The Annals of Mathematical Statistics
1971, Vol. 42, No. 2, 650–661

ON A CLASS OF RANK ORDER TESTS FOR REGRESSION WITH PARTIALLY INFORMED STOCHASTIC PREDICTORS[1]

By Malay Ghosh[2] and Pranab Kumar Sen
University of North Carolina

0. Summary. Hájek (1962) has obtained asymptotically most powerful rank order tests for simple linear regression with non-stochastic predictors. His findings are extended here to the multiple linear regression model with stochastic predictors, including the situations where the predictors are partially informed. The proposed tests are shown to be conditionally distribution-free. Their asymptotic properties and efficiencies are studied, and the asymptotic optimality is established under the conditions of Wald (1943).

1. Introduction. Consider a sequence of stochastic matrices $\mathbf{Z}_\nu = (\mathbf{Z}_{\nu 1}, \cdots, \mathbf{Z}_{\nu\nu})$, $\nu \geqq 1$, where $\mathbf{Z}'_{\nu i} = (Y_{\nu i}, \mathbf{X}'_{\nu i}) = (Y_{\nu i}, X_{\nu i}^{(1)}, \cdots, X_{\nu i}^{(p)}) (p \geqq 1)$, $1 \leqq i \leqq \nu$ are i.i.d. rv (independent and identically distributed random vectors) having an absolutely continuous df (distribution function) $\mathbf{F}_\nu(y, \mathbf{x}; \boldsymbol{\beta})$, where $\mathbf{x} \in R^p$, the p-dimensional Euclidean space, and $\boldsymbol{\beta} = (\beta_1, \cdots, \beta_p)'$ is some (fixed) point in R^p on which the df F_ν depends in the following manner. We assume that $F_\nu(\infty, \mathbf{x}; \boldsymbol{\beta})$ is independent of $\boldsymbol{\beta}$ and denote it by $F_{01,\nu}(\mathbf{x})$, $\mathbf{x} \in R^p$; the corresponding density function is denoted by $f_{01,\nu}(\mathbf{x})$. The conditional density of $Y_{\nu i}$, given $\mathbf{X}_{\nu i} = \mathbf{x}$, depends on $\boldsymbol{\beta}$ and is assumed to be of the form $f_{20,\nu}(y \mid x; \boldsymbol{\beta}) = f_{20}([y - \beta_0 - \nu^{-\frac{1}{2}}\boldsymbol{\beta}'\mathbf{x}]/\sigma)$, where β_0 and $\sigma(>0)$ are nuisance parameters. Thus, the density function corresponding to the df $F_\nu(y, \mathbf{x}; \boldsymbol{\beta})$ is

$$f_\nu(y, \mathbf{x}; \boldsymbol{\beta}) = f_{01,\nu}(\mathbf{x}) f_{20}([y - \beta_0 - \nu^{-\frac{1}{2}}\boldsymbol{\beta}'\mathbf{x}]/\sigma), \qquad \nu \geqq 1. \tag{1.1}$$

We want to test the null hypothesis of no regression on $Y_{\nu i}$ on $\mathbf{X}_{\nu i}$, i.e.,

$$H_0 : \boldsymbol{\beta} = \mathbf{0} \quad \text{against the alternatives} \quad \boldsymbol{\beta} \neq \mathbf{0}. \tag{1.2}$$

Note that under H_0, $f_\nu(y, \mathbf{x}; \mathbf{0}) = f_{01,\nu}(\mathbf{x}) f_{20}([y - \beta_0]/\sigma)$, and hence $Y_{\nu i}$ and $\mathbf{X}_{\nu i}$ are stochastically independent. Also, the sequence of observations $\{\mathbf{Z}_\nu, \nu \geqq 1\}$ together with the sequence of df's $\{F_\nu(y, \mathbf{x}; \boldsymbol{\beta}), \nu \geqq 1\}$ generates a sequence of probability spaces $\{(\mathfrak{X}_\nu, \mathscr{A}_\nu, P_\nu(\boldsymbol{\beta})), \nu \geqq 1\}$, where $\mathfrak{X}_\nu$ is the $(p+1)$ ν-dimensional Euclidean space, $\mathscr{A}_\nu$ is the σ-field of Borel subsets of $\mathfrak{X}$ and for each (fixed) $\boldsymbol{\beta}(\in R^p)$, $P_\nu(\boldsymbol{\beta})$ is a probability measure defined on $(\mathfrak{X}_\nu, \mathscr{A}_\nu)$ determined uniquely by $\prod_{i=1}^{\nu} F_\nu(y_i, \mathbf{x}_i; \boldsymbol{\beta}) = \prod_{i=1}^{\nu} F_\nu(y_i, x_i^{(1)}, \ldots, x_i^{(p)}; \boldsymbol{\beta})$. Consequently, $P_\nu(\mathbf{0})$ is the probability measure corresponding to H_0. Similarly, $\{\mathbf{X}_\nu, \nu \geqq 1\}$ together with $\{F_{01,\nu}(\mathbf{x}), \nu \geqq 1\}$ generates a sequence of probability measures $\{P_\nu^*\}$ on the $p\nu$-dimensional Euclidean space.

Received April 7, 1969; revised August 8, 1970.

[1] Work supported by the National Institutes of Health, Grant GM-12868-04.

[2] Currently at the IBM Indian Statistical Institute, Calcutta.

We denote by $\mathscr{D}$ the class of all densities with finite Fisher information (cf. [8] page 17) and assume that $f_{20} \in \mathscr{D}$. For f_{20} and a df G with density $g \in \mathscr{D}$, we define

$$\phi(u) = -f_2{}'_0(F_{20}^{-1}(u))/f_{20}(F_{20}^{-1}(u)), \qquad \phi^*(u) = -g'(G^{-1}(u))/g(G^{-1}(u)), \quad 0 < u < 1; \tag{1.3}$$

$$I(f_{20}) = \int_0^1 \phi^2(u)\,du(<\infty) \quad \text{and} \quad I(g) = \int_0^1 \phi^{*2}(u)\,du(<\infty). \tag{1.4}$$

We introduce a class of scores $a_\nu(\mathrm{i})$, $1 \leqq i \leqq \nu$, generated by the score function $\phi^*(u): 0 < u < 1$, in any one of the three ways considered in [8] pages 157, 164–165. Then, we have

$$\lim_{\nu\to\infty} \int_0^1 \{a_\nu(1+[u\nu]) - \phi^*(u)\}^2\,du = 0, \tag{1.5}$$

where $[s]$ denotes the largest integer contained in s. Also, assume that

$$E(\mathbf{X}_{\nu 1}) = 0 \quad \text{and or some} \quad \delta > 0,\ \sup_\nu E[\mathbf{X}'_{\nu 1}\mathbf{X}_{\nu 1}]^{2+\delta} < c < \infty; \tag{1.6}$$

$$E(\mathbf{X}_{\nu i}\mathbf{X}'_{\nu i}) = \mathbf{\Sigma}_\nu \text{ (positive definite [pd])}, \qquad \text{and} \quad \lim_{\nu\to\infty} \mathbf{\Sigma}_\nu = \mathbf{\Sigma} \text{ is finite.} \tag{1.7}$$

Further, we consider a class of scores $b_{k\nu}(\mathrm{i})$, $1 \leqq i \leqq \nu (\nu \geqq 1)$, generated by the score function $\psi_k{}^*(u)$: $0 < u < 1$, where we assume that

$$\int_0^1 \psi_k{}^*(u)\,du = 0, \qquad \int_0^1 |\psi_k{}^*(u)|^r\,du < \infty \quad \text{for some} \quad r > 2; \tag{1.8}$$

$$\lim_{\nu\to\infty} \int_0^1 \{b_{k\nu}(1+[u\nu]) - \psi_k{}^*(u)\}^2\,du = 0, \qquad \text{for } k = 1, \cdots, p. \tag{1.9}$$

For later use, we define here

$$F_{k,\nu}(x) = P[X_{\nu i}^{(k)} \leqq x],\ F_{kk',\nu}(x, x') = P[X_{\nu i}^{(k)} \leqq x,\ X_{\nu i}^{(k')} \leqq x'],$$

and let $H_{kk'}^{(\nu)}(u, v) = P[F_{k,\nu}(X_{\nu i}^{(k)}) \leqq u, F_{k',\nu}(X_{\nu i}^{(k')}) \leqq v]$, $0 < u, v < 1$, $1 \leqq k \neq k' \leqq p$; let then

$$\sigma_{kk,\nu}^{(0)} = \sigma_{kk}^{(0)} = \int_0^1 \psi_k{}^{*2}(u)\,du, \qquad k = 1, \cdots, p \tag{1.10}$$

$$\sigma_{kk',\nu}^{(0)} = \int_0^1\int_0^1 \psi_k{}^*(u)\psi_{k'}^*(v)\,dH_{kk'}^{(\nu)}(u, v), \qquad \text{for } k \neq k' = 1, \cdots, p. \tag{1.11}$$

Finally, we assume that

$$\mathbf{\Sigma}_\nu{}^{(0)} = ((\sigma_{kk',\nu}^{(0)})) \text{ is pd for all } \nu \geqq \nu_0; \qquad \lim_{\nu\to\infty} \mathbf{\Sigma}_\nu{}^{(0)} = \mathbf{\Sigma}^{(0)} \text{ exists.} \tag{1.12}$$

In Section 2, a class of permutationally invariant (conditionally) distribution-free tests is proposed and studied. These tests are useful when the $\mathbf{X}_{\nu i}$ are partially informed, that is, they are not observable, but only the ranks on them are available —a case that arises in many educational or psychometric problems involving ranked data. In Section 3, the asymptotic permutation distribution of the proposed class of test statistics (under H_0) is developed. The asymptotic non-null distribution theory is developed, along the lines of Hájek (1962), in Section 4. In Section 5, the optimality properties of the proposed tests are established under the conditions of Wald (1943), and the asymptotic efficiency results are briefly presented.

2. A class of conditionally distribution-free rank order tests. Let $R_{\nu i}^{(0)}$ be the rank of $Y_{\nu i}$ among $Y_{\nu 1}, \cdots, Y_{\nu\nu}$ and let $R_{\nu i}^{(k)}$ be the rank of $X_{\nu i}^{(k)}$ among $X_{\nu 1}^{(k)}, \cdots, X_{\nu\nu}^{(k)}$ for $i = 1, \cdots, \nu$, and $k = 1, \cdots, p$. Thus, we use separate ranking for the different rows of $\mathbf{Z}_\nu$. We define the *collection* (*−rank*) *matrix* by

$$\mathbf{R}_{\nu i} = ((R_\nu^{(j)}))_{\substack{j=0,1,\cdots,p \\ i=1,\cdots,\nu}} = (\mathbf{R}_\nu^{(0)\prime}, \mathbf{R}_\nu^{(1)\prime}, \cdots, \mathbf{R}_\nu^{(p)\prime})'. \tag{2.1}$$

Each row of $\mathbf{R}_\nu$ is some permutation of the numbers $1, \cdots, \nu$. Let

$$\mathbf{S}_\nu = (S_{1\nu}, \cdots, S_{p\nu})'; \quad S_{k\nu} = \nu^{-\frac{1}{2}} \sum_{i=1}^{\nu} [b_{k\nu}(R_{\nu i}^{(k)}) - \bar{b}_{k\nu}][a_\nu(R_{\nu i}^{(0)}) - \bar{a}_\nu], \quad k = 1, \cdots, p, \tag{2.2}$$

where $\bar{b}_{k\nu} = \nu^{-1} \sum_{i=1}^{\nu} b_{k\nu}(i)$, $k = 1, \cdots, p$, and $\bar{a}_\nu = \nu^{-1} \sum_{i=1}^{\nu} a_\nu(i)$. The test statistic to be considered is a quadratic form in $\mathbf{S}_\nu$ and is based on the rank permutation principle of Chatterjee and Sen (1964) (see also [12]).

Since each row of R_ν is a random permutation of the numbers $1, \cdots, \nu$, $\mathbf{R}_\nu$ is a random matrix having $(\nu!)^{p+1}$ possible realizations. Now, under H_0 in (1.2), the $\mathbf{Y}_\nu$ are i.i.d. rv distributed independently of the $\mathbf{X}_{\nu i}$. Hence, under H_0, the $\nu!$ possible permutations of $Y_{\nu i}$'s are independent of the permutations of $X_{\nu i}^{(k)}$'s and are equiprobable among themselves, each permutation having the probability $1/(\nu!)$. Let $\mathbf{R}_{\nu(0)} = (\mathbf{R}_\nu^{(1)\prime}, \cdots, \mathbf{R}_\nu^{(p)\prime})'$. Two such matrices, say, $\mathbf{R}_{\nu(0)}$ and $\mathbf{R}_{\nu(0)}^*$ are said to be permutationally equivalent when it is possible to arrive at one from the other only by interchanging its columns. So if $\mathbf{R}_{\nu(0)}^*$ is a matrix having the same column vectors as of $\mathbf{R}_{\nu(0)}$ but so permuted that the first row of it consists of the numbers $1, \cdots, \nu$ in the natural order, i.e.,

$$\mathbf{R}_{\nu(0)}^* = ((R_{\nu i}^{*(k)}))_{\substack{k=1,\cdots,p \\ i=1,\cdots,\nu}}; \quad R_{\nu i}^{*(1)} = i: \quad 1 \leqq i \leqq \nu, \tag{2.3}$$

then $\mathbf{R}_{\nu(0)}^*$ is permutationally equivalent to $\mathbf{R}_{\nu(0)}$. Thus, corresponding to each $\mathbf{R}_{\nu(0)}^*$, there will be a set $\mathscr{S}(\mathbf{R}_{\nu(0)}^*)$ of $\nu!$ realizations of $\mathbf{R}_{\nu(0)}$ such that any member of the set is permutationally equivalent to $\mathbf{R}_{\nu(0)}^*$. The probability distribution of $\mathbf{R}_{\nu(0)}$ over the $(\nu!)^p$ possible realizations will depend on $F_{01,\nu}(\mathbf{x})$ even under H_0, (unless $F_{01,\nu}(\mathbf{x}) = \prod_{k=1}^{p} F_{k,\nu}(x_k)$); thus, in general, $\mathbf{S}_\nu$ is not distribution-free under H_0. However, given a particular set $\mathscr{S}(\mathbf{R}_{\nu(0)}^*)$ (of $\nu!$ realizations), the conditional distribution of $\mathbf{R}_{\nu(0)}$ over the $\nu!$ permutations of the columns of $\mathbf{R}_{\nu(0)}^*$ would be uniform under H_0, i.e.,

$$P[\mathbf{R}_{\nu(0)} = \mathbf{r}_{\nu(0)} \mid \mathscr{S}(\mathbf{R}_{\nu(0)}^*), H_0] = (\nu!)^{-1} \quad \text{for all} \quad \mathbf{r}_{\nu(0)} \in \mathscr{S}(\mathbf{R}_{\nu(0)}^*) \tag{2.4}$$

irrespective of $F_{01,\nu}(\mathbf{x})$. Let $\mathscr{P}_\nu$ denote the permutational (conditional) probability measure generated by the conditional law in (2.4). Then, we arrive at the following results by some standard computations:

$$E[\mathbf{S}_\nu \mid \mathscr{P}_\nu] = \mathbf{0}, \qquad E[\mathbf{S}_\nu \mathbf{S}_\nu' \mid \mathscr{P}_\nu] = \mathbf{V}_\nu = ((v_{kk',\nu})); \tag{2.5}$$

$$v_{kk',\nu} = \{(\nu-1)^{-1} \textstyle\sum_1^\nu [b_{k\nu}(R_{\nu i}^{(k)}) - \bar{b}_{k\nu}][b_{k'\nu}(R_{\nu i}^{(k')}) - \bar{b}_{k'\nu}]\}\{1/\nu \sum_1^\nu [a_\nu(i) - \bar{a}_\nu]^2\}, \tag{2.6}$$

for $k, k' = 1, \cdots, p$. We propose the following test statistic

(2.7) $$M_\nu = \mathbf{S}_\nu' \mathbf{V}_\nu^* \mathbf{S}_\nu,$$

where $\mathbf{V}_\nu^*$ is a generalized inverse of $\mathbf{V}_\nu$. Our test procedure is based on the following test function:

(2.8) $\zeta_1(\mathbf{Z}_\nu)$ is equal to 1, $\delta_{\nu,\varepsilon}$ or 0 according as $M_\nu >, =$ or $< M_{\nu,\varepsilon}$, where $M_{\nu,\varepsilon}$ and $\delta_{\nu,\varepsilon}$ are so chosen that $E[\zeta_1(\mathbf{Z}_\nu) \mid \mathscr{P}_\nu] = \alpha$, the level of significance.

This implies that $E[\zeta_1(\mathbf{Z}_\nu) \mid H_0] = \alpha$, i.e., $\zeta_1(\mathbf{Z}_\nu)$ is a similar size α test. However, this procedure requires the evaluation of $\nu!$ possible realizations of M_ν. The task becomes prohibitively laborious for large ν, and for this reason, in the next section, we simplify the large sample permutation distribution of M_ν.

To illustrate the proposed test procedure, we consider the simple case of *Wilcoxon scores* for all the $p+1$ variates. In this case, $a_\nu(i) = b_{1\nu}(i) = \cdots = b_{p\nu}(i) = i/(\nu+1)$, $1 \leqq i \leqq \nu$. We denote the Spearman rank-correlation matrix (of order $p+1$)

$$\mathbf{D}_\nu = \begin{pmatrix} 1 & \mathbf{d}'_{0\nu} \\ \mathbf{d}_{0\nu} & \mathbf{D}_{\nu(0)} \end{pmatrix},$$

where $\mathbf{d}_{0\nu} = (d_{01\nu}, \cdots, d_{0p\nu})'$, $\mathbf{D}_{\nu(0)} = ((d_{kk'\nu}))_{k,k'=1,\cdots,p}$ and

(2.9) $$d_{rs\nu} = [12/\nu(\nu^2-1)] \textstyle\sum_{i=1}^{\nu} (R_{\nu i}^{(r)} - (\nu+1)/2)(R_{\nu i}^{(s)} - (\nu+1)/2), \quad r, s = 0, 1, \cdots, p.$$

Then, after some algebraic simplifications, one gets from (2.2), (2.6) and (2.7) that

(2.10) $$\mathbf{S}_\nu = [\nu^{\frac{1}{2}}(\nu-1)/(12(\nu+1))]\mathbf{d}_{0\nu};$$

(2.11) $$\mathbf{V}_\nu = [\nu(\nu-1)/(144(\nu+1)^2)]\mathbf{D}_{\nu(0)};$$

(2.12) $$M_\nu = (\nu-1)\mathbf{d}'_{0\nu}\mathbf{D}^*_{\nu(0)}\mathbf{d}_{0\nu},$$

where $\mathbf{D}^*_{\nu(0)}$ is a generalized inverse of $\mathbf{D}_{\nu(0)}$. Thus, the proposed statistic is defined explicitly in terms of the elements of the Spearman rank-correlation matrix. The conditional (small sample) distribution can be constructed as above, while, as we shall see in Section 3, M_ν has asymptotically (under H_0) a central chi-square distribution with p d.f. (degrees of freedom). Thus, for large samples, the test procedure consists in rejecting the null hypothesis H_0 when M_ν exceeds the upper 100α per cent point of the χ^2 with p d.f. Some other special cases of M_ν can be constructed similarly.

3. Asymptotic permutation distribution of M_ν. We shall show that for large ν, one can approximate the permutation distribution of M_ν by the central χ^2 distribution with p d.f.; the later distribution function is denoted by $\Omega_p(x)$:

$0 \leqq x \leqq \infty$. We denote by $\Omega(x; \mathbf{R}^*_{\nu(0)}) = P[M_\nu \leqq x \mid \mathbf{R}^*_{\nu(0)}, \mathscr{P}_\nu]$, where $\mathscr{P}_\nu$ denotes that the conditional probability is computed under (2.4). Our main result (Theorem 3.3) is based on the following two theorems.

THEOREM 3.1. *Under* (1.4), (1.5), (1.8) *and* (1.9), *whenever* $\mathbf{V}_\nu$ *is asymptotically* pd

$$(3.1) \qquad \lim_{\nu\to\infty} \{\sup_{x \geq 0} |\Omega(x; \mathbf{R}^*_{\nu(0)}) - \Omega_p(x)|\} = 0.$$

PROOF. By (2.5) and (2.7), it suffices to show that as $\nu \to \infty$, the permutation distribution of $\mathbf{V}_\nu^{-\frac{1}{2}}\mathbf{S}_\nu$ converges to the multinormal distribution with null mean vector and dispersion matrix $\mathbf{I}_p$, whenever $\mathbf{V}_\nu$ is pd. For any fixed $\mathbf{e} = (e_1, \cdots, e_p)'$ ($\neq \mathbf{0}$), consider the linear compound $\mathbf{e}'\mathbf{S}_\nu$, which by (2.2) can be written as

$$(3.2) \quad \nu^{-\frac{1}{2}} \sum_{i=1}^{\nu} [a_\nu(R_{\nu i}^{(0)}) - \bar{a}_\nu] c_{\nu i}; \qquad c_{\nu i} = \sum_{k=1}^{p} e_k [b_{k\nu}(R_{\nu i}^{(k)}) - \bar{b}_{k\nu}], \qquad i = 1, \cdots, \nu.$$

Thus, by Theorems 4.1 and 4.2 of Hájek (1961), we have only to show that $\{c_{\nu i}\}$ satisfies the Noether condition (cf. [6] (3.3)) and the $\{a_\nu(i)\}$ satisfies the Hájek condition (cf. [6] (4.10), page 514). Since the $c_{\nu i}$ are linear compounds of the $[b_{k\nu}(R_{\nu i}^{(k)}) - \bar{b}_{k\nu}]$, $k = 1, \cdots, p$, which by (1.8) satisfy (coordinate-wise) the Noether condition, it follows by some standard computations that whenever $\mathbf{V}_\nu$ is pd, the $\{c_{\nu i}\}$ satisfies the Noether condition. Also, for any $\{k_\nu\}$,

$$\max_{1 \leqq i_1 < \cdots < i_{k_\nu} \leqq \nu} \{\nu^{-1} \sum_{j=1}^{k_\nu} [a_\nu(i_j) - \bar{a}_\nu]^2\} = \int_{1-\nu^{-1}k_\nu}^{1} \eta_\nu(u)\, du,$$

where $\eta_\nu(u) = a_\nu^*(i)$, $(i-1)/\nu < u \leqq i/\nu$, $i = 1, \cdots, \nu$, and $a_\nu^*(i)$ are the ordered values of the $[a_\nu(j) - \bar{a}_\nu]^2$, $1 \leqq j \leqq \nu$. Since by (1.4) and (1.5), $\int_0^1 \eta_\nu(u)\, du$ is finite, if we let $\nu^{-1}k_\nu \to 0$ as $\nu \to \infty$, it follows that

$$(\max_{1 \leqq i_1 < \cdots < i_{k_\nu} \leqq \nu} \sum_{j=1}^{k_\nu} [a_\nu(i_j) - \bar{a}_\nu]^2 / \{\sum_{j=1}^{\nu} [a_\nu(i) - a_\nu]^2\}$$
$$= \{\int_{1-\nu^{-1}k_\nu}^{1} \eta_\nu(u)\, du\} / \{\int_0^1 \eta_\nu(u)\, du\} \to 0$$

as $\nu \to \infty$. Thus, (4.10) of Hájek (1961) holds. □

In the next theorem we prove that $\mathbf{V}_\nu$ is pd in P_ν^*-probability as $\nu \to \infty$; this extends Theorem 4.2 of Puri and Sen (1966), without using their conditions on the existence and boundedness of the first derivatives of the score functions.

THEOREM 3.2. *Under* (1.4), (1.5), (1.8) *and* (1.9), $[\mathbf{V} - I(g)\, \mathbf{\Sigma}_\nu^{(0)}] \to 0$, *in* P_ν^*-*probability, as* $\nu \to \infty$.

PROOF. Since by (1.4) and (1.5), $\nu^{-1} \sum_{i=1}^{\nu} [a_\nu(i) - \bar{a}_\nu]^2 \to I(f_{20})$, as $\nu \to \infty$, and by (1.8) and (1.9), $\bar{b}_{k\nu} \to 0$, $\nu^{-1} \sum_{i=1}^{\nu} [b_{k\nu}(i) - \bar{b}_{k\nu}]^2 \to \sigma_{kk}^{(0)}$, $k = 1, \cdots, p$, it suffices to show that for $k \neq k'$, as $\nu \to \infty$,

$$(3.3) \qquad \nu^{-1} \sum_{i=1}^{\nu} b_{k\nu}(R_{\nu i}^{(k)}) b_{k'\nu}(R_{\nu i}^{(k')}) - \sigma_{kk',\nu}^{(0)} \to 0, \quad \text{in } P_\nu^*\text{-probability.}$$

Let now $H_{k\nu}(u) = \nu^{-1}$ [Number of $F_{k,\nu}(X_{\nu i}^{(k)}) \leqq u$], and $H_{kk'\nu}(u, v) = \nu^{-1}$ [Number of $F_{k,\nu}(X_{\nu i}^{(k)}) \leqq u$, $F_{k',\nu}(X_{\nu i}^{(k')}) \leqq v]$, $0 < u$, $v < 1$, for $k \neq k' = 1, \cdots, p$. Also, let $C_{k\nu}(i/[\nu+1]) = b_{k\nu}(i)$, $i = 1, \cdots, \nu$. Then,

$$
\begin{aligned}
\nu^{-1} \textstyle\sum_{i=1} b_{k\nu}(R_{\nu i}^{(k)}) b_{k'\nu}(R_{\nu i}^{(k')}) &= \textstyle\int_0^1 \int_0^1 \psi_k^*(u) \psi_{k'}^*(v)\, dH_{kk'\nu}(u, v) \\
&+ \textstyle\int_0^1 \int_0^1 [C_{k\nu}(\nu H_{k\nu}(u)/(\nu+1)) - \psi_k^*(\nu H_{k\nu}(u)/(\nu+1))] \\
&\qquad \cdot C_{k'\nu}(\nu H_{k'\nu}(v)/(\nu+1))\, dH_{kk'\nu}(u, v) \\
&+ \textstyle\int_0^1 \int_0^1 [C_{k'\nu}(\nu H_{k'\nu}(v)/(\nu+1)) - \psi_{k'}^*(\nu H_{k'\nu}(v)/(\nu+1))] \\
&\qquad \cdot \psi_k^*(\nu H_{k\nu}(u)/(\nu+1))\, dH_{kk'\nu}(u, v) \qquad (3.4) \\
&+ \textstyle\int_0^1 \int_0^1 [\psi_k^*(\nu H_{k\nu}(u)/(\nu+1)) - \psi_k^*(u)] \psi_{k'}^*(\nu H_{k'\nu}(v)/(\nu+1))\, dH_{kk'\nu}(u, v) \\
&+ \textstyle\int_0^1 \int_0^1 [\psi_{k'}^*(\nu H_{k'\nu}(v)/(\nu+1)) - \psi_{k'}^*(v)] \psi_k^*(u)\, dH_{kk'\nu}(u, v) \\
&= (\text{I}) + (\text{II}) + (\text{III}) + (\text{IV}) + (\text{V}), \text{ say.}
\end{aligned}
$$

Since $(\text{I}) = \nu^{-1} \sum_{i=1}^{\nu} \psi_k^*(F_{k,\nu}(X_{\nu i}^{(k)})) \psi_{k'}^*(F_{k',\nu}(X_{\nu i}^{(k')})) = \nu^{-1} \sum_{i=1}^{\nu} W_{\nu i}^*$, say, where by (1.8), $E|W_{\nu i}^*|^{r/2} \leqq \{[\int_0^1 |\psi_k^*(u)|^r du][\int_0^1 |\psi_{k'}^*(v)|^r dv]\}^{\frac{1}{2}} < \infty$ $(r > 2)$, using the Markov law of large numbers, we obtain that as $\nu \to \infty$, $|(\text{I}) - \sigma_{kk',\nu}^{(0)}| \to 0$, in P_ν^*-probability. Again using the Schwarz inequality on II and following some standard manipulations [with the aid of (1.8) and (1.9)], it can be shown that as $\nu \to \infty$, both (II) and (III) tend to 0, in P_ν^*-probability. Also,

$$
|(\text{IV})| \leqq \left\{ \left[\int_0^1 \left(\psi_k^* \left(\frac{\nu}{\nu+1} H_{k\nu}(u) \right) - \psi_k^*(u) \right)^2 dH_{k\nu}(u) \right] \cdot \left[\int_0^1 \psi_{k'}^{*2} \left(\frac{\nu}{\nu+1} H_{k'\nu}(v) \right) dH_{k'\nu}(v) \right] \right\}^{\frac{1}{2}}, \tag{3.5}
$$

where by (1.9), the second factor is finite and it converges to $\sigma_{k'k'}^{(0)}$ as $\nu \to \infty$. Further, on defining $U_1, \cdots, U_\nu$ as independent observations from a rectangular (0,1) distribution, and $\psi_k^{**}(u) = \psi_k^*(i/[\nu+1])$, for $(i-1)/\nu < u \leqq i/\nu$, $i = 1, \cdots, \nu$, we have

$$
\begin{aligned}
E&\left[\int_0^1 \left(\psi_k^* \left(\frac{\nu}{\nu+1} H_{k\nu}(u) \right) - \psi_k^*(u) \right)^2 dH_{k\nu}(u) \right] \\
&= E[\nu^{-1} \textstyle\sum_{i=1}^{\nu} \{\psi_k^*(R_{\nu i}^{(k)}/[\nu+1]) - \psi_k^*(U_i)\}^2] \qquad (3.6) \\
&= E[\nu^{-1} \textstyle\sum_{i=1}^{\nu} (\{\psi_k^{**}(R_{\nu i}^{(k)}/[\nu+1]) - \psi_k^{**}(U_i)\} + \{\psi_k^{**}(U_i) - \psi_k^*(U_i)\})^2].
\end{aligned}
$$

Using then the inequality that $(a+b)^2 \leqq 2(a^2+b^2)$, we obtain that (3.6) is bounded by

$$
\begin{aligned}
&2\{E[\nu^{-1} \textstyle\sum_{i=1}^{\nu} \{\psi_k^{**}(R_{\nu i}^{(k)}/[\nu+1]) - \psi_i^{**}(U_i)\}^2] \\
&\quad + E[\nu^{-1} \textstyle\sum_{i=1}^{\nu} \{\psi_k^{**}(U_i) - \psi_k^*(U_i)\}^2]\} \qquad (3.7) \\
&= 2\{E[\psi_k^{**}(R_{\nu i}^{(k)}/[\nu+1]) - \psi_k^{**}(U_i)]^2 + \textstyle\int_0^1 [\psi_k^{**}(u) - \psi_k^*(u)]^2 du\}.
\end{aligned}
$$

By Lemma 2.1 of [6] and by Lemma V.1.6 on page 164 of [8], the right-hand side of (3.7) $\to 0$ as $\nu \to \infty$. This implies that (IV) $\to 0$ in $P_\nu{}^*$-probability as $\nu \to \infty$. Similarly, it follows that (V) $\to 0$ in $P_\nu{}^*$-probability as $\nu \to \infty$. □

An immediate consequence of the preceding two theorems is the following:

THEOREM 3.3. *Under* (1.4), (1.5), (1.8), (1.9) *and* (1.12),

$$\sup_{x \geqq 0} \left|\Omega(x; \mathbf{R}^*_{\nu(0)}) - \Omega_p(x)\right| = 0, \text{ in } P_\nu{}^*\textit{-probability as } \nu \to \infty.$$

It follows from Theorem 3.3 that $M_{\nu,\varepsilon}$ and $\delta_{\nu,\varepsilon}$ defined by (2.8) converge in $P_\nu{}^*$-probability, as $\nu \to \infty$, to $\chi^2_{p,\alpha}$ and 0, respectively, where $\Omega_p(\chi^2_{p,\alpha}) = 1 - \alpha$.

4. Asymptotic non-null distribution of M_ν. The asymptotic non-null distribution of M_ν is obtained along the lines of Hájek (1962) (see also [8]). This is achieved by showing the "asymptotic equivalence" (in the sense that the difference converges in probability to 0) of M_ν and another statistic to be defined subsequently. The distribution of the latter is obtained by using the "contiguity" of arguments of Hájek (1962). In view of the basic difference that our predictors are stochastic, while Hájek's are non-stochastic, we need extensions and modifications of Lemmas VI.2.1.a, VI.2.1.b (page 211), Theorem VI.2.1 (page 213) and Theorem VI.2.4 (page 216) of [8]. In deriving most of these results, first, we condition with respect to the $\mathbf{X}_{\nu i}$ and proceed as in [8], and then pass on to the unconditional results by using some probabilistic arguments. In this context, we first prove the following lemma which forms a basis for subsequent results.

LEMMA 4.1. *Under* (1.6) *and* (1.7), *as* $\nu \to \infty$, (i) $\nu^{-\frac{1}{2}} \max_{1 \leqq i \leqq \nu} |X^{(k)}_{\nu i}| \to 0$ a.e.[3], *and* (ii) $\nu^{-1} \sum_{i=1}^{\nu} (X^{(k)}_{\nu i} - \bar{X}_\nu{}^{(k)})^2 - E(X^{(k)}_{\nu i})^2 \to 0$ a.e. (*where* $\bar{X}_\nu{}^{(k)} = \sum_{i=1}^{\nu} X^{(k)}_{\nu i}/\nu$), *for* $k = 1, \cdots, p$.

PROOF. For a given ν, $(X^{(k)}_{\nu i})^2$, $1 \leqq i \leqq \nu$ are i.i.d., and by (1.6),

$$E\{|X^{(k)}_{\nu i}|^2\}^{2+\delta/2} < c < \infty.$$

Using a result of Brillinger (1962), we get that for every $\varepsilon > 0$, there exists an integer $\nu_0 = \nu_0(\varepsilon)$, such that

$$P\{|\nu^{-1} \textstyle\sum_{i=1}^{\nu} (X^{(k)}_{\nu i})^2 - E(X^{(k)}_{\nu 1})^2| > \varepsilon\} \leqq K\varepsilon^{2+\delta/2}/\nu^{1+\delta/4},$$

for $\nu \geqq \nu_0$. Hence, $\nu^{-1} \sum_{i=1}^{\nu} (X^{(k)}_{\nu i})^2 - E(X^{(k)}_{\nu 1})^2 \to 0$ a.e., as $\nu \to \infty$. Similarly, from (1.6), $\bar{X}_\nu{}^{(k)} \to 0$, a.e., as $\nu \to \infty$. Hence, (ii) is proved. Again, for some $\delta'(0 < \delta' < \delta/2)$ and $\delta'' = (\delta - 2\delta')/(2+\delta')$, it follows that $E(|X^{(k)}_{\nu 1}|^{2+\delta'})^{2+\delta''} = E|X^{(k)}_{\nu 1}|^{4+\delta} < c < \infty$. Now, as before $\nu^{-1} \sum_{i=1}^{\nu} |X^{(k)}_{\nu i}|^{2+\delta'} - E|X^{(k)}_{\nu 1}|^{2+\delta'} \to 0$ a.e. Using then the inequality $\nu^{-1} \sum_{i=1}^{\nu} |X^{(k)}_{\nu i} - \bar{X}_\nu{}^{(k)}|^{2+\delta'} \leqq 2^{1+\delta'}\{\nu^{-1} \sum_{i=1}^{\nu} |X^{(k)}_{\nu i}|^{2+\delta'} + |\bar{X}_\nu{}^{(k)}|^{2+\delta'}\}$, and proceeding as in (ii), we get

$$\nu^{-1} \textstyle\sum_{i=1}^{\nu} |X^{(k)}_{\nu i} - \bar{X}_\nu{}^{(k)}|^{2+\delta'} - E|X^{(k)}_{\nu 1}|^{2+\delta'} \to 0 \text{ a.e., as } \nu \to \infty. \tag{4.1}$$

[3] If we start with a basic probability space $(\Gamma, \mathscr{B}, Q)$ and regard $\{\{P_\nu{}^*\}$ as a sequence of measures induced by some sequence of measurable transformations $\{X_\nu\}$ on this space, by convergence a.e. we mean (in the sequel) convergence a.e. (Q).

Since, by (1.6), $\sup_\nu E|X_{\nu 1}^{(k)}|^{2+\delta'} \leqq c^{(2+\delta')/(4+\delta)} < \infty$, (i) can be obtained from (ii), (4.1) and the following relation due to Hoeffding (1951):

$$(4.2)\qquad \lim_{\nu\to\infty} \max_{1\leqq i\leqq \nu}(X_{\nu i}^{(k)}-\bar{X}_\nu^{(k)})^2/\sum_{i=1}^{\nu}(X_{\nu i}^{(k)}-\bar{X}_\nu^{(k)})^2 = 0 \quad \text{a.e.} \quad \Leftrightarrow$$
$$\lim_{\nu\to\infty}[\nu^{-\delta'/2}\{\nu^{-1}\sum_{i=1}^{\nu}|X_{\nu i}^{(k)}-\bar{X}_\nu^{(k)}|^{2+\delta'}\}]$$
$$\times\{\nu^{-1}\sum_{i=1}^{\nu}(X_{\nu i}^{(k)}-\bar{X}_\nu^{(k)})^2\}^{1+\delta'/2} = 0 \text{ a.e.}$$

Note that the lemma implies that for each $k(=1,\cdots,p)$, $\{X_{\nu i}^{(k)}, 1\leqq i\leqq \nu\}$ satisfies the Noether condition a.e., a basic requirement of Hájek (1962) with non-stochastic predictors.

Define the statistic $\mathbf{T}_\nu = (T_{1\nu},\cdots,T_{p\nu})'$ as

$$T_{k\nu} = \nu^{-\frac{1}{2}}\sum_{i=1}^{\nu} X_{\nu i}^{(k)}\phi(F_{20}(Y'_{\nu i})) = \nu^{-\frac{1}{2}}\sum_{i=1}^{\nu} X_{\nu i}^{(k)}[-f'_{20}(Y'_{\nu i})/f_{20}(Y'_{\nu i})]$$
$$(4.3)\qquad = -2\nu^{-\frac{1}{2}}\sum_{i=1}^{\nu} X_{\nu i}^{(k)}[s'_{20}(Y'_{\nu i})/s_{20}(Y'_i)];\ s_{20}(x) = f_{20}^{\frac{1}{2}}(x)$$
$$\text{and}\quad Y'_{\nu i} = (Y_{\nu i}-\beta_0)/\sigma.$$

We first derive the asymptotic null distribution of $\mathbf{T}_\nu$.

LEMMA 4.2. *Under* (1.2), (1.4) *and* (1.5), T_ν *is asymptotically multinormal* $(\mathbf{0}, I(f_{20})\mathbf{\Sigma})$ a.e.

PROOF. Define $\mathbf{e} = (e_1,\cdots,e_p)'$ $(\neq \mathbf{0})$ as in Theorem 3.1, and write

$$(4.4)\qquad \mathbf{e}'\mathbf{T}_\nu = \sum_{i=1}^{\nu}[\sum_{i=1}^{p} e_k X_{\nu i}^{(k)}]\phi(F_{20}(Y'_{\nu i}))/\nu^{\frac{1}{2}}.$$

It follows from Lemma 4.1 that the coefficients $\sum_{k=1}^{p} e_k X_{\nu i}^{(k)}/\nu^{\frac{1}{2}}$, $1\leqq i\leqq \nu$, satisfy the Noether condition a.e., while $\phi(F_{20}(Y'_{\nu i}))$, $1\leqq i\leqq \nu$, are i.i.d. with 0 mean and variance $I(f_{20})(<\infty)$. From Theorem V.1.2 (page 153) of [8], it follows that conditional on the $\mathbf{X}_{\nu i}$, $1\leqq i\leqq \nu$ held fixed, $\mathbf{e}'\mathbf{T}_\nu$ is asymptotically normal $(0, \gamma'(\nu^{-1}\sum_{i=1}^{\nu}\mathbf{X}_{\nu i}\mathbf{X}'_{\nu i})\gamma)$ a.e., where $\gamma = \sigma^{-1}\boldsymbol{\beta}$. The result follows on noting that $\nu^{-1}\sum_{i=1}^{\nu}\mathbf{X}_{\nu i}\mathbf{X}'_{\nu i}\to\mathbf{\Sigma}$ a.e. (by Lemma 4.1). □

For convenience, we stick to the notations of [8], as far as possible. Define $r_{\nu i}(\mathbf{Z}_{\nu i};\boldsymbol{\beta}) = f_\nu(\mathbf{Z}_{\nu i};\boldsymbol{\beta})/f_\nu(\mathbf{Z}_{\nu i};\mathbf{0})$, $h_{\nu i} = \nu^{-\frac{1}{2}}\gamma'\mathbf{X}_{\nu i}$, $1\leqq i\leqq \nu$, $W_\nu = 2(\sum_{i=1}^{\nu}(r_{\nu i}^{\frac{1}{2}}-1))$, $L_\nu = \prod_{i=1}^{\nu} r_{\nu i}$ and $\xi = I(f_{20})(\gamma'\mathbf{\Sigma}\gamma)$. We intend to extend LeCam's second lemma (see [8] page 205) to the case of stochastic predictors, that is, to show that (a) $\log L_\nu - W_\nu + \xi/4 \to 0$ in $P_\nu(\mathbf{0})$-probability as $\nu\to\infty$, (b) $\log L_\nu$ is asymptotically normal $(-\xi/4, \xi)$, and (c) the probability measures $P_\nu(\boldsymbol{\beta})$ are contiguous to $P_\nu(\mathbf{0})$ The proof is based on Lemmas 4.3–4.5. We omit the proofs of the first two lemmas as they run parallel to those of Lemmas VI.2.1a and VI.2.1b (pages 211–212) of [8].

LEMMA 4.3. *Under* (1.2)–(1.7), $E(W_\nu)\to -\xi/4$ *as* $\nu\to\infty$.

LEMMA 4.4. *Under* (1.2)–(1.7), $\text{Var}(W_\nu-\gamma'\mathbf{T}_\nu)\to 0$ *as* $\nu\to\infty$.

LEMMA 4.5. *Under* (1.2), (1.4), (1.6) *and* (1.7),

(4.5) $$\lim_{\nu\to\infty} \max_{1\leqq i\leqq \nu} P_\nu(\mathbf{0})[|f_{\nu i}(\mathbf{Z}_{\nu i};\, \boldsymbol{\beta})/f_{\nu i}(\mathbf{Z}_{\nu i};\, \mathbf{0})-1| > \varepsilon] = 0,$$

where for every A_ν *in* $\mathscr{A}_\nu$, *we denote by* $P_\nu(\mathbf{0})(A\nu)$, *the probability of* A_ν *under* $P_\nu(\mathbf{0})$-*measure.*

PROOF. Along the lines of [8], it follows by some simple manipulations that

(4.6) $$P_\nu(\mathbf{0})[f_{\nu i}(\mathbf{Z}_{\nu i};\, \boldsymbol{\beta})/f_{\nu i}(\mathbf{Z}_{\nu i};\, \mathbf{0})-1\,| > \varepsilon] \leqq \varepsilon I^{\frac{1}{2}}(f_{20})E[\max_{1\leqq i\leqq \nu}|h_{\nu i}|].$$

Now, by Lemma 4.1,

$$\max_{1\leqq i\leqq \nu}|h_{\nu i}| \leqq \textstyle\sum_{k=1}^{p}(\nu^{-\frac{1}{2}}\max_{1\leqq i\leqq \nu}|X_{\nu i}^{(k)}|)(\max_{1\leqq k\leqq p}|\gamma_k|) \to 0 \text{ a.e.}$$

Further, $E(\max_{1\leqq i\leqq \nu}|h_{\nu i}|)^2 \leqq (\max_{1\leqq k\leqq p}|\gamma_k|)^2\cdot p\nu^{-1}\sum_{k=1}^{p}E(\max_{1\leqq i\leqq \nu}|X_{\nu i}^{(k)}|)^2$, and $E(\max_{1\leqq i\leqq \nu}|X_{\nu i}^{(k)}|)^2 \leqq (\nu/[2\nu-1]^{\frac{1}{2}})\{E(X_{\nu i}^{(k)})^4\}^{\frac{1}{2}} \leqq \nu^{\frac{1}{2}}c^{2/(4+\delta)}$, for all $k = 1, \cdots, p$. Hence,

$$\sup_\nu E[\max_{1\leqq i\leqq \nu}|h_{\nu i}|]^2\cdot \nu^{\frac{1}{2}} \leqq p^2c^{2/(4+\delta)} < \infty.$$

Thus, we can use the Lebesgue dominated convergence theorem (viz., [11] page 162) to obtain $E(\max_{1\leqq i\leqq \nu}|h_{\nu i}|)\to 0$ as $\nu\to\infty$. The lemma follows now from (4.6).

We are now in a position to derive the asymptotic non-null distribution of $\mathbf{S}_\nu$ by using LeCam's third lemma (see [8] page 208). We first define the following statistics:

(4.7) $$S_{k\nu}^{*} = \nu^{-\frac{1}{2}}\textstyle\sum_{i=1}^{\nu}[b_{k\nu}(R_{\nu i}^{(k)})-\bar{b}_{k\nu}]\phi^{*}(F_{20}(Y_{\nu i}')),$$
$$k = 1, \cdots, p;\ \mathbf{S}_\nu{}^{*} = (S_{1\nu}^{*}, \cdots, S_{p\nu}^{*})';$$

(4.8) $$T_{k\nu}^{*} = \nu^{-\frac{1}{2}}\textstyle\sum_{i=1}^{\nu}\psi_k{}^{*}(U_i)\phi^{*}(F_{20}(Y_{\nu i}')),$$
$$k = 1, \cdots, p;\ \mathbf{T}_\nu{}^{*} = (T_{1\nu}^{*}, \cdots, T_{p\nu}^{*})',$$

where the U_i are defined before (3.6). We prove that $\mathbf{S}_\nu{}^*-\mathbf{T}_\nu{}^*\to 0$ and $\mathbf{S}_\nu-\mathbf{S}_\nu{}^*\to 0$ in $P_\nu(\mathbf{0})$-(and hence by contiguity, in $P_\nu(\boldsymbol{\beta})$-) probability as $\nu\to\infty$; this implies that

(4.9) $\mathbf{S}_\nu$–$\mathbf{T}_\nu{}^{*}\to 0$ in $P_\nu(\mathbf{0})$–(as well as $P_\nu(\boldsymbol{\beta})$–) probability, as $\nu\to\infty$.

It is sufficient to show that as $\nu\to\infty$, for each $k(=1, \cdots, p)$

(4.10) $$E[S_{k\nu}^{*}-T_{k\nu}^{*})^2\,|\,P_\nu(\mathbf{0})]\to 0 \quad\text{and}\quad E[(S_{k\nu}-S_{k\nu}^{*})^2\,|\,P_\nu(\mathbf{0})]\to 0,$$

(as the rest follows by the Bonferroni inequality). Since

$$E[(S_{k\nu}^{*}-T_{k\nu}^{*})^2\,|\,P_\nu(\mathbf{0})] = \nu^{-1}\textstyle\sum_{i=1}^{\nu}E[b_{k\nu}(R_{\nu i}^{(k)})-\bar{b}_{k\nu}-\psi_k{}^{*}(U_i)]^2$$
$$\leqq 2\{\nu^{-1}\textstyle\sum_{i=1}^{\nu}E[b_{k\nu}(R_{\nu i}^{(k)})-\psi_k{}^{*}(U_i)]^2+2\bar{b}_{k\nu}^2\}I(g),$$

the first part of (4.10) can be proved by proceeding as in Theorem VI.1.5.a (page 157) and Theorem VI.1.6.a (page 163) of [8] and using (1.5) and (1.9). A similar proof follows for the second part of (4.10). Thus, (4.9) is proved.

It follows now from (a), Lemma 4.3 and 4.4 and (4.6) that for every $\mathbf{e} \neq \mathbf{0}$ of fixed real constants, $(\log L_\nu, \mathbf{e}'\mathbf{S}_\nu)$ has asymptotically the same distribution as of $(\gamma'\mathbf{T}_\nu - \xi/2, \mathbf{e}'\mathbf{T}_\nu^*)$. Now, define

$$\Sigma_{00\nu} = ((\textstyle\int_0^1 \int_0^1 \psi_k^*(u)\psi_{k'\nu}(v)\, dH_{kk'\nu}(u, v))); \tag{4.11}$$

$$\psi_{k\nu}(u) = F_{k\nu}^{-1}(u),\ k, k' = 1, \cdots, p,$$

and assume that

$$\lim_{\nu\to\infty} \Sigma_{00\nu} = \Sigma_{00} \quad \text{exists and is pd.} \tag{4.12}$$

Then, proceeding analogously as in Theorem VI.2.4 (pages 216–218) of [8], we may show that under H_0, $(\gamma'\mathbf{T}_\nu - \xi/2, \mathbf{e}'\mathbf{T}_\nu)$ is asymptotically bivariate normal $(-\xi/2, 0; \zeta, (\mathbf{e}'\Sigma_\nu^{(0)}\mathbf{e})I(g), (\mathbf{e}'\Sigma_{00\nu}\gamma)(\int_0^1 \phi(u)\phi^*(u)\, du))$ a.e. Using now LeCam's third lemma, it follows that under $P_\nu(\boldsymbol{\beta})$ measure, $\mathbf{e}'\mathbf{T}\nu^*$ is asymptotically normal $((\mathbf{e}'\Sigma_{00\nu}\gamma) \int_0^1 \phi(u)\phi^*(u)\, du, (\mathbf{e}'\Sigma_\nu^{(0)}\mathbf{e})I(g))$ a.e. whence $\mathbf{T}_\nu^*$ (and hence $\mathbf{S}_\nu$) is under $P_\nu(\boldsymbol{\beta})$, asymptotically multinormal $(\Sigma_{00\nu}\gamma(\int_0^1 \phi(u)\phi^*(u)\, du), \Sigma_\nu^{(0)} \cdot I(g))$ a.e. Since [by (1.12)], $\Sigma_\nu^{(0)} \to \Sigma^{(0)}$, and $\Sigma_{00\nu} \to \Sigma_{00}$ (both pd), recalling the definition of M_ν in (2.7), it follows from Theorem 3.2 and Slutsky's theorem that under $P_\nu(\boldsymbol{\beta})$, M_ν is distributed asymptotically (a.e.) as a noncentral chi-square with p d.f. and noncentrality parameter

$$(\gamma'\Sigma_{00}'(\Sigma^{(0)})^{-1}\Sigma_{00}\gamma)(\textstyle\int_0^1 \phi(u)\phi^*(u)\, du)^2/I(g).$$

5. Asymptotic optimality and ARE of the proposed tests. Let $\hat{\boldsymbol{\beta}}_\nu = (\beta_{1\nu}, \cdots, \beta_{p\nu})'$ be the maximum likelihood estimate (m.l.e.) of $\boldsymbol{\beta}$ and let $\hat{\gamma}_\nu = \sigma^{-1}\hat{\boldsymbol{\beta}}_\nu$, $\nu \geqq 1$. Define $L(\mathbf{Z}_\nu; \boldsymbol{\beta}) = \prod_{i=1}^{\nu} f_\nu(\mathbf{Z}_{\nu i}; \boldsymbol{\beta})$ and the likelihood ratio (l.r.) criterion

$$\lambda_\nu = L(\mathbf{Z}_\nu; \mathbf{0})/L(\mathbf{Z}_\nu; \hat{\boldsymbol{\beta}}). \tag{5.1}$$

Consider the statistic

$$M_\nu^* = \mathbf{T}_\nu'\Sigma_\nu^{-1}\mathbf{T}_\nu, \qquad \text{where } \mathbf{T}_\nu \text{ is defined by (4.3).} \tag{5.2}$$

We make assumptions on γ_ν and $L(\mathbf{Z}_\nu; \boldsymbol{\beta})$ as in Wald [14]. It may be noted that Wald's assumptions include some uniformity conditions on the m.l.e. as well as a moment condition slightly more restrictive than the finite Fisher information. These conditions, however, are met by most of the well-known distributions. Then, we have the following:

THEOREM 5.1. *$M_\nu^* + 2\log_e \lambda_\nu \to 0$ in $P_\nu(\mathbf{0})$ as well as $P_\nu(\boldsymbol{\beta})$ probability, as $\nu \to \infty$.*

Outline of the proof. Define $\mathbf{I}_{\gamma\nu} = E_\gamma((-(\partial^2/\partial\gamma_k\, \partial\gamma_{k'}) \log_e L(\mathbf{Z}_\nu; \boldsymbol{\beta})))_{k,k'=1,\cdots p}$. Then, it follows from the results of Wald (1943) that $2\log_e \lambda_\nu + \hat{\gamma}_\nu'\mathbf{I}_{0\nu}\hat{\gamma}_\nu \to 0$ in $P_\nu(\mathbf{0})$ probability as $\nu \to \infty$. Also, it follows by some routine computations that $\mathbf{T}_\nu - \mathbf{I}_{0\nu}\hat{\gamma}_\nu \to 0$ in $P_\nu(\mathbf{0})$ probability as $\nu \to \infty$. The first part of the theorem can now be proved by the Slutsky theorem and the fact that $\mathbf{I}_{0\nu} = \Sigma_\nu$. The second part follows from the contiguity arguments. □

Thus, a test procedure similar to the one considered in (2.8), but based on $M_\nu{}^*$ instead of M_ν, possesses the asymptotic optimal properties of the likelihood ratio test as described in Theorem VIII (page 478) of Wald (1943). In particular, the test will be asymptotically most stringent and have best average power over suitable ellipsoids in $\boldsymbol{\beta}$. Further, using LeCam's third lemma, it can be shown that under $P_\nu(\boldsymbol{\beta})$, $M_\nu{}^*$ is distributed asymptotically as a noncentral chi-square with p d.f. and noncentrality parameter $(\gamma'\Sigma\gamma)I(f_{20})$. Hence, the ARE (in the sense of Hannan [9]) of M_ν with respect to $M_\nu{}^*$ is given by

$$e = [(\gamma'\boldsymbol{\Sigma}_{00}'(\boldsymbol{\Sigma}^{(0)})^{-1}\boldsymbol{\Sigma}_{00}\gamma)/(\gamma'\boldsymbol{\Sigma}\gamma)]\rho_2{}^2, \tag{5.3}$$

where $\rho_2 = (\int_0^1 \phi(u)\phi^*(u)\,du)/[(\int_0^1 \phi^2(u)\,du)(\int_0^1 \phi^{*2}(u)\,du)]^{\frac{1}{2}}$. Note that in the particular case, $\phi^* = \phi$, $\psi_k{}^* = \psi_k$, for $k = 1, \cdots, p$, $e = 1$, i.e., the test based on the statistic M_ν is asymptotically optimal in the sense described above.

We may remark that in general the ARE in (5.3) depends on γ as well as on the underlying df through the dispersion matrices $\boldsymbol{\Sigma}_{00}$ and $\boldsymbol{\Sigma}^{(0)}$. However, using a well-known theorem of Courant on the extrema of the ratio of quadratic forms (see [2]), we obtain that

$$\min_\gamma(\max_\gamma)\, e = \text{minimal (maximal) eigenvalue of } \boldsymbol{\Sigma}^{-1}\boldsymbol{\Sigma}_{00}'(\boldsymbol{\Sigma}^{(0)})^{-1}\boldsymbol{\Sigma}_{00}. \tag{5.4}$$

We shall have occasion to study (5.4) in some special cases later on.

It is interesting to observe that instead of considering the rank statistic S_ν, as defined in (2.2), if we consider the *mixed statistic* $S_\nu{}^0 = (S_{1\nu}^0, \cdots, S_{p\nu}^0)'$, where $S_{k\nu}^0 = \nu^{-\frac{1}{2}}\sum_{i=1}^{\nu} X_{\nu i}^{(k)}[a_\nu(R_{\nu i}^{(0)}) - \bar{a}_\nu]$, $k = 1, \cdots, p$, (which would be more logical to consider when the $\mathbf{X}_{\nu i}$ are observable), the appropriate statistic will be the corresponding quadratic form in $\mathbf{S}_\nu{}^0$, and the ARE of $\mathbf{S}_\nu{}^0$ with respect to $M_\nu{}^*$ will be equal to $\rho_2{}^2$. Thus, in particular when $\phi^* = \phi$, the resulting test is asymptotically optimal. However, as already mentioned, the mixed statistics cannot be used when the $\mathbf{X}_{\nu i}$ are not observable. Also, unlike M_ν the statistic based on $\mathbf{S}_\nu{}^0$ is not unconditionally distribution-free when we have $p = 1$. Note that the permutation distribution of $S_{1\nu}$ agrees with its unconditional null distribution, and $v_{11,\nu}$ is non-stochastic. Hence, for $p = 1$, $M_\nu = S_{1\nu}^2/v_{11,\nu}$ is unconditionally distribution-free under (1.2). Actually, here, we have from (5.3) that the ARE is equal to $\rho_1{}^2\rho_2{}^2$, where ρ_1 is defined as in ρ_2, with ϕ and ϕ^* being replaced by ψ_1 and $\psi_1{}^*$ respectively. Thus, if $\phi = \phi^*$ and $\psi_1 = \psi_1{}^*$, the test based on M_ν has Wald-optimality, while the one-sided test based on $S_{1\nu}$ is asymptotically most powerful (as in [7]).

We now study the special case when $F_\nu(y, \mathbf{x}; \boldsymbol{\beta})$ is a nondegenerate $(p+1)$-variate multinormal df, and for $\phi^*(u)$ or $\psi_k{}^*(u)$ $(k = 1, \cdots, p)$, we use either $u - \frac{1}{2}$ (the Wilcoxon scores) or $\Phi^{-1}(u)$, the inverse of the standard normal df (the normal scores). If for all the $p+1$ variates, normal scores are used, it follows readily from (5.3) (by some standard computations) that $e = 1$, for all $p \geqq 1$ and all possible dispersion matrices. This proves the asymptotic optimality of the normal scores procedure when the underlying df is normal. Also, if for all the $\psi_k{}^*$, $k = 1, \cdots, p$, we use the normal scores, while for ϕ^* we use the Wilcoxon scores, e reduces to $3/\pi = 0.955$. Again, if for ϕ^*, we use the normal scores, while we use the Wilcoxon scores for the other k-variates, we have $\rho_2 = 1$, and hence, e equals the first factor on the right-hand side of (5.3). Thus, when $p = 1$, $e = 3/\pi$; for $p = 2$, the bounds in (5.4) depend on the grade correlations of the $\mathbf{X}_{vi}$, and the minimum of the lower bound and the maximum of the upper bound (over the variation of the parent dispersion matrix) when computed agree, incidentally, with similar bounds (0.866, 0.965) in [1] for the bivariate (one-sample) location problem, while for $p \geqq 2$, by proceeding as in [13], we have $e \leqq 1$, uniformly in the parent dispersion matrix. Finally, if for all the $(p+1)$-variates, we use the Wilcoxon scores, $\rho_2 = 3/\pi$, and hence, all the values and bounds obtained above need to be multiplied by $3/\pi$. Thus, for $p = 1$, $e = 9/\pi^2$, for $p = 2$, the two bounds are (0.827, 0.922), while for $p \geqq 2$, $e \leqq 3/\pi$, uniformly in the underlying dispersion matrices.

Acknowledgments. The authors wish to express their thanks to the referee for his useful comments on the paper and to Professor G. Simons for pointing out to the authors the paper by Brillinger [3].

REFERENCES

[1] Bickel, P. J. (1965). On some asymptotically nonparametric competitors of Hoetlling's T^2. *Ann. Math. Statist.* **36** 160–173. Correction: 36 1583.

[2] Bodewig, E. (1956). *Matrix Calculus.* Interscience, New York.

[3] Brillinger, D. R. (1962). A note on the rate of convergence of a mean. *Biometrika* **49** 574–576.

[4] Chatterjee, S. K. and Sen, P. K. (1964). Nonparametric tests for the bivariate two-sample location problem. *Calcutta Statist. Assoc. Bull.* **13** 18–58.

[5] Ghosh, M. (1969). Asymptotically optimal nonparametric tests for miscellaneous problems in linear regression. Ph.D. dissertation, Univ. of North Carolina.

[6] Hájek, J. (1961). Some extensions of the Weld–Wolfowitz–Noether Theorem. *Ann. Math. Statist.* **32** 506–523.

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

[8] Hájek, J. and Sidák, Z. (1967). *Theory of Rank Tests.* Academic Press, New York.

[9] Hannan, E. J. (1956). The asymptotic power of tests based on multiple correlation. *J. Roy. Statist. Soc. Ser. B* **18** 227–233.

[10] Hoeffding, W. (1951). A combinatorial central limit theorem. *Ann. Math. Statist.* **22** 558–566.

[11] Loève, M. (1963). *Probability Theory* (3rd ed.). Van Nostrand, Princeton.

[12] Puri, M. L. and Sen, P. K. (1966). On a class of multivariate multisample rank order tests. *Sankhya Ser. A* **28** 353–376.

[13] Puri, M. L. and Sen, P. K. (1969). Analysis of covariance based on general rank scores. *Ann. Math. Statist.* **40** 610–618.

[14] Wald, A. (1943). Tests for statistical hypotheses concerning several parameters when the number of observations is large. *Trans. Amer. Math. Soc.* **54** 426–483.