Tests of Linear Hypotheses Based on Regression Rank Scores

C. Gutenbrunner
J. Jurčeková
R. Koenker
S. Portnoy
Tests of Linear Hypotheses
Based on Regression Rank Scores

C. Gutenbrunner
J. Jurečková
R. Koenker
S. Portnoy

Dedicated to the memory of Jaroslav Hájek

Philipps Universität, Marburg, Germany
Charles University, Prague, Czechoslovakia
University of Illinois at Urbana-Champaign, USA

The work was partially supported by NSF grants 88-02555 and 89-22472 to S. Portnoy and R. Koenker and by support from the Australian National University to J. Jurečková and R. Koenker.
ABSTRACT

We propose a general class of asymptotically distribution-free tests of a linear hypothesis in the linear regression model. The tests are based on regression rank scores, recently introduced by Gutenbrunner and Jurečková (1990) as dual variables to the regression quantiles of Koenker and Bassett (1978). Their properties are analogous to those of the corresponding rank tests in location model. Unlike the other regression tests based on aligned rank statistics, however, our tests do not require preliminary estimation of nuisance parameters.

AMS 1980 subject classifications: 62G10, 62J05
Key words and phrases: Regression quantile, regression rank scores
1. Introduction

Several authors including Koul (1970), Puri and Sen (1985) and Adichie (1978) have proposed and investigated asymptotically distribution-free tests of some types of linear hypothesis for the linear regression model based upon aligned rank statistics. A good review of these results including extensions to multivariate models may be found in Puri and Sen (1985) and Adichie (1984). The hypothesis under consideration typically involves nuisance parameters which should be estimated by a preliminary estimate; the aligned (or signed) rank statistics are then based on residuals from the preliminary estimates. Alternative approaches to inference based on rank estimation have been considered by McKean and Hettmansperger (1978), Aubuchon and Hettmansperger (1988) and Draper (1988).

In this paper we explore an alternative approach to the construction of rank tests for the linear regression model based on the regression rank scores introduced in Gutenbrunner and Jurečková (1990). Regression rank scores represent a natural extension of location rank scores introduced in Hájek and Šidák (1967, Section V.3.5) which stand in one-to-one relation to the ranks of sample observations. The tests based on regression rank scores offer a natural extension of rank-based methods of testing to the general linear model and avoid many of the difficulties introduced by preliminary estimation of nuisance parameters in prior proposals.

The next section of the paper surveys our results and establishes notation. Section 3 develops some theory of the regression rank score process. Section 4 treats the theory of
simple linear rank statistics based on this process. Section 5 contains a formal treatment of the proposed tests. And Section 6 describes an example.

2. Notation and preliminary considerations

Consider the classical linear regression model

\[ Y = X\beta + E \]  \hspace{1cm} (2.1)

which we will partition as

\[ Y = X_1\beta_1 + X_2\beta_2 + E \]  \hspace{1cm} (2.2)

where \( \beta_1 \) and \( \beta_2 \) are \( p \)– and \( q \)–dimensional parameters, \( X = X_n \) is a known, \( n \times (p+q) \) design matrix with rows \( x_m \cdot = x_i \cdot = (x_i1 \cdot, x_{i2} \cdot) \in \mathbb{R}^{p+1}, \ i=1,\ldots, n \). We will assume throughout that \( x_{i1} = 1 \) for \( i = 1,\ldots,n \). \( Y \) is a vector of observations and \( E \) is an \( n \times 1 \) vector of \( i.i.d. \) errors with common distribution function \( F \). The precise form of \( F \) need not be known but we shall generally assume that \( F \) has an absolutely continuous density \( f \) on \( (A, B) \) where \( -\infty \leq A = \sup \{x: F(x) = 0\} \) and \( +\infty \geq B = \inf \{x: F(x) = 1\} \). Moreover, we shall impose some conditions on the tails of \( f \) assuming, among other conditions, that \( f \) monotonically decreases to 0 when \( x \rightarrow A+ \), or \( x \rightarrow B- \). Denoting \( D_n = n^{-1}X_1'X_1 \) and

\[ H_1 = X_1(X_1'X_1)^{-1}X_1' \quad \text{and} \quad Q_n = n^{-1}(X_2 - \hat{X}_2)'(X_2 - \hat{X}_2) \]  \hspace{1cm} (2.3)

with \( \hat{X}_2 = H_1X_2 \) being the projection of \( X_2 \) on the space spanned by the columns of \( X_1 \), we shall also assume

\[ \lim_{n \to \infty} D_n = D, \quad \lim_{n \to \infty} Q_n = Q \]  \hspace{1cm} (2.4)

where \( D \) and \( Q \) are positive definite \( (p \times p) \) and \( (q \times q) \) matrices, respectively.

We are interested in testing the hypothesis

\[ H_0: \beta_2 = 0, \quad \beta_1 \text{ unspecified} \]  \hspace{1cm} (2.5)

versus the Pitman (local) alternatives.
\[ H_n : \beta_{2n} = n^{-1/2} \beta_0 \] (2.6)

with \( \beta_0 \) being a fixed vector from \( \mathbb{R}^q \).

The regression rank scores introduced in Gutenbrunner and Jurečková (1990) arise as a vector of solutions

\[ \hat{\alpha}_n(\alpha) = (\hat{\alpha}_{n1}(\alpha), ..., \hat{\alpha}_{nq}(\alpha))^\top, \quad 0 < \alpha < 1 \] (2.7)

of the dual form of the linear program required to compute the regression quantile statistics of Koenker and Bassett (1978). More precisely, the vector \( \hat{\alpha}_1(\alpha) = (\hat{\beta}_1(\alpha), ..., \hat{\beta}_p(\alpha))^\top \in \mathbb{R}^p \) of regression quantiles corresponding to the submodel

\[ Y = X_1 \beta_1 + E \] (2.8)

is any solution of the minimization

\[ \sum_{i=1}^{n} \rho_\alpha(Y_i - x_{1i}^\top t), \quad t \in \mathbb{R}^q \] (2.9)

where

\[ \rho_\alpha(x) = |x| \cdot ((1-\alpha)I[x < 0] + \alpha I[x > 0]), \quad x \in \mathbb{R}^q. \] (2.10)

Koenker and Bassett (1978) characterized the regression quantile as the component \( \hat{\beta} \) of the optimal solution \( (\hat{\beta}, r^+, r^-) \) of the linear program

\[ \alpha I_{1n} \hat{\beta}^\top r^+ + (1-\alpha) I_{1n} \hat{\beta}^\top r^- : = \min \]

\[ X_1 \hat{\beta} + r^+ - r^- = Y \] (2.11)

\[ \beta \in \mathbb{R}^p, \quad r^+, r^- \in \mathbb{R}_+^p \]

and \( 1_n = (1, ..., 1)^\top \in \mathbb{R}^n, \quad 0 < \alpha < 1 \). Finite-sample as well as asymptotic properties of \( \hat{\beta}(\alpha) \) are studied in Koenker and Bassett (1978), Ruppert and Carroll (1980), Jurečková (1984), Gutenbrunner (1986), Koenker and Portnoy (1987), Gutenbrunner and Jurečková (1990) among others.
The dual program to (2.10) can be written in the form

\[ Y'\hat{\alpha}(\alpha) : = \max \]

\[ X_1'\hat{\alpha}(\alpha) = (1-\alpha)X_1'1_n \]  

(2.12)

\[ \hat{\alpha}(\alpha) \in [0, 1]^n, \quad 0 < \alpha < 1 \]

As shown in Gutenbrunner and Jurečková (1990) many aspects of the duality of order statistics and ranks in the location model generalize naturally to the linear model through (2.11) and (2.12).

To motivate the approach, let us illustrate \( \hat{\alpha}_n(\alpha), 0 < \alpha < 1 \) in the location model which in the present notation may be viewed as (2.11) with \( X_1 = 1_n \). Then \( \hat{\alpha}_n(\alpha) \) specializes to

\[ a_n^* (\alpha) = a_n^*(R_i, \alpha) = \begin{cases} 1 & \text{if } \alpha \leq (R_i-1)/n \\ R_i-\alpha n & \text{if } (R_i-1)/n < \alpha \leq R_i/n \\ 0 & \text{if } R_i/n < \alpha \end{cases} \]  

(2.13)

where \( R_i \) is the rank of \( Y_i \) among \( Y_1, \ldots, Y_n \). The function \( a_n(j, \alpha), j=1, \ldots, n, \ 0 < \alpha < 1, \) coincides with that introduced in Hájek and Šidák (1967, Section V.3.5). Under the general model (2.1) the finite-sample as well as asymptotic properties of the regression rank scores and of the process \( \{\hat{\alpha}_n(\alpha), 0 < \alpha < 1\} \) are studied in the next section.

As in the classical theory of rank tests, we shall consider a score-function \( \phi : (0, 1) \to R^1 \) which is nondecreasing and square-integrable on \( (0, 1) \). We may then construct scores in the following way:

\[ \tilde{b}_n = -\int_0^1 \phi(t) d\hat{\alpha}_n(t), \quad i = 1, \ldots, n. \]  

(2.14)

Defining

\[ S_n = n^{-1/2}(X_{n2} - \hat{X}_{n2})'\hat{b}_n \]  

(2.15)

where \( \hat{b}_n = (\hat{b}_{n1}, \ldots, \hat{b}_{nn})' \), we propose the following statistic for testing \( H_0 \) against \( H_n \):
\[ T_n = S_n \cdot Q_n \cdot S_n / A^2(\phi) \] \hspace{1cm} (2.16)

with

\[ A^2(\phi) = \int_0^1 (\phi(t) - \bar{\phi})^2 dt, \quad \bar{\phi} = \int_0^1 \phi(t) dt \] \hspace{1cm} (2.17)

and with \( Q_n \) defined as in (2.3). An important feature of the test statistic \( T_n \) is that it requires no estimation of nuisance parameters, since the functional \( A(\phi) \) depends only on the score function and not on \( F \). This is familiar from the theory of rank tests, but stands in sharp contrast with other methods of testing in the linear model where typically some estimation of a scale parameter of \( F \) is required to compute the test statistic. See for example the discussion in Aubuchon and Hettmansperger (1988) and Draper (1988).

As we shall show in Section 5, the asymptotic distribution of \( T_n \) under \( H_0 \) is central \( \chi^2 \) with \( q \) degrees of freedom while under \( H_n \) it is noncentral \( \chi^2 \) with \( q \) degrees of freedom and noncentrality parameter

\[ \eta^2 = [\gamma^2(\phi, F) / A^2(\phi)] \beta_0 \cdot Q \beta_0 \] \hspace{1cm} (2.18)

where

\[ \gamma(\phi, F) = -\int_0^1 \phi(t) df (F^{-1}(t)) \] \hspace{1cm} (2.19)

The quantities \( \gamma \) and \( A \) are familiar from the theory of rank tests. The test based on \( T_n \) is asymptotically distribution free in the sense that, under \( H_0 \), neither \( T_n \) nor its asymptotic distribution depend on \( F \). Moreover, it follows from (2.18) that the Pitman efficiency of the test based on \( T_n \) with respect to the classical \( F \) test of \( H_0 \) coincides with that of the two-sample rank test of shift in location with respect to the \( t \)-test. For \( f \) unimodal, we obtain an asymptotically optimal test if we take

\[ \phi(t) = \phi_f(t) = -\frac{F^{-1}(t)}{f(F^{-1}(t))}, \quad 0 < t < 1. \] \hspace{1cm} (2.20)
Thus the asymptotic relative efficiency of the test based on $T_n$, relative to the classical $F$ test is, for Wilcoxon scores $(\phi(u) = u - 1/2)$. is $3/\pi = .955$ at the normal distribution and is bounded below by .864 for all $F$. When $F$ is heavy tailed this asymptotic efficiency is generally greater than one, and can in fact be unbounded. For normal (van der Waerden) scores $(\phi(u) = \Phi^{-1}(u))$ the situation is even more striking. Here the test based on $T_n$ has asymptotic efficiency greater than one, relative to the classical $F$ test, for all symmetric $F$, attaining one at the normal distribution. See e.g. Lehmann (1959, p. 239), and Lehmann (1983, pp 383-87).

Let us now look at the scores (2.14), which can be written as

$$\hat{b}_{ni} = -\int \phi(t) \hat{\alpha}_{ni}^\prime(t) dt \quad i = 1, \ldots, n$$

(2.21)

where the functions $\alpha_{ni}^\prime(t)$ are piecewise constant on $[0,1]$. In the location model, using (2.13) this reduces to

$$\hat{b}_{ni} = n \int_{(R_{i-1})/n}^{R_i/n} \phi(t) dt, \quad i = 1, \ldots, n$$

There are three typical choices of $\phi$:

(i) **Wilcoxon scores**: $\phi(t) = t - 1/2, \quad 0 < t < 1$. The scores are

$$\hat{b}_{ni} = -\int_0^1 (t - 1/2) d \hat{\alpha}_i(t) = \int_0^1 \hat{\alpha}_i(t) dt - 1/2$$

while $A^2(\phi) = 1/12$, and $\gamma(\phi, F) = \int f^2(x) dx$. Wilcoxon scores are optimal when $f$ is the logistic distribution.

(ii) **Normal (van der Waerden) scores**: $\phi(t) = \Phi^{-1}(t), \quad 0 < t < 1$, $\Phi$ being the d.f. of standard normal distribution. Here $A^2(\phi) = 1$ and $\gamma(\phi, F) = \int f (F^{-1}(\Phi(x))) dx$. These scores are asymptotically optimal when $f$ is normal.

(iii) **Median (sign) scores**: $\phi(t) = \frac{1}{2} \text{sign}(t - \frac{1}{2}), \quad 0 < t < 1$, then (2.14) leads to the form

$$\hat{b}_{ni} = \hat{\alpha}_{ni}^\prime(\frac{1}{2})$$

which is non-zero if and only if the $l_i$ estimate passes through the $i$th observation.
REMARK. Using the standard reduction to canonical form e.g. Scheffé (1959, Section 2.6) or Amemiya (1985, Section 1.4.2), we may consider a more general form of the linear hypothesis

\[ R' \beta = r \in \mathbb{R}^q \]  

(2.22)

where \( R \) is a \((p + q) \times q\) matrix of rank \( q < p \). Let \( V \) be a \((p + q) \times p\) matrix such that \( A = [V : R]' \) is nonsingular and \( R' V = 0 \). Set \( \gamma = A \beta \) and \( Z = X A^{-1} \). Partitioning \( \gamma = [\gamma_1', \gamma_2']' \) where \( \gamma_1 = V' \beta \) and \( \gamma_2 = R' \beta \), under the hypothesis (2.22) we have

\[ Y - X R (R'R)^{-1} r = X V (V'V)^{-1} \gamma_1 + E. \]

Thus, in view of the equivariance of regression quantiles we may define \( \bar{Y} = Y - X R (R'R)^{-1} r \), \( \bar{X}_1 = X V (V'V)^{-1} \), \( \bar{X}_2 = X R (R'R)^{-1} \), and proceed as previously discussed with \((\bar{Y}, \bar{X}_1, \bar{X}_2)\) playing the roles of \((Y, X_1, X_2)\). By this device the tests described above and detailed in Section 5 below may be extended to a wide range of applications including, for example, the hypotheses of parallelism and coincidence of regression lines discussed by Adichie (1984) and others.

3. Properties of regression rank scores

Consider the linear regression model (2.1) with design \( X_n \) of dimension \( n \times p \). Let \( \hat{\beta}(\alpha) \in \mathbb{R}^p \) be the \( \alpha \)-regression quantile in (2.11) and \( \hat{a}(\alpha) \in \mathbb{R}^n \) be the vector of \( \alpha \)th regression rank scores defined in (2.15). We see from (2.12) that the regression rank scores are regression invariant, i.e.,

\[ \hat{a}(\alpha, Y + X b) = \hat{a}(\alpha, Y), \quad b \in \mathbb{R}^p. \]  

(3.1)

Moreover, in view of (2.12), we may assume

\[ \sum_{i=1}^{n} x_{ij} = 0, \quad j = 2, \ldots, p \]  

(3.2)

without loss of generality. The formal duality between \( \hat{\beta}(\alpha) \) and \( \hat{a}(\alpha) \) implies that for \( i = 1, \ldots, n \)
\[ \hat{\alpha}_n(\alpha) = \begin{cases} 
1 & \text{if } Y_i \geq \sum_{j=1}^{p} x_{ij} \hat{\beta}_j(\alpha) \\
0 & \text{if } Y_i < \sum_{j=1}^{p} x_{ij} \hat{\beta}_j(\alpha) 
\end{cases} \tag{3.3} \]

while the components of \( \hat{\alpha}_n(\alpha) \) corresponding to \( \{i \mid Y_i = x_i \hat{\beta}(\alpha)\} \) are determined by the equality constraints of (2.12).

Our primary interest in this section will be the properties of the regression rank scores process

\[ \{\hat{\alpha}_n(t) : 0 \leq t \leq 1\}. \tag{3.4} \]

Gutenbrunner and Jurečková (1990) studied the process

\[ W^d_n = \{W^d_n(t) = \sqrt{n} \sum_{i=1}^{n} d_{ia} \hat{\alpha}_n(t) : 0 \leq t \leq 1\} \tag{3.5} \]

and showed that \( W^d_n(t) = U^d_n(t) + o_p(1) \) where

\[ U^d_n(t) = n^{-1/2} \sum_{i=1}^{n} d_{ia} I[E_i > F^{-1}(t)] \tag{3.6} \]

as \( n \to \infty \) uniformly on any fixed interval \([\epsilon, 1-\epsilon]\), where \( 0 < \epsilon < 1/2 \) for any properly standardized triangular array \( \{d_{ia} : i=1, \ldots, n\} \) of vectors from \( \mathbb{R}^p \). They also showed that the process (3.4) (and hence (3.5)) has continuous trajectories and, under the standardization \( \sum_{i=1}^{n} d_{ia} = 0 \), (3.5) is tied-down to 0 at \( t = 0 \), and \( t = 1 \). The same authors also established the weak convergence of (3.5) to the Brownian bridge over \([\epsilon, 1-\epsilon]\). Note however that Theorem V.3.5 in Hájek and Šidák (1967) establishes the weak convergence of (3.5) to the Brownian bridge over the entire interval \([0, 1]\) in the special case of location submodel. Here we extend the results of Gutenbrunner and Jurečková (1990) into the tails of \([0,1]\), in order to find the asymptotic behavior of the rank scores and the test statistics (2.14) and (2.15).

To this end, we will assume that the errors \( E_1, \ldots, E_n \) in (2.1) are independent and identically distributed according to the distribution function \( F(x) \) which has an absolutely
continuous density \( f \). We will assume that \( f \) is positive for \( A < x < B \) and decreases monotonically as \( x \to A^+ \), and \( x \to B^- \) where \(-\infty < A = \sup \{x: F(x) = 0\}\) and \( +\infty > B = \inf \{x: F(x) = 1\}\). For \( 0 < \alpha < 1 \) denote \( \psi_{\alpha} \) the score function corresponding to (2.10)

\[
\psi_{\alpha}(x) = \alpha - I[x < 0], \quad x \in \mathbb{R}^1.
\]

(3.7)

We shall impose the following conditions on \( F \):

(F.1) \[ |F^{-1}(\alpha)| \leq c(\alpha(1-\alpha))^{-a} \quad \text{for} \quad 0 < \alpha \leq \alpha_0, \quad 1-\alpha_0 \leq \alpha < 1, \quad \text{where} \quad 0 < a \leq \frac{1}{4} - \epsilon, \quad \epsilon > 0 \]

and \( c > 0 \).

(F.2) \[ \frac{1}{f}(F^{-1}(\alpha)) \leq c(\alpha(1-\alpha))^{-a} \quad \text{for} \quad 0 < \alpha \leq \alpha_0 \text{ and } 1-\alpha_0 \leq \alpha < 1, \quad c > 0. \]

(F.3) \( f(x) > 0 \) is absolutely continuous, bounded and monotonically decreasing as \( x \to A^+ \) and \( x \to B^- \). The derivative \( f' \) is bounded a.e.

(F.4) \[ \left| \frac{f'(x)}{f(x)} \right| \leq c |x| \quad \text{for} \quad |x| \geq K \geq 0, \quad c > 0. \]

REMARK. These conditions are satisfied, for example, by the normal, logistic, double exponential and \( t \) distributions with 5, or more, degrees of freedom.

The following design assumptions will also be employed.

(X.1) \[ x_{i1} = 1, \quad i = 1, \ldots, n \]

(X.2) \[ \lim_{n \to \infty} D_n = D \quad \text{where} \quad D_n = n^{-1}X_n'X_n \quad \text{and} \quad D \text{ is a positive definite } p \times p \text{ matrix.} \]

(X.3) \[ n^{-1} \sum_{i=1}^{n} \|x_i\|^2 = O(1) \text{ as } n \to \infty. \]

(X.4) \[ \max_{1 \leq i \leq n} \|x_i\| = O(n^{(2(b-a)-\delta)/(1+4b)}) \quad \text{for some} \quad b > 0 \text{ and} \delta > 0 \text{ such that} \quad 0 < b-a < \epsilon/2 \quad (\text{hence} \quad 0 < b < \frac{1}{4} - \epsilon/2). \]

We may now define

\[ \alpha_n^* = n^{-1/(1+4b)} \]

(3.8)

and
\[ \sigma_{\alpha} = \frac{(\alpha(1-\alpha))^{1/2}}{f(F^{-1}(\alpha))}, \quad 0 < \alpha < 1. \] (3.9)

and prove the following crucial lemma.

**LEMMA 3.1**  Assume that \( F \) satisfies (F.1) - (F.4) and that \( X_n \) satisfies (X.1) - (X.3). Then, as \( n \to \infty \),

\[ \sup(|r_n(t, \alpha)|: \|t\| \leq C, \alpha_n^* \leq \alpha \leq 1 - \alpha_n^*) \to 0 \] (3.10)

for any fixed \( C > 0 \), where

\[ r_n(t, \alpha) = (\alpha(1-\alpha))^{-1/2} \sigma_{\alpha}^{-1} \sum_{i=1}^{n} [\rho_\alpha(E_{i\alpha} - n^{-1/2} \sigma_{\alpha} x_i \cdot t) - \rho_\alpha(E_{i\alpha})] \]

\[ + n^{-1/2}(\alpha(1-\alpha))^{-1/2} \sum_{i=1}^{n} x_i \cdot t \psi_\alpha(E_{i\alpha}) - \frac{1}{2} t^\top D_n t \] (3.11)

and

\[ E_{i\alpha} = E_i - F^{-1}(\alpha), \quad i = 1, \ldots, n. \] (3.12)

**PROOF.**

(i) First fix \( \alpha \in [\alpha_n^*, 1 - \alpha_n^*] \) and \( t \) such that \( \|t\| \leq C \).

Define

\[ B_n = \min(n^{-(b-a)/(2(1+4b))}, n^{-2a/(1-4b)}) \] (3.13)

We wish to show that for any \( \lambda > 0 \)

\[ P(|r_n(t, \alpha)| \geq (\lambda+1)B_n) \leq Kn^{-\lambda} \] (3.14)

with a fixed \( K > 0 \). To do this, we will use the Markov inequality

\[ P(|r_n(t, \alpha)| \geq s_n) \leq \exp(-us_n)(M(u) + M(-u)), \quad u > 0 \] (3.15)

where \( M(u) = E\exp(ur_n(t, \alpha)) \).
Denote
\[ \epsilon_m = \epsilon_m(t, \alpha) = n^{-1/2} \sigma_x \gamma t \] (3.16)
and
\[ R_i(t, \alpha) = (\alpha(1-\alpha))^{-1/2} \sigma_x^{-1}[\rho(a(E_{i\alpha} - n^{-1/2} \sigma_x \gamma t) - \rho(a(E_{i\alpha})]
+ n^{-1/2}(\alpha(1-\alpha))^{-1/2} x_i \gamma t \psi_{i \alpha}(E_{i\alpha}) - \frac{1}{2} n^{-1}(x_i \gamma t)^2 \quad i = 1, \ldots, n. \] (3.17)

By definition of \( E_{i\alpha}, \sigma_x, \rho_a \) and \( \psi_{i \alpha}, \)
\[ R_i(t, \alpha) + \frac{1}{2} n^{-1}(x_i \gamma t)^2 = (\alpha(1-\alpha))^{-1/2} \sigma_x^{-1}((E_{i\alpha} - \epsilon_m) I[\epsilon_m < E_{i\alpha} < 0]
+ (\epsilon_m - E_{i\alpha}) I[0 < E_{i\alpha} < \epsilon_m]) \] (3.18)
and hence, uniformly for \( \alpha_n^* \leq \alpha \leq 1 - \alpha_n^*, \|\| \leq C \) and \( i = 1, \ldots, n, \)
\[ |R_i(t, \alpha) + \frac{1}{2} n^{-1}(x_i \gamma t)^2| \leq 2n^{-1/2}(\alpha(1-\alpha))^{-1/2} |x_i \gamma t| = O(n^{-2a+6}/(1+4b)). \] (3.19)

If \( u R_i \) is bdd, that is \( 0 < u < n^{(2a+6)/(1+4b)} \) may be expanded to obtain.
\[ \log M_{R_i}(u) \leq u E R_i(t, \alpha) + cu^2 \text{Var}(R_i(t, \alpha)) \] (3.20)
for some constant \( c > 0. \) By (3.18) and conditions F.1-F.4, for \( \epsilon_m > 0 \) and for \( \alpha_n^* \leq \alpha \leq \alpha_0, \)
\[ 1 - \alpha_n^* \geq \alpha \geq 1 - \alpha_0, \]
\[ E R_i(t, \alpha) + \frac{1}{2} n^{-1}(x_i \gamma t)^2 \leq (\alpha(1-\alpha))^{-1/2} \sigma_x^{-1} \int \epsilon_m^{\beta} (\epsilon_m - z) \int_0^\beta (|F^{-1}(\alpha)| + u) f(F^{-1}(\alpha)) du dz \]
\[ \leq c(\alpha(1-\alpha))^{-1/2-2a} n^{-3/2} |x_i \gamma t|^3 + c(\alpha(1-\alpha))^{-1/2-2a} n^{-2} |x_i \gamma t|^4 \] (3.21)
and we get the same inequality for \( \epsilon_m < 0. \) The same expressions are
\( O(n^{-3/2} |x_i \gamma t|^3) + O(n^{-2} |x_i \gamma t|^4) \) if \( \alpha_0 \leq \alpha \leq 1 - \alpha_0. \) Hence,
\[ \frac{\sum_{i=1}^n} n E R_i(t, \alpha) + \frac{1}{2} n t^{1/4} D_n t = O \left( \frac{n^{-(2b-\alpha)}}{n^{1+4b}} \right). \] (3.22)

Similarly,
\[
\sum_{i=1}^{n} \text{Var} R_i(t, \alpha) \leq \sum_{i=1}^{n} ER_i^2(t, \alpha)
\]
\[
\leq \left( \frac{f(F^{-1}(\alpha))}{\alpha(1-\alpha)} \right) \sum_{i=1}^{n} \int_{0}^{\left|\epsilon_n\right|} (|\epsilon_n|-y)^2 f(F^{-1}(\alpha)+y) dy
\]
\[
= O(n^{-1/2}(\alpha(1-\alpha))^{1/2}/f(F^{-1}(\alpha))) = O(n^{-\frac{(2b-a)}{1+4b}}).
\]

These results hold uniformly in \( \alpha, t \).

Hence, using (3.15) and (3.13) with \( u = \log n/B_n = 0(n^{2a/(1+4b)}) \), so that 3.20 holds,

\[
P(|r_n(t, \alpha)| \geq (\lambda+1)B_n) \leq \exp \left(-\frac{(\lambda+1)\log n + (K \log n / B_n)^{1/4b} + (K \log^2 n / B_n^2) \cdot n^{\frac{(2b-a)}{1+4b}}}{1+b} \right)
\]
\[
\leq n^{-\lambda}
\]

(3.24)

for \( n \geq n_0 \) where \( K > 0 \) and \( n_0 \) do not depend on \( \alpha \) and \( t \).

(ii) Following the proof of Lemma A.2 in Koenker and Portnoy (1987), choose intervals \( [\alpha_i, \alpha_{i+1}] \) of length \( 1/n^5 \) covering \( [\alpha_n^*, 1-\alpha_n^*] \) and balls of radius \( 1/n^5 \) covering \( \{t: \|t\| \leq C\} \).

Let \( (\alpha_1, \alpha_2) \subset [\alpha_i, \alpha_{i+1}] \) and \( t_1, t_2 \) lie in one of the balls covering \( \{t: \|t\| \leq C\} \). For \( i \notin \{l_1, l_2\} \) we use (3.17) (and the boundedness of \( f \) and \( f' \)) to obtain

\[
| R_i(t_1, \alpha_1) - R_i(t_2, \alpha_2) | \leq \left| \frac{f(F^{-1}(\alpha_1)) - f(F^{-1}(\alpha_2))}{\alpha_1(1-\alpha_2)} \right| | F^{-1}(\alpha_1^*) + \epsilon_n | + \frac{f(F^{-1}(\alpha_1^*) + \epsilon_n)}{\alpha_1(1-\alpha_1)} \left| \frac{1}{\alpha_1(1-\alpha_1)} - \frac{1}{\alpha_2(1-\alpha_2)} \right|
\]
\[
+ \frac{f(F^{-1}(\alpha_2))}{\alpha_1(1-\alpha_2)} \left| F^{-1}(\alpha_1) - F^{-1}(\alpha_2) \right|
\]
\[
+ n^{-1/2}(\alpha_1(1-\alpha_1))^{-1/2}\|x_i\| \|t_1-t_2\|
\]
+ Cn^{-1/2}\|x_i\| |(\alpha_1(1-\alpha_1))^{-1/2} - (\alpha_2(1-\alpha_2))^{-1/2}| = O(n^{-2});

here $\alpha^* \in (\alpha_1, \alpha_{i+1})$ and we have used the inequalities $F^{-1}(\alpha_1) - F^{-1}(\alpha_2) \leq \frac{|\alpha_1 - \alpha_2|}{f(F^{-1}(\alpha^*))} = O(n^{-6+\frac{1+\alpha}{1+4b}}) = O(n^{-0.76})$. Hence, on any ball in the covering, $|r_n(t_1, \alpha_1) - r_n(t_2, \alpha_2)| \leq K/n$. Since the number of sets needed to cover the set $S = [\alpha^*, 1-\sigma_n^*] \times \{ \| t \| \leq C \}$ is bounded by $n^{5(p+1)}$ we obtain from (3.14) for $\lambda > 5(p+1)$

$$P(\sup_{(\alpha) \in S} |r_n(t, \alpha)| \geq (\lambda + 1)B_n + K/n) \leq n^{5(p+1)}n^{-\lambda} \rightarrow 0$$

**LEMMA 3.2.** Assume the conditions of Lemma 3.1 and let $d_n = (d_{n1}, \cdots, d_{nn})$ be a sequence of vectors satisfying

$$X_n^*d_n = 0, \quad \frac{1}{n} \sum_{i=1}^{n} d_{ni}^2 \rightarrow \Delta^2, \quad 0 < \Delta^2 < \infty$$

(D.1)

$$n^{-1} \sum_{i=1}^{n} |d_{ni}|^3 = O(1) \quad \text{as} \quad n \rightarrow \infty$$

(D.2)

$$\max_{1 \leq i \leq n} |d_{ni}| = O\left(n^{(\frac{2(b-a)}{\alpha} - \delta)/(1+4b)}\right)$$

(D.3)

then

$$\sup_{(\alpha) \in S} ((\alpha(1-\alpha))^{-1/2}n^{-1/2} \sum_{i=1}^{n} |d_{ni}[\psi_a(E_{i\alpha} - n^{-1/2} \sigma\alpha x_i^* t) - \psi_a(E_{i\alpha}) + n^{-1/2} \sigma\alpha x_i^* t]|) \rightarrow 0 \quad (3.25)$$

as $n \rightarrow \infty$ for any fixed $C > 0$ and for $\alpha^*$ given in (3.8).

**PROOF.** Consider the model

$$Y = X^*\beta^* + E$$

(3.26)

where $X^* = (X_n^* d_n), \quad \beta^* = (\beta_1, \cdots, \beta_p, \beta_{p+1})$. Then
\[
X^\prime X^* = \begin{pmatrix}
X_n^\prime X_n & 0 \\
0 & d_n^\prime d_n
\end{pmatrix}
\]
and the conditions of Lemma 3.1 are satisfied even when replacing \( X \) by \( X^* \). Computing the right derivative of (3.11) with respect to \( t \in \mathbb{R}_{p+1} \), we arrive at (3.25).

Let \( \hat{\beta}_n(\alpha) \) be the \( \alpha \)-regression quantile corresponding to the model (2.1) with the design matrix of order \((n \times p) \); i.e., \( \hat{\beta}_n(\alpha) \) is a solution of the minimization

\[
\sum_{i=1}^{n} \rho_{\alpha}(Y_i - x_i^\prime t) = \min, \quad t \in \mathbb{R}^p.
\]  
(3.27)

The following theorem establishes the rate of consistency of regression quantiles, and is needed for the representation of the dual process.

**THEOREM 3.1.** Under the conditions (F.1) - (F.4) and (X.1) - (X.4),

\[
n^{1/2} \sigma_{\alpha}^{-1}(\hat{\beta}(\alpha) - \beta(\alpha)) = n^{1/2} (\alpha(1-\alpha))^{-1/2} D_n \sum_{i=1}^{n} \psi_{\alpha}(E_{i\alpha}) + o_p(1)
\]  
(3.28)

uniformly in \( \alpha_n^* \leq \alpha \leq 1 - \alpha_n^* \). Consequently,

\[
\sup_{\alpha_n^* \leq \alpha \leq 1 - \alpha_n^*} \| n^{1/2} \sigma_{\alpha}^{-1}(\hat{\beta}_n(\alpha) - \beta(\alpha)) \| = O_p(1).
\]  
(3.29)

**PROOF.** If \( \hat{\beta}_n(\alpha) \) minimizes (3.27), then

\[
T_{n\alpha} = n^{1/2} \sigma_{\alpha}^{-1}(\hat{\beta}_n(\alpha) - \beta(\alpha))
\]  
(3.30)

minimizes the convex function

\[
G_{n\alpha}(t) = (\alpha(1-\alpha))^{-1/2} \sigma_{\alpha}^{-1} \sum_{i=1}^{n} [\rho_{\alpha}(E_{i\alpha} - n^{-1/2} \sigma_{\alpha} x_i^\prime t) - \rho_{\alpha}(E_{i\alpha})]
\]  
(3.31)

with respect to \( t \in \mathbb{R}^p \). By Lemma 3.1, for any fixed \( C > 0 \)

\[
\min _{\| t \| < C} G_{n\alpha}(t) = \min _{\| t \| < C} (-t^\prime Z_{n\alpha} + \frac{1}{2} t^\prime D_n t) + o_p(1)
\]  
(3.32)
uniformly in $\alpha_n^* \leq \alpha \leq 1 - \alpha_n^*$, where
\[
Z_{n\alpha} = n^{-1/2}(\alpha(1 - \alpha))^{-1/2} \sum_{i=1}^{n} x_i \psi_\alpha(E_{i\alpha}).
\] (3.33)

Denoting
\[
U_{n\alpha} = \arg\min_{t \in \mathbb{R}^p} (-t'Z_{n\alpha} + \frac{1}{2}t'D_n t),
\] (3.34)
we immediately get
\[
U_{n\alpha} = D_n^{-1}Z_{n\alpha} = O_p(1)
\] (3.35)
uniformly in $\alpha_n^* \leq \alpha \leq 1 - \alpha_n^*$ and
\[
\min_{t \in \mathbb{R}^p} (-t'Z_{n\alpha} + \frac{1}{2}t'D_n t) = -\frac{1}{2}Z_{n\alpha}D_n^{-1}Z_{n\alpha}.
\] (3.36)

From (3.35) and (3.36), we can write
\[
t'Z_{n\alpha} + \frac{1}{2}t'D_n t = \frac{1}{2}[(t-U_{n\alpha})'D_n(t-U_{n\alpha})-U_{n\alpha}'D_nU_{n\alpha}]
\] (3.37)
and hence we could rewrite (3.10) in the form
\[
sup_{(\alpha,t) \in S} |r_n(t, \alpha)| = \sup_{(\alpha,t) \in S} \{G_{n\alpha}(t) - \frac{1}{2}[(t-U_{n\alpha})'D_n(t-U_{n\alpha})-U_{n\alpha}'D_nU_{n\alpha}] \} \rightarrow 0.
\] (3.38)

Inserting $U_{n\alpha} = O_p(1)$, for $t$, we further obtain
\[
\sup_{\alpha_n \leq \alpha \leq 1 - \alpha_n} \{G_{n\alpha}(U_{n\alpha}) + \frac{1}{2}U_{n\alpha}'D_nU_{n\alpha} \} = o_p(1).
\] (3.39)

We would like to show that
\[
\sup_{\alpha_n \leq \alpha \leq 1 - \alpha_n} \{T_{n\alpha} - U_{n\alpha} \} = o_p(1).
\] (3.40)

Consider the ball $B_{n\alpha}$ with center $U_{n\alpha}$ and radius $\delta > 0$. This ball lies in a compact set with probability exceeding $(1 - \epsilon)$ for $n \geq n_0$; actually, for $t \in B_{n\alpha}$,
\[
\|t\| \leq \|t-U_{n\alpha}\| + \|U_{n\alpha}\| \leq \delta + K_1
\]
for some $K_1$ with probability exceeding $1 - \epsilon$ for $n \geq n_0$. Hence, by (3.10),
Following Pollard (1988), consider the behavior of $G_{na}(t)$ outside $B_{na}$. Suppose $t_a = U_{na} + k \xi$, $k > \delta$ and $\|\xi\| = 1$. Let $t_a^*$ be the boundary point of $B_{na}$ that lies on the line from $U_{na}$ to $t_a$, i.e., $t_a^* = U_{na} + \delta \xi$. Then $t_a^* = (1 - (\delta/k))U_{na} + (\delta/k)t_a$ and hence, by (3.38) and (3.39),

$$
\frac{\delta}{k} G_{na}(t) + (1-\frac{\delta}{k}) G_{na}(U_{na}) \geq G_{na}(t_a^*) \geq \frac{1}{2} \delta^2 \lambda_0 + G_{na}(U_{na}) - 2 \Delta_{na}
$$

where $\lambda_0$ is the minimal eigenvalue of $D$. Hence,

$$
\inf_{|t - U_{na}| \geq \delta} G_{na}(t) \geq G_{na}(U_{na}) + (k / \delta)(\frac{1}{2} \delta^2 \lambda_0 - 2 \Delta_{na}).
$$

Using (3.39) the last term is positive with probability tending to one uniformly in $\alpha$ for any fixed $\delta > 0$. Hence, given $\delta > 0$ and $\epsilon > 0$, there exist $n_0$ and $\eta > 0$ such that for $n \geq n_0$,

$$
P(\inf_{a_n \leq \alpha \leq 1 - a_n} \inf_{|t - U_{na}| \geq \delta} G_{na}(t) - G_{na}(U_{na}) > \eta) > 1 - \epsilon
$$

and hence (since the event in (3.43) implies that $G_{na}$ must be minimized inside the ball of radius $\delta$) $P(\sup_{a_n \leq \alpha \leq 1 - a_n} \|T_{na} - U_{na}\| \leq \delta) \rightarrow 1$ for any fixed $\delta > 0$, as $n \rightarrow \infty$. $\blacksquare$

The following theorem approximates the regression rank score process by an empirical process.

**THEOREM 3.2.** Let $d_n$ satisfy (D.1) - (D.3), $X_n$ satisfy (X.1) - (X.4) and $F$ satisfy (F.1) - (F.4). Then

$$
\sup_{a_n \leq \alpha \leq 1 - a_n} \left\{ |n^{-1/2}(\alpha(1-\alpha))^{-1/2} \sum_{i=1}^{n} d_n(\hat{\theta}_n(\alpha) - \bar{\theta}_i(\alpha))| \right\} \rightarrow 0
$$

as $n \rightarrow \infty$, where

$$
\bar{\theta}_i(\alpha) = I[E_i \geq F^{-1}(\alpha)], \quad i = 1, \ldots, n.
$$

**PROOF.** Insert $n^{1/2} \sigma_{-1}^{-1}(\hat{\theta}_n(\alpha) - \bar{\theta}(\alpha))$ for $t$ in (3.25) and notice (3.29) and the fact that
from which (3.44) follows. ■

The following theorem which follows from Theorem 3.2 is an extension of Theorem V.3.5 in Hájek and Šidák (1967) to the regression rank scores. Some applications of this result to Kolmogorow-Smirnov type tests will appear in Jurečková (1990).

**THEOREM 3.3.** Under the conditions of Theorem 3.2, as \( n \to \infty \),

\[
\sup_{0 \leq \alpha \leq 1} \left( n^{-1/2} \sum_{i=1}^{n} d_{\alpha} (\hat{\alpha}_{\alpha}(\alpha) - \bar{\alpha}_{\alpha}(\alpha)) \right) \to 0
\]  

(3.47)

Moreover, the process

\[
\{\Delta^{-1} n^{-1/2} \sum_{i=1}^{n} d_{\alpha} \hat{\alpha}_{\alpha}(\alpha) : 0 \leq \alpha \leq 1\}
\]

(3.48)

converges to the Brownian bridge in the Prokhorov topology on \( C[0, 1] \).

**PROOF.** By Theorem 3.2,

\[
\sup_{0 \leq \alpha \leq 1} \left( n^{-1/2} \sum_{i=1}^{n} d_{\alpha} (\hat{\alpha}_{\alpha}(\alpha) - \bar{\alpha}_{\alpha}(\alpha)) \right) \to 0.
\]  

(3.49)

Further,

\[
\sup_{0 \leq \alpha \leq 1} \left( n^{-1/2} \sum_{i=1}^{n} d_{\alpha} \hat{\alpha}_{\alpha}(\alpha) \right) = \sup_{0 \leq \alpha \leq 1} \left( n^{-1/2} \sum_{i=1}^{n} d_{\alpha} (1 - \hat{\alpha}_{\alpha}(\alpha)) \right) \leq n^{1/2} \max_{1 \leq i \leq n} d_{\alpha} |\alpha_{n}|^\alpha
\]

(3.50)

and we obtain an analogous conclusion for \( \sup_{1 - \alpha \leq \alpha \leq 1} \left( n^{-1/2} \sum_{i=1}^{n} d_{\alpha} \hat{\alpha}_{\alpha}(\alpha) \right) \). On the other hand,

\[
\sup_{0 \leq \alpha \leq 1} \left( n^{-1/2} \sum_{i=1}^{n} d_{\alpha} \bar{\alpha}_{i}(\alpha) - \hat{\alpha}_{\alpha}(\alpha) \right) = \sup_{0 \leq \alpha \leq 1} \left( n^{-1/2} \sum_{i=1}^{n} d_{\alpha} (I[E_i < F^{-1}(\alpha)] - \alpha) \right)
\]

(3.51)
and analogously

\[
\sup_{1 - \alpha_n \leq \alpha \leq 1} |n^{-1/2} \sum_{i=1}^{n} d_i \lambda_i'(\alpha)| = o_p(1).
\]

Thus (3.47) follows, and consequently (3.48).

4. Asymptotic properties of simple linear regression rank scores statistics

Maintaining the notation of Section 3, let \( \phi(t): 0 < t < 1 \) be a nondecreasing square-integrable score-generating function and let \( \hat{\beta}_{i,n}, i = 1, \ldots, n \) be the scores defined by (2.14). Let \( \{d_n\} \) be a sequence of vectors satisfying (D.1) - (D.3).

Following Hájek and Šidák (1967), we shall call the statistics

\[
S_n = n^{-1/2} \sum_{i=1}^{n} d_i \hat{\beta}_{i,n}
\]

the simple linear regression rank-score statistics, or just simple linear rank statistics. Our primary objective in this section is to investigate the conditions on \( \phi \) under which we may integrate (3.47) and obtain an asymptotic representation for \( S_n \) of the form

\[
S_n = n^{-1/2} \sum_{i=1}^{n} d_i \phi(E_i) + o_p(1).
\]

We shall prove (4.2) for \( \phi \) satisfying a condition of the Chernoff-Savage (1958) type; thus our results will cover Wilcoxon, van der Waerden, and median scores, among others.

**THEOREM 4.1.** Let \( \phi(t): 0 < t < 1 \), be a nondecreasing square integrable function such that \( \phi'(t) \) exists for \( 0 < t < \alpha_0, \ 1 - \alpha_0 < t < 1 \) and satisfies

\[
|\phi'(t)| \leq c (t(1-t))^{-1-\delta}
\]
for some \( \delta^* < \delta \) where \( \delta \) is given in condition (X.4), and for \( t \in (0, \alpha_0) \cup (1-\alpha_0, 1) \). Then, under (F.1) - (F.4), (X.1) - (X.4) and (D.1) - (D.3), the statistic \( S_n \) admits the representation (4.2) and hence is asymptotically normally distributed with zero expectation and with variance

\[
\Delta^2 \left( \int_0^1 \phi^2(t) dt - \bar{\phi}^2 \right), \quad \bar{\phi} = \int_0^1 \phi(t) dt.
\]

(4.4)

**PROOF.** Let us consider \( S_n \) defined in (4.1) with the scores (2.14). Integrating by parts (notice that \( \hat{a}_m(t) - \bar{a}_i(t) = 0 \) for \( t = 0, 1 \)), we obtain

\[
-n^{-1/2} \sum_{i=1}^n d_m \int_0^1 \phi(t) d(\hat{a}_m(t) - \bar{a}_i(t)) = n^{-1/2} \sum_{i=1}^n d_m \int_0^1 (\hat{a}_m(t) - \bar{a}_i(t)) d \phi(t).
\]

(4.5)

which we must show is \( o_p(1) \). We shall split the domain of integration into the intervals \((0, \alpha_n^*), (\alpha_n^*, \alpha_0), [\alpha_0, 1-\alpha_0], (1-\alpha_0, 1-\alpha_n^*), [1-\alpha_n^*, 1)\) and denote the respective integrals by \( I_1, \ldots, I_5 \). Regarding Theorem 3.2, we immediately get that \( I_5 \nrightarrow 0 \) by the dominated convergence theorem. Similarly,

\[
|I_2| \leq \int_{\alpha_n^*}^{\alpha_0} |\phi'(t)| \left| n^{-1/2} \sum_{i=1}^n d_m (\hat{a}_m(t) - \bar{a}_i(t)) \right| dt
\]

\[
\leq c \int_{\alpha_n^*}^{\alpha_0} (t(1-t))^{-1/2} \cdot |n^{-1/2} t(1-t))^{-1/2} \sum_{i=1}^n d_m (\hat{a}_m(t) - \bar{a}_i(t)) | dt
\]

\[
= c \int_{\alpha_n^*}^{\alpha_0} (t(1-t))^{-1/2} dt \cdot o_p(1) = o_p(1).
\]

Finally,

\[
|I_1| \leq n^{-1/2} \max_{1 \leq i \leq n} |d_m| \int_0^{\alpha_n^*} |\phi'(t)| \sum_{i=1}^n |\hat{a}_m(t) - \bar{a}_m(t)| dt \leq I_{11} + I_{12}
\]

where

\[
I_{11} = n^{-1/2} \max_{1 \leq i \leq n} |d_m| \int_0^{\alpha_n^*} |\phi'(t)| \sum_{i=1}^n (1 - \hat{a}_m(t)) dt
\]

(4.6)

and
\[ I_{12} = n^{-1/2} \max_{1 \leq i \leq n} |d_m| \int_0^{\alpha_n^*} |\phi'(t)| \sum_{i=1}^n (1 - \overline{\alpha}_i(t)) dt. \] (4.7)

Then

\[ I_{11} \leq n^{1/2} \max_{1 \leq i \leq n} |d_m| \int_0^{\alpha_n^*} t^{-\delta^*} dt = O(n^{1/2 + \frac{2(b-a)-\delta}{1+4b} - \frac{(1-\delta^*)}{1+4b}}) = O(n^{-2(\delta-\delta^*)}). \]

Finally,

\[ I_{12} = n^{-1/2} \sum_{i=1}^n d_m \int_0^{\alpha_n^*} \phi'(t) I[t > F(E_i)] dt = n^{-1/2} \sum_{i=1}^n d_m [\phi(\alpha_n^*) - \phi(F(E_i))] I(F(E_i) < \alpha_n^*) \]

Now we may assume that \( \phi(\alpha_n^*) < 0 \) for \( n \geq n_0 \), since otherwise if \( \phi \) were bounded from below then \( I_{12} \to 0 \). Hence

\[ \text{Var}(I_{12}) \leq n^{-1} \sum_{i=1}^n d_m^2 E[(2\phi(F(E_i)))^2] I(F(E_i) < \alpha_n^*) \leq \int_0^{\alpha_n^*} \phi^2(u) du \cdot O(1) \to 0 \]

due to the square-integrability of \( \phi \). Treating the integrals \( I_4, I_6 \) analogously, we arrive at (4.5) and this proves the representation (4.2).

5. Tests of linear subhypotheses based on regression rank scores

Returning to the model (2.2), assume that the design matrix \( X = (X_1 ; X_2) \) satisfies the conditions (X.1) - (X.4), (2.3) and (2.4). We want to test the hypothesis \( H_0: \beta_2 = 0 \) (\( \beta_1 \) unspecified) against the alternative \( H_n: \beta_{2n} = n^{-1/2} \beta_0 \) (\( \beta_0 \in R^q \) fixed).

Let \( \hat{\alpha}_n(\alpha) = (\hat{\alpha}_{n1}(\alpha), ..., \hat{\alpha}_{nm}(\alpha)) \) denote the regression rank scores corresponding to the submodel

\[ Y = X_1 \beta_1 + E \quad \text{under} \quad H_0. \] (5.1)

Let \( \phi(t): (0, 1) \to R^1 \) be a nondecreasing and square integrable score-generating function. Define the scores \( \hat{\delta}_m, i=1,...,n \) by the relation (2.14), and consider the test statistic

\[ T_n = S_n Q_n^{-1} S_n / A^2(\phi) \] (5.2)
where

$$S_n = n^{-1/2}(X_n - \hat{X}_n)/\hat{b}_n$$  \hspace{1cm} (5.3)$$

and where $Q_n$ and $A^2(\phi)$ are defined in (2.4) and (2.17), respectively. The test is based on the asymptotic distribution of $T_n$ under $H_0$, given in the following theorem. Thus, we shall reject $H_0$ provided $T_n \geq \chi_q^2(\omega)$, i.e. provided $T_n$ exceeds the $\omega$ critical value of the $\chi^2$ distribution with $q$ d.f. The same theorem gives the asymptotic distribution of $T_n$ under $H_n$ and thus shows that the Pitman efficiency of the test coincides with that of the classical rank test.

**THEOREM 5.1.** Assume that $X_i$ satisfies (X.1) - (X.4) and $(X_1; X_2)$ satisfies (2.3) and (2.4). Further assume that $F$ satisfies (F.1) - (F.4). Let $T_n$ defined in (5.3) and (5.4) be generated by the score function $\phi$ satisfying (4.3), and nondecreasing and square-integrable on $(0, 1)$.

(i) Then, under $H_0$, the statistic $T_n$ is asymptotically central $\chi^2$ with $q$ degrees of freedom.

(ii) Under $H_n$, $T_n$ is asymptotically noncentral $\chi^2$ with $q$ degrees of freedom and with non-centrality parameter,

$$\eta^2 = \beta_0'Q\beta_0 \cdot \gamma(\phi, F)/A^2(\phi)$$  \hspace{1cm} (5.4)$$

with

$$\gamma(\phi, F) = -\int_0^1 \phi(t)df(F^{-1}(t)).$$  \hspace{1cm} (5.5)$$

**PROOF.**

(i) It follows from Theorem 4.1 that, under $H_0$, $S_n$ has the same asymptotic distribution as

$$\bar{S}_n = n^{-1/2}(X_n - \hat{X}_n)/\bar{b}_n$$

where $\bar{b}_n = (\bar{b}_1, \ldots, \bar{b}_n)'$ and $\bar{b}_n = \phi(F(E_i)), i = 1, \ldots, n$. The asymptotic distribution of $\bar{S}_n$ follows from the central limit theorem and coincides with $q$-dimensional normal distribution with expectation 0 and the covariance matrix $Q \cdot A^2(\phi)$. 
(ii) The sequence of local alternatives $H_n$ is contiguous with respect to the sequence of null distributions with the densities \( \prod_{i=1}^{n} f(e_i) \). Hence, (4.1) holds also under $H_n$ and the asymptotic distributions of $\tilde{S}_n$ under $H_n$ coincide. The proposition then follows from the fact that the asymptotic distribution of $\tilde{S}_n$ under $H_n$ is normal $N_\theta(\gamma(\phi, F)Q_0, QA^2(\phi))$.

6. An Example

To illustrate the tests proposed above we consider briefly an example taken from Adichie (1984, Example 3). The log of the leaf burn (in seconds) of 30 batches of tobacco is thought to depend upon the percent composition of nitrogen, chlorine, and potassium. Adichie suggests testing the potassium effect and describes an aligned rank version of the test. We are unable to reproduce some details of his calculations, however, using his approach we get least squares estimates of the nitrogen and chlorine effects of -.529 and -.290 with an intercept of 2.653. With these preliminary estimates we obtain aligned ranks

\[
\begin{array}{cccccccccccc}
7 & 17 & 2 & 18 & 6 & 1 & 11 & 3 & 30 & 13 \\
25 & 16 & 4 & 29 & 26 & 27 & 21 & 23 & 19 & 12 \\
28 & 10 & 8 & 15 & 24 & 20 & 22 & 5 & 14 & 9 \\
\end{array}
\]

which yield a test statistic of 13.59 highly significant relative to the 1% $\chi^2$ critical value of 6.63.

In contrast the Wilcoxon regression rank scores computed as

\[ \hat{b}_i = -\int_0^1 (t - 1/2)d\hat{a}_i(t) = \int_0^1 \hat{a}_i(t) dt - 1/2 \]

and based on the restricted model excluding potassium, are

\[
\begin{array}{cccccccccccc}
-0.27 & 0.06 & -0.41 & 0.09 & -0.32 & -0.48 & -0.17 & -0.38 & 0.48 & -0.06 \\
0.23 & 0.04 & -0.37 & 0.42 & 0.28 & 0.37 & 0.19 & 0.41 & 0.15 & -0.26 \\
0.38 & -0.16 & -0.23 & -0.01 & 0.33 & 0.12 & 0.15 & -0.42 & -0.10 & -0.06 \\
\end{array}
\]

and yield a test statistic of 13.17. The full set of regression rank scores $\hat{a}_i(t)$ for this data are illustrated in Figure 6.1 with the plots ordered according to their Wilcoxon rank score. Note
that as a practical matter when $\Phi = \int_0^1 \phi(t)dt = 0$, we may omit the $\hat{X}_2$ term in the computation of $S_n$ in (5.3) since $\hat{b}_n$ is orthogonal to $X_1$. This is in contrast with the aligned rank situation where the use of $X_2 - \hat{X}_2$ is essential.

Corresponding calculations for the normal scores using

$$\hat{b}_i = -\int_0^1 \Phi^{-1}(t)d\hat{a}_i(t) = \sum_{j=1}^{J_n} \hat{a}_i '(t_j)[\phi(\Phi^{-1}(t_j)) - \phi(\Phi^{-1}(t_{j-1}))]$$

yields

<p>| | | | | | | | | |</p>
<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>-0.74</td>
<td>0.15</td>
<td>-1.41</td>
<td>0.23</td>
<td>-0.91</td>
<td>-2.13</td>
<td>-0.45</td>
<td>-1.17</td>
<td>2.08</td>
</tr>
<tr>
<td>0.63</td>
<td>0.10</td>
<td>-1.25</td>
<td>1.44</td>
<td>0.78</td>
<td>1.15</td>
<td>0.50</td>
<td>1.35</td>
<td>0.40</td>
</tr>
<tr>
<td>1.41</td>
<td>-0.40</td>
<td>-0.61</td>
<td>-0.03</td>
<td>0.94</td>
<td>0.30</td>
<td>0.39</td>
<td>-1.45</td>
<td>-0.26</td>
</tr>
</tbody>
</table>

and a test statistic of 12.87. The corresponding normal score aligned rank statistic is 11.72.

Finally, regression rank score version of the sign test yields the scores

<p>| | | | | | | | | |</p>
<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>-1.00</td>
<td>1.00</td>
<td>-1.00</td>
<td>1.00</td>
<td>-1.00</td>
<td>-1.00</td>
<td>-1.00</td>
<td>1.00</td>
<td>-1.00</td>
</tr>
<tr>
<td>1.00</td>
<td>1.00</td>
<td>-1.00</td>
<td>1.00</td>
<td>1.00</td>
<td>1.00</td>
<td>1.00</td>
<td>0.16</td>
<td>-1.00</td>
</tr>
<tr>
<td>1.00</td>
<td>-1.00</td>
<td>-1.00</td>
<td>-0.37</td>
<td>1.00</td>
<td>1.00</td>
<td>1.00</td>
<td>-1.00</td>
<td>-0.79</td>
</tr>
</tbody>
</table>

and a test statistic of 8.42 while the aligned rank sign scores yield 10.20. Obviously, all versions of the tests lead to a decisive rejection of the null. Note that for the sign scores the test coincides with the $L^1$ Lagrange multiplier test discussed in Koenker and Bassett (1982).

Since an important objective of the proposed rank tests is robustness to outlying observations, it is interesting to observe the effect of perturbing the first $y$ observation of the Adichie data set on the aligned and rank scores versions of the test statistic. This sensitivity analysis is illustrated in Figure 6.2. Even a modest perturbation in $y_1$ is enough to confound the initial least squares estimate and reverse the conclusion of the aligned rank test. However the regression rank score version of the test is seen to be relatively insensitive to such perturbations. One should be aware that comparable perturbations in the design observations may wreck havoc even with the rank score form of the test. Recent work of Antoch and Jurečková (1985) and deJongh, deWet, and Welsh (1988) contain suggestions on robustifying
regression quantiles to the effect of influential design points.

Computation of the tests was carried out in S+ using the algorithm described in Koenker and d'Orey (1988, 1990) to compute regression quantiles.

REFERENCES


Figure 6.1

Regression Rank Scores for Tobacco Data

![Graphs showing regression rank scores for tobacco data]
Figure 6.1
(continued)
Sensitivity Curves for Rank Tests

Value of the test statistic

Perturbation of $y(1)$