Asian Journal of Applied Sciences (ISSN: 2321 – 0893)
Volume 02 – Issue 06, December 2014

Selection of Variables in Quantile Regression (Linear Lasso-Goal Programming)

Neveen Sayed Ahmed, Elham Abdul-Razik Ismail

Faculty of Commerce, Al-Azhar University (Girls' Branch), Egypt
Corresponding author's email: alazhar_statistic {at} hotmail.com

ABSTRACT---- ***Quantile regression is a statistical technique intended to estimate, and conduct inference about the conditional quantile functions. Since Koenker and Bassett (1978) introduced quantile regression, which models conditional quantiles as functions of predictors. The quantile regression can give complete information about the relationship between the response variable and covariates on the entire conditional distribution, and has no distributional assumption about the error term in the model. The study evaluates the performance of three methods; two methods of linear programming linear lasso (L_1-Lasso, L_2-Lasso) and one method of Goal programming. The three methods are used to select the best subset of variables and estimate the parameters of the quantile regression equation when four error distributions, with two different sample sizes and two different parameters for each error distribution. The study found that the estimated risk and relative estimated risk which produced from Goal programming method is less than ER and ERE of (L_1-Lasso and L_2-Lasso methods.***

Keywords--- Quantile Regression - Linear Lasso- Selection of Variables - goal programming - estimated risk - relative estimated risk.

1. INTRODUCTION

Koenker and Bassett (1978) introduced quantile regression, which models conditional quantiles as functions of predictors. The quantile regression model is a natural extension of the linear regression model. Quantile regression studies the relationship between the response variable and the independent variables at any quantile of the conditional distribution function. The quantile regression can give complete information about the relationship between the response variable and covariates on the entire conditional distribution, and has no distributional assumption about the error term in the model. Quantile regression is very useful for visualizing changes in the conditional distribution of longitudinal data sets over time. The quantile regression estimates are reliable in the presence of extreme outliers. Quantile regression also goes beyond the location shift model to determine the effect of covariates on the shape and scale of the entire response distribution. The spacing of the quantile lines indicates whether the distribution is skewed to the right or left. Quantile regression is a statistical technique intended to estimate, and conduct inference about the conditional quantile functions. Just as the classical linear regression methods estimate models for conditional mean function, quantile regression offers a mechanism for estimating models for conditional median function, and the full range of other conditional quantile functions.

Variable selection is an important problem in quantile regression when the number of predictor variables is large. Variable selection in linear regression is a problem of great practical importance. There are various methods for subset selection and various selection criteria. While there is no clear consensus regarding which method is the best and which criterion is the most appropriate, there are a general agreement an effective method is needed. The primary purpose of this research is to provide a review of the concepts and methods associated with variable selection in linear regression model. Some of the reasons for using only a subset of the available predictor variables given by Miller (1990) are to estimate or predict at a lower cost by reducing the number of variables on which data are to be collected, they increase the complexity and number of parameters of the model, to predict more accurately by eliminating uninformative variables, To describe multivariate data sets parsimoniously, to estimate regression coefficients with smaller standard errors (par-
ticularly when some of the predictors are highly correlated).

Lasso penalization or regularization methods are use full tools for estimating quantile regression parameters and selection variables. The Lasso reduces the variability of the estimates by shrinking the coefficients and at the same time produces interpretable models by shrinking some coefficients to exactly zero. The key strength of Lasso lies in its ability to do simultaneous parameter estimation and variable selection. The L_1 penalty and L_2 penalty was used in the Lasso for variable selection. The least square (L_2) and least absolute deviation (L_1) regression are a useful method for robust regression, and the least absolute shrinkage and selection operator Lasso is a popular choice for shrinkage estimation and variable selection.

The organization of the study is as follows: In Section 2 the study described least absolute shrinkage and selection operator which used in this study. In Section 3 the study applied simulation study to evaluate the performance of the method under consideration. The numerical results and the discussion are given in Section 4. Finally, concluding remarks are provided in Section 5.

The aim of this study is to evaluate linear lasso (L_1-Lasso, L_2-Lasso) and goal by using linear programming. The three methods can do parameter estimation and variable selection simultaneously.

2. LEAST ABSOLUTE SHRINKAGE AND SELECTION OPERATOR

Since Tibshirani (1996) proposed the least absolute shrinkage and selection operator lasso, which can effectively select important explanatory variables and estimate regression parameters simultaneously. The combination of the quantile regression and Lasso penalty is computationally easy to implement via the standard linear programming. Simulation studies are conducted to assess the finite sample performance of the proposed method. In the general linear model with independent and identically distributed errors, the least absolute deviation (LAD) or L_1 method has been a viable alternative to the least squares method especially for its superior robustness properties. Consider the linear regression model,

$$Y_i = \beta_\circ' x_i + e_i \qquad 1 \le i \le n$$

Where x_i are known p-vectors, $\beta_\circ'$ the unknown p-vector of regression coefficients, and e_i the i.i.d random errors. The L_1 estimator $\hat{\beta}_{L1}$ is defined as a minimize of the L_1 loss function

$$L_n(\beta) \sum_{i=1}^{n} |Y_i - \beta' x_i|$$

As Efron et al.(2004) the least squares estimate $\beta^{Ls} = (X'X)^{-1}X'Y$ uniquely minimizes the squared loss

$$\sum_{i=1}^{n} (Y_i - \beta' x_i)^2$$

Lasso estimate is defined as the minimize of

$$\sum_{i=1}^{n} (Y_i - \beta' x_i)^2 \qquad subject\ to \sum |\beta_j| \le s$$

Where $0 \le s \le 1$controls the amount of shrinkage that is applied to the estimates.

The study, a parallel approach borrowing is proposed the ideas from Lasso by using the L_1 penalty and L_2 penalty, but with the least squares loss replaced by the L_1 loss in quantile regression model. In doing so, we gain advantages in two fronts. First, it allows us to penetrate the difficult problem of variable selection for the L_1 regression. Appealingly, the shrinkage property of the Lasso estimator continues to hold in L_1 regression. Second, the single criterion function with both components
being of L_1-type reduces (numerically) the minimization to a strictly linear programming problem, making any resulting methodology extremely easy to implement. To be specific, our proposed estimator is a minimize of the following criterion function

$$\sum_{i=1}^{n} \rho |Y_i - \beta' x_i| + \lambda \sum_{j=1}^{n} \| \beta_j \|_1$$

and

$$\sum_{i=1}^{n} \rho |Y_i - \beta' x_i| + \lambda \sum_{j=1}^{n} \| \beta_j \|_2$$

It can be equivalently defined as a minimize of the objective function

$$\sum_{i=1}^{n} \rho |Y_i - \beta' x_i|$$

$$subject\ to \sum_{j=1}^{n} \| \beta_j \|_1 \le \lambda$$

and

$$\sum_{i=1}^{n} \rho |Y_i - \beta' x_i|$$

$$subject\ to \sum_{j=1}^{n} \parallel \beta_j \parallel_2 \leq \lambda$$

where $\parallel \beta_j \parallel_1$ is the usual L_1 estimator and $\parallel \beta_j \parallel_2$ is the usual L_2 estimator. The tuning parameter β there plays a crucial role of striking a balance between estimation of β_j and variable selection. Large values of β tend to remove variables and increase bias in the estimation aspect while small values tend to retain variables. Thus it would be ideal that a large β be used if a regression parameter is zeor (to be removed) and small values be used if it not zero. To this end, it becomes clear that we need a separate β for each parameter component β_j.

Least absolute shrinkage and selection operator (Lasso) proposed by Tibshirani (1996) is closely related to the NN Garrote. Lasso also shrinks some coefficients and sets others exactly to zero, but it does not rely on the L_2 solution.

Tibshirani (1996) noticed that the Lasso constraint $\sum_{i=1}^{n} |\beta_i| \leq t$ is equivalent to the addition of a regularization term $\lambda \sum_{i=1}^{n} |\beta_i|$ to the MSE cost function, i.e. there is a one-to-one correspondence between the parameters t and λ. However, the non differentiability of the constraints complicates the optimization in both cases.

Schmidt *et al*. (2007) reviewed and compared several optimization strategies to solve the regularized formulation for given. Osborne *et al* (2000) showed that the optimal solution trajectory of the Lasso problem is a piecewise linear as a function of t and propose an efficient algorithm to follow the trajectory.

3. SIMULATION STUDY

This section discusses the numerical simulation of the three models under consideration, operations research methods linear lasso (L_1-Lasso, L_2-Lasso) and Goal programming for estimation and selection of variables of the quantile regression model if the error distribution has heavy tailed, skewed normal distribution, and long tails. The simulation study discusses steps which applied to evaluate the performance of the three approaches L_1-Lasso, L_2-Lasso and goal programming, for selection of variables and estimation of parameters are as following:

-Quantile regression model was used as $y = x_i'\beta + \varepsilon_i$ or ($y = 3x_1 + 1.5x_2 + 0x_3 + 0x_4 + 2x_5 + 0x_6 + 0x_7 + 0x_8$), where the true value for the $\beta' s$ are set as $(3,1.5,0,0,2,0,0,0)'$. The previous value where arbitrary chosen and taking from many studies in Li, Xi and Lin (2010) and Alhamzawi et al. (2012). Where $x_i' s$ is generated from normal distribution (0, 1) during the simulation study.

-The simulation study applied $p \in (0{:}1;\ 0{:}25;\ 0{:}60;\ 0{:}95)$, depend on that the quantile nation generalizes specific terms like quartile, quintile, decile and percentile. Where p is quantile values which are arbitrary chosen from the previous specific terms.

-The study calculated the quantile regression models with intercept $\beta_\circ$ that for the intercept important in economic application but the tables omit the intercept for brevity. The study depend on Zou and Yuan, (2008) whom suggested that.

-The study when applied the (L_1-Lasso, L_2-Lasso) linear programming methods used a regularization or penalty parameter (¸ λ= 2) as a constant referred to Fan and Lv (2010).

-Where study generated ε_i from different distributions with different parameter, so as to explain the influence of the change in the error distributions on the quantile regression equation, which is the basis for choosing through the linear approaches and Goal programming.

-The Lognormal, Cauchy, Chi-square and skewed normal were employed to generate a long tailed distribution to estimate the parameters of the quantile regression model.

-Random samples of size n=50 and 100 are generated using the GAMS 2.25. The sample sizes of 50 represent moderate case while the sample size of 100 represent large once.

-This study introduced a program by using GAMS 2.25 statistical package to calculate the L_1-Lasso, L_2-Lasso, and Goal programming estimators.

-For each error distribution with two different shape parameters and for each sample size, the estimates of β_j where j = 1,………………, 8 where calculated for L_1-Lasso, L_2-Lasso, and Goal programming.

-Three suggested linear, goal method used to estimate quantile regression model with error distribution and follow each of the following four distributions and their parameters respectively,

Lognormal ~ log (0, 0.9) and (0, 1.5);
Cauchy ~ C (0, 0.9) and (0, 1.7);
Chi-square ~ χ^2 (4) and (5);
Skewed normal ~ N (0, 12) and (0, 15).

-These distributions have been generated using the above parameters that were chosen arbitrarily and taken from many previous studies to calculate performance the methods under considerations.

-In this study the estimated risk and relative estimated risk are used as criteria to compare between the solutions of L_1-Lasso, L_2-Lasso and Goal programming.

-The estimated risk and relative estimated risk of the estimators $\hat{\beta}_j$where j = 1,..., 8 are used to measure its performance, where the ER's and RER's of the estimators $\hat{\beta}_j$ for the parameters β_j the true value which suggested is defined by

$$ER(\hat{\beta}) = MSE = \sum_{r=1}^{R}(\hat{\beta} - \beta)^2 / R$$

And relative mean square errors is defined by

$$RER(\hat{\beta}) = RMSE = \sqrt{MSE(\hat{\beta})} / \beta$$

Where R is the number of repeated samples and $\widehat{\beta}$ are the estimates calculated from the sample for β (true value) for each parameter (see Ismail, 2003).

-The study applied sampling runs (number of repeated) 500 replications for each distribution with the two different parameters and two different sample sizes to be sure of consistency of the results.

-For all sample sizes, for all approaches, and for all distribution parameters for the four distributions, the ER's and RER's for each parameter β_j were calculated using each method separately.

-The criteria to evaluate the performance for the three methods under considerations depend on the approach, which produce a small ER and small ERE, for all parameters then it would be considered more suitable when the objective is to select the variables and estimate the parameters.

-Table (1) to Table (8) showed all ER's and ERE's for the three methods L_1- Lasso method, L_2-Lasso method and Goal programming method.

4. RESULTS

This section concerned with the results related with simulation study for three methods under consideration; the three methods of linear programming (L_1 Lasso, L_2-Lasso) and Goal programming. The three methods are used to select the best subset of variables and estimate the parameters of the quantile regression equation when four error distributions, with two different sample sizes and two different parameters for each error distribution.

First: The three methods for variable selection in quantile regression.

The first aim of this section is to discuss the result of the comparison between L_1-Lasso, L_2-Lasso and Goal programming estimators when it used to select best subset of variables of the quantile regression models. The three methods under consideration are used in selection of best subset of variables in quantile regression models. When the three methods are applied to select variables considering the true parameters β = (3, 1.5, 0, 0, 2, 0, 0, 0) and fixed value of quantile p = (0.10, 0.25, 0.60, 0.95*)* and four different error distributions are considered (Cauchy, skewed normal, log normal and Chi-square) with two different parameters for each distribution.

The results for the first aim demonstrated that:

-three methods under consideration are tends to produce the same coefficient that are exactly (x_3 , x_4 , x_6 , x_7 , x_8) zero and selected the same variables which coefficients that are not exactly not zero (x_1 , x_2 , x_5).

-Three suggested methods deleting the same variables and selected the same values. The estimators are calculated for each variable which selected.

-Estimated risk and relative estimated risk for L_1-Lasso L_2-Lasso and Goal programming estimators are calculated for the selection variables. Tables (1) to (8) collect the results of ER, ERE and named the variable which the three methods selected.

Second: The three methods for estimating the parameters

The second aim of this section is to discuss the results of the comparison between L_1-Lasso, L_2-Lasso and Goal programming estimators when the three methods under consideration are used to estimate the parameters for quantile regression models. Estimator which produced by L_1-Lasso, L_2-Lasso and Goal programming used to calculate the ER and RER. The results for the second aim as follows:

-Estimated risk and relative estimated risk which produced by Goal programming is less than ER and RER which calculated by L_1-Lasso, L_2- Lasso.

Third: The three methods with the quantile regression values

The third aim of this section is to discuss the result of the comparison between L_1-Lasso, L_2-Lasso and goal programming estimators when different value of quantiles are used to estimate the parameters and selection of variables of the quantile regression models. When the values of quantile are p = (0.1, 0.25, 0.60, 0.95). When four different error distributions are used, with two sample size and each distribution have two different parameters. Table (1) to (8) appears the quantile regression values which used in this study. One, major advantage of quantile regression over classical mean regression is its exibility in assessing the effect of predictors on different locations of the response distribution.

Regression at multiple quantiles provides more comprehensive statistical views than analysis is at mean or at single quantile level. When the distribution is highly skewed, the mean can be challenging to interpret while the median remains highly informative.

The study showed ER and ERE with values of quantile regression (0.1, 0.95) is less than ER and ERE with values (0.25, 0.60) as shown in all tables. More generally, values of quantile regression (0.1, 0.2, 0.60, and 0.95) can be used to describe non-central positions of a distribution.

Fourth: The three methods with different sample size

The fourth aim of this section is to discuss the results of suggested three methods when different sample sizes are applied. The simulation study applied L_1-Lasso, L_2-Lasso and goal programming when two different sample sizes are generated when four different error distributions are used with two different parameters for quantile regression models. The simulation study generated two sample sizes (n=50, n=100) to explain if the size of sample effect or not about the estimators and the behavior of the three methods with different samples sizes.

The results for the fourth aim shown that the estimators values which used to calculate ER and RER are improving with the increase in sample size when the (ε_i) generated fat-long tailed distribution.

In case of sample size (n=50) with respect to L_1-Lasso and L_2-Lasso the results of ER and ERE are almost the same. While ER and ERE for the Goal programming method is less than ER and ERE of L_1-Lasso and L_2-Lasso. In sample size (n=100) the results of L_1-Lasso better than L_2- Lasso, also the results of Goal programming better than of both L_1-Lasso and L_2-Lasso.

Fifth: Three methods with different distribution

The fifth aim of this section is to discuss the results of the comparison between L_1-Lasso, L_2-Lasso and goal programming when different distribution with different parameters is used. The study considered with four distributions (fat-long tailed, skewed normal and chi- square distribution) for each distribution two different parameters. The simulation study applied with four different distribution and two different parameters for each to evaluate the performance for the three methods with existence of fat or long tailed distribution. If the estimate model effected by the existence of fat-long tailed distribution. The results for this aim are:

Cauchy distribution: From the results in Table (1) and (2) with two different sample sizes and two different parameters (0, 0.9), (0, 1.7) for Cauchy distribution. This study observed that the results of ER and ERE for L_1-Lasso, L_2-Lasso are almost the same. While ER and ERE for the Goal programming method is less than ER and ERE of L_1-Lasso, L_2- Lasso. Goal programming for estimation of parameters when Cauchy distribution is used much better than the L_1-Lasso, L_2- Lasso.

skewed normal distribution: Tables (3) and (4) showed the results when two different samples sizes, with two different parameter for skewed normal distribution, the results demonstrated that the ER and RER for Goal programming method is less than ER and RER for L_1-Lasso, L_2-Lasso. Goal programming for estimation of parameters when skewed normal distribution is used much better than linear programming.

Log normal distribution: Tables (5) and (6) showed the results when two sample sizes (n=50, n=100) with two different parameters (0, 0.9) and (0, 1.5) for Log normal distribution. The results demonstrated that when two different sample sizes with parameter (0, 09) ER and RER are different and the results indicated that L_2-Lasso, *is* better than L_1-Lasso. When parameter (0, 1.5) the results of ER and RER are different and the results indicated that L_2-Lasso is better than L_1-Lasso. While ER and RER for the Goal programming method is less than of both L_1-Lasso, L_2-Lasso. Goal programming method is better than the L_1-Lasso, L_2-Lasso for estimation of parameters when log normal distribution is used.

Chi-square distribution: From the results in Tables (7) and (8) with two sample sizes (n=50, n=100) with two different degree of freedom (4) and (5) for Chi-square distribution. This study observed that the results of ER and RER in degree of freedom (4) L_1-Lasso is less than L_2-Lasso. In degree of freedom (5) the results are almost the same in L_1-Lasso, L_2-Lasso. While the Goal programming is better than both L_1-Lasso, L_2-Lasso. Goal programming is better than the linear programming and for estimation of parameters when Chi-square distribution is used.

In this study, the three methods for estimating quantile regression parameter through Lasso linear programming and Goal programming are proposed. A simulation study has been made to evaluate the performance of the proposed estimators based on the estimated risk (ER) criterion and the relative estimated risk (RER). Quantile regression is an approach that allows us to examine the behavior of the response variable beyond its average of the Gaussian distribution, e.g., median (50th percentile), 10th percentile, 25th percentile, 60th percentile, and 95th percentile which applied in this study. Examining the different percentiles using quantile regression may be more beneficial for continuous improvement and cost savings.

Lasso quantile regression is a regularization technique for simultaneous estimation and variable selection where the classical variable selection methods are often highly time consuming and maybe suffer from instability. L_1 and L_2 penalized estimation methods shrink the estimates of the regression coefficients towards zero relative to the maximum likelihood estimates. The purpose of this shrinkage is to prevent over fit arising due to either collinearity of the covariates or high-dimensionality.

The study evaluates the performance for the three methods; Lasso linear (L_1-Lasso, L_2-Lasso) and Goal programming. The three methods are used to select the best subset of variables and estimate the parameters of the quantile regression equation when four error distributions, with two sample sizes and two different parameters for each error distribution. Estimated risk and relative estimated risk for which produced from Goal programming method is less than ER and ERE of L_1-Lasso, L_2-Lasso methods.

All results showed superiority of Goal method compared with L_1-Lasso, L_2-Lasso linear programming methods.

Table (1) Estimated risks of the $\boldsymbol{L_1}$-Lasso, $\boldsymbol{L_2}$-Lasso and Goal programming Estimates for Cauchy (0, 0.9)

			N=50			N=100		
			L_1-Lasso	L_2-Lasso	GP	L_1-Lasso	L_2-Lasso	GP
			0.350	0.350	0.019	0.098	0.099	0.010
			0.091	0.091	0.006	0.032	0.033	0.004
			0.202	0.202	0.010	0.046	0.047	0.005
			0.197	0.197	0.046	0.104	0.105	0.033
			0.201	0.201	0.053	0.120	0.121	0.039
			0.225	0.225	0.044	0.108	0.108	0.034
			0.357	0.357	0.042	0.099	0.099	0.042
			0.073	0.073	0.012	0.033	0.033	0.012
			0.201	0.201	0.020	0.046	0.046	0.020
			0.199	0.199	0.069	0.105	0.105	0.069
			0.180	0.180	0.074	0.121	0.121	0.074
			0.224	0.224	0.071	0.107	0.107	0.071
			0.793	0.793	0.084	0.099	0.099	0.084
			0.083	0.083	0.023	0.036	0.036	0.023
			0.150	0.150	0.037	0.044	0.044	0.037
			0.297	0.297	0.097	0.105	0.105	0.097
			0.192	0.192	0.100	0.126	0.126	0.100
			0.193	0.193	0.096	0.105	0.105	0.096
			0.347	0.347	0.014	0.098	0.098	0.014
			0.106	0.106	0.006	0.032	0.032	0.006
			0.203	0.203	0.008	0.045	0.045	0.008
			0.196	0.196	0.040	0.104	0.104	0.040
			0.218	0.218	0.050	0.119	0.119	0.050
			0.226	0.226	0.044	0.107	0.107	0.044

Table (2) Estimated risks of the $\boldsymbol{L_1}$-Lasso, $\boldsymbol{L_2}$-Lasso and Goal programming Estimates for Cauchy (0, 1.7)

			N=50			N=100		
			L_1-Lasso	L_2-Lasso	GP	L_1-Lasso	L_2-Lasso	GP
			0.743	0.762	0.029	0.098	0.099	0.015
			0.204	0.209	0.010	0.053	0.054	0.006
			0.488	0.500	0.016	0.045	0.046	0.008
			0.287	0.291	0.057	0.104	0.105	0.041
			0.301	0.305	0.067	0.154	0.155	0.051
			0.349	0.353	0.063	0.106	0.107	0.044
			0.773	0.773	0.051	0.100	0.101	0.028
			0.129	0.129	0.015	0.061	0.062	0.009
			0.492	0.492	0.025	0.046	0.047	0.013
			0.293	0.293	0.076	0.106	0.106	0.056
			0.239	0.239	0.081	0.165	0.166	0.064
			0.351	0.351	0.079	0.108	0.108	0.058
			0.233	0.239	0.086	0.102	0.103	0.051
			0.153	0.157	0.023	0.013	0.073	0.014
			0.307	0.315	0.039	0.046	0.046	0.022
			0.509	0.516	0.098	0.106	0.107	0.075
			0.261	0.264	0.100	0.180	0.181	0.079
			0.277	0.281	0.099	0.107	0.107	0.074
			0.677	0.826	0.025	0.088	0.088	0.012
			0.264	0.327	0.009	0.058	0.058	0.005
			0.692	0.844	0.013	0.312	0.312	0.007
			0.274	0.303	0.052	0.099	0.099	0.037
			0.345	0.382	0.064	0.160	0.160	0.047
			0.416	0.459	0.058	0.279	0.279	0.040

Table (3) Estimated risks of the $\boldsymbol{L_1}$-Lasso, $\boldsymbol{L_2}$-Lasso and Goal programming Estimates for Skewed Normal (0, 12)

			N=50			N=100		
			L_1-Lasso	L_2-Lasso	GP	L_1-Lasso	L_2-Lasso	GP
			0.223	0.223	0.087	0.231	0.104	0.049
			0.071	0.071	0.025	0.067	0.030	0.014
			0.098	0.098	0.042	0.099	0.045	0.022
			0.157	0.157	0.098	0.160	0.107	0.074
			0.177	0.177	0.106	0.172	0.116	0.079
			0.157	0.157	0.102	0.157	0.106	0.074
			0.212	0.212	0.122	0.224	0.101	0.070
			0.074	0.074	0.029	0.076	0.034	0.017
			0.102	0.102	0.052	0.105	0.047	0.029
			0.153	0.153	0.117	0.158	0.106	0.088
			0.182	0.182	0.114	0.183	0.123	0.088
			0.160	0.160	0.114	0.162	0.109	0.085
			0.206	0.206	0.141	0.216	0.097	0.084
			0.091	0.091	0.033	0.096	0.043	0.019
			0.117	0.117	0.058	0.118	0.053	0.034
			0.151	0.151	0.125	0.155	0.104	0.096
			0.201	0.201	0.121	0.206	0.138	0.092
			0.171	0.171	0.121	0.172	0.115	0.092
			0.219	0.219	0.077	0.229	0.103	0.043
			0.074	0.074	0.023	0.068	0.031	0.013
			0.098	0.098	0.039	0.102	0.046	0.020
			0.156	0.156	0.092	0.160	0.107	0.069
			0.182	0.182	0.102	0.174	0.117	0.076
			0.157	0.157	0.099	0.160	0.107	0.071

Table (4) Estimated risks of the $\boldsymbol{L_1}$-Lasso, $\boldsymbol{L_2}$-Lasso and Goal programming Estimates for Skewed Normal (0, 15)

			N=50			N=100		
			L_1-Lasso	L_2-Lasso	GP	L_1-LASS0	L_2-LASS0	GP
			0.225	0.225	0.044	0.231	0.104	0.057
			0.088	0.088	0.018	0.076	0.034	0.015
			0.109	0.109	0.026	0.107	0.048	0.024
			0.158	0.158	0.070	0.160	0.107	0.079
			0.198	0.198	0.088	0.184	0.123	0.081
			0.165	0.165	0.081	0.164	0.110	0.077
			0.208	0.208	0.132	0.215	0.097	0.078
			0.089	0.089	0.031	0.098	0.044	0.019
			0.115	0.115	0.056	0.117	0.053	0.032
			0.152	0.152	0.121	0.154	0.104	0.093
			0.199	0.199	0.118	0.209	0.140	0.091
			0.170	0.170	0.118	0.171	0.115	0.089
			0.207	0.207	0.141	0.217	0.098	0.086
			0.127	0.127	0.033	0.135	0.061	0.020
			0.144	0.144	0.060	0.139	0.063	0.035
			0.152	0.152	0.125	0.155	0.104	0.098
			0.238	0.238	0.121	0.245	0.164	0.093
			0.189	0.189	0.122	0.186	0.125	0.093
			0.226	0.226	0.088	0.231	0.104	0.088
			0.084	0.084	0.025	0.072	0.033	0.025
			0.111	0.111	0.043	0.113	0.051	0.043
			0.159	0.159	0.099	0.160	0.108	0.099
			0.194	0.194	0.104	0.179	0.120	0.104
			0.167	0.167	0.104	0.168	0.113	0.104

Table (5) Estimated risks of the $\boldsymbol{L_1}$-Lasso, $\boldsymbol{L_2}$-Lasso and Goal programming Estimates for Log Normal (0, 0.9)

			N=50			N=100		
			L_1-Lasso	L_2-Lasso	GP	L_1-Lasso	L_2-Lasso	GP
			0.223	0.218	0.010	0.102	0.102	0.004
			0.057	0.053	0.004	0.027	0.027	0.001
			0.104	0.095	0.006	0.046	0.046	0.002
			0.158	0.155	0.033	0.106	0.106	0.022
			0.159	0.154	0.043	0.110	0.110	0.025
			0.161	0.154	0.039	0.107	0.107	0.022
			0.221	0.221	0.025	0.099	0.099	0.014
			0.055	0.055	0.007	0.025	0.025	0.004
			0.098	0.098	0.012	0.044	0.044	0.006
			0.157	0.157	0.023	0.105	0.105	0.039
			0.157	0.157	0.057	0.106	0.106	0.041
			0.156	0.156	0.055	0.105	0.105	0.039
			0.220	0.220	0.059	0.099	0.099	0.033
			0.054	0.054	0.017	0.024	0.024	0.009
			0.097	0.097	0.027	0.043	0.043	0.015
			0.157	0.157	0.081	0.105	0.105	0.061
			0.154	0.154	0.087	0.103	0.103	0.063
			0.155	0.155	0.083	0.104	0.104	0.062
			0.101	0.101	0.009	0.101	0.101	0.004
			0.026	0.028	0.004	0.028	0.028	0.002
			0.046	0.047	0.006	0.047	0.047	0.002
			0.106	0.106	0.032	0.106	0.106	0.020
			0.108	0.111	0.044	0.111	0.111	0.026
			0.107	0.109	0.039	0.109	0.109	0.022

Table (6) Estimated risks of the $\boldsymbol{L_1}$-Lasso, $\boldsymbol{L_2}$-Lasso and Goal programming Estimates for Log Normal (0, 1.5)

			N=50			N=100		
			L_1-Lasso	L_2-Lasso	GP	L_1-Lasso	L_2-Lasso	GP
			0.222	0.223	0.015	0.098	0.103	0.006
			0.058	0.058	0.007	0.022	0.027	0.003
			0.105	0.104	0.010	0.049	0.046	0.003
			0.105	0.151	0.041	0.105	0.107	0.026
			0.100	0.160	0.056	0.100	0.109	0.034
			0.111	0.161	0.049	0.111	0.107	0.029
			0.098	0.223	0.024	0.098	0.101	0.013
			0.022	0.055	0.009	0.022	0.025	0.004
			0.049	0.098	0.013	0.049	0.044	0.006
			0.105	0.157	0.052	0.105	0.106	0.038
			0.100	0.156	0.062	0.100	0.105	0.040
			0.111	0.157	0.056	0.111	0.105	0.038
			0.222	0.222	0.053	0.098	0.099	0.030
			0.057	0.054	0.015	0.022	0.024	0.008
			0.105	0.097	0.025	0.049	0.043	0.013
			0.157	0.156	0.077	0.105	0.105	0.057
			0.154	0.154	0.082	0.100	0.103	0.060
			0.162	0.156	0.079	0.111	0.104	0.058
			0.222	0.225	0.015	0.098	0.104	0.006
			0.058	0.057	0.008	0.022	0.028	0.003
			0.105	0.101	0.011	0.049	0.047	0.004
			0.157	0.158	0.041	0.105	0.107	0.026
			0.160	0.160	0.059	0.100	0.111	0.038
			0.162	0.159	0.051	0.111	0.109	0.030

Table (7) Estimated risks of the $\boldsymbol{L_1}$-Lasso, $\boldsymbol{L_2}$-Lasso and Goal programming Estimates for Chi Square (0, 4)

			N=50			N=100		
			L_1-Lasso	L_2-Lasso	GP	L_1-Lasso	L_2-Lasso	GP
		β	0.227	0.223	0.003	0.102	0.101	0.017
		β	0.056	0.056	0.002	0.029	0.026	0.008
		β	0.102	0.102	0.002	0.047	0.047	0.011
		β	0.159	0.157	0.018	0.106	0.106	0.044
		β	0.158	0.158	0.028	0.113	0.108	0.059
		β	0.160	0.160	0.022	0.108	0.109	0.053
		β	0.220	0.221	0.031	0.100	0.100	0.018
		β	0.054	0.055	0.014	0.025	0.025	0.006
		β	0.099	0.100	0.019	0.044	0.045	0.010
		β	0.156	0.157	0.059	0.106	0.105	0.044
		β	0.156	0.157	0.080	0.106	0.105	0.053
		β	0.157	0.158	0.070	0.105	0.106	0.049
		β	0.220	0.220	0.046	0.099	0.099	0.027
		β	0.052	0.055	0.015	0.024	0.024	0.008
		β	0.095	0.097	0.023	0.043	0.043	0.012
		β	0.156	0.156	0.071	0.105	0.105	0.055
		β	0.152	0.157	0.082	0.104	0.104	0.058
		β	0.154	0.156	0.076	0.104	0.104	0.056
		β	0.223	0.225	0.034	0.104	0.103	0.018
		β	0.053	0.055	0.018	0.028	0.026	0.008
		β	0.099	0.099	0.026	0.049	0.047	0.012
		β	0.157	0.158	0.062	0.107	0.107	0.045
		β	0.153	0.156	0.089	0.112	0.108	0.061
		β	0.158	0.158	0.081	0.110	0.108	0.055

Table (8) Estimated risks of the $\boldsymbol{L_1}$-Lasso, $\boldsymbol{L_2}$-Lasso and Goal programming Estimates for Chi Square (0, 5)

			N=50			N=100		
			L_1-Lasso	L_2-Lasso	GP	L_1-Lasso	L_2-Lasso	GP
			0.224	0.224	0.047	0.102	0.102	0.024
			0.056	0.056	0.019	0.027	0.027	0.011
			0.102	0.102	0.028	0.047	0.047	0.014
			0.158	0.158	0.072	0.106	0.106	0.051
			0.158	0.158	0.092	0.109	0.109	0.69
			0.160	0.160	0.084	0.108	0.108	0.058
			0.221	0.221	0.042	0.100	0.100	0.021
			0.055	0.055	0.017	0.026	0.026	0.009
			0.097	0.097	0.024	0.045	0.045	0.011
			0.157	0.157	0.068	0.105	0.105	0.049
			0.156	0.156	0.088	0.107	0.107	0.063
			0.156	0.156	0.078	0.106	0.106	0.053
			0.220	0.220	0.048	0.099	0.099	0.027
			0.052	0.052	0.017	0.024	0.024	0.008
			0.095	0.095	0.025	0.044	0.044	0.012
			0.156	0.156	0.073	0.105	0.105	0.055
			0.153	0.153	0.086	0.104	0.104	0.061
			0.154	0.154	0.079	0.105	0.105	0.055
			0.224	0.224	0.049	0.102	0.102	0.025
			0.056	0.056	0.020	0.28	0.28	0.011
			0.100	0.100	0.030	0.047	0.047	0.015
			0.158	0.158	0.074	0.106	0.106	0.053
			0.157	0.157	0.094	0.111	0.111	0.071
			0.158	0.158	0.086	0.108	0.108	0.060

5. CONCLUSION

1- The three methods are used to select the best subset of variables. Three suggested methods deleting the same variables that lead to the motivation for variable selection that the deleting variables from the model can improve the precision of parameter estimates.
2- The three methods are used to estimate the parameters of the quantile regression equation. Estimated risk and relative estimated risk are used to measure the performance of the methods. Goal programming i much better than L1-Lasso, L2-Lasso methods.
3- Three methods are used to select the best subset of variables and to estimate the parameters with different quantile regression values. Different quantile regression may be more beneficial for continuous improvement and cost savings.
4- The performance for the three methods when are used to select the best subset of variables, to estimate the parameters much better with different sample sizes.
5- The performance for the three methods when are used to select the best subset of variables, to estimate the parameters with four error distributions, with two different parameters for each error distribution much better with fat-long tailed distribution.

6. REFERENCES

[1] Alhamzawi, R., Yu, K. and Benoit, T. (2012). Bayesian adaptive lasso quantile regression. Statistical Modelling, 12(3) , 279-297.
[2]Efron, B., Hastie, T., Johnstone, I. and Tibshirani, R. (2004). Least angle regression. *The Annals of Statistics*, 32, 407-499.
[3] Fan, J. and Lv, J. (2010). Aselective overview of variable selection in high dimensional feature space. Statistica Sinica, 20, 101-148.
[4] Ismail, A. R. (2003). Curve fitting for data with outliers using fuzzy linear programming. Unpublished Ph.D. thesis, Faculty of Commerce, Al-Azhar University-Girls' Branch.
[5] Koenker, R. and Bassett, G. (1978). Regression quantiles. Econometrica, 46, 33-50.
[6] Li, Q., Xi, R. and Lin, N. (2010). Bayesian regularized quantile regression. Bayesian Analysis, 5, 1-24.
[7] Li, Y. and Zhu, J. (2008). L1-Normal quantile regression. Journal of Computational and Graphical Statistics, 17, 163-185.
[8] Miller, A. (1990). Subset selection regression. Chapmanand Hall.
[9] Osborne, M., Presnell, B. and Turlach, B. (2000). On the Lasso and its dual. Journal of Computational and Graphical Statistics, 9, 319-337.
[10] Schmidt, E., Berg, M., Fried, L. and Murphy, k. (2007). Group sparsity via linear time projection. Technical report. TR, Department of Computer Science, University of British Columba, Vancouver, July.
[11] Tibshirani, R. (1996). Regression shrinkage and selection via the Lasso. Journal of the Royal Statistical Society, Series B, 58, 267-288.
[12] Zou, H. and Yuan, M. (2008). Regularized simultaneous model selection in multiple quantiles regression. Computational Statistics and Data Analysis, 52, 5296-5304.