CHARLES H. KRAFT
CONSTANCE VAN EEDEN

Linearized Rank Estimates and Signed-Rank Estimates for the General Linear Hypothesis

Publications des séminaires de mathématiques et informatique de Rennes, 1969-1970, fascicule 2
« Séminaire de probabilités et statistiques », , exp. n° 4, p. 1-31
<http://www.numdam.org/item?id=PSMIR_1969-1970___2_A4_0>

LINEARIZED RANK ESTIMATES AND SIGNED - RANK ESTIMATES FOR THE GENERAL LINEAR HYPOTHESIS

Charles H. KRAFT and Constance VAN EEDEN

Université de Montréal

1. INTRODUCTION.

The development of methods of estimation from ranks for the parameters of the general linear hypothesis has proceeded rapidly since the work of Hodges and Lehmann [5] on estimates for one - sample and two - sample problems. Univariate extensions of these estimates to k - sample problems have been given by Lehmann [11], and Bhuchongkul and Puri [2] ; to linear regression by Adichie [1] ; and to regression on monotone functions by Rao and Thornby [14]. Koul [7] studied rank estimates for a wide class of sequences of design matrices which are assumed to be perpendicular to a vector of constants. He used an approximation theorem of Jureckova [6] for some of the asymptotic properties. In [9], [10] the present authors utilized the theorem of Jureckova to study linearized versions of rank estimates for one - and two - , sample problems. These linearized versions are, in most cases, simpler to compute as well as asymptotically equivalent to the non - linearized versions.

In the present paper linearized rank estimates are described for a sub - class of the sequence of design matrices studied by Koul [7]. When Koul's estimates exist the estimates here can be considered as their linearized versions. However, the proofs given here do not require their existence. Linearized signed-rank estimates are given for an analogous sequence of designs and supposing the observations have symmetric distributions. Koul [8] has studied estimates based on signed - rank statistics for more general sequences of designs but with stronger assumptions on the distributions of the observations.

----

*) This work was partially supported by the Mathematics Research Center, United States Army, Madison, Contract # DA-31-124-ARD-D-462 and partially supported by the National Research Council of Canada. The manuseryt was written in final form while the authors were visiting the Department of Mathematics, University of Rennes.

The sequences of design matrices considered here have, at least asymptotically, fixed rank. Thus, the results do not apply to sequences of designs in which there are an increasing number of nuisance parameters as well as a fixed number of parameters of interest. Some of the recent results concerning rank estimates for these more complicated designs can be found in Lehmann [12] , Greenberg [3] , and Puri and Sen [13] .

The conditions under which it is shown here that linearization is possible for multiparameter problems are stronger than those proposed by Jureckova [6] . However the conditions here are notationally simpler and can be simpler to verify.

Section 2 contains the assumptions and theorems concerning estimates based on rank statistics. Section 3 contains the same for estimates based on signed - rank statistics. The results of these two sections require certain initial estimates and estimates of scale. Theorems establishing the existence and construction of such estimates are given in section 4. Section 5 contains the proofs of the theorems in section 2 and of those in section 4 concerning estimates based on rank statistics. Section 6 contains the proofs of the theorems in sections 3 and 4 concerning estimates based on signed - rank statistics.

The basis of linearized estimates is the fundamental theorem of Jureckova [6] . Section 7 gives a particular extension of this theorem to multiparameter problems for rank - statistics and a multiparameter extension of Van Eeden ' s [15] analogue, for signed - rank statistics, of Jureckova ' s theorem. In section 8 the relation between the extension to multiparameter problems of Jureckova ' s theorem used here and the extension suggested by her in [6] is discussed.

2. LINEARIZED RANK ESTIMATES

Suppose that, for each $\nu = 1, 2, \ldots,$ for an $n_\nu \times 1$ vector of observations $Y^{(\nu)} = (Y_1^{(\nu)}, \ldots, Y_{n_\nu}^{(\nu)})'$, there exists an $n_\nu \times (p + q)$ design matrix, $Z^{(\nu)}$, of known constants and a $(p + q) \times 1$ vector β of unknown constants such that the components of $Y^{(\nu)} - Z^{(\nu)}\beta$ are independently and identically distributed as $F(\frac{y}{b})$ $(b > o)$ where $F(y)$ is a completely specified distribution function. p and q will be fixed and limits will be as $\nu \longrightarrow \infty$. (Super- and subscripts ν will not be written).

The following standard reduction of the parameters will be convenient. For the sequence of design matrices, Z, let $Z - \bar{Z} = (Z_{ij} - \frac{1}{n} \sum_{i=1}^{n} Z_{ij})$ and let p be the rank of $Z - \bar{Z}$. Then, if $Z_1 - \bar{Z}_1$ is a set of p linearly independent columns of $Z - \bar{Z}$ and $Z_2 - \bar{Z}_2$ is the rest of the columns of $Z - \bar{Z}$, $Z - \bar{Z}$ can (after, if necessary, rearranging some of the columns) be written as $Z - \bar{Z} = (Z_1 - \bar{Z}_1 , Z_2 - \bar{Z}_2)$ where $Z_1 - \bar{Z}_1$ is of size $n \times p$ and rank p and $Z_2 - \bar{Z}_2 = (Z_1 - \bar{Z}_1)\, c$, where c is a $p \times q$ matrix. Hence $Z\beta = (Z_1 - \bar{Z}_1)(\beta_1 + c\beta_2) + (\bar{Z}_1 \beta_1 + \bar{Z}_2 \beta_2)$, where $\beta = (\beta_1' , \beta_2')'$ corresponds to $Z = (Z_1 , Z_2)$. Let $(Z_1 - \bar{Z}_1)(\beta_1 + c\beta_2) + (\bar{Z}_1 \beta_1 + \bar{Z}_2 \beta_2) = (Z_1 - \bar{Z}_1)\,\theta + \theta_o$ with θ_o a vector of constants and the Θ parameter to be estimated.

The distribution function F of single observations will be assumed to satisfy the regularity conditions of Hajek and Sidak [4], namely

Assumption A

i) $f(y) = \frac{dF(y)}{dy}$ exists and is absolutely continuous on $(-\infty,\infty)$

ii) the function $\varphi_F(u) = -\frac{f'}{f}(F^{-1}(u))$ can be written as the sum of two monotone functions each of which is square integrable on $o < u < 1$.

Let any two vectors u and v be called similarly ordered if $(u_i - u_j)(v_i - v_j) \geq o$ for all i, j. For the sequence Z of design matrices let $z = Z_1 - \bar{Z}_1$. It is supposed that the sequence $\{z = (z_{ij}^{(v)})\}$ satisfies

Assumption B

i) $$\frac{\max\limits_{1 \leq i \leq n} z_{ij}^2}{\sum\limits_{i=1}^{n} z_{ij}^2} \to 0 \qquad j = 1, \ldots, p$$

ii) $$\frac{1}{n} z' z \longrightarrow \Sigma \text{ where } \Sigma \text{ is positive definite.}$$

iii) For each j_1, j_2 ($j_1 \neq j_2$, j_1, $j_2 = 1, \ldots, p$) there exists a number $\gamma_{j_1 j_2} \neq 0$ such that, for $n > n_o$, z_{j_1} and $z_{j_1} + \gamma_{j_1, j_2} z_{j_2}$ are similarly ordered, where $z_1, \ldots, z_p$ are the column vectors of z.

Assumption C

It will be supposed that there exists a sequence $\hat{\theta}_1$ of initial estimates of θ which satisfies

i) $$\hat{\theta}_1\left(\frac{Y - z\theta}{a}\right) = \frac{\hat{\theta}_1(Y) - \theta}{a} \text{ for all } \theta \text{ and all } a > o$$

ii) $$P_\theta\left\{\sqrt{n}\,(\hat{\theta}_1 - \theta) \in A\right\} \longrightarrow P(A) \text{ for some fixed p-dimensional distribution } P.$$

Note that Ci) is satisfied for the least squares estimates $\hat{\theta}_1$ and that, under assumption Bi) and ii), Cii) will also be satisfied if $\int y^2 \, d\,F(y) < \infty$. In section 4 a class of designs is given for which a sequence $\hat{\theta}_1$ satisfying C can be constructed from certain medians.

Define now an n x 1 vector

$$\Phi_F(\theta) = \left\{ \varphi_F\left(\frac{R_{(Y - z\,\theta)_i}}{n+1} \right) \right\}$$

where $R_{(Y - z\,\theta)_i}$ is the rank of the i^{th} component of $Y - z\,\theta$ among all n components. A linearized rank estimate $\hat{\theta}$ will be defined by

$$(2,1) \qquad \hat{\theta} = \hat{\theta}_1 + \frac{\hat{b}}{K_{FF}} (z'\, z)^{-1}\, z'\, \Phi_F(\hat{\theta}_1),$$

where $K_{FF} = \int_0^1 \varphi_F^2(u)du$ and where $\hat{b}$ is a consistent estimate of the scale parameter b.

In section 5 the following theorem will be proved.

<u>Theorem 2 ; 1. If the components of $Y - Z\beta$ have common distribution function $F(\frac{y}{b})$, if F satisfies A, if z satisfies B, if $\hat{\theta}_1$ satisfies C and if $\hat{b}$ is a consistent estimate of b, then $\sqrt{n}\,(\hat{\theta} - \theta)$, where $\hat{\theta}$ is given by (2 ; 1), has asymptotically a normal distribution with mean zero and covariance $\frac{b^2}{K_{FF}} \Sigma^{-1}$.</u>

In order to find the asymptotic distribution of the estimate (2 ; 1) when the components of $Y - Z\beta$ are independently and identically distributed with a common distribution function G(y), the following assumption A_1 concerning G(y) and assumption D concerning the initial estimate $\hat{\theta}_1$ will be needed.

Assumption A_1

i) assumption Ai)

ii) $\int_0^1 \varphi_G^2(u)\,du < \infty$.

Let, for two distribution functions F_1 and F_2,

$K_{F_1F_2} = \int_0^1 \varphi_{F_1}(u)\ \varphi_{F_2}(u)\,du$ and call two sequences of estimates $\hat{t}_1$ and $\hat{t}_2$ G-equivalent if $P_G\{\sqrt{n}\ \|\hat{t}_1 - \hat{t}_2\| > \varepsilon\} \to o$. It will be supposed that the initial estimate $\hat{\theta}_1$ satisfies

Assumption D

i) $\hat{\theta}_1\left(\dfrac{Y - z\,\theta}{a}\right) = \dfrac{\hat{\theta}_1(Y) - \theta}{a}$ for all θ and all $a > o$

ii) if $\theta = o$, $\hat{\theta}_1$ is G-equivalent to $\dfrac{1}{K_{SG}}\ (z'\ z)^{-1} z'\ \Phi_S(o)$
for some distribution function S satisfying assumption A.

Theorem 2 ; 2. If the components of $y - Z\beta$ have common distribution function $G(y)$, if F and S satisfy A, if G satisfies A_1, if z satisfies B, if $\hat{\theta}_1$ satisfies D, then, for $\hat{\theta}$ defined by (2 ; 1), $\sqrt{n}(\hat{\theta} - \theta)$ has asymptotically a normal distribution with mean o and covariance

$$(2\ ;\ 2)\quad \left\{ \frac{K_{SS}}{K_{SG}^2}\left[1 - \frac{cK_{FG}}{K_{FF}}\right]^2 + \frac{2K_{SF}\,c}{K_{SG}K_{FF}}\left[1 - \frac{cK_{FG}}{K_{FF}}\right] + \frac{c^2}{K_{FF}} \right\} \Sigma^{-1}\ ,$$

where $c = P_G - \lim \hat{b}$.

In section 4 examples of initial estimates $\hat{\theta}_1$ satisfying assumption D will be given ; section 4 also contains a method of constructing estimates $\hat{b}$ which are consistent estimates of b if the components of $Y - Z\beta$ have distribution function $F(\frac{y}{b})$ and for which c can easily be found when the components of $Y-Z\beta$ have distribution $G(y)$.

In section 8 it is shown that the assumption Biii) can be replaced by an alternate assumption proposed by Jureckova [6].

3. LINEARIZED SIGNED-RANK ESTIMATES

Let now, for each $\nu = 1, 2, \ldots,$ $Y^{(\nu)} = (Y_1^{(\nu)}, \ldots, Y_{n_\nu}^{(\nu)})'$ be an $n_\nu \times 1$ vector of observations, let $Z^{(\nu)}$ be an $n_\nu \times (p_1 + q_1)$ design matrix and let β be a $(p_1 + q_1) \times 1$ vector of unknown constants such that the components of $Y^{(\nu)} - Z^{(\nu)}\beta$ are independently and identically distributed as $F(\frac{y}{b})$, where $F(y)$ is a completely specified distribution function. p_1 and q_1 will be fixed and limits are as $\nu \longrightarrow \infty$. Super- and subscripts ν will not be written.

Let p_1 be the rank of Z. Then Z can be written as $Z = (x, x_1)$, where x is a set of p_1 linearly independent columns of Z and $x_1 = x\, d$, where d is a $p_1 \times q_1$ matrix.

Let $\beta = (\beta'_3\, , \beta'_4)'$ correspond to $Z = (x\, , x_1)$ then $Z\beta = x(\beta_3 + d\, \beta_4)$. The parameter to be estimated is $\mu = \beta_3 + d\, \beta_4$.

Note that, in section 2, $Z\beta = (Z_1 - \bar{Z}_1\, , 1)\, (\theta_1, \ldots, \theta_p, \theta_o)'$, where $(Z_1 - \bar{Z}_1, 1)$ is the $n \times (p + 1)$ matrix consisting of the p columns of $Z_1 - \bar{Z}_1$ and a column of 1's ; this matrix $(Z_1 - \bar{Z}_1, 1)$ is of rank $p + 1$. The estimation procedure to be given in this section can thus be used to estimate the parameter $(\theta_1, \ldots, \theta_p\, , \theta_o)'$ of section 2. This leads to two different estimates for $(\theta_1, \ldots, \theta_p)'$ which, as will be seen, have asymptotically the same distribution if the underlying distributions are symmetric.

The distribution function F of single observations will be assumed to satisfy

Assumption A'

i) $f(y) = \dfrac{d\,F(y)}{dy}$ exists and is absolutely continuous on $(-\infty,\infty)$

ii) $\psi_F(u) \equiv \varphi_F(\frac{u+1}{2})$ can be written as the sum of two square integrable functions $\psi_1(u)$ and $\psi_2(u)$, where $\psi_1(u)$ is nondecreasing and nonnegative and $\psi_2(u)$ is nonincreasing and nonpositive

iii) $f(y) = f(-y)$ for all y.

For the sequence of design matrices it will be supposed that x satisfies

Assumption B'

i) $\dfrac{\max\limits_{1 \leq i \leq n} x_{ij}^2}{\sum\limits_{i=1}^{n} x_{ij}^2} \longrightarrow o$ for each $j = 1,\ldots, p_1$

ii) $\frac{1}{n} x'\, x \longrightarrow \sum_1$, where $\sum_1$ is a positive definite matrix.

iii) for each pair (j_1, j_2) $(j_1 \neq j_2 \,;\, j_1, j_2 = 1,\ldots, p_1)$

there exists a number $\gamma_{j_1 j_2} \neq o$ such that, for $n>n_o$,

$$\begin{cases} 1.\ x_{ij_1} (x_{ij_1} + \gamma_{j_1 j_2} x_{ij_2}) \geq o \text{ for all } i \\ 2.\ |x_{j_1}| \text{ and } |x_{j_1} + \gamma_{j_1 j_2} x_{j_2}| \text{ are similarly} \end{cases}$$

ordered, where $x_1,\ldots, x_{p_1}$ are column vectors of x

Assumption C'

It will be supposed that there exists a sequence of initial estimates $\hat{\mu}_1$ of μ satisfying

i) $\hat{\mu}_1(\frac{Y-x\mu}{a}) = \frac{\hat{\mu}_1(Y)-\mu}{a}$ for all μ and all $a>o$

ii) $P_\mu(\sqrt{n}(\hat{\mu}_1-\mu)\in A) \rightarrow P(A)$ for some fixed p_1-dimensional distribution P.

Let $\Psi_F(\mu)$ be the $n \times 1$ vector

$$\Psi_F(\mu) = \left\{ \psi_F\left(\frac{R|(Y-x\mu)_i|}{n+1}\right) \operatorname{sgn}(Y-x\mu)_i \right\},$$

where $R|(Y-x\mu)_i|$ is the rank of the absolute value of the i^{th} component $(Y-x\mu)_i$ of $Y-x\mu$ among the absolute values of all its components and

$$\operatorname{sgn} x = \begin{cases} 1 & \text{if } x>o \\ -1 & \text{if } x<o. \end{cases}$$

A linearized estimate $\hat{\mu}$ of μ will be defined by

$$(3;1) \qquad \hat{\mu} = \hat{\mu}_1 + \frac{\hat{b}}{K_{FF}} (x'x)^{-1} x' \Psi_F(\hat{\mu}_1) \ ,$$

where $\hat{b}$ is a consistent estimate of b.

In section 6 the following theorem will be proved.

<u>Theorem 3;1</u> : <u>If the components of $Y-Z\beta$ have common distribution function $F(\frac{y}{b})$, if F satisfies A', if x satisfies B', if $\hat{\mu}_1$ satisfies C', if $\hat{b}$ is a consistent estimate of b, then $\sqrt{n}(\hat{\mu}-\mu)$, with $\hat{\mu}$ given by (3;1), has asymptotically a normal distribution with mean o and covariance $\frac{b^2}{K_{FF}} \sum_1^{-1}$</u> .

In order to find the asymptotic distribution of the estimate (3;1) when the components of $Y-Z\beta$ are independently and identically distributed as $G(y)$, the following assumption A_1' concerning $G(y)$ and assumption D' concerning the initial estimate $\hat{\mu}_1$ are needed.

<u>Assumption A_1'</u>

i) assumption A'i)

ii) $\int_0^1 \varphi_G^2(u)\, du < \infty$

iii) assumption A'iii).

Assumption D'

i) $\hat{\mu}_1(\frac{Y-x\mu}{a}) = \frac{\hat{\mu}_1(Y)-\mu}{a}$ for all μ and all $a>o$

ii) if $\mu = o$, $\hat{\mu}_1$ is G-equivalent to $\frac{1}{K_{SG}} (x'x)^{-1} x' \Psi_S(o)$

for some distribution function S satisfying A'.

Theorem 3;2 : If the components of $Y-Z\beta$ have common distribution function $G(y)$, if F and S satisfy A', if G satisfies A_1', if x satisfies B', if $\hat{\mu}_1$ satisfies D' then, for $\hat{\mu}$ defined by (3;1), $\sqrt{n}(\hat{\mu}-\mu)$ has asymptotically a normal distribution with mean o and covariance

$$(3;2) \qquad \left\{ \frac{K_{SS}}{K_{SG}^2} \left[1-c\,\frac{K_{FG}}{K_{FF}}\right]^2 + \frac{2\,K_{SF}\,c}{K_{SG}\,K_{FF}} \left[1-c\,\frac{K_{FG}}{K_{FF}}\right] + \frac{c^2}{K_{FF}} \right\} \Sigma_1^{-1}$$

where $c = P_G - \lim \hat{b}$.

Examples of initial estimates $\hat{\mu}_1$ satisfying D' are given in section 4.

In section 8 it is shown that assumption B'iii) can be replaced by an alternate assumption.

4. INITIAL ESTIMATES OF Θ AND μ AND ESTIMATES OF THE SCALEPARAMETER b.

INITIAL ESTIMATES.

Perhaps the two most well known choices for initial estimates of Θ and μ are those corresponding to the mean and the median. The resulting relative efficiency of the linearized estimate can be found from Theorem 2;2 (resp. Theorem 3;2) if it is known that the initial estimate satisfies D (resp. D') for some Ψ_S. Identifying such initial estimates $\hat{\Theta}_1$ (resp. $\hat{\mu}_1$) and the corresponding Ψ_S is the purpose of the following four theorems which will be proved in section 5 for $\hat{\Theta}_1$ and in section 6 for $\hat{\mu}_1$.

Theorem 4;1 : If the components of $Y-Z\beta$ have common distribution function $G(x)$, where G satisfies A_1 and has a variance, if z satisfies Bi) and ii) then $\hat{\theta}_1 = (z'z)^{-1} z'Y$ satisfies D with $\varphi_S(u) = G^{-1}(u)$.

A construction of an initial estimate $\hat{\theta}_1$ corresponding to the median can be most easily described for replicated designs. Suppose $Z' = (Z_1', Z_2', \ldots, Z_n')$ where, for each i, $Z_i = Z_o$ where Z_o is a $k \times (p+q)$ matrix. Let z_o span $Z_o - \bar{Z}_o$ so that $z_o' z_o > o$. Then z_o is $k \times p$ so that the total number of observations is nk. For simplicity suppose that the k rows of z_o are distinct. Then the n observations corresponding to a given row in z_o are a sample from a population with the same location (If z_o has some equal rows there will be available more observations for a given "row") Let $m = (m_1, m_2, \ldots, m_k)'$ be the medians of the observations corresponding to each of the k rows of z_o.

Theorem 4;2 : If the components of $Y-Z\beta$ have common distribution function $G(x)$, where $G(x)$ satisfies A_1 and has a positive density at its median, then $\hat{\theta}_1 = (z_o' z_o)^{-1} z_o m$ satisfies D with S the double exponential distribution.

The corresponding statement for an initial estimate, based on the mean, of μ is Theorem 4;3.

Theorem 4;3 : If the components of $Y-Z\beta$ have common distribution $G(x)$, where $G(x)$ satisfies A_1' and has a variance, if x satisfies B'i) and ii), then $\hat{\mu}_1 = (x'x)^{-1} x'Y$ satisfies D' with $\psi_S(u) = G^{-1}(\frac{u+1}{2})$.

For an initial estimate $\hat{\mu}_1$ based on medians, consider again an n-times repeated fixed design matrix. Let $x = (x_o', \ldots, x_o')'$ with x_o a $k \times p_1$ matrix and $x_o' x_o > o$. Let $t = (t_1, t_2, \ldots, t_k)'$ be the medians of the observations corresponding to each of the k rows of x_o.

Theorem 4;4 : If the components of $Y-Z\beta$ have common distribution function $G(x)$, where G satisfies A_1' and has a positive density at its median, then $\hat{\mu}_1 = (x_o' x_o)^{-1} x_o' t$ satisfies D' with S the double exponential distribution.

Estimates of the scale parameter

Estimates of the scale parameter b can e.g. be obtained as follows. Most measures of dispersion D, defined for distribution functions H, H_n on $(-\infty,\infty)$, have the following properties

i) $b\, D(H(y)) = D(H(\frac{y-a}{b}))$ for all a and all $b>o$

ii) $D(H_n(y)) \to D(H(y))$ whenever $\sup_y |H_n(y)-H(y)| \to o$ and $D(H(y))<\infty$.

Given such a measure of dispersion D, $\hat{b}$, in section 2, can be taken as $\frac{D(\hat{F}_n(y))}{D(F(y))}$, where $\hat{F}_n(y)$ is the empirical distribution function of the components of $Y-z\hat{\theta}_1$ and $F(y)$ is the distribution function from which $\varphi_F(u)$ is computed. Then, if the components of $Y-Z\beta$ have common distribution $F(\frac{y}{b})$, if $\hat{\theta}_1$ satisfies C and if $D(F(y))<\infty$, $\hat{b}$ is a consistent estimate of b. If the components of $Y-Z\beta$ have common distribution $G(y)$, if $\hat{\theta}_1$ satisfies D, if $D(F(y))<\infty$ and $D(G(y))<\infty$ then, in Theorem 2;2, $c = \frac{D(G(y))}{D(F(y))}$. The same remarks hold for estimating b in section 3.

D can be taken, for instance, as an interpercentile range or, if the observations have a variance, as the standard deviation.

In [10] some numerical values of the relative efficiencies of linearized estimates are given ; these relative efficiencies are computed as the ratio of the Cramér-Rao lower bound $\frac{1}{\int_o^1 \varphi_G^2(u)\,du}$, for the estimation problem, to

$$\frac{K_{SS}}{K_{SG}^2}\left[1-c\,\frac{K_{FG}}{K_{FF}}\right]^2 + \frac{2\,K_{SF}\,c}{K_{SG}\,K_{FF}}\left[1-c\,\frac{K_{FG}}{K_{FF}}\right] + \frac{c^2}{K_{FF}} .$$

These efficiencies are given in [10] for several choices of F and G, for $\hat{b}$ as the standard deviation or as the interquartile range, and for both choices of the initial estimate given above.

5 - PROOF OF THEOREM 2;1, 2;2, 4;1, and 4;2.

Proof of Theorem 2;1.

Since $\hat{b}$ is a consistent estimate of b, it is sufficient to prove that, for the estimate (2;1) with $\hat{b}$ replaced by b, the distribution of $\sqrt{n}(\hat{\theta} - \theta)$ converges to a normal distribution with mean zero and covariance $\frac{b^2}{K_{FF}} \Sigma^{-1}$.

The asymptotic distribution of $\sqrt{n}\,(\hat{\theta} - \theta)$ with $\hat{b}$ replaced by b can be found as follows.

a) For $c = (c_1, \dots, c_p)' \neq 0$ and $\theta = 0$ it follows from Hajek and Sidak [4] (p. 163) that $\frac{b}{\sqrt{n}\, K_{FF}} c'z' \Phi_F(0)$ is asymptotically normal with mean zero and variance $\frac{b^2}{K_{FF}} c'\Sigma c$ provided that c'z' satisfies Bi) and Bii). That it does if z satisfies Bi) and Bii) is immediate upon noting that, for

$$\frac{\frac{1}{n} \max\limits_{1\leq i\leq n} \left(\sum\limits_{j=1}^{p} c_j z_{ij}\right)^2}{\frac{1}{n} \sum\limits_{i=1}^{n} \left(\sum\limits_{j=1}^{p} c_j z_{ij}\right)^2}$$

Bii) implies that the denominator converges to $c'\,\Sigma\, c > 0$. Hence, by taking $c' = (0, \dots, 0, 1, 0, \dots, 0)$, it follows from Bi) that $\frac{1}{n} \max\limits_{1\leq i\leq n} z_{ij}^2$ approaches zero for each j.

But $\max\limits_{1\leq i\leq n} \left(\sum\limits_{j=1}^{p} c_j z_{ij}\right)^2 \leq M^2 p^2 \max\limits_{1\leq j\leq p} \max\limits_{1\leq i\leq n} z_{ij}^2$, where $M^2 = \max\limits_{1\leq j\leq p} c_j^2$.

b) It follows from Ci) that $\hat{\theta}(Y - z\,\theta) = \hat{\theta}(Y) - \theta$ so we can suppose that $\theta = 0$. If $\hat{\theta}_\circ$ is defined by $\frac{b}{K_{FF}} (z'z)^{-1} z' \Phi_F(0)$ it is immediate from a) that $\sqrt{n}\,\hat{\theta}_\circ$ is asymptotically normal with mean 0 and covariance $\frac{b^2}{K_{FF}} \Sigma^{-1}$.

c) Assuming $\theta = 0$ it remains to show that $\sqrt{n}\, \|\hat{\theta} - \hat{\theta}_\circ\|$ converges to zero and hence that $\sqrt{n}\,\hat{\theta}$ and $\sqrt{n}\,\hat{\theta}_\circ$ have asymptotically the same distribution. However

$$(5;1) \qquad \sqrt{n}\,\|\hat{\Theta}-\hat{\Theta}_o\| = \|\sqrt{n}\,\hat{\Theta}_1 + \frac{b\sqrt{n}}{K_{FF}}(z'z)^{-1}\{z'\,\Phi_F\,(\hat{\Theta}_1) - z'\,\Phi_F(o)\}\|\cdot$$

By (ii) a number d can be chosen so that $P\{\|\hat{\Theta}_1\| \le \frac{d}{\sqrt{n}}\}$ is arbitrarily close to one for all sufficiently large n. Hence the right hand side of (5;1) will be, with arbitrarily high probability, bounded by

$$\sup_{\|\xi\| \le \frac{d}{\sqrt{n}}} \|\sqrt{n}\,\xi + \frac{b\sqrt{n}}{K_{FF}}(z'z)^{-1}\{z'\,\Phi_F(\xi) - z'\,\Phi_F(o)\}\|\ ,$$

which can also be written as

$$\sup_{\|\xi\| \le \frac{d}{\sqrt{n}}} \|\frac{\sqrt{n}\,(z'z)^{-1}}{K_{FF}}\ \{z'\,\Phi_F(\xi) - z'\,\Phi_F(o) + z'z\,\frac{K_{FF}}{b}\,\xi\}\|$$

Further, by an extension of the theorem of Jureckova [6], (see section 7 and 8),

$$\sup_{\|\xi\| \le \frac{d}{\sqrt{n}}} \|\frac{1}{\sqrt{n}}\ \{z'\,\Phi_F\,(\xi) - z'\,\Phi_F\,(o) + \frac{K_{FF}}{b}\,z'z\,\xi\}\|$$

converges to zero in probability if $\Theta = o$. Since $n\,(z'z)^{-1} \longrightarrow \Sigma^{-1}$, it follows that $\sqrt{n}\ \|\hat{\Theta} - \hat{\Theta}_o\| \longrightarrow 0$. This completes the proof.

Proof of Theorem 2;2

As in the proof of Theorem 2;1, we can suppose that $\Theta = o$. By the extension of the theorem of Jureckova [6] (see section 7 and 8) we have, for $\Theta = o$,

$$(5;2) \qquad P_G\left\{\sup_{\|\xi\| \le \frac{d}{\sqrt{n}}} \|\frac{1}{\sqrt{n}}(z'\,\Phi_F(\xi) - z'\,\Phi_F(o) + K_{FG}\,z'z\xi)\| > \varepsilon\right\} \to o$$

If $\hat{\Theta}_{oo} = (1-c\,\frac{K_{FG}}{K_{FF}})\,\hat{\Theta}_1 + \frac{c}{K_{FF}}(z'z)^{-1}\,z'\,\Phi_F(o)$ and $\hat{\Theta}_{o1} = \frac{1}{K_{SG}}(z'z)^{-1}\,z'\,\Phi_S(o)$ it follows from (5;2) and the fact that $\hat{b} \xrightarrow{P_G} c$, as in the proof of Theorem 2;1, that $P_G\,\{\sqrt{n}\,\|\hat{\Theta}_{oo} - \hat{\Theta}\| > \varepsilon\} \to o\cdot$

Further, by assumption D, $P_G\ \{\sqrt{n}\ \|\hat{\theta}_{o1} - \hat{\theta}_1\| > \varepsilon\} \to o$. Hence the asymptotic distribution of $\sqrt{n}\ \hat{\theta}$ is that of $\sqrt{n}\ \hat{\theta}_{o2}$, where

$$\hat{\theta}_{o2} = (1-c\ \frac{K_{FG}}{K_{FF}})\ \frac{(z'z)^{-1}}{K_{SG}}\ z'\ \Phi_S(o) + \frac{c}{K_{FF}}\ (z'z)^{-1}\ z'\ \Phi_F(o).$$

It follows from Hajek and Sidak [4] (p. 163) that the asymptotic distribution of $\sqrt{n}\ \hat{\theta}_{o2}$, and hence that of $\sqrt{n}\ \hat{\theta}$, is normal with mean o and covariance given by (2;2). Q.E.D.

Proof of Theorem 4;1.

Obviously , $\hat{\theta}_1$ satisfies Di). Further $G^{-1}(u)$ is nondecreasing in u and $\int_0^1\ (G^{-1}(u))^2\ du = \int_{-\infty}^{+\infty} y^2\ g(y)\ dy < \infty$ so that S satisfies A if G satisfies A_1 and has a variance. Further it follows from Hajek and Sidak [4] (p. 160) that $(z'z)^{-1}\ z'\ \Phi_S(o)$ is, if $\theta = o$, G-equivalent to

$$(z'z)^{-1}\ z'\ (\ \varphi_S(G(Y_1)),\ \ldots,\ \varphi_S\ (G(Y_n))' = (z'z)^{-1}\ z'\ Y.$$

Since $K_{SG} = 1$ the result follows.

For the proof of Theorem 4;2. the following lemma is needed.

Lemma 5;1.

If the components of $Y-Z\theta$ have common distribution function $G(x)$, where G satisfies A_1 and has a positive density at its median η , then, for $\theta = o$, each median m_j is G-equivalent to $\eta + \frac{1}{nK_{SG}}\ \delta_j$ where S is the double exponential distribution and where δ_j is the sum of ±1's according as the observations corresponding to the j^{th} row of z_o are $\gtrless \eta$.

Proof : It is sufficient to show that, assuming $\eta = 0$,

$$\mathcal{E}_G\ \left\{[\ \sqrt{n}\ (2g(o)\ m_j - \frac{\delta_j}{n})]^2\ \Big|\ m_j\right\} \xrightarrow{P_G} o \text{ since } K_{SG} = 2g(o) > o.$$

Let n_j^* be the number of observations corresponding to the j^{th} row of z_o which are between o and m_j. Then $\delta_j = \pm 2n_j^*$ according as $m_j \gtrless o$. The conditional, given m_j, distribution of n_j^* is B $(\frac{n}{2}, p_j)$ where

$$p_j = \frac{|G(m_j) - G(o)|}{G(m_j)} \quad . \text{ Hence}$$

$$\mathcal{E}_G \left\{ n\left[2g(o)\, m_j - \frac{\delta_j}{n}\right]^2 \Big|\, m_j \right\} = n\left[2g(o)\,|m_j| - p_j\right]^2 + 2p_j(1-p_j)$$

which can be writen as

$$\frac{n\, m_j^2}{G^2(o)} \left\{ g(o) - \frac{|G(m_j) - G(o)|}{|m_j|} \cdot \frac{G(o)}{G(m_j)} \right\}^2 + 2p_j\ (1-p_j).$$

Since $\sqrt{n}\, m_j$ has an asymptotic distribution and

$$\frac{|G(m_j) - G(o)|}{|m_j|} \xrightarrow{P_G} g(o) \text{ the result follows.}$$

<u>Proof of Theorem 4;2.</u>

Since, for the double exponential distribution,

$$\varphi_S(u) = \begin{cases} 1 \text{ if } u > 1/2 \\ -1 \text{ if } u < 1/2 \end{cases}$$

$\Phi_S(o)$ is a vector of $\pm$ 1's according as $Y_i \gtrless \text{med}\ (Y_1,\ldots,Y_{nk})$. Letting $\delta' = (\delta_1,\ldots\ldots,\delta_k)$, with δ_j as in Lemma 5;1, it follows from the lemma that $\hat{\Theta}_1$ is G-equivalent to (note that $z_o'\,\eta = 0$)

$$\frac{1}{K_{SG}} \quad \frac{(z_o' z_o)^{-1}}{n} \quad z_o'\,\delta. \text{ However } z_o'\,\delta = z'\Delta \text{ where } \Delta \text{ is an } nk \times 1$$

vector of $\pm$ 1's according as $Y_i \gtrless \eta$. The conclusion of the theorem will follow if it is true that

$$(5;3) \qquad \frac{1}{\sqrt{nk}} \quad \| z' (\Delta - \Phi_S(o)) \| \xrightarrow{P_G} o.$$

From Hajek and Sidak [4] (p. 61) it follows that the conditional, given Y, expectation of the square of each element of $z'(\Delta - \Phi_S(o))$ is bounded by

$$4 \left\{ \frac{\#\text{ of } Y_i \text{ between } \eta \text{ and } M}{nk} \right\} \sum_{i=1}^{nk} z_{ij}^2 ,$$

where $M = \text{med}\,(Y_1, \ldots, Y_{nk})$. Since $\frac{1}{nk} \sum_{i=1}^{nk} z_{ij}^2 \longrightarrow \Sigma_{ij}$, (5;3) and the theorem follow.

6. PROOF OF THEOREM 3;1, 3;2, 4;3, and 4;4.

The following proofs of theorem 3;1 and 3;2 are the analogues for signed rank statistics to those of theorem 2;1 and 2;2 for rank statistics. Accordingly they require a linearization theorem for signed rank statistics. Such a linearization theorem has, for $p_1 = 1$, been given in [15] ; for the extension to $p_1 > 1$ see section 7 and 8.

Proof of Theorem 3;1.

Since $\hat{b}$ is a consistent estimate of b, it is sufficient to prove that, for the estimate (3;1) with $\hat{b}$ replaced by b, $\sqrt{n}\,(\hat{\mu}-\mu)$ has asymptotically a normal distribution with mean o and covariance $\frac{b^2}{K_{FF}}\ \Sigma_1^{-1}$.

The asymptotic distribution of $\sqrt{n}\,(\hat{\mu}-\mu)$ with $\hat{b}$ replaced by b can be found as follows.

a) For $c = (c_1, \ldots, c_{p_1})' \neq 0$ and $\mu = o$, it follows from Hajek and Sidak [4] (p. 166) and the assumptions B'i) and ii) that $\frac{b}{\sqrt{n}\,K_{FF}}\ c'\ x'\ \Psi_F(o)$ is asymptotically normal with mean o and variance $\frac{b^2}{K_{FF}}\ c'\ \Sigma_1\ c.$

b) It follows from C'i) that $\hat{\mu}(Y-x\mu) = \hat{\mu}(Y)-\mu$ so we can suppose $\mu = o$. With $\hat{\mu}_o$ defined by $\frac{b}{K_{FF}} (x'x)^{-1} x' \Psi_F(o)$ it follows from a) that $\sqrt{n}\ \hat{\mu}_o$ has asymptotically a normal distribution with mean o and covariance $\frac{b^2}{K_{FF}} \Sigma_1^{-1}$.

c) Assuming $\mu = o$, it remains to show that $\sqrt{n}\ \|\hat{\mu}-\hat{\mu}_o\|$ converges to zero However

$$\sqrt{n}\ \|\hat{\mu}-\hat{\mu}_o\| = \|\sqrt{n}\ \hat{\mu}_1 + \frac{b\sqrt{n}}{K_{FF}} (x'x)^{-1} \{x' \Psi_F(\hat{\mu}_1) - x' \Psi_F(o)\}\|$$

so that, using assumption C'ii) it is sufficient to show that

$$(6;1) \qquad \sup_{\|\xi\| \leq \frac{d}{\sqrt{n}}} \| \frac{1}{\sqrt{n}} \{x' \Psi_F(\xi) - x' \Psi_F(o) + \frac{K_{FF}}{b} x' x\xi\}\| \xrightarrow{P_G} o \text{ if } \mu=o.$$

(6;1) follows from the linearization theorem for signed rank statistics proved in section 7 and 8 (see also [15]).

<u>Proof of Theorem 3;2.</u>

As in the proof of Theorem 3;1 we can suppose that $\mu = o$.

Let

$$\hat{\mu}_{oo} = (1-c\frac{K_{FG}}{K_{FF}})\ \hat{\mu}_1 + \frac{c}{K_{FF}} (x'x)^{-1} x' \Psi_F(o)$$

$$\hat{\mu}_{o1} = \frac{1}{K_{SG}} (x'x)^{-1} x' \Psi_S(o)$$

$$\hat{\mu}_{o2} = (1-c\frac{K_{FG}}{K_{FF}})\ \frac{(x'x)^{-1}}{K_{SG}} x' \Psi_S(o) + \frac{c}{K_{FF}} (x'x)^{-1} x' \Psi_F(o)$$

then it follows from (see theorem 7;2).

$$(6;2) \qquad P_G \left\{ \sup_{\|\xi\| \leq \frac{d}{\sqrt{n}}} \| \frac{1}{\sqrt{n}} (x' \Psi_F(\xi) - x' \Psi_F(o) + K_{FG}\ x'x\xi\| > \varepsilon \right\} \longrightarrow o$$

and the fact that $\hat{b} \xrightarrow{P_G} c$ that the asymptotic distribution of $\sqrt{n}\ \hat{\mu}$ is the same as that of $\sqrt{n}\ \hat{\mu}_{o2}$. From Hajek and Sidak [4] (p. 166) it follows that the asymptotic distribution of $\sqrt{n}\ \hat{\mu}_{o2}$ is normal with mean o and covariance given by (3;2).

Proof of Theorem 4;3.

Obviously $\hat{\mu}_1$ satisfies D'i). That S satisfies A' follows from the fact that $G^{-1}(\frac{u+1}{2})$ is non decreasing and non negative,

$$\int_0^1 \left[G^{-1}\left(\frac{u+1}{2}\right)\right]^2 du = \int_{-\infty}^{+\infty} y^2 g(y)\, dy < \infty$$ and that symmetry for G implies symmetry for S.

From Hajek and Sidak [4] (p. 166) it follows that $(x'x)^{-1} x'\Psi_S(o)$ is, if $\mu = o$, G equivalent to

$$(x'x)^{-1} x'(\psi_S(2G(Y_1)-1), \ldots, \psi_S(2G(Y_n)-1))' = (x'x)^{-1} x'Y.$$

The result then follows from the fact that $K_{SG} = 1$.

Proof of Theorem 4;4.

Obviously $\hat{\mu}_1$ satisfies D'i) and the double exponential distribution satisifies A'.

To prove D'ii) it needs to be shown that, if $\mu = o$,

$$\left\| \sqrt{n}\, \{(x_o'x_o)^{-1} x_o't - \frac{1}{K_{SG}} (x'x)^{-1} x' \Psi_S(o)\} \right\| \xrightarrow{P_G} o.$$

Let ε_j be the sum of $\pm$ 1's according as the observations in the j^{th} row of x_o are $\gtrless o$, let $\varepsilon = (\varepsilon_1, \ldots, \varepsilon_k)'$, then $x' \Psi_S(o) = x_o'\varepsilon$ and

$$\sqrt{n}\, \{(x_o' x_o)^{-1} x_o't - \frac{1}{K_{SG}} (x'x)^{-1} x' \Psi_S(o)\} =$$

$$= \frac{n(x'x)^{-1}}{\sqrt{n}} x_o' (nt - \frac{1}{2g(o)} \varepsilon).$$

Hence it is sufficient to prove that

$$(6;3) \qquad \frac{1}{n} \mathscr{E}_G \left\{ (nt_j - \frac{1}{2g(o)} \varepsilon_j)^2 \middle| t_j \right\} \xrightarrow{P_G} o$$

and (6;3) follows, as in the proof of Lemma 5;1, from the fact that the conditional, given t_j, distribution of $\frac{\varepsilon_j}{2} \frac{t_j}{|t_j|}$ is B $(\frac{n}{2}, p_j)$ where

$$p_j = \frac{|G(t_j) - G(o)|}{G(t_j)}$$

7. AN EXTENSION AND AN ANALOGUE OF A THEOREM OF JURECKOVA [6].

The following theorem is an extension to more dimensions of Theorem 3;1 of Jureckova [6].

Theorem 7;1.

If the components of Y have common distribution function G(x), if F satisfies A, if G satisfies A_1, if z satisfies B, if

$$S_j(\xi) = \sum_{i=1}^{n} z_{ij} \, \varphi_F \left(\frac{R_{Y_i - \sum_{\ell=1}^{p} z_{i\ell}\xi_\ell}}{n+1} \right)$$

then, for each j=1,...,p,

$$(7;1) \qquad \lim_{\nu\to\infty} P \left\{ \sup_{\|\xi\| \le d} \frac{1}{\sqrt{n}} \left| S_j\left(\frac{\xi}{\sqrt{n}}\right) - S_j(o) + \frac{K_{FG}}{\sqrt{n}} \sum_{\ell=1}^{p} \xi_\ell \sum_{i=1}^{n} z_{ij} z_{i\ell} \right| > \varepsilon \right\} = 0$$

for each d > o and each ε > 0.

Proof.

For p=1 Theorem 7;1, is a special case of Theorem 3;1 of Jureckova [6]. In the following it will be supposed that p > 1.

The proof will be given for j=1. As $\varphi_F(u)$ is the sum of two monotone square integrable functions it is sufficient to prove (7;1) for the case where $\varphi_F(u)$ is non decreasing. The proof consists of two parts. It will first be shown that, under A and Bi) and ii), for any fixed set of r points $(\xi_1^{(k)},\dots,\xi_p^{(k)})$, k=1,...,r

$$(7;2) \qquad P \left\{ \frac{1}{\sqrt{n}} \left| S_1\left(\frac{\xi^{(k)}}{\sqrt{n}}\right) - S_1(o) + \frac{K_{FG}}{\sqrt{n}} \sum_{\ell=1}^{p} \xi_\ell^{(k)} \sum_{i=1}^{n} z_{i1} z_{i\ell} \right| \le \varepsilon \text{ for each } k=1,\dots,r \right\} \longrightarrow 1.$$

Jureckova [6] proves (7;2) for p=1 in her Lemmas 3;1 – 3;8.

That (7;2) holds for $p > 1$ can be seen by noting that Jureckova's lemmas 3;1 -3;8 hold for $S_1\left(\frac{\xi}{\sqrt{n}}\right)$ if z satisfies

$$(7;3)\qquad \begin{cases} \frac{1}{n} \max\limits_{1 \le i \le n} z_{ij}^2 \longrightarrow 0 & \text{for each } j=1,\dots,p \\ \left|\frac{1}{n} \sum\limits_{i=1}^{n} z_{ij}^2\right| \le M & \text{for each } j=1,\dots,p \text{ where } M \text{ is a positive constant.} \end{cases}$$

Then (7;2) follows from the fact that (7;3) is implied by Bi) and ii).

In the second part of the proof it will be shown that for each $d > 0$ there exists a set of r fixed points $\xi^{(k)}$, $k=1,\dots,r$ such that, for $n > n_o$,

$$(7;4)\qquad \left[\frac{1}{\sqrt{n}} \left|S_1\left(\frac{\xi^{(k)}}{\sqrt{n}}\right) - S_1(o) + \frac{K_{FG}}{\sqrt{n}} \sum_{\ell=1}^{p} \xi_\ell^{(k)} \sum_{i=1}^{n} z_{i1} z_{i\ell}\right| \le \varepsilon \quad \text{for each } k=1,\dots,r\right] \Longrightarrow$$

$$\left[\sup_{\|\xi\| \le d} \frac{1}{\sqrt{n}} \left|S_1\left(\frac{\xi}{\sqrt{n}}\right) - S_1(o) + \frac{K_{FG}}{\sqrt{n}} \sum_{\ell=1}^{p} \xi_\ell \sum_{i=1}^{n} z_{i1} z_{i\ell}\right| \le 2^{p-1}\varepsilon\right]$$

The theorem then follows from (7;2) and (7;4).

The set of points $\xi^{(k)}$, $k=1,\dots,r$ satisfying (7;4) can be found as follows. By Biii) there exists, for each $j=2,\dots,p$, a number $\gamma_j \neq o$ such that, for $n > n_o$,

$$(7;5)\qquad (z_{i_1 1} - z_{i_2 1})\,(z_{i_1 1} - z_{i_2 1} + \gamma_j\,(z_{i_1 j} - z_{i_2 j})) \ge o \quad \text{for all } i_1, i_2.$$

(For simplicity of notation the first subscript on $\gamma_{1,j}$ is omitted).

By the transformation

$$(7;6)\qquad \begin{cases} \eta_1 = \xi_1 - \sum\limits_{j=2}^{p} \dfrac{\xi_j}{\gamma_j} \\ \eta_\ell = \dfrac{\xi_\ell}{\gamma_\ell} \quad \ell = 2,\dots,p \end{cases}$$

$S_1(\frac{\xi}{\sqrt{n}})$ can be written as

$$S_{10}(\frac{\eta}{\sqrt{n}}) \overset{def}{=} \sum_{i=1}^{n} z_{i1} \ \varphi_F \left(\frac{R_{Y_i} - \frac{1}{\sqrt{n}} (z_{i1}\eta_1 + \sum_{\ell=2}^{p} (z_{i1} + \gamma_\ell z_{i\ell})\eta_\ell)}{n+1} \right)$$

By (7;5) and theorem 2;1 of Jureckova [6], $S_{10}\ (\frac{\eta}{\sqrt{n}})$ is, for $n > n_o$, for fixed values of $\eta_1, \ldots, \eta_{j-1}, \eta_{j+1}, \ldots, \eta_p$, with probability one, a non increasing step function of η_j $(j=1,\ldots,p)$. Now choose the r fixed points $\xi^{(k)}$ as follows. Let C and ε be fixed positive numbers. Let R be an integer and let $r = (2R+1)^p$. Divide the cube $-C \leq \eta_j \leq C$ $(j=1,\ldots,p)$ into $(2R)^p$ cubes by dividing each axis into 2R equal pieces and choose $(2R+1)^p$ points $\eta^{(k)}$ on the corners of these cubes. These $(2R+1)^p$ points $\eta^{(k)}$ define, by (7;6), $(2R+1)^p$ points $\xi^{(k)}$. By choosing R in such a way that

$$(7;7) \qquad \begin{cases} |K_{FG}| \ \frac{1}{n} \sum\limits_{i=1}^{n} z_{i1}^2 \cdot \frac{C}{R} \leq \varepsilon \\ |K_{FG}| \ \left|\frac{1}{n} \sum\limits_{i=1}^{n} z_{i1} \ (z_{i1} + \gamma_\ell \ z_{i\ell})\right| \ \frac{C}{R} \leq \varepsilon \text{ all } \ell = 2,\ldots,p \end{cases}$$

these points $\xi^{(k)}$ satisfy, for $n > n_o$,

$$(7;8) \qquad \begin{gathered} \left[\frac{1}{\sqrt{n}} \left|S_1 \ (\frac{\xi^{(k)}}{\sqrt{n}}) - S_1(o) + \frac{K_{FG}}{\sqrt{n}} \sum_{\ell=1}^{p} \xi_\ell^{(k)} \sum_{i=1}^{n} z_{i1}z_{i\ell}\right| \leq \varepsilon \text{ for each } k=1,\ldots,r\right] \Longrightarrow \\ \left[\sup_{\substack{|\eta_j| \leq C \\ j=1,\ldots,p}} \ \frac{1}{\sqrt{n}} \left|S_1(\frac{\xi}{\sqrt{n}}) - S_1(o) + \frac{K_{FG}}{\sqrt{n}} \sum_{\ell=1}^{p} \xi_\ell \sum_{i=1}^{n} z_{i1}z_{i\ell}\right| \leq 2^{p-1}\varepsilon\right] . \end{gathered}$$

That (7;8) holds if R satisfies (7;7) can be seen by using the above mentioned monotonicity of $S_{10}\ (\frac{\eta}{\sqrt{n}})$ and by using the fact that (see also Jureckova [6]) if, for a monotone function $h(\xi)$ of one variable, $|h(\xi) - m\xi| \leq \varepsilon$ for $\xi = \xi_1$ and for $\xi = \xi_2$ $(\xi_1 < \xi_2)$, then $\sup\limits_{\xi_1 \leq \xi \leq \xi_2} |h(\xi) - m\xi| \leq 2\varepsilon$ provided $|m| \ (\xi_2 - \xi_1) \leq \varepsilon$.

That R can, for $n > n_1$, be chosen such that (7;7) is satisfied can be seen as follows. Let

$$\gamma = \max_{2 \leq j \leq p} |\gamma_j|$$

$$\sigma = \max_{1 \leq j \leq p} |\Sigma_{1j}|, \text{ where } \Sigma = (\Sigma_{ij})$$

then, by Bii) there exists n_1 such that for $n > n_1$

$$\frac{1}{n} \sum_{i=1}^{n} z_{i1}^2 \leq 2\sigma$$

$$\left| \frac{1}{n} \sum_{i=1}^{n} z_{i1}(z_{i1} + \gamma_\ell z_{i\ell}) \right| \leq 2\sigma(1+\gamma)$$

so that, by choosing R such that

$$R \geq \frac{|K_{FG}| \; 2\sigma \; C(1+\gamma)}{\varepsilon}$$

(7;7) is satisfied for $n > n_1$.

Further (7;4) follows from (7;9) by choosing d such that

$$\text{(7;9)} \qquad \left[\sum_{i=1}^{p} \xi_i^2 \leq d^2 \right] \Longrightarrow \left[|\eta_j| \leq C \text{ for all } j=1,\dots,p \right]$$

and a $d > o$ satisfying (7;8) is given by

$$d^2 = C^2 \frac{\left[\min\limits_{2 \leq j \leq p} \gamma_j \right]^2}{1 + \left[\min\limits_{2 \leq j \leq p} \gamma_j \right]^2}$$

The next theorem is a linearization theorem for signed rank statistics and is an extension of Theorem 3;2 in [15].

<u>Theorem 7;2.</u>

<u>If the components of Y have common distribution G(x), if F satisfies A', if G satisfies A_1', if x satisfies B', if</u>

$$T_j(\xi) = \sum_{i=1}^{n} x_{ij}\ \psi_F\ \left(\frac{R\left|Y_i - \sum_{\ell=1}^{p} x_{i\ell}\xi_\ell\right|}{n+1}\right) \operatorname{sgn}\ (Y_i - \sum_{\ell=1}^{p} x_{i\ell}\xi_\ell)$$

then, for each $j=1,\ldots,p_1$,

$$\lim_{\nu\to\infty} P\left\{\sup_{\|\xi\|\leq d} \frac{1}{\sqrt{n}}\left|T_j\ (\frac{\xi}{\sqrt{n}}) - T_j(o) + \frac{K_{FG}}{\sqrt{n}} \sum_{\ell=1}^{p} \xi_\ell \sum_{i=1}^{n} x_{ij}x_{i\ell}\right| > \varepsilon\right\} = 0 \tag{7;10}$$

for each $d > 0$ and each $\varepsilon > 0$.

Proof :

The following proof is analogous to the proof of Theorem 7;1. For $p_1=1$ the theorem is a special case of Theorem 3.2 of [15] and in the following it will be supposed that $p_1 > 1$. The proof will be given for $j=1$. As $\psi_F(u)$ is the sum of two square integrable functions, one non decreasing and non negative, the other non increasing and non positive, it is sufficient to prove (7;10) for the case where $\psi_F(u)$ is non decreasing and non negative.

It can be shown, analogously to Jureckova's lemmas (3;1) - (3;8) and using the results of Hajek and Sidak [4] (p. 219-221) that, under the assumptions A', A_1' and B'i) and ii), for any fixed set of points $\xi^{(k)}, k=1,\ldots,r$,

$$P\left\{\frac{1}{\sqrt{n}}\left|T_1(\frac{\xi^{(k)}}{\sqrt{n}}) - T_1(o) + \frac{K_{FG}}{\sqrt{n}} \sum_{\ell=1}^{p} \xi_\ell^{(k)} \sum_{i=1}^{n} x_{i\ell}x_{i\ell}\right| \leq \varepsilon \text{ for each } k=1,\ldots,r\right\} \to 1$$

Further, by B'iii), there exists, for each $j=2,\ldots,p_1$, a number γ_j such that

$$\begin{cases} 1.\ x_{i1}\ (x_{i1} + \gamma_j\ x_{ij}) \geq 0 \quad \text{for all}\ i \\ 2.\ (|x_{i_1 1}| - |x_{i_2 1}|)\ (|x_{i_1 1} + \gamma_j\ x_{i_1 j}| - |x_{i_2 1} + \gamma_j\ x_{i_2 j}|) \geq 0 \end{cases} \tag{7;11}$$

for all i_1, i_2

By the transformation (7;6) $T_1\ (\frac{\xi}{\sqrt{n}})$ can be written as

$$T_{10}(\frac{\eta}{\sqrt{n}}) = \sum_{i=1}^{n} x_{i1}\psi_F\left(\frac{R\left|Y_i - \frac{1}{\sqrt{n}}(x_{i1}\eta_1 + \sum_{\ell=2}^{p}(x_{i1}+\gamma_\ell x_{i\ell})\eta_\ell)\right|}{n+1}\right)$$

$$\operatorname{sgn}\ (Y_i - \frac{1}{\sqrt{n}}\ (x_{i1}\eta_1 + \sum_{\ell=2}^{p}(x_{i1} + \gamma_\ell x_{i\ell})\eta_\ell\ ))$$

and it follows from (7;11) and Theorem 3.1 in [15] that, for $n > n_o$, $T_{10}(\frac{n}{\sqrt{n}})$ is, for fixed values of $n_1, \ldots, n_{j-1}, n_{j+1}, \ldots, n_{p_1}$ with probability 1 **a non increasing** step function of n_j $(j=1,\ldots,p_1)$.

The rest of the proof is identical to that of Theorem 7;1

8. ALTERNATE ASSUMPTIONS.

In section 7 an extension of Jureckova's theorem was proved under the assumptionsA and B. A different set of assumptions has been suggested by Jureckova in her remark on page 1897 of [6]. The following paragraphs contain, first, a proof for p=2 of the multiparameter Jureckova theorem under these assumptions and, second, a proof that the conditions of section 7 imply those here. The application of the stronger approximate linearity theorem of this section to find linearized estimates is completely analogous to those of section 2 and 3.

Suppose F satisfies A, G satisfies A_1 and let z satisfy Bi) and ii).

For each n, z_{i1} can be written as

$$z_{i1} = z^*_{i1} + z^{**}_{i1}$$

such that

$$(8;1)\qquad \begin{cases} \sum_1^n z^*_{i1} = \sum_{i=1}^n z^{**}_{i1} = 0 \\ (z_{i_1 2} - z_{i_2 2})(z^*_{i_1 1} - z^*_{i_2 1}) \geq 0 \quad \text{for all } i_1, i_2 \\ (z_{i_1 2} - z_{i_2 2})(z^{**}_{i_1 1} - z^{**}_{i_2 1}) \leq 0 \quad \text{for all } i_1, i_2 \\ \sum_{i=1}^n (z^*_{i1})^2 > 0 \quad \text{or} \quad \sum_{i=1}^n (z^{**}_{i1})^2 > 0 \end{cases}$$

Then $S_1(\frac{1}{\sqrt{n}}\xi)$ can be written as the sum of

$$S^*_1(\frac{1}{\sqrt{n}}\xi) \stackrel{\text{def}}{=} \sum_{i=1}^n z^*_{i1}\, \varphi_F\left(\frac{R_{Y_i - \frac{1}{\sqrt{n}} z_{i1}\xi_1 - \frac{1}{\sqrt{n}} z_{i2}\xi_2}}{n+1}\right)$$

and

$$S_1^{**}\left(\frac{1}{\sqrt{n}}\xi\right) \overset{\text{def}}{=} \sum_{i=1}^{n} z_{i1}^{**}\, f_F\left[\frac{R_{Y_i - \frac{1}{\sqrt{n}} z_{i1}\xi_1 - \frac{1}{\sqrt{n}} z_{i2}\xi_2}}{n+1}\right]$$

Now suppose

<u>Assumption B iv</u>

$$\begin{cases} 1.\ \frac{1}{n} \max\limits_{1\leq i\leq n} (z_{i1}^{*})^2 \longrightarrow 0\ ,\ \frac{1}{n} \max\limits_{1\leq i\leq n} (z_{i1}^{**})^2 \longrightarrow 0 \\ 2.\ \text{there exists } n_o \text{ such that for } n > n_o \\ \quad \frac{1}{n}\sum\limits_{i=1}^{n} (z_{i1}^{*})^2 \leq M \text{ and } \frac{1}{n}\sum\limits_{i=1}^{n} (z_{i1}^{**})^2 \leq M \\ \quad \text{for some positive constant } M \end{cases}$$

That, for j=1, the extension of Jureckova's theorem holds if assumption Biii) is replaced by B iv) can be seen as follows. Choose, for a fixed $d > o$, $r = (2R+1)^2$ points $\xi^{(k)}$, $k=1,\dots,r$, on the corners of $(2R)^2$ squares obtained by dividing the square $-C \leqq \xi_j \leqq C$ $(j=1,2)$ into $(2R)^2$ equal squares. Then as in the proof of Theorem 7;1

$$(8;1)\quad P\left\{\frac{1}{\sqrt{n}}\left|S_1^{*}\left(\frac{\xi^{(k)}}{\sqrt{n}}\right) - S_1^{*}(o) + \frac{K_{FG}}{\sqrt{n}}\left(\xi_1^{(k)}\sum_{i=1}^{n} z_{i1}^{*} z_{i1} + \xi_2^{(k)}\sum_{i=1}^{n} z_{i1}^{*} z_{i2}\right)\right| \leqq \varepsilon \text{ for all } k=1,\dots,r\right\} \longrightarrow 1.$$

and

$$(8;2)\quad P\left\{\frac{1}{\sqrt{n}}\left|S_1^{**}\left(\frac{\xi^{(k)}}{\sqrt{n}}\right) - S_1^{**}(o) + \frac{K_{FG}}{\sqrt{n}}\left(\xi_1^{(k)}\sum_{i=1}^{n} z_{i1}^{**} z_{i1} + \xi_2^{(k)}\sum_{i=1}^{n} z_{i1}^{**} z_{i2}\right)\right| \leqq \varepsilon \text{ for all } k=1,\dots,r\right\} \longrightarrow 1.$$

Note that $\sum_{i=1}^{n} (z_{i1}^{*})^2$ and $\sum_{i=1}^{n} (z_{i1}^{**})^2$ are not necessarily both positive for all ν. However (8;1) follows, as in the proof of Theorem 7;1 for the subsequence of ν for which $\sum_{i=1}^{n} (z_{i1}^{*})^2 > o$ and (8;1) is obvious for the subsequence of ν for which $\sum_{i=1}^{n} (z_{i1}^{*})^2 = o$. The same holds for (8;2).

Then by choosing (see the proof of Theorem 7;1) R such that

$$|K_{FG}|\ \left|\frac{1}{n}\sum_{i=1}^{n} z_{i1}^{*}\ z_{i2}\right|\ \frac{d}{R} \leqq \varepsilon$$

$$|K_{FG}|\ \left|\frac{1}{n}\sum_{i=1}^{n} z_{i1}^{**}\ z_{i2}\right|\ \frac{d}{R} \leqq \varepsilon$$

$$|K_{FG}|\ \frac{1}{n}\sum_{i=1}^{n} z_{i1}^{2}\ \frac{d}{R} \leqq \varepsilon$$

one finds that

$$\left[\frac{1}{\sqrt{n}}\left|S_1^{*}\left(\frac{\xi^{(k)}}{\sqrt{n}}\right) - S_1^{*}(o) + \frac{K_{FG}}{\sqrt{n}}\left(\xi_1^{(k)}\sum_{i=1}^{n} z_{i1}^{*}\ z_{i1} + \xi_2^{(k)}\sum_{i=1}^{n} z_{i1}^{*}\ z_{i2}\right)\right| \leqq \varepsilon \text{ and}\right.$$

$$\left.\frac{1}{\sqrt{n}}\left|S_1^{**}\left(\frac{\xi^{(k)}}{\sqrt{n}}\right) - S_1^{**}(o) + \frac{K_{FG}}{\sqrt{n}}\left(\xi_1^{(k)}\sum_{i=1}^{n} z_{i1}^{**}\ z_{i1} + \xi_2^{(k)}\sum_{i=1}^{n} z_{i1}^{**}\ z_{i2}\right)\right| \leqq \varepsilon \quad \text{for all } k=1,\dots,r\right] \Longrightarrow$$

$$\left[\sup_{\substack{|\xi_j|\leq d\\ j=1,2}} \frac{1}{\sqrt{n}}\left|S_1\left(\frac{\xi}{\sqrt{n}}\right) - S_1(o) + \frac{K_{FG}}{\sqrt{n}}\left(\xi_1\sum_{i=1}^{n} z_{i1}^{2} + \xi_2\sum_{i=1}^{n} z_{i1}z_{i2}\right)\right| \leqq 8\ \varepsilon\right]$$

which proves the extension of Jureckova's theorem for j=1 under the assumption A, A_1 and Bi), ii) and iv). By analogously writing $z_{i2} = z_{i2}^{*} + z_{i2}^{**}$, an extension can be proved for j=2 under the assumption A, A_1, Bi), ii) and an assumption on z_{i2}^{*} , z_{i2}^{**} analogous to B iv).

That assumption B iv) follows from Biii) can be seen as follows. By Biii) there exists a number $\gamma_{2,1} \neq o$ such that, for all $n > n_o$, z_2 and $z_2 + \gamma_{2,1}\ z_1$ are similarly ordered. Now choose

$$\begin{cases} z_{i1}^{*} = \dfrac{z_{i2} + \gamma_{2,1}\ z_{i1}}{\gamma_{2,1}} \\[2ex] z_{i1}^{**} = -\dfrac{1}{\gamma_{2,1}}\ z_{i,2} \end{cases} \qquad \text{if } \gamma_{2,1} > o$$

and

$$\begin{cases} z^{*}_{i1} = -\dfrac{1}{\gamma_{2,1}} z_{i,2} \\ \\ z^{**}_{i1} = \dfrac{z_{i2} + \gamma_{2,1} z_{i1}}{\gamma_{2,1}} \end{cases} \qquad \text{if } \gamma_{2,1} < o$$

then, for $n > n_o$, z^{*}_{i1} and z^{**}_{i1} satisfy (8;1). Further from the fact that z satisfies Bi) and ii) it follows that z^{*}_{i1} and z^{**}_{i1} satisfy B iv).

The alternate assumptions for Theorem 7;2 are, for $p_1=2$, as follows. Let F satisfy A', G satisfy A_1' and let x satisfy B'i) and ii). In [15] it is shown that x_{i1} can be written as $x_{i1} = \sum_{\ell=1}^{4} x^{(\ell)}_{i1}$, such that

$$\begin{cases} 1. \quad x^{(\ell)}_{i1} \, x_{i2} \geqq o \quad \text{for each } i \text{ and } \ell = 1,2 \\ \\ \quad\;\; x^{(\ell)}_{i1} \, x_{i2} \leqq o \quad \text{for each } i \text{ and } \ell = 3,4 \\ \\ 2. \quad |x_2| \text{ and } |x^{(\ell)}_1| \text{ are simarly ordered for each } \ell = 1,\ldots,4 \\ \\ 3. \quad \sum_{i=1}^{n} (x^{(\ell)}_{i1})^2 > o \quad \text{for at least one } \ell \end{cases}$$

Then $T_1 \left(\frac{\xi}{\sqrt{n}}\right)$ can be written as $\sum_{\ell=1}^{4} T^{(\ell)}_1 \left(\frac{\xi}{\sqrt{n}}\right)$ and, as in the above proof, it can be seen that, for j=1, assumption B'iii) can be replaced by

<u>Assumption B'iv</u>

$$\begin{cases} 1. \; \dfrac{1}{n} \max\limits_{1 \leq i \leq n} (x^{(\ell)}_{i,1})^2 \longrightarrow o \quad \text{for each } \ell = 1,\ldots,4 \\ \\ 2. \; \dfrac{1}{n} \sum\limits_{i=1}^{n} (x^{(\ell)}_{i,1})^2 \leq M \quad \text{for } n > n_o \, , \; \ell = 1,\ldots,4 \end{cases}$$

By analogously writing $x_{i2} = \sum_{\ell=1}^{4} x^{(\ell)}_{i2}$ Theorem 7;2 can be proved for j=2 under the assumptions A', A_1', B'i) and ii) and an assumption on the $x^{(\ell)}_{i,2}$ analogous to B'iv).

REFERENCES

[1] Adichie, J.N.,

" Estimates of regression parameters based on rank tests ",

Ann. Math. Stat. 38 (1967), 894-904

[2] Bhuchongkul, S. and Puri, M.L.,

" On the estimation of contrasts in linear models ",

Ann. Math. Stat. 36 (1965), 198-202

[3] Greenberg, V.L.,

" Robust estimation in incomplete block designs ",

Ann. Math. Stat. 37 (1966), 1331-1337

[4] Hajek, J. and Sidak, Z.,

" Theory of rank tests ",

Academic Press, New York, 1967

[5] Hodges, J.L., Jr. and Lehmann, E.L.,

" Estimates of location based on rank tests ",

Ann. Math. Stat. 34 (1963), 598-611

[6] Jureckova, J.,

" Asymptotic linearity of a rank statistic in regression parameter ",

Ann. Math. Stat. 40 (1969), 1889-1900

[7] Koul, H.L.,

" Asymptotic behavior of a class of regions based on ranks ",

Technical Report, Department of Statistics and Probability, Michigan

State University, 1969

[8] Koul, H.L.,

" Asymptotic behavior of Wilcoxon type confidence regions in multivariate

linear regression ",

Ann. Math. Stat. 40 (1969), 1950-1979

[9] Kraft, C.H., and van Eeden, C.,
" Efficient linearized estimates based on ranks ",
Proceedings of the First International Symposium on Nonparametric Techniques in Statistical Inference,
Bloomington, Indiana, 1969 . To appear.

[10] Kraft, C.H., and van Eeden, C.,
" Relative efficiencies of quick and efficient methods of computing estimates from rank tests ", submitted.

[11] Lehmann, E.L.,
" Robust estimation in the analysis of variance ",
Ann. Math. Stat. 34 (1963), 957-966

[12] Lehmann, E.L.,
" Asymptotically nonparametric inference in some linear models with one observation per cell "
Ann. Math. Stat. 35 (1964), 726-734

[13] Puri, M.L. and Sen, P.K.,
" On robust estimation in incomplete block designs ",
Ann. Math. Stat. 38 (1967), 1587-1591

[14] Rao, P.V. and Thornby, J.I.,
" A robust point estimator in a generalized regression model ",
Ann. Math. Stat. 40 (1969), 1784-1790.

[15] van Eeden, Constance,
" A one - sample analogue of a theorem of Jureckova ",
Submitted to Ann. Math. Stat.