ALIGNED RANK STATISTICS FOR REPEATED MEASUREMENT MODELS WITH ORTHONORMAL DESIGN, EMPLOYING A CHERNOFF-SAVAGE APPROACH

By J.H.J. Einmahl, B.O. Omolo, M.L. Puri, F.H. Ruymgaart

Discussion Paper

CentER

TILBURG ◆ UNIVERSITY

ALIGNED RANK STATISTICS FOR REPEATED MEASUREMENT MODELS WITH ORTHONORMAL DESIGN, EMPLOYING A CHERNOFF-SAVAGE APPROACH

J.H.J. Einmahl, B.O. Omolo, M.L. Puri, and F.H. Ruymgaart
Tilburg University, Texas Tech University, Indiana University

Abstract

In this paper aligned rank statistics are considered for testing hypotheses regarding the location in repeated measurement designs, where the design matrix for each set of measurements is orthonormal. Such a design may, for instance, be used when testing for linearity in a partially linear model. It turns out that the centered design matrix is not of full rank, and therefore doesn't quite satisfy the usual conditions in the literature. The number of degrees of freedom of the limiting chi-squared distribution of the test statistic under the null hypothesis, howerer, is not affected, unless rather special hypotheses are tested. An independent derivation of this limiting distribution is given, using the Chernoff-Savage approach. In passing it is observed that independence of the choice of aligner, which in the location problem is well-known to be due to cancellation, may in scale problems occur as a result of the type of score function suitable for scale tests. A possible extension to multivariate data is briefly indicated.

———————————

1 Introduction

In 1958 Chernoff and Savage published their landmark paper on asymptotic normality for a large class of rank statistics for two-sample problems. They established asymptotic normality under fixed alternatives (including

1

the null hypothesis) and proved this convergence to be uniform over a large class of alternatives so that asymptotic normality under local alternatives could be derived as a corollary. A few years later Hájek (1961, 1962) proved asymptotic normality of rank statistics of more general type under the null hypothesis as well as local alternatives employing LeCam's (1960) results on contiguity and local asymptotic normality, a very different technique. On the one hand, the latter method is very elegant; on the other it does not yield asymptotic normality under fixed alternatives - as the first method does - and this may be of interest in its own right.

Rank tests are only distribution free in a limited number of linear models. It is well-known, however, that for general linear models alignment can be applied to get rid of the nuisance parameter and to obtain asymptotically distribution-free procedures. This kind of result has been obtained by, for instance, Jurečková (1971), Kraft and van Eeden (1972) and Adichie (1978). These authors employed essentially Hájek's approach. It turns out that the limiting distribution of the test statistic does not depend on the choice of the aligners.

In this paper we want to apply the Chernoff-Savage method to deal with the asymptotics in the special case where the linear model has an orthonormal design and repeated measurements are given. It will be seen below that this set-up allows for testing linearity in a partially linear model (Eubank and Whitney (1989)), even when repetitions are not present but enough data are collected to do some grouping. In principle this approach could also provide the asymptotics under fixed alternatives, but in this paper we will restrict ourselves to the null hypothesis, although the basic asymptotics (Section 4 and Appendix) will be of a general nature. Because the error distribution is allowed to be heavy-tailed the aligners might be linear combinations of order statistics.

Following Adichie (1978) the statistic in this repeated measurement model turns out to be the difference of two quadratic forms of a vector with - not surprisingly - two-sample type components. It will be seen in Section 5 that the asymptotic distributions of these components depend on the choice of the aligner. This should not surprise either: it is more surprising that the limiting distribution of the aforementioned difference of quadratic forms does not depend on this choice. This is due to cancellation in the present model where the focus is on differences in location. It will also be seen, however, that when differences in scale were to be tested, independence of the choice of the aligner would already occur at the level of the components, due to the properties of the score functions suitable for scale problems (see Raghavachari (1965) for a related result).

Before proceeding with the general discussion let us give a precise formulation of the model. Our data will consist of n independent copies of a random vector $\mathbf{Y}$ of dimension $\nu \in \mathbb{N}\,(\nu \geqslant 2)$ satisfying

$$\mathbf{Y} = \mathbf{X}\theta + \varepsilon, \tag{1.1}$$

where $\mathbf{X}$ is a known $\nu \times \mu$ dimensional design matrix, with $2 \leqslant \mu \leqslant \nu, \theta \in \mathbb{R}^\mu$ an unknown parameter, and ε a random vector with ν independent and identically distributed components. These error random variables are of the continuous type but are not assumed to be normally distributed. If $\mathbf{x}_j = (x_{j1}, \ldots, x_{j\nu})^*, j = 1, 2, \ldots, \mu$, are the columns of $\mathbf{X}$ it will be assumed that

$$\mathbf{x}_1 = (\frac{1}{\sqrt{\nu}}, \ldots, \frac{1}{\sqrt{\nu}})^*, \text{ so that } ||\mathbf{x}_1|| = 1, \tag{1.2}$$

and that for all j and k

$$\mathbf{x}_j^* \mathbf{x}_k = \delta_{jk}, \tag{1.3}$$

so that the matrix $\mathbf{X} = [\mathbf{x}_1 \cdots \mathbf{x}_\mu]$ is orthonormal. Let $\mathbf{1}_r$ denote the vector of $r \in \mathbb{N}$ components 1, and $\mathbf{I}_r$ the $r \times r$ identity matrix. Statistically condition (1.2) entails the usual assumption that the ν locations contain a common component, and mathematically it follows that

$$\mathbf{1}_\nu{}^* \mathbf{x}_1 = \sqrt{\nu}; \quad \mathbf{1}_\nu{}^* \mathbf{x}_j = 0, \ j = 2, ..., \mu. \tag{1.4}$$

To describe the hypothesis to be tested let us write

$$\mathbf{H} = [\mathbf{x}_1 \cdots \mathbf{x}_p], \quad \mathbf{A} = [\mathbf{x}_{p+1} \cdots \mathbf{x}_\mu], \tag{1.5}$$

for some integer $1 \leqslant p < \mu, \ q = \mu - p$,

$$\eta = (\theta_1, ..., \theta_p)^*, \quad \text{and} \quad \alpha = (\theta_{p+1}, ..., \theta_\mu)^*, \tag{1.6}$$

so that

$$\mathbf{Y} = [\mathbf{H} \ \mathbf{A}] \begin{bmatrix} \eta \\ \alpha \end{bmatrix} + \varepsilon = \mathbf{H}\eta + \mathbf{A}\alpha + \varepsilon. \tag{1.7}$$

The null hypothesis to be tested is

$$\mathcal{H}_0 : \ \alpha = 0. \tag{1.8}$$

Remark 1. It is of some importance to note that the matrix $\mathbf{A}$ does not contain the special vector $\mathbf{x}_1$ in (1.2). Although it seems rather unrealistic not to include the common mean in the null hypothesis, it is a formal possibility. If $\mathbf{A}$ would contain $\mathbf{x}_1$, however, the limiting chi-squared distribution of the test statistic would have $q-1$ rather than q degrees of freedom (cf. (4.25)). To see this we also refer to Remark 3. This differs slighty from Adichie (1978). We will also see that the centered design matrices are not of full rank as required in Kraft and van Eeden (1972) and Adichie (1978). Our derivation of the limiting distribution of the test statistic will be independent, employing the Chernoff-Savage method.

Remark 2. Application of the transformation $\mathbf{X}^*$ to the model in (1.7) would reduce it to a canonical form:

$$\widetilde{\mathbf{Y}} = \mathbf{X}^*\mathbf{Y} = \begin{bmatrix} \eta \\ \alpha \end{bmatrix} + \mathbf{X}^*\varepsilon = \begin{bmatrix} \eta \\ \alpha \end{bmatrix} + \widetilde{\varepsilon}. \tag{1.9}$$

Given n independent copies of $\widetilde{\mathbf{Y}}$, testing (1.8) could be done using a completely distribution free Kruskal-Wallis test, based on the last q coordinates, if it were not for the transformed error vector. Because we do not assume the coordinates of ε to be normally distributed, the linear transformation $\mathbf{X}^*$ may induce considerable dependence among the coordinates of $\widetilde{\varepsilon}$, as the following example will show.

Example 1. Let us take $\nu = 2$ and assume that ε has independent standard Cauchy components $\varepsilon_1, \varepsilon_2$. Rotating ε over $\pi/4$ yields the vector $\tilde{\varepsilon}$. For this vector we have $P\{\tilde{\varepsilon}_1 \leq -5, \tilde{\varepsilon}_2 \leq -5\} = \pi^{-2} \int_{-\infty}^{-5} \int_{-\infty}^{-5} \{1 + \frac{1}{2}(x + y)^2\}^{-1}\{1 + \frac{1}{2}(x-y)^2\}^{-1}dxdy = 0.036$. This probability is almost 5 times as big as when computed under independence of the $\tilde{\varepsilon}_j$, in which case it equals $\{\pi^{-1} \int_{-\infty}^{-5} \sqrt{2}(2 + x^2)^{-1}dx\}^2 = 0.0077$. Here it is used that the $\tilde{\varepsilon}_j$ have the same (non-standard) Cauchy density.

Repeated measurement designs are quite common, particularly in ANOVA models. Aligned rank statistics for such models have been studied, for instance, by Hodges and Lehmann(1962) and Thompson(1991). This section will be concluded with an example where repeated measurements are not required in the strict sense, as that situation can be approximated by grouping.

Example 2: a partly linear model. Consider a linear function $l_{a,b}(x) = a + bx$ and a function $\varphi \in C([0,1]) \cap L^2([0,1])$ such that $\varphi \perp l_{a,b}$

in $L^2([0,1]), \forall\, a, b \in\ \mathbb{R}$. We observe

$$\widetilde{Y}_{ji} = l_{a,b}(\frac{j-1}{\nu} + \frac{i}{\nu n}) + \varphi(\frac{j-1}{\nu} + \frac{i}{\nu n}) + \varepsilon_{ji}, \qquad (1.10)$$
$$i = 1, \ldots, n,\ j = 1, \ldots, \nu,$$

where the ε_{ji} are i.i.d. and have density with respect to Lebesgue measure with median at 0. For sufficiently large ν we have that the $\widetilde{Y}_{j1}, \ldots, \widetilde{Y}_{jn}$ are a random sample from a density with median $a + b\frac{j}{\nu} + \varphi(\frac{j}{\nu})$, and where for $j = 1, \ldots, \nu$ these samples are independent. Also, if L is the linear subspace spanned by $(1, \ldots, 1)^*$ and $(\frac{1}{\nu}, \ldots, \frac{\nu}{\nu})^*$, the vector $(\varphi(\frac{1}{\nu}), \ldots, \varphi(\frac{\nu}{\nu}))^*$ will be approximately orthogonal to L. By orthonormalising the vectors spanning L we obtain the vectors

$$\mathbf{x_1} = (\frac{1}{\sqrt{\nu}}, \ldots, \frac{1}{\sqrt{\nu}})^*, \qquad (1.11)$$

$$\mathbf{x_2} = \Big\{ \sum_{j=1}^{\nu} \Big(\frac{2j-1-\nu}{2\nu}\Big)^2 \Big\}^{-1/2} \Big(\frac{1-\nu}{2\nu}, \ldots, \frac{\nu-1}{2\nu}\Big)^*. \qquad (1.12)$$

In this context it is natural to test whether the regression function consists of the linear part only. A lack-of-fit test then will lead to the null hypothesis $\mathcal{H}_0$ in (1.8) with $p = 2$ and $\mathbf{H} = [\mathbf{x}_1 \quad \mathbf{x}_2]$, where $\mathbf{x}_1$ and $\mathbf{x}_2$ are given above. The full model will be as in (1.1) with $\mathbf{X}$ a $\nu \times \nu$ matrix ($\mu = \nu$ in this case) with columns $\mathbf{x}_1, \ldots, \mathbf{x}_\nu$, where $\mathbf{x}_3, \ldots, \mathbf{x}_\nu$ is an orthonormal basis of $L^\perp$.

2 Construction of the aligned test statistic

The repeated measurements will be collected in the νn-vector of observations $\mathbf{Y}_n = (Y_{11}, \ldots, Y_{1n}, \ldots, Y_{\nu 1}, \ldots, Y_{\nu n})^*$ given by

$$\begin{aligned} Y_{ji} &= \sum_{k=1}^{\mu} x_{kj}\theta_k + \varepsilon_{ji} \\ &= m_j + \varepsilon_{ji}, \qquad i = 1, \ldots, n, j = 1, \ldots \nu, \end{aligned} \qquad (2.1)$$

where $\mathbf{m} = (m_1, \ldots, m_\nu)^*$ is the vector of medians. In matrix notation this boils down to

$$
\begin{aligned}
\mathbf{Y}_n &= \mathbf{X}_n\theta + \varepsilon_n = \\
&= \begin{bmatrix} \mathbf{1}_n & 0 & \cdots & 0 \\ 0 & \mathbf{1}_n & \cdots & 0 \\ . & . & \cdots & . \\ 0 & 0 & \cdots & \mathbf{1}_n \end{bmatrix} \mathbf{X}\theta + \varepsilon_n = \\
&= \begin{bmatrix} \mathbf{1}_n & 0 & \cdots & 0 \\ 0 & \mathbf{1}_n & \cdots & 0 \\ . & . & \cdots & . \\ 0 & 0 & \cdots & \mathbf{1}_n \end{bmatrix} [\mathbf{H}\ \mathbf{A}] \begin{bmatrix} \eta \\ \alpha \end{bmatrix} + \varepsilon_n = \\
&= [\mathbf{H}_n\ \mathbf{A}_n] \begin{bmatrix} \eta \\ \alpha \end{bmatrix} + \varepsilon_n.
\end{aligned}
\tag{2.2}
$$

The centered design matrix is defined as

$$
\widetilde{\mathbf{X}}_n = (\mathbf{I}_{\nu n} - \frac{1}{\nu n}\mathbf{1}_{\nu n}\mathbf{1}_{\nu n}^*)\mathbf{X}_n.
\tag{2.3}
$$

The first matrix on the right is a projection. It follows that

$$
\begin{aligned}
\widetilde{\mathbf{X}}_n^*\widetilde{\mathbf{X}}_n &= \\
&= \mathbf{X}^* \begin{bmatrix} \mathbf{1}_n^* & 0 & \cdots & 0 \\ 0 & \mathbf{1}_n^* & \cdots & 0 \\ . & . & \cdots & . \\ 0 & 0 & \cdots & \mathbf{1}_n^* \end{bmatrix} (\mathbf{I}_{\nu n} - \frac{1}{\nu n}\mathbf{1}_{\nu n}\mathbf{1}_{\nu n}^*) \begin{bmatrix} \mathbf{1}_n & 0 & \cdots & 0 \\ 0 & \mathbf{1}_n & \cdots & 0 \\ . & . & \cdots & . \\ 0 & 0 & \cdots & \mathbf{1}_n \end{bmatrix} \mathbf{X} = \\
&= n\mathbf{X}^*(\mathbf{I}_\nu - \frac{1}{\nu}\mathbf{1}_\nu\mathbf{1}_\nu^*)\mathbf{X} = n \begin{bmatrix} 0 & 0 & \cdots & 0 \\ 0 & & & \\ \vdots & & \mathbf{I}_{\mu-1} & \\ 0 & & & \end{bmatrix},
\end{aligned}
\tag{2.4}
$$

due to (1.2) and (1.3). Note that the matrix on the right in (2.4) is a projection of rank $\mu - 1$. Generalized inverses of this matrix are, for instance, this matrix itself (Moore-Penrose) or the identity $\mathbf{I}_\mu$. We choose (see Remark 3) the latter for easy calculations and thus obtain

$$
(\widetilde{\mathbf{X}}_n^*\widetilde{\mathbf{X}}_n)^- = \frac{1}{n}\mathbf{I}_\mu.
\tag{2.5}
$$

Similarly the centered null hypothesis matrix is given by

$$
\widetilde{\mathbf{H}}_n = (\mathbf{I}_{\nu n} - \frac{1}{\nu n}\mathbf{1}_{\nu n}\mathbf{1}_{\nu n}^*)\mathbf{H}_n,
\tag{2.6}
$$

and this yields

$$
\begin{aligned}
\widetilde{\mathbf{H}}_n^*\widetilde{\mathbf{H}}_n &= n\mathbf{H}^*(\mathbf{I}_\nu - \frac{1}{\nu}\mathbf{1}_\nu\mathbf{1}_\nu^*)\mathbf{H} = \\[2mm]
&= n\begin{bmatrix} \mathbf{x}_1^* \\ \vdots \\ \mathbf{x}_p^* \end{bmatrix}(\mathbf{I}_\nu - \frac{1}{\nu}\mathbf{1}_\nu\mathbf{1}_\nu^*)[\mathbf{x}_1 \ldots \mathbf{x}_p] = \\[2mm]
&= n\left(\mathbf{I}_p - \frac{1}{\nu}\begin{bmatrix} \sqrt{\nu} \\ 0 \\ \vdots \\ 0 \end{bmatrix}[\sqrt{\nu}\,0\ldots0]\right) = \\[2mm]
&= n\left[\begin{array}{c|ccc} 0 & 0 & \cdots & 0 \\ \hline 0 & & & \\ \vdots & & \mathbf{I}_{p-1} & \\ 0 & & & \end{array}\right],
\end{aligned}
\tag{2.7}
$$

where again this last matrix is a projection, but now of rank $p-1$. Choosing $\mathbf{I}_p$ for its generalized inverse we arrive at

$$
(\widetilde{\mathbf{H}}_n^*\widetilde{\mathbf{H}}_n)^- = \frac{1}{n}\,\mathbf{I}_p.
\tag{2.8}
$$

In order to get rid of the nuisance parameter η the data will be aligned and Y_{ji} in (2.1) will be replaced with the statistic $\hat{Y}_{ji}$. How this alignment is performed will be specified in Section 3 and is of no importance at this moment. Assuming there are no equals among the components of the vector $\hat{\mathbf{Y}}_n = (\hat{Y}_{11}, \ldots, \hat{Y}_{1n}, \ldots, \hat{Y}_{\nu 1}, \ldots, \hat{Y}_{\nu n})^*$, the vector of ranks will be denoted by $\hat{\mathbf{R}}_n = (\hat{R}_{11}, \ldots, \hat{R}_{1n}, \ldots, \hat{R}_{\nu 1}, \ldots, \hat{R}_{\nu n})^*$. (See, however, Ruymgaart (1980) for a simple way to deal with ties when the Chernoff-Savage method is used.) Let $J : [0,1] \to \mathbb{R}$ denote a score function with finite

$$
\mu_J = \int_0^1 J(t)dt, \quad \text{and} \quad \sigma_J^2 = \int_0^1 (J(t) - \mu_J)^2 dt.
\tag{2.9}
$$

Let us write

$$
\hat{\mathbf{w}}_n = (J(\frac{\hat{R}_{11}}{\nu n}), \ldots, J(\frac{\hat{R}_{1n}}{\nu n}), \ldots, J(\frac{\hat{R}_{\nu 1}}{\nu n}), \ldots, J(\frac{\hat{R}_{\nu n}}{\nu n}))^*,
\tag{2.10}
$$

and introduce the basic two-sample statistics

$$
W_j = \frac{1}{\sqrt{n}}\sum_{i=1}^n(\hat{w}_{ji} - \overline{w}), \overline{w} = \frac{1}{\nu n}\sum_{i=1}^{\nu n}J\left(\frac{i}{\nu n}\right), \mathbf{W} = (W_1, \ldots, W_\mu)^*.
\tag{2.11}
$$

The test statistic is the difference of two quadratic forms, the first of which equals

$$
\begin{aligned}
Q_1 &= \hat{\mathbf{w}}_n^* \widetilde{\mathbf{X}}_n (\widetilde{\mathbf{X}}_n^* \widetilde{\mathbf{X}}_n)^- \widetilde{\mathbf{X}}_n^* \hat{\mathbf{w}}_n = \\
&= \hat{\mathbf{w}}_n^* (\mathbf{I}_{\nu n} - \frac{1}{\nu n} \mathbf{1}_{\nu n} \mathbf{1}_{\nu n}^*)
\begin{bmatrix}
\mathbf{1}_n & 0 & \cdots & 0 \\
0 & \mathbf{1}_n & \cdots & 0 \\
. & . & \cdots & . \\
0 & 0 & \cdots & \mathbf{1}_n
\end{bmatrix} \mathbf{X} \\
&\times \frac{1}{n} \mathbf{I}_\mu \mathbf{X}^*
\begin{bmatrix}
\mathbf{1}_n^* & 0 & \cdots & 0 \\
0 & \mathbf{1}_n^* & \cdots & 0 \\
. & . & \cdots & . \\
0 & 0 & \cdots & \mathbf{1}_n^*
\end{bmatrix} (\mathbf{I}_{\nu n} - \frac{1}{\nu n} \mathbf{1}_{\nu n} \mathbf{1}_{\nu n}^*) \hat{\mathbf{w}}_n = \\
&= \mathbf{W}^* \mathbf{X} \mathbf{X}^* \mathbf{W} = \|\mathbf{X}^* \mathbf{W}\|^2 = \\
&= \|\mathbf{T}\|^2 = \sum_{j=1}^{\mu} T_j^2,
\end{aligned}
\tag{2.12}
$$

where

$$
\mathbf{T} = (T_1, \ldots, T_\mu)^* = \mathbf{X}^* \mathbf{W}.
\tag{2.13}
$$

The second quadratic form equals

$$
\begin{aligned}
Q_2 &= \hat{\mathbf{w}}_n^* \widetilde{\mathbf{H}}_n (\widetilde{\mathbf{H}}_n^* \widetilde{\mathbf{H}}_n)^- \widetilde{\mathbf{H}}_n^* \hat{\mathbf{w}}_n = \\
&= \hat{\mathbf{w}}_n^* (\mathbf{I}_{\nu n} - \frac{1}{\nu n} \mathbf{1}_{\nu n} \mathbf{1}_{\nu n}^*)
\begin{bmatrix}
\mathbf{1}_n & 0 & \cdots & 0 \\
0 & \mathbf{1}_n & \cdots & 0 \\
. & . & \cdots & . \\
0 & 0 & \cdots & \mathbf{1}_n
\end{bmatrix} \mathbf{H} \\
&\times \frac{1}{n} \mathbf{I}_p \mathbf{H}^*
\begin{bmatrix}
\mathbf{1}_n^* & 0 & \cdots & 0 \\
0 & \mathbf{1}_n^* & \cdots & 0 \\
. & . & \cdots & . \\
0 & 0 & \cdots & \mathbf{1}_n^*
\end{bmatrix} (\mathbf{I}_{\nu n} - \frac{1}{\nu n} \mathbf{1}_{\nu n} \mathbf{1}_{\nu n}^*) \hat{\mathbf{w}}_n = \\
&= (\mathbf{W}^* \mathbf{X})(\mathbf{X}^* \mathbf{H} \mathbf{H}^* \mathbf{X})(\mathbf{X}^* \mathbf{W}) = \\
&= (\mathbf{W}^* \mathbf{X})
\begin{bmatrix}
\mathbf{I}_p & \mathbf{O} \\
\mathbf{O} & \mathbf{O}
\end{bmatrix} (\mathbf{X}^* \mathbf{W}) = \\
&= \sum_{j=1}^{p} T_j^2.
\end{aligned}
\tag{2.14}
$$

Finally, the test statistic is given by

$$
Q = \frac{1}{\sigma_J^2}(Q_1 - Q_2) = \frac{1}{\sigma_J^2} \sum_{j=p+1}^{\mu} T_j^2.
\tag{2.15}
$$

Remark 3. The centering with $\overline{w}$ in (2.11) entails that $T_1 = 0$ and that the statistic Q is independent of the choice of the generalized inverses.

3   Further notation and assumptions

Our aim is to employ reasonable assumptions. We realize that some assumptions could be weakened at the price of more technicalities. Throughout the remainder of this paper $C \in (0, \infty)$ will denote a generic constant that does not depend on the sample size n or any other parameter.

The common c.d.f. F_0 of the errors ε_{ji} has a density f_0 with

$$F_0(0) = \frac{1}{2}, \ f_0(0) > 0, f_0' \text{ bounded and continuous on } \mathbb{R}. \tag{3.1}$$

Recall that $\mathbf{m} = (m_1, \ldots, m_\nu)^* = \mathbf{H}\eta + \mathbf{A}\alpha$ is the vector of medians in the general model (2.1). Let us write

$$\mathbf{m_H} = \mathbf{Pm}, \ \mathbf{m_A} = (\mathbf{I}_\nu - \mathbf{P})\mathbf{m}, \text{ where } \mathbf{P} = \sum_{j=1}^{p} \mathbf{x}_j \mathbf{x}_j^*, \tag{3.2}$$

is the projection onto the linear span of $\mathbf{x}_1, \ldots, \mathbf{x}_p$. Clearly the unobservable random variables

$$\check{Y}_{ji} = Y_{ji} - m_{\mathbf{H},j} = m_{\mathbf{A},j} + \varepsilon_{ji}, \tag{3.3}$$

have c.d.f.'s, respectively 'pseudo-empirical' c.d.f.'s

$$F_j(x) = F_0(x - m_{\mathbf{A},j}), \check{F}_j(x) = \frac{1}{n}\sum_{i=1}^{n} 1_{(-\infty,x]}(\check{Y}_{ji}), \tag{3.4}$$

$x \in \mathbb{R}$. Under the null hypothesis(1.8) we have $\mathbf{A}\alpha = \mathbf{m_A} = 0$ and hence these reduce to

$$F_j(x) = F_0(x), \ \ \check{F}_j(x) = \frac{1}{n}\sum_{i=1}^{n} 1_{(-\infty,x]}(\varepsilon_{ji}), . \tag{3.5}$$

Each median m_j can be estimated, for instance, by the median of the j-th sample or by a suitable linear combination of order statistics when f_0 is symmetric. As usual the choice of estimator does not matter provided only that the estimators $\hat{m}_j$, say, satisfy

$$M_j = \sqrt{n}(\hat{m}_j - m_j) = O_p(1), \text{ as } n \to \infty, \text{ for } j = 1, \ldots, \nu. \tag{3.6}$$

A natural estimator of the nuisance parameter $\mathbf{m_H}$ is:

$$\hat{\mathbf{m}}_{\mathbf{H}} = \mathbf{P}\hat{\mathbf{m}}, \ \hat{\mathbf{m}} = (\hat{m}_1, \ldots, \hat{m}_\nu)^*. \tag{3.7}$$

We are now ready to specify the aligned observations

$$
\begin{aligned}
\hat{Y}_{ji} &= Y_{ji} - \hat{m}_{\mathbf{H},j} = \\
&= \check{Y}_{ji} + m_{\mathbf{H},j} - \hat{m}_{\mathbf{H},j}, \ i = 1, \ldots, n, j = 1, \ldots, \nu.
\end{aligned}
\tag{3.8}
$$

The corresponding empirical c.d.f.'s will be written $\hat{F}_j$. We see from (3.4) that

$$\hat{F}_j(x) = \check{F}_j(x + \hat{m}_{\mathbf{H},j} - m_{\mathbf{H},j}), \ x \in \mathbb{R}. \tag{3.9}$$

We will also need respectively the pooled true, empirical, and pseudo-empirical c.d.f.'s

$$H = \frac{1}{\nu}\sum_{j=1}^{\nu} F_j, \ \ \hat{H} = \frac{1}{\nu}\sum_{j=1}^{\nu} \hat{F}_j, \ \ \check{H} = \frac{1}{\nu}\sum_{j=1}^{\nu} \check{F}_j. \tag{3.10}$$

It will next be assumed that the scores generating function $J = J^{(0)} : [0,1] \to \mathbb{R}$ has two continuous derivatives, so that

$$\max{}_{k=0,1,2} \ \max{}_{0\leq t\leq 1} |J^{(k)}(t)| \leq C. \tag{3.11}$$

4   Basic Asymptotics

In order to deal with the asymptotics in such a way that in principle fixed alternatives could be included we should center the W_j in (2.11) with $\int J(H)dF_j$ rather than with $\overline{w}$. Employing the well-known Chernoff-Savage representation we will then consider

$$W_j = \sqrt{n} \left\{ \int J(\hat{H})d\hat{F}_j - \int J(H)dF_j \right\}. \tag{4.1}$$

It should be noted that under H_0 this is asymptotically equivalent with the original centering because

$$
\sqrt{n} \left\{ \int J(\hat{H})d\hat{F}_j - \int J(H)dF_j \right\} = \tag{4.2}
$$

$$
= \sqrt{n} \left\{ \int J(\hat{H})d\hat{F}_j - \mu_J \right\} =
$$

$$
= \frac{1}{\sqrt{n}} \sum_{i=1}^{n} (\hat{w}_{ji} - \overline{w}) + O\left(\frac{1}{\sqrt{n}}\right).
$$

Applying integration by parts we arrive at the decomposition

$$W_j = S_j + B_j + r_j, \tag{4.3}$$

where $S_j = \sum_{k=1}^2 S_{jk}$, $B_j = \sum_{k=1}^2 B_{jk}$, $r_j = \sum_{k=0}^4 r_{jk}$, and

$$S_{j1} = -\sqrt{n} \int (\check{F}_j - F_j) J'(H) dH, \tag{4.4}$$

$$S_{j2} = \sqrt{n} \int (\check{H} - H) J'(H) dF_j, \tag{4.5}$$

$$B_{j1} = -(\mathbf{PM})_j \int f_j J'(H) dH, \tag{4.6}$$

$$B_{j2} = \frac{1}{\nu} \sum_{k=1}^{\nu} (\mathbf{PM})_k \int f_k J'(H) dF_j, \tag{4.7}$$

$$r_{j0} = \sqrt{n} \left[\int \{J(\hat{H}) - J(H)\} d\hat{F}_j - \int (\hat{H} - H) J'(H) dF_j \right], \tag{4.8}$$

$$r_{j1} = -\sqrt{n} \int (\check{F}_j - F_j) J'(H(\bullet - \frac{1}{\sqrt{n}} (\mathbf{PM})_j)) dH(\bullet - \frac{1}{\sqrt{n}} (\mathbf{PM})_j)$$
$$- S_{j1}, \tag{4.9}$$

$$r_{j2} = \frac{1}{\nu} \sum_{k=1}^{\nu} \sqrt{n} \int (\check{F}_k - F_k) J'(H(\bullet - \frac{1}{\sqrt{n}} (\mathbf{PM})_j)) dF_j(\bullet - \frac{1}{\sqrt{n}} (\mathbf{PM})_j)$$
$$- S_{j2}, \tag{4.10}$$

$$r_{j3} = -\sqrt{n} \int \{F_j(\bullet + \frac{1}{\sqrt{n}} (\mathbf{PM})_j) - F_j\} J'(H) dH - B_{j1}, \tag{4.11}$$

$$r_{j4} = \frac{1}{\nu} \sum_{k=1}^{\nu} \sqrt{n} \int \{F_k(\bullet + \frac{1}{\sqrt{n}} (\mathbf{PM})_k) - F_k\} J'(H) dF_j - B_{j2}, \tag{4.12}$$

Both under the null hypothesis and local alternatives the S_j are sums of independent random variables and mutually independent, so that their asymptotics can easily be dealt with. As we will see, the B_j will not appear in the test statistics so that they need not be considered. The remainder terms r_j converge to zero in probability, as will be sketched in the Appendix. Let us henceforth focus on the test statistics under the null hypothesis.

Under (1.8) we have (cf. $(3.3), (4.1)$)

$$S_{j1} =_d -\frac{1}{\sqrt{n}} \sum_{i=1}^{n} \int_0^1 \{1_{[0,t]}(U_{ji}) - t\} dJ(t) = \tag{4.13}$$

$$= \frac{1}{\sqrt{n}} \sum_{i=1}^{n} \{J(U_{ji}) - \mu_J\},$$

where the $U_{ji} = F_0(\varepsilon_{ji})$ are i.i.d. Uniform$(0,1)$ random variables. Furthermore we have that the vector $\mathbf{S}_1 + \mathbf{S}_2$ can be written as

$$\mathbf{S}_1 + \mathbf{S}_2 = \mathbf{S}_1 - \frac{1}{\nu} \mathbf{S}_1^* \mathbf{1}_\nu \mathbf{1}_\nu. \tag{4.14}$$

Next observe that under (1.8)

$$\int f_k J'(H) dF_j = \int f_0 dJ(F_0) = a_{J,0}. \tag{4.15}$$

The vector $\mathbf{B}_1 + \mathbf{B}_2$ can be written as

$$\mathbf{B}_1 + \mathbf{B}_2 = a_{J,0} \{ \mathbf{PM} - \frac{1}{\nu} (\mathbf{PM})^* \mathbf{1}_\nu \mathbf{1}_\nu \}. \tag{4.16}$$

It follows that (cf. (2.13))

$$\mathbf{T} = \mathbf{X}^* \mathbf{W} =$$
$$= \mathbf{X}^* (\mathbf{S}_1 - a_{J,0} \mathbf{PM}) - \frac{1}{\nu} (\mathbf{S}_1 - a_{J,0} \mathbf{PM})^* \mathbf{1}_\nu \mathbf{X}^* \mathbf{1}_\nu + o_p(1). \tag{4.17}$$

Let us introduce the $q \times \mu$-matrix

$$\mathbf{U} = \begin{bmatrix} \mathbf{O} \ \mathbf{I}_q \end{bmatrix}, \tag{4.18}$$

and note that

$$\mathbf{U}\mathbf{X}^* \mathbf{1}_\nu = \sqrt{\nu} \mathbf{U}\mathbf{X}^* \mathbf{x}_1 = \mathbf{U}(\sqrt{\nu}, 0, \dots, 0)^* = (0, \dots, 0)^*. \tag{4.19}$$

This entails that

$$\begin{aligned}
(T_{p+1}, \dots, T_\mu)^* &= \mathbf{U}\mathbf{X}^* \mathbf{W} = \\
&= \mathbf{U}\mathbf{X}^* (\mathbf{S}_1 - a_{J,0} \mathbf{PM}) \\
&\quad + \frac{1}{\nu} (\mathbf{S}_1 - a_{J,0} \mathbf{PM})^* \mathbf{1}_\nu \mathbf{U}\mathbf{X}^* \mathbf{1}_\nu + o_p(1) = \\
&= \mathbf{U}\mathbf{X}^* (\mathbf{S}_1 - a_{J,0} \mathbf{PM}) + o_p(1). \tag{4.20}
\end{aligned}$$

Finally, let us return to the quadratic form in (2.15), and for a suitable representation note that $\frac{\mathbf{MM}^*}{\|\mathbf{M}\|^2}$ is the projection $\mathbf{P_M}$, say, onto the line generated by the random vector $\mathbf{M}$. ($\mathbf{P_M}$ should not be confused with $\mathbf{PM}$.) In the usual ordering of semi-definite positive matrices we have $\mathbf{P_M} \leqslant \mathbf{I}_\nu$, and hence $\mathbf{AP_MA}^* \leqslant \mathbf{AA}^*$, and, consequently, $tr(\mathbf{AP_MA}^*) \leqslant tr(\mathbf{AA}^*)$ for any $\nu \times \nu$-matrix $\mathbf{A}$ (see, for instance, Rao and Toutenburg (1995)). This leads to (recall(3.2))

$$
\begin{aligned}
0 \leqslant \|\mathbf{UX}^*\mathbf{PM}\|^2 &= tr(\mathbf{M}^*\mathbf{PXU}^*\mathbf{UX}^*\mathbf{PM}) = \\
&= tr(\mathbf{UX}^*\mathbf{PMM}^*\mathbf{PXU}^*) \leqslant \\
&\leqslant \|\mathbf{M}\|^2 tr(\mathbf{UX}^*\mathbf{PXU}^*) \leqslant \\
&= \|\mathbf{M}\|^2 \sum_{j=1}^p tr(\mathbf{UX}^*\mathbf{x}_j\mathbf{x}_j^*\mathbf{XU}^*) = 0, \quad (4.21)
\end{aligned}
$$

and consequently

$$
(T_{p+1},\ldots,T_\mu)^* = \mathbf{UX}^*\mathbf{S}_1 + o_p(1). \tag{4.22}
$$

It follows from (4.11) that

$$
\mathbf{S}_1 \to_d N(0, \sigma_J^2 \mathbf{I}_\nu), \tag{4.23}
$$

and, because $\mathcal{E}\mathbf{UX}^*\mathbf{S}_1\mathbf{S}_1^*\mathbf{XU}^* = \sigma_J^2 \mathbf{I}_q$,

$$
(T_{p+1},\ldots,T_\mu)^* \to_d N(0, \sigma_J^2 \mathbf{I}_q). \tag{4.24}
$$

Theorem: asymptotic distribution of the test statistic under the null hypothesis. It follows from the results obtained above that under $\mathcal{H}_0$

$$
Q = \frac{1}{\sigma_J^2} \sum_{j=p+1}^\mu T_j^2 \to_d \chi_q^2, \text{ as } n \to \infty, \tag{4.25}
$$

for each error distribution F_0 satisfying the assumptions.

5 Some remarks and possible extensions

5.1 Location versus scale alternatives

The formal mathematical equivalent for scales of the location model in (2.1) is

$$
\begin{aligned}
Y_{ji} &= (\Pi_{k=1}^\nu \theta_k^{x_{kj}}) \varepsilon_{ji} = \\
&= (\Pi_{k=1}^\nu \eta_k^{x_{kj}})(\Pi_{k=p+1}^\nu \alpha_k^{x_{kj}}) \varepsilon_{ji} = \\
&= S_{\mathbf{H},j} \cdot S_{\mathbf{A},j} \cdot \varepsilon_{ji}, \quad i=1,\ldots,n, j=1,\ldots,\nu, \quad (5.1)
\end{aligned}
$$

where as before $\theta = (\eta, \alpha)^*$ but now $\theta_k > 0, k = 1, \ldots, \nu$, and

$$\mathcal{H}_0 : \alpha = \mathbf{1}_q, \text{ or } S_{\mathbf{A}} = \mathbf{1}_\nu. \tag{5.2}$$

The pseudo-observables are now of the form

$$\check{Y}_{ji} = \frac{1}{S_{\mathbf{H},j}} Y_{ji} = S_{\mathbf{A},j} \ \varepsilon_{ji}. \tag{5.3}$$

Estimation of $S_{\mathbf{H},j}$ yields the observables

$$\hat{Y}_{ji} = \frac{1}{\hat{S}_{\mathbf{H},j}} Y_{ji}. \tag{5.4}$$

Although we will not engage in an attempt to construct an asymptotically distribution free test for $\mathcal{H}_0$ in (5.2), we may assume it will be based on the components in (4.1), where the scores generating function J is now suitable for testing differences in scale. Typically, such functions are nonnegative and symmetric about the line $t = \frac{1}{2}$. Since the expansion in (4.3)-(4.12) will remain valid with $(\mathbf{PM})_j$ replaced by

$$\sqrt{n} \ (\frac{1}{\hat{S}_{\mathbf{H},j}} - \frac{1}{S_{\mathbf{H},j}}) = -\sqrt{n} \ \frac{\hat{S}_{\mathbf{H},j} - S_{\mathbf{H},j}}{\hat{S}_{\mathbf{H},j} S_{\mathbf{H},j}}, \tag{5.5}$$

the B-terms are similar in structure as before. However, if in addition to choosing J as above we assume f_0 to be symmetric about 0 it follows that

$$\int f_k J'(H) dF_j = \int f_0(F_0^{-1}(t)) J'(t) dt = 0, \tag{5.6}$$

under the current null hypothesis. This means that the B-terms simply are zero, so that the choice of aligner does not even play a role in the asymptotics of the components W_j. It should be noted that (5.6) does not hold true for scores generating functions J suitable for location, because such J are typically symmetric with respect to the point $(\frac{1}{2}, 0)$. In the location case, as we have seen, the B-terms cancel out in the final statistic, although they are not zero themselves. See Raghavachari (1965) for some scale tests when locations are unknown.

5.2   Multivariate data

A possible multivariate extension of the model in (1.7) is

$$\mathbf{Y} = \begin{bmatrix} Y_1^{(1)} & \cdots & Y_1^{(m)} \\ \vdots & \cdots & \vdots \\ Y_\nu^{(1)} & \cdots & Y_\nu^{(m)} \end{bmatrix} =$$

$$= \left[[\mathbf{HA}] \begin{bmatrix} \eta^{(1)} \\ \alpha^{(1)} \end{bmatrix} \cdots [\mathbf{HA}] \begin{bmatrix} \eta^{(m)} \\ \alpha^{(m)} \end{bmatrix} \right] + \begin{bmatrix} \varepsilon_1^{(1)} & \cdots & \varepsilon_1^{(m)} \\ \vdots & \cdots & \vdots \\ \varepsilon_\nu^{(1)} & \cdots & \varepsilon_\nu^{(m)} \end{bmatrix}, \qquad (5.7)$$

where $\mathbf{X} = [\mathbf{H} \ \mathbf{A}]$ is an orthonormal $\nu \times \mu$ matrix as before, $\eta^{(1)}, \ldots, \eta^{(m)} \in \mathbb{R}^p$ and $\alpha^{(1)}, \ldots, \alpha^{(m)} \in \mathbb{R}^q$ are unknown vector-parameters, and the $(\varepsilon_j^{(1)}, \ldots, \varepsilon_j^{(m)})^*, j = 1, \ldots, \nu$, are i.i.d. random vectors with a certain dependence structure.

The hypothesis to be tested is

$$\mathcal{H}_0 : \alpha^{(1)} = \ldots = \alpha^{(m)} = 0, \qquad (5.8)$$

and the actual data will consist of n independent copies $\mathbf{Y}_1, \ldots, \mathbf{Y}_n$ of $\mathbf{Y}$.

Following Roy's(1953) union-intersection principle choose $\mathbf{e} \in \mathbb{R}^m$ with $\|\mathbf{e}\| = 1$, and consider

$$
\begin{aligned}
\mathbf{Y_e} = \mathbf{Ye} \ &= \ \sum_{k=1}^m e_k [\mathbf{H} \ \mathbf{A}] \begin{bmatrix} \eta^{(k)} \\ \alpha^{(k)} \end{bmatrix} = \\
&= \ [\mathbf{H} \ \mathbf{A}] \begin{bmatrix} \sum_{k=1}^m e_k \eta^{(k)} \\ \sum_{k=1}^m e_k \alpha^{(k)} \end{bmatrix} = \\
&= \ [\mathbf{H} \ \mathbf{A}] \begin{bmatrix} \eta_\mathbf{e} \\ \alpha_\mathbf{e} \end{bmatrix}.
\end{aligned}
\qquad (5.9)
$$

Note that $\mathcal{H}_0$ in (5.8) is equivalent with $\bigcap_{\|\mathbf{e}\|=1} \mathcal{H}_\mathbf{e}$, where $\mathcal{H}_\mathbf{e}$ is the hypothesis that $\alpha_\mathbf{e} = 0$. This hypothesis can be tested using the quadratic form $Q_\mathbf{e}$, say, computed from the independent copies $\mathbf{Y}_{\mathbf{e},1}, \ldots, \mathbf{Y}_{\mathbf{e},n}$ of $\mathbf{Y}_\mathbf{e}$ in the manner described in the preceding sections. The question how these tests can be combined to obtain a test for the overall hypothesis (5.8) is beyond the scope of this paper. For a simple instance of this method see Buhrman and Ruymgaart (1981).

Appendix

Let $\mathcal{F}$ denote the class of all univariate c.d.f.'s F of the form $F(\bullet) = F_0(\bullet - \mu), \mu \in \mathbb{R}$, where F_0 satisfies assumption (3.1). Suppose that $\varepsilon_1, \ldots, \varepsilon_n$ are i.i.d. with common c.d.f. $F \in \mathcal{F}$. Denote the empirical process of the ε_i by

$$E_{n,F}(x) = \frac{1}{\sqrt{n}} \sum_{i=1}^n \{1_{(-\infty, x]}(\varepsilon_i) - F(x)\}, \ x \in \mathbb{R}. \qquad (A.1)$$

Lemma. Let $S_n = S_{n,F}$ denote a real-valued random variable such that $S_n = O_p(n^{-\frac{1}{2}})$, as $n \to \infty$, uniformly for $F \in \mathcal{F}$. Then we have

$$\sup_{x \in \mathbb{R}} |E_{n,F}(x + S_n) - E_{n,F}(x)| = O_p \left(\frac{(\log n)^{1/2}}{n^{1/4}} \right), \tag{A.2}$$

as $n \to \infty$, uniformly for $F \in \mathcal{F}$.

Proof. Using $F(x + S_n) = F(x) + S_n f(Y_n)$, with Y_n between S_n and x, we obtain, with α_n the empirical process of the $F(\varepsilon_i)$,

$$\sup_{x \in \mathbb{R}} |E_{n,F}(x + S_n) - E_{n,F}(x)| = \tag{A.3}$$
$$= \sup_{x \in \mathbb{R}} |\alpha_n(F(x + S_n)) - \alpha_n(F(x))| =$$
$$= \sup_{x \in \mathbb{R}} |\alpha_n(F(x) + S_n f(Y_n)) - \alpha_n(F(x))| =$$
$$= O_p \left(\frac{(\log n)^{1/2}}{n^{1/4}} \right),$$

where for the last transition the results for the oscillation modulus of the uniform empirical process in Stute (1982) are used.

Asymptotic neglibility of r_{j0}. After applying a Taylor expansion to the first term on the right in (4.6) it can be shown that

$$r_{j0} = \sqrt{n} \int (\hat{H} - H) J'(H) d(\hat{F}_j - F_j) + o_p(1), \tag{A.4}$$

uniformly for all $F_j \in \mathcal{F}$. The integral expression in (A.4) can be dealt with in a similar manner as in Ruymgaart *et al.* (1972, Corollary 5.5) by partitioning the real line in a finite number of subintervals.

Asymptotic negligibility of r_{j1} and r_{j2}. We will restrict ourselves to r_{j1} because r_{j2} can be dealt with similarly. Note that in the notation of $(A.1)$ and $(A.2)$

$$|r_{j1}| \leqslant \int |E_{n,F_j}(x + \frac{1}{\sqrt{n}}(\mathbf{PM})_j) - E_{n,F_j}(x)| \, |J'(H(x))| h(x) dx. \tag{A.5}$$

The random variable $n^{-\frac{1}{2}}(\mathbf{PM})_j$ satisfies the condition of S_n in the lemma. Furthermore, assumption (3.11) entails that $\sup_{x \in \mathbb{R}} \|J'(H(x))\| \leqslant C$, uniformly for all $F_j \in \mathcal{F}$. Application of the lemma yields

$$r_{j1} = O_p(n^{-1/4}(\log n)^{1/2}) \text{ as } n \to \infty, \tag{A.6}$$

uniformly for all $F_j \in \mathcal{F}$.

Asymptotic negligibility of r_{j3} and r_{j4}. Again let us restrict ourselves to one of the two, r_{j3}, since the other can be dealt with in a similar manner. The desired result follows at once from the expansion $F_j(x + (1/\sqrt{n})(\mathbf{PM})_j) = F_j(x) + (1/\sqrt{n})(\mathbf{PM})_j f_j(x) + (1/n)(\mathbf{PM})_j^2 f_j'(\tilde{x})$, where $\tilde{x}$ is a point between x and $(1/\sqrt{n})(\mathbf{PM})_j$, and the boundedness of f'.

References

[1] Adichie, J.N. (1978). Rank tests of sub-hypotheses in the general linear regression. *Ann.Math.Statist.* **6**, 1012-1026.

[2] Buhrman, H. and Ruymgaart, F.H. (1981). An application of linearization in nonparametric multivariate analysis. *Sankhyā A* **43**, 52-66.

[3] Chernoff, H. and Savage, I.R. (1958). Asymptotic normality and efficiency of certain nonparametric test statistics. *Ann. Math.Statist.* **29**, 972-994.

[4] Eubank, R.L. and Whitney, P. (1989). Convergence rates for estimation in certain partially linear models. *J. Statist. Pl. Inf.* **23**, 33-43.

[5] Hájek, J. (1961). Some extensions of the Wald-Wolfowitz-Noether theorem. *Ann.Math.Statist.* **32**, 506-523.

[6] Hájek, J. (1962). Asymptotically most powerful rank-order tests. *Ann.Math.Statist.* **33**, 1124-1147.

[7] Hodges, J.L., Jr. and Lehmann, E.L.(1962). Rank methods for combination of independent experiments in analysis of variance. *Ann.Math.Statist.* **33**, 482-497.

[8] Jurečková, J. (1971). Nonparametric estimates of regression coefficients. *Ann.Math.Statist.* **42**, 1328-1338.

[9] Kraft, C.H. and van Eeden, C. (1972). Linearized rank estimates and signed-rank estimates for the general linear hypothesis. *Ann. Math. Statist.* **43**, 42-57.

[10] LeCam, L. (1960). Locally asymptotically normal families of distributions. *University of California Publications in Statistics* **3**, 37-98.

[11] Raghavachari, M. (1965). The two-sample scale problem when the locations are unknown. *Ann. Math. Statist.* **36**, 1236-1242.

[12] Rao, C.R. and Toutenburg, H. (1995). *Linear Models.* Springer,New York.

[13] Roy, S.N. (1953). On a heuristic method of test construction and its use in multivariate analysis. *Ann. Math. Statist.* **24**, 220-238.

[14] Ruymgaart, F.H. (1980). A unified approach to the asymptotic distribution of certain midrank statistics. In: Statistique non Paramétrique Asymptotique (J.P. Raoult, ed.; pp 1-18). *Lecture Notes in Mathematics* **821**. Springer, New York.

[15] Ruymgaart, F.H., Shorack, G.R., and van Zwet, W.R. (1972). Asymptotic normality of nonparametric tests for independence. *Ann. Math. Statist.* **43**, 1122-1135.

[16] Stute, W. (1982). The oscillation behavior of empirical processes. *Ann. Probab.* **10**, 86-107.

[17] Thompson, G.L. (1991). A unified approach to rank tests for multivariate and repeated measures designs. *J. Amer. Statist. Assoc.* **86**, 410-419.

J.H.J. Einmahl
Department of Econometrics and Operations Research
Tilburg University
P.O. Box 90153
5000 LE Tilburg
THE NETHERLANDS

B.O. Omolo
Department of Mathematics and Statistics
Texas Tech University
Lubbock, TX 79409
USA

M.L. Puri
Department of Mathematics
Indiana University
Bloomington, IN 47405
USA

F.H. Ruymgaart
Department of Mathematics and Statistics
Texas Tech University
Lubbock, TX 79409
USA