Bernoulli **5**(4), 1999, 659–676

Regression rank-scores tests against heavy-tailed alternatives

JANA JUREČKOVÁ

Department of Probability and Statistics, Charles University, Sokolovská 83, CZ-18675 Prague 8, Czech Republic. E-mail: jurecko@karlin.mff.cuni.cz

Statistical inference in the linear model based on the concept of regression rank scores is invariant to the nuisance regression; hence regression rank-scores tests need no estimation of the nuisance parameters. Such tests, already available in the literature, are manageable, asymptotically distribution-free and have many convenient properties, but they are either censored or applicable only to light-tailed distributions. To extend the universality of regression rank-scores tests, we propose modified tests applicable to heavy-tailed distributions including Cauchy. Depending on the alternative we want to treat by the test, we censor the score generating function but the censoring is asymptotically negligible. The proposed tests, being asymptotically distribution-free, are as efficient as the ordinary rank tests without nuisance parameters, for a broad class of density shapes.

Keywords: linear regression model; regression quantile; regression rank scores; regression rank test

1. Introduction

Consider the linear regression model

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

where $\mathbf{Y}_n = \mathbf{Y} = (Y_1, \ldots, Y_n)'$ is a vector of observations, $\mathbf{X}_n = \mathbf{X}$ is an $(n \times p)$ known design matrix, $\boldsymbol{\beta} \in \mathbb{R}^p$ is an unknown parameter and $\mathbf{E}_n = (E_1, \ldots, E_n)'$ is the vector of i.i.d. errors with (generally unknown) distribution function F. We assume that β_1 is an intercept, that is, the first column of $\mathbf{X}_n$ is $\mathbf{1}_n$.

Regression rank (RR) scores, introduced by Gutenbrunner (1986) and studied by Gutenbrunner and Jurečková (1992), are defined as the optimal solution $\hat{\mathbf{a}}_n(\alpha) = (\hat{a}_{n1}(\alpha), \ldots, \hat{a}_{nn}(\alpha))$ of the parametric linear program

$$\begin{aligned} \mathbf{Y}'\hat{\mathbf{a}}(\alpha) &:= \max \\ \mathbf{X}'(\hat{\mathbf{a}}(\alpha) - (1-\alpha))\mathbf{1}_n &= \mathbf{0} \\ \hat{\mathbf{a}}(\alpha) \in [0, 1]^n, &\qquad 0 < \alpha < 1. \end{aligned} \tag{1.2}$$

They are dual to the *regression quantiles* of Koenker and Bassett (1978) in the sense of linear programming: the regression α-quantile $\hat{\boldsymbol{\beta}}_n(\alpha)$ $(0 < \alpha < 1)$ for the model (1.1) is a solution of the minimization

1350–7265

$$\sum_{i=1}^{n} \rho_\alpha(Y_i - \mathbf{x}_i'\mathbf{t}) := \min \tag{1.3}$$

with respect to $\mathbf{t} \in \mathbb{R}^p$, where $\mathbf{x}_i'$ is the ith row of $\mathbf{X}_n$ and

$$\rho_\alpha(z) = z\Psi_\alpha(z),\ \Psi_\alpha(z) = \alpha - I[z < 0], \qquad z \in \mathbb{R}^1; \tag{1.4}$$

the minimization (1.3) could be further written in the parametric linear programming form

$$\begin{gathered} \alpha\mathbf{1}_n'\mathbf{r}^+ + (1-\alpha)\mathbf{1}_n'\mathbf{r}^- := \min \\ \mathbf{X}\boldsymbol{\beta} + \mathbf{r}^+ - \mathbf{r}^- = \mathbf{Y} \\ (\boldsymbol{\beta}, \mathbf{r}^+, \mathbf{r}^-) \in \mathbb{R}^p \times \mathbb{R}_+^n \times \mathbb{R}_-^n, \qquad 0 < \alpha < 1 \end{gathered} \tag{1.5}$$

and the regression quantile $\hat{\boldsymbol{\beta}}(\alpha)$ coincides with the component $\boldsymbol{\beta}$ of the optimal solution.

In the location submodel with $\mathbf{X} = \mathbf{1}_n$ we have $\hat{a}_{ni}(\alpha) = a_n^*(R_i, \alpha)$, where R_i is the rank of Y_i among $Y_1, \ldots, Y_n$ and

$$a_{ni}^*(R_i, \alpha) = \begin{cases} 0 & \text{if } R_i < n\alpha \\ R_i - n\alpha & \text{if } (R_i - 1)/n \leqslant \alpha \leqslant R_i/n \\ 1 & \text{if } \alpha < (R_i - 1)/n,\ i = 1, \ldots, n. \end{cases} \tag{1.6}$$

Rank scores $a_n^*(R_i, \alpha)$ were first used by Hájek (1965) as a starting point for nonlinear rank tests. Summarizing, we could say that the RR scores and regression quantiles are dual in the linear programming sense, but this duality also extends the duality of ranks and of order statistics from the location to the linear regression model.

Gutenbrunner and Jurečková (1992) proposed some tests based on RR scores generated by truncated score functions. A general class of RR tests, parallel to such classical rank tests as the Wilcoxon, normal scores and median, was constructed in Gutenbrunner *et al.* (1993). The same paper also considers the computational aspects of regression quantiles and RR scores, applies the proposed tests to the tobacco data of Steel and Torrie (1960) and numerically compares the tests with the aligned rank tests of Adichie (1984), along with a sensitivity study of the situation when the nuisance parameters in the aligned tests are estimated by means of least squares. The detailed description of the computational algorithms can be found in Koenker and d'Orey (1987; 1994) and in Osborne (1992). Moreover, RR tests of the Kolmogorov–Smirnov type were proposed by Jurečková (1992) and tests of homoscedasticity in the linear model based both on RR scores and regression quantiles were proposed by Gutenbrunner (1994). Typically, the tests based on RR scores apply to the model

$$\mathbf{Y} = \mathbf{X}\boldsymbol{\beta} + \mathbf{Z}\boldsymbol{\delta} + \mathbf{E} \tag{1.7}$$

with $\mathbf{X}$ of order $n \times p$ and $\mathbf{Z}$ of order $n \times q$, where one wants to test the hypothesis

$$H_0\colon \boldsymbol{\delta} = \mathbf{0},\ \boldsymbol{\beta} \text{ unspecified.} \tag{1.8}$$

The test criteria depend on RR scores calculated under H_0, that is, under the submodel (1.1). By (1.2), $\hat{\mathbf{a}}_n(\alpha)$ is invariant to the regression with the matrix $\mathbf{X}$ in the following sense:

$$\hat{\mathbf{a}}_n(\alpha, \mathbf{Y} + \mathbf{X}\mathbf{b}) = \hat{\mathbf{a}}_n(\alpha, \mathbf{Y}) \qquad \forall\, \mathbf{b} \in \mathbb{R}^p; \tag{1.9}$$

this property is parallel to the invariance of the ordinary ranks to the shift in location. Hence, the tests based on $\hat{\mathbf{a}}_n(\alpha)$ are equally invariant and as such they do not require an estimation of the nuisance $\boldsymbol{\beta}$. The parallel with the rank tests goes even further; choosing a non-decreasing, square-integrable score generating function φ: $(0, 1) \mapsto \mathbb{R}^1$ as the Wilcoxon, median or inverse normal, we compute the *scores* $\hat{b}_{n1} \ldots, \hat{b}_{nn}$ generated by φ in the following way

$$\hat{b}_{ni} = -\int_0^1 \varphi(u)\, \mathrm{d}\hat{a}_{ni}(u), \qquad i = 1, \ldots, n. \tag{1.10}$$

The tests are then based on the *linear regression rank-scores statistics* $S_n = n^{-1/2}\sum_{i=1}^n d_{ni}\hat{b}_{ni}$ with appropriate coefficients d_{ni}. We are able to prove that, under some regularity conditions, the asymptotic behaviour of S_n is analogous to that of the simple linear rank statistics, a basis of rank tests; in fact both statistics admit the same asymptotic representations. This in turn implies that the Pitman efficiencies of the tests based on RR scores coincide with those of the corresponding rank tests being used under $\boldsymbol{\beta} = \mathbf{0}$ (or under known $\boldsymbol{\beta}$). We believe that the regularity conditions leading to our results could still be weakened; simulation studies (see Hallin *et al*. 1997) show that the tests work well even under densities not covered by our conditions.

While the tests proposed by Gutenbrunner and Jurečková (1992), Gutenbrunner *et al*. (1993) and Jurečková (1992) are constructed under a deterministic regression matrix, the situation that either of the matrices $\mathbf{X}$, $\mathbf{Z}$ in (1.7) is random is treated by Picek (1997). Koul and Saleh (1995) extended the RR scores and regression quantiles to the autoregressive time series model, introducing the *autoregression rank scores*. Hallin and Jurečková (1996) considered the tests of the linear hypotheses in the autoregressive models, based on autoregression rank scores, and derived their asymptotic properties under the innovation densities with exponentially decreasing tails. The good performance of the tests is illustrated in Hallin *et al*. (1997) on the simulated AR series with the normal, Laplace and Cauchy innovation densities; the tests are then applied to the dataset of daily maximum temperatures, measured in three stations in South Moravia in the period 1961–90. The tests of independence of two time series based on autoregression rank scores, extensions of the Spearman rank correlation and other tests, are constructed in Hallin *et al*. (1999).

The tests described in Gutenbrunner and Jurečková (1992) and Gutenbrunner (1994) are generated by censored score functions and thus cannot compete with the usual rank tests. On the other hand, the regularity conditions of Gutenbrunner *et al*. (1993) and of Jurečková (1992) exclude distribution shapes with the tails of the t distribution with 4 degrees of freedom and heavier and Hallin and Jurečková (1996) cover only the distributions with exponentially decreasing tails. Noting this situation, we take as a primary goal of the present paper the extension of the universality of the RR tests as close as possible to that of the rank tests. A modified definition of RR criteria, along with refined asymptotics, makes the tests applicable to heavy-tailed distributions including Cauchy. In this respect, the RR tests could compete well with the aligned rank tests studied by Adichie (1984) and by Puri and Sen (1985) or with the tests based on R-estimates studied by Hettmansperger (1984); and this not only theoretically, but also in applications.

We restrict our considerations to the heavy-tailed alternatives with tails heavier than the t

distribution with 5 degrees of freedom. The score function of the test is censored according to the tails which we want to cover by the test. The censoring is asymptotically negligible as $n \to \infty$, and due to this fact the tests are as asymptotically efficient (and distribution-free) as the ordinary rank tests without nuisance parameters. Notice that, in the case of heavy-tailed alternatives, the eventual estimator of the nuisance parameter cannot be the least-squares estimator but a highly robust estimator. If it is sufficient to test (1.8) against the alternatives with lighter tails, we use the uncensored tests described in Gutenbrunner *et al*. (1993).

The asymptotic properties of $\hat{\mathbf{a}}_n(\alpha)$ and of linear RR statistics are derived in Section 3. The proofs of two technical lemmas are postponed to the Appendix. The main tool is the chaining argument combined with probability inequalities for bounded random variables specified to heavy-tailed densities. The tests of H_0 and their asymptotic behaviour are described in Section 4.

2. Asymptotic behaviour of regression rank scores

Consider the linear model (1.1) with i.i.d. errors $E_1, \ldots, E_n$ with a common distribution function F, on which we impose the following regularity conditions:

(F.1) F is absolutely continuous with absolutely continuous, positive and bounded density $f(x)$ and the derivatives f', f'' of f are bounded almost everywhere in $x \in \mathbb{R}^1$.
(F.2) f is monotonically decreasing to 0 as $x \to -\infty$ and $x \to \infty$,

$$\lim_{x\to-\infty} \frac{-a\log F(x)}{\log|x|} = 1, \qquad \lim_{x\to\infty} \frac{-a\log(1-F(x))}{\log x} = 1, \tag{2.1}$$

for some a (the same in each tail), $0 < a < \infty$.

Fix b satisfying $0 < \delta \leqslant b - a \leqslant a + \delta$, and denote

$$\alpha_n^* = n^{-1/(1+2b)} \tag{2.2}$$

and

$$\sigma_\alpha = \frac{(\alpha(1-\alpha))^{1/2}}{f(F^{-1}(\alpha))}, \qquad 0 < \alpha < 1. \tag{2.3}$$

The following regularity conditions will be imposed on the design matrix $\mathbf{X}$:

(X.1) $x_{i1} = 1, \quad i = 1, \ldots, n$.
(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) $n^{-1}\sum_{i=1}^n \|\mathbf{x}_{ni}\|^4 = O(1)$ as $n \to \infty$.
(X.4) $\max_{1\leqslant i\leqslant n}\|\mathbf{x}_{ni}\| = O(n^\Delta)$ as $n \to \infty$, where

$$\Delta = \frac{b-a-\delta}{1+2b}. \tag{2.4}$$

Notice that $\Delta < \frac{1}{4}$; actually,

$$\frac{b-a-\delta}{1+2b}-\frac{1}{4}\leqslant -\frac{1+2\delta}{1+2b}<0.$$

While a could be characterized as the tail-exponent of the distribution (for example, $a=1$ for the Cauchy distribution and $a=1/m$ for the t distribution with m degrees of freedom), censoring at the point (2.2) with $b>a$ allows us to be less restrictive to $\|\mathbf{x}_{ni}\|$s (cf. (X.4)), which is convenient in some situations and models.

The following lemma describes some properties of densities satisfying (F.1) and (F.2), which we shall use below.

Lemma 2.1. *Let the density f satisfy (F.1) and (F.2). Then*

(i) $$\lim_{u\to 0}\frac{f(F^{-1}(u)}{u^{a+1}}=\frac{1}{a},\qquad \lim_{u\to 1}\frac{f(F^{-1}(u)}{(1-u)^{a+1}}=\frac{1}{a}; \tag{2.5}$$

(ii) $$\lim_{x\to-\infty}(-x)^{1/a}F(x)=1,\qquad \lim_{x\to\infty}x^{1/a}(1-F(x))=1; \tag{2.6}$$

(iii) $$\lim_{u\to 0}u\frac{f'(F^{-1}(u))}{f^2(F^{-1}(u))}=a+1,\qquad \lim_{u\to 1}(1-u)\frac{-f'(F^{-1}(u))}{f^2(F^{-1}(u))}=a+1. \tag{2.7}$$

Proof. (i) We obtain, from (F.2),

$$\lim_{u\to 0}\frac{-a\log u}{\log(-F^{-1}(u))}=1, \tag{2.8}$$

hence using l'Hôpital's rule leads to the first part of (i); the second part is analogous.

(ii) By (2.5) and (2.8),

$$\lim_{x\to-\infty}\frac{f(x)}{(F(x))^{a+1}}=\frac{1}{a},\qquad \lim_{x\to-\infty}\left(-ax\frac{f(x)}{F(x)}\right)=1,$$

hence

$$\frac{1}{a}=\lim_{x\to-\infty}\frac{f(x)}{(F(x))^{a+1}}=\lim_{x\to-\infty}\frac{1}{-ax(F(x))^{a}},$$

and this gives the first part of (2.6); the second part is analogous.

(iii) By (2.5) and by l'Hôpital's rule,

$$\lim_{u\to 0}u\frac{f'(F^{-1}(u))}{f^2(F^{-1}(u))}=\frac{a+1}{a}\lim_{u\to 0}\frac{u^{a+1}}{f(F^{-1}(u))}=a+1,$$

and this gives the first part of (2.7). The second part is analogous. □

Let $\hat{\boldsymbol{\beta}}_n(\alpha)$ be the α-regression quantile in the model $\mathbf{Y}=\mathbf{X}\boldsymbol{\beta}+\mathbf{E}$. The first theorem gives an asymptotic representation of $\hat{\boldsymbol{\beta}}_n(\alpha)$, uniform over $\alpha_n^*\leqslant\alpha\leqslant 1-\alpha_n^*$, and the rate of consistency of $\hat{\boldsymbol{\beta}}_n(\alpha)$. It extends Theorem 3.1 of Gutenbrunner *et al*. (1993) to a broad class of distributions covered by conditions (F.1)–(F.2). Generally, the heavier the tails of the distribution we admit, the more restricted is the interval $[\alpha_n^*, 1-\alpha_n^*]$.

Theorem 2.1. *Under (F.1)–(F.2) and (X.1)–(X.4),*

$$\sup_{\alpha_n^* \leqslant \alpha \leqslant 1-\alpha_n^*} \|\sigma_\alpha^{-1}(\hat{\boldsymbol{\beta}}_n(\alpha) - \boldsymbol{\beta}(\alpha)\| = O_p(n^{-1/2}C_n) \tag{2.9}$$

and

$$n^{1/2}\sigma_\alpha^{-1}(\hat{\boldsymbol{\beta}}_n(\alpha) - \boldsymbol{\beta}(\alpha)) = n^{-1/2}(\alpha(1-\alpha))^{-1/2}\mathbf{D}_n^{-1}\sum_{i=1}^{n}\mathbf{x}_{ni}\psi_\alpha(E_{i\alpha}) + o_p(1) \tag{2.10}$$

uniformly in $\alpha_n^* \leqslant \alpha \leqslant 1-\alpha_n^*$, *where* $\boldsymbol{\beta}(\alpha) = (\beta_1 + F^{-1}(\alpha), \beta_2, \ldots, \beta_p)'$, $E_{i\alpha} = E_i - F^{-1}(\alpha)$ *and* $C_n = C(\log\log n)^{1/2}$, $0 < C < \infty$.

The second theorem gives an approximation of the RR scores process by an empirical process, uniform on $[\alpha_n^*, 1-\alpha_n^*]$. Let $\mathbf{d}_n = (d_{n1}, \ldots, d_{nn})'$ be a sequence of vectors satisfying

(D.1) $\Gamma_n^2 = n^{-1}\|\mathbf{d}_n - \hat{\mathbf{d}}_n\|^2 \to \Gamma^2$ as $n \to \infty$, $0 < \Gamma < \infty$ where $\hat{\mathbf{d}}_n = \mathbf{X}_n(\mathbf{X}_n'\mathbf{X}_n)^{-1}\mathbf{X}_n'\mathbf{d}_n$.
(D.2) $n^{-1}\sum_{i=1}^{n}|d_{ni} - \hat{d}_{ni}|^4 = O(1)$ as $n \to \infty$.
(D.3) $\max_{1\leqslant i\leqslant n}|d_{ni} - \hat{d}_{ni}| = O(n^{\Delta})$ with $\Delta(< 1/4)$ of (X.4).

Theorem 2.2. *Under (F.1)–(F.2), (X.1)–(X.4) and (D.1)–(D.3),*

$$\sup_{\alpha_n^* \leqslant \alpha \leqslant 1-\alpha_n^*}\left\{n^{-1/2}\left|\sum_{i=1}^{n}[d_{ni}(\hat{a}_{ni}(\alpha) - (1-\alpha)) - (d_{ni} - \hat{d}_{ni})\tilde{a}_i(\alpha)]\right|\right\} = o_p(C_n(\log n)^{1/4}n^{-\frac{1}{8}+\frac{\Delta}{4}})$$

$$\text{as } n \to \infty, \tag{2.11}$$

where

$$\tilde{a}_i(\alpha) = I[E_i > F^{-1}(\alpha)], \qquad 0 < \alpha < 1, \qquad i = 1, \ldots, n. \tag{2.12}$$

The third theorem extends the uniform approximation of the RR process to the whole segment [0, 1]; the price paid for this extension is a more restrictive condition on the distribution tails – but still weaker than in Gutenbrunner *et al*. (1993). Fortunately, the tests constructed in Section 4 do not need this stronger result and thus their asymptotics hold under the conditions of Theorem 2.2 (and under a Chernoff–Savage type condition on the score generating function).

Theorem 2.3. *Assume conditions (F.1)–(F.2), (X.1)–(X.4) and (D.1)–(D.3) with constants a, b, δ satisfying additional restrictions*

$$0 < a < b < \tfrac{1}{2} \qquad 0 < b - a - \delta < \tfrac{1}{2} - b. \tag{2.13}$$

Then, as $n \to \infty$,

$$\sup_{0\leqslant\alpha\leqslant 1}\left\{n^{-1/2}\left|\sum_{i=1}^{n}[d_{ni}(\hat{a}_{ni}(\alpha) - (1-\alpha)) - (d_{ni} - \hat{d}_{ni})\tilde{a}_i(\alpha)]\right|\right\} \xrightarrow{p} 0. \tag{2.14}$$

Moreover, the process

$$\left\{\Gamma^{-1} n^{-1/2} \sum_{i=1}^{n} d_{ni}(\hat{a}_{ni}(\alpha) - (1-\alpha)) \colon 0 \leqslant \alpha \leqslant 1\right\} \tag{2.15}$$

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

Before proving Theorems 2.1–2.3, we shall first state a crucial approximation of the criterion in (1.3) by a quadratic function of $\mathbf{t}$, uniform in an appropriate neighbourhood of $\boldsymbol{\beta}$ and for $\alpha \in [\alpha_n^*, 1-\alpha_n^*]$.

Lemma 2.2. *For* $\mathbf{t} \in \mathbb{R}^p$ *and* $\alpha \in (0, 1)$*, denote*

$$r_n(\mathbf{t}, \alpha) = \sigma_\alpha^{-1} \sum_{i=1}^{n} [\rho_\alpha(E_{i\alpha} - n^{-1/2}\sigma_\alpha \mathbf{x}_i'\mathbf{t}) - \rho_\alpha(E_{i\alpha})] + n^{-1/2}\mathbf{t}' \sum_{i=1}^{n} \mathbf{x}_i \psi_\alpha(E_{i\alpha}) - \frac{1}{2}(\alpha(1-\alpha))^{1/2}\mathbf{t}'\mathbf{D}_n\mathbf{t}, \tag{2.16}$$

with

$$\psi_\alpha(z) = \alpha - I[z < 0], \qquad z \in \mathbb{R}^1 \tag{2.17}$$

and

$$E_{i\alpha} = E_i - F^{-1}(\alpha), \qquad i = 1, \ldots, n, \qquad 0 < \alpha < 1. \tag{2.18}$$

Then, as $n \to \infty$,

$$\sup\{|r_n(\mathbf{t}, \alpha)| \colon \alpha_n^* \leqslant \alpha \leqslant 1-\alpha_n^*, \|\mathbf{t}\| \leqslant C_n\} = O_p(C_n^{3/2}(\log n)^{1/2} n^{-\frac{1}{4}+\frac{\Delta}{2}}). \tag{2.19}$$

As a consequence, we obtain the following approximation:

Lemma 2.3. *Let* $\{\mathbf{d}_n\}_{n=1}^{\infty}$ *be a sequence of vectors satisfying (D.1)–(D.3). Then, under the (F.1)–(F.2) and (X.1)–(X.4),*

$$\sup_{\|\mathbf{t}\| \leqslant C_n, \alpha_n^* \leqslant \alpha \leqslant 1-\alpha_n^*} n^{-1/2} \left| \sum_{i=1}^{n} (d_{ni} - \hat{d}_{ni})[\psi_\alpha(E_{i\alpha} - n^{-1/2}\sigma_\alpha \mathbf{x}_i'\mathbf{t}) - \psi_\alpha(E_{i\alpha})] \right| = o_p(C_n(\log n)^{1/4} n^{-\frac{1}{8}+\frac{\Delta}{4}}) \text{ as } n \to \infty. \tag{2.20}$$

Lemmas 2.2 and 2.3 are proved in the Appendix.

Proof of Theorem 2.1. The theorem follows from Lemma 2.2 by convexity arguments due to Pollard (1991) similarly as in the proof of Theorem 3.1 in Gutenbrunner *et al*. (1993). Hence, we omit the details. □

Proof of Theorem 2.2. First, note that $\sum_{i=1}^n \hat{d}_{ni} = 0$ by (D.1) and that (1.2) implies that $\sum_{i=1}^n \hat{d}_{ni}(\hat{a}_{ni}(\alpha) - (1-\alpha)) = 0$. Then we obtain (2.11) if we insert $n^{1/2}\sigma_\alpha(\hat{\boldsymbol{\beta}}(\alpha) - \boldsymbol{\beta}(\alpha))$ $[= O_p((\log\log n)^{1/2})]$ in (2.20) and note that

$$\sup\left\{\left|n^{-1/2}\sum_{i=1}^n d_{ni} I[Y_i = \mathbf{x}_i'\hat{\boldsymbol{\beta}}]\right| : \alpha_n^* \leqslant \alpha \leqslant 1-\alpha_n^*\right\} = o_p(n^{-\frac{1}{2}+\Delta}). \tag{2.21}$$

□

Proof of Theorem 2.3. By Theorem 2.2, it is sufficient to consider the behaviour of the process in (2.14) on the intervals $[0, \alpha_n^*]$ and $[1-\alpha_n^*, 1]$, where we have (with $d_{ni}^* = d_{ni} - \hat{d}_{ni}$)

$$\sup_{0\leqslant\alpha\leqslant\alpha_n^*}\left|n^{-1/2}\sum_{i=1}^n d_{ni}^*\hat{a}_{ni}(\alpha)\right| = \sup_{0\leqslant\alpha\leqslant\alpha_n^*}\left|n^{-1/2}\sum_{i=1}^n d_{ni}^*(1-\hat{a}_{ni}(\alpha))\right|$$
$$= O(n^{\frac{1}{2}+\Delta}\alpha_n^*) = o(1) \tag{2.22}$$

and

$$\sup_{0\leqslant\alpha\leqslant\alpha_n^*}\left|n^{-1/2}\sum_{i=1}^n d_{ni}^*\tilde{a}_i(\alpha)\right| = \sup_{0\leqslant\alpha\leqslant\alpha_n^*}\left|n^{-1/2}\sum_{i=1}^n d_{ni}^*(1-\tilde{a}_i(\alpha)+\alpha)\right|$$
$$\leqslant \max_{1\leqslant i\leqslant n}|d_{ni}^*| O_p([\alpha_n^*(1-\alpha_n^*)]^{1/2}) = o_p(1); \tag{2.23}$$

we obtain analogous bounds for $1-\alpha_n^* \leqslant \alpha \leqslant 1$. □

3. Linear regression rank statistics and regression rank tests

The close correspondence of RR scores to Hájek scores (defined in (1.6)), calculated for the (unobservable) errors E_i, explains why tests based on RR scores are as Pitman efficient as the corresponding ordinary rank tests. This relation is characterized in the following crucial theorem.

Theorem 3.1. *Let $R_n, \ldots, R_n$ denote the ranks of errors $E_1, \ldots, E_n$ and let $a_n^*(R_i, \alpha)$, $i = 1, \ldots, n$, $0 \leqslant \alpha \leqslant 1$, denote the Hájek scores defined in (1.6). Then:*

(i) Under the conditions of Theorem 2.2,

$$\sup_{\alpha_n^*\leqslant\alpha\leqslant 1-\alpha_n^*}\left|n^{-1/2}\sum_{i=1}^n [d_{ni}\hat{a}_{ni}(\alpha) - (d_{ni}-\hat{d}_{ni})a_n^*(R_i,\alpha)]\right| \xrightarrow{p} 0, \qquad \text{as } n\to\infty. \tag{3.1}$$

(ii) Under the conditions of Theorem 2.3,

$$\sup_{0\leqslant\alpha\leqslant 1}\left|n^{-1/2}\sum_{i=1}^n [d_{ni}\hat{a}_{ni}(\alpha) - (d_{ni}-\hat{d}_{ni})a_n^*(R_i,\alpha)]\right| \xrightarrow{p} 0, \qquad \text{as } n\to\infty. \tag{3.2}$$

***Proof*.** The propositions follow from Theorems 2.2 and 2.3 and from approximations in Hájek (1965). □

We are now in a position to define the modified tests based on linear RR statistics and show that they are asymptotically distribution-free for the class of distributions covered by (F.1)–(F.2). Choose a score generating function $\varphi(t)$, non-decreasing and square-integrable on (0, 1); let φ_n be φ censored at α_n^*, $1-\alpha_n^*$, that is,

$$\varphi_n(t) = \begin{cases} \varphi(\alpha_n^*), & \text{if } 0 \leqslant t < \alpha_n^*, \\ \varphi(t), & \text{if } \alpha_n^* \leqslant t \leqslant 1-\alpha_n^*, \\ \varphi(1-\alpha_n^*), & \text{if } 1-\alpha_n^* < t \leqslant 1. \end{cases} \tag{3.3}$$

Let $\hat{\mathbf{a}}_n(\alpha) = (\hat{a}_{n1}(\alpha), \ldots, \hat{a}_{nn}(\alpha))'$ be the RR corresponding to the model $\mathbf{Y} = \mathbf{X}\boldsymbol{\beta} + \mathbf{E}$ and let $\hat{\mathbf{b}}_{ni} = (\hat{b}_{n1}, \ldots, \hat{b}_{nn})'$ be the scores generated by φ_n in the following way:

$$\hat{b}_{ni} = -\int_0^1 \varphi_n(t)\,\mathrm{d}\hat{a}_{ni}(t), \qquad i = 1, \ldots, n. \tag{3.4}$$

Let $\{\mathbf{d}_n\}_{n=1}^{\infty}$ be a sequence of n-dimensional vectors $\mathbf{d}_n$ orthogonal to $\mathbf{X}_n$, $n = 1, 2, \ldots$. Consider the linear RR statistics

$$S_{nn} = n^{-1/2}\sum_{i=1}^{n} d_{ni}\hat{b}_{ni}, \qquad n = 1, 2, \ldots. \tag{3.5}$$

The results of Section 2 enable us to derive an asymptotic representation of S_{nn} by a sum of independent summands, parallel to that derived by Hájek (1961). This, in combination with Theorem 3.1, enables us to incorporate Hájek's results into the asymptotic theory of tests based on regression rank scores.

The asymptotic representation of linear RR statistics will be derived for score generating functions satisfying a condition of Chernoff–Savage type, including the inverse normal distribution function.

Theorem 3.2. *Let $\varphi(t)$, $0 < t < 1$, be a non-decreasing square-integrable function such that $\varphi'(t)$ exists for $0 < t < \alpha_0$ and $1-\alpha_0 < t < 1$, $0 < \alpha_0 < \frac{1}{2}$; in this domain $\varphi'(t)$ satisfies*

$$|\varphi'(t)| \leqslant c(t(1-t))^{-1-\delta^*}, \tag{3.6}$$

where $c > 0$ and $0 < \delta^* < (1+2a)/8$. Then, under (F.1)–(F.2), (X.1)–(X.4) and (D.1)–(D.3),

$$S_{nn} = T_n + o_p(1) \qquad \text{as } n \to \infty, \tag{3.7}$$

where

$$T_n = n^{-1/2}\sum_{i=1}^{n} d_{ni}\varphi(F(E_i)). \tag{3.8}$$

***Proof*.** Notice that $\hat{a}_{ni}(t) - \tilde{a}_i(t) = 0$ at $t = 0, 1$. Integrating by parts, we obtain

$$-\int_0^1 \varphi_n(t)\,\mathrm{d}(\hat{a}_{ni}(t)-\tilde{a}_i(t)) = \int_0^1 (\hat{a}_{ni}(t)-\tilde{a}_i(t))\,\mathrm{d}\varphi_n(t)$$

$$= \int_{\alpha_n^*}^{1-\alpha_n^*} (\hat{a}_{ni}(t)-\tilde{a}(t))\,\mathrm{d}\varphi(t). \tag{3.9}$$

By Theorem 2.2 and by the dominated convergence theorem,

$$n^{-1/2}\sum_{i=1}^n d_{ni}\int_{\alpha_0}^{1-\alpha_0} (\hat{a}_{ni}(t)-\tilde{a}_i(t))\,\mathrm{d}\varphi_n(t) = o_p(1). \tag{3.10}$$

Moreover,

$$\left| n^{-1/2}\sum_{i=1}^n d_{ni}\int_{\alpha_n^*}^{\alpha_0} (\hat{a}_{ni}(t)-\tilde{a}_i(t))\,\mathrm{d}\varphi_n(t)\right| \leqslant \left|\int_{\alpha_n^*}^{\alpha_0} (t(1-t))^{-1-\delta^*} n^{-1/2}\sum_{i=1}^n d_{ni}[\hat{a}_{ni}(t)-\tilde{a}(t)]\,\mathrm{d}t\right|$$

$$\leqslant K\left|[t^{-\delta^*}]_{\alpha_n^*}^{\alpha_0}\right| o_p(C_n^2(\log n)^{1/2} n^{-\frac{1}{8}+\frac{\Delta}{4}}) = o_p(1), \tag{3.11}$$

and we obtain an analogous conclusion for the integral over $(1-\alpha_0, 1-\alpha_n^*)$. Thus, combining (3.9)–(3.11), we obtain

$$S_{nn} = T_{nn} + o_p(1) \tag{3.12}$$

as $\to \infty$, where

$$T_{nn} = n^{-1/2}\sum_{i=1}^n d_{ni}\int_0^1 \tilde{a}_i(t)\,\mathrm{d}\varphi_n(t) = n^{-1/2}\sum_{i=1}^n d_{ni}\varphi_n(F(E_i)). \tag{3.13}$$

Furthermore,

$$\operatorname{var}(T_{nn}-T_n)^2 \leqslant n^{-1}\sum_{i=1}^n d_{ni}^2\left\{\int_0^{\alpha_n^*}[\varphi(t)-\varphi(\alpha_n^*)]^2 dt + \int_{1-\alpha_n^*}^1 [\varphi(t)-\varphi(1-\alpha_n^*)]^2\mathrm{d}t\right\}$$

$$= \Gamma^2 o(1) \qquad \text{as } n\to\infty. \tag{3.14}$$

(3.12) and (3.14) then imply (3.7). □

Let us now consider the extended model (1.7) and the problem of testing the hypothesis H_0: $\boldsymbol{\delta} = 0$, $\boldsymbol{\beta}$ unspecified. Let $\hat{\mathbf{a}}_n(\alpha)$ denote the regression rank scores corresponding to the model (1.1) under the hypothesis. Choose a score generating function φ satisfying the conditions imposed in Theorem 3.1; let φ_n be the function defined in (3.3) and let $\hat{b}_{ni}$ be the scores defined in (3.4). Consider the vector of linear regression rank-scores statistics

$$\mathbf{S}_{nn} = n^{-1/2}(\mathbf{Z}_n - \hat{\mathbf{Z}}_n)'\hat{\mathbf{b}}_n, \tag{3.15}$$

where $\hat{\mathbf{Z}}_n = \mathbf{X}_n(\mathbf{X}_n'\mathbf{X}_n)^{-1}\mathbf{X}_n'\mathbf{Z}_n$. Furthermore, denote

$$\mathbf{Q}_n = n^{-1}(\mathbf{Z}_n - \hat{\mathbf{Z}}_n)'(\mathbf{Z}_n - \hat{\mathbf{Z}}_n). \tag{3.16}$$

We propose the statistic

$$V_n = \mathbf{S}'_{nn}\mathbf{Q}_n^{-1}\mathbf{S}_{nn}/A^2(\varphi), \tag{3.17}$$

with $A^2(\varphi) = \int_0^1(\varphi(u) - \overline{\varphi})^2\,\mathrm{d}u$, $\overline{\varphi} = \int_0^1 \varphi(u)\,\mathrm{d}u$, as the test criterion for H_0. The following theorem shows that the test of H_0 based on V_n, which rejects H_0 provided V_n exceeds the critical value of the χ^2_q distribution, is asymptotically distribution-free for the class of distributions satisfying (F.1)–(F.2) and its Pitman efficiency coincides with that of the ordinary rank test with the same score function.

Theorem 3.3. *Consider the model* $\mathbf{Y}_n = \mathbf{X}_n\boldsymbol{\beta} + \mathbf{Z}_n\boldsymbol{\delta} + \mathbf{E}_n$, *where* $\mathbf{X}_n$ *is of order* $(n \times p)$, $\mathbf{Z}_n$ *is of order* $(n \times q)$, $\mathbf{X}_n$ *satisfies (X.1)–(X.4), and* $\mathbf{X}_n$ *and* $\mathbf{Z}_n$ *satisfy*

$$\max_{1\leqslant i\leqslant n} \|\mathbf{z}_{ni}\| = O(n^{\Delta}), \tag{3.18}$$

with Δ *from (X.4),* $\mathbf{z}'_{ni}$ *being the ith row of* $\mathbf{Z}_n$, *and*

$$\mathbf{Q}_n \to \mathbf{Q} \qquad \text{as } n \to \infty, \tag{3.19}$$

where $\mathbf{Q}$ *is a positive definite* $(q \times q)$ *matrix. Assume that the distribution function F of the errors satisfies (F.1)–(F.2). Then:*

(i) *under* H_0, *the statistic* V_n *defined in (3.17) has asymptotically* χ^2 *distribution with q d.f.;*

(ii) *if, moreover, F has positive and finite Fisher information,*

$$0 < I(F) = \int\left(\frac{f'(x)}{f(x)}\right)^2 \mathrm{d}F(x) < \infty,$$

then, under H_n: $\boldsymbol{\delta} = n^{-1/2}\boldsymbol{\delta}_0$, $\boldsymbol{\delta}_0 \in \mathbb{R}^q$ *fixed,* $\boldsymbol{\beta} \in \mathbb{R}^p$ *unspecified,* V_n *has an asymptotic non-central* χ^2 *distribution with q d.f. and non-centrality parameter*

$$\eta^2 = \boldsymbol{\delta}'_0\mathbf{Q}\boldsymbol{\delta}_0[\gamma^2(\varphi, F)/A^2(\varphi)], \tag{3.20}$$

with

$$\gamma(\varphi, F) = -\int_0^1 \varphi(t)\,\mathrm{d}f(F^{-1}(t)). \tag{3.21}$$

Proof. Part (i) follows from Theorem 3.2. Part (ii) follows from Theorem 3.1 with an application of the contiguity and the asymptotic theory of rank tests under contiguous alternatives (Hájek 1962). □

As an illustration, apply the above results to the k-sample model

$$Y_{ij} = \beta_0 + \beta_i + E_{ij}, \qquad j = 1, \ldots, n_i, \quad i = 1, \ldots, k, \tag{3.22}$$

with $\beta_1 = 0$. We want to test the hypothesis H: $\beta_k = 0$, and admit that the distribution of the errors may be as heavy as Cauchy. The regression rank scores under the submodel of (3.22) with $\beta_1 = \beta_k = 0$ are such that $(\hat{a}_{i1}(\alpha), \ldots, \hat{a}_{in_i})$ generate the same ranks of the ith sample $Y_{i1}, \ldots, Y_{in_i}$, $i = 2, \ldots, k-1$, as the Hájek scores, while

$$(\hat{a}_{11}(\alpha), \ldots, \hat{a}_{1n_1}(\alpha), \hat{a}_{k1}(\alpha), \ldots, \hat{a}_{kn_k}(\alpha))$$

analogously generate the ranks in the combined first and kth samples. On the other hand, $a = 1$ in (F.2) for the Cauchy tails and hence $\alpha_n^* = n^{-1/(3+2\delta)}$, $\delta > 0$. If we take the Wilcoxon score function $\varphi(\alpha) = \alpha - \frac{1}{2}$, $0 \leqslant \alpha \leqslant 1$, we finally obtain the two-sample Wilcoxon test of the shift between the first and the kth samples, censored for the observations with the ranks below $[n\alpha_n^*] + 1$ or above $n - [n\alpha_n^*]$.

Appendix: Proofs of Lemmas 2.2 and 2.3

Proof of Lemma 2.2. For a fixed $\mathbf{t} \in \mathbb{R}^p$, denote

$$\varepsilon_{ni} = \varepsilon_i = n^{-1/2}\sigma_\alpha \mathbf{x}_i'\mathbf{t}, \qquad i = 1, \ldots, n. \tag{A.1}$$

Notice that $\max_{1\leqslant i\leqslant n}|\varepsilon_i\sigma_\alpha^{-1}| = O(n^{-\frac{1}{2}+\Delta})$ by (2.4). Moreover, we obtain from Lemma 2.1 that

$$\sigma_\alpha(\alpha(1-\alpha))^{\frac{1}{2}+a} \to a, \qquad \textit{as } \alpha \to 0, 1, \tag{A.2}$$

and hence, noting (X.3) and (X.4),

$$\max_{1\leqslant i\leqslant n}|\varepsilon_i| = O(C_n n^{-\frac{1}{2}+\Delta}(\alpha(1-\alpha))^{-\frac{1}{2}-a}) \qquad \text{as } \alpha \to 0, 1. \tag{A.3}$$

Denote, for $i = 1, \ldots, n$,

$$Q_i(\mathbf{t}, \alpha) = Q_i = \sigma_\alpha^{-1}[\rho_\alpha(E_{i\alpha} - \varepsilon_i) - \rho_\alpha(E_{i\alpha}) + \varepsilon_i\psi_\alpha(E_{i\alpha})]. \tag{A.4}$$

Then we obtain by simple arithmetic that

$$Q_i = \sigma_\alpha^{-1}\{(E_{i\alpha} - \varepsilon_i)I[\varepsilon_i < E_{i\alpha} < 0] + (\varepsilon_i - E_{i\alpha})I[0 < E_{i\alpha} < \varepsilon_i]\}. \tag{A.5}$$

Hence, by (A.1)–(A.5), for $\varepsilon_i > 0$,

$$\begin{aligned}
\sigma_\alpha \mathrm{E}Q_i &= \int_{F^{-1}(\alpha)}^{F^{-1}(\alpha)+\varepsilon_i} (\varepsilon_i - x + F^{-1}(\alpha))\,\mathrm{d}F(x) \\
&= \int_{F^{-1}(\alpha)}^{F^{-1}(\alpha)+\varepsilon_i}\int_{F^{-1}(\alpha)}^{x} f(y)\,\mathrm{d}y\,\mathrm{d}x \\
&= f(F^{-1}(\alpha))\int_{F^{-1}(\alpha)}^{F^{-1}(\alpha)+\varepsilon_i}(x - F^{-1}(\alpha))\,\mathrm{d}x + \int_{F^{-1}(\alpha)}^{F^{-1}(\alpha)+\varepsilon_i}\int_{F^{-1}(\alpha)}^{x}\int_{F^{-1}(\alpha)}^{y} f'(z)\,\mathrm{d}z\,\mathrm{d}y\,\mathrm{d}x \qquad \text{(A.6)} \\
&= f(F^{-1}(\alpha))\frac{\varepsilon_i^2}{2} + \int_{F^{-1}(\alpha)}^{F^{-1}(\alpha)+\varepsilon_i}\int_{F^{-1}(\alpha)}^{x}\int_{F^{-1}(\alpha)}^{y} f'(z)\,\mathrm{d}z\,\mathrm{d}y\,\mathrm{d}x.
\end{aligned}$$

Thus, by Lemma 2.1(iii), given $\eta > 0$, there exists α_0 such that, for $0 < \alpha < \alpha_0$,

$$\left|\sigma_\alpha \mathrm{E}Q_i - f(F^{-1}(\alpha))\frac{\varepsilon_i^2}{2}\right| \leqslant (1+\eta)(1+a)\int_{F^{-1}(\alpha)}^{F^{-1}(\alpha)+\varepsilon_i}\int_{F^{-1}(\alpha)}^{x}\int_{\alpha}^{F(y)}\frac{f(F^{-1}(u)}{u}\,\mathrm{d}u\,\mathrm{d}y\,\mathrm{d}x; \quad \text{(A.7)}$$

and further by Lemma 2.1(i), for $0<\alpha<\alpha_0$,

$$\begin{aligned}\left|\sigma_\alpha \mathrm{E}Q_i - f(F^{-1}(\alpha))\frac{\varepsilon_i^2}{2}\right| &\leqslant K_1(1+\eta)^2\int_{F^{-1}(\alpha)}^{F^{-1}(\alpha)+\varepsilon_i}\int_{F^{-1}(\alpha)}^{x}\int_{\alpha}^{F(y)} u^a\,\mathrm{d}u\,\mathrm{d}y\,\mathrm{d}x \\ &\leqslant K_2[F(F^{-1}(\alpha)+\varepsilon_i)]^{1+2a}\varepsilon_i^3 \\ &\leqslant K_3\alpha^{(1+a)}\varepsilon_i^3,\end{aligned} \quad \text{(A.8)}$$

where K_1, K_2, K_3 are positive constants. We obtain analogous inequalities for $\varepsilon_i<0$, $i=1,\ldots,n$.

Hence, combining (A.1), (A.2), (A.6)–(A.8), we arrive at

$$\left|\sum_{i=1}^{n}\left[\mathrm{E}Q_i - \frac{\varepsilon_i^2}{2}\sigma_\alpha^{-1}f(F^{-1}(\alpha))\right]\right| = O(C_n^3 n^{-1/2}(\alpha(1-\alpha))^{-a}) \quad \text{(A.9)}$$

uniformly for $\alpha_n^* \leqslant \alpha \leqslant 1-\alpha_n^*$, as $n\to\infty$. We shall first prove that

$$P\left\{\left|\sum_{i=1}^{n}(Q_i - \mathrm{E}Q_i)\right| \geqslant \eta B_n\right\} \leqslant 2n^{-\eta^2/3}, \quad \text{(A.10)}$$

for any $\eta>0$ and $n\geqslant n_0$, where

$$B_n = n^{-1/4+\Delta/2}C_n^{3/2}(\log n)^{1/2}. \quad \text{(A.11)}$$

Actually, by the Bernstein–Bennett inequality (see (2.13) in Hoeffding 1963),

$$P\left\{\sum_{i=1}^{n}(Q_i - \mathrm{E}Q_i) > nt\right\} \leqslant \exp\{-\tau h(\lambda)\}, \quad \text{(A.12)}$$

for $t<b$, provided $Q_i - \mathrm{E}Q_i \leqslant b$, $i=1,\ldots,n$, where

$$\tau = \frac{nt}{b}, \qquad \lambda = \frac{bt}{\sigma^2}, \qquad \sigma^2 = \frac{1}{n}\sum_{i=1}^{n}\operatorname{var}Q_i, \qquad h(\lambda) = \frac{\lambda}{2(1+\frac{1}{3}\lambda)}. \quad \text{(A.13)}$$

By (A.5) and (A.8), as $n\to\infty$,

$$Q_i \leqslant n^{-1/2+\Delta}C_n, \qquad \mathrm{E}Q_i \leqslant n^{-1+2\Delta}C_n^2(1+\mathrm{o}(1)) \quad \text{(A.14)}$$

uniformly in $\alpha_n^* \leqslant \alpha \leqslant 1-\alpha_n^*$, hence

$$b = b_n = K_1 n^{-1/2+\Delta}C_n \qquad \text{for } n \geqslant n_0. \quad \text{(A.15)}$$

Moreover, (A.5) and (A.8) imply

$$\operatorname{var}Q_i \leqslant \mathrm{E}Q_i^2 \leqslant \sigma_\alpha^{-1}|\varepsilon_i|\mathrm{E}Q_i, \qquad i=1,\ldots,n,$$

hence

$$\sigma^2 \leqslant K_2 n^{-3/2} C_n^3 \qquad \text{for } n \geqslant n_0. \tag{A.16}$$

Put $nt_n = \eta B_n$, that is, $t = t_n = \eta n^{-5/4+\Delta/2} C_n^{3/2} (\log n)^{1/2}$. Then $t_n < b_n$, $\lambda = b_n t_n / \sigma^2$ and (4.12) gives

$$P\left\{\sum_{i=1}^{n} (Q_i - \mathrm{E}Q_i)) \geqslant \eta B_n)\right\} \leqslant \exp\left\{-\frac{nt_n^2}{2} \cdot \frac{1}{\sigma^2 + \dfrac{bt}{3}}\right\}$$

$$\leqslant \exp\left\{-\frac{2}{3}\frac{\eta^2}{2} \cdot \log n\right\} \leqslant n^{-\eta^2/3} \tag{A.17}$$

for $n \geqslant n_0$. Because Q_i are non-negative random variables, we obtain an analogous inequality for $P(\sum_{i=1}^n (Q_i - \mathrm{E}Q_i) \leqslant -\eta B_n) \leqslant n^{-\eta^2/3}$ and thus we arrive at (A.10). Hence, finally, regarding (2.16), (A.1), (A.9), (A.11) and (A.10), we get

$$P\{|r_n(\mathbf{t}, \alpha)| \geqslant (\eta + 1) B_n\} \leqslant 2n^{-\eta^2/3} \tag{A.18}$$

for $n \geqslant n_0$, any $\eta > 0$ and B_n of (A.11).

Let us now choose intervals $[\alpha_\nu, \alpha_{\nu+1}]$ of length n^{-5} covering $[\alpha_n^*, 1 - \alpha_n^*]$ and balls of radius n^{-5} covering $\{\mathbf{t}: \|\mathbf{t}\| \leqslant C_n\}$. Let $(\alpha_1, \alpha_2) \in (\alpha_\nu, \alpha_{\nu+1})$ and let $\mathbf{t}_1, \mathbf{t}_2$ lie in the same ball. Then, by (A.2),

$$|(\sigma_{\alpha_1}/\sigma_{\alpha_2}) - 1| = O(n^{-4-(10b/(1+2b))}). \tag{A.19}$$

For fixed i, $1 \leqslant i \leqslant n$, write

$$|Q_i(\mathbf{t}_2, \alpha_2) - Q_i(\mathbf{t}_1, \alpha_1)| \leqslant |Q_i(\mathbf{t}_2, \alpha_2) - Q_i(\mathbf{t}_1, \alpha_2)| + |Q_i(\mathbf{t}_1, \alpha_2) - Q_i(\mathbf{t}_1, \alpha_2)| \tag{A.20}$$

and consider the terms on the right-hand side separately. By (A.20), (A.1) and (2.4),

$$|Q_i(\mathbf{t}_2, \alpha_2) - Q_i(\mathbf{t}_1, \alpha_2)| \leqslant \sigma_\alpha^{-1} |\varepsilon_{i\alpha_2 t_2} - \varepsilon_{i\alpha_2 t_1}| \leqslant n^{-1/2} |\mathbf{x}_i'(\mathbf{t}_2 - \mathbf{t}_1)| = O(n^{-5.25}). \tag{A.21}$$

For the corresponding centring term we obtain the bound

$$\left|\frac{1}{2}(\varepsilon_{i\alpha_2 t_2}^2 - \varepsilon_{i\alpha_2 t_1}^2) f(F^{-1}(\alpha_2)) \sigma_{\alpha_2}^{-1}\right| = O(C_n n^{-5.5+\Delta}) = O(n^{-5.25}). \tag{A.22}$$

Consider the second term on the right-hand side of (A.20), which we denote Q^* for the sake of brevity. We should distinguish two cases:

(i) If $\varepsilon_{i\alpha_2 t_1} < E_{i\alpha_2} < 0$ and $\varepsilon_{i\alpha_1 t_1} < E_{i\alpha_1} < 0$ (or $0 < E_{i\alpha_2} < \varepsilon_{i\alpha_2 t_1}$ and $0 < E_{i\alpha_1} < \varepsilon_{i\alpha_1 t_1}$), then

$$\begin{aligned} |Q^*| &\leqslant |\sigma_{\alpha_2}^{-1} - \sigma_{\alpha_1}^{-1}| |\varepsilon_{i\alpha_2 t_1}| + \sigma_{\alpha_1}^{-1}(|F^{-1}(\alpha_2) - F^{-1}(\alpha_1)| + |\varepsilon_{i\alpha_2 t_1} - \varepsilon_{i\alpha_1 t_1}|) \\ &\leqslant 2n^{-1/2} |1 - (\sigma_{\alpha_2}/\sigma_{\alpha_1})| |\mathbf{x}_i'\mathbf{t}_1| + \sigma_{\alpha_1}^{-1}(|\alpha_2 - \alpha_1| / f(F^{-1}(\alpha_1)) + o(n^{-5}) \qquad \text{(A.23)} \\ &= O(C_n n^{-4.5+\Delta}) = O(n^{-4.2}). \end{aligned}$$

(ii) If $\varepsilon_{i\alpha_2 t_1} < E_{i\alpha_2} < 0$ and $E_{i\alpha_1} < \varepsilon_{i\alpha_1 t_1} < 0$ (or $\varepsilon_{i\alpha_1 t_1} < E_{i\alpha_1} < 0$ and $E_{i\alpha_2} < \varepsilon_{i\alpha_2 t_1} < 0$), then

$$|Q^*| = \sigma_{\alpha_2}^{-1}|E_{i\alpha_2} - \varepsilon_{i\alpha_2 t_1}| = \sigma_{\alpha_2}^{-1}|E_{i\alpha_1} + F^{-1}(\alpha_1) - F^{-1}(\alpha_2) - \varepsilon_{i\alpha_2 t_1}| \qquad \text{(A.24)}$$
$$\leqslant \sigma_{\alpha_2}^{-1}(|\varepsilon_{i\alpha_1 t_1} - \varepsilon_{i\alpha_2 t_2}| + |F^{-1}(\alpha_2) - F^{-1}(\alpha_1)|) = O(n^{-4.2})$$

by analogous considerations as in (i).

Moreover, we obtain for the centring terms in both cases

$$\tfrac{1}{2}|\sigma_{\alpha_2}^{-1}\varepsilon_{i\alpha_2 t_1} f(F^{-1}(\alpha_2)) - \sigma_{\alpha_1}^{-1}\varepsilon_{i\alpha_1 t_1} f(F^{-1}(\alpha_1))| = O(n^{-3}). \qquad \text{(A.25)}$$

Let us fix one set S_ν in the decomposition of the set $[\alpha_n^*, 1-\alpha_n^*] \times \{\mathbf{t}\colon \|\mathbf{t}\| \leqslant C_n\}$; the number of such sets is at most $(2C_n)^p n^{5(p+1)}$. It follows from (A.19)–(A.25) that

$$\sup_{S_\nu}|r_n(\mathbf{t}_2, \alpha_2) - r_n(\mathbf{t}_1, \alpha_1)| \leqslant K_1 n^{-3}, \qquad \text{(A.26)}$$

where $0 < K_1 < \infty$. By (A.18),

$$P\{\sup_{S_\nu}|r_n(\mathbf{t}, \alpha)| \geqslant (\eta+1)B_n\} \leqslant 2n^{-\eta^2/3}, \qquad \text{(A.27)}$$

and finally

$$P\Big\{\sup_{\|\mathbf{t}\|\leqslant C_n, \alpha_n^* \leqslant \alpha \leqslant 1-\alpha_n^*}|r_n(\mathbf{t}, \alpha)| \geqslant 2(\eta+1)B_n\Big\} \leqslant \sum_\nu P\Big\{\sup_{S_\nu}|r_n(\mathbf{t}, \alpha)| \geqslant 2(\eta+1)B_n\Big\}$$
$$\leqslant 4C_n^p n^{5(p+1)} n^{-\eta^2/3} = o_p(1) \qquad \text{(A.28)}$$

for $\eta^2 > 15(p+1)$; and this entails

$$\sup\{|r_n(\mathbf{t}, \alpha)|\colon \|\mathbf{t}\| \leqslant C_n,\ \alpha_n^* \leqslant \alpha \leqslant 1-\alpha_n^*\} = O_p(B_n)$$
$$= O_p(C_n^{3/2}(\log n)^{1/2} n^{-1/4+\Delta/2}) \qquad \text{(A.29)}$$

as $n \to \infty$. □

Proof of Lemma 2.3. Consider the model $\mathbf{Y} = \mathbf{X}^*\boldsymbol{\beta}^* + \mathbf{E}$ with $\mathbf{X}^* = (\mathbf{X} \quad (\mathbf{d}_n - \hat{\mathbf{d}}_n))$ and $\boldsymbol{\beta}^* = (\beta_1, \ldots, \beta_p, \beta_{p+1})'$. Then

$$\mathbf{X}^* = \begin{pmatrix} \mathbf{X}'\mathbf{X} & 0 \\ 0 & (\mathbf{d}_n - \hat{\mathbf{d}}_n)'(\mathbf{d}_n - \hat{\mathbf{d}}_n) \end{pmatrix}$$

and the conditions of Lemma 2.2 are also satisfied when replacing $\mathbf{X}$ by $\mathbf{X}^*$ and $\mathbf{t}$ by $\mathbf{t}^* \in \mathbb{R}^{p+1}$. Now, denoting

$$A_n = C_n^{1/2}B_n = C(\log n)^{1/2}(\log\log n)n^{-1/4+\Delta/2} \qquad \text{(A.30)}$$

we obtain from Lemma 2.2 (see (A.29)) that

$$\Big\{\sup\Big|\sigma_\alpha^{-1}\sum_{i=1}^{n}[\rho_\alpha(E_{i\alpha}-n^{-1/2}\sigma_\alpha(\mathbf{x}_i^*)'\mathbf{t}^*)-\rho_\alpha(E_{i\alpha})]+n^{-1/2}(\mathbf{t}^*)'\sum_{i=1}^{n}\mathbf{x}_i^*\psi_\alpha(E_i\alpha)$$

$$-\tfrac{1}{2}n^{-1}(\alpha(1-\alpha))^{1/2}(\mathbf{t}^*)'(\mathbf{X}^*)'\mathbf{X}^*\mathbf{t}^*\Big|:\|\mathbf{t}^*\|\leqslant C_n,\ \alpha_n^*\leqslant\alpha\leqslant 1-\alpha_n^*\Big\}=o_p(A_n) \quad (A.31)$$

as $n\to\infty$. Hence, also (denoting $d_i^*=d_i-\hat{d}_i$, $i=1,\ldots,n$, for the sake of brevity),

$$\Big|\sigma_\alpha^{-1}\sum_{i=1}^{n}[\rho_\alpha(E_{i\alpha}-n^{-1/2}\sigma_\alpha(\mathbf{x}_i'\mathbf{t}+d_i^*t_{p+1}))-\rho_\alpha(E_{i\alpha}-n^{-1/2}\sigma_\alpha\mathbf{x}_i'\mathbf{t})]$$

$$+n^{-1/2}t_{p+1}\sum_{i=1}^{n}d_i^*\psi(E_{i\alpha})-\frac{1}{2}n^{-1}(\alpha(1-\alpha))^{1/2}t_{p+1}^2(\mathbf{d}^*)'\mathbf{d}^*\Big|=o_p(A_n) \quad (A.32)$$

uniformly in $\|\mathbf{t}\|\leqslant C_n$, $\alpha_n^*\leqslant\alpha\leqslant 1-\alpha_n^*$. Expression (A.32) further implies that, for any $0<\delta<C_n$,

$$\sup_{\|\mathbf{t}\|\leqslant C_n,\alpha_n^*\leqslant\alpha\leqslant 1-\alpha_n^*}\Big|\int_0^\delta\Big\{-n^{-1/2}\sum_{i=1}^{n}d_i^*[\psi_\alpha(E_{i\alpha}-n^{-1/2}\sigma_\alpha(\mathbf{x}_i'\mathbf{t}+d_i^*u))-\psi_\alpha(E_{i\alpha})]\Big\}\,\mathrm{d}u$$

$$-(\alpha(1-\alpha))^{1/2}\Gamma_n\int_0^\delta u\,\mathrm{d}u\Big|\leqslant A_n\varepsilon \quad (A.33)$$

with probability at least $1-\eta$ for $n\geqslant n_0$, where $\varepsilon,\eta>0$ are arbitrary numbers; we could give similar statement for the integration over $(-\delta, 0)$. Notice that both integrands in (A.33) are non-decreasing in u. Then (A.33) implies

$$-n^{-1/2}\sum_{i=1}^{n}d_i^*[\varphi_\alpha(E_{i\alpha}-n^{-1/2}\sigma_\alpha\mathbf{x}_i'\mathbf{t})-\psi_\alpha(E_{i\alpha})] \quad (A.34)$$

$$\leqslant\delta^{-1}\int_0^\delta\Big\{-n^{-1/2}\sum_{i=1}^{n}d_i^*[\psi_\alpha(E_{i\alpha}-n^{-1/2}\sigma_\alpha(\mathbf{x}_i'\mathbf{t}+d_i^*u))-\psi_\alpha(E_{i\alpha})]\Big\}\,\mathrm{d}u$$

$$\leqslant(\alpha(1-\alpha))^{1/2}\Gamma_n\frac{\delta}{2}+\frac{A_n\varepsilon}{\delta}\leqslant K\delta+\frac{A_n\varepsilon}{\delta},$$

and analogously we obtain

$$-n^{-1/2}\sum_{i=1}^{n}d_i^*[\psi_\alpha(E_{i\alpha}-n^{-1/2}\sigma_\alpha\mathbf{x}_i'\mathbf{t})-\psi_\alpha(E_i\alpha)]\geqslant -K\delta-\frac{A_n\varepsilon}{\delta}. \quad (A.35)$$

Hence, if we put $\delta=(\mathrm{A}_n\varepsilon)^{1/2}$ we obtain that, for $n>n_0$,

$$P\Big\{\Big|n^{-1/2}\sum_{i=1}^{n}d_i^*[\psi_\alpha(E_i\alpha-n^{-1/2}\sigma_\alpha\mathbf{x}_i'\mathbf{t})-\psi_\alpha(E_i\alpha)]\Big|>(K+1)(A_n\varepsilon)^{1/2}\Big\}<\eta.$$

□

Acknowledgement

The author's research was supported by Czech Republic grant no 201/96/0230. The author thanks two referees and the Editor and Associate Editor for their careful reading and for their comments, which helped to improve the text.

References

Adichie, J.N. (1984) Rank tests in linear models. In P.R. Krishnaiah and P.K. Sen (eds), *Handbook of Statistics*, Vol. 4, pp. 229–257. New York: Elsevier Science.

Gutenbrunner, C. (1986) Zur Asymptotik von Regressions-Quantil-Prozessen und daraus abgeleiten Statistiken. PhD thesis, Universität Freiburg.

Gutenbrunner, C. (1994) Tests for heteroscedasticity based on regression quantiles and regression rank scores. In P. Mandl and M. Hušková, (eds), *Asymptotic Statistics*, pp. 249–260. Heidelberg: Physica-Verlag.

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

Gutenbrunner, C., Jurečková, J., Koenker, R. and Portnoy, S. (1993) Tests of linear hypotheses based on regression rank scores. *J. Nonparametr. Statist.*, **2**, 307–331.

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

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

Hájek, J. (1965) Extension of the Kolmogorov–Smirnov test to the regression alternatives. In L. Le Cam and J. Neyman (eds), *Bernoulli (1713) – Bayes (1763) – Laplace (1813): Proceedings of an International Research Seminar*, pp. 45–60. Berlin: Springer-Verlag.

Hallin, M. and Jurečková, J. (1996) Optimal tests for autoregressive models based on autoregression rank scores. *Submitted.*

Hallin, M., Zahaf, T., Jurečková, J., Kalvová, J. and Picek, J. (1997) Non-parametric tests in AR models with application to climatic data. *Environmetrics*, **8**, 651–660.

Hallin, M., Jurečková, J., Picek, J. and Zahaf, T. (1999) Nonparametric tests of independence of two autoregressive time series based on autoregression rank scores. *J. Statist. Plann. Inference*, **75**, 319–330.

Hettmansperger, T.P. (1984) *Statistical Inference Based on Ranks.* New York: Wiley.

Hoeffding, W. (1963) Probability inequalities for sums of bounded random variables. *J. Amer. Statist. Assoc.*, **56**, 13–30.

Jurečková, J. (1992) Tests of Kolmogorov–Smirnov type based on regression rank scores. In J.Á. Víšek, (ed.), *Transactions of the 11th Prague Conference on Information Theory, Statistical Decision Functions and Random Processes*, pp. 41–49. Prague: Academia, and Amsterdam: Kluwer.

Koenker, R. and Bassett, G. (1978) Regression quantiles. *Econometrica*, **46**, 466–476.

Koenker, R. and d'Orey, V. (1987) Algorithm AS 229: Computing regression quantiles. *Appl. Statist.*, **36**, 383–393.

Koenker, R. and d'Orey, V. (1994) A remark on algorithm AS 229: Computing the dual regression quantiles and regression rank scores. *Appl. Statist.*, **43**, 410–414.

Koul, H.L. and Saleh, A.K.Md.E. (1995) Autoregression quantiles and related rank scores processes. *Ann. Statist.,* **23**, 670–689.

Osborne, M.R. (1992) An effective method for computing regression quantiles. *IMA J. Numer. Anal.*, **12**, 151–166.

Pollard, D. (1991) Asymptotics for least absolute deviation regression estimation. *Econom. Theory*, **7**, 186–200.

Picek, J. (1997) Tests in linear models based on regression rank scores. PhD thesis (in Czech), Charles University, Prague.

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

Steel, R.G.D. and Torrie, J.H. (1960) *Principles and Procedures of Statistics*. New York: McGraw-Hill.

Received March 1996 and revised October 1997.