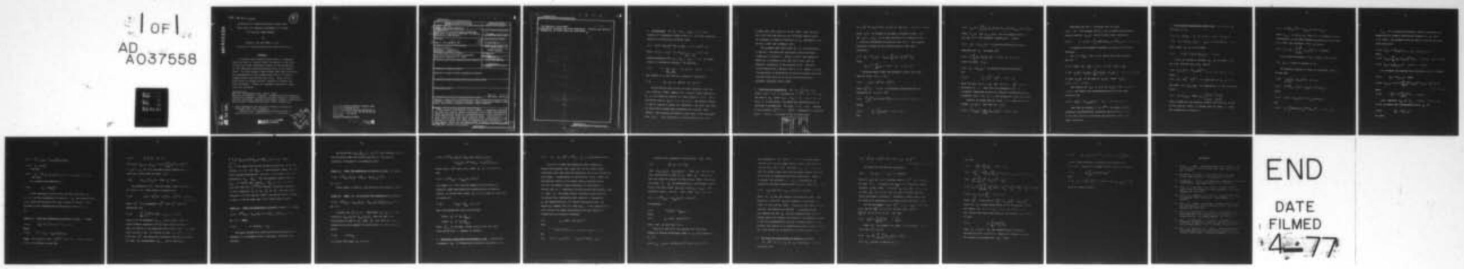
AD-A037 558 INDIANA UNIV BLOOMINGTON DEPT OF MATHEMATICS F/G 12/1
ASYMPTOTICALLY DISTRIBUTION-FREE ALIGNED RANK ORDER TESTS FOR C--ETC(U)
1977 M L PURI, P K SEN AF-AFOSR-2927-76
UNCLASSIFIED AFOSR-TR-77-0167 NL



ASYMPTOTICALLY DISTRIBUTION-FREE ALIGNED RANK ORDER TESTS FOR COMPOSITE HYPOTHESES FOR GENERAL MULTIVARIATE LINEAR MODELS*

by

Pranab K. Sen and Madan L. Puri

University of North Carolina and Indiana University

ABSTRACT

For general multivariate linear models, a composite hypothesis does not usually induce invariance of the joint distribution under appropriate groups of transformations, so that genuinely distribution-free tests do not usually exist. For this purpose, some aligned rank order statistics are incorporated in the proposal and study of a class of asymptotically distribution-free tests. Tests for the parallelism of several multiple regression surfaces are also considered. Finally the optimal properties of these tests are discussed.

AMS 1970 Classification No: 62G10, 62J05

Keywords and Phrases: Alignment, asymptotically distribution-free tests, asymptotic linearity of rank statistics, asymptotic relative efficiency, composite hypotheses, general linear models, parallelism of regression surfaces, robust estimation.

* Work supported by the Air Force Office of Scientific Research, AFSC USAF, Contract AFOSR-76-2927. Reproduction in whole or part is permitted for any purpose of the U. S. Government.

UNCLASSIFIED

SECURITY CLASSIFICATION OF THIS PAGE (When Data Entered)

REPORT DOCUMENTATION PAGE

READ INSTRUCTIONS BEFORE COMPLETING FORM

1. REPORT NUMBER: 19 18 AFOSR - TR - 77 - 0167
2. GOVT ACCESSION NO.:
3. RECIPIENT'S CATALOG NUMBER:
4. TITLE (and Subtitle): 6 ASYMPTOTICALLY DISTRIBUTION-FREE ALIGNED RANK ORDER TESTS FOR COMPOSITE HYPOTHESES FOR GENERAL MULTIVARIATE LINEAR MODELS.
5. TYPE OF REPORT & PERIOD COVERED: 9 Interim repto
6. PERFORMING ORG. REPORT NUMBER:
7. AUTHOR(s): 10 Madan L. Puri ~~and E.~~ Pranab K. Sen
8. CONTRACT OR GRANT NUMBER(s): 15 ~~AF~~ AFOSR ~~76~~-2927-76
9. PERFORMING ORGANIZATION NAME AND ADDRESS: Indiana University, Department of Mathematics, Bloomington, Indiana 47401
10. PROGRAM ELEMENT, PROJECT, TASK AREA & WORK UNIT NUMBERS: 16 61102F 2304 A5 17
11. CONTROLLING OFFICE NAME AND ADDRESS: Air Force Office of Scientific Research/NM, Bolling AFB, Washington, DC 20332
12. REPORT DATE: 11 1977
13. NUMBER OF PAGES: 21 12 24p.
14. MONITORING AGENCY NAME & ADDRESS (if different from Controlling Office):
15. SECURITY CLASS. (of this report): UNCLASSIFIED

15a. DECLASSIFICATION/DOWNGRADING SCHEDULE:

16. DISTRIBUTION STATEMENT (of this Report)

Approved for public release; distribution unlimited.

17. DISTRIBUTION STATEMENT (of the abstract entered in Block 20, if different from Report)

18. SUPPLEMENTARY NOTES

402 522

19. KEY WORDS (Continue on reverse side if necessary and identify by block number)

alignment, asymptotically distribution-free tests, asymptotic linearity of rank statistics, asymptotic relative efficiency, composite hypotheses, general linear models, parallelism of regression surfaces, robust estimation

20. ABSTRACT (Continue on reverse side if necessary and identify by block number)

For general multivariate linear models, a composite hypothesis does not usually induce invariance of the joint distribution under appropriate groups of transformations, so that genuinely distribution-free tests do not usually exist. For this purpose, some aligned rank order statistics are incorporated in the proposal and study of a class of asymptotically distribution-free tests. Tests for the parallelism of several multiple →

next page

DD FORM 1 JAN 73 1473 EDITION OF 1 NOV 65 IS OBSOLETE

UNCLASSIFIED

SECURITY CLASSIFICATION OF THIS PAGE (When Data Entered)

UNCLASSIFIED

SECURITY CLASSIFICATION OF THIS PAGE(When Data Entered)

20. Abstract (Continued)

regression surfaces are also considered. Finally the optimal properties of these tests are discussed.

UNCLASSIFIED

SECURITY CLASSIFICATION OF THIS PAGE(When Data Entered)

1. Introduction. Let $\underset{\sim}{X}_i = (X_{1i}, \ldots, X_{pi})'$, $i \geq 1$ be a sequence of independent random vectors (i. rvs) with continuous cumulative distribution functions (cdfs)

$$(1.1) \quad F_i(\underset{\sim}{x}) = P[\underset{\sim}{X}_i \leq \underset{\sim}{x}] = F(\underset{\sim}{x} - \underset{\sim}{\alpha} - \underset{\sim}{\beta}\underset{\sim}{c}_i) \ , \ i \geq 1 \ , \ \underset{\sim}{x} \in R^p \ , \ p \geq 1$$

where $\underset{\sim}{\alpha} = (\alpha_1, \ldots, \alpha_p)'$, $\underset{\sim}{\beta} = ((\beta_{jk}))_{\substack{j=1,\ldots,p \\ k=1,\ldots,q}}$, $q \geq 1$ are unknown parameters and $\underset{\sim}{c}_i = (c_{1i}, \ldots, c_{qi})'$, $i \geq 1$ are known vectors of regression constants. We partition

$$(1.2) \qquad \underset{\sim}{\beta} = (\underset{p\times q_1}{\underset{\sim}{\beta}_1}, \underset{p\times q_2}{\underset{\sim}{\beta}_2}) \ , \ q_1 + q_2 = q \ , \ q_i \geq 0 \ , \ i = 1,2 \ .$$

The problem is to test the null (composite) hypothesis

$$(1.3) \qquad H_0 : \underset{\sim}{\beta}_2 = \underset{\sim}{0} \quad \text{against} \quad H_1 : \underset{\sim}{\beta}_2 \neq \underset{\sim}{0} \ .$$

We may mention that by the classical canonical reduction [viz. Anderson (1958), Chapter 8)], a general linear hypothesis on $\underset{\sim}{\beta}$ can always be reduced to a form similar to (1.3). For a particular case of $q_2 = q$ i.e. $H_0 : \underset{\sim}{\beta} = \underset{\sim}{0}$, the problem reduces to that of testing a simple null hypothesis, the rank order tests for which have already been studied by Puri and Sen (1969). However, the technique developed in that paper is not applicable when $q_2 < q$. This difficulty is circumvented here by using aligned rank order tests [as in Sen (1969), and, Puri and Sen (1973) both dealing with the univariate models] where the alignment is based on estimates of $\underset{\sim}{\beta}_1$ developed in Sen and Puri (1969) and Jurečková (1971).

The proposed rank order tests for H_0 are considered in section 3 following the preliminary notions and basic assumptions in section 2. Section 4 deals with asymptotic comparison of parametric and rank order tests, and the asymptotic optimality of the proposed tests. The last section deals with a special case of (1.3), namely, testing the hypothesis of parallelism of several multiple regression surfaces which turn out to be the multivariate multi-parameter analogue of Sen (1969).

2. <u>Notations and assumptions</u>. Let $R_{ji} = \sum_{\alpha=1}^{n} u(X_{ji} - X_{j\alpha})$, (where $u(t) = 1$ or 0 according as t is $\geq$ or < 0) be the rank of X_{ji} among $X_{j1}, \ldots, X_{jn}$; $i = 1, \ldots, n$; $j = 1, \ldots, p$. Since F is continuous, ties among the observations may be neglected in probability. For each $j\ (= 1, \ldots, p)$, consider a set of scores $a_n^{(j)}(1), \ldots, a_n^{(j)}(n)$, generated by a function $\phi_j(u)$, $0 < u < 1$, in either of the following ways.

(2.1) $a_n^{(j)}(i) = \varphi_j(i/(n+1))$ or $a_n^{(j)}(i) = E\varphi_j(U_{ni})$, $1 \le i \le n$; $1 \le j \le p$

where $\varphi_j(u)$ is assumed to be square integrable inside $(0,1)$, and $u_{n1} < \ldots < u_{nn}$ is an order statistic of a sample of size n from the rectangular distribution over $(0,1)$. Our proposed procedure is based on the following type of rank order statistics.

$$(2.2) \quad \underset{\sim}{S}_n = ((S_{n,jk})) \ , \ S_{n,jk} = \sum_{i=1}^{n} (c_{ki} - \bar{c}_{kn}) a_n^{(j)}(R_{ji})$$

where

$$\bar{c}_{kn} = n^{-1} \sum_{i=1}^{n} c_{ik} \ ; \quad k = 1,\ldots,q \ ; \quad j = 1,\ldots,p \ .$$

Following Hájek (1968) and Hoeffding (1973), we assume that for every $j(=1,\ldots,p)$,

$$(2.3) \qquad \varphi_j(u) = \varphi_j^{(1)}(u) - \varphi_j^{(2)}(u)$$

where $\varphi_j^{(s)}(u)$, $s = 1,2$ is absolutely continuous and non-decreasing in $u \in (0,1)$ and

$$(2.4) \qquad \int_0^1 |\varphi_j^{(s)}(u)| \{u(1-u)\}^{-\frac{1}{2}} du < \infty \ ; \quad s = 1,2 \ ; \quad j = 1,\ldots,p \ .$$

Denote

$$(2.5) \qquad \bar{\varphi}_j = \int_0^1 \varphi_j(u)\,du \ , \quad j = 1,\ldots,p \ ,$$

and

$$\lambda_{jj'}(F) = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} \varphi_j(F_{[j]}(x))\varphi_{j'}(F_{[j']}(y))\,dF_{[jj']}(x,y) - \bar{\varphi}_j\bar{\varphi}_{j'}, \tag{2.6}$$

where $F_{[j]}(x)$ and $F_{[jj']}(x,y)$ are the marginal cdfs of jth and (j,j') the components respectively. Assume

(2.7) $\underset{\sim}{\Lambda}(F) = ((\lambda_{jj'}(F)))$ is positive definite and finite.

Regarding the $\underset{\sim}{c}_i$, we assume that

$$n^{-1}\sum_{i=1}^{n}(\underset{\sim}{c}_i - \bar{\underset{\sim}{c}}_n)(\underset{\sim}{c}_i - \bar{\underset{\sim}{c}}_n)' = n^{-1}\underset{\sim}{C}_n \to \underset{\sim}{C} \quad \text{as} \quad n \to \infty \tag{2.8}$$

where for every $n \geq n_0$,

(2.9) $\underset{\sim}{C}_n = ((C_{n,kk'}))$ is positive definite and finite

and

$$\underset{\sim}{c}_i = \underset{\sim}{c}_i^{(1)} - \underset{\sim}{c}_i^{(2)} , \quad i = 1,\ldots,n \tag{2.10}$$

where for each $k(=1,\ldots,q)$ and $s(=1,2)$, $c_{ki}^{(s)}$ is non-decreasing in i . (Note that the assumption (2.7) is a slightly simplified version of a parallel assumption made by Jurečkova' (1971). For $q = 1$, this assumption is not necessary).

Finally, we assume that for every $\varepsilon > 0$, there exists an integer $n_0 = n_0(\varepsilon)$ such that for $n \geq n_0$,

$$n^{-1}C_{n,kk} > \varepsilon \{ \max_{1\leq i\leq n} |c_{ki} - \bar{c}_{kn}|^2 \} , \quad k = 1,\ldots,q . \tag{2.11}$$

Regarding the cdf F, we assume that for each $j(=1,\dots,p)$, the marginal cdf $F_{[j]}$ has an absolutely continuous density function $f_{[j]}(x)$ with a finite Fisher information

$$(2.12)\qquad I_j = I_j(f_j) = \int_{-\infty}^{\infty} \{(d/dx)\log f_{[j]}(x)\}^2 dF_{[j]}(x), \quad j=1,\dots,p.$$

To explain the alignment procedure, we need the following notations.

Let $\underset{\sim}{B} = ((b_{jk}))$ be a $p \times q$ matrix with real elements and let

$$(2.13)\qquad \underset{\sim}{X}_i(\underset{\sim}{B}) = \underset{\sim}{X}_i - \underset{\sim}{B}\underset{\sim}{C}_i \; ; \; i=1,\dots,n \; ; \; \underset{\sim}{B}' = (\underset{\sim}{b}_1',\dots,\underset{\sim}{b}_p') \; ;$$

$$(2.14)\qquad R_{ji}(\underset{\sim}{B}) = R_{ji}(\underset{\sim}{b}_j) = \sum_{\alpha=1}^{n} u(X_{ji}(\underset{\sim}{b}_j) - X_{j\alpha}(\underset{\sim}{b}_j)), \quad 1 \le i < n, \quad 1 \le j < p$$

so that $R_{ji}(\underset{\sim}{B})$ is the rank of $X_{ji}(\underset{\sim}{b}_j)$ among $X_{j\alpha}(\underset{\sim}{b}_j)$, $\alpha=1,\dots,n$, $1 \le i \le n$.

Now replace the R_{ji} in (2.2) by $R_{ji}(\underset{\sim}{b}_j)$ for $1 \le i \le n$, $1 < j \le p$ and denote the corresponding matrix of rank order statistics by

$$(2.15)\qquad \underset{\sim}{S}_n(\underset{\sim}{B}) = ((S_{n,jk}(\underset{\sim}{b}_j))), \quad j=1,\dots,p \; ; \; k=1,\dots,q.$$

Note that by varying $\underset{\sim}{B}$ on $R^{p\times q}$, we obtain a multi-parameter multidimensional stochastic process which is used in the next section to introduce the proposed aligned rank order statistics.

3. <u>The Proposed Aligned Rank Order Tests</u>. As in (1.2), we partition $\underset{\sim}{B}$ as

$$(3.1)\quad \underset{\sim}{B} = (\underset{\sim}{B}_1, \underset{\sim}{B}_2)\ ,\quad B_i \text{ is } p \times q_i\ ;\ i = 1,2\ ;\ q_1 + q_2 = q$$

$$(3.2)\quad \underset{\sim}{c}_i' = (\underset{\sim}{c}_{i(1)}', \underset{\sim}{c}_{i(2)})\ ,\quad \underset{\sim}{c}_{i(s)} \text{ is a } q_s\text{-vector},\quad s = 1,2\ .$$

Then, under H_0 in (1.3), we have

$$(3.3)\qquad F_i(\underset{\sim}{x}) = F(\underset{\sim}{x} - \underset{\sim}{\alpha} - \underset{\sim}{\beta}_1 \underset{\sim}{c}_{i(1)})\ ,\quad 1 \le i < n$$

First, we proceed to estimate $\underset{\sim}{\beta}_1$ for the model (3.3). For this, consider the $p \times q_1$ matrix

$$(3.4)\quad \underset{\sim}{S}_{n(1)}(\underset{\sim}{B}_1) = ((S_{n,jk}(\underset{\sim}{b}_j^{(1)})))_{j=1,\ldots,p\ ;\ k=1,\ldots,q_1}$$

where

$$(3.5)\quad \underset{\sim}{b}_j' = (\underset{\sim}{b}_j^{(1)\prime}, \underset{\sim}{b}_j^{(2)\prime}) \text{ is a partition of } \underset{\sim}{b}_j \text{ by (3.1)}$$

Now under (3.3), $\underset{\sim}{S}_{n(1)}(\underset{\sim}{\beta}_1)$ has expectation $\underset{\sim}{o}$, and dispersion matrix

$$(3.6)\quad \underset{\sim}{\Lambda}(F) \otimes \underset{\sim}{C}_{n(11)} \quad \text{where} \quad \underset{\sim}{C}_n = \begin{pmatrix} \underset{\sim}{C}_{n(11)}, & \underset{\sim}{C}_{n(12)} \\ \underset{\sim}{C}_{n(21)}, & \underset{\sim}{C}_{n(22)} \end{pmatrix}$$

(and $\otimes$ stands for the Kronecker product) and from the results of Puri and Sen (1969), it follows that for large n, under the assumptions of section 2, (3.7) $\mathcal{L}(n^{-\frac{1}{2}}\underset{\sim}{S}_{n(1)}(\underset{\sim}{\beta}_1)) \to \mathcal{N}_{p\times q_1}(\underset{\sim}{0}, \underset{\sim}{\Lambda}(F)\otimes \underset{\sim}{C}_{(11)})$

where $\underset{\sim}{C}_{(11)}$ is the $q_1 \times q_1$ minor of $\underset{\sim}{C}$ defined in (2.8). Consequently, by the same alignment procedure as in Sen and Puri (1969) and Jurečkova′ (1971), we define

(3.8) $\underset{\sim}{D}_n = \left\{\underset{\sim}{B}_1 : \sum_{j=1}^{p}\sum_{k=1}^{q_1} |S_{n,jk}(\underset{\sim}{b}_j^{(1)})| = \text{minimum}\right\}$.

Our proposed estimator of $\underset{\sim}{\beta}_1$ (under (3.3)) is then

(3.9) $\hat{\underset{\sim}{B}}_{1,n}$ = center of gravity of D_n .

By arguments parallel to those of Jurečkova′ (1971), it follows that

(3.10) $\sup_{\underset{\sim}{B}_1 \in \underset{\sim}{D}_n} \|\underset{\sim}{\beta}_1 - \hat{\underset{\sim}{\beta}}_{1,n}\| \xrightarrow{p} 0$, as $n\to\infty$

(3.11) $\mathcal{L}(n^{\frac{1}{2}}[\hat{\underset{\sim}{\beta}}_{1,n} - \underset{\sim}{\beta}_1]) \to \mathcal{N}_{p\times q}(0, \underset{\sim}{\Gamma}(F)\otimes \underset{\sim}{C}_{(11)})$

where

(3.12) $\underset{\sim}{T}(F) = ((\tau_{jj'}(F))) = ((\lambda_{jj'}(F)/A_jA_{j'}))$

and

(3.13) $A_j = \int_{-\infty}^{\infty}(d/dx)\varphi_j(F_{[j]}(x))\,dF_{[j]}(x)$, $j = 1,\ldots,p$.

$\hat{\beta}_{1,n}$ is a translation-invariant, robust, consistent and asymptotically normally distributed estimator of $\underset{\sim}{\beta}_1$ when (3.3) holds. Our proposed tests are based on the aligned rank order statistics

$$(3.14)\quad \hat{S}_{n(2)} = ((\hat{S}_{n,jk}))_{j=1,\dots,p\ ;\ k=q_1+1,\dots,q}$$

where

$$(3.15)\quad \hat{S}_{n,jk} = \sum_{i=1}^{n}(c_{ki} - \bar{c}_{k,n})a_n^{(j)}(\hat{R}_{ji})\ ,\quad 1 \le j \le p\ ,\quad q_1+1 \le k \le q\ ,$$

$$(3.16)\quad \hat{R}_{ji} = R_{ji}(\hat{\underset{\sim}{\beta}}_{1,n}, \underset{\sim}{0})\ ,\quad 1 \le i \le n\ ,\quad 1 \le j \le p$$

To introduce the proposed test statistics, we first define

$$(3.17)\qquad \underset{\sim}{M}_n = ((m_{jj',n}))\ \text{ where}$$

$$(3.18)\quad m_{jj',n} = (n-1)^{-1}\left\{\sum_{i=1}^{n} a_n^{(j)}(R_{ji})a_n^{(j')}(R_{j'i}) - \bar{a}_n^{(i)}\bar{a}_n^{(j')}\right\}$$

$$j, j' = 1,\dots,p$$

where

$$(3.19)\qquad \bar{a}_n^{(j)} = n^{-1}\sum_{i=1}^{n} a_n^{(j)}(i)\ ,\quad 1 \le j \le p$$

Also, replacing R_{ji} by $\hat{R}_{ji}$, $1 \le i \le w$, $1 \le j \le p$ in (3.18), we denote the corresponding matrix $\underset{\sim}{M}_n$ by

$$(3.20)\qquad \hat{\underset{\sim}{M}}_n = ((\hat{m}_{jj',n}))\ .$$

Let then,

(3.21) $\underset{\sim}{C}_n^* = \underset{\sim}{C}_{n(22)} - \underset{\sim}{C}_{n(21)}\underset{\sim}{C}_{n(11)}^{-1}\underset{\sim}{C}_{n(12)}$

(3.22) $\underset{pq_2 \times pq_2}{\hat{\underset{\sim}{G}}_n} = \hat{\underset{\sim}{M}}_n \otimes \underset{\sim}{C}_n^*$

(3.23) $\underset{pq_2 \times bq_2}{\underset{\sim}{H}_n} = ((\hat{S}_{n,jk} \hat{S}_{n,j'k'}))_{j,j'=1,\ldots,p\ ;\ k,k'=1,\ldots,q}$

Our proposed test statistic is

(3.24) $$\mathcal{L}_n = \mathrm{Tr}[\underset{\sim}{H}_n\hat{\underset{\sim}{G}}_n^{-1}] ;$$

In the remainder of the section, we show that under H_0 in (1.3) and the assumptions of section 2, $\mathcal{L}_n$ has asymptotically a chi square distribution with pq_2 degrees of freedom. This provides an ADF (asymptotically distribution free) test for H_0.

<u>Lemma 3.1</u>. <u>Under the assumptions of section 2, when H_0 holds,</u>

(3.25) $$n\hat{\underset{\sim}{G}}_n^{-1} \xrightarrow{p} \underset{\sim}{\Lambda}^{-1}(F) \otimes \underset{\sim}{C}^{*-1}, \text{ as } n \to \infty$$

<u>where</u>

(3.26) $$\underset{\sim}{C}^* = C_{(22)} - \underset{\sim}{C}_{(21)}\underset{\sim}{C}_{(11)}^{-1}\underset{\sim}{C}_{(12)}$$

<u>**Proof**</u>. By virtue of (2.8), $\underset{\sim}{C}_n^* \xrightarrow{p} \underset{\sim}{C}^*$, as $n \to \infty$. Thus to prove (3.25), it suffices to show that

(3.27) $$\hat{\underset{\sim}{M}}_n \xrightarrow{p} \underset{\sim}{\Lambda}(F)\ , \text{ as } n \to \infty$$

Also since $\hat{m}_{jj',n} = m_{jj,n} = (n-1)^{-1}\left\{\sum_{i=1}^{n}[a_n^{(j)}(i) - \bar{a}_n^{(j)}]^2\right\}$

$\to \lambda_{jj}(F) = \lambda_{jj}$ by (2.1) and some routine computations, we need only to show that for every $j \neq j'$,

(3.28) $$\hat{m}_{jj',n} \xrightarrow{p} \lambda_{jj'}(F) \text{ when } H_0 \text{ holds .}$$

By assumption (2.3), (see also Hájek (1968), section 5) for every $\epsilon > 0$, there exists a decomposition

(3.29) $$\varphi_j(u) = \varphi_j^{(1)}(u) + \varphi_j^{(2)}(u) - \varphi_j^{(3)}(u)\ ,\ 0 < u < 1\ ,\quad j = 1,\ldots,p$$

where $\varphi_j^{(1)}$ is a polynomial, $\varphi_j^{(2)}$ and $\varphi_j^{(3)}$ are non-decreasing, and

(3.30) $$\sum_{k=2}^{3} \int_0^1 [\varphi_j^{(k)}(u)]^2 du < \epsilon\, \lambda_{jj}\ ,\quad 1 < j < p\ .$$

Using (3.29) we decompose $\hat{m}_{jj',n}$ into 9 terms. Using the Cauchy-Schwarz inequality for the eight terms for which at least one factor is non polynomial along with (3.30), it follows that to prove (3.28), it suffices to take $\varphi_j = \varphi_j^{(1)}$, $1 < j < p$. Since the $\varphi_j^{(1)}$ are absolutely continuous and are polynomials, for them, the corresponding $\hat{m}_{jj',n}$ can be written as

$$\int_{-\infty}^{\infty}\int_{-\infty}^{\infty} \varphi_j^{(1)}(\hat{H}_{nj}(x))\varphi_{j'}^{(1)}(\hat{H}_{nj'}(y))\,d\hat{H}^*_{njj'}(x,y)) + o(1) \quad \text{where}$$

$\hat{H}_{nj}$ is the sample cdf for the aligned observations on the jth variate, $1 \le j \le p$, and $\hat{H}^*_{njj'}$ is the bivariate sample cdf for these aligned observations. By (2.6), (2.14) and (3.11), on denoting by H_{nj}, $H^*_{njj'}$ the corresponding sample cdfs for $\hat{\underset{\sim}{B}}_{1,n}$, it follows that $\sup_{x,y}|H^*_{njj'} - \hat{H}^*_{n,jj'}| \to 0$, as $n \to \infty$. Also note that the $\varphi_j^{(i)}$'s are bounded, continuous functions. So first replacing $\hat{H}_n$ by H_n, $\hat{H}^*_n$ by H^*_n, and then using theorem 4.1 of Puri and Sen (1969), the desired result follows. In fact, it can be shown that (3.27) holds almost surely.

Lemma 3.2. *Under the assumptions of section 2, when* H_0 *holds,*

$$(3.31)\quad n^{-\frac{1}{2}}[\hat{\underset{\sim}{S}}_{n(2)} - \underset{\sim}{S}_{n(1)}(\underset{\sim}{\beta}_1, \underset{\sim}{0}) + \underset{\sim}{A}(\hat{\underset{\sim}{\beta}}_{1,n} - \underset{\sim}{\beta}_1)\underset{\sim}{C}_{n(12)}] \xrightarrow{p} 0$$

as $n \to \infty$, *where*

$$(3.32)\qquad \underset{\sim}{A} = \mathrm{Diag}(A_1, \ldots, A_p)\ .$$

The proof follows as a direct multivariate extension of Theorem 3.1 of Jurečková (1971), and hence, the details are omitted.

By noting that $\underset{\sim}{S}_{n(1)}(\hat{\underset{\sim}{\beta}}_{1,n}, 0) = \underset{\sim}{o}_p(n^{\frac{1}{2}})$. (see Jurečkova' (1971)), the following lemma also follow directly as a multivariate extension of Theorem 3.1 of Jurečkova' (1971).

<u>Lemma 3.3</u>. <u>Under the assumptions of section 2, when</u> H_0 <u>holds</u>,

$$(3.33) \quad n^{-\frac{1}{2}}[\underset{\sim}{S}_{n(1)}(\underset{\sim}{\beta}_1, \underset{\sim}{0}) - \underset{\sim}{A}(\hat{\underset{\sim}{\beta}}_{1,n} - \underset{\sim}{\beta}_1)\underset{\sim}{C}_{n(11)}] \overset{p}{\to} 0 ,$$

as $n \to \infty$.

Using Lemmas 3.2 and 3.3, we arrive at the following result.

<u>Lemma 3.4</u>. <u>Under</u> H_0 <u>in (1.3) and the assumptions of section 2</u>,

$$(3.34) \quad n^{-\frac{1}{2}}[\hat{\underset{\sim}{S}}_{n(2)} - \underset{\sim}{S}_{n(2)}(\underset{\sim}{\beta}_1, \underset{\sim}{0}) + \underset{\sim}{S}_{n(1)}(\underset{\sim}{\beta}_1, \underset{\sim}{0})\underset{\sim}{C}^{-1}_{n(11)}\underset{\sim}{C}_{n(12)}] \overset{p}{\to} 0 ,$$

<u>as</u> $n \to \infty$.

Consider now $H_0^* : \underset{\sim}{\beta} = \underset{\sim}{0}$. Then under $H_0 : \underset{\sim}{\beta}_2 = \underset{\sim}{0}$, the statistics $[\underset{\sim}{S}_{n(2)}(\underset{\sim}{\beta}_1, \underset{\sim}{0}), \underset{\sim}{S}_{n(1)}(\underset{\sim}{\beta}_1, \underset{\sim}{0})]$ have the same joint distribution as that of $\underset{\sim}{S}_n$ under H_0^* , and since the later is asymptotically multi-normal with mean vector $\underset{\sim}{0}$ and dispersion matrix

$$(3.35) \qquad \Lambda(F) \otimes \underset{\sim}{C}_n ,$$

it follows that under H_0 in (1.3),

$$(3.36)\quad \mathcal{L}(n^{-\frac{1}{2}}[\underset{\sim}{S}_{n(2)}(\underset{\sim}{\beta}_1,\underset{\sim}{0}) - \underset{\sim}{S}_{n(1)}(\underset{\sim}{\beta}_1,\underset{\sim}{0})\underset{\sim}{C}^{-1}_{n(11)}\underset{\sim}{C}_{n(12)}])$$

$$\rightarrow \mathfrak{N}_{p\times q_2}(\underset{\sim}{0},\ \underset{\sim}{\Lambda}(F)\otimes\underset{\sim}{C}_{(22)} - \underset{\sim}{C}_{(21)}\underset{\sim}{C}^{-1}_{(11)}\underset{\sim}{C}_{(12)}]) .$$

Hence using (3.34) and (3.36), under H_0 in (1.3), we find that

$$(3.37)\quad \mathcal{L}(n^{-\frac{1}{2}}\hat{S}_{n(2)}) \rightarrow \mathfrak{N}_{p\times q_2}(\underset{\sim}{0},\ \underset{\sim}{\Lambda}(F)\otimes\underset{\sim}{C}^{*})$$

From Lemma 3.1, (3.37) and the asymptotic distribution of quadratic forms associated with asymptotically multinormal vectors, it follows that (under H_0 in (1.3) and the conditions of section 2),

$$(3.38)\qquad \mathcal{L}(\mathcal{L}_N) \rightarrow \chi^2_{pq_2}\ ,\ \text{as}\ \ n\rightarrow\infty$$

Thus the proposed ADF test is as follows:

Reject H_0 if $\mathcal{L}_N \geq \chi^2_{pq_2,\alpha}$

Accept H_0 if $\mathcal{L}_N < \chi^2_{pq_2,\alpha}$

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

4. <u>Asymptotic comparison with parametric test</u>. Consider now a sequence $\{K_n\}$ of Pitman-type alternative hypotheses, viz.

(4.1) $\quad K_n : \underset{\sim}{\beta}_2 = \underset{\sim}{\beta}_2^{(n)} = n^{-\frac{1}{2}} \underset{\sim}{\gamma}_2 , \quad \underset{\sim}{\gamma}_2$ is fixed and non-null.

Our aim is to make the asymptotic power comparisons between the proposed rank order test and the normal theory likelihood ratio test when the underlying cdf is not necessarily multinormal. Proceeding as in Sen and Puri (1970). (Where the distribution theory of the normal theory likelihood ratio test for the general linear hypotheses is considered), it follows that if F possesses a finite second order moments, then (i) under H_0, the normal theory likelihood ratio statistic [actually-2 log (likelihood ratio statistic)], denoted by L_n has asymptotically a chi square distribution with pq_2 degrees of freedom, and (ii) under $\{K_n\}$, it has asymptotically a non central chi square distribution with pq_2 degrees of freedom and non-centrality parameter

(4.2) $$\Delta_L = \mathrm{Tr}[\bar{\Gamma} \cdot (\underset{\sim}{\Sigma}(F) \otimes \underset{\sim}{C}^*)^{-1}] ,$$

where

$$\bar{\Gamma} = ((\gamma_{jk}\gamma_{j'k'}))_{j,j'=1,\ldots,p\ ;\ k,k'=q_2+1,\ldots,q} ,$$

and

(4.3) $$\underset{\sim}{\Sigma}(F) = ((\sigma_{jj'}(F))) , \quad \sigma_{jj'}(F) = \mathrm{Cov}(X_{ji}, X_{j'i})$$

Consider now a sequence of alternatives $\{K_n^*\}$, where

$$(4.4) \qquad K_n^* : \underset{\sim}{\beta} = (\underset{\sim}{0},\ n^{-\frac{1}{2}}\underset{\sim}{\gamma}_2)$$

then, $(\underset{\sim}{S}_{n(1)}(\underset{\sim}{\beta}_1,\underset{\sim}{0})\ ,\ \underset{\sim}{S}_{n(2)}(\underset{\sim}{\beta}_1,\underset{\sim}{0}))$, under K_n , has the same joint distribution as that of $\underset{\sim}{S}_n$ under K_n^* . Noting this fact and using the results of Puri and Sen (1969), it follows that under K_n^* , $\underset{\sim}{S}_n$ has asymptotically a multinormal distribution with mean vector $\underset{\sim}{A}[\underset{\sim}{0},\underset{\sim}{\gamma}_2]\underset{\sim}{C} = \underset{\sim}{A}\underset{\sim}{\gamma}[\underset{\sim}{C}_{(21)},\ \underset{\sim}{C}_{(22)}]$, and dispersion matrix $\underset{\sim}{\Lambda}(F) \otimes \underset{\sim}{C}$. Thus, under $\{K_n\}$, as $n \to \infty$,

$$(4.5) \qquad \mathcal{L}(n^{-\frac{1}{2}}\hat{\underset{\sim}{S}}_{n(2)}) \to \mathcal{N}_{pq_2}(\underset{\sim}{A}\,\underset{\sim}{\gamma}\,\underset{\sim}{C}^*\ ,\ \underset{\sim}{\Lambda}(F) \otimes \underset{\sim}{C}^*)$$

Consequently

$$(4.6) \qquad \mathcal{L}(\mathfrak{L}_n | K_n) \to \chi^2_{pq_2,\Delta\mathfrak{L}}$$

where

$$(4.7) \qquad \Delta_{\mathfrak{L}} = \mathrm{Tr}[\bar{\Gamma} \cdot (\underset{\sim}{T}(F) \otimes \underset{\sim}{C}^*)^{-1}]$$

where $\underset{\sim}{T}(F)$ is given by (3.12).

From (4.2) and (4.7), we conclude that the Pitman Asymptotic Relative Efficiency (ARE) of $\mathfrak{L}_n$ with respect to L_n is

$$(4.8) \qquad e_{\mathfrak{L},L} = \Delta\mathfrak{L}/\Delta L = \mathrm{Tr}[\bar{\Gamma}(\underset{\sim}{T}(F) \otimes \underset{\sim}{C}^*)^{-1}]/\mathrm{Tr}[\bar{\Gamma}(\underset{\sim}{\Sigma}(F) \otimes \underset{\sim}{C}^{*-1})]$$

which depends on $\bar{\Gamma}, F$ and $\underset{\sim}{C}^*$. If F is a multinormal cdf and if we use the normal scores, then it can easily be checked that $\underset{\sim}{T}(F) = \underset{\sim}{\Sigma}(F)$ and hence $\Delta_{\mathfrak{L}} = \Delta_L$. In such a case the normal scores test and the normal theory likelihood ratio tests are asymptotically power equivalent. However, in general for arbitrary F, $\ell_{\mathfrak{L},L}$ is bounded by the minimum and maximum characteristic roots of $\underset{\sim}{\Sigma}(F)\underset{\sim}{T}^{-1}(F)$, i.e.

$$(4.9) \qquad \mathrm{Ch}_p[\underset{\sim}{\Sigma}(F)\underset{\sim}{T}^{-1}(F)] \leq \ell_{\mathfrak{L},L} \leq \mathrm{Ch}_1[\underset{\sim}{\Sigma}(F) \cdot \underset{\sim}{T}^{-1}(F)]$$

where ch_i is the ithe largest characteristic root. (The bounds of $\underset{\sim}{\Sigma}(F)\underset{\sim}{T}^{-1}(F)$ may be studied as in Sen and Puri (1967) or Puri and Sen (1969). Because of the similarity of the work, the details are omitted). In passing we may also remark that the $\mathfrak{L}_N$ test has asymptotically the best average power with respect to surfaces in the parameter space; it has also asymptotically the best constant power on such surfaces and finally it is asymptotically most stringent test. The proof follows as in Theorem 6.2 of Puri and Sen (1969).

5. ADF Tests for parallelism of regression surfaces.

Let $\underset{\sim}{X}_i^{(k)}$, $k = 1,\ldots,n_k$ be n_k independent rvs with continuous cdfs

(5.1) $F_i^{(k)}(\underset{\sim}{x}) = P[\underset{\sim}{X}_i^{(k)} \leq \underset{\sim}{x}] = F(\underset{\sim}{x} - \underset{\sim}{\alpha}_k - \underset{\sim}{\beta}_k \underset{\sim}{c}_i^{(k)})$,

$1 \leq i \leq n_k$, $k = 1,\ldots,s \cdot 2$

We desire to test the null hypothesis

(5.2) $H_0 = \underset{\sim}{\beta}_1 = \ldots = \underset{\sim}{\beta}_s = \underset{\sim}{\beta}$ (unknown)

Here the $\underset{\sim}{\beta}_k$'s are $p \times t$ matrices and the $\underset{\sim}{c}_i^{(k)}$ are t-vectors for some $t \geq 1$. A special case of $p = t = 1$ has been studied in detail in Sen (1969). If we let $\underset{\sim}{\beta}_k = \underset{\sim}{\beta}_1 + \underset{\sim}{\beta}_k^*$, $k = 1,\ldots,s$, (so that $\underset{\sim}{\beta}_1^* = \underset{\sim}{0}$) , $q = st$, then the result follows from the theory developed in section 3. Therefore, without going into the details of derivation, we briefly present the theory here.

For the kth sample {i.e. $\underset{\sim}{X}_i^{(k)}$, $i = 1,\ldots,n_k$} , define the $p \times t$ matrix $\underset{\sim}{S}_{nk}^{(k)}$ as in (2.2) and for every $\underset{\sim}{B} \in R^{pt}$, $\underset{\sim}{S}_{nk}^{(k)}(\underset{\sim}{B})$ as in (2.13)-(2.15). Let then

(5.3) $\underset{\sim}{\bar{S}}_n(\underset{\sim}{B}) = \sum_{k=1}^{s} \underset{\sim}{S}_{n_k}^{(k)}(\underset{\sim}{B})$, $n = \sum_{k=1}^{s} n_k$

Under H_0 , we estimate the common $\underset{\sim}{B}$ as follows: as in (3.8) and (3.9), we let

(5.4) $\underset{\sim}{D}_n = \left\{\underset{\sim}{B} \cdot \sum_{j=1}^{p} \sum_{s=1}^{t} |\bar{S}_{n,jr}(\underset{\sim}{b}_j)| = \min\right\}$,

(5.5) $\hat{\beta}_n$ = center of gravity of $\underset{\sim}{D}_n$.

Let then

$$(5.6) \qquad \hat{\underset{\sim}{S}}{}^{(k)}_{n_k} = S^{(k)}_{n_k}(\hat{\underset{\sim}{\beta}}_n) \ , \quad k = 1,\ldots,s \ ,$$

$$(5.7) \qquad \underset{\sim}{H}{}^{(k)}_{n_k} = ((\hat{S}^{(k)}_{n_k,jr}\ \hat{S}^{(k)}_{n_k,j's'}))_{j,j'=1,\ldots,p\ ;\ r,r'=1,\ldots,t}$$

$$(5.8) \qquad \underset{\sim}{C}{}^{(k)}_{n_k} = \sum_{i=1}^{n_k} [\underset{\sim}{c}{}^{(k)}_i - \bar{\underset{\sim}{c}}_{n_k}][\underset{\sim}{c}{}^{(k)}_i - \bar{\underset{\sim}{c}}_{n_k}]'$$

$$(5.9) \qquad \hat{\underset{\sim}{M}}_n = \left(\left(\sum_{k=1}^{s}\sum_{i=1}^{n_k}\{a^{(j)}_{n_k}(\hat{R}^{(k)}_{ji}) - \bar{a}^{(j)}_{n_k}\}\{a^{(j')}_{n_k}(\hat{R}^{(k)}_{j'i}) - \bar{a}^{(j')}_{n_k}\}/(n-s)\right)\right)$$

$$(5.10) \qquad \underset{\sim}{C}_{n_k} = \hat{\underset{\sim}{M}}_n \otimes \underset{\sim}{C}{}^{(k)}_{n_k} \ , \quad k = 1,\ldots,s \ .$$

where $\hat{R}^{(k)}_{ji}$ is the rank of $X^{(k)}_{ji} - \hat{\beta}_{n,j1}c^{(k)}_{i1},\ldots,\hat{\beta}_{n,jt}c^{(k)}_{it}$ among the n_k aligned observations on the jth variate in the kth sample, for $i = 1,\ldots,n_k$; $j = 1,\ldots,p$; $k = 1,\ldots,s$. The aligned rank order test statistic for testing H_0 in (5.2) is then

$$(5.11) \qquad \hat{\mathcal{L}}_N = \sum_{k=1}^{s} \mathrm{Tr}[\underset{\sim}{H}{}^{(k)}_{n_k}\underset{\sim}{G}{}^{-1}_{n_k}]$$

Under H_0 in (5.2), $\hat{\mathcal{L}}_N$ has asymptotically chi square distribution with $p(s-1)(t-1)$ degrees of freedom and under the sequence of alternatives $\{K_n\}$, where

$$(5.12) \qquad K_n : \underset{\sim}{\beta}_k = \underset{\sim}{\beta} + n^{-\frac{1}{2}} \gamma_k \ , \ k = 1,\ldots,s \ ; \ \sum_{k=1}^{t} \underset{\sim}{C}_{nk}^{(k)} \underset{\sim}{\gamma}_k = \underset{\sim}{0} \ ,$$

it has a non-centrality chi square distribution with $p(s-1)(t-1)$ degrees of freedom and non centrality parameter

$$(5.13) \qquad \Delta_{\hat{\mathfrak{L}}} = \sum_{k=1}^{S} \mathrm{Tr}[\bar{\Gamma}_k (\underset{\sim}{T}(F) \otimes \underset{\sim}{C}_k)^{-1}]$$

where

$$(5.14) \qquad \bar{\Gamma}_k = ((\gamma_{jr}^{(k)} \ \gamma_{j'r'}^{(k)})), \ 1 \le k \le S \quad \text{and} \quad \underset{\sim}{C}_k = \lim_{n\to\infty} n^{-1} \underset{\sim}{C}_{n_k}^{(k)}$$

which we assume to exist.

REFERENCES

[1] Hájek, J. (1968). Asymptotic normality of simple linear rank statistics under alternatives. Ann. Math. Statist 39, 325-346.

[2] Hoeffding, W. (1973). On the centering of a simple linear rank statistics. Ann. Statist. 1, 54-66.

[3] Jurecková, J. (1971). Non parametric estimates of regression coefficients. Ann. Math. Statist. 42, 1328-1338.

[4] 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.

[5] Puri, M. L. and Sen, P. K. (1973). A note on ADF tests in multiple linear regression. Ann. Statist. 1, 553-556.

[6] Sen, P. K. (1969). On a class of rank order tests for the parallelism of several refusion lines. Ann. Math. Statist. 40, 1668-1683.

[7] Sen, P. K. and Puri, M. L. (1967). On the theory of rank order tests for location in the multivariate one sample problem. Ann. Math. Statist. 38, 1216-1228.

[8] Sen, P. K. and Puri, M. L. (1969). On robust non-parametric estimation in some multivariate linear models. Multivariate Analysis, Vol. 2. Academic Press. 33-52.

[9] Sen, P. K. and Puri, M. L. (1970). Asymptotic theory of likelihood ratio and rank order tests in some multivariate linear models. Ann. Math. Statist. 41, 87-100.