Minimum distance tests and estimates based on ranks

Authors: Radim Navrátil
– Department of Mathematics and Statistics, Masaryk University Brno, Czech Republic (navratil@math.muni.cz)

Abstract:

- It is well known that the least squares estimate in classical linear regression model is very sensitive to violation of the assumptions, in particular normality of model errors. That is why a lot of alternative estimates has been developed to overcome these shortcomings. Quite interesting class of such estimates is formed by R-estimates. They use only ranks of response variable instead of their actual value.

  The goal of this paper is to extend this class by another estimates and tests based only on ranks. First, we will introduce a new rank test in linear regression model. The test statistic is based on a certain minimum distance estimator, but unlike classical rank tests in regression it is not a simple linear rank statistic. Then, we will return back to estimates and generalize minimum distance estimates for various type of distances.

  We will show that in some situation these tests and estimates have greater power than the classical ones. Theoretical results will be accompanied by a simulation study to illustrate finite sample behavior of estimates and tests.

Key-Words:

- *Minimum distance estimates; Ranks; Robustness; Tests;*

AMS Subject Classification:

- 62G05, 62G10, 62G35.

1. INTRODUCTION

Consider the model of regression line

$$Y_i = \beta_0 + x_i\beta + e_i, \quad i = 1, \dots, n, \tag{1.1}$$

where β_0 and β are unknown parameters, $x_1, \dots, x_n$ are regressors, model errors $e_1, \dots, e_n$ are assumed to be i.i.d. with an unknown distribution function F and uniformly continuous density f. Our aim is to estimate the slope parameter β and test the hypothesis

$$\mathbf{H}_0 : \ \beta = 0 \quad \text{against} \quad \mathbf{K}_0 : \ \beta \neq 0.$$

There is a lot of methods described in the literature. Ordinary least squares estimate and the corresponding t-test which are optimal for normal model errors. Unfortunately, normality assumption is often in practice not satisfied. Its violation may cause that the estimate or test fails.

We do not put any assumptions on the shape of the distribution function F. Generally, F is unknown; therefore we should use a nonparametric approach. We will focus on rank tests and estimates that instead of original response variables Y_i's use their ranks.

Rank tests form a class of statistical procedures which have the advantage of simplicity combined with surprising power. Modern development of rank tests began in the 1930's, see e.g. [2] and [4]. Well known is also Wilcoxon [11] who introduced popular Wilcoxon test for comparing two treatments. At first, it was believed that a high price in loss of efficiency when using rank tests has to be paid. However, it turned out that efficiency of rank tests behaves quite well under the classical assumption of normality. In addition these tests remain valid and have high efficiency when the assumption of normality is not satisfied. These facts were first brought out by Pitman [8]. Recently rank tests have been still very popular and widely used, see [1] and [5].

Let us briefly show the classical approach based on linear rank statistic (see e.g. [3]). It was developed more than fifty years ago and it is still being used thanks to its simplicity and robustness. Denote

$$Q_n = \frac{1}{n}\sum_{i=1}^{n}(x_i - \overline{x})^2, \quad \text{with} \quad \overline{x} = \frac{1}{n}\sum_{i=1}^{n} x_i.$$

Let R_i be the rank of Y_i among $Y_1, \dots, Y_n$ and define linear rank statistic

$$S_n = \frac{1}{\sqrt{n}}\sum_{i=1}^{n}(x_i - \overline{x})\varphi\left(\frac{R_i}{n+1}\right)$$

for some nondecreasing, nonconstant, square integrable score function $\varphi : (0,1) \mapsto \mathbb{R}$. Test criterion for $\mathbf{H}_0$ is then

$$T_n^2 = \frac{S_n^2}{A^2(\varphi)Q_n}, \tag{1.2}$$

where

$$A^2(\varphi) = \int_0^1 (\varphi(t) - \overline{\varphi})^2 dt, \quad \overline{\varphi} = \int_0^1 \varphi(t)dt.$$

T_n^2 has under $\mathbf{H}_0$ asymptotically (under very mild conditions) χ^2 distribution with 1 degree of freedom.

Remark 1.1. The choice $\varphi(u) = u$, for $0 < u < 1$, leads to Wilcoxon rank test in regression. Hájek in [3] proved that such test in locally most powerful linear rank test for logistic model errors. In this case it has even greater power than t-test.

2. EMPIRICAL PROCESSES IN SIMPLE LINEAR REGRESSION

Koul [6] considered a class of estimates in linear regression model based on minimization of certain type of distances. Let us remind his approach. Define

$$T_{g,n}(s,t) = \frac{1}{\sqrt{n}} \sum_{i=1}^n g(x_i)\mathbb{I}\{R_{i,t} \le ns\}, \quad 0 \le s \le 1, \ t \in \mathbb{R} \tag{2.1}$$

$$K_{g,n}(t) = \int_0^1 T_{g,n}^2(s,t)dL(s), \quad t \in \mathbb{R}, \tag{2.2}$$

where $R_{i,t}$ is the rank of the residual $Y_i - x_i t$ among $Y_1 - x_1 t, \ldots, Y_n - x_n t$. L is a distribution function on $[0,1]$ and g a real (weight) function such that $\sum_{i=1}^n g(x_i) = 0$.

The minimum distance estimator $\widehat{\beta}_{g,n}$ is then defined as

$$\widehat{\beta}_{g,n} = \text{argmin}\{K_{g,n}(t) : t \in \mathbb{R}\}.$$

Koul [6] showed that such estimates might have in some situations greater efficiency then corresponding R-estimates and LSE respectively. He also proved their asymptotic unbiasedness and normality. We will develop his idea and introduce a class of test statistics based on these estimates. We will investigate their finite sample as well as asymptotic behavior. Finally, we will return back to the estimates, generalize them and show that some have greater efficiency then original Koul's estimates.

3. TEST IN SIMPLE LINEAR REGRESSION

Recall that we want to test whether regression is present, i.e. we test the null hypothesis

$$\mathbf{H}_0: \ \beta = 0 \quad \text{against} \quad \mathbf{K}_0: \ \beta \neq 0.$$

We put the hypothetical value $\beta = 0$ into (2.1) and (2.2) and get the test statistic

$$K_{g,n}(0) = K^*_{g,n} = \int_0^1 T^2_{g,n}(s,0)dL(s). \tag{3.1}$$

Discuss some computation aspects of (3.1). First, have a look at the formula (3.1) for $K^*_{g,n}$. Inserting (2.1) into (2.2) for $t = 0$ we have

$$\begin{aligned} K^*_{g,n} &= \frac{1}{n}\sum_{i=1}^{n}\sum_{j=1}^{n} g(x_i)g(x_j)\int_0^1 \mathbb{I}\{R_i \leq ns\}\mathbb{I}\{R_j \leq ns\}dL(s) \\ &= \frac{1}{n}\sum_{i=1}^{n}\sum_{j=1}^{n} g(x_i)g(x_j)\int_{\max\left\{\frac{R_i}{n},\frac{R_j}{n}\right\}}^1 1dL(s). \end{aligned}$$

L is a distribution function, hence $L(\max\{a,b\}) = \max\{L(a), L(b)\}$, it also remains true for limits from the left

$$K^*_{g,n} = \frac{1}{n}\sum_{i=1}^{n}\sum_{j=1}^{n} g(x_i)g(x_j)\left(1 - \max\left\{L\left(\frac{R_i}{n}-\right), L\left(\frac{R_j}{n}-\right)\right\}\right).$$

Since $\sum_{i=1}^n g(x_i) = 0$ we get

$$K^*_{g,n} = -\frac{1}{n}\sum_{i=1}^{n}\sum_{j=1}^{n} g(x_i)g(x_j)\max\left\{L\left(\frac{R_i}{n}-\right), L\left(\frac{R_j}{n}-\right)\right\}.$$

Using the fact

$$2\max\{a,b\} = a + b + |a-b|, \quad \forall\, a,b \in \mathbb{R}$$

and $\sum_{i=1}^n g(x_i) = 0$ we have

$$K^*_{g,n} = -\frac{1}{2n}\sum_{i=1}^{n}\sum_{j=1}^{n} g(x_i)g(x_j)\left|L\left(\frac{R_i}{n}-\right) - L\left(\frac{R_j}{n}-\right)\right|,$$

which is much more convenient for practical computations.

Since $K^*_{g,n}$ depends on Y_i's only through their ranks R_i's, it is a rank statistic. However, unlike the classical rank test statistic T^2_n defined in (1.2), $K^*_{g,n}$ is not a linear function of the ranks. That may cause some computation issues, but we can profit from its greater power in some situations.

Under $\mathbf{H}_0$ ($\beta = 0$) model (1.1) reduces to

$$Y_i = \beta_0 + e_i, \quad i = 1, \dots, n. \tag{3.2}$$

Since distribution of model errors e_i is absolutely continuous, there can be any ties in ranks with probability 0. Thanks to invariance of ranks with respect to the location, distribution of $R_1, \dots, R_n$ under null hypothesis is uniform over all $n!$ permutations of numbers $\{1, \dots, n\}$. Therefore distribution of $K^*_{g,n}$ given $x_1, \dots, x_n$ under $\mathbf{H}_0$ is distribution-free and may be even computed directly. To do it, we have to compute all values of the test statistic $K^*_{g,n}$ for each of $n!$ permutations of numbers $\{1, \dots, n\}$. From there we can get $(1-\alpha)$-quantile or the corresponding p-value.

However, for large sample size n computation of exact (conditional) distribution may be time consuming, that is why we will investigate asymptotic distribution of $K^*_{g,n}$.

For $s \in [0, 1]$ define empirical processes

$$\widehat{V}_{g,n}(s) = \frac{1}{\sqrt{n}} \sum_{i=1}^{n} g(x_i) \mathbb{I}\{e_i \le F_n^{-1}(s)\},$$

$$V_{g,n}(s) = \frac{1}{\sqrt{n}} \sum_{i=1}^{n} g(x_i) \mathbb{I}\{e_i \le F^{-1}(s)\},$$

where $F_n(s) = \frac{1}{n} \sum_{i=1}^{n} \mathbb{I}\{e_i \le s)\}$ is empirical distribution function.

Now, state assumptions needed for proofs of asymptotic properties of $K^*_{g,n}$. Note that all limits are considered as $n \to \infty$:

$$x_1, \dots, x_n \text{ are not all equal,} \tag{3.3}$$

$$\max_{i=1,\dots,n} \frac{(x_i - \bar{x})^2}{\sum_{j=1}^{n} (x_j - \bar{x})^2} \to 0, \tag{3.4}$$

$$g(x_i) \ne 0 \text{ for some } \quad i = 1, \dots, n, \tag{3.5}$$

$$\text{there exists } \alpha_1 > 0, \text{ such that } \quad \frac{1}{n} \sum_{i=1}^{n} g(x_i)(x_i - \overline{x}) \to \alpha_1, \tag{3.6}$$

$$\max_{i=1,\dots,n} g^2(x_i) \to 0, \tag{3.7}$$

$$\sup_{n \in \mathbb{N}} \max_{i=1,\dots,n} |g(x_i)| \le c \quad \text{for some} \quad 0 < c < \infty, \tag{3.8}$$

$$\text{there exists } \gamma^2 > 0, \text{ such that } \quad \frac{1}{n} \sum_{i=1}^{n} g^2(x_i) \to \gamma^2. \tag{3.9}$$

Remark 3.1. Assumptions (3.3) and (3.4) state that the design points $x_1, \dots x_n$ are well-defined. Remaining assumptions put conditions on the g function. If there exists a limit $\lim_{n \to \infty} \frac{1}{n} \sum_{i=1}^{n} (x_i - \overline{x})^2$, then the natural choice $g(x_i) = x_i - \overline{x}$ meets the above assumptions.

Lemma 3.1. *Under (3.3) – (3.6) it holds*

$$\Big|K^*_{g,n} - \int \widehat{V}^2_{g,n}(s)dL(s)\Big| = o_p(1), \quad \text{as} \quad n \to \infty.$$

Proof: For convenience we will drop off an index g in $K^*_{g,n}$ and $\widehat{V}_{g,n}$. Adding and subtracting $\widehat{V}_n(s)$ in the first integral, squaring and using Cauchy-Schwarz inequality we get

$$\begin{aligned}
&\Big|\int T^2_n(s)dL(s) - \int \widehat{V}^2_n(s)dL(s)\Big| \\
&= \Big|\int \Big[T_n(s) - \widehat{V}_n(s)\Big]^2 dL(s) + 2\int \widehat{V}_n(s)(T_n(s) - \widehat{V}_n(s))dL(s)\Big| \\
&\leq \sup_{0\leq s\leq 1} |T_n(s) - \widehat{V}_n(s)|^2 + 2\sqrt{\int \widehat{V}^2_n(s)dL(s)\int (T_n(s) - \widehat{V}_n(s))^2 dL(s)}.
\end{aligned}$$

The fact

$$\sup_{0\leq s\leq 1} |T_n(s) - \widehat{V}_n(s)| \leq 2 \max_{i=1,\dots,n} |g(x_i)| = o_p(1)$$

together with $\int \widehat{V}^2_n(s)dL(s) = O_p(1)$ proves the Lemma. □

Lemma 3.2. *Under (3.3) – (3.6) it holds*

$$\Big|K^*_{g,n} - \int V^2_{g,n}(s)dL(s)\Big| = o_p(1), \quad \text{as} \quad n \to \infty.$$

Proof:

$$\begin{aligned}
&\Big|\int T^2_n(s)dL(s) - \int V^2_n(s)dL(s)\Big| \\
&= \Big|\int [T_n(s) - V_n(s)]^2 dL(s) + 2\int V_n(s)(T_n(s) - V_n(s))dL(s)\Big|. \qquad (3.10)
\end{aligned}$$

Using Minkowski inequality

$$\begin{aligned}
&\int [T_n(s) - V_n(s)]^2 dL(s) = \int [T_n(s) - \widehat{V}_n(s) + \widehat{V}_n(s) - V_n(s)]^2 dL(s) \\
&\leq 2\int [T_n(s) - \widehat{V}_n(s)]^2 dL(s) + 2\int [\widehat{V}_n(s) - V_n(s)]^2 dL(s). \qquad (3.11)
\end{aligned}$$

By Cauchy-Schwarz inequality

$$\begin{aligned}
\Big|\int V_n(s)(T_n(s) - V_n(s))dL(s)\Big| &\leq \sqrt{\int V^2_n(s)dL(s)\int [T_n(s) - V_n(s)]^2 dL(s)} \\
&= o_p(1), \qquad (3.12)
\end{aligned}$$

because $\int V_n^2(s)dL(s) = O_p(1)$ and $\int [T_n(s) - V_n(s)]^2 dL(s) = o_p(1)$.

Observe that

$$V_n(FF_n^{-1}(s)) = \sum_{i=1}^n g(x_i)\mathbb{I}\{e_i \le F^{-1}FF_n^{-1}(s)\} = \sum_{i=1}^n g(x_i)\mathbb{I}\{e_i \le F_n^{-1}(s)\} = \widehat{V}_n(s).$$

Therefore

$$\sup_{0\le s\le 1} |\widehat{V}_n(s) - V_n(s)| = \sup_{0\le s\le 1} |V_n(FF_n^{-1}(s)) - V_n(s)| = o_p(1),$$

because

$$\begin{aligned}
&\sup_{0\le s\le 1} |FF_n^{-1}(s) - s| = \sup_{0\le s\le 1} |FF^{-1}(s) - F_nF_n^{-1}(s) + F_nF_n^{-1}(s) - s| \\
&\le \sup_{x\in\mathbb{R}} |F(x) - F_n(x)| + \sup_{0\le s\le 1} |F_nF_n^{-1}(s) - s| = o_p(1).
\end{aligned}$$

Now, combining previous result, Lemma 3.1 and (3.10), (3.11) and (3.12) we have proven the Lemma. □

Remark 3.2. The previous lemma states that the asymptotic distribution of $K^*_{g,n}$ is the same as $\int V^2_{g,n}(s)dL(s)$ that is easier to investigate. Now, we are able to state the theorem about asymptotic null distribution of $K^*_{g,n}$.

Theorem 3.1. *Under (3.3) – (3.9) in model (1.1) under* $\mathbf{H}_0$

$$K^*_{g,n} \xrightarrow{d} \gamma^2 \cdot Y_L, \quad \textit{with} \quad Y_L = \int_0^1 B^2(s)dL(s),$$

where $B(s)$ is a Brownian bridge in $\mathcal{C}[0,1]$.

Proof: Recall that

$$\begin{aligned}
V_{g,n}(s) &= \frac{1}{\sqrt{n}} \sum_{i=1}^n g(x_i)\mathbb{I}\{e_i \le F^{-1}(s)\} = \frac{1}{\sqrt{n}} \sum_{i=1}^n g(x_i)\mathbb{I}\{F(e_i) \le s\} \\
&= \frac{1}{\sqrt{n}} \sum_{i=1}^n g(x_i)\mathbb{I}\{U_i \le (s)\},
\end{aligned}$$

where $U_1, \ldots, U_n$ are i.i.d. random variables with uniform $\mathcal{U}(0,1)$ distribution.

By [6] we have

$$V_{g,n}(s) \Rightarrow \gamma \cdot B(s) \quad \text{in} \quad \mathcal{D}[0,1]$$

and therefore $\int V^2_{g,n}(s)dL(s) \xrightarrow{d} \gamma^2 \int B^2(s)dL(s)$. That together with Lemma 3.2 proves Theorem 3.1. □

α	0.90	0.95	0.99	0.999
α - quantile	0.34730	0.46136	0.74346	1.16786

Table 1: Quantiles of distribution Y_L for $L(s) = s$.

Distribution of random variable Y_L for $L(s) = s$ was first investigated by Smirnov [9]. Values of its distribution function may be found for example in [10], some quantiles are listed in Table 1. For other choices of function L one has to use simulated values.

In [7] we also investigated the behavior of $K^*_{g,n}$ under the local alternative

$$\mathbf{K}_{0,n}: \ \beta = n^{-1/2}\beta^*, \quad 0 \neq \beta^* \in \mathbb{R} \text{ fixed.}$$

The resulting distribution cannot be expressed in a closed formula, that is why we omit it here. Power of the test will be illustrated later in the simulation study.

4. GENERALIZATION OF THE TEST

In the definition of the test statistic $K^*_{g,n}$ (3.1) we used second power of the L^2-norm of the empirical process $T_{g,n}(s, 0)$. Instead, we may use any norm on $\mathcal{D}[0, 1]$. For simplicity, we will consider only the class of L^p-norms for $p \in [1, \infty]$.

For $p \in [1, \infty)$ define

$$K^{(p)}_{g,n} = \left(\int_0^1 |T_{g,n}(s, 0)|^p dL(s) \right)^{\frac{1}{p}}, \tag{4.1}$$

for $p = \infty$ define

$$K^{(\infty)}_{g,n} = \max\{|T_{g,n}(s, 0)| : s \in [0, 1]\}. \tag{4.2}$$

Remark 4.1. Obviously, for $p = 2$ we have $\left(K^{(2)}_{g,n}\right)^2 = K^*_{g,n}$.

From computation point of view, formulas (4.1) and (4.2) might be simplified. Obviously, $T_{g,n}(0, 0) = 0$ and $T_{g,n}(s, 0)$ is piecewise constant:

$$T_{g,n}(s, 0) = \frac{1}{\sqrt{n}} \sum_{i:R_i \leq j} g(x_i), \quad \frac{j-1}{n} < s \leq \frac{j}{n}, \ j = 1, \dots, n.$$

Therefore,

$$K_{g,n}^{(\infty)} = \frac{1}{\sqrt{n}} \max_{i=1,\dots,n} \left| \sum_{i:R_i \le j} g(x_i) \right|,$$

$$K_{g,n}^{(1)} = \frac{1}{n^{3/2}} \sum_{i=1}^{n} \left| \sum_{i:R_i \le j} g(x_i) \right|,$$

for $L(s) = s$.

Again, since $K_{g,n}^{(p)}$ depends on Y_i's only through their ranks R_i's, it is a rank statistic, but not linear like (1.2). That may cause some computation issues, but we can profit from its greater power in some situations.

Now, focus on the distribution under the null hypothesis. Under $\mathbf{H}_0$ ($\beta = 0$) model (1.1) reduces to (3.2)

$$Y_i = \beta_0 + e_i, \quad i = 1, \dots, n.$$

Thanks to the same arguments as in the previous section, the distribution of $K_{g,n}^{(p)}$ given $x_1, \dots, x_n$ under the null hypothesis is distribution-free and can be easily computed directly the same way. For large sample sizes n the following asymptotic approximation might be used.

Theorem 4.1. *Under (3.3) – (3.9) in model (1.1) under* $\mathbf{H}_0$

$$K_{g,n}^{(p)} \xrightarrow{d} \gamma \cdot Y_L^{(p)}, \quad \textit{with} \quad Y_L^{(p)} = \left(\int_0^1 |B(s)|^p dL(s) \right)^{\frac{1}{p}},$$

where $B(s)$ is a Brownian bridge in $\mathcal{C}[0,1]$.

Proof: The proof is analogous to the proof of Theorem 3.1. □

5. CHOICE OF THE PARAMETERS IN PRACTISE

In the previous section, we derived a class of minimum distance tests. In practise there arises a question how to choose optional parameters of the test.

Function g is in fact a weight function for regressors, so it can downweight outlying observations to robustify these tests against extreme values of x_i (if g is bounded for example). Anyway, if we are not afraid of leverage observations x_i, then the optimal choice of the g function is $g(x_i) = x_i - \overline{x}$. This choice leads to the test with the greatest power among all test with different g functions.

Function L has similar interpretation as score-function φ in standard rank tests theory, optimal L could be chosen based on the estimate of unknown model errors. Anyway, the simplest choice $L(s) = s$ gives very reasonable results (see the simulations).

And finally, the choice of L^p-norm depends on the model errors e_i. From computational point of view, one should consider $p = 1, 2, \infty$ for that we have a simple formula. Power comparisons are made in the simulation study.

6. GENERALIZATIONS

In [7] we investigated behavior the test in measurement error model:

$$\begin{aligned} Y_i &= \beta_0 + \beta x_i + e_i, \\ w_i &= x_i + v_i, \quad i = 1, \ldots, n, \end{aligned}$$

where instead of actual regressors x_i we observed w_i affected by measurement errors v_i.

We showed that the test is still valid in this model, the presence of measurement errors decreases power of the test, because we do not use values of function g in optimal points $x_1, \ldots, x_n$ but in w_i's.

In Section 4 we showed extension of the test using various norms for the empirical process. Analogously, we may define generalization of Koul's estimate defined in Section 2.

Consider empirical process $T_{g,n}(s,t)$ defined in (2.1) and for $p \in [1, \infty)$ define

$$K^{(p)}_{g,n}(t) = \left(\int_0^1 |T_{g,n}(s,t)|^p dL(s) \right)^{\frac{1}{p}}, \quad t \in \mathbb{R}$$

and for $p = \infty$

$$K^{(\infty)}_{g,n}(t) = \max\{|T_{g,n}(s,t)| : s \in [0,1]\}, \quad t \in \mathbb{R}.$$

Minimum distance estimator $\widehat{\beta}^{(p)}_{g,n}$ is then defined as

$$\widehat{\beta}^{(p)}_{g,n} = \operatorname{argmin}\{K^{(p)}_{g,n}(t) : t \in \mathbb{R}\}.$$

In the similar way, thanks to duality of rank tests and estimates, we may show favorable properties and good performance of the estimates. Detailed analysis will be part of our future study.

7. SIMULATIONS

To support previous theoretical results we conducted a large simulation study, let us present a few interesting results. Let us start with model (1.1) for moderate sample size $n = 30$. We have compared empirical power of our test based on the test statistic $K^*_{g,n}$ with $g(x_i) = x_i - \bar{x}$ and $L(s) = s$ (call it *minimum distance test*) with Wilcoxon test for regression (based on (1.2) with $\varphi(u) = u$) and standard t-test for regression.

Regressors $x_1, \ldots, x_{30}$ were once generated from uniform $\mathcal{U}(-2, 10)$ distribution and then considered fixed, model errors e_i were generated from normal, logistic, Laplace and t-distribution with 6 degrees of freedom, respectively, always with 0 mean a variance 3/2. The empirical powers of the tests were computed as a percentage of rejections of $\mathbf{H}_0$ among 10 000 replications, at significance level $\alpha = 0.05$. The results are summarized in Table 2.

$\beta \setminus e_i$	$\mathcal{N}\left(0, \frac{3}{2}\right)$			$Log\left(0, \frac{\sqrt{2}\pi}{3}\right)$			$Lap\left(0, \frac{\sqrt{3}}{2}\right)$			$t(6)$		
	MD	W	t	MD	W	t	MD	W	t	MD	W	t
0	4.98	4.42	5.00	5.06	4.55	5.00	5.00	4.55	5.04	5.00	4.32	4.93
0.1	28.7	28.3	31.5	32.7	31.4	32.0	42.4	39.0	33.5	34.6	33.1	32.9
−0.1	28.3	28.2	30.9	32.7	31.2	32.2	42.5	39.0	33.7	33.3	32.1	31.9
0.2	78.2	78.8	82.3	82.5	81.8	81.9	88.3	86.6	82.0	84.5	83.9	82.6
−0.2	78.3	78.7	82.9	83.3	82.7	82.9	89.2	87.5	83.1	84.0	83.4	82.5

Table 2: Percentage of rejections of hypothesis $\mathbf{H}_0: \beta = 0$ of minimum distance test (MD), Wilcoxon test for regression (W) and t-test for regression (t); $n = 30$, $\alpha = 0.05$.

For normal model errors t-test achieves (not surprisingly) the largest power, but the differences among the three tests are not much distinct. For distributions with heavier tails than normal our test has the largest power, even for logistic distribution (for which Wilcoxon test is locally most powerful rank test). It is caused by the slow convergence of Wilcoxon test statistic to its asymptotic distribution.

In Table 3 comparison of tests based on various norms (L^2, L^1, L^∞) is made.

Bad performance of the test based on L^∞-norm is caused by slow convergence of corresponding test statistic to its limit distribution. For large sample size n test preserves prescribed significance level α under null hypothesis and

e_i	$\mathcal{N}\left(0, \frac{3}{2}\right)$			$Log\left(0, \frac{3}{\sqrt{2}\pi}\right)$			$Lap\left(0, \frac{\sqrt{3}}{2}\right)$			$t(6)$		
β	L^2	L^1	L^∞	L^2	L^1	L^∞	L^2	L^1	L^∞	L^2	L^1	L^∞
0	5.15	5.03	2.73	5.34	5.31	2.61	5.03	5.18	2.71	4.90	4.97	2.64
0.1	30.0	31.0	17.8	33.4	34.1	20.6	43.6	43.0	30.6	35.8	36.6	22.4
−0.1	29.2	30.2	17.2	34.0	34.7	21.4	44.2	43.4	30.4	35.0	35.8	21.6
0.2	81.2	82.6	62.4	84.9	85.7	68.1	90.5	90.3	78.9	87.8	88.5	72.2
−0.2	80.6	82.6	62.2	84.9	86.0	68.8	90.5	90.4	78.9	86.8	87.4	71.2

Table 3: Percentage of rejections of hypothesis $\mathbf{H}_0 : \beta = 0$ of minimum distance test based on L^2, L^1 and L^∞-norm; $n = 30$, $\alpha = 0.05$.

under the alternative its power is quite similar to other tests. Tests based on L^2 and L^1-norm perform very similar. Test based on L^1-norm might have slightly greater power which is caused by faster convergence of the test statistic. On the other hand, computation of the test statistic based on L^2-norm is easier than those with L^1-norm.

We performed more simulations for various design points x_i, sample sizes n and model errors e_i. We also compared the tests according to the choice of functions L and g. However, the corresponding results are very similar to those presented in Tables 2 and 3.

Finally, we studied the finite sample behavior of generalized estimates from Section 6. Because of the duality of rank tests and estimates corresponding results and conclusions were the same as for the tests. That is why we omit it here.

8. CONCLUSIONS

We introduced a class of new rank tests in linear regression model. Unlike the classical ones introduced by Hájek and Šidák, our tests are not linear functions of the ranks. Thanks to that they can achieve greater power. Our tests are robust, we do not need to assume normality of model errors. Anyway, under normality our tests has similar power as classical t-test; for model errors with heavy tails our test has significantly greater power.

Our test may be also robust with respect to leverage observations. The right choice of the weight function leads to the test that is not sensitive to outlying regressors. We also generalized Koul's minimum distance estimates when considering various L^p-norms instead of L^2. Corresponding estimates have the same favorable properties as the tests.

ACKNOWLEDGMENTS

The author's research was supported by Student Project Grant at MU (specific research, rector's programme) MUNI/A/1204/2017.

REFERENCES

[1] Feng, L.; Zou, C. and Wang, Z. (2012). Rank-based inference for the single-index model. *Statistics & Probability Letters*, **82**, 535–541.

[2] Friedman, M. (1937). The use of ranks to avoid the assumption of normality implicit in the analysis of variance, *J. Amer. Statist. Assoc.*, **32**, 675–701.

[3] Hájek, J.; Šidák, Z. and Sen, P.K. (1999). *Theory of Rank Tests*, Academic Press, New York.

[4] Hotteling, H. and Pabst, M.R. (1936). Rank correlation and tests of significance involving no assumptions of normality, *Ann. Math. Statist.*, **7**, 29–43.

[5] Jurečková, J; Sen, P.K. and Picek, J. (2013). *Metodological tools in robust and nonparametric statistics*, Chapman & Hall/CRC Press, Boca Raton, London.

[6] Koul, H.L. (2002). *Weighted empirical processes in dynamic nonlinear models*, Springer, New York.

[7] Navrátil, R. (2015). Rank tests in regression model based on minimum distance estimates, *Kybernetika : The Journal of the Czech Society for Cybernetics and Informatics*, **51**, 6, 909–922.

[8] Pitman, E.J.G. (1948). *Lecture notes on nonparametric statistics*, Columbia University, New York.

[9] Smirnov, N.V. (1936). Sur la distribution de ω^2 (criterium de m. r. v. mises), *C. R. Akad. Sci. Paris*, **202**, 449–452.

[10] Tolmatz, L. (2002). On the distribution of the square integral of the brownian bridge, *The Annals of Probability*, **30**, 253–269.

[11] Wilcoxon, F. (1945). Individual comparisons by ranking methods. *Biometrics*, **1**, 80–83.