International Journal of Research Publication and Reviews

Journal homepage: www.ijrpr.com ISSN 2582-7421

Performance Improvement of Some Robust Estimators Using Huber Redecent Weights in the Presence of Outliers and High Leverage Points

U. F. Salisui[1], U. F. Abbas[2], K. E. Lasisi[3]*

[1]Federal Polytechnic Bauchi, Bauchi State, Nigeria
[2]Federal Polytechnic Bauchi, Bauchi State, Nigeria
[3]Abubakar Tafawa Balewa University Bauchi, State, Nigeria

ABSTRACT

This research work was conducted to improve the performance of S, M and MM robust estimators in the presence of outliers in regression model using a default and adjusted turning constant values of the Huber redescent weights, so as to see the impact of increasing or reducing the turning constant values on the resistibility of the robust estimators. Since it has been observed from the literature, the previous researchers stick to the used of the default turning constant values despite the fact that the turning constant can be adjusted. A simulation study was used to compare the performance of the robust methods with five and ten independent variables on the sample sizes 20, 30, 150, 300 and 20,50, 200, 500 respectively. The result of the study showed that increasing or decreasing the turning constant values for S and MM estimators have a significant impact on the resistibility of the estimators on the sample sizes used, while adjusting the turning constant value for M-estimator has no significant impact on the robust regression result. The study also recommends that analyst should use smaller and higher value than the default turning constant for MM-estimation and S-estimator respectively to have more resistibility against outlying observations in the regression models, because the smaller the turning constant value the more resistant the MM-estimator. Secondly the study recommended that analyst should use a value greater than the default value for the S-estimation, because the higher the turning constant the more resistant the S-estimator against outlying observations.

Keywords: Outlier, Turning Constant, S-Estimation, M-estimation, MM-estimation

1. Introduction

In regression modeling,the relationship between the dependent; y and a set of independent variables;$x_1, x_2, \dots, x_n$can be approximated by the model:

$$y = \mathrm{f}(x_1, x_2, \dots, x_n)\beta + e_{ij} \quad \dots \quad 1.1$$

Where $e_{ij}$is assumed to be a random error representing the discrepancy in the approximation, it accounts for the failure of the model to fit the data exactly. To judge how well the estimated regression model fits the data, we can look at the size of the residuals, which is:

$$e_{ij} = \mathrm{y_i} - (\mathrm{f}(x_1, x_2, \dots, x_n)\beta) \quad \dots \quad 1.2$$

A point which lies far from the line and often has a large residual value is called an outlier. However, it is well known that the OLS estimate is extremely sensitive to the outliers. When such violation occur, robust regression that is resistant to the influence of outliers may be the reasonable recourse to alternate these shortcomings (Yuliana et al., 2014), papageogiou,Bouboulis, and Theodoridis (2015); Barnish and Lagoa(1997); Hansen and sargent(2008); Andrews(1974); Street, Carroll, and Ruppert(1988); hogg(1974);(1979); Huber(1996); Krasker and Welsch(1982). The robust methods have been defined to deal with the influential points in regression analysis, where the value of the estimation by using this method is not much affected with outliers. Recent years have seen a dynamic development in statistical methods for analyzing data contaminated with outliers. One of the more important techniques that can deal with outlying observations is robust regression, which represents four decades of research. Until recently the implementation of robust regression methods, such as M-estimation or MM-estimation, was limited owing to their iterative nature. With advances in computing power and the growing availability of statistical packages, such as R, SAS and Stata, the applicability of robust regression methods has increased considerably (Abonazel and Rabie, 2019).

According to Abonazel and Rabie, (2019); (Bulletin and 1984 nd); gupta et al, 2014; Montgomery,Peck, and Vining (2001) when the dataset is contaminated with a single or few outliers, the problem of identifying such observations is serious problem.

A common method of robust regression is the M estimate, introduced by Huber (1964), which is as efficient as Ordinary Least Square (OLS), and is considered the simplest approach. The S estimation is used to minimize the dispersion of residuals. The MM estimation is a special type of M-estimation introduced by Yohai (1987) which combines high breakdown value estimation and efficient estimation. The M estimation has a higher breakdown value and greater statistical efficiency than the S estimation. The well-known methods of robust estimation are M-estimation, S- estimation and MM-estimation.

1.2 Statement of the Problem

Controversy surrounding the default value of redescent weight and turning constant (C) that reaches its highest breakdown point and performance in robust regression; in other word, which robust regression will produce the highest breakdown point on a data set that is contaminated with outliers and high leverage point(s) from among the techniques selected (S, M and MM). A number of researchers posited that live with what you have (use the default value of 4.865) while others opined that such value is subject to change for higher break point and performance. A lot of researchers stick to the used of default turning constant value C of the redescent weights for the respective robust regression methods despite the fact that the C value can be adjusted. Calandra (2021) they studied the default value and recommended for adjustment for better performance. This work is the extension of their work where value of the redescent weight is modified and adjusted.

In his work, Marison (2021) he evaluated the performance of some robust estimators in the presence of outliers using a default turning constant values, where he evaluated the constant value for M-estimation, S-estimation, and MM-estimation method which led to improved efficiency and reduce bias in estimating the contaminated data set.The question is what will happen when the default value is modified and adjusted, as recommended by Justo and Calandra (2021)? Which method will produce the best estimate when there is change in value in both direction (increase or decrease). Since outliers and high leverage point often causes a huge explanatory mishap in regression modeling, studies on how to remedy it is required so as to know the possible approaches in reaching its highest efficiency and break down point. This is where I based my research. We want to study a situation when all the estimators are operating under the same redescent weight and modifying the tuning constants and measure the performance of such operators.

2. Methodology

Several methods have been developed to overcome the deficiencies of modeling data set that violate the assumptions of multiple regression. The current study explores S, M and MM estimators. Hence these methods will be discussed along with their described algorithms and a numerical example using the same data set

2.1 M - Estimator

One of the robust regression estimation methods is the M estimation. The letter M indicates that M estimation is an estimation of the maximum likelihood type. If estimator at M estimation is $\hat{\beta}_M = \beta_n(x_1, x_2, \ldots, x_n)$ then:

$E\left[\hat{\beta} = \beta_n(x_1, x_2, \ldots, x_n)\right] = \beta$

This implies that equation (1) is an unbiased estimator with its corresponding variance given as:

$$Var(\hat{\beta}) = \frac{\left[\tilde{\beta}\right]^2}{n.E\left[\frac{d}{d\beta} Inf(x_i;\beta)\right]^2}$$

Where $\tilde{\beta}$is the linear unbiased estimator for β.M estimation is an extension of the maximum likelihood estimate methodand a robust estimation (Yuliana & Susanti, 2008); Martin and Zamar(1989); Mohamed,Abdullah and muthu(1989); Dalalyan and Thompson (2019); Deng et al, 2014; Susanti et al (2014)

2.2 S - Estimator and Algorithms

The weakness of M estimation 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. This method uses the residual standard deviation to overcome the weaknesses of M- estimator. According to Salibian and Yohai (2006), the S-estimator is defined by $\beta_S = Min_\beta \hat{\sigma}_s(e_1, e_2, \ldots, e_n)$ with determining minimum robust scale estimator.

$$\hat{\sigma}_i = \begin{cases} \dfrac{Median.\,e_i.\,(Median)(e_i)}{0.6745}, & Iteration = 1 \\ \dfrac{1}{nK}\sum w_i e_i^2, & Iteration > 1 \end{cases}$$

Where;

$$w_i = \begin{cases} \begin{cases} \left[1 - \left(\dfrac{u_i}{1.547}\right)^2\right]^2, & |u_i| \leq 1.547 \\ 0, & |u_i| > 1.547 \end{cases} & Iteration = 1 \\ \dfrac{\rho(u)}{u^2} & Iteration > 1 \end{cases}$$

2.3 MM – Estimator and Algorithms

MM - estimation provide a high breakdown value and more efficient than M and S estimators. Breakdown value is a common measure of the proportion of outliers that can be addressed before these observations affect the model. MM-estimator is the solution of;

$$\sum_{i=1}^{n} \rho_1(u_i)X_{ij} = 0 \; or \sum_{i=1}^{n} \rho_1 \left(\frac{Y_i - \sum_{j=0}^{k} X_{ij}\beta_j}{S_{MM}}\right) X_{ij} = 0$$

Where S_{MM} is the standard deviation of S estimation and ρ is the Turkey's biweight function defined by:

$$\rho(u_i) = \begin{cases} \frac{u_i^2}{2} - \frac{u_i^4}{2c^2} + \frac{u_i^6}{6c^2}, & -c \le u_i \le c \\ \frac{c^2}{6}, & u_i < -c \; or \; u_i > c \end{cases} \qquad 3.4$$

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

2.4 Modified Algorithm

1. Estimate regression coefficients on the data using OLS.

2. Test assumptions of the 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 $\hat{\sigma}_i = \hat{\sigma}_{sn}$

6. Calculate $u_i = \frac{e_i}{\hat{\sigma}_i}$

7. Calculate the weighted value w_i:

$$w_i = \begin{cases} \left[1 - \left(\frac{u_i}{3.685: 6.685}\right)^2\right]^2, & |u_i| \le 3.685: 6.685 \\ 0, & |u_i| > 3.685: 6.685 \end{cases}$$

8. Calculate $\hat{\beta}_{MM}$ using weighted least square method with weighted w_i

9. Repeat step 5 – 8 to obtain a convergent value of $\hat{\beta}$ and test whether the independent variables have a significant effect on the dependent variable.

3. Simulation: Monte Carlo Simulation Study

A Monte Carlo simulation will be used to assess the merit of our proposed method over the existing method in terms of its ability to excellently separate the data set according to regular observations, vertical outliers (regression outliers), collinearity-enhancing observations with large residuals, bad leverage collinearity-enhancing observations, good leverage collinearity-enhancing observations and collinearity enhancing observations. To achieve this aim, we follow the simulation procedure used by (Bagheri and Habshah, 2015) by considering a linear relation$y_i = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \beta_3 x_3 + e_i$where $i = 1,2,\ldots,n$. Three explanatory variables $(p = 3)$ were to be generated from Uniform (0,1) to produce non-collinear data sets in which the true parameters were set at $\beta_0 = \beta_1 = \beta_2 = \beta_3 = 1$ and $\varepsilon_i \sim N(0, \delta_I^2)$in such a way that different scenarios were created, namely, high leverage collinearity-enhancing/reducing observations and vertical outliers. In each scenario, small, medium and large samples of size 20, 40 and 100 with three different levels of high leverages points (HLP) of 10%, 15% and 20% for all the sample sizes considered at 500 replications.

However, to generate high leverage collinearity-enhancing observations, each variable was firstly generated from Uniform (0,1) to produce non-collinear data sets. This generated data is referred to as the regular observations. The last 100%α observations of the regular observations of each regressor were then replaced with certain percentage of high leverage points to create high leverage collinearity-enhancing observations. To generate the high leverage points as collinearity-enhancing observations with unequal weights in non-collinear data sets, the values corresponding to the first high leverage point were kept fixed at 10 and those of the successive values are created by multiplying the observations index, i, by 10.

As per (Abu Sayed et al., 2020), high leverage collinearity-reducing observations are created by generating three collinear regressors on the outset:

$X_{ij} = (1 - \rho^2)Z_{ij} + \rho Z_{i4} \qquad i = 1, \ldots, n; \qquad j = 1, \ldots, 3.$

Where, the X and Z are independent standard normal random numbers. The value of ρ^2 represents the correlation between the two explanatory variables, is set to 0.95 which cause high collinearity between regressors. High leverage collinearity-reducing observations in collinear data sets were then created by replacing the first $100(\frac{\alpha}{2})$ percent observations of X_1 and the last $100(\frac{\alpha}{2})$ percent observations of X_2 with high leverage points.

Factors and levels for the simulated data sets

Number of Independent Variables 'p'	Sample Size 'n'
5	20 30150 300
10	20 50 200 500

4. Result and Comparison

Comparison of Residual Standard Error with P = 5; n =20

Table 1: Result of M-estimation, S-estimation and MM-estimation on simulated data with five levels on sample size 20

	RSE	**RSE**	**RSE**
	C=4.685	C=3.685	C=6.685
M-estimation	0.3291	0.3224	0.3296
MM-estimation	0.2994	0.2994	0.2994
S-estimation	k=1.547	k=0.547	k=2.547
	0.2996	0.8474	0.2134

Source: Authors' computation using R package v 4.1.1.

Comparison of Residual Standard Error P = 5; n = 30

Table 2: Result of M-estimation, S-estimation and MM-estimation on simulated data with five levels on sample size 30

	RSE	**RSE**	**RSE**
	C=4.685	C=3.685	C=6.685
M-estimation	0.3784	0.3224	0.3296
MM-estimation	0.3013	0.3015	0.3007
S-estimation	k=1.547	k=0.547	k=2.547
	0.5536	0.4732	0.195

Source: Authors' computation using by R package v 4.1.1.

Comparison of Residual Standard Error P = 5; n = 150

Table 3: Result of M-estimation, S-estimation and MM-estimation on simulated data with five levels on sample size 150

	RSE	**RSE**	**RSE**
	C=4.685	C=2.685	C=5.685
M-estimation	0.3402	0.3368	0.3423
MM-estimation	0.3208	0.3208	0.3208
S-estimation	k=1.547	k=0.547	k=2.547
	0.321	0.9078	0.1820

Source: Authors' computation using by R package v 4.1.1.

Comparison of Residual Standard Error P = 5; n = 300

Table 4: Result of M-estimation, S-estimation and MM-estimation on simulated data with five levels on sample size 300

	RSE	**RSE**	**RSE**
	C=4.685	C=2.685	C=5.685
M-estimation	0.3282	0.3103	0.3258
MM-estimation	0.2993	0.2993	0.2993
S-estimation	k=1.547	k=0.547	k=2.547
	0.2995	0.8471	0.1719

Source: Authors' computation using by R package v 4.1.1.

Comparison of Residual Standard Error P = 10; n =20

Table 5: Result of M-estimation, S-estimation and MM-estimation on simulated data with five levels on sample size 20

	RSE	**RSE**	**RSE**
	C=4.685	C=3.685	C=6.685

M-estimator	0.3669	0.3633	0.3634
MM-estimator	0.3445	0.3433	0.3359
S-estimator	k=1.547	k=0.547	k=2.547
	0.3524	0.9897	0.2219

Source: Authors' computation using by R package v 4.1.1.

Comparison of Residual Standard Error P = 10; n = 50

Table 6: Result of M-estimation, S-estimation and MM-estimation on simulated data with five levels on sample size 50

	RSE	**RSE**	**RSE**
	C=4.685	C=3.685	C=6.685
M- estimator	0.3899	0.3989	0.3786
MM- estimator	0.3543	0.3598	0.3545
S- estimator	k=1.547	k=0.547	k=2.547
	0.3587	0.9945	0.2231

Source: Authors' computation using by R package v 4.1.1.

Comparison of Residual Standard Error P = 10; n = 200

Table 7: Result of M-estimation, S-estimation and MM-estimation on simulated data with five levels on sample size 200

	RSE	**RSE**	**RSE**
	C=4.685	C=2.685	C=5.685
M- estimator	0.3766	0.3678	0.3765
MM- estimator	0.3407	0.3354	0.5433
S- estimator	k=1.547	k=0.547	k=2.547
	0.3459	0.9876	0.3126

Source: Authors' computation using by R package v 4.1.1.

Comparison of Residual Standard Error P = 10; n = 500

Table 8: Result of M-estimation, S-estimation and MM-estimation on simulated data with five levels on sample size 500

	RSE	**RSE**	**RSE**
	C=4.685	C=2.685	C=5.685
M-estimator	0.3987	0.3984	0.3995
MM-estimator	0.3576	0.3756	0.3674
S-estimator	k=1.547	k=0.547	k=2.547
	0.3678	0.9856	0.2556

Source: Authors' computation using by R package v 4.1.1.

Table: 9 Comparing the performance of the robust methods with 5 levels when n=20, 30,150and 300

	P =5 n = 20			P = 5 n = 30			P = 5 n = 150			P = 5 n = 300		
Method	RSE			RSE			RSE			RSE		
M	0.3291	0.3224	0.3296	0.3784	0.3224	03216	0.3402	0.3368	0.3423	0.3282	0.3103	0.3258
Mm	0.2994	0.2994	0.2994	0.3013	0.3015	0.3007	0.3208	0.3208	0.3208	0.2993	0.2993	0.2993
S	0.2996	0.8474	0.2134	0.5536	0.4732	0.195	0.3210	0.9078	0.1820	0.2995	0.8471	0.1719

Table: 10 Comparing the performance of the robust methods with 10 levels when n= 20, 50,200and 500

	P =10 n = 20			P = 10 n = 50			P = 10 n = 200			P = 10 n = 500		
Method	RSE			RSE			RSE			RSE		
M	0.3669	0.3633	0.3634	0.3899	0.3989	03786	0.3766	0.3678	0.3765	0.3987	0.3984	0.3995
Mm	0.3445	0.3433	0.3359	0.3543	0.3598	0.3545	0.3407	0.3354	0.5433	0.3576	0.3756	0.3674
S	0.3524	0.9897	0.2219	0.3587	0.9945	0.2231	0.3359	0.9876	0.3126	0.3678	0.9856	0.2556

5. Discussion

Table4.1 to 4.4 showed the result of M, MM and S estimators methods of robust regression with five (5) independent variables on sample sizes 20, 30 150 and 300 respectively. While Table 5.1 to 5.4 showed the result of M, MM and S estimators' robust regression with 10 independent variables on sample sizes 20, 50, 200 and 500 respectively. Table 6 compared the performance of the three methods with 10 independent variables on sample sizes 20, 50, 200 and 500 respectively.

As we can see from the table 6 above, for n=20 as we used the default value (4.685) and decrease the value to (3.686), the MM estimator is the best, with minimum RSE value. While when we increased the default value S – estimator is the best with minimum value of RSE (0.2134). AlsoTABLE we have seen that by default and decreasing the default value to (3.685) the MM – estimator is also the best with minimum RSE (0.3015), and when we increased the default value to (6.686), the S – estimator is also the best with minimum RSE (0.1950). TABLE and TABLE also show that the MM – estimator is the best when we used default and decreased the constant value respectively, While S - estimator remain the best when we increases default value from (4.686) to (6.686) respectively.

When the number of independent variables is 10 with varied sample sizes of 20, 50, 200 and 500 respectively, table 8 shows a comparison between the methods when the turning constant is at default, increased and decreased respectively. When the sample size, n = 20 decreasing the **C** value improves the performance of M – estimator and MM-estimator with the RSE value of (0.3633) and (0.3433) respectively while increasing the **C** value also increases the RSE, while for MM-estimator increasing and decreasing the turning constant did not make any significant impact and for the S-estimator increasing the **C** value to (6.685) gives a better result than the default value (4.685).

Therefore, at default values of the redescent weight of robust regression there is no significant impact of decreasing the value, rather there is significant impact or improvement when we increase the value of the redescent weight.

6. Conclusion

Based on the findings of this study, we can say that adjusting the turning constant value for M, MM and S estimators have a significant impact on the resistibility of the estimators against outlying observations. From the result of the study, we can conclude that decreasing the default turning constant value for MM-estimation from (4.685) to (3.685) will increase the resistibility of the estimator against outlying observations on all sample sizes with few or more independent variables. The study also concludes that changing the default **c** value for M-estimator has no significant impact. It also reveal that increasing the default constant value of S-estimator from (4.685) to (6.685) increases the resistibility of the estimator by having smaller RSE on all the sample sizes 20,30, 50, 150, 200, 300 and 500 respectively, with p = 5 and p = 10 respectively.

7. Reference

Almongy, Hisham Mohamed, and Ehab Mohamed Almetwaly.n.d. (2021) "Comparison between Methods of Robust Estimation to Reduce the Effect of Outliers": Researchgate.Net: https://www.researchgate.net/profile/Ehab_Almetwally/publication/326252179_Estimation_of_the_ Generalized_Power_Weibull_Distribution_Parameters_Using_Progressive_Censoring_ Schemes/links/5bf47bf2299bf1124fe0e6dd/Estimation-of-the-Generalized-Power-Weibull- Distribution-Parameters-Using-Progressive-Censoring-Schemes.pdf.

AND MM ESTIMATION IN ROBUST REGRESSION. International Journal of Pure and Applied Mathematics Volume 91 No. 3 2014, 349-360: ISSN: 1311-8080 (printed version); ISSN: 1314-3395 (on-line version); url: http://www.ijpam.eu: doi: http://dx.doi.org/10.12732/ijpam.v91i3.7

Anderson, Cynthia, and Randall E. Schumacker. (2003). "A Comparison of Five Robust Regression Methods With Ordinary Least Squares Regression: Relative Efficiency, Bias, and Test of the Null Hypothesis." Understanding Statistics 2 (2): 79–103. https://doi.org/10.1207/s15328031us0202_01.

Andrews, D F. (1974). "A Robust Method for Multiple Linear Regression." Technometrics 16 (4): 523– 31. https://doi.org/10.1080/00401706.1974.10489233.

Barmish, B. R., and C. M. Lagoa. (1997). "The Uniform Distribution: A Rigorous Justification for Its Use in Robustness Analysis." Mathematics of Control, Signals, and Systems 10 (3): 203–22. https://doi.org/10.1007/BF01211503.

Brown, B. M., and J. S. Maritz.(1982). "Distribution-Free Methods in Regression." Australian Journal of Statistics 24 (3): 318–31. https://doi.org/10.1111/j.1467-842X.1982.tb00837.x.

bulletin, JP Stevens - Psychological, and undefined 1984. n.d. "Outliers and Influential Data Points in Regression Analysis." Psycnet.Apa.Org. Accessed April 4, 2021. 43: https://psycnet.apa.org/record/1984-13956-001.

Cheng, CL, JW Van Ness (1992) The Annals of Statistics, and undefined 1992: n.d. "Generalized - Estimators for Errors-in-Variables Regression." Projecteuclid.Org. Accessed April 4, 2021. https://projecteuclid.org/euclid.aos/1176348528.

Chun, &Welxin, () Robust Regression and Outlier Detection with the ROBUSTREG Procedure, Statistics and Data Analysis, paper 265-27, SAS Institute Inc., Cary, NC.

Computational and Graphical Statistics, 15, No. 2 (2006), 414-427, doi: 10.1198/106186006X113629. 360 Y.

Cook, RD, S Weisberg (1982) Sociological methodology, and undefined 1982. n.d. "Criticism and Influence Analysis in Regression."JSTOR.Accessed April 4, 2021. https://www.jstor.org/stable/270724.

Dalalyan, Arnak S., and Philip Thompson.(2019). "Outlier-Robust Estimation of a Sparse Linear Model Using ℓ1-Penalized Huber's M-Estimator."ArXiv.arXiv. Data Analysis. John Wiley, 1983

Deng, Lu, Li Shang, ShouminBai, Ji Chen, Xueyan He, Rachel Martin-Trevino, Shanshan Chen, et al. (2014). "Tumor and Stem Cell Biology MicroRNA100 Inhibits Self-Renewal of Breast Cancer Stem-like Cells and Breast Tumor Development."AACR.https://doi.org/10.1158/0008-5472.CAN-13-3710.

Draper, N. R. & Smith, H. (1998). Applied Regression Analysis, Third Edition, Wiley Interscience Publication, United States, 1998.

Gupta, Manish, Jing Gao, Charu C Aggarwal, and Jiawei Han. (2014). "Outlier Detection for Temporal Data: A Survey." IEEE TRANSACTIONS ON KNOWLEDGE AND DATA ENGINEERING 26 (9). https://doi.org/10.1109/TKDE.2013.184.

Gürünlü Alma, Özlem. (2011). "Comparison of Robust Regression Methods in Linear Regression Poisson-Exponential Distribution: Problems of Estimation and Prediction View Project Özlem GÜRÜNLÜ ALMA MuglaÜniversitesi Comparison of Robust Regression Methods in Linear Regression." Int. J. Contemp. Math. Sciences.Vol. 6. https://www.researchgate.net/publication/284946872.

Hampel, Frank R. (1992). "Introduction to Huber (1964) Robust Estimation of a Location Parameter."In , 479–91. https://doi.org/10.1007/978-1-4612-4380-9_34.

Hansen, LP, and TJ Sargent.(2008). Robustness. https://books.google.com/books?hl=en&lr=&id=7wwV27TvR8EC&oi=fnd&pg=PP2&dq= The+notation+of+robustness+&ots=QLJTQp6H- E&sig=D1Tz9DXnaygqoRcFdJd4g5hJgTU.

Hawkins, DM. (1980a). Identification of Outliers.https://link.springer.com/content/pdf/10.1007/978- 94-015-3994-4.pdf.

Hoaglin, D.C., Mosteller, F. and Tukey, J.W. (1983) Understanding Robust and Exploratory Kafadar, K. (1983). The efficiency of the biweight as a robust estimator of location. Journal of Math-Info, 1, No. 11 (2008), 8-16.

Mosteller, F. &Tukey, J.W. (1977). Data Analysis and Regression: A Second Course in Research of the National Bureau of Standards, 88(2):105{116, 1983.

Salibian, M. &Yohai, V.J. (2006). A Fast Algoritm for S-Regression Estimates, Journal of Statistics. Addison-Wesley Publishing Company, 1977.

Yuliana and Y. Susanti, Estimasi M dansifat-sifatnyapadaRegresi Linear Robust, Jurnal Yuliana, S., Hasih, P.,Sri, S.H., & Twenty, L. (2014) M ESTIMATION, S ESTIMATION,