Research Journal of Recent Sciences ISSN 2277-2502
Vol. **2(8),** 79-91, August **(2013)** *Res.J.Recent Sci.*

A New Efficient Redescending M- Estimator: Alamgir Redescending M- estimator

Alamgir[1], Amjad Ali[2], Sajjad Ahmad Khan[2], Dost Muhammad Khan[2] and Umair Khalil[2]
[1]Department of Statistics, University of Peshawar, PAKISTAN
[2]Department of Statistics, Islamia College University Peshawar, PAKISTAN

Available online at: www.isca.in
Received 19th March 2013, revised 11th May 2013, accepted 4th July 2013

Abstract
The ordinary least squares (OLS) estimators are very sensitive to the presence of outliers in the data. Several robust methods have been suggested by researchers to cope with this problem. In this paper we propose a new redescending M-estimator, called Alamgir redescending M- estimator. Its performance is compared with other robust estimators and also with OLS using simulation studies. Real data examples have been presented to evaluate the performance of the proposed estimator.

Keywords: Regression, OLS, Redescending M- estimator, ALARM estimator, tuning constant.

Introduction

It is evident from the literature that Least Squares estimates are extremely sensitive to outliers. Even, single badly placed outlier can distort the LS estimates that lead to unreliable and misleading results and hence will provide useful information about majority of the data points in the data set. To cope with such a situation, researchers have developed "Robust Regression" procedures as an improvement to the classical least squares estimation in the presence of outliers in the data. The OLS estimates are badly affected by the extreme leverage points having arbitrarily very large residuals[1].

Robust estimation is considered to be a possible alternative to classical estimation procedures to compensate for its sensitivity to outliers, especially, when dealing with multivariate contaminated datasets. The main aim of robust estimators is to reduce the influence of these outliers and provide stable results. A major disadvantage of robust methods is that the robust procedures do not have rigorous statistical basis.

The first step in this direction came in1887 when L_1-criterion or Least Absolute Deviation was introduced[2,3].For the regression model,

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

The objective function for the L_1-norm is based on the minimization of sum of absolute residuals instead of sum of squared residuals, that is

$$min \sum_{i=1}^{n} |r_i|$$

where $r_i$is the *i*th residual.

Another important class of robust estimators is the M-estimator[4]. M-estimator is an extension of least squares procedure[5]defined as follows

$$f(\beta) = \sum_{i}^{n} \rho(y_i - g(x_i, \beta)) \Rightarrow \min \tag{2}$$

The form of $\rho(.)$ defines various forms of robust M-estimators. When $\rho(r_i) = r_i^2$, it becomes simple least-squares.The minimization of the function is done iteratively using iteratively reweighted least squares (IRLS or IWLS).A reasonable $\rho(.)$ should possess the following properties: i. $\rho(r_i) \geq 0$ (non-negativity), ii. $\rho(0) = 0$, iii. $\rho(r_i) = \rho(-r_i)$ (symmetry), iv. $\rho(r_i) \geq \rho(r_j)$ for $|r_i| \geq |r_j|$ (monotonicity), v. ρ is continuous (ρ is differentiable).

A summary of such iterative methods is given by Holland and Welch[6].

One of the popularly known M-estimators is the Huber estimator. The $\rho(.)$ -function of Huber estimator is given by

$$\rho(r) = \begin{cases} \dfrac{r^2}{2} & |r| < c \\ c(|r| - \dfrac{c}{2}) & |r| \geq c \end{cases} \tag{3}$$

and the ψ-function is defined as

$$\psi(r) = \begin{cases} r & |r| < c \\ c & |r| \ge c \end{cases} \quad (4)$$

where *c* is an arbitrary value known as tuning constant. For Huber function, a tuning constant of *c* =1.345 yields 95% efficiency on the normal distribution.

It is to be noted that the Huber proposal does not downweight the large residuals which is a major drawback of Huber estimator. This drawback can be overcome by considering the next class of robust estimators, called as redescending M-estimators for which the score function $\psi(r) \to 0 \;\; as \;\; r \to \infty$.

Redescending M-estimators

Another important class of estimators is the "smoothly redescending" M-estimators with ψ-functions redescending to zero. "These estimators are constructed in order to remove clear outlying observations, to avoid the hard jump of elimination procedures and to substitute it by continuous transition of treating the data from fully good to fully bad"[7]. Redescending M-estimators are those estimators that are designed in such a way so that extreme outliers are completely rejected. In other words, an M-estimator is said to be a redescending M-estimator, for which the ψ-function redescends. That is, if its ψ-function satisfies $\lim_{r_i \to \pm\infty} \psi(r_i) = 0$.

Several researchers have proposed various redescending M-estimators[8]. Practically, all these redescending M-estimators are very similar computationally and yield almost similar outcomes. A three-part redescending function was proposed by Hampel in the "Princeton Robustness Study"[9] and is defined as follows

$$\psi(r) = \begin{cases} r & ,|r| \le a \\ a \times sign(r) & ,a < |r| \le b \\ a\dfrac{c-|r|}{c-b} sign(r) & ,b < |r| \le c \\ 0 & ,|r| > c \end{cases} \quad (5)$$

where the constant a, b, c are called tuning constants satisfying 0$< a \le b < c < \infty$. Hampel's three part function has revealed very good performance in the Princeton study. The $\psi(r)$ function of Hampel estimator indicates that it is still not a very good one as there are sudden changes in the slope of its $\psi(r)$-function. There is still a need of smoothed $\psi(r)$-function. The Hampel's $\psi(r)$-function lacks the property of differentiability, and thus a smooth redescending $\psi(r)$-function would be desired. According to Winsor's principle, any $\psi(r)$-function which is linear in the center, yields better efficiency in case of normal distribution[10]. To achieve this objective, Andrew's sine function[9] was developed with $\psi(r)$ function given by

$$\psi(r) = \begin{cases} c\sin(r/c) & ,|r| < \pi \\ 0 & ,|r| > \pi \end{cases} \quad (6)$$

This function is continuous and hence differentiable. Another popular redescending M-estimator is Tukey Bisquare[11] with ψ given by

$$\psi(r) = \begin{cases} r\{1-(r/c)^2\}^2 & ,|r| < c \\ 0 & ,|r| > c \end{cases} \quad (7)$$

For Tukey Biweight function, *c* = *4.685* gives 95% efficiency on normal distribution. M-estimators have a low break down point[12], that is, 1/*n*. It shows that these estimators can be biased even by one outlier. The scenario totally changes if outliers are only in the y- direction but not in the x- direction[13,14].

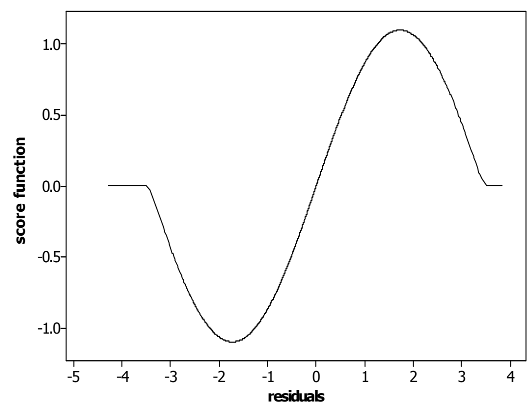

Figure-1
Andrew score function

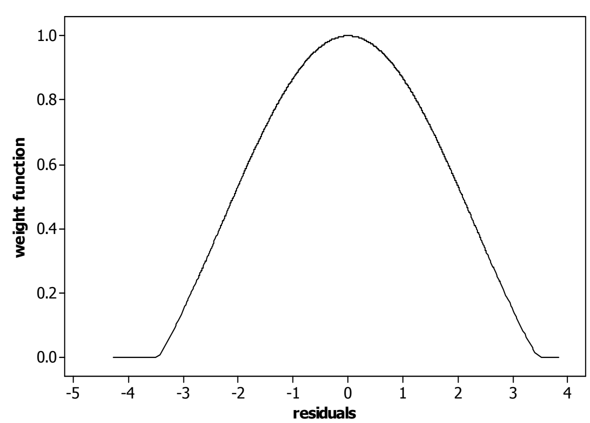

Figure-2
Andrew weight function

For a comparison of the most popular robust estimators including, Huber's M-estimator, Tukey's bisquare and Hampel estimator etc. through simulation study see Muthukrishnan and Radha[15]. They discussed some asymptotic properties of M-estimators. They argued that while choosing the redescending ψ-functions, one should take care that the ψ-functions of the redescending M-estimator does not descend very steeply to zero because that may have a ruthless influence on the asymptotic variance. In this paper, we propose a more robust and efficient redescending M-estimator that does not descend very steeply and is compared with the existing M and redescending M-estimators.

The Proposed Estimator

The proposed estimator, known as "Alamgir Redescending M-Estimator abbreviated as (ALARM)", is based on a modified tangent hyperbolic (tanh) type weight function.Consider tanh function given by

$$\tanh(r) = \frac{1 - e^{-2r}}{1 + e^{2r}} \qquad (8)$$

Replacing $2r$ by $kr^2$to have a modified tangent hyperbolic function given by

$$g(r) = \frac{1 - e^{-kr^2}}{1 + e^{kr^2}} \qquad (9)$$

where k is the normalizing constant. The derivative of this function is given by

$$\psi_1(r) = \frac{4re^{-r^2}}{(1 + e^{-r^2})^2}, \qquad k = 1 \qquad (10)$$

The corresponding weight function, $w_1(r)$, is given by

$$w_1(r) = \psi_1(\mathrm{r})/\mathrm{r} = \frac{4e^{-r^2}}{(1 + e^{-r^2})^2} \qquad (11)$$

Introduce a tuning constant parameter "*C*" in the above weight function to have

$$\mathrm{w}_1(\mathrm{r}) = \frac{4e^{-(r/c)^2}}{(1 + e^{-(r/c)^2})^2} \qquad (12)$$

The behavior of the weight function, $\mathrm{w}_1(\mathrm{r})$, is described in figure 4.

The drawback of this weight function is that it assigns rapidly decreasing weights to even good observations in the center of the data, thereby, reducing efficiency of the estimator. Tukey Bisquare weight function also suffers from the same problem.To overcome the drawback of these weight functions, we propose an alternative and a more efficient weight function given below

$$\mathrm{w}(\mathrm{r}) = (\mathrm{w}_1(r))^2 = \begin{cases} \dfrac{16e^{-2(r/c)^2}}{(1 + e^{-(r/c)^2})^4} & ,|r| \le c \\ 0 & ,|r| > c \end{cases} \qquad (13)$$

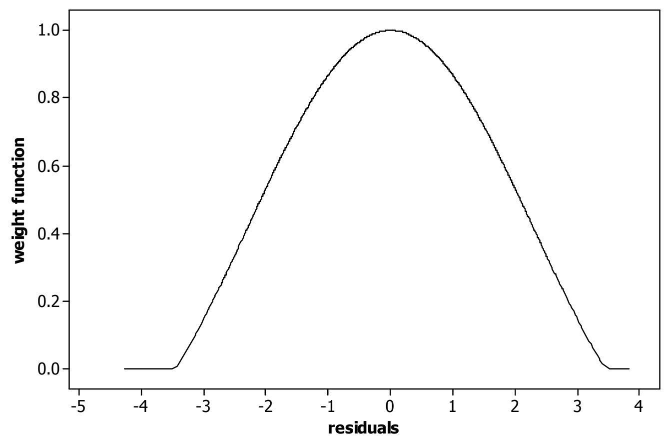

Figure-3
Graph of the W1(r), weight function

The above proposed function assigns weights close to "1" to majority of the good observations in the data set and consequently slow decreasing weights to outlying observation as it exhibits more linearity in the middle. The leverage points (extreme outliers) are assigned weight zero.

The weight function, given in (14), corresponds to the proposed new linear and slow decaying ψ-function given by

$$\psi(\mathrm{r}) = \begin{cases} \dfrac{16re^{-2(r/c)^2}}{(1 + e^{-(r/c)^2})^4} & ,|r| \le c \\ 0 & ,|r| > c \end{cases} \qquad (14)$$

The functional relationship between ψ is given by

$$\psi(r) = \frac{d}{dr}\rho(r) \qquad (15)$$

Integrating out the ψ-function under the initial conditions, we get the corresponding $\rho(r)$, given by

$$\rho(\mathrm{r}) = \begin{cases} \dfrac{2c^2}{3}[1 - \dfrac{2(1 + 3e^{(\frac{r}{c})^2})}{(1 + e^{(\frac{r}{c})^2})^3}] & ,|r| \le c \\ \dfrac{2c^2}{3}[1 - \dfrac{2(1 + 3e)}{(1 + e)^3}] & ,|r| > c \end{cases} \qquad (16)$$

The newly proposed score function $\rho(r)$ satisfies the standard properties of an ideal score function $\rho(r)$. The proposed weight function $w(r) = \psi(r)/r$, is as follows:

$$w(r) = \begin{cases} \dfrac{(4e^{-(r/c)^2})^2}{(1 + e^{-(r/c)^2})^4} & ,|r| \le c \\ 0 & ,|r| > c \end{cases} \qquad (17)$$

Where "*r*" denotes residual and "*c*" is tuning constant.

Robustness and efficiency are the two properties of a robust procedure which are inversely related. So one should choose an estimator with maximum resistance and the one that results in minimum loss of efficiency. Nobody can choose a highly robust estimator which is resistant to outlier at the cost of decreasing efficiency. There should be a compromise between these two properties. The proposed weight function is a symmetric function that assigns weights to the residuals. Figure (4-6) clearly shows symmetry of the function. The weight function ensures maximum weight (close to 1) assigned to the residuals corresponding to majority of the good observations. The $\rho(r)$, *w(r)* and $\psi(r)$ functions of ALARM estimator are plotted in figure (4-6).

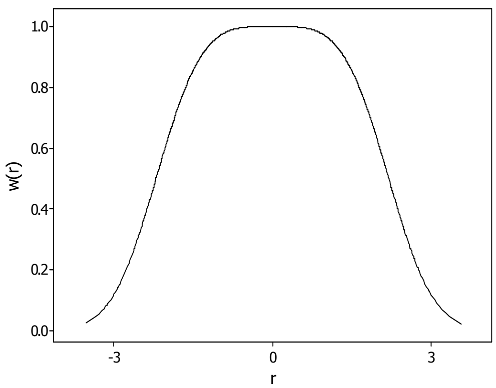

Figure-4
Proposed Weight function

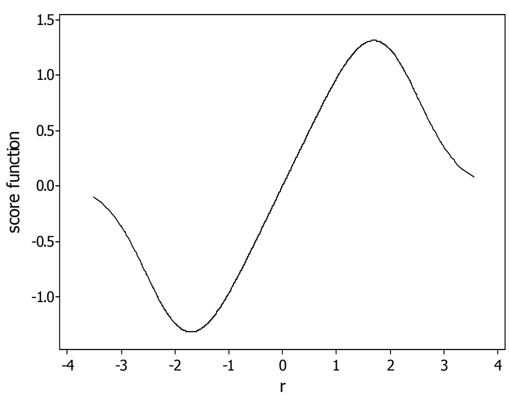

Figure-5
Proposed ψ -function

This fact differentiates the proposed estimator from all the other estimators and enables the proposed estimator to be preferred over others in majority of its applications. Only extremely outlying observations (bad leverage points) receive zero weights. The proposed weight function is a continuous and differential function. The proposed estimator performs soft trimming. Unlike Tukey's weight function and many more, the proposed weight decreases too slowly in the center and thus, having more linearity in the middle ensuring utilization of full information from good observations and less relying on extreme outliers. The proposed ψ -function is highly linear in the center yielding enhanced efficiency. The proposed estimator treats the middle observations nearly like OLS and then redescends. For empirical study of the proposed estimator, we set the tuning constant c = 3as it gives approximately 95% efficiency at normal case.

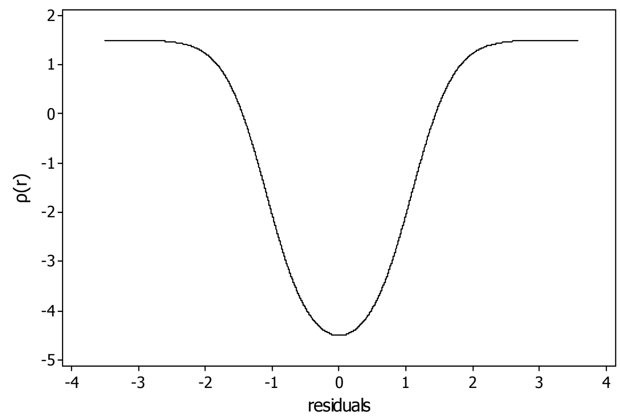

Figure-6
Proposed objective function

Computational Algorithm

Consider the regression model given in (1). We adopt the following algorithm for all competing estimators being considered in the study.

Step1: Choose initial estimates of β. We use LTS as starting value which is a high break down point estimator.

Step2: Compute the residuals, r_i , and also the corresponding weights based on the weight function from the previous iteration. Usually the residuals are scaled by applying some suitable scale estimate. In practice the variance, σ^2, is unknown. A very good choice for scale estimate due to high degree of robustness is the Median Absolute Deviation (MAD) is given by

$$MAD = Median\left|r_i - Median(r_i)\right| / 0.6745$$

Step3: Calculate new estimates of the regression coefficients by performing weighted least squares, that is

$$\hat{\beta}_i = (X' W^{(i-1)} X)^{-1} X' W^{(i-1)} Y$$

whereW is a diagonal matrix of weights based on specified weight function.

Step4: Repeat step 2- 3 until convergence or using some stopping criteria. In our case, we terminate the process when the maximum relative change is less than 0.000001.

Simulation Studies

Simulation studies are carried out to take into account various factors affecting the model fit and to be able to generalize the results to almost all types of data sets on the basis of artificially generated data. Our simulation study is based on the following simulation design, each consisting of 3000 runs.

The proposed (ALARM) estimator is compared with Huber, Tukey Biweight, Andrew sine, Hampel estimators and also with the classical OLS estimator.

Consider all simulations of one estimate, $\hat{\beta}_i = (\hat{\beta}_{i1}, \hat{\beta}_{i2}, \ldots, \hat{\beta}_{ip}), 1 \le i \le g, g = 3000.$ For comparison purpose, total absolute bias (TAB) and the TMSE[16] are computed for the estimates and are averaged over 3000 simulations for various sample sizes and number of predictors *p* = 1, 2, 5. The total bias, TAB, is computed from all replications as

$$TAB = \frac{1}{g}\sum_{i}^{g}\sum_{j}^{p}\left|\hat{\beta}_{ij} - \beta_{ij}\right| \qquad (18)$$

And the total MSE, denoted by TMSE, is calculated from all replications as

$$TMSE = \frac{1}{g}\sum_{i}^{g}\sum_{j}^{p}(\hat{\beta}_{ij} - \beta_{ij})^2 \qquad (19)$$

We generated data from various symmetric distributions of which two are heavy tailed distributions. Different simulation scenarios are as follows:

Generating both X and Y independently from: i. N (0, 1). ii. Student t distribution, T(3), iii. Double exponential distribution L (0, 1).

Next we consider samples containing $\lambda = 10\%$, 20% and 40% outliers in either Y- direction or in both XY- directions. The contaminations are done from normal distributions with different values of the parameters.

Results and Discussion

Numerical (simulation): Based on the simulation scenarios, the simulated AB's and TMSE's results are presented in tables (1-9).

Efficiency and Break Down Point: Bias, efficiency and break down point are the three desiring properties for almost all robust methods. Table1 summarizes the results based on clean data from normal distribution. In case of simulation from Normal distribution, $\lambda = 0\%$, OLS estimates have lowest total bias (simulation bias) and have smallest TMSE. ALARM estimator is a good competitor of the OLS in terms of total bias and TMSE as compared to all other estimators. None of the remaining estimators performs as efficiently as does the ALARM estimator.

When data is generated from t(3) (table 2), the ALARM estimator yielded lowest total bias as compared to OLS and all other robust estimators. Hampel, Tukey and Andrew yielded on the average approximately the same bias. The total bias of all these estimators tend to decrease as the sample size increases. Here, Hampel, Huber, Andrew and Tukey estimators have almost the same TMSE and hence equally more efficient estimators than ALARM estimator. The performance of ALARM estimator improves, however, with increasing sample size.

Table-1
Bias and efficiency of the estimates (X from N(0, 1) and Y from N(0,1))

N	P	Estimate	OLS	Huber	Hampel	Andrew	Tukey	ALARM
n=30	P=1	TAB TMSE	0.008 0.300	0.009 0.300	0.009 0.300	0.011 0.300	0.011 0.300	0.008 0.300
	P=2	TAB TMSE	0.003 0.340	0.004 0.351	0.005 0.343	0.007 0.370	0.007 0.370	0.004 0.343
	P=5	TAB TMSE	0.005 0.524	0.007 0.548	0.006 0.551	0.009 0.742	0.008 0.732	0.005 0.528
n=50	P=1	TAB TMSE	0.013 0.198	0.015 0.206	0.015 0.202	0.015 0.210	0.015 0.210	0.013 0.199
	P=2	TAB TMSE	0.002 0.250	0.003 0.259	0.003 0.252	0.003 0.263	0.003 0.263	0.002 0.251
	P=5	TAB TMSE	0.003 0.366	0.010 0.386	0.008 0.376	0.009 0.371	0.008 0.411	0.006 0.367
n=100	P=1	TAB TMSE	0.005 0.138	0.006 0.144	0.006 0.141	0.006 0.144	0.005 0.144	0.005 0.138
	P=2	TAB TMSE	0.002 0.178	0.004 0.183	0.004 0.179	0.004 0.183	0.004 0.183	0.003 0.179
	P=5	TAB TMSE	0.002 0.242	0.004 0.258	0.004 0.253	0.004 0.261	0.004 0.261	0.003 0.247

Table-2
Bias and efficiency of the estimates (X and Y from t-distribution, t(3))

n	P	Estimate	OLS	Huber	Hampel	Andrew	Tukey	ALARM
n=30	P=1	TAB TMSE	0.009 (0.376)	0.009 (0.302)	0.006 (0.313)	0.009 (0.308)	0.008 (0.306)	0.004 (0.318)
	P=2	TAB TMSE	0.008 0.443	0.007 0.364	0.005 0.372	0.012 0.384	0.012 0.382	0.006 0.385
	P=5	TAB TMSE	0.011 0.671	0.009 0.507	0.007 0.532	0.009 0.660	0.009 0.652	0.007 0.540
n=50	P=1	TAB TMSE	0.002 0.290	0.002 0.214	0.002 0.224	0.003 0.216	0.002 0.215	0.001 0.239
	P=2	TAB TMSE	0.006 0.340	0.009 0.262	0.009 0.267	0.009 0.264	0.009 0.263	0.009 0.265
	P=5	TAB TMSE	0.007 0.458	0.004 0.362	0.0033 0.371	0.004 0.379	0.005 0.373	0.002 0.380
n=100	P=1	TAB TMSE	0.007 0.199	0.003 0.155	0.003 0.161	0.002 0.153	0.002 0.153	0.002 0.162
	P=2	TAB TMSE	0.004 0.229	0.004 0.174	0.004 0.179	0.004 0.175	0.004 0.175	0.004 0.178
	P=5	TAB TMSE	0.005 0.302	0.003 0.235	0.002 0.242	0.002 0.240	0.002 0.240	0.001 0.242

Table-3
Bias and efficiency of the estimates (X from L(0,1) and Y from L(0,1))

n	P	Estimate	OLS	Huber	Hampel	Andrew	Tukey	ALARM
n=30	P=1	TAB TMSE	0.007 0.339	0.002 0.297	0.002 0.308	0.001 0.302	0.001 0.301	0.001 0.324
	P=2	TAB TMSE	0.007 0.411	0.004 0.351	0.004 0.367	0.006 0.376	0.006 0.372	0.004 0.369
	P=5	TAB TMSE	0.008 0.566	0.005 0.509	0.005 0.534	0.005 0.687	0.006 0.677	0.005 0.530
n=50	P=1	TAB TMSE	0.006 0.245	0.004 0.217	0.004 0.228	0.003 0.223	0.003 0.223	0.003 0.223
	P=2	TAB TMSE	0.012 0.306	0.009 0.261	0.009 0.272	0.009 0.263	0.009 0.263	0.009 0.263
	P=5	TAB TMSE	0.007 0.566	0.005 0.509	0.007 0.534	0.005 0.687	0.006 0.677	0.005 0.530
n=100	P=1	TAB TMSE	0.004 0.174	0.002 0.150	0.003 0.159	0.003 0.152	0.003 0.152	0.003 0.159
	P=2	TAB TMSE	0.010 0.202	0.003 0.174	0.002 0.183	0.004 0.175	0.004 0.175	0.002 0.174
	P=5	TAB TMSE	0.004 0.278	0.004 0.237	0.003 0.248	0.004 0.241	0.004 0.240	0.003 0.240

Table-4
Bias and efficiency of the estimates (90% X from N(0, 1) and 10% X from N(20,1); 90% Y from N(0,1) and 10% Y from N(100,1))

N	P	Estimate	OLS	Huber	Hampel	Andrew	Tukey	ALARM
n=30	P=1	TAB TMSE	2.574 4.984	2.577 4.988	0.011 0.298	0.012 0.306	0.010 0.300	0.010 0.295
	P=2	TAB TMSE	1.704 3.632	1.715 3.698	0.020 0.368	0.050 0.382	0.050 0.381	0.010 0.366
	P=5	TAB TMSE	0.846 2.502	0.811 3.001	0.020 0.701	0.020 0.800	0.010 0.702	0.020 0.700
n=50	P=1	TAB TMSE	3.000 5.100	3.001 5.100	0.003 0.200	0.003 0.200	0.003 0.205	0.003 0.200
	P=2	TAB TMSE	2.001 4.000	2.000 4.001	0.003 0.300	0.004 0.301	0.004 0.301	0.003 0.300
	P=5	TAB TMSE	0.838 2.374	0.801 2.121	0.006 0.401	0.006 0.402	0.006 0.400	0.005 0.400
n=100	P=1	TAB TMSE	2.574 4.984	3.002 4.712	0.004 0.151	0.004 0.151	0.004 0.154	0.004 0.151
	P=2	TAB TMSE	1.691 3.533	2.001 4.012	0.0060 0.201	0.0060 0.201	0.0060 0.201	0.0060 0.201
	P=5	TAB TMSE	0.840 2.293	0.851 2.002	0.003 0.274	0.002 0.278	0.002 0.278	0.003 0.273

Table-5
Bias and efficiency of the estimates (80% X from N(0,1), 20% X from N(20,1), 80% Y from N(0,1), 20% Y from N(100,1))

N	P	Estimate	OLS	Huber	Hampel	Andrew	Tukey	ALARM
n=30	P=1	TAB TMSE	2.609 5.054	2.618 5.094	0.007 0.298	0.008 0.301	0.008 0.301	0.006 0.300
	P=2	TAB TMSE	1.719 3.659	1.721 3.701	0.001 0.400	0.001 0.410	0.001 0.500	0.001 0.400
	P=5	TAB TMSE	0.849 2.508	0.843 2.577	0.101 1.101	0.101 1.102	0.101 1.102	0.101 1.101
n=50	P=1	TAB TMSE	2.605 4.989	2.609 5.012	0.010 0.200	0.020 0.203	0.020 0.203	0.010 0.200
	P=2	TAB TMSE	1.720 3.596	1.719 3.626	0.003 0.331	0.003 0.331	0.003 0.332	0.003 0.292
	P=5	TAB TMSE	0.842 2.386	0.836 2.428	0.012 0.492	0.013 0.503	0.013 0.503	0.012 0.492
n=100	P=1	TAB TMSE	2.609 5.054	2.628 4.983	0.004 0.156	0.004 0.158	0.004 0.158	0.004 0.156
	P=2	TAB TMSE	1.719 3.659	1.713 3.567	0.005 0.197	0.005 0.199	0.005 0.199	0.004 0.197
	P=5	TAB TMSE	0.849 2.508	0.841 2.320	0.003 0.281	0.003 0.284	0.003 0.283	0.003 0.280

Table-6
Bias and efficiency of the estimates (80% X from N(0, 1) and 20% from N(20,1); 80% Y fromN(0,1) and 20% Y from N(10,5))

N	P	Estimate	OLS	Huber	Hampel	Andrew	Tukey	ALARM
n=30	P=1	TAB TMSE	0.255 0.549	0.244 0.539	0.007 0.310	0.007 0.315	0.008 0.315	0.006 0.309
	P=2	TAB TMSE	0.170 0.629	0.166 0.516	0.116 0.489	0.102 0.490	0.102 0.489	0.102 0.477
	P=5	TAB TMSE	0.087 1.037	0.085 0.693	0.067 0.655	0.062 0.770	0.062 0.760	0.061 0.661
n=50	P=1	TAB TMSE	0.250 0.520	0.242 0.513	0.209 0.485	0.131 0.408	0.137 0.414	0.3155 0.434
	P=2	TAB TMSE	0.168 0.526	0.169 0.449	0.143 0.435	0.109 0.401	0.111 0.403	0.121 0.401
	P=5	TAB TMSE	0.084 0.769	0.085 0.526	0.074 0.522	0.067 0.499	0.068 0.500	0.066 0.500
n=100	P=1	TAB TMSE	0.257 0.510	0.250 0.501	0.245 0.497	0.170 0.406	0.175 0.411	0.169 0.402
	P=2	TAB TMSE	0.171 0.629	0.170 0.399	0.164 0.398	0.137 0.352	0.130 0.356	0.131 0.347
	P=5	TAB TMSE	0.087 1.037	0.083 0.383	0.079 0.389	0.072 0.359	0.084 0.361	0.067 0.359

Table-7
Bias and efficiency of the estimates (60% X from N(0,1) and 40% from N(20,1); 60% Y from N(0,1), 40% Y from N(100,1))

N	P	Estimate	OLS	Huber	Hampel	Andrew	Tukey	ALARM
n=30	P=1	TAB TMSE	2.677 5.121	2.694 5.161	0.0203 0.5232	0.0195 0.5237	0.0196 0.5237	0.0212 0.5246
	P=2	TAB TMSE	1.737 3.701	1.736 3.727	0.1836 1.266	0.1831 1.280	0.1832 1.279	0.1839 1.265
	P=5	TAB TMSE	0.8487 2.5346	0.8487 2.572	0.7946 2.5029	0.7883 2.7316	0.7872 2.723	0.7944 2.519
n=50	P=1	TAB TMSE	2.6663 5.0447	2.669 5.0619	0.0140 0.2578	0.0147 0.2582	0.0147 0.2582	0.0139 0.2578
	P=2	TAB TMSE	1.7318 3.6292	1.7280 3.6451	0.0209 0.5216	0.0210 0.5286	0.0210 0.5282	0.0207 0.5196
	P=5	TAB TMSE	0.8487 2.5346	0.8413 2.4170	0.6460 2.1273	0.6466 2.1594	0.6465 2.1571	0.6461 2.1231
n=100	P=1	TAB TMSE	2.6612 5.0143	2.6792 5.0234	0.0018 0.1817	0.0018 0.1820	0.0018 0.1820	0.0018 0.1817
	P=2	TAB TMSE	1.7265 3.5766	1.7332 3.5860	0.0040 0.2295	0.0042 0.2296	0.0042 0.2296	0.0040 0.2295
	P=5	TAB TMSE	0.8467 2.3106	0.8452 2.3215	0.3626 1.5480	0.3626 1.5519	0.3626 1.5516	0.3624 1.5472

Table-8
Bias and efficiency of the estimates (X from N(0, 1); 80% Y from N(0,1) and 20% Y from N(100,1))

N	P	Estimate	OLS	Huber	Hampel	Andrew	Tukey	ALARM
n=30	P=1	TAB TMSE	10.282 21.402	0.235 0.613	0.007 0.310	0.008 0.315	0.008 0.315	0.006 0.309
	P=2	TAB TMSE	6.908 22.81	0.170 0.725	0.005 0.395	0.006 0.403	0.006 0.402	0.005 0.393
	P=5	TAB TMSE	3.602 27.764	0.145 3.275	0.006 0.603	0.006 0.643	0.007 0.640	0.006 0.601
n=50	P=1	TAB TMSE	10.290 20.864	0.233 0.556	0.015 0.225	0.016 0.227	0.016 0.227	0.014 0.224
	P=2	TAB TMSE	6.766 21.785	0.166 0.641	0.002 0.285	0.002 0.288	0.002 0.288	0.002 0.284
	P=5	TAB TMSE	3.600 24.246	0.096 0.854	0.007 0.422	0.008 0.427	0.008 0.426	0.006 0.421
n=100	P=1	TAB TMSE	10.132 20.464	0.249 0.538	0.004 0.156	0.004 0.158	0.004 0.158	0.003 0.155
	P=2	TAB TMSE	6.767 20.885	0.166 0.565	0.004 0.200	0.004 0.202	0.004 0.202	0.003 0.199
	P=5	TAB TMSE	3.378 22.152	0.089 0.665	0.004 0.283	0.004 0.285	0.004 0.285	0.004 0.282

Table-9
Bias and efficiency of the estimates (X from N(0, 1); 60% Y from N(0,1) and 40% Y from N(100,1))

N	P	Estimate	OLS	Huber	Hampel	Andrew	Tukey	ALARM
n=30	P=1	TAB TMSE	20.092 41.297	2.364 12.455	0.005 0.356	0.005 0.356	0.005 0.356	0.004 0.355
	P=2	TAB TMSE	13.660 42.357	3.718 23.847	0.005 0.466	0.005 0.467	0.005 0.467	0.004 0.465
	P=5	TAB TMSE	7.178 46.581	5.111 48.647	0.008 0.753	0.008 0.752	0.008 0.752	0.007 0.750
n=50	P=1	TAB TMSE	20.225 40.617	1.318 5.583	0.014 0.258	0.015 0.258	0.015 0.258	0.013 0.257
	P=2	TAB TMSE	13.644 41.376	2.116 14.146	0.003 0.339	0.003 0.339	0.004 0.401	0.003 0.339
	P=5	TAB TMSE	7.011 43.40	3.451 34.80	0.008 0.499	0.008 0.499	0.008 0.499	0.008 0.499
n=100	P=1	TAB TMSE	20.23 40.36	1.092 2.295	0.002 0.182	0.002 0.182	0.002 0.182	0.001 0.181
	P=2	TAB TMSE	13.48 40.63	0.850 4.248	0.006 0.231	0.006 0.231	0.009 0.269	0.005 0.231
	P=5	TAB TMSE	6.873 41.637	0.905 11.334	0.005 0.331	0.005 0.332	0.005 0.332	0.004 0.331

Based on the data generated from heavy tailed Laplace distribution, the results in table 3 show that the performance of ALARM estimator gets better for increasing sample size and increasing *p*. OLS has the maximum possible total bias in estimating the model parameters whereas ALARM estimator has the least total bias among all estimators considered and its bias further tends to decrease with *n*.

The results summarized in table4 shows that OLS gives worst results with regards to TAB and TMSE. The remaining robust estimators offer negligible amount of bias and are almost equally efficient. However, ALARM estimator is the most efficient one as its TMSE is the smallest and becomes more and more efficient as sample size increases. Table 5 presents the results for more extreme contamination, that is, for 20% outliers. The proposed estimator outperforms all other estimators for all *n* and *p* as it has the smallest TMSE among all estimators. Except OLS and Huber estimators, all other robust estimators offer the same amount of TAB but are comparatively more biased contrary to the case of 10% outliers present in the data. Also the TMSE's are inflated for the robust estimators due to increased percentage of outliers. Table 6 presents results based on 20% contamination in Y-direction**.** Due to larger variance of the contaminating distribution, the performance of all estimators is affected yielding comparatively more deviating estimates of the parameters yielding larger TAB's and TMSE's, particularly, when *p* is large. But again ALARM estimator comparatively performs better than rest of the estimators. However, the results are not very promising for any of the estimators under consideration.

Table 7 contains results based on the highest percentage of contamination in XY- direction ($\lambda=$40%). The results for all estimators are not much promising as they are in case of lower percentage of outliers. Among all estimators, even for $\lambda=$40% in both XY- directions, ALARM estimator provides comparatively better results in terms of TMSE for almost all *n* and *p*. With increasing sample size, it provides more efficient results with regards to TMSE. Hampel estimator is the second best performer. Thus, the results indicate that ALARM estimator has highest break down point among the competing estimators in case of outliers in XY- direction.

Tables (8-9) summarize the results based on percentage of outliers $\lambda=$ 20% and 40% in Y-direction only. The results clearly indicate that ALARM estimator provides very close estimates as compared to other robust redescending estimators for all *n* and *p*. However, the results are not very promising for large *p*. Hampel estimator is a close competitor of ALARM estimator. The performance of ALARM estimator gets improved with increasing sample size. The results presented in Table9 ($\lambda=$40%) are almost similar to those of table 8 ($\lambda=$ 20%) for all robust estimators. But OLS provides worst estimates due to increased percentage of outliers. The winner of all the estimators, even in case of $\lambda=$40%, is ALARM estimator. Thus our proposed redescending estimator has the highest break down point in case of outliers in y-direction which is similarto argumentsmade researchers[13,14].

Figure 7 describes how the efficiency is affected by sample size and the number of predictor variables. It is obvious from the figure that the asymptotic efficiency of ALARM estimator is a non-decreasing monotonic function of the tuning constant. The efficiency gets its maximum around *c* = 3 for all sample sizes and all values of *p*.

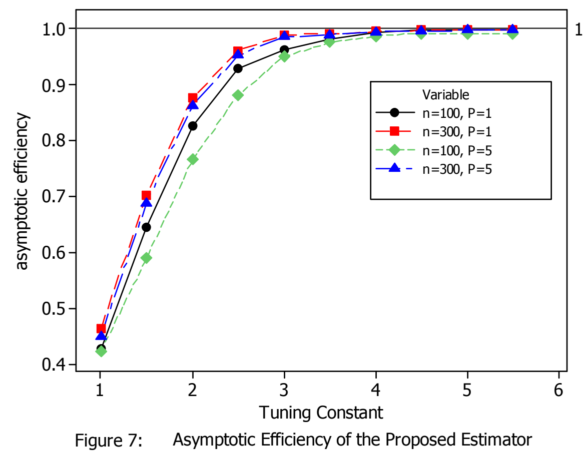

Figure-7
Asymptotic Efficiency of the Proposed Estimator

The efficiency of the proposed estimator is highest for *n* = 300 and *p* = 1. The larger the value of *p*, the lesser is the efficiency of ALARM estimator but the efficiency increases as *n* increases even for larger value of *p*.

Real Data Examples: We evaluate the performance of our proposed estimator along with five other estimators on real data sets. These real data sets are available in the literature. These data sets have widely been used by researchers in the context of robust statistics and robust regression.

Example 1 (Belgian phone calls data): Outliers in Y-space: This data set[17] is based on the number of international calls from Belgian in the year 1950 to 1973. The data contain 24 data points on 2 variables. The two variables are: the number of calls received (Y) in tens of millions and the year(X). This data set has widely been analyzed in the literature by various researchers in the context of robust regression and detection of outliers. The data set contain outliers in Y-direction. The results presented in table 10 indicate that all the robust methods, including Huber estimator, provide much closed estimates.

Table-10
Estimates of the model parameters for Phone calls data

Method	Coefficient		Data points Used	WRSS*
	intercept	X		
OLS	-26.006	0.504	24	695.435
Huber	-7.203	0.146	17	0.600
Hampel	-5.223	0.110	17	0.202
Andrew	-5.211	0.109	17	0.114
Tukey	-5.211	0.109	17	0.114
ALARM	-5.242	0.110	17	0.201

The estimates provided by our proposed estimator are very close to Hampel, Andrew and Tukey estimators. The robust fit and the residuals plots based on ALARM estimator (figure 8 and figure 9) clearly indicate that ALARM estimator makes use of good data points and clearly detecting 6 bad outliers and one good outlier. Majority of the data points are represented by our proposed estimator.

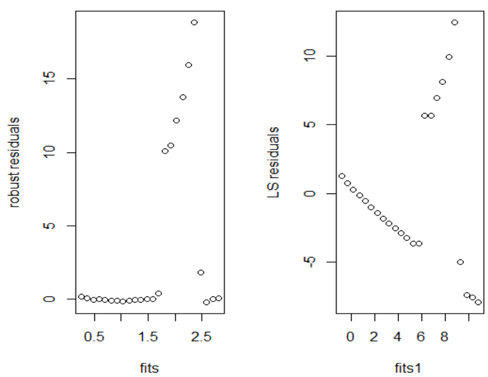

Figure-8
Robust and OLS residuals versus fits

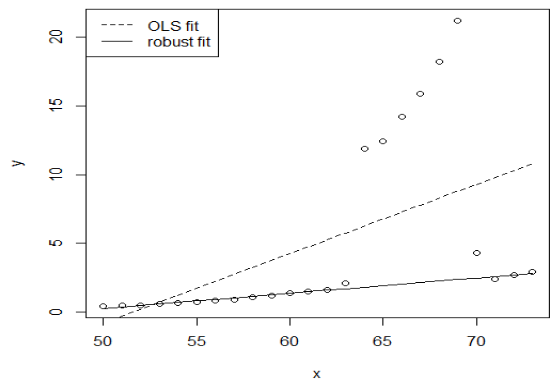

Figure-9
Robust residual versus fits

Example 3 (Heart Catherization Data): Outliers in X-space: This data set have been analyzed by various researchers for outliers and multicollinearity[18,19]. This data set contains 12 observations on three variables. The three variables are: catheter length (Y) in centimeters, patient height (X_1) in inches and patient weight (X_2) in pounds.

Table-11
Estimates of the model parameters for Heart data

Method	Coefficient			Data points used	WRSS*
	Intercept	X_1	X_2		
OLS	20.376	0.211	0.191	12	128.479
Huber	26.521	0.012	0.248	12	7.095
Hampel	63.362	-1.227	0.688	8	2.887
Andrew	63.525	-1.229	0.689	8	2.599
Tukey	63.523	-1.229	0.689	8	2.917
ALARM	63.359	-1.227	0.688	8	2.884

*Weighted Residual Sum of Squares

Table-12
Fitted Values based on OLS and various redescending M-estimators

Y	OLS	Huber	Hampel	Andrew	Tukey	ALARM
37	37.03	36.97	38.39	38.45	38.45	38.39
50	51.62	50.49	49.83	49.86	49.86	49.89
34	35.06	35.79	41.79	41.87	41.87	41.73
36	34.43	34.45	35.55	5.62	35.62	35.68
43	39.90	39.98	43.34	43.40	43.40	43.31
28	31.73	31.22	27.83	27.90	27.89	27.89
37	36.79	36.60	37.11	37.18	37.17	37.05
20	26.74	28.91	41.61	41.71	41.71	41.61
34	34.47	35.16	40.69	40.76	40.76	40.69
30	27.14	29.17	41.07	41.17	41.17	41.07
38	31.34	32.14	37.33	37.41	37.41	37.43
47	47.69	46.82	46.59	46.64	46.63	46.59

Table 11 and table 12 present the estimates obtained and the fitted values using OLS, ALARM estimator and other robust methods. OLS and Huber estimators are severely affected by the presence of 4 outliers and offer very poor estimates of the parameters. On the other hand, ALARM, Andrew, Hampel and Tukey estimators show resistance to those 4 outliers. ALARM estimator shows good promising results.

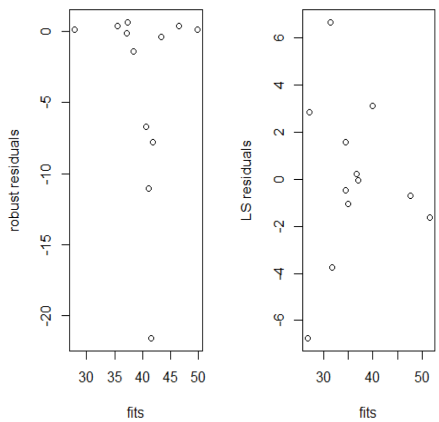

Figure-10
Robust and OLS residuals versus fits

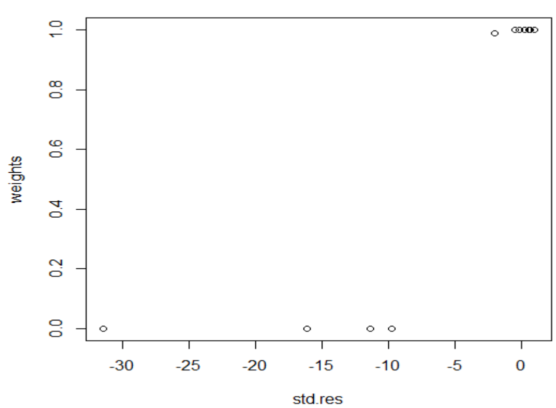

Figure-11
Weights versus Std. residuals

Figure 10 and figure 11 indicate the influence of outliers on the OLS and ALARM. The above figures also show that our proposed estimator detects all 4 outliers present in the data and shows resistant against them. The proposed estimator assigns maximum possible weight (very near to 1) to the good data points.

Conclusion

We proposed a new redescending M-estimator, called, ALARM estimator which is more linear in the middle, smooth and continuous. Hence, according to Winsor's principle, we conclude that our proposed estimator possesses optimal properties.

Simulation study and real data exampleswere used to evaluate and compare the performance of our proposed ALARM estimator with some estimators existing in the literature. In case of clean data from Normal distribution, the performance of ALARM was almost the same as that of OLS which and it outperformed rest of the estimators. In case of contaminated data (outliers in both XY- direction), ALARM estimator was found to be the best among all estimator for all *n* and *p* and for all percentages of outliers in the data as it offered promising results in all scenarios. The performance of ALARM estimator improved with increasing sample size. The breakdown point of ALARM estimator was found to be higher than rest of the estimators. The efficiency of ALARM estimator was also better than all other estimators. All robust estimators performed very well in case of all levels of contaminations in y-direction but ALARM estimator performed better than all redescending M-estimators.

References

1. Chauhan J.D., Modi C.K. and Pithadiya K.J., Comparison of Redescending and Monotone M Estimator for Surface Roughness Estimation Using Machine Vision, *Second International Conference in Emerging Trends in Engineering and Technology*, ICETET- 09, 464- 469 **(2009)**
2. Edgeworth F.Y., On Observations Relating to Several Quantities, *Hermathena*, **6**, 279-285 **(1887)**
3. Rousseeuw P.J. and Yohai V.J., Robust regression by means of S-estimators, in Robust and Nonlinear Time Series Analysis, (Franke, J., Hardle, W., and Martin, R. eds.), *Lecture Notes in Statistics*, No. **26**, Springer-Verlag, New York, 256–272 **(1984)**
4. Huber P.J., Robust estimation of a location parameter, *The Annals of Mathematical Statistics,* **35**, 73–101 **(1964)**
5. Black M.J. and Rangarajan A., On the unification of line processes, outlier rejection, and robust statistics with applications in early vision, *International Journal of Computer Vision*, **25(19),** 57–92 **(1996)**
6. Holland P.W. and Welch R.E., Robust regression using iteratively reweighted least squares, *Comm. in Statistics*, A**6**, 813-827 **(1977)**
7. Hampel F.R., Some Additional Notes on the 'Princeton Robustness Year', In: The Practice of Data Analysis: Essays in Honor of Tukey, J. W. eds. Brillinger, D. R., and Fernholz, L. T. Princeton: Princeton University Press, 133–153 **(1997)**

8. Wu C.J.F., Jackknife, bootstrap and other resampling methods in regression analysis (with discussions), *Ann. Statist.*, **14**, 1261- 1350 **(1986)**

9. Andrews D.F., Bickel P.J., Hampel F.R., Huber P.J., Rogers W.H. and Tukey J.W., Robust Estimates of Location: Survey and Advances, Princeton University Press, Princeton, New Jersey **(1972)**

10. Tukey J.W., A survey of sampling from contaminated distributions, In: I. Olkin (Ed.), Contributions to Probability and Statistics, (448-485), Palo Alto: Stanford University Press **(1960)**

11. Beaton A.E. and Tukey J.W., The fitting of power series, meaning polynomials, illustrated on banned-spectroscopic data, *Technometrics*, **16**, 147-185 **(1974)**

12. Maronna R.A., Bustos O.H. and Yohai V.J., Bias-and efficiency-robustness of general M-estimators for regression with random carriers in Smoothing Techniques for Curve Estimation (Gasser, T. and Rosenblatt, M. eds.), *Lecture Notes in Mathematics*, **757**, 91-116, Springer, Berlin **(1979)**

13. Ellis S.P. and Morgenthaler S., Leverage and breakdown in L1-regression, *J. Amer. Statist. Assoc.,* **87**, 143–148 **(1992)**

14. He X., Jureckova Koenker, R. and Portnoy S., Tail behavior of regression estimators and their breakdown points, *Econometrica*, **58**, 1195-1214 **(1990)**

15. Muthukrishnan R. and Radha M., M- Estimators in Regression models, *Journal of Mathematics Research*, **2(4)**, 23-27 **(2010)**

16. Bianco A.M., Ben M.G. and Yohai V.J., Robust Estimation for Linear Regression with Asymmetric Errors, *The Canadian Journal of Statistics*, **33(4)**, 511-528 **(2005)**

17. Rousseeuw P.J. and Leroy M.A., *Robust Regression and Outlier Detection*, John Wiley and Sons, New York **(1987)**

18. Weisberg S., *Applied Linear Regression,* Wiley, New York **(1980)**

19. Chambers J.M., Cleveland W.S., Kleiner B. and Tukey P.A., *Graphical Methods for Data Analysis,* Belmont, CA: Wadsworth **(1983)**