DOI: http://dx.doi.org/10.18187/pjsor.v18i2.2996

Pakistan Journal of Statistics and Operation Research

Huber M-estimator for Cumulative Odds Model with Application to the Measurement of Students' Final Exam Grades

Faiz Zulkifli[1], Zulkifley Mohamed[2*], Nor Afzalina Azmee[3]

* Corresponding Author

1. Faculty of Computer and Mathematical Sciences, Universiti Teknologi MARA, Perak Branch, Tapah Campus, 35400 Tapah Road, Perak, Malaysia, email: faiz7458@uitm.edu.my
2. Department of Mathematics, Faculty of Science and Mathematics, Universiti Pendidikan Sultan Idris, 35900 Tanjong Malim, Perak, Malaysia, email: zulkifley@fsmt.upsi.edu.my
3. Department of Mathematics, Faculty of Science and Mathematics, Universiti Pendidikan Sultan Idris, 35900 Tanjong Malim, Perak, Malaysia, email: afzalina@fsmt.upsi.edu.my

Abstract

The Huber M-estimator is proposed in this study as a robust method for estimating the parameters of the cumulative odds model, which includes a logistic link function and polytomous explanatory variables. With the help of an intensive Monte Carlo simulation study carried out using the statistical software *R*, this study evaluates the performance of the maximum likelihood estimator (MLE) and the robust technique developed. Bias, RMSE, and the Lipsitz Statistic are used to measure comparisons. When conducting the simulation study, different sample sizes, contamination proportions, and error standard deviations are considered. Preliminary findings indicate that the M-estimator with Huber weight estimates produces the best results for parameter estimation and overall model fitting compared to the MLE. As an illustration, the procedure is applied to real-world data of students' final exam grades as measured by two different estimators.

Key Words: Cumulative Odds Model, Maximum Likelihood Estimator, Ordinal Response Model, Robust M-estimator, Students' Final Exam Grades.

Mathematical Subject Classification: 62G32, 62G35

1. Introduction

The Fourth Industrial Revolution requires more precise data analysis to produce new technology and apps that improve people's lives. Many academics have recently moved to qualitative data to do more profound research. Qualitative data are frequently measured using nominal and ordinal scales. Ordinal data is commonly used in scientific research, education, sociological psychology, and economics (Agresti, 2010). The emergence of item instruments has raised interest in this form of data. Generally, category data appear when the item is expressed as an opinion, judgment, or rating. An online solution also reduces the cost of data collection.

Past scholars have built numerous ordinal data models based on Agresti (2010) and Tutz (2014). Ordinal data can be analysed using log-linear and logit models. Unlike the logit model, which treats one variable as a response and another as a covariate, the log-linear model solely looks at the relationship pattern between ordinal category variables. Previous researchers created various logit models. The models include polytomous, cumulative, partial proportional, adjacent category, continuation ratio, and stereotype ordinal regression.

As demonstrated by Huber (1981), Hampel et al. (1986), Maronna et al. (2006), and Huber & Ronchetti (2009), literature has made a significant contribution to the robustness of discrete and continuous data. However, the ordinal model's robustness has overshadowed its gains. Occasionally, respondents purposefully or mistakenly choose the wrong category (Jiang et al., 2019). This situation changed the estimator's nature and modelling. Only a few research have revealed a robust ordinal regression model. Recently, Scalera et al. (2021) presented requirements for evaluating the robustness of link functions when the covariates are outlier free. Iannario et al. (2017) used a robust M-estimator instead of maximum likelihood estimates of ordinal response models. They evaluated the estimator's performance in five different models with dichotomous and continuous explanatory variables. To test the Cub model's resilience, Iannario et al. (2016) used uniform random variables and binomial shifts. Croux et al. (2013) propose a robust estimator that includes a weighting step for an ordinal response model with a logistic link function. Moustaki & Victoria-Feser (2006) created a robust LISREL alternative and a robust estimator for latent variable models. On the other hand, the Bayesian technique has become increasingly popular among researchers (Al-Taweel & Sadeek, 2020). Albert & Chib (1993) and Albert & Chib (1995) developed residual Bayesian for categorical and ordinal data to detect outliers. However, none of the researchers employs M-estimator to estimate ordinal models with polytomous explanatory variables.

Thus, this study only estimates the cumulative odds model (COM) parameter with a logistic link function and polytomous explanatory factors to overcome outlier data. This model was chosen because it is extensively utilised by researchers who prefer its outcomes to other ordinal models. This paper developed a robust M-estimator based on Huber's weighting method. Monte Carlo simulation will be used to compare the maximum likelihood estimator (MLE) and the suggested M-estimator. Fitting models will evaluate the performance of these estimators to simulation data for various sample sizes, contamination proportions, and outlier point distances. The model's overall performance is assessed using the Lipsitz Statistic. Next, the approach will be applied to students' final exam grades estimated by two estimators.

2. Cumulative Odds Model

McCullagh (1980) proposed the COM, known as an odds ratio model. The latent regression model is translated into the Equation using variable Y as a multinomial categorical response and variable X as an m-dimensional vector for covariates with polytomous categories:

$$P(Y_i \leq j) = Y_i^* = \sum_{j=1}^{m} x'_{ij}\beta_j + \varepsilon_i, i = 1,2,\dots,n; j = 1,2,\dots,m; \alpha_{j-1} < Y_i^* < \alpha_j, \quad (1)$$

where β is the parameter of slope and x' is matrix transpose for covariates.

The probability distribution of Y_i is denoted by:

$$P(Y_i = j|x_{ij}) = P(\alpha_{j-1} < Y_i^* < \alpha_j) = F\left(\alpha_j - \sum_{j=1}^{m} x'_{ij}\beta_j\right) - F\left(\alpha_{j-1} - \sum_{j=1}^{m} x'_{ij-1}\beta_{j-1}\right), \quad (2)$$

where $-\infty = \alpha_0 < \alpha_1 < \cdots < \alpha_m = +\infty$ and $F(.)$ is a cumulative distribution function for a random variable ε_i. $P(Y_i \geq j)$ is another similar equation proposed by scholars and software (Croux et al., 2013). The estimator coefficient in the second equation has the opposite sign.

According to (2), the regression coefficient vector is not dependent on i. This implies that the link between x_i and Y_i is independent of the variable i. McCullagh (1980) defines it as a proportional odds assumption of equality in logarithmic proportions that crosses k point deductions.

The COM is the most extensively used ordinal regression model due to its simplicity. It can also be used on continuous variables that have been made discrete after data gathering (Iannario et al., 2017).

2.1. Maximum Likelihood Estimator

The parameters in (2) can be estimated using the MLE. This is because the logit function is a complex function that requires cell calculations, which results in the probability function not always having the approximation form.

The log-likelihood function defined by

$$L(y_i, x_{ij}; \alpha_j, \beta_j) = \sum_{i=1}^{n}\sum_{j=1}^{m} I(y_i = j) log P(Y_i = j | x_{ij})$$
$$= \sum_{i=1}^{n}\sum_{j=1}^{m} I(y_i = j) log\left[F\left(\alpha_j - \sum_{j=1}^{m} x'_{ij}\,\beta_j\right) - F\left(\alpha_{j-1} - \sum_{j=1}^{m} x'_{ij-1}\,\beta_{j-1}\right)\right], \quad (3)$$

where $I(y_i = j) = \begin{cases} 1, y_i = j \\ 0, otherwise \end{cases}$.

The first derivative of the log-likelihood can be represented as:

$$\frac{\partial}{\partial \beta_j} L(y_i, x_{ij}; \alpha_j, \beta_j)$$
$$= -\sum_{i=1}^{n}\sum_{j=1}^{m} I(y_i = j)\, log\left[F\left(\alpha_j - \sum_{j=1}^{m} x'_{ij}\,\beta_j\right) - F\left(\alpha_{j-1} - \sum_{j=1}^{m} x'_{ij-1}\,\beta_{j-1}\right)\right] \quad (4)$$
$$= -\sum_{i=1}^{n}\sum_{j=1}^{m} I(y_i = j)\frac{f\left(\alpha_j - \sum_{j=1}^{m} x'_{ij}\beta_j\right) - f\left(\alpha_{j-1} - \sum_{j=1}^{m} x'_{ij-1}\beta_{j-1}\right)}{F\left(\alpha_j - \sum_{j=1}^{m} x'_{ij}\beta_j\right) - F\left(\alpha_{j-1} - \sum_{j=1}^{m} x'_{ij-1}\beta_{j-1}\right)} x_{im},$$

where

$$\frac{\partial F\left(\alpha_j - \sum_{j=1}^{m} x'_{ij}\,\beta_j\right) - F\left(\alpha_{j-1} - \sum_{j=1}^{m} x'_{ij-1}\,\beta_{j-1}\right)}{\partial \beta_j}$$
$$= \left[f\left(\alpha_j - \sum_{j=1}^{m} x'_{ij}\,\beta_j\right) - f\left(\alpha_{j-1} - \sum_{j=1}^{m} x'_{ij-1}\,\beta_{j-1}\right)\right] x_{im}.$$

Franses & Paap (2010) defined generalised errors as follows:

$$e_{ij}(\alpha_j, \beta_j) = \frac{f\left(\alpha_j - \sum_{j=1}^{m} x'_{ij}\beta_j\right) - f\left(\alpha_{j-1} - \sum_{j=1}^{m} x'_{ij-1}\beta_{j-1}\right)}{F\left(\alpha_j - \sum_{j=1}^{m} x'_{ij}\beta_j\right) - F\left(\alpha_{j-1} - \sum_{j=1}^{m} x'_{ij-1}\beta_{j-1}\right)}, \quad (5)$$

where for $j = 1$ and $j = m$ yields

$e_{i1}(\alpha_1, \beta_1) = \frac{f(\alpha_1 - x_i\beta)}{F(\alpha_1 - x_i\beta)}, e_{im}(\alpha_m, \beta_m) = \frac{f(\alpha_{m-1} - x_i\beta)}{1 - F(\alpha_{m-1} - x_i\beta)}$.

Equating (4) with zero yields

$$\frac{\partial}{\partial \beta_j} L(y_i, x_{ij}; \alpha_j, \beta_j) = 0$$
$$-\sum_{i=1}^{n}\sum_{j=1}^{m} I(y_i = j)\, e_{ij}(\alpha_j, \beta_j) x_{im} = 0. \quad (6)$$

In the linear regression model, (6) has the same structure as the maximum likelihood equation of the Gaussian, which has the form $\sum_{i=1}^{n} r_i x_i$ and can be solved using Newton-Raphson iterations.

Respondents can be dissatisfied with their choices or data to be collected incorrectly. There may be outliers in the response *Y* that have limited responses, such as $\{1,2,...,m\}$ that affect the generalised residuals in equation (5). The resulting residual leaves an indelible imprint in (6).

3. Robustness

The breakdown point examines the robustness of estimated procedures by identifying the lowest breakdown of the data before producing unacceptable estimates. The breakdown point size can be used to test the robustness of an estimator. The estimates are more robust in overcoming extreme data if the breakdown point is higher. When $n \times \alpha$ approaches infinity and the estimator becomes obsolete, the finite sample breakdown for the estimator is the lowest fraction α of data points. This is the breakdown point formula:

$$\varphi(\hat{\beta}_m; y_1, y_2, \dots, y_m) = \frac{m^*}{n} \times 100, \quad (7)$$

where: $\hat{\beta}_m$ is the estimate for a set of m outlier data points,

$m^* = max\{m \geq 0; \hat{\beta}_m < \infty\}$ and n is a sample size.

The breakdown point should not be higher than 50% of the total in most cases. According to Rousseeuw & Leroy (1987), when more than 50% of the points are contaminated, it is impossible to discern between the original distribution and the contaminating distribution of the points.

3.1. Robust Estimation

The MLE is typically used to estimate the COM. However, it is essential to note that this strategy works best when the terms of the error match the predetermined assumptions exactly. When these assumptions are not followed, estimators are more sensitive. When dealing with data availability extremes and leverage points that affect the estimator's performance, there should be vigilance. Even in the regression line, good leverage points significantly impact on the standard residuals but are difficult to identify (Huber, 1981). The MLE for the COM with polytomous explanatory variables has issues, and this study presents a robust technique to solve those concerns. The least absolute deviations, the least median of squares, the M-estimator, and the MM-estimator are among the most popular robust regression model estimates.

The M-estimator follows the same procedures as the MLE. Adding weights to Equation (6) creates this estimator's concept

$$-\sum_{i=1}^{n}\sum_{j=1}^{m} I(y_i = j)\, e_{ij}(\alpha_j, \beta_j) x_{im} w(y_i, x_i; \alpha_j, \beta_j) = 0. \quad (8)$$

The M-estimator error functions must meet three requirements: they cannot have a negative value, they cannot be a decreasing function, and they must be symmetrical. It is essential to adjust the errors of the M-estimator to solve the minimization problem, as it does not change scale with the M-estimator.

The M-estimator has a breakdown point that is somewhat close to 0.5. Huber is one of the most widely used objective functions among the many others available (Jiang et al., 2019). This method is utilised for a robust estimate because of its widespread use and excellent performance. The weight function of Huber can be expressed as follows:

$$w(y_i, x_i; \alpha_j, \beta_j) = \begin{cases} 1 & , \sum_{j=1}^{m} I(y_i = j)\, |e_i(\alpha_j, \beta_j)| \leq c \\ \dfrac{c}{\sum_{j=1}^{m} I(y_i = j)\, |e_i(\alpha_j, \beta_j)|} & , \sum_{j=1}^{m} I(y_i = j)\, |e_i(\alpha_j, \beta_j)| > c \end{cases} \quad (9)$$

Where c is set to the best option of 1.345σ, where σ is the errors' standard deviation. The standard deviation should be calculated using a robust measure of dispersion for the errors. Estimator $\hat{\sigma} = \frac{AMR}{0.6745}$ is commonly employed, where AMR is the median absolute residuals.

4. Procedures for Monte Carlo Simulation

Monte Carlo simulation is used to test the robustness of model estimators on simulation data with varying levels of contamination and error standard deviations. Monte Carlo simulation studies can evaluate the accuracy of existing statistical models when they are subjected to adverse circumstances (Mahdizadeh & Strzalkowska-Kominiak, 2017). In this study, the ordinal model was $y_{ij} = \alpha_j + \beta_1 x_{i1} + \beta_2 x_{i2} + \varepsilon_i$ where $i = 1,2, \dots, n \; and \; j = 1,2, \dots, 12$ - using a high number of categories to evaluate polytomous explanatory variable as well as matching the actual data for the illustration. The ordinal discrete scale is used for all variables.

The steps in the simulation are:
First, choose from 100, 250, or 500 samples. Divide the sample into two groups: good and contaminated. The number of outlier data points is $n \times p$, where p is the proportion of contamination (5, 10, 20, 40, and 50%). The rest of the sample will be good.
Second, the X values of all independent variables should be generated with a range of 0 to 12. Then obtain the error standard deviation for both the good and contaminated data points, $e_{cont.}$. A logistic distribution is used for both, with the same mean and standard deviation for each, as shown by $e \sim logistic(\mu = 0, \sigma = 1)$ and $e_{cont.} \sim logistic(\mu = 0, \sigma = k)$ where k=2,3, 4 and 5 and μ is the mean.
Third, generate true slope parameters from the uniform distribution over intervals 0 to 1, then log-odds values using the ordinal model's equation. The intercept parameters α_j are set to a constant value.
Forth, by employing a logistic distribution, calculate the cumulative probability for each category of Y: $P(Y \le y) = \frac{exp(\log - odds)}{1 + exp(\log - odds)}$. Then variable Y is generated by taking a value between 0 and 12 from the set of cumulative probability obtained.
Fifth, find the standardised residuals of model $\tilde{e}$ by fitting it with MLE and then repeat fitting the model with Huber weighting, as described therein (9).
Sixth, calculate the p-values for each estimate $\hat{\beta}$. Essentially, all p-values must be less than or equal to 0.05. (at 5 per cent significance level). This simulation only accepts statistically significant models on each of their variables.
Seventh, the process is repeated 5000 times, and then the bias and root mean square error (RMSE) of $\hat{\beta}$ for the proposed MLE and the M-estimator are calculated. Both measurements have the following formulas: $Bias = \frac{\sum_{i=1}^{200}(\hat{\beta}_i - \beta)}{5000}$ and $RMSE = \sqrt{\frac{\sum_{i=1}^{200}(\hat{\beta}_i - \beta)^2}{5000}}$. The best estimator for estimating the parameters β is the one with the lowest bias and RMSE values.
Finally, the value of the Lipsitz statistic, X^2 will be used to determine whether or not a model is overall good. The optimal model fitting will result in the average test statistic having the least value possible. The formula is given by $X^2 = 2 \sum_{g=1}^{G-1} \sum_{j=1}^{r} O_{gj} log \frac{O_{gj}}{E_{gj}}$ where O is the number of subjects, E is an estimate of the number of subjects, r is the number of responses, and G is the area based on the predicted mean score's percentile.

These simulation procedures will yield data that may be used to evaluate the model estimators. The simulation data anticipate the contamination proportion and the distance between the most outlier points. There are three methods used to evaluate the model estimators: bias, RMSE, and Lipsitz statistic. Bias and RMSE are the most accurate estimator for estimating model parameters as a rule of thumb. No statistical significance in the model is found if there is no statistically significant difference between observed and predicted frequencies. Monte Carlo simulation using the statistical software R was used to derive the study's findings. The analysis section will go into greater detail about the measurement findings.

 DOI: http://dx.doi.org/10.18187/pjsor.v18i2.2996

5. Simulation Results

In order to evaluate the estimation of the MLE and the proposed M-estimator, Monte Carlo simulation was utilized with a variety of data sets. The produced random data should contain three constant values that change and match the conditions of ordinal regression. The constants are the proportion of contamination, error standard deviation, and sample size. The proportions of selected contamination were 5%, 10%, 20%, 40%, and 50%. Errors are logistically distributed with standard deviations (S.D.) of 2s, 3s, 4s and 5s, where s is S.D. for the most non-extreme points. The sample sizes were 100, 250, and 500 with 5000 replications. The parameter estimations were measured via bias and RMSE. Meanwhile, the Lipsitz statistic was used to assess the overall goodness of fit. The model employed two parameters: β_1 and β_2.

5.1. The MLE

Figure 1 shows how the MLE approach could adapt to data with varying proportions of contamination. The combination of S.D. and sample size has produced a more consistent pattern of bias and RMSE. With more samples, the RMSE decreased until it was practically zero for both β_1 and β_2. However, the bias and RMSE values grow when the outlier point are placed further away from the rest of the data. The COM parameters can be affected by influencing points, according to this.

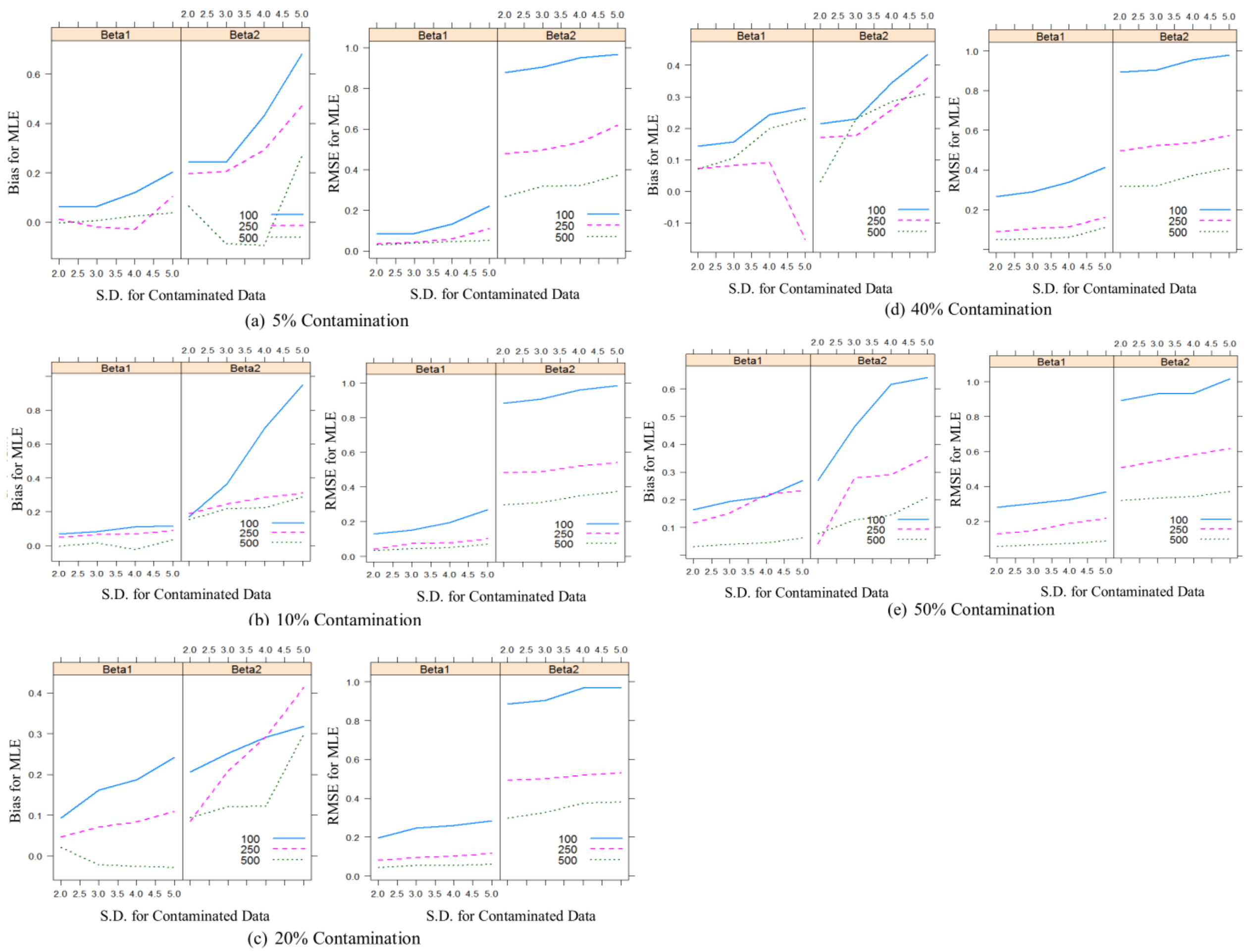


Figure 1: Bias and RMSE for MLE.

Figure 2 shows the Lipsitz statistic for the MLE against various combinations of simulated factors. The statistic numbers demonstrate an upward tendency as contamination levels rise. As the sample size and the S.D. of the outlier data increase, the pattern appears more frequently. Increases in the distance of outlier data, contamination proportion, and sample size have impacted parameter estimation and the accuracy of response variable predictions. These findings serve as a basis for developing a more robust estimator approach for the COM, as proposed by Croux et al. (2013) and Iannario et al. (2017).

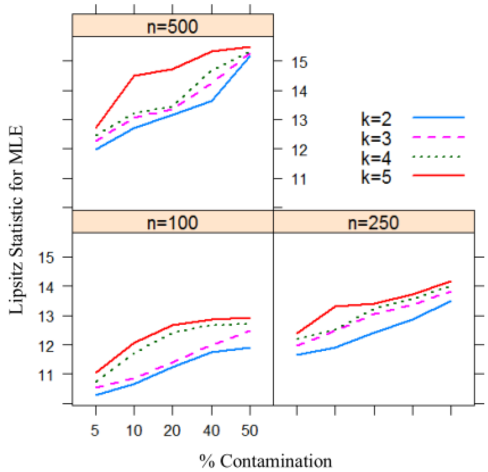
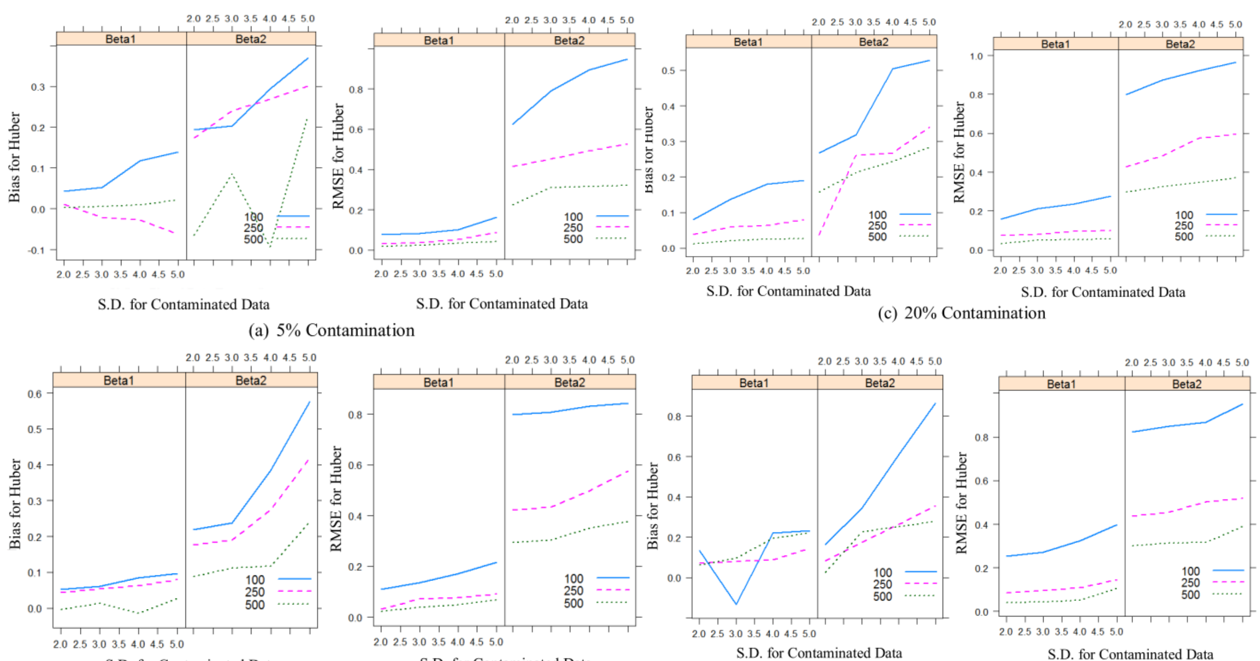


Figure 2: Lipsitz Statistic for MLE.

5.2. The Huber M-estimator

The Bias and RMSE patterns for the M-estimator with Huber weights have shown a more consistent trend, as illustrated in Figure 3. The bias pattern exhibits slight fluctuation since it permits values to be taken in either a positive or a negative direction. When the sample size is large, bias and RMSE trends for both parameters are close to zero. Conversely, the values of both measures increase in parallel with the rise in the outlier data's S.D.



(a) 5% Contamination (b) 10% Contamination (c) 20% Contamination (d) 40% Contamination

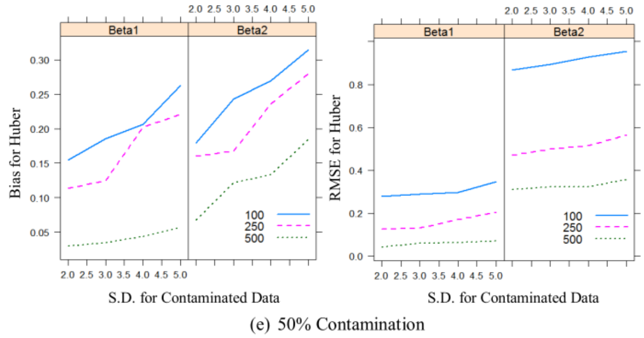


Figure 3: Bias and RMSE for Huber M-estimator.

The Huber M-estimator was also affected by changes in the S.D. of the outlier data, the proportion of contamination, and the sample size, as seen in the Lipsitz statistic plot from Figure 4. When all three components are raised, Lipsitz statistics demonstrate an upward tendency. In the comparison section, the performance of the M-estimator versus the MLE approach will be discussed.

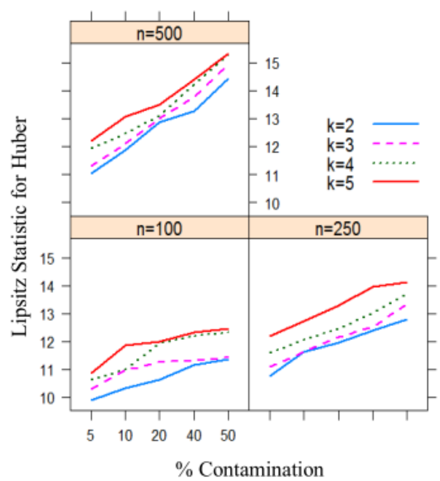


Figure 4: Lipsitz Statistic for Huber.

5.3. Bias, RMSE and Lipsitz Statistic between MLE and Huber M-estimator

The previous discussion patterns of bias, RMSE, and Lipsitz statistics for the two estimators focused primarily on simulation factors without identifying which estimator may produce the best measurement results. Figure 5 depicts the relationship between bias or RMSE and the proportion of contaminated data according to sample size and the S.D. of outlier data.

The bias and RMSE patterns for β_1 and β_2 tend to increase as contaminate levels rise. Both MLE and Huber estimators exhibit the same pattern. Huber weights, however, always have the lowest curve line or measurement value compared to the MLE. Outlier data S.D. values and sample sizes had no effect on this trend. The M-estimator can produce more accurate estimation values than the MLE, according to these data. Lipsitz statistics will examine the M-estimator's performance in more detail.

DOI: http://dx.doi.org/10.18187/pjsor.v18i2.2996

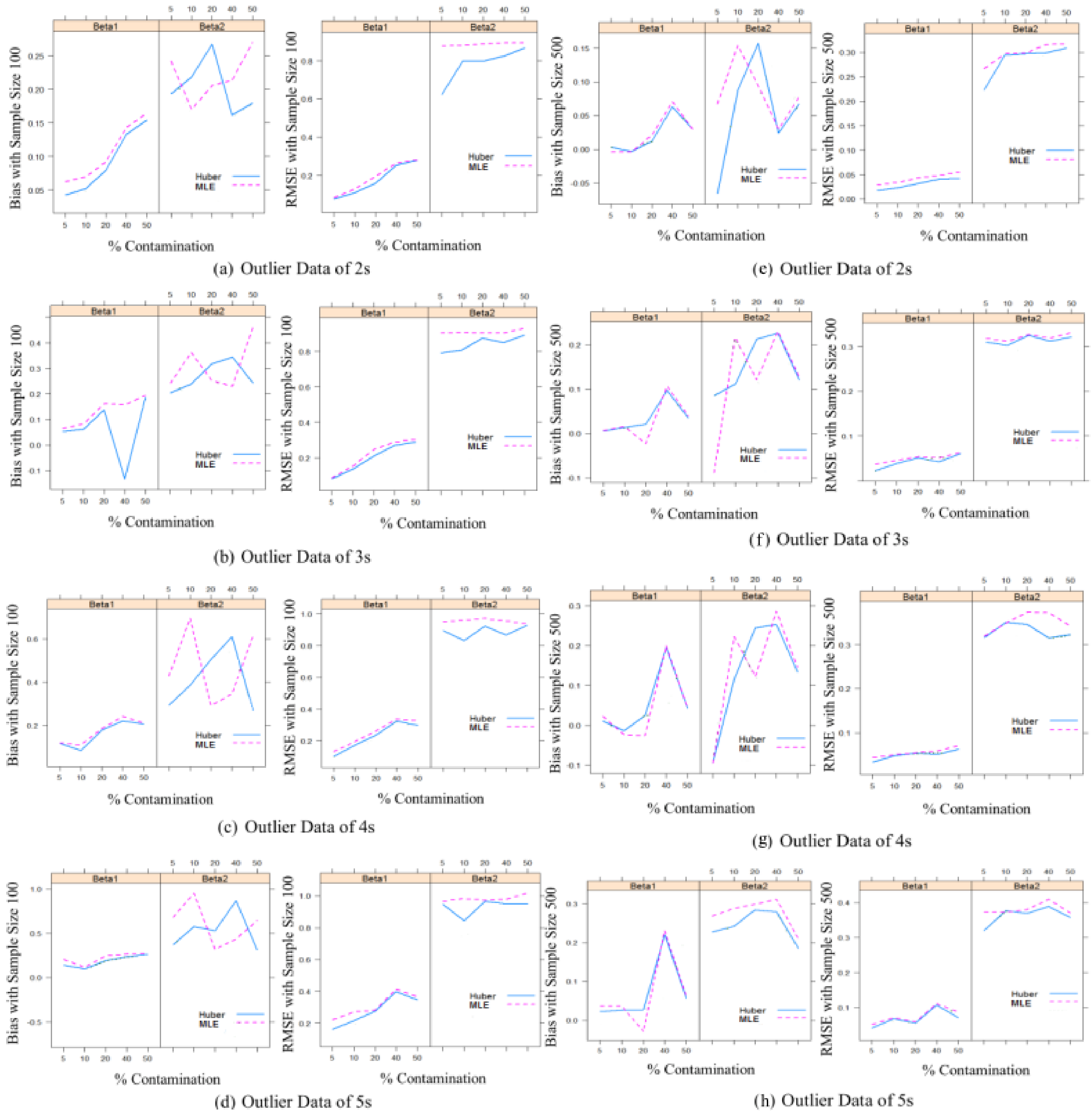


Figure 5: Bias and RMSE Between the MLE and the M-estimator.

Figure 6 illustrates that the Lipsitz statistical plot reveals that the M-estimator is better than the MLE. All estimates show an upward trend as contamination levels rise based on Lipsitz statistics. Huber weights produced the lowest Lipsitz statistic compared to the MLE, as evidenced by measurements of bias and RMSE. Variability in the S.D. of outlier data and sample size did not affect the performance.

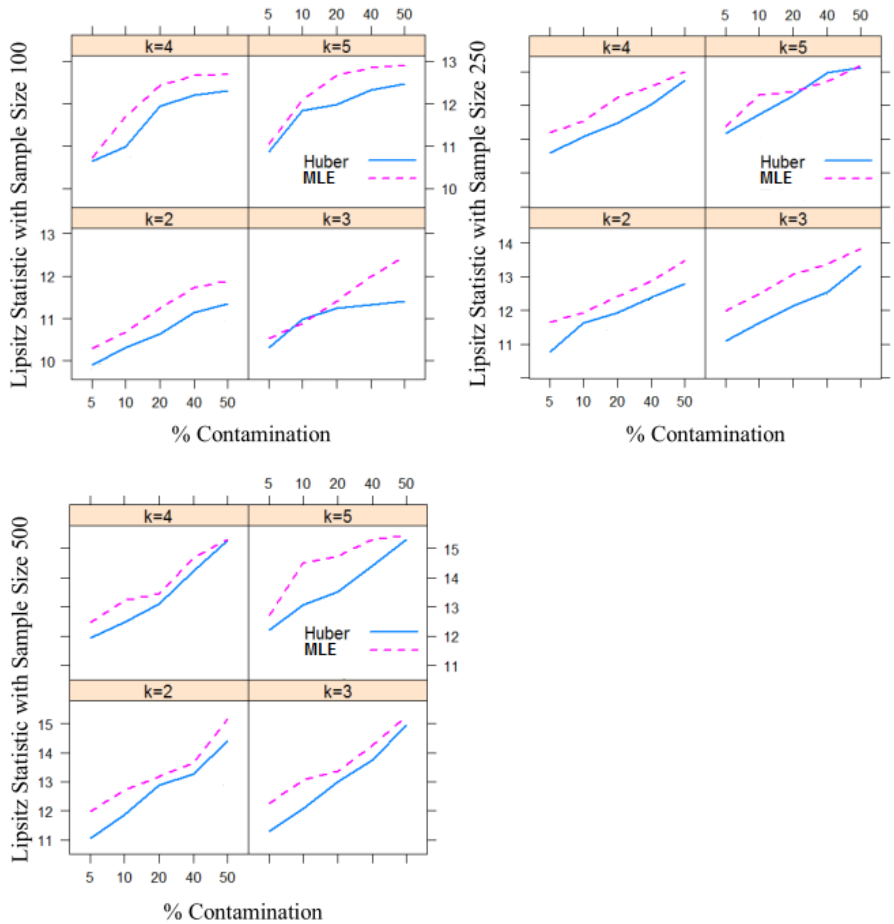


Figure 6: Lipsitz Statistic Between the MLE and the M-estimator.

6. Numerical Example

As an illustration, this study uses data from the final exam grades of 326 students as an example of how this research might be applied. Four explanatory variables that have the potential to impact the response variable have been discovered. The variables are the student's continuous assessment performance taken throughout the semester. All variables were graded into a 13-level system to satisfy the COM and categorical polytomous assumptions. Meanwhile, the selected course is "Introduction to Statistics and Probability", which collects data using cluster sampling. Lipsitz statistics will be employed once more to assess the accuracy of the estimators established for this study.

Table 1 shows the model fitting results using the MLE and the Huber M-Estimator. As seen in the table, Lipsitz's statistic fell dramatically when the Huber M-estimator was applied to the COM. This shows that the M-estimator is more effective than the MLE at reducing the impacts of outlier data on the model estimation.

Table 1: Lipsitz Statistic for Different Estimators

Estimator	Lipsitz Statistic
MLE	549.94
Huber M-estimator	64.70

Conclusion

In this study, an M-estimator with Huber weights was tested for its estimation parameter of the COM with logistic link function and polytomous explanatory variables. This study effectively evaluated the accuracy of the MLE and the Huber M-estimator using simulation data with a variety of sample sizes, contamination proportions, and error standard deviations. Both estimators' Monte Carlo simulation results indicate that the Huber estimator produced the best results for parameter estimation and overall model fitting. Both estimators obtained a 50% breakdown point for data, including outlier points that are very distant from the mean. Additionally, extreme points that are merely five times the distance from the most points do not affect the MLE. This means that if the model's error S.D. is between -5 and 5, the MLE and the Huber M-estimator are likely to produce the same result. This condition is also possible for data with less than a 5% contamination rate. As an illustration, this approach has been applied to the final exam grades data, and the results are consistent with those obtained in prior studies.

To ensure that the analysis's results are reliable and consistent across different data types, the researchers recommend doing simulations on data with multiple explanatory variables, a higher percentage of contamination, a larger sample size, and broader extreme point distances. In order to achieve this, simulation results can only be generated using high-end computers capable of increasing the test's speed and accuracy. Additionally, where many more knowledge gaps remain unexplored, the usage of other weights or robust methods can be suggested.

Acknowledgement

The authors would like to express their gratitude to the UiTM and the UPSI for allowing this study to be completed successfully.

References

1. Agresti, A. (2010). *Analysis of Ordinal Categorical Data* (2nd ed.). John Wiley & Sons, Inc.
2. Al-Taweel, Y., & Sadeek, N. (2020). A comparison of different methods for building Bayesian kriging models. *Pakistan Journal of Statistics and Operation Research*, *16*(1), 73–82.
3. Albert, J., & Chib, S. (1995). Bayesian Residual Analysis for Binary Response Regression Models. *Biometrika*, *82*(4), 747.
4. Albert, J. H., & Chib, S. (1993). Bayesian Analysis of Binary and Polychotomous Response Data. *Journal of the American Statistical Association*, *88*(422), 669–679.
5. Croux, C., Haesbroeck, G., & Ruwet, C. (2013). Robust estimation for ordinal regression. *Journal of Statistical Planning and Inference*, *143*(9), 1486–1499.
6. Franses, P. H., & Paap, R. (2010). *Quantitative models in marketing research*. Cambridge University Press.
7. Hampel, F. R., Ronchetti, E., Rousseeuw, P. J., & Stahel, W. A. (1986). *Robust statistics : the approach based on influence functions*. Wiley.
8. Huber, P. . (1981). *Robust Statistics*. J. Wiley & Sons.
9. Huber, P. J., & Ronchetti, E. (2009). *Robust Statistics* (2nd ed.). J. Wiley & Sons.
10. Iannario, M., Clara, A., & Piccolo, D. (2016). Robustness issues for CUB models. *TEST*.
11. Iannario, M., Monti, A. C., Piccolo, D., & Ronchetti, E. (2017). *Robust inference for ordinal response models*. *11*, 3407–3445.
12. Jiang, Y., Wang, Y. G., Fu, L., & Wang, X. (2019). Robust Estimation Using Modified Huber's Functions With New Tails. *Technometrics*, *61*(1), 111–122.
13. Mahdizadeh, M., & Strzalkowska-Kominiak, E. (2017). Resampling based inference for a distribution function using censored ranked set samples. *Computational Statistics*, *32*(4), 1285–1308.
14. Maronna, R. A., Martin, R. D., & Yohai, V. (2006). *Robust Statistics: Theory and Methods*. J. Wiley & Sons.
15. McCullagh, P. (1980). Regression models for ordinal data. *Journal of the Royal Statistical Society. Series B*, *42*, 109–142.
16. Moustaki, I., & Victoria-Feser, M.-P. (2006). Bounded-Influence Robust Estimation in Generalized Linear Latent Variable Models. *Journal of the American Statistical Association*, *101*(474), 644–653.
17. Rousseeuw, P. J., & Leroy, A. M. (1987). *Robust Regression and Outlier Detection*. John Wiley & Sons, Inc.
18. Scalera, V., Iannario, M., & Monti, A. C. (2021). Robust link functions. *A Journal of Theoretical and Applied Statistics*, *55*(4), 963–977.
19. Tutz, G. (2014). *Regression for Categorical Data* (Issue May). Cambridge University Press.