Rank scores tests in measurement error models

Jan Picek(1) and Jana Jurečková(2)
(1) Technical University of Liberec, (2) Charles University, Prague, Czech Republic
contact: jan.picek@tul.cz

The rank and regression rank score tests of linear hypothesis in the linear model are defined for measurement error models. The modified tests are still distribution free. Some tests of linear subhypotheses are invariant to the nuisance parameter, others are based on the aligned ranks using the R-estimators. The efficiency changes caused by the measurement errors is also illustrated by a simulation study.

Introduction

Consider the linear regression model

\[ Y_i = \beta_0 + \beta_1 x_i + \epsilon_i, \quad i = 1, \ldots, n, \]

with unknown parameters \( \beta_0 \in \mathbb{R}, \beta_1 \in \mathbb{R}, \delta \in \mathbb{R} \).

We want to test the hypotheses:

- \( H_0: (\beta_1, \delta) = (0, 0) \), nuisance,
- \( H_1: \beta_1, \delta \) nuisance,
- \( H_2: \delta = 0 \), nuisance.

Consider the situation that the regressors \( x_i \) are deterministic, but are affected by the random errors, so that instead of \( x_i \) we measure \( w_i = x_i + \epsilon_i, \quad i = 1, \ldots, n, \) where \( \epsilon_1, \ldots, \epsilon_n \) are \( p \)-dimensional random errors, identically distributed with an unknown distribution function \( G \).

Linear rank test of regression of hypothesis \( H_{n1} \)

Without loss of generality we can consider the model \( Y_i = \beta_1 + X_{ni}\beta + e \) Jurečková et al. (2010) proposed the test of \( H_0 \) based on the criterion

\[ \hat{T}_n = (A(\hat{\beta}))^{-1} S_n(Q_n)^{-1} S_n, \]

where

\[ \hat{S}_n = n^{-1/2} \sum_{i=1}^n (w_i - \hat{w}_i) a_i(R_n) \]

and

\[ Q_n = \sum_{i=1}^n (w_i - w_i) (w_i - \bar{w})^\top. \]

The scores \( \hat{e}_i \) are generated by \( \varphi \) based on the regression rank scores. Under \( H_{n0} \) \( \hat{T}_n \) has asymptotic \( \chi^2 \) distribution with \( q \) degrees of freedom.

Aligned rank tests of hypothesis \( H_{n0} \)

Unfortunately, the parallel test based on regression rank scores is not asymptotically distribution free. Replace the nuisance slope parameter \( \beta \) with an estimator \( \hat{\beta} \), and then construct the test based on aligned ranks of the residuals.

\[ \hat{\beta} = \arg \min \{ ||L_n(b)|| : b \in \mathbb{R}^q \} \]

where \( || \cdot || \) can be the \( L_1, L_2 \) norms, or eventually the sup-norm. The vector of aligned rank statistics

\[ L_n(b) = n^{-1/2} \sum_{i=1}^n (w_i - w_i)b(R_n(b)) \]

where \( R_n(b) \) stands for the aligned rank of \( Y_i - w_i b \) among \( Y_1 - w_1 b, \ldots, Y_n - w_n b \). \( b \in \mathbb{R}^q \).

The test criterion for the hypothesis \( H_{n0} \):

\[ T_n = (A(\hat{\beta}))^{-1} \left[ S_n(\hat{\beta})^\top C_n S_n(\hat{\beta}) \right] \]

where

\[ S_n(\hat{\beta}) = n^{-1/2} \sum_{i=1}^n (z_i - \bar{z}_{\hat{\beta}}) a_i(R_n(\hat{\beta})) \]

and

\[ C_n = n^{-1} \sum_{i=1}^n (z_i - \bar{z}_{\hat{\beta}}) (z_i - \bar{z}_{\hat{\beta}})^\top. \]

Under \( H_{n0} \) the asymptotic distribution of \( T_n \) is central \( \chi^2 \) with \( q \) degrees of freedom.

Regression rank score tests of hypothesis \( H_{n2} \)

In the model without measurement errors, Gutenbrunner et al. (1993) constructed a class of tests of \( H_0 \) based on regression rank scores, that are defined as the vector

\[ a_i(\tau) = (\bar{a}_i(\tau), \ldots, \bar{a}_{ni}(\tau))^{\top} \]

of solutions of the parametric linear programming problem

\[ \sum_{j=1}^q a_j(x_j - \bar{a}_j(\tau)) = \max \]

subject to

\[ \sum_{j=1}^q a_j = (1 - \tau)n, \]

\[ \sum_{j=1}^q z_j a_j(x_j - \bar{a}_j(\tau)) = (1 - \tau) \sum_{j=1}^q z_j, \quad j = 1, \ldots, q, \]

\[ a_j(x_j - \bar{a}_j(\tau)) \in [0,1]^q, \quad 0 \leq \tau \leq 1. \]

The test criterion \( T_n \) is:

\[ T_n = (A(\hat{\beta}))^{-1} \sum_{i=1}^n (Q_n - \hat{S}_i)^{-1} \hat{S}_i, \]

where

\[ Q_n = n^{-1/2}(W_n - W_n)(Z_n - \hat{Z}_n), \]

and

\[ \hat{S}_i = n^{-1/2}(W_n_i - W_n_i) \hat{b}_i, \]

with \( W_n = H_i W_n - \) the projection on the space spanned by the columns of \( X_n \).

The scores \( \bar{b}_i \) are generated by \( \varphi \) based on the regression rank scores. Under \( H_{n2} \) \( T_n \) has asymptotic \( \chi^2 \) distribution with \( q \) degrees of freedom.


case of \( \hat{\beta} \):

- \( \hat{\beta} = \arg \min \{ \| L_n(b) \| : b \in \mathbb{R}^q \} \)

where \( \| \cdot \| \) can be the \( L_1, L_2 \) norms, or eventually the sup-norm. The vector of aligned rank statistics

\[ L_n(b) = n^{-1/2} \sum_{i=1}^n (w_i - w_i)b(R_n(b)) \]

where \( R_n(b) \) stands for the aligned rank of \( Y_i - w_i b \) among \( Y_1 - w_1 b, \ldots, Y_n - w_n b \). \( b \in \mathbb{R}^q \).

The test criterion for the hypothesis \( H_{n0} \):

\[ T_n = (A(\hat{\beta}))^{-1} \left[ S_n(\hat{\beta})^\top C_n S_n(\hat{\beta}) \right] \]

where

\[ S_n(\hat{\beta}) = n^{-1/2} \sum_{i=1}^n (z_i - \bar{z}_{\hat{\beta}}) a_i(R_n(\hat{\beta})) \]

and

\[ C_n = n^{-1} \sum_{i=1}^n (z_i - \bar{z}_{\hat{\beta}}) (z_i - \bar{z}_{\hat{\beta}})^\top. \]

Under \( H_{n0} \) the asymptotic distribution of \( T_n \) is central \( \chi^2 \) with \( q \) degrees of freedom.

Conclusion

The results show a good performance of the linear, regression rank score and aligned rank tests under the measurement error models. The measurement errors do not seriously affect the power even for a small sample size, provided their variance is not too large. The power of the test mainly depends on the distribution of the model errors \( \epsilon_i, i = 1, \ldots, n \).

References


Acknowledgement: Research was supported by the Grant GAČR 201/09/1031 and The Project LO1424.