International Journal of Mathematical Archive-9(11), 2018, 55-63
Available online through www.ijma.info ISSN 2229 – 5046

COMPARISON BETWEEN M-ESTIMATION, S-ESTIMATION, AND MM ESTIMATION METHODS OF ROBUST ESTIMATION WITH APPLICATION AND SIMULATION

EHAB MOHAMED ALMETWALLY*
Demonstrator of Statistics,
Higher Institute of Computer and Management Information Systems.

HISHAM MOHAMED ALMONGY
Lecturer of Applied Statistics, Faculty of Commerce, Mansoura University

(Received On: 08-10-18; Revised & Accepted On: 26-11-18)

ABSTRACT

In regression analysis the use of ordinary least squares, (OLS) method would not be appropriate in solving problem containing outlier or extreme observations. Therefore, we need a method of robust estimation where the value of the estimation is not much affected with these outlier or extreme observations. In this paper, six methods of estimation will be compared in order to reach the best estimation, and these methods are M.Humpel estimation method, M.Bisquare estimation method, M.Huber estimation method, S-estimation method, MM(S)-estimation method, and MM estimation method in robust regression to determine a regression model. We find that, the best three method, through this study, are M-estimation method, MM(S)-estimation method and MM estimation method. Since M-estimation method is an extension of the maximum likelihood method, while MM estimation method is the development of M-estimation method and MM(S) estimation method is the development of S-estimation method. Robust regression methods can considerably improve estimation precision, but should not be applied automatically instead of the classical methods.

Keywords: *Ordinary Least Squares, Robust Estimation, M-estimation, S-estimation, MM estimation and Monte Carlo simulation.*

1. INTRODUCTION

Zioutas et al (2005), Discussed linear regression models are commonly used to analyze data from many fields of study. These data often contain outliers and influential observations. Hample (2002) introduced the alternative methods to ordinary least squares (OLS), which are known as "Robust Regression". Robust regression analysis provides good alternative method of a least squares regression model, when fundamental assumptions are unfulfilled by the nature of the data. Robust methods have been defined to deal with the influential points in regression analysis. Hample (2011) introduced robust inference is more precision, because it is insensitive to (smaller or larger) deviations from the assumptions under which it is derived. Some very commonly used assumptions in statistics are normality, independence, identical distributions, linearity, and stationary of stochastic processes. Almongy and Almetwaly (2018) discussed comparisons between the method of Least Absolute Deviations (LAD) estimation, the method of Least Median of Squares (LMS) estimation, the method of Least Quantile of Squares (LQS) estimation, the method of Least Trimmed Squares (LTS) estimation, the method of Reweighted Least Squares (LTS.RLS) estimation, the method of M.Huber (MH) estimation and the method of S-estimation in robust regression to determine a suitable regression model. As a rule, such assumptions are only approximations to reality, and the questions arise what deviations tend to occur in practice, what effects they have unknown statistical procedures, and how to develop better, “more robust” procedures? In this paper, this question will be answered by introducing alternative methods of robust estimation and more comparison between M-estimation method, S-estimation method, MM estimation method and MM(S) estimation method in robust regression.

Corresponding Author: Ehab Mohamed Almetwally*
Demonstrator of Statistics, Higher Institute of Computer and Management Information Systems.

2. THE LEAST SQUARES METHOD

Consider the standard linear regression model:

$$Y = X\beta + \varepsilon. \tag{1}$$

$$\begin{pmatrix} y_1 \\ \vdots \\ y_n \end{pmatrix} = \begin{pmatrix} x_{11} & \cdots & x_{1k} \\ \vdots & \ddots & \vdots \\ x_{n1} & \cdots & x_{nk} \end{pmatrix} \begin{pmatrix} \beta_0 \\ \vdots \\ \beta_k \end{pmatrix} + \begin{pmatrix} \varepsilon_1 \\ \vdots \\ \varepsilon_n \end{pmatrix} \tag{2}$$

Regression analysis aims to find the best relationship between one or more independent variables and a dependent variable. The method of least squares is one of the oldest techniques, the least square methods (LS) is probably the most popular technique in statistical methods. Abdi (2007) discussed the resulted OLS estimators which have unbiased, minimum variance, minimum mean square error, efficiency and best linear unbiased estimator (BLUE). The Least Squares (LS) is widely used to estimate the numerical values of the parameters to a function, OLS is called ordinary least squares (OLS), which defined as:

$$\hat{\beta}_{OLS} = (X'X)^{-1}X'Y \tag{3}$$

The OLS estimates for regression models are highly sensitive to (not robust against) outliers. So that is no precise definition of an outlier, outliers are observations which do not follow the pattern of the other observations. This is not normally a problem when the outlier observation is simply an extreme observation drawn from the tail of a normal distribution; however, if the outlier results from non-normal measurement error or some other violation of standard OLS assumptions, then it compromises the validity of the regression results when a non-robust estimation technique is used. \

3. ROBUST ESTIMATION METHOD

When the data are contaminated with a single or few outliers, the problem of identifying such observations is serious problem. We note that, in most cases data sets contain more outliers or a group of influential observations. Alma (2011) discussed robust estimation is an important method for analyzing data that are contaminated with outliers, robust estimation method is a form of regression analysis designed to circumvent some limitations of traditional parametric and non-parametric methods, Robust estimation methods are designed to be not overly affected to outliers. Under these conditions, robust regression is resistant to the influence of outliers is the best method. Therefore, we introduce a comparison between robust methods to get the best method.

3.1 M-Estimation Method

Fox (2002) discussed the most common general method of robust regression is M-estimation method, introduced by Huber (1964), the method of M-estimation method as a generalization to maximum likelihood estimation in context of location models. That is nearly as efficient as OLS. Rather than minimizing the sum of squared errors, as the objective, M-estimation method principle is minimizing the residual function. The M-estimate objective function is:

$$\hat{\beta}_M = min \sum_{i=1}^{n} \rho\left(y_i - \sum_{j=0}^{k} x'_{ij}\beta_j \right) \tag{4}$$

we have to solve

$$min \sum_{i=1}^{n} \rho(u_i) = \min \sum_{i=1}^{n} \rho\left(\frac{e_i}{\hat{\sigma}_{MAD}}\right) = min \sum_{i=1}^{n} \rho\left(\frac{y_i - \sum_{j=0}^{k} x'_{ij}\beta_j}{\hat{\sigma}_{MAD}}\right) \tag{5}$$

$$\hat{\sigma}_{MAD} = \frac{median|e_i - median(e_i)|}{0.6745}. \tag{6}$$

where $\hat{\sigma}$ (median absolute deviation) is an estimate of scale often formed from linear combination of the residuals, the constant 0.6745 makes S an approximately unbiased estimate of σ if n is large and the distribution is normal. For ρ function, we use the table (M)

$$w_i = \frac{\psi\left(\frac{y_i - \sum_{j=0}^{k} x'_{ij}\beta}{\hat{\sigma}_{MAD}}\right)}{\left[\frac{y_i - \sum_{j=0}^{k} x'_{ij}\beta}{\hat{\sigma}_{MAD}}\right]} \tag{7}$$

where $\psi = \acute{\rho}$ is derivative of , x_{ij} is i-th observation on the j-th independent variable. Iteratively Reweighted Least Squares (IRLS) are the two methods to solve the M-estimates nonlinear normal equations. Since the weights depend on the unknown parameter β (and σ), we cannot calculate the weighted mean explicitly. But this weighted-means representation of M-estimators leads to a simple iterative algorithm for calculating the M-estimator.

1. We start with the median as an initial estimate of β and then estimate s. Calculate the weights w_i .
3. Calculate a new estimate of β using equation (4).
4. Repeat Step 2 and 3 until the algorithm converges. Ruckstuhl (2014)

Hampel *et al.* (2011) introduced the system of normal equations to solve this minimization problem which is found by taking partial derivatives with respect to β and setting them equal to zero, $x'wx\beta' = x'wy$. Where w is an (nn) diagonal matrix of weight, popular functions for M-estimators.

$$\hat{\beta}_M = (x'wX)^{-1}(x'w\,Y). \qquad (8)$$

Table-1: A detailed description of M-estimations method

	Objective Function	Score Function	Weight Function (w)
Huber (1964)	$\begin{cases} \frac{1}{2}e^2, & if\ \lvert e\rvert < a \\ a\lvert e\rvert - \frac{1}{2}a^2, & if\ \lvert e\rvert \geq a \end{cases}$	$\begin{cases} e\,, & if\ \lvert e\rvert < a \\ a\ sign\ e\,, & if\ \lvert e\rvert \geq a \end{cases}$	$\begin{cases} 1, & \lvert e\rvert < a \\ \frac{a}{\lvert e\rvert}, & \lvert e\rvert \geq a \end{cases}$
Hampel (2002)	$\begin{cases} \frac{1}{2}e^2\,,\ \lvert e\rvert < a \\ a\lvert e\rvert - \frac{1}{2}a^2,\ a \leq \lvert e\rvert \leq b \\ a\frac{c\lvert e\rvert - \frac{1}{2}e^2}{c-b} - \frac{7a^2}{6},\ b \leq \lvert e\rvert \leq c \end{cases}$	$\begin{cases} e\,, & \lvert e\rvert < a \\ a\ sign\ e\,, & a \leq \lvert e\rvert \leq b \\ a\frac{asign\ e - e}{c-b}, & b \leq \lvert e\rvert \leq c \end{cases}$	$\begin{cases} 1 \\ a/\lvert e\rvert \\ a\frac{c/\lvert e\rvert - 1}{c-b} \end{cases}$
Tukey Bisquare (1987)	$\begin{cases} \frac{a^2}{6}(1-(\frac{e}{a})^2)^3 & if\ \lvert e\rvert \leq a \\ \frac{1}{6}a^2 & if\ \lvert e\rvert > a \end{cases}$	$\begin{cases} e-(1-(\frac{e}{a})^2)^3 \\ 0 \end{cases}$	$\begin{cases} (1-(\frac{e}{a})^2)^3 \\ 0 \end{cases}$

3.2 S-Estimation Method

The regression estimators associated with M-scales is the S-estimators which proposed by Yohai (1987), S-estimation method is based on residual scale of M-estimation method. S-estimators are a generalization of LMS and LTS. And they have the same asymptotic properties corresponding to M-estimators and also handle 50% of the outliers appearing in the data. The weakness of M-estimation method is the lack of consideration on the data distribution and not a function of the overall data because only using the median as the weighted value. S-estimator refers to the fact that this estimator essentially is based on the minimization of a (robust) Scale M-Estimator. Susanti and Pratiwi (2014) discussed this method uses the residual standard deviation to overcome the weaknesses of median, the S-estimator is defined by

$$\hat{\beta}_S = min_\beta\ \hat{\sigma}_s(e_1, e_2, \dots, e_n). \qquad (9)$$

with determining minimum robust scale estimator $\hat{\sigma}_s$ and satisfying

$$min \sum_{i=1}^{n} \rho\left(\frac{y_i - \sum_{i=1}^{n} x_{ij}\beta}{\hat{\sigma}_s}\right)$$

where

$$\hat{\sigma}_s = \begin{cases} \frac{median\lvert e_i - median(e_i)\rvert}{0.6745} & ;\text{iteration} = 1 \\ \sqrt{\frac{1}{nk}\sum_{i=1}^{n} w_i e_i^2} & ;\text{iteration} > 1 \end{cases}. \qquad (10)$$

$$\sum_{i=1}^{n} x_{ij}\psi\left(\frac{y_{i-}\sum_{i=0}^{k} x_{ij}\beta}{\hat{\sigma}_s}\right) = 0 \quad ,j = 0,1,\dots,k$$

ψ is a function as derivative of ρ:

$$\psi(u_i) = \rho'(u_i) = \begin{cases} u_i\left(1-(\frac{u_i}{c})^2\right)^2, & \lvert u_i\rvert \leq c \\ 0 & ,\lvert u_i\rvert > c \end{cases}$$

S-estimators are more robustly than the M-estimator, because S-estimators have smaller asymptotic bias and smaller asymptotic variance in the case contaminated data. Rousseeuw and Leroy (2005), and Pitselis, (2013).

3.3 MM Estimate

MM estimation method is a special type of M-estimation method developed by Yohai (1987). MM estimation method is a combination of high breakdown value estimation method and efficient estimation method Yohai's MM estimator, which was the first estimation with a high breakdown point and high efficiency under normal error. The so-called regression MM estimator (Modified M estimator)

$$\sum_{i=1}^{n} \rho'_1\left(\frac{y_i - \sum_{j=0}^{k} x_{ij}\beta_j}{\boldsymbol{s}_{MM}}\right)x_{ij} = 0$$

where $\boldsymbol{s}_{MM}$ is the standard deviation obtained from the residual of S-estimation method. MM estimation method aims to obtain estimators that have a high breakdown value and more efficient, we will simply extend this approach to mixed linear models. An MM estimator of β is then defined as any solution of an M-type equation where

$$\psi_{MM}(y, x, \beta) = u_{MM}(d)x'\hat{\Sigma}_S^{-1}(y - x\beta)\,. \qquad (11)$$

This looks similar to the previous proposal. The difference with the Huber estimator lies in the definition of the weight function $u_{MM}(d)$ now based on a redescending score.

$$\sum \mathrm{u_{MM}}(\mathrm{d}) x'_{\mathrm{i}} \hat{\Sigma}_S^{-1}(y_i - x_i\beta) = 0.$$

It is likely that their approach can be extended to *MM* estimators in mixed effects models. The formal derivation of the breakdown point of *MM*-estimators in this setting is however beyond the scope of this paper. Some properties of MM-estimator are follows as they are highly efficient when the errors have normal distribution. Their BP is 0.5. Susanti and Pratiwi (2014) introduced the algorithm of computing MM-estimator can be illustrated in detail as follow:

1. Estimate regression coefficients on the data using the OLS.
2. Test assumptions of the classical regression model.
3. Detect the presence of outliers in the data.
4. Calculate residual value $e_i = y_i - \hat{y}_i$ of S estimate.
5. Calculate value of $\hat{\sigma}_i = \hat{\sigma}_{sn}$.
6. Calculate value $u_i = \frac{e_i}{\sigma_i}$
7. Calculate weighted value

$$\omega_i = \begin{cases} \left[1 - \left(\frac{u_i}{4.685}\right)^2\right]^2, & |u_i| \le 4.685; \\ 0 \qquad , & |u_i| > 4.685. \end{cases}$$

8. Calculate $\hat{\beta}_M\ M$ using WLS method with weighted ω_i .
9. Repeate steps $5 - 8$ to obtain a convergent value of $\hat{\beta}_M\ M$.
10. Test to determine whether independent variables have significant effect on the dependent variable.

For more information, see Yohai, et al (1987) and Ruckstuhl (2014)

4. THE SIMULATION STUDY

We make Mote Carlo simulation compare. Least Squares Estimators (OLS), M-Huber (MH), M-Bisquare, MM-Estimator robust regression, S-Estimator, and MM based initials of coefficient S-Estimator, (MM(S). We use R language to create our program to set up Monte Carlo simulation and this program is available if requested.

4.1 Design of the Simulation

Monte Carlo experiments were carried out based on the following data-generating process: Obtain the error term (ε) using normal distribution$(n, 0, \sigma)$. σ is stander deviation of Normal distribution, $\sigma = 1, 5$. X is distributed Uniform distribution on interval (0,1), (1,3), (2,4), (3,6), and (0,6) where (k-1) is number of the Variables of X Selecting $K = (3,6)$, samples of size $n = 50, 100, and\ 150$ and consider that these samples may contain outliers, To investigate the robustness of the methods against outliers, we randomly generate different percentages of outliers (P= 5%, 10%, 15%, 20%, 25% and 30%). Setting the coefficients β equal 1, all simulation results are based on 1500 replications. All computations are obtained based on the R language. The simulation methods are compared using the criteria of estimation method parameters, bias and mean square errors (MSE). When comparing to the MSE of the OLS for such robust methods.

$$MSE = Mean(\hat{\beta} - \beta)^2 \qquad (12)$$

where $\hat{\beta}$ is the estimated value of β.

4.2 The Simulation Results

The simulation results are presented in tables (2), (3), (4), (5), (6) and (7), displaying the properties of different robust estimation methods for different percentages of outliers (P), different number of parameters (k), different of standard normal distribution of error term (σ) and different sample sizes (n). Note that the higher the value of σ and the value of sample size and the ratio of outliers, the lower the value of MSE for the following methods of Robust MM(S), MM and S compared with OLS. As k increases, the bias increases and MSE increases. As σ increases, the bias increases and MSE increases. As n increases, the bias decreases and MSE decreases. As p increases, the bias decreases and MSE increases.

Table 2 Indicates that, in general, the value of bias and MSE are the smallest for the following methods of robust (M.Huber, M.Hampel, M.Bisqare, S, MM(S) and MM estimations). When $\sigma \ge 1$, $k = 3$, $n = 50$ and $0.05 < p < 0.20$, the best method is M.Hampel estimation and the next is MM estimation. If $p > 0.20$ then the MM estimation method is the smallest in bias and MSE compared with other methods. Followed by MM(S) estimation method, and S-estimation method. When $p = 30\%$ M-estimators are not robust.

Table 4 Indicates that, in general, the value of bias and MSE are the smallest for the following methods of robust M-Estimations (M.Huber, M.Hampel, M.Bisqare), S, MM(S) and MM estimations. When $\sigma \ge 1$, $k = 6$, $n = 50$ and $0.05 < p < 0.20$, the best method is M.Hampel estimation and the next is MM(S) estimation. If $p > 0.20$ then the MM estimation method is the smallest in bias and MSE compared with other methods. Followed by MM(S) estimation method, and S-estimation method. When $p = 30\%$ M-estimators are not robust.

Table-2: Bias and MSE values for different estimation method when n=50 and k=3

$\sigma = 1$								
P		OLS	M.Huber	M.Hampel	M.Bisquare	S	MM(S)	MM
0.05	Bias	0.2758	0.0081	-0.0021	-0.0023	-0.0007	-0.0025	-0.0014
	MSE	15.794	0.2378	0.2087	0.2215	0.676	0.2191	0.2366
0.10	Bias	1.8319	0.082	-0.0023	-0.0019	-0.0021	-0.0024	-0.0027
	MSE	24.1761	0.4588	0.219	0.2274	0.6173	0.2255	0.2297
0.15	Bias	3.072	0.1498	-0.0027	-0.0025	-0.0035	-0.0026	-0.0019
	MSE	45.8409	0.3366	0.2241	0.2316	0.5805	0.2285	0.2326
0.20	Bias	4.0357	0.2496	-0.0033	-0.0033	-0.0016	-0.0034	-0.0035
	MSE	85.2381	0.6171	0.2466	0.2519	0.5386	0.2498	0.2488
0.25	Bias	2.4862	0.181	2.3447	-0.003	-0.0067	-0.003	-0.0028
	MSE	12.4908	0.2996	10.962	0.2402	0.4779	0.2381	0.2361
0.30	Bias	2.9892	1.2072	2.9892	0.6614	-0.0012	-0.0036	-0.0035
	MSE	49.3169	11.4253	49.3169	9.661	0.4602	0.2677	0.2673
$\sigma = 5$								
P		OLS	M.Huber	M.Hampel	M.Bisquare	S	MM(S)	MM
0.05	Bias	1.3788	0.0406	-0.0103	-0.0117	-0.0033	-0.0115	-0.0068
	MSE	394.8489	5.9454	5.2187	5.5366	16.8988	5.4699	5.9155
0.10	Bias	9.1594	0.4101	-0.0113	-0.0095	-0.0103	-0.0118	-0.0136
	MSE	2604.402	11.4692	5.4759	5.6847	15.4333	5.6438	5.7416
0.15	Bias	15.3601	0.7492	-0.0134	-0.0124	-0.0173	-0.0132	-0.0095
	MSE	1146.024	8.4158	5.6024	5.7902	14.5122	5.7135	5.816
0.20	Bias	20.1784	1.2478	-0.0165	-0.0167	-0.0081	-0.0169	-0.0176
	MSE	2130.953	15.428	6.1642	6.2974	13.4644	6.2449	6.2188
0.25	Bias	12.4312	0.9049	11.7234	-0.015	-0.0335	-0.0151	-0.0139
	MSE	312.2702	7.4893	274.0496	6.0057	11.9486	5.952	5.9019
0.30	Bias	14.9459	6.036	14.9459	3.3069	-0.0062	-0.0179	-0.0173
	MSE	1232.923	285.6326	1232.923	241.5242	11.5045	6.692	6.6825

Table-3: Bias and MSE values for different estimation method when n=100 and k=3

$\sigma = 1$								
P		OLS	M.Huber	M.Hampel	M.Bisquare	S	MM(S)	MM
0.05	Bias	0.6596	0.0274	0.003	0.0035	0.0029	0.0036	0.0039
	MSE	27.6315	0.1461	0.1057	0.1097	0.3504	0.1105	0.1158
0.10	Bias	1.1647	0.0486	0.0044	0.0051	0.0079	0.0048	0.0047
	MSE	17.3342	0.1391	0.1083	0.1116	0.3172	0.1111	0.1154
0.15	Bias	3.2944	0.1612	0.004	0.0045	0.0047	0.0043	0.0041
	MSE	32.717	0.1928	0.1127	0.1158	0.2978	0.115	0.1179
0.20	Bias	4.3209	0.2623	0.0034	0.004	0.0039	0.004	0.0034
	MSE	83.9513	0.4288	0.1188	0.1212	0.2861	0.1205	0.1215
0.25	Bias	5.0449	0.451	5.0488	0.004	0.0094	0.0039	0.0037
	MSE	138.5631	1.2479	139.0472	0.1244	0.2661	0.1236	0.1237
0.30	Bias	6.5279	6.6844	6.5279	5.9648	0.0044	0.0036	0.0035
	MSE	179.4393	194.6137	179.4393	155.6399	0.2568	0.1299	0.1294
$\sigma = 5$								
P		OLS	M.Huber	M.Hampel	M.Bisquare	S	MM(S)	MM
0.05	Bias	3.2982	0.1371	0.0148	0.0175	0.0143	0.018	0.0194
	MSE	690.7878	3.6531	2.6437	2.7428	8.7601	2.7623	2.896
0.10	Bias	5.8234	0.2429	0.022	0.0254	0.0395	0.0239	0.0236
	MSE	433.3552	3.4786	2.7081	2.7898	7.9299	2.7781	2.8843
0.15	Bias	16.4718	0.806	0.0198	0.0224	0.0236	0.0216	0.0203
	MSE	817.9255	4.821	2.818	2.8943	7.444	2.8741	2.9487
0.20	Bias	21.6043	1.3113	0.0172	0.0202	0.0193	0.02	0.0172
	MSE	2098.7834	10.7192	2.9704	3.0295	7.1524	3.0113	3.0385
0.25	Bias	25.2244	2.255	25.2441	0.0198	0.0469	0.0193	0.0183
	MSE	3464.0777	31.1964	3476.1806	3.1097	6.6536	3.0896	3.0936
0.30	Bias	32.6397	33.4218	32.6397	29.8238	0.0222	0.0182	0.0176
	MSE	4485.9816	4865.3425	4485.9816	3890.9974	6.4194	3.2481	3.2349

Table-4: Bias and MSE values for different estimation method when n=50 and k=6

$\sigma = 1$								
P		OLS	M.Huber	M.Hampel	M.Bisquare	S	MM(S)	MM
0.05	Bias	1.6641	0.0717	0.0032	0.0035	0.0089	0.0036	0.003
	MSE	56.6139	0.4741	0.3557	0.3818	1.3212	0.3862	0.4292
0.10	Bias	3.5349	0.171	-0.0031	-0.0029	0.0072	-0.0038	-0.0041
	MSE	284.8458	1.1116	0.4133	0.4356	1.3606	0.4372	0.449
0.15	Bias	2.309	0.1217	0.0023	0.0026	0.0067	0.0034	0.002
	MSE	51.8402	0.5466	0.3937	0.4087	1.1176	0.4066	0.4231
0.20	Bias	2.3863	0.1476	-0.0034	-0.0035	0.0088	-0.0035	-0.0037
	MSE	47.4764	0.6354	0.4587	0.4712	1.2981	0.4637	0.4659
0.25	Bias	1.5679	0.1168	1.5646	0.0018	0.0095	0.0017	0.0018
	MSE	8.3258	0.4727	8.1673	0.4275	1.1215	0.4229	0.4198
0.30	Bias	1.5584	1.4923	1.5584	1.3647	0.0019	-0.0002	-0.0003
	MSE	23.3054	32.7936	23.3054	26.8426	1.4133	0.5487	0.547
$\sigma = 5$								
P		OLS	M.Huber	M.Hampel	M.Bisquare	S	MM(S)	MM
0.05	Bias	8.3206	0.3586	0.0161	0.0176	0.0447	0.0212	0.0152
	MSE	1415.3479	11.8517	8.8933	9.5442	33.0292	9.609	10.73
0.10	Bias	17.6743	0.8549	-0.0154	-0.0146	0.0359	-0.0198	-0.0204
	MSE	7121.1457	27.7911	10.3325	10.8906	34.0148	11.0429	11.2249
0.15	Bias	11.5449	0.6086	0.0116	0.0132	0.0336	0.017	0.0098
	MSE	1296.0047	13.6639	9.8435	10.2166	27.9407	10.1672	10.5775
0.20	Bias	11.9315	0.7381	-0.0171	-0.0174	0.0441	-0.0175	-0.0184
	MSE	1186.9093	15.8843	11.4674	11.781	32.4526	11.593	11.6477
0.25	Bias	7.8395	0.5838	7.8229	0.0089	0.0477	0.0085	0.0089
	MSE	208.1448	11.8172	204.1834	10.6866	28.0365	10.5749	10.4948
0.30	Bias	7.7922	7.4613	7.7922	6.8234	0.0096	-0.0009	-0.0014
	MSE	582.636	819.841	582.636	671.0652	35.3315	13.7183	13.6751

Table-5: Bias and MSE values for different estimation method when n=100 and k=6

$\sigma = 1$								
P		OLS	M.Huber	M.Hampel	M.Bisquare	S	MM(S)	MM
0.05	Bias	1.7659	0.0686	0.0035	0.0029	-0.007	0.0027	0.0028
	MSE	62.7539	0.2587	0.1615	0.168	0.6212	0.1742	0.1754
0.10	Bias	1.0529	0.0446	0.0033	0.0029	0.0024	0.0035	0.0031
	MSE	17.3685	0.2115	0.1735	0.1787	0.6364	0.1803	0.1827
0.15	Bias	2.4574	0.129	0.0031	0.003	-0.0017	0.0024	0.0028
	MSE	48.8491	0.3426	0.192	0.1963	0.7028	0.1984	0.1979
0.20	Bias	3.7896	0.2597	0.0052	0.0051	0.0071	0.0049	0.0053
	MSE	119.5279	0.8259	0.2066	0.2095	0.73	0.2087	0.2091
0.25	Bias	5.2723	0.6284	5.4063	0.0052	0.0074	0.0052	0.0053
	MSE	290.1303	4.741	307.9456	0.225	0.7977	0.2237	0.2224
0.30	Bias	4.7001	4.9983	4.7001	4.6767	0.006	0.0045	0.0045
	MSE	221.0939	272.7856	221.0939	235.2693	0.8118	0.232	0.231
$\sigma = 5$								
P		OLS	M.Huber	M.Hampel	M.Bisquare	S	MM(S)	MM
0.05	Bias	8.8296	0.3429	0.0176	0.0146	-0.0348	0.0128	0.0141
	MSE	1568.8467	6.4685	4.0366	4.1988	15.5305	4.3549	4.3859
0.10	Bias	5.2646	0.223	0.0166	0.0145	0.012	0.0177	0.0154
	MSE	434.2128	5.2873	4.337	4.4669	15.9102	4.5423	4.5678
0.15	Bias	12.2869	0.6448	0.0154	0.0148	-0.0084	0.0126	0.0139
	MSE	1221.227	8.5643	4.7996	4.9078	17.5691	4.9625	4.9474
0.20	Bias	18.9481	1.2986	0.0261	0.0254	0.0356	0.0245	0.0265
	MSE	2988.1963	20.6466	5.1655	5.2383	18.2497	5.2173	5.2269
0.25	Bias	26.3614	3.1421	27.0314	0.0258	0.0369	0.0259	0.0264
	MSE	7253.2582	118.5256	7698.6407	5.6256	19.9418	5.5925	5.561
0.30	Bias	23.5007	24.9913	23.5007	23.3836	0.0299	0.0223	0.0227
	MSE	5527.3468	6819.639	5527.3468	5881.7322	20.2949	5.8002	5.7749

Table-6: Bias and MSE values for different estimation method when n=150 and k=3

$\sigma = 1$								
P		OLS	M.Huber	M.Hampel	M.Bisquare	S	MM(S)	MM
0.05	Bias	0.3424	0.0107	-0.0006	-0.0002	0.0031	-0.0003	-0.0002
	MSE	4.2576	0.0913	0.0832	0.0864	0.2974	0.0864	0.0904
0.10	Bias	1.9375	0.0747	0.0007	0.0011	0.0055	0.0011	0.0003
	MSE	12.6273	0.1018	0.0808	0.0825	0.2601	0.0823	0.085
0.15	Bias	2.7999	0.1194	-0.0014	-0.0011	0.0012	-0.0011	-0.0017
	MSE	19.7873	0.1315	0.0959	0.0986	0.2882	0.0981	0.0992
0.20	Bias	4.126	0.2256	0.0026	0.003	0.0048	0.003	0.0026
	MSE	43.7382	0.2218	0.0971	0.0984	0.2447	0.0979	0.0988
0.25	Bias	4.043	0.2853	3.972	-0.001	0.0005	-0.0009	-0.0011
	MSE	27.6015	0.2121	24.5644	0.1075	0.2662	0.1069	0.1065
0.30	Bias	3.592	1.6741	3.592	1.8948	0.005	0.0037	0.0036
	MSE	61.1725	19.0827	61.1725	29.895	0.2548	0.1103	0.11
$\sigma = 5$								
P		OLS	M.Huber	M.Hampel	M.Bisquare	S	MM(S)	MM
0.05	Bias	1.7119	0.0535	-0.0028	-0.0009	0.0155	-0.0014	-0.001
	MSE	106.4398	2.2824	2.0796	2.1593	7.436	2.1605	2.2611
0.10	Bias	9.6875	0.3734	0.0034	0.0054	0.0273	0.0054	0.0014
	MSE	315.6821	2.545	2.0206	2.0625	6.5022	2.0576	2.1261
0.15	Bias	13.9997	0.597	-0.0071	-0.0054	0.0061	-0.0053	-0.0086
	MSE	494.6815	3.2869	2.3976	2.4657	7.205	2.4533	2.4799
0.20	Bias	20.6301	1.1281	0.0132	0.0152	0.0241	0.0149	0.0129
	MSE	1093.4541	5.5452	2.4285	2.4589	6.1166	2.4473	2.4711
0.25	Bias	20.2149	1.4265	19.86	-0.0048	0.0024	-0.0047	-0.0053
	MSE	690.0381	5.3025	614.1107	2.6871	6.6551	2.6726	2.6614
0.30	Bias	17.9601	8.3705	17.9601	9.4742	0.0248	0.0184	0.0178
	MSE	1529.3124	477.0668	1529.3124	747.376	6.3697	2.7576	2.7501

Table-7: Bias and MSE values for different estimation method when n=150 and k=6

$\sigma = 1$								
P		OLS	M.Huber	M.Hampel	M.Bisquare	S	MM(S)	MM
0.05	Bias	-0.812	-0.0313	-0.0026	-0.0025	0.0007	-0.0029	-0.0024
	MSE	9.2295	0.1214	0.105	0.1093	0.5296	0.1148	0.1144
0.10	Bias	0.242	0.0052	-0.0029	-0.0029	0.0011	-0.0032	-0.0033
	MSE	3.9336	0.1203	0.1082	0.1116	0.5392	0.1133	0.1149
0.15	Bias	1.3113	0.0562	-0.0029	-0.0027	0.0081	-0.0027	-0.0028
	MSE	15.7825	0.1645	0.1212	0.1247	0.6025	0.1241	0.1266
0.20	Bias	2.9832	0.1652	-0.0035	-0.0036	0.0005	-0.0037	-0.0039
	MSE	66.0396	0.3547	0.1327	0.1353	0.5942	0.1349	0.1337
0.25	Bias	4.0064	0.3415	4.0092	-0.0045	0.0046	-0.0045	-0.0046
	MSE	119.628	1.0735	119.8884	0.1371	0.6563	0.1363	0.1357
0.30	Bias	3.5952	3.57	3.5952	3.2383	-0.0062	-0.0049	-0.0049
	MSE	145.8346	169.3027	145.8346	140.2173	0.7458	0.1483	0.1474
$\sigma = 5$								
P		OLS	M.Huber	M.Hampel	M.Bisquare	S	MM(S)	MM
0.05	Bias	-4.06	-0.1565	-0.013	-0.0127	0.0037	-0.0166	-0.0121
	MSE	230.7363	3.0359	2.6248	2.7336	13.2412	2.8877	2.8611
0.10	Bias	1.2099	0.0261	-0.0144	-0.0146	0.0055	-0.0159	-0.0165
	MSE	98.3398	3.007	2.7039	2.7894	13.4799	2.8526	2.8722
0.15	Bias	6.5566	0.2811	-0.0147	-0.0137	0.0403	-0.0135	-0.014
	MSE	394.5614	4.1118	3.0292	3.1167	15.0634	3.1138	3.1656
0.20	Bias	14.9161	0.8258	-0.0176	-0.0181	0.0024	-0.0185	-0.0193
	MSE	1650.991	8.8675	3.3181	3.3833	14.8542	3.3727	3.3431
0.25	Bias	20.0319	1.7077	20.046	-0.0224	0.0231	-0.0227	-0.023
	MSE	2990.6992	26.8378	2997.2108	3.4275	16.4083	3.4086	3.3923
0.30	Bias	17.9762	17.8502	17.9762	16.1915	-0.0312	-0.0243	-0.0243
	MSE	3645.865	4232.5687	3645.865	3505.4335	18.6451	3.7081	3.6843

4.3 Summary and Conclusions From Simulation Results:

For all sample sizes, MSE decreases with increasing sample size for all estimation methods for sample sizes 50, 100 and 150. In the case of errors of normal distribution. We note that S, M-Estimations, MM(S) and MM estimation methods are the most efficient compared with other methods, but MM is more efficient than S, MM(S) and M-estimations, and it is better in generating outlier (distribution error). When the number of parameters (K) is equal 3 and the σ value is equal 5, the higher the ratio of the outliers (P) the better method is S, MM(S) or MM. When the number of parameters (K) is equal 6 and the σ value is equal 1, the higher the ratio of the outliers (P) the better method is the MM(S) method followed by the MM method if the sample size is less than 100 observations. But in case of sample size greater than 100 observations the best method is MM and the next method is MM(S). When the number of parameters (K) is equal 6 and the σ value is equal 5, the higher the value of P, the better method is MM method followed by the MM(S) method. when the number of parameters (K) changes and the value of σ changes, the greater the value of (P) than 20%, the better method is MM and the next is MM(S) method if the sample size changes.

5. THE APPLICATION OF REAL DATA

We present the numerical results of robust regression estimators of real data. The Grunfeld's Investment Data, Description the total number of observations 200 of production units in United States, since data is gross investment value is value of the firm and capital is stock of plant and equipment by Baltagi (2013).

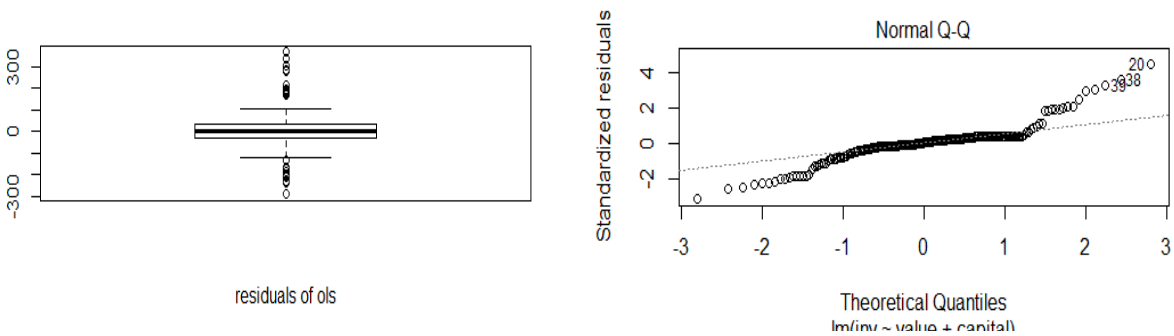

Figure-1: Residuals of OLS estimator

From the previous drawing, we note the extent to which the variable is abnormal and that because there are outliers and conform to that test of Shapiro test of normality since the result is:

$$W = 0.8798, \quad p - value = 1.552e^{-11}.$$

Table-2: Robust estimation method of real data in case 1

methods	coefficients			std
	(Intercept)	value	capital	
OLS	-42.714	0.116	0.231	80.242
M.Huber	-24.816	0.119	0.148	74.943
M.Hampel	-20.989	0.119	0.125	74.434
M.Bisquare	-19.699	0.124	0.120	77.375
S	3.820	0.053	0.092	24.602
MM(S)	6.12689	0.052	0.0815	24.433
MM	10.636	0.051	0.072	24.132

We note from the previous results that if OLS estimation method is used then not robust is produced and the standard deviation is increased, compared with the robust methods where the standard deviation is reduced we can use methods M, MM and MM based initials of coefficient S (MM(S). The result of application then the best method estimation method is MM estimation method based on initial (S).

6. CONCLUSION

We have discussed procedures to estimate robust regression model using OLS estimation method, M.Huber-estimation method, M.Hampel-estimation method, M.Bisquare-estimation method, S-estimation method, MM(S)-estimation method and MM estimation method. The use of the method of robust estimation method in the presence of outliers tends to improve the efficiency and reduce the bias compared with the classical methods of estimation. MM(S) is the most efficiency compared with other methods, but MM is the more efficiency compared with S method and it is better in outlier generating error distribution. The M.Humpel-estimator is not robust with respect to high leverage points, so it should be used in situations where high leverage points do not occur, but we can use methods MM and MM(S). The result of simulation Monte Carlo then the best method estimation method is MM estimation method.

ACKNOWLEDGMENT

The authors would like to thank Dr. Mohamed R. Abonazel for Assistance, encouragement and guidance in generating some of the simulation results presented in this manuscript.

REFERENCES

1. Abdi, H. (2007). The method of Least Squares. Encyclopedia of Measurement and Statistics. CA, USA: Thousand Oaks.
2. Almongy, H., & Almetwaly, E. (2018). Comparison between methods of robust estimation for reducing the effect of outliers. The Egyptian Journal for Commercial Studies, Faculty of Commerce, Mansoura University, Egypt, 4(41), 1-23.
3. Alma, Ö. G. (2011). Comparison of Robust Regression Methods in Linear Regression. International Journal of Contemporary Mathematical Sciences, 6(9), 409-421.
4. Baltagi, Badi H. (2013) Econometric Analysis of Panel Data, 5th ed., John Wiley and Sons.
5. Fox, J. (2002). Robust Regression. An R and S-Plus Companion to Applied Regression.
6. Hampel, F. (2002). Robust Inference. Encyclopedia of Environmetrics.
7. Hampel, F. R., Ronchetti, E. M., Rousseeuw, P. J., & Stahel, W. A. (2011). Robust Statistics: the Approach Based on Influence Functions (Vol. 114). John Wiley & Sons.
8. Huber, P. J. (1964). Robust Estimation of a Location Parameter. The Annals of Mathematical Statistics, 35 (1), 73-101.
9. Pitselis, G. (2013). A Review on Robust Estimators Applied to Regression Credibility. Journal of Computational and Applied Mathematics, 239, 231-249.
10. Rousseeuw, P. J., & Leroy, A. M. (2005). Robust Regression and Outlier Detection (Vol. 589). John wiley & sons.
11. Ruckstuhl, A. (2014). Robust Fitting of Parametric Models Based on M-Estimation. Lecture notes.
12. Susanti, Y., & Pratiwi, H. (2014). M-estimation, S-estimation, and MM- estimation in robust regression. International Journal of Pure and Applied Mathematics, 91(3), 349-360.
13. Yohai, V. J. (1987). High Breakdown-point and High Efficiency Robust Estimates for Regression. The Annals of Statistics, 642-656.
14. Zioutas, G., Avramidis, A., & Pitsoulis, (2005) L. A Penalized Trimmed Squares Method for Deleting Outliers in Robust Regression.