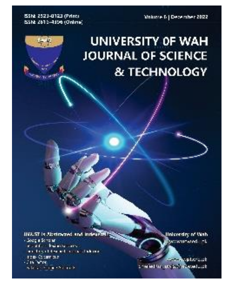
University of Wah Journal of Science and Technology Vol. 8 (2024),PP 66-76.
ISSN: 2523-0123 (Print) 2616-4396 (Online)

University of Wah
Journal of Science and Technology

www.uow.edu.pk



Monte-Carlo Study of Some Robust Estimators: The Simple Linear Regression Case.

Adewole, Ayoade I*[1]. Bodunwa, Oluwatoyin K[2] . and Oseni, Kazeem O[3].

*[*1]Department of Mathematics, Tai Solarin University of Education Ijagun Ogun State Nigeria.*

[2]Department of Statistics, Federal University of Technology Akure .Ondo State Nigeria.

[3]Department of Computer Science, Richmond American International University London.

*Corresponding author email : hayorhade2005@gmail.com,

Abstract

In this study, Least Trimmed Squares (LTS), Theil's Pair-wise Median (Theil) and Bayesian estimation methods (BAYES) are compared relative to the OLSE via Monte-Carlo Simulation. Variance, Bias, Mean Square Error (MSE) and Relative Mean Square Error (RMSE) were calculated to evaluate the estimators' performance. The Simple Linear Regression model is explored for the conditions in which the error term is assumed to be drawn from three error distributions: unit normal, lognormal and Cauchy. Theil's non-parametric estimation procedure was found to have the strongest and most reliable performance. The subsequent-best results are acquired from LTS approach Though it was observed that the Bayesian estimators are affected by deviation of the dataset from normality, yet it is established from the results that the Bayesian estimators performed optimally more than all other competitors, even under non normal situations (especially under the standard lognormal distribution) in some cases, except whenever the error is drawn from a heavy tail distribution (Lognormal and Cauchy)..OLSE is most effective reliable as long as the normality assumptions preserve

Keywords

Robust estimation, Monte-Carlo, Lognormal and Cauchy.

1. Introduction

Frequentist approach of linear regression yields a single estimate for the model parameters based only on the assumption that the model is absolutely influenced by the nature of the data set involved, Nevertheless, the methods of classical estimations of regression models' squares is not robust to violations of any of its underlying assumptions. Estimates of Ordinary least square regression was heavily influence by any presence of outliers in data sets [1] Thus, researchers [2-4] among others developed practical alternative methods, which have been considered with numerous advantages over frequentist method of estimation.

This research delved into other alternatives to Ordinary least square regression estimation procedures such as the Least Trimmed Squares (LTS) introduced by [5], Theil's pairwise median procedure introduced in 1950. that assumes to performs better exclusive of the distribution of the error terms and the Bayesian inference method developed by [6], with further advancement by researchers such as [7-9] among others.

The Bayesian approach methods of estimation necessitate the usage of probability distributions rather than point estimates in the expressions of linear regression models. It establishes the update of prior beliefs in the evidence of new data. Bayesian analysis aim is to determine the posterior distribution, and never to ascertain the "best" estimate for the parameters of the model which is the fundamental bedrock of Bayesian Inference. Trace back to 1980s, Bayesian methods have been widely used among researchers within statistics and have found usage in many fields due to the detection of Markov chain Monte Carlo methods [9]. [10] worked on simulation studies that compared the OLSE regression estimation with the Theil pairwise median and weighted Theil estimators using one hundred replications per situation.[11] used MCMC simulation to examine the regression estimates of OLSE, Least Trimmed Squares (LTS) and Theil, along other frequentist regression estimators for the Simple Linear Regression model that have a Generalized Logistic distribution error term.[8] considered different approaches to robust Bayesian inferencefor set identified structural vector autoregressions using numerical and empirical illustrations. [12] employed Bayesian estimation in determining the ranges of the quantiles that revealed the optimal hyperparameters from samples with vague facts in empirical Bayesian inference [13] expatiated the shortcomings of datasets showing multicollinearity using frequentist approach and discusses the Bayesian approach which serve as a relief to some of the problems outfaced by the multiple regression approach. [14] developed a new non parametric hypothesis testing with reliability analysis applications to model some real data using lifetime distribution theory and Monte Carlo Simulations.

Very little research from the literature captures works in which the Bayesian method to linear regression is directly compared to Robust-non-parametric alternatives. The current study investigates the conduct of robust regression and Bayesian inference techniques to Simple Linear Regression under three different conditions in relation to contaminated data and non-normal error distributions.

2. Methodology

2.1 Simulation Design

The study design employs a Monte-Carlo simulation approaches, the scheme of the explanatory variable X is created as a consecutive model of the structure $X_t = t;\ t = 1,2,3,\dots,n$, while the response variable Y is generated as a form of linear model $Y_t = \alpha + \beta X_t + \varepsilon_t$ (for $\alpha = 0$ and $\beta = 1$). 150,000 sets of random data of sample sizes $n = 30$ and 500 were simulated, the random component ε_t was assumed to be drawn three different types of error distribution namely Normal, Lognormal and Cauchy. The Monte Carlo simulations procedures using MATLAB generates random values with the original data by setting the number of simulations and the distribution parameters according to distribution type.

The following iterative procedures was adopted;

i. Set up the predictive model and identify both the dependent variable and the drive of the prediction which is the independent variables and define every inputs.
ii. Specify the probability distributions.
iii. Set up the number of simulations.
iv. Run simulations repeatedly by generating the random values of the independent variables.
v. Aggregate and assess the outputs from the simulations.
vi. Estimate the parameters such as the mean, standard deviation among others.
vii. Draw conclusions.

2.2 Choice of Error Distributions

Two heavy tailed distributions namely Cauchy and Lognormal distributions were used to inspect the efficiency of each estimator as the dataset deviates from normality. However, the Cauchy distribution has the properties of heavier tails than the Log-normal distribution. Sensitivity of each estimation methods to outliers were examined by alternative forms of the error distributions (mixture, outlier and contamination). Detailed information on this technique, and procedures for drawing random deviates from each of the error distributions, under this study had been discussed in [15].

2.3 Estimation Procedures

For all simulated records set the estimate of α and β were estimated using the four estimation approaches earlier described, for every estimator, the validation statistic such as the average, the variance, the bias and the mean square error were computed. The mean square error (MSE) was estimated as:

$MSE\ (\beta) = Var\ (\beta)\ + [Bias\ (\beta)]^2$.

Previous literatures such as [16-17] described the details on the algorithms and method of estimations using these estimators. However, brief procedure of each estimator is reviewed below:

2.4 Ordinary Least Squares Method

OLS method of estimation is a classical technique for estimating coefficients in a linear regression model by minimizing the sum of squared errors. This yields an estimation of the mean function of the dependent variables which is conditionally distributed. Ordinary Least Squares Method accomplished the property of Best, Linear and Unbiased Estimator (BLUE), if

$$Y = X_t'\beta + e_i \tag{1}$$

where

Y is the response variable,

$X_t'\beta$ is the t^{th} of the matrix X,

e_i is the error terms.

then following assumptions hold.

- The relationship between Y and X requires that the dependent variable (Y) is a linear combination of explanatory variable and error term
- The independent variable X_i is non-Stochastic.
- Homoscedacity of the residuals.
- No Serial correlation of the error terms.
- Normally distributed residuals.

However, frequently one or more of these assumptions are violated, it results in OLS not any more the best linear unbiased estimator. However, these assumptions are stringent such that if any one of the assumptions is not met, OLSE procedure breaks down.

2.5 The Least Trimmed Squares (LTS)

The Least Trimmed Squares (LTS) introduced by [5] aims at minimizing $\sum_{i=1}^{h}\left(y_i - \hat{\alpha} - \hat{\beta}x_i\right)^2$ by choosing a subsample of h observations, computing some α and β that minimize the sum of squared errors for the chosen subsample after which deleting sets of data related to a selected percent of the prevalent residuals underneath an preliminary OLSE to reduce their adverse results on the inferences[11]. Consequently, the significant disparity amid OLSE and LTS estimation is that LTS method of estimation is not really affected by the outliers due to presence of large squares errors.[11]

2.6 Theil's Pair-Wise Median Methods

The Theil's method is a non-parametric procedure based on using the ranks of the observed data rather than using the actual values of the observed data [11],[4] pairwise style comparison is done to each data pair and all other in computing the complete Theil's slope estimate. A data set of $n(X,Y)$ pairs will result in N $= \binom{n}{2} = \frac{n(n-1)}{2}$ pairwise comparisons, for each of these a slope $\frac{\Delta Y}{\Delta X}$ is computed. The median of all possible pairwise slopes is taken as the non-parametric Theil's slope estimate, β^{THEIL}, where; $b_{ij} = \frac{\Delta Y}{\Delta X} = \frac{y_i - y_j}{x_j - x_i}$; $x_j \neq x_i; 1 \leq i \leq j \leq n$. The y-interceptis obtained by calculating; $a_{ij} = \frac{x_j y_i - x_i y_j}{x_j - x_i}$; $x_j \neq x_{i;i<j}$ and taking the median of these a_{ij} values as the y-intercept.

2.7 Bayesian Inference Method

The application of Bayesian estimation requires a probabilistic reformulation of the Simple Linear Regression model based on major fundamental assumptions of classical regression models. To achieve this, the response variable Y, and the model parameters α, β and ε are assumed to come from a predetermined (prior) distribution. Moreover, an appropriate specified likelihood function component is the part that incorporates the data.

The posterior probability of the model parameters is conditional upon the training inputs and outputs using the prior belief and the likelihood of the model given the data.

$$P(\beta|y,X) = P(y|\beta,X) * P(\beta|X)P(y|X). \tag{2}$$

Here; $P(\beta|y,X)$ is the posterior probability distribution of the model parameters. This is equal to the likelihood of the data, $P(y|\beta,X)$, multiplied by the prior probability of the parameters and divided by a normalization constant.

The joint likelihood of the i^{th}, denoted L_i:

$$\mathrm{L_i}\ (\alpha_{\bar{x}}, \beta) \propto \exp(-\frac{1}{2}(\frac{\left(y_i - (\alpha_{\bar{x}} + \beta(x_i - \bar{x}))\right)^2}{\sigma^2}) \tag{3}$$

The joint likelihood of the whole sample of all observations is the product of the independent likelihoods

$$\mathrm{L_{sample}}\ (\alpha_{\bar{x}}, \beta) \propto \Pi_{i=1}^{n} \exp(-\frac{1}{2}(\frac{\left(y_i - (\alpha_{\bar{x}} + \beta(x_i - \bar{x}))\right)^2}{\sigma^2}). \tag{4}$$

Which is simplified to

$$SS_y + 2\beta SS_{xy} + \beta^2 SS_x + n(\alpha_{\bar{x}} - \bar{y})^2, \tag{5}$$

Where

$$SS_y = \Sigma_{i=1}^{n}(y_i - \bar{y})^2; \tag{6}$$

$$SS_{xy} = \Sigma_{i=1}^{n}(y_i - \bar{y})(x_i - \bar{x}); \tag{7}$$

$$SS_x = (x_i - \bar{x})^2 \tag{8}$$

Thus, the joint likelihood of sample is;

$$\mathrm{L_{sample}}\ (\alpha_{\bar{x}}, \beta)\ \alpha \exp(-\frac{1}{2}\left(\frac{(SS_y - 2\beta SS_{xy} + \beta^2 SS_x + n(\alpha_{\bar{x}} - \bar{y})^{2^2}}{\sigma^2}\right)$$

$$\alpha \exp(-\frac{1}{2}(\frac{(SS_y - 2\beta SS_{xy} + \beta^2 SS_x)}{\sigma^2}) \exp(-\frac{1}{2}(\frac{n(\alpha_{\bar{x}} - \bar{y})^2}{\sigma^2}) \tag{9}$$

Factorizing βSS_x and completing the squares;

$$\mathrm{L_{sample}}\ (\alpha_{\bar{x}}, \beta)\ \alpha \exp(-\frac{1}{\frac{2\sigma^2}{SS_x}}(\beta - \mathrm{B})^2 \text{ x } \exp(-\frac{1}{\frac{2\sigma^2}{n}}(\alpha_{\bar{x}} - \mathrm{A}_{\bar{x}})^2, \tag{10}$$

where; $\frac{SS_{xy}}{SS_x} = \mathrm{B}$ (the least squares slope); and $\bar{y} = \mathrm{A}_{\bar{x}}$ (the least squares intercept)

The joint prior of $\alpha_{\bar{x}}$ and β is the product of the individual prior defined as

$$G(\alpha_{\bar{x}}, \beta) = G(\alpha_{\bar{x}})\ x\ G(\beta). \tag{11}$$

By Bayes' rule, the joint posterior is proportional to the joint likelihood multiplied by the joint prior

$$G(\alpha_{\bar{x}}\ ,\beta|data)\ \alpha\ G(\alpha_{\bar{x}}\ ,\beta)\ x\ L_{sample}\ (\alpha_{\bar{x}}\ ,\beta) \tag{12}$$

where the data is the set of ordered pair (X_i, Y_i).

$$G(\alpha_{\bar{x}}\ ,\beta|data)\ \alpha\ G(\alpha_{\bar{x}}|data)\ x\ G(\beta|data) \tag{13}$$

The marginal posteriors are independent and can be found by the simple updating rules for normal distributions:
Given (m_β, s_β^2) prior for β, then the posterior is; $N(m'_\beta, (s'_\beta)^2)$; where

$$\frac{1}{(s'_\beta)^2} = \frac{1}{s_\beta^2} + \frac{SS_x}{\sigma^2} \tag{14}$$

$$m'_\beta = \frac{\frac{1}{s_\beta^2}}{\frac{1}{(s'_\beta)^2}} \text{ x } m_\beta + \frac{\frac{SS_x}{\sigma^2}}{\frac{1}{(s'_\beta)^2}} \text{ x B} \tag{15}$$

Also, Given $(m_{\alpha_{\bar{x}}}, s_{\alpha_{\bar{x}}}^2)$ prior for $\alpha_{\bar{x}}$, then the posterior is;
$N(m'_{\alpha_{\bar{x}}}, (s'_{\alpha_{\bar{x}}})^2)$
where

$$\frac{1}{(s'_{\alpha_{\bar{x}}})^2} = \frac{1}{s_{\alpha_{\bar{x}}}^2} + \frac{n}{\sigma^2}; \tag{16}$$

and

$$m'_{\alpha_{\bar{x}}} = \frac{\frac{1}{s_{\alpha_{\bar{x}}}^2}}{\frac{1}{(s'_{\alpha_{\bar{x}}})^2}} \text{ x } m_{\alpha_{\bar{x}}} + \frac{\frac{n}{\sigma^2}}{\frac{1}{(s'_{\alpha_{\bar{x}}})^2}} \text{ x } A_{\bar{x}} \tag{17}$$

For the explicit cases of this work, the probabilistic reframe of the regression model parameters presumes a conjugate Normal prior distribution that has mean μ with the regression slope parameter β and the regression intercept parameter α. The prior standard deviation was calculated as

$$\left(\frac{(m'+3\sigma)-(m'-3\sigma)}{6}\right)^2 \tag{18}$$

where σ is the OLSE sample estimate, while employing the Normal probability distribution density

$$(\mathrm{N}(\mu, \sigma)) \sim \frac{\exp(\frac{-(x-\sigma)^2}{2\sigma^2})}{\sigma(2\pi)^{\frac{1}{2}}}) \tag{19}$$

as the likelihood of the data across the simulation.

3. Simulation Results and Discussions

Tables 1A, 1B, and 2A, 2B shows the results across sample sizes. Estimator variances for both the intercept and the slope parameters decreased with increasing sample size. All forms of the error distribution records decrease in σ^2 and bias, the pattern is also revealed in the mean square error's values.
Selected slope estimators are approximately unbiased. There is a consistent improvement in the pattern of performance of error distribution and respective alternate models as there is a noticeable increase in the size of the sample. Results of the research study reveals that OLSE and Bayesian slope estimators are always biased with Cauchy distributed error term regardless of sample size (though the performance improve with increasing sample size)

TABLE 1A: GENERAL ESTIMATORS' PERFORMANCE RESULTS CHART FOR SOME REGRESSION ESTIMATORS (SAMPLE SIZE $N = 30$)

Tables 1A gives the general estimators' performance results chart for Regression Estimators of standard model and the models with outlier of some error distribution with sample size $n = 30$. The Least Trimmed Squares estimation approach followed Theil's underneath the normal distribution for sample size $n = 30$, the result discovered the father away the error distribution diverges from normality, the more efficient LTS turs into as it acquires more accuracies following closely after Theil's while replacing Bayesian estimator and OLSE absolutely. Also, it was observed that the Theil's estimator revealed the most satisfactory performance across all the cells of the simulation followed closely by LTS estimator (except whenever the sample size is very small). The Bayesian estimator maintained consistent relative unbiased-ness only under the standard and mixture normal error model. This confirms the elusiveness of the OLSE when the data set is heteroscedastic.

STANDARD MODEL												
NORMAL												
N=30	α					β					**CREDIBLE INTERVAL**	
METHOD	$\bar{x}$	$\boldsymbol{\sigma^2}$	**Bias**	**MSE**	**RMSE**	$\bar{x}$	$\boldsymbol{\sigma^2}$	**Bias**	**MSE**	**RMSE**	**LCL**	**UCL**
OLSE	-0.0196	0.047	0.0196	0.477	0.000	0.995	0.012	-0.005	0.012	0.000	0.808	1.180
LTS	0.0314	0.454	0.0314	0.452	-0.218	0.991	0.015	0.008	0.015	-0.218	0.790	1.190
THL	0.0059	0.083	0.0059	0.083	-0.039	0.996	0.004	-0.005	0.004	-0.039	-5.571	7.700
BAYES	0.0188	0.375	0.0188	0.376	0.0031	0.995	0.015	-0.005	0.015	0.031	0.800	1.190
LOGNORMAL												
N=30	α					β					**CREDIBLE INTERVAL**	
METHOD	$\bar{x}$	$\boldsymbol{\sigma^2}$	**Bias**	**MSE**	**RMSE**	$\bar{x}$	$\boldsymbol{\sigma^2}$	**Bias**	**MSE**	**RMSE**	**LCL**	**UCL**
OLSE	1.660	2.180	1.660	2.180	1.020	0.997	0.052	0.003	0.052	1.000	0.024	1.760
LTS	1.020	0.750	1.020	0.750	0.504	1.010	0.025	0.033	0.025	0.050	0.022	2.000
THL	1.130	0.674	1.130	0.674	0.699	1.000	0.015	0.001	0.015	0.009	-2.85	4.850
BAYES	1.560	2.160	1.560	2.160	0.690	0.997	0.053	0.003	0.053	0.027	0.602	1.390
CAUCHY												
N=30	α					β					**CREDIBLE INTERVAL**	
METHOD	$\bar{x}$	$\boldsymbol{\sigma^2}$	**Bias**	**MSE**	**RMSE**	$\bar{x}$	$\boldsymbol{\sigma^2}$	**Bias**	**MSE**	**RMSE**	**LCL**	**UCL**
OLS	-0.390	1.520	-3.900	1.520	0.000	7.260	3.160	6.260	3.160	0.000	-2.130	3.590
LTS	-0.009	4.130	-0.009	4.130	1.000	0.972	5.090	-0.028	5.090	1.000	-4.331	4.520
THL	-0.130	-4.390	-0.133	-4.390	0.000	1.030	1.450	0.026	1.450	0.000	-3.271	3.470
BAYES	-0.901	1.501	-3.920	1.501	1.080	7.260	3.160	6.218	3.160	0.010	4.840	9.680
OUTLIERS MODEL												
NORMAL												
N=30	α					β					**CREDIBLE INTERVAL**	
METHOD	$\bar{x}$	$\boldsymbol{\sigma^2}$	**Bias**	**MSE**	**RMSE**	$\bar{x}$	$\boldsymbol{\sigma^2}$	**Bias**	**MSE**	**RMSE**	**LCL**	**UCL**
OLSE	-0.141	2.170	-0.142	2.170	0.000	1.010	0.385	0.207	0.385	0.000	-8.220	2.860
LTS	-0.034	6.320	-0.031	6.320	0.653	1.000	0.134	0.005	0.134	0.654	-1.690	3.700
THL	0.011	-4.400	0.011	-4.400	0.796	1.000	0.079	0.003	0.079	0.797	-2.440	4.440
BAYES	-0.104	2.130	-0.104	2.130	0.022	1.020	0.394	0.021	0.394	0.021	0.484	1.640
LOGNORMAL												
N=30	α					β					**CREDIBLE INTERVAL**	
METHOD	$\bar{x}$	$\boldsymbol{\sigma^2}$	**Bias**	**MSE**	**RMSE**	$\bar{x}$	$\boldsymbol{\sigma^2}$	**Bias**	**MSE**	**RMSE**	**LCL**	**UCL**
OLSE	1.260	2.190	1.260	2.190	0.000	-1.720	4.070	-1.720	4.070	0.000	0.000	1.920
LTS	1.680	6.650	1.680	6.650	1.000	-2.240	1.301	-2.340	1.301	1.000	-2.260	2.210
THL	1.030	-7.330	1.830	-7.330	-7.000	0.918	0.090	-8.260	0.090	0.000	0.244	1.590
BAYES	1.260	2.193	1.261	2.193	-0.001	-1.721	4.070	-1.720	4.070	-0.002	0.001	1.720
CAUCHY												
N=30	α					β					**CREDIBLE INTERVAL**	
METHOD	$\bar{x}$	$\boldsymbol{\sigma^2}$	**Bias**	**MSE**	**RMSE**	$\bar{x}$	$\boldsymbol{\sigma^2}$	**Bias**	**MSE**	**RMSE**	**LCL**	**UCL**
OLSE	7.510	1.000	7.510	1.000	0.003	-0.004	1.892	-1.012	1.896	0.000	-2.569	2.564
LTS	0.102	3.559	0.102	3.559	0.996	0.986	7.512	0.0145	7.513	0.996	-6.074	6.096
THL	-0.166	0.283	-0.107	0.283	0.004	1.025	0.439	0.0240	0.440	0.000	-5.245	5.261
BAYES	7.514	1.003	7.515	1.003	0.003	-0.004	1.892	-0.011	1.089	0.003	-7.251	7.247

TABLE 1B: GENERAL ESTIMATORS' PERFORMANCE RESULTS CHART FOR SOME REGRESSION ESTIMATORS (SAMPLE SIZE N=30)

Tables 1B gives the general estimators' performance results chart for Regression Estimators of mixture models and the models

MIXTURE MODEL												
NORMAL												
N=30	α					β					**CREDIBLE INTERVAL**	
METHOD	$\bar{x}$	$\boldsymbol{\sigma^2}$	**Bias**	**MSE**	**RMSE**	$\bar{x}$	$\boldsymbol{\sigma^2}$	**Bias**	**MSE**	**RMSE**	**LCL**	**UCL**
OLSE	0.009	0.466	0.003	0.464	0.000	1.000	0.121	0.000	0.121	0.100	0.731	1.270
LTS	0.000	0.435	0.005	0.435	-0.218	1.000	0.014	0.002	0.014	0.022	0.635	1.360
THL	0.002	0.063	0.000	0.063	-0.140	1.000	0.002	0.000	0.002	-0.142	-4.0000	6.000
BAYES	0.008	0.303	0.001	0.368	0.003	1.000	0.012	0.000	0.012	0.032	0.766	1.230
LOGNORMAL												
N=30	α					β					**CREDIBLE INTERVAL**	
METHOD	$\bar{x}$	$\boldsymbol{\sigma^2}$	**Bias**	**MSE**	**RMSE**	$\bar{x}$	$\boldsymbol{\sigma^2}$	**Bias**	**MSE**	**RMSE**	**LCL**	**UCL**
OLSE	1.708	2.380	1.750	2.380	0.000	0.9851	0.059	-0.014	0.059	0.000	0.421	1.550
LTS	1.080	0.878	1.080	0.878	5.301	1.000	0.028	0.003	0.028	0.530.	0.044	1.960
THL	1.190	0.766	1.190	0.766	1.130	0.990	0.078	-0.007	0.078	0.003	-2.020	4.000
BAYES	1.650	2.370	1.650	2.370	6.453	0.985	0.058	-0.015	0.058	0.006	0.646	1.320
CAUCHY												
N=30	α					β					**CREDIBLE INTERVAL**	
METHOD	$\bar{x}$	$\boldsymbol{\sigma^2}$	**Bias**	**MSE**	**RMSE**	$\bar{x}$	$\boldsymbol{\sigma^2}$	**Bias**	**MSE**	**RMSE**	**LCL**	**UCL**
OLSE	0.690	6.280	0.690	6.280	0.000	0.978	1.120	-0.022	1.120	0.000	2.700	2.890
LTS	0.076	6.390	0.076	6.390	0.998	1.000	0.238	0.0083	0.238	0.998	-0.462	4.620
THL	0.046	3.500	0.045	3.500	0.000	0.980	0.020	-0.002	0.020	0.000	-3.310	3.510
BAYES	0.689	6.280	0.689	6.280	0.000	0.938	1.120	-0.025	1.120	0.000	-1.420	3.370
CONTAMINATIONS MODEL												
NORMAL												
N=30	α					β					**CREDIBLE INTERVAL**	
METHOD	$\bar{x}$	$\boldsymbol{\sigma^2}$	**Bias**	**MSE**	**RMSE**	$\bar{x}$	$\boldsymbol{\sigma^2}$	**Bias**	**MSE**	**RMSE**	**LCL**	**UCL**
OLSE	0.012	0.455	0.012	0.455	0.000	0.998	0.011	-0.002	0.011	0.000	0.808	1.190
LTS	0.005	0.422	0.008	0.422	-0.248	1.000	0.014	0.001	0.014	-0.248	0.781	1.220
THL	0.007	0.039	0.008	0.039	-0.105	1.000	0.013	0.001	0.013	-0.147	-4.150	6.150
BAYES	0.008	0.358	0.009	0.358	0.032	0.998	0.011	-0.002	0.011	0.032	0.802	1.190
LOGNORMAL												
N=30	**ALPHA**					β					**CREDIBLE INTERVAL**	
METHOD	$\bar{x}$	$\boldsymbol{\sigma^2}$	**Bias**	**MSE**	**RMSE**	$\bar{x}$	$\boldsymbol{\sigma^2}$	**Bias**	**MSE**	**RMSE**	**LCL**	**UCL**
OLSE	1.600	2.430	1.600	2.430	0.000	1.010	0.066	0.0101	0.066	0.000	-0.283	2.300
LTS	0.982	0.783	0.982	0.783	0.541	1.030	0.027	0.026	0.027	0.549	-0.777	2.830
THL	1.130	0.729	1.130	0.729	0.004	1.010	0.017	0.006	0.017	0.044	-2.950	4.970
BAYES	1.510	2.410	1.510	2.410	0.007	1.010	0.066	0.017	0.066	0.057	0.497	1.530
CAUCHY												
N=30	α					β					**CREDIBLE INTERVAL**	
METHOD	$\bar{x}$	$\boldsymbol{\sigma^2}$	**Bias**	**MSE**	**RMSE**	$\bar{x}$	$\boldsymbol{\sigma^2}$	**Bias**	**MSE**	**RMSE**	**LCL**	**UCL**
OLSE	-4.400	1.550	-4.400	1.550	0.000	1.650	4.570	0.065	4.570	0.000	-0.631	3.930
LTS	0.104	4.060	0.104	4.060	1.000	0.9920	0.214	-0.008	0.214	1.000	-2.170	4.160
THL	0.099	0.340	-0.099	0.340	1.000	0.9901	0.144	-0.010	0.144	0.000	-8.020	1.000
BAYES	-4.400	1.554	-4.400	1.554	0.000	1.6500	4.570	0.650	4.570	0.000	0.967	2.330

with contaminations of some error distribution with sample size $n = 30$. From the results, Theil's slope estimator outperformed every other estimator regardless of sample size or distribution, OLSE slope estimator was better than LTS when sample size is small $(n<50)$.; Bayesian point estimator is more efficient with the normal distribution but relented whenever the error term is lognormal or Cauchy distributed. Theils estimator performed well regardless of the distribution type and under non-normal error distribution situation

TABLE 2A: GENERAL ESTIMATORS' PERFORMANCE RESULTS CHART FOR SOME REGRESSION ESTIMATORS (SAMPLE SIZE $N = 500$)

STANDARD MODEL												
NORMAL												
N=500	α					β					**CREDIBLE INTERVAL**	
METHOD	$\bar{x}$	σ^2	**Bias**	**MSE**	**RMSE**	$\bar{x}$	σ^2	**Bias**	**MSE**	**RMSE**	**LCL**	**UCL**
OLSE	-0.006	0.012	-0.006	0.012	0.000	1.000	0.000	0.000	0.000	0.000	0.999	1.000
LTS	-0.011	0.012	-0.011	0.012	-0.227	1.000	0.000	0.000	0.000	-0.227	0.999	1.000
THL	-0.001	0.005	-0.008	0.005	-0.027	1.000	0.000	0.000	0.000	-0.027	-1.500	1.700
BAYES	-0.006	0.012	-0.006	0.012	0.000	1.000	0.000	0.000	0.000	0.000	0.999	1.000
LOGNORMAL												
N=500	α					β					**CREDIBLE INTERVAL**	
METHOD	$\bar{x}$	σ^2	**Bias**	**MSE**	**RMSE**	$\bar{x}$	σ^2	**Bias**	**MSE**	**RMSE**	1.000	**UCL**
OLSE	1.600	0.048	1.620	0.048	0.000	1.000	0.000	0.000	0.000	0.000	1.000	1.000
LTS	1.090	0.023	1.000	0.023	0.396	1.000	0.000	0.000	0.000	0.320	1.000	1.000
THL	1.000	0.078	1.000	0.078	0.002	1.000	0.000	0.000	0.000	0.000	-1.000	1.900
BAYES	1.610	0.038	1.61	0.038	0.000	1.000	0.000	0.000	0.000	0.000	1.000	1.000
CAUCHY												
N=500	α					β					**CREDIBLE INTERVAL**	
METHOD	$\bar{x}$	σ^2	**Bias**	**MSE**	**RMSE**	$\bar{x}$	σ^2	**Bias**	**MSE**	**RMSE**	**LCL**	**UCL**
OLSE	-2.481	8.853	-2.481	8.853	0.000	1.100	0.100	0.111	0.100	0.000	0.454	1.770
LTS	-0.701	1.410	-0.470	1.410	1.000	1.000	0.000	0.002	0.000	0.000	0.249	1.760
THL	0.002	0.050	0.100	0.050	0.000	1.000	0.000	0.001	0.000	0.000	-2.170	2.371
BAYES	2.401	8.850	2.401	8.850	0.000	1.200	0.000	0.113	0.000	0.000	1.090	1.140
OUTLIERS MODEL												
NORMAL												
N=500	α					β					**CREDIBLE INTERVAL**	
METHOD	$\bar{x}$	σ^2	**Bias**	**MSE**	**RMSE**	$\bar{x}$	σ^2	**Bias**	**MSE**	**RMSE**	**LCL**	**UCL**
OLSE	-0.001	0.012	-0.001	0.012	0.000	1.0000	0.000	0.000	0.000	0.000	0.999	1.000
LTS	-0.010	0.158	-0.010	0.158	0.224	1.0000	0.000	0.000	0.000	-0.224	0.999	1.000
THL	-0.005	0.015	-0.005	0.015	0.001	1.0000	0.000	0.000	0.000	0.000	-1.450	1.650
BAYES	-0.001	0.280	-0.001	0.280	0.000	1.0000	0.000	0.000	0.000	0.000	0.999	1.000
LOGNORMAL												
N=500	α					β					**CREDIBLE INTERVAL**	
METHOD	$\bar{x}$	σ^2	**Bias**	**MSE**	**RMSE**	$\bar{x}$	σ^2	**Bias**	**MSE**	**RMSE**	**LCL**	**UCL**
OLSE	1.660	0.004	1.660	0.004	0.000	1.000	0.000	0.000	0.000	0.000	0.999	1.000
LTS	1.140	0.037	1.140	0.037	0.095	1.000	0.000	0.000	0.000	0.095	0.999	1.000
THL	1.020	0.011	1.020	0.011	0.008	1.050	0.000	0.000	0.000	0.008	-0.142	1.150
BAYES	1.660	0.040	1.660	0.040	0.000	1.000	0.000	0.000	0.000	0.000	0.999	1.000
CAUCHY												
N=500	α					β					**CREDIBLE INTERVAL**	
METHOD	$\bar{x}$	σ^2	**Bias**	**MSE**	**RMSE**	$\bar{x}$	σ^2	**Bias**	**MSE**	**RMSE**	**LCL**	**UCL**
OLSE	-3.180	1.610	-3.180	1.610	0.000	1.020	0.005	0.022	0.005	0.000	0.802	1.240
LTS	-0.107	0.801	-0.107	0.801	0.990	1.000	0.000	0.000	0.000	0.996	0.747	1.260
THL	-0.028	0.047	-0.028	0.047	0.000	1.000	0.000	0.000	0.000	0.000	-1.510	1.711
BAYES	-3.180	1.610	-3.180	1.610	0.000	1.020	0.005	0.022	0.005	0.000	1.010	1.040

Tables 2A gives the general estimators' performance results chart for Regression Estimators of standard model and the models with outlier of some error distribution with sample size $n = 500$. Theil's and LTS estimators gave negative RMSE values without regards for the sample size except under outlier error model as long as the error distribution is normal, but as the sample size increases and the error term deviates further from normality, Theil's gains precision and its MSE values were approximately equal with that of OLSE when the sample size is very large.

TABLE 2B: GENERAL ESTIMATORS' PERFORMANCE RESULTS CHART FOR SOME REGRESSION ESTIMATORS

(SAMPLE SIZE $N = 500$)

Tables 2B gives the general estimators' performance results chart for Regression Estimators of mixture models and the models

MIXTURE MODEL												
NORMAL												
N=500	α					β					CREDIBLE INTERVAL	
METHOD	$\bar{x}$	σ^2	Bias	MSE	RMSE	$\bar{x}$	σ^2	Bias	MSE	RMSE	LCL	UCL
OLSE	3.580	0.540	3.580	0.540	0.000	0.983	0.000	-0.000	0.000	0.000	0.958	1.010
LTS	1.440	0.200	1.440	0.200	0.836	0.993	0.002	-0.001	0.048	8.360	0.960	1.030
THL	0.909	0.077	0.909	0.077	0.931	0.996	0.000	-0.004	0.000	0.093	-2.470	2.670
BAYES	3.580	0.520	3.580	0.520	0.000	0.983	0.000	-0.001	0.000	0.465	0.978	0.989
LOGNORMAL												
N=500	α					β					CREDIBLE INTERVAL	
METHOD	$\bar{x}$	σ^2	Bias	MSE	RMSE	$\bar{x}$	σ^2	Bias	MSE	RMSE	LCL	UCL
OLSE	4.821	7.432	4.821	7.432	0.000	-2.211	C	-2.211	-2.211	0.000	-6.071	5.631
LTS	2.730	7.678	2.730	7.678	1.000	-1.280	1.681	-1.280	-1.280	1.000	-8.231	8.231
THL	1.780	0.098	1.780	0.098	0.000	0.996	0.000	-0.003	0.996	0.000	-5.921	5.921
BAYES	4.821	7.432	4.821	7.432	0.000	-2.2110	1.572	-2.211	-2.210	0.000	-2.210	2.211
CAUCHY												
N=500	α					β					CREDIBLE INTERVAL	
METHOD	$\bar{x}$	Var	Bias	MSE	RMSE	$\bar{x}$	Var	Bias	MSE	RMSE	LCL	UCL
OLSE	3.800	1.275	3.800	1.275	0.000	0.827	3.010	-0.170	3.010	0.000	-1.780	1.780
LTS	1.370	3.800	1.370	3.800	1.000	0.995	0.001	-0.053	0.001	1.000	-2.500	2.530
THL	1.200	0.405	1.200	0.405	0.006	0.994	0.000	-0.057	0.000	0.006	-1.805	1.800
BAYES	3.800	1.270	3.800	1.270	0.000	0.827	3.010	-0.173	3.010	0.000	-1.320	1.480
CONTAMINATION MODEL												
NORMAL												
N=500	α					BETA					CREDIBLE INTERVAL	
METHOD	$\bar{x}$	σ^2	Bias	MSE	RMSE	$\bar{x}$	σ^2	Bias	MSE	RMSE	LCL	UCL
OLSE	0.675	0.023	0.675	0.023	0.000	0.997	0.000	-0.003	0.000	0.000	0.995	0.998
LTS	0.435	0.018	0.435	0.018	0.553	0.998	0.000	-0.002	0.000	0.553	0.996	1.000
THL	0.381	0.024	0.381	0.024	0.037	0.998	0.000	-0.002	0.000	0.037	-1.540	1.740
BAYES	0.673	0.023	0.673	0.023	0.000	0.997	0.000	-0.003	0.000	0.000	0.996	0.998
5LOGNORMAL												
N=500	α					β					CREDIBLE INTERVAL	
METHOD	$\bar{x}$	σ^2	Bias	MSE	RMSE	$\bar{x}$	σ^2	Bias	MSE	RMSE	LCL	UCL
OLSE	1.240	0.053	1.240	0.053	0.000	1.000	0.000	0.003	0.000	0.000	0.990	1.010
LTS	0.776	0.075	0.776	0.075	0.025	1.000	0.000	0.008	0.000	0.000	0.998	1.010
THL	0.776	0.015	0.776	0.015	0.584	1.000	0.000	0.049	0.000	0.000	-0.176	1.960
BAYES	1.240	0.053	1.240	0.053	0.000	1.000	0.002	0.002	0.002	0.000	0.996	1.000
CAUCHY												
N=500	α					β					CREDIBLE INTERVAL	
METHOD	$\bar{x}$	σ^2	Bias	MSE	RMSE	$\bar{x}$	σ^2	Bias	MSE	RMSE	LCL	UCL
OLSE	1.940	2.420	1.940	2.420	0.000	0.976	0.024	-0.024	0.024	0.000	0.799	1.020
LTS	0.633	0.124	0.633	0.124	0.996	0.997	1.780	-0.034	1.780	0.999	0.869	1.130
THL	0.509	0.036	0.509	0.036	1.000	0.997	0.226	-0.000	0.226	1.000	-9.980	1.200
BAYES	1.940	2.422	1.940	2.422	1.000	0.976	0.024	-0.024	0.024	1.000	0.889	0.939

with contaminations of some error distribution with sample size $n = 500$. With the insertion of contaminations into the dataset, it's far noticed that every estimator remained noticeably unbiased. However, OLSE and Bayesian estimators were found to maintain steady unbiased-ness handiest for both large and small trial size (for outliers' model) and regular biased-ness (for contamination model) pattern length notwithstanding. Under this circumstance, LTS estimator stayed on pinnacle maximum of the time (maximum specifically each time the error time period is Cauchy allotted) with the least bias value, accompanied closely by means of Theil's estimator and then the Bayesian factor estimator.

3.1. Based on variance and RMSE criteria

The Bayesian point estimator of the slope parameter β, accompanied by the OLSE estimator, had the smallest σ^2 and therefore the minimum of mean square error with Normal distribution throughout all cells of the simulation and all sample sizes, aside the exception of under mixture and contamination error model for large sample size ($n = 500$) wherein Theil's and LTS estimators led the way. As sample size increases, Bayesian and OLSE received precision i.e. the variance and the MSE values of each estimator lower with growing sample size. Ultimately the Bayesian point estimator converges to the OLSE as the sample size has a tendency to infinity. However, the Bayesian factor estimator maintained steady continuity in efficiency than the OLSE pattern regardless of the sample size.

As the dataset diverge from normality, however, Bayesian estimator and OLSE looses precision to LTS and Theil's estimators with nearly four hundred percent increment in the values of mean square error for samples with their distribution are lognormal and Cauchy. OLSE and Bayesian point estimator gave excessive large value for mean square errors compared with other estimators. This result affirms deviations from normality assumptions in reasons in which the OLSE is considered not to be an efficient estimator and thus its inappropriateness under non-normal conditions.

The Least Trimmed Squares estimation method came next in view to Theil's under the normal distribution for sample size n=30 and as has been usually observed, it experiences a consistent decrease in its MSE as sample size gets larger.

Theil's and LTS estimators gave negative RMSE values without regards for the sample sizes besides under outlier error model. But as the sample size increases, the error term deviates further from Theil's gain precision and its MSE values had been approximately equal with that of OLSE whilst the sample size is very massive. The Least Trimmed Squares estimation approach followed Theil's underneath the normal distribution for sample size n=30.the result discovered the father away the error distribution diverges from normality, the more efficient LTS turs into as it acquires more accuracies following closely after Theil's while replacing Bayesian estimator and OLSE absolutely.

3.2 Effect of contamination

3.2.1 Y-intercept estimators' performance

The overall performance of the regression y-intercept estimators for every of the methods was observed to obey similar outline of slope estimators, but for some remarkable variations. The results shows that the intercept of Bayesian and OLSE approach of estimations gives more accurate precisions than the intercept of LTS and Theils methods of estimation irrespective of their sizes as much as normality assumptions holds. Theils and LTS is still efficient with non-normal assumptions especially with Cauchy error distribution. The Bayesian estimators consistently maintains a higher precision about the true value of the regression parameters compared to all other estimators under consideration.

3.3 Empirical Data Analysis

Regression analysis was carried out on the dataset for $n = 10$ years $(2010 - 2021)$, $n = 30$ years $(1991 - 2021)$ and $n = 50$ years $(1971 - 2021)$ respectively of Nigeria Gross Domestic Product and Gross National Income. Table 6 shows the regression parameter estimates for the dataset using the previously discussed methods. The prior parameter for the intercept and slope parameters is obtained as the OLSE estimates for the intercept and slope parameter of the data ten years before the years in view. The variances for both parameters were calculated as:

$$\left(\frac{(m'+3\sigma)-(m'-3\sigma)}{6}\right)^2 \tag{20}$$

where σ is the OLSE sample estimate.

TABLE 3: EMPIRICAL ANALYSIS OF GROSS DOMESTIC PRODUCT AND GROSS NATIONAL INCOME USING SELECTED REGRESSION TECHNIQUES

	DATA WITH NO CONTAMINATION						**DATA WITH 20% CONTAMINATION**					
	α	β	**C. INTER (β)**		**MODEL**		α	β	**C. INTER (β)**		**MODEL**	
ESTIMATORS	**Mean (S.Dev)**	**Mean (S.Dev)**	**LCL**	**UCL**	**BIAS**	**MSE**	**Mean (S.Dev)**	**Mean (S.Dev)**	**LCL**	**UCL**	**BIAS**	**RMSE**
	n =10											
OLSE	676.2 **(345.0)**	0.6994 **(0.138)**	-48.2	49.6	-0.0	92494.	-5471.3 **(2550.)**	3.3231 **(1.022)**	-2669	2676	-0.0	208801
LTS	392.0 **(318.5)**	0.7741 **(0.149)**	-59.6	61.1	-100.	12435.	-1101.9 **(3188.)**	1.4479 **(1.641)**	-5083	5086	-238.	453453
THELS	660.8 **(1030.)**	0.7033 **(0.418)**	-48.2	436.6	-5.9	93557.	-519.2 **(450.4)**	1.1894 **(1.103)**	-2669	8393	-291.	352327
BAYESIAN	676.2 **(60.1)**	0.6994 **(0.139)**	0.442	0.956	-0.0	92494.	-5471.3 **(444.5)**	3.3321 **(1.022)**	1.422	5.224	-0.0	2088001
	n =30											
OLSE	-5.7 **(0.27)**	0.9469 **(0.094)**	-6.6	8.5	-0.0	787175	-45.7 **(359.3)**	1.1506 **(0.219)**	-419.	421.9	-0.0	1162307
LTS	11.6 **(1.42)**	**0.9056** **(0.036)**	-9.0	10.8	-38.1	721485	-31.7 **(386.8)**	0.9846 **(0.306)**	-602.	604.5	-208.	894615.7
THELS	11.316 **(O.167**	0.9031 **(0.211)**	-6.6	756.2	-41.8	717783	4.6 **(985.9)**	0.9059 **(0.419)**	-419.	1543	-278.	797771
BAYESIAN	-5.7 **(0.034)**	0.9469 **(0.029)**	0.8969	0.9970	0.0	78175.8	-45.7 **(205.7)**	1.1506 **(0.219)**	0.772	1.542	-0.0	1162307.2
	n = 50											
OLSE	43.045 **(0.932**	0.8960 **(0.036)**	-12.1	13.9	-0.0	592418	291.4 **(284.8)**	0.9523 **(0.206)**	-434.	436.2	0.0	669197
LTS	33.341 **(0.662)**	0.8934 **(0.041)**	-14.1	15.8	-12.6	58122.	49.5 **(308.7)**	0.9238 **(0.245)**	-527	530.7	-272.	704069.5
THELS	9.800 **(0.452)**	0.9101 **(0.213)**	-12.1	466.5	-17.9	611519.5	53.7 **(408.2)**	0.8690 **(0.296)**	-434	895	-327	664607
BAYESIAN	43.027 **(30.06)**	0.8960 **(0.036)**	0.836	0.956	0.0	592418.	291.4 **(177.2)**	0.9523 **(0.206)**	0.605	1.300	-0.0	669197

3.3.1 Intercept Parameter Estimators' Performance

For a very small sample size, $n = 10$, the Bayesian estimates of the regression intercept parameter α has the least standard deviation (60.115). The LTS estimator put up a better performance with standard deviation 318.5068 than the OLSE estimator with standard deviation 345.0132. The standard deviation (1030.231) of the Theil's estimator is particularly very large compared to OLSE and LTS. However, as the sample size increases, OLSE and LTS improved significantly with almost 85.6% and 83.5% reduction in standard deviation respectively while Theil's estimator improved with 53.9% reduction in standard deviation ($for\ n = 30$). When the sample size increased to 50, Theil's estimator further improved with 71.2% reduction in standard deviation, OLSE and LTS had no significant improvement, while the Bayesian estimator remained consistently more efficient than all others regardless of sample size. With 20% contamination in the data set, however, OLSE and LTS estimator lost significant precision with 468.0% and 526.8% increase in standard deviation respectively, while Theil's estimators gave a better performance with 38.7% and the Bayesian estimator seems not to be significantly affected with just about 7.7% increase in the standard deviation. This ascertained the robustness characteristic of the Bayesian estimator to contamination in the dataset.

3.3.2 Slope Parameter Estimators' Performance

For a very small sample size, $n = 10$, it is observed from Table 3 that the Bayesian estimate for the regression slope parameter β has the least standard deviation (0.13825) outperforming the LTS and Theil's estimators by 12.4% and 206.1% respectively. The OLSE estimator for slope parameter seems to agree with the Bayesian estimator for all sample sizes. With significant increase in sample size ($n = 30$), OLSE, LTS and Bayesian estimates improved significantly with 78.0%, 73.8% and 76.2% decrease in standard deviation respectively while the Theil's estimator slightly improved with just about 29.7% reduction in standard deviation.

With further increase in sample size ($n = 50$), the Theil's estimator gained precision with a further 18.3% reduction in its standard deviation while the OLSE, LTS and Theil's seem to be relatively consistent in their precision. With 20% contaminations in the dataset, however, the Theil's estimator proved to be more robust than all others with just 37.6% reduction in its precision. This affirms the fact that OLSE estimator breaks down in the presence of outliers in the data set, thus, the Theil's slope estimator maintained an overall and consistent robustness over all other slope estimators across board.

3.3.3 Test of Hypothesis

A two-sided hypothesis was erected for the regression slope parameter β as follows:

$$H_{a\beta} : \beta = \xi$$

Versus

$$H_{1\beta} : \beta \neq \xi \tag{21}$$

(where $\xi = 0.8458, 0.2818, 0.4544$ for $n = 10,30,50$ respectively)

Considering range of the confidence interval, in Table 3 above constructed for all estimators under consideration at $\alpha = 0.05$, the confidence interval for OLSE, LTS and Theil's estimator spans the corresponding value of ξ for each sample size. Furthermore, the intervals showed that the value of β could possibly be zero and as well could be negative. Thus, the test is not significant at $\alpha = 0.05$ and we therefore cannot reject H_a ($\beta = \xi$) and conclude that there exists a significant predictive relationship between GDP and GNI at $\alpha = 0.05$ level of significance though the relationship is weakly quantified by the estimated values of the slope parameter respectively.

However, the Bayesian credible interval provided strong evidence, in the light of both the data and the prior knowledge that the values $\xi = 0.2818$ and 0.4544 are not tenable values of β, since the test is significant for both $\xi = 0.2818$ and 0.4544, but $\xi = 0.8458$ is a possible value of β since the test is not significant for $\xi = 0.8458$ for both large and very small sizes. Also, the credible interval revealed a very strong opinion that β can neither possibly be zero nor negative in value. Thus, $0.8361 \leq \beta \leq 0.997$ is a credible range of β at $\alpha = 0.05$ and we could therefore conclude that there exists a significant predictive relationship between GDP and GNI at $\alpha = 0.05$ level of significance and the relationship is strongly quantified by the estimated values of the slope parameter respectively.

Hence, the fitted Simple Linear Regression model for each of the estimation procedure is as follows:

Ordinary Least Square Estimation: GNI = 43.00191 + 0.89603GDP

Least Trimmed Square Estimation: GNI = 33.25113 + 0.89342GDP

Theil's Estimation: GNI = 09.82935 + 0.91013GDP

Bayesian Estimation: GNI = 43.00191 + 0.89603GDP

4. Conclusion.

Linear regression is characterized with the underlying assumption that error terms have a normal distribution, which leads OLSE procedure to give good inferences. However, in real life it is nearly impossible to discover set of data that satisfies the normality assumption. Moreso, with these conditions, OLSE yields lack of efficiency, alternative regression approaches are required. In this paper, some robust procedures for a simple linear regression model under normal and non-normal error situations were studied. Results from the study shows that the Theils method demonstrates the strongest performance gains as compared to OLSE in terms of model evaluation error. The lower the value of mean square error, the better the fit of the model with the estimator. This study revealed that Bayesian linear regression is a more optimal and safer alternative when compared with OLSE and provides inferences that are conditional on the data and are exact without reliance on asymptotic approximation. Theil estimator has high small sample efficiency across board, but especially when the variance of the error terms is not constant. More so, Theils estimation is the most efficient and reliable which can be applied in numerous situations. Also, the empirical results affirm the robustness characteristic of the Bayesian estimator to contamination in the dataset as it remains robust with violations of linear regressions assumptions.

Least Trimmed Squares is most especially applicable when the error term is confirmed to come from a heavy tailed distribution and the sample size is large. OLSE is only consistence if it is best linear and unbiased.

The current study has also confirmed the applicability of the Theils method to varying circumstances and its robustness over many other robust methods, especially the LTS and Bayesian methods.

References

[1] N. S. Namita and R. Tandon, "A comprehensive study of outliers," *International Journal of Engineering Research and Technology*, vol. 11, no. 1, pp. 58–62, 2022.

[2] D. Birkes and Y. Dodge, *Alternative Methods of Regression*. New York, NY: Wiley, 1993.

[3] A. I. Adewole, O. K. Bodunwa, and F. B. Adebola, "Bayesian causal relation effect in quantiles regression models," *Kathmandu University Journal of Science, Engineering and Technology*, vol. 16, no. 1, pp. 1–6, 2022.

[4] Y. Lio, D. G. Chen, H. T. Ng, and T. R. Tsai, *Bayesian Inference and Computation in Reliability and Survival Analysis*. Springer, 2022.

[5] P. J. Rousseeuw, "Least median of squares regression," *Journal of the American Statistical Association*, vol. 79, pp. 871–880, 1984.

[6] T. Bayes, "An essay towards solving a problem in the doctrine of chances," *Philosophical Transactions of the Royal Society*, vol. 53, pp. 370–418, 1763.

[7] M. Sami, *Nonparametric Regression Models: Theories and Applications in R*. Hal Open Science, 2023.

[8] R. Giacomini, T. Kitangwa, and M. Read, "Robust Bayesian analysis for econometrics," *CEPR Discussion Paper No. 16488*, CEPR Press, 2021.

[9] A. I. Adewole and O. K. Bodunwa, "Sensitivity analysis of prior distributions in regression model estimation," *Al Bahir Journal of Engineering and Pure Sciences*, vol. 4, no. 1, pp. 88–96, 2024.

[10] W. M. Bolstad, *Introduction to Bayesian Statistics*. Hoboken, NJ: Wiley, 2004.

[11] S. S. Hussain and P. Sprent, "Nonparametric regression," *Journal of the Royal Statistical Society. Series A (General)*, 1983.

[12] I. A. Ayoade and O. A. Fashoranbaku, "Determination of quantile range of optimal hyperparameters using Bayesian estimation," *Tanzania Journal of Science*, vol. 47, no. 1, pp. 981–987, 2021.

[13] L. Al-Essa, E. Endris, and Y. Mergiaw, "Bayesian regression modeling and inference of energy efficiency data: The effect of collinearity and sensitivity analysis," *Frontiers in Energy Research*, vol. 12, 2024.

[14] M. E. Bakr, B. M. Golam Kibra, and A. M. Gadallah, "A new non-parametric hypothesis testing with reliability analysis applications to model some real data," *Journal of Radiation Research and Applied Sciences*, vol. 16, no. 4, 2023.

[15] J. Zhu, Z. Ge, and Z. Song, "Variational Bayesian Gaussian mixture regression for soft sensing key variables in non-Gaussian industrial processes," *IEEE Transactions on Control Systems Technology*, vol. 25, no. 3, pp. 1092–1099, 2017.

[16] A. Ibrahim, I. Dike, and A. Badawaire, "Comparative study of some estimators of linear regression models in the presence of outliers," *FUDMA Journal of Sciences*, vol. 6, no. 1, pp. 368–376, 2022.

[17] J. Guerard, A. Saxena, and M. Gultekin, "Regression analysis and estimating regression models," in *Quantitative Corporate Finance*. Springer, 2021.