0} \qquad \text{against} \qquad H_1: \mathbf{A}\boldsymbol{\beta} = \mathbf{0}, \tag{3.1}$$

where $\boldsymbol{\beta}$ is defined in (2.1) and $\mathbf{A}$ is a full rank, a pq matrix, with $a < pq$. In [7], we developed a statistic for a more general restricted alternative problem from U_N. However, for the special case in this section, the sophisticated nonlinear programming approach to locating a UI statistic is not necessary. Hence, we just demonstrate the UI technique for this special case.

For each $\mathbf{b} \in E^{pq}$, $\mathbf{b} \neq \mathbf{0}$, we define the univariate statistic

$$U_N(\mathbf{b}) = (\mathbf{b}'\mathbf{U}_N)/(\mathbf{b}'\mathbf{\Sigma}_N\mathbf{b})^{1/2}, \tag{3.2}$$

where the stochastic matrix $\mathbf{\Sigma}_N$ is defined by

$$\mathbf{\Sigma}_N = \mathbf{D}_N^{-1} \otimes (\hat{\mathbf{\Gamma}}_N^{-1}\mathbf{V}_N\hat{\mathbf{\Gamma}}_N^{-1}), \tag{3.2}$$

and $\mathbf{D}_N$, $\mathbf{V}_N$, and $\hat{\mathbf{\Gamma}}_N$ are defined in (2.3), (2.12), and (2.19), respectively. From the results in Section 2,

$$\mathbf{\Sigma}_N \xrightarrow{P} \mathbf{\Sigma} = \mathbf{D}^{-1} \otimes (\mathbf{\Gamma}^{-1}\mathbf{v}\mathbf{\Gamma}^{-1}) \tag{3.4}$$

under $H_0: \boldsymbol{\lambda} = \mathbf{0}$ and the contiguous sequence $\{K_N\}$, where $\mathbf{D}$, $\mathbf{v}$, and $\mathbf{\Gamma}$ are defined in Assumptions I(ii), (2.10), and (2.17), respectively. Therefore,

$$U_N(\mathbf{b}) \xrightarrow{\mathscr{D}} \mathscr{N}_1((\mathbf{b}'\boldsymbol{\lambda})/(\mathbf{b}'\mathbf{\Sigma}\mathbf{b})^{1/2}, 1), \tag{3.5}$$

under H_0 and the contiguous sequence $\{K_N\}$.

If we let $\rho_N(\mathbf{b})$ be the asymptotic size α test function $(0 < \alpha < 1)$ which rejects H_0: $\boldsymbol{\lambda} = \mathbf{0}$ for large values of $U_N(\mathbf{b})$, then from (3.5),

$$P[\rho_N(\mathbf{b}) = 1 \mid H_0] = P[U_N(\mathbf{b}) > \tau_\alpha \mid H_0] \to \alpha, \tag{3.6}$$

and the asymptotic power of the test $\rho_N(b)$ is

$$P[\rho_N(b) = 1 \mid K_N] \to 1 - \Phi(\tau_\alpha - (\mathbf{b}'\boldsymbol{\lambda})/(\mathbf{b}'\boldsymbol{\Sigma} b)^{1/2}), \tag{3.7}$$

where Φ is the standar normal distribution function and $\Phi(\tau_\alpha) = 1 - \alpha$. Note that the right-hand side of (3.7) is large if $\mathbf{b}'\boldsymbol{\lambda}/(\mathbf{b}'\boldsymbol{\Sigma} b)^{1/2}$ is positive and large; the value of $\mathbf{b}$ which maximizes it is $\mathbf{b}^* = M\boldsymbol{\Sigma}^{-1}\boldsymbol{\lambda}$, $M > 0$. With this in mind, we partition the parameter space Ω $(= \{\boldsymbol{\lambda} \in E^{pq}\colon \mathbf{A}\boldsymbol{\lambda} = \mathbf{0}\})$ into subspaces $\Omega(\mathbf{b}) = \{\boldsymbol{\lambda} \in E^{pq}\colon \boldsymbol{\lambda} = M\boldsymbol{\Sigma}\mathbf{b},\ M > 0\}$ and define the set B as

$$B = \{\mathbf{b} \in E^{pq}\colon \mathbf{A}\boldsymbol{\Sigma}\mathbf{b} = \mathbf{0},\ \mathbf{b}'\boldsymbol{\Sigma}\mathbf{b} = 1\}, \tag{3.8}$$

so that $\Omega = \bigcup_{\mathbf{b} \in B} \Omega(\mathbf{b})$. Note that $\boldsymbol{\Sigma}$ is unknown, and hence, as in [7], we define

$$B_N = \{\mathbf{b} \in E^{pq}\colon \mathbf{A}\boldsymbol{\Sigma}_N\mathbf{b} = \mathbf{0},\ \mathbf{b}'\boldsymbol{\Sigma}_N\mathbf{b} = 1\}. \tag{3.9}$$

Then, analogous to the Type I UI-test of Roy [14], with the modifications in [7], we reject H_0: $\boldsymbol{\lambda} = \mathbf{0}$ in favor of H_1: $\mathbf{A}\boldsymbol{\lambda} \doteq \mathbf{0}$, if $\rho_N(\mathbf{b}) = 1$ for some $\mathbf{b} \in B_N$; this leads us to the UI-test statistic

$$Q_N = \sup\{U_N(\mathbf{b})\colon \mathbf{b} \in B_N\}. \tag{3.10}$$

From the theory of Lagrangian multipliers (viz., [1, p. 152]), the solution to Q_N in (3.10) is easily found to be

$$Q_N^2 = \mathbf{U}_N'\{\boldsymbol{\Sigma}_N^{-1} - \mathbf{A}'(\mathbf{A}\boldsymbol{\Sigma}_N\mathbf{A}')^{-1}\,\mathbf{A}\}\,\mathbf{U}_N. \tag{3.11}$$

Note that the matrix $\{\boldsymbol{\Sigma}^{-1} - \mathbf{A}'(\mathbf{A}\boldsymbol{\Sigma}\mathbf{A}')^{-1}\,\mathbf{A}\} = \mathbf{I} - \mathbf{A}'(\mathbf{A}\boldsymbol{\Sigma}\mathbf{A}')^{-1}\,\mathbf{A}$ is idempotent and of rank $pq - a$. Hence, by an appeal to (2.22), (3.4) and Lemma 3.1 of Chatterjee and De [5], we arrive at the following.

THEOREM 3.1. *Under the sequence of alternatives* $\{K_N\}$ *and the regularity conditions in Section* 2, Q_N^2 *has asymptotically a noncentral chi-squared distribution with pq-a degrees of freedom and noncentrality parameter*

$$\delta^2 = \boldsymbol{\lambda}'\{\boldsymbol{\Sigma}^{-1} - \mathbf{A}'(\mathbf{A}\boldsymbol{\Sigma}\mathbf{A}')^{-1}\}\,\boldsymbol{\lambda}. \tag{3.12}$$

The advantage of using the statistic Q_N^2 in (3.11) for the testing problem in (3.1), instead of using the unrestricted test statistic (viz., [12])

$$R_N^2 = \mathbf{U}_N'\boldsymbol{\Sigma}_N^{-1}\mathbf{U}_N, \tag{3.13}$$

is that Q_N^2 is asymptotically more powerful than R_N^2 in the region Ω. To force this point, note that [12] R_N^2, under $\{K_N\}$, has asymptotically a noncentral chi-squared distribution with pq degrees of freedom and noncentrality parameter $\boldsymbol{\lambda}'\boldsymbol{\Sigma}^{-1}\boldsymbol{\lambda}$; therefore, if q_α^2 and r_α^2 be so defined that

$$P\{Q^2 \geqslant q_\alpha^2 \mid H_0\} = P\{R^2 \geqslant r_\alpha^2 \; H_0\} = \alpha \qquad (0 < \alpha < 1), \tag{3.14}$$

then for any $\boldsymbol{\lambda} \neq \mathbf{0}$,

$$P\{Q^2 \geqslant q_\alpha^2 \mid \boldsymbol{\lambda} \in \Omega\} > P\{R^2 \geqslant r_\alpha^2 \mid \boldsymbol{\lambda} \in \Omega\}. \tag{3.15}$$

A proof of a result more general than (3.15) is provided in Section 3.2 of [6]. It may be remarked that in the parametric case (viz., [11]), the current testing problem can e reduced to testing $H_0: \boldsymbol{\beta}_1 = \mathbf{0}$ against $H_1: \boldsymbol{\beta}_1 \neq \mathbf{0}$ after a proper reparametrization (with $\boldsymbol{\beta}_1$ being a $(pq-a)$-dimensional vector). However, in the nonparametric case, the linear rank statistics $\mathbf{T}_N$ in (2.5) and (2.6) may not remain invariant under nonsingular transformation on the p variates. Hence, this reparameterization will generally lead to some lack of uniqueness in the resulting test statistic. However, given such a reduction, the theory developed in Sen and Puri [17] will remain applicable and will agree with the one presented here.

4. Profile Analysis

Profile analysis is a special collection of testing problems for the p-response q-sample model: $\mathbf{X}_i = c_{Ni1}\boldsymbol{\beta}_1 + \cdots + c_{Niq}\boldsymbol{\beta}_q + \boldsymbol{\varepsilon}_i$, $i = 1,\ldots, N$, where the $\boldsymbol{\beta}_j$ are all p-vectors, c_{Nik} is equal to 1 or 0 according as the $\mathbf{X}_i$ is from the kth sample or not $(1 \leqslant k \leqslant q;\ 1 \leqslant i \leqslant N)$, the $\boldsymbol{\varepsilon}_i$ are i.i.d.r.v.'s and the p-responses on each individual are comparable. We wish to test that the q-samples are equivalent, with the possible information that the response–sample interactions are null. Typically, such a case arises in many educational testing problems. If we define

$$\mathbf{G} = (\mathbf{I}_{p-1,p-1}, -\mathbf{J}_{p-1,1}) \qquad \text{and} \qquad \mathbf{M}' = (\mathbf{I}_{q-1,q-1}, -\mathbf{J}_{q-1,1}) \tag{4.1}$$

(where $\mathbf{J}$ has all elements equal to 1), then the sample main effects are mathematically represented by $\boldsymbol{\beta}\mathbf{M}$ and the response–sample interactions by $\mathbf{G}\boldsymbol{\beta}\mathbf{M}$, where $\boldsymbol{\beta} = (\boldsymbol{\beta}_1,\ldots, \boldsymbol{\beta}_q)$.

Profile analysis consists of the following hypothesis testing problems:

$$H_{RS}^0: \mathbf{G}\boldsymbol{\beta}\mathbf{M} = \mathbf{0} \qquad \text{against} \qquad H_{RS}^*: \mathbf{G}\boldsymbol{\beta}\mathbf{M} \neq \mathbf{0}, \tag{4.2}$$

$$H_S^0: \boldsymbol{\beta}\mathbf{M} = \mathbf{0} \qquad \text{against} \qquad H_S^*: \boldsymbol{\beta}\mathbf{M} \neq \mathbf{0}, \tag{4.3}$$

$$H_S^0: \boldsymbol{\beta}\mathbf{M} = \mathbf{0} \qquad \text{against} \qquad H_{RS}^*: \mathbf{G}\boldsymbol{\beta}\mathbf{M} \neq \mathbf{0}. \tag{4.4}$$

The strategy is to conduct a *preliminary test* for the response–sample interaction in (4.2), at significance level α_1. If H_{RS} is not rejected, then the testing problem of sample equivalence in (4.3) is conducted at significance level α_2; otherwise, the testing problem in (4.4) is implemented at significance level α_3. The appropriateness and implications of such a strategy will be discussed alter, after the development of the statistics and their (asymptotic) distributions.

We let $n_1, ..., n_q$ denote the respective sizes of the q samples with $\sum_{k=1}^{q} n_k = N$. Then for each k, $\bar{c}_{Nk} = n_k/N$ and we assume that there exist $\bar{c}_k$: $0 < \bar{c}_k < 1$, such that $\bar{c}_{Nk} \to \bar{c}_k$, for $k = 1, ..., q$. As in (2.2), we define $\mathbf{d}_{Ni} = N^{-1/2}(\mathbf{c}_{Ni} - \bar{\mathbf{c}}_N)$, $i = 1, ..., N$, so that $\mathbf{D}_N$ in (2.3) reduces to $((\bar{c}_{Nk}(\delta_{kk'} - \bar{c}_{Nk'})))$, where $\delta_{kk'}$ is the Kronecker delta. Also, $\mathbf{D} = \lim_{N\to\infty} \mathbf{D}_N = ((\bar{c}_k(\delta_{kk'} - \bar{c}_{k'})))$. Thus, Assumption I in Section 2 holds and we assume that Assumptions II, III and IV also hold. Since $\mathbf{D}_N$ (or $\mathbf{D}$) is not of full rank, we partition $\mathbf{c}_{Ni}$, $\mathbf{d}_{Ni}$, $\mathbf{D}_N$ and $\mathbf{D}$ according to the first $q-1$ components (i.e., $\mathbf{c}'_{Ni} = (\mathbf{c}'_{Ni(1)}, \mathbf{c}_{Niq})$,

$$\mathbf{D}_N = \begin{pmatrix} \mathbf{D}_{N(11)} & \mathbf{D}_{N(12)} \\ \mathbf{D}_{N(21)} & \mathbf{D}_{N(22)} \end{pmatrix}$$

and so on), write $\boldsymbol{\lambda} = N^{1/2}\boldsymbol{\beta}$ and $\boldsymbol{\beta}_0^{(N)} = \boldsymbol{\beta}\bar{\mathbf{c}}_N$. Then, (2.15) may be written here as

$$\mathbf{X}_i = \boldsymbol{\beta}_0^{(N)} + \boldsymbol{\lambda}\mathbf{M}\mathbf{d}_{Ni(1)} + \boldsymbol{\varepsilon}_i, \qquad i = 1, ..., N, \tag{4.5}$$

where $\mathbf{M}$ is defined in (4.1) and the corresponding $\mathbf{D}_{N(11)}$ is of full rank $(= q-1)$. For this model, we may virtually repeat the steps in Section 3 and arrive at the following. The UI-statistics for the testing problems in (4.2) and (4.3) are

$$Q_{N,1}^2 = \mathbf{U}'_N[\mathbf{D}_{N(11)} \otimes \{\mathbf{G}'(\mathbf{G}\hat{\boldsymbol{\Gamma}}_N^{-1}\mathbf{V}_N\hat{\boldsymbol{\Gamma}}_N^{-1}\mathbf{G}')^{-1}\,\mathbf{G}\}]\,\mathbf{U}_N \tag{4.6}$$

and

$$Q_{N,2}^2 = \mathbf{U}'_N\{\mathbf{D}_{N(11)} \otimes (\hat{\boldsymbol{\Gamma}}_N^{-1}\mathbf{V}_N\hat{\boldsymbol{\Gamma}}_N^{-1})^{-1}\}\,\mathbf{U}_N, \tag{4.7}$$

respectively, where $\mathbf{V}_N$ and $\hat{\boldsymbol{\Gamma}}_N$ are defined in (2.12) and (2.19). The UI-statistic for the restricted alternative problem in (4.4) is [by (3.11)]

$$Q_{N,3}^2 = Q_{N,2}^2 - Q_{N,1}^2. \tag{4.8}$$

For the model (4.5), with $\{K_N\}$ defined as in after (2.15), we have then

$$Q_{N,1}^2 \xrightarrow{\mathscr{D}} \chi^2((p-1)(q-1), \delta_1^2),$$

$$Q_{N,2}^2 \xrightarrow{\mathscr{D}} \chi^2(p(q-1), \delta_2^2)$$

and

$$Q^2_{N,3} \xrightarrow{\mathscr{D}} \chi^2(q-1, \delta_3^2), \tag{4.9}$$

where the noncentrality parameters δ_1^2, δ_2^2 and δ_3^2 $(= \delta_2^2 - \delta_1^2)$ are found by substituting the corresponding values of $\boldsymbol{\lambda}\mathbf{M}$, $\mathbf{D}_{(11)}$, $\boldsymbol{\Gamma}$ and $\mathbf{v}$ for $\mathbf{U}_N$, $\mathbf{D}_{N(11)}$, $\hat{\boldsymbol{\Gamma}}_N$ and $\mathbf{V}_N$ in (4.6), (4.7) and (4.8).

For the sake of comparison, we now develop an equivalent set of statistics under a normality assumption. Suppose that $\boldsymbol{\varepsilon}_i \sim \mathcal{N}_p(\mathbf{0}, \boldsymbol{\Psi})$ where $\boldsymbol{\Psi}$ is a known $p \times p$ matrix of full rank. It is not essential that the covariance matrix $\boldsymbol{\Psi}$ be specified (cf. [6, Sects. 4.6 and 5.6]; however, we make this simplifying assumption because we are only interested in asymptotic comparisons. From (4.5) and the maximum likelihood estimator (MLE) of $\boldsymbol{\lambda}\mathbf{M}$,

$$\hat{\boldsymbol{\lambda}}_N\mathbf{M} = \left(\sum_{i=1}^{N} \mathbf{X}_i \mathbf{d}'_{Ni(1)}\right) \mathbf{D}^{-1}_{N(11)}, \tag{4.10}$$

the likelihood ratio statistics for the three testing problems of (4.2)–(4.4) are

$$W^2_{N,1} = (\hat{\boldsymbol{\lambda}}_N\mathbf{M})' \; [\mathbf{D}_{N(11)} \otimes \{\mathbf{G}'(\mathbf{G}\boldsymbol{\Psi}\mathbf{G}')^{-1}\,\mathbf{G}\}](\hat{\boldsymbol{\lambda}}_N\mathbf{M}), \tag{4.11}$$

$$W^2_{N,2} = (\hat{\boldsymbol{\lambda}}_N\mathbf{M})' \; \{\mathbf{D}_{N(11)} \otimes \boldsymbol{\Psi}^{-1}\}(\hat{\boldsymbol{\lambda}}_N\mathbf{M}) \quad \text{and} \quad W^2_{N,3} = W^2_{N,2} - W^2_{N,1}, \tag{4.12}$$

respectively. Each of these statistics has an exact chi-squared distribution with an appropriate noncentrality parameter, found by substituting $\boldsymbol{\lambda}$ for $\hat{\boldsymbol{\lambda}}_N$, and degrees of freedom $(p-1)(q-1)$, $p(q-1)$ and $(q-1)$, respectively. Their asymptotic distributions remain the same even if the $\boldsymbol{\varepsilon}_i$ have non-normal distributions, but with finite second-order moments (see, [16]). Thus, individually, the asymptotic relative efficiency (ARE) of $Q_{N,t}$ with respect to $W_{N,t}$ $(t = 1, 2, 3)$ can be computed simply by the ratio of the corresponding noncentrality parameters (see [13, Sect. 3.8]). However, in general, these depend on the direction of $\boldsymbol{\lambda}$ and, in many cases, one may have to be satisfied with some lower or upper bounds easily obtainable from the characteristic roots of the two matrices appearing in the noncentrality parameters. In general, the rank statistics fare quite well compared to their parametric counterparts, more noticably for distributions with heavy tails. In particular, if the $\boldsymbol{\varepsilon}_i$ have normal distribution and for the rank procedure, we use the normal scores (i.e., in (2.4), we take $\phi_j(u) = \Phi^{-1}(u)$, the inverse normal d.f.), then this ARE is equal to 1.

The rank statistics set for profile analysis provides a suitable alternative to the likelihood ratio tests because (i) it is ADF, (ii) it is computationally simple on a computer and (iii) it fares well with regards to asymptotic power. Although the U-statistics set for profile analysis, proposed in [4], also has the properties (i) and (iii), it is nearly impossible to make any type of analytical power comparison with the likelihood ratio type tests.

We now return to the question of whether or not there is any advantage of using the strategy decision rule for testing sample differences. For convenience, let us assume that we are using the asymptotic analogues Q_1^2, Q_2^2 and Q_3^2 of the rank UI-statistics $Q_{N,1}^2$, $Q_{N,2}^2$ and $Q_{N,3}^2$ in (4.6)–(4.8); these statistics have exact chi-squared distributions, given by the right hand sides of (4.9). If P_0 denotes the probability under appropriate null hypothesis and we define the s_t by

$$\alpha_t = P_0\{Q_t^2 \geqslant s_t\} \qquad \text{for} \quad t = 1, 2, 3, \tag{4.13}$$

then the overall significance level of the strategic decision rule is

$$\alpha = P_0\{Q_1^2 \geqslant s_1,\ Q_2^2 \geqslant s_2\} + P_0\{Q_1^2 < s_1,\ Q_3^2 \geqslant s_3\}. \tag{4.14}$$

We would like to compare α with α_2, because Q_2^2 would be the statistic used for testing sample differences if no attention were paid to the response–sample interaction. Note that $Q_2^2 = Q_1^2 + Q_3^2$ and Q_1 and Q_3 are independent. Hence, if we set $\alpha_2 = \alpha_3$ (a reasonable condition since Q_2^2 and Q_3^2 test the same null hypothesis), then by (4.13), (4.14) and the fact that $s_1 + s_3 \geqslant s_2$, we have

$$\begin{aligned} \alpha &= P_0\{Q_1^2 + Q_3^2 \geqslant s_2,\ Q_1^2 \geqslant s_1\} + \alpha_3(1-\alpha_1) \\ &= \alpha_1 P_0\{Q_1^2 + Q_3^2 \geqslant s_2 \mid Q_1^2 \geqslant s_1\} + \alpha_2(1-\alpha_1) \\ &\geqslant \alpha_1\alpha_2 + \alpha_2(1-\alpha_1) = \alpha_2. \end{aligned} \tag{4.15}$$

This shows that α, the overall significance level for the decision rule, is larger than α_2, the significance level for just testing (4.3). However, if the difference is small enough [as is usually the case, see Sen and Saleh [18]), the decision rule would still have a power edge over Q_2^2 [because of (3.14) and (3.15)].

References

[1] Apostol, T. M. (1964). *Mathematical Analysis.* Addison-Wesley, Reading, Mass.

[2] Barlow, R. E., Bartholomew, D. J., Bremner, J. M., and Brunk, H. D. (1972). *Statistical Inference Under Order Restrictions.* Wiley, New York.

[3] Bhapkar, V. P. and Patterson, K. W. (1976). A Monte Carlo study of some multivariate nonparametric statistics for profile analysis of several samples. Dept. Statist. Univ. Kentucky, Tech. Report No. 101.

[4] Bhapkar, V. P. and Patterson, K. W. (1977). On some nonparametric tests for profile analysis of several multivariate samples. J. Multivariate Anal. 7 265–277.

[5] Chatterjee, S. K. and De, N. K. (1972). Bivariate nonparametric location tests against restricted alternatives. *Calcutta Statist. Assoc. Bull.* 21 1–20.

[6] CHINCHILLI, V. M. (1979). *Rank Tests for Restricted Alternatives Problems in Multivariate Analysis*. Dissertation, Univ. of North Carolina, Chapel Hill.

[7] CHINCHILLI, V. M. AND SEN, P. K. (1981). Multivariate linear rank statistic and the Union–Intersection Principle for hypothesis testing under restricted alternatives. *Sankhyā Ser. B* **43** 135–151.

[8] CHINCHILLI, V. M. AND SEN, P. K. (1981). Multivariate linear rank statistics and the Union–Intersection Principle for the orthant restriction problem. *Sankhyā Ser. B* **43** 152–171.

[9] JUREČKOVÁ, J. (1969). Asymptotic linearity of a rank statistic in regression parameter. *Ann. Math. Statist.* **40** 1889–1900.

[10] JUREČKOVÁ, J. (1971). Nonparametric estimates of regression coefficients. *Ann. Math. Statist.* **42** 1328–1338.

[11] MORRISON, D. F. (1976). *Multivariate Statistical Methods*. McGraw–Hill, New York.

[12] PURI, M. L. AND SEN, P. K. (1969). A class of rank order tests for a general linear hypothesis. *Ann. Math. Statist.* **40** 1325–1343.

[13] PURI, M. L. AND SEN, P. K. (1971). *Nonparametric Methods in Multivariate Analysis*. Wiley, New York.

[14] ROY, S. N. (1957). *Some Aspects of Multivariate Analysis*. Wiley, New York.

[15] SEARLE, S. R. (1971). *Linear Models*. Wiley, New York.

[16] SEN, P. K. AND PURI, M. L. (1970). Asymptotic theory of liklihood ratio and rank order tests in some multivariate linear models. *Ann. Math. Statist.* **41** 87–100.

[17] SEN, P. K. AND PURI, M. L. (1977). Asymptotically distribution-free aligned rank order tests for composite hypotheses for general multivariate linear models. *Z. Wahrsch. Verw. Gebiete* **39** 175–186.

[18] SEN, P. K. AND SALEH, A. K. MD. E. (1979). Nonparametric estimation of location parameter after a preliminary test on regression in the multivariate case. *J. Multivariate Anal.* **9** 322–331.