

## On Robust Estimation for Slope in Linear Functional Relationship Model (Penggangan Teguh bagi Kecerunan dalam Model Linear Hubungan Fungsian)

AZURAINI MOHD ARIF, YONG ZULINA ZUBAIRI\* & ABDUL GHAPOR HUSSIN

### ABSTRACT

*In this paper, we propose a robust parameter estimation method for the linear functional relationship model. We improved the maximum likelihood estimation using robust estimators and robust correlation coefficients to estimate the slope parameter. The performance of the propose method, MMLE, is compared with the standard maximum likelihood estimation (MLE) and the nonparametric method in terms of mean square error. The results for simulation studies suggested the performance of the MMLE and nonparametric methods gives better estimate than the standard MLE in the presence of outliers. The novelty of the proposed method is that it is not affected by the presence of outliers and is simple to use. To illustrate practical application of the methods, we obtain the estimate of the slope parameter in a study of body-composition techniques for children.*

*Keywords: Linear functional relationship model; mean square error; modified maximum likelihood estimation; outliers; robust*

### ABSTRAK

*Dalam kertas ini, kami mencadangkan kaedah penganggaran parameter teguh bagi model linear hubungan fungsian. Kami menambah baik kaedah kebolehdajian maksimum menggunakan penganggar teguh dan pekali korelasi teguh bagi menganggarkan parameter kecerunan. Kuasa pretasi diukur bagi kaedah yang disyorkan iaitu MMLE, MLE dan kaedah tidak berparameter menggunakan ralat kuasa dua min. Keputusan simulasi menunjukkan prestasi bagi kaedah yang disyorkan, MMLE dan kaedah tidak berparameter adalah lebih teguh daripada kaedah kebolehdajian maksimum apabila terdapat data terpencil. Kepentingan kaedah yang dicadangkan adalah ia tidak terjejas dengan kehadiran data terpencil dan juga mudah digunakan. Penggunaan kesemua kaedah yang dicadangkan ditunjukkan melalui data set sebenar dengan kaedah untuk menganggarkan kecerunan model bagi data komposisi badan untuk kanak-kanak.*

*Kata kunci: Kebolehdajian maksimum yang diubah suai; min ralat kuasa dua; model linear hubungan fungsian; teguh; terpencil*

### INTRODUCTION

Errors-in-variable model (EIVM) or measurement error model was first introduced in the 19th century by Adcock, R.J. Since then, many authors have worked on estimating the parameter of EIVM (Fuller 1987; Kendall & Stuart 1979; Lindley 1947). Suppose the variables  $X$  and  $Y$  are related by  $Y = \alpha + \beta X$ . If both  $X$  and  $Y$  are observed correctly, there is no statistical problem in obtaining values of  $\alpha$  and  $\beta$ . If  $Y$  only is observed with error, then regression model is formulated. However, when both  $X$  and  $Y$  are subject to error, the errors-in-variable model is applied. In real situations, measurement errors arise when both the variables involved cannot be recorded exactly (Gençay & Gradojevic 2011; Ghapor et al. 2017; Patefield 1985). Ignorance of measurement errors directly affects the desirable criteria of an estimator in which in this case, EIVM is more applicable rather than regression model.

In this study, we focus on the linear functional relationship model (LFRM) which is one of the branch in errors-in-variable model. It is categorized as functional relationship model for  $X$  and  $Y$ , when  $X$  is a mathematical

variable (Kendall 1951; Lindley 1947; Moran 1971). Linear functional relationship model (LFRM) can be expressed by,

$$Y_i = \alpha + \beta X_i, \quad \text{for } i = 1, 2, 3, \dots, n \quad (1)$$

where both the variables  $X$  and  $Y$  are linearly related but observed with error, with  $\alpha$  is the intercept, and  $\beta$  is the slope parameters. For any fixed  $X_i$ , we observe  $x_i$  and  $y_i$  from continuous linear variable subject to errors  $\delta_i$  and  $\varepsilon_i$ , respectively, i.e.

$$x_i = X_i + \delta_i \text{ and } y_i = Y_i + \varepsilon_i \quad (2)$$

where the error terms  $\delta_i$  and  $\varepsilon_i$  are assumed to be mutually independent and normally distributed random variables, i.e.

$$\delta_i \sim N(0, \sigma_\delta^2) \text{ and } \varepsilon_i \sim N(0, \sigma_\varepsilon^2) \quad (3)$$

In LFRM, there are  $(n + 4)$  parameters that need to be estimated, namely  $\alpha$ ,  $\beta$ , the two error variances and the

incidental parameters  $X_1, X_2, \dots, X_n$ , respectively. The log likelihood function is given by,

$$\begin{aligned} &\log L(\alpha, \beta, \sigma_\delta^2, \sigma_\epsilon^2, X_1, \dots, X_n; x_1, \dots, x_n, y_1, \dots, y_n) = \\ &- n \log (2 \pi) - \frac{n}{2} (\log \sigma_\delta^2 + \log \sigma_\epsilon^2) - \\ &\frac{\sum (x_i - X_i)^2}{2\sigma_\delta^2} - \frac{\sum (y_i - \alpha - \beta X_i)^2}{2\sigma_\epsilon^2} \end{aligned} \tag{4}$$

However, the estimation will lead to inconsistencies with the existence of this incidental parameters and an assumption must be made in order to avoid this problem which is the ratio of the two variances is known,  $\sigma_\epsilon^2 = \lambda \sigma_\delta^2$  (Abdullah 1989; Moran 1971; Solari 1969). In this case, the log likelihood function can be expressed as,

$$\begin{aligned} &\log L(\alpha, \beta, \sigma_\delta^2, X_1, \dots, X_n; \lambda, x_1, \dots, x_n, y_1, \dots, y_n) = \\ &- n \log (2 \pi) - \frac{n}{2} \log \lambda - \frac{1}{2\sigma_\delta^2} \left\{ \sum (x_i - x_i)^2 + \frac{\sum (y_i - \alpha - \beta X_i)^2}{\lambda} \right\} \end{aligned} \tag{5}$$

Numerous methods of estimation of linear functional relationship model have been suggested using normality assumption namely Fuller (1987), Kendall and Stuart (1979) and Moran (1971). However, according to Al-Nasser and Ebrahim (2005) and Ghapor et al. (2015), when data contain outliers, the normality assumption is invalid. To circumvent this problem, some methods such as nonparametric methods or robust method have been proposed where normality assumption can be ignored and can diminish the effect of the outliers in the data.

In this paper, we propose a new parameter estimation method based on the robust estimator and robust coefficient correlation in estimating the slope parameter. This paper is organized as follows: Next section describes the maximum likelihood estimation including the nonparametric method (Ghapor et al. 2015) and the proposed modified maximum likelihood method. This is followed by subsequent section where a simulation study is conducted to compare existing methods of maximum likelihood estimation and nonparametric method (Ghapor et al. (2015) with the proposed method (MMLE). The results and discussion are given in the section that follows. A practical example is highlighted using published data set next. Lastly, conclusion is presented in the last section.

MAXIMUM LIKELIHOOD ESTIMATION METHOD (MLE)

Maximum likelihood estimation method (MLE) is the common method used in LFRM. Based on the assumption when the ratio of error variances is known,  $\sigma_\epsilon^2 = \lambda \sigma_\delta^2$ , there are  $(n+3)$  parameters to be estimated which are  $\alpha, \beta, \sigma_\delta^2$  and  $X_1, \dots, X_n$  (Fuller 1987; Kendall & Stuart 1979). The parameters may be obtained by differentiating the log

likelihood function as given in equation (5) with respect to  $\hat{\alpha}, \hat{\beta}, \hat{\sigma}_\delta^2$  and  $\hat{X}_i$ , respectively and equating to zero. Thus, we can obtain the parameters given by,

$$\hat{\beta} = \frac{S_{yy} - \lambda S_{xx} + \sqrt{(S_{yy} - \lambda S_{xx})^2 + 4\lambda S_{xy}^2}}{2S_{xy}} \tag{6}$$

$$\hat{\alpha} = \bar{y} - \hat{\beta} \bar{x},$$

$$\hat{\sigma}_\delta^2 = \frac{1}{(n-2)} \left\{ \sum (x_i - \hat{X}_i)^2 + \frac{1}{\lambda} \sum (y_i - \hat{\alpha} - \hat{\beta} \hat{X}_i)^2 \right\}$$

$$\hat{X}_i = \frac{\lambda x_i + \hat{\beta} (y_i - \hat{\alpha})}{\lambda + \hat{\beta}^2}$$

where

$$\bar{x} = \frac{1}{n} \sum x_i, \quad \bar{y} = \frac{1}{n} \sum y_i$$

$$S_x^2 = \sum (x_i - \bar{x})^2, \quad S_y^2 = \sum (y_i - \bar{y})^2$$

and

$$S_{xy} = \sum (x_i - \bar{x})(y_i - \bar{y})$$

However, as mentioned before, in the presence of the outliers, the value of the parameters using maximum likelihood estimation may be affected (Abdullah 1989).

NONPARAMETRIC METHOD

The nonparametric method as proposed by Ghapor et al. (2015) uses median to obtain the estimated slope value,  $\hat{\beta}_G$ . As mentioned earlier, in this method, the normality assumption can be ignored. The steps in estimating  $\hat{\beta}_G$  are as follows (Ghapor et al. 2015):

Step 1

The observations are first arranged in ascending order, based on  $x$  value namely

$$x_{(1)} \leq x_{(2)} \leq \dots \leq x_{(n)}$$

The associated values of  $y$  which may not be in ascending order are taken namely,

$$y_{[1]} \leq y_{[2]} \leq \dots \leq y_{[n]}$$

The new pairs will be  $(x_{(i)}, y_{[i]})$

Step 2

All the data are divided into  $m$ -subsamples. These subsamples contains  $r$  elements such that  $m * r = n$  where  $m$  is the maximum divisor of  $n$ , such that  $m \leq r$ .

Step 3

Find all the possible slopes.

$$\left\{ b_x(k)_{ij} = \frac{y_{(j)} - y_{(i)}}{x_{(j)} - x_{(i)}}; i = 1, 2, \dots, j-1; j = 2, 3, \dots, r \right\}; k = 1, 2, \dots, m$$

Step 4

Repeat Steps 1 to 3 by interchanging  $y$  and  $x$  to get possible paired of  $b_y(k)_{ij}$

$$\left\{ b_y(k)_{ij} = \frac{y_{(j)} - y_{(i)}}{x_{(j)} - x_{(i)}}; i = 1, 2, \dots, j-1; j = 2, 3, \dots, r \right\}; k = 1, 2, \dots, m$$

Step 5

Find the median of all slopes.

$$\hat{\beta}_G = \text{median} \{ b_x(k)_{ij}, b_y(k)_{ij} \}$$

In this method, only the slope parameter  $\hat{\beta}_G$  will be estimated. Other parameters will be estimate using the traditional method, MLE.

PROPOSED METHOD (MODIFIED MAXIMUM LIKELIHOOD ESTIMATION (MMLE))

In this section, a modification of maximum likelihood estimation method is proposed to overcome the presence of outlier. As mentioned earlier, some standard statistics such as mean, variance, covariance in the maximum likelihood estimation in (6) are sensitive to the outliers. To overcome the presence of outliers, we introduce a robust estimator  $Q_n$  as proposed by Rousseeuw and Croux (1993) in the formulation of the MLE.

To construct the modified maximum likelihood, we replace the sample variances  $S_x^2$  and  $S_y^2$  as given in (6) with a robust estimator  $\{Q_n(x)\}^2 = S_x^{*2}$  and  $\{Q_n(y)\}^2 = S_y^{*2}$  respectively, where  $Q_n(x) = 1.0483 \{ |x_i - x_j|; i < j \}_{(k)}$  and  $Q_n(y) = 1.0483 \{ |y_i - y_j|; i < j \}_{(k)}$  where 1.0483 is a constant factor chosen to provide consistency of estimation of the standard deviation of a normal distribution where  $k = \binom{h}{2} \approx \binom{n}{2} / 4$  and  $h = (n / 2) + 1$  is a roughly half the number of observations (Rousseeuw & Croux 1993).

This means the sample covariance  $S_{xy}$  is replaced by  $S_{xy}^*$  which,  $S_{xy}^* = r_{Qn} \times S_x^* \times S_y^*$  where  $r_{Qn}$  is the robust correlation coefficient proposed by Shevlyakov and Smirnov (2011) and defined as,

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

where  $u$  and  $v$  are the robust principle variables defined by,

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

and

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

Now, we have the new slope parameter  $\hat{\beta}_{MMLE}$  and replace the estimation in (6) to obtain the modified maximum likelihood estimator given as:

$$\hat{\beta}_{MMLE} = \frac{S_y^{*2} - \lambda S_x^{*2} + \sqrt{(S_y^{*2} - \lambda S_x^{*2})^2 + 4\lambda S_{xy}^{*2}}}{2S_{xy}^*} \tag{7}$$

SIMULATION STUDY

A simulation study was carried out using R software in order to evaluate the performance of the proposed method, MMLE, with the existing method, MLE and the nonparametric method (Ghapor et al. (2015)), in the presence of the outliers. The observations are then simulated using our model,

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

where  $X_i = 10 \frac{i}{n}$  and  $\delta_i, \varepsilon_i \sim N(0, 0.1)$ .

Without loss of generality, the slope and intercept parameters are fixed at  $\alpha = 1$  and  $\beta = 1$ . We also consider when the observation has no outlier, single outlier and certain percentages of outliers namely 10% and 20% outliers, respectively. Here, we contaminate data points as suggested by Al-Nasser and Ebrahim (2005) and Ghapor et al. (2015) using this relationship,  $y_c = 1 + X_c + \varepsilon_c$  with  $\varepsilon_c \sim N(0, 25)$ . Using 10000 trials, the performance of these three methods is measured based on mean square error (MSE) given by  $MSE = \frac{1}{s} (\hat{\beta}_j - \beta)^2$  where  $\beta$  is the slope parameter and  $s$  is the number of trials. In each trials, a sample size of 20, 50 and 100 are generated using relationship described earlier. Additionally, the errors term  $\delta_i$  and  $\varepsilon_i$  are generated from three non-normal distribution namely, Beta (2,9) for right-skewed case, Beta (9,2) for left-skewed case and Beta (3,3) for non-normal symmetric case in order to investigate the robustness of the proposed method. Simulation results are presented in Tables 1-4.

RESULTS AND DISCUSSION

For the simulation results in Table 1, where the errors  $\delta_i$  and  $\varepsilon_i$  are normally distributed, as expected there is no much difference among the three methods in estimating the slope when the data have no outlier as the mean square error (MSE) of the proposed method, MMLE, the nonparametric method (Ghapor et al. (2015)) and the traditional method, MLE is somewhat similar to each other. However, when a single outlier is present, the MLE method starts to break down and has higher MSE value for the slope parameter

compared to the MMLE method and the nonparametric method. Furthermore, the MSE value are not much affected by 10% and 20% outliers using our proposed method, MMLE, and nonparametric method.

Next, from Table 2, where the errors term  $\delta_j$  and  $\varepsilon_j$  are skewed to the right with Beta (2, 9), when the data have

no outliers, the MSE value for three methods are somewhat similar for each other. When the data gets contaminated, from single outlier to 10% and 20%, respectively, the value of MSE for MMLE method and the nonparametric method are much smaller compared to the MLE method. The value of MSE for MLE method become huge as the outlier

TABLE 1. MSE of the slope: Normal-case (0,0.1)

Contamination	Method	$n = 20$	$n = 50$	$n = 100$
No outlier	MLE	1.179E-05	4.614E-05	2.442E-05
	MMLE	1.362E-03	2.071E-04	7.479E-05
	NONPARAMETRIC	1.546E-04	5.567E-05	2.767E-05
Single outlier	MLE	4.436E+01	6.474E-01	8.722E-02
	MMLE	5.607E-03	7.287E-04	2.073E-04
	NONPARAMETRIC	2.241E-04	6.697E-05	3.007E-05
10%	MLE	1.581E+02	1.598E+02	1.601E+02
	MMLE	1.169E-02	8.499E-03	7.643E-03
	NONPARAMETRIC	4.864E-04	4.458E-04	4.067E-04
20%	MLE	3.996E+01	4.006E+01	4.007E+01
	MMLE	1.930E-02	1.305E-02	1.087E-02
	NONPARAMETRIC	4.356E-03	3.349E-03	3.268E-03

TABLE 2. MSE of the slope: right skewed-case: Beta (2,9)

Contamination	Method	$n = 20$	$n = 50$	$n = 100$
No outlier	MLE	1.513E-04	6.083E-05	3.025E-05
	MMLE	1.441E-03	2.391E-04	8.523E-05
	NONPARAMETRIC	1.906E-04	6.648E-05	3.214E-05
Single outlier	MLE	4.452E+01	6.477E-01	8.728E-02
	MMLE	5.701E-03	7.682E-04	2.193E-04
	NONPARAMETRIC	2.769E-04	7.999E-05	3.497E-05
10%	MLE	1.596E+02	1.605E+02	1.604E+02
	MMLE	1.204E-02	8.465E-03	7.689E-03
	NONPARAMETRIC	6.009E-04	5.273E-04	4.731E-04
20%	MLE	4.002E+01	4.008E+01	4.008E+01
	MMLE	1.979E-02	1.307E-02	1.092E-02
	NONPARAMETRIC	5.278E-03	4.016E-03	3.890E-03

TABLE 3. MSE of the slope: left skewed-case: Beta (9,2)

Contamination	Method	$n = 20$	$n = 50$	$n = 100$
No outlier	MLE	1.505E-04	5.940E-05	2.996E-05
	MMLE	1.444E-03	2.408E-04	8.287E-05
	NONPARAMETRIC	1.909E-04	6.594E-05	3.178E-05
Single outlier	MLE	4.452E+01	6.480E-01	8.730E-02
	MMLE	5.780E-03	7.634E-04	2.166E-04
	NONPARAMETRIC	2.758E-04	7.921E-05	3.505E-05
10%	MLE	1.597E+02	1.605E+02	1.603E+02
	NONPARAMETRIC	1.213E-02	8.432E-03	7.650E-03
	NONPARAMETRIC	5.931E-04	5.240E-04	4.799E-04
20%	MLE	3.997E+01	4.007E+01	4.008E+01
	MMLE	1.995E-02	1.299E-02	1.092E-02
	NONPARAMETRIC	5.284E-03	4.006E-03	3.909E-03

TABLE 4. MSE of the slope: Non-normal symmetric-case: Beta(3,3)

Contamination	Method	$n = 20$	$n = 50$	$n = 100$
No outlier	MLE	4.321E-04	1.706E-04	8.419E-05
	MMLE	1.823E-03	4.950E-04	1.938E-04
	NONPARAMETRIC	5.613E-04	2.028E-04	1.011E-04
Single outlier	MLE	4.583E+01	6.501E-01	8.733E-02
	MMLE	5.278E-03	1.065E-03	3.363E-04
	NONPARAMETRIC	8.197E-04	2.440E-04	1.124E-04
10%	MLE	1.629E+02	1.618E+02	1.611E+02
	MMLE	1.273E-02	9.024E-03	7.816E-03
	NONPARAMETRIC	1.769E-03	1.588E-03	1.483E-03
20%	MLE	4.004E+01	4.009E+01	4.009E+01
	MMLE	2.409E-02	1.427E-02	1.168E-02
	NONPARAMETRIC	1.487E-02	1.149E-02	1.132E-02

increases. This suggest the superiority of both MMLE and nonparametric methods.

From Table 3, where the errors term  $\delta_i$  and  $\varepsilon_i$  are skewed to the left with Beta (9, 2), the MSE value gives similar conclusion in which all three methods perform well. As data gets contaminated, MLE method fails to perform while MMLE and nonparametric methods remain unaffected. The same can be said for the case when error terms  $\delta_i$  and  $\varepsilon_i$  are non-normal symmetric with Beta (3,3) distribution in Table 4.

In summary, in all cases, the MMLE and nonparametric methods are superior than the traditional MLE in estimating the slope parameter  $\beta$ , when there is a presence of outliers. This implies that both MMLE and nonparametric estimator are both robust to outliers. In comparison to the two superior methods namely MMLE and nonparametric methods, each has got its own merits and limitation. For nonparametric method for parameter estimation, it does not require any distributional assumption but it may lack of power and less efficient when the underlying populations are normal compared with traditional methods (Kendall & Stuart 1979). Also, the steps involved in getting the parameter estimate can be quite cumbersome.

The proposed MMLE method is very robust to outliers and provide efficient estimates of the parameters as the standard estimation such as mean and variance which are sensitive to outliers (Hampel et al. 1986). In our proposed model, a simple modification to the covariance and consequently to the slope estimate has made the estimator robust to outliers.

In short, both methods provide viable alternatives when data are contaminated, when sample size is small or when the sampling distribution cannot be derived analytically.

#### PRACTICAL EXAMPLE

To illustrate the practicality of the method, we used real life data that can be modelled using linear functional relationship model. The data is obtained from a study that measures the accuracy of some widely used body-composition techniques

for children between the ages 4 and 10 years by two different techniques, namely skinfold thickness (ST) and bioelectrical resistance (BR) (Goran et al. 1996). As measurement error can occur in both variables for this experiment, we note that we can describe the relationship by LFRM as given in (1). Here, we assume that the error terms follow a normal distribution. The data consists of 97 observations. Nevertheless, in the nonparametric method particularly in step 2, the observations cannot be divided into  $m$ -subsamples as the  $n = 97$  is a prime number. Thus, in this case, we choose  $m = 1$  and proceed to step 5. In examining the slope effect by these three different methods, some original  $y$  values were replaced by the values of the outliers namely, single outlier, 10% and 20% outliers to create different situations by following Imon & Hadi (2008) and Kim (2000). The estimated slopes and standard deviations by these three different methods were shown in Table 5.

From Table 5, it can be seen that both the proposed method, MMLE and nonparametric are more robust than the MLE method when outliers are present in the data. The value of slope parameter of MLE, MMLE and nonparametric methods are quite similar when the data has no outlier. However, the slope parameter using MLE method starts to break down and change significantly when the percentage of data contaminated increased from single outlier to 10% and 20% outliers compared with the proposed method, MMLE, and the nonparametric method.

#### CONCLUSION

In this paper, we propose a robust method namely modified maximum likelihood estimation (MMLE) method in estimating the slope parameter for linear functional relationship model. The simulation studies suggest when there is an outlier or multiple outliers exists, both the MMLE method and the nonparametric method are robust to outliers unlike the MLE method. However, the nonparametric method has a limitation when the sample size is a prime number, the steps cannot be applied wholly. The proposed MMLE method, on the other hand is simple as it only requires some modification to the covariance estimate.

Additionally, we illustrate the relevance of the method using real data set. In summary, the proposed robust MMLE is a good when estimating the slope parameter of the linear functional relationship model.

#### ACKNOWLEDGEMENTS

We are most grateful to Universiti Malaya for the financial assistance (PG128-2015B, BKS010-2016 & GPF006H-2018) and Ministry of Higher Education (MOHE), Malaysia for the financial support. We also wish to thank to referee for their helpful comments and suggestions.

#### REFERENCES

- Abdullah, M.B. 1989. On robust alternatives to the maximum likelihood estimators of a linear functional relationship. *Pertanika* 12(1): 89-98.
- Al-Nasser, A.D. & Ebrahem, M.A.H. 2005. A new nonparametric method for estimating the slope of simple linear measurement model in the presence of outliers. *Pak. J. Statist.* 21(3): 265-274.
- Fuller, W.A. 1987. *Measurement Error Models*. New York: John Wiley & Sons.
- Gençay, R. & Gradojevic, N. 2011. Errors-in-variables estimation with wavelets. *Journal of Statistical Computation and Simulation* 81(11): 1545-1564.
- Ghapor, A.A., Zubairi, Y.Z. & Imon, A.H.M.R. 2017. Missing value estimation methods for data in linear functional relationship model. *Sains Malaysiana* 46(2): 317-326.
- Ghapor, A.A., Zubairi, Y.Z., Mamun, A.S.M. & Imon, A.H.M.R. 2015. A robust nonparametric slope estimation in linear functional relationship model. *Pak. J. Statist.* 31(3): 339-350.
- Goran, M., Driscoll, P., Johnson, R., Nagy, T. & Hunter, G. 1996. Cross-calibration of body-composition techniques against dual-energy X- ray absorptiometry in young children. *Am. J. Clin. Nutr.* 63(3): 299-305.
- Hampel, F.R., Ronchetti, E.M., Rousseeuw, P.J. & Stahel, W.A. 1986. *Robust Statistics: The Approach Based on Influence Functions*. New York: John Wiley & Sons.
- Imon, A.H.M.R. & Hadi, A.S. 2008. Identification of multiple outliers in logistic regression. *Communications in Statistics-Theory and Methods* 37(11): 1697-1709.
- Kendall, M.G. 1951. Regression, structure and functional relationship. Part I. *Biometrika* 38(1/2): 11-25.
- Kendall, M.G. & Stuart, A. 1979. *The Advanced Theory of Statistics*. Vol. 2. London: Griffin.
- Kim, M.G. 2000. Outliers and influential observations in the structural errors-in-variables model. *Journal of Applied Statistics* 27(4): 451-460.
- Lindley, D.V. 1947. Regression lines and the linear functional relationship. *Supplement to the Journal of the Royal Statistical Society* 9(2): 218-244.
- Moran, P.A.P. 1971. Estimating structural and functional relationships. *Journal of Multivariate Analysis* 1(2): 232-255.
- Patefield, W.M. 1985. Information from the maximized likelihood function. *Biometrika* 72(3): 664-668.
- Rousseeuw, P.J. & Croux, C. 1993. Alternatives to the median absolute deviation. *Journal of the American Statistical Association* 88(424): 1273-1283.
- Shevlyakov, G. & Smirnov, P. 2011. Robust estimation of the correlation coefficient: An attempt of survey. *Austrian Journal of Statistics* 40(1): 147-156.
- Solari, M.E. 1969. The "maximum likelihood solution" of the problem of estimating a linear functional relationship. *Journal of the Royal Statistical Society: Series B (Methodological)* 31(2): 372-375.

Azuraini Mohd Arif  
Institute of Graduate Studies  
Universiti Malaya  
50603 Kuala Lumpur, Federal Territory  
Malaysia

Yong Zulina Zubairi\*  
Centre for Foundation Studies in Science  
Universiti Malaya  
50603 Kuala Lumpur, Federal Territory  
Malaysia

Abdul Ghapor Hussin  
National Defense University Malaysia  
Sungai Besi Camp  
57000 Kuala Lumpur, Federal Territory  
Malaysia

\*Corresponding author; email: yzulina@um.edu.my

Received: 29 August 2017

Accepted: 3 August 2018