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

December 1992

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 (1992) 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, indeed they are invariant with respect to a regression shift of the nuisance parameters.

AMS 1980 subject classifications: 62G10, 62J05

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.

Key words and phrases: Ranks, Regression quantiles, Regression rank scores

1. Introduction

Several authors including Koul (1970), Puri and Sen (1985) and Adichie (1978) have developed asymptotically distribution-free tests of linear hypotheses for the linear regression model based upon aligned rank statistics. Excellent reviews of these results including extensions to multivariate models may be found in Puri and Sen (1985) and the survey paper of Adichie (1984). The hypothesis under consideration typically involves nuisance parameters which require preliminary estimation; the aligned (or signed) rank statistics are then based on residuals from the preliminary estimate. Alternative approaches to inference based on rank *estimation* have been considered by McKean and Hettmansperger(1978), Aubuchon and Hettmansperger (1988) and Draper (1988) among others.

A completely new approach to the construction of rank statistics for the linear model has recently been introduced by Gutenbrunner and Jurečková (1992). Their approach is based on the dual solutions to the regression quantile statistics of Koenker and Bassett (1978). These regression rank scores represent a natural extension of the "location rank scores" introduced by Hájek and Šidák (1967, Section V.3.5), which play a fundemental role in the classical theory of rank statistics. In this paper we consider tests of a general linear hypothesis for the linear regression model based upon regression rank scores. These tests have the advantages of more familiar rank tests: they are robust to outliers in the response variable and they are asymptotically distribution free in the sense that no nuisance parameter depending on the error distribution need be estimated in order to compute the test statistic. Furthermore, they are considerably simpler than many of the proposed aligned rank tests which require preliminary estimation of the linear model by computationally demanding rank estimation methods. The robustness of the proposed tests and the sensitivity of the aligned rank procedures to response outliers is illustrated in the sensitivity analysis of the example discussed in Section 2.

In the classical linear model,

$$\mathbf{Y} = \mathbf{X}\beta + \mathbf{E}\,, \tag{1.1}$$

the vector $\hat{\beta}(\alpha) \equiv (\hat{\beta}_1(\alpha),\ldots, \hat{\beta}_p(\alpha))' \in \mathbf{R}^p$ of αth regression quantiles is any solution of the problem

$$\min \sum_{i=1}^{n} \rho_\alpha(Y_i - \mathbf{x}_{1i}'\mathbf{t}), \qquad \mathbf{t} \in \mathbf{R}^p \tag{1.2}$$

where

$$\rho_\alpha(u) = |u| \,\{(1-\alpha) I[u<0] + \alpha\, I[u>0]\}, \qquad u \in \mathbf{R}^1. \tag{1.3}$$

Least absolute error regression corresponds to the median case with $\alpha = ½$. In the one-sample location model, with $\mathbf{X} = \mathbf{1}_n$, solutions to (1.2) are the ordinary sample quantiles: when $n\alpha$ is an integer we have an interval of solutions between two adjacent order statistics. Computation of the regression quantiles is greatly facilitated by expressing (1.2) as the linear program

$$\alpha \mathbf{1}_n'\mathbf{u}^+ + (1-\alpha)\mathbf{1}_n'\mathbf{u}^- \; := \min$$

$$\mathbf{X}\hat{\beta} + \mathbf{u}^+ - \mathbf{u}^- = \mathbf{Y} \tag{1.4}$$

$$\beta \in \mathbf{R}^p, \quad \mathbf{u}^+, \mathbf{u}^- \in \mathbf{R}_+^n$$

and $\mathbf{1}_n = (1,..., 1)' \in \mathbf{R}^n$, with $0 < \alpha < 1$. Even in this form, the problem of finding *all* the regression quantile solutions may appear computationally demanding, since there would appear to be a distinct problem to solve for each $\alpha \in (0,1)$. Fortunately, there are only a few *distinct* solutions. In the location model we know, of course, that there are at most *n* distinct quantiles. In regression, Portnoy(1991) has shown that the number of distinct solutions to (1.2) is $O_p(n \log n)$. Finding all the regression quantiles is a straightforward exercise in *parametric* linear programming. From any given solution for fixed α we may compute the interval containing α for which is solution remains optimal, and one simplex pivot brings us to a new solution at either endpoint of the interval. Proceeding in this way we may compute the entire path $\hat{\beta}(\cdot)$ which is a piecewise constant function from $[0,1]$ to $\mathbf{R}^p$. Detailed descriptions of algorithms to compute the regression quantiles may be found in Koenker and d'Orey(1990), and Osborne(1992). 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 (1992), and Portnoy(1991b).

The regression rank scores introduced in Gutenbrunner and Jurečková (1992) arise as a *n*-vector $\hat{\mathbf{a}}_n(\alpha) = (\hat{a}_{n1}(\alpha),..., \hat{a}_{nn}(\alpha))'$ of solutions to the dual form of the linear program required to compute the regression quantiles. The formal dual program to (1.4) can be written in the form

$$\begin{aligned} &\mathbf{Y}'\hat{\mathbf{a}}(\alpha) \ := \max \\ &\mathbf{X}'\hat{\mathbf{a}}(\alpha) = (1-\alpha)\mathbf{X}'\mathbf{1}_n \\ &\hat{\mathbf{a}}(\alpha) \in [0,\ 1]^n, \quad 0 < \alpha < 1 \end{aligned} \tag{1.5}$$

As shown in Gutenbrunner and Jurečková (1992), many aspects of the duality of order statistics and ranks in the location model generalize naturally to the linear model through (1.4) and (1.5). Moreover, as pointed out there, $\hat{\mathbf{a}}$ is regression invariant with respect to $\mathbf{X}_1$, in the sense that $\hat{\mathbf{a}}(\alpha)$ is unchanged if Y is transformed to $Y + \mathbf{X}_1\gamma$ for any $\gamma \in \mathbf{R}^p$.

To motivate our approach, consider $\{\hat{\mathbf{a}}_n(\alpha),\ 0 < \alpha < 1\}$ in the location model with $\mathbf{X} = \mathbf{1}_n$. In this case, $\hat{a}_{ni}(\alpha)$ specializes to

$$\hat{a}_{ni}(\alpha) \equiv a_n^*(R_i,\ \alpha) \equiv \begin{cases} 1 & \text{if } \alpha \le (R_i-1)/n \\ R_i-\alpha n & \text{if } (R_i-1)/n < \alpha \le R_i/n \\ 0 & \text{if } R_i/n < \alpha \end{cases} \tag{1.6}$$

where R_i is the rank of Y_i among $Y_1,..., Y_n$. The function $a_n^*(j,\ \alpha)$, $j=1,..., n$, $0 < \alpha < 1$, coincides exactly with that introduced in Hájek and Šidák (1967, Section V.3.5). Under the general model (1.1), both the finite-sample and asymptotic properties of the regression rank scores and of the process $\{\hat{\mathbf{a}}_n(\alpha),\ 0 < \alpha < 1\}$ are described in the next section. The regression rank score process may be efficiently computed by standard parametric linear programming techniques, essentially as a byproduct of the regression quantile computation requiring no additional computational effort and only some additional storage. See Koenker and d'Orey(1990) for algorithmic details.

The formal duality between $\hat{\beta}(\alpha)$ and $\hat{\mathbf{a}}(\alpha)$ implies that for i=1, ..., n

$$\hat{a}_{ni}(\alpha) = \begin{cases} 1 \text{ if } & Y_i > \sum_{j=1}^{p} x_{ij}\hat{\beta}_j(\alpha) \\ 0 \text{ if } & Y_i < \sum_{j=1}^{p} x_{ij}\hat{\beta}_j(\alpha) \end{cases} \tag{1.7}$$

while the components of $\hat{\mathbf{a}}_n(\alpha)$ corresponding to $\{i \mid Y_i = \mathbf{x}_i'\hat{\beta}(\alpha)\}$ are determined by the equality constraints of (1.5). Thus, as in the location model, the regression rankscore for observation i is one while y_i is above the αth quantile regression plane, and zero when y_i falls below this plane, and taking an intermediate value while y_i falls on the αth plane. Integrating the regression rankscore function for each observation over [0,1] yields a vector of (Wilcoxon) ranks: observations falling "below" most of the others receiving small ranks, while those falling "above" the others, and thus having rankscore one over a wide interval, receive large ranks. This observation is completely transparent in the location model where "above" and "below" have an obvious interpretation. In regression, the interpretation of these terms relies on the optimization problem defining the regression quantiles. The resulting rank scores illustrated, for example, in Figure 6.1, are, we believe, a useful graphical diagnostic in linear regression in addition to their role in formal hypothesis testing.

The next section of the paper surveys our results, establishes some notation, and provides an illustrative example. 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, and Section 5 contains a formal treatment of the proposed tests.

2. Notation and preliminary considerations

We will partition the classical linear regression model

$$\mathbf{Y} = \mathbf{X}\beta + \mathbf{E} \tag{2.1}$$

as

$$\mathbf{Y} = \mathbf{X}_1\beta_1 + \mathbf{X}_2\beta_2 + \mathbf{E} \tag{2.2}$$

where β_1 and β_2 are $p-$ and $q-$dimensional parameters, $\mathbf{X} = \mathbf{X}_n$ is a known, $n\times(p+q)$ design matrix with rows $\mathbf{x}_{ni}' = \mathbf{x}_i' = (\mathbf{x}_{1i}', \mathbf{x}_{2i}') \in R^{p+q}$, i=1,..., n . We will assume throughout that $\mathbf{x}_{i1} = 1$ for $i = 1,...,n$. $\mathbf{Y}$ is a vector of observations and $\mathbf{E}$ is an $n\times 1$ vector of *i.i.d.* errors with common distribution function F. As in the familiar two-sample rank test, our test statistic is shift-invariant and hence independent of location. Thus like other rank tests, hypotheses on the intercept cannot be tested. This is immediately apparent from the regression invariance of the test statistic noted above. 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 $-\infty \le A = \sup\{x: F(x) = 0\}$ and $+\infty \ge B = \inf\{x: F(x) = 1\}$. Moreover, we shall impose some conditions on the tails of f assuming, among other conditions, that f monotonically decreases to 0 when $x \to A+$, or $x \to B-$. Define $\mathbf{D}_n = n^{-1}\mathbf{X}_1'\mathbf{X}_1$,

$$\mathbf{H}_1 = \mathbf{X}_1(\mathbf{X}_1{}'\mathbf{X}_1)^{-1}\mathbf{X}_1{}' \quad \text{and} \quad \mathbf{Q}_n = n^{-1}(\mathbf{X}_2 - \hat{\mathbf{X}}_2)'(\mathbf{X}_2 - \hat{\mathbf{X}}_2) \tag{2.3}$$

with $\hat{\mathbf{X}}_2 = \mathbf{H}_1\mathbf{X}_2$ being the projection of $\mathbf{X}_2$ on the space spanned by the columns of $\mathbf{X}_1$. We shall also assume

$$\lim_{n\to\infty} \mathbf{D}_n = \mathbf{D}, \quad \lim_{n\to\infty} \mathbf{Q}_n = \mathbf{Q} \tag{2.4}$$

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

We are interested in testing the hypothesis

$$H_0 : \ \beta_2 = 0, \quad \beta_1 \text{ unspecified} \tag{2.5}$$

versus the Pitman (local) alternatives

$$H_n : \ \beta_{2n} = n^{-1/2}\beta_0 \tag{2.6}$$

with β_0 being a fixed vector in $\mathbf{R}^q$.

As in the classical theory of rank tests, we shall consider a score-function $\varphi : \ (0, 1) \to \mathbf{R}$ which is nondecreasing and square-integrable on (0, 1). We may then construct scores based on the regression rankscore process following Hájek and Šidák, (1967) as,

$$\hat{b}_{ni} = -\int_0^1 \varphi(t)d\hat{a}_{ni}(t), \quad i=1,\ldots, n. \tag{2.7}$$

Defining

$$\mathbf{S}_n = n^{-1/2}(\mathbf{X}_{n2} - \hat{\mathbf{X}}_{n2})'\hat{\mathbf{b}}_n \tag{2.8}$$

where $\hat{\mathbf{b}}_n = (\hat{b}_{n1},\ldots, \hat{b}_{nn})'$, we propose the following statistic for testing H_0 against H_n:

$$T_n = \mathbf{S}_n{}'\mathbf{Q}_n^{-1}\mathbf{S}_n \,/\, A^2(\varphi) \tag{2.9}$$

where

$$A^2(\varphi) = \int_0^1 (\varphi(t) - \bar{\varphi})^2 dt, \quad \bar{\varphi} = \int_0^1 \varphi(t)dt \tag{2.10}$$

and with $\mathbf{Q}_n$ 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 $A(\varphi)$ depends only on the score function and not on (the unknown) 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).

We shall show in Section 5, that the asymptotic distribution of T_n under H_0 is central χ^2 with q degrees of freedom while under H_n it is noncentral χ^2 with q degrees of freedom and noncentrality parameter

$$\eta^2 = [\gamma^2(\varphi, F) \,/\, A^2(\varphi)]\beta_0{}'\mathbf{Q}\beta_0 \tag{2.11}$$

where

$$\gamma(\varphi, F) = -\int_0^1 \varphi(t) df\,(F^{-1}(t)). \tag{2.12}$$

Like A, γ is also familiar from the classical theory of rank tests. The test statistic T_n is first-order asymptotically distribution free in the sense that the first-order term in its asymptotic representation is exactly distribution free, as follows from (4.2). Moreover, it follows from (2.11) 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

$$\varphi(t) = \varphi_f(t) = -\frac{f'(F^{-1}(t))}{f\,(F^{-1}(t))}, \quad 0 < t < 1. \tag{2.13}$$

Thus for Wilcoxon scores (see below) the asymptotic relative efficiency of the test based on T_n relative to the classical F test 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 ($\varphi(u) = \Phi^{-1}(u)$) 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 examine more closely the scores (2.7), which can be written as

$$\hat{b}_{ni} = -\int \varphi(t)\hat{a}_{ni}{}'(t)dt \quad i=1,\ldots, n \tag{2.14}$$

where the functions $a_{ni}{}'(t) \equiv da_{ni}(t)/dt$ are piecewise constant on [0,1]. The piecewise linearity of the regression rank scores follows immediately from the linear programming formulation (1.5) of the dual, greatly simplifying the computation in (2.21). In the location model, using (2.13) this reduces to the well-known Hájek and Šidák (1967) scores

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

There are three typical choices of φ:

(i) *Wilcoxon scores*: $\varphi(t) = t - 1/2$, $0 < t < 1$. The scores are $\hat{b}_{ni} = -\int(t - 1/2)d\hat{a}_i(t) = \int\hat{a}_i(t)dt - 1/2$ while $A^2(\varphi) = 1/12$, and $\gamma(\varphi, F) = \int f^2(x)dx$. Wilcoxon scores are optimal when f is the logistic distribution.

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

(iii) *Median (sign) scores*: $\varphi(t) = ½\text{sign}(t–½)$, $0 < t < 1$, then (2.7) leads to the form $\hat{b}_{ni} = \hat{a}_{ni}(½) - ½$ which is ½ if the ith l_1 residual is positive and $-½$ if it is negative, and between $-½$ and ½ otherwise.

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

$$\mathbf{R}'\beta = \mathbf{r} \in \mathbf{R}^q \tag{2.}$$

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

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

Thus, in view of the equivariance of regression quantiles, see Koenker and Bassett(1978), Theorem 3.2, we may define $\tilde{\mathbf{Y}} = \mathbf{Y} - \mathbf{X}\mathbf{R}(\mathbf{R}'\mathbf{R})^{-1}\mathbf{r}$, $\tilde{\mathbf{X}}_1 = \mathbf{X}\mathbf{V}(\mathbf{V}'\mathbf{V})^{-1}$, $\tilde{\mathbf{X}}_2 = \mathbf{X}\mathbf{R}(\mathbf{R}'\mathbf{R})^{-1}$, and proceed as previously discussed with $(\tilde{\mathbf{Y}}, \tilde{\mathbf{X}}_1, \tilde{\mathbf{X}}_2)$ playing the roles of $(\mathbf{Y}, \mathbf{X}_1, \mathbf{X}_2)$. 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.

To illustrate the tests proposed above we consider briefly an example taken from Adichie (1984, Example 3) dealing with the combustion of tobacco. 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 (Wilcoxon) ranks

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

which yield a test statistic of 13.59 highly significant relative to the 1% χ^2_1 critical value of 6.63.

The full set of regression rank scores $\hat{a}_i(t)$ for the restricted model ecluding potassium for this data are illustrated in Figure 6.1. There are 34 distinct regression quantile solutions and therefore each $\hat{a}_{ni}(t)$ is a piecewise linear function with at most 34 distinct segments. Recall that $\hat{a}_{ni}(t) = 1$ while the observed y_i is above the tth regression quantile plane, 0 while below, and takes some intermediate value when y_i falls on the tth plane. The plots ordered according to their Wilcoxon rank score, which may be computed as $\hat{b}_i = -\int_0^1 (t - 1/2)\, d\hat{a}_i(t) = \int_0^1 \hat{a}_i(t)\,dt - 1/2$. While the Wilcoxon rank scores provide an unambiguous ranking of the observations, since the regression rank score functions typically cross in regression applications, in contrast to the location model, this ranking depends upon the score function employed. The regression rank score plots give some further visual evidence concerning the ranking of the sample observations. Note that if $\hat{a}_{ni}(t) \geq \hat{a}_{nj}(t)$ for all t, then $\hat{b}_{ni} \geq \hat{b}_{nj}$ for any montone score function φ. Numerical calculations give Wilcoxon ranks

-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

and yield a test statistic of 13.17. In view of Theorem 5.1 the approximate p-value is .0003. The two vectors of Wilcoxon ranks correspond closely. Observation 6 is smallest in both rankings and observations 14 and 9 are largest in both. The simple correlation between the two rankings is .978. Note that as a practical matter when $\bar{\varphi} = \int_0^1 \varphi(t)dt = 0$, we may omit the $\hat{\mathbf{X}}_2$ term in the computation of $\mathbf{S}_n$ in (5.3) since $\hat{\mathbf{b}}_n$ is orthogonal to $\mathbf{X}_1$. This is in contrast with the aligned rank situation where the use of $\mathbf{X}_2 - \hat{\mathbf{X}}_2$ is essential.

Corresponding calculations for the normal scores using

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

where ϕ denotes the standard normal density, and t_i is the ith regression quantile breakpoint yields

-0.74	0.15	-1.41	0.23	-0.91	-2.13	-0.45	-1.17	2.08	-0.15
0.63	0.10	-1.25	1.44	0.78	1.15	0.50	1.35	0.40	-0.72
1.41	-0.40	-0.61	-0.03	0.94	0.30	0.39	-1.45	-0.26	-0.18

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

-1.00	1.00	-1.00	1.00	-1.00	-1.00	-1.00	-1.00	1.00	-1.00
1.00	1.00	-1.00	1.00	1.00	1.00	1.00	1.00	0.16	-1.00
1.00	-1.00	-1.00	-0.37	1.00	1.00	1.00	-1.00	-1.00	-0.79

and a test statistic of 8.42 while the aligned rank sign scores yield 10.20. Note that we have multiplied the sign scores by 2 to conform to conventional useage. 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 one of the y observations 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. Adding 10 to the first response, for example, alters the aligned Wilcoxon test statistic from 13.58 to 5.7, which is no longer significant at 1%. and the vector of ranks based on the perturbed data has a correlation of only .48 with the aligned ranks based on the original data. The same perturbation of y_1 changes the Wilcoxon regression rankscore test statistic from 13.17 to 14.70 with a correlation between the two rank vectors of .87. A more robust initial estimator would improve the performance of the aligned rank test somewhat. 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 $\mathbf{X}_2$ 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 and therefore the corresponding regression rank scores 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 (1987, 1990) to compute regression quantiles.

3. Properties of regression rank scores

Consider the linear regression model (2.1) with design $\mathbf{X}_n$ of dimension $n \times p$. Let $\hat{\beta}(\alpha) \in \mathbf{R}^p$ be the α–regression quantile and $\hat{\mathbf{a}}(\alpha) \in R^n$ be the vector of αth regression rank scores defined in (2.7). We see from the form of the linear constraints in (1.5) that the regression rank scores are *regression invariant*, i.e.,

$$\hat{\mathbf{a}}_n(\alpha, \mathbf{Y}+\mathbf{Xb}) = \hat{\mathbf{a}}_n(\alpha, \mathbf{Y}), \qquad \mathbf{b} \in \mathbf{R}^p. \tag{3.1}$$

Moreover, in view of the invariance, we may assume

$$\sum_{i=1}^{n} x_{ij} = 0, \quad j=2, \ldots, p \tag{3.2}$$

without loss of generality.

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

$$\{\hat{\mathbf{a}}_n(t): \; 0 \le t \le 1\}. \tag{3.4}$$

Gutenbrunner and Jurečková(1992) studied the process

$$\mathbf{W}_n^d = \{\mathbf{W}_n^d(t) = \sqrt{n}\, \sum_{i=1}^{n} \mathbf{d}_{ni} \hat{a}_{ni}(t): \; 0 \le t \le 1\} \tag{3.5}$$

and showed that $W_n^d(t) = U_n^d(t) + o_p\,(1)$ where

$$\mathbf{U}_n^d(t) = n^{-1/2} \sum_{i=1}^{n} \mathbf{d}_{ni} I[E_i > F^{-1}(t)] \tag{3.6}$$

as $n \to \infty$ uniformly on any fixed interval $[\varepsilon, 1-\varepsilon]$, where $0 < \varepsilon < 1/2$ for any appropriately standardized triangular array $\{\mathbf{d}_{ni} : i=1,\ldots, n\}$ of vectors from $\mathbf{R}^{q.}$ They also showed that the process (3.4) (and hence (3.5)) has continuous trajectories and, under the standardization $\sum_{i=1}^{n} \mathbf{d}_{ni} = 0$, (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 $[\varepsilon, 1-\varepsilon]$. 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 the location submodel. Here we extend the results of Gutenbrunner and Jurečková (1992) into the tails of [0,1], in order to find the asymptotic behavior of the rank scores and the test statistics (2.7) and (2.8), for which the score functions are not constant in the tails.

It may be noted that this extension is rather delicate. If the rank scores involved integration from ε to $1-\varepsilon$ (i.e., if φ were constant near 0 and 1), then the earlier Gutenbrunner-Jurečková (1992) representation theorem could be used to obtain the asymptotic distribution theory here under somewhat weaker hypotheses (see the remark following Theorem 5.1). It is the desirability of treating such tests as the Wilcoxon and Normal Scores Tests that requires the extensions here. Nonetheless, the fact shown here that the rank score process can be represented uniformly on an interval $(\alpha_n^*,\ 1-\alpha_n^*)$ with α_n^* decreasing as a negative power of n (precisely, $\alpha_n^* = n^{-1/(1+4b)}$ for some $b>0$) is rather remarkable and of independent theoretical interest.

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

$$-\infty \le A \equiv \sup\ \{x\colon F(x) = 0\} \quad \text{and} \quad +\infty \ge B \equiv \inf\ \{x\colon F(x) = 1\}\ .$$

For $0 < \alpha < 1$, let ψ_α denote the score function corresponding to (1.2):

$$\psi_\alpha(x) = \alpha - I[x < 0], \qquad x \in \mathbf{R}^1. \tag{3.7}$$

We shall impose the following conditions on F:

(F.1) $|F^{-1}(\alpha)| \le c(\alpha(1-\alpha))^{-a}$ for $0 < \alpha \le \alpha_0$, $1-\alpha_0 \le \alpha < 1$, where $0 < a \le 1/4 - \varepsilon$, $\varepsilon > 0$ and $c > 0$.

(F.2) $1/f(F^{-1}(\alpha)) \le c(\alpha(1-\alpha))^{-1-a}$ for $0 < \alpha \le \alpha_0$ and $1-\alpha_0 \le \alpha < 1$, $c > 0$.

(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) $\left| \dfrac{f'(x)}{f(x)} \right| \le c\,|x|$ for $|x| \ge K \ge 0$, $c > 0$.

REMARK. These conditions are satisfied, for example, by the normal, logistic, double exponential and t distributions with 5, or more, degrees of freedom. Condition (F.1) implies $\int |t|^{4+\delta} dF(t) < +\infty$ for some $\delta > 0$. Hence using (F.4) also, F has finite Fisher Information, a fact to be applied in Theorem 5.1.

The following design assumptions will also be employed.

(X.1) $x_{i1} = 1$, $i=1, \ldots, n$

(X.2) $\lim_{n\to\infty} \mathbf{D}_n = \mathbf{D}$ where $\mathbf{D}_n = n^{-1}\mathbf{X}_n'\mathbf{X}_n$ and $\mathbf{D}$ is a positive definite $p \times p$ matrix.

(X.3) $n^{-1} \sum_{i=1}^{n} \|x_i\|^4 = O(1)$ as $n \to \infty$.

(X.4) $\max_{1\le i\le n} \|x_i\| = O(n^{(2(b-a)-\delta)/(1+4b)})$ for some $b>0$ and $\delta>0$ such that $0 < b-a < \varepsilon/2$ (hence $0 < b < 1/4 - \varepsilon/2$).

We may now define

$$\alpha_n^* = n^{-1/(1+4b)} \quad \text{and} \quad \sigma_\alpha = \frac{(\alpha(1-\alpha))^{1/2}}{f(F^{-1}(\alpha))},\ \ 0 < \alpha < 1\ . \tag{3.8}$$

Let C be a fixed constant and define

$$C_n = C\ (\log_2 n)^{1/2}\ . \tag{3.9}$$

We now prove the following crucial lemma:

LEMMA 3.1 Assume that F satisfies (F.1) - (F.4) and that $\mathbf{X}_n$ satisfies (X.1) - (X.3).

Then, as $n \to \infty$,

$$\sup\{\,|r_n(\mathbf{t},\,\alpha)|\,:\|\mathbf{t}\|\le C_n,\ \alpha_n^*\le\alpha\le 1-\alpha_n^*\}\overset{P}{\to} 0 \tag{3.10}$$

for C_n given by (3.9), where

$$r_n(\mathbf{t},\,\alpha)=(\alpha(1-\alpha))^{-1/2}\sigma_\alpha^{-1}\sum_{i=1}^{n}[\rho_\alpha(E_{i\alpha}-n^{-1/2}\sigma_\alpha\mathbf{x}_i{}'\mathbf{t})-\rho_\alpha(E_{i\alpha})]$$

$$+\,n^{-1/2}(\alpha(1-\alpha))^{-1/2}\sum_{i=1}^{n}\mathbf{x}_i{}'\mathbf{t}\psi_\alpha(E_{i\alpha})-\tfrac{1}{2}\mathbf{t}'\mathbf{D}_n\mathbf{t} \tag{3.11}$$

and

$$E_{i\alpha}=E_i-F^{-1}(\alpha),\ \ i=1,\,...,\,n. \tag{3.12}$$

PROOF.
(i) First fix $\alpha\in[\alpha_n^*,\,1-\alpha_n^*]$ and $\mathbf{t}$ such that $\|\mathbf{t}\|\le C_n$.
Define for some $0<\gamma<b$,

$$B_n=\max\left(n^{-2a/(1+4b)},\,n^{-(2-\gamma)(b-a)/(1+4b)},\,n^{-(b-\gamma)/(1+4b)}\right)\ . \tag{3.13}$$

We wish to show that for any $\lambda>0$

$$P(\,|r_n(\mathbf{t},\,\alpha)|\ge(\lambda+1)B_n)\le Kn^{-\lambda} \tag{3.14}$$

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

$$P(\,|r_n(\mathbf{t},\,\alpha)|\ge s_n)\le\exp(-us_n)(M(u)+M(-u)),\quad u>0 \tag{3.15}$$

where $M(u)=E\exp(ur_n(\mathbf{t},\,\alpha))$.

Denote

$$\varepsilon_{ni}=\varepsilon_{ni}(\mathbf{t},\,\alpha)=n^{-1/2}\sigma_\alpha\mathbf{x}_i{}'\mathbf{t} \tag{3.16}$$

and

$$R_i(\mathbf{t},\,\alpha)=(\alpha(1-\alpha))^{-1/2}\sigma_\alpha^{-1}[\rho_\alpha(E_{i\alpha}-n^{-1/2}\sigma_\alpha\mathbf{x}_i{}'\mathbf{t})-\rho_\alpha(E_{i\alpha})] \tag{3.17}$$

$$+\,n^{-1/2}(\alpha(1-\alpha))^{-1/2}\mathbf{x}_i{}'\mathbf{t}\psi_\alpha(E_{i\alpha})-\tfrac{1}{2}n^{-1}(\mathbf{x}_i{}'\mathbf{t})^2\quad i=1,\,.\,.\,.\,,n.$$

By definition of $E_{i\alpha}$, σ_α, ρ_α and ψ_α,

$$R_i(\mathbf{t},\,\alpha)+\tfrac{1}{2}n^{-1}(\mathbf{x}_i{}'\mathbf{t})^2=(\alpha(1-\alpha))^{-1/2}\sigma_\alpha^{-1}\{(E_{i\alpha}-\varepsilon_{ni})I[\varepsilon_{ni}<E_{i\alpha}<0]$$

$$+\,(\varepsilon_{ni}-E_{i\alpha})I[0<E_{i\alpha}<\varepsilon_{ni}]\} \tag{3.18}$$

and hence, uniformly for $\alpha_n^*\le\alpha\le 1-\alpha_n^*$, $\|\mathbf{t}\|\le C_n$ and $i=1,...,n$,

$$|R_i(\mathbf{t},\,\alpha)+\tfrac{1}{2}n^{-1}(\mathbf{x}_i{}'\mathbf{t})^2|\le 2n^{-1/2}(\alpha(1-\alpha))^{-1/2}|\mathbf{x}_i{}'\mathbf{t}|=O\big(n^{-(2a+\delta)/(1+4b)}\,(\log_2 n)^{1/2}\big). \tag{3.19}$$

If uR_i is bounded, that is, $0<u<n^{(2a+\delta)/(1+4b)}(\log_2 n)^{-1/2}$, Taylor series expansion yields

$$\log M_{R_i}(u)\le uER_i(\mathbf{t},\,\alpha)+cu^2\mathrm{Var}(R_i(\mathbf{t},\,\alpha)) \tag{3.20}$$

for some constant $c>0$. By (3.18), for $\varepsilon_{ni}>0$ and for $\alpha_n^*\le\alpha\le\alpha_0$, $1-\alpha_n^*\ge\alpha\ge 1-\alpha_0$,

$$ER_i(\mathbf{t},\ \alpha) = -\tfrac{1}{2}n^{-1}(\mathbf{x}_i{}'\mathbf{t})^2 + (\alpha(1-\alpha))^{-1/2}\sigma_\alpha^{-1}\int_0^{\varepsilon_{ni}}(\varepsilon_{ni}-z)f(z+F^{-1}(\alpha))dz\ .$$

Now,

$$\log\ f(z+F^{-1}(\alpha)) = \log f(F^{-1}(\alpha)) + \int_0^z \frac{d}{du}\log\ f(u+F^{-1}(\alpha))\ du\ \ ;$$

or, by condition F.4,

$$f(z+F^{-1}(\alpha)) \le f(F^{-1}(\alpha))\ \exp\Big\{\int_0^z \big(u + |F^{-1}(\alpha)|\big)\ du\Big\}\ \ .$$

Also by (3.8), 3.16) and the conditions,

$$\varepsilon_{ni}\ |F^{-1}(\alpha)| = O\Big(n^{-1/2}\ \alpha_n^{-1/2-2a}\ n^{\frac{2(b-a)-\delta}{1+4b}}\ (\log_2 n)^{1/2}\Big) \to 0\ \ .$$

Hence,

$$\exp\Big\{\int_0^z\big(u + |F^{-1}(\alpha)|\big)\ du\Big\} \le 1 + c\int_0^z\big(|F^{-1}(\alpha)| + u\big)\ du\ .$$

Therefore,

$$ER_i(\mathbf{t},\ \alpha) = -\tfrac{1}{2}n^{-1}(\mathbf{x}_i{}'\mathbf{t})^2 + (\alpha(1-\alpha))^{-1/2}\sigma_\alpha^{-1} f(F^{-1}(\alpha))\ \Big\{\int_0^{\varepsilon_{ni}}(\varepsilon_{ni}-z)dz$$

$$+\ O(1)\int_0^{\varepsilon_{ni}}(\varepsilon_{ni}-z)\int_0^z\big(|F^{-1}(\alpha)| + u\big)\ du\ dz\Big\}\ \ . \tag{3.21}$$

By (3.8) and (3.16), the first integral in (3.21) exactly cancels $-\tfrac{1}{2}n^{-1}(\mathbf{x}_i{}'\mathbf{t})^2$; and, therefore, using conditions F.1 - F.4,

$$ER_i(\mathbf{t},\ \alpha) \le c(\alpha(1-\alpha))^{-1/2-2a}n^{-3/2}|\mathbf{x}_i{}'\mathbf{t}|^3 + c(\alpha(1-\alpha))^{-1-2a}n^{-2}|\mathbf{x}_i{}'\mathbf{t}|^4\ \ . \tag{3.21}$$

We get the same inequality for $\varepsilon_{ni} < 0$. The same expressions are $O(n^{-3/2}|\mathbf{x}_i{}'\mathbf{t}|^3)$ + $O(n^{-2}|\mathbf{x}_i{}'\mathbf{t}|^4)$ if $\alpha_0 \le \alpha \le 1-\alpha_0$. Hence,

$$\sum_{i=1}^n E\,|R_i(\mathbf{t},\ \alpha)| = O\left(n^{\frac{-2(b-a)}{1+4b}}\right). \tag{3.22}$$

Similarly, using (3.18) and (3.21),

$$\mathrm{Var}R_i(\mathbf{t},\ \alpha) \le \left(\frac{f(F^{-1}(\alpha))}{\alpha(1-\alpha)}\right)^2\Big\{\ f(F^{-1}(\alpha))\int_0^{|\varepsilon_{ni}|}(|\varepsilon_{ni}|-z)^2\Big[1 + \int_0^z(|F^{-1}(\alpha)|+y)\ dy\ \Big]\ dz$$

$$+\ f^2(F^{-1}(\alpha))\left\{\int_0^{|\varepsilon_{ni}|}(|\varepsilon_{ni}|-z)\Big[1 + \int_0^z(|F^{-1}(\alpha)|+y)\ dy\ \Big]\ dz\right\}^2\ \Big\}$$

Therefore, using (3.8),

$$\sum_{i=1}^{n} \mathrm{Var} R_i(\mathbf{t},\ \alpha) \le c\ n^{-3/2}\ (\alpha(1-\alpha))^{-1/2} \sum_{i=1}^{n} |\mathbf{x}_i{}'\mathbf{t}|^3 \ =\ O\left(n^{\frac{-2b}{1+4b}} \right)\ . \tag{3.23}$$

These results hold uniformly in α and **t**.
Hence, using (3.15) and (3.13) with $u = \log n/B_n = O(n^{2a/(1+4b)})$, so that 3.20 holds,

$$P(|r_n(t,\ \alpha)| \ge (\lambda+1)B_n) \le \exp\{-(\lambda+1)\log n$$

$$+(K\log n/B_n)\cdot n^{\frac{-2(b-a)}{1+4b}} + (K\log^2 n/B_n^2)\cdot n^{\frac{-2b}{1+4b}}\}$$

$$\le n^{-\lambda} \tag{3.24}$$

for $n \ge n_0$ where $K > 0$ and n_0 do not depend on α and **t**.

(ii) Now apply the chaining argument to extend (3.24) uniformly in (**t**, α). Following the proof of Lemma A.2 in Koenker and Portnoy (1987), choose intervals of length $1/n^5$ covering $[\alpha_n^*,\ 1-\alpha_n^*]$ and balls of radius $1/n^5$ covering $\{\mathbf{t}:\ \|\mathbf{t}\| \le C_n\}$. Let $\{\alpha_1,\ \alpha_2\}$ lie in one of the intervals and $\{\mathbf{t}_1,\ \mathbf{t}_2\}$ lie in one of the balls covering $\{\mathbf{t}:\|\mathbf{t}\| \le C_n\}$. We now use (3.18) to bound $\Delta_i \equiv |R_i(\mathbf{t}_1,\ \alpha_1) - R_i(\mathbf{t}_2,\ \alpha_2)|$. So define intervals $J_l^{\pm}$ as follows for $l = 1,\ 2$:

$$J_l^+ = [F^{-1}(\alpha_l),\ \varepsilon_{ni}(\mathbf{t}_l,\alpha_l) + F^{-1}(\alpha_l)] \qquad J_l^- = [\varepsilon_{ni}(\mathbf{t}_l,\alpha_l) + F^{-1}(\alpha_l),\ F^{-1}(\alpha_l)]\ .$$

Also define (for $l = 1,\ 2$):

$$G_l(E_i) \equiv \frac{f(F^{-1}(\alpha_l))}{\alpha_l(1-\alpha_l)}\left(E_i - F^{-1}(\alpha_l) - \varepsilon_{ni}(\mathbf{t}_l,\alpha_l)\right)$$

Then, from (3.18),

$$\Delta_i = -\frac{1}{2n}(\mathbf{x}_i{}'\mathbf{t}_1)^2 + \frac{1}{2n}(\mathbf{x}_i{}'\mathbf{t}_2)^2\ + H(E_i), \tag{3.25}$$

where

$$H(E_i) = \begin{cases} G_1(E_i) - G_2(E_i) & E_i \in J_1^+ \cap J_2^+ \\ G_2(E_i) - G_1(E_i) & E_i \in J_1^- \cap J_2^- \\ \Delta_i^* & E_i \in (J_1^+ \cap J_2^-) \cup (J_1^- \cap J_2^+) \\ 0 & \text{otherwise} \end{cases}$$

and

$$\Delta_i^* \le \max_{l=1,\ 2}\left(\frac{f(F^{-1}(\alpha_l))}{\alpha_l(1-\alpha_l)}\ |\ \varepsilon_{ni}(\mathbf{t}_l,\alpha_l)| \right) = \max_{l=1,\ 2}\left(\frac{n^{-1/2}\,|\mathbf{x}_i\mathbf{t}_l|}{(\alpha_l(1-\alpha_l))^{1/2}} \right)$$

$$\le \frac{C_n\, n^{-1/2}\, \|\mathbf{x}_i\|}{(\alpha_l(1-\alpha_l))^{1/2}} \le C_n\, n^{-1/2}\, n^{\frac{2(b-a)}{1+4b}}\, n^{\frac{1}{2(1+4b)}} = C_n\, n^{-\left(\frac{2a}{1+4b}\right)} \to 0\,.$$

Now note that for $\|\mathbf{t}\| \le C_n$ and $\alpha_n^* \le \alpha_l \le 1-\alpha_n^*$,

$$|\, \varepsilon_{ni}(\mathbf{t}_1,\alpha_1) - \varepsilon_{ni}(\mathbf{t}_2,\alpha_2)\,| \le \frac{C_n\, n^{-1/2}\, \|\mathbf{x}_i\|}{\min_\alpha f(F^{-1}(\alpha))} \left|\, (\alpha_1(1-\alpha_1))^{1/2} - (\alpha_2(1-\alpha_2))^{1/2}\, \right|$$

$$+\ C_n\, n^{-1/2}\, \|\mathbf{x}_i\| \left|\, \frac{1}{f(F^{-1}(\alpha_1))} - \frac{1}{f(F^{-1}(\alpha_2))}\, \right| + \frac{c\, n^{-1/2}\, \|\mathbf{x}_i\|}{\min_\alpha f(F^{-1}(\alpha))} \|\mathbf{t}_1 - \mathbf{t}_2\|\,.$$

By conditions F.1 and F.2, for $\alpha \in [\alpha_n^*,\ 1-\alpha_n^*]$, $1/\alpha(1-\alpha) \le n$,

$$\frac{1}{\min_\alpha f(F^{-1}(\alpha))} \le \frac{c}{(\alpha(1-\alpha))^{1+a}} \le n^{5/4}\,,$$

and

$$|\, F^{-1}(\alpha_1) - F^{-1}(\alpha_2)\, | \le \frac{|\, \alpha_1 - \alpha_2\, |}{\min_\alpha f(F^{-1}(\alpha))} \le n^{-3.75}\,.$$

Therefore, using X.4 and the fact that $|\, \alpha_1 - \alpha_2\, | \le n^{-5}$ and $\|\mathbf{t}_1 - \mathbf{t}_2\| \le n^{-5}$, it is straightforward to show that the contributions to (3.25) excluding Δ_i^* are all $o(1)$. Since, $H_l(E_i) = \Delta_i^*$ only if E_i is between $F^{-1}(\alpha_1)$ and $F^{-1}(\alpha_2)$, (for otherwise, the intersections of intervals defining Δ_i^* must be empty), then

$$\sup_{S_\nu} \left|\, \sum_{i=1}^n R_i(\mathbf{t}_1,\ \alpha_1) - \sum_{i=1}^n R_i(\mathbf{t}_2,\ \alpha_2)\, \right| \le o(1) + K o(1)$$

where S_ν denotes the covering set containing $(\alpha_1,\ \mathbf{t}_1)$ and $(\alpha_2,\ \mathbf{t}_2)$, and K is the number of times E_i lies between $F^{-1}(\alpha_1)$ and $F^{-1}(\alpha_2)$. Now, $K \sim \text{binomial}(n,\ p)$ where p is the probability that E_i lies between $F^{-1}(\alpha_1)$ and $F^{-1}(\alpha_2)$ with $|\, \alpha_1 - \alpha_2\, | \le n^{-5}$. Thus, since f is bounded, $p \le c^* n^{-3.75} \le n^{-2}$. Therefore,

$$P\left\{ \sup_{S_\nu} |\, r_n(\mathbf{t}_1,\ \alpha_1) - r_n(\mathbf{t}_2,\ \alpha_2)\, | \ge o(1) + \lambda\, o(1) \right\} \le \sum_{k=\lambda}^n \binom{n}{k} \left(\frac{c^*}{n^2}\right)^k \left(1 - \frac{c^*}{n^2}\right)^{n-k} \le c' n^{-\lambda}\,.$$

Since the number of sets needed to cover the set $S \equiv [\alpha_n^*,\ 1-\alpha_n^*] \times \{\mathbf{t} :\ \|\mathbf{t}\| \le C_n\}$ is bounded by $n^{5(p+1)}$ we obtain from (3.24) for $\lambda > 5(p+1)$

$$P\left\{ \sup_{(\alpha,\mathbf{t})\in S} |\, r_n(\mathbf{t},\ \alpha)\, | \ge (\lambda+1)B_n + \mathbf{o}(1) + \lambda\, \mathbf{o}(1) \right\} \le n^{5(p+1)} n^{-\lambda} \to 0 \qquad \square$$

LEMMA 3.2. Assume the conditions of Lemma 3.1 and let $\mathbf{d}_n = (d_{n1},\ \cdots,\ d_{nn})'$ be a sequence of q-vectors satisfying

$$\mathbf{X}_n{}'\mathbf{d}_n = \mathbf{0}, \qquad \frac{1}{n}\sum_{i=1}^n d_{ni}^2 \to \Delta^2,\ \ 0 < \Delta^2 < \infty \tag{D.1}$$

$$n^{-1} \sum_{i=1}^{n} |d_{ni}|^3 = O(1) \quad \text{as} \quad n \to \infty \tag{D.2}$$

$$\max_{1 \le i \le n} |d_{ni}| = O\left(n^{(2(b-a)-\delta)/(1+4b)} \right). \tag{D.3}$$

Then, with $S^* = \{(\mathbf{t}, \alpha) : \ \|\mathbf{t}\| \le C_n, \ \alpha_n^* \le \alpha \le 1 - \alpha_n^* \}$,

$$\sup_{(\alpha, \mathbf{t}) \in S^*} \{ (\alpha(1-\alpha))^{-1/2} n^{-1/2} \, \Big| \sum_{i=1}^{n} d_{ni} [\psi_\alpha(E_{i\alpha} - n^{-1/2} \sigma_\alpha \mathbf{x}_i{}' \mathbf{t}) - \psi_\alpha(E_{i\alpha})] \Big| \ \} \xrightarrow{P} 0 \tag{3.26}$$

as $n \to \infty$ for α_n^* given in (3.8), and for C_n given in (3.9).

PROOF. Consider the model

$$\mathbf{Y} = \mathbf{X}^* \beta^* + \mathbf{E}$$

where $\mathbf{X}^* = (\mathbf{X}_n \ \ | \ \ \mathbf{d}_n)$, $\quad \beta^* = (\beta_1, \cdots, \beta_p, \beta_{p+1}, \cdots, \beta_{p+q})$. Then

$$\mathbf{X}^{*\prime} \mathbf{X}^* = \begin{pmatrix} \mathbf{X}_n{}' \mathbf{X}_n & \mathbf{0} \\ \mathbf{0} & \mathbf{d}_n{}' \mathbf{d}_n \end{pmatrix}$$

and the conditions of Lemma 3.1 are satisfied even when replacing $\mathbf{X}$ by $\mathbf{X}^*$ and taking $\mathbf{t} \in \mathbf{R}^{p+q}$. Now, the quantity in brackets in (3.26) is just the right derivative of (3.11) with respect to the last q coordinates of $\mathbf{t}$ (evaluated when the last q coordinates of $\mathbf{t}$ are zero). To obtain the desired uniform convergence, let $f_n(tt, \alpha)$ denote the right hand side of (3.11) without the last term, $\frac{1}{2}\mathbf{t}' D_n \mathbf{t}$, and let $g(t) \equiv \frac{1}{2}\mathbf{t}' D \mathbf{t}$. Note that $\frac{1}{2}\mathbf{t}' D_n \mathbf{t}$ can be replaced by $g(\mathbf{t})$ since $D_n \to D$. By Lemma 3.1, choose δ_n so that

$$\sup_{(\alpha, \mathbf{t}) \in S^*} | f_n(\mathbf{t}, \alpha) - g(\mathbf{t}) | \le \delta_n^2 \ .$$

Following Rockafellar (1970, Thm. 25.7, p. 248), the convexity of f_n makes the difference quotients monotonic. That is, with $\mathbf{u}$ a properly chosen coordinate vector,

$$\frac{\partial}{\partial \mathbf{t}_j} f_n(\mathbf{t}, \alpha) \le \frac{1}{\delta_n} \big(f_n(\mathbf{t} + \delta_n \mathbf{u}, \alpha) - f_n(\mathbf{t}, \alpha) \big)$$

$$\le \frac{1}{\delta_n} \big(g(\mathbf{t} + \delta_n \mathbf{u}) - g(\mathbf{t}) \big) \ + \ \frac{2}{\delta_n} \ \mathbf{O}(\delta_n^2) \ .$$

Replacing $\mathbf{u}$ by $-\mathbf{u}$, the reverse inequality follows similarly (with minus signs on the right side). Therefore,

$$\Big| \frac{\partial}{\partial \mathbf{t}_j} (f_n(\mathbf{t}, \alpha) - g(\mathbf{t}) \Big| \ \le \left| \frac{g(\mathbf{t} + \delta_n \mathbf{u}) - g(\mathbf{t})}{\delta_n} \ - \ \frac{\partial}{\partial \mathbf{t}_j} g(\mathbf{t}) \right| + \mathbf{O}(\delta_n) \ .$$

Since3 g is a quadratic function, this last term tends to zero as a constant times δ_n (uniformly on S^*). This gives (3.26), since the contribution of the final term of (3.11) vanishes when differentiating with respect to the last q coordinates and setting them to zero. □

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

$$\sum_{i=1}^{n} \rho_\alpha(Y_i - \mathbf{x}_i{}'\mathbf{t}) := \min, \quad t \in R^p. \tag{3.27}$$

Analogously, define $\beta(\alpha) = \left(F^{-1}(\alpha), 0, 0, \ . \ . \ . \ , 0\right)$; that is, the solution to (3.27) when the summation is replaced by expectation. 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),

$$n^{1/2}\sigma_\alpha^{-1}(\hat{\beta}(\alpha) - \beta(\alpha)) = n^{-1/2}(\alpha(1-\alpha))^{-1/2}\mathbf{D}_n^{-1} \sum_{i=1}^{n} \mathbf{x}_i \psi_\alpha(E_{i\alpha}) + \mathbf{o}_p(1) \tag{3.28}$$

uniformly in $\alpha_n^* \le \alpha \le 1-\alpha_n^*$. Consequently,

$$\sup_{\alpha_n^* \le \alpha \le 1-\alpha_n^*} \| n^{1/2}\sigma_\alpha^{-1}(\hat{\beta}_n(\alpha) - \beta(\alpha)) \| = O_p\left((\log_2 n)^{1/2}\right). \tag{3.29}$$

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

$$\mathbf{T}_{n\alpha} = n^{1/2}\sigma_\alpha^{-1}(\hat{\beta}_n(\alpha) - \beta(\alpha)) \tag{3.30}$$

minimizes the convex function

$$G_{n\alpha}(\mathbf{t}) = (\alpha(1-\alpha))^{-1/2}\sigma_\alpha^{-1} \sum_{i=1}^{n} [\rho_\alpha(E_{i\alpha} - n^{-1/2}\sigma_\alpha \mathbf{x}_i{}'\mathbf{t}) - \rho_\alpha(E_{i\alpha})] \tag{3.31}$$

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

$$\min_{\|\mathbf{t}\| < C_n} G_{n\alpha}(\mathbf{t}) = \min_{\|\mathbf{t}\| < C_n} \{-\mathbf{t}'\mathbf{Z}_{n\alpha} + \tfrac{1}{2}\mathbf{t}'\mathbf{D}_n\mathbf{t}\} + o_p(1) \tag{3.32}$$

uniformly in $\alpha_n^* \le \alpha \le 1-\alpha_n^*$, where

$$\mathbf{Z}_{n\alpha} = n^{-1/2}(\alpha(1-\alpha))^{-1/2} \sum_{i=1}^{n} \mathbf{x}_i \psi_\alpha(E_{i\alpha}). \tag{3.33}$$

It will be necessary to provide a probabilistic bound for $B \equiv \sup\{ \mathbf{Z}_{n\alpha} : \ \alpha_n^* \le \alpha \le 1-\alpha_n^* \}$. Using the fact that

$$\mathbf{Z}_{n\alpha} \sim \frac{n^{-1/2}}{(\alpha(1-\alpha))^{1/2}} \sum_{i=1}^{n} \mathbf{x}_i \left\{(1-\alpha)\left(I\{F(E_i) \le \alpha\} - \alpha\right) + \alpha\left(I\{F(E_i) \le 1-\alpha\} - (1-\alpha)\right)\right\},$$

the invariance theorem of Shorack (1991) can be applied. By conditions X.3, X.4, and the fact that $\alpha_n^* > n^{-1/4}$, equation (1.10) or (1.11) of Shorack (1991) implies that

$$B \le O_p(1) + c \sup\left\{\left(s(1-s)\right)^{-1/2} \mathbf{W}(s) : \ \alpha_n^* \le s \le 1-\alpha_n^* \right\}$$

for some constant c, where $\mathbf{W}(s)$ is a Brownian Bridge. This last supremum is bounded by $c(\log_2 n)^{1/2} + O_p(1)$ (see, for example, Shorack and Wellner (1986), p. 599). Thus $\mathbf{Z}_{n\alpha} = O_p\left((\log_2 n)^{1/2}\right)$ uniformly on $\alpha_n^* \le \alpha \le 1-\alpha_n^*$. Therefore, denoting

$$\mathbf{U}_{n\alpha} = \arg\min_{\mathbf{t} \in R^p}\{-\mathbf{t}'\mathbf{Z}_{n\alpha}+\tfrac{1}{2}\mathbf{t}'\mathbf{D}_n\mathbf{t}\}, \tag{3.34}$$

we immediately get

$$\mathbf{U}_{n\alpha} = \mathbf{D}_n^{-1}\mathbf{Z}_{n\alpha} = O_p\big((\log_2 n)^{1/2}\big) \tag{3.35}$$

uniformly in $\alpha_n^* \le \alpha \le 1-\alpha_n^*$ and

$$\min_{\mathbf{t} \in R^p}\{-\mathbf{t}'\mathbf{Z}_{n\alpha}+\tfrac{1}{2}\mathbf{t}'\mathbf{D}_n\mathbf{t}\} = -\tfrac{1}{2}\mathbf{Z}_{n\alpha}\mathbf{D}_n^{-1}\mathbf{Z}_{n\alpha}. \tag{3.36}$$

From (3.35) and (3.36), we can write

$$-\mathbf{t}'\mathbf{Z}_{n\alpha} + \tfrac{1}{2}\mathbf{t}'\mathbf{D}_n\mathbf{t} = \tfrac{1}{2}\{(\mathbf{t}-\mathbf{U}_{n\alpha})'\mathbf{D}_n(\mathbf{t}-\mathbf{U}_{n\alpha})-\mathbf{U}_{n\alpha}'\mathbf{D}_n\mathbf{U}_{n\alpha}\} \tag{3.37}$$

and hence we could rewrite (3.10) in the form

$$\sup_{(\alpha,\mathbf{t})\in S}|r_n(\mathbf{t},\alpha)| = \sup_{(\alpha,\mathbf{t})\in S}\Big|\{G_{n\alpha}(\mathbf{t})-\tfrac{1}{2}[(\mathbf{t}-\mathbf{U}_{n\alpha})'\mathbf{D}_n(\mathbf{t}-\mathbf{U}_{n\alpha})-\mathbf{U}_{n\alpha}'\mathbf{D}_n\mathbf{U}_{n\alpha}]\}\Big| \xrightarrow{p} 0 \tag{3.38}$$

Inserting $\mathbf{U}_{n\alpha} = O_p\big((\log_2 n)^{1/2}\big)$, for $\mathbf{t}$, we further obtain

$$\sup_{\alpha_n^* \le \alpha \le 1-\alpha_n^*}\{|G_{n\alpha}(\mathbf{U}_{n\alpha}) + \tfrac{1}{2}\mathbf{U}_{n\alpha}'\mathbf{D}_n\mathbf{U}_{n\alpha}|\} = o_p(1). \tag{3.39}$$

We would like to show that

$$\sup_{\alpha_n^* \le \alpha \le 1-\alpha_n^*}\{\|\mathbf{T}_{n\alpha}-\mathbf{U}_{n\alpha}\|\} = o_p(1). \tag{3.40}$$

Consider the ball $\mathbf{B}_{n\alpha}$ with center $\mathbf{U}_{n\alpha}$ and radius $\delta > 0$. Then, for $\mathbf{t} \in \mathbf{B}_{n\alpha}$,

$$\|\mathbf{t}\| \le \|\mathbf{t}-\mathbf{U}_{n\alpha}\| + \|\mathbf{U}_{n\alpha}\| \le \delta + K_1 (\log_2 n)^{1/2}$$

for some K_1 with probability exceeding $1-\varepsilon$ for $n \ge n_0$. Hence, by (3.10),

$$\Delta_{n\alpha} = \sup_{\alpha_n^* \le \alpha \le 1-\alpha_n^*}\ \sup_{\mathbf{t}\in\mathbf{B}_{n\alpha}}|r_n(\mathbf{t},\alpha)| \xrightarrow{P} 0. \tag{3.41}$$

Following Pollard (1991), consider the behavior of $G_{n\alpha}(\mathbf{t})$ outside $\mathbf{B}_{n\alpha}$. Suppose $\mathbf{t}_\alpha = \mathbf{U}_{n\alpha} + k\xi$, $k > \delta$ and $\|\xi\| = 1$. Let $\mathbf{t}_\alpha^*$ be the boundary point of $\mathbf{B}_{n\alpha}$ that lies on the line from $\mathbf{U}_{n\alpha}$ to $\mathbf{t}_\alpha$, i.e., $\mathbf{t}_\alpha^* = \mathbf{U}_{n\alpha} + \delta\xi$. Then $\mathbf{t}_\alpha^* = (1-(\delta/k))\mathbf{U}_{n\alpha} + (\delta/k)\mathbf{t}_\alpha$ and hence, by (3.38) and (3.39),

$$\delta/k\,G_{n\alpha}(\mathbf{t}) + (1-\delta/k)G_{n\alpha}(\mathbf{U}_{n\alpha}) \ge G_{n\alpha}(\mathbf{t}_\alpha^*) \ge \tfrac{1}{2}\delta^2\lambda_0 + G_{n\alpha}(\mathbf{U}_{n\alpha}) - 2\Delta_{n\alpha}$$

where λ_0 is the minimal eigenvalue of $\mathbf{D}$. Hence,

$$\inf_{\|\mathbf{t}-U_{n\alpha}\|\ge\delta} G_{n\alpha}(\mathbf{t}) \ge G_{n\alpha}(\mathbf{U}_{n\alpha}) + (k/\delta)(\tfrac{1}{2}\delta^2\lambda_0 - 2\Delta_{n\alpha}). \tag{3.42}$$

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

$$P\Big(\inf_{\alpha_n^* \le \alpha \le 1-\alpha_n^*}\Big[\inf_{\|\mathbf{t}-\mathbf{U}_{n\alpha}\|\ge\delta} G_{n\alpha}(\mathbf{t}) - G_{n\alpha}(\mathbf{U}_{n\alpha})\Big] > \eta\Big) > 1-\varepsilon \tag{3.43}$$

and hence (since the event in (3.43) implies that $G_{n\alpha}$ must be minimized inside the ball of radius δ) $P\big(\sup_{\alpha_n^* \le \alpha \le 1-\alpha_n^*}\|\mathbf{T}_{n\alpha} - \mathbf{U}_{n\alpha}\| \le \delta\big) \to 1$ for any fixed $\delta > 0$, as $n \to \infty$.

Finally, equation (3.29) follows using the argument after (3.33). □

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

THEOREM 3.2. Let $\mathbf{d}_n$ satisfy (D.1) - (D.3), $\mathbf{X}_n$ satisfy (X.1) - (X.4) and F satisfy (F.1) - (F.4). Then

$$\sup_{\alpha_n^* \le \alpha \le 1-\alpha_n^*} \{ | n^{-1/2} (\alpha(1-\alpha))^{-1/2} \sum_{i=1}^{n} d_{ni} (\hat{a}_{ni}(\alpha) - \tilde{a}_i(\alpha)) | \} \overset{p}{\to} 0 \tag{3.44}$$

as $n \to \infty$, where

$$\tilde{a}_i(\alpha) = I[E_i \ge F^{-1}(\alpha)], \quad i=1,\ldots,n. \tag{3.45}$$

PROOF. Insert $n^{1/2}\sigma_\alpha^{-1}(\hat{\beta}_n(\alpha) - \beta(\alpha))$ for $\mathbf{t}$ in (3.26) and notice (3.29) and the fact that

$$\sup_{\alpha_n^* \le \alpha \le 1-\alpha_n^*} \{ n^{-1/2} (\alpha(1-\alpha))^{-1/2} \sum_{i=1}^{n} d_{ni} I[Y_i = \mathbf{x}_i' \hat{\beta}(\alpha)] \} \overset{p}{\to} 0, \tag{3.46}$$

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 Kolmogorov-Smirnov type tests appears in Jurečková (1991).

THEOREM 3.3. Under the conditions of Theorem 3.2, as $n \to \infty$,

$$\sup_{0 \le \alpha \le 1} \{ | n^{-1/2} \sum_{i=1}^{n} d_{ni} (\hat{a}_{ni}(\alpha) - \tilde{a}_{ni}(\alpha)) | \} \overset{P}{\to} 0 \tag{3.47}$$

Moreover, the process

$$\{ \Delta^{-1} n^{-1/2} \sum_{i=1}^{n} d_{ni} \hat{a}_{ni}(\alpha) : \ 0 \le \alpha \le 1 \} \tag{3.48}$$

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

PROOF. By Theorem 3.2,

$$\sup_{\alpha_n^* \le \alpha \le 1-\alpha_n^*} | n^{-1/2} \sum_{i=1}^{n} d_{ni} (\hat{a}_{ni}(\alpha) - \tilde{a}_{ni}(\alpha)) | \overset{p}{\to} 0. \tag{3.49}$$

Further, using the fact that $\sum_{i=1}^{n} \big(1 - \hat{a}_{ni}(\alpha)\big) = n\alpha$, due to the linear constraints in (1.5),

$$\sup_{0 \le \alpha \le \alpha_n^*} | n^{-1/2} \sum_{i=1}^{n} d_{ni} \hat{a}_{ni}(\alpha) | = \sup_{0 \le \alpha \le \alpha_n^*} | n^{-1/2} \sum_{i=1}^{n} d_{ni} (1 - \hat{a}_{ni}(\alpha)) | \le n^{1/2} \max_{1 \le i \le n} | d_{ni} | \alpha_n^*$$

$$= O\left(n^{1/2 + \frac{2(b-a)-\delta}{1+4b} - \frac{1}{1+4b}} \right) = O(n^{-2\delta}) \tag{3.50}$$

and we obtain an analogous conclusion for $\sup_{1-\alpha_n^* \le \alpha \le 1} | n^{-1/2} \sum_{i=1}^{n} d_{ni} \hat{a}_{ni}(\alpha) |$. On the other hand,

$$\sup_{0\le\alpha\le\alpha_n^*} |n^{-1/2}\sum_{i=1}^{n} d_{ni}\tilde{a}_i(\alpha)| = \sup_{0\le\alpha\le\alpha_n^*} |n^{-1/2}\sum_{i=1}^{n} d_{ni}(I[E_i < F^{-1}(\alpha)] - \alpha)|$$

$$\le \max_{1\le i\le n} |d_{ni}| \cdot O_p(\alpha_n^*(1-\alpha_n^*))^{1/2} = o_p(1) \quad (3.51)$$

and analogously

$$\sup_{1-\alpha_n^*\le\alpha\le 1} |n^{-1/2}\sum_{i=1}^{n} d_{ni}\tilde{a}_i(\alpha)| = o_p(1).$$

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 $\varphi(t): 0 < t < 1$ be a nondecreasing square-integrable score-generating function and let b_{ni}, $i=1,...,n$ be the scores defined by (2.7). Let $\{\mathbf{d}_n\}$ be a sequence of vectors satisfying (D.1) - (D.3) . Following Hájek and Šidák (1967), we shall call the statistics

$$S_n = n^{-1/2}\sum_{i=1}^{n} d_{ni}\hat{b}_{ni} \quad (4.1)$$

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

$$S_n = n^{-1/2}\sum_{i=1}^{n} d_{ni}\varphi(F(E_i)) + o_p(1). \quad (4.2)$$

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

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

$$|\varphi'(t)| \le c(t(1-t))^{-1-\delta^*} \quad (4.3)$$

for some $\delta^* < \delta$ where δ 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

$$\Delta^2\left(\int_0^1 \varphi^2(t)dt - \bar{\varphi}^2\right), \quad \bar{\varphi} = \int_0^1 \varphi(t)dt. \quad (4.4)$$

PROOF. Let us consider S_n defined in (4.1) with the scores (2.7). Integrating by parts (notice that $\hat{a}_{ni}(t) - \tilde{a}_i(t) = 0$ for $t = 0, 1$), we obtain

$$-n^{-1/2}\sum_{i=1}^{n} d_{ni}\int_0^1 \varphi(t)d(\hat{a}_{ni}(t)-\tilde{a}_i(t)) = n^{-1/2}\sum_{i=1}^{n} d_{ni}\int_0^1(\hat{a}_{ni}(t)-\tilde{a}_i(t))d\varphi(t)\ , \qquad (4.5)$$

which we must show is $o_p(1)$. We shall split the domain of integration into the intervals $(0, \alpha_n^*]$, (α_n^*, α_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 $I_3 \xrightarrow{P} 0$ by the dominated convergence theorem. Similarly, for some $\delta^* > ½$,

$$|I_2| \le \int_{\alpha_n^*}^{\alpha_0} |\varphi'(t)|\ |n^{-1/2}\sum_{i=1}^{n} d_{ni}(\hat{a}_{ni}(t)-\tilde{a}_i(t))|dt$$

$$\le c\int_{\alpha_n^*}^{\alpha_0} (t(1-t))^{-1-\delta^*}(t(1-t))^{1/2}\cdot |n^{-1/2}(t(1-t))^{-1/2}\sum_{i=1}^{n} d_{ni}(\hat{a}_{ni}(t)-\tilde{a}_i(t))|dt$$

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

Finally,

$$|I_1| \le n^{-1/2}\max_{1\le i\le n}|d_{ni}|\int_0^{\alpha_n^*}|\varphi'(t)|\sum_{i=1}^{n}|\hat{a}_{ni}(t)-\tilde{a}_{ni}(t)|dt \le I_{11}+I_{12}$$

where

$$I_{11} = n^{-1/2}\max_{1\le i\le n}|d_{ni}|\int_0^{\alpha_0^*}|\varphi'(t)|\sum_{i=1}^{n}(1-\hat{a}_{ni}(t))dt \qquad (4.6)$$

and

$$I_{12} = n^{-1/2}\max_{1\le i\le n}|d_{ni}|\int_0^{\alpha_n^*}|\varphi'(t)|\sum_{i=1}^{n}(1-\tilde{a}_i(t))dt. \qquad (4.7)$$

Then

$$I_{11} \le n^{1/2}\max_{1\le i\le n}|d_{ni}|\int_0^{\alpha_n^*} t^{-\delta^*}dt = O(n^{1/2+\frac{2(b-a)-\delta}{1+4b}-\frac{(1-\delta^*)}{1+4b}}) = O(n^{-2(\delta-\delta^*)}).$$

Finally,

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

Now we may assume that $\varphi(\alpha_n^*) < 0$ for $n \ge n_0$, since otherwise if φ were bounded from below then $I_{12} \xrightarrow{P} 0$. Hence

$$Var(I_{12}) \le n^{-1}\sum_{i=1}^{n} d_{ni}^2 E([2\varphi(F(E_i))]^2 I[F(E_i) < \alpha_n^*]) \le \int_0^{\alpha_n^*}\varphi^2(u)du\cdot O(1) \to 0$$

due to the square-integrability of φ. Treating the integrals I_4, I_5 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 $\mathbf{X} = (\mathbf{X}_1 \vdots \mathbf{X}_2)$ satisfies the conditions (X.1) - (X.4), (2.3) and (2.4). We want to test the hypothesis $H_0 : \beta_2 = \mathbf{0}$ (β_1 unspecified) against the alternative $H_n : \beta_{2n} = n^{-1/2}\beta_0$ ($\beta_0 \in R^q$ fixed).

Let $\hat{\mathbf{a}}_n(\alpha) = (\hat{a}_{n1}(\alpha), \ldots, \hat{a}_{nn}(\alpha))$ denote the regression rank scores corresponding to the submodel

$$\mathbf{Y} = \mathbf{X}_1\beta_1 + \mathbf{E} \quad \text{under } H_0. \tag{5.1}$$

Let $\varphi(t) : (0, 1) \to R^1$ be a nondecreasing and square integrable score-generating function. Define the scores $\hat{b}_{ni}$, $i=1,\ldots,n$ by the relation (2.7), and consider the test statistic

$$T_n = \mathbf{S}_n{}'\mathbf{Q}_n^{-1}\mathbf{S}_n / A^2(\varphi) \tag{5.2}$$

where

$$\mathbf{S}_n = n^{-1/2}(\mathbf{X}_{n2} - \hat{\mathbf{X}}_{n2})'\hat{\mathbf{b}}_n \tag{5.3}$$

and where $\mathbf{Q}_n$ and $A^2(\varphi)$ are defined in (2.4) and (2.10), 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 ω critical value of the χ^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 $\mathbf{X}_1$ satisfies (X.1) - (X.4) and $(\mathbf{X}_1 \vdots \mathbf{X}_2)$ 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 φ satisfying (4.3), and nondecreasing and square-integrable on (0, 1).

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

(ii) Under H_n, T_n is asymptotically noncentral χ^2 with q degrees of freedom and with noncentrality parameter,

$$\eta^2 = \beta_0{}'\mathbf{Q}\beta_0 \cdot \gamma^2(\varphi, F)/A^2(\varphi) \tag{5.4}$$

with

$$\gamma(\varphi, F) = -\int_0^1 \varphi(t)\,df(F^{-1}(t)). \tag{5.5}$$

REMARKS.

(i) If φ is of bounded variation and is constant near 0 and 1, the representation given in Theorem 2 (ii) of Gutenbrunner and Jurečková (1992) could be used to provide the conclusion of Theorem 5.1 under somewhat weaker hypotheses; namely, (X.1), (X.2), $\max_i \|x_i\| = o(n^{1/2})$, F has finite Fisher Information, and $0 < f < \infty$ on $\{0 < F < 1\}$.

(ii) The analogy between the location and regression models concerning the noncentrality parameter $\gamma(\varphi, F)$ may be extended in the following way: instead of defining local alternatives via (2.6), the definition of Behnen (1972) can be generalized to the regression model. That is, with $F_i(t) = F(t - x_{1i}'\beta_1)$ and $G_i = L(Y_i)$, consider

$$H_0 : G_i = F_i \qquad \text{vs.} \qquad H_n : \frac{dG_i}{dH_i} = 1 + x_{2i}'\beta_{2n} h_n(F_i)$$

where

$$\beta_{2n} = n^{-1/2}\beta_0, \quad h_n \underset{L^2}{\to} h \in L^2(0, 1), \quad \text{and} \quad \max_i \| x_{2i} \| \, \| h_n \|_\infty^3 = o(n^{1/2}).$$

In this setting, even without the assumption of finite Fisher Information, (4.2) implies that the conclusion of Theorem 5.1 holds with $\gamma(\varphi, F)$ in (5.4) replaced by the F−independent constant

$$\gamma^*(\varphi, h) = \frac{\int(\varphi(u) - \bar{\varphi})(h(u) - \bar{h})du}{(\int(\varphi(u) - \bar{\varphi})^2 du \int(h(u) - \bar{h})^2 du)^{1/2}},$$

i.e., the correlation of the functions φ and h. Such local alternatives provide insight into the structure of the regions of constant efficiency for regression rank tests.

PROOF.

(i) It follows from Theorem 4.1 that, under H_0, $\mathbf{S}_n$ has the same asymptotic distribution as

$$\tilde{\mathbf{S}}_n = n^{-1/2}(\mathbf{X}_{n2} - \hat{\mathbf{X}}_{n2})'\tilde{\mathbf{b}}_n$$

where $\tilde{\mathbf{b}}_n = (\tilde{b}_{n1},...,\tilde{b}_{nn})'$ and $\tilde{b}_{ni} = \varphi(F(E_i))$, $i=1,...,n$. The asymptotic distribution of $\tilde{\mathbf{S}}_n$ follows from the central limit theorem and coincides with q−dimensional normal distribution with expectation 0 and the covariance matrix $\mathbf{Q} \cdot A^2(\varphi)$.

(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 $\tilde{\mathbf{S}}_n$ under H_n coincide. The proposition then follows from the fact that the asymptotic distribution of $\tilde{\mathbf{S}}_n$ under H_n is normal $N_q(\gamma(\varphi, F)\mathbf{Q}\beta_0, \mathbf{Q}A^2(\varphi))$. □

REFERENCES

Adichie, J.N. (1978). "Rank tests for sub-hypotheses in the general linear regression," *Am. Statist.*, 6, 1012-1026.

Adichie, J.N. (1984). "Rank tests in linear models," in P.R. Krishnaiah and P.K. Sen (eds.), *Handbook of Statistics,*, vol. 4, Elsevier, New York.

Amemiya, T. (1985). *Advanced econometrics*, Harvard University Press, Cambridge.

Antoch, J. and Jurečková, J. (1985). "Trimmed LSE resistant to leverage points," *Comp. Statist. Quarterly*, 4, 329-339.

Aubuchon, J.C. and Hettmansperger, T.P. (1989). "Rank based inference for linear models: Asymmetric errors," *Statistics and Probability Letters*, 8, 97-107.

Behnen, K. (1972). "A characterization of certain rank-order tests with bounds for the asymptotic relative efficiency," *Ann. Math. Statist.*, 43, 1839-1851.

Chernoff, H. and Savage, I.R. (1958). "Asymptotic normality and efficiency of certain non-parametric test statistics," *Ann. Math. Statist.*, 29, 972-994.

Draper, D. (1988). "Rank-based robust analysis of linear models. I. Exposition and Review," *Statistical Science*, 3, 239-271.

Gutenbrunner, C. (1986). Zur Asymptotik von Regressionquantileprozessen und daraus abgeleiten Statistiken, Ph.D. Dissertation, Universität Freiburg.

Gutenbrunner, C. and Jurečková, J. (1992). "Regression rank-scores and regression quantiles," *Ann. Statist., 20*, 305-330.

Hájek, J. and Šidák, Z. (1967). "Theory of Rank Tests," *Academia*, Prague.

deJongh, P.J.,deWet, T. and Welsh, A.H. (1988). "Mallows type bounded-influence-regression trimmed means," *J. Amer. Statist. Assoc.*, 83, 805-810.

Jurečková, J. (1984). "Regression quantiles and trimmed least squares estimator under a general design", *Kybernetika*, 20, 345-357.

Jurečková, J. (1991). "Tests of Kolmogorov-Smirnov type based on regression rank scorens. *Transactions of the 11th Prague Conf. on Information Theory, Statist.Decis.* Functions and Random Processes, (J.A.Visek, ed.), pp.41-49. ACADEMIA, Prague.

Koenker, R. and Bassett, G. (1978). "Regression quantiles," *Econometrica*, 46, 33-50.

Koenker, R. and Bassett, G. (1982). "Tests of linear hypotheses and l_1–estimation," *Econometrica*, 50, 1577-83.

Koenker, R. and d'Orey, V. (1987). "Computing regression quantiles," *Applied Statistics*, 36, 383-393.

Koenker, R. and d'Orey V. (1990). "Remark on algorithm 229," to appear: *Applied Statistics.*

Koenker, R.W. and Portnoy, S. (1987). L-Estimation for the Linear Model., *Journal of the American Statistical Association, 82*, 851-857.p

Koul, H.L. (1970). "A class of ADF tests for subhypotheses in the multiple linear regression," *Ann. Math. Statist.*, 41, 1273-1281.

Lehmann, E.L. (1959). *Testing Statistical Hypotheses*, Wiley, New York.

Lehmann, E.L. (1983). *Theory of Point Estimation*, Wiley, New York.

McKean, J.W. and Hettmansperger, T.P. (1978). "A robust analysis of the general linear model based on one-step $R-$estimates," *Biometrika*, 65, 571-79.

Osborne, M.R. (1992) "An effective method for computing regression quantiles", *IMA Journal of Numerical Analysis*, 12, 151-166.

Pollard, D. (1991). "Asymptotics for least absolute deviation regression estimators," *Econometric Theory*, 7, 186-200.

Portnoy, S. (1991). Asymptotic behavior of the number of regression quantile breakpoints, *J. Sci. Statist. Computing.* 12, 867-883.

Portnoy S. (1991b). "Asymptotic behavior of regression quantiles in non-stationary, dependent cases, *J. Mult. Anal.* 38, 100-113.

Puri, M.L. and Sen, P.K. (1985). *Nonparametric Methods in General Linear Models.* J. Wiley, New York.

Rockafellar, R.T. (1970). *Convex Analysis*, Princeton University Press, Princeton, New Jersey.

Ruppert, D. and Carroll, R.J. (1980). "Trimmed least squares estimation in the linear model," *J. Amer. Statist. Assoc.*, 75, 828-838.

Scheffé, H. (1959). *The Analysis of Variance*, Wiley, New York.

Shorack, G. (1991). Embedding the finite sampling process at a rate, *Ann. Probab., 19*, 826-842.

Shorack, G. and Wellner, J. (1986). *Empirical Processes with Applications to Statistics*, Wiley, New York.