An Improved Simple Averaging Approach for Estimating Parameters in Simple Linear Regression Model

Jetsada Singthongchai¹, Noppakun Thongmual²*, Nirun Nitisuk²

¹Department of Informatics and Computer Science, Faculty of Sciences and Health Technology, Kalasin University, 46000, Thailand
²Department of Science and Mathematics, Faculty of Sciences and Health Technology, Kalasin University, 46000, Thailand

Received July 15, 2021; Revised October 3, 2021; Accepted November 11, 2021

Cite This Paper in the following Citation Styles

(a): [1] Jetsada Singthongchai, Noppakun Thongmual, Nirun Nitisuk, "An Improved Simple Averaging Approach for Estimating Parameters in Simple Linear Regression Model," Mathematics and Statistics, Vol. 9, No. 6, pp. 939 - 946, 2021. DOI: 10.13189/ms.2021.090610.

(b): Jetsada Singthongchai, Noppakun Thongmual, Nirun Nitisuk (2021). An Improved Simple Averaging Approach for Estimating Parameters in Simple Linear Regression Model. Mathematics and Statistics, 9(6), 939 - 946. DOI: 10.13189/ms.2021.090610.

Copyright©2021 by authors, all rights reserved. Authors agree that this article remains permanently open access under the terms of the Creative Commons Attribution License 4.0 International License

Abstract This research is about estimating parameters in simple linear regression model. Regression model is applied for predictive in many filed. Ordinary least square (OLS) approach and Maximum likelihood (ML) approach are employed for estimating parameter in simple linear regression model when the assumption is not violated. This research interested in simple linear regression model when the assumption is violated. Simple Averaging (SA) approach is an alternative for estimating parameters in simple linear regression model where assumptions are not successfully used. We improved SA approach based on the median which is called the improved Simple Averaging (ISA) approach. For comparing the two approaches for estimating parameter in simple linear regression model, ISA approach is compared with SA approach under Root Mean Square Error (RMSE) which reflected accuracy of prediction in simple linear regression. By using the sample, the results showed that ISA approach is better than SA approach where the value of RMSE of ISA approach is less than the value of RMSE of SA approach. Therefore, ISA approach is better than SA approach. Our study suggests ISA approach to estimating parameter on simple linear regression because ISA approach accuracy than SA approach and ISA approach simplify the estimation of parameters in the simple linear regression model. Hence, ISA approach an alternative for estimating parameters in simple linear regression model when the assumptions are not successfully used.

Keywords Simple Averaging, Simple Linear Regression Model, Improved Simple Averaging, Ordinary Least Square, Root Mean Square Error

1. Introduction

Simple linear regression model is a statistic for prediction which is used in many filed such as sciences, engineering, agricultural and education etc. Simple linear regression model describes the relationship between two continuous variables by fitting a line to the observed data. Simple linear regression model is

\[ Y = \alpha + \beta X + \epsilon, \]  

(1)

where \( Y \) is dependent variable and \( X \) is independent variable, \( \alpha \) and \( \beta \) are parameters in simple linear regression model, and \( \epsilon \) is error term under assumptions as normal distribution with \( E(\epsilon) = 0 \) and \( \text{Var}(\epsilon) = \sigma^2 \). Hence, the assumptions in simple linear regression model are about to error term.

In general, estimating parameter in simple linear regression is OLS approach and ML approach based on
assumption on simple linear regression model. Sometimes, the assumption is not examined. OLS approach and ML approach require many assumptions to complete. The assumptions are frequently found that there is an unconditional assumption that both methods are not successfully used. Hence, OLS approach and ML approach are not appropriated for using estimating parameter in simple linear regression model under unconditional assumption. OLS approach and ML approach are used in many works for estimating parameters $\alpha$ and $\beta$ in simple linear regression model such as Chioma [2], Cliff et al. [3], Prabitha et al. [12], Jain et al. [6], Gustiarisanie et al. [4], Permai et al. [11] and Singh et al. [15] used OLS approach and ML approach for estimating parameters $\alpha$ and $\beta$ in simple linear regression model.

In this research is not require assumption in simple linear regression model. Sometimes, the assumption may be violated. An alternative for estimating parameter on simple linear regression model proposing by Agung et al. [1] presents an approach to estimating parameters in simple linear regression model under unconditional assumption which is simple averaging (SA) approach. SA approach is a process for estimating parameter supporting by Cliff et al. [3] where have developed a new method in solving the simple and multiple linear regression equation. Estimators for parameters in simple linear regression model by SA approach are

$$\bar{\gamma}_{SA} = \frac{1}{n-1} \sum_{i=2}^{n} b_i = \frac{1}{n-1} \sum_{i=2}^{n} y_i - y_{i-1}$$

(2)

$$a_{SA} = \frac{1}{n-1} \sum_{i=2}^{n} \left( y_i - \frac{1}{i-2} \sum_{j=2}^{i-1} b_j \right).$$

(3)

where, $a_{SA}$ and $\bar{\gamma}_{SA}$ are the estimators for parameters $\alpha$ and $\beta$, respectively. SA approach is an interested process for estimating parameters in simple linear regression model under unconditional assumption since it simplifies for estimating parameter. SA approach in simple regression model for estimating parameter can be written as

$$Y = a_{SA} + \bar{\gamma}_{SA} X.$$  

(4)

In this research interests SA approach for estimating parameter in simple linear regression model under unconditional assumptions. Where SA approach used the average of the slop $b_i$ for $i = 2, 3, \ldots, n$ which are $\bar{\gamma}_{SA}$ for parameters $\beta$. The median is an alternative for estimating parameter $\beta$ where the median is used in this research for improving estimator under unconditional assumption. The median is the value separating the higher half from the lower half of a data sample, and it is middle point of a data sample. Furthermore, the median is used for nonparametric when the assumptions are violated. Many researchers employed median for inference statistics such as Nick [9] compared the nonparametric Wilcoxon–Mann–Whitney and Robust Rank-Order Tests when the median is used, Przemysław [13] used the k-sample median test for vague data, Qing [14] study explores issues related to one-sample nonparametric tests for the median of a continuous distribution when the sample is collected via size-bias of a known order, Mohammad et al. [8] propose a nonparametric shrinkage estimator for the median survival times from several independent samples of right-censored data, Hyo [5] used median test procedure to multivariate in censored data, Mohammad [7] proposed and investigated the performance of a new non-parametric test procedure for the median of a non-normal population. Therefore, we estimated parameter $\alpha$ and $\beta$ in simple linear regression under unconditional assumption based on median which is an alternative approach for estimation parameter under unconditional assumption.

The aim of this research is improving SA approach based on median which is called the improved simple averaging (ISA) approach, and compared SA approach and ISA approach based on root mean square error (RMSE) which reflected accuracy of estimator in simple linear regression model under unconditional assumption. The paper is organized as follows. Materials and methods are described in Section 2. Section 3 shows results of comparison of the performance of SA approach and ISA approach based on RMSE. Section 4 contains discussion. Finally, section 5 contains conclusion.

### 2. Materials and Methods for SA Approach and ISA Approach

Comparing SA approach and ISA approach based on RMSE, we used two data samples for examining accuracy of predictive of the two approaches as SA approach and ISA approach, and the two data samples are shown in Table 1 and Table 2. The process of SA approach and ISA approach is shown in 2.2 and 2.3, respectively. The equation for calculating RMSE of SA approach and ISA approach under unconditional assumption is

$$RMSE = \sqrt{\frac{\sum_{i=1}^{n} (y_i - \hat{y}_i)^2}{n}} = \sqrt{\frac{\sum_{i=1}^{n} e_i^2}{n}}.$$  

(5)

Where $e_i$ for $i = 1, 2, 3, \ldots, n$ are the error team for predictive in simple linear regression. RMSE in (5) is one of indicator for accuracy of predictive in simple linear regression for each approach.
Table 1. Data pairs by Cliff et al. [3] for examining RMSE under unconditional assumption of SA approach and ISA approach

| i | $x_i$ | $y_i$ |
|---|---|---|
| 1 | 6 | 526 |
| 2 | 3 | 421 |
| 3 | 6 | 581 |
| 4 | 9 | 630 |
| 5 | 3 | 412 |
| 6 | 9 | 560 |
| 7 | 6 | 434 |
| 8 | 3 | 443 |
| 9 | 9 | 590 |
| 10 | 6 | 570 |
| 11 | 3 | 346 |
| 12 | 9 | 672 |

Table 2. Data pairs by Nwabueze [10] for examining RMSE under unconditional assumption of SA approach and ISA approach

| i | $x_i$ | $y_i$ |
|---|---|---|
| 1 | 911,100 | 6,511,330 |
| 2 | 1,977,700 | 6,785,900 |
| 3 | 2,823,900 | 8,493,320 |
| 4 | 29,395,000 | 8,931,120 |
| 5 | 2,837,200 | 8,558,990 |
| 6 | 83,400,600 | 36,025,620 |
| 7 | 97,375,000 | 44,485,000 |
| 8 | 93,249,000 | 50,394,000 |
| 9 | 108,450,000 | 56,394,000 |
| 10 | 115,650,000 | 62,708,000 |
| 11 | 134,000,000 | 68,206,000 |
| 12 | 182,210,000 | 81,994,500 |
| 13 | 171,510,000 | 76,195,000 |
| 14 | 259,050,000 | 95,966,250 |

2.1. Estimating Parameter on Simple Linear Regression with SA Approach under Unconditional Assumption

Estimating parameter in simple linear regression model with SA approach is as follows.

1. Calculate slope $b_i = \frac{y_i - y_{i-1}}{x_i - x_{i-1}}$ for $i = 2, 3, \ldots, n$.

2. Fine the average of the slope $b_i$ which is

$$\overline{b}_i = \frac{1}{n-1} \sum_{i=2}^{n} b_i.$$  

3. Write $\overline{b}_i \cdot x_i$ for $i = 2, 3, \ldots, n$.

4. Denote $A = \frac{1}{n-1} \sum_{i=2}^{n} \overline{b}_i \cdot x_i$.

5. Calculate $y_i - A$ for $i = 2, 3, \ldots, n$.

6. Fine the average of $y_i - A$ for $i = 2, 3, \ldots, n$ which are

$$a_s = \frac{1}{n-1} \sum_{i=2}^{n} (y_i - A).$$

The process with SA approach for estimating parameters $\alpha$ and $\beta$ in simple linear regression under unconditional assumption is $\overline{b}_i$ and $a_s$, respectively. Therefore, simple linear regression with SA approach under unconditional assumption is

$$Y = a_s + \overline{b}_i X$$

The equation in (8) is simple linear regression with SA approach for prediction under unconditional assumption.

2.2. Estimating Parameter in Simple Linear Regression with ISA Approach under Unconditional Assumption

Estimating parameter in simple linear regression model with ISA approach is as follows.

1. Calculate slope $b_i = \frac{y_i - y_{i-1}}{x_i - x_{i-1}}$ for $i = 2, 3, \ldots, n$.

2. Fine the median of the slope $b_i$ for $i = 2, 3, \ldots, n$.

Where the slope $b_i$ for $i = 2, 3, \ldots, n$ are listed in order from smallest to greatest. Then considering the slope $b_i$, if the slope $b_i$ for $i = 2, 3, \ldots, n$ has an odd number of observations then the median can be defined as $\overline{b}_{ISA} = b_i \left( \frac{1}{2} \right)$ for $i = 2, 3, \ldots, n$, and if an even number of observations then the median can be defined as $\overline{b}_{ISA} = \frac{b_i + b_{i-1}}{2}$ for $i = 2, 3, \ldots, n$. Therefore, $\overline{b}_{ISA}$ is the estimator of parameter $\beta$ in simple linear regression under unconditional assumption.

7. Calculate $\overline{b}_{ISA} \cdot x_i$ for $i = 2, 3, \ldots, n$.

8. Denote $A_{ISA} = \frac{1}{n-1} \sum_{i=2}^{n} \overline{b}_{ISA} \cdot x_i$.

9. Calculate $y_i - A_{ISA}$ for $i = 2, 3, \ldots, n$.

10. Compute the average of $y_i - A_{ISA}$ for $i = 2, 3, \ldots, n$ which are
\[ a_{ISA} = \frac{1}{n-1} \sum_{i=2}^{n} (Y_i - A_{ISA}). \]  \hfill (9)

Therefore, \( a_{ISA} \) is the estimator of parameter \( \beta \) in simple linear regression under unconditional assumption.

The process with ISA approach to estimating parameters \( \alpha \) and \( \beta \) in simple linear regression are \( \beta_{ISA} \) and \( a_{ISA} \), respectively. Therefore, simple linear regression with SA approach is

\[ Y = a_{ISA} + \beta_{ISA}X. \]  \hfill (10)

The equation in (10) is simple linear regression with ISA approach for prediction. The process of ISA approach and SA approach simplify for estimating parameters \( \alpha \) and \( \beta \) in simple linear regression under unconditional assumption.

3. Results

The results for calculating the value of RMSE in simple linear regression under unconditional assumption by using the Data pairs in Table 1 of SA approach and ISA approach. For calculating the value of RMSE in simple linear regression of SA approach shows in Table 3. For calculating the value of RMSE in simple linear regression under unconditional assumption of ISA approach shows in Table 4. In additional, Table 3 and Table 4 present the process of SA approach and ISA approach. For OLS approach, the estimators for parameters \( \alpha \) and \( \beta \) in simple linear regression are showed in Cliff et al. [3].

The results for calculating the value of RMSE in simple linear regression under unconditional assumption from the Data pairs are shown in Table 2 and Table 3 for SA approach and ISA approach, respectively. In additional, Table 2 and Table 3 present the process of SA approach and ISA approach for estimating parameter in simple linear regression under unconditional assumption. The pairs data in Table 1 is used for calculating the value of RMSE for comparing the accuracy for predictive in simple linear regression under unconditional assumption. The results indicate that ISA approach with 47.1575 of RMSE is accuracy more than SA approach with 47.2740 of RMSE. However, ISA approach and SA approach simplify for estimating parameters \( \alpha \) and \( \beta \) in simple linear regression under unconditional assumption.

From the results in Table 3, the solution with SA approach for estimating parameters \( \alpha \) and \( \beta \) in simple linear regression under unconditional assumption is \( \beta_{SA} = 315.4545 \) and \( a_{SA} = 33.1667 \), respectively. Hence the simple linear regression equation of SA approach is presented as

\[ Y = 315.4545 + 33.1667X, \]  \hfill (9)

where \( Y \) is dependent variable and \( X \) is independent variable.

From the results in Table 4, the solution with SA approach for estimating parameters \( \alpha \) and \( \beta \) in simple linear regression under unconditional assumption is \( \beta_{ISA} = 304.4545 \) and \( a_{ISA} = 35.0000 \), respectively. Hence the simple linear regression equation ISA approach is presented as

\[ Y = 304.4545 + 35.0000X. \]  \hfill (10)

where \( Y \) is dependent variable and \( X \) is independent variable.

Therefore, ISA approach and SA approach are the process of estimating parameter in simple linear regression under unconditional assumption when we consider OLS approach which requires the fulfilment of some assumptions in the data samples, that is, the data samples are considered from the normally distributed population with \( E(\varepsilon) = 0 \) and \( \text{Var}(\varepsilon) = \sigma^2 \) in error term. OLS approach for estimating parameter \( \alpha \) and \( \beta \) in simple linear regression by using the Data pairs in Table 1, the parameter \( \alpha \) estimated by \( a \) and parameter \( \beta \) is estimated by \( b \) where the solution with OLS approach results in value of the estimator \( a = 307.9167 \) and the estimator \( b = 34.5833 \) that the simple linear regression with OLS approach is

\[ Y = 307.9167 + 34.5833X. \]  \hfill (11)

where \( Y \) is dependent variable and \( X \) is independent variable. For, OLS approach has 47.1367 of RMSE by supporting Cliff et al. [3]. However, OLS approach need to base on assumption in simple linear regression model. If we used parameter \( \alpha \) and \( \beta \) in simple linear regression. The assumptions need to examine where the error term is distributed as normal distribution with \( E(\varepsilon) = 0 \) and \( \text{Var}(\varepsilon) = \sigma^2 \).
### Table 3. Process and RMSE of SA approach by Data pairs of Cliff et al. [3]

| i | x | y   | $b_i = \frac{y_i - y_{i+1}}{x_i - x_{i+1}}$ | $\bar{x}_A = \frac{1}{n} \sum_{i=1}^{n} x_i$ | $\bar{x}_A \cdot x_i$ | $\sum_{i=1}^{n} \bar{x}_A \cdot x_i - \frac{n}{n-1} y_i - A$ | $\sum_{i=1}^{n} (\hat{y}_i - y_{iA})^2$ | $\epsilon^2$ |
|---|---|-----|-----------------------------------------|---------------------------------|-----------------|-------------------------------------------------|----------------------------------|------|
| 1 | 6 | 526 | 35.0000 | 33.16667 | 99.5000 | 222 | 315.4545 | 514.4545 | 133.2975 |
| 2 | 3 | 421 | 35.0000 | 33.16667 | 99.5000 | 222 | 315.4545 | 514.4545 | 133.2975 |
| 3 | 6 | 581 | 53.3333 | 199.0000 | 419 | 431 | 613.9545 | 514.4545 | 442.2890 |
| 4 | 9 | 630 | 16.3333 | 298.5000 | 341 | 213 | 414.9545 | 8.7293 |
| 5 | 3 | 412 | 36.3333 | 99.5000 | 199 | 235 | 514.4545 | 2911.0930 |
| 6 | 9 | 560 | 24.6667 | 298.5000 | 361 | 235 | 514.4545 | 6472.9340 |
| 7 | 6 | 434 | 42.0000 | 199.0000 | 235 | 235 | 514.4545 | 786.5475 |
| 8 | 3 | 443 | -3.0000 | 99.5000 | 244 | 244 | 514.4545 | 573.8202 |
| 9 | 9 | 590 | 24.5000 | 298.5000 | 391 | 391 | 514.4545 | 3085.298 |
| 10 | 6 | 570 | 6.6667 | 199.0000 | 371 | 371 | 514.4545 | 4754.729 |
| 11 | 3 | 346 | 74.6667 | 74.6667 | 147 | 147 | 514.4545 | 3369.275 |
| 12 | 9 | 672 | 54.3333 | 298.5000 | 473 | 473 | 514.4545 | 3369.275 |

RMSE = 47.2740

### Table 4. Process and RMSE of ISA approach by Data pairs of Cliff et al. [3]

| i | x | y | $b_i = \frac{y_i - y_{i+1}}{x_i - x_{i+1}}$ | $\bar{x}_{IA} = \frac{1}{n} \sum_{i=1}^{n} x_i$ | $\bar{x}_{IA} \cdot x_i$ | $\sum_{i=1}^{n} \bar{x}_{IA} \cdot x_i - \frac{n}{n-1} y_i - A$ | $\sum_{i=1}^{n} (\hat{y}_i - y_{IA})^2$ | $\epsilon^2$ |
|---|---|-----|-----------------------------------------|---------------------------------|-----------------|-------------------------------------------------|----------------------------------|------|
| 1 | 6 | 526 | 35.0000 | 35 | 105.000 | 210 | 211 | 304.4545 | 514.4545 | 133.2975 |
| 2 | 3 | 421 | 35.0000 | 35 | 105.000 | 210 | 211 | 304.4545 | 514.4545 | 133.2975 |
| 3 | 6 | 581 | 53.3333 | 210.000 | 420 | 420 | 619.4545 | 111.2066 |
| 4 | 9 | 630 | 16.3333 | 315.000 | 350 | 350 | 619.4545 | 3534.843 |
| 5 | 3 | 412 | 36.3333 | 105.000 | 224 | 224 | 514.4545 | 6472.934 |
| 6 | 9 | 560 | 24.6667 | 315.000 | 380 | 380 | 619.4545 | 867.5702 |
| 7 | 6 | 434 | 42.0000 | 210.000 | 360 | 360 | 514.4545 | 3085.298 |
| 8 | 3 | 443 | -3.0000 | 105.000 | 360 | 360 | 514.4545 | 3085.298 |
| 9 | 9 | 590 | 24.5000 | 315.000 | 360 | 360 | 514.4545 | 4026.479 |
| 10 | 6 | 570 | 6.6667 | 210.000 | 136 | 136 | 514.4545 | 2761.025 |
| 11 | 3 | 346 | 74.6667 | 210.000 | 462 | 462 | 514.4545 | 2761.025 |
| 12 | 9 | 672 | 54.3333 | 315.000 | 462 | 462 | 514.4545 | 2761.025 |

RMSE = 47.1575
For ISA approach and SA approach may become an alternative method in simple linear regression parameter estimation. The advantage of ISA approach and SA approach are free from all assumptions required by regression, such as error normality assumption while the data should be from the normal distribution. terms of RMSE is indicated that ISA approach and OLS approach are quite well compared with SA approach. gain, observed that ISA approach is comparable with OLS approach in simple linear regression.

For the data samples by Nwabueze [10], process and RMSE of SA approach and ISA approach are shown in Table 5 and Table 6.

### Table 5. Process and RMSE of SA approach by Data pairs of Nwabueze [10]

| i  | x     | y     | b_i = (y_i - \bar{y}) / (x_i - \bar{x}) | \hat{y}_{SA} = \frac{1}{n-1} \sum_{i=1}^{n} b_i x_i - \bar{x} | A_{SA} = \frac{1}{n-1} \sum_{i=1}^{n} (y_i - \hat{y}_{SA})^2 | \hat{y}_{ISA} | e_i^2 |
|----|-------|-------|---------------------------------------|----------------------------------------------------------------------|-----------------------------------------------------------------|----------------|-------|
| 1  | 911,100 | 6,511,330 |                                      |                                                                       |                                                                 | -2.459 \times 10^9 | 6.077 \times 10^{18} |
| 2  | 1,977,700 | 6,785,900 | 0.2574 | 30.4364 | 2.530 \times 10^9 | -2.486 \times 10^9 | 5.920 \times 10^{18} |
| 3  | 2,823,900 | 8,493,320 | 2.0177 |                                      | 4.773 \times 10^9 | 2.278 \times 10^{17} |
| 4  | 29,395,000 | 8,931,120 | 0.0165 |                                      | 2.484 \times 10^9 | 6.421 \times 10^{18} |
| 5  | 2,837,200 | 8,558,990 | 0.0140 |                                      | 2.484 \times 10^9 | 6.421 \times 10^{18} |
| 6  | 83,400,600 | 36,025,620 | 0.3409 |                                      | 5.197 \times 10^9 | 2.543 \times 10^{14} |
| 7  | 97,375,000 | 44,485,000 | -2.5748 |                                      | 4.692 \times 10^9 | 2.072 \times 10^{19} |
| 8  | 93,249,000 | 50,394,000 | -0.5176 |                                      | 2.484 \times 10^9 | 6.421 \times 10^{18} |
| 9  | 108,450,000 | 56,394,000 | 0.0546 |                                      | 1.034 \times 10^9 | 9.425 \times 10^{17} |
| 10 | 115,650,000 | 62,708,000 | 0.2996 |                                      | 1.592 \times 10^9 | 2.322 \times 10^{18} |
| 11 | 134,000,000 | 68,206,000 | 0.3409 |                                      | 4.320 \times 10^9 | 5.155 \times 10^{13} |
| 12 | 182,210,000 | 81,994,500 | 0.2860 |                                      | 2.734 \times 10^9 | 7.062 \times 10^{18} |
| 13 | 171,510,000 | 76,195,000 | 0.5420 |                                      | 5.398 \times 10^9 | 2.811 \times 10^{19} |
| 14 | 259,050,000 | 95,966,250 | 0.2259 |                                      | 5.398 \times 10^9 | 2.811 \times 10^{19} |

\( RMSE = 2,486,633,230 \)

### Table 6. Process and RMSE of SA approach by Data pairs of Nwabueze [10]

| i  | x     | y     | b_i = (y_i - \bar{y}) / (x_i - \bar{x}) | \bar{y}_{ISA} = \frac{1}{n} \sum_{i=1}^{n} y_i | A_{ISA} = \frac{1}{n} \sum_{i=1}^{n} (y_i - \bar{y}_{ISA})^2 | \hat{y}_{ISA} | e_i^2 |
|----|-------|-------|---------------------------------------|------------------------------------------------|-----------------------------------------------------------------|----------------|-------|
| 1  | 911,100 | 6,511,330 |                                      | 21970441                                      | 2.390 \times 10^{14}                                           |
| 2  | 1,977,700 | 6,785,900 | 0.2574 | 21394652.083 | 21735901 | 2.390 \times 10^{14} |
| 3  | 2,823,900 | 8,493,320 | 2.0177 | 23394652.083 | 22245011 | 2.390 \times 10^{14} |
| 4  | 29,395,000 | 8,931,120 | 0.0165 | 22462844 | 1.951 \times 10^{14} |
| 5  | 2,837,200 | 8,558,990 | 0.0140 | 29302922 | 4.150 \times 10^{14} |
| 6  | 83,400,600 | 36,025,620 | 0.3409 | 22466268 | 1.934 \times 10^{14} |
| 7  | 97,375,000 | 44,485,000 | -2.5748 | 43205339 | 5.155 \times 10^{13} |
| 8  | 93,249,000 | 50,394,000 | -0.5176 | 46802705 | 2.186 \times 10^{15} |
| 9  | 108,450,000 | 56,394,000 | 0.0546 | 21759905 | 8.199 \times 10^{14} |
| 10 | 115,650,000 | 62,708,000 | 0.2996 | 21763818 | 1.199 \times 10^{15} |
| 11 | 134,000,000 | 68,206,000 | 0.2996 | 51507156 | 1.255 \times 10^{14} |
| 12 | 182,210,000 | 81,994,500 | 0.2860 | 68641394 | 1.783 \times 10^{14} |
| 13 | 171,510,000 | 76,195,000 | 0.5420 | 65886942 | 1.063 \times 10^{14} |
| 14 | 259,050,000 | 95,966,250 | 0.2259 | 88421967 | 5.692 \times 10^{13} |

\( RMSE = 20,957,301 \)
From the results in Table 5, the solution with SA approach for estimating parameters $\alpha$ and $\beta$ in simple linear regression under unconditional assumption is $\beta_{\text{sa}} = 30.4364$ and $a_{\text{sa}} = -2.486 \times 10^9$, respectively. Hence the simple linear regression equation of SA approach is presented as

$$Y = -2.486 \times 10^9 + 30.4364 X,$$  \hspace{1cm} (12)

where $Y$ is dependent variable and $X$ is independent variable.

From the results in Table 4, the solution with SA approach for estimating parameters $\alpha$ and $\beta$ in simple linear regression under unconditional assumption is $\beta_{\text{isa}} = 0.2574$ and $a_{\text{isa}} = 21,735,901$, respectively. Hence the simple linear regression equation ISA approach is presented as

$$Y = 21,735,901 + 0.2574X,$$  \hspace{1cm} (13)

where $Y$ is dependent variable and $X$ is independent variable.

From the results in Table 5 and Table 6, we observed that RMSE of ISA approach is less than RMSE of SA approach. Therefore, ISA approach is better than SA approach in terms of accuracy based on RMSE.

4. Conclusions

This research is about estimating parameters in simple linear regression model under unconditional assumption for prediction. We improved SA approach by using median for replacing the average of the slope $\beta_{\text{sa}}$ for estimating parameter in simple linear regression model under unconditional assumption and it is called ISA approach. ISA approach has high accuracy for prediction under unconditional assumption when ISA approach is compared with SA approach based on the value of RMSE. For calculating the value of RMSE in Table 3 and Table 4, the RMSE values in SA approach and ISA approach are 47.2740 and 47.1575, respectively. For, the average of the slope $\beta_{\text{sa}}$ is replaced with the median of the slope $\beta_{\text{isa}}$ in SA approach which is ISA approach has accuracy for predictive in simple linear regression model under unconditional assumption by supporting Mohammad [7], and he used the median for non-normal distribution in population. When ISA approach and SA approach are compared to OLS approach with 47.1367 of RMSE supporting by Cliff et al. [3] where OLS approach is based on assumption of simple linear regression. ISA approach and OLS approach are similarly in terms of RMSE. Again observed that in Table 5 and Table 6 are RMSE of SA approach less than ISA approach. ISA approach is an alternative approach for estimating parameter in simple linear regression when assumption is violated. Our results suggest that ISA approach has high accuracy in terms of RMSE more than SA approach under unconditional assumption. Therefore, ISA approach should be used as an alternative approach for prediction on simple linear regression model.

Acknowledgements

This study is supported by Faculty of Sciences and Health Technology, Department of Informatics and computer Science, Department of Science and Mathematics, Kalasin University, Thailand.

REFERENCES

[1] Agung P., Agus S., Agustini T., Mustafa M., Sukono R. B., A New Method to Estimate Parameters in the Simple Regression Linear Equation, Mathematics and Statistics, Vol. 8, no. 2, pp. 75-81, 2020. DOI: 10.13189/ms.2020.080201.

[2] Chioma N. J., Causal Relationship between Gross Domestic Product and Personal Consumption Expenditure of Nigeria, African Journal of Mathematics and Computer Science Research, Vol. 2, no. 8, pp. 179-183, 2009.

[3] Cliff K. R., Billy K. M., Estimation of the Parameters of Linear Regression System using the Simple Averaging Method, Global Journal of Pure and Applied Mathematics, Vol. 13, No. 11, pp. 7749-7758, 2017.

[4] Gustiarisanie A., Rahardjo M. F., Emawati Y., Hubungan Panjang-Bobot dan Faktor Kondisi Ikan LidahCynoglossus cyanoglossus, Hamilton1882 (Pisces: Cynoglosiade) di Teluk Pabean Indramayu, Jawa Barat, Journal of applied dental and medical studies, Journal of Applied Dental and Medical Science, Vol. 2, no. 1, pp. 199-208, 2016.

[5] Hyo I. P., Median test for multivariate and right censored data, Journal of Nonparametric Statistics, Vol. 16, no. 5, pp. 753-760, 2014.

[6] Jain S., Chourse S., Dubey S., Jain S., Kamakoty J., Jain D., Regression Analysis – Its Formulation and Execution in Dentistry, Journal of Applied Dental and Medical Science, Vol. 2, no. 1, pp. 199-208, 2016.

[7] Mohammad I. S. G., A New Non-parametric Test Procedure for the Median of an Asymmetrical Population, American Journal of Theoretical and Applied Statistics, Vol. 5, no. 6, pp. 376-386, 2016.

[8] Mohammad H. R., Sangbum C., Chuan H., Liang Z., Sangchoon J., Joseph C. G., Nonparametric estimation of median survival times with applications to multi-site or multicenter studies, PLoS one, Vol. 13, no. 5, 2018.

[9] Nick F., Nonparametric Tests of Differences in Medians: Comparison of the Wilcoxon–Mann–Whitney and Robust Rank-Order Tests, Experimental Economics, Vol. 6, no. 3, pp. 273-297, 2003.
[10] Nwabueze J. C., Causal relationship between gross domestic product and personal consumption expenditure of Nigeria, African Journal of Mathematics and Computer Science Research, vol. 2, no. 8, pp. 179-183, 2009.

[11] Permai S. D., Tanty H., Linear Regression Model using Bayesian Approach for Energy Performance of Residential Building, Procedia Computer Science, Vol. 135, pp. 671-677, 2018.

[12] Prabitha J., Archana S., Application of Regression Analysis in Numerous Times, International Journal of Science Engineering and Technology Research, Vol. 4, no. 4, pp. 1002-1005, 2005.

[13] Przemyslaw G., K-Sample Median Test for Vague Data, International Journal of Intelligent Systems, vol. 24, no. 5, pp. 529-539, 2009.

[14] Qing K., Paul I. N., Nonparametric tests for the median from a size-biased sample, Journal of Nonparametric Statistics, Vol. 20, No. 1, pp. 19–37, 2008.

[15] Singh A. S., Masuku M. B., Application of Modeling and Statistical Regression Techniques in Research, Research Journal of Mathematical and Statistical Sciences, Vol. 1, no. 16, pp. 14-20, 2013.