AUT Journal of Mathematics and Computing

AUT J. Math. Comput., 3(1) (2022) 77-91
DOI: 10.22060/AJMC.2021.20246.1062

Comparing regression methods with non-Gaussian stable errors

Reza Alizadeh Noughabi[a], Adel Mohammadpour*[a]

[a]*Department of Statistics, Faculty of Mathematics and Computer Science, Amirkabir University of Technology (Tehran Polytechnic), Iran*

ABSTRACT: Nolan and Ojeda-Revah in [16] proposed a regression model with heavy-tailed stable errors. In this paper, we extend this method for multivariate heavy-tailed errors. Furthermore, A likelihood ratio test (LRT) for testing significant of regression coefficients is proposed. Also, confidence intervals based on fisher information for [16] method, called NOR, and LRT are computed and compared with well-known methods. In the end, we provide some guidance for various error distributions in heavy-tailed cases.

Review History:
Received:08 July 2021
Accepted:27 October 2021
Available Online:01 February 2022

Keywords:
Regression
Quantile regression
Stable distribution
Ordinary least squares
Maximum likelihood.

AMS Subject Classification (2010):
60E07; 62G08

1. Introduction

Stable distribution was introduced by Paul Lévy in his study of sums of independent and identically distributed terms in the 1920's. The stable distribution provides heavy-tails, and a dataset including outliers can be well modeled by stable distributions.
A d-dimensional random vector $\boldsymbol{E}$ is called stable if $\boldsymbol{E}$ has a characteristic function, $\varphi_{\boldsymbol{E}}(\boldsymbol{t})$, of the form

$$\exp\left(-\int_{\mathbb{S}_d}\psi_\alpha\left(\langle\boldsymbol{t},\boldsymbol{s}\rangle\right)\Lambda\left(d\boldsymbol{s}\right)+i\left\langle\boldsymbol{t},\boldsymbol{\delta}\right\rangle\right)$$

where

$$\psi_\alpha(u)=\begin{cases}|u|^\alpha\left(1-i\mathrm{sgn}\left(u\right)\tan\left(\frac{\pi\alpha}{2}\right)\right), & \alpha\neq 1,\\ |u|\left(1+i\frac{2}{\pi}\mathrm{sgn}\left(u\right)\log|u|\right), & \alpha=1.\end{cases}$$

Also Λ is a finite measure on Borel subsets of $\mathbb{S}_d=\{\boldsymbol{t}\in\mathbb{R}^d:|\boldsymbol{t}|=1\}$, $\boldsymbol{\delta}\in\mathbb{R}^d$ is a shift vector, $\alpha\in(0,2]$ is the tail index and $\mathrm{sgn}\,u$ is the sign function. Multivariate stable distributions, in the general case, are not computational by available computers. Fortunately, several special cases are computationally available, elliptical, isotropic and independent components.

Elliptical random vectors is a well known parametric sub-class with the following characteristic function

$$\exp\left(\left(\boldsymbol{t}'R\boldsymbol{t}\right)^{\frac{\alpha}{2}}+i\left\langle\boldsymbol{t},\boldsymbol{\delta}\right\rangle\right),$$

*Corresponding author.
E-mail addresses: re_al@aut.ac.ir, adel@aut.ac.ir

where R is a positive definite matrix. $R = \gamma^2 I$ is known as isotropic stable.

In univariate case ($d = 1$), characteristic function of the stable random variable ε, reduced as follows

$$\varphi_\varepsilon(t) = \begin{cases} \exp\left\{-\gamma^\alpha |t|^\alpha \left[1 - i\beta\left(\tan\frac{\pi\alpha}{2}\right)(\operatorname{sgn} t)\right] + i\delta t\right\} & \alpha \neq 1, \\ \exp\left\{-\gamma |t| \left[1 + i\beta\frac{2}{\pi}(\operatorname{sgn} t)\log(\gamma |t|)\right] + i\delta t\right\} & \alpha = 1, \end{cases}$$

where $-1 \leq \beta \leq 1$ is the skewness parameter, $\gamma > 0$ and $\delta \in \mathbb{R}$ are the scale and location parameters, respectively, Nolan (2019). If components are independent with univariate stable, called independent components.

Ordinary Least Squares (OLS) is the most popular method for performing regression when error terms have normally been distributed. If errors of regression have a heavy-tailed distribution, especially non-Gaussian stable distribution, $\alpha \in (0, 2)$, the OLS method is very poor. We focus on the cases that the mean or variance of errors' distribution does not exist.

The OLS is a method for estimating a mean of a variable conditional on the levels or values of independent variables. However, we don't always have to estimate the conditional mean. That's where quantile regression comes in. The algorithm for constructing median regression was proposed by [3].

[4] and [7] did some valuable research about quantile regression and expressed some advantages of this regression versus OLS. Several recent studies have investigated quantile regressions; for example, a quantile regression memoir is provided by [12] and [13]. Also, [8] have a long history on quantile regression in their book. Furthermore, [17] discussed the best linear unbiased estimator. Unfortunately, this paper is not available, and we are unable to comment on this paper. [2] assumed errors with symmetric distribution, and they minimized the sum of absolute errors. [14] used the maximum likelihood approach to estimate the linear regression coefficients when the errors follow from a symmetric stable distribution. Furthermore, in recent years, [1] studied autoregressive time series models for α-stable models. Improving the method of [2] by an estimator, based on the ranked set sampling, is studied in [9]. [10] applied the ranked methods to fit a linear model in the case of infinite variance.

[16] introduced a method for regression with stable errors. Nolan's method used a trimmed regression for estimating the initial values of regression parameters. Two quantiles are used to select a trimmed data set. The suggestions are 0.1 and 0.9. By the residuals, the maximum likelihood estimator (MLE) of parameters of the stable distribution that is considered as an error term is computed. The initial estimate for all parameters is used as the starting value for a numerical optimization to find the maximum of the likelihood function. In other words, Nolan's method evaluates parameters of linear regression by maximum likelihood when the errors have a stable distribution. As we know, non-Gaussian stable distributions do not have a variance (and in some cases, mean). Therefore, we decided to compare this algorithm with quantile regression.

The plan of this paper is as follows. In section 2, we generalized Nolan's method (GNOR) to the multivariate regression. In section 3, linear regression is studied and simulation results are presented. In section 4, we investigate linear regression with no stable heavy-tailed errors. In section 5, nonlinear regression with stable error is discussed. The last sections are a brief discussion on the likelihood ratio test and the confidence intervals of the regression parameters.

2. Generalized Nolan's Method to Multivariate Regression

Now we want to develop this method for multivariate regression. The multivariate linear regression model has the form

$$\boldsymbol{Y} = \boldsymbol{X}\mathbf{B} + \mathbf{E},$$

where $\boldsymbol{Y} = (\boldsymbol{Y}_1, \ldots, \boldsymbol{Y}_m)_{n\times m}$ is the response matrix and $\boldsymbol{Y}_k = (Y_{1k}, \ldots, Y_{nk})^T$ is k-th response vector, $\mathbf{B} = (\boldsymbol{b}_1, \ldots, \boldsymbol{b}_m)$ is the matrix of regression coefficients, i.e, $\boldsymbol{b}_k = (b_{0k}, b_{1k}, \ldots, b_{pk})^T$ is the k-th coefficient vector, and $\mathbf{E} = (\boldsymbol{\epsilon}_1, \ldots, \boldsymbol{\epsilon}_m)$ is the residuals matrix with normal distributed, i.e. $\mathbf{E}$ has $N(\mathbf{0}_{nm}, \boldsymbol{\Sigma} \otimes \boldsymbol{I}_n)$ where $\otimes$ denotes the Kronecker product. Now, we assume that the $\mathbf{E}$ has a multivariate stable distribution.

The algorithm that introduced by [16] for ordinary regression can be generalized in the multivariate regression by a few changes. We propose the following algorithm.

Algorithm 2.1.

1. Perform an initial multivariate OLS and compute the residuals.
2. Omit the residuals that are large and small. For example, if we have bivariate regression, sort the residuals of each variable, then the lowest and highest 5% are trimmed away.
3. Remove the data corresponding to the omitted residuals.
4. Perform a multivariate OLS by the new data set.
5. Use the estimation of step 4 as the initial values and optimize the multivariate distribution parameters that it considered for the residuals.

For the simulation, we consider two responses and four explanatory variables. The error terms are independent multivariate stable distribution with different values of parameters α and β; we have the following regression model:

$$(\boldsymbol{Y}_1, \boldsymbol{Y}_2) = (\boldsymbol{X}_1, \boldsymbol{X}_2) \begin{pmatrix} b_{11} & b_{12} \\ b_{21} & b_{22} \end{pmatrix} + \mathbf{E}.$$

For simulation we consider the following values $b_{11} = 10$, $b_{12} = 7$, $b_{21} = 0.5$ and $b_{22} = 4$. $\boldsymbol{X}_1$ and $\boldsymbol{X}_2$ are two vectors of size 50 such that $\boldsymbol{X}_1$ and $\boldsymbol{X}_2$ are uniformly distributed in (0, 100) and (100, 300), respectively. As mentioned $\mathbf{E}$ is a bivariate stable distribution with independent components. A multivariate stable distribution with dimension d and independent components requires an index of stability, $\alpha \in (0, 2]$, and vectors β, γ and δ of length d. Simulation results are reported in Tables 1 and 2 . From reported values in Table 1 and Figures 1 and 2, it is clear that the proposed algorithm for multivariate regression is very efficient and has little MAD. We recall that Mean Absolute Deviations (MAD) of estimators are computed as follows for 1000 iterations:

$$\hat{\theta} = \frac{1}{1000} \sum_{i=1}^{1000} \hat{\theta}_i,$$

and

$$\text{MAD}\left(\hat{\theta}\right) = \frac{1}{1000} \sum_{i=1}^{1000} \left|\hat{\theta}_i - \theta\right|,$$

where $\hat{\theta}_i$ is an estimation of the regression coefficients in the i-th iteration.

Table 1: MAD and estimation (EST) of the regression parameters with independent bivariate stable error in 1000 iterations and sample size $n = 30$.

			GNOR				OLS			
α	β		b_1	b_2	b_3	b_4	b_1	b_2	b_3	b_4
0.7	(0, 0)	MAD	0.0072	0.0984	0.0020	0.0325	4.3351	0.5898	3.2572	1.6018
		EST	10.0011	0.4997	7.0758	3.9736	13.4867	0.8770	5.8755	4.8587
0.7	(0.5, 0.5)	MAD	0.0067	0.0961	0.0019	0.0216	563.4902	115.4902	2.7753	1.1907
		EST	10.0005	0.50002	6.9354	4.0125	573.3112	-114.9013	9.2304	3.0898
0.7	(-0.5, -0.5)	MAD	0.0063	0.0146	0.0017	0.0040	1.1132	0.2475	2.2593	0.6130
		EST	10.0002	0.4996	6.9962	4.0013	9.5989	0.3939	5.4619	4.3275
0.7	(0.9, -0.9)	MAD	0.0046	0.0680	.0016	0.0310	0.6651	0.1525	5.4006	0.5758
		EST	9.9996	0.5011	7.0347	3.9824	10.3714	0.5409	2.4264	4.0384
0.7	(-0.9, -0.9)	MAD	0.0054	0.0224	0.0016	0.0085	7.1286	1.3446	4.2353	1.1099
		EST	9.9994	0.4989	7.0107	3.9946	4.0235	1.2441	6.4865	3.5814

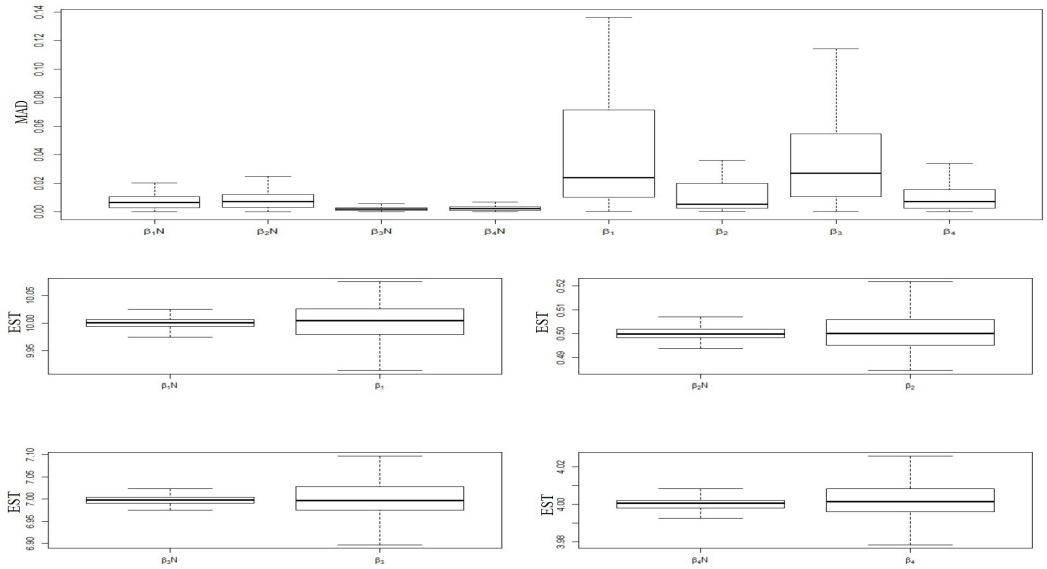

Figure 1: Box plot of MAD (up) and the estimation of each regression coefficient when residuals are simulated from multivariate stable with parameters $\alpha = 1$ and $\beta = (0, 0)$, for two different methods.

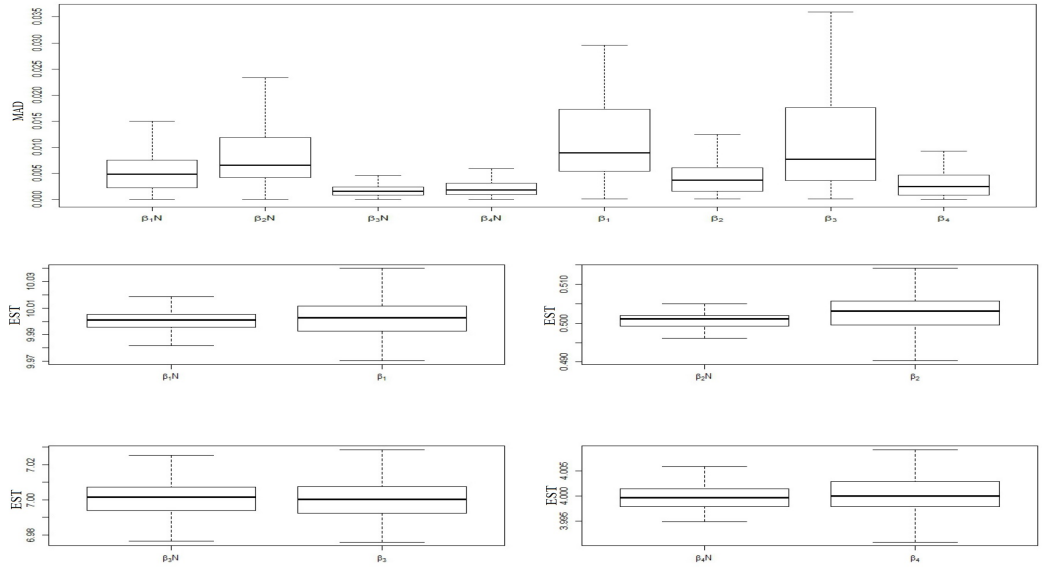

Figure 2: Box plot of MAD (top) and the estimation of each regression coefficient when residuals are simulated from multivariate stable with parameters $\alpha = 1.5$ and $\beta = (0.9, 0)$, for two different methods.

3. Linear Regression

The linear regression model can be written as

$$\boldsymbol{Y} = \boldsymbol{X\theta} + \boldsymbol{\epsilon},$$

where $\boldsymbol{X}_{n\times(k+1)}$ is the design matrix, $\boldsymbol{\theta} = (\theta_0, \theta_1, \ldots, \theta_k)^T$ are the regression coefficients, and $\boldsymbol{\epsilon} = (\epsilon_1, \ldots, \epsilon_n)^T$ are i.i.d. random variables. In OLS, ϵ_i must be normally distributed, but this paper assumes ϵ_i has a non-Gaussian stable distribution.

In this section, we consider two classes of stable distributions for error terms of regression: class I when $\alpha \geq 1$, $\{1, 1.5\}$ and class II when $\alpha < 1$, i.e., 0.7. In each class, β is chosen in such a way as to have symmetric and asymmetric stable distributions, $\beta = 0, -0.5, 0.5$.

To perform quantile regression three quantiles are considered, $q = 0.2$, 0.5 and 0.8. It can be noted that $q = 0.5$ is the median.

Table 2: MAD and estimation (EST) of the regression parameters with isotropic bivariate stable error in 1000 iterations and sample size $n = 100$.

		GNOR				OLS			
α		b_1	b_2	b_3	b_4	b_1	b_2	b_3	b_4
1.5	MAD	0.0022	0.0021	0.0007	0.0007	0.0089	0.0022	0.008	0.0023
	EST	10.0001	0.5001	6.9997	4.00003	10.0005	0.4996	7.0023	3.9992
1	MAD	0.0028	0.0026	0.0007	0.0007	0.0834	0.0229	0.0635	0.0173
	EST	10.0004	0.4999	6.9993	4.0001	10.0212	0.4930	7.0224	3.9924
0.7	MAD	0.0031	0.0031	0.0013	0.0012	8.5564	1.3438	16.8908	9.2170
	EST	10.0004	0.5010	7.0011	4.0010	13.0278	1.1671	22.6068	-4.6629

3.1. Simple Linear Regression

For simulation, we consider the following regression model

$$\boldsymbol{Y} = 0.7\boldsymbol{X} + \boldsymbol{\epsilon}, \tag{1}$$

where $\boldsymbol{X}$ is the vector of 50 numbers uniformly distributed in (0,100) as an explanatory variable and $\boldsymbol{\epsilon}$ is the vector of errors from a stable distribution.

Table 3, Figures 3 and 4 show the simulation results. Furthermore, we have got the error terms of regression and have drawn the figure of residuals in some cases (see Figures 5-7). Furthermore, by the residuals of regression, we obtained the maximum likelihood estimators of stable distribution parameters. The results are summarized in Table 4, which include OLS, quantile regression with quantiles 0.2, 0.5 and 0.8, ($Q(0.2)$, $Q(0.5)$, $Q(0.8)$) and Nolan's method (NOR).

It is observed that from Tables 3 and 4, and Figures 3-7, for the mentioned model the quantile regression for both of classes is the best. However, we should clarify which quantile of quantile regressions gives the best result. In other words, we introduce a few rules for selecting a quantile regression based on the tail and skewness parameters of error terms in the following.

Quantile selection.
In quantile regression, the best quantile selection is an important problem. In other words, using an appropriate quantile enables us to build a better model and prediction and so, we have less error in prediction. If we select an inappropriate quantile, the estimation of regression coefficients may have greater error, so the length of the confidence interval increases for the slope parameter. We demonstrate this fact through a simulation study and then propose a few advices in practical problems.

Table 3: MAD and estimation (EST) of the regression coefficients in two classes in 1000 iterations. The best values are bolded.

		NOR	$Q(0.2)$	$Q(0.5)$	$Q(0.8)$	OLS
$S(1.5,0)$	EST	0.6935	0.6747	0.6937	0.7129	0.6990
	MAD	0.0111	0.0205	**0.0035**	0.0197	0.0147
$S(1.5,0.5)$	EST	0.6964	0.6713	0.6923	0.7193	0.6913
	MAD	0.0165	0.0242	**0.0058**	0.0166	0.0113
$S(1.5,-0.5)$	EST	0.6957	0.6842	0.7052	0.7250	0.6928
	MAD	0.0102	0.0166	**0.0059**	0.0243	0.0081
$S(1,0)$	EST	0.6937	0.6791	0.6951	0.7354	0.7361
	MAD	0.1050	0.0244	**0.0032**	0.0246	0.1484
$S(1,0.5)$	EST	0.7191	0.6845	0.6997	0.7277	0.7637
	MAD	0.0191	0.0138	**0.0047**	0.0409	0.0927
$S(1,-0.5)$	EST	0.7026	0.6600	0.6978	0.7162	0.6151
	MAD	0.0131	0.0422	**0.0050**	0.0141	0.1171
$S(0.7,0)$	EST	0.6926	0.6699	0.6972	0.7382	-
	MAD	0.0082	0.0353	**0.0026**	0.0348	-
$S(0.7,0.5)$	EST	0.7118	0.7033	0.7196	0.8371	-
	MAD	0.0118	**0.0059**	0.0212	0.0948	-
$S(0.7,-0.5)$	EST	0.6924	0.6415	0.6760	0.6917	-
	MAD	0.0150	0.0966	0.0215	**0.0059**	-

Table 4: Maximum likelihood estimator of tail index (α) and skewness (β) with their's MAD in 1000 iterations. The best values are bolded.

		NOR		$Q(0.2)$		$Q(0.5)$		$Q(0.8)$		OLS	
		α	β	α	β	α	β	α	β	α	β
$S(1.5,0)$	EST	1.4712	-0.0016	1.5822	0.0903	1.5413	0.0317	1.5851	0.0726	1.5384	0.0569
	MAD	**0.1733**	**0.1883**	0.2104	0.4848	0.1996	0.4674	0.2098	0.4919	0.2028	0.4676
$S(1.5,0.5)$	EST	1.6637	0.2486	1.5862	0.5484	1.5279	0.5202	1.5663	0.4941	1.5313	0.5059
	MAD	0.2363	**0.3467**	0.1977	0.4019	0.1909	0.3834	0.1980	0.4068	**0.1740**	0.3865
$S(1.5,-0.5)$	EST	1.5122	-0.2215	1.5788	-0.3891	1.5380	-0.4295	1.5903	-0.3943	1.5466	-0.4349
	MAD	0.3210	**0.2784**	0.2015	0.4704	0.1984	0.4165	0.2037	0.4791	**0.1849**	0.4218
$S(0.7,0)$	EST	0.8511	0.0515	0.8376	0.0236	0.7128	0.0025	0.8604	-0.0091	1.8196	0.0366
	MAD	0.2527	**0.0735**	0.1619	0.2512	**0.0955**	0.1667	0.1798	0.2490	0.3877	0.4113
$S(0.7,0.5)$	EST	0.7969	0.3844	0.7274	0.5200	0.7928	0.4796	1.0624	0.4570	1.1821	0.5583
	MAD	0.0969	0.1915	**0.0948**	**0.1556**	0.1260	0.1826	0.3641	0.3176	0.4846	0.3619
$S(0.7,-0.5)$	EST	0.9904	-0.4883	1.0674	-0.4451	0.7939	-0.4700	0.7146	-0.5039	1.1684	-0.5178
	MAD	0.2904	0.2227	0.3687	0.3193	0.1288	0.1851	**0.0909**	**0.1571**	0.4716	0.3727

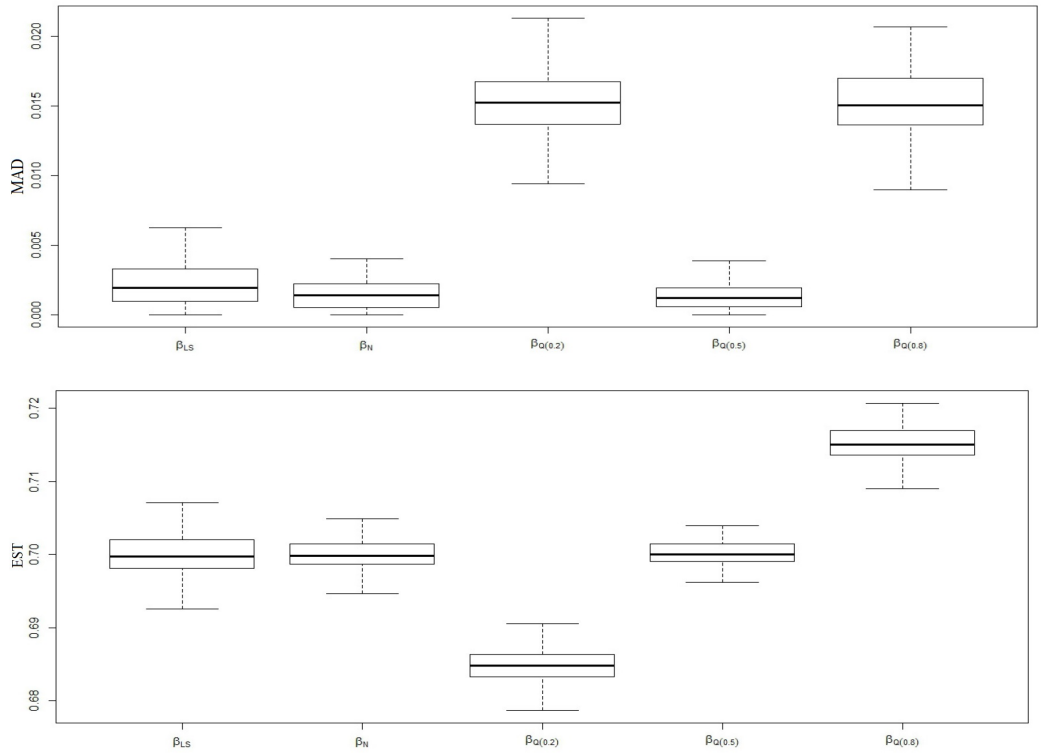

Figure 3: Box plot of MAD (top) and Estimate (bottom) of the regression coefficients in 1000 iterations when errors are simulated from $S(1.5, 0)$ and sample size $n = 200$ for different methods.

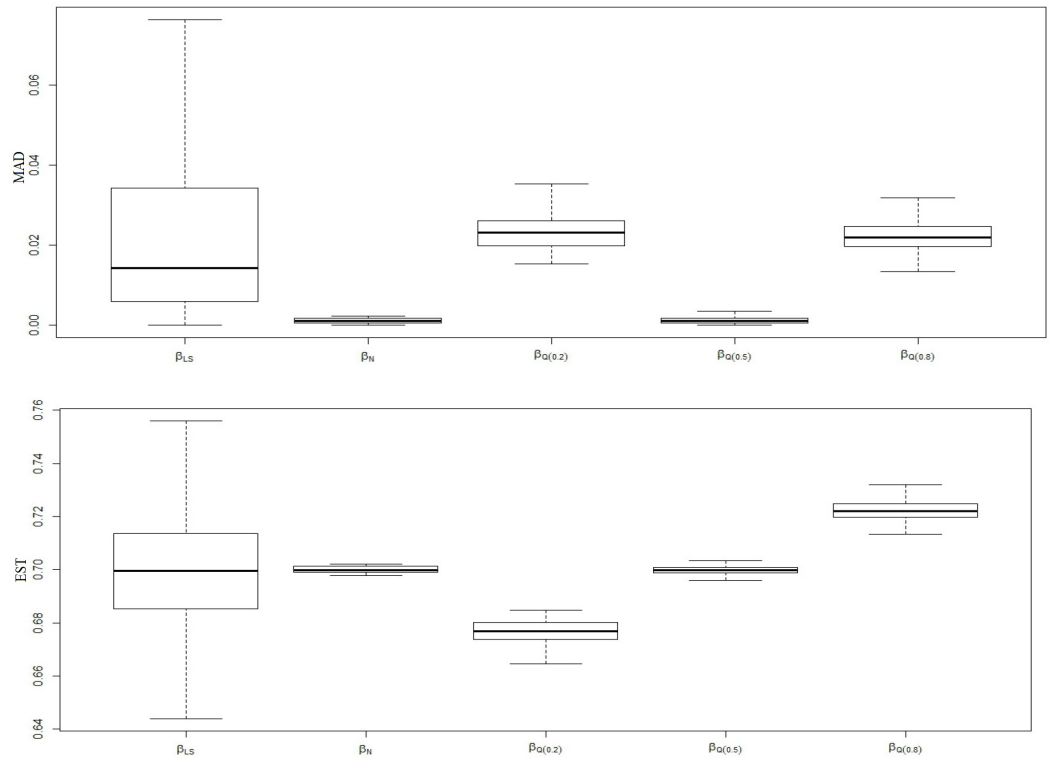

Figure 4: Box plot of MAD (top) and Estimate (bottom) of the regression coefficients in 1000 iterations when errors are simulated from $S(1, 0)$ and sample size $n = 300$ for different methods.

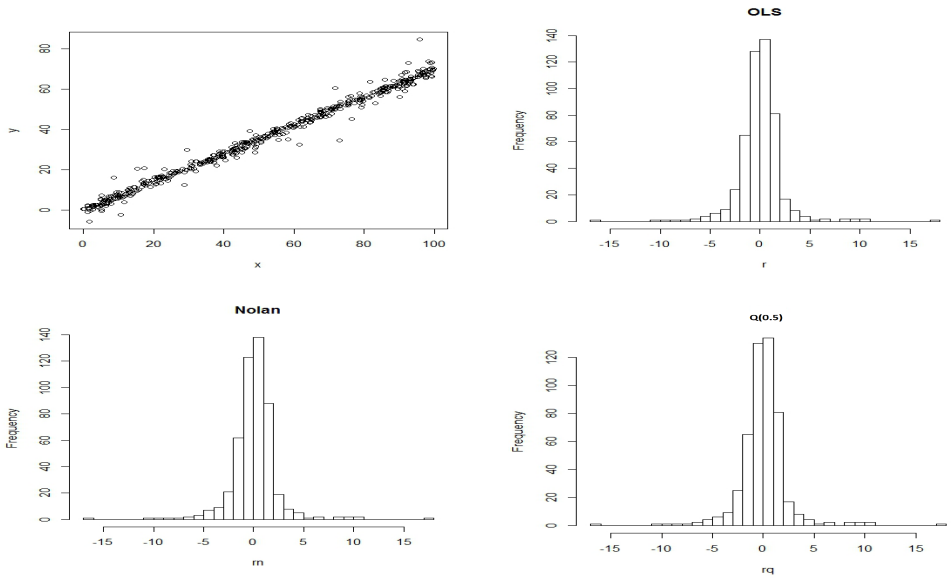

Figure 5: Data scatter plot and residual plot when residuals are simulated from $S(1.5, 0)$ with sample size $n = 100$ for different methods.

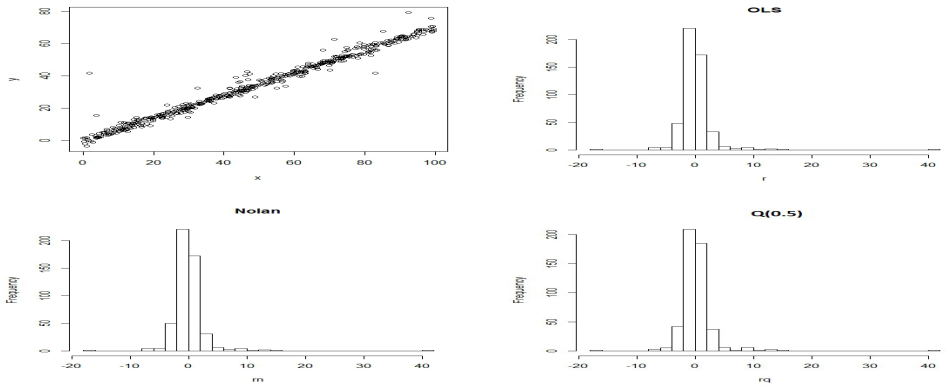

Figure 6: Data scatter plot and residual plot when errors are simulated from $S(1.5, 0.5)$ and sample size $n = 100$ for different methods.

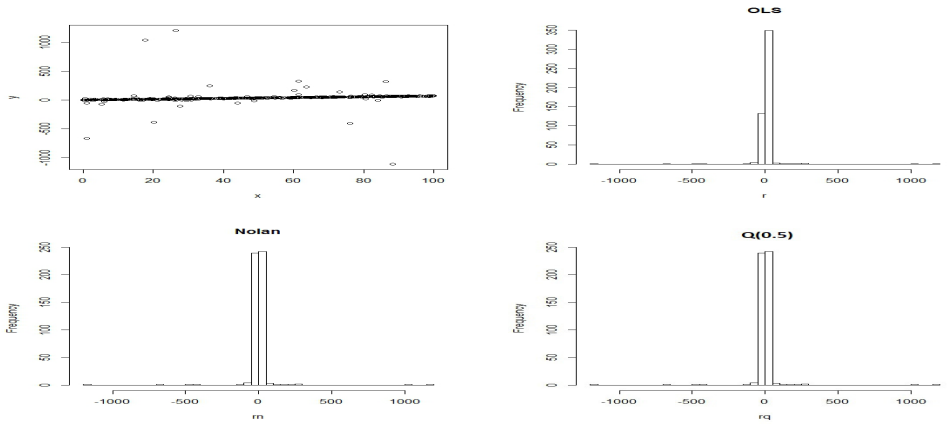

Figure 7: Data scatter plot and residual plot when errors are simulated from $S(0.7, 0)$ and sample size $n = 100$ for different methods.

To show the role of the tail index and skewness parameter of error terms in a quantile regression, we reconsider the mentioned regression model in equation (1). The OLS estimation of slope in the linear regression is computed. Furthermore, for different quantiles, the regression coefficient using quantile regression is estimated. The confidence intervals are computed and compared, see Figure 8. Recall that confidence intervals have fixed lengths in an OLS regression. One can observe that, for $\alpha \geq 1$, median is better than the other quantiles. However, for $\alpha < 1$, depends on $\beta \approx 0$, $\beta > 0$, or $\beta < 0$; median, quantiles less than 0.5, or graters than 0.5 are better, respectively.

In Figure 9, estimation of regression coefficients for different quantiles and some values of skewness and tail index are plotted. Graphs in Figure 9 confirm the previous results. In general, if most residuals are negative, we prefer the quantiles that greater than 0.5. If most of the residuals are positive, then the quantiles less than 0.5 are preferred. If the number of positive residuals and negative residuals are almost equal, then the quantiles near 0.5 are more efficient.

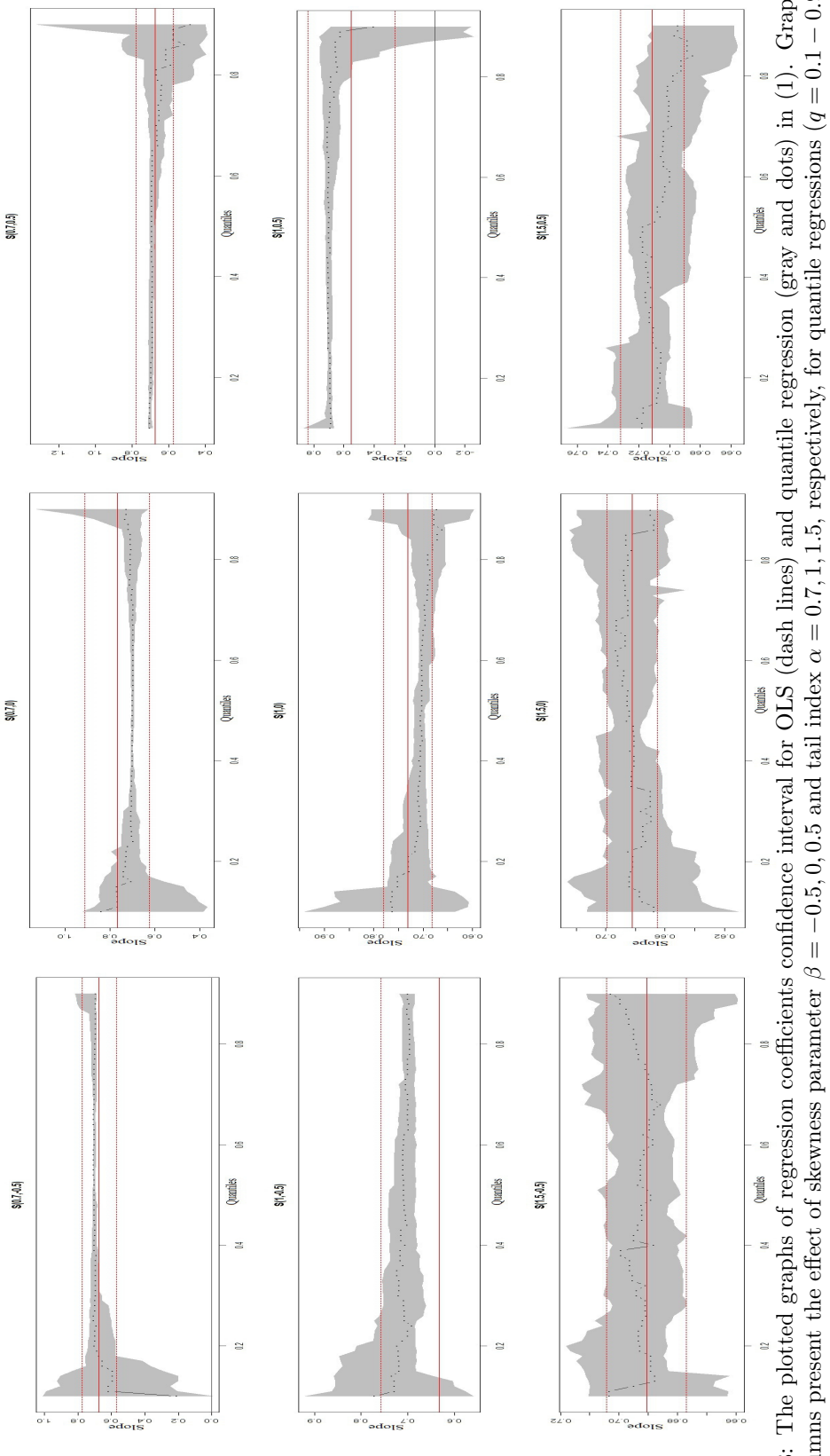

Figure 8: The plotted graphs of regression coefficients confidence interval for OLS (dash lines) and quantile regression (gray and dots) in (1). Graphs in rows and columns present the effect of skewness parameter $\beta = -0.5, 0, 0.5$ and tail index $\alpha = 0.7, 1, 1.5$, respectively, for quantile regressions ($q = 0.1 - 0.9(0.01)$).

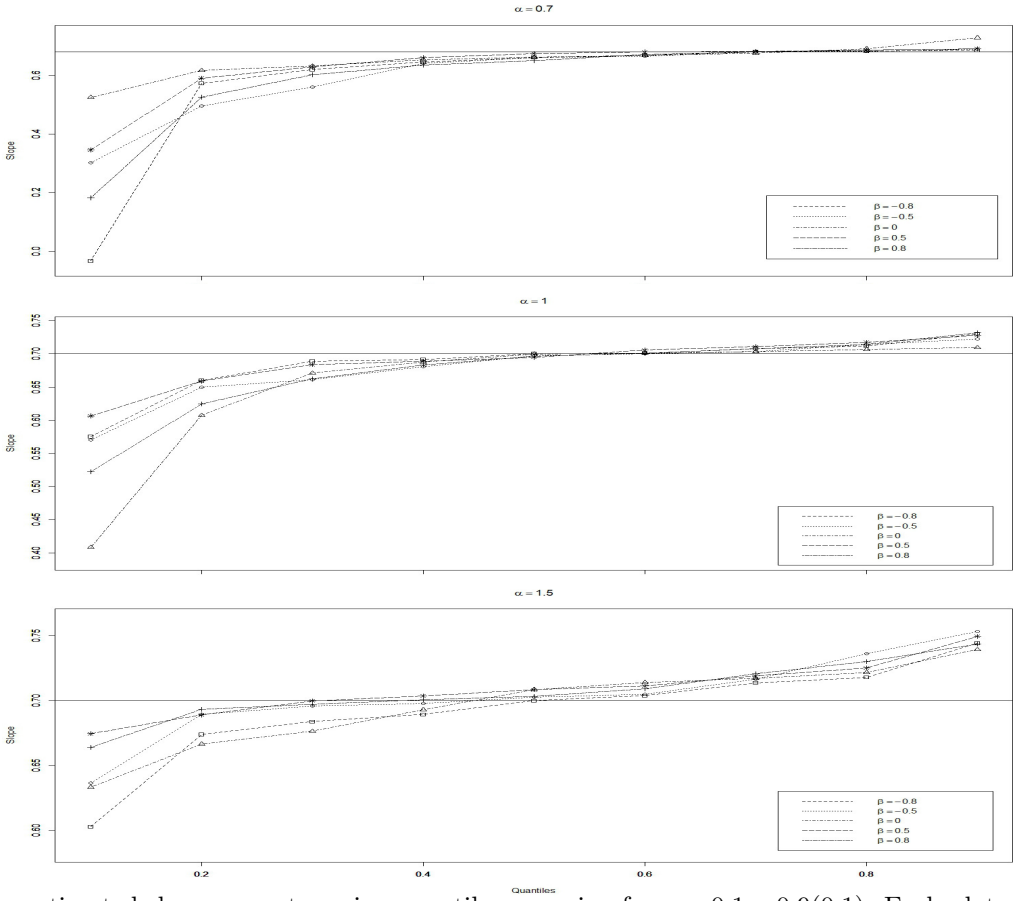

Figure 9: The estimated slop parameter using quantile regression for $q = 0.1 - 0.9(0.1)$. Each plot correspond to a tail index $\alpha = 0.7, 1, 1.5$ and each graph correspond to an skewness $\beta = -0.8, -0.5, 0, 0.5, 0.8$.

3.2. Multiple Linear Regression

In multiple linear regression, we have one response variable and several explanatory variables. Same as simple linear regression, we consider two classes of stable distributions as error terms. For the simulation study, we consider the following multiple regression model

$$\boldsymbol{Y} = 0.5\boldsymbol{X}_1 + 2\boldsymbol{X}_2 + \boldsymbol{\epsilon}, \tag{2}$$

where $\boldsymbol{X}_1$ and $\boldsymbol{X}_2$ are two random vectors of size 50, from uniform distributions in (0,100) and (200,300), respectively. Furthermore, $\boldsymbol{\epsilon}$ is a vector of error terms. Simulation results are reported in Table 5. We can estimate the parameters of stable distribution that are considered as an error term because it is similar to Table 4, we remove this table. In Tables 5 and 6, b_1 and b_2 are estimations of regression coefficients. We observe that, for estimation of coefficients, in multiple regression, we prefer quantile regression. See Tables 3 and 4.

4. Linear Regression with Non-Stable Heavy-Tailed Errors

We know that OLS is not very sensitive about tails that are heavier than a normal distribution. In Table 7, some simulation studies presented for $\boldsymbol{Y} = 0.5\boldsymbol{X} + \boldsymbol{\epsilon}$, like previous section $\boldsymbol{X}$ is a vector of size 50, that uniformly distributed in (0, 100). Also, $\boldsymbol{\epsilon}$ is a vector of heavy-tailed distribution; here we have taken the Pareto distribution. Let us recall that the Pareto distribution with shape a and scale s parameters has a density

$$f(x) = as^a \Big/ (x+s)^{(a+1)},$$

for $x > 0$, $a > 0$ and $s > 0$.

Table 5: MAD and estimation (EST) of the regression coefficients in class I in 1000 iterations. The best values are bolded.

		$S(1.5,0)$		$S(1.5,0.5)$		$S(1.5,-.5)$		$S(1,0)$		$S(1,0.5)$		$S(1,-0.5)$	
		MAD	EST	MAD	EST	MAD	EST	MAD	EST	MAD	EST	MAD	EST
NOR	b_1	0.0049	0.4978	**0.0029**	0.5029	**0.0042**	0.4986	0.0062	0.4942	0.0053	0.5020	0.0074	0.5014
	b_2	0.0012	2.0008	**0.0011**	1.9995	0.0020	1.9996	0.0029	2.0002	0.0065	2.0051	0.0096	2.0002
$Q(0.2)$	b_1	0.0065	0.4991	0.0055	0.4989	0.0077	0.4992	0.0102	0.4987	0.0052	0.4993	0.0158	0.4984
	b_2	0.0047	1.9952	0.0057	1.9942	0.0040	1.9960	0.0055	1.9944	0.0032	1.9968	0.0092	1.9907
$Q(0.5)$	b_1	**0.0048**	0.4998	0.0049	0.4997	0.0048	0.5002	**0.0043**	0.4998	**0.0052**	0.5002	0.0051	0.4997
	b_2	**0.0011**	1.9999	0.0016	1.9986	**0.0016**	2.0013	**0.0010**	2.0001	**0.0014**	2.0009	**0.0013**	1.9990
$Q(0.8)$	b_1	0.0066	0.5009	0.0080	0.5006	0.0054	0.5010	0.0102	0.5010	0.0161	0.5020	**0.0050**	0.5006
	b_2	0.0047	2.0047	0.0041	2.0040	0.0057	2.0057	0.0056	2.0056	0.0091	2.0092	0.0032	2.0032
OLS	b_1	0.0117	0.05012	0.0110	0.4999	0.0112	0.4973	0.2851	0.6465	0.1605	0.5079	0.1655	0.4700
	b_2	0.0027	1.9996	0.0028	1.9995	0.0026	2.002	0.0731	1.9551	0.0382	2.0148	0.0382	1.9886

Table 6: MAD and estimation (EST) of the regression coefficients in class II in 1000 iterations. The best values are bolded.

		$S(0.7,0)$		$S(0.7,0.5)$		$S(0.7,-0.5)$	
		MAD	EST	MAD	EST	MAD	EST
NOR	b_1	0.0132	0.5108	0.0216	0.5173	0.0062	0.4942
	b_2	0.0050	1.9966	0.0068	2.0068	0.0029	2.0002
$Q(0.2)$	b_1	0.0182	0.4978	**0.0054**	0.5001	0.0102	0.4987
	b_2	0.0073	1.9926	**0.0016**	2.0009	0.0055	1.9944
$Q(0.5)$	b_1	**0.0038**	0.50002	0.0060	0.5008	**0.0043**	0.4998
	b_2	**0.0009**	1.9999	0.0049	2.0049	**0.0010**	2.0001
$Q(0.8)$	b_1	0.0185	0.5021	0.0357	0.5043	0.0102	0.5010
	b_2	0.0073	2.0073	0.0196	2.0196	0.0056	2.0056
OLS	b_1	-	-	-	-	0.2851	0.6465
	b_2	-	-	-	-	0.0731	1.9551

Table 7: MAD and estimation (EST) of the regression coefficients with heavy-tailed errors in 1000 iterations. The best values are bolded.

	$Pa(1.5,3)$		$Pa(3,3)$		$Pa(1.5,1)$	
	EST	MAD	EST	MAD	EST	MAD
NOR	0.5200	0.2041	0.5118	0.0121	0.5093	0.0094
$Q(0.2)$	0.5077	**0.0078**	0.5036	**0.0036**	0.5025	**0.0026**
$Q(0.5)$	0.5289	0.0289	0.5126	0.0126	0.5099	0.0095
$Q(0.8)$	0.6053	0.1054	0.5376	0.0376	0.5343	0.0343
OLS	0.5862	0.8599	0.5228	0.2301	0.5309	0.0297

5. Nonlinear Regression

The methods in the previous sections can be used for nonlinear regression. In this section, we compare Nolan's method and quantile regression for nonlinear regression. The formula that is considered as follows

$$Y_i = b_1 e^{(b_2 X_i)}(1+\epsilon_i), \tag{3}$$

where X_i's are from the interval (0,20) with grids 0.025, similar to [16]. For simulation $b_1 = 30$ and $b_2 = 0.1$ are considered. Results are represented in Table 8.

Table 8: MAD and estimate (EST) of the regression parameters for $y_i = b_1 e^{(b_2 x_i)}(1+\epsilon_i)$ in 1000 iterations. The best values are bolded.

		$S(1.5,0)$		$S(1.5,0.5)$		$S(1.5,-0.5)$	
		EST	MAD	EST	MAD	EST	MAD
NOR	b_1	28.0635	6.3175	29.5640	2.7623	**28.8393**	7.2099
	b_2	0.1045	0.0117	0.1055	0.0235	0.1106	0.0203
$Q(0.5)$	b_1	30.7315	3.4977	**19.2269**	10.8852	40.8788	10.9897
	b_2	**0.0992**	0.0063	**0.1006**	0.0088	**0.1003**	0.0041
$Q(0.8)$	b_1	64.2929	34.3500	56.6853	27.2557	36.4383	12.3948
	b_2	0.1009	0.0051	01036	0.0076	0.1325	0.0480
OLS	b_1	**27.7318**	7.8714	25.6657	11.5985	36.4383	12.3948
	b_2	0.1051	0.1405	0.1188	0.0276	0.1325	-0.0480

6. Compare Confidence Intervals

In this section, we intend to compare the confidence intervals of regression coefficients for quantile regression and Nolan's method. For this mean, consider the vector $\boldsymbol{\psi} = (\alpha, \beta, \gamma, \theta_1, \ldots, \theta_k)$ and $\boldsymbol{\phi} = (\alpha, \beta, \gamma)$, where $\mathbf{J}$ is the information matrix of $\boldsymbol{\psi}$. The components of $\boldsymbol{\psi}$ is

$$\mathbf{J}_{ij} = E\left[\frac{\partial}{\partial \psi_i}\log f(\varepsilon|\boldsymbol{\phi})\,\frac{\partial}{\partial \psi_j}\log f(\varepsilon|\boldsymbol{\phi})\right].$$

so, we can write Fisher information matrix is as follows

$$\mathbf{J} = \begin{pmatrix} n\mathbf{I}_{1:3,1:3} & \mathbf{I}_{4,1:3}x_{.1}\mathbf{I}_{4,1:3}x_{.2}\cdots\mathbf{I}_{4,1:3}x_{.k} \\ (\mathbf{I}_{4,1:3}x_{.1}\mathbf{I}_{4,1:3}x_{.2}\cdots\mathbf{I}_{4,1:3}x_{.k})^T & \mathbf{I}_{4,4}\mathbf{X}^T\mathbf{X} \end{pmatrix},$$

where

$$\mathbf{I}_{1:3,1:3} = \begin{pmatrix} I_{1,1} & I_{1,2} & I_{1,3} \\ I_{2,1} & I_{2,2} & I_{2,3} \\ I_{3,1} & I_{3,2} & I_{3,3} \end{pmatrix}, \mathbf{I}_{4,1:3} = \begin{pmatrix} I_{4,1} \\ I_{4,2} \\ I_{4,3} \end{pmatrix}, x_{.l} = \sum_{j=1}^{n} x_{jl}.$$

Using Fisher information matrix, we can construct confidence intervals for the parameters.
For example, assume that $\boldsymbol{x}$ be the vector of size 500 numbers from the interval $(0, 100)$, that uniformly distributed, and $\boldsymbol{y}$ is

$$\boldsymbol{y} = 10 + 0.1\boldsymbol{x} + \boldsymbol{\varepsilon},$$

where $\boldsymbol{\varepsilon}$ is simulated from $S(1.5, 0.5)$.

Table 9: Estimation of the regression coefficients and their confidence interval length (CIL) for $\boldsymbol{y} = 10 + 0.1\boldsymbol{x} + \boldsymbol{\varepsilon}$. The best values are bolded.

	Intercept	CIL	Slop	CIL
OLS	10.4838	0.9469	0.0996	0.1261
$Q(0.5)$	10.1697	**0.3774**	0.10024	**0.0068**
NOR	9.9826	0.4293	0.10057	0.0073

Table 9 confirms the previous results and states that the confidence interval length is typically shorter in the two cases of Nolan and quantile regression.

7. Likelihood Ratio Test for Significant Regression

One way to form a test statistic is to compare the value of the likelihood functions for two hypotheses $H_0 \in \Omega_0$ versus $H_A \in \Omega_1$. Let $\Lambda(.)$ be the likelihood ratio statistic:

$$\Lambda(y) = \frac{\sup\limits_{\theta \in \Omega_0} [f_\theta(y)]}{\sup\limits_{\theta \in \Omega} [f_\theta(y)]},$$

where $\Omega = \Omega_0 \cup \Omega_1$ and we will reject H_0 if $\Lambda(x) \leq c$ for some critical value c.
Now we want to perform a significant test for regression coefficients, in other words it is desired to test the following hypotheses:

$$H_0 : \theta = 0 \quad \text{versus} \quad H_A : \theta \neq 0,$$

where θ is the regression coefficient in simple linear regression.

Table 10: Percentage that H_0 is rejected with LRT in 1000 iterations and significant level 0.1.

	$\theta = 0.1$		$\theta = 0.001$	
	NOR	OLS	NOR	OLS
$S(1.5, 0)$	**100**	99.2	**0.3**	10.4
$S(1.5, 0.5)$	**100**	99.1	**0.4**	47.8
$S(0.7, 0)$	**100**	22.7	**0.8**	7.3
$S(0.7, 0.5)$	**100**	49.2	**3.6**	26.1

Furthermore, we can build a confidence interval based on LRT. This means determining whether the test statistic based on the data falls in the critical region for various null hypothesis values for θ. Those values of $H_0 = 0$ that reject H_0 should not be in the confidence interval. Table 10 implies that Nolan's method is better than OLS in terms of LRT.

8. Conclusion

From the simulation results for linear regression, we prefer to use the quantile regression or Nolan's method and there is not a significant difference between them. For convenience, as respects existing in most statistical software, we offer quantile regression. Furthermore, in quantile regression, we are open to select a quantile according to the data that give a better answer. In the nonlinear case, usually, Nolan's method is better than the other. In multivariate regression, generalized Nolan's method is better than the multivariate OLS and it is very efficient.

Acknowledgment

The authors thank the associate editor and anonymous referees for comments that helped to improve the manuscript.

References

[1] B. Andrews, M. Calder, R. A. Davis, Maximum likelihood estimation for α-stable autoregressive processes, The Annals of Statistics, 37(4) (2009), 1946-1982.

[2] R. Blattberg, T. Sargent, Regression with non-Gaussian stable disturbances: Some sampling results, Econometrica, 39(3) (1971), 501-510.

[3] R. J. Boskovich, De literaria expeditione per pontificiam ditionem et synopsis amplioris operis, Bononiensi Scientiarum et Artum Instituto Atque Academia Commentarii, 4 (1757), 353-396.

[4] M. Buchinsky, Changes in the US wage structure 1963-1987: Application of quantile regression, Econometrica: Journal of the Econometric Society,(1994), 405-458.

[5] M. Buchinsky, Estimating the asymptotic covariance matrix for quantile regression models a Monte Carlo study, Journal of Econometrics, 68(2) (1995), 303-338.

[6] M. Buchinsky, Recent advances in quantile regression models: a practical guideline for empirical research, Journal of human resources, (1998), 88-126.

[7] M. Buchinsky, P. Leslie, Educational attainment and the changing US wage structure: dynamic implications on young individuals choices, Journal of Labor Economics, 28(3) (2010), 541-594.

[8] C. Davino, M. Furno, D. Vistocco, Quantile Regression: Theory and Applications, UK: John Wiley, 2014.

[9] S. Dorniani, A. Mohammadpour, N. Nematollahi, Raked set sampling method for estimation of linear regression model with stable errors, The Eighth World Congress in Probability and Statistics, Turkey, 2012.

[10] M. Hallin, Y. Swan, T. Verdebout, D. Veredas, One-step R-estimation in linear models with stable errors, Journal of Econometrics, 172(2) (2013), 195-204.

[11] R. Koenker, G. Bassett, Robust Tests of Heteroscedasticity based on Regression Quantiles, Econometrica, 50 (1982), 43-61.

[12] R. Koenker, Quantile regression: 40 years on, Annual Review of Economics, 9 (2017), 155-176.

[13] R. Koenker, V. Chernozhukov, X. He, L. Peng, Handbook of Quantile Regression, New York: Chapman and Hall/CRC, 2018.

[14] J. H. McCulloch, Linear regression with stable disturbances. In R. Adler, R. Feldman, and M. Taqqu, (Eds.), A Practical Guide to Heavy Tailed Data, Birkhauser, Boston, MA (1998), 359-378.

[15] J. Nolan, Stable distributions: models for heavy-tailed data, Birkhauser, 2013.

[16] J. P. Nolan, D. Ojeda-Revah, Linear and nonlinear regression with stable errors, Journal of Econometrics, 172(2) (2013), 186-194.

[17] J. Wise, Linear estimators for linear regression systems having infinite residual variances, University of Hawai, Unpublished manuscript, 1966.

Please cite this article using:

Reza Alizadeh Noughabi, Adel Mohammadpour, Comparing regression methods with non-Gaussian stable errors, AUT J. Math. Comput., 3(1) (2022) 77-91
DOI: 10.22060/AJMC.2021.20246.1062