Malaysian Journal of Mathematical Sciences 14(2): 209–220 (2020)

MALAYSIAN JOURNAL OF MATHEMATICAL SCIENCES

Journal homepage: http://einspem.upm.edu.my/journal

A Modified Maximum Likelihood Estimator for the Parameters of Linear Structural Relationship Model

Mamun, A.S.M.A.[1], Hussin, A.G.[2], Zubairi, Y.Z. [3], and Rana, S.*[4]

[1] *Department of Statistics, University of Rajshahi, Bangladesh*
[2] *Centre for Foundation Studies in Science, Universiti Malaya, Malaysia*
[3] *Faculty of Defence Science and Technology, National Defence University of Malaysia, Malaysia*
[4] *Department of Mathematical & Physical Sciences, East West University, Bangladesh*

E-mail: srana@ewubd.edu
* *Corresponding author*

Received: 31 December 2018
Accepted: 20 March 2020

ABSTRACT

The maximum likelihood method is the best method for estimating the parameters of a linear structural relationship model. However, if the data contains outliers then sample estimates and subsequent results of the maximum likelihood method could be unreliable. Robust estimation techniques have become popular due to its resistance to outliers. Thus, we proposed a modified maximum likelihood method to estimate the parameters of a linear structural relationship model where the non-robust components of the maximum likelihood methods are replaced by their corresponding robust alternatives. The simulation study and real life examples show that the proposed method performs very well in estimating the parameters in terms of estimated bias and mean square error.

Keywords: Linear structural relationship model, maximum likelihood method, outliers, robustness.

1. Introduction

In linear regression analysis, the explanatory variables are assumed to be fixed and measured without error. But in reality, this assumption often does not hold due to many practical reasons and inherent measurement errors arise into the observations. Ignorance of measurement errors directly affects the desirable criteria of point estimators or interval estimators. A large body of literature has been developed over the years (see Madansky (1959), Moran (1971), Kendall and Stuart (1973), Fuller (1987) and Cheng and Van Ness (1994)) on the errors-in-variables model (EIVM). The linear structural relationship model (LSRM) is one of the families in the EIVM which also includes functional, ultrafunctional and ultrastructural relationship models. Over the past thirty years, a large number of works have been done in LSRM (see, Birch (1964),Barnett (1967), Chan and Mak (1979), Lakshminarayanan and Gunst (1984), Reilman et al. (1985), Bolfarine and Cordani (1993), Hood et al. (1999)).

Suppose two random variables X and Y are linearly related as

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

It is assume that both X and Y variables are measured without error. However, in real life, sometimes these two variables can not measure without error. Suppose X_i and Y_i are measured with errors δ_i, ε_i respectively, then we can write these variable as follows.

$$\begin{aligned} x_i &= X_i + \delta_i \\ y_i &= Y_i + \varepsilon_i \end{aligned} \tag{2}$$

where $\delta_i \sim N(0, \sigma_\delta^2)$ and $\varepsilon_i \sim N(0, \sigma_\varepsilon^2)$

This states that the variances of error terms are homoscedastic. Kendall and Stuart (1973) described the structural model considering as normal distribution with mean μ and σ_X^2 variance. Now we use equation 2 to rewrite equation 1 as

$$y_i = \alpha + \beta x_i + (\varepsilon_i - \beta\delta_i)$$

which shows that both x_i and y_i are correlated with the error term $(\varepsilon_i - \beta\delta_i)$ and depends on β.

Most of the authors use the maximum likelihood method for estimating parameters in EIVM, which is non-robust. To make estimator robust many authors, Lieberman (2005), Koláček (2008), Alfaro and Ortega (2009), Liu (2012) used some modification of traditional method, where classical estimators are replaced by their corresponding robust estimators. In this study, we also propose a modification of maximum likelihood estimation method to estimate the parameters in a linear structural relationship model. This paper is organized as follows: In Section 2 we review the estimation of the parameters for LSRM using the maximum likelihood estimation method and propose a modified maximum likelihood estimation method. We report a Monte Carlo simulation experiment in Section 3 which is designed to investigate the performance of the new estimate of parameters in the presence of outlier(s). In Section 4, we apply the new method on the Iron in Slag data as given in Hand et al. (1993). Serum Kanamycin data taken from Kelly (1984) to investigate the performance of the proposed method.

2. Estimation of parameters in LSRM

In this section, we first review the maximum likelihood estimation (MLE) method for the estimation of slope and intercept in LSRM and then introduce the modified maximum likelihood estimation method for the same.

2.1 Maximum likelihood method

In LSRM, we assumed that the errors are normal and hence the bivariate normal distribution of x_i and y_i is given by

$$\begin{pmatrix} x_i \\ y_i \end{pmatrix} \sim N\left(\begin{bmatrix} \mu \\ \alpha + \beta\mu \end{bmatrix}, \begin{bmatrix} \sigma_X^2 + \sigma_\delta^2 & \beta\sigma_X^2 \\ \beta\sigma_X^2 & \beta^2\sigma_X^2 + \sigma_\epsilon^2 \end{bmatrix} \right)$$

Kendall and Stuart (1973) have shown that the above distribution yields five normal equations with six unknowns $(\mu, \alpha, \beta, \sigma_X^2, \sigma_\delta^2, \sigma_\varepsilon^2)$, hence an additional assumption is required for the unique and consistent solutions for the parameters of the model (1). In particular, Hood et al. (1999) discussed in detail the procedure to estimate the parameters of model (1) under various assumptions. However, for the case when ratio of error variances $(\lambda = \sigma_\varepsilon^2/\sigma_\delta^2)$

is assumed to be known Hood et al. (1999), the MLE for the parameters are given by

$$\hat{\beta}_{MLE} = \frac{(S_y^2 - \lambda S_x^2) + \sqrt{(S_y^2 - \lambda S_x^2)^2 + 4\lambda S_{xy}^2}}{2S_{xy}} \qquad (3)$$
$$\hat{\alpha}_{MLE} = \bar{y} - \hat{\beta}_{MLE}\bar{x}$$

where, S_x^2, S_y^2 and S_{xy} are the corrected sums of squares and products can be written in a convenient format as

$$S_x^2 = \frac{1}{n}\sum(x_i - \bar{x})^2, S_y^2 = \frac{1}{n}\sum(y_i - \bar{y})^2, S_{xy} = \frac{1}{n}\sum(x_i - \bar{x})(y_i - \bar{y})$$

2.2 Modified maximum likelihood method

The estimation of slope and intercept by the MLE method given in equation 3 contains some standard statistics such as mean, variance and covariance which are sensitive to outliers. In order to avoid the unfortunate consequences of the effect of outliers we propose a modified maximum likelihood estimation (MMLE) method here by replacing the usual estimators in equation (3) with robust estimators. To construct the MMLE for slope and intercept, the sample mean is replaced by sample median, i.e. $\bar{x}_{(Robust)} = median(x)$ and $\bar{y}_{(Robust)} = median(y)$, and the sample variances S_x^2 and S_y^2 given in 3 are replaced by estimated Q_n, which is proposed by Rousseeuw and Croux (1993), i.e.

Replace $S_x^2 = \frac{1}{n}\sum(x_i - \bar{x})^2$ by $\{Q_n(x)\}^2 = S_x^2$(Robust) and

Replace $S_y^2 = \frac{1}{n}\sum(y_i - \bar{y})^2$ by $\{Q_n(y)\}^2 = S_y^2$(Robust)

where, $Q_n(x) = d\{|x_i - x_j|; i < j\}_{(k)}$ and $Q_n(y) = d\{|y_i - y_j|; i < j\}_{(k)}$, where d is a constant factor chosen to provide consistency of estimation of the standard deviation of a normal distribution and $k = \binom{h}{2} \approx \binom{n}{2}/4$, where $h = [n/2] + 1$ is roughly half the number of observations (Rousseeuw and Croux (1993)) and the sample covariance S_{xy} is replaced by S_{xy}(Robust) , which can be written in the following form

$$S_{xy}(Robust) = r_{Q_n} * S_x(Robust) * S_y(Robust)$$

where r_{Q_n} is the robust correlation coefficient proposed by Shevlyakov and Smirnov (2011) defined as

$$r_{Q_n} = \frac{{Q_n}^2(u) - {Q_n}^2(v)}{{Q_n}^2(u) + {Q_n}^2(v)}$$

and u and v are the robust principle variables defined as

$$u = \frac{x - median(x)}{\sqrt{2}Q_n(x)} + \frac{y - median(y)}{\sqrt{2}Q_n(y)}$$

and

$$v = \frac{x - median(x)}{\sqrt{2}Q_n(x)} - \frac{y - median(y)}{\sqrt{2}Q_n(y)}$$

Now we replace the robust estimators in place of classical estimators in equation 3 and obtain the modified maximum likelihood estimator (MMLE) for the slope and intercept of LSRM as

$$\begin{aligned} \hat{\beta}_{MMLE} &= \frac{(S_y^2(Robust) - \lambda S_x^2(Robust)) + \sqrt{((S_y^2(Robust) - \lambda S_x^2(Robust))^2 + 4\lambda S_{xy}^2(Robust)}}{2S_{xy}(Robust)} \\ \hat{\alpha}_{MMLE} &= \bar{y}_{(Robust)} - \hat{\beta}_{MMLE}\bar{x}_{(Robust)} \end{aligned} \quad (4)$$

2.3 Simulation study

In this section, we carry out a simulation study to compare the performance of the MMLE method with the existing MLE method in the presence of outliers. Following Mamun et al. (2019), we simulate observations from the following model,

$$Y_i = 1 + X_i, x_i = X_i + \delta_i, y_i = Y_i + \varepsilon_i$$

where $X_i \sim N(20, 3)$ and $\delta_i, \varepsilon_i \sim N(0, 0.1)$. In the next step, we contaminate the data at each stage by replacing original observations by contaminated observations. We generate the contaminated data points, for example at point c for variable y in a way that the observation y_c is given by ${y_c}^* = 1 + X_c + \varepsilon_c + v$ where v is constant Mamun et al. (2012).We generate samples of size 20, 50 and 100 from the sampling distribution as mentioned earlier.In order to investigate the robustness of MMLE, simulation study is extended to non-normal error terms. Thus,the error terms δ_i and ε_i are generated from three different distributions, symmetric beta distribution with parameters (3, 3), right skewed beta distribution (2, 9) and left skewed beta distribution (9, 2). Furthermore, on the basis of estimated bias (EB) and mean square error (MSE) in 15,000 trials; we examine the properties of these two methods. Simulation results are given in Tables 1 to 8.

Table 1: EB and MSE of the slope: Normal-Case

Contamination	Methods	$n = 20$		$n = 50$		$n = 100$	
		EB	MSE	EB	MSE	EB	MSE
Without outlier	MLE	4.89E-05	1.32E-04	3.86E-05	4.77E-05	2.28E-05	2.19E-05
	MMLE	5.99E-02	1.35E-01	1.75E-02	3.58E-02	5.94E-03	1.22E-02
Single outlier	MLE	4.44E+00	5.15E+03	9.52E-01	9.75E-01	4.12E-01	1.79E-01
	MMLE	3.92E-01	2.43E-01	1.26E-01	1.26E-01	1.98E-01	5.46E-02
10%	MLE	8.21E+00	9.09E+04	6.76E+00	7.58E+02	5.99E+00	4.35E+01
	MMLE	4.04E-01	2.14E-01	3.87E-01	1.61E-01	3.84E-01	1.53E-01
20%	MLE	1.39E+01	4.17E+05	1.27E+01	1.88E+05	1.11E+01	9.04E+02
	MMLE	2.94E-01	1.13E-01	2.70E-01	7.85E-02	2.67E-01	7.39E-02
30%	MLE	8.48E+01	1.15E+08	2.10E+01	2.62E+06	1.82E+01	5.89E+04
	MMLE	2.16E-01	5.97E-02	2.00E-01	4.28E-02	1.99E-01	4.09E-02

Table 2: EB and MSE of the slope: Right Skewed-Case, Beta (2, 9)

Contamination	Methods	$n = 20$		$n = 50$		$n = 100$	
		EB	MSE	EB	MSE	EB	MSE
Without outlier	MLE	1.65E-04	1.65E-04	1.44E-05	6.00E-05	1.13E-05	2.77E-05
	MMLE	6.27E-02	1.34E-01	1.82E-02	3.43E-02	6.21E-03	1.17E-02
Single outlier	MLE	4.43E+00	6.37E+04	9.51E-01	9.77E-01	4.10E-01	1.78E-01
	MMLE	3.85E-01	2.32E-01	2.90E-01	1.17E-01	1.84E-01	4.87E-02
10%	MLE	6.95E+00	1.66E+06	5.56E+00	3.90E+04	3.01E+00	4.40E+01
	MMLE	3.99E-01	2.08E-01	3.86E-01	1.60E-01	3.81E-01	1.51E-01
20%	MLE	3.70E+01	1.15E+07	1.62E+01	5.13E+06	1.30E+01	2.56E+03
	MMLE	2.93E-01	1.11E-01	2.70E-01	7.88E-02	2.67E-01	7.39E-02
30%	MLE	9.78E+01	5.29E+08	6.55E+01	8.31E+07	2.10E+01	6.23E+05
	MMLE	2.16E-01	5.94E-02	2.01E-01	4.31E-02	1.99E-01	4.09E-02

Table 3: EB and MSE of the slope: Left Skewed-Case, Beta (9, 2)

Contamination	Methods	$n = 20$		$n = 50$		$n = 100$	
		EB	MSE	EB	MSE	EB	MSE
Without outlier	MLE	1.51E-04	1.65E-04	1.33E-05	6.00E-05	1.11E-05	2.77E-05
	MMLE	5.05E-02	1.27E-01	1.46E-02	3.41E-02	6.99E-03	1.18E-02
Single outlier	MLE	2.34E+00	1.15E+05	9.51E-01	9.76E-01	4.10E-01	1.78E-01
	MMLE	3.79E-01	2.28E-01	2.87E-01	1.15E-01	1.84E-01	4.86E-02
10%	MLE	4.33E+01	6.67E+05	7.29E+00	1.61E+03	6.09E+00	1.57E+02
	MMLE	3.97E-01	2.06E-01	3.86E-01	1.60E-01	3.81E-01	1.50E-01
20%	MLE	1.08E+02	7.89E+07	1.12E+01	3.86E+06	1.06E+01	1.84E+05
	MMLE	2.93E-01	1.11E-01	2.71E-01	7.90E-02	2.67E-01	7.39E-02
30%	MLE	1.53E+02	2.89E+08	1.44E+01	1.49E+07	1.24E+01	3.72E+06
	MMLE	2.80E+00	5.88E-02	2.01E-01	4.31E-02	1.99E-01	4.09E-02

Table 4: EB and MSE of the slope: Non-Normal Symmetric-Case, Beta (3, 3)

Contamination	Methods	$n = 20$		$n = 50$		$n = 100$	
		EB	MSE	EB	MSE	EB	MSE
Without outlier	MLE	4.77E-04	4.73E-04	1.59E-04	1.71E-04	2.40E-05	8.19E-05
	MMLE	4.75E-02	1.02E-01	1.29E-02	2.35E-02	5.20E-03	8.10E-03
Single outlier	MLE	4.14E+00	3.03E+03	9.52E-01	9.83E-01	4.10E-01	1.79E-01
	MMLE	3.07E-01	1.66E-01	1.92E-01	6.12E-02	1.11E-01	2.19E-02
10%	MLE	1.85E+01	1.72E+06	6.91E+00	4.07E+02	6.03E+00	6.38E+01
	MMLE	3.57E-01	1.69E-01	3.58E-01	1.37E-01	3.57E-01	1.31E-01
20%	MLE	3.19E+01	3.43E+06	9.17E+00	7.11E+05	7.02E+00	2.13E+04
	MMLE	2.85E-01	1.03E-01	2.66E-01	7.62E-02	2.63E-01	7.12E-02
30%	MLE	3.55E+01	1.08E+07	1.34E+01	5.22E+06	1.18E+01	7.54E+05
	MMLE	2.14E-01	5.78E-02	1.99E-01	4.25E-02	1.97E-01	4.01E-02

Table 5: EB and MSE of the intercept: Normal-Case

Contamination	Methods	$n = 20$		$n = 50$		$n = 100$	
		EB	MSE	EB	MSE	EB	MSE
Without outlier	MLE	1.09E-03	5.42E-02	2.78E-04	1.95E-02	5.13E-05	8.96E-03
	MMLE	1.20E+00	5.39E+01	3.50E-01	1.44E+01	1.19E-01	4.92E+00
Single outlier	MLE	8.66E+01	1.99E+06	1.85E+01	3.72E+02	7.98E+00	6.77E+01
	MMLE	7.65E+00	9.45E+01	6.01E+00	4.96E+01	3.92E+00	2.15E+01
10%	MLE	1.61E+02	3.67E+07	1.33E+02	3.03E+05	1.17E+02	1.68E+04
	MMLE	7.67E+00	8.01E+01	7.33E+00	5.86E+01	7.26E+00	5.49E+01
20%	MLE	2.70E+02	1.62E+08	2.50E+02	7.06E+06	2.44E+02	3.58E+05
	MMLE	4.94E+00	3.68E+01	4.45E+00	2.27E+01	4.39E+00	2.06E+01
30%	MLE	1.73E+03	4.83E+10	4.12E+02	4.73E+07	3.57E+02	2.33E+07
	MMLE	2.63E+00	1.45E+01	2.31E+00	7.36E+00	2.29E+00	6.16E+00

Table 6: EB and MSE of the intercept: Right Skewed-Case, Beta (2, 9)

Contamination	Methods	$n = 20$		$n = 50$		$n = 100$	
		EB	MSE	EB	MSE	EB	MSE
Without outlier	MLE	4.09E-03	6.83E-02	3.95E-04	2.49E-02	7.09E-05	1.14E-03
	MMLE	1.26E+00	5.48E+01	3.67E-01	1.40E+01	1.25E-01	4.79E+00
Single outlier	MLE	8.77E+01	2.72E+07	1.17E+01	3.79E+02	8.03E+00	6.85E+01
	MMLE	7.57E+00	9.20E+01	5.77E+00	4.67E+01	3.67E+00	1.96E+01
10%	MLE	8.27E+01	6.86E+08	1.90E+01	1.55E+07	9.19E+00	1.74E+04
	MMLE	7.63E+00	7.92E+01	7.36E+00	5.91E+01	7.28E+00	5.52E+01
20%	MLE	1.56E+02	4.79E+09	3.48E+01	2.05E+08	2.57E+01	1.05E+05
	MMLE	4.96E+00	3.66E+01	4.50E+00	2.32E+01	4.44E+00	2.11E+01
30%	MLE	7.09E+03	8.86E+10	3.81E+02	2.65E+09	3.35E+02	2.89E+07
	MMLE	2.68E+00	1.47E+01	2.35E+00	7.61E+00	2.33E+00	6.38E+00

Table 7: EB and MSE of the intercept: Left Skewed-Case, Beta (9, 2)

Contamination	Methods	$n = 20$		$n = 50$		$n = 100$	
		EB	MSE	EB	MSE	EB	MSE
Without outlier	MLE	7.29E-02	7.90E-03	2.66E-02	8.25E-04	1.51E-03	1.23E-04
	MMLE	1.05E+00	5.53E+01	3.03E-01	1.48E+01	1.45E-01	5.13E+00
Single outlier	MLE	4.98E+01	4.75E+07	1.93E+01	4.04E+02	4.10E-01	7.31E+01
	MMLE	7.70E+00	9.60E+01	5.90E+00	4.91E+01	3.80E+00	2.08E+01
10%	MLE	8.64E+02	2.84E+08	7.49E+01	7.00E+05	1.24E+01	7.12E+04
	MMLE	7.86E+00	8.34E+01	7.61E+00	6.32E+01	7.52E+00	5.89E+01
20%	MLE	2.23E+03	3.37E+10	6.02E+02	1.65E+09	3.40E+02	7.72E+07
	MMLE	5.15E+00	3.92E+01	4.67E+00	2.50E+01	4.61E+00	2.27E+01
30%	MLE	3.31E+03	1.37E+11	7.74E+02	1.10E+10	5.12E+02	9.39E+08
	MMLE	2.80E+00	1.56E+01	2.48E+00	8.35E+00	2.45E+00	7.03E+00

Table 8: EB and MSE of the intercept: Non-Normal Symmetric-Case, Beta (3, 3)

Contamination	Methods	$n = 20$		$n = 50$		$n = 100$	
		EB	MSE	EB	MSE	EB	MSE
Without outlier	MLE	1.01E-02	2.03E-01	3.81E-03	7.33E-02	5.14E-04	3.51E-03
	MMLE	9.79E-01	4.30E+01	2.66E-01	9.94E+00	1.07E-01	3.45E+00
Single outlier	MLE	8.33E+01	1.20E+06	1.90E+01	3.94E+02	8.16E+00	7.10E+01
	MMLE	6.10E+00	6.78E+01	3.86E+00	2.52E+01	2.24E+00	9.07E+00
10%	MLE	3.83E+02	7.47E+08	1.39E+02	1.70E+05	1.21E+02	2.61E+04
	MMLE	6.90E+00	6.60E+01	6.92E+00	5.19E+01	6.89E+00	4.92E+01
20%	MLE	6.36E+02	1.36E+09	1.79E+02	3.19E+08	1.15E+02	1.08E+06
	MMLE	4.90E+00	3.47E+01	4.50E+00	2.32E+01	4.43E+00	2.09E+01
30%	MLE	7.23E+02	4.43E+09	5.69E+02	6.07E+08	3.36E+02	3.65E+06
	MMLE	2.71E+00	1.46E+01	2.40E+00	7.84E+00	2.35E+00	6.48E+00

Results presented in Tables 1 to 8 clearly show the advantage of using MMLE method. When the data are free from outliers, the EB and MSE of the proposed modified maximum likelihood estimator are similar to those of the maximum likelihood estimator. But we observe a marked difference when the data set gets contaminated. The EB and MSE of slope and intercept using MLE break down cheaply and become huge. It tends to deteriorate with the increase in the level of contamination. But the performance of the MMLE is solid here. The EB and MSE values are not affected by outliers irrespective of the sample sizes and percentages of contamination.

3. Numerical Examples

In this section we consider two real world data to investigate the performance of the MMLE method. In order to make the relationship as model 1, we assume that measurement error can occur in both the variables of these two examples.

3.1 Iron in slag data

At first we consider the Iron in Slag data taken from Hand et al. (1993). This data contains 50 results of crushed blast-furnace slag measured by two different techniques, which are chemical test (dependent variable) and magnetic test (independent variable). The original data do not contain any outlier. However, we purposely insert few outliers in this data set to create few different situations like single outlier, 10%, 20% and 30% outlier cases Mamun et al. (2012). We employ both the MLE and the MMLE to estimate the parameters of LSRM and the results are shown in Table 9.

Table 9: Estimated parameters using two different methods from Iron in Slag data

Contamination	Methods	Intercepts	Slopes
Without outlier	MLE	2.16	0.9303
	MMLE	1.5	1
Single outlier	MLE	-3.5142	1.2225
	MMLE	-1.107	1.1336
10%	MLE	-86.2355	5.4494
	MMLE	1.0911	1.0722
20%	MLE	-245.4938	13.5884
	MMLE	2.7288	1.0651
30%	MLE	2724.891	-132.4453
	MMLE	4.5106	1.0507

We observe from Table 9 that in absence of outlier, the performance of the MMLE is very similar to the MLE. But the MLE of slopes and intercepts start breaking down in the presence of outliers and gets bigger and bigger with the increase in the percentages of outliers, whereas the MMLE was not affected at all by the outliers.

3.2 Serum Kanamycin data

Our final example is the Serum Kanamycin data taken from Kelly (1984). In order to measure the serum kanamycin levels in blood samples a simultaneous pairs of measurements were taken from twenty premature babies. These two measurements were obtained by heelstick method (x) and umbilical catheter method (y). Likewise the previous example, the original data do not contain any outlier. We insert few outliers in this data set to create few different situations like single outlier, 10%, 20% and 30% outlier cases. The estimates of slope and intercept for this data are computed by the MLE and the MMLE under a variety of situations and the results are shown in Table 10.

Table 10: Estimated parameters using two different methods from Serum Kanamycin data

Contamination	Methods	Intercept	Slope
Without outlier	MLE	-1.16	1.0697
	MMLE	-0.388	0.9419
Single outlier	MLE	-23.5304	2.2023
	MMLE	-1.0586	0.9754
10%	MLE	-30.2891	2.5967
	MMLE	-1.9299	1.0189
20%	MLE	-84.5344	5.3948
	MMLE	-0.6488	1.0574
30%	MLE	-180.6768	10.2556
	MMLE	1.8972	1.0801

Table 10 shows the improvement of using our proposed MMLE method. For the original data set where there was no outlier, the modified maximum likelihood estimators are similar to the maximum likelihood estimators. But we observe a marked difference when outliers are present in the data. The estimated parameters using MLE break down completely with the increase in the percentages of outliers, whereas the MMLE's are not at all affected by the outliers.

4. Conclusion

In this paper, we proposed a modified maximum likelihood estimation method based on robust estimator to estimate parameters for the LSRM. Both the real-world examples and simulation results show that traditional MLE method gives marginally better result than our proposed method when no outlier is present in the data. However, the MLE method breaks down completely when the degree of contamination increases, while the modified maximum likelihood estimation method performs very well in every situation. Since our proposed method is able to produce satisfactory results even in the presence of a large amount of outliers, so the modified maximum likelihood estimator should be considered as a good alternative to the traditional maximum likelihood estimator.

Acknowledgement

The authors are thankful to the referees and the editor of the Journal for their very helpful comments and suggestions.

References

Alfaro, J. and Ortega, J. F. (2009). A comparison of robust alternatives to hotelling's T^2 control chart. *Journal of Applied Statistics*, 36(12):1385–1396.

Barnett, V. (1967). A note on linear structural relationships when both residual variances are known. *Biometrika*, 54:670–672.

Birch, M. (1964). A note on the maximum likelihood estimation of a linear structural relationship. *Journal of the American Statistical Association*, 59:1175–1178.

Bolfarine, H. and Cordani, L. K. (1993). Estimation of a structural linear regression model with a known reliability ratio. *Annals of the Institute of Statistical Mathematics*, 45(3):531–540.

Chan, L. K. and Mak, T. K. (1979). On the maximum likelihood estimation of a linear structural relationship when the intercept is known. *Journal of Multivariate Analysis*, 9(2):304–313.

Cheng, C. L. and Van Ness, J. W. (1994). On estimating linear relationships when both variables are subject to errors. *Journal of the Royal Statistical Society*, Series B (Methodological):167–183.

Fuller, W. (1987). *Measurement Error Models*. New York: John Wiley & Sons.

Hand, D. J., Daly, F., McConway, K., Lunn, D., and Ostrowski, E. (1993). *A Handbook of Small Data Sets*, volume 1. London, U.K.: Chapman and Hall.

Hood, K., Nix, B. A., and Iles, T. C. (1999). Asymptotic information and variance–covariance matrices for the linear structural model. *Journal of the Royal Statistical Society: Series D (The Statistician)*, 48(4):477–493.

Kelly, G. (1984). The influence function in the errors in variables problem. *The Annals of Statistics*, 12(1):87–100.

Kendall, M. G. and Stuart, A. (1973). *The Advanced Theory of Statistics*. London: Charles Griffin.

Koláček, J. (2008). Plug-in method for nonparametric regression. *Computational Statistics*, 23(1):63–78.

Lakshminarayanan, M. Y. and Gunst, R. F. (1984). Estimation of parameters in linear structural relationships: Sensitivity to the choice of the ratio of error variances. *Biometrika*, 71(3):569–573.

Lieberman, O. (2005). On plug-in estimation of long memory models. *Econometric Theory*, 21(2):431–454.

Liu, C. A. (2012). A plug-in averaging estimator for regressions with heteroskedastic errors (August 10, 2012). Available at SSRN: https://ssrn.com/abstract=2138013 or http://dx.doi.org/10.2139/ssrn.2138013.

Madansky, A. (1959). The fitting of straight lines when both variables are subject to error. *Journal of the American Statistical Association*, 54(285):173–205.

Mamun, A. S. M. A., Hussin, A. G., Zubairi, Y. Z., and Imon, A. H. M. R. (2012). A nonparametric robust estimator for slope of linear structural relationship model. *Pakistan Journal of Statistics*, 28:385–394.

Mamun, A. S. M. A., Zubairi, Y. Z., Hussin, A. G., Imon, A. H. M. R., Rana, S., and Carrasco, J. (2019). Identification of influential observation in linear structural relationship model with known slope. *Communications in Statistics - Simulation and Computation*. https://doi.org/10.1080/03610918.2019.1645172.

Moran, P. (1971). Estimating structural and functional relationships. *Journal of Multivariate Analysis*, 1(2):232–255.

Reilman, M. A., Gunst, R. F., and Lakshminarayanan, M. Y. (1985). Structural model estimation with correlated measurement errors. *Biometrika*, 72(3):669–672.

Rousseeuw, P. J. and Croux, C. (1993). Alternatives to the median absolute deviation. *Journal of the American Statistical Association*, 88(424):1273–1283.

Shevlyakov, G. and Smirnov, P. (2011). Robust estimation of the correlation coefficient: An attempt of survey. *Austrian Journal of Statistics*, 40(1&2):147–156.