Applied Mathematics, 2022, 13, 672-682
https://www.scirp.org/journal/am
ISSN Online: 2152-7393
ISSN Print: 2152-7385

Goodness of Fit Test Based on BLUS Residuals for Error Distribution of Regression Model

Jianxin Zhao[1], Xinmin Li[2]

[1]Navy Submarine Academy, Qingdao, China
[2]School of Mathematics and Statistics, Qingdao University, Qingdao, China
Email: bit_zjx@163.com

How to cite this paper: Zhao, J.X. and Li, X.M. (2022) Goodness of Fit Test Based on BLUS Residuals for Error Distribution of Regression Model. *Applied Mathematics*, **13**, 672-682.
https://doi.org/10.4236/am.2022.138042

Received: July 14, 2022
Accepted: August 20, 2022
Published: August 23, 2022

Abstract

The error distribution testing plays an important role in linear regression as distribution misspecification seriously affects the validity and efficiency of regression analysis. The least squares (OLS) residuals are often used to construct test statistics; in order to overcome the non-independent and identically residuals, the best linear unbiased scale (BLUS) residuals are applied in this paper, which, unlike OLS residuals, the residuals vector is identically and independently distributed. Based on the BLUS residuals, a new test statistic is constructed by using the sample random distance between sample quantile and quasi sample quantile derived from the null distribution, and the goodness-of-fit test of error distribution in the linear model is studied. The powers of the new tests under certain alternatives are examined. They are more powerful tests for the hypotheses concerned.

Keywords

Goodness of Fit, Cramér-Von Mises Statistic, Kolmogorov-Smirnov Statistic, Anderson-Darling Statistic, Sample Quantiles, Stochastic Sample Quantiles

1. Introduction

Parametric and nonparametric regression models are widely applied to the fields such as biology, chemistry and economics. The general form is often written as

$$Y = m(x) + \varepsilon, \tag{1}$$

where m is the regression function, the error ε satisfies $\mathrm{E}(\varepsilon) = 0$ and $\mathrm{E}(\varepsilon^2) > 0$, x are explanatory variables. The following general assumption is that the errors corresponding to different observations are independent and identically distributed. The most popular model is the linear model, that is, $m(\boldsymbol{x})=\boldsymbol{x}^{\mathrm{T}}\boldsymbol{\beta}$, $\boldsymbol{\beta}$ assumed as unknown parameters, and the error terms of the model are distributed according to the normal distribution. Under the assumption, the linear model has been attracting practical and theoretical workers because of its simplicity and validity.

Obviously, it is important to test whether the error is distributed as the assumed distribution before using the linear model to analyze the data under a certain assumption of the error distribution. On the other hand, even in general nonparametric models, additional knowledge of the distribution of errors can also improve the effectiveness of statistical analysis. For example, under the assumption of normal error, accurate or optimal tests can be obtained in many cases. A typical example is the goodness-of-fit test of regression function (See ref. [1] [2] [3]).

From the above analysis and the existing literature, it is easy to see that there are two kinds of important goodness of fit tests to the model. One is the goodness-of-fit test of the error distribution; the other is about the goodness of fit test of the regression function. That is to test the hypothesis:

$$H:\varepsilon\sim N\left(0,\sigma^2\right) \tag{2}$$

and

$$H:m(x)\in\mathcal{M}, \tag{3}$$

where $\mathcal{M}$ is a class of regression functions with certain properties.

There are many kinds of literatures about the two kinds of tests mentioned above, especially about the second kind of tests. For example, Of the first kind of test are ref. [4] [5] [6], etc., the second kind of test are ref. [7]-[17], etc.

For Linear regression models, there are usually two kinds of residual construction tests. One is based on the ordinary least square (OLS) residuals, under the normal assumption, the residual vector of OLS has a singular normal distribution, and the components of a vector are no longer independently identically distributed. The other is best linear unbiased scale residuals (BLUS) (See ref. [18]). It is a kind of residuum given by Theil in consideration of such a fact where the residuals are neither independent nor identical, even if the error distribution is independent and identical. Using OLS residuals directly to test the sequence independence or the same variance is impossible (See ref. [19]). However, the BLUS residual vector, under normal assumptions, is different from the OLS residual. There is a non-singular normal distribution, and the components of the vector are identically and independently distributed. It is a fact that most of the existing literature uses the OLS residuals to construct the test. But as you can see from the above analysis, using residuals as a new sample, it is not natural to construct a test with the existing test, such as the Shapiro-Wilk test, because the test of the original structure is based on independent samples from the same distribution, at the same time, it was proved that the skewness and Kurtosis of OLS residuals cannot exceed the skewness and Kurtosis of the error term (See ref. [20]). So, if the distribution of error terms is not normal, the distribution of OLS residuals is always close to the normal form, not the probability distribution of error terms. This shows that when the null hypothesis is not true, any normal test using the OLS residuals directly seems to have a tendency not to reject the null hypothesis. In view of the above analysis, this paper uses BLUS residuals to construct test statistics.

This paper is organized as follows. Section 2 constructs the new tests. Power comparisons are given in Section 3. Simulated powers for the new statistics are tabulated in this section. Section 4 gives some comments.

2. Test Statistics Based on BLUS Residuals

Let's have a linear model as follows:

$$\boldsymbol{y} = \boldsymbol{X}\boldsymbol{\beta} + \boldsymbol{\varepsilon}, \mathrm{E}\boldsymbol{\varepsilon} = \boldsymbol{0}, \mathrm{E}\boldsymbol{\varepsilon}\boldsymbol{\varepsilon}^{\mathrm{T}} = \sigma^2 \boldsymbol{I}, \tag{4}$$

where $\boldsymbol{y}$ is a $n\times 1$ random vector, $\boldsymbol{X}$ is a non-random $n\times k$ matrix with a known rank of *k*, $\boldsymbol{\beta}$ is a $k\times 1$ unknown parametric vector. For the above model, the Normality test of the error term distribution is important. Many scholars have studied this problem. The early literature includes ref. [20] [21] and so on. For the past twenty years, the two papers (See ref. [22] [23]) are based on Shapiro & Wilk test statistics (See ref. [24]) to construct tests using OLS residuals. Furthermore, ref. [25] constructs test statistics by the normalized residuals based on Shapiro & Wilk test statistics. While ref. [26] using Bootstrap method and based on the process of empirical residuals, constructs a new test statistic by using KS and AD statistics. Recently, ref. [27] proposed several tests based on partial sums of residuals where the test statistics are based on sums of a subset of the (ordered and standardized) residuals (Also see ref. [28] [29]).

For a linear model (4), the least square estimate of the Regression Coefficient is $\hat{\boldsymbol{\beta}} = \left(\boldsymbol{X}^{\mathrm{T}}\boldsymbol{X}\right)^{-1}\boldsymbol{X}^{\mathrm{T}}\boldsymbol{y}$, using $\boldsymbol{e}$ to represent the OLS residuals associated with the error vector $\boldsymbol{\varepsilon}$, we have

$$\boldsymbol{e} = \boldsymbol{y} - \boldsymbol{X}\hat{\boldsymbol{\beta}} = \left(\boldsymbol{I} - \boldsymbol{X}\left(\boldsymbol{X}^{\mathrm{T}}\boldsymbol{X}\right)^{-1}\boldsymbol{X}^{\mathrm{T}}\right)\boldsymbol{y} = \boldsymbol{M}\boldsymbol{y} = \boldsymbol{M}\boldsymbol{\varepsilon}. \tag{5}$$

And from that, $\mathrm{E}\boldsymbol{\varepsilon} = \boldsymbol{0}, \mathrm{E}\boldsymbol{\varepsilon}\boldsymbol{\varepsilon}^{\mathrm{T}} = \sigma^2\boldsymbol{M}$.

According to Theil (See ref. [18]), the BLUS residuals are obtained by the following steps:

First, select the smallest K elements in the main diagonal elements of the matrix $\boldsymbol{M}$ and rearrange the observed value Y according to the position of the row in which the K elements are located. Might as well be, place them in the first K position (See ref. [30]). The original model is then divided into blocks:

$$\begin{pmatrix} \boldsymbol{y}_0 \\ \boldsymbol{y}_1 \end{pmatrix} = \begin{pmatrix} \boldsymbol{X}_0 \\ \boldsymbol{X}_1 \end{pmatrix}\boldsymbol{\beta} + \begin{pmatrix} \boldsymbol{\varepsilon}_0 \\ \boldsymbol{\varepsilon}_1 \end{pmatrix} \tag{6}$$

$$\boldsymbol{M} = \begin{bmatrix} \boldsymbol{I} - \boldsymbol{X}_0\left(\boldsymbol{X}^{\mathrm{T}}\boldsymbol{X}\right)^{-1}\boldsymbol{X}_0^{\mathrm{T}} & -\boldsymbol{X}_0\left(\boldsymbol{X}^{\mathrm{T}}\boldsymbol{X}\right)^{-1}\boldsymbol{X}_1^{\mathrm{T}} \\ -\boldsymbol{X}_1\left(\boldsymbol{X}^{\mathrm{T}}\boldsymbol{X}\right)^{-1}\boldsymbol{X}_0^{\mathrm{T}} & \boldsymbol{I} - \boldsymbol{X}_1\left(\boldsymbol{X}^{\mathrm{T}}\boldsymbol{X}\right)^{-1}\boldsymbol{X}_1^{\mathrm{T}} \end{bmatrix}, \tag{7}$$

where the $\boldsymbol{I}$ in the upper left-hand corner in (7) is the $k \times k$ unit matrix, and the $\boldsymbol{I}$ in the lower right-hand corner is the $(n-k)\times(n-k)$ unit matrix.

Secondly, compute the eigenvalues, denoted by $d_1^2,\cdots,d_k^2$, of the matrix $\boldsymbol{X}_0\left(\boldsymbol{X}^{\mathrm{T}}\boldsymbol{X}\right)^{-1}\boldsymbol{X}_0^{\mathrm{T}}$, and the corresponding eigenvectors by $q_1,\cdots,q_k$.

Finally, calculate the BLUS residuals $\hat{\boldsymbol{\varepsilon}} = \boldsymbol{e}_1 - \boldsymbol{X}_1\boldsymbol{X}_0^{-1}\left[\sum_{i=1}^{k}\frac{d_i}{1+d_i}\boldsymbol{q}_i\boldsymbol{q}_i^{\mathrm{T}}\right]\boldsymbol{e}_0$, here, $\boldsymbol{e}_0,\boldsymbol{e}_1$ is the block residuals vector corresponding to the error term in (6). At this point, under the original assumption (2), the BLUS residual vector is

$$\hat{\boldsymbol{\varepsilon}} \sim N\left(0,\sigma^2\boldsymbol{I}_{n-k}\right). \tag{8}$$

The tests in this section and the next are constructed from this new sample. Denote $\hat{\boldsymbol{\varepsilon}}_{(1)},\cdots,\hat{\boldsymbol{\varepsilon}}_{(m)}$ be an order statistic of BLUS residuals, $m = n-k$, $F_0(x)$ be the normal distribution.

BLUS residuals were mainly applied to multivariate models (See ref. [31]), the cusum and the cusum-of-squares tests (See ref. [19]) which have higher power than those based on the more popular recursive residuals for structural break, and the Dickey-Fuller unit root test which is based on BLUS residuals (See ref. [32]), etc.

The test statistics constructed in this paper and the competition test statistics available in the literature are described below.

The competition test statistics based on the empirical distribution function are KS statistics

$$\hat{D}_m = \sqrt{m}\max_{1\le i\le m}\left\{\left|\frac{i}{m}-F_0(\tilde{\varepsilon})\right|,\left|\frac{i-1}{m}-F_0(\tilde{\varepsilon})\right|\right\} \tag{9}$$

and AD statistics

$$\hat{A}_m^2 = -\frac{1}{m}\sum_{i=1}^{m}\left[(2i-1)\log F_0(\tilde{\varepsilon})+(2m+1-2i)\log\left(1-F_0(\tilde{\varepsilon})\right)\right]-m, \tag{10}$$

here $\tilde{\varepsilon} = \hat{\varepsilon}_{(i)}\Big/\sqrt{S_m^2}$, since the mean of the original hypothesis is 0, the parameter estimation of the variance of the error is $S_m^2 = 1/m\sum_{i=1}^{m}\hat{\varepsilon}_i^2$.

The test statistic based on sample order statistics or sample quantiles is Shapiro & Wilk test statistic (See ref. [33]):

$$W = \frac{\left(\sum_{i=1}^{m}a_i\hat{\varepsilon}_{(i)}\right)^2}{m\cdot S_m^2} \tag{11}$$

where $\boldsymbol{a} = (a_1,\cdots,a_m)^{\mathrm{T}} = \frac{\boldsymbol{V}^{-1}\mathrm{E}\tilde{\boldsymbol{W}}}{\sqrt{\mathrm{E}\tilde{\boldsymbol{W}}^{\mathrm{T}}\boldsymbol{V}^{-1}\boldsymbol{V}^{-1}\mathrm{E}\tilde{\boldsymbol{W}}}}$, $V = (v_{ij}) = Var\left(\tilde{\boldsymbol{W}} - \mathrm{E}\tilde{\boldsymbol{W}}\right)$, $\tilde{\boldsymbol{W}} = \left(\tilde{W}_{(1)},\cdots,\tilde{W}_{(m)}\right)$ be the sample order statistics from $F_0(\cdot)$.

The test statistic constructed based on De Wet and Venter' idea (See ref. [34]) is

$$\hat{T}_2 = \sum_{i=1}^{m}\left(\frac{\hat{\varepsilon}_{(i)}}{\hat{\theta}} - F_0^{-1}\left(\frac{i}{m+1}\right)\right)^2, \tag{12}$$

where $F_0^{-1}(u) = \inf\{x : F_0(x) \geq u\}$, $\hat{\theta} = \sqrt{S_m^2}$ be the estimation of variance of standard error.

Del Barrio, Couesta-Albertos, Matrán, Rodríguez-Rodríguez (See ref. [35]) proposed that the statistic constructed based on the distance of L_2-Wasserstein is

$$\widehat{BCMR} = 1 - \frac{\left(\sum_{i=1}^{m}\hat{\varepsilon}_{(i)}\int_{(i-1)/m}^{i/m}\Phi^{-1}(t)\,\mathrm{d}t\right)^2}{S_m^2}, \tag{13}$$

Let $X_{(1)},\cdots,X_{(m)}$ be the order statistics of $X_1,\cdots,X_m$ and $U_{(1)},\cdots,U_{(m)}$ are the order statistics from $U(0,1)$. If $X_1,\cdots,X_m$ is an iid sample from the continuous cumulative distribution function $F(x)$, the following equation holds (See ref. [36]).

$$\left(X_{(1)},\cdots,X_{(m)}\right) \overset{d}{=} \left(F^{-1}\left(U_{(1)}\right),\cdots,F^{-1}\left(U_{(m)}\right)\right), \tag{14}$$

where $\overset{d}{=}$ stands for equality in distribution.

Using this conclusion, Zhao (See ref. [37] [38]) thinks that if $X_1,\cdots,X_m$ doesn't come from the null distribution $F(x)$, (14) doesn't hold and then there are differences between $\left(X_{(1)},\cdots,X_{(m)}\right)$ and $\left(F^{-1}\left(U_{(1)}\right),\cdots,F^{-1}\left(U_{(m)}\right)\right)$. The larger the differences are, the greater the evidence against the null hypothesis is. Using this idea, Zhao (See ref. [37]) constructed criteria to describe the differences based on random distance.

$$ZR = \sum_{i=1}^{m}\left(X_{(i)} - F^{-1}\left(U_{(i)}\right)\right)^2 \Big/ \sigma_i^2 \tag{15}$$

In this paper, the null hypothesis is that the distribution function is a normal distribution function with only scale parameters. So (14) is reduced to

$$X_{(i)} \overset{d}{=} \sigma Z_{(i)}, i = 1,\cdots,m. \tag{16}$$

And then, with respect to the parameter σ^2, take the smallest value of (15), it can be obtained that:

$$ZR = \sum_{i=1}^{m} Z_{(i)}^2 + \frac{\left[\sum_{i=1}^{m} X_{(i)} Z_{(i)}\right]^2}{\sum_{i=1}^{m} X_{(i)}^2} \tag{17}$$

where $Z_{(1)},\cdots,Z_{(m)}$ is the order statistics for a sample of m size from the normal distribution. Note that (8) and above analysis, using $\hat{\varepsilon}_{(i)}$ instead of $X_{(i)}$, we obtain the following test statistics in this paper.

$$\widehat{ZR} = \sum_{i=1}^{m} Z_{(i)}^2 + \frac{\left[\sum_{i=1}^{m} \hat{\varepsilon}_{(i)} Z_{(i)}\right]^2}{\sum_{i=1}^{m} \hat{\varepsilon}_{(i)}^2}, \tag{18}$$

where $\widehat{ZR}$ is scale-invariant. For the same reason as in Zhao (See ref. [37]), we can take the q quantile and expectation of $\widehat{ZR}$ as the test statistics. Let $\widehat{ZR}_q$

and $\widehat{ZR}_{\mu}$ be the qth quantile and the expectation of $\widehat{ZR}$. Here we select only $q = 0.05$, $q = 0.50$, and $q = 0.95$ as quantile statistics. Using BLUS residuals to solve the critical value of each test statistic is the same as Zhao's (See ref. [37]) algorithm, which is omitted here.

3. Power Comparisons

In order to compare the power of the above tests, the following linear models are considered.

$$y_i = 1 + 2x_{1i} - 3x_{2i} + \varepsilon_i, \quad i = 1, 2, \cdots, n. \tag{19}$$

The variable x is independent of ε_i, which comes from a uniform distribution, these values are constant for a given sample size. The error variable ε follows an alternative distribution.

Many continuous alternative distributions are chosen which were used in the power studying by Shapiro *et al.* (See ref. [33]), Gan and Koehler (See ref. [39]) and Eva Krauczi (See ref. [40]).

The first group: alternative distributions are the non-normal distribution, such as the chi-square distribution with a degree of freedom of 3, denoted by $\chi^2(3)$; the exponential distribution with a mean of 1, denoted by *Exp* (1); the gamma distribution functions with shape parameters of 0.8 and 1.5, *Gamma* (0.8), *Gamma* (1.5), respectively; the double exponential distribution *laplace* (0,1); the log-normal distribution *lognormal* (0,1); the Cauchy distribution *Cauchy* (0,1); the student distribution with degree of freedom of 3, *t* (3); the unbounded Johnson's distribution of the random variable sinh (*Z*), denoted as *SU* (0,1), where $Z \sim N(0,1)$.

The second group: in heteroscedasticity case, the he model under consideration at this point is

$$y_i = 1 + 2x_{1i} - 3x_{2i} + \sigma_i \bar{\varepsilon}_i, \quad i = 1, 2, \cdots, n, \tag{20}$$

where σ_i be distributed as uniform distribution $U(0,50)$, $i = 1, 2, \cdots, n$. At the same time, the error variable $\bar{\varepsilon}$ follows *logistic* (0,1), *Cauchy* (0,1), *N* (0,1) and *Laplace* (0,1), respectively.

The third group: the case of outlier with non-zero mean of regression error, such as

$$\varepsilon_1 \sim N(2,1), \varepsilon_i \sim N(0,1), i = 2, \cdots, n,$$

and

$$\varepsilon_i \sim N(2,1), i = 1, \cdots, 5, \ \varepsilon_i \sim N(0,1), i = 6, \cdots, n,$$

and also in high Leverage outlier case, like that

$$\varepsilon_1 \sim N(8,1), \varepsilon_i \sim N(0,1), i = 2, \cdots, n, \ x_{11} = x_{21} = 2.$$

In this section, the selected test level is 5%, the sample size is $n = 20, 50$, and the empirical power of the test is based on 10,000 simulations. The results are placed in **Table 1** and **Table 2**. From these two tables, you can see that:

Table 1. Powers for testing the error distribution is normal, against different non-normal alternative distributions, at significance level 5% and $n = 20, n = 50$.

Alternatives	$\chi^2(3)$	$exp(1)$	$G(0.8)^a$	$G(1.5)^a$	$laplace$	$F_1(x)$ [b]	$Cauchy$	$t_3(x)$	$F_2(x)$ [c]
					$n = 20$				
$\widehat{ZR}_{q=0.05}$	0.2875	0.3968	0.4744	0.2811	0.1917	0.6167	0.7496	0.2533	0.3166
$\widehat{ZR}_{q=0.50}$	0.1623	0.2237	0.2881	0.1580	0.1034	0.4444	0.6277	0.1619	0.1994
$\widehat{ZR}_{\mu}$	0.1349	0.1869	0.2421	0.1320	0.0923	0.3912	0.5880	0.1431	0.1746
$\widehat{ZR}_{q=0.95}$	0.0710	0.0858	0.1073	0.0709	0.0707	0.2026	0.4134	0.0915	0.1014
D_n	0.0972	0.1160	0.1335	0.1029	0.0756	0.1872	0.3686	0.0792	0.1047
A_n^2	0.0851	0.0997	0.1195	0.0898	0.0666	0.1892	0.3942	0.0790	0.0926
$BCMR$	0.1233	0.1722	0.2235	0.1209	0.0836	0.3698	0.5553	0.1273	0.1532
T_2	0.1415	0.1959	0.2511	0.1393	0.0990	0.4031	0.6099	0.1507	0.1856
W	0.1201	0.1677	0.2161	0.1173	0.0811	0.3609	0.5438	0.1214	0.1463
					$n = 50$				
$\widehat{ZR}_{q=0.05}$	0.7992	0.9136	0.9473	0.7902	0.4586	0.9754	0.9837	0.5928	0.7087
$\widehat{ZR}_{q=0.50}$	0.6621	0.8316	0.8963	0.6489	0.2760	0.9593	0.9709	0.4426	0.5480
$\widehat{ZR}_{\mu}$	0.6072	0.7903	0.8692	0.5970	0.2407	0.9478	0.9648	0.4058	0.5080
$\widehat{ZR}_{q=0.95}$	0.3692	0.5743	0.6866	0.3605	0.1329	0.8736	0.9166	0.2698	0.3318
D_n	0.2440	0.3877	0.4918	0.2503	0.1390	0.7121	0.8728	0.1892	0.2433
A_n^2	0.2518	0.4270	0.5470	0.2566	0.1044	0.7870	0.8969	0.1835	0.2394
$BCMR$	0.6081	0.7920	0.8697	0.5974	0.2040	0.9473	0.9570	0.3694	0.4656
T_2	0.6078	0.7890	0.8655	0.5957	0.2740	0.9489	0.9707	0.4435	0.5468
W	0.6058	0.7928	0.8730	0.5942	0.1579	0.9453	0.9397	0.3106	0.3938

[a] $G(0.8), G(1.5)$: the Gamma distribution with shape parameter 0.8, 1.5. [b] $F_1(x)$: the LogNormal(0, 1) distribution. [c] $F_3(x)$: $X = \sinh(Z)$, $Z \sim N(0,1)$.

1) When the alternative distribution is a non-normal distribution function, the power of the test $\widehat{ZR}_{q=0.05}$ is significantly higher than that of other tests. When the capacity is 20, the power of $\widehat{ZR}_{\mu}$ and T_2 is similar. However, when the capacity is 50, for an asymmetric alternative distribution, $BCMR$, T_2, W and $\widehat{ZR}_{\mu}$ are not much different, for a symmetric alternative distribution, T_2 is the best, coming next is $\widehat{ZR}_{\mu}$. In short, Among the opponents of the new tests, the order of superiority and inferiority of the tests is T_2, $BCMR$, Shapiro-Wilk test, A_n^2 and D_n.

2) For an alternate distribution of $N(0,1)$ the power of the test is approximately equal to the probability of the first type of error. As can be seen from Table 2, under the different sample sizes, all the tests make full use of the approximately 5% test level.

3) For the alternative distribution of the second group, the heteroscedasticity case, the power of $\widehat{ZR}_{q=0.05}$ test is significantly higher than the other tests, the

Table 2. Powers for testing the error distribution is normal, against Heteroscedasticity, outlier alternative distribution and normal alternative distributions, at significance level 5% and $n = 20, n = 50$.

Alternatives	$N(0,1)$	$F_1(x)$ [a]	$F_2(x)$ [a]	$F_3(x)$ [a]	$F_4(x)$ [a]	$F_5(x)$ [b]	$F_6(x)$ [b]	$F_7(x)$ [b]
				$n = 20$				
$\widehat{ZR}_{q=0.05}$	0.0462	0.7786	0.2011	0.4041	0.2837	0.0461	0.4500	0.0562
$\widehat{ZR}_{q=0.50}$	0.0483	0.6500	0.1004	0.2233	0.1531	0.0493	0.5529	0.0674
$\widehat{ZR}_{\mu}$	0.0485	0.6114	0.0896	0.1909	0.1324	0.0496	0.5561	0.0678
$\widehat{ZR}_{q=0.95}$	0.0484	0.4239	0.0671	0.1037	0.0881	0.0478	0.5561	0.0684
D_n	0.0509	0.3821	0.0772	0.1203	0.0992	0.0482	0.5365	0.0647
A_n^2	0.0504	0.4066	0.0628	0.0967	0.0816	0.0477	0.5921	0.0694
$BCMR$	0.0474	0.5769	0.0809	0.1661	0.1184	0.0485	0.5528	0.0684
T_2	0.0482	0.6308	0.0968	0.2086	0.1465	0.0485	0.5560	0.0672
W	0.0479	0.5628	0.0779	0.1564	0.1122	0.0483	0.5518	0.0685
				$n = 50$				
$\widehat{ZR}_{q=0.05}$	0.0534	0.9851	0.4994	0.8355	0.6521	0.1143	0.6871	0.2844
$\widehat{ZR}_{q=0.50}$	0.0497	0.9778	0.2760	0.6614	0.4302	0.1392	0.7832	0.3717
$\widehat{ZR}_{\mu}$	0.0502	0.9731	0.2351	0.6102	0.3818	0.1414	0.7865	0.3761
$\widehat{ZR}_{q=0.95}$	0.0506	0.9344	0.1195	0.3626	0.1951	0.1427	0.7933	0.3857
D_n	0.0520	0.9062	0.1757	0.3782	0.2203	0.1245	0.7096	0.3036
A_n^2	0.0505	0.9242	0.1102	0.3260	0.1632	0.1394	0.7944	0.3726
$BCMR$	0.0500	0.9672	0.1951	0.5548	0.3341	0.1413	0.7868	0.3772
T_2	0.0504	0.9781	0.2777	0.6611	0.4314	0.1410	0.7852	0.3748
W	0.0519	0.9536	0.1493	0.4584	0.2559	0.1427	0.7870	0.3838

[a]The error $\sigma\bar{\varepsilon}$, $F_1(x)$: $\bar{\varepsilon}$ the Cauchy distribution; $F_2(x)$: $\bar{\varepsilon}$ the Normal distribution; $F_3(x)$: $\bar{\varepsilon}$ the Laplace distribution; $F_4(x)$: $\bar{\varepsilon}$ the Logistic distribution $\sigma \sim U(0,50)$. [b] $F_5(x)$: $\varepsilon_1 \sim N(2,1)$, $\varepsilon_i \sim N(0,1)$, $i = 2,\cdots,n$; $F_6(x)$: $\varepsilon_i \sim N(2,1)$, $i = 1,\cdots,5$, $\varepsilon_i \sim N(0,1)$, $i = 6,\cdots,n$; $F_7(x)$: $\varepsilon_1 \sim N(8,1)$, $\varepsilon_i \sim N(0,1)$, $i = 2,\cdots,n$, $x_{11} = x_{12} = 2$.

other results are similar to those in (1).

4) For the alternative distribution of the third group, where there are outliers. In the new test, $\widehat{ZR}_{q=0.95}$ performs best, but not much different from $\widehat{ZR}_{q=0.50}$, $\widehat{ZR}_{\mu}$. Compared with the comparative test, the difference is not significant, A_n^2 is slightly better, while $\widehat{ZR}_{q=0.05}$ performs worse.

4. Comments

Based on the BLUS residuals, using the difference between the residuals order statistics and the pseudo-random sample order statistics, the quantile-type and the conditional expectation-type test statistics are constructed, which are used in the error distribution Normality test of the linear regression model. The simulation results show that the power of the tests provided in this paper is better than some tests in the literature. Of course, T_2, $BCMR$ and W are also good tests.

Conflicts of Interest

The authors declare no conflicts of interest regarding the publication of this paper.

References

[1] Eubank, R.L. and Hart, J.D. (1992) Testing Goodness-of-Fit in Regression via Order Selection Criteria. *Annals of Statistics*, **20**, 1412-1425. https://doi.org/10.1214/aos/1176348775

[2] Milhem, H. (2003) Testing the Goodness-of-Fit in a Gaussian Regression. *Proceedings of the European Young Statisticians Meeting* (*EYSM*), Ovronnaz, 21-26 September 2003, 101-111.

[3] Young, S.G. and Bowman, A.W. (1995) Non-Parametric Analysis of Covariance. *Biometrics*, **51**, 920-931. https://doi.org/10.2307/2532993

[4] Durbin, J. (1973) Distribution Theory for Tests Based on the Sample Distribution Function. Society for Industrial and Applied Mathematics, Philadelphia. https://doi.org/10.1137/1.9781611970586

[5] Durbin, J., Knott, M. and Taylor, C.C. (1975) Components of Cramérvon Mises statistics II. *Journal of the Royal Statistical Society*: *Series B* (*Methodological*), **37**, 216-237. https://doi.org/10.1111/j.2517-6161.1975.tb01537.x

[6] D'Agostino, R.B. and Stephens, M.A. (1986) Goodness-of-Fit Techniques. Marcel Dekker, Inc., New York.

[7] Cox, D., Koh, E., Wahba, G. and Yandell, B.S. (1988) Testing the (Parametric) Null Model Hypothesis in (Semiparametric) Partial and Generalized Spline Models. *Annals of Statistics*, **16**, 113-119. https://doi.org/10.1214/aos/1176350693

[8] Gu, C. (1992) Diagnostics for Nonparametric Regression Models with Additive Terms. *Journal of the American Statistical Association*, **87**, 1051-1058. https://doi.org/10.1080/01621459.1992.10476260

[9] Eubank, R.L. and Hart, J.D. (1993) Commonality of Cusum, von Neumann and Smoothing-Based Goodness-of-Fit Tests. *Biometrika*, **80**, 89-98. https://doi.org/10.1093/biomet/80.1.89

[10] Eubank, R.L. and La Riccia, V.N. (1993) Testing for No Effect in Nonparametric Regression. *Journal of Statistical Planning and Inference*, **36**, 1-14. https://doi.org/10.1016/0378-3758(93)90097-P

[11] Stute, W. (1997) Nonparametric Model Checks for Regression. *Annals of Statistics*, **25**, 613-641. https://doi.org/10.1214/aos/1031833666

[12] Stute, W. and Gonzáilez Manteiga, W. (1996) NN Goodness-of-Fit Tests for Linear Models. *Journal of Statistical Planning and Inference*, **53**, 75-92. https://doi.org/10.1016/0378-3758(95)00144-1

[13] Stute, W., González Manteiga, W. and Presedo Quindimil, M. (1998) Bootstrap Approximations in Model Checks for Regression. *Journal of the American Statistical Association*, **93**, 141-149. https://doi.org/10.1080/01621459.1998.10474096

[14] Stute, W., Thies, S. and Zhu, L.X. (1998) Model Checks for Regression: An Innovation Process Approach. *Annals of Statistics*, **26**, 1916-1934. https://doi.org/10.1214/aos/1024691363

[15] Stute, W. and Zhu, L.X. (2002) Model Checks for Generalized Linear Models. *Scandinavian Journal of Statistics*, **29**, 535-545. https://doi.org/10.1111/1467-9469.00304

[16] Koul, H.L. and Ni, P. (2004) Minimum Distance Regression Model Checking. *Jour-*

nal of Statistical Planning and Inference, **119**, 109-141.
https://doi.org/10.1016/S0378-3758(02)00415-9

[17] Zhu, L., Zhu, R. and Song, S. (2008) Diagnostic Checking for Multivariate Regression Models. *Journal of Multivariate Analysis*, **99**, 1841-1859.
https://doi.org/10.1016/j.jmva.2008.01.022

[18] Theil, H. (1968) A Simplification of the BLUS Procedure for Analyzing Regression Disturbances. *Journal of the American Statistical Association*, **63**, 242-251.
https://doi.org/10.1080/01621459.1968.11009238

[19] Magnus, J.R. and Sinha, A.K. (2005) On Theil's Errors. *Econometrics Journal*, **8**, 39-54. https://doi.org/10.1111/j.1368-423X.2005.00150.x

[20] White, H. and MacDonald, G. (1980) Some Large-Sample tests for Non-Normality in the Linear Regression Model. *Journal of the American Statistical Association*, **75**, 16-28. https://doi.org/10.1080/01621459.1980.10477415

[21] Cliff, H.J. and Ben, B.W. (1974) On the Testing of Regression Disturbances for Normality. *Journal of the American Statistical Association*, **69**, 330-335.
https://doi.org/10.1080/01621459.1974.10482948

[22] Jurečková, J., Picek, J. and Sen, P.K. (2003) Goodness-of-Fit Test with Nuisance Regression and Scale. *Metrika*, **58**, 235-258. https://doi.org/10.1007/s001840300262

[23] Sen, P.K., Jurečková, J. and Picek, J. (2003) Goodness-of-Fit Test of Shapiro-Wilk Type with Nuisance Regression and Scale. *Austrian Journal of Statist*, **32**, 163-177.

[24] Shapiro, S.S. and Wilk, M.B. (1965) An Analysis of Variance Test for Normality. *Biomatrika*, **52**, 591-611. https://doi.org/10.1093/biomet/52.3-4.591

[25] Jurečková, J. and Picek, J. (2007) Shapiro-Wilk Type Test of Normality under Nuisance Regression and Scale. *Computational Statistics & Data Analysis*, **51**, 5184-5191. https://doi.org/10.1016/j.csda.2006.08.026

[26] Neumeyer, N., Dette, H. and Nagel, E.-R. (2006) Bootstrap Tests for the Error Distribution in Linear and Nonparametric Regression Models. *Australian & New Zealand Journal of Statistics*, **48**, 129-156.
https://doi.org/10.1111/j.1467-842X.2006.00431.x

[27] Hattab, M.W. and Christensen, R. (2018) Lack of Fit Tests Based on Sums of Ordered Residuals for Linear Models. *Australian & New Zealand Journal of Statistics*, **60**, 230-257. https://doi.org/10.1111/anzs.12231

[28] Christensen, R. and Sun, S.K. (2010) Alternative Goodness-of-Fit Tests for Linear Models. *Journal of the American Statistical Association*, **105**, 291-301.
https://doi.org/10.1198/jasa.2009.tm08697

[29] Christensen, R. and Lin, Y. (2015) Lack-of-fit Tests Based on Partial Sums of Residuals. *Communications in Statistics - Theory and Methods*, **44**, 2862-2880.
https://doi.org/10.1080/03610926.2013.844256

[30] Ramsey, J.B. (1969) Tests for Specification Errors in Classical Linear Least Squares Regression Analysis. *Journal of the Royal Statistical Society: Series B* (*Methodological*), **31**, 350-371. https://doi.org/10.1111/j.2517-6161.1969.tb00796.x

[31] Chinchilli, V.M. (1987) Blus Residuals in Multivariate Linear Models. In: Bozdogan, H. and Gupta, A.K., Eds., *Multivariate Statistical Modeling and Data Analysis*, Springer, Dordrecht, 61-75. https://doi.org/10.1007/978-94-009-3977-6_4

[32] Vougas, D.V. (2008) Unit Root Testing Based on BLUS Residuals. *Statistics & Probability Letters*, **78**, 1943-1947. https://doi.org/10.1016/j.spl.2008.01.063

[33] Shapiro, S.S., Wilk, M.B. and Chen, H.J. (1968) A Comparative Study of Various Tests for Normality. *Journal of the American Statistical Association*, **63**, 1343-1372.

https://doi.org/10.1080/01621459.1968.10480932

[34] De Wet, T. and Venter, J.H. (1972) Asymptotic Distributions of Certain Test Criteria of Normality. *South African Statistical Journal*, **6**, 135-149.

[35] Del Barrio, E., Couesta-Albertos, J.A., Rodríguez-Rodríguez, J.M. and del Barrio, E. (1999) Tests of Goodness of Fit Based on the L_2-Wasserstein Distance. *The Annals of Statistics*, **27**, 1230-1239. https://doi.org/10.1214/aos/1017938923

[36] Scheffe, H. and Tukey, J.W. (1945) Non-Parametric Estimation. I. Validation of Order Statistics. *The Annals of Mathematical Statistics*, **16**, 187-192. https://doi.org/10.1214/aoms/1177731119

[37] Zhao, J., Xu, X. and Ding, X. (2009) Some New Goodness-of-Fit Tests Based on Stochastic Sample Quantiles. *Communications in Statistics - Simulation and Computation*, **38**, 571-589. https://doi.org/10.1080/03610910802585832

[38] Zhao, J. and Xu, X. (2014) Goodness-of-Fit Tests for Location?Cscale Families Based on Random Distance. *Journal of Statistical Computation and Simulation*, **84**, 739-752. https://doi.org/10.1080/00949655.2012.725246

[39] Gan, F.F. and Koehler, K.J. (1990) Goodness-of-Fit Tests Based on P-P Probability Plots. *Technometrics*, **32**, 289-303. https://doi.org/10.2307/1269106

[40] Krauczi, É. (2007) A Study of the Quantile Correlation Test for Normality. *Test*, **18**, 156-165. https://doi.org/10.1007/s11749-007-0074-6